व्यावसायिक सांख्यिकी

बी. कॉम. II

दूरस्थ शिक्षा निदेशालय

महर्षि दयानन्द विश्वविद्यालय

रोहतक — 124 001

Maharshi Dayanand University
ROHTAK — 124 001

Developed & Produced by EXCEL BOOKS PVT. LTD., A-45, Naraina, Phase 1, New Delhi-110028.

विषय सूची

SYLLABUS

Paper - IV : Business Statistics

Max. Marks : 100
Time : 3 Hours

Note : 1. *Ateast ten questions shall be set in the question paper within minimum of three questions from each unit. The candidate shall be required to attempt five questions in all, selecting atleast one question but not more than two from each unit.*

Note : 2. *One weightage should given be given to theory portion.*

Unit-I Introduction : Statistics as a Subject, Statistical Data-Meaning and Types, Collection and Rounding of Data, Classification and Presentation of Data, Diagrammatic Presentation of Data, Graphic Presentation of Data, Statistical Averages, Measures of Dispersion.

Unit-II Method of Measurement of Correlation, Rank Correlation, Method of Concurrent Deviation, Coefficient of Determination, Association of Attributes, Regression Analysis (Linear), Uses of Regression Analysis, Regression Lines, Regression Equations, Standard Error of Estimate.

Unit-III Index Number : Definition and Characteristics, Problems involved in the construction of Index numbers, the uses of averages, Construction of different type of indices, Simple aggregate method, Simple average of relatives, Weighted aggregate method, Test of adequacy, Time reversal test, Factor reversal test and the Circular test, Consumer price index, Time Series Analysis, Definition, Utility of Time Series Analysis, Components of time and concepts series-secular trend, Seasonal variations, Cyclical variations, irregular variations, Measurement of trend, Moving average and Least Square Methods, Interpolation and Extrapolation.

Unit-IV Probability concept and various approaches of defining probability, Additive rule, Applicative theorem, Conditional probability and Bayes Theorem, Probability distributions : Binomial, Poisson and Normal distributions.

अध्याय - 1

परिचय : सांख्यिकी एक विषय
(Introduction : Statistics as a Subject)

सांख्यिकी का उद्‌गम एवं विकास
(Origin and Growth of Statistics)

यद्यपि आधुनिक अर्थ में, 'सांख्यिकी' शब्द बहुत पुराना नहीं है, तथापि यह शब्द प्राचीन काल से प्रयुक्त होता रहा है। इस शब्द की उत्पत्ति किस प्रकार हुई यह एक रोचक विषय है।

आंग्ल भाषा के 'Statistics' तथा 'Statistical' शब्दों की उत्पत्ति लैटिन भाषा के 'Status', इटैलियन भाषा के 'Statsta' और जर्मन भाषा के शब्द 'Statistsik' से हुई। इन सभी शब्दों का अर्थ राजनैतिक राज्य (Political State) है। 'Statistics' शब्द का सर्वप्रथम प्रयोग करने का श्रेय जर्मनी के प्रसिद्ध गणिताचार्य गॉटफ्रायड एचेनवाला को प्राप्त है। इन्हें सांख्यिकी का जन्मदाता भी कहा जाता है। सत्रहवीं शताब्दी में प्रसिद्ध कवि सम्राट विलियम शेक्सपीयर तथा महाकवि मिल्टन ने अपनी रचनाओं में इस शब्द का प्रयोग इसी अर्थ में किया था। शेक्सीपयर ने अपने प्रसिद्ध नाटक 'हैम्लेट' तथा 'साइम्बलाइन में और मिल्टन ने अपने महाकाव्य 'पैराडाइज रीगेन्ड' में इस शब्द का प्रयोग किया था। 1770 में बेरन जे. एफ. वॉन बॉपलफील्ड ने अपनी पुस्तक 'The Element of Universal Erudition' में सांख्यिकी पर एक अध्याय सम्मिलित किया था। उन्होंने सांख्यिकी को 'ज्ञात संसार' के आधुनिक राज्यों की राजनैतिक अवस्था सिखाने वाले विज्ञान के रूप में परिभाषित किया था।

पिछली तीन शताब्दियों में सांख्यिकी विज्ञान का बहुमुखी विकास हुआ था। इस विषय में 'अब्राहम डी मॉयर, लाप्लेस, गॉस, गाल्टन, कार्ल पियर्सन, विलियम एस. गोसेट, आर. ए. फिशर आदि के नाम उल्लेखनीय है। भारत में प्रो. पी. सी. महालानोबिस, डॉ वी. के. आर. राव एवं डॉ. पी. वी. सुखात्मक के नाम सांख्यिकी के विकास के क्षेत्र में विशेष तौर पर उल्लेखनीय हैं।

सांख्यिकी विज्ञान की उत्पत्ति दो स्त्रोतों से हुई है :-

(1) सरकारी अभिलेख; तथा

(2) गणित

(1) सरकारी अभिलेख (Governemnt Records) : यह सबसे प्राचीन स्त्रोत है। सरकारी कार्यों के लिये जनसंख्या की गणना की जाती थी। उदाहरणार्थ, 'मिस्त्र में पुलिस के पास प्रत्येक परिवार के मुखिया का लेखा होता है। इसी प्रकार जूडिया में जनसंख्या की गणना ईसा से 2030 वर्ष पूर्व हुई थी। जिसके अनुसार जनसंख्या का अनुमान 3,800,000 लगाया गया था। रोम में पहली जनगणना ईसा में 430 वर्ष पूर्व हुई थी। रोम में जनगणना के अतिरिक्त फौजी शक्ति, कर क्षमता, जन्म तथा मत्यु -दर सम्बन्धी समंक एकत्रित किये जाते थे। चूँकि समंक अधिकांशतः सरकारी कार्यों के लिये एकत्र किये जाते थे, इसलिये इसको 'सरकारी-विज्ञान' या 'राजतन्त्र की कला' कहा जाता था। सोलहवीं शताब्दी के आरम्भ में सांख्यिकी पर कई पुस्तकें प्रकाशित हुई।

(2) गणित (Mathematics) : आधुनिक सांख्यिकी विधियाँ, विशेषतया न्यादर्श के आधार पर समग्र का ज्ञान, गणित के संभावना सिद्धान्त पर आधारित है। इंग्लैंड व फ्रांस में प्रचलित जुए की प्रवत्ति से इस सिद्धान्त का उद्‌गम हुआ। सत्रहवीं शताब्दी में डी मॉयर, फरमेट, गैलिलियों और प्रमुख गणितशास्त्री 'जेम्स बरनौली, डैनियल बरनौली, इत्यादि ने हार-जीत के अवसर जानने के लिये सम्भावना सिद्धान्त की रचना की। आधुनिक संख्या-शास्त्री, प्राचीन काल के जुआरी की भाँति, किसी कार्य में निहित जोखिम को मापने में लगा है। किसी एक घटना से क्या होगा, बताना कठिन है परन्तु घटनाओं की संख्या अधिक हो तो सम्भावना अथवा अनुमान-संबंधी सिद्धान्त की सहायता से ठीक-ठीक अनुमान लगाया जा सकता है। प्रसिद्ध गणितशास्त्री डी मौरे ने सामान्य वक्र की खोज की, जो आधुनिक सांख्यिकी

सिद्धान्त का महत्त्वपूर्ण अंग है। पिछली शताब्दी में सांख्यिकी रीतियों में काफी प्रगति हुई है। प्रमुख संख्या-शास्त्रियों फ्रांसीसी गैलटर्न, ए. एल. बाउले, एजवर्थ कार्ल पियर्सन, आर. ए. फिशर, येट्स आदि ने सांख्यिकी विधियों के विकास में महत्त्वपूर्ण सहयोग दिया। इन्हीं के प्रयत्नों के सांख्यिकी एक समद्धिशाली शास्त्र बन सकी है।

सांख्यिकी का विकास (Growth of Statistics) : सांख्यिकी की उत्पत्ति राज्य विज्ञान (Science of Statecraft) के रूप में हुई। प्राचीन काल में सांख्यिकी का प्रयोग राज्यों के शासन प्रबन्ध को उचित रूप से चलने के लिये किया जाता था। आधुनिक समय में सांख्यिकी को निर्णय लेने का एक महत्त्वपूर्ण साधन माना जाता है। वास्तव में, प्रत्येक विज्ञान सांख्यिकी का प्रयोग किसी न किसी रूप में अवश्य करता है। सांख्यिकी के विकास के निम्नलिखित दो प्रमुख कारण हैं :-

(1) सांख्यिकी की बढ़ती हुई माँग (Increasing Demand for Statistics) : आधुनिक युग में व्यापार, वाणिज्य एवं सरकारी कार्यों में बहुत प्रगति हुई है। सांख्यिकी उचित नीतियों के निर्धारण में सहायता देती है इसलिये इसकी आवश्यकता बहुत बढ़ जाती है। व्यापार के कार्य क्षेत्र में वद्धि होती जा रही है तथा व्यापारिक समस्यायें जटिल होती जा रही हैं। इस जटिलता को वास्तविक तथ्यों के आधार पर सुगमता से सुलझाया जा सकता है।

सरकार को उचित रूप से शासन चलाने के लिये सांख्यिकी का अधिकाधिक प्रयोग करना पड़ रहा है। प्राचीन काल में सरकारी कार्य केवल आन्तरिक कानून व्यवस्था तथा देश को विदेशी आक्रमणों से बचाना था। सरकार आर्थिक समस्याओं से कम-से-कम हस्तक्षेप करने की नीति अपनाती थी; परन्तु आज सरकार का कार्य केवल कानून व्यवस्था को बनाये रखना अथवा विदेशी आक्रमणों से बचाना ही नहीं है। अब सरकार प्रत्येक क्षेत्र में सक्रिय है और उन सब कार्यों में अग्रसर है जो मानव-कल्याण के लिए आवश्यक है। उपयुक्त नीतियों के निर्माण में सांख्यिकी का अत्यधिक योगदान है।

विज्ञान के क्षेत्र में बहुत विकास हुआ है। अन्वेषण कार्य बहुत तेजी से बढ़ रहा है। सांख्यिकी विज्ञान अन्वेषण कार्य के लिये एक प्रकार का उपकरण है। इस प्रकार अनेक दिशाओं में सांख्यिकी की बढ़ती हुई माँग से सांख्यिकी विधियों के विकास को बहुत अधिक प्रोत्साहन मिला है।

(2) सांख्यिकी की घटती हुई लागत (Decreasing Cost of Statistics) : सांख्यिकी का प्रयोग समंकों को एकत्र करने में लगे समय और लागत पर निर्भर करता है। विद्युतगणक (Electronic Computers) के अविष्कार से समंक एकत्र करने के समय व लागत में बहुत बचत हुई है। उदाहरण के लिये, बड़े-बड़े अंकों का योग, गुणा, भुजा, वर्ग व वर्गमूल कुछ ही क्षणों में मशीनों द्वारा ज्ञान किया जा सकता है। फलस्वरूप सांख्यिकी का प्रयोग समस्याओं को हल करने में दिन-प्रति-दिन बढ़ता जा रहा है।

सांख्यिकी सिद्धान्तों के विकास में समंकों का एकत्र तथा विश्लेषण करने की लागत में बहुत कमी हुई है। उदाहरण के लिए निदर्शन रीति सहायता से समग्र का अध्ययन बहुत सरल हो गया है। 1935 में सांख्यिकी की एक शाखा परीक्षण के प्रारूप में भी काफी विकास हुआ है। इसके द्वारा समंकों का एकत्र करने तथा उनके विश्लेषण में लागत एवं समय की बहुत बचत हुई है। यद्यपि अनेक विद्वानों ने सांख्यिकी के विकास में योगदान दिया था तथापि आर. ए. फिशर तथा कार्ल पियर्सन के नाम विशेष तौर पर उल्लेखनीय हैं।

सांख्यिकी की परिभाषायें
(Definitions of Statistics)

'सांख्यिकी' शब्द की अनेक परिभाषायें दी गई है। कुछ विशेषज्ञों ने इसे तथ्यों के समूह (facts and figures or data) के रूप में माना है तथा कुछ ने सांख्यिकीय विधियों (Statistical Methods) के रूप में। कुछ प्रसिद्ध विद्वानों द्वारा दी गई परिभाषाओं का यहाँ विवेचन किया गया है।

वेब्स्टर (Webster) सांख्यिकी की परिभाषा इस प्रकार दी है :-

''समंक एक राज्य की जनता की स्थिति से सम्बन्धित वर्गीकृत तथ्य हैं - विशेष रूप से वे तथ्य जिनको अंकों के रूप में अथवा किसी भी सारणी या वर्गित पद्धति द्वारा प्रस्तुत किया गया है।

उपर्युक्त परिभाषा बहुत ही संकीर्ण है। इसके द्वारा सांख्यिकी का क्षेत्र केवल राजकीय लाभ के अंक संकलन में है। वास्तव में सांख्यिकी का क्षेत्र अब अत्यन्त विस्तत हो गया है।

डॉ. बाउले के अनुसार, ''किसी अनुसन्धान से सम्बन्धित अंकों में व्यक्त किये गये उन तत्वों के विवरण को समंक कहते हैं जिन्हें एक-दूसरे के सम्बन्ध में रखा जा सकता है।''

इस परिभाषा के अनुसार समंकों की तीन विशेषतायें होती हैं :-

(*i*) समंक तथ्यों को संख्यात्मक रूप से व्यक्त करते हैं;

(*ii*) समंक किसी अनुसन्धान से सम्बन्धित होते हैं; तथा

(*iii*) तुलना करने के लिए समंकों को एक-दूसरे के सम्बन्ध में रखा जाता है।

इस परिभाषा में समंकों की कुछ ही विशेषताओं पर प्रकाश डाला गया है - समस्त विशेषताओं पर नहीं; अर्थात् यह परिभाषा भी संकीर्ण है।

प्रो. यूल तथा केण्डल के अनुसार, ''समंकों का तात्पर्य उन संख्यात्मक तथ्यों से है। जो पर्याप्त सीमा पर अनेक प्रकार के कारणों से प्रभावित होते हैं।'' यह परिभाषा भी पूर्ण नहीं है क्योंकि इसमें समंकों की केवल दो विशेषताओं का ही उल्लेख किया गया है। (क) संख्या में व्यक्त होना तथा (ख) अनेक कारणों से प्रभावित होना।

प्रो. हौरेस सेक्रिस्ट (Prof. Horace Secrist) ने सांख्यिकी शब्द की परिभाषा इस प्रकार दी है :-

''सांख्यिकी से हमारा अभिप्राय तथ्यों के उस समूह से है, जो अनेक कारणों से पर्याप्त मात्रा में प्रभावित होते हैं, जिन्हें अंकों में व्यक्त किया जाता है, जिनकी गणना या अनुमान परिशुद्धता के एक उचित स्तर के अनुसार की जाती है, जिन्हें पूर्व-निश्चित उद्देश्य के लिए एक व्यवस्थित ढंग से संग्रहीत किया जाता है तथा तुलना के लिए एक-दूसरे के सम्बन्ध में रखा जा सकता है।''

इस परिभाषा के अनुसार समंकों के निम्न प्रमुख लक्षण कहे जो सकते हैं :-

(1) तथ्यों के समूह (Aggregate of Facts) : अकेला अथवा असम्बद्ध अंक समंक नहीं कहलाता। व्यक्तिगत अथवा असम्बद्ध समंकों की तुलना नहीं की जा सकती क्योंकि उनका आपस मे कोई सम्बन्धी नहीं होता। उदाहरण के लिए, यह वाक्य 'राम की आय 50,000 रू. वार्षिक है' सांख्यिकी नहीं कहलायेगा यद्यपि यह तथ्य अंकों में व्यक्त है। इसी प्रकार उत्पादन, बिक्री, जन्म, मत्यु , रोजगार, क्रय, आदि के अकेले अंक समंक नहीं हो सकते। पिछले पाँच या दस वर्षों के उत्पादन, बिक्री आदि में समंकों को सांख्यिकी कहेंगे क्योंकि इनकी तुलना की जा सकती है।

(2) विविध कारणों से प्रभावित (Affected to a Market Extent by Multiple of Causes) : साधारणतया, समंकों पर कई कारणों का प्रभाव पड़ता है। उदाहरण के लिए, चावल की उपज बहुत बातों पर निर्भर है, जैसे वर्षा ठीक हुई है या नहीं, भूमि की क्षमता, बीज की किस्म, खाद का प्रयोग, खेती का ढंग, इत्यादि। निश्चित रूप से यह कहना बहुत कठिन है कि चावल की उपज पर इन सब कारणों का अलग-अलग क्या प्रभाव पड़ा। चावल की उपज इन सब कारणों के सम्मिलित प्रभाव का परिणाम है। भौतिक शास्त्र तथा रसायन शास्त्र में अलग-अलग कारणों का प्रभाव आसानी से ज्ञात किया जा सकता है। सांख्यिकी विज्ञान में इस दिशा में कुछ कार्य अवश्य हो रहा है। तो भी घटनाओं का अलग-अलग प्रभाव जानना बहुत कठिन है क्योंकि बहुत से प्रभाव है जो नाप-तौल के योग्य नहीं है।

(3) संख्यात्मक रूप में अभिव्यक्त (Numerically Expressed) : गुणात्मक वाक्य जैसे 'भारत की जनसंख्या बहुत तेजी से बढ़ रही है'' या ''गेहूँ का उत्पादन काफी नहीं है'' आदि स्पष्ट हैं तथा इनका कोई अर्थ नहीं निकलता। सांख्यिकी तथ्यों को संख्यात्मक रूप में व्यक्त करती है जिसके अनेक लाभ हैं।

(4) यथोचित परिशुद्धता (Reasonable Standard Accuracy) : समंक दो विधियों से एकत्र किये जा सकते हैंः गिनती अथवा माप, और अनुमान से। अनुमान विधि पूर्ण गणना विधि की तुलना में इतनी शुद्ध नहीं है। उदाहरण के लिये अनुमान है कि 1992 में दस लाख व्यक्तियों ने 26 जनवरी की परेड देखी; इसका अर्थ है कि कुछ हजार व्यक्ति अधिक हो सकते हैं या कुछ हजार कम हो सकते हैं। यदि हम कक्षा में बैठे विद्यार्थियों को गिनकर कहें कि 60 विद्यार्थी बैठे हैं तो यह वाक्य शत-प्रतिशत शुद्ध होगा कुछ परिस्थितियों में शत-प्रतिशत शुद्धता होना असम्भव है, अतः उनमें अनुमान विधि अपनानी पड़ती है। जाँच के उद्देश्य एवं क्षेत्र के अनुसार ही शुद्धता होना असम्भव है, अतः उनमें अनुमान विधि अपनानी पड़ती है। जाँच के उद्देश्य एवं क्षेत्र के अनुसार ही शुद्धता के स्तर का निर्णय किया जाता है। उदाहरण के लिए, मनुष्यों की ऊँचाई सेन्टीमीटर में नापी जा सकती है। परन्तु दो शहरों के बीच की दूरी नापने के लिए हम किलोमीटर में ही नाम ज्ञात करेंगे। जैसा कि ऊपर कहा गया है प्रत्येक सांख्यिकी अध्ययन में शत-प्रतिशत शुद्धता नहीं हो सकती, तो भी यथोचित परिशुद्धता होना आवश्यक है अन्यथा समंक भ्रामक होंगे।

(5) सुनियोजित ढंग से संग्रहीत (Collected in a Systematic Manner) : समंक एकत्र करने से पहले तथा सुव्यवस्थित योजना बना लेना चाहिये। अव्यवस्थित ढंग से समंक एकत्र करने पर गम्भीर परिणाम होने की आशंका हमेशा बनी रहेगी।

(6) पूर्व निश्चित उद्देश्य (Pre-determined Purpose) : समंक पूर्व-निश्चित उद्देश्य के लिए एकत्र किये जाने चाहियें। बिना उद्देश्य संग्रहीत किये गये समंक प्रायः निरर्थक होते हैं।

(7) पारस्परिक सम्बन्ध (Placed in Relation to each other) : समंक सम्बद्ध होने चाहियें ताकि उनकी तुलना सुगमता से की जा सके। प्रायः उनकी तुलनाएं समयानुसार अथवा क्षेत्रानुसार की जाती है। उदारहण के लिए, भारत की 1991 की जनसंख्या की तुलना पिछले वर्षों की जनसंख्या से की जा सकती है। भारत की जनसंख्या की तुलना अमरीका और इंगलैंड की जनसंख्या के साथ की जा सकती थी। सार्थक तुलना तभी हो सकती है जबकि समंकों का आपस में सम्बद्ध हो। ऊँटों की ऊँचाई की तुलना मनुष्यों की ऊँचाई से की जायें तो यह सार्थक नहीं होगी।

उपर्युक्त वर्ण से यह स्पष्ट है कि सभी समंक संख्यात्मक होते हैं किन्तु सभी संख्यात्मक तथ्य समंक नहीं होते (All statistics are numeral statements of facts but all numerical statements of facts are not statistics) अतः अंकों को समंक तभी कहा जा सकता है जबकि उसमें उपर्युक्त लक्षण विद्यमान हो।

सांख्यिकी विधियाँ
(Statistical Methods)

कई लेखकों ने, जिनमें बाउले, बोडिन्गटन, सैलिंगमैन तथा क्रोक्सटन और कॉउटन के नाम उल्लेखनीय हैं, सांख्यिकी की परिभाषा बहुवचन के रूप में की है।

प्रो. ए. बाउले ने तीन परिभाषायें दी है : एक स्थान पर बाउले ने कहा कि कि "सांख्यिकी गणना का विज्ञान है।" यह परिभाषा बहुत ही संकीर्ण है, क्योंकि यह केवल एक ही पक्ष प्रस्तुत करती है। विश्लेषण, निर्वचन, आदि पक्षों का उल्लेख इस परिभाषा से नहीं मिलता।

बोडिंगटन (Boddington) के अनुसार, "सांख्यिकी अनुमानों और सम्भावनाओं का विज्ञान है।" यह परिभाषा भी ठीक नहीं है क्योंकि इसमें अनुमान तथा सम्भावनाओं के एक पक्ष पर ही प्रकाश डाला गया है, अन्य पक्षों को इसमें स्थान नहीं मिलता।

कॉक्सटन (Croxton) तथा **कॉउडन (Cowden)** ने एक बहुत सरल तथा विस्तत परिभाषा इस प्रकार दी है : "सांख्यिकी को संख्यात्मक तथ्यों के संग्रहण, प्रस्तुतीकरण, विश्लेषण और निर्वचन से सम्बन्धित कहा जा सकता है।"

यह परिभाषा सांख्यिकी अनुसन्धान (Statistical Investigation) की चार अवस्थाओं का वर्णन करती है। (*i*) समंकों का संकलन, (*ii*) प्रस्तुतीकरण, (*iii*) विश्लेषण एवं (*iv*) निर्वचन।

उपर्युक्त परिभाषा में व्यवस्था (Organisation) का समावेश नहीं है। यह कहना अधिक उचित होगा कि सांख्यिकी समंकों के एकत्रीकरण उनकी व्यवस्था, प्रस्तुतीकरण विश्लेषण तथा निर्वचन का विज्ञान है।

(1) समंकों का संकलन (Collection of Data) : सांख्यिकी अनुसन्धान में समंकों का संकलन सबसे प्रमुख कार्य है। समंक सांख्यिकी विश्लेषण तथा निर्वचन की आधारशिला की तरह होते हैं। यदि समंक अशुद्ध अपर्याप्त हैं तो उनसे निकाले जाने वाले निष्कर्ष भी भ्रमात्मक होंगे। समंक दो प्रकार के होते हैं — प्राथमिक व द्वितीयक। प्राथमिक समंक अनुसन्धानकर्ता द्वारा आरम्भ से अन्त तक नये सिरे से एक किये जाते हैं। द्वितीयक समंक पहले से ही अन्य व्यक्तियों व संस्थाओं द्वारा एकत्रित एवं प्रकाशित किये जा चुके होते हैं, अनुसन्धानकर्ता तो केवल उनका प्रयोग करता है। चाहे समंक प्राथमिक हो या द्वितीयक, उनके संकलन व प्रयोग में बहुत सावधानी की आवश्यकती होती है।

(2) एकत्रित समंकों को व्यवस्थित करना (Organising Collected Data) : समंकों को व्यवस्थित करने में सबसे पहला कार्य सम्पादन का है। सम्पादन का कार्य बहुत ही ध्यानपूर्वक किया जाना चाहियें ताकि अशुद्धियों, अतर्कपूर्ण और अनावश्यक तथ्यों को अलग किया जा सके। जब समंकों का सम्पादन कार्य पूरा हो जाता है तब उनका वर्गीकरण किया जाता है। वर्गीकरण इस मुख्य उद्देश्य एक ही विशेषता वाले समंकों के साथ रखना होता है। इन समंकों के आधार पर सारणी बनाई जाती है। सारणी बनाने का उद्देश्य समंकों तथा पंक्तियों में बाँटना है ताकि प्रस्तुत समंक आसानी से समझ में आ सकें तथा सांख्यिकीय विश्लेषण के लिए प्रयोग किये जा सकें।

(3) **व्यवस्थित समंकों को प्रस्तुत करना (Presenting the Organised Data) :** समंकों को एकत्र करने तथा व्यस्थित करने के पश्चात् प्रस्तुत किया जाता है। ठीक प्रकार से प्रस्तुत समंक सांख्यिकी विश्लेषण में बहुत सहायक होते हैं। एकत्रित समंकों को प्रस्तुत करने की तीन विधियाँ हैं :-

(*i*) सांख्यिकी, सारणी

(*ii*) चित्र, तथा

(*iii*) रेखाचित्र

(4) **समंकों का विश्लेषण (Analysis of Data) :** समंकों को एकत्र, व्यवस्थित तथा प्रस्तुत करने के पश्चात उनका विश्लेषण किया जाता है इस पुस्तक में अधिकांश समंकों के विश्लेषण की रीतियों का ही वर्णन किया गया है। विश्लेषण रीतियाँ सरल से सरल एवं कठिन से कठिन हैं लेकिन इस पुस्तक में केवल उन्हीं विश्लेषण रीतियों का उल्लेख किया गया है जो साधारणतया प्रयोग में आती हैं, जैसे केन्द्रीय प्रवत्ति के माप , विचरण के माप, सहसम्बन्ध, प्रतीपगमन आदि।

(5) **समंकों का निर्वचन (Interpretation of Data) :** सांख्यिकी अनुसंधान में अंतिम कार्य निर्वचन करना अर्थात एकत्रित या विश्लेषण समंकों से निष्कर्ष निकालना है। निर्वचन का कार्य बहुत की कठिन होता है जिसके लिये उच्चकोटि के अनुभव एवं योग्यता की आवश्यकता है। यदि विश्लेषित समंकों का ठीक ढंग से निर्वचन किया जाये तो अनुसंधान का उद्देश्य विफल हो जायेगा और निष्कर्ष अशुद्ध होंगे। उचित निर्वचन के द्वारा ही सही निष्कर्षों पर पहुँच कर ठीक निर्णय लिया जा सकता है।

स्मरण रहे कि सांख्यिकी अनुसंधान की पाँचों अवस्थायें एक-दूसरे से सम्बन्धित हैं। इसका विभाजन सांख्यिकी रीतियों के क्रमबद्ध अध्ययन में सहायता देता है।

सांख्यिकी की परिभाषा में नवीन प्रवत्ति
(New Trends in the Definition of Statistics)

पिछले कुछ वर्षों से सांख्यिकी विज्ञान में गणितीय सम्भावना सिद्धान्त का बहुत प्रयोग किया जा रहा है, इसलिये सांख्यिकी को 'अनिश्चितता में निर्णय लेने की कला व विज्ञान' माना जाने लगा है। निम्न परिभाषाओं से इसका स्पष्ट आभास होता हैः-

हैडले (Hadley) के शब्दों में, ''आधुनिक अर्थ में सांख्यिकी अनिश्चितता वाली किसी स्थिति के बारे में संकलित समंकों का प्रयोग करके उस स्थिति के विषय में ठोस निष्कर्ष निकाले की विधि है।''

वॉलिस तथा **राबर्ट** के अनुसार, ''सांख्यिकी अनिश्चितता की परिस्थितियों में बुद्धिमत्तापूर्ण निर्णय लेने की विधि है।''

सांख्यिकी निष्कर्ष शायद ही कभी पूर्ण सत्य होते हैं। जब उनका प्रयोग किया जाता है तो अनिश्चितता बनी रहती है। सांख्यिकी निष्कर्षों में अनिश्चितता का विचार करके हमें सांख्यिकी विज्ञान को ठुकराना नहीं चाहिये। जॉन आई. गिफिन ने ठीक ही कहा है : - ''यद्यपि सांख्यिकी ज्ञान अपूर्णता तथा अनिश्चितता की विशेषता से सदा युक्त रहता है, तथापि इस अपूर्ण तथा अनिश्चित ज्ञान को बेकार समझकर अस्वीकार नहीं करना चाहिये क्योंकि सांख्यिकी विधियाँ अपूर्णतया तथा अनिश्चितता के मापन की व्याख्या करती है।''

सांख्यिकी — विज्ञान अथवा कला
(Statistics — Science or Art)

सांख्यिकी विज्ञान है अथवा कला या दोनों ही, यह विवाद का विषय है विज्ञान का अभिप्राय एक क्रमबद्ध ज्ञान के भण्डार से है। विज्ञान कारणों एवं प्रभावों का अध्ययन करके वैज्ञानिक सिद्धान्तों एवं रीतियों का सामान्ययन करता हैः यह तथ्य का निष्पक्ष वर्णन करता है तथा पक्षपातपूर्ण निर्णयों को दूर करने का भरसक प्रयत्न करता है। विज्ञान उस प्रकाश-स्तम्भ (Light House) के समान है जो आते-जाते जलयानों को अपना मार्ग ढूँढ़ने में सहायता तो करता है लेकिन उन्हें यह नहीं बताता कि वे किसी ओर जायें। प्रकाश की सहायता से जलयानों को स्वयं ही अपना मार्ग खोजना पड़ता है।

कार्ल पियर्सन के मतानुसार वह ज्ञान जो (*i*) नागरिकों को मानसिक शिक्षा दे, (*ii*) महत्वपूर्ण सामाजिक समस्याओं पर प्रकाश डाले, (*iii*) व्यावहारिक जीवन में आनन्द प्रदान करें, (*iv*) हमारी कलात्मक भावनाओं को सन्तुष्टि प्रदान करें, विज्ञान कहा जा सकता है। सांख्यिकी में इन सभी बातों का समावेश है — अतः इसे विज्ञान माना जा सकता है।

सांख्यिकी विज्ञान अवश्य है परन्तु इसे भौतिकशास्त्र, रसायनशास्त्र, अर्थशास्त्र तथा समाजशास्त्र जैसा विज्ञान नहीं माना जा सकता। इसे वैज्ञानिक पद्धतियों का विज्ञान कहा जा सकता है क्योंकि यह अन्य विज्ञानों के सिद्धान्तों के प्रतिपादन में सहायता देती है। सांख्यिकी विज्ञान तथा अन्य विज्ञान में भी मूल अन्तर है। सांख्यिकी विधियाँ अन्य क्षेत्रों का ज्ञान प्राप्त करने में सहायक होती हैं इसलिये कहा जाता है कि ''सांख्यिकी की सहायता के बिना विज्ञान रूपी वक्ष में फल नहीं लगता, उसी प्रकार विज्ञान की सहायता के बिना सांख्यिकी की जड़े मज़बूत नहीं होती।''

कला किसी निश्चित उद्देश्य को पूरा करने का ढंग बताती है। यह व्यावहारिक ज्ञान है अर्थात् — विज्ञान यदि ज्ञान (Knowledge) है तो कला क्रिया है। कला के निम्न लक्षण होते हैं :-

(1) कला उन क्रियाओं का समूह है जिनके द्वारा समस्याओं का हल किया जाता है।

(2) कला तथ्यों का वर्णन ही नहीं करती बल्कि उसके गुण दोषों तथा लक्ष्यों को प्राप्त करने के उपाय भी बताती है।

(3) कला की साधना में विशेष निपुणता और अनुभव आदि की आवश्यकता होती है।

सांख्यिकी में सभी उपर्युक्त विशेषतायें पाई जाती हैं। सांख्यिकी में हम विभिन्न सांख्यिकीय रीतियों और नियमों का ही अध्ययन नहीं करते अपितु यह भी सीखते हैं कि निश्चित उद्देश्यों की पूर्ति के लिये व्यवहार में इनका प्रयोग किया प्रकार किया जाये। विभिन्न सांख्यिकीय रीतियों का प्रयोग करने के लिए निपुणता और अनुभव अनिवार्य है। अतः सांख्यिकी कला भी है।

यह कहना उचित होगा कि सांख्यिकी विज्ञान तथा कला दोनों ही है — टिप्पेट ने ठीक ही कहा है कि, ''सांख्यिकी विज्ञान तथा कला दोनों है - यह विज्ञान है क्योंकि इसकी रीतियाँ मौलिक रूप से व्यवस्थित हैं और उनका सर्वत्र प्रयोग होता है, यह एक कला है क्योंकि इसकी रीतियों का सफल प्रयोग पर्याप्त सीमा तक सांख्यिकी की योग्यता व विशेष अनुभव तथा उनके प्रयोग क्षेत्र, जैस अर्थशास्त्र के ज्ञान पर निर्भर होता है।''

सांख्यिकी के कार्य
(Functions of Statistics)

सांख्यिकी के निम्नलिखित मुख्य कार्य हैं :-

(1) तथ्यों को स्पष्ट रूप में प्रकट करना (To Present Facts in a Definite Form) : अंकों में कही गई बात विश्वसनीय होती हैं। साधारण कथनों को संक्षिप्त तथा स्पष्ट रूप से प्रकट करना सांख्यिकी का मुख्य कार्य है। अंकों द्वारा प्रकट तथ्य, अस्पष्ट कथनों की तुलना में कही अधिक विश्वसनीय होते हैं। सांख्यिकी तथ्यों को संक्षिप्त तथा स्पष्ट रूप से प्रस्तुत करके शीघ्रता से समझने योग्य बनाने में सहायक होती है। उदाहरण के लिये, यह वाक्य कि, ''भारत में चावल का उत्पादन बढ़ रहा है'' निरर्थक है क्योंकि इसमें स्पष्टता का अभाव है। लेकिन यह तथ्य कि भारत में चावल का उत्पादन 1990-91 में 74.1 m. tonnes था स्पष्ट सूचना प्रदान करता है। इसी प्रकार यह वाक्य कि ''भारत में बहुत बेकारी है'', भारत की जनसंख्या तेजी से बढ़ रही है'' वस्तुओं की कीमत तेजी से बढ़ रही हैं'' आदि कोई स्पष्ट सूचना नहीं देते क्योंकि जब तक समस्या अंकों में व्यक्त न की गई हो तब तक समस्या की वास्तविक स्थिति का ज्ञान नहीं होता।

(2) तथ्यों को सरल तथा संक्षिप्त रूप में प्रकट करना (To Present Facts in a Simple and Condensed Form) : सांख्यिकी तथ्यों को सरल तथा संक्षिप्त रूप में प्रस्तुत करती है। बिखरे हुए समंकों के अपार समूह से कोई परिणाम नहीं निकाले जो सकते। उदाहरणार्थ, एक कारखाने में काम करने वाले 20,000 मजदूरों के वेतन के समंक अपने आप में निरर्थक हैं लेकिन उनका प्रयोग औसत वेतन ज्ञात करने में किया जा सकता है। कुल वेतन को व्यक्तियों की संख्या से भाग देकर औसत वेतन ज्ञात होगा। मान लिजिये कारखाने में काम करने वाले मज़दूरों का औसत वेतन 1275 रू. है। वेतन समंकों को विस्तत अध्ययन के लिये विभिन्न वर्गों में इस प्रकार प्रस्तुत किया जा सकता है : -

मासिक वेतन (रू. में)	मज़दूरों की संख्या
800 से कम	5,280
800-1000	8,320
1000-1200	3,122
1200-1400	1,887
1400 से अधिक	1,391
योग	**20,000**

विभिन्न सांख्यिकी रीतियों जैसे वर्गीकरण, केन्द्रीय प्रवत्ति के माप , अपकिरण के माप, आदि का प्रयोग करके जटिल सांख्यिकी सामग्री को सरल बनाया जा सकता है जिसके आधार पर महत्वपूर्ण निष्कर्ष निकाले जा सकते हैं।

(3) **तथ्यों का तुलनात्मक अध्ययन (To Facilitate Comparison of Data) :** समंकों की पारस्परिक तुलना के बिना कोई परिणाम नहीं निकाले जा सकते। भारत में 1988-89 में प्रचलित कीमतों के आधार पर राष्ट्रीय आय 3875.2 करोड़ थी जो बढ़कर 1989-90 में 4252.4 करोड़ हो गई।

(4) **सिद्धान्तों को बनाने व जाँच करने में सहायता प्रदान करना (To help in formulating and testing hypothesis):** सिद्धान्तों को बनाने, उनकी जाँच करने तथा नये सिद्धान्तों को खोजने में सांख्यिकी विधियों का बहुत योग है। उदाहरण के लिये, कोई सिक्का निष्पक्ष है या नहीं, क्लोरोमाइसीन मलेरिया बुखार रोकने में समर्थ है या नहीं, मुद्रा की मात्रा कम करने से कीमतों की वद्धि रूकती है या नही , आदि-आदि परिकल्पानओं की जाँच सांख्यिकी के उचित उपकरणों द्वारा की जा सकती है।

(5) **व्यक्तिगत ज्ञान व अनुभव में वद्धि करना (To Enlarge Individual Knowlede and Experience) :** अन्य विज्ञानों की भाँति सांख्यिकी भी व्यक्तिगत व अनुभव में वद्धि करती है। सांख्यिकी के अध्ययन से विचारों को स्पष्टता और निश्चयात्मकता मिलती है, तर्क शक्ति का विकास होता है तथा समस्याओं के समाधान के लिये उचित दष्टिकोण स्वयं ही विकसित होने लगता है। सांख्यिकी का ज्ञान न होने पर अनेक विद्वान स्वयं ही अपने आप में कुछ कमी महसूस करते हैं। सत्य तो यह है कि अच्छे नागरिक के लिये सांख्यिकी का ज्ञान उतना की आवश्यक है जितना लिखने या बोलने की शक्ति।

(6) **अन्य विज्ञानों के नियमों की जाँच करना (To Test Laws of Other Sciences) :** सांख्यिकी विधियाँ विभिन्न विज्ञानों के विकास में महत्वपूर्ण योगदान प्रदान करती है। अधिकांश विज्ञानों के नियम निगमन प्रणाली पर आधारित होते हैं। जहाँ नियम निगमन प्रणाली के आधार पर नहीं बनाये जा सकते वहाँ सांख्यिकीय विधियों का प्रयोग किया जा सकता है।

(7) **पूर्वानुमान में सहायता प्रदान करना (To Help in Forecasting) :** व्यापारिक संगठनों की नीतियाँ तथा योजनायें कार्यान्वित करने से पहले ही बना ली जाती है। भविष्य की प्रवत्ति का ज्ञान योजनाएँ तथा नीतियाँ बनाने में बहुत ही सहायक सिद्ध होता है। सांख्यिकीय विधियाँ भविष्य में होने वाली घटनाओं के बारे में सही जानकारी देती हैं। उदाहरण के लिये, एक व्यापारी को यह जानने के लिए कि 1992 में कितना उत्पादन करना चाहिए इस बात का अनुमान आवश्यक है कि उस वर्ष में कितनी बिक्री की आशा है। वह बिक्री के बारे में कल्पना कर सकता है लेकिन श्रेष्ठ यह होगा कि वह गत वर्षों की बिक्री के समंकों का विश्लेषण करके अगले वर्ष की बिक्री का सांख्यिकी विधियों के प्रयोग द्वारा अनुमान लगाये।

(8) **उचित नीतियों को निर्माण में सहायता प्रदान करना (To Help in Formulating Suitable Policies) :** उचित नीतियों को निर्धारित करने में सांख्यिकी आधारभूत सामग्री प्रदान करती हैं। उदाहरण के लिए, इस बात का निर्णय कि ''1992 में भारत कितना चावल निर्यात करे'' भारत के आन्तरिक उत्पादन तथा माँग पर निर्भर करेगा। जब तक चावल के आन्तरिक उत्पादन तथा माँग का अनुमान नहीं लगाया जायेगा तब तक निश्चित रूप से यह कहना कि भारत से 1992 में कितना चावल निर्यात किया जाए कठिन है। अगर हमें यह ज्ञान हो कि चावल का उत्पादन 15 करोड़ टन होगा तथा माँग 13 करोड़ टन, तब हम आसानी से यह कह सकते हैं कि 2 करोड़ टन चावल निर्यात किया जा सकता है। वास्तव में, सांख्यिकीय तथ्य तथा विधियाँ उचित निर्णय लेने में बहुत सहायक होते हैं।

सांख्यिकी का महत्त्व
(Importance of Statistics)

सांख्यिकी का प्रयोग प्राचीन काल से ही चला आ रहा है। तब सांख्यिकी को राजनीतिक अंकगणित कहा जाता था। आधुनिक काल में सांख्यिकी की महत्ता इतनी बढ़ गई है और दिन-प्रतिदिन बढ़ती जा रही है कि सांख्यिकीय विधियों का प्रयोग उचित नीतियों के निर्माण में लगभग प्रत्येक क्षेत्र में किया जाता है : जैसे शासन प्रबन्ध, व्यवसाय एवं वाणिज्य, आर्थिक नियोजन, जीवनशास्त्र, वनस्पतिशास्त्र, ज्योतिषशास्त्र, भौतिकशास्त्र, रसायनशास्त्र, शिक्षा, औषधि, समाजशास्त्र, मनोविज्ञान, नक्षत्रविज्ञान तथा ऋतुविज्ञान। कुछ प्रमुख क्षेत्रों में सांख्यिकी की महत्ता का वर्णन नीचे दिया गया है।

शासन प्रबन्ध में महत्त्व
(Importance in Administration)

प्राची काल से ही सांख्यिकी का प्रयोग शासन-प्रबन्ध को सुचारु रूप से चलाने में होता रहा है। प्राचीन काल में सरकार का प्रमुख कार्य कानून-व्यवस्था को बनाये रखना तथा देश को विदेशी आक्रमणों से बचाना समझा जाता था। अधिकांश समंक फौजी शक्ति, कर, जनसंख्या, अपराध आदि से ही सम्बन्धित होते थे। आधुनिक काल में राजकीय क्रियाओं का अत्यधिक विकास हो चुका है। सरकार उन सब कार्यों में भाग लेती है जिनका सम्बन्ध देशवासियों के कल्याण में वद्धि करना हो। समंक तथा सांख्यिकीय विधियाँ मानव की कल्याण-वद्धि में बहुत सहायक सिद्ध हो रही हैं। आजकल सांख्यिकी केवल जनशक्ति , फौजी शक्ति अथवा जन्म तथा मत्यु के समंकों तक ही सीमित नहीं है अपितु प्रत्येक क्षेत्र में समस्याओं के समाधान हेतु इसका प्रयोग हो रहा है। सरकार के सभी मंत्रालयों तथा विभागों जैसे यातायात, रक्षा, रेलवे, खाद्य, वाणिज्य, डाक व तार, कृषि आदि के कार्य को कुशलतापूर्वक करने के लिए सांख्यिकी पर निर्भर रहना पड़ता है। उदाहरण के लिए, यातायात विभाग द्वारा दिल्ली की यातायात समस्या को तब तक हल नहीं किया जा सकता जब तक यह पता न हो कि कितनी बसें इस समय चल रही हैं, और कितनी बसों की आवश्यकता है। इन समंकों के आधार पर यह निर्णय लिया जा सकता है कि कितनी बसें और मँगाई जाएँ। युद्ध के समय भी सांख्यिकी की बहुत आवश्यकता पड़ती है क्योंकि दुश्मन की शक्ति का सही अनुमान लगाए बिना युद्ध में सफलता पाना असम्भव ही प्रतीत होता है।

सरकार द्वारा नियुक्त विभिन्न आयोगों व समितियों की रिपोर्ट समंकों पर ही आधारित होती हैं। सरकार को पुराने कानूनों में संशोधन करने के लिये तथा नये कानूनों का निर्माण करने के लिए भी सांख्यिकीय सामग्री का सहारा लेना पड़ता है। राष्ट्रीय स्तर पर अधिकांशतः सरकार द्वारा ही समंक एकत्रित किए जाते हैं। इस समंकों को सरकार के अनेक शोधकर्त्ता तथा शोध संस्थाएँ भी प्रयोग में लाती हैं।

व्यवसाय तथा वाणिज्य में महत्व
(Importance in Business and Commerce)

व्यापार के बढ़ते हुए आकार तथा प्रतिस्पर्धा ने व्यापारिक समस्याओं को और भी अधिक कठिन बना दिया है। प्राचीन काल में जब व्यापार का आकार छोटा होता था तो व्यापारी का व्यापार के क्षेत्र में सीधा सम्बन्ध होता था। एक छोटे व्यापार का मालिक स्वयं ही प्रबन्धक, हिसाब लेखक, विक्रेता अथवा क्रेता का कार्य करता था। उसके लिए अपने ग्राहकों से निजी सम्बन्ध रखना सम्भव था तथा व जान सकता था कि उसके ग्राहकों की क्या आवश्यकताएं हैं। व्यापार का आकार बढ़ने से व्यापारी के लिए हजारों, लाखों ग्राहकों से निजी सम्बन्ध रखना लगभग असम्भव सा हो गया है। प्रबन्ध-व्यवस्था एक विशिष्ट कार्य माना जाता है तथा प्रबन्धक को व्यापारिक कार्यों की योजना, उनकी व्यवस्था, निरीक्षण तथा नियन्त्रण करना पड़ता है। चूँकि ग्राहक से निजी सम्बन्ध रखना अत्यन्त कठिन हो गया है इसलिए व्यापार को पहले की तुलना में कहीं अधिक अनिश्चितताओं का समाना करना पड़ता है। साथ-साथ अब उत्पादन भविष्य की माँग को ध्यान में रखकर किया जाता है। इसलिए जब तक वस्तु की माँग का ठीक-ठीक अनुमान न लगा लिया जाए तब तक व्यापार में लाभोपार्जन आदि असम्भाव नहीं तो कठिन अवश्य है। अतः सफल निर्णय करने के लिए व्यापारी को सांख्यिकीय विधियों का प्रयोग करना चाहिए। व्यापार वास्तव में अनुमानों तथा सम्भावनाओं पर आधारित होता है। अतः व्यापार की सफलता अनुमानों की शुद्धता पर आधारित होती है।

व्यापारिक गतिविधियाँ मुख्य रूप से निम्नलिखित भागों में बाँटी जा सकती हैं :-

(1) उत्पादन (Production);

(2) विक्रय (Sales);

(3) क्रय (Purchases);

(4) वित्त (Finance);

(5) श्रम (Personnel);

(6) लेखा (Accouting);

(6) बाजार शोध, (Market Research)

सांख्यिकीय विधियों की सहायता से प्रत्येक क्षेत्र में उचित नीति निर्धारण के लिय संख्यात्मक सूचना प्राप्त की जा सकती है। उदाहरणार्थ मंडी की अनुसंधान करने वाले ग्राहकों की पसन्द जानने के लिए न्यादर्श विधि का प्रयाग करते हैं। इसी प्रकार माल की किस्म को ऊँचा रखने के लिए सांख्यिकीय गुण नियंत्रण का प्रयोग किया जाता है। बिक्री के संख्यात्मक तथ्यों को प्रस्तुत करने के लिये सांख्यिकीय सारणियों तथा चार्टों का प्रयोग किया जाता है। इसी प्रकार वस्तुओं की कीमतें निर्धारित करने के लिए समंकों से बहुत सहायता मिलती है।

व्यापारिक पूर्वानुमान की सहायता से व्यापारी संभावित घटनाओं का प्रभाव बड़ी शुद्धता से जान सकता है। आधुनिक व्यापार सांख्यिकीय विश्लेषण व नियन्त्रण के द्वारा ही व्यवस्थित किया जा रहा है। हमारे देश में वैज्ञानिक, प्रबन्ध आन्दोलन ने आंकड़ों को एकत्र करने तथा उनका ध्यानपूर्वक निर्वचन करने की आवश्यकता पर विशेष जोर दिया है।

यद्यपि सांख्यिकीय विधियाँ निर्णय लेने में सहायक हैं, तथापि वे व्यावाहारिक ज्ञान का स्थान नहीं ले सकतीं। एक व्यावहारिक व्यापारी को अपने व्यापार की तकनीकि विशेषताओं, व्यापार सम्बन्धी पूर्ण ज्ञान, व्यापारिक ज्ञान व सांख्यिकीय विधियों की निर्वचन योग्यता में सामंजस्य स्थापित करना चाहिये।

अर्थशास्त्र में महत्व

(Importance in Economics)

1890 में प्रसिद्ध अर्थशास्त्री प्रो. मार्शल ने कहा था कि ''समंक वे तण हैं जिनसे प्रत्येक अर्थशास्त्री की भांति मुझे भी ईटें बनानी पड़ती है।'' अर्थात् अर्थशास्त्र के भिन्न-भिन्न नियमों का आधार ही समंक हैं। अर्थशास्त्र सम्पत्ति के उत्पादन, वितरण उपभोग तथा आय की बचत और विनियोजिन सम्बन्धी जटिल ढांचे से है। सांख्यिकीय समंक तथा सांख्यिकीय विधियाँ आर्थिक समस्या को समझने व आर्थिक नीतियों के निर्धारण में बहुत सहायक है। वास्तव में ये अर्थशास्त्री की परीक्षणशाला के यंत्र-सयंत्र हैं। उदाहरण के लिए, क्या उत्पन्न करना है, कैसे उत्पन्न करना है और किसके लिए उत्पन्न करना है, आदि प्रश्न सांख्यिकीय आंकड़ों की सहायता से ठीक तरह हल किये जा सकते हैं। उपभोग सम्बन्धी समंकों की सहायता से हम पता लगा सकते हैं कि किस प्रकार समाज के भिन्न-भिन्न समुदाय के लोग अपनी आय को खर्च करते हैं। इस प्रकार के आंकड़ों से लोगों के रहन-सहन के जीवन-स्तर तथा कर योग्य क्षमता का पता लगाया जा सकता है। विनियम के क्षेत्र में हम माँग तथा पूर्ति पर आधारित मूल्यों के सिद्धान्त, उत्पादन की लागत, बैंक तथा साख-पत्र का अध्ययन करते हैं। किसी वस्तु की पूर्ति के घटने अथवा बढ़ने से उसके मूल्य पर क्या प्रभाव होगा या अधिकाधिक लाभ कमाने की दष्टि से एकाधिकारी कौन -सी कीमत लें, आदि प्रश्नों का उत्तर सांख्यिकी की सहायता से सरलता से दिया जा सकता है।

सांख्यिकीय विधियाँ केवल आर्थिक नीतियों के बारे में सहायक नहीं है, अपितु सफलताओं का मूल्यांकन भी उन पर निर्भर है। उदाहरण के लिए, यदि बढ़ती हुई जनसंख्या को रोकने के लिए परिवार नियोजन पर जोर दिया गया तो सांख्यिकीय विधियों द्वारा यह पता लगाया जा सकता है कि परिवार नियोजन कितना सफल रहा है। अर्थशास्त्र के क्षेत्र में सांख्यिकी का प्रयोग इतना बढ़ गया है। 1926 में आर. ए. फिशर ने कहा ''सांख्यिकी को अर्थशास्त्र की शाखा मान लेना एक दुखद मिथ्या भ्रम है।''

आजकल विश्व के लगभग सभी देश आर्थिक नियोजन का पालन कर रहे हैं। देश की आर्थिक योजना का निर्माण पर्याप्त तथा विश्वसनीय समंकों के आधार पर ही किया जा सकता है। समंकों से यह ज्ञात होता है कि देश की अल्पकालीन तथा दीर्घकालीन आवश्यकतायें क्या हैं तथा किन-किन समस्याओं को प्राथमिकता दी जाये। यदि योजनाएँ समंकों के आधार पर नहीं बनाई गई तो उनमें लाभ के स्थान पर हानि हो सकती है।

हमारे देश में पंचवर्षीय योजनाओं का प्रादुर्भाव 1951 से हुआ। देश में अभी तक सात पंचवर्षीय योजनाएँ पूरी हो चुकी हैं तथा आठवीं येाजना तैयार की जा रही है। इन योजनाओं का आधर समंक ही है। दुर्भाग्यवश भारतीय समंक स्वतंत्रता के 44 वर्ष बाद भी बहुत दोषों से युक्त हैं, जैसे अपर्याप्त, अविश्वसनीयता आदि। परिणामस्वरूप, योजनाओं में निर्धारित बहुत से अनुमान गलत सिद्ध हुए और हम लक्ष्य तक नहीं पहुँच सके। योजना आयोग ने यह स्वीकार किया कि, ''अपर्याप्त और अशुद्ध समंकों के आधार पर किया जाने वाला नियोजन, अर्थव्यवस्था के न होने से भी बुरा है।

सांख्यिकी तथा भौतिक विज्ञान

(Statistics and Physical Sciences)

भौतिक विज्ञान, विशेषतया खगोल विद्या, भूगर्भ-शास्त्र तथा भौतिक-शास्त्र आदि क्षेत्रों में सांख्यिकीय विधियों का प्रयोग सर्वप्रथम हुआ। जीव विज्ञान तथा सामाजिक विज्ञान में सांख्यिकी का जितना प्रयोग बीसवीं शताब्दी में हुआ है उतना भौतिक

विज्ञान में नहीं हुआ। अब भौतिक विज्ञान के क्षेत्र में, विशेषतया खगोल विद्या, रसायन-शास्त्र, इंजीनियरिंग भूगर्भ, मौसम विज्ञान आदि में सांख्यिकी का प्रयोग निरन्तर बढ़ रहा है।

सांख्यिकी तथा प्राकृतिक विज्ञान
(Statistics and Natural Sciences)

सभी प्राकृतिक विज्ञानों जैसे, खगोलविद्या, जीवशास्त्र, ऋतु विज्ञान, वनस्पति विज्ञान आदि में सांख्यिकीय विधियाँ बहुत ही उपयोगी सिद्ध हुई हैं। उदाहरण के लिये, किसी बीमारी का पता लगाने में डॉक्टर को शरीर का तापमान, नाड़ी की गति, आदि वास्तविक तथ्यों पर निर्भर करना पड़ता है। इस प्रकार यह पता लगाने के लिए कि अमुक औषधि ठीक है अथवा नहीं, डॉक्टर को परीक्षण करने पड़ेंगे। परीक्षणों की सहायता से यह पता लगाना पड़ेगा कि कितने रोगी उस औषधी से ठीक हुए। वनस्पति-शास्त्र में मौसम का पौधों पर प्रभाव, मिट्टी की किस्म, आदि परीक्षणों के लिये सांख्यिकी का अधिक प्रयोग होता रहा है। वास्तव में, ऐसा कोई क्षेत्र नहीं है जहाँ सांख्यिकी समंकों एवं विधियों का प्रयोग न होता हो।

सांख्यिकी की अनुसंधान में महत्ता
(Importance of Statistics in Research)

सांख्यिकीय अनुसंधानों के द्वारा ही ज्ञान के भंडार में वद्धि हुई है। उदाहरण के लिये , फसलों में कितनी उपज हुई तथा विभिन्न प्रकार की खाद तथा मिट्टी सम्बन्धी सांख्यिकीय विधियों के द्वारा की निश्चित एवं विश्लेषित किये जाते हैं। वास्तव में, ऐसा कोई भी अनुसंधान-कार्य नहीं है जहाँ सांख्यिकीय समंकों एवं विधियों का प्रयोग न होता हो।

सांख्यिकी के अन्य लाभ
(Statistics and Other Uses)

पिछले कुछ पष्ठों में यह वर्णन किया गया है कि सांख्यिकी का विभिन्न क्षेत्रों में किस प्रकार प्रयोग होता है। इसके अतिरिक्त सांख्यिकी कई संस्थाओं जैसे बैंकर्स, दलाल, बीमा कम्पनियों, लेखा-निरीक्षकों, सामाजिक कार्यकर्त्ताओं, मजूदर संगठनों, व्यापारिक संगठनों तथा वाणिज्य चैम्बरों आदि के लिये भी बहुत लाभदायक है। बैंक अधिकारी नकद रुपये की आवश्यकता का अनुमान लगाते हैं ताकि बैंक सुचारु रूप से चल सके। बीमा कम्पनियों की बीमे की दर भी मत्यु सम्बन्धी समंकों पर आधारित होती है।

सांख्यिकी के उपयोगों की उपर्युक्त सूची अन्तिम नहीं है। यह सूची तो केवल सुझाव मात्र है। वास्तव में, सांख्यिकी वह जो सांख्यिकी करती है। मानव-कल्याण वद्धि में सांख्यिकी एक प्रकार से कार्य -साधक सिद्ध हुई है तथा प्राणीमात्र की लगभग प्रत्येक समस्या का समाधान सांख्यिकी के द्वारा की सम्भव है। वास्तव में, सांख्यिकी का ज्ञान प्रत्येक अच्छे नागरिक के लिये आवश्यक है।

यह स्मरण रखना चाहिये कि सांख्यिकी की उपयोगिता यद्यपि सार्वभौमिक है तथापि प्रत्येक समस्या का समाधान उसकी स्थिति विशेष को ध्यान में रखकर ही करना चाहिये। ''बिना पर्याप्य सोच-विचार किये सांख्यिकी का प्रयोग भयानक सिद्ध होगा।''

सांख्यिकी की सीमायें
(Limitations of Statistics)

यद्यपि सांख्यिकी का उपयोग दिन-प्रतिदिन बढ़ता जा रहा है तथापि यह कोई जादू नहीं है जो प्रत्येक समस्या का समाधान कर सके। जब तक समंक सुव्यवस्थित ढंग से संग्रहीत व निर्वाचित न हों तब तक शुद्ध निष्कर्ष पर पहुँचने की सम्भावना बनी रहती है। इसलिये सांख्यिकी की सीमाओं तथा उसके दुरुपयोग के विषय में जानकारी नितान्त आवश्यक है। सांख्यिकी की कुछ सीमायें निम्नलिखित हैं :-

(1) **सांख्यिकी व्यक्तिगत इकाइयों का अध्ययन नहीं करती (Statistics does not deal with individuals) :** सांख्यिकी तथ्यों का समूह है। सांख्यिकीय निष्कर्ष सामूहिक व्यवहार का अध्ययन करने में सहायक होते हैं, समूह की व्यक्तिगत इकाइयों के विषय में जानकारी प्रदान नहीं करते। उदाहरण के लिये, यदि हम कहें कि किसी परीक्षा में विद्यार्थियों के औसत अंक 60 थे तो इसका तात्पर्य यह नहीं कि प्रत्येक विद्यार्थी ने 60 अंक प्राप्त किये — कुछ विद्यार्थियों ने

60 से कम अर्थात 0, 5, 25 आदि तथा कुछ ने 60 से अधिक अर्थात 70, 78, 90 आदि अंक भी प्राप्त किये होंगे। सांख्यिकी में व्यक्तिगत इकाइयों का अध्ययन नहीं किया जाता जैसे किसी एक व्यक्ति की आय, भार आदि।

(2) सांख्यिकी गुणात्मक तथ्यों का अध्ययन नहीं करती (Statistics does not deal with qualitative phenomenon): सांख्यिकी केवल संख्यात्मक तथ्यों का ही अध्ययन करती है गुणात्मक तथ्यों का नहीं। इस प्रकार के गुणात्मक तथ्यों जैसे ईमानदारी, कुशलता, बुद्धिमता, अन्धत और बहरापन, आदि का अध्ययन प्रत्यक्ष रूप में नहीं हो सकता। इन समस्याओं का विश्लेषण तभी हो सकता है जब हम उनको संख्यात्मक रूप से प्रस्तुत करें। उदाहरण के लिये, छात्रों की बौद्धिक योग्यता का अनुमान परीक्षा में प्राप्त अंकों के आधार पर किया जा सकता है।

(3) सांख्यिकीय परिणाम केवल औसतन सत्य होते हैं (Statistical results are ture only on an average) : सांख्यिकी के नियम केवल औसत रूप से ही सत्य होते हैं। भौतिकशास्त्र और रसायनशास्त्र जैसे विज्ञानों के नियमों की भाँति सांख्यिकी के नियम प्रत्येक दशा में लागू नहीं होते। उदाहरण के लिये, यदि यह कहा जाये कि भारत में पुरुषों की औसत आयु 52.5 वर्ष है, इसका यह अर्थ यह नहीं कि प्रत्ये पुरुष 52.5 वर्ष तक जीवित रहता है, अपितु कुछ पुरुष सौ साल तक जीवित रहते हैं तथा कुछ पैदा होते ही मर जाते हैं। अतः यह बात औसत रूप से ठीक है।

(4) सांख्यिकीय रीति किसी समस्या के अध्ययन की विभिन्न रीतियों में से एक है (Statistics in only one of the methods of studying a problem) : एक समस्या के समाधान अथवा अध्ययन के लिये सांख्यिकी प्रत्येक परिस्थिति में सर्वश्रेष्ठ हल नहीं बताती। प्रायः किसी एक विशेष समस्या का उस देश की सभ्यता, धर्म तथा दर्शन-शास्त्र के अनुसार की अध्ययन किया जाता है। सांख्यिकी इस प्रकार की समस्यायें हल करने में असमर्थ है। इसलिये सांख्यिकी द्वारा प्राप्त निष्कर्ष की अन्य तथ्यों द्वारा पुष्टि कर लेनी चाहिये।

(5) सांख्यिकी का दुरुपयोग सम्भव है (Statistics can be misused) : सांख्यिकी की सबसे प्रमुख सीमा यह है कि इसका सरलता से दुरुपयोग हो सकता है। सांख्यिकी के द्वारा सही अथवा गलत बातों की पुष्टि की जा सकती है। इसके अतिरिक्त, कोई भी व्यक्ति सांख्यिकी का उचित प्रयोग नहीं कर सकता — इसके लिये अभ्यास कुशलतापूर्वक निष्कर्ष निकालने की योग्यता अनिवार्य है, अन्यथा निष्कर्ष गलत होने की सम्भावना बहुत अधिक रहती है। अदक्ष (inexpe-rienced) व्यक्ति सांख्यिकी का उचित प्रयोग नहीं कर सकता।

(6) सांख्यिकी केवल साधनमात्र है, समस्या का समाधान नहीं (Statistics only provides a means and not a solution to the problem) : सांख्यिकी किसी समस्या के समाधान में केवल एक साधन मात्र है — समस्या का समाधान नहीं है। निष्कर्ष निकालते समय पक्षपातरहित तथा स्वार्थहीन होना अत्यन्त आवश्यक है अन्यथा सांख्यिकी जो कच्ची सामग्री के रूप में समंकों को प्रस्तुत करती है, दुरुपयोग किया जा सकता है।

(7) सांख्यिकी का प्रयोग वही व्यक्ति कर सकता है जिसे सांख्यिकीय रीतियों का पूर्ण ज्ञान हों जो उन रीतियों के प्रयोग से भली-भाँति परिचित हो और उनमें दक्ष हो। सांख्यिकी के क्षेत्र में दक्षता न होने पर समंकों का प्रयोग करना व उनका विश्लेषण करके उनसे निष्कर्ष निकालना अत्यन्त खतरनाक होता है क्योंकि ऐसे निष्कर्ष प्रायः अशुद्धियों से परिपूर्ण तथा भ्रमात्मक होते हैं।

❑❑

अध्याय - 2

समंकों का संकलन, अर्थ, प्रकार, संग्रह (Statistical Data-Meaning and Types, Collection and Rounding of Data)

समंक-संकलन एक महत्वपूर्ण क्रिया है क्योंकि अनुसन्धान के लिये समंक मूलाधार हैं और उसकी सफलता अथवा असफलता पूरी तरह से समंकों पर निर्भर करती है। समंकों की पर्याप्तता और उसमें शुद्धता का होना भी जरूरी है ताकि उनसे सही निष्कर्ष प्राप्त किये जा सकें। अतः यह कहना अनावश्यक न होगा कि समंकों का संकलन करते समय अत्यन्त सावधानी, सतर्कता, दढ़ता , विश्वास, निष्पक्षता और धैर्य से काम लिया जाना चाहिये ताकि समंकों के रूप में एकत्रित कच्ची-सामग्री कहीं अशुद्ध व अविश्वसनीय न हो जाये। स्मरण रहे, अविश्वसनीय एवं अशुद्ध समंकों से भ्रमात्मक निष्कर्ष निकलने की सम्भावना सदैव बनी रहती है।

प्राथमिक तथा द्वितीयक समंक (Primary and Secondary Data)

समंक दो प्रकार के होते हैं (1) प्राथमिक तथा (2) द्वितीयक समंक

(1) **प्राथमिक समंक (Primary Data) :** प्राथमिक समंकों से अभिप्राय उन समंकों से है जिन्हें अनुसंधानकर्त्ता स्वयं अपने किसी विशिष्ट उद्देश्य से एकत्र करता है। दूसरे शब्दों में, वे समंक, जो किसी सांख्यिकीय अनुसन्धान के लिये, किसी अनुसन्धानकर्त्ता या एजेन्सी द्वारा मौलिक रूप से पहली बार एकत्र किये जाते हैं प्राथमिक समंक कहलाते हैं। स्पष्ट है कि प्राथमिक समंकों के संकलनों की योजना मौलिक होती है और प्राथमिक समंक मौलिक अनुसन्धान का ही परिणाम होते हैं।

(2) **द्वितीयक समंक (Secondary Data) :** द्वितीयक समंक वे हैं जो पहले से ही अन्य व्यक्तियों या संस्थाओं द्वारा किसी उद्देश्य के लिए एकत्रित किये जा चुके होते हैं और अनुसंधानकर्त्ता केवल उनका प्रयोग करता है। अतः ऐसे समंक नये एवं मौलिक नहीं होते। अनुसंधानकर्त्ता उन्हें स्वयं एकत्र भी नहीं करता, बल्कि दूसरे लोगों द्वारा एकत्र समंकों का स्वयं के लिये प्रयोग करता है। उपरोक्त उदाहरण में यदि अनुसन्धानकर्त्ता, सरकार द्वारा संकलित औद्योगिक श्रम संबंधी समंकों का प्रयोग कर लेता है तो वे उसके लिये द्वितीयक समंक माने जायेंगे।

प्राथमिक तथा द्वितीयक समंकों में अंतर (Distinction Between Primary and Secondary Data)

उपरोक्त विवरण से स्पष्ट है कि प्राथमिक तथा द्वितीयक समंकों में अन्तर केवल अवस्था (Degree) या सापेक्षता (Relativity) का है, प्रकृति का नहीं। एक ही प्रकार के समंक यदि एक व्यक्ति के लिये प्राथमिक हैं तो दूसरे के लिए द्वितीयक बन जाते हैं। सामान्यता समंक, जिस एजेन्सी द्वारा पहली बार एकत्र किये जाते हैं उसके लिये वे प्राथमिक सामग्री का कार्य करते हैं। किन्तु जब वही समंक किसी अन्य व्यक्ति द्वारा प्रयोग किया जाते हैं तो उसके लिये द्वितीयक सामग्री का रूप ले लेते हैं। प्राथमिक तथा द्वितीयक समंकों के बीच मुख्य अंतर इस प्रकार है :-

(1) प्राथमिक समंक मौलिक होते हैं और वे सांख्यिकीय विधियों के लिये कच्चे माल की भाँति हैं। इसके विपरीत द्वितीयक समंक मौलिक नहीं होते क्योंकि उन पर सांख्यिकीय यन्त्रों का एक बार प्रयोग हो चुका होता है। इसलिये वे निर्मित माल की भाँति होते हैं।

(2) प्राथमिक समंक रीति के अनुसार सम्पूर्ण क्षेत्र या समय में से एकत्र किये जाते हैं, जबकि द्वितीयक समंक अन्य व्यक्तियों या संस्थाओं द्वारा पूर्व-संकलित होते हैं।

(3) प्राथमिक समंकों में अधिक छान-बीन या संशोधन करने की आवश्यकता नहीं होती; जबकि द्वितीयक समंकों का प्रयोग करने से पूर्व उनकी शुद्धता की पूरी जाँच पड़ताल, और आवश्यकतानुसार उनमें संशोधन करने पड़ते हैं।

(4) प्राथमिक समंकों के संकलन में अधिक समय, परिश्रम व धन की आवश्यकता होती है क्योंकि योजना को नये सिरे से प्रारम्भ करना होता है। इसके विपरीत द्वितीयक समंकों को पत्र-पत्रिकाओं, सरकारी व गैर सरकारी प्रकाशनों से सिर्फ उद्धत कर लिया जाता है। इसलिये उन पर धन , समय व परिश्रम की बहुत कम आवश्यकता होती है।

प्राथमिक समंकों के संकलन की रीतियाँ

(Methods of Collecting Primary Data)

उपरोक्त विवेचन से यह स्पष्ट हो ही चुका है कि संकलन की समस्या मुख्य रूप से केवल प्राथमिक समंकों से ही सम्बद्ध है। प्राथमिक समंकों का संकलन करने के लिये मुख्यतः निम्नलिखित रीतियाँ प्रयोग में लाई जाती हैं :-

(1) प्रत्यक्ष व्यक्तिगत अनुसन्धान (Direct Personal Investigation) : इस रीति के अंतर्गत अनुसन्धानकर्त्ता स्वयं अनुसन्धान क्षेत्र में जाकर सम्बद्ध लोगों से सम्पर्क स्थापित करके सूचना प्राप्त करता है। इस रीति को 'प्रत्यक्ष' रीति इसलिये कहा जाता है क्योंकि जिस व्यक्ति के बारे में सूचना प्राप्त करनी होती है, वह सूचना स्वयं उसी व्यक्ति-विशेष से प्राप्त की जाती है, न किसी अन्य व्यक्ति से फिर, इस 'व्यक्तिगत अनुसंधान' इसलिए कहा जाता है क्योंकि अनुसन्धानकर्त्ता ''स्वयं'' क्षेत्र में जाकर सूचना प्राप्त करता है, न कि अपने सहायकों द्वारा। इस रीति का यूरोप में सर्वप्रथम प्रयोग श्री ली प्ले और भारत में आर्थर यंग ने किया था।

इस रीति का प्रयोग कब किया जाना चाहिये ?

(*i*) जहाँ अनुसन्धान का क्षेत्र सीमित हो।

(*ii*) जब समंकों की मौलिकता व शुद्धता की अधिक आवश्यकता हो।

(*iii*) जब अनुसंधान-योजना की सूक्ष्मता को देखते हुए अनुसन्धानकर्त्ता के वैयक्ति अनुभव, तीक्ष्ण बुद्धि व सतत निरीक्षण की आवश्यकता समझी जाये।

(*iv*) जब समंकों को गोपनीय रखना हो।

प्रणाली के गुण (Merits of the Method)

(*i*) **शुद्धता :** अनुसंधानकर्त्ता कार्य-क्षेत्र में स्वयं उपस्थित रहता है इसलिये उसके द्वारा प्रत्यक्ष सूचनाएँ निश्चित रूप से शुद्ध एवं मौलिक होती हैं।

(*ii*) **विश्वसनीयता :** अनुसंधानकर्त्ता स्वयं कार्य-क्षेत्र में उपस्थित रहता है और जाँच-कार्य के दौरान, यदि सूचना देने वालों के मन में कोई संदेह है तो दूर करके, सही-सही सूचना प्राप्त करने में सफल रहता है। इतना ही नहीं, अनुसन्धानकर्त्ता को इस बात का संदेह होने पर कि संसूचक जान-बूझ कर गलत सूचनाएँ दे रहा है, वह इसकी जाँच प्रश्नों को घुमा फिरा कर आसानी से कर सकता है।

(*iii*) **विस्तत सूचनाओं की प्राप्ति :** इस रीति का एक अन्य गुण, मुख्य सूचना के साथ-साथ अनेक सम्बन्धित सूचनाओं का सहज की प्राप्त हो जाना है। उदाहरणार्थ, श्रमिकों की आय-व्यय संबंधी जाँच करते समय, उनकी रोजगार-स्थिति रहन-सहन की दशाएँ, उन्हें प्राप्त सुविधाएँ आदि के बारे में अनेक उपयोगी सूचनाएँ प्राप्त की जा सकती हैं।

(*iv*) **सजातीयता :** चूँकि इस रीति के अंतर्गत सूचनाएँ एक ही व्यक्ति द्वारा एकत्र की जाती हैं इसलिये समंकों में एकरूपता बनी रहती है।

(*v*) **लोचदार प्रणाली :** यह रीति लोचदार है। अनुसन्धानकर्त्ता आवश्यकतानुसार प्रश्नों में थोड़ा बहुत परिवर्तन करके मनचाही सूचना प्राप्त कर सकता है।

(*vi*) **सार्थकता :** इस प्रणाली में सूचना देने वालों से अनुसन्धानकर्त्ता व्यक्तिगत रूप से सम्पर्क करता है, जिससे जाँच-कार्य काफी उत्साहवर्धक रहता है और फलतः उसकी सार्थकता बढ़ जाती है।

प्रणाली के दोष (Demerits of the Method)

(*i*) **सीमित क्षेत्र :** इस रीति का सबसे बड़ा दोष, विस्तत अनुसन्धान-क्षत्रों के लिये इसका अनुपयुक्त होना है।

(*ii*) **अपव्यय :** यह रीति अपव्ययी है क्योंकि इसमें धन, समय व परिश्रम अधिक लगाना पड़ता है।

(*iii*) **पक्षपात :** इस प्रणाली में अनुसंधानकर्त्ता के व्यक्तिगत पूर्वग्रह पक्षापात तथा सनक के कारण, परिणामों के दूषित होने की पूरी सम्भावना बनी रहती है।

(*iv*) **भ्रमात्मक निष्कर्ष :** कार्य-क्षेत्र के सीमित होने के कारण, यह सम्भव है कि प्राप्त समंक पूरे समग्र का सही प्रतिनिधित्व न करें और फलतः भ्रमात्मक निष्कर्ष प्राप्त हों।

(*v*) **अनुसन्धान की विषयगत प्रकृति :** इस प्रणाली की सफलता मुख्य रूप से अनुसन्धानकर्त्ता की बुद्धिमता, कार्य-दक्षता, चतुराई व दूरदर्शिता पर निर्भर करती है। यदि अनुसन्धानकर्त्ता में ये सभी गुण नहीं है तो फिर परिणामों के असंतोषजनक व अविश्वसनीय होने की सम्भावना बढ़ जाती है।

रीति के प्रयोग सम्बन्धी सावधानियाँ (Precautions) : इस रीति के प्रयोग करते समय निम्न बातों पर ध्यान देना अति आवश्यक है।

(*i*) अनुसंधानकर्त्ता अनुसंधान-कला में प्रवीण, व्यवहार-कुशल, परिश्रमी व धैर्यवान् हो।

(*ii*) अनुसंधानकर्त्ता उस क्षेत्र की विशेषताओं, रीति-रिवाजों तथा भाषा इत्यादि से पूरी तरह परिचित हो।

(*iii*) पूछे जाने वाले प्रश्न संक्षिप्त व सारगभित हो।

(*iv*) अनुसन्धानकर्त्ता को बिना पक्षपात व पूर्वाग्रह के, जाँच-कार्य करना चाहिये।

(*v*) उसे इस बात की पुष्टि कर लेनी चाहिये कि प्रश्न सही व्यक्ति से पूछे गए है और वह व्यक्ति प्रश्नों का उत्तर देने की योग्यता भी रखता है।

(2) अप्रत्यक्ष मौखिक अनुसन्धान (Indirect Oral Investigation) : इस रीति कें अन्तर्गत, समस्या से प्रत्यक्ष रूप से सम्बन्ध रखने वाले व्यक्तियों से सूचना प्राप्त नहीं की जाती बल्कि समस्या से अप्रत्यक्ष रूप से सम्बन्धित व्यक्तियों से मौखिक पूछ-ताछ करके सूचनाएँ एकत्रित की जाती हैं। इसलिये इसे 'अप्रत्यक्ष अनुसन्धान' कहते हैं।

प्रणाली का प्रयोग कब किया जाये :-

(*i*) यह रीति वहाँ प्रयोग की जाती है जहाँ क्षेत्र अत्यन्त विस्तत हो।

(*ii*) प्रत्यक्ष रूप से सूचना देने वाले लोगों से व्यक्तिगत सम्पर्क करना सम्भव न हो सके या अज्ञान, रुचिहीनता अथवा जान-बूझकर वह व्यक्ति सूचना देने में असमर्थ हो या ऐच्छिक रूप से असमर्थता प्रकट करे।

(*iii*) जब विषय से सम्बन्धित व्यक्तियों से प्रश्न करना उचित ही न समझा जाये, यह इसे गुप्त रखना हो या उनसे पक्षपातपूर्ण व्यवहार की आशा हो।

इस रीति का प्रयोग अधिकतर सरकारी स्तर पर नियुक्त तथा सीमितयों द्वारा किया जाता है। अपराधों की जाँच पड़ताल के मामले में भी इस रीति का अधिक प्रयोग किया जाता है।

प्रणाली के गुण (Merits of the Method)

(*i*) **मिव्ययिता :** यह प्रणाल मितव्ययी है क्योंकि इसमें समय, धन, परिश्रम आदि कम लगता है।

(*ii*) **सरल तथा सुविधाजनक :** प्रत्यक्ष व्यक्तिगत अनुसन्धान की तुलना में यह एक सरल सुविधाजनक प्रणाली है क्योंकि इसमें कार्य शीघ्रता से हो जाता है और अधिक परेशानी नहीं उठानी पड़ती।

(*iii*) **विस्तत क्षेत्र :** यह प्रणाली विस्तत क्षेत्र वाले इस अनुसन्धानों के लिये अधिक उपयुक्त है जहाँ सूचकों से प्रत्यक्ष सम्पर्क करना सम्भव न हो या लाभप्रद न हो।

(*iv*) **विशेषज्ञों की सम्मति प्राप्त होना :** इस प्रणाली का एक महत्वपूर्ण गुण यह है कि अनुसन्धान-प्रणाली पर विशेषज्ञों की राय तथा उनके सुझाव अनायास की प्राप्त हो जाते हैं जिससे अनुसन्धान को प्रभावी ढंग से पूरा करने में सहायता मिलती है।

(*v*) **निष्पक्षता :** इस प्रणाली के अन्तर्गत अनुसन्धानकर्ता के व्यक्तिगत पक्षपात की सम्भावना कम होती हैं, क्योंकि जाँच-कार्य प्रगणकों की सहायता से पूरा कराया जाता है।

प्रणाली के दोष (Demerits of the Method)

(*i*) **अशुद्ध परिणामों की सम्भावना :** इस रीति में सूचनाएँ, समस्या से परोक्ष सम्बन्ध रखने वाले लोगों से प्राप्त की जाती हैं जिससे परिणामों के अशुद्ध होने की सम्भावना अधिक रहती है। फिर, अनुसन्धानकर्त्ता को, प्रत्यक्ष निरीक्षण तथा व्यक्तिगत सम्पर्क के अभाव में पूरी तरह से प्रगणकों पर आश्रित होना पड़ता है, जिससे परिणाम अशुद्ध होने की संभावना और भी बढ़ जाती है।

(*ii*) **साक्षियों के दोष :** जिन साक्षियों से सूचनाएँ प्राप्त की जाती हैं उनकी थोड़ी सी लापरवाही या पक्षपात व अज्ञानता के कारण समंक दूषित हो जाते हैं। फिर, साक्षियों का गलत चयन, अनुसन्धान के परिणामों की और भी अशुद्ध बना देता है।

रीति के प्रयोग सम्बन्धी सावधानियाँ (Precautions) : इस रीति के प्रयोग में सम्भवतः सबसे अधिक सावधानी तथा सतर्कता बरतने की आवश्यकता होती है। इसलिये निम्न बातों को सदैव ध्यान में रखना चाहिये :-

(*i*) साक्षियों की बातों पर न तो पूर्ण विश्वास और न ही अविश्वास करना चाहिये।

(*ii*) साक्षियों की संख्या पर्याप्त होनी चाहिये अन्यथा निष्कर्ष गलत या अभिनत हो सकते हैं।

(*iii*) सूचाकों की मनोवत्ति तथा मनोविज्ञान का ध्यान रखना आवश्यक है। कुछ लोग जन्मजात जरूरत से ज्यादा आशावादी होते हैं तो कुछ लोग निराशावादी, जिसके लिये पहले से ही समायोजन कर लेना आवश्यक हे।

(*iv*) यह भी सोच लेना चाहिये कि साक्षियों को समस्या से सम्बद्ध सभी पहलुओं का पूर्ण ज्ञान नहीं है। परन्तु उनका विश्वास प्राप्त के लिये उनकी जानकारी पर अनावश्यक सन्देह करना भी ठीक नहीं।

(*v*) साक्षियों से पूछ-ताछ करते समय जाँचकर्ता को धैर्य, विनम्रता, चतुराई और निष्पक्षता से काम लेना चाहिये।

(3) स्थानीय स्त्रोतों से सूचना प्राप्त करना (Information through Local Sources or Correspondents) : इस रीति के अन्तर्गत अनुसन्धानकर्त्ता द्वारा विभिन्न स्थानों पर, स्थानीय व्यक्ति नियुक्त कर दिये जाते हैं जो समय-समय पर अपने अनुभवों के आधार पर अनुमानतः सूचनायें भेजते रहते हैं। उल्लेखनीय बात यह है कि ये सम्वाददाता सूचनाओं का संकलन अपने ही तौर-तरीकों, रुचि, निर्णय आदि के आधार पर करते हैं। इसलिये ऐसे समंकों में अशुद्धियों की सम्भावना अधिक होती है।

रीति का प्रयोग : इस रीति का प्रयोग अधिकतर समाचार-पत्रों, पत्रिकाओं और आकाशवाणी द्वारा किया जाता है। यह रीति ऐसे जाँच-विषयों के लिये ही अधिक उपयुक्त है जिनमें शुद्धता की कम आवश्यकता होती है।

प्रणाली के गुण (Merits of the Method)

(*i*) **विस्तत क्षेत्र :** यह प्रणाली विस्तत क्षेत्र के लिये अधिक उपयुक्त है।

(*ii*) **मितव्ययी :** यह प्रणाली धन, समय व परिश्रम की दष्टि से मितव्ययी है।

प्रणाली के दोष (Demerits of the Method)

(*i*) **शुद्धता व मौलिकता की कमी :** इस प्रणाली द्वारा संकलित समंकों में शुद्धता व मौलिकता की कमी होती है क्योंकि ये मुख्यतया अनुमानों पर आधारित होते हैं।

(*ii*) **एकरूपता का अभाव :** अलग-अलग सम्वाददाताओं द्वारा सूचना एकत्र करने के आधार व तरीके भिन्न-भिन्न होते हैं। फलस्वरूप समंकों में एकरूपता नहीं आ पाती।

(*iii*) **अभिनत निष्कर्ष :** सम्वाददाताओं की मनोवत्तियाँ और पूर्व -धारणायें समंकों को प्रभावित किये बिना नहीं रह पाती, जिससे निष्कर्ष अभिनत तथा एकांगी हो जाते हैं।

(*iv*) **विलम्बपूर्ण रीति :** इस प्रणाली में विलम्ब अधिक होता है। कभी-कभी सम्वाददाताओं द्वारा सूचनायें इतनी देर से भेजी जाती हैं कि विषय का महत्व ही समाप्त हो जाता है।

रीति के प्रयोग सम्बन्धी सावधानियाँ (Precautions)

(*i*) सम्वाददाताओं की नियुक्ति पूर्ण सतर्कता और सोच-समझकर की जानी चाहिये।

(*ii*) सम्वाददाताओं को अपनी व्यक्तिगत राय का कम-से-कम प्रयोग करना चाहिये।

(*iii*) एक निश्चित क्षेत्र में सम्वाददाताओं की संख्या पर्याप्त होनी चाजिये ताकि अभिनतियों की क्षतिपूर्ति होती रहे।

4. सूचकों द्वारा प्रश्नावली भरवाकर सूचना प्राप्त करना (Information through Questionnaire to be Filled in by Informants) : यह रीति को ''डाक-प्रश्नावली प्रणाली'' (Mailed Questionnaire Method) भी कहते है। इस रीति के अन्तर्गत सर्वप्रथम अनुसन्धानकर्त्ता जाँच से सम्बन्धित प्रश्नों की एक सूची (प्रश्नावली) तैयार करता है। फिर, एक-एक प्रति सूचकों के पास डाक भिजवा दी जाती है जो उनके उत्तर भरकर निश्चित तिथि तक वापिस लौटा देते हैं। प्रश्नावली के साथ एक अनुरोध-पत्र भी लगाया जाता है जिसमें सूचना एकत्र करने का उद्देश्य बताने के साथ-साथ सूचक को यह भी आश्वासन दिया जाता है उसके द्वारा दी गई सूचनायें पूर्णतया गुप्त रखी जायेंगी।

रीति के प्रयोग सम्बन्धी सावधनियाँ (Precautions) : यह रीति सरल अवश्य है किन्तु इसका प्रयोग करते समय भिन्न-भिन्न सावधानियाँ ध्यान में रखना आवश्यक है :-

(*i*) सूचना देने वाले व्यक्ति का सहयोग प्राप्त करने के लिये पहले उसे जाँच के उद्देश्य आदि के बारे में स्पष्ट जानकारी करा देनी चाहिये।

(*ii*) प्रश्नावली में निम्नलिखित गुण होने चाहियें : (*a*) प्रश्न मधुर भाषा हों, (*b*) प्रश्न, सरल, स्पष्ट तथा छोटे हों, (*c*) प्रश्न दो अर्थी नहीं होने चाहियें, (*d*) प्रश्नों की संख्या कम हो, (*e*) प्रश्नों की रचना इस प्रकार होनी चाहिये कि उनका उत्तर 'हाँ' या 'नहीं' में दिया जा सके, (*f*) प्रश्न ऐसे हो कि जिनकी गलत सत्यता की आड़ी-जाँच या प्रति-जाँच हो सके, (*g*) प्रश्न, शंका या विरोध उत्पन्न करने वाले न हों अर्थात् उनसे सूचना देने वाले व्यक्ति की मानसिक, धार्मिक, साम्प्रदायिक तथा सामाजिक भावनाओं को ठेस न पहुँचती हो।

(*iii*) प्रबन्ध इस प्रकार का हो कि सूचनायें शीघ्र से शीघ्र प्राप्त हो जायें।

(*iv*) सूचकों ने प्रश्नों का उत्तर बिना किसी पक्षपात या पूर्वाग्रह के दिया है, इसकी प्रतिदर्श-जाँच करा लेना सदैव हितकर होगा।

प्रणाली के गुण (Merits of the Method)

(*i*) **मौलिकता :** इस रीति में प्राप्त सूचनायें मौलिक तथा विश्वसनीय होती है क्योंकि सूचनायें स्वयं सूचकों द्वारा उपलब्ध करायी जाती हैं।

(*ii*) **मितव्ययिता :** यह एक मितव्ययी रीति है क्योंकि इसमें परिश्रम, समय व धन कम खर्च होता है।

(*iii*) **विस्तत क्षेत्र :** यह रीति विस्तत क्षेत्र के अनुसन्धानों के लिये अधिक उपयुक्त है।

प्रणाली के दोष (Demerits of the Method)

(*i*) **सीमित व्याप्ति :** यह रीति केवल शिक्षित समाज के लिये उपयुक्त है। चूँकि अशिक्षित व्यक्तियों से इस रीति द्वारा सूचनायें प्राप्त नहीं की जा सकती। फलतः सीमित व्याप्ति (Limited Coverage) इस रीति का प्रमुख दोष है।

(*ii*) **अपर्याप्त तथा अपूर्ण सूचना :** यह प्रायः देखने में आया है कि सूचक अधिकतर प्रश्नावलियाँ वापिस ही नहीं भेजते। फिर, कुछ प्रश्नावलियाँ अधूरी प्राप्त होती हैं क्योंकि सूचकों की उदासीनता, आलग्य या शंका के कारण

अनेक प्रश्नों के उत्तर भरे ही नहीं जाते अथवा अस्पष्ट व अबाधगम्य ढंग से दिये जाते हैं। स्मरण रहे, अधूरी सूचनायें अभिनति तथा विभ्रमों को जन्म देती हैं जिससे परिणाम भ्रमात्मक निकलते हैं।

(*iii*) **विश्वसनीयता की कमी :** इस रीति द्वारा प्राप्त समंक कई कारणों से कम विश्वसनीय होते हैं। प्रथम, कई बार सूचक जानबूझकर सही बात छिपा लेते हैं और गलत सूचना भर देते हैं। द्वितीय, यदि प्रश्नावली सावधानी से तैयार नहीं की गई है और पूछे गये प्रश्न अस्पष्ट व दो-अर्थी हैं तो प्राप्त सूचनाओं के दूषित होने की पूरी सम्भावना रहती है।

(*iv*) **प्रति-जाँच का अभाव :** इस रीति में प्रश्नावली स्वयं सूचकों द्वारा भरी जाती हैं। अतः दिये गये उत्तरों की प्रति-जाँच (Cross Checking) हेतु, अनुपूरक प्रश्न पूछे जाने की गुँजाइश बहुत कम होती है। फिर, सूचना देने वाले व्यक्ति के मन में यदि किसी प्रश्न-विशेष के बारे में सन्देह है तो उसके समाधान का कोई अवसर नहीं होता और फलस्वरूप सूचनायें गलत तथा अधूरी प्राप्त होती हैं।

(*v*) **अव्यावहारिकता :** यह रीति अव्यावहारिक भी है क्योंकि अधिकांशत सूचक अपने कुछ व्यक्तिगत मामलों, जैसे आय, सम्पत्ति, आदतें, परिवार आदि के बारे में लिखित रूप से सूचना देने के लिये तैयार नहीं होते।

5. **प्रगणकों द्वारा अनुसूचियां भिजवाना (Schedules Sent Through Enumerators) :** सूचकों द्वारा प्रश्नावली भरवाने में अनेक कठिनाइयाँ हैं और प्राप्त सूचनायें भी पर्याप्त, अपूर्ण, अधुरी व अशुद्ध होती हैं। इन कठिनाइयों को दूर करने के लिये यह रीति अपनाई जाती है। इस रीति के अनुसार नियुक्त किये गये प्रगणक, अपने-अपने क्षेत्र में अनुसूचियाँ लेकर घर-घर जाते हैं और सूचकों से व्यक्तिगत रूप से प्रश्न पूछकर, उनके उत्तर दर्ज कर लेते हैं। इस प्रकार इस रीति में और पहली वाली रीति में जहाँ थोड़ी-सी समानता है। (दोनों रीतियों में प्रश्नों के उत्तर स्वयं सूचकों से लिये जाते हैं, अन्य व्यक्तियों या साक्षियों से नहीं), वहाँ इन दोनों में आधारभूत अन्तर यह है कि पहली रीति में प्रश्नावली सूचक के पास डाक से भेजी जाती है और सूचक स्वयं उसे भरकर डाक से ही वापिस भेज देता है। इसके विपरीत इस रीति में तैयार की गयी अनुसूची को संचकों के पास प्रगणक स्वयं ले जाता है और उनसे प्रश्नों के उत्तर पूछ कर उन्हें अनुसूची से स्वयं भरता है।

रीति के प्रयोग सम्बन्धी सावधानियाँ (Precautions) : इस प्रणाली का प्रयोग करते समय निम्नलिखित बातों पर ध्यान देना आवश्यक है :-

(*i*) प्रश्नावली-सम्बन्धी सभी आवश्यक बातों, जिनका वर्णन इससे पहले वाली प्रणाली में किया गया है, का अनुसूची तैयार करते समय भी पालन किया जाना चाहिये।

(*ii*) इस रीति की सम्पूर्ण प्रगणकों के सही चयन और उनकी योग्यता पर निर्भर करती है। अतः प्रगणकों की नियुक्ति करते समय इस बात का ध्यान रखना चाहिये कि वे — (क) बुद्धिमान, निपुण, धैर्यवान, ईमानदान, परिश्रमी तथा व्यवहार-कुशल हों, (ख) भली-भाँति प्रशिक्षित हों, (ग) पक्षपात की भावना से अछूते हों, (घ) अनुसंधान-कार्य में रुचि रखते हो तथा (ड) उस क्षेत्र की भाषा, परम्पराओं, रीति-रिवाजों व परिस्थितियों से भली-भाँति परिचित हों।

(*iii*) प्रगणकों के लिये प्रशिक्षण की पूर्व व्यवस्था करनी चाहिये जिससे कि वे अनुसूची भरने में निपुण हो जायें और सम्भावित कठिनाइयों का उन्हें ज्ञान हो सके।

(*iv*) प्रगणकों के लिये प्रशिक्षण की पूर्व व्यवस्था करनी चाहिये जिससे कि वे अनुसूची भरने में निपुर्ण हो जायें और सम्भावित कठिनाइयों का उन्हें ज्ञान हो सके।

रीति का प्रयोग : इस रीति का प्रयोग प्रायः बड़े व्यापार गहों , सार्वजनिक उपक्रमों, शोध-संस्थाओं, सरकारी तथा गैर सरकारी संस्थाओं द्वारा किया जाता है। भारत में एक अत्यन्त महत्त्वपूर्ण समस्या, जनगणना अनुसन्धान इसी प्रणाली द्वारा किया जाता है।

प्रणाली के गुण (Merits of the Method)

(*i*) **विस्तत क्षेत्र :** यह प्रणाली विस्तत क्षेत्र वाले अनुसन्धानों के लिये सर्वाधिक उपयुक्त है।

(*ii*) **शुद्धता :** इस रीति द्वारा प्राप्त सूचनाओं में शुद्धता होती है क्योंकि अनुसन्धान-कार्य योग्य, प्रशिक्षित तथा अनुभवी प्रगणकों द्वारा किया जाता है।

(*iii*) **विश्वसनीयता :** इस रीति के अन्तर्गत प्रगणकों का सूचकों से व्यक्तिगत सम्पर्क होने के कारण प्राप्त सूचनायें अत्यधिक विश्वसनीय होती हैं।

(*iv*) **निष्पक्षता :** इस प्रणाली में व्यक्तिगत पक्षपात की सम्भावना बहुत कम होती है।

(*v*) **असीमित व्याप्ति :** इस रीति की व्याप्ति काफी विस्तत है क्योंकि 'डाक प्रश्नावली प्रणाली' के विपरीत इस रीति का प्रयोग अशिक्षित सूचकों वाले क्षेत्रों के लिये भी आसानी से किया जा सकता है।

प्रणाली के दोष (Demerits of the Methods)

(*i*) **महंगी प्रणाली :** यह एक मंहगी प्रणाली है जिससे समय तथा धन काफी खर्च होता है। इसलिये इस रीति का प्रयोग पर्याप्त वित्तीय साधनों वाली संस्थायें या फिर सरकार ही कर सकती है।

(*ii*) **जटिलता :** यह अनुसन्धान की एक जटिल रीति है। विशेष रूप से योग्य प्रगणकों का चयन, उनकी प्रशिक्षण-व्यवस्था व उनके कार्य का निरीक्षण आदि सरल कार्य नहीं है।

(*iii*) **स्वाभाविक अभिनति :** प्रगणकों के व्यक्तित्व में जन्मदाता अन्तर होने के कारण, उनके द्वारा संकलित सूचनाओं में भी अन्तर या विचरण का होना स्वाभाविक है, जिससे परिणाम सही नहीं निकलते।

उपयुक्त रीति का चुनाव : उपर्युक्त सभी प्रणालियाँ अपनी-अपनी दष्टि से उचित हैं और प्रत्येक प्रणाली के अपने लाभ व दोष हैं। हाँ कहना कि इसमें से कौन सी प्रणाली अधिक उचित है अत्यन्त कठिन है क्योंकि एक उपयुक्त रीति का चुनाव काफी हद तक (*i*) अनुसंधान की प्रकृति अर्थात् समस्या के स्वरूप, (*ii*) उद्देश्य एवं क्षेत्र, (*iii*) आर्थिक साधन, (*iv*) अभीष्ट शुद्धता की मात्रा तथा (*v*) उपलब्ध समय पर निर्भर करता है। इसलिये अलग-अलग प्रकार की समस्याओं के लिये अलग-अलग रीतियाँ ही उपयुक्त हो सकती हैं।

अनुसूची तथा प्रश्नावली (Schedule and Questionnaire)

प्राथमिक प्रकृति के अनुसन्धानों में आवश्यक सूचनायें अनुसूचियाँ भरवाकर प्राप्त की जाती हैं। यद्यपि अनुसूची तथा प्रश्नावली में अन्तर हैं, किन्तु व्यवहार में दोनों का एक ही अर्थ लिया जाता है। प्रश्नावली एक ऐसा फार्म या प्रपत्र है जिसमें अनुसन्धान विषय से सम्बन्धित अभीष्ट तथा विस्तत जानकारी प्राप्त करने हेतु प्रश्नों का क्रमानुसार तथा प्राथमिकतानुसार ब्यौरा होता है जिसका उत्तर सूचकों द्वारा दिया गया है।

उत्तम प्रश्नावली के गुण (Essentials of a Good Questionnaire)

सांख्यिकी अनुसंधान की सफलता मुख्य रूप से प्रश्नावली की उत्तमता पर निर्भर करती है। अतः इस दष्टि से एक प्रश्नावली की रचना करते समय निम्न बातों को विशेष रूप से ध्यान में रखना चाहिये :-

(1) **कम प्रश्न (Less Questions) :** जहाँ तक संभव हो सके प्रश्नों की संख्या कम रखनी चाहिये। प्रश्नों की आदर्श 15 से 20 मानी जाती है। हाँ ! प्रश्न इतने कम भी नहीं होने चाहियें कि पर्याप्त सूचना की प्राप्त न हो सके और उनकी संख्या इतनी अधिक भी नहीं होनी चाहिये कि सूचक को घबराहट महसूस हो।

(2) **उचित कम (Logical Sequence) :** प्रश्नों को उनके महत्व या प्राथमिकता के क्रम में रखना चाहिये और परस्पर सम्बन्धित प्रश्नों को आगे या पीछे न देकर, एक-ही स्थान पर क्रमबद्ध किया जाना चाहिए।

(3) **सरलता व स्पष्टता (Simplicity and Clarity) :** प्रश्न सरल व स्पष्ट होने चाहिएँ और वे लम्बे, जटिल तथा दो-अर्थों वाले नहीं होने चाहिएँ। फिर, अनिश्चितता उत्पन्न करने वाले शब्द जैसे 'शायद', 'अक्सर', 'कभी-कभी' आदि का प्रयोग कदापि नहीं करना चाहिये।

(4) **संक्षिप्तता (Conciseness or Concision) :** प्रश्न ऐसे होने चाहिएँ जिनका उत्तर 'हा' या 'नहीं' में दिया जा सके।

(5) **शिष्ट शैली (Courteous Style) :** प्रश्नों की भाषा तथा शैली मधुर व सम्मानजनक होनी चाहिये। असम्मानजनक शब्दों जैसे, 'चपरासी', 'नौकर' आदि का प्रयोग नहीं करना चाहिये।

(6) **आपत्तिजनक या वर्जित प्रश्न (Undesirable Questions) :** ऐसे प्रश्न कभी नहीं पूछने चाहियें जो सूचकों के आत्मसम्मान और उनकी धार्मिक व सामाजिक भावनाओं को ठेस पहुंचाते हों, अथवा उनसे सूचना देने वालों के मन में शंका, उत्तेजना या विरोध उत्पन्न होता हो।

(7) **प्रश्नों के स्वरूप (Types of Questions) :** राबर्ट वैलेस एवं एडवर्ड विलेट के अनुसार प्रश्नों का स्वरूप निम्न हो सकता है :-

(अ) **बन्द प्रश्न (Shut Questions) :** ऐसे प्रश्नों के सम्भाव्य उत्तर, अनुसन्धानकर्त्ता द्वारा स्वयं सुझाये जाते हैं और सूचक को उनमें से केवल एक को टिक (✓) करना होता है। बन्द प्रश्न भी कई प्रकार के हो सकते हैं :-

(*i*) सरल विकल्प वाले प्रश्न जैसे, हाँ या नहीं

(*ii*) त्रिविकल्प वाले प्रश्न जैसे, हाँ, नहीं, कुछ नहीं कह सकता

(*iii*) बहु-विकल्प वाले प्रश्न जैसे क्या आप धूम्रपान करते हैं ? हाँ, नहीं, कभी-कभी, कभी-कभार

(ख) **खुले प्रश्न (Open Questions) :** ये ऐसे प्रश्न होते हैं जिनका उत्तर सूचक को स्वयं अपने शब्दों में देना होता है। इन प्रश्नों को उद्देश्य सूचक के व्यक्तिगत विचार जानना है और इसलिये उसे उत्तर के रूप में कोई विकल्प नहीं बताया जाता। खुले प्रश्नों का प्रयोग कम-से-कम करना चाहिये क्योंकि बहुत कम लोग इनका उत्तर दे पाते हैं। उदाहरणार्थ — आप की राय में बरोजगारी को कम करने के क्या उपाय किये जाने चाहिएँ ? महँगाई किस प्रकार रोकी जा सकती है ?

(स) **विशिष्ट सूचना देने वाले प्रश्न (Specific Informations Questions) :** ये वह प्रश्न होते हैं जिनका उद्देश्य सूचकों से विशिष्ट जानकारी प्राप्त करना है। उदाहरणार्थ, आपकी आय कितनी है ? आपका शैक्षिक स्तर क्या है ? आपके बच्चे कितने हैं ?

(8) **सूचना प्रश्नों से बचना (To avoid leading questions) :** उदाहरण के तौर पर 'आप पौण्डस पाउडर ही क्यों इस्तेमाल होना चाहिये जिनसे उत्तरों की सत्यता की परस्पर जाँच की जा सके।

(9) **क्रास-जाँच (Cross Check) :** शुद्धता के उच्च-स्तर को बनाये रखने की दष्टि से प्रश्नावली में कुछ प्रश्नों का भी समावेश होना चाहिये जिनसे उत्तरों की सत्यता की परस्पर जाँच की जा सके।

(10) **पूर्व परीक्षण व संशोधन (Pre-testing and Rectification) :** जब प्रश्नावली तैयार हो जाये तो अनुसंधान कार्य प्रारम्भ करने से पूर्व उसका कुछ लोगों पर परीक्षण कर लेना चाहिये। प्रश्नावली का पूर्व-परीक्षण एक लाभप्रद क्रिया है क्योंकि इससे प्रश्नावली के दोष और प्रश्नों सम्बन्धी कठिनाइयों का पहले से ही ज्ञान हो जाता है।

(11) **निर्देश व टिप्पणी (Instructions and Footnote) :** प्रश्नावली में सूचक के लिये आवश्यकतानुसार निर्देश दे देने चाहियें। जैसे, प्रश्नावली कब कहाँ, और किस पते पर लौटानी है। इसके अलावा यदि किसी प्रश्न में व्याख्या की आवश्यकता है तो उसकी चिन्हित करके उसकी टिप्पणी फुटनोट के रूप में कर देनी चाहिये।

(12) **व्याख्या-पत्र या निवेदन पत्र (Covering Letter) :** जाँच के आयोजकों की तरफ से प्रश्नावली के साथ एक प्रपत्र लगा होना चाहिए जिसमें निम्न बातें शामिल हों — (*i*) अपना परिचय तथा अनुसन्धान का उद्देश्य, (*ii*) सूचकों को इस बात का आश्वासन कि उनकी सूचनाओं को गुप्त रखा जायेगा और न ही दुरुपयोग किया जायेगा, (*iii*) डाक प्रश्नावली रीति की स्थिति में पता लिखा लिफाफा संलग्न होना चाहिये, (*iv*) शीघ्र तथा सही उत्तरों के लिये, सूचकों को उपहार-कूपन्स आदि के रूप में प्रोत्साहन देने की व्यवस्था होनी चाहिए, (*v*) इच्छुक सूचकों को यदि वे माँग, तो जाँच रिपोर्ट की एक कापी देने का आश्वासन भी होना चाहिये।

प्रश्नावली का नमूना

(Model)

सुपर बाजार से वस्तुओं को खरीदने सम्बन्धी ''ग्राहकों की पसन्द'' का सर्वेक्षण करने के लिये एक उपयुक्त प्रश्नावली का उदाहरण निम्न हैं :-

सुपर बाजार सम्बन्धी ग्राहकों की पसन्द

नोट : कृपया नीचे दिये वर्ग में आवश्यक सूचनायें भरिये और अपने उत्तर को 'सही' का निशान लगाकर प्रकट कीजिए।

I सामान्य

1. नाम .... 3. आयु ....
2. पता .... 4. लिंग ..... स्त्री ☐ पुरुष ☐
5. व्यवसाय * नौकरी ☐ व्यापार ☐ अन्य ☐
6. मासिक आय 7. परिवार की सदस्य-संख्या

100 से 500 तक	☐	1 से 3 तक	☐
500 से 1000 तक	☐	4-6 तक	☐
1000 से ऊपर	☐	7 से ऊपर	☐

II विशिष्ट

1. आप सुपर बाजार क्यों जाते हैं ?
   (*i*) घर के निकट है। हाँ ☐ नहीं ☐
   (*ii*) घर के रास्ते में पड़ता है। हाँ ☐ नहीं ☐
2. सुपर बाजार की प्राथमिकता देने का क्या कारण है ?
   (*i*) सभी वस्तुओं की एक स्थान पर पूर्ति होना हाँ ☐ नहीं ☐
   (*ii*) माल की अच्छी किस्म हाँ ☐ नहीं ☐
   (*iii*) अधिक अच्छे डिजाइन हाँ ☐ नहीं ☐
   (*iv*) उचित व निश्चित मूल्य हाँ ☐ नहीं ☐
3. क्या आप सभी वस्तुयें सुपर बाजार से खरीदते हैं ? हाँ ☐ नहीं ☐
4. आप एक महीनें में कितनी बार सुपर बाजार जाते हैं ?
   1 से 6 दिन ☐ 6 से 10 दिन ☐ 19 से ऊपर ☐
5. वहाँ के 'सेल्समैन' का व्यवहार आप को कैसा लगता है ?
   (*i*) विनम्र ☐ (*iv*) रूखापन ☐
   (*ii*) सहायताशील ☐ (*v*) असहयोगी ☐
   (*iii*) कुशल ☐ (*vi*) अकुशल ☐
6. वस्तुयें प्राप्त करने में कोई कठिनाई तो नहीं होती ? हाँ ☐ नहीं ☐
7. क्या आप पुरुष सेल्समैन की बजाय महिला सेल्समैन पसन्द करते हैं ? हाँ ☐ नहीं ☐

द्वितीयक सामग्री : स्रोत और उसका प्रयोग
(Secondary Data : Sources and its Uses)

इस तथ्य को ऊपर स्पष्ट किया जा चुका है कि द्वितीयक समंक के समंक होते हैं जो पहले से ही किसी अन्य व्यक्ति या संस्था द्वारा एकत्रित किये गये होते हैं। एम. एस. ब्लेयर के अनुसार, द्वितीयक समंक वे हैं जो पहले से ही अस्तित्व में हैं और जो वर्तमान प्रश्नों के उत्तर के लिए न होकर, बल्कि किसी दूसरे उद्देश्य के लिए एकत्र किये गये हैं।'' इस प्रकार स्पष्ट है कि द्वितीयक सामग्री का संकलन नहीं किया जाता बल्कि उसको 'उद्धत' करके उसका केवल अपने उद्देश्य हेतु प्रयोग किया जाता है।

द्वितीयक समंकों के स्रोत
(Sources of Secondary Data)

द्वितीयक समंकों के निम्न दो स्त्रोत हो सकते हैं :-

(1) **प्रकाशित स्त्रोत (Published Sources) :** प्रायः देखने में आता है कि विभिन्न विषयों पर सरकारी व गैर-सरकारी संस्थाएँ अथवा व्यक्ति, प्राथमिक अनुसन्धान द्वारा प्राप्त समंकों को समय-समय पर प्रकाशित करते रहते हैं जिनका प्रयोग प्रायः अन्य लोगों द्वारा द्वितीयक सामग्री के रूप में किया जाता है। प्रकाशित समंकों के स्त्रोत इस प्रकार हैं :-

(*i*) **अन्तर्राष्ट्रीय प्रकाशन :** विदेशी सरकारों तथा अन्तर्राष्ट्रीय संस्थाओं द्वारा प्रकाशित समंक जैसे U.N. Statistical Year Book, Demographic Year Book; Annual Reports of I.M.F., I.L.O., I.B.R.D. etc.

(*ii*) **सरकारी प्रकाशन :** केन्द्रीय तथा राज्य सरकारों द्वारा भी विभिन्न विषयों से सम्बन्धित समंकों का प्रकाशन किया जाता है और इन समंकों में विश्वसनीयता की मात्रा भी अधिक होती है। उदाहरणार्थ, Five Year Plans (Draft and Progress Reports), R.B.I. Bulletin, Economic Survey (Annual); Census Reports etc.

(*iii*) **अर्द्ध-सरकारी प्रकाशन :** अर्द्ध-सरकारी संस्थाएँ जैसे नगरी-निगम, नगर पालिकाएँ, जिला परिषदें, पंचायतों आदि के द्वारा भी विभिन्न समस्याओं (जन्म स्मरण व स्वास्थ्य आदि) सम्बन्धी समंकों को प्रकाशन किया जाता है।

(*iv*) **आयोगों व समितियों के प्रतिवेदन :** सरकार द्वारा समय-समय पर नियुक्त किये गये आयोगों व समितियों की रिपोर्टों का भी प्रकाशन किया जाता है। उदाहरणार्थ — वित्त-आयोग, एकाधिकार आयोग, गोरवाला कमेटी, झा कमेटी, अलेक्जेन्डर समिति, हजारी समिति रिपोर्ट आदि।

(*v*) **अनुसन्धन व शोध संस्थाओं के प्रकाशन** : अनेक शोध संस्थायें समय-समय पर अपने शोध-परिणामों को प्रकाशित करती रहती हैं। उदाहरणार्थ — भारतीय सांख्यिकी अनुसन्धान (I.S.I.), राष्ट्रीय प्रतिदर्श जाँच संगठन (N.S.S.O), व्यवहारिक आर्थिक शोध की राष्ट्रीय परिषद (NCAER) इत्यादि।

(*vi*) **व्यापारिक व वित्तीय संस्थाओं के प्रकाशन :** व्यापारिक-वित्तीय संस्थाएँ तथा व्यावसायिक परिषदों जैसे FICCI, Trade Unions, Ahmedabad Mill Owners Association, Bank Bodies, Institute of Charted Accountants; Stock Exchanges, कुछ महत्वपूर्ण विषयों के प्रतिवेदन तथा समंक प्रकाशित करती हैं।

(*vii*) **समाचार पत्र तथा पत्रिकाएँ :** इसके अलावा समाचार-पत्र, सामयिक पत्रिकाएँ, मैगजीन्स आदि जैसे (Economic Times, Eastern Economists, The Financial Express, Indian Journal of Economics, Capital, Commerce) आदि द्वारा भी वर्तमान आर्थिक-सामाजिक विषयों पर महत्वपूर्ण समंक प्रकाशित किये जाते हैं।

(2) **अप्रकाशित स्त्रोत (Unpublished Sources) :** कुछ ऐसी सांख्यिकी सामग्री भी होती है जिसका विधिवत प्रकाशन तो नहीं हो पाता परन्तु फिर भी वह काफी उपयोगी होती है और उसका द्वितीयक सामग्री के रूप में प्रयोग किया जाता है। जैसे प्राध्यापकों द्वारा किये गये शोध-कार्य, कालिजों के रिकार्ड आदि।

द्वितीयक समंकों का प्रयोग करते समय सावधानियाँ
(Precautions in the Use of Secondary Data)

द्वितीयक समंकों की मुख्य समस्या उनके संकलन की न होकर, उनके सही और संशोधित रूप में प्रयोग करने की होती है। इसमें यह कभी नहीं भूलना चाहिये कि दूसरे लोगों द्वारा संकलित समंकों में त्रुटियों तथा सीमाएँ हो सकती हैं। ऐसे समंकों की आलोचनात्मक जाँच एवं विस्तत संपादन किये बिना , उनका प्रयोग करना सदैव खतरनाक होता है। अतः द्वितीयक समंकों का प्रयोग करने से पूर्व अनुसन्धानकर्त्ता को इस बात की पूरी जाँच कर लेनी चाहिये कि उपलब्ध समंक वर्तमान उद्देश्य के लिये (*i*) विश्वसीनयता, (*ii*) अनुकूलता तथा (*ii*) पर्याप्तता का गुण रखते हैं अथवा नहीं।

सावधानियाँ (Precautions) : द्वितीयक समंकों का प्रयोग करने से पूर्व उनके बारे में निम्न तथ्यों की जाँच-पड़ताल अवश्य कर लेनी चाहिये :-

(1) **पिछले अनुसन्धानकर्त्ता की योग्यता :** सर्वप्रथम यह पता लगाना चाहिए कि सामग्री का संकलनकर्त्ता क्या अनुभवी, योग्य, ईमानदार व निष्पक्ष व्यक्ति था। यदि उत्तर 'हाँ' में है तो उसका प्रयोग करना चाहिये अन्यथा नहीं।

(2) **उद्देश्य व क्षेत्र :** दूसरा उन समंकों के अनुसन्धान का उद्देश्य व क्षेत्र क्या था ? यदि वर्तमान और पिछले अनुसन्धान के उद्देश्य व क्षेत्र में काफी हद तक समानता है तो समंकों को प्रयोग करना चाहिये अन्यथा नहीं।

(3) **समंक संकलन की रीति :** इस बात की जाँच करना कि समंक संकलन की जो रीति पहले अपनाई गया थी, (*i*) वह कहाँ तक उपयुक्त थी, (*ii*) और वर्तमान प्रयोग के लिये कहाँ तक उपयुक्त तथा विश्वसनीय है। साथ ही, यदि प्रतिदर्श अनुसन्धान किया गया था तो क्या प्रतिदर्श, यथेष्ट था या अपर्याप्त था।

(4) **इकाई की उचित परिभाषा :** द्वितीयक सामग्री के प्रयुक्तकर्त्ता को इन बात की भी जाँच कर लेनी चाहिए कि पूर्व-अनुसंधान की इकाइयाँ, अर्थ, परिभाषा व समरूपता की दष्टि से वर्तमान -प्रयोग के अनुकूल है या नहीं।

(5) **अनुसंधान का समय व परिस्थितियाँ :** इस बात की भी जाँच कर लेनी चाहिए, कि उपलब्ध समंक किस काल व समय में तथा 'किन' परिस्थितियों में संकलित किये गये थे। स्मरण रहे, मन्दी काल में परिवर्तन होने पर समंकों की उपादेयता कम हो जाती है। अतः उनमें उचित संशोधन करके ही उनका प्रयोग किया जाना चाहिये।

(6) **शुद्धता की मात्रा का स्तर :** इस सम्बन्ध में एक सावधानी यह भी जरूरी है कि उपलब्ध सामग्री में पिछले अनुसन्धानकर्त्ता द्वारा शुद्धता का स्तर क्या रखा गया था ? शुद्धता की मात्रा कम होने पर समंक अविश्वसनीय हो जाते हैं और उनका प्रयोग कदापि नहीं करना चाहिए।

(7) **तुलना व जाँच :** अनुसन्धानकर्त्ता द्वारा समंकों की परीक्षात्मक-जाँच अवश्य की जानी चाहिये इसके दो तरीके हो सकते हैं। प्रथम, उपलब्ध समंकों में से कुछ की परीक्षात्मक-जाँच करके उनकी विश्वसनीयता का पता लगाया जाए। द्वितीय, यदि वर्तमान समस्या (उद्देश्य) से सम्बन्धित समंक, अपने स्त्रोतों से उपलब्ध है, तो उनकी परस्पर तुलना द्वारा शुद्धता की जाँच कर लेनी चाहिये। यदि उनमें काफी अन्तर है तो ऐसी दशा में नये सिरे से स्वयं अनुसन्धान करना अच्छा होगा, बजाय इसके कि द्वितीयक सामग्री पर भरोसा किया जाये।

समंकों का सम्पादन
(Editing of Data)

समंकों के संकलन के पश्चात उनका ठीक ढंग से सम्पादन अत्यन्त आवश्यक है। सम्पादन का मुख्य उद्देश्य त्रुटियों का संशोधन करना है। सम्पादन का काम ठीक रूप से करने के लिये बहुत अनुभव की आवश्यकता है अन्यथा समंकों का संग्रह करने में की गई मेहनत बेकार सिद्ध हो सकती है।

प्राथमिक सामग्री का सम्पादन
(Editing or Primary Data)

प्राथमिक सामग्री का सम्पादन करते समय निम्नलिखित बातों पर ध्यान देना चाहिये।

(*i*) समंक पूर्ण होने चाहिये।

(*ii*) समंकों में सामंजस्य होना चाहिये।

(*iii*) समंक शुद्ध होने चाहिये।

(*iv*) समंकों में सजातीयता होनी चाहिये।

(1) **पूर्णता की जाँच (Editing for Completeness) :** सम्पादक को देखना चाहिये कि प्रश्नावलियाँ तथा अनुसूचियाँ सब प्रकार से पूर्ण हों अर्थात प्रत्येक प्रश्न का उत्तर दिया गया हो। यदि कुछ महत्वपूर्ण प्रश्नों का उत्तर प्राप्त नहीं हुआ है तो सूचकों से सम्पर्क स्थापित करके अनुसूचियों को पूर्ण कर लेना चाहिये। सम्भवतः भरसक प्रयत्न के बाद भी कुछ प्रश्नों का सही उत्तर न मिला हो; ऐसी दशा में उस प्रश्नावली को छोड़ देना चाहिये।

(2) **संगति की जाँच (Editing for Consistency) :** सम्पादन करते समय यह देखना चाहिये कि प्रश्नों के उत्तर परस्पर विरोधी तो नहीं हैं। यदि उत्तर परस्पर विरोधी हों तो उस प्रश्नावली का दुबारा देखना अथवा जहाँ सम्भव हो सूचना देने वाले से सम्पर्क करके उससे सही उत्तर प्राप्त करने का प्रयत्न करना चाहिये।

(3) शुद्धता की जाँच (Editing for Accuracy) : निष्कर्षों की विश्वसनीयता सूचना की शुद्धता पर निर्भर करती है। यदि सूचना ही गलत है तो निष्कर्ष भी गलत होगें इसलिये सम्पादक को यह देखना आवश्यक है कि सूचना हर प्रकार से सही है। यदि अशुद्धता गणित सम्बन्धी है तो उसका आसानी से पता लगाया जा सकता है और उसको सही किया जा सकता है। यदि सूचना ही अशुद्ध है जैसे आयु के बारे में गलत सूचना तो इस गलती का पता लगाना व सही करना बहुत कठिन होगा।

(4) एकरूपता की जाँच (Editing for Uniformity) : सम्पादक को एकरूपता के लिये सारे प्रश्नों की जाँच करनी चाहिये। प्रायः प्रश्नावली में दिये गये उत्तरों में एकरूपता का अभाव होता है। उदाहरणार्थ, आय सम्बन्धी सूचना वित्तीय वर्ष के स्थान पर 'कलैण्डर वर्ष' या 'सम्वत वर्ष' के आधार पर दी जा सकती है। ऐसी परिस्थिति में यदि एकरूपता बनाये रखने के लिये समंकों का सम्पादन न किया जाये तो प्राप्त परिणाम भ्राम (भ्रामक) होंगे।

द्वितीयक समंकों के प्रयोग में सावधानियाँ

(Precautions in the Use of Secondary Data)

क्योंकि द्वितीयक समंक अन्य व्यक्तियों द्वारा संकलित होते हैं इसलिये यह आवश्यक है कि प्रयोग करने से पहले उनकी भली-भाँति जाँच कर ली जाये। वास्तव में उनके प्रयोग में बहुत ही सावधानी से काम लेना चाहिये। इस सम्बन्ध में प्रो. बाउले का कथन उल्लेखनीय है : ''प्रकाशित समंकों को बिना उनका अर्थ व सीमायें जाने जैसे का तैसा स्वीकार कर लेना खतरे से खाली नहीं है।'' ऐसे समंक कई कारणों से अशुद्ध हो सकते हैं, जैसे पक्षपातपूर्ण अपर्याप्त न्यादर्श का प्रयोग करते समय अनुसंधानकर्ता की निम्नलिखित बातें ध्यान में रखनी चाहिएँ :-

(1) क्या समंक अनुसंधान के उद्देश्य के अनुकूल हैं ? (Whether the data are suitable for the purpose of investigation) : समंकों को प्रयोग करने से पूर्व अनुसंधानकर्त्ता को देख लेना चाहिये कि समंक अनुसंधान के उद्देश्य के अनुकूल है। समंकों की उपयुक्तता का निश्चय अनुसंधान के स्वरूप तथा क्षेत्र को ध्यान में रखकर किया जा सकता है। उदाहरण के लिये, यदि उद्देश्य एवं क्षेत्र में भिन्नता है तो यह निर्भर करना पड़ेगा कि वर्तमान अनुसंधान में वह कहाँ तक उपयोगी सिद्ध हो सकते हैं। मान लीजिये अनुसंधान का उद्देश्य 'मजदूरी के वेतन' (भत्तों-सहित) का अध्ययन करना हैं। यदि संकलित समंक केवल मूल वेतन (Basic Wages) से संबंधित हैं तो उपर्युक्त अनुसंधान के उद्देश्य के अनुकूल नहीं हैं।

(2) क्या समंक अनुसंधान के उद्देश्य के लिये पर्याप्त हैं ? (Whether the data are adequate for the investigation) : समंक की उपयुक्तता का निश्चय करने के पश्चात यह देखना आवश्यक है कि क्या ये समंक अनुसन्धान के उद्देश्य के लिये पर्याप्त हैं। परीक्षण की आवश्यकता तथा भौगोलिक क्षेत्र के सन्दर्भ में समंकों की पर्याप्तता को निश्चित किया जायेगा। उदाहरण के लिए, यदि प्राप्त समंक केवल उत्तर प्रदेश की चीनी मिल तक ही सीमित है तो हम सारे भारत में चीनी मिलों के मज़दूरों वेतन का अध्ययन इन समंकों के आधार पर नहीं कर सकते। पर्याप्तता का प्रश्न समंकों के समय को ध्यान में रखकर भी निश्चित किया जा सकता है। उदाहरण के लिये, हम पिछले आठ-दस वर्षों की कीमतों का अध्ययन करना चाहें और हमारे पास प्राप्त समंक केवल दा या तीन वर्ष के लिये हैं तो ये समंक हमारे उद्देश्य के लिय अपर्याप्त होंगे।

(3) क्या समंक विश्वसनीय हैं ? (Whether the data are reliable) : समंकों की विश्वनीयता की जाँच कई दष्टिकोणों से की जा सकती है, जैसे किस संस्था द्वारा समंक एकत्रित किये गये, समंक संकलन में किसी रीति का प्रयोग हुआ; शुद्धता का स्तर क्या था, इत्यादि सब बातों से सम्बन्धित सूचना का पता लगाा सम्भव नहीं होता। अतः विश्वसनीयता की जाँच करना सबसे कठिन कार्य है।

द्वितीयक समंकों की विश्वनीयता, पर्याप्तता तथा उपयुक्तता की जाँच करने के लिये निम्न बातों को ध्यान में रखना चाहिये।

(1) उद्देश्य एवं क्षेत्र (Object and Scope) : यह जान लेना आवश्यक है कि क्या प्राथमिक रूप से जब समंक एकत्र किये गये थे जो अनुसन्धान का उद्देश्य व क्षेत्र था। अनुसन्धान के उद्देश्य व क्षेत्र, जिनके लिये अब उनका द्वितीयक समंक के रूप में प्रयोग किया जा रहा है, भिन्न थे तो समंक अनुपयुक्त होंगे।

(2) संकलन रीति (Method of Collection) : समंक संकलन के लिये कौन-सी रीति का प्रयोग किया गया था इसका ज्ञान आवश्यक है। यदि निदर्शन रीति का प्रयोग किया गया था तो इस बात का पता लगाना चाहिये कि न्यादर्श यथेष्ट

था और समग्र का ठीक प्रतिनिधित्व करता था। इस बात की भी जानकारी होनी चाहिये किस समंक संकलन के संगठन एवं निरीक्षण का उपयुक्त प्रबन्ध किया गया था या नहीं।

(3) **अनुसन्धानकर्त्ता व प्रगणकों की योग्यता (The ability and integrity of investigation and enumerators) :** क्या समंकों का संकलन किन्हीं विशेष तथ्यों को प्रमाणित करने के लिये पक्षपातपूर्ण ढंग से किया गया था ? यदि हाँ, तो ऐसे समंकों पर विश्वास नहीं किया जा सकता और उनका प्रयोग अनुचित होगा। उन्हीं समंकों का प्रयोग किया जाना चाहिये जो अनुभवी, निष्पक्ष, ईमानदार, अनुसन्धानकर्ता द्वारा संकलित किये गये हों।

(4) **अनुसन्धान में प्रयुक्त इकाइयों की परिभाषा (Definition of units) :** यह देख लेना चाहिये कि पूर्व अनुसन्धान में प्रयुक्त सांख्यिकीय इकाइयों के अर्थ वर्तमान प्रयोग के अनुकूल हैं या नहीं — यदि इकाइयों की परिभाषा में अन्तर है तो द्वितीयक समंकों का प्रयोग बिना आवश्यक समायोजन किये नहीं करना चाहिये।

(5) **विभिन्न स्त्रोतों से प्राप्त द्वितीयक समंकों की तुलना (Comparison of secondary data obtained from different sources) :** यदि एक ही विषय पर अनेक साधनों से द्वितीयक समंक उपलब्ध हैं तो उनकी सत्यता की जाँच करने के लिये उनमें तुलना कर लेनी चाहिये। यदि उनमें काफी अन्तर है तो सबसे अधिक विश्वसनीय स्त्रोत से प्राप्त समंकों का प्रयोग करना चाहिये।

(6) **समंक-संकलन का समय और परिस्थितियाँ (Time and conditions of collection of data) :** यह जान लेना आवश्यक है कि उपलब्ध सामग्री किस समय से सम्बन्धित है तथा किन परिस्थितियों में एकत्र की गई थी। यदि परिस्थितियों में बहुत अंतर है तो द्वितीयक समंकों का प्रयोग अनुचित होगा। उदाहरणार्थ, युद्ध काल में संग्रहीत समंकों का प्रयोग शान्तिकाल में नहीं करना चाहिये। इसी प्रकार अपवाद काल (Depression Period) में संकलित समंकों का उपयोग अभिवद्धि काल (Boom Period) में नहीं करना चाहिये।

(7) **शुद्धता का स्तर (Degree of Accuracy) :** द्वितीयक समंकों के प्रयोग से पूर्व यह जान लेना चाहिये कि प्रस्तुत समंकों में शुद्धता का स्तर क्या रखा गया था और उसे प्राप्त करने में कहाँ तक सफलता हुई। प्रकाशित समंकों की शुद्धता का स्तर उपयोगकर्त्ता द्वारा किये जाने वाले अनुसन्धान के स्तर से कम नहीं होना चाहिये।

(8) **परीक्षात्मक जाँच (Test Checking) :** अनुसंधानकर्ता को द्वितीयक समंकों में से कुछ की परीक्षात्मक जाँच कर लेनी चाहिये ताकि यह ज्ञात हो जाये कि वे कहाँ तक विश्वसनीय हैं।

यदि उपर्युक्त बातों को ध्यान में नहीं रखा जाएगा तो द्वितीयक समंकों के आधार पर जो निष्कर्ष निकाले जायेंगे वे अशुद्ध और अविश्वसनीय हो सकते हैं।

❑❑

अध्याय - 3

समंकों का वर्गीकरण तथा सारणीयन
(Classification and Presentation of Data)

वर्गीकरण
(Classification)

वर्गीकरण अध्याय में समंक संग्रह करने की विभिन्न रीतियों का विवेचन किया गया। संग्रहीत समंक अव्यवस्थित दशा में होते हैं और उनका विश्लेषण एवं निर्वचन बहुत कठिन होता है। जब तक इस सामग्री को व्यवस्थित करके सरल व समझने योग्य न बनाया जाये, उससे कोई निष्कर्ष नहीं निकाला जा सकता। उदाहरण के लिये, जनसंख्या सम्बन्धी समंकों (Census Data) से आयु, लिंग, वैवाहिक स्थिति, आदि पर एकत्र की गई सूचना व्यर्थ सिद्ध होगी, जब तक कि इसे संक्षिप्त तथा सुव्यस्थित ढंग से प्रस्तुत न किया जाये। यह तथ्य निम्न सारणी से स्पष्ट हे जायेगा :-

आयु श्रेणी (वर्षों में)	व्यक्तियों की संख्या (करोड़ में)
10 वर्ष से कम	6.5
10-20	8.0
20-30	10.0
30-40	15.0
40-50	13.0
50-60	11.0
60 से अधिक	5.3
	योग 68.8

उपर्युक्त सारणी 68.8 करोड़ व्यक्तियों की आयु केवल 7 श्रेणियों में प्रस्तुत करती है। इस सारणी द्वारा बहुत सुगमता से निष्कर्ष निकाले जा सकते हैं। इससे यह स्पष्ट होता है कि वर्गीकरण तथा सारणीयन में समंकों को इस प्रकार से प्रस्तुत किया जाता है कि वे सरलता से समझे जा सकें और उनका विश्लेषण एवं निर्वचन सुगम हो जाये।

वर्गीकरण का अर्थ

(Meaning of Classification)

समंकों का संकलन तथा सम्पादन करने के पश्चात् उनका वर्गीकरण किया जाता है। समंकों को समान गुणों के आधार पर व्यवस्थित करके वर्गों या विभागों में प्रस्तुत करने की क्रिया को वर्गीकरण कहते हैं। वर्गीकरण की क्रिया बहुत कुछ डाक-घर में पत्र छाँटने की क्रिया के समान है। डाक-घर में, एकत्रित पत्रों को, भौगोलिक आधार अर्थात् उनके गन्तव्य स्थानों, उनके पिन कोड, जैसे मुम्बई, मैसूर, पटना आदि के अनुसार छाँटा जाता हैं, इसके पश्चात उन्हें एक समान गुण के आधार पर जैसे एक ही गन्तव्य स्थान के पत्रों को एक थैले में अन्य स्थान के पत्रों को अन्य थैलों में रख देते हैं। ठीक इसी प्रकार वर्गीकरण की क्रिया है। विभिन्न मदों से समान गुणों के आधार पर व्यवस्थित किया जाता है।

वर्गीकरण के उद्देश्य

(Objects of Classification)

वर्गीकरण के मुख्य उद्देश्य निम्नलिखित हैं :-

(1) **सांख्यिकी सामग्री को सरल एवं संक्षिप्त बनाना (To Condense and Simplify the Collected Data) :** वर्गीकरण का मुख्य उद्देश्य सांख्यिकीय सामग्री की जटिलता को दूर करके उसे सरल व संक्षिप्त बनाना है। लाखों और करोड़ों समंकों को कुछ ही श्रेणियों में व्यवस्थित करके इस प्रकार प्रस्तुत किया जाता है कि वे आसानी से समझे जा सकें। उदाहरणार्थ, एक कारखाने में यदि 1,000 कर्मचारी काम करते हैं तो वेतन को इस प्रकार व्यक्त किया जा सकता हैः-

वेतन (रू. में)	कर्मचारियों की संख्या
800 से कम	50
800-1000	150
1000-1200	350
1200-1400	200
1400-1600	150
16001-1800	80
1800 से अधिक	20
	योग 1,000

(2) **समंकों की समानता व असमानता को स्पष्ट करना (To Bring Out Points of Similarities and Dissimilarities of Data) :** वर्गीकरण में सांख्यिकीय तथ्यों की समानता व असमानता स्पष्ट हो जाती है। समान गुण वाले मद एक साथ रखे जाते हैं, 'स्त्री', 'विवाहित', अविवाहित', 'साक्षर' आदि।

(3) **तुलना में सहायक होना (To Facilitate Comparison of Data) :** वर्गीकरण से समंकों का तुलनात्मक निर्वचन सरल हो जाता है। उदाहणार्थ, दो कक्षाओं के विद्यार्थियों द्वारा किसी एक में प्राप्त अंक आगे सारणी में दिये गये हैं।

इस सारणी से तुलनात्मक अध्ययन बहुत सुगम हो जाता है क्योंकि इसे देखकर यह कहा जा सकता है कि कक्षा 'ख' के विद्यार्थी कक्षा 'क' से अच्छे हैं।

विद्यार्थियों द्वारा प्राप्त अंक

अंक	कक्षा 'क'	कक्षा 'ख'
0-10	5	5
10-20	10	6
20-30	18	25
30-40	25	32
40-50	22	24
50-60	8	3
60-70	7	2
योग	**95**	**97**

(4) **सारणीयन का आधार प्रस्तुत करना (To Reveal the Basis of Tabulation) :** वर्गीकरण द्वारा सारणीयन तथा भावी विश्लेषण की अन्य क्रियाओं का समुचित आधार प्रस्तुत किया जाता है।

वर्गीकरण की मुख्य रीतियाँ
(Main Methods of Classification)

समंक मुख्यतः निम्नलिखित आधार पर वर्गीकृत किये जा सकते हैं :-

(1) भौगोलिक (Geographical), (2) समयानुसार (Chronological), (3) गुणात्मक (Qualitative), तथा (4) संख्यात्मक (Quantitative).

(1) भौगोलिक वर्गीकरण (Geographical Classification) : इस रीति के अंतर्गत समंकों का वर्गीकरण भौगोलिक आधार पर किया जाता है। उदाहरणार्थ, खाद्य उत्पादन की मात्रा भारत में राज्यानुसार इस प्रकार प्रस्तुत की जा सकती है :-

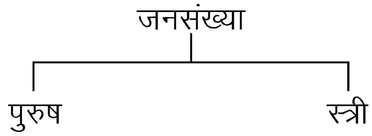
1989-90 में खाद्यान्न का उत्पादन

राज्यों के नाम	उत्पादन (हजारों टनों में)
हिमाचल प्रदेश	1410.9
कर्नाटक	7127.5
मध्य प्रदेश	14805.2
पंजाब	18985.6
तमिलनाडु	8124.4

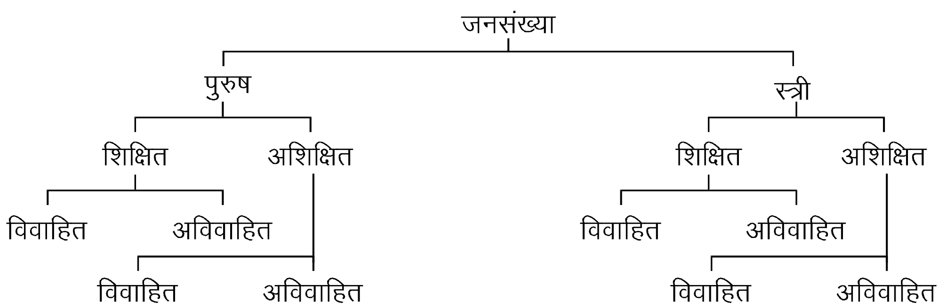
(2) समयानुसार वर्गीकरण (Chronological Classification) : जब समंक समय के अनुसार व्यवस्थित किये जायें तो समयानुसार इस प्रकार प्रस्तुत कर सकते हैं :-

वर्ष	भारत की जनसंख्या (करोड़ में)
1921	24.80
1931	27.60
1941	31.30
1951	35.70
1961	43.80
1971	54.90
1981	68.38
1991	83.44

(3) गुणात्मक वर्गीकरण (Qualitative Classification or Classification According to Attributes) : गुणात्मक वर्गीकरण में समंक गुणों अथवा विशेषताओं के आधार पर वर्गीकृत किये जाते हैं, जैसे — लिंग, आयु, व्यवसाय, साक्षरता, धर्म आदि। ध्यान देने योग्य यह बात है कि गुणात्मक समंकों की प्रत्यक्ष रूप से माप नहीं हो सकती। केवल यह पता लगाया जा सकता है कि प्रस्तुत अध्ययन में अमुक गुण विद्यमान है या नहीं। उदाहरण के लिये, अन्धता का अध्ययन करते समय यह ज्ञात कर सकते हैं कि एक निश्चित संख्या में कितने व्यक्ति अन्धे हैं लेकिन हम उनकी अन्धता की सीमा का पता नहीं लगा सकते। इस प्रकार हम केवल एक गुण या विशेषता का अध्ययन कर रहे हैं तो दो वर्ष या श्रेणियाँ बनेंगी — एक जिसमें वह गुण या विशेषता विद्यमान है, तथा दूसरी जिसमें वह गुण या विशेषता विद्यमान नहीं है, उदाहरणार्थ, जनसंख्या को लिंगानुसार निम्न दो भागों में विभाजित किया जा सकता है :-



इसी प्रकार जनसंख्या का वर्गीकरण, साक्षरता, धर्म आदि के आधार पर किया जा सकता है। जिस वर्गीकरण में केवल दो वर्ग या श्रेणियाँ होती हैं उसे साधारण (Simple) अथवा द्वन्द भाजन वर्गीकरण (Dichotomous Classification) कहते हैं। यदि समंकों को दो श्रेणियों के स्थान पर कई वर्गों या श्रेणियों में विभाजित करते हैं तो वह बहुगुणीय वर्गीकरण (Manifold Classification) कहलाता है। उदाहरण के लिये, हम जनसंख्या के पहले लिंगानुसार पुरुषों एवं स्त्रियों में विभाजित कर लें, फिर प्रत्येक वर्ग को साक्षरता के आधार पर शिक्षित अथवा अशिक्षित में बाँट लें। तत्पश्चात, अन्य गुणों के आधार पर उन्हें उपवर्गों में विभाजित कर दें। इस प्रकार यह क्रम किसी अन्य गुण अथवा विशेषता के आधार पर आगे बढ़ाया जा सकता है। बहुगुण वर्गीकरण का एक नमूना आगे दिया गया है।



(बहुगुणीय वर्गीकरण)

(4) वर्गान्तरों के अनुसार या संख्यात्मक वर्गीकरण (Quantitative Classification) : संख्यात्मक वर्गीकरण सामान्यतः वर्गान्तरों (Class Intervals) के अनुसार किया जाता है। सबसे छोटी और सबसे बड़ी संख्या को ध्यान में रखते हुए सभी समंकों को विद्यार्थियों द्वारा प्राप्त अंकों को निम्न रूप में प्रस्तुत कर सकते हैं :-

अंक	**विद्यार्थियों की संख्या**
0-20	8
20-40	12
40-60	25
60-80	15
80-100	10
	योग 70

ऐसे वर्गीकरण में दो तत्व होते हैं :-

(1) चर (Variable) तथा

(2) वर्ग-आवत्ति (Class Frequency)

(1) चर (Variable) : 'चर' शब्द का अभिप्राय उन तथ्यों से है जिनका अंकात्मक माप या प्रत्यक्ष माप सम्भव होता है और जो मात्रा अथवा आकार में घटते-बढ़ते रहते हैं। एक चर खण्डित या विच्छिन्न (Discrete) अथवा अखण्डित या

अविच्छिन्न (Continuous) हो सकता है। अखण्डित चर को एक निश्चित सीमा में गणितीय शुद्धता के साथ मापा जा सकता है। जब एक व्यक्ति की ऊँचाई 4 फुट से बढ़कर 5 फुट होती है तो वह 4 फुट से 5 फुट तक ही प्रत्येक ऊँचाई से गुजरेगा। लेकिन खण्डित चर में ऐसा नहीं होता। एक खण्डित चर कुछ निश्चित सीमाओं से ही गुजर सकता है। उदाहरणार्थ एक घर में कमरों की संख्या पूर्णांक में ही हो सकती है जैसे 1, 2 या 3 आदि, डेढ़, या ढाई संख्या नहीं हो सकती। इसी प्रकार, यदि एक घर में दो बच्चे और अन्य बच्चा जन्म लेता है तो कुल संख्या 3 होगी न कि $2\frac{1}{4}$, $2\frac{1}{2}, 2\frac{3}{4}$ आदि। व्यावहारिक दष्टि से अधिकांश चर अखण्डित होते हैं। अखण्डित एवं खण्डित आवत्ति वितरणों के दो उदाहरण नीचे दिये गये हैं :-

खण्डित आवत्ति विवरण		अखण्डित आवत्ति वितरण	
बच्चों की संख्या	परिवारों की संख्या	वजन (पौण्ड में)	व्यक्तियों की संख्या
0	10	100-110	10
1	40	110-120	15
2	80	120-130	40
3	100	130-140	45
4	250	140-150	20
5	150	150-160	4
6	50	160-170	2
योग	680	**योग**	136

(2) **वर्ग आवत्ति (Class Frequency) :** वर्ग बना लेने के पश्चात् यह जानना आवश्यक है कि इस समूह में से कितने मद किस वर्ग विशेष में आते हैं। इन मदों या अवलोकनों (Observations) की संख्या उस वर्ग की आवत्ति कहलाती है। ऊपर दिये गये उदाहरण में 10 व्यक्ति ऐसे हैं जिनका वजन 100-110 पौण्ड के बीच है। इस संख्या 10 को 100-110 वर्ग की आवत्ति कहते हैं। इस प्रकार पाँच -पाँच मदों के खण्ड (Block) बन जाते हैं। ऐसा करने से गणना में बहुत सहायता मिलती है। अन्त में हम इन रेखाओं की संख्या गिनकर मद के सामने आवत्ति वाले स्तम्भ में लिख देते हैं।

खण्डित या विच्छिन्न आवत्ति वितरण की रचना
(Formation of a Discrete Frequency Distribution)

इस प्रकार के वितरण का बनाना बहुत ही सरल है। इसमें हम केवल मदों की पुरावत्ति की संख्या गिनते हैं। इस संख्या की उस श्रेणी की आवत्ति कहते हैं। गिनने के लिये मिलान रेखाओं (Tally Bars) का प्रयोग किया जाता है : ''प्रत्येक वर्ग में आने वाले एक मद या इकाई के लिये एक खड़ी रेखा (।) उस वर्ग के सामने लगा दी जाती है। इस प्रकार पाँच-पाँच समूहों में आवत्ति की गणना करने से वर्गीकरण का कार्य सरल हो जाता है। इस रीति को अनुमेलन विधि (Four and Cross Method) कहते हैं। अंत में, इन रेखाओं को गिन लिया जाता है और वर्ग के सामने लिख दिया जाता है। इन संख्या को उस वर्ग की आवत्ति कहते हैं। निम्न उदाहरण से यह स्पष्ट हो जायेगी :

उदाहरण : एक कक्षा के 25 विद्यार्थियों द्वारा प्राप्त अंक इस प्रकार हैं :-

10	20	20	30	40	25	25	30	40	20	25	25	50
15	25	30	40	50	40	50	30	25	25	15	40	

एक विच्छिन्न आवत्ति वितरण की रचना कीजिये।

हल : आवत्ति वितरण की रचना

अंक	मिलान रेखायें (Tally Bars)	आवत्ति
10	|	1
15	||	2
20	|||	3
25	~~||||~~ ||	7
30	||||	4
40	~~||||~~	5
50	|||	3
		योग 25

इस सारणी से स्पष्ट है कि 1 विद्यार्थी ने 10 अंक प्राप्त किये तथा 3 विद्यार्थी ऐसे हैं जिनमें से प्रत्येक ने 20 अंक प्राप्त किये, 5 विद्यार्थी ऐसे हैं जिनमें से प्रत्येक ने 40 अंक प्राप्त किये, आदि। यह रीति उसी दशा में उपयोगी है जहाँ मदों की बहुत अधिक पुरावत्ति होती है अन्यथा सारणीयन को कोई विशेष लाभ नहीं होता। व्यवहार में इस रीति का बहुत कम प्रयोग होता है।

उदाहरण 2 : निम्न कथन के प्रत्येक शब्द में अक्षरों की संख्या लिखिये तथा विच्छिन्न (Discrete) आवत्ति वितरण की रचना कीजियेः-

"Without an adequate understanding of the statistical methods, the investigator in the social sciences may be like the blind man groping in dark room for a black cat that is not there. The methods of statistics are useful in a ever-widening range of human activities in any field of thought in which numericals data may be had".

हल : आवत्ति वितरण की रचना

प्रत्ये शब्द में अक्षरों की संख्या	मिलान रेखायें (Tally Bars)	आवत्ति
1	||	2
2	~~||||~~ ~~||||~~ ||||	14
3	~~||||~~ ~~||||~~ |||	13
4	~~||||~~	5
5	~~||||~~ |	6
6	||	2
7	~~||||~~	5
8	|||	3
9	|	1
10	||	2
11	|	1
12	|	2
13	|	1
		योग 57

अखण्डित या अविच्छिन्न वितरण की रचना

(Formation of a Continuous Frequency Distribution)

अखण्डित आवत्ति वितरण समंकों को प्रस्तुत करने की सबसे अधिक प्रचलित रीति है। इस प्रकार के वर्गीकरण में निम्न विशेष शब्दों का प्रयोग होता है :-

(1) वर्ग सीमायें (Class Limits) : एक वर्ग (Class) में दो सीमायें होती है। उनमें से एक को निम्न सीमा (Lower Limit) तथा दूसरी को उच्च सीमा (Upper Limit) कहते हैं। उदाहरण के लिये, 20-40 वर्ग में 20 निम्न सीमा तथा 40 उच्च सीमा है। कभी-कभी वर्ग सीमायें अनिश्चित-सी रहती हैं। इन्हें खुले सिरों वाली सारणी या विवर्तमुखी सारणी (Open-end Table) कहते हैं, जैसे नीचे उदाहरण से स्पष्ट होगा :-

वेतन (रू. में)	कर्मचारियों की संख्या
1000 से कम	10
1000-1200	60
1200-1400	120
1400-1600	50
1600 से अधिक	5
	योग 245

(2) वर्ग-विस्तार (Magnitude of Class-interval) : किसी वर्ग की उच्च तथा निम्न सीमा निम्न सीमा के अन्तर को वर्ग-विस्तार या वर्गान्तर कहते हैं। उदाहरणार्थ, 20-40 का वर्ग-विस्तार 20 है (40-20)। सूत्र के रूप में :-

वर्ग-विस्तार = उच्च सीमा (Upper Limit) – निम्न सीमा (Lower Limit)

(3) मध्य-मूल्य या मध्य-बिन्दु (Mid-value or Mid-point) किसी वर्ग की सीमाओं के मध्य स्थान को मध्य-मूल्य या मध्य-बिन्दु कहते हैं। इसे ज्ञात करने के लिये वर्ग की निम्न सीमा व उच्च सीमा को जोड़कर दो (2) से भाग देते हैं। सूत्र के रूप में :

$$\text{मध्य-बिन्दु} = \frac{\text{निम्न सीमा}\quad\text{उच्च सीमा}}{2} \left[\text{Mid Point} = \frac{\text{Lower Limit + Upper Limit}}{2}\right]$$

यदि एक वर्ग 10-20 है तो इसका मध्य बिन्दु 15 होगा, अर्थात्

$$\text{मध्य-बिन्दु} = \frac{20+10}{2} \quad 15$$

(4) वर्ग आवत्ति (Class Frequency) : किसी वर्ग-विशेष से सम्बन्धित मदों की संख्या को वर्ग-आवत्ति कहते हैं। पीछे दिये गये उदाहरण में 120 से 130 के वर्ग की आवत्ति है। जिसका अर्थ है कि 40 व्यक्तियों का वजन 120 से 130 पौण्ड के बीच में है। यदि एक प्रत्येक वर्ग की आवत्ति का योग करें तो हम कुल आवत्ति का पता लगा सकते हैं। इस प्रकार , उक्त उदाहरण में कुल आवत्ति 136 हैं, जिसका अर्थ है कि कुल 136 व्यक्तियों के वजन का अध्ययन किया गया है। वर्गान्तर द्वारा वर्गीकरण करने की दो रीतियाँ हैं :-

(क) अपवर्जी रीति (Exclusive Method);

(ख) समावेशी रीति (Inclusive Method).

(क) अपवर्जी रीति (Exclusive Method) : जब वर्गों की रचना इस प्रकार हो कि एक वर्ग की उच्च सीमा (Upper Limit) दूसरे वर्ग की निम्न सीमा हो तो इसको अपवर्जी रीति कहते हैं। निम्न समंक इस रीति द्वारा प्रस्तुत किये गये हैं :-

आय (रू. में)	व्यक्तियों की संख्या
700-800	50
800-900	100
900-1000	200
1000-1100	150
1100-1200	40
1200-1300	10
	योग 550

इस सारणी से यह स्पष्ट हो जाता है कि अपवर्जी रीति में एक वर्ग की उच्च सीमा उससे अलग वर्ग की निम्न सीमा बनती है। इस प्रकार उपर्युक्त उदाहरण में 50 व्यक्ति 700 और 800 रुपये के वर्ग में हैं, लेकिन जिस व्यक्ति की आय पूरे 800 रू. है उसे 800-900 के वर्ग में सम्मिलित किया जायेगा। यह रीति व्यवहार में सबसे अधिक प्रचलित है। लेकिन इस रीति के प्रयोग में एक साधारण व्यक्ति, जिसे सांख्यिकी का ज्ञान नहीं है, भ्रम में पड़ सकता है। उदाहरण के लिये, यदि किसी व्यक्ति से पूछा जाये कि वह एक मास में कितनी बार सुपर बाजार गया तथा उससे कहा जाये कि वह 0-5, 5-10, 10-20, 25-20 के सामने निशान लगा दे। यदि वह 10 बार सुपर बाजार गया तो वह भ्रम में पड़ सकता है कि निशान 5 से 10 वाले वर्ग में लगाये अथवा 10 से 15 वाले वर्ग में। किसी स्पष्ट निर्देश के न होने पर कुछ व्यक्ति 5 से 10 वाले वर्ग तथा कुछ 10 से 15 वाले वर्ग में निशान लगायेंगे। अतः जब भी इस रीति का प्रयोग किया जाये प्रश्नावली में स्पष्ट निर्देशों का उल्लेख अवश्य होना चाहिये। जब कभी वर्गान्तर 0-10, 10-20, 20-30 आदि दिये हों पाठक को यह सोच लेना चाहिये कि इस सारणी मे अपवर्जी रीति का प्रयोग किया गया है।

(ख) समावेशी रीति (Inclusive Method) : इस रीति के अनुसार किसी वर्ग की उच्च सीमा उसी वर्ग में सम्मिलित होती है तथा एक वर्ग की ऊपरी सीमा और उससे अगले वर्ग की निचली सीमा में अंतर होता है निम्नलिखित उदाहरण से यह रीति स्पष्ट हो जायेगी :-

आय (रू. में)	व्यक्तों की संख्या
700-799	50
800-899	100
900-999	200
1000-1099	150
1100-1199	40
1200-1299	10
	योग 550

700 से 799 के वर्ग में उन व्यक्तियों को सम्मिलित किया जायेगा जिनकी आय 700 से 799 के बीच में है। यदि किसी व्यक्ति की आय 800 रू. है तो उसे दूसरे वर्ग में सम्मिलित किया जायेगा। उपर्युक्त उदाहरण से यह स्पष्ट है कि अपवर्जी रीति के प्रयोग से उत्पन्न भ्रम समावेशी रीति में नहीं है। यदि हम चाहे तो वर्ग 700-799.5 या 700-799.9 आदि बना सकते हैं। यह स्मरण रखना चाहिये कि समावेशी तथा अपवर्जी रीतियों द्वारा वर्ग आवत्ति एक समान ही होती है यद्यपि वर्गान्तर दोनों रीतियों में भिन्न होते हैं।

वर्गीकरण के सिद्धान्त

(Principles of Classification)

वर्गीकरण करते समय निम्न बातों को ध्यान में रखना आवश्यक है :-

(1) यथासम्भव वर्गों की संख्या कम से कम होनी चाहिये। जहाँ तक हो सके वर्गों की संख्या 5 या 15 के बीच हो। यदि समंक अधिक होने के कारण वर्गों की संख्या 15 से अधिक हो जाए तो कोई बात नहीं 5 से कम नहीं होनी चाहिये, अन्यथा वर्गीकरण समंकों की मुख्य विशेषता को भली-भाँति प्रस्तुत नहीं कर पाएगा। वर्गों की संख्या मूल रूप से निम्नलिखित बातों पर निर्भर करती है :-

(क) वर्गीकरण के लिये समंकों की संख्या (Number of figures to be classified), तथा

(ख) समंकों का विस्तार (Magnitude of Figures):

प्रोफेसर स्टर्जेस (Prof. Sturges) ने वर्गों की संख्या ज्ञात करने के लिये निम्न सूत्र का सुझाव दिया है :-

$n = 1 + 3.22 \log N$

n = वर्गों की संख्या (Number of Classes);

N = मदों की संख्या (Number)।

इस सूत्र के अनुसार अवलोकनों की संख्या बहुत अधिक हो जाने पर भी वर्गों की संख्सा अधिक नहीं होती। शायद ही कभी 20 वर्ग बनाये जाये। निम्न उदाहरण से यह बात स्पष्ट हो जायेगी :-

मदों की संख्या	स्टर्जेस का नियम	वर्गों की संख्या
50	$n = 1 + 3.322 \log (50)$	1+ 5.6 = 6.6 या 71
	$= 1 + 3.322 (1.6990)$	
500	$n = 1 + 3.322 \log 500$	1 + 8.9 = 9.9 या 10
	$= 1 + 3.322 (2.6999)$	
5,000	$n = 1 + 3.322 \log 5{,}000$	1 + 12.2 = 13.2 या 11
	$= 1 + 3.322 (3.6990)$	
50,000	$n = 1 + 3.322 \log 50{,}000$	1 + 15.5 = 16.5 या 17
	$= 1 + 3.322 (4.6992)$	

वर्गों की संख्या पूर्णांक में ही हो सकती है इसलिये 6.6, 9.9 आदि को 7.10 लिखा गया है।

(2) जहाँ तक सम्भव हो विषय संख्या (Odd Numbers) वाले वर्गान्तर नहीं लेने चाहिये जैसे 3, 11, 17 आदि। यथासम्भव 5 से अथवा 5 के गुणांक का वर्गान्तर रखना चाहिये। उदाहरण के लिये 5, 10, 20, 50 और 100 आदि। ऐसा करने से समंकों का विश्लेषण सरल हो जाता है।

(3) प्रथम वर्ग की निम्न सीमा जहाँ तक हो सके शून्य (Zero) अथवा 5 के गुणक

$$i = \frac{L \quad S}{1 \quad 3.22 \log N}$$

500 से आरम्भ करनी चाहिए। उदाहरण के लिये, यदि दिये हुए समंकों में न्यूनतम संख्या 33 है तथा वर्गान्तर 10 है तो पहला वर्ग 60 से 70 का लेना ठीक रहेगा न कि 63 से 73 का। इसी प्रकार यदि सबसे कम संख्या 76 है तो 76-81 वर्ग के स्थान 75-80 का वर्ग अधिक उचित रहेगा।

(4) सततता (Continuity) बनाये रखने तथा ठीक वर्गान्तर ज्ञात करने के लिये अपवर्जी रीति का प्रयोग उचित रहेगा। जहाँ समावेशी रीति का प्रयोग किया गया है वहाँ ठीक वर्गान्तर निकालने के लिये वर्ग सीमाओं में समायोजन (Adjustment) की आवश्यकता होती है। इसके लिये प्रथम वर्ग की उच्च सीमा तथा द्वितीय वर्ग की निम्न सीमा का

अन्तर ज्ञात करके उसे 2 से विभाजित कर देते हैं। प्राप्त भागफल को प्रत्येक वर्ग की निम्न सीमा का अन्तर ज्ञात करके उसे 2 से विभाजित कर देते हैं। प्राप्त भागफल को प्रत्येक वर्ग की निम्न सीमा में से घटाकर उच्च सीमा में जोड़ देते हैं। सूत्र के रूप में :-

इन एण्ड (Correction Factor) = $\frac{\text{प्रथम वर्ग की उच्च सीमा} \quad \text{द्वितीय वर्ग की निम्न सीमा}}{2}$

ऐसा करने से समावेशी वर्गान्तर अपवर्जी बन जाते हैं।

नीचे लिखे उदहारण से यह बात स्पष्ट हो जायेगी :-

प्राप्तांक	**विद्यार्थियों की संख्या**
10-19	5
20-29	10
30-39	15
40-49	8
50-59	2
	योग 40

शोधन खंड = $\frac{20 \quad 19}{2} \quad 0.5$

इस राशि को प्रत्येक वर्ग की निम्न सीमा में से घटायेंगे और उच्च सीमा में जोड़ देंगे। परिवर्तित वर्ग इस प्रकार बनेंगे :-

प्राप्तांक	**विद्यार्थियों की संख्या**
9.5-19.5	5
19.5-29.5	10
29.5-39.5	15
39.5-49.5	5
49.5-59.5	2
	योग 40

स्मरण रखना चाहिये कि समायोजन से पहले वर्गान्तर 9 था तथा समायोजन के पश्चात् 10 है। एक और उदाहरण लीजिये :-

प्राप्तांक	**विद्यार्थियों की संख्या**
5-9.5	8
10-14.5	10
15-19.5	2
	योग 20

शोधन खंड = $\frac{10 \quad 9.5}{2} \quad 0.25$

परिवर्तित वर्ग-सीमायें इस प्रकार होंगी :-

चर	आवत्ति
4.75-9.75	8
9.75-14.75	10
14.75-19.75	2
योग	**20**

अब वर्गान्तर 4.5 के स्थान पर 5 है।

एक और उदाहरण लीजिये।

चर	आवत्ति
5-9.99	8
10-14.99	10
15-19.99	5
योग	**20**

शोधन खण्ड = $\frac{10 \quad 9.99}{2} \quad \frac{0.01}{2} \quad 0.005$

समायोजन के पश्चात् वर्ग-सीमायें इस प्रकार होंगी :-

चर	आवत्ति
4.995-9.995	8
9.995-14.995	10
14.995-19.995	2
योग	**20**

(5) विवर्तमुखी या खुले सिर वाली सारणी (Open-end Table) का प्रयोग, जहाँ तक सम्भव हो, नहीं किया जाना चाहिये। खुले सिरे वाली सारणी वह है जिसमें प्रथम वर्ग की निम्न सीमा तथा अंतिम वर्ग की उच्च सीमा न दी हो, जैसे वेतन 500 रू. से कम 500-600, 600-700, 700-800, 800-900, 900 से अधिक। उच्च सीमा के न देने पर बिना कल्पना किये वर्गान्तर ज्ञात करना कठिन हो जाता है और विश्लेषण में भी परेशानियाँ होती हैं। इसलिये खुले सिरे वाली सारणी का जहाँ तक हो प्रयोग नहीं करना चाहिये।

निम्नलिखित उदाहरणों से उपर्युक्त सिद्धान्तों का प्रयोग स्पष्ट हो जायेगा :-

उदाहरण 3 : एक कक्षा के 40 विद्यार्थियों द्वारा सांख्यिकी की एक परीक्षा देने में प्राप्त अंक नीचे दिये गये हैं। प्रथम वर्ग 65-70 लेकर एक आवत्ति वितरण की रचना कीजिए।

सांख्यिकी में प्राप्त अंक

83, 80, 91, 81, 88, 82, 87, 97, 93, 97, 75, 85, 72, 92

84, 90, 87, 78, 83, 98, 86, 80, 93, 86, 88, 88, 82, 72

89, 82, 85, 94, 80, 89, 84, 92, 76, 81, 65, 94

हल : **आवत्ति वितरण की रचना**

सांख्यिकी में प्राप्त अंक	मिलान रेखायें	आवत्ति
65-70	|	1
70-75	||	2
75-80	|||	3
80-85	卌 卌 ||	12
85-90	卌 卌 |	11
90-95	卌 ||	7
95-1000	||||	4
		योग 40

उदाहरण 4 : निम्न समंकों का उचित वर्गान्तर लेते हुए वर्गीकरण कीजिये :-

35 कर्मचारियों द्वारा प्राप्त मासिक वेतन (रू. में)

1370	1950	1130	1780	1240	1350	1920
1970	1875	1598	1580	1460	1352	1430
1282	1313	1461	1390	1260	1612	1213
1314	1415	1518	908	1104	1982	1876
1392	1451	1260	1111	1226	988	1367

हल : न्यूनतम और अधिकतम वेतन में अन्तर 1970-908 = 1062 यदि हम 200 का वर्गान्तर ले तो 6 वर्ग इस प्रकार बनेंगेः-

वेतन (रू. में)	मिलान रेखायें	कर्मचारियों की संख्या
900-1100	|||	3
1100-1300	卌 |||	8
1300-1500	卌 卌 |||	13
1500-1700	卌	5
1700-1900	|||	3
1900-2100	|||	3
		योग 35

सारणीयन

(Tabulation)

समंकों का वर्गीकरण करने के बाद उन्हें व्यवस्थिति ढंग से सारणियों के रूप में प्रस्तुत करना आवश्यक होता है ताकि आंकड़ों से यथोचित निष्कर्ष निकाले जा सके। क्राक्सटन एवं काउडेन का कहना है कि ''स्वयं अपने प्रयोग हेतु या अन्य लोगों द्वारा प्रयोग किये जाने के उद्देश्य से समंकों को किसी उपयुक्त रूप में प्रस्तुत किया जाना चाहिए। आमतौर पर समंक या तो सारणियों में क्रमबद्ध किये जाते हैं या रेखीय (Graphic) विधियों द्वारा उनका चित्रण किया जाता है।''

जहाँ तक सारणीयन के अर्थ का प्रश्न है सारणीयन, समंकों को प्रस्तुत करने की एक व्यवस्थिति पद्धति है। प्रो. ब्लेयर के शब्दों में, ''विस्तत अर्थ में , सारणीय समंकों की खानों (Columns) और पंक्तियों (Rows) के रूप में एक क्रमबद्ध व्यवस्था है।''

प्रो. कौनर के मतानुसार, ''सारणीय किसी विचाराधीन समस्या को स्पष्ट करने के उद्देश्य से किया जाने वाला सांख्यिकीय तथ्यों का क्रमबद्ध एवं सुव्यवस्थित प्रस्तुतीकरण है।''

उद्देश्य (Objects) — सारणीय के तीन उद्देश्य हैं — (*i*) समंकों को व्यवस्थित रूप से प्रस्तुत करना, (*ii*) आंकड़ों को संक्षिप्त ढंग से प्रकट करना, तथा (*iii*) समस्या को अधिक सरल व स्पष्ट बनाना।

सारणीय का महत्व या लाभ
(Importance of Tabulation)

(1) **सरलता :** सारणीयन से समंकों की जटिलता समाप्त हो जाती है और फलस्वरूप आवश्यक सूचनायें जल्दी तथा आसानी से समझ में आ जाती है।

(2) **स्थान व समय की बचत :** सारणीयन द्वारा विशाल तथ्यों को थोड़े व संक्षिप्त रूप में व्यक्त किया जाता है, जिससे समय तथा स्थान की बचत होती है।

(3) **तुलनात्क अध्ययन की सुविधा :** सारणीयन-क्रिया तुलनात्मक अध्ययन को सम्भव बनाती है क्योंकि इसमें समान व तुलना-योग्य समंकों को परस्पर निकटवर्ती खानों में रखा जाता है।

(4) **आकर्षण प्रदर्शन :** सारणीयन के फलस्वरूप नीरस आंकड़े भी आकर्षक लगने लगते हैं।

(5) **सांख्यिकी विवेचना में सहायक :** सारणीबद्ध समंकों का सांख्यिकीय विवेचन जैसे माध्य, अपकिरण, विषमता, सह-सम्बन्ध आदि आसानी से ज्ञान किया जा सकता है।

सारणीयन व वर्गीकरण में अन्तर
(Distinction Between Classification And Tabulation)

वर्गीकरण तथा सारणीयन में काफी अंतर है।

(1) सारणीयन समंकों के वर्गीकरण के बाद की एक स्थिति है। सर्वप्रथम आंकड़ों को वर्गीकृत किया जाता है, तत्पश्चात उन्हें सारणियों में प्रस्तुत किया जाता है। इस प्रकार वर्गीकरण, सारणीयन का आधार है।

(2) वर्गीकरण में संकलित समंकों को उनके समान व असमान गुणों के आधार पर विभिन्न वर्गों (Classes) या श्रेणियों (Series) में बांटा जाता है जबकि सारणीयन में उन्हीं वगीकृत तथ्यों को खानों और पंक्तियों में प्रस्तुत किया जाता है। इस दष्टि से सारणीयन वर्गीकरण का एक यन्त्रात्मक पहलू (Mechanical Aspect) है।

(3) वर्गीकरण सांख्यिकीय विश्लेषण की एक विधि है जबकि सारणीयन समंकों के प्रस्तुतीकरण की एक प्रक्रिया है।

(4) वर्गीकरण के अन्तर्गत समंकों को वर्गों व उपवर्गों में बांटा जाता है जबकि सारणीयन में उन्हें शीर्षक व उपशीर्षकों में रक्खा जाता है।

सारणी की रचना के नियम या सारणी के प्रमुख अंग
(Rules or Main Parts of Table)

(1) **शीर्षक (Title) :** प्रत्येक सारणी का एक संक्षिप्त, स्पष्ट एवं पूर्ण शीर्ष होना चाहिये ताकि समंकों की एक दष्टि में जानकारी हो सके।

(2) **खाने की पंक्तियाँ (Columns and Rows) :** खानों व पंक्तियों की संख्या सारणीयन के उद्देश्य एवं प्रस्तुत सामग्री के आकार को ध्यान में रखकर पहले से ही निश्चित कर लेनी चाहिये। ध्यान रहे, खानों की संख्या अधिक होने से समस्या जटिल व अस्पष्ट हो सकती है। प्रत्येक खाने का उपशीर्षक (Caption) देना आवश्यक है। उपशीर्षक यथासम्भव संक्षिप्त होना चाहिए और प्रत्येक खाने पर क्रम खाने पर क्रम संख्या लिख देनी चाहिये।

(3) **खानों की रूलिंग (Ruling of Columns) :** विषय-सामग्री का महत्वपूर्ण भाग मोटी या दोहरी रेखाओं से बनाना चाहिए जिससे कि द्रष्टा का ध्यान तुरन्त आकर्षित हो सके। कम महत्व के खानों को हल्की रेखाओं द्वारा प्रदर्शित किया जाना चाहिए।

(4) तुलनात्मक अध्ययन (Comparative Study) : ध्यान रहे, जिन समंकों की परस्पर तुलना करनी हो वे पास-पास रक्खे जायें।

(5) पदों की व्यवस्था (Arrangement of Items) : एक आदर्श सारणी की दष्टि से यह आवश्यक होगा कि विभिन्न पदों को उनके महत्व के अनुसार सारणी में स्थान दिया जाय अर्थात् महत्वपूर्ण पदों को पहले और कम महत्व वाले पदों को बाद में रखा जाये। अधिकतर उन पदों को पहला स्थान दिया जाता है जिनको कई वर्गों में विभक्त करना होता है।

(6) टिप्पणियाँ (Foot-notes) : यदि समंकों से सम्बन्धित कोई आवश्यक सूचना सारणी में देने से रह गई है, अथवा किसी तथ्य से सम्बन्धित विशेष स्पष्टीकरण की आवश्यकता है तो उसके लिए सारणी के नीचे व्याख्यात्मक टिप्पणी (Explanatory) दे देनी चाहिए।

(7) स्त्रोत (Source) : सारणी को संदेहरहित व प्रभावशाली बनाने के लिये समंकों का स्त्रोत अवश्य स्पष्ट कर देना चाहिये।

(8) योग (Total) : सारणी में प्रयुक्त होने वाले समंकों के योग व अन्तर्योग की व्यवस्था इस प्रकार की जानी चाहिए कि खानों व पंक्तियों के योग की जाँच व स्पष्टीकरण (Verification) स्वतः ही हो सके।

(9) इकाई एवं व्युत्पन्न समंक (Unit and Derivatives) समंकों के माप की इकाई को सम्बन्धित खानों के ऊपर लिख देना चाहिये। इसी प्रकार प्रतिशत, अनुपात, गुणक व माध्य आदि व्युत्पन्न समंकों को मूल्य समंकों के पास वाले खाने में रखना चाहिये।

(10) सामान्य नियम (General Rules) : इसके अलावा सारणी का आकर्षक होना आवश्यक तथा उसका आकार दिये हुए स्थान अर्थात् कागज के अनुकूल हो। सारणी स्वयं में समंकों की एक मुँह बोलती तस्वीर होनी चाहिये। डॉ. बाउले का कहना है कि ''एक सामान्य व्यक्ति का सारणी को समझने के लिये किया गया विशेष प्रयत्न, सही अर्थों में, सारणी की दोषपूर्ण रचना समझी जायेगी''। वास्तव में सारणी समंकों का एक ऐसा दर्पण है जिसमें झांकने पर वस्तु-स्थिति की पूर्ण जानकारी तुरन्त हो जानी चाहिये।

ध्यान रहे, सारणीय एक सरल नहीं वरन् तान्त्रिक कार्य है। श्री हैरी जरोम (Harry Jerome) के मतानुसार, ''एक आदर्श सांख्यिकीय सारणी निपुणता व प्राविधि की कसौटी है और स्पष्ट रूप से प्रस्तुत की गई सूचनाओं तथा स्थान की मितव्ययिता की सर्वोत्कृष्ट कला-कृति है।'' संक्षेप में सारणीयन के कार्य में निपुणता, अनुभव व सामान्य विवेक का होना अत्यन्त आवश्यक है। डॉ. बाउले का कहना है कि 'संकलन तथा सारणीयन में सामान्य विवेक एक प्रमुख आवश्यकता है तो अनुभव प्रमुख शिक्षक।''

सारणी के प्रमुख 'भाग' निम्न प्रारूप में स्पष्ट किये गये हैं :-

तालिका शीर्षक (Title)
(शीर्ष संकेत)

<table>
<tr><th rowspan="3">Sub Heading</th><th colspan="4">Caption</th><th rowspan="3">Total</th></tr>
<tr><th colspan="2">Sub-heads</th><th colspan="2">Sub-Heads</th></tr>
<tr><th>Column Head</th><th>Column Head</th><th>Column Head</th><th>Column Head</th></tr>
<tr><td>Total</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

Foot note

सारणियों के प्रकार
(Types of Tables)

सांख्यिकीय सारणियों का वर्गीकरण मुख्य रूप से (क) उद्देश्य, (ख) मौलिकता एवं (ग) रचना से आधार पर किया जाता है।

(अ) उद्देश्य के आधार पर सारणीयन

उद्देश्य के आधार पर सारणियाँ दो प्रकार की होती हैं — (1) **सामान्य उद्देश्य वाली सारणी** (General Purpose Table) तथा (2) **विशेष उद्देश्य वाली सारणी** (Special Purpose Table)। सामान्य उद्देश्य वाली अथवा सन्दर्भ सारणी का प्राथमिक उद्देश्य समंकों को इस प्रकार प्रस्तुत करना होता है व्यक्तिगत इकाइयों का पाठकों द्वारा तुरन्त पता लगाया जा सके (क्राक्स्टन एवं काउडेन)। यह सारणी अत्यधिक विस्तत होने के कारण अधिक उपयुक्त नहीं समझी जाती है। यही कारण है कि इसका प्रयोग अधिकतर सरकारी रिपोर्टों के सन्दर्भ में किया जाता है। इसके विपरीत विशेष उद्देश्य वाली अथवा सारांश सारणी (Summary Table) किसी उद्देश्य विशेष की पूर्ति के लिए तैयार की जाती है तथा इसका आधार सामान्य सारणी से अपेक्षाकृत छोटा होता है।

(ब) मौलिकता के आधार पर सारणीयन

मौलिकता के आधार पर सारणियाँ दो प्रकार की होती हैं — (1) **मौलिक या प्राथमिक सारणी** (Original or Primary Table) तथा (2) **व्युत्पन्न सारणी** (Derivative Tabel)। मौलिक सारणी में समंकों को मौलिक रूप में रखा जाता है जबकि व्युत्पन्न सारणी में समंकों के योग, प्रतिशत, अनुपात, गुणांक व माध्य आदि मूल्यों को प्रस्तुत किया जाता है।

(स) रचना के आधार पर सारणीयन

रचना अथवा बनावट के आधार पर सारणियाँ दो प्रकार की होती है :-

(1) **सरल सारणी (Simple Table) :** जब समंकों को केवल एक ही गुण अथवा विशेषता के आधार पर प्रस्तुत किया जाता है, तो उसे सरल सारणी करते हैं, जैसे जनसंख्या का आयु (Age) अथवा लिंग (Sex) अथवा राज्यों के अनुसार वितरण। सरल अर्थात् एक-गुण सारणी का उदाहरण नीचे दिया गया है :-

Illustration 1.

Distribution of Population by Age-groups.

Age-grops (Years)	Number of Persons (Millions)
0-20	— —
20-5	— —
over 50	— —
Total	**— —**

(2) **जटिल सारणी (Complex Table) :** जब समंकों को एक से अधिक विशेषताओं के आधार पर प्रस्तुत किया जाता है तो उसे जटिल सारणी कहते हैं। जटिल सारणी पुनः तीन रूपों में विभाजित की जा सकती है :-

(*i*) **द्विगुण सारणी (Double or Two-way Table) :** द्विगुण सारणी उसे कहते हैं जिसमें समंकों की केवल दो विशेषताओं को समावेश किया जाता है जैसे जनसंख्या का आयु तथा लिंग (Sex) के आधार पर वितरण। इस दष्टि से द्विगुण सारणी का उदाहरण इस प्रकार होगा : -

Illustration 2.

Distribution of Population by Age and Sex

Age-groups	Number of Persons (Millions)		
(Years)	Male	Female	Total
0-20	—	—	—
20-50	—	—	—
over 50	—	—	—
Total	**—**	**—**	**—**

(*ii*) **त्रिगुण सारणी (Treble or Three-way Table) :** त्रिगुण सारणी में एक तथा तीन गुणों के आधार पर समंकों को प्रदर्शित किया जाता है जैसे जनसंख्या का आयु, लिंग तथा साक्षरता (Age, Sex and Literacy) के अनुसार वितरण। उदाहरण इस प्रकार है :-

Illustration 3 : Draw up a blank table showing the distribution of population in a district A, according to five age-groups from 0 to 100 years, sex and literacy.

Distribution of Population by Age, Sex and Literacy

Age-Groups (Years)	**Number of Persons (Millions)**								
	Male			**Female**			**Total**		
	Literate	**Illiterate**	**Total**	**Literate**	**Illiterate**	**Total**	**Literate**	**Illiterate**	**Total**
0-20									
20-40									
40-60									
60-80									
80-100									
Total									

(*iii*) **बहुगुण सारणी (Manifold or Higher Order Table) :** जब समंकों को तीन से अधिक विशेषताओं के आधार पर प्रस्तुत किया जाता है जो उसे बहुगुण सारणी करते हैं। उदाहरण के लिए जनसंख्या को आयु, लिंग, साक्षरता तथा राज्यों में वितरण के आधार पर इस प्रकार दिखाया जायेगा :-

Illustration 4 :

जनसंख्या का राज्य, आयु, लिंग व साक्षरता के अनुसार वितरण

(Distribution of Population by States, Age, Sex and Literacy)

राज्य State	आयु-वर्ग (Age-Groups) (years)	व्यक्तियों की संख्या (मिलियन)								
		पुरुष (Male)			स्त्री (Female)			योग (Total)		
		L	IL	T	L	IL	T	साक्षर	निरक्षर	योग
1. Assam असम	0-20									
	20-50									
	Over 50									
	Total									
2. Bihar बिहार	0-20									
	20-50									
	Over 50									
	Total									

नोट — इस सारणी को इसी प्रकार अन्य प्रान्तों के लिए बढ़ाया जा सकता है।

Illustration 5 : एक रिक्त सारणी बनाइये जिसमें जनशक्ति समंकों का वितरण आयु (Age), लिंग (Sex) तथा ग्रामीण-शहरी निवास (Rural-Urban Character of Residence) के आधार पर दिखाया जा सके।

Solution : प्रस्तुत प्रश्न के लिये त्रिगुण सारणी तैयार की जायेगी। आयु वर्ग चार माने गये हैं।

जनशक्ति का आयु, लिंग तथा ग्रामीण-शहरी निवास के अनुसार वितरण

(Distribution of Manpower by Age, Sex and Character of Residence)

Age-Groups आयु-वर्ग	**Urban (शहरी)**			**Rural (ग्रामीण)**			**Total (योग)**		
	Male	**Female**	**Total**	**Male**	**Female**	**Total**	**Male**	**Female**	**Total**
0-20									
20-40									
40-60									
Over 60									
Total									

Illustration 6 : एक समाचार-पत्र में एक ही परिवार में रहने वाले क्षय-ग्रस्त व्यक्तियों (Tubercular Persons) में फ्लू (Influenza) के प्रभाव से सम्बन्धित निम्न विवरण प्रकाशित हुआ :-

'एक लाख निवासियों में से ठीक पांचवें भाग के बराबर व्यक्तियों में क्षयरोग (T.B.) के लक्षण प्रकट हुए और उनमें से 5000 व्यक्ति फ्लु से पीड़ित हुए परन्तु उनमें से केवल 1000 व्यक्ति ही अदूषित (Uninfected) घरों में रहते थे। इसके विपरीत जिन्हें फ्लू नहीं हुआ, ऐसे क्षय-ग्रस्त व्यक्तियों के 15 वें भाग पर अभी भी स्पर्श अर्थात् संक्रमण (Infection) का प्रभाव था। कुल मिलाकर 21 हजार व्यक्ति फ्लु से पीड़ित हुए और 41 हजार संक्रमण से प्रभावित थे परन्तु ऐसे व्यक्तियों की संख्या केवल 2000 थी जो फ्लु के शिकार हुए और क्षय रोग से प्रभावित नहीं हुए तथा जो ऐसे घरों में रहते थे जहाँ फ्लू का कोई और रोगी नहीं था ? उपर्युक्त सूचना को संक्षिप्त सारणी में पुनर्व्यवस्थित कीजिए।

Solution :

Table showing the incidence of influenza among tubercular persons living in same family

	फ्लू से पीड़ित Having Influenza			फ्लू से पीड़ित नहीं Not Having Incluenza			Grand Total महायोग
	दूषित घर Infected Houses	अदूषित Uninfected Houses	योग Total	दूषित घर Infected Houses	अदूषित Uninfected Houses	योग Total	
क्षयग्रस्त Having T.B.	1,000	4,000*	5,000	1,000	14,000¶	15,000	20,000
क्षयग्रस्त नहीं Not Having T.B.	14,000	2,000	16,000§	25,000**	39,000	64,000	80,000
योग (Total)	15,000	6,000	21,000	26,000	53,000	79,000	1,00,000

* Persons having influenza and living in uninfected houses (out of tubercular persons) are 5000 – 1000 = 4000

¶ Persons not having influenza and living in uninfected houses (out of tubercular persons) are 15000 – 1000 = 14000 ($\because$ 20,000 – 5,000 = 15,000)

§ No. of persons not having T.B. but having influenza are 21,000 – 5,000 = 16,000.

