इकाई 1 सांख्यिकी
(Statistics)

1.1 प्रस्तावना (Introduction)

मनुष्य के जीवन में सांख्यिकी का बहुत महत्त्व होता है। हम प्रतिदिन अपने आसपास विभिन्न प्रकार के आँकड़ों से रूबरू होते है और इन आँकडों का प्रयोग कर अपने दैनिक जीवन में विभिन्न प्रकार के निर्णय लेते है। जैसे- खरीदारी करते समय बजट तय करना, वित्तीय निवेश के फैसले लेना आदि। सामान्यतः आँकडे दो प्रकार के होते है- मात्रात्मक (quantitative) और गुणात्मक (qualitative)। इन दो प्रकार के आँकडों का प्रयोग कर सांख्यिकी की मदद से हम अपनी आर्थिक समस्याओं का समाधान कर बेहतर निर्णय लेने में सहायक होते है। अत: मनुष्य जाने अनजाने में अपने दैनिक जीवन में कहीं न कहीं सांख्यिकी का प्रयोग अवश्य करता है।

1.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के बाद आप

- ✓ सांख्यिकी के अर्थ को समझ सकेंगे।
- ✓ सांख्यिकी की विभिन्न परिभाषाओं को समझने में सक्षम होंगे।
- ✓ सांख्यिकी के महत्त्व या उपयोगिता के बारे में अवगत होंगे।
- ✓ सांख्यिकी की सीमाओं को जान सकेंगे।
- ✓ सांख्यिकी की कुछ आधारभूत अवधारणाओं से परिचित हो सकेंगे।

1.3 सांख्यिकी का अर्थ (Meaning of Statistics)

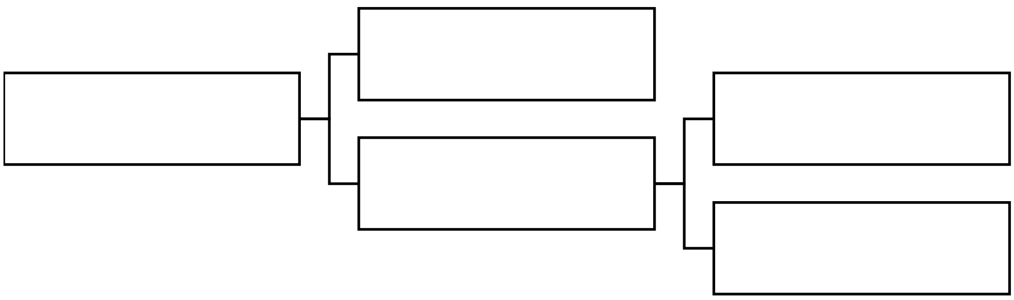
सांख्यिकी शब्द लैटिन भाषा के *'Status'*, या इटालियन भाषा के *'Statista'* शब्द से लिया गया है। सांख्यिकी का शाब्दिक अर्थ होता है- संख्याओं से संबंधित शास्त्र या विज्ञान। सांख्यिकी मुख्य रूप से मात्रात्मक (quantitative) आँकड़ों के संकलन, विश्लेषण और प्रस्तुतीकरण से संबंधित है। अर्थशास्त्र, मनोविज्ञान, समाजशास्त्र, इंजीनियरिंग आदि कई अन्य क्षेत्रों में भी सांख्यिकी का उपयोग किया जाता है। *Statistics* शब्द का प्रयोग सर्वप्रथम जर्मन विद्वान ***गॉटफ्रायड एकेनवाल*** **(Gottfried Achenwall)** ने वर्ष 1749 में किया था। इन्हें सांख्यिकी का जन्मदाता भी कहा जाता है।

ज्ञान-विज्ञान की किसी भी शाखा से संबंधित आँकडों को मात्रात्मक रूप में संकलित कर प्रस्तुत करने, उनका वैज्ञानिक विश्लेषण करने और उनसे तर्कपूर्ण निष्कर्ष निकालने की क्रियाओं का विधिवत् अध्ययन सांख्यिकी के अंतर्गत किया जाता है।

1.4 सांख्यिकी की परिभाषा (Definitions of Statistics)

विभिन्न विद्वानों द्वारा सांख्यिकी की अलग-अलग परिभाषाएँ दी गई हैं। व्याकरण के अनुसार सांख्यिकी शब्द का प्रयोग मुख्य रूप से दो प्रकार से प्रयोग में लाया जाता है – बहुवचन एवं एकवचन में।

बहुवचन में इसका अभिप्राय किसी निश्चित उद्देश्य से एकत्रित/संकलित मात्रात्मक आँकडों से होता है जबकि एकवचन में इसका अर्थ '**सांख्यिकीय विज्ञान**' **(Statistical Science)** से होता है जिसमे आँकड़ों के संकलन, प्रस्तुतीकरण, वर्गीकरण आदि सांख्यिकीय विधियों (Statistical Methods) का अध्ययन किया जाता है। इस वर्गीकरण के आधार पर सांख्यिकी की परिभाषाओं को दो वर्गों में बाँटा जा सकता है-



1.4.1 बहुवचन के रूप में सांख्यिकी की परिभाषा (Definition of Statistics in Plural Sense)

बहुवचन रूप में सांख्यिकी शब्द का आशय मात्रात्मक आँकडों से होता है। जैसे- जनगणना के आँकड़े, किसी सेमेस्टर में पंजीकृत शिक्षार्थियों की संख्या आदि। इस वर्ग में विभिन्न विद्वानों द्वारा सांख्यिकी को निम्न रूप में परिभाषित किया गया है-

बाउले (Bowley) के अनुसार, ***"सांख्यिकी का आशय अनुसंधान के किसी विभाग में संख्यात्मक तथ्यों के समूहों से है जो एक-दूसरे से संबंधित रूप में प्रकट किए जाते है। (Statistics are numerical statement of facts in any department of enquiry placed in relation to each other.)"***

यूल और केन्डाल (Yule and Kendall) *के अनुसार,* ***"सांख्यिकी से हमारा आशय ऐसे संख्यात्मक तथ्यों से है जो एक बड़ी सीमा तक अनेक कारणों से प्रभावित होता है। (By statistics we mean quantitative data affected to a market extent by multiplicity of causes.)"***

एच. सैक्रिस्ट (H. Secrist) के शब्दों में, ***"सांख्यिकी से हमारा आशय तथ्यों के उस समूह से है जो एक बड़ी सीमा तक अनेक कारणों से प्रभावित होतें है, संख्याओं के रूप में रखे जाते है, जिनकी गणना या अनुमान शुद्धता के एक उचित स्तर तक की जाती है, जिन्हें पूर्व निश्चित उद्देश्यों के लिए क्रमबद्ध रूप में एकत्रित किया जाता है एवं एक-दूसरे से संबंधित रूप में प्रस्तुत किया जाता है। (By statistics we mean aggregate of facts affected to a market extent by multiplicity of causes, numerically expressed, enumerated or estimated according to a reasonable standards of accuracy collected in a systematic manner for a predetermined purpose and placed in relation to each other.)"***

1.4.2 एकवचन के रूप में सांख्यिकी की परिभाषा (Definition of Statistics in Singular Sense)

एकवचन के रूप में सांख्यिकी का तात्पर्य सांख्यिकीय विधियों के अध्ययन से होता है। सांख्यिकी के निरंतर विकास को देखते हुए इस वर्ग की परिभाषाओं को दो उपशीर्षकों में विभाजित किया जा सकता है- परंपरागत दृष्टिकोण और आधुनिक दृष्टिकोण।

A. परंपरागत दृष्टिकोण (Traditional Approach)

परंपरागत दृष्टिकोण के अंतर्गत सांख्यिकी की सांख्यिकीय विधियों का अध्ययन किया जाता है। इस मत की परिभाषाएँ संकीर्ण (narrow) मानी जाती है। इस दृष्टिकोण के आधार पर सांख्यिकी की कुछ प्रमुख परिभाषाएँ निम्न हैं-

बाउले (Bowley) ने विज्ञान के रूप में सांख्यिकी को तीन प्रकार से परिभाषित किया है-

1. ***"सांख्यिकी गणनाओं का विज्ञान है। (Statistics is the science of counting.)"***
2. ***"सांख्यिकी को उचित रूप से माध्यों का विज्ञान कहा जा सकता है। (Statistics may rightly be called the science of averages.)"***
3. ***"सांख्यिकी वह विज्ञान है जो सामाजिक व्यवस्था को सम्पूर्ण मानकर उसके सभी रूपों का मापन करती है। (Statistics is the science of measurement of the social organism, regarded as a whole, in all its manifestations.)"***

पहली परिभाषा में सांख्यिकी को गणना करने की विधि बताया गया है लेकिन सांख्यिकी केवल गणना तक ही सीमित नहीं है। इसमें विश्लेषण एवं प्रस्तुतीकरण जैसी विधियों का अभाव दिखाई देता है। इसी प्रकार **द्वितीय** परिभाषा में सांख्यिकी को माध्यों का विज्ञान कहा गया है जबकि इसमें माध्यों के अतिरिक्त सहसंबंध, विषमता, अपकिरण, सूचकांक आदि विधियों का उपयोग भी होता है। **तृतीय** परिभाषा में सांख्यिकी का क्षेत्र मनुष्य और उसकी सामाजिक क्रियाओं तक ही सीमित किया गया है परन्तु वर्तमान समय में सांख्यिकी का उपयोग भौतिक विज्ञान, अंतरिक्ष विज्ञान, शिक्षाशास्त्र, चिकित्साशास्त्र आदि में भी होता है। **डॉ बाउले** स्वयं यह स्वीकार करते है कि ***"सांख्यिकी न तो केवल राज्य-अर्थशास्त्र की एक शाखा मात्र है और न ही वह किसी विज्ञान तक सीमित है।"***

क्रॉक्सटन एवं कॉउडेन (Croxton and Cowden) के अनुसार, ***"संख्यात्मक तथ्यों के संकलन, प्रस्तुतीकरण, विश्लेषण और निर्वचन को सांख्यिकी कहा जा सकता है। (Statistics may be defined as collection, presentation, analysis and interpretation of numerical data)"***

B. आधुनिक दृष्टिकोण (Modern Approach)

आधुनिक दृष्टिकोण की परिभाषाएँ व्यापक (broad) है एवं इसमें विभिन्न विधियों को सम्मिलित किया गया है। इस दृष्टिकोण के अंतर्गत कुछ प्रमुख परिभाषाएँ हैं-

वालिस एवं रॉबर्टस (Wallis and Roberts) के अनुसार, ***"सांख्यिकी रीतियों का एक समूह है जिसमे अनिश्चितता की परिस्थितियों में बुद्धिमत्तापूर्ण निर्णय लिया जा सके। (Statistics is a body of methods for making wise decision in the face of uncertainty.)"***

या-लुन-चाऊ (Ya-Lun-Chau) के अनुसार, ***"सांख्यिकी निर्णयन की एक विधि है जिसमें संख्यात्मक आँकडों एवं गणना की गई जोखिम के आधार पर अनिश्चितता की स्थिति में निर्णय लिए जाते हैं। (Statistics is a method of decision making in the face of uncertainty on the basis of numerical data and calculated risks.)"***

लॉरेन्स लापिन (Lawrence Lapin) के शब्दों में, ***"सांख्यिकी ऐसी रीतियों एवं सिद्धान्तों का समूह है जिसका प्रयोग संख्यात्मक प्रमाणों पर समय किया जाता है जबकि अनिश्चितिता की दशा में निर्णय लिया जाता है। (Statistics is a body of methods and theory that is applied to numerical evidences when making inferences in the face of uncertainty.)"***

सांख्यिकी की आधुनिक दृष्टिकोण की परिभाषाओं में इस बात पर अधिक बल दिया गया है कि सांख्यिकी विभिन्न विधियों एवं सिद्धांतों का समूह है जो अनिश्चितिता की स्थिति में सही निर्णय निकालने एवं विवेकपूर्ण निर्णय से संबंधित है।

उपर्युक्त परिभाषाओं का अध्ययन करने के बाद यह कहा जा सकता है की विभिन्न विद्वानों ने सांख्यिकी की अलग-अलग परिभाषाएँ दी हैं जिसमे कोई भी परिभाषा सम्पूर्ण प्रतीत नहीं होती है। सांख्यिकी की विभिन्न परिभाषाओं का अध्ययन करने के बाद निष्कर्ष के रूप में यह कहा जा है-

"सांख्यिकी विज्ञान और कला दोनों है, जो ज्ञान या अनुसंधान के किसी भी क्षेत्र में निश्चित उद्देश्य के लिए मात्रात्मक आँकडों के संकलन, प्रस्तुतीकरण, विश्लेषण और पूर्वानुमान में प्रयुक्त सिद्धान्तों एवं विधियों के अध्ययन या प्रयोग से संबंधित है, जिससे अनिश्चितिता की स्थिति में ठोस एवं विवेकपूर्ण निर्णय लिया जा सके।"

1.5 सांख्यिकी एवं अर्थशास्त्र के मध्य संबंध (Relationship between Statistics and Economics)

प्राचीन काल में सांख्यिकी और अर्थशास्त्र के बीच कोई संबंध नहीं था, क्योंकि न तो अर्थशास्त्र, न ही सांख्यिकी विषय इतना अधिक विकसित हुआ था। लेकिन सांख्यिकीय आँकड़ों का प्रयोग राजनीतिक दृष्टिकोण से अवश्य किया जाता था। आर्थिक क्रियाओं से संबंधित अध्ययन के विकास के साथ-साथ अर्थशास्त्र में आँकडों के संकलन एवं सांख्यिकी विधियों का प्रयोग बढ़ने लगा। वर्तमान समय में किसी भी आर्थिक आँकडों को व्याख्या करने के लिए संकलन, प्रस्तुतीकरण एवं उनके विश्लेषण में सांख्यिकी विधियों का अधिक प्रयोग किया जाता है। आज अर्थशास्त्र के स्पष्टीकरण के लिए सांख्यिकी का प्रयोग अनिवार्य हो गया है। इसलिए कहा गया है कि, 'आज बिना सांख्यिकी की सहायता के अर्थशास्त्र का ज्ञान अधूरा है।'

प्रो. मार्शल (Prof. Marshall) सांख्यिकी के महत्त्व को बताते हुए कहते है कि ***"आँकड़े वे तिनके है जिनसे प्रत्येक अर्थशास्त्री की भाँति मुझे भी ईंट बनानी पडती है।(Statistics are the straw out of which I, like every other economist, have to make the bricks.)"*** यानि आर्थिक नियम रूपी ईंटें तैयार करने के लिए सांख्यिकी रूपी तिनके आवश्यक है। आजकल अर्थशास्त्र के लगभग सभी क्षेत्रों में सांख्यिकी का प्रयोग अधिकाधिक बढ़ता जा रहा है। जैसे- आर्थिक नीतियों का प्रभाव जाँचने के लिए सांख्यिकी ही उपयुक्त साधन है।

अर्थशास्त्र के अध्ययन की आगमन प्रणाली, उत्पादन की दर, जनसंख्या घनत्व, प्रतिव्यक्ति आय, उपयोगिता ह्रास नियम, **माल्थस (Malthus)** का जनसंख्या सिद्धान्त आदि सभी नियमों की पुष्टि के लिए तथा इनके स्पष्टीकरण के लिए सांख्यिकी का प्रयोग अनिवार्य है। इसलिए सिद्धान्त तथा व्यवहार दोनों पक्षों के लिए अर्थशास्त्री को सांख्यिकी की सहायता लेना आवश्यक है।

अर्थशास्त्र में सांख्यिकी का प्रयोग सर्वप्रथम **ग्रेगरी किंग (Gregory King)** ने सत्रहवीं शताब्दी में किया। इन्होंने वस्तुओं की कीमत एवं पूर्ति के बीच संबंध का पता लगाने के लिए सांख्यिकी का प्रयोग किया था। इसके पश्चात 19 वीं शताब्दी में **जेवन्स (Jevons), लेस्ली (Lesly)** आदि अर्थशास्त्रियों ने तथा 20 वीं शताब्दी में **मार्शल (Marshall), कीन्स (Keynes), पेरेटो (Pareto)** आदि ने सांख्यिकी का उपयोग किया। प्रसिद्ध अर्थशास्त्री **जेवन्स (Jevons)** ने सांख्यिकी की अर्थशास्त्र में

उपयोगिता बताते हुए कहा है कि ***"मैं यह नही जानता कि हम कब पूर्ण सांख्यिकी व्यवस्था को प्राप्त कर सकेंगें परन्तु उसकी कमी ही अर्थशास्त्र को पूर्ण विज्ञान बनाने में एक मात्र अजेय बाधा है।"***

प्रसिद्ध संख्याशास्त्री **टिप्पेट (Tippett)** अर्थशास्त्र में सांख्यिकी के महत्त्व को बताते हुए कहते हैं कि ***"एक दिन ऐसा हो सकता है कि विश्वविद्यालयों के अर्थशास्त्र विभाग कोरे सिद्धान्तवादियों के प्रभुत्व में न रहकर सांख्यिकी प्रयोगशालाओं के प्रभाव में उसी प्रकार आ जायेगें जिस प्रकार भौतिक तथा रसायनशास्त्र विभाग प्रयोगशालाओं के प्रभाव में है।"*** अर्थशास्त्र में माल्थस का जनसंख्या सिद्धान्त अर्थशास्त्र में सांख्यिकी के प्रयोग का एक नवीनतम उदाहरण है। माल्थस ने इंग्लैंड की बुरी आर्थिक दशा का अध्ययन किया और बताया कि खाद्यान की वृद्धि अंकगणितीय दर से एवं जनसंख्या वृद्धि रेखागणितीय दर से होती है। इसके अतिरिक्त माँग का नियम, आर्थिक विकास के विभिन्न सिद्धांत आदि अर्थशास्त्र में सांख्यिकी के बढ़ते प्रयोग के उदाहरण हैं।

अर्थशास्त्र के अंतर्गत ही एक नए विषय अर्थमिति (Econometrics) का प्रारंभ हुआ जिसके अंतर्गत सांख्यिकी तथा गणित के प्रयोग से आर्थिक नियमों की जाँच की जाती है। इस विषय का मुख्य उद्देश्य यह है कि सांख्यिकी को अर्थशास्त्र के क्षेत्र में अधिक व्यावहारिक एवं उपयोगी बनाया जाना चाहिए। आर्थिक क्रियाओं का मापन व पूर्वानुमान करने में इस विषय के अंतर्गत प्रतिरूपों (Models), समीकरणों (Equations) एवं फलनों (Functions) का उपयोग किया जाता है।

1.6 सांख्यिकी का महत्त्व (Importance of Statistics)

1. **लोक प्रशासन (Public Administration) :** लोक प्रशासन को सुचारू रूप से चलाने के लिए प्राचीन काल से ही सांख्यिकी का उपयोग किया जाता रहा है। सरकारी बजट प्रचलित वर्ष और आगामी वर्ष के अनुमानों के आधार पर ही बनाया जाता है। जनसंख्या, राष्ट्रीय आय, उत्पादन, आयात-निर्यात आदि के आँकडों आधार पर ही यह तय किया जाता है कि शिक्षा, प्रतिरक्षा (Defence), स्वास्थ्य, परिवहन आदि पर कितना व्यय किया जाए। इसके अतिरिक्त करों के निर्धारण, सरकारी विभागों एवं मंत्रालयों में अपव्यय को कम करने के लिए भी विभिन्न आँकडों का प्रयोग किया जाता है।
2. **आर्थिक नियोजन (Economic Planning) :** किसी भी देश के आर्थिक नियोजन या नीति निर्धारण में वहाँ उपलब्ध संसाधनों, स्थानीय समस्याओं एवं आवश्यकताओं से संबंधित सांख्यिकीय आँकडों की आवश्यकता होती है। आँकड़ों के आधार पर ही अल्पकालीन एवं दीर्घकालीन आवश्यकताओं एवं समस्याओं को ध्यान में रखकर अर्थव्यवस्था में लक्ष्य तय किए जाते है एवं लक्ष्य प्राप्ति हेतु उपर्युक्त वित्तीय संसाधनों के अनुमान लगाए जाते है। योजनाओं की प्रगति का मूल्यांकन करने में भी सांख्यिकीय विधियों का ही प्रयोग किया जाता है।
3. **अर्थशास्त्र (Economics) :** अर्थशास्त्र में आर्थिक समस्याओं का अध्ययन करने में सांख्यिकी महत्वपूर्ण भूमिका निभाती है। अर्थव्यवस्था के प्रत्येक क्षेत्र के आँकड़ों की तुलना, व्याख्या एवं अंतर्संबंध को ज्ञात करने में सांख्यिकी ज्ञान की आवश्यकता होती है। उदाहरण के लिए उपभोग के आँकडों के आधार पर लोगों के जीवन स्तर, उपभोग प्रवृत्ति और बचत आदि के बारे ने अनुमान लगाना संभव हो पाता है। व्यावहारिक अर्थशास्त्र के क्षेत्र में भी बेरोजगारी, मुद्रा-स्फीति जैसी अनेक आर्थिक समस्याओं के अध्ययन तथा समाधान में आँकडों के संकलन एवं विश्लेषण की आवश्यकता होती है।
4. **अर्थमिति (Econometric) :** सांख्यिकी एवं गणितीय विधियों का उपयोग करते हुए अर्थशास्त्र के विभिन्न सिद्धांतों के विश्लेषण अर्थमिति के अंतर्गत किया जाता है। यह बात जानने योग्य है कि सांख्यिकी द्वारा अर्थमिति के लिए मात्रात्मक आँकड़े उपलब्ध करवाए जाते है लेकिन सांख्यिकी स्वयं आर्थिक परिणामों के संबंधों का मापन नहीं कर सकती।

4\. **अनुसंधान कार्यों में (In Research Work) :** वर्तमान समय में प्रत्येक क्षेत्र के अनुसंधान कार्य में सांख्यिकी का उपयोग परम आवश्यक हो गया है। सांख्यिकी का उपयोग अनुसंधानकर्ता द्वारा आँकडों के प्रस्तुतीकरण एवं विश्लेषण आदि में किया जाता है जिससे उन्हें अपने अध्ययन के लिए सही निष्कर्ष निकलने में मदद मिलती है। जैसे जनांकिकी में जनसंख्या से संबंधित अनुसंधान में, वाणिज्य में वस्तुओं की खरीद-बिक्री या आय-व्यय से संबंधित अनुसंधान में इत्यादि। सांख्यिकी के बिना अनुसंधान अधूरा एवं अविश्वसनीय हो सकता है क्योंकि यह आँकडों को समझने एवं उससे निष्कर्ष निकालने की प्रक्रिया को प्रभावी बनाता है।

5\. **व्यापार और वाणिज्य में (In Trade and Commerce) :** व्यापार एवं वाणिज्य में सांख्यिकी का प्रयोग उत्कृष्ट निर्णय लेने में महत्वपूर्ण भूमिका निभाता है। यह व्यवसायिक संगठनों को बाजार के विभिन्न पहलुओं को समझने एवं उनका विश्लेषण करने में मदद करता है। जैसे उत्पादों की माँग एवं पूर्ति, कीमत स्तर एवं बाजार की प्रवृत्तियों, व्यापार चक्रों, उपभोक्ता की अभिरूचि आदि को समझने एवं इनका

विश्लेषण करने हेतु। सांख्यिकी का उपयोग कर निकाले गए निष्कर्षो द्वारा उत्पाद एवं सेवा संबंधी विपणन (marketing) रणनीति बनाई जाती है जिससे व्यापारिक संगठन अपने निर्धारित लक्ष्यों को प्राप्त करने में सफल होते है। यदि व्यापारी का अनुमान गलत हो गया तो उनको हानि हो सकती है। इसके लिए व्यापार एवं वाणिज्य में आँकडों के आधार पर सही अनुमान का होना आवश्यक होता है जोकि सांख्यिकी की मदद से ही संभव है।

1.7 सांख्यिकी की सीमाएँ (Limitations of Statistics)

हालांकि वर्तमान समय में सांख्यिकी का बहुत अधिक महत्त्व है परन्तु इसकी भी कुछ सीमाएँ है। सांख्यिकी की कुछ प्रमुख सीमाएँ निम्नलिखित है-

1. **केवल मात्रात्मक आँकडों का अध्ययन (Study of quantitative data only)**
   सांख्यिकी में केवल उन्हीं आँकडों का अध्ययन किया जाता है जिनको मात्रात्मक रूप में व्यक्त किया जा सकता है। उदाहरण के लिए किसी व्यक्ति की आयु, ऊंचाई, वेतन, परिवार के सदस्यों की संख्या आदि। सांख्यिकी में गुणात्मक आँकडों का अध्ययन संभव नहीं है हालाँकि गुणात्मक आँकडों को भी संख्यात्मक रूप में प्रकट करके अध्ययन करना संभव है। उदाहरण के लिए, गरीबी को संख्यात्मक रूप में मापा नहीं जा सकता परन्तु लोगों की आय एवं जीवन-स्तर का अध्ययन करके गरीबी का अनुमान लगाया जा सकता है। इसी प्रकार किसी सेमेस्टर में नामांकित शिक्षार्थियों का बौद्धिक स्तर परीक्षा में प्राप्त अंकों के आधार पर ज्ञात किया जा सकता है।
2. **अध्ययन हेतु एकमात्र विधि नहीं है। (It is not the only method for study.)**
   सभी देश में अलग अलग प्रकार की समस्याएँ व्याप्त होती है जिसके आँकड़े मात्रात्मक रूप में उपलब्ध नहीं होते या जिनकी प्रवृति गुणात्मक होती है। जैसे किसी देश में प्रचलित रीति-रिवाज गुणात्मक प्रवृति है। इस प्रकार के गुणात्मक आँकडों का प्रयोग सांख्यिकी द्वारा संभव नहीं है। अतः किसी समस्या के अध्ययन में सांख्यिकी ही एकमात्र विधि नहीं है। **क्रॉक्सटन एवं कॉउडेन (Croxton and Cowden)** के अनुसार, ***"यह नहीं मान लेना चाहिए कि सांख्यिकी विधि ही अनुसंधान कार्य में प्रयोग की जाने वाली एकमात्र विधि है, न ही इस विधि को प्रत्येक प्रकार की समस्या का सर्वोत्तम हल समझना चाहिए।"***
3. **सामूहिक अध्ययन को अधिक महत्व (More importance to group study)**
   सांख्यिकी में सामूहिक अध्ययन के द्वारा प्राप्त मात्रात्मक आँकडों को अधिक महत्त्व दिया जाता है। उदाहरण के लिए, किसी देश की प्रति व्यक्ति आय ज्ञात करने के लिए उस देश के सभी व्यक्तियों की आय को जोड़ कर व्यक्तियों की संख्या से भाग दिया जाता है परंतु यह किसी व्यक्ति (गरीब या अमीर) की व्यक्तिगत आय को ना बताकर समूह की औसत आय को प्रदर्शित करता है। अतः सांख्यिकी निष्कर्ष में व्यक्तिगत विशेषताओं को कोई महत्त्व नहीं दिया जाता है।
4. **सांख्यिकीय नियम केवल दीर्घकाल एवं औसत रूप में ही सही होते है। (Statistical laws are true only in the long run and on average.)**
   विज्ञान के नियम सभी परिस्थितियों में समान रहते है। जैसे गुरुत्वाकर्ष के नियम के अनुसार प्रत्येक वस्तु जिसे ऊपर की ओर उछाला जाता है सदैव पृथ्वी की ओर नीचे गिरती है। यह नियम सम्पूर्ण पृथ्वी पर समान रूप से लागू होता है परंतु सांख्यिकी के नियम अन्य विज्ञानों की तरह सभी परिस्थितियों में समान रूप से लागू नहीं होते है। सांख्यिकी के नियम केवल दीर्घकाल एवं औसत रूप में समूहों पर ही सही होते है। उदाहरण के लिए जब हम सिक्के को एक बार उछालते है तो चित (head) और पट (trail) आने की प्रायिकता आधी (½) होती है। जब हम सिक्के को 20 बार उछालते है तो हो सकता है 15 बार चित आए और 5 बार पट आए परंतु जब हम उसी सिक्के को 10,000 बार उछालने पर चित और पट आने की संभावना लगभग आधी (½) होगी। टॉस की संख्या और बढ़ाने पर संभावना आधी (½) के निकट होती जाती है।
5. **विशिष्ट ज्ञान की आवश्यकता (Specific knowledge required)**
   सांख्यिकी में इकट्ठा किए गए आँकडों से उचित परिणाम निकालने के लिए उस विषय का विशिष्ट ज्ञान एवं अनुभव होना आवश्यक होता है। एक अनुभवी व्यक्ति द्वारा ही आँकडों का सही तरीके से इकट्ठा करने के बाद उसका विश्लेषण कर उससे सही निष्कर्ष निकाल पाना संभव है। यदि ऐसा ना हो तो भ्रामक या यूँ कहे कि गलत परिणाम निकलने की अधिक संभावना हो जाती है।

1.8 सांख्यिकी की आधारभूत अवधारणाएँ (Basic Concepts of Statistics)

सांख्यिकी की प्रमुख आधारभूत अवधारणाएँ निम्न है-

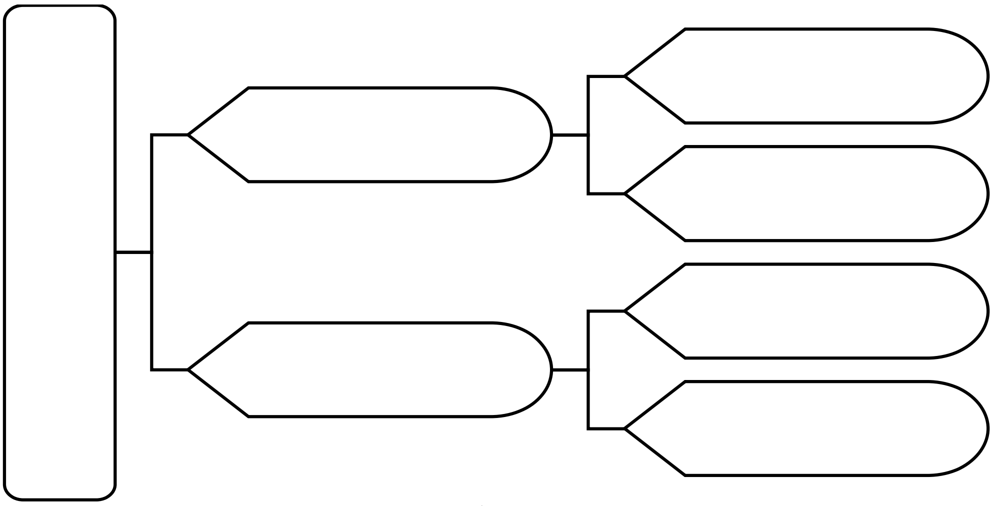
1.8.1 समग्र (Universe or Population)

अनुसंधान के लिए निर्धारित क्षेत्र या सभी इकाईयाँ जिनके बारे में जानकारी एकत्रित करनी होती है, **समग्र** (Universe or population) कहलाती है।

हैमबर्ग (Hamburg) के अनुसार, ***"एक सांख्यिकीय अनुसंधान के अंतर्गत आने वाले पदों अथवा तत्वों के सम्पूर्ण समूह को समग्र कहते है। (A universe of population consists of the total collection of items or element that fall within the slope of statistical investigation.)"***

सिम्पसन एवं काफ्का (Simpson and Kafka) के अनुसार, **"समग्र से आशय किसी अनुसन्धान क्षेत्र की सभी इकाइयों के समूह को समग्र कहते हैं। (A universe or population may be defined as an aggregate of items possessing a common trait or traits.)"**

समग्र को दो आधारों पर विभाजित किया जा सकता है-
1. संख्या निर्धारण के आधार पर 2. आस्तित्व के आधार पर



1. संख्या निर्धारण के आधार पर समग्र दो प्रकार के होते है:

अ. सीमित समग्र (Finite Population): ऐसे समग्र जिनकी इकाइयों की संख्या निश्चित होती है एवं इनकी गणना कर पाना संभव होता है, सीमित समग्र कहलाते है। उदाहरण के लिए, उत्तराखण्ड मुक्त विश्वविद्यालय में पंजीकृत शिक्षार्थियों की संख्या।

ब. असीमित समग्र (Infinite Population): ऐसे समग्र जिनकी कुल इकाइयों की संख्या अनिश्चित होती है अथवा इनकी गणना कर पाना संभव नहीं होता, असीमित समग्र कहलाते है। उदाहरण के लिए, आकाश में तारों की संख्या, किसी पेड़ पर उपलब्ध पत्तों की संख्या आदि।

2. आस्तित्व (Existence) के आधार पर भी समग्र दो प्रकार के होते है:

अ. वास्तविक समग्र (Real or Existant Population): ऐसे समग्र जिनकी सभी इकाइयाँ ठोस या मूर्त रूप में विद्यमान होती है, वास्तविक समग्र कहलाती है। उदाहरण के लिए, किसी सेमेस्टर में नामांकित शिक्षार्थियों की संख्या।

ब. काल्पनिक या सैद्धांतिक समग्र (Hypothetical/Theoretical Population): ऐसे समग्र जो वास्तव में या मूर्त रूप में विद्यमान नहीं होते है एवं जिनकी इकाइयों की केवल कल्पना की जा सकती है। उदाहरण के लिए, पासा फेंकने पर 1 से 6 तक किसी भी संख्या का आना।

1.8.2 न्यादर्श या प्रतिदर्श या नमूना (Sample)

जब किसी समग्र में से कुछ प्रतिनिधि इकाइयों का चयन किया जाता है जो सम्पूर्ण समग्र का प्रतिनिधित्व करती है, ऐसी इकाईयों को **न्यादर्श** या **प्रतिदर्श** (Sample) कहा जाता है। उदाहरण के लिए, किसी विश्वविद्यालय के कुल नामांकित शिक्षार्थियों में से केवल 100 शिक्षार्थियों को चयनित कर अध्ययन किया जाए तो इसे न्यादर्श या प्रतिदर्श कहा जाएगा। इसी प्रकार दैनिक जीवन में जब हम बाजार में कोई सामान खरीदने जाते है तो हम उनके नमूनों को देखकर हम उनकी गुणवत्ता के बारे में पता लगा सकते है।

गिल्बर्ट (Gilbert) के अनुसार, ***"एक न्यादर्श समग्र का प्रतिनिधित्व करने के लिए वैज्ञानिक ढंग से चुना गया एक अपेक्षाकृत छोटा समूह होता है। (A sample is a relatively small group scientifically chosen so as to represent the population.)"***

सिम्पसन एवं काफ्का (Simpson and Kafka) के अनुसार, ***"न्यादर्श अथवा प्रतिदर्श समग्र का वह भाग है, जिसे हम अनुसंधान के उद्देश्य से चुनते हैं। (A sample is that part of the universe which which we select for the purpose of investigation.)"***

सांख्यिकी में निदर्शन (sampling) तकनीकों के बारे में आप विस्तारपूर्वक अगले अध्याय में पढ़ेंगे।

1.8.3 प्राचल (Parameter)

समग्र की सभी इकाइयों के आधार पर निकाले गए सांख्यिकीय माप (जैसे- माध्य, बहुलक, सहसंबंध आदि) को **प्राचल** (parameter) कहा जाता है। उदाहरण के लिए, माना किसी शहर की जनसंख्या 1,00,000 है। इन सभी व्यक्तियों के आँकडों के आधार पर निकाले गए सांख्यिकीय माप को प्राचल कहा जाएगा क्योंकि इनसे सम्पूर्ण जनसंख्या के आँकडों को शामिल किया गया है।

1.8.4 प्रतिदर्शज (Statistics)

समग्र में से चुने गए प्रतिदर्श की इकाइयों के आँकडों के आधार पर निकाले गए सांख्यिकीय माप को **प्रतिदर्शज** (Statistics) कहा जाता है। उदाहरण के लिए, माना किसी शहर की जनसंख्या 1,00,000 है। यदि हम इनमें से 10,000 व्यक्तियों के आँकडों के आधार पर सांख्यिकीय माप को निकाले तो इन्हें प्रतिदर्शज कहा जाएगा।

प्राचल एवं प्रतिदर्शज में अंतर दर्शाने के लिए प्राचल में ग्रीक एवं बड़े अक्षरों (Capital letters) का प्रयोग किया जाता है वहीं प्रतिदर्शज में छोटे रोमन अक्षरों (lower case Roman letters) का प्रयोग किया जाता है।

प्राचल (Parameter)		**प्रतिदर्शज (Statistics)**	
समग्र का आकार (Population size)	N	न्यादर्श का आकार (sample size)	n
समग्र का माध्य (Population mean)	μ (म्यू)	न्यादर्श का माध्य (Sample mean)	(एक्स बार)
समग्र का प्रमाप विचलन (S.D. of population)	σ (सिग्मा)	न्यादर्श का प्रमाप विचलन (S.D. of sample)	S

1.8.5 वर्णनात्मक सांख्यिकी (Descriptive Statistics)

सांख्यिकी एक ऐसा विषय है जो संख्याओं और आँकड़ों से संबंधित है और कई तथ्यों और स्थितियों का वर्णन करता है। वर्णनात्मक सांख्यिकी की अवधारणा विभिन्न तथ्यों के विवरण की इसी विशेषता पर आधारित है और विभिन्न लेखकों द्वारा इसे इस प्रकार परिभाषित किया गया है:-

रोनाल्ड ई. वालपोल (Ronald E. Walpole) के अनुसार, ***"वर्णनात्मक सांख्यिकी में आँकड़ों के एक समूह को इकट्ठा करने और उसका वर्णन करने से संबंधित वे विधियाँ हैं जिनसे सार्थक जानकारी प्राप्त की जा सके। (Descriptive statistics comprises those methods concerned with collecting and describing a set of data so as to yield meaningful information.)"***

एम. सी. शुक्ला (M. C. Shukla) के अनुसार, ***"वर्णनात्मक आँकड़ों में वे सांख्यिकीय विधियाँ शामिल होती हैं जो हमें बताती हैं कि आँकड़ों के एक समूह की विशेषताओं का वर्णन कैसे किया जाए। (Descriptive statistics consist of those statistical methods which tell us how to describe the characteristics of a body of data.)"***

वर्णनात्मक सांख्यिकी उन विधियों से संबंधित है जिनका उपयोग संख्यात्मक तथ्यों की मूल विशेषताओं का वर्णन करने के उद्देश्य से किया जाता है। उनकी बुनियादी विशेषताओं का वर्णन करने के उद्देश्य में संख्यात्मक तथ्यों का संपादन, वर्गीकरण, सारणीकरण, आरेखीय या चित्रमय प्रस्तुति, केंद्रीय प्रवृत्ति के माप, विस्तार, सहसंबंध आदि सम्मिलित होते हैं।

वर्णनात्मक आँकड़ों को भी दो भागों में विभाजित किया जाता है: एकचर (univariate) और द्विचर (bivariate)। एकचर (univariate) वर्णनात्मक आँकड़ो में आवृत्ति वितरण (frequency distributions), केंद्रीय प्रवृत्ति के माप (measures of central tendency) , अपकिरण की माप (measures of dispersion) आदि शामिल हैं, जबकि सहसंबंध (correlation), प्रतिगमन (regression), गुण-साहचर्य (association of attributes) आदि, द्विचर वर्णनात्मक आँकड़ों में शामिल हैं।

1.8.6 अनुमानात्मक सांख्यिकी (Inferential Statistics)

सांख्यिकी के अध्ययन के आधुनिक दृष्टिकोण में 'अनुमानात्मक सांख्यिकी' (Inferential Statistics) पर विशेष जोर दिया गया है, जिसे 'आगमनात्मक सांख्यिकी' (Inductive Statistics) भी कहा जा सकता है। इसकी कुछ परिभाषाएँ इस प्रकार हैं-

रोनाल्ड ई. वालपोल (Ronald E. Walpole) के अनुसार, ***"सांख्यिकीय अनुमान में वे विधियाँ शामिल होती हैं जो आँकड़ों के उप-समूचय/समूह के विश्लेषण से संबंधित होती हैं, जिससे आँकड़ों के पूरे समूह के बारे में भविष्यवाणियाँ या अनुमान लगाया जाता है। (Statistical inference comprises those methods which are concerned with the analysis of a sub-set of data leading to predictions or inferences about the entire set of data.)"***

आर. पी. हुडा (R. P. Hooda) के अनुसार, ***"अनुमानात्मक आँकड़ों को उन तरीकों के एक समूह के रूप में परिभाषित किया जा सकता है जो न्यादर्श के प्रतिदर्शज के आधार पर समग्र के प्राचल के बारे में सामान्यीकरण करने में सक्षम है। (Inferential statistics may be defined as the body of methods capable of making generalisations about the population parameters on the basis of sample statistics.)"***

संक्षेप में उन सभी विधियों को अनुमानात्मक सांख्यिकी में रखा जाता है जो न्यादर्श के प्रतिदर्शज के आधार पर समग्र के प्राचल के बारे में सामान्यीकरण करने में मदद करते हैं। इसमें न्यादर्श द्वारा प्राप्त प्रतिदर्शज के विश्लेषण के आधार पर समग्र से प्राप्त प्राचल के बारे में उचित और तर्कसंगत निष्कर्ष निकालने, अनुमान लगाने और पूर्वानुमान लगाने और निर्णय लेने के विभिन्न तरीकों की प्रक्रिया शामिल है। इन विधियों में संभाव्यता सिद्धांत (probability theory), नमूना परीक्षण (sampling tests), सामान्य वक्र मॉडल (normal curve models), परिकल्पना परीक्षण (hypothesis testing) टी-वितरण (t-distribution) आदि भी शामिल हैं।

1.9 अभ्यास प्रश्न (Practice Questions)

बहुविकल्पीय प्रश्न:

1. 'सांख्यिकी' शब्द का प्रयोग किस रूप में किया जाता है?
   a.एकवचन में b. बहुवचन में
   c. एकवचन एवं बहुवचन दोनों में d. सभी विकल्प गलत है।
2. "सांख्यिकी गणना का विज्ञान है।" यह कथन निम्न में से किस विद्वान ने दिया है?
   a. वैबस्टर b. बाउले c. मार्शल d. सैलिगमैन
3. वर्णनात्मक सांख्यिकी में निम्नलिखित में से कौन सी विधि शामिल होती हैं?
   a. आवृत्ति वितरण b. केंद्रीय प्रवृत्ति के माप c. गुण-साहचर्य d. उपरोक्त सभी
4. अनुमानात्मक सांख्यिकी में निम्नलिखित में से कौन सी विधि शामिल होती हैं?
   a. संभाव्यता सिद्धांत b. परिकल्पना परीक्षण c. टी-वितरण d. उपरोक्त सभी
5. निम्न विकल्पों में से सही कथन का चयन कीजिए:

कथन 1: जब किसी समग्र में से कुछ प्रतिनिधि इकाइयों का चयन किया जाता है जो सम्पूर्ण समग्र का प्रतिनिधित्व करती है, ऐसी इकाइयों को समग्र कहा जाता है।

कथन 2: जब किसी समग्र में से कुछ प्रतिनिधि इकाइयाँ का चयन किया जाता है जो सम्पूर्ण समग्र का प्रतिनिधित्व करती है, ऐसी इकाइयों को न्यादर्श या प्रतिदर्श कहा जाता है।

विकल्प:

a. केवल कथन 1 सही है।
b. केवल कथन 2 सही है।
c. 1 और 2 दोनों सही कथन है।
d. 1 और 2 दोनों गलत कथन है।

1.10 सारांश (Summary)

सांख्यिकी का शाब्दिक अर्थ होता है- संख्याओं से संबंधित शास्त्र/ विज्ञान। सांख्यिकी मुख्य रूप से मात्रात्मक (quantitative) आँकड़ों के संकलन, विश्लेषण और प्रस्तुतीकरण से संबंधित है। व्याकरण के अनुसार सांख्यिकी शब्द का प्रयोग मुख्य रूप से दो प्रकार से प्रयोग में लाया जाता है – बहुवचन एवं एकवचन में। **बहुवचन** में इसका अभिप्राय किसी निश्चित उद्देश्य से एकत्रित मात्रात्मक आँकडों से होता है जबकि एकवचन में इसका अर्थ '**सांख्यिकीय विज्ञान**' **(Statistical Science)** से होता है जिसमें आँकड़ों के संकलन, प्रस्तुतीकरण, वर्गीकरण आदि सांख्यिकीय विधियों (Statistical Methods) का अध्ययन किया जाता है। विभिन्न परिभाषाओं के निष्कर्ष के रूप में यह कहा जा सकता है कि ***"सांख्यिकी विज्ञान और कला दोनों है, जो ज्ञान या अनुसंधान के किसी भी क्षेत्र में निश्चित उद्देश्य के लिए मात्रात्मक आँकडों के संकलन, प्रस्तुतीकरण, विश्लेषण और पूर्वानुमान में प्रयुक्त सिद्धान्तों एवं विधियों के अध्ययन या प्रयोग से संबंधित है, जिससे अनिश्चितिता की स्थिति में ठोस एवं विवेकपूर्ण निर्णय लिया जा सके।"*** सांख्यिकी लोक प्रशासन, आर्थिक नियोजन, अर्थशास्त्र, अर्थमिति एवं अनुसंधान कार्यों के अलावा व्यापार एवं वाणिज्य में भी महत्वपूर्ण भूमिका निभाती है परंतु सांख्यिकी की अपनी कुछ सीमाएँ भी होती है इसमें केवल मात्रात्मक आँकडों का अध्ययन किया जाता है। सांख्यिकी में संकलित आँकडों से उचित परिणाम निकालने हेतु विशेष ज्ञान एवं अनुभव का होना आवश्यक होता है।

1.11 शब्दावली (Glossary)

- **समग्र (Population):** अनुसंधान के लिए निर्धारित क्षेत्र या सभी इकाइयाँ जिनके बारे में जानकारी एकत्रित करनी होती है, समग्र (population) कहलाती है।
- **न्यादर्श या प्रतिदर्श या नमूना (Sample):** जब किसी समग्र में से कुछ प्रतिनिधि इकाइयों का चयन किया जाता है जो सम्पूर्ण समग्र का प्रतिनिधित्व करती है, ऐसी इकाईयों को न्यादर्श या प्रतिदर्श (Sample) कहा जाता है।
- **प्राचल (Parameter):** समग्र की सभी इकाइयों के आधार पर निकाले गए सांख्यिकीय माप (जैसे- माध्य, बहुलक, सहसंबंध आदि) को प्राचल कहा जाता है।
- **प्रतिदर्शज (Statistics):** समग्र में से चुने गए प्रतिदर्श की इकाइयों के आँकडों के आधार पर निकाले गए सांख्यिकीय माप को प्रतिदर्शज कहा जाता है।
- **वर्णनात्मक सांख्यिकी (Descriptive Statistics):** उन विधियों से संबंधित है जिनका उपयोग संख्यात्मक तथ्यों की मूल विशेषताओं का वर्णन करने के उद्देश्य से किया जाता है। उनकी बुनियादी विशेषताओं का वर्णन करने के उद्देश्य में संख्यात्मक तथ्यों का संपादन, वर्गीकरण, सारणीकरण, आरेखीय या चित्रमय प्रस्तुति, केंद्रीय प्रवृत्ति के माप, विस्तार, सहसंबंध आदि सम्मिलित होते हैं।
- **अनुमानात्मक सांख्यिकी (Inferential Statistics):** उन सभी विधियों को अनुमानात्मक सांख्यिकी में रखा जाता है जो न्यादर्श के प्रतिदर्शज के आधार पर समग्र के प्राचल के बारे में सामान्यीकरण करने में मदद करते हैं। इसमें न्यादर्श द्वारा प्राप्त प्रतिदर्शज के विश्लेषण के आधार पर समग्र से प्राप्त प्राचल के बारे में उचित और तर्कसंगत निष्कर्ष निकालने, अनुमान लगाने और पूर्वानुमान लगाने और निर्णय लेने के विभिन्न तरीकों की प्रक्रिया शामिल है।

1.12 अभ्यास प्रश्नों के उत्तर (Answers of Practice Questions)

बहुविकल्पीय प्रश्नों के उत्तर:

1. c 2. b 3.d 4. d 5.b

1.13 सन्दर्भ ग्रन्थ सूची (Reference/Bibliography)

- Singh, S.P. (2010) ***Principles of Statistics***, S. Chand and Company Pvt. Ltd. New Delhi.
- Bose, D., (2003), ***An Introduction to Mathematical Economics***, Himalaya Publishing House, New Delhi.
- Richard J. Larsen and Morris L. Marx, ***An Introduction to Mathematical Statistics and its Applications***, Prentice Hall, 2011.
- K. Sydsaeter and P. Hammond, ***Mathematics for Economic Analysis***, Pearson Educational Asia: Delhi, 2002.
- Allen, R.G.D. (1974), ***Mathematical Analysis for Economists***, Macmillan Press and ELBS, London.
- Chiang, A.C. (1986), ***Fundamental Methods of Mathematical Economics***, McGraw Hill, New York.
- Gupta, S.C. (1993), ***Fundamentals of Applied Statistics***, S. Chand & Sons, New Delhi.
- Monga, G.S. (1972), ***Mathematics and Statistics for Economists***, Vikas Publishing House, New Delhi.
- Ramendu, Roy (2017), ***Principles of Statistics,*** Prayag Pustak Bhawan, Allahabad.
- Bhardwaj, R.S. (2000), ***Mathematics for Economics and Business***, Excel Books.
- Gupta, Jain, Aggrawal (2006), ***Basic Quantitative Methods for Economics***, Navyug Sahitya Sadan, Agra.
- Kumar, Anil (2008), ***Statistical Research Methodology***, Alpha Publishing House.

1.14 सहायक / उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

- एस.एन. लाल, एल.के. चतुर्वेदी (2010),***परिमाणात्मक विश्लेषण***, शिव पब्लिशिंग हाउस, इलाहाबाद।
- नागर, कैलाश नाथ (2008), ***सांख्यिकी के मूलतत्व***, मीनाक्षी प्रकाशन, मेरठ।
- सिंह, एस. पी. (2010), ***सांख्यिकी सिद्धांत एवं व्यवहार***, एस. चन्द एण्ड कंपनी लिमिटेड, दिल्ली।

1.15 निबंधात्मक प्रश्न (Essay type question)

1. सांख्यिकी से आप क्या समझते है? एकवचन में सांख्यिकी से क्या आशय है?
2. सांख्यिकी की सीमाओं का वर्णन कीजिए?
3. सांख्यिकी की आधारभूत अवधारणाओं से आप क्या समझते है?
4. समग्र से आप क्या समझते है? इसके प्रमुख प्रकार कौन कौन से है?
5. वर्णनात्मक एवं अनुमानात्मक सांख्यिकी से आप क्या समझते है?

इकाई 2 आँकडों के संकलन की विधियाँ
(Methods of Data Collection)

2.1 प्रस्तावना (Introduction)

पिछले अध्याय में आपने सांख्यिकी का अर्थ, उसकी विभिन्न परिभाषाएँ, सीमाएँ तथा कुछ प्रमुख आधारभूत अवधारणाओं जैसे- समग्र, न्यादर्श, प्राचल एवं प्रतिदर्शज को जाना। इसी क्रम को आगे बढ़ाते हुए इस अध्याय में आप व्याप्ति और स्रोत के आधार पर आँकडों के संकलन की विधियों के बारे में विस्तारपूर्वक पढ़ेंगे। इसके अतिरिक्त आप प्राथमिक आँकडों को संकलित करने की पाँच विधियों के बारे में विस्तारपूर्वक पढ़ेंगे। आँकडों के संकलन के प्रकाशित और अप्रकाशित प्रकार के द्वितीयक आँकडों के बारे में विस्तार से अध्ययन करेंगे।

2.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप

- ✓ आँकडों के संकलन की विधियों का अर्थ एवं प्रकारों को समझ सकेंगे।
- ✓ समग्र विधि एवं न्यादर्श विधि में अंतर करने में सक्षम हो सकेंगे।
- ✓ व्याप्ति और स्रोत के आधार पर आँकडों के संकलन के प्रकारों से परिचित हो सकेंगे।
- ✓ प्राथमिक आँकडों को संकलित करने की विधियों से अवगत हो सकेंगे।
- ✓ प्रकाशित और अप्रकाशित प्रकार के द्वितीयक आँकडों के बारे में अवगत हो सकेंगे।

2.3 आँकडों के संकलन की विधियों का अर्थ एवं प्रकार (Meaning and Types of Methods of Data Collection)

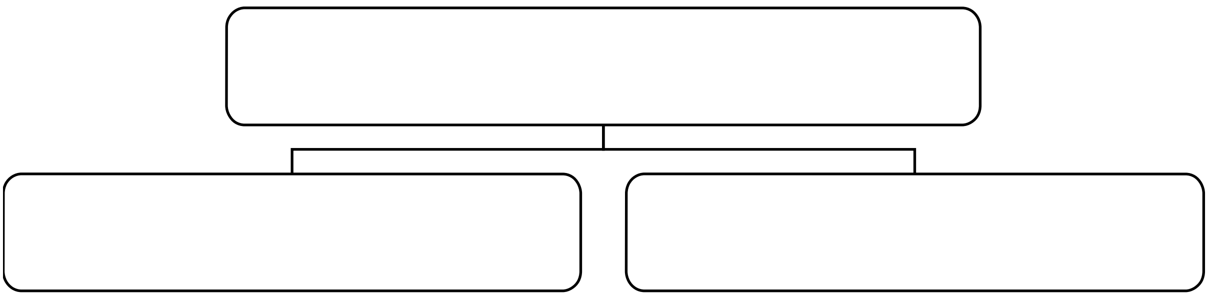
आँकडों के संकलन का आशय विभिन्न प्रकार की विधियों का उपयोग करके मात्रात्मक रूप में अर्थपूर्ण जानकारियों को इकट्ठा करने से हैं। अनुसन्धान में आँकड़ों के संकलन की विधि का चुनाव इस बात पर निर्भर करता है कि आपके अनुसन्धान की विषय-सामग्री में कितनी संवेदनशीलता एवं सटीकता की आवश्यकता हैं। अत: आँकड़ों का संकलन पूरी कुशलता, योग्यता एवं शुद्धता से करना चाहिए। यदि इसमें कोई त्रुटि होती है तो यह अनुसन्धान के निष्कर्षों को प्रभावित कर त्रुटिपूर्ण कर देता हैं।

आँकडों का संकलन किसी भी प्रकार से क्यों ना किया जाए इसमें कमी रह जाना स्वाभाविक है। इसी कमी को दूर करने के लिए जो क्रिया अपनाई जाती है उसे **सम्पादन** (Editing)कहते हैं। सम्पादन की प्रक्रिया में आँकडों का क्रमबद्ध आयोजन, जाँच तथा संशोधन, शुद्वता का स्तर आदि निश्चित किये जाते हैं। आँकडों के संकलन के बाद सम्पादन का कार्य किया जाता है।

आँकडों के संकलन की विधियों का अध्ययन मुख्यतः दो प्रकार से किया जा सकता है। पहला है व्याप्ति के आधार पर एवं दूसरा है स्रोत के आधार पर। आईये इनकों विस्तार से जानने का प्रयास करे।

2.4 व्याप्ति के आधार पर आँकडों का संकलन (Data Collection on the basis of Coverage)

अनुसन्धान करते समय आँकडों के संकलन की इस विधि के अंतर्गत अनुसन्धानकर्ता को यह तय करना होता है कि उसके अनुसन्धान में सभी इकाईयों को अनुसन्धान में सम्मिलित करने की आवश्यकता है या फिर वह कुछ इकाईयों को ही शामिल कर निष्कर्ष प्राप्त कर सकता हैं। अध्ययन क्षेत्र की व्याप्ति (coverage) इसका मुख्य आधार है। जिसके आधार पर हम आँकडों का संकलन दो प्रकार से करते है। पहला हैं समग्र विधि और दूसरा हैं न्यादर्श विधि।



2.4.1 समग्र विधि (Census Method)

किसी अनुसन्धान की प्रक्रिया में जब सभी इकाईओं को शामिल करके आँकडों को इकट्ठा करके निष्कर्ष निकाले जाते हैं तो इस विधि को **समग्र विधि** (Census Method) कहते हैं। भारत में प्रत्येक 10

वर्ष में होने वाली जनगणना इसका सबसे प्रसिद्ध उदाहरण है जिसके द्वारा भारत में व्याप्त सभी इकाईयों (प्रत्येक व्यक्ति और परिवार) के बारे में जानकारी इकट्ठा करने के बाद निष्कर्ष निकाले जाते हैं। इस विधि का प्रयोग केवल ऐसे अनुसंधानों में किया जाता है जहाँ अनुसन्धान का क्षेत्र सीमित हो, प्रत्येक इकाई का अध्ययन करना आवश्यक हो एवं जहाँ निष्कर्षों में अधिक की सटीकता की आवश्यकता हो।

क. समग्र विधि के गुण (Merits of Census Method) - समग्र विधि के कुछ प्रमुख लाभ निम्नलिखित हैं-

1. **शुद्धता एवं विश्वसनीयता (Accuracy and Reliability) :** समग्र विधि से प्राप्त आँकड़े अधिक शुद्ध एवं विश्वसनीय होते है क्योंकि इसमें समग्र की प्रत्येक इकाई से आँकडों का इकट्ठा करके अध्ययन एवं विश्लेषण किया जाता है।
2. **विस्तृत सूचनाएँ (Detailed Information) :** इस विधि से प्राप्त आँकडों से समग्र के बारे में विस्तृत सूचनाएँ उपलब्ध हो जाती हैं। जैसे देश में होने वाली जनगणना से व्यक्तियों की कुल संख्या के साथ-साथ उनकी आयु, वैवाहिक स्थिति, व्यवसाय, आय, परिवार में सदस्यों की संख्या आदि की जानकारी भी प्राप्त होती है।
3. **समग्र में विविधता (Diversity in Population) :** यदि अनुसन्धान का क्षेत्र सीमित हो एवं समग्र की इकाइयाँ भिन्न-भिन्न विशेषताओं वाली हो तो समग्र विधि उपयुक्त होती है।

ख. समग्र विधि के दोष (Demerits of Census Methods) - समग्र विधि के कुछ प्रमुख दोष निम्नलिखित हैं-

1. **अधिक खर्चीली पद्धति (More Expensive Method):** सांख्यिकी की यह विधि बहुत खर्चीली होती है। इसमें अनुसन्धानकर्ता को निर्धारित क्षेत्र की सभी इकाइयों से आँकड़े इकट्ठा करने होते है जिसमें धन की अधिक आवश्यकता होती है। इसी वजह से यह विधि अधिकतर संस्थागत या सरकार के स्तर पर प्रयुक्त होती है।
2. **अधिक समय (More Time):** इस विधि से आँकड़े इकट्ठा करने में अधिक समय लगता है क्योंकि समग्र की प्रत्येक इकाई से आँकड़े इकट्ठा किए जाते है। अधिक आँकड़े होने के कारण निष्कर्ष निकालने में भी अधिक समय लगता है।
3. **कुछ परिस्थितिओं में असंभव (Impossible in Some Circumstances) :** यदि अनुसन्धान का क्षेत्र असीमित (Infinite) हो या समग्र की प्रकृति कुछ इस प्रकार की हो कि अनुसन्धान से उनका अस्तित्व ही समाप्त हो जाए तो इस विधि का प्रयोग संभव नहीं है। उदाहरण के लिए, किसी तालाब में जल की मात्रा को मापने हेतु जल को तालाब से बाहर निकालने पर तालाब का अस्तित्व ही समाप्त हो जाएगा।

2.4.2 न्यादर्श विधि (Sample Method)

किसी अनुसन्धान की प्रक्रिया में जब अध्ययन निर्धारित क्षेत्र की सम्पूर्ण इकाइयों में से कुछ प्रतिनिधि इकाइयाँ का चुनाव करके आँकडों को इकट्ठा कर निष्कर्ष निकाले जाते हैं तो इस विधि को **न्यादर्श विधि** (Sample Method) कहते हैं। जैसे- डॉक्टर द्वारा शरीर से खून की कुछ बूँदें निकालकर पूरे रक्त समूह (Blood Group) का पता लगाना।

क. न्यादर्श विधि के गुण (Merits of Sample Method) - न्यादर्श विधि के कुछ प्रमुख लाभ निम्नलिखित है-