** Total no. of persons living in infected houses are 41,000. Figures known are 15000 + 1000. So persons living in such houses with T.B. effect and no influenza effect are 41,000 – (15,000 + 1,000) = 25,000.

द्विचर आवत्ति सारणी

(Bivariate or Two-directional Frequency Table)

कभी-कभी दो आवत्ति वितरण , जिनमें परस्पर कुछ सम्बन्ध होता है, एक-साथ द्विचर आवत्ति सारणी के रूप में रखे जाते हैं। यह वितरण, जिसे आड़ी वर्गीकृत-सारणी (Cross-classified Table) भी कहते हैं, दो चरों के मध्य बहत सम्बन्ध (Borad Relationship) का विश्लेषण करने में सहायक सिद्ध होता है। उदाहरणार्थ, सेवाकाल और मजदूरी के आधार पर किया गया श्रम-शक्ति का वर्गीकरण, यह स्पष्ट करता है कि क्या मजदूरी की दरें सेवा की अवधि से सम्बन्धित हैं अर्थात् उससे प्रभावित होती हैं। इसी प्रकार व्यक्तियों की लम्बाई तथा भार के आधार पर तैयार की गई द्विमुखी सारणी दोनों चरों के मध्य कार्यात्मक-सम्बन्ध (Functional Relationship) को स्पष्ट करती है। ध्यान रहे, ऐसी सारणी सम्बन्ध-स्थापन का कोई सांख्यिकीय माप प्रस्तुत नहीं करती बल्कि एक सामान्य चाक्षुष प्रभाव (Visual Impact) रखती है।

रचना विधि : इसकी रचना विधि नीचे एक उदाहरण द्वारा स्पष्ट की गई है।

Illustration 7. नीचे 20 पतियों और पत्नियों की आयु के अंक दिये गये हैं। इन समंकों को एक द्विचर वर्गीकृत वितरण सारणी के रूप में प्रस्तुत कीजिए, जिनमें वर्ग-विस्तार 5 का अर्थात् वर्गान्तर 20-25, 25-30 आदि हों।

क्रम संख्या	:	1	2	3	4	5	6	7	8	9	10
पतियों की आयु	:	28	37	42	25	29	49	37	35	23	41
पत्नियों की आयु	:	23	30	40	26	25	41	35	25	21	38
क्रम संख्या	:	11	12	13	14	15	16	17	18	19	20
पतियों की आयु	:	27	39	23	33	36	32	22	29	38	48
पत्नियों की आयु	:	24	34	20	31	29	35	23	27	34	47

Solution :

(1) यहाँ प्रथम चर, पतियों की आयु है और दूसरा चर पत्नियों की आयु। दोनें चरों में न्यूनतम आयु 20 वर्ष तथा अधिकतम आयु 49 वर्ष है। अतः दोनों के लिये 20-25, 25-30, 30-35, 35-40, 40-45, 45-50 वर्गान्तर पर्याप्त रहेंगे।

नोट : ध्यान रहे, दोनों चरों के लिये समान वर्ग-विस्तार या एक-समान वर्गान्तर होना आवश्यक नहीं है। क्योंकि परिस्थिति अनुसार यह तो चर मूल्यों पर निर्भर करेगा।

(2) पतियों की आयु (प्रथम चर) के वर्गान्तर पहले खाने या कॉलम (Column) में रखे जायेंगे और पत्नियों की आयु (दूसरे चर) के वर्ग ऊपर की तरफ छः खाने बनाकर उनमें रख दियें जायेंगे।

नोट : खाने (Column) कितने बनाये जाएं, यह दूसरे चर के वर्गान्तरों की संख्या पर निर्भर करेगा। इसी प्रकार पंक्तियों (Rows) कितनी हों, यह प्रथम चर के वर्गान्तरों की संख्या द्वारा तय होगा।

(3) अंतिम पंक्ति और अंतिम खाना कुल आवत्तियों के लिये होते हैं।

(4) अब प्रश्न पर आइए। क्रम संख्या 1 पद दोनों चरों का संयोग 28, 23 है। यदि 28 वर्ष पतियों (प्रथम चर) के आयु वर्ग 25-30 में आता है और 23 वर्ष पत्नियों के आयु-वर्ग 20-25 से सम्बद्ध है। अतः बायें से 25-30 वर्ग के सामने, और ऊपर से 20-25 वर्ग के नीचे उभयनिष्ठ कोष्ठक (Common Cell) में एक मिलान-रेखा (Tally-bar) खींच दी जायेगी। फिर, दूसरे नम्बर पर 37, 30 का संयोग है। इसे बायें से 35-40 के आयु-वर्ग में और ऊपर से 30-35 आयु वर्ग के नीचे के कोष्ठक में रखा जायेगा। इसी प्रकार शेष संयोगों के लिये उपयुक्त कोष्ठकों में मिलान रेखाएं बना दी जायेंगी। (देखिये सारणी)।

(5) फिर, मिलान रेखाओं की गिनती करके उनकी संख्या लिख दी जायेगी। यह कोष्ठक आवत्तियाँ (Call Frequencies) कहलाती है। (देखिये सारणी)।

(6) अन्त में कॉलम-वार और पंक्ति-वार कोष्ठकों के योग किये जाते हैं और उन्हें अंतिम कॉलम और अंतिम पंक्ति में रखा जाता है। जैसे कॉलम-वार योग 5, 5, 4, 3, 2, 1 है और पंक्ति-वार योग 3, 5, 2, 6, 2, 2 आया है।

(7) फिर इनका महायोग (Grand Total) किया जाता है। इसे कुल आवत्ति (S*f*) कहते हैं स्मरण रहे, दोनों चरों की आवत्तियों का योग महायोग के बराबर होना चाहिए। जैसे इस उदाहरण में दोनों तरफ से योग 20 आया है। यदि यह महायोग दोनों तरफ से समान नहीं है तो समझ लीजिए कि कोई त्रुटि हो चुकी है।

द्विचर आवत्ति वितरण सारणी का अंतिम प्रारूप निम्न है : -

मिलान-रेखा तालिका (Tally Sheet)

पतियों की आयु	**पत्नियों की आयु**						**योग**
20-25	25-30	30-35	35-40	40-45	45-50		
20-25	||| 3						3
25-30	|| 2	||| 3					5
30-35				1		1			2
35-40		|| 2	||| 3		1			6
40-45					1		1		2
45-50						1		1	2
योग	5	5	4	3	2	1	20

❑❑

अध्याय - 4

समंकों का चित्रमय प्रदर्शन
(Diagrammatic Presentation of Data)

सांख्यिकी का एक महत्वपूर्ण कार्य जटिल समंकों को इस प्रकार प्रस्तुत करना है कि उनको सरलता से समझा जा सके और उन्हें देखकर तुरन्त कोई निर्णय लिया जा सके। यद्यपि वर्गीकरण, सारणीयन, माध्य इत्यादि रीतियों द्वारा समंकों को सरल और बोधगम्य बनाने का प्रयास किया जाता है, परन्तु सामान्य व्यक्ति के लिए यह प्रस्तुतीकरण अधिक उपयोगी सिद्ध नहीं होता क्योंकि ये सभी रीतियां तथ्यों को समंकों में ही व्यक्त करती हैं। फिर, एक सामान्य व्यक्ति नीरस अंकों में न तो कोई रुचि रखता है और न ही उसमें समंकों के जमघट को समझने की क्षमता होती है वास्तव में, उसे तो विषय-सामग्री का एक ऐसा चित्र चाहिए जिसे वह मस्तिष्क पर बिना जोर डोल समझ सके। अतः इस दष्टि से आवश्यक है कि सांख्यिकीय तथ्यों को दष्टिगत विधियों (Visual Method) द्वारा इस ढंग से प्रदर्शित किया जाये किस साधारण-से-साधारण बुद्धि का व्यक्ति भी उसे एक-ही नजर में समझ सके। सांख्यिकी में समंकों के आकर्षक प्रदर्शन को ऐसी दो विधियां हैं — (*i*) समंकों का चित्रमय प्रदर्शन तथा (*ii*) बिन्दू-रेखीय प्रदर्शन। इस अध्याय में हम चित्रमय प्रदर्शन का और अगले अध्याय में बिन्दु-रेखीय प्रदर्शन का विवेचन करेंगे।

चित्रमय प्रदर्शन का अर्थ (Meaning) : चित्रमय प्रदर्शन से तात्पर्य सांख्यिकीय समंकों को सरल, रोचक व आकर्षक ज्यामितीय आकृतियों (Geometrical Figures), जैसे — दण्ड चित्र, आयत, वत्त अथवा मानचित्रों के रूप में प्रस्तुत करना है। इस प्रकार चित्रमय प्रदर्शन अन्य सांख्यिकीय रीतियों के विपरीत, तथ्यों को अधिक रोचक एवं सरल बनाने की एक प्रक्रिया है।

चित्रमय प्रदर्शन की उपयोगिता एवं लाभ
(Utility and Advantages of Diagrammatic Presentation)

मानव मस्तिष्क अत्यधिक सूक्ष्म तथा उपजाऊ होने के बावजूद समंकों को भली-भांति समझने में असमर्थ है। वैसे भी मनुष्य कोई कम्प्यूटर नहीं और न ही उससे यह आशा की जा सकती है कि वह विशाल तथ्यों को मानस-पटल पर अंकित कर सके। फिर, डॉ. बाऊले के अनुसार ''संख्याओं की कोई भी सूची, जैसे-जैसे उसकी लम्बाई बढ़ती जाती है, कम ग्राह्य होती है।'' वास्तव में, चित्रमय प्रदर्शन समस्या का एक सरल समाधान है। चित्र, समंकों के आकर्षक प्रदर्शन में अत्यधिक उपयोगी होते हैं। सांख्यिकीय चित्रों के मुख्य लाभ इस प्रकार हैं :-

(1) आकर्षक और प्रभावी साधन (Attractive and Effective Media of Presentation) : चित्र रोचक तथा आकर्षक होते हैं और ये उन लोगों का ध्यान भी अपनी ओर बरबस खींच लेते हैं, जिन्हें अंकों में कोई रुचि नहीं होती। प्रायः कहा जाता है ''एक चित्र हजार शब्दों के बराबर होता है।'' (A Picture is Worth in Thousand Words)। चित्रों का अंकों की तुलना में मस्तिष्क पर अधिक स्थायी प्रभाव पड़ता है अर्थात् ये अधिक समय तक याद रहते हैं।

(2) सरल एवं सुबोध प्रस्तुतीकरण (Simple and Intelligible Presentation) : चित्र नीरस, जटिल तथा अव्यवस्थित समंकों को सरल और बोधगम्य बना देते हैं। फिर, चित्रों को समझने के लिये दिमाग पर अधिक जोर नहीं डालना पड़ता क्योंकि ये विषय की तत्काल जानकारी देते हैं। ये विशालता को संक्षिप्तता और जटिलता को सरलता में बदलते हैं। प्रो. मोरोने के अनुसार ''अधिकांश व्यक्तियों के लिये निरे समंक नीरस होते हैं। चित्र, किसी जटिल स्थिति के स्वरूप को दिखाने में हमारी सहायता करते हैं। जिस प्रकार एक मानचित्र हमें एक विशाल देश का विहंगम दश्य प्रदान करता है ठीक उसी प्रकार चित्र एक ही दष्टि में समंकों से सम्बन्धित जटिल तथ्यों का सम्पूर्ण सार समझने में हमारे लिये सहायक सिद्ध होते हैं।''

(3) समय तथा श्रम की बचत (Saving of Time and Labour) : चित्रों से समय तथा श्रम की बचत होती है जहां समंकों को समझने और उनके निष्कर्ष निकालने में अधिक समय व परिश्रम लगाना पड़ता है वहां चित्र सहज ही तथ्यों की सारी विशेषतायें हमारे सामने खोलकर रख देते हैं। फिर, चित्रों को समझने के लिये किसी विशेष या तकनीकी ज्ञान की आवश्यकता नहीं होती।

(4) **तुलना में सहायक (Helpful in Comparison) :** चित्र विभिन्न तथ्यों के बीच तुलना करने में सहायक होते हैं। संख्यात्मक तुलना उतनी प्रभावी नहीं होती जितनी की चित्रात्मक रूप में। उदाहरणार्थ, दस वर्षों के उत्पादन के आंकड़े और इन आंकड़ों का चित्र, अगर साथ-साथ दिये हुए हैं तो देखने वाले का ध्यान चित्र की ओर अधिक जाता है और यह तुरन्त बताया जा सकता है कि किस वर्ष का उत्पादन सबसे अधिक या सबसे कम है।

(5) **व्यापक उपयोग (Universal Use) :** चित्रों के उपयोग का क्षेत्र अति व्यापक है। विभिन्न लाभों के कारण सांख्यिकीय चित्रों का सभी शास्त्रों में प्रयोग किया जाता है। विशेष रूप से व्यापार, वाणिज्य तथा विज्ञापन के क्षेत्र में तो चित्रों का और भी अधिक महत्व है। सच तो यह है कि चित्र जन-जीवन की सार्वभौमिक शैली का रूप ले चुके हैं।

(6) **मनोरंजनक एवं जानकारी के साधन (Source of Entertainment and Knowledge) :** चित्र हमें तथ्यों की जानकारी देने के साथ-साथ हमारा मनोरंजन भी करते हैं। चित्रों में जहां गहनता होती है वहां उनमें हास्य-विनाद भी छिपा होता है। जिसके कारण हम चित्रों के प्रति अधिक आकर्षित होते हैं। फिर, चित्र उत्सुकता को भी बढ़ाते हैं और उत्सुकता ज्ञान की पहली सीढ़ी है।

चित्रमय प्रदर्शन की परिसीमाएं
(Limitations)

चित्र तथ्यों का एक प्रभावी प्रदर्शन अवश्य है परन्तु वास्तविक नहीं। चित्रों की सहायता से विषय-सामग्री का केवल एक मोटा-सा अंदाजा (Rough Idea) ही लगाया जा सकता है। फिर, यह एक ऐसा कोमल यंत्र है जिसके त्रुटिपूर्ण निर्वचन (Wrong Interpretation) की सम्भावना हमेशा बनी रहती है।

(1) **सीमित उपयोग :** चित्रमय प्रदर्शन की उपयोगिता केवल एक सामान्य व्यक्ति के लिए है, एक विशेषज्ञ के लिये है, एक विशेषज्ञ के लिये नहीं।

(2) **सीमित शुद्धता :** चित्र सन्निकट मूल्यों (Approximate Values) पर आधारित होने के कारण तथ्यों का यथार्थ विवेचन नहीं कर पाते।

(3) **अनुपयुक्ता :** चित्रों की सहायता से विभिन्न मूल्यों का सूक्ष्म अन्तर अथवा बहत् अन्तर प्रदर्शित करना सम्भव नहीं है। उदाहरण के लिये दो मूल्यों 6,000 तथा 6,015 अथवा 50 तथा 5,4000 का अन्तर चित्रों द्वारा स्पष्ट नहीं किया जा सकता, भले ही माप-दण्ड कुछ भी हो।

(4) **दुरुपयोग :** चित्रों की सरलता से दुरुपयोग को सकता है क्योंकि गलत माप-दण्ड पर बने चित्र भ्रमात्क होते हैं।

(5) **सीमित तुलनीयता :** चित्रों द्वारा तुलनात्मक अध्ययन तभी सम्भव है जब समंक दो या दो से अधिक हों और वे सजातीय हों अन्यथा निष्कर्ष भ्रमात्मक हो सकते हैं।

(6) **बहुगुण-प्रदर्शन की कठिनाई :** चित्रों की सबसे बड़ी सीमा यह है कि इनके द्वारा बहु-गुणी सूचनाएं प्रदर्शित नहीं की जा सकती। ऐसी सूचनाओं के चित्रों के स्थान पर सारणीयन एक अधिक प्रभावशाली रीति है। ध्यान रहे, चित्र सारणियों के अनपूरक तो हैं स्थापनन्न नहीं।

चित्र-रचना के सामान्य नियम
(General Rules for Constructing Diagrams)

चित्रों की रचना अत्यन्त सावधानी और कुशलता के साथ की जानी चाहिये ताकि व समंकों के यथार्थ-भाव को सही रूप में प्रकट कर सके। चित्र-रचना का कार्य, कोई साधारण-कार्य न होकर विशिष्ट-योग्यता का कार्य हैं अतः इस दष्टि से चित्रों को आकर्षक तथा प्रभावशाली बनाने के लिये निम्न नियमों का पालन करना आवश्यक है :-

(1) **आकर्षण एवं स्वच्छता (Attractive and Neatness) :** जैसाकि हम जानते हैं चित्र सांख्यिकीय समंकों के प्रस्तुतीकरण के चक्षुष-उपकरण (Visual Aids) हैं और ये आंखों को लुभाने के साथ-साथ मस्तिष्क पर स्थायी प्रभाव छोड़ते हैं। इसलिये यह आवश्यक है कि चित्र इतने स्वच्छ, आकर्षक तथा रोचक बनाए जाएं ताकि वे बरबस ही लोगों का ध्यान अपनी ओर आकर्षित कर लें। चित्रों को आकर्षक बनाने के लिये विभिन्न प्रकार के रंगों, शेड्स, बिन्दुओं (डाट्स), रेखाओं, चारखानों आदि उपायों का प्रयोग किया जा सकता है।

(2) **शुद्धता (Accuracy) :** चित्र-रचना में शुद्धता पर ध्यान देना आवश्यक है। विशेष रूप से चित्रों में आकर्षण पैदा करते समय, उनकी शुद्धता का परित्याग नहीं करना चाहिए क्योंकि अशुद्ध चित्रों से भ्रमात्मक निष्कर्ष निकलते हैं।

(3) **उपयुक्त आकार (Suitable Size) :** चित्र के आकार के बारे में कोई दढ़ व निश्चित नियम नहीं है। हाँ ! चित्र का आकार कागज के आकार के अनुरूप होना चाहिए। चित्र न तो बड़ा हो औन न ही बहुत छोटा और यह कागज के मध्य में बनाया जाना चाहिये। वैसे प्रो. लूट्स ने चित्र के आकार के सम्बन्ध में 'Root Two' नामक नियम तैयार किया है जिसके अनुसार चित्र की छोटी तथा बड़ी भुजा के बीच $1 : \sqrt{2}$ या $1 : 1.414$ का अनुपात होना चाहिये।

(4) **शीर्षक तथा फुटनोट (Title and Footnote) :** चित्र के ऊपर एक उपयुक्त, स्पष्ट तथा संक्षिप्त शीर्षक होना चाहिये जिससे कि उसे देखते ही चित्र की विषय-सामग्री और उसमें दिखाये गये तथ्यों की तुरन्त जानकारी हो जाये। यदि शीर्षक से कोई तथ्य स्पष्ट नहीं होता तो उसकी व्याख्यात्मक टिप्पणी चित्र के नीची बायीं तरफ दे देनी चाहिये।

(5) **मापदण्ड का चयन (Selection of Scale) :** चित्र-रचना का सबसे महत्वपूर्ण पहलू उचित मानदण्ड या पैमाने का चयन करना हैं। मापदंड के चयन के बारे में कोई निश्चित नियम नहीं है, तथापि इस सम्बन्ध में कुछ विचारणीय बातें इस प्रकार हैं :-

(*i*) कागज का आकार और समंकों की प्रकृति को देखते हुए पैमाना इस प्रकार निर्धारित किया जाना चाहिये कि चित्र न तो बहुत बड़ा और न ही बहुत छोटा हो।

(*ii*) उसमें समंकों की सभी महत्वपूर्ण विशेषताएं स्पष्ट हो जाएं।

(*iii*) पैमान समसंख्या (Even Numbers) अथवा 5 या 10 के गुणितों (Multiples) में होना चाहिए।

(*iv*) उदग्र (Vertical) तथा क्षैतिज (Horizontal) दोनों तरफ के पैमानों का स्पष्ट संकेत कर देना चाहिए।

(*v*) यदि दो या दो से अधिक चित्रों की तुलना करनी है तो सही निष्कर्ष निकालने के लिये उनका मापदंड एक-समान रखना चाहिये।

(6) **चित्र खींचना (Sketching of Diagram) :** चित्रों की रचना सदैव पैमाने या ज्यामितीय उपकरणों की सहायता से करनी चाहिये। मुक्त-हस्त (Free Hand) चित्र देखने में भद्दे लगते हैं इसलिये इस आदत से बचना चाहिए। इसके अलावा चित्र के विभागों या उप-विभागों के अंतर को बिन्दुओं, खड़ी या पड़ी रेखाओं, चारखानों आदि संकेतों द्वारा स्पष्ट कर देना चाहिए।

(7) **संकेत (Index) :** चित्र की रचना में प्रयुक्त विभिनन प्रकार के चिह्नों जैसे शेड्स, बिन्दुओं, रेखाओं, चारखानों आदि के अर्थ को स्पष्ट करने के लिये, चित्र के ऊपर दायें कोने में संकेत देने चाहिएं जिससे कि चित्र और उसके विभाग आसानी से समझ में आ सके।

(8) **स्त्रोत-टिप्पणी (Source-note) :** जिस तथ्य को चित्रांकित किया गया है उससे सम्बन्धित आंकड़ें जहां से प्राप्त किये गये है, उसका स्त्रोत चित्र के नीचे दे देना चाहिए। इससे चित्र की विश्वसनीयता बढ़ जाती है।

(9) **सरलता (Simplicity) :** जहां तक सम्भव हो सके, चित्र सरल होने चाहियें ताकि उन्हें साधारण-से-साधारण व्यक्ति भी आसानी से समझ सके। यदि बहुत सारी विशेषताओं को एक ही जटिल चित्र बनाकर दिखाया गया है तो उसे न केवल समझना मुश्किल होगा बल्कि उससे दिमाग में भ्रम पैदा होगा।

(10) **उपयुक्त चित्र का चुनाव (Choice of a Suitable Diagram)** एक उपयुक्त चित्र का चयन करना सबसे कठिन कार्य है क्योंकि समंकों के प्रदर्शन हेतु अनेक प्रकार के चित्रों का प्रयोग किया जाता है। वैसे उचित चित्र का चयन, मुख्य रूप से समंकों की प्रकृति (Nature), समंकों का विस्तार (Magnitude), चित्र का क्षेत्र (अर्थात् जिन लोगों के लिये चित्र बनाया जाना है) आदि बातों पर निर्भर करता है और साथ-साथ ही इसके लिये पर्याप्त निपुणता, अनुभव, ज्ञान तथा विवेक की आवश्यकता होती है।

चित्रों के प्रकार
(Kinds of Diagrams)

सांख्यिकी चित्र मुख्यता पांच प्रकार के होते हैं — (1) एक-बिमा चित्र, (2) द्वि-बिमा चित्र, (3) त्रि-बिमा चित्र, (4) चित्र-लेख, (5) मानचित्र।

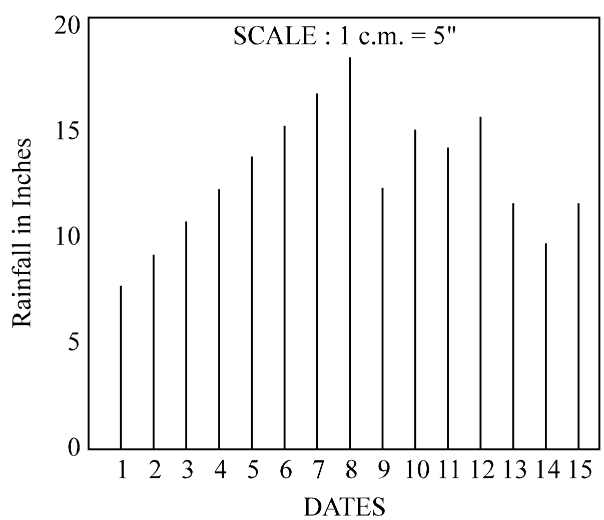
एक विस्तार वाले या एक-विमा चित्र
(One Dimensional Diagrams)

एक-विमा चित्र उन चित्रों को कहते हैं जिनके बनाने में केवल एक ही विस्तार अर्थात् ऊँचाई का प्रयोग किया जाता है और चौड़ाई अथवा मोटाई का नहीं। यह चित्र मुख्य रूप से रेखाओं (Lines) तथा दण्डों (Bars) के रूप में बनाये जाते हैं। ध्यान रहे, दण्ड-चित्रों में दण्ड की चौड़ाई भी रक्खी जाती है परन्तु उसका माप-दण्ड से कोई सम्बन्ध नहीं होता। ये चित्र रचना व समझने की दष्टि से सरल होते हैं परन्तु इनका प्रयोग केवल उन स्थितियों में किया जाता है जहां न्यूनतम व अधिकतम मूल्यों में अधिक अन्तर न हो। नीचे हम एक-विमा चित्र के विभिन्न स्वरूपों का अध्ययन करेंगे :-

(1) रेखा-चित्र (Line Diagram) : रेखा-चित्र का प्रयोग तब किया जाता है, जब पद-मूल्यों की संख्या अधिक हो और श्रेणी के सबसे छोटे और सबसे बड़े पद मूल्य के बीच अन्तर कम हो। चित्र-रचना के लिये प्रत्येक पद-मूल्य के बराबर लम्बाई की खड़ी रेखा खींची जाती है और सभी रेखाओं के बीच समान अन्तर रक्खा जाता है। ध्यान रहे, इन रेखाओं में मोटाई नहीं होती, जिसके कारण यह चित्र प्रायः कम आकर्षक होते हैं (देखिये चित्र-1)।

Illustration 1. चेरापूंजी क्षेत्र में प्रतिदिन वर्षा का 15 दिनों का रिकार्ड नीचे दिया गया है। इन समंकों को एक उपयुक्त चित्र द्वारा प्रस्तुत कीजिये।

Dates	:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ranifall ('"')	:	7	8	10	12	13	14	15	17	11	15	14	16	10	8	10



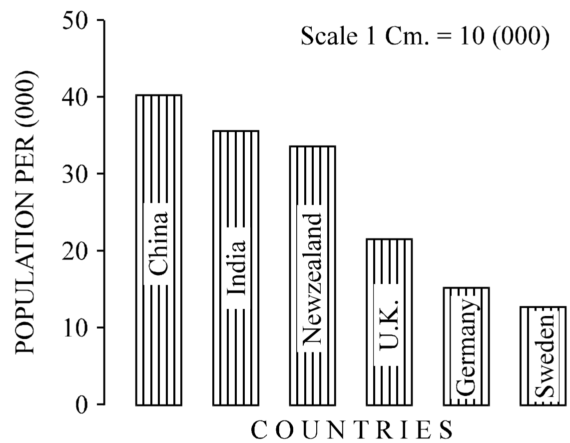
चित्र-1 : रेखाचित्र (Line Diagram)

(1) सरल दण्ड चित्र (Simple Bar Diagram) : पद-मूल्यों की संख्या कम होने की दशा में सरल दण्ड-चित्रों तैयार किये जाते हैं। रेखा-चित्र और दण्ड-चित्र में अन्तर केवल इतना है कि दण्ड-चित्र में रेखाओं की चौड़ाई बना दी जाती है जिससे चित्र का आकर्षण कई गुणा बढ़ जाता है, जबकि शेष सारी बातें रेखा-चित्र के समान ही होती हैं। दण्ड-चित्र की रचना के लिये सबसे बड़े मूल्य के आधार पर पहले एक उपयुक्त पैमाना निश्चित कर लिया जाता है। फिर, सभी दण्ड इस पैमाने के अनुसार बनाये जाते हैं। हाँ ! दण्डों की चौड़ाई व उनके बीच का रिक्त स्थान बराबर रखा जाता है। दण्डों को अधिक आकर्षक बनाने के लिये उनमें रंग व डिजाइनों का प्रयोग किया जा सकता है। दण्ड-चित्र उदग्र अर्थात् खड़े (Vertical) तथा क्षैतिक अर्थात् पड़े (Horizontal) दोनों प्रकार के हो सकते हैं। ये चित्र व्यक्तिगत समंकों, काल श्रेणी तथा स्थानानुसार समंक श्रेणियों के लिये अधिक उपयुक्त होते हैं। (देखिये चित्र 2)।

Illustration 2. The following table gives the birth rate per thousand of different countries over a certain period. Construct a suitable diagram.

Country	**:**	China	India	New Zealand	U.K.	Germany	Sweden
Birth Rate	**:**	40	33	30	20	16	15

विभिन्न देशों में जन्म दर प्रति हजार



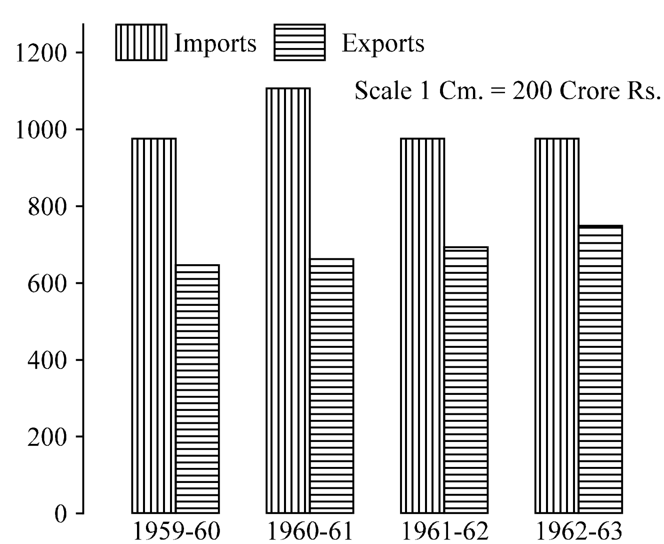
चित्र-2 : सरल दण्ड-चित्र (Simple Bar Diagram)

(3) बहुगुणी दण्ड-चित्र (Multiple Bar Diagrams) : जब एक ही साथ अनेक सूचनाएं जोकि एक ही तथ्य-समूह से सम्बन्धित हों, को प्रदर्शित करना हो तब सरल दण्ड-चित्र के बजाय बहुगुणी दण्ड-चित्र का प्रयोग किया जाता है। एक समूह से सम्बन्धित सभी दण्ड आपस में सटे हुए बनाये जाते हैं। फिर, थोड़ा रिक्त स्थान छोड़कर दूसरे समूह के दण्ड तैयार किये जाते हैं। हाँ ! अध्ययन की सुविधा के लिये विभिन्न तथ्यों को प्रदर्शित करने वाले दण्डों को अलग-अलग रंगों या डिजाइनों द्वारा दिखाया जाता है। दो समूह वाले दण्ड चित्रों को युगल दण्ड-चित्र (Double Bars) और तीन समूह वालों को त्रिदण्ड-चित्र (Treble Bars) कहते हैं (देखिये चित्र संख्या क्रमशः 3 व 4)।

lllustration 3. भारत के विदेशी व्यापार से सम्बन्धित निम्नांकित आंकड़ों को चित्र द्वारा प्रदर्शित कीजिये :-

Year	Imports	Exports
1959-60	961	640
1960-61	1087	643
1961-62	958	677
1962-63	981	709

भारत के आयात एवं निर्यात



चित्र-3 : युगल दण्ड-चित्र (Double Bar Diagram)

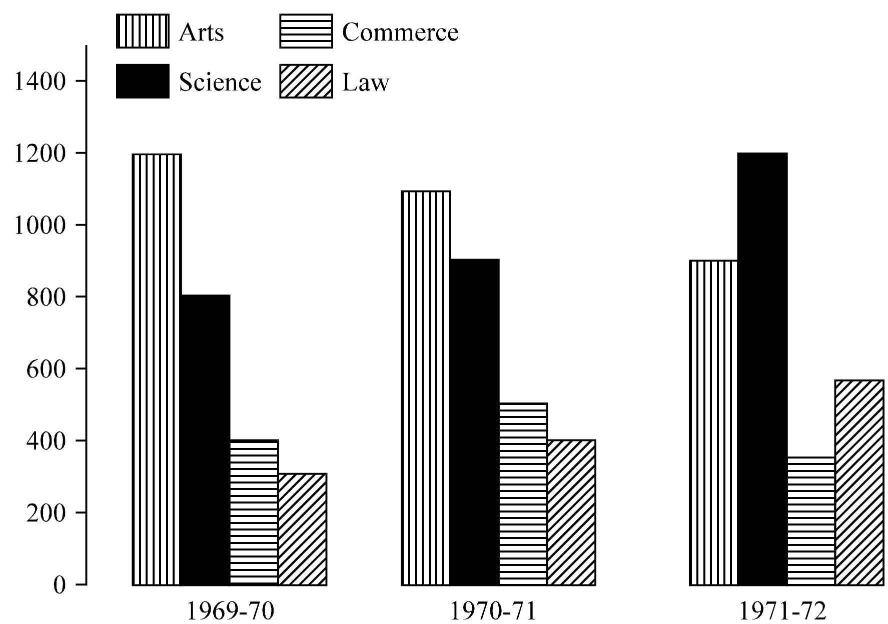
Illustration 4. मेरठ कॉलिज के चार संकायों की तीन वर्षों की छात्र-संख्या को बहुगुणी दण्ड-चित्रों द्वारा प्रस्तुत कीजियेः-

छात्रों की संख्या

सकाय (Faculty)	**1669-70**	**1970-71**	**1971-72**
कला (Arts)	1200	1100	900
विज्ञान (Science)	800	900	1200
वाणिज्य (Commerce)	400	500	350
विधि (Law)	300	400	550

(4) अन्तर्विभक्त दण्ड-चित्र (Sub-Divided Bar Diagram) : जब समंकों के 'योग' (Total) तथा उनके विभिन्न 'विभागों' (Divisions) को एक-साथ दिखाना हो तब अन्तर्विभक्त दण्ड-चित्रों का प्रयोग किया जाता है। इन्हें संघटक दण्ड-चित्र (Component Bar Diagrams) भी कहते हैं। रचना की दष्टि से , सर्वप्रथम दिये हुए पद-मूल्यों के अनुसार विभिन्न दण्ड बनाये जाते हैं और प्रत्येक दण्ड का उसके विभागों के अनुपात में बांट दिया जाता है। और अधिक स्पष्ट करने के

विभिन्न संकायों के छात्र संख्या



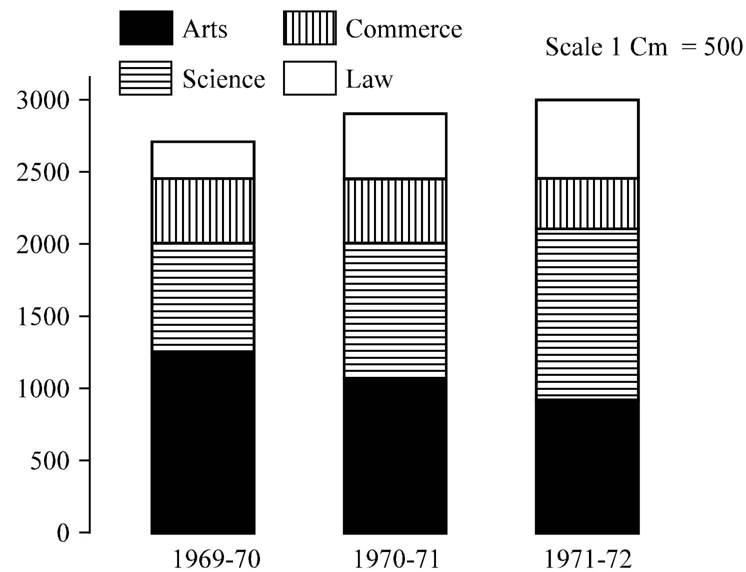
चित्र-4 : बहु-दण्ड चित्र (Multiple Diagram)

लिये पिछले उदाहरण में यदि मेरठ कॉलेज के विभिन्न समंकों में विद्यार्थियों की संख्या के साथ-साथ उनकी कुल संख्या में होने वाले निरपेक्ष परिवर्तनों को भी प्रदर्शित करना हो तो अन्तविभक्त दण्ड-चित्र अधिक उपयुक्त माना जायेगा। चूंकि कॉलेज की तीनों वर्षों में छात्र-संख्या क्रमशः 2700, 2900, 3000 है। अतः 1 cm = 500 का पैमाना अधिक उचित रहेगा (देखिये चित्र-5)।

अन्तर प्रदर्शित करने वाले अन्तविभक्त चित्र (Sub-divided Bars Showing Differences) : इन चित्रों के द्वारा दो प्रकार के समंकों और उनके पारस्परिक अन्तर को भी प्रदर्शित किया जा सकता है, जैसे आयात निर्यात तथा व्यापार-शेष; जीवन-दर, मत्यु -दर एवं अतिजीवन-दर आदि। इन चित्रों की रचना के लिए, सर्वप्रथम दोनों तथ्यों में से बड़े तथ्य (मूल्य) को लेकर सरल दण्ड-चित्र बना लिया जाता हैं फिर, उसमें से छोटे तथ्य के बराबर विभाग काट लिया जाता है। इस प्रकार, बड़े तथा छोटे दोनों मूल्यों के बीच का 'अन्तर' (Difference) दण्ड के ऊपर भाग में प्रदर्शित हो जाता है। हाँ ! उल्लेखनीय यह है कि अन्तर वाले विभाग को, उस चिन्ह से दिखाया या अंकित किया जाता है जो दोनों में बड़ा हो। (देखिये चित्र-6)।

निम्न उदाहरण में 1981 में निर्यात आयात से अधिक हैं। अतः 'निर्यात-मूल्य (73) के आधार पर दण्ड बनाया गया है और उसमें से आयात-मूल्य (70) के बराबर हिस्सा (काले रंग का भाग) काट दिया गया है। दण्ड का शेष भाग (सफेद भाग) निर्यात्-

विभिन्न संकायों के छात्र संख्या



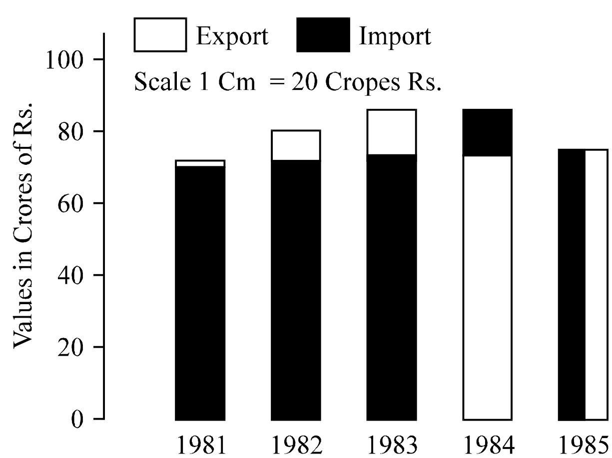
चित्र-5 : अन्तर्विभक्त दण्ड चित्र (Componment Bar Diagram)

आयात के अन्तर अर्थात व्यापार-क्षेत्र को दिखलाता है। 1984 में चूंकि आयात की मात्रा निर्यात से अधिक है। अतः दण्ड 'आयात-मूल्य' के आधार पर बनाया गया हैं। 1985 में आयात और निर्यात दोनों बराबर हैं अतः दण्ड को दो बराबर हिस्सों में उदग्र रूप में बांट दिया गया है।

Illustration 5. Represent the following data by means of sub divided bars :-

Year	Imports	Exports	Balance of Trade
1981	70	73	+ 3
1982	72	80	+ 8
1983	74	85	+ 11
1984	85	80	– 5
1985	75	75	–

आयात, निर्यात तथा व्यापार-शेष (करोड़ रू.)



चित्र-6 : अंतर वाले विभाजित दण्ड-चित्र।

(5) प्रतिशत अन्तर्विभक्त दण्ड-चित्र (Percentage Sub-divided Bars) : यदि एक तथ्य के विभिन्न विभागों से सम्बन्धित समंकों में होने वाले सापेक्ष परिवर्तनों की आपस में तुलना करनी हो तो उसके लिए प्रतिशत आधार पर अन्तर्विभक्त चित्र बनाये जाते हैं। इन चित्रों की रचना के लिए सर्वप्रथम मूल्यों के कुल योग को 100 मानकर सभी विभागों के प्रतिशत निकाल लिये जाते हैं। फिर, उनका संचयी प्रतिशत निकाला जाता है। इसके बाद सरल दण्ड-चित्र बनाकर उसमें संचयी प्रतिशतों के बराबर विभाग काट लिये जाते हैं।

लाभ-हानि चित्र (Profit or Loss Diagram) : जब अन्तर्विभक्त दण्ड-चित्रों का प्रयोग किसी वस्तु के **लागत-तत्वों तथा लाभ-हानि** में होने वाले सापेक्ष परितर्वनों को दिखाने के लिये किया जाता है, तब ऐसे चित्रों को **'लाभ-हानि चित्र'** कहते हैं। ये चित्र प्रतिशत अन्तर्विभक्त चित्रों की तरह ही बनाए जाते हैं परन्तु अन्तर की दष्टि से , इनमें विभिन्न विभागों को **'नीचे से ऊपर'** काटने के बजाय **'ऊपर से नीचे'** काटा जाता है।

Illustration 6. Represent the following by sub-divided bars drawn on the percentage basis :-
निम्न समंकों को प्रतिशत आधार पर बनाये जाने वाले अन्तर्विभक्त दण्ड चित्रों द्वारा प्रदर्शित कीजिये :-

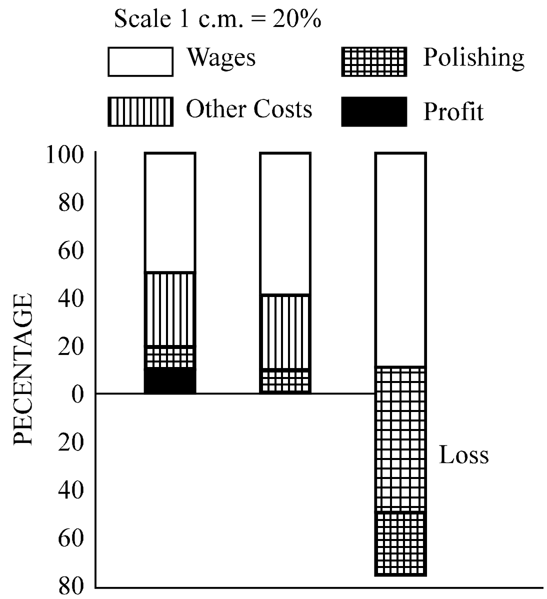
		लागत प्रति कुर्सी (रू. में)	
Particulars	**1938**	**1939**	**1940**
(1) मजदूरी (Wages)	4.5	7.5	10.5
(2) अन्य लगात (Other Costs)	3.0	5.1	7.5
(3) पालिस व्यय (Polishing)	1.5	2.4	3.0
कुल लागत (Total Cost)	9.0	15.0	2.10
विक्रय मूल्य प्रति कुर्सी	10.0	15.0	12.0
लाभ (+) या हानि (–)	1.0	—	(–) 9.0

Solution : सर्वप्रथम, प्रति-कुर्सी आमद (proceed per chair) को 100 मानते हुये सभी तथ्यों को प्रतिशत में बदला जायेगा। फिर, उन्हें संचयी प्रतिशत बनाकर लाभ-हानि चित्र तैयार किया जायेगा (देखिये चित्र-7)

Particulars	**1938**			**1939**			**1940**		
	Rs.	**%**	**Cum. %**	**Rs.**	**%**	**Cum.%**	**Rs.**	**%**	**Cum.%**
Wages	4.5	45	45	7.5	50	50	10.5	87.5	87.5
Other Costs	3.0	30	75	5.1	34	84	7.5	62.5	150
Polishing	1.5	15	90	2.4	16	100	3.0	25.0	175
Total Costs	9.0	90		15.0	100		21.0	175	
Profit/Loss	+ 1	+ 10		—	0		– 9	– 75	
Proceeds	10	100		15	100		12	100	

(6) द्वि-दिशा दण्ड-चित्र (Duo-Directional Bar Diagram) : समंकों के दो परस्पर विरोधी स्वरूपों अर्थात् विपरीत प्रकृति के तथ्यों का प्रदर्शन करने के लिए द्वि-दिशा दण्ड चित्रों का प्रयोग किया जाता है। ये चित्र 'आधार रेखा' के दोनों ओर अर्थात ऊपर और नीचे तैयार किये जाते हैं (देखिये चित्र-8)।

प्रतिशत लागत, विक्रय-मूल्य व लाभ-हानि (प्रति कुर्सी)



चित्र-7 : प्रतिशत दण्ड-चित्र (Percentage Bars)

Illustration 7. निम्न समंकों को द्वि-दिशा दण्ड-चित्रों की सहायता से प्रदर्शित कीजिये :-

भारतीय रेलवे के लाभ करोड़ (रूपये)

Year	Net Profit	Expenses	Gross Profit
1967	6	2	8
1968	5	3	8
1969	3	2	5
1970	6	3	9
1971	4	2	6

Solution : उत्तर के लिये चित्र 8 देखिये।

दो विस्तार वाले या द्वि-विमा-चित्र
(Duo-Dimensional Diagrams)

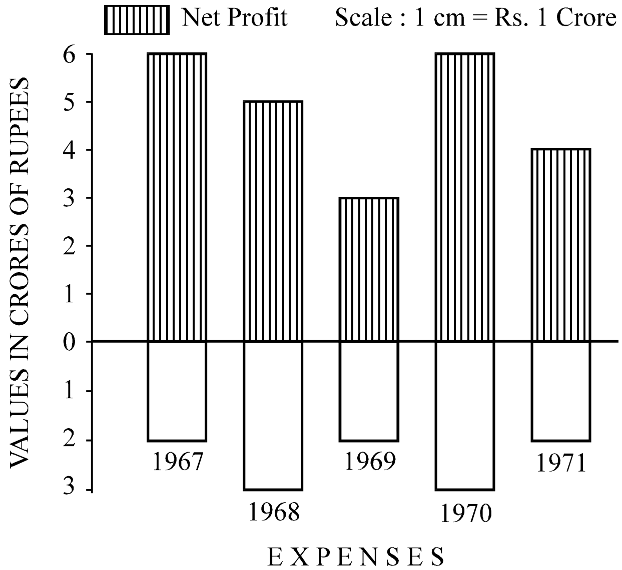
एक-विमा चित्रों अर्थात् दण्ड-चित्रों में केवल एक ही विस्तार का ध्यान रक्खा जाता है जबकि द्वि-विमा चित्रों में दो विस्तारों-ऊंचाई तथा चौड़ाई के आधार पर चित्रों का निर्माण किया जाता है। इन चित्रों के क्षेत्रफल पद-मूल्यों के अनुपात में होते हैं, इसीलिये इनको 'क्षेत्रफल-चित्र' (Area Diagram) अथवा 'धरातल-चित्र' (Surface Diagram) भी कहते हैं। द्वि-विमा चित्र मुख्यतया तीन प्रकार के होते हैं :-

(1) आयत चित्र (Rectangular Diagrams)

(2) वर्ग चित्र (Square Diagrams)

(3) वत्तीय चित्र (Circular or Pie Diagrams)

भारतीय रेलवे का शुद्ध लाभ तथा व्यय



चित्र-8 : द्वि-दिशा दण्ड चित्र (Duo-Directional Bars)

(1) आयत चित्र (Rectangular Diagrams) : दो या दो से अधिक राशियों की पारस्परिक तुलना करने के लिये आयत चित्रों का प्रयोग किया जाता है। आयत चित्र दो प्रकार के होते हैं :-

(*i*) **प्रतिशत अन्तर्भिक्त आयत चित्र (Percentage Sub-divided Rectangle)** : जब किसी विभिन्न परिवारों के पारिवारिक बजट की तुलना करनी हो, तो उसके लिये आयत चित्रों का ही प्रयोग किया जाता है। ऐसे चित्रों की रचना करने के लिये सबसे पहले, दिये हुये परिवारों के कुल आय को 100 मानकर विभिन्न मदों पर होने वाली व्यय-राशियों को प्रतिशत में बदल लिया जाता है। इसके बाद 100 के बराबर मापदण्ड लेकर प्रत्येक परिवार के लिये एक-समान ऊंचाई वाले आयत बना लिये जाते हैं। हाँ ! इन आयतों की चौड़ाई कुल-व्यय के अनुपात में रखी जाती है। तत्पश्चात् व्यय की प्रतिशत राशियों के अनुसार, आयत को 'नीचे से ऊपर' की ओर विभिन्न खण्डों में बांट दिया जाता है।

Illustration 8 : नीचे दिये गये परिवारों के मासिक व्यय (रू. में), सम्बन्धी विवरण को द्वि-विमा द्वारा प्रस्तुत कीजिये :-

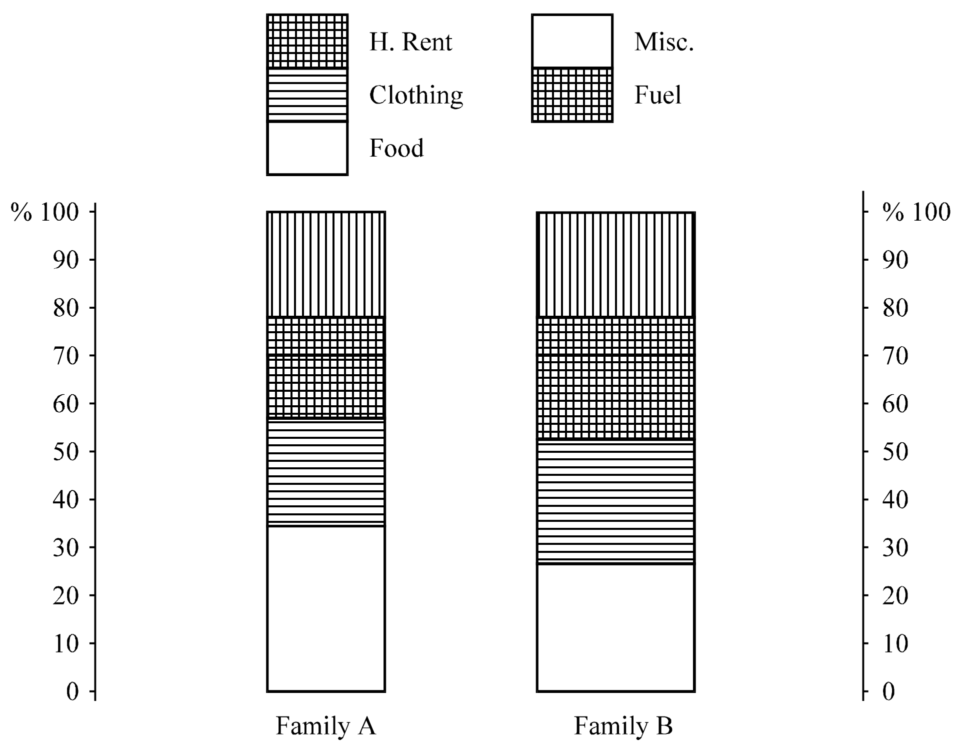
व्यय की मदें (Items of Expenditure)	**परिवार (A)**	**परिवार (B)**
भोजन (Food)	200	250
वस्त्र (Clothing)	100	200
मकान किराया (House Rent)	80	100
ईंधन (Fuel)	40	50
विविध (Miscellaneous)	80	200
कुल व्यय (Total Expenditure)	500	800

Solution : दोनों परिवार की आय को 100 मानते हुये विभिन्न मदों को निम्न ढंग से प्रतिशतों में बदला जायेगा और इसके बात अन्तविभक्त आयत-चित्र की रचना की जायेगी। दोनों आयतों की चौड़ाई उनके आयतों की चौड़ाई उनके मासिक व्यय 500 : 800 अर्थात् 5 : 8 के अनुसार में रक्खी जायेगी (देखिये चित्र 9)।

Items of Expenditure	Family A (Rs. 500)			Family B (Rs. 800)		
	Rs.	%	Cum. %	Rs.	%	Cum. %
Food	200	40	40	250	31.3	31.3
Clothing	100	20	60	200	25.0	56.3
House Rent	80	16	76	100	12.5	68.8
Fuel & Lighting	40	8	84	50	6.2	75.0
Miscellaneous	80	16	100	200	25.0	100.0
Total	**500**	**100**		**800**	**100**	

(1) विभाजित आयत-चित्र (Sub-divided Rectangles) : जब तीन विभिन्न किन्तु परस्पर संबंधित तथ्यों का चित्रमय प्रदर्शन करना हो तब विभाजित आयत-चित्रों का प्रयोग किया जाता है जैसे किसी वस्तु का प्रति इकाई मूल्य, बिक्री की मात्रा तथा विक्रय-राशि का एक साथ प्रदर्शित किया जाना। ध्यान रहे, ऐसे चित्रों में चौड़ाई सदैव प्रति-इकाई मूल्य के अनुपात में ली जाती है जबकि ऊंचाई बिक्री की मात्रा के अनुपात में रखी जाती है। आयत का क्षेत्रफल (ऊंचाई × चौड़ाई) कुल विक्रय-मूल्य को व्यक्त करता है। शेष क्रिया प्रतिशत अन्तर्विभक्त आयत चित्र के समान ही है।

दो परिवारों का मासिक व्यय



चित्र-9 : प्रतिशत अन्तर्विभक्त आयत

(2) वर्ग चित्र (Square Diagrams) : जब तथ्यों के न्यूनतम और अधिकतम मूल्यों में काफी अन्तर होता है तो उनसे दण्ड-चित्र नहीं बनाये जा सकते। उदाहरणार्थ यदि न्यूनतम और अधिकतम मूल्य क्रमशः 4000 और 100 हैं तो सबसे बड़ा दण्ड, सबसे छोटे दण्ड से 40 गुना होगा। भले ही कोई-सा पैमाना क्यों न ले लिया जाये, इस अनुपात को व्यक्त करना संभव नहीं है। अतः ऐसी स्थिति में दण्ड-चित्रों के स्थान पर 'वर्ग-चित्रों' का निर्माण किया जाता है। **रचना विधि** — (*i*) सबसे पहले समंकों का वर्गमूल निकाला जाता है। (*ii*) फिर, वर्गमूलों के अनुपात में वर्गों की रचना कर ली जाती है। हाँ ! उल्लेखनीय यह है कि सभी वर्ग सदैव एक ही धरातल पर बनाये जायें ताकि वे देखने में सुन्दर लगें और उनकी सरलता से तुलना की जा सके।

Illustration 9. चार देशों में गेहूँ का उत्पादन (करोड़ टन) में नीचे दिया गया है। इन समंकों को एक उर्पयुक्त चित्र द्वारा प्रस्तुत कीजिये :-

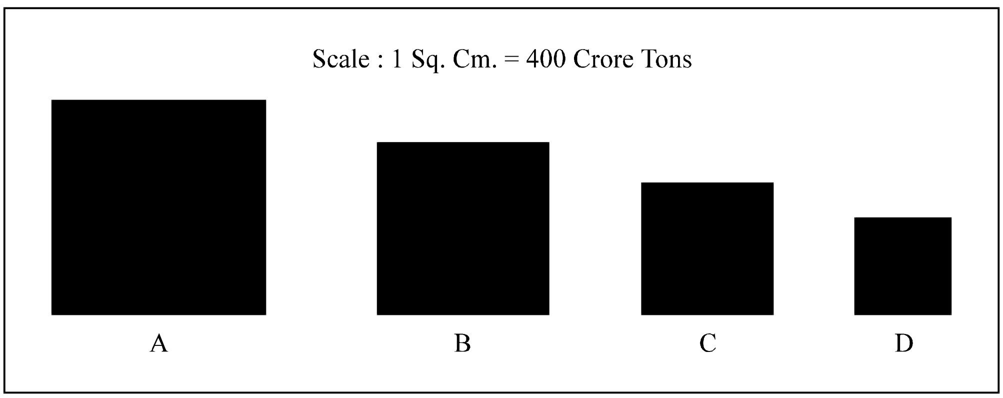
Countries	:	A	B	C	D
Quantity (Crore Tons)	:	3600	2000	1200	800

Solution : चूँकि अधिकतम मूल्य (3600) और न्यूनतम मूल्य (800) में काफी अन्तर है। अतः इन समंकों को वर्ग-चित्र द्वारा प्रदर्शित करना अधिक उपर्युक्त होगा।

देश (Country)	उपज (करोड़ टन) (Quantity)	वर्गमूल (Square-root)	वर्ग की भुजा (In cms.)
A	3600	60	3
B	2000	44.7	2.2
C	1200	24.6	1.7
D	800	28.3	1.4

वर्गों की भुजाओं की लम्बाई ज्ञात करना : वर्गों की रचना करते समय एक प्रश्न यह उठता है कि वर्गों की भुजाओं की लम्बाई क्या रखी जाए ? यह तथ्य उपलब्ध स्थान अर्थात् कागज के आकार पर निर्भर करता है। मान लीजिये, वर्ग रचना के लिये हमारे पास कुल 12 cm का स्थान उपलब्ध है जिसमें से लगभग 4 cm वर्गों के बीच के रिक्त स्थान के लिये छोड़ना होगा।

विभिन्न देशों मे गेहूँ का उत्पादन



चित्र-10 : वर्ग–चित्र (Squre Diagram)

यदि शेष बचे 8 cm से, वर्गमूलों के जोड़ को विभाजित कर दिया जाये तो इससे वह संख्या निकल जाती है जिससे प्रत्येक वर्गमूल को भाग देने पर, उस वर्ग की वांछित लम्बाई ज्ञात हो जाती है। उपर्युक्त उदाहरण में वर्गमूलों का जोड़ 167.6 है जिसको 8 से भाग देने पर लगभग 20.9 मूल्य आता है जिसके सन्निकट (Approximated) सरल मूल्य 20 से सभी वर्गमूलों को भाग देकर भुजाओं की लम्बाई ज्ञात की गई है (देखिये चित्र 10)।

वर्गों का पैमाना ज्ञात करना (Scales of Squares) : वर्गों का पैमाना निकालने के लिये कभी भी एक वर्ग की भुजा की लम्बाई का वर्ग करके, उसका क्षेत्रफल निकाल लिया जाता है। उदाहरणार्थ 3 cm की भुजा वाले वर्ग का क्षेत्र $3 \times 3 = 9$ cm हुआ जो 3600 करोड़ टन की मात्रा को प्रदर्शित करता है अतः

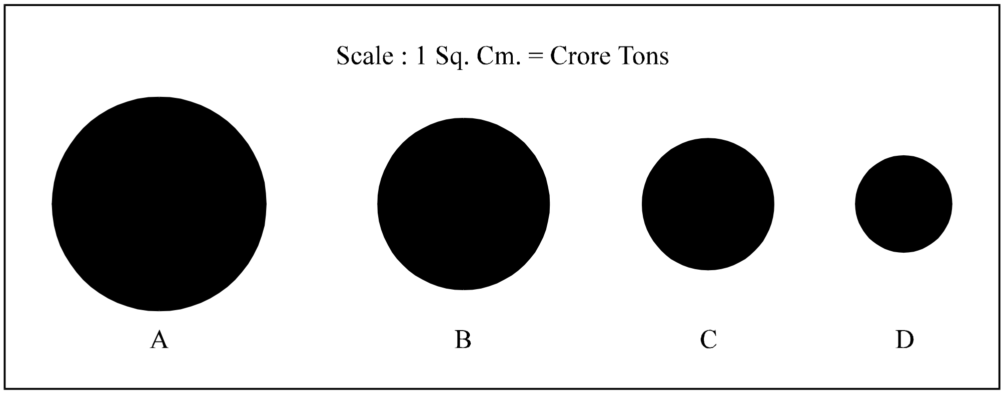
$$1 \text{ sq. cm} = 3600 \div 9 = 400 \text{ crore tons.}$$

(3) क्तीय चित्र (Circular or Pie Diagrams) : जिन तथ्यों के लिये वर्ग-चित्रों का प्रयोग उपर्युक्त समझा जाता है ठीक उन्हीं के लिये वत्त चित्रों का भी उपयोग किया जा सकता है। वत्त चित्र की रचना वर्ग -चित्र की भांति ही की जाती है। वत्त -चित्रों का निर्माण करने के लिये पहले मूल्यों के वर्गमूल निकाले जाते हैं। फिर, वर्गमूलों के अनुपात में वत्तों की त्रिज्याएं (Radius) ज्ञात कर ली जाती है जिनकों आधार मानकर वत्त तैयार कर दिये जाते हैं। ध्यान रहे , वत्तों को सदैव ही धरातल पर बनाना चाहिये और उनके बीच का अन्तर भी समान रखना चाहिये ताकि चित्र देखने में सुन्दर लग सकें।

Illustration 10. Represent the data given in illustration 9 by a Circular or Pie Diagram.

पिछले उदाहरण में वर्ग की भुजायें क्रमशः 3, 2.2, 1.7, 1.4 cm हैं जिनका आधा करने पर चारों वत्तों की त्रिज्याएं क्रमशः 1.5, 1.1, 0.85, 0.7 होंगी (देखिये चित्र 11)।

विभिन्न देशों में गेहूँ का उत्पादन



चित्र-11 : क्त -चित्र (Pie Diagram)

क्त -चित्र का पैमाना ज्ञात करना (Scale of Pie-Diagram) : पहले किसी एक वत्त की त्रिज्या के आधार पर उसका क्षेत्रफल निकाल लिया जाता है। वत्त का क्षेत्र πr^2 के बराबर होता है। जहां, p (Pie) = $\frac{22}{7}$ और r = radius अर्थात् त्रिज्या। उपर्युक्त उदाहरण में प्रथम, क्त , जिसकी त्रिज्या 1.5 cm है, का क्षेत्रफल इस प्रकार निकलेगा :-

क्षेत्रफल = $\frac{22}{7} \times (1.5)^2$ या $\frac{22}{7} \times 2.25 = \frac{49.5}{7}$ वर्ग सेमी. $\frac{3600 \quad 7}{7 \quad 49.5}$

वर्ग सेमी. दिखाता है 3600 करोड़ टन को

\ 1 वर्ग सेंमी. दिखायेगा = = 509 करोड़ टन

कोणीय चित्र या वत्त -खंड चित्र (Angular or Sector Diagram) : जिस प्रकार तथ्यों के उपविभागों को प्रदर्शित करने के लिए दण्ड और आयत को विभक्त किया जाता है, उसी प्रकार वत्त को भी अन्तर्विभक्त किया जा सकता है और ऐसे चित्रों को ही वत्त -खण्ड चित्र कहते हैं। इन चित्रों की रचना के लिये समंकों को कुल योग को 360 डिग्री मानते हुए विभिन्न विभागों के कोणों (Angles) का माप निकाल दिया जाता है क्योंकि वत्त के केन्द्र में 360° का कोण होता है। हाँ ! यदि दो कोणिय क्त एक -साथ बनाने हों तो उनकी त्रिज्याएं उनके कुल जोड़ के वर्गमूल के अनुपात में रखी जाती है (देखिये चित्र 12)।

Illustration 11. Represent by suitable diagram the following data of the outlay in the second and third five year plans.

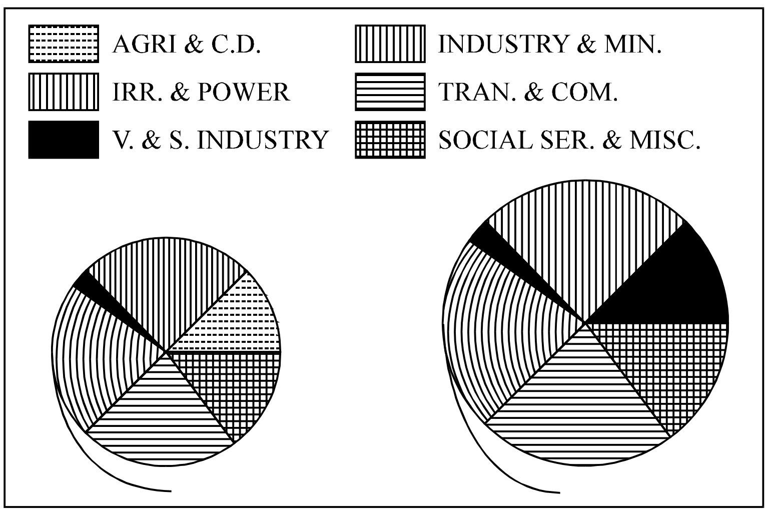
परियोजन व्यय का आबंटन (करोड़ रू.)

मद (Head)	Second Plan	Third Plan
Agriculture and C.D.	530	1068
Irrigation of Power	865	1662
Village and Small Industry	175	264
Industry and Minerals	900	1520
Transport and Communiction	1300	1486
Social Services and Misc.	830	1500
	4600	**7500**

Solution : सर्वप्रथम दोनों योजनाओं के कुल व्यय का वर्गमूल निकालकर (अर्थात् $\sqrt{4600} = 67.8$ व = $\sqrt{7500} = 86.6$), उनका अनुपात निकाला जायेगा ताकि त्रिज्याओं की लम्बाई ज्ञात हो सके। तत्पश्चात् 360º के आधार पर विभिनन मदों के कोणीय माप ज्ञात किये जायेंगे (देखिये चित्र 12)।

Head	**Second Plan**		**Third Plan**	
	Rs. Crores	**Degrees**	**Rs. Crores**	**Degrees**
Agriculture and C.d.	530	41	1068	51
Irrigation and Power	865	68	1662	80
Villages and Small Industry	175	14	246	13
Industry and Minerals	900	70	1520	73
Trans and Communication	1300	102	1486	71
Social Service and Misc.	830	65	1500	72
Radius	4600	360º	7500	360º
	67.8	2.26 cm	86.6	2.90 cm.

द्वितीय एवं ततीय योजना में व्यय -वितरण



चित्र-12 : कोणीय चित्र (Angular or Pie Diagram)

त्रि-विमा चित्र
(Three-Dimentional Diagram)

जब कभी मूल्यों में अत्यधिक अन्तर होता है तो उनसे वर्ग या वत्त -चित्र बनाना भी कठिन हो जाता है क्योंकि मूल्यों के वर्गमूलों में काफी अंतर बना रहता है। ऐसी दशा में सदैव तीन विस्तार वाले चित्र अर्थात् त्रि-विमा चित्रों का प्रयोग करना चाहिये। इन चित्रों के निर्माण में तीनों विस्तारों — ऊंचाई, चौड़ाई व मोटाई का एक-साथ इस्तेमाल किया जाता है। त्रि-विमा चित्रों में घन (Cubes), इष्टका (Blocks), गोलाकार (Spheres) तथा बेलनाकार चित्र इत्यादि तैयार किये जाते हैं। त्रि-विमा काफी जटिल होते हैं। नीचे केवल घन-चित्र की रचना विधि स्पष्ट की गयी है।

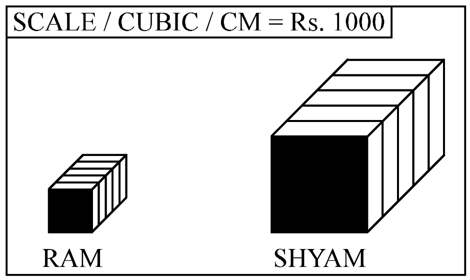
घन-चित्र की निर्माण विधि — (*i*) सर्वप्रथम, मूल्यों के घनमूल (Cubes Roots) निकाले जाते हैं। घनमूल निकालने के लिए मूल्य का लघुगणक (log) ज्ञात करके उसे यदि 3 से भाग दिया जाय और प्राप्त भजनफल का Antilog ले लिया जाय तो वह घनमूल होता हे। (*ii*) तत्पश्चात घनमूलों के अनुपात में घनों की भुजाएं ली जाती हैं। (*iii*) घन की भुजा के आधार पर पहले एक वर्ग तैयार किया जाता है, फिर उसी क्षेत्रफल में दूसरा वर्ग इस प्रकार तैयार किया जाता है कि उसका बायां निचला

कोना, पहले वाले वर्ग के केन्द्र-बिन्दु पर स्थित हो और उनकी भुजाएं परस्पर समानान्तर बनी रहें। (*iv*) अन्त में इन दोनों वर्गों के कोनों को मिला देने पर घनचित्र पूरा हो जाता है (देखिये चित्र 13)।

Illustration 12. राम और श्याम का मासिक वेतन क्रमशः 512 रू. व 4096 रू. है। इन समंकों को त्रि-विमा चित्र द्वारा प्रकट कीजिए।

Solution :

		वेतन (रू.)	घनमूल	घन की भुजा (सेंमी.)
Ram	:	512	8	1
Shyam	:	4096	16	2

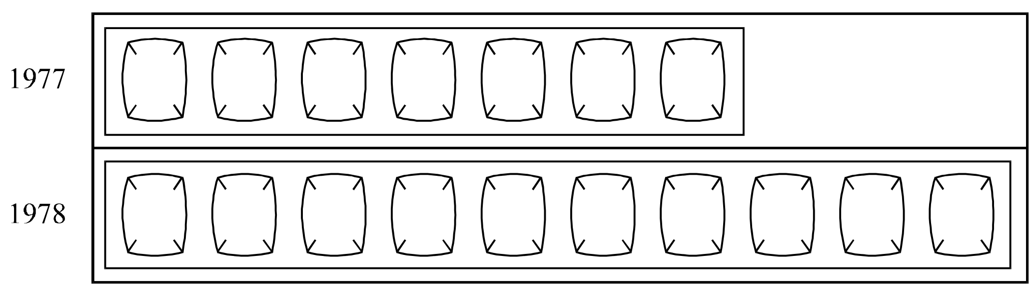


चित्र-13 : घन-चित्र (Cube Diagram)

चित्रलेख एवं मानचित्र
(Pictograms and Cartograms)

चित्र-लेख — चित्र-लेख पद्धति का विकास डॉ. ऑटो न्यूरथ (Otto Neurath) ने किया था। इस रीति के अन्तर्गत समंकों को सम्बन्धित वस्तुओं की आकृति या चित्र बनाकर प्रदर्शित किया जाता है। उदाहरणार्थ जनसंख्या समंकों को मनुष्य-आकृति द्वारा, योजना-व्यय को रूपयों की थैली द्वारा और खाद्यान्न उत्पादन को बोरियों के चित्र बनाकर प्रदर्शित किया जा सकता है। इन चित्रों की रचना समंकों के अनुपात में की जाती है। ये चित्र काफी आकर्षक होते हैं और देखने पर अधिक अच्छा प्रभाव डालते हैं। यही कारण है कि विज्ञापन आदि के लिये इन चित्रों का प्रयोग अधिक किया जाता है। चित्र 14 में दो वर्षों के खाद्यान्न उत्पाद को बोरियाँ बनाकर दिखाया गया है।

1977 व 1978 में खाद्यान्न उत्पादन
(1 बोरी = 10 मि. टन)



चित्र-14 : चित्र लेख (Pictogram)

मानचित्र : मानचित्र भौगोलिक या प्रादेशिक समंकों के लिये अधिक उपर्युक्त होते हैं। उदाहरणार्थ किसी देश में जनसंख्या-घनत्व, वर्षा, वनस्पति-वितरण, कृषि उपज, खनिज-स्थिति, आदि को प्रदर्शन उस देश के मानचित्र पर आकर्षक ढंग से किया जा सकता है। हाँ ! मानचित्र में जिन-जिन तथ्यों को दिखाया गया है उनके स्पष्ट संकेत या चिन्ह अवश्य बना देने चाहिएं।

उपयुक्त चित्र का चुनाव

समंकों का चित्रमय निरूपण करते समय किस चित्र-विशेष का प्रयोग किया जाये, यह वास्तव में एक कठिन समस्या है। इसका कारण यह है कि एक-ही प्रकार के समंकों के लिये कई प्रकार के चित्र बनाए जा सकते हैं।

❑❑

अध्याय - 5

समंकों का बिन्दुरेखीय प्रदर्शन
(Graphic Presentation of Data)

पिछले अध्याय में विभिन्न प्रकार के चित्रों का वर्णन किया गया। चित्रों का प्रयोग विशेष रूप से स्थान संबंधी श्रेणियों (Spatial Series) के चित्रण के लिये किया जाता है। लेकिन काल श्रेणियों (Time Series) तथा आवत्ति वितरणों के चित्रण के लिये बिन्दु रेखीय विधि (Graphical Method) अधिक उपयुक्त है। समकों के बिन्दुरेखीय प्रदर्शन के लिये बिन्दूरेखीय पत्र (Graph Paper) का प्रयोग किया जाता है। दिये हुए समंकों को किसी मापदण्ड के आधार पर एक बिन्दुरेखीय पत्र पर अंकित कर दिया जाता है और प्रांकित विभिन्न बिन्दुओं को मिला दिया जाता है।

चित्रों तथा बिन्दुरेखीय चित्रों में अंतर
(Difference Between Diagrams and Graphs)

(1) बिन्दुरेखीय चित्र बनाने के लिये बिन्दुरेखीय पेपर का प्रयोग किया जाता है जबकि चित्र साधारण कागज पर बनाया जा सकता है।

(2) चित्र अधिक आकर्षक होते हैं, इसलिये प्रचार तथा विज्ञापन के लिये बहुत उपयुक्त हैं। चित्रों का प्रत्यक्ष रूप से सांख्यिकी का अर्थ समझने में कोई योग नहीं होता, चित्र केवल प्रदर्शन मात्र होते हैं परन्तु रेखाचित्रों का प्रयोग अनुसंधानकर्त्ता अंकों के विश्लेषण में भी करते हैं।

(3) काल श्रेणियों तथा आवत्ति वितरणों को प्रदर्शित करने के लिये बिन्दुरेखीय चित्र , चित्रों की उपेक्षा अधिक उपयोगी हैं।

(4) विशेष प्रकार के आवत्ति रेखाचित्रों द्वारा भूयिष्ठक (Mode) , मध्यका (Median) तथा विभाजन मूल्यों (Partition Values) का निर्धारण किया जा सकता है परन्तु चित्रों द्वारा ऐसा संभव नहीं है।

(5) बिन्दुरेखीय विधि द्वारा समंकों का आन्तर्गणन (Interopolation) तथा बाह्यगणन (Extrapolation) सरलता एवं शीघ्रता से किया जा सकता है। यदि एक वक्र की ठीक प्रकार से रचना किए जाने पर कोई आकस्मिक परिवर्तन नहीं हुआ है तो ये अनुमान बहुत कुछ ठीक सिद्ध होते हैं। चित्रों द्वारा ऐसा संभव नहीं।

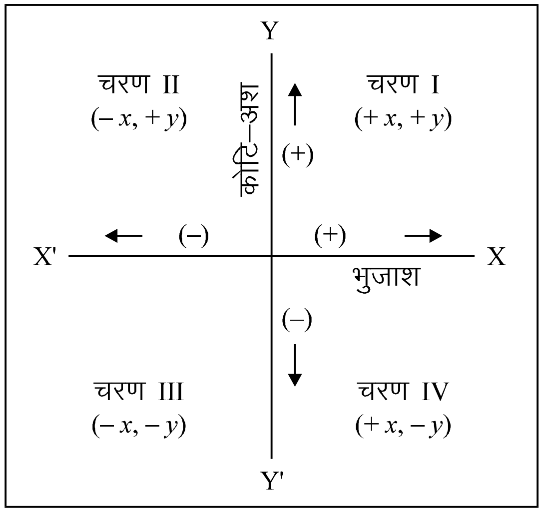
रेखाचित्र कई प्रकार के होते हैं। यहां केवल प्रमुख रूप से प्रयुक्त होने वाले रेखाचित्रों का विवरण दिया गया है। सुगमता के लिये विभिन्न रेखाचित्रों को दो भागों में बांटा गया है :-

(1) काल मालाओं के रेखाचित्र (Graph of Time Series); तथा

(2) आवत्ति -वितरणों के रेखाचित्र (Graphs of Frequency Distribution)।

रेखाचित्र बनाने की विधि
(Technique of Constructing Graphs)

रेखाचित्र बनाने के लिये प्राय: बिन्दु रेखीय पत्र (Graph Paper) का प्रयोग किया जाता है। दो सरल रेखायें बिन्दुरेखीय पत्र पर इस प्रकार डाली जाती है कि वे एक दूसरे को 90 डिग्री के कोण पर काटें। जिस बिन्दु पर दोनों रेखायें एक-दूसरे को काटती है उसे मूल-बिन्दु (Point of Origin) अथवा शून्य बिन्दु (Zero Point) कहते हैं। पड़ी रेखा (Horizontal Line) को क्षैतिज माप श्रेणी (X-Axis) या भुजाक्ष (Abscissa) कहते हैं और खड़ी रेखा (Vertical Line) को उदग्र माप श्रेणी (Y-Axis) या कोटि-अक्ष (Ordinate) कहते हैं। ये दो रेखायें निम्न चित्र में दिखाई गई हैं :-



ऊपर की आकृति में 'O' मूल-बिन्दु है। XOX' भुजाक्ष (X-Axis) तथा YOY' कोटि (Y-Axis) है। धनात्मक (Positive) तथा ऋणात्मक (Negative) दोनों प्रकार के समंक रेखाचित्र द्वारा भी प्रदर्शित की जा सकती है। कटाव बिन्दु के दायीं ओर तथा ऊपर धनात्मक मूल्यों तथा नीचे व बायीं ओर ऋणात्मक मूल्यों को नापा जाता है।

बिन्दुरेखीय पत्र को चार भागों में बांटा जाता है जिसमें प्रत्येक भाग को चरण (Quadrant) कहते हैं। पहले चरण में X तथा Y दोनों के मान ऋणात्मक होते हैं। सामान्यत:, समंक धनात्मक होते हैं अत: व्यवहार में अधिकांशत: रेखाचित्र प्रथम चरण (First Quardrant) में बनाये जाते हैं।

प्रथा के अनुसार स्वतंत्र चर (Independent Variable) जैसे समय, आकार आदि को पड़ी रेखा (Horizontal Scale) तथा आश्रित चर (Dependent Variable) को खड़ी रेखा (Vertical Scale) पर दिखाया जाता है। काल मालाओं में समय क्षैतिज पैमाने (Horizontal Scale) पर तथा चर की उदय पैमाने (Vertical) पर दिखाया जाता है। प्रत्येक अक्षर के लिये उचित मापदंड निश्चित कर लिया जाता है। मापदंड भुजांक्ष के नीचे तथा कोटि अक्ष की बायीं ओर लिख दिया जाता है साधारण तथा रेखाचित्र में प्राकृतिक पैमाने का प्रयोग होता है। आवश्यकतानुसार "आनुपातिक पैमाने" (Ratio Scale) का प्रयोग किया जा सकता है।

मापदण्ड निर्धारण के पश्चात् अन्तिम कार्य दिये गये समंकों को आलेखित करना होता है। इस प्रकार उपलब्ध चिह्नों को सीधी रेखाओं के द्वारा मिला दिया जाता है।

(क) काल-श्रेणी के रेखाचित्र या कालिक-चित्र
(Graphs of Time Series of Historigrams)

समय के आधार पर अवलोकनों की श्रेणी 'काल-श्रेणी' कहलाती है। बिन्दुरेखीय प्रणाली समय-समय पर होने वाले परिवर्तनों का विश्लेषण करने में अत्यन्त सहायक X-अक्ष पर प्राय: समय दिया जाता है तथा Y-अक्ष पर चर के विभिनन मान। एक से अधिक चर भी एक रेखाचित्र पर दिखाये जा सकते हैं और उनमें आपस में तुलना की जा सकती है।

काल मालाओं के रेखाचित्र प्राकृतिक पैमाने अथवा आनुपातिक पैमाने के आधार पर बनाये जा सकते हैं। प्राकृतिक पैमाने में एक समय से दूसरे समय में निरपेक्ष परिवर्तन तथा आनुपातिक पैमने में एक समय से दूसरे समय के आनुपातिक परिवर्तन या परिवर्तन की दर (Relative Change or Rate of Change) को प्रदर्शित किया जाता है।

प्राकृतिक माप श्रेणी के कालिक चित्र निम्न दो प्रकार के हो सकते हैं :-

(1) **निरपेक्ष कालिक चित्र (Absolute Historigram) :** निरपेक्ष कालिक चित्र में समंक श्रेणी की मूल राशियों या समंकों को प्रांकित किया जाता है ये राशियां माप की विभिन्न इकाइयों जैसे टन, किलोग्राम, करोड़ रुपये आदि में व्यक्त की जा सकती है।

(2) **निर्देशांक कालिक चित्र (Index Historigram) :** निर्देशांक कालिक चित्र में वास्तविक मूल्यों के स्थान पर उनके

निर्देशांकों (Index Numbers) को बिन्दुरेखीय पत्र पर अंकित किया जाता है। निर्देशांकों या सूचकांकों द्वारा मूल्यों के सापेक्ष या तुलनात्मक अन्तरों का मापन होता है।

प्राकृतिक पैमाने पर रेखाचित्र बनाने के नियम
(Rules of Constructing the Line Graphs on Natural Scale)

(1) **शीर्षक (Heading) :** प्रत्येक रेखाचित्र का एक उपर्युक्त शीर्षक होना चाहिए। शीर्षक संक्षिप्त तथा स्पष्ट होना चाहिये ताकि उसको पढ़ते की विषय-सामग्री का ज्ञान हो जाये।

(2) **मापदण्ड का चुनाव (Choice of Scale) :** मापदण्ड का चुनाव ऐसा होना चाहिये जिससे सम्पूर्ण समंकों का प्रदर्शन किया जा सके। मापदण्ड प्रदर्शित करने वाले अंकों को स्पष्ट रूप से भुजाओं के नीचे तथा कोटि-अक्ष के बायीं ओर लिख देना चाहिए। रेखाचित्र के ऊपर की ओर किसी विशिष्ट स्थान पर मापदण्ड का वितरण भी दे देना चाहिये ताकि पाठक को रेखाचित्र समझने में कोई कठिनाई न हो।

(3) **अक्षों का अनुपात (Proportion of Axis) :** जहां तक हो सके भुजाक्ष की लम्बाई कोटि-अक्ष की लम्बाई से लगभग डेढ़ गुनी रखनी चाहिये।

(4) **कृत्रिम आधार रेखा का प्रयोग (Use of False Base Line) :** उदग्र पैमान एक से प्रारम्भ होता है लेकिन यदि आश्रित चर व मूल्य अधिक आकार के हैं तो मूल बिन्दु से कुछ ऊपर किसी अन्य संस्था से पैमाना आरम्भ किया जा सकता है। ऐसी स्थिति में कृत्रिम आधार रेखा का प्रयोग किया जाता है।

(5) **समंकों की सारणी (Table of Data) :** रेखाचित्र के साथ-साथ सम्बन्धित सारणी भी देनी चाहिये ताकि विस्तत अध्ययन तथा शुद्धता की जांच की सुरक्षा बनी रहे।

(6) **विभिन्न प्रकार की रेखायें (Different Types of Curves)** रेखाचित्र में प्रत्येक बिन्दु को स्पष्ट रूप से अंकित करना चाहिये। यदि एक से अधिक वक्रों का प्रदर्शन करना हो तो विभिन्न प्रकार की रेखायें खींचनी चाहियें, जैसे

———————— सरल रेखा (Simple Line)

- - - - - - - - - - - - - - - - - टूटी रेखा (Broken Line)

.... बिन्दु रेखा (Dotted Line)

-.-.-.-.-.-.-.-.-.-.-.-.-.-.-.-. बिन्दु-विराम रेखा (Dot-dash Line)

विभिन्न प्रकार के रंगों आदि का भी प्रयोग किया जा सकता है।

(7) **संकेत (Index Indicator) :** विभिन्न रेखाओं का अन्तर स्पष्ट करने के लिये रेखाचित्र में ऊपर की ओर उनसे सम्बन्धित संकेत दे देना चाहिये।

एक चर के रेखाचित्र
(Graphs of One Variable)

एक चर को प्रदर्शित करने के लिये X-अक्ष (X-axis) पर समय दिखाना चाहिये तथा Y-अक्ष (Y-axis) पर चर के विभिन्न मानों को प्रदर्शित करना चाहिये। फिर दिये हुए समंकों को समय और चर के मानों के अनुसार रेखाचित्र पर अंकित कर दिया जाता है एवं सब बिन्दुओं को रेखाओं से मिला देते हैं। इस रेखा का उतार-चढ़ाव (Fluctuating) उस चर में दिये हुए परिवर्तन को बहुत की सुचारू रूप में दर्शाता है।

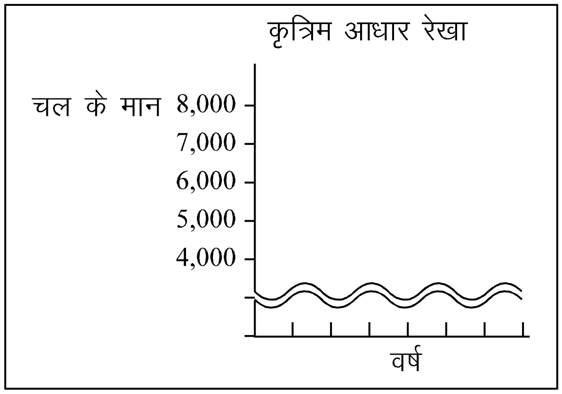
कृत्रिम आधार रेखा
(False Base Line)

रेखाचित्र बनाने का एक नियम यह है कि Y-अक्ष शून्य से शुरू हो चाहे दिये दिये हुये समंकों में चर का मूल्य कितना भी हो। इस नियम का पालन किया जाये तो समंकों को दिखाने के लिये बहुत अधिक स्थान की आवश्यकता पड़ती है फिर भी समंकों के उतार-चढ़ाव को भली प्रकार से प्रदर्शित नहीं किया जा सकता। जब सभी शून्य तथा चर के न्यूनतम मान (Minimume Value) में बहुत अन्तर हो तो कृत्रिम आधार रेखा का प्रयोग किया जाता है। कृत्रिम आधार के प्रयोग के दो प्रमुख उद्देश्य हैं :-

(1) समंकों के उतार-चढाव को रेखाचित्र द्वारा इस प्रकार प्रदर्शित करना चाहिये कि रेखाचित्र देखने वाला इस उतार-चढ़ाव के महत्व को आसानी से समझ सके।

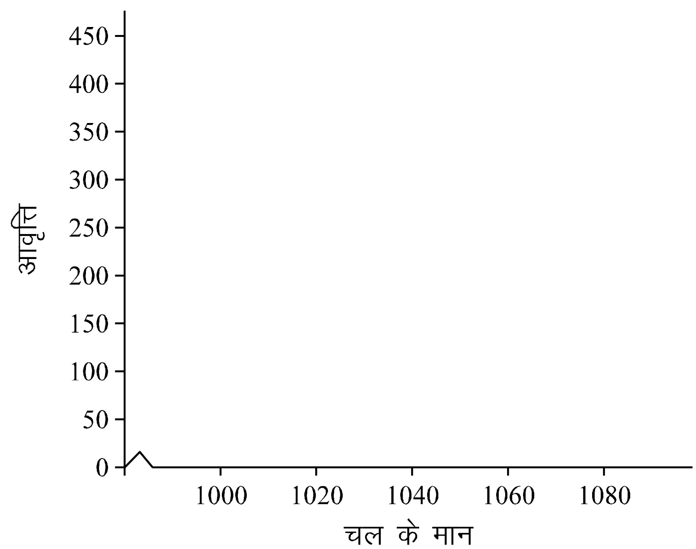
(2) रेखा-चित्र बनाने में स्थान कम लगे।

जब कृत्रिम आधार का प्रयोग किया जाता है तो शून्य और चल के न्यूनतम मान के बीच दो वक्र रेखायें खींच दी जाती है जैसा कि निम्न उदाहरण से स्पष्ट हो जायेगा :



मान लीजिये चल का न्यूनतम मान 4,250 है। हम 0 और 4,000 के बीच दो रेखायें इस प्रकार खींचेंगे :-

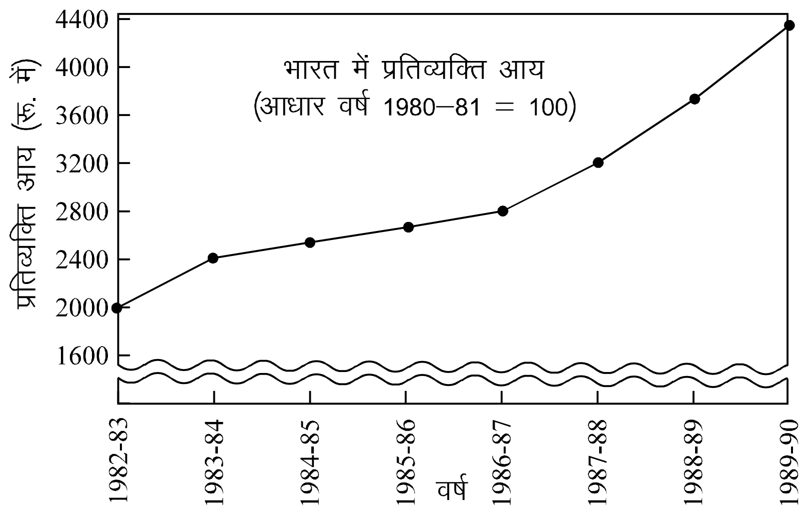
यदि X अक्ष पर हम शून्य से प्रारम्भ करना चाहें जबकि श्रेणी का न्यूनतम मान शून्य से बहुत अधिक है तो आधार रेखा का प्रारम्भ से थोड़ा ऊपर ले जाकर आधार रेखा के साथ मिला दिया जाता है। ऐसा करने से यह स्पष्ट हो जाता है कि कृत्रिम आधार पर प्रयोग किया गया है। निम्न चित्र द्वारा इस बात का स्पष्टीकरण हो जायेगा :



यह अनिवार्य नहीं है कि आधार रेखा शून्य से ही प्रारम्भ हो यद्यपि प्राय: ऐसा करते हैं। यदि आधार को श्रेणी के न्यूनतम मान से शुरू किया जाये तो इस क्रिया की आवश्यकता नहीं रहती जो ऊपर के चित्र में दिखाई गई है।

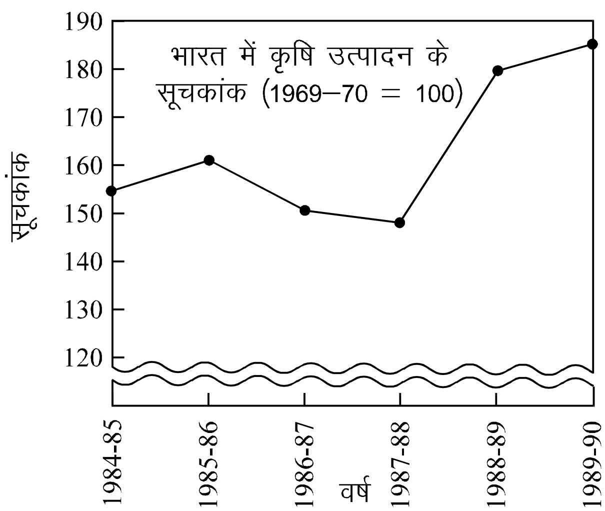
उदाहरण 1 : निम्न समंकों को रेखाचित्र द्वार प्रस्तुत कीजिये :-

वर्ष	प्रति व्यक्ति आय (रू. में) (1980-81 के आधार पर)	वर्ष	प्रति व्यक्ति आय (रू. में) (1980-1981 के आधार पर)
1982-83	2000.1	1986-87	2953.6
1983-84	2300.4	1987-88	3286.1
1984-85	2504.2	1988-89	2875.2
1985-86	2726.0	1989-90	4252.4



उदाहरण 2 : निम्न समंकों को रेखाचित्र द्वारा प्रदर्शित कीजिये :-

वर्ष	कृषि उत्पादन का सूचकांक	वर्ष	कृषि उत्पादन का सूचकांक
1984-85	154.6	1987-88	151.3
1985-86	158.4	1988-89	183.2
1986-87	152.5	1989-90	186.4

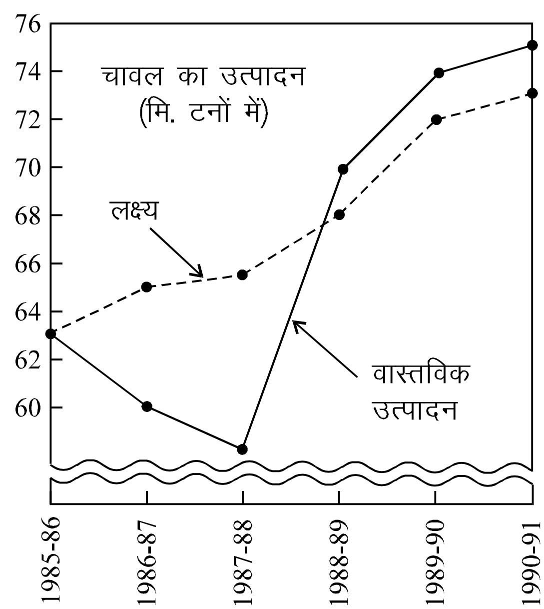


दो यो दो से अधिक चरों के रेखा-चित्र
(Graphs of Two or More Variables)

यदि मापदण्ड की इकाई (Unit of Measurement) समान है तो एक ही रेखाचित्र पर एक से अधिक चरों को प्रदर्शित किया जा सकता है। इससे तुलनात्मक अध्ययन में बहुत सहायता मिलती है लेकिन 5 या 6 से अधिक चरों को एक ही रेखा-चित्र पर नहीं दिखाना चाहिये अन्यथा चित्र को समझने में बहुत कठिनाई होगी क्योंकि एक रेखा दूसरी को कई जगह से काट सकती है जिससे चलों की प्रवत्ति (Behaviour) का ठीक रूप से पता लगाना कठिन होगा। जहां एक से अधिक चरों को रेखा-चित्र पर दिखाया जाता हैं वहां विभिन्न प्रकार की रेखाओं, रंगों, आदि का प्रयोग करना चाहिये ताकि चरों की प्रवत्ति पहचानने में सरलता रहे।

उदाहरण 3 : निम्न समंक 1985-86 से 1990-91 के भारत में चावल उत्पादन से संबंधित हैं :-

वर्ष	लक्ष्य (Target)	वास्तविक उत्पादन (Actal Production Million Tonnes)
1985-86	63.5	63.8
1986-87	65.0	60.6
1987-88	65.2	56.9
1988-89	68.0	70.5
1989-90	72.5	74.1
1990-91	73.7	75.0



हल : इन समंकों को उचित रूप से प्रदर्शित करने के लिये कृत्रिम आधार का प्रयोग उपयुक्त होगा। X-अक्ष पर वर्ष तथा Y-अक्ष पर दोनों चरों के मूल्य दिखाये गये हैं। निर्यात को बिन्दु रेखा (Dotted Line) तथा आयात को मोटी रेखायें (Thick Line) से दिखाया गया है।

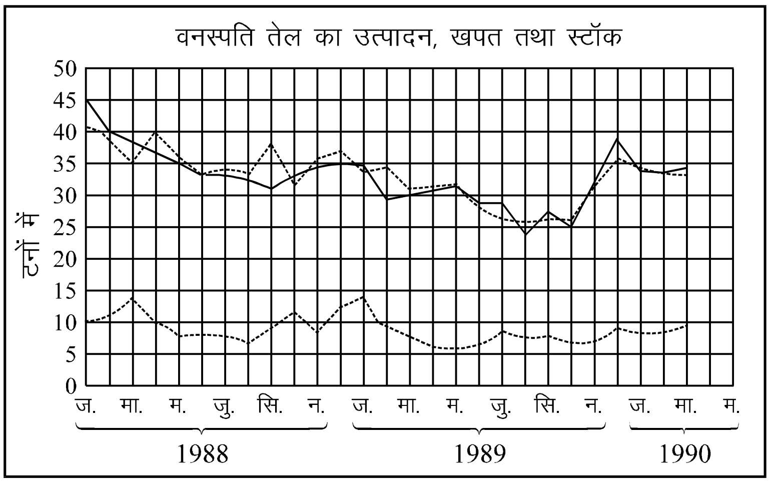
उदाहरण 4 : निम्न अंकों को रेखा-चित्रों द्वारा प्रस्तुत कीजिये।

वनस्पति तेल के पदार्थ (हजारों मीटर टनों में)

विवरण	जन.	फर.	मार्च	अ.	मई	जून	जुलाई	अ.	सि.	अक्तू.	न.	दि.
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
उत्पादन												
1988	43.8	39.8	37.8	36.0	36.0	35.4	32.6	32.8	31.6	33.4	33.8	41.1
1989	34.2	29.2	29.9	31.1	31.1	31.7	27.6	27.5	28.1	24.9	32.3	38.6
1990	33.7	33.1	34.3									

बचत												
1988	40.5	39.1	34.9	40.1	36.4	32.6	33.4	33.0	38.2	31.3	35.5	36.2
1989	33.7	33.7	30.8	32.0	32.3	27.7	24.4	25.5	25.4	26.5	32.0	35.8
1990	33.4	32.8	33.1									
स्टॉक												
1988	0.9	10.6	13.5	9.4	8.3	8.5	7.9	6.9	8.4	10.5	8.8	12.6
1989	13.6	9.2	8.3	3.6	5.7	6.0	7.7	7.1	7.7	6.1	6.4	9.2
1990	8.6	8.9	10.1									

हल : तीनों चलों को एक ही रेखा-चित्र पर इस प्रकार प्रदर्शित किया जायेगा :



कोटिबन्ध या विस्तार चार्ट

(Range Charts)

कभी-कभी किसी चल के किसी समय के अधिकतम व न्यूनतम उतार-चढ़ाव को अंकित करने की आवश्यकता पड़ती है जैसे किसी रोग को अधिकतम या न्यूनतम, किसी वस्तु का न्यूनतम व अधिकतम मूल्य आदि। ऐसी स्थिति में अधिकतम व न्यूनतम मूल्यों के वक्र अलग-अलग खींच लिये जाते हैं और उनके बीच के स्थान को किसी रंग या पेंसिल से भर देते हैं। ऐसे रेखा-चित्र बनाने की विधि इस प्रकार है :-

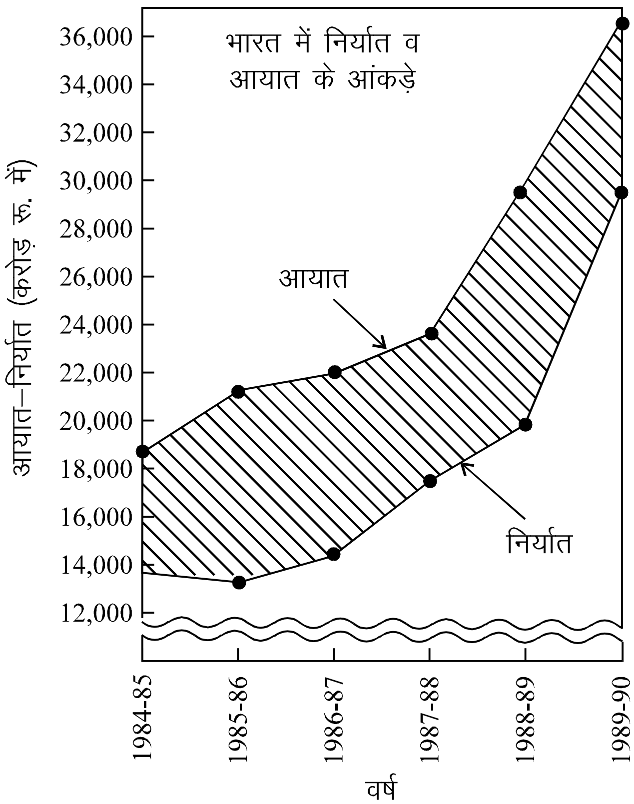
(1) समय को X-अक्ष तथा चल को Y-अक्ष पर लीजिये।

(2) दो वक्र, एक अधिकतम मूल्यों को तथा दूसरी न्यूनतम मूल्यों को प्रदर्शित करती हुई बनाइए।

(3) अधिकतम तथा न्यूनतम वक्रों के बीच की जगह को किसी रंग से भर दीजिए ताकि चित्र आकर्षक और समझने में सरल बन सके।

उदाहरण 5 : निम्नलिखित समंकों को उपयुक्त चित्र द्वारा प्रस्तुत कीजिये :-

वर्ष	आयात (करोड़ रू. में)	निर्यात (करोड़ रू. में)
1984-85	17,134	11,744
1985-86	19,658	10,895
1986-87	20,096	12,452
1987-88	22,244	15,674
1988-89	28,235	20,232
1989-90	35,412	27,681



हल : महीनों को X-अक्ष पर तथा सोने की कीमत Y-अक्ष पर लीजिये दो वक्र — एक अधिकतम तथा दूसरी न्यूनतम मूल्य का प्रदर्शित करती हुई खीचियें जैसा कि अगले उदाहरण में दिये गये चित्र द्वारा विदित होगा।

पट्टीदार वक्र या संघटक भाग-चित्र
(Band Curve or Component Chart)

जब काल श्रेणी पर आधारित सम्पूर्ण राशि अंशों में विभक्त हो तो पट्टीदार वक्र का प्रयोग किया जाता है। विभिन्न अंशों को एक के ऊपर एक करके चिन्हित कर दिया जाता है। बीच कें अंतर में रंग भर देते हैं जिससे एक पट्टीदार श्रेणी बन जाती है। इस प्रकार के चित्र सम्पूर्ण लागत को विभिन्न मदों में विभक्त करके प्रदर्शित करने में प्रयुक्त होते हैं जैसे कुल बिक्री विभाग अथवा जिले के अनुसार या कुल उत्पादन राज्यानुसार या उद्योगानुसार (Industry-wise), आदि।

उत्पादन दस लाख टनों में
(Production in Tonnes)

वर्ष (Year)	चावल (Rice)	गेहूँ (Wheat)	ज्वार (Jawar)	मक्का (Maize)	बाजरा (Bajra)	अन्य खाद्यान्न (Other-cereals)
1980-81	53.63	36.31	10.43	6.96	5.43	6.20
1981-82	53.25	37.45	12.06	6.90	5.54	6.59
1982-83	47.12	42.79	10.75	6.55	5.13	5.32
1983-84	60.10	45.48	11.92	7.92	5.72	6.34
1984-85	58.34	44.07	11.40	8.44	6.05	5.28
1985-86	62.82	47.05	10.20	6.64	3.66	5.70
1986-87	60.56	44.32	91.19	7.59	4.51	5.54
1987-88	56.43	45.10	11.85	5.83	3.28	5.08

हल : चावल, गेहूं आदि के आंकड़ों का योग ज्ञात कीजिये, जैसा निम्न सारणी में दिया गया है :-

वर्ष	चावल	गेहूं	ज्वार	मक्का	बाजरा	कुल योग
1980-81	53.63	89.94	100.37	107.33	112.76	118.97
1981-82	53.25	90.70	102.76	109.66	115.20	121.79
1982-83	47.12	89.91	100.66	107.21	112.34	117.66
1983-84	60.10	105.58	117.50	125.42	133.15	139.48
1984-85	58.34	102.41	113.81	122.25	128.30	133.58
1985-86	62.82	110.87	121.07	127.71	131.37	137.07
1986-87	60.56	104.88	114.07	121.66	126.17	131.71
1987-88	56.43	101.53	113.38	119.01	122.29	127.37

एक पट्टीदार वक्र उपर्युक्त समंकों को ठीक ढंग से प्रस्तुत करेगा। ऐसा वक्र अग्र प्रकार से बनाये जायेगा :

(1) वर्षों को X-अक्ष पर तथा उत्पादन को Y-अक्ष पर दिखायें।

(2) चावल से संबंधित समंकों को चिन्हित करके एक रेखा द्वारा मिलाइये जैसा चित्र में रेखा "A" द्वारा विदित है।

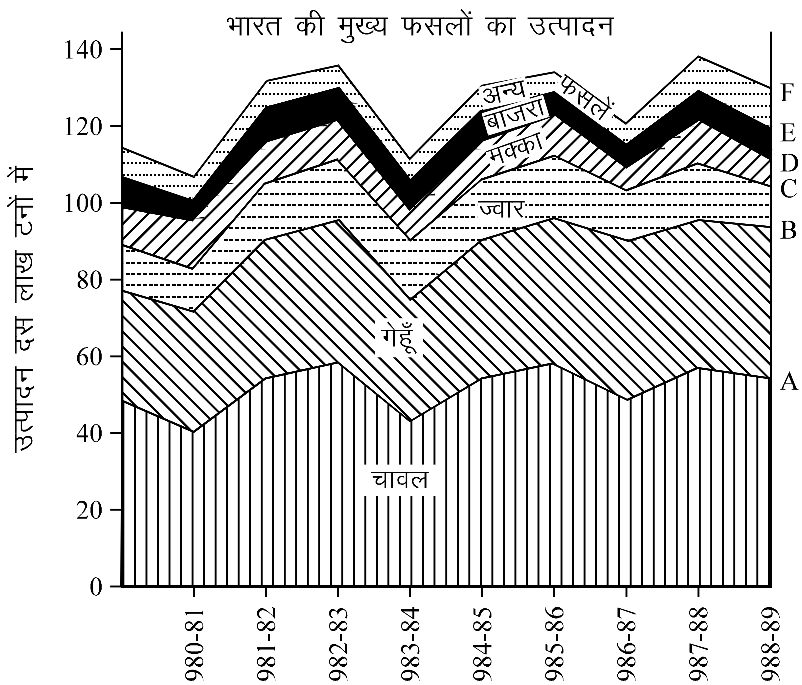
(3) चावल और गेहूं के समंकों का योग करके चिन्हित कीजिये जैसा "B" द्वारा विदित है।

(4) चावल, गेहूं तथा ज्वार के समंकों का योग करके उन्हें चिन्हित कीजिये तथा एक रेखा द्वारा मिलाइये — इस चित्र में देखिये रेखा "C"।

(5) चावल, गेहूं ज्वार तथा मक्का के समंकों को योग करके चिन्हित कीजिये और रेखा द्वारा मिलाइये — देखिये रेखा "D"।

(6) चावल, गेहूं, ज्वार तथा मक्का तथा अन्य खाद्यान्न के समंकों का योग करके चिन्हित कीजिये — देखिये रेखा "E"।

(7) चावल, गेहूं, ज्वार, मक्का तथा बाजरे के समंकों का योग करके चिन्हित कीजिये — देखिये रेखा "F"।



'जी' चित्र पर 'जी' वक्र
(Zee Chart or Zee Curve)

''जी'' चित्र एक बहु वक्र चित्र होता है, जो अंग्रेजी के अक्षर ''जेड'' (Z) से मिलता हुआ होता है। इसलिये इसे जी "Zee" या "Z" चित्र कहते हैं। इसमें तीन वक्र तीन बातों को प्रदर्शित करते हुयें खींचे जाते हैं जो इस प्रकार हैं :-

(1) मौलिक समंकों का वक्र (Curve of Original Data);

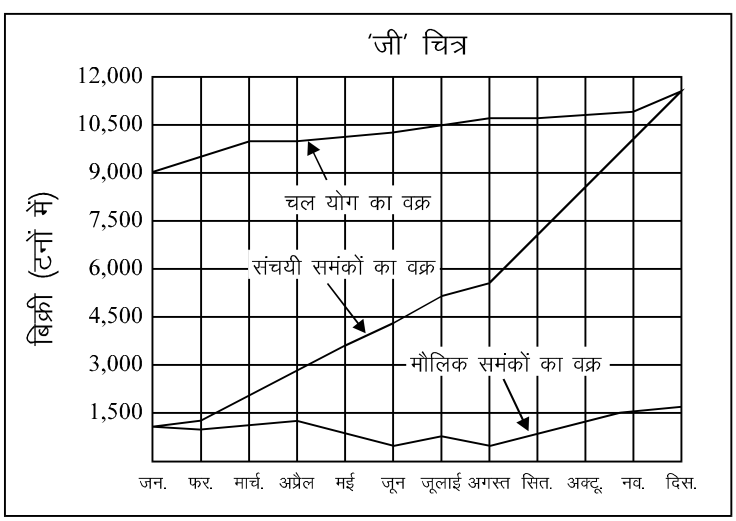
(2) संचयी समंकों का वक्र (Curve of Cumulative Data); तथा

(3) चल योगों का वक्र (Curve of Moving Totals).

मौलिक समंकों का वक्र सबसे नीचे दिखाया जाता है। चल योग वक्र सबसे ऊपर तथा संचयी वक्र मध्य में होता है। व्यापारिक समंकों को प्रस्तुत करने में ''जी'' चित्र का विशेष महत्व है। चल योग वक्र व्यापार की दशा बताता है — ऊपर उठती रेखा व्यापार से उन्नति का प्रदर्शन करती है। मौलिक समंकों का वक्र मौसमी परिवर्तन प्रदर्शित करता है तथा संचयी वक्र समय-समय की स्थिति से परिचित कराता है।

उदाहरण 7 : निम्न समंकों को ''जी'' चित्र द्वारा प्रस्तुत कीजिये :-

माह	मासिक बिक्री (टनों में)	मासिक संचयी बिक्री (टनों में)	वार्षिक चल योग
जनवरी	500	500	9,000
फरवरी	550	1,050	9,300
मार्च	800	1,850	9,800
अप्रैल	1,000	2,850	10,000
मई	850	3,700	10,400
जून	600	4,300	10,500
जुलाई	650	4,950	10,700
अगस्त	700	5,650	10,900
सितम्बर	1,150	6,800	11,000
अक्टूबर	1,400	8,200	11,300
नवम्बर	1,500	9,700	11,400
दिसम्बर	1,800	11,500	11,500



आनुपातिक रेखा चित्र
(Ratio Chart)

अब तक जो बिन्दु-रेखीय चित्र इस अध्याय में बनाये गये हैं उनमें प्राकृतिक माप-श्रेणी का प्रयोग किया गया है। इस प्रकार के चित्रों का प्रयोग वास्तविक अथवा निरपेक्ष अन्तरों को दिखाने या तुलना करने के लिये किया जाता है। जहां सापेक्ष (Relative) परिवर्तनों को चित्र द्वारा दिखाना हो वहां प्राकृतिक माप-श्रेणी के स्थान पर आनुपातिक माप-श्रेणी का प्रयोग किया जाता है। उदाहरणार्थ, यदि हम यह जानना चाहें कि जनसंख्या कीमतों, आदि में किसी अनुपात से सिद्ध हो रही हैं तो प्राकृतिक माप-श्रेणी के स्थान पर आनुपातिक माप-श्रेणी का प्रयोग करना पड़ेगा। निम्न उदाहरण द्वारा इन दोनों विधियों में अन्तर स्पष्ट हो जायेगा :-

वर्ष	उत्पादन (टनों में)	निरपेक्ष वद्धि	सापेक्ष वद्धि
1998	20	10	—
1989	30	10	50 प्रतिशत
1990	40	10	33½ प्रतिशत
1991	50	10	25 प्रतिशत
1992	60	10	20 प्रतिशत

उपर्युक्त उदाहरण से स्पष्ट है कि प्रतिवर्ष निरपेक्ष वद्धि समान है परन्तु आनुपातिक वद्धि की दर घट रही है अत : जहां आनुपातिक परिवर्तन दिखाना हो वहां प्राकृतिक माप श्रेणी का प्रयोग ठीक निष्कर्ष पर पहुंचने में बाधक होगा। इसलिये ऐसी परिस्थिति में आनुपातिक माप-श्रेणी का ही प्रयोग करना चाहिये।

प्राकृतिक माप-श्रेणी व आनुपातिक माप-श्रेणी में अन्तर

(Difference Between Natural Scale and Ratio Scale)

(1) प्राकृतिक माप-श्रेणी में निरपेक्ष (Absolute) परिवर्तनों को दिखाया जाता है जबकि आनुपातिक-श्रेणी में सापेक्ष परिवर्तनों को प्रदर्शित किया जाता है।

प्राकृतिक माप दण्ड	अनुपाति माप दण्ड			
50	32	320	3,200	32,000
40	16	160	1,600	16,000
30	8	80	800	8,000
20	4	40	400	4,000
10	2	20	200	2,000
0	1	10	100	1,000

(2) प्राकृतिक माप-श्रेणी अंकगणितीय वद्धि (Arithmetic Progression) पर आधारित है जबकि आनुपातिक माप-श्रेणी गुणोत्तर वद्धि (Geometric Progression) के आधार पर होती है। निम्नांकित चित्र द्वारा यह बात स्पष्ट हो जायेगी।

(3) प्राकृतिक माप-श्रेणी में उदग्र मापदण्ड शून्य से आरम्भ होता है और आवश्यकतानुसार कृत्रिम आधार का प्रयोग किया जाता है। आनुपातिक माप-श्रेणी में उदग्र मापदण्ड शून्य से आरम्भ नहीं होता इसलिये इस प्रकार के चित्रों में कृत्रिम आधार रेखा की आवश्यकता नहीं होती।

(4) प्राकृतिक माप-श्रेणी में ऋणात्मक मूल्यों का प्रदर्शन किया जा सकता है परन्तु आनुपातिक माप-श्रेणी में ऋणात्मक-मूल्यों का प्रदर्शन नहीं हो सकता।

(5) प्राकृतिक माप-श्रेणी में कोटि-अक्ष पर निरपेक्ष मान प्रदर्शित किये जाते हैं। आनुपातिक माप-श्रेणी में या तो मानों के लघुगणक (Logarithms) अंकित किये जाते हैं या लघुगणकीय बिन्दुरेखीय पत्र (Semi-logarithmic Paper) का प्रयोग किया जाता है।

(6) जहां चर के विभिन्न मूल्यों में अधिक अन्तर हो वहाँ आनुपातिक माप-चित्र का प्रयोग उचित है।

आनुपातिक माप-श्रेणी तथा छेदा वक्र
(Ratio Scale and Logarithmic Curves)

निम्न दो रीतियों द्वारा आनुपातिक माप-श्रेणी पर बिन्दु रेखा की रचना की जा सकती है।

(1) दी हुई श्रेणी के लघुगणकों को साधारण बिन्दुरेखीय पत्र पर चिन्हित करके; तथा

(2) मूल्यों को लघुगणकीय पत्र पर चिन्हित करके। यह एक विशेष प्रकार का बिन्दुरेखीय पत्र होता है जो अनुपात के आधार पर निर्मित होता है।

पहली रीति कें अन्तर्गत दिये हुये मूल्यों के लघुगणक, लघुगणक सारणी (Logarithmic Table) द्वारा ज्ञात कर लिये जाते हैं। X-अक्ष पर समय लिया जाता है और Y-अक्ष पर विभिन्न मूल्यों के लघुगणक। विभिन्न बिन्दुओं को रेखाओं द्वारा मिला दिया जाता है। दूसरी रीति के अंतर्गत चर मूल्यों का लघुगणक ज्ञात नहीं किया जाता अपितु वास्तविक मानों को लघुगणकीय बिन्दु रेखा-चित्र पर चिन्हित कर लेते हैं। ये पत्र बाजार में सरलता से मिल जाते हैं। पहली रीति की तुलना में यह रीति कहीं अधिक सरल व लोकप्रिय है।

इन दोनों रीतियों का प्रयोग निम्न उदाहरण से स्पष्ट हो जाएगा :-

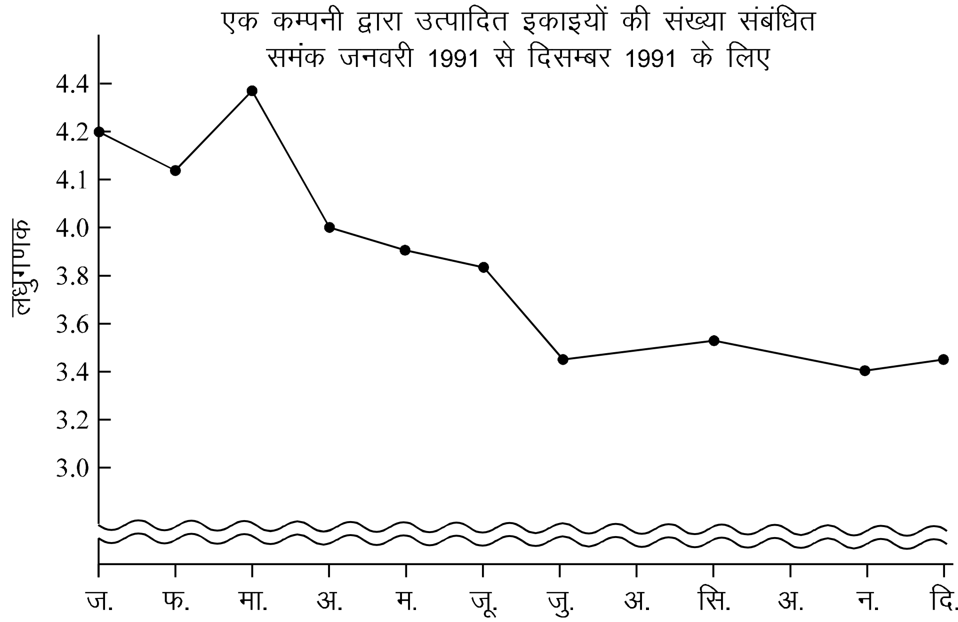
उदाहरण 8 : निम्न समंक एक कम्पनी द्वारा प्रति माह उत्पादन से संबंधित हैं :-

माह	इकाइयों की संख्या	माह	इकाइयों की संख्या
जनवरी 1991	15,560	जुलाई 1991	3,250
फरवरी 1991	13,170	अगस्त 1991	3,570
मार्च 1991	18,740	सितम्बर 1991	3,620
अप्रैल 1991	12,450	अक्टूबर 1991	3,140
मई 1991	8,320	नवम्बर 1991	2,580
जून 1991	5,540	दिसम्बर 1991	2,540

उपर्युक्त समंकों को एक बिन्दुरेखीय पत्र पर लघुगणकीय माप-श्रेणी द्वारा प्रदर्शित कीजिये।

हल : दिये हुये समंकों का लघुगणक लेकर प्राकृतिक पैमाने पर अंकित कीजिए :

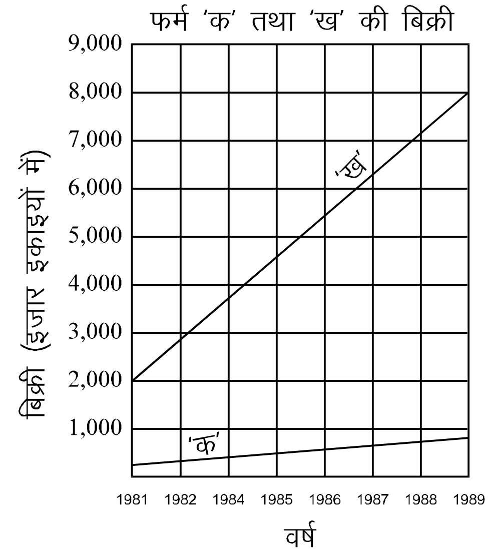
माह	इकाइयों की संख्या	लघुगणक	माह	इकाइयों की संख्या	लघुगणक
जनवरी 1991	15,560	4.1920	जुलाई 1991	3,250	3.5119
फरवरी 1991	13,170	4.1200	अगस्त 1991	3,570	2.5527
मार्च 1991	18,740	0.2728	सितम्बर 1991	3,630	3.5587
अप्रैल 1991	12,450	4.0952	अक्टूबर 1991	3,140	3,4987
मई 1991	8,320	3.9201	नवम्बर 1991	2,580	2.4116
जून 1991	5,540	3.8774	दिसम्बर 1991	2,540	3.4048



उदाहरण 9 : अग्रांकित सारणी फर्म 'क' और फर्म 'ख' की 1982 से 1989 तक की बिक्री के समंक प्रस्तुत करती है। इन समंकों को एक बिन्दुरेखीय चित्र द्वारा प्रस्तुत कीजिये :-

वर्ष	बिक्री फर्म 'क' (हजार इकाइयों में)	बिक्री फर्म 'ख' (हजार इकाइयों में)
1982	200	2,000
1983	300	3,000
1984	400	4,000
1985	500	5,000
1986	600	6,000
1987	700	7,000
1988	800	8,000
1989	900	9,000

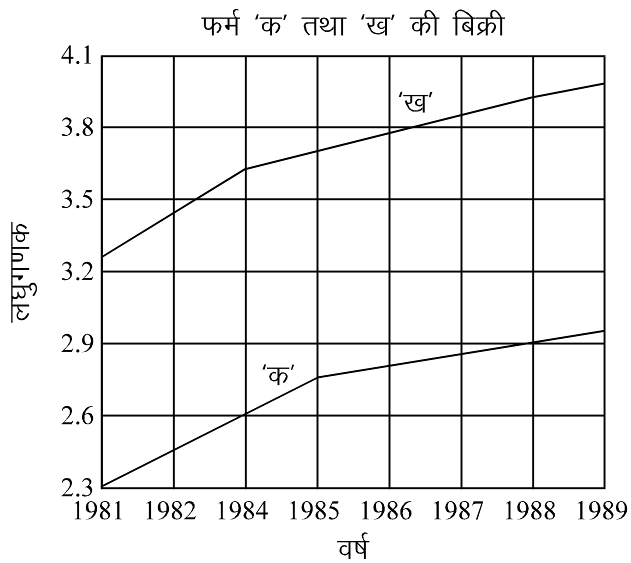
हल : दिये गये समंकों को प्राकृतिक तथा आनुपातिक पैमानों पर अंकित किया गया है। पहला चित्र प्राकृतिक पैमाने पर है, दूसरा आनुपातिक पर।



आनुपातिक पैमाने पर इन समंकों को प्रस्तुत करने के लिये इनका लघुगणक ज्ञात करेगें जैसा कि निम्न सारणी से विदित होगा :

वर्ष (हजार इकाइयों में)	बिक्री फर्म 'क' (हजार इकाइयों में)	लघुगणक	बिक्री फर्म 'ख'	लघुगणक
1982	200	2.3010	2,000	3.3010
1983	300	2.4771	3,000	3.4771
1984	400	2.6021	4,000	3.6021
1985	500	2.6990	5,000	3.6990
1986	600	2.7782	6,000	3.7782
1987	700	2.8451	7,000	3.8451
1988	800	2.9031	8,000	3.9031
1989	900	2.9542	9,000	3.9542

इन समंकों को निम्न चित्र में अंकित किया हैं :



इन दोनों रेखाचित्रों की तुलना से विदित होगा कि समंक प्राकृतिक पैमाने पर पस्तुत किये गये तो फर्म 'क' की तुलना में फर्म 'ख' की बिक्री में कहीं अधिक वद्धि प्रतीत हुई ; जबकि आनुपातिक माप श्रेणी के आधार पर फर्म 'क' और फर्म 'ख' दोनों की वद्धि दर एक ही है। वास्तव में फर्म 'क' की बिक्री प्रतिवर्ष 100 इकाई से बढ़ रही और फर्म 'ख' की बिक्री 1,000 इकाई से। इस प्रकार यह कहा जा सका है कि फर्म 'ख' फर्म 'क' से 10 गुनी अच्छी है परन्तु आनुपातिक रेखाचित्र से पता चलता है कि दोनों की वद्धि दर समान है।

लघुगणकीय वक्रों का निर्वचन (Interpretation of Logarithmic Curve) : लघुगणकीय वक्रों का निर्वचन सावधानी से करना चाहिये, अन्यथा भ्रामक परिणामों पर पहुंचने की आशंका रहती है। लघुगणकीय वक्रों को पढ़ते समय निम्न बातों पर ध्यान देना आवश्यक है :-

(1) यदि वक्र ऊपर की ओर उठ रही है तो परिवर्तन की दर में वद्धि का संकेत करती है।

(2) यदि वक्र नीचे की ओर झुक रही है तो परिवर्तन की दर में कमी करती है।

(3) यदि एक सरल रेखा के समान है तो परिवर्तन एक ही दर से हो रहा है।

(4) यदि वक्र ऊपर की ओर बढ़ रही है लेकिन लगभग सरल ही है तो परिवर्तन की दरे में लगभग समान गति में वद्धि का पता लगता है।

(5) यदि वक्र नीचे की ओर झुक रही है लेकिन लगभग सरल है तो परिवर्तन की दर लगभग समान गति से घट रही है।

(6) यदि वक्र एक स्थान से दूसरे स्थान की तुलना में कहीं अधिक झुका हुआ है तो पहले स्थान पर परिवर्तन की दर एक-दूसरे की तुलना में अधिक होती है।

(7) यदि एक ही रेखाचित्र पर दो वक्र एक-दूसरे के समान्तर (Parallel) हैं तो दोनों में प्रतिशत परिवर्तन समान होता है।

(8) यदि एक ही रेखाचित्र पर एक वक्र दूसरे की तुलना में अधिक झुकी हुई है तो पहले वक्र में परिवर्तन की दर दूसरे की तुलना में अधिक होती है।

लघुगणकीय माप श्रेणी के विशिष्ट प्रयोग (Specific Uses of Semi-log Graphs) : इस प्रकार के रेखाचित्र निम्न परिस्थितियों में अधिक लाभदायक होते हैं।

(1) जनसंख्या, उत्पादन, बिक्री आदि चरों की विकास की गति (Rate of Growth) प्रदर्शित करने के लिये।

(2) विभिन्न इकाइयों में प्रस्तुत समंकों की तुलना करने में।

(3) एक ही श्रेणी में विभिन्न समयों पर होने वाले परिवर्तनों की तुलना करने के लिये।

लघुगणकीय रेखाचित्रों की सीमायें (Limitations of Ratio Charts) : यद्यपि लघुगणकीय चित्रों का बहुत महत्व है तथापि व्यापार व अर्थशास्त्र के क्षेत्र में उनका प्रयोग लगभग ''नहीं'' के बराबर है। इनकी कुछ सीमायें निम्नलिखित हैं :-

(1) शून्य अथवा ऋणात्मक समंकों को आनुपातिक माप श्रेणियों के द्वारा प्रस्तुत नहीं किया जा सकता।

(2) उनको समझने तथा ठीक निष्कर्ष निकालने के लिये बहुत कुशलता की आवश्यकता है, अन्यथा निष्कर्षों के अशुद्ध होने की सम्भावना बहुत अधिक होती है।

(3) आनुपातिक माप श्रेणी द्वारा निरपेक्ष परिवर्तनों का माप सम्भव नहीं होता।

(4) एक समूह (Aggregate) को अंशों में विभाजित करके उसका अध्ययन करना सम्भव नहीं।

जहां प्रचलित रीति के द्वारा समस्या का समाधान न होता हो वहीं पर इस नीति का प्रयोग करना चाहिये।

आवत्ति वितरणों के रेखाचित्र
(Graphs of Frequency Distribution)

एक अखण्डित श्रेणी को निम्नलिखित रेखाचित्रों के रूप में प्रस्तुत किया जा सकता है :-

(1) आवत्ति आयत चित्र (Histogram);

(2) आवत्ति बहुभुत (Frequency Polygon);

(3) सरलित-आवत्ति वक्र (Smoothed Frequency Curve);

(4) संचयी-आवत्ति वक्र (Ogive or Cumulative Frequency Curve)।

(1) आवत्ति -आयत चित्र (Histograms) : आवत्ति -आयत आवत्ति वितरण को प्रस्तुत करने की सबसे उत्तम व प्रचलित रीति है। आवत्ति -आयत चित्र के अंतर्गत आवत्ति वितरण की वर्ग आवत्तियों को खड़े हुये आयातों की पंक्ति के रूप में प्रस्तुत करते हैं।

आवत्ति आय चित्र बनाते समय वर्गान्तर X-अक्ष पर तथा उप पर आश्रित आवत्तियां Y-अक्ष पर ली जाती हैं। यदि वर्गान्तर समान है तो आयतों की चौड़ाई में अन्तर होगा। आयत की ऊँचाई वर्ग आवत्ति पर आधारित होगी। इस प्रकार हमारे पास आयतों की एक श्रंखला सी बन जोयगी। आवत्ति आयत चित्र का क्षेत्रफल कुल आवत्ति प्रदर्शित करेगा।

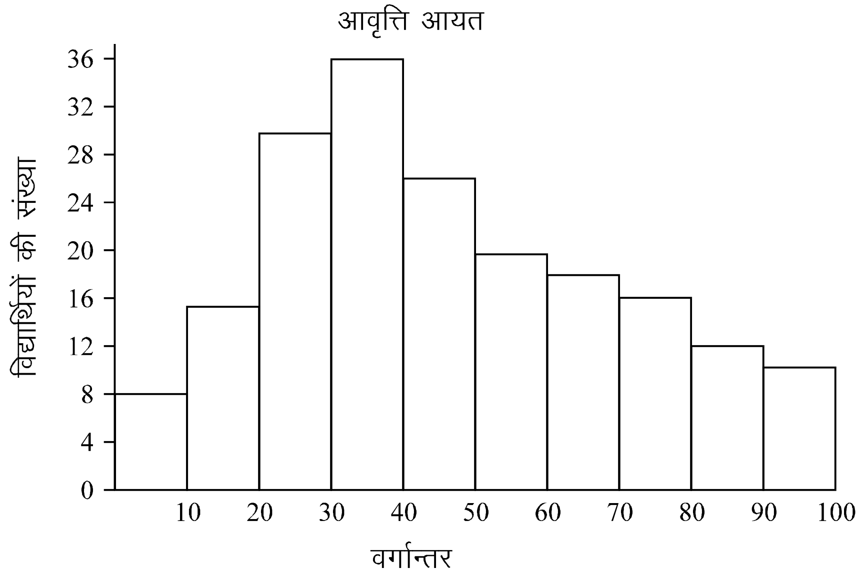
आवत्ति -आयत चित्र तथा दण्ड में अन्तर भली-भांति समझ लेना चाहिये। जबकि दण्ड-चित्र में दण्ड की लम्बाई का ही महत्व होता है, आवत्ति -आयत चित्र में लम्बाई तथा चौड़ाई दोनों का महत्व होता है। अर्थात् दण्ड-चित्र एक-विमा चित्र हैं और आवत्ति -आयत चित्र द्वि-विमा चित्र है।

आवत्ति -आयत चित्र बनाने की विधि का विवरण दो भागों में किया गया है। (1) जब आवत्ति वितरण के वर्गान्तर समान हों, तथा (2) आवत्ति वितरण के वर्गान्तर भिन्न हों।

आवत्ति -आयत चित्र की रचना जब वर्गान्तर समान हों (Histogram When Class Intervals are Equal) : जब वर्गान्तर समान हों तो X-अक्ष (भुजाक्ष) पर चर तथा Y-अक्ष (कोटि-अक्ष) पर आवत्ति लेकर एक से एक मिलते हुये आयत बनाये जाते हैं और उनमें कोई परिवर्तन करने की आवश्यकता नहीं होती जैसा कि निम्न उदाहरण से स्पष्ट हो जायेगा :-

उदाहरण 10 : निम्न समंकों से एक आवत्ति -आयत चित्र बनाइये :-

वर्गान्तर	विद्यार्थियों की संख्या	वर्गान्तर	विद्यार्थियों की संख्या
0-10	5	50-60	20
10-20	15	60-70	18
20-30	30	70-80	16
30-40	36	80-90	12
40-50	26	90-100	10



हल : X-अक्ष पर वर्गान्तर लीजिये और Y-अक्ष पर आवत्ति। वर्गान्तर की आवत्ति के अनुसार आयत बनाइये जो एक -दूसरे से मिलते हुये हों।

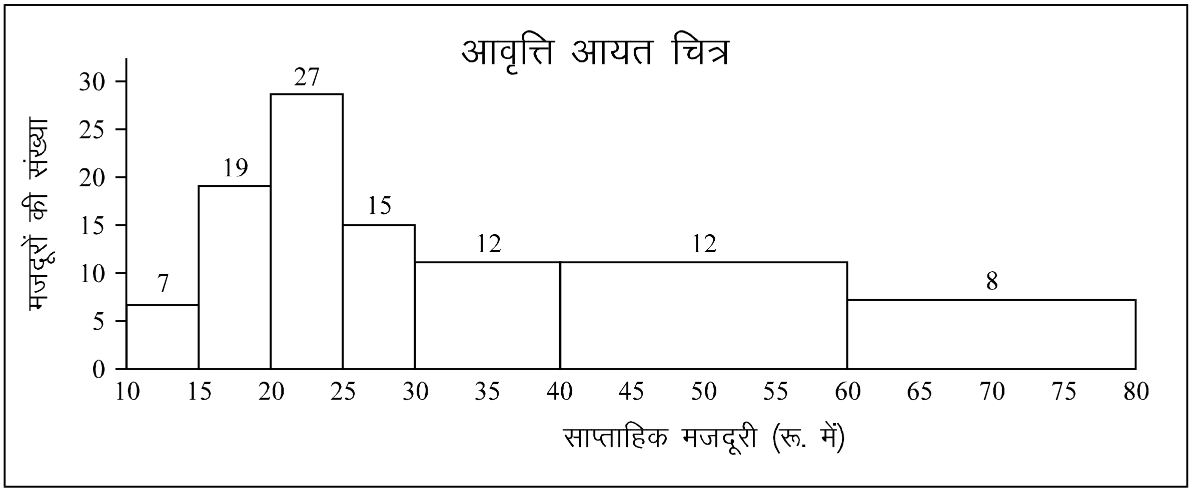
आवत्ति आयात चित्र की रचना जब वर्गान्तर भिन्न हों (Construction of Histogram When Class-intervals are Equal) : जब वर्गान्तर भिन्न हों आवत्ति आयत चित्र बनाने से पहले आवत्तियों में समायोजन (Adjustment) करना आवश्यक हो जाता है। सबसे कम वर्गान्तर वाले वर्ग को लेकर अन्य वर्गान्तरों की आवत्ति को उसी अनुपात में समायोजिक कर लिया जाता है। उदाहरण के लिये, एक वर्ग का अन्तर दूसरे की तुलना में दुगुना या तिगुना है तो उसके आयात की ऊंचाई क्रमश: आधी या तिहाई कर देंगे अर्थात् प्रत्येक आयत की ऊंचाई उसके वर्गान्तर के अनुपात में होती है। निम्न उदाहरण से यह बात स्पष्ट हो जायेगी।

उदाहरण 11 : निम्न समंकों को आवत्ति -आयत-चित्र द्वारा प्रस्तुत कीजिये :-

सात्पाहितक मजदूरी (रुपयों में)	मजदूरों की संख्या	साप्ताहिक मजदूरी (रुपयों में)	मजदूरों की संख्या
10-15	7	30-40	12
15-20	19	40-50	12
20-25	27	50-60	8
25-30	15		

हल : वर्गान्तर समान नहीं हैं, इसलिये आवत्तियों में समायोजन करना आवश्यक है। न्यूनतम वर्गान्तर 5 है और जहां वर्गान्तर 5 से अधिक है वहां आवत्तियों को इसी अनुपात में कम कर दिया गया है। उदाहरणार्थ पांचवी श्रेणी 30-40 है और इसका वर्गान्तर 10 है। क्योंकि वर्गान्तर दुगना है इसलिये आवत्ति को आधा कर दिया जायेगा। अर्थात् 30-40 की श्रेणी पर आयत की ऊंचाई 6 होगी न कि 12।

आवत्ति -आयत-चित्र की रचना जब केवल मध्य-बिन्दु दिये हों (Construction Histogram When Only Mid-points are Given) : जब मध्य-बिन्दु और आवत्ति दी हों तो आयत चित्र को बनाने के लिये प्रत्येक श्रेणी की निम्न सीमा तथा उच्च सीमा निकाल दी जाती है, फिर ऊपर दी गई विधि के अनुसार आवत्ति आयत बनाया जाता है। जैसा निम्न उदाहरण से स्पष्ट होगा :-

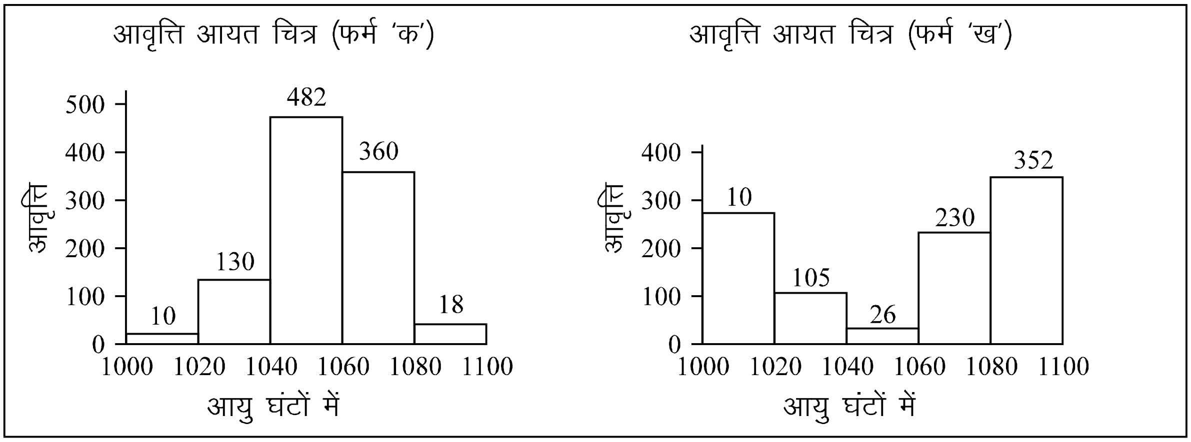


उदाहरण 12 : निम्न समंकों को आवत्ति -आयत चित्र द्वारा प्रस्तुत कीजिये :-

बिजली के बल्बों के जलने की क्षमता (घंटों में)	फर्म 'क'	फर्म 'ख'
1,010	10	287
1,030	30	105
1,050	82	26
1,070	60	230
1,090	18	352

हल : दिये गये मध्य बिन्दुओं से प्रत्येक श्रेणी की उच्च सीमा तथा निम्न सीमा निकालिये जो इस प्रकार होगी :-

बिजली के बल्बों के जलने की क्षमता (घंटों में)	आवत्ति फर्म 'क'	आवत्ति फर्म 'ख'
1,000-1,020	10	287
1,020-1,040	130	105
1,040-1,060	482	26
1,060-1,080	360	230
1,080-1,100	18	352



(2) **आवत्ति -बहुभुज (Frequency Polygon) :** आवत्ति वितरण (Frequency Distribution) को आवत्ति -बहुभुज द्वारा भी प्रस्तुत किया जा सकता है। आवत्ति -बहुभुज की चार से अधिक भुजायें होती हैं। इसको बनाने की दो विधियाँ हैं :-

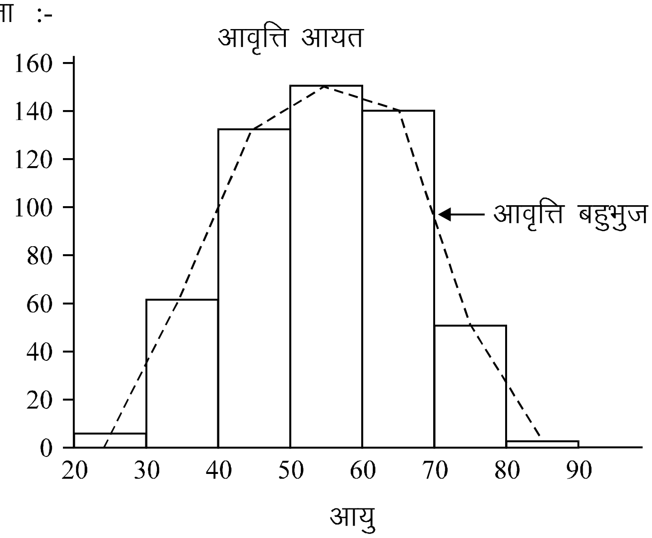
(*i*) दिये गये समंकों से एक आवत्ति -आयत-चित्र बनाइये। प्रत्येक आयत की ऊपर की भुजा के मध्य बिन्दु को साथ वाले आयतों के मध्य बिन्दुओं से सरल रेखा द्वारा मिलाइये। इस प्रकार आवत्ति -बहुभुज बन जायेगा। अब इसके दोनों किनारों को आधार तक बढ़ाइये। इसके लिये दो कल्पित वर्ग (Hypothetical Classes) लेने पड़ेगें - एक बहुभुज के बायीं ओर तथा दूसरा बायीं ओर। आवत्ति बहुभुज के दोनों किसानों को जब आधार से मिला दिया जाता है तो आवत्ति आयत तथा बहुभुज दोनों के क्षेत्रफल बराबर हो जायेंगे।

(*ii*) दूसरी विधि के अनुसार विभिन्न वर्गान्तरों के मध्य बिन्दु लेकर बिन्दु रेखीय पत्र पर चिन्हित कर दिये जाते हैं तथा इन बिन्दुओं को सरल रेखाओं द्वारा मिला दिया जाता है। दोनों विधियों से एक ही प्रकार की आकृति बनेगी।

उदाहरण 13 : निम्न समंकों की सहायत से एक आवत्ति बहुभुज बनाइए :-

आयु	सदस्यों की संख्या	आयु	सदस्यों की संख्या
20-30	8	60-70	140
30-40	61	70-80	51
40-50	132	70-80	2
40-60	153		

हल : आवत्ति बहुभुज की रचना :-



दोनों चित्रों से विदित होता है कि आवत्ति एक ही है चाहे हम आवत्ति बहुभुज बनायें या मध्य बिन्दु लेकर ; लेकिन आवत्ति आयत से बहुभुज बनाना अधिक उपयोगी होता है :-

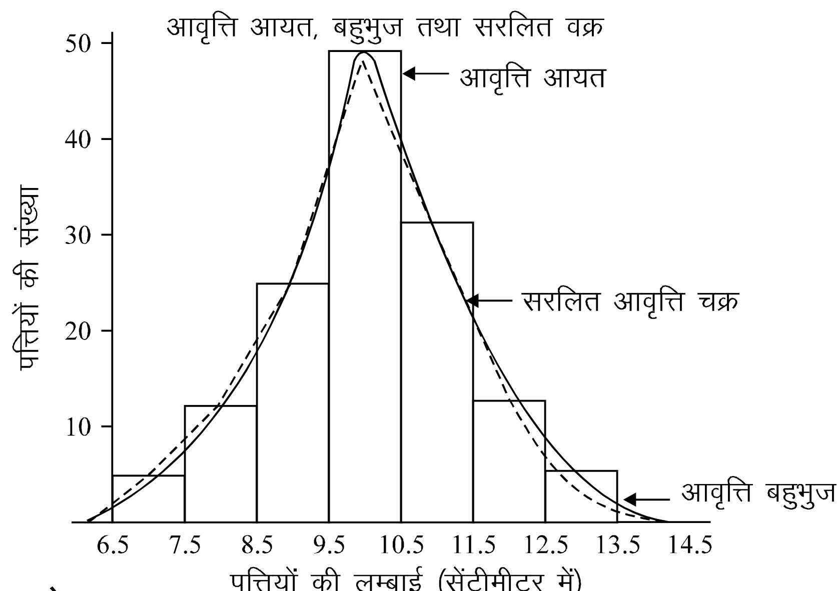
आवत्ति आयत और आवत्ति बहुभुज में अन्तर (Histogram and Polygo Distinguished) : आवत्ति बहुभुज का एक विशेष लाभ यह है कि कई आवत्ति वितरणों की तुलना भी चित्र द्वारा जा सकती है लेकिन आवत्ति से ऐसा संभव नहीं। यही कारण है कि आवत्ति वितरणों की तुलना के लिये आवत्ति बहुभुज का प्रयोग किया जाता है।

(3) **सरलित आवत्ति वक्र (Smoothed Frequency Curve) :** सरलित आवत्ति वक्र बनाने से पहले आवत्ति बहुभुज बनाना आवश्यक है। आवत्ति वक्र को बनाने का उद्देश्य यथासंभव समंकों के अकस्मात् परिवर्तनों को कम करना है। इस वक्र में यह प्रयत्न किया जाता है कि आवत्ति बहुभुज के कोण समाप्त होकर वह एक सरलित वक्र (Smoothed Curve) बन जाये। वक्र का सरलीकरण समंकों की प्रकृति अथवा उनके स्वरूप पर निर्भर करेगा। यदि समंक प्राकृतिक विषयों से संबंधित हैं जैसे सिक्के का उछालना, तो उस दशा में काफी सीमा तक सरलीकरण किया जा सकता है। यदि समंक सामाजिक या आर्थिक विषयों से संबंधित हैं तो सरलीकरण अधिक नहीं किया जा सकता। निम्न उदाहरण सरलित आवत्ति की रचना स्पष्ट हो जायेगी।

उदाहरण 14 : निम्न समंकों से एक सरलित आवत्ति वक्र बनाइये :-

पत्तियों की लम्बाई (सें.मी. में)	पत्तियों की संख्या
6.5-7.5	5
7.5-8.5	12
8.5-9.5	25
9.5-10.5	48
10.5-11.5	32
11.5-12.5	6
12.5-13.5	1

हल : दिये हुये समंकों का पहले एक आवत्ति आयत और फिर आवत्ति बहुभुज बनाये। इसके पश्चात आवत्ति बहुभुज का सरलीकरण कीजिये जैसा कि निम्न चित्र से स्पष्ट है :

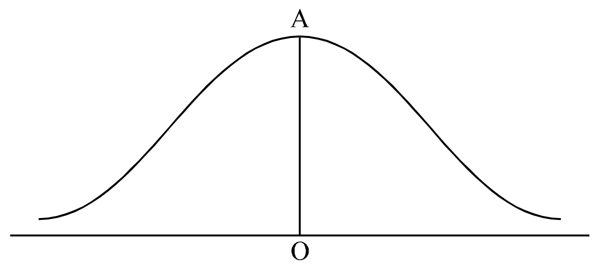


प्रवत्ति वितरण वक्र के आकार

(Frequency Distributions and Shapes of Frequency Curves)

सामान्यत: आवत्ति वितरण (Frequency Distribution) पांच प्रकार के होते हैं जिसके अनुसार आवत्ति वक्र के पांच प्रमुख आकार हैं उनके वर्णन नीचे किया गया है :-

(*i*) **सम्मित वितरण वाले वक्र (Curves of Symmetrical Distributions) :** एक सम्मित वितरण वक्र पूर्णरूप से घंटाकार (Bell-Shaped) होता है। इसमें आवत्तियां धीरे -धीरे बढ़ती हैं और अधिकतम आवत्ति पर पहुंचकर इसी क्रम व गति से घटती है। इस प्रकार के वक्र को सामान्य वक्र (Normal Curve) कहा जाता है। यदि अधिकतम आवत्ति के बिन्दु के आधार पर एक लम्ब (Perpendicular) डाला जाये तो यह वक्र के दो समान भागों में विभाजित कर देता है। यह वक्र निम्न प्रकार का होता है :-



सामान्य वक्र

ऐसे वक्र व्यवहार में बहुत कम पाये जाते हैं। लेकिन सांख्यिकी में इनका महत्व है क्योंकि न्यादर्श संबंधी समस्याओं को समझने के लिये इनका ज्ञान अत्यन्त आवश्यक है।

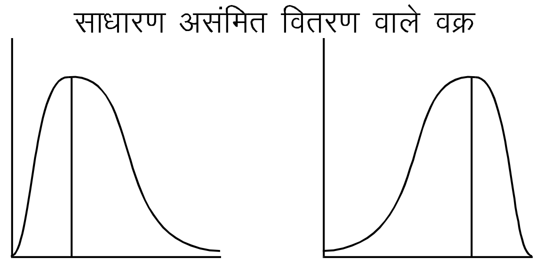
(*ii*) **साधारण असम्मित वितरण वाले वक्र (Curves of Moderately Asymmetrical Distribution) :** इस प्रकार के वितरणों में वक्र-आवत्तियों के बढ़ने व घटने का क्रम भिन्न होता है। ऐसे वक्रों को विषम वक्र (Skewed Curves) कहते हैं। इनमें वक्र का सिरा एक दूसरे से भिन्न प्रकार का अर्थात बहुत लम्बा या छोटा होता है। ऐसे वक्र दो प्रकार के होते हैं :-

(क) धनात्मक विषम वक्र (Positively Skewed Curves), तथा

(ख) ऋणात्मक विषम वक्र (Negatively Skewed Curves)।

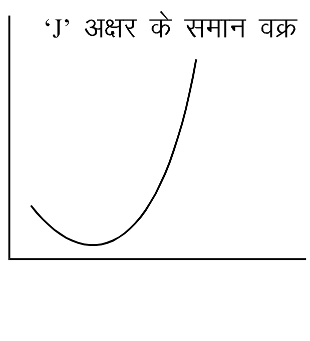
यदि वक्र का लम्बा सिरा दायीं ओर है तो वक्र धनात्मक विषय वक्र होगा, यदि बायीं ओर है तो ऋणात्मक विषम वक्र होगा। सम्मित वितरण वाले वक्रों की अपेक्षा इस प्रकार के वक्र व्यवहार से बहुत अधिक पाये जाते हैं। नीचे दो चित्र दिये गये हैं — पहला चित्र धनात्मक विषय वक्र तथा दूसरा ऋणात्मक विषम वक्र को प्रदर्शित करता है :-

साधारण असंमित वितरण वाले वक्र

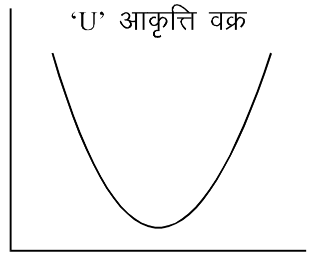

(अ) धनात्मक विषम वक्र (अ) ऋणात्मक विषम वक्र

(*iii*) **'J' अक्षर के समान वक्र (J-shaped Curves) :** इस प्रकार के वक्र अंग्रेजी के अनुसार 'J' के आकार के होते हैं। ऐसे वक्रों में अधिकतम आवत्ति प्रारम्भ में या अन्त में होती है। यदि अधिकतम आवत्ति श्रेणी के अन्त में होती है तो वक्र निम्न प्रकार का बनता है:-

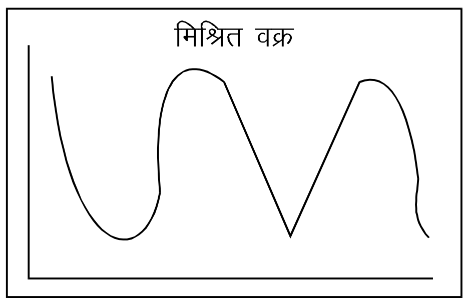
'J' अक्षर के समान वक्र



(*iv*) **'U' की आवत्ति वाले वक्र (U-shaped Curves) :** कुछ वक्र ऐसे होते है जो अंग्रेजी के अक्षर 'U' के आकार का वक्र तब बनता है जब अधिकतम आवत्तियां श्रेणी के प्रारम्भ में या अन्त में होती है , मध्य में आवत्तियां बहुत कम होती है। जैसा कि निम्न चित्र से स्पष्ट होगा।



(*v*) **मिश्रित वक्र (Mixed Curves) :** कुछ ऐसे वक्र भी होते हैं जो ऊपर दिये दो य दो से अधिक वक्रों का समिश्रण जैसे प्रतीक होते हैं, जैसे निम्न चित्र से स्पष्ट होगा।



(4) संचयी आवत्ति वक्र (Cumulative Frequency Curve or 'Ogive') : कभी-कभी ऐसे प्रश्नों का उत्तर जानने की आवश्यकता होती है, जैसे कितने मजदूर 1000 रूपये मासिक से कम वेतन पाते हैं अथवा कितने मजदूर 2000 रुपये मासिक से अधिक वेतन पाते हैं। इस प्रकार के प्रश्नों का उत्तर देने के लिये आवत्तियों को जोड़ना पड़ता है। जब आवत्तियों को जोड़ा जाता है तो उन्हें संचयी आवत्ति (Cumulative Frequency) कहते हैं। इन संचित आवत्तियों को एक सारणी द्वारा प्रस्तुत करते हैं जिसे आवत्ति सारणी कहते हैं। इस प्रकार की सारणी से बने वक्र को संचयी आवत्ति वक्र या ओगिव वक्र भी कहते हैं।

संचयी आवत्ति वक्र दो रीतियों से बनाया जा सकता है :-

(*i*) ऊपरी सीमाओं और बढ़ती हुई संचयी आवत्तियों के आधार पर ; तथा

(*ii*) निचली सीमाओं और घटती हुई संचयी आवत्तियों के आधार पर।

पहली रीति के अनुसार हम ऊपरी सीमाओं से आरम्भ करते हैं, तथा आवत्तियों को जोड़ते जाते हैं। जब इन आवत्तियों को रेखाचित्र पर अंकित किया जाता है तो एक बढ़ता हुआ वक्र (Rising Curve) बन जाता है। दूसरी रीति के अनुसार हम निचली सीमाओं से आरम्भ करते हैं तथा आवत्तियों को घटाते जाते हैं। यदि इन आवत्तियों को रेखाचित्र पर परस्तुत करें तो घटता हुआ वक्र (Decreasing Curve) बन जायेगा।

निम्न आवत्ति वितरण को पहले ऊपरी सीमाओं से आरम्भ करके तथा आवत्तियों को जोड़कर संचयी आवत्ति वितरण बनायेंगे। बाद में निचली सीमाओं से आरम्भ करके आवत्तियों को घटा कर संचयी आवत्ति वितरण बनायेंगे :

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10-20	4	40-50	20
20-30	6	50-60	18
30-40	10	60-70	2

संचयी आवत्ति सारणी

(Cumulative Frequency Table)

अंक इतने से कम	विद्यार्थियों की संख्या	अंक इतने से कम	विद्यार्थियों की संख्या
20	4	10	60
30	10	20	56
40	20	30	50
50	40	40	40
60	58	50	20
70	60	60	2
		70	0

उपर्युक्त सारणी से यह सुगमता से ज्ञात किया जा सकता है कि कितने विद्यार्थियों के अंक किसी विशेष संख्या से कम या अधिक हैं। उदाहरणार्थ, 20 विद्यार्थियों ने 40 कम अंक प्राप्त किये हैं तथा 50 विद्यार्थियों ने 30 से अधिक अंक प्राप्त किये हैं।

संचयी आवत्ति वक्रों की उपयोगिता

(Utility of Ogives)

संचयी आवत्ति वक्रों का बहुत अधिक महत्व है। सारणी को देखकर तुरन्त बताया जा सकता है कि कितने व्यक्ति ऐसे हैं जिनका वेतन, आयु आदि किसी विशेष संस्था से कम या अधिक है। ऊपर की सारणी से स्पष्ट होगा कि 40 विद्यार्थी ऐसे है जिनके अंक 50 से कम है तथा 20 विद्यार्थी ऐसे हैं जिनके अंक 60 से अधिक है। ओगिव को प्रतिशत के आधार पर भी व्यक्त किया जा सकता है। उस दशा में हम यह बता सकते है कि कितने प्रतिशत विद्यार्थी ऐसे हैं जिनके अंक एक विशेष संस्था से कम या अधिक है। जब ओगिव को प्रतिशत के आधार पर व्यक्त किया जाता है तो ऐसे वक्र को शतमक वक्र कहते हैं।

संचयी आवत्ति वक्र का प्रयोग भूष्ठिक (Mode) चतुर्थक (Quartile) तथा दशमक (Deciles), आदि ज्ञान करने के लिये किया जाता है।

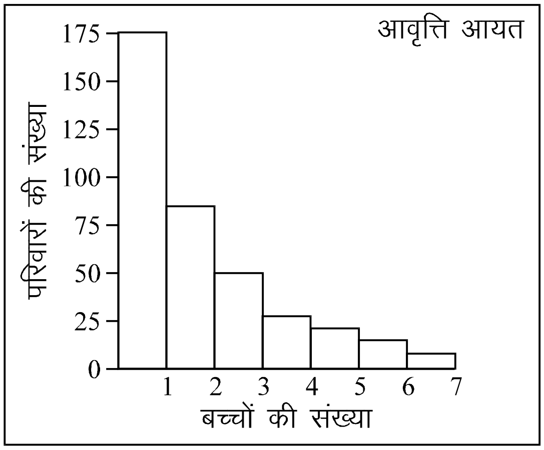
उदाहरण 15 : निम्न समंकों को एक आवत्ति आयत तथा संचयी आवत्ति वक्र :

बच्चों की संख्या	परिवारों की संख्या	बच्चों की संख्या	परिवारों की संख्या
0	171	4	13
1	82	5	7
2	50	6	2
3	25		

हल : ऊपर दिये गये समंकों को इस प्रकार व्यवस्थित कीजिये :-

बच्चों की संख्या	परिवारों की संख्या	बच्चों की संख्या	परिवारों की संख्या
0-1	171	4-5	13
1-2	82	5-6	7
2-3	50	6-7	2
3-4	25		

दिये हुए समंकों का आवत्ति -वितरण इस प्रकार होगा।



अब दिये हुये समंकों को 'Less Than' विधि के अनुसार व्यविस्थत कीजिये :-

बच्चों की संख्या	परिवारों की संख्या
1 से कम	171
2 से कम	253
3 से कम	303
4 से कम	328
5 से कम	341
6 से कम	348
7 से कम	350

❑❑

अध्याय - 6

सांख्यिकीय माध्य (Statistical Averages)

पिछले अध्याय में समकों को एकत्र करने तथा ठीक ढंग से प्रस्तुत करने की विभिन्न रीतियों का वर्णन किया गया। समंकों को एकत्र करने के पश्चात उनका विश्लेषण किया जाता है। विश्लेषण का उद्देश्य समकों की प्रमुख विशेषताओं का ज्ञान प्राप्त करना है। समंकों के विश्लेषण के लिये अनेक विधियों का प्रयोग किया जाता है जैसे सांख्यिकीय माध्य (Statistical Averages), अपरिकण (Dispersion), विषमता और पथुशीर्षत्व (Skewness and Kurtosis) आदि।

माध्य वह मान है, जो श्रेणी के समस्त समंकों का प्रतिनिधित्व करता है। इसके द्वारा हजारों और करोड़ों समंकों की विशेषता एक ही संख्या में व्यक्त की जाती है। उदाहरणार्थ, एक कारखाने में 2,000 व्यक्ति काम करते हैं, प्रत्येक कर्मचारी के वेतन को याद रखना एक असंभव सी बात है। लेकिन यदि कर्मचारियों को दिये गये कुल वतेन के कर्मचारियों के वेतन का प्रतिनिधित्व करेगा। यदि 2,000 व्यक्तियों को कुल मिलाकर 20,00,000 रू. मासिक दिया जाता है तो प्रति-व्यक्ति औसत वेतन 1000 रू. होगा। इसका अर्थ यह नहीं कि कारखाने के प्रत्येक व्यक्ति को 1,000 रू. मासिक मिलते हैं — कुछ व्यक्तियों को 1,000 से कम मिलते होंगे, कुछ को 1,000 से अधिक, और कुछ को पूरे 1,000 ही। इसी प्रकार जब हम कहते हैं कि भारत में प्रति-व्यक्ति औसत आयु 55 वर्ष है तो इसका अर्थ यह नहीं कि प्रत्येक व्यक्ति 55 वर्ष तक जीवित रहता है; कुछ व्यक्ति की 55 वर्ष से पहले ही मत्यु हो जाती है , कुछ 55 वर्ष के बहुत बाद तक जीवित रहते हैं कुछ व्यक्ति लगभग 55 वर्ष की ही आयु पाते हैं। अर्थात् माध्य वह मान है जो श्रेणी के समस्त समंकों का प्रतिनिधित्व करता है। माध्य की कुछ परिभाषायें इस प्रकार हैं।

''माध्य केवल एक गणितीय धारणा है — जैसा कि विभिन्न प्रकार की जनसंख्या के जीवन की माध्य अवधि, जो किसी विशेष समूह के अनुरूप नहीं होती परन्तु अंकगणितीय निष्कर्ष को अभिव्यक्ति करने का संक्षिप्त ढंग है।''

''माध्य समंकों को विस्तार के अंतर्गत स्थित एक ऐसा मान है, जिसका प्रयोग श्रेणी के सभी मूल्यों का प्रतिनिधित्व करने के लिये किया जाता है। एक माध्य समंकों के विस्तार अन्तर्गत ही किया जाता है, इसलिये इसे केन्द्रीय मूल्य का माप भी कहते हैं।

माध्यों को 'केन्द्रीय प्रवत्ति का माप' भी कहते हैं। इसका कारण यह है कि समूह के व्यक्तिगत मद माध्य मूल्य के चारों ओर केन्द्रित रहते हैं। माध्य को स्थान संबंधी, प्रतिरूपी मूल्य या सारांश अंक भी कहते हैं।

माध्यों का सांख्यिकी विज्ञान में बहुत महत्वपूर्ण स्थान है। सांख्यिकी विश्लेषण की अन्य बहुत-सी रीतियां माध्यों पर आधारित है। यही कारण है कि डॉ. बाउले ने सांख्यिकी को 'माध्यों का विज्ञान' कहा है। माध्यों की सहायता से समंक श्रेणी के सभी मूल्यों का सार प्रकट किया जाता है। सांख्यिकी की व्यक्तिगत इकाइयों का अलग-अलग कोई महत्व नहीं है। माध्यों द्वारा सभी इकाइयों में सामूहिक रूप से पाये जाने वाले मुख्य लक्षण स्पष्ट हो जाते हैं, तथा उनकी तुलना भी सरल हो जाती है।

सांख्यिकीय माध्यों के उद्देश्य एवं कार्य
(Objects and Functions of Statistical Averages)

सांख्यिकीय माध्यों के उद्देश्य व कार्य निम्नलिखित हैं :-

(1) **समंकों का संक्षिप्त चित्र प्रस्तुत करना (To Present a Brief Picture of the Entire Data) :** माध्यों द्वारा जटिल और व्यवस्थित समंकों की मुख्य विशेषतओं का एक सरल, स्पष्ट एवं संक्षिप्त चित्र प्रस्तुत किया जाता है। इससे उन समंकों का समझना व याद रखना बहुत सुगम हो जाता है। उदाहरणार्थ 80 करोड़ भारतवासियों की अलग-अलग आयु को याद रखना एक असंभव-सी बात है, परन्तु औसत आयु प्रत्येक व्यक्ति याद रख सकता है। इसी प्रकार 80 करोड़ व्यक्तियों की आय के समंक याद रखना असंभव है, लेकिन औसत आय सुगमता से याद रखी जा सकती है। अत: माध्य समंकों का बिहंगम दश्य प्रस्तुत करते हैं। मोरोने ने ठीक ही कहा है , ''माध्य व उद्देश्य व्यक्तिगत मूल्यों के समूह का सरल और संक्षिप्त रूप से प्रतिनिधित्व करना है, जिससे मस्तिष्क समूह की इकाइयों के सामान्य आकार को शीघ्रता से ग्रहण कर सके।''

(2) **तुलना में सहायक होना (To Facilitate Comparison) :** माध्य समंकों की उस राशि को संक्षिप्त व सरल करके तुलना योग्य बनाते हैं समकों की तुलना से बहुत महत्वपूर्ण निष्कर्ष निकाले जा सकते हैं। उदाहरणार्थ विभिन्न देशों की औसत आय की गणना से ज्ञात किया जा सकता है कि कौन-सा देश सबसे अधिक समद्धशाली है तथा कौन सबसे कम।

(3) **उपयुक्त नीतियों के निर्धारण में सहायक होना (To Help in the Formulation of Suitable Policies)** माध्य उपयुक्त नीतियों के निर्धारण में बहुत अधिक सहायक होते हैं। उदाहरणार्थ - यदि किसी विद्यालय में किसी विषय (Subject) में औसत नम्बर इस प्रकार हैं — 60, 58, 40 एंव 55 तो इससे यह निष्कर्ष निकलेगा कि कक्षा 'ग' के विद्यार्थी इस विषय में बहुत कमजोर हैं और उनकी इस कमी को दूर करने के लिये कोई विशेष प्रबन्ध करना आवश्यक है।

(4) **सांख्यिकीय विश्लेषण के आधार (Basis of Statistical Analysis) :** सांख्यिकीय विश्लेषण की अनेक क्रियायें माध्यों पर आधारित है।

आदर्श माध्य के आवश्यक गुण
(Properties of an Ideal Average)

माध्य पूरे समूह का प्रतिनिधित्व करता है। ऐसे मूल्य में निम्न गुण होने चाहिये ताकि समंकों का ठीक रूप से प्रतिनिधित्व हो सके :

(1) **समझने में सरल (Easy to Understand)** सांख्यिकीय विधियों का प्रयोग समंकों को संक्षिप्त तथा सरल बनाने के लिये किया जाता है। अत: माध्य ऐसा होना चाहिये जो सुगमता से समझा जा सके, अन्यथा इसका प्रयोग बहुत की सीमित होगा।

(2) **निर्धारण में सुगम (Easy to Compute) :** माध्य की गणन-क्रिया सरल होनी चाहिये ताकि इसका प्रयोग व्यापक रूप से हो सके। यद्यपि माध्य का निर्धारण यथा-सम्भव सरल होना चाहिये तथापि विशेष परिस्थितियों में परिणामों की शुद्धता के लिये अधिक कठिन माध्यों का प्रयोग भी किया जा सकता है।

(3) **श्रेणी के सभी मूल्यों पर आधारित (Based on all the Items of the Series) :** माध्य श्रेणी के सभी मूल्यों पर आधारित होना चाहिये ताकि एक या अधिक मूल्यों में परिवर्तन होने के माध्य में भी परिवर्तन हो सके। यदि माध्य श्रेणी के सभी मूल्यों पर आधारित नहीं है तो वह पूरे समूह का ठीक प्रकार से प्रतिनिधित्व नहीं कर सकता।

(4) **न्यूनतम तथा अधिकतम मूल्यों पर अनुचित प्रभाव से बचाव (Should not be Unduly Affected by Extreme Items) :** यद्यपि माध्य सभी मूल्यों पर आधारित होना चाहिये तथापि किसी विशेष मूल्य का माध्य पर अधिक प्रभाव नहीं पड़ना चाहिये अन्यथा माध्य समंकों का सही प्रतिरूप व्यक्त नहीं करेगा।

(5) **स्पष्ट व स्थिर (Rigidly Defined) :** माध्य की परिभाषा स्पष्ट शब्दों में व्यक्त होनी चाहिये ताकि जो भी व्यक्ति दिये हुये समंकों से माध्य निकाले वह एक ही निष्कर्ष पर पहुंचे; इसलिये यह आवश्यक है कि माध्य गणितीय सूत्र के रूप में दिया जाये। यदि माध्य के परिगणन में व्यक्तिगत प्रवत्तियों का प्रभाव पड़ा तो फल भ्रामक तथा अशुद्ध होंगे।

(6) **बीजगणितीय विवेचन संभव (Capable of Algebraic Treatment) :** एक अच्छे माध्य का बीजगणितीय विवेचन संभव होना चाहिये। उदाहरणार्थ, यदि दो कारखानों में मजदूरों की संख्या तथा औसत आय से संबंधित समंक दिये गये हों तो दोनों कारखानों के मज़दूरों की आय का सामूहिक माध्य निकालना संभव होना चाहिये।

(7) **न्यादर्शो की भिन्नता का कम-से-कम प्रभाव (Less Effect of Fluctuations of Sampling) :** यदि एक ही समग्र में से उचित रीति द्वारा विभिन्न न्यादर्श लेकर माध्य निकाले जायें तो उन माध्यों में बहुत अधिक अंतर नहीं होना चाहिये। उदाहरणार्थ, यदि एक विश्वविद्यालय के विद्यार्थियों को दस भागों में बांटकर 10 न्यादर्श लिये गये हैं तो उनके परिणामों में बहुत अधिक समानता नहीं होनी चाहिये।

सांख्किीय माध्यों के प्रकार
(Kind of Statistical Averages)

माध्य कई प्रकार के होते हैं। सुविधा की दष्टि से उन्हें दो भागों में बांटा गया है। — (1) गणितीय माध्य (Mathematical Averages), तथा (2) स्थिति संबंधी माध्य (Averages of Position)।

(1) गणितीय माध्य — गणितीय माध्य मुख्य रूप से तीन प्रकार के होते है।

(क) सामान्तर माध्य का मध्यक (Arithmetic Average or Mean),

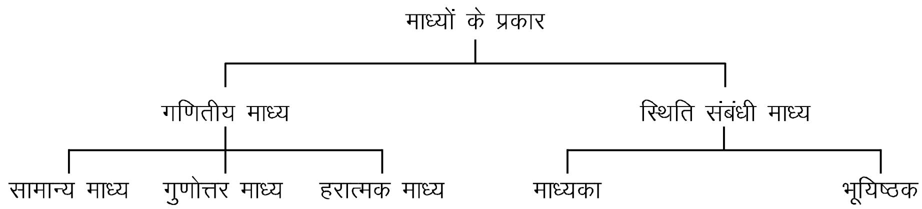
(ख) गुणोत्तर माध्य (Geometric Mean), तथा

(ग) हरात्मक माध्य (Harmonic Mean).

(2) स्थिति संबंधी माध्य — इन माध्यों का निर्धारण अधिकतर निरीक्षण मात्र से ही हो जाता है। अत: इन्हें निरीक्षण संबंधी माध्य (Inspectional Averages) भी कहते हैं। स्थिति संबंधी माध्य दो प्रकार के होते हैं :-

(क) माध्यका (Median), तथा

(ख) भूयिष्ठक (Mode)।



इन प्रमुख माध्यों के अतिरिक्त कुछ और माध्य भी होते हैं, जिनका प्रयोग बहुत अधिक प्रचलित नहीं हैं, जैसे चल माध्य (Moving Average), प्रमागी, वर्गकरणी माध्य (Quadratic Average), आदि।

आगे कुछ प्रमुख माध्यों का विस्तार से वर्णन किया गया है।

समान्तर माध्य
(Arithmetic Mean)

समान्तर माध्य (या मध्यक) का व्यवहार में सबसे अधिक प्रयोग होता है। वास्तव में 'औसत' शब्द का प्रयोग इसी माध्य के लिये किया जाता है। समान्तर माध्य वह मूल्य है जो किसी श्रेणी के समस्त मदों के मूल्यों के योग को उनकी संख्या से भाग देने पर प्राप्त होता हे।

समान्तर माध्य दो प्रकार का होता है : (1) सरल समान्तर माध्य तथा (2) भारित समान्तर माध्य।

सरल समान्तर माध्य में श्रेणी के समस्त मदों को समान महत्व दिया जाता है जबकि भारित समान्तर माध्य में प्रत्येक मद को उसकी व्यक्तिगत महत्ता के अनुसार भारन प्रदान करते हैं।

सरल समान्तर माध्य के परिगणन की रीति
(Method of Calculating Simple Arithmetic Mean)

सरल समान्तर माध्य का निर्धारण अत्यन्त सुगम है। इसकी गणन-क्रिया कि व्याख्या निम्न भागों में की गई है :-

(1) व्यक्तिगत श्रेणी में समान्तर माध्य का परिगणन

(2) विच्छिन्न श्रेणी में समान्तर माध्य का परिगणन

(3) अविच्छिन्न श्रेणी में समान्तर माध्य का परिगणन

व्यक्तिगत श्रेणी में समान्तर माध्य का परिगणन
(Calculation of Arithmetic Mean in Individual Series)

व्यक्तिगत श्रेणी में समान्तर माध्य दो प्रकार से ज्ञात किया जा सकता है (1) प्रत्यक्ष रीति (Direct Method) तथा (2) लघु रीति (Short-cut Method) ।

(1) प्रत्यक्ष रीति (Direct Method) : इस रीति के अन्तर्गत सब मदों के मानों को जोड़कर मदों की संख्या से भाग दे दिया जाता है। इस प्रकार जो अंक प्राप्त होता है उसे समान्तर माध्य कहते हैं। सूत्र के रूप में :-

$$\overline{X} = \frac{X_1 \quad X_2 \quad X_3 \quad X_4 \quad ......X_n}{N}$$

X_1, X_2 आदि श्रेणी वे विभिन्न मान

N = मदो की संख्या (Number of Observations)

संक्षिप्त रूप में $$\overline{X} = \frac{X}{N}$$

$\overline{X}$ समान्तर माध्य (Arithmetic Mean), N = मदों की संख्या, SX = चर X के विभिन्न मानों का योग।

उदाहरण 1 : एक कक्षा में 19 विद्यार्थी हैं। सांख्यिकी की एक मासिक परीक्षा में अंक इस प्रकार हैं :-

रोल नम्बर	**:**	1	2	3	4	5	6	7	8	9	10
अंक	**:**	80	65	45	58	62	72	70	40	25	33

समान्तर माध्य ज्ञात कीजिये।

हल : **समान्तर माध्य का परिगणन प्रत्यक्ष रीति द्वारा**

रोल नम्बर (Roll No.)	**अंक X**
1	80
2	65
3	45
4	58
5	62
6	72
7	70
8	40
9	25
10	33
N = 19	**SX = 550**

$$\overline{X} = \frac{X}{N} = \frac{650}{10} = 55 \text{ अर्थात औसत अंक } = 55$$

(2) लघु रीति (Short-cut Method) : लघु रीति से समान्तर माध्य के परिगणन की विधि निम्नलिखित है :-

(*i*) श्रेणी के किसी भी मूल्य का अन्य किसी संख्या को कल्पित मध्यक (Assumed Mean) : मान लेते हैं। चाहे कोई भी मूल्य कल्पित मध्यक के रूप में लिया जाये, उत्तर एक ही आयेगा परन्तु जितना समान्तर माध्य केन्द्रीय माप के समीप होगा उतने की विचलन कम आयेंगे और गणना में सरलता रहेगी।

(*ii*) इस कल्पित समान्तर माध्य को श्रेणी के प्रत्येक मद में से घटा दिया जाता है और उसके सामने विचलन लिखते जाते हैं। घटाते समय ऋणात्मक (–) तथा धनात्मक (+) चिन्हों का ध्यान रखा जाता है। मान लीजिये श्रेणी का पहला पद 40 है और कल्पित समान्तर माध्य 50 है तो हम (40 – 50) अर्थात् – 10 लिखेंगे। इसी प्रकार यदि श्रेणी का दूसरा मद 72 है और कल्पित समान्तर माध्य 50 तो (70 – 50) = 22 लिखेंगे। इससे स्पष्ट होगा कि यदि मूल्य कल्पित समान्तर माध्य से कम है तो विचलन ऋणात्मक और अधिक है तो धनात्मक होगा।

(*iii*) सभी विचलनों को जोड़ लिया जाता है।

(*iv*) विचलनों के योग को मदों की संख्या से भाग दिया जाता है।

(*v*) भागफल यदि धनात्मक है तो कल्पित मध्यक में जोड़ देते हैं, ऋणात्मक है तो घटा देते हैं। इस प्रकार प्राप्त संख्या समान्तर होगी।

सूत्र के रूप में :

$$\overline{X} = A + \frac{d}{N}$$

A = कल्पित समान्तर माध्य (Assumed Mean), d = (X – A)

उपर्युक्त उदाहरण से समान्तर माध्य का परिगणन लघु रीति द्वारा इस प्रकार होगा।

समान्तर माध्य का परिगणन लघु रीति द्वारा

रोल नम्बर (Roll No.)	अंक (X)	कल्पित समान्तर माध्य 50 से विचलन (X – 50) *d*
1	80	+ 30
2	65	+ 15
3	45	– 5
4	58	+ 8
5	62	+ 12
6	72	+ 22
7	70	+ 20
8	40	– 10
9	25	– 25
10	33	– 17
N = 10		**S*d* = 50**

$$\overline{X} = A + \frac{d}{N}$$

A = 50; Sd = 50; N = 10

\ $$\overline{X} = 50 + \frac{50}{10} = 50 = 5 = 55 \text{ अंक}$$

इससे यह स्पष्ट होगा कि दोनों रीतियों से उत्तर एक ही आयेगा। इन रीतियों को ध्यान में रखते हुए पाठक को ऐसा प्रतीत होगा की लघु रीति प्रत्यक्ष रीति से कठिन होते हुये भी इसका नाम लघु रीति रखा गया है। यह बात व्यक्तिगत श्रेणी के साथ बिल्कुल ठीक है। इसलिये व्यक्तिगत श्रेणी में लघु रीति का प्रयोग नहीं करना चाहिये। लेकिन विच्छिन्न व अविच्छिन्न श्रेणी में लघु रीति का प्रयोग गणन क्रिया को बहुत सुगम बना देता है।

विच्छिन्न श्रेणी में समान्तर माध्य का परिगणन

(Calculation of Arithmetic Mean in Discrete Series)

उपर्युक्त दोनों रीतियों द्वारा विच्छिन्न श्रेणी में भी समान्तर माध्य ज्ञात किया जा सकता है। परन्तु विच्छिन्न श्रेणी में आवत्तियाँ होती हैं इसलिये सूत्र भिन्न हैं। जब प्रत्यक्ष रीति का प्रयोग किया जाता है जो सूत्र इस प्रकार होगा :-

$$\overline{X} = \frac{fx}{N}$$

Sfx = चर तथा आवत्तियों की गुणाओं का योग , N = आवत्तियों का योग

विधि :

(1) चल के प्रत्येक मूल्य को उसकी आवत्ति से गुणा कीजिये और सभी गुणनफलों को जोड़ लीजिये।

(2) प्राप्त गुणनफल को आवत्तियों के कुल योग से भाग दीजिये।

उदाहरण 2 : एक कक्षा के 60 विद्यार्थियों द्वारा प्राप्त निम्न अंकों के औसत अंक निकालिये :-

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
20	8	50	10
30	12	60	6
40	20	70	4

हल : अंकों को X से तथा विद्यार्थियों की संख्या को f से प्रदर्शित कीजिये :-

समान्तर माध्य का परिगणन लघु रीति द्वारा

अंक X	विद्यार्थियों की संख्या f	fX
20	8	160
30	12	360
40	20	800
50	10	500
60	6	360
70	4	280
	N = 60	**SfX = 2,460**

$$\overline{X} = \frac{f\,X}{N} = \frac{2,460}{60} = 41 \text{ अंक}$$

लघु रीति (Short-cut Method) : जब लघु रीति का प्रयोग किया जाता है सूत्र इस प्रकार होगा :-

$$= A + \frac{fd}{N}$$

विधि :

(1) एक कल्पित समान्तर माध्य लीजिये।

(2) चर मूल्यों को कल्पित समान्तर माध्य से घटाकर विचलन निकालिये।

(3) प्राप्त विचलनों को उनकी आवत्तियों से गुणा कीजिये और जोड़ लीजिये अर्थात Sfd निकालिये।

(4) इस प्रकार प्राप्त योग को कुल आवत्ति से विभाजित कीजिये।

(5) भागफल यदि धन में आये तो समान्तर माध्य में जोड़ दीजिये और ऋण में आये तो कल्पित समान्तर माध्य में से घटा दीजिये। इस प्रकार प्राप्त संख्या समान्तर माध्य होगी।

उपर्युक्त प्रश्न को लघु रीति द्वारा हल किया गया है।

अंक X	विद्यार्थियों की संख्या f	(X – 40) d	fd
20	8	– 20	– 160
30	12	– 10	– 120
40	20	0	0
50	10	+ 10	+ 100
60	6	+ 20	+ 120
70	4	+ 30	+ 120
	N = 60		**S*fd* = 60**

$$\overline{X} = A + \frac{fd}{N} = 40 + \frac{60}{60} = 40 + 1 = 41$$

अविच्छिन्न श्रेणी में समान्तर माध्य का परिगणन
(Calculation of Arithmetic Mean in Continuous Series)

अविच्छिन्न श्रेणी में समान्तर माध्य ज्ञात करने की निम्न विधियां हैं :-

(1) प्रत्यक्ष रीति (Direct Method),

(2) लघु रीति (Short-cut Method), तथा

(3) पद-विचलन (Step Deviation Method)

(1) प्रत्यक्ष रीति (Direct Method) : प्रत्यक्ष रीति द्वारा समान्तर

$$\overline{X} = \frac{fm}{N}$$

$$= \text{मध्य-बिन्दु (mid-point)}$$

विधि :

(1) श्रेणी के प्रत्येक वर्ग का मध्य-बिन्दु ज्ञात कीजिये :

$$\text{मध्य-बिन्दु} = \frac{\text{निम्न सीमा} + \text{उच्च सीमा}}{2}$$

(2) प्रत्येक श्रेणी के मध्य-बिन्दु को उसकी आवत्ति से गुणा कीजिये और कुल योग ज्ञात कीजिये अर्थात S*fm* निकालिये।

उदाहरण 3- प्रत्यक्ष रीति से निम्नलिखित समंकों का समान्तर माध्य ज्ञात कीजिए :-

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
0-10	5	30-40	30
10-20	10	40-50	20
20-30	25	50-60	10

हल : समान्तर माध्य का परिगणन (प्रत्यक्ष रीति द्वारा)

अंक X	मध्य-बिन्दु *m*	विद्यार्थियों की संख्या *f*	*fm*
0-10	5	5	25
10-20	15	10	150
20-30	25	25	625
30-40	35	30	1,050
40-50	45	20	900
50-60	55	10	550
		N = 100	**S*fm* = 3,300**

$$\overline{X} = \frac{fm}{N} = \frac{3,300}{100} = = 33 \text{ अंक}$$

लघु रीति (Short-cut Method) : जब लघु रीति का प्रयोग किया जाता है सूत्र इस प्रकार होगा :-

$$= A + \frac{fd}{N}$$

$d = (X - A)$ अर्थात् मध्य बिन्दुओं का कल्पित समान्तर माध्य से विचलन।

विधि :

(1) प्रत्येक वर्ग सीमा का मध्य-बिन्दु निकालिये। $\overline{X}$

(2) मध्य बिन्दुओं में से एक कल्पित माध्य लीजिये।

(3) प्रत्येक मध्य-बिन्दु से कल्पित समान्तर माध्य घटाइये और विचलनों को एक अलग स्तम्भ में लिखिये।

(4) प्रत्येक श्रेणी की आवत्ति को उसके विचलन से गुणा कीजिये और इस स्तम्भ का कुल योग निकालिये अर्थात S*fd* ज्ञात कीजिये।

(5) इस योग को आवत्तियों के कुल योग से भाग दीजिये। यदि भागफल धनात्मक (+) में आये तो कल्पित समान्तर माध्य में जोड़ दों अन्यथा घटा दें। इस प्रकार प्राप्त संख्या मध्यक होगी।

उदाहरण 4 : उदाहरण 3 के समंकों से लघु रीति द्वारा समान्तर माध्य ज्ञात कीजिये।

हल : समान्तर माध्य का परिगणन (प्रत्यक्ष रीति द्वारा)

अंक	मध्य-बिन्दु *m*	विद्यार्थियों की संख्या *f*	(*m* – 35) *d*	*fd*
0-10	5	5	– 30	– 150
10-20	15	10	– 20	– 200
20-30	25	25	– 10	– 250
30-40	35	30	0	0
40-50	45	20	+ 10	+ 200
50-60	55	10	+ 20	+ 200
		N = 100		**S*fd* = – 200**

$$\overline{X} = A - \frac{fd}{N} = 35 - \frac{200}{100} = 35 - 2 = 33$$

पद-विचलन रीति (Step Deviation Method) : यह रीति गणन-क्रिया को और भी सरल बना देती है। इस रीति में तथा लघु रीति में थोड़ा ही अंतर है। जब इसका प्रयोग होता है तो विचलनों को उभयनिष्ठ गुणक (Common Factor) से भाग देकर पद विचलन ज्ञात किये जाते हैं। इन पद-विचलनों को उनकी आवत्ति से गुणा करके कुल योग ज्ञात कर देते हैं। फिर निम्न सूत्र को प्रयोग किया जाता है।

$$\overline{X} = A + \frac{fd}{N} \times i$$

$$d = (m - A)/i \text{ वर्ग अन्तराल}$$

विधि :

(1) वर्ग सीमाओं के मध्य-बिन्दु ज्ञात कीजिये।

(2) मध्य बिन्दुओं में से एक को कल्पित समान्तर माध्य मान लीजिये।

(3) प्रत्येक मध्य-बिन्दु से कल्पित समान्तर माध्य घटाकर अन्तर को उभयनिष्ठ गुणक से भाग दीजिये।

(4) पद-विचलनों को आवत्तियों से गुणा करके गुणफलों को जोड़ ज्ञात कीजिए और उपर्युक्त सूत्र का प्रयोग कीजिये।

उदाहरण 5 : उदाहरण 4 के समंकों से पद-विचलन रीति का प्रयोग करके समान्तर माध्य निकालिये।

हल : **समान्तर माध्य का परिगणन पद-विचलन रीति द्वारा**

अंक	मध्य-बिन्दु m	विद्यार्थियों की संख्या f	$(m - 35)$ d	$(m - 35)/10$ fd	
0-10	5	5	– 30	– 3	– 15
10-20	15	10	– 20	– 2	– 20
20-30	25	25	– 10	– 1	– 25
30-40	35	30	0	0	0
40-50	45	20	+ 10	+ 1	+ 20
50-60	55	10	+ 20	+ 2	+ 203
		N = 100			**Σfd = – 20**

$$\overline{X} = A = \frac{fd}{N} \times i = 35 - \frac{20}{100} \times 10 = 35 - 2 = 33 \text{ अंक}$$

उपर्युक्त उदाहरणों से यह स्पष्ट है कि तीनों रीतियों द्वारा एक ही उत्तर आता है। प्रत्यक्ष रीति सबसे सुगम प्रतीत होती है। परन्तु जहां मध्य-बिन्दु और आवत्तियों का आकार बड़ा हो वहां गणन -क्रिया जटिल हो जाती है। ऐसी परिस्थिति में पद-विचलन रीति अधिक उपर्युक्त रहती है जैसा निम्न सारणी से स्पष्ट है।

आय (रुपयों में)	व्यक्तियों की संख्या	आय (रुपयों में)	व्यक्तियों की संख्या
700-800	1,068	1000-1100	1,567
800-900	1,472	1100-1200	904
900-1000	2,969		

टिप्पणी : अविच्छिन्न श्रेणी में समान्तर माध्य ज्ञात करते समय विभिनन वर्गों के मध्य-बिन्दुओं को उस वर्ग का प्रतिनिधित्व मान लिया जाता है। इसका कारण यह है कि जब समंक श्रेणी-बद्ध (grouped) होते हैं तब तब यह ज्ञात करना असंभव सा

होता है कि व्यक्तिगत आवत्ति क्या है। हम केवल वर्ग की सीमा जानते हैं जिसमें वह आवत्ति सम्मिलित है। उदाहरण के लिये , जब हम यह कहते हैं कि 500-600 के वर्ग में 50 व्यक्ति हैं तब सही रूप से कहना कि 50 में से कितने व्यक्ति 501, 502, 503 रू. आदि प्राप्त करते हैं संभव नहीं। इसलिये समान्तर माध्य निकालते समय हम कल्पना करते हैं कि श्रेणी की आवत्तियां वर्गान्तर के विस्तार के समान रूप से बिखरी हुई हैं। अर्थात् मध्य-बिन्दु से नीचे तथा ऊपर मदों की संख्या समान है। जब तक इस प्रकार की कल्पना न की जाये, समान्तर माध्य का मूल्य अखंडित श्रेणी में नहीं निकाला जा सकता।

उदाहरण 6 : एक निर्वाचन अनुमान संबंधी संस्था ने 200 व्यक्तियों का साक्षात्कार (Interview) किया। निम्न आवत्ति - वितरण साक्षात्कार किये गये व्यक्तियों की आयु प्रस्तुत करता है।

आयु (वर्षों में)	आवत्ति	आयु (वर्षों में)	आवत्ति
80-89	2	40-49	56
70-79	2	30-39	40
60-69	6	20-29	42
50-59	20	10-19	32

इस आवत्ति -वितरण से समान्तर माध्य ज्ञात कीजिये।

हल : **समान्तर माध्य का परिगणन पद-विचलन रीति द्वारा**

आयु (वर्षों में)	*m*	*f*	**(*m* – 44.5)**	**(*m* – 44.5)/10** *d*	*fd*
10-19	14.5	32	– 30	– 3	– 96
20-29	24.5	42	– 20	– 2	– 84
30-39	34.5	40	– 10	– 1	– 40
40-49	44.5	56	0	0	0
50-59	54.5	20	+ 10	+ 1	+ 20
60-69	64.5	6	+ 20	+ 2	+ 12
70-79	74.5	2	+ 30	+ 3	+ 6
80-89	84.5	2	+ 40	+ 4	+ 8
		N = 200			**S*fd* = – 174**

$$\overline{X} = A + \frac{fd}{N} \times i = 44.5 - \frac{174}{200} \times 10 = 44.5 - 8.7 = 35.8 \text{ वर्ष}$$

टिप्पणी : जब समंक सम्मिलित रीति (Inclusive Method) द्वारा प्रस्तुत किये गये हों तो समान्तर माध्य ज्ञात करते समय वर्गों में संयोजन (Adjustment) आवश्यक नहीं क्योंकि मध्य-बिन्दु का संयोजन का कोई प्रभाव नहीं पड़ता, परन्तु मध्यक तथा भूयिष्ठक (Mode) निकालते समय वर्गों में संयोजन आवश्यक है।

उदाहरण 7 : एक फर्म के 30 कर्मचारियों के मासिक वेतन के समंक निम्न हैं।

140	139	126	114	100	88	62	77	99
108	129	144	148	138	63	79	148	132
141	116	123	104	95	80	85	106	123

फर्म ने विभिन्न वेतन वर्गों के 60 से अधिक किन्तु 75 से अधिक नहीं, 76 से अधिक 90 से अधिक नहीं और इसी प्रकार 135 से अधिक किन्तु 150 से अधिक नहीं, के बोनस की राशि प्रदान की। बोनस की औसत राशि ज्ञात कीजिये।

हल: **बोनस की औसत राशि का परिगणन**

मासिक वेतन (रू. में)	मिलान रेखायें	आवत्ति f	बोनस X	कुल बोनस fX
61-75	||	2	10	20
76-90	𝍸	5	15	75
91-405	𝍸	5	20	100
106-120	𝍸	5	25	125
121-135	𝍸 |	6	30	180
136-150	𝍸 ||	7	35	245
		N = 30		**SfX = 745**

$$\overline{X} = \frac{fx}{N} = \frac{745}{30} = 24.83$$

अर्थात बोनस की औसत राशि = 24.83

उदाहरण 8 : 100 विद्यार्थियों के प्राप्तांकों का औसत 40 था, किन्तु बाद में यह पता लगा कि एक विद्यार्थी के प्राप्तांक 53 के स्थान पर गलती से 83 पढ़ लिये गये थे। सही प्राप्तांकों के अनुरूप सही औसत ज्ञात कीजिये।

हल :
$$\overline{X} = \frac{X}{N} \qquad \backslash\ N\overline{X} = SX$$

$$N = 100,\ \overline{X} = 40, \qquad \backslash\ SX = 100 \times 40 = 4{,}000$$

लेकिन यह सही SX नहीं है। सही SX ज्ञात करने के लिये गलत मद को घटाना पड़ेगा और ठीक मद को जमा करना पड़ेगा।

सही SX + गलत SX – गलत मद + सही मद

$$= 4{,}000 - 83 + 53 = 3{,}970$$

$$\text{सही समान्तर माध्य} = \frac{3{,}970}{100} = 39.7$$

सही प्राप्तांक के अनुरूप सही औसत 39.7 है।

उदाहरण 9 : 200 मदों के समान्तर माध्य 50 थे। बाद में, यह ज्ञात हुआ कि दो मदों का मूल्य 192 और 88 के स्थान पर 92 और 8 ले लिया गया है। सही समान्तर माध्य ज्ञात कीजिये।

हल :
$$= \frac{X}{N} \qquad \backslash\ SX = NX$$

$$\overline{X} = 50 \qquad \backslash\ N = 200$$

$$\backslash \qquad SX = 50 \times 200 = 10{,}000$$

सही SX + गलत SX – गलत मद + सही मद

$$= 10{,}000 - 92 - 8 + 192 + 88$$

$$= 10{,}000 - 100 + 280 = 10{,}180$$

$$\overline{X} = \frac{10{,}181}{200} = 50.9$$

उदाहरण 10 : यदि किसी अचल (Constant) को श्रेणी के प्रत्येक मूल्य में जोड़े अथवा उसमें से घटायें तो समान्तर माध्य पर उसका क्या प्रभाव पड़ेगा ?

हल : मान लीजिये, एक कक्षा के 5 विद्यार्थियों द्वारा प्राप्तांक इस प्रकार हैं :-

10, 15, 18, 22, 25

$$= \frac{10 \quad 15 \quad 18 \quad 22 \quad 25}{5} = \quad = 18$$

यदि प्रत्येक लड़के को 5 नम्बर और दिये जायें तो ये मद इस प्रकार होंगे :-

15, 20, 23, 27, 30

$$\overline{X} = \frac{15 \quad 20 \quad 23 \quad 27 \quad 30}{5} = \quad = 23$$

इससे विदित होता है कि यदि हम एक श्रेणी के सब पदों को एक अचल से बढ़ाये तो समान्तर माध्य भी उतना ही बढ़ जायेगा जितना अचल का मूल्य है। इस प्रकार यदि एक श्रेणी के सभी मदों में से कोई अचल घटाया जाये तो समान्तर माध्य का मूल्य भी उतना ही कम हो जायेगा।

ऊपर दिये उदाहरण में यदि मद में से 5 घटाया जाये तो विभिन्न मद इस प्रकार होंगे :-

5, 10, 13, 17, 20

$$\overline{X} = \frac{5 \quad 10 \quad 13 \quad 17 \quad 20}{5} = \quad = 13$$

इस प्रकार समान्तर माध्य 5 से कम हो गया।

खुले-सिरे वाले आवत्ति वितरणों में समान्तर माध्य का निर्धारण
(Calculation of Arithmetic Mean in Case of Open-end Frequency Distribution)

खुल-सिरे वाले आवत्ति वितरण वे होते हैं जिनमें प्रथम वर्ग की निम्न सीमा तथा अन्तिम वर्ग की उच्च सीमा नहीं दी होती। इस दशा में समान्तर माध्य ज्ञात नहीं किया जा सकता जब तक हम अज्ञात सीमाओं के बारे में कल्पना न कर लें, खुले-सिरे वाली श्रेणियों के वर्गान्तर की कल्पना पहली श्रेणी के बाद वाली श्रेणी तथा अंतिम श्रेणी से पहले वाली श्रेणी के वर्गान्तर पर निर्भर करेगी। उदाहरणार्थ, निम्न सारणी देखिये :-

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10 से कम	4	30-40	15
10-20	6	40-50	8
20-30	10	50 से अधिक	7

उपर्युक्त उदाहरण में वर्गान्तर समान है। इसलिये यह कल्पना करना उचित होगा कि प्रथम वर्ग की निम्न सीमा शून्य है तथा अन्तिम वर्ग की उच्च सीमा 60 है।

एक और उदाहरण लीजिये।

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10 से कम	4	60-100	7
10-30	6	100 से अधिक	3
30-60	8		

उपर्युक्त उदाहरण में दूसरी श्रेणी में वर्गान्तर 20 है, तीसरी में 30, चौथी में 40 अर्थात् प्रत्येक वर्ग में वर्गान्तर 10 से बढ़ रहा है। ऐसी स्थिति में यह कल्पना उपयुक्त होगी कि पहली श्रेणी 0 से 10 तथा अंतिम श्रेणी 100 से 150 है।

यदि वर्गान्तर भिन्न-भिन्न विस्तार के हों तो खुले सिरे वाले वर्गों का निर्धारण प्रथम वर्ग के बाद वाले तथा अंतिम वर्ग के पहले वर्ग-विस्तार को ध्यान में रखकर किया जायेगा जैसा कि निम्न उदाहरण से स्पष्ट होगा :-

आय (रुपयों में)	व्यक्तियों की संख्या	आय (रुपयों में)	व्यक्तियों की संख्या
50 से कम	8	100-110	30
50-70	12	110-120	7
70-100	20	120 से अधिक	3

क्योंकि वर्गान्तर भिन्न हैं इसलिये प्रथम वर्ग की सीमा दूसरे वर्ग की सीमा के अनुसार निश्चित की जायेगी अर्थात् यह 30-50 होगी। इसी प्रकार अन्तिम वर्ग की सीमा उससे पहले वर्ग की सीमा को ध्यान में रखकर निश्चित की जायेगी अर्थात् यह 120-130 होगी।

टिप्पणी : खुले-सिरे वाली सारणी में समान्तर माध्य के स्थान पर मध्यक या भूयिष्ठक ज्ञात करना उपयुक्त होगा क्योंकि समान्तर माध्य निकालते समय वर्ग की निम्न सीमा तथा अन्तिम वर्ग की उच्च सीमा के बारे में कल्पना करना आवश्यक होगा, लेकिन मध्यक, भूयिष्ठक खुले-सिरे वाली सारणी में सुगमता से ज्ञात किये जा सकते हैं।

समान्तर माध्य के गणितीय गुण
(Mathematical Properties of Arithmetic Mean)

मध्यक के कुछ प्रमुख गणितीय गुण इस प्रकार हैं :-

(1) मध्यम से विभिन्न मदों के विचलनों का योग शून्य होता अर्थात $S(X-\bar{X})=0$ यह तथ्य अग्रांकित उदाहरण से स्पष्ट हो जायेगा :-

X	**$(X-\bar{X})$**
10	– 20
20	– 10
30	0
40	+ 10
50	+ 20
$SX = 150$	**$S(X-\bar{X})=0$**

उपर्युक्त उदाहरण में समान्तर माध्य 30 है। जब विभिन्न मदों का विचलन 30 से लिया गया तो योग $S(X-\bar{X})$ शून्य आया। इस गुण के कारण समान्तर माध्य को संतुलन बिन्दु (Point of Balance) कहते हैं, अर्थात् धनात्मक विचलनों का योग ऋणात्मक विचलनों के योग के बराबर होता है। लघु रीति भी इसी गुण पर आधारित है।

(2) यदि हम विभिन्न मदों को मध्यक से घटायें और इन विचलनों के वर्गों का योग निकालें तो यह योग उस योग से कम होगा जो समान्तर माध्य के अतिरिक्त किसी और मूल्य से निकाला जाये अर्थात $Sd^2 = \text{Minimum}$। इस तथ्य की पुष्टि निम्न उदाहरण द्वारा हो जायेगी :-

X	**$(X-\bar{X})$** **$\bar{X}=4$**	**$(X-\bar{X})^2$**	**$(X-3)^2$**
2	– 2	4	1
3	– 1	1	0
4	0	0	1
5	+ 1	1	4
6	+ 2	4	9
$SX = 20$	**$S(X-\bar{X})=0$**	**$S(X-\bar{X})^2=10$**	**$S(X-3)^2=15$**

उपर्युक्त उदाहरण से स्पष्ट होगा कि जब हमने समान्तर माध्य से लिये गये विचलनों के वर्गों का योग किया तो यह 10 आया और जब एक कल्पित माध्य जोकि समान्तर माध्य से कम है तो विचलनों के वर्गों का योग किया तो वह 15 आया।

(3) समान्तर माध्य का सामान्य विभ्रम (Standard Error) अन्य माध्य की तुलना में कम होता है।

(4) $\overline{X} = \frac{X}{N}$ या N $\overline{X}$ X

अर्थात् यदि हम श्रेणी के प्रत्येक मूल्य के स्थान पर मध्यक का मूल्य रख दें तो इन परिवर्तित मूल्यों का योग व्यक्तिगत मूल्यों के योग के बरारब होगा। उदाहरणार्थ, प्रथम गुण के विवेचन में SX 150 था तथा मध्यक 30; यदि सब मदों के स्थान पर हम 30 लिखें तो यही योग प्राप्त हो जायेगा अर्थात् (30 + 30 + 30 + 30 + 30) = 150

इस गुण का बहुत अधिक व्यवहारिक महत्व है। यदि हमें औसत तथा मदों की संख्या दी हो तो कुल योग सरलता से ज्ञात कर सकते हैं। उदाहरणार्थ, यदि एक कारखाने में मज़दूरों का औसत वेतन 200 रुपये है और मज़दूरों की संख्या 50 है तो हम मज़दूरों को दिये गये कुल वेतन का पता इस प्रकार लगा सकते हैं : $SX = N\overline{X}$ इस उदाहरण में कुल वेतन 200 × 50 अर्थात् 10,000 रुपये होगा।

यदि एक समूह के दो या दो से अधिक भागों के समान्तर माध्य तथा उनके मदों की संख्या ज्ञात है तो उनके आधार पर सामूहिक समान्तर माध्य (Combined Arithmetic Mean) प्राप्त किया जा सकता है। सामूहिक समान्तर माध्य ज्ञात करते समय निम्न सूत्र का प्रयोग होता है।

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2}{N_1 \quad N_2}$$

$\overline{X}_{12}$ = सामूहिक समान्तर माध्य (Combined Mean)

N_1 = पहले भाग के मदों की संख्या (Number of Items in the First Group)

N_2 = दूसरे भाग के मदों की संख्या (Number of Items in the Second Group)

$$\overline{X}\frac{57,500 \quad 42,000}{180} \quad \frac{99,500}{180}$$

$\overline{X}_1$ = पहले भाग का समान्तर माध्य (Arithmetic Mean of the First Group)

$\overline{X}_2$ = दूसरे भाग का समान्तर माध्य (Arithmetic Mean of the Second Group)

उदाहरण 11 : एक फर्म की दो शाखाओं में क्रमश: 100 और 80 व्यक्ति काम करते हैं। यदि दोनों शाखाओं में दिये गये वेतन का समान्तर माध्य 575 रुपये तथा 525 रू. हो तो इस फर्म के सब कर्मचारियों के वेतन का समान्तर माध्य निकालिये।

हल : सामूहिक समान्तर माध्य निम्न सूत्र से ज्ञात करेंगे :-

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2}{N_1 \quad N_2}$$

$N_1 = 100; N_2 = 80; \overline{X}_1 = 575; \overline{X}_2 = 525$

\ $$\overline{X}_{12} = \frac{(100 \quad 575) \quad (80 \quad 525)}{100 \quad 80}$$

= = = 552.78 रु.

यदि हमें एक समूह के तीन भागों का समान्तर माध्य निकालना हो तो सूत्र इस प्रकार होगा :-

$$_{123} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2 \quad N_3\overline{X}_3}{N_1 \quad N_2 \quad N_3} \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2 \quad N_3\overline{X}_3}{N_1 \quad N_2 \quad N_3}$$

समान्तर माध्य के गुण व दोष

(Merits and Demerits of Arithmetic Mean)

गुण (Merits) : समान्तर माध्य के प्रमुख गुण निम्नलिखित हैं :-

(1) सांख्यिकीय माध्यों में समान्तर माध्य सबसे अधिक सरल है। इसका समझना अत्यन्त सुगम है, इसलिए व्यवहार में इसका सबसे अधिक प्रयोग होता है।

(2) समान्तर माध्य न केवल समझने में ही सरल है अपितु इसके निर्धारण की गणन-क्रिया भी बहुत सुगम है।

(3) समान्तर माध्य श्रेणी के सभी मदों पर आधारित होता है। किसी एक मद में परिवर्तन कर देने पर समान्तर माध्य के मूल्य में भी परिवर्तन हो जायेगा। मध्यका व भूयिष्ठक में यह गुण भी नहीं पाया जाता।

(4) सामान्तर माध्य पर निदर्शन के परिवर्तनों का सबसे कम प्रभाव पड़ता है।

(5) सामान्तर माध्य में अनेक बीजगणितीय गुण होते हैं जिनके कारण उच्चतर सांख्यिकीय विश्लेषण में इसको बहुत प्रयोग में आया जाता है।

दोष (Demerits) : समान्तर माध्य के प्रमुख दोष निम्नलिखित हैं :-

(1) समान्तर माध्य पर श्रेणी में असाधारण व सीमान्तर मूल्यों (Extreme Items) का बहुत प्रभाव होता है। उदाहरणार्थ, यदि एक कारखाने के मैनेजर को 5,000 रू. मासिक मिलते हैं और कारखाने में काम करने वाले पांच और व्यक्तियों का वेतन इस प्रकार है : 500 रू., 560 रू., 600 रू., 650 रू., 700 रू. तो 6 व्यक्तियों के वेतन का समान्तर माध्य

$\frac{5{,}000 \quad 500 \quad 560 \quad 600 \quad 650 \quad 700}{6} = \qquad = 1348.33$ रू. होगा लेकिन यह वेतन कारखाने के कर्मचारियों का ठीक प्रतिनिधित्व नहीं करता क्योंकि एक मद अर्थात् 5,000 रू0 का इस पर बहुत गहरा पड़ा है।

(2) अनुपात, दर-प्रतिशत, आदि का अध्ययन करने के लिये समान्तर माध्य उपयुक्त नहीं है।

(3) बिन्दु रेखा द्वारा समान्तर माध्य का प्रदर्शन संभव नहीं।

(4) कभी-कभी इस माध्यम के द्वारा विचित्र परिणाम निकलते हैं; जैसे यदि एक परिवार में 6 बच्चे हैं और दूसरे में 3 तो समान्तर माध्य 4.5 होगा।

8090

(5) समान्तर माध्य से समस्त श्रेणी की रचना को ठीक-ठीक पता नहीं चलता जिससे कभी-कभी गलत निष्कर्ष लिये जाते है। उदाहरणार्थ, दो कारखानों 'क' और 'ख' के पिछले तीन वर्ष के शुद्ध लाभ इस प्रकार हैं :-

वर्ष	कारखाना 'क' (शुद्ध लाभ लाख रुपयों में)	कारखाना 'ख' (शुद्ध लाभ लाख रुपयों में)
1990	80,000	1,90,000
1991	1,20,000	1,20,000
1992	1,90,000	80,000

दोनों कारखानों के लाभों का समान्तर माध्य 1,30,000 है जिससे यह परिणाम निकाला जा सकता है कि दोनों एक ही स्तर पर हैं। लेकिन इन समंकों को ध्यानपूर्वक विश्लेषण करने पर यह स्पष्ट होगा कि कारखाना 'क' शीघ्रता से उन्नति कर रहा है और कारखाना 'ख' अवनति पर है।

सरल समान्तर माध्य का उपयोग : व्यवहार में सरल समान्तर माध्य का सबसे अधिक प्रयोग होता है तथा इसे एक आदर्श माना जाता है। सामाजिक व आर्थिक समस्याओं के विश्लेषण के लिये यह माध्य बहुत उपर्युक्त है जैसे औसत आय, औसत उत्पादन, औसत आयु, औसत आयात, औसत निर्यात, औसत व्यय, आदि। समान्तर माध्य विशेष रूप से ऐसी श्रेणियों के लिये उपयोगी है जिनमें विभिन्न मूल्यों का लगभग समान महत्व होता है।

भारित समान्तर महत्व

(Weighted Arithmetic Mean)

सरल समान्तर माध्य में, जिसका विवेचन ऊपर किया गया है, श्रेणी के सभी मानों को समान महत्व दिया जाता है। व्यवहार में श्रेणी के सभी मूल्यों का समान महत्व नहीं होता, किसी मद का अधिक महत्व होता है और किसी का कम। ऐसी दशा में सरल समान्तर माध्य का प्रयोग उचित नहीं है। भारित समान्तर माध्य वह माध्य होता है जिसमें मदों को उनके सापेक्षिक महत्व

के अनुसार भाग देकर माध्य की मणना की जाती है। भारित समान्तर माध्य का प्रयोग उन परिस्थितियों में किया जाता है जहां श्रेणी से समस्त मानों का महत्व समान नहीं है। भारत समान्तर माध्य ज्ञात करने के लिये निम्न सूत्र का प्रयोग होता है।

$$\overline{X}_w = \frac{WX}{W}$$

वहां $\overline{X}_w$ = भारित समान्तर माध्य (Weighted Arithmetic Mean)

W = भार (Weights)

जैसा की उपर्युक्त सूत्र से विदित होगा भारित समान्तर माध्य का गणनक्रिया बहुत सरल है। गणना करते समय इकाइयों के मूल्य 'X' और उनके भार 'W' को गुणा किया जाता है और मूल्य व भार की गुणांओं का जोड़ 'SWX' निकाल लिया जाता है। इस योग को भारों के योग SW से भाग दिया जाता है जिससे प्राप्त भागफल भारित समान्तर माध्य कहलाता है।

वास्तविक तथा अनुमानित भार (Actual and Estimated Weight) : भारित समान्तर माध्य का प्रयोग करते समय सही भारों का निश्चिय कर लेना आवश्यक है। भार दो प्रकार के होते हैं : वास्तविक और अनुमानित। वास्तविक भार या तो स्पष्ट रूप से दिये होते हैं। उदाहरणार्थ परीक्षाफल की तुलना करने के लिये विद्यार्थी की संख्या को भार के रूप में प्रयुक्त किया जाता है। इसी प्रकार एक कॉलेज के प्राध्यापकों व अन्य कर्मचारियों का औसत वेतन ज्ञात करने के लिए उनकी वास्तविक संख्या को भार मान लिया जाता है। यदि वास्तविक भार ज्ञात न हो सके तो विभिन्न चर मूल्यों के सापेक्षिक महत्व को ध्यान में रखते हुये भारों का उचित अनुमान लगा लिया जाता है। विभिन्न व्यक्तियों द्वारा कल्पित भार भिन्न हो सकते हैं, किन्तु यदि ये अनुमान तर्कयुक्त हैं तो इनके आधार पर निकाले गये समान्तर माध्यों में कोई विशेष अन्तर नहीं होगा।

भारित समान्तर माध्य के प्रयोग (Uses of Weighted Arithmetic Mean) : भार समान्तर माध्य निम्न दशाओं में विशेष रूप से उपयुक्त होता है :-

(1) सूचकांकों के निर्माण में

(2) सामान्य तथा प्रमाणिक मत्यु -दर और जन्म-दर निकालने $\overline{X}$ के लिये; तथा

(3) अनुपातों, प्रतिशत एवं दरों का माध्य निकालते समय।

निम्न उदाहरणों से भारित समान्तर माध्य के प्रयोग स्पष्ट हो जायेंगे।

उदाहरण 12 : एक ठेकेदार तीन प्रकार के मज़दूरों को काम पर लगाता है : पुरुष, स्त्री व बच्चे। एक पुरुष को 30 रू. स्त्री को 20 रू. तथा बच्चे का 10 रू. प्रतिदिन के हिसाब से मजदूरी मिली है। औसत मजदूरी ज्ञात कीजिये।

हल : यदि सरल समान्तर माध्य द्वारा इस प्रश्न को हल किया जाये तो औसत मजदूरी = $\frac{30 \quad 20 \quad 10}{3}$ = 20 रू. होगी।

परन्तु यह उत्तर उस परिस्थिति में ठीक होगा जब पुरुष, स्त्री तथा बच्चों की संख्या समान हो। व्यवहार में प्राय: तीन प्रकार के मजदूरों की संख्या भिन्न-भिन्न होती है, इसलिये सरल समान्तर माध्य के स्थान पर भारित समान्तर माध्य अधिक उपर्युक्त है। इस प्रश्न में वास्तविक मजदूरों की संख्या नहीं दी गई है इसलिये भार कल्पित होंगे। मान लीजिये एक ठेकेदर 20 पुरुष, 15 स्त्री तथा 5 बच्चे काम पर लगाता है। इस दशा में भारित समान्तर इस प्रकार होगा :

मजदूर	प्रतिदिन मजदूरी (रू. में) X	मजदूरों की संख्या W	WX
पुरुष	30	20	600
स्त्री	20	15	300
बच्चे	10	15	50
		SW = 40	**SWX = 950**

$$_w = \frac{WX}{W} = \frac{950}{40} = 23.75 \text{ रू.}$$

उदाहरण 13 : दो विद्यालयों 'क' और 'ख' के निम्नलिखित परीक्षाफलों से ज्ञात

परीक्षा	विद्यालय 'क'		विद्यालय 'ख'	
	परीक्षा में बैठे	उत्तीर्ण हुये	परीक्षा में बैठे	उत्तीर्ण हुये
M.A.	100	90	240	200
M.Sc.	60	45	200	160
B.A.	120	75	160	100
B.Sc.	200	150	200	140

हल : सर्वप्रथम, दोनों विद्यालयों को प्रतिशत पास संख्या ज्ञात कीजिय, तत्पश्चात् विद्यालय 'क' का भारित समान्तर माध्य निकालिये। इसी प्रकार विद्यालय 'ख' का भारित निरन्तर माध्य निकालिये जिसके भार वही होंगे जो विद्यालय 'क' के हैं क्योंकि तुलना का स्तर होना चाहिये।

विद्यालय 'क'					विद्यालय 'ख'			
परीक्षा	परीक्षा में बैठे W (2)	उत्तीर्ण हुये (3)	प्रतिशत फल X (4)	परीक्षा में बैठे WX (5) (2 × 4)	उत्तीर्ण हुये W (6)	प्रतिशत फल (7)	X (8)	WX (9) (2 × 8)
M.A.	100	90	90	9,000	240	200	83.00	8,333
M.Sc.	60	45	75	4,500	200	160	80.00	4,800
B.A.	120	75	62.5	7,500	160	100	62.5	7,500
B.Sc.	200	150	75	15,000	200	140	70.0	14,000
	SW = 480			SWX = 36,000	800	600		SWX=34,633

विद्यालय 'क' :

$$\overline{X}_w = \frac{WX}{W} = \frac{36,000}{480} = 75.0$$

विद्यालय 'ख' :

$$_w = \frac{WX}{W} = \frac{34,633}{480} = 72.15$$

अत: विद्यालय 'क' के परिणाम 'ख' से अच्छे हैं।

उदाहरण 14 : निम्न सारणी तीन विश्वविद्यालयों के परीक्षा परिणाम प्रस्तुत करती है इनमें से कौन-सा विश्वविद्यालय उत्तम है ? उत्तर कारण सहित दीजिये।

परीक्षा	प्रतिशत परिणाम		
	विश्वविद्यालय		
	'क'	'ख'	'ग'
1. एम. ए.	80	75	70
2. एम. एस. सी.	70	70	60
3. एम. कॉम.	65	80	70
4. बी. एस. सी.	60	70	80
5. बी. कॉम.	75	65	75

हल : इस प्रश्न में भारित समान्तर माध्य का प्रयोग उपयुक्त होगा परन्तु विद्यार्थियों की वास्तविक संस्था यहां पर नहीं दी गई इसलिए विद्यार्थियों की संख्या के बारे में कल्पना करनी पड़ेगी। यह कल्पना निम्न आधार पर की गई है :-

(1) एम. ए. की कक्षाओं में बी. ए. की कक्षाओं की अपेक्षा कम विद्यार्थी होते हैं।

(2) विज्ञान के विषयों में कला (Arts) की तुलना में कम विद्यार्थी होते हैं।

(3) विभिन्न विद्यालयों की विभिन्न परीक्षाओं में विद्यार्थियों की संख्या विभिन्न होती है।

	विश्वविद्यालय 'क'			**विश्वविद्यालय 'ख'**			**विश्वविद्यालय 'ग'**		
परीक्षा	**प्रतिशत परिणाम**	**विद्यार्थियों की संख्या (सैकड़ों में)**		**प्रतिशत परिणाम**	**विद्यार्थियों की संख्या (सैकड़ों में)**		**प्रतिशत परिणाम**	**विद्यार्थियों की संख्या (सैकड़ों में)**	
	X	**W**	**WX**	**X**	**W**	**WX**	**X**	**W**	**WX**
एम. ए.	80	3	240	75	2	150	70	2	140
एम. एस. सी.	70	2	140	70	3	210	60	2.5	150
एम. कॉम.	65	6	325	80	6	480	70	7.0	490
बी. एस. सी.	60	4	240	70	3	210	80	4.5	360
बी. कॉम.	75	6	450	65	6	390	75	7.0	525
		SW = 20	**SWX = 1,395**	$\overline{X}$	**SW = 20**	**SWX = 1,440**		**SW = 23**	**SWX = 1,665**

विद्यालय 'क' : $_w = \frac{WX}{W} = \frac{1,396}{20} = 69.75$

विद्यालय 'ख' : $_w = \frac{WX}{W} = \frac{1,440}{20} = 72.00$

विद्यालय 'ग' : $_w = \frac{WX}{W} = \frac{1,665}{23} = 72.39$

क्योंकि विश्वविद्यालय 'ग' की औसत सबसे अधिक है इसलिये हम कह सकते हैं कि विश्वविद्यालय 'ग' सर्वश्रेष्ठ है।

उदाहरण 15 : किसी विश्वविद्यालय में छात्रवत्ति (Scholarship) प्रदान करने के लिये एक परीक्षा ली गई है। विभिन्न विषयों को अलग-अलग भाग दिये गये। चार विद्यार्थियों द्वारा प्राप्त अंक इस प्रकार थे :-

विषय	**भार**	**विद्यार्थी 'क'**	**विद्यार्थी 'ख'**	**विद्यार्थी 'ग'**	**विद्यार्थी 'घ'**
सांख्यिकी	4	70	80	85	92
अर्थशास्त्र	3	90	75	75	40
व्यवसाय-व्यवस्था	1	50	60	45	70
मुद्रा, बैंकिग व व्यापार	2	60	45	65	50

यदि छात्रवत्ति इस विद्यार्थी को दी जाये जिसने सबसे अधिक अंक प्राप्त किये हों तो उपर्युक्त समंकों के अनुसार छात्रवत्ति किसे दी जानी चाहिये।

हल : **चारों विद्यार्थियों द्वारा प्राप्त समंकों का भारित समान्तर माध्य**

विषय	भार	'क' के अंक		'ख' के अंक		'ग' के अंक		'घ' के अंक	
		X	**WX**	**X**	**WX**	**X**	**WX**	**X**	**WX**
सांख्यिकी	4	70	280	80	320	85	340	92	368
अर्थशास्त्र	3	90	270	75	225	75	225	40	120
व्यवसाय-व्यवस्था	1	50	50	60	60	45	45	70	70
मुद्रा, बैंकिग व व्यापार	2	60	120	45	90	65	130	50	100
			SWX = 720		**SWX = 695**		**SWX = 740**		**SWX = 658**

विद्यालय 'क' : $_w = \frac{720}{10} = 72.0$

विद्यालय 'ख' : $\overline{X}_w = \frac{695}{10} = 69.5$

विद्यालय 'ग' : $\overline{X}_w = \frac{740}{10} = 74.0$

विद्यालय 'घ' : $\overline{X}_w = \frac{658}{10} = 65.8$

क्योंकि विद्यार्थी 'ग' के औसत अंक सबसे अधिक है इसलिये विद्यार्थी 'ग' को छात्रवत्ति मिलनी चाहिये।

सामान्य व प्रमापित मत्यु -दरें

(Crude and Standarised Death Rate)

दो नगरों की औसत मत्यु -दरों की तुलना करने के लिये भारित समान्तर का माध्य प्रयोग किया जाता है और मत्यु -दर दो प्रकार की होती है :-

(1) सामान्य मत्यु -दर Crude of General Death Rate)

(2) प्रमापित मत्यु -दर (Standardised Death Rate)

मत्यु -दरें प्रति हजार (Per Thousand0 में व्यक्त की जाती हैं।

लेकिन दो नगरों की औसत मत्यु -दरों की तुलना करने के लिये सामान्य मत्यु दर लेना ठीक नहीं है क्योंकि विभिन्न आयु -श्रेणियों में व्यक्तियों की संख्या तथा मत्यु -दर भिन्न होगी। ऐसी स्थिति में प्रभावित मत्यु -दर (Standardised Death Rate) का प्रयोग किया जाता हैं दो नगरों की औसत मत्यु -दर की तुलना करने में एक प्रमापित नगर की जनसंख्या को दोनों नगरों के लिये भार मान लिया जाता है। स्थानीय नगर की अलग-अलग प्रति हजार मत्यु -दर को प्रभावित नगर की जनसंख्या से भाग देने पर जो भारित समान्तर माध्य दर ज्ञात होती है वह स्थानीय नगर की प्रमापित या संशोधित मत्यु -दर कहलाती है। सूत्र के रूप में:-

$$\text{S.D.R. (प्रमापित मत्यु -दर)} = \frac{WX}{W}$$

जहां W = प्रमापित जनसंख्या (Standard Population); X = विशेष मत्यु -दर (Specific Death Rate)

स्थानीय नगर की प्रमापित मत्यु -दर और प्रमापित नगर की सामान्य मत्यु दर की आपस में तुलना करके दोनों नगरों की स्वास्थ्य संबंधी का पता लग जाता है क्योंकि दोनों मत्यु -दरों की गणन क्रिया में एक ही नगर की जनसंख्या को भार मानते है जिस नगर की औसत मत्यु -दर कम होती है वह अधिक स्वस्थ माना जाता है। यदि दिये हुए समंकों से यह ज्ञात न हो कि प्रमापित नगर कौन-सा है तो पहले नगर को ही प्रमापित मान लिया जाता है।

विवाह-दरों, जन्म-दरों, बेरोजगारी-दरों, परीक्षाफलों की प्रतिशत आदि के तुलनात्मक अध्ययन में सामान्य एवं प्रमापित दरों के सिद्धान्त का प्रयोग किया जाता है।

उदाहरण 16 : निम्न समंकों के आधार पर यह निर्णय कीजिये कि कौन से शहर के व्यक्ति अधिक स्वस्थ हैं ?

आयु वर्षों में	प्रमापित जनसंख्या शहर 'क'		स्थानीय जनसंख्या शहर 'ख'	
	जनसंख्या	मतकों की संख्या	जनसंख्या	मतकों की संख्या
5 से कम	15,000	360	40,000	1,000
5-30	20,000	400	52,000	1,040
30 से अधिक	10,000	280	8,000	240

हल : शहर 'क' की सामान्य मत्यु -दर तथा शहर 'ख' की प्रमापित मत्यु -दर की परिगणना

आयु (वर्षों में) (1)	प्रमापित जनसंख्या शहर 'क' जनसंख्या W (2)	मतकों की संख्या (3)	मत्यु दर प्रति हजार X (4)	WX (5) (2 × 4)	स्थानीय जनसंख्या शहर 'ख' जनसंख्या X (6)	मतकों की संख्या (7)	मत्यु दर प्रति हजार X (8)	WX (9) (2 × 8)
5 से कम	15,000	360	24	3,60,000	40,000	1,000	35	3,75,000
5-30	20,000	400	20	4,00,000	52,000	1,040	20	4,00,000
30 से अधिक	10,000	280	28	2,80,000	8,000	240	30	3,00,000
	SW = 45.000			**SWX = 10,40,000**				**SWX = 10,75,000**

शहर 'क' सामान्य मत्यु -दर $= \frac{WX}{W} = \frac{10,40,000}{45,000} = 23.11$

शहर 'ख' सामान्य मत्यु -दर $= \frac{WX}{W} = \frac{10,75,000}{45,000} = 23.89$

उदाहरण 17 : जनसंख्या और बेरोजगारी के निम्नलिखित समंकों से दोनों जनसंख्याओं के लिये सामान्य बेरोजगारी दर एवं स्थानीय जनसंख्या के लिये प्रमापित बेरोजगारी पर निकालिये तथा बताइये कि किस जनसंख्या में रोजगार के अच्छी स्थिति पाई जाती है।

आयु वर्षों में	प्रमापित जनसंख्या		स्थानीय जनसंख्या	
	जनसंख्या	बेरोजगार	जनसंख्या	बेरोजगार
15-30	2,000	125	3,000	120
30-45	3,500	280	3,000	270
45-60	3,000	360	3,500	420
60 से अधिक	1,000	150	500	100

30-45	3,500	280	3,000	270
45-60	3,000	360	3,500	420
60 से अधिक	1,000	150	500	100

हल :

आयु (वर्षों में)	प्रमापित जनसंख्या				स्थानीय जनसंख्या			
	जनसंख्या	बेरोजगार	बेरोजगारी		जनसंख्या	बेरोजगार	बेरोजगारी	
	W		**दर X**	**WX**	**X**		**दर X**	**WX**
(1)	**(2)**	**(3)**	**(4)**	**(5)** **(2 × 4)**	**(6)**	**(7)**	**(8)**	**(9)** **(2 × 8)**
15-30	2,000	125	62.5	1,25,000	3,000	120	40	80,000
30-45	3,500	280	80.0	2,80,000	3,000	270	90	3,15,000
45-60	3,000	360	120.0	3,60,000	3,500	420	120	3,60,000
60 से अधिक	1,000	150	150.0	1,50,000	500	100	200	2,00,000
	SW = 9,500			**SWX = 9,15,000**				**SWX = 9,55,000**

प्रमापित जनसंख्या : सामान्य बेरोजगारी दर $= \dfrac{WX}{W} = \dfrac{9{,}15{,}000}{9{,}500} = 93.62$

स्थानीय जनसंख्या : प्रमापित बेरोजगारी दर $= \dfrac{WX}{W} = \dfrac{9{,}55{,}000}{9{,}500} = 100.53$

इससे यह स्पष्ट है कि प्रमापित जनसंख्या में रोजगार की अच्छी दशा पायी जाती है।

मध्यका
(Median)

किसी समंक श्रेणी को आरोही (चढ़ते हुए) या अवरोही (गिरते हुए) क्रम में व्यवस्थित करने पर उस श्रेणी के मध्य में जो मूल्य आता है उसे मध्यका कहते हैं। अर्थात् मध्यका स्थिति का माध्य है। कौनर के शब्दों में, ''मध्यका समंक श्रेणी का वह चर-मूल्य है जो समूह को दो बराबर भागों में इस प्रकार बांटता है कि एक भाग में सारे मूल्य मध्यका से अधिक और दूसरे भाग में उससे कम हों।''

उदाहरण के लिए, यदि पांच व्यक्तियों की आय 100, 200, 150, 160, 180 रुपये है तो मध्यका 150 होगा अर्थात् दो व्यक्ति ऐसे हैं जिनकी आय 150 रुपये कम है और व्यक्ति ऐसे हैं जिनकी आय 150 रुपये से अधिक; इस प्रकार मध्यका श्रेणी के बिल्कुल बीच में स्थित होता है और वह मूल्य श्रेणी को दो बराबर भागों में बांट देता है — मध्यका से पहले तथा बाद की आवत्तियां सदा समान रहती हैं। यदि श्रेणी के मदों की संख्या सम या युग्म (Even) है तो उसमें कोई भी मूल्य बीच में नहीं होगा। ऐसी स्थिति में मध्यका निकालने के लिये बीच के दो मूल्यों का औसत निकाल लेते हैं। उदाहरणार्थ, यदि 6 व्यक्तियों की आय 100, 120, 150, 170, 180 तथा 250 रुपये है तो मध्यका 150 और 170 के बीच अर्थात् 160 होगा।

मध्यका का निर्धारण-व्यक्तिगत श्रेणी में
(Calculating Median in Individual Series)

व्यक्तिगत मूल्यों में मध्यका ज्ञात करने के लिये निम्न क्रियायें की जाती हैं :-

(1) सर्वप्रथम, दिये गये मूल्यों को आरोही या अवरोही क्रम में अनुविन्यसित (Arrange) किया जाता है। दोनों क्रमों के अनुसार केन्द्र बिन्दु एक ही होता है। मूल्यों की क्रम संख्यायें (Serial Numbers) भी साथ-साथ लिख देनी चाहिये।

(2) मूल्यों को क्रमबद्ध करने के पश्चात् निम्न सूत्र का प्रयोग किया जाता है।

उदाहरण 18 : 7 कर्मचारियों को दिया हुआ वेतन इस प्रकार है :-

वेतन (रुपयों में) : 1500 1550 1680 1690 1660 1700 1640

मध्यका निकालिये।

हल : **मध्यका परिगणन**

क्रम संख्या में अनुविन्यसित	वेतन आरोही क्रम में अनुविन्यसित	क्रम संख्या	वेतन आरोही क्रम
1	1500	5	1680
2	1550	6	1690
3	1640	7	1700
4	1660		

$$M = \text{Size of } \frac{N \quad 1}{2} \text{th item} = \quad = 4\text{th item}$$

चौथे पद का मूल्य = 1660

अर्थात् मध्यका मूल्य = 1660 रूपये

उदाहरण 19 : निम्न समंकों से मध्यका मूल्य ज्ञात कीजिये :-

10 व्यक्तियों के वेतन (रूपयों में) : 1391 1384 1591 1407 1672 1522 1777 1733 2488 2390

हल **मध्यका का परिगणन**

क्रम संख्या में अनुविन्यसित	समंक आरोही क्रम में अनुविन्यसित	क्रम संख्या	समंक आरोही क्रम
1	1384	6	1672
2	1391	7	1733
3	1407	8	1777
4	1522	9	2390
5	1591	10	2400

$$M = \text{Size of } \quad \text{th item} = \text{Size of } \quad = 5.5\text{th item}$$

$$\text{Size of 5.5th item} = \frac{5\text{th item} + 6\text{th item}}{2} = \frac{1591 \quad 1672}{2} = 1631.5$$

अर्थात् माध्यका मूल्य = 1631.5 रुपये।

विच्छिन्न श्रेणी में मध्यका का निर्धारण

(Calculating Median in Discrete Series)

विच्छिन्न आवत्ति वितरण में मध्यका ज्ञात करने की विधि निम्नलिखित हैं :-

(1) सर्वप्रथम, दिये हुये मूल्यों को आरोही या अवरोही क्रम में अनुविन्यासित किया जाता है।

(2) संचयी आवत्तियां (Cumulative Frequencies) निकालकर श्रेणी को संचयी आवत्ति माला में बदल लिया जाता है।

(3) निम्न सूत्र द्वारा मध्यका की क्रम संख्या ज्ञात कर ली जाती है।

M = Size of th item; जहां N आवत्तियों का योग

(4) मध्यका की क्रम संख्या का मूल्य आवत्ति की सहायता से ज्ञात कर लिया जाता है। जिस संचयी आवत्ति में यह क्रम - संख्या प्रथम बार सम्मिलित होती है उसका मूल्य ही मध्यका होता है।

उदाहरण 20 : निम्न समंकों से मध्यका ज्ञात कीजिये :-

आयु :	20	21	22	23	24	25	26	27	28
कर्मचारियों की संख्या :	8	10	11	16	20	25	15	9	6

हल : **मध्यका का परिगणन**

आयु	आवत्ति	संचयी आवत्तियां
20	8	8
21	10	18
22	11	29
23	16	45
24	20	65
25	25	90
26	15	105
27	9	114
28	6	120

M = Size of th item = Size of = 60.5th item

Size of 60.5 item = 24 अर्थात मध्यका आयु = 24 वर्ष

मध्यका का निर्धारण—अविच्छिन्न श्रेणी में

(Calculating Median in Continuous Series)

अविच्छिन्न श्रेणी में मध्यका ज्ञात करने की प्रणाली निम्नलिखित है :-

(1) सर्वप्रथम, संचयी आवत्तियां ज्ञात की जाती हैं।

(2) निम्न सूत्र द्वारा केन्द्रीय मद ज्ञात किया जाता है।

M = Size of th item

अविच्छिन्न श्रेणी में मध्यका $\frac{N}{2}$th item का ही मूल्य होता $\frac{N \quad 1}{2}$th item का नहीं। यद्यपि कुछ लेखकों ने अविच्छिन्न श्रेणी में का प्रयोग किया है लेकिन ऐसा करना ठीक प्रतीत नहीं होता क्योंकि अविच्छिन्न श्रेणी में आवत्ति वक्र के क्षेत्रफल को ठीक दो भागों में विभाजित करता है $\frac{N \quad 1}{2}$ नहीं।

(3) जिस संचयी आवत्ति में मध्यका की संख्या सबसे पहली बार आती है उससे सम्बन्धित वर्ग को ले लिया जाता है। यह वह वर्ग होता है जिससे मध्यका का ठीक मूल्य निकाला जाता है। इसकी निम्न व उच्च सीमाओं के अंतर्गत ही कही मध्यका होगी।

(4) मध्यका-वर्ग से मध्यका का मूल्य निर्धाति करने के लिये निम्न सूत्र का प्रयोग किया जाता है।

$$M = L + \qquad \times i$$

जहाँ L = मध्यका-वर्ग की निम्न सीमा (Lower Limit of Median Class),

c.f. = मध्यका-वर्ग के पूर्व वाले वर्ग को संचयी आवत्ति (Cumulative Frequency of the Class Preceding the Median Class),

f = मध्यका-वर्ग की आवत्ति (Frequency of the Median Class),

i = मध्यका-वर्ग का वर्गान्तर (Class Interval of the Median Class);

टिप्पणी (Note) : अविच्छिन्न श्रेणी में मध्यका-मूल्य ज्ञात करते समय यह कल्पना की जाती है कि प्रत्येक वर्ग की इकाइयों का उसके पूरे वर्गान्तर में समानरूप से वितरण हुआ है।

उदाहरण 21 : निम्न समंकों से मध्यका तथा भूयिष्ठक ज्ञात कीजिये :-

अनुपस्थित दिनों की संख्या	विद्यार्थियों की संख्या	अनुपस्थित दिनों की संख्या	विद्यार्थियों की संख्या
5 से कम	29	30 से कम	644
10 से कम	224	35 से कम	650
15 से कम	465	40 से कम	653
20 से कम	682	45 से कम	655
25 से कम	634	$\frac{N/2 \quad c.f.}{2 \quad f}$	

हल : **मध्यका का परिगणन**

अनुपस्थित दिनों की संख्या	विद्यार्थियों की संख्या *f*	*c.f.*
0-5	29	29
5-10	195	224
10-15	241	465
15-20	117	582
20-25	52	634
25-30	10	644
30-35	6	650
35-40	3	653
40-45	2	655

$$M = \text{Size of} \quad \text{th item} = \frac{655}{2} = 327.5\text{th item}$$

अर्थात मध्यका-वर्ग 15-20 है।

L = 10, N/2 = 327.2, *c.f.* = 224, *f* = 241, *i* = 5

$$\text{Med.} = 10 + \frac{327.5 \quad 224}{241} \times 5 = 10 + 2.15 = 12.15$$

उदाहरण 22 : निम्न समंकों से मध्यका ज्ञात कीजिये :-

वर्ग अन्तराल :	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
बारम्बारता :	4	16	56	97	124	137	146	150

हल : हमें संचयी आवत्ति विवरण दिया है पहले साधारण आवत्ति विवरण बनायें :-

वर्ग अन्तराल	*f*	*c.f.*
10-20	4	4
20-30	12	16
30-40	40	56
40-50	41	97
50-60	27	124
60-70	13	137
70-80	9	146
80-90	4	150

$$\text{Med.} = \text{Size of} \quad \text{th item} = \text{Size of } \frac{150}{2} = 75\text{th item}$$

अर्थात् मध्यका-वर्ग 40-50 है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 40, N/2 = 75, *c.f.* = 56, *f* = 41, *i* = 10

$$= 40 + \frac{75 \quad 56}{41} \times 10 = 40 + 4.63 = 44.63$$

मध्यका के गुण व दोष

(Merits and Demerits of Median)

गुण (Merits) :

(1) गुणात्मक तथ्यों (Qualitative Facts) जैसे ईमानदारी, बुद्धिमता, क्षमता आदि का माध्य ज्ञात करने के लिये मध्यका सर्वोत्तम मानी जाती है।

(2) मध्यका को समझना और ज्ञात करना बहुत सरल है।

(3) मध्यका पर चरम मूल्यों या सीमान्त मदों को कोई प्रभाव नहीं पड़ता।

(4) खुले-सिरे वाली सारणी में मध्यका निकालने के लिये प्रथम वर्ग की निम्नतम सीमा तथा अन्तिम वर्ग की उच्चतम सीमा निश्चित करना आवश्यक नहीं जबकि मध्यका निकालते समय ये सीमायें निश्चित करनी पड़ती हैं।

(5) रेखा-चित्र खींचकर मध्यका मूल्य निर्धारित किया जा सकता है जबकि मध्यक में ऐसा सम्भव नहीं है।

(6) माध्यका एक स्पष्ट और निश्चिम माध्य है; भूयिष्ठक की भांति अनिश्चित नहीं है।

दोष (Demerits) :

(1) मध्यका में बीजगणितीय गुणों का अभाव है इसलिये उच्चतर गणितीय क्रियाओं में इसका प्रयोग नहीं किया जा सकता। उदाहरणार्थ, मध्यका मूल्य और मदों की संख्या को गुणा करने से सभी मदों के मूल्यों का जोड़ ज्ञात नहीं किया जा सकता।

(2) मध्यका-मूल्य निर्धारित करने से पूर्व मदों को आरोही या अवरोही क्रम में अनुविन्यसति करना पड़ता है जिसमें काफी समय लगता है।

(3) अविच्छिन्न श्रेणी में मध्यका ज्ञात करते समय मान्यता की जाती है कि प्रत्येक वर्ग में आवत्तियां समान रूप से वितरित है लेकिन यह मान्यता सदैव सत्य नहीं होती।

(4) मध्यका-मूल्य निकालते समय श्रेणी के सभी मदों को समान महत्व दिया जाता है।

विभाजन-मूल्य
(Partition Values)

ऊपर कुछ पष्ठों में मध्यका का वर्णन किया गया है जो श्रेणी को दो बराबर भागों में बांटता है। माध्यका के सिद्धान्त के आधार पर समंकमाल को चार, पांच, आठ, दस या सौ बराबर भागों में बांटा जा सकता है। समंक श्रेणी को अनेक भागों में बांटने वाले मूल्यों के विभाजन-मूल्य (Partition Values) कहते हैं। चार बराबर भागों में बांटने वाले मूल्य चतुर्थक (Quartiles) पांच, भागों में बांटने वाले मूल्य पंचमक (Quantiles) आठ भागों में बांटने वाले अष्टमक (Octiles) दस भागों में बांटने वाले मूल्य दशमक (Deciles) तथा सौ बराबर भागों में बांटने वाले मूल्य शतमक (Percentile) कहलाते हैं। जिस प्रकार एक बिन्दु श्रेणी को दो बराबर भागों में बांटता है उसी प्रकार तीन बिन्दु चार भागों में, 9 बिन्दु दस भागों में बांटेंगे, आदि, अर्थात् एक समंकमाला में 3 चतुर्थक, 7 अष्टमक, 9 दशमक तथा 99 शतमक होते हैं। मध्यका-मूल्य द्वितीय चतुर्थक, चौथे अष्टमक, 5वें दशमक तथा 50 वें शतमक के मूल्य बराबर होता है। (Med. = $Q_2 = Q_4 = D_5 = P_{50}$)। विभाजन मूल्यों में चतुर्थक तथा शतमक अधिक महत्वपूर्ण है। इनसे समंक श्रेणी की रचना का आभास हो जाता है। अपकिरण (Dispersion) तथा विषमता (Skewness) ज्ञात करते समय इन मूल्यों का प्रयोग होता है।

विभाजन मूल्यों का निर्धारण (Determination of Partition Values) : मध्यका नियम के आधार पर ही विभाजन-मूल्यों का निर्धारण किया जाता है व्यक्तिगत तथा विच्छिन्न श्रेणी में इन मूल्यों को ज्ञात करने के लिये (N + 1) को उस अंक से भाग दिया जाता है जितने भागों में मूल्य श्रेणी को बांटते हैं। इस प्रकार जो मद संख्या निश्चित होती है उसका मूल्य समंक श्रेणी में देख लिया जाता है। अविच्छिन्न श्रेणी में $\left(\frac{N \quad 1}{2}\right)$ के स्थान पर $\frac{N}{2}$ का प्रयोग होता है।

Q_1 = Size of $\left(\frac{N \quad 1}{4}\right)$th item – व्यक्तिगत तथा विच्छिन्न श्रेणी में

Q_2 = Size of $2\left(\frac{N \quad 1}{4}\right)$the item – व्यक्तिगत तथा विच्छिन्न श्रेणी में

Q_3 = Size of $3\left(\frac{N \quad 1}{4}\right)$th item – व्यक्तिगत तथा विच्छिन्न श्रेणी में

Q_1 = Size of $\frac{N}{4}$th item अविच्छिन्न श्रेणी में

Q_3 = Size of $\frac{3N}{4}$th item अविच्छिन्न श्रेणी में

Q_1 = Size of $\left(\frac{N \quad 1}{4}\right)$th item – व्यक्तिगत तथा विच्छिन्न श्रेणी में

D_4 = Size of $4\left(\frac{N+1}{10}\right)$th item – व्यक्तिगत तथा विच्छिन्न श्रेणी में

$D_6 = \text{Size of } \frac{6N}{10} \text{th item}$ – अविच्छिन्न श्रेणी में

$P_{20} = \text{Size of } 20\left(\frac{N+1}{100}\right) \text{th item}$ – व्यक्तिगत तथा विच्छिन्न श्रेणी में

$D_6 = \text{Size of } \frac{60N}{100} \text{th item}$ – अविच्छिन्न श्रेणी में

निम्न उदाहरण द्वारा विभाजन-मूल्यों की गणन-क्रिया स्पष्ट हो जायेगी।

उदाहरण 23 : (*i*) निम्न समंकों से प्रथम व ततीय चतुर्थक (Q_1 तथा Q_3) के मूल्य ज्ञात कीजिये।

7 व्यक्तियों के वेतन (रुपयों में) : 600 620 700 580 560 650 780

हल : सर्वप्रथम, इन समंकों को आरोही क्रम में व्यवस्थित कीजिये।

क्रम संख्या	वेतन (रुपयों में)	क्रम संख्या	वेतन (रुपयों में)
1	560	5	650
2	580	6	700
3	600	7	780
4	620		

$Q_1 = \text{Size of } \frac{N+1}{4} \text{th item} = \frac{7+1}{4} = \text{2nd item}$

$= \text{Size of 2nd item} = 580$ अर्थात् $Q_1 = 580$ रू.

$\frac{\text{5th item + 6th item}}{2}$

$Q_3 = \text{Size of } 3\left(\frac{N+1}{4}\right) \text{th item} = \frac{3 \quad 8}{4} = \text{6th item}$

$= \text{Size of 6th itme} = 700$ अर्थात् $Q_3 = 700$ रू.

(*ii*) निम्न समंकों से मध्यका, प्रथम व ततीय चतुर्थक , आठवां दशमक तथा बीसवां शतमक ज्ञात कीजिये।

10 विद्यार्थियों द्वारा प्राप्त अंक : 32 12 30 18 11 41 35 26 22 14

हल : सर्वप्रथम, दिये हुए समंकों को आरोही क्रम में व्यवस्थित कीजिये :-

क्रम संख्या	अंक	क्रम संख्या	अंक
1	11	6	26
2	12	7	30
3	14	8	32
4	18	9	35
5	22	10	40

$M = \text{Size of} \qquad \text{th item} = \qquad = \text{5.5th item}$

$\text{Size of 5.5th item} = \qquad = \frac{22 \quad 26}{2} = 24$ अंक

$Q_1 = \text{Size of} \qquad \text{th item} = \qquad = \text{2.75 item}$

$$\text{Size of 2.75th item} = \text{2nd item} + 0.75\,(\text{Size of 3rd item} - \text{Size of 2nd item})$$
$$= 12 + 0.75\,(14 - 12) = 13.5 \text{ अंक}$$

$$Q_1 = \text{Size of } 3 \quad \text{th item} = \frac{3 \quad 11}{4} = 8.25 \text{ item}$$

$$\text{Size of 8.25th item} = \text{8th item} + 0.25\,(\text{Size of 9th item} - \text{Size of 8th item})$$
$$= 32 + 0.25\,(35 - 32) = 32.75 \text{ अंक}$$

$$D_8 = \text{Size of } 8 \quad \text{th item} = \frac{8 \quad 11}{10} = 8.8\text{th item}$$

$$\text{Size of 8.8th item} = \text{8th item} + 0.8\,(\text{Size of 9th item} - \text{Size of 8th item})$$
$$= 32 + 0.8\,(35 - 32) = 34.4 \text{ अंक}$$

$$P_{20} = \text{Size of } 20 \quad \text{th item} = \frac{20 \quad 11}{100} = 2.2\text{th item}$$

$$\text{Size of 2.2th item} = \text{2nd item} + 0.2\,(\text{Size of 3rd item} - \text{Size of 2nd item})$$
$$= 12 + 0.2\,(14 - 12) = 12 + 0.4 = 12.4 \text{ अंक}$$

उदाहरण 24 : निम्न समंकों से प्रथम ततीय , चतुर्थक, छठे, दशमत तथा 70वें शतमक के मूल्य ज्ञात कीजिये :-

अंक	:	10	20	30	40	50	60
विद्यार्थियों की संख्या	:	4	10	20	8	6	3

हल : **Q_1, Q_3, D_6 तथा P_{70} की मूल्यों का परिगणन**

अंक	विद्यार्थियों की संख्या f	संचयी आवत्ति
10	4	4
20	10	14
30	20	34
40	8	42
50	6	48
60	3	51

$$Q_1 = \text{Size of} \quad \text{th item} = \frac{51 \quad 1}{4} = 13\text{th item}$$

Size of 13th item = 20 अर्थात $Q_1 = 20$ अंक

$$Q_3 = \text{Size of } 3 \quad \text{th item} = \frac{3 \quad 52}{4} = 39\text{th item}$$

Size of 39th item = 40 अर्थात् $Q_3 = 40$ अंक

$$D_6 = \text{Size of } 6 \quad \text{th item} = \frac{6 \quad 52}{10} = 31.2\text{th item}$$

Size of 31.2th item = 30 अर्थात् $D_6 = 30$ अंक

$$P_{70} = \text{Size of } 70 \qquad \text{th item} = \frac{70 \quad 52}{100} = 36.4\text{th item}$$

$$\text{Size of } 36.4\text{th item} = 40\text{th अर्थात्} \; P_{70} = 40 \text{ अंक}$$

उदाहरण 25 : निम्न समंकों से प्रथम तथा ततीय चतुर्थक , 6वें दशमक (D_6) तथा 80वें शतमक (P_{80}) के मूल्य ज्ञात कीजिये:-

अंक	आवत्ति	अंक	आवत्ति
0-10	10	40-50	90
10-20	25	50-60	40
20-30	35	60-70	30
30-40	70		

हल : **Q_1, Q_3, D_6 तथा P_{80} के मूल्यों का परिगणन**

अंक	विद्यार्थियों की संख्या f	संचयी आवत्ति *c.f.*
0-10	10	10
10-20	25	35
20-30	35	70
30-40	70	140
40-50	90	230
50-60	40	270
60-70	30	300

प्रथम चतुर्थक

$$Q_1 = \text{Size of} \quad \text{th item} = \frac{300}{4} = 75\text{th item}$$

अर्थात् Q_1 30-40 के वर्ग में स्थित है।

$$Q_1 = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 30, N/4 = 75; *c.f.* = 70; *f* = 70, *i* = 10

$$Q_1 = 30 + \qquad \times 10 = 30 + 0.714 = 30.714$$

$$Q_3 = \text{Size of } \frac{3N}{4} \text{th item} = \frac{3 \quad 300}{4} = 225\text{th item}$$

अर्थात् तुतीय चतुर्थक वर्ग = 40-50 है।

$$Q_3 = L + \qquad \times i$$

L = 40, 3N/4 = 225; *c.f.* = 140; *f* = 90; *i* = 10

\ $$Q_1 = 40 + \quad \times 10 = 40 + 9.45 = 48.45$$

$$D_6 = \text{Size of (6N/10th) item} = \quad = 180\text{th item}$$

अर्थात् D_6 40-50 वर्ग में स्थित है।

$$D_6 = L + \quad \times i$$

L = 40, 6N/10 = 180; *c.f.* = 140; *f* = 90; *i* = 10

$$D_6 = 40 + \frac{180 \quad 140}{90} \times 10 = 40 + 4.44 = 44.44$$

$$P_{80} = \text{Size of} \quad \text{th item} = \frac{8 \quad 300}{100} = 240\text{th item}$$

अर्थात् P_{80} वर्ग 50-60 में स्थित है।

$$P_{80} = L + \quad \times i$$

L = 50, 80N/100 = 240; *c.f.* = 230; *f* = 40; *i* = 10

$$P_{80} = 50 + \frac{240 \quad 230}{40} \times 10 = 50 + 2.5 = 52.5$$

बिन्दुरेखीय पद्धति द्वारा मध्यका तथा अन्य विभाजन मूल्यों का निर्धारण

(Determination of Median, Quartiles, etc. Graphically)

बिन्दुरेखीय पद्धति द्वारा मध्यका की गणना दो प्रकार से की जाती है :'

(1) बिन्दुरेखीय पत्र पर दो 'आगिव वक्र' (Ogive Curves) एक ऊपरी सीमाओं तथा बढ़ती हुई संचयी आवत्ति के आधार पर ('Less Than' Ogive) तथा दूसरी निचली सीमाओं और घटती संचयी आवत्तियों के आधार पर ('More Than' Ogive) बनाइये जिस बिन्दु पर ये दोनों वक्र मिलते हैं वहां से एक लम्ब (Perpendicular) X-अक्ष पर डालिए। यह लम्ब जिस स्थान पर X-अक्ष से मिलता है वही मध्यका मूल्य है।

(2) एक ही ओगिव वक्र बिन्दुरेखीय पत्र पर दिखाई जाती है जिससे मध्यका ज्ञात जाता है। चर को X-अक्ष पर तथा आवत्तियों को Y-अक्ष पर लिया जाता है। N/2 के आधार पर मध्यका की संख्या निकाल ली जाती है। इस मूल्य को Y-अक्ष पर देखा जाता है और वहां से 'ओजाइव वक्र' पर एक लम्ब डाला जाता है। जिस बिन्दु पर यह लम्ब 'ओगिव वक्र' को मिलाता है वहां से एक लम्ब X-अक्ष पर डाला जाता है। इस लम्ब के स्पर्श-बिन्दु का मूल्य पढ़ लिया जाता है। यही मूल्य मध्यका-मूल्य होगा। इस पद्धति द्वारा चतुर्थक, दशमक तथा अन्य विभाजन मूल्य भी ज्ञात किये जा सकते हैं। निम्न उदाहरण द्वारा ये दोनों विधियां स्पष्ट हो जायेगी :-

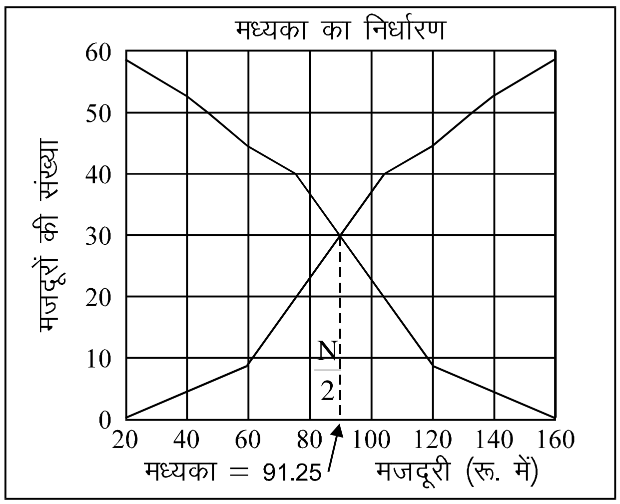
उदाहरण 26 : निम्न समंकों से बिन्दुरेखीय पद्धति द्वारा माध्यका ज्ञात कीजिये :-

वेतन (रुपयों में)	मजदूरों की संख्या *f*	वेतन (रुपयों में)	मजदूरों की संख्या *f*
20-40	4	100-120	12
40-60	6	120-140	7
60-80	10	140-160	3
80-100	16		

हल : **प्रथम रीति-मध्यका का निर्धारण**

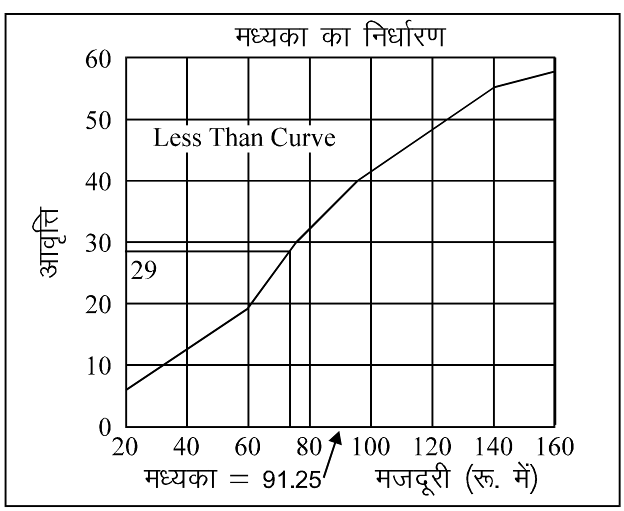
मजदूरी इतने रू. से कम	मजदूरों की संख्या	मजदूरी इतने रू. से अधिक	मजदूरों की संख्या
40	4	20	58
60	10	40	54
80	20	60	48
100	36	80	38
120	48	100	22
140	55	120	10
160	58	140	3

बिन्दुरेखीय पत्र पर इन समंकों को इस प्रकार प्रस्तुत करेंगे :-



दूसरी रीति : $M = \text{Size of} \quad \text{th item} = \frac{58}{2} = 29\text{th item}$

एक ही वक्र 'Less Than' Method से बनाई जायेगी और आवत्ति वाले अक्ष पर अर्थात Y-अक्ष पर 29 से एक लम्ब इस वक्र पर डाला जायेगा। जिस जगह यह लम्बाई इस वक्र को काटेगा वहां से एक और लम्ब X-अक्ष पर डालेंगे। यह लम्ब जिस स्थान पर X-अक्ष में मिलेगा वहीं मध्यका मूल्य होगा जैसा कि निम्न चित्र से विदित है :-



उदाहरण 27 : सांख्यिकी की एक परीक्षा में 60 विद्यार्थियों द्वारा प्राप्त अंक इस प्रकार है :-

75	86	66	86	50	66	78	79	68	60	80	83
87	79	80	77	81	92	57	52	58	82	72	95
66	60	84	80	79	62	80	88	58	84	96	87
72	65	79	80	86	68	76	41	80	40	63	90
83	94	76	66	74	76	68	42	59	75	35	80

एक आवत्ति वितरण बनाइये जिसमें वर्ग 30-39, 40-49 आदि हों। इस आवत्ति से एक वक्र बनाइयें और मध्यका ज्ञात कीजिये।

हल : **60 विद्यार्थियों द्वारा प्राप्त अंकों का आवत्ति वितरण**

अंक	मिलान रेखायें	आवत्ति	संचयी आवत्ति
30-39	I	1	1
40-49	III	3	4
50-59	卌 I	6	10
60-69	卌 卌 II	12	22
70-79	卌 卌 IIII	14	36
80-89	卌 卌 卌 IIII	19	55
90-99	卌	5	60
		योग 60	

$$M = \text{Size of } \frac{N}{2}\text{th item} = \frac{60}{2} = 30\text{th item}$$

निम्नांकित चित्र द्वारा यह स्पष्ट है कि मध्यका का मूल्य 75.5 है।

उदाहरण 28 : निम्नलिखित आवत्ति से एक संचयी आवत्ति वक्र बनाइये मध्यका और दोनों चतुर्थकों का मूल्य ज्ञात कीजिये।

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
0-5	4	20-25	25
5-10	6	25-30	22
10-15	10	30-35	18
15-20	10	35-40	5

हल : संचयी आवत्ति वक्र से मध्यका तथा चतुर्थकों का परिगणन

अंक (इतने से कम)	विद्यार्थियों की संख्या (f)	*c.f.*
5	4	4
10	6	10
15	10	20
20	10	30
25	25	55
30	22	77
35	18	95
40	5	100

$$M = \text{Size of } \frac{N}{2}\text{th item} = \frac{117}{2} = 58.35\text{th items}$$

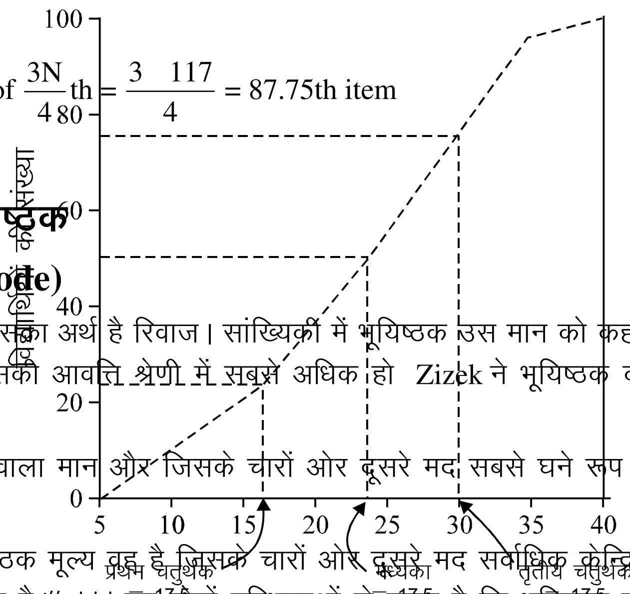
चित्र से स्पष्ट है कि मध्यका का मूल्य 37 है।

$$Q_1 = \text{Size of } \frac{N}{4}\text{th item} = \frac{117}{4} = 29.25\text{th item}$$

चित्र से स्पष्ट कि प्रथम चतुर्थक का मूल्य 24.5 है।

$$Q_3 = \text{Size of } \frac{3N}{4}\text{th} = \frac{3 \quad 117}{4} = 87.75\text{th item}$$

चित्र से स्पष्ट है कि ततीय चतुर्थक का मूल्य 48 है।



भूयिष्ठक
(Mode)

'Mode' शब्द की उत्पत्ति फ्रेंच भाषा के 'La Mode' से हुई जिसका अर्थ है रिवाज। सांख्यिकी में भूयिष्ठक उस मान को कहते हैं जो समंक माला में सबसे अधिक बार आता है अर्थात् जिसकी आवत्ति श्रेणी में सबसे अधिक हो Zizek ने भूयिष्ठक की परिभाषा इस प्रकार दी है :-

''मदों की किसी श्रेणी (या समूह) में सबसे अधिक बार आने वाला मान और जिसके चारों ओर दूसरे मद सबसे घने रूप में विपरित होते है।''**

क्रॉक्सटन व कॉउडेन के मतानुसार, ''किसी वितरण का भूयिष्ठक मूल्य वह है जिसके चारों ओर दूसरे मद सर्वाधिक केन्द्रित हों। वह मूल्यों की श्रेणी में सर्वाधिक प्रतिरूप माना जा सकता है।'' *** इन दोनों परिभाषाओं से स्पष्ट है कि भूयिष्ठक वह मूल्य है जिसकी आवत्ति सबसे अधिक हो। भूयिष्ठक को सर्वाधिक घनत्व की स्थिति या मूल्यों के अधिक केन्द्रीकरण का बिन्दु भी कहा जाता है। निम्न उदाहरण द्वारा भूयिष्ठक का अर्थ स्पष्ट हो जायेगा।

कॉलर का नाम (सेंमी. में)	:	25	28	30	32	35	38
व्यक्तियों की संख्या	:	13	15	26	28	24	10

उपर्युक्त उदाहरण से यह स्पष्ट हो जाता है कि भूयिष्ठक 32 सेंमी. है क्योंकि सबसे अधिक व्यक्ति अर्थात 28 व्यक्ति इस नाप का कॉलर पहनते हैं।

व्यक्गित श्रेणी में भूयिष्ठक का निर्धारण
(Determination of Mode in Individual Series)

व्यक्तिगत श्रेणी में भूयिष्ठक ज्ञात करने के लिये या निश्चित करना पड़ता है कि कौन-सा मद सबसे अधिक श्रेणी में आ रहा है — जो मद सबसे अधिक बार आता है उसे भूयिष्ठक कहते हैं। निम्न उदाहरण द्वारा गणन-क्रिया स्पष्ट हो जायेगी :

उदाहरण 29 : निम्न समंकों से भूयिष्ठक ज्ञात कीजिये :-

10 विद्यार्थियों द्वारा प्राप्त अंक

क्रम संख्या	प्राप्तांक	क्रम संख्या	प्राप्तांक
1	10	6	27
2	27	7	20
3	24	8	18
4	12	9	15
5	27	10	30

* "The value of the variable which occurs most frequently in a distribution is called the mode."

** "The value occuring most frequently in a series (or group) of item and around which the other items are distributed most desely."

*** "The mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded as the most typical of a series *pf* values."

हल : **भूयिष्ठक का निर्धारण**

प्राप्तांक	आवत्ति	प्राप्तांक	आवत्ति
10	1	20	1
12	1	24	1
15	1	27	3
18	1	30	1

क्योंकि अंक 27 श्रेणी में सबसे अधिक बार (अर्थात् 3 बार) आया है इसलिये भूयिष्ठक 27 है।

टिप्पणी : गणन-क्रिया से यह विदित है कि व्यक्तिगम श्रेणी में भूयिष्ठक निकालने के लिये पहले उसे खंडित श्रेणी में बदलना पड़ता है।

यदि दो या दो से अधिक चर-मूल्यों की आवत्तियों अधिकतम हों तो यह कहना कठिन हो जाता है कि भूयिष्ठक कौन -सा है क्योंकि श्रेणी, में उतने ही भूयिष्ठक होगें जितनी अधिकतम आवत्तियां हैं। ऐसी समंकमालायें दो -भयिष्ठक वाली तीन-भूयिष्ठक वाली (Tri-model) या अनेक-भूयिष्ठक वाली (Multi-model) श्रेणियां कहलाती हैं। ऐसी परिस्थिति में भूयिष्ठक निकालने के लिये निम्न सूत्र का प्रयोग किया जाता है :-

(Mode = 3 Median – 2 Mean)

उदाहरण 30 : 10 व्यक्तियों द्वारा प्राप्त मासिक वेतन इस प्रकार हैं :-

वेतन (रु. में) **:** 1600 1620 1630 1620 1630 1620 1630 1620 1640 1650

भूयिष्ठक ज्ञात कीजिये।

हल : भूयिष्ठक का निर्धारण

वेतन	आवत्ति	वेतन	आवत्ति
1600	1	1640	2
1620	3	1650	1
1630	3		
1630	3		

क्योंकि 1620 और 1630 दोनों मूल्यों की समान अधिकतम आवत्तियां हैं यह कहना है कठिन कि भूयिष्ठक 1620 है या 1630।

भूयिष्ठक का अनुमान ऐसी परिस्थिति में निम्न सूत्र द्वारा लगाया जाता है :-

Z = 3 median – 2 mean

भूयिष्ठक का निर्धारण

वेतन आरोही क्रम में अनुविन्यसित	वेतन आरोही क्रम में अनुविन्यसित
1600	1630
1620	1630
1620	1640
1620	1640
1630	$\frac{5\text{th item}}{2} + \frac{6\text{th item}}{2}$ 1650

$$M = \text{Size of } \frac{N \quad 1}{2}\text{th item} = \qquad = 5.5\text{th item}$$

$$= \qquad\qquad = \frac{1630 \quad 1630}{2} = 1630$$

$$= \frac{X}{N} = \frac{16280}{10} = 1628 \text{ रू.}$$

$$\text{Mode} = (3 \times 1630) - (2 \times 1628) = 4890 - 3256 = 1634 \text{ रू.}$$

अर्थात् भूयिष्ठक-मूल्य = 1634 रू.

विच्छिन्न श्रेणी में भूयिष्ठक का निर्धारण

(Determination of Mode in Discrete Series)

विच्छिन्न श्रेणी में भूयिष्ठक ज्ञात करने की दो रीतियां हैं :- (*i*) निरीक्षण रीति (Inspection Method), तथा (*ii*) समूहीकरण रीति (Grouping Method)।

(*i*) **निरीक्षण रीति (Inspection Method) :** विच्छिन्न श्रेणी में केवल निरीक्षण द्वारा भूयिष्ठक ज्ञात किया जा सकता है लेकिन यह तभी सम्भव है जब आवत्तियां नियमित हों अर्थात् श्रेणी के आरम्भ में आवत्तियां निरन्तर बढ़ती रहें , अधिकतम आवत्ति लगभग केन्द्र में हो और उसके बाद आवत्तियां फिर निरन्तर घटने लगे। ऐसे समूह में अधिकतम आवत्ति बिल्कुल स्पष्ट होती है। निरीक्षण द्वारा उसका मूल्य ज्ञात कर लिया जाता है। निम्न उदाहरण से यह रीति स्पष्ट हो जायेगी :-

उदाहरण 31 : निम्न सारणी 50 व्यक्तियों की आयु प्रदर्शित करती है। भूयिष्ठक ज्ञात कीजिये :-

आयु (वर्ष में)	आवत्ति *f*	आयु (वर्ष में)	आवत्ति *f*
20	4	23	18
21	6	24	7
22	12	25	3

हल : उपर्युक्त सारणी में आवत्तियां नियमित हैं अत : निरीक्षण द्वारा भूयिष्ठक ज्ञात किया जा सकता है। श्रेणी का निरीक्षण करने से पता लगता है कि अधिकतम आवत्ति 18 है जिसका मान 23 है अर्थात् भूयिष्ठक आयु 23 वर्ष है।

(ii) **समूहीकरण रीति (Grouping Method) :** जहां आवत्तियों में अनियमितता (Irregularity) हो वहां निरीक्षण द्वारा भूयिष्ठक ज्ञात करना कठिन है। कभी-कभी दो या दो से अधिक मूल्यों की आवत्ति सबसे अधिक होती है और यह निश्चित करना कठिन होता है कि मद को भूयिष्ठक माना जाये। सर्वाधिक आवत्ति एक होने पर पता भी भूयिष्ठक मद सर्वाधिक आवत्ति वाला न होकर दूसरा हो सकता है। भूयिष्ठक का ठीक लगाने के लिये समूहीकरण रीति (Grouping Method) का प्रयोग किया जाता है। जब यह रीति प्रयोग में लाई जाती है तो दो सारणियां बनाई जाती हैं, जो इस प्रकार हैं : (1) सामूहीकरण सारणी (Grouping Table) तथा (2) विश्लेषण (Analysis Table)।

इन दोनों सारणियों के आधार पर भूयिष्ठक ज्ञात किया जाता है।

(1) समूहीकरण सारणी (Grouping Table) : समूहीकरण का उद्देश्य अनियमित आवत्ति वाले वितरण में आवत्तियों का सर्वाधिक घनत्व निश्चित करना होता है। समूहीकरण सारणी बनाने की क्रिया इस प्रकार है :-

एक सारणी बनाई जाती है जिसमें चर-मानों के अतिरिक्त 6 स्तम्भ होते हैं। इन स्तम्भों में आवत्तियों का दो -दो और तीन-तीन के समूहों में समूहन (Grouping in two'o and three's) निम्न क्रम से किया जाता है :-

(i) **प्रथम स्तम्भ :** प्रथम स्तम्भ में प्रश्न में दी हुई आवत्तियां होती हैं। इन आवत्तियों में से अधिकतम आवत्ति को चिन्हित किया जाता है।

(ii) **द्वितीय स्तम्भ :** दूसरे स्तम्भ में प्रथम स्तम्भ में दी हुई पहली दो आवत्तियों का योग , फिर इसके आगे वाली दो आवत्तियों का योग , और इसी प्रकार अन्त तक दो-दो आवत्तियों का योग लिया जाता है। *

(iii) **ततीय स्तम्भ :** प्रथम स्तम्भ में दी हुई आवत्तियों में से पहले वाली आवत्ति को छोड़कर दो -दो आवत्तियों के योग लिये जाते हैं।*

(iv) **चतुर्थ स्तम्भ :** प्रथम स्तम्भ में दी हुई आवत्तियों में से पहली तीन आवत्तियों का योग , फिर आगे की तीन आवत्तियों का योग और इसी प्रकार आगे भी तीन -तीन आवत्तियों का योग लेते हैं। *

(v) **पंचम स्तम्भ :** प्रथम स्तम्भ में दी हुई आवत्तियों में से पहली आवत्ति को छोड़ कर अगली तीन आवत्तियों का योग , फिर अगली तीन आवत्तियों का योग और इसी प्रकार अगले भी तीन -तीन आवत्तियों का योग लेते हैं। *

(vi) **षष्ठम स्तम्भ :** प्रथम स्तम्भ में दी हुई आवत्तियों में से प्रथम आवत्तियों को छोड़कर अगली तीन आवत्तियों का योग और इसी प्रकार आगे तीन-तीन आवत्तियों का योग लेते हैं।

* यह याद रखना अत्यन्त आवश्यक है कि समूहन (Grouping) केवल आवत्तियों का ही होता है , चर (Variable) का नहीं।

उपयुक्त स्तम्भों में समूहीकरण की संख्यायें लिखने के बाद प्रत्येक खाने की सबसे बड़ी संख्या को मोटे अक्षरों में वत्त में लिख दिया जाता है ताकि वह अन्य संख्याओं से भिन्न लगे और सुगमता से पहचानी जा सके।

(2) विश्लेषण सारणी (Analysis Table) : यह सारणी उप अधिकतम आवत्तियों (Maximum Frequencies) के आधार पर बनाई जाती है जिन्हें उपयुक्त समूहीकरण वाली सारणी में रंखांकित (Underline) किया गया है या वत्त में लिखा गया है इस सारणी में छहों स्तम्भों के सामने अधिकतम आवत्तियों के चर -मूल्यों पर चिन्ह लगाकर उनकी गणना कर ली जाती है जिस मूल्य के सामने सबसे अधिक चिन्ह होते हैं वही भूयिष्ठक मूल्य होता है। विश्लेषण सारणी का प्रारूप नहीं दिया गया है :-

विश्लेषण सारणी का प्रारूप

स्तंभ संख्या	चल के मान						
I							
II							
III							
IV							
V							
VI							
योग							

उदाहरण 32 : निम्नलिखित समंकमाला से समूहीकरण रीति द्वारा भूयिष्ठक ज्ञात कीजिये :-

चर :	40	44	48	52	56	60	64	68	72	76
आवत्ति :	10	12	14	20	15	20	18	10	8	4

हल : भूयिष्ठक का निर्धारण (समूहीकरण सारणी)

	I	II		III	IV	V	VI
40	10						
44	12	22					
48	14		26	36	46	46	
52	20	34					49
56	15		35				
60	20	35		55	53	53	
64	18		38				
68	10	28		36			48
72	8		18		22	52	
76	4	12					

विश्लेषण सारणी

स्तम्भ संख्या	**40**	**44**	**48**	**52**	**56**	**60**	**64**	**68**	**72**	**76**
	चर मूल्यसंख्या									
I				1		1				
II					1	1				
III						1	1			
IV				1	1	1				
V					1	1	1			
VI			1	1	1					
योग			**1**	**3**	**4**	**5**	**2**			

अर्थात् सारणी से यह स्पष्ट है कि चर मूल्य 60 सबसे अधिक (पांच) बार आया है अर्थात् यही भूयिष्ठक मूल्य है।

अविच्छिन्न श्रेणी में भूयिष्ठक का निर्धारण
(Determination of Mode in Continuous Series)

अविच्छिन्न श्रेणी में भूयिष्ठक ज्ञात करने के लिये सर्वप्रथम भूयिष्ठक (Mode Group) का निश्चय करना पड़ता है। यदि आवत्तियां नियमित हैं तो निरीक्षण द्वारा हो। भूयिष्ठक वर्ग का पता चल जाता है अन्यथा समूहीकरण रीति द्वारा भूयिष्ठक वर्ग निर्धारित किया जाता है। भूयिष्ठक वर्ग से यह निश्चित हो जाता है कि माध्य इसी वर्ग की निम्नतम तथा उच्चतम सीमा के बीच में है। सीमाओं के अंतर्गत भूयिष्ठक का मूल्य निर्धारित करने के लिये निम्न सूत्र का प्रयोग किया जाता है :-

$$M_0 = L + \frac{1}{i \quad 2} \times i$$

L = भूयिष्ठक वर्ग की निम्न सीमा (Lower limit of the modal class) $D_i = f_1 - f_0$ (ignoring the signs)

f_1 = भूयिष्ठक वर्ग की आवत्ति (Frequency of the Modal Class)

f_0 = भूयिष्ठक वर्ग से पहले वर्ग की आवत्ति (Frequency of the Class Preceeding the modal class)

$D_2 = f_1 - f_2$ (ignoring the signs)

f_2 = भूयिष्ठक वर्ग के बाद वाले वर्ग की आवत्ति (Frequency of the class succeeding the modal class)

i = भूयिष्ठक वर्ग का वर्गान्तर (Class interval of the modal class)

उपयुक्त सूत्र को इस प्रकार भी लिखा जा सकता है :-

$$M_0 = L + \frac{f_1 \quad f_0}{2f_1 \quad f_0 \quad f_2} \times i$$

टिप्पणी : उपर्युक्त सूत्रों का प्रयोग करते समय यह ध्यान रखना चाहिये कि भूयिष्ठक वर्ग तथा भूयिष्ठक से पहले वाले और बाद वाले वर्गों के वर्गान्तर समान हैं। यदि वर्गान्तर समान नहीं हैं तो पहले उन्हें समान कर लेना चाहिये।

उदाहरण 33 : निम्नलिखित सूचना को एक साधारण आवत्ति -वितरण में परिवर्तित कीजिये और उससे भूयिष्ठक ज्ञात कीजिये:-

5 विद्यार्थी	3 से कम अंक प्राप्त करते हैं।
12 विद्यार्थी	6 से कम अंक प्राप्त करते हैं।
25 विद्यार्थी	9 से कम अंक प्राप्त करते हैं।
30 विद्यार्थी	12 से कम अंक प्राप्त करते हैं।

हल : भूयिष्ठक का निर्धारण

अंक	विद्यार्थियों की संख्या (f)
0-3	5
3-6	7
6-9	13
9-12	5
योग	**30**

निरीक्षण द्वारा यह स्पष्ट है कि भूयिष्ठक वर्ग 6-9 है।

$$M_0 = L + \frac{f_1 \quad f_0}{2f_1 \quad f_0 \quad f_2} \times i$$

$L = 6$; $D_1 = (13 - 7) = 6$; $D_2 = (13 - 5) = 8$; $i = (9 - 6) = 3$

$$M_0 = 6 + \frac{6}{6 \quad 8} \times 3 = 6 + 1.3 = 7.3$$

उदाहरण 34 : एक कारखाने में काम करने वाले मजदूरों की मजदूरी के निम्न समंकों से एक आवत्ति -वितरण 10 का वर्गान्तर लेकर बनाइये, जिसमें वर्ग 10-19.99; 20-29.99; 30-39.99 आदि हों। इस आवत्ति -वितरण से भूयिष्ठक ज्ञात कीजिये।

मजदूरी रू. में

10	100	90	30	99	25	70	32	76	15
68	31	75	39	89	40	66	27	109	42
93	53	97	43	29	92	28	95	36	105
67	55	47	108	37	86	46	112	44	68
47	81	77	48	50	87	41	88	59	80
52	85	56	61	58	72	69	118	82	78
69	54	71	60	63	73	65	79	64	61

हल : आवत्ति वितरण से भूयिष्ठक का निर्धारण

<table>
<tr><th>मजदूरी (रुपयों में)</th><th>मिलान रेखायें</th><th>आवत्ति</th></tr>
<tr><td>10-19.99</td><td>||</td><td>2</td></tr>
<tr><td>20-29.99</td><td>||||</td><td>4</td></tr>
<tr><td>30-39.99</td><td>~~||||~~ |</td><td>6</td></tr>
<tr><td>40-49.99</td><td>~~||||~~ ||||</td><td>9</td></tr>
<tr><td>50-59.99</td><td>~~||||~~ |||</td><td>8</td></tr>
<tr><td>60-69.99</td><td>~~||||~~ ~~||||~~ ||</td><td>12</td></tr>
<tr><td>70-79.99</td><td>~~||||~~ ||||</td><td>9</td></tr>
<tr><td>80-89.99</td><td>~~||||~~ |||</td><td>8</td></tr>
<tr><td>90-99.99</td><td>~~||||~~ |</td><td>6</td></tr>
<tr><td>100-109.99</td><td>||||</td><td>4</td></tr>
<tr><td>110-119.99</td><td>||</td><td>2</td></tr>
</table>

भूयिष्ठक वर्ग 60-69.99 हैं जिसकी वास्तविक सीमायें 59.99-69.99 हैं।

$$M_0 = L + \frac{1}{1 \quad 2} \times i$$

$L = 59.995$; $D_1 = (12 - 8) = 4$; $D_2 = (12 - 9) = 3$; $i = 10$

$$Z = 59.995 + \frac{4}{4 \quad 3} \times 10 = 59.995 + 5.714 = 65.7 \text{ रू.}$$

उदाहरण 35 : निम्नलिखित समंक 122 व्यक्तियों का वजन प्रस्तुत करते हैं।

वजन (पौंडों में)	व्यक्यिों की संख्या	वजन (पौंडों में)	व्यक्तियों की संख्या
100-110	4	140-150	33
110-120	6	150-160	17
120-130	20	160-170	8
130-140	32	170-180	2

हल : निरीक्षण द्वारा यह कहना कठिन है कि भूयिष्ठक वर्ग कौन-सा है। समूहीकरण सारणी तथा विश्लेषण सारणी द्वारा भूयिष्ठक वर्ग ज्ञात किया जायेगा।

सामूहीकरण सारणी

वजन (पौंडों में)	आवत्ति			व्यक्तियों की संख्या		
	I	II	III	IV	V	VI
100-110	4	10	26	30	58	85
110-120	6					
120-130	20	52				
130-140	32		65			
140-150	33	50		82	58	27
150-160	17		25			
160-170	8	10				
170-180	2					

विश्लेषण सारणी

वर्ग जिसमें भूयिष्ठक स्थित होने की सम्भावना है।

क्रम सं.	120-130	(61 30) 310 / 32 23 130-140	140-150
I		1	1
II	1	1	
III		1	1
IV			1
V	1	1	1
VI	1	1	1
योग	**3**	**5**	**5**

यह दो भूयिष्ठक वाली (Bimodal) श्रेणी है इसमें भूयिष्ठक का निर्धारण निम्न सूत्र द्वारा किया जायेगा :-

$$M_0 = 3 \text{ median} - 2 \text{ mean}$$

$$M = \text{Size of } \frac{N}{2}\text{th item} = \frac{122}{2} = 61\text{st item}$$

अर्थात् मध्यका वर्ग 130-140 है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 30; N/2 = 61; *c.f.* = 30; *f* = 32; *i* = 10

\

$$M = 130 + \qquad \times 10$$

$$= 130 + \qquad = 130 + 9.69 = 139.69$$

सामान्तर माध्य का परिगणन

वजन पौंडों में	*m*	*f*	(*m* – 135)/10	*f*	*c.f.*
100-110	105	4	– 3	– 12	4
110-120	115	6	– 2	– 12	10
120-130	125	20	– 1	– 20	30
130-140	135	32	0	0	62
140-150	145	33	+ 1	+ 33	95
150-160	155	17	+ 2	+ 34	112
160-170	165	8	+ 3	+ 24	120
170-180	175	2	+ 4	+ 8	122
	N = 122			**S*fd* = 55**	

$$\overline{X} = A + \frac{fd}{N} \times i$$

$$\overline{X} = 135 + \frac{55}{122} \times 10 = 135 + 4.51 = 139.51$$

$$M_0 = 3 \text{ median} - 2 \text{ mean}$$

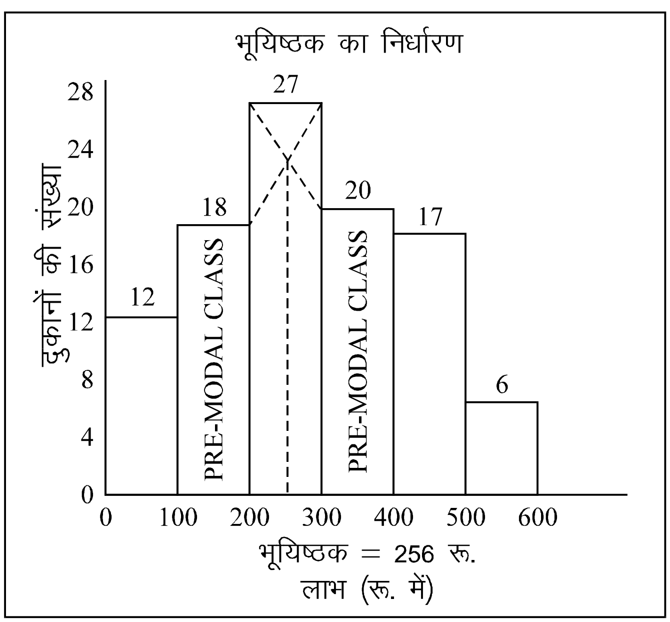
$$M_0 = (3 \times 130.59) - (2 \times 139.51) = 419.07 - 279.01 = 140.05$$

उदाहरण 36 : दुकानों के मासिक लाभ का वितरण इस प्रकार है :-

लाभ (लाख रू. में)	दुकानों की संख्या	लाभ (लाख रू. में)	दुकानों की संख्या
0-100	12	300-400	20
100-200	18	400-500	17
200-300	27	500-600	6

रेखा-चित्र द्वारा भूयिष्ठक ज्ञात कीजिये और प्राप्त मूल्य की जांच सूत्र द्वारा भूयिष्ठक निकाल कर कीजिये।

हल : रेखाचित्र द्वारा भूयिष्ठक का निर्धारण



सूत्र द्वारा भूयिष्ठक का निर्धारण :

निरीक्षण द्वारा स्पष्ट है कि भूयिष्ठक-वर्ग 200-300 है।

$$M_0 = L + \frac{1}{1 \quad 2} \times i$$

L = 200; D_1 = (27 – 18) = 9; D_2 = (27 – 20) = 7; i = 100

$$M_0 = 200 + \frac{9}{9 \quad 7} \times 100 = 200 + 56.25 = 256.25 \text{ लाख रु.}$$

उपर्युक्त उदाहरण से स्पष्ट होता है कि सूत्र द्वारा तथा रेखाचित्र द्वारा भूयिष्ठक का मूल्य एक ही आयेगा।

उदाहरण 37 : निम्नलिखित समंकों से बहुलक ज्ञात कीजिये। आयत चित्र द्वारा भी बहुलक की स्थिति बताइये।

मूल्य	बारम्बारता	मूल्य	बारम्बारता
0-5	328	20-25	598
5-10	350	25-32	524
10-15	720	30-35	378
15-20	664	35-42	244

हल : निरीक्षण द्वारा स्पष्ट है कि बहुलक 10-15 वर्ग में स्थित है।

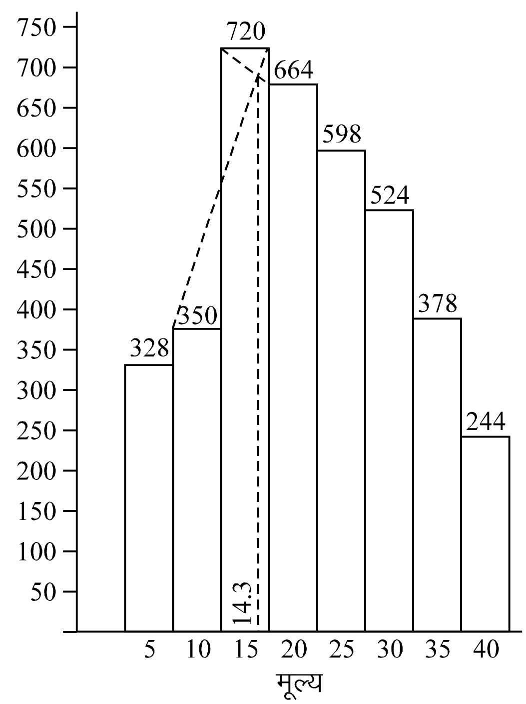
$$M_0 = L + \frac{1}{1 \quad 2} \times i$$

L = 10; $D_1 = (f_1 - f_0) = 720 - 350 = 370$

$D_2 = (f_1 - f_2) = 720 - 664 = 56$; $i = 5$

$$M_0 = 10 + \frac{370}{370 \quad 56} \times 5 = 10 + 4.34 = 14.34$$

आवत्ति आयत चित्र द्वारा बहुलक का निर्धारण



उदाहरण (ख) : एक सामान्य विषयम वितरण में भूयिष्ठक एवं समान्तर माध्य क्रमशः 32.1 और 35.4 हैं। मध्यका ज्ञात कीजिये।

हल :

$$\text{Mode} = 3\ \text{Median} - 2\ \text{Mean}$$
$$32.1 = 3\ \text{Median} - 2\ (35.4)$$
$$3\ \text{Median} - 70.8 = 102.9$$

\ Median = 34.3 अर्थात् मध्यका मूल्य 34.3 है।

गुणोत्तर माध्य
(Geometric Mean)

किसी समंक श्रेणी का गुणोत्तर माध्य उसके सभी मानों के गुणनफल का वह मूल होता है जिसकी उस श्रेणी में इकाइयां हैं। उदाहरणार्थ; यदि दो संख्याओं के मूल्य 4 और 16 हैं तो उनका गुणोत्तर माध्य $\sqrt{(4\quad 16)}$ अर्थात 8 होगा। इसी प्रकार, यदि तीन संख्याओं के मूल्य 4, 16, 8 हैं तो उनका गुणोत्तर माध्य = 8 होगा। $(8 \times 8 \times 8 = 512)$। सूत्र के रूप में :-

G.M. =

G.M. = गुणोत्तर माध्य (Geometric Mean)

X_1, X_2, X_3, etc. = चर के विभिन्न मान;

N = मदों की संख्या

यदि मदों की संख्या 4 या 4 से अधिक है तो उपर्युक्त सूत्र के गुणोत्तर माध्य की गणना कठिन हो जायेगी। ऐसी स्थिति में लघुगणकों (Logarithms) तथा प्रतिलघुगणकों का प्रयोग किया जाता है। इस प्रकार गुणोत्तर माध्य निकालने का निम्न सूत्र होगा :-

$$\text{G.M.} = \text{Antilog}\ \frac{\text{Log}\ X_1 + \text{Log}\ X_2 \ldots \text{Log}\ X_n}{N} \text{ or A.L. } \frac{\text{Log X}}{N}$$

व्यक्तिगत श्रेणी में गुणोत्तर माध्य का निर्धारण
(Calculating Geometric Mean in Individual Series)

व्यक्ति श्रेणी में गुणोत्तर माध्य की गणना निम्न सूत्र द्वारा की जाती है :-

$$\text{G.M.} = \text{A.L. } \frac{\text{Log X}}{N}$$

विधि :

(1) चर के सभी मदों का लघुगणक निकालियें और उनका योग कीजिये अर्थ S Log X निकालिये।

(2) S Log X को मदों की संख्या से भाग दीजिये और प्राप्त राशि का प्रति लघुगुंणक निकालिये। यह मूल्य गुणोत्तर माध्य है।

उदाहरण 38 : निम्न समंकों से गुणोत्तर ज्ञात कीजिये।

15 250 15.7 157 105.7 10.5 1.06 25.7 0.257

हल : गुणोत्तर माध्य का परिगणन

X	Log X
15.000	1.1761
250.000	2.3879
15.700	1.1959
157.000	2.1959
1.570	0.1959
105.700	2.0241

X	log X
10.500	1.0212
1.060	0.0253
25.700	1.4099
0.257	1.4099
N = 10	**S log X = 11.0521**

$$\text{G.M.} = \text{A.L.}\left(\frac{\text{Log X}}{\text{N}}\right) = \text{A.L.}\left(\frac{11.0521}{10}\right) = \text{A.L. } 1.1052 = 1.27$$

विच्छिन्न श्रेणी में गुणोत्तर माध्य का परिगणन
(Calculating Geometric Mean in Discrete Series)

$$\text{G.M.} = \text{Antilog}\left(\frac{f\text{Log X}}{\text{N}}\right)$$

विधि :

(1) श्रेणी के प्रत्येक मद का लघुगणक ज्ञात कीजिये।

(2) इन लघुगणकों को उनकी आवत्ति से गुणा कीजिये और S (*f* Log X) निकालिये।

(3) (S *f* Log X) को आवत्ति के कुल योग से भाग दीजिये।

(4) प्राप्त भागफल का प्रतिलघुगणक ज्ञात कीजिये। इस प्रकार जो राशि आयेगी उसे गुणोत्तर माध्य कहेंगे।

उदाहरण 39 : निम्न समंकों से गुणोत्तर माध्य ज्ञात कीजिये :-

अंक	:	20	22	40	45	50	10	8
विद्यार्थियों की संख्या	:	4	6	10	20	8	6	4

नोट : जहां ऋणात्मक (Negative) तथा धनात्मक (Positive) Characteristics हों वहां पहले (Mentissa) अर्थात् दशमलव के सीधे हाथ पर दिये मानों का योग कर लेना चाहिये और फिर इस योग द्वारा योग की अन्य धनात्मक Characteristics में जोड़कर कुल योग में से ऋणात्मक Characteristics को घटा देना चाहिये। उदाहरण 39 में Mantissa का योग 2.052 है तथा धनात्मक और एक ऋणात्मक Characteristics है अर्थात् कुल योग 10 + 2.052 = अर्थात 12.052 होगा।

गुणोत्तर माध्य का निर्धारण

अंक X	विद्यार्थियों की संख्या *f*	log X	*f* log X
20	4	1.3010	5.2040
22	6	1.3424	8.0544
40	10	1.6021	16.0210
45	20	1.6532	33.0640
50	8	1.6990	13.5920
10	6	1.0000	6.0000
8	4	1.9031	3.6124
60	2	1.7782	3.5564
	N = 60		**S *f* log X = 89.1042**

$$\text{G.M.} = \text{Antilog}\left(\frac{f \text{ Log X}}{\text{N}}\right) = \text{A.L.}\left(\frac{89.1042}{60}\right)$$

$$= \text{Antilog } 1.4851 = 30.56$$

अविच्छिन्न श्रेणी में गुणोत्तर माध्य का निर्धारण

(Calculating Geometric Mean in Continuous Series)

अविच्छिन्न श्रेणी में गुणोत्तर माध्य निम्न सूत्र द्वारा ज्ञात किया जाता है :-

$$\text{G.M.} = \text{A.L.}\left(\frac{f \text{ Log } m}{\text{N}}\right)$$

$$m = \text{मध्य बिन्दु}$$

विधि :

(1) श्रेणी के सब मदों के मध्य-बिन्दु निकालिये और उनका लघुगणक लीजिये।

(2) इन लघुगणकों को उनकी आवत्ति से गुणा कीजिये और ($Sf \log m$) निकालिये।

(3) ($Sf . \log m$) को आवत्ति के कुल योग से भाग दीजिये।

(4) प्राप्त भागफल का प्रतिलघुणक लीजिये। इस प्रकार जो मूल्य आयेगा उसे गुणोत्तर माध्य कहेंगे।

उदाहरण 40 : निम्न समंकों से गुणोत्तर माध्य ज्ञात कीजिये :-

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
0-10	8	30-40	6
10-20	12	40-50	4
20-30			

हल : **गुणोत्तर माध्य का निर्धारण**

अंक	मध्य बिन्दु m	विद्यार्थियों की संख्या f	**log m**	$f \times \log m$
0-10	5	8	0.6990	5.5920
10-20	15	12	1.1761	14.1132
20-30	25	20	1.3979	27.9580
30-40	35	6	1.6532	6.6128
40-50	45	4	1.6532	6.6128
	N = 50			**S ($f \times \log m$) = 63.546**

$$\text{G.M.} = \text{A.L.}\left(\frac{f \text{ Log } m}{\text{N}}\right) \text{A.L.}\left(\frac{63.5406}{50}\right)$$

$$= \text{A.L. } 1.2708 = 18.65$$

भारित गुणोत्तर माध्य

(Weighted Geometric Mean)

यदि विभिन्न मूल्यों का सापेक्ष महत्व समान न हो तो समानन्तर माध्य की भांति भारित (Weighted) गुणोत्तर माध्य किया जा सकता है। भारित गुणोत्तर माध्य निकालते समय निम्न सूत्र का प्रयोग किया जाता है :-

$$\text{Weighted G.M.} = \text{A.L.}\left[\frac{(\text{W.log X})}{\text{W}}\right]$$

$$\text{W} = \text{भार (Weights)}$$

विधि :

(1) श्रेणी की प्रत्येक मद का लघुगणक निकालिये।

(1) प्रत्येक मद के लघुगणक को भार 'W' से गुणा करके गुणनफलों का योग S (W. log X) निकालिये।

(3) S (W. log X) को भारों के योग से भाग दीजिये और भागफल का प्रतिलघु गुणंक निकालिये। प्राप्त मूल्य भारित गुणोत्तर माध्य होगा।

उदाहरण 41 : निम्न समंक से भारित गुणोत्तर माध्य ज्ञात कीजिये :-

मदें	सूचकांक	भार
भोजन	125	7
वस्त्र	133	5
ईंधन और रोशनी	141	4
मकान का किराया	173	1
अन्य	182	3

हल : **भारित गुणोत्तर माध्य का परिगणन**

मदें	सूचकांक	भार	Log X W. Log X	
भोजन	125	7	2.0969	14.6783
वस्त्र	133	5	2.1239	10.6195
ईंधन और रोशनी	141	4	2.1492	8.5968
मकान का किराया	173	1	2.2380	6.2380
अन्य	182	3	2.2601	6.7803
		SW = 20		**S (*w*.Log X) = 42.9219**

$$\text{Weighted G.M.} = \text{A.L.}\left[\frac{(\text{W.log X})}{\text{W}}\right]$$

$$= \text{A.L.} = \text{A.L.}\ (2.1456) = 139.8$$

गुणोत्तर माध्य के विशेष प्रयोग
(Special Uses of Geometric Mean)

गुणोत्तर माध्य का प्रमुख उपयोग प्रतिशत वद्धि -दरों तथा अनुपातों की औसत निकालने में किया जाता है। विशेषत: जनसंख्या की वद्धि , चक्रवद्धि ब्याज , मूल्यों में होने वाले प्रतिशत परिवर्तनों आदि की औसत दरें गुणोत्तार माध्य पर आधारित, 'चक्रवद्धि ब्याज' (Compound Interest Formula) के प्रयोग द्वारा ज्ञात की जाती है। सूत्र इस प्रकार हैं :-

$$P_n = P_0 (1 + r)^n$$

जहां P_0 = निश्चित अवधि के बार चर-मूल्य की राशि (Value of variable at the end of a period)

P_n = अवधि के आरम्भ में चर-मूल्य (Value of variable in the beginning)

n = वर्षों, आदि की संख्या (Number of years or other points of time)

r = प्रति इकाई परिवर्तन की दर (Average rate of change per units)

प्रति इकाई परिवर्तन की दर ज्ञात करनी हो तो निम्न सूत्र का प्रयोग करते हैं :-

$$r = n\sqrt{\left(\frac{P_n}{P_0}\right) \quad 1}$$

उदाहरण 42 : एक देश की जनसंख्या सन् 1968 में 30 करोड़ थी। सन् 1986 में जनसंख्या बढ़कर 52 करोड़ हो गई। वार्षिक प्रतिशत वद्धि दर ज्ञात कीजिये।

हल : यहां चक्रवद्धि ब्याज सूत्र का प्रयोग किया जायेगा।

$$P_n = P_0 (1 + r)^n$$

$$P_n = 52 \text{ करोड़},\ P_0 = 30 \text{ करोड़ } n = 18 \text{ वर्ष}$$

$$52 = 30 (1 + r)^{18}$$

$$(1 + r)^{16} - \frac{52}{30} = (1.733)^{1/18}$$

$$\text{Log X} = 1/73^{1/18}$$

$$\text{Log X} = \frac{1}{18} \text{ Log } 1.731 = \frac{1}{18} \times 0.2387 = 0.01326$$

$$X = \text{Antilog } 0.01326 = 1.031$$

$$(1 + r) = 1.031$$

$$r = 1.031 - 1 = 0.031 \text{ or } 3.1\%$$

चक्रवद्धि वार्षिक दर = 3.1 प्रतिशत।

उदाहरण 43 : 5 वर्षों में कारखाने के उत्पादन की वार्षिक वद्धि दर क्रमश : 5, 7.5, 2.5, 5 और 10 प्रतिशत थी। उत्पादन की चक्रवद्धि वार्षिक वद्धि दर (Compount Rate of Growth) ज्ञात कीजिये।

हल : गुणोत्तर माध्य द्वारा इस प्रश्न को हल किया जायेगा।

गुणोत्तर माध्य का परिगणन

वार्षिक वद्धि दर	वर्ष के अन्त में उत्पादन	Log X
5.0	105.0	2.0212
7.5	107.5	2.0314
2.5	102.5	2.0107
5.0	105.0	2.0212
10.0	110.0	2.0414
		S Log X = 101259

$$\text{G.M.} = \text{A.L.}\left(\frac{\text{Log X}}{\text{N}}\right) = \text{A.L.}\left(\frac{10.1259}{5}\right) = \text{A.L. } 2.0251 = 105.9$$

चक्रवद्धि वार्षिक वद्धि दर = 105.9 – 5.9 प्रतिशत।

उदाहरण 44 : एक मशीन का मूल्य पहले वर्ष 40 प्रतिशत कम होता है, दूसरे वर्ष में 25 प्रतिशत तथा अगले तीन वर्षों में 10 प्रतिशत (प्रत्येक प्रतिशत को घटते हुए मूल्य के आधार पर निर्धारित किया गया है) पांचों के लिये औसत अवमूलक (Average Percentage Depreciation) ज्ञात कीजिये।

हल : गुणोत्तर माध्य द्वारा इस प्रश्न को हल किया जायेगा।

गुणोत्तर माध्य का परिगणन

वर्ष	X	Log X
I	100 – 40 = 60	1.7782
II	100 – 25 = 75	1.8751
III	100 – 10 = 90	1.9542
IV	100 – 10 = 90	1.9542
V	100 – 10 = 90	1.9542
		S Log X = 9.5159

$$\text{G.M.} = \text{A.L.}\left(\frac{\text{Log X}}{\text{N}}\right) = \text{A.L.}\left(\frac{9.5159}{5}\right)$$

$$= \text{A.L. } 1.90318 = 80.025$$

5 वर्षों का औसत अवमूल्यन (average depreciation) = 100 – 80.0 = 19.98 या 20 प्रतिशत।

उदाहरण 45 : यदि एक वस्तु की कीमत 4 वर्षों में दुगनी हो जाती है तो प्रतिशत औसत वद्धि (Average Percentage Increase per Annum) ज्ञात कीजिये।

हल : मान लीजिये वस्तु का मूल्य 100 रुपये है। 4 वर्ष बाद मूल्य 200 रुपये हो जायेगा तो वार्षिक औसत वद्धि दर निम्न सूत्रानुसार ज्ञात की जायेगी।

वार्षिक औसत वद्धि = 19 प्रतिशत।

$$r = n\sqrt{\left(\frac{P_n}{P_0}\right) \quad 1}$$

$$P_0 = 100; P_n = 200; n = 4$$

\

$$r = 4\sqrt{\frac{200}{100} \quad 1}$$

$$= \text{AL}\left(\frac{\text{Log } 2}{4}\right) - 1 = \text{AL}\left(\frac{0.3010}{4}\right) - 1$$

$$= \text{AL } [0.07525] - 1$$

$$= 1.09 - 1 = 0.19 \text{ या } 19 \text{ प्रतिशत।}$$

गुणोत्तर माध्य के गुण व दोष
(Merits and Demerits of Geometric Mean)

गुण (Merits) : गुणोत्तर माध्य के प्रमुख गण निम्नलिखित हैं :-

(1) समान्तर माध्य की भांति गुणोत्तर माध्य भी समंकमाला के सभी मूल्यों पर आधारित होता है।

(2) यह माध्य समंक श्रेणी के छोटे मूल्यों को अधिक और बड़े मूल्यों को कम महत्व देता है इस प्रकार एक संतुलित स्थिति स्पष्ट हो जाती है।

(3) अन्य माध्यों की अपेक्षा गुणोत्तर माध्य पर सीमान्त मूल्यों का बहुत कम प्रभाव पड़ता है।

(4) अनुपातों व प्रतिशत वद्धि दरों की औसत निकालने में गुणोत्तर माध्य विशेष रूप से उपर्युक्त होता है।

(5) समान्तर माध्य की भांति सामूहिक गुणोत्तर माध्य निकाला जा सकता है।

(6) इस माध्य का बीजगणितीय विवेचन सम्भव है।

दोष (Demerits) : गुणोत्तर माध्य के प्रमुख दोष निम्नलिखित हैं :-

(1) इस माध्य की गणन-क्रिया अन्य माध्यों की अपेक्षा अधिक कठिन है, इसलिये इसका प्रयोग बहुत कम होता है।

(2) यदि किसी मद का मूल्य शून्य (0) या ऋणात्मक है तो गुणोत्तर माध्य की गणना नहीं की जा सकती।

हरात्मक माध्य
(Harmonic Mean)

''यदि किसी श्रेणी के मदों की संख्या को उन मदों के व्युत्क्रमों (Reciprocals) के योग से भाग दिया जाये तो भागफल प्राप्त होता है उसे उस श्रेणी का हरात्मक माध्य कहते हैं।'' दूसरे शब्दों में हरात्मक माध्य किसी श्रेणी के विभिन्न मदों के व्युत्क्रमों के समान्तर माध्य का व्युत्क्रम होता है। किसी मूल्य का व्युत्क्रम वह संख्या है जो एक 1. को उस मूल्य से भाग देने पर, उपलब्ध होती है। जैस (10 का) व्युत्क्रम, $\frac{1}{10}$ है तथा 8.67 का $\frac{1}{8.67}$। किसी संख्या का व्युत्क्रम सारणी (Reciprocal Table) की सहायता से अत्यन्त सरलता से प्राप्त किया जा सकता है।

व्यक्तिगत श्रेणी में हारात्मक माध्य का परिगणन
(Calculating Harmonic Mean in Individual Series)

व्यक्तिगत श्रेणी में निम्न सूत्र द्वारा हरात्मक माध्य ज्ञात किया जाता है।

$$\text{H.M.} = \frac{N}{\frac{1}{X_1} \quad \frac{1}{X_2} \quad \frac{1}{X_3} \quad \frac{1}{X_4}} \text{ या } \frac{N}{(1/X)}$$

$$X_1, X_2, X_3, \text{etc.} = \text{चर से विभिन्न मूल्य}$$

उदाहरण 46 : निम्न समंकों से हरात्मक माध्य ज्ञात कीजिये।

1 0.5 10 45 175 0.01 4.0 11.2

हल : हरात्मक माध्य का निर्धारण

X	$\frac{1}{X}$
1	1.0000
0.5	2.0000
10	0.1000
45	0.0222
175	0.0057
0.01	100.0000
4.0	0.2500
11.2	0.0893
N = 8	$S\left(\frac{1}{X}\right) = 103.4672$

$$\text{H.M.} = \frac{N}{\left(\frac{1}{X}\right)} = \frac{8}{103.4672} = 0.0773$$

विच्छिन्न श्रेणी में हरात्मक माध्य का निर्धारण
(Calculating Harmonic Mean in Discrete Series)

विच्छिन्न श्रेणी में हरात्मक माध्य ज्ञात करने के लिय निम्न सूत्र का प्रयोग होता है :-

$$\text{H.M.} = \frac{N}{\left(f \quad \frac{1}{X}\right)}$$

विधि :

(1) श्रेणी के विभिन्न मूल्यों के व्युत्क्रम (Reciprocals) निकालिये।

(2) व्युत्क्रमों को आवत्ति से गुणा कीजिये और गुणनफलों का योग $\Sigma\left(f \quad \frac{1}{X}\right)$ ज्ञात कीजिये।

(3) इस गुणनफल से योग को मदों की संख्या से भाग दीजिये। इस प्रकार जो उत्तर प्राप्त होगा वह हरात्मक माध्य है।

उदाहरण 47 : निम्न समंकों से हरात्मक माध्य निकालिये।

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10	20	40	15
20	30	50	5
25	50		

हल : हरात्मक माध्य का निर्धारण

अंक X	आवत्ति f	$\frac{1}{X}$	$f \times \frac{1}{X}$
10	20	0.100	2.000
20	30	0.050	1.500
25	50	0.040	2.000
40	15	0.025	0.375
50	5	0.020	0.100
	N = 120		$\Sigma f \times \frac{1}{X} = 5.975$

$$\text{H.P.} = \frac{N}{\left(f \quad \frac{1}{X}\right)} = \frac{120}{5.975} = 20.083$$

अविच्छिन्न श्रेणी में हरात्मक माध्य का निर्धारण
(Calculating Harmonic Mean in Continuous Series)

अविच्छिन्न श्रेणी में हरात्मक माध्य निकालने की विधि वही है जो ऊपर दी गई है केवल इतना अन्तर है कि अखंडित श्रेणी में हम माध्य बिन्दुओं के व्युत्क्रम निकालते हैं।

$$\text{H.M.} = \frac{N}{\left(f \quad \frac{1}{m}\right)}$$

m = मध्य बिन्दु

उदाहरण 48 : निम्न समंकों से हरात्मक माध्य ज्ञात कीजिये।

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10-20	4	40-50	7
20-30	6	50-60	3
30-40	10		

हल : हरात्मक माध्य का निर्धारण

अंक	मध्य बिन्दु m	आवत्ति f	$\frac{1}{m}$	$f \times \frac{1}{m}$
10-20	15	4	0.067	0.268
20-30	25	6	0.040	0.240
30-40	35	10	0.029	0.290
40-50	45	7	0.022	0.154
50-60	55	3	0.018	6.054
		N = 30	$\left[f \quad \frac{1}{m}\right]$	

$$\text{H.M.} = \frac{1}{\left[f \quad \frac{1}{m}\right]} = \frac{30}{1.006} = 29.82$$

हरात्मक माध्य के विशेष प्रयोग
(Special Uses of Harmonic Mean)

$$\left(\frac{30 \quad 20}{2}\right)$$

व्यवहार में हरात्मक माध्य कुछ विशेष परिस्थितियों में ही प्रयुक्त होता है। समयदर, गति, चलन वेग (Velocity), आदि की औसत ज्ञात करने के लिये हरात्मक माध्य विशेष रूप से उपयुक्त है। निम्न उदाहरणों से हरात्मक माध्य के विशेष प्रयोग स्पष्ट हो जायेंगे।

उदाहरण 49 : एक व्यक्ति शहर से हिल स्टेशन 100 किमी. की दूरी गाड़ी द्वारा 30 कि.मी. प्रतिघंटा की रफ्तार से तय करता है। लौटते समय वह 20 किमी. प्रति घंटा की रफ्तार से वही दूरी (100 कि.मी.) तय करता है, उसकी औसत गति ज्ञात कीजिये।

हल : इस प्रश्न में औसत गति निकालने के लिये हरात्मक माध्य का प्रयोग अधिक उपयुक्त होगा।

$$\text{H.M.} = \frac{N}{\frac{1}{X_1} \quad \frac{1}{X_2}}$$

$X_1 = 30, X_2 = 20, N = 2$

\

$$\text{H.M.} \frac{1}{\frac{1}{30} \quad \frac{1}{20}} = \frac{2 \quad 120}{10} = 24 \text{ कि.मी. प्रति घंटा}$$

यदि हम समान्तर माध्य निकालें तो गति 25 कि.मी. प्रति घंटा आयेगी। 25, लेकिन यह उत्तर अशुद्ध है जैसाकि निम्न सारणी से विदित है :-

दूरी (कि.मी. में)	**औसत गति (कि.मी. प्रति घंटा)**	**समय लगा**
जाते समय 100 कि.मी.	30	3 घंटे 20 मिनट
आते समय 100 कि.मी.	20	5 घंटे
योग 200 कि.मी.	8 घण्टे 20 मिनट	

8 घंटे 20 मिनट में कुल दूरी 200 कि.मी. तय की गई। इस प्रकार औसत गति प्रति घंटा 24 कि.मी. $\left(\frac{200 \quad 3}{25}\right)$ आती है।

अत: इस प्रश्न से हरात्मक माध्य ही उपयुक्त है।

उदाहरण 50 : एक हवाई जहाज वर्ग की चारों भुजाओं की उड़ान करता है। वर्ग की एक भुजा 1,000 कि.मी. है। हवाई जहाज वर्ग की चारों भुजाओं को 1,000; 2,000; 3,000; 4,000 कि.मी. प्रति घंटा की रफ्तार से तय करता। हवाई जहाज की औसत गति निकालिये और यह भी सिद्ध कीजिये कि समान्तर माध्य इस प्रश्न में उपयुक्त नहीं हैं।

हल : इस प्रश्न में हरात्मक माध्य द्वारा ही ठीक उत्तर आयेगा।

$$\text{H.M.} = \frac{N}{\frac{1}{X_1} \quad \frac{1}{X_2} \quad \frac{1}{X_3} \quad \frac{1}{X_4}}$$

$X_1 = 1{,}000; X_2 = 2{,}000; X_3 = 3{,}000; X_4 = 4{,}0000, N = 4$

\

$$\text{H.M.} = \frac{N}{\frac{1}{1000} \quad \frac{1}{2000} \quad \frac{1}{3000} \quad \frac{1}{4000}}$$

$$= \frac{1/25}{12{,}000} = \frac{4 \quad 12{,}000}{25} = 1.920 \text{ कि.मी. प्रति घंटा}$$

यदि हम, समान्तर माध्य निकालें तो गति प्रति घंटा इस प्रकार होगी :

$$X = \qquad = \qquad = 2{,}500 \text{ कि.मी.}$$

लेकिन यह गलत है जैसा कि निम्न सारणी से विदित होगा।

1000 N2000 3000 4000

$\frac{1}{X_1} \quad \frac{1}{X_2} \quad \ldots \quad 4$

दूरी (कि.मी. में)	गति कि.मी. प्रति घंटा	समय लगा
1,000	1,000	60 मिनट
1,000	2,000	30 मिनट
1,000	3,000	20 मिनट
1,000	4,000	15 मिनट
योग 4,000		**125 मिनट**

125 मिनट में दूरी तय की = 4,000 कि.मी.

60 मिनट में दूरी तय होगी $= \frac{4{,}000 \quad 60}{125}$ 1,920 कि.मी.

अर्थात् ठीक उत्तर 1,920 कि.मी. प्रति घंटा है न कि 2,500 कि.मी. प्रति घण्टा।

उदाहरण 51 : एक टैक्सी वाला शहर से पहाड़ी की ओर 60 कि.मी. की दूरी 10 कि.मी. प्रति पैट्रों के हिसाब से तय करता हैं वापसी पर वहीं दूरी 15 कि.मी. प्रति गैलन के हिसाब से तय करता है। प्रति गैलन औसत दूरी निकालिये और सिद्ध कीजिये कि आपका उत्तर ठीक है।

हल : इस प्रश्न में भी हरात्मक माध्य उपयुक्त औसत है।

$$\text{H.M.} =$$

$N = 2; X_1 = 10; X_2 = 15$

$$\text{H.M.} = \frac{2}{\frac{1}{10}\quad\frac{1}{15}} = \frac{2\quad 30}{5} = 12 \text{ कि.मी.}$$

यदि हम मध्यक निकालें तो उत्तर = 12.5 कि.मी. प्रति गैलन आयेगा। लेकिन यह उत्तर ठीक नहीं हैं क्योंकि शहर से पहाड़ी की ओर लौटते समय 6 गैलन पैट्रोल खर्च होगा (10 कि.मी. एक गैलन में आता है तो 60 कि.मी. 6 गैलन में आयेगा) तथा लौटते समय 4 गैलन खर्च होगा। (15 कि.मी. 1 गैलन में आता है तो 60 कि.मी. 4 गैलन में आयेगा)। इस प्रकार कुल पैट्रोल 6 + 4 अर्थात् 10 गैलन खर्च हुआ और दूरी 120 कि.मी. तय हुई अर्थात् औसत दूरी प्रति गैलन = 12 कि.मी.।

भारित हरात्मक माध्य

(Weighted Harmonic Mean)

जब विभिन्न गतियों से यह किया जाने वाला फासला भी अलग-अलग होता है तो औसत गति ज्ञान करने के लिय भारित हरात्मक माध्य का प्रयोग किया जाता है। भारित हरात्मक माध्य ज्ञात करते समय विभिन्न दूरियों को भार और गतियों को मूल्य माना जाता है, फिर निम्न सूत्र का प्रयोग किया जाता है :-

$$\text{Weighted H.M.} = \frac{W}{\left(W \quad \frac{1}{X}\right)}$$

विधि : भारित हरात्मक माध्य की गणना-क्रिया इस प्रकार है :-

(1) प्रत्येक मद का व्युत्क्रम ज्ञात किया जाता है।

(2) प्रत्येक मद के व्युत्क्रम को उसके भार से गुणा किया $\frac{120\quad 15}{102}$ जाता है अर्थात् $\left(W \quad \frac{1}{X}\right)$ ज्ञात किया जाता हैं।

(3) इस गुणनफलों का योग करते हैं अर्थात् $\left(W \quad \frac{1}{X}\right)$ ज्ञात करते हैं।

(5) भार के योग को क्रम 3 से प्राप्त योग से विभाजित किया जाता है।

निम्न उदाहरण से यह गणन-क्रिया स्पष्ट हो जायेगी :-

उदाहरण 52 : एक गाड़ी पहले 16 कि.मी., 20 कि.मी. प्रति घंटा के जवाब से दूसरे 20 कि.मी., 40 कि.मी. प्रति घंटा के हिसाब से और अंतिम 10 कि.मी., 15 कि.मी. प्रति घंटा के हिसाब से तय करती है। पूरी यात्रा की माध्य गति क्या है ?

हल : औसत गति ज्ञात करने के लिए भारित हरात्मक माध्य अधिक उपयुक्त रहेगा। विभिन्न दूरियों को भार तथा गतियों का चर माना जायेगा।

गति X	भार	$\times \frac{1}{X}$
20	16	$16 \times \frac{1}{20} = \frac{4}{5}$
40	20	$20 \times \frac{1}{40} \quad \frac{1}{2}$
15	10	$10 \times \frac{1}{15} \quad \frac{2}{3}$

$$\text{Weighted H.M.} = \frac{W}{\left(W \quad \frac{1}{X}\right)} = \frac{46}{\left(\frac{4}{5} \quad \frac{1}{2} \quad \frac{2}{3}\right)}$$

$$= \frac{46}{\frac{59}{30}} = \frac{46 \quad 30}{59} = 23.39 \text{ कि.मी. प्रति घंटा।}$$

हरात्मक के गुण व दोष

(Merits and Demerits of Harmonic Mean)

गुण (Merits)

(1) हरात्मक माध्य श्रेणी के सभी मूल्यों पर आधारित होता है।

(2) इस माध्य का समान्तर माध्य की भांति बीजगणितीय विवेचन हो सकता है।

(3) इस माध्य श्रेणी में बड़े मदों को कम और छोटे मदों को अधिक भार देता है इसलिये, इस माध्य द्वारा प्राप्त परिणाम अन्य सभी माध्यों की अपेक्षा कम होता है।

(4) जिन श्रेणियों में विषमता अधिक हो, वहां इस माध्य का प्रयोग अधिक उपयुक्त है।

(5) समय, गति, दर आदि समस्याओं का अध्ययन करते समय यह माध्य विशेष रूप से प्रयुक्त होता है।

दोष (Demerits) : इस माध्य की गणन क्रिया समान्तर माध्य की अपेक्षा अधिक जटिल है क्योंकि श्रेणी के सभी मदों का व्युत्क्रम निकालना पड़ता है। अत: व्यवहार में इसका प्रयोग कम होता है।

विविध उदाहरण

(Miscellaneous Illustrations)

उदाहरण 53 : 8 परिवारों की मासिक आय इस प्रकार है :-

परिवार	आय (रुपयों में)	$\left(\frac{1}{X}\right)$ परिवार	आय (रुपयों में)
क	70	छ	8
ख	10	च	250
ग	500	छ	8
घ	75	ज	42

इन समंकों से मध्य, गुणोत्तर माध्य तथा हरात्मक माध्य ज्ञात कीजिये और यह सिद्ध कीजिए कि A.M. > G.M. > H.M.

हल : गुणोत्तर माध्य तथा हरात्मक माध्य का निर्धारण

परिवार	आय (रुपयों में) X	log X	1/X
क	70	1.8451	0.0143
ख	10	1.0000	0.1000
ग	500	2.6990	0.0020
घ	75	1.8751	0.0133
छ	8	0.9031	0.1250
च	**250**	**2.3979**	0.0040
छ	8	0.9031	0.1250
ज	42	1.6232	0.0238
N = 8	**SX = 963**	**S Log X = 13.2465**	**S = 0.4074**

$$\therefore \qquad \bar{X} = \frac{963}{8} = 120.375 \text{ रू.}$$

$$\text{G.M.} = \text{A.L.}\left(\frac{\text{Log X}}{\text{N}}\right) = \text{A.L.}\left(\frac{13.2465}{8}\right)$$

$$= \text{A.L. } 1.6558 = 45.27$$

$$\text{H.M.} = \frac{\text{N}}{\left(\frac{1}{\text{X}}\right)} = \frac{8}{0.4074} = 19.637$$

$$\bar{X} = 120.375;\ \text{G.M.} = 45.27;\ \text{H.M.} = 19.637$$

स्त्रियों की ऊंचाई (इंचों में) :	60	61	62	63	64	65	66
स्त्रियों की संख्या :	27	146	435	398	210	128	98

हल : माध्य एवं मध्यका का परिगणन

ऊंचाई X	*f*	(X – 63) *d*	*fd*	*c.f.*
60	27	– 3	– 81	27
61	146	– 2	– 292	173
62	435	– 1	– 435	608
63	398	0	0	1006
64	210	+ 1 $\frac{1442 \quad 1}{2}$	+ 210	1216
65	128	+ 2	+ 256	1344
66	98	+ 3	+ 294	1442
	N = 1442		**Σ*fd* = – 48**	

$$\bar{X} = A + \frac{fd}{N}$$

A = 63, Σ*fd* = – 48, N = 1442

$$\therefore \qquad \bar{X} = 63 - \frac{48}{1442} = 63 - 0.03 = 62.97$$

$$\text{Med} = \text{Size of } \frac{N \quad 1}{2} \text{th item} = \qquad = 7215.5\text{th item}$$

$$\text{Size of 721.5th item} = 63$$

अर्थात् मध्यका मूल्य = 63

उदाहरण 54 : 50 विद्यार्थी एक परीक्षा में बैठे। उत्तीर्ण विद्यार्थियों का परीक्षाफल इस प्रकार है।

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
4	8	7	6
5	10	8	4
6	9	9	3

50 छात्रों के कुल प्राप्तांक का औसत 5.16 था। अनुत्तीर्ण छात्रों के प्राप्तांक का औसत ज्ञात कीजिये।

हल : **औसत प्राप्तांक का परिगणन**

X	f	fX
4	8	32
5	10	50
6	9	54
7	6	42
8	4	32
9	3	27
	N = 40	**Sfx = 237**

50 विद्यार्थियों के प्राप्तांक का औसत = 5.16

50 विद्यार्थियों द्वारा प्राप्त कुल अंक = 50 × 5.16 = 258

उत्तीर्ण विद्यार्थियों द्वारा प्राप्त कुल अंक = 237

शेष 10 विद्यार्थियों द्वारा प्राप्त अंक = 258 – 237 = 21

अनुत्तीर्ण विद्यार्थियों के प्राप्तांक का औसत = = 2.1

उदाहरण 55 : निम्न सूचना से ज्ञान कीजिये कि :

(क) कौन-सी फैक्ट्री अधिक मजदूरी देती है ?

(ख) दोनों फैक्ट्रीयों में काम करने वालों की सम्मिलित औसत मजदूरी क्या है ?

	फैक्ट्री 'क'	फैक्ट्री 'ख'
मज़दूरों की संख्या	250	200
औसत प्रतिदिन मज़दूरी	28 रु.	25 रु.

हल : फैक्ट्री 'क' में मज़दूरों को दी गई धनराशि = 250 × 28 = 7000

फैक्ट्री 'ख' में मज़दूरों को दी गई धनराशि = 200 × 25 = 5000

अर्थात् फैक्ट्री 'क' अधिक मज़दूरी देती है।

$$X_{12} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2}{N_1 \quad N_2}$$

$N_1 = 250; \overline{X}_1 = 18; N_2 = 200; \overline{X}_2 = 25$

$$X_{12} = \frac{(250 \quad 8) \quad (200 \quad 25)}{250 \quad 200} = \qquad = 2.67 \text{ रु.}$$

उदाहरण 56 : नीलोफर एक पत्र को पांच मिनट में टाइप कर सकती है, नीलू उसी को दस मिनट में और नीना पन्द्रह मिनट में। इस प्रकार प्रत्रों को टाइप करने की प्रति घंटे प्रति टाइपिस्ट औसत क्या होगा ?

हल : नीलोफर एक घण्टे में टाइप करेगी = = 12 पत्र

नीलू एक घण्टे में टाइप करेगी = $\frac{60}{10}$ = 6 पत्र

नीनर एक घण्टे में टाइप करेगी = $\frac{60}{15}$ = 4 पत्र

पत्रों को टाइप करने की प्रति घण्टा औसत = $\frac{12 \quad 6 \quad 4}{3}$ = = 7.33

उदाहरण 57 : सांख्यिकी की परीक्षा में 50 विद्यार्थियों द्वारा प्राप्त अंकों का विवरण इस प्रकार है :-

अंक निम्नांकित से अधिक	विद्यार्थियों की संख्या
0	50
10	46
20	40
30	20
40	10
50	3

मध्यका ज्ञात कीजिये। यदि परीक्षा में 60 प्रतिशत विद्यार्थी सफल हो रहे हों तो सफल विद्यार्थियों द्वारा प्राप्त न्यूनतम अंकों का पता लगाइये।

हल : माध्यका का परिगणन

अंक	विद्यार्थियों की संख्या	*c.f.*
0-10	4	4
10-20	6	10
20-30	20	30
30-40	10	40
40-50	7	47
50-60	3	50

$$M = \text{Size of } \frac{N}{2}\text{th item} = \frac{50}{2} = 25\text{th item}$$

मध्यका 20-30 वर्ग में स्थित है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 20; N/2 = 25; *c.f.* = 10; *f* = 20; *i* = 10

$$M = 20 + \qquad \times 10 = 20 + 7.5 = 27.5$$

यदि 60 प्रतिशत विद्यार्थी परीक्षा में सफल रहे है तो सफल विद्यार्थियों द्वारा प्राप्त निम्नतम अंकों का पता P_{60} द्वारा लगाया जा सकता है।

$$P_{60} = \text{Size of} \qquad \text{th item} = \frac{60 \quad 50}{100} = 30\text{th item}$$

P_{60} वर्ग 20-30 में स्थित है।

$$M = L + \qquad \times i$$

$$= 20 + \qquad \times 10 = 20 + 10 = 30$$

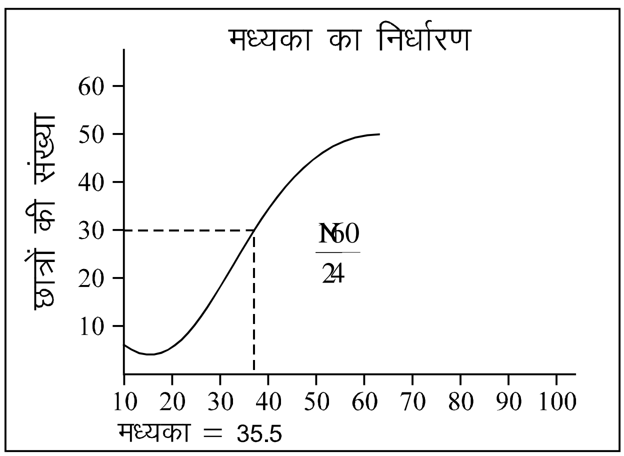
उदाहरण 58 : किसी कक्षा में 60 छात्रों के एक विषय में प्राप्तांक नीचे दिये गये हैं। इस सूचना को रेखाचित्र द्वारा प्रदर्शित कीजिये और रेखाचित्र में मध्यका अंक प्राप्त कीजिये :-

अंक	छात्रों की संख्या
10 से कम	4
20 से कम	8
30 से कम	25
40 से कम	34
50 से कम	41
60 से कम	46
70 से कम	50
100 से कम	60

हल :

$$M = \text{Size of} \quad \text{th item} = \frac{60}{2} = 30\text{th item}$$

उपर्युक्त समंकों को रेखाचित्र पर अंकित किया गया है और माध्यका अंक ज्ञात किये गये हैं। मध्यका अंक 35.5 है।



उदाहरण 59 : मोहन ने चार दिन तक कार से यात्रा की। वह प्रतिदिन 10 घंटे चला। पहले दिन उसने 45 किलोमीटर प्रति घंटे, दूसरे दिन 40 किलोमीटर प्रति घंटे, तीसरे दिन 38 किलोमीटर प्रति घंटे और चौथे दिन 37 किलोमीटर प्रति घंटे की चाल से यात्रा की। उसकी औसत चाल बताइये।

हल : क्योंकि वह चार दिन समान समय अर्थात दस-दस घंटे चलता है इसलिये उपर्युक्त माध्य समान्तर माध्य होगा।

$$\overline{X} = \frac{45 \quad 40 \quad 38 \quad 37}{4} = \qquad = 40$$

अर्थात् औसत चाल = 40 किलोमीटर प्रति घंटा।

इस हल की शुद्धता की जांच इस प्रकार की जा सकती है।

पहले दिन 10 घण्टे में चला $= 45 \times 10 = 450$ किमी.

दूसरे दिन 10 घण्टे में चला $= 40 \times 10 = 400$ किमी.

तीसरे दिन 10 घण्टे में चला $= 38 \times 10 = 380$ किमी.

चौथे दिन 10 घण्टे में चला $= 37 \times 10 = 370$ किमी.

अर्थात् 40 घण्टों में चला = 1600 किमी.

1 घण्टे में चला $= \frac{1600}{40} = 40$ किमी.

उदाहरण 60 : निम्न श्रेणी से समान्तर माध्य तथा मध्यका की गणना कीजिये :-

साप्ताहिम मजदूरी (रुपयों में)	मजदूरों की संख्या f	साप्ताहिक मज़दूरी (रुपयों में)	मज़दूरों की संख्या f
100-105	200	130-135	410
105-110	210	135-140	320
110-115	230	140-145	280
115-120	320	145-150	210
120-125	350	150-155	160
125-130	520	155-150	90

हल : समान्तर माध्य तथा मध्यका का परिगणन

साप्ताहिक मज़दूरी (रु. में)	मध्य बिन्दु m	f	d	$(m - 127.5)$ fd	संचयी आवत्ति
100-105	102.5	200	– 5	– 1,000	200
105-110	107.5	210	– 4	– 840	410
110-115	112.5	230	– 3	– 690	640
115-120	117.5	320	– 2	– 640	960
120-125	122.5	350	– 1	– 350	1,310
125-130	127.5	520	0	0	1,830
130-135	132.5	410	+ 1	+ 410	2,240
135-140	137.5	320	+ 2	+ 640	2,560
140-145	142.5	280	+ 3	+ 840	2,840
145-150	147.5	210	+ 4	+ 840	3,050
150-155	152-5	160	+ 5	+ 800	3,210
155-160	157.5	90	+ 6	+ 540	3,300
		N = 3,300		**Σfd = 550**	

$$\overline{X} = A + \frac{fd}{N} \times i = 127.5 + \frac{550}{3,300} \times 5$$

$$= 127.5 + 0.833 = 127.33 \text{ रु.}$$

$$M = \text{Size of } \frac{N}{2}\text{th item} = \frac{3,300}{2} = 1,650\text{th item}$$

मध्यका 125-130 वर्ग में स्थित है।

$$M = L + \frac{N/2 - c.f.}{f} \times i$$

$$= 125 + \frac{1650 - 1310}{520} \times 5 = 125 + 3.27 = 128.27 \text{ रु.}$$

उदाहरण 61 : निम्न आवत्ति वितरण से भूयिष्ठक ज्ञात कीजिये :-

अनुपस्थिति के दिन	विद्यार्थियों की संख्या	अनुपस्थिति के दिन	विद्यार्थियों की संख्या
0-05	29	25-30	10
05-10	195	30-35	6
10-15	241	35-40	3
15-20	117	40-45	2
20-25	52		

हल : **भूयिष्ठक का परिगणन**

अनुपस्थिति के दिन	f	अनुपस्थिति के दिन	f
0-05	29	25-30	10
05-10	195	30-35	6
10-15	241	35-40	3
15-20	117	40-45	2
20-25	52		

निरीक्षण से स्पष्ट है कि भूयिष्ठक वर्ग 10-15 है।

$$M_0 = L + \frac{{}_1}{{}_1 \quad {}_2} \times i$$

$L = 10; D_1 = (f_1 - f_0) = (241 - 195) = 46$

$D_1 = (f_1 - f_2) = (241 - 117) = 124; i = 5$

\

$$M_0 = 10 + \frac{46}{46 \quad 124} \times 5 = 10 + 1.35 = 11.5$$

उदाहरण 62 : निम्न समंकों से मध्यका तथा भूयिष्ठक निकालिये :-

खेतों का एकड़ में मध्यवर्ती माप :	10	20	30	40	50	60
खेतों की संख्या :	7	12	17	29	31	5

हल : क्योंकि हमें मध्यवर्ती माप दिया है इसलिये विभिन्न श्रेणियों ज्ञात करनी होगी। प्रथम श्रेणी 5-15 होगी, दूसरी 15-25 आदि।

मध्यका तथा भूयिष्ठक का परिगणन

खेतों की माप	f	संचयी आवत्ति
5-15	4	7
15-25	12	19
25-35	17	36
35-45	29	65
45-55	31	96
55-65	5	101
65-75	3	104
	N = 104	

मध्यका : $\quad M = \text{Size of } \frac{N}{2} \text{th item} = \frac{104}{2} = 52\text{th item}$

मध्यका वर्ग 35-45 में स्थित है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i$$

$L = 35; N/2 = 52; c.f. = 36; f = 29; i = 10$

\ $\quad M = 35 + \qquad \times 10 = 35 + 5.52 = 40.52$

भूयिष्ठक : $\quad M = L + \qquad \times i$

यद्यपि 31 अधिकतम आवत्ति है लेकिन समूहीकरण तथा विश्लेषण सारणी बनाने पर यह ज्ञात होता है कि भूयिष्ठक 35-40 वर्ग में स्थित है।

$L = 35; D_1 = |f_1 - f_0| = (29 - 17) = 12; D_2 = (29 - 31) = 2; i = 10$

$$Z = 35 + \frac{12}{12 \quad 2} \times 10 = 3.5 + 8.57 = 43.47$$

उदाहरण 63 : निम्न मज़दूरी वितरण (Wage Distribution) का मध्यका और भूयिष्ठक क्रमश: 33.5 रुपये और 34 रुपये हैं किन्तु सारणी में तीन आवत्तियां अज्ञात है। अज्ञात आवत्तियों को ज्ञात कीजिये :

मजदूरी (रुपयों में) :	0-10	10-20	20-30	30-40	40-50	50-60	60-70	**योग**
आवत्तियां :	4	14	?	?	?	6	4	**230**

हल : मध्यका तथा भूयिष्ठक का मान क्रमश: 33.5 तथा 34 $\frac{52 \quad 36}{29}$ है अर्थात् दोनों 30-40 वर्ग में स्थित हैं। 30-40 की आवत्ति f_1 इससे पहले वर्ग अर्थात 20-30 की आवत्ति को f_0 तथा इसके बाद के वर्ग अर्थात् 40-50 की आवत्ति को f_2 मानकर निम्न आवत्ति सारणी बनाई जायेगी :

मजदूरी (रुपयों में)	आवत्तियां	संचयी आवत्ति
0-10	4	7
10-20	16	20
20-30	f_0	$10 + f_0$
30-40	f_1	$20 + f_0 + f_1$
40-50	f_2	$20 + f_0 + f_1 + f_2$
50-60	6	226
60-70	4	130

$$f_0 + f_1 + f_2 = 230 - (4 + 16 + 6 + 4) = 200$$

$$f_2 = 200 - (f_0 + f_1) = 200 - f_0 - f_1$$

$$M_0 = L + \frac{1}{1 \quad 2} \times i$$

$$30 = 30 + \frac{f_1 \quad f_0}{2f_1 \quad f_0 \quad (200 \quad f_0 \quad f_1)} \times 10$$

$$\frac{4}{10} = \frac{f_1 \quad f_0}{2f_1 \quad f_0 \quad 203 \quad f_0 \quad f_1}$$

$$10f_1 - f_0 = = 8f_1 - 3f_0 - 800 + 4f_0 - 4f_1$$
$$-2f_1 - 10f_0 = -800$$
$$f_1 + 5f_0 = 400 \qquad ...(i)$$

$$M = L + \frac{N/2 \quad c.f.}{f} + i$$

$$33.5 = 30 + \qquad \times 10$$

$$3.5f_1 = 950 - 10f_0 \text{ (2से गुणा करने पर)}$$
$$7f_1 = 1900 - 20f_0 \qquad ...(ii)$$

(i) तथा (ii) समीकरणों को हल करने पर f_0 तथा f_1 का मान ज्ञान हो जायेगा।

$$f_1 + 5f_0 = 400 \qquad ...(i)$$
$$7f_1 + 20f_0 = 1900 \qquad ...(ii)$$

पहले समीकरण को 4 से गुणा कीजिये

$$4f_1 + 20f_0 = 1600$$
$$7f_1 + 20f_0 = 1900$$
$$-3f_1 = -300$$

\ $$f_1 = 100$$

समीकरण (i) में f_1 का मान रखने पर

$$100 + 5f_0 = 400$$
$$5f_0 = 400 - 100$$
$$f_0 = \qquad = 60$$
$$f_2 = 200 - f_0 - f_1$$
$$= 200 - 60 - 100 = 40$$

अर्थात् $$f_0 = 60; f_1 = 100, f_2 = 40$$

उदाहरण 64 : एक प्रिंसिपल ने अपने विद्यालय में 1987 में विद्यार्थियों की संख्या 15% बढ़ाई। 1988 में फिर 5% बढ़ाई। परन्तु 1989 में उनको कुछ कठिनाइयों के कारण 20 प्रतिशत संख्या कम करनी पड़ी। इस तरह उनके विद्यालय में 1987 से पहले की संख्या फिर हो गई। क्या आप इस कथन से सहमत हैं ? कारण बताइये।

हल : हम इस कथन से सहमत नहीं हैं। मान लीजिये 1986 में विद्यालय में 100 विद्यार्थी थें।

1987 में वद्धि की दर 15% अर्थात् विद्यार्थियों की संख्या = 115

1988 में वद्धि की दर 5% अर्थात् विद्यार्थियों की संख्या = $115 + \frac{115 \quad 5}{100}$

= 120.75 या 121

1989 में 20 प्रतिशत कम करने पर विद्यार्थियों की संख्या = 121 – = 121 – 24.2 = 96.8

अर्थात् 1989 में विद्यार्थियों की संख्या 1987 की तुलना में कम हो गई।

गुणोत्तर माध्य द्वारा भी इस परिणाम की पुष्टि की जा सकती है।

वर्ष	प्रतिशत वद्धि	X	log X
1987	+ 15	115	2.0607
1988	+ 5	105	2.0212
1989	– 20	80	1.9031
			S log X = 5.98552

$$\text{G.M.} = \text{AL} \qquad = \text{AL}\left(\frac{5.985}{3}\right) = \text{AL } 1.995 = 98.86$$

उदाहरण 65 : 100 पौधों पर पूर्ण विकसित टमाटरों की गणना करके निम्न परिणाम प्राप्त हुए :

पौधे	टमाटर	पौधे	टमाटर
2	0	12	6
5	1	8	7
7	2	6	8
11	3	3	9
18	4	4	10
24	5		

(*i*) कुल कितने टमाटर थे ?

(*ii*) प्रति पौधा टमाटर की औसत संख्या क्या है ?

(*iii*) टमाटरों की बहुलक संख्या (Modal Number) क्या थी ? $\left(\frac{\log X}{N}\right)$

हल : **मध्यका तथा भूयिष्ठक का परिगणन**

X	*f*	*f*X
0	2	0
1	5	5
2	7	14
3	11	33
4	18	72
5	24	120
6	12	72
7	8	56
8	6	48
9	4	36
10	3	30
	N = 100	**S*fx* = 486**

(*i*) टमाटर की कुल संख्या = 486

(*ii*) प्रति पौधा टमाटर की औसत संख्या $\frac{486}{100} = 4.86$

(*iii*) निरीक्षण द्वारा स्पष्ट है कि बहुलक 5 है।

उदाहरण 66 : दिल्ली में टैक्सी का भाड़ा पहले किलोमीटर पर एक रुपया और उसके बाद 60 पैसा प्रति किलोमीटर है। भाड़ा प्रत्येक किलोमीटर के आरम्भ में ही मीटर पर आ जाता है। इससे प्रत्येक यात्री को पूरे किलोमीटर का भाड़ा देना पड़ता है। बताइये $3\frac{3}{4}$ किलोमीटर का औसत भाड़ा क्या होगा ?

हल : पहले किलोमीटर का भाड़ा = 100 पैसे

दूसरे किलोमीटर का भाड़ा = 60 पैसे

तीसरे किलोमीटर का भाड़ा = 60 पैसे

3 या 3 से कम किलोमीटर का भाड़ा = 220 पैसे

वास्तविक दूरी = $\frac{11}{4}$ किलोमीटर

$\frac{11}{4}$ किलोमीटर का भाड़ा = 220 पैसे

इसलिये एक किलोमीटर का भाड़ा = $220 \times \frac{4}{11} = 80$ पैसे

अर्थात् औसत भाड़ा = 80 पैसे प्रति किलोमीटर।

उदाहरण 67 : एक कारखाने में काम करने वाले कर्मचारियों का वार्षिक वेतन का समान्तर माध्य 500 रु. है। पुरुषों को दिये गये वेतन का समान्तर माध्य 520 रु. तथा महिलाओं को दिये गये वेतन का समान्तर माध्य 420 रु. है। कारखाने में काम करने वाले पुरुषों तथा महिलाओं की प्रतिशत ज्ञात कीजिये।

हल : मान लीजिये N_1 प्रतिशत पुरुष तथा N_2 प्रतिशत महिलाओं को व्यक्त करता है।

$$N_1 + N_2 = 100$$

हमें दिया गया है : $\overline{X}_{12} = 500, \overline{X}_1 = 520, \overline{X}_2 = 420$

$$X_{12} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2}{N_1 \quad N_2}$$

$$500 = \frac{N_1\,(520) \quad N_2\,(420)}{100}$$

$$50{,}000 = 520\,N_1 + (100 - N_1)\,420 \qquad \text{क्योंकि } N_1 + N_2 = 100$$

$$50{,}000 = 520\,N_1 + 42000 - 420\,N_1 \qquad \backslash\ N_2 = (100 - N_1)$$

$$100\,N_1 = 8000$$

$$N_1 = 80 \text{ तथा } N_2 = (100 - 80) = 20$$

अर्थात् पुरुष = 80 प्रतिशत तथा महिलायें = 20 प्रतिशत।

उदाहरण 68 : सिले हुए वस्त्रों का व्यापार करने वाली एक फर्म पुरुषों तथा स्त्रियों के वस्त्र (कमीजें) बनाती हैं। इसका औसत लाभ बिक्री का 6 प्रतिशत है; पुरुषों के वस्त्रों पर औसत लाभ बिक्री का 8 प्रतिशत है और स्त्रियों के वस्त्रों का उत्पादन 60 प्रतिशत के बराबर है। स्त्रियों के वस्त्रों पर औसत लाभ बिक्री-रुपया क्या है।

हल : मान लीजिये कुल कमीजें 100 हैं।

स्त्रियों की कमीजें = 60 (60 प्रतिशत)

पुरुषों तथा स्त्रियों की कमीजों का लाभ = 6 प्रतिशत

पुरुषों की कमीजों पर लाभ = 8 प्रतिशत।

पुरुषों की कमीजों पर कुल लाभ = 40 × 3.20 रु.

स्त्रियों की कमीजों पर लाभ = 6 – 3.20 = 2.80 रु.

स्त्रियों की 60 कमीजों पर लाभ = 2.8

स्त्रियों की 100 कमीजों पर लाभ = $\frac{2.8}{60} \times 100 = 4$ प्रतिशत

उदाहरण 69 : निम्नलिखित श्रेणी से मध्यका एवं समान्तर माध्य का परिकलन कीजिये :-

05 व्यक्ति 5 रु. से कम पाते हैं।

12 व्यक्ति 10 रु. से कम पाते हैं।

22 व्यक्ति 15 रु. से कम पाते हैं।

30 व्यक्ति 20 रु. से कम पाते हैं।

36 व्यक्ति 25 रु. से कम पाते हैं।

40 व्यक्ति 30 रु. से कम पाते हैं।

हल : हमें संचयी आवत्ति वितरण दिया गया है। समान्तर माध्य ज्ञात करने के लिये सर्वप्रथम इसे साधारण आवत्ति वितरण में परिवर्तित करना होगा।

समान्तर माध्य तथा मध्यका का परिगणन

मज़दूरी (रु. में)		**(*m* – 12.5)/5**			
	m	***f***	***d***	***fd***	***cf***
0-5	2.5	5	28 12 – 2	– 10	5
5-10	7.5	7	3000 – 1	– 7	12
10-15	12.5	10	0	0	22
15-20	17.5	8	+ 1	+ 8	830
20-25	22.5	6	+ 2	+ 12	36
25-30	27.5	4	+ 3	+ 12	40
		N = 40		**S*fd* = 15**	

$$\overline{X} = A + \frac{fd}{N} \times i = 12.5 + 1.875 = 14.375$$

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th item} = \frac{40}{2}\ 20\text{th item}$$

मध्यका 10-15 वर्ग में स्थित है।

$$\text{Med.} = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 10; N/2 = 20; $c.f.$ = 12; f = 10; i = 5

\ $\text{Med.} = 10 + \quad \times 5 = 10 + 4 = 14$

उदाहरण 70 : 150 छात्रों का समान्तर माध्य वजन 56 किलोग्राम हैं इनमें से 100 छात्रों का वनज निम्न बारंबारता बंटन में दिया हुआ है। शेष 52 छात्रों का समाना माध्य वजन ज्ञात कीजिये।

वजन (किलोग्राम में)	**:**	30-40	40-50	50-60	60-70	70-80
बारंबारता	**:**	10	30	40	10	10

हल : **समान्तर माध्य का परिगणन**

वजन (किलोग्राम में)	m	बारम्बारता f	$(m-55)/10$ d	fd
30-40	35	10	– 2	– 20
40-50	45	30	– 1	– 30
50-60	55	40	0	0
60-70	65	10	+ 1	+ 10
70-80	75	10	+ 2	+ 20
		N = 100		**S*fd* = – 20**

$$= A + \frac{fd}{N} \times i = 55 - \frac{22}{100} \times 10 = 55 - 2 = 53$$

$$\overline{X}_{12} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2}{N_1 \quad N_2}$$

$$56 = \frac{(100 \quad 53) \quad (150 \quad 100)\,\overline{X}_2}{150}$$

$$8,400 = 5300 + 50\,\overline{X}_2$$

$$50\,\overline{X}_2 = 3,100 \qquad \backslash\ \overline{X}_2 = 62$$

अर्थात् शेष 50 छात्रों का समान्तर माध्य वजन 62 किलोग्राम होगा।

उदाहरण 71 : निम्न आंकड़ों से भूयिष्ठक ज्ञात कीजिये :-

प्राप्तांक	विद्यार्थियों की संख्या	प्राप्तांक	विद्यार्थियों की संख्या
10 से कम	5	$\overline{X}$ 60 से कम	86
20 से कम	6	70 से कम	96
30 से कम	24	80 से कम	99
40 से कम	46	90 से कम	100
50 से कम	67		

हल : सर्वप्रथम संचयी आवत्तियों से साधारण आवत्तियां ज्ञात कीजिये।

प्राप्तांक	:	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
विद्यार्थियों की संख्या	:	4	2	18	22	21	19	10	3	1

प्रश्न से यह स्पष्ट नहीं है कि भूयिष्ठक किस वर्ग में स्थित है। अत: एक समूहीकरण तथा विश्लेषण सारण की रचना की जायेगी।

समूहीकरण सारणी

प्राप्तांक	आवत्ति I	II	III	IV	V	VI
0-10	4	6	20			
10-20	2			24	42	61
20-30	18	40	43			
30-40	22			62		
40-50	21	40	29		61	32
50-60	19			14		
60-70	10	13	4			
70-80	3					
80-90	1					

विश्लेषण सारणी

स्तंभ संख्या	20-30	30-40	40-50	50-60
I		1		
II	1	1	1	1
III		1	1	
IV		1	1	1
V			1	1
VI	1	1	1	
योग	**2**	**5**	**5**	**3**

प्रत्यक्ष रीति द्वारा इस श्रेणी से भूयिष्ठक ज्ञात नहीं किया जा सकता। निम्न सूत्र द्वारा भूयिष्ठक ज्ञात किया जायेगा :

$$\text{Mode} = 3\ \text{Median} - 2\ \text{Mean}$$

समान्तर माध्य तथा मध्यका का परिगणन

प्राप्तांक	*m*	*f*	**(*m* – 45)/10**	*fd*	*c.f.*
0-10	5	4	– 4	– 16	4
10-20	15	2	– 3	– 6	6
20-30	35	18	– 2	– 36	24
30-40	35	22	– 1	– 22	46
40-50	45	21	0	0	67
50-60	55	19	$\frac{50 \quad 46}{21}$ + 1	+ 19	86
60-70	65	10	+ 2	+ 20	96
70-80	75	3	+ 3	+ 9	99
80-90	85	1	+ 4	+ 4	100
		N = 100		**S*fd* = – 28**	

$$\overline{X} = A - \frac{fd}{N} \times i = 45 - \frac{28}{100} \times 10 = 42.2$$

$$M = \text{Size of } \frac{100}{2} = 50\text{th item}$$

मध्यका 40-50 वर्ग में स्थित है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i = 40 + \qquad \times 10 = 40 + 1.9 + 41.9$$

$$\text{Mode} = 3\,(41.9) - 2\,(42.2) = 125.7 - 84.4 = 41.3$$

उदाहरण 72 : एक आवत्ति वितरण में अवलोकनों की संख्या 100 है। वितरण की मध्यका 30 है। अज्ञात आवत्तियां ज्ञात कीजिये।

प्राप्तांक	:	0-10	10-20	20-30	30-40	40-50	50-60
विद्यार्थियों की संख्या	:	10	?	25	30	?	10

हल : मान लीजिये किस 10-20 वर्ग की आवत्ति f_1 है तथा 40-50 वर्ग की आवत्ति f_2 है।

ज्ञात आवत्तियों का योग $= 10 + 25 + 30 + 10 = 75$

अज्ञात आवत्तियों का योग अर्थात् $f_1 + f_2 = 100 - 75 = 25$

$$M = L + \qquad \times i$$

$$M = \text{Size of} \quad \text{th item} = \frac{100}{2} = 50\text{th item}$$

मध्यका 30-40 वर्ग में स्थित है।

$$30 = 30 + \frac{50 \quad (10 \quad f_1 \quad 25)\,30}{30} \times 10$$

$$900 = 900 + (50 - 35 - f) = 10$$

$$150 - 10f_1 = 0$$

$$f_1 = 15$$

क्योंकि $f_1 + f_2 = 25$

\ $$f_2 = 25 - 15 = 10$$

अर्थात् वर्ग 10-20 की आवत्ति 15 तथा वर्ग 40-50 की आवत्ति 10 है।

उदाहरण 73 : (क) रविवार को छोड़कर एक सप्ताह की औसत वर्षा 10 सें.मी. थी। रविवार को भारी वर्षा के कारण सप्ताह की औसत 15 सें.मी. हो गई। रविवार को कितनी वर्षा हुई।

हल : छ: दिन की औसत वर्षा = 10 सें.मी.

छ: दिन की कुल वर्षा = $10 \times 6 = 60$ सें.मी.

7 दिन की औसत वर्षा = 15 सें.मी.

7 दिन की कुल वर्षा = $15 \times 7 = 105$ सेंमी.

रविवार वाले दिन वर्षा = $105 - 60 = 45$ सें.मी.