1. **कम खर्चीली विधि (Less Expensive Method) :** न्यादर्श विधि में आँकडों के एक हिस्से का ही अध्ययन कर निष्कर्ष निकाले जाते है। अतः इसमें अपेक्षाकृत कम धन खर्च होता है।
2. **समय एवं श्रम की बचत (Saving of Time and Labour) :** इस प्रकार की विधि में सम्पूर्ण समग्र का अध्ययन ना करके केवल प्रतिनिधि इकाइयों का अध्ययन किया जाता है। इसलिए यह विधि समय एवं श्रम दोनों की बचत करती है।
3. **विस्तृत जाँच की सुविधा (Facility of Detailed Investigation) :** न्यादर्श विधि में इकाइयों की संख्या कम होने की वजह से उनकी विस्तृत एवं गहन (Intensive) जाँच संभव है।
4. **विशेष अनुसन्धान में एकमात्र विधि (The Only Method in Special Investigation) :** अनुसन्धान के कुछ क्षेत्र ऐसे होते है जहाँ समग्र विधि से अध्ययन संभव नहीं

होता है। उन परिस्थितियों में केवल न्यादर्श विधि ही उपयुक्त होती है। उदाहरण के लिए, जहाँ समग्र की इकाइयों की प्रकृति नाशवान होने की होती है वहाँ केवल न्यादर्श विधि ही संभव होती है।

5. **सुविधाजनक (Convenience) :** इस विधि में प्रयुक्त आँकडों का आकार छोटा होने के कारण यह विधि अनुसन्धान में सुविधाजनक होती है।

ख. न्यादर्श विधि के दोष (Demerits of Sample Method)

1. **कम सटीकता (Low Accuracy):** न्यादर्श विधि में केवल कुछ चयनित इकाइयों के अध्ययन के आधार पर निकाले गए निष्कर्षों में सटीकता की कमी पाती जाती है।
2. **भ्रामक निष्कर्ष (Misleading Conclusion):** इस विधि में यदि न्यादर्श का चुनाव ठीक प्रकार से नही होगा तो निष्कर्ष भ्रामक प्राप्त होते है। अतः न्यादर्श का चुनाव करने में विशेष सावधानियाँ अपनाने की आवश्यकता होती है।
3. **विशिष्ट ज्ञान की आवश्यकता (Need of Specific Knowledge):** यह एक जटिल विधि है। अतः इसके अध्ययन के लिए एक अनुभवी एवं विशिष्ट ज्ञान वाले व्यक्ति की आवश्यकता होती है।
4. **न्यादर्श लेने की असंभवता (Impossibility of Sampling):** यदि समग्र का आकार बहुत छोटा हो तो उसमें न्यादर्श का चयन करना असंभव होता है।
5. **अनुपयुक्तता (Unsuitablity):** यदि समग्र की इकाइयों में विविधता पाई जाती है तो न्यादर्श विधि अनुपयुक्त होती है।

2.4.3 समग्र विधि एवं न्यादर्श विधि में अंतर (Difference between Census Method and Sample Method)

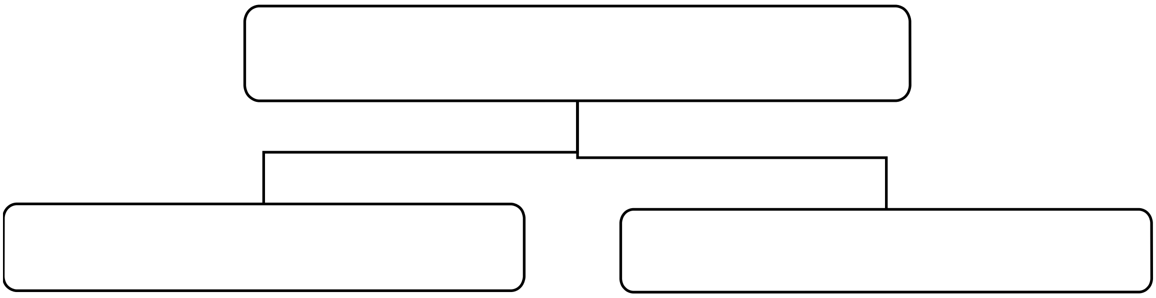
समग्र विधि एवं न्यादर्श विधि के बीच कुछ प्रमुख अंतर निम्नलिखित हैं-

1. **जाँच का स्वरुप (Nature of Enquiry) :** समग्र विधि में समग्र (Universe) की सभी इकाइयों का अध्ययन किया जाता है जबकि न्यादर्श विधि में समग्र (Universe) में से चयनित इकाइयों का अध्ययन किया जाता है।
2. **समय (Time) :** समग्र विधि में निष्कर्ष ज्ञात करने में अधिक समय लगता है जबकि न्यादर्श विधि में निष्कर्ष कम समय में निकाले जा सकते है।
3. **लागत (Cost) :** समग्र विधि में इकाइयों की संख्या अधिक होने के कारण अधिक लागत आती है जबकि न्यादर्श विधि में इकाइयों की संख्या कम होने से लागत कम आती है।
4. **शुद्धता का स्तर (Level of Accuracy) :** समग्र विधि में प्रत्येक इकाई का अध्ययन किया जाता है जिससे शुद्धता का स्तर ऊँचा होता है जबकि न्यादर्श विधि में केवल समग्र से चयनित इकाइयों का अध्ययन होता है जिसकी वजह से शुद्धता का स्तर अपेक्षाकृत कम होता है। साथ ही न्यादर्श विधि में प्राप्त शुद्धता का स्तर भी निदर्शन के चयन की पद्धति पर निर्भर करती है। निदर्शन के चयन की विभिन्न विधियों के बारे में विस्तार से आप 5 वें अध्याय में पढ़ेंगे।
5. **अध्ययन का आधार (Basis of Study) :** समग्र विधि में अध्ययन निश्चितता (Certainty) के आधार पर होता है जबकि न्यादर्श विधि में अध्ययन संभावनाओं (Probability) के आधार पर होता है।
6. **उपयुक्तता (Suitability) :** समग्र विधि वहाँ उपर्युक्त होती है जहाँ अनुसन्धान का क्षेत्र सीमित हो या समग्र की सभी इकाइयों में बारे में जानकारी इकट्ठा करने की संवेदनशीलता एवं सटीकता की आवश्यकता हो। वहीं दूसरी ओर न्यादर्श विधि का प्रयोग वहाँ किया जाता है जहाँ अनुसन्धान का क्षेत्र विस्तृत या व्यापक हो अथवा इकाइयाँ सजातीय (Homogeneous) या नाशवान (mortal) हो।

2.5 स्रोत के आधार पर आँकडों का संकलन (Data Collection on basis of Sources)

अनुसन्धान करते समय आँकडों के संकलन की इस विधि के अंतर्गत अनुसन्धानकर्ता को यह तय करना होता है कि उसके अनुसन्धान में कितनी मात्रा में मौलिकता, सटीकता और प्रमाणिकता को प्राप्त करने की आवश्यकता है और उस विषय पर आँकड़ों की उपलब्धता की वास्तविक स्थिति कैसी हैं। आँकड़ों के संकलन की

स्थिति ही इसका मुख्य आधार है। संकलन की दृष्टि से आँकड़े दो प्रकार के होते है- पहला हैं प्राथमिक आँकड़े और दूसरा हैं द्वितीयक आँकड़े।



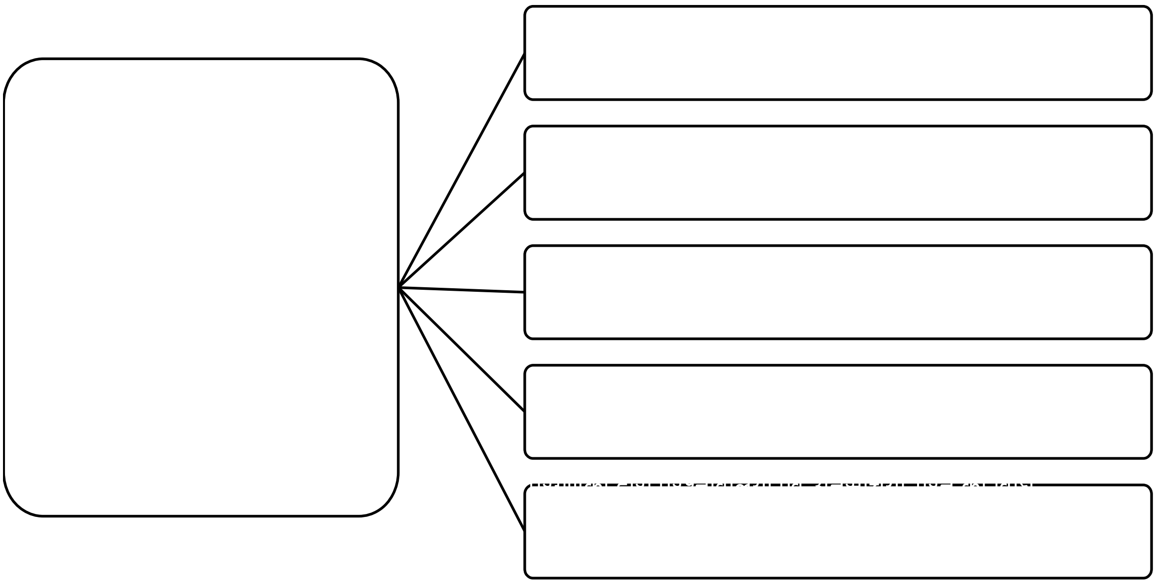
2.5.1 प्राथमिक आँकड़े (Primary Data)

अनुसन्धानकर्ता जिन आँकड़ों को पहली बार मौलिक रूप से संकलित करता है उन आँकड़ों को प्राथमिक आँकड़े कहते है। उदाहरण के लिए, उत्तराखंड के किसानों की आर्थिक स्थिति के बारे में जानने के लिए यदि अनुसन्धानकर्ता किसानों से मिलकर नए सिरे से आँकड़ों को इकट्ठा करता है इन आँकड़ों को प्राथमिक आँकड़े कहा जाता है।

सैक्राइस्ट (Secrist) के शब्दों में, ***"प्राथमिक आँकडों से आशय यह है कि वे मौलिक है, अर्थात जिनका समूहीकरण बहुत ही कम या नहीं हुआ है। घटनाओं की गणना या उनका अंकन उसी प्रकार किया जाता है, जैसा पाया जाता है। (By primary data are meant those which are original, that is, those in which little or no grouping has been made, the instance being recorded or itemised as encountered. They are essentially raw materials.)"***

प्राथमिक आँकडों के संकलन की विधियाँ (Methods of Collection of Primary Data)

प्राथमिक आँकडों को संकलित करने की प्रमुख विधियाँ निम्नलिखित है-



1. प्रत्यक्ष व्यक्तिगत साक्षात्कार विधि (Direct Personal Interview Method)

इस विधि में अनुसन्धानकर्ता स्वयं अनुसन्धान क्षेत्र में जाकर संबंधित व्यक्तियों से प्रत्यक्ष व्यक्तिगत संपर्क करके आवश्यक जानकारी एवं आँकड़ें संकलित करता है। उदाहरण के लिए, यदि अनुसन्धानकर्ता किसी गाँव में व्यक्तियों की आर्थिक स्थिति ज्ञात करने के लिए वह उस गाँव में जाकर

व्यक्तियों से संपर्क करके आँकड़े संकलित करता है तो इसे **प्रत्यक्ष व्यक्तिगत साक्षात्कार विधि** कहा जाता है।

प्रत्यक्ष व्यक्तिगत साक्षात्कार विधि के गुण (Merits of Direct Personal Interview Method) :

क. **शुद्धता (Accuracy) -** इस विधि में अनुसन्धानकर्ता स्वयं आँकडों को संकलित करता है। अत: आँकडों में उच्चस्तर की शुद्धता पाई जाती है।

ख. **सजातीयता (Homogeneity):** इस विधि में एक ही व्यक्ति या अनुसन्धानकर्ता द्वारा आँकड़ें संकलित किए जाने के कारण आँकडों में सजातीयता एवं एकरूपता का गुण पाया जाता है।

ग. **लोचशीलता (Elasticity):** इस विधि में अनुसन्धानकर्ता आवश्यकता पड़ने पर प्रश्नों को बदल सकता है तथा प्रश्नों का अर्थ एवं अर्थ स्पष्ट रूप से समझा सकता है। अत: इस विधि में लचीलेपन का गुण अधिक मात्रा में पाया जाता है।

घ. **प्रमाणिकता (Authenticity):** इस विधि में अनुसन्धानकर्ता स्वयं अनुसंधान क्षेत्र में जाकर सूचना का अवलोकन कर उसकी प्रामाणिकता की जांच कर जानकारी इकट्ठा करता है। जिसके कारण यह विधि अधिक प्रमाणित मानी जाती है।

प्रत्यक्ष व्यक्तिगत अनुसन्धान विधि के दोष (Demerits of Direct Personal Investigation Method):

क. **सीमित क्षेत्र (Limited Area):** बड़े या व्यापक क्षेत्रों में यह विधि उपयुक्त नहीं है क्योंकि इसमें अधिक धन एवं समय लगता है। इस विधि का प्रयोग इसीलिए सीमित क्षेत्र में ही संभव है।

ख. **पूर्वाग्रहो (Prejudices):** प्रत्येक अनुसन्धानकर्ता का अपना दृष्टिकोण, पूर्वाग्रह या मान्यताओं के प्रति झुकाव होता है जिसके कारण परिणामों पर प्रभाव पड़ने की संभावना बनी रहती है। अत: इस विधि में पक्षपात की संभावना पाई जाती है।

ग. **अधिक खर्चीली (More Expensive):** इस विधि के अंतर्गत अनुसन्धानकर्ता को स्वयं अनुसन्धान के क्षेत्र में जाकर आँकड़ें इकट्ठा करने पड़ने है जिसमें अधिक धन की आवश्यकता होती है। अत यह विधि अधिक खर्चीली है।

घ. **भ्रामक निष्कर्ष (Misleading Conclusions):** इस विधि में सीमित क्षेत्र होने के कारण इस बात की संभावना भी बनी रहती है कि इकट्ठा किए गए आँकड़े समग्र का सही प्रतिनिधित्व करेगें या नहीं और यदि आँकड़े समग्र का सही प्रतिनिधित्व नहों करते होगे तो उनसे प्राप्त निष्कर्ष भ्रामक ही होगें।

प्रत्यक्ष व्यक्तिगत साक्षात्कार विधि के प्रयोग में बरती जाने वाली सावधानियाँ (Precautions to be taken while using Direct Personal Investigation Method)

क. अनुसन्धानकर्ता को व्यवहारकुशल, परिश्रमी व धैर्यवान होना चाहिए ताकि वह सूचना देने वालों का विश्वास व सहयोग प्राप्त कर सके।

ख. प्रश्न थोड़े, सरल, स्पष्ट और ऐसे होने चाहिए जिससे उत्तर देने वाले को आहात ना पहुँचे।

ग. अनुसन्धानकर्ता को अपनी व्यक्तिगत भावनाओं एवं पूर्वाग्रहों को यथासंभव दूर रखना चाहिए ताकि इनका प्रभाव अनुसन्धान पर ना पड़े।

2. अप्रत्यक्ष मौखिक अनुसन्धान विधि (Indirect Oral Investigation Method):

इस विधि में अनुसन्धान से संबंधित व्यक्ति से सीधे जानकारी इकट्ठा ना करके उसकी जानकारी रखने वाले व्यक्ति या साथी से मौखिक बातचीत करके आँकड़े संकलित किए जाते है। उदाहरण के लिए, किसी फैक्ट्री में कार्य करने वाले श्रमिकों की आय के बारे में जानकारी संकलित करने के लिए श्रमिकों से ना पूछकर उनके अधिकारियों से पूछा जाए। इसी प्रकार यदि किसी विद्यालय के विद्यार्थियों के बारे में जानकारी संकलित करने के लिए विद्यार्थियों से ना पूछकर अध्यापकों से जानकारी संकलित की जाए तो वह **अप्रत्यक्ष मौखिक अनुसन्धान विधि** कहलाती है।

अप्रत्यक्ष मौखिक अनुसन्धान विधि के गुण (Merits of Indirect Oral Investigation Method)

क. **व्यापक क्षेत्र (Broad Area) :** इस विधि में अनुसन्धान का क्षेत्र व्यापक होता है। जहाँ प्रत्यक्ष व्यक्तिगत अनुसन्धान संभव या लाभकारी नहीं होता है वहाँ इस विधि का प्रयोग किया जा सकता है।

ख. धन एवं समय की बचत (Money and Time Saving) : इस विधि के प्रयोग से धन एवं समय की काफी बचत होती है।

ग. विशेषज्ञों की राय (Experts opinion) - इस पद्धति में विशेषज्ञों की राय एवं सुझावों का लाभ सहजता से प्राप्त हो जाता है।

घ. पूर्वाग्रहों का कम प्रभाव (Less impact of Prejudices) - अन्य विधियों की तुलना में इस विधि में अनुसन्धानकर्ता के व्यक्तिगत पूर्वाग्रह का कम प्रभाव पड़ता है।

ङ. गुप्त जानकारी प्राप्त करने के लिए (To obtain secret information) - इस विधि के अन्तर्गत वे सूचनाएं भी प्राप्त की जा सकती हैं जिन्हें सूचनादाता या तो देने को तैयार नहीं होता या सही जानकारी नहीं देता।

अप्रत्यक्ष मौखिक अनुसन्धान विधि के दोष (Demerits of Indirect Oral Investigation Method)

क. सटीकता की कमी (Lack of Accuracy) - इस विधि में अनुसन्धानकर्ता प्रत्यक्ष रूप से सूचनादाता के संपर्क में नहीं आता है जिसके कारण प्राप्त परिणामों में सटीकता की कमी रह जाती हैं।

ख. सूचनादाता की पक्षपात-भावना (Bias of the Respondent) - जिन व्यक्तियों की सहायता से आँकड़ें इकट्ठा किये जाते हैं उनकी पक्षपात की भवना का प्रभाव अनुसन्धान पर पड़ता है।

ग. सूचनादाता का अरुचिकर व्यवहार (Unpleasant behaviour of the Respondent) - जब सूचना देने वालों की प्रश्न में कोई प्रत्यक्ष व्यक्तिगत हिस्सेदारी या रुचि नहीं होती है तब वे उदासीन व्यवहार करते हैं और ज्यादातर टाल-मटोल कर उत्तर देते हैं। यदि जिन लोगों से जानकारी इकट्ठा की जाती है वे प्रश्नों का उत्तर देने में लापरवाही बरतते हैं तो इससे निष्कर्ष प्रभावित होते हैं और परिणामों की सटीकता कम हो जाती है।

घ. एकरूपता की कमी (Lack of Uniformity) - जब अलग-अलग सूचना दाताओं द्वारा अलग-अलग व्यक्तियों के बारे में जानकारी इकट्ठा की जाती हैं तो प्राप्त किए गए आँकड़ों की एकरूपता समाप्त हो जाती है जिससे सही परिणाम पर पहुंचने में बाधा आती है।

अप्रत्यक्ष मौखिक अनुसन्धान विधि के प्रयोग में बरती जाने वाली सावधानियाँ (Precautions to be taken while using Indirect Oral Investigation Method)

क. जिनकी सहायता से आँकड़े इकट्ठा किए जा रहे हों उनकी बातों पर बिना जाँच किए पूर्ण विश्वास नहीं करना चाहिए।

ख. अनुसंधान में यह पूर्णत: सुनिश्चित किया जाना चाहिए कि सूचनादाता को तथ्यों की पूर्ण जानकारी हो तथा वह सूचना देने में रुचि रखता हो।

ग. इस बात को ध्यान में रखना आवश्यक है कि जिस व्यक्ति की सहायता से सामग्री इकट्ठा की जा रही है वह उस विषय के पक्ष व विपक्ष में पक्षपातपूर्ण धारणाएं नहीं रखता है। यदि ऐसा हुआ तो प्राप्त निष्कर्ष भ्रामक होगें एवं सटीकता की उनमें कमी पाई जाएगी।

3. संवाददाताओं द्वारा आँकड़ें इकट्ठा करना (Method of Collecting data by correspondents)

इस विधि को स्थानीय स्रोतों द्वारा सूचना-प्राप्ति भी कहा जाता है। इस विधि के अन्तर्गत अनुसन्धानकर्ता विभिन्न स्थानों पर स्थानीय लोगों को नियुक्त करता है जो समय-समय पर अपने अनुभवों के आधार पर अपेक्षित सूचनाएं भेजते हैं। उन व्यक्तियों को संवाददाता कहा जाता है।

संवाददाताओं द्वारा आँकड़ें इकट्ठा करने वाली विधि के गुण (Merits of the method of collecting data by correspondents) :

क. विस्तृत क्षेत्र (Broad area) - इस विधि के माध्यम से दूर-दूर तक फैले अनुसन्धान क्षेत्रों से सूचनाएं प्राप्त की जा सकती हैं।

ख. धन एवं समय की बचत (Money and Time Saving) : इस विधि के प्रयोग से धन एवं समय की काफी बचत होती है।

ग. नियमितता (Regularity) - इस विधि द्वारा सूचनाएं नियमित रूप से प्राप्त की जा सकती हैं।

संवाददाताओं द्वारा आँकड़ें इकट्ठा करने वाली विधि के दोष (Demerits of the method of collecting data by correspondents) :

क. मौलिकता का अभाव (Lack of originality) - यदि संवाददाता अपने अनुमान को अधिक महत्व देकर सूचना भेजता है तो उसमें मौलिकता का अभाव होता है।

ख. एकरूपता का अभाव (Lack of Uniformity) - विभिन्न संवाददाताओं द्वारा विभिन्न स्थानों तथा विभिन्न विधियों द्वारा सूचना एकत्र किये जाने के कारण आँकडों में एकरूपता का अभाव रह जाता है।

ग. पूर्वाग्रह (Prejudice) - संवाददाताओं के पक्षपात व पूर्वाग्रहों के कारण निकले गए निष्कर्ष पक्षपातपूर्ण हो सकते हैं।

घ. सूचना मिलने में देरी (Delay in receiving Information) - कभी-कभी सम्वाददाता सूचना भेजने में इतना विलंब कर देते हैं कि उस सूचना का महत्व ही समाप्त हो जाता है।

संवाददाताओं द्वारा आँकड़ें इकट्ठा करने वाली विधि के प्रयोग में बरती जाने वाली सावधानियाँ (Precautions to be taken while using the method of collecting data by correspondents)

क. केवल निष्पक्ष, कुशल एवं अनुभवी तथा व्यक्तिगत धारणाओं की भावना से दूर रहने वाले व्यक्तियों को ही संवाददाता नियुक्त किया जाना चाहिए।

ख. संवाददाता को अनुसन्धान विषय और उसकी गहराई के साथ-साथ संवेदनशीलता से भी परिचित होना चाहिए।

ग. अनुसंधान में अशुद्धियों को दूर करने के लिए क्षतिपूरक त्रुटियों के सिद्धान्त को लागू करते हुए अधिक से अधिक सम्वाददाताओं को नियुक्त किया जाना चाहिए।

घ. सामान्य मानसिकता से कार्य करने वाला संवाददाता को नियुक्त करना चाहिए जोकि ज्यादा आशावादी या निराशावादी नहीं होना चाहिए।

4. संवाददाताओं द्वारा प्रश्नावली एवं साक्षात्कार अनुसूची भरने की विधि (Method of filling Questionnaire and Interview Schedule by correspondents)

इस विधि को 'डाक प्रश्नावली पद्धति' (Mailed Questionnaire Method) के नाम से जाना जाता हैं। इस विधि के अन्तर्गत अनुसन्धानकर्ता सबसे पहले एक प्रश्नावली या अनुसूची तैयार करता है। इसकी एक-एक प्रति डाक द्वारा (यदि सम्भव हो तो व्यक्तिगत रूप से भेजकर भी) सूचनादाता को भेज दी जाती है। यहाँ यह बात ध्यान देने योग्य है कि वर्तमान समय में डाक के स्थान पर इमेल के माध्यम से सूचना प्राप्त करने का चलन व्यापक रूप से उपयोग किया जाता हैं। सामान्यत: प्रश्नावली के साथ एक अनुरोध-पत्र भी लगा होता है जिसमें अनुसन्धान के उद्देश्य तथा महत्त्व का उल्लेख होता है और यह अनुरोध किया जाता है कि निश्चित अवधि में प्रश्नावली भरकर वापस लौटा दी जाए। इसमें आम तौर पर सूचनादाता को यह आश्वासन दिया जाता है कि भेजी गई जानकारी पूरी तरह से गोपनीय रखी जाएगी और इसका उपयोग केवल इस अनुसन्धान के लिए किया जाएगा।

संवाददाताओं द्वारा प्रश्नावली एवं साक्षात्कार अनुसूची भरने की विधि के गुण (Merits of the method of filling Questionnaire and Interview Schedule by correspondents) :

क. विस्तृत क्षेत्र (Broad Area) - एक विस्तृत क्षेत्र में कम खर्चे पर सूचनाएँ प्राप्त करने के लिए यह विधि बहुत उपयुक्त है।

ख. मितव्ययी (Frugal) - इस विधि में प्रश्नावली छपवाने तथा डाक टिकटों में धन व्यय होता है लेकिन जाने-आने में होने वाले भारी व्ययों से बचत हो जाती है। इसीलिए यह प्रणाली मितव्ययी है।

ग. मौलिक एवं सटीकता (Originality and Accuracy) - इस विधि में प्रश्नावली एवं साक्षात्कार अनुसूची सूचकों द्वारा स्वयं भरकर भेजी जाती है इसीलिए प्राप्त की गए आँकड़ें मौलिक एवं सटीक होते है।

संवाददाताओं द्वारा प्रश्नावली एवं साक्षात्कार अनुसूची भरने की विधि के दोष (Demerits of the method of filling Questionnaire and Interview Schedule by correspondents) :

क. प्रश्नावलियों का वापस न आना (Non-return of questionnaires) - इस प्रणाली का सबसे बड़ा दोष यह है कि अधिकांश व्यस्त सूचक प्रश्नावलियों को कोई महत्त्व नहीं देते और वे इन्हें भरकर वापस नहीं भेजते।

ख. अपूर्ण सूचना (Incomplete) - जब सूचनादाता कुछ प्रश्नों को समझ नहीं पाता तथा कुछ प्रश्नों के उत्तर जो व्याख्यात्मक होते हैं उनके उत्तर अस्पष्ट एवं अपूर्ण देता है तो प्रश्नावलियों के आधार पर प्राप्त किए गए आँकड़ें अपूर्ण एवं अशुद्ध होते है।

ग. लोच का अभाव (Lack of Elasticity) - सूचनादाता को प्रश्नावली डाकया इमेल से भेज देने के बाद उसमें कोई संशोधन करना सम्भव नहीं होता। इसी प्रकार, प्रश्नों के उत्तर प्राप्त हो जाने पर उन्हीं के स्पष्टीकरण हेतु यदि कुछ अन्य प्रश्न उठते हैं तो उनका उत्तर पुन: प्राप्त करना लगभग असम्भव होता है।

घ. संदिग्ध निष्कर्ष (Dubious Conclusion) - यदि पर्याप्त मात्रा में प्रश्नावली भरकर नहीं आ पाती तो इन अपर्याप्त प्रश्नावलियों के आधार पर निकाले गये निष्कर्ष अविश्वसनीय हो सकते हैं।

ड. केवल साक्षरों पर लागू (Applicable only on literate) - इस विधि का प्रयोग केवल वहीं हो सकता है जहाँ सूचनादाता साक्षर हों और प्रश्नावली या अनुसूची भर सके।

च. व्यावहारिक कठिनाइयाँ (Practical Difficulties) - व्यवहार में इस विधि द्वारा आँकड़ें इकट्ठा करने में यह कठिनाई आती है कि सूचनादाता अनेक सूचनाओं को मौखिक रूप से बता सकते हैं लेकिन अपने हाथ से लिखकर नहीं देना चाहते हैं जैसे अपनी आय, सम्पत्ति, आदतें।

5. प्रगणकों द्वारा प्रश्नावली एवं साक्षात्कार अनुसूची भरने की विधि (Method of filling Questionnaire and Interview Schedule by Enumerators)

इस विधि में आँकड़ों को इकट्ठा करने के लिए प्रगणकों की नियुक्ति की जाती है और वे प्रगणक एक निश्चित प्रश्नावली के आधार पर आँकड़ों को इकट्ठा करते हैं। भारत में प्रत्येक दस वर्ष के पश्चात् की जाने वाली जनगणना इस विधि का महत्त्वपूर्ण उदाहरण है।

प्रगणको द्वारा प्रश्नावली एवं साक्षात्कार अनुसूची भरने की विधि के गुण (Merits of the method of filling Questionnaire and Interview Schedule by Enumerators)

क. विस्तृत क्षेत्र (Broad Area) - एक वृहत् क्षेत्र में समग्र अथवा प्रतिदर्श अनुसन्धान के लिए यह विधि उपयुक्त है। डाक प्रश्नावली पद्धति के विपरीत इसका प्रयोग अशिक्षित सूचकों के क्षेत्र में भी हो सकता है।

ख. सटीक एवं पूर्ण सूचना (Accurate and complete Information) - इस विधि में सूचनादाताओं के समक्ष प्रश्न प्रस्तुत करने और उसे सही ढंग से समझाने के लिए प्रगणक मौजूद रहता है जिसके कारण सभी सूचनादाताओं से सही एवं पूर्ण जानकारी प्राप्त कर ली जाती है।

ग. एकरूपता(Uniformity) - सभी प्रगणको का प्रशिक्षण एक ही तरीके से किया जाता हैं जिसके कारण उनके द्वारा इकट्ठा किए गए आँकड़ें एवं सूचनाओं में एकरूपता पाई जाती है।

घ. पक्षपात रहित (Unbiased) – यह विधि पक्षपात रहित है क्योकि प्रगणक प्रश्नावली के आधार पर तथा समझाये गये ढंग से ही प्रश्नों के उत्तर प्राप्त करता है। अत: प्राप्त सूचनाओं में प्रगणक अथवा सूचनादाता के पक्षपात की सम्भावना बहुत कम हो जाती है।

प्रगणको द्वारा प्रश्नावली एवं साक्षात्कार अनुसूचियाँ भरने की विधि के दोष (Demerits of the method of filling Questionnaire and Interview Schedule by correspondents) :

क. खर्चीली पद्धति (Expensive Method) - इस विधि में प्रश्नावली तैयार कराने के साथ साथ अनेक प्रगणक नियुक्त करने पड़ते हैं जिसके कारण यह विधि बहुत खर्चीली है। अत: यह विधि केवल छोटे अनुसन्धान या व्यक्तिगत अनुसन्धानों के लिए ही उपयोगी है।

ख. प्रगणकों की समस्या (Problem of Enumerators) - इस विधि में प्रगणकों का चयन करने और उन्हें प्रशिक्षित करने की समस्या आती है। इसके साथ ही प्रगणकों की विचारधारा और पक्षपात का अनुसन्धान के निष्कर्षों पर प्रभाव पड़ सकता है।

प्राथमिक आँकड़ों को इकट्ठा करने के लिए किसी भी विधि का चयन अनुसंधान की प्रकृति, उद्देश्य, शुद्धता का स्तर, उपलब्ध आर्थिक साधन एवं समय के आधार पर ही किया जाता हैं।

2.5.2 द्वितीयक आँकड़ें (Secondary Data)

द्वितीयक आँकड़ों का अर्थ

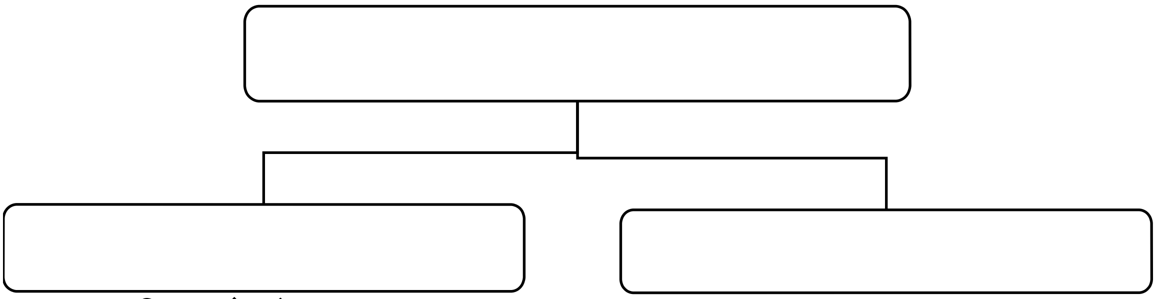
वे आँकड़े जो किसी व्यक्ति या संस्था द्वारा पहले संकलित किए जा चुके है एवं अनुसन्धानकर्ता उनका उपयोग अपने अनुसन्धान हेतु करता है उन आँकड़ों को द्वितीयक आँकड़ें कहते है। अथार्त द्वितीयक आँकड़ें वे आँकड़े हैं जो अनुसन्धानकर्ता के द्वारा उपयोग तो किए जाते हैं परन्तु इकट्ठा नहीं किए जाते। उदाहरण के लिए, उत्तराखंड के कृषकों की आर्थिक स्थिति के बारे में जानने के लिए यदि अनुसन्धानकर्ता सरकार या किसी संस्था द्वारा प्रकाशित आँकडों का प्रयोग करता है तो ऐसे आँकडों को द्वितीयक आँकड़ें कहा जाता है।

द्विवितीयक आँकड़ों की परिभाषा

एम. एम.ब्लेयर (M.M. Blair) के शब्दों में, ***"द्विवितीयक आँकड़ें वे है, जो पहले से अस्तित्व में हैं और जिन्हें वर्तमान के उत्तर में नहीं, वरन किसी अन्य उद्देश्य के लिए एकत्र किया गया होता है। (Secondary data those already in existence and which have been collected for some other purpose than the answering of the question at hand.)"***

द्विवितीयक आँकड़ों के प्रकार

द्वितीयक आँकड़ों के मुख्यतः प्रकाशित एवं अप्रकाशित स्रोत के रूप में विभाजित किया जाता हैं।



क. प्रकाशित आँकड़े –

सरकारी एवं गैर-सरकारी संगठन एवं अन्य अनुसन्धानकर्ता विभिन्न विषयों एवं अध्ययन के आयामों पर महत्वपूर्ण आँकड़ें इकट्ठा करते हैं और उन्हें समय-समय पर प्रकाशित करते हैं। प्रकाशित आँकड़ों के प्रमुख स्रोत निम्नलिखित हैं -

1. **अन्तर्राष्ट्रीय प्रकाशन (International Publications) -** संयुक्त राष्ट्र संघ (UNO), अन्तर्राष्ट्रीय श्रम संघ (ILO), अन्तर्राष्ट्रीय मुद्रा कोष (IMF), संयुक्त राष्ट्र विकास कार्यक्रम (UNDP) जैसी अन्तर्राष्ट्रीय संस्थाएं समय-समय पर आँकड़ों का संकलन तथा प्रकाशन करती रहती हैं। उदाहरण के रूप में संयुक्त राष्ट्र विकास कार्यक्रम (UNDP) द्वारा प्रत्येक वर्ष मानव विकास के आँकडें संकलित कर मानव विकास सूचकांक की रिपोर्ट प्रकाशित करती हैं।
2. **सरकारी प्रकाशन (Government Publication)** - प्रत्येक देश की सरकारें देश की आर्थिक, सामाजिक, राजनैतिक एवं अन्य प्रकार की दशाओं की जानकारी प्राप्त करने के लिए आँकड़ें इकट्ठा कर उनको प्रकाशित करती हैं। ये आँकडें बहुत विश्वसनीय और सटीक होते हैं। आजकल भारत में लगभग सभी मन्त्रालयों द्वारा अनेक प्रकार की सूचनाएं व रिपोर्ट प्रकाशित किए जाते हैं। जैसे भारतीय रिजर्व बैंक बुलेटिन (Reserve Bank of India Bulletin), आर्थिक सर्वेक्षण (Economic Survey) और कृषि सांख्यिकी एक नज़र में (Agricultural Statistics at a Glance)।
3. **अर्ध-सरकारी संस्थाओं के प्रकाशन (Semi-Government Publicatios)** - नगरपालिकाएं, नगर निगम, जिला बोर्ड आदि विभिन्न प्रकार के आँकड़ें इकट्ठा कराकर प्रकाशित करवाते हैं। जैसे-जन्म-मरण, स्वास्थ्य, शिक्षा से सम्बन्धित आँकड़ें।
4. **आयोग व समितियो की रिपोर्ट (Reports of Commissions and Committees)** - देश की विभिन्न समस्याओं के अध्ययन के लिए सरकार या किसी अन्य संस्था द्वारा आयोग या समितियां नियुक्त की जाती रहती हैं। ये आयोग या समितियां समस्याओं के अध्ययन उपरान्त सम्बन्धित आँकड़ें संकलित करके अपना प्रतिवेदन प्रस्तुत करती हैं। जैसे महिलाओं के मुद्दों से संबंधित विभिन्न पहलुओं पर राष्ट्रीय महिला आयोग (NCW) अनुसंधानों को प्रायोजित कर उन्हें प्रकाशित करता है।
5. **वाणिज्यिक व वित्तीय संस्थाओं के प्रकाशन (Publications of Commercial and Financial Institions) -** देश की वाणिज्यिक व वित्तीय संस्थाओं द्वारा विभिन्न प्रकार के

आँकड़ें इकट्ठा कर करके प्रकाशित किया जाता हैं। जैस-भारतीय उद्योग व वाणिज्य संघ, जूट मिल्स एसोसिएशन, हिन्दुस्तान लिवर लिमिटेड, बिड़ला ग्रुप, टाटा एण्ड सन्स लिमिटेड, रिलायंस इंडस्ट्रीज लिमिटेड नामक व्यापार परिषदें साथ ही स्कन्ध विपणियां, उपज विपणियां भी अनेक प्रकार के आँकड़ें इकट्ठा करके प्रकाशित करती हैं।

6. **विश्वविद्यालय एवं शोध संस्थाओं के प्रकाशन** (Publications of University and Research Institutions) - विश्वविद्यालय एवं शोध संस्थाओं द्वारा अनेक प्रकार के आँकड़ें इकट्ठा किये जाते हैं और प्रकाशित किये जाते हैं। जैसे भारत के विभिन्न विश्वविद्यालयों के शोध ग्रंथों को हम शोध गंगा की आधिकारिक वेबसाइट पर पढ़ सकते हैं।
7. **समाचार-पत्र एवं पत्रिकाएं (News papers and Magazines) -** बहुत से पत्र एवं पत्रिकाएं अनेक प्रकार के आंकड़े एकत्र करके प्रकाशित करती हैं। जैसेः इकॉनॉमिक टाइम्स एवं इकॉनोमिक एंड पॉलिटिकल वीकली।

ख. अप्रकाशित आँकड़े -

अप्रकाशित रूप से भी द्वितीयक आँकड़े उपलब्ध हो जाते हैं। अनेक व्यापारिक संघों एवं अनुसंधानकर्ताओं द्वारा विभिन्न निजी उद्देश्यों की पूर्ति के लिए आँकड़े इकट्ठा किए जाते हैं जिनका प्रकाशन नहीं किया जाता।

द्वितीयक आँकड़ो की संवीक्षा और प्रयोग (Scrutiny and Use of Secondary Data)

द्वितीयक आँकड़ों का प्रयोग करने से पूर्व आलोचनात्मक समीक्षा द्वारा उनका विस्तृत सम्पादन कर लेना बहुत जरुरी है क्योकि द्वितीयक आँकड़ों में कई प्रकार की कमियाँ पाई जाती हैं। अतः उसका उपयोग सावधानीपूर्वक करना चाहिए। **कौनर (Connor)** के अनुसार, ***"आँकड़ें, विशेष रूप से अन्य व्यक्तियों द्वारा एकत्रित आँकड़ें, प्रयोगकर्ता के लिए अनेक त्रुटियों से पूर्ण होते हैं। (Statistics, especially other people's statistics, are full of pitfalls for the users.)"*** ये कमियाँ अनेक कारणों से हो सकती हैं जैसे सांख्यिकीय इकाई में परिवर्तन, सूचना की अपर्याप्तता व अपूर्णता, पक्षपात, उद्देश्य व क्षेत्र की भिन्नता आदि। अतः अनुसन्धानकर्ता को द्वितीयक आँकड़ों को प्रयोग करने से पहले यह भली-भाँति जाँच लेना चाहिए कि प्रस्तुत द्वितीयक सामग्री में विश्वसनीयता (reliability), अनुकूलता (suitability) तथा पर्याप्तता (adequacy) आदि आवश्यक गुण पर्याप्त मात्रा में उपलब्ध हैं या नहीं।

सावधानियाँ -

द्वितीयक सामग्री की विश्वसनीयता, उपयुक्तता व पर्याप्तता की जाँच करने के लिए निम्न बातों का ध्यान रखना चाहिए-

1. **पूर्व अनुसन्धानकर्ता की योग्यता (Qualification of previous researcher) -** सबसे पहले, यह देखा जाना चाहिए कि द्वितीयक आँकड़े मुख्य रूप से किस अनुसन्धानकर्ता द्वारा इकट्ठा किए गए थे। यदि उसकी योग्यता, ईमानदारी, अनुभव और निष्पक्षता संतोषजनक है तो उन द्वितीयक आँकड़ों का उपयोग किया जा सकता है।
2. **सांख्यिकीय इकाइयों का अर्थ (Meaning of Statistical Units)** - अनुसन्धानकर्ता को यह भी देख लेना चाहिए कि पूर्व-अनुसन्धान में प्रयुक्त सांख्यिकीय इकाइयों का क्या अर्थ है और क्या वे उसके अनुसन्धान में प्रयुक्त सांख्यिकीय इकाइयों की परिभाषा से मेल खाते हैं? और यदि वह इसका प्रयोग अपने अनुसंधान में करेगा तो क्या वह सही निष्कर्ष निकाल पाएगा या नहीं?
3. **उद्देश्य व क्षेत्र (Objective and Scope)** - यह भी ध्यान दिया जाना चाहिए कि जब प्रस्तुत आँकड़ें प्राथमिक रूप से इकट्ठा किया गया था तो अनुसंधान के उद्देश्य और क्षेत्र ज्ञात नहीं थे जिनके लिए अब उन्हें द्वितीयक आँकड़ें के रूप में उपयोग किया जाना है। यदि उद्देश्य और दायरे में अंतर है तो आँकड़ें अनुपयुक्त और अविश्वसनीय होगे।
4. **सटीकता का स्तर (Level of Accuracy) -** इस बात पर भी विचार करना आवश्यक है कि प्रस्तुत आँकड़ों में किस स्तर की सटीकता रखी गई और उसे प्राप्त करने में किस हद तक सफलता प्राप्त हुई। आँकड़ों में जितनी अधिक सटीकता होगी, वह उतना ही अधिक विश्वसनीय होगा। साथ ही यह भी देख लेना चाहिए कि आँकड़ों में अत्यधिक सन्निकटन (approximation) तो नहीं किया गया है। आँकड़ों में जितनी कम मात्रा में सन्निकटन होता है उनमें सटीकता का स्तर उतनी ही अधिक होता है।

5. **प्रायोगिक जाँच (Experimental Investigation) -** अनुसन्धानकर्ता को प्रस्तुत आँकड़ों में से कुछ की प्रायोगिक जाँच करके यह देख लेना चाहिए कि वे आँकड़ें विश्वसनीय हैं अथवा नहीं।
6. **जाँच का समय और उसकी परिस्थितियाँ (Time and circumstances of Investigation) -** अनुसंधानकर्ताओं को यह भी निश्चय कर लेना चाहिए कि उपलब्ध सामग्री किस समय से सम्बन्धित है तथा किन परिस्थितियों में इकट्ठा की गई थी। आँकड़ों के प्रारम्भिक संकलन और उनके उपयोग के समय की परिस्थितियों में अन्तर होने के कारण उनकी उपयोगिता कम हो सकती है। जैसे अंतर्राष्ट्रीय दशा (युद्ध, आपातकाल), लोगों के रहन-सहन व रीति-रिवाज में होने वाले परिवर्तन, मूल्यों में अन्तर आदि को ध्यान में रखकर ही प्रकाशित आँकड़ों का प्रयोग करना चाहिए।
7. **संकलन की विधि (Method of Collection) -** पूर्व-अनुसन्धान में संकलन की जो विधि पहले अपनाई गई थी वह आँकड़ों के वर्तमान प्रयोग के लिए कहाँ तक उपयुक्त और विश्वसनीय है? यदि प्रतिदर्श विधि का प्रयोग करके अनुसन्धान किया गया हो तो यह निश्चित कर लेना चाहिए कि प्रतिदर्श यथेष्ट (enough) है और पूर्ण रूप से समग्र का प्रतिनिधित्व करता है अथवा नहीं। इन सब बातों के बारे में सन्तुष्ट हो जाने पर ही द्वितीयक आँकड़ों का प्रयोग करना चाहिए।
8. **तुलना (Comparison)** - यदि एक ही विषय पर कई स्रोतों से द्वितीयक आँकड़े प्राप्त होते हैं तो उनकी प्रमाणिकता की जाँच करने के लिए उनकी तुलना की जानी चाहिए। यदि इनमें कोई महत्वपूर्ण अंतर है तो सबसे विश्वसनीय स्रोत से प्राप्त आँकड़ों को ही स्वीकार करना चाहिए या नए सिरे से अनुसंधान का कार्य शुरू करना चाहिए।

द्वितीयक आँकड़ों का प्रयोग करने से पहले अनुसंधानकर्ताओं को उन आँकड़ों की आलोचनात्मक जाँच उपर्युक्त बातों का ध्यान रखकर करनी चाहिए। यदि परीक्षण के बाद द्वितीयक आँकड़ें विश्वसनीय, उपयुक्त व यथेष्ट प्रतीत हो तभी उसका प्रयोग वर्तमान अनुसन्धान के लिए करना चाहिए। बिना जाँचे द्वितीयक आँकड़ों का उपयोग करना पूर्णत: अनुचित है। डॉ. बाउले ने अपनी पुस्तक An elementary Manual of Statistics में लिखा है, ***"प्रकाशित आंकड़ों को ऊपर से ही देखकर उनके बाह्य मूल्य पर ग्रहण कर लेना कभी सुरक्षित नहीं है जब तक उनका अर्थ व उनकी सीमाएँ अच्छी तरह ज्ञात न हो जाएँ; और यह सदैव आवश्यक है कि उन तर्कों की आलोचनात्मक समीक्षा की जाए जो उन पर आधारित हैं। (It is never safe to take published statistics at their face value, without knowing their meaning and limitations, and it is always necessary to criticize arguments that can be based on them.)"***

2.6 अभ्यास प्रश्न (Practice Questions)

बहुविकल्पीय प्रश्न:

1. व्याप्ति के आधार पर आँकडों को निम्नलिखित में से किस विधि के द्वारा इकट्ठा किया जाता है?
   a. समग्र विधि b. न्यादर्श विधि
   c. समग्र एवं न्यादर्श विधि d. सभी विकल्प ग़लत है
2. स्रोत के आधार पर निम्न में से किस प्रकार के आँकड़ें इकट्ठा किया जाता है?
   a. प्राथमिक b. द्वितीयक
   c. प्राथमिक एवं द्वितीयक d. सभी विकल्प ग़लत है
3. प्राथमिक आँकड़ें, द्वितीयक आँकड़ों की तुलना में कैसे होते हैं?
   a. कम विश्वसनीय b. अधिक विश्वसनीय
   c. समान रूप से विश्वसनीय d. सभी विकल्प ग़लत है

रिक्त स्थान भरिए

4. जब सूचनादाता .... (साक्षर/निराक्षर) हो तब प्रश्नावली विधि का प्रयोग नहीं किया जा सकता।
5. किसी भी बैंक की वार्षिक प्रतिवेदन से प्राप्त आँकड़ों को ....(प्राथमिक/द्वितीयक) आँकड़ें कहा जाता हैं।

2.7 सारांश (Summary)

इस अध्याय में आपने आँकड़ों के संकलन का अर्थ एवं इनके विभिन्न प्रकारों को जाना। आँकडों को मुख्यत: दो आधारों पर वर्गीकृत किया जाता है- व्याप्ति और स्रोत के आधार पर। जहाँ व्याप्ति के आधार पर आँकडों को समग्र एवं न्यादर्श विधि द्वारा वहीँ स्रोत के आधार पर आँकडों को प्राथमिक एवं द्वितीयक विधि द्वारा

इकट्ठा किया जाता है। इन सभी विधियों में कुछ गुण और दोष पाए जाते हैं। लेकिन अनुसंधान में सभी विधियाँ अपना महत्व रखती हैं। अनुसंधान में आँकड़ों को इकट्ठा करने के लिए किसी भी विधि का चयन उसकी प्रकृति, उद्देश्य, शुद्धता का स्तर, उपलब्ध आर्थिक साधन एवं समय के आधार पर ही करना चाहिए।

2.8 शब्दावली (Glossary)

- **समग्र विधि (Census Method)** - अनुसन्धान हेतु जब सभी इकाईओं को सम्मिलित करके आँकडों को इकट्ठा कर निष्कर्ष प्राप्त किए जाते हैं तो इसे समग्र विधि कहा जाता हैं।
- **न्यादर्श विधि (Sample Method)** - अनुसन्धान हेतु जब अध्ययन निर्धारित क्षेत्र की सम्पूर्ण इकाइयों में से कुछ प्रतिनिधि इकाइयों का चुनाव करके आँकडों को इकट्ठा कर निष्कर्ष प्राप्त किए जाते हैं तो से इस विधि को न्यादर्श विधि कहा जाता हैं।
- **प्राथमिक आँकड़े (Primary Data)** - अनुसन्धानकर्ता जिन आँकड़ों को पहली बार मौलिक रूप से संकलित करता है उन आँकड़ों को प्राथमिक आँकड़े कहते है।
- **द्वितीयक आँकड़ें (Secondary Data)** - वे आँकड़े जो किसी व्यक्ति या संस्था द्वारा पहले संकलित किए जा चुके है एवं अनुसन्धानकर्ता उनका उपयोग अपने अनुसन्धान हेतु करता है उन आँकड़ों को द्वितीयक आँकड़ें कहते है।
- **प्रतिनिधि इकाइयाँ (Representative Units)** - न्यादर्श विधि के अंतर्गत जब अध्ययन क्षेत्र की सभी इकाइयों में से कुछ प्रतिनिधि इकाइयों का चयन किया जाता है जोकि सम्पूर्ण समग्र का प्रतिनिधित्व करते हैं। इस दृष्टि से केवल वे इकाइयाँ जो समग्र में व्याप्त अधिकतम विशेषताओं का प्रतिनिधित्व करती हैं प्रतिनिधि इकाइयाँ कहलाती हैं ।
- **सूचनादाता (Respondent) -** किसी बात या विषय के बारे में जानकारी या सूचना देने वाले व्यक्ति को सूचनादाता कहा जाता है।
- **संवाददाता (Correspondent)** - अनुसन्धानकर्ता विभिन्न स्थानों पर स्थानीय लोगों को नियुक्त करता है जो समय-समय पर अपने अनुभवों के आधार पर अपेक्षित सूचनाएं भेजते हैं। उन व्यक्तियों को संवाददाता कहा जाता है।

2.9 अभ्यास प्रश्नों के उत्तर (Answers of Practice Questions)

बहुविकल्पीय प्रश्नों के उत्तर:

1. c 2. c 3.b

रिक्त स्थान भरिए

4. साक्षर 5. द्वितीयक

2.10 सन्दर्भ ग्रन्थ सूची (Reference/Bibliography)

- Singh, S.P. (2010) ***Principles of Statistics***, S. Chand and Company Pvt. Ltd. New Delhi.
- Bose, D., (2003), ***An Introduction to Mathematical Economics***, Himalaya Publishing House, New Delhi.
- Richard J. Larsen and Morris L. Marx, ***An Introduction to Mathematical Statistics and its Applications***, Prentice Hall, 2011.
- K. Sydsaeter and P. Hammond, ***Mathematics for Economic Analysis***, Pearson Educational Asia: Delhi, 2002.
- Allen, R.G.D. (1974), ***Mathematical Analysis for Economists***, Macmillan Press and ELBS, London.
- Chiang, A.C. (1986), ***Fundamental Methods of Mathematical Economics***, McGraw Hill, New York.
- Gupta, S.C. (1993), ***Fundamentals of Applied Statistics***, S. Chand & Sons, New Delhi.
- Monga, G.S. (1972), ***Mathematics and Statistics for Economists***, Vikas Publishing House, New Delhi.

- Ramendu, Roy (2017), ***Principles of Statistics,*** Prayag Pustak Bhawan, Allahabad.
- Bhardwaj, R.S. (2000), ***Mathematics for Economics and Business***, Excel Books.
- Gupta, Jain, Aggrawal (2006), ***Basic Quantitative Methods for Economics***, Navyug Sahitya Sadan, Agra.
- Kumar, Anil (2008), ***Statistical Research Methodology***, Alpha Publishing House.

2.11 सहायक / उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

- एस.एन. लाल, एल.के. चतुर्वेदी (2010),***परिमाणात्मक विश्लेषण***, शिव पब्लिशिंग हाउस, इलाहाबाद।
- नागर, कैलाश नाथ (2008), ***सांख्यिकी के मूलतत्व***, मीनाक्षी प्रकाशन, मेरठ।
- सिंह, एस. पी. (2010), ***सांख्यिकी सिद्धांत एवं व्यवहार***, एस. चन्द एण्ड कंपनी लिमिटेड, दिल्ली।

2.12 निबंधात्मक प्रश्न (Essay type question)

1. सांख्यिकीय सामग्री के संग्रहण में प्रयुक्त विभिन्न रीतियों को समझाइए। इनमें से आप किसको ठीक समझते हैं।
2. प्राथमिक आँकड़ें इकट्ठा करने की प्रत्यक्ष व्यक्तिगत अनुसन्धान तथा अप्रत्यक्ष मौखिक अनुसन्धान विधियों के गुण-दोष बतलाइए।
3. द्वितीयक समंक क्या होते हैं? इनके प्रमुख स्रोत बताइए। द्वितीयक समंकों को अनुसंधान में प्रयोग करते समय क्या सावधानियां रखी जानी चाहिए?
4. व्याप्ति के आधार पर समग्र विधि एवं न्यादर्श विधि का अर्थ, गुण एवं दोष लिखिए। दोनों के बीच अंतर भी स्पष्ट कीजिए?

इकाई 3 आँकडों का वर्गीकरण एवं सारणीयन
(Classification and Tabulation of Data)

3.1 प्रस्तावना (Introduction)

पिछले अध्याय में आपने आँकडों के संकलन की विधियों के बारे में जान चुके है। इस इकाई में आप आँकडों के वर्गीकरण का अर्थ एवं परिभाषा, इसके उद्देश्य को जानेंगे। इसके अलावा आप इसके आँकडों के वर्गीकरण कार्य, महत्त्व तथा आदर्श वर्गीकरण के तत्वों को समझ सकेंगे। साथ ही आप आँकडों के सारणीयन के अर्थ, उद्देश्य एवं आवश्यक तत्वों के बारे में पढेंगे तथा सारणी के वर्गीकरण के मुख्य आधारों को समझ सकेंगे।

संकलित आँकड़ें अत्यधिक जटिल, अस्पष्ट एवं अव्यवस्थित रूप में होते है। उन्हें ऊपरी नज़र से देखने पर समझना एवं कोई निष्कर्ष निकालना कठिन होता है। अतः एकत्रित आँकडों को सरल एवं समझने योग्य बनाने हेतु किसी निश्चित एवं तार्किक आधार पर व्यवस्थित करना वर्गीकरण एवं सारणीयन कहलाता है।

ए.आर.एलेर्सिक (A.R. Ilersic) के अनुसार, ***"सांख्यिकीविद् का पहला कार्य विवरणों को इस तरह से कम करना और सरल बनाना है कि मुख्य विशेषताएँ स्पष्ट रूप से दृष्टिगोचर हो सकें, साथ ही एकत्रित आँकडों की व्याख्या की सुविधा भी मिल सके। इस प्रक्रिया को डेटा को वर्गीकृत और सारणीबद्ध करने के रूप में जाना जाता है। (The statistician's first task is to reduce and simplify the details into such a form that the salient features may be brought out, while still facilitating the interpretation of the assembled data. This procedure is known as classifying and tabulating the data.)"***

3.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के बाद आप

- ✓ आँकड़ों के वर्गीकरण का अर्थ, उसकी परिभाषा, उद्देश्य आदि से परिचित हो सकेंगे।
- ✓ आँकड़ों के वर्गीकरण के कार्य, महत्व एवं सीमाओं एवं को समझ सकेंगे।
- ✓ आदर्श वर्गीकरण के आवश्यक तत्व तथा वर्गीकरण की विधियों से अवगत होंगे।
- ✓ सारणीयन का अर्थ, उद्देश्य एवं आवश्यक तत्वों को जान सकेंगे।
- ✓ सारणी के वर्गीकरण के मुख्य आधारों को समझने में सक्षम होंगे।

3.3 आँकड़ों के वर्गीकरण का परिचय (Introduction to Classification of Data)

आँकड़ों के वर्गीकरण का अर्थ

संकलित आँकडों को किसी गुण के आधार पर समान व असमान कर अलग-अलग वर्गों में बांटने की प्रक्रिया को वर्गीकरण कहा जाता है। उदाहरण के लिए, उत्तराखंड मुक्त विश्वविद्यालय में पंजीकृत कुल विद्यार्थियों का आयु, लिंग, लम्बाई, धर्म आदि के आधार पर वर्गीकरण किया जा सकता है।

आँकड़ों के वर्गीकरण की परिभाषा

कॉनर (Connor) के शब्दों में, ***''वर्गीकरण आँकडों को (वास्तविक रूप से या काल्पनिक रूप से) समानता तथा सदृश्यता के आधार पर वर्गों या विभागों में क्रमानुसार रखने की क्रिया है और इससे व्यक्तिगत इकाइयों की विविधता में पाई जाने वाली गुणों की एकता व्यक्त हो जाती है। (Classification is the process of arranging things (either actually or notionally) in groups or classes according to their resemblances and affinities, and given expression to the unity of attributes that may subsist amongst a diversity of individuals.)"***

सैक्रिस्ट (Secrist)** के शब्दों में, "वर्गीकरण आँकडों को उनकी सामान्य विशेषताओं के आधार पर क्रम अथवा समूहों में क्रमबद्ध और विभिन्न, लेकिन संबंधित भागों में अलग-अलग करने की प्रक्रिया है। (Classification is the process of arranging data into sequences and groups according to their common characteristics or separating them into different but related parts.)"

जे.आर.हिक्स (J.R. Hicks) के अनुसार, ***"वर्गीकरण एवं क्रमबद्ध तथ्य स्वयं बोलते है, जबकि अव्यवस्थित रूप में वे मांस की तरह मृत होते है। (Classified and***

arranged facts speak themselves, unarranged they are as dead as mutton.)"