(ख) **उस जनसंख्या की औसत वद्धि दर निर्धारति कीजिये , जिसमें पहले दशम में 20% दूसरे दशक में 30% तथा तीसरे दशम में 40% वद्धि हुई।**

हल : गुणोत्तर माध्य द्वारा जनसंख्या की औसत वद्धि दर ज्ञात की जा सकती है।

दशक के अंत में जनसंख्या (X)	Log (X)
120	2.0792
130	2.1139
140	2.1461
	Slog X = 6.3392

$$\text{G.M.} = \text{AL} \qquad = \text{AL}\left(\frac{6.3992}{3}\right) = \text{AL}\ (2.1131) = 129.7$$

अर्थात् जनसंख्या की औसत वद्धि दर 29.7 प्रतिशत प्रति दशक है।

उदाहरण 74 : विधि की परीक्षा में 50 विद्यार्थियों द्वारा प्राप्तांकों का वितरण इस प्रकार है :-

अंक (निम्नलिखित से अधिक)	विद्यार्थियों की संख्या	अंक (निम्नलिखित से अधिक)	विद्यार्थियों की संख्या
0	50	30	20
10	46	40	10
20	40	50	3

मध्यका-अंकों की गणना कीजिये।

हल : सर्वप्रथम संचयी आवत्ति वितरण को साधारण आवत्ति वितरण में बदले और इसके बाद मध्यका का परिगणन करें।

अंक	**विद्यार्थियों की संख्या *f***	***c.f.***	**अंक**	**विद्यार्थियों की संख्या *f***	***c.f.***
0-10	4	4	30-40	10	40
10-20	6	10	40-50	7	47
20-30	20	30	50-60	3	50

$$\text{M= Size of } \frac{N}{2} \text{th item} = \frac{50}{2} = \text{25th item}$$

अर्थात् मध्यका 20-30 वर्ग में स्थित है।

$$M = L + \frac{N/2 \quad c.f.}{f} \times i = 20 + \qquad \times 10 = 20 + 7.5 - 27.5$$

उदाहरण 75 : निम्नलिखित समंकों से समान्तर माध्य एवं भारित समान्तर माध्य ज्ञात कीजिये।

वस्तुयें	**उपभोग की मात्रा (किलोग्राम)**	**भाव प्रति किलो (रु. में)**
आटा	10	2.5
ईंधन	50	1.4
शक्कर	2	3.5
तेल	3	4.2

$$X \frac{25 \quad 10}{20}$$

हल : समान्तर माध्य तथा भारित समान्तर माध्य का परिगणन

वस्तुयें	**उपभोग की मात्रा**	**भाव**	***wx***
आटा	10	2.5	2.5
ईंधन	50	1.5	75
शक्कर	2	3.5	7
तेल	3	4.5	13.5
	S*w* = 65	**S*x* =12**	**S*wx* = 120.5**

$$= \frac{X}{N} = \frac{12}{4} = 3 \text{ रु.}$$

$$\overline{X}_w = \frac{wx}{w} = \frac{120.5}{65} = 1.85$$

उदाहरण 76 : निम्नलिखित श्रेणी में अज्ञात मूल्य ज्ञात कीजिये; यदि श्रेणी का समान्तर माध्य 8.52 रु. हैं:-

मज़दूरी	:	4	6	?	10	12	14
व्यक्तियों की संख्या	:	5	10	12	15	6	2

हल : मान लीजिये कि अज्ञात मूल्य का मान Z है। अज्ञात मूल्य का परिगणन

x	*f*	*fx*
4	5	20
6	10	60
Z	12	12Z
10	15	150
12	6	72
14	2	28
	N = 50	**S*fx* = 330 + 12Z**

$$= \frac{fx}{N}$$

$$8.52 = \frac{330 \quad 12Z}{50}$$

$$330 + 12Z = 426$$

$$12Z = 426 - 330 = 96 \text{ या } Z = 8$$

उदाहरण 77 : पुरुषों तथा महिलाओं के एक मिश्रित समूह की माध्य आयु 25 वर्ष है यदि समूह में पुरुषों की माध्य आयु 26 वर्ष तथा महिलाओं की 21 वर्ष है तो समूह में पुरुषों एवं महिलाओं का प्रतिशत ज्ञात कीजिये।

हल : हमें दिया गया है :

$$\bar{X}_{12} = 25, \bar{X}_1 = \bar{X}_1 = 21$$

मान लीजिये पुरुषों का प्रतिशत N_1 है तथा महिलाओं का N_2

अर्थात् $N_1 + N_2 = 100$

$$\bar{X}_{12} = \frac{N_1\bar{X}_1 \quad N_2\bar{X}_2}{N_1 \quad N_2}$$

$$25 = \frac{N_1(26) \quad (100 \quad N_1)\,21}{100}$$

$$2500 = 26N_1 + 2100 - 21N_1$$

$$5N_1 = 2500 - 2100 \text{ या } N_1 = 80$$

$$N_2 = 100 - 80 = 20$$

अर्थात् पुरुषों का प्रतिशत = 80

स्त्रियों का प्रतिशत = 20

उदाहरण 78 : निम्नलिखित समंकों से समान्तर माध्य, मध्यका तथा बहुलक ज्ञात कीजिये :-

मूल्य	:	0-9	10-19	20-29	30-39	40-419	50-59	60-96
बारम्बारता	:	8	10	15	20	15	10	8

वितरण की प्रकृति पर भी व्याख्या कीजिये।

हल : **समान्तर माध्य, मध्यका तथा बहुलक का परिगणन**

मूल्य	आवत्ति f	*m.p.* m	$(m - 34.5)/10$ d	fd	*c.f.*
0-19	8	4.5	– 3	– 24	8
10-19	10	14.5	– 3	– 20	18
20-29	15	24.5	– 1	– 15	33
30-39	20	34.5	0	0	53
40-49	15	44.5	+ 1	+ 25	68
60-69	8	64.5	+ 3	+ 24	86
	N = 86			**S*fd* = 0**	

समान्तर माध्य : $\overline{X} = A + \frac{fd}{N} \times i$

A = 34.5; S*fd* = 0; N = 76; i = 10

\ $\overline{X} = 34.5 + \frac{0}{86} \times 10 = 34.5$

मध्यका : M = Size of $\frac{N}{2}$th item = $\frac{86}{2}$ = 43rd item. मध्यका 30-39 वर्ग में स्थित है। लेकिन इसकी वास्तविक सीमा 29.5-39.5 है।

$$M = L + \frac{N/2 - c.f.}{f} \times i$$

L = 29.5; N/2 = 42; *c.f.* = 30; f = 20; i = 10

$$M = 29.5 + \qquad \times 10$$

$$= 29.5 + 5 = 34.5$$

बहुलक : $M_0 = L + \qquad \times i$

बहुलक 30-39 वर्ग में स्थित है लेकिन इसकी वास्तविक सीमा 29.5-39.5 है।

L = 29.5; D_1 = 20 – 15 = 5; D_2 = 20 – 15 = 5; i = 10

$$M_0 = 29.5 + \frac{5}{5 \quad 5} \times 10$$

$$= 29.5 + 5 = 34.5$$

उदाहरण 79 : निम्न समंकों से समान्तर माध्य मध्यका तथा भूयिष्ठक ज्ञात कीजिये :-

दैनिक मज़दूरी (रु. में)	मज़दूरों की संख्या	दैनिक मज़दूरी (रु. में)	मज़दूरो की संख्या
12.5-17.5	2	37.5-42.5	4
17.5-22.5	22	42.5-47.5	6
22.5-27.5	10	47.5-52.5	1
27.5-32.5	14	52.5-57.5	1
32.5-37.5	3		

हल : समान्तर माध्य, मध्यका तथा भूयिष्ठक का परिगणन

दैनिक मज़दूरी	*m*	*f*	(*m* – 35)/5 *d*	*fd*	*c.f.*
12.5-17.5	15	2	– 4	– 8	2
17.5-22.5	20	22	– 3	– 66	24
22.5–27.5	25	10	– 2	– 20	34
27.5-32.5	30	14	– 1	– 14	48
35.5-37.5	35	3	0	+ 0	51
37.5-42.5	40	4	+ 1	+ 4	55
42.5-47.5	45	6	+ 2	+ 12	61
47.5-52.5	50	1	+ 3	+ 3	62
52.5-57.5	55	1	+ 4	+ 4	63
		N = 62		**S*fd* = – 85**	

माध्य का परिगणन :

$$\overline{X} = A + \frac{fd}{N} \times i$$

A = 35; S*fd* = – 85; N = 63; *i* = 5

\

$$\overline{X} = 35 - \frac{85}{63} \times 5 = 35 - 6.75 = 28.25$$

मध्यका का परिगणन :

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th item } \frac{63}{2} = 31.5\text{th item}$$

मध्यका 22.5-27.5 वर्ग में है।

$$\text{Med.} = L + \frac{N/2 \quad c.f.}{f} \times i$$

L = 225; N/2 = 31.5; *c.f.* = 24; f = 10; *i* = 5

$$\text{Med.} = 22.5 + \qquad \times 5 = 22.5 + 3.75 = 26.25$$

भूयिष्ठक का परिगणन :

$$M_0 = L + \qquad \times i$$

प्रश्न देखकर यह करना मुश्किल है भूयिष्ठक वर्ग कौन-सा है। अत: समूहीकरण तथा विश्लेषण सारणी बनायें।

समूहीकरण सारणी

I	II	III	IV	V	VI
2	24	32	34		
22				46	27
10	24	17			
14			21		
3	7	10			
4				13	11
6	7	2	9		
1					

विश्लेषण सारणी

स्तंभ संख्या 12.5-17.5	17.5-22.5	22.5-27.5	चर मूल्य 27.5-32.5	32.5-37.5	
I		1			
II	1	1	1	1	
III		1	1		
IV		1	1		
V	1	1	1	1	
VI			1	1	1
योग	**2**	**5**	**5**	**3**	**1**

यह दो भूयिष्ठक वाली श्रेणी है अर्थात् निम्न सूत्र से भयिष्ठक ज्ञात किया जायेगा।

$$M_0 = 2\text{ Med} - 2\,\overline{X}$$
$$= 3\,(26.25) - 2\,(28.25) = 78.75 - 56.5 = 22.25$$

उदाहरण 80 : एक व्यक्ति तीन प्रकार की पैंसिलें खरीदता हैं सम्बद्ध आंकड़े नीचे दिये गये हैं :-

प्रकार	प्रति पैंसिल मूल्य (रु.)	व्यय किया धन (रु.)
A	1.00	50
B	1.50	30
C	2.00	20

प्रति पैंसिल औसत मूल्य की गणना कीजिये।

हल : व्यय को प्रति पैंसिल मूल्य से भाग देकर मात्रा ज्ञात करें।

प्रति पैंसिल मूल्य X	W	WX
1.0	50	50
1.5	20	30
2.0	10	20
	SW = 80	**SWX = 100**

$$\overline{X} = \frac{WX}{W} = \frac{100}{80}$$

अर्थात् प्रति पैंसिल औसत मूल्य = 1 रु. 80 पैसा।

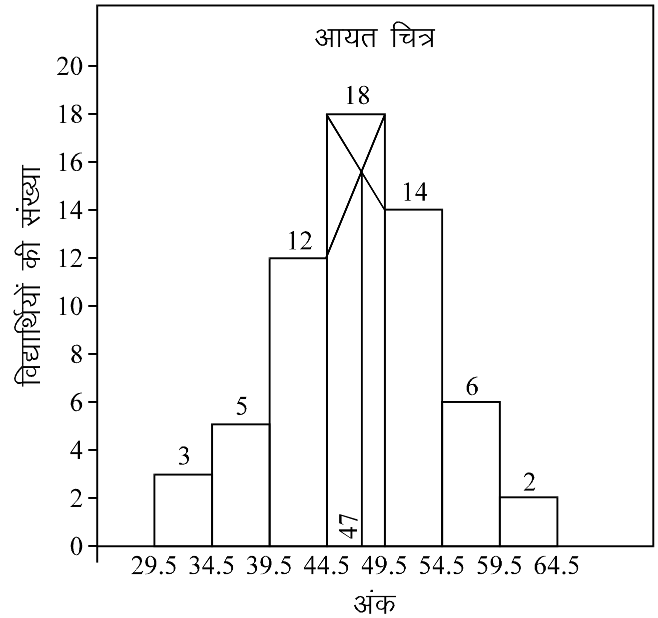
उदाहरण 81 : किसी महाविद्यालय में एक कक्षा में 60 विद्यार्थियों के प्राप्तों का आवत्ति बंटन नीचे दिया गया है :

अंक :	30-34	35-39	40-44	45-49	50-54	55-59	60-64
विद्यार्थियों की संख्या :	3	5	12	18	14	6	2

(*i*) इस बंटन का आयत चित्र बनाइयें और भूयिष्ठक ज्ञात कीजिये।

(*ii*) संचयी आवत्ति वक्र बनाइये और मध्य के 50% विद्यार्थियों की अंक सीमायें ज्ञात कीजिये।

हल : आयत चित्र बनाने से पहले वर्गों की वास्तविक सीमायें ज्ञात करें, जो इस प्रकार होंगी 2935-3435, 34.5-39.5 आदि।



एक उपयुक्त माध्य का चुनाव
(Choice of a Suitabe Average)

ऊपर पांच विभिन्न प्रकार के माध्यों का उल्लेख किया गया है। अब प्रश्न यह उठता है कि किस परिस्थिति में कौन-से माध्य का प्रयोग किया जाये। कोई भी माध्य प्रत्येक स्थिति में उपयुक्त नहीं होता। इसलिये ठीक माध्य का चुनाव न होने पर परिणाम भ्रामक होंगे। माध्य का चुनाव अधिकांशत: निम्न दो बातों पर निर्भर करता है :-

(1) माध्य का उद्देश्य, तथा

(2) दिये हुए समंकों की प्रकृति।

एक उपयुक्त माध्य का चुनाव नीचे दिये गये विभिन्न माध्यों के विशेष उपयोगों को ध्यान में रखकर करना चाहिये :-

समान्तर माध्य (Arithmetic Mean) : इसमें एक आदर्श माध्य के लगभग सभी गुण पाये जाते हैं इसलिये व्यापार में इस माध्य का सबसे अधिक प्रयोग होता है। केवल उन परिस्थितियों को छोड़कर जिनमें किसी अन्य माध्य के प्रयोग का विशेष लाभ हो, बाकि सभी स्थितियों में समान्तर माध्य का प्रयोग किया जाना चाहिये। आर्थिक सामाजिक समस्याओं के अध्ययन के लिये अधिकतम माध्य का ही प्रयोग किया जाता है। निर्यात, आयत, लागत, उत्पादन, उपभाग, आदि की केन्द्रीय प्रवत्ति का विश्लेषण समान्तर माध्य के प्रयोग द्वारा ही उचित रूप से किया जा सकता है। यदि समंकमाला में विभिन्न मदों का अलग-अलग सापेक्षिक महत्व हो तो भारित समान्तर माध्य उपयुक्त होता है।

मध्यका (Median) : खुले सिरे वाले आवत्ति -वितरण में माध्य निकालते समय मध्यका का विशेष स्थान है। गुणात्मक प्रकृति के समंकों का माध्य निकालने समय मध्यका का प्रयोग लाभप्रद है। उदाहरणार्थ, ईमानदारी, योग्यता, विद्यार्थियों के बौद्धिक स्तर आदि का अध्ययन करते समय मध्यका अधित उचित है।

भूयिष्ठक (Mode) : यह ज्ञात करने के लिये कि किसी श्रेणी में सबसे अधिक बार कौन-सा मूल्य पाया जाता है, भूयिष्ठक का प्रयोग उचित होता है। ॠतु-विज्ञान, जीव-शास्त्र, व्यापार व उद्योग के क्षेत्र मे विभिन्न समस्याओं का अध्ययन करने के लिये भूयिष्ठक बहुत उपयोगी सिद्ध हुआ है। उदाहरणार्थ, 'कॉलर का माध्य आकार', 'जूतों का माध्य आकार', 'तैयार कपड़ों का माध्य आकार', 'प्रति व्यक्ति या प्रति मशीन औसत उत्पादन', आदि ज्ञात करते समय भूयिष्ठक का प्रयोग अधित उचित है।

गुणोत्तर माध्य (Geometric Mean) गुणोत्तर माध्य उन परिस्थितियों में अधिक उचित हैं जहाँ श्रेणी के मूल्यों में अत्यधिक असमानता हो या तथ्यों में होने वाले सापेक्ष अपरिवर्तनों का अध्ययन करना हो। प्रतिशतों, अनुपातों व चक्रवद्धि दरों की औसत इसी माध्य द्वारा निकाली जाती है। सूचकांकों की रचना तथा जनसंख्या की वद्धि दर ज्ञात करने में यही अधिक संतोषजनक है।

हरात्मक माध्य (Harmonic Mean) : यह माध्य मात्रा-मूल्यों, गति तथा चलन वेग आदि समस्याओं के लिये सर्वश्रेष्ठ समझा जाता है।

माध्यों की सामान्य सीमायें
(General Limitation of Averages)

(1) किसी श्रेणी का माध्य ज्ञात करते समय हमें ऐसा मूल्य प्राप्त हो सकता है जो श्रेणी का ठीक प्रतिनिधित्व न करें। उदाहरणार्थ 100, 300, 250, 1500, 1000 का समान्तर माध्य $\frac{3150}{5}$ अर्थात् 630 है, यह मूल्य श्रेणी का ठीक प्रतिनिधित्व नहीं करता।

(1) कुछ परिस्थितियों में माध्य निकालते समय एक ऐसा मूल्य प्राप्त होता है जो बड़ा अजीब-सा होता है। उदाहरणार्थ, कुछ परिवारों के औसत बच्चों की संख्या निकालते समय हमें मान 5.67 प्राप्त हों सकता है लेकिन बच्चे अंशों में नहीं हो सकते ध्यान रहे कि माध्य का उपसादन नहीं करना चाहिये अत्यथा परिणामों में बहुत अशुद्धता आ सकती है। यदि हम 5.67 के स्थान पर बच्चों की संख्या 6 कर दें और 100 परिवारों के आधार पर हमने माध्य ज्ञात किया तो कुल बच्चों की संख्या 600 हो जायेगी जबकि वास्तविक संख्या 567 है।

(1) माध्यों से समंकमाला की बनावट, व्यक्तिगत मानों का माध्य-मान से औसत अन्तर तथा उनकी विषमता का आभास नहीं हो पाता। उदाहरणार्थ, निम्न दो समंक मालाओं को देखिये :-

श्रेणी (क)	**श्रेणी (ख)**
150	300
170	500
190	20
210	78
180	2
योग 900	**900**
$\overline{X}$ **= 180**	$\overline{X}$ **= 180**

दो श्रेणियों का समान्तर माध्य एक ही है लेकिन इससे निष्कर्ष निकालना गलत होगा कि दोनों श्रेणियां एक-सी हैं क्योंकि उनकी बनावट में बहुत अन्तर है।

(4) क्योंकि माध्य समंक श्रेणी का प्रतिनिधित्व करता है इसलिये इसका ठीक निर्वचन अत्यन्त आवश्यक है अन्यथा गलत निष्कर्षो पर पहुंचा जा सकता है। एक कहानी द्वारा यह बात स्पष्ट हो जायेगी। एक व्यक्ति नदी पार करना चाहता था। उसे बताया गया कि नदी की औसत गहराई 5'-8" है। एक वह व्यक्ति 5'-11" लम्बा था इसलिये उसने सोचा कि वह सुगमता से नदी पार कर सकता है क्योंकि वह नदी के पानी से ऊपर ही रहेगा। लेकिन जब वह नदी पार करने लगा तो डूब गया क्योंकि शुरू में पानी कम था और बीच में बहुत गहरा जबकि शुरू से आखिर तक पानी की औसत गहराई 5'-8" ही थी।

❑❑

अध्याय - 7

अपकिरण के माप (Measures of Dispersion)

पिछले अध्याय में केन्द्रीय प्रवत्ति के मापन की विभिन्न विधियों का वर्णन किया गया है। इन विधियों द्वारा एक ऐसा मान ज्ञात किया जाता है जो समंकमाला का प्रतिनिधित्व करता है। इस मान का बहुत महत्व है लेकिन यह समंकमाला की सब विशेषताओं को स्पष्ट करने में असमर्थ है। उदाहरणार्थ, माध्य द्वारा यह पता नहीं लग सकता है कि श्रेणी के विभिन्न व्यक्तिगत मानों का इससे औसत अंतर क्या है और श्रेणी की रचना तथा स्वरूप क्या है। निम्न उदाहरण से यह तथ्य स्पष्ट हो जायेगाः

	श्रेणी 'क'	श्रेणी 'ख'	श्रेणी 'ग'
	100	100	480
	100	105	10
	100	95	5
	100	105	3
	100	95	2
योग	**500**	**500**	**500**
मध्यक	**100**	**100**	**100**

तीन श्रेणियों में मध्यक समान है अर्थात 100 है। सांख्यिकी के ज्ञान से अनभिज्ञ व्यक्ति इससे यह निष्कर्ष निकाल सकता है कि तीनों श्रेणियों में कोई अंतर नहीं है लेकिन यह निष्कर्ष निकालना उचित न होगा। क्योंकि इन तीनों श्रेणियों में बहुत असमानता है। श्रेणी 'क' के मध्यक प्रत्येक मद का पूर्णरूप से प्रतिनिधित्व करता है क्योंकि मदों के मानों तथा मध्यक में कोई अंतर नहीं है। श्रेणी 'ख' में केवल एक मद ऐसा है जिसका मान मध्यक के बराबर है — अन्य मान्य मध्यक से भिन्न हैं लेकिन मध्यक मान और मदों के इन मानों में बहुत कम अंतर है। श्रेणी 'ग' में मदों में आपस में बहुत भिन्नता है और मध्यक किसी भी मान का प्रतिनिधित्व नहीं करता। इस उदाहरण से यह स्पष्ट है कि केवल माध्य को प्राप्त करके हम ठीक परिणाम पर नहीं पहुँच सकते। अत्यन्त रोचक घटना इस बात को स्पष्ट कर देगी। एक व्यक्ति ने सपरिवार नदी पार करने से पहले यह अनुमान लगाया कि नदी की औसत गहराई से उसके कुटुम्ब के सदस्यों की औसत ऊँचाई अधिक है। परन्तु एक स्थान पर नदी की गहराई बहुत अधिक थी इसलिए उसका पूरा कुटुम्ब डूब गया। यदि उस नदी की अधिकतम गहराई और उसकी औसत गहराई में अंतर तथा अपने परिवार के सदस्यों की न्यूनतम ऊँचाई, आदि की गणना कर लेता तो इस दुर्घटना से बच जाता। इससे स्पष्ट है कि संमक श्रेणी के बारे में यथेष्ट ज्ञान प्राप्त करने के लिए, न केवल उसका माध्य जानना आवश्यक है बल्कि विभिन्न व्यक्तिगत मानों का उस माध्य से औसत अंतर और श्रेणी की रचना तथा स्वरूप आदि के बारे में पूरी जानकारी प्राप्त करना भी परमावश्यक है। अर्थात् यह जानना आवश्यक है कि श्रेणी का प्रत्येक मद माध्य से कितनी दूरी पर या कितना बड़ा या छोटा होता है। विचलन की दूरी, फैलाव, बिखराव या विस्तार को अपकिरण कहते हैं। अपकिरण की कुछ विशेषताऐं इस प्रकार हैं :

''अपकिरण मदों के विचरण का माप है।'' - **बाउले**

''अपकिरण व्यक्तिगत मदों से भिन्नता की मात्रा का माप है।'' - **कोनर**

''किसी समंकमाला के माध्य से समंकों के बिखराव के माप को अपकिरण कहते हैं।'' - **कॉफका**

अपकिरण माप द्वितीय श्रेणी का माप (average of the second order) कहा जाता है। इसका कारण यह है कि अपकिरण का माप ज्ञान करते समय पहले समंकमाला का माध्य निकाला जाता है फिर उस माध्य से विभिन्न मूल्यों के विचलनों में अंतरों का माध्य ज्ञात किया जाता है अर्थात् अपकिरण, माध्य से निकाले गये विचलन का माध्य है।

निरपेक्ष तथा सापेक्ष अपकिरण
(Absolute and Relative Dispersion)

अपकिरण निरपेक्ष (Absolute) माप में किसी समंकमाला के प्रसार, बिखराव या विचरण को उस श्रेणी की इकाई में ही व्यक्त किया जाता है। उदाहरणार्थ, व्यक्तियों की आयु, भार आदि के अपकिरण के निरपेक्ष माप क्रमशः रुपये, किलोग्राम तथा वर्ष के रूप में प्रकट किये जायेंगे। निरपेक्ष माप द्वारा दो या दो से अधिक श्रेणियों का तुलनात्मक अध्ययन सम्भव नहीं क्योंकि विभिन्न श्रेणियों की इकाइयाँ भिन्न हो सकती हैं। तुलनात्मक अध्ययन के लिये निरपेक्ष मापों में बदलना आवश्यक है। अपकिरण निरपेक्ष माप को संबंधित माध्य से भाग देने पर जो अनुपात आता है वह अपकिरण का सापेक्ष माप कहलाता है। यह समंकमाला की इकाई में व्यक्त नहीं किया जाता वरन् एक अनुपात के रूप में होता है। **इसे अपकिरण गुणांक** (coefficient of dispersion) कहते हैं।

अपकिरण के उद्देश्य या महत्व
(Objects and Importance of Dispersion)

अपकिरण का माप निम्न उद्देश्यों की पूर्ति के लिये किया जाता है :

1. श्रेणी बनावट के बारे में सूचना प्राप्त करने तथा यह पता लगाने के लिए कि माध्य के दोनों और मूल्यों का फैलाव या बिखराव कैसा है।
2. श्रेणी के माध्य से विभिन्न पद-मानों की औसत दूरी ज्ञात करने के लिये।
3. दो या अधिक समंकमालाओं में पाई जाने वाली असमानताओं की तुलना करके यह निश्चित करने के लिये किसमें विचरण की मात्रा अधिक है।
4. यह देखने के लिये कि माध्य द्वारा श्रेणी का किस सीमा तक प्रतिनिधित्व होता है। जब अपकिरण कम होता है तो माध्य विभिन्न मदों का ठीक रूप से प्रतिनिधित्व करता है और अपकिरण के अधिक होने पर माध्य समंकों का ठीक प्रतिनिधित्व नहीं कर सकता।
5. अपकिरण के मापन का उद्देश्य विचरण (variation) की प्रकृति तथा कारणों की जानकारी प्राप्त करना भी है ताकि इसका नियंत्रण किया जा सके। उदाहरणार्थ, डॉक्टर के लिये बीमारी ठीक पता लगाने, तथा चिकित्सा करने के लिए शरीर के तापमान, नाडी-स्फुरण (pulse beat) तथा रक्त चाप (blood pressure) के विचरण वे समंक अत्यन्त सहायक होते हैं। औद्योगिक मापों का किस्म नियन्त्रण की विभिन्न सांखिकीय विधियों में प्रयोग होता है। आर्थिक व सामाजिक क्षेत्र में आय या सम्पत्ति के वितरण की असमानताओं का माप और तुलनात्मक अध्ययन के लिए अपकिरण के विभिन्न माप बहुत उपयोगी सिद्ध होते हैं। अर्थात् प्रत्येक क्षेत्र में मद-मानों के माध्यम से प्राप्त आंशिक एक अपूर्ण सूचना की पूर्ति अपकिरण माप द्वारा की जाती है।

एक आदर्श माप की विशेषतायें
(Properties of a Good Measure of Variations)

एक आदर्श विचरण-माप में निम्नलिखित विशेषतायें होनी चाहिये :

1. समझने में सरल,
2. निर्धारण में सरल,
3. श्रेणी के सभी मूल्यों पर आधारित,
4. निदर्शन के परिवर्तन का न्यूनतम प्रभाव,
5. स्पष्ट व स्थिर परिभाषा, तथा
6. बीजगणित विवेचन सम्भव।

अपकिरण-मापन की विभिन्न विधियाँ
(Different Methods of Measuring Dispersion)

अपकिरण-मापन की विभिन्न विधियाँ निम्नलिखित हैं :

1. विस्तार (range)
2. अर्द्ध-अन्तर चतुर्थक विस्तार (semi-interquartle range or quartile deviation)
3. माध्य-विचलन (mean deviation)
4. प्रमाप विचलन (standard deviation)
5. लॉरेन्ज वक्र (Lorenz curve)

I. विस्तार
(Range)

किसी समंकमाला के सबसे बड़े और सबे छोटे मान के अंतर को उसका विस्तार (range) कहते हैं। अपकिरण ज्ञान करने की सबसे सरल रीति है। विस्तार की गणन क्रिया इस प्रकार है :

1. श्रेणी का अधिकतम मान (largest value) और न्यूनतम मान (smallest value) ज्ञात किया जाता है। अविच्छित श्रेणी में न्यूनतम वर्ग की निचली सीमा को न्यूनतम मान और अधिकतम वर्ग की उच्च सीमा को अधिकतम मान माना जाता है।
2. निम्न सूत्र का प्रयोग किया जाता है :

   $R = L - S$

   R = विस्तार (range)

   L = श्रेणी का सबसे बड़ा मूल्य (largest value)

   S = श्रेणी का सबसे छोटा मूल्य (smallest value)

विस्तार गुणक (Coefficient of Range) : विस्तार के सापेक्ष माप (relative pressure) को विस्तार गुणक (coefficient of range) कहते हैं। तुलनात्मक अध्ययन के लिये विस्तार गुणक निकालना पड़ता है। इसकी गणना निम्न सूत्र द्वारा की जाती है।

$$\text{Coeff. of range} = \frac{L \quad S}{L \quad S}$$

उदाहरण 1 : एक कम्पनी के शेयरों के मूल्य (price of share) सोमवार से शनिवार तक इस प्रकार थे :

दिन	मूल्य (रू. में)	दिन	मूल्य (रू. में)
सोमवार	200	बहस्पतिवार	160
मंगलवार	210	शुक्रवार	220
बुधवार	208	शनिवार	250

इन समंकों से विस्तार व विस्तार-गुणक ज्ञात कीजिये।

हल : Range = L – S; L = 250; S = 160

\ Range = 250 – 160 अर्थात 90 रू.

$$\text{Coeff. of range} = \frac{L \quad S}{L \quad S} = \frac{250 \quad 160}{250 \quad 160} = \frac{90}{410} = 0.22$$

उदारहण 2 : निम्न समंकों से विस्तार गुणक ज्ञात कीजिये :

विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या	
20—29	8	50—59	7
30—39	12	60—69	3
40—49	20		

हल : सर्वप्रथम; समावेशी वर्गान्तरों को अपवर्जी वर्गान्तरों में बदलिये। उक्त उदाहरण में न्यूनतम सीमा 19.5 तथा अधिकतम सीमा 69.5 है।

$$\text{Coeff. of range} = \frac{L \quad S}{L \quad S} = \frac{69.5 \quad 19.5}{69.5 \quad 19.5} = 0.562$$

उदाहरण 3 : निम्न सूचना से न्यूनतम मूल्य ज्ञात कीजिये :

अधिकतम मूल्य = 120; विस्तार-गुणक = 0.5

हल :

$$\text{Coeff. of range} = \frac{L \quad S}{L \quad S}$$

$$0.5 = \frac{120 \quad S}{120 \quad S}$$

$$0.5\,(120 + S) = 120 - S$$

या

$$60 + 0.5\,S = 120 - S$$

$$1.5\,S = 120 - 60 = 60;$$

$$S = \frac{60}{1.5} = 40 \text{ अर्थात न्यूनतम मूल्य } = 40$$

विस्तार के गुण व दोष
(Merits and Demerits of Range)

गुण (Merits)

अपकिरण मापन की विभिन्न रीतियों में से विस्तार का समझना तथा निर्धारण सबसे सुगम है। गणन-क्रिया सरल होने के कारण विस्तार का मूल्य ज्ञात करने में बहुत कम समय लगता है।

दोष (Demerits)

1. विस्तार श्रेणी के सभी मद-मूल्यों पर आधारित नहीं होते। अधिकतम व न्यूनतक मूल्यों के बीच के मदों में होने वाले परिवर्तनों का विस्तार पर कोई प्रभाव नहीं पड़ता।
2. विस्तार मूल्य से समंकमाला की बनावट या चरम् मूल्यों के मध्य-मूल्यों के फैलाव या बिखराव का पता नहीं चलता। दो संमक श्रेणी का विस्तार समान होने पर भी उनकी बनावट में बहुत अंतर हो सकता है।

निम्न उदाहरण से यह बात स्पष्ट हो जायेगी :

श्रेणी 'क'	6,	46,	46,	46,	46,	46,	46,	46
श्रेणी 'ख'	6,	6,	6,	6,	46,	46,	46,	46
श्रेणी 'ग'	6,	10,	15,	25,	30,	32,	40,	46

तीनों श्रेणियों में विस्तार-मूल्य समान है अर्थात (46 – 6) या 40 है। लेकिन इसका अर्थ यह नहीं है कि तीनों श्रेणियाँ एक समान है।

3. विवर्तमुखी आवत्ति वितरणों (open-end frequency distributions) में विस्तार का मूल्य ज्ञात नहीं किया जा सकता।

विस्तार के प्रमुख प्रयोग
(Main Uses of Range)

निम्न परिस्थितियों में विस्तार अत्यन्त लाभप्रद है :

1. **सांख्यिकीय गुण नियन्त्रण** (Statistical Quality Control) : सांख्यिकीय गुण नियंत्रण की विभिन्न रीतियों में विस्तार का बहुत अधिक प्रयोग होता है।
2. **मौसम का पूर्वनूमान** (Weather Forecasting) : न्यूनतम तथा अधिकतम तापमान में अंतर निकालते समय विस्तार का प्रयोग किया जाता है।
3. वस्तुओं के स्टॉक और शेयर (Stocks and Shares) के मूल्यों में परिवर्तन का अध्ययन करने के लिए विस्तार का प्रयोग होता है। उदाहरणार्थ यदि 1990 में 10 ग्राम सोने की न्यूनतम कीमत 3000 रू. तथा अधिकतम कीमत 3259 रू. थी तो इससे विस्तार का पता लगता है अर्थात (3250 – 3000) = 250 रू.। समाचार-पत्रों में स्टॉक व शेयर की न्यूनतम व अधिकतम कीमतें दी जाती हैं। जिससे बहुत लाभप्रद सूचना मिलती है।

II. चतुर्थक विचलन
(Quartile Deviation)

किसी भी श्रेणी के ततीय व प्रथम चतुर्थकों के आधे को चतुर्थक -विचलन (Quartile Deviation), अर्द्ध या अंतर-चतुर्थक विस्तार (Semi-inter Quartile Range) कहते हैं। चतुर्थक-विचलन ज्ञात करने के लिये निम्न सूत्र का प्रयोग किया जाता है :

$$\frac{\frac{Q_3 \quad Q_1}{2}}{\frac{Q_3 \quad Q_1}{2}} \qquad \frac{Q_3 \quad Q_1}{Q_3 \quad Q_1}$$

$$\text{Q.D.} = \frac{Q_3 \quad Q_1}{2}$$

Q.D. = चतुर्थक विचलन (Quartile deviation)

Q_1 = प्रथम चतुर्थक (First quartile)

Q_3 = ततीय चतुर्थ (Third quartile)

चतुर्थक-विचलन गुणक
(Coefficient of Quartile Deviation)

चतुर्थक-विचलन अपकिरण का निरपेक्ष माप है। इसका सापेक्ष माप चतुर्थक-विचलन गुणक कहलाता है। इसे ज्ञात करने के लिये चतुर्थक-विचलन के निरपेक्ष माप को दोनों चतुर्थकों के माध्य से भाग दे दिया जाता है। सूत्र के रूप में :

Coeff. of Q.D. =

चतुर्थक-विचलन तथा चतुर्थक गुणक का मूल्य ज्ञात करने के लिये प्रथम चतुर्थक तथा ततीय चतुर्थक का मूल्य ज्ञात करना पड़ता है, फिर उपयुक्त सूत्रों का प्रयोग होता है।

उदाहरण 4 : निम्न समंकों से चतुर्थक-विचलन गुणक ज्ञात कीजिये :

9 कर्मचारियों का वेतन (रू. में)

170, 82, 110, 100, 150, 120, 200, 116, 250

हल : Coeff. of range $= \dfrac{Q_3 \quad Q_1}{Q_3 \quad Q_1}$

अर्थात Q_1 और Q_3 का मूल्य निकालने से चतुर्थक-विचलन गुणक की गणना की जा सकती है। Q_1 और Q_3 का मूल्य ज्ञात करने के लिये सर्वप्रथम समंकों को आरोही

वेतन आरोही क्रम में अनुविन्यिसित	
82	150
100	170
110	200
116	250
120	

$$Q_1 = \text{Size of } \frac{N \quad 1}{4} \text{th item} =$$
$$= 2.5\text{th item}$$

$$\text{Size of 2.5th item} = \quad = \frac{100 \quad 110}{2}$$
$$= 105$$

$$Q_3 = \text{Size of } 3 \quad \text{th item} =$$
$$= 7.5\text{th item}$$

$$\text{Size of 7.5 item} = \quad = \frac{170 \quad 200}{2}$$
$$= 185$$

$$Q_3 = 185; Q_1 = 105$$

$$\text{Coeff. of Q.D.} = \quad = \frac{80}{290} = 0.276$$

उदाहरण : निम्न समंकों से चतुर्थक-विचलन (Quarti Deviation) तथा चतुर्थक-विचलन गुणक (Coefficient of Quarti Deviation) ज्ञात कीजिये :

ऊँचाई (इंचों में)	विद्यार्थियों की संख्या	ऊँचाई (इंचों में)	विद्यार्थियों की संख्या
58	15	63	22
59	20	64	20
60	32	65	10
61	35	66	8
62	33		

हल : चतुर्थक-विचलन का परिगणन

ऊँचाई (इंचों में)	आवत्ति f	संचयी आवत्ति $c.f.$	ऊँचाई (इंचों में)	आवत्ति f	संचयी आवत्ति $c.f.$
58	15	15	63	22	157
59	20	35	64	20	177
60	32	67	65	10	187
61	35	102	66	8	195
62	33	135			

$$\text{Q.D.} = \frac{Q_3 \quad Q_1}{2}$$

$$Q_1 = \text{Size of} = \quad \text{th item} \quad = 49\text{th item}$$

Size of 49th item = 60; अर्थात् $Q_1 = 60$

$$Q_3 = \text{Size of } 3 \quad \text{th item} \quad = 147\text{th item}$$

Size of 147th item = 63, अर्थात $Q_3 = 63$

\ $\text{Q.D.} = \quad = 1.5$

$$\text{Coeff. of Q.D.} = 888 = \frac{Q_3 \quad Q_1}{Q_3 \quad Q_1} = \frac{63 \quad 60}{63 \quad 60} = \frac{3}{123} = 0.024$$

उदाहरण 6 : निम्न समंकों से चतुर्थक-विचलन (Quarti Deviation) तथा चतुर्थक विचलन गुणांक (Coefficient of Mean Deviation) ज्ञात कीजिये :

चर-मूल्य	आवत्ति	चर-मूल्य	आवत्ति
4 — 8	6	24 — 28	12
8 — 12	10	28 — 32	10
12 — 16	18	32 — 36	6
20 — 24	30	36 — 40	2
20 — 24	15		

हल : चतुर्थक विचलन तथा चतुर्थक-विचलन गुणांक का निर्धारण

चर-मूल्य	आवत्ति	संचयी आवत्ति	चर-मूल्य	आवत्ति	संययी आवत्ति
4 — 8	6	6	24 — 28	12	91
8 — 12	10	16	28 — 32	10	101
12 — 16	18	34	32 — 36	6	107
20 — 24	30	64	36 — 40	2	109
20 — 24	15	79			

$$\text{Q.D.} = \frac{Q_3 \quad Q_1}{2}$$

$$Q_1 = \text{Size of} \quad \text{th item}$$

$$= \frac{109}{4}$$

$$= 27.25\text{th item}$$

अर्थात् Q_1 12 – 16 वर्ग में स्थित है :

$$Q_1 = L + \frac{\frac{N}{4} \quad c.f.}{f} \times i$$

$L = 12$; $\frac{N}{4} = 27.25$; $c.f. = 16$; $f = 18$; $i = 4$

\

$$Q_1 = 12 + \frac{27.25 \quad 16}{18} \times 4$$

$$= 12 + 2.5$$

$$= 14.5$$

$$Q_3 = \text{Size of} \quad \text{th item} - \frac{3 \quad 109}{4}$$

$$= 81.75\text{th item}$$

Q_3 = 24 – 28 वर्ग में स्थित है।

$$Q_3 = L + \qquad \times i$$

$L = 24$; $\frac{3N}{4} = 81.75$; $c.f. = 79$; $f = 12$; $i = 4$

$$Q_3 = 24 + \frac{81.75 \quad 79}{12} \times 4$$

$$= 24 + \qquad = 24.917$$

\

$$\text{Q.D.} = \frac{24.917 \quad 14.5}{2}$$

$$=$$

$$= 5.208$$

$$\text{Coeff. of Q.D.} = \qquad = \frac{24.917 \quad 14.5}{24.917 \quad 14.5} = \frac{10.417}{39.415}$$

$$= 0.264$$

उदारहण 7 : निम्नलिखित आँकड़ों में चतुर्थक विचलन के गुणाँक की गणना कीजिये :

अंक	छात्र संख्या
20 से कम	**8**
40 से कम	**20**
60 से कम	**50**
80 से कम	**60**
100 से कम	**80**

हल : चतुर्थक विचलन तथा विचरण गुणांक का परिगमन

अंक	*m.p.* *m*	छात्र संख्या *f*	(*m* – 50)/20 *d*	*fd*	fd^2	*c.f.*
0 — 20	10	8	– 2	– 16	32	8
20 — 40	30	12	– 1	– 12	12	20
40 — 60	**50**	30	0	0	0	0
60 — 80	70	20	+1	+20	20	70
80 — 100	90	10	+2	+20	40	80
	N = 80			**S*fd* = 12**	**Sfd^2 = 104**	

विचरण गुणांक (Coeff. of Variation) $= \frac{\ }{X} \times 100$

$$\overline{X}X = A + \frac{fd}{N} \times i$$

$$= 50 + \frac{12}{80} \times 20$$

$$= 50 + 3 = 53$$

$$s = \sqrt{\frac{fd^2}{N} \quad \left(\frac{fd}{N}\right)^2} \times i$$

$$= \sqrt{\frac{104}{80} \quad \left(\frac{12}{80}\right)^2} \times 20$$

$$= \sqrt{1.3 \quad 0.0225} \times 20$$

$$= 1.13 \times 20 = 22.6$$

चतुर्थक विचलन गुणांक (Coeff. of Q.D.) $= \frac{Q_3 \quad Q_1}{Q_3 \quad Q_1}$

$$Q_1 = \text{Size of } \frac{N}{4}\text{nth item}$$

$$= \frac{80}{4} = 20\text{th item}$$

Q_1 20 – 40 वर्ग में स्थित है।

$$Q_1 = L + \frac{\frac{N}{4} \quad c.f.}{f} \times i$$

$L = 20;\ \frac{N}{4} = 20;\ c.f. = 8, f = 12;\ i = 20$

$$Q_1 = 20 + \frac{20 \quad 8}{12} \times 20$$

$$= 20 \ + 20$$

$$= 40$$

$$Q_3 = \text{Size of} \qquad \text{th item}$$

$$= \frac{3 \quad 80}{4} = 60\text{th item}$$

Q_3 60 – 80 वर्ग में स्थित है।

$$Q_3 = L + \frac{\frac{3N}{4} \quad c.f.}{f} \times i \qquad \frac{Q_3 \quad Q_1}{Q_3 \quad Q_1}$$

$L = 60;\ \frac{3N}{4} = 60;\ c.f. = 50; f = 20;\ i = 20$

$$Q_3 = 60 + \frac{60 \quad 50}{20} \times 20$$

$$= 70$$

$$\text{Coeff. of Q.D.} =$$

$$= \frac{70 \quad 40}{70 \quad 40}$$

$$= \frac{30}{110} = 0.273$$

चतुर्थक-विचलन के गुण व दोष
(Merits and Demerits of Quartile Deviation)

गुण (Merits)

1. चतुर्थक-विचलन का समझना व निर्धारण करना सरल है।
2. अपकिरण के इस माप पर चरम मूल्यों का बहुत कम प्रभाव पड़ता है।

3. जहाँ श्रेणी के मध्य भाग का ही अध्ययन करना हो वहां इस माप का प्रयोग होता है।

दोष (Demerits)

1. चतुर्थक-विचलन श्रेणी के सभी मूल्यों पर आधारित नहीं है।
2. इसका बीचगणितीय विवेचना सम्भव नहीं है।
3. निदर्शन-परिवर्तन का इस पर बहुत प्रभाव पड़ता है।
4. इससे समंकमाला की बनावट का ठीक पता नहीं चलता है। यदि दो श्रेणियों में चतुर्थक समान है तो चतुर्थक-विचलन भी बराबर होगी जबकि दोनों में मूल्यों का बिखराव या श्रेणी की रचना भिन्न हो सकती है।

   इन दोषों के कारण चतुर्थक-विचलन का सांख्यिकी में बहुत कम प्रयोग होता है।

III. माध्य विचलन या औसत विचलन (Mean Deviation on Average Deviation)

समंक श्रेणी के किसी सांख्यिकीय माध्य (मध्यक, मध्यका तथा भूयिष्ठक) से निकाले गये विभिन्न मूल्यों का समान्तर माध्य, उसका माध्य-विचलन (mean deviation) कहलाता है। यह एक ऐसा मूल्य है जो किसी श्रेणी के माध्य से विभिन्न मूल्यों के औसत विचलन को प्रकट करता है। माध्य विचलन जितना अधिक होता है, पदमाला में अपकिरण या फैलाव उतना ही अधिक होता है।

माध्य-विचलन ज्ञात करने में निम्न तथ्यों को ध्यान में रखा जाता है :

1. **माध्य का चुनाव** (Selection of Appropriate Average) : माध्य-विचलन भूयिष्ठक, मध्यका एवं मध्यक किसी से भी निकाला जा सकता है परन्तु मध्यका को प्रधानता दी जानी चाहिए क्योंकि इससे निकाले गये विचलनों का प्रयोग सबसे कम होता है। व्यवहार में माध्य विचलन निकालने के लिये मध्यक का भी प्रयोग होता है। भूयिष्ठक का प्रयोग यथासम्भव नहीं करना चाहिये क्योंकि यह बहुत अनिश्चित होता है। यदि प्रश्न में स्पष्ट रूप से यह निर्देश नहीं दिया है किस माध्य से माध्य-विचलन निकाला है तो मध्यका से ही विचलन निकालने चाहियें।
2. **बीजगणित चिन्हों की उपेक्षा** (Ignoring Algebric Signs) : माध्य के चुनाव के पश्चात प्रत्येक मूल्य का माध्य से विचलन निकाल लेते हैं। माध्य-विचलन निकालते सम + तथा – चिन्हों को छोड़ दिया जाता है अर्थात् ऋणात्मक विचलनों (negative deviation) को भी धनात्मक (positive) मान लेते हैं। उदाहरणार्थ, यदि 20 में से 25 घटायेंगे तो – 5 के स्थान पर 5 लिखेंगे। जब विचलन + तथा – को छोड़कर लिखे जाते हैं तो उसके योग को सीधी रेखाओं के बीच इस प्रकार लिखते हैं S | D | इसका अर्थ है कि विचलनों का योग करते सम + या – का ध्यान नहीं रखा गया है।
3. **विचलनों का माध्य** (Averge of Deviations) : विचलनों के योग की मदों की संख्या से भाग देकर जो मूल्य आता हैं उसे माध्य-विचलन कहते हैं। विच्छिन्न एवं अविच्छिन्न आवत्ति श्रेणियों से विचलनों में आवतियों को गुणा देकर योग निकाला जाता है और आवत्ति के योग से भाग दिया जाता है।
4. **माध्य-विचलन गुणांक** (Coefficeint of Mean Deviation) : तुलनात्मक अध्ययन के लिए माध्य-विचलन का सापेक्ष माप निकाला जाता है जिसे माध्य-विचलन का गुणक कहते हैं। इसके लिए माध्य विचलन के निरपेक्ष माप से उस माध्य के विचलनों के कुल योग को भाग दिया जाता है जिससे विचलन निकाले गया हैं अर्थात्

   (क) Coeff. of mean deviation from median or median coefficient of dispersion = $\frac{\text{Mean deviation}}{\text{Median}}$

   (ख) Coeff. of mean deviation from mean or mean coefficient of dispersion = $\frac{\text{Mean deviation}}{\text{Median}}$

व्यक्तिगत श्रेणी में माध्य विचलन का निर्धारण

(Computing Mean Deviation in Individual Series)

निम्न सूत्र द्वारा व्यक्तिगम श्रेणी में माध्य-विचलन का निर्धारण किया जाता है :

$$\text{M.D.} = \frac{|D|}{N}$$

| D | = (माध्य विचलन) मध्यक का मध्यका से विचलन (+ या – चिन्ह को छोड़ते हुए)। | D | = Deviation from mean or median ignoring signs)।

विधि :

1. उस माध्य का परिगणन किया जाता है जिससे माध्य विचलन निकालना है - अधिकतर मध्यका का प्रयोग किया जाता है।
2. + और – चिन्हों को छोड़ते हुए मध्यका या अन्य माध्य से विभिन्न मूल्यों के विचलन निकाले जाते हैं।
3. इन विचलनों का जोड़ (S | D |) प्राप्त किया जाता है।
4. विचलनों के योग को मदों की संख्या से भाग देते है। प्राप्त भागफल माध्य विचलन होता है।

   माध्य विचलन गुणक निकालने के लिए माध्य विचलन को संबंधित माध्य से भाग दे दिया जाता है।

उदाहरण 8 : 10 दुकानों पर किसी एक प्रकार के रेडियो की कीमत इस प्रकार हैं :

कीमत रू. में।

210, 220, 225, 225, 225, 235, 240, 250, 270, 280

माध्य-विचलन तथा माध्य-विचलन गुणक ज्ञात कीजिये।

हल : **माध्य-विचलन का परिगणन**

X	**230 से विचलन | D |**
210	20
220	10
225	5
225	5
225	5
235	5
240	10
250	20
270	40
280	50
N = 10	**Σ | D | = 170**

$$\text{M.D.} = \frac{|D|}{N}$$

$$M = \text{size of } \frac{N \quad 1}{2}\text{th item}$$

$$= \qquad = 5.5\text{th item}$$

$$= \qquad = \frac{225 \quad 235}{2} = 230$$

$$\text{M.D.} = \qquad = 17$$

$$\text{Coeff. of M.D.} = \frac{\text{M.D.}}{\text{Median}} = \frac{17}{230} = 0.074$$

विच्छिन्न श्रेणी में मध्य विचलन का परिगणन

(Calculating Mean Deviation in Discrete Series)

विच्छिन्न श्रेणी में निम्न सूत्र द्वारा माध्य-विचलन ज्ञात किया जाता है।

$$\text{माध्य विचलन} = \frac{f\ |D|}{N}$$

विधि :

1. जिस माध्य से माध्य-विचलन निकालना है उसका निर्धारण किया जाता है अर्थात् मध्यक, मध्यका, भूयिष्ठक आदि।
2. प्रत्येक चर-मूल्य का उस माध्य से विचलन लिया जाता है अर्थात् | D | ज्ञात किया जाता है।
3. विचलनों को आवत्तियों से गुणा करके जोड़ निकाला जाता है, अर्थात S*f* | D | ज्ञात करते हैं।
4. S*f* | D | को आवत्ति के योग से भाग देते हैं। प्राप्त भागफल माध्य -विचलन कहलाता है।

$$\frac{5\text{th item} + 6\text{th item}}{2} \qquad \frac{100}{10}$$

माध्य-विचलन गुणक निकालने के लिए, निरपेक्ष माप को उस माध्य से भाग दे दिया जाता है जिससे विचलन ज्ञात किये गये हैं।

उदाहरण 9 : निम्न समंकों से माध्य विचलन तथा माध्य-विचलन गुणांक का परिगणना कीजिए :

X	**:**	**2.5**	**3.5**	**4.5**	**5.5**	**6.5**	**7.5**	**8.5**	**9.5**	**10.5**
f	**:**	**2**	**3**	**5**	**6**	**6**	**3**	**6**	**4**	**14**

हल : माध्य-विचलन का मध्यका से परिगणन

X	*f*	*c.f.*	| X – 7.5 | | D |	*f* | D |
2.5	2	2	5	10
3.5	3	5	4	12
4.5	5	10	3	15
5.5	6	16	2	12
6.5	6	22	1	6
7.5	4	26	0	0
8.5	6	32	1	6
9.5	4	36	2	8
10.5	14	50	3	42
	N = 50			*****f* | D | = 111**

Med = Size of $\frac{N \quad 1}{2}$th item = = 25.5th item

अर्थात मध्यका मूल्य 7.5 है।

माध्य-विचलन (Mean Deviation) = $= \frac{111}{50} = 2.22$

माध्य-विचलन गुणांक (Coefficeint of M.D.) = $\frac{M.D.}{Median} = \frac{2.22}{7.5} = 0.296$

उदाहरण 9 : निम्न आवत्ति -वितरण से समान्तर माध्य तथा (समान्तर माध्य सापेक्ष) माध्य-विचलन ज्ञात कीजिए :

X :	**2**	**4**	**6**	**8**	**10**
आवत्ति :	**1**	**4**	**6**	**4**	**1**

हल : **सामान्तर माध्य तथा माध्य-विचलन का निर्धारण**

X	***f***	***f*.X**	**(X – 6) | D |**	***f* | D |**
2	1	2	4	4
4	4	16	2	8
6	6	36	0	0
8	4	32	2	8
10	1	10	4	4
	N = 16	**$\Sigma fX = 96$**		**Σf | D | = 24**

$$\overline{X} = \frac{f\ X}{N} = \frac{96}{16} = 6$$

$$M.D. = \frac{f\ |D|}{N} = \frac{24}{16} = 1.5$$

अविच्छिन्न श्रेणी में माध्य-विचलन का परिगणन
(Calculating Mean Deviation in Continuous Series)

अविच्छिन्न श्रेणी में दो रीतियों द्वारा माध्य-विचलन ज्ञात किया जा सकता है :

1. प्रत्यक्ष (Direct method), तथा
2. लघु रीति (Short-cut method)।

प्रत्यक्ष रीति (Direct Method) **:** प्रत्यक्ष रीति द्वारा माध्य विचलन निकालते समय प्रत्येक वर्ग के मध्य-बिन्दु ज्ञात किये जाते हैं। मध्य-बिन्दु निकालते ही अखण्डित श्रेणी खण्डित श्रेणी में बदल जाती है और शेष सभी क्रियायें पूर्ववत् रहती हैं माध्य-विचलन के निर्धारण के लिये निम्न सूत्र का प्रयोग करते हैं :

$$M.D. = \frac{f\ |D|}{N}$$

| D | = माध्य-बिन्दुओं का मध्यक से विचलन (+ या – चिन्ह छोड़कर)

(Deviation of midpoints from median, ignoring signs)

निम्न उदाहरण द्वारा या रीति स्पष्ट हो जायेगी :

यदि हम कोई उभयनिष्ठ गुणांक (common factor) लेते हैं तो सूत्र इस प्रकार होगा :

$$\text{M.D.} = \frac{f\,|D|}{N} \times i$$

उदाहरण 10 : निम्नलिखित समंकों का मध्यका की सहायता से माध्य-विचलन गुणांक ज्ञात कीजिये :

अंक	विद्यार्थियों की संख्या	अंक	विद्यार्थियों की संख्या
10 — 20	2	50 — 60	25
20 — 30	6	60 — 70	20
30 — 40	12	70 — 80	10
40 — 50	18	80 — 90	7

हल : माध्य-विचलन गुणांक का परिगणन

अंक	*m*	*f*	*c.f.*	| *m* – 54.8| / 10	
				| D |	*f* | D |
10 — 20	15	2	2	3.98	7.96
20 — 30	25	6	8	2.98	17.88
30 — 40	35	12	20	1.98	23.76
40 — 50	45	18	$\frac{f\,\|D\|}{N}$ 38	0.98	17.64
50 — 60	55	25	63	0.02	0.50
60 — 70	65	20	83	1.02	20.40
70 — 80	75	10	93	2.02	20.20
80 — 90	85	7	100	3.02	21.14
	N = 100				**S*f* | D | = 129.48**

Med. = Size of $\frac{N}{2}$th item = $\frac{100}{2}$ = 50th item

मध्यका 50 — 60 वर्ग में स्थित है।

$$\text{Med.} = L = \frac{\frac{N}{2} \quad c.f.}{f} \times i$$

L = 50, $\frac{N}{2}$ = 50, *c.f.* = 38, *f* = 25, *i* = 10

$$\text{Med.} = 50 + \frac{50 \quad 38}{25} \times 10$$

$$= 50 + 4.8 = 54.8$$

$$\text{M.D.} = \qquad \times i$$

$$= \frac{129.48}{100} \times 10 = 12.948$$

Coeff. of M.D. $= \dfrac{\text{M.D.}}{\text{Median}}$

$$= \frac{12.948}{54.8} = 0.236$$

लघु रीति (short-cut Method) : माध्य विचलन निकालते समय जब मध्य या मध्यका मूल्य अंशों में आये तो गणन-क्रिया को सरल बनाने के लिए लघु रीति का प्रयोग उचित रहता है। सूत्र इस प्रकार है :

$$\text{M.D. (from Median)} = \frac{f_{ma} \quad f_{mb} \quad (\ f_a \quad f_a) \text{ Med.}}{\text{N}}$$

$$\text{M.D. (from Mean)} = \frac{f_{ma} \quad f_{mb} \quad (\ f_a \quad f_a) \ \overline{\text{X}}}{\text{N}}$$

Sf_{ma} तथा Sf_{ma} = माध्य-मूल्य से अधिक एवंक कम मूल्य वाले मध्य बिन्दुओं और तत्संबंधी आवतियों के गुणनफल के योग (Total of products of midpoint and frequencies corresponding to mid-points above and below the average value respectively).

Sf_a and Sf_b = माध्य-मूल्य से अधिक व कम मध्य-बिन्दुओं से संबंधित आवतियों के योग (Total of frequencies pertaining to mid-points above and below the average value).

विधि :

1. वह माध्य (माध्यक/मध्यका) निकालते हैं जिससे माध्य-विचलन निकालना होता है।
2. मध्य-बिन्दु (m) और संबंधित आवत्ति (f) को गुणा करते हैं।
3. अपेक्षित माध्य से अधिक मूल्यों व उनकी आवत्तियों के गुणनफलों का योग प्राप्त करते हैं अर्थात Sf_{ma} ज्ञात करते हैं।
4. अपेक्षित माध्य से कम मूल्यों तथा उनकी आवत्तियों के गुणनफलों का योग प्राप्त करते हैं अर्थात Sf_{mb} ज्ञात करते हैं।
5. अपेक्षित माध्य के बराबर के मूल्य को छोड़ देते हैं।
6. माध्य-मूल्य से अधिक आवत्तियों का योग निकाल लेते हैं अर्थात Sf_a ज्ञात करते हैं।
7. माध्य-मूल्य से कम आवत्तियों का योग निकाल लेते हैं अर्थात Sf_b ज्ञात होते हैं।

उदाहरण 12 : उदाहरण 10 के समंकों से लघु रीति का प्रयोग करके माध्य-विचरण ज्ञात कीजिये :

हल :

$$\text{M.D.} = \frac{f_{ma} \quad f_{mb} \quad (\ f_a \quad f_a) \text{ Median}}{\text{N}}$$

$Sf_{ma} = 619$; $Sf_{mb} = 274$; $Sf_a = 62$; $Sf_b = 38$; Med. = 9; N = 100

$$\text{M.D.} = \frac{619 \quad 274 \quad (62 \quad 38) \quad 9}{100}$$

$$=$$

$$= \quad = 1.29$$

इससे स्पष्ट है कि दोनों विधियों द्वारा उत्तर एक ही जायेगा।

उदाहरण 12 : निम्नलिखित समंकों से मध्यका द्वारा विचलन ज्ञात कीजिए :

उम्र वर्षों में	व्यक्तियों की संख्या	उम्र वर्षों में	व्यक्तियों की संख्या
10 — 20	15	50 — 60	15
20 —30	16	60 — 70	11
30 — 40	16	70 — 80	9
40 — 50	20		

हल : मध्यका से माध्य विचलन का परिगणन

उम्र	मध्य बिन्दु (m)	f	$c.f.$	$\lvert m - 42.5 \rvert$ $\lvert D \rvert$	$f \lvert D \rvert$
10 — 20	15	15	15	27.5	412.5
20 — 30	25	14	29	17.5	245.0
30 — 40	35	16	45	7.5	120.0
40 — 50	45	20	95	2.5	50.0
50 — 60	55	15	80	12.5	187.5
60 — 70	65	11	91	22.5	247.5
70 — 80	75	9	100	32.5	292.5
		N = 100	$\frac{f \lvert D \rvert}{N}$		**Sf | D | = 1555**

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th item}$$
$$= \text{50th item}$$

मध्यका = 40 – 50 वर्ग में स्थित है।

$$M = L + \frac{\frac{N}{2} \quad c.f.}{f}$$

$L = 40,\ \frac{N}{2} = 50,\ c.f. = 45,\ f = 20,\ i = 10$

$$\text{Med.} = 40 + \frac{50 \quad 45}{20} \times 10$$
$$= 40 + 2.5$$
$$= 42.5$$

$$\text{M.D.} =$$

$$= \frac{1555}{100}$$
$$= 1.555$$

उदाहरण 14 : उदाहरण 10 से लघु रीति द्वारा मध्य विचलन गुणांक ज्ञात कीजिए :

माध्य विचलन गुणांक का लघुरीति द्वारा परिगणन

m	*f*	*fm*		
15	2	39		(b)
25	6	150	$Sf_b = 38$	मध्यका
35	12	420	$Sf_{mb} = 1410$	वर्ग से
45	18	810		नीचे
55	25	1375		(a)
65	20	1300	$Sf_a = 62$	मध्यका
75	10	750	$Sf_{mb} = 4020$	वर्ग
85	7	895		के ऊपर

$$\text{M.D.} = \frac{f_{ma} \quad f_{mb} \quad (f_a \quad f_a)\ \text{Median}}{\text{N}}$$

$$= \frac{4020 \quad 1410 \quad (62 \quad 38)54.8}{100}$$

$$= \frac{2610 \quad (24 \quad 54.8)}{100}$$

2610 M.D. 1313.2
100
Median 100

$$=$$

$$= \frac{1294.8}{100}$$

$$= 12.498$$

$$\text{Coeff. of M.D.} =$$

$$= \frac{12.948}{54.8} = 0.236$$

इस विधि द्वारा गणन क्रिया बहुत सरल हो जाती है और उत्तर भी वही आता है जो प्रत्यक्ष रीति द्वारा प्राप्त होता है। जहाँ सामान्तर माध्य या मध्यका अंशों में आये वहाँ लघुरीति का प्रयोग करना चाहिए।

मध्य-विचलन के गुण व दोष
(Merits and Demerits of Mean Deviation)

गुण (Merits)

1. माध्य-विचलन श्रेणी के सभी मूल्यों पर आधारित है।
2. इसका समझना व निर्धारण प्रमाप विचलन की अपेक्षा बहुत सरल है।
3. इस विचलन पर चरम (extreme) मदों का कम प्रभाव पड़ता है।

4. इसकी गणना किसी भी माध्य से अर्थात् मध्यक, मध्यका व भूयिष्ठक से हो सकती है, लेकिन मध्यका का प्रयोग अन्य माध्यों से श्रेष्ठ है।

दोष (Demerits)

1. माध्य-विचलन का सबसे बड़ा दोष यह है कि विचलन निकालते समय बीजगणित चिन्ह + और – को छोड़ दिया जाता है अर्थात् सब विचलनों को धनात्मक माना जाता है। चिन्हों का परित्याग कर देने से यह माप गणितीय दष्टिकोण से अशुद्ध एवं अवैज्ञानिक हो जाता है तथा इसका उच्च स्तरीय प्रयोग नहीं रहता।
2. इसका बीजगणितीय विवेचन सम्भव नहीं।
3. माध्य-विचलन अधिक विश्वसनीय नहीं है क्योंकि भूयिष्ठक के अनिश्चित होने के कारण उससे विचलन निकालना ही अनुपयुक्त है जबकि चरम सीमाओं से अधिक प्रभावित हो सकता है।

   उपरोक्त दोषों के कारण सांख्यिकी में इसका प्रयोग बहुत कम होता है।

IV. प्रमाप विचलन
(Standard Deviation)

अपकिरण से मापों में प्रमाप विचलन का सांख्यिकी में प्रयोग अधिक होता है। प्रमाप विचलन अपकिरण मापन की विभिन्न विधियों के दोषों को दूर करता है। विस्तार श्रेणी के न्यूनतम तथा अधिकतम मानों पर आधारित है — दो मूल्यों के अतिरिक्त अपमूल्यों का इस पर कोई प्रभाव नहीं पड़ता। चतुर्थक विचलन में श्रेणी के आधे मूल्यों को छोड़ दिया जाता है। माध्य विचलन में सभी मूल्यों का प्रयोग होता है, लेकिन यह रीति गणितीय दष्टि से अशुद्ध है क्योंकि इसमें सभी विचलनों को धनात्मक मान लिया जाता है। प्रमाप विचलन श्रेणी के सभी मूल्यों पर आधारित है। इसके निर्धारण में तथा — चिन्हों को छोड़ा नहीं जाता। इसके अतिरिक्त विचलन सदैव समान्तर माध्य से ही लिये जाते है क्योंकि वही केन्द्रीय प्रवत्ति का सर्वश्रेष्ठ माप है।

परिभाषा : किसी श्रेणी के समान्तर माध्य से निकाले गये उसके विभिन्न मद-मूल्यों के विचलनों के वर्गों का माध्य वर्गमूल उस श्रेणी का प्रमाप विचलन कहलाता है। प्रमाप-विचलन को द्वितीय घात का अपकिरण (second moment of dispersion) या मूल मध्यक वर्ग विचलन (root mean square deviation) भी कहते हैं।

व्यक्तिगत श्रेणी में प्रमाप विचलन का निर्धारण

(Calculating Standard Deviation in Individual Series)

व्यक्तिगत श्रेणी में प्रमाप विचलन ज्ञात करने की निम्न दो रीतियाँ हैं।

1. प्रत्यक्ष रीति (direct method); तथा
2. लघु रीति (short-cut method)

प्रत्यक्ष रीति : प्रत्यक्ष रीति द्वारा प्रमाप विचलन ज्ञात करने में निम्न सूत्र का प्रयोग होता है :

$$s^{+} = \sqrt{\frac{x^2}{N}}$$

s = प्रमाप विचलन (standard deviation)

$x = (X - \overline{X})$ अर्थात् मदों का मध्यक से विचलन

N = मदों की संख्या

इस रीति का प्रयोग वहीं करना चाहिये जहाँ मध्यक का मूल्य पूर्णांक में आता है अन्यथा गणन क्रिया जटिल हो जायेगी। उदाहरणार्थ, यदि किसी श्रेणी में मध्यक का मध्य 50 है तो विचलन और विचलनों के वर्ग निकालने में अधिक कठिनाई नहीं होगी। लेकिन यदि मध्यक मूल्य 50.346 है तो विवेचन लेने पर तथा विचलनों के वर्ग निकालने में बहुत कठिनाई होगी।

ऐसी स्थिति में मध्यक का उपसादन करना अर्थात 50 मानकर उपर्युक्त सूत्र का प्रयोग वांछनीय नहीं है। जहाँ मध्यक मूल्य पूर्णांक में नहीं होता वहाँ लघु रीति का प्रयोग करना चाहिये।

विधि :

1. सर्वप्रथम, श्रेणी का समान्तर माध्य ज्ञात कीजिये।
2. श्रेणी के प्रत्येक मूल्य में से समान्तर माध्य घटाइये और विचलनों को प्रदर्शित कीजिये।
3. विचलनों के वर्गों का योग Sx^2 ज्ञात कीजिये।
4. इस प्रकार प्राप्त वर्ग के जोड़ को मदों की संख्या से भाग दीजिये।
5. प्राप्त मूल्य का वर्गमूल निकालिये। यही प्रमाप विचलन है।

उदाहरण 15 : निम्न समंकों से प्रमाप विचलन ज्ञात कीजिये :

दस व्यक्तियों की मासिक मजदूरी (रू. में)

480 485 600 610 582 597 593 495 588 540

हल : **प्रमाप विचलन का परिगणन**

मजदूरी (रू. में) X	$(X - \overline{X})$ $\overline{X} = 557$ x	x^2
480	– 77	5929
485	– 72	5184
600	+ 43	1849
610	+ 53	2809
582	+ 25	625
597	+ 40	1600
593	+ 36	1396
495	– 62	3844
588	+ 31	961
540	– 17	289
SX = 5570	**Sx = 0**	**Sx^2 = 24386**

$$s = \sqrt{\frac{x^2}{N}}$$

$$= \sqrt{\frac{24386}{10}}$$

$$= \sqrt{2438.6}$$

$$= 49.382$$

****

$$s = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2}$$

d = (X – A) अर्थात् मदों का कल्पित माध्य से विचलन।

विधि :

1. पदमाला के मूल्यों में से किसी मूल्य को समान्तर माध्य मान लीजिये।
2. इस कल्पित माध्य श्रेणी के प्रत्येक मूल्य का विचलन निकालियें और इन विचलनों का योग कीजिये अर्थात Sd ज्ञात कीजिये।
3. प्रत्येक विचलन का वर्ग निकाल कर योग ज्ञात कीजिये अर्थात Sd^2 निकालिये।
4. उपर्युक्त सूत्र का प्रयोग कीजिये।

उदाहरण 16 : निम्नलिखित समंकों से प्रमाप विचलन ज्ञात कीजिये :

8 विद्यार्थियों का वजन (पौण्डों में)

110 118 125 133 140 142 135 148

हल : प्रमाप विचलन का परिगणन

वजन (पौण्डों में) X	(X – A) A = 130 d	d^2
110	– 20	400
118	– 12	144
125	– 5	25
133	+ 3	9
140	+ 10	100
142	+ 12	144
135	+ 5	25
148	+ 18	324
SX = 1,051	**Sd = 11**	**Sd^2 = 1,171**

मध्यक पूर्णांक नहीं है, इसलिये लघु रीति का प्रयोग लाभप्रद है। यहाँ 130 को कल्पित माना जाता है।

$$s = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2}$$

$$= \sqrt{\frac{1171}{8} \quad \left(\frac{11}{8}\right)^2}$$

$$= \sqrt{146.375 \quad 1.89}$$

$$= \sqrt{144.485}$$

$$= 12.02$$

विच्छिन्न श्रेणी में प्रमाप विचलन का परिगणन

(Calculating Standard Deviation in Discrete Series)

विच्छिन्न श्रेणी में प्रमाप-विचलन प्रत्यक्ष तथा लघु दोनों रीतियों से ज्ञात किया जा सकता है लेकिन व्यवहार में लघु रीति का ही अधिकतर प्रयोग किया जाता है। यदि मध्यक पूर्णांक नहीं है तो प्रत्यक्ष रीति के प्रयोग से गणना क्रिया जटिल हो जाती है इससे पहले से ही कल्पित मध्यक लेकर लघु रीति का प्रयोग करते हैं। जब प्रत्यक्ष रीति का प्रयोग होता है तो सूत्र इस प्रकार होगाः

$$\sigma = \sqrt{\frac{fx^2}{N}} \text{ and } x = (X - \overline{X})$$

पदमाला के प्रत्येक मूल्य का मध्यक से विचलन लेकर उसका वर्ग निकाला जाता है। इन वर्गों को प्रत्येक मद की आवत्ति से गुणा किया जाता है और Σfx^2 ज्ञात किया जाता है। Σfx^2 को आवत्ति के योग अर्थात् N से भाग देकर भागफल का वर्गमूल निकालते हैं और यही प्रमाप विचलन होता है।

जब लघु रीति का प्रयोग किया जाता है तो सूत्र इस प्रकार होता है :

$$\sigma = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2}$$

d = (X – A) अर्थात् मूल्यों का कल्पित मध्यक से विचलन

N = Σf अर्थात मदों की संख्या

विधि :

1. पदमाला के मूल्यों में से किसी मूल्य को कल्पित मध्यक लेते हैं।
2. इस कल्पित मध्यक से पदमाला के प्रत्येक मूल्य का विचलन निकालते हैं।
3. विचलनों को उनसे संबंधित आवत्ति से गुणा करके गुणनफलों का योग लेते हैं अर्थात Σfd निकालते हैं।
4. विचलनों व आवत्तियों की गुणाओं में फिर विचलनों को गुणा देकर इन गुणनफलों का भी जोड़ निकाल लेते हैं अर्थात् Σfd^2 ज्ञात करते है।
5. अन्त में उपर्युक्त सूत्र का प्रयोग किया जाता है।

उदाहरण 16 : निम्न सूचना से प्रमाप विचलन निकालिये :

मूल्य	:	**6**	**7**	**8**	**9**	**10**	**11**	**12**
बारम्बारता	:	**3**	**6**	**9**	**13**	**8**	**5**	**4**

हल : **प्रमाप विचलन का परिगणन**

X	***f***	**(X – 9)** ***d***	***fd***	***fd*²**
6	3	– 3	– 9	27
7	6	– 2	– 12	24
8	9	– 1	– 9	9
9	13	0	0	0
10	8	+ 1	+ 9	8
11	5	+ 2	+ 10	20
12	4	+ 3	+ 12	36
	N = 48		**$\Sigma fd = 0$**	**$\Sigma fd^2 = 124$**

$$s = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2}$$

$$= \sqrt{\frac{124}{48} \quad \left(\frac{0}{48}\right)^2}$$

$$= \sqrt{2.583} = 1.607$$

अविच्छिन्न श्रेणी में प्रमाप विचलन का निर्धारण

(Calculating Standard Deviation in Continuous Series)

अविच्छिन्न श्रेणी में प्रमाप विचलन ज्ञात करने की निम्न रीतियाँ हैं :

1. प्रत्यक्ष रीति (direct method)
2. लघु रीति (short-cut method), तथा
3. पद-विचलन रीति (step-deviation method)।

व्यवहार में सबसे अधिक पद-विचलन रीति का ही प्रयोग किया जाता है क्योंकि इससे गणन-क्रिया बहुत सरल हो जाती है।

प्रत्यक्ष रीति (Direct Method) : इस रीति के अनुसार सर्वप्रथम मध्यक ज्ञात किया जाता है। प्रत्येक मध्य-बिन्दु में से मध्यक घटाकर विचलन ज्ञात किये जाते हैं। शेष सभी क्रियायें वैसी ही रहती हैं जैसे विच्छिन्न श्रेणी में। सूत्र इस प्रकार है :

$$s = \sqrt{\frac{fx^2}{N}}$$

$x = (m - \overline{X})$ अर्थात मध्य-बिन्दुओं का समान्तर माध्य से विचलन

लघु रीति (Short-cut Method) : इस रीति के अनुसार गणन-क्रिया वही है जो खण्डित श्रेणी में है। केवल पदमाला के विभिन्न मूल्यों के स्थान पर मध्य बिन्दुओं का प्रयोग होता है। सूत्र वही जो खंडित श्रेणी में प्रयुक्त होता है, अर्थात्

$$s = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2}$$

$d = (m - A)$ अर्थात् मध्य-बिन्दुओं का कल्पित सामान्तर माध्य से विचलन

पद विचलन रीति (Step-deviation Method) : पद विचलन रीति से प्रमाप विचलन ज्ञात करते समय एक उभयनिष्ठ गुणक लिया जाता है। प्रत्येक मध्य-बिन्दु से कल्पित मध्यक घटाकर उभयनिष्ठ गुणक से भाग दिया जाता है अर्थात् $\left(\frac{m \quad A}{i}\right)$ ज्ञात करते हैं। इन विचलनों को d से प्रदर्शित करते हैं। शेष क्रिया ऊपर के सूत्र के समान ही है। पद-विचलन से प्रमाप विचलन निकालते समय निम्न सूत्र का प्रयोग किया जाता है :

$$s = \sqrt{\frac{d^2}{N} \quad \left(\frac{d}{N}\right)^2} \times i$$

$$d = \left(\frac{m \quad A}{i}\right)$$

अधिकांशतः, अविच्छिन्न श्रेणी में पद-विचलन रीति का प्रयोग होता है।

उदाहरण 18 : निम्न बंटन का माध्य व मानक विचलन ज्ञात कीजिये :

आयु (वर्षों में)	कर्मचारियों की संख्या	आयु (वर्षों में)	कर्मचारियों की संख्या
25 — 30	70	40 — 45	31
30 — 35	51	45 — 50	29
35 — 40	47	50 — 55	22

हल : **माध्य व प्रमाप विचलन का परिगणन।**

आयु	*m*	*f*	**(*m* – 42.5)/5** ***d***	*fd*	*fd* ²
25 — 30	27.5	70	– 3	– 210	630
30 — 35	32.5	51	– 2	– 10	204
35 — 40	37.5	47	– 1	– 47	47
40 — 45	42.5	31	0	0	0
45 — 50	47.5	29	+ 1	+ 29	29
50 — 55	52.5	22	+ 2	+ 44	88
		N = 250		**S*fd* = – 286**	**S*fd*² = 998**

$$\overline{X} = A + \frac{fd}{N} \times i$$

A = 42.5, S*fd* = – 286, N = 250, *i* = 5

\

$$\overline{X} = 42.5 - \frac{286}{250} \times 5$$
$$= 42.5 - 5.72$$
$$= 36.78$$

$$s = \sqrt{\frac{fd^2}{N} \quad \left(\frac{fd}{N}\right)^2} \times i$$

$$= \sqrt{\frac{998}{250} \quad \left(\frac{286}{250}\right)^2} \times 5$$

$$= \sqrt{3.992 \quad 1.309} \times 5$$

$$= 1.638 \times 5 = 8.19$$

प्रमाप विचलन की गणितीय विशेषतायें
(Mathematical Properties of Standard Deviation)

निम्नलिखित गणितीय विशेषताओं के कारण प्रमाप विचलन का सांख्यिकी में बहुत महत्वपूर्ण स्थान है :

1. **सामूहिक प्रमाप विचलन** (Combined Standard Deviation) : जिस प्रकार अलग समूहों के समान्तर माध्य तथा मदों की संख्या की सहायता से सामूहिक माध्य निकाला जा सकता है उसी प्रकार विभिन्न समूहों के प्रमाप विपलनों, माध्यों व मद संख्याओं के आधार पर सामूहिक प्रमाप विचलन का परिगणन किया जा सकता है। सामूहिक प्रमाप विचलन ज्ञात करने का सूत्र निम्नलिखित हैः

$$s_{12} = \sqrt{\frac{N_1 \ {}^2_1 \quad N_2 \ {}^2_2 \quad N_1 d_1^2 \quad N_2 d_2^2}{N_1 \quad N_2}}$$

s_{12} = सामूहिक प्रमाण विचलन (combined standard deviation)

s_1 = प्रथम भाग का प्रमाप विचलन (standard deviation of the first group)

s_2 = द्वितीय भाग का प्रमाप विचलन (standard deviation of the second group)

N_1 = पहले भाग में मदों की संख्या (number of items in the first group)

N_2 = दूसरे भाग में मदों की संख्या (number of items in the second group)

$d_1 = \left(\overline{X}_1 \quad \overline{X}_{12}\right)$ = प्रथम वर्ग के सामान्तर माध्य तथा सामूहिक समान्तर माध्य में अंतर (difference between the arithmetic mean of the first group and the combined arithmetic average)

यदि तीन भागों का सामूहिक प्रमाप विचलन ज्ञात करना हो तो सूत्र इस प्रकार होगा।

$$s_{123} = \sqrt{\frac{N_1 \ {}^2_1 \quad N_2 \ {}^2_2 \quad N_3 \ {}^2_3 \quad N_1 d_1^2 \quad N_2 d_2^2 \quad N_3 d_3^2}{N_1 \quad N_2 \quad N_3}}$$

यहाँ $d_1 \left(\overline{X}_1 \quad \overline{X}_{123}\right), d_2 \quad \left(\overline{X}_2 \quad \overline{X}_{123}\right), d_3 \quad \left(\overline{X}_3 \quad \overline{X}_{123}\right)$

उदाहरण 19 : (क) निम्नलिखित आँकड़ों से ज्ञात कीजिए :

(अ) सामूहिक माध्य, (ब) सामूहिक प्रमाप विचलन

$\sqrt{\frac{31500}{630}}$

	Firm A	Firm B
कुल संख्या	**280**	**350**
माध्य	**45**	**54**
विचरण मापांक	**36**	**16**

हल : (अ) सामूहिक माध्य

$$\overline{X}_{12} = \frac{N_1 \overline{X}_1 \quad N_2 \overline{X}_2}{N_1 \quad N_2}$$

$N_1 = 280, \overline{X}_1 = 45, N_2 = 350, \overline{X}_2 = 54$

$$\therefore \quad \overline{X}_{12} = \frac{280(45) \quad 350(54)}{280 \quad 350}$$

$$\text{अर्थात सामूहिक माध्य} = \frac{12600 \quad 18900}{630} = \qquad = 50$$

(ब) सामूहिक प्रमाप विचलन :

$$s_{12} = \sqrt{\frac{N_1 \ {}^2_1 \quad N_2 \ {}^2_2 \quad N_1 d_1^2 \quad N_2 d_2^2}{N_1 \quad N_2}}$$

$N_1 = 280, s_1 = \sqrt{36}, = 6, N_2 = 350$

$$s = \sqrt{16} = 4, d_1$$

$$= \left(\overline{X}_1 \quad \overline{X}_{12}\right)$$

$$= (45 - 50) = -5$$

$$d_2 =$$

$$= 54 - 50 = 4$$

$$s_{12} =$$

$$= \sqrt{\frac{10080 \quad 5600 \quad 7000 \quad 5600}{630}}$$

$$= \sqrt{\frac{28280}{630}} = 6.7$$

उदाहरण : (ख) निम्न समंकों से सिद्ध कीजिये कि समान्तर माध्य 16 तथा प्रमाप विचलन 7.2 है :

वर्ग	N	$\overline{X}$	s
1	**200**	**25**	**3**
2	**250**	**10**	**4**
3	**300**	**15**	**5**

हल :

$$\left(\overline{X}_2 \quad \overline{X}_{123}\right) \quad (10 \quad 16) \quad 6 \qquad \frac{5000 \quad 2500 \quad 4500}{750 \quad 750}$$

$$\sqrt{\frac{280\,(36) \quad 350\,(16) \quad 280\,(5)^2 \quad 350\,(4)^2}{280 \quad 350}}$$

$$\overline{X}_{123} = \frac{N_1\overline{X}_1 \quad N_2\overline{X}_2 \quad N_3\overline{X}_3}{N_1 \quad N_2 \quad N_3}$$

$\overline{X}_1 = 25$, $\overline{X}_2 = 10$, $\overline{X}_3 = 15$, $N_1 = 200$, $N_2 = 250$, $N_3 = 300$

$$\overline{X}_{123} = \frac{(200 \quad 25) \quad (250 \quad 10) \quad (300 \quad 15)}{200 \quad 250 \quad 300}$$

$$=$$

$$=$$

$$s_{123} = \sqrt{\frac{N_1 \;{}_1^2 \quad N_2 \;{}_2^2 \quad N_3 \;{}_3^2 \quad N_1 d_1^2 \quad N_2 d_2^2 \quad N_3 d_3^2}{N_1 \quad N_2 \quad N_3}}$$

$$d_1 = \left(\overline{X}_1 \quad \overline{X}_{123}\right) \quad (25 \quad 16) \quad 9$$

$$d_2 =$$

$$d_3 = \left(\overline{X}_3 \quad \overline{X}_{123}\right) \quad (15 \quad 16) \quad 1$$

$$= \sqrt{\frac{200\,(3)^2 \quad 250\,(4)^2 \quad 300\,(5)^2 \quad 200\,(9)^2 \quad 250(\;6)^2 \quad 300(\;1)^2}{200 \quad 250 \quad 300}}$$

$$= \sqrt{\frac{1800 \quad 4000 \quad 7500 \quad 16200 \quad 9000 \quad 300}{750}}$$

$$= \sqrt{\frac{38800}{750}}$$

$$= \sqrt{51.733}$$

$$= 7.19$$

2. **क्रमानुसार दी गई संख्याओं का प्रमाप विचलन** (Standard Deviation of Natural Numbers) : यदि क्रमानुसार दी गई संख्याओं का प्रमाप विचलन ज्ञात करना हो तो नीचे दिये सूत्र का प्रयोग किया जा सकता है :

$$s = \sqrt{\frac{N^2 \quad 1}{12}}$$

इस सूत्र द्वारा विचलन निकालते समय समान्तर माध्य व विचलन, आदि निकालने की आवश्यकता नहीं है। उदाहरणार्थ, 1 से 10 तक की संख्याओं का प्रमाप विचलन निम्न प्रकार से ज्ञात किया जा सकता है :

$$s = \sqrt{\frac{N^2 \quad 1}{12}} = \sqrt{\frac{99}{12}} = 2.87$$

3. एक संमित (normal or symmetrical) या साधारण रूप से असंमित वितरण (moderately asymmetrical distribution) में माध्य एवं प्रमाप विचलन के आधार पर वे सीमायें निर्धारित की जा सकती हैं, जिनमें निश्चित प्रतिशत मूल्य पाये जाने की सम्भावना है। ये सीमायें इस प्रकार हैं :

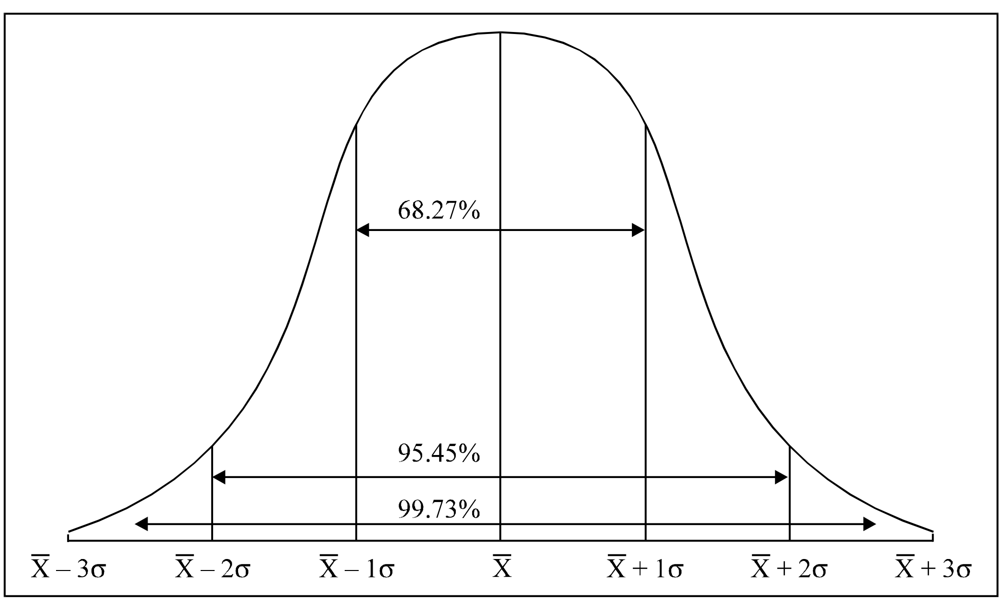
   Mean ± 1s covers 68.27% items

   Mean ± 2s covers 95.45% items

   Mean ± 3s covers 99.73% items

   निम्न चित्र में ये सीमायें स्पष्ट रूप से प्रदर्शित की गई हैं। निदर्शन में ये सीमायें बहुत महत्वपूर्ण हैं।

4. समान्तर माध्य के विचलन लिये जाने के कारण प्रमाप विचलन में विचलन वर्गों का जोड़ न्यूनतम होता है, *i.e.*, Sd^2 = minimum, अर्थात् यदि समान्तर माध्य के अतिरिक्त किसी अन्य मूल्य के मदों का विचलन लेकर वर्ग लिये जायें तो यह योग अधिक होगा। यही कारण है कि प्रमाप विचलन हमेशा समान्तर माध्य से ही विचलन लेकर ज्ञात किया जाता है।



विचरक गुणक

(Coefficient of Variation)

प्रमाप विचलन अपकिरण का निरपेक्ष माप है इसलिए इस आधार पर दो या दो से अधिक श्रेणियों की तुलना सम्भव नहीं, श्रेणियों के तुलनात्मक अध्ययन के लिय विचरण का गुणक निकाला जाता है। लेकिन विचरण गुणक प्रायः दशमलव अंकों में आता है, इसलिये विचलन के अंतर का ठीक अनुमान नहीं हो पाता। इस कठिनाई के समाधान के लिए विचरण-गुणक का सहारा लिया जाता है। इस माप का प्रयोग सर्वप्रथम प्रसिद्ध वैज्ञानिक पियर्सन (Karl Pearson) ने किया और इसलिये इसे कार्ल पियर्सन का विचरण गुणक भी कहते हैं। विचरण गुणक निकालने के लिए प्रमाप विचलन के गुणक को 100 से गुणा कर देते हैं अर्थात निम्न सूत्र का प्रयोग किया जाता है :

विचरण गुणक (Coefficient of Variation) $= \frac{}{X} \times 100$

विचरण-गुणक का प्रयोग दो समूहों की स्थिरता (variability), संगति (consistency), सजातीयता (homogenity), स्थिरता (stability), तथा एकरूपता (uniformity) की तुलना करने में किया जाता है। जिस समंक श्रेणी का विचरण-गुणक अधिक होता है वह अधिक अस्थिर (more variable); कम स्थिर (less stable), कम एकरूपता (less uniform या less homogeneous) तथा कम संगत (less consistent) कहलाती है। इसके विपरीत जिस समंक श्रेणी का विचरण-गुणक सम होता है वह अधिक स्थिर, एकरूप संजातीय तथा संगत कहलाती है।

अपकिरण के विभिन्न मापों के बीच संबंध

1. माध्य-विचलन प्रमाप विचलन का 0.7979 या $\frac{4}{5}$ होता है।

$$\text{M.D.} = \frac{4}{5}\text{s}$$

2. चतुर्थक विचलन प्रमाप विचलन का 0.6745 या $\frac{2}{3}$ होता है।

$$\text{Q.D.} = \frac{2}{3}\text{s}$$

3. चतुर्थक विचलन माध्य-विचलन का $\frac{5}{6}$ होता है।

$$\text{Q.D.} = \frac{5}{6}\text{ M.D.}$$

4. प्रमाप विचलन का 6 गुणा, चतुर्थक विचलन का 9 गुणा और माध्य विचलन का 7.5 गुणा आपस में बराबर होते हैं:

$$6\text{s} = 9\text{ Q.D.} = 7.5\text{ M.D.}$$

उदाहरण 21 : 1 किसी श्रेणी के लिए माध्य-विचलन 15 है। चतुर्थक विचलन ज्ञात कीजिये।

2. एक आवत्ति वितरण में जो कि बहुत विषम नहीं है, माध्य -विचलन का मूल्य 12.4 है। प्रमाप विचलन ज्ञात कीजिये।

3. विचरण-गुणांक ज्ञात कीजिए यदि ततीय चतुर्थक (Q_3) तथा प्रथम चतुर्थक (Q_1) 25 हो।

4. विचरण-गुणांक ज्ञात कीजिए यदि विचरण मापांक (variance) का मूल्य 225 तथा सामान्तर माध्य का 40 हो।

हल : 1. $\text{MD} = 15$

$$\text{Q.D.} = \frac{5}{6}\text{ M.D.}$$

$$\text{Q.D.} = \frac{5}{6} \times 15 = 12.5$$

2.
$$s = \frac{5}{6} \text{ M.D.}$$

$$\text{M.D.} = 12.4$$

$$s = \frac{5}{4} \times 12.4 = 15.5$$

3.
$$\text{Q.D.} = \frac{Q_3 \quad Q_1}{2}$$

$$= \qquad = 7.5$$

$$\text{Q.D.} = \quad \text{M.D.}$$

$$\frac{5}{6} \text{ M.D.} = 7.5$$

$$\text{M.D.} = \frac{7.5 \times 6}{5} = 9$$

4.
$$\text{C.V.} = \frac{\quad}{\overline{X}} \quad 100 \quad \frac{40 \quad 25}{2} \quad \frac{5}{6} \quad \sqrt{(225)} \quad 15$$

$$s = \sqrt{(\text{Variance})}$$

$$=$$

$$\text{C.V.} = \frac{15}{40} \times 100 = 37.5$$

प्रमापविचलन के गुण व दोष

(Merits and Demerits of Standard Deviation)

गुण (Merits)

1. यह माप श्रेणी के सभी मूल्यों पर आधारित होता है।
2. यह अपकिरण का स्पष्ट और निश्चित माप है।
3. अपकिरण के अन्य मापों की अपेक्षा प्रमाप विचलन पर निर्देशन परिवर्तनों का सबसे कम प्रभाव पड़ता है।
4. प्रमाप विचलन निकालते समय बीजगणितीय नियमों का पूर्ण रूप से पालन किया होता है, इसलिये इसका उच्चस्तरीय अध्ययन में बहुत प्रयोग होता है।
5. प्रमाप विचलन अपकिरण का सवेश्रेष्ठ माप है।

दोष (Demerits)

1. अन्य मापों की अपेक्षा इसकी गणन-क्रिया कठिन है।
2. इसके मध्यक की सहायता से निकाला जाता है इसलिये इस पर चरम मूल्य का अधिक प्रभाव पड़ता है।

जिस प्रकार मध्यक की केन्द्रीय प्रवत्ति के मापों में सबसे अधिक प्रयोग होता है उसी प्रकार प्रमाप विचलन भी अपकिरण के मापों में सबसे अधिक प्रचलित है। सांख्यिकी में मध्यक तथा प्रमाप विचलन का विशेष स्थान है क्योंकि बहुत से सूत्र इन पर आधारित है।

लारेंज वक्र

(Lorenz Curve)

अपकिरण को प्रदर्शित करने की यह एक बिन्दुरेखीय रीति है। सर्वप्रथम इस प्रकार के वक्र का प्रयोग डॉ. मैक्स ओ लॉरेज (D. Max O' Lorenz) ने किया और उन्हीं के नाम पर इस वक्र का नाम 'लॉरेंज वक्र' पड़ा।

लॉरेंज वक्र एक संचयी प्रतिशत वक्र (cumulative percentage curve) है, इसे बनाने की क्रिया इस प्रकार है :

1. मूल्यों (वर्गान्तर होने पर मध्य-बिन्दुओं) के संचयी योग ज्ञात करके अंतिम संचयी योग को 100 मानकर प्रत्येक संचयी मूल्य को प्रतिशत में बदलिए।
2. मूल्यों की भाँति ही आवत्तियों को संचयी करके अंतिम संचयी आवत्ति को 100 मानते हुए सभी आवत्तियों को प्रतिशत में बदल लीजिये।
3. संचयी मूल्यों के प्रतिशत कोटि अक्ष (Y-axis) पर तथा संचयी आवत्ति को 100 मानते हुए सभी आवत्तियों को प्रतिशत में बदल लीजिये।
4. भुजाक्ष पर प्रतिशतों का क्रम 100 से प्रारम्भ होकर (0) तक जाता है तथा कोटि पर 0 से 100 तक जाता है।
5. 0 से 100 को एक सीधी रेखा से मिला देते हैं। इस रेखा को समान-वितरण की रेखा (line of equal distribution) कहते हैं।
6. अब संचयी मूल्यों के प्रतिशत और संचयी आवत्तियों के प्रतिशत को क्रमानुसार अंकित कीजिए। इस प्रकार से अंकित बिन्दुओं को आपस में मिला देने से जो वक्र बनता है वही लॉरेज वक्र होता है।

लॉरेज वक्र समान-वितरण रेखा के जितना पास होगा अपकिरण की मात्रा उतनी ही कम होगी। अर्थात् वितरण में उतनी ही कम असमानतायें होंगी। लॉरेज वक्र जितना समान वितरण रेखा से दूर हो जायेगा, अपकिरण या असमानता की मात्रा उतनी ही बढ़ती जायेगी। यदि दो लॉरेज वक्रों में तुलना करनी हो तो वक्र समान-वितरण रेखा से अधिक दूरी पर होंगे, उस समंकमाला में अपकिरण की मात्रा अधिक होगी।

लॉरेज वक्र के गुण व दोष

(Merits and Demerits of Lorenz Curve)

गुण (Merits)

1. चित्र देखने मात्र से ही अपकिरण का ज्ञान हो सकता है।
2. लॉरेज वक्र आकर्षक होता है इसलिये पाठक को इसके अध्ययन में समंको की अपेक्षा अधिक रुचि रहती है।
3. दो या दो से अधिक श्रेणियों के अपकिरण की मात्रा की तुलना बड़ी संख्या से सम्भव हो जाती है।
4. धन, आय, मजदूरी, लाभ आदि के वितरण की असमानताओं को एक प्रकार का माना जा सकता है।

दोष (Demerits)

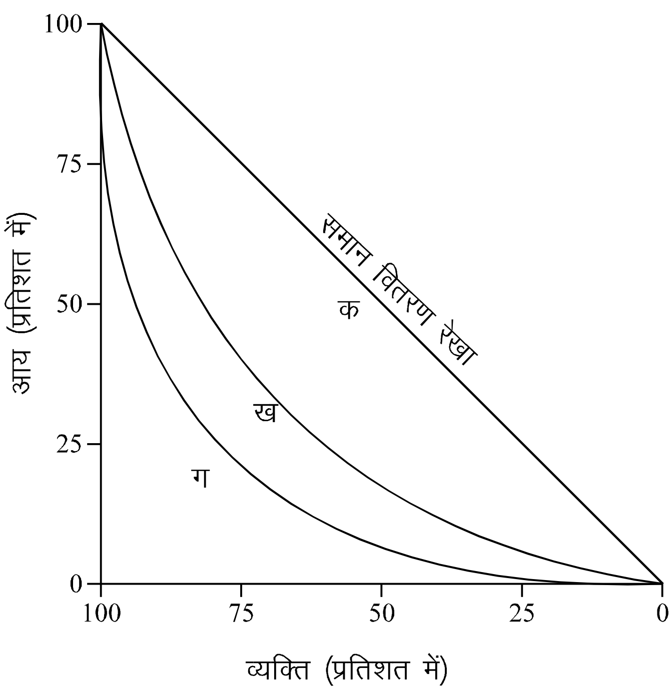
1. इसका प्रमुख दोष यह है कि इससे अपकिरण का अंकात्मक माप नहीं होता।
2. इसके बनाने की क्रिया कुछ कठिन है :

उदाहरण 22 : निम्न समंकों से एक लॉरेंज वक्र बनाइये :

आय	व्यक्तियों की संख्या हजारों में		
(हजार रू. में)	वर्ग 'क'	वर्ग 'ख'	वर्ग 'ग'
20	10	16	30
40	20	14	12
80	40	10	4
100	50	6	2
160	80	4	2

हल :

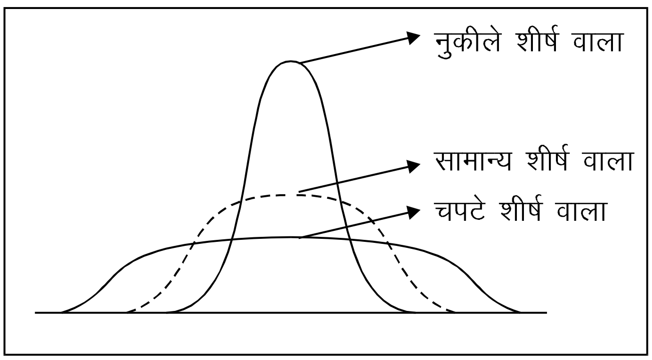
आय			वर्ग 'क'			वर्ग 'ख'			वर्ग 'ग'		
आय हजारों में	संचयी आय	संचयी प्रतिशत	व्यक्तियों की संख्या (हजारों में)	संचयी नम्बर	संचयी प्रतिशत	व्यक्तियों की संख्या (हजारों में)	संचयी नम्बर	संचयी प्रतिशत	व्यक्तियों की संख्या (हजारों में)	संचयी नम्बर	संचयी प्रतिशत
20	20	5	10	10	5	16	16	32	30	30	60
40	60	15	20	30	15	14	30	60	12	42	84
80	140	35	40	70	35	10	40	80	4	46	92
100	240	60	50	120	60	6	46	92	2	48	96
160	400	100	80	200	200	4	50	100	2	50	100



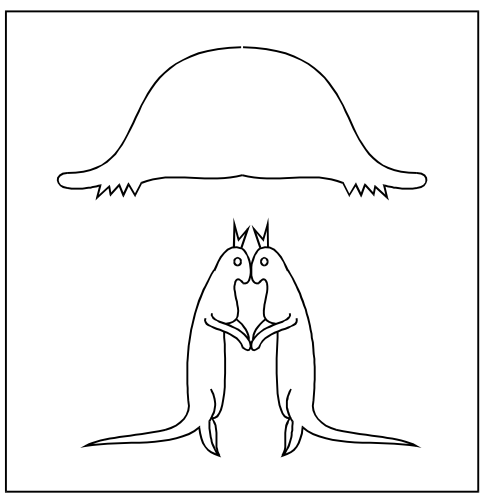
पथुशीर्षत्व
(Kurtosis)

पथुशीर्षत्व वह सांख्यिकीय माप है जिसके आधार पर आवत्ति -वितरण से संबंधित आवत्ति वक्र (frequency curve) के शीर्ष की बनावट का ज्ञान किया जाता है। आवत्ति वक्र सामान्य शीर्ष वाला है अथवा नुकीले शीर्ष वाला है अथवा चपटे शीर्ष वाला, इस बात की जानकारी हम पथुशीर्षत्व माप के आधार पर कर सकते हैं। सरल शब्दों में इस माप से हमें इस बात का आभास हो जाता है कि समंक श्रेणी के मध्य भाग में आवत्तियों का जमाव किस प्रकार का है। यदि आवत्तियों का जमाव सामान्य है तो संबंधित आवत्ति वक्र सामान्य शीर्ष वाला (Normal or Mesokurtic) कहलाता है। यदि आवत्तियों का जमाव श्रेणी के मध्य भाग में बहुत अधिक होता है तो संबंधित आवत्ति वक्र नुकीले शीर्ष वाला (Leptokurtic) कहलाता है और यदि आवत्तियों का जमाव श्रेणी के मध्य भाग में बहुत कम होता है तो संबंधित आवत्ति वक्र चपटे शीर्ष वाला (platykurtic) कहलाता है। क्रॉक्सटन एवं काउर्डन के अनुसार ''पथुशीर्षत्व का माप उस मात्रा को प्रकट करता है जिसमें आवत्ति वितरण का वक्र नोंकदार या चपटे शीर्ष वाला होता है'' (A measure of Kurtosis indicates the degree to which curve of a frequency distribution is peaked of flat topped).

तीनों अवस्थाओं को हम भिन्न चित्र द्वारा प्रदर्शित कर सकते हैं :



प्रसिद्ध अंग्रेज सांख्यिक विलियम एस. गोसट ने चपटे शीर्ष वाले वक्र की तुलना छोटी पूँछ और चपटी पीठ वाले जानवर ''प्लैटिपस'' (Platypus) से और नुकीले शीर्ष वाले वक्र की तुलना ऊँचे शीर्ष व लम्बी पूँछ वाले जानवर 'कंगारू' (Kangaroo) से की है, जिसको कि चित्र में निम्न प्रकार दिखाया जा सकता है :



शीर्षत्व का माप (Measure of Kurtosis) : शीर्षत्व का माप चतुर्थ एवं द्वितीय परिघातों के आधार पर किया जाता है। कार्ल पियर्सन ने निम्न सूत्र का प्रयोग किया है :

$$\text{Kurtosis of } b_2 = \frac{{}_4}{{}_2^{\,2}},$$

यदि b_2 का मूल्य 3 के बराबर है तो वक्र सामान्य अथवा (Mesokurtic) होगा।

यदि b_2 का मूल्य 3 से अधिक है तो वक्र शीघ्र का होगा।

और यदि b_2 का मूल्य 3 से कम है तो वक्र चपटे शीर्ष वाला होगा।

बीटा (b_2) के अतिरिक्त शीर्ष निकालने के लिए गामा (g) का भी प्रयोग किया जाता है। उस संबंध में अग्रांकित सूत्र है :

$$g_2 = b_2 - 3 \text{ or } \frac{\;_4 \quad 3\;_2^2}{\;_2^2}$$

उपर्युक्त सूत्र से यदि :

g_2 का मूल्य शून्य होता है तो वक्र सामान्य शीर्ष वाला वक्र होता है।

g_2 ऋणात्मक होता है तो वक्र चपटा होता है।

g_2 धनात्मक होता है तो वक्र नुकीला होता है।

उदाहरण 25 : उदाहरण 5 में दिये गये आवत्ति -वितरण के शीर्ष की बनावट के संबंध मे अपने विचार प्रस्तुत कीजिऐ।

हल : उदाहरण 5 के हल के आधार पर :

$$m_2 = 636.96;$$
$$m_4 = 13{,}38{,}557.9$$
$$b_4 = \frac{\;_4}{\;_2^2}$$
$$= \frac{13{,}38{,}557.9}{(636.96)^2} = 3.29$$

क्योंकि b_2 का मूल्य 3 से अधिक है इसलिए वक्र नुकीले शीर्ष वाला (Leptokurtic) है।

उदाहरण 24 : निम्नलिखित सामग्री आर्थिक विश्लेषण हेतु एक अर्थशास्त्री को दी गई। ये समंक प्रतिदर्श के रूप में च न गए कुछ डनलप टायर्स (Dunlop Tyres) की आयु से संबंधित है। बटन की शीर्ष बनावट के संबंध में अपने विचार प्रस्तुत कीजिए।

N = 100, S*fdx* = 50, S*fd*²*x* = 1967.0; S*fd*³*x* = 2925.8 तथा S*fd*⁴*x* = 86650.2;

हल : दी गई सूचना के आधार पर

$$n_1 = \frac{fdx}{N} = \frac{50}{100} = 0.5$$

$$n_2 = \frac{fd^2x}{N} = \frac{1967.2}{100} = 19.672$$

$$n_3 = \frac{fd^3x}{N} = \frac{2925.8}{100} = 29.258$$

$$n_4 = \frac{fd^4x}{N} = \frac{86650.2}{100} = 866.502$$

$$m_2 = n_2 - (n_1)^2$$
$$= 19.672 - (0.5)^2$$

$$= 19.672 - 0.25$$
$$= 19.422$$
$$m_4 = m_4 - 4n_3n_1 + 6n_2(n_1)^2 - 3(n_1)^4$$
$$= 866.502 - 4(29.258)(.5) + 6(19.672)(.5)^2 - 3(.5)^4$$
$$= 866.502 - 58.516 + 29.5080 - .1875$$
$$= 837.3065$$

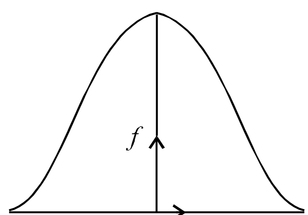
पथुशीर्षत्व माप (अर्थात b_2)

$$= \frac{\mu_4}{(\mu_2)^2} = \frac{837.3065}{(19.422)^2} = 2.22$$

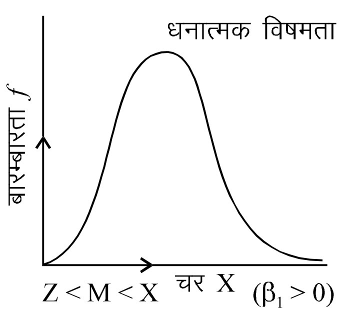
क्योंकि b_2 का परिकलित मूल्य 3 से कम है, अतः संबंधित आवत्ति वक्र चपटे शीर्ष वाला (Playkurtic) है।

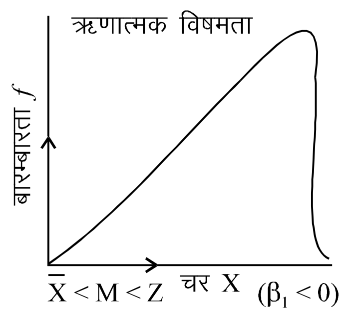
विषमता
(Skewness)

किसी समंक माला में सममितता के अभाव को विषमता कहते हैं। इसके द्वारा श्रेणी के स्वरूप का पता चलता है। यह एक ऐसा संख्यात्मक माप होता है जो कि समंक माला की असममित प्रकृति को प्रकट करता है।



X
$Z < M < X \rightarrow (\beta_1 > 0)$
सममित बंटन





विषमता अंक श्रेणी में समान्तर माध्य के दोनों ओर विभिन्न समंको के विचरणों का तुलनात्क अध्ययन करती है। इसके द्वारा यह ज्ञात किया जा सकता है कि समंक माला में समान्तर माध्य से दोनों भागों का विचरण समान है अथवा एक भाग अधिक है और एक कम। दूसरे शब्दों में अंक श्रेणी सममिता (Symmetry) से कितनी दूर है एवं उसका सामूहिक झुकाव किसी दिशा में है। यदि अंक श्रेणी में सममितता है तो इसकी वक्र रेखा दोनों दिशाओं में समान झुकाव बतलाएगी, यदि वक्र रेखा की दोनों दिशाओं में समान झुकाव नहीं है तो इसका अर्थ है कि अंकों में असममितता (Asymmetry) है, अर्थात विषमता है। एक विषम अंक श्रेणी की वक्र रेखा का झुकाव दांई अथवा पश्चिम की ओर अधिक हो तो विषमता धनात्मक होगी और यदि इसका झुकाव बायीं ओर अथवा पूर्व की ओर अधिक होगा तो ऋणात्मक विषमता होगी, अतः विषमता धनात्मक (Positive) अथवा ऋणात्मक (Negative) होती है।

विषमता का माप : विषमता का माप निरपेक्ष (Absolute) या सापेक्ष (Relative) दोनों की प्रकार से ज्ञात किया जा सकता है। सापेक्ष विषमता को विषमता गुणक (Coefficient of Skewness) भी कहते हैं। निरपेक्ष माप की अपेक्षा सापेक्ष माप ही अधिक उपयोग में आता है। अतः विषमता ज्ञात करने के लिए, प्रायः विषमता गुणांक का ही परिकलन किया जाता है। विषमता के दो प्रमुख माप हैं।

विषमता का प्रथम माप : इस ज्ञात करने के लिए कार्ल पियर्सन द्वारा बताए गए नीचे दिये सूत्रों का उपयोग किया जाता है। जिस श्रेणी में भूयिष्ठक स्पष्ट हो वहाँ

विषमता (Skewness) $= \overline{X} - Z$ (निरपेक्ष माप)

विषमता गुणक (Coefficient of Skewness) या $j = \dfrac{\left(\overline{X} \quad Z\right)}{}$ (सापेक्ष माप)

इस सूत्र का प्रयोग करने पर सापेक्ष विषमता माप का मूल्य +1 से –1 के बीच होगा इसके बाहर नहीं। यदि भूयिष्ठक किन्हीं कारणों से अस्पष्ट हो या उसका निर्धारण सम्भव न हो तो

विषमता (Skewness) $= 3\,(\overline{X} - M)$ (निरपेक्ष माप)

विषमता गुणक (Coefficient of Skewness) या $j = \dfrac{3\left(\overline{X} \quad M\right)}{}$ (सापेक्ष माप)

उदाहरण 25 : निम्न श्रेणी के लिए विषमता गुणक ज्ञात कीजिए :

प्राप्तांक	**:**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
छात्र संख्या	**:**	**1**	**3**	**6**	**10**	**5**	**4**	**3**

प्राप्तांक *x*	**छात्रों की संख्या** *f*	**कल्पित माध्य (5)** *dx*	*fdx*	*fxd*^{2}
2	1	– 3	– 3	9
3	3	– 2	– 6	12
4	6	– 1	6	6
5	10	0	0	0
6	5	+ 1	+ 5	5
7	4	+ 2	+ 8	16
8	3	+ 3	+ 9	27
	32		**+ 7**	**75**

उपरोक्त श्रेणी के लिए :

समान्तर माध्य या $\overline{X}$ $= A + \dfrac{fdx}{N}$

$= 5 + \dfrac{7}{32}$

$= 5 + 0.219$

$= 5.219$ अंक

भूयिष्ठक या Z = 5 अंक (अवलोकन से)

$$\text{प्रमाप विचलन या } s = \sqrt{\frac{fdx^2}{N} \quad \left(\frac{fdx}{N}\right)^2}$$

$$= \sqrt{\frac{75}{32} \quad \left(\frac{7}{32}\right)^2} = \sqrt{(2.344 \quad 0.219)^2\, 9}$$

$$= \sqrt{2.344 \quad 0.048} \quad 8 = \sqrt{2.296}$$

$$= 1.52 \text{ अंक}$$

$$\text{विषमता या } Sk = \overline{X} - Z$$

$$= 5.219 - 5 = 0.219 \text{ अंक}$$

$$\text{विषमता गुणक या } j = \frac{(\overline{X} \quad Z)}{} = \frac{5.219 \quad 5}{1.52}$$

$$\text{या} \quad j = \frac{0.219}{1.52} = 0.14$$

उपरोक्त श्रेणी के लिए विषमता गुणक 0.14 है।

उदाहरण 26 : निम्न श्रेणी के लिए विषमता-गुणक निकालिए :

मूल्य :	**0-5**	**5-10**	**10-15**	**15-20**	**20-25**	**25-30**	**30-35**	**35-40**
बारंबारता :	**2**	**5**	**7**	**13**	**21**	**66**	**8**	**3**

वर्ग	**बारम्बारता** f	**मध्य मूल्य**	**कल्पित माध्य (22.5) पद विचलन** dx_1*	$fd'x$	$fd'x^2$	**सामूहिक बारम्बारता**				
0-5	2	2.5	– 4	– 8	32					
5-10	5	7.5	– 3	– 15	45	7				
							12	14		
10-15	7	12.5	– 2	– 14	28				25	
15-20	$13f^0$	17.5	– 1	– 13	13	20				
							34			41
								50		
20-25	$21f^1$	22.5	0	0	0					
25-30	$16f^2$	27.5	1	16	16	37			45	
							24			27
30-35	8	32.5	2	16	32					
35-40	3	37.5	3	9	27	11				27
योग	**75**			**– 9**	**193**					

*** पद विचलन 5 का भाग देकर ज्ञात किये गये है।**

सारिणी के अवलोकन से तथा बारंबारताओं का समूहीकरण (Grouping) करने से ज्ञात होता है कि भूयिष्ठक 20-25 वाले वर्ग में है।

$$\text{Coeff. of range} = \frac{\text{L} \quad \text{S}}{\text{L} \quad \text{S}}$$

$$\text{भूयिष्ठक या Z} = \text{I}_1 + \frac{1}{1 \quad 2} \times i$$

$$= 20 + \frac{(21 \quad 13)}{(21 \quad 13) \quad (21 \quad 16)} \times 5$$

$$= 20 + \frac{8}{8 \quad 5} \times 5$$

$$= 20 + \frac{8}{13} \times 5$$

$$= 20 + 3.1$$

$$= 23.1$$

$$\text{सामान्तर माध्य या } \overline{\text{X}} = \text{A} + \frac{fd'x}{\text{N}} \times i$$

$$= 22.5 + \frac{9}{75} \times 5$$

$$= 22.5 - 0.6 \quad \frac{1.2}{8}$$

$$= 21.9$$

$$\text{प्रमाप विचलन या s} = \sqrt{\frac{fx'x^2}{\text{N}} \quad \left(\frac{fd'x}{\text{N}}\right)^2} \times i$$

$$= \sqrt{\frac{193}{75} \quad \left(\frac{9}{75}\right)^2} \times 5$$

$$= \sqrt{2.5733 \quad 0.0144 \quad 5}$$

$$= \sqrt{2.5589} \times 5$$

$$= 1.6 \times 5$$

$$= 8.0$$

$$j = \frac{\overline{\text{X}} \quad \text{Z}}{}$$

$$= \frac{21.9 \quad 23.1}{8}$$

$$=$$

$$= 0.15$$

यह विषमता ऋणात्मक है।

उदाहरण 27 : निम्न श्रेणी विषमता गुणक ज्ञात कीजिए :

प्राप्तांक	**:**	**0-10**	**10-20**	**20-30**	**30-40**	**40-50**	**50-60**	**60-70**	**70-80**
छात्र. सं.	**:**	**10**	**40**	**20**	**0**	**10**	**40**	**16**	**14**

उपर्युक्त श्रेणी में भूयिष्ठक एक नहीं है। यह 10-20 व 50-60 दोनों ही वर्गों में है। ऐसी अवस्था में $j = \frac{3\overline{X} \quad M}{\sigma}$ सूत्र का प्रयोग करना चाहिये।

प्राप्तांक	छात्र सं f	संचयी बारम्बारता	मध्य मूल्य x	कल्पित माध्य (35) से पद विचलन $d'x$*	$fd'x$	$fd'x^2$
0-10	10	10	5	– 3	– 30	90
10-20	40	50	15	– 2	– 80	160
20-30	20	70	25	– 1	– 20	20
30-40	0	70	35	0	0	0
40-50	10	80	45	1	10	10
50-60	40	120	55	2	80	160
60-70	16	136	65	3	48	144
70-80	14	150	75	4	56	224
योग	**150**			**64**	**808**	

* विचलनों में 10 का भाग देकर ज्ञात किए गए हैं।

$$\text{सामान्तर माध्य } \overline{X} = A + \frac{fd'x}{N} \times i$$

$$= 35 + \frac{64}{150} \times 10$$

$$= 35 + 4.27$$

$$= 39.27 \text{ अंक}$$

मध्यका क्रमांक $\frac{N}{2} = \frac{150}{2} = 75$ यह 40-50 वाले वर्ग में है, अतः

$$\text{मध्यका का } M = I_1 + \frac{i}{f}\left(\frac{N}{2} \quad C_0\right)$$

$$= 40 + \frac{10}{10}(75 - 70)$$

$$= 40 + 5$$

$$= 5 \text{ अंक}$$

$$\text{प्रमाप विचलन या } s = \sqrt{\frac{fx'x^2}{N} \quad \left(\frac{fd'x}{N}\right)^2} \times i$$

$$= \sqrt{\frac{808}{150} \quad \left(\frac{64}{150}\right)^2} \times 10$$

$$= \sqrt{5.3867 \quad 0.1820} \times 10$$

$$= \sqrt{5.2047} \times 10$$

$$= 2.28 \times 10$$

$= 22.8$ अंक

विषमता गुणक या $j = \frac{3\left(\overline{X} \quad M\right)}{} = \frac{3\,(39.27 \quad 5)}{22.8}$

$= -0.75$ approx.

विषमता का द्वितीय माप : इस माप में मध्यका एवं चतुर्थकों का प्रयोग किया जाता है। इसके द्वारा दिया गया सूत्र इसके परिकलन के लिए उपयोग में आता है। यह सूत्र निम्न है :

विषमता $= (Q3 - M) - (M - Q1)$

$= Q3 + Q1 - 2M$ (निरपेक्ष)

विषमता गुणक या $j = \frac{(Q_3 \quad M)\,(M \quad Q_1)}{(Q_3 \quad M)\,(M \quad Q_1)} = \frac{Q_3 \quad Q_1 \quad 2M}{Q_3 \quad Q_1}$ सापेक्ष

उदाहरण 28 : निम्न श्रेणी के विषमता गुणक का द्वितीय माप ज्ञान कीजिए।

मासिक आय (रू. में)	**100-110**	**110-120**	**120-130**	**130-140**
श्रमिकों की संख्या	**3**	**9**	**15**	**24**
मासिक आय (रू. में)	**140-150**	**150-160**	**160-170**	**170-180**
श्रमिकों की संख्या	**12**	**8**	**7**	**5**

मासिक आय (रूपयों में)	श्रमिकों की संख्या	संचयी बारम्बारता
100-110	3	3
110-120	9	12
120-130	15	27
130-140	24	51
140-150	12	63
150-160	8	71
161-170	7	78
170-180	5	83
योग	**83**	

मध्य का क्रमांक $\frac{N}{2} = \frac{83}{2} = 41.5$ (130-140) वर्ग

प्रथम चतुर्थक क्रमांक $\frac{N}{4} = \frac{83}{4} = 20.76$ (120-130) वर्ग

$$= 41.5 \ (130\text{-}140) \text{ वर्ग}$$

प्रथम चतुर्थक क्रमांक $\frac{N}{4} = \frac{83}{4}$

$$= 20.76 \ (120\text{-}130) \text{ वर्ग}$$

ततीय चतुर्थक क्रमांक $\frac{3N}{4} = \frac{3 \quad 83}{4}$

$$= 62.25 \ (140\text{-}150) \text{ वर्ग}$$

$$Q_1 = I_1 +$$

$$= 120 + \frac{10}{15}(20.75 - 12)$$

$$= 120 + \frac{10 \quad 8.75}{15}$$

$$= 120 + 5.833 + 125.833 \text{ रू.}$$

$$Q_3 = \frac{i}{f}\left(\frac{3N}{4} \quad C_0\right)$$

$$= 140 + \frac{10}{12}(62.25 - 51)$$

$$= 140 + \frac{10 \quad 11.15}{12} \quad \frac{i}{f}\left(\frac{N}{4} \quad C_0\right)$$

$$= 140 + 9.375$$

$$= 149.375 \text{ रू.}$$

$$M = \frac{i}{f}\left(\frac{N}{2} \quad C_0\right)$$

$$= 130 + \frac{10}{24}(41.5 - 27)$$

$$= 130 + \frac{10 \quad 14.5}{24}$$

$$= 130 + 6.042$$

$$= 136.042 \text{ रू. लगभग}$$

$$j = \frac{Q_3 \quad Q_1 \quad 2M}{Q_3 \quad Q_1}$$

$$= \frac{149.375 \quad 125.833 \quad 2 \quad 136.042}{149.375 \quad 125.833}$$

$$= \frac{3.124}{23.542} = + 0.133 \text{ approx.}$$

यदि बंटन में समान्तर माध्य ($\overline{X}$) मध्यका (M) व बहुलक (Z) के मूल्य बराबर है तो विषमता नहीं होगी अर्थात बंटन सममित होगा। अन्यथा विषमता होगी। बंटन में यदि समान्तर माध्य का मूल्य सबसे अधिक है तथा बहुलक का मूल्य सबसे कम है तो धनात्मक विषमता होगी, यदि इसके विपरीत है अर्थात् बहुलक का मूल्य सबसे अधिक तथा समान्तर माध्य का मूल्य

सबसे कम है तो विषमता ऋणात्मक होगी। बंटन में माध्य व बहुलक में जितना अधिक का माप किया जा सकता है। इसके लिए हमको (b) गुणकों का प्रयोग करना पड़ता है। कार्ल पियर्सन के अनुसार परिघातों के द्वारा विषमता गुणांक निम्न सूत्र से निकाला जा सकता है।

$$j = b_1$$

$$b_1 = \frac{{}_{3}^{2}}{{}_{2}^{3}}$$

यदि b_1 का मूल्य शून्य (0) है तो अंक श्रेणी में सममितता (Symmetry) है अर्थात विषमता नहीं है। यदि b_1 का मूल्य शून्य नहीं है तो इस बात का द्योतक है कि समंकमाला में विषमता है। यदि m_3 का मान ऋणात्मक है तो विषमता ऋणात्मक होगी और यदि m_3 का मान धनात्मक है तो विषमता भी धनात्मक होगी।

प्रो. फिशर ने विषमता के अध्ययन के लिए ग्रीक अक्षर 'गामा' (g) का प्रयोग किया है, उन्होंने विषमता का निम्न सूत्र दिया है:

$$g_1 = \sqrt{{}_1}$$

उदाहरण 29 : एक बंटन के प्रथम चार परिघातों के मान 0, 2.5, 0.7 तथा 18.75 हैं। बंटन की विषमता तथा पथुशीर्षत्व ज्ञात कीजिए।

हल : दिया गया है :

$$m_1 = 0$$
$$m_2 = 2.5$$
$$m_3 = 0.7$$
$$m_4 = 18.75$$

$$\text{विषमता } b_1 = \frac{{}_{3}^{2}}{{}_{2}^{3}}$$

$$= \frac{(0.7)^2}{(2.5)^3}$$

$$= \frac{0.49}{15.63}$$

$$= 0.031$$

चूँकि b_1 का मूल्य +0.031 है, इसलिए बंटन में थोड़ी-सी धनात्मक विषमता है।

$$\text{पथुशीर्षत्व } b_1 = \frac{{}_{4}}{{}_{2}^{2}}$$

$$= \frac{18.75}{(2.5)^2}$$

$$= \frac{18.75}{6.25}$$

$$= 3$$

चूँकि b_2 का मान पूरा 3 है, इसलिए इस बंटन का शीर्ष सामान्य वक्र है।

उदाहरण 30 : निम्न आवत्ति से विषमता गुणक तथा पथुशीर्षत्व (परिघात पर आधारित) ज्ञात कीजिए :

x	*y*
0-9	1
10-19	5
20-29	12
30-39	22
40-49	17
50-59	9
60-69	4
70-79	3
80-89	1
90-99	1

हल : Calculation of Skewness and Kurtosis

	M.V. X	*f*	$\left(\frac{X \quad 44.5}{10}\right)$ *d'*	*fd'x*	*fd'x*^{2}	*fd'x*^{3}	*fx'x*^{4}
0-9	4.5	1	– 4	– 4	16	– 64	256
10-19	14.5	5	– 3	– 15	45	– 153	403
20-29	24.5	12	– 2	– 24	48	– 96	192
30-39	34.5	22	– 1	– 22	22	– 22	22
40-49	44.5	17	0	0	0	0	0
50-59	54.5	9	+ 1	9	9	9	9
60-69	64.5	4	+ 2	8	16	52	64
70-79	74.5	3	+ 3	9	27	81	243
80-89	84.5	1	+ 4	4	16	64	236
90-99	94.5	1	+ 5	5	25	123	625
		N = 75		**S*fd'x* = – 30**	**S*fd'x*^{2} = 224**	**S*fd'x*^{3} = – 6**	**S*fd'x*^{4} = 2,072**

$$n_1 = \frac{fd'x}{N} \times 10$$

$$= \frac{300}{75} = -4$$

$$n_2 = \frac{fd'^2x}{N} \times 10^2$$

$$= \frac{22400}{75}$$

$$= 229$$

$$n_3 = \frac{fd'^3 x}{N} \times 10^3$$

$$= \frac{6}{75} \times 10^3$$

$$= -80$$

$$n_4 = \frac{fd'^4 x}{N} \times 10^4$$

$$= \frac{2,072}{75} \times 10^4$$

$$= 273600$$

$$m_1 = 0$$

$$m_2 = n_2 - (n_1)^2$$

$$= 299 - 16$$

$$= 283$$

$$m_3 = n_3 - 3n_2n_1 + 2\,(n_1)^3$$

$$= -.80 - 3\,(-4)\,(299) + 2\,(-4)^3$$

$$= -.\,80 + 3588 - 128$$

$$= 3380$$

$$m_4 = n_4 - 4n_3n_1 + 6n_2\,(n_1)^2$$

$$= 276300 - 4\,(-80)\,(-4) + 6\,(299)\,(-4)^2 - 3\,(-4)^4$$

$$= 276300 - 1280 + 28704 - 768$$

$$= 303356$$

Skewness

$$b_1 = \frac{{}_3^2}{{}_2^3}$$

$$= \frac{(3380)^2}{(283)^3}$$

$$= \frac{114240}{22665187}$$

$$= 0.504$$

For Kurtosis we have to compute the value of b_2

$$b_2 = \frac{{}_4}{{}_2^2}$$

$$= \frac{303356}{(283)^2}$$

$$= 3.787$$

चूँकि b_1 का मान धनात्मक है अतः बंटन में थोड़ी विषमता है एवं b_2 का मूल्य 3 से अधिक है इसलिए ऊँचे शीर्ष वाला है।

उदाहरण 31 :

(*a*) The standard deviation of a symmetrical distribution is 3. What must be the fourth moment about the mean in order that the distribution be mesokurtic ?

(*b*) If the fourt moments of distribution about the value 5 are equal to –4, 22, – 117 and 560, determine the corresponding moments :

(*i*) about the mean, and (*ii*) about zero.

Solution : (*a*) For a mesokurtic distribution $b_2 = 3$

$$b_2 = \frac{4}{\frac{2}{2}}$$

We are given

$s = 3, m_2 = s^2 = (3)^2 = 9, b_2 = 3, m_2 = 9$

\

$$3 = \frac{4}{9^2} \text{ or } m_4$$
$$= 243$$

Thus the fourth moment about mean must be 243 in order that the distribution be mesokurtic.

(*b*) We are given moments about an arbitary origin 5.

Thus, $n_1 = -4, n_2 = 22, n_3 = 177, n_4 = 560.$

Moments about Mean :

From these we can find out the moments about mean from the following relationship :

$$m_1 = n_1 - n_1$$
$$= 0$$
$$m_2 = n_2 - (n_1)^2$$
$$m_3 = n_3 - 3\,n_2\,n_1 + 2\,(n_1)^3$$
$$m_4 = n_4 - 4\,n_1\,n_3 + 6\,n_2\,(n_1)^2 - 3\,(n_1)^4$$

Substituting the values,

$$m_2 = 22 - (-4)^2$$
$$= 22 - 16$$
$$= 6$$
$$m_3 = -177 - 3(-4)(22) + 2(-4)^3$$
$$= -117 + 264 - 128$$
$$= 19$$
$$m_4 = 560 - 4(-4)(-117) + 6(22)(-4)^2 - 3(-4)^4$$
$$= 560 - 1872 + 2112 - 768$$
$$= 32$$

Thus, the moments about mean are $m_1 = 0, m_2 = 6, m_3 = 19$ and $m_4 = 32$

Moments about Zero :

Let the moments about zero be denoted by $n_1', n_2', n_3', = 0$, etc.

The first moment about zero, *i.e.* $n_1' = A + n_1$

The second moment about zero, *i.e.* $n_2' = m_2 + (n_1)^2$

The third moment about zero, *i.e.* $n_3' = m_3 + 3m_2n_1 + (n_1)^2$

The fourth moment about zero, *i.e.* $n_4' = m_2 + 4m_2n_1 + 6m_2n_1^2 (n_1)^4$

Substituting these values.

$$n_1' = 5 + (-4) = 1$$

$$n_2' = 5 + (1)^2 = 7$$

$$n_3' = 19 + 3(1)(6) + (1)^3 = 19 + 18 + 1 = 38$$

$$n_4' = 31 + 4(1)(19) + 6(1)^3(6) + (1)^4 = 32 + 76 + 36 + 1 = 45$$

The moments about zero are :

$n_1' = 1$, $n_2' = 7$, $n_3' = 38$, $n_4' = 145$

उदाहरण 32 : Find out Kurtosis from data given below :-

Income Rs.	**:**	**20-40**	**40-60**	**60-80**	**80-100**
No. of workers	**:**	**2**	**4**	**3**	**1**

Solution :

Income Rs.	No. of Workers f	Mid Value X	(X-56) x	fx	x^2	fx^2	x^3	fx^3	x^4	fx^4
20-40	2	30	– 26	– 52	676	1352	– 17576	– 35152	456976	913952
40-60	4	50	– 6	– 24	36	144	– 216	– 864	1296	5184
60-80	3	70	+ 14	42	196	588	2744	8232	38416	115242
80-100	1	90	+ 34	34	1156	1156	39304	39304	1336336	1336336
Total	**10**		**0**		**3240**		**+ 11520**			**2370720**

$$m_1 = \frac{fx}{N} = \frac{0}{10} \quad 0$$

$$m_2 = \frac{fx^2}{N} = \frac{3240}{10} = 324$$

$$m_3 = \frac{fx^3}{N}$$

$$= \frac{11520}{10} = 1152$$

$$m_4 = \frac{fx^4}{N}$$

$$= \frac{2370720}{10} = 237072$$

$$\overline{X} = \frac{560}{10} = 56$$

$$b_2 = \frac{{}_4}{{}_2{}^{2}} = 2.258$$

चूँकि b_2 का मान 3 से छोटा है अतः बंटन चपटे शीर्ष वाला है।

शैपर्ड संशोधन

(Sheppard's Corrections for Grouping)

व्यक्तिगत (Individual) एवं खंडित श्रेणी में शैपर्ड की शुद्धि की आवश्यकता नहीं है, क्योंकि आवत्ति निश्चित मूल्य (Size) संबंधी होती है अतः प्रत्येक मद का मूल्य ठीक-ठीक माप जा सकता है। आपने अविच्छिन्न समंक श्रेणी संबंधी प्रस्तुत किये गये उदाहरण में देखा गया कि परिघात निकालते समय यह मान लिया है कि प्रत्येक वर्ग की आवत्ति उसके मध्य-मूल्य (Mid-value) पर केन्द्रित है। इसका अर्थ यह हुआ कि विभिन्न वर्गों में आने वाली सभी इकाइयों का मूल्य वर्ग सीमाओं (Class Limits) के माध्य-मूल्य के बराबर ही है। इस संबंध में वास्तविकता यह है कि विभिन्न वर्गों की आवत्तियाँ उनके केवल मध्य मूल्य से ही संबंधित नहीं है वरन् उन सभी मूल्यों से संबंधित हो सकती है जो वर्गों की उच्चतम एवं न्यूनतम सीमा के अंतर्गत आते हों।

उदाहरण के लिए यदि 20-30 प्राप्तांक-वर्ग की आवत्ति 30 है तो यह मान लिया जाता है कि इन सभी 30 विद्यार्थियों में से प्रत्येक को $\frac{20 \quad 30}{2} = 25$ अंक प्राप्त हुए हैं जबकि वास्तविकता यह हो सकती है कि कुछ विद्यार्थियों को 21, कुछ को 22 और कुछ को 23 आदि-आदि अंक प्राप्त हुए हों, अतः इस प्रकार की मान्यता (जिसके लिए बिना परिकलन कार्य असम्भव रहता है) के कारण परिकलित मापों में कुछ विभ्रम हो जाता है। ऐसे विभ्रम को दूर करने अथवा कम करने की दष्टि से ही शैपर्ड ने संबंधित मापों के सूत्रों में कुछ परिवर्तन अथवा समायोजन सुझाये हैं जिन्हें शैपर्ड के संशोधन के नाम से पुकारा जाता है।

शैपर्ड के द्वारा बतलाये गये संशोधन निम्नलिखित हैं :

1. m_1 कोई संशोधन नहीं।

2. m_2 के परिकलित मूल्य में से $\frac{i}{12}$ को घटा लेने पर संशोधित m_2 का मूल्य ज्ञात हो जाता है, अर्थात्

   संशोधित $m_2 = \left({}_2 \quad \frac{i^2}{12} \right)$

   जहाँ m_2 = समान्तर माध्य पर आधारित द्वितीय परिघात

   i = वर्ग विस्तार

3. m_3 कोई संशोधन नहीं।

4. m_4 के परिकलित मूल्य में $\frac{1}{2}\ {}_2 i^2 \quad \frac{7}{240} i^4$ का समायोजन करने में संशोधित m_4 का मूल्य ज्ञात हो जाता है अर्थात्

संशोधित $m_4 = \left({}_4 \quad \frac{1}{2}\ {}_2 i_1^2 \quad \frac{7}{240} i^4 \right)$

जहाँ m_4 = समानान्तर माध्य पर आधारित चतुर्थ परिघात

i = वर्ग विस्तार

प्रथम तथा ततीय परिघातों में कोई संशोधन करने की आवश्यकता नहीं है क्योंकि घनात्मक (positive) एवं ऋणात्मक (negative) चिन्ह बने रहते हैं जिससे विभ्रम पूरक हो जाता है, द्वितीय तथा चतुर्थ परिघातों में विचलनों के वर्ग तथा चतुर्थ घात हो जाने से विनम्र संचयी प्रकृति (cumulative type) का हो जाता है, क्योंकि सभी मान धनात्मक हो जाते हैं। अतः संशोधन करना आवश्यक है। उदाहरण 3 में जो हमने परिघात निकाले हैं, उनमें से द्वितीय एवं चतुर्थ परिघातों को इन उपर्युक्त सूत्रों की सहायता से शुद्ध कर सकते हैं। संशोधित परिघात निम्न प्रकार के होंगे :

1. m_2

$$
\begin{aligned}
\text{Corrected } m_2 &= m_2 - \frac{i^2}{12} \\
&= 100 - \frac{10^2}{12} \\
&= 100 - 8.3 \\
&= 91.7
\end{aligned}
$$

2. m_4

$$
\begin{aligned}
\text{Corrected } m_4 &= m_4 - \frac{1}{2}\ {}_2 \ i^2 \quad \frac{7}{240} \ i^4 \\
&= 22000 - \frac{1}{2} \times 100 \times 100 + \frac{7}{240} \times 10{,}000 \\
&= 22000 - 5000 + 291.7 \\
&= 22291.7 - 5000 \\
&= 17291.7
\end{aligned}
$$

शैपर्ड के संशोधन लागू करने संबंधी आवश्यक शर्तें :

1. आवत्ति -बंटन विभिन्न वर्गों के रूप में होना चाहिए।
2. सभी वर्गों के वर्ग-विस्तार समान होने चाहिए।
3. वर्ग-विस्तार प्रदत्त मूल्यों के विस्तार (Range) के लगभग $\frac{1}{12}$ से अधिक नहीं होना चाहिए।

शैपर्ड के संशोधन में यह मान्यता रहती है कि आवत्ति बंटन सममित या थोड़ा -सा ही असममित होता है। मदों की संख्या 1,000 तक होने पर ही शैपर्ड के संशोधन उपयुक्त समझे जाते हैं। यदि मदों की संख्या 1,000 से अधिक हो तो शैपर्ड संशोधन की आवश्यकता नहीं होती है।

उदाहरण 33 : आपको परिघातों का मूल्य निम्न प्रकार दिया गया है :

$m_2 = 43.353, m_3 = -9.774, m_4 = 5508.567$

यदि वर्ग विस्तार 3 हो तो संशोधित परिघातों का परिकलन कीजिए।

हल : 1. संशोधित $m_2 = \left[{}_2 \quad \frac{i^2}{12} \right]$

$$= 43.353 - \frac{(3)^2}{12} \text{ या}$$

$$= 43.353 - 0.75$$

$$= 42.603$$

2. संशोधित $m_3 = -9.774$ (संशोधन अनावश्यक)

3. $m_4 =$

$$= 5508.567 - \frac{1}{2} (3)2 \, (43.353) + \frac{7\,(3)^4}{240}$$

$$= 5508.567 - 195.0885 + 2.3625$$

$$= 5315.841$$

❑❑

$$\left[{}_4 \quad \frac{1}{2} \;\; {}_2 i^2 \quad \frac{7}{240} \quad i^4 \right]$$

अध्याय - 8

सहसम्बन्ध (Correlation)

विवरणात्मक सांख्यिकी में (Descriptive Statistics) में हम एक ही चर (Variable) के विश्लेषणात्मक अध्ययन जैसे सांख्यिकीय माध्य (Statistical Averages) व अपकिरण तथा विषमता के माप (Measures of dispersion & Skewness) का अध्ययन करते हैं। लेकिन अपनी रोजमर्रा की जिन्दगी में हमें दो या दो से अधिक चरों के बीच एक सम्बन्ध का अध्ययन करना पड़ता है, जो हमारे निर्णय करने में मदद करते हैं। सहसम्बन्ध सांख्यिकी की एक महत्त्वपूर्ण तकनीक है जो हमें ऐसे निर्णय लेने में मदद करती है। प्रस्तुत पाठ हमें दो चरों के बीच सहसम्बन्ध स्थापित करने की विधियाँ दर्शाता है जिससे कि हम उचित निर्णय ले सकें।

सहसम्बन्ध एक सांख्यिकी माप है जो दो समंक-समूहों के मध्य पाये जाने वाले सम्बन्ध का विवरण देती है। कुछ प्रकार के चरों (Variables) में पारस्परिक सम्बन्ध पाया जाता है, जैसे पूर्ति एवं मूल्य, माँग एवं वत्ति , उत्पादन तथा आयात आदि चरों में सम्बन्ध होता है। दिन-प्रति-दिन के अनुभव से यह स्पष्ट होता है कि किस प्रकार विभिन्न तथा पारस्परिक रूप से सम्बन्धित होते हैं। उदाहरणार्थ, पति-पत्नियों की आयुओं में, पिता-पुत्र की ऊँचाइयों में, युवा व्यक्तियों की ऊँचाई एवं भार में, मकान के आकार एवं उसकी व्यवस्था करने की लागत में, विनियोजित पूंजी एवं अर्जित लाभ में तथा अन्य ऐसे ही तथ्यों में निकट का सम्बन्ध पाया जाता है। सहसम्बन्ध एक सांख्यिकीय तकनीक है जो यह मापती एवं विश्लेषण करती है कि दो चर अथवा तथ्य एक दूसरे के सन्दर्भ में किस सीमा तक परिवर्तित होते हैं। यह दो चरों अथवा तथ्यों के मध्य सम्बन्ध का अध्ययन करती है। यह माप यह बतलाती है कि चरों के सम्बन्ध की मात्रा क्या है। सहसम्बन्ध दो चरों के मध्य अन्तरनिर्भरता (Interdependence) इंगित करता है।

सहसम्बन्ध की परिभाषा

जब दो तथ्यों में एक साथ एक ही दिशा में अथवा विपरीत दिशाओं में परिवर्तन हो तथा एक तथ्य में परिवर्तन दूसरे तथ्य में परिवर्तन का कारण हो तो यह कहा जाता है कि उन दोनों तथ्यों में सहसम्बन्ध है। किंग में शब्दों में, सहसम्बन्ध का यह अर्थ है कि दो समंकमालाओं अथवा तथ्य समूहों के कारण व परिणाम का सम्बन्ध पाया जाता है। एक अन्य स्थान पर उन्होंने मत व्यक्त किया है कि, ''यदि यह सत्य प्रमाणित हो जाता है कि अधिकांश उदाहरणों में दो चर सदैव एक ही दिशा में या विपरीत दिशाओं में घटने-बढ़ने की प्रवत्ति रखते हैं तो हम यह मानते हैं कि उनमें एक सम्बन्ध पाया जाता है। यह सम्बन्ध ही सहसम्बन्ध कहलाता हैं कौनर के अनुसार, ''यदि दो या अधिक राशियाँ सहानुभूति में परिवर्तित हों जिससे एक में होने वाले परिवर्तनों के फलस्वरूप दूसरी राशि में भी परिवर्तन होने की प्रवत्ति पायी जाए , तो वे राशियाँ सहसम्बन्धित कहलाती है। प्रो. बोडिंगटन के शब्दों में, ''जब कभी दो या अधिक समूहों अथवा वर्गों अथवा समंकमालाओं में निश्चित सम्बन्ध विद्यमान हो, तो उसमें सहसम्बन्ध का होना कहा जाता है।

डेयरपोर्ट के मतानुसार, ''सहसम्बन्ध का सम्पूर्ण विषय पथक विशेषताओं के मध्य पाये जाने वाले उस पारस्परिक सम्बन्ध की ओर संकेत करता है जिसके अनुसार वे कुछ सीमा तक साथ-साथ परिवर्तन होने की प्रवत्ति रखती है।'' तारों यामने के अनुसार, ''सहसम्बन्ध विश्लेषण दो चरों के मध्य सम्बन्ध की घनिष्टता की मात्रा का विवेचन है। 'ग्रोहमेन के अनुसार, ''जब दो या अधिक संख्याएँ सहानुभूति में परिवर्तित होती हैं, जिससे एक में परिवर्तन का परिणाम दूसरी में उसी अथवा विपरीत दिशा में परिवर्तन हो, तो उन संख्याओं के मध्य सहसम्बन्ध होता है।'' सहसम्बन्ध दो चरों के सम्बन्ध के परिणाम को व्यक्त करता है। यह एक सांख्यिकीय तकनीक है जो दो सम्बन्धित चरों के मध्य सम्बन्ध की प्रकृति व मात्रा मापती है।

सांख्यिकी में सहसम्बन्ध सिद्धान्त एवं तकनीक का बहुत महत्त्व है। इस सिद्धान्त के मूल तत्वों का प्रतिपादन सर्वप्रथम फ्रांस के खगोल शास्त्री ब्रावेस (Bravais) ने किया था, परन्तु बिन्दु रेखीय रूप में सहसम्बन्ध-तकनीक का अन्वेषण सर्वप्रथम सर फ्रांसिस गाल्टन (Sir Francis Galton) ने किया था। 1896 में, प्रसिद्ध संख्या शास्त्री कार्ल पियर्सन (Karl Pearson) ने

सहसम्बन्ध गुणक (Coeffecient of Correlation) द्वारा सहसम्बन्ध ज्ञात करने की गणितीय विधि का प्रतिपादन किया। इन दोनों संख्या शास्त्रियों (गाल्टर तथा पियर्सन) ने इस तकनीक की सहायता से प्राणिशास्त्र (Biology) तथा जनन-विद्या (Genetics) की अनेक समस्याओं का विवेचन किया। अर्थशास्त्र में भी इस तकनीक का विशेष महत्त्व है। अर्थशास्त्र में सहसम्बन्ध के उपयोग के बारे में नीसवेंजर लिखते हैं, ''सहसम्बन्ध-विशलेषण आर्थिक व्यवहार को समझने में योग देता है, विशेष महत्त्वपूर्ण चरों, जिन पर अन्य चर निर्भर करते हैं, को खोजने में सहायता देता है; अर्थशास्त्री उन सम्बन्धों को स्पष्ट करता है जिनमें गड़बड़ी फैलती है तथा उसे उन उपायों का सुझाव देता है जिनके द्वारा स्थिरता लाने वाली शक्तियाँ प्रभावी हो सकती है।'' प्रतीपगमन (regression) तथा विचरण-अनुपात (ratio of variation) के विचार सहसम्बन्ध की माप पर ही आधारित है। सहसम्बध की माप यह भी आश्वस्त करती है कि सम्बन्धित चरों में आन्तरगणन अथवा बाह्यगणन एवं पूर्वानुमान विश्वसनीय होगा। टिपेट का कथन है कि, ''सहसम्बन्ध का प्रभाव हमारी भविष्यवाणी की अनिश्चितता के विस्तार को कम करना है।'' सहसम्बन्ध-विश्लेषण पर आधारित पूर्वानुमान अधिक विश्वसनीय एवं वास्तविकता के निकट होते हैं।

सहसम्बन्ध के प्रकार

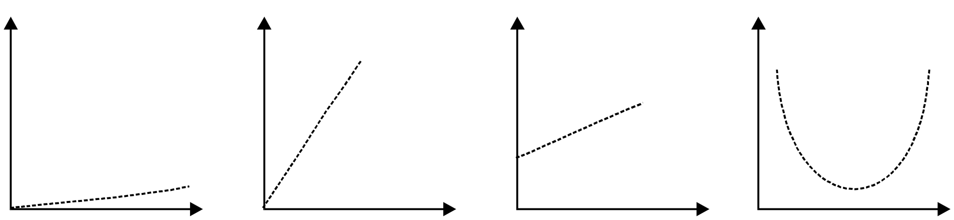
सम्बद्ध चरों के मध्य परिवर्तनों की दिशा, अनुपात आदि के आधार पर सहसम्बन्ध निम्नलिखित प्रकारों का हो सकता है :

(1) **धनात्मक अथवा ऋणात्मक सहसम्बन्ध** (Positive or Negative Correlation) : दो चरों में यदि एक ही दिशा अथवा विपरीत दिशाओं में परिवर्तन होते हों तो उनके मध्य सहसम्बन्ध होता है। यदि एक चर-मूल्य घटने पर दूसरा चर-मूल्य भी घटे अथवा एक चर-मूल्य के बढ़ने पर दूसरा चर-मूल्य भी बढ़े तो ऐसा सहसम्बन्ध धनात्मक (Positive) होता है। मूल्य एवं पूर्ति में इस प्रकार का सम्बन्ध पाया जाता है। यदि किसी वस्तु का मूल्य बढ़ता है तो उसकी पूर्ति (Supply) भी बढ़ जाती है और वस्तु का मूल्य घटने पर उनकी पूर्ति भी घट जाती हैं।

ऋणात्मक सहसम्बन्ध उस दशा में होता है जब एक चर-मूल्य के घटने पर दूसरा चर-मूल्य बढ़ता हो तथा चर-मूल्य के बढ़ने पर दूसरे चर-मूल्य में कमी होती हो। इस प्रकार के सहसम्बन्ध को विलोम (Inverse) सहसम्बन्ध भी कहते हैं। इस प्रकार का सहसम्बन्ध मूल्य एवं माँग में पाया जाता है। किसी वस्तु का मूल्य बढ़ने पर उसकी माँग (Demand) कम हो जाती है और मूल्य कम हो जाने पर माँग बढ़ जाती है।

(2) **सरल, आंशिक अथवा बहुगुणी सहसम्बन्ध** (Simple, Partial or Multiple Correlation) : स्वतंत्र एवं आश्रित चर मूल्यों (Independent and Dependent Variables) की संख्या के आधार पर सहसम्बन्ध सरल, आंशिक अथवा बहुगुणी प्रकार का हो सकता हैं दो चर-मूल्यों के सहसम्बन्ध को सरल सहसम्बन्ध (Simple Correlation) कहते हैं। इनमें से एक श्रेणी, जिसे आधार श्रेणी (Subject Series) कहते हैं, के चर-मूल्य स्वतंत्र (Independent Variables) होते हैं तथा दूसरी श्रेणी, जिसे सम्बद्ध श्रेणी (Relative Series) कहते हैं, के चर-मूल्य आश्रित (Dependent Variables) होते हैं। आंशिक सहसम्बन्ध (Partial Correlation) में दो मूल्यों में एक अन्य स्वतंत्र चरमूल्य का समावेश करके, सहसम्बन्ध ज्ञात किया जाता है। बहुगुणी सहसम्बन्ध (Multiple Correlation) में तीन या अधिक चर-मूल्यों कें मध्य सहसम्बन्ध अध्ययन किया जाता है।

(3) **रेखीय अथवा अ-रेखीय सहसम्बन्ध** (Linear and Non-Linear Correlation) : रेखीय अथवा अ-रेखीय सहसम्बन्ध के मध्य अंतर का आधार विचारगत चर-मूल्यों के मध्य परिवर्तन-अनुपात की नियमितता होती है। यदि दो चर-मूल्यों के मध्य परिवर्तन का अनुपात समान होता है तो उनमें रेखीय सहसम्बन्ध होगा। इन चर-मूल्यों को यदि बिन्दु-रेखीय पत्र पर अंकित किया जाए तो बिन्दु एक सीधी रेखा के रूप में होंगे। अ-रेखीय सहसम्बन्ध, जिसे वक्र-रेखीय सहसम्बन्ध (Curvilinear Correlation) भी कहते हैं, में एक चर-मूल्य के परिवर्तनों की मात्रा व दूसरे चर-मूल्य के परिवर्तनों की मात्रा एक अनुपात में नहीं होगी। इन चर-मूल्यों को बिन्दु रेखा पर अंकित करने पर वक्र बन जाती है। निम्न चित्र इन दोनों प्रकार के सहसम्बन्ध के वक्र स्पष्ट करता है।



Linear Correlation **Non-linear Correlation**

सहसम्बन्ध का परिणाम
(Degree of Correlation)

जब दो सम्बन्धित चरों में बिल्कुल आनुपातिक परिवर्तन होते हैं, तो उनमें पूर्ण सहसम्बन्ध (Perfect Correlation) होता है। यदि परिवर्तन आनुपातिक नहीं होते तो उन चरों में सीमित सहसम्बन्ध (Limited Correlation) होता है। यदि समान आनुपातिक परिवर्तन एक ही दिशा में होते हैं तब दोनों चरों के मध्य पूर्ण धनात्मक सहसम्बन्ध (Perfect Positive Correlation) होता है। इसके विपतरीत यदि समान आनुपातिक परिवर्तन विपरीत दिशाओं में होते हैं तो उन दोनों चरों के मध्य पूर्ण ऋणात्मक सहसम्बन्ध (Perfect Negative Correlation) होता है। यदि दो चरों के मध्य असमान परिवर्तन एक ही दिशा में होते हैं तो उनमें सीमित धनात्मक (Limited Positive) तथा असमान परिवर्तन विपरीत दिशाओं में होने पर सीमित ऋणात्मक (Limited Negative) सहसम्बन्ध होता है।

कार्ल पियर्सन (Karl Pearson) ने सहसम्बन्ध की माप का एक सूत्र दिया है। उस सूत्र का परिणाम ± के बीच में आता है। पूर्ण धनात्मक सहसम्बन्ध (Perfect Positive Correlation) होने की दशा में परिणाम + 1 तथा पूर्ण ऋणात्मक सहसम्बन्ध (Perfect Negative Correlation) होने की दशा में परिणाम – 1 प्राप्त होता है। यदि परिणाम 'शून्य' आता है। तो उन दोनों चरों में सहसम्बन्ध की अनुपस्थिति (Absence of Correlation) होती है। परिणाम की अन्य मात्राओं को उच्च (High), मध्यम (Moderate) तथा निम्न (Low) स्तरों में इस प्रकार निश्चित करते हैं :

सहसम्बन्ध-परिणाम के निर्वचन की तालिका

सहसम्बन्ध का परिणाम	धनात्मक (Positive)	ऋणात्मक (Negative)
सहसम्बन्ध की अनुपस्थिति	0	0
पूर्ण सहसम्बन्ध	+ 1	– 1
उच्च-परिणाम (High Degree)	+ .75 से लेकर + 1 तक	– .75 से लेकर – 1 तक
मध्यम-परिणाम (Moderate Degree)	+ .25 से लेकर + .75 तक	– .25 से लेकर – .75 तक
निम्न-परिणाम (Low Degree	0 से लेकर + .25 तक	0 से लेकर – .25 तक

सहसम्बन्ध और कारण-कार्य सम्बन्ध
(Correlation and Causation)

सह-सम्बन्ध विश्लेषण में यदि दो चरों के बीच सह-सम्बन्ध स्थापित हो जाता है तो यह आवश्यक नहीं कि उनके बीच कारण-कार्य सम्बन्ध हो। कुछ दशाओं में हम चरों को स्वतन्त्र (Independent or Cause) व आश्रित (Dependent or Effect) चरों के रूप में पहचान सकते हैं जैसे आयु व लम्बाई, आय और बचत, बिक्री व लाभ आदि। लेकिन अनेक दशाओं में दोनों चरों में कारण-कार्य सम्बन्ध होना आवश्यक नहीं है। ऐसी तीन स्थितियों में यह सह-सम्बन्ध पाया जाता है जो निम्नलिखित हैं :

(1) **संयोगवश सह-सम्बन्ध** (Chance Correlation) : कई बार दो चरों में सह-सम्बन्ध पाया जाता है जबकि वास्तविकता में इनमें कोई सह-सम्बन्ध नहीं होता। उदाहरण के लिए, यदि स्टील के उत्पादन व एल्युमिनियम की बिक्री में उच्च स्तरीय सह-सम्बन्ध पाया जाए तो यह भ्रामक होगा।

(2) **दोनों चर किसी तीसरे चर से प्रभावित हों** (Both the Variables may be Influenced by a Third Variable) : दो चरों में उच्चस्तरीय ऋणात्मक या धनात्मक सहसम्बन्ध इसलिए भी हो सकता है कि दोनों ही चर किसी तीसरे चर से प्रभावित हों। जैसे ब्याज दर तथा निवेश के बीच उच्चस्तरीय ऋणात्मक सह-सम्बन्ध सरकारी उधार नीति के कारण हो सकता है।

(3) दोनों चर एक दूसरे को प्रभावित कर रहे हों (Both the Variables Affect Each Other) : कभी-कभी दो चरों को स्वतंत्र व आश्रित चरों के रूप में पहचान करने में कठिनाई आती है। जैसे किसी भी वस्तु के मूल्य व मांग के बीच सहसम्बन्ध होते हुए भी यह कहना कठिन है कि कौन-सा चर स्वतंत्र है व कौन-सा आश्रित।

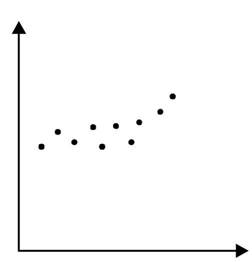
सहसम्बन्ध ज्ञात करने की रीतियां
(Methods of Determining Correlation)

सहसम्बन्ध ज्ञात करने की विभिन्न रीतियाँ निम्नलिखित हैं :

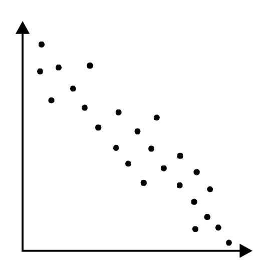
1. विक्षेप चित्र या बिन्दु चित्र (Scatter Diagram or Dotogram or Scattergram)
2. सहसम्बन्ध बिन्दु रेखा (Correlation Graph)
3. कार्ल पियर्सन का सहसम्बन्ध गुणक (Karl Pearson's Coefficient of Correlation)
4. स्पियरमैन का अनुपस्थिति सहसम्बन्ध गुणक (Spearman's Rank Coeffcient of Correlation)
5. संगामी विचलन गुणक (Coefficient of Concurrent Deviations)
6. न्यूनतम वर्ग रीति (Least Square Method)

1. विक्षेप-चित्र अथवा बिन्दु चित्र
(Scatter Diagram or Dotogram or Scattegram)

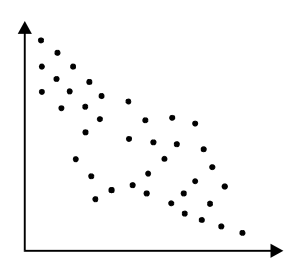
दो चरों के मूल्यों को विक्षेप चित्र पर अंकित करके उनके मध्य विद्यमान सहसम्बन्ध का अनुमान लगाया जा सकता है। विक्षेप चित्र बनाने के लिए स्वतंत्र चर मूल्यों (x) को बिन्दुरेखीय पत्र (Graph Paper) के भुजाक्ष (x-axis) पर तथा तत्सम्बन्धित आश्रित चर मूल्यों (y) को कोटि अक्ष (y-axis) पर प्रांकित किया जाता है। x श्रेणी तथा y श्रेणी के सम्बन्धित दो मूल्यों के लिए एक बिन्दु प्रांकित किया जाता है। श्रेणी में जितने पद-युग्म (Prais of Values) होते हैं उतने ही बिन्दु बिन्दुरेखीय पत्र पर प्रांकित हो जाते हैं। इस प्रकार समस्त समंकसमूह बिन्दुओं के रूप में परिवर्तित हो जाता है। बिन्दुरेखीय पत्र पर प्रांकित बिन्दुओं की प्रवत्ति के आधार पर दोनों चरों के मध्य के सम्बन्ध का अनुमान लगाया जाता है। यदि बिन्दुओं की प्रवत्ति एक निश्चित दिशा में जाने वाले प्रवाह या धारा की भाँति है तो दोनों चर-मूल्यों में सहसम्बन्ध होगा। इस प्रवाह में विभिन्न बिन्दु एक दूसरे के जितने निकट होंगे सहसम्बन्ध की मात्रा उतनी ही अधिक होगी। यदि बिन्दुओं का प्रवाह बायें कोने से दायीं तरफ ऊपर की ओर होता है तो दोनों चरों में धनात्मक सहसम्बन्ध होता है। बिन्दुओं का प्रवाह बायीं ओर से दायीं तरफ नीचे की ओर होने पर, दोनों चरों के मध्य ऋणात्मक सहसम्बन्ध होता है। यह निम्न विक्षेप चित्रों से स्पष्ट हो जाएगा :



(a) उच्च धनात्मक सहसम्बन्ध
(High Positive Correlation)

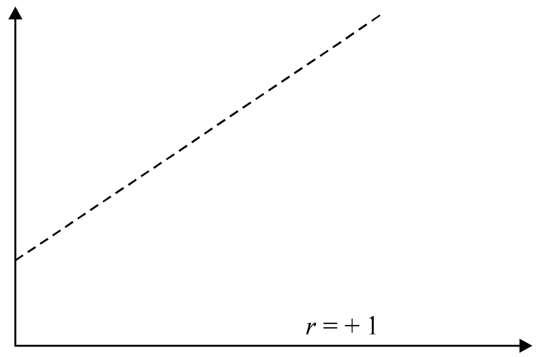


(b) ऋणात्मक सहसम्बन्ध
(Negative Correlation)

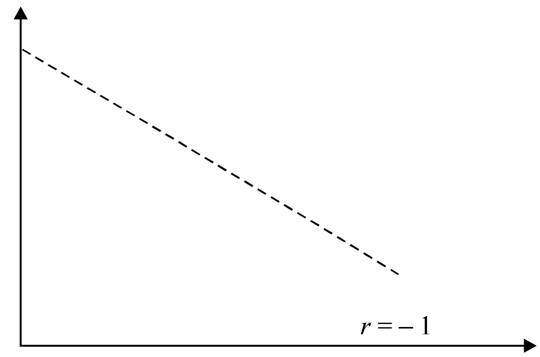


(c) सहसम्बन्ध का अभाव
(No Correlation)

यदि प्रांकित बिन्दु बायीं ओर के निचले कोने से दाहिनी ओर के ऊपर वाले कोने तक एक सीधी रेखा में हों तो दोनों चरों में पूर्ण धनात्मक सहसम्बन्ध होता है। इसके विपरीत, यदि प्रांकित बिन्दु ऊपर से नीचे की ओर एक सीधी रेखा में हों तो उन चर-मूल्यों में ऋणात्मक सहसम्बन्ध होता है। यह नीचे दिये गये चित्रों से स्पष्ट होता है :



(a) पूर्ण धनात्मक सहसम्बन्ध
(Perfect High Positive Correlation)



(b) पूर्ण ऋणात्मक सहसम्बन्ध
(Perfect Negative Correlation)

विक्षेप-चित्रांकन दो चरों में सहसम्बन्ध की प्रकृति ज्ञात करने की सरल व आकर्षक विधि है। इन चित्रों से एक ही दष्टि में यह ज्ञात हो जाता है कि चर-मूल्यों में सहसम्बन्ध है या नहीं और यदि है तो धनात्मक है अथवा ऋणात्मक। विक्षेप चित्रों का एक बड़ा दोष यह है कि उनसे सहसम्बन्ध की मात्रा ज्ञात नहीं की जा सकती।

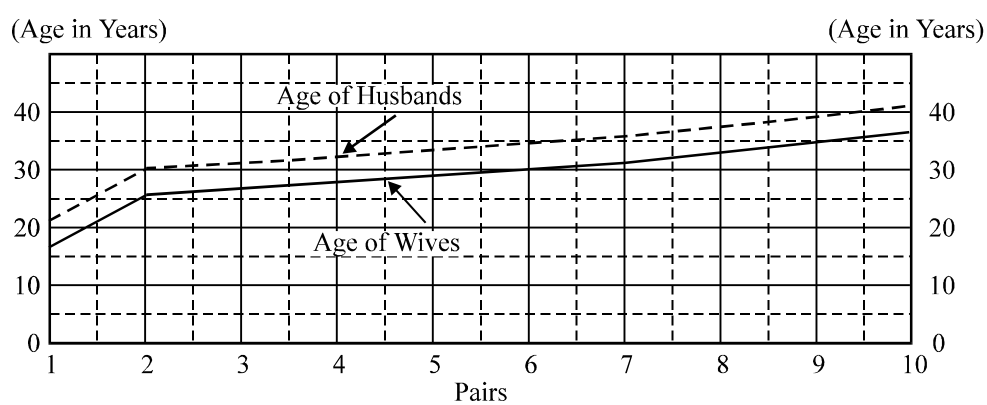
2. सहसम्बन्ध बिन्दुरेखा
(Correlation Graph)

बिन्दुरेखा द्वारा भी दो चर-मूल्यों के मध्य सहसम्बन्ध का अनुमान लगाया जा सकता है। दोनों श्रेणियों को बिन्दुरेखीय पत्र पर प्रांकित करके दोनों वक्रों (Curves) की प्रवत्ति (Direction) तथा निकटता (Closeness) के आधार पर उनके मध्य विद्यमान सहसम्बन्ध का अनुमान लगाया जा सकता है। यदि दोनों वक्र समानान्तर हों तो उनमें धनात्मक सहसम्बन्ध होता है। यदि दोनों वक्र विपरीत दिशाओं में हों तो उनमें ऋणात्मक सहसम्बन्ध होता है। दोनों वक्रों के उच्चावन की प्रवत्ति जितनी समान होगी, सहसम्बन्ध की मात्रा उतनी ही अधिक होगी। यदि दोनों वक्रों में एक ही दिशा या विपरीत दिशाओं में परिवर्तित होने की कोई प्रवत्ति दष्टिगोचर नहीं होती , तो दोनों में सहसम्बन्ध का अभाव समझना चाहिए।

Illustration 1 : Find out graphically, if there is any correlation between ages of husbands and wives, given below :

No. of Pairs	1	2	3	4	5	6	7	8	9	10
Husband's age	23	27	28	29	30	31	33	35	36	39
Wife's age	18	22	23	24	25	26	28	29	30	32

Solution.



बिन्दुरेखाओं से स्पष्ट होता है कि पतियों तथा पत्नियों की उम्रों में उच्च परिमाणीय धनात्मक सहसम्बन्ध (High Degree of Positive Correlation) है।

3. कार्ल पियर्सन का सहसम्बन्ध गुणक
(Karl Pearson's Coefficient of Correlation)

प्रसिद्ध संख्याशास्त्री कार्ल पियर्सन ने सहसम्बन्ध ज्ञात करने के लिए एक सूत्र का प्रतिपादन किया है। यह सूत्र अंकगणित माध्य (Arithmetic Mean) तथा प्रमाप विचलन (Standard Deviation) पर आधारित है जो क्रमशः केन्द्रीय प्रवत्ति तथा अपकिरण की आदर्श मापें हैं। सहसम्बन्ध की प्रकृति (धनात्मक है अथवा ऋणात्मक) तथा परिणाम की सन्तोषजनक अंकात्मक माप ज्ञात हो जाती है। कार्ल पियर्सन द्वारा प्रतिपादित सूत्र से 'सहसम्बन्ध गुणक' (Coefficient of Correlation) ज्ञात किया जाता है तो ± 1 $(-1 \leq r \leq +1)$ के अंतर्गत ही होता है। सहसम्बन्ध ज्ञात करने की यह सर्वोत्तम रीति है क्योंकि इससे सहसम्बन्ध गुणक $+1$ होने पर दोनों श्रेणियों में पूर्ण धनात्मक तथा -1 होने पर पूर्ण ऋणात्मक सहसम्बन्ध पाया जाता है। यदि गुणक शून्य (0) हो तो दोनों श्रेणियों में सहसम्बन्ध का अभाव होता है। जैसे-जैसे गुणक की माप 0 से 1 की ओर बढ़ती है, सहसम्बन्ध का परिणाम भी बढ़ता जाता है।

कार्ल पियर्सन का सहसम्बन्ध सूत्र निम्नलिखित मान्यताओं पर आधारित है :

(1) दोनों श्रेणियों, जिनके मध्य सहसम्बन्ध ज्ञात करना है, अनेक स्वतंत्र कारणों से प्रभावित होती है, जो उन श्रेणियों में सामान्य वितरण लाते हैं।

(2) दोनों श्रेणियों के पद-मूल्यों को प्रभावित करने वाली शक्तियाँ एक दूसरे से 'कारण तथा परिणाम' के रूप में सम्बन्धित होती है।

(3) दोनों श्रेणियों के मध्य रेखीय सम्बन्ध (Linear Relationship) होता है। इसका अर्थ यह है कि यदि श्रेणियों के पद-मूल्यों को विक्षेप चित्र पर प्रांकित किया जाए तो प्रांकित बिन्दुओं से एक रेखा बन जाएगी।

कार्ल पियर्सन का सहसम्बन्ध गुणांक सहविचरण (Co-variance) की माप से ज्ञात किया जाता है।

$$r = \frac{\text{covariance } (x, y)}{x \cdot \ y}$$

$$\text{Covariance} = \frac{(x \quad \bar{x})(y \quad \bar{y})}{\text{N}} \quad \frac{xy}{\text{N}}$$

$$s_x = \sqrt{\frac{x^2}{\text{N}}},\ s_y = \sqrt{\frac{y^2}{\text{N}}}$$

इस तरह से

$$r = \frac{xy}{\text{N}\sqrt{\frac{x^2}{\text{N}}}\cdot\sqrt{\frac{y^2}{\text{N}}}} = \frac{xy}{\sqrt{\ x^2}\cdot\sqrt{\ y^2}}$$

x व y के बीच कार्ल पियर्सन का सहसम्बन्ध निकालने का प्रथम सूत्र है। इस सूत्र में

Sxy = दोनों श्रेणियों के माध्यों से निकाले गये संबंधित विचलनों के गुणनफलों का योग

(Total of the products of deviations of x and y series from theri respective deviations)

N = पदों की सख्या (Number of items)

s_x = x श्रेणी का प्रमाप विचलन (Standard deviation of x series)

s_y = y श्रेणी का प्रमाप विचलन (Standard deviations of y series)

x = x श्रेणी का चर-मूल्य (Values in the x series)

$\bar{x}$ = x श्रेणी का अंकगणितीय माध्य (Arithmetic average of x series)

y = y श्रेणी के चर मूल्य (Values in the y series)

= y श्रेणी का अंकगणितीय माध्य (Arithmetic average of y series)

यह सूत्र केवल उन्हीं स्थिति में प्रयुक्त किया जाता है जब x तथा y दोनों के अंकगणितीय माध्यम पूर्णांक हों।

Illustration 1 : Find Karl Pearson's co-efficient of correlation between x and y series

X	**:**	90	104	93	91	104	107	109	113	128	131
Y	**:**	62	61	56	64	67	65	68	76	83	88

Solution.

X	x (X-30)	x^2	Y	y Y-69	y^2	xy
90	– 17	289	62	– 7	49	119
104	– 3	9	61	– 8	64	24
93	– 14	196	56	– 13	169	182
91	– 16	256	64	– 5	25	80
104	– 3	9	67	– 2	4	6
107	0	0	65	– 4	16	0
109	2	4	68	– 1	1	2
113	6	36	76	7	49	42
128	21	441	83	14	196	294
131	24	576	88	19	365	456
SX = 1070	**Sx = 0**	**Sx^2 = 1816**	**SY = 690**	**Sy = 0**	**Sy^2 = 934**	**Sxy = 1201**

इस सूत्र में

$$\overline{X} = \frac{X}{N} = \frac{1070}{10} = 107, \quad \overline{Y} = \frac{Y}{N} = \frac{690}{10} = 69$$

$$r = \frac{xy}{\sqrt{x^2, \quad y^2}} \quad \frac{1201}{\sqrt{1816 \quad 934}} = 0.8281$$

इस प्रकार हमें x व y के बीच उच्चस्तरीय घनात्मक सह-सम्बन्ध मिलता है।

द्वितीय सूत्र : यदि x तथा y में से एक या दोनों चरों के अंकगणितीय माध्य पूर्णांक नहीं हैं तो प्रथम सूत्र का उपयोग नहीं किया जाता है क्योंकि इससे विचलन ज्ञात करने व उनका वर्ग तथा गुणनफल ज्ञात करने में कठिनाई आती है। इसलिए लघु विधि (short-cut method) उपयोग में लाई जाती है। इसमें 'r' निकलने के सूत्र इस प्रकार है :

(*i*)
$$r = \frac{N \quad dx\,dy \qquad dx\,dy}{\sqrt{N \quad dx^2 \quad (\quad dx)^2} \,.\, \sqrt{N \quad dy^2 \quad (\quad dy)^2}}$$

इस सूत्र में

S$dxdy$ = कल्पित माध्यों से दोनों श्रेणियों के विचलनों के गुणनफलों का योग (sum of the products of deviations of two series from assumed means)

Sdx तथा Sdy = क्रमश: x तथा y श्रेणियों के विचलनों का योग (sum of deviations from assumed means)

Sdx^2 तथा Sdy^2 = क्रमश: x तथा y श्रेणियों के विचलनों के वर्गों का योग (sum of squares of deviations from assumed mean)

A_x = x श्रेणी का कल्पित माध्य (assumed mean of x series)

A_y = y श्रेणी का कल्पित माध्य (assumed mean of y series)

Illustration 2 : Calculate the coefficient of correlation for the following ages of husbands and wives :

Husband's age	**:**	**23**	**27**	**28**	**29**	**30**	**31**	**33**	**35**	**36**	**39**
Wife's age	**:**	**18**	**22**	**23**	**24**	**25**	**26**	**28**	**29**	**30**	**32**

Solution :

Husband's age x	**Deviation from assumed Mean = 30** dx	**Deviations Squared** dx^2	**Wife's age** y	**Deviations from assumed Mean = 25** dy	**Deviations squared** dy^2	**Product of deviations** $dxdy$
23	– 7	49	18	– 7	49	+ 49
27	– 3	9	22	– 3	9	+ 9
28	– 2	4	23	– 2	4	+ 4
39	– 1	1	24	– 1	1	+ 1
30	0	0	25	0	0	0
31	+ 1	1	26	+ 1	1	+ 1
33	+ 3	9	28	+ 3	9	+ 9
35	+ 5	25	29	+ 4	16	+ 20
36	+ 6	36	30	+ 5	25	+ 30
39	+ 9	81	32	+ 7	49	+ 63
Total N = 10	**11**	**215**		**+ 7**	**163**	**186**

$$r = \frac{N \quad dx\,dy \qquad dx\,dy}{\sqrt{N \quad dx^2 \quad (\quad dx)^2}.\sqrt{N \quad dy^2 \quad (\quad dy)^2}}$$

$$= \frac{10 \quad 186 \quad 11 \quad 7}{\sqrt{10 \quad 215 \quad (11)^2}.\sqrt{10 \quad 163 \quad (7)^2}}$$

$$= \frac{1860 \quad 77}{\sqrt{2150 \quad 121}\sqrt{1630 \quad 49}}$$

$$= \frac{1783}{\sqrt{2029}.\sqrt{1581}}$$

$$= \frac{1783}{\sqrt{3207849}}$$

$$= \frac{1783}{1791} + 0.99$$

$$r = \frac{dxdy \quad N\left(\bar{x} \quad A_x\right)\left(\bar{y} \quad A_y\right)}{N. \quad _x \cdot \quad _y}$$

Illustration 3.

Given	X	Y
Arithmatic Mean	74.5	125.5
Assumed Mean	69	112
Standard Deviation	13.07	15.85

No. of pairs of *x* and *y* = 8

Summation of products of corresponding deviation of X and Y series = 2176

Calculate the co-efficeint of correlation.

Solution.

$$N = 8,\quad \overline{X} = 74.5,\quad \overline{Y} = 125.5,\quad \Sigma dxdy = 2176$$

$$\sigma_x = 13.07,\ \sigma_y = 15.85,\quad A_x = 69,\quad A_y = 112$$

$$r = \frac{2176 \quad 8(74.5 \quad 69)(1255 \quad 112)}{10 \quad 13.07 \quad 15.85}$$

$$= \frac{2176 \quad 594}{1657.276}$$

$$= \frac{1582}{1657.276}$$

$$= +0.96$$

ततीय सूत्र -प्रत्यक्ष विधि (Direct Method)

इस विधि में ना तो अंकगणितीय माध्य से और ना ही काल्पनिक माध्य से विचलन निकाले जाते हैं। दी हुई मदों से प्रत्यक्ष रूप से सह-सम्बन्ध गुणांक निकाला जाता है। इसका सूत्र इस प्रकार है :

$$r = \frac{N \quad xy \qquad x \quad y}{\sqrt{N \quad x^2 \quad (\quad x)^2}\sqrt{N \quad y^2 \quad (\quad y)^2}}$$

Illustration 4. Calculate Karl Pearson's co-efficient of correlation for the following X and Y series :

X	**:**	12	6	18	13	25	22	8
Y	**:**	32	20	40	33	42	48	22

Solution :

X	X^2	Y	Y^2	XY
12	144	32	1024	384
6	36	20	400	120
18	324	40	1600	720
13	169	33	1089	429
25	625	42	1764	1050
22	484	48	2304	1056
8	64	22	484	176
104	**1846**	**237**	**8665**	**3935**

$$r = \frac{N \quad XY \qquad X \quad Y}{\sqrt{N \quad X^2 \quad (\quad X)^2}\,.\sqrt{N \quad Y^2 \quad (\quad Y)^2}}$$

$$= \frac{7 \quad 3935 \quad 104 \quad 237}{\sqrt{7 \quad 1846 \quad (104)^2}.\sqrt{7 \quad 8665 \quad (237)^2}}$$

$$= \frac{27545 \quad 24648}{\sqrt{12922 \quad 10816}\sqrt{60655 \quad 56169}}$$

$$= \frac{2897}{\sqrt{2106 \quad 4486}} = 0.94$$

वर्गीकृत श्रेणी में सहसम्बन्ध
(Correlation in Grouped Series)

यदि दो चरों के मूल्यों में वर्गीकृत करके विभिन्न वर्गों की आवत्तियाँ दी गयी हों तो सहसम्बन्ध गुणक की गणना की जा सकती है। इस प्रकार वर्गीकृत समंक-सारणी की सहसम्बन्ध सारणी (Correlation Table) अथवा द्वि-चर आवत्ति सारणी (Bivariate Frequency Table) कहते हैं। इसमें प्रत्येक कोष्ठ आवत्ति का सम्बन्ध दोनों चरों से होता है।

ऐसी श्रेणी में सहसम्बन्ध गुणक निकालने के दो तरीके हैं। प्रथम विधि के अनुसार दोनों चरों के प्रमाप विचलन पथक -पथक निकालकर, एक विस्तत तालिका द्वारा S*dxdy* ज्ञात की जाती है। दूसरी रीति के अनुसार तालिका को इस प्रकार समायोजित किया जाता है कि उसी तालिका से प्रमाप विचलन भी निकाले जा सकते हैं। वर्गीकृत श्रेणी में सहसम्बन्ध ज्ञात करने के लिए सूत्र में आवत्ति (*f*) के लिए समायोजन करना पड़ता हैं वगीकृत श्रेणी में सहसम्बन्ध गुणक ज्ञात करने का सूत्र इस प्रकार है:

$$r = \frac{fdxdy \quad N(a_1 \quad x)(a_2 \quad x_2)}{N \quad {}_1 \quad {}_2}$$

व्यवहार में इस सूत्र का प्रयोग नहीं किया जाता है क्योंकि इसमें माध्य (*a*) तथा प्रमाप विचलन (s) निकालने पड़ते हैं। अन्य सूत्र इस प्रकार हैं :

$$\frac{N \quad fdxdy \quad fdx \quad fdy}{\sqrt{\left[N \quad fdx^2 \quad (\ fdx)^2\right]\left[N \quad fdy^2 \quad (\ fdy)^2\right]}}$$

इस सूत्र में S*fdxdy* कोष्ठ-आवत्तियों तथा तत्सम्बन्धी विचलनों के गुणनफलों का योग होता है।

यदि विचलनों में पद-विचलन किये जाते हैं, तो *i* (समापवर्तक) से गुणा करने की आवश्यकता नहीं है, क्योंकि सूत्र में अंश और हर दोनों में उभयनिष्ठ गुणक ($ix \times iy$) से गुणा करने पर अंश व हर का अनुपात पूर्ववत् रहता है।

Illustration 5 : Calculate the co-efficient of correlation from the data given below in the table showing days lost by workers according to their ages :

	Number of Workers per Age Group (years)					
Days lost	**20-30**	**30-40**	**40-50**	**50-60**	**60-70**	**Total**
1 - 2	4	—	4	—	—	8
3 - 4	8	8	4	4	—	24
5 - 6	4	10	16	—	—	32
7 - 8	—	4	8	4	—	16
9 - 10	—	—	12	—	—	12
11 - 12	—	—	—	7	1	8
Total	**16**	**24**	**44**	**15**	**1**	**100**

Solution : सह सम्बन्ध गुणांक का मान निकालने के लिए नीचे दिखाई गई तालिका बनाई जाएगी।

X →			Class	20-30	30-40	40-50	50-60	60-70				
			m.v.	25	35	45	55	65				
			dx (45)	– 20	– 10	0	+ 10	+ 20				
Y ↓			*dx'* \ *i*	– 2	– 1	0	+ 1	+ 2	Total	*fdy'*	fdy^2	*fdxdy*
Class	*m.v.*	*dy* (65)	*dy'* (*i*)									
1 — 2	1.5	– 4	– 2	+ 4 4 + 16	—	0 4 0	—	—	8	– 16	32	+ 16
3 — 4	3.5	– 2	– 1	+ 2 8 + 16	+ 1 8 + 8	0 4 0	– 1 4 – 4	—	24	– 24	24	+ 20
5 — 6	5.5	0	0	0 4 0	0 12 0	0 16 0	—	—	32	0	0	0
7 — 8	7.5	+ 2	+ 1	—	– 1 4 –4	0 8 0	+ 1 4 + 4	—	16	+ 16	16	0
9 — 10	9.5	+ 4	+ 2	—	—	0 12 0	—	—	12	+ 24	48	+ 0
11 — 12	11.5	+ 6	+ 3	—	—	—	+ 3 7 + 21	+ 6 1 + 6	8	+ 24	72	+ 27
			Total	16	24	44	15	1	100	+ 24	192	+ 63
			fdx'	– 32	– 24	0	15	2	– 39			
			fdx^2	64	24	0	15	4	107			
			fdxdy	+ 32	+ 4	0	+ 21	+ 6	+ 63			

इस विस्तत तालिक बनाने की रीति निम्नांकित प्रकार है :-

(1) x तथा y श्रेणियों के मध्य बिन्दु ज्ञात करके, किसी संख्या को कल्पित माध्य लेकर विचलन निकाले जाते हैं जो क्रमश: dx तथा dy होते हैं।

(2) x श्रेणी की आवत्तियों को dx से गुणा करके $Sfdx$ ज्ञात किया जाता है। इसी प्रकार y श्रेणी की आवत्तियों को dy से गुणा करके $Sfdx$ ज्ञात किया जाता है।

(3) fdx को dx से गुणा करके $Sfdx^2$ तथा fdy से dy का गुणा करके $Sfdy^2$ ज्ञात किये जाते हैं।

(4) ***dxdy* ज्ञात करने की विधि :** प्रत्येक कोष्ठ (cell) के ऊपर जो dx तथा बायीं ओर dy है उनका गुणा करके गुणनफल कोष्ठ के ऊपर बायें लिख लिया। जिन कोष्ठों में आवत्तियाँ हैं उनको ऊपर लिखे $dxdy$ को कोष्ठ आवत्ति से गुणा करके नीचे दायें ओर लिख लिया। $\boxed{dx \times dy \rightarrow f \rightarrow f \times dx \times dy}$ बाद में कॉलम-वार तथा रेखा (row) बार $dxdy$ का गुणा करके योग के कालम में लिख लिया। बाद में सभी योगों को जोड़कर $Sfdxdy$ ज्ञात किया जाता है।

(5) एक लाइन के कोष्ठों में ऊपरी कोनों में रखी गयी संख्याओं को जोड़कर $fdxdy$ के खाने में रख दी जाती है। इन खाने का योग ही $Sfdxdy$ होता है।

(6) सूत्र में मूल्यों को रखकर 'r' का मान ज्ञात किया जाता है।

$$r = \frac{\text{N}\;\; fdxdy \quad fdx.\; fdy}{\sqrt{\text{N}\;\; fdx^2 \quad (\;\; fdx)^2}\sqrt{\text{N}\;\; fdy^2 \quad (\;\; fdy)^2}}$$

$$= \frac{100 \quad 63 \quad (\;\; 39) \quad 24}{\sqrt{100 \quad 107 \quad (\;\; 39)^2}\sqrt{100 \quad 192 \quad (24)^2}}$$

$$= \frac{6300 \quad 936}{\sqrt{10700 \quad 1521}.\sqrt{19200 \quad 576}}$$

$$= \frac{7236}{\sqrt{9179}\sqrt{18624}}$$

$$= 0.55$$

कार्ल पियर्सन के सहसम्बन्ध गुणक की मान्यताएँ

(Assumptions of Karl Pearson's Coefficient of Correlation)

कार्ल पियर्सन द्वारा प्रतिपादित सहसम्बन्ध गुणक निम्नलिखित तीन मान्यताओं पर आधारित हैं :

(1) सामान्यता (Normality) : सहसम्बन्धित समंक मालाएँ अनेक कारणों से प्रभावित होती है, जिससे उनमें सामान्यता आ जाती है।

(2) कारण-परिणाम सम्बन्ध (Casual Relationship) : समंक मालाओं को प्रभावित करने वाले स्वतंत्र कारणों में परस्पर कारण एवं परिणाम (Cause and Effect) का सम्बन्ध होता है। यदि इस प्रकार का सम्बन्ध न हो तो सहसम्बन्ध अर्थहीन होता है।

(3) रेखीय प्रकृति (Linear Nature) : यह मान्यता होती है कि दोनों समंक मालाओं में पारस्परिक रेखीय सम्बन्ध है, अर्थात् यदि दोनों पद-युग्मों को बिन्दु रेखीय पत्र पर प्रांकित किया जाये तो बिन्दु चित्र एक सरल रेखा के रूप में होगा।

कार्ल पियर्सन के सह-सम्बन्ध गुणांक की विशेषताएँ
(Characteristics of Karl Pearson's Coefficient of Correlation)

(1) सहसम्बन्ध गुणांक का मान हमेशा ± 1 के बीच में ही होता है। किसी भी दशा में इसका मान -1 से कम व $+1$ से अधिक नहीं हो सकता।

(2) उदगम (Origin) एवं पैमाने (Scale) के परिवर्तन का इस गुणांक के मान पर कोई प्रभाव नहीं पड़ता। यदि x चर की सभी मदों में समान अंक जोड़ दिया जाए या उन सब मदों में से घटा दिया जाए (Change of Origin) अथवा सभी मदें एक अंक से गुणा या भाग कर दिया जाए (Change of Scale) तो भी सहसम्बन्ध गुणांक के मान में कोई परिवर्तन नहीं आएगा।

(3) सहसम्बन्ध गुणांक का प्रतीपगमन गुणांकों (Regression Coefficients) से भी सीधा सम्बन्ध है जैसा कि निम्नलिखित सूत्र से विदित है :

$$r^2 = bxy \times byx$$

सम्भाव्य-विभ्रम
(Probable Error)

कार्ल पियर्सन (Karl Pearson) के सहसम्बन्ध गुणक की महत्ता (Significance) ज्ञात करने के लिए सम्भाव्य विभ्रम (Probable Error) का प्रयोग किया जाता है। सम्भाव्य विभ्रम, विभ्रम की वह मात्रा है जिसे यदि किसी विशिष्ट सांख्यिकीय माप (जैसे माध्य, सहसम्बन्ध गुणक आदि) में जोड़ देने पर घटा देने से वे दो सीमाएँ ज्ञात हो जाती हैं जिनके अंतर्गत मूल समग्र (Original Universe) में से लिए अन्य दैव-न्यादर्शों (Random Samples) के कथित सांख्यिकीय माप पाये जाने की सम्भावना होती है। होरेस सीक्रीस्ट के अनुसार, ''सहसम्बन्ध गुणक की सम्भाव्य विभ्रम वह राशि होती है जिसे यदि सहसम्बन्ध गुणक में जोड़ दिया जाए तथा घटा दिया जाये, तो ऐसी संख्या ज्ञात हो जाती है जिनके अन्तर्गत दैव-निदर्शन के आधार पर छाँटे गये पदों के सहसम्बन्ध गुणक पाये जाने की समान संभावनाएँ होती हैं। ''ह्वेल्डन (Wheldon) के शब्दों में, ''सम्भाव्य विभ्रम निर्धारित गुणक के ऊपर व नीचे सीमाओं को परिभाषित करती है जिनके अन्तर्गत अन्य न्यादर्शों का उसी प्रकार निर्धारित सहसम्बन्ध गुण पाये जाने का समान अवसर होता है।'' सहसम्बन्ध गुणक (r) के सम्भाव्य विभ्रम ज्ञात करने के लिए निम्नलिखित सूत्र का प्रयोग किया जाता है :

$$\text{Probable Error of '}r\text{'} = .6745 \times \frac{1 \quad r^2}{\sqrt{N}}$$

इस सूत्र में, 6745 स्थिर संख्या (Constant), r = सहसम्बन्ध गुणक तथा N = पदों की संख्या के लिए प्रयुक्त किया गया है। उदाहरण 5 का सहसम्बन्ध गुणक + 0.55 है, इसका सम्भाव्य विभ्रम इस प्रकार ज्ञात किया जायेगा :

$$\text{P.E.} = .6745 \times \frac{1 \quad r^2}{\sqrt{N}} = .6745 \times \frac{1 \quad .55^2}{\sqrt{100}}$$

$$= .6745 \times \frac{1 \quad .3025}{10} = .6745 \times \frac{.6975}{10} = .0465$$

सहसम्बन्ध गुणक के साथ सम्भाव्य विभ्रम इस प्रकार लिखा जाता है :

$r = + .55 \pm .0465$, अर्थात्, उक्त सहसम्बन्ध गुणक की सीमाएँ $+ .55 + .0465 = + .5965$ तथा $+ .55 - .0465 = .5035$ होंगी। इसका तात्पर्य यह है कि उसी समग्र में से 100 और पदों का दैव-न्यादर्श लेकर उनका सहसम्बन्ध-गुणक ज्ञात किया जाए तो उसे + .5965 तथा .5035 के बीच ही होने की सम्भावना होगी।

सम्भाव्य विभ्रम का सहसम्बन्ध गुणक की महत्ता ज्ञात करने के रूप में प्रयोग

सम्भाव्य विभ्रम को सहसम्बन्ध गुणक की महत्ता की माप (Measure of Significance) के रूप में माना जाता है। सम्भाव्य विभ्रम पर आधारित कार्ल पियर्सन के सहसम्बन्ध गुणक के निर्वचन के सम्बन्ध में कुछ नियम इस प्रकार हैं :-

(1) यदि सहसम्बन्ध गुणक सम्भाव्य विभ्रम से कम ($r <$ P.E.) है तो दोनों श्रेणियों में सहसम्बन्ध की उपस्थिति का कोई प्रमाण नहीं है।

(2) यदि सहसम्बन्ध गुणक सम्भाव्य विभ्रम के छः गुने से अधिक ($r > 6$ P.E.) हैं तो सहसम्बन्ध गुणक महत्वपूर्ण हैं तथा दोनों श्रेणियों में सहसम्बन्ध का होना निश्चित है।

(3) यदि सहसम्बन्ध गुणक .3 से कम हो तथा उसका सम्भाव्य विभ्रम अपेक्षाकृत कम हो तो सहसम्बन्ध महत्वपूर्ण नहीं माना जाना चाहिए।

(4) यदि सहसम्बन्ध गुणक .5 से अधिक हो तथा उसका सम्भाव्य विभ्रम बहुत कम हो तो सहसम्बन्ध का अस्तित्व लगभग निश्चित होता है।

अवधारणा गुणांक
(Coefficient of Determination)

यदि हमें दो चरों के सहसम्बन्ध का निर्वचन (Interpretation) करना है तो अवधारणा गुणांक सहसम्बन्ध गुणांक से बेहतर विधि है। सहसम्बन्ध गुणांक से हमें यह पता लगता है कि दो चरों के बीच में सहसम्बन्ध है या नहीं। अवधारणा गुणांक से हमें ये पता लगता है कि एक चर के मान के मूल्यों के अपकिरण में दूसरे चर का कितना योगदान है। सह-सम्बन्ध गुणांक का वर्ग (r^2) अवधारणा गुणांक कहलाता है। उदाहरण के लिए, यदि दो चरों के बीच y का मान 0.6 है तो अवधारणा गुणांक का मान 0.36 [$(0.6)^2$] होगा। इसका अर्थ यह है कि x (or y) के मूल्यों में जितना भी अपकिरण है, उसमें से 36% y (or x) के कारण है। बाकी बचा हुआ 64% अपकिरण किन्हीं और चरों के कारण है, जिन्हें हमने सहसम्बन्ध विश्लेषण में शामिल नहीं किया।

4. स्पियरमैन की अनुपस्थिति रीति द्वारा सहसम्बन्ध गुणक की गणना
(Spearman's Ranking Method)

पदों की अनुस्थिति (Ranks) के आधार पर सहसम्बन्ध गुणक की गणना करने की रीति प्रतिपादन प्रोफेसर चार्ल्स स्पियरमैन (Charles Spearman) द्वारा किया गया। यह रीति श्रेणी के पदों की अनुस्थिति अथवा क्रम (Rank) पर आधारित है। पद मूल्यों के आकार (Size) के अनुसार उनकी अनुपस्थिति निश्चित की जाती है।

सहसम्बन्ध ज्ञात करने की इस रीति का उपयोग उस दशा में उपयोगी होता है जबकि :

(1) तथ्यों की प्रत्यक्ष संख्यात्मक माप सम्भव न हो परन्तु पदों को एक निश्चित क्रम में रखा जाना सम्भव हो। उदाहरणार्थ, बुद्धिमत्ता, सुन्दरता आदि गुणात्मक तथ्यों को प्रत्यक्ष अंकों में नहीं नापा जा सकता परन्तु विभिन्न इकाइयों को इन गुणों के आधार पर क्रमबद्ध अवश्य किया जा सकता है।

(2) समकों में अनियमितता हो।

(3) सीमान्त पदों मे अस्पष्टता अथवा त्रुटिपूर्ण व्यवहार पाया जाता है।

इस रीति का उपयोग व्यक्तिगत श्रेणी में ही हो सकता है, आवत्ति वितरणों (Frequency Distribution) में नहीं। सहसम्बन्ध गुणक ज्ञात करने के लिए मौलिक पद-मूल्यों को विचारगत नहीं किया जाता, केवल उनकी अनुस्थिति के ही आधार पर सहसम्बन्ध गुणक की गणना की जाती है। इस रीति से सहसम्बन्ध गुणक ज्ञात करने का सूत्र निम्नलिखित है :

'p' (Pronounced as 'rho') or take $r = 1 - \dfrac{6\ \ d^2}{N\left(N^2\ \ \ 1\right)}$ or $- \dfrac{6\ \ d^2}{N^3\ \ \ N}$

इस सूत्र के संकेताक्षरों का अर्थ इस प्रकार है :-

'p' or r = सहसम्बन्ध गुणक (Coefficient of Correlation)

d^2 = अनुस्थिति के अंतरों के वर्गों का जोड़ (Total of Squares of Rank Differences)

N = पद-युग्मों की संख्या (Number of Pairs of Items)

अनुस्थिति निश्चित करने की रीति
(Method of Ranking)

पद-मूल्यों की अनुस्थिति निश्चित करने में सबसे बड़े आकार वाले मूल्य को 1, उससे कम आकार वाले मूल्य को 2 और इसी प्रकार अनुस्थिति-क्रम निश्चित किये जाते हैं। अनुस्थिति निश्चित करने में उस समय कठिनाई उपस्थित होती है जबकि एक ही मूल्य के दो या अधिक पद उस श्रेणी में हों। ऐसी दशा में निम्न विधियों में से एक को अपनाया जा सकता है :

(1) **कोष्ठक-अनुस्थिति रीति** (Bracket Rank Method) **:** इस रीति द्वारा समान मूल्य के सभी पदों के समान क्रम या अनुस्थिति निश्चित की जाती है। अगले मूल्य को, अगला क्रम दिया जाता है। यह निम्न उदाहरण से स्पष्ट हो जाता है :

Items	:	30	32	35	35	40	42
Rank	:	6	5	3	3	2	1

(2) **औसत-अनुस्थिति रीति** (Average Rank Method) **:** इस रीति द्वारा समान मूल्य के सभी पदों को उनकी अनुस्थिति की औसत के बराबर क्रम निश्चित किये जाते हैं, उदाहरणार्थ :

Items	**:**	30	32	35	35	40	42
Rank	**:**	6	5	3.5	3.5	2	1

स्पियरमैन की रीति से सहसम्बन्ध गुणक ज्ञात करने में अनुस्थिति निश्चित करने के पश्चात x श्रेणी के क्रमों में से y श्रेणी के क्रमों को घटाकर क्रमान्तर (Rank Difference) निकाल कर उनका वर्ग (Squares) करके उनका योग ज्ञात किया जाता है। इस विधि से ज्ञात 'r' का मान भी ± 1 के बीच होता है। कार्ल पियर्सन के सूत्र से ज्ञात किया गया सहसम्बन्ध गुणक तथा स्पियरमैन के सूत्र से ज्ञात किया सहसम्बन्ध गुणक के मूल्यों में कोई विशेष अन्तर नहीं होता है। प्रोफेसर थर्सटोन (Thurstone) ने अपनी पुस्तक 'Fundamentals of Statistics' में एक सूची दी है जिससे अनुस्थिति रीति से निकाले गये सहसम्बन्ध गुणकों से सम्बन्धित कार्ल पियर्सन के सहसम्बन्ध गुणक दिये गये हैं। यह सूचि निम्नलिखित है :-

Correlation between Rank and Pearson's coefficient.

Rank Correlation Coefficient	Pearson's Coefficient	Rank Correlation Coefficient	Pearson's Coefficient
.00	.000	.55	.568
.05	.052	.60	.618
.10	.105	.65	.668
.15	.157	.70	.717
.20	.210	.75	.765
.25	.261	.80	.813
.30	.313	.85	.861
.35	.364	.90	.908
.40	.416	.95	.954
.45	.467	1.00	1.000
.50	.518		

Illustration 6 : A survey was conducted selecting a sample of 12 workers. The long servicing clerks feel that they should have a seniority increment based on the length of service. An assessment of their efficiency by the scholar with the assistance of the management produces a ranking of efficiency. Data of the above is given below :

Calculate Spearman's coefficient of rank correlation.

Rank according to length of service	**:**	1	2	3	4	5	6	7	8	9	10	11	12
Rank according to efficiency	**:**	2	3	5	1	9	10	11	12	8	7	6	4

Solution : Calculation of rank correlation coefficient for claiming a seniority increment. Let us denote ranks according to length of service and to efficiency by R_1 and R_2 respectively.

R_1	R_2	D ($R_1 - R_2$)	D^2
1	2	1	1
2	3	1	1
3	5	2	4
4	1	3	9
5	9	4	16
6	10	4	16
7	11	4	16
8	12	4	16
9	8	1	1
10	7	3	9
11	9	5	25
12	4	8	64
			SD^2 = 178

Applying the Spearman formula,

$$r = 1 - \frac{6\quad D^2}{n^3\quad n}$$

$$= 1 - \frac{6\quad 178}{12^3\quad 12}$$

$$= 1 - \frac{1068}{1716}$$

$$= 1 - 0.6223$$

$$= 0.3777$$

Illustration 7. In an evaluation of answer-scripts, the following marks were awarded by the examiners :

1st examiner	**:**	88	95	70	60	50	80	75	85
2nd examiner	**:**	80	86	88	55	48	85	82	75

Do you agree the evaluation by the two examiner is fair ?

Solution : Let us denote 1st examiner and 2nd examiner awarding marks by X and Y respectively. Calculate the value of square of difference to find whether the awarding marks by the two examiners is fair or not.

X	R_1	Y	R_2	D	D^2
88	2	80	5	3	9
95	1	86	2	1	1
70	6	88	1	5	25
60	7	55	7	0	0
50	8	48	8	0	0
80	4	85	3	1	1
75	5	82	4	1	1
85	3	75	6	3	9
					SD^2 = 46

Applying formula,

$$r = 1 - \frac{6 \quad D^2}{n^3 \quad n}$$

$$= 1 - \frac{6 \quad 46}{8^3 \quad 8}$$

$$= 1 - \frac{276}{504} = 50$$

$$= 1 - 0.5476$$

$$= 0.4524 \text{ or,}$$

$$= 0.45$$

The value of rank correlation coefficient derived as above shows not much fair in awarding marks in the sense discrepancy has arisen in evaluating the answer-scripts between the two examiners.

समान क्रम के लिए संशोधन

(Correction for Equal Ranks)

श्रेणी में दो या दो से अधिक पदों से अधिक पदों का समान मूल्य होने पर उन्हें औसत क्रम प्रदान किये जाते हैं। ऐसी स्थिति में अनुस्थिति सहसम्बन्ध गुणांक निकालने के लिए सूत्र को संशोधित करना पड़ता है।

संशोधित सूत्र इस प्रकार है :

$$r = 1 - \frac{6\left[\ d^2 \quad \frac{1}{12}\left(m_1^3 \quad m\right) \quad \frac{1}{12}\left(m_2^3 \quad m_2\right) \quad .....\right]}{\mathrm{N}\left(\mathrm{N}^2 \quad 1\right)}$$

m उन पद मूल्यों की संख्या, जिनके क्रम समान है, का प्रतिनिधित्व करते हैं।

Illustration 8. Calculate the rank coefficient of correlation of the following data :-

x	**:**	80	78	75	75	68	67	60	59
y	**:**	12	13	14	14	14	16	15	17

Solution :

X	Rank	Y	Rank	Rank difference *d*	d^2
80	1	12	8	– 7	49
78	2	13	7	– 5	25
75	3.5	14	5	– 1.5	2.25
75	3.5	14	5	– 1.5	2.25
68	5	14	5	0	0
67	6	16	2	+ 4	16
60	7	15	3	+ 4	16
59	8	17	1	+ 7	49
N = 8		**N = 8**			**S = 159.50**

$$r = 1 - \frac{6\left[\ d^2 \quad \frac{1}{12}\left(m^3 \quad m\right)\right]}{N\left(N^2 \quad 1\right)} = 1 - \frac{6\left[159.50 \quad \frac{1}{12}\left(2^3 \quad 2\right) \quad \frac{1}{12}\left(3^3 \quad 3\right)\right]}{8\left(8^2 \quad 1\right)}$$

$$= 1 - \frac{6 \quad (159.50 \quad 0.5 \quad 2)}{8\,(64 \quad 1)} = 1 - \frac{6 \quad 162}{8 \quad 63} = 1 - \frac{972}{504} = 1 - 1.19 = -.93$$

अनुस्थिति रीति के गुण व दोष
(Merits and Demerits of Rank Method)

गुण (Merits)

(1) यह विधि गणना में सरल व समझने में आसान है।

(2) जब सूचना गुणात्मक हो जैसे सुंदरता, कार्यकुशलता, ईमानदारी, साक्षात्कार में प्रदर्शन आदि तो यही विधि प्रयुक्त की जाती है।

(3) यदि दो या अधिक मद समान मूल्य के नहीं हों तो कार्ल पियर्सन गुणांक तथा स्पियरमैन गुणांक में ज्यादा अंतर नहीं होता जैसा कि Thurstone का सारणी से भी सिद्ध होता है।

दोष (Demerits)

(1) वगीकृत श्रेणियों के लिए हम यह विधि प्रयुक्त नहीं कर सकते।

(2) इस विधि से निकाला गया मान परिशुद्ध नहीं होता क्योंकि इसमें चरों के वास्तविक मद की जगह उनके क्रम लिए जाते हैं।

5. संगामी विचलन रीति
(Concurrent Deviation Method)

संगामी विचलन रीति सहसम्बन्ध ज्ञात करने की सरल तथा गणना करने में आसान विधि है। इस रीति का उपयोग उस दिशा में किया जाता है जब यह ज्ञान करना हो कि दो समंकमालाओं के मध्य किस प्रवत्ति —धनात्मक अथवा ऋणात्मक—का सहसम्बन्ध है। इस रीति में अंकगणितीय माध्य से विचलन नहीं निकाले जाते बल्कि प्रत्येक पद का विचलन उससे पूर्व वाले पद के मूल्य से निकाला जाता है। विचलन निकालने में भी उनकी मात्रा की उपेक्षा करके केवल विचलन की दिशा (+ अथवा –) को ही ध्यान में रखा जाता है। यदि विचलन शून्य हो, अर्थात् पूर्व पद के मूल्य के बराबर ही सम्बन्धित पद का मूल्य हो, वहाँ = का निशान लगा दिया जाता है। x तथा y श्रेणियों के तत्सम्बन्धी विचलन-चिन्हों (Corresponding Deviation Signs) को गुणा करके संगामी विचलन (Concurrent Deviation) वाले कॉलम में लिख दिया जाता है। गुणा के नियम के अनुसार दो समान चिन्हों का गुणा + होता है अर्थात् + + तथा – – का गुणा + होता है। असमान चिन्हों अर्थात् + – अथवा – + का गुणा – होता है। संगामी विचलन के कॉलम में लिखे + चिन्हों को जोड़ लिया जाता है तथा निम्न सूत्र से सहसम्बन्ध गुणक ज्ञात किया जाता है :

$$r_c = \pm\sqrt{\quad \frac{2C \quad N}{N}}$$

r_c = संगामी विचलन गुणक (Coefficient of Concurrent Deviations)

C = संगामी विचलनों की संख्या (Number of Concurrent Deviations)

N = विचलन-युग्मों की संख्या जो पद-युग्मों की संख्या से कम होती है। (Number of pairs of Deviations)

सहसम्बन्ध गुणक में + अथवा – चिह्न को $\left[\frac{2C \quad N}{N}\right]$ के चिह्न के अनुसार लगाया जाता है। यदि इसका परिणाम + होगा तो r भी + और – होगा तो r भी – होता है।

Illustration 9 : Following are the marks scored by a group of 11 students in Commerce and Economics :

Student	:	A	B	C	D	E	F	G	H	I	J	K
Marks in Commerce	:	65	40	35	75	63	80	35	20	80	60	50
Marks in Economics	:	60	55	50	56	30	70	40	35	80	75	80

Calculate the coefficient of correlation by the method of Concurrent Deviations.

Solution :

Student X	Marks in Commerce *dx*	Deviation from Preceding Students Y	Marks in Economics *dy*	Deviation from Preceding Students C	Concurrent Deviation	Disagreement
A	65		60			
B	40	–	55	–	+	
C	35	–	50	–	+	
D	75	+	56	+	+	
E	63	–	30	–	+	
F	80	+	70	+	+	
G	35	–	40	–	+	
H	20	–	35	–	+	
I	80	+	80	+	+	
J	60	–	75	–	+	
K	50	–	80	+		–
		N = 10		**N = 10**	**SC = 9**	**1**

Substituting these values in the formula

$$r = \pm\sqrt{\frac{2C \quad N}{N}} \text{ we get}$$

$$r = +\sqrt{\frac{18 \quad 10}{10}} = +\sqrt{\quad 80} = +0.89$$

अन्य उदाहरण

Illustration 10. In a city, a survey was conducted to find out the correlation between age and blindness. The relevant data obtained are as follows :-

Age (years)	No. of Persons (000')	No. of Blind persons
0-10	100	55
10-20	60	40
20-30	40	38
30-40	36	36
40-50	24	34
50-60	11	22
60-70	8	18

Solution : In this problem, first no. of blind persons is to be calculated for a common base, like 100000. Thus no. of blind persons per 100000 population for each age group. The new nos. are as follows :-

Age Group	No. of blind
10-20	$\frac{40}{60000} \times 100000 = 67$
20-30	$\frac{38}{40000} \times 100000 = 95$
30-40	$\frac{36}{36000} \times 100000 = 100$
40-50	$\frac{34}{24000} \times 100000 = 142$
50-60	$\frac{22}{11000} \times 100000 = 200$
60-70	$\frac{18}{8000} \times 100000 = 225$

Age (X)	Mid values	*dx* = 35	dx^2	Y	*dy* = 126	dy^2	*dxdy*
0-10	5	– 30	900	55	– 71	5041	2130
10-20	15	– 20	400	67	– 59	3481	1180
20-30	25	– 10	100	95	– 31	961	310
30-40	35	0	0	100	– 26	676	0
40-50	45	10	100	142	16	256	160
50-60	55	20	400	200	74	5476	1480
60-70	65	30	900	225	99	9801	2970
		0	**2800**		**2**	**25692**	**8230**

Now

$$S dx = 0, \quad S dx^2 = 2800, \quad S dy = 2$$

$$S dy^2 = 25692, \quad S dxdy = 8230$$

So substituting these values in the

$$r = \frac{N \quad dxdy \qquad dx \quad dy}{\sqrt{N \quad dx^2 \quad (\ dx)^2}\sqrt{N \quad dy^2 \quad (\ dy)^2}}$$

$$= \frac{7 \quad 8230 \quad 0 \quad 2}{\sqrt{7 \quad 2800 \quad (0)^2}\sqrt{7 \quad 25692 \quad (2)}}$$

$$= \frac{57610}{\sqrt{19600}\quad\sqrt{179840}}$$

$$= 0.97$$

Illustration 11 : A computer while calculating correlation coefficient between two variables X and Y obtained the following values :-

$$n = 30, \quad Sx = 120, \quad Sx^2 = 600, \quad Sy = 90, \quad Sy^2 = 350, \quad Sxy = 356$$

It was, however, later discovered that it had wrongly copied two pairs of values as

x	*y*
8	12
10	8

instead of

x	*y*
8	10
12	7

. Find the correct value of *r*.

Solution : सर्वप्रथम सही मान निकाले जाएंगे :

$$\text{Correct } Sx = 120 - (8 + 10) + (8 + 12) = 122$$

$$\text{Correct } Sy = 90 - (12 + 8) + (10 + 7) = 87$$

$$\text{Correct } Sx^2 = 600 - (8^2 + 10^2) + (8^2 + 12^2) = 664$$

$$\text{Correct } Sy^2 = 350 - (12^2 + 8^2) + (10^2 + 7^2) = 291$$

$$\text{Correct } Sxy = 356 - (8 \times 12 + 10 \times 8) + (8 \times 10 + 12 \times 7) = 344$$

Now

$$r = \frac{30 \quad 344 \quad 122 \quad 87}{\sqrt{30 \quad 644 \quad (122)^2}\sqrt{30 \quad 291 \quad (87)^2}}$$

$$= \frac{10320 \quad 10614}{\sqrt{19320 \quad 14884}.\sqrt{8730 \quad 7569}}$$

$$= \frac{294}{\sqrt{4436}.\sqrt{1161}} = -0.129$$

Illustration 12 : Calculate coefficient of correlation for the following data :

$$n = 10, Sx = 140, Sy = 150, S(x - 10)^2 = 180, S(y - 15)^2 = 215, S(x - 10)(y - 15) = 60$$

Solution : इस प्रश्न में प्रथम सूत्र का प्रयोग नहीं किया जा सकता क्योंकि अंकगणितीय माध्य से विचलनों का मान नहीं दिया गया है। अत: तीसरा सूत्र (प्रत्यक्ष विधि) का प्रयोग किया जाएगा।

$$S(x - 10)^2 = 180$$

or $$S(x^2 - 20 \times + 100) = 180$$

or $$Sx^2 - 20\, S \times + 10 \times 100 = 180$$

or $$Sx^2 = 180 - 1000 + 2800 = 1980$$

$$S(y - 15)^2 = 215$$

or $$S(y^2 - 30y + 225) = 215$$

$$Sy^2 - 30 \times 150 + 10 \times 225 = 215$$

$$Sy^2 = 215 + 4500 + 2250 = 2465$$

$$S(x - 10)(y - 15) = 60$$

or $$S(xy - 15x - 10y + 150) = 60$$

or $$Sxy - 15 \times 140 - 10 \times 150 + 10 \times 150 = 60$$

or $$Sxy = 60 + 2100 + 1500 - 1100$$

$$= 2160$$

Hence

$$r = \frac{10 \quad 2160 \quad 140 \quad 150}{\sqrt{10 \quad 1980 \quad (140)^2}\sqrt{10 \quad 2465 \quad (150)^2}}$$

$$= \frac{21600 \quad 21000}{\sqrt{19800 \quad 19600}\sqrt{24650 \quad 22500}}$$

$$= \frac{600}{\sqrt{200}\sqrt{2110}} = 0.915$$

Illustration 13 :

(*a*) For a given problem, *r* = 0.8, probable error is 0.09. Is the value of *r* significant ?

(*b*) A student calculates the value of *r* as + 0.7 when the value of N is 5 and concludes that '*r*' is highly significant. Is he correct ?

(*c*) Show that *r* is significant if —

N = 16, P.E. = 0.125 and 0.6745 = 2/3

Solution :

(*a*) $\frac{r}{\text{P.E.}} = 9$ (App.)

Since the value or '*r*' is 9 times the value of P.E., so it's value is significant.

(*b*) $r = +.7, \text{N} = 5$

$$\text{P.E.} = 0.6745 \times \frac{1 \quad r^2}{\sqrt{\text{N}}} = .6745 \times \frac{1 \quad 0.49}{\sqrt{5}} = 0.6745 \times \frac{0.51}{2.236}$$

$$= 0.6745 \times .228 = .152$$

सहसम्बन्ध गुणक सम्भाव्य विभ्रम के छः गुने से कम है (.7 < 6 + .152) अतः सहसम्बन्ध गुणक महत्त्वपूर्ण नहीं है। निकाला गया निष्कर्ष गलत है।

(*c*) $\text{N} = 16, \text{P.E.} = .125, 0.6745 = \frac{2}{3}$

पहले P.E. तथा N की सहायता से *r* का मूल्य ज्ञात किया जाना चाहिए।

$$\text{P.E.} = 0.6745 \times \frac{1 \quad r^2}{\sqrt{\text{N}}} \qquad = 0.125 = \frac{2}{3} \times \frac{1 \quad r^2}{\sqrt{16}}$$

$$= 0.125 = \frac{2}{3} \times \frac{1 \quad r^2}{4} \qquad = 0.125 = \frac{2 \quad 2r^2}{12}$$

$$= 1.8 = 2 - 2r^2 \qquad = 2r^2 = 2 - 1.8 \qquad = 2r^2 = .2$$

$$= r^2 = 0.1 \qquad = r = \sqrt{0.1} = 0.3$$

$$\text{P.E.} \times 6 = 1.25 \times 6 = .75$$

इस उदाहरण में $r = 0.3$ है तो $\text{P.E.} \times 6 = 0.75$ से कम है, अतः सहसम्बन्ध गुणक महत्त्वपूर्ण नहीं है।

Illustration 14 : Two judges in a baby-competition rank the 12 entries as follows :-

Entry	:	A	B	C	D	E	F	G	H	I	J	K	L
X Judge	:	1	2	3	4	5	6	7	8	9	10	11	12
Y Judge	:	12	9	6	10	3	5	4	7	8	2	11	1

What degree of agreement is there between the Judges ?

Solution :

Entry	**Rank by *x***	**Rank by *y***	***d***	***d*²**
A	1	12	– 11	121
B	2	9	– 7	49
C	3	6	– 3	9
D	4	10	– 6	36
E	5	3	+ 2	4
F	6	5	+ 1	1
G	7	4	+ 3	9
H	8	7	+ 1	1
I	9	8	+ 1	1
J	10	2	+ 8	64
K	11	11	0	0
L	12	1	+ 11	121
N = 12				**S = 416**

$$r = 1 - \frac{6 \quad d^2}{N^3 \quad N} = 1 - \frac{6 \quad 416}{12^3 \quad 12} = 1 - \frac{2496}{1716} = -\frac{780}{1716} = -.455$$

it indicates that the judges have fairly strongly divergent likes and dislikes so far as ranking of the babies is concerned.

सारांश
(Summary)

— सहसम्बन्ध विश्लेषण सांख्यिकी का एक महत्त्वपूर्ण माप है जिसकी सहायता से हम किन्हीं भी दो या अधिक चरों के मध्य सम्बन्ध ज्ञात कर सकते हैं।

— दो चरों के बीच सहसम्बन्ध मुख्यत: दो प्रकार का होता है — ऋणात्मक या धनात्मक।

— दो चरों के बीच सहसम्बन्ध निकालने की सभी विधियों में कार्ल पियर्सन की गुणांक की विधि (Carl Pearson's coefficient of correlation) सबसे लोकप्रिय व उपयुक्त है। इस गुणांक हो हम '*r*' से प्रदर्शित करते हैं।

— *r* का मान हमेशा – 1 से + 1 के बीच में रहेगा।

— यदि प्राप्त सूचना गुणात्मक हो, जिसको हम कोई मान न दे सकें, तो उस अवस्था में *r* का मान ज्ञात नहीं किया जा सकता। ऐसी स्थिति में Spearman's coefficient of rank correlation निकाला जाता है।

— ये सहसम्बन्ध गुणों के क्रम अथवा अनुस्थिति (Rank) पर आधारित है।

— यदि हमें ये मालूम करना हो कि एक चर को कुल अपकिरण में दूसरे चर का कितना प्रतिशत योगदान है तो हम अवधारणा गुणांक निकालते हैं जो कि सहसम्बन्ध गुणांक (*r*) पर आधारित है।

— यदि हमें ये मालूम करना हो कि एक चर को कुल अपकिरण में दूसरे चर का कितना प्रतिशत योगदान है तो हम अवधारणा गुणांक निकालते हैं जो कि सहसम्बन्ध गुणांक (*r*) का वर्ग होता है।

— यदि दोनों चरों के उदगम व पैमान में परिवर्तन कर दिया जाए तो सभी '*r*' के मान में कोई परिवर्तन नहीं होगा। (Correlation coefficient is independent of change of origion and change of scale).

प्रश्नावली

(Exercise)

(1) सहसम्बन्ध की परिभाषा दीजिए और सांख्यिकी में इसकी उपयोगिता का विवेचन कीजिए। क्या यह सदा दो चर-मूल्यों के बीच कारण परिणाम सम्बन्ध को व्यक्त करता है ?

Define correlation and discuss its usefulness in Statistics. Does it always show cause and effect relationship between two variables ?

(2) सहसम्बन्ध क्या है ? इसके कितने प्रकार हैं ? व्याख्या कीजिए।

What is correlation ? What are its types ? Explain.

(3) सहसम्बन्ध ज्ञात करने की विभिन्न विधियों की व्याख्या कीजिए।

Explain different methods of measuring correlation.

(4) कार्ल पियर्सन का सहसम्बन्ध गुणांक क्या है ? इसे कैसे निकाला जाता है ? इसकी विशेषताएँ व सीमाएं क्या हैं ?

What is Karl Pearson's coefficient of correlation ? How it is measured ? What are its merits and limitations ?

(5) सम्भाव्य विभ्रम की परिभाषा दीजिए व इसकी महत्ता बताइए।

Define probable error and discuss its importance.

(6) सहसम्बन्ध तथा संगामी विचलन में अंतर स्पष्ट कीजिए। संगामी विचलन गुणांक निकालने की विधि की व्याख्या कीजिए।

Distinguish between correlation and concurrent deviation. Explain the method for calculating coefficient of concurrent deviation.

(7) कोटि अन्तर सहसम्बन्ध क्या है ? किन परिस्थितियों में उसे उपयोग में लाया जाता है ? इसका गुणांक कैसे निकाला जाता है ?

What is rank correlation ? In what situations is it used ? How is its co-efficient determined ?

(8) निम्नलिखित पर संक्षिप्त टिप्पणियाँ लिखिए :-

(क) विक्षेप चित्र (ख) निरर्थक सहसम्बन्ध

(ग) बिन्दुरेखीय विधि (घ) सम्भाव्य विभ्रम

Write short notes on the following :-

(a) Scatter diagram (b) Spurious or non-sense correlation

(c) The graphic method (d) Probable error

(9) निम्न समंकों को विक्षेप चित्र पर चिह्नित कीजिए :-

Draw a scatter diagram to represent the following data :

X	**:**	25	29	38	43	45	49	55	59
Y	**:**	52	57	69	76	79	84	92	96

[**Ans.** $r = +.99$]

(10) निम्न समंकों को बिन्दु-रेखीय विधि से प्रस्तुत करें :

Represent the following data with the help of graphic method :

X	**:**	42	44	58	55	89	98	66
Y	**:**	56	49	53	58	65	76	51

[**Ans.** $r = +.904$]

(11) निम्नलिखित दस कंपनियों के खर्च तथा बिक्री से सहसम्बन्ध गुणांक ज्ञात कीजिए :-

Find the coefficient of correlation between the sales and expenses of the following 10 firms :-

Firms	:	1	2	3	4	5	6	7	8	9	10
Sales	:	50	50	55	60	65	65	65	60	60	50
Expenses	:	11	13	14	16	16	15	15	14	13	13

[**Ans.** $r = + .79$]

(12) निम्न समंकों से पिताओं तथा पुत्रों की ऊंचाई के मध्य सहसम्बन्ध गुणांक ज्ञात कीजिए :-

Find out the correlation coefficient between the heights of father and son from the following data :-

Height of father (inches) x	:	65	63	67	64	68	62	70	66	68	67	67	71
Height of Son (inches) y	:	68	66	68	65	69	66	68	65	71	67	68	70

[**Ans.** $r = + .7$]

(13) निम्न तालिका 10 विद्यार्थियों द्वारा लेखाकर्म व सांख्यिकी में प्राप्त अंकों को दिखाती है। सहसम्बन्ध गुणांक ज्ञात कीजिए :-

The following table shows the marks obtained by 10 students in Accountancy and Statistics. Find the coefficient of correlation.

विद्यार्थी (Student)	:	1	2	3	4	5	6	7	8	9	10
लेखाफर्म में अंक (Marks in Accountancy)	:	45	70	65	30	90	40	50	75	85	60
सांख्यिकी में अंक (Marks in Statistics)	:	35	90	70	40	95	40	60	80	80	50

[**Ans.** $r = + 0.9$]

(14) निम्न समंकों से x तथा y के मध्य सहसम्बन्ध गुणांक ज्ञात कीजिए :-

From the following data, compute the coefficient of correlation between x and y.

	***x*-series**	***y*-series**
Number of items	15	15
Arithmetic Mean	25	18
Square of deviations from arithmetic mean	136	138
Summation of products of deviations of x any y series from their respective arithemtic means		122

[**Ans.** $r = + .89$]

(15) निम्न समंकों से सहसम्बन्ध गुणांक और इसका सम्भाव्य विभ्रम ज्ञात कीजिए :-

Calculate the correlation coefficient fromt the following data and also calculate its probable error :-

Year	No. of Building Permits	Value
1977	21	41
1978	18	34
1979	23	38
1980	34	67
1981	36	68
1982	38	84
1983	38	76
1984	36	72
1985	32	99
1986	33	67
1987	32	58

[**Ans.** $r = +.82$, P.E. = .07]

(16) निम्न समंकों से X तथा Y में सह-सम्बन्ध गुणांक ज्ञात कीजिए।

	X-श्रेणी	Y-श्रेणी
मदों की संख्या	10	10
समान्तर माध्य	107	69
माध्य से विचलनों के वर्गों का योग	1816	934

X और Y के उनके सम्बन्धित माध्यों से क्रमशः विचलनों की पारस्परिक गुणांकों का कुल योग = 1201

[**Ans.** $r = +0.83$]

From the following data, compute the coefficient of correlation between X and Y.

	X-series	Y-series
No. of items	10	10
Arithmetic Mean	107	69
Sum of squares of deviations from Mean	1816	934

Summation of product of deviations of X and Y series from their respective means = 1201

[**Ans.** $r = +0.83$]

(17) निम्न समंकों से X तथा Y में सह-सम्बन्ध ज्ञात कीजिए :-

	X-श्रेणी	Y-श्रेणी
मदों की संख्या	10	10
समान्तर माध्य	45	15
माध्य से विचलनों के वर्गों का योग	1602	296

X और Y के उनके सम्बन्धित माध्यों से क्रमशः विचलनों की पारस्परिक गुणांकों का कुल योग = 684

	X-series	Y-series
No. of items	10	10
Mean	45	15
Sum of squares of deviations from Mean	1602	396

Summation of product of deviations of *x* and *y* series = 684 [**Ans.** $r = 0.99$, P.E. = 0.023]

(18) नीचे दी गई X व Y श्रेणियों के बीच सहसम्बन्ध गुणांक की गणना कीजिए :-

Calculate coefficient of correlation between X and Y series given below :-

X	:	51	63	73	46	50	60	47	36	60
Y	:	49	72	74	44	58	66	50	30	55

[**Ans.** $r = 0.93$]

(19) समंकों से आय और व्यय के बीच सहसम्बन्ध गुणांक ज्ञात कीजिए :-

Calculate coefficient of correlation between income and expenditure from the following data :-

	Income (Rs.)				
Expenditure (Rs.)	**0-1000**	**1000-2000**	**2000-3000**	**3000-4000**	**4000-5000**
0-200	3	3	4	1	—
200-400	3	6	4	4	2
400-600	2	5	7	8	6
600-800	—	4	5	7	5
800-1000	—	3	8	6	4

[**Ans.** $r = 0.33$]

(20) नीचे दी गई श्रेणियों से 'लाभांश' व 'प्रतिभूति का मूल्य' में कार्ल पियर्सन सहसम्बन्ध गुणांक की गणना कीजिए :-

Calculate Karl Pearson's co-efficient of correlation between dividends and prices of securities given below in a bi-variate series :

प्रतिभूति का मूल्य	**वार्षिक लाभांश रूपयों में (Annual Dividends in Rs. (X))**						**योग (Total)**
(Prices of Securities) (Rs.) (Y)	**16-18**	**18-20**	**20-22**	**22-24**	**24-26**	**26-28**	
200-210	—	—	1	3	4	2	10
190-200	—	1	3	3	3	1	11
180-190	—	1	2	3	2	—	8
170-180	—	2	3	2	—	—	7
160-170	2	2	1	1	—	—	6
150-160	3	1	1	—	—	—	5
140-150	2	1	—	—	—	—	3
योग Total	**7**	**8**	**11**	**12**	**9**	**3**	**50**

[**Ans.** $r = 0.755$]

(21) एक शोधार्थी ने x तथा y के 12 जोड़ों के समंक से सहसम्बन्ध निकालने के लिए निम्नलिखित मान निकालें :-
$Sx = 30$, $SX = 5$, $Sx^2 = 670$, $Sy^2 = 285$ तथा $Sxy = 344$. बाद में ये पाया गया कि x तथा y के एक जोड़े के मान $X = 11$, $y = 4$ लिखे गये थे जबकि सही मान $x = 10$, $y = 14$ थे। सहसम्बन्ध गुणांक का सही मान ज्ञात करो।

A scholar in order to find the correlation coefficient between x and y from 12 pairs of observations, the following calculations were made : The sum of $x = 30$, $y = 5$, $x^2 = 670$, $y^2 = 285$ and $xy = 334$. On subsequent verification it was found that the pair $x = 11$ and $y = 4$ was copied wrongly, the correct value being $x = 10$ and $y = 14$. Find the correct value of correlation coefficient. [**Ans.** $r = 0.75$]

(22) विद्यार्थियों द्वारा हिन्दी व अंग्रेजी में प्राप्त क्रमान्तर में प्राप्त नीचे दिए गए हैं। सहसम्बन्ध गुणांक की गणना कीजिए:

Following are the marks obtained by 10 students in Hindi and English. Calculate the coefficient of correlation.

हिन्दी (Hindi) (R_1)	:	5.5	2.5	8	4	5.5	7	2.5	1	10	9
अंग्रेजी (English) (R_2)	:	4	6	9	2	6	8	1	3	10	6

[**Ans.** $rk = 0.76$]

(23) 10 विद्यार्थियों में लेखांकन व सांख्यिकी में प्राप्त अंकों के मध्य क्रमान्तर सहसम्बन्ध गुणांक का माप 0.5 निकाला गया। बाद में ये पाया गया कि एक विद्यार्थी द्वारा दोनों विषयों में प्राप्त क्रमांकों का अंतर गलती से 7 की जगह 3 ले लिया गया। क्रमांतर सहसम्बन्ध गुणांक का सही मान ज्ञात करो।

The rank correlation of marks obtained by ten students in statistics and accountancy was found to be 0.5. It was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 3 instead of 7. Find the correct rank correlation coefficient.

(24) एक सौन्दर्य प्रतियोगिता में 10 प्रतियोगियों को तीन निर्णायकों ने निम्न कोटियाँ प्रदान कीं :-

1st Judge (प्रथम निर्णायक)	:	1	5	6	10	2	3	4	9	8	7
2nd Judge (द्वितीय निर्णायक)	:	3	5	8	7	4	10	2	1	6	9
3rd Judge (ततीय निर्णायक)	:	6	4	9	8	1	7	5	10	3	2

क्रमान्तर सह-सम्बन्ध का प्रयोग करते हुए बताइए कि निर्णायकों के किस जोड़े की सुन्दरता के प्रति निकटतम समान रुचि है ?

Use the rank correlation to determine which pair of Judges has the nearest approach to common tastes in beauty.

[**Ans.** rk I and II Judge = + .115, rk II and IIIrd Judge = – 0.78. rk I and IIIrd Judge = + .345. So, I and IIIrd Judges have the nearst approach]

(25) सह-सम्बन्ध क्या है ? निम्न समंकों से संगामी विचलन विधि द्वारा सह-सम्बन्ध गुणांक ज्ञात कीजिए।

What is correlation ? Calculate the coefficient of correlation from the following data by the method of concurrent deviations.

Year	:	**1960**	**1960**	**1962**	**1963**	**1964**	**1965**	**1966**
Export	:	70	68	85	80	75	69	72
Import	:	22	18	25	16	15	12	20

Year	:	**1967**	**1968**	**1969**	**1970**
Export	:	76	82	77	64
Import	:	29	18	24	17

[M.D.U., B.Com. 1985] [**Ans.** $rc = + .755$]

(26) निम्नलिखित से संगामी विचलन गुणांक का परिकलन कीजिए :

Calculate the coefficient of concurrent deviations from the following :

वर्ष	**1973**	**1974**	**1975**	**1976**	**1977**	**1978**	**1979**	**1980**
आपूर्ति	160	164	172	182	166	170	178	192
कीमत	222	280	260	224	266	254	230	190

[B.Com., K.U., 1998] [**Ans.** $rc = -.845$]

❑❑

अध्याय - 9

प्रतीपगमन विश्लेषण
(Regression Analysis)

पिछले अध्याय में दो चरों के बीच सहसम्बन्ध ज्ञात करने की विधियों का वर्णन किया गया था। सहसम्बन्ध से चरों का कारण व परिणाम में विभाजन स्पष्ट नहीं होता। दूसरे एक चर के किसी ज्ञात मूल्य से दूसरे चर के अज्ञात मूल्य का अनुमान लगाना भी संभव नहीं है। प्रतीपगमन विश्लेषण इन दोनों तथ्यों को बताने वाली सांख्यिकी की एक महत्वपूर्ण तकनीक है।

'प्रतीपगमन' (Regression) शब्द का अर्थ पीछे लौटना अथवा पीछे हटना (to regress or to step back) है। उन्नीसवीं शताब्दी के अन्त में, पिताओं तथा पुत्रों की ऊँचाइयों के सम्बन्ध का अध्ययन करने में सर फ्रांन्सिस गाल्टन (Sir Francis Galton) ने सांख्यिकीय विज्ञान में प्रतीपगमन तकनीक का विकास किया था। उन्होंने इस शब्द का प्रयोग अपने शोध-लेख 'पैतक ऊँचाई में मध्यता की ओर प्रतीपगमन' (Regression Towards Mediocrity in Heriditary Stature) में किया था। अपने उक्त शोध-लेख में उन्होंने यह स्पष्ट किया था कि सामान्यतः लम्बे पिताओं के पुत्र भी लम्बे होते हैं, परन्तु पुत्रों में पुरुषों के समग्र की माध्य ऊंचाई (Average Male Height) की ओर वापस आने की प्रवत्ति पायी जाती है। अन्य शब्दों में, पिता -पुत्रों की ऊंचाइयों में घनिष्ट सहसम्बन्ध था परन्तु सामान्य-माध्य से दोनों के विचलनों में पर्याप्त अन्तर था। समस्त पुरुष जाति की माध्य ऊंचाई से पिताओं की ऊंचाई के विचलनों की तुलना में पुत्रों की ऊंचाई के विचलन कम थे। यदि पिताओं की माध्य ऊंचाई समग्र की माध्य ऊंचाई से एक सेंटीमीटर से कम थी तो उनके पुत्रों की माध्य ऊंचाई समग्र की माध्य ऊंचाई से एक सेंटीमीटर अधिक थी। गाल्टन ने लगभग एक सहस्त्र पिताओं तथा पुत्रों की माध्य ऊंचाइयों का अन्वेषण करके इस महत्वपूर्ण प्रवत्ति को ज्ञात किया था। लम्बे पिताओं के पुत्र भी लम्बे होते हैं तथा छोटे कद के पिताओं के पुत्र भी छोटे कद के होते हैं, परन्तु लम्बे पिताओं के पुत्रों की औसत ऊंचाई उनके पिताओं की ऊंचाई की औसत से कम होती है, तथा छोटे कद के पिताओं के पुत्रों की औसत ऊंचाई उनके पिताओं की औसत से अधिक होती है।'' अन्य शब्दों में, पिताओं की ऊंचाई समग्र की सामान्य ऊंचाई से अधिक या कम होती थी परन्तु पुत्रों की ऊंचाई समग्र की ऊंचाई के पर्याप्त निकट होती जाने की प्रवत्ति पायी जाती है। पुत्रों की ऊंचाई का सामान्य माध्य के निकट वापस आने की प्रवत्ति को गाल्टन ने 'प्रतीपगमन' की संज्ञा दी है।

सांख्यिकी विज्ञान में 'प्रतीपगमन' का प्रयोग उन समस्त क्षेत्रों में किया जाता है, जिनमें दो या अधिक सम्बन्धित श्रेणियों के विभिन्न चर-मूल्यों में सामान्य माध्य की ओर वापस जाने की प्रवत्ति पायी जाती हो। ''जबकि सहसम्बन्ध विश्लेषण दो या अधिक घटनाओं के सहपरिवर्तन की घनिष्टता की जाँच करता है, प्रतीपगमन विश्लेषण इस सम्बन्ध की प्रकृति एवं मात्रा की माप करके हमें भावी अनुमान लगाने की क्षमता प्रदान करता है''। औसत सम्बन्ध के आधार पर पूर्वानुमान करने में प्रतीपगमन तकनीक विशेष उपयोगी है। प्रतीपगमन की सहायता से एक श्रेणी में एक निश्चित मात्रा में परिवर्तन होने पर दूसरी श्रेणी में होने वाले संभावित औसत परिवर्तन को ज्ञात किया जा सकता है। वालिस एवं रॉबर्टस् (Wallis Roberts) के शब्दों में, प्रायः यह ज्ञात करना अधिक महत्वपूर्ण होता है कि (दो चर मूल्यों में) वास्तविक सम्बन्ध क्या है जिससे एक चर मूल्य (स्वतंत्र चर मूल्य) ज्ञात होने पर दूसरे चर मूल्य (आश्रित चर मूल्य) का पूर्वानुमान किया जा सके; ऐसी दशा में प्रयोग की जाने वाली उपयुक्त तकनीक प्रतीपगमन विश्लेषण कहलाती है''। आर्थिक एवं व्यापारिक जगत में प्रतीपगमन की अत्यन्त व्यावहारिक उपयोगिता है। इस तकनीक के प्रयोग से अनेक प्रकार के तथ्यों के पूर्वानुमान किये जा सकते हैं। उदाहरणार्थ, किसी वस्तु के उत्पादन में निश्चित परिवर्तन होने पर उसके मूल्य में क्या संभावित परिवर्तन होगा, सामान्य मूल्य स्तर में निश्चित वद्धि होने पर जीवन-निर्वाह व्यय कितना बढ़ जाएगा, वर्षा की मात्रा में निश्चित वद्धि अथवा कमी होने पर उत्पादन में कितनी वद्धि अथवा कमी होगी आदि।

रेखीय प्रतीपगमन
(Linear Regression)

दो चर मूल्यों के मध्य रेखीय सम्बन्ध (Linear Relationship) उस दशा में होता है जब चर मूल्य (स्वतंत्र) में एक इकाई से परिवर्तन होने पर दूसरे चर मूल्य (आश्रित) में एक निश्चित परिणाम में परिवर्तन (Constant Absolute Change) हो। दो चर-मूल्यों में रेखीय सम्बन्ध होने की दशा में स्वतंत्र चर-मूल्यों के दिये गये मान के आधार पर आश्रित चर मूल्य का मान ज्ञात करने के लिए प्रतीपगमन रेखाओं का प्रयोग किया जाता है। दो चर-मूल्यों (x तथा y) को बिन्दुरेखीय पत्र पर प्रांकित करने पर एक विक्षेप चित्र बन जाता है। प्रांकित बिन्दुओं के मध्य से गुजरती हुई दो 'सर्वोचित्त रेखाएं' (Lines of the Best Fit) खींची जाती हैं। ये रेखाएँ ही प्रतीपगमन रेखाएं कहलाती हैं। रेखीय प्रतीपगमन में ये रेखाएँ सीधी (Straight)) होती हैं। प्रतीपगमन रेखाओं के आधार पर एक घातीय दो समीकरण होते हैं। इन समीकरणों को x on y तथा 'y' on 'x' कहते हैं। 'X on Y' समीकरण से 'y' के दिये गये मूल्य के लिए 'x' का सर्वाधिक सम्भव मूल्य ज्ञात किया जा सकता है। यह समीकरण $y = a + bx$ होता है। 'y on x' समीकरण से 'X' के दिये मूल्ये के लिये 'y' का सर्वाधिक सम्भव मूल्य ज्ञात किया जा सकता है। यह समीकरण $x = a + by$ होता है।

प्रतीपगमन रेखाएँ
(Regression Lines)

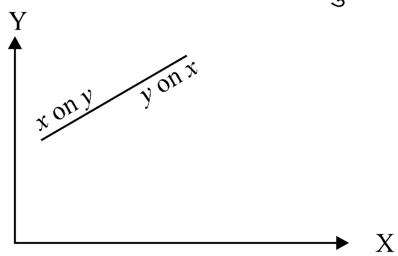
''एक चर-मूल्य से दूसरे चर-मूल्य का पूर्वानुमान करने में प्रयुक्त विधि में एक रेखा, जो प्रांकित बिन्दुओं के मध्य इस प्रकार खींची जाती है दोनो चर-मूल्यों के औसत सम्बन्ध का प्रदर्शन करें, का प्रयोग किया जाता है। इस प्रकार की रेखा को प्रतीपगमन रेखा कहते हैं ''प्रतीपगमन रेखाएँ दो होती हैं जो दो श्रेणियों के पारस्परिक माध्य सम्बन्ध को व्यक्त करती हैं। परन्तु यदि दोनों श्रेणियों के मध्यपूर्ण सहसम्बन्ध (Perfect Correlation, or $r = \pm1$) होता है तो दोनो रेखाएँ एक-दूसरे को ढक लेती हैं, अर्थात् दोनों रेखाएँ एक रेखा का स्वरूप प्राप्त कर लेती हैं। दोनों प्रतीपगमन रेखाएँ एक दूसरे को x तथा y श्रेणियों के माध्यों पर एक-दूसरे को काटती हैं। ये रेखाएँ एक-दूसरे के जितने निकट होंगी, x तथा y श्रेणियों के मध्य सहसम्बन्ध की मात्रा भी उतनी ही अधिक होगी।

प्रतीपगमन की एक रेखा x का y पर प्रतीपगमन (Regression of x and y) व्यक्त करती है। इस रेखा की रचना y को स्वतंत्र चर-मूल्य (Independent Variable) तथा x को आश्रित चर-मूल्य (Dependent Variable) मान कर की जाती है। दोनों चर-मूल्यों के माध्य सम्बन्ध के आधार पर इस रेखा से 'y' को दिये गये मूल्य को समकक्ष 'x' का सर्वोचित मूल्य ज्ञात किया जाता है। प्रतीपगमन की दूसरी रेखा y का x पर प्रतीपगमन (Regression of y on x) व्यक्त करती है। इस रेखा की रचना में x का स्वतंत्र चर-मूल्य तथा y को आश्रित चर-मूल्य माना जाता है। ये रेखा 'x' को दिये गये मूल्य को समकक्ष 'y' को सर्वोचित मूल्य ज्ञात करने में योग देती है।

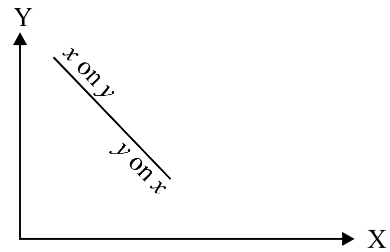
प्रतीपगमन रेखाएँ व सहसम्बन्ध गुणांक
(Regression Lines and Co-efficient of Correlation)

प्रतीपगमन रेखाओं की सहायता से हम सहसम्बन्ध का अनुमान लगा सकते हैं। यदि दो श्रेणियों के मध्य धनात्मक सहसम्बन्ध होता है तो दोनों प्रतीपगमन रेखाएँ रेखाचित्र पर बायें निचले कोने से दाहिने ऊपर के कोने की ओर बढ़ती हैं। ऋणात्मक सहसम्बन्ध होने की दशा में ये रेखाएँ ऊपर से नीचे की ओर आती हैं। दोनों श्रेणियों के मध्यपूर्ण सहसम्बन्ध होने की दशा में दोनों रेखाएँ पूरी तरह से एक दूसरे पर ढक लेती हैं, अर्थात् पूर्ण सहसम्बन्ध होने पर केवल एक ही प्रतीपगमन रेखा बनती है। यदि दो श्रेणियों में सहसम्बन्ध का अभाव होता है, तो दोनो प्रतीपगमन रेखाएँ एक दूसरे को समकोण (Right Angle) पर काटती हैं। दो श्रेणियों से सीमित सहसम्बन्ध होने की दशा में, दोनों प्रतीपगमन रेखाओं की दूरी सहसम्बन्ध की मात्रा को इंगित करती हैं। ये रेखाएँ एक दूसरे के जितने निकट होंगी, दोनों श्रेणियों में सहसम्बन्ध उतना ही अधिक होगा। ये रेखाएँ एक-दूसरे से जितनी अलग होती जायेंगी, सहसम्बन्ध की मात्रा उतनी ही कम होती जायेगी।

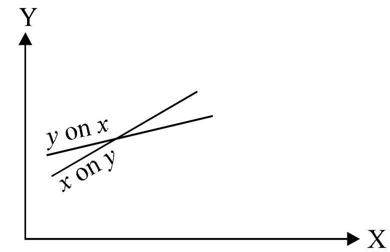
उपरोक्त स्थितियों को हम बिन्दुरेखीय चित्रण से भी दिखा सकते हैं, जैसा कि नीचे दिखाया गया है।



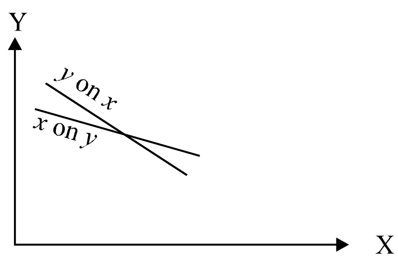
(*i*) Perfect Positive Correlation



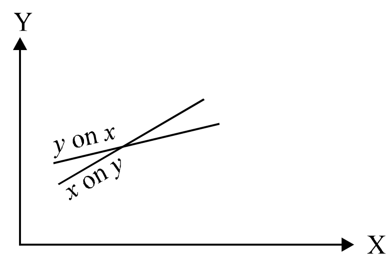
(*ii*) Perfect Negative Correlation



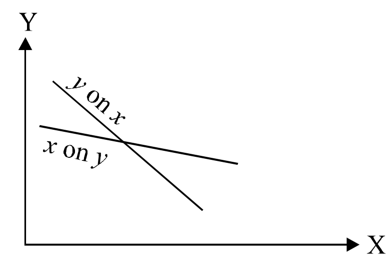
(*iii*) High Degree of Positive Correlation



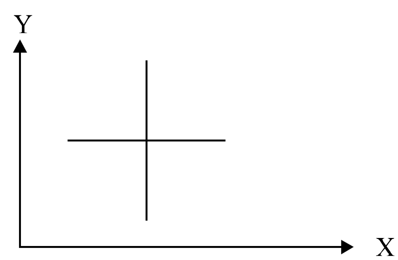
(*iv*) High Degree of Negative Correlation



(*v*) Low Degree of Positive Correlation



(*vi*) Low Degree of Negative Correlation



(*vii*) No Correlation

प्रतीपगमन समीकरण
(Regression Equation)

प्रतीपगमन समीकरण, प्रतीपगमन को व्यक्त करते हैं। प्रतीपगमन रेखाएँ दो होती है, अतः प्रतीपगमन समीकरण भी दो होते हैं :-

x on *y* Regression Equation : $x = a + by$

Regression Equation : $y = a + bx$

उक्त समीकरणों में x तथा y चर-मूल्य हैं तथा 'a' और 'b' अचर (Constant) हैं। अचर 'a' अन्तःखण्ड (Intercept) होता है, अर्थात् वह बिन्दु होता है जहाँ प्रतीपगमन रेखा कोटि अक्ष (y-axis) को स्पर्श करती है। अन्य शब्दों में मूल बिन्दु (Point of Origin) से प्रतीपगमन रेखा द्वारा कोटि-अक्ष को स्पर्श करने वाले बिन्दु की दूरी होती है। 'a' धनात्मक होने पर प्रतीपगमन रेखा से ऊपर तथा ऋणात्मक होने पर मूल बिन्दु से नीचे कोटि अक्ष को स्पर्श करती है। पूर्व में दिये गये चित्र में इसे स्पष्ट किया गया है। $x = a + by$ समीकरण में $a = x - by$ तथा $y = a + bx$ समीकरण में $x = y - bx$ होता है। (x तथा y सम्बन्धित श्रेणियों के माध्यमों का प्रतिनिधित्व करते हैं)।

अचर 'b' प्रतीपगमन रेखा का ढाल (Slope of the Line) प्रदर्शित करता है। (The Constant 'b' is the value of the tangent to the angle formed between the line and any horizontal drawn from it) इसे प्रतीपगमन गुणांक भी कहते हैं। इससे यह ज्ञात होता है कि 'x' में एक इकाई का परिवर्तन होने पर 'y' में कितना परिवर्तन होगा। यदि 'b' का मूल्य धनात्मक होता है तो प्रतीपगमन रेखा का ढलान बायें से दायें ऊपर की ओर होगा। 'b' का मूल्य ऋणात्मक होने पर रेखा नीचे की ओर होगी। प्रतीपगमन समीकरणों को निम्नलिखित विधियों से हल किया जाता है :-

(1) समान्य समीकरण (Normal Equation)

(2) x तथा y के अंकगणितीय माध्य से विचलन (Deviation taken form arithmetic mean of $x + y$)

(3) x तथा y के काल्पनिक माध्य से विचलन (Deviation taken from assumed means of x and y)

(1) सामान्य समीकरण (Normal Equation)

(A) x का y पर प्रतीपगमन समीकरण

$$x = a + by$$

a तथा b के मान निकालने के लिये दो समीकरण बनाये जाएंगे

$$\Sigma x = Na + b\Sigma y \qquad ...(1)$$

$$\Sigma xy = a\Sigma y + b\Sigma y^2 \qquad ...(2)$$

प्रथम समीकरण को Σy से गुणा करें व द्वितीय समीकरण को N से गुणा करने के बाद उसे प्रथम से घटाये तो हमारे पास b का मान आ जाता है।

$$\Sigma x\Sigma y = Na\Sigma y + b\,(\Sigma y)^2$$

$$N\,\Sigma xy = Na\Sigma y + bN\,\Sigma y^2$$

or
$$\frac{x \;\; y \quad N \;\; xy \quad b\left[\left(\;\; y^2\right) \quad N \;\; y^2\right]}{b\left[N \;\; y^2 \quad \left(\;\; y^2\right)\right] \quad N \;\; xy \quad\; x \;\; y}$$

or
$$b_{xy} = \frac{N \;\; xy \quad\; x \;\; y}{N \;\; Y^2 \quad (\;\; y)^2} \qquad ...(3)$$

(B) y का x पर प्रतीपगमन समीकरण

$$y = a + bx$$

इसमें निम्नलिखित दो समीकरण बनेंगे

$$\Sigma y = Na + b\,\Sigma x$$

$$\Sigma xy = a\Sigma x + b\,\Sigma x^2$$

ऊपर वाली विधि अपनाकर हम b_{yx} का मान निकाल सकते हैं।

$$b_{yx} = \frac{N \;\; xy \quad\; x \;\; y}{N \;\; x^2 \quad (\;\; x)^2}$$

a का मान हम किसी एक समीकरण में b का मान रखकर निकाल सकते हैं।

Illustration 1 : Obtain two regression equations for the following data :-

x	8	9	10	11	12	13	14	15	16	17
y	7	7	8	8	8	9	9	10	11	11

Solution :

x	y	x^2	y^2	xy
8	7	64	49	56
9	7	81	49	63
10	8	100	64	80
11	8	121	64	88
12	8	144	64	96
13	9	169	81	117
14	9	196	81	126
15	10	225	100	150
16	11	256	121	176
17	11	289	121	187
Sx = 125	**Sy = 88**	**Sx^2 = 1645**	**Sy^2 = 794**	**Sxy = 1139**

(*i*) Equation of x and y

$$x = a + by$$

Two normal equations are

$$\mathrm{S}x = \mathrm{N}a + b\mathrm{S}y$$

$$\mathrm{S}xy = a\mathrm{S}y + b\mathrm{S}y^2$$

$$b_{xy} = \frac{\mathrm{N} \;\; xy \quad\; x \;\; y}{\mathrm{N} \;\; y^2 \quad (\;\; y)^2} = \frac{10 \quad 1139 \quad 125 \quad 88}{10 \quad 794 \quad (88)^2} = \frac{390}{196} = 1.99$$

Now

$$\mathrm{S}x = \mathrm{N}a + b\mathrm{S}y$$

$$125 = 10a + 1.99 \times 88$$

$$= 10a + 175$$

\ $$10a = -50$$

or $$a = -5$$

So equation of x any y is $\boxed{x = -5 + 1.99\ \mathrm{Y}}$

(*ii*) Equation of y on x

Two normal equations are :

$$\mathrm{S}y = \mathrm{N}a + b\mathrm{S}x$$

$$\mathrm{S}xy = a\mathrm{S}x + b\,(\mathrm{S}x)^2$$

by $$x = \frac{\mathrm{N} \quad \mathrm{Y} \quad\; \mathrm{X.} \;\; \mathrm{Y}}{\mathrm{N} \;\; \mathrm{X}^2 \quad (\;\; \mathrm{X})^2} = \frac{10 \quad 1139 \quad 125 \quad 88}{10 \quad 16450 \quad (125)^2}$$

Now $$\mathrm{S}y = \mathrm{N}a + b\mathrm{S}x$$

$$88 = 10a + 0.47 \times 125$$

$$10a = 29.25$$

or $$a = 2.925$$

\ Equation of y on x is $\boxed{y = 2.925 + 0.47\,x}$

Illustration 2. Obtain two regression equations from the data obtained below :-

x	6	3	9	4	8
y	9	8	5	8	6

Solution :

x	y	x^2	y^2	xy
6	9	36	81	54
3	8	9	64	24
9	5	81	25	45
4	8	16	64	32
8	6	64	36	48
30	**36**	**206**	**270**	**203**

(*i*) Two normal equations to find regression line of *x* and *y*

$$Sx = na + bSy$$

$$Sxy = aSy + bSy^2$$

Substituting the values in the normal equations,

$$30 = 5a + 36b$$

$$203 = 36a + 270b$$

Multiply equation (3) by 7.2

$$216 = 36a + 259.2b$$

Sub : $$203 = 36a + 270.0b$$

$$\frac{59.88}{5}$$

$$13 = -10.8\,b$$

$$b = \frac{10.8}{13}$$

$$= (-)\ 0.83$$

Interpret the value of *b* in equation (3)

$$30 = 5a + 36\,(-0.83)$$

$$30 = 5a - 29.88$$

$$-5a = -29.88 - 30.00$$

$$5a = 59.88$$

$$a =$$

$$= 11.976$$

Putting the values of *a* and *b*, the form of regression equation of *x* on *y* is :

$$x = 11.976 - 0.83\,y$$

(*ii*) Two normal equations to find out regression line of *y* on *x* is :

$$Sy = na + bSx$$

$$Sxy = aSx + bSx^2$$

Substituting the values of equations (7) and (8)

$$36 = 5a + 30b$$

$$203 = 30a + 206b$$

Multiplying equation (9) by 6

$$216 = 30a + 180b$$

Sub :

$$216 = 30a + 206b$$

$$13 = -26b$$

$$b = \quad = -2$$

Interpret the value of b in equation (9)

$$36 = 5a + 30(-2)$$

$$36 = 5a - 60$$

$$-5a = -66 - 36$$

$$a = \frac{96}{5} = 19.2$$

Putting the values of a and b, the regression equation of y on x is :

$$y = 19.2 - 2.0x$$

(2) x तथा y के अंकगणितीय माध्यों से विचलन

(Deviations taken from Arithmetic Means of x and y)

प्रथम विधि में हमने $x = a + by$ तथा $y = a + bx$ में a व b के मान निकालने के लिए दो-दो समीकरण बनाये थे। इनमें से एक-एक समीकरण लेकर हम यह दूसरी विधि प्रयोग करेंगे :-

(i) Equation of x any y

$\frac{\bar{y}\ 26\ xy}{13\ y^2}$

$$x = a + by \qquad ...(i)$$

इसके लिये एक सामान्य समीकरण ये था।

$$Sx = Na + bSy$$

दोनों पक्षों को N से भाग देने पर नया समीकरण यह है।

$$\frac{N}{N} = a + b.\frac{Y}{N}$$

or

$$\bar{x} = a + b \qquad ...(2)$$

समीकरण (1) में से (2) घटाएं

$$x - \quad = b(y - \quad) \qquad ...(3)$$

यह हमारी दूसरी विधि के समीकरण का सूत्र है। इसमें हमें b यानि b_{xy} का मान निकालना है। इस मान को उपरोक्त समीकरण में रखकर हमें x or y का समीकरण मिलता है। जैसा कि हम जानते हैं।

$$x - \quad = x \text{ तथा } y - \quad = y \qquad ...(3)$$

इस तरह समीकरण का परिवर्तित रूप है।

$$x = bxy.y$$

or $xy = bxy.y^2$ (दोनों पक्षों को y से गुणा करने पर)

or $Sxy = bxy\ Sy^2$

इस तरह $bxy =$

इस तरह से y on x का समीकरण के लिये

$$y = a + bx$$
$$\bar{y} = a + b$$
$$y - \quad = b\,(x - \quad) \qquad ...(4)$$

and $\quad b_{yx} =$

byx के इस मान को समीकरण (4) में रखने पर हमें y on x का समीकरण प्राप्त होता है।

Illustration 3 : Obtain two regression equations for the following data relating to sales (in crore Rs.) and advertisement expenditure (in lakh Rs.) of a company.

Sales (x)	**:**	10	11	13	15	16	19	14
Adv. Exp. (y)	**:**	60	62	65	70	73	75	71

Solution :

x	x $(x - 14)$	x^2	y	y $(y - 68)$	y^2	xy
10	– 4	16	60	– 8	64	32
11	– 3	9	62	– 6	36	18
13	– 1	1	65	– 3	9	3
15	1	1	70	2	4	2
16	2	4	73	5	25	10
19	5	25	75	7	49	35
14	0	0	71	3	9	0
98	**0**	**56**	**476**	**0**	**196**	**100**

$$\bar{x} = 98/7 = 14$$
$$= 476/7 = 68$$

(i) Regression equation of x and y $\quad bxy = \quad = \frac{100}{196} = 0.51$

$$x - 14 = 0.51\,(y - 68)$$
$$x - 14 = 0.51\,y - 34.68$$
$$x = 0.51\,y - 34.68 + 14.0$$

or $\quad x = -\,20.68 + 0.51\,y$

(ii) Regression equation of y and x

$$byx = \frac{xy}{x^2} = \frac{100}{56}$$
$$= 1.7857$$
$$y - 68 = 1.7857\,(x - 14)$$
$$y - 68 = 1.7857\,x \;\; 24.9998$$
$$y = 1.7857\,x - 24.9998 + 68.0$$

or $\quad y = 43.002 + 1.7857\,x$

(3) काल्पिनिक माध्यों से विचलन

(Deviations taken from assumed means of *x* and *y*)

ऐसी स्थिति में जबकि x तथा y के अंकगणितीय माध्य पूर्णांक ना हो तो हम इस विधि का प्रयोग करते हैं। यदि x सारणी का काल्पनिक माध्य Ax है व y का Ay तो हम विचलन dx व dy निकालते हैं।

$$dx = x - \mathrm{A}x$$

$$dy = y - \mathrm{A}y$$

इसके बाद हम प्रतीपगमन गुणांकों का मान निम्नलिखित सूत्रों से निकालते हैं।

$$bxy = \frac{\mathrm{N}\ \ dxdy \qquad dx\ \ dy}{\mathrm{N}\ \ dy^2 \quad (\ \ dy)^2}$$

और

$$byx = \frac{\mathrm{N}\ \ dxdy \qquad dxdy}{\mathrm{N}\ \ dx^2 \quad (\ \ dx)^2}$$

x on y के समीकरण को प्राप्त करने का सूत्र है

$$x - \bar{x} = bxy\,(y - \quad)$$

तथा y on x का समीकरण P प्राप्त करने का सूत्र है

$$y - \quad = byx\,(x - \quad)$$

Illustration 4 : Obtain the two regression equation for the following data relating to income (Rs.) and expenditure on foods (Rs.) of 10 families.

$\bar{y}$

Income (Rs.)	120	90	80	150	130	140	110	95	75	105
Expenditure of Food (Rs.)	40	36	40	45	40	44	45	38	50	35

Solution : Let us denote income by x and food expenditure by y. Further let Ax = 109 and Ay = 41. So $dx = x - 109$ and $dy = y - 41$.

x	dx	dx^2	y	dy	dy^2	$dx.dy$
120	11	121	40	– 1	1	– 11
90	– 19	361	36	– 5	25	95
80	– 29	841	40	– 1	1	29
150	41	1681	45	4	16	164
130	21	441	40	– 1	1	– 21
140	31	961	44	3	9	93
110	1	1	45	1	16	4
95	– 14	196	38	– 3	9	42
75	– 34	1156	50	9	81	– 306
105	– 4	16	35	– 6	36	24
1095	**5**	**5775**	**413**	**3**	**195**	**113**

$$= \quad = 109.50 \qquad \bar{y} = \quad = 41.30$$

$$bxy = \frac{10 \quad 113 \quad 5 \quad 3}{10 \quad 195 \quad 9} \qquad byx = \frac{10 \quad 113 \quad 5 \quad 3}{10 \quad 5775 \quad (5)^2}$$

$$= \frac{1130 \quad 15}{1950 \quad 9} \qquad = \frac{1130 \quad 15}{57751 \quad 25}$$

$$= \frac{1115}{1941} = 0.57 \qquad = \frac{1115}{57725} = 0.019$$

(*i*) Regression equation of *x* and *y*

$$x - \bar{x} = bxy\,(y - \quad)$$
$$x - 109.50 = 0.57\,(y - 41.30)$$
$$= 0.574 - 23.72$$
$$\boxed{x = 85.78 + 0.574}$$

(*ii*) Regression equation of *y* and *x*

$$y - \quad = bxy\,(x - \quad)$$
$$y - 41.30 = 0.019\,(x - 109.50)$$
$$= 0.019\,x - 2.11$$
$$\boxed{y = 39.19 + 0.019\,x}$$

द्विचर वर्गित आवत्ति बंटन में प्रतीपगमन समीकरण

(Regression Equations for Grouped Frequency Distribution)

द्विचर वर्गित आवत्ति बंटन में प्रतीपगमन गुणक ज्ञात करने के लिए किसी प्रकार की तालिका बनायी जाएगी जिस प्रकार ऐसी श्रेणी में सहसम्बन्ध गुणांक ज्ञात करने के लिए बनायी जाती है। यह उग्र उदाहरण में स्पष्ट किया गया है।

Illustration 5 : In the following table, A's stand respectively for marks obtained in Economics in intervals of 4-8, 8-12, 12-16 and 16-20 and B's stand for marks obtained in Statistics in intervals of 8-14, 14-20 and 20-26. Calculate the equation to the two lines of regression.

	A_1	A_2	A_3	A_4	**Total**
B_1	11	6	2	1	20
B_2	5	12	15	8	40
B_3	—	2	3	15	20
Total	16	20	20	24	80

Solution :

Marks in Statistics *y* *m.v.*		**Marks in Economics** *x* 4-8 (6)	8-12 (10)	12-16 (14)	16-20 (18)	**Total** *fy*	*d*…
8-14	11	2 / 11 / +22	1 / 6 / +6	2 / 0	–1 / 1 / –1	20	–
14-20	17	5 / 0	12 / 0	15 / 0	8 / 0	40	
20-26	23	—	–1 / 2 / 2	3 / 0	1 / 15 / 15	20	+
Total *fx*		16	20	20	24	80 *n*	
dx/(i_x) = 4/*i*		–2	–1	0	+1		
fdx		–32	–20	0	+24	–28	Σ…
fd^2x		64	20	0	24	108	Σ…
fdxdy		22	4	0	14	40	Σ…

$$\text{Average of } x = A_x + \frac{\Sigma fdx}{N} \times i_x = 14 + \frac{-28}{80} \times 4 = 14 - 1.4 = 12.6$$

$$\text{Average of } y = A_y + \frac{\Sigma fdy}{N} \times i_y = 17 + \frac{0}{80} \times 6 = 17$$

Regression Coefficients :

x on *y*

$$b_{xy} = \frac{\left[\Sigma fd_x d_y - \frac{\Sigma fd_x \Sigma fd_y}{N}\right] i_x}{\left[\Sigma fd_y^2 - \frac{(\Sigma fd_y)^2}{N}\right] i_y}$$

$$= \frac{\left[40 - \frac{(-28)(0)}{80}\right] 4}{\left[40 - \frac{(0)^2}{80}\right] 6}$$

$$= \frac{160}{240} = 0.67$$

x on *y*

$$b_{yx} = \frac{\left[\Sigma fd_2 d_y - \frac{\Sigma fd_x \Sigma fd_y}{N}\right] i_y}{\left[\Sigma fd_y^2 - \frac{(\Sigma fd_2)^2}{N}\right] i_x}$$

$$= \frac{\left[40 - \frac{(-28)(0)}{80}\right] 6}{\left[108 - \frac{(-28)^2}{80}\right] 4}$$

$$= \frac{240}{392.8} = 0.611$$

प्रतीपगमन गुणाकों, सहसम्बन्ध गुणांक व प्रमाप विचलन के मध्य सम्बन्ध

(Relationship Between Regression Co-efficients, Correlation Co-effcient and Standard Deviations)

जब हम अंकगणितीय माध्यों से विचलन निकालने की विधि का प्रयोग करते है, तो सहसम्बन्ध गुणांक का सूत्र होता है।

$$r = \frac{xy}{\sqrt{x^2 . \; y^2}}$$

जब हम इसी विधि से प्रतीपगमन गुणांक के मान निकालते हैं तो निम्नलिखित सूत्र प्रयोग करते हैं :-

$$bxy = \frac{xy}{y^2}, \quad byx = \frac{xy}{x^2}$$

Now

$$bxy \times byx = \frac{xy}{y^2} \times \frac{xy}{x^2}$$

$$= \frac{(\;xy)^2}{\left(\sqrt{\;y^2}\right)^2}, \left(\sqrt{\;x^2}\right)^2$$

$$= \left[\frac{xy}{\left(\sqrt{\;x^2 . \; y^2}\right)}\right]^2$$

$$= r^2$$

इस प्रकार दो प्रतीपगमन गुणांकों को गुणनफल सहसम्बन्ध गुणांक के वर्ग के बराबर होता है।

आगे

$$bxy = \frac{xy}{y^2} = \frac{xy}{\sqrt{\;y^2} . \sqrt{\;y^2}} \times \frac{\sqrt{\;x^2}}{\sqrt{\;x^2}}$$

$$= \frac{xy}{\sqrt{\;y^2} . \sqrt{\;x^2}} \quad \frac{\sqrt{\;x^2}}{\sqrt{\;y^2}}$$

$$= r . \frac{\sqrt{\;x^2}/N}{\sqrt{\;y^2}/N}$$

$$= r. \frac{x}{y}$$

इसी तरह से हम ये सिद्ध कर सकते हैं कि

$$byx = r. \frac{x}{y}$$

उपरोक्त सम्बन्धों से प्रतीपगमन गुणांकों के विषय में कुछ महत्वपूर्ण निष्कर्ष निकलते हैं जो कि निम्नलिखित हैं:-

(1) क्योंकि r का मान -1 से $+1$ के बीचक में होता है तो r^2 का अधिकतम मान 1 होगा। इसलिए दोनों प्रतीपगमन गुणांकों के गुणनफल का मान कभी भी 1 से ज्यादा नहीं हो सकता।

(2) तीनों गुणांकों (r, bxy, byx) के चिन्ह हमेशा एक समान होंगे। या तो तीनों धनात्मक होंगे या तीनों ऋणात्मक होंगे। उदाहरण के लिए, यदि $bxy = -0.8$, $byx = -0.45$ हों, तो r का मान $\sqrt{0.8 \quad 0.45}$ होगा न कि 0.6.

(3) सहसम्बन्ध गुणांक दो प्रतीपगमन गुणांकों का गुणात्तर माध्य (Geometrical Mean) होता है।

(4) यदि $s_x = s_y$ हो, तो गुणांकों का मान बराबर होगा, अर्थात् $r = bxy = byx$

(5) अवधारणा का गुणांक (r^2) दोनों प्रतीपगमन गुणांकों के गुणनफल के बराबर होता है।

(6) यदि $byx = bxy$ हो, तो r का मान भी उनके बराबर होगा।

Illustration 6 : The following equations were obtained in a certain investigation :

$$x = 19.13 - 0.87y$$
$$y = 11.64 - 1.50x$$

Find the mean of x and y and correlation coefficient.

Solution :

$$x = 19.13 - 0.87\,y$$

or $$x + .87\,y = 19.13 \qquad ...(1)$$

$$y = 11.64 - 5x$$

or $$.5x + y = 11.64 \qquad ...(2)$$

Multiplying the equation No. (2) by 2

$$x + 2y = 23.28 \qquad ...(3)$$
$$x + 87y = 19.13 \qquad ...(1)$$

(deduct) $$1.13y = 4.15 \quad \sqrt[y]{bxy \quad byx}$$

or $$y = 3.67$$

or $$\bar{y} = 3.67$$

Substituting the value of y in equation no. (3)

$$x + 2y = 23.28$$

or $$x + 2 \times 3.67 = 23.28$$

or $$x + 7.34 = 23.28$$

or $$x = 23.28 - 7.34 = 15.94 \text{ or } x = 15.94$$

$$x = -.87y + 19.13$$

or $$bxy = -.87$$

$$y = -.5x + 11.64$$

or $$bxy = -.5$$

$$r = \quad = \sqrt{.87 \quad .5} = \sqrt{.435} = 0.66$$

Illustration 7 : Two lines of regression are given by

$$\mathbf{x + 2y - 5 = 0}$$

and $$\mathbf{2x + 3y - 8 = 0}$$

and $$^{2}_{x} = \mathbf{12}$$

Calculate the value of $\bar{x}$, , sy and r.

Solution :

$$x + 2y - 5 = 0 \quad ...(1)$$

or $$x + 2y = 5 \quad ...(2)$$

$$2x + 3y - 8 = 0$$

or $$2x + 3y = 8$$

Multiplying equation no. 1 and 2

$$\begin{aligned} 2x + 4y &= 10 \\ 2x + 3y &= 8 \end{aligned}$$

(deduct) $y = 2$ or $y = 2$

Substituting the value of y in equation no. 1

$x + 2y = 5$ or $x + 4 = 5$

or $x = 5 - 4 = 1$ or $= 1$

Assuming that 1st equation stands for y on x and 2nd equation stands for x and y.

then

$$\begin{aligned} 2x + 4y &= 10 \\ 4y &= -2x + 10 \\ y &= -.5x + 2.5 \end{aligned}$$

or $byx = -.5$

and

$$\begin{aligned} 2x + 3y &= 8 \\ 2x &= -3y + 8 \\ x &= -1.5 + 4 \\ bxy &= -1.5 \end{aligned}$$

For calculating s_y

$$bxy = r \quad = 1.5 = .86 \times \frac{3.46}{y} = 1.5$$

$$\begin{aligned} 1.5\, s_y &= 3 \\ s_y &= 2 \end{aligned}$$

$$r = \sqrt{bxy \quad byx} \quad \sqrt{1.5 \quad .5}$$

$$r = -0.86$$

Illustration 8 : A study by the Commerce Department in the effect of bus ticket prices upon the number of passenger produced the following results :-

Ticket price increase (paise)	:	15	20	25	30	40	50
Passenger per 100 kms.	:	440	430	450	370	340	370

Predict the number of passenger per km if the ticket price is 35 paise increase. And estimate the ticket price if the number of passenger per km is 400.

Solution : Let X denote the ticket price and Y the number of passenger per km.

X	x	x^2	Y	y	y^2	xy
15	– 15	225	440	40	1600	– 600
20	– 10	100	430	30	900	– 300
25	– 5	25	450	50	2500	– 250
30	00	0	370	– 30	900	0
40	10	100	340	– 60	3600	– 600
50	20	400	370	– 30	900	– 600
180	**0**	**850**	**2400**	**0**	**10400**	**– 2350**

$$\bar{X} = \frac{180}{6} = 30 \quad \text{and } \bar{Y} = \frac{2400}{6} = 400$$

Regression equation of x and y

$$bxy = \frac{xy}{y^2}$$

$$= \frac{2350}{10400}$$

$$= -0.22586$$

$$x - \bar{x} = bxy\,(y - \quad)$$

$$x - 30 = -0.22586\,(y - 400)$$

$$x - 30 = -0.22586\,y + 90.344$$

$$x = -0.22586\,y + 90.344 + 30.00$$

$$x = 120.344 - 0.22580\,y$$

when $y = 400$

$$x = 120.344 - .22586 \times 400$$

$$= 120.344 - 90.344 = 30.00$$

Regression equation of y and x

$$byx =$$

$$y - \bar{y} = byx\,(x - \quad) =$$

[illegible]

$$y - 400 = -2.7529\,(x - 30) = -2.7529$$

$$y - 400 = -2.7529\,x + 82.575$$

$$y = 482.575 - 2.7529x$$

when $x = 35$

$$y = 482.575 - 2.7529 \times 35$$

$$= 386.22$$

Illustration 9 : For 50 students of M.Com class, the regression equation of marks in statistics (x) on the marks in accountancy (y) is $3y = -5x = 180$. The mean of marks of accountancy is 44 and variance of marks in statistics is nine-sixteenth of the variance of marks in accountancy. Find the mean marks of statistics and the coefficient of correlation between marks in two subjects.

Solution :

We are given $\bar{y} = 44$

$$3y - 5x = 180$$

Now, $3\,(44) - 5x = 180$

$$132 - 5x = 180$$

$$5x = 180 + 132$$

$$x =$$

$$= 62.4$$

Hence the mean marks in statistics is 62.4

From given equation, we can calculate regression coefficient of x and y.

$$\frac{x}{y} = bxy, \qquad 3y - 5x = 180$$

$$-5x = 180 - 3y$$

$$x = \frac{180}{5} \quad \frac{3}{5}y$$

$$= -36 + 0.6y$$

Hence, $bxy = 0.6$

$$r\frac{x}{y} = bxy = 0.6$$

We are given that $\frac{x^2}{y^2} = \frac{9}{16}$

or $\frac{x}{y} = \frac{3}{4}$

So $r \times \frac{3}{4} = 0.6$

or $r = 0.6 \times \frac{4}{3} = 0.8$

Illustration 10. The following results for height and weight of 1000 men were calculated :

	Mean	Standard Deviation	Coefficient of Correlation
Weight	150 lb	20 lbs	0.6
Height	68"	2.5"	

Find an estimate of (*i*) the weight of a man whose height is 5 *ft.* and (*ii*) the height of a man whose weight is 200 *lbs.*

Solution :

Let weight be represented by x

height be represented by y

Then $\bar{x} = 150$, $\quad$ s$x = 20$ $\quad r = .6$

$= 68$ $\quad$ s$y = 2.5$

(*i*) Regression of x and y

$$(x + \quad) = .6 \times r \qquad (y - \bar{y})$$

or $(x - 150) = .6 \times \qquad (y - 68)$

or $(x - 150) = \frac{12}{2.5}(y - 68)$

or $x - 150 = 4.8\,(y - 68)$

or $x - 150 = 4.8y - 326.4$

or $x = 4.8y - 326.4 + 150$

or $x = 4.8y - 176.4$

Value of x when y is 5ft or 60"

$$x = (4.8 \times 60) - 176.4$$
$$= 288 - 176.4 = 111.6 \text{ lbs}$$

(*ii*) Regression of y on x

$$(y - \bar{y}) = r \qquad (x - \bar{x})$$

or $$(y - 68) = .6 \times \frac{2.5}{20} (x - 150)$$

or $$(y - 68) = .075\ (x - 150)$$

or $$(y - 68) = .075x - 11.25$$

or $$y = .075x - 11.25 + 68$$

or $$y = .075x + 56.75$$

Value of y when x is 200 lbs

or $$y = (.075 \times 200) + 56.75$$

or $$y = 15 + 56.75$$
$$= 71.75"$$

Thus the estimated weight of a person whose weight is 5ft is 111.6 lbs and the estimated height of a person whose height is 100 lbs is 71.75".

अनुमान का प्रमाप-विभ्रम
(Standard Error of Estimate)

एक श्रेणी में प्रतीपगमन रेखाओं से एक श्रेणी के स्वतंत्र चर मूल्यों के समकक्ष आश्रित चर मूल्यों का सर्वोचित अनुमान लगाया जाता है। अनुमानिक मूल्य यथार्थ मूल्य के कितने निकट है, यह ज्ञात करने के लिए 'अनुमान का प्रमाप विभ्रम' ज्ञात किया जाता है। 'अनुमान का प्रमाप विभ्रम' आश्रित श्रेणी के वास्तविक मूल्यों (Actual Values) तथा अनुमानित मूल्यों (Computed Values or Trend Values) के विचलनों के वर्गों के औसत का वर्गमूल होता है। इसकी गणना प्रमाप विचलन की भांति की जाती है, अन्तर केवल इतना है कि इसमें अपेक्षा माध्य के वास्तविक मूल्यों के अनुमानिक अथवा प्रवत्ति -मूल्यों से विचलन निकाले जाते हैं।

दोनों प्रतीपगमन रेखाओं के अनुमान के प्रमाप विभ्रम के सूत्र निम्नलिखित हैं :-

x on y : $\qquad$ y on x

$$S_{xy} = \sqrt{\frac{(x \quad x')^2}{N}} \qquad S_{yx} = \sqrt{\frac{(y \quad y')^2}{N}}$$

यदि x' व y' की गणना न की जाय तो r, s_x व s_y से भी प्रमाप विभ्रम निकाले जा सकते हैं :-

$$S_{xy} = {}_x\sqrt{1 \quad r^2} \qquad S_{yx} = {}_y\sqrt{1 \quad r^2}$$

इससे तात्पर्य यह है कि वास्तविक मूल्यों के 68.27% बिन्दु y की प्रतीपगमन रेखा के $\pm S_{yx}$ के बराबर दोनों ओर के क्षेत्र में बिखरे होंगे। इसी प्रकार रेखा के दोनों ओर $\pm 2S_{yx}$ में 95.45% तथा $\pm 3S_{yx}$ में 99.73% मूल्य बिखरे होंगे।

Illustration 11 : Given the following data :

x	**:**	6	3	9	4	8
y	**:**	8	6	11	6	9

Calculation of standard error of estimate.

Solution : Calculation of standard error of estimate.

X	***x***	***x*²**	**Y**	***y***	***y*²**	***xy***
6	0	0	8	0	0	0
3	–3	9	6	–2	4	6
9	3	9	11	3	9	9
4	–2	4	6	–2	4	4
8	2	4	9	1	1	2
30	**0**	**26**	**40**	**0**	**18**	**21**

$$\overline{X} = \frac{x}{N} = \frac{30}{5} = 6 \qquad \overline{Y} = \frac{y}{N} = \frac{40}{5} = 8$$

$$b_{xy} = \frac{xy}{y^2} \qquad b_{yx} = \frac{xy}{x^2}$$

$$= \frac{21}{18} = 1.17 \qquad = \frac{21}{26} = 0.81$$

(*i*) Regressin equation of *x* on *y*

$$x - \overline{x} = bxy\,(y - \overline{y})$$
$$x - 6 = 1.17\,(y - 8)$$
$$= 1.17y - 9.36$$
$$x = -3.36 + 1.17\,y$$

(*ii*) Regression equation of *y* on *x*

$$y - \quad = byx\,(x - \quad)$$
$$y - 8 = 0.81\,(x - 6)$$
$$= 0.81x - 4.86$$
$$y = 3.14 + 0.81x$$

Now for calculating S.E.E., we have to find trend values of *x* and *y* (*x*′, *y*′) using the above two equations.

x	***x*′**	**(*x* – *x*;)**	**(*x* –)²**	***y***	***y*′**	**(*y* – *y*′)**	**(*y* –)²**
6	6	0	0	8	8	0	0
3	3.67	– 0.67	1.45	6	5.58	0.42	0.18
9	9.50	– 0.50	0.25	11	10.42	0.58	0.34
4	3.67	0.33	0.11	6	6.38	– 0.38	0.14
8	7.16	0.84	0.70	9	9.62	– 0.62	0.38
			1.51				**1.04**

$$Sxy \ =$$

$$Syx \ = \sqrt{\frac{(y \quad y')^2}{N}} \quad \sqrt{\frac{1.04}{5}} \quad 0.46$$

Alternative Method :

$$sx \ = \sqrt{\frac{x^2}{N}} = \sqrt{\frac{26}{5}} = 2.28; \qquad s_y \ = \sqrt{\frac{y^2}{N}} = \sqrt{\frac{18}{5}} = 1.90$$

$$r \ = \frac{xy}{\sqrt{\ x^2.\ \ y^2}} = \frac{21}{\sqrt{18 \quad 26}} = 0.97$$

$$Sxy \ = sx\sqrt{1 \quad r^2} \qquad Syx \ = sy\sqrt{1 \quad r^2}$$

$$= 2.28\sqrt{1 \quad (0.97)^2} \qquad = 1.90$$

$$= 0.55 \qquad = 0.46$$

प्रतीपगमन विश्लेषण की उपयोगिता

(Uses of Regression Analysis)

$$\sqrt{\frac{1.\ ((0.97)^2}{N}} \quad \sqrt{\frac{1.51}{5}} \quad 0.55$$

प्रतीपगमन विश्लेषण सांख्यिकी की एक महत्वपूर्ण तकनीक है जिसके द्वारा दो या अधिक चरों के मध्य सहसम्बन्ध की प्रवत्ति तथा मात्रा का माप करके उनके अज्ञात मानों का अनुमान लगाया जाता है। निम्नलिखित दशाओं में ये बहुत ज्यादा उपयोगी होता है :-

(1) **व्यवसायिक पूर्वानूमान (Business Forecasting) :** व्यवसाय में पूर्वानुमान बहुत जरूरी होता है। भविष्य में बिक्री, लाभ, लागत आदि का पूर्वानुमान लगाकर एक कम्पनी अपनी नीतियाँ बना सकती है। प्रतीपगमन इस पूर्वानुमान की मुख्य तकनीक है।

(2) **शुद्ध व वाणिज्य सम्बन्धित अनुसंधान (Pure and Applied Research) :** जैसा की इस पाठ के आरम्भ में ही गाल्टन की पिता और पुत्रों की ऊंचाई के सम्बन्धी अनुसंधान का उल्लेख किया गया है। इसी प्रकार अनेक प्रकार की शुद्ध तथा वाणिज्य सम्बन्धी अनुसंधान में प्रतीपगमन विश्लेषण एक महत्वपूर्ण स्थान रखता है।

(3) **नियन्त्रण उपकरण (Control Tool) :** इस विश्लेषण के द्वारा निश्चित उत्पादन पर कीमत-स्तर, निश्चित मूल्य पर मांग की मात्रा, प्रयुक्त होने वाले कच्चे काल की कीमत आदि के अनुमान द्वारा महत्वपूर्ण व्यवसायिक नीतियां निर्धारित की जाती है।

इसके अतिरिक्त वर्षा व उत्पादन, खाद व उत्पादन, बचत और आर्थिक विकास दर आदि में भी प्रतीपगमन विश्लेषण का काफी प्रयोग किया जाता है।

प्रतीपगमन और सहसम्बन्ध विश्लेषण में अन्तर

(Difference Between Regression and Correlation Analysis)

सहसम्बन्ध	प्रतीपगमन
(1) सहसम्बन्ध दो चरों के बीच में सम्बन्ध का अध्ययन है जिससे कि एक चर के मानों में हलचल से दूसरे चर के मानों में भी हलचल होती है।	(1) प्रतीपगमन का अर्थ औसतन की तरफ पीछे लौटना है यानि ये विश्लेषण दो चरों के बीच औसत सम्बन्ध दर्शाता है।
(2) सहसम्बन्ध में दो चरों में कारण और असर (Cause and Effect) सम्बन्ध अनिवार्य नहीं है।	(2) प्रतीपगमन में कारण और असर सम्बन्ध साफ नजर आते हैं। कारण स्वतन्त्र चर है व असर आश्रित चर है।
(3) इसमें दो चरों के बीच में निरर्थक या भ्रामक (Non-sense or spurious) सहसम्बन्ध हो सकता है।	(3) इसमें कोई भी प्रतीपगमन समीकरण निरर्थक या भ्रामक नहीं होता।
(4) सहसम्बन्ध के मान पर x का y पर आश्रित होने या y का x पर आश्रित होने का कोई प्रभाव नहीं पड़ता।	(4) प्रतीपगमन दो चरों के बीच में फलन सम्बन्ध (Functional Relationship) दर्शाता है। यह सम्बन्ध एक चर को दूसरे चर पर आश्रिती दर्शाता है।
(5) सहसम्बन्ध गुणांक एक सापेक्ष माप (Relative Measure) है। इसकी कोई इकाई नहीं होती।	(5) प्रतीपगमन गुणांक निरपेक्ष माप (Absolute Measure) है जो कि एक चर की कीमतों में परिवर्तन दर्शाते हैं। यदि हमें स्वतन्त्र चर का मूल्य ज्ञात हो तो हम आश्रित चर के मूल्य का अनुमान लगा सकते हैं।
(6) सहसम्बन्ध गुणांक उद्‌गम तथा पैमाने में परिवर्तन से प्रभावित नहीं होता।	(6) प्रतीपगमन गुणांक उद्‌गम के प्रति स्वतन्त्र होते हैं $\sqrt{b_{xy} \cdot b_{yx}}$ परन्तु पैमाने के प्रति नहीं। अर्थात् यदि पैमाने में परिवर्तन कर दिया जाए तो गुणांक के मूल्य में भी परिवर्तन हो जाएगा।
(7) यदि दो चरों के बीच रेखीय सम्बन्ध (Linear Relationship) तक सीमित है।	(7) यह चरों के बीच रेखीय तथा गैर-रेखीय दोनो प्रकार के सम्बन्ध का अध्ययन करता है।

सारांश

- प्रतीपगमन विश्लेषण दो या अधिक चरों के बीच औसत सम्बन्ध दिखाने की एक महत्वपूर्ण विधि है।
- इस विश्लेषण में एक चर (x or y) स्वतन्त्र होता है तथा दूसरा आश्रित (y or x) होता है। x तथा y के बीच सम्बन्ध दिखाने के लिए दो समीकरण बनाये जाते हैं। $y = a + bx$ तथा $x = a + by$।
- इन समीकरणों में a तथा b के मान निकालने के लिए हम प्रत्येक समीकरण के लिए दो समीकरण और बनाते हैं जैसे

  (1) y on x ($y = a + bx$)

  $Sy = Na + bSx$

  $Sxy = aSx + bSx^2$

  (2) x on y ($x = a + by$)

  $Sx = Na + bSy$

  $Sxy = aSy + bSy^2$
- इन समीकरणों की सहायता से स्वतंत्र चर के किसी एक मूल्य के लिए आश्रित चर के अज्ञात मूल्य का अनुमान लगाते हैं। इसलिए प्रतीपगमन विश्लेषण पूर्वानुमान लगाने की एक मुख्य विधि है।
- प्रतीपगमन गुणांकों का सहसम्बन्ध गुणांक से गहरा सम्बन्ध है। जैसे $r =$ यानि दोनों प्रतीपगमन गुणांकों के गुणनफल के वर्गमूल का मान कार्ल पियर्सन के सहसम्बन्ध गुणांक के बराबर होता है।

- तीनों गुणांकों (r, b_{yx}, b_{xy}) का चिन्ह हमेशा एक ही रहेगा। या तो तीनो धनात्मक होंगे या तीनों ऋणात्मक होंगे।
- जिस तरह से सहसम्बन्ध विश्लेषण में हम निर्धारण गुणांक का मान निकालते हैं, उसी तरह से प्रतीपगमन विश्लेषण में हम अनुमान का प्रमान विभ्रम (S_{xy} or S_{yx}) का मान निकालते हैं :-

$$S_{xy} = sx\sqrt{1 \quad r^2} \qquad \text{or} \qquad \sqrt{\frac{(x \quad x_c)^2}{N}}$$

तथा

$$S_{yx} = sy\sqrt{1 \quad r^2} \qquad \text{or} \qquad \sqrt{\frac{(y \quad y_c)^2}{N}}$$

- इस अनुमान से हमें अनुमानित मूल्यों व वास्तविक मूल्यों का एक-दूसरे की निकटता के बारे में अंदाज होता है। दोनों मूल्यों में जितना कम अन्तर होगा, अनुमान का मूल्य भी कम होगा।

प्रश्नावली
(Exercise)

(1) प्रतीपगमन से क्या आश्य है ? प्रत्येक द्विचर वितरण के लिए सामान्यतः दो प्रतीपगमन रेखाएं क्यों होनी चाहिए ?

What is meant by regression ? Whay should there be in general two lines of regression for each bivariate distribution?

(2) रेखीय प्रतीपगमन के विचार को समझाइए ? प्रतीपगमन रेखाएँ दो क्यों होती हैं ? क्या वे एक दूसरे को काटती हैं? यदि हाँ, तो कहाँ ?

Explain the concept of linear regression. Why are there two regression lines ? Do they cut each other ? Is so, where ?

(3) 'प्रतीपगमन का क्या अर्थ है ? द्विचर वितरण के लिए, सामान्यतः दो प्रतीपगमन रेखाओं का होना क्यों आवश्यक है? आपके विचारानुसार दो चर-मूल्यों का सहसम्बन्ध गुणक कितना होना चाहिए, यदि दोनों प्रतीपगमन रेखाएँ एक दूसरे को समकोण पर काटें तथा यदि वे दोनों एक दूसरे को ढक लें ?

What is meant by 'regression' ? Why should there be, in general, two lines of regression for each bivariate distribution ? What do you think the coefficient of correlation between the two variables would be if two regression lines cut at right angles, and what if they coincide ?

(5) सहसम्बन्ध तथ प्रतीपगमन का अन्तर स्पष्ट कीजिए। प्रतीपगमन मापने की विभिन्न विधियों का भी वर्णन कीजिए।

Distinguish between correlation and regression. Also discuss the various methods of measuring regression.

(6) ''प्रतीपगमन रेखा सम्बन्धित राशि (Quantity) के मूल्य का केवल 'सर्वोत्तम अनुमान' देती है। हम इस अनुमान में अनिश्चितता की मात्रा को एक परिणाम, जिसको अनुमान का प्रमाप विभ्रम कहते हैं, कि गणना की आंक सकते है.......'' स्पष्ट कीजिए।

"The regression lines gives only a 'best estimate' of the value of the quantity in question. We may assess the degree of uncertainty in this estimate by calculating a quantity known as the Standard Error of Estimate...." Elucidate.

(7) प्रतीपगमन विश्लेषण की धारणा की व्याख्या कीजिए और इसकी उपयोगिता पर प्रकाश डालिए।

Explain the concept of regression analysis and comment on its utility.

(8) निम्न आंकड़े 20 दंपतियों में पति (x) व पत्नी (y) की आयुओं को दर्शाते हैं। दोनों प्रतीपगमन रेखाएँ ज्ञात करें व प्रतीपगमन गुणांकों से सहसम्बन्ध गुणांक का मान ज्ञात करो।

The following data represent the ages of husband (x) and wife (y) for twenty couples. Calculate both the regression equations and coefficient of correlation from regression coefficients :

x	22	24	26	26	27	27	28	28	29	30	30	30	31	32	33	34	35	35	36	37
y	18	20	20	24	22	24	27	24	21	25	29	32	27	27	30	27	30	31	30	32

$y = 0.89x - 0.57,\ x = 0.82y + 8.6,\ r = \overline{\text{Y}}.856$

(9) निम्न दिए हुए समंकों से $\bar{x}$, , σ_x, σ_y, r तथा प्रतीपगमन रेखाएँ ज्ञात करो।

Given the following :

$\Sigma x = 150{,}000$ $\Sigma y = 70{,}000$ $\Sigma xy = 10{,}522{,}500$

$\Sigma x^2 = 22{,}725{,}000$ $\Sigma y^2 = 49{,}36{,}000$ $N = 1000$

Find , , σ_x, σ_y, r and lines of regressin.

[**Ans.** = 150, = 70, $\sigma_x = 15$, $\sigma_y = 6$, $r = .26$, $y = .1x + 55$, $x = .625y + 106.25$]

(10) अर्थशास्त्र व सांख्यिकी में प्राप्त अंकों से निम्नलिखित परिणाम प्राप्त किए गए :-

The following results were obtained from scores in Economics and Statistics.

	अर्थशास्त्र में अंक (Scores in Economics) (X)	सांख्यिकी में अंक (Scores in Statistics) (Y)
माध्य (Mean)	50	100
प्रमाप विचलन (Standard Deviation)	5	10

X और Y के बीच कार्ल पियर्सन सहसम्बन्ध गुणांक = 0.5

Karl Pearson's correlation coefficient between X and Y = 0.5

दोनों प्रतीपगमन रेखाएँ ज्ञात कीजिए। इन प्रतीपगमनों का प्रयोग करके X का मूल्य ज्ञात कीजिए जबकि Y = 80 और Y का मूल्य ज्ञात कीजिए जबकि X = 90 हो।

Obtain both the regression lines, using regressions estimate the value of X when Y = 80 and Y for X = 90.

(11) निम्न दिए हुए आंकड़ों से y का अनुमानिक मूल्य ज्ञात करें जबकि $x = 12$ व x का अनुमानिक मूल्य ज्ञात करें जबकि $y = 20$.

Given the following data, calculate the expected value of y when x is 12 and of x when y is 20.

	x	y
माध्य (Mean Value)	10.5	25.5
प्रमाप विचलन (Standard Deviation)	2.2	4.4

x तथा y के बीच कार्ल पियर्सन सहसम्बन्ध गुणांक = + .9

Coefficient of correlation between x and y = + .9

[**Ans.** $y = 28.2$, $x = 8.025$]

(12) कीमत x तथा आपूर्ति y के 10 निरीक्षणों में निम्नलिखित आंकड़े लिए गए : $\Sigma x = 130$, $\Sigma y = 200$, $\Sigma x^2 = 2288$, $\Sigma y^2 = 5506$, $\Sigma xy = 3467$ y की x पर प्रतीपगमन रेखा बनाओं तथा 16 इकाई कीमत के लिए आपूर्ति का अनुमान लगाएं।

For 10 observations on price (x) and suppy (y) the following data were obtained (in approximate units) : $\Sigma x = 130$, $\Sigma y = 200$, $\Sigma x^2 = 2288$, $\Sigma y^2 = 5506$, $\Sigma xy = 3467$ obtain the line of regression of y on x and estimate the supply when the price is 16 units.

[**Ans.** $y = 1.44\,x + 1.28$, 24.32]

(13) सांख्यिकी तथा लेखाकर्म की एक विशेष परीक्षा के अंकों के सम्बन्ध में निम्न आँकड़े दिए गए हैं :-

	सांख्यिकी	लेखाकर्म
मध्यक	39.5	47.5
प्रमाप विचलन	10.8	16.8

सांख्यिकी तथा लेखाकर्म के अंकों के मध्य सह-सम्बन्ध गुणांक + .42 है।

(*i*) लेखाकर्म में ऐसे विद्यार्थी के सम्भावित औसत प्राप्तांक ज्ञात कीजिए जिन्होंने सांख्यिकी में 40 अंक प्राप्त किए हों।

(*ii*) सांख्यिकी में ऐसे विद्यार्थी के सम्भावित औसत प्राप्तांक ज्ञात कीजिए जिन्होंने लेखाकर्म में 50 अंक प्राप्त किए हों।

[**Ans.** Expected makrs in Accountancy = 47.825
Expected marks in Statistics = 40.175]

The following data are given for marks in Statistics and Accountancy in a certain examination :

	Statistics	Accountancy
Mean	39.5	47.5
Standard deviation	10.8	16.8

Coefficient of correlation between marks in Statistics and Accountancy is + .42.

(*i*) Give the expectations of marks in Accountancy for candidates who secured 40 marks in Statistics.

(*ii*) Give the expectations of marks in Statistics for candidates who secured 50 marks in Accountancy.

Find out two regression equations from the following table :

निम्न सारणी से दो प्रतीपगमन समीकरण ज्ञात कीजिए :

X \ Y	**18**	**19**	**20**	**21**	**22**	**Total**
0-5	—	—	—	3	1	4
5-10	—	—	—	3	2	5
10-15	—	—	7	10	—	17
15-20	—	5	4	—	—	9
20-25	3	2	—	—	—	5
Total	3	7	11	16	3	40

[**Ans.** $X_c = 101.66 - 4.37\,Y$, $Y_c = 22.35 - .16\,X$]

(15) एक द्विमुखी समंकों के लिए, निम्न सूचना दी हुई है :-

$$S\,(x-44) = -5,\; S\,(x-44)^2 = 255,$$
$$S\,(y-26) = -6,\; S\,(y-6)^2 = 704$$
$$S\,(x-44)(y-26) = -306$$

निरीक्षणों के जोड़ो की संख्या = 12

(*i*) दोनों प्रतीपगमन रेखाओं व

(*ii*) x तथा y के बीच सहसम्बन्ध गुणांक ज्ञात करो।

For a bivariate data, you are given the following information :-

$$S\,(x-44) = -5,\ S\ (x-44)^2 = 255,$$
$$S\ (y-26) = -6,\ S\ (y-6)^2 = 704$$
$$S\ (x-44)\,(y-26) = -306$$

Number of pairs of obsevation = 12

Find out (*i*) the two regression equations and (*ii*) the coefficient of correlation between x and y series.

[**Ans.** $x = 54.80 - .44\,y$, $y = 78.67 - 1.22\,x$ and $r = -.733$]

(16) सहसम्बन्ध विश्लेषण के आंशिक रूप से जले हुए लेख से प्राप्त आंकड़े नीचे दिये हुए हैं।

X का प्रसरण = 9

प्रतीपगमन रेखाएं :
$$8X - 10y + 64 = 0$$
$$40x - 18x = 240$$

इस सूचना के आधार पर (*i*) X तथा Y के माध्य (*ii*) X तथा y के बीच सहसम्बन्ध गुणांक तथा (*iii*) y का प्रमाप विचलन निकालो।

Given below results are available from the partially destroyed records of an analysis of correlation data.

Variance of $x = 9$

Regression equations,
$$8x - 10y + 64 = 0$$
$$40x - 18y = 240$$

On the basis of information, find (*i*) mean of x and y, (*ii*) correlation coefficient between x and y and (*iii*) standard deviation of y.

(17) नीचे दो प्रतीपगमन समीकरण दिए गए हैं :-

आपको निकालना है :-

(*i*) x तथा y के माध्य;

(*ii*) y का प्रमाप विचलन

(*iii*) x एवं y के मध्य सहसम्बन्ध गुणांक

प्रतीपगमन समीकरण हैं :

$8x - 10y + 70 = 0$; $15x - 6y = 60$, x का प्रसारण = 9

Two regression equations are given below :

Find out :

(*i*) Mean values of x and y;

(*ii*) Standard deviation of y.

(*iii*) Co-efficient of correlation between x and y.

The regression equations are $8x - 10y + 70 = 0$, $15x - 6y = 60$; Variance of $x = 9$.

❑❑

अध्याय - 10

गुणसम्बन्ध (Association of Attributes)

गुणात्मक संमक (Statistics of Attributes) : संख्यिकीय विधियों (Statistical Methods) का प्रयोग संख्यात्मक तथ्यों तक ही सीमित रहता है। संख्यात्मक तथ्य दो प्रकार से संकलित किये जा सकते हैं :-

(1) अबलोकनकर्त्ता (Observer) प्रत्येक इकाई अथवा पद की किसी विशेषता की माप का लेखा करता है। व्यक्तियों की आयु, विद्यार्थियों की ऊँचाई, दुर्घटनाएँ, श्रमिकों का पद आदि ऐसे तथ्य हैं जिनकी संख्यात्मक माप की जा सकती हैं। इस प्रकार के समंकों को चर-मूल्यों के समंक (Statistics of Variable) कहते हैं।

(2) अवलोकनकर्त्ता प्रत्येक इकाई अथवा पद में किसी गुण की उपस्थिति (Presence) अथवा अनुपस्थिति (Absence) के आधार पर भी समंकों का संकलन कर सकता है। कुछ ऐसे तथ्य होते हैं जिनकी अंकात्मक माप सम्भव नहीं होती जैसे योग्यता, आंख का रंग, वैवाहिक स्थिति आदि। ''प्रयोगात्मक कार्य में बहुधा हम कुछ ऐसी विशेषताओं अथवा गुणों से व्यवहार करते हैं जिनका ठीक मापन सम्भव नहीं होता, यद्यपि समग्र को उसके गुणों के आधार पर दो या अधिक वर्गों से विभक्त करना सम्भव होता है'' ऐसे गुणात्मक तथ्यों को गिनकर ही संख्यात्मक रूप दिया जाता है। इस प्रकार संकलित समंकों को गुणात्मक समंक (Statistics of Attributes) कहते हैं।

गुण सम्बन्ध तथा सह-सम्बन्ध

(Association of Attributes and Co-efficient of Correlation)

इन दोनों में मुख्य अंतर चर और गुण के कारण है। यदि हम दो या उससे अधिक चरों के मध्य सम्बन्ध का अध्ययन करना चाहते हैं तो यह सहसम्बन्ध है। दूसरी तरफ यदि हमें दो गुणों में सम्बन्ध का अध्ययन करना है तो गुण सम्बन्ध है। दोनों में कुछ समानताएं हैं व कुछ अन्तर हैं। जैसे दोनों के ही चिन्ह ऋणात्मक या धनात्मक हो सकते हैं। दोनों के ही गुणंकों का मान -1 से $+1$ के बीच होता है। सहसम्बन्ध की तरह गुण सम्बन्ध भी मात्रक हो सकते हैं।

'उपरोक्त मुख्य अंतर के अतिरिक्त भी दोनों में कुछ और अंतर भी है। जैसे गुण सम्बन्ध में किसी एक गुण को हम दो वर्गों में विभाजित करते हैं — एक जो गुण से सम्बन्ध रखता है, दूसरा जो उस गुण से सम्बन्ध नहीं रखता। सहसम्बन्ध में इस प्रकार का वर्गीकरण संभव नहीं है। उसमें हम इकाइयों का संख्यात्मक आधार पर वर्गीकरण करते हैं। गुण सम्बन्ध का विश्लेषण बिन्दु-रेखीय विधि से संभव नहीं है लेकिन सहसम्बन्ध का विश्लेषण हम इस विधि से कर सकते हैं।

वर्गीकरण (Classification) अवलोकित इकाई अथवा पद की विशेषता 'गुण' (Attribute) कहलाती है। जब हम निश्चित गुणों के आधारों पर समंकों का संकलन करके एक वर्ग में रखते हैं तो इस प्रक्रिया को गुणात्मक वर्गीकरण (Classification According to Attributes) कहते हैं। विशेष गुणों की उपस्थिति तथा अनुपस्थिति के आधार पर समंकों को वगीकृत किया जाता है। साधारण रूप में, यदि किसी समग्र (Universe or Population) में केवल एक गुण का अध्ययन किया जाए, तो पारम्परिक अपवर्जी (Mutually Exclusive) दो वर्गों में समग्र विभक्त हो जाएंगा — एक वर्ग में वे पद अथवा इकाइयाँ होंगी जिनमें वह गुण उपस्थित है तथा दूसरे वर्ग में वे पद अथवा इकाइयाँ होंगी जिनमें वह गुण अनुपस्थित हैं। उदाहरणार्थ, यदि किसी समग्र में साक्षरता (Literacy) गुण का अध्ययन किया जाए, तो दो वर्ग बनेंगे — एक वर्ग उनका जो साक्षर हैं तथा दूसरा वर्ग उनका जो साक्षर नहीं हैं। यदि एक से अधिक गुणों का अध्ययन किया जाए तो वर्गों की संख्या दो से अधिक होगी। यदि साक्षरता के साथ-साथ अपराधिता (Criminality), गुण का भी अध्ययन किया जाए तो 'साक्षर' (Literates) निरक्षर (Not Literates), अपराधी (Criminals), निर-अपराधी (Not Criminal), साक्षर-अपराधी (Literate Criminals), साक्षरनिर अपराधी (Literate not Criminals), निरक्षर-अपराधी (Not Literate Criminals) तथा निरक्षर-निर अपराधी (Not Literate not Criminals) के वर्गों की रचना करनी पड़ेगी।

यदि एक गुण (Attributes) का ही अध्ययन किया जाता है तो समग्र को दो वर्गों में विभक्त किया जाता है — एक वर्ग में विवेचित गुण उपस्थिति होता है तथा दूसरे वर्ग में वह गुण उपस्थित नहीं होता है। ये वर्ग पारस्परिक रूप से अपवर्जी (Mutually Exclusive) होते हैं। इस प्रकार के वर्गीकरण को द्वन्द-भाजन (Classification of Division by Dichotomy) कहते हैं। यदि उन्हीं गुणों के आधार पर समग्र को विभाजित तथा उप-विभाजित किया जाता है तो इस प्रकार के वर्गीकरण को बहुगुणी-वर्गीकरण (Manifold Classification) कहते हैं।

संकेताक्षर (Notation and Terminology) सैद्धान्तिक उद्देश्यों के लिए यह आवश्यक हो जाता है कि गुणों तथा बनाये गए वर्गों को व्यक्त करने के लिए संकेताक्षरों का प्रयोग किया जाय। सामान्यतः गुणों की उपस्थिति के लिए अंग्रेजी वर्णमाला के बड़े अक्षरों A, B,C आदि का प्रयोग किया जाता है। ग्रीक वर्णमाला के अक्षर 'a' (alpha), 'b' (Beta), 'g' (Gamma) आदि का प्रयोग गुणों की अनुपस्थिति के लिए किया जाता है। ग्रीक वर्णमाला के अक्षरों के स्थान पर अंग्रेजी वर्णमाला के छोटे अक्षरों *a, b, c* आदि का प्रयोग भी किया जा सकता है। उदाहरणार्थ, यदि 'A' संकेताक्षर का प्रयोग किसी गुण, जैसे साक्षरता का प्रतिनिधित्व करने के लिए किया जाता है तो '*a*' निरक्षरता का प्रतिनिधित्व करेगा। यदि 'B' संकेताक्षर अपराधिता का प्रतिनिधित्व करता है तो '*b*' नि-अपराधिता का संकेत करेगा। यदि 'C' संकेताक्षर 'दण्ड प्राप्त होने' (Punishment) का प्रतिनिधित्व करता है तो '*c*' दण्ड प्राप्त न होने का प्रतिनिधित्व करेगा। N समग्र (Universe) का प्रतिनिधित्व करता है।

गुण-संयोग (Combination of Attributes) : गुणों के संयोग का प्रतिनिधित्व अक्षरों के संयोग द्वारा किया जाता है। उदाहरणार्थ, यदि 'A' साक्षरता का तथा 'B' अपराधिता का प्रतिनिधित्व करता है तो 'AB' साक्षरता तथा अपराधिता दोनों का प्रतिनिधित्व करेगा। इसी प्रकार 'A*b*' साक्षरता तथा निर अपराधिता '*a*B' निरक्षरता अपराधिता, तथा '*ab*' निरक्षरता तथा निरअपराधिता का प्रतिनिधित्व करेंगे।

वर्ग आवत्ति (Class Frequency) : विभिन्न गुण-वर्गों की बारंबारात वर्ग-आवत्ति कहलाती है। वर्ग-संकेताक्षरों को कोष्ठकों (Brackets) के अन्दर लिख देने से उस वर्ग की आवत्तियों का संकेत मिलता है जैसे (AB), (*a*B), (A*b*), (*ab*) आदि सम्बन्धित आवत्तियों को प्रदर्शित करेंगे।

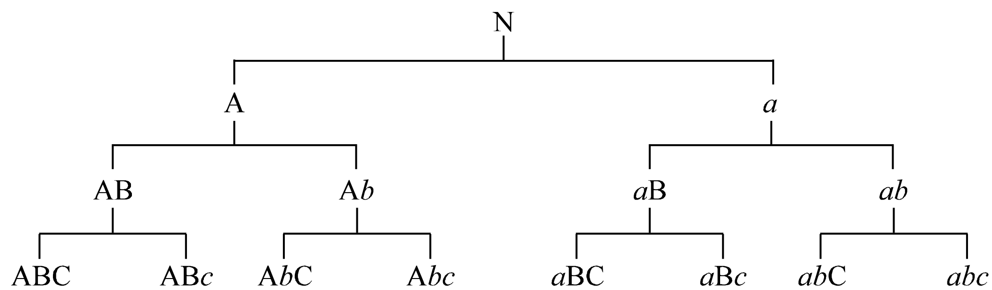
वर्गों की संख्या (Number of Classes) : गुणों की संख्या के आधार पर गठित सम्पूर्ण वर्गों की संख्या 3^n सूत्र से ज्ञात की जा सकती है। इस सूत्र में n = गुणों की संख्या (Number of Attributes) है। यदि केवल एक गुण का अध्ययन किया जाता है तो वर्गों की संख्या $3^1 = 3$ (N, A, *a*) होगी। यदि दो गुणों का अध्ययन किया जाता है तो सम्पूर्ण वर्गों की संख्या 32 = 9 (N, A, B, *a*, *b*, AB, A*b*, *a*B, *ab*) होगी। तीन गुणों के अध्ययन करने पर वर्गों की संख्या $3^3 = 27$ [N, A, B, C, *a*, *b*, *c*, AB, AC, BC, A*b*, B*c*, *a*B, *a*C, *bc*, *ab*, *ac*, *bc*, ABC, AB*c*, A*b*C, *a*BC, *a*B*c*, *ab*C, *abc*) होगी।

बड़े अक्षरों (Capital Letters) द्वारा जिन गुणों का प्रतिनिधित्व होता है उन्हें अनुलोम गुण (Positive Attributes) कहते हैं। अनुलोम गुण वाले वर्गों जैसे (A, AB, ABC आदि) को अनुलोम वर्ग (Positive Attributes) कहते हैं। छोटे अक्षरों (Small Letters) द्वारा जिन गुणों का प्रतिनिधित्व होता है उन्हें विलोम गुण (Negative Attributes) कहते हैं तथा विलोम गुण वाले वर्गों जैसे (*a, ab, abc* आदि) को विलोम वर्ग (Negative Attributes) कहते हैं। अनुलोम तथा विलोम गुणों के संयोग से बने वर्गों जैसे A*b*, *a*B, *a*B*c* आदि की विपरीत वर्गों के युग्म (Pairs of Contrary Classes) कहते हैं।

वर्गों का क्रम (Order of Classes) : अवलोकनों (Observations) की पूर्ण संख्या का समग्र (Universe of Population) कहते हैं तथा उसे N संकेताक्षर द्वारा व्यक्त किया जाता है। इसे शून्य क्रम का वर्ग (Class of the Zero Order) कहते हैं। यदि केवल दो गुणों का ही अध्ययन किया जाता है तो A, *a,* B, *b* प्रथम क्रम के वर्ग (Classes of the First Order) AB, A*b,* *a*B, *ab* द्वितीय क्रम के वर्ग (Classes of the second order) कहलाते हैं। यदि उच्चतर क्रम के अन्य वर्ग न हों तो अन्तिम क्रम के वर्गों को 'अन्तिम क्रम के वर्ग' (Classes of the Ultimate Order) कहलाते हैं। यदि तीन गुणों का अध्ययन किया जाता है तो ABC, AB*c* आदि ततीय क्रम के वर्ग (Classes of the Third Order) कहलाते हैं। अन्य वर्ग न होने पर इन्हीं को अन्तिम क्रम के वर्ग कहते हैं। अन्तिम क्रम में कितने वर्ग होंगे यह 2^n सूत्र से ज्ञात किया जा सकता है। इस सूत्र में n गुणों की संख्या का प्रतिनिधित्व करता है। यदि दो गुणों का अध्ययन किया जाय तो अन्तिम क्रम के वर्गों की संख्या $2^2 = 2 \times 2 = 4$ होगी। यदि तीन गुणों का अध्ययन किया जाय तो अन्तिम क्रम के वर्गों की संख्या $2^3 = 2 \times 2 \times 2 = 8$ होगी। इन वर्गों को अन्तस्थ वर्ग (Ultimate Classes) भी कहा जाता है।

इसको निम्न भांति भी प्रदर्शित किया जा सकता है।

Order	0	N			
Order	1	(A)	(B)		(C)
		(a)	(b)		(c)
Order	2	(AB)	(AC)		(BC)
		(Ab)	(Ac)		(Bc)
		(aB)	(aC)		(bC)
		(ab)	(ac)		(bc)
Order	3	(ABC)	(aBC)	(AbC)	(abC)
		(ABc)	(abc)	(Abc)	(abc)



उपर्युक्त वर्गीकरण के आधार पर :-

$$\begin{aligned}
N &= (A) + (a) \\
&= (B) + (b) \\
(A) &= (AB) + (Ab) \\
(a) &= (aB) + (ab) \\
(B) &= (AB) + (aB) \\
(b) &= (Ab) + (ab)
\end{aligned}$$

यदि तीन गुणों का अध्ययन किया जाय तो :'

$$\begin{aligned}
(AB) &= (ABC) + (ABc) \\
(Ab) &= (AbC) + (Abc) \\
(aB) &= (aBC) + (aBc) \\
(ab) &= (abC) + (abc)
\end{aligned}$$

इसी प्रकार अन्य सम्बन्ध भी ज्ञात किए जा सकते हैं।

Illustration 1 : From the following ultimate class frequencies, find the frequencies of the positive and negative clases and the total number of observations :

$$\mathbf{(AB) = 100 \qquad (aB) = 80}$$
$$\mathbf{(Ab) = 50 \qquad (ab) = 40}$$

Solution :

Since the given problem has two attributes A and B, so total no. of values = $3^2 = 9$

We are given (AB) = 100; (aB) = 80, (Ab) = 50, (a, b) = 40

We have to find

(A) Frequency of — Frequency of

(B) Positive classes, a — b Negative classes and N

(A) $\quad (A) = (AB) + (Ab) = 100 + 50 = 150 \qquad (a) = (aB) + (ab) = 80 + 40 = 120$

(B) $\quad = (BA) + (Ba) = 100 + 80 = 180 \qquad (b) = (bA) + (ba) = 50 + 40 = 90$

$$N = (A) + (a) = 150 + 120 = 270$$

or $$(B) + (b) = 180 + 90 = 270$$

We can also nine-square table (for two attributes) to find out the remaining frequencies

	A	*a*	Total
B	(AB) 100	(*ab*) 80	(B) 180
b	(A*b*) 50	(*ab*) 40	(*b*) 90
Total	(A) 150	(*a*) 120	N 270

Illustration 2 : Given the following frequencies of the positive classes, find the frequencies of the ultimate classes :-

$$\mathbf{A = 160, (B) = 200, (AB) = 140, N = 500}$$

Solution :

We have to find out :-

A*b*, *a*B, and *ab*

$$\begin{aligned}(Ab) &= (A) - (AB)\\ &= 160 - 140 = 20\\ (aB) &= (B) - (AB)\\ &= 200 - 140 = 60\\ (ab) &= a - (aB)\\ &= (N - A) - (A - AB)\\ &= N - A - B + AB\\ &= 500 - 160 - 200 + 140 = 280\end{aligned}$$

उपरोक्त आवत्तियों को निम्न सूत्र की सहायता से ज्ञात कर सकते हैं।

(*i*) शून्य क्रम की आवत्ति को उससे उच्च क्रम की धनात्मक व ऋणात्मक आवत्ति के योग के रूप में व्यक्त करेंगे। अतः

$$N = A + a; N = B + b; N = C + c$$

(*ii*) इसी प्रकार प्रथम क्रम की आवत्ति ज्ञात करने के लिए ज्ञात करने वाली आवत्ति का अन्य गुण के धनात्मक व ऋणात्मक योग से गुणा करने पर प्रथम क्रम की आवत्ति ज्ञात की जा सकती है। जैसे : -

$$A = A\,(B + b) = AB + Ab$$

$$a = a\,(B + b) = aB + ab$$

(*iii*) इसी प्रकार AB ज्ञात करने के लिए AB को (C + *c*) से गुणा करेंगे। अर्थात्

$$AB = AB\,(C + c)$$

$$AB = ABC + ABc$$

प्रथम क्रम की आवत्तियाँ ज्ञात करना : हम जानते हैं कि :-

$$\begin{aligned}N &= A + a\\ A &= A\,(B + b)\\ A &= AB + Ab\\ A &= AB\,(C + c) + Ab\,(C + c)\\ A &= ABC + ABc + AbC + Abc\end{aligned}$$

Similarly,

$$B = B(C + c)$$
$$B = BC + Bc$$
$$B = BC(A + a) + Bc(A + a)$$
$$B = ABC + aBC + ABc + aBc$$

Similarly,

$$C = C(A + a)$$
$$C = AC + aC$$
$$C = AC(B + b) + aC(B + b)$$
$$C = ABC + AbC + aBC + abC$$

Similarly,

$$a = a(B + b)$$
$$= aB + ab$$
$$= aB(C + c) + ab(C + c)$$
$$= aBC + aBC + abC + abc$$

द्वितीय क्रम की आवत्तियाँ ज्ञात करना : यदि हम AB ज्ञात करना चाहते हैं तो AB को ततीय गुण $(C + c)$ से गुणा करें अर्थात्

$$AB = AB(C + c) = ABC + ABc$$
$$BC = BC(A + a) = ABC + aBC$$
$$AC = AC(B + b) = ABC + AbC$$
$$ab = ab(C + c) = abC + abc$$
$$bc = bc(A + a) = Abc + abc$$
$$ac = ac(B + b) = aBc + abc$$
$$Ab = Ab(C + c) = AbC + Abc$$
$$aB = aB(C + c) = aBC + aBc$$
$$Bc = Bc(A + a) = ABc + aBc$$
$$bC = bC(A + a) = AbC + abC$$
$$Ac = Ac(B + b) = ABc + Abc$$
$$aC = aC(B + b) = abC + abC$$

ततीय क्रम की आवत्तियाँ ज्ञात करना : हम जानते हैं कि :-

$$AB = AB(C + c)$$
$$AB = ABC + ABc$$
$$\backslash \quad ABC = AB - ABc$$
$$ABc = AB - ABC$$

ततीय गुण वाले समग्र में ततीय क्रम की आवत्ति ABC, abc, ABc, Abc, aBc, aBC, abC, AbC ज्ञात करने के लिए दो गुण आवत्ति में से तीन गुण घटाएंगे। ध्यान रहे जो गुण तीसरे गुण तीसरे गुण आवत्ति में होना चाहिए। जैसे : -

$$AB = AB(C + c) = ABC + ABc$$
$$\text{and} \quad ABC = AB - ABc$$
$$\backslash \quad ABc = AB - ABC$$

Simlarly,

$$aBC = BC - ABC$$
$$aBc = aB - aBC$$
$$Abc = Ab - AbC$$
$$abc = ab - abC$$
or
$$abc = ab - aBc$$
$$AbC = AC - ABC$$
$$abC = ab - abc$$

Illustration 3 : Find out the running frequencies from the following information.

(ABC) = 75,	**(A*bc*) = 106**
(*a*B*c*) = 98,	**(*ab*C) = 74,**
(AB*c*) = 310,	**(A*bc*) = 489,**
(*a*B*c*) = 702,	**(*abc*) = 4815.**

Solution :-

In this problem, we are given the frequencies of ultimate classes (3rd order). First we will calculate the frequencies of 2nd order classes, then of the Ist order classes and in the last the value of N.

Frequencies of 2nd order classes

$$(AB) = (ABC) + (ABC) = 075 + 310 = 385$$
$$(Ab) = (AbC) + (Abc) = 106 + 489 = 595$$
$$(AC) = (ABc) + (AbC) = 075 + 106 = 181$$
$$(Ac) = (ABC) + (Abc) = 310 + 489 = 799$$
$$(aB) = (aBC) + (aBc) = 098 + 702 = 800$$
$$(ab) = (abC) + (abC) = 074 + 4815 = 4889$$
$$(aC) = (aBC) + (abC) = 098 + 074 = 172$$
$$(ac) = (aBC) + (abc) = 702 + 4815 = 5517$$
$$(Bc) = (ABC) + (aBc) = 075 + 098 = 173$$
$$(Bc) = (ABc) + (aBc) = 310 + 702 = 1012$$
$$(bC) = (AbC) + (abc) = 106 + 074 = 180$$
$$(bc) = (Abc) + (abc) = 489 + 4815 = 5304$$

Frequencies of 1st order class

$$(A) = (AB) + (Ab) = 385 + 595 = 980$$
or $$(A) = (AC) + (Ac) = 81 + 799 = 980$$
$$(a) = (aB) + (ab) = 800 + 4889 = 5689$$
or $$(a) = (aC) + (ac) = 172 + 5517 = 5689$$
$$(B) = (AB) + (aB) = 595 + 800 = 1185$$
or $$(B) = (BC) + (Bc) = 173 + 1012 = 1185$$
$$(b) = (Ab) + (ab) = 285 + 4889 = 5484$$
or $$(b) = (bC) + (bC) = 180 + 5304 = 5484$$
$$(C) = (Ac) + (bC) = 181 + 172 = 353$$
or $$(C) = (Bc) + (bC) = 173 + 180 = 353$$
$$(c) = (Ac) + (Ac) = 799 + 5517 = 6316$$
or $$(c) = (Bc) + (bc) = 1012 + 5304 = 6316$$

Frequency of Zero order

$$N = (A) + (a) = 980 + 5689 = 6669$$

or $(B) + (b) = 1185 + 5484 = 6669$

or $(C) + (c) = 353 + 6316 = 6669$

Illustration 6 : Given the following frequencies of the positive classes find out remaining class frequencies.

(N) = 12000		**(AB) = 453**	
(A) = 977		**(AC) = 284**	
(B) = 1185		**(BC) = 250**	
(C) = 596		**(ABC) = 127**	

Solution :

We have to find out :

(a), (b), (c), (Ab), (aB), $(a\ b)$, (Ac), (aC), (ac), (Bc), (bC), (bc), (ABc), (AbC), (Abc), (aBC), (aBC), (aBc), (abC), (abc)

$$(a) = N - (A) = 12000 - 977 = 11023$$

$$(b) = N - (B) = 12000 - 1185 = 10815$$

$$(c) = N - (C) = 12000 - 596 = 11404$$

$$(ABc) = (AB) - (ABC) = 453 - 127 = 326$$

$$(AbC) = (AC) - (ABC) = 284 - 127 = 157$$

$$(aBC) = (BC) - (ABC) = 250 - 127 = 123$$

$$(Abc) = (Ab) - (AbC) = (A) - (AB) - (AbC) = 977 - 453 - 157 = 367$$

$$(aBc) = (aB) - (aBC) = (B) - (AB) - (aBC) = 1185 - 453 - 123 = 609$$

$$(abC) = (bC) - (AbC) = (C) - (BC) - (AbC) = 596 - 250 - 157 = 189$$

$$(Ab) = (AbcC + (Abc) = 157 + 367 = 524$$

$$(aB) = (aBC) + (aBc) = 123 + 609 = 732$$

$$(Ac) = (ABc) + (Abc) = 326 + 367 = 693$$

$$(aC) = (aBC) + (abC) = 123 + 189 = 312$$

$$(Bc) = (ABc) + (aBc) = 326 + 609 = 935$$

$$(bC) = (AbC) + (abC) = 157 + 189 = 346$$

$$(abc) = (ab) - (abC) = (b) - (Ab) - (abC) = 10815 - 524 - 189 = 10102$$

$$(ab) = (abC) + (abc) = 189 + 10102 = 10291$$

$$(ac) = (aBc) + (abc) = 606 + 10102 = 10711$$

$$(bc) = (Abc) + (abc) = 367 + 10102 = 10469$$

रिक्त आवत्तियां हम 27 वर्गीय तालिका (27 Squared Table) से निकाल सकते हैं। उदाहरण 6 के लिये निम्नलिखित तालिका का प्रयोग कर सकते हैं :-

	A			***a***			**Total**		
	B	***b***	**Total**	**B**	***b***	**Total**	**B**	***b***	**Total**
C	(ABC) 127	(A*b*C) 157	(AC) 284	(*a*BC) 123	(*abc*) 189	(*a*C) 321	(BC) 250	(*b*C) 346	(C) 596
c	(ABC) 326	(A*bc*) 367	(A*c*) 693	(*a*B*c*) 609	(*abc*) 10102	(*ac*) 10711	(B*c*) 935	(*bc*) 10469	(*c*) 11404
Total	(AB) 453	(A*b*) 524	(A) 977	(*a*B) 732	(*ab*) 10291	(*a*) 11023	(B) 1185	(*b*) 10815	N 12000

There were 400 students in the B.Com class of a University. Their results in the various terminal examinations are given below :-

I Terminal — passed 180
II Terminal — passed 140
III Terminal — passed 180

60 passed in all the terminals;
80 failed in all the three;
40 passed in the Ist and IIIrd terminals and failed in the IIIrd;
70 failed in the Ist and IInd terminals and passed in the IIIrd;
Find out how many students passed atleast two examinations.

Solution :

Let Success in Ist terminal be denoted by A and failure by '*a*'
Success in IInd terminal be denoted by B and failure by '*b*'
Success in IIIrd terminal be denoted by C and failure by '*c*'

Then the given data

N = 400
(A) = 180
(B) = 140
(C) = 180
(ABC) = 60
(*abc*) = 80
(AB*c*) = 40
(*ab*C) = 70

We have to find out the values of :-

(ABC) + (AB*c*) + (A*b*C) + (*a*BC)

Now —

(*a*BC) + (A*b*C) + (ABC) + (*ab*C) = C
(*a*BC) + (A*b*C) = (C) – (ABC) – (*ab*C)
= 180 – 60 – 70 = 50
(ABC) + (AB*c*) + (A*b*C) + (*a*BC)
60 + 40 + 50 = 150

Thus, the number of students who passed atleast two examinations is 150.

Illustration 8 : In a war between White and Red Forces there are more Red Soldiers than white; there are more armed Whites than unarmed Reds, there are fewer armed Reds with ammunition than unarmed Whites without ammunition. Show that there are more armed Reds without ammunition than unarmed Whites and ammunition.

Let 'A' stand for a White soldier and '*a*' for Red Soldier

'B' stand for armed and '*b*' for unarmed

and 'C' stand for possession of ammunition.

and '*c*' for non-possession of ammunition.

Then the data can be denoted as

$$(a) > (\mathrm{A}) \qquad (a)$$

$$(\mathrm{AB}) > (ab) \qquad (b)$$

$$(\mathrm{A}bc) > (a\mathrm{BC}) \qquad (\mathrm{c})$$

Whe have to prove that $(a\mathrm{B}c) > (\mathrm{A}bc)$

From (*a*) considering that dichotomy of each side according to 'B' we have.

$$(a) > (\mathrm{A})$$

or $$(\mathrm{aB}) + (ab) > (\mathrm{AB}) + (\mathrm{A}b)$$

As (AB) is greater than (*ab*) as given in (*b*) if (*ab*) is substituted for (AB) then it will become still greater

or $$(a\mathrm{B}) + (ab) > (ab) + (\mathrm{A}b)$$

or $$(a\mathrm{B}) > (\mathrm{A}b) \qquad \{(ab) \text{ is common}\}$$

From this, considering the dichotomy of each side according to C we have.

$$(a\mathrm{BC}) + (a\mathrm{B}c) > (\mathrm{A}b\mathrm{C}) + (\mathrm{A}bc)$$

As (A*bc*) is greater then (*a*BC) as given in (*c*) then if (*a*BC) is substituted for (A*bc*) then it will become still greater —

$$(a\mathrm{BC}) + (a\mathrm{B}c) > (\mathrm{A}b\mathrm{C}) + (a\mathrm{BC})$$

hence $$(a\mathrm{B}c) > (\mathrm{A}b\mathrm{C}) \qquad \{(a\mathrm{BC}) \text{ is common}\}$$

समंकों की सगतिता

(Consistency of Data)

समंकों की संगतित की जाँच करने के लिए सभी वर्गों की आवत्तियों का परीक्षण करके यह देखा जाता है कि उनमें से किसी वर्ग की आवत्ति ऋणात्मक (Negative) है या नहीं। यदि किसी भी वर्ग की आवत्ति ऋणात्मक नहीं पाई जाती है जो समंक संगत (Consistent) होते हैं। दूसरी ओर, यदि किसी भी वर्ग की आवत्ति ऋणात्मक पायी जाती है जो संकलित समंक असंगत (Inconsistent) होंगे। समंकों को असंगत होना अवश्य ही इस बात का द्योतक है कि समंकों में कही त्रुटि अवश्य है। समंकों की संगतिता का पता लगाने के लिए अन्तिम वर्गों की आवत्तियाँ ज्ञात करनी चाहिए , क्योंकि एक भी अंतिम वर्ग की आवत्ति ऋणात्मक होने पर असंगत होना आवश्यक है।

समंकों की असमंति का पता लगाने के लिए निम्नलिखित सूत्रों से जाँच की जाती है :-

एक गुण (Attribute) का अध्ययन करने पर

(1) $(\mathrm{A}) \lhd \mathrm{O}$ otherwise ('A') will be negative

(2) (A)   A otherwise (*ab*) will be negative because $\mathrm{N} = (\mathrm{A}) + (a)$

दो गुणों (Attributes) का अध्ययन करने पर

(3) $(\mathrm{AB}) \lhd \mathrm{O}$ otherwise (AB) will be negative

(4) $(\mathrm{AB}) \rhd \mathrm{A}$ otherwse (*ab*) will be negative
because $(\mathrm{A}) = (\mathrm{AB}) + (\mathrm{A}b)$

(5) $(\mathrm{AB}) \rhd \mathrm{B}$ otherwise (*a*B) will be negative
because $(\mathrm{B}) = (\mathrm{AB}) + (a\mathrm{B})$

(6) (AB) $\not<$ (A) + (B) – N otherwise (*ab*) will be negative

$$\begin{aligned}
\text{because} \quad (ab) &= (a) - (aB) \\
\text{or} \quad (ab) &= [N - (A)] - [(B) - (AB)] \\
\text{or} \quad (ab) &= N - (A) - (B) + (AB) \\
\text{or} \quad -(AB) &= N - (A) - (B) - (ab) \\
\text{or} \quad (AB) &= -N + (A) + (B) - (ab) \\
\text{or} \quad (AB) &= (A) + (B) - N + (ab)
\end{aligned}$$

If (AB) is less than (A) + (B) – N it is obvious that (*ab*) will be negative.

तीन गुणों (Attributes) का अध्ययन करने पर

(7) (ABC) $\not<$ O otherwise (ABC) will be negative

(8) (ABC) $\not<$ (AB) + (BC) – (A) otherwise (A*bc*) will be negative.

$$\begin{aligned}
\text{because} \quad (Abc) &= (Ab) - (AbC) \\
\text{or} \quad (Abc) &= (A) - (AB) - \{(AC) - (ABC)\} \\
\text{or} \quad (Abc) &= (A) - (AB) - (AC) + (ABC) \\
\text{or} \quad -(ABC) &= (A) - (AB) - (AC) - (Abc) \\
\text{or} \quad (ABC) &= (AB) + (AC) - (A) + (Abc)
\end{aligned}$$

Now if (ABC) is less than (AB) + (AC) – (A) it is obvious that (A*bc*) will be negative.

(9) (ABC) $\not<$ (AB) + (BC) – (B) otherwise (*a*B*c*) will be Negative

(10) (ABC) $\not<$ (BC) + (AC) – (C) otherwise (*a*B*c*) will be Negative

(11) (ABC) $\not>$ (AB) otherwise (AB*c*) will be negative

because (AB) = (ABC) + (AB*c*)

(12) (ABC) $\not>$ (AC) otherwise (A*b*C) will ne negative

because (AC) = (ABC) + (A*b*C)

(13) (ABC) $\not>$ BC otherwise (*a*BC) will be negative

because (BC) = (ABC) × (*a*BC)

(14) (ABC) $\not>$ (AB) + (BC) + (AC) – (A) – (B) – (C) + N

otherwise (*abc*) will be negative

$$\begin{aligned}
\text{because} \quad (abc) &= (ab) - (abC) \\
\text{or} \quad (abc) &= (a) - (aB) - \{(bC) - (AbC)\} \\
\text{or} \quad (abc) &= (a) - (aB) - (bC) + (AbC) \\
\text{or} \quad (abc) &= \{N - (A)\} - \{(B) - (AB)\} - \{(C) - (BC)\} + (AC) - (ABC)\} \\
\text{or} \quad (abc) &= N - (A) - (B) + (AB) - (C) + (BC) + (AC) - (ABC) \\
\text{or} \quad (ABC) &= (AB) + (AC) + (BC) - (A) - (B) - (C) + N - (abc)
\end{aligned}$$

Now if (ABC) is more than (AB) + (AC) + (BC) – (A) – (B) – (C) + N, it is obvious that (*abc*) will be negative.

From the above rules, it is clear that when three attributed are studies then

(ABC) $\not<$ O

(ABC) $\not>$ (AB) + (AC) + (BC) – (A) – (B) – (C) + N

The upper limit of (ABC) cannot be less than the lower limit. Therefore,

(AB) + (AC) + (BC) – (A) – (B) – (C) + N $\not<$ O

or(AB) + (AC) + (BC) $\not<$ (A) + (B) + (C) – N

(15) (AB) + (AC) + (BC) $\not<$ (A) + (B) + (C) – N

Similarly if we combine the other sets of lower and upper limit then

$(ABC) \triangleleft (AB) + (AC) - A$

$(ABC) \triangleright (BC)$

Hence $\quad (AB) + (AC) - (A) \triangleright (BC)$

or $\quad (AB) + (AC) - (BC) \triangleright A$

(16) $\quad (AB) + (AC) - (BC) \triangleright A$

(17) $\quad (AB) - (AC) + (BC) \triangleright B$

(18) $\quad -(AB) + (AC) + (BC) \triangleright C$

(19) $\quad (AB) + (AC) - (A) \triangleright (BC)$

(20) $\quad (AB) + (BC) - (B) \triangleright (AC)$

(21) $\quad (AC) + (BC) - (C) \triangleright (AB)$

Illustration 9 : If a report give the following frequencies as acutally observed, show that there must be misprint or mistake of some sort, and that possibly the misprint in the dropping of a '1' before the 85 given as the frequency (BC).

N = 1000

(A) = 510	**(AB) = 189**
(B) = 490	**(AC) = 140**
(C) = 417	**(BC) = 85**

Solution :

$(AB) + (AC) + (BC) \triangleleft (A) + (B) + (C) - N$

or $\quad (BC) \triangleleft (A) + (B) + (C) - (AB) - (AC) - N$

or $\quad (BC) \triangleleft 510 + 490 + 427 - 189 - 140 - 1000$

or $\quad (BC) \triangleleft 98$

but (BC) = 85 < 98, therefore it is not the correct value of (BC). The value cannot be less than 98. If the figure is made 185 then all the conditions are fulfilled.

Illustration 10 : A market investigator returns the following data of 1000 people consulted. 811 liked chocolates 752 liked toffee and 418 liked boiled sweets; 570 liked chocolated and toffee, 356 liked chocolated and boiled sweets and 348 liked toffee and boiled sweets; 297 liked all three. Show that this information as its stands must be incorrect.

Solution :

Let A be denoted for liking of chocolates.

B be denoted for liking of toffee

C be denoted for liking of boiled sweets; then the data as returned woule be :

N = 1000,	(B) = 752,	(C) = 418
(A) = 811,	(AC) = 356,	(BC) = 348
(AB) = 570	(ABC) = 297	

Now

$(ABC) \triangleright (AB) + (AC) + (BC) - (A) - (B) - (C) + N$

or $\quad (ABC) \triangleright 570 + 356 + 348 + - 811 - 752 - 418 + 1000$

or $\quad (ABC) \triangleright 293$

But (ABC) is greater than 293, hence data are inconsistent.

Illustration 11 : Test the inconsistency of data by inequality method.

गुण संगति की जाँच असमानता पद्धति से करो।

N = 1000, A = 500, B = 550, C = 450, AB = 200,

BC = 250, AC = 150 and ABC = 120

Solution.

There are 12 condition to test if there are 3 attributes. Since all the positive values of one, two and three attributes have been given so we can test all the 12 conditions in this data.

Condition I

$ABC \triangleleft 0$ or $120 \triangleleft 0$ satisfied

Condition II

$ABC \triangleright AB$ or $120 \triangleleft 200$ satisfied

Condition III

$ABC \triangleright BC$ or $120 \triangleleft 250$ satisfied

Condition IV

$ABC \triangleright AC$ or $120 \triangleleft 150$ satisfied

Condition V

$ABC \triangleright AB + BC + AC - A - B - C + N$

$120 \triangleright 200 + 250 + 150 - 500 - 550 - 450 + 1000$

or $120 \triangleright 100$ not satisfied

Condition VI

$ABC \triangleleft AB + AC - A$

$120 \triangleleft 200 + 150 - 500$

or $120 \triangleleft -150$ satisfied

Condition VII

$ABC \triangleleft AB + BC - B$

$120 \triangleleft 200 + 250 - 550$

or $120 \triangleleft -100$ satisfied

Condition VIII

$ABC \triangleleft AB + BC - C$

$120 \triangleleft 200 + 250 - 450$

or $120 \triangleleft -50$ satisfied

Condition IX

$AB + AC + BC \triangleleft A + B + C + N$

$200 + 150 + 250 \triangleleft 500 + 550 + 450 + 1000$

or $600 \triangleleft 500$ satisfied

Condition X

$AB + AC - BC \triangleright A$

$200 + 150 - 250 \triangleright 500$

or $150 \triangleright 500$ satisfied

Condition XI

$AB + BC - AC \triangleright B$

$200 + 250 - 150 \triangleright 550$

or $250 \triangleright 500$ satisfied

Condition XII

$AC + BC - AB \triangleright C$

$150 + 250 - 200 \triangleright 450$ satisfied

Data is in inconsistent as condition No. V, is not satisfied.

$ABC \triangleright AB + BC + AC - A - B - C + N$

गुणसम्बन्ध
(Association of Attributes)

सांख्यिकी विज्ञान में गुणसम्बन्ध (Association) शब्द का प्रयोग एक विशिष्ट अर्थ में होता है। दो गुणों (Attributes) के मध्य गुणसम्बन्ध का होना उसी दशा में कहा जाएगा जबकि वे स्वतंत्र होते, इस प्रत्याशा (Expectation) से अधिक बार एक साथ प्रकट हों। यदि दो गुणों — A तथा B में किसी प्रकार का गुणसम्बन्ध नहीं है तो A गुण B में उसी अनुपात में होगा जिस अनुपात में वह non-B में है।

दो गुणों के मध्य गुणसम्बन्ध ज्ञात करने के लिए यह आवश्यक होता है कि उनके एक साथ प्राप्त होने (Simultaneous Occurence) की प्रत्याशा (Expectation) का पता लगाया जाय।

गुण-सम्बन्ध के प्रकार
(Types of Association)

गुणों के मध्य निम्न तीन प्रकार का सम्बन्ध हो सकता है :-

(1) **धनात्मक गुणसम्बन्ध अथवा साहचर्य (Positive Association or Association) :** जब दो गुण साथ-साथ उपस्थिति हों अथवा साथ-साथ अनुपिस्थित हों तो उनमें **धनात्मक गुण-सम्बन्ध अथवा साहचर्य** का होना कहा जाता है। धनात्मक गुण-सम्बन्ध की स्थिति में, **वास्तविक आवत्ति** प्रत्याशित आवत्ति से अधिक होती है। इस प्रकार का गुण -सम्बन्ध धूम्रपान एवं कैंसर, शिक्षा एवं रोजगार आदि गुणों में पाया जाता है।

(2) **ऋणात्मक गुण-सम्बन्ध अथवा असाहचर्य (Negative Association or Disassociation) :** जब एक गुण की उपस्थिति के कारण दूसरा गुण अनुपस्थित पाया जाए तो उनमें **ऋणात्मक गुण-सम्बन्ध अथवा असाहर्च** का होना कहा जाता है। ऋणात्मक गुण-सम्बन्ध की स्थिति में **वास्तविक आवत्ति** प्रत्याशित आवत्ति से कम होती है।

(3) **स्वतंत्र गुण (Independence Attributes) :** जब दो गुण साथ-साथ उपस्थित अथवा अनुपस्थित न हों तथा एक की उपस्थिति के साथ दूसरे गुण की अनुपस्थिति भी न हो, तो इन्हें **स्वतंत्र गुण** कहा जाता है। ऐसी स्थिति में, **वास्तविक आवत्ति** प्रत्याशित आवत्ति के बराबर होती है।

विशेषताएँ/लक्षण (Characteristics) : गुण सम्बन्ध की प्रमुख विशेषताएँ निम्न हैं :-

(1) गुण-सम्बन्ध की प्रमुख विशेषता-लक्षण दो गुणों के साथ-साथ **उपस्थित** अथवा **अनुपस्थित** होना या एक गुण की उपस्थिति के साथ दूसरे गुण का अनुपस्थित होना है।

(2) गुण-सम्बन्ध के अध्ययन के लिए दो गुणों में **आर्यकारण सम्बन्ध** (Causal Connection) होना आवश्यक है। किसी कार्य पर प्रस्थान करते समय बिल्ली का रास्ता काटने का या छींक आ जाने को असफलता का कारण नहीं माना जा सकता, क्योंकि इनमें कोई कार्यकारण सम्बन्ध विद्यमान नहीं है, यह केवल अन्धविश्वास है।

(3) गुण-सम्बन्ध में हर समय तुलनात्मक अध्ययन होता है। केवल इस आधार पर कि दो गुण परस्पर अधिक अनुपात में एक साथ पाए जाते हैं, वह निष्कर्ष कि दोनों में गुण-सम्बन्ध है, प्राप्त नहीं किया जा सकता।

''चेचक का टीका लगवाने वालों में से 90 प्रतिशत व्यक्तियों को चेचक का प्रकोप नहीं हुआ, अतः चेचक के प्रकोप से बचने के लिए चेचक का टीका लगवाना उपयोगी व प्रभावशाली है।''

चेचक का टीका लगवाने वालों में से 90 प्रतिशत व्यक्तियों को चेचक का प्रकोप नहीं हुआ, केवल इस आधार पर टीका लगवाने की उपयोगिता को सिद्ध नहीं किया जा सकता, इसके लिए हमें टीका न लगवाने वालों में से उन व्यक्तियों का प्रतिशत ज्ञात करना होगा जिन पर चेचक का प्रकोप नहीं हुआ, इसके बाद तुलनात्मक अध्ययन से ही टीका लगवाने और उपयोगिता को सिद्ध किया जा सकता है।

गुण-सम्बन्ध निर्धारण की विधियाँ
(Methods of Determining Association)

दो गुणों के मध्य सम्बन्ध है या नहीं, यह ज्ञात करने की विधियाँ निम्न हैं :-

(1) वास्तविक तथा प्रत्याशित आवत्तियों की तुलना विधि (Comparison of Observed and Expected Frequencies Method)

(2) अनुपात तुलना विधि (Comparison of Proportion Method)

(3) यूल का गुण-सम्बन्ध गुणांक (Yule's Coefficient of Association)

(4) सम्बद्ध (तथ्यानुबन्धन) गुणांक (Coefficient of Colligation)

(5) सम्भाव्यता गुणांक (Coefficient of Contingency)

(1) वास्तविक तथा प्रत्याशित आवत्तियों की तुलना विधि
(Comparison of Observed and Expected Frequencies Method)

जब दो गुणों के इस मध्य इस विधि से सम्बन्ध ज्ञात किया जाता है तब उसकी वास्तविक आवत्ति की तुलना प्रत्याशित आवत्ति से की जाती है। दो गुणों की वास्तविक एक साथ उपस्थिति या अनुपस्थिति, प्रत्याशित अनुपात से अधिक अनुपात में होनी चाहिए। अतः दो गुणों के मध्य गुणसम्बन्ध ज्ञात करने के लिए यह आवश्यक होता है कि उनके एक साथ प्राप्त होने (Simultaneous Occurence) की प्रत्याशा (Expectation) का पता लगाया जाय।

सम्भावना सिद्धान्त (Theory of Probability) के अनुसार किसी घटना के घटने की सम्भावना :-

$$= \frac{\text{Number of favourable cases}}{\text{Total number of cases}} \quad \text{Number of observations}$$

सूत्र के आधार पर ज्ञात की जा सकती है। यदि किसी सिक्के को उछाला जाय तो उसके चित्त (Head Upward) गिरने की सम्भावना $\frac{1}{2}$ होती है। यदि उस सिक्के को 100 बार उछाला जाय तो उसके चित्त गिरने की प्रत्याशा $\frac{1}{2} \times 100 = 50$ होगी। इस नियम के आधार पर

(A) की सम्भावना $= \frac{(A)}{N}$, (B) की सम्भावना $= \frac{(B)}{N}$

(B) तथा (B) की संयुक्त $= \frac{(A)}{N} \times \frac{(B)}{N}$

तथा (A) व (B) के 'N' बार में एक साथ होने की प्रत्याशा

$$= \frac{(A)}{N} \times \frac{(B)}{N} \times N \text{ or } \frac{(A)\quad(B)}{N}$$

इसी प्रकार (ab) की प्रत्याशा

इसी ढंग से अन्य वर्गों की प्रत्याशित आवत्तियाँ ज्ञात की जा सकती हैं। गुणसम्बन्ध धनात्मक (Positive) अथवा ऋणात्मक (Negative) हो सकता है। A व B दो गुणों में धनात्मक गुणसम्बन्ध होगा, यदि (AB) की वास्तविक आवत्ति प्रत्याशित आवत्ति

से अधिक हो। यदि (AB) की वास्तविक आवत्ति प्रत्याशित आवत्ति से कम है तो इन दोनों गुणों के माध्य ऋणात्मक गुणसम्बन्ध होगा। यदि (AB) की वास्तविक आवत्ति प्रत्याशित आवत्ति के समान ही हो तो दोनों गुण स्वतंत्र कहे जाते हैं।

निम्न तालिका से यह स्पष्ट हो जायेगा :-

Attributes Independent	**Positive Association**	**Negative Association or Dissassociation**
A and B $(AB) = \frac{(A)\quad(B)}{N}$	$(AB) >$	$(AB) <$
A and *b* $(Ab) =$	$(Ab) > \frac{(A)\quad(b)}{N}$	$(Ab) < \frac{(A)\quad(b)}{N}$
a and B $(aB) = \frac{(a)\quad(B)}{N}$	$(aB) > \frac{(a)\quad(B)}{N}$	$(aB) < \frac{(a)\quad(B)}{N}$
a and *b* $(ab) = \frac{(a)\quad(b)}{N}$	$(AB) > \frac{(a)\quad(b)}{N}$	$(AB) < \frac{(a)\quad(b)}{N}$

यदि दो गुणों के अध्ययन में अन्तिम वर्गों की आवत्तियाँ ज्ञात हों तो निम्नलिखित सूत्र के द्वारा उनमें गुणसम्बन्ध ज्ञात किया जा सकता है :-

$$(AB) \times (ab) = (Ab) \times (aB)$$

यदि यह समान हाते हैं तो दोनों गुणों के मध्य गुणसम्बन्ध का अभाव है। यदि $(AB) \times (ab) > (Ab) \times (aB)$, तो दोनों गुणों के मध्य धनात्मक गुणसम्बन्ध है। $\frac{(A)\quad(B)}{N}$

यदि $(AB) \times (ab) < (Ab) \times (aB)$ तो दोनों गुणों के मध्य ऋणात्मक गुणसम्बन्ध है। इसको इस प्रकार सिद्ध किया जा सकता है :-

$$(AB) \times (ab) = (Ab) \times (aB)$$

or $$\frac{(A)\quad(B)}{N} \times \quad = \frac{(A)\quad(b)}{N} \times \frac{(a)\quad(B)}{N}$$

or $$\frac{(A)\quad(B)\quad(a)\quad(b)}{N\quad N} = \frac{(A)\quad(b)\quad(a)\quad(B)}{N\quad N}$$

or $$\frac{(A)\quad(B)\quad(a)\quad(b)}{N\quad N} = \frac{(A)\quad(b)\quad(a)\quad(b)}{N\quad N}$$

Illustration 12 : In an anti-malaria campaign in a certain area, quinine was administered to 812 persons out of a total population of 3,248.

The number of fever cases is shown below :-

Treatment	**Fever**	**No Fever**
Quinine	20	792
No-Quinine	220	2216

Discuss the usefulness of quinine in checking malaria.

Solution :

Denoting A for quinine treatment,

a for No Quinine treatment,

B for No attack of fever,

b for attack of fever.

Then :

$$(AB) = 792, (Ab) = 20, (aB) = 2216, (ab) = 220$$

To test asssociation :

$$(AB) \times (ab) = (Ab) \times (aB)$$

$$(7920 \times 220) = (20 \times 2216)$$

$$174240 > 44320$$

अतः A तथा B में धनात्मक सहसम्बन्ध है, इसका आश्य है कि कुनने मलेरिया रोकने में सहायक है।

Illustration 13 : नीचे दिये गये समंकों से ज्ञात कीजिए कि A तथा B स्वतंत्र हैं, उनमें धनात्मक या ऋणात्मक गुणसम्बन्ध हैं :-

Find if A and B are independent, positively associated or negatively associated from the data given below:

(A) = 470, (B) = 620, (AB) = 320, N = 1000

Solution : Attributes A and B shall be called independent if

$$(AB) = \frac{(A)(B)}{N}$$

$$\frac{470 \quad 620}{1000}$$

$$\left[\frac{(AB)}{(A)} \quad 100\right]$$

$$(AB) = 320, (A) = 470, (B) = 620, N = 1000$$

$$\text{Expectation of } (AB) = \qquad = 291.4$$

Since (AB) acutal observation (320) is more than the expectation (291.4) attributes A and B are positively associated.

(2) अनुपात तुलना विधि

(Comparison of Proportion Method)

इस विधि के अनुसार दो प्रतिशतों की गणना की जाती है :-

(*i*) B की A में उपस्थिति या अनुपस्थिति (Presence or absence of B in A) :

(*ii*) B की Non-A में उपस्थिति या अनुपस्थिति (Presence or absence of B in Non-A) :

$$\left[\frac{(\ B)}{(\ )} \quad 100\right]$$

A व B में गुणसम्बन्ध ज्ञात करने के लिए इन दो प्रतिशतों में तुलना की जाती है। यदि हम B_s का A_s में वही अनुपात पाते हैं तो non-A(a_s) में है, तो दोनों गुण (A तथा B) स्वतंत्र माने जाते हैं। यदि B_s का A_s में अनुपात non-A_s (a_s) में अनुपात से अधिक होता है तो दोनों गुणों (A तथा B) में धनात्मक गुणसम्बन्ध माना जाता है। इसके विपरीत यदि B_s का A_s में अनुपात non-A_s (a_s) में अनुपात से कम होता है तो A तथा B में ऋणात्मक गुणसम्बन्ध होता है।

	गुण (Association)	**B का B तथा a में अनुपात (Proportion of B in A and a)**	**A का B तथा b में अनुपात (Proportion of A in B and b)**
(*i*)	स्वतंत्र (independent)	$\frac{(AB)}{(A)} \quad 100 = \frac{(\ B)}{(\ )} \quad 100$	$\frac{(AB)}{(B)} \quad 100 = \frac{(A\ )}{(\ )} \quad 100$
(*ii*)	धनात्मक (Positive)	$\frac{(AB)}{(A)} \quad 100 > \frac{(\ B)}{(\ )} \quad 100$	$\frac{(AB)}{(B)} \quad 100 > \frac{(A\ )}{(\ )} \quad 100$
(*iii*)	ऋणात्मक (Negative)	$\frac{(AB)}{(A)} \quad 100 < \frac{(\ B)}{(\ )} \quad 100$	$\frac{(AB)}{(B)} \quad 100 < \frac{(A\ )}{(\ )} \quad 100$

Illustration 14 : Out of 500 students, 200 were married. Total students failed were 150 out of which 50 were married. Find whether the attributes marriage and failure are positively or negatively associated.

Solution :

Let A denoted married, $\frac{100}{300}$

\ (*a*) represents unmarried

Let B denoted failure students

\ *b* represents passed students

\ We are given :-

N = 500, A = 200, B = 150, AB = 50

Using proportion method, we have proportion of married students failed.

$$\frac{AB}{A} = \frac{50}{200} = .25$$

Proportion of Unmarried who failed.

We need

$$aB = B - AB = 150 - 50 = 100$$

$$a = N - A = 500 - 200 = 300$$

\ $$\frac{aB}{a} = \quad = .33$$

∵ $$< \frac{aB}{a}$$

i.e. $$.25 < .33$$

We can conclude thate there is negative association or disassociation between marriage and failure.

Illustration 15. किसी इलाके में 1000 साक्षरों में से 10 अपराधी हैं। उसी इलाके में 20,000 निरक्षरों में से 1500 अपराधी हैं। क्या आपको निरक्षरता व अपराध में कोई सम्बन्ध प्रतीत होता है ?

Out of 1000 literates in a particular area 10 were criminal. Out of 20,000 illiterates in the same area 1500 were criminals. Do you find association between illiteracy and criminality ?

Solution :

Let	'A'	= Illiteracy
\	a	= Literacy
Let	B	= Criminals
\	b	= Non-criminals
(Given)	A	= 20,000 a = 1000
	AB	= 1500, aB = 10

Using proportion method, we have

$$= \frac{1500}{20,000} = .75$$

$$\because \quad = \quad = .01$$

$$\because \quad < .\frac{aB}{a}$$

Hence, there is positive association between illiteracy and criminality.

(3) गुणसम्बन्ध-गुणक
(Coeffcient of Association)

उपर्युक्त वर्णित विधियों से दो गुणों के मध्य सम्बन्ध का केवल अनुमान ही लगाया जा सकता है परन्तु गुणसम्बन्ध की मात्रा नहीं जानी जा सकती। गुणसम्बन्ध की मात्रा ज्ञात करने के लिए प्रोफेसर यूल (Prof. Yule) ने एक सूत्र का प्रतिपादन किया है। जिससे 'गुणसम्बन्ध गुणक' (Coefficient of Association) ज्ञात किया जाता है। यह सूत्र है :-

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)}$$

गुणसम्बन्ध गुणक (Q) का मान ± 1 के अन्तर्गत आता है तथा उसका निर्वाचन ठीक उसी प्रकार किया जाता है जिस प्रकार कार्ल पियर्सन के सहसम्बन्ध गुणक (Coefficient of Correlation) का निर्वाचन किया जाता है। यदि गुणसम्बन्ध गुणक '0' होता है तो दोनों गुणों के मध्य गुणसम्बन्ध का अभाव होता है। गुणक + 1 होने पर दोनों गुणों के मध्य पूर्ण धनात्मक गुणसम्बन्ध (Perfect Positive Association) तथा –1 होने पर दोनों गुणों के मध्य पूर्ण ऋणात्मक गुणसम्बन्ध (Perfect Negative Association) होता है।

Illustration 16 : In an assortive study to find whether tall husbands tend to marry tall wives, the following information about the wives of 125 tall and 125 short statured husbands was published.

	Tall Husbands	**Short Husbands**
Tall Wives	56%	13%
Short Wives	11%	48%

Find the coeffficient of association between the stature of wives and husbands.

Solution :

Let A denote tall husbands

a denote short husbands

B denote tall wives

b denote short wives

Then the data are :-

(AB) = 56% (A*b*) = 11%, (*a*B) = 13%, (*ab*) = 48%

$$\text{Coefficient of Association} = \frac{(\text{AB})(ab) \quad (\text{A}b)(a\text{B})}{(\text{AB})(ab) \quad (\text{A}b)(a\text{B})}$$

By substituting the values we get :-

$$\frac{(56 \quad 48) \quad (11 \quad 13)}{(56 \quad 48) \quad (11 \quad 13)} = \frac{2688 \quad 143}{1688 \quad 143} = \frac{2545}{2831} + .9$$

It shows a very high degree of positive association between stautres of husbands and wives.

Illustration 17 : A list of 1000 items classified according to three attributed is given below. A, B and C denote the presence and *a*, *b* and *c* denote absence of the attributes. Show if there is any association between A and B, and A and C.

(ABC) = 300, (AB*c*) = 200, (A*b*C) = 20, (A*bc*) = 80

(*a*BC) = 150, (*a*B*c*) = 50, (*ab*C) = 120, (*abc*) = 80

Solution :

Coefficient of Association between A and B (Q_{AB})

$$= \frac{(\text{AB})(ab) \quad (\text{A}b)(a\text{B})}{(\text{AB})(ab) \quad (\text{A}b)(a\text{B})}$$

Coefficient of Association between A and C (Q_{AC})

$$= \frac{(\text{AC})(ac) \quad (\text{A}c)(a\text{C})}{(\text{AC})(ac) \quad (\text{A}c)(a\text{C})}$$

$$(\text{AB}) = (\text{ABC}) + (\text{AB}c) = 300 + 200 = 500$$

$$(ab) = (ab\text{C}) + (abc) = 120 + 80 = 200$$

$$(\text{A}b) = (\text{A}b\text{C}) + (\text{A}bc) = 20 + 80 = 100$$

$$(a\text{B}) = (a\text{BC}) + (a\text{B}c) = 150 + 50 = 200$$

Hence

$$Q_{AB} = \frac{(500 \quad 200) \quad (100 \quad 200)}{(500 \quad 200) \quad (100 \quad 200)}$$

$$= \frac{10000 \quad 200000}{100000 \quad 200000} \quad \frac{80000}{1200000} \quad .67$$

$$(\text{AC}) = (\text{ABC}) + (\text{A}b\text{C}) = 300 + 20 = 320$$

$$(ac) = (a\text{B}c) + (abc) = 50 + 80 = 130$$

$$(\text{A}c) = (\text{AB}c) + (\text{A}bc) = 200 + 80 = 280$$

$$(a\text{C}) = (a\text{BC}) + (ab\text{C}) = 150 + 120 = 270$$

$$Q_{AC} = \frac{(320 \quad 130) \quad (280 \quad 270)}{(320 \quad 130) \quad (280 \quad 270)}$$

$$= \frac{41600 \quad 75600}{41600 \quad 75600} \quad \frac{34000}{117200} \quad 2.9$$

A + B have positive association while *a* & *c* have negative association.

Illustration 18 : Investigate the association between eye colour of husbands and eye colour of wives from the data given below :-

Husbands with light eyes and wives with light eyes = 309

Husbands with light eyes and wives with non-light eyes = 214

Husbands with non-light and wives with light eyes = 132

Husbands with non-light eyes and wives with non-light eyes = 119

Solution : Since we have to find out the association between eye colour of husbands and that of wife one attribute we would as A and another as B.

Let A denote husbands with light eyes.

\ a would denote husbands with not-light eyes

Let B denote wives with light-eyes.

\ b would denote wives with not-light eyes.

The given data in terms of these symbols is :

(AB) = 309, (Ab) = 214, (aB) = 132, (ab) = 119

Applying Yule's method : $$Q = \frac{(AB)(\quad)\quad(\quad)(\ B)}{(AB)(\quad)\quad(A\ )(\ B)}$$

Substituting the above values in the formula

$$Q = \frac{(309 \quad 119) \quad (214 \quad 132)}{(309 \quad 119) \quad (214 \quad 132)}$$

$$= \frac{8523}{65019} = 0.131$$

Thus there is very little association between the eye colour of husband and wife.

(4) सम्बद्ध-गुणक
(Coefficient of Collignation)

प्राफेसर यूल ने एक अन्य महत्वूपर्ण सूत्र प्रस्तुत किया है जिसे 'Coefficient of Collignation' कहते हैं। इसका गुणक संकेताक्षर 'y' है तथा इसका सूत्र निम्न हैं :-

$$y = \frac{1 \quad \sqrt{\dfrac{(Ab) \quad (aB)}{(AB) \quad (ab)}}}{1 \quad \sqrt{\dfrac{(Ab) \quad (aB)}{(AB) \quad (ab)}}}$$

Coefficient of Collignation (y) से Coefficient of Association (Q) की गणना भी की जा सकती है। इसका सूत्र निम्न है :-

$$Q = \frac{2y}{1 \quad y^2}$$

Illustration 19 : We are given (AB) = 56, (Ab) = 11, (aB) = 13, (ab) = 45, Calculate coefficient of colligation

Solution :

$$g = \frac{1 \quad \sqrt{\frac{11 \quad 13}{56 \quad 45}}}{1 \quad \sqrt{\frac{11 \quad 13}{56 \quad 45}}} = \frac{1 \quad \sqrt{\frac{143}{2520}}}{1 \quad \sqrt{\frac{143}{2520}}} = \frac{0.76}{1.238} \qquad + 0.61$$

'y' से गुणसम्बन्ध गुणक की गणना इस प्रकार की जायेगी :-

$$Q = \frac{2}{1 \quad ^2} \text{ or } Q = \frac{2 \quad 0.61}{1 \quad 0.16^2} = \frac{1.22}{1.3721} = +0.89$$

(5) सम्भाव्यता गुणक
(Coefficient of Contingency)

'A' व 'B' गुणों को अनेक उप-विभागों में विभाजित होने पर, उनके मध्य रूपेण गुण सम्बन्ध ज्ञात करने के लिए कार्ल पियर्सन का माध्य वर्ग सम्भाव्यता गुणक (Karl Pearson's Coeffcient of Mean Square Contingency) ज्ञात किया जाता है। इसको सम्भाव्यता गुणक भी कहते हैं और इसका संकेताक्षर 'C' है। सम्भाव्यता गुणक की गणना में हम शून्य कल्पना (Null Hypothesis) मान कर चलते हैं जिसके अनुसार हम यह जानते हैं कि दोनों गुण स्वतंत्र हैं और उनके बीच गुणसम्बन्ध नहीं है। इस मान्यता के आधार पर हम विभिन्न कोष्ठों की आवत्तियाँ (Cell Frequencies) जैसे (A_1B_1), (A_1B_2) आदि ज्ञात करते हैं। कोष्ठों की आवत्तियाँ निम्न सूत्र से ज्ञात की जाती हैं : -

$$(A_1B_1) = \frac{(A_1) \quad (B_1)}{N}, \; (A_1B_2) = \frac{(A_1) \quad (B_1)}{N} \text{ etc.}$$

यदि प्रत्येक कोष्ठ (cell) की वास्तविक आवत्ति प्रत्याशित आवत्ति के बराबर होती है , तो दोनों स्वतंत्र होते हैं। वास्तविक आवत्ति तथा प्रत्याशित आवत्ति में अन्तर होने पर दोनों गुणों में गुणसम्बन्ध होती है। दोनों गुणों के मध्य गुणसम्बन्ध की मात्रा निर्धारित करने के लिए समभाव्यता गुणक की गणना की जा सकती है। इसका सूत्र इस प्रकार है :-

$$C = \sqrt{\frac{X^2}{N \quad X^2}}$$

'C' की गणना करने के लिए 'X^2' (Pronounced as Chi-squares) का मान ज्ञात किया जाता है। 'X^2' की गणना के पद्दोष इस प्रकार हैं :-

(1) प्रत्येक कोष्ठ (Cell) की प्रत्याशित आवत्ति (Expected Frequency) 'E' की गणना की जाती है।

अवलोकित आवत्ति तालिका
(Table of Observed Frequencies)

B® / A¯	B_1	B_2	B_3	Total
A_1	(A_1B_1)	(A_1B_2)	(A_1B_3)	(A_1)
A_2	(A_2B_1)	(A_2B_2)	(A_2B_3)	(A_2)
A_3	(A_3B_1)	(A_3B_2)	(A_3B_3)	(A_3)
Total	(B_1)	(B_2)	(B_3)	N

अप्रत्याक्षित आवत्ति तालिका
(Table of Expected Frequencies)

B ® / A ¯	B_1	B_2	B_3	Total
A_1	$\frac{(A_1)\ (B_1)}{N}$	$\frac{(A_1)\ (B_2)}{N}$	$\frac{(A_1)\ (B_3)}{N}$	(A_1)
A_2	$\frac{(A_2)\ (B_1)}{N}$	$\frac{(A_2)\ (B_2)}{N}$	$\frac{(A_2)\ (B_3)}{N}$	(A_2)
A_3	$\frac{(A_3)\ (B_1)}{N}$	$\frac{(A_3)\ (B_2)}{N}$	$\frac{(A_3)\ (B_3)}{N}$	(A_3)
Total	(B_1)	(B_2)	(B_3)	N

(2) विभिन्न कोष्ठों की अवलोकित (O) एवं प्रत्याशित (E) आवत्तियों का अन्तर $(O-E)$ निकाला जाता है।

(3) अवलोकित एवं प्रत्याशित आवत्तियों के अन्तरों के वर्ग किये जाते हैं $(O-E)^2$ तथा अन्तरों के वर्गों में उनकी सम्बन्धित प्रत्याशित आवत्ति से भाग दिया जाता है :-

$$\left[\frac{(O-E)^2}{E}\right]$$

(4) इस प्रकार प्राप्त संख्याओं का योग ज्ञात किया जाता है। यह योग $\left\{\frac{(O-E)^2}{E}\right\}$ ही X^2 का मान होता है। संकेताक्षरों के अनुसारः

$$X^2 = \Sigma \left[\frac{(O-E)^2}{E}\right]$$

X^2 का मान निकाल लेने पर, सम्भाव्यता गुणक (C) की गणना की जा सकती है।

Illustration 20 : Calculate coefficient of contingency from the following table :-

PARENT

Off Spring	**Very Tall**	**Tall**	**Medium**	**Short**	**Total**
Very Tall	20	30	20	2	72
Tall	14	125	85	12	236
Medium	3	140	165	125	433
Short	3	37	68	151	259
Total	40	332	338	290	1000

(I.C.S.)