आँकडों के वर्गीकरण के मुख्य विशेषताएं (Main Features of Classification of Data)

सांख्यिकीय अनुसंधान के उद्देश्य, क्षेत्र एवं स्वरूप के अनुसार वर्गीकरण के अंतर्गत संकलित आँकडों को विभिन्न वर्गों में बांटा जाता है-

1. वर्गीकरण किसी गुण या विशेषता या माप के आधार पर होता है।
2. वर्गीकरण वास्तविक या काल्पनिक हो सकता है।
3. वर्गीकरण पदों की विभिन्नता के बीच उनकी एकता (Unity in Diversity) को स्पष्ट करता है।

आँकड़ों के वर्गीकरण के उद्देश्य (Objectives of Classification)

आँकडों के वर्गीकरण के प्रमुख उद्देश्य निम्न हैं-

1. **आँकडों को सरल व संक्षिप्त बनाना (Simplify and Concise the Data)-** वर्गीकरण का मुख्य उद्देश्य आँकडों की जटिलता को दूर करके उन्हें सरल तथा संक्षिप्त रूप देना है।
2. **समानता व असमानता में स्पष्टता लाना-** वर्गीकरण से सांख्यिकीय तथ्यों की समानता स्पष्ट रूप से प्रकट होती है इससे उन्हें समझने में सहायता मिलती है।
3. **तर्कपूर्ण एवं वैज्ञानिक व्यवस्था करना (To Make Logical And Scientific Arrangements)-** वर्गीकरण की सहायता से आँकडों की मौलिक विशेषताओं के अनुसार उनको वैज्ञानिक एवं तर्कपूर्ण ढंग से प्रस्तुत करने की व्यवस्था की जाती है। जैसे-जनसंख्या आँकडों को आयु, जाति, धर्म, लिंग आदि वर्गों में व्यक्त करना एक तर्कपूर्ण क्रिया है।
4. **पारस्परिक सम्बन्ध स्पष्ट करन` म` सहायक (Helpful in Clarifying Interpersonal Relationships)-** आँकडों का वर्गीकरण उनके बीच पाये जाने वाले सम्बन्ध के अध्ययन को सम्भव बनाता है। जैसे- चेचक की बीमारी से आँकडों को वर्गीकृत करने से यह ज्ञात किया जा सकता है कि जिन व्यक्तियों को टीका लगाया गया था उनकी अपेक्षा जिन व्यक्तियों को टीका नहीं लगाया गया था, कम चेचक निकली या नहीं।
5. **एक मानसिक चित्र प्रस्तुत करना (To Present a Mental Picture)-** वर्गीकरण से यह सम्भव हो जाता है कि अनुभव तथा विचार का एक मानसिक चित्र खींचा जा सके। इससे मस्तिष्क का बोझ कम हो जाता है तथा आँकडों के ढेर को संक्षिप्त रूप से समझा व याद रखा जा सकता है तथा उन्हें गणितीय विवेचना के योग्य बना लिया जाता है।
6. **अन्य सांख्यिकीय विधियों का आधार प्रस्तुत करना (Laying the Foundation for Other Statistical Methods)-** वर्गीकरण द्वारा सांख्यिकीय सामग्री के सारणीयन तथा विश्लेषण के लिए आधार तैयार किया जाता है। बिना वर्गीकरण के आँकडों का सारणीयन असम्भव है और सारणीयन के अभाव में सांख्यिकीय विश्लेषण अव्यावहारिक है।

आँकड़ों के वर्गीकरण का महत्व एवं सीमाएँ (Importance and Limits of Classification)

संकलित आँकड़े अपने मूलरूप में अव्यवस्थित होते हैं। इन्हें सरल व समझने योग्य बनाने के लिए वर्गीकरण की आवश्यकता होती है। वर्गीकरण के अभाव में समंकों का प्रस्तुतीकरण, विश्लेषण एवं उनकी परस्पर तुलना असम्भव हैं। वर्गीकरण, सारणीयन का आधार है।

वर्गीकरण की एक सीमा आँकडों के संक्षिप्तीकरण की प्रक्रिया में कुछ विवरणों का समाप्त हो जाना है। संक्षिप्तीकरण जितना अधिक होगा उतना ही अधिक विवरणों के छूट जाने की सम्भावना रहेगी। अतः सांख्यिक को आवश्यक विवरण समाप्त होने से बचने के लिए संक्षिप्तीकरण बहुत ही सावधानी से करना चाहिए।

आदर्श वर्गीकरण के आवश्यक तत्व (Essential Elements of Ideal Classification)

अनुसंधान की प्रकृति, उद्देश्य, क्षेत्र आदि को ध्यान में रखते हुए यह आवश्यक है कि वर्गीकरण में निम्न विशेषताएं हो:

1. **असंदिग्धता तथा स्पष्टता (Unambiguity and Clarity)**- विभिन्न वर्गों की परिभाषा अथवा निर्धारण अथवा योजना स्पष्ट, सरल व निश्चित होनी चाहिए जिससे किस इकाई को किस वर्ग में रखा जाना है इस प्रकार की संदिग्धता व अनिश्चितता की गुंजाइश न रहे।
2. **निःशेषी (Exhaustive)** - वर्गीकरण के लिए यह आवश्यक है कि प्रत्येक इकाई किसी न किसी वर्ग में सम्मिलित हो।

3. **पारस्परिक पृथक्ता (Mutually Exclusive)** - विभिन्न वर्गों में पारस्परिक पृथक्ता होनी चाहिए अर्थात् विभिन्न वर्ग परस्पर अपवर्जी (non-overlapping) होने चाहिए।
4. **स्थिरता (Stability)** - वर्गीकरण एक ही सिद्धान्त के आधार पर किया जाए और प्रत्येक स्तर पर उसी आधार को बनाये रखा जाये।
5. **अनुकूलता (Suitability)** - वर्गीकरण अनुसंधान के अनुरूप होना चाहिए।
6. **लचीलापन (Flexibility)** - वर्गीकरण लोचदार होना चाहिए जिससे नवीन परिस्थितियों के अंतर्गत आवश्यकतानुसार संशोधन किया जा सके।
7. **सजातीयता (Homogeneity)** -प्रत्येक वर्ग की इकाइयों में सजातीयता होनी चाहिए।
8. **गणितीय शुद्धता (Mathematical Precision)** - मापन, योग आदि शुद्ध एवं सही होने चाहिए।

आँकड़ों के वर्गीकरण के कार्य (Functions of Classification of Data)

वर्गीकरण के कार्यों को निम्नलिखित रूप में व्यक्त किया जा सकता है-

1. **यह आँकडों को संक्षिप्त करता है-** वर्गीकरण बड़े अव्यवस्थित एवं जटिल आँकडों को संक्षिप्त रूप में व्यक्त करता है जिसकी वजह से आँकडों को समझने में आसानी होती है। वर्गीकरण आँकडों के उल्लेखनीय तथ्यों पर प्रकाश डालने का कार्य भी करता है।
2. **तुलना करने में सहायक (Helpful in Comparison)-** वर्गीकरण से आँकडों का तुलनात्मक अध्ययन सम्भव होता है। जैसे- दो महाविद्यालयों के छात्रों में बौद्धिक स्तर की तुलना प्राप्तांकों को प्रथम, द्वितीय व तृतीय श्रेणी के आधार पर वर्गों में विभाजित करके सरलता से की जा सकती है।
3. **संबंधों का अध्ययन करने में सहायक (Helpful in Studying Relationships):** आँकडों का वर्गीकरण दो या उससे अधिक कसौटियों के बीच संबंध का अध्ययन करने में सहायक होते है।

आँकड़ों के वर्गीकरण के आधार (Basis of Classification of Data):

यद्यपि आँकडों के वर्गीकरण के अनेक सम्भव आधार हैं परन्तु अधिकतर जिन आधारों पर आँकडों का वर्गीकरण किया जाता है वे इस प्रकार हैंः

1. **आकार (Size or magnitude) के आधार पर**- आकार के आधार पर वर्गीकरण को संख्यात्मक वर्गीकरण (Quantitative Classification) कहा जाता है।
2. **स्थान (Place or location) के आधार पर**- स्थान के आधार पर वर्गीकरण को भौगोलिक वर्गीकरण (Geographical Classification) कहा जाता है।
3. **प्रकार या गुण (Kind or Type or Attribute) के आधार पर**- प्रकार व गुण आदि के आधार पर वर्गीकरण को वर्णनात्मक वर्गीकरण (Descriptive Classification) कहा जाता है।
4. **समय (Time) के आधार पर-** समय के आधार पर वर्गीकरण को कालानुसार वर्गीकरण (Chronological Classification) कहा जाता है।

आँकड़ों के वर्गीकरण की विधियाँ (Methods Of Classification of Data):

किसी पूर्व उद्देश्य के लिए सुव्यवस्थित रूप से सांख्यिकीय विषय-सामग्री जिनमें प्रायः संख्यात्मक तथ्य होते हैं, का संकलन किया जाता है। ये तथ्य जिन्हें संख्याओं में व्यक्त किया जाता है, दो प्रकार के हो सकते हैंः

1. अनुसन्धान में अध्ययनरत इकाइयों में व्याप्त किसी गुणात्मक विशेषता (गरीबी, ईमानदारी, जाति, साक्षरता, अन्धापन, बुद्धिमत्ता आदि) को प्रायः गुण (Attribute) कहा जाता है। जैसे- किसी देश में गरीबी की स्थित क्या हैं कितने लोग गरीबी में जीवनयापन कर रहें हैं, बेरोजगार कितने हैं?
2. इकाइयों की किसी मापन या गणना योग्य विशेषता जिसे प्रायः चर (variable) कहा जाता है, के आधार पर गणना या मापन द्वारा प्राप्त आँकड़े। जैसे- आयु, बच्चों की संख्या, वस्तुओं के मूल्य, ऊंचाई, भार, दूरी आदि।

इन दो प्रकार के तथ्यों के आधार पर वर्गीकरण की दो विधियाँ हैंः

(क) गुणात्मक वर्गीकरण (Attributes/Qualitative Classification),
(ख) संख्यात्मक वर्गीकरण (Numerical/Quantitative Classification)

(क) गुणात्मक वर्गीकरण (Attributes/Qualitative Classification)

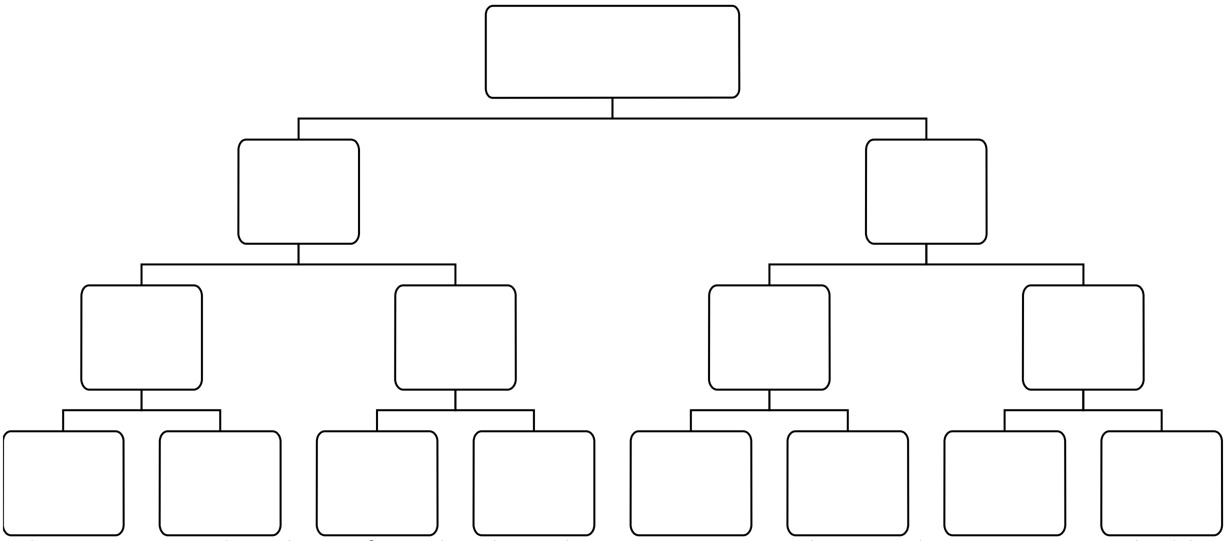
गुणों की उपस्थिति अथवा अनुपस्थिति के आधार पर किया गया वर्गीकरण गुणात्मक वर्गीकरण कहलाता है। गुणात्मक वर्गीकरण दो प्रकार का हो सकता हैः

1. **द्वन्द्वभाजन वर्गीकरण (Dichotomy Classification**)- यदि इकाइयों के समूह को एक गुण के आधार पर उसकी उपस्थिति या अनुपस्थिति के अनुसार दो वर्गों (Classes) में बांटते हैं तो ऐसे

गुणात्मक वर्गीकरण को द्वन्द्वभाजन वर्गीकरण या साधारण वर्गीकरण (Simple Classification) या एक गुण वर्गीकरण कहते हैं। जैसे- पुरुष और महिला, ग्रामीण एवं शहरी, विवाहित एवं अविवाहित आदि।

2. **बहुगुणी वर्गीकरण या लगातार द्वन्द्वभाजन वर्गीकरण (Manifold Classification or Classification as a Series of Dichotomies)** - यदि एक गुण (माना जनसंख्या) के आधार पर वर्गीकरण किया जाता है तो दो वर्ग (शहरी एवं ग्रामीण) बनेंगे। फिर प्रत्येक वर्ग को दूसरे गुण (माना पुरुष एवं महिला) के आधार पर वर्गीकरण किया जाये तो प्रत्येक वर्ग के दो उपवर्ग होंगे। इस प्रकार दो गुणों के आधार पर लगातार द्वन्द्वभाजन द्वारा चार वर्ग बनेंगे।

उदाहरण के लिए:



गुणों की उपस्थिति को अंग्रेजी वर्णमाला के बड़े अक्षरों (A,B,C,.. आदि) में तथा गुणों की अनुपस्थिति को छोटे अक्षरों (a, b,c,...आदि) अथवा ग्रीक अक्षरों (α,β,..आदि) में व्यक्त किया जाता है।

सावधानियां (Precautions) - इस प्रकार गुणात्मक वर्गीकरण करना सरल है परन्तु निम्न सावधानियां रखना वांछनीय हैः

(1) आधार का स्पष्ट होना- गुण की उपस्थिति अथवा अनुपस्थिति का आधार स्पष्ट रूप से निश्चित होना चाहिए। जैसे- यदि वयस्क और अवयस्क दो वर्गों में बांटना है तो यह निश्चित होना चाहिए कि किस आयु तक व्यस्क तथा किस आयु तक अवयस्क माना जायेगा।

(2) परिवर्तनों को ध्यान रखना- एकत्रित आँकडों में परिवर्तन होता रहता है जैसे-अशिक्षित शिक्षित हो जाते हैं। इसका ध्यान रखना बहुत आवश्यक है।

(ख) संख्यात्मक वर्गीकरण (Numerical/Quantitative Classification)

आँकडों का संख्या/मात्रा के आधार पर वर्गीकरण संख्यात्मक वर्गीकरण कहलाता है। इस वर्गीकरण की प्रमुख विधियाँ निम्नलिखित है-

1. **कालक्रमबद्ध वर्गीकरण (Chronological Classification):** जब आँकडों को काल या समय (घंटे, दिन, महीने, साल आदि) के अनुसार वर्गीकृत किया जाता है तो उसे कालक्रमबद्ध वर्गीकरण कहते है। उदाहरण के लिए, किसी देश में विगत 10 वर्षों के मोटर वाहनों के विनिर्माण के आँकड़े।
2. **भौगोलिक वर्गीकरण (Geographical Classification):** आँकडों का क्षेत्र या स्थान के अनुसार वर्गीकरण भौगोलिक वर्गीकरण कहलाता है। इसे क्षेत्रीय वर्गीकरण (Regional Classification) भी कहा जाता है। उदाहरण के लिए, प्रमुख देशों के सकल घरेलु उत्पाद (GDP) के आँकड़े।

3. चर-मूल्य वर्गीकरण (Variable Classification): आँकडों का चरों के आधार पर वर्गीकरण चर-मूल्य वर्गीकरण कहलाता है। चर दो प्रकार के होते है- खण्डित चर (Discrete variable) एवं अखण्डित चर (Continuous variable)।

i. खण्डित चर (Discrete Variable)- वह चर जिसके मान परिमित हों या अपरिमित परन्तु अलग-अलग गणनीय हों, खण्डित या असतत् चर कहा जाता है। जैसे- छात्रों की संख्या, प्रति पृष्ठ गलतियों की संख्या, कर्मचारियों का मासिक वेतन आदि। ये सभी पूर्णांक रूप में होते है अर्थात छात्रों की संख्या 20 या 30 हो सकती है, 30.5 नहीं।

ii. अखण्डित चर (Continuous Variable)- वह चर जो किसी अन्तराल या अन्तरालों में सैद्धान्तिक रूप से (व्यावहारिक रूप से सम्भव हो या न हो) प्रत्येक मान ले सकता है, सतत् चर या अखण्डित चर कहलाता है। जैसे- छात्रों की आयु, कर्मचारियों की कुल आय, व्यक्तियों की ऊंचाई आदि। छात्रों की आयु 18 वर्ष भी हो सकती है, 18.5 वर्ष भी हो सकती है। इसी प्रकार व्यक्तियों की ऊंचाई 150 सेंटीमीटर भी हो सकती है और 150.35 सेंटीमीटर भी।

आवृति वितरण के प्रकार (Types of Frequency Distribution)

आवृति वितरण एक ऐसी तालिका होती है जिसमे चर के विभिन्न मूल्यों को आकार के क्रम मे अलग अलग या समूहों में संबद्ध आवृतियों के साथ साथ दिखाया जाता है। इसे दो प्रकार से वर्गीकृत किया जा सकता है- एक चरीय एवं द्विचरीय आवृति वितरण।

1. एक चरीय आवृति वितरण (Univariate Frequency Distribution): जब एक चर मूल्य के आधार पर आवृतियाँ निर्धारित की जाती हैं तो इसे एक चरीय आवृति वितरण कहते है। उदाहरण के लिए, अंकों के आधार पर विद्यार्थियों की संख्या या वेतन के अनुसार कर्मचारियों की संख्या। इसके भी दो प्रकार होते है- i. खंडित आवृति वितरण ii. सतत् या अखण्डित आवृति वितरण

i. खंडित आवृति वितरण (Discrete Frequency Distribution): जब खण्डित चरों के आधार पर आवृति तालिका का निर्माण किया जाता है तो उसे खंडित आवृति वितरण कहते हैं। उदाहरण के लिए, किसी कक्षा के 50 विद्यार्थियों के कुल 10 अंकों में से प्राप्तांकों का आवृति वितरण निम्न प्रकार से दर्शाया जा सकता है-

प्राप्तांक (Marks Obtains)	मिलान चिन्ह (Tallies)	विद्यार्थियों की संख्या/आवृतियाँ (Numbers of Student /Frequency)
0	II	2
1	III	3
2	~~IIII~~ III	8
3	IIII	4
4	~~IIII~~ I	6
5	~~IIII~~ ~~IIII~~	10
6	IIII	4
7	~~IIII~~ II	7
8	I	1
9	III	3
10	II	2
		कुल संख्या= 50

ii. सतत या अखण्डित आवृति वितरण (Continuous Frequency Distribution): जब अखंडित या सतत चरों के आधार पर आवृति तालिका का निर्माण किया जाता है तो उसे अखंडित आवृति वितरण कहते हैं। यह वर्गान्तर (Class-Interval) पर आधारित होता है। उदाहरण के लिए, किसी कक्षा के 50 विद्यार्थियों की ऊँचाई के अनुसार आवृति वितरण निम्न प्रकार से दर्शाया जा सकता है-

वर्गान्तर (Class-Interval) (cm में)	**विद्यार्थियों की संख्या/आवृतियाँ (Numbers of Student /Frequency)**
110-120	10
120-130	7

130-140	4
140-150	9
150-160	3
160-170	15
170-180	2
	कुल संख्या= 50

2. द्विचरीय आवृति वितरण (Bivariate Frequency Distribution): जब दो चर मूल्य के आधार पर आवृतियों का वर्गीकरण किया जाता हैं तो इसे दो चरीय आवृति वितरण कहते है। उदाहरण के लिए, किसीकक्षा के 20 विद्यार्थियों के अर्थशास्त्र विषय में 20 से लेकर 24 प्राप्तांक एवं सांख्यिकी में 10 से 14 प्राप्तांको के आधार पर निम्न सारणी बनेगी-

अर्थशास्त्र एवं सांख्यिकी मे 20 विद्यार्थियों के अंकों को प्रदर्शित करती द्विचर आवृति सारणी

सांख्यिकी में प्राप्तांक	अर्थशास्त्र में प्राप्तांक					
	20	21	22	23	24	कुल
10	II		I	II		5
11		I	I		I	3
12	I			I		2
13	II	I	II		I	6
14		II		II		4
कुल	5	4	4	5	2	20

3.4. सारणीयन का परिचय (Introduction to Tabulation)

सारणीयन अथवा आँकडों को सारणियों के रूप में प्रदर्शन सांख्यिकीय विश्लेषण का एक अत्यन्त महत्वपूर्ण भाग है। सारिणी (Table) से हमारा अभिप्राय आँकडों की पंक्तियों (Columns) में क्रमबद्ध व्यवस्था से है जिससे समंकों की विशेषताएँ उभर कर स्पष्ट रूप से सामने आ जायें। सारणीयों के रूप में प्रदर्शित आँकड़े न केवल सांख्यिकीय विश्लेषण के दृष्टिकोण से सुविधाजनक होते हैं, वरन् यह मूल आँकडों की जटिलताओं को कम करने में भी सहायक होते हैं। **कॉनर** के अनुसार, ***''सारणीयन आंकिक सामग्री को किसी व्यवस्थित एवं क्रमबद्ध ढंग से प्रदर्शित करने की एक रीति है जिसका उद्देश्य किसी विचाराधीन समस्या पर पर्याप्त प्रकाश डालना है।''***

वास्तव में आँकडों को व्यवस्थित रूप में प्रदर्शित करने की तीन विधियाँ हैं- उद्धरण रूप में (Textual Representation), सारिणी के रूप में (Tabular Representation) एवं ग्राफ अथवा आरेखों के रूप में (Graphical or Diagrammatic Representation)

सारिणीयन के उद्देश्य (Objective of Tabulation)

सारणीयन के प्रमुख उद्देश्य निम्नलिखित हैं-

(1) सारणीयन का प्रमुख उद्देश्य विशाल तथा बिखरे समंकों को संक्षिप्त रूप में व्यवस्थित क्रम से सामने लाना।
(2) अनुसंधान का उद्देश्य सुलभ रूप में प्रस्तुत करना।
(3) न्यूनतम स्थान में तथ्यों को सामने रखना।
(4) सांख्यिकीय विधियों के प्रयोग को सुगम बनाना।

सारणीयन करते समय निम्नांकित बातों को ध्यान में रखना आवश्यक है-

(1) प्रत्येक सारणी के ऊपर स्पष्ट रूप से शीर्षक लिखा होना चाहिए जिससे यह स्पष्ट हो जाए कि सारिणी में दी गयी सूचनाएँ किस विषय से संबंधित हैं।
(2) सारणी उचित मात्रा में स्तम्भों तथा पंक्तियों में विभाजित होनी चाहिए, जहाँ तक संभव हो पंक्तियों की संख्या अधिक रखनी चाहिए।
(3) सारिणी की रचना के पूर्व खानों की संख्या तथा विस्तार आदि को निश्चित कर लेना चाहिए।
(4) प्रत्येक स्तंभ पर उपशीर्षक तथा उसकी संख्या स्पष्ट रूप से दी जानी चाहिए।
(5) जहाँ तक सम्भव हो सारणीयन की प्रक्रिया सरल, श्रम को बचाने वाली तथा कम खर्चीली हो।
(6) सारणी इस प्रकार से तैयार की जानी चाहिए जिससे प्रस्तुत तथ्यों की जाँच दूसरी ओर से भी की जा सके जिससे अशुद्धियाँ कम हो।

आदर्श सारणीयन के आवश्यक तत्व (Essential Elements of Ideal Tabulation)

इन सामान्य उद्देश्यों को दृष्टिगत रखते हुए एक सांख्यिकीय सारिणी में निम्न तत्वों का होना आवश्यक है-

1. सारिणी को सरल, संक्षिप्त, आकर्षक तथा स्वतः स्पष्ट (Self-Explanatory) होना चाहिए।
2. सारिणी का शीर्षक (Title) सारिणी के ऊपर स्पष्ट रूप से अंकित होना चाहिए तथा इसमें किसी प्रकार की अस्पष्टता अथवा सन्दिग्धता नहीं होना चाहिए।
3. शीर्षक के नीचे सारिणी में दर्शायी गई राशियों की माप की इकाइयों का उल्लेख करना चाहिए।
4. सारिणी में प्रयुक्त पंक्तियों के शीर्षक जिन्हें कि अनुशीर्षक (Sub Heading) कहा जाता है, स्पष्ट रूप से दर्शाये जाने चाहिए। इस प्रकार स्तम्भों (Columns) के शीर्षक जिन्हें कि उपशीर्षक (Captions) कहा जाता है स्पष्ट रूप से अंकित होने चाहिए।
5. सारिणी के मुख्य भाग (Main Body of the Table) में प्रविष्टियों अथवा संबंधित आँकडों को पंक्तियों तथा स्तम्भों के शीर्षकों के अनुसार भरा जाना चाहिए। यदि किसी तथ्य से संबंधित आँकड़े उपलब्ध नहीं हैं तो सारिणी में संबंधित स्थान/खानों पर स्पष्ट रूप में 'N.A.' (Not Available) अंकित करना चाहिए। यदि आँकडों के सापेक्षित मानों अथवा प्रतिशतों को भी विश्लेषण में प्रयुक्त किया जाना है तो इन्हें मूल आँकडों के नीचे कोष्टकों (Brackets) में दर्शाया जाता है।
6. सारिणी में दर्शाए गए वे आँकड़े जिनकी व्याख्या मुख्य शीर्षक (Title) अनुशीर्षक (Sub Headings) अथवा उपशीर्षक (Captions) के अंतर्गत नहीं की जा सकी है, उन्हें स्पष्ट करने के लिये सारिणी के मुख्य-भाग के नीचे व्याख्यात्मक टिप्पणी (Footnote) लिखी जानी चाहिए।
7. सारिणी के नीचे आँकडों के स्रोत (Source of Data) अथवा सन्दर्भ ग्रन्थों (References) के विषय में जानकारी अंकित होनी चाहिए।

सारिणी का प्रारूप:

सारणी शीर्षक (Table Title)

खाने (Column)

पंक्तियाँ (Rows)

पंक्ति शीर्षक	उप शीर्षक (Captions)	उप शीर्षक (Captions)	योग (Total)
अनुशीर्षक (Sub Headings)			
अनुशीर्षक (Sub Headings)			
योग (Total)			

टिप्पणी (Note):

आँकडों के स्रोत (Source of Data):

अनुसंधान में सारणी के प्रमुख तत्त्व (Main Element of Table in Research)

एक अनुसंधान में सारणी के कुछ प्रमुख तत्व होने चाहिए जो निम्नलिखित हैं-

1. **उचित आकार (Proper Size):** अनुसंधान में सारणी का आकार इस प्रकार होना चाहिए कि उसको देखने, समझने में आसानी हो। एक आदर्श सारणी उचित आकार की होनी चाहिए। यदि सारणी में अधिक जानकारियाँ प्रस्तुत करनी हो तो एक बड़ी सारिणी के स्थान पर अनेक छोटी-छोटी सारणियाँ बनाना ज्यादा उचित माना जाता है।
2. **उद्देश्य के अनुसार (According to the Objecive):** अनुसंधान में सारणी बनाते समय यह ध्यान रखना आवश्यक है कि उसकी विषय-सामग्री उद्देश्य के हिसाब से होनी चाहिए। उद्देश्य के अतिरिक्त विषय-सामग्री को सारणी में शामिल नहीं करना चाहिए।
3. **तुलनात्मक (Comparable):** अनुसंधान में सारणी में आँकडों को इस तरह व्यवस्थित करना चाहिए जिससे की आँकडों का तुलनात्मक अध्ययन कर आसानी से निष्कर्ष ज्ञात किया जा सके।
4. **सारणीयन के नियमों का उपयोग (Use of Rules of Tabulation):** अनुसंधान में सारणी बनाते समय सारणी संख्या, शीर्षक, उप-शीर्षक, आँकडों के स्रोत आदि के विषय में सारणीयन के विभिन्न नियमों को ध्यान में रखकर उनका पालन करना आवश्यक होता है।

सारणीयन के वर्गीकरण के मुख्य आधार (Main Basis of Classification of Tabulation)

सारणियों को तीन प्रमुख आधारों पर निम्नलिखित प्रकार से वर्गीकृत किया जा सकता है-

1. निर्माण के आधार पर (On the Basis of Construction): निर्माण के आधार पर सारिणी को दो प्रकार से बाँटा जा सकता है- सरल एवं जटिल सारिणी

i. सरल सारिणी (Simple Table): जब सारिणी को एक गुण के आधार पर आँकडों को वर्गीकृत किया जाता है तो वह सरल सारिणी कहा जाता है। इसे एक गुण सारिणी या प्रथम क्रम की सारिणी भी कहा जाता है। जैसे- महाविद्यालय के विद्यार्थियों का संकाय के अनुसार वर्गीकरण आदि। उदाहरण-

उत्तराखंड मुक्त विश्वविद्यालय में विभिन्न विद्याशाखाओं में विद्यार्थियों की संख्या (वर्ष 2024)

विद्याशाखा	विद्यार्थियों की संख्या	विद्याशाखा	विद्यार्थियों की संख्या
कला		शिक्षा	
वाणिज्य		विज्ञान	
कृषि		आयुर्वेद	
विधि		पर्यटन	

ii. जटिल सारिणी (Complex Table): जब सारिणी में एक से अधिक गुणों के आधार पर आँकडों को वर्गीकृत किया जाता है तो उसे जटिल सारिणी कहते है। जटिल सारिणी के भी तीन प्रकार होते है-

अ. द्विगुण सारिणी (Two way Table): इस सारिणी में दो गुणों के आधार पर आँकडों को विभाजित किया जाता है। उदाहरण-

साक्षरता के अनुसार जनसंख्या का वर्गीकरण:

राज्य/जनसंख्या	साक्षर	निरक्षर	योग
योग			

ब. त्रिगुण सारिणी (Three Way Table): इस सारिणी में तीन परस्पर संबंधित गुणों के आधार पर आँकडों को वर्गीकृत किया जाता है। उदाहरण-

जनसंख्या का लिंग एवं साक्षरता के अनुसार वर्गीकरण

राज्य	पुरुष			महिला		
	साक्षर	निरक्षर	कुल	साक्षर	निरक्षर	कुल
कुल						

स. बहुगुण सारिणी (Manifold Table): इस सारिणी में तीन से अधिक परस्पर संबंधित गुणों के आधार पर आँकडों को वर्गीकृत किया जाता है। उदाहरण-

राज्य	पुरुष				महिला				कुल
	ग्रामीण		शहरी		ग्रामीण		शहरी		
	साक्षर	निरक्षर	साक्षर	निरक्षर	साक्षर	निरक्षर	साक्षर	निरक्षर	
कुल									

2. उद्देश्य के आधार पर (On the Basis of Purpose): उद्देश्य के आधार पर सारिणी को दो प्रकार से वर्गीकृत किया जाता है- सामान्य उद्देश्य एवं विशेष उद्देश्य वाली सारिणी।

i. सामान्य उद्देश्य वाली सारिणी (General Purpose Table): इस प्रकार की सारिणी में आँकडों का व्यापक भंडार होता है इसलिए इसे संदर्भ सारिणी (Reference Table) भी कहते है। उदाहरण के लिए- किसी देश में जनसंख्या के आँकडों वाली सारिणी में विभिन्न पहलुओं के बारे में सामान्य एवं व्यापक जानकारी होती है।

ii. विशेष उद्देश्य वाली सारिणी (Special Purpose Table): इस प्रकार की सारिणी को किसी विशेष उद्देश्य या जानकारी पर प्रकाश डालने हेतु बनाया जाता है। इसे सारांश सारिणी (Summary Table) भी कहा जाता है। उदाहरण के लिए, किसी फैक्ट्री के विभिन्न वर्षों में हुए लाभ की राशि को दर्शाती हुई सारिणी। इस प्रकार की सारिणी सामान्यः छोटे आकार की होती है।

3. मौलिकता के आधार पर (On the Basis of Originality): मौलिकता के आधार पर सारिणियों को दो प्रकार से विभाजित किया जा सकता है-

i. मौलिक सारिणी (Original Table): जिस सारिणी में सारे आँकड़े मौलिक होते है वह मौलिक सारिणी (Original Table) कहलाती है। इन्हें प्राथमिक सारिणी भी कहा जाता है।

ii. व्युत्पन्न सारिणी (Derivative Table): इस प्रकार की सारिणी में मौलिक आँकडों के आधार पर निकाले गए योग, माध्य, प्रमाप विचलन आदि को दर्शाया जाता है।

अतः उपरोक्त के आधार पर हम कह सकते है कि सारणियों का वर्गीकरण तीन मुख्य आधारों- निर्माण, उद्देश्य तथा मौलिकता के आधार पर किया जा सकता है। निर्माण के आधार पर सारिणियों को सरल और जटिल में बांटा जाता है। सरल सारिणी एक गुण पर आधारित होती है जबकि जटिल सारिणियाँ एक से अधिक गुणों पर आधारित होती हैं, जिन्हें द्विगुण, त्रिगुण और बहुगुण सारिणियों में विभाजित किया जाता है। द्विगुण सारिणी में दो गुण, त्रिगुण सारिणी में तीन गुण, और बहुगुण सारिणी में तीन से अधिक गुण होते हैं। उद्देश्य के आधार पर सारिणियाँ सामान्य और विशेष उद्देश्य वाली हो सकती हैं। सामान्य उद्देश्य वाली सारिणियों में व्यापक जानकारी होती है जबकि विशेष उद्देश्य वाली सारिणियाँ किसी विशिष्ट जानकारी के लिए बनाई जाती हैं। मौलिकता के आधार पर सारिणियाँ मौलिक और व्युत्पन्न होती हैं। मौलिक सारिणियों में सारे आँकड़े मूल होते हैं, जबकि व्युत्पन्न सारिणियों में मौलिक आँकड़ों से निकाले गए योग, माध्य, प्रमाप विचलन आदि दिखाए जाते हैं।

वर्तमान समय में वर्गीकरण एवं सारणीयन के लिए कम्प्यूटर का उपयोग अधिक प्रचलित है। इसके लिए कई प्रकार के सॉफ्टवेयर जैसे एम.एस. ऑफिस, आदि का उपयोग किया जाता है वर्तमान समय में वर्गीकरण और सारणीयन के लिए कम्प्यूटर का उपयोग अत्यधिक प्रचलित हो गया है। कम्प्यूटर के माध्यम से आँकड़ों का त्वरित और सटीक विश्लेषण संभव हो गया है जिससे शोधकर्ताओं, डेटा विश्लेषकों और विभिन्न संगठनों के लिए यह एक अनिवार्य उपकरण बन गया है। कम्प्यूटर का उपयोग समय की बचत करता है और मानवीय त्रुटियों को कम करता है जिससे निर्णय लेने की प्रक्रिया अधिक प्रभावी और भरोसेमंद हो जाती है। एक्सेल, एस.पी.एस.एस. (SPSS) आर और पायथन जैसी सॉफ्टवेयर प्रणालियाँ आँकड़ों को वर्गीकृत करने, सारणीबद्ध करने और विश्लेषण करने के लिए व्यापक रूप से उपयोग की जाती हैं। इसमे कुछ प्रमुख सॉफ्टवेयर प्रणालियों के बारे में आप अगले सेमेस्टर में विस्तार से पढ़ेंगे।

3.5 अभ्यास प्रश्न (Practice Questions)

बहुविकल्पीय प्रश्न -

1. तथ्यों के आधार पर निम्न में से वर्गीकरण की प्रमुख विधियाँ हैं:

   A. गुणात्मक वर्गीकरण　　B. संख्यात्मक वर्गीकरण

   C. दोनों गुणात्मक एवं संख्यात्मक वर्गीकरण　　D. सभी विकल्प गलत हैं

2. निम्नलिखित में से सही विकल्पों का चयन कीजिए:

   कथन 1: गुणों की उपस्थिति अथवा अनुपस्थिति के आधार पर किया गया वर्गीकरण गुणात्मक वर्गीकरण कहलाता है।

   कथन 2: आँकडों का संख्या/मात्रा के आधार पर वर्गीकरण संख्यात्मक वर्गीकरण कहलाता है।

   विकल्प:

   A. कथन 1 सही है　　B. कथन 2 सही है

   C. कथन 1 और 2 दोनों गलत है　　D. कथन 1 और 2 दोनों सही हैं

3.6 सारांश (Summary)

संकलित आँकड़ें अत्यधिक जटिल, अस्पष्ट एवं अव्यवस्थित रूप में होते है। उन्हें ऊपरी नज़र से देखने पर समझना एवं कोई निष्कर्ष निकालना कठिन होता है। अतः एकत्रित आँकडों को सरल एवं समझने योग्य बनाने हेतु किसी निश्चित एवं तार्किक आधार पर व्यवस्थित करना वर्गीकरण एवं सारणीयन कहलाता है। संकलित आँकडों को किसी गुण के आधार पर समान व असमान कर अलग-अलग वर्गों में बांटने की प्रक्रिया को वर्गीकरण कहा जाता है। वर्गीकरण बड़े अव्यवस्थित एवं जटिल आँकडों को संक्षिप्त रूप में व्यक्त करता है जिसकी वजह से आँकडों को समझने में आसानी होती है। वर्गीकरण से आँकडों का तुलनात्मक अध्ययन सम्भव होता है। आँकडों का वर्गीकरण आकार, स्थान, गुण एवं समय (Time) आदि के आधार पर हो सकता है। तथ्यों के आधार पर वर्गीकरण की दो प्रमुख विधियाँ हैं: गुणात्मक एवं संख्यात्मक वर्गीकरण। गुणों की उपस्थिति अथवा अनुपस्थिति के आधार पर किया गया वर्गीकरण गुणात्मक वर्गीकरण कहलाता है जबकि आँकडों का संख्या/मात्रा के आधार पर वर्गीकरण संख्यात्मक वर्गीकरण कहलाता है। सारणीयन अथवा आँकडों को सारणियों के रूप में प्रदर्शन सांख्यिकीय विश्लेषण का एक अत्यन्त महत्वपूर्ण भाग है। सारिणी (Table) से हमारा अभिप्राय आँकडों की पंक्तियों

(Columns) में क्रमबद्ध व्यवस्था से है जिससे समंकों की विशेषताएँ उभर कर स्पष्ट रूप से सामने आ जायें। एक आदर्श सारणी के प्रमुख तत्व निम्नलिखित है- उचित आकार, उद्देश्य के अनुसार तथा सारणीयन के नियमों का उपयोग। सारणियों को तीन प्रमुख आधारों पर वर्गीकृत किया जा सकता है- निर्माण के आधार पर, उद्देश्य के आधार पर एवं मौलिकता के आधार पर।

3.7 शब्दावली (Glossary)

- **कालक्रमबद्ध वर्गीकरण (Chronological Classification):** जब आँकडों को काल या समय (घंटे, दिन, महीने, साल आदि) के अनुसार वर्गीकृत किया जाता है तो उसे कालक्रमबद्ध वर्गीकरण कहते है। उदाहरण के लिए, किसी देश में विगत 10 वर्षों के मोटर वाहनों के विनिर्माण के आँकड़े।
- **भौगोलिक वर्गीकरण (Geographical Classification):** आँकडों का क्षेत्र या स्थान के अनुसार वर्गीकरण भौगोलिक वर्गीकरण कहलाता है। इसे क्षेत्रीय वर्गीकरण (Regional Classification) भी कहा जाता है। उदाहरण के लिए, प्रमुख देशों के सकल घरेलु उत्पाद (GDP) के आँकड़े।
- **चर-मूल्य वर्गीकरण (Variable Classification):** आँकडों का चरों के आधार पर वर्गीकरण चर-मूल्य वर्गीकरण कहलाता है।
- **खण्डित चर (Discrete Variable)-** वह चर जिसके मान परिमित हों या अपरिमित परन्तु अलग-अलग गणनीय हों, खण्डित या असतत् चर कहा जाता है। जैसे- छात्रों की संख्या, प्रति पृष्ठ गलतियों की संख्या, कर्मचारियों का मासिक वेतन आदि।
- **अखण्डित चर (Continuous Variable)-** वह चर जो किसी अन्तराल या अन्तरालों में सैद्धान्तिक रूप से (व्यावहारिक रूप से सम्भव हो या न हो) प्रत्येक मान ले सकता है, सतत् चर या खण्डित चर कहलाता है। जैसे- छात्रों की आयु, कर्मचारियों की कुल आय, व्यक्तियों की ऊंचाई आदि।

3.8 अभ्यास प्रश्नों के उत्तर (Answer for Practice Questions)

बहुविकल्पीय प्रश्न -

1. C 2. D

3.9 सहायक / उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

- Roy Ramendu,(2017), ***Principles of Statistics,*** Prayag Pustak Bhawan, Allahabad.
- Gupta S. P. & Gupta M. P.(2010), ***Business Statistics,*** Sultan Chand & Sons, New Delhi.
- Bhardwaj, R.S. (2000), ***Mathematics for Economics and Business***, Excel Books.
- Gupta, Jain, Aggrawal (2006), ***Basic Quantitative Methods for Economics***, Navyug Sahitya Sadan,Agra.
- Bose, D., (2003), ***An Introduction to Mathematical Economics***, Himalaya Publishing House.
- Kumar, Anil (2008), ***Statistical Research Methodology***, Aifa Publishing House.

3.10 सहायक / उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

- एस०एन०लाल, एल०के०चतुर्वेदी, (2010), ***परिमाणात्मक विश्लेषण***, शिव पब्लिशिंग हाउस, इलाहाबाद।
- नागर, कैलाश नाथ (2008), ***सांख्यिकी के मूल तत्व***, मीनाक्षी प्रकाशन, मेरठ।
- सिंह, एस 0 पी0 (2010), ***सांख्यिकी सिद्धांत एवं व्यवहार***, एस चन्द एण्ड कंपनी लिमिटेड, दिल्ली।

3.11 निबंधात्मक प्रश्न (Essay Type Questions)

1. वर्गीकरण से आप क्या समझते है? आँकडों के वर्गीकरण के उद्देश्य तथा विधियों की विवेचना कीजिए?
2. वर्गीकरण एवं सारणीयन में अंतर स्पष्ट कीजिए?
3. सारणी बनाते समय किन किन बातों का ध्यान रखना आवश्यक होता है?

इकाई 4 - आंकड़ों का प्रस्तुतीकरण की विधियाँ

4.1 प्रस्तावना

वास्तव में आँकडों को व्यवस्थित रूप में प्रदर्शित करने की तीन विधियाँ हैं- उद्धरण रूप में (Textual Representation), सारिणी के रूप में (Tabular Representation) एवं ग्राफ/ बिन्दुरेखीय अथवा आरेखों के रूप में (Graphical or Diagrammatic Representation)। पिछली इकाई में आप उद्धरण रूप एवं सारिणी के रूप में आँकडों को प्रदर्शित करने की विधियों के बारे में जान चुके है। इस इकाई मे आप बिन्दुरेखीय एवं चित्र विधि के अर्थ, परिभाषा एवं इनके गुण और दोषों के बारे मे विस्तारपूर्वक पढ़ेंगे।

4.2 उद्देश्य

इस इकाई के अध्ययन के पश्चात आप

- ✓ आंकड़ों के प्रस्तुतीकरण की सबसे मुख्य इकाई बिन्दुरेखीय विधि का अर्थ जान सकेंगें।
- ✓ चित्र तथा बिन्दुरेखीय में अंतर कर सकेंगे।
- ✓ बिन्दुरेखीय प्रदर्शन की उपयोगिता जानेंगे।
- ✓ बिन्दुरेखीय विधि से प्रदर्शन के गुण और दोषों की विवेचना कर सकेंगे।
- ✓ बिन्दुरेखीय प्रदर्शन की उपयोगिता प्राप्त कर सकेंगे।

4.3 बिन्दुरेखीय विधि (Graphical Method)

सांख्यिकी में आँकड़ों को प्रदर्शित करने की दो विधियों हैं -

(अ) बिन्दुरेखीय प्रदर्शन (ब) चित्रमय प्रदर्शन

4.3.1 बिन्दुरेख का अर्थ एवं परिभाषा (Meaning and Definition of Graph)

सांख्यिकीय आंकड़ों का ग्राफ पेपर पर प्रदर्शन बिन्दुरेख (Graph) कहलाता है। अधिकांश सांख्यिकीय आंकड़ें इतने विशाल और जटिल होते हैं कि जन-सामान्य के लिए उनका समझना अत्यन्त कठिन है। वर्गीकरण व सारणीयन समंकों को व्यवस्थित व सुन्दर ढंग से प्रस्तुत करते हैं, परन्तु इनके द्वारा आंकड़ों की विशेषताओं को ठीक प्रकार से प्रदर्शित नहीं किया जा सकता।

एम. एम. ब्लेयर (M.M. Blair) के अनुसार ***''समझने में व रचना में सरलतम, सर्वाधिक चल और सबसे अधिक प्रयोग में लाया जाने वाला चित्र बिन्दुरेख है। (The simplest to understand, the easiest to make, the most variable and the most widely used type of chart is the line graph.)''***

आई. आर. वेसेलो (I.R. Vessolo) के कथन से स्पष्ट है ***-''संख्यात्मक पाठन की सबसे सरल एवं सामान्य विधि बिन्दुरेख है। यह संख्याओं का चित्र इस प्रकार प्रस्तुत करती है कि नेत्रों को उनके सम्बन्ध तत्काल पता लग जाते हैं। संख्याओं को स्पष्ट बनाने में इसका सर्वाधिक महत्त्व है। (The simplest and commonest aid to the numerical reading is the graph. This gives a picture of numbers in such a way that relation between them strikes the eyes. It is of the greatest value in making numbers clear.)"***

4.3.3.1 बिन्दुरेखीय प्रदर्शन के गुण (Merits of Graphical Representation)

1- **आकर्षक व प्रभावशाली (Attractive and Impressive)** - बिन्दुरेख बहुत आकर्षक होते हैं। सुन्दर ढंग से बनाकर उन्हें और भी आकर्षक बना लिया जाता है।

2- **समझने मे सरल (Easy to Understand)** - समंकों की अव्यवस्थित और विशाल राशि बिन्दुरेख के द्वारा सरल व सुबोध बन जाती है और वह जन-सामान्य के समझने योग्य हो जाती है।

3- **समय व श्रम की बचत (Time and Labor Saving)** - इस रीति द्वारा आंकड़ों को प्रस्तुत करने में समय व श्रम अपेक्षाकृत कम लगता है। जो आंकड़ों का अध्ययन करते हैं। उनका भी समय व श्रम बचता है।

4- **तुलनात्मक अध्ययन में सरलता (Simplicity in Comparative Study)-** रेखाओं द्वारा दो प्रकार के समंकों की तुलना में बहुत सुविधा रहती है। दोनों प्रकार से समंकों की दिशा का ठीक-ठीक ज्ञान सरलता से हो जाता है और उनके तुलनात्मक अध्ययन में भी सरलता रहती है।

डिकसन हार्टबेल के अनुसार, **''चार्ट एवं बिन्दुरेखा के उदाहरण सांख्यिकीय सामग्री एवं प्रवृत्तियों की तुलना को सरलता से स्पष्ट करते हैं।''**

5- स्थायी भाव (Permanent Aspect) - संख्या सम्बन्धी सूचनाओं को प्रायः हम लोग कुछ समय के उपरान्त भूल जाते हैं क्योंकि सभी संख्याओं को याद रखना सरल नहीं। परन्तु बिन्दुरेखों का प्रभाव पर्याप्त अंशों में स्थायी होती है तथा उन्हें हम जल्दी नहीं भूलते हैं।

6- आन्तरगणन, ब्राह्मगणन व पूर्वानुमान में सुविधा (Convenience in Interpolation, Extrapolation and Forecasting)- बिन्दुरेखों की सहायता से आन्तरगणन, वाह्मगणन व पूर्वानुमान सरलता व शीघ्रता से किया जा सकता है। इनके द्वारा इन क्रियाओं के करने में बहुधा सरलता होती है। न सूत्रों का प्रयोग करना पड़ता है और न संख्या सम्बन्धी अधिक क्रियाएं ही करनी पड़ती हैं।

7- सहसम्बन्ध का अनुमान (Correlation Estimation) - बिन्दुरेखों की सहायता से सहसम्बन्ध का बहुत अंशों में अनुमान लगाया जा सकता है। वक्रों की गति इसे स्पष्ट रूप से प्रकट करती है।

8- भूयिष्ठक, मध्यका एवं चतुर्थंकों का ज्ञान (knowledge of Mode, Median and Quartiles)- बिन्दुरेखीय प्रदर्शन द्वारा भूयिष्ठक (Mode), मध्यका (Median) तथा चतुर्थंकों (Quartiles) का ज्ञान सरलता से हो जाता है।

9- ऐतिहासिक तथा कालिक सूचनाएं (Historical and Periodic Information)- ऐतिहासिक तथा कालिक सूचनाएं, जो कि आंकड़ों के द्वारा प्रकट की जाती हैं, बिन्दुरेखीय प्रदर्शन द्वारा अधिक प्रभावशाली रूप में दिखायी जा सकती हैं।

4.3.3.2 बिन्दुरेखीय प्रदर्शन के दोष/सीमाएं (Demerits/Limits of Graphical Representation)

सरल रेखा चित्रों से भी उन्हें पूर्णतः समझे बिना गलत निष्कर्ष निकाले जा सकते हैं। बिन्दुरेखीय प्रदर्शन के प्रमुख दोष निम्नलिखित हैं -

1- शुद्धता की जांच न होना (Lack of Accuracy Check)- वक्रों के द्वारा विगत वर्षों का प्रदर्शन होता है, परन्तु वास्तविक मूल्य का अनुमान नहीं हो पाता है। इसलिए शुद्धता की जांच नहीं हो पाती।

2- तर्कसंगत न होना (Not Being Logical) - बिन्दुरेखों का प्रभाव कभी-कभी आंखों तक ही रहता है।

3- दुरुपयोग सम्भव (Misuse Possible) - मापदण्ड में थोड़ा परिवर्तन कर देने पर वक्र के आकार में बहुत अन्तर पड़ जाता है, इसलिए मापदण्डों को लेकर समंकों को विभिन्न ढंगों से प्रस्तुत किया जा सकता है और इनका दुरुपयोग भी किया जा सकता है।

4- उद्धरण के रूप में प्रस्तुत न किया जाना (Not Presented as a Quote) - किसी तथ्य की पुष्टि के लिए बिन्दुरेखों को उद्धरण के रूप में प्रस्तुत नहीं किया जा सकता।

5- अपर्याप्त सूचना देना (Giving Inadequate Information) - बिन्दुरेख के द्वारा सभी सांख्यिकीय सामग्री को प्रस्तुत नहीं किया जा सकता है और न ये सभी प्रकार की समस्याओं के समाधान में सहायक हो सकते हैं इसलिए इनकी सूचनाएं अपर्याप्त होती हैं।

4.3.3.4 बिन्दुरेख बनाने के सामान्य नियम (General rules for Plotting a Graph)

बिन्दुरेखीय प्रदर्शन करते समय निम्न नियमों का पालन करना आवश्यक है:-

1- उपयुक्त व पूर्ण शीर्षक (Appropriate and Complete Title) - प्रत्येक रेखाचित्र का उपयुक्त व पूर्ण शीर्षक होना चाहिए ताकि देखते ही यह समझ में आ जाए कि वह किससे सम्बन्धित है।

2- बिन्दुरेखों की गति- बिन्दुरेखों की गति क्षैतिज पैमाने पर सामान्यतः बायें से दायें और उदग्र पैमाने पर नीचे से ऊपर होती है अतः मूलबिन्दु को यथास्थान रखना चाहिए।

3- कृत्रिम आधार रेखा (Artificial Baseline)- उदग्र मापदण्ड का चुनाव ऐसा होना चाहिए कि शून्य रेखा-पत्र पर दिखायी दे। यदि किसी कारण ऐसा करना सम्भव न हो तो मूलबिन्दु के पास शून्य रेखा से प्रारम्भ करके कुछ ऊपर जाकर इसे तोड़कर कृत्रिम आधार रेखा बना लेनी चाहिए और फिर उसके ऊपर अपनी आवश्यकतानुसार निश्चित किये हुए पैमाने के अनुसार अंकित कर लेनी चाहिए।

4-मापदण्ड का चुनाव (Choice of Criteria)- मापदण्ड का चुनाव एक महत्वपूर्ण कार्य है।

5-भुजाक्ष की लम्बाई - सामान्यतः इस बात का ध्यान रखना चाहिए कि लम्बाई में भुजाक्ष कोटि-अक्ष की डेढ़ गुनी है।

6-मापदण्ड का विवरण (Description of Criteria)- मापदण्ड का विस्तृत विवरण दिया जाना चाहिए ताकि यह सरलता से समझ में आ जाए कि आकार क्या प्रकट करता है।

7-स्पष्ट प्रदर्शन (Clear Display)- रेखाचित्रों में बिन्दुओं का प्रदर्शन स्पष्ट होना चाहिए, जो रेखा विभिन्न बिन्दुओं को मिलाये वह भी स्पष्ट था समान रूप में पतली या मोटी होना चाहिए। यदि एक ही चित्र में

कई रेखाओं का प्रदर्शन करना हो तो वहां विभिन्न प्रकार की रेखाएं खींची जा सकती है या अलग-अलग रंग व्यवहार मे प्रयोग किये जाते हैं।

8-क्षैतिज मापदण्ड या उदग्र मापदण्ड (Horizontal Scale or Vertical Scale)- क्षैतिज मापदण्ड व उदग्र मापदण्डड अलग-अलग लिए जा सकते हैं।

9-समंकों का प्राप्ति स्थान व आवश्यक टिप्पणियां (Source of Data and Necessary Notes)- जहां आवश्यकता हो वहां समंकों का प्राप्ति स्थान तथा आवश्यक टिप्पणियां भी दे देनी चाहिए ताकि उनका स्रोत ठीक से पता रहे और उनकी शुद्धता की जांच की जा सके।

10-संकेत (Sign)- यदि कुछ संकेत हैं तो उन्हें ऊपर की ओर दे देना चाहिए।

11-परिणाम कोट-अक्ष पर (Result on coat-axis)- सामान्यतः समय, स्थान, परिस्थिति, आकार आदि की इकाइयों को भुजाओं पर और समंकों के परिणाम, आश्रित चलों व आवृत्ति को कोटि-अक्ष पर प्रदर्शित करना चाहिए।

12-वक्रों के पास समंकों को देना (Giving Data Near Curves)- रेखाचित्र के साथ समंकों को पास ही सारणी में देना चाहिए ताकि यदि कभी चाहे तो विस्तृत अध्ययन कर सके या शुद्धता की जांच कर सके।

13-अनुपात माप (Ratio Measurement)- आनुपातिक श्रेणियों को प्रदर्शित करने के लिए अनुपात मापदण्ड का प्रयोग करना चाहिए।

14-धनात्मक संख्याएं (Positive Numbers)- जहां संख्याएं केवल धनात्मक हों वहां भुजाक्ष के नीचे या कोटि -अक्ष की बायीं ओर का भाग बिन्दुरेखीय-पत्र पर दिखाना व्यर्थ है।

4.3.2 चित्र तथा बिन्दुरेख

यद्यपि चित्र तथा बिन्दुरेख (या रेखाचित्र) मुख्य रूप से एक ही उद्देश्य की पूर्ति करते हैं। फिर भी दोनों में अन्तर निम्न प्रकार किया जा सकता है -

आधार	चित्र	रेखाचित्र
प्रयोग सम्बन्धी अन्तर	चित्रों का प्रयोग विशेष रूप से स्थानीय श्रेणियों के प्रदर्शन के लिए किया जाता है।	रेखाचित्रों का प्रयोग काल-श्रेणी तथा आवृत्ति वितरणों के प्रदर्शन के लिए किया जाता है।
रचना सम्बन्धी अन्तर	चित्रों की रचना सादे कागज पर की जाती है और इनका निरूपण दण्डों, आयतों, वर्गों, वृत्तों आदि द्वारा किया जाता है। इनकी रचना कठिन है।	रेखाचित्र प्रायः बिन्दुरेखीय पत्र पर बनाये जाते हैं तथा इनको बनाने के लिए विभिन्न प्रकार के बिन्दुओं या रेखाओं आदि का प्रयोग किया जाता है। इनकी रचना सरल है। इनमें श्रम तथा समय कम लगता है।
कार्य सम्बन्धी अन्तर	चित्रों की कार्य तुलनात्मक अध्ययन को सम्भव बनाना है।	बिन्दुरेखों का कार्य तुलनात्मक अध्ययन की सम्भव बनाना है।
तुलना सम्बन्धी अन्तर	चित्रों द्वारा एक ही श्रेणी के विभिन्न मूल्यों की तुलना की जाती है।	रेखाचित्र द्वारा एक से अधिक श्रेणियों में आने वाले परिवर्तन की तुलना की जा सकती है।
माध्यों का निर्धारण	चित्रों से संख्यिकीय मापों का अनुमान सम्भव नहीं है।	रेखाचित्र से मध्यका, बहुलक आदि के मूल्य ज्ञात किया जा सकते हैं।
सार्थकता सम्बन्धी अन्तर	चित्र किसी विषय की केवल सन्निकट सूचना प्रकार करते हैं। ये आंकड़ों के अर्थ में किसी प्रखार की वृद्धि नहीं करते, अतः अनुसन्धान में विवेचन के लिए उपयोगी सिद्ध नहीं होते।	बिन्दुरेख अधिक स्पष्ट, संक्षिप्त तथा शुद्ध होते हैं और अनुपात, ढाल, परिवर्तन दर आदि के अध्यय में काफी सहायक होते हैं।

4.3.3 बिन्दुरेखीय प्रदर्शन के कार्य (Function of Graphical Presentation)

1- बिन्दुरेख विश्लेषक को अनुसन्धान से सम्बन्धित सामान्य विधि, गणना तथा नियोजन में मार्गदर्शन करती है।
2- बिन्दुरेख का प्रयोग गणितीय गणना के स्थान पर समय व श्रम बचाने के लिए किया जाता है।
3- गणितीय वक्र के स्थान पर मुख्त हस्त वक्र समंकों की प्रवृत्ति के अधिक अनुरूप बनायी जा सकती है।
4- बिन्दुरेख जटिल समंकों को चित्रित करके उन्हें सरल व बोधगम्य बनाता है।
5- सम्बन्धित तथ्यों को पास-पास प्रदर्शित करके तुलना को सरल बना देता है।

4.3.3.3 बिन्दुरेख की रचना (Construction of Graph)

बिन्दुरेखों की रचना सामान्यतः बिन्दुरेखीय-पत्र पर अंकित किये जाने वाले बिन्दुओं को आपस में मिला देने से होती है। उदग्र रेखा को उदग्र माप श्रेणी या कोटि अक्ष और क्षैतिज रेखा को क्षैतिज रेखा पर क्षैतिज माप श्रेणी कहते हैं। भुजाक्ष के लिये यय ;ग्श्द्ध तथा कोटि अक्ष के लिए रर ;ल्श्द्ध संकेतों का प्रयोग प्रचलन में है।

इस प्रकार बिन्दुरेखीय-पत्र चार भागों में बंट जाता है जिनमें से प्रत्येक भाग को चरण ;फनंकतंदजद्ध कहते हैं। बिन्दुरेखीय-पत्र पर किसी भी बिन्दु को प्रांकित करते समय उदग्र एवं क्षैतिज श्रेणियों का अध्ययन करके उसे निश्चित करते हैं।

4.3.3.5 बिन्दुरेखीय वक्रों का प्रयोग (Use of Graphic Curves)

बिन्दुरेखीय वक्रों का प्रयोग दो प्रकार से किया जाता है:-

1-कालमालाओं (Time Series) के प्रदर्शन के लिए

2-आवृत्ति वितरण (Frequency Distribution) के प्रदर्शन के लिए

कालमालाओं के रेखाचित्र बिन्दुरेख (Graph of Time Series) - कालमालाओं को प्रदर्शित करने के लिए जो बिन्दुरेख बनता है, उसे कालिक चित्र कहते हैं। कालिक चित्र की रचना में समय (वर्ष, माह इत्यादि) को सदा भुजाक्ष पर तथा मूल्यों को कोटि-अक्ष पर अंकित किया जाता है। काल श्रेणी के रेखाचित्र दो प्रकार के मापदण्ड पर बनाये जा सकते हैं- i. प्राकृतिक या साधारण मापदण्ड पर या ii. आनुपातिक मापदण्ड पर।

प्राकृतिक मापदण्ड पर कालमाला चित्र-इस विधि के अन्तर्गत बिन्दुरेखीय-पत्र में मापदण्ड प्राकृतिक होता है। प्राकृतिक मापदण्ड पर भी कालमाला के रेखाचित्र निम्न दो प्रकार के हो सकते हैं-

i). निरपेक्ष कालिक चित्र (Absolute Time Frame)- जब कालिक चित्रों के लिए मौलिक समंकों को ही प्रांकित किया जाता है तो उसे निरपेक्ष कालिक चित्र कहते हैं। ये एक चर-मूल्य को या दो या अधिक चर-मूल्यों को प्रदर्शित करने के लिए बनाये जा सकते हैं।

ii). निर्देशांक कालिक चित्र (Coordinate Time Diagram)- जब वास्तविक मूल्यों या मौलिक समंकों के स्थान पर उनके निर्देशांकों अर्थात् सापेक्षिक मूल्यों को बिन्दुरेखीय-पत्र पर प्रांकित किया जाता है तो वह निर्देशांक कालिक अक्ष कहलाता है।

4.3.3.6 कृत्रिम आधार रेखा (Artificial Baseline)

बिन्दुरेख बनाते समय इस महत्वपूर्ण नियम का पालन करना आवश्यक है कि उदग्र माप पर शून्य मूलबिन्दु से प्रारम्भ किया जाये। यह नियम क्षैतिज माप के लिए आवश्यक नहीं। इस नियम के अनुसार उदग्र माप पर शून्य से प्रारम्भ करने में कभी-कभी कठिनाइयां आती हैं। जैसे यदि वे मूल्य जो उदग्र पाल पर शून्य से प्रारम्भ करने हों, बहुत बड़े हों और उनमें आपस में अन्तक कम हो तो शून्य से प्रारम्भ करने पर निम्न सुविधाएं सामने आयेंगी।