Solution :

Off Spring A	**Parent B**				**Total**
	Very Tall (B_1)	**Tall (B_2)**	**Medium (B_3)**	**Short (B_4)**	
Very Tall (A_1)	20 (A_1B_1)	30 (A_1B_2)	20 (A_1B_3)	2 (A_1B_4)	72 (A_1)
Tall (A_2)	14 (A_2B_1)	125 (A_2B_2)	85 (A_2B_3)	12 (A_2B_4)	236 (A_2)
Medium (A_3)	3 (A_3B_1)	140 (A_3B_2)	165 (A_3B_3)	125 (A_3B_4)	433 (A_1)
Short (A_4)	3 (A_4B_1)	37 (A_4B_2)	68 (A_4B_3)	151 (A_4B_4)	259 (A_4)
Total	40 (B_1)	332 (B_2)	338 (B_3)	290 (B_4)	1000 N

Expected frequencies will be calculated :-

$$(A_1B_1) = \frac{(A_1)\quad(B_1)}{N} \quad \frac{72 \quad 40}{1000} \quad 2.9$$

$$(A_1B_2) = \frac{(A_1)\quad(B_2)}{N} \quad \frac{72 \quad 332}{1000} \quad 23.9$$

$$(A_1B_3) = \frac{(A_1)\quad(B_3)}{N} \quad \frac{72 \quad 338}{1000} \quad 24.3$$

$$(A_1B_4) = \frac{(A_1)\quad(B_4)}{N} \quad \frac{72 \quad 290}{1000} \quad 20.9$$

$$(A_2B_1) = \frac{(A_2)\quad(B_1)}{N} \quad \frac{236 \quad 40}{1000} \quad 9.4$$

$$(A_2B_2) = \frac{(A_2)\quad(B_2)}{N} \quad \frac{236 \quad 332}{1000} \quad 78.4$$

$$(A_2B_3) = \frac{(A_2)\quad(B_3)}{N} \quad \frac{236 \quad 338}{1000} \quad 79.8$$

$$(A_2B_4) = \frac{(A_2)\quad(B_4)}{N} \quad \frac{236 \quad 290}{1000} \quad 68.4$$

$$(A_3B_1) = \frac{(A_3)\quad(B_1)}{N} \quad \frac{433 \quad 40}{1000} \quad 17.3$$

$$(A_3B_2) = \frac{(A_3)\quad(B_2)}{N} \quad \frac{433 \quad 332}{1000} \quad 143.8$$

$$(A_3B_3) = \frac{(A_3)\quad(B_3)}{N} \quad \frac{433 \quad 338}{1000} \quad 146.4$$

$$(A_3B_4) = \frac{(A_3)\quad(B_4)}{N} \quad \frac{433 \quad 290}{1000} \quad 125.5$$

$$(A_4B_1) = \frac{(A_4)\quad(B_1)}{N} \quad \frac{259 \quad 40}{1000} \quad 10.4$$

$$(A_4B_2) = \frac{(A_4)\quad(B_2)}{N} \quad \frac{259 \quad 332}{1000} \quad 85.9$$

$$(A_4B_3) = \frac{(A_4)\quad(B_3)}{N} \quad \frac{259 \quad 338}{1000} \quad 87.5$$

$$(A_4B_4) = \frac{(A_4)\quad(B_4)}{N} \quad \frac{259 \quad 290}{1000} \quad 75.2$$

	O	**E**	**(O – E)**	**$(O – E)^2$**	**$\frac{(O \quad E)^2}{E}$**
(A_1B_1)	20	2.9	17.1	292.41	100.83
(A_1B_2)	30	23.9	6.1	37.21	1.55
(A_1B_3)	20	24.3	– 4.3	18.49	0.76
(A_1B_4)	2	20.9	– 18.9	357.21	17.09
(A_2B_1)	14	9.4	4.6	21.16	2.25
(A_2B_2)	125	78.4	46.6	2171.56	27.70
(A_2B_3)	85	89.8	5.2	27.04	0.34
(A_2B_4)	12	68.4	– 56.4	3180.96	46.50
(A_3B_1)	3	17.3	– 14.3	204.49	11.82
(A_3B_2)	140	143.8	– 3.8	14.44	0.10
(A_3B_3)	165	146.4	18.69	345.96	2.36
(A_3B_4)	125	125.5	– 0.5	00.25	0.002
(A_4B_1)	3	10.4	– 7.4	54.76	5.26
(A_4B_2)	37	85.9	– 48.9	2391.21	27.83
(A_4B_3)	68	87.5	– 19.5	380.25	4.34
(A_4B_4)	151	75.2	75.8	5745.64	76.40
					S = x^2 = 325.132

Square Contingency or = x^2 = 325.132

$$\text{Square Contingency or C} = \quad = \sqrt{\frac{325.13}{1000 \quad 325.13}} = .495$$

'C' का निर्वचन

गुण सम्बन्ध की अनुपस्थिति की दशा में 'C' सदैव शून्य होगा, परन्तु 'C' का मूल्य कभी भी 1 नहीं होगा। 'C' का अधिकतम मान अवलोकनों के वर्गों पर निर्भर करता है। अधिकतम मान $\sqrt{\frac{t \quad 1}{t}}$ (t = उपवर्गों की संख्या) के बराबर होगा। अग्रलिखित तालिका 'C' का अधिकतम मान बनाती है।

Contingency Table	Maximum Value of 'C'	Contingency Table	Maximum Value of 'C'
2×2	0.707	7×7	0.926
3×3	0.816	8×8	0.935
4×4	0.866	9×9	0.743
5×5	0.895	10×10	9.949
6×6	0.913		

गणना से ज्ञात 'C' के मान की तुलना तालिका में दिये गये मान से की जा सकती है।

एक अन्य माप भी होती है जिसे मध्य-मार्ग सम्भाव्यता (Mean Square Contingency) अथवा f^2 (phi square) कहते हैं। इसका सूत्र निम्न है :

$$\text{Mean square contingency of } f^2 \text{ (phi square)} = \frac{X^2}{N}$$

उपर्युक्त उदाहरण में f^2 = = = .32513

आंशिक गुणसम्बन्ध

(Partial Association)

अभी तक हमने एक समग्र में दो गुणों के मध्य गुणसम्बन्ध का अध्ययन किया है तथा इस बात का बिल्कुल ध्यान नहीं रखा है कि उस समग्र में अन्य कोई गुण भी विद्यमान हो सकता है। इस प्रकार के गुणसम्बन्ध को 'पूर्ण गुणसम्बन्ध' (Total Association) कहते हैं। दो गुणों के मध्य किसी उप-समग्र (Sub-universe) से गुणसम्बन्ध ज्ञात करना 'आंशिक-गुणसम्बन्ध' (Partial Association) कहलाता है।

आंशिक गुणसम्बन्ध निम्न विधि से ज्ञात किया जा सकता है :-

A तथा B गुणों के मध्य 'C' उप-समग्र में धनात्मक गुण-सम्बन्ध होगा यदि :-

$$(ABC) > \frac{(AC) \quad (BC)}{(C)}$$

A तथा B गुणों के मध्य 'C' उप-समग्र में ऋणात्मक गुण-सम्बन्ध होगा यदि :-

$$(ABC) <$$

A तथा B गुणों के मध्य 'c' उप-समग्र में धनात्मक गुण-सम्बन्ध होगा यदि :-

$$(ABc) >$$

A तथा B गुणों के मध्य 'c' उप-समग्र में ऋणात्मक गुण-सम्बन्ध होगा यदि :-

$$(ABc) <$$

आंशिक गुणसम्बन्ध का गुणक (Coefficient of Partial Association) भी निम्न सूत्र से ज्ञात किया जा सकता है :-

A तथा B गुणों के मध्य 'C' उप-समग्र में आंशिक गुणसम्बन्ध गुणक :-

$$Q\ ABC =$$

A तथा B गुणों के मध्य 'c' उप-समग्र में आंशिक गुणसम्बन्ध गुणक :-

$$Q\ AB.c = \frac{(ABc)(abc) \quad (Abc) \quad (aBc)}{(ABc)(abc) \quad (Abc) \quad (aBc)}$$

Illustration 21 : Find coefficient of partial association between A and B in the Universe 'C' and '*c*' from the following data :-

(ABC) = 300, (AB*c*) = 200, (A*b*C) = 20, (A*bc*) = 80

(*a*BC) = 150, (*a*B*c*) = 50, (*ab*C) = 120, (*abc*) = 80

$$\frac{(ABc)(abc) \quad (Abc)(aBc)}{(ABc)(abc) \quad (Abc)(aBc)}$$

Solution :-

$$Q.AB.c = - \frac{(ABC)(abC) \quad (AbC)(aBC)}{(ABC)(abC) \quad (AbC)(aBC)}$$

$$= \frac{(300 \quad 120) \quad (20 \quad 150)}{(300 \quad 120) \quad (20 \quad 150)} = \frac{36000 \quad 3000}{36000 \quad 3000}$$

$$= \frac{33000}{39000} = .846$$

$$Q.AB.c = \frac{(ABc)(abc) \quad (Abc)(aBc)}{(ABc)(abc) \quad (Abc)(aBc)}$$

$$= \frac{(200 \quad 80) \quad (80 \quad 50)}{(200 \quad 80) \quad (80 \quad 50)} = \frac{16000 \quad 4000}{16000 \quad 4000}$$

$$= \frac{12000}{20000} = .75$$

भ्रमात्मक गुणसम्बन्ध (Illusory Association) : आंशिक गुणसम्बन्ध ज्ञात करने में कभी-कभी दो गुणों के मध्य गुणसम्बन्ध होने का भ्रम उत्पन्न हो सकता है। दो स्वतंत्र गुणों A तथा B में किसी उप-समग्र में सम्बन्ध स्थापित हो जाने से इस प्रकार का भ्रम उत्पन्न हो सकता है क्योंकि वास्तव मे उन गुणों के मध्य कोई प्रत्यक्ष सम्बन्ध नहीं होता है। अनुसन्धानकर्त्ताओं के पक्षपात, आँकड़े मिला देने, निर्वचन की त्रुटि आदि के कारण भी भ्रमात्क गुणसम्बन्ध ज्ञात हो सकते हैं।

Illustration 22 : Comment on the following statements :-

(*i*) Road accidents resulted in 4513 deaths in 1958 and 5250 in 1968, while the number of women drivers increased in the period. Hence women make bad drivers.

(*ii*) 99% of the people who drink beeer die before reaching 100 years of age. Therefore, drinking beer is bad for longivity.

(*iii*) Not cultivating land owner should be deprived of their ownership rights without payments of compensation, because 90% of the land owned by them is inherited property for which they have paid nothing.

Solution :

(*i*) यह भ्रमात्मक गुणसम्बन्ध का एक उदाहरण है। स्त्री-चालकों को उसी दशा में बुरा कहा जा सकता है जबकि प्रति स्त्री-चालक दुर्घटनाएँ प्रति पुरुष-चालक दुर्घटनाओं से अधिक हों। दुर्घटनाओं में अधिक संख्या में मत्यु के अन्य कारण भी हो सकते है जैसे नगर जनसंख्या में वद्धि अथवा सड़कों पर अधिक भीड़ -भाड़।

(*ii*) यह निष्कर्ष उसी दशा में सत्य कहा जा सकता है जब यह ज्ञात हो वे सभी व्यक्ति जो बीयर नहीं पीते हैं। 100 वर्ष की उम्र प्राप्त कर लेते हैं। बीयन न पीने वाले व्यक्तियों की जीवन-अवधि ज्ञात न होने की दशा में यह निष्कर्ष भ्रमात्मक हो सकता है। यह सम्भव हो सकता है कि वे व्यक्ति भी जो बीयर-पीने के सम्बन्ध में पूर्ण तथ्य ज्ञात न होने की दशा में कोई निश्चित निष्कर्ष नहीं निकाला जा सकता है।

(*iii*) इस विवरण में दो गुण निहित हैं — (*a*) भूमि के स्वामी जो कृषिकार-स्वामी (cultivating owners) हों अथवा गैर-कृषिकार स्वामी (non-cultivating owners) हों, तथा (*b*) भूमिगत सम्पत्ति जो विरासत (inherited) में प्राप्त की गयी हो तथा जो विरासत में प्राप्त न की गयी हो।

यह कथन कि गैर कृषिकार स्वामियों को, जिन्होंने भूमि विरासत में प्राप्त की है, कोई क्षतिपूर्ति नहीं की जानी चाहिए क्योंकि 90% गैर कृषिकार स्वामियों के पास ऐसी भूमि है; इस तथ्य की अवहेलना करता है कि कृषिकार-स्वामियों के पास भी विरासत में प्राप्त भूमि हो सकती है। 'विरासत मे प्राप्त भूमि' का गुण कृषिकार स्वामियों तथा गैर-कृषिकार स्वामियों दोनों में ही समान अनुपात में हो सकता है। अतः दोनों गुण स्वतंत्र हो सकते हैं और कृषिकार-स्वामित्व एवं विरासत में प्राप्त भूमि के मध्य कोई गुणसम्बन्ध न हो। अतः इस कथन भ्रमात्क गुणसम्बन्ध पर आधारित है।

सारांश

- गुण वह तथ्य होते हैं जिनकी अंकात्मक माप सम्भव नहीं है जैसे योग्यता, साक्षरता, आंख का रंग आदि। इनकी गणना करके की संख्यात्मक रूप दिया जा सकता है।
- जिस तरह से दो चरों के मध्य सहसम्बन्ध निकाला जाता है, उसी तरह दो गुणों के मध्य गुण सम्बन्ध निकाला जाता है।
- गुण सम्बन्ध का अध्ययन करते समय किसी भी गुण को केवल दो श्रेणियों में विभाजित किया जाता है — एक श्रेणी उनकी होगी जिनमें वह गुण उपस्थित हैं तथा दूसरी श्रेणी वह होगी जिनमें वह गुण अनुपस्थित हैं। सांख्यिकी में इन्हें A, B, C etc. तथा a, b, g या *a, b, c* से दिखाया जाता है।
- यदि हम केवल एक गुण की उपस्थिति (A) या अनुपस्थिति (*a*) दिखाते हैं तो इसके तीन वर्ग बनेंगे [*n* (A), *n* (*a*) तथा N] समग्र संख्या N के बराबर होती है।
- यदि दो गुणों की उपस्थिति (A, B) या अनुपस्थिति (*a, b*) दिखानी है तो वर्गों की कुल संख्या 9 होगी। [*n* (A), *n* (*a*), *n* (B), *n* (*b*), *n* (AB), *n* (A*b*), *n* (*a*B), *n* (*ab*), N]
- यदि तीन गुणों की उपस्थिति (A, B) या अनुपस्थिति (*a, b*) दिखानी है तो कुल 27 वर्ग बनेंगे। [*n* (A), *n* (B), *n* (C), *n* (*a*), *n* (*b*), *n* (*c*), *n* (AB), *n* (AC), *n* (BC), *n* (A*b*), *n* (A*c*), *n* (*a*B), *n* (*ab*), *n* (*a*C), *n* (*ac*), *n* (B*c*), *n* (*b*C), *n* (*bc*), *n* (ABC), *n* (AB*c*), *n* (A*b*C), *n* (A*bc*), *n* (*a*BC), *n* (*a*B*c*), *n* (*ab*C), *n* (*abc*), N]

- गुण सम्बन्ध गुणांक किन्हीं दो गुणों के बीच सम्बन्ध स्थापित करने की विधि है। यह सम्बन्ध धनात्मक भी हो सकता है तथा ऋणात्मक भी। यदि दो गुणों के बीच सम्बन्ध नहीं है तो उन्हें हम स्वतंत्र गुण कहते हैं।
- गुण सम्बन्ध का अध्ययन आर्थिक, सामाजिक तथा स्वास्थ्य संबंधी समस्याओं को समझने में बहुत उपयोगी साबित हो सकता है। जैसे यदि हमें निरीक्षता और अपराधीकरण में सम्बन्ध का पता करना है तो किसी एक शहर, जिला या राज्य से निरक्षरों व साक्षरों की संख्या तथा उनमें से कितने अपराधी हैं इस पर आंकड़े इकट्ठे करके गुण सम्बन्ध गुणांक निकल कर किस निष्कर्ष पर पहुंच सकते हैं।

प्रश्नावली
(Exercise)

(1) सांख्यिकी में प्रयुक्त 'गुणसंबंध' तथा 'सहसम्बन्ध' शब्दों में आप किस प्रकार अन्तर करेंगे ?
How would you distinguish between 'Association' and 'Correlation' as the terms are used in Statistics ?

(2) आर्थिक समंकों के विश्लेषण में गुणसम्बन्ध-गुणक के उपयोग पर एक टिप्पणी लिखिए।
Write a note on the use of coefficient of Association in analysing economic statistics.

(3) 'गुणसम्बन्ध' तथा 'आंशिक गुणसम्बन्ध' के विचारों का विवेचन कीजिए तथा निम्नलिखित दलीलों को आलोचनात्मक समीक्षा कीजिए :-
(अ) लगभग A के सभी B के हैं अतः A तथा B में अवश्य ही गुणसम्बन्ध है।
(ब) एक विधान सभा में भूमिक विधेयक के पक्ष में मतदाताओं ने 99 प्रतिशत किसान थे। अतः यह मानना अनुचित है कि मतदान पक्षपात रहित था।
Discuss the concept of Association and Partial Association and examine critically the following arguments :-
(*a*) Nearly all the A's and B's and, therefore, A and B must be associated.
(*b*) 99 percent of the members in a legislature who voted for the tenancy bill were farmers. Therefore, it was unfair to suppose that voting was unbiassed.

(4) गुणसम्बन्ध की परिभाषा दीजिए और यह बतलाइए कि यह किस प्रकार मापा जाता है ? 'आंशिक गुणसम्बन्ध' का क्या तात्पर्य है और इसके अध्ययन का क्या महत्व है ?
Define 'Association' and discuss the methods by which it is measured ? What is 'partial association' and what is the significance of its study ?

(5) Discuss the meaning and significance of association of attributes.
गुणों सम्बन्ध का अर्थ व महत्व बतायें।

(6) What is the test of determining that attributes are independent, associated and disassociated ?
गुणों की स्वतन्त्रता, धनात्मक व ऋणात्मक सम्बन्ध परीक्षण विधि क्या है ?

(7) Define association of attributes. How would you measure it ?
गुण सम्बन्ध परिभाषित करें। इसे कैसे मापते हैं ?

(8) What do you mean by 'Association and Attributes? How will you examine the consistency of data classified according to different attributes ?
गुण सम्बन्ध से आप क्या समझते हैं ? विभिन्न गुणों के होने पर समंकों की संगति की जाँच आप कैसे करेंगे ?
निम्नलिखित उदाहरण से सभी वर्गों की आवत्तियां ज्ञात करो।

(9) एक परीक्षा में जिसमें 1600 उम्मीदवारों ने परीक्षा दी, लड़कों की संख्या लड़कियों की संख्या से 166 ज्यादा थी। उत्तीर्ण परीक्षार्थियों की संख्याा अनुत्तीर्ण परिक्षार्थीयों से 316 अधिक थी। उत्तीर्ण हुए लड़ाकें में से 300 ने साइंस चुनी जबकि उन लड़कियों में से जिन्होंने कला विषय चुन थे 25 अनुतीर्ण हो गई। कुल मिलाकर 135 ने कला विषय चुने थे जिनमें से 33 अनुत्तीर्ण हुए थे। कुल अनुतीर्ण लड़कों की संख्या 18 थी।
Obtain all the class frequencies in the following example :-
At an examination at which 600 candidates took the examination boys outnumbered girls by 16%. These passing the examination exceeded in number those failing by 310. The number of successful boys choos-

ing science subjects was 300 while among the girls offering Arts subjects there are 25 failures. Altogether only 135 offered Arts any 33 among them failed. Boys failing in the examination numbered 18.

[If 'A' denoted boys, '*a*' girls. 'B' denotes passing and '*b*' failing and 'C' denotes choosing science subjects and '*c*' choosing Arts subject then —

[N = 600, (A) = 348, (*a*) = 252, (B) = 455, (*b*) = 145, (C) = 465, (*c*) 135, (AB) = 330, (AC) = 310, (BC) = 353, (A*b*) = 18, (A*c*) = 38, (B*c*) = 102, (*a*B) = 125, (*a*C) = 155, (*b*C) = 112, (*ab*) = 127, (*ac*) = 97, (*bc*) = 33, (ABC) = 300, (AB*c*) = 30, (A*b*C) = 10, (A*bc*) = 8, (*a*BC) = 53, (*a*B*c*) = 72, (*ab*C) = 102, (*abc*) = 25]

(10) With the help of the following ultimate class frequencies, find the frequencies of positive and negative classes and also the total number of observations :

निम्नलिखित अन्तस्थ वर्गों की सहायता से धनात्मक, ऋणात्मक व कुल अवलोकन संख्या ज्ञात करो :-

AB = 200, *ab* = 100, *a*B = 150 and *ab* = 50

(11) In a college there are 200 students and B.Com. II. They appeared 3 times in examinations. Their result is as under :-

80 got through in Ist exam, 122 could not succeed in second exam and 96 got through in third exam. 18 students passed in first two but could not pass the third exam, 38 students failed in first two but passed in third exam.

एक कॉलेज में बी.काम — II के 200 छात्र हैं और वे तीन बार परीक्षा में बैठे। उनका परीक्षाफल निम्न प्रकार है। 80 छात्र प्रथम परीक्षा में उत्तीर्ण व 122 दूसरी परीक्षा में अनुत्तीर्ण रहे। 96 तीसरी परीक्षा में उत्तीर्ण रहे। 18 छात्र प्रथम दो परीक्षाओं में पास व तीसरी में फेल रहे जबकि 38 छात्र प्रथम दो परीक्षाओं में फेल व तीसरी में उत्तीर्ण रहे हों तो उन विद्यार्थियों की संख्या बताओ जो कम से कम 2 परीक्षाओं में उत्तीर्ण रहे।

(12) निम्न आंकड़ों से (ABC) की कम से कम संभावित सीमाएं ज्ञात करें।

From the data given below determine the smallest possible limits for the value of (ABC)

N = 100, (A) = 60, (B) = 70, (C) = 78

(AB) = 30, (BC) = 50 and (AC) = 45

[Between 17 and 23]

(13) 100 बच्चों ने तीन परीक्षाएँ दीं प्रथम परीक्षा में 40 पास हुए, द्वितीय में 39 और ततीय परीक्षा में 48 बच्चे पास हुए। 10 बच्चे तीनों परीक्षाओं में पास हुए। 21 बच्चे तीनों परीक्षाओं में फेल हुए। 9 बच्चे पहली दो में पास हुए और ततीय परीक्षा में फेल हुए। 19 बच्चे पहली दो में फेल और ततीय परीक्षा में पास हुए। कितने बच्चों ने कम -से-कम दो परीक्षाएँ पा कीं ?

100 children took three examination, 40 passed the first, 39 passed the second and 48 passed the third, 10 passed all the three. 21 failed all the three, 9 passed first two and failed third, 19 failed first two and passed the third. Find how many children passed at least two examinations ?

[**Ans.** 38 passed atleast two examinations]

(14) एक शहर में व्यस्कों की कुल संख्या में से 50% पुरुष हैं, 60% कमाने वाले हैं व 50-50 साल से ऊपर की आयु के हैं। पुरुषों में से 10% कमाने वाले नहीं है तथा 40% पुरुष 50% से कम साल के हैं। क्या हम 50 साल की अधिक आयु वाले कमाने वाले पुरुषों के बारे में कुछ नतीजा निकाल सकते हैं।

Among the adult population of a certain town 50 per cent of the population are males, 60 per cent wage earners, and 50 per cent are 50 years of age or over. 10 per cent of the males and not wage earners and 40 per cent of the males are under 50. Can we infer anything about what percentage of the population of 50 years or over are wage earners ?

[The percentage of wage earning population of 50 years or over must lie between 25 and 45]

(15) उत्तर प्रदेश के किसी जिले में 170,000 शिक्षितों में 5000 अपराधी हैं। इसी जिले में 1330000 अशिक्षितों में 120000 अपराधी हैं। इन आँकड़ों के आधार पर क्या आप अशिक्षा और अपराध में गुण-सम्बन्ध पाते हैं।

Out of 170000 literates in a particular district of India, number of criminals was 5000. Out of 1330 thousand of illiterate in the same district number of criminals was 120000. On the basis of these figures do you find any association between illiteracy and criminality ?

(16) Following information related to employment of males and females :-
निम्न सूचना पुरुषों व स्त्रियों के रोजगार से सम्बन्धित है :

Males Employed (पुरुष रोजगार)	4000
Females Employed (पुरुष रोजगार)	3500
Males Unemployed (पुरुष रोजगार)	1500
Females Unemployed (पुरुष रोजगार)	1000

Does the information reveal any association between sex and his/her employment ?
क्या यह सूचना लिंग व रोजगार में गुण सम्बन्ध प्रकट करती है ?

(17) निम्न तीन यू.पी. के तीन शहरों में शिक्षित तथा अपराधियों की संख्या के आँकड़े दिखाती है :-
The following table gives the number of literates and criminals in three cities of U.P. :-

		Kanpur	**Allahabad**	**Agra**
कुल संख्या	Total Number (in thousands)	244	184	230
शिक्षित	Literates (in thousands)	40	47	33
शिक्षित अपराधी	Literate Criminals (in thousands)	3	2	2
अशिक्षित अपराधी	Illiterate Criminals (in thousands)	40	20	24

हर शहर में शिक्षा तथा अपराधीकरण में बीच में गुणसम्बन्ध गुणांक की पदवी ज्ञात करें।
Compare the degree of association between criminality and illiteracy in each of the three towns.

[Qs. Kanpur = + .45, Alld. + .55 and Agra = + .34]

(18) यह कहा जाता है कि महाविद्यालयों का अध्यापक वर्ग देश की राजनीति में अधिक रुचि लेता है। इस तथ्य की प्रामाणिकता का पता लगाने के लिए 2,000 ऐसे लोगों का अध्ययन किया गया तो 5,000 रू. मासिक या इससे अधिक वेतन पाते हैं। उनसे राजनीति में उनकी रुचि के सम्बन्ध में प्रश्न पूछे गए और यह ज्ञात हुआ कि 2,000 लोगों में से 920 अध्यापक थे। इन अध्यापकों में से 480 अध्यापकों की राजनीति में दिलचस्पी थी और गैर-अध्यापकों में से 75 व्यक्तियों की राजनीति में कोई दिलचस्पी नहीं थी। इस गणात्मक वर्गीकरण को एक सारणी के रूप में प्रदर्शित कीजिए तथा अज्ञात वर्ग-आवत्तियों का निर्धारण कीजिए तथा अध्यापकों और उनके द्वारा देश की राजनीति में अधिक रुचि लेने पर मध्य साहचर्य गुणांक भी ज्ञात कीजिए।
It is commented that teachers of degree colleges take more active interest in the politics of the nation. To verify the truthfulness of the matter a survey of 2,000 persons getting Rs. 5,000 p.m. or more was conducted and it was found that out of the same of 2,000 persons 920 were teachers. Out of these teachers 480 were found to be keenly interested in politics. Of the non-teachers 750 persons had no interest in politics. Represent the above qualitative classification in the form of a table and determine the unknown class-frequencies and also find out coefficient of association between teachers and active interest in the politics of the nation.

[**Ans.** Q_{AB} = + 0.425; (AB) = 440, (*a*B) = 330, (*a*) = 1,080, (B) = 810, (*b*) = 1,190]

(19) 127 व्यक्तियों में से जिन्हें टीका नहीं लगा था, 10 प्लेग से प्रभावित हुए, उनमें से 6 की मत्यु हुई। 147 व्यक्ति जिन्हें टीका लगा था, 3 प्लेग से प्रभावित हुए — और उनमें से किसी की मत्यु नहीं हुई। गुण -सम्बन्ध ज्ञात कीजिए।

(*i*) टीका लगने एवं बीमार होने के बीच तथा

(*ii*) टीका लगने एवं प्लेग होने वाले व्यक्तियों में मत्यु होने के बीच।

Of 127 persons who were uninoculated, 10 were affected by plague, 6 of them died. Of 14 persons who were inoculated, 3 were affected by plague and there were no deaths. Find the association :-

(*i*) Between inoculation and illness and

(*ii*) Between inoculation and mortality among the persons who were affected by the plague.

[**Ans.** (*i*) $Q_{AB} = -0.61$; (*ii*) $Q_{AB} = -1$].

अध्याय - 11

निर्देशांक (Index Numbers)

निर्देशांक एक ही समह के विभिन्न परन्तु सम्बन्धित चलों के सामान्य स्तर की दशाओं में तुलना करने की सांख्यिकीय विधि है। यदि हम भारत में 1970 के सामान्य मूल्य स्तर की तुलना 1960 में प्रचलित मूल्य स्तर से करना चाहते हैं तो अनेक प्रकार की वस्तुओं, जैसे गेहूँ, चावल, कपड़ा, मकान, दूध, तेल आदि के मूल्यों को विचारमत करना पड़ेगा। यदि इसी विधि में सभी वस्तुओं के मूल्यों में समान अनुपात में ही परिवर्तन हों तो मूल्य-स्तर के परिवर्तन का पता लगाने में कोई विशेष कठिनाई नहीं होती है। परन्तु व्यवहार में यह पाया जाता है कि विभिन्न वस्तुओं के मूल्यों में परिवर्तन विभिन्न दिशाओं में तथा विभिन्न अनुपातों में होते हैं। इसके अतिरिक्त दूसरी कठिनाई या उपस्थित होती है कि विभिन्न प्रकार की वस्तुओं के मूल्य अलग-अलग मापदण्डों में व्यक्त किये गये हैं, उदाहरणार्थ, गेहूँ, चावल आदि का भाव किलोग्राम में, कपड़े का भाव मीटर में, दूध का भाव लिटर में व मकान का किराया कमरों के आधार पर व्यक्त किये जाते हैं। मापदण्ड की इकाइयाँ समान न होने के कारण तथा विभिन्न वस्तुओं के मूल्यों में अलग-अलग दिशाओं तथा अनुपातों में परिवर्तन होने के कारण वस्तुओं के मूल्यों का साध्य ज्ञात करने में कठिनाई आती है। आवश्यकता एक ऐसी निर्देशक संख्या की होती है जो विभिन्न वस्तुओं के मूल्यों में नए परिवर्तन की दिशा तथा मात्रा को सामूहिक रूप से व्यक्त कर सके। यह कार्य निर्देशांक-विधि से किया जाता है। वास्तव में निर्देशांक एक विशिष्ट प्रकार के माध्य होते है।[1] इनका प्रयोग सामान्य मूल्य स्तर की प्रवत्ति की माप तक ही सीमित नहीं है बल्कि जीवन स्तर , उत्पादन, राष्ट्रीय आय, औद्योगिक क्रियाएँ, उत्पादकता आदि, जिनकी प्रत्यक्ष माप सम्भव नहीं होती है की सापेक्ष माप के लिए भी निर्देशकों की सहायता ली जाती है।[2]

परिभाषा एवं विशेषताएं
(Definition and Characteristics)

सरलतम रूप में निर्देशांक, प्रतिशत के रूप में व्यक्त दो संख्याओं का अनुपात है। विभिन्न सांख्यिकी ने निर्देशांक को उग्र प्रकार परिभाषित किया है।

वॉडिन्गटन के अनुसार, ''निर्देशांक, जैसा कि उसके नाम से स्पष्ट है, संख्याओं के एक समूह की सामान्य प्रकृति का द्योतक है।''[3] वॉडिन्गटन ने निर्देशांक को ऐसा इसलिए परिभाषित किया है क्योंकि उनसे किसी समूह में होने वाले परिवर्तनों के आधार पर उस समूह की सामान्य प्रवत्ति का पता चलता है। कॉक्सटन तथा कॉउडन के अनुसार , ''निर्देशांक एक समूह के सम्बन्धित पर मूल्यों के आकार में होने वाले अन्तरों का माप करने की विधि है।'' निर्देशांकों से मूल्यों के सापेक्ष या तुलनात्मक अन्तरों का मापन होता है। यह माप समय या स्थान (Over Time and Space) के आधार पर किया जाता है। होरेस सीकारइट लिखते हैं, ''निर्देशांक अंकों की एक ऐसी श्रेणी है जिनके द्वारा किसी भी तथ्य के परिणाम में करने वाले परिवर्तनों को समय या स्थान के आधार पर मापा जा सकता है।'' डॉ. बाउले के अनुसार, ''निर्देशांकों की एक श्रेणी ऐसी श्रेणी होती है जो अपनी प्रवत्ति एवं उच्चावचनों के द्वारा इस परिभाषा में होने वाले परिवर्तनों को जिससे वह सम्बन्धित है, स्पष्ट करती है।''

स्पाइगेल के अनुसार, ''निर्देशांक एक सांख्यिकीय माप है जो समय, भौगोलिक स्थिति अन्य विशेषता के आधार पर किसी चर मूल्य अथवा सम्बन्धित चर-मूल्यों के समूह में होने वाले परिवर्तनों को प्रदर्शित करती है।'' पेडन तथा लिन्डनिवस्ट के अनुसार, ''एक निर्देशांक परिवर्तन, जो समयानुसार क्रमबद्ध मूल्यों की एक श्रेणी में हुए हों, की एक माप है तथा सामान्यतः इन्हें प्रतिशतों से व्यक्त किया जाता है।'' स्पर तथा दोमिनी के अनुसार, ''निर्देशांक किसी आधार, जिसे 100 माना जाता है, की तुलना में एक घर में होने वाले सापेक्ष परिवर्तनों को व्यक्त करते हैं।'' बंसल विलेट तथा साइमोन के शब्दों में, ''एक निर्देशांक विशिष्ट प्रकार का भाव्य है जो समय-समय अथवा स्थान-स्थान के सापेक्ष परिवर्तनों को माप करता है।'' क्लार्क तथा शकार्ड के शब्दों में, ''निर्देशांक एक प्रतिशत अनुपात (Percentage Relative) है जो एक दी गयी समयावधि की आर्थिक मापों को अतीत की स्थिर

समयावधि की उन्हीं मापों से तुलना करता है।'' टटिल के अनुसार, ''निर्देशांक एकमात्र अनुपात (सामान्यतः प्रतिशतों में) है जो विभिन्न चरों में दो विभिन्न समयों, स्थानों अथवा स्थितियों के सामूहिक (अर्थात् औसत रूप में) परिवर्तनों को मापता है।''

सूचकांकों की विशेषताएं
(Characteristics of Index Numbers)

सूचकांकों की प्रमुख विशेषताएं निम्नलिखित हैं :-

(1) **सूचकांक एक संख्यात्मक माप है (Index Number is a Numerical Measure) :** सूचकांक सदैव संख्या के रूप में व्यक्त किए जाते हैं।

(2) **व्यापकता (Universality) :** सूचकांक का प्रयोग केवल मूल्य-स्तर के माप के लिए ही नहीं किया जाता बल्कि किसी भी ऐसी घटना के सापेक्ष माप के लिए किया जा सकता है जिसका प्रत्यक्ष माप सम्भव न हो।

(3) **तुलना का आधार (Basis of Comparison) :** सूचकांकों द्वारा समय अथवा स्थान के आधार पर तुलना की जाती है, परन्तु साधारणतया यह तुलना समय के आधार पर ही की जाती है। तुलना के लिए जिस निश्चित वर्ष के स्तर को आधार मान लिया जाता है उसे आधार वर्ष (Base Year) कहते हैं तथा जिस वर्ष की तुलना की जाती है वह चालू वर्ष (Current Year) कहलाता है।

(4) **प्रतिशतों का माध्य (Average of Percentage)** इसमें आधार वर्ष या आधार स्थानों के मूल्यों को 100 मानकर प्रत्येक दूसरे वर्षों के मूल्यों के प्रतिशतों में बदला जाता है जिन्हें मूल्यानुपात कहते हैं। फिर भी मूल्यानुपातों का माध्य निकाला जाता है। यही माध्य सूचकांक कहलाता है। अतः सूचकांक प्रतिशतों का माध्य है।

(5) **परिवर्तनों का सापेक्ष माप (Relative Measure of Changes)** सूचकांकों के द्वारा दो समयों अथवा स्थानों के बीच तुलनात्मक या सापेक्ष परिवर्तनों को ही मापा जाता है। उदाहरण के लिए, यदि 1990 में थोक मूल्य सूचकांक 100 हो तथा 1997 में 180 हो जाए तो इसका अर्थ यह हुआ कि 1990 की तुलना में थोक मूल्य स्तर 80% बढ़ गया।

सूचकांकों के उपयोग
(Uses of Index Numbers)

आज के समग्र में चूकांकों का महत्व बढ़ता जा रहा है। आज इसका प्रयोग वाणिज्य, उद्योग, प्रशासन, राजनीतिक एवं सामाजिक सभी क्षेत्रों में किया जाने लगा है। सूचकांकों के निम्नलिखित उपयोग हैं :-

(1) **जटिल तथ्यों को सरल बनाना (To Simplify Complexities) :** सूचकांक द्वारा ऐसे जटिल परिवर्तनों का माप संभव हो जाता है, जिनका प्रत्यक्ष रूप से माप नहीं हो सकता।

(2) **व्यावसायिक क्षेत्र में उपयोगिता (Utility in Business) :** सूचकांक व्यावसायिक परिस्थितियों में होने वाले परिवर्तनों को मापते हैं तथा उन परिवर्तनों का तुलनात्मक अध्ययन करने में काफी उपयोगी सिद्ध होते हैं। जैसे बिक्री, उत्पादन तथा मूल्यों आदि में होने वाले परिवर्तन। अतः सूचकांक व्यवसायी के लिए वायुमापक यन्त्र (Barometer) की तरह कार्य करते हैं।

(3) **मुद्रा में मूल्य परिवर्तन का मापन (To Measure the Changes in the Value of Money) :** सामान्य मूल्य सम्बन्धी-निर्देशांक मूद्रा के मूल्य में होने वाले परिवर्तनों के बारे में जानकारी देता है। जब मूल्यों में कमी होती है तो मुद्रा का मूल्य बढ़ जाता है और जब मूल्यों में वद्धि होती है तो मुद्रा का मूल्य घट जाता है।

(4) **तुलनात्मक अध्ययन को सम्भव बनाना (To Facilitate Comparative Study) :** निर्देशांक मुद्रा-मूल्य के सम्पूर्ण या निर्पेक्ष (Absolute) अध्ययन को सम्भव नहीं बनाते बल्कि मूल्यों में परिवर्तन के तुलनात्मक अध्ययन की ओर संकेत करते हैं। जैसे — यदि कहा जाए कि मूल्यों का निर्देशांक 120 था तो इसका तात्पर्य यह है आधार वर्ष से चालू वर्ष तक वस्तुओं के मूल्यों में 20% की वद्धि हुई है। यदि 1995 के मूल्यों की तुलना 1994 के मूल्यों से की जाए जो 1995 को चालू वर्ष (Current Year) एवं 1994 का आधार वर्ष (Base Year) कहा जाएगा। इस प्रकार निर्देशांकों की सहायता से तुलनात्मक अध्ययन काफी सरल हो जाता है।

(5) **वेतन एवं भत्ते निर्धारित करने में सहायक (Helpful in Fixation of Salary and Allowances) :** उपभोक्ता मूल्य निर्देशांक (Consumer Price Index Numbers) देश के उपभोक्ता वर्ग विशेषतः वेतन पाने वाले व्यक्ति के घटते या बढ़ते

हुए व्यय की ओर संकेत करते हैं, जिससे श्रमिकों की न्यूनतम मज़दूरी एवं महंगाई भत्ते निर्धारित करने में सहायता मिलती है।

(6) कर-नीति में सहायता (Helpful in Tax Policy) : मूल्य, उपभोक्ता व्यय एवं औद्योगिक उत्पादन तथा लाभ निर्देशांक सरकार को उचित कर-नीति निर्धारित करने में सहायता प्रदान करते हैं।

(7) विदेशी व्यापार सम्बन्धी ज्ञान (Knowledge Regarding Foreign Trade) : विदेशी व्यापार सम्बन्धी निर्देशांकों से विदेशी व्यापार की स्थिति की जानकारी प्राप्त होती है और इसके पश्चात् विदेशी व्यापार के भुगतान में सन्तुलन की स्थिति लाने की कोशिश की जा सकती है।

(8) आर्थिक नीति का आधार (Basis of Economic Policy) : उत्पादन सम्बन्धी निर्देशांक इस बात की ओर संकेत करते हैं कि किन उद्योगों का उत्पादन बढ़ रहा है एवं किन का घट रहा है। इस जानकारी के आधार पर सरकार अपनी आर्थिक नीति का निश्चय सही रूप से कर सकती है। विभिन्न उद्योग-धन्धे भी इस बात की जानकारी प्राप्त कर सकते हैं कि उनका कार्य सन्तोषजनक स्थिति में है या नहीं।

(9) अन्य लाभ (Other Uses) : निर्देशांकों के कुछ अन्य लाभ भी हैं, जैसे :-

(*i*) यातायात से सम्बन्धित सूचनाओं के आधार पर रेलवे विभाग या निर्णय लेता है कि विशेष समय में कितनी गाड़ियाँ चलानी चाहिएँ।

(*ii*) बीमा कम्पनियों को प्रीमियम की दर निर्धारित करने में सहायक।

(*iii*) इनके आधार पर ही प्रतिभूतियों में सट्टा करने वाले स्टोरिए सट्टे का जोखिम लेते हैं।

(*iv*) बैंक अधिकारियों को ब्याज की दर निर्धारित करने में सहायक।

इस प्रकार वर्तमान युग में निर्देशांकों का प्रयोग निर्विवाद सत्य है।

निर्देशाकों की रचना में समस्याएँ
(Problems in the Construction of Index Numbers)

निर्देशांक तैयार करते समय विभिन्न समस्याएँ हमारे सामने आती हैं, वे निम्न प्रकार से हैं :-

(1) निर्देशांक-रचना के उद्देश्य को परिभाषित करना (Definition of the Purpose).

(2) निर्देशांक पदों का चुनाव (Selection of the Items of 'Regimen')

(3) समंक प्राप्त करने के स्रोतों का चुनाव (Selection of Sources of Data).

(4) समंक-संकलन (Collection of Data)

(5) आधार का चुनाव (Selection of the Base)

(6) माध्य का चुना (Selection of an Average)

(7) भारांकन विधि (System of Weighting)

(1) निर्देशांक का उद्देश्य परभाषित करना (Defining the Purpose of an Index Number) : निर्देशांक की रचना करने से पूर्व यह परम आवश्यक है कि उसके उद्देश्य को स्पष्ट रूप से परिभाषित कर लिया जाय। कोई भी निर्देशांक सर्व-उद्देशीय (All-purpose) नहीं होता है। पूर्व में ही यह जान लेना आवश्यक होता है कि हम किसका माप करने जा रहे हैं तथा उस माप का प्रयोग किस उद्देश्य के लिए किया जाएगा। ''उद्देश्य का सूक्ष्म रूप से निर्धारण किये बिना .... यह जानना असम्भव होता है कि निर्देशांक की रचना में निहित पदक्षेपों को किस प्रकार उचित ढंग से सम्पन्न किया जाय।'' निर्देशांक विशिष्ट यन्त्र होते हैं, अतः उनके उचित प्रयोग पर ही उनकी क्षमता तथा उपयोगिता निर्भर करती है। वस्तुओं का चुनाव, उनके मूल्य-उद्धहरण (Price Quotation) तथा उनको भार देने आदि के बारे में निर्णय लेने में निर्देशांक का उद्देश्य आधारभूत तथ्य होता है। उदाहरणार्थ, एक सूक्ष्मग्राही मूल्य निर्देशांक (Sensitive Price Index) में उन वस्तुओं को शामिल किया जाएगा जिनके मूल्यों में तेजी से परिवर्तन होते हैं : इसके विपरीत सामान्य उद्देश्य वाले मूल-निर्देशांक में अधिक से अधिक वस्तुओं को लिया जायगा। जीवन-निर्वाह निर्देशांक (Cost of Living Index Number) यदि श्रमिकों से सम्बन्धित होगा तो सामान्य मज़दूरों द्वारा उपभोग की जाने वाली वस्तुओं का समावेश किया जायेगा। अतः निर्देशांकों

के उद्देश्य का पूर्व-निर्धारण आवश्यक होता है।

(3) पदों का चुनाव (Selection of the Items or 'Regimen') : निर्देशांक में शामिल की जाने वाली वस्तुओं की सूची को 'Regimen' या 'Basket' कहते हैं। किसी भी निर्देशांक में सभी पदों या वस्तुओं को शामिल कर लेना न तो सम्भव ही होता है और न आवश्यक ही। प्रत्येक निर्देशांक का उद्देश्य एक विशेष वर्ग से सम्बन्धित परिवर्तनों को मापना होता है। निर्देशांक में सभी परिवर्तन शामिल नहीं होते बल्कि कुछ चुने हुए परिवर्तनों पर ही विचार किया जाता है। उदाहरणस्वरूप, उपभोक्ता मूल्य निर्देशांक में उपभोग में आने वाली समस्त वस्तुओं का समावेश नहीं किया जा सकता है। आजकल उपभोक्ता पदार्थ इतने उपलब्ध हैं कि उन्हें उपभोक्ता मूल्य निर्देशांक में शामिल करना असम्भव होता है। इसके अतिरिक्त नये उपभोक्ता पदार्थों का चलन शुरु होता रहता है। कॉनर लिखते हैं कि निर्देशांक के लिए चुन गये पद सम्बन्धित (Relevant) प्रतिनिधि, विश्वसनीय तथा तुलना योग्य होनी चाहिए। पदों का चुनाव निर्देशांक के उद्देश्य पर निर्भर करता है। निर्देशांक में शामिल करने के लिए वस्तुओं अथवा पदों का चुनाव करते समय निम्नलिखित व्यापक सिद्धान्तों को आधार बनाया जाना चाहिए।

(*i*) चुनी हुई वस्तुएँ अपने वर्ग का उचित प्रतिनिधित्व कर सकें।

(*ii*) चुनी हुई वस्तुएँ आसानी से पहचानी (Easy Recognition) जा सकें।

(*iii*) चुनी हुई वस्तुएँ स्थान या क्षेत्र में विशेष लोकप्रिय (Popular) हों।

(*iv*) चुनी हुई वस्तुएँ के गुणों में कोई अन्तर न हो।

(3) समंक-स्रोतों का चुनाव (Selection of Sources of Data) : निर्देशांक की रचना के लिए मूल्य-उद्धरण (Price Quotations) या तो नियमित रूप से प्रकाशित बाजार भावों में लिये जा सकते हैं अथवा व्यापारी, उत्पादक, निर्यातक आदि के प्राप्त किये जा सकते हैं। दोनों ही दिशाओं में इस बात का ध्यान रखा जाना चाहिए कि भाव उसी वस्तु के हैं। जिसे निर्देशांक में शामिल किया जाता है। संकलित की जाने वाली जानकारी सत्य तुलनीय, पर्याप्त तथा प्रतिनिधि होनी चाहिए। अतः उसी के अनुकूल जानकारी का स्रोत चुनना चाहिए।

(4) समंक-संकलन (Collection of Data) : वस्तुओं के मूल्यों से सम्बन्धित समंक संकलित करने में निम्नलिखित बातों पर विचार करना पड़ता है :-

(अ) मूल्य व्यक्त करने का ढंग — वस्तुओं के मूल्य दो प्रकार से व्यक्त किये जाते हैं (1) मुद्रा में व्यक्त मूल्य (Money Prices) — जिसमें वस्तु का मूल्य मुद्रा की इकाइयों में व्यक्त किया जाता है, जैसे 100 रुपये प्रति क्विंटल (100 किलोग्राम) गेहूँ, 2 रुपये प्रति मीटर कपड़ा, 24 रु. प्रति कमरे मकान किराया आदि। (2) वस्तु परिणाम में व्यक्त मूल्य (Quantity Price) — जिनमें वस्तु का मूल्य प्रति मुद्रा-इकाई में आने वाले परिणाम में व्यक्त किया जाता है, जैसे 1 रुपये का 1 किलोग्राम गेहूँ, 1 रुपये का 50 सेन्टीमीटर कपड़ा आदि। निर्देशांक की रचना में 'मुद्रा में व्यक्त' मूल्यों का ही प्रयोग होता है। यदि मूल्य परिणाम के रूप में व्यक्त हो तो उन्हें मुद्रा में व्यक्त कर लेना चाहिए।

(ब) थोक या फुटकर मूल्य — वस्तुओं के मूल्य थोक या फुटकर हो सकते हैं। सामान्य उद्देश्यीय निर्देशांक की रचना में थोक मूल्य ही लिये जाते हैं। इसका कारण यह है कि थोक मूल्य फुटकर मूल्यों की अपेक्षा कम परिवर्तनशील होते हैं तथा उनमें स्थान-स्थान के आधार पर कम अन्तर पाया जाता है। परन्तु जीवन-निर्वाह निर्देशांक की रचना में फुटकर मूल्यों को लेना उपयुक्त होता है।

(स) बाजार का चुनाव — मूल्य-उद्धरण (Price Quotations) प्रमुख बाजारों से प्राप्त करने चाहिए। सभी बाजारों से, जहाँ वस्तुओं का क्रय तथा विक्रय होता है, मूल्य-उद्धरण प्राप्त करना न तो सम्भव ही है और न आवश्यक ही, अतः बाजारों का भी न्यादर्श (Sample) लिया जाना चाहिए। जिन बाजार में उन वस्तुओं का काफी मात्रा में क्रय-विक्रय होता है, उसी बाजार का न्यादर्श में चुना जाता है।

(द) मूल्य-उद्धरण देने वालों का चुनाव — बाजार का चुनाव कर लेने के पश्चात यह समस्या उत्पन्न होती है कि मूल्य-उद्धरण किन से प्राप्त किये जायें। सदैव विश्वसनीय साधनों से ही मूल्य-उद्धरण प्राप्त करने चाहिए। कई साधनों से मूल्य-उद्धरण प्राप्त करके उनका सत्यापन (Verification) कर लेना चाहिए।

(य) प्रमाणिक (Standard Price) मूल्य को लेना — विश्वसनीय तथा प्रमाणित निर्देशांकों की रचना के लिए 'प्रमापित

मूल्य' ही लिये जाने चाहिए। प्रमापित मूल्य से तात्पर्य उस प्रतिनिधि मूल्य से है जो विचाराधीन समयान्तर में उसे वस्तु का होता है। यदि मासिक निर्देशांक की रचना करनी ही तो किस सप्ताह के मूल्य लिये जाये ? यदि साप्ताहिक निर्देशांक की रचना करनी हो तो किस दिन मूल्य लिये जाये ? यह एक जटिल समस्या होती है। माह अथवा सप्ताह के प्रारम्भ के मूल्य उद्धरण लिये जाये या मध्य के या अन्त के ? यदि साप्ताहिक निर्देशांक की रचना करनी हो तो यह उचित होगा कि सप्ताह के सभी मूल्य उद्धरण लेकर उनके माध्य के आधार पर निर्देशांक की रचना की जाये। मासिक निर्देशांकों की रचना के लिए साप्ताहिक मूल्यों के माध्य लिए जाने चाहिए। वार्षिक निर्देशांकों में मासिक मूल्यों के माध्य उचित रहते हैं। इस परम्परा से असाधारण उच्चावचनों का प्रभाव निर्देशांकों पर नहीं पड़ने पाता।

समंक स्रोतों तथा संकलन विधि की जाँच बड़ी सावधानी से करनी चाहिए जिससे प्रत्येक पदक्षेप तथा हर समय परिशुद्धता के प्रति आवश्वस्त रह जा सके।

(क) **आगणकों का चुनाव** — समंक संग्रहण के लिए उत्तरदायी आगणकों के चुनाव में भी विशेष सावधानी बरतनी चाहिए क्योंकि उनकी योग्यता तथा निष्पक्षता पर ही निर्देशांकों की विश्वसनीयता तथा उपयोगिता निर्भर रहती है।

(5) आधार का चुनाव (Selection of Base) : निर्देशांकों की रचना में उचित आधार का चुनाव महत्त्वपूर्ण होता है। प्रत्येक निर्देशांक का एक आधार होता है जिससे परिवर्तनों को व्यक्त किया जाता है। "एक निर्देशांक का आधार वर्ष वह वर्ष होता है जिससे अन्य वर्षों की तुलना की जाती है" अन्य शब्दों में, निर्देशांकों का आधार एक वर्ष (Base Year) होता है जिसके स्तर के आधार पर आगे के वर्षों के परिवर्तनों को मापा जाता है। यह वर्ष एक सामान्य वर्ष (Normal Year) होना चाहिए अर्थात् उस वर्ष न तो मूल्य स्तर असाधारण रूप से ऊँचा हो और न नीचा ही। आधार वर्ष बहुत पुराना भी नहीं होना चाहिए। यह इसलिए आवश्यक होता है क्योंकि व्यक्ति सामान्यतः वर्तमान दशाओं की तुलना उस आधार-अवधि से करना चाहते हैं जो अधिक पुरानी न हो। आधार वर्ष वर्तमान अवधि के निकट का ही होना चाहिए जिससे उपभोक्ताओं की आदतों एवं उनकी व्यय-संरचना में मूलभूत परिवर्तन न हो पाये। आधार वर्ष कभी भी 100 या 50 अथवा 25 वर्ष पुराना नहीं होना चाहिए। वर्तमान गतिमान परिस्थितियों में आधार वर्ष कभी भी एक दशक से पूर्व का नहीं होना चाहिए। सिम्पसन एवं कापका लिखते हैं, "क्योंकि निर्देशांकों के आधार पर व्यावहारिक निर्णय लिये जाते हैं तथा आर्थिक व्यवहार प्रायः अल्पकालीन ढंग के होते हैं, हम उस आधार से तुलना करना चाहेंगे जो उसी सामान्य आर्थिक ढाँचे में से हो जिसमें से वर्तमान वर्ष। अतः अध्ययन वर्षों के निकट का ही वर्ष हम चुनते हैं।" हाल का ही आधार वर्ष लेने से यह निश्चित हो जाता है कि तुलना अपेक्षित सजातीय तथ्यों के आधार पर ही की गयी है। हाल का आधार वर्ष लेने से यह निश्चित हो जाता है कि तुलना अपेक्षित सजातीय तथ्यों के आधार पर ही की गयी है। हाल का आधार वर्ष लेने से यह मनोवैज्ञानिक कारण भी है जैसा कि लुईस एवं फोक्स ने लिखा है, 'उचित निकट का आधार लाभदायक है क्योंकि वह निर्देशांक को 100 के अति निकट रखता है और इस प्रकार निर्देशांक के परिवर्तनों को समझने में सहायता देता है। उदाहरणार्थ, यदि आधार वर्ष पर्याप्त दूर का हो और वर्तमान निर्देशांक — मान 400, 500 अथवा 600 के निकट होगा। यदि दो लगातार निर्देशांक 573 तथा 604 हों तो परिवर्तन की प्राप्त ज्ञात करने के लिए पुनः गणना करनी होगी। इसके अतिरिक्त इसका मनोवैज्ञानिक प्रभाव भी होता है। यदि मूल्य निर्देशांक 604 हो तो लोगबाग यह सोचते हैं कि मूल्यों का स्तर ऊँचा है।" वर्तमान में प्रत्येक दशक के प्रथम वर्ष (जैसे 1951, 1961, 1971) को आधार वर्ष मानकर आगे के वर्षों की तुलना करने की प्रवत्ति पायी जाती है।

आधार वर्ष निम्नलिखित चार प्रकार का हो सकता है :-

(*i*) स्थिर आधार (Fixed Base)

(*ii*) औसत आधार (Average Base)

(*iii*) श्रृंखला आधार (Chain Base)

(*iv*) लासपेयरे की विधि (Laspeyre's Method)

(*i*) स्थिर आधार नीति (Fixed Base Method) : इस रीति के अनुसार एक सामान्य वर्ष को आधार वर्ष चुना लिया जाता है तथा अन्य वर्षों के मूल्य-स्तरों की तुलना इस आधार वर्ष के आधार पर की जाती है। यह वर्ष हर प्रकार से सामान्य होना चाहिए अर्थात् बाढ़, महामारी, अकाल, युद्ध, अभिवद्धि काल , अवसाद काल उस वर्ष में न हों। जहाँ

तक सम्भव हो सभी प्रकार के निर्देशांकों के लिए समान आधार वर्ष रखना चाहिए जिससे विभिन्न क्षेत्रों में तुलना की जा सके।

(*ii*) **औसत आधार रीति (Average Base Method) :** प्रायः कोई भी वर्ष पूर्णतः सामान्य नहीं होता, जिसको आधार मानकर अगले वर्षों के मूल्य स्तरों की तुलना की जा सके। क्रॉक्सटन तथा कॉउडन के शब्दों में, ''यह सम्भावना हो सकती है कि कोई भी एक वर्ष पर्याप्त रूप से सामान्य न हो जिसे तुलना का श्रेष्ठ आधार बनाया जा सके। व्यापारिक चक्र के साथ-साथ व्यापार तथा मूल्य सदैव घटते तथा बढ़ते रहते हैं, ऐसी दशा में अनेक वर्षों का माध्य श्रेष्ठ आधार रहता है।'' कई वर्षों के मूल्यों के औसत को आधार मानने से अनेक विषमताएँ समाप्त हो जाती है, अतः औसत-आधारप सर्वश्रेष्ठ आधार होता है। कम से कम पाँच वर्षों के मूल्यों का औसत आधार अवश्य बनाया जाना चाहिए।

(*iii*) **श्रंखला आधार रीति (Chain Base Method) :** यदि वर्ष-प्रतिवर्ष के मूल्य स्तरों के परिवर्तनों की तुलना करनी हो तो श्रंखला -आधार रीति अपनायी जाती है। इस रीति के अनुसार प्रत्येक वर्ष के लिए उसके पूर्व का वर्ष आधार होता है, उदाहरणार्थ 1970 वर्ष के लिए 1969, 1969 वर्ष के लिए 1968, 1968 वर्ष के लिए 1967, 1967 वर्ष के लिए 1966 आधार वर्ष रखा जाता है। इस रीति के अनुसार आधार वर्ष बदलता रहता है। अतः इसे बदलते आधार की रीति (Shifting Base System) भी कहते हैं। होरेस सीक्रीस्ट के अनुसार ''अपेक्षाकृत पुराने वर्ष को स्थिर वितरण, रखकर निकाले गये मूल्यनुपातों के वितरण के, गत वर्ष के आधार पर निकाले गये मूल्यानुपातों का वितरण, 'सामान्य वक्र' (Normal Curve or Error) के अधिक अनुरूप होता है। यदि मूल्यानुपात अथवा प्रतिशत विधि के आधार पर मूल्य-परिवर्तन की माप करनी हो तो दूर के वर्ष की अपेक्षा निकट वर्ष के आधार को प्राथमिक दी जानी चाहिए। अतः श्रंखला मूल्यानुपात जिन्हें बाद में श्रंखलाबद्ध किया जाता है , इस उद्देश्य पूर्ति के लिए प्रयुक्त किये जाते हैं।

तथ्य का चुनाव (Form of Average to be Used) : सैद्धान्तिक रूप से निर्देशांक की गणना में प्रत्येक माध्य जैसे अंकगणितीय माध्य, मध्यका, भूयिष्ठक, गुणोत्तर, माध्य तथा हरात्मक माध्य, का प्रयोग किया जा सकता है। परन्तु व्यावहारिक दष्टिकोण से मध्यका तथा भूयिष्ठक का प्रयोग नहीं किया जाता क्योंकि इनसे भ्रमात्मक परिणाम होते हैं। गुणोत्तर माध्य तथा हरात्मक माध्य की गणना करना कठिन होता है अतः अंकगणितीय माध्य का ही प्रयोग किया जाता है। विद्युत-चलित गणना करने वाली मशीनों का प्रयोग बढ़ने से निर्देशांकों की गणना में गुणोत्तर माध्य का प्रयोग बढ़ रहा है। अंकगणीतीय माध्य के प्रयोग से जहाँ निर्देशांकों की गणना करना सरल हो जाता है वहाँ गुणोत्तर माध्य से निर्देशांकों की गणना करने में निम्नलिखित गुण होते हैं।

(*i*) गुणोत्तर माध्य से सापेक्ष परिवर्तनों की माप श्रेष्ठ रूप से होती है।

(*ii*) गुणोत्तर माध्य कम मूल्यों को अधिक तथा अधिक मूल्यों को कम महत्व देता है। इससे वस्तुस्थिति का सन्तुलित चित्र प्रस्तुत हो जाता है।

(*iii*) सन्तोषप्रद निर्देशांकों में उत्काम्यता (reversibility) का गुण होना आवश्यक होता है तथा गुणोत्तर माध्य के प्रयोग से ही निर्देशांकों में यह गुण आ सकता है।

(*iv*) गुणाोत्तर माध्य से निर्मित निर्देशांकों में आधार वर्ष अथवा वस्तुओं में परिवर्तन करना सरल होता है।

(7) भारांकन विधि (System of Weighting) : ''सरल शब्दों में, भारांकन आंगिक-श्रेणियों को उनकी वास्तविक महत्ता के सम्बन्ध में महत्व देने के लिए किया जाता है।'' प्रत्येक वस्तु का उसके महत्व के अनुसार निर्देशांक पर प्रभाव डालने के लिए भारांकन विधि का प्रयोग किया जाता है। आभारित अथवा साधारण निर्देशांक (Unweighted of Simple Index Number) में प्रत्येक वस्तु का समान महत्व दिया जाता है। परन्तु व्यावहारिक जीवन में अलग-अलग वस्तुओं का अलग-अलग सापेक्षिक महत्व होता है। यदि गेहूँ का मूल्य दुगना तथा तम्बाकू का मूल्य आधा हो जाय तो अभारित निर्देशांक पर कोई प्रभाव नहीं पड़ेगा जबकि उपभोक्ता को तम्बाकू का मूल्य आधा रह जाने से उतना लाभ नहीं होगा जितना गेहूँ का मूल्य दुगना हो जाने से नुकसान होगा। निर्देशांकों की रचना में विभिन्न वस्तुओं को उनके सापेक्षिक महत्व के अनुसार भारांकित करने का यही प्रमुख कारण है।

निर्देशांक के प्रकार

मुख्य रूप से तीन प्रकार के निर्देशांक होते हैं :-

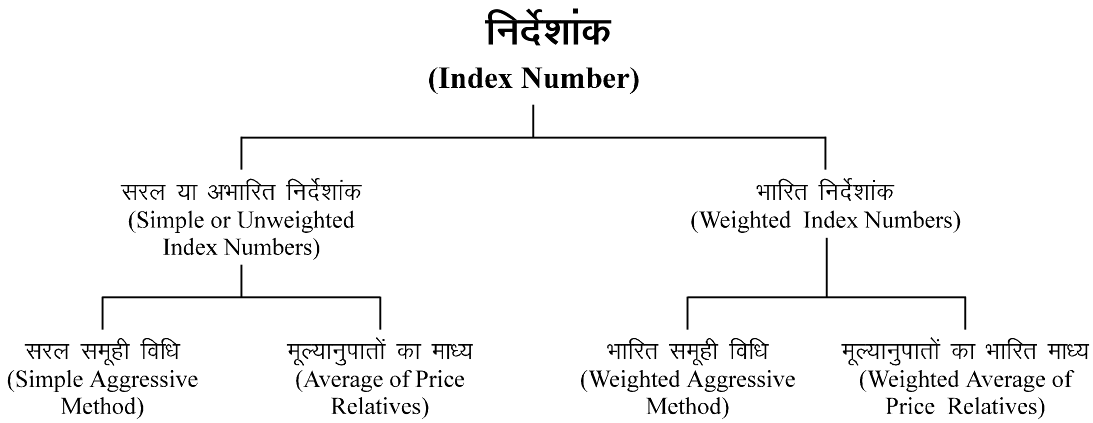
(1) **कुल मूल्य निर्देशांक (Value Index) :** इस प्रकार के निर्देशांक इस बात की तुलना करते हैं कि आधार वर्ष के कुल मूल्य (मूल्य × मात्रा) की तुलना में प्रचलित वर्ष के कुल मूल्य में (मूल्या × मात्रा) कितनी वद्धि अथवा कमी हुई है।

(2) **मूल्य निर्देशांक (Price Index) :** सामान्यतः निर्देशांक से तात्पर्य मूल्य निर्देशांक से ही होता है। ये निर्देशांक दो समयान्तरों के मध्य मूल्य-स्तर के परिवर्तनों को नापते हैं।

(3) **परिमाण निर्देशांक (Quantity Index) :** किसी वस्तु के उत्पादित, विक्रय अथवा उपभोग किए जाने वाले परिणाम के परिवर्तनों को मापने के लिए परिमान निर्देशांक बनाये जाते हैं।

सूचकांक रचना की विधियाँ
(Methods of Constructing Index Numbers)

निर्देशांक बनाने की अनेक विधियां हैं। सरलता की दष्टि से एक चाट निम्न प्रकार से प्रदर्शित किया जा सकता है : -



(*i*) **सरल समूही विधि (Simple Aggressive Method) :** निर्देशांक बनाने की यह सरलतम विधि है। इस विधि के अन्तर्गत प्रचलति वर्ष के लिए निर्देशांक की गणना निम्न सूत्र द्वारा की जाती है :-

$$P01 = \frac{P_1}{P_0} \times 100$$

जहाँ (Where)

P_{01} = प्रचलित वर्ष का निर्देशांक (Price Index of the Current Year)

SP_1 = प्रचलित वर्ष के मूल्यों का योग (Total of Price of Current Year)

SP_0 = आधार वर्ष के मूल्यों का योग (Total of Price of Base Year)

Illustration 1 : Given below are the prices of the five commodities for three years. Construct the price index with 1985 as base.

Commodities	**Price per unit**		
	1985	**1986**	**1987**
A	12	15	18
B	20	24	28
C	55	85	90
D	44	52	50
E	28	40	44
Solution :			

Construction of price index for 1986 with 1985 as base year. Let the price of commodities in 1985 be denoted by P_0; and in 1986 and 1987 be denoted by P_1 and P_2 respectively.

Commodities	Price per unit		
	1985	1986	1987
	(P_0)	(P_1)	(P_2)
A	12	15	18
B	20	24	28
C	55	85	90
D	44	52	50
E	28	40	44
	159	216	230

We have

$SP_0 = 159$, SP_1 216 and $SP_2 = 230$

Price index for 1986 $= P_{01} = \frac{P_1}{P_0} \times 100$

$$= \frac{216}{159} \times 100 = 135.85$$

Price index for 1987,

$$P_{02} = \frac{P_2}{P_0} \times 100$$

$$= \frac{230}{159} \times 10$$

$$= 144.65$$

Illustration 2 : Construt price index numbers, using simple aggregative method, for the data given below by taking :-

(*i*) 1993 as base year

(*ii*) Five years average as base year

(*iii*) 1990 to 1992 as base year

Year	Price (Rs.)
1990	7
1991	8
1992	9
1993	10
1994	21

Solution :

(*i*) Construction of Index Number taking 1993 as base year.

Years **Price** **Index Number of Price Relatives**

	(Rs.)	(1995 = 100)
1990	7	$\frac{7}{10}$ 100 70
1991	8	$\frac{8}{10} \times 100 = 80$
1992	9	$\frac{9}{10} \times 100 = 90$
1993	10	100
1994	21	$\frac{21}{10} \times 100 = 210$

(*ii*) Construction of Index Number taking five years average as base year

$$\text{Average} = \frac{7 \quad 8 \quad 9 \quad 10 \quad 21}{5} = \quad = 11$$

Years	Price (Rs.)	Index Number of Price Relatives Five years Averages (II) as base year
1990	7	$\frac{7}{11} \times 100 = 63.64$
1991	8	$\frac{8}{11} \times 100 = 72.73$
1992	9	$\frac{55}{5}$ $\frac{9}{11} \times 100 = 81.82$
1993	10	$\frac{10}{11} \times 100 = 90.91$
1994	21	$\frac{21}{11} \times 100 = 191.91$

(*ii*) Construction of Index Number taking 1990 to 1992 as base year.

$$\text{Average of prices from 1990 to 1992} = \frac{7 \quad 8 \quad 9}{3} = \quad = 8$$

Years	Price	Index Number of Price Relatives 1990 to 1992 average price as base
1990	7	$\frac{7}{8} \times 100 = 87.5$
1991	8	$\frac{8}{8} \times 100 = 100.00$
1992	9	$\frac{9}{8} \times 100 = 112.5$
1993	10	$\frac{10}{8} \times 100 = 125.0$
1994	21	$\frac{21}{8} \times 100 = 262.5$

(*ii*) **मूल्यानुपात माध्य विधि (Average of Price Relatives Method) :** इस विधि के अनुसार निर्देशांकों की रचना करने

के लिए सर्वप्रथम वस्तुओं या मदों के मूल्यानुपात (Price Relatives) ज्ञात किए जाते हैं। इसके पश्चात् समान्तर माध्य, मध्यका या गुणोत्तर माध्य में से किसी भी माध्य का प्रयोग कर निर्देशांक ज्ञात किए जाते हैं

जब समान्तर माध्य (Arithmetic Mean) का प्रयोग किया जाता है तो निम्न सूत्र का प्रयोग होता है।

$$P_{01} = \frac{\left(\frac{P_1}{P_0} \quad 100\right)}{N}$$

जहाँ (Where) N = Number of Items or Commodities

जब गुणोत्तर माध्य (Geometric Mean) का प्रयोग किया जाता है तो निर्देशांक ज्ञात करने के लिए निम्न सूत्र का प्रयोग होता है :-

$$\log P_{01} = \frac{\log\left[\frac{P_1}{P_0} \quad 100\right]}{N} \quad \frac{\log P}{N} \qquad \left(\text{जहाँ } P \quad \frac{P_1}{P_0} \quad 100\right)$$

या $$P_{01} = \text{Antilog}\left[\frac{\log P}{N}\right]$$

Illustration 3 : From the data given below, construct index for 1988 taking 1987 as base year by the average or relatives method using (a) arithmetic mean and (b) geometric mean methods.

Commodities	**Price 1987**	**Price 1988**
A	70	90
B	60	95
C	120	135
D	150	180
E	20	20
F	15	16

Solution : (*a*) Construction of index number by arithmetic mean method.

Commodities	**Price 1987 (P_0)**	**Price 1988 (P_1)**	**Price Relatives ($P_1 \times P_0 \times 100$)**
A	70	90	128.57
B	60	95	158.33
C	120	135	112.50
D	150	180	120.00
E	20	20	100.00
F	15	16	106.67
		$\Sigma P_1/P_0 \times 100 = 726.07$	

$$\text{Price index for 1988, } P_{01} = \frac{P_1 / P_0 \quad 100}{N}$$

=

= 121.01

(***b***) Construction of index number by geometric mean method.

Commodities	Price (P_0)	Price (P_1)	Price Relatives ($P_1/P_0 \times 100$)	Log P
A	70	90	128.57	2.1091
B	60	95	158.33	2.1996
C	120	135	112.50	2.0511
D	150	180	120.00	2.0792
E	20	20	100.00	2.0000
F	15	16	106.67	2.0280
				S log P = 12.4670

P_{01} = Antilog

P_{01} = Antilog S Log p/N

= 12.460/6

= antilog 2.0778

= 119.62

$\frac{726.07}{6}$ $\frac{P_1/P_0 \quad 100)}{N}$

Illustration 4 : Given below is the data of three groups of commodities price for the year 1984 to 1988. Find out the price index using arithmetic mean and geometric mean methods.

Group	1984	1985	1986	1987	1988
A	4	6	8	10	12
B	16	20	24	30	36
C	8	10	16	20	24

Solution : Computation of price index number based on arithmetic mean and geometric mean methods. Assume 1984 as base year (1984 = 100).

Group	1984	1985	1986	1987	1988
A	100	150	200	250	300
B	100	125	150	187.5	225
C	100	125	200	250	300
Total Relatives	**300**	**400**	**550**	**687.5**	**825**
Price Index based on A.M.	**100**	**133.33**	**183.33**	**229.17**	**275.00**
Price Index based on G.M.	**100**	**132.83**	**181.71**	**227.14**	**272.56**

Illustration 5 : Find out the index number for 1992, 1993 and 1994 based on 1991 using mean, median and geometric mean as base :

वस्तुएँ Commodities	1991	1992	1993	1994
A	15	130	20	24
B	10	12	16	13
C	12	18	8	10
D	8	8	12	16
E	17	15	16	20

Solution : Construction of Index Numbers.

Commodities	1991 Price	1992 Price Relatives	1993 Price	1994 Price Relatives	Price	Price Relatives	Price	Price Relatives
A	15	100	30	200.00	20	133.33	24	160.0
B	10	100	12	120.00	16	160.00	13	130.0
C	12	100	18	150.00	8	66.70	10	83.3
D	8	100	8	100.00	12	150.00	16	200.0
E	17	100	15	88.20	16	94.10	20	117.7
Total of Relatives		500		658.2		604.1		691.0
Mean of Relatives		100		131.6		120.8		138.2
Median of Relatives		100		120.0		133.3		130.0
Geometric Mean of Relatives		100		126.0		115.0		125.9

Illustration 6 : Calculate Index Number from the following data taking average price as index :-

वस्तुएँ (Articles)	मूल्य (रु. में) Price (in Rs.)		
1999	2000	2001	
A	4	5	9
B	3	6	9
C	2	4	6

Solution :

Articles	Average Price (P_0)	1999 Price	1999 Price Relatives	2000 Price	2000 Price Relatives	2001 Price	2001 Price Relatives
A	6	4	66.67	5	83.33	9	150
B	6	3	50.00	6	100.00	9	150
C	4	2	50.00	4	100.00	6	150
Total of Price Relatives			166.67		283.33		450
Average or Price			55.67		94.44		150

Relatives

Calculation of Price Relatives

For 1999	**For 2000**	**For 2001**
$\times 100 = 66.67$	$\frac{5}{6} \times 100 = 83.33$	$\frac{9}{6} \times 100 = 150$
$\frac{3}{6} \times 100 = 50.00$	$\frac{6}{6} \times 100 = 100.00$	$\frac{9}{6} \times 100 = 150$
$\frac{2}{4} \times 100 = 50.00$	$\frac{4}{4} \times 100 = 100.00$	$\frac{6}{4} \times 100 = 150$

(2) श्रंखला आधार विधि (Chain Base Method) : जैसा कि हम स्थायी आधार विधि में वर्णन कर चुके हैं कि एक बार चुना हुआ आधार स्थिर रहता है और प्रतिशत मूल्यानुपात (Price Relatives) उसी आधार वर्ष या वर्षों के औसत पर आधारित होते हैं। श्रंखला आधार विधि इससे कुछ भिन्न है। श्रंखला आधार विधि के अनुसार प्रत्येक वर्ष के लिए उससे पिछला वर्ष आधार माना जाता है। दूसरे शब्दों में, प्रत्येक वर्ष अपने अगले वर्ष के लिए आधार वर्ष होता है या कहें कि आधार निरन्तर बदलता रहता है जैसे 1989 के लिए 1988, 1990 के लिए 1989, 1991 के लिए 1990 आदि। श्रंखला आधार के निर्देशांकों की रचना की क्रिया-विधि निम्न प्रकार से है :-

1. प्रथम वर्ष के निर्देशांकों को 100 मान लिया जाता है।
2. प्रथम वर्ष के मूल्य को आधार मान कर द्वितीय वर्ष के मूल्यानुपातों की गणना की जाती है जिन्हें श्रंखला मूल्यानुपात (Link Relatives) कहते है। इसी प्रकार द्वितीय वर्ष के मूल्य को आधार मान कर ततीय वर्ष के मूल्यानुपातों की गणना की जाती है। यही क्रम अन्त तक चलता रहता है।

$$\text{श्रंखला मूल्यानुपात (Link Relative)} = \frac{\text{चालू वर्ष का मूल्य (Price of Current Year)}}{\text{पिछले वर्ष का मूल्य (Price of Previous Year)}} \times 100$$

3. प्रत्येक वर्ष के श्रंखला मूल्यानुपातों का योग करके वस्तुओं की संख्या से भाग देकर श्रंखला मूल्यानुपातों का माध्य (Average of Link Relatives) ज्ञात कर लिया जाता है।
4. श्रंखला मूल्यानुपात केवल दो वर्षों के मध्य प्रतिशत अनुपात को प्रकट करता है इस प्रकार दो निकटवर्ती वर्षों में सम्बन्ध स्थापित हो जाता है। सभी श्रंखला मूल्यानुपातों के मध्य सम्बन्ध (Link) स्थापित करने के लिए तथा एक श्रंखला का निर्माण करने के लिए सभी श्रंखला मूल्यानुपातों को सामान्य आधार से श्रंखलाबद्ध श्रंखला निर्देशांक (Chain Indices Chained to a Common Base) इसकी गणना के लिए निम्न सूत्र का प्रयोग करते हैं:-

$$\text{(Chain Index No. for Current Year)} = \left(\frac{\begin{array}{l}\text{Chain Index No. of Previous Year} \\ \quad \text{Average Link Relative of Current Year}\end{array}}{100} \right)$$

Illustration 7 : From the following data, calculate Chain Base Index Numbers.

निम्न समंकों से श्रंखला सूचकांक ज्ञात कीजिए : -

Price in Rs

Items	**1986**	**1987**	**Years** **1988**	**1989**	**1990**
A	5	8	10	12	15
B	3	6	8	10	12

Items	C		2		3		5	7		10.5
	1986		1987			1988		1989		1990
	Price	**Link Relatives**	**Price**	**Link Relatives**	**Price**	**Link Relatives**	**Price**	**Link Relatives**	**Price**	**Link Relatives**
A	5	100	8	$\frac{8}{5}\times 100 = 160$	10	$\frac{10}{8}\times 100 = 125$	12	$\frac{12}{10}\times 100 = 120$	15	$\frac{15}{12}\times 100 = 125$
B	3	100	6	$\frac{6}{3}\times 100 = 200$	8	$\frac{8}{6}\times 100 = 133.33$	10	$\frac{10}{8}\times 100 = 125$	12	$\frac{12}{10}\times 100 = 120$
C	2	100	3	$\frac{3}{2}\times 100 = 150$	5	$\frac{5}{3}\times 100 = 166.67$	7	$\frac{7}{5}\times 100 = 140$	10.5	$\frac{10.5}{7}\times 100 = 150$
Total Link Relatives		300		510		425		385		395
Average of Link Relatives		100		170		141.67		128.33		131.67
Chain Indices		100		$\frac{100 \quad 170}{100}$				$\frac{240.84 \quad 128.33}{100}$		
				$=170$		$=240.84$		$=309.07$		$=406.95$

Aggregative Method

भारांकन तर्क-संगत विधि पर ही आधारित होना चाहिए। संख्याशास्त्रियों द्वारा प्रतिपादन भारांकन के विभिन्न सूत्र इस प्रकार हैं :-

(1) आधार वर्ष की मात्रा के अनुसार भारांकन (Base Period Quantities as Weights) : इस विधि को लासपियरे (Laspeyre's) विधि भी कहते हैं। इसमें आधार वर्ष की मात्राओं के अनुसार भाग दिये जाते हैं। सूत्रानुसार :-

$$\text{Index Number} = \quad \times 100$$

इस सूत्र में

P_1 = वर्तमान वर्ष का मूल्य (Price of the Current Year)

P_0 = आधार वर्ष का मूल्य (Price of the Base Year)

q_0 = आधार वर्ष कह मात्रा (Quantity of the Base Year)

(2) वर्तमान वर्ष की मात्रा के अनुसार भारांकन (Current Year's Quantities as Weights) : इस विधि को पाश्चर्य की विधि (Pasche's Method) भी कहते हैं। इस विधि के अनुसार वर्ष का मात्राओं के अनुसार भार दिये जाते हैं। इस विधि का सबसे बड़ा गुण है कि हर वर्ष नये भार निश्चित किये जाते हैं। सूत्रानुसार :-

$$\text{Index Number} = \frac{P_1q_1}{P_0q_1} \times 100$$

संकेतों के अर्थ इस प्रकार हैं :-

P_1 = वर्तमान वर्ष का मूल्य (Price of the Current Year)

P_0 = आधार वर्ष का मूल्य (Price of the Base Year)

q_0 = आधार वर्ष कह मात्रा (Quantity of the Base Year)

(3) मार्शल-एजवर्थ की विधि (Marshall-Edgeworth's Method) इस विधि के अन्तर्गत वर्तमान तथा आधार दोनों वर्षों की मात्राओं के माध्य अथवा योग के बराबर भार दिये जाते हैं। इस प्रसिद्ध अर्थशास्त्री-मार्शल तथा एजवर्थ ने पहले दोनों सूत्रों को मिला कर यह सूत्र प्रस्तुत किया है। उनका सूत्र है :-

$$\text{Index Number} = \frac{P_1\,(q_0 \quad q_1)}{P_0\,(q_0 \quad q_1)} \times 100 \text{ अथवा } \frac{P_1\,q_0 \qquad P_1\,q_0}{P_0\,q_0 \qquad P_0\,q_1} \times 100$$

(क) निर्देशांक में शामिल समस्त वर्षों की मात्राओं के माध्य के आधार पर भी भार निश्चित किये जा सकते हैं। ऐतिहासिक अध्ययन के लिए भारांकन की यह विधि सर्वोचित है परन्तु इसके प्रयोग में व्यावहारिक कठिनाइयाँ हैं।

(ख) महत्वपूर्ण वर्षों की मात्राओं के मध्य के आधार पर भी भार निश्चित किये जा सकते हैं।

(ग) **मात्राओं के महत्तम समापवर्त के आधार पर भारांकन (Highest Common Factor of the Quantities as Weights) :** भारांकन की इस विधि का प्रतिपादन जे. एम. कीन्स (J. M. Keynes) ने किया है। इस विधि के अनुसार, प्रत्येक वस्तु की समस्त वर्षों अथवा आधार वर्ष तथा वर्तमान वर्ष की मात्रा की समान्य मात्रा (Common Quantity) के आधार पर भार निश्चित किये जाते हैं।

(4) फिशर की विधि (Fisher's Method) : इस विधि के अनुसार आधार वर्ष तथा वर्तमान वर्ष की मात्राओं के आधार पर भार देकर दो निर्देशांक निर्मित किये जाते हैं और उन दोनों का गुणोत्तर माध्य (Geometri Mean) निकाल लिया जाता है। इस विधि को Crossed Weights Formula भी कहते हैं। इसका सूत्र निम्न है :-

$$\text{Index Number} = \sqrt{\frac{P_1\,q_0}{P_0\,q_0} \quad \frac{P_1\,q_1}{P_0\,q_1}} \times 100$$

संकेतों का अर्थ इस प्रकार है :-

P_1 = वर्तमान वर्ष का मूल्य (Price of the Current Year)

P_0 = आधार वर्ष का मूल्य (Price of the Base Year)

q_1 = वर्तमान वर्ष कह मात्रा (Quantity of the Current Year)

q_0 = आधार वर्ष कह मात्रा (Quantity of the Base Year)

(5) डोरबिश तथा बाउले का सूत्र (Dorbish and Bowley Formula) : यह सूत्र फिशर के सूत्र की तरह ही है। परन्तु इसमें गुणोत्तर माध्य के स्थान पर अंकगणितीय माध्य निकाला जाता है। इनका सूत्र निम्न है :-

$$\text{Index Number} = \text{—}\left\{\left(\frac{P_1\,q_0}{P_0\,q_0} \quad \frac{P_1\,q_1}{P_0\,q_1}\right) \quad 2\right\} \quad 100$$

(6) वॉल्श का सूत्र (Walsh's Formula) :-

$$\text{Index Number} = \frac{\sqrt{q_0(q_1 P_1)}}{\sqrt{q_0(q_1 P_0)}}$$

Illustration 8 : From the following data, calculate Laspeyre's, Passche's-Dorbish-Bowley's. Fisher's Index and Marshall-Edgeworth's, and Walsch's Index Numbers.

Item (Rs.)	Base Year Price (Rs.)	Base Year Quantity (Rs.)	Current Year Price (Rs.)	Current Year Quantity
A	6	50	10	56
B	2	100	2	120
C	4	60	6	60
D	10	30	12	24
E	8	40	12	36

(Madras Uni. B.Com March 1978)

Solution : Computation of Laspeyre's and Paashe's Index Numbers.

Item	P_0	q_1	P_1	q_1	P_0q_0	P_0q_1	P_1q_0	P_1q_1
A	6	50	10	56	300	336	500	560
B	2	100	2	120	200	240	200	240
C	4	60	6	60	240	240	360	360
D	10	30	12	24	300	240	360	288
E	8	40	12	36	320	288	480	432
					SP_0q_0 = 1360	SP_0q_1 = 1344	SP_1q_0 = 1900	SP_1q_1 = 1880

(*i*) **Laspeyre's Price Index is given by** 2.7959 / 2 $\left[\frac{P_1\,q_0}{P_0\,q_0} \quad \frac{P_1\,q_1}{P_0\,q_1}\right]$

$$P_{01} = \frac{P_1\,q_0}{P_0\,q_0} \times 100 = \frac{1900}{1360} \times 100 = 1.3971 \times 100 = \mathbf{139.71}$$

(*ii*) **Paasche's Price Index is given by**

$$P_{01} = \frac{P_1\,q_1}{P_0\,q_1} \times 100 = \frac{1880}{1344} \times 100 = 1.3988 \times 100 = \mathbf{139.88}$$

(*iii*) **Dorbish and Bowley's Index**

$$P_{01} = \frac{\frac{P_1\,q_0}{P_0\,q_0} \quad \frac{P_1\,q_1}{P_0\,q_1}}{2} \times 100 = \frac{1.3971 \quad 1.3988}{2} \times 100$$

$$= \quad \times 100 = 1.39.795 \times 100 = \mathbf{139.795}$$

\ P_{01} = **139.80**

(*vi*) **Marshall-Edgeworth Method**

$$P_{01} = \quad \times 100 = \frac{1900 \quad 1880}{1360 \quad 1344} \times 100$$

$$= \frac{3780}{2704} \times 100$$

\ P_{01} = 1.39 × 100 = **139.8**

(*v*) Fisher's Ideal Formula

$$P_{01} = \sqrt{\frac{P_1\,q_0}{P_0\,q_0}\;\frac{P_1\,q_1}{P_0\,q_1}} \times 100 = \sqrt{\frac{1900}{1360}\;\;\frac{1880}{1344}} \times 100$$

$$= \sqrt{1.9543} \times 100$$

(*vi*) Walsch's Index

Calculation of Walsch's Price Index

Item	$q_0\ q_1$	$\sqrt{q_0 \cdot q_1}$	$P_1\sqrt{q_0 \cdot q_1}$	$P_0\sqrt{q_0 \cdot q_1}$
A	2800	52.9	529	317.4
B	12000	109.5	219	219.0
C	3600	60.0	360	240
D	720	26.8	321.6	268
E	1440	37.9	454.8	303.2
			1884.4	1347.6
			$S\ P_1\sqrt{q_0 \cdot q_1} = S\ P_1\sqrt{q_0 \cdot q_1}$	

$$P01 = \frac{P_1\sqrt{q_0 \cdot q_1}}{P_0\sqrt{q_0 \cdot q_1}} \times 100 = \frac{1884.4}{1347.6} \times 100$$

$$= 1.398 \times 100$$

\ $$S\ P_{01} = 139.08$$

भारित निर्देशांक

(Weighted Index Numbers)

(1) Weighted Average of Price Relatives

यद्यपि सरल या अभारित निर्देशांकों की रचना करनी काफी सरल होत है परन्तु फिर भी इन्हें सन्तोषजनक नहीं माना जा सकता। अभारित निर्देशांकों की रचना करते समय प्रत्येक वस्तु को समान महत्व दिया जाता है जबकि वास्तव में समस्त वस्तुएँ समान महत्व की नहीं होतीं। जैसे टमाटर के मूल्य में वद्धि होना उतना महत्वपूर्ण नहीं है जितना कि गेहूँ के मूल्य में वद्धि होना महत्वपूर्ण है। अतः ऐसी परिस्थिति में यह अनिवार्य हो जाता है कि प्रत्येक वस्तु को उसके महत्व के अनुसार भारांकन किया जाए। ऐसा करने से हम शुद्ध निर्देशांक की प्राप्ति कर सकते हैं।

किसी वस्तु के लिए दिए जाने वाले भार की मात्रा का अनुमान लगाने के लिए यह देखना होता है कि वर्ग विशेष के उपभोक्ता अपनी आय का कितना भाग उस वस्तु के लिए खर्च करते हैं। जिस वस्तु के लिए आय का जितना अधिक भाग खर्च किया जाता है वह उतनी ही महत्वपूर्ण होती है। कुल खर्च में से प्रत्येक वस्तु का प्रतिशत खर्च ज्ञात किया जाता है और उसी के आधाार पर भारांकन किया जाता है। इस सम्बन्ध में यह बात ध्यान रखने योग्य है कि वस्तुओं का महत्व, स्थान, समय, उपयोगिता एवं व्यापारियों की रुचियों के अनुसार परिवर्तित होता रहता है। इस प्रकार भार दो प्रकार के हो सकते हैं :-

(*i*) स्थिर (Fixed)

(*ii*) परिवर्तनशील (Fluctuating)

स्थिर भाव वे होते हैं जो एक बार निश्चित कर दिये जाते हैं और जब भी निर्देशांकों की रचना कर दी जाती है उन्हीं भारों (Weights) को प्रयोग में लाया जाता है। इसके विपरीत, **परिवर्तनशील भार** वे होते हैं जो समय-समय पर परिवर्तित किए जाते

हैं। आजकल विभिन्न प्रकार की नई-नई वस्तुएँ बाजार में आ रही हैं और पुरानी वस्तुओं की समाप्ति हो रही है। इन परिवर्तनों के कारण या आवश्यक हो जाता है कि भार देने की विधि का समय-समय पर संशोधन किया जाए। निर्देशांकों की रचना में परिवर्तनशील भार अधिक उपयोगी होते हैं क्योंकि वस्तुओं के सापेक्षिक महत्व में होने वाले परिवर्तनों का मापन सम्भव बनाया जा सकता है।

क्रिया विधि

(अ) जब विभिन्न वस्तुओं के मूल्य निर्देशांक एवं भार दिए गए हों (When Price Index No. and Weights of different Articles are given) :

(1) मूल्य निर्देशांकों (Price Relatives) को भार (Weights) से गुणा करके (PW) की गणना कीजिए।

(2) मूल्य निर्देशांकों एवं भारों की गुणा का योग करके (S PW) ज्ञात कीजिए।

(3) भारों का योग करके (S W) से भाग देकर भारित निर्देशांकों की गणना कर ली जाती है।

(ब) जब विभिन्न वस्तुओं के मूल्य एवं भार दिए गए हों (When Price and Weights of different Articles are given):

(1) मूल्यानुपात (Price Relatives) ज्ञात कीजिए।

$$P = \frac{P_1}{P_0} \times 100$$

(2) मूल्यानुपातों (P) को भार (W) से गुणाा करके (S PW) ज्ञात कीजिए।

(3) (S PW) को (S W) से भाग देकर भारित निर्देशांकों की गणना कर ली जाती है।

Illustration 9 : An enquirey into the budgets of middle class families in a city in England gave the following information.

खर्चा (Expense on)	**भोजन (Food)**	**किराया (Rent)**	**कपड़े (Clothing)**	**ईंधन (Fuel)**	**विविध (Miscellaneous)**
	35%	15%	20%	10%	20%
मूल्य (Price) (2000)	Rs. 150	Rs. 30	Rs. 75	Rs. 25	Rs. 40
मूल्य (Price) (2001)	Rs. 145	Rs. 30	Rs. 65	Rs. 23	Rs. 45

2000 की तुलना में 2001 में निर्वाह व्यय सूचकांक में क्या परिवर्तन होगा ?

What changes in the cost of living figures of 2001 as compared with 2000 are seen ?

हल (Solution) : Construct of Weighted Index No.

Expense on	**Price in 2000** (P_0)	**Price in 2001** (P_1)	**Price Relative** $\left(\frac{P_1}{P_2} \quad 100\right)$ **P**	**W**	**PW**
Food	150	145	96.7	35	3384.5
Rent	30	30	100.0	15	1500.0
Clothing	75	65	86.7	20	1734.0
Fuel	25	26	92.0	10	920.0
Misc.	40	45	112.5	20	2250.0
				S W = 100	S PW = 9788.5

$$\text{Cost of Living Index No.} = \frac{PW}{W} = \frac{9788.5}{100}\ 97.9$$

Illustration 10 : Construct a Price Index Number for the following data :

Raw Material	Price (1968)	Price (1970)	Weights
A	16	19	5
B	24	25	1
C	13	18	3
D	8	9	6
E	12	14	4
F	4	8	3

Solution :

Raw Material	Weights W	1968			1970		
		Price P_0	Relative R_0	R_0 XW	Price P_1	Relative R_1	$R_1 \times W$
A	5	16	100	500	10	118.75	593.75
B	1	24	100	100	25	104.17	104.17
C	3	13	100	300	18	138.46	415.38
D	6	8	100	600	9	112.50	675.00
E	4	12	100	400	14	116.67	466.68
F	3	4	100	300	8	200.00	600.00
Total	22			2200			2854.95

$$\text{Index Number for 1970} = \frac{\Sigma(R_1 W)}{\Sigma W} = \frac{2854.95}{22} = 129.77$$

आदर्श निर्देशांक के परीक्षण (Test of an Ideal Index) : यदि कोई निर्देशांक निम्नलिखित परीक्षणों पर खरा उतरता है तो आदर्श कहा जाएगा अन्यथा नहीं।

(1) समय-उत्क्राम्यता परीक्षण (Time Reversal Test)

(2) तत्त्व-उत्क्राम्यता परीक्षण (Factor Reversal Test)

(3) चक्रीय परीक्षण (Circular Test)

समय-उत्क्राम्यता परीक्षण (Time Reversal Test) : फिशर के कथानुसार, ''इस परीक्षण का अभिप्राय यह है कि निर्देशांक-रचना का सूत्र ऐसा होना चाहिए जो तुलनात्मक व्याख्या के दो बिन्दुओं के मध्य समान अनुपात व्यक्त करे चाहे उनमें से किसी को भी आधार माना जाए।'' इसे समीकरण रूप में इस प्रकार सिद्ध किया जा सकता है :-

$$P_{01} \times P_{10} = 1$$

P_{10} ज्ञात करने के P_{01} वाले सूत्र में 0 के स्थान पर 1 और 1 के स्थान पर 0 लिख दिया जाता है।

फिशर निर्देशांक के लिए

$$P_{01} = \sqrt{\frac{P_1 q_0}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_1}}$$

$$P_{10} = \sqrt{\frac{P_0 q_1}{P_1 q_1} \quad \frac{P_0 q_0}{P_1 q_0}}$$

$$P_{01} \times P_{10} = \sqrt{\frac{P_1 q_0}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_1} \quad \frac{P_1 q_0}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_1}} = 1$$

इसका अभिप्राय यह है कि यदि 1994 के आधार पर ज्ञात किया हुआ 1995 का निर्देशांक या प्रदर्शित करे कि गेहूँ का मूल्य दोगुना हो गया है तो 1995 के आधार पर ज्ञात किया हुआ 1994 का निर्देशांक यह स्पष्ट करे कि 1990 में गेहूँ का मूल्य 1995 की तुलना में आधा रह गया है अर्थात् $2 \times \frac{1}{2} = 11$

फिशर का आदर्श सूत्र इस परीक्षण पर खरा उतरता है।

तत्त्व-उत्क्राम्यता परीक्षण (Factor Reversal Test) : फिशर के कथानानुसार, ''जिस तरह हमारे सूत्र के अनुसार यह सम्भव होना चाहिए कि दो समयों के अन्र्तपरिवर्तन से असंगत फल न प्राप्त हो उसी प्रकार यह भी सम्भव होना चाहिए कि मूल्य एवं मात्राओं के प्रतिस्थान करने पर भी असंगत परिणाम न प्राप्त हों अर्थात् दोनों निष्कर्षों के परस्पर गुणा करने पर वास्तविक मूल्य अनुपात (True Value Ratio) प्राप्त हो।''

इस समीकरण-रूप से इस प्रकार सिद्ध किया जा सकता है :-

$$P_{01} \times q_{01} = \frac{P_1 q_1}{P_0 q_0}$$

$$P_{01} = \sqrt{\frac{P_1 q_0}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_1}}$$

$$q_{01} = \sqrt{\frac{q_1 P_0}{q_0 P_0} \quad \frac{q_1 P_1}{q_0 P_1}}$$

$$P_{01} \times q_{01} = \sqrt{\frac{P_1 q_0}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_1} \quad \frac{q_1 P_0}{q_0 P_0} \quad \frac{q_1 P_1}{q_0 P_1}}$$

$$= \sqrt{\frac{P_1 q_1}{P_0 q_0} \quad \frac{P_1 q_1}{P_0 q_0}} = \sqrt{\frac{(P_1 q_1)^2}{(P_0 q_0)^2}} — = \frac{P_1 q_1}{P_0 q_0}$$

फिशर द्वारा प्रतिपादित सूत्र तत्त्व उत्क्राम्यता परीक्षण (Factor Reversal Test) को भी पूरा करता है, इसीलिए उनके सूत्र को 'आदर्श सूत्र' (Ideal Formual) माना जाता है। फिर भी बोडिंगटन का मत है कि, ''अभाग्यवश, जबकि यह सूत्र पूर्ण निर्देशांक सूत्र को अधिकांश गणितीय विशेषताओं को समान्य रूप से पूरा करता है, परन्तु इसका विरोध इसलिए किया जाता है कि यह स्पष्ट नहीं है कि यह क्या माप करता है, अर्थात् परिणाम में मूल्य तथा मात्रा दोनों के परिवर्तन शामिल होते हैं जबकि हम सामान्यतः एक को दूसरे से अलग रखना चाहते हैं।''

निर्देशांक सूत्र की पर्याप्तता का एक ही परीक्षण होता हे जिसे चक्री परीक्षण कहते हैं। सर्वप्रथम इस परीक्षण का सुझाव वेस्टरगार्ड ने दिया था तथा वाल्श ने इसको अत्यधिक पसन्द किया था जिसने इसको चक्रीय परीक्षण की संज्ञा दी थी। इस परीक्षण को निर्देशांक का कोई भी सूत्र पूरा नहीं करता है। यह समय-उत्क्राम्यता परीक्षण आ ही विस्तत स्वरूप है। इस परीक्षण के अनुसार यदि 1980 (वर्तमान वर्ष '1') का निर्देशांक 1960 के आधार (आधार वर्ष '0') पर बनाया जाए तथा एक अन्य निर्देशांक असंगत नहीं होना चाहिए। दूसरे शब्दों में यदि 1970 का निर्देशांक 1960 का दुगना हो तथा 1960 का निर्देशांक 1950 के निर्देशांक का चार गुना हो तो 1950 की तुलना में 1970 का निर्देशांक $2 \times 4 = 8$ गुना होना चाहिए।

यदि P_{01} = Index No. for the current year based on '0' period.

P_{20} = Index No. for the '0' period based on '2' period.

P_{21} = Index No. for the current year based on '2' period.

तो $= P_{01} \times P_{20} \times P_{21}$

यह परीक्षण निम्न दो परिस्थितियों में पूरा होता है :-

(1) जब निर्देशांक रचना में भार का प्रयोग न किया गया हो।

(2) जब निर्देशांक में स्थिर भार का प्रयोग किया गया हो।

फिशर का आदर्श सूत्र इस घटना पर खरा नहीं उतरता है।

सूचकांक रचना सम्बन्धी विविध समस्याएँ

(Miscellaneous Problem Regarding Construction of Index Number)

(I) आधार परिवर्तन के नियम (Rules of Base Conversion)

कभी-कभी अनुसन्धान या अध्ययन के लिए स्थिर आधार के सूचकांकों को श्रंखला के आधार और श्रंखला आधार के सूचकांकों को स्थिर आधार में परिवर्तित करना आवश्यक हो जाता है। इसके सम्बन्ध में नियम इस प्रकार है :-

(1) स्थिर आधार से श्रंखला आधार (From Fixed Base to Chain Base) : स्थिर आधार सूचकांकों को श्रंखला आधार सूचकांकों में बदलने के लिए निम्न प्रक्रिया प्रयोग में लाई जाती है :-

(*i*) पहले वर्ष का सूचकांक 100 माना जाता है।

(*ii*) अगले वर्षों के सूचकांक के लिए निम्न सूत्र का प्रयोग किया जाता है।

$$\text{Chain Base Index Number} = \frac{\text{Current Year Fixed Base Index Number}}{\text{Previous Year Fixed Base Index Number}} \times 100$$

Illustration 11 : Convert into chain base Index No. from fixed base Index No.

निम्न स्थिर आधार सूचकांकों को श्रंखला आधार सूचकांकों में परिवर्तित कीजिए :

Year (वर्ष)	:	1981	1982	1983	1984	1985
Fixed Base Indices (स्थिर आधार सूचकांक)	:	100	95	114	228	285

Solution : Fixed Base to Chain Base.

Year	Fixed Base Indices	Conversion into Chain base Index Number
1981	100	100
1982	95	$\frac{95}{100} \times 100 = 95$
1983	114	$\frac{114}{95} \times 100 = 120$
1984	228	$\frac{228}{114} \times 100 = 200$
1985	285	$\frac{285}{228} \times 100 = 125$

(2) श्रंखला आधार में स्थिर आधार (From Chain Base to Fixed Base) : श्रंखला आधार सूचकांकों को स्थिर आधार सूचकांकों में बदलने के लिए निम्न प्रक्रिया अपनाई जाती है :-

(*i*) पहले वर्ष के लिए स्थिर आधार पर मूल्य सूचकांक वही माना जाता है जो उस वर्ष का श्रंखला आधार सूचकांक होता है। परन्तु यदि प्रश्न में पहले वर्ष को आधार मानने के लिए कहा गया है तो सूचकांक 100 माना जाएगा।

(*ii*) अगले वर्षों के सूचकांकों के लिए निम्न सूत्र का प्रयोग किया जाएगा :

$$\text{Fixed Base Index No.} = \frac{\begin{array}{c}\text{Current Year Chain Base Index Number}\\ \text{Previous Year Fixed Base Index Number}\end{array}}{100}$$

Illustration 12 : Convert into fixed base from the following chain base Indices :-

Year	**:**	1984	1985	1986	1987	1988
Chain Index	**:**	80	110	120	90	140

Solution : Computation of fixed based index numbers from chain based index numbers.

Year	Chain Index		Fixed Index
1984	80		80.00
1985	110	$110 \times 80/100$	88.00
1986	120	$120 \times 88/100$	105.60
1987	90	$90 \times 105.6/100$	95.04
1988	140	$140 \times 95.04/100$	133.06

(3) आधार परिवर्तन (Base-Shifting) : कभी-कभी यह आवश्यक हो जाता है कि निर्देशांक के आधार में परिवर्तन किया जाय। यह या तो इसलिए किया जाता है कि पहला आधार अनुपयुक्त हो गया है तथा अब उससे तुलना करने में कोई महत्व दिखाई नहीं देता अथवा दो निर्देशांक श्रेणियों में, जिनकी रचना अलग-अलग आधार से की गई है, तुलना करना आवश्यक हो; आधार परिवर्तन दो प्रकार से किया जा सकता है :-

(*i*) नये आधार वर्ष के मूल्यों को 100 मानकर फिर से सभी वर्षों के मूल्यानुपात ज्ञात करके निर्देशांकों की गणना की जाय। इसमें काफी गणना-क्रिया करनी पड़ती है। $\frac{100}{50}$

(*ii*) लघु रीति द्वारा, नये आधार वर्ष के निर्देशांक से सभी वर्षों के निर्देशांकों में भाग देकर प्राप्त परिणामों को गुणा किया जाय। यदि निर्देशांकों की रचना में गुणोत्तर माध्य का प्रयोग किया गया हो तो इस रीति से सही परिणाम प्राप्त होते हैं।

Illustration 13 : From the index numbers given below, find out index numbers by shifting base from 1950 to 1960.

Year	**:**	1950	1951	1952	1960	1961	1962	1963
Index No.	**:**	100	76	68	50	60	70	75

Solution :

Year	Index No.	Index No. with base 1960
1950	100	$100 \times \quad = 200$
1951	76	$76 \times \frac{100}{50} = 152$
1952	68	$68 \times \frac{100}{50} = 136$
1960	50	$50 \times \frac{100}{50} = 100$
1961	60	$60 \times \frac{100}{50} = 120$
1962	70	$70 \times \frac{100}{50} = 140$
1963	75	$75 \times \frac{100}{50} = 150$

(4) शिरोबन्धन (Spacing) : कभी-कभी एक विशेष आधार वर्ष पर रचित निर्देशांक श्रेणी की रचना बन्द कर दी जाती है तथा उस निर्देशांक के बन्द होने वाले वर्ष को आधार मान कर कई निर्देशांक श्रेणी की रचना प्रारम्भ कर दी जाती है। ऐसी स्थिति में तुलनात्मक अध्ययन के लिए यह आवश्यक हो जाता है कि नवीन निर्देशांक श्रेणी का पुरानी श्रेणी से शिरोबन्ध (Splicing) किया जाय। शिरोबन्धन के लिए दोनों श्रेणियों के सामान्य-वर्ग के निर्देशांक का अनुपात ज्ञात कर नई श्रेणी के निर्देशांकों से गुणा कर दिया जाता है।

Illustration 14 : Given below are two sets of indices one with 1939 as base and the other with 1947 as base :-

(A)	Year	1939	1940	1941	1942	1943	1944	1945	1946	1947
Index Nos.		100	120	130	200	300	350	370	380	400
(B)	Year	1947	1948	1949	1950	1951	1952	1953	1954	
Index Nos.		100	110	90	98	101	110	98	96	

It is desired to splice the second index number (B) with 1947 base to the first index (A) for the sake of continuity. You are to prepare a combined series with 1939 as base.

Solution :

Year	Old Index Nos.	New Index Nos.	Splicing Technique Index No. Spliced to Old
1939	100	—	
1940	120	—	
1941	130	—	
1942	200	—	
1943	300	—	
1944	350	—	
1945	370	—	
1946	380	—	
1947	400	100	$\frac{400}{100} \times 100 = 400$
1948	—	110	$\frac{400}{100} \times 110 = 440$
1949	—	90	$\frac{400}{100} \times 90 = 360$
1950	—	98	$\frac{400}{100} \times 98 = 392$
1951	—	101	$\frac{400}{100} \times 1014 = 404$
1952	—	110	$\frac{400}{100} \times 110 = 440$
1953	—	98	$\frac{400}{100} \times 98 = 392$

1954	—	96	$\frac{400}{100} \times 96 = 384$

(5) निर्देशांकों की अपस्फीति (Deflating Index Nos.) मूल्य-स्तर अथवा निर्वाह-व्यय में होने वाले परविर्तनों के आधार पर मौद्रिक आय या मजदूरी आदि से वास्तविक आय या मजदूरी (Rage Wage or Income) ज्ञात करना तथा वास्तविक आय या मजदूरी निर्देशांक गणना, अर्ध-सांख्यिकी में आवश्यक होता है। वास्तविक मजदूरी निम्न सूत्रानुसार ज्ञात की जाती है।

$$\text{Real Wage} = \frac{\text{Money Wage}}{\text{Price Index}} \times 100$$

तथा $$\text{Real Wage or Income Index No.} = \frac{\text{Index of Money Wages}}{\text{Price Index Number}} \times 100$$

अथवा

$$= \frac{\text{Real Wages of the Current Year}}{\text{Real Wage of the Base Year}} \times 100$$

Illustration 15 : The following table gives the annual income of a teacher and the general index number of price for nine years.

Year	1939	1940	1941	1942	1943	1944	1945	1946	1947
Income (Rs.)	360	420	500	550	600	640	680	720	750
General Index Number of Prices	100	104	115	160	280	290	300	320	330

Prepare the Index Numbers to show changes in the real income of the teacher and discuss the effects of a rise in the general level of prices of his real income.

Solution

Year	**Income in Rs.**	**I.No.**	**Conversion**	**Real Income**	**Conversion**	**Real Income Index No.**	**Money Income**	**Real Income**	**Money Income** $\frac{\textbf{I. No.}}{\textbf{Price Index no}}\times 100$ **= Real Income I. No.**
1939	360	100	$\frac{360}{100} \times 100$	360	—	100	—	100	$\frac{100}{100} \times 100 = 100$
1940	420	104	$\frac{420}{104} \times 100$	403.8	$\frac{403.8}{360} \times 100$	112.2	$\frac{420}{360} \times 100$	116.7	$\frac{116.7}{104} \times 100 = 112.2$
1941	500	115	$\frac{500}{115} \times 100$	434.8	$\frac{434.8}{360} \times 100$	120.8	$\frac{500}{360} \times 100$	138.9	$\frac{238.9}{115} \times 100 = 120.8$
1942	550	160	$\frac{550}{160} \times 100$	343.75	$\frac{343.75}{360} \times 100$	95.5	$\frac{550}{360} \times 100$	152.8	$\frac{152.8}{160} \times 100 = 95.5$
1943	600	280	$\frac{600}{280} \times 100$	214.3	$\frac{2143.3}{360} \times 100$	59.5	$\frac{600}{360} \times 100$	166.7	$\frac{166.7}{280} \times 100 = 59.5$
1944	640	290	$\frac{640}{290} \times 100$	220.7	$\frac{220.7}{360} \times 100$	61.3	$\frac{640}{360} \times 100$	177.8	$\frac{177.8}{290} \times 100 = 61.3$
1645	680	300	$\frac{680}{300} \times 100$	226.7	$\frac{226.7}{360} \times 100$	63.0	$\frac{680}{360} \times 100$	188.9	$\frac{188.9}{300} \times 100 = 63$
1946	720	320	$\frac{720}{320} \times 100$	225.0	$\frac{225}{360} \times 100$	62.5	$\frac{720}{360} \times 100$	200	$\frac{200}{320} \times 100 = 62.5$
1947	750	330	$\frac{750}{330} \times 100$	227.39	$\frac{227.3}{360} \times 100$	63.1	$\frac{750}{360} \times 100$	208.5	$\frac{208.3}{330} = 6.31$

मात्राओं का सूचकांक

(Index Number of Quantities)

मात्राओं के सूचकांक उत्पादन में होने वाले परिवर्तनों को ज्ञात करने के लिए बनाए जाते हैं। इन सूचकांकों की रचना की विधि मूल्य सूचकांकों जैसी ही है। अन्तर केवल इतना है कि इनके निर्माण में मूल्य के स्थान पर मात्रा (Quantity) का प्रयोग किया जाता है। सर्वप्रथम निम्न सूत्र द्वारा मात्रानुपात (Quantity Relative) ज्ञात किए जाते हैं :-

$$\text{Quantity Relative} = \frac{q_1}{q_0} \times 100$$

Where, q_1 = Quantity of the Current year
q_0 = Quantity of the Base Year

इसके पश्चात् निम्न सूत्र का प्रयोग किया जाता है :-

$$Q_{01} = \frac{R}{N}$$

Where, R = Quantity Relatives
N = Quantity of Items

यदि मूल्य भी दिए हैं तो निम्नलिखित सूत्रों का प्रयोग करके सूचकांक प्राप्त किए जा सकते हैं :-

(1) **लेसपियर की सूत्र** $Q_{01} = \frac{q_1p_0}{q_0p_0} \times 100$

(2) **पासचे का सूत्र** $Q_{01} = \frac{q_1p_1}{q_0p_1} \times 100$

(3) **फिशर का सूत्र** $Q_{01} = \sqrt{\frac{q_1p_0}{q_0p_0} \quad \frac{q_1p_1}{q_0p_1}} \times 100$

(4) **कैली का सूत्र** $Q_{01} = \frac{q_1p}{q_0p} \times 100$

Illustration 16 : From the following data, construct Quantity Index Number by using (a) Lapeyre's Formula (b) Paasche's Formula (c) Fisher's Formula

Commodity (वस्तुएँ)	**Base year** (आधार वर्ष)		**Current Year** (वर्तमान वर्ष)	
	Quantity (मात्रा)	**Price** (कीमत)	**Quantity** (मात्रा)	**Price** (कीमत)
A	8	5	10	7
B	12	4	15	6
C	15	3	20	4

Solution :

Commodity	**Base Year**		**Current Year**		q_0p_0	q_0p_1	q_1p_1	q_1p_0
	Quantity (q_0)	**Price (p_0)**	**Quantity (q_1)**	**Price (p_0)**				
A	8	5	10	7	40	56	70	50
B	12	4	15	6	48	72	90	60
C	15	3	20	4	45	60	80	60
					$\Sigma q_0p_0 = 133$	$\Sigma q_0p_1 = 188$	$\Sigma q_1p_1 = 240$	$\Sigma q_1p_1 = 170$

उपभोक्ता मूल्य निर्देशांक (Consumer Prie Index Number)

कभी-कभी किसी वर्ग विशेष के जीवन-निर्वाह सम्बन्धी निर्देशांकों की रचना की जाती है। इनका उद्देश्य मूल्य-स्तर जानना नहीं बल्कि यह जानना है कि निर्वाह व्यय में कितनी वद्धि या कमी हुई। इस प्रकार के निर्देशांकों की रचना इस बात की जानकारी प्राप्त करने के लिए की जाती है कि देश की आर्थिक प्रगति का किसी वर्ग विशेष के जीवन-निर्वाह पर क्या प्रभाव पड़ा है।

उपभोक्ता मूल्य-निर्देशांकों की रचना सामान्य जीवन में प्रयोग की जाने वाली वस्तुओं के मूल्यों को आधार मानकर की जाती है। सामान्य जीवन में प्रयोग (या उपभोग) की जाने वाली वस्तुओं को निम्न प्रमुख पाँच श्रेणियों में विभक्त किया जा सकता है :-

(*i*) खाद्य सामग्री (Food)

(*ii*) वस्त्र (Clothing)

(*ii*) ईंधन एवं प्रकाश (Fuel and Lighting)

(*iv*) मकान का किराया (House Rent)

(*v*) विविध (Miscellaneous)

उपभोक्ता-मूल्य निर्देशांकों की रचना से श्रमिकों या सामान्य नौकरी पेशा लोगों के मंहगाई भत्ते एवं वेतन आदि का निर्धारण करने में, राशनिंग व्यवस्था को लागू करने में तथा मूल्य नियन्त्रण में (यदि मूल्य में वद्धि हो रही हो) सहायता मिलती है।

उपभोक्ता-मूल्य निर्देशांक रचना की विधियां (Methods of Construction of Consumer Price Index Number) :

उपभोक्ता-मूल्य निर्देशांक रचना की निम्न दो विधियाँ हैं :-

(1) **समूही व्यय विधि या भारित समूही विधि** (Aggrerative Expenditure Method or Weighted Aggregative Method)

(2) **पारिवारिक आय व्यय विधि या भारित मूल्यानुपात विधि** (Family Budget Method or Weighted Average of Price Relatives Method)

समूही व्यय निधि (Aggregative Expenditure Method) : इस विधि द्वारा निर्देशांक ज्ञात करने के लिए निम्न सूत्र प्रयोग किया जाता है :-

$$P_{01} = \frac{p_1q_0}{p_0q_0} 1 \times 100$$

जहाँ (Where) P_{01} = Consumer Price Index Number

पारिवारिक आय व्यय निधि (Family Budget Method) : इस विधि द्वारा निर्देशांक ज्ञात करने के लिए निम्न सूत्र प्रयोग किया जाता है :-

$$P_{01} = \frac{PV}{V}$$

$$P = \text{Price Relatives}$$

$$\text{Price Relatives} = \frac{P_1}{P_0} \times 100$$

$$V = p_0q_0$$

$$\text{S}PV = \text{Total of product of price relatives (P) and weights V } (p_0q_0)$$

$$\text{S}V = \text{Total of weights}$$

नीचे दिये गय आंकड़ों में

(*i*) सामूहिक व्यय रीति तथा

(*ii*) पारिवारिक बजट रीति से उपभोक्ता मूल्य निर्देशांक बनाएं।

Illustraion 13 : From the data given below calculate the cost of living index number for the current year by the (1) Aggregate Expenditure Method and (2) Family Budget Method :

Article in Base Year	Quantity Consumed Year	Unit Year	Price in Base	Price in Current
Rice	5 quintal	Quintal	24	30
Wheat	1 quintal	Quintal	16	20
Pulses	2 quintal	Quintal	12	18
Ghee	4 kg.	kg.	5	6.25
Oil	20 kg.	per 40 kg.	40	50
Clothing	40 meters	Meter	1	1.50
Firewood	10 quintals	Quintal	2	2.50
House Rent	—	House	20	25

Solution :-

Aggregate Expenditure Method

Article	Quantity Consumed In Base Year	Unit	Price in Base Year	Price in Current Year	Aggregate Expenditure in Base year	Aggregate Expenditure in Current year
	q_0		P_0	P_1	P_0q_0	P_1q_0
Rice	5 quintals	Quintal	24	30	120	150
Wheat	1 quintal	Quintal	16	20	16	20
Pulses	2 quintals	Quintal	12	18	24	36
Ghee	4 kg.	1 kg.	5	6.25	20	25
Oil	20 gk.	per 40 kg.	40	50	20	25
Clothing	40 metres	Meter	1	1.50	40	60
Firework	10 quintals	Quintal	2	2.50	20	25
House Rent	1	House	20	25	20	25
					280	366

Index Number for the current year $= \frac{p_1q_0}{p_0q_0} \times 100$

$$= \frac{366}{280} \times 100$$

$$= 130.7$$

Family Budge Method

Article	Quantity Consumed in Base Year	Unit	Price in Base Year	Price in Current Year	Price Relatives for Current Year	Weights (Values Consumed in base year)	Product of Price Relative and values
			P_0	P_1	P	V	PV
Rice	5 quintals	Quintal	24	30	125	120	15,000
Wheat	1 quintal	Quintal	16	20	125	16	2,000
Pulses	2 quintal	Quintal	12	18	150	24	3,600
Ghee	4 kg.	kg.	5	6.25	125	20	2,500
Oil	20 kg.	40 kg.	40	50	125	20	2,500
Clothing	40 meters	Meter	1	1.50	150	40	6,000
Firewood	10 quintals	Quintal	2	2.50	125	20	2,500
House Rent	1	House	20	25	125	20	2,500
						280	**36,600**

$$\text{Index Number for the current year} = \frac{PV}{V} = \frac{36600}{280} = 130.7$$

कुल मूल्य निर्देशांक
(Value Index Numbers)

यद्यपि इन निर्देशांकें का व्यापक प्रकाशन नहीं होता, परन्तु फिर भी प्रबन्धकीय तथा व्यापारिक दष्टि से इनका व्यापक प्रयोग होता है। बहुत-सी संस्थाएँ अपनी बिक्री की जानकारी प्रदान करने के लिए निम्न प्रकार से सूचनाएँ प्रस्तुत करती हैं :-

वर्ष (Year)	:	1190	1991	1992	1993	1994
बिक्री (Sales-Lakh Rs.)	:	4.00	4.20	4.56	5.24	6.00
प्रवत्ति प्रतिशत (Trend-%)	:	100	105	114	131	150

उपर्युक्त प्रदर्शित प्रवत्ति प्रतिशत वास्तव में कुल मूल्य का निर्देशांक है। प्रत्येक वर्ष की मात्रा तथा वस्तु के मूल्य के गुणा के योग के बराबर होता है। प्रत्येक वर्ष की बिक्री आधार वर्ष की तुलना में मूल्य निर्देशांक होते हैं। कुल मूल्य निर्देशांक को सूत्र निम्न है :-

$$\text{Value Index} = \frac{p_1q_0}{p_0q_0} \times 100$$

इस स्थिति में, न तो मात्रा ही समान रहती है और न ही मूल्य, अर्थात् दोनों की परिवर्तित होते रहते हैं। निर्देशांकों के परिवर्तन पर मूल्य व मात्रा दोनों का प्रभाव पड़ता है। मूल्य निर्देशांक, मात्रा निर्देशांक तथा कुल मूल्य निर्देशांक में अंतर निम्न तरीके द्वारा स्पष्ट किया जाता है :-

मूल्य निर्देशांक (Price Index)	मात्रा निर्देशांक (Quantity Index)	कुल मूल्य निर्देशांक (Value Index)
Index No. = $\frac{p_1q_0}{p_0q_0} \times 100$	Index No. = $\frac{p_0q_1}{p_0q_0} \times 100$	Index No. = $\frac{p_1q_1}{p_0q_0} \times 100$
मूल्य अस्थिर, मात्रा स्थिर (Price vary, Quantity Constant)	मूल्य स्थिर, मात्रा अस्थिर (Price Constant, Quantity Vary)	मूल्य एवं मात्रा दोनों अस्थिर (Prices and Quantity Both Vary)

Illustration 18 :

Commodity	1985		1988	
	Price	Quantity	Price	Quantity
A	8	10	10	11
B	10	9	12	9
C	16	16	20	13
D	18	15	22	17

Construct the value index for 1988 with 1985 as base.

Solution : Construction of value index number.

Commodity	1985		1988			
	Price p_0	Quantity q_0	Price p_1	Quantity q_0	p_1q_1	p_0q_0
A	8	10	10	11	110	80
B	10	9	12	9	108	90
C	16	16	20	13	260	256
D	18	15	22	17	374	270
					852	**696**

$$\text{Value Index for 1998, } V_{01} = \frac{p_1q_1}{p_0q_0} \times 100 = \frac{852}{696} \times 100 = 122.41$$

सारांश

(Summary)

- निर्देशांक एक विशेष प्रकार के माध्य हैं जिनसे हम किसी भी तथ्य के परिणाम में होने वाले परिवर्तनों का समय अथवा स्थान के आधार पर मापन करते हैं।
- यदि हम समय के आधार पर निर्देशांक बनाते हैं तो किसी एक वर्ष को आधार बनायेंगे। आधार वर्ष में किसी भी तथ्य के परिणाम को 100 के बराबर लेते हैं। उसकी तुलना में अगले किसी भी वर्ष उस परिणाम का जो भी मूल्य होगा वह उस वर्ष के लिए निर्देशांक होता है।
- निर्देशांक मूख्य रूप से तीन तरह के होते हैं - मूल्य निर्देशांक, परिणाम निर्देशांक व कुल मूल्य निर्देशांक।
- निर्देशांक सरकारी व औद्योगिक नीतियों का निर्माण करने में महत्त्वपूर्ण भूमिका अदा करते हैं।
- निर्देशांक निकालने की अनेक विधियाँ हैं जैसे सरल तथा भारित समूही विधि तथा मूल्यानुपातों का माध्य तथा भारित माध्य।
- सरकारी स्तर पर उपभोक्ता मूल्य निर्देशांक अथवा जीवन-निर्वाह-व्यय संबंधी निर्देशांक नियमित तौर पर निकाले जो हैं जिसके आधार पर सरकार महंगाई भत्ते का माप तय करती है।
- इसके अलावा वास्तविक आय निर्देशांक औद्योगिक उत्पादन निर्देशांक, स्टॉक मार्केट का निर्देशांक (BSE Senstivity Index) खाद्यान्न उत्पादन निर्देशांक आदि नियमित तौर पर निकाले जाते हैं।

प्रश्नावली

(Exercise)

(1) ''निर्देशांक का प्रयोग किसी तथ्य के उन परिवर्तनों को नापने के लिए किया जाता है जिनका अवलोकन हम प्रत्यक्ष रूप से नहीं कर सकते हैं।'' उपर्युक्त कथन को समझाइए तथा निर्देशांकों के प्रयोग तथा सीमाओं का इंगित कीजिए।

"Index Number are used to measure the changes in some quantity which we cannot observe directly".

Explain the above statement and point out the uses and limitation of index numbers.

(2) ''निर्देशांक व्यापार के पथ पर संकेतक तथा निर्देश-स्तम्भ हैं जो व्यापारी को अपने कार्य का संचालन करने की विधि बताते है।''

उपर्युक्त कथन को स्पष्ट करिए एवं निर्देशांकों में प्रयुक्त होने वाले माध्यों की उपयोगिता का तुलनात्मक अध्ययन कीजिए। आप किस माध्य को प्राथमिकता देंगे और क्यों ?

"Index Numbers are the signs and guide posts along the business highway that indicate to the businessman how he should drive or manage his affairs". Explain the above statement and also point out the relative advantages of the various averages as applied to index numbers. Which would you prefer and why ?

(3) ''निर्देशंक सम्बन्धित चर-मूल्यों के परिणाम में होने वाले अन्तरों की माप करने का साधन है।'' इस कथन का विवेचन कीजिए तथा निर्देशांक के प्रमुख उपयोगों का वर्णन कीजिए।

"Index Numbers are devices for measuring differences in the magnitude of a group of related variables". Discuss this statement and point out the important uses of Index Numbers.

(4) निर्देशांकों की उपयोगिता समझाइए। सामान्य तथा जीवन-निर्वाह निर्देशांकों की रचना में अपनायी जाने वाली विधि का वर्णन कीजिए।

Explain the uses of index numbers. Describe the procedure followed in the preparation of general and cost of living index numbers.

(5) निर्देशांकों की रचना में

(*i*) आधार-वर्ष का चुनाव, तथा

(*i*) भारों का चुनाव की समस्याओं का विवेचन कीजिए।

Discuss the problems of

(*i*) Selection of the base, and

(*ii*) Selection of weights in the construction of index numbers.

(6) निर्देशांक किस प्रकार बनाये जाते हैं ? जीवन निर्वाह निर्देशांक बनाने में आधार वर्ष और भार के चुनावों के महत्व का वर्णन कीजिए।

How are index numbers constructed ? Discuss the importance of the choice of base year and the selection of weighting in the construction of a cost of living index number.

(7) ''निर्देशांक-निर्माता के लिए वास्तविक समस्या यह है कि यह भारांकन को अवसर पर छोड़ देगा अथवा उसे तर्क संगत बनाने का यत्न करेगा।'' (मिट्चैल) अवसर-भारांकन तथा तर्क-संगत भारांकन में भेद स्पष्ट कीजिए तथा उक्त समस्या का हल सुझाइए। क्या फिशर का आदर्श सूत्र तर्क-संगत भारांकन प्रस्तुत करता है, विवेचन कीजिए।

The real problem for the maker of index numbers is whether he shall leave weighting to chance or seek to rationalist it." (Mitchell) Distinguish clearly between chance weighting and rational weighting and suggest a solution of the above problem. Also discuss whether Fisher's ideal formula offers a rational system of weighting.

(8) ''निर्देशांक का अर्थ समझाइए। एक श्रेष्ठ निर्देशांक से किन परीक्षणों की पूर्ति की आशा की जाती है। निर्देशांक में अभिनति क्या होती है ? मूल्य-निर्देशांकों की रचना में किन समस्याओं का समावेश होता है ? ऐसे निर्देशांक में आधार वर्ष के महत्व को स्पष्ट रूप से समझाइए।

Explain "Index Number". What tests is a good index number expected to satisfy ? What is bias in an index number ? What problems are involved in the construction of an index number of prices ? Point out clearly the importance of base year in such as index.

(9) निर्देशांक की परिभाषा दीजिए। निर्देशांक बनाने की स्थिर श्रंखला आधार विधियों की तुलना कीजिए और उनके गुण - दोषों का विवेचन कीजिए।

Define an "Index Number". Distinguish between the Fixed Base and Chain Base Methods of constructing Index numbers and discuss their relative merits.

(10) ''निर्देशांक आर्थिक बैरोमीटर है।'' इस कथन की व्याख्या कीजिए तथा यह भी बताइए कि किसी प्रकाशित देशनांक का प्रयोग करते समय आप किस प्रकार की सावधानी बरतेंगे।

"Index Numbers are economic barometers." Explain this statement, and mention the precautions that should be taken in making use of any published index numbers.

(11) स्पष्ट रूप में समझाइए :-

(*a*) समय उत्क्रम्यता परीक्षण तथा

(*b*) तत्त्व उत्क्रम्यता परीक्षण

प्रदर्शित कीजिए कि फिशर का आदर्श सूत्र इन दोनों परीक्षणों का पूरा करता है। चूँकि निर्देशांक रचना के अन्य सूत्रों द्वारा इन परीक्षणों को पूरा नहीं किया जाता है, अतः उन्हें अस्वीकार किया जाना चाहिए।

Explain Clearly :-

(*a*) Time Reversa Test, and

(*b*) Factor Reversal Test

Demonstrate how Fisher's Ideal Index Formula satisfies both these tests. Should other formulae for computing index number be rejected because they do not satisfy these tests.

(12) निम्न समंकों से 1995 पर आधारित श्रंखला सूचकांक ज्ञात कीजिए।

वर्ष (Year)	:	1995	1996	1997	1998	1999	2000	2001
उत्पादन	:	20	22	25	28	30	36	40

('000 क्विंटल में)

[Production (In '000 Quintals)

(13) निम्नलिखित आंकड़ों से 1993 को आधार मानकर निर्देशांकों को ज्ञात कीजिए :-

From the following data, construct Index Number by taking 1993 as base year :

वर्ष (Year)	:	1993	1994	1995	1996	1997	1998	1999	2000
मूल्य (रु.)	:	67	90	91	92	87	88	89	90
Price (Rs.)	:	134.3	135.8	137.3	129.8	131.3	132.2	133.3	134.4

(14) निम्न समंकों से 1980 को आधार वर्ष मानकर 1984 के लिए सामूहिक विधि द्वारा सूचकांक की रचना कीजिए।

From the following data construct an Index Number or 1984 taking 1980 as base year by using aggregate method.

Commodity	:	A	B	C	D	E	F
Price in (1980)	:	110	125	415	180	100	70
Prices in (1984)	:	115	140	505	198	95	83

(15) निम्न वस्तुओं के औसत मूल्यों से 1972 से श्रंखलाबद्ध श्रंखला सूचकांक ज्ञात कीजिए :

Calculate the chain base Index Numbers chained to 1972 from the average price of following commodities:

Commodities (वस्तुएँ)	1972	1973	1974	1975	1976
Wheat (गेहूँ)	4	6	8	10	12
Rice (चावल)	16	20	24	30	36
Sugar (चीनी)	8	10	16	20	24

[**Ans.** 100, 133.33, 183.7, 229,63, 275.56]

(16) निम्न समंकों से दिए गए सूत्रों का प्रयोग करके सूचकांक ज्ञात कीजिए :-

(अ) लेसपियर सूत्र

(ब) पासचे सूत्र

(स) फिशर का आदर्श सूत्र

(ग) मार्शल एजवर्थ का सूत्र

Compute Index numbers from the following data using :-

(*a*) Laspeyre's formula,

(*b*) Paasche's formula,

(*c*) Fisher's Ideal Index Formula

(*d*) Marshall Edgeworth Formula

Commodities (वस्तुएँ)	**Base Year (आधार वर्ष)**		**Current Year (वर्तमान वर्ष)**	
	Quantity (मात्रा)	**Price** (कीमत)	**Quantity** (मात्रा)	**Price** (कीमत)
A	5	100	6	150
B	4	80	5	100
C	2.5	60	5	72
D	12	30	9	33

[**Ans.** 129.8, 131.02, 130.45, 130.41]

(17) कर्मचारियों की मासिक मजदूरी 1991 से 1997 तक निम्न तालिका में दी गयी है। वास्तविक मजदूरी का सूचकांक ज्ञात कीजिए :-

From the following table showing montly wages of workers from 1991 to 1997 construct the real wages index numbers :-

वर्ष (Year)	:	1991	1992	1993	1994	1995	1996	1997
मजदूरी (Wage)	:	800	819	825	876	920	938	924
कीमत निर्देशांक (Price Index No.)	:	100	105	110	120	125	140	140

[**Ans.** 100, 97.5, 93.5, 91.25, 92, 83.75, 82.5]

(18) एक औसत भारतीय मजूदर के पारिवारिक बजट से सम्बन्धित समूह निर्देशांक तथा उनके भार निम्न सारणी में दर्शाये गये हैं। भारित निर्देशांक निकालिए :-

The following table gives group index number and their weights relating to family budget of an average Indian labour. Prepare weighted index number :-

वर्ग (Groups)	**निर्देशांक (Index No.)**	**भार (Weight)**
खाद्य पदार्थ (Food)	350	48
ईंधन एवं रोशनी (Fuel and Lighting)	220	10
कपड़ा (Clothing)	230	8
किराया (Rent)	160	12
विविध (Miscellaneous)	190	22

[**Ans.** 269.4]

(19) नीचे दिये हुए आंकड़ों से फिशर आदर्श निर्देशांक परिकलित कीजिए तथा समय उत्क्राम्यता और तत्त्व उत्क्राम्यता परीक्षणों का सत्यापन कीजिए :-

Calculate Fisher's ideal index number from the data given below and verify Time Reversal and Factor Reversal Tests :-

वस्तु (Commodity)	**आधार वर्ष (Base Year)**		**चालू वर्ष (Current Year)**	
	मूल्य (Value)	**मात्रा (Quantity)**	**मूल्य (Value)**	**मात्रा (Quantity)**
A	300	50	495	9
B	200	100	375	3
C	240	60	390	6
D	500	50	350	14

[**Ans.** $P_{01} = 146$, $P_{10} = 68.1$, $Q_{01} = 88.4$, $V_{01} = 129.8$]

(20) निम्न आंकड़ों से 2001 वर्ष के लिए निम्न विधियों को अपनाते हुए जीवन-निर्वाह-व्यय निर्देशांक ज्ञात कीजिए :-

(*i*) समूह व्यय विधि

(*ii*) पारिवारिक बजट विधि

Construct cost of Living Index Number from the following data for 2001 using :-

(*i*) Aggregative Expenditure Method

(*ii*) Family Budget Method.

Article	**Quantity consumed**	**Unit**	**Price in Rs. 2001**	**Price in Rs. 2001**
Wheat	12 mds.	per md.	18	27
Pulses	1 md.	” ”	25	30
Ghee	0.5 md.	” ”	150	200
Fuel	15 md.	” ”	3	3.50
House	—	—	15	20
Misc.	—	—	0.50	0.75

[**Ans.** Index No. 140]

(21) निम्न स्थिर आधार-सूचकांकों को श्रंखला आधार सूचकांकों में परिवर्तित कीजिए : -

Convert into Chain Base Index No. from Fixed Base Index Number.

Years	:	1980	1981	1982	1983	1984
Fixed Base Indices	:	100	98	102	140	190

[**Ans.** 100, 98, 104.08, 137,25, 135.71]

(22) नीचे दिए गए श्रंखला आधार सूचकांकों से स्थिर आधार सूचकांक बनाइए : -

From the chain base index number given below, prepare fixed base index numbers :-

Years	**:**	1992	1993	1994	1995	1996	1997
Index	**:**	92	102	104	98	103	101

[**Ans.** 92, 03.84, 97.59, 95.64, 98.51, 99.50]

(23) निम्नलिखित मूल्य सूंचकांक दिए हुए हैं (आधार वर्ष = 1980) आधार वर्ष 1980 से 1990 से परिवर्तित करो।

The following are the index number of price (base 1980)

Year	1980	1981	1982	1983	1984	1985	1986	1987
Index	100	107	108	95	115	130	125	165
Year	1988	1989	1990	1991	1992	1993	1994	1995
Index	199	220	240	225	265	268	258	290

Shift the base from 1980 to 1990

[**Ans.** 41.6, 44.5, 45, 39.5, 47.9, 54.1, 52.08, 68.7, 82.9, 91.6, 100, 93.7, 110.4, 111.6, 107.5, 120.8]

(24) निम्न आंकड़ों से जीवन-निर्वाह-व्यय निर्देशांक ज्ञात कीजिए :-

Construct the cost of Living Index Number from the data given below :-

समूह (Group)	निदेशांक (Index)	व्यय (Expenditure)
Food	550	46%
Clothing	215	10%
Fuel and Lighting	220	7%
House Rent	150	12%
Miscellaneous	275	25%

[**Ans.** Cost of Living Index No. = 376.65]

(25) निम्न आंकड़ो से जीवन-निर्वाह-व्यय निर्देशांक ज्ञात कीजिए :-

Construct the cost of Living Index from the following data :-

Group	Weights	Group Indices
Food	47	247
Fuel and Lighting	7	293
Clothing	8	289
House Rent	13	100
Miscellaneous	14	236

[**Ans.** Index No. = 231.19]

❑❑

अध्याय - 12

काल-श्रेणियों का विश्लेषण (Analysis of Time Series)

अर्थ-व्यवस्था गतिशील (Dynamic) होती है, स्थिर नहीं, तथा गतिशीलता का सम्बन्ध समय-घटक (Time Factor) से होता है। आर्थिक विश्लेषण क्रमबद्ध गति (Chronological Movement) एक आधारभूत विचार होता है। उत्पादन, राष्ट्रीय आय, बिक्री आदि अनेक आर्थिक तथ्यों का हम गतिमय चित्र लेना चाहते हैं, स्थिर चित्र नहीं। समय के आधार पर अवलोकनों (Observations) की श्रेणी 'काल श्रेणी' (Time Series) कहलाती है। संख्यात्मक तथ्य जिनका संकलन समयान्तर से किया जाता है काल श्रेणी का रूप ले लेते हैं। 'एक काल श्रेणी को किसी आर्थिक चर अथवा मिश्रित चरों, जिनका सम्बन्ध विभिन्न समयाविधियों से होता है, के अवलोकनों के संकलन के रूप में परिभाषित किया जा सकता है।'' **प्रो. कैने तथा कीपिंग** के अनुसार, समय पर आधारित समंक-समूह काल-श्रेणी कहलाते हैं।'' **वर्नर हिर्श** के शब्दों में, ''समय के क्रमिक बिन्दुओं के तत्सवादी उसी चर के मूल्यों का व्यवस्थित अनुक्रम ही काल श्रेणी कहलाता है।'' **सिसिल सी. मेयर्स** लिखते हैं, ''काल श्रेणी को एक चर के समयावधिनुसार पुनवर्तित मापों के क्रम के रूप में परिभाषित किया जा सकता है।'' **पी. जी. मूर** के अनुसार, मूल्यों की एक समयावधि की श्रेणी, एक काल-श्रेणी कहलाती है।'' अतः समय के क्रमागत बिन्दुओं के अनुसार किसी चर के क्रमबद्ध मूल्यों को काल-श्रेणी कहते हैं।

आर्थिक एवं व्यापारिक क्षेत्रों में समंकों के व्यावहारिक उपयोग की दिशा में एक संख्या-शास्त्री के लिए काल-श्रेणी के विश्लेषण का विशेष महत्त्व है। इसका कारण यह है, कि ''काल-श्रेणी का विश्लेषण करने का एक मुख्य उद्देश्य भावी घटनाओं की गतिविधि का यथार्थ अनुमान लगाने के लिये आर्थिक तथ्यों में होने वाले परिवर्तनों को समझना, समझाना एवं मूल्यांकित करना है।''

निम्न आँकड़ों से काल श्रेणी का अर्थ स्पष्ट होता है :-

Time	Wheat Production of Haryana in ('000 tons)
1990	12
1991	15
1992	18
1993	16
1994	22
1995	25
1996	23

काल श्रेणी विश्लेषण का महत्त्व

(Utility of Analysis of Time Series)

काल श्रेणी का विश्लेषण केवल अर्थशास्त्रियों व व्यावसायिक संस्थाओं के लिए ही उपयोगी नहीं अपितु सरकार, वैज्ञानिक शोधकर्ताओं, भूगोल शास्त्रियों आदि के लिए भी निम्न कारणों से बहुत उपयोगी है :-

(1) **पिछले अनुभव से लाभ (Advantages from Previous Experience) :** भूतकाल में हुए परिवर्तनों की दशा का अध्ययन करने से अनेकों महत्वपूग्र निष्कर्ष निकाले जा सकते हैं और भूतकाल में हुए परिवर्तनों के कारणों को विश्लेषण द्वारा जाना जा सकता है।

(2) **भविष्य के बारे में अनुमान (Prediction for Future) :** इन श्रेणियों के अध्यन से भविष्य में आँकड़ों में होने वाले परिवर्तनों के बारे में अनुमान लगाया जा सकता है। उदाहरणार्थ — भविष्य की बिक्री के आधार पर उत्पादन नीति में उपयोगी परिवर्तन किए जा सकते हैं व सम्भावित हानियों से बचा जा सकता है।

(3) **वास्तविक आँकड़ों का मूल्यांकन (Evaluation of Actual Data) :** वास्तविक आँकड़ों का काल श्रेणी विश्लेषण से प्राप्त अनुमानित आँकड़ों से तुलनात्मक अध्ययन कर सकते हैं और वास्तविक व अनुमानिक आँकड़ों के अन्तर के विभिन्न कारणों को जाना जा सकता है।

(4) **व्यापार चक्र का अनुमान (Prediction of Trade Cycle) :** व्यापारिक चक्रीय उच्चावनों (Cyclical Fluctuations) की सहायता से व्यापार में होने वाले परिवर्तन जैसे — समद्धि (Boom), अवनति (Recession), अवसाद (Depression), व पुनरुत्थान (Recovery) का पता चलता है। व्यवसायी इस जानकारी से अपनी क्रियाओं को नियन्त्रित कर सकता है, जिससे सम्भावित हानि से बचा जा सके।

(5) **विभिन्न काल श्रेणियों का तुलनात्मक अध्ययन (Comparison of Various Series) :** विभिन्न काल मालाओं के पारस्परिक अध्ययन से अनेकों निष्कर्ष निकाले जा सकते हैं जैसे — विभिन्न देशों की मत्यु -दर, जन्म-दर, कृषि उपज दर प्रति हेक्टेयर, शिक्षा दर आदि को जानकर विभिन्न देशों के बारे में जाना जा सकता है।

(6) **सार्वभौमिक उपयोगिता (Universal Utility) :** इन श्रेणियों की उपयोगिता केवल व्यवसायी व अर्थशास्त्रियों के लिए ही नहीं बल्कि प्रायः सभी क्षेत्रों में है, जैसे — सरकार, कृषक, राजनैतिक, सामाजिक संस्थाएँ, शोधकर्ताओं, वैज्ञानिकों के लिए आदि।

काल श्रेणी के संघटक
(Components of Time Series)

काल श्रेणी पर कालान्तर में अनेक कारणों का प्रभाव पड़ता है, जिनके कारण काल श्रेणी के चल-मूल्य समय परिवर्तन पर बदलते रहते हैं। इन कारणों को काल श्रेणी के संघटक भी कहते हैं। मुख्यतः इन कारणों को चार प्रमुख भागों में बाँटा जा सकता है, जिनका विवरण पहले भी दिया जा चुका है। ये चार घटक इस प्रकार हैं :-

(1) दीर्घकालीन प्रवत्ति (Secular Trend or Trend)

(2) अल्पकालीन उच्चावचन (Short time Oscillations)

  (*i*) आर्तव विचरण (Seasonal Variation)

  (*ii*) चक्रीय उच्चावचन (Cyclical Fluctuations)

(3) अनियमित या दैव उच्चावचन (Irregular or Random Fluctuations)

(1) **दीर्घकालीन प्रवत्ति (Secular Trend or Trend)** काल श्रेणी की दीर्घकाल में बढ़ने, घटने अथवा स्थिर रहने की प्रवत्ति को ही आँकड़ों की दीर्घकालीन प्रवत्ति कहते हैं। कालान्तर में आँकड़ों में विभिन्न उतार -चढ़ाव होते रहते हैं, किन्तु दीर्घकाल में यह परिवर्तन प्रायः एक निश्चित दिशा में लेते हैं। जैसे - जनसंख्या में वद्धि का क्रम पाया जाता है। इसी प्रकार मूल्य वद्धि , व्यापारियों की बिक्री व उत्पादन में भी पिछले वर्षों मे बढ़ने की ही प्रवत्ति पायी जाती है। प्रायः चलों में वद्धि की ही प्रवत्ति होती है किन्तु कुछ चल पारिवारिक आय में वद्धि होने पर भी स्थिर प्रवत्ति को ही होते हैं, जैसे — परिवार में नमक का प्रयोग। इसी प्रकार कुछ चलों में घटने की प्रवत्ति भी पायी जाती है , जैसे — मत्यु दर पिछले दो दशकों में घटी है। इसी प्रकार पिछले दशकों में बैलों द्वारा कृषि की प्रवत्ति घटी है।

दीर्घकाल में आँकड़ों की प्रवत्ति में परिवर्तन होने के अनेकों कारण हैं , जिनमें जनसंख्या का लगातार बढ़ना, उत्पादन में तकनीक सुधार, प्राकृतिक साधनों का अधिकतम उपयोग, व्यावसायिक संगठन में सुधार, स्वचालित मशीनों का उत्पादन करने में प्रयोग, उत्पादन प्राथमिकताओं में सरकारी हस्तक्षेप प्रमुख हैं। दीर्घकाल में बढ़ने की प्रवत्ति से यह अभिप्राय कदापि नहीं है कि आँकड़े सदैव ही बढ़ेंगे बल्कि आँकड़े किसी वर्ष बढ़ेगे व किसी वर्ष घटेंगे, किन्तु यह घटने-बढ़ने की प्रवत्ति सामान्य रूप से बढ़ने की ओर अग्रसर होगी।

दीर्घकालीन प्रवत्ति या उपनति का अध्ययन दो प्रकार से किया जा सकता है :

(1) रेखीय प्रवत्ति (Linear Trend)

(2) अरेखीय प्रवत्ति (Non-linear Trend)

जब दीर्घकाल में काल श्रेणी मूल्यों के बढ़ने या घटने की दर निश्चित होती है तो इसे रेखीय प्रवत्ति (Linear Trend or Straight Trend) कहते हैं। इसे निम्न समीकरण द्वारा व्यक्त करते हैं :-

$$Y_t = a + bX$$

where, Yt = trend value of a variable

'a' and 'b' are constants

'a' is the intercept of Y and

'b' is the slope of the line

जिन परिस्थितियों में काल श्रेणी के मूल्यों के बढ़ने अथवा घटने की दर असमान होती है, ऐसी दशा को अरेखीय प्रवत्ति (Non-linear Trend) कहते हैं। इसे निम्न समीकरण द्वारा व्यक्त करते हैं :-

$$Y_t = a + bX + cX^2$$

दीर्घकालीन प्रवत्ति अध्ययन करने से समंकों की भूतकालीन प्रवत्ति का पता चलता है। इसके द्वारा भविष्य के समंकों का अनुमान लगाया जा सकता है। हम विभिन्न स्थानों की काल श्रेणियों का तुलनात्मक अध्ययन कर सकते हैं। इसी प्रकार हम विभिन्न समय की दो या इससे अधिक काल श्रेणियों का भी तुलनात्मक अध्ययन कर सकते हैं। काल श्रेणी से यदि दीर्घकालीन उपनति को हटा दिया जाय तो हम लघु कालीन परिवर्तनों का भी अध्ययन कर सकते हैं।

(2) अल्पकालीन परिवर्तन (Short Term Fluctuations) : अल्पकालीन परिवर्तन नियमित अथवा अनियमित हो सकते हैं।

नियमित अथवा सामयिक परिवर्तन (Regular or Periodic Movements) : अनेक आर्थिक तथ्यों में प्राकृतिक अथवा रीति-रिवाजों से सम्बन्ध के कारण नियमित परिवर्तन होते रहते हैं। सामयिक परिवर्तन में निश्चित समयावधि के पश्चात परिवर्तनों की पुनरावत्ति होती है। सामयिक परिवर्तन दो प्रकार के होते हैं : -

(*i*) आर्तव परिवर्तन (Seasonal Changes) : इस प्रकार के परिवतनों का सम्बन्ध मौसम में परिवर्तन से होता है। अनेक आर्थिक काल श्रेणियों में इस प्रकार के परिवर्तन पाये जाते हैं। साप्ताहिक, मासिक अथवा त्रैमासिक समयसान्तर से संकलित समंकों में आर्तव परिवर्तन स्पष्ट हो जाते हैं। आर्तव-परिवर्तनों की मात्रा में भले ही परिवर्तन हो जाये, परन्तु परिवर्तनों का समय निश्चित होता है। वार्षिक समंक मालाओं में आर्तव परिवर्तन दष्टिगोचर नहीं होते हैं।

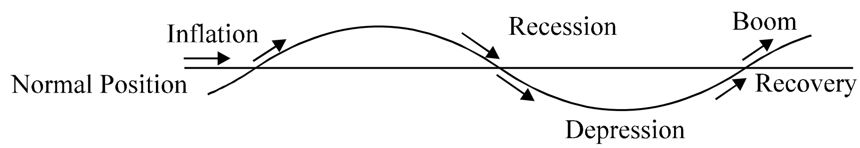
ऐसे परिवर्तन 12 महीनों के अन्दर होते हैं और यह क्रम निश्चित रूप से चलता है। प्रायः प्रत्येक प्रकार के व्यवसायों पर इन परिवर्तनों का असर कम व अधिक होता है और यह असर समान्य है। यहाँ मौसमी परिवर्तनों का अभिप्राय शाब्दिक रूप में केवल मौसम तबदीली के कारण हुए परिवर्तनों से ही नहीं है, बल्कि और सभी प्रकार के परिवर्तन जो प्रतिवर्ष एक विशेष समय पर होते हैं और उनकी पुनरावत्ति प्रति वर्ष होती है। अतः मौसमी परिवर्तन — सर्दी, गर्मी, वर्षा, वसन्त आदि के होने पर हमारे खान-पान व पहनावे में परिवर्तन होता है। उदाहरणार्थ — गर्मियों में मनुष्य शीतल पेय अधिक पसन्द करते हैं किन्तु सर्दियों में गर्म पेय जैसे — चाय, काफी आदि की माँग अधिक होती है। इसी प्रकार गर्मियों में पहिनने वाले व सर्दियों में पहनने वाले कपड़ों में अन्तर होता है अतः कपड़ों की माँग मौसम के परिवर्तन से बदलती रहती है।

वस्तुओं की माँग में परिवर्तन हमारे त्योहार, रीतिरिवाज, स्वभाव, आदतों, फैशन परम्परा आदि के कारण भी होता है। उदाहरणार्थ — त्योहारों के अवसर पर मिठाई की माँग बढ़ जाती है। विवाह के अवसरों पर कपड़ो, फर्नीचर, जेवरों आदि की माँग अधिक होती है। इसी प्रकार हम अपनी मासिक आय का अधिकतम खर्चा मास के प्रथम सप्ताह में ही करते हैं अतः दुकानदारों की बिक्री मास के प्रथम सप्ताह में ही अधिक होती है।

उपरोक्त सभी प्रकार के परिवर्तन प्रतिवर्ष नियमित रूप से होते हैं अतः इनका पूर्वानुमान करना आसान है उदाहरणार्थ — वर्षा ऋतु आने से पूर्व दुकानदार छतरी की बिक्री का पूर्वानुमान लेते हैं। इसी प्रकार राखी का त्यौहार आने से पूर्व की दुकानदार राखियों की माँग का पूर्वानुमान लगा लेते हैं। यह अनुमान पिछले वर्षों के आँकड़ों के अध्ययन व विश्लेषण से लगाया जा सकता है।

काल श्रेणी के विश्लेषण में मौसमी परिवर्तन के माप का महत्त्वपूर्ण स्थान है। पूर्व नियोजित योजना बनाते समय व्यवसायी इन परिवर्तनों को अधिक महत्व देते हैं, जिससे कि उत्पादन रहतिया, विज्ञापन आदि की लागत को नियन्त्रित रखा जा सके। मौसमी माँग की पूर्ति के लिए व्यवसायी लघु कालीन ऋण की व्यवस्था पहले से ही कर सकते हैं, जैसे — कपड़े का व्यापारी सर्दी शुरु होने से पहले ही ऊनी वस्त्रों की व्यवस्था करने के लिए ऋण की व्यवस्था करता है।

(*ii*) **चक्रीय परिवर्तन (Cyclical Changes) :** आर्थिक क्रियाओं के स्तर में क्रमानुसार उतार-चढ़ाव को अर्थशास्त्री द्वारा ''व्यापारिक चक्र'' की संज्ञा दी जाती है। आर्थिक क्रियाओं के उतार-चढ़ाव की चार क्रमानुसार अवस्थाएँ होती हैं — अभिवद्धि (Inflation), ह्रास (Recession), अवसाद (Depression) तथा पुनरुत्थान (Recovery)। प्रत्येक व्यापारिक चक्र में ये अवस्थाएँ पायी जाती हैं। उतार-चढ़ाव की मात्रा तथा व्यापारिक चक्र की अवधि में भले ही नियमितता न हो परन्तु आर्थिक क्रियाओं से सम्बन्धित श्रेणियों में आठ या नौ वर्षों के व्यापारिक चक्र सामान्यतः पाये जाते हैं। लगभग 3 वर्ष की अवधि के लघु व्यापारिक चक्र भी पाये जा सकते हैं।



उपरोक्त चित्र में विभिन्न स्थितियों को दर्शाया गया है। प्रथम अवस्था में व्यावसायिक क्रियाएँ जैसे — उत्पादन, मूल्य, वद्धि , रोजगार आदि समद्धि की ओर अग्रसित होते हैं अतः समद्धि चिन्ह पाए जाते हैं। किन्तु द्वितीय अवस्था में उपरोक्त व्यावसायिक क्रियाओं में अवनति पाई जाती है। जब सरकार मुद्रा स्फीति को नियन्त्रित करने की कोशिश करती है तो मूल्यों में कमी आने लगती हैं। किन्तु ऐसा करने पर उत्पादक उत्पादन कम कर देते हैं। अतः रोजगार के अवसर प्रभावित होते हैं। इन सभी कारणों से अवनति (Disinflation or Recession) की स्थिति उत्पन्न होती है। यदि यह नियन्त्रण सामान्य स्थिति (Normal Position) से नीचे चला जाता है तो अवसाद (Depression) की स्थिति उत्पन्न हो जाती है। ऐसी स्थिति से निबटने के लिए सरकार बैंक दर में कमी करती है व अन्य कदम उठाती है, सामान्य स्थिति में न आकर पुनः सामान्य स्थिति से बढ़ जाते हैं तो पुनः तेजी आ जाती है अर्थात् समद्धि की स्थिति उत्पन्न होने लगती है। यह व्यापार चक्र निरन्तर इसी क्रम से चलता रहता है। सामान्यतः यह चक्र 3 वर्ष से 8 वर्षों में पूरा हो जाता है किन्तु यह अनिवार्य नहीं कि व्यापार चक्र की चारों स्थितियाँ बराबर समय लें। यह अवधि सरकार की नियन्त्रण नीति पर निर्भर करती है।

व्यापारिक प्रक्रिया को स्थिर बनाए रखने के लिए नीति निर्धारण में चक्रीय उच्चावचन को समझना अनिवार्य है ताकि समद्ध व अवसाद की परिस्थितियों से दूर रहा जा सके क्योंकि देश की आर्थिक प्रगति के लिए दोनों ही स्थितियाँ प्रतिकूल हैं, विशेषकर अवसाद स्थिति। चक्रीय उच्चावचन की इतनी महत्ता होते हुए भी इसको मापना बहुत कठिन है क्योंकि व्यापार चक्र के विभिन्न तत्त्वों में परिवर्तन की न तो अवधि और न ही मात्रा समान होती है। प्रायः चक्रीय उच्चावचन अनियमित परिवर्तनों (Irregular Variations) के साथ-साथ होता है, अतः दोनों प्रकार के परिवर्तनों को अलग-अलग मापना असम्भव भी हो जाता है।

(3) अनियमित परिवर्तन (Irregular Variations) अनियमित अथवा भ्रमात्क (Erratic) अथवा दैव (Random) परिवर्तन दो प्रकार के होते हैं — पूर्णतः दैव परिवर्तन, इनसे श्रेणी में कभी एक ओर तो कभी दूसरी ओर परितर्वन केवल अवसर (Chance) के कारण ही होते हैं। दूसरे, ऐसे परिवर्तन जो समय-समय पर विशेष परिस्थितियों के कारण होते हैं तथा जिन्हें पथक किया जा सकता है। ऐसे परिवर्तन हड़ताल , राजनैतिक हलचल, युद्ध आदि घटनाओं के कारण होते हैं अतः इन्हें घटनात्मक परिवर्तन (Episodic Movement) भी कहते हैं।

अल्पकालीन परिवर्तनों के विश्लेषण में प्रवत्ति मूल्यों को मौलिक समंकों से पथक किया जाता है। एक काल श्रेणी के पद मूल्यों में से प्रवत्ति -मूल्यों को घटाकर अल्पकालीन परिवर्तन ज्ञात किये जा सकते हैं।

काल श्रेणी निदर्श

(Time Series Models)

काल श्रेणी के संघटकों के अध्ययन से यह स्पष्ट है कि काल श्रेणी के मूल संमक जो काल श्रेणी में (Y) के नाम से जाने जाते हैं, प्रवत्ति (Trend or T), चक्रीय उच्चावचन (Cyclical Variation or C), मौसमी परिवर्तन (Seasonal Variations or S) तथा अनियमित परिवर्तन (Irregular Variaton or I) द्वारा सामूहिक रूप से प्रभावित होते हैं। इस प्रकार हम कह सकते हैं कि Y का मूल्य इन संघटकों (T, C, S and I) का मिश्रण है।

काल माला के इन संघटकों को अलग-अलग व्यक्त कर उनका गहन अध्ययन ही काल-माला का विश्लेषण (Analysis of Time Series) कहलाता है। अतः किसी एक घटक का अध्ययन करना हो तो सम्पूर्ण समंक से अन्य तीन घटकों को पथक करना होगा। उदाहरण स्वरूप प्रवत्ति (Trend or T) ज्ञात करने के लिए कुल समंक (Y) में से चक्रीय उच्चावचन (C), मौसमी परिवर्तन (S) व अनियमित परिवर्तन (I) को पथक करना होगा।

इस प्रकार हमें काल श्रेणी के संघटकों को विश्लेषण करना होगा। प्रायः विश्लेषण प्रक्रिया में दो रीतियों का प्रयोग किया जाता है।

काल श्रेणी निदर्श (Time Series Model)

योगात्मक निदर्श (Additive Model) | गुणक निदर्श (Multiplicative Model)

(1) योगात्मक निदर्श (Additive Model) : योगात्मक निदर्श के अनुसार काल माला के किसी भी वर्ष का मूल्य उसके चारों संघटकों अर्थात् दीर्घकालीन प्रवत्ति , चक्रीय उच्चावचन, मौसमी परिवर्तन व अनियमित परिवर्तन के योग के बराबर होगा। इसे निम्न सूत्र के रूप में भी व्यक्त किया जा सकता है :-

सूत्र

$$Y = T + S + C + I$$

where, Y = Value of Original Time Series

T = Trend Value

S = Seasonal Variation

C = Cyclical Variation

I = Irregular Variation

योगात्मक निदर्श के अन्तर्गत काल श्रेणी के चारों घटकों को इकाई के रूप में व्यक्त करते हैं। उदाहरणतया इकाई — रुपये, टन, किलो इत्यादि के रूप में व्यक्त की जा सकती है। इस सिद्धान्त के अनुसार किसी भी विशेष घटक का ज्ञात करने के लिए कुल मूल्य में से बाकी तीनो घटकों को घटाना होगा। निम्न समीकरण से इस तथ्य की पुष्टि की जा सकती है :-

$$Y = T + S\,C + I$$

\ $$T = Y - (S + C + I)$$

or $$S = T - (T + C + I)$$

or $$C = Y - (T + S + I)$$

or $$I = Y - (T + S + C)$$

(2) गुणक निदर्श (Mulplicative Mode) : काल श्रेणी विश्लेषण के शास्त्रीय मत के अनुसार काल माला के किसी भी वर्ष का मूल्य उसके चारों संघटकों अर्थात् दीर्घकालीन प्रवत्ति , चक्रीय उच्चावचल, मौसमी परिवर्तन व अनियमित परिवर्तन के गुणक के बराबर होता है। सूत्र के रूप में :-

$$Y = T \times S \times C \times I$$

where, Y = Value of Original Time Series

T = Trend Value

S = Seasonal Variation

C = Cyclical Variation

I = Irregular Variation

उपरोक्त सूत्र के दीर्घकालीन प्रवत्ति (T) को दी हुइ इकाई के रूप में ही व्यक्त करते हैं किन्तु मौसमी परिवर्तन (S) चक्रीय उच्चावचन (C) व अनियमित परिवर्तन (I) को अनुपात (Ratios) में व्यक्त करते हैं।

गुणक निदर्श के अन्तर्गत किसी विशेष घटक को निकालने के लिए कुल मूल्य (Y) को बाकि घटकों से विभाजित करना होगा। इसे निम्न सूत्र से व्यक्त कर सकते हैं :-

$$Y = T \times S \times C \times I$$

$$T = \frac{Y}{S \quad C \quad I}$$

$$S = \frac{Y}{T \quad C \quad I}$$

$$C = \frac{Y}{S \quad C \quad I}$$

As

\ $$S \times C \times I = \frac{Y}{T}$$

प्रायः काल श्रेणी विश्लेषण में गुणक, निदर्श का ही अधिक प्रयोग किया जाता है।

दीर्घकालीन प्रवत्ति अथवा उपनति
(Secular Trend)

दीर्घकालीन प्रवत्ति (Secular Trend) से आश्य काल-श्रेणी के सामान्य दीर्घकालीन परिवर्तनों से है। यह श्रेणी की आधारभूत प्रवत्ति होती है। **वर्नर जेड. हिर्श्च** (Werner Z. Hirsh) के अनुसार, ''प्रवत्ति , जिसे कभी-कभी दीर्घकालीन प्रवत्ति भी कहते हैं , से हमारा आश्य एक श्रेणी में दीर्घकालीन में धीरे-धीरे होने वाली वद्धि अथवा कमी से है जो जनसंख्या वद्धि , तकनीकी ज्ञान एवं उत्पादकता में सुधार, पूँजी उपकरणों की पूर्ति में वद्धि तथा उपभोग की आदतों में परिवर्तन आदि आधारभूत शक्तियों को व्यक्त करती है।'' **सिम्पसन तथा कापका** के शब्दों में, ''प्रवत्ति जिसे उपनति अथवा दीर्घकालीन प्रवत्ति भी कहा जाता है, श्रेणी की एक समयावधि के बढ़ने अथवा घटने की आधारभूत प्रवत्ति होती है। प्रवत्ति के विचार के अन्तर्गत अल्पकालीन परिवर्तन शामिल नहीं होते हैं बल्कि दीर्घकाल में हुए नियमित परिवर्तन शामिल होते हैं।''

कुछ तथ्यों की दीर्घकालीन प्रवत्ति वद्धि की ओर तथा कुछ की ह्रास की ओर होती है। उदाहरणार्थ , अधिकांश देशों में सप्ताह के काम करने के घंटों की प्रवत्ति घटने की ओर तथा जनसंख्या , कृषि तथा उद्योगों के उत्पादन की प्रवत्ति वद्धि की ओर पायी जाती है। सभी दीर्घकालीन परिवर्तन समान गति से नहीं होते हैं। प्रवत्ति से अस्थायी विचरणों के बावजूद , समंक-माला में एक निश्चित दिशा में स्पष्ट रूप से दीर्घकालीन प्रवत्ति दष्टिगोचर होती है।

दीर्घकालीन प्रवत्ति को इंगित करने के लिए (1) एक उद्योग, (2) एक उपक्रम, तथा (3) सम्पूर्ण अर्थ-व्यवस्था के लिए प्रवत्ति ज्ञात की जा सकती है। प्रत्येक की प्रवत्ति ज्ञात करने में अलग -अलग सांख्यिकीय तथा सैद्धान्तिक समस्याएँ उत्पन्न होती हैं।

दीर्घकालीन प्रवत्ति के अनुमान की रीतियाँ

(Methods of Estimating Trend)

काल श्रेणियों के सांख्यिकीय विश्लेषण में उन विधियों का प्रयोग किया जाता है जिनका विकास गणित तथा आर्थिक विश्लेषण के संयोग से हुआ है। दीर्घकालीन प्रवत्ति का अनुमान करने की प्रमुख रीतियाँ निम्नांकित हैं : -

(1) स्वतंत्र हस्त वक्र रीति (Freehold Curve Method)

(2) माध्यों की रीति (Method of Averages)

- (अ) चुने गये बिन्दुओं की रीति (Selected Points Method)
- (ब) अर्द्ध-माध्य रीति (Semi-Average Method)
- (स) चल माध्य रीति (Moving Average Method)

(3) न्यूनतम वर्ग रीति (Method of Least Squares)

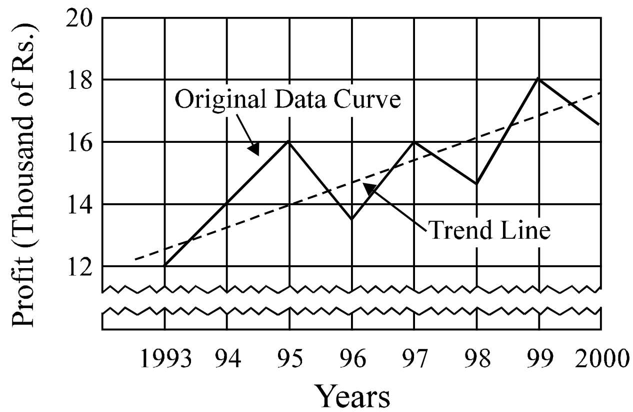
- (अ) अंकगणितीय सरल रेखा (Arithmetic Straight Line)
- (ब) लधुगणकीय सरल रेखा (Logarithmic Straight Line)
- (ग) एकेन्द वक्र (Parabolic Curve)

(1) स्वतंत्र हस्त वक्र रीति (Freehand Curve Method) : स्वतन्त्र हस्त वक्र रीति से दीर्घकालीन प्रवत्ति का अनुमान लगाने के लिए काल श्रेणी के मौलिक समंकों (Original Data) को बिन्दुरेखीय पत्र (Graph Paper) पर प्रांकित किया जाता है; तत्पश्चात् प्रांकित बिन्दुओं के मध्य से गुजरती हुई एक सरलित वक्र (Smoothed Curve) खींची जाती है। यही सरलित वक्र संख्याशास्त्री के मतानुसार दीर्घकालीन प्रवत्ति को इंगित करती है। सरलित वक्र खींचने में लचीले - पैमाने (Flexible Rulers) की सहायता ली जा सकती है अथवा बिन्दुरेखीय पत्र पर एक डोरा फैला का प्रांकित बिन्दुओं द्वारा प्रदर्शित प्रवत्ति के अनुसार उसे समायोजित करके फिर वक्र खींचा जा सकता है।

Illustration 1 : नीचे दिये समंकों से स्वतंत्र हस्त-वक्र विधि द्वारा उपनति मूल्य ज्ञात कीजिए :-
Calculate the value of Trend by Freehand Curve Method from the data given below :-

वर्ष (Year)	**लाभ (Profit) (Thousands of Rs.)**	**वर्ष (Year)**	**लाभ (Profit) (Thousanda of Rs.)**
1993	12	1997	16
1994	14	1998	15
1995	16	1999	18
1996	13	2000	17

Solution : Trend by Freehand Curve Method.



स्वतंत्र हस्त-वक्र विधि एवं दोष (Merits and Demerits of Freehand Curve Method) :-

गुण (Merits) : स्वतंत्र हस्त-वक्र विधि के निम्न गुण हैं :-

(1) यह दीर्घकालीन प्रवत्ति या उपनति ज्ञात करने की सरलतम विधि है।

(2) वक्र खींचने में अधिक समय नहीं लगता।

(3) इस विधि में गणितीय क्रिया का प्रयोग आवश्यक नहीं है।

दोष (Demerits) : स्वतंत्र हस्त-वक्र विधि के निम्न दोष हैं :-

(1) इसमें निश्चितता का गुण नहीं होता क्योंकि एक ही आँकड़ों से विभिन्न वक्र खींचे जा सकते हैं।

(2) इस विधि में शुद्धता का अभाव रहता है।

(3) यदि प्रयोग करने वाले के मन में पक्षपात की भावना हो तो वक्र खींचते समय उसका प्रभाव पड़ना स्वाभाविक है।

इन दोषों के होने के कारण इस विधि को बहुत कम प्रयोग में लाया जाता है।

(ब) अर्द्ध-माध्य रीति (Semi-Average Method)

यदि यह स्पष्ट होता है कि एक सरल रेखा दीर्घकालीन प्रवत्ति को उचित रूप से व्यक्त कर सकेगी , तब अर्द्ध-माध्य रीति का प्रयोग किया जाता है। काल-श्रेणी को दो समान भागों में विभक्त कर दिया जाता है। यदि श्रेणी में पदों की संख्या विषम (Odd) होती है तो मध्य के पद को छोड़ दिया जाता है। प्रथम अर्द्ध भाग के चर मूल्यों को जोड़कर सम्मिलित पदों की संख्या से भाग देकर उस भाग का माध्य ज्ञात किया जाता है। इसी प्रकार दूसरे अर्द्ध भाग के चर मूल्यों को जोड़कर पदों की संख्या से भाग देकर माध्य ज्ञात किया जाता है। काल-श्रेणी के मौलिक समंकों को बिन्दुरेखीय पत्र पर प्रांकित कर लिया जाता है। दोनों अर्द्ध-भागों के माध्यों को भी प्रत्येक भाग के मध्य में प्रांकित कर लेते हैं तथा उन दोनों बिन्दुओं को सरल रेखा द्वारा जोड़ दिया जाता है। यही सरल रेखा दीर्घकालीन प्रवत्ति को व्यक्त करती है।

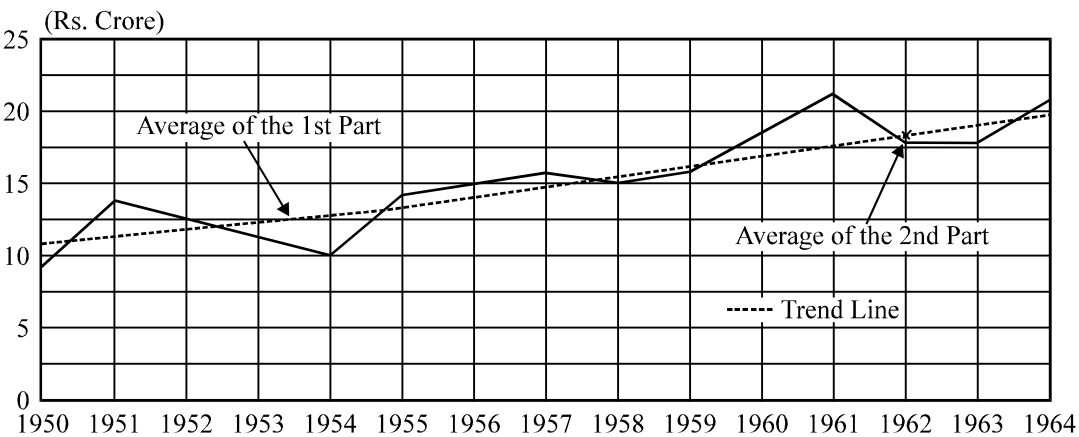
यदि उचित समझा जाता है और यदि सम्भव हो तो काल-श्रेणी को दो से अधिक भागों में भी विभक्त किया जा सकता है। प्रत्येक भाग के माध्य ज्ञात करके माध्य मूल्यों को सम्बन्धित भाग के मध्य में प्रांकित कर दिया जाता है। प्रांकित बिन्दुओं को सरल रेखा द्वारा जोड़कर प्रवत्ति रेखा खींची जाती है।

Illustration 2 : Determine the trend of the following by semi-average method.

Year	**Export (Rs. crore)**	
1950	9.5	
1951	14.36	
1952	12.5	
1953	11.9	S = 88.2 ÷ 7 = 12.6 (Average)
1954	10.1	
1955	14.9	
1556	15.0	
1957	15.7	

Year	Export (Rs. crore)	
1958	15.0	
1959	14.4	
1960	18.9	
1961	20.8	S = 126.2 ÷ 7 = 18.6 (Average)
1962	17.8	
1963	17.6	
1964	20.5	

Solution :



अर्द्ध-माध्य विधि के गुण एवं दोष (Merits and Demerits of Semi-Average Method)

गुण (Merits) : अर्द्ध-माध्य विधि के निम्न गुण हैं :-

(1) स्वतंत्र हस्त-वक्र विधि की तरह वह विधि भी सरल है।

(2) वक्र खींचने में अधिक समय और श्रम नहीं लगता।

(3) यदि इस विधि के प्रयोग करने वाले के मन में पक्षपात की भावना हो तो वक्र खींचते समय इसका कोई प्रभाव नहीं पड़ता।

(4) इस विधि के अन्तर्गत बिन्दुओं को सरलित करने का एक निश्चित एवं स्पष्ट आधार होता है।

दोष (Demerits) : अर्द्ध-माध्य विधि के निम्न दोष हैं :-

(1) चरम मूल्यों (Extreme Values) की उपस्थिति के कारण यदि समान्तर माध्य श्रेणी का उचित प्रतिनिधित्व नहीं करता तो इस विधि द्वारा प्रस्तुत की गई प्रवत्ति भी वास्तविक नहीं होगी।

(2) यह विधि तभी प्रयोग में लाई जा सकती है जब आँकड़ों की आवत्ति लगभग सीधी रेखा की हो।

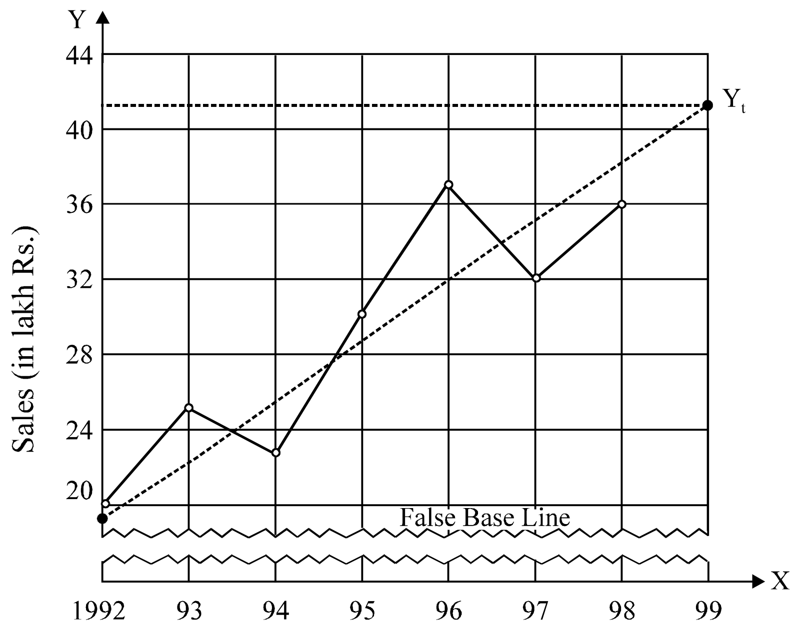
Illustration 3 : Draw a trend line by the method of semi-averages for the following data. Also predit sales for (1999).

निम्न समंकों से अर्ध-माध्य रीति द्वारा दीर्घकालीन प्रवत्ति ज्ञात कीजिए। 1999 की बिक्री का अनुमान भी लगाइए।

Years	:	1992	1993	1994	1995	1996	1997	1998
Sales (in Lakh Rs.)	:	21	25	23	30	37	32	36

Solution :

X Years	Sales	Semi-Totals	Semi-Average
1992 1993 1994	21 25 → 23	69 →	$\frac{21\quad 25\quad 23}{3}$ $\frac{69}{3}$ 23
1995	30	X (ignored)	
1996 1997 1998	37 32 → 36	105 →	$\frac{37\quad 32\quad 36}{3}$ $\frac{105}{3}$ 35



The trend value of 1994 is 41 approx.

1999 की दीर्घकालीन प्रवत्ति का अनुमान लगाने के लिए y_t रेखा का बढाएंगे व 1999 से इस रेखा का लम्ब (^) डालेंगे। जहाँ भी यह लम्ब y_t को काटेगा उस स्थान पर y-axis द्वारा प्राप्त मूल्य की 1999 की अनुमानित बिक्री होगी।

गणितीय विधि द्वारा भी उपनति मूल्य ज्ञात किया जा सकता है।

$$\text{Annual Change} = \frac{35 \quad 23}{1997 \quad 1993} = \frac{12}{4} = 3$$

$$\text{Semi Average value of 1997} = 35 \text{ (Given)}$$

$$\text{Trend Value of 1998} = 35 + 3 = 38$$

$$\text{Trend Value of 1999} = 38 + 3 = 41$$

चल-माध्य विधि (The Moving Average Method) : व्यवहार में, दीर्घकालीन प्रवत्ति ज्ञात करने के लिए चल -माध्य विधि को अधिक प्रयोग में लाया जाता है। यह विधि उच्चावचनों को एक बड़ी सीमा तक कम करने की आसान विधि है। चल-विधि की गणना करते समय सबसे पहले एक अवधि निश्चित कर ली जाती है। जैसे चल-माध्य की अवधि तीन वर्ष, चार वर्ष, छः वर्ष, सात वर्ष इत्यादि। मान लीजिए चल-माध्य की अवधि तीन वर्ष है तो सबसे पहले श्रेणी के पहले तीन वर्षों के मूल्यों को जोड़कर उनका माध्य ज्ञात कीजिए और इस माध्य को दूसरे वर्ष के समक्ष लिख दिया जाएगा (क्योंकि पहले तीन वर्षों का औसत वर्ष दूसरा वर्ष होगा) अब अगला माध्य मूल्य ज्ञात करने के लिए वर्ष के मूल्यों को छोड़कर अगले चौथे वर्ष के मूल्य को जोड़कर माध्य मूल्य ज्ञात किया जाता है और इस माध्य-मूल्य को तीसरे वर्ष के समक्ष रखा जाएगा (क्योंकि दूसरे, तीसरे एवं चौथे वर्ष का माध्य वर्ष तीसरा वर्ष होगा) इस प्रकार अगले वर्षों के माध्य मूल्य प्राप्त किए जा सकते हैं।

चल माध्य ज्ञात करने के लिए निम्न सूत्रों का प्रयोग किया जाता है :-

(*i*) तीन वर्षीय चल-माध्य (Three-yearly Moving Averages)

$$\frac{x_1 \quad x_2 \quad x_3}{3}, \frac{x_2 \quad x_3 \quad x_4}{3}, \frac{x_3 \quad x_4 \quad x_5}{3} \text{ इत्यादि}$$

(*ii*) पंचवर्षीय चल-माध्य (Five-yearly Moving Averages)

$$\frac{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5}{5}, \frac{x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6}{5}, \frac{x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7}{5} \text{ इत्यादि}$$

(*iii*) सातवर्षीय चल-माध्य (Seven-yearly Moving Averages)

$$\frac{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7}{7}, \frac{x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8}{7} \text{ इत्यादि}$$

उदाहरण (Illustration) 4 : निम्न समंकों से तीन वर्षीय चल-माध्य ज्ञात कीजिए :-

From the following data find out 3-yearly moving averages :-

वर्ष (Year)	:	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
बिक्री (Sales) ('000 Rs.)	:	30	25	35	20	24	25	28	26	28	32

हल (Solution) : Three-yearly Moving Averages.