(1) वक्र आधार रेखा से बहुत दूर बनेगा और आधार रेखा के बीच का बिन्दुरेख-पत्र बेकार रहेगा।

(2) यदि मूल्य बड़े हों परन्तु इनमें होने वाले परिवर्तन बहुत कम हों अर्थात् आपसी अन्तर कम हों तो उसे भी स्पष्ट रूप में प्रदर्शित नहीं किया जा सकेगा।

इन असुविधाओं को दूर करने और बिन्दुरेख को प्रभावशाली बनाने के उद्देश्य से कृत्रिम आधार रेखा का सहारा लिया जाता है। इसमें उदग्र माप-श्रेणी का वह भाग छोड़ दिया जाता है जो मूलबिन्दु से लेकर निम्नतम मूल्य, जिसे प्रांकित करना है, तक है।

4.3.3.7 दो मापदण्डों के रेखाचित्र (Graph of Double Scales)

कहीं-कहीं कोटि-अक्ष पर दो मापदण्ड लेकर संख्याओं को प्रांकित करना पड़ता है क्योंकि वे दोनों ही विभिन्न इकाइयों को प्रकट करती हैं। उनकी इकाइयां सजातीय हों, किन्तु उनके विस्ताओं में बहुत अन्तर होता है तो भी वहां दो मापदण्डों का सहारा लेना पड़ता है। ऐसी स्थिति में बायीं ओर के कोटि-अक्ष पर एक चर-मूल्य तथा दाहिनी ओर को कोटि-अक्ष पर दूसरा चल-मूल्य दिखाया जाता है। ऐसा करने के लिए दोनो चर-मूल्यों के समान्तर माध्य एक सीध में रेखा-पत्र के साम्य मे रखकर मापदण्ड निर्धारित किये जाते हैं।

निम्न को रेखाचित्रों द्वारा प्रस्तुत कीजिए -

4.3.3.8 अधिकतम व न्यूनतम मूल्यों के रेखाचित्र: कटिबन्ध चित्र तथा वक्र (Graphs of Maximum and Minimum Values: Zone Chart and Zone Curve)

कभी-कभी किसी चर के किसी समय के अधिकतम व न्यूनतम उतार-चढ़ाव को अंकित करने की आवश्यकता पड़ती है; जैसे किसी दिन या माह में किसी वस्तु का निम्नतम व अधिकतम भाव या किसी रोगी का आधिकतम व निम्नतम तापमान। ऐसी दशा में अधिकतम मूल्यों का वक्र और न्यूनतम मूल्यों का वक्र अलग-अलग खींचकर फिर उनके बीच के स्थान को किसी रंग या चिह्न से भर देते हैं। इन्हें कटिबन्ध वक्र कहते हैं।
उदाहरण - मुम्बई में सोने की अधिकतम न न्यूनतम कीमत निम्न प्रकार है। इनमें रेखाचित्र बनाओं -

4.3.3.9 जी-रेखाचित्र या Z-वक्र (Zee or Z-Curve)

यह वक्र एक प्रकार का रेखाचित्र है जो अधिकतम व्यापारिक क्षेत्रों में प्रयोग होता है। यह वक्र अंग्रेजी के अक्षर जेड (Z) के आकार का होता है। इसीलिए इसे जी (Zee) या Z-वक्र कहते हैं। इसमें तीन वक्र तीन बातों को प्रदर्शित करते हुए खींचे जाते हैं। तीनों के लिए अलग-अलग पैमाने लिए जाते हैं। ये तीन वक्र निम्न प्रकार होते हैं :-

1-मौलिक समंकों का वक्र (Curve of Original Data)
2-संचयी समंकों का वक्र (Curve of Cumulative Data)
3-चल योगों का वक्र (Curve of Moving Totals)

4.4 चित्रों की तुलना में बिन्दुरेखों के गुण :-

चित्रों की तुलना में बिन्दुरेखों में निम्न गुण हैं -

1-लोकप्रिय (Popular) - बिन्दुरेखों का प्रयोग चित्रों की अपेक्षा अधिक होता है। यह बहुत लोकप्रिय है और लगभग सभी प्रकार के अध्यनों में प्रयुक्त होता है।

2- गणितीय प्रश्न का हल सम्भव (Possible Solution to Mathematical Question)- बिन्दुरेखों की सहायता से कई प्रकार के गणितीय प्रश्न भी हल किये जा सकते हैं, इसलिए गणित की दृष्टि से ये चित्रों की अपेक्षा अधिक महत्वपूर्ण हैं।

3-भूयिष्टक, चतुर्थक आदि निकालना सम्भव - बिन्दुरेखों की सहायता से भूयिष्ठक, चतुर्थक, दशमक, शतमक आदि निकाले जा सकते हैं।

4-सबके लिए लाभप्रद (Profitable for Everyone)- बिन्दुरेख की रचना करने वाला स्वयं अपने लाभ के लिए भी उनकी रचना कर सकता है क्योंकि किसी भी अध्ययन के लिये ये बड़े लाभप्रद होते हैं। परन्तु चित्र सामान्यतः दूसरों के लिए बनाये जाते हैं।

5-समय-श्रेणी का अच्छा प्रदर्शन (Good Performance of Time Series)- समय श्रेणी या काल-माला के प्रदर्शन के लिए बिन्दुरेख बहुत आवश्यक है ताकि परिवर्तन को ठीक प्रकार से देखा जा सके। चित्रों की सहायता से यह उतना सम्भव नहीं है।

4.5 बिन्दुरेख की तुलना में चित्रों के गुण (Properties of Pictures Compared to Graphs) -

बिन्दुरेख की तुलना मे चित्रों में निम्न विशेष गुण होते हैं :

(अ) समझने में सरल (Simple to Understand)- चित्र बिन्दुरेखों की अपेक्षा समझने मे अधिक सरल होते हैं। देखते ही वे समझ में आ जाते हैं।

(ब) स्थायी प्रभाव (Lasting Impact) - चित्रों का प्रभाव मस्तिष्क पर बिन्दुरेखों की अपेक्षा अधिक स्थायी होता है।

(स) आकर्षण तत्व - चित्रों में आकर्षण तत्व अधिक होता है क्योंकि ये कई आकृतियों मे तथा कई रंगों या चिन्हों की सहायता से बनाये जाते हैं।

4.6 आंकड़ों का प्रदर्शन (ग्राफीय)

पिछले खण्ड के अध्ययन के पश्चात् आप इस बात को भली-भाँति समझ चुके होंगे कि 'जैसे-जैसे संख्याओं की किसी भी सूची की लम्बाई बढ़ती जाती है, वैसे-वैसे वह कम ग्राह्य होती जाती है। इसी प्रकार का विचार प्रो0 किंग ने भी दिया।'' संख्याओं के अपने सबसे अच्छे रूप में भी तुलना करने के उद्देश्य से मस्तिष्क के लिए समझना और देर तक भाव रखना सरल बात नहीं है।

अतः यदि हम अपनी बातों को अंकों द्वारा प्रस्तुत करने के बजाय किसी अन्य सरल साधन द्वारा प्रस्तुत करें जहां अंकों का प्रयोग कम से कम किया जाए तो हमारी बात जन-साधारण के लिए सरल, समझने तथा याद करने योग्य हो जाती है।

4.6 चित्र एवं ग्राफ

''अधिकांश व्यक्तियों के लिए नीरस संख्याएं प्ररेणा शून्य होती हैं। चित्र किसी जटिल स्थिति के स्वरूप को देखने समझने में हमारी सहयता करते हैं। जिस प्रकार एक मानचित्र हमारे सामने किसी विशाल देश का

विहंगम दृश्य प्रस्तुत करता है, ठीक उसी प्रकार चित्र एक दृष्टि में संख्यात्मक जटिल तथ्यों का सम्पूर्ण अर्थ समझने में हमारी सहयता करते हैं।'' - **मोरोने**

जटिल आंकड़ों को इस प्रकार प्रस्तुत किया जाये कि वे देखने में सुन्दर तथा समझने में बहुत सुन्दर बन जायें। वर्गीकरण व सारणीयन इसी उद्देश्य को लेकर किये जाते हैं।

सांख्यिकी में समंकों को प्रदर्शित करने की दो विधियों हैं -

(अ) चित्रमय प्रदर्शन (ब) बिन्दुरेखीय प्रदर्शन

4.6.1 चित्रों की उपयोगिता एवं लाभ -

1- चित्र समंकों को सरल व सुबोध बनाते हैं - चित्रों के द्वारा जटिल, अव्यवस्थित और विशाल समंक राशि सरल हो जाती है और वह जनसाधारण के समझने के योग्य हो जाती है। केवल अंकों को देखकर कोई निष्कर्ष निकालना कठिन होता है।

प्रसिद्ध विद्वान **प्रो. स्टीफेन जुल्क** के शब्दों में, ''एक चित्र अधिक स्पष्ट तथा चित्र को सीध्ो आकर्षित करने वली तस्वीर प्रदान करता है।''

2-अधिक समय तक स्मरणीय - अंकों को बहुत समय तक याद करना अत्यन्त कठिन है। कुछ समय बाद मनुष्य अंको को भूल जाता है, परन्तु चित्रों द्वारा आंकड़ों की एक अमिट छाप मस्तिष्क पर पड़ती है जो बहुत दिनों तक बनी रहती है।

3- चित्रों को समझने के लिए विशेष शिक्षा या ज्ञान की आवश्यकता नहीं - चित्रों की समझना जनसामान्य के लिए बहुत सरल है।

4.समय या श्रम की बचत - चित्रों की सहायता से आंकड़ों के समझने व उनसे निष्कर्ष निकालने में बहुत कम समय तथा परिश्रम की आवश्यकता होती है। बिना किसी परिश्रम के बड़ी सरलता से और एक दृष्टि में आंकड़ों पर्याप्त मात्रा में समझ में आ जाते हैं।

5-आकर्षक एवं प्रभावशाली - चित्र बहुत आकर्षक होते हैं। ये बरबस ध्यान अपनी ओर खींच ल्ोते हैं। सी. डब्लू लोव के शब्दों में ''सभी प्रकार के समंकों को अत्यन्त प्रभावशाली रूप में निरूपण करने के लिए चित्रों का प्रयोग किया जा सकता है।''

6-सूचना के साथ-साथ मनोरंजन होना - सुन्दर चित्र सूचना तो देते ही हैं, परन्तु साथ ही साथ मनोरंजन भी देते हैं। इनकी सहायता से विभिन्न सूचनाओं के अध्ययन में थकावट प्रतीत नहीं होती है।

7-तुलना करने में सहायक - चित्रों की सहायता से विभिन्न सूचनाओं की प्रभावशाली तुलना की जा सकती है।

8- समंको का संक्षिप्त रूप - चित्र समंको को संक्षिप्तता प्रदान करते हैं। इनका निर्माण समस्या पर विचार करने एवं पर्याप्त विश्लेषण करने के उपरान्त होता है।

चित्रों द्वारा प्रदर्शन की परिसीमाएं अथवा हानियां

चित्रमय प्रदर्शन उन व्यक्तियों के लिए भ्रमात्मक होते हैं जो सावधानीपूर्ण अध्ययन के बिना ही उनसे निष्कर्ष निकालते हैं। **एम. जे. मोरोने** के अनुसार ***'किसी चित्र का अध्ययन कनरे के लिए पर्याप्त चौकन्ना रहना आवश्यक होता है। वह इतना सरल, स्पष्ट तथा मनभावी होता है कि असावधान व्यक्ति बड़ी आसानी से मूर्ख बन जाता है।'*** इस तकनीकी का प्रयोग करते समय तथा निष्कर्ष निकालते समय विशेष सावधानी की आवश्यकता है। इन चित्रों की परिसीमाएं निम्नलिखित हैं -

1-तुलना के लिए गुण व स्वभाव की समान आवश्यकता-चित्रों में तुलना तभी ठीक होगी जब वे समान गुण के आधार पर बनाये जाएं। यदि वे दो विभिन्न गुणों के आधार पर बनाये जायें तो उनमें तुलना करना भ्रामक व अशुद्ध होगा।

2-केवल तुलनात्मक अध्यय सम्भव - चित्रों की सहायता से केवल तुलमात्मक अध्ययन सम्भव हो पाता है।

3-सूक्ष्म अन्तर दिखाना सम्भव नहीं- चित्रों द्वारा बहुत सूक्ष्म अन्तर को प्रदर्शित करना सम्भव नहीं है।

4-बहुमुखी सूचनाओं का प्रदर्शन सम्भव नहीं - चित्रों द्वारा बहुमुखी विशेषताओं को प्रदर्शित नहीं किया जा सकता।

5-संख्यात्मक प्रदर्शन असम्भव - चित्रों द्वारा आंकड़ों का पूर्ण शुद्ध रूप में प्रदर्शन सम्भव नहीं होता है। चित्र अनुमानित रूप से आंकड़ों का प्रदर्शन करते हैं। चित्र वहीं के लिए उपयुक्त होते हैं जहाँ संख्या में मूल्य प्राप्त करना उद्देश्य न हो बल्कि उनके मूल्य का अनुमान चित्रों को देखर लगाया जा सके।

6-सरलतापूर्वक दुरुपयोग-अनुचित और अशुद्ध चित्र बनाकर उनका दुरुपयोग किया जा सकता है।

7-निष्कर्ष निकालने का एक उचित साधन- चित्रों को देखर पूर्ण सत्य निष्कर्ष निकाला जाना सम्भव नहीं है।

8-पर्याप्त ज्ञान के अभाव में भ्रमोत्पादक-यदि चित्र बनाने वाले को विषय का पर्याप्त ज्ञान नहीं है तो चित्र ऐसा बन सकता है जो वस्तुस्थिति का ठीक ज्ञान न करा सके और भ्रम पैदा करे।

9-आंखों का धोखा-यदि आंकड़ों के अनुरूप चित्र न बनाये जायें तो इससे आंखों को भारी धोखा हो सकरता है और देखने वाला गलत परिणाम पर पहुंच सकता है।

10-आगे विश्लेषण असम्भव-चित्रों की एक यह भी सीमा है कि इनका आगे विश्लेषण सम्भव नहीं है।

4.6.2 चित्र बनाने का सामान्य नियम

सांख्यिकीय चित्रों को आकर्षक व प्रभावशाली बनाने के लिए निम्न बातों को ध्यान मे रखना आवश्यक है :

1-शुद्धता - चित्र आकर्षक व कलात्मक हों, शुद्धता उनकी जान है। चाहे कितना भी आकर्षक चित्र क्यों न हो, यदि शुद्धता नहीं तो वह व्यर्थ है।

2-आकर्षक-चित्रों को आकर्षक बनाना सबसे अधिक आवश्यक है।

3-रेखापत्र का प्रयोग -चित्र बनाते समय रेखापत्र का प्रयोग ठीक रहता है। इससे सुन्दरता व शुद्धता दोनों की रक्षा हो जाती है।

4-शीर्षक-यह अत्यन्त आवश्यक है कि चित्र के ऊपर उसकी संख्या व शीर्षक दिया जाए।

5-आकार-चित्र का आकार प्राप्त स्थान के अनुसार होना चाहिए ताकि वह देखने में सुन्दर लगे।

6-मापदण्ड-चित्र बनाने से पहल्ो मापदण्ड निश्चित कर ल्ोना आवश्यक होता है। मापदण्ड निश्चित करते समय प्राप्त स्थन व अंकित करने वाली सूचना दोनों को ध्यान में रखा जाता है। मापदण्ड ऐसा होना चाहिए कि चित्र स्थान को ध्यान में रखते हुए न तो बहुत बड़े बन जायें और न बहुत छोटे रहें।

7- चिह्नों या रंगों का प्रयोग-चित्रों में आवश्यकतानुसार विभिन्न प्रकार की सूचनाओं को प्रदर्शित करने के लिए विभिन्न प्रकार की चिह्नों व रंगों का प्रयोग करना चाहिए और उनके विषय में संकेत चित्र के नीचे बायें कोने पर दे देना चाहिए।

8-चित्रों को घ्ोरना-चित्रों को मीटी या दोहरी रेखाओं से घ्ोर देना चाहिए ताकि वे देखने में अधिक आकर्षक लगें।

9-उपयुक्त चित्र का चुनाव-चित्र कई प्रकार के होते हैं ओर इस प्रकार की चित्र सभी प्रकार के समंकों के लिए उपयुक्त नहीं हो सकते।

10-सरलता-चित्र ऐसा होना चाहिए कि वह सरलता से एक बार देखने से समझ में आ जाय।

11-महत्वपूर्ण बातें गहरे रंग से-चित्र बनाते समय यह भी ध्यान मे रखना आवश्यक है कि आंकड़ों के महत्वपूर्ण अंशों को गहरे रंग से या ऐसे ढंग से प्रदर्शित किया जाए।

12-मितव्ययिता-यह भी ध्यान मे रखना आवश्यक है कि चित्र बनाते समय धन व अन्य साधनों का दुरुपयोग न हो।

4.6.3 चित्रों के प्रकार

सांख्यिकीय में साधारणतः निम्न प्रकार के चित्रों का प्रयोग किया जाता हैः

1-एक-विमा या एक-विस्तार वाले चित्र,
2- द्विविमा या दो-विस्तार वाले चित्र,
3- त्रिविमा या तीन-विस्तार वाले चित्र,
4-मानचित्र,
5-चित्र-लेख

1-एक-विमा या एक-विस्तार वाले चित्र-जब पदमाला विच्छिन्न रहती है और केवल गुण की तुलना करनी होती है तो एक विमा या एक-विस्तार वाले चित्रों का रचना की जाती है। एक विमा चित्र निम्न प्रकार के होते हैं -(क) रेखाचित्र (ख) दण्ड चित्र

(क) रेखाचित्र-इन रखाओं की रचना विभन्न पदों के मूल्यों के अनुसार होती हैं। रेखाचित्रों की रचना तब की जाती है जबकि सम्बन्दित पद-मूल्यों की संख्या अधिक हों तथा न्यूनतम व अधिकतम मूल्यों का अनुपात कम हो।

(ख) दण्ड चित्र - दण्ड चित्र एवं रेखा चित्र में बहुत साधारण अन्तर होता है और यह कि यहां रेखाओं को मोटा बना देते हैं। मोटा बनाते समय मूल्य का कोई ध्यान नहीं रखा जाता है।

विभिन्न प्रकार के दण्ड चित्र

1. सरल दण्ड चित्र-ये दो प्रकार के होते हैं

(क) उदग्र दण्ड, (ख) क्षैतिज दण्ड

(क) उदग्र दण्ड-जब दण्ड सीधे खड़े बनाये जाते हैं तो उदग्र कहलाते हैं। यथा सम्भव यह प्रयास होना चाहिए कि सबसे ऊंचा दण्ड वायें और ऊँचाई के क्रम में अन्य दण्ड बनाते हुए सबसे छोटा दण्ड दांये बनाना चाहिए। यह क्रम विपरीत भी हो सकता है। परन्तु जब समंक समय या किसी अन्य महत्वपूर्ण क्रम में दिये हों तब छोटे या बड़े का विचार किये बिना दण्ड इसी क्रम में बनाया जाना चाहिए।

4.7 आंकड़ों का विश्लेषण

आंकड़ों का विश्लेषण दो प्रकार से किया जाता है -

1- आरेखों द्वारा
2- सांख्यिकीय गणना के द्वारा

आरेखात्मक विश्लेषण के अन्तर्गतहम आंकड़ों को क्रमबद्ध रूप से व्यवस्थित करने के बाद उन्हें उपयुक्त आरेखों के द्वारा प्रदर्शित करते हैं। तत्पश्चात आरेखो के आधार पर आंकड़ों से सम्बन्धित निष्कर्ष ज्ञात किये जाते हैं। गणनात्मक विश्लेषण के अन्तर्गत हम सांख्यिकीय गुणांकों के मान ज्ञात करते हैं। तथा गुणांकों के मानों के आधार पर आंकड़ों से संबन्धित निष्कर्ष निकालते हैं। ये सांख्यिकीय गुणांक हैं-माध्य, आंकड़ों का क्षितराव, आवृत्तियों का संकेन्द्रण इत्यादि।

सबसे पहल्ो हम आंकड़ों के आरेखात्मक विश्लेषण की चर्चा करेंगे। परन्तु इससे पूर्व आंकड़ों तथा चरराशिओं से सम्बन्धित कुछ संकल्पनाओं की व्याख्या करना आवश्यक है। आंकड़ों को प्रायः निम्न श्रेणियों में विभक्त किया जाता है-

i) मात्रात्मक तथा गुणात्मक आंकड़े

ii) काल श्रृंखला आंकड़े तथा अनुप्रस्थ आंकड़े

मात्रात्मक तथा गुणात्मक आंकड़ें की चर्चा हम पहल्ो भी कर चुके हैं।

काल-श्रृंखला आंकड़ें हम उन आंकड़ों को कहते हैं जिनका प्रत्येक मान अलग-अलग समय बिन्दु से अथवा अलग-अलग माह, वर्ष अथवा समयावधियों से सम्बन्धित हों, जैसे-किसी देश की जनसंख्या के आंकड़े, किसी फैक्ट्री के उत्पादन के वार्षिक आंकड़े, किसी अस्पताल में उपचार के लिये आने वाले रोगियों की साप्ताहिक संख्या, किसी नगर में घण्टेवार तापमान के आंकड़े इत्यादि-इत्यादि।

अनुप्रस्थ आंकड़े - ऐसे आंकड़ों को कहते हैं, जिनका प्रत्येक मान एक ही समय बिन्दु एक ही तिथि, एक ही माह, वर्ष अथवा एक ही समाधावधि से सम्बन्धित हों. अनुप्रस्थ आंकड़ें के उदाहरण हैं- सिविल सेवा प्रारम्भिक परीक्षा में सम्मिलित परीक्षार्थियों के प्राप्तांकों के आंकड़े, सिविल सेवा परीक्षा में सम्मिलित परीक्षार्थियों की आयु के आंकड़े, व्यक्तियों की ऊंचाई, वजन, आय सम्बन्धी आंकड़े।

स्पष्ट है कि सिविल सेवा प्रारम्भिक परीक्षा, 29 सितम्बर 1991 को आयोजित की गई थी, तथा इस परीक्षा में सम्मिलित विद्यार्थियों के प्राप्तांकों के सभी आंकड़े इसी तिथि से सम्बन्धित होंगे। अतः इन आंकड़ों को हम अनुप्रस्थ आंकड़ें कहेंगे।

इसी प्रकार हम जानते हैं कि सिविल सेवा परीक्षा में सम्मिलित होने वाले परीक्षार्थियों की अधिकतम आयु सीमा 28 वर्ष है, (परीक्षा, वर्ष की तिथि 1 अगस्त को) तथा परीक्षा में बैठने के लिये आवेदन करते समय प्रत्येक परीक्षार्थी प्रवेश फार्म में वर्ष के 1 अगस्त को अपनी आयु का विवरण (वर्ष, माह तथा दिनों से) देता है। अन्य शब्दों में संघ लोक सेवा आयोग के पास परीक्षार्थियों की आयु के सभी आंकड़ों एक ही तिथि (परीक्षा वर्ष के 1 अगस्त) से सम्बन्धित होंगे। अर्थात् आयु के इन आंकड़ों को हम अनुप्रस्थ आंकड़े कहेंगे। इसी प्रकार आयकर विभाग के पास लोगों की आय के आंकड़ें एक वित्तीय वर्ष विशेष (विशिष्ट वित्तीय वर्ष में 1 अप्रैल से परवर्ती 31 मार्च की समयावधि) से सम्बन्धित हैं, इसलिए ये अनुप्रस्थ आंकड़े होंगे। इसी प्रकार लोगों की ऊंचाई तथा वचन के आंकड़े भी अनुप्रस्थ आंकड़ों की श्रेणी मे आते हैं।

काल श्रृंखला तथा आरेखीय प्रदर्शन -

काल श्रृंखला आंकड़ों को तीन प्रकार के आरेखों के द्वारा प्रदर्शित किया जाता है-

i) बिन्दु रेखीय अथवा रेखाचित्रीय आरेख

ii) दण्ड चित्र

iii) प्रतीक चित्र

1 रेखाचित्रीय आरेख - उपर्युक्त जनंख्या आंकड़ों का रेखाचित्रीय आरेख बनाने के लिये सर्वप्रथम हम चरराशियों के परिसर को ज्ञात करते हैं। आंकड़ें को देखने से स्पष्ट है कि स्वतंत्र चरराशि समय के मान 1930 से 1980 के मध्य तथा आश्रित चरराशि जनसंख्या में मान 20 करोड़ (चरराशि का न्यूनतम मन) तथा 54.96 करोड़ (चरराशि का अधिकतम) मान के मध्य हैं।

अब एक ग्राफ पेपर (सेन्टीमीटर अथवा इंच ग्राफ पेपर) पर हम सबसे पहल्ो अक्षों को अंकित करते हैं। ग्-अक्ष पर हम स्वतंत्र चरराशि समय (वर्षों में) के मानों को तथा इसी प्रकार ल्-अक्ष पर आश्रित चरराशि, जनसंख्या (करोड़ों में) के मानों को इनके दिये हुए परिसरों के अनुसार अंकित करते हैं।

स्पष्ट है कि समय का न्यूनतम मान, 1930 चूंकि शून्य से अत्यधिक दूरी पर है,. अतः समय के न्यूनतम मान को हम मूल बिन्दु पर ही प्रदर्शित करते हैं। इस प्रक्रिया से ग्राफ पेपर का एक बड़ा भाग बर्बाद होने से बच जाता है। जनसंख्या के मानों में न्यूनतम मान (20 करोड़) चूंकि शून्य से अधिक दूरी पर नहीं है, अतः ल्-अक्ष पर जनसंख्याके मानों को शून्य से लगभग 60 करोड़ तक अंकित किया गया है।

अब आंकड़ों में दिये हुए वर्षों एवं सम्बन्धित जनसंख्या के आंकड़ों को ग्राफ बिन्दुओं के रूप में अंकित किया जाता है। प्राप्त बिन्दुओं को क्रम से सीधी रेखाओं द्वारा जोड़ने पर टेढ़ी-मेढ़ी रेखाओं का जो आरेख प्रापत होता है उसे रेखाचित्रीय आरेख अथवा बिन्दुरेखीय आरेख कहते हैं।

रेखाचित्रीय आरेख हमें तीन तथ्यों की जानकारी देते हैं - (1) आंकड़ों की सामान्य अथवा दीर्घकालीन प्रवृत्ति, (2) किन्हीं दो बिन्दुओं के मध्य रेखा चित्रीय आरेख के रेखीय खण्ड जनसंख्या के वार्षिक परिवर्तन की दर को प्रदर्शित करते हैं। (3) रेखीचित्रीय आरेख चरराशि के अनुमान अथवा पूर्वानुमान त्रात करने में भी सहायक होते हैं।

चित्र से स्पष्ट है कि समयोपरि जनसंख्या के मानों का आरेख दाहिनी ओर उठता हुआ है, अतः जनसंख्या की सामान्य (अथवा दीर्घकालीन) प्रवृत्ति बढ़ने की है। आरेख उच्चावचनों अथवा अत्यधिक उतार-चढ़ाव से रहित है, अन्य शब्दों में समयावधि 1930 से 1980 के मध्य देश की जनसंख्या स्थाई दर से बढ़ती रही है। विभिन्न दशकों में जनसंख्या की वार्षिक संवृद्धि दर (अर्थात् जनसंख्या के मानों में प्रतिवर्ष बढ़ोत्तरी) ज्ञात करने के लिये हम रखा चित्रीय आरेख के विभिन्न रेखीय खण्डों (अर्थात् आरेख पर बिन्दुओं को जोड़ने वाली सीधी रेखाओं) के मध्य जनसंख्या के वार्षिक परिवर्तनों को ज्ञात करते हैं। इनका सूत्र निम्न प्रकार है-
समयावधि विशेष में जनसंख्या का वार्षिक परिवर्तन =
अब हम समयावधि 1930 - 1940 के मध्य जनसंख्या के औसत वार्षिक परिवर्तन को ज्ञात कर सकते हैं। आरेख से -
वर्ष 1930 की जनसंख्या = 20 करोड़
समयावधि में जनसंख्या का परिवर्तन = 4.9 करोड़
समयावधि (1930-1940) के मध्य वर्षों की संख्या= 4.9/10 = 0.49 करोड़ प्रतिवर्ष
अर्थात् समाधावधि 1930-1940 मे देश की जनसंख्या औसतन 0.49 करोड़ अथवा 29 लाख प्रतिवर्ष बढ़ती रही है। इसी प्रकार अन्य समयावधियों में हम जनसंख्या के औसत वार्षिक परिवर्तन की दरों को भी ज्ञात कर सकते हैं। इनकी गणना को निम्न सारिणी में प्रदर्शित किया गया है -
वर्ष जनसंख्या
(करोड़ों में) समयावधि समयावधि में जनसंख्या परि. (करोड़ों में) वर्षों की संख्या औसत वार्षिक परिवर्तन
(करोड़ों में)
1930
1940
1950
1960
1970
1980 20
24.9
35.1
38
45
54.9
1930-40
1940-50
1950-60
1960-70
1970-80
4.9
10.2
2.9
7.0
9.9
10
10
10
10
10

0.49
1.02
0.29
0.70
0.99

गणितीय अवधारणाओं के परिचित विद्यार्थी समझ गए होंगे कि विभिन्न समयावधियों में जनसंख्या की निरपेक्ष दर वास्तव में रेखाचित्रीय आरेख के संगत रखीय खण्डों के ढाल के बराबर है तथा अधिक तिरछी रेखाओं का ढाल अधिक तथा कम तिरछी रेखाओं का ढाल कम होता है। आरेख मं समयावधि 1930-40 से सम्बन्धित रेखीय खण्ड, 1940-50 समयावधि से सम्बन्धित रेखीय खण्ड की तुलना में कम तिरछा ह (अथवा अपेक्षाकृत समानान्तर है) तथा जनसंख्या परिवर्तन की निम्न वार्षिक वृद्धि (49 लाख प्रतिवर्ष) को प्रदर्शित करता है। समयावधि 1940-50 मं जनसंख्या की वार्षिक वृद्धि 1.02 करोड़ प्रतिवर्ष है।

2 दण्ड चित्र - दण्ड चित्र का तात्पर्य है - आंकड़ों का दण्डों के द्वारा प्रदर्शन। आंकड़ों को प्रदर्शित करने की यह सर्वाधिक प्रचलित विधि है।

वास्तव में सांख्यिकीय आरेखों का रचना का मुख्य उद्देश्य आम व्ति को सांख्यिकीय तथ्यों की जानकारी देना होता है। आरेखों का प्रभाव मस्तिष्क पर अधिक स्थाई तथा प्रभावपूर्ण होता है, तथा ये सांख्यिकीय तथ्यों की सहज व्याख्या करने में सुलभ होते हैं।

रेखाचित्रीय आरेख की भाँति ही यहां भी हम स्वतन्त्र चरराशि के मानों (वर्षों को) ग्-अक्ष पर तथा आश्रित चरराशि, जनसंख्या के मानो को ल्-अक्ष पर दर्शाते हैं। अब प्रत्येक समय बिन्दु (वर्ष) पर हम सम्बन्धित जनसंख्या के बराबर ऊंचाई के दण्ड निर्मित करते हैं। सभी दण्डों की चौड़ाइयां परस्पर समान होती हैं तथा विभिन्न दण्डों के बीच लगभग दण्डों की चौड़ा के बराबर दूरी छोड़ी जाती है। जनसंख्या आंकड़ों से सम्बन्धित दण्ड चित्र को चित्र द्वारा प्रदर्शित किया गया है। दण्ड चित्र आंकड़ों की आवृत्ति के अतिरिक्त और कोई महत्वपूर्ण जानकारी नहीं देता है - न तो यह विभिन्न समयावधि में चरराशि के परिवर्तन की दर को दर्शाता है, तथा नहीं दण्ड चित्र का प्रयोग आंकड़ों के अनुमान अथवा पूर्वानुमान ज्ञात करने के लिये किया जा सकता है। दण्ड चित्र का महत्व दृष्टिगत अधिक है, विश्लेषण के दृष्टिकोण से इस आरेख की उपयोगिता सीमित है।

यहां इस तथ्य की ओर संकेत करना आवश्यक है कि यहां सारिणी में आंकड़ों के वास्तविक मानों को ही दर्शाया जाता है, वहाँ चित्रों अथवा आरेखों में केवल उनके सन्निकट मानों को ही दर्शाया जा सकता है।

3 प्रतीक चित्र - प्रतीक चित्र का अर्थ है, आंकड़ों की प्रतीकों अथवा चिह्नों के द्वारा प्रदर्शित करना है। इस विधि के अन्तर्गत विभिन्न समय बिन्दुओं पर चरराशि के परिमाण को प्रतीकों अथवा आकृतियों का संख्या के द्वारा प्रदर्शित किया जाता है।

इस प्रस्तुतीकरण का उद्देश्य सांख्यिकीय आंकड़ों के मुख्य तथ्यों को सहज तथा सरलतम रूप में व्यक्त करना होता है। आरेख बनाने के लिये सर्वप्रथम हम आंकड़ों से सम्बन्धित एक उपयुक्त प्रतीक का चुनाव करते हैं, तत्पश्चात् आंकड़ों के मानं को प्रतीकों की संख्या के द्वारा प्रदर्शित करते हैं। जैसे-मानव जनसंख्या को प्रदर्शित करने के लिये सर्वाधित उपयुक्त प्रतीक है-मनुष्य की आकृति।

इस प्रकार के आरेखों का प्रयोग समाचार-पत्र पत्रिकाओं तथा टी.वी. परिचर्चाओं में प्रचुर रूप से किया जाता है। इस आरेख में बहुधा तकनीकों का प्रयोग भी किया जाता है, जैसे चरराशि के मानों को गड्डियों की लम्बाई अथवा संख्या के द्वारा भी प्रदर्शित किया जाता है।

यदि समय अन्तराल के आंकड़ों में एक चरराशि के मानों को दर्शाया गया हो तो इस प्रकार के आंकड़ों को एक चरीय कालश्रृंखला आंकड़े कहा जाता है। यदि समयोपरि दो चरराशियों के मानों से सम्बन्धित आंकड़ो दिये हुए हों तो आंकड़ों को द्विचरीय काल श्रृंखला आंकड़े कहते हैं। द्विचरीय काल-श्रृंखला आंकड़ों को मुख्यतः निम्न आरेखों के द्वारा प्रदर्शित किया जाता है -

i) रेखाचित्रीय आरेख
ii) बहुदण्ड चित्र
iii) संघटक-भाग दण्ड चित्र
iv) प्रतिशत संघटक भाग दण्ड चित्र

i) रेखाचित्रीय आरेख - आंकड़ों से सम्बन्धित रेखाचित्रीय आरेख निम्न प्रकार है- आरेख में ग्.अक्ष पर स्वतंत्र चरराशि समय (वर्षों में) को तथा ल्-अक्ष पर सभी प्रकार के वाहनों की संख्या (लाखो में) दर्शाया गया है। चारों चरराशियों के आरेखों को अलग-अलग प्रकार की रेखाओं के द्वारा प्रदर्शित किया गया है तथा इनके सूचक को ग्राफ पर दाहिनी ओर ऊपर की दिशा में अंकित किया गया है।

आरेख में स्कूटर एवं वाहनों के कुल उत्पादन की दीर्घकालीन प्रवृत्ति बढ़ने की है, जबकि अन्य वाहनों का उत्पादन समयोपरि लगभग स्थिर बना हुआ है। स्कूटर तथा मोपेड के आरेख निरन्तर मोटर साइकिल के आरेख के ऊपर हैं, जो इस तथ्य को प्रदर्शित करते हैं। कि मोटर साइकिल की तुलना में इन वाहनों का उत्पादन सदैव अधिक रहा है, जो कि जनसाधारण में इन वाहनों की लोकप्रियता को प्रदर्शित करता है। लोकप्रियता की दृष्टि से स्कूटर का स्थान श्रेष्ठ है तथा वर्ष 81-82 से स्कूटर का आरेख ऊपर की दिशा में तिरछा होता गया है जो इस बात का द्योतक है कि वर्ष 83-84 से स्कूटर का वार्षिक दर में तीव्र वृद्धि हुई है।

इसी प्रकार के और भी बहुत सारे तथ्य रेखाचित्रीय आरेख के आधार पर ज्ञात किये जा सकते हैं, इसके अतिरिक्त इस आरेख की चरराशियों के मानों के अनुसार मान पूर्वानुमान ज्ञात करने के लिये भी प्रयुक्त किया जा सकता है।

ii) बहुदण्ड चित्र - त्रिचरीय काल-श्रृंखला आंकड़ों को दण्ड आरेख के द्वारा प्रदर्शित करने के लिये हम प्रत्येक समय बिन्दु तक तीन दण्ड (प्रत्येक चरराशि के मान के लिये एक दण्ड) खींचते हैं। दण्डों की ऊंचाइयाँ सम्बन्धित चरराशियों के मानों को प्रदर्शित करती हैं।

चूँकि प्रत्येक समयबिन्दु पर आंकड़ों को एक से अधिक दण्ड के द्वारा प्रदर्शित किया गया है, इसलिये इस आरेख को बहुदण्ड चित्र कहा जाता है।

बहुदण्ड चित्र चरराशियों की दीर्घकालीन प्रवृत्ति के अलावा अन्य कोई जानकारी नहीं हेता, अतः विश्लेषण की दृष्टि से रेखाचित्रीय आरेख बहुदण्ड से श्रेष्ठ है। बहुदण्ड आरेख का महत्व केवल दृष्टिगत है तथा इसे मुख्यतः जनसाधारण को सांख्यिकीय जानकारी देने के उद्देश्य से ही निर्मित किया जाता है।

iii) संघटक-भाग दण्ड चित्र - संघटक भाग दण्ड चित्र से सभी तथ्यों की जानकारी हमें एक ही आरेख से प्राप्त हो जाती है। वास्तव में संघटक भाग दण्ड चित्र यह प्रदर्शित करता है कि समय राशि, समयोपरि अपने संघटकों में किस प्रकार विभाजित हो रही है।

आरेख खींचने के लिये सबसे पहल्ो हम विभिन्न समयों में समग्र राशि अथवा वस्तु का कुल उत्पादन (यदि यह पहल्ो से न दिया हो) ज्ञात करते हैं। तत्पश्चात् समग्र राशि को प्रदर्शित करने वाला एक सरल दण्ड चित्र बनाते हैं, इसके उपरान्त प्रत्येक दण्ड को (कुल अथवा समग्र राशि को) तीन भागों अथवा संघटकों में विभाजित करते हैं। दण्डों के ये संघटक विभिन्न चरराशियों के समयोपरि मानों को प्रदर्शित करते हैं। चित्र निम्न प्रकार है।

iv) प्रतिशत संघटक भाग दण्ड चित्र - संघटक भाग चित्र हमें निरपेक्ष मानों के बारे में इनकी मूल प्रवृत्तियों के बारे में तो जानकारी देता है परन्तु कुल लक्ष्य में संघटक राशियों के अनुपातों के विषय में कोई जानकारी नहीं देता। अनुपातों से सम्बिन्धित जानकारी प्राप्त करने के लिये निर्मित किये जाने वाले आरेख को हम प्रतिशत संघटक भाग दण्ड चित्र कहते हैं।

प्रतिशत संघटक भाग दण्ड चित्र तथा संघटक भाग दण्ड चित्र में केवल इतना अन्तर है कि संघटक भाग दण्ड चित्र जहां निरपेक्ष मानों के परिवर्तन की व्याख्या करता है, वहीं प्रतिशत संघटक भाग दण्ड चित्र संगत प्रतिशतों की व्याख्या करता है। प्रतिशतों को ज्ञात करने के पश्चात्, इन्हें एक संघटक भाग दण्ड चित्र के द्वारा प्रदर्शित करते हैं। यह आरेख चूंकि प्रतिशतों की व्याख्या करता है, अतः इसे हम प्रतिशत संघटक भाग दण्ड चित्र कहते हैं।

4 पाई आरेख अथवा वृत्त चित्र

पाई आरेख अथवा वृत्त चित्र वास्तव में यह दर्शाता है कि एक समग्र राशि अथवा कुल राशि (जिसे 100 प्रतिशत के बराबर माल लिया जाता है) विभिन्न संघटक राशियों में किस प्रकार वर्गीकृत अथवा विभाजित है। अन्य शब्दों में यह समग्र राशि (100 प्रतिशत) के संघटक राशियों में सापेक्षिक अथवा प्रतिशत वितरण की व्याख्या करता है।

इस आरेख के अन्तर्गत समग्र अथवा कुल राशि को एक वृत्त के क्ष्ोत्रफल के द्वारा तथा संघटक राशियों को वृत्तांओं के द्वारा प्रदर्शित किया जाता है। आरेख में बड़े वृत्तांश, बड़े प्रतिशतों (अनुपातों को) तथा छोटे वृत्तांश छोटे प्रतिशतों को प्रदर्शित करते हैं। वृत्तांशों का आकार अथवा क्ष्ोत्रफल वृत्तांश कोण के ऊपर निर्भर करता है - यदि आरेख में किसी संघटक राशि से सम्बन्धित वृत्तांश का वृत्तांश कोण 900 है, तो वृत्तांश क्ष्ोत्रफल वृत्त क्ष्ोत्रफल के 1/4 के बराबर होगा इसका अर्थ यह होगा कि संघटक राशि का मान, समग्र राशि के 1/4 अथवा 25 प्रतिशत के बराबर है। इसी भांति 600 वृत्तांश कोण वृत्त के 1/6 क्ष्ोत्रफल को अर्थात्् इस तथ्य को प्रदर्शित करता हो कि समग्र राशि में संघटक राशि का अनुपात 1/6 है।

आरेख में सबसे बड़ा वृत्तांश व्यय की सर्वाधिक महत्वपूर्ण मद को तथा क्रम से घटते हुए वृत्तांश क्रमशः मदों के घटते हुए महत्त्व को प्रदर्शित करते हैं। अन्य शब्दों में आरेख में टेलीविजन विज्ञापन से सम्बन्धित वृत्तांश

सबसे बड़ा है, अर्थात व्यापारिक प्रतिष्ठान टेलीविजन विज्ञापन को अन्य विज्ञापन माध्यमों की तुलना में श्रेष्ठतम मानता है। महत्व के अनुसार क्रम में दूसरा स्थान अखबारी विज्ञापनों का है, तत्पश्चात् ख्ोल प्रतिस्पर्धाओं में समर्थन राशि का तथा सबसे कम महत्व 'अन्य' विज्ञापन माध्यमों का है।

5 आवृत्ति बहुभुज - बहुभुज से हमारा आशय एक ऐसी ज्यामितीय आकृति से है, जिसमें अनेक भुजायें होती हैं, तता आकृति बन्द होती है। स्पष्ट है कि आवृत्ति बहुभुज भी एक अनेकों भुजाओं वाली बन्द आकृति होगी।

आवृत्ति बहुभुज निर्मित करने के लिये हगम वर्गान्तरों के मध्य मानों को ग्.अक्ष पर तथा आवृत्तियों को ल्.अक्ष पर अंकित करते हैं। वर्गान्तर के मध्य मान से तात्पर्य वर्गान्तर सीमाओं के माध्य से है।

मध्यमानों एवं आवृत्तियों के संगत मानों को ग्राफ पर बिन्दुओं के रूप में अंकित करने के बाद उन्हें क्रम से सीधी रेखाओं के द्वारा जोड़ दिया जाता है। इस प्रकार जो प्रकृति प्राप्त होती है, वह अनेकों भुजाओं वाली आकृति तो है, परन्तु बन्द नहीं है, वरन् यह बायों तथा दाहिनी ओर खुली है।

आकृति को बन्द करने के लिये हम प्रथम वर्गान्तर से पहल्ो एक और वर्गान्तर ल्ोते हैं। वर्गान्तर के मध्यमान तथा आवृत्ति शून्य को ग्रराफ में अंकित करने पर हमें
ग्.अक्ष पर बिन्दु प्राप्त हो जाता है, जिसे सीधी रेखा में जोड़ने पर आकृति बायीं ओर बन्द हो जाती है। आकृति को दाहिनी ओर बन्द कनरे के लिये भी इसी प्रक्रिया को अपनाया जाता है।

आवृत्ति बहुभुज तथा ग्.अक्ष के मध्य क्ष्ोत्रफल आंकड़ों की कुल संख्या को प्रदर्शित करता है।

6 आवृत्ति वक्र -

आवृत्ति वक्र तथा आवृत्ति बहुभुज बनाने की समस्त प्रक्रिया पूर्णतया समान है। यहां भी ग्.अक्ष पर वर्गान्तरों के मध्यमानों को तथा ल्.अक्ष पर आवृत्तियों को अंकित करते हैं। प्राप्त बिन्दुओं को एक सतत वक्र से जोड़ देने पर आवृत्ति वक्र प्राप्त हो जाता है।

आवृत्ति वक्र वास्तव में आवृत्ति बिन्दुओं अर्थात्् ग्राफ पर अंकित बिन्दुओं का प्रवृत्ति पथ है - अन्य शब्दों में यह आवश्यक नहीं है कि आरेख का प्रत्येक बिन्दु आवृत्ति वक्र पर स्थित हो। आवृत्ति वक्र, आंकड़ों से सम्बन्धित बहुत सारी जानकारी देने में सहायक होते हैं। आवृत्ति वक्र को देखने मात्र से हमें चर राशि के परिसर, अधिकतम आवृत्ति वाला चरराशि का मान, आवृत्तियों के संकेन्द्रण की स्थिति आंकड़ों का अनुमानित औसत इत्यादि जानकरियां प्राप्त हो जाती हैं।

आयत चित्र तथा आवृत्ति बहुभुज की भाँति है, आवृत्ति वक्र तथा ग्.अक्ष के मध्य क्ष्ोत्रफल आंकड़ों की कुल संख्या को प्रदर्शित करता है। आवृत्ति वक्र को विभिन्न चरराशियों के मानों की आवृत्तियों को अनुमानित करने के लिये प्रयुक्त किया जा सकता है।

7 संचयी आवृत्ति वक्र -

संचयी आवृत्ति का अर्थ है-आवृत्तियों का योग। आवृत्तियों का योग दो प्रकार से ज्ञात किया जा सकता है ;पद्ध ऊपर से ;पपद्ध नीचे से। आवृत्तियों का योग ऊपर से लेने पर प्राप्त संचयी आवृत्तियों का क्रम वर्द्धमान होता है, अतः इन्हें वर्द्धमान अथवा आरोही संचयी आवृत्ति कहते हैं। इसकी प्रकार आवृत्तियों का योग सारिणी में नीचे से ल्ोने पर, जो संचयी आवृत्तियों प्राप्त होती हैं, इनका क्रम ह्रासमान होता है तथा इन्हें ह्रासमान अथवा अवरोही संचयी आवृत्तियां कहते हैं।

4.8 अभ्यास प्रश्न (Practice Questions)

सत्य / असत्य

रेखा चित्र द्वारा एक से अधिक श्रेणियों में आने वाले परिवर्तन की तुलना की जा सकती है -

(।) सत्य (ठ) असत्य ।दे. ।

बहुविकल्पीय प्रश्न

1. बिन्दुरेखीय प्रदर्शन में आवश्यक है -
   A. शीर्षक B. लम्बाई
   C. मापदण्ड का चुनाव D. इनमें से सभी
2. ग्राफीय विधि से आंकड़ें -
   A. सुबोध बनते हैं B. स्मरणीय बनते हैं
   C. आकर्षक बनते हैं D. उपर्युक्त सभी ।

4.9 सारांश (Summary)

इस प्रकार इस इकाई में हमने बिन्दुरेख से सम्बन्धित उन तमाम बातों की जानकारी प्राप्त की जो कि एक अच्छे आंकड़ों के प्रस्तुतीकरण के लिये आवश्यक है। चूंकि बिन्दुरेखीय विधि सरल एवं आकर्षक है इसीलिए आज लगभग हर क्ष्ोत्र में आंकड़ों के प्रदर्शन हेतु इस विधि का प्रयोग किया जा रहा है। अन्त में यह कहा जा सकता है कि

सांख्यिकीय आंकड़ों को चित्र द्वारा ग्राफीय निरूपण करने से यह अधिक सरल और सुग्राह्य बन जाते हैं। चित्रों द्वारा प्रदर्शित आंकड़ें कठिन प्रकिया से सरल प्रक्रिया में परिवर्तित हो जाते हैं। अतः यह इकाई आप सभी के लिए अत्यन्त महत्वपूर्ण है।

4.10 संदर्भ ग्रन्थ

- एस०एन०लाल, एल०के०चतुर्वेदी, (2010), ***परिमाणात्मक विश्लेषण***, शिव पब्लिशिंग हाउस, इलाहाबाद।
- नागर, कैलाश नाथ (2008), ***सांख्यिकी के मूल तत्व***, मीनाक्षी प्रकाशन, मेरठ।
- सिंह, एस 0 पी0 (2010), ***सांख्यिकी सिद्धांत एवं व्यवहार***, एस चन्द एण्ड कंपनी लिमिटेड, दिल्ली।

4.11 निबंधात्मक प्रश्न (Essay Types Questions)

1. निम्नलिखित आंकड़ों को बिन्दुरेख द्वारा प्रदर्शित कीजिए -

वर्ष	1981	1982	1983	1984	1985	1986	1987	1988	1989
चावल का उत्पादन (मिलियन टन)	10	15	18	20	22	30	32	35	38
गेहूँ का उत्पादन (मिलियन टन)	15	18	20	25	28	32	34	36	40

2. ध्यान आकर्षित करने की दृष्टि से रेखाचित्र आंकड़ों के प्रदर्शन की अन्य रीतियों की अपेक्षा अधिक प्रभावशाली होते हैं।' इस तथ्य की उदाहरण साहित व्याख्या कीजिए।

इकाई 5

प्रतिदर्श का परिचय व प्रकार

(Introduction and Types of Sampling)

उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ नमूनाकरण की मूल आवधारणा की व्याख्या कर सकें।
- ✓ आंकड़े एकत्रित करने की विधियों के प्रकार की व्याख्या कर सकें।
- ✓ नमूनाकरण विधियों के प्रकार की व्याख्या कर सकें।
- ✓ नमूनाकरण एवं गैर नमूनाकरण त्रुटियों की व्याख्या कर सकें।
- ✓ नमूनाकरण वितरण की अवधारणाओं का वर्णन कर सकें।

5.1 प्रस्तावना (Introduction)

जीवन के सभी क्षेत्रों के लिए (जैसे आर्थिक, सामाजिक और व्यापार) सांख्यिकीय जांच और संमकों के विश्लेषण की आवश्यकता दिन प्रतिदिन बढ़ रही है। सांख्यिकीय आँकड़ों को एकत्र करने के दो तरीके हैं: 1. जनगणना विधि और 2. नमूना विधि। जनगणना विधि के तहत, पूरी जानकारी के लिए संबंधित जाँच के दायरे या जनसंख्या इकाइयों को एकत्र किया जाता है जबकि नमूना विधि के तहत, इसके विपरीत समग्र की ~~में~~ केवल चयनित इकाइयों से संबंधित जानकारी एकत्र की जाती है। आधुनिक समय में नमूना विधि सांख्यिकीय जाँच का एक महत्वपूर्ण और लोकप्रिय तरीका है। आर्थिक और व्यापार की दुनिया के अलावा, इस विधि का व्यापक रूप से दैनिक जीवन में प्रयोग किया जाता है। उदाहरण के लिए, एक ~~घर पत्नी~~ गृहणी को पकवान का स्वाद एक छोटे से नमूने को चखने पर समग्र के सुगंध, नमक, मिर्च खुशबू आदि का पता चलता है। चिकित्सक मरीज के रक्त की जाँच रक्त की केवल एक या दो बूँद द्वारा कर सकता है। उसी तरह हम दैनिक उपयोग की वस्तुओं जैसे गेहूँ, चावल, दाल आदि खरीदने से पहले इन चीजों के नमूनों को देखकर चीजों के गुणवत्ता के बारे में जानते हैं।

कारखानों में सांख्यिकीय गुणवत्ता नियंत्रक उत्पादन की कुछ इकाईयों का परीक्षण कर समग्र की गुणवत्ता का परीक्षण करता है। एक शिक्षक कुछ छात्रों से प्रश्न पूछकर अपने शिक्षण की प्रभावकारिता के बारे में जान लेता है। वास्तविकता में, शायद ही कोई ऐसा क्षेत्र हो जहाँ नमूने विधि का प्रयोग नहीं किया जाता है।

5.2 नमूने की बुनियादी अवधारणाएँ (Basic Concepts of Sampling)

इससे पहले कि आप नमूने के पहलुओं का विस्तृत अध्ययन करें आप को नमूने से संबधित बुनियादी अवधारणाओं को समझना चाहिए जो कि निम्नवत हैः

5.2.1 समग्र (Census/Population)

सांख्यिकी में समग्र का आशय संग्रहित (कुल) वस्तुएँ/चीजें जिसके बारे में आप जानकारी प्राप्त करते है। समग्र का आशय जाँच के तहत पूरे क्षेत्र से है जो ज्ञान की मांग करता है। उदाहरण के लिए यदि आप विद्यालय के 2,000 विद्यार्थियों के औसत मासिक व्यय के बारे मे सूचना चाहते हैं, तो उसके दो प्रकार हैः 1. सीमित 2. असीमित।

सीमित समग्र में वस्तुओं की संख्या निश्चित होती है जैसे- कालेज में छात्रों या शिक्षकों की संख्या निश्चित है। दूसरी ओर असीमिति समग्र में, वस्तुओं की संख्या अनिश्चित होती है। जैसे आसमान में तारों की संख्या, समुद्र में पानी की बूँदें, पेड़ मे पत्तियों की संख्या, सिर पर बालों की संख्या।

5.2.2 नमूना/प्रतिदर्श (Sample)

समग्र के चयन का एक हिस्सा नमूना/प्रतिदर्श कहा जाता है। दूसरे शब्दों में समग्र से चयनित या वर्गीकृत इकाईयों को नमूना/प्रतिदर्श कहते हैं। वास्तव में, नमूना/प्रतिदर्श समग्र का वह हिस्सा है जिसे जाँच करने के उददेश्य से चयनित करते हैं। उदाहरण के लिए एक अन्वेषक (Researcher) एक कालेज के 2000 छात्रों में से 200 छात्रों को चयनित करता है जो कि 2000 छात्रों का प्रतिनिधित्व करते हैं, तो इन 200 छात्रों को एक नमूने के रूप मे करार दिया नमूना/प्रतिदर्श कहा जाएगा। इस प्रकार, नमूने/प्रतिदर्श का आशय समग्र से चयनित उन इकाईयों से है जो कि समग्र का प्रतिनिधित्व करते हैं।

5.3 संमकों को संग्रह करने की विधियाँ (Methods of Data Collection)

सांख्यिकी संमकों को संग्रह करने के दो तरीके निम्नवत हैं

5.3.1. जनगणना विधि (Population/Census Method): जनगणना विधि वह तरीका है जिसमें जाँच के लिए जानकारी या समंक जो कि समग्र से सम्बन्धित है के प्रत्येक एवं हर इकाई को एकत्रित करके एवं उनके आधार पर निष्कर्ष तैयार किया जाता है। उदाहरण के लिए, (मासिक व्यय, औसत ऊँचाई, औसत वजन इत्यादि) यदि कालेज के 2000 छात्रों के बारे में कुछ जानकारी एकत्र की जा रही है उस उद्देश्य में यदि आप कालेज प्रत्येक एवं हर छात्र से प्रश्नमय करते हैं तो इस विधि को जनगणना विधि कहते हैं। इस उदाहरण में सम्पूर्ण कालेज या सभी 2000 छात्र समग्र समझे जायेंगे और व्यक्तिगत रूप से प्रत्येक छात्र समग्र इकाई कहलायेगा। भारत में जनगणना विधि या पूर्ण गणना विधि का उपयोग हर 10 वर्षों बाद किया जाता है।

5.3.1.1 जनगणना विधि के गुण (Merits of Census Method) :

विश्वसनीय और सटीक समंक : जनगणना विधि द्वारा प्राप्त आँकड़े अधिक विश्वसनीय एवं सटीक होते हैं क्योंकि इस विधि में समंकों को समग्र के प्रत्येक एवं हर इकाई से सम्पर्क कर एकत्रित किया जाता है।

व्यापक जानकारी : यह विधि समग्र के प्रत्येक इकाई के बारे में विस्तृत जानकारी देता है। उदाहरण के लिए, भारतीय जनगणना केवल व्यक्तियों की संख्या के बारे मे ज्ञान नहीं देता अपितु व्यक्तियों की आयु, व्यवसाय, आय, शिक्षा, वैवाहिक स्थिति के बारे में भी जानकारी देता है।

उपयुक्तता : यह विधि सीमित दायरे एवं विविध विशेषताओं वाले समग्र के लिए अधिक उपयुक्त है। इस विधि का उपयोग गहन अध्ययन मे भी वांछित है।

5.3.1.2 जनगणना विधि के दोष (Demerits of Census Method):

अधिक महंगा : जनगणना विधि एक महंगी विधि है। समग्र के प्रत्येक इकाई से जानकारी एकत्र करने के लिए ज्यादा धन की आवश्यकता होती है। यही कारण है कि सरकार द्वारा इस विधि का ज्यादातर उपयोग महत्वपूर्ण मुद्दों के लिए किया जाता है जैसे -जनगणना।

ज्यादा समय का लगना : इस विधि में आँकड़ों के सग्रह में ज्यादा समय लगता है क्योंकि आँकड़े सम्रग के प्रत्येक एवं हर इकाई से संग्रहित किये जाते हैं इस कारण से सांख्यिकी निष्कर्ष बनाने मे देरी होती है।

अधिक श्रम का लगनाः आँकड़ें संग्रहित करने की इस विधि में बहुत ज्यादा परिश्रम लगता है। इसके लिए प्रगणक को बड़ी संख्या की आवश्यकता होती है।

विशिष्ट समस्याओं के लिए उपयुक्त नहीं : यह विधि कुछ विशिष्ट समस्या और अनंत समग्र के सम्बन्ध में उपयुक्त नहीं है। उदाहरण के लिए, अगर समग्र अनन्त है या समग्र की इकाईयाँ खराब या प्रकृति में बहुत जटिल हैं तो जनगणना विधि उपयुक्त नहीं हैं।

5.3.2 नमूना विधि (Sampling Method):

नमूना विधि वह विधि है जिसमें समग्र से चयनित नमूना इकाईयों से आँकडें एकत्रित कर समग्र के लिए निष्कर्ष निकाला जाता है। उदाहरण के लिए यदि कालेज के 2000 छात्रों के मासिक व्यय का अध्ययन किया जाता है तो आप 2000 छात्रों से जानकारी एकत्र करने के बजाय कुछ चयनित छात्रों से जैसे 100 छात्रों से जानकारी एकत्र कर निष्कर्ष निकाल सकते हैं, तो इस विधि को नूना विधि कहा जायेगा।

नमूना विधि के आधार पर कालेज के सभी छात्रों के मासिक व्यय का ध्यान करना सम्भव है। नमूना विधि के तीन महत्वपूर्ण चरण हैंः

नमूने का चयन करना
नमूने से जानकारी एकत्र करना

समग्र के लिए निष्कर्ष निकालना

5.3.2.1 नमूने विधि के गुण (Merits of Sampling Method)

कम खर्चीली : यह विधि कम खर्चीली है। इस विधि मे धन एवं श्रम दोनों की बचत होती है क्योंकि इसमे समग्र की कुछ इकाईयों का अध्ययन किया जाता है।

समय की बचत : इस विधि मे आँकडे जल्दी से एकत्र किए जा सकते हैं क्योंकि आँकडे समग्र की कुछ इकाईयों से प्राप्त किए जाते हैं जिससे ज्यादा समय की बचत होती है।

गहन अध्ययन : नमूना विधि में इकाईयों की संख्या कम होती है जिससे समग्र का गहन अध्ययन किया जा सकता है।

संगठनात्मक सुविधा : इस विधि में अनुसंधान कार्य का आयोजन एवं क्रियान्वयन अधिक आसानी से किया जा सकता है। अधिक कुशल और सक्षम जाँचकर्ता नियुक्त किए जा सकते हैं।

अधिक विश्वसनीय परिणाम : यदि समग्र से नमूनों का चयन इस तरह से किया जाता है कि चयनित नमूने सम्पूर्ण समग्र का प्रतिनिधित्व करते हैं, तो इससे उत्पन्न परिणाम अधिक सटीक एवं विश्वसनीय होंगें।

अधिक विज्ञान संबंधी :- नमूना विधि ज्यादा विज्ञान संबंधी है क्योंकि इसमें आँकड़ों की पूछताछ अन्य नमूनों के साथ की जा सकती है।

एकलौती विधि : कुछ ऐसे क्षेत्र जहाँ पूछताछ जनगणना विधि से सम्भव नहीं है उन परिस्थितियों में केवल नमूना विधि ही आँकड़े एकत्रित करने के लिए उपयुक्त हैं । यदि समग्र अनन्त या बड़े पैमाने पर या खराब होने की प्रकृति का है, तो केवल नमूना विधि का प्रयोग इस तरह के मामलों में किया जाता है।

5.3.2.2 नमूने विधि के दोष (Demerits of Sampling Method) :

कम सटीकता : नमूने विधि मे कम सटीकता होती है क्योंकि समग्र की प्रत्येक इकाई मे पूछताछ करने के बजाय इसमे केवल चयनित इकाईयों से आंशिक पूछताछ की जाती है।

गलत निष्कर्ष : यदि एक नमूना चयन की विधि निष्पक्ष या इसके चयन में सावधानी नहीं बरती गई है तो निश्चित रूप से परिणाम गुमराह करते हैं।

कम विश्वसनीय : जनगणना विधि की तुलना में, इस विधि में अन्वेषक के पक्षपात की ज्यादा सम्भावना होती है, जो परिणाम को कम विश्वसनीय बनाता है।

निर्दिष्ट ज्आन की आवश्यकता : यह एक जटिल विधि है जिसमें नमूने के चयन के लिए निर्दिष्ट ज्ञान की आवश्यकता होती है।

उपयुक्तता का अभावः नमूना विधि समग्र के इकाईयों के मध्य विविधता के मामले में उपयुक्त नहीं है।

5.3.3 जनगणना विधि एवं नमूना विधि में अन्तर: (Difference Between Census and Sampling Method)

जनगणना विधि एंव नमूना विधि के बीच मुख्य अन्तर निम्नवत है-

विस्तार : जनगणना विधि मे समग्र से सम्बन्धित सभी इकाईयों की जाँच की जाती है जबकि नमूना विधि में केवल कुछ इकाईयों से पूछताछ की जाती है।

कीमत : जनगणना विधि समय, धन एवं श्रम की दृष्टि से कीमती है जबकि इन मामलों में नमूना विधि किफायती है।

जाँच का क्षेत्र : जाँच के लिए जनगणना विधि का प्रयोग सीमित क्षेत्र तक किया जाता है जबकि नमूने विधि का प्रयोग बड़े क्षेत्र तक किया जाता है।

समरूपता : जनगणना विधि का प्रयोग वहाँ पर उपयोगी है जहाँ समग्र के इकाईयों मे विविधता होती है जबकि नमूना विधि का प्रयोग समग्र की इकाईयों मे समरूपता होने पर किया जाता हे।

समग्र का प्रकार : जनगणना विधि उन क्षेत्रों में ज्यादा उपयुक्त है जहाँ समग्र के प्रत्येक एवं सभी इकाईयों का अध्ययन आवश्यक है। इसके विपरीत, जब समग्र अनंत या विशाल या पूर्ण गणना के परिणाम में नष्ट किया जा रहा है तो नमूना विधि को अधिक उपयुक्त माना जाता है।

5.4 नमूनाकरण विधियाँ (Sampling Methods):

दिए गए समग्र में से नमूना चयन की विधि को नमूनाकरण कहा जाता है। दूसरे शब्दों में, नमूनाकरण संग्रहित सांख्यिकीय सामाग्री के चयन के उस भाग को दर्शाता है जिसके बारे में समग्र की दृष्टि से जानकारी प्राप्त की जाती है। विभिन्न आवश्यकताओं के अनुसार समग्र में से नमूना चयन की कई विधियाँ है :

5.4.1 प्रायिकता नमूनाकरण विधिया ँ (Probability Sampling Methods)

सरल यादृच्छिक नमूनाकरण

स्तरीयकृत यादृच्छिक नमूनाकरण

क्रमबद्ध यादृच्छिक नमूनाकरण

बहुचरणीय यादृच्छिक नमूनाकरण

गुच्छीय नमूनाकरण

5.4.1 प्रायिकता नमूनाकरण विधिया ँ (Probability Sampling Methods)

प्रायिकता नमूना विधियाँ समग्र में से नमूने चयन की ऐसी विधियाँ हैं जिसमें समग्र की सभी इकाईयों को समान अवसर देते हुए नमूने मे शामिल किया जाता है। प्रायिकता नमूनाकरण विधियों के विभिन्न रूप होते हैं, जो नीचे दिये गए है :

5.4.1.1. सरल यादृच्छिक नमूनाकरण (Simple Random Sampling):

सरल यादृच्छिक नमूना वह पद्धति है जिसमें समग्र की प्रत्येक इकाई का नमूने में चयनित होने का समान अवसर होता है। कौन सा तत्व नमूने मे शामिल होगा और कौन सा नहीं, इस तरह के निर्णय जांचकर्ता द्वारा अपनी इच्छा से नहीं बनाये जाते बल्कि नमूनों का चयन संयोगवश होता है। यादृच्छिक नमूना चयन की दो विधियाँ होती हैं :

अ) लाटरी विधि (Lottery method):- इस विधि मे समग्र के प्रत्येक इकाई को कागज के एक टुकड़े में नामित या क्रमांकित किया जाता है। इन पर्चीयों को मोड़कर कलश या थैले में डाला जाता है। तद्पश्चात कई इकाईयों को एक नमूने मे शामिल करने के लिए कई पर्चीयों को किसी व्यक्ति द्वारा चुना जाता है।

ब) यादृच्छिक अंकों की तालिका (Table of Random Numbers):- कुछ विशेषज्ञों ने यादृच्छिक अंक तालिकाओं का निर्माण किया है। ये तालिकाऐं नमूने के चयन मे सहायता करती है। इन सभी विभिन्न तालिकाओं मे टिपेट तालिका सबसे प्रसिद्ध एवं उपयोग में है। टिपेट ने 10400 संख्याओं की, 41600 संख्याओं के रूप द्वारा चार अंकों वाली तालिका का निर्माण किया है। इस विधि में, सबसे पहले, समग्र की सभी इकाईयों को क्रमिक रूप से लिखा जाता है। तद्पश्चात नमूने के आकार के अनुसार टिपेट तालिका का प्रयोग करते हुए, अंकों का चयन किया जाता है। एक उदाहरण द्वारा, टिपेट तालिका की सहायता से, नमूना चयन को स्पष्ट रूप से समझा जा सकता है :

टिपेट की तालिका का एक उदाहरण :

2952 6641 3992 9792 7979 5911

3170 5224 4167 9525 1545 1396

7203 4356 1300 2693 2370 7483

3408 2762 3563 6107 6913 7691

0560 5246 1112 9025 6008 8127

उदाहरण के लिए, मान लें कि 5000 इकाईयों में से 12 इकाईयों को चुना जाना है। इन इकाईयों को निर्धारित करने के लिए, टिपेट तालिका के लिए, पहले 5000 इकाईयों को 1 से 5000 तक क्रमबद्ध करना होगा और तब टिपेट तालिका से 12 अंकों का चयन प्रारम्भ से, जो कि 5000 से कम होगा किया जायेगा। ये 12 अंक निम्नवत हैं :

2952 4156 4356 2370

3992 1545 1300 3408

3170 1396 2693 2762

इकाईयों की यह क्रम संख्या को नमूने मे शामिल किया जायेगा। यदि समग्र की इकाईयाँ 100 से कम हो, तब 4 अंक के यादृच्छिक संख्या को, 2 अंकों की संख्या में (छोटा) संक्षिप्त किया जायेगा, और तब इन दो अंकों की संख्या का चयन होगा। इसी तरह 60 इकाईयों में से यदि 6 इकाईयाँ चयनित करनी है, तो क्रम संख्या 29, 39, 31, 41, 15 और 13 को नमूने में शामिल किया जायेगा।

गुण (Merits) :

1. इस पद्धति (विधि) में निजी पूर्वाग्रह की कोई संभावना नहीं होती है। दूसरे शब्दों में, यह विधि व्यक्तिगत पूर्वाग्रह से ~~मुक्त~~ मुक्त होती है।
2. इस विधि के अन्तर्गत, समग्र की प्रत्येक इकाई के चयन का अवसर एक समान होता है।
3. इस विधि के प्रयोग से समय, धन तथा श्रम की बचत होती है ।

दोष (Demerits):

1. यदि नमूने का आकार छोटा है, तो नमूना पर्याप्त रूप से समग्र का प्रतिनिधित्व नहीं करता है।
2. यदि समग्र बहुत छोटा है, तो यह विधि उपयुक्त नहीं है।
3. यदि समग्र में कुछ वस्तुएँ इतनी महत्वपूर्ण हैं कि नमूने में उनका शामिल किया जाना बहुत जरूरी है तो इस विधि का प्रयोग उचित नहीं होगा।
4. जब समग्र की इकाईयाँ विविध लक्षणों के साथ हो तो इस विधि का प्रयोग उचित नहीं होगा।

5.4.1.2 स्तरीय यादृच्छिक नमूना (Stratified Random Sample) :

जब समग्र के इकाईयों मे समरूपता के बजाय विविधता होती है तो इस विधि का प्रयोग किया जाता है। इस विधि के अन्तर्गत समग्र के भी इकाईयों को पहले उनकी विशेषताओं के अनुसार अलग अलग हिस्सों में विभाजित किया जाता है। उसके बाद यादृच्छिक नमूने का उपयोग करके प्रत्येक परत से नमूना इकाई का चयन किया जाता है। उदाहरण के लिए, यदि किसी कालेज के 1500 छात्रों में 150 छात्रों का चयन करना है तो सबसे पहले कालेज के विद्यार्थियों को कला, व्यवसाय एंव विज्ञान विषय के आधार पर तीन वर्गों में विभाजित करना होगा। माना इन तीन संकायों में क्रमशः 500, 700, 300 छात्र हैं और 10% नमूने लेने है। तब यादृच्छिक नमूना विधि के प्रयोग के आधार पर क्रमशः 50, 70 और 30 छात्र चयनित किये जा सकेंगें। इस तरह, इस विधि मे प्रत्येक कक्षा या वर्ग की आनुपातिक प्रतिनिधित्व की अवधारणा रहती है और समग्र की सभी इकाईयों को नमूने में चयनित किये जाने का बराबर का मौका मिलता है ।

गुण (Merits) :

1. इस विधि में इकाईयों के चयनित होने की ज्यादा सम्भावनाएं होती है।
2. तथ्यों के आधार पर विभिन्न स्तर पर इस विधि के तहत तुलनात्मक अध्ययन संभव है।
3. इस विधि में ज्यादा शुद्धता होती हैं।

दोष (Demerits):

1. इस विधि का सीमित दायरा है क्योंकि इस विधि को तभी अपनाया जा सकता है जब केवल समग्र और उसके विभिन्न तबको का ज्ञान हो।
2. इस विधि में पूर्वाग्रह की संभावना हो सकती है यदि समग्र ठीक से स्तरीकृत न हो।
3. यदि समग्र आकार में भी बहुत छोटा हो तो इस स्तरीकृत करने में कठिनाई होती है।

5.4.1.3 क्रमबद्ध यादृच्छिक नमूना (Systematic Random Sample):

इस विधि मे समग्र के सभी इकाईयों को क्रमबद्ध तरीके से व्यवस्थित और गिना जाता है और तब नमूना इकाई को बराबर के अन्तराल में चयनित किया जाता है। उदाहरण के लिए, यदि 50 छात्रों में से 5 को नमूने के लिए चयनित कर रहे हैं तो 50 छात्रों को गिनकर क्रमबद्ध तरीके से व्यवस्थित किया जायेगा। पहले दस में से एक इकाई को यादृच्छिक तरीके से चयनित किया जायेगा। इसके बाद चयनित संख्या से प्रत्येक 10 वीं इकाई का चयन नमूना बनावट के लिए होगा। यदि प्रथम चयनित अंक 5 वीं इकाई है तो उसके बाद के अंक 15 वीं इकाई, 25 वीं इकाई, 35 वीं इकाई और 45 वीं इकाई होंगें।

गुण (Merits) :

1. यह एक सरल विधि है। इसके द्वारा नमूने आसानी से प्राप्त किये जा सकते हैं।
2. इस विधि द्वारा नमूना चयन में बहुत कम समय लगता है और परिणाम लगभग सटीक होता है।

दोष (Demerits):

1. इस विधि में, प्रत्येक इकाई को चयनित होने के बराबर मौके नहीं है क्योंकि केवल पहली इकाई का चयन यादृच्छिक नमूना विधि पर आधारित है।
2. यदि सभी इकाइयाँ लक्षणों में भिन्न हैं तो परिणाम उचित नहीं होगें।

5.4.1.4 बहु-स्तरीय यादृच्छिक नमूनाः(Multi-Stage Random Sampling:)-

जब नमूना पद्धति विभिन्न स्तरों से गुजरती है तो इसे बहुस्तरीय यादृच्छिक नमूना कहा जाता है। इस विधि मे सर्वप्रथम सम्पूर्ण समग्र को स्तरों या उप स्तरों में विभाजित किया जाता है। प्रत्येक स्तर पर कुछ इकाईयों का चयन यादृच्छिक नमूने के आधार पर किया जाता है। तद्पश्चात इन इकाईयों का उप विभाजन किया जाता है और फिर से यादृच्छिक नमूना विधि के आधार पर कुछ उप इकाइयों को चयनित किया जाता है। उदाहरण के लिए एक राज्य में प्रौढ़ शिक्षा के अध्ययन के उद्देश्य को जानने के लिए, सर्वप्रथम यादृच्छिक आधार कुछ जिलों को चयनित किया जायेगा। तद्पश्चात चयनित जिलों से कुछ तहसीलें और तहसीलों से कुछ वार्ड और वार्डों से कुछ परिवारों का चयन किया जायेगा जिनसे समस्या के विषय में पूछताछ की जा सकेगी।

गुण (Merits) :

1. क्षेत्रीय आधार पर समग्र के अध्ययन की यह सर्वोत्तम विधि है।
2. यह विधि उन समस्याओं के लिए उपयुक्त है जहाँ अकेले नमूने के आधार पर निर्णय नहीं लिया जा सकता है।

दोष (Demerits):

1. सही ढंग से सटीकता के स्तर का अनुमान लगाने के लिए इस विधि में कई परीक्षणों जिसमें अधिक समय एवं श्रम शामिल हैं, की आवश्यकता होती है।
2. इस विधि में अनुमानित सटीकता का स्तर पूर्व में ही तय होता है जोकि तक्र तर्क संगत प्रतीत नहीं होता है।

5.4.1.4 गुच्छ नमूना (Bunch Sample) :

इस विधि में सीधे समग्र को कई हिस्सों में विभाजित किया जाता है जिन्हें गुच्छ कहते हैं जिनमें से कुछ गुच्छों को यादृच्छिक आधार पर चयनित किया जाता है तब इन गुच्छों का पूर्ण रूप से गणना की जाती है। यह विधि सामान्यतया उद्योग जगत में प्रयोग की जाती है- जैसे भेषजीय उद्योग, एक मशीन जो कि प्रत्येक 100 के खेप में दवा की गोलियाँ बनाती है, तो गुणवत्ता निरीक्षण के लिए कुछ यादृच्छिक खेपों को चयनित कर जाँच की जाती है।

5.4.2 गैर-प्रायिकता नमूनाकरण विधिया ँ (Non-Probability Sampling Methods)

आलोचनात्मक नमूनाकरण

नियतांश नमूनाकरण

विस्तृत नमूनाकरण

गैर प्रायिकता नमूना विधियाँ वह विधियाँ हैं जिसमें इकाईयों का चयन प्रायिकता या संभाविता के आधार पर न होकर अन्वेषक के सुविधा या निर्णय के अनुसार होता है। इस तरह की विधियों में इकाईयों का चयन विशेष उद्देश्य के साथ एवं अन्वेषक की सुविधानुसार होता है।

5.4.2.1 निर्णय नमूना (Decision Model) :- इस विधि के अन्तर्गत, नमूना इकाईयों का चयन पूरी तरह से अन्वेषक के निर्णय पर निर्भर करता है। दूसरे शब्दों में, अन्वेषक अपने (उसका/उसकी) निर्णय से पसन्द के नमूने का प्रयोग करता है और अध्ययन के अर्न्तगत समग्र से केवल उन नमूना इकाईयों को शामिल करता है जिनमें विशिष्ट लक्षण हैं।

उदाहरण के लिए यदि एक कक्षा के 80 छात्रों में से 20 छात्रों का एक नमूना चयनित करना है जिससे 10 छात्रों की खर्चीली प्रवृत्ति का मनोविश्लेषण किया जा सके, अन्वेषक उन्हीं 20 छात्रों को चयनित करेगा जिनकी, उसका/उसकी राय कक्षा में उक्त अध्ययन पर प्रतिनिधित्व करेंगी।

गुण (Merits):-

1. यह विधि कम खर्चीली है।
2. यह विधि बहुत सरल और आसान है।
3. इस विधि का प्रयोग उन क्षेत्रों में उपयोगी है जहाँ लगभग एक ही तरह की इकाईयाँ मौजूद हैं या कुछ इकाईयाँ आवश्यक हैं जिन्हें नमूने से बाहर नहीं निकाला जा सकता है।

दोष (Demerits):

1. इस विधि में अन्वेषक के स्वयं के पूर्वाग्रह का एक बड़ा मौका होता है।
2. यह विधि बहुत सटीक एवं विश्वसनीय नहीं है।

5.4.2.2 नियतांश नमूना (Quota Sampling):-

इस विधि में जाँचकर्ता कुछ मापदण्डों (कोटा) के अनुसार निश्चित कोटा आवंति आबंटित करते हैं। उन्हें अपेक्षित संख्या प्राप्त कर प्रत्येक कोटा को भरने के लिए निर्देश दिये जाते हैं। जानकारी एकत्र करने के लिए अन्वेषकों को व्यक्तियों या (नमूना इकाईयाँ) का चयन अपने निर्णय से कोटा के भीतर किया जाता है। जब सम्पूर्ण नियतांश के सभी या आंशिक प्रतिक्रियादाता उपलब्ध या सुगम्य नहीं होते है तो नियतांश को नये प्रतिक्रियादाता के पूरक में पूर्ण कर लिया जाता है। नियतांश नमूना, निर्णय नमूने का एक प्रकार है।

गुण (Merits) -

1. इस विधि मे महत्वपूर्ण इकाईयों को सम्मिलित करने का बड़ा मौका होता है।
2. नियतांश की निर्धारित इकाईयों के कारण इस विधि में सांख्यिकी जाँच ज्यादा संगठित होती है।

दोष (Demerits):

1. प्रायिकता की पूर्वाग्रह पहले की तरह रहेगी।
2. इस विधि में नमूना त्रुटियों की ज्यादा सम्भावना होती है।

5.4.2.3 सुविधा नमूना (Convenience Sample) :-

इस गैर प्रायिकता नमूना मे अन्वेषक को सुविधानुसार पूर्ण रूप से नमूना पंसद करने की छूट दी जाती है। अन्वेषक कालेज विवरणिका के आधार पर अध्यापकों की सूची से पसंदित नमूना प्राप्त करता है और (उसका/उसकी) अपने प्रकाशन के संदर्भ में जानकारी प्राप्त करता है। यह विधि कम कीमती एवं सरल है लेकिन

अवैज्ञानिक एवं अविश्वसनीय है। इस विधि का परिणाम गणनाओं पर ज्यादा निर्भर करता है। जहाँ पर समग्र को स्पष्ट रूप से परिभाषित नहीं किया जाता है या इकाइयों की सूची उपलब्ध नहीं है या नमूना इकाईयों की सूची उपलब्ध नहीं है या नमूना इकाईयाँ स्वयं में स्पष्ट नहीं है तो यह विधि नमूना चयन के लिए उपयुक्त है।

5.4.2.3 विस्तृत नमूना (Detailed Sample):-

इस विधि मे नमूने का आकार समग्र के ही रूप में लगभग बड़ा लिया जाता है। जैसे समग्र का 90% केवल उन्हीं इकाईयों को छोड दिया जाता है जिनसे आँकड़े एकत्र करने में ज्यादा कठिनाई या लगभग असम्भव होते हैं। बहुत बड़ा नमूना आकार होने के कारण इस विधि में बड़े स्तर की सटीकता होती है। समस्या का विस्तृत अध्ययन सम्भव है। इस विधि के निष्कासन में भारी संसाधनों की जरूरत होती है।

5.5 नमूनाकरण और गैर नमूनाकरण त्रुटिया ँ (Sampling and Non-Sampling Errors)

हालाँकि एक नमूने की पसंद अत्यन्त सावधानी से की जा सकती है, फिर भी निश्चित रूप से इसमें दो तरह की त्रुटियाँ शामिल होती हैंः (1) नमूना त्रुटियाँ (2) गैर नमूना त्रुटियाँ
ये त्रुटियाँ आँकड़ों के संग्रह, प्रसंस्करण और विश्लेषण में घटित हो सकती हे। आँकड़ों के नमूना वितरण की एक महत्वपूर्ण विशेषता यह होती है कि समग्र से बड़े आकार का यादृच्छिक नमूना ($n>30$) लिया गया हो जोकि सामान्य रूप से वितरित है या नहीं लेकिन आँकड़ों का नमूना वितरण, सामान्य वितरण के समीप होगा।

5.5.1 नमूनाकरण त्रुटियाँ (Sampling Errors):-

नमूना गलतियाँ वे हैं जो कि नमूना विधि के कारण पैदा होती है। नमूना गलतीयाँ निम्न कारणों की वजह से मुख्य रूप से उत्पन्न होती हैंः-

1. नमूना विधि का गलत चयन।
2. नमूना एकत्रित होने की समस्याओं की वजह से एक नमूना इकाई को दूसरे नमूने इकाई के साथ प्रतिस्थापित करने से।
3. नमूने इकाईयों का गलत सीमांकन करने से
4. समग्र के विभिन्न लक्षणों में परिवर्तनशीलता से भिन्नता ।

5.5.2 गैर नमूनाकरण त्रुटियाँ (Non-Sampling Errors):-

गैर नमूना त्रुटियाँ वो है जो मानवीय कारकों से घटित होती हैं जो एक अन्वेषक से लेकर दूसरे अन्वेषक तक बदलती है। ये निम्न कारकों में से किसी एक कारक के वजह से उत्पन्न होती है :-

खराब योजना।
नमूना इकाईयों का गलत चयन
कर्मचारी जो आँकड़े एकत्रित करते हैं उनमें प्रशिक्षण एवं अनुभव का अभाव
प्रतिक्रियादाता की तरफ से लापरवाही एवं गैर प्रतिक्रिया होने पर
संकलन में त्रुटियाँ होने पर
गलत सांख्यिकी मापों की वजह से
गलत प्रश्नावली की बनावट से
नमूना निरीक्षण की जाँच अपूर्ण होने पर

5.6 प्रतिस्थापन के साथ और इसके बिना नमूना (Sample With and Without Replacement) :-

नमूना, समग्र से चयनित नमूने की एक विधि है। नमूना प्रतिस्थापन के साथ या इसके बिना भी किया जा सकता है। नमूना जहाँ समग्र से प्रत्येक इकाई को एक बार से ज्यादा चुन लिया जाता है तो उसे नमूना प्रतिस्थापन के साथ कहा जाता है। यदि प्रत्येक इकाई को एक बार से ज्यादा नहीं चुन सकते उसे नमूना प्रतिस्थापन के बाहर कहा जाता है। प्रतिस्थापन के साथ नमूना के संदर्भ में समग्र का आकार N होने पर n आकार के नमूनों की कुल सम्भावित संख्या N^n लेकिन यदि प्रतिस्थापना के बाहर नमूने के लिए कल्पना कुल सम्भावित नमूनों की संख्या $NC_n = m$ होगी ।

5.7 आँकड़ों का नमूनाकरण वितरण (Sampling Distribution of Data)

नमूना वितरण सांख्यिकीय अनुमान के सैद्धान्तिक आधार का गठन करता है और इसका व्यापार निर्णय लेने के लिए काफी महत्व है।

ऑकड़ों का नमूना वितरण आवृत्ति वितरण है जो एक ही आकार मे एक ही समग्र से तैयार की गई विभिन्न नमूनों से गणना कर विभिन्न मान्यता के साथ निकाली जाती है। माना आप समग्र (N) से प्रतिस्थापना के साथ या बाहर n आकार के सभी सम्भावित नमूनों को निकालते हैं। समग्र से निकले हुए सम्भावित प्रत्येक नमूने के लिए आप आँकड़ों के लिए जैसे माध्य, माध्यिका, मानक विचलन, परिवर्तनशीलता आदि की गणना करते हैं। तब आंकड़ों के सभी सम्भव मूल्यों को आवृत्ति विभाजन या प्रायिकता विभाजन में वर्गीकृत किया जाता है। इस तरह के आँकड़ों के वितरण से प्राप्त वितरण को नमूना वितरण कहते हैं। आँकड़ों की प्रकृति के आधार पर आपके द्वारा गणना किये हुए विभिन्न नमूना वितरण हो सकते हैं। जैसे यदि कोई विशेष आँकड़ा जिसका नमूना माध्य (Sample Mean) ज्ञात किया जाता है तो वह वितरण माध्य का नमूना वितरण कहलाएगा। यदि आप प्रत्येक नमूने की परिवर्तनशीलता की गणना करते हैं तो इसे परिर्वतनशीलता का नमूना वितरण कहते हैं। इसी तरह आप अनुपात, माध्यिका मानक विचलन आदि के नमूने वितरण की गणना कर सकेंगें।

5.8 आँकड़ों की मानक त्रुटियाँ (Standard Error of the Data)

आँकड़ों के नमूने विवरण के मानक विचलन को आँकड़ों का मानक त्रुटियाँ कहते हैं। जैसा कि विभिन्न प्रकार के नमूना वितरण होते हैं, नमूना वितरण के प्रकृति के आधार पर आपके पास विभिन्न प्रकार की मानक त्रुटियाँ हो सकती हैं। नमूना वितरण के माध्य के मानक विचलन को माध्य की मानक त्रुटियाँ कहते हैं। माध्य की मानक त्रुटि नाप की सीमा है नमूना माध्य, समग्र माध्य से पृथक किया जाता है। इस प्रकार, मानक विचलन एवं माध्य मानक त्रुटि के बीच मूलभूत अन्तर यह है कि मानक विचलन जिसमें व्यक्तिगत इकाईयों की मापों की सीमा को मिलान केन्द्र के मूल्य से पृथक किया जाता है और माध्य मानक त्रुटि वह सीमा है जिसमें व्यक्तिगत नमूना माध्य को समग्र माध्य से पृथक किया जाता है। माध्य की मानक त्रुटियों की तरह आप के पास माध्यिका, मानक त्रुटियाँ, मानक विचलन, अनुपात, परिवर्तनशीलता आदि हो सकते हैं।

मानक त्रुटि का उपयोग बड़ी संख्या में समस्याओं के लिए किया जाता है जिनका वर्णन निम्नवत किया जाता हैः-

1) नमूने की विश्वसनीयता के लिए :- मानक त्रुटि नमूने की विश्वसनीयता एवं यर्थाथता के बारे में एक धारणा देती है अर्थात अनुमानित मान प्रेषित मान से कितना ज्यादा भिन्न है। ज्यादा मानक त्रुटि होने पर, अनुमानित एवं प्रेषित मानों के बीच में ज्यादा विचलन होता है और नमूने की विश्वसनीयता बहुत कम होती है। मानक त्रुटि बहुत कम होने पर अनुमानित एवं प्रेषित मानों के मध्य बहुत कम विचलन होता है और नमूने की विश्वसनीयता बहुत ज्यादा होती है।

2) परीक्षणों का महत्व :- मानक त्रुटियों का उपयोग छोटे एवं बड़े नमूनों से प्राप्त विभिन्न परिणामों के परीक्षण के महत्व में भी किया जाता है। यदि प्रेक्षित एवं अनुमानित मानों के मध्य अंतर मानक त्रुटि की तुलना मे 1.96 से ज्यादा होता है तो आप 5% में परिकल्पना को अस्वीकार करते हैं और निष्कर्ष निकालते हैं कि नमूना व्यापक रूप में समग्र से भिन्न है। लेकिन यदि प्रेक्षित एवं अनुमानित मानों के मध्य अन्तर मानक त्रुटि तुलना में 2.58 से ज्यादा है तो आप शून्य परिकल्पना को 1% में अस्वीकार करते हैं और निष्कर्ष निकालते हैं कि नमूना व्यापक रूप में समग्र से भिन्न होता है।

3) अज्ञात समग्र माध्य की विश्वास सीमाओं को निर्धारित करने के लिए :- मानक त्रुटि विश्वास सीमाओं के भीतर जिसमें विश्वास की निश्चित घात के अस्तित्व की अपेक्षा समग्र प्राचल से निर्धारित कर हमें सक्षम बनाती है।

बड़े नमूने के लिए :

μ के 95% विश्वास सीमाओं के लिए - $\bar{X} \pm 1.96$ मानक त्रुटि

μ के 99% विश्वास सीमाओं के लिए - $\bar{X} \pm 2.58$ मानक त्रुटि

छोटे नमूने के लिए :

μ के 95% विश्वास सीमाओं के लिए - $\bar{X} \pm t_{0.05}$ मानक त्रुटि

μ के 99% विश्वास सीमाओं के लिए - $\bar{X} \pm t_{0.01}$ मानक त्रुटि

5.9 माध्यों का नमूना वितरण (Sampling Distribution of Means)

यह ध्यान देना आवश्यक है कि नमूना वितरण का प्रयोग नमूना सिद्धान्त में व्यापक रूप से किया जाता है। प्रतिस्थापन के साथ या के रहित N आकार समग्र जिसका माध्य μ और परिवर्तनशीलता σ^2 है से सभी सम्भव नमूने जिनका आकार n है निकालते हैं। समग्र से निकाले हुए प्रत्येक सभी संभव नमूनों के लिए आप प्रत्येक नमूने के माध्य $\bar{X}$ की गणना करते हैं। माध्य नमूना, नमूना के लिए परिवर्तित होगा। विभिन्न नमूनों से प्राप्त समस्त सम्भव माध्यों की सूची को माध्यों का नमूना वितरण कहते हैं।

माध्यों के नमूना वितरण की कुछ महत्वपूर्ण विशेषताएँ निम्नवत हैंः

1. माध्यों के नमूना वितरण का माध्य समग्र माध्य (μ) के बराबर होता है।

लक्षणिक रूप से, $\mu\bar{x} = \mu$ or $E(\bar{x}) = \mu$ इस गुण को निम्नवत सिद्ध किया जा सकता हैः
N आकार के सीमित समग्र जिसका माध्य μ और परिवर्तनशीलता σ^2 है से $x_1, x_2, x_3, \ldots\ldots x_n$ n आकार का यादृच्छिक नमूना (प्रतिस्थापन के साथ) प्रदर्शित करता है, तो-

$\bar{x} =$

$=$

इस प्रकार माध्यों के नमूना वितरण का माध्य समग्र माध्य के बराबर होता है।

2. **माध्यों के नमूना वितरण की मानक त्रुटि को इस तरीके से प्राप्त किया जाता हैः**

$S.E._{\bar{x}}$ or $\sigma_{\bar{x}} = \quad =$ समग्र का नमूना वितरण

इस गुण को निम्नवत सिद्ध किया जा सकता हैः

$Var(\bar{x}) = Var = Var$

$=$

$=$

$= \quad . n\sigma^2 =$

जहाँ σ^2 परिवर्तनशीलता है, n नमूना आकार है।

क्योंकि, $n>1$, स्पष्टतः $\sigma^2 = > V(\bar{x}) <$

समग्र परिवर्तनशीलता

$\therefore S.E._{\bar{x}}$ or $\sigma_{\bar{x}} = = \quad =$

इस सूत्र का प्रयोग तभी करते हैं जब नमूना वितरण का ध्यान प्रतिस्थापन के साथ रखा जाता है ।

टिप्पणी : जब समग्र निश्चित है और नमूने प्रतिस्थापना के बगैर निकाले जाते है तब $S.E._{\bar{x}}$ को इस तरीके से प्राप्त करते हैं :

$S.E._{\bar{x}} =$

माध्यों का नमूना विवरण लगभग सामान्य विवरण माध्य μ और परिवर्तनशीलता σ^2 के साथ होता है,बर्शते बडा नमूना हो $(n>30)$।

टिप्पणीयाँः माध्यों के नमूना वितरण की प्रायिकता को ज्ञात करने के लिए निम्नलिखित सूत्र का प्रयोग किया जाता है। $Z =$

5.10 बड़ी संख्याओं का नियम एवं केन्द्रीय सीमा प्रमेय (Law of Large Numbers and Central Limit Theorem)

बडी संख्याओं का नियम और केन्द्रीय सीमा प्रयोग दोनों आंकड़ों के विकास की नींव में उपयुक्त होते हैं।

5.10.1 बडी संख्याओं का नियम (Law of Large Numbers):-

बडी संख्याओं का नियम यह बताता है कि जैसे नमूना आकार बढता है नमूना माध्य, समग्र माध्य के और करीब होगा। इसकी आशा नहीं कर सकते कि यदि नमूना आकार पर्याप्त रूप से बढ़ रहा है तो नमूना माध्य समग्र माध्य के बराबर होगा। बडी संख्याओं के नियम की दो उलझने होती है।

नमूना आकार को बढाकर, नमूना माध्य और समग्र माध्य के बीच के अन्तर को कम किया जा सकता है।

एक नमूना माध्य की परिवर्तनशीलता को दूसरे नमूना माध्य (जो समान आकार के हैं) को नमूना आकार बढ़ाकर भी कम किया जाता है।

5.10.2 केन्द्रीय सीमा प्रमेय (Central Limit Theorem):-

इस विधि का प्रयोग व्यापक रूप से अनुमान एवं निष्कर्ष के क्षेत्र मे किया जाता है। यह प्रमेय बताता है कि यदि आप किसी समग्र से यादृच्छिक बड़े आकार का नमूना n चयनित करते हो जिसका माध्य μ और मानक विचलन σ है और प्रत्येक नमूने के माध्य की गणना करते हैं तो माध्य का नमूना वितरण $\bar{x}$ सामान्य वितरण जिसका माध्य μ और मानक विचलन के समीप होता है। यदि समग्र अपने में सामान्य नहीं है तो यह सत्य है। इस प्रमेय की उपयोगिता यह है कि इसमें बिना प्रतिबन्ध के यर्थाथ रूप में समग्र के वितरण तरीके की आवश्यकता होती है।

5.11 सारांश (Summary)

कुल मिलाकार सांख्यिकी आँकड़ों को एकत्रित करने की दो विधियाँ होती हैं- जनगणना विधि और नमूना विधि। वर्तमान में नमूना विधि सांख्यिकी पूछताछ की एक महत्वपूर्ण एवं चर्चित विधि है। अन्वेषक के द्वारा समय और धन की बचत के लिए विभिन्न नमूना विधियों का प्रयोग कर चुनिंदा नमूना लिया जाता है। इसलिए यह कहा जा सकता है कि नमूनाविधि जीवन के सभी क्षेत्रों मे अधिक उपयोगी है।

5.12 शब्दावली (Glossary)

- **गैर प्रायिकता नमूना विधियाँ** : वह विधियाँ हैं जिसमें इकाईयों का चयन प्रायिकता या संभाविता के आधार पर न होकर अन्वेषक के सुविधा या निर्णय के अनुसार होता है।
- **समग्रः** सांख्यिकी में समग्र का आशय संग्रहित (कुल) वस्तुऐं/चीजें जिसके बारे में आप जानकारी प्राप्त करते हैं।

5.13 अभ्यास प्रश्न (Practice Questions)

रिक्त स्थान भरें

1. ....के तहत, इसके विपरीत समग्र की सभी इकाइयों के बारे में जानकारी एकत्रित करने की तुलना में केवल चयनित इकाइयों से संबंधित जानकारी एकत्र की जाती है।

2. समग्र के चयन का एक हिस्से को .... कहा जाता है।

3. जनगणना विधि का प्रयोग वहाँ पर उपयोगी है जहाँ समग्र के इकाईयों मे ....होती है ।

4. ....विधि मे सर्वप्रथम सम्पूर्ण समग्र को स्तरों या उप स्तरों में विभाजित किया जाता है।

5. गैर नमूना त्रुटियाँ वो है जो................................ कारकों से घटित होती हैं ।

5.14 अभ्यास प्रश्नों के उत्तर (Answers to Practice Questions)

1. नमूना विधि 2. नमूना 3. विविधता 4. बहुस्तरीय यादृच्छिक 5. मानवीय

5.15 सन्दर्भ पुस्तकें (Reference Books)

1. Roy Ramendu, ***'Principles of Statistics'***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***"Business Statistics"*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***"Advanced Statistics",*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***"Basic Statistics"*** World Press Limited, Calcutta.
5***. "Fundamentals of Business Statistics"*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **"Quantitative Methods in Management".**

—

इकाई-6
बिन्दु अनुमानक एवं अंतराल अनुमानक संरचना
(Structure of Point Estimators and Interval Estimators)

इकाई की रुपरेखा

उद्देश्य (Objective)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ अनुमान की बुनियादी अवधारणाओं की व्याख्या कर सकें।
- ✓ अच्छे अनुमानों की विशेषताओं की व्याख्या कर सकें।
- ✓ अनुमानों के प्रकारों का वर्णन कर सकें।
- ✓ लक्ष्य आंकलनों की उपयोगिता का वर्णन कर सकें।
- ✓ अन्तर आंकलन की उपयोगिता का वर्णन कर सकें।

6.1 प्रस्तावना (Introduction)

रोजमर्रा की जिन्दगी में, नमूना आँकड़ा से समग्र प्राचल के बारे मे आंकलन करने की आवश्यकता पड़ती है। उदाहरण के लिए, मान लो कि आप किसी विश्वविद्यालय के छात्रों द्वारा प्रतिदिन औसत मात्रा मे पिया हुआ कोका कोला को ज्ञात करने के इच्छुक हैं। सभी छात्रों के औसत मात्रा को ज्ञात करने पर कठिनाई होती है। इस समस्या का हल निकालने के लिए, आप एक नमूना ले सकते हैं कर औसत कोका कौला के औसत मात्रा का पता लगायेंगें सकते है। तब इस नमूना माध्य का प्रयोग समग्र का औसत ज्ञात करने में करेंगें। वास्तव में, नमूना औसत के आधार पर आप समग्र औसत को अनुमानित कर सकते है।

आगणन का सिद्धान्त अज्ञात समग्र प्राचलों के अनुमानो के साथ (जैसे -समग्र माध्य और समग्र परिवर्तनशीलता) समरूपी ऑकड़े नमूनों से (जैसे नमूने माध्य नमूने परिवर्तनशीलता) का वर्णन करता है।

6.2 सांख्यिकी आगणन की बुनियादी अवधारणाएँ (Basic Concepts of Statistical Inference)

सांख्यिकी आगणन के अध्ययन में निम्नलिखित पदों का प्रयोग किया जाता हैः

6.2.1 आगणक एवं अनुमान (Estimators and Estimates)

समग्र प्राचलों का अनुमान लगाने के लिए आप विभिन्न नमूना आंकड़ों का प्रयोग कर सकते हैं जो नमूना आँकड़े जैसे नमूना माध्य $\bar{x}$, नमूना माध्यिका m, नमूना परिवर्तनशीलता S^2 इत्यादि जो अज्ञात समग्र प्राचलों जैसे समग्र माध्य μ, समग्र परिवर्तनशीलता σ^2 आदि का अनुमान लगाते हैं उन्हें आगणक कहा जाता है और आगणक से वास्तविक मान लेने को अनुमानित कहा जाता है।

यदि θ (ठीटा पडें) समग्र पाचल θ का आगणक है।

टिप्पणीयाँ :- तथ्य अनुमान एवं अन्तराल अनुमान

6.2.2 समग्र प्राचल का अनुमान

समग्र प्राचल का अनुमान दो तरीको से किया जा सकता है-

(1) तथ्य अनुमान :- आँकड़ों का एक एकल मान जिसे अज्ञात समग्र प्राचल के अनुमान के लिए प्रयोग किया जाता है उसे तथ्य अनुमान कहते हैं। जैसे नमूना माध्य $\bar{x}$,जिसे μ के तथ्य आगणन के लिए समग्र माध्य μ को अनुमानित करने मे प्रयोग कर सकते है। इसी तरह S^2 ऑकडा σ^2 का तथ्य आंगणक है जबकि S^2 के मान की गणना यादृच्छिक नमूने से करते है। तथ्य आंकलन वास्तविक अंक प्रणाली मे और जिसे तथ्य आगणक कहते हैं एक एकल तथ्य है।

(2) अंतराल अनुमान :- एक अंतराल का अनुमान संभावित सीमा के भीतर प्राचल को संदर्भित करता है जो वास्तविक मान को असत्य ठहराने के लिए अपेक्षित है । इस तरह के विस्तार की दो चरम सीमाओं को विश्वस्त हो विश्वास सीमाएँ कहा जाता है और इस विस्तार को विश्वास अंतराल कहा जाता है।

इनका निर्धारण समग्र के नमूना अध्ययनों के आधार पर किया जाता है। इस प्रकार, नमूना अध्ययनों के आधार पर जब आप छात्रावास में रह रहे छात्रों की औसत मासिक व्यय का अनुमान लगाते हैं जो रू0 5000 और रू0

6000 के बीच है। यह अन्तर अनुमान का एक उदाहरण होगा। और रू0 5,000 और रू0 6,000 की राशि छात्रों की वास्तविक व्यय के भीतर की चरम सीमाओं के अस्तित्व में होगी।

6.3 एक अच्छे आगणक की विशेषताएँ (Properties of a Good Estimator)

समग्र प्राचल में एक से ज्यादा आगणक हो सकते हैं, जैसे समग्र माध्य (μ) का अनुमान या तो नमूना माध्य ($\bar{x}$) या नमूना माध्यिका (m) या नमूना बहुलक के द्वारा किया जा सकता है। इसी तरह, समग्र परिवर्तनशलता (σ^2) का अनुमान या तो नमूना परिवर्तनशीलता (s^2), नमूना मानक विचलन (s), नमूना माध्य विचलन द्वारा किया जा सकता है। इसलिए, प्रापय आगणकों की संख्या के बाहर एक अच्छे आगणक को निर्धारित करना आनिवार्य होता है। एक अच्छा आगणक वह है जो प्राचल के सही सम्भव मानों के समीप होता हैं। अच्छे आगणक में निम्नलिखित लक्षण या विशेषताएँ होती हैं।

6.3.1 निष्पक्ष आगणक (Unbiased Estimator)

निष्पक्ष आगणक को समग्र पाचल θ का निष्पक्ष आगणक कहा जायेगा यदि आगणक के नमूना वितरण का माध्य समग्र प्राचल θ के तुल्य एक समान है।

तथ्य आगणक और अन्तराल आगणक

टिप्पणीयाँ

लाक्षणिक रूप से , $\mu_{\hat{\theta}} = \theta$

गणितीय अपेक्षाओं के मामले में $\hat{\theta}$ एक निष्पक्ष आकलनकर्ता है यदि आकलनकर्ता का अपेक्षित मान प्राचल के अनुमानित मान के बराबर है।

लाक्षणिक रूप से , $E(\hat{\theta}) = \theta$

उदाहरण 6.1 :- नमूना माध्य $\bar{x}$, समग्र माध्य μ, का एक निष्पक्ष आकलनकर्ता है क्योंकि, माध्यों के नमूने वितरण का माध्य $\mu_{\bar{x}}$ या $E(\bar{x})$ समग्र माध्य μ के बराबर है।

लाक्षणिक रूप से , $\mu_{\bar{x}} (\bar{x}) = \mu$ या $E(\bar{x}) = \mu$

उदाहरण 6.2% नमूना परिवर्तनशीलता S^2, समग्र परिवर्तनशीलता σ^2 के झुकाव का एक आकलनकर्ता हे। क्योंकि नमूना वितरण की परिवर्तनशीलता समग्र परिवर्तनशलता के बराबर नहीं है।

लाक्षणिक रूप से $\mu^2_s \neq \sigma^2$ or $E(S^2) \neq \sigma^2$

फिर भी, परिवर्तित नमूना परिवर्तनशीलता (S^2), समग्र परिवर्तनशीलता का एक निष्पक्ष आकलनकर्ता है, क्योंकि

$E(S^2) \neq \sigma^2$ Where $\hat{S}^2 = n/(n-1) \times S^2$

उदाहरण 6.3 :- नमूना अनुपात p समग्र अनुपात P का एक निष्पक्ष आकलनकर्ता है क्योंकि, नमूने वितरण का आनुपातिक माध्य, समग्र आनुपातिक माध्य के बराबर हे।

लाक्षणिक रूप से , $\mu_p = p$ or $E(p) = p$

6.3.2 समान आकलनकर्ता (Uniform Estimator)

एक आकलनकर्ता को एक समान कहेंगें, यदि नमूना आकार बढ़े तो आकलनकर्ता का दृष्टिकोण समग्र प्राचल है। दूसरे शब्दों में, एक आकलनकर्ता $\hat{\theta}$ को समग्र प्राचल θ का एक समान आकलनकर्ता कहेंगें यदि n बहुत बडा है तो $\hat{\theta}$ की प्रायिकता का दृष्टिकोण θ के लिए 1 है।

लाक्षणिक रूप से , $P(\hat{\theta} \to \theta) \to 1$ as $n \to \infty$

टिप्पणी :- एक समान आकलनकर्ता की निष्पक्ष होने की जरूरत है। एक आकलनकर्ता के एक समान होने का पर्याप्त लक्षण यह है कि

1. $E(\hat{\theta}) \to \theta$
2. $Var(\hat{\theta}) \to 0$ as $n \to \infty$

उदाहरण 6.4 :- नमूना $\bar{x}$, समग्र माध्य μ का एक समान आकलन कर्ता है क्योंकि नमूने माध्य के अपेक्षित मान का दृष्टिकोण समग्र माध्य है और यदि नमूने का आकार पर्याप्त रूप से बढाये तो नमूने माध्य की परिवर्तनशीलता का दृष्टिकोण शून्य है।

लाक्षणिक रूप से ,

i. $E(\bar{x}) \to 1$
ii. $Var(\bar{x}) = \sigma^2/n \quad \to 0 \quad as \quad n\to\infty$

उदाहरण 6.5 : नमूना माध्यिका भी समग्र माध्य का एक समान आकलनकर्ता है क्योंकि

i. $E(m) \to 1$
ii. $Var(m) = 0 \quad as \quad n \to \infty$

6.3.3 सक्षम आकलनकर्ता () :-

सक्षमता एक सापेक्ष शब्द है। आमतौर पर एक आकलनकर्ता की सक्षमता को दूसरे आकलनकर्ता द्वारा तुलना कर परिभाषित किया जाता है। मान लो कि आप θ के दो निष्पक्ष आकलनकर्ता $\hat{\theta}_1$ और $\hat{\theta}_2$ का संज्ञान ले रहे हैं। इनमे से $\hat{\theta}_1$ आकलनकर्ता को θ को सक्षम आकलनकर्ता कहा जायेगा यदि $\hat{\theta}_1$ की परिवर्तनशीलता, $\hat{\theta}_2$ की परिवर्तनशीलता से कम है।
लाक्षणिक रूप से , $Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$
तब $\hat{\theta}_1$ को एक सक्षम आकलनकर्ता कहा जाता है।

6.3.4 सक्षम आकलनकर्ता (Competent Estimator)

एक अच्छे आगणक की अन्तिम विशेषता उसकी सक्षमता होनी चाहिए। एक आकलनकर्ता $\hat{\theta}$ को θ का सक्षम आकलनकर्ता कहा जाता है। यदि यह प्राचल के संदर्भ में, नमूने में सभी सूचनाओं को सम्मलित करता है। दूसरे शब्दों में एक सक्षम आकलनकर्ता, समग्र के बारे मे नमूने मे उपस्थिति सभी सूचनाओं का प्रयोग करके उसे प्रस्तुत करता है। नमूना माध्य $\bar{x}$ को समग्र माध्य μ का एक सक्षम आकलनकर्ता कहा जाता है।

6.4 तथ्य आकलनकर्ता का उपयोग (Use of Fact Estimator)

अब आप तथ्य आकलनकर्ता के उपयोगों का अध्ययन करेंगें जो निम्नवत है:

6.4.1 एकल नमूने की स्थिति में तथ्य आकलनकर्ता (Fact Estimator in Single Sample Situation)

जब अज्ञात समग्र से एक एकल स्वतन्त्र यादृच्छिक नमूना निकालते हैं तो उसे एकल नमूना कहते है। समग्र पाचल के तथ्य आकलनकर्ता की व्याख्या निम्न उदाहरणों से दी जा सकती है।
उदाहरण 6.6% एक गोले के व्यास (मोटाई) के 10 नामों के नमूने का माध्य $\bar{x} = 4.38$ इंच और परिवर्तनशीलता = 0.06 इंच दी गई है।
(अ) सच्चे/माध्य (अर्थात् समग्र माध्य) और (ब) सही परिवर्तनशीलता (अर्थात समग्र परिवर्तनशलता) के निष्पक्ष एवं सखम अनुमान ज्ञात कीजिए ।-
हल : आप को $n=10$, $\bar{x} = 4.38$, $s^2 = 0.6$ इंच दिया गया है।

(अ) सही माध्य (μ) का निष्पक्ष एवं सक्षम अनुमान $\bar{x} = 4.38$ होगा।
(ब) सही परिवर्तनशीलता σ2 का निष्पक्ष एवं सक्षम अनुमान
$\hat{s}^2 = n/n-1.s^2$ है।
इसमें मानों को रखकर आप

$\hat{s}^2 = 10/10-1 \times 0.06 = 1.11\times 0.06 = 0.066$ प्राप्त करते है।

इस प्रकार $\mu = 4.38$, $\sigma'^2 = 0.666$

6.4.2 नमूना पुनरावृत्ति की घटना में तथ्य आंकलन (Fact Estimation in Case of Sample Repetition)

प्रतिस्थापना के साथ या बिना, समग्र से जब एक ही आकार के एक से ज्यादा यादृच्छिक नमूने निकाले जाते हैं तो उन्हें नमूना पुनरावृत्ति कहा जाता है। इसे निम्नलिखित उदाहरणों से समझा जा सकता है।

उदाहरण 6.7: एक समग्र में पांच मान: 3,4,5,6 और 7 शामिल है। प्रतिस्थापना के बिना समग्र से आकार 3 के सभी सम्ीव नमूनों की सूची बनाएं और प्रत्येक नमूने के माध्य $\bar{x}$ की गणना करें। वह नमूना माध्य समग्र माध्य का एक निष्पक्ष अनुमान है, की जाँच करें।

हलः समग्र मे 5 मान : 3,4,5,6 और 7 शामिल है। प्रतिस्थापना के बिना आकार 3 के सभी सम्भव नमूनों की संख्या $^5c_3$ = 10 है जो निम्नलिखित तालिका में प्रदर्शित किए जा रहे हैं।

Sample No.	Sample value	Sample Mean ($\bar{X}$)
1	(3,4,5,)	1/3(3+4+5) =12/3*4
2	(3,4,6)	1/3(3+4+6)= 13/3*4.33
3	(3,4,7)	1/3(3+4+7) = 14/3*4.67
4	(3,5,6)	1/3(3+5+6) =14/3*4.67
5	(3,5,7)	1/3(3+5+7) = 15/3*5
6	(3,6,7)	1/3(3+6+7) = 16/3*5.33
7	(3,5,6)	1/3(3+5+6) = 15/3*5
8	(3,5,7)	1/3(3+5+7) = 16/3*5.33
9	(3,6,7)	1/3(3+6+7) = 17/3*5.67
10	(3,6,7)	1/3(3+6+7) =18/3*6
Total	K=10	$\sum \bar{x}$ = 50

माध्यों के नमूना वितरण का माध्य = $\mu_{\bar{x}} = \sum \bar{x}/x$ = 50/10 = 5s

माध्यों के नमूने वितरण का माध्य = $\mu_{\bar{X}} = \sum \bar{x}/k$ = 50/10 = 5

समग्र माध्य (μ) = (3+4+5+6+7)/5 = 5

इसलिए निश्चित रूप से कहा जा सकता है कि $\mu_{\bar{x}} = \mu$ नमूना माध्य $\bar{x}$ समग्र माध्य का एक निष्पक्ष अनुमान है।

उदाहरण 6.8: दर्शाए कि नमूना माध्य $\bar{x}$, समग्र माध्य का एक निष्पक्ष अनुमान है।

या

एक स्वतन्त्र यादृच्छिक नमूना $x_1, x_2, x_3, \ldots, x_n$ समग्र से जिसका माध्य μ है से निकाला जाता है। सिद्ध करें नमूना माध्य $\bar{x}$ का अपेक्षित मान समग्र माध्य μ के बराबर है।

हलः यादृच्छिक नमूना वह होता है जहाँ प्रत्येक नमूने के चयनित होने के बराबर मौके होते हैं आप आकार n के यादृच्छिक नमूने प्राप्त कर सकते है। तब

$$E(\bar{x}) = E\{(x_1 + x_2 + x_3 + \ldots + x_n)/n\}$$

जहाँ x_1 नमूना प्रेक्षण हैं।

$= 1/n \ [E(x_{1)} + E(x_2) + \cdots E(x_n)]$

अब x_i (जो समग्र का एक सदस्य हैं) का अपेक्षित मान समग्र माध्य μ हैं। इसलिए

$E(\overline{x}) = 1/n \ [\mu + \cdots + \mu]$, क्योंकि $[E(x_{1)} = E(x_2) = E(x_{n)} = \mu]$, $= 1/n * [n \ \mu] = \mu$

क्योंकि $[\sum C = c_1 + c_2 + \ldots\ldots + c_n = nC]$,

इस प्रकार नमूना माध्य $\overline{x}$ समग्र माध्य का एक निष्पक्ष अनुमान है।

6.5 अन्तराल आंकलन या अन्तराल विश्वास (Interval Estimator or Confidence Interval)

अन्तराल आंकलन के सिद्धान्त में, आप एक अन्तराल या दो अंकों के भीतर जिसमें अज्ञात समग्र प्राचल के अपेक्षित मान का अस्तित्व प्रायिकता के साथ दर्शाते हुए ज्ञात कर सकते हैं।

दो नियत राशियों t_1 और t_2 के निर्धारण मे अन्तराल आंकलन विधि इस तरीके से शामिल होती है कि $[t_1 < \theta < t_2$, t के दिये हुए मान के लिए], $= 1-\propto$ जहाँ $\propto$ एक स्तर का महत्व है।

t_1 और t_2, का अंतराल जिसके भीतर प्राचल θ के अपेक्षित अज्ञात मान का अस्तित्व हो को विश्वास अन्तराल कहते हैं और ज्ञात की गई सीमाएँ t_1 और t_2 को विश्वास सीमाएँ कहते हैं और $1-\propto$ को अनुमान का वांछित यर्थाथमापी आधारित विश्वास गुणांक कहते हैं। जैसे 2= 0.5 (या 0.01) 95% (या 99%) विश्वास सीमाऐं देता है। अब, आप विश्वास सीमा (या अंतराल आंकलन) की पद्धति की स्थापना का या समग्र प्राचल की सीमाओं का अध्ययन करेंगें।
समग्र प्राचल θ के लिए नमूना आकड़ा ज से संबन्धित विश्वास अंतराल या विश्वास सीमाओं की गणना के उद्देश्य से निम्नलिखित चरण आपको सक्षम बनाते हैं।

उपयुक्त नमूना आँकड़ों ज की गणना करें या लें।

मानक त्रुटि ज नमूना आंकडे की मानक त्रुटि ज ज्ञात करें और

विश्वास स्तर का चयन करें और समरूपी दर्शाये गये विश्वास स्तर का आंकड़ों ज के समीक्षात्मक मान को लिखें।

6.6 अन्तराल आंकलन के अनुप्रयोग (Application of Interval Estimators)

अन्तराल आंकलन (या विश्वास अन्तराल) से संबन्धित अनुप्रयोगों को निम्नलिखित शीर्षकों के अर्न्तगत अध्ययन करते हैं।

6.6.1 बडे नमूनों ($n \geq 30$) के लिए अन्तराल आंकलन (या विश्वास अन्तराल)

बडें नमूनों के लिए आकलन अंतराल को और आगे निम्नलिखित शीर्षकों के अर्न्तगत विभाजित किया जा सकता है।

विश्वास अंतराल या समग्र माध्य के लिए सीमाऐं

विश्वास अंतराल या समग्र अनुपात के लिए सीमाऐं

विश्वास अंतराल या समग्र मानक विचलन के लिए सीमाऐं

μ या p के अनुमान हेतु उचित नमूना आकार का निर्धारण

(1) विश्वास अंतराल या समग्र माध्य μ के लिए सीमाऐं जब ($n \geq 30$)

बडें नमूने की स्थिति में ($n \geq 30$) की सीमाओं के निर्धारण मे सामान्य वितरण के प्रयोग की आवश्यकता होती है।

(1) μ के लिए (1-$\propto$) 100% विश्वास सीमाऐं $\bar{x} \pm Z_{a/2}$ S.E.$_x$ से दी जाती है।

टिप्पणीयाँ या $\bar{x} \pm Z_{a/2}$. s/√n जहाँ σ ज्ञात है।

या $\bar{x} \pm Z_{a/2}$. s/√n σ अज्ञात है। (बडे नमूनों के लिए, σ=s)

(2) μ के लिए (1- a) 100% विश्वास सीमाऐं

$\bar{x} \pm Z_{a/2}$. s/√n < μ < $\bar{x} \pm Z_{a/2}$. s/√n जहाँ σ ज्ञात है।

or $\bar{x} \pm Z_{a/2}$. s/√n < i < $\bar{x} \pm Z_{a/2}$. s/√n जहाँ σ अज्ञात है।

विशेष रूप से, μ के लिए 95% विश्वास सीमाऐं

$\bar{x} \pm 1.96$ ó/√n [बडें नमूनों के लिए ó =s]

इसी तरह μ के लिए 99% विश्वास सीमाऐं

$\bar{x} \pm 2.58$ ó/√n

कार्य विधिः समग्र माध्य μ के लिए विश्वास अंतराल के निर्माण हेतु निम्नलिखित चरण शामिल हैं :

1. $\bar{x}$ की गणना करें या $\bar{x}$ लें।
2. निम्नलिखित सूत्र का प्रयोग करते हुए S.E.$_{\bar{x}}$
   a) S.E.$_{\bar{x}}$ = ó/√n , जब σ ज्ञात है।
   b) S.E.$_{\bar{x}}$ = s/√n, जब σ अज्ञात है।

आप $Z_{\propto/2}$ का मान ज्ञात करेंगे तो इसके लिए

3. वांछित विश्वास अंतराल या समतुल्य विश्वास स्तर चयनित करें।

   उपरोक्त वर्णित सूत्र में वि $\bar{x}$, S.E.$_{\bar{x}}$ और $Z_{\propto/2}$ और $Z_{\propto/2}$ के मानो को प्रतिस्थिपत करें।

टिप्पणीयाँः (1) यदि समग्र S.D. अज्ञात है तो नमूना S.D.(S) को बडें नमूनों में प्रयोग किया जाता है।

2) $Z_{\propto/2}$ के मानों (बडें नमूनों के लिए) को विभिन्न विश्वास स्तर पर निम्नवत दिया जा रहा हैः

विश्वास स्तर (1-α) 100%	90%	95%	96%	98%	99%	बिना किसी संदर्भ के विश्वास स्तर ±3
Z-Value	±1.64	±1.96	±2.06	±2.33	±2.58	

टिप्पणी : जहाँ विश्वास अन्तराल का संदर्भ नहीं दिया गया हो तो आपको $Z_{\propto/2}$ = 3 लेना चाहिए। यह मान 99.73% विश्वास स्तर के समतुल्य है।

उदाहरण 6.9: 100 प्रेक्षणों का यादृच्छिक नमूना, नमूना माध्य $\bar{x}$ = 150 और नमूना परिवर्तनशीलता S^2 = 400 देता है। समग्र माध्य के लिए 95 प्रतिशत और 99 प्रतिशत विश्वसनीयता अंतराल की गणना करें।
हलः आपकों n =100, $\bar{x}$ = 150, s^2= 400 ⇒ s =20 दिया गया है।
S.E.$_{\bar{x}}$ = s/√n, (बडे नमूनों के लिए σ= s)
= 20/√100 = 2
95% विश्वसनीयता स्तर पर $Z_{\propto/2}$ मान = 1.96
99% विश्वसनीयता स्तर पर $Z_{\propto/2}$ का मान = 2.58