Year ('000 Rs.) (*i*)	**Sales (Moving Totals) (*ii*)**	**Three-yearly Averages (Col. *iii* + 3) (*iii*)**	**Three-yearly Moving (*iv*)**
1990	30	—	—3
1991	25	30 + 25 + 35 = 90	30.00
1992	35	35 + 20 + 24 = 79	26.67
1993	20	35 + 20 + 24 = 79	26.33
1994	24	20 + 24 + 25 = 69	23.00
1995	25	24 + 25 + 28 = 77	25.67
1996	28	25 + 28 + 26 = 79	26.33
1997	26	28 + 26 + 28 = 82	27.33
1998	28	26 + 28 + 32 = 86	28.67
1999	32	—	—

Illustration 5 : Find the 7-yearly moving averages from the following data :-

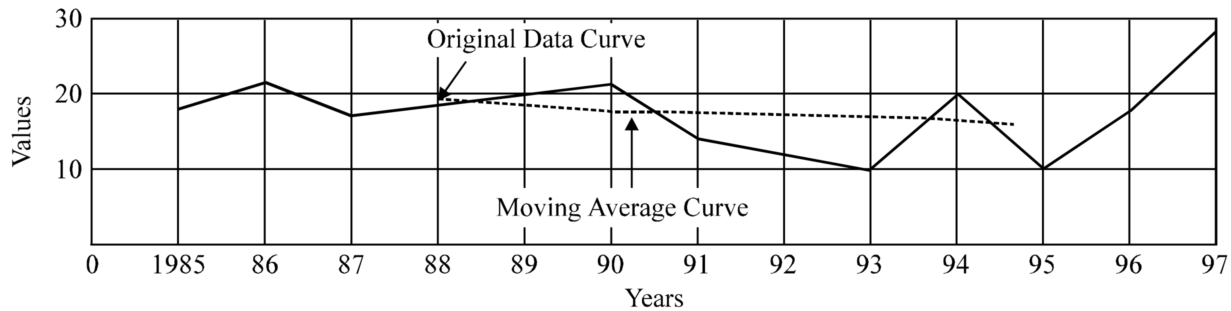
Year	**:**	1985	1986	987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Value	**:**	19	22	17	18	20	21	16	12	10	23	12	15	30

हल (Solution) : Seven-yearly Moving Averages.

Year (*i*)	**Value (*ii*)**	**Seven-yearly Moving (Totals) (*iii*)**	**Seven-yearly Moving Averages (Col. *iii* + 5) (*iv*)**
1985	19	—	—
1986	22	—	—
1987	17	—	—
1988	18	19 + 22 +17 +18 + 20 + 21 + 16 = 133	19.0

1989	20	22 + 17 + 18 + 20 + 21 + 16 + 12 = 126	18.0
1990	21	17 + 18 + 20 + 21 + 16 + 12 + 10 = 114	16.3
1991	16	18 + 20 + 21 + 16 + 12 + 10 + 23 = 120	17.1
1992	12	20 + 21 + 16 + 12 + 10 + 23 + 12 = 114	16.3
1993	10	21 + 16 + 12 + 10 + 23 + 12 + 15 = 109	15.6
1994	23	16 + 12 + 10 + 23 + 12 + 15 + 30 = 118	16.9
1995	12	—	—
1996	15	—	—
1997	30	—	—

Graph Showing Seven-Yearly Moving Average.



(2) चल-माध्य अवधि सम (Even) होने पर : यदि चल-माध्य सम वर्षों का हो अर्थात 4, 6, 8 अथवा 10 वर्षों का हो तो चल माध्यों को कन्द्रित भी करना पड़ेगा, क्योंकि सम वर्षों का चल-योग व चल माध्य दो वर्षों के मध्य में लिखा जाएगा। उदाहरणार्थ — 1990, 1991, 1992 व 1993 के समंकों का योग 1191 व 1992 के माध्य में लिखा जाएगा और 4 वर्षों का माध्य भी 1991 व 1992 के मध्य में ही होगा। इसी प्रकार प्रथम समंक छोड़कर अगले चार वर्षों का चल-योग व चल-माध्य भी दो वर्षों के मध्य ही आएगा। ऐसी परिस्थिति सम वर्षों के चल-माध्यों निकालते समय आती है। अतः इन वर्षों के चल-माध्य ज्ञात करने के उपरान्त इन्हें केन्द्रित करना पड़ता है। अर्थात् यदि हम दो वर्ष का चल-माध्य निकालें तो समस्त वर्षों के सम्मुख दीर्घकालीन प्रवत्ति ज्ञात की जा सकती है। इस प्रकार सम -वर्षों के चल-माध्य में चल-माध्यों को केन्द्रित करना अनिवार्य हो जाता है। जैसा कि निम्न उदाहरण से स्पष्ट है।

Illustration 6 : Calculate trend to moving average method assuming 4-yearly cycle from the following data relating to the production of tea.

चार वर्षीय चक्र मानकर चल-माध्य रीति द्वारा चाय के उत्पादन से सम्बन्धित निम्न समंकों से दीर्घकालीन प्रवत्ति ज्ञात करें।

Years	Production of Tea (in Quintals)	Years	Production of Tea (in Quintals)
1988	462	1995	560
1989	510	1996	580
1990	525	1997	620
1991	470	1998	645
1992	515		
1993	550		
1994	575		

Solution : Computation of Trendy by 4 yearly moving average.

Column I	2	3	4	5	6
Year	**Production**	**4 yearly moving Total**	**4 yearly moving Average**	**Centring Total**	**Average Y_t**
1998	462		—	—	—
1989	510		—	—	—
		1967	491.75		
1990	525			996.75	498.4
		2020	505.0		
1991	470			1020	510.0
		2060	515.0		
1992	515			1042.5	521.2
		2110	527.5		
1993	550			1077.5	538.75
		2200	550.0		
1994	575			1116.25	558.13
		2265	566.25		
1995	560			1150.0	575.0
		2335	583.75		
1996	580			1185.0	592
		2405	601.25		
1997	620	—	—	—	
1998	645	—	—	—	

Steps to Compute 4 yearly moving average

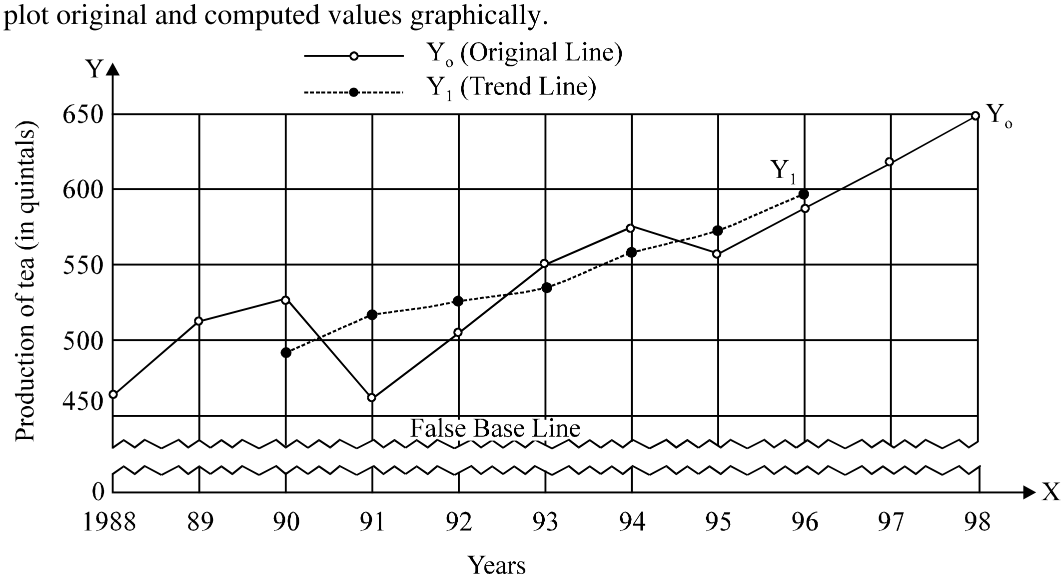
(1) In column number (13) compute 4 yearly moving data moving totals from data given in column number (2) and write it in the middle of first 4 years and so on.

(2) Divide the totals obtained in column 3 by 4 to get 4 yearly moving average.

(3) Since the moving averages come in the middle of 2 years, so compute 2 yearly moving totals of data obtained in column (4) to get the requisite total against respective years. This is called centring total.

(4) Now divide figures given in column number 5 by 2 to get the centring averages.

(5) Now plot original and computed values graphically.



गुण (Merits) :

(1) यह रीति न्यूनतम वर्ग रीति की तुलना में सरल है।

(2) समंकों की दीर्घकालीन प्रवत्ति जानने में यह विधि अधिक लचीला (Flexible) है क्योंकि यदि काल श्रेणी में कुछ और वर्षों के आँकड़े सम्मिलित कर लिए जाएं तो दीर्घकालीन प्रवत्ति की समस्त गणना नहीं बदलती , बल्कि पुरानी गणना में ही कुछ गणना की और वद्धि हो जाती है।

(3) यदि दीर्घकालीन प्रवत्ति निकालते समय चल -माध्य का समय चक्रीय उच्चावचन के समान हो तो चक्रीय उच्चावचन का उतार-चढ़ाव स्वतः ही दूर हो जाता है।

(4) चल-माध्य का स्वरूप आँकड़ों के स्वरूप पर निर्भर है न कि गणितीय आधार पर।

(5) यदि दीर्घकालीन प्रवत्ति में बहुत अनियमितता है तो चल -माध्य रीति द्वारा प्राप्त निष्कर्ष अधिक विश्वसनीय होंगे।

(6) इस रीति के प्रयाग में मानव की पक्षपात भावना का कोई प्रभाव नहीं पड़ता।

दोष (Demerits or Limitations) :

(1) इस पद्धति का सबसे बड़ा दोष यह है कि इस रीति में सभी दिए हुए वर्षों का दीर्घकालीन मूल्य प्राप्त नहीं होता। चल-माध्य की अवधि जितनी अधिक होगी, उतने ही अधिक वर्षों के चरम मूल्य (Extreme Values) का मान मालूम नहीं होगा। उदाहरणार्थ, 3 वर्षीय चल माध्य में एक वर्ष प्रथम व एक वर्ष अन्त की दीर्घकालीन प्रवत्ति ज्ञात नहीं होगी। इसी प्रकार 7 वर्षीय चल माध्य में प्रथम तीन वर्षों व अन्तिम तीन वर्षों की दीर्घकालीन प्रवत्ति ज्ञात नहीं की जा सकती।

(2) चल-माध्य अवधि का चयन करते समय बड़ी सावधानी बरतने की आवश्यकता है।

(3) इस रीति से भविष्य के बारे में अनुमानित मूल्य ज्ञात नहीं किए जा सकते।

(4) सैद्धान्तिक रूप से हम यह कह सकते हैं कि चल-माध्य व चक्रीय उच्चावचन अवधि समान होने पर चक्रीय उच्चावचन का असर स्वतः ही समाप्त हो जाता है किन्तु वास्तविक जीवन में चक्रीय उच्चावचन का समय प्रत्येक स्थिति में समान नहीं होता। अतः किसी भी अवधि का चल-माध्य चक्रीय उच्चावचन के प्रभाव को पूर्ण रूप से समाप्त नहीं कर सकता।

(5) यह विधि समान्तर माध्य की परिसीमाओं से प्रभावित होती है।

(2) न्यूनतम वर्ग रीति (Methods of Least Square) दीर्घकालीन प्रवत्ति को मापने की यह सर्वोत्तम विधि है। यह विधि गणितीय होने के कारण विश्लेषण करने के लिये भी सर्वोत्तम है। यह विधि आर्थिक व व्यावसायिक काल श्रेणियों पर प्रयोग की जा सकती है और इस प्रकार भविष्य के उपनति मूल्यों को भी ज्ञात किया जा सकता है। दीर्घकालीन उपनति मूल्य (*i*) रेखीय (linear) व (*ii*) अरेखीय (non-linear) भी हो सकते हैं।

यदि दीर्घकालीन तक काल श्रेणी में समंकों के बढ़ने व घटने की प्रवत्ति एक निश्चित दर से पायी जाती है तो इसे रेखीय उपनति (Straight Line Trend) कहा जाता है और इसे निम्न प्रकार व्यक्त करते हैं :-

Strainght Line Trend

$$Y_t = a + bX$$

where,

Y = '*y*' represents the trend value and is, therefore, also called as Y_t, or computed trend value (Y_c).

a = '*a*' is a constant and represent value of Y variable when X = 0.

b = '*b*' is another constant and represent slope of trend line. This discloses the change in value of '*y*' variable when there is a unit change in X [*i.e.,* 'X' changes by 1].

Generally X represents change in time period by 1 year.

उपरोक्त रेखा समीकरण **($Y_t = a + bX$) द्वारा प्राप्त दीर्घकालीन प्रवत्ति को सर्वाधिक उपयुक्तता की रेखा Line of Best Fit) भी कहते हैं।** यह रेखा समीकरण समंकों की प्रवत्ति को इस प्रकार स्पष्ट करती है कि -

(1) Y व इससे सम्बन्धित प्रवत्ति मूल्य (Y_t or Y_c) के अन्तर का योग सदैव शून्य (Zero) होता है। सूत्र के रूप में $S\,(Y - Y_t) = 0$।

(2) इसी प्रकार Y व इससे सम्बन्धित प्रवत्ति मूल्यों के अन्तर के वर्गों का योग सदैव कम से कम होगा। सूत्र के रूप से $S\,(Y - Y_t)^2 = 0$ will be the best. इसी कारण से इसे न्यूनतम वर्ग रीति कहते हैं

यदि दीर्घकाल में काल श्रेणी के समंकों में बढ़ने व घटने की दर असमान है तो ऐसी दशा में अरेखीय उपनति (Non-linear Trend) का प्रयोग किया जाता है। इस समीकरण को निम्न प्रकार से व्यक्त किया जाता है :-

$$\mathbf{Y_t = a + bX + cX^2}$$

इसी द्वितीय घात परवलयिक वक्र (Second Degree Parabolic Curve) भी कहते हैं।

उपरोक्त रेखा माध्य-मूल्यों को प्रदर्शित करती है। अतः इसे रेखा द्वारा प्राप्त प्रवत्ति किसी अन्य रेखा द्वारा प्राप्त प्रवत्ति से अच्छी होगी। अतः इस विधि से प्राप्त भविष्य के अनुमानित प्रवत्ति मूल्य भी अधिक विश्वसनीय होते हैं। प्रवत्ति का सरल (Straight) व परवलयिक वक्र (Parabolic Curve) दोनों ही प्रकार की हो सकती हैं। इस अध्याय में हम सरल रेखा का ही वर्णन करेंगे।

सरल रेखीय प्रवत्ति बनाने की विधि (Procedure of Drawing Straight Line Trend) न्यूनतम वर्ग रीति में सरल रेखा प्रवत्ति ज्ञात करने का समीकरण निम्न है :

$$Yt = a + bX$$

इस समीकरण के विभिन्न मूल्यों को पहले ही स्पष्ट किया जा चुका है। इसी समीकरण में 'a' व 'b' दोनों अचर मूल्यों (Constant) को ज्ञात करने के लिए निम्न दो समीकरणों की आवश्यकता होती है :-

$$SY = aN + bSX \quad ...(i)$$

$$SXY = aNS + bSX^2 \quad ...(ii)$$

प्रथम समीकरण $Y = a + bX$ के सभी समंकों का योग है, जैसा कि निम्नलिखित समीकरणों से स्पष्ट है :-

पहली समीकरण	$Y_1 = a + bX_1$
द्वितीय समीकरण	$Y_2 = a + bX_2$
ततीय समीकरण	$Y_3 = a + bX_3$
	:
	:
	$Y_n = a + bX_n$
योग करने पर	$SY_n = aN + bSX_n$

इसी प्रकार दूसरी समीकरण $Y = a + bX$ को दोनों ओर X से गुणा करने पर प्राप्त।

प्रथम समीकरण	$X_1Y_1 = aX_1 + bX_1^2$
द्वितीय समीकरण	$X_2Y_2 = aX_1 + bX_2^2$
ततीय समीकरण	$X_3Y_3 = aX_1 + bX_3^2$
	:
	:
योग करने पर	$SXY = aSX + bSX^2$

इन दोनों समीकरणों को हल करने व हमें a व b का मान मिलता है। इन मानों को $Y_t = a + bX$ समीकरण में रखने पर X की भिन्न कीमतों के लिए हमें Y के अलग-अलग मूल्य मिलेंगे। इन्हीं मूल्यों को हम प्रवत्ति मूल्य कहते हैं। निम्न उदाहरण से इस रीति का प्रयोग समझा जा सकता है :-

Illustration 7 : Fit a straight line trend to the following data by the method of least squares. Also estimate the trend value for 2000 :

Year	:	1993	1994	1995	1996	1997
Profit (in lakhs of Rs.)	:	45	56	78	46	75

Solution :

Year	**Profits (in lakhs of Rs.) Y**	**Deviation from 1995 X**	**Square of Devs X^2**	**Profits × Deviation XY**	**Trend Values (Y_c)**
1993	45	– 2	4	– 90	50
1994	56	– 1	1	56	55
1995	78	0	0	0	60
1996	46	+ 1	1	46	65
1997	75	+ 2	4	150	70
	SY = 300	**SY = 0**	**SX^2 = 10**	**SXY = 50**	**300**

The equation of the straight line trend in $Y_c = a + bX$

Since SX = 0; $a = \frac{Y}{N}$, $b = \frac{XY}{X^2}$

Hence SY = 300, N = 5, SXY = 50, SX^2 = 10

\ $a = \frac{Y}{N} = \frac{300}{5} = 60$, $b = \frac{XY}{X^2} = \frac{50}{10} = 5$

Now the trend value for the different years can be calculated in the following manner :-

Trend value for 1993 = 60 + 5 (– 2)
= 60 – 10 = 50

Trend value for 1994 = 60 + 5 (– 1)
= 60 – 5 = 55

Trend value for 1995 = 60 + 5 (0)
= 60 + 0 = 60

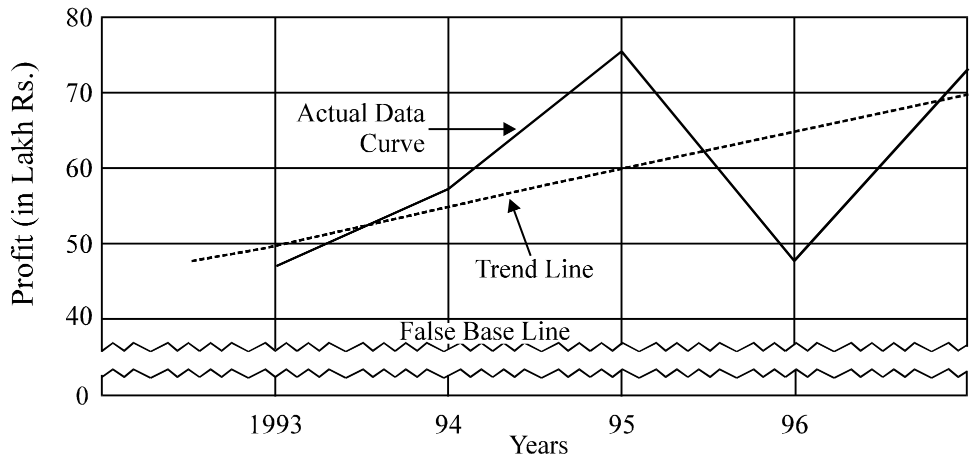
Trend value for 1996 = 60 + 5 (1)
= 60 + 5 = 65

Trend value for 1997 = 60 + 5 (2)
= 60 + 10 = 70

For 2000 the value of X will be 5.

\ Trend value for 2000 = 60 + 5 (5)
= 60 + 25 = 85

Trend line by the method of least square.



Steps to solve question :-

(1) First of all write the years of data in terms of unit as 1, 2, 3 etc. in order of occurence of various years.

(2) Square the value of X written in terms of unit.

(3) Multiply 'X' in terms of units with corresponding values given by Y.

(4) Add Y, X, X^2 and XY and put these values in two normal equations to find out the values of constants '*a*' and '*b*'.

(5) The value of '*b*' is called as regression coeffcSiennt of Y and X and is denoted with b_{yx}.

(6) Put the value of X in terms of units 1, 2, 3 etc. in trend line equations $Y_t = a + bX$ to determine the trend values of various years given in the data. These computed values of Y are generally denoted with Y_t or Y_c *i.e.* trend values of Y or computed values of Y.

(7) Now plot the values of y_0 and y_t graphically in usual way. Take the help of false base line if necessary.

लघु रीति (Short Cut Method)

(*i*) समान्तर माध्य से विचलन (Deviations taken from mean)

(*ii*) कल्पित माध्य से विचलन (Deviations taken from assumed mean)

(*i*) लघु रीति (समान्तर माध्य से विचलन) Short Cut Method (Deviation taken from mean).

(1) इस रीति से प्रश्न को हल करते समय समंकों के वर्षों का मान इकाई (unit) 1, 2, 3, ... आदि के रूप में लिखा जाता है। तत्पश्चात् X में लिखी इकायों का समान्तर माध्य ($\overline{X}$) निकालकर X का माध्य से विचलन (Deviation of X from mean) ज्ञात करेंगे।

(2) विचलन ज्ञात करने के उपरान्त इन विचलनों का वर्ग करेंगे व विचलन का Y से गुणा करेंगे।

(3) इस प्रकार हमें Y, x, x^2, xY का योग प्राप्त हो जाएगा।

इस विधि के प्रयोग से हमें दोनों सामान्य समीकरणों को '*a*' व '*b*' का मान ज्ञात करने के लिए, हल नहीं करना पड़ेगा। क्योंकि Sx का योग शून्य (Zero) होगा, अतः '*a*' व '*b*' का मान बिना समीकरणों के हल किए ही ज्ञात किया जा सकता है। जैसे :-

प्रथम समीकरण $\quad SY = Na + bSx$ में

यदि Sx का योग शून्य है तो $\quad SY = Na$ or $a = \dfrac{Y}{N}$

इसी प्रकार से दूसरे समीकरण $\quad SXY = aSX + bSX^2$ में $SX = 0$ तो

$$bSX^2 = SXY$$

या $$b = \frac{XY}{X^2}$$

इस तरह से ‘a’ व ‘b’ दोनों का मान बिना समीकरण हल किये मिल जाते हैं।

निम्न उदाहरणों से ये रीति स्पष्ट हो जाएगी।

(A) Odd number of years

Illustration 8 : Fit a straight-line trend by method of least-squares.

Year	1980	1981	1982	1983	1984	1985	1986
Sales	83	60	54	21	22	13	23

Solution :

X	**Y**	**X in term of units X**	$\bar{X} = 4$ $(X - \bar{X})$	x^2	xY	$Y_t = a + bx$ $Y_t = 39.43 - 10.93\,x$
1980	82	1	– 3	9	– 249	39.43 – 10.93 (–3) = 72.22
1981	60	2	– 2	4	– 120	39.43 – 10.93 (–2) = 61.29
1982	54	3	– 1	1	– 54	39.43 – 10.93 (–1) = 50.36
1983	21	4	0	0	0	39.43 – 10.93 (0) = 39.43
1984	22	5	1	1	22	39.43 – 10.93 (1) = 28.50
1985	13	6	6	4	25	39.43 – 10.93 (2) = 17.57
1986	23	7	3	9	69	39.43 – 10.93 (3) = 6.64
N = 7	**SY = 276**	**SX = 28**	**Sx 0**	**Sx^2 = 28**	**SxY = – 306**	

$$\bar{X} = \frac{X}{N} \quad \frac{28}{7} \quad 4$$

We know that straight line trend equation is —

$$Y = a + bx$$

To find the value of ‘a’ and ‘b’ constants, we have two normal equations namely —

$$SY = aN + bSx \quad ...(i)$$

$$SxY = aSx + bSx^2 \quad ...(ii)$$

In equation (i), $Sx = 0$

\ $$a = \frac{Y}{N} \quad \frac{276}{7} \quad 39.43$$

In equation (ii), $Sx = 0$

\ $$b = \frac{xY}{x^2} = \frac{306}{28} = -10.93$$

The trend equations is —

$$Y_t = a + bx$$

$$Y_t = 393.43 - 0.93\,x$$

Illustration 9 : Find the trend values from the following data and estimate the value of Y for year 2000.

Year	1990	1991	1992	1993	1994	1995	1996
Production (Million Tonnes)	55	62	65	58	65	72	78

Solution.

Year	X	Y	X^2	XY	Y_t
1990	– 3	55	9	– 165	55.46
1991	– 2	62	4	– 124	58.64
1992	– 1	65	1	– 65	61.82
1993	0	58	0	0	65.00
1994	1	65	1	65	68.18
1995	2	72	4	144	71.36
1996	3	78	9	234	74.54
	SX = 0	**SY = 455**	**SX² = 28**	**SXY = 89**	

$$a = \frac{Y}{N}$$

$$= \frac{455}{7} = 65$$

$$b = \frac{XY}{X^2}$$

$$= \frac{89}{28} = 3.18$$

यदि वर्षों की संख्या विषम हो तो इस बीच वाले वर्ष (Middle Year) जोकि $\left(\frac{N \quad 1}{2}\right)^{th}$ वर्ष होता के लिए X का मूल्य शून्य ले लेते हैं।

$$\begin{aligned} \text{Trend equation is, } Y &= 65 + 3.18\ X \\ \text{for year 2000, } X &= 7 \\ \backslash \qquad Y_t \text{ for 2000} &= 65 + 3.18 \times 7 \\ &= 87.26 \end{aligned}$$

Short-cut Method (with even numbes of years)

यदि प्रश्न में दिए गऐ विभिन्न वर्षों का योग सम (even) हो तो ऐसी स्थिति में X का माध्य सदैव ही भिन्न (Fraction) में आएगा और X का माध्य से विचलन (Deviation for mean) ज्ञात करने पर x भी भिन्न (Fraction) में ही आएगा। यह प्रायः – 3.5, – 2.5, – 1.5, –.5, +.5, +1.5, + 2.5 व + 3.5 आदि आएंगे। अतः ऐसी दशा में यदि x को 0.5 से भाग करें तो दशमलव x से हट जाएंगे और एक नया खाना x' के नाम से बनेगा जिसमें x' – 5, –3, –1, 1, 3, 5 आदि आएंगे। ऐसी दशा में $Y_t = a = bx'$ का मूल्य ज्ञात करते समय x' का मूल्य ही भरना पड़ेगा। ऐसा करने से प्रश्न आसानी से हल किया जा सकता है। निम्न उदाहरण से इसका प्रयोग देखा जा सकता है।

Illustration 10 : Fit a straight line trend by the method of Least Squares in the following Series :-

निम्न श्रेणी से न्यूनतम वर्ग विधि द्वारा सरल रेखा उपनति की रचना कीजिए :-

Year	1954	1955	1956	1957	1958	1959
Production	7	10	12	14	17	24
(in crores of rupees)						

[B.Com, K.U. 1980 and M.D.U. 1996]

Solution :

Years X	Production Y	X in terms of units x	$\overline{X} = 35$ $(X - \overline{X})$	Dividing x by $0.5 = x'$	x'^2	x'^3	$Y = a + bx'$ $Y_t = 14 + 1.54\,x$
1954	7	1	– 2.5	– 5	25	– 35	14 + 1.54 (–5) = 6.3
1955	10	2	– 1.5	– 3	9	– 30	14 + 1.54 (–3) = 9.38
1956	12	3	– 0.5	– 1	1	– 12	14 + 1.54 (–1) = 12.46
1957	14	4	0.5	1	1	14	14 + 1.54 (1) = 15.54
1958	17	5	1.5	3	9	51	14 + 1.54 (3) = 18.62
1959	24	6	2.5	5	25	120	14 + 1.54 (5) = 21.7
N = 6	**84 SY**	**21 = SX**		**0 = Sx'**	**70 = Sx'^2**	**108 = Sx'Y**	

$$a = \frac{Y}{N}$$

$$= \frac{84}{6} = 14$$

$$b = \frac{XY}{X^2} = \frac{108}{70} = 1.54$$

\ Trend equation is Y = 14 + 1.54 X

Illustration 11. Calculate trend values from the following data by applying the equation $Y_t = a + bX$

Year	1991	1992	1993	1994	1995	1996
Sales (in crores)	100	108	128	140	182	190

Estimate the likely sales for the year 1998.

Solution :

Years X	Sales in croers (Y)	X in term of units X	$\overline{X} = 3.5$ $(X - \overline{X})$ x	xY	x^2	$Y_t = a + bx$ $Y_t = 141.33 + 19.54\,x$
1991	100	1	– 2.5	– 250	6.25	141.33 + 19.54 (– 2.5) = 92.48
1992	108	2	– 1.5	– 162	2.25	141.33 + 19.54 (– 1.5) = 112.02
1993	128	3	– .5	– 64	. 25	141.33 + 19.54 (– 1.5) = 131.56
1994	140	4	+ .5	+ 70	.25	141.33 + 19.54 (+ .5) = 151.10
1995	182	5	+ 1.5	+ 273	2.25	141.33 + 19.54 (1.5) = 170.64
1996	190	6	+ 2.5	+ 475	6.25	141.33 + 19.54 (2.5) = 190.18
N = 6	**SY = 848**	**SX = 21**	**Sx = 0**	**Sxy = 342**	**Sx^2 = 17.50**	

$$\overline{X} = \frac{X}{N} = \frac{21}{6} = 3.5$$

We know the straight line trend equation is —

$$Y = a + bx$$

To solve '*a*' and '*b*' constants, we have two normal equations, namely —

$$SY = aN + bSx \quad ...(i)$$

$$SxY = aSx + bSx^2 \quad ...(ii)$$

Since, $Sx = 0$ in eq. (*i*) and (*ii*)

\ $$a = \frac{Y}{N} = \frac{848}{6} = 141.00$$

$$b = \frac{XY}{X^2} = \frac{342}{17.5} = 19.54 \quad [\because \; x \quad 0]$$

The trend equation is

$$Y_t = a + bx$$

or $$Y_t = 141.33 + 19.54\,x$$

The trend value for the year 1998.

Here, value of x is 4.5 \ $$Y_t = a + bx$$

$$= 141.33 + 19.54\,(4.5)$$

$$= \text{Rs. } 229.26 \text{ crores.}$$

लघु रीति (कल्पित माध्य से विचलन) **Short-cut Method** (Deviations taken from Assumed Mean) : प्रायः लघु रीति का प्रयोग करते समय हम समान्तर माध्य से ही विचलन लेते हैं। किन्तु यदि समान्तर माध्य दशमलव भिन्न में आ जाये तो हम कल्पित माध्य से भी विचलन ले सकते हैं। ऐसी दशा में क्योंकि [Sx 0] शून्य नहीं होगा अतः हमें '*a*' व '*b*' का मान ज्ञात करने के लिए दोनों समीकरणों को प्रत्यक्ष रीति से समान ही हल करना पड़ेगा। उदाहरण — नीचे कल्पित माध्य से विचलन लेकर समझाया गया है।

Illustration 12 : Find the trend value, by taking deviation from assumed mean, 11, from the following data :

Year	1951	1961	1971	1981	1991
Y	34	50	67	75	85

Solution :

X	Y	X in term of units X	A = 11 (X –) x	x^2	xY	$Y_t = a + bx$ $Y_t = 49.5 + 1.27\,x$
1951	34	1	– 10	100	– 340	19.5 + 1.27 (– 10) = 36.8
1961	50	11	0	0	0	19.5 + 1.27 (0) = 49.5
1971	67	21	10	100	670	19.5 + 1.27 (10) = 62.2
1981	75	31	20	400	1500	19.5 + 1.27 (20) = 74.9
1991	85	41	30	900	2550	19.5 + 1.27 (30) = 87.6
N = 5	**311 = SY**		**50 = Sx**	**1500 = Sx^2**	**4380 = Sxy**	

Note : Here deviation has been taken from assumed mean = 11

We know that straight line equation is :-

$$y = a + bx$$

To solve '*a*' and '*b*' constants, we have 2 normal equations :-

$$\text{S}Y = aN + b\text{S}x \quad ...(i)$$

$$\text{S}xY = a\text{S}x + b\text{S}x^2 \quad ...(ii)$$

By putting the values in equations, we have —

$$311 = 5a + 50b \quad ...(i)$$

$$4380 = 50a + 1500b \quad ...(ii)$$

Multiply eq. (*i*) by 10 both sides, we have —

$$3110 = 50a + 500b \quad ...(iii)$$

On subtracting eq. (*iii*) from eq. (*ii*), we have —

$$4380 = 50a + 1500b \quad ...(ii)$$

$$3110 = 50a + 500b \quad ...(iii)$$

$$1270 = 1000b$$

$$b = \frac{1270}{1000} = 1.27$$

By putting the value of *b* in eq. (*i*) we have

$$311 = 5a + 50\,(1.27)$$

or

$$5a = 311 - 63.5 = 247.5$$

$$a = \frac{247.5^{\overline{X}}}{5} = 49.5$$

So the trend equation is

$$Y_t = 49.5 + 12.7\,b$$

The trend values obtained are shown in the table above.

Method of Least Square (when there is a gap in the given years).

यदि प्रश्न में दिए गए विभिन्न वर्षों में कहीं अन्तराल (gap) हो तो ऐसी परिस्थति में प्रश्न में दिए गए वर्षों को इकाई में परिवर्तित करते समय वर्षों को क्रमानुसार अभिव्यक्त करेंगे, उदाहरणतया X — 1, 2, 4, 5, 8, 9 आदि। उपरोक्त उदाहरण में इकाई 3, 6 व 7 नहीं दी गई हैं। अब X का योग (SX) करके देखना चाहिए कि प्रश्न में माध्य ( ) दशमलव (Decimals) में तो नहीं आ रहा है। यदि माध्य ($\overline{X}$) पूर्ण है तो लघु रीति (Short-cut Method) अधिक उपयुक्त होगी, किन्तु यदि माध्य ($\overline{X}$) दशमलव में है तो प्रत्यक्ष विधि (Direct Method) अथवा कल्पित माध्य से विचलन (लघु रीति) अधिक उपयुक्त रहेगी। निम्न उदाहरणों से इसे अच्छी प्रकार समझा जा सकता है।

Illustration 13 : Fit a straight line trend to the following data by method of least square.

निम्न आंकड़ों से सरल रेखा उपनति की रचना कीजिए।

Year	**:**	1991	1992	1994	1996	1997
Sales (in lakh Rs.)	**:**	5	8	12	20	25

Also estimate the trend value of 1999 and find out monthly increase in sales volume.

वर्ष 1999 का उपनति मूल्य ज्ञात कीजिए तथा बिक्री मात्रा से मासिक वद्धि ज्ञात कीजिए।

Solution : **SHORT-CUT METHOD**

Years X	Sales (in lakhs Y	X in term of units	$(X - \overline{X})$ Take Deviations from $\overline{X} = x$	x^2	xY	$Y_t = a + bx$ $Y_t = 14 + 3.23\,x$
1991	5	1	– 3	=	– 15	14 + 3.23 (– 3) = 4.31
1992	8	2	– 2	4	– 16	14 + 3.23 (– 2) = 7.54
1994	12	4	0	0	0	14 + 3.23 (0) = 14.00
1996	20	6	2	4	40	14 + 3.23 (2) = 20.49
1997	25	7	3	9	75	14 + 3.23 (3) = 23.69
N = 5	**70 = SY**	**20 = SX**	**Sx = 0**	**26 = Sx^2**	**84 = Sxy**	

$$a = \frac{Y}{N} = \frac{70}{5} = 14$$

$$b = \frac{XY}{X^2} = \frac{84}{26} = 3.23$$

Trend equation is

$$Y_t = 14 + 3.23\,X$$

For year 1999, X = 11

$$\backslash \qquad Y_t \text{ for } 1999 = 14 + 3.23 \times 11$$
$$= 14 + 35.53$$
$$= \text{Rs. } 49.53 \text{ lakhs}$$

एकेन्द्र वक्र (Parabolic Curve) : सरल रेखा प्रवत्ति समीकरण साधारण Polynomials से सम्बन्धित होते हैं। अधिक लोच के लिये तथा वक्रीय-रेखा प्रवत्ति (Curvilinear Trend) प्रदर्शित करने के लिऐ दूसरे तथा तीसरे स्तर के एकेन्द्र वक्रों का प्रयोग किया जाता है। दूसरे स्तर के एकेन्द्र वक्र (Second Degree of Parabola) का समीकरण निम्नांकित है —

$$Y = a + bx + cx^2$$

इस समीकरण में a, b तथा c अचर हैं जिनका मान ज्ञात करने के लिए निम्नलिखित सामान्य समीकरणों को प्रयोग किया जाता है :-

$$Sy = Na + bSx + cSx^2$$
$$Sxy = aNx + bSx^2 + cSx^3$$
$$Sx^2y = aNx^2 + bSx^3 + cSx^4$$

यदि काल श्रेणी में मध्य बिन्दु से विचलन ज्ञात किये जाते हैं तो सामान्य समीकरण निम्न प्रकार हो जाते हैं :-

$$Sy = Na = cSx^2$$
$$Sxy = bSx^2$$
$$Sx^2y = aSx^2 + cSx^4$$

a, b तथा c अचरों का मान निम्नलिखित संक्षेप-सूत्रों से ज्ञात किया जा सकता है :-

$$a = \frac{y \quad c \quad x^2}{N},\ b = \frac{xy}{x^2},\ c = \frac{x^2y \qquad x^2 \quad y}{N \quad x^4 \qquad \left(x^2\right)^2}$$

Illustration 14 : Fit a parobola of the second order to the following data :-

Year :	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965
y :	17	20	19	24	24	40	35	55	51	74	69

Solution :

Year	(y)	Time Deviation from 1960 (x)	(xy)	(x^2)	(x^2y)	(x^3)	(x^4)	Computed 'y' or trend ordinate
1955	17	– 5	– 85	25	425	– 125	625	16.55
1956	20	– 4	– 80	16	320	– 64	256	18.48
1957	19	– 3	– 57	9	171	– 27	81	21.27
1958	26	– 2	– 52	4	104	– 8	16	24.92
1959	24	– 1	– 24	1	24	– 1	1	29.43
1960	40	0	0	0	0	0	0	34.80
1961	35	+ 1	+ 35	1	35	+ 1	1	41.03
1962	55	+ 2	+ 110	4	220	+ 8	16	48.12
1963	51	+ 3	+ 153	9	459	+ 27	81	56.07
1964	74	+ 4	+ 296	16	1184	+ 64	256	64.88
1965	69	+ 5	+ 335	25	1725	+ 125	625	74.55
***n* = 11**	**S*y* = 430**	**S*x* = 0**	**S*xy* = + 641**	**Sx^2 = 110**	**Sx^2y = 4667**	**Sx^3 0**	**Sx^4 = 1958**	

Normal equations are :

$$\text{S}y = \text{N}a + b\text{S}x + c\text{S}x^2 \quad \text{or} \quad 430 = 11a + b\,(0) + 110c$$
$$\text{S}xy = a\text{S}x + b\text{S}x^2 + c\text{S}x^3 \quad \text{or} \quad 641 = a\,(0) + 110b + c\,(0)$$
$$\text{S}x^2y = a\text{S}x^2 + b\text{S}x^3 + c\text{S}x^4 \quad \text{or} \quad 4667 = 110a + b\,(0) + 1958c$$

उपर्युक्त समीकरणों की सहायता से a, b तथा c के मान इस प्रकार निकलते हैं।

$$a = 34.8,\ b = 5.8,\ c = .43$$

द्वितीय स्तर एकेन्द्र वक्र का समीकरण $y = a + bx + cx^2$

$$\text{Y} = 34.8 + 5.8\,x + .43\,x^2$$

अतः x के मूल्य इस प्रकार निकाले जा सकते हैं :-

when $x = –5$, $y = 34.8 – 29 + 10.75 = 16.5$

$y = – 4$, $y = 14.48$ तथा इसी प्रकार अन्य मूल्य ज्ञात किये जा सकते हैं।

ततीय -स्तर एकेन्द्र वक्र (Third Degree Parobola) : समीकरण में चतुर्थ 'd' का समावेश करके ततीय स्तर एकेन्द्र वक्र बनाया जाता है। उसका समीकरण निम्न है :-

$$y = a + bx + cx^2 + dx^3$$

'd' चतुर्थ अचर के समावेश से वक्र में दो मोड़ आते हैं। उक्त समीकरण में अचर-मूल्यों को ज्ञात करने में निम्नांकित सामान्य समीकरणों का प्रयोग किया जाता है :-

$$\text{S}y = \text{N}a + b\text{S}x + c\text{S}x^2 + d\text{S}x^3$$
$$\text{S}xy = a\text{N}x + b\text{S}x^2 + c\text{S}x^3 + d\text{S}x^4$$
$$\text{S}x^2y = a\text{N}x^2 + b\text{S}x^3 + c\text{S}x^4 + d\text{S}x^5$$

$$\Sigma x^3y = aNx^3 + b\Sigma x^4 + c\Sigma x^5 + d\Sigma 5^6$$

When middle year is taken as origin, then

$$\Sigma y = Na + c\Sigma x^2$$
$$\Sigma xy = b\Sigma x^2 + d\Sigma x^4$$
$$\Sigma x^2y = a\Sigma x^2 + c\Sigma x^4$$
$$\Sigma x^3y = a\Sigma x^4 + d\Sigma x^6$$

सारांश

- किसी भी चर के मूल्यों का समय के आधार पर संकलन को काल श्रेणी कहते हैं।
- किसी भी काल श्रेणी में समय हमेशा स्वतंत्र चर होता है व दूसरा चर जैसे बिक्री, लाभ, जनसंख्या, आयात, निर्यात, उत्पादन, आय, व्यय, बचत इत्यादि आश्रित चर होते हैं।
- काल श्रेणी के चार संघटक होते हैं — दीर्घकालीन प्रवत्ति, आर्तव विचरण, चक्रीय उच्चावचन तथा अनियमित उच्चावचन। काल श्रेणी के विश्लेषण में हम किसी भी चर मूल्यों में वद्धि या कमी में इन चारों संघटकों के योगदान का अध्ययन करते हैं।
- काल श्रेणी विश्लेषण से हमें भविष्य के बारे में अनुमान लगाने में सहायता मिलती है जिससे भविष्य के लिए नीति निर्धारण का कार्य अच्छी तरह से किया जा सके। इसके अतिरिक्त वास्तविक व अनुमानित आंकड़ों का तुलनात्म्क अध्ययन, व्यापार चक्र का अनुमान भी कर सकते हैं।
- दीर्घकालीन प्रवत्ति निकालने की मुख्य विधियाँ चल माध्य रीति व न्यूनतम रीति विधि है।
- बाकि बचे संघटकों का मूल्य निकालने के लिए दो निर्देश हैं, योगात्मक (Y = T + S + C + I) तथा गुणात्मक (Y = T × S × C × 1) ।

  काल श्रेणी के विश्लेषण में गुणात्मक निदर्श ज्यादा प्रयोग किया जाता है।
- काल श्रेणी विश्लेषण सरकारी नीति निर्धारण में भी सहायता करती है, जैसे यदि जनसंख्या की काल श्रेणी के विश्लेषण से यह पाया जाए कि पुरुष तथा स्त्री का अनुपात कम होता जा रहा है तो सरकार कुछ ऐसी नीतियाँ बनायेगी जिससे आने वाले वर्षों में ये अनुपात बढ़ता रहे।

प्रश्नावली

(Exercise)

(1) 'काल श्रेणी विश्लेषण' को स्पष्ट रूप में समझाइए। इस प्रकार के विश्लेषण का व्यापार में महत्व इंगित कीजिए।

Explain clearly the meaning of 'Time Series Analysis'. Indiate the importance of such analysis in business.

(2) एक काल श्रेणी में पाये जाने वाले विभिन्न प्रकार के उच्चावचनों को समझाइए तथा कुछ उदाहरण देकर उनकी महत्ता प्रदर्शित कीजिए।

Explain the various kinds of fluctuations in a time series and show their significance by taking a few example.

(3) एक काल श्रेणी में उच्चावचनों के विभिन्न संघटक कौन-कौन होते हैं ? उपनित, संघटक के मापने के लिए उपलब्ध विधियों को समझाइऐ तथा उनके सापेक्षिक गुण-दोषों को बताइए।

What are the different components of fluctuations in a time series ? Elucidate the methods available for measuring the trend component and their relative merits and demerits.

(4) 'दीर्घकालीन प्रवत्ति' से आप क्या समझते हैं ? एक काल श्रेणी में उपनति मूल्यों को पथक करने की किन्हीं दो विधियों का विवेचना कीजिए तथा उनकी त्रुटियों को दर्शाइए।

What is meant by 'Secular Trend' ? Discuss any two methods of isolating trend values in a time series and point out their shortcoming.

(5) ''आर्थिक काल श्रेणी के सांख्यिकीय विश्लेषण का प्रमुख उद्देश्य समयावधि में समंकों में परिवर्तनों की विशेषता स्पष्ट करने वाली किन्हीं की खोज करना व उसकी माप करना है।'' इस कथन की विवेचना कीजिए तथा संक्षेप में दीर्घकालीन प्रवत्ति ज्ञात करने की विभिन्न तकनीकों का वितरण दीजिए।

"The primary purpose in the statistical analysis of economics time series is to discover and measure any regularities which characterise the movement of the data through time". Discuss this statement and describe briefly the several techniques of computing trend.

(6) एक काल श्रेणी में न्यूनतम वर्ग रीति द्वारा आप उपनति-मूल्य कैसे ज्ञात करेंगे ? एक संख्यात्मक उदाहरण लेकर अपना उत्तर स्पष्ट कीजिए।

How would you find out the trend values in a time series by the method of least squares ? Illustrate your answer by a numerical example.

(7) काल श्रेणी के संघटक क्या हैं और क्यों ? क्या प्रत्येक संघटक शक्तियों का अलग व स्वतंत्र समूह के कारण होता है? व्यापार चक्र को कैसे निकालते हैं ?

What are the components of a time series and why ? Is each component caused by a separate and independent set of forces ? How is the trade cycle is isolated ?

(8) उपनति, सामयिक परिवर्तनों तथा चक्रीय उच्चावचनों में अन्तर कीजिए। दिये गये समंकों में उपनति का माप आप किस प्रकार करेंगे ?

Distinguish between secular trend, seasonal variations and cyclical fluctuations. How would you measure secular trend in any given data ?

(9) निम्न समंकों से स्वतंत्र-हस्त वक्र रीति द्वारा दीर्घकालीन प्रवत्ति ज्ञात कीजिए।

Fit a trend line to the following data by free hand curve method.

Year	:	1990	1991	1992	1993	1994	1995	1996
Production	:	77	88	94	85	91	98	90

(10) नीचे दिए गए समंकों से अर्द्ध-माध्य रीति द्वारा दीर्घकालीन प्रवत्ति रेखा ज्ञात कीजिए।

Fit a straight line trend by method of semi-average from the following data.

(11) निम्न समंकों से अर्द्ध-माध्य रीति का प्रयोग करते हुए दीर्घकालीन प्रवत्ति ज्ञात कीजिए तथा 1999 के मूल्य का अनुमान लगाइए।

Apply the method of semi-averages to depict trend in the following data and also estimate the value of 1999.

Year	:	1991	1992	1993	1994	1995	1996	1997
Sales	:	51	54	57	55	54	58	56

(in thousands)

(12) निम्न आंकड़ों से तीन-वर्षीय चल माध्य विधि द्वारा प्रवत्ति ज्ञात कीजिए।

Find out the trend values by taking three-yearly moving averages from the following data :

वर्ष (Year)	Index	वर्ष (Year)	Index
1991	105	1996	95
1992	115	1997	85
1993	110	1998	75
1994	90	1999	60
1995	80	2000	70

आंकड़ों को ग्राफ पेपर पर भी अंकित करें।

Also plot the data on the graph paper.

(13) चार-वर्षीय चल माध्य के आधार पर उपनति की जानकारी करें एवे उन्हें ग्राफ पेपर पर प्रदर्शित करें।

Find out the value of the trend by 4-yearly moving average and plot them on a graph paper :

वर्ष (Year) :	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
मूल्य (Value) :	10	15	12	18	15	22	19	24	20	26	22	30	25

[**Ans.** 12.3, 15, 15, 18.3, 18.7, 21.7, 21, 23.5, 22.7, 26, 25.7, —]

(14) निम्न आंकड़ों से पाँच-वर्षीय चल माध्य ज्ञात करें तथा मौलिक एवं उपनति मूल्यों का ग्राफ पर अंकित करें :-

Find five-yearly moving averages from the following data and represent original and trend values on the graph :

वर्ष (Year) :	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
संख्या (Number) :	1,332	1,317	1,375	1,392	1,402	1,405	1,410	1,427	1,405	1,438	1,450	1,470

[**Ans.** —, —, 1,360.00, 1,374.60, 1,393.20, 1,407.00, 1,426.00, 1,438.00, —, —]

(15) निम्नलिखित समंकों से न्यूनतम वर्ग विधि द्वारा दीर्घकालीन प्रवत्ति मूल्य ज्ञात कीजिए :

From the following data determine the trend values by the method of least squares :-

Year :	1961	1962	1963	1964	1965	1966	1967
Profit (Rs. 000) :	57	65	63	72	69	78	82

[**Ans.** 57.97, 61.79, 65.61, 69.43, 73.25, 77.07, 80.89]

(16) निम्न समंकों से न्यूनतम वर्ग रीति द्वारा सरल रेखा का आसंजन करो तथा प्रवत्ति मूल्यों को बिन्दुरेखा पर अंकित करो। 2003 वर्ष के लिए उपनति मूल्य का अनुमान भी लगाइए :-

For straight line trend by the method of least squares, to the following data. Plot the values on a graph paper and also estimate the year for 2003.

Year :	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Size of item :	110	125	115	135	150	165	155	175	180	200

[**Ans.** 108.16, 117.68, 127.20, 136.72, 146.24, 155.76, 165.28, 174.80, 184.32, 193.84]

(17) निम्न आंकड़ों से द्वितीय घात के परवलय को आसंजित कीजिए तथा उपनति मूल्यों को परिकलित कीजिए। 1998 के लिए अनुमानित मूल्य भी ज्ञात कीजिए :

Fit a parabola of second degree to the following data and compute trend values. Also estimate the value for 1998 :

वर्ष (Year) :	1991	1992	1993	1994	1995	1996	1997
मूल्य (Price) :	13	13	22	21	54	60	83

[**Ans.** 12, 14, 20, 30, 44, 62, 84, 1998 ® 110]

(18) निम्न श्रेणी से न्यूनतम वर्ग रीति द्वारा प्रवत्ति मूल्य ज्ञात कीजिए :

Calculate trend values from the following data by applying the method or least squares :-

वर्ष (Year)	:	1973	1974	1975	1976	1977	1978	1979
बिक्री (Sales) (Rs. Crores)	:	20	23	22	25	26	29	30

Estimate the likely sales for the year 1982.

सन् 1982 के लिए बिक्री का अनुमान लगाइए।

[**Ans.** $Y_t = 25 + 1.643\,x$ (Short-cut Method) Trend Value of 1982 = Rs. 34.86 Crores]

(19) न्यूनतम वर्ग रीति द्वारा उपनति मूल्य ज्ञात कीजिए।

Find trend values by least square method :

Year	:	1985	1986	1987	1988	1989	1990	1991	1992
Production	:	102.3	101.9	105.8	112.0	114.8	118.7	124.5	102.9

[**Ans.** $Y_t = 101.86 + 1.89$ X (Direct Method and $Y_t = 110.36 + .94\,x'$
(By short cut with step deviation method)]

(20) निम्न समंकों से दीर्घकालीन प्रवत्ति रेखा बनाएं : -

Fit a trend line to the data :

Year	:	1981	1982	1983	1984	1985	1986
Production	:	7	17	12	19	22	27

[**Ans.** $Y_t = 5.12 + 3.49$ X (Direct Method)]

❑❑

अध्याय - 13

आन्तरगणन तथा बाह्यगणन (Interpolation and Extropolation)

अर्थ — आन्तरगणन वह सांख्यिकीय विधि है, जिसकी सहायता से किसी स्वतंत्र चरमूल्य (Subject Variable) के लिए आश्रित चरमूल्य (Dependent Variable) का सर्वोचित मान ज्ञात किया जाता है। यदि x तथा y दो चर-मूल्य दिये गये हों, तथा दिये गये x के लिए y का सर्वोचित मूल्य ज्ञात करना हो तो उसके लिए आन्तरगणन तकनीक का प्रयोग किया जाता है। उदाहरणार्थ, यदि भारत की जनंसख्या 1901, 1911, 1921, 1931, 1941, 1951 तथा 1961 जनगणना वर्षो की ज्ञात हो तथा 1945 या 1955 की जनसंख्या का अनुमान लगाना हो तो यह आन्तरगणन कहलाता है। यदि 1971 की जनसंख्या का अनुमान लगाना हो तो वह बाह्यगणन कहलाता है, क्योंकि 1971 दिये गये मूल्यों के बाद आता है। आन्तरगणन श्रेणी के मध्य की कड़ी का मूल्य ज्ञात करने में सहायता देता है जबकि बाह्यगणन भविष्य के पूर्वानुमान में सहायक होता है। हार्पर (W.M. Harper) के शब्दों में, ''**दो सीमान्त बिन्दुओं के मध्य बिन्दु का मूल्य ज्ञात करना 'आन्तरगणन' कहलाता है। बाह्यगणन का अर्थ दो सीमान्त बिन्दुओं के बाहर के किसी बिन्दु का मूल्य ज्ञात करना है।** आन्तरगणन तकनीक एक महत्वपूर्ण सांख्यिकीय विधि है। वर्गीकृत समंकों अथवा सतत् श्रेणी (Continuous Series) में मध्यका (Median) तथा भूयिष्ठक (Mode) का मूल्य ज्ञात करने में आन्तरगणन की सरल विधि ही अपनायी जाती है। समंकों में अज्ञात अथवा खोई हुई संख्या को ज्ञात करने में इस तकनीक से बड़ी सहायता मिलती है।

उपरोक्त दोनों विधियों में चर (Variable)-X सदैव ही स्वतंत्र होता है तथा चर-Y सदैव ही आश्रित होता है। अतः हम स्वतंत्र चर के समक्ष केवल आश्रित चर (Y) का ही अनुमान लगा सकते हैं।

आन्तरगणन एवं बाह्यगणन में अन्तर
(Difference Between Interpolation and Extrapolation)

आन्तरगणन एवं बाह्यगणन में अन्तर यह है कि आन्तरगणन विधि का प्रयोग श्रेणी के बीच के मूल्यों का अनुमान लगाने के लिए किया जाता है। जबकि बाह्यगणन की विधि श्रेणी के बाहर के मूल्य अर्थात भावी समंकों का पूर्वानुमान करने के लिए प्रयोग में लाई जाती है। इन दोनों का अन्तर निम्न उदाहरण से स्पष्ट हो जाता है :-

Year	:	1931	1941	1951	1961	1971	1981	1991
Population (in crores)	:	27.9	31.87	36.11	43.92	54.82	68.3	84.39

यदि इन उपलब्ध समंकों के आधार पर 1931 से 1991 तक के किसी भी वर्ष की (जिसकी जनसंख्या मालूम नहीं है) जनसंख्या का अनुमान लगाया जाए तो इसे आन्तर्गणन कहा जाएगा। इसके विपरीत यदि दी हुई सीमाओं 1931 से 1991 के बाहर जैसे वर्ष 2000 के लिए जनसंख्या का अनुमान लगाया जाए तो इसे बाह्यगणन कहा जाएगा।

मान्यताएँ
(Assumptions)

आन्तरगणन एवं बाह्यगणन विधियों द्वारा अज्ञात मूल्यों का अनुमान निम्नलिखित मान्यताओं पर आधारित है :-

(1) आकस्मिक उतार-चढ़ाव का न होना (Absence of any Jump or Fluctuation) : अन्तगर्णन एवं बाह्यगणन करते समय यह मान लिया जाता है कि दी हुई अवधि के समंकों में आकस्मिक उतार-चढ़ाव नहीं हुए। जैसे यदि हम 1921, 1931. 1941, 1951, 1961, 1971 तथा 1981 की जनगणना के आधार पर सन् 1977 की जनसंख्या अनुमान लगाना चाहते हैं तो हमारी मान्यता होगी के जनसंख्या के इन समंकों में अधिक उतार-चढ़ाव नहीं हुए हैं।

(2) परिवर्तनों की दर में एकरूपता (Uniformity in the Rate of Change) : आन्तरगणन एवं बाह्यगणन करते समय यह मान लिया जाता है कि समंकों में परिवर्तन सदैव नियमित एवं समान दर से होते हैं।

(3) पद श्रेणियों का पारस्परिक सम्बन्ध (Mutual Relationship of Series) : यह मान लिया जाता है कि दिए हुए समंकों के बीच एक सुनिश्चित तथा स्थाई सम्बन्ध है।

आन्तरगणन एवं बाह्यगणन का महत्त्व

(Importance of Interpolation and Extrapolation)

सांख्यिकी विश्लेषण के लिए आन्तगर्णन एवं बाह्यगणन का विशेष महत्व है, जिसके निम्नलिखित कारण हैं :-

(1) समंकों का अभाव (Lack of Data) : कभी-कभी ऐसा होता है कि भूतकाल के आंकड़े या तो उपलब्ध नहीं होते या यदि उपलब्ध होते हैं तो अपर्याप्त मात्रा में। अपर्याप्त आंकड़ों की सहायता से उचित एवं अविश्वसनीय निष्कर्ष नहीं निकाले जा सकते। ऐसी परिस्थिति में आन्तरगणन एवं बाह्यगणन का प्रयोग अपेक्षित है।

(2) समंकों का समाप्त हो जाना (If Data are Finished) : कभी-कभी ऐसा होता है कि आंकड़े किसी कारण जैसे आग लग जाने से, भूकंप एवं बाढ़ इत्यादि से नष्ट या समाप्त हो जाते हैं। ऐसी परिस्थिति में आन्तरगणन विधि द्वारा आंकड़ों का अनुमान लगाया जा सकता है।

(3) भविष्य के बारे में अनुमान (Future Estimate) : भविष्य के बारे में अनुमान लगाने के लिए बाह्यगणन विधि का प्रयोग किया जाता है। उद्योगपति, अर्थशास्त्री एवं व्यापारी के लिए ये अनुमान बहुत ही लाभकारी सिद्ध होते हैं, क्योंकि ये भविष्य के बारे में अनुमान लगाकर ही नीति निर्धारण करते हैं। जब तक हमें भावी आंकड़ों का ज्ञान नहीं होगा तब तक कुशल नीति का निर्माण सम्भव नहीं है। सरकार एवं राजनीतिज्ञ भी इन्हीं अनुमानों के आधार पर अपनी योजनाएँ बनाते हैं एवं वित्त आदि का प्रबन्ध करते हैं।

(4) तुलनात्मक अध्ययन सम्भव होना (Comparative Study is Possible) : आन्तरगणन एवं बाह्यगणन की सहायता से दो स्थानों के बारे में उपलब्ध समंकों में तुलनात्मक अध्ययन सम्भव हो जाता है। यदि दो स्थानों के समंक भिन्न-भिन्न अवधि के लिए दिए गए हों तो वे तुलना योग्य नहीं होते। ऐसी परिस्थिति में आन्तर्गणन की सहायता से तुलनात्मक अध्ययन सम्भव हो जाता है।

(5) बीच की तिथि के लिए समंक प्राप्त करना (To Find out Middle Value) : कभी-कभी समंक एकत्र किए जाने वाली तिथियों के मध्य किसी तिथि के लिए समंक प्राप्त करने की जरूरत होती है। ऐसी परिस्थिति में आन्तरगणन का प्रयोग अपेक्षित है। उदाहरणार्थ, भारत में जनसंख्या से समंक दस वर्ष बाद एकत्रित किए जाते हैं। यदि किसी मध्य के वर्ष की जनसंख्या को प्राप्त करना हो तो आन्तर्गणन की सहायता प्राप्त करनी होगी।

आन्तरगणन व बाह्यगणन की रीतियाँ

(Methods of Interpolation and Extrapolation)

आन्तरगणन की रीतियाँ निम्नलिखित हैं :-

(1) बिन्दुरेखीय रीति (Graphic Method)

(1) बीजगणितीय रीति (Algebraic Method)

बीजगणितीय रीतियाँ निम्न हैं :-

(*i*) द्विपद्-विस्तार रीति (Bionomial Expansion Method),

(*ii*) अन्वायोजन वक्र रीति (Parabolic Curve Method)

(*iii*) न्यूटन की परिमित अथवा प्रगामी अन्तर रीति (Newton's Method of Finite or Advancing Differences),

(*iv*) न्यूटस-गॉस (अग्रिम) सूत्र [Newton-Gauss (forward) Formula], (*v*) स्टर्लिंग का सूत्र (Sterling's Formula)

(*vi*) न्यूटस-गॉस (प्रष्ठगामी) सूत्र [Newton-Gauss (backward) Formula]

(*vii*) लैगरेंज की रीति (Lagrange's Method)

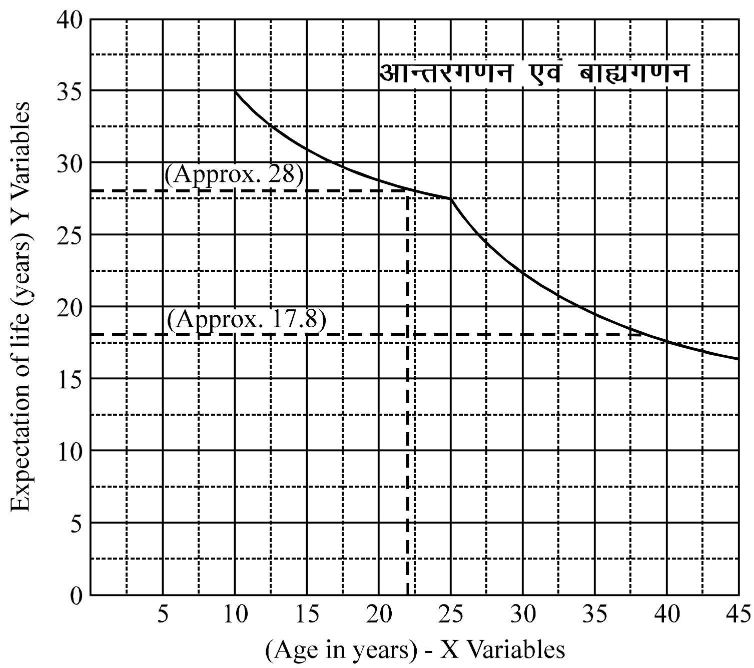
आन्तरगणन की बिन्दुरेखीय रीति
(Graphic Method of Interpolation)

नीसवेंजन (Neiswanger) के अनुसार, ''जब मध्य-स्थित मूल्यों का अनुमान लगाने के उद्देश्य से दो ज्ञात बिन्दुओं को एक रेखा द्वारा जोड़ा जाता है तो वह क्रिया आन्तर्गणन कहलाती है।'' बिन्दुरेखीय रीति आन्तरगणन की सबसे सरल विधि है, परन्तु इससे प्राप्त परिणाम अधिक विश्वसनीय नहीं होते हैं। इस रीति के अनुसार उपलब्ध समंकों को, x-variables को भुजाक्ष (x-axis) पर लेकर y-variable को कोटिअक्ष (y-axis) पर लेकर, बिन्दुरेखीय-पत्र (Graph Paper) पर प्रांकित किया जाता है। विभिन्न प्रांकित बिन्दुओं को क्रमशः मिलाकर एक सतत्-वक्र बनाया जाता है। भुजाक्ष (x-axis) पर एक बिन्दु लेकर, जिसके लिए y का मूल्य ज्ञात करना हो, कोटिअक्ष (y-axis) को समान्तर एक रेखा खींच दी जाती है। यह रेखा वक्र को जिस बिन्दु पर काटती है उस बिन्दु से भुजाक्ष (x-axis) के समान्तर रेखा खींच दी जाती है। यह रेखा कोटि-अक्ष (y-axis) को जहाँ स्पर्श करती है, उससे 'y' का मूल्य ज्ञात कर लिया जाता है।

Illustration 1 : Estimate graphically the expectation of life at age 22 and 40 from the following data :-

Age in Years	10	15	20	25	30	35
Expectation of Life in Years	35.4	32.2	29.1	26.0	23.1	20.4

Solution :



(*i*) द्विपद-विस्तार विधि
(Binomial Expansion Method)

आन्तरगणन एवं बाह्यगणन की यह विधि सरल है, यह विधि द्विपद प्रमेय (Binomial Theorem) पर आधारित है। यह विधि तभी प्रयोग में लाई जा सकती है जब निम्न दो शर्तें पूरी हो जाती हों :-

(1) श्रेणी में स्वतंत्र चल x के मूल्य समान अन्तर से बढ़ते हों।

(2) श्रेणी में x से सम्बन्धित y के मूल्य को ज्ञात करना हो।

निम्न उदाहरणों से यह जाना जा सकता है कि द्विपद-विस्तार विधि का प्रयोग होगा या नहीं।

Firm A		Firm B	
Years	**Sales (Rs.) ('000)**	**Years**	**Sales (Rs.) ('000)**
1982	10	1979	12
1984	12	1981	13
1986	12	1986	?
1988	?	1988	16
1990	18	1989	20
1992	20	1992	21
1994	22	1994	23

उपर्युक्त उदाहरण से स्पष्ट है कि 'Firm A' में 1998 की बिक्री ज्ञात करने के लिए द्विपद विस्तार विधि का प्रयोग किया जा सकता है क्योंकि यहाँ पर दोनों शर्तें पूरी हो जाती हैं परन्तु 'Firm B' को 1986 में बिक्री ज्ञात करने के लिए द्विपद-विस्तार विधि को प्रयोग में नहीं लाया जा सकता क्योंकि यहाँ पर दोनों शर्तें पूरी नहीं होती।

विधि : इस विधि से आन्तरगणन की निम्न क्रिया है :

(1) स्वतंत्र चल x के मूल्यों को क्रमानुसार $x_0, x_1, x_2, x_3, x_4, ...,$ आदि निर्देशांकों द्वारा लिखा जाता है और आश्रित श्रेणी y के मूल्यों को $y_0, y_1, y_2, y_3, y_4, ...,$ आदि निर्देशांकों द्वारा लिखा जाता है।

(2) आश्रित श्रेणी y के जितने भी मूल्य होते हैं उनका प्रमुख अन्तर (leading difference) शून्य होगा। यदि 4 मूल्य ज्ञात हों तो चतुर्थ मुख्य अन्तर शून्य होगा। (If the known values are four, the fourth leading difference will be zero)

(3) इस विधि का सूत्र $(y-1)^n = 0$ के विस्तार पर आधारित है। इस सूत्र का विस्तार इस प्रकार किया जाता है :-

$$(y-1)^n = y_n - n, y_{n-1} + \frac{n(n \quad 1)}{2!} y_{n-2} + \frac{n(n \quad 1)(n \quad 2)}{3!} y_{n-3} + ... = 0$$

इस सूत्र में n, y के ज्ञात मूल्यों की संख्या है :

संक्षेप में यह विस्तार इस प्रकार होता है —

$$(y-1)^2 = 0 : y^2 - 2^{y1} + y^0 = 0$$
$$(y-1)^3 = 0 : y^3 - 3^{y2} + 3y^1 - y^0 = 0$$
$$(y-1)^4 = 0 : y^4 - 4y^3 + 6y^2 - 4y^1 + y^0 = 0$$
$$(y-1)^5 = y^5 - 5y^4 + 10y^3 - 10y^2 + 5y^1 - y^0 = 0$$
$$(y-1)^6 = y^6 - 6y^5 + 15y^4 - 20y^3 + 15y^2 - 6y^1 + y^0 = 0$$
$$(y-1)^7 = y^7 - 7y^6 + 21y^5 - 35y^4 + 35y^3 - 21y^2 + 7y^1 - y^0 = 0$$
$$(y-1)^8 = y^8 - 8y^7 + 28y^6 - 56y^5 + 70y^4 - 56y^3 + 28y^2 - 8y^1 + y^0 = 0$$

संक्षेप में सूत्र-विस्तार की रीति

(अ) (*i*) पास्कल द्वारा दिये गये त्रिभुज (Pascal's Triangle), जो नीचे दिया गया है, से विस्तार के Coefficients ज्ञात किये जा सकते हैं।

No.	Pascals' Triangle	Sum
1	1 1	2
2	1 2 1	8
3	1 3 3 1	6
4	1 4 6 4 1	16
5	1 5 10 10 5 1	32
6	1 6 15 20 15 6 1 1	64
7	1 7 21 35 35 21 7 1	128
8	1 8 28 56 70 56 28 8 1	256
9	1 9 36 84 126 126 84 36 9 1	512
10	1 10 45 120 210 252 210 120 45 10 1	1024

उपर्युक्त द्विपद विस्तार की प्रक्रिया काफी कठिन है इसलिए निम्न प्रक्रिया द्वारा द्विपद विस्तार को हल कर लेना चाहिए।

(*i*) जितने मूल्य ज्ञात हों, उस क्रम के y को पहले लिखा जाता है जैसे — यदि 3 मूल्य ज्ञात हैं तो द्विपद विस्तार करने में y_3 से समीकरण शुरू होगा।

(*ii*) इसके बाद y का घात एक-एक कम करते जाते हैं ताकि अन्त में y_0 आ जाए। यथा —

$$y_3, y_2, y_1, y_0$$

(*iii*) प्रथम मूल्य (+) दूसरा मूल्य (–) फिर + ... क्रम से होगा। जैसे —

$$+\ y_3, -y_2, +y_1, -y_0 \qquad \frac{3}{1} y_2 \quad \frac{2\ \ 3}{2} y_1 \quad \frac{1\ \ 3}{3} y_0 \quad 0$$

(*iv*) इन पदों के गुणक ज्ञात करने के लिए जिस पद का गुणक ज्ञात करना है उसके पूर्व के y की घात को उसके गुणक से गुणा करके उसके स्थान क्रम से भाग दिया जाएगा।

सूत्र के रूप में :

$$= \frac{\text{पूर्व } y \text{ की घात} \quad \text{पूर्व पद का गणक}}{\text{पूर्व पद का स्थान क्रम}}$$

उपरोक्त उदाहरण में :

$$= y_3 -$$

$$= y_3 - 3y_2 + 3y_1 - y_0 = 0$$

Illustration 1 : The sales of a businessman is given below for different years. Estimate the sales for 1996.

एक व्यापारी की विभिन्न वर्षों की बिक्री नीचे दी गई है। वर्ष 1996 की बिक्री ज्ञात कीजिए।

Years	:	1993	1994	1995	1996	1997
Sales (Rs. in Laksh)	:	56	58	62	?	68

Solution : Estimate of sales for the year 1996.

Byexpanding : $y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$

By putting the values, we get

$$68 - 4y_3 + (6 \times 62) - (4 \times 58) + 56 = 0$$

$$68 - 4y_3 + 372 - 232 + 56 = 0$$

$$-4y_3 = -68 - 372 + 232 - 56$$
$$-4y_3 = -264$$
$$-y_3 = -66$$
$$-y_3 = 66$$

The estimate sales for the year 1996 is Rs. 66 Lakhs.

Illustration 2 : निम्नलिखित सारणी के अज्ञात मूल्य का आन्तरगणन कीजिए।

Interpolate the missing value in the following table.

X	:	1	2	3	4	5	6	7	8	9	10
Y	:	485	445	356	315	—	278	239	232	209	189

Solution : **Estimate of *y* when *x* = 5**

(X)		(Y)	
1	x_0	485	y_0
2	x_1	445	y_1
3	x_2	356	y_2
4	x_3	315	y_3
5	x_4	?	y_4
6	x_5	278	y_5
7	x_6	239	y_6
8	x_7	232	y_7
9	x_8	209	y_8
10	x_9	189	y_9

Since the known values are nine, hence the ninth leading will be zero $^{9}_{0} = 0$ or $(Y - 1)^9 = 0$

By expanding : $y_9 - 9y_8 + 36y_7 - 84y_6 + 126y_5 - 126y_4 + 84y_3 - 36y_2 + 9y_1 - y_9 = 0$

By putting the values, we get

$189 - (9 \times 2209) + (36 \times 232) - (84 \times 239) + (126 \times 278) - 126y_4 + (84 \times 315) - (36 \times 356) + (9 \times 445) - 485 = 0$

$189 - 1881 + 8352 - 20076 + 35028 + 126y_4 + 26460 - 12816 + 4005 - 485 = 0$

$$74034 - 35258 - 126y_4 = 0$$
$$126y_4 = 74034 - 35258$$
$$126y_4 = 38776$$
$$y_4 = 307.75$$

Illustration 3 : The following figures show the contribution of mining, manufacturing and small enterprises towards the national income of India from 1955-56 to 1959-60.

Year	:	1955-56	1956-57	1957-58	1958-59	1959-60
Rs. abja (100 crores)	:	18.5	20.0	21.2	21.7	23.0

Estimate the contribution of this sector for 1960-61 by any suitable algebraic method.

Solution :

Year X	1955-56	1956-57	1957-58	1958-59	1959-60	1960-61
Rs. (abja)	$18.5y_0$	$20.0y_1$	$21.2y_2$	$21.7y_3$	$23.0y_4$	$-y_5$

$(y - 1)^5 = 0$

$$(y-1)^5 = 0$$

or $y^5 - 5y^4 + 10y^3 - 10y^2 - 5y^1 - y^0 = 0$

By substituting the values we get

$$y^5 - (5 \times 23.0) + (10 \times 21.7) - (10 \times 21.2) + (5 \times 20.0) - 18.5 = 0$$

or $y^5 - 115 + 217 - 212 + 100 - 18.5 = 0$

or $y^5 = 115 - 217 + 212 - 100 + 18.5$

or $y^5 = 28.5$

दो अज्ञात मूल्य (Two Unknown Values) : यदि प्रश्न में दो अज्ञात मूल्य हों तो उनको ज्ञात करने के लिए दो समीकरणों (equations) की आवश्यकता पड़ेगी। प्रथम समीकरण को उपर्युक्त विधि के अनुसार लिखिए और दूसरा समीकरण बनाने की दो विधियाँ हैं। **प्रथम,** यह कल्पना करेंगे कि एक ही पद मूल्य अज्ञात हैं अतः अन्तिम उपलेख (Subscrit) y का समीकरण बनायेंगे। इस प्रकार दोनों ही समीकरणों (Equations) के अन्त में y_0 आएगा जैसे पद ज्ञात है और 2 पद अज्ञात है तो पहले $^6_0$ का समीकरण बनाएंगे फिर $^7_0$ का समीकरण बनाएँगे। इन दोनों समीकरणों को हल करके अज्ञात मूल्यों की गणना कर लेंगे। दूसरे समीकरण ज्ञात करने की **द्वितीय विधि** यह है कि प्रथम समीकरण के उपलेख (Subscript) में I की वद्धि करते हैं, इस प्रकार अन्त में y_0 के स्थान पर y_1 आता है। इस प्रकार दोनों समीकरण में ज्ञात मूल्यों को रखकर अज्ञात मूल्यों को ज्ञात किया जाता है।

Illustration 4 : Interpole the missing figures in the following table of rice cultivation :

Year	Acres in Millions	
1961	76.6	y_0
1962	78.7	y_1
1963	—	y_2
1964	77.7	y_3
1965	78.7	y_4
1966	—	y_5
1967	80.6	y_6
1968	77.6	y_7
1969	78.7	y_8

Solution :

Note : As two values are to be interpolated, last item will be left in one equation, so that values of both may be calculated.

$$(y-1)^8 = 0, (y-1)^7 = 0$$

$$= y^8 - 8y^7 + 28y^6 - 56y^5 + 70y^4 - 56y^3 + 28y^2 - 8y^1 + y^0 = 0$$

$$= y^7 - 7y^6 + 21y^5 - 35y^4 + 35y^3 - 21y^2 + 7y^1 - y^0 = 0$$

By substituting values, we get

$$78.7 - (8 \times 77.6) + (28 \times 80.6) - 56y^5 + (70 \times 78.7) - (56 \times 77.7) + 28y^2 - (8 \times 78.7) + 76.6 = 0$$

$$77.6 - (7 \times 80.6) + 21y^5 - (35 \times 78.7) + (35 \times 77.7) - 21y^2 + (7 \times 78.7) - 76.6 = 0$$

taking first

$$-56y^5 + 28y^2 + 7921 - 5601.6 = 0$$

or $\underline{-56y^5 + 28y^2 = -2319.5}$ ...(*i*)

taking 2nd

$$21y^5 + 21y^2 + 3348 - 3395.3 = 0$$

or $\underline{21y^5 - 21y^2 = 47.3}$ ...(*ii*)

$$84y^5 - 84y^2 = 189.2 \quad ...(ii \times 4)$$
$$-168y^5 + 84y^2 = -6958.5 \quad (i \times 3)$$
$$-84y^2 = -6769.3 \text{ (Added)}$$
$$y^5 = \frac{6769.3}{84} = 80.6$$

Substituting the value of y^5 in (ii)

$$1692.6 - 21y^2 = 47.3$$
$$-21y^2 = 47.3 - 1692.6$$
$$-21y^2 = -1645.3$$
$$y^2 = \frac{1645.3}{21} = 78.35$$

The average in 1963 = 78.35 millions

1966 = 80.6 millions

Illustration 5 : निम्नलिखित आँकड़ों के लिए वर्ष 1976 तथा 1986 के लिए उत्पादन का आँकलन कीजिए :

Estimate the production for the years 1976 and 1986 from the following data :

Year	:	1971	1976	1981	1986	1991	1996
Production (In Tonnes)	:	180	—	250	—	320	400

Solution : Estimate of production for the years 1976; and 1986.

Year		**Production (in tonnes)**	
1971	x_0	180	y_0
1976	x_1	—	y_1
1981	x_2	250	y_2
1986	x_3	—	y_3
1991	x_4	320	y_4
1996	x_5	400	y_5

Since, the known values of Y are 4, hence, the fourth leading difference will be zero. In this question, two values of Y are unknown and hence, two equations are required to determine them. These are :-

$$= y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0 \quad ...(i)$$

$$^4_1 = y_5 - 4y_4 + 6y_3 - 4y_2 + y_1 = 0 \quad ...(ii)$$

By putting the values in two equations, we get

$$320 - 4y_3 + (6 \times 250) - 4y_1 + 180 = 0 \quad ...(i)$$
$$400 - (4 \times 320) + 6y_3 - (4 \times 250) + y_1 = 0 \quad ...(ii)$$

By solving eq. (i) $\quad -4y_3 - 4y_1 = -2000$

or $\quad 4y_3 + 4y_1 = 2000 \quad ...(iii)$

By solving eq. (ii) $\quad 6y_3 - y_1 = 1880 \quad ...(iv)$

By multiplying equation (*iv*) by 4, we get

$$24y_3 + 4y_1 = 7520 \quad ...(v)$$

By deducting equation (*iii*) from (*v*), we get

$$\begin{aligned} 24y_3 + 4y_1 &= 7520 \\ 4y_3 + 4y_1 &= 7520 \\ - \quad - \quad &\quad - \\ \hline 20y_3 &= 5520 \\ y_3 &= 276 \end{aligned}$$

Putting the value of y_3 in the equation (*iii*), we get

$$\begin{aligned} (4 \times 276) + 4y_1 &= 2000 \\ 1104 + 4y_1 &= 2000 \\ 4y_1 &= 896 \\ y_1 &= 224 \end{aligned}$$

Thus, the estimated production for 1976 and 1986 is 224 and 276 tonnes respectively.

(*ii*) **अन्वायोजन वक्र रीति (Parabolic Curve Method) :** इस विधि को युगपत समीकरणों की विधि (Method of Simultaneous Equation) भी कहते हैं। इस विधि से हर दशा में आन्तरगणन किया जा सकता है। अन्वायोजन वक्र का समीकरण निम्न है :

$$y = a + bx + cx^2 + dx^3 + ex^4 \text{ ...... } nx^n$$

अन्वायोहन की घात श्रेणी के ज्ञात पदों की संख्या पर निर्भर करती है। ज्ञात पदों की संख्या से 1 कम घात का वक्र लगाया जाता है। समीकरण में a, b, c, d, e आदि अचर पद (Constant) होते हैं। इस समीकरण से सींचा गया वक्र सर्वोत्तम होता है क्योंकि वह सभी बिन्दुओं से गुजरता है।

Illustration 6 : The following values are given in a table :-

x	**:**	1	2	3	4	5
y	**:**	21600	226981	—	250047	262144

Using any suitable algebraic method find out the value of *y* for *x* = 3.

Solution : ज्ञात पदों की संख्या 4 है अतः वक्र $n - 1$ or $4 - 1 = 3$ घात का ज्ञात किया जायेगा।

$$y = a + bx + cx^2 + dx^3$$

x	dx (3)	y
1	-2	216000
2	-1	226981
3	0	$-y_0$
4	$+1$	250047
5	$+2$	262144

Let us substitue the value of X in the parabola : $Y = a + bx + cx^2 + dx^3$

when $x = -2$, then $= 216000 = a - 2b + 4c - 8d$ ...(*i*)

when $x = -1$, then $= 226981 = a - b + c - d$ ...(*ii*)

when $x = 0$, then $= y_0 = a$ ...(*iii*)

when $x = +1$, then $= 250047 = a + b + c + d$ ...(*iv*)

when $x = -2$, then $= 262144 = a + 2b + 4c + 8d$ ...(*v*)

उक्त समीकरणों की सहायता से यदि '*a*' का मान ज्ञात हो जाये तो '*y*' का आवश्यक मूल्य ज्ञात हो जायेगा।

(*ii*) तथा (*iv*) समीकरणों को जोड़ा :

$$\begin{aligned} 226981 &= a - b + c - d \\ 250047 &= a + b + c + d \\ \hline 477028 &= 2a + 2c \quad ...(vi) \end{aligned}$$

(i) तथा (v) समीकरणों को जोड़ा :

$$\begin{aligned} 216000 &= a - 2b + 4c - 8d \\ 262144 &= a + 2b + 4c + 8d \\ \hline 478144 &= 2a + 8c \end{aligned} \qquad ...(viii)$$

(vii) समीकरण में से (vi) को घटाया

$$\begin{aligned} 478144 &= 2a + 8c \\ 477028 &= 2a + 2c \\ - \quad & \quad - \quad - \\ \hline 1116 &= 6c \\ 186 &= c \end{aligned}$$

c का मान (vi) समीकरण में रखा :

$$477028 = 2a + (2 \times 186)$$

or $$477028 = 2a + 372$$

or $$477028 = 372 = 2a$$

$$476656 = 2a$$

$$238328 = a$$

238328 ही y_0 का (when $x = 3$) का मान है।

Illustration 7 : निम्न तालिका एक व्यापारिक संस्था की चार वर्षों की बिक्री बताती है। अन्वायोजन वक्र विधि का प्रयोग करके वर्ष 1995 की बिक्री का अनुमान लगाइए :

The following table gives the sales of a trading concern for four years. Estimate the sales for the year 1995 by parabolic curve method :

Year	:	1993	1994	1996	1997
Sales (Rs. in Thousand)	:	20	22	30	36

Solution : Since four values of the dependent variable are known, we will fit the parabola of 3rd order, *i.e.* $(n - 1) = 4 - 1 = $ 3rd order *i.e.*

$$\boldsymbol{y = a + bx + cx^2 + dx^3}$$

Years (X)	**Deviation from 1995 (X)**	**Y**
1993	– 2	20
1994	– 1	22
1995	0	y_0
1996	+ 1	30
1997	+ 2	36

The equation is

$$y = a + bx + cx^2 + dx^3$$

Substituting the above values of x in the equations, we get

$$20 = a - 2b + 4c - 8d \qquad ...(i)$$

$$22 = a - b + c - d \qquad ...(ii)$$

$$y_0 = a \qquad ...(iii)$$

$$30 = a + b + c + d \qquad ...(iv)$$

$$36 = a + 2b + 4c + 8d \qquad ...(v)$$

y_0 is the missing figure required from equation (*iii*), it is equal to *a*. We, therefore, wish to find out the value of *a*.

Adding equations (*i*) and (*v*), we get

$$20 = a - 2b + 4c - 8d \quad ...(i)$$
$$36 = a - 2b + 4c + 8d \quad ...(v)$$
$$56 = 2a + 8c \quad ...(vi)$$

Again adding equations (*ii*) and (*iv*), we get

$$22 = a - b + c - d \quad ...(i)$$
$$30 = a + b + 4c + d \quad ...(iv)$$
$$52 = 2a + 2c \quad ...(vii)$$

Multiplying equation (*vii*) by 4 and subtracting equation (*vi*) from it,

We get,

$$208 = 8a + 8c$$
$$56 = 2a + 8c$$
$$152 = 6a$$
$$a = 25.33$$

Therefore, the estimate sales for 1995 is 25.33 (thousand)

यद्यपि अन्तर्गणन की इस विधि का प्रयोग सभी दशाओं में किया जा सकता है फिर भी इस विधि का प्रयोग सीमित है, क्योंकि श्रेणी में संख्याएं अधिक हों तो बहुत सारे समीकरणों को हल करना पड़ता है जो बहुत कठिन कार्य है तथा समय भी अधिक लगता है। अतः इस विधि का प्रयोग उस स्थिति में ही करना चाहिए जहाँ पदों की संख्या कम हो।

(*iii*) **न्यूटन की परिमित अथवा प्रमामी अन्तर रीति (Newton's Method of Finite or Advancing Differences) :** न्यूटन की परिमित प्रमाकी अन्तर रीति का प्रयोग उस दशा में किया जाता है जब *x*-variables में समान अन्तर होता है। श्रेणी के प्रारम्भ में आन्तर्गणन करने के लिए यह रीति सर्वोच्च है। इस रीति को परिमित अथवा प्रगामी अन्तर रीति इसलिए कहा जाता है क्योंकि *y* के मूल्यों में पारस्परिक अन्तर निकालने के पश्चात अन्तरों के भी पारस्परिक अन्तर उस स्थिति तक निकाले जाते हैं जब तक केवल एक अन्तर शेष न रह जाये। अन्तर ज्ञात करते समय बीजगणित चिन्हों (+ तथा –) का ध्यान रखा जाता है। अन्तरों को निकालने का तरीका इस प्रकार है :-

परिमित अथवा प्रगामी अन्तर निकालने की विधि-तालिका

First x	Second y	Third Differences D^1	Fourth Differences D^2	Fifth Differences D^3	Differences D^4	Differences D^5
x_0	y_0					
x_1	y_1	$y_1 - y_0 = {}^{1}_{0}$				
x_2	y_2	$y_2 - y_1 = {}^{1}_{1}$	${}^{1}_{1} - {}^{1}_{0} = {}^{2}_{0}$			
x_3	y_3	$y_3 - y_2 = {}^{1}_{2}$	${}^{1}_{2} - {}^{1}_{1} = {}^{2}_{1}$	${}^{2}_{1} - {}^{2}_{0} = {}^{3}_{0}$		
x_4	y_4	$y_4 - y_3 = {}^{1}_{3}$	${}^{1}_{3} - {}^{1}_{2} = {}^{2}_{2}$	${}^{2}_{2} - {}^{2}_{1} = {}^{3}_{1}$	${}^{3}_{1} - {}^{3}_{0} = {}^{4}_{0}$	
x_5	y_5	$y_5 - y_4 = {}^{1}_{4}$	${}^{1}_{4} - {}^{1}_{3} = {}^{2}_{3}$	${}^{2}_{3} - {}^{2}_{2} = {}^{3}_{2}$	${}^{3}_{2} - {}^{3}_{1} = {}^{4}_{1}$	${}^{4}_{1} - {}^{4}_{0} = {}^{5}_{0}$

न्यूटन के निम्न सूत्र से इस रीति द्वारा आन्तरगणन किया जाता है :-

$$y_x \;=\; y \quad x\ {}^{1}_{0} \quad \frac{x(x\quad 1)}{1\quad 2}\ {}^{2}_{0} \quad \frac{x(x\quad 1)(x\quad 2)}{1\quad 2\quad 3}\ {}^{3}_{0}$$

इस सूत्र में संकेताक्षरों के अर्थ इस प्रकार हैं :-

$y_x = y$ का वह मूल्य जिसका आन्तरगणन करना है। (Figure to be interpolated)

$y_0 =$ प्रारम्भिक वर्ष का मूल्य (The value of the year of origin)

${}^{1}_{0}$, ${}^{2}_{0}$, ${}^{3}_{0}$, ${}^{4}_{0}$, ${}^{5}_{0}$ आदि 'O' कम के प्रथम, द्वितीय, ततीय आदि अन्तर

(1st, 2nd, 3rd and so on difference of the zero order)

$$x \;=\; \frac{\text{Year of interpolation} \quad \text{Year of origin}}{\text{Time distance between adjoining years}}$$

Illustration 8 : The following table shows the expectation of life at different ages. Find the expectation of life at age 16.

Age in Years	10	15	20	25	30	35
Expectation of life in years	35	30	29	27	22	20

Solution :

$\frac{x(x\quad 1)(x\quad 2)(x\quad 3)}{5\quad 1\quad 2\quad 3\quad 4}\ {}^{4}_{0} \quad \frac{x(x\quad 1)(x\quad 2)(x\quad 3)(x\quad 4)}{1\quad 2\quad 3\quad 4\quad 5}\ {}^{5}_{0}\ldots$

Age in years x		Expectation of life in years y		Time or Advancing Difference D[1]		D[2]		D[3]		D[4]		D[5]	
10	x_0	35	y_0										
15	x_1	30	y_1	–5									
20	x_2	29	y_2	– 1	${}^{1}_{1}$	+ 4	${}^{2}_{0}$						
25	x_3	27	y_3	–2	${}^{1}_{2}$	– 1	${}^{2}_{1}$	– 5	${}^{3}_{0}$				
30	x_4	22	y_4	–5	${}^{1}_{3}$	– 3	${}^{2}_{2}$	– 2	${}^{3}_{1}$	+ 3	${}^{4}_{0}$		
35	x_5	20	y_5	–2	${}^{1}_{4}$	+3	${}^{2}_{3}$	+ 6	${}^{3}_{2}$	+ 8	${}^{4}_{1}$	+5	${}^{5}_{0}$

$$x \;=\; \frac{\text{Year of interpolation} \quad \text{Year of origin}}{\text{Difference between adjoining values of } x} = \frac{16 \quad 10}{5} = \quad 1.2$$

The formula is :

$$y_x \;=\; y \quad x\ {}^{1}_{0} \quad \frac{x(x\quad 1)}{1\quad 2}\ {}^{2}_{0} \quad \frac{x(x\quad 1)(x\quad 2)}{1\quad 2\quad 3}\ {}^{3}_{0}$$

By substituting the values, we get

$$y_x = 35 \quad (1.2 \quad 5) \quad \frac{1.2\,(1.2 \quad 1)}{2} \quad 4 \quad \frac{1.2\,(1.2 \quad 1)\,(1.2 \quad 2)}{6} \quad 5$$

$$\frac{1.2\,(1.2 \quad 1)\,(1.2 \quad 2)\,(1.2 \quad 3)}{24} \quad 3 \qquad \frac{1.2\,(1.2 \quad 1)\,(1.2 \quad 2)\,(1.2 \quad 3)\,(}{120}$$

$$y_x = 35 - 6 + .48 + .16 + .0432 - .04032$$
$$= 29.64288 \text{ or } 29.6 \text{ Years}$$

आवत्ति -वितरण (Frequency Distribution) में आन्तरगणन-आवत्ति वितरण में, आवत्तियों को संचयी (Cumulating the Frequencies) बनाकर आन्तरगणन किया जाता है। शेष क्रिया पूर्ववत् ही की जाती है।

Illustration 9 : From the following table, find the number of students who obtained less than 45 marks :-

Marks	:	30-40	40-50	50-60	60-70	70-80
No. of Students	:	31	42	51	35	31

Solution :

Marks		**No. of Students**		**Finite or Advancing Differences**							
x		***y***		D^1		D^2		D^3		D^4	
Below 40	x_0	31	y_0								
Below 50	x_1	73	y_1	+ 42	$_1^1$						
Below 60	x_2	124	y_2	+ 51	$_1^1$	+ 9	$_0^2$				
Below 70	x_3	159	y_3	+ 35	$_2^1$	– 16	$_1^2$	– 25	$_0^3$		
Below 80	x_4	190	y_4	+ 31	$_3^1$	– 4	$_2^2$	+ 12	$_1^3$	+ 37	$_0^4$

$$x = \frac{x \text{ to be interpolated} \quad x \text{ at origin}}{\text{Time Difference}} = \frac{45 \quad 40}{10} \quad .5$$

$$y_x = y \quad x \;{}_0^1 \quad \frac{x\,(x \quad 1)}{1 \quad 2}\;{}_0^2 \quad \frac{x\,(x \quad 1)\,(x \quad 2)}{1 \quad 2 \quad 3}\;{}_0^3 \quad \frac{x\,(x \quad 1)\,(x \quad 2)\,(x \quad 3)}{1 \quad 2 \quad 3 \quad 4}\;{}_0^4$$

$$= 31 \quad (.5 \quad 42) \quad \frac{.5\,(5 \quad 1)}{2} \quad 9 \quad \frac{5\,(.5 \quad 1)\,(.5 \quad 2)}{6} \quad 25 \quad \frac{.5\,(.5 \quad 1)\,(.5 \quad 5)\,(.5 \quad 3)}{24} \quad 37$$

$$= 31 + 21 - 1.125 - 1.5625 - 1.4453$$
$$= 47.8672 \text{ or } 48 \text{ students.}$$

Illustration 10 : Estimate the probable number of persons earning between Rs. 40 and 50 and Rs. 30 and Rs. 50.

40 रू. से 50 तथा 30 रू. से 50 रू. अर्जित करने वाले वयक्तियों की सम्भावित संख्या ज्ञात कीजिए :-

Income (in Rs.)	No. of persons
0-20	120
0-40	265
0-60	465
0-80	715
0-100	865

Solution :

Income (in Rs.)	No. of Persons
(X)	**(Y)**
Below 20	120
Below 40	265
Below 60	465
Below 80	715
Below 100	865

In this question, we have to interpolate No. of persons for Income below Rs. 30 and Rs. 50 separately by using Newton's method :—

Table Showing Differences

Income (X)		No. of Persons (Y)		Differences D^1		D^2		D^3		D^4	
Below 20	x_0	120	y_0								
Below 40	x_1	265	y_1	145							
Below 60	x_2	465	y_2	200	${}^{1}_{1}$	55	${}^{2}_{0}$				
Below 80	x_3	715	y_3	250	${}^{1}_{2}$	50	${}^{2}_{1}$	– 5	${}^{3}_{0}$		
Below 100	x_4	965	y_4	150	${}^{1}_{3}$	– 100	${}^{2}_{2}$	– 150	${}^{3}_{1}$	+ 145	${}^{4}_{0}$

(*i*) **Interpolation of No. of persons earning below Rs. 30.**

$$x = \frac{\text{Item to be Interpolated} \quad \text{Item at the origin}}{\text{Difference between two adjoining items}}$$

$$=$$

Newton's Formula

$$y_x = y \quad x \; {}^{1}_{0} \quad \frac{x(x \quad 1)}{2!} \; {}^{2}_{0} \quad \frac{x(x \quad 1)(x \quad 2)}{3!} \; {}^{3}_{0} \quad \frac{x(x \quad 1)(x \quad 2)(x \quad 3)}{4!} \; {}^{4}_{0}$$

By putting the values, we get

$$y_x = 120 + 0.5\,(145) + \frac{0.5\,(0.5 \quad 1)}{2 \quad 1} \times (55) + \qquad \times (-5) +$$

$$\times (-145)$$

$$= 120 + 72.5 - 6.875 - 0.3125 + 5.664$$

$$= 190.976 \text{ or } 191$$

(*ii*) Interpolation of No. of persons earning below Rs. 50.

$$x = \qquad\qquad = 1.5$$

Newton's Formula

$$y_x = y \quad x \;{}^{1}_{0} \quad \frac{x(x \quad 1)}{2!} \;{}^{2}_{0} \quad \frac{x(x \quad 1)(x \quad 2)}{3!} \;{}^{3}_{0} \quad \frac{x(x \quad 1)(x \quad 2)(x \quad 3)}{4!} \;{}^{4}_{0}$$

By putting the values, we get

$$y_x = 120 + 1.5\,(145) + \frac{1.5(1.5 \quad 1)}{2 \quad 1} \times (55) + \qquad\qquad \times (-5)$$

$$+ \qquad\qquad \times (-145)$$

$$= 120 + 217.5 + 20.625 + 0.3125 - 3.3984$$

$$= 355.039 \text{ or } 355$$

Therefore, (*i*) the probable number of persons earning beween

Rs. 40 and 50 = 355 – 265 = 90 [below 40 = 265 (given)]

(*ii*) the probable number of persons earning between

Rs. 30 and 50 = 355 – 191 = 164 $\frac{1.5(1.5 \quad 1)(1.5 \quad 2)(1.5 \quad 3)}{4 \quad 3 \quad 2 \quad 1}$ $\frac{50 \quad 40}{20}$ $\frac{30 \quad 20}{20}$ $\frac{(x \quad 1)\,(x)\,(x \quad 1)}{3!}\;{}^{3}_{y\ 1}$ $\frac{(x \quad 1)\,(x)\,(x}{4!}$

(*iii*) **न्यूटन-गॉस अग्रगामी विधि (Newton-Gauss Forward Method) :** इस विधि का प्रयोग उस स्थिति में किया जाता है जब स्वतंत्र चर-मूल्य समान अन्तर वाले हों और ज्ञात की जाने वाली संख्या श्रेणी कें मध्य में हो। इस विधि का निम्न सूत्र हैं :-

$$y_x =$$

इस विधि में x के अन्तर्गणन-पद से तुरन्त पिछले पद को x_0 और इससे पहले के दो पदों के क्रमानुसार x_{-1}, x_{-2} तथा इसके बाद के पदों को क्रमानुसार x_1, x_2, ... आदि द्वारा व्यक्त करते हैं। इन पदों के सापेक्ष चर y के मूल्यों को y_0, y_{-1}, y_{-2}, x_0, y_1, y_2, ... आदि से व्यक्त करते हैं। न्यूटन की प्रगामी विधि की भाँति इसमें भी अन्तर-सारणी (Difference Table) की रचना की जाती है। अन्तरों के संकेत चिन्ह ${}^{1}{}_{y0}$, ${}^{2}{}_{y\ 1}$, ${}^{3}{}_{y\ 1}$, ${}^{4}{}_{y\ 2}$ आदि y के चिन्हों के अनुरूप होते हैं।

$$x = \frac{\text{Item to be interpolated} \quad \text{Preceding Item}}{\text{Difference between two adjoining items}}$$

इस विधि द्वारा ज्ञात किया गया मूल्य न्यूटन की प्रगामी अन्तर विधि के परिणाम के बराबर आता है।

Illustration 11 : Estimate the number of living at the age 13 from the following position of a life table :-

Age in years	10	12	14	16
Number of living	10,000	99,223	98,540	97,843

Solution :

Age in Years		**No. of Living**		**Differences**					
				D^1		D^2		D^3	
10	$x-1$	1,00,000	$y-1$						
12	$x\ 0$	99,223	$y\ 0$	– 777	$D^1 y-1$				
14	$x+1$	98,540	$y+1$	– 683	$D^1 y\ 0$	+ 94	$D^2 y-1$		
14	$x+2$	97,843	$y+2$	– 697	$D^1 y+1$	– 14	$D^2 y\ 0$	– 108	$D^3 y-1$

(origin is taken just above the year for which we have to interpole)

$$x = \qquad = .5$$

$$y_x = y_0 + x\,D^1 y_0 + \frac{x(x-1)}{1\cdot 2} D^2 y_{-1} + \frac{(x+1)(x)(x-1)}{1\cdot 2\cdot 3} D^3 y_{-1}$$

$$= 99223 + (.5 \times -683) + \frac{.5(.5-1)}{2} \times 94 + \qquad \times -108$$

$$= 99223 - 341.5 - 11.75 + 6.75 = 98{,}876.5 \text{ or } 98{,}877$$

The number of living at the age 13 is 98877.

Illustration 12 : Interpolate in expectation of life at the age of 25.

25 वर्ष की आयु के समय जीवन की प्रत्याशा को आन्तर्गणन विधि से ज्ञात करें :-

आयु वर्षों में (Age in Years)	:	10	20	30	40
जीवन की प्रत्याक्षा [Expectation of life (yr.)]	:	26	22	18	14

Solution : Estimation of life expectation at the age of 25.

Age (in Years)		**Expectation of life (in years)**		**Differences**					
				First	D^1	**Second**	D^2	**Third**	D^3
10	$x-1$	26	$y-1$						
20	x_0	22	y_0	– 4	$D^1 - 1$				
30	$x+1$	18	$y+1$	– 4	$D^1 y_0$	0	$D^2 y-1$	0	$D^2 y-1$
40	$x+2$	14	$y+2$	– 4	$D^1 + 1$	0	$D^2 y_0$		

Substituting the values of the above formula, we get

$$y_x =$$

$$x = \frac{\text{Item of interpolation} \quad \text{Preceding Item}}{\text{Class Intercept}}$$

$$= \qquad = \qquad = 0.5$$

Substituting the values in the above formula, we get

$$y_x = 22 + (0.5 \times -4) + \frac{0.5(0.5 \quad 1)}{2 \quad 1} \quad 0 \quad \frac{0.5(0.5 \quad 1)(0.5 \quad 1)}{3 \quad 2 \quad 1} \quad 0$$

$$y_x = 22 - 2 = 20$$

Hence, expectation of life at the age of 25 is 20 years.

(*v*) **स्ट्रलिंग का सूत्र (Sterling's Formula) :** स्ट्रलिंग का सूत्र न्यूटन-गॉस (अग्रिम) सूत्र की भाँति ही उस दशा में प्रयुक्त किया जाता है जबकि (*i*) *x*-variable समान अन्तर पर हो तथा (*ii*) श्रेणी के मध्य में आन्तगरणन करना हो। उनका सूत्र इस प्रकार है :-

$$y_x = \qquad \frac{x^2}{24}\left(x^2 \quad 1^2\right) \; {}^4y \quad 2$$

Illustration 13 : Using the Sterling's Formula interpolate the value of *y* when *x* = 15

x	:	10	12	14	16	18	20
y	:	50	60	75	95	120	150

Solution :

x		*y*		Differences					
				D^1		D^2		D^3	
10	$x-2$	50	$y-2$						
12	$x-1$	60	$y-1$	10	D[illegible]				
14	$x \quad 0$	75	$y \quad 0$	15	D$^1y-1$	5	D$^2y-2$		
16	$x+1$	95	$y+1$	20	D$^1y \quad 0$	5	D$^2y-1$	0	D$^3y-1$
18	$x+2$	120	$y+2$	25	D$^1y+1$	5	D$^2y \quad 0$	0	D$^3y-1$
20	$x+3$	150	$y+3$	30	D$^1y+2$	5	D$^2y+1$	0	D$^3y \quad 0$

By substituting the values, we get

$$y_x = \frac{15 \quad 14}{2} \quad \frac{1}{2} \quad .5$$

$$y_x = y_0 \quad x\frac{{}^1y \quad 1 \quad {}^1y0}{2} \quad \frac{x^2}{2} \; {}^2_y \quad 1 \quad \frac{x\,(x^2 \quad 1^2)}{6} \quad \frac{{}^3y \quad 2 \quad {}^3y \quad 1}{2}$$

By substituting the values, we get

$$y_x = 75 + .5\,\frac{15 \quad 20}{2} \quad \frac{.25}{2} \times 5 \; +$$

$$= 75 + 8.75 + .625 + 0$$

$$= 84.375$$

The value of *y* when *x* = 84.375

Illustration 14 : Interpolate y_{35} by using Sterling formula from the following data :

स्टर्लिंग विधि से आन्तरगणन द्वारा निम्न समंकों से y_{35} को ज्ञात कीजिए।

$$y_{20} = 512, \; y_{30} = 439, \; y_{40} = 346, \; y_{50} = 243$$

If y_x represents the number of person living at the age of x in a life table.

यदि y_x जीवन तालिका में x आयु वाले व्यक्तियों को प्रस्तुत करता है।

Solution :

X	Y	Differences		
		D^1	D^2	D^3
$20\ x-1$	$512\ y-1$			
$30\ x_0$	$439\ y_0$	-73 D$^1y-1$	-20 D$^2y-1$	
$40\ x+1$	$346\ y+1$	-93 D$^1y_0$	-10 D$^2y_0$	10 D$^3y-1$
$50\ x+2$	$243\ y+2$	-103 D$^1y+1$		

$$y_x =$$

$$y_x = \frac{35 \quad 30}{30 \quad 20}$$

$$= \frac{5}{10}$$

$$= 0.5$$

$$y_{35} = 439 + 0.5\left(\frac{73 \quad (\ 93)}{2}\right) + \qquad \times -20$$

$$= 439 - 41.5 - \qquad \times 20$$

$$= 439 - 41.5 - 2.5 = 439 - 44$$

$$= 395$$

Hence, the value of $y_{35} = 395$.

(*vi*) **न्यूटन-गॉस (प्रष्ठगामी) सूत्र (Newton Gauss Backward Formula) :** यह सूत्र उस दशा में प्रयुक्त होता है जबकि (*i*) x-variable समान अन्तर पर हो तथा (*ii*) श्रेणी के अन्तिम भाग के किसी मूल्य का आन्तरगणन करना हो। इस सूत्र में x_0 आन्तरगणन करने वाले मूल्य से आगे वाले पद को माना जाता है। यह सूत्र इस प्रकार है :-

$$y_x =$$

Illustration 15 : Estimate the value of y when $x = 23$ from the following data :-

x	:	5	10	15	20	25	303
y	:	25	32	40	47	55	64

Solution :

x		*y*		Differences							
				D^1		D^2		D^3		D^4	
5	$x-4$	25	$y-4$	7	Dy^1-4						
10	$x-3$	32	$y-3$	8	Dy^1-3	1	Dy^2-4	-2	Dy^3-4		
15	$x-2$	40	$y-2$	7	Dy^1-1	-1	Dy^2-3	2	Dy^3-3	4	Dy^4-4
20	$x-1$	47	$y-1$	8	Dy^1-1	1	Dy^2-2	0	Dy^3-2	-2	Dy^4-3
25	$x-0$	55	$y \quad 0$	9	Dy^1-0	1	Dy^2-1				
30	$x+1$	64	$y+1$								

$$x = \frac{35 \quad 23}{25 \quad 20} = \frac{2}{5} = .4$$

(*vii*) **लेगरेंज की नीति (Lgrange's Method) :** लेगरेंज एक प्रसिद्ध फ्रेन्च गणितज्ञ थे। इनका नाम दो गणितीय क्रियाओं 'LAG' तथा 'RANGE' के नाम से मिल कर बनता है। इन्होंने आन्तरगणन का एक सूत्र प्रस्तुत किया है। यह सूत्र उस दशा में प्रयुक्त किया जाता है जबकि *x*-variable के मध्य असमान अन्तर होते हैं। इनका सूत्र इस प्रकार है :-

$$y_x = y_0 \frac{(x \quad x_1)(x \quad x_2)(x \quad x_3)(x \quad x_4).......(x \quad x_n)}{(x_0 \quad x_1)(x_0 \quad x_2)(x_0 \quad x_3)(x_0 \quad x_4).......(x_0 \quad x_n)} +$$

$$y_1 \frac{(x \quad x_0)(x \quad x_2)(x \quad x_3)(x \quad x_4).......(x \quad x_n)}{(x_1 \quad x_0)(x_1 \quad x_2)(x_1 \quad x_3)(x_1 \quad x_4).......(x_1 \quad x_n)} +$$

$$y_2 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_2)(x \quad x_3).......(x \quad x_n)}{(x_2 \quad x_0)(x_2 \quad x_2)(x_1 \quad x_3)(x_2 \quad x_4).......(x_2 \quad x_n)} +$$

$$y_3 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_2)(x \quad x_3).......(x \quad x_n)}{(x_3 \quad x_0)(x_3 \quad x_2)(x_3 \quad x_3)(x_3 \quad x_4).......(x_3 \quad x_n)} +$$

$$y_4 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_2)(x \quad x_3).......(x \quad x_n)}{(x_4 \quad x_0)(x_4 \quad x_2)(x_4 \quad x_3)(x_4 \quad x_4).......(x_4 \quad x_n)}$$

where *x* stands for the figure for which interpolation is to be done.

x_0, x_1, x_2, x_3, x_4 and y_0, y_1, y_2, y_3, y_4 etc. are the variable of *x* and *y* series respectively.

Illustration 16 : The following table gives the number of income tax assesees in the U.P.

Income not exceeding	No. of assessess
Rs. 2500	7166
Rs. 3000	10576
Rs. 5000	17200
Rs. 7500	20505
Rs. 10000	21925

Estimate the number of assessess with income not exceeding Rs. 400.

Solution :

x		y	
2500	x_0	7166	y_0
3000	x_1	10576	y_2
5000	x_2	17200	y_3
7500	x_3	20505	y_4
10000	x_4	21975	y_5

$x = 4000$

Illustration 17 : Find the value of y when $x = 6$.

$x = 6$ के लिए y का मूल्य ज्ञात कीजिए :-

X	**:**	3	7	9	10
Y	**:**	168	120	72	63

Solution :

X		Y	
3	x_0	168	y_0
7	x_1	120	y_1
9	x_2	72	y_2
10	x_3	63	y_3

By using the Lagrange's formula :

$$y_x = y_0 \frac{(x \quad x_1)(x \quad x_2)(x \quad x_3)}{(x_0 \quad x_1)(x_0 \quad x_2)(x_0 \quad x_3)} + y_1 \frac{(x \quad x_0)(x \quad x_2)(x \quad x_3)}{(x_1 \quad x_0)(x_1 \quad x_2)(x_1 \quad x_3)}$$

$$+ y_2 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_3)}{(x_2 \quad x_0)(x_2 \quad x_1)(x_2 \quad x_3)} + y_3 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_2)}{(x_3 \quad x_0)(x_3 \quad x_1)(x_3 \quad x_2)}$$

Here, $x = 6$

By putting the values, we get

$$y_x = 168 \frac{(6 \quad 7)(6 \quad 9)(6 \quad 10)}{(3 \quad 7)(3 \quad 9)(3 \quad 10)} + 120 \frac{(6 \quad 3)(6 \quad 9)(6 \quad 10)}{(\ \ 3)(7 \quad 9)(7 \quad 10)}$$

$$+ 72 \frac{(6 \quad 3)(6 \quad 7)(6 \quad 10)}{(9 \quad 3)(9 \quad 7)(9 \quad 10)} + 63 \frac{(6 \quad 3)(6 \quad 7)(6 \quad 9)}{(10 \quad 3)(10 \quad 7)(10 \quad 9)}$$

$$= 168 \frac{(\ \ 1)(\ \ 3)(\ \ 4)}{(\ \ 4)(\ \ 6)(\ \ 7)} + 120 \frac{(3)(\ \ 3)(\ \ 4)}{(4)(\ \ 2)(\ \ 3)} + 72 \frac{(3)(\ \ 1)(\ \ 4)}{(6)(2)(\ \ 1)} + 63 \frac{(3)(\ \ 1)(\ \ 3)}{(7)(3)(1)}$$

$$= 12 + 180 - 72 + 27$$

$$= 147$$

Illustration 18 : Find out the number of students who secured more than 20 marks but not more than 30 marks from the following data :

निम्न समंकों के आधार पर उन विद्यार्थियों की संख्या ज्ञात करो जो 20 से अधिक लेकिन 30 से अधिक अंक प्राप्त न करें।

Marks	:	Below 10	10-20	20-40	40-50
No. of Students	:	12	18	25	5

Solution :

Marks (X)		No. of Students (Y)	
Below 10	x_0	12	y_0
Below 20	x_1	30	y_1
Below 40	x_2	55	y_2
Below 50	x_3	60	y_3

By using the Lagrange's formula :

$$y_x = \qquad + y_1 \frac{(x \quad x_0)(x \quad x_2)(x \quad x_3)}{(x_1 \quad x_0)(x_1 \quad x_2)(x_1 \quad x_3)}$$

$$+ y_2 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_3)}{(x_2 \quad x_0)(x_2 \quad x_1)(x_2 \quad x_3)} + y_3 \frac{(x \quad x_0)(x \quad x_1)(x \quad x_2)}{(x_3 \quad x_0)(x_3 \quad x_1)(x_3 \quad x_2)}$$

Here, $x = 30$

By putting the values, we get

$$y_0 \frac{(x \quad x_1)(x \quad x_2)(x \quad x_3)}{(x_0 \quad x_1)(x_0 \quad x_2)(x_0 \quad x_3)}$$

$$y_x = 12 \frac{(30 \quad 20)(30 \quad 40)(30 \quad 50)}{(10 \quad 20)(10 \quad 40)(10 \quad 50)} + 30 \frac{(30 \quad 10)(30 \quad 40)(30 \quad 50)}{(20 \quad 10)(20 \quad 40)(20 \quad 50)}$$

$$+ 55 \frac{(30 \quad 10)(30 \quad 20)(30 \quad 50)}{(40 \quad 10)(40 \quad 20)(40 \quad 50)} + 63 \frac{(30 \quad 10)(30 \quad 20)(30 \quad 40)}{(50 \quad 10)(50 \quad 20)(50 \quad 40)}$$

$$= 12 \frac{(10)(\ \ 10)(\ \ 20)}{(\ \ 10)(\ \ 30)(\ \ 40)} + 30 \frac{(20)(\ \ 10)(\ \ 20)}{(10)(\ \ 20)(\ \ 30)} + 55 \frac{(20)(10)(\ \ 20)}{(30)(20)(\ \ 10)}$$

$$+ 60 \frac{(20)(10)(\ \ 10)}{(40)(30)(10)}$$

$$y_x = -2 + 20 + 36.67 - 10$$

$$y_x = 44.67 \text{ or } 45 \text{ students}$$

Number of students secured below 30 marks = 45 (Interpolated)

Number of students secured below 20 marks = 30 (given)

Therefore, number of student secured marks between

20 – 30 is (45 – 30) = 15 students

(vii) **न्यूटन की विभाजित अन्तर विधि (Newton's Method for Divided Differences) :** इस विधि का प्रयोग उस स्थिति में किया जाता है जब स्वतंत्र चर-मूल्य समान अन्तर वाले न हों। इस विधि के अन्तर्गत विभाजित अन्तर सारणी (Table of Divided Difference) बनाई जाती है तथा निम्न सूत्र का प्रयोग किया जाता है।

$$y_x = y_0 + (x - x_0) \qquad + (x - x_0)(x - x_1) \ {}^{2}_{0} + (x - x_0)(x - x_1)(x - x_2) \ {}^{3}_{0} + ...$$

where, y_x is the value to be interpolated and x is the value of X variabe for which the value y is to be interpolated. Table for the calculation of Divided Differences is as under :

Table Showing Leading Differences

X	Y	Differences					
		D^1		D^2		D^3	
x_0	y_0	$\frac{y_1 \quad y_0}{x_1 \quad x_0}$	$^1_0$				
x_1	y_1			$\frac{^1_1 \quad ^1_0}{x_2 \quad x_0}$	$^2_0$		
		$\frac{y_2 \quad y_1}{x_3 \quad x_2}$	$^1_1$			$\frac{^2_1 \quad ^2_0}{x_3 \quad x_0}$	$^3_0$
x_2	y_2			$\frac{^1_1 \quad ^1_1}{x_3 \quad x_1}$	$^2_1$		
		$\frac{y_3 \quad y_2}{x_3 \quad x_2}$	$^1_2$				
x_3	y_3						

इस विधि द्वारा ज्ञात किया गया मूल्य लाग्रैंज विधि के परिणाम के बराबर आता है।

Illustration 19 : न्यूटन की विभाजित अन्तर विधि द्वारा y का मूल्य ज्ञात करें जबकि $x = 12$ है।
Using Divided Difference Formula estimate the value of y when $x = 12$

X	:	5	10	20	$^3_0$ 25
Y	:	120	90	70	55

Solution : विभाजित अन्तर तालिका (Divided Difference Table)

X	Y	Didvided Differences							
		First			Second			Third	
5 x_0	120 y_0	$\frac{90 \quad 120}{10 \quad 5}$	6	$^1_0$					
					$\frac{2 \quad (\ 6)}{20 \quad 5}$	0.267	$^2_0$		
10 x_1	90 y_1	$\frac{70 \quad 90}{20 \quad 10}$	2	$^1_1$				$\frac{0.067 \quad (0.267)}{25 \quad 5} = -0.0167$	
					$\frac{3 \quad (\ 2)}{25 \quad 10}$	0.067	$^2_1$		
20 x_2	70 y_2	$\frac{55 \quad 70}{25 \quad 20}$	3	$^1_2$					
25 x_3	55 y_3								

$$y_x = y_0 + (x - x_0) \ ^1_0 + (x - x_0)(x - x_1) \ ^2_0 + (x - x_0)(x - x_1)(x - x_2) \ ^3_0$$

$= 120 - (12 - 5) \times -6 + (12 - 5)(12 - 10) \times .267 + (12 - 5)(12 - 10)(12 - 20) \times -0.0167$

$= 120 - 42 + 3.738 + 1.8704$

$= 125.6084 - 42 = 83.6084$

Hence, the value of y when x 12 is 83.16 (Approx.)

सारांश

- आन्तरगणन तथा बाह्यगणन वो विधियाँ हैं जिनके प्रयोग से किसी स्वतंत्र चरमूल्य के लिए आश्रित चर मूल्य के मान का अनुमान लगाया जाता है।
- श्रेणी के मूल्यों के बीच के मूल्यों का अनुमान लगाने की विधि को आन्तरगणन कहते हैं तथा श्रेणी के बाहर के मूल्यों का अनुमान लगाने की विधि को बाह्यगणन कहते हैं।
- यदि भूतकाल के आँकड़े उपलब्ध नहीं हों या आग, बाढ़ एवं भूकम्प इत्यादि के कारण नष्ट हो गये हों तो उन मूल्यों का अनुमान इन विधियों से निकाला जा सकता है। इस तरह से ये भविष्य के बारे में अनुमान लगाने में भी सहायक सिद्ध होते हैं।
- आन्तरगणन व बाह्यगणन की रीतियों को दो भागों में विभाजित किया जा सकता है :-

  (*i*) बिन्दु रेखीय रीति तथा

  (*ii*) बीजगणितीय रीति। बीजगणीतीय रीति A को भी अनेक रीतियों में विभाजित किया गया है।

प्रश्नावली

(Exercise)

(1) आन्तरगणन एवं बाह्यगणन से आप क्या समझते हैं ? आन्तरगणन की विभिन्न विधियों का वर्णन कीजिए।
What do you understand by Interpolation and Extrapolation ? Explain the different methods of Interpolation.

(2) एक व्यवसायी के लिए आन्तरगणन एवं बाह्यगणन के महत्व का वर्णन कीजिए।
Discuss the importance of Interpolation and Extrapolation to a businessman.

(3) आन्तरगणन की विधियां किन मान्यताओं पर आधारित है ? विवेचना कीजिए।
What are the assumptions upon which methods of Interpolation are based ? Discuss.

(4) बाह्यगणन से आप क्या समझते हैं ? सांख्यिकी में बाह्यगणन का प्रयोग किन-किन परिस्थितियों में किया जाता है ? उदारहण सहित स्पष्ट कीजिए।
What do you understand by extrapolation ? Under what ciucumstances is extrapolation used in statistics ? Illustrate your answer.

(5) निम्न समंकों से बिन्दु रेखीय विधि से 22 तथा 40 साल की आयु पर जीवन की उम्मीद निकालें :
Estimate graphically the expectation of life at age 22 and 40 from the following data :

Age in years	10	15	20	25	30	35
Expectation of life in years	35.4	32.2	29.1	26.0	23.1	20.4

(6) नीचे दी गई तालिका भारत में किसी एक वस्तु का उत्पादन बताती है। बिन्दु रेखीय विधि द्वारा वर्ष 1995 का उत्पादन ज्ञात कीजिए :-
The following table gives the production of a commodity in India. Estimate the production in 1995 by Graphic Method :

Year	:	1991	1992	1993	1994	1996
Production	:	260	330	380	525	965

(in Tonnes) [**Hint :** Use False Base Line]

[**Ans.** 745 Tonnes]

(7) निम्न सारणी से एक फर्म द्वारा विभिन्न वर्षों में अर्जित लाभ दिए हैं। 1996 में फर्म का लाभ ज्ञात करें :-

The following table gives the profits of a firm for different years. Estimate the profits for 1996.

Year	:	1991	1992	1993	1994	1995
Profits	:	125	163	204	238	282

(Rs. in Lakhs) [**Hint :** Binomial Method]

[**Ans.** 380]

(8) निम्न तालिका से वर्ष 1991 में होने वाली जनसंख्या का अनुमान लगाइए :

From the following table, estimate the population for 1991 :

Year	:	1931	1941	1951	1961	1971	1981
Population	:	75401	82984	86686	44947	93091	127327

[**Hint :** First interpolate the value for 1961 = 86546.9 and then extrapolate for 1991]

[**Ans.** 220750]

(9) निम्न आंकड़ों से उन प्राणियों की संख्या का आंकलन करो जिनकी आय 400 तथा 500 रू. के बीच है :

Estimate the number of persons where incomes are between Rs. 400 and Rs. 500 from the following figures :-

Income in Rs.	:	Below 200	200-400	400-600	600-800	800-1000
No. of persons in thousands.	:	120	145	200	250	150

[**Ans.** Below 500 = 355.039 thousand, Between 400-500 = 90.039 thousand—Newton's Method]

(10) निम्नलिखित आंकड़ों से $x = 2$ के लिए आन्तरगणन द्वारा तथा $x = 5$ के लिए बाह्यगणन द्वारा y के मान ज्ञात कीजिए:

Obtain the value of y for $x = 2$ by interpolation and for $x = 5$ by extrapolation :

x	:	0	1	2	3	4	5
y	:	8	10	?	15	17	?

[**Ans.** 12.5 and 18.1]

(11) किसी परीक्षा में 492 विद्यार्थियों ने निम्नलिखित अंक प्राप्त किए। उन विद्यार्थियों की संख्या बताइए जिन्होंने 42 से अधिक लेकिन 45 तक अंक प्राप्त किए हों।

The following are the marks obtained by 492 students in a certain examination. Find out the number of students who secured than 42 but not more than 45 marks.

Marks	**No. of Students**
Not more than 40	212
Not more than 45	296
Not more than 50	368
Not more than 55	429
Not more than 60	460
Not more than 65	481
Not more than 70	490
Not more than 75	492

[**Ans.** No. of Students = 40]

(12) नीचे दिये गए आँकड़ों से 35-40 वर्ष की आयु के बीच माताओं से उत्पन्न बच्चों की औसत संख्या का अनुमान लगाइए :-

From the data given below, estimate the average number of children born per mother aged 35-40 years :

Age of mother in years	**Average no. of children born**
15-20	0.7
20-25	2.1
25-30	3.5
30-35	4.8
40-45	5.8

[**Ans.** 5.72]

(13) निम्न आंकड़ों से 20 रू. से 21 रू. के बीच कमाने वाले प्राणियों की सम्भाव्य संख्या का आन्तरगणन करें :

Interpolate the probable number of persons earning between 20 and 25 rupees from the following figures:-

Income in Rs.	:	Less than 10	10-20	20-30	30-40	40-50
No. of Persons	:	150	170	200	250	180

(14) किसी उपयुक्त बीजगणितीय विधि से $x = 3$ के लिए y का मान ज्ञात करें :-

Using any suitable algebric method find out the value of y for x =3

The following values are given in a table :

x	:	1	2	3	4	5
y	:	216000	226981	—	250047	262144

(15) निम्न सारणी से 45 से 48 अंक प्राप्त करने वाले विद्यार्थियों की संख्या ज्ञात कीजिए :

From the following table find the number of students who obtained the marks between 45 to 48 :

प्राप्तांक (Marks)	:	35-40	35-45	35-50	35-55	35-60
विद्यार्थियों की संख्या (No. of Persons)	:	138	230	284	320	350

[**Ans.** 36]

(16) निम्न तालिका में, लैगरेंज विधि की सहायता से अज्ञात मूल्य का आन्तरगणन कीजिए :

Interpolate the missing figure in the table given below with the help of Lagrange's method :

Year	**Acres of Lakhs**
1961	1331
1962	1758
1963	2197
1964	?
1965	3375
1966	4096
1967	4913

[**Hint :** Revise X taking 1960 as common]

[**Ans.** 2744 Acres in Lakhs]

❑❑

अध्याय - 14

सम्भावना सिद्धान्त (Probability)

सिद्धान्त का विकास

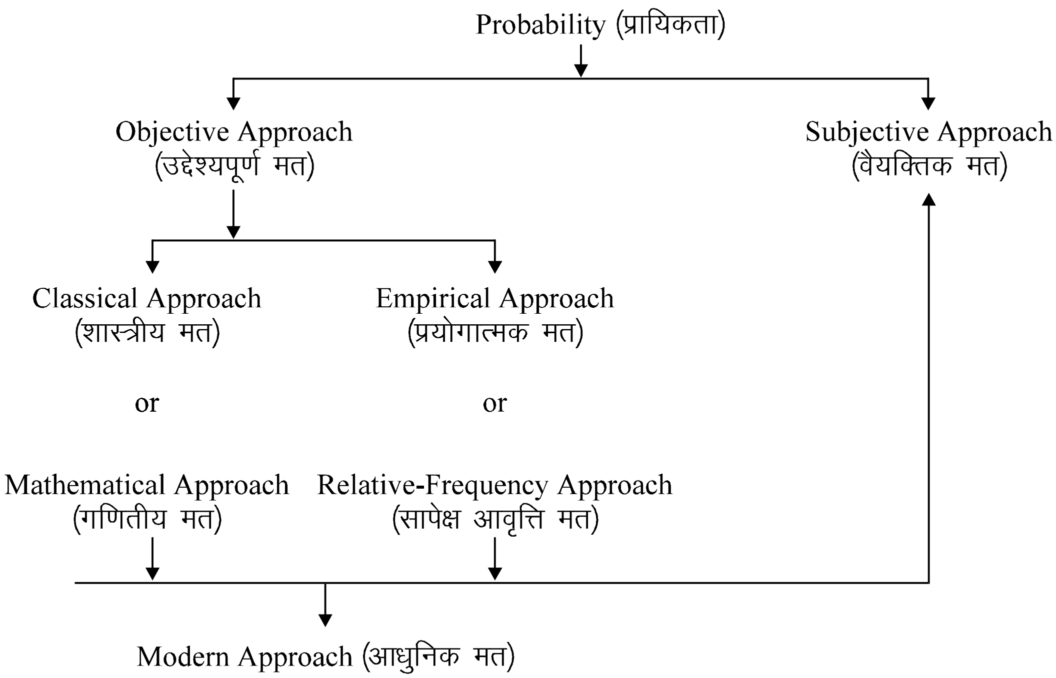
''सम्भावना सिद्धान्त का महत्व केवल ताश अथवा पासा खेलने वालों के लिए ही नहीं है जो कि उसके आदि'' देव थे, अपितु उन सभी कार्यशील व्यक्तियों, उद्योगों के अध्यक्षों, सेनानायकों आदि के लिए उसका महत्व है जिनकी सफलता निर्णयों पर निर्भर करती है, जो स्वयं दो प्रकार के कारकों पर निर्भर करते हैं, प्रथम ज्ञात अथवा गणना-योग्य तथा दूसरे अनिश्चित तथा सम्भावित।'' वास्तव में सम्भावना सिद्धान्त (Theory of Probability) के अध्ययन की प्रेरणा सर्वप्रथम जुआरियों से प्राप्त हुई थी। भाग्य देवता (Goddess Fortune) अथवा अन्धविश्वास (Supersitions) पर आस्था रखने वाले सत्रहवीं शताब्दी के कुछ जुआरी जुए में जीतने के अवसरों से सम्बन्धित समस्याओं का समाधान कराने के लिए तत्कालीन प्रसिद्ध गणितज्ञों के पास जाते थे। इस प्रकार गेलिलियो (Galileo) पॉस्कल (Pascal) फर्मेट (Fermat) नाम प्रसिद्ध गणितज्ञों ने अपनी योग्यता का उपयोग अवसरों की समस्या (Problem of Chance) को सुलझाने में किया जिससे सम्भावना सिद्धान्त के विकास का रास्ता खुला। पॉस्कल (Psacal) उन्नीसवीं शताब्दी के प्रारम्भ के फ्रांस के निवासी लौप्लेस (Laplace) तथा जर्मनी निवासी गॉस (Gauss) नामक गणितज्ञों ने 'अवसर के नियमों' (Law of Chance) के ज्ञान में महत्वपूर्ण वद्धि की थी। राष्ट्रीय अर्थ -व्यवस्था के विस्तार के साथ-साथ अनेक महान गणितज्ञों, जैसे डी मायवरे (De Moivre) निकोलस डेनियल बर्नोली यूलर डी एलमबर्ट आदि को भी सम्भावना सिद्धान्त के विकास करने की प्रेरणा मिली। इन व्यक्तियों ने इस सिद्धान्त का प्रयोग वित्तीय, जन-स्वास्थ्य, सैनिक तथा राजनैतिक क्षेत्र की समस्याओं के स्पष्टीकरण में किया। आधुनिक काल में फिशर (R.A. Fisher) पिर्यसन-पिता एवं पुत्र (Pearson, Father and son both) तथा जे. नेमेन (J. Neyman) ने सम्भावना सिद्धान्त पर आधारित न्यादर्श सिद्धान्त (Sampling Theory) का विकास किया और आज एक व्यापक सम्भावना सिद्धान्त विद्यमान है और, ''यह सम्भव प्रतीत होता है कि अगले दशकों में सम्भावना सिद्धान्त का और आगे विकास विज्ञान के इतिहास में एक महत्वपूर्ण अध्याय का श्रीगणेश करेगा।'' क्योंकि, ''प्रकृति में अवसर तथा कारणों की ऐसी व्यवस्था पायी जाती है जो इस प्रकार कार्य करती है कि उन कारणों के प्रभावों को परम्परागत सम्भावना सिद्धान्त के आधार पर, जिसमें सीमित सांख्यिकी अर्थ में गणितीय सम्भावना का प्रतिस्थापन व्यावहारिक सम्भावना से किया जाता है, पूर्व-व्यक्त किया जा सकता है।'' जेम्स जी. स्मिथ ने मत व्यक्त किया है कि, ''वैज्ञानिक पद्धति-दर्शन का एक नया यंत्र-का यह कार्य रहा है कि वह पारिकाल्पिक ज्ञान क्षेत्र में गहरा बैठे। मुख्यतः इसी कारण सम्भावना सिद्धान्त का विकास सम्भव हुआ है।'' सम्भावना सिद्धान्त का प्रयोग भौतिकशास्त्र, खगोलशास्त्र आदि की तुलना में व्यापार तथा अर्थशास्त्र में नया ही है। सांख्यिकी विज्ञान के विस्तार तथा समंकों के अधिकारिक व्यवहार से सिद्धान्त के विकास की सम्भावना और बढ़ी है। हेरोल्ड होटेलिंग के अनुसार, ''पिछले दो दशकों से समंकों के प्रयोग के अत्यधिक विस्तार का सम्बन्ध सम्भावना सिद्धान्त पर आधारित गणितीय विकास से है तथा मुख्यतः इसी से सम्भव भी हुआ है।''

सांख्यिकी में सम्भाविता सिद्धान्त का बहुत महत्व है। सांख्यिकी निष्कर्षों की व्याख्या करने में इसका बड़ा योगदान है। सांख्यिकीय नियमितता नियम (Law of Statistical Regularity) व महांक जड़ता का नियम (Law of Inertia of Large Numbers) इस सिद्धान्त पर आधारित है। सार्थकता के परीक्षण जैसे (Tests of Significance) भी इसी सिद्धान्त पर आधारित हैं।

अर्थ (Meaning) — ''सांख्यिकीय सिद्धान्त में सम्भावना का विचार महत्वपूर्ण पार्ट अदा करता है फिर भी एक सन्तोषजनक ढंग से इसको परिभाषित करना कठिन है।'' 'सम्भावना' का प्रयोग दिन-प्रतिदिन की बोलचाल में होता है; हम एक प्रकार के कथन सुनते रहते हैं : ''अमुक व्यक्ति की जीतने की सम्भावना कम है''; यह सम्भावना है कि आज शाम तक पानी बरसे''; ''सम्भवतः तुम्हारी बात सही हो''; ''सफलताक की सम्भावना पचास-पचास है'' आदि-आदि। इन समस्त कथनों में अनिश्तिता

की भावना को व्यक्त किया गया है। गौथे ने कहा है कि, अज्ञानता में कार्य करने से अधिक भय की बात और कोई नहीं है।''[7] सम्भावनाओं के रूप में तर्क करने की अनिश्चितता अथवा ज्ञान को कम करने का प्रयत्न किया जाता है। सांख्यिकी विज्ञान में 'सम्भावना' शब्द का पथक अर्थ है , यह अर्थ सामान्य प्रयोग के अर्थ में अधिक सूक्ष्म है। लैप्लेस के शब्दों में, ''सम्भावना अनुकूल घटनाओं का समस्त समान रूप से घटित घटनाओं के साथ अनुपात है।

किसी दी हुई घटना के घटने की सम्भावना ही उस घटना की प्रायिकता है। प्रायिकता सदैव शून्य (0) तथा एक (1) के मध्य होती है। यदि कोई घटना निश्चित रूप से नहीं घटेगी तो उसने घटने की प्रायिकता शून्य होगी और यदि किसी घटना का होना प्रायः निश्चित है तो उसकी प्रायिकता एक (1) होगी। प्रायिकता के विषय में विद्वानों के विभिन्न मत हैं। मुख्यतः यह मत निम्न प्रकार के हैं :-



(1) **उद्देश्यपूर्ण सम्भावना या प्रायिकता (Objective Probability Approach) :** उद्देश्यपूर्ण प्रायिकता प्रकृति के सर्वमान्य नियमों पर आधारित हैं, जिनके विषय में कोई दो मत नहीं हैं या परीक्षणों (Trials) व प्रयोगों (Experiments) पर आधारित हैं। इस प्रकार की प्रायिकता वैयक्तिक मत (Subjective Approach) से भिन्न है जो केवल किसी व्यक्ति विशेष के वैयक्तिक मत पर ही आधारित है। अतः उद्देश्यपूर्ण मत पक्षपात की परिधि से बाहर है और इसमें उद्देश्य पूरा करने की अभिलाषा प्रदर्शित होती है। इसके अन्तर्गत निम्नलिखित दो मत आते हैं:-

(अ) **शास्त्रीय मत या गणितीय मत (Classical Approach or Mathematical Apprach) :** सम्भावना शब्द का यह प्रारम्भिक वर्णन है। इसे स्वयं सिद्ध सम्भावना अथवा सर्वमान्य प्राकृतिक नियमों पर आधारित नियम (Based on Priori Law) भी कहा जाता है। गणितीय तर्क पर आधारित होने के कारण इसे गणितीय सम्भावना भी कहते हैं। **सन् 1954** में फ्रांस के स्टोरियों ने **विद्वान गणितज्ञ पॉस्कल (Pascal) व उनके समकालीन विद्वान गणितज्ञ फर्मेट (Fermat)** से सट्टे व अवसर सम्बन्धी समस्याओं का सम्भावना के माध्यम से हल निकलवाया, जिससे सम्भाविता के शास्त्रीय विचार को बल मिला। इस समय में प्रायिकता सिद्धान्त का उपयोग, सर्वमान्य उद्देश्यों जैसे — सिक्के (Coin) व पासे (Dice) के उछालने, ताश की गड्डी (Pack of Cards0 से पत्ते निकालने, किसी बर्तन से विभिन्न रंगों की गेंदें निकालने पर आने वाली विशेष रंग की गेंद सम्बन्धी कार्यों के लिए किया गया।

लैप्लेस (Laplace) : लैप्लेस के द्वारा दी गयी सम्भावना की परिभाषा प्रायः सभी गणितज्ञ मानते हैं।
इस परिभाषा को यदि हम **सूत्र के रूप में लिखें** तो ये इस प्रकार होती हैं :

$$p = \frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

जहाँ p = घटना घटित होने की प्रायिकता (Probability that the event will take place)

तथा
$$q = \frac{\text{Number of cases not favourable}}{\text{Total number of equally likely cases}}$$

जहाँ q = घटना न घटित होने की प्रायिकता (Probability that the event will not take place)

किसी घटना के घटित होने (p) तथा न घटित होने (q) की सम्भावनाओं का योग 1 होता है,

अर्थात् $p + q = 1$

गणितीय रूप में :
$$p = \frac{a}{a \quad b}, \quad q = \frac{b}{a \quad b}$$

where,
a = Number of favourable cases
b = Number of non-favourable cases
$a + b$ = Total number of equally likely cases.

Illustration 1 : From a bag containing 10 black and 5 red balls, a ball is drawn at random. What is probability that it will be red ?

Solution : Total number of balls in the bag = 10 + 5 = 15

Number of red balls = 5

Probability of getting a red balls or

$$p\,(\text{A}) = \frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

$$= \frac{5}{15} = \frac{1}{3}$$

\ Probability of getting a red ball is $\frac{1}{3}$ or .33

Illustration 2 : From a bag containing 5 white, 8 black and 17 red balls, a ball is drawn at random. What is the probability that it is white and it is not white ?

Solution. Total number of balls in the bag = 5 + 8 + 17 = 30

Number of white balls = 5

Probability of getting a white ball or

$$p\,(\text{A}) = \frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

$$= \frac{5}{30} \quad \frac{1}{6}$$

\ For the probability of not drawing white ball or

$$q = \frac{25}{30} \quad \frac{5}{6}$$

Thus $$p+q = \frac{1}{6} \quad \frac{5}{6} \quad 1$$

Illustration 3 : One card is drawn from a standard pack of 52 cards. What is the chance that it is a queen ?

Solution : There are four queens in a pack of 52 cards

\ The probability of drawing a queen or

$$P(A) = \frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

$$= \frac{4}{32} \quad \frac{1}{13}$$

गणितीय (शास्त्रीय) सम्भावित सिद्धान्त के अन्तर्गत घटनाओं में निम्न तीन विशेषताएँ होनी चाहिए :

(1) समान रूप से घटित घटनाएँ (Equally likely events)

(2) सर्वग्राही घटनाएँ (Collectively exhaustive events)

(3) पारस्परिक अपवर्जी घटनाएँ (Mutually exclusive events)

(1) समान रूप से घटित घटनाएँ (Equally Likely Events) : सम्भावना सिद्धान्त के विवेचन के 'समान रूप से घटने वाली घटनाएँ' वाक्यांश को प्रयोग किया गया है। ये वे घटनाएँ होती हैं जिनमें से कोई घटना घट सकती है। उदाहरणार्थ यदि किसी सिक्के को उछाला (Toss) जाय तो या तो वह चित्त गिरेगा या पट, अतः समान रूप से घटने वाली घटनाएँ दो हुई। निम्न उदाहरणों से यह और अधिक स्पष्ट हो जायेगा :-

1. पहली उछाल में चित्त (H) दूसरी में भी चित्त (H)
2. पहली उछाल में चित्त (H) दूसरी में पट (T)
3. पहली उछाल में पट (T) दूसरी में चित्त (H)
4. पहली उछाल में पट (T) दूसरी में भी पट (T)

इस तरह हमारे पास चार सम्भावित घटनाएं हैं (HH, HT, TH, TT) । ये चारों ही समान रूप से घटने वाली हैं।

इसी तरह से जब हम एक पासा फेंकते हैं (Roll a Dice) तो हमें 1 से 6 तक कोई भी संख्या मिल सकती है। इस तरह से हमारे पास समान रूप से घटने वाली संभावित घटनाएं हैं। यदि हम दो पासे फेंके तो संभावित घटनाओं के कुल संख्या 36 होगी जो कि निम्नलिखित होगी :-

(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)
(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)
(3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6)
(4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6)
(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6)
(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)

उपरोक्त कोष्ठकों में प्रथम संख्या पहला पासा फेंकने पर मिली संख्या व दूसरी संख्या, दूसरा पासा फेंकने पर मिली संख्या। उदाहरण के लिए (4, 5) का अर्थ है कि पहले पासे में 4 अंक मिला व दूसरे पासे में 5 अंक मिला।

यदि हम ताश के 52 पत्तों में से एक पत्ता निकालते हैं तो हमारे पास कुल 52 संभावित घटनाएं हैं जो समान रूप से घटने वाली हैं।

(2) **सर्वग्राही घटनाएँ (Collectively Exhaustive Events) :** समस्त प्रतिकूल व अनुकूल घटित घटनाओं का योग कुल घटनाओं के घटने के योग के बराबर होता है। अतः अनुकूल घटनाएं (Favourable Events) कभी भी कुल घटनाओं के योग से अधिक नहीं हो सकती। इससे अभिप्राय है कि हमें समस्त सम्भव घटनाओं का प्रयोग किए बिना भी पूर्व आभास (Intuition) होता है। उदाहरणतया एक अनभिनत सिक्के को उछालने पर केवल पट या चित्त ही सम्भव हो सकते हैं। इसी प्रकार एक पासे (Dice) को फेंकने पर 1, 2, 3, 4, 5 और 6 बिन्दु (Dots) ही आ सकते हैं। अतः यह सभी सर्वग्राही घटनाएँ हैं। इसी प्रकार एक ताश की गड्डी में चिड़ी, पान, ईंट व हुकुम के पत्ते ही हो सकते हैं, अतः यहाँ कुल सर्वग्राही घटनाएँ 52 हैं।

(3) **पारस्परिक अपवर्जी घटनाएँ (Mutually Exclusive Events) :** पारस्परिक अपवर्जी घटनाएँ वे घटनाएँ होती हैं जिनके घटने पर अन्य घटनाओं के घटने की सम्भावना समाप्त हो जाती है तथा इसके विपरीत अन्य घटना के घट जाने पर पूर्व घटना के घटने की सम्भावना समाप्त हो जाती है। अन्य शब्दों में, पारस्परिक अपवर्जी घटनाएँ एक साथ ही नहीं घट सकतीं। उदाहरणार्थ, यदि एक सिक्के को उछाला जाय तो चित्त गिर जाने पर पट गिरने अथवा पट गिर जाने पर चित्त गिर जाने की सम्भावना समाप्त हो जाती है। इसी प्रकार यदि पासे को फेंका जाय तो उसकी कोई भी साइड ऊपर आ जाने पर अन्य साइडों के ऊपर आने की सम्भावना समाप्त हो जाती है। पासे की छः साइड होती हैं परन्तु एक बार में एक ही साइड ऊपर रह सकती है अतः ये पारस्परिक अपवर्जी घटनाएँ हैं।

(II) सापेक्ष आवत्ति सम्भाविता सिद्धान्त (Relative Frequency Theory of Probability)

प्रयोगात्मक सम्भाविता (Empirical Approach) : सम्भाविता सिद्धान्त की शास्त्रीय परिभाषा (गणितीय परिभाषा) केवल सिक्कों के उछाल, पासे के फेंकने, ताश की गड्डी से पत्ते निकालने व अन्य अवसर सम्बन्धी समस्याओं को हल करने तक की सीमित है। अतः ऐसे सभी परीक्षण, जिन्हें अधिक बार दोहराया जाता है, वहाँ सापेक्ष आवत्ति सिद्धान्त कार्य करता है। इस सिद्धान्त के अनुसार कोई घटना कुल समान रूप से घटित 'n' प्रकार हो सकती है यह घटना 'A' यदि 'm' प्रकार से घटित हो सकती है, तो इसकी आवत्ति $\left(\frac{m}{n}\right)$ होगी। किन्तु यदि कुल परीक्षण (trial) अनन्त (infinite) हो तो $\frac{m}{n}$ से जो मान आएगा, उसे ही सापेक्ष आवति सम्भाविता सिद्धान्त की सीमा कहेंगे। इसे निम्न प्रकार व्यक्त किया गया है : -

$$P(A) = \underset{n}{\text{Limit}} \frac{m}{n}$$

इस प्रमेय या प्रतिपादन वाने माईस (Von Mises) ने किया। यह सिद्धान्त निम्नलिखित दो मुख्य मान्यताओं (Assumptions) पर आधारित है :

(1) परीक्षण (trial) या अवलोकन (observation) दैव निदर्शन (random sampling) पर आधारित हों तथा प्रत्येक घटना के घटने के समान अवसर (equally likely chances) हों।

(2) अवलोकनों की संख्या अनेक बार हो।

इन दोनों मान्यताओं के पूर्ण होने पर ही सापेक्ष आवत्ति स्थिर (Stable) होगी क्योंकि यह सांख्यिकीय नियमितता नियम (Law of Statistical Regularity) व महाँक जड़ता नियम (Law of Inertia of Large Numbers) का सिद्धान्त है। इसी कारण से इस सिद्धान्त को सांख्यिकीय प्रायिकता परिभाषा (Statistical Definition of Probability) भी कहते हैं। सम्भावना सिद्धान्त की इस परिभाषा का प्रयोग व्यवहार में सम्भव प्रतीत नहीं होता, क्योंकि अवलोकन प्रायः परिमित (Finite) ही होते हैं, अनन्त (Infinite) नहीं। अतः व्यवहार में इस सिद्धान्त को निम्न प्रकार से व्यक्त करते हैं :

$$P(A) = \frac{m}{n} = \frac{\text{Number of favourable cases}}{\text{Total number of equally likely cases}}$$

उदाहरण के लिए, यदि हम किसी सिक्के को 100 बार उछालें और 55 बार चित्त (Head) आए तो चित्त की सापेक्ष आवत्ति $\frac{55}{100} = 0.55$ होगी। इसी प्रकार सिक्के को 100 बार और उछालने पर यदि चित्त (Head) की संख्या 48 हो तो सापेक्ष आवत्ति

$\frac{48}{100} = 0.48$ होगी। अतः कुल 200 परीक्षणों में सापेक्ष आवत्ति $\frac{0.55+0.48}{2}$ $\frac{1.03}{2}$ 0.515 होगी। अतः जैसे-जैसे परीक्षणों की संख्या बढ़ाई जायेगी चित्त की सापेक्ष आवत्ति में स्थिरता आती रहेगी और परीक्षणों की संख्या (n) अनन्त (infinite) होने पर सापेक्ष आवत्ति (r/n) = 0.50 की ओर प्रवत्त होगी।

सीमाएँ (Limitations) : सांख्यिकी की इस परिभाषा की निम्न सीमाएँ हैं :-'

(1) विभिन्न परीक्षणों की परिस्थितियों का आवश्यक रूप से समान होना सम्भव नहीं है। प्रायः वैज्ञानिक परीक्षणों में विभिन्न शोध कार्य (Research) होने के परिणामस्वरूप परीक्षणों की परिस्थितियाँ समान नहीं रह पातीं।

(2) परीक्षणों की संख्या का अनन्त की ओर प्रवत्त होना (n tending to infinity) एक सैद्धान्तिक धारणा मात्र है।

प्रायिकता की उपर्युक्त दोनों परिभाषाएँ देखने में एक-जैसी प्रतीत होती हैं परन्तु इनमें बहुत अंतर है।

प्रथम परिभाषा में P (A), तथा $\frac{r}{n}$ में कोई अन्तर नहीं है परन्तु **द्वितीय परिभाषा में** p (A), n के अनन्त होने पर $\frac{r}{n}$ की सीमा व्यक्त करती है अर्थात् अवलोकनों के असीमित होने पर प्रायिकता सापेक्ष आवत्ति की सीमा व्यक्त करती है।

प्रथम परिभाषा द्वारा प्राप्य प्रायिकता को स्वयं सिद्ध प्रायिकता (a priori probability) **भी कहते हैं क्योंकि यह तर्कपूर्ण होती है तथा इसमें प्रयोग** (experiment) **की आवश्यकता नहीं होती परन्तु सांख्यिकी में तर्क द्वारा प्राप्त प्रायिकता बहुत कम पाई जाती है। सापेक्ष आवत्ति प्रायिकता सिद्धान्त** (relative frequency theory of probability) **द्वारा प्राप्त प्रायिकता को वास्तविक प्रायिकता** (a posteriori probability) **भी कहते हैं।** यह तर्क के स्थान पर प्रयोग (experiment) पर आधारित होती है।

(III) वैयक्तिक सम्भावना (Subjective Approach or Personalistic View or Probability)

यद्यपि सापेक्ष आवत्ति सम्भाविता सिद्धान्त ही अधिक प्रचलित व मान्य है , किन्तु फिर भी द्वितीय महायुद्ध के बाद से वैयक्तिक सम्भावना को भी बल मिला है। विंकलयर व हेयज के अनुसार, ''वैयक्तिक सम्भावना की व्याख्या एक व्यक्ति के विश्वास के माप को अथवा एक व्यक्ति के संख्यात्मक विचार को कहते हैं।''

इस प्रकार की सम्भावना की उपयोगिता ऐसी परिस्थिति में अधिक है जहाँ न तो सर्वमान्य प्राकृतिक नियम (a priori probability) लागू होते हैं और न ही घटना के घटित होने का भार (weight) दिया जाता है जो 0 और 1 के माध्य हो सकता है। यदि घटना के घटने के लिए विभिन्न प्रयोग किए जा सकते हैं। उदारहणार्थ — यदि एक अध्यापक बी. कॉम परीक्षा में किसी छात्र के प्रथम आने की सम्भावना को 0.99 व्यक्त करता है व अन्य छात्र के फेल होने की सम्भावना 0.55 व्यक्त करता है तो यह उसके व्यक्तिगत विश्वास पर ही आधारित है। इसी प्रकार एक व्यवसायी अपने विश्वास व व्यक्तिगत अनुभव के आधार पर कह सकता है कि अगले माह उसे 'अ' वस्तु का आदेश प्राप्त हो सकता है अथवा नहीं। वैयक्तिक सम्भावना में घटना के घटने की सम्भावना अधिक है तो भार 1 से समीप होगा और यदि घटना के घटने की सम्भावना नहीं है तो भार 0 के नजदीक होगा। अतः इस सिद्धान्त का क्षेत्र व्यापक व लचीला है। जिन क्षेत्रों में उद्देश्यपूर्ण आंकड़े (Objective Data) उपलब्ध नहीं है अथवा जहाँ वैयक्तिक व उद्देश्यपूर्ण आँकड़ों का मिश्रण पाया जाता है, वहाँ इस प्रकार की सम्भावना ही उत्तम है, किन्तु व्यक्ति को कड़ी जाँच के बाद ही भार व्यक्त करने चाहिए अन्यथा सम्भावना निष्कर्ष भ्रामक हो सकते हैं। व्यवसाय की विभिन्न परिस्थितियों में निर्णय लेने के लिए वैयक्तिक सम्भावना का बहुत महत्व है।

(IV) सम्भाविता सिद्धान्त की आधुनिक परिभाषा (Modern Apprach of Probability Theory)

सम्भावना की आधुनिक परिभाषा **रूस के गणितज्ञ ए. एन. कोलमोग्रोव ने सन् 1933 में दी।** इस परिभाषा में वैयक्तिक (Subjective) व उद्देश्यपूर्ण (Objective) आँकड़ों का सम्मिश्रण है। इसमें सम्भावना सिद्धान्त की कोई निश्चित परिभाषा नहीं है, बल्कि कुछ मान्यताओं (Postulated or axioms) का वर्णन है, जिन पर सम्भावना आधारित है, जो निम्न प्रकार है :-

(i) किसी घटना को घटने की प्रायिकता सदैव 0 व 1 के मध्य रहती है।

(ii) समस्त न्यादर्श के घटित होने की सम्भावना 1 है।

(*iii*) यदि 'अ' तथा 'ब' दो पारस्परिक अपवर्जी घटनाएँ (mutually exclusive events) हों, तो 'अ' अथवा 'ब' घटना के घटने की सम्भावना निम्न प्रकार व्यक्त होगी :

$$P\,(A \text{ or } B) = P\,(A) + P\,(B)$$

कोलमोग्रोव ने सम्भावना निकालने की एक नई पद्धति निकाली जो चित्रों के माध्यम (Theory of sets called venn diagram) पर आधारित है। कोलमोग्रोव ने सभी सम्भावना तथ्य जैसे समान रूप से घटित घटनाएँ, अनुकूल घटनाएँ, सर्वग्राही घटनाएँ, पारस्परिक अपवर्जी घटनाएँ, स्वतंत्र व आश्रित घटनाएँ सभी चित्रों द्वारा (Theory of Sets) समझाई हैं। वैन चित्र (Venn Diagram) तथा बिन्दुरेखीय प्रदर्शन (Graphic Presentation) का प्रयोग भी कोलमोग्रोव ने सम्भावना निकालने में किया है, इसीलिए इस सिद्धान्त को प्रायिकता का आधुनिक रूप समझा जाता है।

उपर्युक्त चारों सिद्धान्त के अपने-अपने गुण व दोष हैं। सम्भावना निकालते समय सरल व उपयुक्त प्रणाली का प्रयोग ही सर्वश्रेष्ठ है। सांख्यिकी की सहायता से ही अनिश्चितता में उचित निर्णय सम्भाविता सिद्धान्त के माध्यम से निकाले जा सकते हैं, क्योंकि सम्भावना ही अनिश्चितता का माप व भाषा है। अतः सम्भावना सिद्धान्त के व्यावहारिक प्रयोग की जानकारी होना अनिवार्य है।

प्रायिकता का महत्व अथवा उपयोगिता
(Importance or Utility of Probability)

प्रायिकता का प्रारम्भ 17वीं शताब्दी में सटोरियों से सम्बन्धित प्रश्नों का गणितीय उत्तर देने के लिए हुआ। बाद में इसका प्रयोग अवसर सम्बन्धी समस्याओं (Problems related to chance) जैसे सिक्के के उछालने, पासा फेंकने, ताश की गड्डी से पत्ते निकालने व थैले से विभिन्न रंगों की गेंदें निकालने सम्बन्धी समस्याओं का उत्तर निकालने के लिए हुआ। आजकल सम्भावना सिद्धान्त का उपयोग आर्थिक समस्याओं को सुलझाने में प्रबन्ध में योजना व नियन्त्रण सम्बन्धी निष्कर्ष निकालने में, सभी प्रकार की दुर्घटनाओं के घटने से सम्बन्धी प्रश्नों में, उत्पन्न होने वाले बच्चों के लिंग — लड़का अथवा लड़की उत्पन्न होने की सम्भावना आदि को जानने में किया जाता है। यहाँ तक कि इसका उपयोग विभिन्न अनुसंधानों व दैनिक जीवन से सम्बन्धी समस्याओं को सुलझाने से भी किया जाता है। आज के वैज्ञानिक युग में सम्भावना सिद्धान्त को निश्चित (Certainty) के स्थानापन्न के रूप में जाना जाने लगा है।

संभावना की उपयोगिता को निम्नलिखित तथ्यों से और भी स्पष्ट किया जा सकता है :-

(1) सांख्यिकी के आधारभूत नियम जैसे सांख्यिकी नियमितता नियम (Law of Statistical Regularity) व महांक जड़ता का नियम (Law of Inertia of Large Numbers) सम्भाविता सिद्धान्त पर ही आधारित है।

(2) यही नहीं विभिन्न सार्थकता के परीक्षण (Test of Significance) जेसै T-Test, F-Test, Z-Test, Chi-Square Test आदि भी सम्भाविता सिद्धान्त पर ही आधारित है।

(3) प्रायिकता या संभावना सिद्धान्त अवसर सम्बन्धी समस्या (Problems related to chance) का समाधान करने में उपयोगी है।

(4) निर्णय सिद्धान्त (Decision Theory) भी संभावना के आधारभूत नियमों पर ही आधारित है।

(5) यह सिद्धान्त प्रायः आर्थिक व व्यावसायिक समस्याओं के समाधान में प्रयोग किया जाता है। जिन स्थितियों में जोखिम व अनिश्चितता होती है, वहाँ प्रायिकता सिद्धान्त अधिक उपयोगी होता है।

(6) संभावना व वैयक्तिक मत ऐसी स्थितियों में अधिक उपयोगी है, जिनमें संभावना को वास्तविक रूप से नहीं मापा जा सकता। अतः इस मत ने संभावना का एक नया द्वार खोला है। बाद में अनुभव के आधार पर संभावना के वैयक्तिक मत में परिवर्तन भी किया जा सकता है।

यालुन-चाऊ (Ya-Lun Chou) ने सांख्यिकी में प्रायिकता सिद्धान्त के बारे में यहाँ तक कहा है कि ''सांख्यिकी अनिश्चितता में निर्णय लेने की पद्धति है।'' (Statistics is a method of decision making under uncertainty)".

आधारभूत धारणाएँ
(Fundamental Concepts)

प्रायिकता की गणना विधि का वर्णन करने से पहले कुछ शब्दो/धारणाओं/घटनाओं की जानकारी अति आवश्यक है, क्योंकि इनकी ठीक जानकारी होने से प्रश्न को समझना सरल हो जाता है और अशुद्धि की सम्भावना भी कम हो जाती है। इनका वर्णन निम्न प्रकार से किया गया है :

देव प्रयोग या परीक्षण
(Random Experiment)

यह एक प्रयोग (Experiment) है जो यदि समरूप शर्तों या दशाओं (Homogeneous Conditions) के आधार पर बार-बार दोहराया जाए तो समान परिणाम प्रदान नहीं करता। विभिन्न सम्भव परिणामों में से कोई भी एक परिणाम हो सकता है। यहाँ, परिणाम सदैव एक-समान नहीं होता। उदाहरण के लिए, यदि एक पासा (Dice) फेंका जाए तो यह हमेशा नं. 6 ही ऊपर हो इस प्रकार नहीं गिरेगा। यह 6 सम्भव ढंगों में से किसी भी प्रकार से गिरेगा अर्थात् पासे पर अंकित 6 अंकों में से कोई भी एक ऊपर आ सकता है।

निदर्शन समूह
(Sample Space)

एक दैव प्रयोग (Random Experiment) के सभी सम्भव परिणामों के समूह को **निदर्शन समूह** कहते हैं। दूसरे शब्दों में, दैव प्रयोग की सभी सम्पूर्ण घटनाओं (Exhaustive Events) का समूह निदर्शन समूह है। प्रयोग के प्रत्येक सम्भव परिणाम को अवयव या निदर्शन बिन्दु (Element or Sample Point) भी कहा जाता है। एक सिक्के को उछालने पर निदर्शन समूह के केवल दो परिणाम हो सकते हैं — चित्त या पट्, इस प्रकार निदर्शन समूह = (चित्त, पट)। एक पासे को फेंकने पर निदर्शन समूह (1, 2, 3, 4, 5, 6) होगा। यदि एक साथ दो सिक्के फेंके जाएं तो प्रयोग के चार सम्भव परिणाम होंगे :-

सिक्का 2 (Coin 2) / सिक्का 1 (Coin 1)	H	T
H	HH	HT
T	TH	TT

इस प्रयोग का निदर्शन समूह = (HH, HT, TH, TT)

सरल घटनाएँ (Simple Events)

जब एक समय में एक ही घटना घटित होती है, तो उसे सरल घटना कहते हैं। उदाहरण के लिए, एक छः पहलुओं वाले पासे (Six-faced Dice) के फेंके जाने पर बिन्दु 6 प्राप्त होना।

संयुक्त घटनाएँ (Compound Events)

जब दो या दो से अधिक घटनाएँ एक साथ घटित होती हैं तो उनके संयुक्त रूप से घटित होने को संयुक्त घटना कहा जाता है। उदाहरण के लिए, दो पासों को एक साथ फेंकना, दो सिक्कों को एक साथ उछालना, इत्यादि।

स्वतंत्र घटनाएँ (Independent Events)

यदि किसी घटना के घट जाने से आगे किसी घटना के घटने पर प्रभाव नहीं पड़ता तो ऐसी घटनाएँ स्वतंत्र होती हैं। जब

दो घटनाओं का प्रभाव एक दूसरे पर नहीं पड़ता, तो वे घटनाएँ स्वतंत्र होती हैं। उदाहरणार्थ, एक बार सिक्का उछालने के परिणाम का प्रभाव दुबारा सिक्का उछालने के परिणाम पर नहीं पड़ता।

आश्रित घटनाएँ (Dependent Events)

यदि एक घटना के घट जाने पर प्रभाव दूसरी घटना के घटने पर प्रभाव पड़ता है तो उन घटनाओं को आश्रित घटनाएँ (Dependent Events) कहा जाता है। उदाहरणार्थ, ताश की एक गड्डी में से बादशाह निकालने की सम्भावना $\frac{4}{52}$ या $\frac{1}{13}$ है, यदि पहली बार ताश निकालने में बादशाह निकल आता है तो दूसरी बार बादशाह निकलने की सम्भावना $\frac{3}{51}$ होगी। अतः पहले वाली घटना का प्रभाव दूसरी घटना पर पड़ा, ये आश्रित घटनाएँ हैं।

सम्भावना निकालने के नियम
(Rules of Calculating Probability)

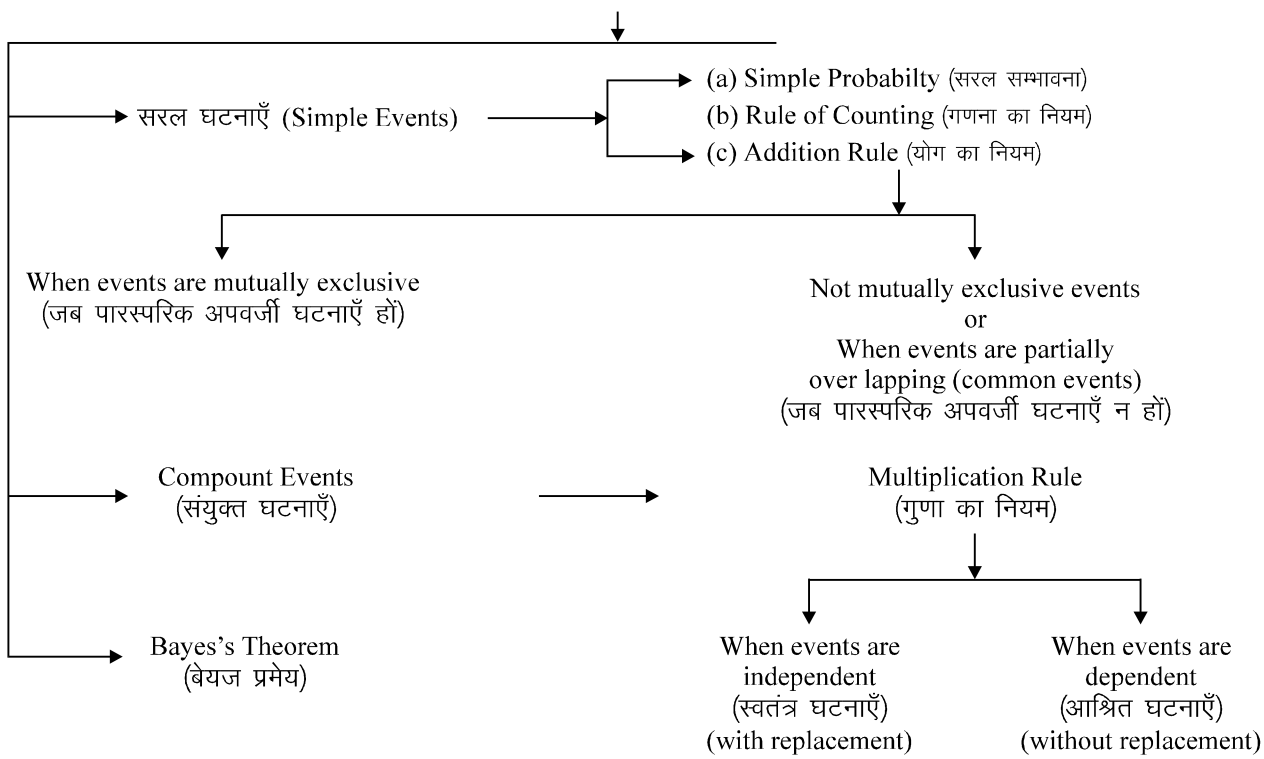


Illustration 4 : From a bag containing 12 white and 18 black balls, a ball is drawn at random. (*i*) What is the probability that it is the black ball ?

Solution :

Number of white balls = 12

Number of black balls = 18

Total Number of balls = 30

$$P \text{ (Probability of getting black ball)} = \frac{a}{n} \quad \frac{18}{30} \quad \frac{3}{5} \quad 0.6 = 0.6$$

$$q \text{ (Probability of not getting black ball)} = \frac{12}{30} \quad \frac{2}{5} = 0.4$$

Thus, $P + q = 0.6 + 0.4 = 1$

उदाहरण (Illustration) 5 : 52 पत्तों की ताश की गड्डी में से एक पत्ता निकालने पर बादशाह आने की प्रायिकता क्या हैं ?

What is the chance of drawings a king in a draw from a pack of 52 cards ?

हल (Solution) : Total number of cases that can happen = 52 (N)

Number of favourable cases *i.e.*, the total number of kings in a pack of cards = 4 (P)

$$\text{The probability} = \frac{P}{N} \quad \frac{4}{52} \quad \frac{1}{13}$$

शब्दावली (Terminology) : (in case of problems based on cards)

1.	ईंट	= Diamond	1.	इक्का	= Ace
2.	पान	= Hearts	6.	बादशाह	= King
3.	हुकम	= Spades	7.	बेगम	= Queen
4.	चिड़ी	= Clubs	8.	गुलाम	= Knave or jack

Illustration 6 : What is the probability of getting number 5 in a single throw of a dice ?

Solution : Total no. of cases = 6

Number of favourable cases = 1

\ The probability of getting number S $= \frac{1}{6}$

Illustration 7 : Find the chance of picking an even number from a sereis of natural numbers 1 to 100.

Solution : Number of even numbers in a series 1 to 100 = 50

Chance or probability of picking an even number $= \frac{50}{100} \quad \frac{1}{2}$

Illustration 8 : A single letter is Nicked at random few the word Mathematics. What is the probability that it is M ?

Solution : Total number of cases = 11

Number of favourable cases = 2

\ The probability $= \frac{2}{11}$

Illustration 9 : The data concerning promotion and academic qualification regarding 100 employees of a company is as under :-

नीचे एक कम्पनी के 100 कर्मचारियों के समंक पदोन्नति व शैक्षणिक योग्यता के आधार पर दिए गए हैं :

	C.A.'s	**Non-C.A.**	**Total**
Promoted	12	40	52
Not Promoted	13	35	48
Total	25	75	100

An employee is selected at random, find the probability that

(*i*) He is C.A.

(*ii*) He is promoted.

(*iii*) An employee who is C.A. is promoted.

(*iv*) A promoted employee is C.A.

Solution : Let various events are :-

'A' = an employee is C.A.

'B' = an employee is promoted.

'C' = C.A. is promoted.

'D' = A promoted employee is C.A.

(*i*) $P(A) = \frac{m}{n} \quad \frac{25}{100} \quad \frac{1}{4}$

(*ii*) $P(B) = \frac{52}{100} \quad \frac{13}{25}$

(*iii*) $P(C) = \frac{12}{15}$

(*iv*) $\frac{12}{15} = \frac{3}{13}$

Illustration 10 : Find the probability of getting 35 is randomly forming 2-digits numbers from out of 1, 3, 5, 7, 9 repetitons of a digit being allowed in forming the number.

Solution : Number of possible ways of getting 2-digit number is 5 × 5 = 25.

The number 37 can be formed only in one way. Hence, the required probability is $\frac{1}{25}$.

Illustration 11 : A bag contains 10 red, 6 white and 4 black balls. Two balls are drawn randomly, one by one, without replacement.

Find the probability that :

(*i*) **Both are red**

(*ii*) **None is white**

(*iii*) **2nd ball is black**

Solution :1st ball can be drawn in 20 (10 + 6 + 4) ways.

2nd ball can be drawn in 19 (20 – 1) ways.

\ Total number of possible outcomes = 19 × 20 = 380

(*i*) There are 6 red black balls. 1st red ball can be drawn in 6 ways and 2nd red ball can be drawn in 5 ways.

\ Number of favourable cases = 6 × = 30

Thus, the required probability = $\frac{30}{380} \quad \frac{3}{38}$

(*ii*) Total number of possible outcomes remain the same *i.e.* 380

Since, out of 2 ball drawn, none is white, so these 2 balls are to be drawn from the remaining 16 balls.

\ Number of favourable cases = 16 × 15 = 210

Thus the required probability = $\frac{210}{380} \quad \frac{21}{38}$

(*iii*) Total number of possible outcomes = 380

Probability of 2nd ball being black can not be directly found without considering the outcome of the first-draw, which has two possible outcomes :

(*a*) 1st ball is black. It can be drawn in 4 ways. Number of ways of drawing 2nd black is 3.

\ Number of favourable cases = 4 × 3 = 12 (*b*) 1st ball is not red. It can be drawn in 16 ways. 2nd ball is black.

It can be drawn in 4 ways.

\ Number of favourable cases = $16 \times 4 = 64$

Thus total number of favourable cases = $64 + 12 = 76$

\ The required probability = $\frac{76}{380} = \frac{1}{5}$

Illustration 12 : In a single throw with two uniform dice, find the probability of throwing (*i*) Five, and (*ii*) Eight.

Solution : Possible ways in a single throw with the uniform dice = $6 \times 6 = 36$

(*i*) Number of ways of getting the sum of 5 on throwing two uniform dice at a time is 4, *i.e.*

[(1, 4), (4, 1), (2, 3), (3, 2)]

Required probability = $\frac{4}{36}$ $\frac{1}{9}$

(*ii*) Number of ways of getting the sum of 8 on two dice is 5, *i.e.* (2, 6), (3, 5), (4, 4), (5, 3), (6, 2).

Required probability = $\frac{5}{36}$

Illustration 13 : परिवार में कम से कम एक पुत्र होने की प्रायिकता बताए यदि औसत रूप से एक परिवार में दो बच्चे हों। Find the probability of having at least one son in a family if there are two children in a family on an average.

Solution : Two children in a family may be either :

or (1) Both sons

or (2) Son and daughter

or (3) Daughter and Son

or (4) Both daughters

Thus, total number of equally likely

cases = $n = 4$

At least one son implies that a family may have one son or two sons.

Thus, favourable number of cases = $m = 3$ (*i.e.* option Nos. 1, 2 and 3)

$$P(A) = \frac{m}{n} = \frac{\text{Number of favourable cases}}{\text{Total Number of equally likely cases}} = \frac{3}{4}$$

Illustration 14 : In a single toss of your coins find the probability of getting :-

(*i*) Two heads (*ii*) Three tails (*iii*) Not more than three heads.

Solution : In a toss of 4 coins total number of possible outcomes = 16

(HHHH, HHHT, HHTH, HHTT, HTHH, HTHT, HTTH, HTTT, THHH, THHT, THTH, THTT, TTHH, TTHT, TTTH, TTTT)

(*i*) No. of favourable cases = 6

\ Required probability = $\frac{6}{16}$ $\frac{3}{8}$

(*ii*) No. of favourable cases = 4

\ Required probability = $\frac{4}{16}$ $\frac{1}{4}$

(*iii*) No. of favourable cases = 16

No. of cases where there are 4 heads = 16 – 1 = 15

\ Required probability = $\frac{15}{16}$

गिनती का नियम (Rule of Counting) : सम्भावना सिद्धान्त को समझने में गिनती के नियम का बड़ा योगदान है। किसी घटना की सम्भावना जानने के लिए अनुकूल घटित घटनाओं व कुल समान रूप से घटित घटनाओं का जानना आवश्यक है, जिसके लिए गिनती का नियम जानना अनिवार्य है। इस नियम के अनुसार यदि कोई कार्य 'm' प्रकार से किया जा सकता है तथा दूसरा कार्य 'n' प्रकार से पूर्ण किया जा सकता है तो दोनों कार्यों को एक साथ पूरा करने में कुल $(m \times n)$ तरीकें हो सकते हैं।

हल : माना पहला कार्य (a) है व दूसरा कार्य (b) है। दोनों कार्यों को एक साथ पूर्ण करेंगे।

$$\begin{array}{llll}
a_1b_1, & a_1b_2, & a_1b_3, \text{------------------} & a_1b_n, \\
a_2b_1, & a_2b_2, & a_2b_3, \text{------------------} & a_2b_n, \\
\quad | & \quad | & \quad | & \quad | \\
a_mb_1, & a_mb_2, & a_mb_3, \text{------------------} & a_mb_n,
\end{array}$$

अतः दोनों कार्यों को एक साथ पूरा करने के सम्भव तरीके $m \times n$ होंगे।

उदाहरणार्थ — दो सिक्कों के एक साथ $2 \times 2 = 4$ तरीके से उछाल सकते हैं।

अर्थात HH, HT, TH, TT, where H = Head and T = Tail.

Illustration 15 : Three dice are rolled simultaneously. Find the probability of getting a total of (*i*) not more than 5, (*ii*) atleast 15, and (*iii*) exactly 8.

तीन पाँसें एक साथ फेंक गये हों तो (*i*) 5 से अधिक के योग नहीं (*ii*) कम से कम 15 का योग (*iii*) 8 के योग की संभावना बतायें।

Solution : First dice can be tossed in = 6 ways

Second dice can be tossed in = 6 ways

Third dice can be tossed in = 6 ways

\ All the three dice can be tossed in = $6 \times 6 \times 6 = 216$ ways

(By using rule of counting)

Thus, total number of ways of throwing 3 dice= $n = 216$ ways.

With 3 dice, we can get a minimum total of 3 (*i.e.* 1 + 1 + 1) and maximum total of 18 (*i.e.* 6 + 6 + 6)

(*i*) Sum of not more than 5 with 3 dice means getting a sum of 3 or 4 or 5.

Sum of 3 with 3 dice can come in = 1 way
(1 + 1 + 1)

Sum of 4 with 3 dice can come in = 3 ways
(1 + 1 + 2, 1 + 2 + 1, 2 + 1 + 1)

Sum of 5 with 3 dice can come in = 6 ways
(1 + 1 + 3, 1 + 2 + 2, 1 + 3 + 1,
2 + 1 + 2, 3 + 1 + 1, 2 + 2 + 1)

Total number of favourable ways = m = 10 ways

We know that,

$$P(A) = \frac{m}{n} = \frac{\text{No. of favourable cases}}{\text{Total no. of equally likely cases}} = \frac{10}{216} = \frac{5}{108}$$

(*ii*) Sum of at least 15 with 3 dice means getting a sum of either 15 or 16 or 17 or 18.

Sum of 15 with 3 dice can come in = 10 ways

6 + 3 + 6, 3 + 6 + 6, 6 + 6 + 3, 4 + 5 + 6,

4 + 6 + 5, 5 + 6 + 4, 5 + 4 + 6, 6 + 4 + 5

Sum of 16 with 3 dice can come in = 6 ways

6 + 6 + 4, 6 + 5 + 5, 6 + 4 + 6, 5 + 5 + 6, 5 + 6 + 5, 4 + 6 + 6

Sum of 17 with 3 dice can come in = 3 ways

6 + 5 + 6, 6 + 6 + 5, 5 + 6 + 6

Sum of 18 with 3 dice can come in = 1 way

Total number of favourable cases = 20 ways

\ $$P(A) = \frac{20}{216} \quad \frac{5}{54}$$

योग-प्रमेय

(Addition Theorem)

दो या दो से अधिक परस्पर अपवर्जी घटनाओं (Mutually Exclusive Events) में से किसी एक घटना (Either A or B) के घटने की प्रायिकता उन घटनाओं के अलग-अलग घटित होने की प्रायिकताओं का योग है।

सूत्र के रूप में

दो घटनाएँ :

$$P(A \text{ or } B) = P(A) \overset{a_1}{\underset{n}{+}} P(B)$$

दो से अधिक घटनाएँ

$$P(A \text{ or } B \text{ or } C \ldots) = P(A) + P(B) + P(C) + \ldots\ldots$$

योग-प्रमेय का प्रमाण (Proof of Addition Theorem) : यदि घटना A 'a_1' तरीकों तथा घटना A 'a_2' तरीकों से घटित हो सकती है तो घटना A या B, '$a_1 + a_2$' तरीकों से घटित हो सकती है। यदि घटना के घटित होने की कुल सम्भावनाएँ 'n' के बराबर हैं तो :

$$\frac{a_1 \quad a_2}{n} = \qquad + \frac{a_2}{n}$$

परन्तु $\frac{a_1}{n} = P(A)$ तथा $\frac{a_2}{n} = P(B)$

अर्थात् $P(A \text{ or } B) = P(A) + (B)$

योग-प्रमेय उसी दशा में सत्य होगा जबकि :

(*i*) घटनाएँ परस्पर अपवर्जी (mutually exclusive) हों, तथा

(*ii*) घटनाएँ एक ही प्रकार की हों।

Illustration 16 : One card is drawn from a standard pack of 52. What is the probability that it it either a king or knave ?

Solution : There are 4 kings and 4 knaves in a pack of 52 cards.

\ The probability that the card drawn is a king is $P(A) = \frac{4}{52}$

and the probability that the card drawn of Jack is P (B) = $\frac{4}{52}$

Since the events are mutually exclusive, the probability that the drawn card is either king or knave is

$$P\ (A\ \text{or}\ B) = P\ (A) + P\ (B)$$

$$= \frac{4}{52} \quad \frac{4}{52} \quad \frac{8}{52} \quad \frac{2}{13}$$

Illustration 17 : A perfect dice is thrown. What is the probability of throwing 3 or 6 ?

Solution : The probability of throwing 3 is P (A) = $\frac{1}{6}$

The probability of throwing 6 is P (B) = $\frac{1}{6}$

\ The probability of throwing 3 or 6 is

$$P\ (A\ \text{or}\ B) = P\ (A) + P\ (B)$$

$$= \frac{1}{6} \quad \frac{1}{6} \quad \frac{2}{6} \quad \frac{1}{3}$$

Illustration 18 : One card is drawn from a standard packis 52. What is the chance that it is either a document or a heart ?

Solution : There are 13 diamond cards and 13 hearts cards in a pack of 52 cards.

\ The probability of drawing a diamond is P (A) = $\frac{4}{12}$

The probability of drawing a heart is P (B) = $\frac{3}{12}$

Since the events are mutually exclusive, the chance the card being a diamond or a heart is

$$P\ (A\ \text{or}\ B) = P\ (A)\ \text{or}\ P\ (B)$$

$$= \frac{13}{52} \quad \frac{13}{52} \quad \frac{26}{52} \quad \frac{1}{2}$$

Illustration 19 : A box contans 4 red, 2 white, 3 black and 3 yellow balls. What is the probability of getting a red or yellow ball at random in single draw of one ball.