(अ) μ के लिए 95% विश्वसनीयता अंतराल या सीमाऐं

$$\bar{X} \pm 1.96\ S.E._x$$

मानों को रखने, आप प्राप्त करेंगें

150 ± 1.96 ×2 = 150± 3.92 =153.92 or 146.08

इस प्रकार $146.08 < \mu < 153.92$

(ब) μ के लिए 99% विश्वसनीयता अंतराल या सीमाऐं

$$\bar{X} \pm 2.58\ S.E._x$$
$$= 150 \pm 2.58 * 2$$
$$= 150 + 5.16$$
$$= 155.16 \text{ or } 144.84$$

इस प्रकार $144.84 < \mu < 155.16$

2. समग्र अनुपात p के लिए विश्वसनीयता अंतराल या सीमाऐं

यद्यपि नमूना वितरण अनुपातों के साथ द्विपद वितरण से संबंधित है लेकिन, सामान्य वितरण निकटता का प्रयोग किया जा सकता है बर्शतें नमूना बडा है (जैसे $n \geq 30$) और np और $nq > 5$ (जब n नमूने का आकार है , p सफलता का अनुपात है और $q = 1-p$)

(1) P के लिए (1-á) 100% विश्वसनीयता सीमाऐं

$$p \pm Z_{\alpha/2}\ .\ S.E.(p)$$

या $p \pm Z_{\alpha/2} \sqrt{(PQ/n)}$ जब P ज्ञात है।

या $p \pm Z_{\alpha/2} \sqrt{(pq/n)}$ जब P अज्ञात है।

दी जा रही है।

(2) P के लिए (1- á) 100% विश्वसनीयता अंतराल

$p - Z_{\alpha/2} \sqrt{(pq/n)} < P < p + Z_{\alpha/2} \sqrt{(pq/n)}$ दी जा रही है।

P के लिए 95% विश्वसनीयता सीमाऐं

$P \pm 1.96 \sqrt{(pq/n)}$ है।

P के लिए 99% विश्वसनीयता सीमाऐं

$P \pm 2.58 \sqrt{(pq/n)}$ है।

कार्यविधिः- समग्र अनुपात के विश्वसनीयता सीमाऐं या अंतराल के लिए निर्माण के लिए निम्नलिखित चरण शामिल है।

1. p की गणना करें या लें

2. S.E.(p) की गणना निम्नलिखित सूत्र द्वारा करें

$S.E.(p) = \sqrt{(PQ/n)}$ जब θ ज्ञात है।

$S.E.(p) = \sqrt{(pq/n)}$ जब P अज्ञात हैं।

3. आप $Z_{\alpha/2}$ का मान ज्ञात करेंगे तो इसके लिए वांछित विश्वसनीयता अंतराल या समतुल्य विश्वसनीयता स्तर चयनित करें।

4. उपरोक्त वर्णित सूत्र में p, $Z_{\alpha/2}$ के मानों को प्रतिस्थापित करें।

टिप्पणी :-

1) यदि समग्र अनुपात p अज्ञात है, तो बडे नमूनों मे नमूना अनुपात p का प्रयोग किया जाता है।
2) जब विश्वास अन्तराल का संदर्भ नहीं दिया गया है तो हमेशा 99.73% विश्वसनीयता स्तर के लिए $Z_{\alpha/2} = 3$ लें।
उदाहरण 6.10: एक सिक्के को 1200 बार उछालने पर, इसमें से 480 चित् और 720 पट निकले। चित्तों के लिए 95% पर पर विश्वासनीयता अंतरात ज्ञात करें।

हलः आपको n =1200, कुल चिट (np) = 480 दिये गए है।

P = चित्तों का नमूना अनुपात 480/1200 = 0.4

टिप्पणीयाँ चित्तों का समग्र अनुपात भी = p =0.50

q = 1 - p = 1 - 0.50 = 0.50
S.E.(p) = √(PQ/n) (बडे नमूनों के लिए q = p)
=√[(0.5×0.5)/1200]त् = 0.0144
95% विश्वसनीयता स्तर के लिए $Z_{\alpha/2}$ का मान = 1.96
p के लिए 95% विश्वसनीयता अन्तराल:
P± 1.96 $S.E._p$ दिया जा रहा है।
इन मानों को उक्त में रखने पर
= 0.4± 1.96 *0.0144
= 0.4 ± 0.028 = 0.372 or 0.4258 प्राप्त करेंगें
इस प्रकार, 0.372 < P < 0.4258
उदाहरण 6.11: एक बड़ें प्रेषित माल से 600 अन्नानासों का एक यादृच्छिक नमूना लिया गया था। उनमें 75 खराब पाये गये थे। प्रेषित माल में खराब अन्नानासों के अनुपात का आंकलन करें औश्र् अनुमान की मानक त्रुटि बताऐं।

सीमाओं के भीतर जो प्रेषित माल में इक्कठे हुए खराब अन्नानासों के प्रतिशत का निर्धारण करें।

हल : आपको n = 600, खराब अन्नानासों की संख्या (np =75 दी गई है।

नमूना अनुपात

p =75/600 =0.125 = 12.5%

q = 1 -0.125 = 0.875

$S.E._p$ = √(pq/n) =√[(0.125×0.875)/600] = 0.013 (P अज्ञात है।)

क्योंकि विश्वसनीयता स्तर वर्णित नहीं किया गया है तो आप इसके 99.73% विश्वसनीयता स्तर पर $Z_{\alpha/2}$ का मान = 3

P के लिए 99.73% विश्वसनीयता सीमाऐं p ± 3× $S.E._{\bar{x}}$ दी जा रही है।

उक्त में मानों को रखने पर, आप

0.125 ± 3 *0.013
=0.125 ± 0.039
= 0.164 ± 0.086 से प्राप्त करेंगें।
अतः आवश्यक प्रतिशत 16.4% और 8.6% के (मध्य) बीच है।

जब नमूनें को बिना प्रतिस्थापना के परिमित समग्र से निकाला जाता है, तो समग्र अनुपात p का विश्वसनीयता अन्तराल या सीमाएँ :

इस घटना में (1-∝)100% विश्वसनीयता अन्तराल या सीमाऐं निम्न सूत्र से दी जाती है -

$$p \pm Z_{\alpha/2} \cdot \sqrt{(pq/n)} \cdot \sqrt{[(N-n)/N-1]}$$

जहाँ $\sqrt{[(N-n)/N-1]}$ परिमित समग्र शुद्धि गुणक है।

टिप्पणी : यदि m नमूना आकार n की तुलना में पर्याप्त रूप से बड़ा है तो परिमित समग्र के शुद्धि गुणक को उपेक्षित किया जा सकता है।

उदाहरण 6.12: 2,000 ग्राहकों के बहीखातों में से 600 नमूनों को प्रविश्टि एवं संतुलन की शुद्धता के परीक्षण के लिए लिया गया था जिसमें 45 गलतियाँ पायी गई थी । सीमाओं के भतर जिसमें त्रुटिपूर्ण नमूनों की संख्या की अपेक्षा 95% स्तर पर की जा सकती है, निर्धारित करें।

हलः आपको n = 600, N = 20,000, नमूना बहीखाते में गलतियों की संख्या (np=45) दी गई है।

नमूना अनुपात $p = [(np)/n] = 45/600 = 0.075$

$q = 1 - p = 1 - 0.075 = 0.925$

क्योंकि N नमूना आकार n की तुलना में पर्याप्त रूप से बडा है, तो परिमित समग्र शुद्धि गुणक $\sqrt{[(N-n)/N-1]}$ को उपेक्षित किया जा सकता है।

अतः इसे परिमित (बडा) समग्र का एक नमूना समझे, p की मानक त्रुटि

$S.E._p = \sqrt{(pq)/n}$ द्वारा दी जा रही है।

$= \sqrt{[(0.075 \times 0.928)/600]} = \sqrt{0.0001156} = 0.011$ (लगभग)

95% विश्वसनीयता स्तर पर $Z_{\alpha/2}$ का मान = 1.96

समग्र p के लिए 95% विश्वसनीयता सीमाओं को $P \pm 1.96\ S.E._{\bar{x}}$ द्वारा दिया जा रहा है। इन मानों को रखने पर, आप

$= 0.075 \pm 0.022 = (0.053, 0.097)$ प्राप्त करते हैं।

अतः 2,000 समूह के त्रुटिपूर्ण नमूनों की संख्या के घटित होने की अपेक्षा 20,000 * 0.053 और 20,000* 0.095 के मध्य है या 1060 और 1940 ।

टिप्पणी : यदि परिमित शुद्धि गुणक को उपेक्षित नहीं किया जाता है तब p की 95% विश्वसनीयता सीमाऐं

$P \pm 1.96 \sqrt{(pq/n)} \cdot \sqrt{[(N-n)/(N-1)]}$ है

$$= 0.75 \pm 1.96\sqrt{[(0.075 \times 0.925)/600]} \times \sqrt{[(20,000.600)/(20,000.1)]}$$
$$= 0.75 \pm 1.96 \times 0.0108$$
$$= 0.075 \pm 0.021168$$
$$= (0.0538,\ 0.096168)$$

अतः समूह में आवश्यक त्रुटिपूर्ण नमूनों की संख्या 20,000×0.0538 और 20,000×0.096168 या 1076 और 1924 के मध्य होगी।

(3) समग्रों के मानक विचलन की विश्वसनीयता अन्तराल या सीमाऐं

समग्र के मानक विचलन σ का विश्वसनीयता स्तर निर्धारण में मानक वितरण का प्रयोग आवश्यक है जब नमूना बडा (n≥30) होता है।

σ के लिए (1-á) 100% विश्वसनीयता सीमाऐं

$$s \pm Z_{\alpha/2} \cdot (S.E._s)$$

या $s \pm Z_{\alpha/2} \cdot [á/\sqrt{2n}]$ जब σ ज्ञात है।

या $s \pm Z_{\alpha/2} \cdot [s/\sqrt{2n}]$ जब σ अज्ञात है।

σ के लिए (1- á) 100% विश्वसनीयता अन्तराल

$$s - Z_{\alpha/2} \cdot [s/\sqrt{2n}] < \sigma < s + Z_{\alpha/2} \cdot [s/\sqrt{2n}]$$

σ के लिए 95% विश्वसनीयता सीमाऐं

$$s \pm 1.96 \, [s/\sqrt{2n}]$$ [बड़े नमूने के लिए $s = \sigma$, है।

σ के लिए 99% विश्वसनीयता सीमाऐं

$$s \pm 2.58 \, [s/\sqrt{2n}]$$ है।

कार्य विधिः σ के विश्वसनीयता सीमाओं के निर्माण मे निम्नलिखित चरण शामिल हैंः

1. S की गणना करें या s लें।
2. निम्नलिखित सूत्र का प्रयोग करते हुए S.E.(s) की गणना करें
   $S.E.(s) = \sigma/\sqrt{2n}$ या $S.E.(s) = s/\sqrt{2n}$
3. वांछित विश्वसनीयता स्तर चयनित करें और $Z_{\alpha/2}$ के समतुल्य विश्वसनीयता स्तर का मान।
4. उपरोक्त वर्णित सूत्र में s, $Z_{\alpha/2}$ और n के मानों को प्रतिस्थापित करें।

(4) µ या p के आंकलन हेतु नमूने आकार का निर्धारण

ज्ञात नमूना आकार के लिए मान्यताओं के आधार पर आपने विश्वसनीयता अन्तरालों की गणना की है। अब आप जब नमूना आकार अज्ञात है, विश्वसनीयता स्तर की गणना करने में निपुण होंगें।

समग्र माध्य के आंकलन के लिए नमूना आकार :

समग्र माध्य आंकलन के लिए, नमूने आकार के निर्धारण हेतु निम्नलिखित तीन कारक ज्ञात होने चाहिए।

टिप्पणीयाँ :

वांछित विश्वसनीयता स्तर और Z का समतुल्य मान ।

अनुज्ञेय नमूना त्रुटि E

σ का मानक विचलन या σ का एक अनुमान (जैसे $\bar{s}$)

उपरोक्त कारकों को जानने के बाद , नमूना आकार n निम्नवत द्वारा दिया जाता है।

$$n = (Z. \sigma / E)^2$$

टिप्पणी :

(1) Z और E के मान पूर्वनिश्चित होते हैं।

(2) समग्र का मानक विचलन ((S.D.) वास्तविक या अनुमानित हो सकता है।

उदाहरण 6.13: एक सिगरेट उत्पादक (निर्माता) यादृच्छिक नमूने का प्रयोग कर औसत, सम्बाकू यात्रा का अनुमान लगाना चाहता है। नमूने में त्रुटि 99 प्रतिशत विश्वसनीयता स्तर पर सच्चे माध्य की तुलना में एक मिलीग्राम से कम या ज्यादा नहीं होनी चाहिए। समग्र मानक विचलन 4 मिलीग्राम है। इन आवश्यकताओं की पूर्ति के लिए कम्पनी को कितने नमूने आकार का प्रयोग करना चाहिए।

हलः आपको E = 1, 99% विश्वसनीयता स्तर पर $Z_{\alpha/2} = 2.58$ और $\sigma = 4$ दिया गया है। नमूना आकार सूत्र $n = (z^2.\sigma^2)/ E^2$ है।

इन मानों को प्रतिस्थापित करने पर, आप प्राप्त करेंगे।

$$n = (2.58)^2 \times (4)^2 / 1^2 = 106.50 \text{ or } 107$$

अतः कम्पनी की अवश्यकताओं की पूर्ति के लिए आवश्यक नमूना आकार n = 107 होना चाहिए।

(ब) समग्र अनुपात के आंकलन के लिए नमूना आकार :

समग्र अनुपात का अनुमान ज्ञात करने के लिए नमूना आकार हेतु निम्नलिखित तीन कारक ज्ञात होने चाहिए।

1. वांछित विश्वसनीयता स्तर और त्र का समतुल्य मान ।
2. अनुज्ञेय नमूना त्रुटि E ।
3. सफलता p का वास्तविक या अनुमानित सच्चा अनुपात ।

नमूना आकार $n = z^2 \times PQ/E^2$ जहाँ, $Q = 1 - P$ से दिया जाता है।

टिप्पणी :

1. Z और E के मान पूर्वनिश्चित हैं।

2. समग्र अनुपात च का मान वास्तवकि या अनुमानित हो सकता है।

उदाहरण 6.14: एक व्यवसाय अधिकतम मान्य त्रुटि 0.5 के साथ और 98 प्रतिशत विश्वसनीयता स्तर पर उपभोक्ताओं के अनुपात को, जो इनके उत्पादों को पसंद करते हैं, को ज्ञात करना चाहता है। यदि प्रारम्भिक बिक्र्री सूचनाऐं दर्शाते हैं कि सभी उपभोक्ताओं में से 25 प्रतिशत व्यवसाय उत्पाद को पसंद करते हैं तो कितने बड़े नमूने की आवश्यकता होगी?

हलः आपको दिया गया हैः $E = 0.05$, $P = 0.25$, $Q = 1-0.25 = 0.75$

98% विश्वसनीयता स्तर के लिए $Z = 2.33$

नमूना आकार सूत्र $n = z^2 \times PQ/E^2$

इन मानों को उक्त सूत्र में प्रतिस्थापित करने पर आप प्राप्त करेंगें।

$n = (2.33)^2/(0.05)^2 \times (0.025)(0.75)$

$= 5.429/0.0025 \times (0.1875)$ $= 1.0179/0.0025 = 407.16$ or 408

अतः आवश्यक नमूना आकार दत्र 408 होगा ।

6.6.2 छोटे नमूनों (n<30) के लिए अन्तराल

छोटे आकार के नमूनों (n<30) की स्थिति में विश्वसनीयता स्तरों का निर्धारण का अध्ययन दो शीर्षकों के अन्तर्गत किया जाता है।

(1) समग्र माध्य (n<30) के लिए विश्वसनीयता स्तर या सीमाऐं :

जब नमूनो का आकार छोटा (जैसे (n<30) और σ (समग्र मानक विचलन) अज्ञात हैं तो समग्र माध्य μ के लिए वांछित विश्वसनीयता स्तर या सीमाऐं को t- वितरण का प्रयोग कर प्राप्त किया जा सकता है। छोटे नमूनों की स्थिति में Z मानों के वदले में t मानों का प्रयोग करते हैं।

1. समग्र माध्य के लिए (1-∝) 100% विश्वसनीय अन्तराल के द्वारा दी जाती हैः

$\bar{x} \pm t_{\propto/2} \,.\, s\hat{}/n$ जहाँ $s\hat{}$ = परिवर्तित नमूना

$S.D. = \sqrt{[(\Sigma(x-\bar{x})^2/n-1]}$ or $s\hat{} = \sqrt{(n/(n-1)s^2}$

2. μ के लिए (1.ह्यद्ध 100% विश्वसनीय अन्तराल के द्वारा दी जाती हैः

$\bar{x} \pm t_{\propto/2} \,.\, s\hat{}/\sqrt{n} \;<\; I \;<\; \bar{x} \pm t_{\propto/2} \,.\, s\hat{}/\sqrt{n}$

μ के लिए 95% विश्वसनीयता सीमाऐं के द्वारा दिया जाता है।

$\bar{x} \pm t_{0.025} \,.\, s\hat{}/\sqrt{n}$

μ के लिए 99% विश्वसनीयता सीमाऐं के द्वारा दिया जाता है।

$\bar{x} \pm t_{0.005} \,.\, s\hat{}/\sqrt{n}$

कार्य विधिः छोटे नमूनों (n<30) की स्थिति में विश्वसनीयता अंतराल या सीमाओं के निर्माण में निम्नलिखित चरण शामिल हैं।

1. $\bar{x}$ की गणना करें या $\bar{x}$ लें।
2. निम्नलिखित सूत्र का प्रयोग करते हुए परिवर्तित नमूने वितरण की गणना करें

   $S\hat{} = \sqrt{(n/(n-1)s^2}$ जब s दिया गया हो।
3. d.o.f. = v = n-1 सूत्र का प्रयोग करते हुए Degree of Freedom की गणना करें।
4. दिये गये Degree of Freedom के लिए वांछित विश्वसनीयता स्तर और दर्शायी गई समतुल्य विश्वसनीयता स्तर को चयनित करें। आपको $t_{\propto/2}$ का मान t तालिका से देखना चाहिए।
5. उपरोक्त वर्णित सूत्र में ग ऐ और जत्रः(∝ध्2 द्ध के मानों को प्रतिस्थापित करें।

उदाहरण 6.15: 16 आकार के यादृच्छिक नमूनों का मानक विचलन 3 के साथ माध्य 50 है। 98 प्रतिशत विश्वसनीयता सीमाओं पर समग्र का माध्य प्राप्त करें।

हलः आपको दिया गया है : $n = 16$, $\bar{x} = 50$, $s = 3 \Longrightarrow s^2 = 9$

$s\hat{} = \sqrt{(n/(n-1)s^2}$ $= \sqrt{(16/(16-1)9}$

degree of freedom = v = n =16 -1 = 15

98% विश्वसनीयता स्तर पर $\propto$=0.02 so that $\propto/2 = 0.02/2 = 0.01$

t तालिका का प्रयोग करते हुए 15 d.f. के लिए $t_{0.01} = 2.602$

μ के लिए 98% विश्वसनीयता सीमाऐं के द्वारा दी जाती है

$$\bar{x} \pm t_{0.025} . s\hat{}/\sqrt{n}$$

इन मानों को उक्त सूत्र में रखने पर, आप प्राप्त करेंगें

$$=50\pm 2.602\times3.098/\sqrt{16}$$
$$=50\pm2.015$$
$$=52.015 \text{ to } 47.985$$

(2) समग्र परिवर्तनशीलता के लिए (n<30) विश्वसनीयता अंतराल या सीमाऐं

समग्र परिवर्तनशीलता σ^2 के विश्वसनीयता स्तर या सीमाओं के निर्धारण में x^2 (काई वर्ग) वितरण का प्रयोग आवश्यक है। यहाँ पर x^2 मानों का प्रयोग t मानों के बदले में होता है।

समग्र परिवर्तनशीलता σ^2 के $(1-\propto)$ 100% विश्वसनीयता अन्तराल द्वारा दिया जाता हे।

$$[(n-1)s\hat{}^2]/x^2_{\propto/2} < ó^2 < [(n-1)s\hat{}^2]/x^2_{1-\propto/2}$$

विशेष रूप से, समग्र परिवर्तनशीलता σ^2 के लिए 95% विश्वसनीयता अंतराल है।

$$[(n-1)s\hat{}^2]/x^2_{0.025} < ó^2 < [(n-1)s\hat{}^2]/x^2_{0.975}$$

इसी तरह, समग्र परिवर्तनशीलता σ^2 के लिए 99% विश्वसनीयता अंतराल है।

$$[(n-1)s\hat{}^2]/x^2_{0.005} < ó^2 < [(n-1)s\hat{}^2]/x^2_{0.995}$$

कार्य विधिः परिवर्तनशीलता σ^2 के विश्वसनीयता स्तर के निर्माण के लिए निम्नलिखित चरण शामिल हैंः

(1) सूत्र का प्रयोग करते हुए परिवर्तित नमूना परिवर्तनशीलता की गणना करें।

$$s\hat{}^2 = (n/n - 1). s^2 \ [(\sum(x-\bar{x})^2/n-1$$

(2) वांछित विश्वसनीयता स्तर एवं दर्शायी गई समतुल्य विश्वसनीयता स्तर को चयनित करें, आपको विश्वसनीयता गुणांक $x^2_{\propto/2}$ और $x^2_{1-\propto/2}$ के मानों को निश्चित degee of freedom पर x^2 तालिका से लिखना चाहिए।

(3) $s\hat{}^2$, $x^2_{\propto/2}$ और $x^2_{1-\propto/2}$ को उपरोक्त वर्णित सूत्र में रखते हुए σ^2 के लिए विश्वसनीयता अंतराल को निर्मित करें।

उदाहरण 6.16: आकार 15 के एक यादृच्छिक नमूने जिसका मानक विचलन s = 2.5 है को सामान्य समग्र से चयनित किया जाता है। परिवर्तनशीलता σ^2 और मानक विचलन σ के लिए 95% विश्वसनीयता स्तर निर्मित कीजिए।

हलः आपको दिया जा रहा हैः

$$n = 15, \ s= 2.5 \Longrightarrow s^2 = 6.25$$
$$s\hat{}^2 = (n/n - 1). s^2$$

$$= 15/15\text{-}1 \ x6.25 = 6.696$$

95% विश्वसनीयता स्तर के लिए

$\propto = 0.05$ so $\propto/2 = 0.025$ and $1-\propto = 1-0.025 = 0.975$

Degree of freedom (v) = n-1 = 15-1 = 14

14 df के लिए $x^2_{0.0252}$ का तालिका मान = 26.1

14 df के लिए $x^2_{0.0975}$ का तालिका मान = 5.63

(अ) σ^2 के लिए 95% विश्वसनीयता अन्तराल है

$$[(n-1)s\hat{}^2]/x^2_{0.025} < \sigma^2 < [(n-1)s\hat{}^2]/x^2_{0.975}$$

मानों को रखने पर, आप प्राप्त करेंगें।

$$[(15-1)x\ 6.696/26.1 \quad < \quad \sigma^2 \quad < \quad [(15-1)x\ 6.696/5.63\backslash$$

$$3.59 < \sigma^2 < 16.65$$

(ब) σ के लिए 95% विश्वसनीयता अन्तराल है :

या $< \sigma^2 < \sqrt{16.65}$

$$1.89 < \sigma^2 < 4.08$$

6.7 सारांश (Summary)

कभी कभी समग्र के बारे में आप किसी तरह का निष्कर्ष निकालने मे असमर्थ रहते हैं या सांख्यिकी पदों में समग्र परिणाम क्या है को प्रकट करने में आप असमर्थ हैं। उन परिस्थितियों में, नमूना आँकडों के आधार पर समग्र प्राचल के बारे में अनुमान लगाने की आवश्यकता होती है। इसे आंकलन का सिद्धान्त कहा जाता है। समग्र प्राचल के आंकलन के लिए नमूना आँकडों का प्रयोग कियाजाता है और इस प्रकार के आंकलनों में निष्पक्ष, अनुरूप, सक्षम और यथेष्ट तरह के विशिष्ट लक्षण होने चाहिए।

6.8 शब्दावली (Glossary)

- **आगणकः** समग्र प्राचलों का अनुमान लगाने के लिए आप विभिन्न नमूना आँकड़ों का प्रयोग करते हैं जो नमूना आँकड़े जैसे नमूना माध्य $\overline{x}$, नमूना माध्यिका m, नमूना परिवर्तनशीलता S^2 इत्यादि जो अज्ञात समग्र प्राचलों जैसे समग्र माध्य μ, समग्र परिवर्तनशीलता σ^2 आदि का अनुमान लगाते हैं, उन्हें आगणक कहा जाता है

6.9 अभ्यास प्रश्न (Practice Questions)

रिक्त स्थान भरें

1. आँकडों का एक एकल मान जिसे अज्ञात समग्र प्राचल के अनुमान के लिए प्रयोग किया जाता है उसे ....अनुमान कहते हैं।

2. जब अज्ञात समग्र स े एक एकल स्वतन्त्र यादृच्छिक नमूना निकालत े ह ैं तो उसे ....नमूना कहते है।

6.10 अभ्यास प्रश्नों के उत्तर (Answer to Practice Questions)

1. तथ्य 2. एकल

6.11 निबंधात्मक प्रश्न (Essay Type Questions)

1) समूहों के नमूने की मापें 8.3, 10.6, 9.7 , 8.8., 10.2 और 9.4 किलोग्राम ज्ञात की गई थी।
(अ) समग्र माध्य
(ब) समग्र परिवर्तनशीलता
(स) नमूना मानक विचलन अनुमानित समग्र मानक विचलन की तुलना निष्पक्ष एवं सक्षम अनुमानो से निर्धारित करें।
(2) 9 व्यक्तियों के एक यादृच्छिक नमूने में उनकी ऊँचाईयाँ 45, 47, 50, 52, 48, 47, 49, 53 और 51 इंच हैं।
(अ) सच्चा माध्य
(ब) सच्ची परिवर्तनशीलता का निष्पक्ष एवं सक्षम अनुमान ज्ञात करें।
(3) एक कम्पनी द्वारा उत्पादित 10 टेलीविजन ट्यूबों के नमूनों ने औसत जीवन 1200 घंटे और 10 घंटे का मानक विचलन दर्शाया।
(अ) समग्र माध्य
(ब) समग्र परिवर्तनशीलता के निष्पक्षएवं सक्षम अनुमानों को ज्ञात करें।

6.12 सन्दर्भ ग्रन्थ सूची/पुस्तकें (References)

1. Roy Ramendu, ***‘Principles of Statistics’***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***“Business Statistics”*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***“Advanced Statistics”,*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***“Basic Statistics”*** World Press Limited, Calcutta.
5***. “Fundamentals of Business Statistics”*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **“Quantitative Methods in Management”.**

—

इकाई 7 प्रायिकता के दृष्टिकोण
(Approaches to Probability)

7.1 प्रस्तावना (Introduction)

7.2 यादृच्छिक प्रयोग (Randomized Experiment)

7.3 प्रतिदर्श समष्टि (Sample Space)

7.4 विभिन्न पदों की परिभाषा (Definition of Various Terms)

7.5 घटना एवं प्रायिकता (Event and Probability)

7.6 क्रमचय तथा संचय की सहायता से प्रायिकता (Probability with the help of Permutation and Combination)

7.7 सारांश (Summary)

7.8 शब्दावली (Glossary)

7.9 अभ्यास प्रश्न (Practice Question)

7.10 अभ्यास प्रश्नों के उत्तर (Answer for Practice Question)

7.11 निबंधात्मक प्रश्न (Essay Types Questions)

7.12 संदर्भ पुस्तकें (Biblography)

उद्देश्य (Objective)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ यादृच्छिक प्रयोग के अर्थ की व्याख्या कर सकें।
- ✓ यादृच्छिक प्रयोग में संभावना के महत्व का वर्णन कर सकें।
- ✓ एक घटना के लिए प्रतिदर्श समष्टि की व्याख्या कर सकें।
- ✓ विभिन्न प्रकार की घटनाओं जैसे पारस्परिक अपवर्जी, सांप्रदायिक घटनाएं, सर्वांगपूर्ण, स्वतंत्र और आश्रित घटनाओं में अंतर कर सकें।
- ✓ एक घटना के घटित होने की प्रायिकता और क्रमचय तथा संचय की सहायता से प्रायिकता प्रश्नों का हल कर सकें।

7.1 प्रस्तावना (Introduction)

रोजमर्रा की जिन्दगी में आप देखते हैं कि क्रिकेट मैच शुरू होने के पहले दोनों कप्तान सिक्का उछालते हैं। सिक्का उछालना एक प्रक्रिया है और यह मानते हुए कि सिक्का खड़ा न गिरे, चित या पट आना दो संभावित निष्कर्ष है। यदि आप एक पासा फेंकते हैं तो संभावित निष्कर्ष 1,2,3,4,5,6 में से कोई भी हो सकता है। एक प्रक्रिया जो परिणाम या निष्कर्ष दे उसे प्रयोग (Experiment) कहते हैं। सामान्यतः एक प्रयोग में का निष्कर्ष संभावित निष्कर्षों में से कोई एक होता है तथा यह संयोग की बात है कि प्रयोग करते समय कौन सा निष्कर्ष आयेगा। इस अध्याय में आप विभिन्न प्रयोग और उनके निष्कर्षों के बारे में पढ़ेंगे।

आपने आज बारिश हो सकती है या ''*भारत यह मैच जीत सकता है*'' या ''*मैं इस पद के लिए चुना जा सकता हूँ*'' अवसर इस प्रकार के वाक्यों का प्रयोग किया होगा। इस प्रकार के वाक्यों में अनिश्चितता का अंश है। आप इस अनिश्चितता को कैसे मापेंगे? गणित की एक शाखा जिसे प्रायिकता सिद्धान्त (Theory of Probability) कहते हैं। इस प्रकार की अनिश्चितता को मापती है। किसी घटना घटित होने की अनिश्चितता का परिणाम मापने के लिए प्रायिकता सिद्धान्त का निर्माण किया गया है। प्रायिकता शब्द का कोषिक अर्थ है ''संभावित परन्तु अनिश्चित''। अतः जब एक सिक्के को उछालते है, चित आ सकता है परन्तु आता नहीं है। उसी प्रकार जब एक पासा को फेंकते है तो 6 आ सकता है या नहीं आ सकता।

7.2 यादृच्छिक प्रयोग (Randomized Experiment)

निम्नलिखित गतिविधियों पर ध्यान दें-

1. एक सिक्के को उछालें और निष्कर्ष को नोट करें। यहाँ पर दो संभावित निष्कर्ष हैं एक चित या पट।
2. एक पासा (fair dice) को फेंकने पर 6 संभावित निष्कर्ष प्राप्त होते हैं जो हैं 1,2,3,4,5,6 पासों का जो तल पर उपर होता है उसे परिणाम कहते हैं।
3. दो सिक्कों को एक साथ उछालें तथा संभावित निष्कर्षों को लिखें। यहाँ पर चार निम्नलिखित निष्कर्ष संभव हैं, HH, TT, HT, TH ।
4. दो पासों को फेंके निम्नलिखित 36 संभावित निष्कर्ष प्राप्त होंगे-

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

उपर्युक्त प्रत्येक गतिविधि निम्नलिखित दो शर्तों को पूरा करती हैः

क. गतिविधि को एक ही जैसी परिस्थिति में कई बार दोहराया जा सकता है।

ख. चूँकि सभी संभावित निष्कर्षों के चुने जाने की संभावना बराबर हैं इसलिए किसी भी गतिविधि का निष्कर्ष पहले से नहीं बताया जा सकता है। अतः एक गतिविधि जो-

1. एक जैसी परिस्थिति में दोहराया जाए, तथा
2. जिसका निष्कर्ष पहले से न बताया जा सके को एक यादृच्छिक प्रयोग कहते हैं।

उदाहरण 7.1 : क्या अच्छी तरह से फेटे हुए ताश के पत्तों में से एक पत्ता निकालना एक यादृच्दिक प्रयोग है?

हल : क. चँकि एक पत्ता निकालने से पहले ताश के पत्तों की गड्डी को अच्छी तरह फेंटा जा सकता है, अतः इस प्रक्रिया को कई बार दोहराया जा सकता है।

ख. 52 पत्तों मे से कोई भी पत्ता निकाला जा सकता है। अतः निष्कर्ष को पहले से नहीं बताया जा सकता। अतः यह प्रक्रिया एक यादृच्छिक प्रयोग है।

उदाहरण 7.2 : सिद्ध करें 00 कुर्सियों में से एक कुर्सी चुनना एक यादृच्छिक प्रयोग है।

हल : क. इस प्रयोग को एक समान परिस्थितियों में दोहराया जा सकता है।

ख. हर कुर्सी के चुने जाने की संभावना बराबर है।

अतः निष्कर्ष पूर्व निर्धारित नहीं है। इसलिए यह एक यादृच्छिक प्रयोग है।

क्या आप इस प्रकार की अन्य गतिविधियों के बारे में सोच सकते हैं जहाँ संभावना प्रकृति है।

आइए अब कुछ क्रियाओं की चर्चा करें जो कि यादृच्छिक प्रयोग नहीं है।

1. अवनि का जन्मः चँकि किसी व्यक्ति के जन्म की प्रक्रिया दोहरायी नहीं जा सकती अतः यह एक यादृच्छिक प्रयोग नहीं है।

2. कैलकुलेटर पर 4 तथा 8 का गुणा करना : चँकि कैलकुलेटर पर इस प्रक्रिया को कई बार दोहराया जा सकता है परन्तु परिणाम हमेशा 32 हो जायेगा। अतः यह यादृच्छिक प्रयोग नहीं है।

7.3 प्रतिदर्श समष्टि (Sample Space)

जब एक पासा फेंकते है तो संभावित निष्कर्ष क्या हो सकते हैं? पासा फेंकने पर निश्चित रूप से कोई एक भाग (हिस्सा) सबसे उपरी सतह पर होगा। अतः प्रत्येक सतह पर लिखा हुआ अंक (number(1-c) ही संभावित निष्कर्ष है।

सभी संभावित निष्कर्षों का समुच्चय (Set) S इस प्रकार से लिख सकते हैं-

S = (1,2,3,4,5,6)

इसी प्रकार जब एक सिक्के को उछालते हैं तो संभावित निष्कर्ष चित (Head) या पट (Tail) होगा। अतः संभावित निष्कर्षों का समुच्चय S होगा।

S = (H,T)

किसी प्रयोग (Experiment) के संभावित निष्कर्षों का समुच्चय S जो कि निम्न शर्तां को पूरा करें।

1. समुच्चय (Set) का प्रत्येक तत्व (Element) प्रयोग (Experiment) के संभावित निष्कर्ष को दर्शाये।

2. किसी Trail का परिणाम समुच्चय S के सिर्फ एक तत्वों (Element) हो तो ऐसे समुच्चय S को प्रतिदर्श समष्टि कहते हैं तथा इसके तत्वों को प्रतिदर्श तत्व (Sample points) कहते हैं। प्रतिदर्श समष्टि को S से प्रदर्शित किया जा सकता है।

उदाहरण 7.3 : दो सिक्कों को उछालने की क्रिया (Experiment) का प्रतिदर्श समष्टि (Sample space) लिखें।

हल : माना कि H चित तथा T पट को दर्शाता है

S ={(H,H), (H,T), (T,H), (T,T)}

नोट- यदि दो सिक्कों को एक साथ उछाला जाए तो समष्टि प्रतिदर्श को निम्न प्रकार भी लिख सकते हैं।

S ={HH, HT, TH,TT}

उदाहरण 7.4 : एक सिक्का तथा पासा को एक साथ फेंकने की प्रक्रिया का प्रतिदर्श समष्टि लिखें।

हल : एक सिक्का उछालने पर संभावित परिणाम है चित H या पट T।

एक पासा फेंकने पर संभावित निष्कर्ष है- 1,2,3,4,5 तथा 6

माना कि H = 0 T = 1

S = {(1,0), (1,1), (2,0), (2,1), (3,0), (3,1), (4,0), (4,1), (5,0), (5,1), (6,0), (6,1)}

उदाहरण 7.5 : ऐसे परिवारों का चयन करें जिनके सिर्फ 3 बच्चे हैं। प्रथम, द्वितीय और तृतीय बच्चे का लिंग पूछने ही प्रयोग है। इस प्रयोग का प्रतिदर्श समष्टि लिखें।

हल : एक बच्चे का लिंग बालक (Boy = B) अथवा बालिका (Girl = G) हो सकता है अतः
S ={BBB, BBG,BGB, GBB, BGG, GBG,GGB, GGG)

उपरोक्त तरह से प्रतिदर्श समष्टि लिखने का लाभ यह है कि निम्न प्रकार के प्रश्नों क्या दूसरी सन्तान लड़की/कन्या/बालिका थी? या कितने परिवारों में पहली सन्तान बालक है इत्यादि का उत्तर आसानी से दिया जा सकता है।

7.4 विभिन्न पदों की व्याख्या (Explanation of Various Terms)

7.4.1 घटना (Event)

सिक्के उछालने का एक प्रयोग करते है। इस प्रयोग में आने में हमारी दिलचस्पी है। अतः परिणाम में Head आना एक घटना है।

एक पासा फेंकने के प्रयोग में सम संख्या आने में हमारी दिलचस्पी है। अत: 2, 4 तथा 6 परिणाम घटना का निर्माण करते हैं। हमने देखा है जब किसी प्रयोग को एक Head समान परिस्थितियों में कई बार दोहराया जाता है तो हर बार एक ही परिणाम प्राप्त होगा और ये संभावित परिणाम प्रतिदर्श समष्टि (Sample Space) का निर्माण करते हैं।

प्रतिदर्श समष्टि के कुछ परिणाम/निष्कर्ष एक निर्दिष्ट विवरण को पूरा करते हैं, जिन्हें हम घटना कहते हैं। प्रायः घटना को A, B, C इत्यादि (अंग्रेजी के बड़े अक्षरों) शब्दों से प्रदर्शित करते हैं।

उदाहरण 7.6 : माना कि E तीन सिक्कों को एक साथ उछालने की घटना को दर्शाता है। सभी संभावित परिणाम तथा घटनाओं की सूची बनाओ जब-

1. भ्मंक की संख्या Tail की संख्या से अधिक हो।
2. जब दो Head आए।

हल : प्रतिदर्श समष्टि S

S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT)

माना कि E_1 Head की संख्या की संख्या Tail से अधिक होने की घटना दर्शाता है तथा E_2 जब दो Head आने की घटना को दर्शाता है। अतः

E_1 = {HHH, HHT, HTH, THH)

E_2 = {HHT, HTH, THH}

7.4.2 समान रूप से संभावित घटनाएं (Equally Likely Event)

यदि किसी भी कारणवश हम एक परिणाम पर दूसरे दूसरे परिणाम को वरीयता नहीं दे सकते तो Trail के ऐसे परिणामों को समान संभावित (Equally likely) कहते हैं।

उदाहरण : 1. एक निष्पक्ष सिक्के को उछालने पर चित या पट प्राप्त करना समान संभावित घटनाएं हैं।

2. एक पांसे फेंकने के प्रयोग में सभी छः तलों के समान रूप से उपरी तल पर आने में संभावित हैं।

3. एक अच्छी तरह से फेंटी हुई 52 पत्तों वाली ताश की गड्डी से एक पत्ता निकालने के लिए 52 पत्ते समान रूप से संभावित है।

7.4.3 पारस्परिक अपवर्जी घटनाएं (Mutually Exclusive Event)

यदि किसी एक घटना के घटित होने पर बाकी सारी घटनाएं नहीं घटित होंगी तो ऐसी घटनाओं को परस्पर अनन्य घटनाएं कहते हैं। अर्थात् एक ही Trail में दो या दो से अधिक घटनाएं एक साथ घटित नहीं हो सकती।

उदाहरण : 1. एक पांसे फेंकने में 1 से 6 तक अंकित सभी 6 तल परस्पर अनन्य घटनाएं है। अर्थात् यदि कोई भी एक तल उपर आता है तो बाकी सारे तल उस ज्तपंस में उपर नहीं आ सकते।

2. दो सिक्कों को एक साथ उछालने पर दोनां सिक्कों पर tail आने की घटना तथा कम से कम एक head आने की घटना परस्पर अनन्य घटनाएं हैं।

गणितीय भाषा में यदि घटनाओं का प्रतिछेदन (Intersection) Null set है (अर्थात् खाली) तो ऐसी घटनाओं को परस्पर अनन्य घटनाएं कहते हैं।

7.4.4 सर्वांगपूर्ण घटनाएं (Holistic Events)

यदि सारे पासे ऐसी घटनाओं का संग्रह है जिनकी विशेषता यह है कि घटनाओं के संग्रह में से ही कोई घटना घटित होगी तो ऐसी घटनाओं के संग्रह को संपूर्ण घटनाएं कहते हैं। उदाहरण के लिए जब एक पांसे को फेंकते हैं तो सम संख्या आने की घटना तथा विषय संख्या आने की घटना संपूर्ण घटनाएं हैं। या जब दो सिक्कों को उछालते हैं तो कम से कम एक Head आने की घटना तथा कम से कम एक Tail आने की घटना को संपूर्ण घटना कहते हैं।

गणितीय भाषा में घटनाओं के संग्रह को संपूर्ण घटना कहेंगे यदि सभी घटनाओं का union {U set theory) संग्रहण संपूर्ण प्रतिदर्श समष्टि हो।

7.4.5 स्वतंत्र तथा निर्भर आश्रित घटनाएं (Independent and Dependent Events)

एक घटनाओं के सेट को स्वतंत्र घटना कहेंगे यदि किसी एक के घटित होन पर बाकी घटनाओं पर कोई प्रभाव/असर नहीं होगा। जबकि दूसरी तरफ, यदि एक घटना के घटित होने पर दूसरी घटनाओं के घटित होने पर प्रभाव पड़ता है तो ऐसी घटनाओं को निर्भर/ परतंत्र घटनाएं कहते हैं।

उदाहरण : 1. एक सिक्के को उछालने पर पहले टॉस पर भ्मंक आने की घटना दूसरे, तीसरे और आगे आने वाली टॉसों में भ्मंक आने की घटना से स्वतंत्र है।

2. यदि एक अच्छी तरह से फेंटे हुए ताश की गड्डी से एक पत्ता निकालें और दूसरा पत्ता निकालने से पहले उसे वापस गड्डी में रख दें तो दूसरे पत्ता निकालने का परिणाम पहले बार पत्ता निकालने पर आए परिणाम से स्वतंत्र है। परन्तु यदि पहली बार में निकाले गये पत्ते को वापस गड्डी में न रखें तो दूसरी बार में निकाला गया पत्ता पहले बार में निकाले पत्ते पर निर्भर करेगा। (क्योंकि पहली बार में निकला पत्ता दूसरी बार नहीं निकल सकता)।

7.5 घटनाएं तथा उनकी प्रायिकता (Event and Their Probability)

पिछले भाग में हमने सीखा कि कैसे पता करें कोई क्रिया यादृच्छिक प्रयोग है या नहीं। प्रायिकता का अध्ययन यादृच्छिक प्रयोग दर्शाता है। अतः अंक से आगे यादृच्छिक प्रयोग की जगह सिर्फ प्रयोग शब्द का इस्तेमाल करेंगे। इसके पहले भाग में हमने विभिनन प्रकार के घटनाओं जैसे समान रूप से संभावित, परस्पर अनन्य, संपूर्ण, स्वतंत्र तथा निर्भर घटनाएं को उदाहरण सहित परिभाषित किया।

जब हम कोई प्रयोग करते हैं तो हम यह जानने के इच्छुक होते हैं कि कोई निर्धारित घटना घटित होने की क्या संभावना है। आइए कुछ उदाहरण की सहायता से समझते हैं।

एक निष्पक्ष सिक्का उछालने पर Head आने की क्या संभावना है। यहाँ पर दो Head व Tail नाम की समान संभावित परिणाम है। रोजमर्रा की जिन्दगी में हम कहते हैं कि एक सिक्के पर Head आने की संभावना 2 में से 1 है। तकनीकी भाषा में हम कहते हैं Head आने की प्रायिकता 1/2 है।

इसी प्रकार, एक पांसे फेंकने के प्रयोग में 6 समान रूप से संभावित परिणाम है जो 1,2,3,4,5 और 6 है। जिस तल पर 1 अंकित है उसके उपरी तल पर आने की संभावना 6 में से 1 है। अतः 1 आने की प्रायिकता 1/6 है।

उपरी प्रयोग में, माना कि जब पांसा फेंकते है तो उपरी तल पर सम संख्या की प्रायिकता जानने में इच्छुक है। अतः 2, 4 और 6 आने पर सम आने की घटना की संभावना 6 में से 3 है। इसलिए सम संख्या आने की प्रायिकता 3/6 या 1/2 है।

संपूर्णतया, यदि एक प्रयोग जिसमें 'n' संपूर्ण, परस्पर अनन्य, समान रूप से संभावी परिणाम हों तथा उसमें से 'm' परिणाम घटना के घटित होने के पक्ष में हो तो घटना A के घटित होने की प्रायिकता p को ऐसे ज्ञात कर सकते हैं।

p = अनुकूल परिणाम की संख्या/संभावित परिणामों की कुल संख्या

या

$p = m/n$ .... (i)

चूंकि, घटना A के घटिन न होने के अनुकूल परिणामों की संख्या n-m है, अतः A के घटित न होने की प्रायिकता q, को निम्न तरह से ज्ञात कर सकते हैं,

$q + (n-m)/n = 1 – m/n$
$= 1 - p$ {(i) देखें $p+q =1$}

नोट : p तथा q गैर नकारात्मक है तथा 1 से अधिक नहीं हो सकते।

i.e. = 0<_p<_1, 0<_q<_1

अतः किसी भी घटना के घटित होने की प्रायिकता 0 तथा 1 के बीच होती है। (0 तथा 1 को समावेशित किए हुए)।

नोट :

1. किसी घटना के घटित होने की प्रायिकता p को सफलता की प्रायिकता कहते हैं तथा घटना के न घटित होने की प्रायिकता q को असफलता की प्रायिकता कहते हैं।
2. असंभव घटना (impossible event) के घटित होने की प्रायिकता '0' (zero) हैं तथा निश्चित घटना (Sure event) के घटित होने की प्रायिकता '1' (one) है। यदि P(A) =1, तो घटना A निश्चित रूप से घटित होगी। और यदि P(A) = 0, तो घटना निश्चित रूप से घटित नहीं होगी अर्थात् घटना असंभव है।
3. किसी घटना के कुल अनुकूल परिणामों की संख्या (m) कुल संभावित परिणामों की संख्या (n) से अधिक नहीं हो सकती।

उदाहरण 7.7 : एक पांसे को एक बार फेंकने पर 5 आने की प्रायिकता ज्ञात कीजिए।

हल : एक पांसा 6 प्रकार से गिर सकता है जिसमें से केवल एवं घटना (5 आने) के घटित होने के अनुकूल है।

P(5) = 1/6

उदाहरण 7.8 : एक सिक्के को एक बार उछालने पर Head आने की प्रायिकता क्या है?

हल : सिक्का गिरने पर या तो Head (H) या tail (T) उपर आयेगा। अतः कुल संभावित परिणाम दो है तथा उनमें से 1 घटना के अनुकूल है।

अतः P(H) = ½

उदाहरण 7.9 : एक पांसे को एक बार फेंकने पर अभाज्य संख्या आने की प्रायिकता क्या है?

हल : एक पांसा फेंकने पर 6 संभावित परिणाम हो सकते हैं। जिनमें से 2,3 तथा 5 घटना के अनुकूल है।

अतः P (अभाज्य संख्या) = 3/6 = 1/2

उदाहरण 7.10 : एक पांसे को एक बार फेंकने पर 7 आने की प्रायिकता ज्ञात कीजिए। 7 से कम अंक आने की प्रायिकता ज्ञात कीजिए।

हल : एक पांसा फेंकने पर 6 संभावित परिणाम हैं 1,2,3,4,5 तथा 6। और उनमें से किसी भी तल पर 7 अंकित नहीं है।

P(7) = 0/6 = 0

चँकि सभी तलों पर अंकित अंक 7 से कम है।

अतः P(<_7) = 6/6 = 1

उदाहरण 7.11 : एक साथ 2 सिक्के उछालने पर

1. दो Head आने की 2. केवल एक Head आने की प्रायिकता ज्ञात करें।

हल : यहां संभावित परिणाम हैं HH, HT, TH, TT ।

अतः कुल संभावित परिणाम की संख्या = 4

1. दो Head आने की अनुकूल परिणामों की संख्या = 1

(i.e. HH)

P(HH) = 1/4

2. केवल एक Head आने की घटना के दो अनुकूल परिणाम हैं (HT एवं TH)

अतः P(1 Head) = 2/4 = 1/2

उदाहरण 7.12 : एक साथ दो पांसें फेंकने पर योग 9 आने की प्रायिकता ज्ञात कीजिए।
हल : कुल संभावित परिणाम की संख्या $6 \times 6 = 36$ है। जो निम्न तरह से लिख सकते हैं।

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

योग 9 निम्न तरह से ज्ञात कर सकते हैं।
3 + 6 = 9
4 + 5 = 9
5 + 4 = 9
6 + 3 = 9
अतः (3,6), (4,5), (5,4), 6,3) घटना के अनुकूल परिणाम है और अनुकूल परिणामों की संख्या 4 है।
अत P(योग 9) = 4/36 या 1/9

उदाहरण 7.13 : एक बैग जिसमें 10 लाल, 4 नीली तथा 6 काली गेंदे हैं एक गेंद (अकस्मात् यादृच्छिक) निकालो। 1. एक काली 2. एक नीली 3. काली गेंद न निकालने की प्रायिकता ज्ञात कीजिए।
हल : कुल गेंदों की संख्या 20 (10+4+6= 20) है। अतः कुल संभावित परिणामों की संख्या 20। (यादृच्छिक गेंदे निकालने से सभी समान रूप से संभावित परिणाम है।)
1. लाल गेंदों की संख्या 10
   P (लाल गेंद) = 10/20 = 1/2
2. नीली गेंदों की संख्या = 4
   ∴ P(नीली गेंद) = 4/20 = 1/5
3. काली गेंद के अलावा गेंदों की संख्या = 10 + 4 = 14
   ∴ P(न काली गेंद) = 14/20 = 7/10

उदाहरण 7.14 : एक अच्छी तरह से फेंटी हुई ताश की गड्डी के 52 पत्तों में से एक पत्ता यादृच्छिक तरह से निकालिए। यदि रानी आने की घटना है तथा 4 से अधिक तथा 10 से कम वाला पत्ता आने की प्रायिकता ज्ञात कीजिए।
हल : अच्छी तरह से फेंटे हुई ताश की गड्डी होने के कारण हर पत्ता समान रूप से संभावी है।
अतः कुल संभावी परिणामों की संख्या 52 है।
1. एक गड्डी में 4 रानी होती है।
अतः P(A) = 4/52 = 1/13
2. 4 से अधिकत था 10 से कम वाले पत्ते हैं 5,6,7,8 तथा 9। चूंकि हर कार्ड 4 तरह ईंट, पान, चिड़ी तथा हुकुम में उपलब्ध है। अतः कुल अनुकूल परिणामों की संख्या त्र $5 \times 4 = 20$
   ∴ P(B) = 20/52

उदाहरण 7.15 : एक यादृच्छिक रूप से चुने हुए अधिवर्ष (leap year) में 53 रविवार होने की संभावना ज्ञात करें।
हलः एक अधिवर्ष में 366 दिन होते हैं जिस में 52 हफ्तें तथा 2 दिन होते हैं। ये दो अधिक दिन निम्न प्रकार से हो सकते हैं-
1. रविवार तथा सोमवार
2. सोमवार तथा मंगलवार
3. मंगलवार तथा बुधवार
4. बुधवार तथा बृहस्पतिवार

5. बृहस्पतिवार तथा शुक्रवार
6. शुक्रवार तथा शनिवार
7. शनिवार तथा रविवार
अतः उपर लिखी 7 में से 2 परिणाम 1. तथा 7. घटना के अनुकूल है।
अतः P(53 रविवार) = 2/7

7.6 क्रमचय तथा संयोजन द्वारा प्रायिकता (Probability by Permutation and Combination)

पिछले भाग में हमने किसी घटना के कुल संभावित परिणाम तथा घटना के अनुकूल परिणामों की गणना करके घटना की प्रायिकता ज्ञात की। परन्तु यह तभी संभव है जब परिणामों की संख्या कम हो अन्यथा यह प्रक्रिया मुश्किल और इसमें बहुत समय लगेगा। साधारणतया आपको सभी परिणामों की सूची की आवश्यकता नहीं होती परन्तु कुल संभावित परिणामों की संख्या तथा अनुकूल परिणामों की संख्या आवश्यकता होती है। बहुत सी परिस्थितियों में यह क्रमचय तथा संयोजन के ज्ञान की सहायता से ज्ञात किया जा सकता है। जो कि आप पहले ही पढ़ चुके हैं।

उदाहरण 7.16% एक बैग/थैले में 3 लाल, 6 सफेद तथा 7 नीली गेंद है। एक सफेद तथा एक नीली एक साथ गेंद निकालने की प्रायिकता ज्ञात करें।
हल : कुल गेंद की संख्या त्र 3+6+7 = 16
16 गेंद में से 2 गेंद $^{7}c_{2}$ तरह से निकाली जा सकती है।
अतः संपूर्ण घटनाओं की संख्या = $^{16}c_{2}$ = 120
6 सफेद गेंद में से 1 गेंद $^{6}c_{1}$ तरह से तथा 7 नीली गेंदों से 1 गेंद $^{7}c_{1}$ तरह से निकाली जा सकती है। चूंकि प्रत्येक का पहला मामला प्रत्येक के दूसरे मामले से जुड़ा हुआ है। इसलिए अनुकूल परिणामों की संख्या त्र $^{6}c_{1} \times {}^{7}c_{1} = 6 \times 7 = 42$
प्रायिकता = 42/120 = 7/20
उदाहरण 7.17 : एक थैला जिसमें 5 लाल तथा 4 काली गेंदें हो, दो लाल गेंद निकालने की प्रायिकता ज्ञात करें, जब
1. प्रतिस्थापन (पहली गेंद वापस थैले में रख दी जाए)
2. बिना प्रतिस्थापन
हल : 1. कुल गेंदों की संख्या दोनों बार निकालने पर 5+4 = 9 है।
अतः (Fundamental Principle of counting) गणना के मूलभूत सिद्धान्त से, कुल संभावित परिणाम की संख्या $9 \times 9 = 81$
इसी प्रकार, अनुकूल परिणामों की संख्या = $5 \times 5 = 25$
$\therefore$ प्रायिकता (दो लाल गेंद) = 25/81
2. कुल संभावित परिणामों की संख्या बराबर 9 में से 2 गेंद निकालने के तरीके
= $^{9}C_{2}$ = (9*8)/2 =36
अनुकूल परिणामों की संख्या बराबर है 5 में से 2 लाल गेंद निकालने के तरीके
= $^{5}C_{2}$ = (5*4)/2 =10
$\therefore$ P(दो लाल गेंद) = 10/36 = 5/18
उदाहरण 7.18 : 52 पत्तों वाली ताश की गड्डी से यादृच्छिक 6 पत्ते निकाले जाते हैं उनमें 3 लाख तथा 3 काले पत्तो होने की प्रायिकता बताइए।
हल : 52 पत्तों में से 6 पत्ते $^{5}C_{6}$ तरह से निकाले जा सकते हैं।
$\therefore$ कुल संभावित परिणाम = $^{52}C_{6}$

चूंकि ताश की गड्डी में 26 पत्ते लाल तथा 26 पत्ते काले होते हैं अतः
3 लाल पत्ते $^{26}C_{3}$ तरह से तथा 3 काले पत्ते $^{26}C_{3}$ तरह से निकाले जा सकते हैं।
$\therefore$ कुल अनुकूल परिणाम = $^{26}C_{3} \times {}^{26}C_{3}$
$\therefore$ वांछित प्रायिकता = $({}^{26}C_{3} \times {}^{26}C_{3}) / {}^{52}C_{6}$ =13000/39151

उदाहरण 7.19: 3 आदमी, 2 औरत तथा 4 बच्चों के एक समूह से 4 लोगों को यादृच्छिक तरह से चुनते हैं उनमें से दो बच्चे होने की प्रायिकता 10/21 है सिद्ध करो।

हल : समूह में कुल लोगों की संख्या त्र 3+2+4 = 9

4 लोग चुने जाने पर यदि दो बच्चे हों तो बाकी 2 लोग, 3 आदमी तथा 2 औरत अर्थात् 5 लोग में से होंगे।

4 बच्चों में से 2 बच्चे $^4C_2 = 6$ तरह से चुने जा सकते हैं।

बाकी दो लोग 5 में से $^5C_2 = 10$ तरह से चुने जा सकते हैं।

9 में से 4 लोग $^9C_4 = 126$ तरह से चुने जा सकते हैं।

$\therefore$ वांछित प्रायिकता = $(^4C_2 * {}^5C_2)/{}^9C_4 = 6*10/126 = 10/21$

उदाहरण 7.20 : 52 पत्तों वाली ताश की गड्डी से 3 पत्ते निकाले जाते हैं। उनके राजा, रानी तथा गुलाम होने की प्रायिकता ज्ञात करें।

हल : 52 में से 3 पत्ते $^5C_3$ तरह से चुने जा सकते हैं चूंकि सभी समान संभावी हैं।

अतः संपूर्ण संभावों की संख्या = $^{52}C_3$

ताश की गड्डी में 4 राजा, 4 रानी तथा 4 गुलाम होते हैं। 1 राजा, 1 रानी तथा 1 गुलाम $^4C_1$ तरह से निकाला जा सकता है। चूंकि ये सभी स्वतंत्र घटनाएं है अतः अनुकूल परिणामों की संख्या = $^4C_1 \times {}^4C_1 \times {}^4C_1$

$\therefore$ वांछित प्रायिकता = $(^4C_1 \times {}^4C_1 \times {}^4C_1)/{}^{52}C_3 = 16/5525$

उदाहरण 7.21% 1 से 25 तक अंकित 25 टिकटों में से एक टिकट निकालो। निकाले टिकट के 5 के गुणक होने की प्रायिकता ज्ञात करो।

हल : 1 से 25 में 5 के गुणक हैं 5,10,15,20 तथा 25।

अर्थात् संभावित परिणाम की संख्या = 25

अनुकूल परिणाम की संख्या = 5

$\therefore$ वांछित प्रायिकता = 5/25 =1/5

7.7 सारांश (Summary)

एक क्रिया जिसका परिणाम या निष्कर्ष आए प्रयोग कहते हैं। एक क्रिया जिसको एक समान परिस्थितियों में कई बार दोहराया जा सके तथा क्रिया का परिणाम अनुमान लगाने योग्य न हो तो ऐसी क्रिया को यादृच्छिक प्रयोग कहते हैं। यादृच्छिक प्रयोग के कुल संभावित परिणामों के समुच्चय को प्रतिदर्श समष्टि तथा समुच्चय के प्रत्येक तत्व को प्रतिदर्श समंकन बिंदु कहते हैं। प्रतिदर्श समष्टि के कुछ परिणाम एक निर्धारित स्थिति को परिपूर्ण करते हैं जिसे घटना कहते हैं। जब एक घटना को किसी भी कारणवश दूसरी घटना पर वरीयता नहीं दे सकते तो ऐसी घटनाओं को समान संभावी घटना कहते हैं। यदि एक घटना के घटित होने पर दूसरे घटना न घटित हो तो ऐसी घटनाओं को परम्पर अनन्य घटनाएं कहते हैं एक ज्तपंस (एक के प्रयोग को एक बार करने की क्रिया) में कुल संभावित परिणामों को संपूर्ण घटना कहते हैं। यदि एक घटना के घटित होने से बाकी घटनाओं के घटित होने पर कोई प्रभाव न पड़े तो ऐसे घटनाओं के समूह को स्वतंत्र घटनाएं कहते हैं अन्यथा इन्हें परतंत्र/निर्भर घटनाएं कहेंगे।

7.8 शब्दावली (Glossary)

- **घटना (Event):** प्रतिदर्श समष्टि के कुछ परिणाम/निष्कर्ष एक निर्दिष्ट विवरण को पूरा करते हैं, जिन्हें हम घटना कहते हैं।
- **समान रूप से संभावित घटनाएं:** यदि किसी भी कारणवश हम एक परिणाम पर दूसरे दूसरे परिणाम को वरीयता नहीं दे सकते ।

7.9 अभ्यास प्रश्न (Practice Questions)

सत्य/असत्य

1. A तथा B परस्पर अनन्य है।
2. A तथा B परस्पर अनन्य तथ संपूर्ण है।
3. A तथा C परस्पर अनन्य हैं।
4. C तथा D परस्पर अनन्य तथा संपूर्ण है।

7.10 अभ्यास प्रश्नों के उत्तर (Answer for Practice Questions)

1. सत्य 2. सत्य 3. असत्य 4. सत्य

7.11 निबंधात्मक प्रश्न (Essay Type Question)

1. एक स्कूल से बिना किसी प्रभाव के एक विद्यार्थी चुनने की क्रिया एक यादृच्छिक प्रयोग है सिद्ध करो।
2. एक कैलकुलेटर से दो संख्या जोड़ना यादृच्छिक प्रयोग नहीं है। सिद्ध करो।
3. एक साथ 3 सिक्के उछालने का प्रतिदर्श समष्टि लिखो।
4. एक सिक्का तथा एक पासा एक साथ फेंकने पर प्रतिदर्श समष्टि लिखो।
5. दो पांसे एक साथ फेंकने पर दोनों पांसों पर उपरी तल पर 6 आए। क्या दोनों घटनाएं परस्पर अनन्य घटनाएं हैं या नहीं?
6. दो पांसे एक साथ फेंकिए। A, B, C तथा क् घटनाएं हैः
अ. पहले पांसे पर सम संख्या आए
ब. पहले पांसे पर विषम संख्या आए
स. दोनों पांसों के उपरी तल पर आए अंकों को योग 7 से कम हो।
द. दोनों पांसों के अंकों का योग 7 से ज्यादा हो
7. एक थैले में 6 लाल, 4 सफेद तथा 5 नील गेंद हैं इसमें से एक गेंद निकाली जाती है। इसके प्रतिदर्श समष्टि में कितने प्रतिदर्श बिन्दु हैं?
8. एक साथ दो पांसे फेंकने पर प्रतिदर्श समष्टि तथा प्रतिदर्श बिन्दु लिखो।
9. माना कि सिर्फ 2 बच्चों वाले सभी परिवारों को आप चुनते हैं लोगों से उनके पहले तथा दूसरे बच्चे का लिंग पूछने से प्रयोग होता है। इसका प्रतिदर्श समष्टि लिखो।
10. एक पांसे को एक बार फेंकने पर 3 आने की प्रायिकता ज्ञात करो।
11. एक सिक्के को एक बार उछालने पर Tail (पट्ट) आने की प्रायिकता ज्ञात करो।
12. एक पांसा फेंकने पर 3 से अधिक अंक आने की प्रायिकता बताओ।
13. एक साथ दो सिक्के उछालने पर कम से कम एक Tail आने की प्रायिकता बताओ।
14. एक थैला जिसमें 15 लाल तथा 10 नीली गेंदे हो एक गेंद निकाली जाती है। प्रायिकता बताइए निकलने वाली गेंद 1. लाल 2. नीली हो।
15. दो पांसे फेंकने पर योग 1) 6 2) 8 3) 10 4) 12 आने की प्रायिकता बताओ।
16. दो पांसे फेंकने पर उपरी तल पर आए अंको का योग 3 या 4 से विभक्त होने की प्रायिकता बताओ।
17. दो पांसे फेंकने पर दोनों के उपरी तल पर आने वाले अंकों का योग 10 से अधिक होने की प्रायिकता बताओ।
18. 52 पत्तों वाली ताश की गड्डी से एक पत्ता निकालने पर उसके लाल होने की प्रायिकता बताओ।
19. 52 पत्तों वाली ताश की गड्डी से एक कार्ड निकालने पर उसके 1. हुकुम 2. राजा 3. हुकुम का राजा होने की प्रायिकता बताओ।
20. एक साथ दो पांसें फेंकने पर प्रायिकता ज्ञात करो जब
1. योग एक अभाज्य संख्या हो।
2. दोनों पांसों पर एक ही अंक आए।
3. एक पांसे पर 2 का गुणक तथा दूसरे पर 3 का गुणक आए।
21. तीन सिक्कों को एक साथ उछालो। निम्नलिखित घटनाओं का प्रायिकता ज्ञात करो।
1. एक भी Head न आए
2. कम से कम एक Head आए
3. सभी Head आए।
22. एक थैले में 3 लाल, 6 सफेद तथा 7 नीली गेंदे हैं। दो गेंद निकालने पर दोनों सफेद गेंद आने की प्रायिकता बताओ।
23. एक थैले में 5 लाल तथा 8 नीली गेंद हैं। दो गेंद निकालने पर नीली और लाल आने की प्रायिकता बताओ।

24. एक थैले में 20 सफेद तथा 30 काली गेंद है। दो गेंद निकालने पर सफेद गेंद होने की प्रायिकता ज्ञात करो जब
1. प्रतिस्थापन 2. बिना प्रतिस्थापन
25. एक ताश की गड्डी से 3 पत्ते निकलने पर तीनों के गुलाम होने की प्रायिकता ज्ञात करो।

26. एक 52 पत्तों वाली ताश की गड्डी से दो पत्ते निकालने पर सिद्ध कीजिए 2 इक्के निकालने की प्रायिकता 1/221
27. एक स्कूल के 10 उत्कृष्ट बच्चों के एक समूह में 6 लड़के तथा 4 लड़कियाँ हैं। वाद-विवाद प्रतियोगिता के लिए इन में से 3 बच्चे चुनना हैं प्रायिकता ज्ञात करें जब
1. एक लड़का दो लड़की
2. सभी लड़के
3. सभी लड़कियाँ हों।
28. 1 से 21 तक अंकित 21 टिकटों में से 3 टिकट निकालने पर उनके अंक AP (Arithmetic Progression) में होने की प्रायिकता बताओ।
29. 1 से 8 तक अंकित कार्डों में से 2 कार्ड निकालो। दोनों कार्डों के अंकों का योग विषम संख्या होने की प्रायिकता बताओ।
30. 6 लड़कों तथा 8 लड़कियों के समूह में से 5 खिलाड़ियों को चुनना है। 2 लड़के तथा 3 लड़कियों के चयनित होने की प्रायिकता बताओ।
31. प्रथम 200 धनात्मक पूर्णांक में से एक पूर्णांक चुनने पर उसके 6 या 8 से विभक्त होने की प्रायिकता बताओ।

स्वपरख प्रश्नों के उत्तर

1. दोनों विशेषताएं हैं।
2. परिणाम अनुमान लगाने योग्य है।
3. S : {HHH, HHT, HTH,HTT,THH,THT,TTH,TTT}
4. H1, H2,H3,H4, H5, H6, T1,T2, T3, T4, T5,T6}
5. नहीं
6. 15
7.