Solution : The probability of getting one red ball P (A) = $\frac{4}{12}$

The probability of getting one yellow ball P (B) = $\frac{3}{12}$

Since the events are mutually exclusive, the probability of the drawn bell being red or yellow is

$$P\ (A\ \text{or}\ B) = P\ (A) + P\ (B)$$

$$= \frac{4}{12} \quad \frac{3}{12} \quad \frac{7}{12}$$

Illustration 20 : Find out the probability of getting a total of either 7 or 11 in a single roll of two dice.

Solution : A total of 7 come in the following 6 ways.

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \end{bmatrix}$$

A total of 11 can be in 2 different ways $\begin{bmatrix} 5 & 6 \\ 6 & 5 \end{bmatrix}$

The probability of getting a total of 7 is P (A) $= \dfrac{6}{36}$

The probability of getting a total of 1 is P (B) $= \dfrac{2}{36}$

Since the events are mutually exclusive, the probability of getting either 7 or 11 is

$$P(A \text{ or } B) = P(A) + P(B)$$

$$= \frac{6}{36} \quad \frac{2}{36} \quad \frac{8}{36} \quad \frac{2}{9}$$

Example 21. Among 3 events A, B and C only one event can take place. The odds against A are 3 : 2, against B are 4 : 3. Find the odds agains event C.

तीन घटनाओं अ, ब, स में से केवल एक घटना घट सकती है। 'अ' घटना के प्रतिकूल होने की आशा 3 : 2 है तथा 'ब' घटना के प्रतिकूल होने की सम्भावना 4 : 3 है तो स घटना से प्रतिकूल होने की प्रायिकता ज्ञात कीजिए।

Solution : Probability of happenning A event is P (A) $= \dfrac{2}{5}$ [3 विपरीत व 2 अनुकूल है]

Probability of happening B event is P (B) $= \dfrac{3}{7}$ [4 विपरीत व 3 अनुकूल है]

(Since events are mutually exclusive so by using addition rule)

$$\left(\frac{2}{5} \quad \frac{3}{7}\right)$$

$$P(A \text{ or } B \text{ or } C) = P(A) + P(B) + P(C)$$

By substituting values, $1 = \dfrac{2}{5} \quad \dfrac{3}{7} + P(C)$ [∵ Total prob. is = 1]

$$1 - \qquad = P(C)$$

or $P(C) = \dfrac{6}{5}$

$\dfrac{6}{5}$ probability implies 6 in favour out of 35 chances, thus odds against events are 35 – 6 = 29

Thus, odds against even C are = 29 : 6

Example 22 : A bag contains 25 balls, numbered from 1 to 25, one is to be drawn at random. Find the probability that the number of the drawn ball will be a multiple of 5 or of 7.

Solution : The probability of the number being multiple of 5 (5, 10, 15, 20, 25) $= \dfrac{5}{25}$

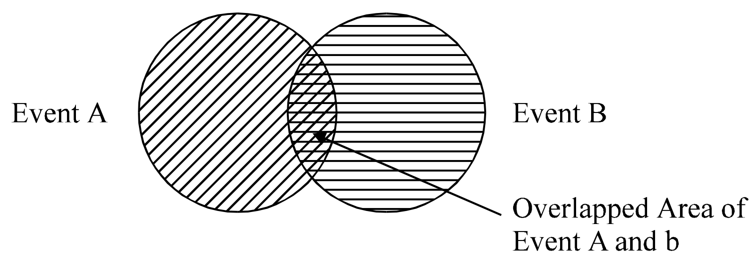
The probability of the number being multiple of 7 (7, 14, 21) $= \dfrac{3}{25}$

Thus the probability of the number being a multiple of 5 or 7 will be

$$\frac{5}{25} \quad \frac{3}{25} \quad \frac{8}{25}$$

(*c*) **योग का नियम यदि घटनाएँ पूर्ण रूप से पारस्परिक अपवर्जी न हों (Addition Rule if events are not mutually exclusive)** यदि दो घटनाएँ पूर्ण रूप से पारस्परिक अपवर्जी नहीं हैं तो ऐसी देशा में समान रूप से घटने वाली घटनाओं को जानने के लिए योग के नियम में संशोधन करना अनिवार्य हो जाता है। उदाहरणार्थ - एक ताश की गड्डी से एक पत्ता निकालने पर बादशाह अथवा हुकुम के पत्ते आने की सम्भावना जानने के लिए योग के सिद्धान्त में संशोधन अनिवार्य है, क्योंकि हुकम के 13 पत्तों में हुकुम का बादशाह भी है और 4 बादशाहों में भी हुकुम का बादशाह सम्मिलित है। अतः हुकम का बादशाह दो बार सम्मिलित हो गया अतः एक बार हुकुम का बादशाह कम करना पड़ेगा। अर्थात् 4 बादशाह + 13 हुकम के पत्ते – 1 हुकम का बादशाह = 16 पत्तें अनुकूल हैं। अतः प्रायिकता यहाँ $\frac{16}{52}$ होगी।

Not Mutually exclusive events or overlapping Events



सूत्र की अभिव्यक्ति निम्न प्रकार करेंगे :

$$P\,(A \text{ or } B) = P\,(A) + P\,(B) - P\,(A \text{ and } B)$$

इस संशोधन के करने से घटनाएँ स्वतः ही पूर्णरूप से पुनः पारस्परिक अपवर्जी हो जाएंगी। निम्न उदाहरणों से इस विधि को समझा जा सकता है।

Illustration 23 : 1 थैले में 30 गेंदें हैं जिन पर 1 से 30 तक नम्बर लिखे हैं। एक गेंद यादच्छिक रूप से चुनी गई। बताओं कि उस गेंद के (अ) 5 या 9 तथा (ब) 5 या 6 के गुणक में होने की क्या प्रायिकता है ?

(*a*) 5 or 9 and **(*b*) 5 or 6**

Solution : (*a*) The probability of the number being multiple of 5

$$5\,(5, 10, 15, 20, 25, 30) = \frac{6}{30}$$

The probability of the number being multiple of 9

$$9\,(9, 18, 27) = \frac{3}{30}$$

Since the events are mutually exclusive the probability of the number being a mutiple of 5 or 9 will be

$$= \frac{6}{30} \quad \frac{3}{30} \quad \frac{9}{30} \quad \frac{3}{10}$$

(*b*) The probability of the number being a multiple of 5

$$5\,(5, 10, 15, 20, 25, 30) = \frac{6}{30}$$

The probability of the number being a multiple of 6

$$6\,(6, 12, 18, 24, 30) = \frac{5}{30}$$

Hence the probability of getting a number either a multiple of 5 or 6 is

$$= \frac{6}{30} \quad \frac{5}{30} \quad \frac{11}{30}$$

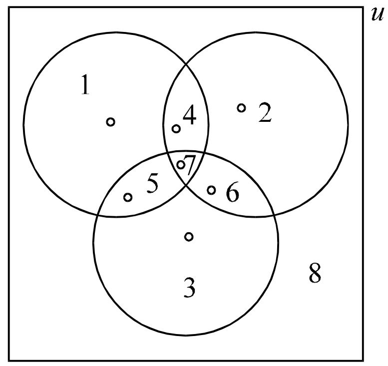
But this is not correct because item No. 30 is a multiple of both 5 and 6.

Hence the correct probability will be $= \frac{11}{30} \quad \frac{1}{30} \quad \frac{10}{30} \quad \frac{1}{3}$

In the case of three events

$$P\,(A \text{ or } B \text{ or } C) = P\,(A) + P\,(B) + P\,(C) - P\,(AB) - P\,(AC) - P\,(BC) + P\,(ABC)$$

यह चित्र से स्पष्ट किया गया है



घटनाओं का पारस्परिक अपवर्जी (Mutually Exclusive) होने के साथ-साथ उनका ही एक प्रकार (Belong to Same Set) का होना भी आवश्यक है। वॉन मिसेस (Von. Mises) के निम्न उदाहरण से इसकी आवश्यकता स्पष्ट हो जाती है :

''यदि यह कल्पना की जाय कि एक मनुष्य के 40-41 वें वर्ष में मरने की सम्भावना 0.011 तथा 41-42 वें वर्ष में विवाह करने की सम्भावना 0.009 है। ये दोनों घटनाएँ पारस्परिक अपवर्जी हैं, परन्तु यह नहीं कहा जा सकता कि मनुष्य के 40-41वें वर्ष में मर जाने तथा 41-42वें वर्ष में विवाह करने की सम्भावना .011 + .009 = .02 है, क्योंकि ये दोनों घटनाएँ एक ही प्रकार की नहीं हैं।''

उपरोक्त चित्र 8 भागों में बांटा गया है जो निम्न प्रकार से है :-

(1) केवल A ($A\overline{BC}$)

(2) केवल A ($\overline{A}B\overline{C}$)

(3) केवल A ($\overline{AB}C$)

(4) केवल A और B ($AB\overline{C}$)

(5) केवल A और C ($A\overline{B}C$)

(6) केवल B और C ($\overline{A}BC$)

(7) A, B तथा C तीनों (ABC)

(8) A, B, C में से कोई भी नहीं ($\overline{ABC}$)

U - Universal set or total number of observations.

इस चित्र से हम ये निष्कर्ष निकाल सकते हैं कि A का कुल योग

$$n\,(A) = n\,(A\overline{BC}) + n\,(AB\overline{C}) + n\,(A\overline{B}C) + n\,(ABC)$$

तथा $$n\,(AB) = n\,(AB\overline{C}) + n\,(ABC)$$

इसी तरह हम और भी सम्बन्ध निकाल सकते हैं।

Illustration 24 : A library has 1000 members out of which 525 read magazine A, 475 read magazine B and 325 read magazine C. 200 read A and B, 150 read B and C, 100 read A and C and 40 read all the three magazines. One member is chosen at random. Find the probability that he/she reads none of the three magazines.

Solution : $n\,(u) = 1000$

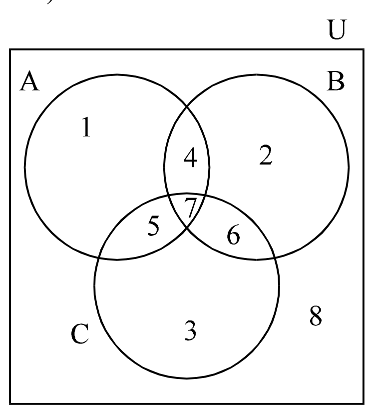
$$n(A) = 525 \qquad P(A) = \frac{n(A)}{n(u)} = 0.525$$
$$n(B) = 475 \qquad P(B) = 0.475$$
$$n(C) = 325 \qquad P(C) = 0.325$$
$$n(AB) = 200 \qquad P(AB) = 0.200$$
$$n(BC) = 150 \qquad P(BC) = 0.150$$
$$n(CA) = 100 \qquad P(CA) = 0.100$$
$$n(ABC) = 40 \qquad P(ABC) = 0.040$$

Now,
$$\begin{aligned} P(A \text{ or } B \text{ or } C) &= P(A) + P(B) + P(C) - P(AB) - P(BC) - P(CA) + P(ABC) \\ &= 0.525 + 0.475 + 0.325 - 0.200 - 0.150 - 0.100 + 0.040 \\ &= 1.365 - 0.450 \\ &= 0.915 \end{aligned}$$

P (reader reads none of three magazines)

$$\begin{aligned} P(\overline{ABC}) &= 1 - P(A \text{ or } B \text{ or } C) \\ &= 1 - 0.915 = 0.085 \end{aligned}$$

Alternative Method — (Using Vendiagram)



हम $n(ABC)$ से शुरू करते हैं।

(1) $$\begin{aligned} n(A\overline{BC}) &= n(A) - [n(AB\overline{C}) + n(A\overline{B}C) + n(ABC)] \\ &= 525 - [160 + 60 + 40] = 265 \end{aligned}$$

(2) $$\begin{aligned} n(\overline{A}B\overline{C}) &= n(B) - [n(AB\overline{C}) + n(\overline{A}BC) + n(ABC)] \\ &= 475 - [160 + 110 + 40] = 165 \end{aligned}$$

(3) $$\begin{aligned} n(\overline{AB}C) &= n(C) - [n(A\overline{B}C) + n(\overline{A}BC) + n(ABC)] \\ &= 325 - [60 + 110 + 40] = 115 \end{aligned}$$

(4) $$\begin{aligned} n(AB\overline{C}) &= n(B) - n(ABC) \\ &= 200 - 40 = 160 \end{aligned}$$

(5) $$\begin{aligned} n(A\overline{B}C) &= n(AC) - n(ABC) \\ &= 200 - 40 = 160 \end{aligned}$$

(6) $$\begin{aligned} n(\overline{A}BC) &= n(BC) - n(ABC) \\ &= 150 - 40 = 110 \end{aligned}$$

कुल पाठकों की संख्या = (1) + (2) + (3) + (4) + (5) + (6) + (7)

$$= 265 + 165 + 115 + 160 + 60 + 110 + 40 = 915$$

(8) उन सदस्यों की संख्या जो कोई भी मैगजीन नहीं पढ़ते $= 1000 - 915 = 85$

Probabilty that a reader does not read any of the three magazines $= \frac{85}{1000} = 0.085$

(2) गुणन-प्रमेय (Multiplication Theorem) : "यदि दो घटनाएं पारस्परिक रूप से स्वतंत्र (Mutually Independent) हों तथा एक घटना के घटने की सम्भावना P_1 तथा दूसरी घटना के घटने की सम्भावना P_2 हो तो दोनों घटनाओं के एक साथ घटने की सम्भावना $P_1 \times P_2$ होगी।" उदाहरणार्थ, एक सिक्का उछालने में उसके चित्त गिरने की सम्भावना $\frac{1}{2}$ है, एक पासा फेंकने में 4 संख्या आने की सम्भावना $\frac{1}{6}$ है। यदि सिक्के तथा पासे दोनों को एक साथ उछाला तथा फेंका जाय तो सिक्के के चित्त गिरने तथा पासे में 4 संख्या आने की सम्भावना

$$\frac{1}{2} \quad \frac{1}{6} = \frac{1}{12} \text{ होगी।}$$

Illustration 25 : What is probability of throwing two 'fours' in two throws of a dice.

Solution : The probability of a 'four' in first throw $= \frac{1}{6}$

The probability of a 'four' in second throw $= \frac{1}{6}$

The probability of two 'four' $= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

Illustration 26 : What is the probability of getting all the heads in four throws of a coin.

Solution : The chance of getting head in the 1st throw $= \frac{1}{2}$

The chance of getting head in the 2nd throw $= \frac{1}{2}$

The chance of getting head in the 3rd throw $= \frac{1}{2}$

The chance of getting head in the 4th throw $= \frac{1}{2}$

Illustration 27 : A problem in statistics is given to three students A, B, C whose chances of solving it are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ respectively. What is the probability that the problem will be solved.

Solution : Probability that student A will fail to solve the problem $= 1 - \frac{1}{2} = \frac{1}{2}$

Probability that student B will fail to solve the problem $= 1 - \frac{1}{3} = \frac{2}{3}$

Probability that student C will fail to solve the problem $= 1 - \frac{1}{4} = \frac{3}{4}$

Since the events are independent, the probability that all the three students A, B, C will fail to solve the problem.

$$= \frac{1}{2} \quad \frac{2}{3} \quad \frac{3}{4} \quad \frac{1}{4}$$

\ The probability that the problem will be solved $= 1 - \frac{1}{4} \quad \frac{3}{4}$

सम्भावना के गुणन प्रमेय के प्रयोग में भी यह आवश्यक है कि सभी घटनाएँ एक ही प्रकार (same set) की होनी चाहिएं। इस आवश्यकता का महत्व करने के लिए मोरोने (Moroney) ने अपनी पुस्तक 'Facts from figures' में एक बड़ा ही रोचक उदाहरण दिया है जो इस प्रकार है :

''उस व्यक्ति के बारे में सोचिए जो अपनी भावी पत्नी में अनेकों असम्बन्धित प्रकृति के गुणों की माँग करता है। कल्पना कीजिए वह चाहता है कि उसकी भावी पत्नी की ग्रीकवासियों जैसी नाक हो, सुनहले बाल हों, आँखों के रंग अलग-अलग हों — एक आँख नीली तथा एक आँख बादामी हो तथा उसे सांख्यिकी विज्ञान को उच्चस्तरीय ज्ञान हो। ऐसी स्त्री मिलने की क्या सम्भावना होगी?''

इस उदाहरण में सम्भावना का गुणन प्रमेय प्रयोग नहीं किया जा सकता क्योंकि सभी घटनाएँ अलग-अलग प्रकार की हैं।

(B) नकारात्मक (Negative Probability)

यदि कई स्वतंत्र रूप से घटित होने वाली घटनाओं की सम्भावना दी हुई हो और उनमें से कम से कम एक घटना के घटने की सम्भावना निकालनी हो तो ऐसी सम्भावना नकारात्मक सम्भावना के द्वारा निकाली जाती है। सूत्र की अभिव्यक्ति इस प्रकार होगी :

Probability of happening at least one of events $= 1 - P$ (happenning of none of the events)

or $= 1 - P(\bar{A}$ and $\bar{C})$

इसे निम्न उदाहरण द्वारा स्पष्ट किया जा सकता है :

Illustration 28 : सांख्यिकी में एक प्रश्न चार विद्यार्थियों A, B, C तथा D को दिया गया जिनकी प्रश्न हल करने की प्रायिकता क्रमशः 1/2, 1/3, 1/4 तथा 1/4 है। प्रश्न के हल होने की प्रायिकता बताएँ।

A problem is statistics is given to four students A, B, C and D. Their chances of solving if are 1/2, 1/3, 1/4 and 1/4. What is the probability that the problem will be solved ?

हल (Solution) : Probability that student 'A' fails to solve the problem $= 1 - 1/2 = 1/2$

Probability that student 'A' fails to solve the problem $= 1 - 1/3 = 2/3$

Probability that student 'A' fails to solve the problem $= 1 - 1/4 = 3/4$

Probability that student 'A' fails to solve the problem $= 1 - 1/4 = 3/4$

Since the events are independent the probability that all the four students fail to solve the problem is

$$= \frac{1}{2} \quad \frac{2}{3} \quad \frac{3}{4} \quad \frac{3}{4} \quad \frac{13}{16}$$

\ The probability that the problem will be solved $= 1 - \frac{3}{16} \quad \frac{13}{16}$

Illustration 29 : Let A and B be the two possible outcomes of an experiment and suppose :-

P (A) = 0.4, P(A or B) = 0.8 and P (B) = P

(*i*) For which choice of P are A and B mutually exclusive ?

(*ii*) For what choice of P are A and B independent ?

Solution : (*i*) We have

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

$$P(AB) = P(A) + P(B) - P(A \text{ or } B)$$

$$= 0.4 + P - 0.8$$
$$= P + 0.4 - 0.8$$
$$P(AB) = P - 0.4$$

If A and B are mutually exclusive

$$P(AB) = 0$$
$$0 = P - 0.4$$
$$P = 0.47$$

(*ii*) A and B are independent if

$$P(AB) = P(A)\,P(B)$$
$$P - 0.4 = 0.4 \times P$$
$$(1 - 0.4)\,P = 0.4$$
$$0.6\,P = 0.4$$
$$P = \frac{0.4}{0.6} = 0.67$$

Illustration 30 : Three coins are tossed simultaneously. What is the probability that they will fall 2 heads and 1 tail ?

तीन सिक्कों को एक साथ उछालने पर 2 पट व 1 चित्त आने की क्या प्रायिकता है ?

Solution : Let prob. of getting head in a throw of a coin is,

$$P(H) = \frac{1}{2}$$

Prob. of getting tail is,

$$P(T) = \frac{1}{2}$$

Two heads and one tail can come in either of the 3 ways,

(1) HHT \ P (H and H and T) $= \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{8}$

(2) HTH \ P (H and T and H) $= \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{8}$

(3) THH \ P (H and H and H) $= \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{8}$

(By applying Multiplication rule as throwing of coin is an independent event)

But all the 3 options are mutually exclusive so by applying addition rule :-

$$P(\text{I or II or III}) \text{ option} = P(\text{I}) + P(\text{II}) + P(\text{III}) \text{ option}$$
$$= \frac{1}{8} \quad \frac{1}{8} \quad \frac{1}{8} \quad \frac{3}{8}$$

Illustration 31 : A bag contains 10 red and 5 white balls. Four balls are drawn one by one without replacing the previous one. Find the probability that they are alternativley of different colour.

एक थैले में 10 लाल तथा 5 सफेद गेंदें हैं। चार गेंदें एक-एक करके निकाली, पिछली गेंद का वापिस नहीं रखा गया। चारों गेंदों के एक के बाद एक अलग रंग होने की प्रायिकता क्या है ?

Solution : 4 balls of alternative colours can be either-red, white, red, white or white, red, white, red

Ist option — If first ball is red

Total number of balls are = 10 red + 5 white = 15

Since events are dependent as drawing of ball is without replacement.

$$\text{Prob. of drawing first red ball} = P(A) = \frac{10}{15}$$

$$\text{Prob. of drawing second white ball} = P(B/A) = \frac{5}{14}$$

$$\text{Prob. of drawing third red ball} = P(C/AB) = \frac{9}{13}$$

$$\text{Prob. of drawing fourth white ball} = P(D/ABC) = \frac{4}{12}$$

(By using Rule of multiplication as all the 4 events are depdendent)

Prob. of (red and white and red and white) is :-

$$P(A \text{ and } B \text{ and } C \text{ and } D) = P(A) \times P(B/A) \times P(C/AB) \times (D/ABC)$$

$$= \frac{10}{15} \quad \frac{5}{14} \quad \frac{9}{13} \quad \frac{4}{12} \quad \frac{5}{91}$$

IInd option — If first drawn ball is white.

$$\text{Prob. of drawing first white ball} = P(A) = \frac{5}{15}$$

$$\text{Prob. of drawing II red ball} = P(B/A) = \frac{10}{14}$$

$$\text{Prob. of drawing III white ball} = P(C/AB) = \frac{10}{14}$$

$$\text{Prob. of drawing IV red ball} = P(D/ABC) = \frac{9}{12}$$

Since events are dependents so by using multiplication rule.

Prob. of (white and red and white and red) is :-

$$P(A \text{ and } B \text{ and } C \text{ and } D) = P(A) \times P(B/A) \times P(C/AB) \times P(D/ABC)$$

$$= \frac{5}{15} \quad \frac{10}{14} \quad \frac{4}{13} \quad \frac{9}{12} \quad \frac{5}{91}$$

Since both options are mutually exclusive so by applying rule of addition :-

$$P(\text{I or II}) \text{ options} = P(\text{I}) + P(\text{II}) \text{ option}$$

$$= \frac{5}{91} \quad \frac{5}{91} \quad \frac{10}{91}$$

Illustration 32 : An University has to appoint examiners to evaluate papers in statistics. Outof panel of 40 examiners, 10 are women; 30 out of them know Hindi and 5 of them are Ph.D. Find the probability of selecting a Hindi knowing Ph.D. women teacher to evaluate the papers.

एक विश्वविद्यालय को सांख्यिकी विषय के पेपर जँचवाने के लिए जाँचकर्त्ता चाहिए। 40 जाँचकर्त्ताओं की सूची में से 10 स्त्रियाँ हैं, 30 हिन्दी विषय जानते हैं व 5 पी. एच. डी. हैं। एक हिन्दी जानने वाली पी. एच. डी. स्त्री के चुन जाने की प्रायिकता बताएँ।

Solution : Total number of examiners are = 40

Prob. of selecting a woman examiner $= P(A) = \frac{10}{40}$

and, Prob. of selecting a Hindi knowing person $= P(B) = \frac{30}{40}$

and, Prob. of selecting a Ph.D examiner $= P(C) = \frac{5}{40}$

Since all attributes are independent so required probability can be obtained by applying multiplication rule —

Prob. of selecting a Hindi knowing, Ph.D., women is —

$$P(A \text{ and } B \text{ and } C) = P(A) \times P(B) \times P(C)$$

$$= \frac{10}{40} \quad \frac{30}{40} \quad \frac{5}{40} \quad \frac{3}{128}$$

Illustration 33 : A husband and wife appear in an interview for two vacancies in the same post. The probability of husband's selection is 1/7 and that of wife's selection is 1/5. What is the probability that

(*i*) both of them will be selected.

(*ii*) both of them will be rejected.

(*iii*) only one of them will be selected.

Solution : Let 'A' denotes husband's selection and 'B' denotes wife's selection.

(*i*) Both of them will be selected.

Probability that both of them will be selected shows that the events are independent.

\ $P(A \text{ and } B) = P(A) \times P(B)$

$$= \frac{1}{7} \quad \frac{1}{5} \quad \frac{1}{35}$$

(*ii*) None of them will be selected.

Probability of wife not being selected.

$$q = 1 - p$$

$$\overline{A} = 1 - \frac{1}{7} = \frac{6}{7}$$

Probability of wife not being selected is

$$\overline{B} = 1 - \frac{1}{5} = \frac{4}{5}$$

Probability that none of them will be selected.

$$P(\overline{A}) \times (\overline{B}) = \frac{6}{7} \quad \frac{4}{5} \quad \frac{24}{35}$$

(*iii*) Only of them is selected.

It means if 'A' is selected then B is rejected or if B is selected then A is rejected.

\ Probability that only of them will be selected.

$$= P[(A \text{ and } \overline{B}) \text{ or } (B \text{ and } \overline{A})]$$

$$= P(A \text{ and } \overline{B}) + P(B \text{ and } \overline{A})$$

$$= P(A) . P(\overline{B}) + P(B) . P(\overline{A})$$

$$= P(A). [(1 - P(B)] + P(B) [1 - P(A)]$$

$$= \frac{1}{7} \quad \frac{4}{5} \quad \frac{1}{5} \quad \frac{6}{7}$$

$$= \frac{4}{35} \quad \frac{6}{35} \quad \frac{10}{35} \quad \frac{2}{7}$$

Illustration 34 : 'A' speaks truth in 60 per cent cases and 'B' in 70 per cent cases. In what percentage of cases are they likely to contradict each other in standing the same fact ?

Solution : It is a case of independent events because both will contradict each other only if one speaks the truth and the other speaks a lie.

$$\text{Now given A speaks the truth} = \frac{60}{100}$$

$$\text{A speaks a lie} = 1 - \frac{60}{100} = \frac{40}{100}$$

$$\text{Similary 'B' speaks a lie} = 1 - \frac{70}{100} = \frac{30}{100}$$

The probability that 'A' speaks the truth and 'B' a lie

$$= \frac{60}{100} \quad \frac{30}{100} \quad \frac{9}{50}$$

Again, the probability that 'B' speaks the truth and 'A' a lie

$$= \frac{70}{100} \quad \frac{40}{100} \quad \frac{14}{50}$$

$$\text{The total probability} = \frac{9}{50} \quad \frac{140}{50} \quad \frac{23}{50}$$

The percentage of cases in which they contradict each other.

$$= \frac{23}{50} \times 100 = 45\%$$

Illustration 35 : A bag contains 5 red and 4 green balls. Another bas contains 4 red and 6 green balls. A ball is drawn from first bag and is placed in second. A ball is then drawn from second bag. What is the probability that it is red ?

एक थैले में 5 लाल व 4 हरी गेंदें हैं। दूसरे थैले में 4 लाल व 6 हरी गेंदें हैं। एक गेंद प्रथम थैले से निकाल कर दूसरे थैले में डाल दी और तब दूसरे थैले से एक गेंद निकाली। इस गेंद के लाल रंग के होने की क्या प्रायिकता है ?

Solution : The ball drawn from first bag and put in another bag may be either red or green.

(*i*) **If first ball drawn happens to be red.**

Probability of drawing a red ball from first bag is —

$$P(R_1) = \frac{5}{5 \quad 4} = \frac{5}{9}$$

Now, second bag possesses 5 red (4 + 1) and 6 green balls, so the probability of drawing red ball from second bag.

$$= P(R_2) = \frac{5}{11}$$

Since both the events are dependent so probability of drawing first red and drawing again red after transfering first one is —

$$P(\text{I red and II red}) = \frac{5}{9} \quad \frac{5}{11} \quad \frac{25}{99}$$

(*ii*) **If first ball drawn happens to be green**

$$\text{Prob. of drawing green ball from 1st bag} = P(G_1) = \frac{4}{9}$$

Now, second bag possesses 4 red and 7 green (6 + 1) balls, so the probability of drawing second red ball is —

$$= P(R_2) = \frac{4}{11}$$

Since events are dependent so probability of drawing first green and second red after transfering first balls is—

$$P(\text{I green and II red}) = \frac{4}{9} \quad \frac{4}{11} \quad \frac{16}{99}$$

Since both options are mutually exclusive so by applying addition rule —

$$P(\text{I or II option}) = P(\text{I}) + P(\text{II}) \text{ option}$$

$$= \frac{25}{99} \quad \frac{16}{99} \quad \frac{41}{99}$$

Illustration 36 : Four students A, B, C and D are given a problem to solve. Their chances of solving the problem are $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ and $\frac{1}{5}$ respectively. Find the probability that :-

(*i*) at least one will solve the problem.

(*ii*) only one will solve the problem.

(*iii*) only two will solve the problem.

Solution : Let

$P(A)$ ® A's probability of solving problem

$P(\overline{A})$ ® A's probability of not solving problem

$P(B)$ ® B's probability of solving problem

$P(\overline{B})$ ® B's probability of not solving problem

$P(C)$ ® C's probability of solving problem

$P(\overline{C})$ ® C's probability of not solving problem

$P(D)$ ® D's probability of solving problem

$P(\overline{D})$ ® D's probability of not solving problem

Now

$$P(A) = \frac{1}{2} \text{ so } P(\overline{A}) = 1 - \frac{1}{2} = \frac{1}{2}$$

$$P(B) = \frac{1}{3} \text{ so } P(\overline{B}) = 1 - \frac{1}{3} = \frac{2}{3}$$

$$P(C) = \frac{1}{4} \text{ so } P(\overline{C}) = 1 - \frac{1}{4} = \frac{1}{2}$$

$$P(D) = \frac{1}{5} \text{ so } P(\overline{D}) = 1 - \frac{1}{5} = \frac{2}{3}$$

(*i*) (at least one will solve the problem)

$$= 1 - \text{None will solve the problem.}$$

$$= 1 - P(\overline{A}) \times P(\overline{B}) \times P(\overline{C}) \times P(\overline{D})$$

$$= 1 - \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} \times \frac{4}{5}$$

$$= 1 - \frac{1}{5} = \frac{4}{5}$$

(*ii*) (only one solved the problem)

This probability can be found out by considering all the cases in which only one will solve the problem and other three can't solve the problem.

$$\text{So required probability} = P(A) \times P(\overline{B}) \times P(\overline{C}) \times P(\overline{D})$$

$$+ P(A) \times P(\overline{B}) \times P(\overline{C}) \times P(\overline{D})$$

$$+ P(A) \times P(\overline{B}) \times P(\overline{C}) \times P(\overline{D})$$

$$+ P(A) \frac{25}{60} \times P(\overline{B}) \times P(\overline{C}) \times P(\overline{D})$$

$$= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} \times \frac{4}{5} + \frac{1}{2} \quad \frac{1}{3} \quad \frac{3}{4} \quad \frac{4}{5}$$

$$= \frac{1}{5} \quad \frac{1}{10} \quad \frac{1}{15} \quad \frac{1}{20}$$

$$= \frac{12 \quad 6 \quad 4 \quad 3}{60}$$

$$= \quad = \frac{5}{12}$$

(*d*) आश्रित घटनाएँ (Dependent Events)

यदि एक घटना के घट जाने का प्रभाव, दूसरी घटना के घटने पर प्रभाव पड़ता है तो उन घटनाओं को आश्रित घटनाएँ (Dependent Events) कहा जाता है। उदाहरणार्थ, ताश की एक गड्डी में से बादशाह निकालने की सम्भावना $\frac{4}{52}$ या $\frac{1}{13}$ है, यदि पहली बार ताश निकालने में बादशाह निकल आता है तो दूसरी बार बादशाह निकलने की सम्भावना $\frac{3}{51}$ होगी। अतः पहले वाली घटना का प्रभाव दूसरी घटना पर पड़ा, ये आश्रित घटनाएँ हैं।

जब घटनाएँ आश्रित (Dependent हों, तो गुणन-प्रमेय का प्रयोग नहीं किया जा सकता। आश्रित घटनाओं में एक घटना के घटने का प्रभाव दूसरी घटना पर पड़ता है। 'B' घटना के घट जाने पर घटना 'A' के घटने की प्रायिकता को **शर्तयुक्त प्रायिकता** (Conditional Probability) कहते हैं।

यदि 'A' तथा 'B' दो आश्रित घटनाएँ हैं तो 'B' के घटने की शर्तयुक्त प्रायिकता निम्न प्रकार से ज्ञात की जाएगी :

$$P\,(B/A) = \frac{P\,(AB)}{P\,(A)} \text{ या } P\,(A) \times P\,(B/A)$$

तथा 'A' के घटने की शर्तयुक्त प्रायिकता :

$$P\,(B/A) = \frac{P\,(AB)}{P\,(A)} \text{ या } P\,(AB) = P\,(B) \times P\,(A/B)$$

आश्रित घटनाएँ : $P\,(ABC) = P\,(A) \times P\,(B/A) \times P\,(C/AB)$

Illustration 37 : The manufacturer of 'Rajdoot' scooter give choice to their customers to have either a double seated scooter or a single seated scooter. On analysis of the booked orders for those scooter, they find that 75% of their customers are men and 25% women. 80% of the men customer prefer double seated scooters and rest one seated. 90% of their women customer prefer one seated scooters and rest two seated scooters. In what proportion the manufacturers should manufacture these two scooters ?

Solution : **Probability-Table**

(*b*)

Happy	**Not Happy**	**Total**	
Request accepted	.24	.04	.28
Request not accepted	.36	.36	.72
Total	.6	.4	1

(*p*) of being happy and accepting the request = .6 × .4 = .24

(*p*) of not being happy and not accepting the request = .1 × .4 = .04

The chances of his accepting the request = .28

and the chances of his accepting the request when he is happy = .24

Hence the probability by his being happy having accepted the request = $\frac{.24}{.28} = \frac{6}{7} = 0.857$

(*c*) Men customers preferring two-seater = .78 × .8 = .600

Women customer preferring two-seater = .25 × .1 = .025

Total = .625

Men customers preferring one-seater = .75 × .2 = .150

Women customer preferring one-seater = .25 × .9 = .225

Total = .625

Ratio .625 : .375 or 5 : 3

Illustration 38 : 52 पत्तों की एक गड्डी में से चार पत्ते एक-एक करके निकाले गए और निकाले गये पत्ते वापिस नहीं रखे गये तो प्रायिकता बताएँ कि

(*i*) वे सभी पत्ते इक्के हैं तथा **(*i*) वे सभी पत्ते अलग-अलग सूट कें हैं।**

Four cards are drawn from a pack of 52 cards without replacement. What is the probability that

(*i*) they all are aces ? and **(*ii*) they all are of different suits ?**

Solution.

(*i*) Since the events are dependent the required probability is

$$= \frac{4}{52} \quad \frac{3}{51} \quad \frac{2}{50} \quad \frac{1}{49} = \frac{1}{2,70,725}$$

(*ii*) Since the events are dependent the required probability is

$$= \frac{52}{52} \quad \frac{39}{51} \quad \frac{26}{50} \quad \frac{13}{49} \quad \frac{2,197}{20,825}$$

सम्भावना सिद्धान्त में क्रमचय तथा संयोग

(Permutations and Combinations in the Theory of Probability)

क्रमचय (Permutations)

कभी-कभी यह जानने की आवश्यकता नहीं पड़ती है कि पदों को कितने प्रकार से विन्यसित (Arrange) किया जा सकता है जबकि कोई भी दो विन्यास एक से न होने पायें। इस प्रकार के अनुविन्यास को क्रमचय (Permutation) कहते हैं। उदाहरणार्थ, सात अक्षरों — A, B, C, D, E, F, G, H के दो-दो अक्षरों के जोड़े से अनुविन्यसित करना हो जबकि किसी भी जोड़े में दोनों एक ही अक्षर नहीं होने चाहिए (जैसे AA, BB आदि) तो निम्नलिखित क्रमचय सम्भव होंगे :-

AB	AC	AD	AE	AF	AG
BA	BC	BD	BE	BF	BG
CA	CB	CD	CE	CF	CG
DA	DB	DC	DE	DF	DG
EA	EB	EC	ED	EF	EG
FA	FB	FC	FD	FE	FG
GA	GB	GC	GD	GE	GF

सूत्र के रूप में :-

$$"P_r = \frac{n!}{(n \quad r)!}$$

Where,

$"P$ = Permutation

n = number of items

r = number of items taken at a time

! = factorial *i.e.* $n\,(n-1)\,(n-2)$ ......

उपरोक्त उदाहरण में क्रमचय होंगे।

$${}^7P_2 = \frac{7!}{(7 \quad 2)!} = \frac{7!}{5!}$$

Illustration 39 : (*a*) There are six doors in a room. Four persons have to enter it. In how many ways they can enter from different doors ?

Perm. $= n\,(n-1)\,(n-2)\,(n-3)$

$= 6\,(6-1)\,(6-2)\,(6-3)$

$= 6 \times 5 \times 4 \times 3 = 360$

$n = 6, r = 4$

$$\frac{n!}{(n \quad r)!} = \frac{6!}{(6 \quad 4)!}$$

$${}^6p_4 = \frac{n!}{(n \quad r)!}$$

$$= \frac{6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}{2 \quad 1}$$

$= 6 \times 5 \times 4 \times 3 = 360$

(*b*) In how many ways first, second and third prizes can be distributed to three of ten competitors ?

$n = 10, r = 3$

$${}^{10}p_3 = \frac{10!}{(10 \quad 3)!}$$

$$= \frac{10 \quad 9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}{7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}$$

$$= 10 \times 9 \times 8 = 207 \text{ ways}$$

(*c*) Four strangers broad a bus in which there are 6 empty seats. In how many different ways can they be seated ?

$n = 6, r = 4$

$$^{6}p_{4} = \frac{n!}{(n \quad r)!} = \frac{6!}{(6 \quad 4)!}$$

$$= \frac{6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}{2 \quad 1}$$

$$= 6 \times 5 \times 4 \times 3$$

$$= 360 \text{ ways}$$

(*d*) In how many ways can 12 seats be occupied by 6 students ?

$n = 12, r = 6$

$$^{12}p_{6} = \frac{n!}{(n \quad r)!} = \frac{12!}{(12 \quad 6)!}$$

$$= \frac{12 \quad 11 \quad 10 \quad 9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}{6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}$$

$$= 12 \times 11 \times 10 \times 9 \times 8 \times 7 = 665280$$

(2) यदि किसी प्रक्रिया 'n' में एक समय में 'p' का प्रकार का 'q' दूसरी प्रकार का था 'r' तीसरी प्रकार के कार्य हैं तो क्रमचयों की संख्या $= \dfrac{n!}{p!q!r!}$ होगी।

Illustration 40 : In how many ways can the word 'ERROR' be arranged ?

शब्द ERROR के कितने विन्यास हो सकते हैं ?

Solution : The word 'ERROR' has 5 words and '*r*' come 3 times, '*e*' is coming once and '*o*' is also coming once.

$$^{n}p_{r} = \frac{n!}{n_1! \quad n_1! \quad n_2! \quad \ldots\ldots} = \frac{5!}{3! \quad 1! \quad 1!}$$

$$= \frac{5\times4\times3\times2\times1}{3\times2\times1} = 20 \text{ ways}$$

संयोग
(Combination)

क्रमचय (Permutations) : में वर्गित पदों के अनुविन्यास का महत्त्व होता है; संयोग में अनुविन्यास महत्त्वहीन होता है। ''संयोग में पदों का अनुविन्यास महत्त्वहीन होता है परन्तु पदों की पुनरावत्ति नहीं होनी चाहिए।''

यदि A B C D E अक्षरों को दो-दो अक्षरों के जोड़े से रखा जाय तो क्रमचय 20 होंगे क्योंकि

Perm = $n\,(n-1) = 5\,(5-1) = 20$ परन्तु संयोग केवल 10 ही होंगे :

$$\left.\begin{array}{cccc} AB & AC & AD & AE \\ BA & BC & BD & BE \\ CA & CB & CD & CE \\ DA & DB & DC & DE \\ EA & EB & EC & ED \end{array}\right\} \text{क्रमचय}$$

$$\left.\begin{array}{llll} AB & AC & AD & AE \\ BA & BD & BE & \\ CD & CE & & \\ DE & & & \end{array}\right\} \text{संयोग}$$

संयोग का सूत्र है —

$$^{n}C_{r} = \frac{n!}{r!(n \quad r)!} \; n = 5, r = 2$$

$$= \frac{5!}{2!(5 \quad 2)!} = \frac{5 \quad 4 \quad 3 \quad 2 \quad 1}{2 \quad 1 \quad 3 \quad 2 \quad 1} = 10$$

$$^{n}C_{r} = \frac{^{n}p_{r}}{r!} \quad \frac{20}{2\times1} \quad 10$$

इस सूत्र में $"C_r$ = number of combination with r things at a time

$"p_r$ = number of permutations with r things at a time

$r!$ = factorial 'r' or number of permutations within a combination

Illustration 41 : A shopkeeper decides that he has surplus stocks of 9 different kind of goods. He decided to make up gifts parcels, each to contain four different goods, which he will sell at a special price. How many different kinds of parcels may be produce ?

Solution : $n = 9, r = 4$

$$C = \frac{n!}{r!(n \quad r)!} = \frac{9!}{4!(9 \quad 4)!}$$

$$= \frac{9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1}{4 \quad 3 \quad 2 \quad 1 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1} = 126$$

Illustration 42 : A और B एक खेल खेलते हैं। A को सर्वप्रथम पासा फेंकना है और नम्बर 6 अंक आने पर जीतेगा। A के हारने पर B पास फेंकेगा और नम्बर 6 या 5 अंक आने पर जीतेगा। B के हारने पर पुनः A पासा फेंकेगा और नम्बर 6 या 5 या 4 आने पर वह जीतेगा। इसी प्रकार यह क्रम चलेगा। A और B के इनाम जीतने की क्रमशः प्रायिकता ज्ञात कीजिए।

A and B play for a prize of Rs. 2,000. A is to throw a dice first and is to win it he throws 6. If he fails B is to throw and he is to win if he throws 6 or 5. If he fails A to throw again and to win if he throws. 6, 5 or 4 and so on. What are their respective probabilities of winning ?

Solution :

Probability of A's winning in the first throw $= \frac{1}{6}$

Probability of B's winning in the second throw $= \frac{5}{6} \quad \frac{2}{6} \quad \frac{10}{36} \quad \frac{5}{18}$

Probability of A's winning in the third throw $= \frac{5}{6} \quad \frac{4}{6} \quad \frac{3}{6} \quad \frac{60}{216} \quad \frac{5}{18}$

Probability of B's winning in the fourth throw $= \frac{5}{6} \quad \frac{4}{6} \quad \frac{3}{6} \quad \frac{4}{6} \quad \frac{240}{1296} \quad \frac{5}{27}$

Probability of A's winning in the fifth throw $= \frac{5}{6} \quad \frac{4}{6} \quad \frac{3}{6} \quad \frac{2}{6} \quad \frac{5}{6} \quad \frac{600}{7776} \quad \frac{25}{324}$

Probability of B's winning in the sixth throw $= \frac{5}{6} \quad \frac{4}{6} \quad \frac{3}{6} \quad \frac{2}{6} \quad \frac{1}{6} \quad \frac{6}{6}$

$$\frac{720}{46656} = \frac{5}{324}$$

$$\text{A's total chances of success} = \frac{1}{6} \quad \frac{5}{18} \quad \frac{25}{324} \quad \frac{169}{324}$$

$$\text{B's total chances of success} = \frac{5}{18} \quad \frac{5}{27} \quad \frac{5}{324} \quad \frac{155}{324}$$

Illustration 44 : A bag contains 10 red, 6 white and 4 black balls. 4 balls are drawn. What is the probability of (*i*)All are white (*ii*) only two red balls(*iii*) 2 black and 1 red ball.

Solution : Number of ways in whcih 4 balls can be drawn out of 20 balls = $^{20}C_4$

$$= \frac{20!}{4!6!} = \frac{20 \quad 19 \quad 18 \quad 17 \quad 16!}{4 \quad 3 \quad 2 \quad 16!} = 4845$$

(*i*) Number of ways in which 4 white balls can be drawn out of 6 white balls = $^{6}C_4$

$$= \frac{6!}{2!4!} = \frac{6 \quad 5 \quad 4!}{2 \quad 4!} = 15$$

\ Required probability = $\frac{15}{4845} \quad \frac{1}{325}$

(*ii*) Probability of getting two red balls = $\frac{^{10}C_2 \quad ^{10}C_2}{^{20}C_4}$

$$= \frac{\frac{^{5}10 \quad 9}{4 \quad 2} \quad \frac{^{5}10 \quad 9^{3}}{2}}{\frac{10 \quad 9}{2 \quad 4845}} \quad \frac{63}{1615} \quad 2 \quad \frac{72}{223} \quad \frac{10 \quad 9}{2} \quad \frac{1}{4845} \quad \frac{135}{323}$$

$$223 \qquad 223 \qquad 969 \qquad 323$$

(*iii*) Probability of getting 2 black and 1 red ball = $\frac{^{4}C_1 \quad ^{10}C_2 \quad ^{6}C_1}{^{20}C_4}$

=

प्रतिलोम प्रायिकता/बेयज प्रमेय

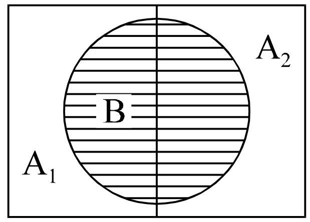
(Inverse Probability/Bayes Theorem)

प्रतिलोम प्रायिकता में हम विपरीत प्रकार की समस्याओं का वर्णन करेंगे। हम यह जानते हैं कि अमुक घटना अनेक कारणों में से किसी एक निश्चित कारण का प्रभाव है। यह ज्ञात करना कि घटना के उस निश्चित कारण का प्रभाव होने की क्या प्रायिकता है - **प्रतिलोम प्रायिकता** कहलाती है। उदाहरण के लिए यदि किसी थैले में 12 काली एवं 7 हरी गेंदें हैं तथा दूसरे थैले में 9 काली एवं 6 हरी गेंदें हैं तथा किसी एक थैले में से 1 काली गेंद निकाली जाये तो यह ज्ञात करना कि उस गेंद के पहले थैले में से निकाले जाने की क्या प्रायिकता होगी, यह **प्रतिलाम प्रायिकता** से सम्बन्धित प्रश्न है।

बेयज प्रमेय (Bayes's Theorem)

बेयज प्रमेय का नाम अंग्रेज गणितज्ञ थॉमस बेयज (Thomas Bayes) के नाम पर पड़ा। इस प्रमेय का प्रचलन 1763 में उनके शोध बेयजियन निर्णय प्रमेय (Bayesian Decision Theory) के आने के बाद ही हुआ। इस प्रमेय को और भी कई अन्य नामों से जाना जाता है। जैसे — नई प्रयोगात्मक संभावना (Posterior Probability) आदि। यदि हमें पारस्परिक अपवर्जी

(Mutually Exclusive) व सर्वग्राही घटना (Exhaustive Events) 'A' दी हो तथा एक दूसरे निदर्श (Sample) से घटना 'B' दी हो तो प्रथम घटना 'A' को इस प्रकार काटती है जैसा कि निम्न चित्र में दर्शाया गया है :



उपरोक्त चित्र में भाग (B) जो कि A_1 का ही भाग है (A_1 व B) को प्रदर्शित कर रहा है और इसी प्रकार भाग (B) जो कि A_2 का ही भाग है, (A_2 व B) को प्रदर्शित कर रहा है।

\ Probability of event A_1 given event B, is

$$P(A_1/B) = \frac{P(A_1 \text{ and } B)}{P(B)}$$

Similarly, probability of event A_2 and B, is

$$P(A_2/B) = \frac{P(A_2 \text{ and } B)}{P(B)}$$

Where
$$P(B) = P(A_1 \text{ and } B) + P(A_2 \text{ and } B)$$
$$P(A_1 \text{ and } B) = P(A_1) \times P(B/A_1) \text{ and}$$
$$P(A_2 \text{ and } B) = P(A_2) \times P(B/A_2)$$

जब हमें घटना (A) की संभावना दी हुई हो तथा एक ही संभावना जो या तो प्रयोग पर आधारित हो अथवा भूतकालीन अनुभव पर आधारित हो, जिसे हमें घटना (B/A) कहते हैं, इन्हें तर्क संगत संभावना (Priori Probability) कहते हैं। जब उपरोक्त संभावनाओं का बेयज प्रमेय द्वारा परिवर्तित (Revised) कर दिया जाता है तो इसे नई प्रयोगात्मक संभावना (Posterior) कहते हैं क्योंकि ऐसा करते समय नई सूचनाओं को भी ध्यान में रखा जाता है। इसे परिवर्तित संभावना (Revised Probability) इसलिए कहा जाता है क्योंकि इसे तर्क संगत संभावना (Priori Probability) को नई सूचना के आधार पर परिवर्तित करके ही ज्ञात किया जाता है। इसे विपरीत संभावना (Inverse Probability) इसलिए कहा जाता है क्योंकि इसमें हम संभावना की ही संभावना ज्ञात करते हैं।

बेयज प्रमेय का उपयोग उद्योग, व्यापार व वाणिज्य में अनिश्चितता होते हुए भी निर्णय लेने के लिए किया जाता है। इस प्रमेय के द्वारा खराब वस्तु अथवा उच्च किस्म की वस्तु की संभावना के बारे में जाना जा सकता है।

तर्क संगत व परिवर्तित संभावना (Priori and Posterior Probability)

जब हमें किसी घटना (A) की संभावना तथा किसी अन्य घटना (B/A) की संभावना जो कि वास्तविक प्रयोग पर आधारित हो अथवा भूतकालीन अनुभव पर आधारित हो तो ऐसी संभावनाओं को तर्क संगत संभावना (Priori Probability) कहते हैं। क्योंकि इन संभावनाओं को बेयज प्रमेय द्वारा परिवर्तित नहीं किया जाता। अतः इन संभावनाओं को अपरिवर्तित (pre-revised) अथवा तर्कसंगत संभावना (priori probability) कहा जाता है।

जबकि अपरिवर्तित अथवा तर्कसंगत संभावना को जब बेयज प्रमेय (Baye's Theorem) द्वारा संशोधित कर दिया जाता है तो इसे परिवर्तित अथवा विपरीत संभावना (Inverse Probability) कहते हैं। परिवर्तित संभावना सदैव ही शर्तयुक्त संभावना (Conditional Probability) होती है क्योंकि निदर्शन सूचना कुछ शर्तों पर आधारित होती है। अतः तर्कसंगत संभावना सदैव ही अथर्तयुक्त (Unconditional) अथवा अपरिवर्तित (Unrevised) होती है जबकि परिवर्तित संभावना सदैव ही शर्तयुक्त होती है और इसी शर्त के अनुरूप बेयज प्रमेय द्वारा इसे परिवर्तित (Revised) किया जाता है।

निम्न उदाहरण से बेयज प्रमेय का प्रयोग देखा जा सकता है :

Illustration 44 : An Insurance Company insured 2000 scooter drivers, 4000 car driver and 6000 truck drivers. The probability of an accident involving a scooter driver, car driver and a truck driver is 0.01, 0.03, 0.15 respectively. One of the insured driver meets with an accident. What is the probability that he is a car driver ?

एक बीमा कम्पनी ने 2000 स्कूटर चालक, 4000 कार चालक व 6000 ट्रक चालकों का बीमा कराया। स्कूटर चालक, कार चालक व ट्रक चालक के दुर्घटनाग्रस्त होने की प्रायिकता क्रमशः 0.01, 0.03 व 0.15 है। यदि एक बीमित चालक दुर्घटनाग्रस्त हुआ तो उसके कार चालक होने की क्या प्रायिकता है ?

Solution : Let A, B and C be the events that the insured person is a scooter driver, a car driver and a trukc driver respectively and let D be the event that the insured person meets an accident.

We are given the information :

$$P\,(A) = \frac{2000}{2000+4000+6000} = \frac{2000}{12000} = \frac{1}{6}$$

$$P\,(B) = \frac{4000}{2000+4000+600} \quad \frac{4000}{12000} \quad \frac{1}{3}$$

$$P\,(C) = \frac{6000}{2000 \quad 4000 \quad 6000} \quad \frac{6000}{12000} \quad \frac{1}{2}$$

The conditional probabilities are :-

$$P\,(D/A) = 0.01 = \frac{1}{100}$$

$$P\,(D/B) = 0.03 = \frac{3}{100}$$

$$P\,(D/C) = 0.15 = \frac{15}{100}$$

Putting the information in the table given below :

Events **(1)**	**Prior Probabilities** **(2)**	**Conditional Probablities** **(3)**	**Joint Probabilities** **(2 × 3)**
A	$P\,(A) = \frac{1}{6}$	$P\,(D/A) = \frac{1}{100}$	$\frac{1}{6} \quad \frac{1}{100} \quad \frac{1}{600}$
B	$P\,(B) = \frac{1}{3}$	$P\,(D/B) = \frac{3}{100}$	$\frac{1}{3} \quad \frac{3}{100} \quad \frac{1}{100}$
C	$P\,(C) = \frac{1}{2}$	$P\,(D/C) = \frac{15}{100}$	$\frac{1}{2} \quad \frac{15}{100} \quad \frac{15}{200}$

We have to find P (B/D) *i.e.* the probability of a car driver given that he meets an accident.

By Bayes Theorem,

$$P\,(B/D) = \frac{P\,(B) \quad P\,(D/B)}{P\,(A) \quad P\,(D/A) \quad P\,(B) \quad P\,(D/B) + P\,(C) \quad P\,(D/C)}$$

$$= \frac{\frac{1}{100}}{\frac{1}{600} \quad \frac{1}{100} \quad \frac{15}{200}} = \frac{\frac{1}{100}}{\frac{1 \quad 6 \quad 45}{600}}$$

$$= \frac{\frac{1}{100}}{\frac{52}{600}} = \frac{1}{100} \quad \frac{600}{52}$$

$$= \frac{6}{52} = \mathbf{0.115}$$

Illustration 45 : The probability of selection of 3 persons for the post of a Principal in a newly started college is 4 : 3 : 2. The probability that they will introduce co-education in the college is 0.2, 0.3 and 0.5 respectively. Find the probability of selection of second person as principal of the college if it is given that co-education will be introduced in the college.

Solution : Let events

A_1 ® 1st person is selected

A_2 ® 2nd person is selected

A_3 ® 3rd person is selected

B ® Co-education introduced in the college.

Now $P(A_1) = \frac{4}{9}$ $P(B/A_1) = 0.2$ $P(A_1B) = .8/9$

$P(A_2) = \frac{3}{9}$ $P(B/A_2) = 0.3$ $P(A_2B) = .9/9$

$P(A_3) = \frac{2}{9}$ $P(B/A_3) = 0.5$ $P(A_3B) = 1.0/9$

$\frac{4}{9}$

$$P(B) = \frac{2.7}{9} = 0.3$$

$$3P(A_1B) = P(A_1) \times P(B/A_1) = \quad \times 0.2 = 0.8/9$$

$$P(A_2B) = P(A_2) \times P(B/A_2) = \frac{3}{9} \times 0.3 = 0.9/9$$

$$P(A_3B) = P(A_3) \times P(B/A_3) = \frac{2}{9} \times 0.5 = 1.0/9$$

Now $P(A_2/B) = \frac{P(A_2B)}{P(B)} = \frac{.1}{.3} = 1/3$

सारांश

(Summary)

- संभावना भविष्य में घटने वाली किसी घटना के घटने की उम्मीद दर्शाती है।
- यदि हम किसी घटना या न घटने के बारे में 100% आश्वस्त हैं तो सम्भावना का मान क्रमशः एक व शून्य होगा। सम्भावना की ये अधिकतम तथा न्यूनतम सीमाएँ हैं।
- सम्भावना के प्रश्नों को हल करने से पहले कुछ आधारभूत धारणाओं को समझाना बहुत जरूरी है।

- जब भी कोई प्रयोग या परीक्षण किया जाता है तो उसके हमें कुछ परिणाम मिलते हैं। इन परिणामों को हम कई तरह से विभाजित कर सकते हैं।

  (*i*) सरल घटना (एक समय में एक घटना घटित होना) तथा (*ii*) संयुक्त घटना (एक समय में एक से ज्यादा घटनाओं का घटित होना) उदाहरण के लिए एक सिक्के को उछालने पर चित्त या पट पाना एक सरल घटना है। लेकिन दो सिक्कों का एक साथ उछालना संयुक्त घटना है।

- दूसरी वर्गीकरण समान रूप से घटने वाली घटनायें हैं। यदि प्रयोग दैव है तो परिणाम समान रूप से घटने वाली घटनाएं हैं जैसे एक सिक्का उछालने पर चित्त या पट मिलना। लेकिन यदि प्रयोग दैव नहीं हैं तो घटनाएं असमान रूप से घटने वाली घटनाएं हैं। यदि एक कार्यालय में 10 रिक्तियों के लिए 100 उम्मीदवार आवेदन करते हैं तो उनका चयन समान रूप से घटने वाली घटनाएं नहीं हैं।

- तीसरी तरह का वर्गीकरण पारस्परिक अपवर्जी घटनाएं हैं। ये वे घटनाएँ हैं जिनके घटने पर अन्य घटनाओं के घटने की सम्भावनाएं समाप्त हो जाती हैं। जैसे सिक्के के उछाल में यदि हमें चित्त मिलता है तो पट मिलने की सम्भावना खत्म हो जाती हैं। लेकिन यदि एक घटना के घटने से दूसरी घटना के घटने की सम्भावना खत्म नहीं होती है तो ये वर्गीकरण पारस्परिक अपवर्जी घटनाएँ नहीं हैं। जैसे ताश की गड्डी में से अगर हम एक पत्ता निकालते हैं तो वो बादशाह हो तथा काले रंग का भी हो तो यो घटनाएं अपारस्परिक अपवर्जी घटना हैं क्योंकि दोनों इकट्ठी घट सकती हैं, जैसे चिड़ी तथा हुकम का बादशाह।

- चौथा वर्गीकरण स्वतंत्र तथा आश्रित घटनाएं हैं। यदि एक प्रयोग के परिणामों का दूसरे प्रयोग के परिणामों पर कोई असर नहीं पड़ता है तो इन्हें हम स्वतंत्र घटनाएं कहते हैं। जैसे एक पासे के फेंकने से जो परिणाम मिलेगा वो दूसरा पासा फेंकने का परिणाम इस परिणाम पर निर्भर नहीं है। दूसरी तरफ यदि प्रथम परिणाम दूसरे प्रयोग की संभावित घटनाओं के घटने की संभावना पर असर डालता है तो उन्हें आश्रित घटनाएं कहते हैं। जैसे एक बैग में 6 सफेद तथा 4 लाल गेंदें हैं। हम दो गेंदे निकालते हैं। दूसरी गेंद के लाल होने की संभावना का मान प्रथम गेंद के रंग पर निर्भर करता है। यदि प्रथम गेंद लाल है तो ये संभावना 3/9 है और यदि प्रथम गेंद सफेद है तो यह संभावना 4/9 है। इस तरह दूसरी गेंद का लाल होना आश्रित घटना है। आश्रित घटनाओं के लिए हम शर्तयुक्त सम्भावना निकालते हैं।

- पारस्परिक अपवर्जी घटनाओं के घटने की संभावना निकालने के लिए हम योग का नियम प्रयोग करते हैं तथा स्वतंत्र घटनाओं के लिए हम गुणा का नियम प्रयोग करते हैं। कुछ विशेष प्रश्नों को हल करने के लिए बेयज प्रमेय का प्रयोग करते हैं।

- कठिन प्रश्नों को हल करने में क्रमचय व संयोग का भी प्रयोग करते हैं।

प्रश्नावली

(Exercise)

(1) प्रायिकता की परिभाषा बताएँ और सांख्यिकी में इसका महत्त्व बताइए।

Define probability and explain the importance of this concept in Statistics.

(2) संभाव्यता में क्लॉसिकी तथा आनुभाविक उपागमों की विस्तारपूर्वक विवेचना कीजिए।

Discuss in details the Classical and Empirical approaches to probability.

(3) प्रायिकता की विभिन्न धारणाओं की व्याख्या कीजिए। क्या ये एक-दूसरे की विरोधी हैं ?

Explain the various approached to probability. Are they contradictory ?

(4) Explain the following :-

निम्नलिखित की व्याख्या कीजिए :-

(*a*) Mutually exclusve events.	(क) परस्पर अपवर्जी घटनाएं
(*b*) Equally likely events.	(ख) समान रूप से संभव घटनाएं
(*c*) Independent events.	(ग) स्वतंत्र घटनाएं
(*d*) Dependent events.	(घ) पराश्रयी घटनाएं

(5) Explain the examples the concept of independent and mutually exclusive events.

पारस्परिक अपवर्जी एवं स्वतंत्र घटनाएं उदाहरण देकर समझाइए।

(6) (*a*) प्रायिकता में स्वतंत्र एवं आश्रित घटनाओं की धारणाओं की व्याख्या कीजिए।

Explain the concepts of Independent and dependent events in probability.

(7) निम्नलिखित को उचित उदाहरण देकर व्याख्या कीजिए :

(*i*) समान रूप से घटने वाली घटनाएं।

(*ii*) प्रायिकता सिद्धान्त में क्रमचय तथा संयोग।

(*iii*) योग तथा गुणन प्रमेय।

Explain the following with suitable examples :-

(*i*) Equally likely events.

(*ii*) Permutation and combination in the theory of probability.

(*iii*) Addition and multiplication theorems of probability.

(8) Explain any two difference between the following concepts :

(*i*) Simple and Compound events.

(*ii*) Mutually exclusive and independent Events.

(*iii*) Permutation and combinations.

(9) शर्तयुक्त प्रायिकता को उदाहरण सहित समझाइए।

Explain conditional probability with example.

(10) आप 'प्रायिकता' शब्द से क्या समझते हैं ? समझाइए। योग व गुणन प्रमेय समझाइए व सिद्ध कीजिए।

Explain what do you understand by the term 'Probability' ? State and prove the addition and multiplication theorems of probability.

(11) बेयज प्रमेय बताओ व सिद्ध करो।

State and prove Baye's theorem.

(12) पूर्व (तर्क संगत) व उत्तर (परिवर्तित) संभावना से आपका क्या अभिप्राय है ? बतायें कि परिवर्तित संभावनायें बेयज प्रमेय के प्रयोग से कैसे निकालते हैं ? बेयज प्रमेय के व्यवसाय में क्या प्रयोग है ?

What do you understand by priori and posterior probabilities ? Show how do we calculate posterior probabilities using Bayes' Theorem. What are the business applications of Bayes' Theorem ?

(13) 52 पत्तों की एक ताश की गड्डी से एक पत्ता निकालने पर बादशाह आने की प्रायिकता क्या हैं ?

What is the probability of drawing a king in a draw from a pack of 52 cards ?

[**Ans.** 1/13]

(14) एक पासा जिस पर 1 से 6 तक अंक लिखे गये हैं, 1 या 2 प्राप्त होने की प्रायिकता क्या है ?

If a dice having six sides and marked 1 to 6 is thrown, what is the probability of getting 1 or 22 ?

[**Ans.** 1/3]

(15) एक थैले में 5 लाल, 8 सफेद तथा 10 काली गेंदें हैं। एक गेंद निकाली गई। वह सफेद या काली होगी, इसकी प्रायिकता क्या है ?

A bag contains 5 red balls, 8 white balls and 10 black balls. A ball is drawn. What is the probability that the ball drawn will be either white or black ? [**Ans.** 18/23]

(16) एक थैले में 20 गेंदें हैं जिन पर 1 से 20 तक नम्बर लिखे हैं। एक गेंद यादच्छिक रूप से चुनी गई। इस गेंद के 3 या 5 के गुणांक होने की प्रायिकता क्या है ?

A bag contains 20 balls numbered from 1 to 20. One ball is drawn at random. Find the probability that a ball drawn is a multiple of 3 or 5. [**Ans.** 9/20]

(17) एक परीक्षा में (अ) के सफलता के अवसर 5/40, (ब) के 7/50 तथा (स) के 10/60 हैं। इनमें से एक के सफलता के अवसर क्या है। ?

In an examination the chance of A's success is $\frac{5}{40}$. B's success $\frac{7}{50}$ and C's success $\frac{10}{60}$. Find the chance that one of them will succeed. [**Ans.** 259/600]

(18) एक थैले में 25 गेंदें है जिन पर 1 से 25 तक नम्बर लिखे हैं। एक गेंद यादच्छिक रूप से चुनी गई। इस गेंद के

(*i*) 5 या 6 (*ii*) 5 या 7 के गुणक होने की प्रायिकता क्या है ?

A bag contains 25 balls numbered from 1 to 25. One ball is drawn at random. Find the probability that the number is a multiple of

(*i*) 5 or 6 (*ii*) 5 or 7 [**Ans.** (*i*) 9/25, (*ii*) 8/25]

(19) आकस्मिक प्रवरण रीति द्वारा चुने एक बिना लीप वर्ष में 53 रविवार होने की क्या प्रायिकता है ?

What is the chance that a non-leap year selected at random should have 53 Sundays ?

[B.Com., K.U., 1992 Sept.] [**Ans.** 1/7]

(20) एक अनिभिनत पासे को फेंकने पर विषम अंक होने की क्या प्रायिकता है ?

What is probability of getting an odd number in a throw of an unbiased die. [**Ans.** 1/2]

(21) What is the probability of getting a spade card in a draw from the pack of 52 cards ? [**Ans.** 1/4]

(22) When a soldier fires at a target, the probability that he hits the target is $\frac{1}{3}$ for soldier A, $\frac{1}{6}$ for soldier B, $\frac{1}{6}$ for soldier C and $\frac{1}{12}$ for soldier D. If all the four soldiers A, B, C and D fire at the target simultaneously calculate the probability that the target is hit by some one of more.

$$\left[(b)\frac{378}{648}\right]$$

(23) Estimate the probability of securing the total 5 from two dice thrown simultaneously. $\left[\frac{1}{9}\right]$

(24) One bag contains 4 white balls and 2 black balls. Another contains 3 white balls and 5 black balls. If one ball is drawn from each bag, find the probability that (*a*) both are white, (*b*) both are black, and (*c*) one is white and one is black. [**Delhi.** M.A., 1964]

$$\left[(a)\frac{1}{4},(b)\frac{5}{24},(c)\frac{13}{24}\right]$$

(25) Find the chance of drawing a king, a queen and a knave in that order from a pack of 52 cards in three consecutive draws, the cards drawn not being replaced.

$$\left[\frac{4}{52} \quad \frac{4}{51} \quad \frac{4}{50}\right]$$

(26) एक कक्षा में 100 विद्यार्थी हैं जिनमें से 55 क्रिकेट खेलते हैं, 40 हॉकी खेलते हैं व 35 फुटबाल खेलते हैं। 20 विद्यार्थी क्रिकेट तथा हॉकी खेलते हैं। 15 विद्यार्थी हॉकी तथा फुटबाल खेलते हैं व 10 बच्चे फुटबाल तथा क्रिकेट खेलते हैं। 3 विद्यार्थी तीनों खेल खेलते हैं। ऐसे कितने विद्यार्थी हैं जो कोई भी खेल नहीं खेलते।

In a class of 100 students, 55 play cricket, 40 plays hockey, and 35 play football. 20 students play cricket and hockey, 15 play hockey and football, 10 play football and cricket while three players play all the three games. Find the number of students who do not play any game.

(27) If there are 2 children on an average in a family, find the probability of :-

(*i*) having one son (*ii*) atleast 1 daughter **[Ans.** (*i*) 1/2, (*ii*) 3/4]

यदि प्रत्येक परिवार में औसतन 2 बच्चे हों तो

(*i*) एक लड़का (*ii*) कम से कम एक लड़की होने की प्रायिकता बतायें।

(28) The data concerning lecturers and their academic qualification is given as under :-

	M.A.	**Ph.D.**	**Total**
Promoted	50	10	60
Not Promted	25	15	40
Total	75	25	100

एक प्रवक्ता को प्रवरण रीति द्वारा चुना हो तो प्रायिकता बतायें कि –

(*i*) वह पी.एच.डी. है।

(*ii*) वह पदोन्नत है।

(*iii*) वह प्रवक्ता जो पी.एच.डी. है को पदोन्नत किया।

(*iv*) एक पदोन्नत प्रवक्ता पी.एच.डी. है।

(*v*) एक प्रवक्ता जो पदोन्नत नहीं है।

A lecturer is selected at random, find the probability that —

(*i*) He is Ph.D.

(*ii*) He is promoted

(*iii*) A lecturer who is Ph.D. is promoted.

(*iv*) A promoted lectuter is Ph.D.

(*v*) A lecturer is not promoted.

[Ans. (*i*) 1/4, (*ii*) 3/5, (*iii*) 2/5, (*iv*) 1/6, (*v*) 2/5]

(29) एक बैग में 12 सफेद, 8 लाल, 6 नीली तथा 4 हरी गेंदें हैं। एक के बाद चार गेंदे निकाली जाती हैं। प्रायिकता बतायें कि –

(*a*) 2 गेंदें सफेद हैं।

(*b*) चारों रंगों की एक–एक गेंद है (*i*) ऊपर वाले क्रम में तथा (*ii*) किसी भी क्रम में

(*c*) 2 लाल तथा एक नीली गेंद है।

(*d*) एक भी सफेद गेंद नहीं है।

(30) किसी स्थान की 50% जनसंख्या हिन्दी का समाचार पत्र पढ़ती है व 20% अंग्रेजी का और 10% दोनों ही प्रकार के समाचार पत्र पढ़ती है। यदि एक पाठक निदर्शन से चुने तो उसके हिन्दी अथवा अंग्रेजी समाचार पत्र पढ़ने की प्रायिकता बताओ।

Fifty per cent population of an area read Hindi newspaper while 20% population reads English newspaper and 10% population read both newspapers. Find the probability that a resident selected at random reads either of the newspaper. [**Ans.** 3/5]

(31) एक लड़के के प्रश्न हल करने की सम्भावना $\frac{3}{4}$ है जबकि एक लड़की द्वारा हल किए जाने की सम्भावना $\frac{4}{5}$ हो तो निम्न प्रायिकताएं ज्ञात कीजिए :

(अ) दोनों ही प्रश्न हल कर लेंगे।

(ब) कोई भी प्रश्न हल नहीं करेगा।

(स) केवल लड़का प्रश्न हल करेगा।

(द) दोनों में कम-से-कम एक प्रश्न हल कर लेगा।

The probability that a boy will solve a problme is 3/4 and that a girl will solve the same problem is 4/5. Find the probability that :-

(*a*) both will solve the problem.

(*b*) none will solve the problem.

(*c*) only boy will solve the problem.

(*d*) at least one of them will solve the problem.

[**Ans.** (*a*) 3/5, (*b*) 1/20, (*c*) 3/20, (*d*) 19/20]

(32) एक बर्तन में 4 लाल व 5 सफेद गेंदें हैं। दूसरे बर्तन में 2 लाल व 3 सफदे गेंदें हैं। प्रथम बर्तन से एक गेंद निकाली और दूसरे बर्तन में डाल दी। फिर दूसरे बर्तन से एक गेंद निकाली। निकाली हुई गेंद के लाल रंग की होने की क्या सम्भावना है ?

An urn contains 4 red and 5 white balls. Another urn contains 2 red and 3 white balls. A ball is drawn from first urn and is placed in second urn. A ball is then drawn from second urn. What is the probability that the drawn ball is red in colour ? [**Ans.** 11/27]

(33) एक कारखाने में किसी प्रकार की वस्तु 3 मशीनों से बनती है। मशीन अ, ब, स का दैनिक उत्पादन क्रमशः 3000, 2500 व 4500 इकाई है। भूतकालीन अनुभव के आधार पर मशीन 'अ' 1% मशीन 'ब' 1.2% तथा मशीन 'स' 2% खराब उत्पादन करती है। एक वस्तु पूरे दिन के उत्पादन से ली गई और खराब पाई हो तो उसके (*i*) मशीन 'अ' से बनने की (*ii*) मशीन 'स' से बनने की प्रायिकता क्या है ?

A factory produced certain type of output by 3 machines.

The respective daily production figures are — machine A — 3,000 units, machine B — 2,500 units and machine C — 45,00 units. Past experience shows that `% of output produced by machine A is defective. The corresponding fraction of defectives for the other two machines are 1.2 and 2.0 percent respectively. An item is drawn from the day's production and is found to be defective. What is the probability that it comes from outputs of

(*i*) machine A (*ii*) machine C ?

[**Ans.** (*i*) 1/5, (*ii*) 3/5]

❑❑

अध्याय - 15

सैद्धान्तिक आवत्ति वितरण
(Theoretical Frequency Distribution)

आवत्ति -वितरणों को निम्न दो शीर्षकों में विभाजित किया जा सकता है :

(1) वास्तविक अवलोकनों पर आधारित आवत्ति -वितरण (Observed or Empirical Frequency Distribution)

(2) प्रत्याशित या सैद्धान्तिक आवत्ति -वितरण (Expected or Theoretical Frequency Distribution)

वास्तविक अवलोकनों पर आधारित आवत्ति -वितरणों में हम समंकों या सामग्री को वास्तविक अवलोकनों (Obervations) तथा प्रयोगों (Experiments) द्वारा प्राप्त करते हैं। उदाहरण के लिए, ऊँचाई-मापन से सम्बन्धित श्रेणी या विद्यार्थियों के प्राप्तांक या मजदूरों की मजदूरी, इत्यादि से सम्बन्धित श्रेणियाँ समंकों के मापन या अवलोकन द्वारा प्राप्त की जाती है। जैसे एक कक्षा के 60 विद्यार्थियों की ऊँचाई का माप करके तथा परीक्षा में उनके प्राप्तांक ज्ञात करके निम्न दो अवलोकित आवत्ति -वितरणों की रचना की जा सकती है :

(1)		**(2)**	
ऊँचाई (सें.मी. में) (Height in cm)	**विद्यार्थियों की संख्या (No. of students)**	**प्राप्तांक (Marks Obtained)**	**विद्यार्थियों की संख्या (No. of students)**
150-155	4	0-10	5
155-160	8	10-20	10
160-165	30	20-30	29
165-170	12	30-40	11
170-175	4	40-50	3
175-180	2	50-60	2
	60	**योग**	60

उपर्युक्त उदाहरण **वास्तविक अवलोकन** पर **आधारित आवत्ति -वितरण** का है।

''किसी तथ्य के माप को वर्गान्तरों (Class Interval) में विभाजित करके उसके आधार पर आवत्ति श्रेणी के रूप में लिखने पर जो वितरण प्राप्त होता है उसे **वास्तविक अवलोकनों पर आधारित आवत्ति -वितरण** कहते हैं।''

उपर्युक्त आवत्ति -वितरण का सांख्यिकीय विधियों जैसे समान्तर माध्य (Arithmetic Average), प्रमाप विचलन (Standard Deviation), विषमता (Skewness) आदि द्वारा विस्तत विवेचन किया जा सकता है तथा महत्वपूर्ण निष्कर्ष निकाले जा सकते हैं।

इसके विपरीत, यदि वितरण की आवत्तियाँ वास्तविक निरीक्षणों या प्रयोगों द्वारा प्राप्त न करके कुछ निश्चित मान्यताओं के आधार ''गणितीय विधियों'' द्वारा प्राप्त की जाती है तो उसे **सैद्धान्तिक वितरण या प्रायिकता वितरण** (Theoretical Distributions of Probability Distribution) कहते हैं। उदाहरण के लिए, यदि एक सिक्के को 100 बार उछाला जाए तो यह आशा की जा सकती है कि 50% परिस्थितियों में अर्थात् 50 बार चित्त (Head) आएगा तथा 50% परिस्थितियों में अर्थात् 50 बार पट (Tail) । व्यवहार में यह सम्भव है कि ठीक 50% परिस्थितियों में चित्त या पट न आए अर्थात् सिक्के को 1000

बार उछाले जाने पर 560 बार चित्त तथा 440 बार पट आ सकता है। सैद्धान्तिक रूप से जो **आवत्तियाँ** किसी परिस्थिति में आनी चाहिएँ और वास्तविक रूप से जो आवत्तियाँ प्राप्त होती हैं उनके अन्तर का सांख्यिकीय विधियों द्वारा अध्ययन करके महत्वपूर्ण निष्कर्ष निकाले जा सकते हैं। उपर्युक्त परिस्थिति में यह ज्ञात किया जा सकता है कि कहीं सिक्का उछालने वाला छल तो नहीं कर रहा है क्योंकि वास्तविकता और मान्यता में काफी अन्तर है।

सैद्धान्तिक आवत्ति वितरण की उपयोगिता

(Uses of Theoretical Distribution)

प्रायिकता या सम्भावना या सैद्धान्तिक आवत्ति वितरण सांख्यिकी के आधार स्तम्भ हैं। इन वितरणों की उपयोगिता निम्न तथ्यों से स्पष्ट होती है :-

(1) सम्भावित समंकों की सहायता से विवेकपूर्ण निर्णय (Wise Decision) लेने में सहायता मिलती है।

(2) ऐसी स्थिति में जहाँ वास्तविक अवलोकन धन के अभाव अथवा ऊँची लागत के कारण सम्भव नहीं होते, सम्भावना आवत्ति वितरण के आधार पर निर्णय लिए जा सकते हैं।

(3) वास्तविक आवत्ति वितरण व प्रत्याशित आवत्ति के तुलनात्मक अध्ययन से दोनों वितरणों के अन्तर का कारण जाना जा सकता है। अर्थात् दोनों वितरणों में अन्तर न्यादर्श के उच्चावचनों (Sampling Fluctuations) के कारण है अथवा किन्हीं अन्य कारणों से है।

(4) प्रायिकता आवत्ति वितरण की सहायता से व्यवसायी भविष्य में वस्तु की अनुमानित माँग जान सकते हैं।

(5) वास्तविक अनुसंधान करने से पूर्व सैद्धान्तिक आवत्ति की सहायता से वास्तविक वितरण की प्रकृति का अनुमान लगाया जा सकता है।

(6) गुण नियन्त्रक (Quality Controller) इन वितरणों की सहायता से यह जान सकते हैं कि एक विशेष उत्पादन विधि ठीक प्रकार से चल रही है अन्यथा नहीं।

(7) इन वितरणों की सहायता से जोखिम (Risk) व अनिश्चितता (Uncertainty) के बारे में जानकारी प्राप्त की जा सकती है।

सैद्धान्तिक आवत्ति वितरण के प्रकार

(Types of Theoretical Frequency Distribution)

वास्तव में सैद्धान्तिक आवत्ति वितरण कई प्रकार के होते हैं , किन्तु यहाँ हम केवल प्रमुख तीन प्रकार के वितरणों का ही अध्ययन करेंगे जो अधिक उपयोगी व व्यवहारिक है। यही तीन विवरण इस प्रकार हैं :-

(A) द्विपद वितरण (Binomial Distribution)

(B) पॉयसन वितरण (Poisson Distribution)

(C) सामान्य वितरण (Normal Distribution)

(A) द्विपद वितरण

(Bionomial Distribution)

द्विपद वितरण के साथ जेम्स बर्नोली (James Bernoulli) (1654-1705) का नाम जुड़ा हुआ है तथा इसे बर्नोली वितरण (Bernoulli Distribution) भी कहते हैं। इस सैद्धान्तिक वितरण का प्रतिपादन बर्नोली ने किया था, और इसका प्रकाशन उनकी मत्यु के 8 वर्ष पश्चात सन् 1713 में हुआ था। द्विपद (Binomial) का अर्थ 'दो नाम' (Two Names) है, अतः आवत्ति वितरण दो वर्गों के अन्तर्गत आता है। द्विपदीय समग्र (Binomial Population) के सभी अवलोकनों (Observations) को एक विशेषता रखने वाले तथा दूसरे विशेषता न रखने वाले दो वर्गों में विभक्त किया जाता है। यह द्विपद-वितरण एक सम्भावित वितरण होता है जो द्वन्द-विकल्पनाओं जैसे सफलता तथा असफलता के एक युग्म की सम्भावनाओं को व्यक्त करता है।

यदि A तथा B दो सिक्के उछाले जाएँ तो सम्भावित परिणाम निम्न प्रकार होंगे :-

A	B
H	H
H	T
T	H
T	T

(H stands for head and T stands for tail)

दोनों सिक्कों के उछलने पर उनके चित्त गिरने (H) की सम्भावना $\frac{1}{4}$ है तथा पट गिरने (T) की सम्भावना भी $\frac{1}{4}$ है। सम्भावना प्रमेय से भी यही परिणाम ज्ञात होता है :

$$\text{प्रथम उछाल में चित्त (H) गिरने की सम्भावना} = \frac{1}{2}$$

$$\text{द्वितीय उछाल में चित्त (H) गिरने की सम्भावना} = \frac{1}{2}$$

दोनों सिक्कों की एक उछाल में दोनों के चित्त (H) गिरने की सम्भावना $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ है। पट गिरने (T) की सम्भावना भी $\frac{1}{4}$ है।

यदि चित्त गिरने को घटना की सफलता माना जाये तो उसकी सम्भावना का संकेत 'p' रखा जाये तथा पट गिरने को घटना की असफलता माना जाये और उसकी सम्भावना का संकेत 'q' रखा जाये तो उपर्युक्त परिणामों को निम्न प्रकार व्यक्त किया जा सकता है :

$$\begin{array}{cccc} HH & HT & TH & TT \\ \underbrace{pp}_{p^2} & \multicolumn{2}{c}{\underbrace{pq \quad qp}_{2\,pq}} & \underbrace{qq}_{q^2} \end{array}$$

$p^2 + 2\,pq + q^2$, $(p+q)^2$ का विस्तार है। अतः दो स्वतंत्र घटनाओं की सामूहिक सम्भावना का सरल द्विपक्षी सूत्र $(p+q)^2$ है।

क्योंकि $\quad p = \frac{1}{2}$ और $pq\ \frac{1}{2}$

अतः $\quad (p+q)^2 = p^2 + 2pq + q^2 = \frac{1}{4} + 2 + \frac{1}{2} + \frac{1}{2} + \frac{1}{4} = T$

इसी प्रकार तीन सिक्के A, B, C उछालने पर 8 सम्भावित परिणाम हो सकते हैं :-

A	**B**	**C**		
H	H	H	=	p^3
H	H	T	=	p^2q
H	T	H	=	p^2q
H	T	T	=	p^2q
T	H	H	=	pq^2
T	H	T	=	pq^2
T	T	H	=	pq^2
T	T	T	=	q^3

यदि $p = \frac{1}{2}, q = \frac{1}{2}$:

$$(p + q)^3 = \left(\frac{1}{2} \quad \frac{1}{2}\right)^3 = p^3 + 3p^2q + 3pq^2 + q^3 = \frac{1}{8} \quad \frac{3}{8} \quad \frac{3}{8} \quad \frac{1}{8}$$

यह विस्तार बीजगणित के निम्न सूत्रों पर आधारित हैं :-

$$(a + b)^2 = a^2 + b^2 + 2ab$$

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^2$$

इस विस्तार-सूत्र को $(p + q)^n$ के रूप में व्यक्त किया जा सकता है। यदि विभिन्न परिणामों की सम्भावित आवत्तियाँ ज्ञात करना चाहते हैं तो $N (p + q)^n$ सूत्र की सहायता से उन्हें ज्ञात किया जा सकता है। इस सूत्र में N योग (number of trials) व n स्वतंत्र घटनाओं की संख्या का प्रतिनिधित्व करता है। यदि आवत्तियों का कुल योग 100 हो तथा स्वतंत्र घटनायें दो हों तो संभावित आवत्ति वितरण इस प्रकार होगा : -

$N (p + q)^n$ or $100 (p + q)^2$

$100 (p^2 + 2pq + q^2)$ ir $p = \frac{1}{2}, q = \frac{1}{2}$

Then $100\left(\frac{1}{4} \quad \frac{1}{2} \quad \frac{1}{4}\right)$

$$= 100\left(\frac{1}{4}\right) + 100\left(\frac{1}{2}\right) \quad 100\left(\frac{1}{4}\right)$$

$$= 25 + 50 + 25 = 100$$

अन्य रूप में

25 — for two successes

50 — for one success and one failure

25 — for no successes.

इसी प्रकार से अगर हम 4 व 5 सिक्कों को उछालें तो हमें निम्नलिखित विवरण मिलेंगे :-

$$(p + q)^4 = p^4 \quad 4p^3q \quad 6p^2q^2 \quad 4pq^3 \quad q^4$$

$$(p + q)^4 = p^5 \quad 5p^4q \quad 10p^3q^2 \quad 10p^2q^3 \quad 5pq^4 \quad q^5$$

यदि n घटनाएं हों तो ये विस्तार निम्न प्रकार दिया जाएगा :

$$(p + q)^n = q^n \quad {}^nC_1q^{n \quad 1}p \quad {}^nC_2q^{n \quad 2}p^2 \quad ..... \quad {}^nC_rq^{n \quad r}p^r \quad ..... \quad p^n$$

इस विस्तार का हर पद हमें सफलता के एक निश्चित मान की सम्भावना बताता है। उदाहरण के लिए ${}^nC_2q^{n-2}p^2$ का जब हम मूल्य निकालेंगे तो ये 2 सफलताओं की प्रायिकता है। सामान्य रूप में

$$\text{Probability of } r \text{ success} = {}^nC_2.q^{n-r}p$$

यहाँ पर p सफलता की प्रायिकता व q असफलता की प्रायिकता को दिखाती है। निम्न उदाहरण से ये स्पष्ट हो जाएगा।

Illustration 1 : A coin is tossed 8 times. Find the probability of getting 4 heads.

Solution : As we know, getting head is success and since probability of getting head in one toss is $\frac{1}{2}$.

So, $$p = \frac{1}{2}, q = 1 - p = \frac{1}{2}$$
$$n = 8, \quad r = 4$$

So applying the formula

$$\text{P } (r \text{ successes}) = {}^{n}\text{C}_{r} . q^{n-r}\, p^{r}$$

\ $$\text{P (4 successes)} = 8_{\text{C}_4} \left(\frac{1}{2}\right)^{8\ \ 4} \left(\frac{1}{2}\right)^{4}$$

$$= \frac{8!}{4!4!} \left(\frac{1}{2}\right)^{4} \left(\frac{1}{2}\right)^{4}$$

$$= \frac{8 \quad 7 \quad 6 \quad 5 \quad 4}{4 \quad 3 \quad 2 \quad 4} \left(\frac{1}{2}\right)^{8}$$

$$= 70 \times \frac{1}{256} \quad \frac{35}{128}$$

बरनौली वितरण की मान्यताएँ
(Assumptions of Bernoulli Distribution)

इस वितरण को भली प्रकार से समझने के लिए इस प्रमेय की मान्यताओं को जानना आवश्यक है। इसकी मान्यताएँ निम्नलिखित हैं :-

(1) प्रत्येक प्रयोग में केवल दो पारस्परिक अपवर्जी घटनाएँ (Mutually Exclusive Events) हो सकती हैं जिनके परिणाम सफलता (Success) व असफलता (Failure) के रूप में व्यक्त करते हैं।

(2) सफलता की अभिव्यक्ति 'p' से व असफलता की अभिव्यक्ति 'q' से करते हैं व ($p + q = 1$) सफलता व असफलता का योग सदैव एक है।

(3) इस प्रमेय में न्यादर्श की संख्या (Size of Sample) निश्चित होती है व इसे (n) के रूप में व्यक्त करते हैं।

(4) घटना स्वतंत्र रूप से घटती है अर्थात् पहले प्रयोग (trail) का प्रभाव दूसरे प्रयोग पर नहीं पड़ता। सफलता व असफलता का परिणाम प्रत्येक प्रयोग पर समान ही रहता है।

जैसे — सिक्के को बार-बार उछालने पर पट या चित्त की सम्भावना क्रमशः $\frac{1}{2}$ व $\frac{1}{2}$ रहेगी। इसी प्रकार ताश की गड्डी से प्रथम पत्ता यदि वापिस रख दिया जाये तो दूसरे पत्ते के निकालने पर प्रथम पत्ते का दूसरी घटना पर कोई प्रभाव नहीं पड़ेगा।

(5) द्विपद वितरण के प्रयोग से एक निश्चित सफलता (r) बार होने की सम्भावना को भी जाना जा सकता है। उदाहरणार्थ, 5 सिक्कों को उछालने पर 2 (head) आने की सम्भावना को भी माना जा सकता है।

द्विपद वितरण के अचल (Constants of Binomial Distribution)

द्विपद विस्तार के विभिन्न (Constants) इस प्रकार हैं :-

(1) समान्तर माध्य अथवा $\overline{\text{X}} = np$

(2) प्रमाप विचलन अर्थात् $s = \sqrt{npq}$ अथवा मापांक (Variance) या द्वितीय परिघात (Second Moment – m_2) = npq

(3) ततीय परिघात (Third Moment – m_3) = $npq\ (q - p)$

(4) चतुर्थ परिघात (Third Moment – m_4) = $3n^2p^2q^2 + npq\ (1 - 6pq)$

(5) विषमता का माप (Measure of Skewness – b_1)

$$= \frac{\frac{2}{3}}{\frac{3}{2}} = \frac{(q \quad p)^2}{npq}$$

(6) पथुशीर्षत्व का माप (Measure of Kurtosis – b_2)

$$= \qquad = 3 + \frac{1 \quad 6pq}{npq}$$

उपर्युक्म द्विपद-प्रमाप से निम्न सामान्य संबंधों (General Relationship) को ध्यान में रखना चाहिएः

(1) **पदों की संख्या (Number of Terms) :** द्विपद विस्तार में $(n + 1)$ पद होते हैं अर्थात् घात (power) से एक अधिक। जैसे $(p + q)^4$ में 5, $(p + q)^7$ में 8 पद होंगे।

(2) **घातों का योग (Sum of Powers) :** प्रत्येक पद में p व q घातों का योग n होगा। $(p + q)^5$ के विस्तार पद में p का घात 5 तथा p का 0 है, दोनों घातों का $5 + 0 = 5$, दूसरे पद में $4 + 1$, तीसरे पद में $3 + 2$, चौथे पद में $2 + 3$, पाँचवें पद में $1 + 4$ तथा छठे पद में $0 + 5$ है। इस प्रकार प्रत्येक पद में घातों का योग 5 अर्थात n ही है।

(3) **घातों का क्रम (Order of Power) :** $(p + q)$ के विस्तार में घात क्रमशः $n, (n - 1), (n - 2)$ ..... 2, 1, 0 होते हैं अर्थात् n से लेकर एक-एक बढ़ते जाते हैं, अंतिम पद में p^0 रह जाता है। इसके विपरीत q के घात 0 से एक-एक बढ़ते जाते हैं तथा अन्तिम पद में n घात होता है अर्थात् 0, 1, 2, ... $(n - 1), n$ 1 $(p + q)^5$ में p के घात 5, 4, 3, 2, 1, 0 तथा q के घात क्रमशः 0, 1, 2, 3, 4, 5 हैं।

(4) **संख्यात्मक गुणांक (Numerical Coefficients) :** विभिन्न$\frac{4}{3}$ पदों के गुणाकों के सम्बन्ध में निम्न बातें उल्लेखनीय हैं :-

(*i*) सम्पूर्ण द्विपद विस्तार के विभिन्न गुणांक हमेशा संमित (Symmetrical) होते हैं, जैसा पासकल त्रिभुज (Pascal's Triangle) से स्पष्ट है। $n = 4$ के लिए 1, 5, 10, 10, 5, 1 व $n = 6$ के लिए 1, 6, 15, 20, 15, 6, 1 तथा इसी प्रकार।

(*ii*) **गुणांक,** सूत्र के अनुसार, संचय (Combination) द्वारा या **पासकल त्रिभुज** की सहायता से प्राप्त कर सकते हैं।

(*iii*) किसी द्विपद वितरण के सभी पदों (Terms) के गुणांकों का योग 2" होता है जिससे उस स्थिति की कुल सम्भावनाओं की संख्याओं का पता चलता है। प्रत्येक क्रमिक घात के लिये यह योग **दो गुना** होता चला जाता है। उदाहरण के लिए, यदि $n = 2$ तो गुणांक (1, 2, 1) का योग $2^2 = 4$ होगा, यदि $n = 3$ हो तो गुणांकों (1, 3, 3, 1) का योग $2^3 = 8$ होगा जो पिछले योग का दो गुना है। इसी प्रकार अन्य गुणांकों को ज्ञात किया जा सकता है।

(5) ***p* तथा *q* क्रम (Order of *p* and *q*) :** द्विपद वितरण का सामान्य रूप $(p + q)^n$ का विस्तार है। परन्तु इस स्वरूप के प्रयोग में यह कठिनाई होती है कि सफलताओं की संख्या **अवरोही क्रम** (Descending Order) में लिखी जाती है अर्थात् पहले सबसे अधिक फिर कम तथा अन्त में 0, सफलताओं की संख्या को **आरोही क्रम** (Ascending Order) में रखकर प्रायिकता ज्ञात करने के लिए $(q + p)^n$ का विस्तार लिखा जा सकता है। $(q + p)^n$ का द्विपद विस्तार उसी प्रकार लिखा जाता है जिस प्रकार $(p + q)n$ का। अन्तर केवल इतना है कि p के स्थान पर q तथा q के स्थान पर p लिख दिया जाता है।

$(q + p)^n$ का विस्तार करने पर $x = 0, 1, 2, 3, ... n$ के विभिन्न मानों को प्रायिकता प्राप्त होती है।

$$(q + p)n = q_n + {}^nC_1q^{n-1}\,p + {}^nC_2q^{n-2}\,p^2 + ... + {}^nC_rq^{n-r}\,p^r + ...\,p^n$$

${}^nC_0, {}^nC_1, {}^nC_2$ आदि को द्विपद गुणांक कहते हैं।

सफलताओं की संभावित आवत्ति निकालना

(Computation of Expectation of Various Successes)

इसी प्रकार यदि हम उपरोक्त द्विपद वितरण का प्रयोग स्वतंत्र रूप से (N) बार करें तो हम (N) परिणामों की सम्भावित आवत्ति भी जान सकते हैं। ऐसी स्थिति में सूत्र को निम्न प्रकार व्यक्त करेंगे :-

$$N\,(q+p)^n\,p \;=\; \text{Probability of Success}$$

where, q = Probability of failure

n = Size of Sample

N = Number of Trials

निम्न उदाहरणों को समझने पर हम द्विपद वितरण को भली प्रकार समझ सकेंगे :-

Illustration 1 : Find the probability of all the values of 'X' for the binomial distribution with parameters $n = 6$ and $p = 1/3$.

द्विपद वितरण में चर (X) के सभी मूल्य ज्ञात करें यदि $x = 6$ व $p = 1/3$ हों।

हल (Solution) : Success $= p = \dfrac{1}{3}$ (given)

\ Failure $= q = 1 - p = 1 - \dfrac{1}{3} = \dfrac{2}{3}$ $[\because q + q = 1]$

Size of sample $= n = 6$

Various probabilities can be known by using the bionomial expansion of $(q + p)^6$

$$(q+p)^6 = q^6 + 6q^5p^1 + 15q^4p^2 + 20q^3p^3 + 15q^2p^4 + 6q^1p^5 + p^6$$

By substituting the values of p and q we have : $\left[\dfrac{64}{729}\;\; \dfrac{192}{729}\;\; \dfrac{240}{729}\;\; \dfrac{160}{729}\;\; \dfrac{60}{729}\;\; \dfrac{12}{729}\;\; \dfrac{1}{729}\right]$

$$= \left(\frac{2}{3}\right)^6 \quad 6\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^1 \quad 15\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^2 \quad 20\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 \quad 15\left(\frac{2}{3}\right)^2\left(\frac{1}{4}\right)^4 \quad 6\left(\frac{2}{3}\right)^1\left(\frac{1}{3}\right) \quad \left(\frac{1}{3}\right)^6$$

=

This information can be written as :-

No. of Success (p) X	Probability
0	64/729
1	192/729
2	240/729
3	160/729
4	60/729
5	12/729
6	1/729

Illustration 2 : Six dice were thrown 128 times. Each 4, 5, or 6 spot appearing was considered to be success and each 1, 2 or 3 spot a failure. The results were :

No. of successes	0	1	2	3	4	5	6
	0	10	26	44	34	14	0

Solution :-

Expected Frequency $= N(p \quad q)^6 = 128\left(p^6 \quad 6p^5q \quad 15p^4q^2 \quad 20p^3q^3 \quad 15p^2q^4 \quad 5pq^5 \quad q^6\right)$

=

= 2, 12, 30, 40, 30, 12, 2

No. of Successes	Actual Frequencies	Expected Frequencies
0	0	2
1	10	12
2	26	30
3	44	40
4	34	30
5	14	12
6	0	2
	128	**128**

उपरोक्त उदाहरण में सफलता की प्रायिकता (Probability of Success) and so $q = 1 - p = \frac{1}{2}$

Illustration 3 : A bowl contains 4 red balls and 6 green balls. Five balls are to be drawn. Find probabilities for different events.

Solution :

The probability of drawing a red ball $= \frac{4}{10} = .4\ (p)$

The probability of drawing a green ball $= \frac{6}{10} = .6\ (q)$

Number of balls to be drawn		Probability		(*p*)
Red	Green			
5	0	$p \quad \frac{5!}{5!(5 \quad 5!)}(.4)^5(.6)^0$	=	.01024
4	1		=	0.7680
3	2		=	.23040
2	3		=	.34560
1	4		=	.25920
0	5		=	.07776
				1.00000

Illustration 4 : A coin is tossed four times. What is the probability to getting two or more heads ?

Solution :

When a coin is tossed

$$\text{the probability of head } q = 1/2$$

$$\text{and the probability of tail } q = (1 - p) = 1 - 1/2 = 1/2$$

The various possibilities for all the events are the terms of the binomial expansion $(q + p)^4$

$$(q + p)4 = q^4 + 4q^3 p + 6q^2 p^2 + 4q p^3 + p^4$$

\ The probability of drawing 2 heads is

$$6q^2p^2 = 6 \times$$

$$= 6 \times \frac{1}{4} \times \frac{1}{4}$$

$$= \frac{3}{8}$$

The probability of getting 3 heads is

$$4qp^3 = 4 \times \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)^3$$

$$\left(\frac{1}{2}\right) \left(\frac{1}{2}\right)^2$$

$$= 6 \times \frac{1}{2} \; \frac{1}{8}$$

$$= \frac{1}{4}$$

The probability of getting 4 heads is

$$p^4 = \left(\frac{1}{2}\right)^4$$

$$= \frac{1}{16}$$

\ The probability of getting 2 or more heads is

$$\frac{3}{8} \quad \frac{1}{4} \quad \frac{1}{16} = \frac{11}{16}$$

Illustration 5 : Assuming that half the population are consumers of wheat in the Northern India, so that the chance of an individual being a consumer of wheat is 1/2 and assuming that 100 investigators each take 10 individual to see whether they are consumers of wheat, how many investigators would you expect to report that three or less were consumers of wheat ?

Solution : We are given :

Probability of an individual being a consumer of wheat $p = 1/2$

$$q = 1 - p$$
$$= 1 - 1/2$$
$$= 1/2$$

N = 100, n = 10

Using the bionomial distribution, we have

$$\text{P}(r) = {}^{n}\text{C}_{r} q^{n-r} p^{r}$$

Substituting various values we get

$$\text{P}(r) = {}^{10}\text{C}_{r} \left(\frac{1}{2}\right)^{10\ \ r} \left(\frac{1}{2}\right)^{r}$$

$$= {}^{10}\text{C}_{r} \qquad = \frac{1}{1024}\ {}^{10}\text{C}_{r}$$

The probability that in a sample of 10, three or less people are consumers of wheat is given by

$$p_0 + p_1 + p_2 + p_3$$

$$= \frac{1}{1024}\left[{}^{10}\text{C}_0 + {}^{10}\text{C}_1 + {}^{10}\text{C}_2 + {}^{10}\text{C}_3\right]$$

$$= \frac{1}{1024}[1 + 10 + 45 + 120] = \frac{176}{1024} \quad \frac{11}{64}$$

Thus out of 100 investigators, the number of investigators who will report 3 or less consumers of wheat in a sample of 10 is $100 \times \frac{11}{64} = 17.2$ $\left(\frac{1}{2}\right)^{10}$

Illustration 6 : The mean of a binomial distribution is 20 and standard deviation is 4. Calculate *n*, *p*, *q*.

Solution : Mean = 20, *i.e.* $np = 20$

s = 4 *i.e.* $\sqrt{npq} = 4$

Now $npq = 16$

i.e., $20q = 16$

\ $q = \frac{16}{20} \quad \frac{4}{5}$

$p = 1 - q = 1 - \frac{4}{5} \quad \frac{1}{5}$

Now $np = 20$

or $n \times p = 20$

or $n \times \frac{1}{5} = 20$

\ $n = 5 \times 20 = 100$

Hence $n = 100, p = \frac{1}{5}$ and $q = \frac{4}{5}$

Illustration 7 : किसी उद्योग में 25% कर्मचारियों में कार्य के कारण बीमार होने का प्रभाव हो तो 6 कर्मचारियों के न्यादर्श में से 4 या अधिक के बीमार होने की प्रायिकता क्या है ?

The incidence of occupational disease in an industry is such that the workmen have a 25% chance of suffering from it. What is the probability that out of six workmen 4 or more will contact the disease ?

हल (Solution) :

Let q denotes the chance of suffering and p denotes the chance of not suffering.

$$\backslash \qquad q = 25\% = \frac{1}{4}, p = 1 - \frac{1}{4} = \frac{3}{4}$$

The binomial expansion is

$$(q + q)^n = \left(\frac{3}{4} \quad \frac{1}{4}\right)^6$$

$$= p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6p^1q^5 + q^6$$

The probability of 4 or more (*i.e.,* 4, 5, and 6 successes is) who will contact the disesase

$$= 15p2q4 + 6pq5 + q6$$

$$= 15\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^4 \quad 6\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^5 \quad \left(\frac{1}{4}\right)^6$$

$$= \frac{135}{4096} \quad \frac{18}{4096} \quad \frac{1}{4096}$$

$$= \frac{135 \quad 18 \quad 1}{4096} = \qquad = \mathbf{0.376.}$$

$\frac{154}{4096}$ $\sqrt{81}$

Illustration 8 : यदि एक बोल्ट के खराब होने की प्रायिकता 0.1 है तो ज्ञात कीजिए (*a*) समान्तर माध्य (*b*) कुल 900 बोल्टों में से खराब बोल्टों का प्रमाप विचलन। विषमता और पथुशीर्षत्व भी ज्ञात कीजिए।

If the probability of a defective bolt is 0.1, find (*a*) the mean and (*b*) the standard deviation of defective bolts in a total of 900. Also calculate skewness and kurtosis.

हल (Solution) : The probability of a defective bolt is 0.1.

$$\backslash \qquad p = 0.1 \text{ and } q = (1 - 0.1) = 0.9$$

(*a*) Mean = np = 900 (0.1) = **90,** *i.e.*, we can expect 90 bolts to be defective.

(*b*) $$\text{Standard Deviation} = \sqrt{npq} = \sqrt{900\,(0.1)\,(0.9)}$$

$$= \qquad = 9$$

Skewness $$(\text{b}_1) = \frac{(q \quad p)^2}{npq} \quad \frac{(0.9 \quad 0.1)^2}{81} \quad \frac{0.64}{81} = 0.0079$$

Kurtosis $$(\text{b}_1) = 3 + \frac{1 \quad 6pq}{npq}$$

$$= 3 + \frac{1 \quad 6\,(0.1)\,(0.9)}{81} = 3 + 0.006 = \mathbf{3.006}$$

By putting the value of p and q, we can find n

$$n \times 0.6 \times 0.4 = 16,\ 0.24\,n = 16$$

$$n = \qquad = 66.67 \text{ or } 67.$$

द्विपद वितरण का अन्वायोजन (Fitting a Binomial Distribution) : जब अवलोकित समंकों (Observed Data) के लिए द्विपद वितरण का अन्वायोजन (Fitting) किया जाता है तो निम्न विधि अपनाई जाती है :

(1) p तथा q का मूल्य ज्ञात कीजिए। यदि इन दोनों मूल्यों (मानों-values) में से एक ही दिया हुआ हो तो दूसरे को एक सरल सम्बन्ध द्वारा ज्ञात किया जा सकता है, $p = (1 - q)$ तथा $q = (1 - p)$ जब p व q दोनों समान होते हैं तो वितरण **संमित** (Symmetrical) होता है। यदि p व q असमान होते हैं तो वितरण विषम (Skewed) होता है। यदि p का मूल्य $\frac{1}{2}$ से कम है तो वितरण धनात्मक विषम (Positively Skewed) है और जब p का मूल्य $\frac{1}{2}$ से अधिक है तो वितरण ऋणात्मक विषम (Negatively Skewed) होता है।

(2) द्विपद $(p + q)^n$ का विस्तार कीजिए। द्विपद विस्तार में घात n (power n) पदों की संख्या से एक कम के बराबर होती है। जब दो सिक्के उछाले जाते हैं तब द्विपद (Binomial) में तीन पद (Terms) होंगे। इसी प्रकार, जब चार सिक्के ($n = 4$) उछाले जाते हैं तब पाँच पद होंगे और आगे इसी प्रकार।

(3) प्रत्तयाशित आवत्तियाँ (Expected Frequencies) ज्ञात करने के लिए द्विपद विस्तार के प्रत्येक पद को N (the Total Frequency) से गुणा कीजिए।

उपर्युक्त उदाहरण से यह विधि (Procedure) पूर्णतया स्पष्ट हो जाएगी।

Illustration 9 : Five coins are tossed 3,200 times, find the frequency of the distribution of heads and tails and tabulate the results. Calculate mean number of successes and standard deviation.

पाँच सिक्कों को 3200 बार उछालने पर सिक्कों के विभिन्न पट व चित्त आने का अनुमानित आवत्ति वितरण बनाओ। इस वितरण का माध्य व प्रमाप विचलन भी ज्ञात करें।

Solution : Let Probability of getting head $= p = 1/2$

Prob. (tail) $= q = 1 - p = 1/2$

Size of sample of $= n = 5$ (given)

Number of trials $= N = 3{,}200$ (given)

$$\left[\frac{1}{32} \quad \frac{5}{32} \quad \frac{10}{32} \quad \frac{10}{32} \quad \frac{5}{32} \quad \frac{1}{32}\right]$$

Various possibilities of getting, 0, 1, 2, 3, 4 and 5 heads can be obtained by applying formula of bionomial distribution N $(q + p)^n$

\ $3200\,[q + p]^5 = 3200\,[q^5 + 5q^4p^1 + 10q^3p^2 + 10q^2p^3 + 5q^1p^4 + q^5]$

$$3200\left[\frac{1}{2} \quad \frac{1}{2}\right]^5 = 3200\left[\left(\frac{1}{2}\right)^5 \quad 5\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^1 \quad 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 \quad 10\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 \quad 5\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 \quad \left(\frac{1}{2}\right)^5\right]$$

$\because$ $p = q$, so product of pq with any power will give $(1/2)^5$ in every case

$= 3200$

$= 100 + 500 + 1{,}000 + 1{,}000 + 500 + 100$

This information can be tabulated as under :-

Number of Possibilities	
No. of Heads	**Frequencies**
0 heads	100
1	500
2	1,000
3	1,000
4	500
5	100

(*ii*) $\overline{X}$ of success $= np = 5 \times \frac{1}{2} = 2.5$

S.D (s) of success $= \sqrt{n.p.q.}$

$= \sqrt{5 \quad \frac{1}{2} \quad \frac{1}{2}}$

$= 1.118$

Illustration 10 : The following data shows the number of seeds germinating out of 10 on damp filter for 80 sets of seeds. Fit a binomial distribution to this data :-

x	**:**	0	1	2	3	4	5	6	7	8	9	10
f	**:**	6	20	28	12	8	6	0	0	0	0	0

हल (Solution) :

Fitting Binomial Distribution

x	*f*	*fx*
0	6	0
1	20	20
2	28	56
3	12	36
4	8	32
5	6	30
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
	***n* = 80**	**S*fx* = 174**

$\overline{X} = \frac{fx}{N} = \frac{174}{80} = 2.175$ But mean $= np = \frac{174}{80} = 10\,p = \frac{174}{80}$ $(n = 10)$

$p = \frac{174}{800} = \mathbf{0.2175} \quad \backslash \quad q = 1 - 0.2175 = \mathbf{0.7825}$

Hence the bionomial distribution to be **fitted** to the data is

$$80\,(0.7825 + 0.2175)^{10}$$

The theoretical frequencies are the succcessive terms in the expansion 80 $(0.7825 + 0.2175)^{10}$ and a tabulated below :-

x	**Theoretical Frequencies** $\mathbf{N \times {}^nC_r\, q^{n-1} p^r}$		***fe***
0	$80 \times (0.7825)^{10}$	=	6.9
1	$80 \times 10\,(0.7825)^{9}\,(0.2175)^{1}$	=	19.1
2	$80 \times 45\,(0.7825)^{8}\,(0.2175)^{2}$	=	24.0
3	$80 \times 120\,(0.7825)^{7}\,(0.2175)^{3}$	=	17.8

4	$80 \times 210\,(0.7825)^6\,(0.2175)^4$	= 8.6
5	$80 \times 252\,(0.7825)^5\,(0.2175)^5$	= 2.8
6	$80 \times 210\,(0.7825)^4\,(0.2175)^6$	= 0.7
7	$80 \times 120\,(0.7825)^3\,(0.2175)^7$	= 0.1
8	$80 \times 45\,(0.7825)^2\,(0.2175)^8$	= 0.0
9	$80 \times 10\,(0.7825)^1\,(0.2175)^9$	= 0.0
10	$80 \times (0.2175)^{10}$	= 0.0
	Total	**= 80.0**

उदाहरण (Illustration) 11 : 6 पांसे 729 बार उछाले जाते हैं। कम-से-कम 3 पांसे 5 या 6 का अंक कितनी बार दिखाने की सम्भावना हो सकती है ?

Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six ?

हल (Solution) : The probability of getting 5 or 6 with one dice.

$$p = \frac{2}{6} \quad \frac{1}{3}; \quad p = (1-p) = 1 - \frac{1}{3} = \frac{2}{3}$$

Since the dice are in sets of 6 and there are 729 times, the binomial distribution is :

$$\mathrm{N}\,(p+q)n = 729\left(\frac{1}{3} \quad \frac{2}{3}\right)^6$$

\ The expected number of times at least *three* dice showing 5 or 6 would be

$$= 729\left[{}^6C_3\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^3 \quad {}^6C_4\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 \quad {}^6C_5\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1 \quad {}^6C_6\left(\frac{1}{3}\right)^6\right]$$

$$= 729\left[\frac{960}{6} \quad \frac{120}{2} \quad \frac{12}{1} \quad 1\right]$$

$$= 160 + 60 + 12 + 1 = \mathbf{233}$$

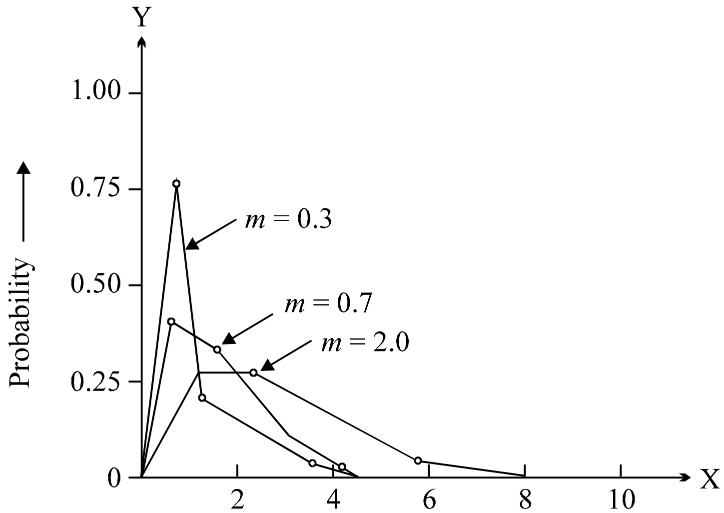
पॉयसन वितरण
(Poisson Distribution)

पॉयसन वितरण का प्रयोग सबसे पहले प्रसिद्ध फ्रांसीसी गणितज्ञ **साइमन जेनिस पॉयसन** (Simon Dennis Poisson —1781 - 1840) ने 1873 में किया। इसे उन्हीं के नाम पर **पॉयसन वितरण** कहा जाता है। पॉयसन वितरण एक खंडित प्रायिकता वितरण (Discrete Probability Distribution) है।

पॉयसन वितरण का उपयोग उन परिस्थितियों में किया जाता है जहाँ प्रायिकता (p) का मान (Value) बहुत की कम (लगभग शून्य के निकट) और n का मान (*value*) अत्यधिक हो। ऐसी परिस्थितियों में जहाँ p का मान बहुत ही कम तथा n का मान अत्यधिक हो तो **द्विपद वितरण** उपयुक्त सैद्धान्तिक आवत्तियाँ (Appropriate Theoretical Frequencies) प्रदान नहीं करता अर्थात् सही स्थिति का निरूपण नहीं कर सकता। ऐसी परिस्थितियों में पॉयसन वितरण ही अधिक उपयुक्त है। पॉयसन वितरण द्विपद वितरण की सीमान्त रूप (Limiting Form) है। जैसे कि p का मान शून्य की ओर (बहुत कम) तथा n का मान अनन्त (Infinity) की ओर (बहुत अधिक) प्रवृत्त हो तथा np अर्थात् माध्य (Mean) एक धनात्मक अचर (Constant) संख्या हो।

पॉयसन वितरण उन घटनाओं पर लागू होता है जो असामान्य तथा दुर्लभ घटनाएँ (Rare Events) होती हैं तथा उनके घटित होने की प्रायिकता (p) बहुत कम होती है। उदाहरण के लिए, एक दिन में सड़क दुर्घटना में मरने वाले व्यक्तियों की संख्या बहुत कम होती है, किसी बड़े कारखाने कें उत्पादन में खराब (Defective) होने वाले माल की सम्भावना बहुत कम होती है, किसी पष्ठ पर छपाई में होने वाली अशुद्धियों की सम्भावना कम होती है आदि।

द्विपद वितरण में समान्तर माध्य $(\overline{X})$ को न्यादर्श संख्या व सफलता की प्रायिकता के गुणा करने से प्राप्त किया जा सकता $(\overline{X} = np)$ है, जबकि पॉयसन वितरण में समान्तर माध्य को (m) से व्यक्त करते हैं जो पिछले अनुभव से प्राप्त होती है।



उपरोक्त वितरण में जैसे-जैसे माध्य (m) का मान बढ़ेगा, वैसे-वैसे पॉयसन वितरण वक्र दायीं आरे चलता रहेगा। जैसा कि उपरोक्त वक्र से स्पष्ट है। अतः पॉयसन वितरण का परिघात विषमता गुणांक (Coefficient of Skewness) सदैव जमा चिन्ह (+ive) ही रहेगा और पथुशीर्षत्व (Kurtosis) सदैव 3 से अधिक रहेगा। [*i.e.* $b_2 > 3$].

पॉयसन वितरण की मान्यताएँ व प्रयोग

(Assumptions and Applications of Poisson Distribution)

इस वितरण की मान्यताएँ संक्षेप में इस प्रकार हैं :-

(1) इस वितरण में न्यादर्श संख्या (Size of Sample) अनन्त होती है, अर्थात् -
['n' approaches infinity (n ® ¥)]

(2) घटना के घटने की प्रायिकता को 'p' से व्यक्त करते हैं व घटना के घटने की प्रायिकता शून्य के समीप होती है, अर्थात्
['p' approaches zero *i.e*, p ® 0]

(3) घटना के न घटने को 'q' से व्यक्त करते हैं, जिसकी प्रायिकता बहुत अधिक होती है और प्रायः एक के समीप ही होती है। अर्थात्

['q' approaches zero *i.e.*, q ® 1]

(4) घटना के घटने का माध्य $(m = np)$ अचर (Constant) होता है।

कुछ व्यवहारिक परिस्थितियाँ, जिनमें पॉयसन वितरण का उपयोग किया जा सकता है, इस प्रकार हैं :-

(1) गुण-नियन्त्रण सांख्यिकी (Quality Control Statistics) में किसी एक मद (Item) की खराबियों (Defects) की संख्या की गणना करने के लिए।

(2) जीव विज्ञान (Biology) में अणुओं (Bacteria) की गणना करने के लिए।

(3) बीमा के क्षेत्र में दुर्घटनाओं की गणना करने के लिए।

(4) छपे हुए पष्ठों पर अशुद्धियों की गिनती करने के लिए।

(5) किसी नगर में प्रति मिनट टेलीफोन कॉल (Telephone Calls) की संख्या ज्ञात करने के लिए।

(6) किसी बड़े कारखाने में निर्मित ब्लेडों में से दोषपूर्ण ब्लेडों की संख्या की गणना के लिए।

(7) एक फुटबाल मैच में एक टीम द्वारा किए गए गोलों की संख्या ज्ञात करने के लिए।

(8) किसी नगर के बड़ चौराहे पर होने वाली दुर्घटनाओं की संख्याओं की गणना करने के लिए।

(9) एक वर्ष में प्रतिष्ठित व्यक्तियों द्वारा की जाने वाली आत्महत्याओं की संख्या की गणना करने के लिए।

पॉयसन के वितरण को निम्न प्रकार से व्यक्त किया जा सकता है :

$$P(r) = e^{-m} \frac{m^r}{r!}$$

जहाँ e = 2.7183 प्राकृतिक लघुगणक का आधार (The base of Natural Logarithms)

m = पॉयसन वितरण का समान्तर माध्य अर्थात् np अथवा किसी घटना के घटने की माध्य संख्या (the mean of the Poisson Distribution *i.e. np* on the average number of occurence of an event)

r = वह संख्या, जिसके लिए सम्भाव्य आवत्ति ज्ञात करती है जैसे 0, 1, 2 .... (the number for which expected frequency is to be calculated.)

इस सूत्र को इस प्रकार की लिखा जा सकता है :-

$$p(r) = e^{-m}\left[1 \quad \frac{m}{1!} \quad \frac{m^2}{2!} \quad \frac{m^3}{3!} \quad .....\right]$$

पॉयसन वितरण की विशेषताएँ या अचर

(Properties or Constants of Poisson Distribution)

इस वितरण के मुख्य अचर इस प्रकार हैं :-

(1) समान्तर माध्य = $\overline{X} = m = np.$

(2) प्रमाप विचलन = $s = \sqrt{m}$

(3) प्रसारण = Variance = m

(4) परिघात (Moments) :0

(1) $m_1 = 0$

(2) $m_2 = m$

(3) $m_3 = m$

(4) $m_4 = m + 3m^2$

(5) $b_1 = \frac{{}^{2}_{3}}{{}^{3}_{2}} \quad \frac{m^2}{m^3} \quad \frac{1}{m}$

(6) $b_2 = \frac{4}{{}^{2}_{2}} \quad 3 \quad \frac{1}{m}$

(7) $g_1 = \sqrt{\ _1} \quad \frac{1}{\sqrt{m}}$

(8) $g_2 = \ _2 \quad 3 \quad \frac{1}{m}$

Illustration 12 : If the mean of a Poisson distribution is 4 find

(*a*) Standard deviation **(*b*) b_1** **(*c*) = b_2** **(*d*) m_3** **(*e*) m_4**

Solution :

$$m = 4$$

$$s = \sqrt{m} = \sqrt{4} = 2$$

$$b_1 = \frac{1}{m} \quad \frac{1}{4} \quad 0.25$$

$$b_2 = 3 \quad \frac{1}{m}$$

$$= 3 + \frac{1}{4}$$

$$= 3.25$$

$$m_3 = m = 4$$

$$m_4 = m + 3m^2 = 4 + 3 \times 4^2 = 52$$

Illustration 19 : It is given that 3% of fountain pens manufactured by a foreign based company are defective. Using Poisson distribution, find the probability that a sample of 100 pens will contain (*a*) no defective (*b*) exactly one defective.

Solution :

No. of defective pens in a sample of 100 = 3

\ $m = 3$

Probability of no defective pen in a sample of 100 is

$$P_{(0)} = e^{-m} = e^{-3} = 0.05$$

$$P_{(1)} = (P_0) \times m = 0.05 \times 3 = 0.15$$

Illustration 13 : An average number of phone calls per minute into the switchboard of Reddy Company Limited between the hours of 10 a.m. and 1 p.m. is 2.5 Find the probability that during one particular minute there will be (*i*) no phone all at all (*b*) exactly 3 calls and (*iii*) at least 5 calls.

$$\frac{c^{0.08208}(2.5)}{3!} \quad \frac{15625}{6}$$

Solution : Let denote the number of a telephone calls per minute by X. The X follows Poisson distribution with mean distribution, $m = 2.5$. The Poisson probability function is :

$$P\,(x = r) = \frac{c^{\;m} mr}{r!}$$

$$= \frac{c^{\;2.5}(2.5)^r}{4!}$$

(*i*) Probability of no call at all is :

$$p\,(x = 0) = c^{-25}\,(2.5) \times 0$$

$$= c^{-25} = c^{-20} \times c^{-0.5}$$

$$= 0.13534 \times 0.6065 = 0.08208$$

(*ii*) Probability of exactly 2 calls is :

$$p\,(x = 2) =$$

$$=$$

$$= 0.21375$$

(*iii*) Probability of at least 5 calls is :

$$p\,(x \geq 5) = 1 - p\,(x \leq 4)$$

$$= 1 = 0.08208$$

$$= 1 - 0.08208\,[1 + 2.5 + 3.125 + 2.6042 + 0.6144]$$
$$= 1 - 0.8208 \times 9.8436$$
$$= 1 - 0.80796 = 0.19204$$
$$= 0.192$$

Illustration 14 : The probability that a particular injection will have reaction to an individual is 0.002. Find the probability that out of 1000 individuals (*i*) exactly 2 individual (*ii*) at least 1 individual will have reaction from injection.

एक व्यक्ति की एक विशेष टीके द्वारा प्रभावित होने की प्रायिकता 0.002 हो तो 1000 व्यक्तियों में से (1) दो व्यक्तियों को (2) कम से कम 1 व्यक्ति की टीके से प्रभावित होने की प्रायिकता बताओ।

Solution :

Let Probability of reaction success $= p = 0.002$

or $p =$

$n = 1000$ (given)

$$\overline{X} = m = np = 1000 \times \frac{2}{1000} = 2$$

(*i*) **We know that P (*r*) success =** $\dfrac{e^{\ m}.m^r}{r!}$ $\left[\dfrac{2}{1000} 2.5 \quad \dfrac{(2.5)^2}{2!} \quad \dfrac{(2.5)^4}{4!} \quad \dfrac{(2.5)^4}{4!}\right]$

\ P (2) individual will have reaction $= \dfrac{e^{\ 2} \quad 2^2}{2!}$

$$= \frac{2^2 \quad 2.7183^{\ 2}}{2 \quad 1} \text{ (as } e = 2.7183\text{)}$$

$$= \frac{4 \quad 1}{2 \quad 2.7183^2} = \frac{2}{7.389} = .2706$$

(*ii*) Probablity of at least 1 individual will have reaction can be known by applying negative probability, *i.e.* 1 – P (non reaction).

Þ 1 – P (0) individual is reacted

P (0) reaction $= e^{-m} = 2.7183^{-2}$

$$\frac{1}{2.7183^2} = \frac{1}{7.389} = 0.1353$$

\ P (at least 1 reaction) $= 1 - 0.1353 = .6847$

Illustration 15 : The probability that a worker working in a mine will be met with an accident is 1/2000 during a year. Find the probability that a miner employing 3000 workers, there will be (1) no accident (2) at least 1 accident (3) at the most 1 accident in a year.

एक खनिक के 1 वर्ष में दुर्घटना के शिकार होने की प्रायिकता 1/2000 है। तो एक खान में 3000 खनिकों में (1) कोई दुर्घटना न होने (2) कम से कम। दुर्घटना होने (3) अधिकतम एक दुर्घटना वर्ष भर में होने की प्रायिकता बताओ।

Solution : Let probability of meeting with an accident is success —

$$p = \frac{1}{2000}$$

$$\text{Size of sample} = n = 3000$$

\ $$\overline{X} = m = np = 3000 \times \frac{1}{2000} = 1.5$$

We know that —

$$\textbf{P (}r\textbf{) success} = \frac{e^{\ m}.m^r}{r!}$$

(*i*) **Probability of no accident is –**

$$\text{P (0) accident} = e^{-m} - e^{-1.5} = e^{-1} \times e^{-.5}$$
$$= .3677 \times .6065 = .2230$$

(**Note :** $e^{-1.5}$ can be written as $e^{-1} \times e^{-.5}$)

This have been done because value of e^{-m} is given in the table for a whole number and for a decimal number.

(*ii*) **Probability of at least 1 accident —**

$$\text{This is same as} = 1 - \text{P (0) accident}$$
$$= 1 - .2230 = 0.7770$$

(**Note :** For P (0) [See solution of part (*i*)])

(*iii*) **Probability of at the most 1 accident —**

This is same as probability of either (0 or 1) accidents.

$$\textbf{P (}r\textbf{)} = \frac{e^{\ m}.m^r}{r!}$$

$$\text{P (1)} = \frac{e^{\ 1.5}.15^1}{1!}$$

$$= \frac{.2330 \quad 1.5}{1!} = .3345$$

Hence Probability of at the most 1 accident.

$$= \text{P (0)} + \text{P (1) accidents}$$
$$= .2230 + .3345 = .5575$$

पॉयसन वितरण का अन्वायोजन (Fitting a Poisson Distribution) : अवलोकित समंकों (observed data) के लिए पॉयसन वितरण का अन्वायोजन (Fitting) निम्न प्रकार किया जा सकता है।

(1) सबसे पहले हम अवलोकित समंकों द्वारा समान्तर माध्य (Arithmetic Mean,) की गणना करते हैं और इसी माध्य को पॉयसन वितरण के अचल (Constant) के रूप में प्रयोग करते हैं अर्थात् $\overline{X} = m$ ।

(2) इसके बाद प्रत्येक r सफलताओं के लिए सम्भावाओं की गणना पॉयसन वितरण $\text{P}(r) = \frac{e^{\ m}.m^r}{r!}$ का प्रयोग करके ज्ञात करते हैं।

(3) उपर्युक्त सम्भावनाओं को कुल आवत्तियों (Total Frequencies) N से गुणा करने पर प्रत्याशित या सैद्धान्तिक आवत्तियाँ (Expected or Theoretical Frequencies) प्राप्त होती हैं, जो तालिका से स्पष्ट है :-

सफलताओं की संख्या (No. of Successes) (r)	प्रत्याशित आवत्तियाँ (Expected Frequencies)
0	$N \times p\ (0) = N \times e^{-m}$
1	$N \times p\ (1) = N \times e^{-m} \times m = N\ (p_0) \times m$
2	$N \times p\ (2) = \frac{Ne^{\ \ m}\ \ m^2}{2!} = N\ (p_0) \times \frac{m^2}{2!}$
:	:
:	:
:	:
:	:
:	:
r	$N \times p\ (r) = \frac{Ne^{\ \ m} m^r}{r!}\quad N\left(p_o\right).\frac{m^r}{r!}$
:	:
:	:
:	:
:	:
:	:
योग (Total)	**N**

Illustration 16 : Fit a Poisson distribution to the following data and calculate the theoretical frequencies:

X	**:**	0	1	2	3	4
Y	**:**	123	59	14	3	1

Solution :

X	***f***	***fx***
0	123	0
1	59	59
2	14	28
3	3	9
4	1	4
	N = 200	**S*fx* = 100**

$$\overline{X} = \frac{fx}{N} = \frac{100}{200} = .5$$

Thus, $\overline{X}$ or mean m = .5

$(P_0) = e^{-m} = e^{-.5} = .6065$ (From the table)

$N\ (P_0)\ 200 \times .6065 = 121.3$

$$N(P_1) = N(P_0) \times m = 121.3 \times 0.5 = 60.65$$

$$N(P_2) = N(P_1) \times \frac{m}{2} = 60.65 \times \frac{0.5}{2} = 15.31$$

$$N(P_3) = N(P_2) \times \quad = 15.31 \times \frac{0.5}{3} = 2.55$$

$$N(P_4) = N(P_3) \times \quad = 2.55 \times \frac{0.5}{4} = 0.32$$

The theoretical frequencies nearest to the unit are :

X	**:**	0	1	2	3	4
Y	**:**	121	61	15	3	0

Illustration 17 : The probability that a man aged 50 years will die within a year is 0.01125. What is the probability that of 12 such men at least 11 will reach their fifty first birthday ?

Solution : Since the probability of death is very small, we use the Poisson Distribution.

Given $p = 0.01125$

$n = 12$

\ $m = np$

$= 12 \times 0.01125$

$= 0.135$

\ Probability that no body will die, *i.e.*, all the 12 persons will survive

$$(P_0) = e^{-m} = e^{-.135} = 0.8737$$

The probability that 1 person will die, *i.e.*, 11 persons will survive

$$(P_1) = P_0 \times m = .8737 \times .135 = 0.1179$$

\ Probability that 11 persons will survive

$= (P_0) + P_1$

$= 0.8737 + 0.1179$

$= 0.9916$

Illustration 18 : Letters were received in a college office on each of 100 days. Assuming the following data to form a random sample from a Poisson Distribution, find the expected frequencies taking e^{-4} = .0183

Number of letters	**:**	0	1	2	3	4	5	6	7	8	9	10
Frequency	**:**	1	4	15	22	21	20	8	6	2	0	1

Solution :

Given $e^{-4} = 0.183$

\ $P(0) = e^{-m} = e^{-4} = .0183$

$$N(P_0) = 100 \times .183 = 1.83$$

$$N(P_1) = N(P_0) \times m = 1.83 \times 4 = 7.32$$

$$N(P_2) = N(P_1) \times \quad = 7.32 \times \frac{4}{2} = 14.64$$

$$N(P_3) = N(P_2) \times \frac{m}{3} = 14.64 \times \frac{4}{3} = 19.52$$

$$N(P_4) = N(P_3) \times \frac{m}{4} = 19.52 \times \frac{4}{4} = 19.52$$

$$N(P_5) = N(P_4) \times \frac{m}{5} = 19.52 \times \frac{4}{5} = 15.62$$

$$N(P_6) = N(P_5) \times \frac{m}{6} = 15.62 \times \frac{4}{6} = 10.41$$

$$N(P_7) = N(P_6) \times \frac{m}{7} = 10.41 \times \frac{4}{7} = 5.95$$

$$N(P_8) = N(P_7) \times \frac{m}{8} = 5.95 \times \frac{4}{8} = 2.97$$

$$N(P_9) = N(P_8) \times \frac{m}{9} = 2.97 \times \frac{4}{9} = 1.32$$

$$N(P_{10}) = N(P_9) \times \frac{m}{10} = 1.32 \times \frac{4}{10} = 0.53$$

The expected frequencies to the nearest unit are :-

No. of letters	0	1	2	3	4	5	6	7	8	9	10
Frequency	2	7	15	20	20	16	10	6	3	1	1

Illustration 19 : A person has 2 cars for hire purpose. The demand of car for hire follows Poisson Distribution with mean 1.5. Compute the proportion of days on which (*i*) no car is used (*ii*) some demand of car is refused.

एक व्यक्ति के पास 2 कारें किराये के लिए हैं। कार की माँग पॉयसन वितरण पर आधारित है और समान्तर माध्य 1.5 है। उन दिनों का अनुपात बताओं जब (*i*) कोई कार प्रयोग नहीं की गयी (*ii*) कार की कुछ माँग अस्वीकार की गयी।

Solution : We known that

$$\textbf{P}(\boldsymbol{r})\textbf{ success} = \frac{\boldsymbol{e^{\ m}.m^2}}{\lfloor \boldsymbol{r}}$$

$$m = \overline{X} = 1.5 \text{ (given)}$$

(*i*) **Probability of no car is used means — P (0) success**

$$\text{P}(0)\text{ success} = e^{-m} = e^{-1.5} = e^{-1} \times e^{-.5}$$
$$= .36788 \times 0.6065$$
$$= 0.2231$$

(*ii*) **Probability that some demand is refused**

If implies that at the most we can give (0), (1) or (2) cars on hire.

If demand comes for 3 or more cars on a particular day, the owner will have to refuse the customers.

$$\text{P}(0)\text{ success} = e^{-m} = e^{-1.5} = .2231$$

$$\text{P}(1)\text{ success} = \frac{e^{\ 1.5}.1.5^1}{\lfloor 1} = .2231 \times 1.5 = .3346$$

$$\text{P}(2)\text{ success} = \frac{e^{\ 1.5}.1.5^2}{\lfloor 2} = \frac{.2231 \quad 2.25}{2 \quad 2} = 0.2510$$

We can accept the demand of cars on hire for 0, 1 and 2 cars and will have to refuse the demand of excess cars.

\ Required probability = 1 – [P (0) + P (1) + P (2) cars]

= 1 – [0.2231 + 0.3346 + 0.2510]

= 1 – .8087 = 0.1913 = 0.1913.

Illustration 20 : If X is a Poisson variate, compute the value of mean of distribution whose P (*r* = 1) = P (*r* = 2) success.

यदि (X) चर पॉयसन वितरण है तो वितरण का समान्तर माध्य ज्ञात कीजिए यदि P (*r* = 1) = P (*r* = 2) हो।

Solution : In Poisson Distribution, we know that —

$$\mathbf{P\ (r)\ success} = \frac{e^{\ m}.m^2}{\lfloor r}$$

$$\text{P (1) success} = \frac{e^{\ m}.m^1}{\lfloor 1}$$

$$\text{P (2) success} = \frac{e^{\ m}.m^2}{\lfloor 2}$$

$$P\ (r = 1) = P\ (r = 2)\ \text{success}$$

$$\frac{e^{\ m}.m^1}{\lfloor 1} = \frac{e^{\ m}.m^2}{\lfloor 2}$$

$$\frac{m^1}{1} = \frac{m^2}{2 \quad 1} \qquad [\because \text{e–m is common}]$$

$$= \frac{m^2}{2}$$

By cross multiplication, we have —

$$2m = m^2$$

or $$2 = \frac{m^2}{m}$$

Hence $$m = 2$$

$$\text{So mean of distribution} = \overline{X} = m = 2$$

Illustration 21 : In a Poisson distribution, probability of 3 successes is 2/3rd of probability of 1 success. Find the mean of the distribution.

Solution.

$$P\ (r) = \frac{e^{\ m}.m^r}{r!}$$

$$P\ (1) = \frac{e^{\ m}.m^1}{1!} = e^{-m}\ .\ m$$

$$P\ (3) = \frac{e^{\ m}.m^3}{3!} = e^{-m}\ .\ \frac{m^3}{6}$$

Now $$P_3 = \frac{2}{3}P_1$$

or $$e^{-m}.\frac{m^3}{6} = \frac{2}{3}\ e^{-m}.m \text{ or } \frac{m^3}{m} = \quad \times 6$$

$$m^2 = 4 \quad \backslash \quad m = 2$$

द्विपद व पॉयसन वितरण में अन्तर
(Difference Between Binomial and Poisson Distribution)

द्विपद व पॉयसन वितरण दोनों ही खण्डित (Discrete) वितरण हैं, किन्तु दोनों की वितरणों में अनेक भिन्नाताएँ हैं, जो निम्न प्रकार हैं —

(1) द्विपद वितरण में न्यादर्श का आकार (n) निश्चित होता है जबकि पॉयसन वितरण में न्यादर्श का आकार (Size of Sample *i.e.* n) बहुत अधिक होता है और यह आकार अनन्तता (Infinity) की ओर अग्रसित होता है।

(2) द्विपद वितरण में दो पारस्परिक अपवर्जी घटनाएं इस प्रकार से घटती हैं कि सफलता की प्रायिकता (p) मालूम होने पर असफलता की प्रायिकता को ($q = 1 - p$) द्वारा जाना जा सकता है। पॉयसन वितरण में प्रायः किसी घटना के घटने की सम्भावना प्रायः शून्य (Zero) ही होती है और घटना के न घटने की संभावना प्रायः एक (One) के बहुत समीप होती है।

(3) द्विपद वितरण में सफलता व असफलता के बारे में जानकारी होती है, जबकि पॉयसन वितरण में घटना की सम्भावना तो प्रायः पिछले अनुभव से मालूम होती है किन्तु घटना के न घटने के बारे में पता नहीं होता।

(4) द्विपद वितरण में वितरण बनाते समय सफलता व असफलता दोनों की ही प्रायिकता की आवश्यकता होती है। जबकि पॉयसन वितरण में यदि केवल समान्तर माध्य (Mean) दिया गया हो तो पॉयसन वितरण बनाया जा सकता है और सफलता व असफलता की प्रायिकता का पता होना अनिवार्य नहीं है।

(5) द्विपद वितरण के विभिन्न अचर (Constants) को जानने के लिए सफलता (p), असफलता (q) व न्यादर्श संख्या (n) का मालूम होना अनिवार्य है। जबकि पॉयसन वितरण के विभिन्न अचरों को केवल माध्य (mean *i.e.*, m) से पता चलते ही निकाला जा सकता है।

(6) पॉयसन वितरण में सफलता की प्रायिकता बहुत कम होती है (जैसे .01, .003, .04 इत्यादि)।

द्विपद वितरण उन परिस्थितियों में प्रयुक्त किया जाता है जब इस प्रायिकता का मान 0.10, 0.18 या इससे ज्यादा होता है। e^{-a} का मान ज्ञात हो जाने पर गणना करना आसान हो जाता है। निम्न सारणी में e^{-a} का मान दिया गया है, जिसको सन्दर्भित किया जा सकता है।

TABLE OF VALUES OF e^{-a}
(a lying between 0 and +1)

a	0	1	2	3	4	5	6	7	8	9
0.0	1.000	.9900	.9802	.9704	.9608	.9512	.9418	.9324	.9231	.9139
0.1	.9048	.8958	.8869	.8781	.8694	.8607	.8521	.8437	.8353	.8207
0.2	.8237	.8106	.8025	.7945	.7866	.7788	.7711	.7634	.7558	.7483
0.3	.7408	.7334	.7261	.7189	.7118	.7047	.6977	.6907	.6839	.6771
0.4	.6703	.6636	.6570	.6505	.6440	.6376	.6313	.6250	.6188	.6126
0.5	.6065	.6005	.5945	.5886	.5827	.5770	.5712	.5655	.5599	.5543
0.6	.5488	.5434	.5379	.5326	.5276	.5220	.5196	.5117	.5066	.5016
0.7	.4933	.4916	.4868	.4819	.4771	.4724	.4677	.4630	.4584	.4538
0.8	.4493	.4449	.4404	.4360	.4317	.4274	.4232	.4190	.4148	.4107
0.9	.4066	.4025	.3985	.3946	.3906	.3867	.3829	.3791	.3753	.3716

सामान्य वितरण
(Normal Distribution)

अँग्रेज गणितज्ञ डी. मॉयर (D. Moivre) (1667- 1754) ने सर्वप्रथम 'सामान्य वितरण' ज्ञात किया था। फ्रेंच गणितज्ञ लैप्लेस (Laplace) (1749-1827) ने इस सिद्धान्त की पुनः खोज की तथा इसका प्रयोग प्राकृतिक तथा सामाजिक विज्ञानों तथा

व्यावहारिक मामलों में किया। जर्मन गणितज्ञ, भौतिकशास्त्री तथा खगोलशास्त्री गॉस (Gauss) (1777-1855) ने इस सिद्धान्त का विकास, विस्तार तथा व्यावहारिक उपयोग किया। बेल्जियन संख्याशास्त्री तथा खगोलशास्त्री क्यूटलेट (Quetelet) न सबसे पहले सामान्य वितरण का प्रयोग सामाजिक समंकों में व्यापक रूप से किया। जीव-विज्ञान (Biology) में इस वितरण का प्रयोग का श्रेय प्रसिद्ध अँग्रेज संख्याशास्त्र पर फ्रान्सिस गाल्टन (Sir Francis Galton) (1822-1911) को है जो चार्ल्स डार्विन के चचेरे भाई थे।

सांख्यिकी सिद्धान्त में सामान्य वितरण सबसे महत्त्वपूर्ण सम्भावना-वितरण (Probability Distribution) है। अधिकांश न्यादर्श समंकों को सम्भावना-वितरण से ही ज्ञात किया जाता है तथा दोनों में घनिष्ठ सम्बन्ध होता हैं सांख्यिकी विज्ञान में सामान्य वितरण का महत्व इसलिए अधिक है कि मौलिक समंकों का सामान्य वितरण न होने पर भी, न्यादर्श से ज्ञात विभिन्न मापों के वितरण की प्रवत्ति सामान्य वितरण की ओर पायी जाती है। सामान्य वितरण वक्र को 'सामान्य वक्र' (Normal Curve) सामान्य सम्भावना वक्र (Normal Probability Curve) अथवा सामान्य त्रुटि वक्र (Normal Curve or Error) भी कहते हैं। आधुनिक संख्याशास्त्री डब्ल्यु. जे. यूडन (W.J. Youden) ने इस वक्र की प्रशंसा करते हुए लिखा है, ''सामान्य वक्र मानव जाति के अनुभव में प्राकृतिक दर्शन का सबसे व्यापक सिद्धान्त है। भौतिक तथा सामाजिक विज्ञानों तथा चिकित्सा, कृषि व इन्जीनियरिंग में यह एक निर्देशक उपकरण का कार्य करता है। अवलोकन तथा प्रयोग से प्राप्त आधारभूत समंकों के विश्लेषण तथा निर्वचन में यह एक आवश्यक यन्त्र है। सर फ्रान्सिस गाल्टन के शब्दों में, ''अतर्क का यह सर्वोच्च नियम है। जब भी मिले-जुले तत्त्वों का एक बड़ा न्यादर्श चुना जाता है तथा उनके मापानुसार उन्हें अनुविन्यसित किया जाता है तो बिना संदेह के एक सर्वसुन्दर नियमितता सिद्ध होती है।''

प्रसामान्य वक्र निम्न समीकरण (Equation) द्वारा व्यक्त किया जाता है —

$$y = \frac{1}{\sqrt{2}} e^{\,x^{2/2}\ \ 2}$$

जहाँ

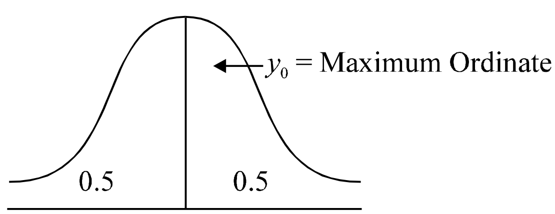
y = माध्य x की दूरी पर एक कोटि की परिकलित ऊँचाई (the computed height of an ordinate at a distance of x from the mean)

s = प्रसामान्य वितरण का प्रमाप वितरण (standard deviation of the normal distribution)

p = (the constant = 22/7 = 3.1417; $\sqrt{2}$ = 25066)

e = (the constant = 2.7183, the base of the system of natural logarithms)

x = $(X - \overline{X})$



Normal Curve

सामान्य वक्र का कुल क्षेत्रफल एक (1) होता है। यदि इसी समीकरण को 'N' से गुणा करें तो समस्त वितरण की विभिन्न आवत्तियों को जाना जा सकता है। ऐसी स्थिति में सूत्र इस प्रकार से व्यक्त करेंगे : -

$$y = \frac{N}{\sqrt{.2}} e^{\frac{1}{2}\left(\frac{x}{\ }\right)^2}$$

अधिकतम कोटि अक्ष (Maximum Ordinate) को y_0 कहते हैं और $y_0 = \frac{N}{\sqrt{.2}}$ होता है। अतः उपरोक्त सूत्र को निम्न सूत्र से भी लिख सकते हैं :

$$y = y_0 \frac{x^2}{2.\ \ ^2}$$

सामान्य वक्र की विशेषताएँ
(Properties of Normal Curve)

(1) यह घंटाकार आकृति (Bell-shaped) की संमितीय वक्र होती है।

(2) सामान्य वक्र में एक ही भूयिष्ठक (Monomodal) होता है तथा वह शीर्ष (Peak) बिन्दु से दोनों ओर संमितीय होती है। यह वक्र दोनों ओर अनन्त तक फैली होती है, अर्थात् वक्र भुजाक्ष के निकट आती जाती है, परन्तु उसको कभी भी स्पर्श नहीं करती है।

(3) केन्द्रीय प्रवत्ति की सभी मापें समान होती हैं तथा वे उच्चतम कोटि अक्ष (Ordinate) पर स्थित होती हैं।

(4) वितरण के सभी अवलोकन वक्र के अन्दर भुजाक्ष के ऊपर ही होते हैं।

(5) माध्य कोटि अक्ष (Mean Ordinate) वक्र को दो समान भागों में विभक्त करता है। आवत्तियों का वितरण वक्र के एक ओर जितना होता है, ठीक उतना की वक्र के दूसरी ओर भी होता है।

(6) माध्य मूल्य के निकट वक्र अवतल (Concave) होता है तथा $\pm 3s$ के निकट समतल आधार के उत्तल (Convex) के रूप में होता है। वक्र $\pm 1s$ के बिन्दुओं पर मुड़ता है।

(7) माध्य के दोनों ओर $.6745\ s$ के विस्तार के अन्तर्गत 50% आवत्तियाँ आ जाती हैं। यही सम्भाव्य विभ्रम होता है।

(8) प्रथम तथा ततीय चतुर्थांश मध्यका के समान दूरी पर होते हैं।

(9) चतुर्थांश विचलन (Quartile Deviation) = सम्भाव्य विभ्रम (Probable Error).

(10) माध्य विचलन (d) प्रमाप विचलन (s) का .7979 या 4/5 होता है। यदि माध्य विचलन को प्रथम चतुर्थांश में जोड़ दिया जाये तो ततीय चतुर्थांश में से घटा दिया जाये तो माध्य = मध्यका = भूयिष्ठका का मान ज्ञात हो सकता है।

(11) सामान्य वक्र तथा भुजाक्ष के मध्य का क्षेत्र 'वक्र के अंतर्गत क्षेत्र' कहलाता है तथा उस क्षेत्र में आवत्तियों के वितरण की संख्या स्पष्ट करता है।

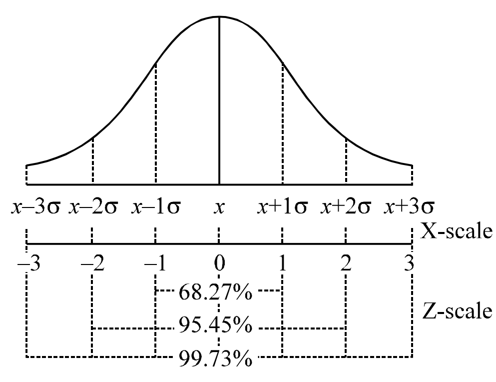
सामान्य वक्र के अन्तर्गत क्षेत्र का माध्य (*a*) तथा प्रमाप विचलन के आधार पर वितरण इस प्रकार होता है :-

(*i*) $\overline{X} \pm 1s$ के अन्तर्गत 68.27% आवत्तियाँ होती हैं। माध्य के प्रत्येक ओर 34.135% आवत्तियाँ होती हैं।

(*ii*) $\overline{X} \pm 2s$ के अन्तर्गत 95.45% आवत्तियाँ होती हैं। माध्य के प्रत्येक ओर 47.725% आवत्तियाँ होती हैं।

(*iii*) $\overline{X} \pm 3s$ के अन्तर्गत 99.73% आवत्तियाँ होती हैं। माध्य के प्रत्येक ओर 49.865% आवत्तियाँ होती हैं।

यह विशेषता निम्न चित्र द्वारा स्पष्ट जो जाएगी —



चित्र : प्रसामान्य वक्र : क्षेत्रफल सम्बन्ध
(Normal-Curve : Area Relationship)

प्रसामान्य वितरण के अचल (Constant of Normal Distribution)

प्रसामान्य वितरण के अचल इस प्रकार हैं :-

(1) समान्तर माध्य $\overline{X}$

(2) प्रमाप विचलन = s

(3) प्रथम परिघात अर्थात $m_1 = 0$

(4) द्वितीय परिघात अर्थात् $m_2 = s^2$

(5) ततीय परिघात अर्थात् $m_3 = 0$

(4) चतुर्थ परिघात अर्थात् $m_4 = 3 \quad {}^2_2 = 3s^4$

(7) $b_1 = \frac{\frac{2}{3}}{\frac{3}{2}} = 0$

(4) $b_2 = \frac{4}{\frac{2}{2}} = \frac{3 \ \frac{2}{2}}{\frac{2}{2}} = \frac{3 \ \ 4}{4} = 3$

प्रसामान्य वितरण के लिए b_2 का मूल्य 3 होना चाहिए। यदि b_2 का मूल्य '3' है तो वक्र 'सामान्य' या 'मध्य शीर्ष वाला' (Mesokurtic) वक्र होता है। यदि b_2 का '3' मूल्य से अधिक है तो वक्र 'लम्बे शीर्ष वाला' (Leptokurtic) वक्र होता है और यदि b_2 का मूल्य '3' से कम है तो वक्र 'चपटे शीर्ष वाला' (Platykurtic) वक्र होता है।

प्रसामान्य वक्र के अन्तर्गत क्षेत्रफल (Area Under the Normal Curve)

प्रसामान्य वक्र समान्तर माध्य ($\overline{X}$) तथा प्रमाप विचलन (s) के मानों पर निर्भर करता है। $\overline{X} = 0$ तथा $= 1$ के लिए प्रमापित प्रसामान्य वक्र (Standard Normal Curve) का निर्माण किया जाता है। ऐसे प्रसामान्य वितरण को जिसमें $\overline{X} = 0$ तथा $s = 1$ प्रमापित प्रसामान्य वितरण (Standard Normal Distribution) या इकाई प्रसामान्य वितरण (Unit Normal Distribution) कहते हैं।

समान्तर माध्य ($\overline{X}$) तथा प्रमाप विचरण (s) वाले प्रसामान्य वितरण को $\overline{X} = 0$, $s = 1$ वाले प्रमापित प्रसामान्य वितरण के रूप में निम्न सूत्रानुसार परिणत किया जा सकता है :

$$Z = \frac{X \quad \overline{X}}{} \text{ or } \frac{x}{6} \qquad \left[\because x \quad X \quad \overline{X}\right]$$

सामान्य वितरण का महत्त्व
(Importance of Normal Distribution)

सैद्धान्तिक वितरणों में सामान्य वितरण का बहुत महत्त्व है। यहाँ तक कि यदि द्विपद वितरण में $p = q = \frac{1}{2}$ हों तो रेखाचित्र द्वारा आवत्तियों को चित्रित करके संमित वक्र (Symmetrical Curve) प्राप्त होगा। किन्तु न्यादर्श संख्या (Size of Sample) के बढ़ने पर द्विपद वितरण के बजाय सामान्य वितरण अधिक उपयोगी सिद्ध होगा। अखण्डित वितरण होने के कारण सामान्य वितरण का महत्त्व और भी बढ़ जाता है। निम्न बातों से इस वक्र के महत्त्व का आभास होता है :-

(1) व्यवहार में द्विपद वितरण व पॉयसन वितरण दोनों को की सामान्य वितरण के आधार पर मापा जा सकता है।

(2) जब न्यादर्श का आकार बहुत बढ़ जाता है तो सामान्य वितरण ही अधिक उपयोगी सिद्ध होता है।

(3) सामान्य वितरण के अनेक गणितीय सम्बन्ध हैं, जिनसे इस वितरण की व्यावहारिक उपयोगिता और भी बढ़ जाती है। गणितीय सम्बन्ध जैसे :

$$\overline{X} \pm 1s = 68.27\% \text{ area}$$

(4) ऐसे वितरण जो असामान्य होते हैं, उन्हें संशोधित करके सामान्य वितरण का आकार दिया जा सकता है।

(5) सामान्य वक्र का सांख्यिकी गुणवत्ता नियन्त्रण (Quality Control) व औद्योगिक प्रयोगों में बहुत महत्व है।

(6) सामान्य वितरण न्यादर्श सिद्धान्त (Theory of Sampling) का आधार है। इसी आधार के माध्यम से यह जाना जा सकता है किस विशाल समग्र (Universe of Census) में से लिए गए न्यादर्श पूर्ण समग्र का प्रतिनिधित्व करते हैं अथवा नहीं।

सामान्य वितरण की मान्यताएं
(Assumptions of Normal Distribution)

सामान्य वितरण निम्न मान्यताओं पर आधारित हैं :-

(1) विभिन्न घटनाओं को प्रभावित करने वाले कारण स्वतंत्र होते हैं।

(2) घटनाओं को प्रभावित करने वाले कारण अनेकों हैं व समान महत्व के हैं।

(3) विभिन्न कारणों का प्रभाव इस प्रकार हो कि अधिकतम आवत्तियाँ समान्तर माध्य ($\overline{X}$) के दोनों ओर हों व समान्तर माध्य से विचलन (Deviation from $\overline{X}$) आपस में सन्तुलित होकर शून्य (Zero) हो जाते हैं।

(4) सामान्य वक्र की सभी कारण राशियाँ समस्त समग्र (Universe) व लिए गए न्यादर्श (Samples) पर समान हों, यद्यपि उनका प्रभाव घटना पर अलग-अलग ही हो।

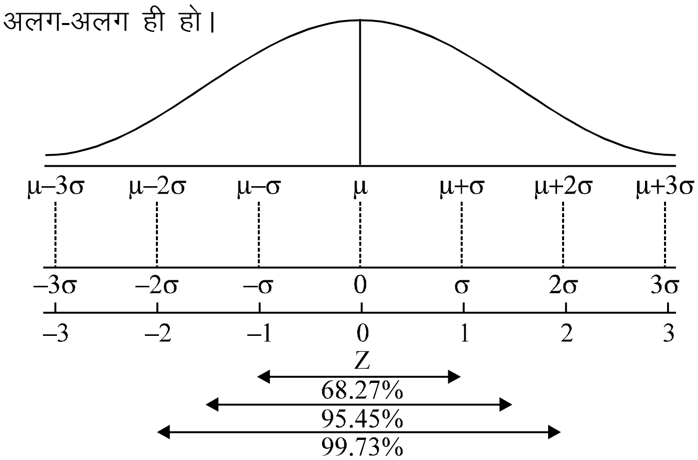


Illustration 22 : Find the area under the normal curve for $z = 1.2$.

Solution : In the table, the entry corresponding to $z = 1.2$ is 0.3849 which measures the shaded area in the following figure between $z = 0$ and $z = 1.2$.

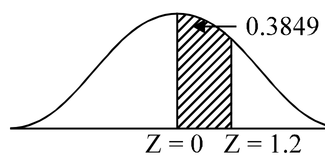


Illustration 23 : Find the area between $z = -0.68$ and $z = 0$

Solution : The table given at the end of the book does not contain entries relating to negative values of z. But since the curve is symmetrical, we can find the area between $z = 0$ and $z = -0.68$ by looking the area relating to $z = 0.68$. Therefore, the entry corresponding to $z = 0.68$ is 0.2157 and its measures the shaded area in the following figure between = 0 and $z = -0.68$.

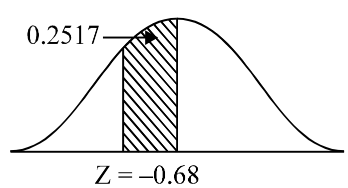


Illustration 24 : Find the area between $z = -0.46$ and $z = 2.21$

Solution : Required area = (area between $z = 0.46$ and $z = 0$)
+ (area between $z = 0$ and $z = 2.21$)
= (area between 0 and $z = 0.46$) + (area between $z = 0$, and $z = 2.21$)
= 0.1772 + 0.4864 = 0.6636 which shows the shaded area in the following figure.

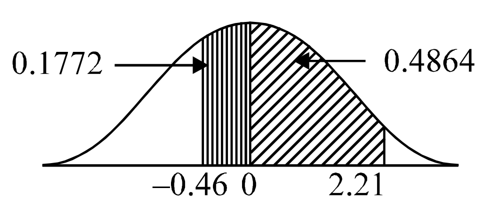


Illustration 25 : Determine the value of *z* in each of the following cases where area refers to that under the normal curve :-

(*i*) Area between 0 and *z* is 0.3770

(*ii*) Area to the left of *z* is 0.8621

Solution :

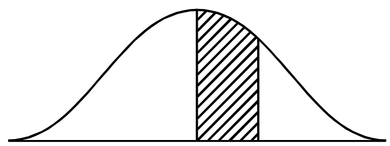
(*i*) Area between 0 and z is 0.3770

In the table at the end of the book the entry 0.3770 is located to the right of the row marked 1.1 and under the column marked 6. Then the required $z = 1.16$.

By Symmetry, $z = 1.16$ is another value of z

Thus $z = \pm 1.16$

The required value of $z = 1.16$ is shown in the following shaded figure.



(*ii*) Area to the left of z is 0.8621

Since the area is greater than 0.5000, z must be positive area between and $z = 0.8621 - 0.5000 = 0.3621$ which is located to right of the row marked and under the column marked 9 in the table at the end of the book. Then $z = 1.0$.

The shaded area in the following figure shows area to the left of z is 0.8621.

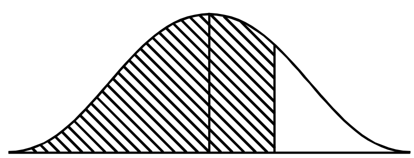


Illustration 26 : A normal curve has mean 70 and standard deviation 20. (*i*) Find the area under normal distribution between 60 and 120. (*ii*) What is the area under normal distribution between 50 and 100 ?

Solution : We are given

$$\overline{X} = 70 \text{ and } q = 20$$

Applying formula of noraml curve, we get

$$Z = X - \overline{X} \div q$$

(*i*) $ZX = 60$

$= 60 - 70 \div 20 = -10 \div 20 = -0.5$

$ZX = 120, 120 - 70 \div 20 = 50 \div 20 = 2.5.$

Referreing to Z-table we find the area corresponding to the Z = (–) 0.5 and Z = 2.5 are 0.1915 and 0.4938. The desired area between 80 and 120 is 0.1915 + 0.49387 = 0.6853.

(*iii*) $ZX = 50, 50 - 70 \div 20 = -20 \div 20 = -1.00$

$ZX = 100, 100 - 70 \div 20 = 30 \div 20 = 1.50.$

The corresponding to Z = – 1.00 and Z = 1.50 are 0.3413 and 0.4332. Thus, the total area under normal curve is 0.3413 + 0.4332 = 0.7745. In other words, 7745 per cent of area lies between 50 and 100.

Illustration 27 : The customer accounts at a certain department store have an average balance of Rs. 120 and standard deviation of Rs. 40. Assuming that the account balances are normally distributed.

(*i*) What proportion of the accounts is over Rs. 150 ?

(*ii*) What proportion of the accounts is between Rs. 100 and Rs. 150 ?

(*iii*) What proportion of the accounts is between Rs. 60 and Rs. 90 ?

Solution : We are given,

$$\overline{X} = 120 \text{ and } q = 40$$

Formula, $$Z = X - \overline{X} \div q$$

(*i*) Proportion of the accounts is over Rs. 150

$$ZX = 150 = 150 - 120 \div 40 = 0.75$$

Referring to Z-table the area under Z = 0.2734.

We have to find probability of items falling to the right of Z = 0.75 *i.e.*, over Rs. 150, deduct the value of 0.2734 from the total probability to the right origin, 0.5 – 0.2734 = 0.2266. Hence, 22.66 per cent of the accounts have a balance in excess of Rs. 150.

(*ii*) Propotion of the accounts between Rs. 100 and Rs. 150

$$ZX = 100 = 100 - 120 \div 40 = -0.50$$

$$ZX = 150 = 150 - 120 \div 40 = 0.75$$

Referring to Z-table, area between

$$Z = -0.50 \text{ and } Z = 0.75$$

$$= 0.1915 + 0.2734 = 0.4649$$

Therefore, 46.49 per cent of the accounts have an average between Rs. 100 and 150.

(*iii*) Proportion of the accounts betweenRs. 60 and Rs. 90

$$ZX = 60 = 60 - 120 \div 40 = -1.5$$

$$ZX = 90 = 90 - 120 \div 40 = -0.75$$

Referring to Z-table

Area corresponding to Z = – 1.50 = 0.4332

Area corresponding to Z = – 0.75 = 0.2734

Area between Z = – 1.50 and Z = – 0.75 is 0.4332 – 0.2734 = 0.1598

Thus, 15.98 per cent of the accounts is an average between Rs. 60 and Rs. 90.

Illustration 28 : If the height of 1000 students in a University Campus closely follows a normal distribution with a mean of 69 inches and a standard deviation of 2.5 inches, how many out of 1000 students would you expect to be atleast 6 feet in height ?

Solution : Assuming the distribution of height to be normal

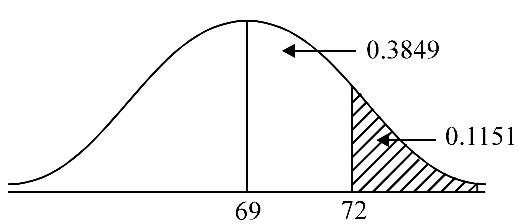
Standard normal variate, $$Z = \frac{X \quad \overline{X}}{}$$

Here $$X = 72 \text{ inches}$$

$$\overline{X} = 69 \text{ inches}$$

$$s = 2.5 \text{ inches}$$

\ $$Z = \frac{72 \quad 69}{2.5} \; 1.2$$



Required Students

$$= \text{area to the right of } Z = 1.2$$
$$= (\text{area to the right of } Z = 0) - (\text{area between } Z = 0 \text{ and } 1.2)$$
$$= 0.5000 - 0.3849 = 0.1151$$

Thus the number of students having at least 6 feet heights

$$= 0.115 \times 1000 = 11.51 = 12 \text{ approximately}$$

Illustration 29 : The Anantapur Municipal Corporation has installed 500 sodium lamps. The life of distribution is normally distributed with mean 4400 hours and standard deviation 525 hours. Find the number of bulbs that will burn out in (*i*) and (*ii*).

(*i*) in the first 3,300 hours of burning, (*ii*) between 3,000 and 5,4000 hours of burning, (*iii*) after what period of time would you expect only 20 per cent of lamps to be still burning.

Solution. We are given

$$\bar{X} = 4{,}400 \text{ hours and } q = 525 \text{ hours.}$$

Formula,

$$Z = \frac{X \quad \bar{X}}{}$$

(*i*) $$Z = (3300 - 4400) \div 525 = (-)\ 2.10$$

Z-table value corresponding Z (–) 2.10=0.4772

Since we have to find the probability of lamps falling to the left of Z = 0.5, deduct 0.4772 from 0.5, then remains 0.0228.

Number of lamps that would fail in 3,300 hours of burning, 0.0228 × 500 = 11

(*ii*) $$ZX = 3000 = 3000 - 4400 \div 525 = (-)\ 2.67$$
$$ZX = 5400 = 5400 - 4400 \div 525 = 190$$

Area between (Z) (–) 2.67 and Z 1.90 is 0.9675 (0.4962 + 0.4713).

Number of lamps that would fail between 3,000 adn 5,400 hours of burning is 0.9675 × 500 = 484.

(*iii*) $$\text{20 per cent of lamps} = 500 \times \frac{20}{100} = 100$$

$$\text{Area under which 100 lamps fall} = \frac{100}{500} = 0.2$$
$$\text{Area under which 400 lamps fall} = 0.5 - 0.2 = 0.3$$
$$\text{Z-value for Z-table value of 0.3} = 0.84$$

Now, $$ZX = \frac{X \quad 4400}{525} = 0.8$$
$$= X - 4400 = 0.84 \times 5.25$$
$$X = 4400 + 441$$
$$= 4{,}841 \text{ hours.}$$

Illustration 30 : Monthly rent is normally distributed about a mean of 56.50 rupees. With a standard deviation of 16.23 rupees. A random sample of 300 housed is taken. If the rents of these houses are arranged as a frequency distribution with classed under 10 rupees, under 20 rupees, under 30 rupees and so on, what frequency would you expect in each class according to Noraml Distribution.

Solution :

Rent in Rs.	Lower Limit of the class		Area Between Lower limit and central ordinate	Area within classes	Expected Frequency col. 5 × 300
under 10	0	= – 3.48	.5000	.0020	0.6
10-20	10	– 2.87	.4980	.0102	3.1
20-30	20	– 2.25	.4878	.0393	11.8
30-40	30	– 1.63	.4485	.1024	30.7
40-50	40	– 1.02	.3461	.1907	57.2
50-60	50	– 0.40	.1554	.2425	72.8
60-70	60	0.22	.0871	.2096	62.9
70-80	70	0.83	.2967	.1298	38.9
80-90	80	1.45	.4265	.0538	16.1
90-100	90	2.06	.4803	.0160	4.8
100-110	100	2.68	.4963	.0037	1.1
Total				**1.0000**	**300.0**

नोट : (*i*) कॉलम 4 में दी गयी संख्याएँ Table-ordinates and Area of the normal curve से ज्ञात की गई हैं।

(*ii*) कॉलम 5 की संख्याएँ कॉलम 4 की संख्याओं के क्रमानुसार घटाकर प्राप्त किया गया है, केवल मध्य वाले वर्ग के क्षेत्रों को जोड़ा गया है।

Illustration 31 : The wages of 10,000 workers were found to be normally distributed with mean = Rs. 750 per month and standard deviation = Rs. 50. Show that out of the workers 95% had wages exceeding Rs. 668 and only 5% had wages exceeding Rs. 832. What was the lowest wage among the richest 100 ?

Solution : We know that $Z = \frac{X \quad \bar{X}}{\ }$

$$= \frac{X \quad 750}{50}$$

When $X = 668$ $Z = \qquad = 1.64$

Area of right of Z = – 1.64 is 0.4495 + 0.5000 = 0.9495

Hene the required proportion, who had wages above Rs. 668.

= 0.9495 × 100 = 95%

Similarly when $X = 832$

$Z = \qquad = 1.64$

Area to the right of Z = 1.64 is 0.5000 – 0.4495 = 0.0505

Hence, the required proportion = 0.0505 × 100

= 5%

$$\text{Probability of getting richest 100} = \frac{100}{10,000} = 0.01$$

$$\text{Standard normal variate having 0.01 area to the right} = 2.33$$

\ $$2.33 = \frac{X \quad 750}{50}$$

or $$2.33 \times 50 = X - 750$$

\ $$116.50 + 750 = X$$

or $$X = 866.5$$

Thus the lowest wage of the richest 100 workers was Rs. 866.50

Illustration 32 : For a normal distribution in which N = 100 mean is 40 and standard deviation is 8. Find (*a*) the first quartile (*b*) the thired quartile and (*c*) the inter-quartile range.

Solution.

$$Q_1 = \quad -\frac{2}{3}s$$

$$= 40 - \frac{2}{3} \times 8$$

$$= 34.67$$

$$Q_2 = \overline{X} + \frac{2}{3}s$$

$$= 40 + \frac{2}{3}X \times 8$$

$$= 40 + \frac{16}{3}$$

$$= 45.33$$

$$\text{Inter-quartile Range} = Q_3 - Q_1$$

$$= 45.33 - 34.67$$

$$= 10.66$$

Illustration 33 : The height of 22-years old man can be considered for recruitment with mean 63" and standard deviation 2.50". A person of this age group is eligible for recruitment if his height is between 62" and 65". Find the expected number of persons ineligible because of excess height in a group of 180 randomly selected men.

Solution : We are given

$$\overline{X} = 63" \qquad q = 2.5"$$

Number of men whose height is more than 65" is :

$$ZX = 65 = 65 - 63 \div 2.5 = 0.08$$

The standard normal variate corresponding to 0.8 is 0.2881. Area to the right of ordinate at 0.8 is

$$0.5 - 0.2881 = 0.2119$$

Hence, number of men rejected because of excess height is 0.2199 × 180 = 38.142 or 38.

Illustration 34 : A student secures 63 marks in Accounting in M.Com. class of which his class average is 51 and the standard deviation 12. He secures 92 marks in Quantitative Methods for which his class average is 62 and the standard deviation 20.

What can we say about the performance of this student with reference to these two examiners.

Solution : Let marks of the students in Accounting and Quantitative Methods be denoted by X and Y respectively. Then we are given

$$X = 63, \ \overline{Y} = 51 \quad qX = 12$$
$$Y = 92, \ \overline{X} = 62 \quad qY = 20$$

Standard score in Accounting is :

$$X - \overline{X} \div qX = 63 - 51 \div 12 = 1.0$$

Standard score in Quantitative Methods is :

$$Y - \overline{Y} \div qY = 92 - 62 \div 20 = 1.5$$

Since the standard score in Quantitative Methods is much greater than the standard score in Accounting, the performance of the candidate is much better in Quantitative Methods as compared with his performance in Accouting.

Illustration 35 : In a normal distribution 31% of the items are under 45 and 8% are over. Find the mean and standard deviation.

Solution : The 31% items are under 45, *i.e.*, the area lying left to the ordinate at X = 45 = .31, *i.e.*

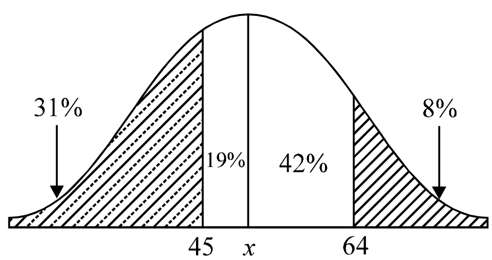
$$\text{The area lying between } 45 - \overline{X} = .5 - .31 = .19$$
$$\text{The value of Z corresponding to this area} = .5$$

Thus, $$\frac{45 \quad \overline{X}}{} = -.5$$

or $$45 - \overline{X} = -.5 \text{ s}$$

or $$- \overline{X} + .5 \text{ s} = 45$$

8% of the items are above 64, *i.e.*, the area to the right of the ordinate at 64 is .08. The area to the left of the ordinate at X = 64 is equal to .5 – .08 = .42, the value of Z corresponding to that area is 1.4.



Thus, $$\frac{64 \quad \overline{X}}{} = 1.4$$

or $$64 - \overline{X} = 1.4 \text{ s}$$

or $$- \overline{X} - 1.4 \text{ s} = 64$$

$$- \overline{X} + .5 \text{ s} = -45$$
$$- \overline{X} - 1.4 \text{ s} = -64$$
$$+ \qquad + \qquad + \qquad \text{(Deduct)}$$
$$1.9\text{s} = 19$$
$$\text{s} = 10$$

$$\overline{X} - .5 \text{ s} = 45 \qquad \text{or} \qquad \overline{X} - .5 \times 10 = 45$$
$$\overline{X} - 5 = 45 \qquad \text{or} \qquad X = 45 + 5 = 50$$

Illustration 36 : The mean and standard deviation of a graduation examination, following normal distribution are 500 and 100 respectively. If 550 students are to be passed out of 746 students, what would be the minimum passing marks ?

किसी स्नातक परीक्षा में प्राप्तांकों का बंटन सामान्य है जिसका माध्य 500 तथा प्रमाप विचलन 100 है। यदि परीक्षा में 674 छात्रों को पास करना हो तो न्यूनतम पास अंक क्या होने चाहिए ?

Solution.

$$\text{Total number of students} = 674$$

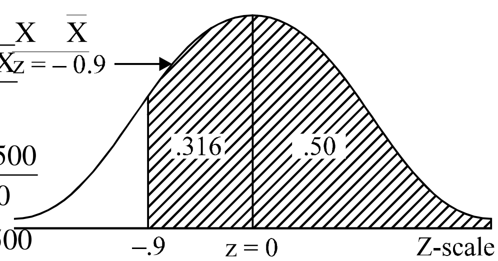
$$\text{Students to be passed} = 550$$

$$\text{Proportion of passed students} = \frac{550}{674} = .816\ [.5 + .316]$$

Total students passed covers 0.816 area of a normal curve from right to left side of $z = 0$. The area of 0.316 on left side of $z = 0$ will be at $z = -0.9$

We know that —

$$Z = \frac{X \quad \bar{X}}{}$$



or $$-0.9 = \frac{X \quad 500}{100}$$

or $$-90 = X - 500$$

$$500 - 90 = X$$

$$X = 410$$

Hence, the minimum passing marks are 410.

Illustration 37 : The mean inside the diameter of a sample of 500 washers produced by a machine 5.02 mm and the standard deviation is 0.05 mm. The purpose for which these washers are intended allows a maximum tolerance in the diameter of 4.96 to 5.08 otherwise washers are considered defective. Determine the percentage of defective washers produced by the machine, assuming the diameters are normally distributed.

Solution.

$$4.96 \text{ in standard units} = \frac{X \quad \bar{X}}{} = \frac{4.96 \quad 5.02}{0.05} = -1.2$$

$$5.08 \text{ in standard units} = \quad = \frac{5.08 \quad 5.02}{0.05} = 1.2$$

The proportion of non-defective washers = area under the normal curve be $x = -1.2$ and $x = 1.2$ or = twice the area between $x = 0$ and $x = 1.2$

$$= 2\ (0.3849) = 0.7698 \text{ or } 77\%$$

It is shown in the following figure.

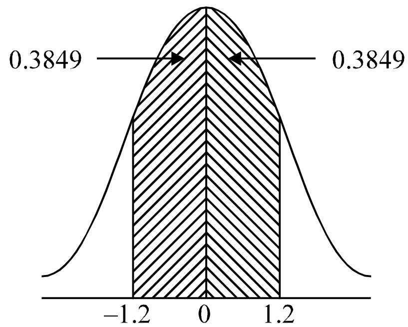


Illustration 38 : The following table give frequencies of occurence of a variable X between certain limits.

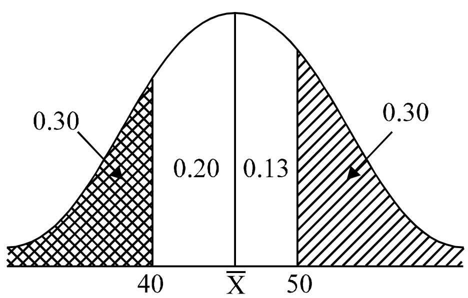
चर X (Variable X)	**आवत्ति (Frequency)**
40 से कम (Less than 40)	30
40 या 50 से अधिक परन्तु 50 से कम (40 or more but less than 50)	33
50 या 50 से अधिक (50 and more)	37
योग Total	100

वितरण पूर्णतया प्रसामान्य है। X का समान्तर माध्य तथा प्रमाप विचलन ज्ञात कीजिए।

The distribution is exactly normal. Find the average and standard deviation.

Solution :

Since 30% of the items are less than 40%, 20% of the items are in between and the mean. Similarly the percentage of items between mean and 50 is 13% (the items are 50 or more). The following figure illustrates this situation.



From the table the Z value corresponding to 0-2 (20% area) = – 0.5244

$$\begin{bmatrix} Z \quad 0.52 \text{ for } 0.1985 \\ Z = 0.53 \text{ for } 0.2019 \\ Z = 0.5244 \text{ for } 0.2000 \end{bmatrix}$$

\ $$Z_1 = \frac{40 \quad \overline{X}}{} = -0.5244 \text{ or } 40 - \overline{X} = -0.5244\ s \qquad ...(i)$$

From the table the Z value corresponding to 0.13 (13% area) = 0.3318.

$$Z_2 = \frac{50 \quad \overline{X}}{} = 0.3318 \text{ or } 50 - \overline{X} = 0.3318\ s \qquad ...(ii)$$

From equations (*i*) and (*ii*), we get

$$40 - \overline{X} = -0.5244s$$
$$50 - \overline{X} = -0.3318s$$
$$-\ +\qquad -$$
$$-10 = -0.8562\ s$$

\ $$s = \frac{10}{0.8562} = 11.68$$

Substituting the value of s in equation (*i*), we get

$$40 - \overline{X} = (-0.5244 \times 11.68)$$
$$40 - \overline{X} = -6.12$$
$$40 + 6.12 = \overline{X},\ \overline{X} = 46.12$$

\ The mean of the distribution is 46.12 and standard deviation is 11.68.

प्रसामान्य वक्र आसंजन/अन्वायोजन (Fitting a Normal Curve) : प्रसामान्य वक्र का आसंजन (Fitting) निम्न दो विधियों द्वारा किया जा सकता है।

(1) कोटि-अक्ष विधि (Method of Ordinates)

(1) क्षेत्रफल विधि (Method of Area)

(1) कोटि-अक्ष विधि (Method or Ordinates) : एक प्रसामान्य वक्र में माध्य-कोटि की ऊँचाई को निम्न सूत्र द्वारा ज्ञात किया जा सकता है।

$$\left(\frac{Ni}{\ }\right)$$
$$y_0 = \frac{Ni}{\sqrt{2}}$$

जहाँ y_0 = माध्य कोटि की ऊँचाई (the ordinate at the mean)

N = कुल आवत्ति (total frequency)

i = वर्गान्तर का अन्तर (magnitude of class-interval)

s = प्रमाप विचलन (Standard Deviation)

p = एक स्थिर है जिसका मान 2.7183 है।

(A constant with a value of 2.7183; $\sqrt{2}\ \ = 2.5066$)

दूसरे शब्दों में माध्य कोटि अर्थात्

$$y_0 = \frac{Ni}{2.5066} \text{ अथवा } 0.3989\left(\frac{Ni}{\ }\right)$$

आसजित वक्र (Fitted Curve) का यह अधिकतम माध्य कोटि माध्य से 1s की दूरी पर कोटि की ऊँचाई (the height of the ordinate at a distance 1s from the mean) निम्न प्रकार से गणना की जाएगी :

$$y_1 = 0.3989\left(\frac{Ni}{\ }\right) e^{\ 1/2(1)^2} \qquad \text{अथवा} = 0.3989 \qquad e^{1/2(x/\ \ )^2}$$

इसी प्रकार माध्य कोटि से s दूरियों पर अन्य कोटियों (ordinates) की ऊँचाई ज्ञात की जा सकती है।