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

8. {MM, MF, FM, FF}
9. 1/6
10. 1/2
11. 1/2
12. 3/4
13. i) 3/5 ii) 2/5
14. i) 3/5 ii) 5/36 iii) 1/12 iv) 1/36
15. 5/9
16. 1/12
17. 1/2
18. i) 1/4 ii) 1/13 iii) 1/52
19. i) 5/12 ii) 1/6 iii) 11/36
20. i) 1/8 ii) 7/8 iii) 1/8

21. 1/8
22. 20/39
23. a) 4/25 b) 38/245
24. 1/5525
25. सिद्ध
26. i) 3/10 ii) 1/6 iii) 1/30
27. 10/133
28. 4/7
29. 60/143
30. 1/4

7.12 सन्दर्भ ग्रन्थ सूची/पुस्तकें (Bibliography)

1. Roy Ramendu, ***'Principles of Statistics'***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***"Business Statistics"*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***"Advanced Statistics",*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***"Basic Statistics"*** World Press Limited, Calcutta.
5***. "Fundamentals of Business Statistics"*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **"Quantitative Methods in Management".**

इकाई 8 प्रायिकता सिद्धान्त
(Probability Theorem)

8.1 प्रस्तावना (Introduction)

इसके पहले का अध्याय के अध्ययन के उपरान्त आप प्रायिकता के विभिन्न पहलू तथा शब्दावली को समझने योग्य हो गये होंगे। आज बारिश हो सकती है। तथा वह आज पहुँच सकता है। इन दोनों वाक्यों में घटित होना निश्चित नहीं है। दूसरे शब्दों में इन दोनों ही वाक्यों में अनिश्चितता का भाव/अंश है। अनिश्चितता का गणितीय मापन प्रायिकता सिद्धान्त द्वारा दिया गया है। प्रायिकता सिद्धान्त का उद्देश्य अनिश्चितता मापन के तरीके प्रदान करना है। प्रायिकता सिद्धान्त की उत्पत्ति अनिश्चितता के अध्ययन जैसे पत्तों का खेल, सिक्कों को उछालना, पांसा फेंकना इत्यादि से हुई है। परन्तु आज के समय में निर्णय लेने की समस्याओं (decision making problems) के समाधान में इसका विशेष महत्व है।

8.2 उद्देश्य (Objective)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ प्रायिकता के योग प्रमेय की व्याख्या कर सकें।
- ✓ प्रायिकता के गुणन प्रमेय की व्याख्या कर सकें।
- ✓ प्रायिकता में बर्नाली प्रमेय के उपयोग का वर्णन कर सकें।
- ✓ बेज प्रमेय की सहायता से प्रश्नों का हल कर सकें।

8.3 योग प्रमेय (Additional theorem)

प्रायिकता के योग प्रमेय को आप निम्न तरह/तरीके से समझ सकते हैं।

8.3.1 पारस्परिक अपवर्जी घटनाओं के लिए योग प्रमेय (Additional Theorem for Mutually Exclusive Events)

योग प्रमेय कहता है कि यदि A तथा B दो पारस्परिक अपवर्जी घटनाएं है तो A या B के घटित होने की प्रायिकता सिर्फ A तथा B के घटित होने की प्रायिकता का योग के बराबर होगा

सांकेतिक भाषा में,

$$P\ (A \text{ या } B) = P(A) + P(B)$$

या

$$P\ (A+B) = P(A) + P(B)$$

साधारणीकरण-

इस प्रमेय को तीन या उससे अधिक पारस्परिक अपवर्जी घटनाओं के लिए बढ़ाया जा सकता है A यदि A, B तथा C तीन पारस्परिक अपवर्जी घटनाएं हैं तो

$$P\ (A+B+C) = P(A) + P(B) + P(C)$$

उदाहरण 8.1: 52 पत्तों में से 1 पत्ता निकालने पर उसके राजा अथवा रानी होनी की प्रायिकता बताइए Ar

हल : 52 पत्तों में 4 राजा तथा 4 रानी होते हैं 1 पत्ता निकालने पर उसके

राजा होने की प्रायिकता P(K) = 4/52

तथा निकाले पत्ते के रानी होने की प्रायिकता P(Q) = 4/52

चूंकि दोनों घटनाएं पारस्परिक अपवर्जी हैं अतः निकाले पत्ते के राजा या रानी होने की प्रायिकता

$$P(K \text{ या } Q) = P(K) + P(Q)$$

$$= 4/52 + 4/52 = 8/52 = 2/13$$

उदाहरण 8.2 : 3 घटनाओं A, B तथा C में से केवल एक घटना घटित हो सकती है।

हल : घटना A के घटित होने की प्रायिकता, P(A) = 3/5

घटना B के घटित होने की प्रायिकता, P(B) = 3/7

चूंकि A, B तथा C पारस्परिक अपवर्जी घटनाएं हैं।

P(A या B या C) = P(A)+P(B)+P(C) = 1

$1 = (2/5 + 3/7) = P(C) = 6/35$

उदाहरण 8.3: एक साथ 2 पांसे फेंकने पर योग 7 या 9 आने की प्रायिकता ज्ञात करो।

हलः 2 पांसे फेंकने पर $6 \times 6 = 36$ संभावित परिणाम हो सकते हैं जो निम्नलिखित हैं।

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

योग 7 निम्न 6 तरह से प्राप्त किया जा सकता है।

(6,1) (5,2) (4,3) (3,4) (2,5) (1,6)

योग 9 निम्न 4 तरह से प्राप्त किया जा सकता है।

(6,3) (5,4) (4,5) (3,6)

योग 7 ज्ञात करने की प्रायिकता $P(A) = 6/36 = 1/6$

योग 9 ज्ञात करने की प्रायिकता, $P(B) = 4/36 = 1/9$

चूंकि दोनां घटनाएं पारस्परिक अपवर्जी हैं, अतः योग 7 या 9 प्राप्त करने की प्रायिकता-

P(A या B) = P(B) + P(B)

$= 6/36 + 4/36 \quad = 10/36 \quad = 5/18$

उदाहरण 8.4 : एक थैले में 4 लाल, 5 काली, 3 पीली और 11 हरी गेंदे हैं। एक गेंद प्रायिकता ज्ञात करो कि निकाली गई गेंद

1. लाल, काली या पीली हो।
2. लाल, काली, पीली या हरी हो।

हल : कुल गेंदों की संख्या = 4R+5B +3Y+11G = 23

लाल गेंद आने की प्रायिकता, $P(A) = 4/23$

काली गेंद आने की प्रायिकता, $P(B) = 5/23$

पीली गेंद आने की प्रायिकता, $P(C) = 3/23$

हरी गेंद आने की प्रायिकता, P(D = 11/23

चूंकि घटनाएं पारस्परिक अपवर्जी हैं अतः

1. लाल, काली या पीली गेंद आने की प्रायिकता

P(A या B या C) = P(A) + P(B) +P(C)

$= 4/23 + 5/23 + 3/23$

$= 12/23$

2. लाल, काली, पीली या हरी गेंद आने की प्रायिकता

P(A या B या C या D) = P(A) + P(B) +P(C) + P(D)

$= 4/23 + 5/23 + 3/23 + 11/23$

$= 23/23 = 1$

उदाहरण 8.5: यदि एक जोड़े पांसे फेंके जाए तो प्रायिकता ज्ञात करें

1. योग न तो 7 है न 11
2. न तो दोनों पांसे पर एक ही संख्या आए और न ही योग 9 आए।

हल : यहाँ पर 36 संभावित परिणाम हैं जो निम्न हैं।

1. योग 7 निम्न 6 तरह से आ सकता है-

(6,1) (5,2) (4,3) (3,4) (2,5) (1,6)

योग 11 निम्न 2 तरह से प्राप्त कर सकते हैं-

(6,5) (5,6)

योग 7 प्राप्त करने की प्रायिकता, P(A) = 6/36

योग 11 प्राप्त करने की प्रायिकता, P(B) = 2/36

चूंकि A तथा B पारस्परिक अपवर्जी घटनाएं हैं, अतः योग 7 या 11 आने की प्रायिकता

P(A या B) = P(A) + P(B)

योग न तो 7 और न ही 11 होने की प्रायिकता

= 1- P(A या B)

= 1 -2/9 = 7/9

2. दोनों ही पांसो में एक समान अंक निम्न 6 तरह से आ सकता है

(1,1) (2,2) (3,3) (4,4) (5,5) (6,6)

योग 9 निम्न 4 तरह से आ सकता हैः

(6,3) (5,4) (4,5) (3,6)

दोनों पासों में एक समान अंक आने की प्रायिकता, P(A) = 6/36

योग 9 आने की प्रायिकता, P(B) = 4/36

चूंकि घटनाएं पारस्परिक अपवर्जी हैं, अतः जोड़ा आने या योग 9 आने की प्रायिकता

P(A या B) = P(A) + P(B)

=6/36 +4/36 =10/36 = 5/18

अतः न तो जोड़ा आए न ही योग 9 आने की प्रायिकता

= 1 - P(A या B)

= 1 – 5/18 = 13/18

उदाहरण 8.6 : A और B पारस्परिक अपवर्जी घटनाएं हैं। P(A) = 0.3 तथा P(B) = च और P(A+B) = 0.5 p का मान ज्ञात करो।

हल : चूंकि A तथा B पारस्परिक अपवर्जी घटनाएं हैं तो

P(A+B) = P(A) + P(B)

0.5 = 0.3 + p

p = 0.5 - 0.3 = 0.2

8.3.2 अपारस्परिक अपवर्जी घटनाओं के लिए योग प्रमेय-

जब घटनाएं पारस्परिक अपवर्जी न हों तो उपर वर्णित योग प्रमेय उपयुक्त नहीं होगा। उदाहरण के लिए एक हुकुम का पत्ता निकालने की प्रायिकता 13/52 है तथा एक राजा निकालने की प्रायिकता 4/52 है, चूंकि ये घटनाएं पारस्परिक अपवर्जी घटनाएं नहीं है अतः एक हुकुम का पत्ता या राजा वाला पत्ता निकालने की प्रायिकता दोनों घटनाओं की एकाकी प्रायिकता जोड़कर नहीं निकाल सकते क्योंकि पत्ता राजा भी हो सकता है और हुकुम भी हो सकता है।

जब घटनाएं पारस्परिक अपवर्जी न हो तो योग प्रमेय में संशोधन करना आवश्यक है।

संशोधित योग प्रमेय कहता है कि यदि A तथा B पारस्परिक अपवर्जी घटनाएं न हों तो A या B या दोनों के घटित होने की प्रायिकता A के घटित होने की प्रायिकता तथा B के घटित होने की प्रायिकता ऋण का योग और A तथा B के एक साथ घटित होने की प्रायिकता को घटाकर बराबर होगा।

सांकेतिक भाषा में,

P (A या B या दोनों) = P(A) + P(B) . P(AB)

इस फार्मूले में हम P(A और B)= P(AB) को घटाते हैं अर्थात् P(A) +P(B) के घटित होने पर जिन घटनाओं की प्रायिकता को दो बार गिना गया है। अतः प्रमेय को इस प्रकार संशोधित किया गया है कि A तथा B को पारस्परिक अपवर्जी प्रस्तुत किया जा सके।

निम्नलिखित चित्र से इसे दर्शाया/समझाया जा सकता है।

साधारणीकरण (Generalisation)

इस प्रमेय को तीन या अधिक घटनाओं के लिए विस्तृत/बढ़ाया जा सकता है। अतः

P(A या B या C) = P(A) +P(B)+P(C)- P(AB) - P(AC) - P(BC)+ P(ABC)

निम्नलिखित उदाहरण संशोधित योग प्रमेय के उपयोग को स्पष्ट करते हैं।

उदाहरण 8.7 : एक अच्छी तरह फेंटी हुई ताश की गड्डी से एक पत्ता निकालिए। पत्ते के राजा या हुकुम होने की प्रायिकता क्या है?

हल : एक हुकुम का पत्ता निकालने की प्रायिकता, P(A) = 13/52

राजा वाला पत्ता निकालने के प्रायिकता, P(B) = 4/52

चूंकि राजा के पत्तों में से एक हुकुम का पत्ता हो सकता है अतः घटनाएं पारस्परिक अपवर्जी नहीं है।

हुकुम के राजा निकालने की प्रायिकता, P(।B)= 1/52

अतः हुकुम या राजा निकालने की प्रायिकता

P(A या B या दोनों) = P(A) + P(B) - P(AB)

= 13/52 + 4/52 - 1/52 = 4/13

उदाहरण 8.8 : एक थैले में 1 से 30 तक अंकित 30 गेंदें हैं। एक गेंद दैव दर्शन से निकाला जाता है। निकाली गई गेंद पर अंकित नंबर 5 या 6 का गुणक होने की प्रायिकता ज्ञात करो।

हल : गेंद के 5 के गुणक होने की प्रायिकता

(5,10,15,20,25,30) P(A) = 6/30

गेंद के 6 के गुणक होने की प्रायिकता

(6, 12, 18, 24, 30) P(B) = 5/30

चूंकि 30, 5 तथा 6 दोनों का गुणक है अतः घटनाएं पारस्परिक अपवर्जी नहीं है।

P(A तथा B) = 1/30

अतः गेंद के 5 या 6 के गुणक होने की प्रायिकता

P (A या B) = P(A) + P(B) . P(A B)

6/30 + 5/30 - 1/30 = 10/30 = 1/3

उदाहरण 8.11: एक पत्तों की गड्डी से दैव दर्शन द्वारा एक पत्ता निकाला जाता है। प्रायिकता बताओ 1. पत्ता राजा या रानी हो 2. पत्ता या तो राजा हो या काला पत्ता हो।

हल : 1. राजा पत्ता निकालने की प्रायिकता P(K) = 4/52

रानी पत्ता निकालने की प्रायिकता, P(Q) = 4/52

चूंकि दोनों घटनाएं पारस्परिक अपवर्जी हैं अतः पत्ते के राजा अथवा रानी होने की प्रायिकता

P(K या Q) = P(K) + P(Q)

2. राजा पत्ता निकालने की प्रायिकता, P(A) = 4/52

काला पत्ता निकालने की प्रायिकता, P(B) = 26/52

चूंकि काले राजा दोनों घटनाओं में common है। अतः घटनाएं पारस्परिक अपवर्जी नहीं है।

P (काला राजा) = 2/52

अतः पत्ते के राजा या काला पत्ता होने की प्रायिकता

P (राजा या काला पत्ता)= P(राजा) + P(काला पत्ता) - P(काला राजा)

=4/52 + 26/52 – 2/52 = 28/52

8.4 गुणन प्रमेय (Multiplication Theorem)

अब आप प्रायिकता के गुणन प्रमेय का दो आयामों के अंतर्गत अध्ययन करेंगे।

8.4.1 स्वतंत्र घटनाओं के लिए गुणन प्रमेय (Multiplication Theorem for Independent Events)

गुणन प्रमेय कहता है यदि A तथा B दो स्वतंत्र घटनाएं हों तो A तथा B के एक साथ घटित होने की प्रायिकता उनकी एकाकी प्रायिकता के गुणन के बराबर होगी। सांकेतिक भाषा में,

P(AB) = P(A) × P(B)

साधारणीकरणः इस प्रमेय को तीन या उससे अधिक घटनाओं के लिए विस्तारित किया जा सकता है। यदि A, B तथा C तीन स्वतंत्र घटनाएं हो, तो

P(ABC) = P(A)× P(B) × P(C)

उदाहरण 8.12: एक सिक्के को 3 बार उछाला जाता है। 3 चित आने की प्रायिकता बताओ।

हलः पहले टॉस में चित आने की प्रायिकता P(A) = 1/2
दूसरे टॉस में चित आने की प्रायिकता, P(B) = 1/2
तीसरे टॉस में चित आने की प्रायिकता, P(C) = 1/2
चूंकि घटनाएं स्वतंत्र है अतः 3 चित आने की प्रायिकता,

$$P(ABC) = P(A) \times P(B) \times P(C)$$
$$= ½ x ½ x ½ = 1/8$$

उदाहरण 8.13: 52 पत्तों की गड्डी से दो पत्ते दैव निदर्शन द्वारा निकाले जाते हैं एक के बाद दूसरा जबकि पहला पुनः गड्डी में शामिल कर दिया जाता है। दोनों पत्तों के राजा होने की प्रायिकता बताओ।
हलः राजा पत्ता निकालने की प्रायिकता, P(A) = 4/52
रिप्लेसमेंट के उपरांत राजा पत्ता निकालने की प्रायिकता, P(B) = 4/52
चूंकि घटनाएं स्वतंत्र है अतः दो राजा निकालने की प्रायिकता,

$$P(AB) = P(A) \times P(B)$$
$$=4/52 x 4/52 = 1/169$$

8.4.1.1 द स्वतंत्र घटनाओं में से कम से कम एक घटना के घटित होने की प्रायिकता

यदि आपके पास n स्वतंत्र घटनाएं $A_1, A_2, A_3, \ldots\ldots A_n$ हो तथा उनके घटित होने की प्रायिकता क्रमशः $p_1, p_2, p_3, \ldots\ldots p_n$ हो, तो स्वतंत्र घटनाओं $A_1, A_2, A_3, \ldots\ldots A_n$ में से कम से कम एक घटना के घटित होने की प्रायिकता

P(कम से कम एक घटना घटित होना)
= 1 - P(किसी घटना का घटित न होना)
$= 1 – [(1-p_1) \quad (1- p_2) \quad (1- p_3) \ldots\ldots\ldots (1- p_n)]$

उदाहरण 8.14: तीन विद्यार्थियों A, B तथा C को एक सांख्यिकी की समस्या/प्रश्न हल करने के लिए दी जाती है तथा उनको इसको हल कर लेने का संयोग ½, 1/3 तथा 1/4 है। समस्या हल कर लेने की प्रायिकता बताइये।

हलः A के समस्या हल कर लेने की प्रायिकता, P(A) = 1/2
B के समस्या हल कर लेने की प्रायिकता, P(B) = 1/3
C के समस्या हल कर लेने की प्रायिकता, P(C) = 1/4
A के समस्या हल न कर लेने की प्रायिकता, $P(\bar{A})$ = 1- 1/2 = 1/2
B के समस्या हल न कर लेने की प्रायिकता, $P(\bar{B})$ = 1- 1/3 = 2/3
C के समस्या हल न कर लेने की प्रायिकता, $P(\bar{C})$ = 1- 1/4 = 3/4
चूँकि सभी घटनाएं स्वतंत्र हैं, अतः

P(कोई भी समस्या हल नहीं कर पाया) = P(A) × P(B) × P(C)
= 1/2 × 1/3 × 1/4 = 1/4
P(समस्या हल हो जाए) = 1 - P(कोई समस्या हल न कर पाए)
= 1- 1/4 = 3/4

उदाहरण 8.15 : एक अभ्यर्थी X ने 3 पदों के लिए साक्षात्कार दिया। प्रथम पद के लिए 3 अभ्यर्थी, द्वितीय के लिए 4 तथा तृतीय के लिए 2 अभ्यर्थी थे। X के चयनित होने की प्रायिकता बताओ।
हलः प्रथम पद के लिए चयनित होने की प्रायिकता, P(A) = 1/3
द्वितीय पद के लिए चयनित होने की प्रायिकता, P(B) = 1/4
तृतीय पद के लिए चयनित होने की प्रायिकता, P(C) = 1/2
प्रथम पद के लिए चयनित न होने की प्रायिकता, P(A) = 1 - 1/3= 2/3
द्वितीय पद के लिए चयनित न होने की प्रायिकता, P(B) = 1 - 1/4= 3/4
तृतीय पद के लिए चयनित न होने की प्रायिकता, P(C) = 1 - 1/2= 1/4

चूँकि सभी घटनाएं स्वतंत्र है अतः X के चयनित न होने की प्रायिकता

$$p(\bar{1}), p(\bar{2}), p(\bar{3}), p(\bar{4}), p(\bar{5}), p(\bar{6})$$

$= 5/6, 5/6, 5/6, 5/6, 5/6, 5/6, 5/6$
$= (5/6)^2$

अतः कम से कम एक पद पर चयनित होने की प्रायिकता $= 1 - (5/6)^2$

8.4.1.2 सशर्त (शर्तयुक्त प्रायिकता) (Conditional Probability)

यदि घटनाएं आश्रित हों तो उपर वर्णित गुणन प्रमेय उपयुक्त नहीं होगा। आश्रित घटनाएं वो हैं जिसमें एक से घटित होने के दूसरे घटनाओं की प्रायिकता प्रभावित हो। घटना B के घटित होने की प्रायिकता जब A पहले ही घटित हो चुका हो को B की शर्त युक्त प्रायिकता कहते हैं। इसे P(B/A) द्वारा प्रदर्शित करते हैं। इसी तरह शर्तयुक्त A जब B पहले घटित हो चुका है को P(A/B) द्वारा प्रदर्शित करते हैं।

शर्तयुक्त प्रायिकता की परिभाषा-

यदि A तथा B आश्रित घटनाएं हो तो B की शर्तयुक्त प्रायिकता जब A पहले ही घटित हो चुका हो इस प्रकार परिभाषित करते हैं और

$P(B/A) = P(AB)/P(A), \quad P(A) > 0$

इसी प्रकार A की शर्तयुक्त प्रायिकता Given B को परिभाषित करते हैं तथा

$P(A/B) = P(AB)/P(B), \quad P(B) > 0$

8.4.2 आश्रित घटनाओं के लिए गुणन प्रमेय या सशर्त प्रायिकता के संदर्भ में गुणन प्रमेय-

जब घटनाएं स्वतंत्र न हों अर्थात् आश्रित घटनाएं हो तो गुणन प्रमेय को संशोधित करने की आवश्यकता है। संशोधित गुणन प्रमेय कहता है कि यदि A और B दो आश्रित घटनाएं हो तो उनके एक साथ घटित होने की प्रायिकता (समसामयिक प्रायिकता) बराबर होगी एक घटना की प्रायिकता और दूसरी घटना की सशर्त प्रायिकता के गुणन के बराबर।

सांकेतिक भाषा में,

$P(A\ B) = P(A) \cdot P(B/A)$

या $P(A\ B) = P(B) \cdot P(A/B)$

जहां $P(B/A) =$ B की सशर्त प्रायिकता given A

$P(A/B) =$ A की सशर्त प्रायिकता given B

उदाहरण 8.16 : एक थैले में 10 सफेद और 5 काली गेंदे हैं। दैव निदर्शन द्वारा दो गेंदों बिना प्रतिस्थापन के निकाली जाती है। दोनों गेंदों के काली होने की प्रायिकता बताओ।

हल : प्रथम प्रयास में काली गेंद निकालने की प्रायिकताए $P(A) = 5/10+5 = 5/10$

पहले प्रयास में काली गेंद निकालते और बिना प्रतिस्थापन दूसरी काली गेंद निकालने की प्रायिकता $= P(B/A) = 4/10+4 = 4/14$

चँकि घटनाएं आश्रित है अतः दोनों गेंद काली होने की प्रायिकता

$P(A\ B) = P(A) \cdot P(B/A)$

$= 5/15 \times 4/14 = 2/21$

उदाहरण 8.17: तीन लगातार प्रयासों में बिना प्रतिस्थापन क्रमशः राजा, रानी और गुलाम के पत्ते निकालने की प्रायिकता बताओ।

हल : राजा निकालने की प्रायिकता $= 4/52$

एक रानी निकालने की प्रायिकता जब एक राजा पहले ही निकाला जा सकता है,

$P(B/A) = 4/51$

एक राजा और रानी निकालने के बाद गुलाम आने की प्रायिकता,

$P(B/AB) = 4/51$

चँकि घटनाएं आश्रित है अतः एक राजा, एक रानी और गुलाम निकालने की प्रायिकता,

$P(ABC) = 4/52 \times 4/51 \times 4/50 = 8/16575$

उदाहरण 8.18 : एक बैग, जिसमें 5 सफेद और 3 लाल गेंद है, चार गेंदे एक के बाद एक करके बिना प्रतिस्थापन निकाली जाती है। प्रायिकता बताओ।

1. सफेद और लाल गेंद एक के बादएक आएगी।
2. लाल और सफेद गेंद वैकल्पिक रूप से आएगी।

हल : 1. 1 सफेद गेंद निकालने की प्रायिकता= 5/8
1 लाल गेंद निकालने की प्रायिकता = 3/7
1 सफेद गेंद निकालने की प्रायिकता = 4/6
1 लाल गेंद निकालने की प्रायिकता = 2/5
चूंकि घटनाएं आश्रित हैं अतः वांछित प्रायिकता
P (1W 1R 1W 1R) = 5/8 x 3/7 x 4/6 x 2/5 = 1/14

2. 1 लाल गेंद निकालने की प्रायिकता = 3/8
1 सफेद गेंद निकालने की प्रायिकता = 5/7
1 लाल गेंद निकालने की प्रायिकता = 2/6
1 सफेद गेंद निकालने की प्रायिकता = 4/5
चूंकि घटनाएं आश्रित हैं अतः वांछित प्रायिकता

P(1R 1W 1R 1W) = 3/8 x5/7 x 2/6 x 4/5 =1/14

8.5 योग तथा गुणन प्रमेय का संयुक्त प्रयोग (Combined Use of Addition and Multiplication Theorems)

प्रायिकता के अन्तर्गत कुछ समस्याएं हैं जहाँ दोनों योग तथा गुणन प्रमेय का एक साथ उपयोग होता है। ऐसी परिस्थितियों में सबसे पहले आप गुणन प्रमेय लागू करेंगे और उसके बाद अंततः योग प्रमेय उपयोग करेंगे।

उदाहरण 8.19: A 80% प्रतिशत मामलों सत्य बोलता है और B 90% कितने प्रतिशत मामलों में एक ही तथ्य को बताने में दोनों में विरोधाभास है।
हल: माना कि A तथा B की सत्य बोलने की प्रायिकता क्रमशः A तथा B है।
$P(A) = 80/100 = 4/5$, $P(\bar{A}) = 1-P(A) = 1-4/5 = 1/5$
$P(B) = 90/100 = 9/10$, $P(\bar{B}) = 1-P(B) = 1-2/10 = 1/10$
उनमें विरोधाभास तभी होगा जब तक सत्य बोले और दूसरा असत्य A अतः निम्न 2 संभावनाएं हैं
1. A सत्य बोले और B असत्य
2. B सत्य बोले और A असत्य
चूंकि घटनाएं स्वतंत्र है अतः गुणन प्रमेय की सहायता से
1. I संभावना की प्रायिकता = P(A).P(B) = 4/5 x 1/10 = 4/50
2. II संभावना की प्रायिकता = P(A).P(B) = 1/5 x 9/10 = 9/50
चूंकि दोनों घटनाएं पारस्परिक अपवर्जी हैं अतः योग प्रमेय की सहायता से
वांछति प्रायिकता = 4/50 +9/10 = 13/50

उदाहरण 8.20: एक थैले में 5 सफेद और 3 लाल गेंदें हैं और क्रमशः 4 गेंद बिना प्रतिस्थापन निकाली जाती है। उनके alternatively अलग होने की प्रायिकता ज्ञात करो।
हल : अल्टरनेटिवली 4 रंग हो सकते हैं।
1. सफेद, लाल, सफेद, लाल (1W 1R 1W 1R)
2. लाल, सफेद, लाल, सफेद (1R 1W 1R 1W)
1. सफेद गेंद से शुरूआत हो
पहली गेंद सफेद होने की प्रायिकता = 5/8
पहली गेंद लाल होने की प्रायिकता = 3/7
पहली गेंद सफेद होने की प्रायिकता = 4/6
पहली गेंद लाल होने की प्रायिकता = 2/5
चूंकि घटनाएं आश्रित हैं अतः गुणन प्रमेय की सहायता से
P(1W 1R 1W 1R) = 3/8 x 5/7 x 2/6 x 4/5 = 1/14

चूंकि 1. तथा 2. पारस्परिक अपवर्जी घटनाएं है अतः योग प्रमेय द्वारा

वांछित प्रायिकता = 1/14 + 1/14 = 2/14 = 1/7

उदाहरण 8.21: कारीगरों के तीन समूह में 3 आदमी और 1 औरत, 2 आदमी और 2 औरत और 1 आदमी और 3 औरत हैं। प्रत्येक समूह से दैव निदर्शन द्वारा एक कारीगर चुना जाता है। चुने गये कारीगरों के समूह में 1 आदमी और 2 और होने की प्रायिकता बताओ।

हल : यहाँ पर 3 संभावनाएं हो सकती हैं

1. पहले समूह से 1 आदमी और दूसरे व तीसरे समूह से एक-एक और चुनी जाए।
2. दूसरे समूह से एक आदमी तथा पहले और तीसरे समूह से एक-एक औरत चुनी जाए।
3. तीसरे समूह से एक आदमी तथा पहले व दूसरे समूह से एक-एक औरत चुनी जाए।

चूंकि घटनाएं स्वतंत्र है अतः गुणन प्रमेय की सहायता से

1. प्रथम केस की प्रायिकता = ¾ x 2/4 x ¾ = 18/64
2. द्वितीय केस की प्रायिकता = 2/4 x ¼ x ¾ = 6/64
3. तृतीय केस की प्रायिकता = ¼ x ¼ x 2/4 = 2/64

चूंकि संभावनाएं पारस्परिक अपवर्जी हैं अतः योग प्रमेय की सहायता से

वांछित प्रायिकता = 18/64 + 6/64 + 2/64 =24/64 =13/32

8.6 प्रायिकता सिद्धान्त में बर्नाली प्रमेय का उपयोग (Use of Bernali's theorem in Probability Theory)

प्रायिकता समस्याओं के समाधान में बर्नाली प्रमेय बहुत उपयोगी है। यह प्रमेय कहता है कि यदि एक घटना या परख या प्रयोग के घटित होने की प्रायिकता ज्ञात हो तो इसके 1,2,3........r बार घटित होने की प्रायिकता निम्न सूत्र से ज्ञात किया जा सकता है।

$$P(r) = {}^nC_r\, p^r q^{n-r} \qquad ; r = 1,2,3,\ldots\ldots,n$$

जहां

P(r) = n परख में r सफलता की प्रायिकता

p = सफलता की प्रायिकता या एक परख में घटना के घटित होने की प्रायिकता

q = असफलता की प्रायिकता या एक परख में घटना के घटित न होने की प्रायिकता

n = कुल परखों की संख्या

निम्नलिखित उदाहरण इस प्रमेय के उपयोग को दर्शाते हैं।

उदाहरण 8.22 : तीन सिक्के एक साथ (समसामयिक) उछाते जाते हैं। एक साथ दो चित आने की प्रायिकता ज्ञात करो।

हल : चूंकि दो चित आने की प्रायिकता ज्ञात करना है अतः बर्नाली प्रमेय की सहायता से यह आसान होगा। इस प्रमेय के अनुसार

$$P(r) = {}^nC_r\, p^r q^{n-r}$$

दिया गया है n = 3, r = 2, p= सिक्के के एक उछाल पर चित आने की प्रायिकता = 1/2

$$P(2H) = {}^3C_2\,(1/2)^2\,(1/2)^{3-2}$$

$$= 3!\,/(3-2)!\,2! \times 1/8 = 3/8$$

उदाहरण 8.23: एक सेना की टुकड़ी में 3/5 सैनिक विवाहित/शादीशुदा है और बाकी के 2/5 सैनिक अविवाहित है। 5 सैनिकों की पंक्ति में 4 सैनिकों के विवाहित होने की प्रायिकता ज्ञात करो।

हल : यहाँ पर n= 5, r= 4; p = विवाहित सैनिक होने की प्रायिकता = 3/5

q = 1-p = 1-3/5 = 2/5

P(4 विवाहित सैनिक) = $^5C_4\,(B/5)^4\,(2/5)^1$

= = 162/655

उदाहरण 8.24: इंग्लैण्ड के खिलाफ भारत के क्रिकेट मैच जीतने की प्रायिकता 1/3 है। यदि भारत और इंग्लैण्ड के बीच 3 टेस्ट मैच खेले जाने हो तो प्रायिकता ज्ञात करो 1. भारत तीनों मैच हार जाए 2. भारत कम से कम एक मैच जीत जाए।

हल : $n = 3$

$p =$ मैच जीतने की प्रायिकता $= 1/3$

$q = 1-1/3 = 2/3$

1. P (सभी मैच हार जाए) = $P(0) = {}^3C_0\,(1/3)^0\,(2/3)^3 = 8/27$

2. P (कम से कम एक मैच जीत जाए)

= 1- P (एक भी मैच न जीते)

$= 1 - {}^3C_0\,(1/3)^0\,(2/3)^3 = 19/27$

8.6 बेज प्रमेय (Bayes Theorem)

बेज प्रमेय का नामकरण ब्रिटिश गणितज्ञ थॉमस बेज (Thomas Bayes) के आधार पर किया गया है और यह 1763 में मुद्रित हुआ। बेज प्रमेय की सहायता से स्वयंसिद्ध प्रायिकताएं (prior probalities) कुछ न्यादर्श जानकारी की सहायता से संशोधित करके प्रतिबंधी प्रायिकताएं प्राप्त की जाती है। इस प्रमेय को विपरीत प्रायिकता प्रमेय भी कहते हैं। माना कि एक फैक्ट्री में A_1 तथा A_2 दो मशीनों द्वारा माल तैयार किया जाता है। माना कि A_1 और A_2 मशीन क्रमशः कुल माल का 70% तथा 30% माल तैयार करती है और क्रमशः 5% तथा 3% की दर से खराब बोल्ट बनाती है। माना कि कुल तैयार माल से एक समान चुना जाता है और यह खराब पाया जाता है और अब हम यह जानना चाहते हैं कि ये माल, सामान A_1 तथा A_2 मशीन द्वारा बनाये जाने की क्या प्रायिकता है तो यह बेज प्रमेय की सहायता से ज्ञात कर सकते हैं। माना कि एक थैले 6 काली और 4 सफेद गेंदें हैं। दूसरे थैले में 45 काली और 6 सफेद गेंद हैं। एक थैले से दैव निदर्शन द्वारा एक गेंद निकालने पर वह काली पायी जाती है यदि आप यह जानना चाहते हैं कि वह गेंद पहल या दूसरे थैले से आयी है तो इसे बेज प्रमेय को उपयोग करके निकाल/ज्ञात कर सकते हैं।

बेज प्रमेय का कथन :

यदि A_1 तथा A_2 परस्पर अपवर्जी और सर्वांगपूर्ण घटनाएं हो तथा B एक ऐसी घटना हो जो कि A_1 तथा A_2 के जोड़े के साथ घटित हो तो A_1 तथा A_2 शर्तयुक्त/सशर्त प्रायिकता जबकि B घटित हो चुका हो इस प्रकार ज्ञात कर सकते हैं।

$$P(A_1/B)=$$

इसी तरह

$$P(A_2/B)=$$

साधारणीकरणः

बेज प्रमेय को तीन या अधिक घटनाओं के लिए विस्तृत कर सकते हैं। यदि तथा तीन पारस्परिक अपवर्जी घटनाएं हो तथा एक ऐसी घटना हो जो तथा के संयोजन के साथ घटित हो तो

$$P(A_1/B)=$$

$$P(A_2/B)=$$

$$P(A_3/B)=$$

उदाहरण 8.25 : एक बोल्ट फैक्ट्री में A, B तथा C मशीनें कुल माल का क्रमशः 25%, 35% तथा 40% तैयार करती है। तैयार माल से क्रमशः 5, 4 तथा 2 प्रतिशत बोल्ट खराब हो जाते हैं। दैव निदर्शन से एक बौल्ट निकाला जाता है और वह खराब पाया जाता है। इसके मशीन C से तैयार होने की प्रायिकता ज्ञात करो।

हल : माना कि A, B तथा C मशीनों द्वारा तैयार बोल्ट को निकालने की घटना क्रमशः A, B ए तथा C है और क् बोल्ट के खराब होने की घटना दर्शाता है।

दी गई सूचना का आधार पर

P(A) = 35% = 25/100 = 0.25

P(B) = 35% = 35/100 = 0.35

P(C) = 40% = 40/100 = 0.40

सशर्त प्रायिकता

P(D/A) 5% = 5/100 = 0 ण्05

P(D/B) = 4% = 4/100 = 0.04

P(D/C) = 2% = 2/200= 0.02

नीचे दिए गए तालिका/टेबल में दी गई जानकारी रखने पर

घटना	स्वयंसिद्ध प्रायिकता	सशर्त प्रायिकता	संधि प्रायिकता
A	P(A) =0.25	P(D/A) = 0.05	0.25 × 0.05
B	P(B) = 0.35	P(D/B) = 0.04	0.35 × 0.04
C	P(C) = 0.40	P(D/C) = 0.02	0.40 ×0.02

हमें ज्ञात करना है कि खराब बोल्ट मशीन द्वारा तैयार किया गया है अर्थात् P(C/D)

P(C/D) = मशीनों की संधि प्रायिकता / तीनों मशीनों की संधि प्रायिकता का योग

= (0.40 x 0.02)/ (0.25x0.05+0.35x0.04+0.40x0.02)

= 0.008/ (0.0125+0.014+0.008) =0.008/0.345

= 0.2318 or 23.18%

8.8 सारांश (Summary)

प्रायिकता का योग प्रमेय पारस्परिक अपवर्जी घटनाओं के लिए उपयोग करते हैं न कि अपारस्परिक अपवर्जी घटनाओं के लिए। प्रायिकता का गुण प्रमेय स्वतंत्र तथा आश्रित घटनाओं के लिए उपयोग करते हैं। अतः इसे सशर्त प्रायिकता में भी उपयोग कर सकते हैं। एक परख या घटना के घटित होने की प्रायिकता ज्ञात करने में बर्नाली प्रमेय बहुत उपयोगी है और बेज प्रमेय का उपयोग प्रायिकता समस्याओं के समाधान में उपयोगी है।

8.9 शब्दावली (Glossary)

- **पारस्परिक अपवर्जी घटनाएः** दो घटनाएँ उस समय परस्पर अपवर्जी होती हैं जब वे एक ही समय घटित नहीं हो सकती हैं

8.10 अभ्यास प्रश्न (Practice Question)

1. ....वो हैं जिसमें एक से घटित होने के दूसरे घटनाओं की प्रायिकता प्रभावित हो।

2. .... = P(A) × P(B)

3. बेज प्रमेय का नामकरण ब्रिटिश गणितज्ञ .... के आधार पर किया।

8.11 अभ्यास प्रश्नों के उत्तर (Answer for Practice Question)

1. आश्रित घटनाएं 2. P(AB) 3. थॉमस बेज

8.12 निबंधात्मक प्रश्न (Essay Type Questions)

1. 52 पत्तों की गड्डी से एक पत्ता निकालने पर उसके पान अथवा चिड़ी की रानी होने की प्रायिकता बताओ।

2. 25 छात्रों की कक्षा में प्रत्येक छात्र को 1 से 25 तक रोल नं0 दिया गया है। एक प्रश्न का उत्तर देने के लिए एक छात्र को दैव निदर्शन द्वारा चुना जाता है। चुने गये छात्र के रोल नं0 का 5 या 7 से गुणक होने की प्रायिकता बताइए?

3. एक थैले में 63 लाल, 6 सफेद, 4 नीली तथा 7 पीली गेंदें हैं। एक गेंद निकालने पर उसके सफेद या पीली होने की प्रायिकता बताओ।

4. एक साथ तीन पांसे फेंकने पर योग 17 या 18 आने की प्रायिकता बताओ।

5. एक साथ 2 पांसे फेंकने पर योग 9 या 11 आने की प्रायिकता बताओ।

6. एक पत्तों की गड्डी में से पान या राजा पत्ता निकालने की प्रायिकता बताओ।

7. एक थैले में 1 से 50 से अंकित 50 गेंदें हैं। दैव निदर्शन से एक गेंद निकालने पर उसके 5 या 7 के गुणक होने की प्रायिकता ज्ञात करो।

8. एक कांट्रैक्टर के मरम्मत कांट्रेक्ट हासिल करने की प्रायिकता 2/3 है तथा बिजली कांट्रैक्ट हासिल न कर पाने की प्रायिकता 5/9 है। कम से कम एक कांट्रैक्ट हासिल करने की प्रायिकता 4/5 है। दोनों कांट्रैक्ट हासिल करने की प्रायिकता ज्ञात करो।

9. एक छात्र दो फर्मों X तथा ल् में नौकरी के लिए आवेदन करता है। X फर्म में चयनित होने की प्रायिकता 0.7 है तथा ल् फर्म में 0.5। दोनों फर्मों में उसका अभ्यर्थन निरस्त होने की प्रायिकता 0.6 है। किसी एक फर्म में चयनित होने की प्रायिकता बताओ।

10. एक कक्षा के 3 पेपरों A, B तथा C की परीक्षा का परिणाम दिया गया है। यह अनुमान लगाया जाता है कि 40 प्रतिशत विद्यार्थी पेपर । में फेल हुए, 30 प्रतिशत पेपर B में फेल हुए तथा 25 प्रतिशत पेपर C में फेल हुए, 15 प्रतिशत पेपर । तथा B दोनों में फेल हुए, 12 प्रतिशत पेपर B तथा C में फेल हुए, 10 प्रतिशत पेपर । तथा C में फेल हुए तथा 3 प्रतिशत सभी पेपरों में फेल हुए। दैव निदर्शन से चयनित छात्र के कम से कम एक पेपर में चयनित होने की प्रायिकता ज्ञात करो।

स्वपरख प्रश्नों के उत्तर

1.	14/52	2.	8/25
3.	13/20	4.	1/54
5.	1/54	6.	4/13
7.	8/25	8.	14/15
9.	0.8	10.	0.39

8.13 संदर्भ पुस्तकें (Biblography)

1. Roy Ramendu, ***‘Principles of Statistics’***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***“Business Statistics”*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***“Advanced Statistics”,*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***“Basic Statistics”*** World Press Limited, Calcutta.
5***. “Fundamentals of Business Statistics”*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **“Quantitative Methods in Management”.**

इकाई 9 द्विपद तथा प्वॉयसन वितरण
(Binomial and Poisson Distribution)

उद्देश्य (Objective)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ वास्तविक आवृत्ति वितरण की व्याख्या कर सकें।
- ✓ सैद्धान्तिक या प्रायिकता वितरण की व्याख्या कर सकें।
- ✓ द्विपद प्रायिकता वितरण की व्याख्या कर सकें।
- ✓ प्वॉयसन वितरण एवं उनके उपयोग का वर्णन कर सकें।

9.1 प्रस्तावना (Introduction)

सांख्यिकी में विभिन्न तरह के वितरणों का अध्ययन किया जाता है। वे प्रमुखतया दो श्रेणी में वर्गीकृत किये गये हैं। पहला वास्तविक आवृत्ति वितरण तथा दूसरा सैद्धान्तिक या प्रायिकता वितरण। पहला वितरण वास्तविक observation समंकों या घटना/प्रयोग पर आधारित है तथा जबकि दूसरा न तो वास्तविक समंकों या प्रयोग पर आधारित है और न ही उससे जनित है।

9.2 वास्तविक आवृत्ति वितरण (Real Frequency Distribution)

वास्तविक आवृत्ति वितरण उन आवृत्ति वितरणों को कहते हैं जो वास्तविक समंक या प्रयोग से प्राप्त किये जाते हैं। उदाहरण के लिए, किसी कक्षा के 70 विद्यार्थियों को प्राप्त हुए अंकों का वास्तविक वितरण निम्न प्रकार है :

अंक	0-10	10-20	20-30	30-40	40-50
विद्यार्थियों की संख्या	5	15	20	25	5

वास्तविक आवृत्ति वितरणों का विश्लेषण सामान्यतया विभिन्न सांख्यिकीय उपकरणों जैसे औसत, अपकिरण तथा विषमता इत्यादि की सहायता से करते हैं।

9.3 सैद्धान्तिक या प्रायिकता वितरण (Theoretical or Probability Distribution)

सैद्धान्तिक आवृत्ति वितरण उन वितरणों को कहते हैं जो वास्तविक समंकों या प्रयोगों से प्राप्त नहीं होते परन्तु निश्चित मान्यताओं के आधार पर गणितीय विधि से प्राप्त किए जा सकते हैं। सैद्धान्तिक आवृत्ति वितरण को प्रायिकता वितरण या प्रत्याशित आवृत्ति वितरण कहते हैं। उदाहरण के लिए यदि 4 सिक्कों को 160 बार टास किया जाए तथा चित आने की घटना को सफलता मानी जाए तो प्रायिकता सिद्धान्त के आधार पर प्रत्याशित आवृत्ति वितरण निम्न प्रकार से है :

सफलता की संख्या	**प्रायिकता**	**प्रत्याशित आवृत्ति**
(X)	**(p)**	**Expected Frequency**
0	1/16	160*1/16 = 10
1	4/16	160*4/16 =10
2	6/16	160*6/16 =10
3	4/16	160*4/16 = 10
4	1/16	160*1/16 = 10
योग	p= 1	160

अतः सैद्धान्तिक आवृत्ति वितरण वास्तविक अवलोकन पर आधारित नहीं है जबकि निश्चित मान्यताओं के आधार पर गणितीय निकाला जाता है।

9.4 सैद्धान्तिक आवृत्ति वितरण के उपयोग (Uses of Theoretical Frequency Distribution)

सैद्धान्तिक वितरण के उपयोग निम्न है :

1. दिये गये वितरण की प्रकृति के विश्लेषण में सैद्धान्तिक वितरण उपयोगी है।
2. सैद्धान्तिक आवृत्ति वितरण से प्राप्त प्रत्याशित आवृत्तियाँ ताकि्रक निर्णय लेने में सहायक है।

3. वास्तविक और प्रत्याशित आवृत्तियों की तुलना करने में सैद्धान्तिक आवृत्ति वितरण सहायक है तथा दोनों के बीच का अंतर सार्थक है या अथवा यह अंतर सभी अन्य कारण से है।
4. भविष्यवाणी, प्रक्षेपण तथा पूर्वानुमान में सैद्धान्तिक वितरण सहायक है।
5. सैद्धान्तिक वितरण कई व्यावसायिक तथा अन्य समस्याओं को हल करने में सहायक है। प्वॉयसन वितरण गुणवत्ता नियंत्रण से संबंधित महत्वपूर्ण निर्णय लेने में सहायक है।
6. ऐसी परिस्थितियों में जब वास्तविक प्रयोग कर पाना संभव न हो अथवा वास्तविक अवलोकन प्राप्त करने में बहुत पैसा लगता हो तो अवलोकित आवृत्ति वितरण के स्थान पर सैद्धान्तिक आवृत्ति वितरण का उपयोग किया जाता है।

9.5 सैद्धान्तिक तथा प्रायिकता वितरण के प्रकार (Theoretical and Types of Probability Distributions)

सैद्धान्तिक वितरण के प्रमुख प्रकार :

अ. प्रायिकता वितरण

1. द्विपद वितरण
2. प्वॉयसन वितरण

ब. सतत प्रायिकता वितरण

1. सामान्य वितरण

इस अध्याय में आप सिर्फ द्विपद तथा प्वॉयसन वितरण का अध्ययन करेंगे। सामान्य वितरण पर अगले अध्याय में चर्चा करेंगे।

9.6 द्विपद वितरण (Binomial Distribution)

द्विपद वितरण एक असतत प्रायिकता वितरण है। इस वितरण की खोज स्विस गणितज्ञ जेम्स बर्नाली ने की। इसका प्रयोग ऐसी परिस्थितियों में करते हैं जब प्रयोग का परिणाम सफलता अथवा असफलता आए। द्विपद वितरण एक असतत प्रायिकता वितरण है जो कि दो संभावितों सफलता तथा असफलता की प्रायिकता बताता है।

9.6.1 द्विपद वितरण की परिभाषा :

द्विपद वितरण को निम्न प्रायिकता फंक्शन से परिभाषित करते हैं

$$P(X = x) = {}^nC_x\, q^{n-x}\, p^x$$

जहाँ p =सफलता की प्रायिकता

q = असफलता की प्रायिकता= 1-p

n= परखों की संख्या

P(X = x) = n परखों में x सफलता की प्रायिकता

द्विपद वितरण के प्रायिकता फंक्शन में X के भिन्न मान -2 रखने पर 0,1,2, ...n सफलता प्राप्त करने की प्रायिकता निम्न है

सफलता की संख्या (X)	सफलता की प्रायिकता p(X=x)
0	${}^nC_0\, q^{n-0}\, p^0 = q^n$
1	${}^nC_1\, q^{n-1}\, p^1 = nq^{n-1}p$
2	${}^nC2\, q^{n-2}\, p^2 = n(n-1)/2 \times (q^{n-2}p)$
….	${}^nC_{..}\, q^{n-..}\, p^{..} = nq^{n+...}p$
X	${}^nC_x\, q^{n-x}\, p^x$
N	${}^nC_n\, q^{n-n}\, p^n = p^n$

9.6.2 द्विपद वितरण लागू करने की मान्यताएँ/परिस्थितियाँ

द्विपद वितरण का प्रयोग केवल निम्न परिस्थितियों में ही कर सकते हैं :

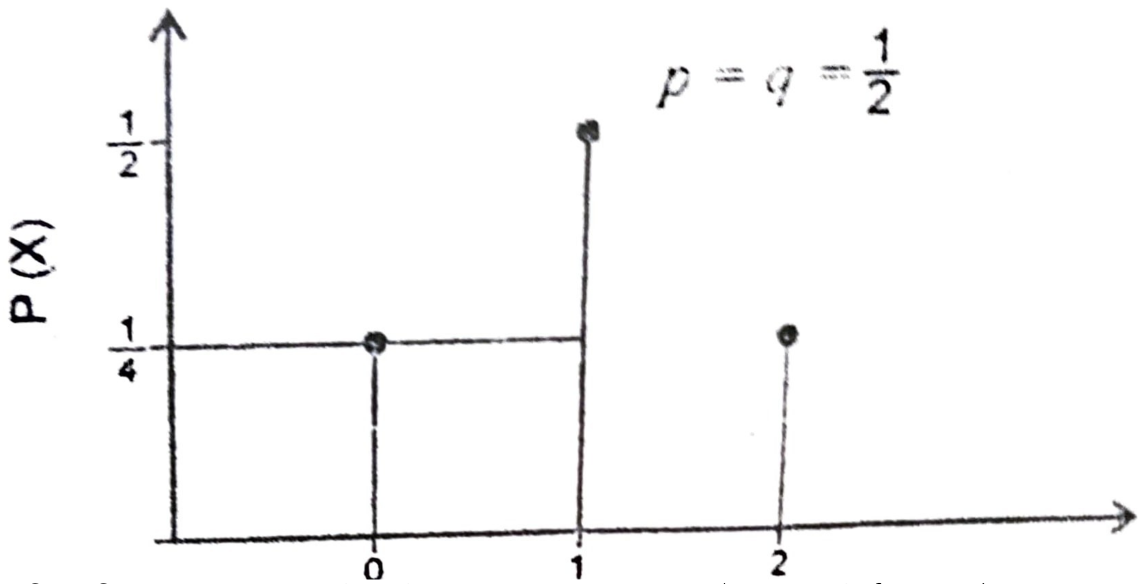
1. **परखों की निश्चित संख्या-** यादृच्छिक प्रयोगों की पुनरावृत्ति एक निश्चित संख्या में होती है। अन्य शब्दों में परख की संख्या n निश्चित एवं स्थिर होती है।
2. **पारस्परिक अपवर्जी परिणाम-** प्रत्येक परख के दो पारस्परिक अपवर्जी परिणाम-सफलता या असफलता होते हैं। उदाहरण के लिए यदि एक सिक्के को उछाला जाए तो या चित आएगा या पट।
3. **प्रत्येक परख में सफलता की प्रायिकता स्थिर है-** सफलता की प्रायिकता जिसे p से प्रदर्शित करते हैं, प्रत्येक परख में स्थिर होती है। दूसरे शब्दों में, सफलता की प्रायिकता प्रत्येक परख में स्थिर रहती है। उदाहरण के तौर पर, एक सिक्के को उछालने पर चित आने की प्रायिकता बराबर रहती है। अतः $p = p(H) = 1/2$।
4. **परख स्वतंत्र है-** द्विपद वितरण में परख स्वतंत्र होते हैं अर्थात् एक परख के परिणाम का अन्य परख के परिणामों पर कोई प्रभाव नहीं डालता।

9.6.3 द्विपद वितरण की विशेषताएँ

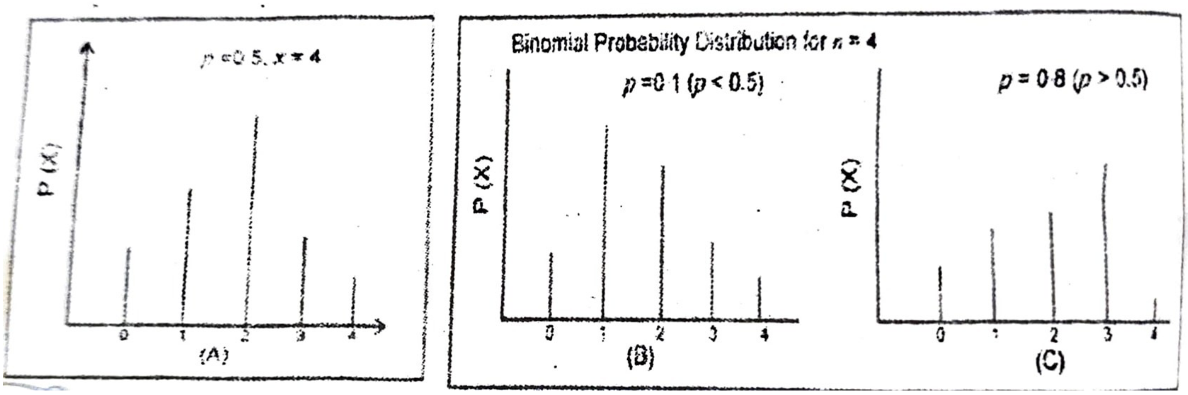
द्विपद वितरण की प्रमुख विशेषताएं निम्नलिखित है :

1. द्विपद वितरण एक सैद्धान्तिक आवृत्ति वितरण है जो कि बीजगणित के द्विपद प्रमेय पर आधारित है।
2. **असतत प्रायिकता वितरण**-द्विपद वितरण एक असतत प्रायिकता वितरण है जिसमें सफलता की संख्या 0, 1, 2, ...n पूर्ण संख्या में दी जाती है।
3. **पंक्ति आरेखः** द्विपद वितरण एक पंक्ति आरेख की सहायता से भी प्रदर्शित कर सकते हैं। सफलता की संख्या (X) को X-अक्ष पर तथा सफलता की प्रायिकता (p) को Y-अक्ष पर दर्शाते हैं। निम्नलिखित पंक्ति आरेख एक सिक्के को दो बार उछालने पर प्राप्त होता है।

चित की संख्या (X)	प्रायिकता
0	$^2C_0\,(1/2)^2 = 1/4$
1	$^2C_1\,(1/2)^2\ (1/2)^1 = 1/2$
2	$^2C_2\,(1/2)^2 = 1/4$



4. **द्विपद वितरण का आकार -** द्विपद वितरण का आकार p तथा q के मान पर निर्भर करता है।
   1. यदि $p = q = 1/2$ तो द्विपद वितरण सममित है। (Symmetrical) (चित्र A देखें)
   2. जब p= ½, तब द्विपद वितरण असममित है (asymmetrical)। यदि $p<q$ अर्थात् $(p< 1/2)$ तो वितरण धनात्मक में विषमता पायी जाती है और $p>q$ अर्थात् $(p> 1/2)$ तो वितरण में ऋणात्मक विषमता पाई जाती है। चित्र (B) तथा (C) देखें



5. प्रमुख पाचाल (Main Parameters):
द्विपद वितरण के दो प्राचाल है n तथा p। इन दो प्राचालों की मदद से पूरा वितरण ज्ञात कर सकते हैं।

6. द्विपद वितरण के स्थिर (constants) :
द्विपद वितरण के स्थिर निम्न सूत्र के आधार पर प्राप्त किया जा सकता है।

माध्य (X) = np
प्रसरण = σ^2 = npq
प्रमाप विचलन = S.D. = $\sqrt{npq}$
संवेग विषमता गुणांक = $\sqrt{\beta_1}$ = q-p/$\sqrt{npq}$
संवेग पृथुशीर्षत्व गुणांक = $\sqrt{\beta_2}$ = 3 + (1-6pq)/npq

7. उपयोग : यह उन क्षेत्रों में उपयोगी है जहाँ निष्कर्ष सफलता और असफलता में वर्गीकृत किया जा सकता है। दूसरे शब्दों में, सिक्कों के प्रयोग पांसा फेंकने, मदों/किसी कंपनी द्वारा चीजों का निर्माण इत्यादि में उपयोगी है।

9.7 द्विपद वितरण के उपयोग (Uses of Binomial Distribution)

अब आप द्विपद वितरण के उपयोग का इस प्रकार अध्ययन करेंगे :

अ. द्विपद वितरण सूत्र का उपयोग :

जब किसी समस्या से संबंधित घटना के घटित होने की प्रायिकता आपको ज्ञात हो अर्थात् p तथा q का मान ज्ञान हो तो n परख में से x सफलता घटित होने की प्रायिकता निम्न सूत्र से ज्ञात करते हैं :

$$p\,[X=x] = = {}^{n}C_{x}\, q^{n-x}\, p^{x}$$

उदाहरण 9.1 : एक सिक्के को तीन बार उछालें जाने पर प्रायिकता ज्ञात करें।

1. पूर्णतया 2 चित
2. कम से कम 2 चित
3. अधिकतम 2 चित

हल : माना कि चित आने की प्रायिकता = p = 1/2

q = पट आने की प्रायिकता = 1/2

और n = 3, $p\,[X=x] = {}^{n}C_{x}\, q^{n-x}\, p^{x}$

(i) p(2H) = ${}^{3}C_{2}\,(1/2)^{2}\,(1/2)^{1}$ = 3 x1/2 x1/4 = 3/8

(ii) P(कम से कम 2 चित) = P(2H) + P(3H)
= ${}^{3}C_{2}\,(1/2)^{2}\,(1/2)^{1} + {}^{3}C_{3}\,(1/2)^{0}\,(1/2)^{3}$
= 3 x 1/8 + 1 x 1/8 = 4/8 =1/2

(iii) P (अधिकतम 2 चित) = P(0H + P(1H) + P(2H)
= 1-P(3H)
=1 - ${}^{3}C_{3}\,(1/2)^{0}\,(1/2)^{3}$ = 1- 1/8 = 7/8

ब. X तथा σ से n, p और q प्राप्त करना

यदि आपके पास द्विपद वितरण का माध्य (X) तथा प्रसरण (σ^2) या प्रमाप विचलन (σ) का मान ज्ञात हो आप दए p तथा q का मान ज्ञात कर सकते हैं। निम्नलिखित उदाहरण इस प्रकिया को समझाएगा।

उदाहरण 9.2 : एक द्विपद वितरण का माध्य 20 तथा प्रमाप विचलन 4 है। n, p तथा q का मान ज्ञात करो।

हल : द्विपद वितरण में, माध्य = np

प्रमाप विचलन, SD = √npq

X = np = 20 ....(1)

σ = √npq = 4 ....(2)

दोनों तरह वर्ग करने पर,

σ = npq = 16 .... (3)

3. को 2 से विभाजित करने पर

npq/np = 16/20

q = 16/20 = 4/5

p = 1 –q = 1-4/5 = 1/5

p का मान 1. में रखने पर

n x 1/5 = 20

अतः n = 100, p = 1/5 तथा q= 4/5

उदाहरण 9.3 : एक द्विपद वितरण जिसका माध्य 6 तथा प्रसरण 2 हो तो 5 सफलता आने की प्रायिकता ज्ञात करो।

हल : द्विपद वितरण में

माध्य = np = 6 (1)

प्रसरण = npq = 2 (2)

2. को 1. से विभाजित करने पर

npq/np = 2/6

q = 1/3

p = 1- 1/3 = 2/3

p का मान 1. में रखने पर

n x 2/3 = 6

यहाँ पर, n = 9, p= 2/3, q 1/3

अतः 126 x 32/19683 = 0.2048

स. X तथा σ का मान ज्ञात करना जब n, p तथा q का मान ज्ञात हो।

उदाहरण 9.4: यदि दोषयुक्त बोल्ट की प्रायिकता 0.1 हो 500 में से दोषयुक्त बोल्ट के वितरण का 1. माध्य तथा 2. प्रमाप विचलन ज्ञात करें। विषमता तथा पृथुशीर्षत्व गुणांक प्राप्त करें।

हल : दिया गया है,

p = 0.1 q = 1- 0.1 = 0.9, n= 500

1. माध्य = np = 500 x 0.1 = 50
2. प्रमाप विचलन = σ = √npq = √500 x 0.1 x 0.9
3. विषमता गुणांक =$\sqrt{\beta_1}$ = q-p/√npq = (0.9 -0.1)/ (500 x 0.1 x 0.9)
 = 0.8/6.70 = 0.119
4. पृथुशीर्षत्व गुणांक = $\sqrt{\beta_2}$ = 3 + (1-6pq)/npq
 = 3+ [1-6(0.1)(0.9)]/[(500 x 0.1 x 0.9)]
 = 3.010

उदाहरण 9.5: एक सिक्के को 100 बार टॉस करने पर चित आने की संख्या के वितरण का माध्य तथा प्रमाप विचलन ज्ञात करो।

हलः $n= 100, P(H) = p= ½, \; q = 1/2$

∴माध्य $= np = 100 \times 1/2 = 50$

प्रमाप विचलन = $\sigma = \sqrt{npq} = \sqrt{100 \times ½ \times ½} = \sqrt{25} = 5$

द. द्विपद वितरण को फिट करना-

वास्तविक डाटा/अवलोकन पर द्विपद वितरण निम्नलिखित प्रक्रिया से फिट करते हैं।

1. दी गई जानकारी के आधार पर p तथा q का मान ज्ञात करो।
2. n तथा N का मान नोट करो, जहां n एक प्रयोग के परखों की संख्या है तथा N सभी प्रयोगों के सभी परखों की कुल संख्या है।
3. दिये गये प्रयोग से आने वाले सभी संभावित सफलताओं की संख्या की प्रायिकता ज्ञात करो।
4. इन प्रायिकताओं को N से गुणा करने पर आप वांछित प्रत्याशित आवृत्ति प्राप्त करते हैं।

उदाहरण 9.6 : 4 सिक्कों को 160 बार उछालने पर निम्नलिखित परिणाम प्राप्त होता हैः

चित की संख्या	0	1	2	3	4
आवृत्ति	17	52	54	31	6

सिक्कों को निष्पक्ष मानते हुए द्विपद वितरण फिट कीजिए।

हल : माना कि सिक्के निष्पक्ष है तो चित आने की प्रायिकता (p) तथा पट आने की प्रायिकता (q), 1/2 तथा 1/2 है।

यदि $n = 4, N = 160$

अतः 0,1,2,3,4 चित आने की प्रायिकता निम्नलिखित सूत्र से ज्ञात करते हैं

$$p\,[X=x] = = {}^{n}C_{x}\, q^{n-x}\, p^{x}$$

प्रत्याशित प्रायिकता ज्ञात करने के लिए प्रायिकता को N से गुणा करो।

प्रत्याशित प्रायिकता इस प्रकार प्राप्त कर सकते हैं।

चित की संख्यां (n)	**प्रत्याशित आवृत्ति** $= N\,{}^{n}C_{x}\, q^{n-x}\, p^{x}$
0	$= 160 \times {}^{4}C_{0}\,(1/2)^{4}\,(1/2)^{0} = 10$
1	$= 160 \times {}^{4}C1\,(1/2)^{3}\,(1/2)^{1} = 40$
2	$= 160 \times {}^{4}C_{2}\,(1/2)^{2}\,(1/2)^{2} = 60$
3	$= 160 \times {}^{4}C_{3}\,(1/2)^{1}\,(1/2)^{3} = 40$
4	$= 160 \times {}^{4}C_{3}\,(1/2)^{0}\,(1/2)^{4} = 10$

9.8 प्वॉयसन वितरण (Poisson Distribution)

प्वॉयसन वितरण एक असतत् प्रायिकता वितरण है और सांख्यिकीय कार्यों में इसका बहुत प्रयोग होता है। 1837 में फ्रेंच गणितज्ञ डा0 सिमन डेनिस प्वॉयसन ने इस वितरण का विकास किया तथा उनके नाम पर इस वितरण का नाम रखा गया। प्वॉयसन वितरण का प्रयोग उन स्थितियों में करते हैं जब किसी घटना के घटित होने की प्रायिकता बहुत छोटी हो अर्थात् घटना बहुत दुर्लभ हो। उदाहरण के लिए, एक निर्माण करने वाली कंपनी में दोषयुक्त मद की प्रायिकता बहुत कम है, एक वर्ष में भूकंप आने की प्रायिकता बहुत कम है, रोड/रास्ता/सड़क पर होने वाले दुर्घटना की प्रायिकता बहुत कम है। ये सभी ऐसी घटनाओं के उदाहरण है जब घटना के होने की प्रायिकता बहुत कम है।

9.8.1 द्विपद वितरण का सीमित रूप में प्वॉयसन वितरण

निम्नलिखित परिस्थितियों में द्विपद वितरण के सीमित रूप में प्वॉयसन वितरण को प्राप्त किया जाता है :

1. n परखों की संख्या अंतत रूप से बड़ी हो, अर्थात् $n \dashrightarrow \infty$

2. p, सफलता की प्रायिकता बहुत कम हो तथा q असफलता की प्रायिकता बहुत अधिक हो अर्थात् p— 0, q--- 1.
3. सफलता की औसत संख्या (np) एक धनात्मक निश्चित संख्या (m) के बराबर है अर्थात् np =̄ m जहाँ m वितरण का प्राचाल (parameter) है।

9.8.2 प्वॉयसन वितरण की परिभाषा :

द्विपद समीकरण से

$^{n}C_{x} = p^{x}q^{n-x}$

चूंकि np =̄ m ⇒ p= m/n अतः q = 1-m/n

$= n^{x}p^{x}q^{n-x}$

$=(1-1/n)(1-2/n)..........[1-(x-1)/n] . (np)^{x}q^{n-x} / x!$

$=(1-1/n)(1-2/n)..........[1-(x-1)/n] [1-m/n]^{n-x}. m^{x} / x!$

जब $n \text{---} \infty = e^{-m} .m^{x} / x!$

प्वॉयसन वितरण निम्नलिखित प्रायिकता फंक्शन से परिभाषित और प्रदर्शित किया जाता है

$$P(X =x) = e^{-m} .m^{x} / x!$$

जहाँ P(X=x) = x सफलता प्राप्त करने की प्रायिकता

m = np = वितरण का प्राचाल

म = 2.7183

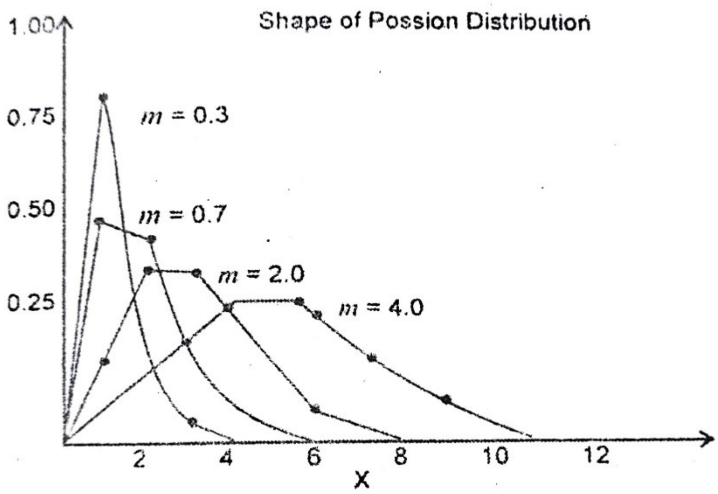
उपरोक्त प्रायिकता वितरण में के विभिन्न मान रखने पर सफलता प्राप्त करने की प्रायिकता इस प्रकार प्राप्त कर सकते हैं :

सफलता की संख्या (X)	प्रत्याशित P(x)
0	$e^{-m} .m^{0} / 0! = e^{-m}$
1	$e^{-m} .m^{1} / 1! = m\ e^{-m}$
2	$e^{-m} .m^{2} / 2! = m^{2}/2 . e^{-m}$
3	$e^{-m} .m^{3} / 3! = e^{-m}$
....	
n	$e^{-m} .m^{n} / n! = m^{x}/x!\ e^{-m}$

9.9 प्वॉयसन वितरण की विशेषताएँ (Properties of Poisson Distribution)

प्वॉयसन वितरण की प्रमुख विशेषताएं निम्नलिखित हैं :

1. **असतत प्रायिकता वितरण :** प्वॉयसन वितरण एक असतत प्रायिकता वितरण है जहाँ पर सफलताओं की संख्या पूर्ण संख्या जैसे 0,1,2............ इत्यादि के रूप में हैं।
2. **p तथा q का मान :** प्वॉयसन वितरण उन स्थितियों में प्रयोग करते हैं जहाँ घटना के घटित होने की प्रायिकता करते हैं जहाँ है अर्थात् (p--0) तथा घटना के घटित न होने की प्रायिकता बहुत अधिक हो अर्थात् (q--0) तथा द का मान अनन्त रूप से बड़ा हो।
3. **प्रमुख प्राचाल :** इसका बस एक ही प्राचाल है और उसका मान np के बराबर है अर्थात् m = np। इस प्राचाल की सहायता से पूरा वितरण जाना जा सकता है।
4. **प्वॉयसन वितरण का आकार :** प्वॉयसन वितरण हमेशा धनात्मक विषम है लेकिन m का मान बढ़ने से विषमता घटती है। m का मान बढ़ाने पर वितरण दायीं और खिसकता है और विषमता स्तर गिरता है। जो कि निम्नलिखित चित्र् द्वारा दर्शाया जाता है :



5. प्वॉयसन वितरण के स्थिर : प्वॉयसन वितरण के निम्नलिखित सूत्र से ज्ञात कर सकते हैं:

माध्य = $X = np$
विषमता गुणांक संवेग = $\sqrt{\beta} = 1/\sqrt{m}$
प्रसरण = $\sigma^2 = m$
प्रमाप विचलन = $S.D. = \sigma = \sqrt{m}$
पृथुशीर्षत्व गुणांक संवेग = $\beta_2 = 3 + 1/m$

6. माध्य तथा प्रसरण की बराबरी : प्वॉयसन वितरण की एक महत्वपूर्ण विशेषता यह है कि इसका माध्य और प्रसरण बराबर है अर्थात् $X = \sigma^2$ अथवा माध्य प्रसरण।

9.10 प्वॉसन वितरण की महत्ता (Importance of Poisson distribution)

प्वॉयसन वितरण निम्नलिखित क्षेत्रों में बहुतायत उपयोग होता है :

1. इसका प्रयोग सांख्यिकीय गुणवत्ता नियंत्रण में दोषयुक्त मदों की गणना करना।
2. बायोलाजी में, बैक्टीरिया की संख्या की गणना करना।
3. इन्श्योरेन्स में, कारणों की गणना में समस्या।
4. एक टाइप किए गए पेज पर टाइपिंग के दौरान होने की गलतियों की संख्या की गणना।
5. एक टाउन में आने वाली फोन काल की संख्या।
6. एक ब्लेड बनाने वाली कंपनी के एक लॉट में दोषयुक्त ब्लेड की संख्या की गणना करना।
7. एक टाउन में एक क्रासिंग पर रोड दुर्घटना में होने वाली मौतों की संख्या।
8. एक वर्ष में लर्क्स प्वाइंट पर होने वाली आत्महत्या की संख्या।

साधारणतया प्वॉयसन वितरण का उपयोग दुर्लभ घटनाओं में होता है जहाँ सफलता की प्रायिकता (p) बहुत कम है और n का मान बहुत अधिक है।

9.11 प्वॉयसन वितरण के उपयोग (Uses of Poisson distribution)

प्वॉयसन वितरण के उपयोग का अध्ययन इस प्रकार है :

अ. प्वॉयसन वितरण सूत्र का प्रयोग :

आप प्वॉयसन वितरण सूत्र के उपयोग का अध्ययन निम्न दो अलग तरह की परिस्थितियों में कर सकते हैं : 1. जब p का मान ज्ञात हो तथा 2. जब m का मान ज्ञात हो।

1. जब p का मान ज्ञात हो :

उदाहरण 9.14 : यह ज्ञात है कि एक पेंp का निर्माण करने वाली कंपनी में 2% पेंच दोषयुक्त निर्मित होते हैं। प्वॉयसन वितरण का प्रयोग करते हुए 100 पेंच वाले एक पैकेट में 1. एक भी दोषयुक्त पेंच न हो 2. एक दोषयुक्त पेंच हो तथा 3. दो या उससे अधिक दोषयुक्त पेंच होने की प्रायिकता ज्ञात करो। (दिया है = 0.135)

हल : माना कि p = पेंच के दोषयुक्त होने की प्रायिकता = 2% = 2/100

p = 2/100 य n = 100

चूंकि m = np = 100 × 2/100 = 2

प्वॉयसन वितरण इस प्रकार

$P(X=x) = P(X=0)= (e^{-2}\ 2^{0})/0!$

$= e^{-2} = 0.135$

2. P (एक दोषयुक्त) = $P(X=1) = (e^{-2}\ 2^{1})/1! = 0.135 \times 2 = 0.270$

3. P(दो या उससे अधिक दोषयुक्त) = $1-[P(0) + P(1)]$

$= 1 - [0.135 + 0.270] = 1-0.405 = 0.595$

उदाहरण 9.15 : एक पिन का निर्माण करने वाला जानता है कि औसतन उसके 5% उत्पाद दोषयुक्त है। वह 100 पिन का पैकेट बनाकर उन्हें बेचता है और विश्वास दिलाता है कि पैकेट में 4 से अधिक दोषयुक्त पिन नहीं है। पैकेट के विश्वसनीय गुणवत्ता पर खरा उतरने की प्रायिकता ज्ञात करो।

हल : माना कि p = दोषयुक्त पिन की प्रायिकता = 5% = 5/100

दिया है n = 100, p= 5/100

चूंकि m = np = 100 × 5/100 = 5

प्वॉयसन वितरण

$P(X=x) = (e^{-m}\ m^{x})/x!$

वांछित प्रायिकता = P[पैकेट गुणवत्ता पर खरा उतरे]

= P [पैकेट में अधिकतम 4 दोषयुक्त]

$= P(0)+P(1)+P(2)+P(3)+P(4)$

$= (e^{-5}\ 5^{0})/0! +(e^{-5}\ 5^{1})/1!+ (e^{-5}\ 5^{2})/2!+ (e^{-5}\ 5^{3})/3!+ (e^{-5}\ 5^{4})/4!$

$= 0.0067 \times 65.374 = 0.438$

2. जब M का मान ज्ञात हो:

उदाहरण 9.16 : दोपहर 2 से 4 के बीच किसी कंपनी के स्विच बोर्ड पर औसतन 2.5 फोन काल प्रति मिनट आता है। प्रायिकता ज्ञात करो जब एक मिनट में 1. एक भी फोन कॉल न आए 2. पूर्णतया 3 काल 3. कम से कम 2 काल।

(दिया है $e^{-2} = 0.1353$, $e^{-5} = 0.6065$)

हल : यह प्वॉयसन वितरण की समस्या है।

$P(X=x) = (e^{-m}\ m^{x})/x!$ where X = 1,2,3,........

औसतन फोन काल की संख्या = X = m = 2.5

प्वॉयसन वितरण $P(X=x) = (e^{-m}\ m^{x})/x!$

1. P(एक भी फोन काल न आए) = $P(x=0)$

$= e^{-2.5}\ (2.5)0\ /\ 0! = e^{-2.5}$

$= e^{-2}\ e^{-0.5}$ $(e^{-2} = 0.1353,\ e^{-0.5} = 0.6065)$

$= 0.1353 \times 0.6065 = 0.0821$

अतः एक मिनट में एक भी फोन काल न आने की प्रायिकता = 0.0821

2. पूर्णतया 3 काल = $P(X=3) = e^{-2.5}\ (2.5)3\ /\ 3!$

$= (0.0821)\ (2.5)3\ /\ 3\times2\times1 = 0.2138$

3. P(कम से कम 2 काल आए)= $1-[P(X)=0 + P(X=1)$

$= 1- [e^{-2.5} +(2.5)\ e^{-2.5}]$

$= 1- [e^{-2.5}\ (1+2.5)] = 1- [(0.0821)\ (3.5)]$

$= 1- 0.28735 = 0.71265$

उदाहरण 9.17 : पूर्व अनुभव के आधार पर ज्ञात है कि एक कंपनी में औसतन औद्योगिक दुर्घटना की संख्या प्रति मासिक 4 है। एक महीने में 4 से कम दुर्घटना होने की प्रायिकता ज्ञात करो। प्वॉयसन वितरण का प्रयोग करके उत्तर की व्याख्या करो। (दिया गया है $e^{-4} = 0.0183$)

हल : औसत दुर्घटना की संख्या = X = m= 4

$P(X) = (e^{-m} m^{x})/x!$

$P(0) = (e^{-m} m^{0})/0! = e^{-m} = e^{-4}$

$P(1) = (e^{-m} m^{1})/1! = (e^{-m} .m) = 4 e^{-4}$

$P(2) = (e^{-m} m^{2})/ 2! = m^{2}/2! e^{-m} = (4)2/2!. e^{-4}$

$P(3) = (e^{-m} m^{3})/ 3! = m^{3}/ 3! e^{-m} = (4)3/3!. e^{-4}$

4 से कम दुर्घटना होने की प्रायिकता

$= P(0) + P(1) + P(2) + P(3)$

$= e^{-4} ({}1 + 4 + 4^{2}/2! + 4^{3} /3!,$

$= 0.0183 (1 + 4 + 8 + 10.67)$ $(e^{-4}= 0.0183)$

$= 0.0183 \times 23.67 = 0.4332$

अतः 4 से कम दुर्घटना की प्रायिकता 0.4332 या 43.32% ।

उदाहरण 5.18 : एक प्वॉयसन प्रायिकता वितरण जिसमें घटना के घटित होने का औसत प्रति समयकाल 2 है।

1. उचित प्वॉयसन प्रायिकता फंक्शन लिखें।
2. 3 समयकाल में औसतन घटित होने की प्रायिकता बताओ।
3. 3 समयकाल में 6 घटना के घटित होने की प्रायिकता ज्ञात करो।

हल : 1. समयकाल में औसतन घटित होने की संख्या = उ = 2

1. प्वॉयसन प्रायिकता फंक्शन = P(ग्त्रग) = (मॅ2 2X)/ X!

2. 3 समयकाल में औसत घटनाओं की संख्या = 2×3= 6

3. च्ग्त्र 6, = मॅ6 (6)6 / 6! = 0.1575

5.12 प्वॉयसन वितरण को फिट करना (Fitting the Poisson Distribution)

अवलोकित डाटा पर प्वॉयसन वितरण फिट करने के लिए निम्नलिखित प्रक्रिया अपनाते हैंः

1. सर्वप्रथम वास्तविक आवृत्ति से माध्य (X) की गणना निम्न सूत्र से करें।

$X = \sum fx / n$

आप माध्य के इस मान को प्वॉयसन वितरण का प्राचाल की तरह उपयोग करेंगे अर्थात् x = m

2. e^{-m} का मान प्राप्त करें। यदि e^{-m} का मान प्रश्न में न दिया गया हो तो निम्न सूत्र से ज्ञात करें :

e^{-m} = reciprocal [antilog(m x 0.4343)],

3. 0,1,2,3, या x सफलता की प्रायिकता निम्न प्वॉयसन प्रायिकता वितरण सूत्र से गणना करते हैं।
$P(X=x) = e^{-m} m^{x} / x!$

4. प्रत्याशित या सैद्धान्तिक आवृत्ति प्राप्त करने के लिए उपर निकाली गई प्रायिकता को N (कुल आवृत्ति) आप मध्य से गुणा करें। अतः

सफलता की संख्या X	प्रायिकता P(X)	प्रत्याशित प्रायिकता fe(X)
0	$P(0)= e^{-m} m^{0} / 0! = e^{-m}$	$N.P(0)= N e^{-m}$
1	$P(1)= e^{-m} m^{1} / 1! = me^{-m}$	$N.P(1)= N m e^{-m}$

2	$P(2) = e^{-m} m^2 / 2! = m^2/2!\ e^{-m}$	$N.P(2) = N\ m^2/2!\ e^{-m}$
......		
x	$P(x) = m^x / x!\ e^{-m}$	$N.P(x) = N\ e^{-m}\ m^x / x!$

वैकल्पिक प्रक्रिया :

प्रत्याशित आवृत्तियों को निम्न प्रकार से आसानी से ज्ञात कर सकते हैं :

1. सर्वप्रथम गणना करें $Fe(0) = N.\ p(0) = N\ e^{-m}$
2. अन्य प्रत्याशित आवृत्तियों को निम्न प्रकार से गणना कर सकते हैं :

$Fe\ (0) = N.\ P(0) = N\ e^{-m}$

$Fe\ (1) = m/1\ .\ Fe\ (0)$

$Fe\ (2) = m/2\ Fe\ (1)$

$Fe\ (3) = m/3\ Fe\ (2)$

$Fe\ (4) = m/4\ Fe\ (3)$ और इसी तरह

उदाहरण 9.19 : निम्न डाटा पर प्वॉयसन वितरण फिट करें तथा सैद्धान्तिक आवृत्ति की गणना करें।

मृत्यु	0	1	2	3	4
आवृत्ति	109	65	22	3	1

उपरी वितरण का माध्य तथा प्रसरण का मान ज्ञात करें (दिया है $e^{-0.61} = 0.5432$)

हल : प्वॉयसन वितरण को फिट करना।

मृत्यु (x)	**आवृत्ति (f)**	**fx**
0	109	0
1	65	65
2	22	44
3	3	9
4	1	4
	$\sum f = 200$	$\sum fx = 122$

$$X = \sum fx / \sum f = 122/200 = 0.61$$

$$m = 0.61$$

अब आप $e^{-0.61}$ का मान या तो तालिका से या निम्न सूत्र से ज्ञात कर सकते हैं :

e^{-m} = reciprocal [antilog(m x 0.4343)],

m = 0.61 का मान रखने पर,

$e^{-0.61}$ = = rec [antilog(0.61 x 0.4343)],

= rec [antilog(026492)],

= rec(1.841) = 0.5432

अब $P(0) = e^{-0.61}.\ (0.61)0\ /0!$

$= e^{-0.61} = 0.5432$

प्रत्याशित आवृत्तियों की गणना :

$fe(0) = N\ P(0) = 200 \times (0.5432) = 108.64 \approx 109$

$fe(1) = fe(0) \times m/1 = 108.64 \times 0.61/1 = 66.22 \approx 66$

$fe(2) = fe(1) \times m/2 = 66.27 \times 0.61/2 = 20.21 \approx 20$

$fe(3) = fe(2) \times m/3 = 20.21 \times 0.61/3 = 4.11 \approx 4$

$fe(4) = fe(3) \times m/4 = 4 \times 0.61/4 = 0.61 \approx 1$

अतः

X	0	1	2	3	4
fe	109	66	20	4	1

माध्यम = X = प्रसरण $\sigma^2 = 0.61$

उदाहरण 9.19 : एक पुस्तक के प्रथम 50 पन्नों की प्रूफरीडिंग के समय औसतन प्रति 5 पन्ना 3 गलती/त्रुटि पाई जाती है। प्वॉयसन वितरण की सहायता से 1000 पन्नों वाली पुस्तक में 0,1,2,3,..... त्रुटियों वाले पन्नों की संख्या ज्ञात करें।

हल : त्रुटियों की औसत संख्या = m = 3/5 = 0.6

जहाँ $P(0)= e^{-m}\, m^0 / 0! = e^{-0.6} = 0.5488$

$P(1)= e^{-0.6}\ (00.6)^1 / 1! = 0.5488 \times 0.6 / 1 = 0.32928$

$P(2)= e^{-0.6}\ (00.6)^2 / 2! = 0.5488 \times 0.36 / 2 = 0.098784$

$P(3)= e^{-0.6}\ (00.6)^3 / 3! = 0.5488 \times 0.216 / 3 = 0.0197568$

$P(x>3)= 1- [P(0) + P(1) + P(2) +P(3)]$

$= 1 – [0.5488 + 0.32928 + 0.098784 + 0.197568]$

$= 0.0033792$

प्वॉयसन वितरण को फिट करना

X	P(X)	fe (x) = n. P(x)
0	0.5488	1000 x0.5488 = 548.8
1	0.32928	1000 x 0.32928 = 329.28
2	0.098784	1000 x 0.98784 = 98.74
3	0.0197568	1000 x 0.0197568 =10.76
3 से अधिक	0.0033792	1000 x 0.0033792 = 3.37
		n = 1000

9.13 सारांश (Summary)

सैद्धान्तिक आवृत्ति वितरण उन वितरणों को कहते हैं जो वास्तविक अवलोकनों या प्रयोगों से प्राप्त नहीं किया जाता बल्कि निश्चित मान्यताओं के आधार पर गणितीय विधि से प्राप्त करते हैं। सैद्धान्तिक आवृत्ति वितरण को प्रायिकता वितरण या प्रत्याशित आवृत्ति वितरण कहते हैं। सैद्धान्तिक वितरण के प्रमुख प्रकार हैं 1. द्विपद वितरण 2. प्वॉयसन वितरण तथा 3. सामान्य वितरण। द्विपद वितरण एक असतत प्रायिकता वितरण है बर्नाली ने की। इसका प्रयोग ऐसी स्थितियों में करते हैं जब प्रयोग दो संभावितों सफलता और असफलता में निष्कर्षित हो। प्वॉयसन वितरण एक असतत प्रायिकता वितरण है और फ्रेंच गणितज्ञ डॉ0 सिमन डेनिस प्वॉयसन ने इसका विकास किया। प्वॉयसन वितरण को उन परिस्थितियों में प्रयोग करते हैं जब घटना के घटित होने की प्रायिकता बहुत म हो अर्थात् घटना बहुत दुर्लभ है।

9.14 शब्दावली (Glossary)

- **सैद्धान्तिक आवृत्ति वितरणः** उन वितरणों को कहते हैं जो वास्तविक समंकों या प्रयोगों से प्राप्त नहीं होते परन्तु निश्चित मान्यताओं के आधार पर गणितीय विधि से प्राप्त किए जा सकते हैं।

9.15 अभ्यास प्रश्न (Practice Questions)

1. .... उन आवृत्ति वितरणों को कहते हैं जो वास्तविक समंक या प्रयोग से प्राप्त किये जाते हैं।
2. सैद्धान्तिक आवृत्ति वितरण को प्रायिकता वितरण या ....कहते हैं।
3. द्विपद वितरण की खोज स्विस गणितज्ञ ....ने की।

4. .... का प्रयोग उन स्थितियों में करते हैं जब किसी घटना के घटित होने की प्रायिकता बहुत छोटी हो अर्थात् घटना बहुत दुर्लभ हो।

9.16 अभ्यास प्रश्नों के उत्तर

1. वास्तविक आवृत्ति वितरण 2. प्रत्याशित आवृत्ति वितरण
3. जेम्स बर्नाली 4. प्वॉयसन वितरण

9.17 स्वपरख प्रश्न

1. एक सिक्के को 6 बार उछाला जाता है। चार या उससे अधिक चित आने की प्रायिकता ज्ञात करें।
2. एक पांसा को 4 बार फेंका जाता है। 2 से अधिक आने पर उसे सफलता माना जाता है। प्रायिकता ज्ञात करो।
   अ. पूर्णतया 1 सफलता
   ब. 3 से कम सफलता
   स. 3 से अधिक सफलता
3. सैद्धान्तिक आवृत्ति वितरण से आप क्या समझते हैं?
4. द्विपद तथा सामान्य वितरण की विशेषताएं बताओ।
5. द्विपद वितरण क्या है? किन परिस्थितियों में द्विपद वितरण का प्रयोग होता है चर्चा करें।
6. प्वॉयसन वितरण क्या है? प्वॉयसन वितरण की विशेषताएं बताओ।
7. द्विपद तथा प्वॉयसन वितरण की प्रमुख विशेषताओं पर चर्चा करें।
8. प्वॉयसन वितरण क्या है? यह कहाँ पर प्रयुक्त किया जा सकता है उदाहरण दो।

स्वपरख प्रश्नों के उत्तर

1. 11/32
2. अ. 0.0988 ब. 0.4074 स. 0.1975

9.18 सन्दर्भ पुस्तकें

1. Roy Ramendu, ***‘Principles of Statistics’***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***“Business Statistics”*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***“Advanced Statistics”,*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***“Basic Statistics”*** World Press Limited, Calcutta.
5***. “Fundamentals of Business Statistics”*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **“Quantitative Methods in Management”.**

इकाई 10 घातांकी, बीटा, और सामान्य वितरण
(Exponential, Beta, and Normal Distributions)

इकाई की रूपरेखा

उद्देश्य (Objective)

इस इकाई के अध्ययन के पश्चात आप इस योग्य हो सकेंगे कि -

- ✓ घातांकी वितरण की व्याख्या कर सकें।
- ✓ बीटा वितरण की व्याख्या कर सकें।
- ✓ सामान्य प्रायिकता वितरण एवं उनके प्रयोग का वर्णन कर सकें।

10.1 प्रस्तावना (Introduction)

प्रायिकता सिद्धांत और सांख्यिकी में, घातांकी, बीटा, और सामान्य वितरण सतत वितरण परिवार का हिस्सा है। सामान्य वितरण बहुत ही महत्वपूर्ण तथा बहुतायत उपयोगी सतत प्रायिकता वितरण है। इसका प्रयोग प्रमुखतया सतत दैव चरों जैसे लंबाई, भार/ वजन और छात्रों के एक समूह के बुद्धिमत्ता के व्यवहार का अध्ययन करने में होता है।

10.2 घातांकी वितरण (Exponential Distribution)

घातांकी वितरण दो घटनाओं के मध्य के समय को प्वायन प्रक्रिया के रूप में प्रदर्शित करता है अर्थात् एक ऐसी प्रक्रिया जिसमें घटनाएं स्वतंत्र तथा सतत रूप से एक निश्चित औसत दर से घटित होती है। यह गुणोत्तर वितरण (ळमवउमजतपब क्पेजतपइनजपवद) का सतत रूप है।

10.2.1 घातांकी वितरण की विशेषताएं (Characteristics of Exponential Distribution)

घातांकी वितरण की विशेषताएं निम्नलिखित हैंः

(अ) प्रायिकता घनत्व फलन [Probability Density Function (pdf)]

घातांकी वितरण का प्रायिकता घनत्व फलन निम्न हैः

$$= \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

वैकल्पिक रूप में इसे Heaviside step function, H(x) की सहयता से इस प्रकार पारिभाषित करते हैं

$$f(x, \lambda) = \lambda e^{-\lambda x} H(x)$$

यहां पर $\lambda > 0$ वितरण का प्राचाल है और प्रायः इसे प्राचाल दर कहते हैं। यह वितरण $[0,\infty)$ अंतराल पर आधारित होता है। यदि एक दैव चर X इस वितरण को लागू करता है तो हम कहते हैं $X \approx Exp\ (\lambda)$

(ब) संचयी वितरण फलन (Cumulative Distribution Function)

संचयी वितरण फलन को इस प्रकार लिख सकते हैं

$$= \begin{cases} 1- e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

वैकल्पिक रूप में इसे Heaviside step function H(x) की सहयता से इस प्रकार पारिभाषित करते हैं

$$f(x. \lambda) = (1- e^{-\lambda})\ H(x)$$

(स) वैकल्पिक प्रचालीकरण

वैकल्पिक प्राचालीकरण का अर्थ है घातांकी वितरण के प्रायिकता घनत्व फलन को इस प्रकार पारिभाषित करना

$$= \begin{cases} 1/\beta\ .e^{-x/\beta}, & x \geq 0 \end{cases}$$

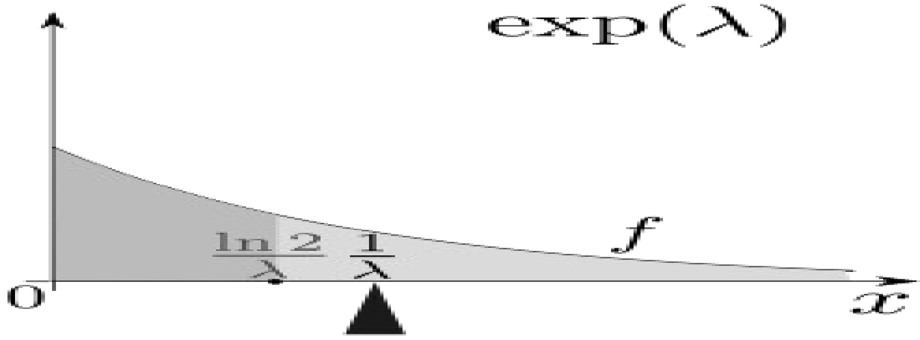
$0, x < 0$

जहां स्केल $\beta>0$ प्राचाल है और यह प्राचाल दर λ का reciprocal है। इस विशिष्टीकरण में, β उत्तर जीविता (survival) प्राचाल कहलाता है जब दैव चर X किसी बायोलाजिकल या मेकैनिकल प्रक्रिया में जीवित रहने कि समयावधि को प्रदर्शित करे और $X \approx Exp(\beta)$ तो $E(X) = \beta$। अत: प्रक्रिया के जीवित रहने की प्रत्याशित अवधि β इकाई समय है। जब दो घटनाओं के मध्य समय की चर्चा करते हैं तो प्राचाल दर λ, प्राचालीकरण का उपयोग करते हैं और इसका माध्य $\beta = \lambda^{-1}$ होता है।

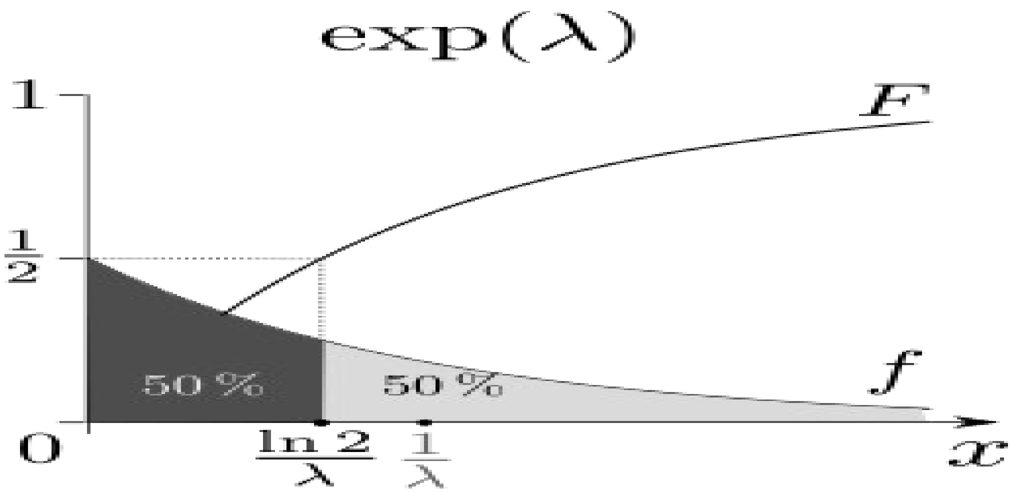
10.2.2 घातांकी वितरण के गुण (Properties of Exponential Distributions)

घातांकी वितरण के गुण निम्न हैंः

(1) माध्य, प्रसरण, संवेग तथा माध्यिका



इसका माध्य प्रायिकता भार केन्द्र (Probability mass centre) है जो कि संवेग है (first moment preimage) (पूर्वचित्र)



$F^{-1}(1/2)$ माध्यिका है।

माध्य अथवा घातांकी वितरित दैव चर का प्रत्याशित मान जिसका प्रचाल दर λ है, इस प्रकार है

$$E(X) = 1/\lambda$$

ऊपर वर्णित उदाहरणों से यह स्पष्ट है कि यदि आप औसतन 2 फोन काल प्रति घंटा करते हैं तो आपको प्रत्येक काल के लिए औसतन आधा घंटा इंतजार करना पड़ता है।

X का प्रसरण, $Var(X) = 1/\lambda^2$

X का संवेग जब $n = 1,2,3,\ldots..$

$m(X), = \ln 2/(\lambda) < E(X)$

जहां पर ln प्राकृतिक लागरिथम (natural logarithm) है। अतः माध्य तथा माध्यिका का वास्तविक अंतर, माध्य- माध्यिका inequality के अंतर्गत

$$E[X] - m[X] = (1- \ln 2)/\lambda < 1/\lambda$$

(ब) स्मृति विभ्रम (Memory Lessness)

घातांकी वितरण की एक महत्वपूर्ण विशेषता यह है कि स्मृति विभ्रम। इसका अर्थ यह है कि यदि एक दैव चर (Random Variable) T घातांकी वितरित है तो सशर्त प्रायिकता

$$P[T > s+t \mid T > s] = P(T > t) \qquad s, t \geq 0$$

इसका अर्थ है कि इंतजार करने की सशर्त प्रायिकता उदाहरणार्थ प्रथम आगमन से पहले 10 सेकेंड से अधिक की प्रतीक्षा करने की सशर्त प्रायिकता जब दिया गया है कि पहला आगमन 30 सेकेंड तक नहीं हुआ है, शुरुआती प्रायिकता कि प्रथम आगमन के लिए 10 सेकेंड से अधिक की प्रतीक्षा करने के बराबर है। अतः यदि आप 30 सेकेंड प्रतीक्षा करने के बाद भी प्रथम आगमन नहीं होता $(T > 30)$ तो पहले आगमन के लिए 10 सेकेंड और इंतजार करने की प्रायिकता $(T > 30+10)$ शुरुआती प्रायिकता जो कि प्रथम आगमन के लिए 10 सेकेंड से अधिक $(T>10)$ का इंतजार करना के बराबर होगी।

वास्तव में, $P(T>40|T>30) = P(T>10)$ का अर्थ यह नहीं है कि घटनाएं $T>40$ तथा $T>30$ स्वतंत्र हैं। सारांशः स्मृति विभ्रम प्रथम आगमन होने तक इंतजार करने का समय का प्रायिकता वितरणका स्मृति विभ्रम का अर्थ है

(सही) $P(T>40|T>30) = P(T>10)$

इसका अर्थ यह नहीं है कि

(गलत) $P(T>40|T>30) = P(T>40)$

वे स्वतंत्र हो सकते हैं। ये घटनाएं स्वतंत्र नहीं हैं। घातांकी वितरण तथा गुणोत्तर वितरण में ही स्मृति विभ्रम विशेषता होती है।

घातांकी वितरण ही सिर्फ एक ऐसा सतत प्रायिकता वितरण है जिसका असफलता दर (failure rate) स्थिर (constant) है।

(स) चतुर्थांश

घातांकी वितरण का चतुर्थांश फलन (quartile function) विपरीत संचयी वितरण फलन (inverse cumulative distribution function)

$$F^{-1}P(p, \lambda)= -\ln(1-p)/\lambda, \qquad 0 \leq p < 1$$

अतः चतुर्थांश निम्न हैंः

प्रथम चतुर्थांश = $\text{Ln}(4/3)/\lambda$

माध्यिका= $\ln(2)/\lambda$

तृतीय चतुर्थांश= $\ln(4)/\lambda$

10.3 बीटा वितरण (Beta Distribution)

बीटा वितरण एक सतत प्रायिकता वितरण है जो कि अंतराल (0,1) पर पारिभाषित है और जिसके दो धनात्मक आकार प्राचाल (shape parameter) A तथा B हैं। इसका प्रयोग समानुपातों के सांख्यिकीय माडलिंग (statistical modelling) में होता है जहां समानुपातों का मान या यॉंग के बराबर नहीं होता। एक दैव चर जो गासियन (Gaussian) वितरित हो तो उसके तथा एक और स्वतंत्र गासियन चर, जिसका स्केल प्राचाल पहले चर के बराबर तथा संभवतः अलग आकार प्राचाल हो, के योग से भाग देने पर प्राप्त अनुपात का वितरण का एक सैद्धांतिक उदाहरण है।

10.3.1 विशेषताएं (Characteristics)

बीटा वितरण की विशेषताएं निम्नलिखित हैंः

(अ) प्रायिकता घनत्व फलन

बीटा वितरण का प्रायिकता घनत्व फलन निम्न हैः

$$f(x,\alpha,\beta) = x^{\alpha-1} (1-x)^{\beta-1} / ^{\alpha-1}(1-u)^{\beta-1} du$$
$$= [\prod(\alpha+\beta)/ \prod(\alpha) \prod(\beta)] . x^{\alpha-1} (1-x)^{\beta-1}$$
$$= 1/ B(\alpha,\beta) . x^{\alpha-1} (1-x)^{\beta-1}$$

जहां पर $\prod(z)$गामा फलन है। बीटा फलन B पूरी प्रायिकता जुड़कर करं होने के लिए सामान्यीकरणधारित है। एक दैव चर जो कि बीटा वितरित है और जिसका प्रचाल α और β है को $X \backsim Be(\alpha,\beta)$ द्वारा प्रदर्शित करते हैं।

(ब) संचयी वितरण फलन

संचयी वितरण फलन को निम्न प्रकार से दर्शाते हैं

$$f(x,\alpha,\beta) = B_x(\alpha,\beta)/B(\alpha,\beta) = I_x(\alpha,\beta)$$

जहां पर ठ X (α,β) अधूरा बीटा फलन(incomplete Beta function) है और $I_x (\alpha,\beta)$ निरंतर बीटा फलन(regularized incomplete Beta function) है।

10.3.2 गुण (Merits)

बीटा वितरण के गुण निम्नलिखित हैंः

बीटा वितरित दैव चर X जिसका प्राचाल $\alpha>1$ तथा $\beta>1$ का बहुलक

बहुलक $=(\alpha.1)/(\alpha+\beta-2)$

प्रत्याशित मान, माध्य तथा प्रसरण(द्वितीय केन्द्रीय संवेग)

$$\mu = E(X) = \alpha/(\alpha+\beta)$$
$$\text{var}(X) = E(X-\mu)^2 = \alpha\beta/(\alpha+\beta)^2 (\alpha+\beta+1))$$

सभी परिस्थितियों में, $Var(X) < 1/4$

विषमताः

$$E(X-\mu)^3/ [E(X-\mu)^2)^{3/2} = (2(\beta-\alpha) \surd \alpha+\beta+1) / (\alpha+\beta-2)\surd\alpha\beta$$

$$E(X-\mu)^4/ [E(X-\mu)^2)^2 - 3 = 6[(\alpha^3- \alpha^2) (2\beta-1)+ \beta^2 (\beta+1) - 2\alpha\beta (\beta+2)] / \beta\alpha(\alpha+\beta+2)(\alpha+\beta+3)$$

$$= 6[(\alpha- \beta)^2 (\alpha+\beta+1) - \alpha\beta (\alpha+\beta-2) / \alpha\beta (\alpha+\beta+2) (\alpha+\beta+3)$$

या

समान्यतया, k^{th} कच्चा संवेग (Raw moment)

$$E(X^k) = B(\alpha+k, \beta) / (\alpha+\beta) = (\alpha)_k /(\alpha+\beta)_k$$

जहां पर $(x)_k$ पोसोम्बर चिन्ह (Pochhammer symbol) है जो कि rising factorial को प्रदर्शित करता है। इसे recursive form में इस प्रकार लिख सकते हैं

$$E(X^k) = (\alpha+k-1) / (\alpha+\beta+k+1) E(x^{k-1})$$

हम यह भी दिखा सकते हैं

$$E(\log X) = \psi(\alpha) - \psi(\alpha+\beta)$$

और

$$E(X^{-1})= (\alpha+\beta-1)/(\alpha-1)$$

(स) सूचना की मात्रा (Quantities of information)

दो बीटा वितरित दैव चर $X \backsim Be(\alpha,\beta)$ और $Y \backsim Be(\alpha',\beta')$ हैं तो X का differential entropy है

h(X) =

डिगामा फलन (Digamma function) कहां है?

क्रास इन्ट्रापी (cross entropy) है

$$H(X,Y) = \ln B(\alpha',\beta') -(\alpha'-1)\,\psi(\alpha)-(\beta'-1)\,\psi(\beta)+ (\alpha'+\beta'-2)\,\psi(\alpha+\beta)$$

10.4 सामान्य प्रायिकता वितरण (Normal Probability Distribution)

सामान्य प्रायिकता वितरण की खोज अंग्रेजी गणितज्ञ अब्राहम डिमोरे ने 1733 में की थी। परन्तु बाद में लाप्लास और कार्ल गास ने इसे पुनः खोजा और उपयोग किया। कार्ल गास के नाम पर सामान्य वितरण को गासियन वितरण भी कहते हैं।

निम्न निश्चित परिस्थितियों में सामान्य वितरण को द्विपद वितरण का सन्निकट समझा जा सकता हैः

(i) परखों (trial) की संख्या (n), बहुत अधिक हो अर्थात् n-- ∞

(ii) p तथा q में से कोई भी बहुत छोटा न हो।

सामान्य वितरण का प्रायिकता घनत्व फलन इस प्रकार पारिभाषित करते हैं

$P\,(X = x) = \;\; 1/\sigma\surd 2\pi\ .\ e^{-1/2(x-\bar{x}/\sigma)2} \qquad -\infty < X < +\infty$

जहां पर $\bar{X}$ = माध्य, σ = प्रमाप विचलन, e= नेचुरल लागिरथम का बेस =2.7183, π = 3.1415

सामान्य वितरण का Standard Normal Variate (SNV) फार्म निम्न हैः

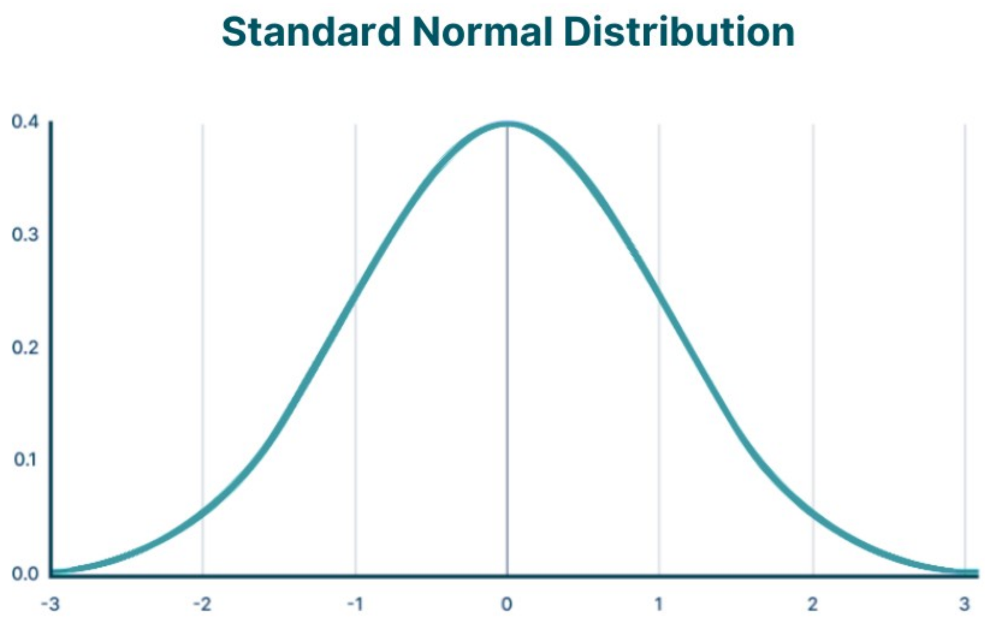
$P\,(Z) = \;\; 1\surd 2\pi\ .\ e^{-1/2(Z)2} \qquad -\infty < X < +\infty$ where $Z = (X - \bar{X})/\,\sigma$

जहां पर $Z = (X - \bar{X}\,)/\,\sigma$

Z का माध्य 0 तथा प्रमाप विचलन σ =1 है अर्थात् सामान्य वितरण है जिसका माध्य 0 तथा प्रमाप विचलन σ इकाई है।

10.4.1 सामान्य वितरण वक्र (Normal Distribution Curve)

सामान्य वितरण के वक्र को सामान्य वितरण वक्र कहते है। सामान्य वक्र सामान्य वितरण का चित्रीय प्रदर्शन है। निम्न चित्र सामान्य वक्र दर्शाता हैः



सामान्य वक्र का आकार माध्य ($\bar{X}$) के मान तथा प्रमाप विचलन (σ) पर निर्भर करता है। माध्य तथा प्रमाप विचलन के भिन्न-2 मानों के लिए सामान्य वक्र का आकार भिन्न होगा।

10.4.2 सामान्य वितरण की मान्यताएं (Assumptions of Normal Distribution)

सामान्य वितरण निम्न मान्यताओं पर आधारित होता है

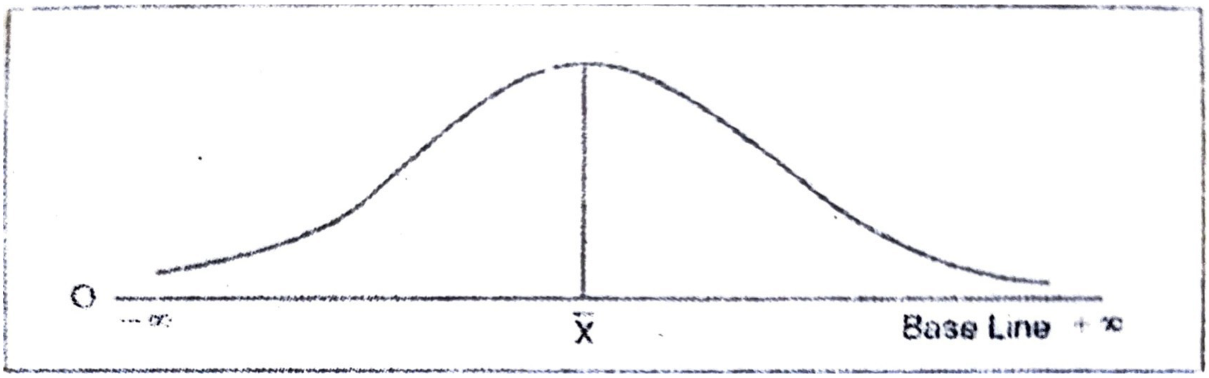
1. **स्वतंत्र कारण-** घटनाओं को प्रभावित करने वाले कारण एक दूसरे से स्वतंत्र होने चाहिए।
2. **विषमता की परिस्थिति -** कारणीय ताकतों का क्रियान्वयन इस प्रकार है कि माध्य से विचलन किसी भी दिशा में संख्या तथा आकार में बराबर हो।
3. **विभिन्न कारण-** कारणीयताकत बहुतायत तथा करीब-करीब बराबर वजन या महत्व हो।

10.4.3 सामान्य वितरण/सामान्य वक्र की चारित्रिक विशेषताएं (Characteristics of Normal Distribution/Normal Curve)

1. **पूर्णतया सममितीय तथा घंटाकार आकार-** सामान्य वक्र माध्य के प्रति पूर्णतया सममित तथा घंटाकार आकार का है। अर्थात यदि हम वक्र को इसके लम्बवत् अक्ष के तहत मोड़ें तो दोनों आधे भाग एक दूसरे पर मेल खाएंगे।
2. **एक बहुलक वितरण-** इसमें सिर्फ एक ही बहुलक होता है। अर्थात् एक बहुलक वितरण है।
3. **माध्य, माध्यिका तथा बहुलक की समानता-** सामान्य वितरण में माध्य, माध्यिका तथा बहुलक बराबर/समान होते हैं अर्थात्

$$\bar{X} = M = Z$$

4. **आधार रेखा (Base Line) से स्पर्शोन्मुखी-** सामान्य वक्र आधार रेखा से दोनों तरफ स्पर्शोन्मुखी है अर्थात् इसकी प्रवृत्ति बेस लाइन से छू जाने की है किन्तु यह कभी छूती नहीं है। यह निम्न चित्र से पूर्णतया साफ है।



5. **प्रसार/विस्तार (Range)-** सामान्य वितरण का प्रसार दोनों ही तरफ अनन्त तक है अर्थात् $-\infty$ से $+\infty$ तक।

6. **पूर्ण क्षेत्रफल**- सामान्य वक्र का क्षेत्रफल 1 है।

7. **कोटि (ordinate)** - सामान्य वक्र का कोटि (ordinate) माध्य के सापेक्ष उच्चतम है।

8. **माध्य कोटि-** माध्य कोटि पूरे क्षेत्रफल को दो बराबर भागों में विभाजित करता है। अर्थात् 50%बायीं तरफ तथा 50% दायीं तरफ ।

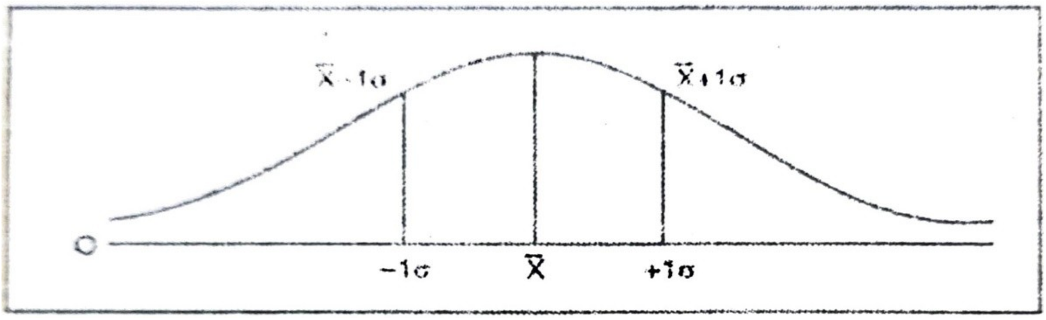
9. **चतुर्थांश का समान दूरी पर होना-** सामान्य वितरण में चतुर्थांश फ 1 तथाफ 3 माध्यिका से समान दूरी पर है। अर्थात $Q_3 - M = M - Q_1$

10. चतुर्थांश विचलन (Q.D.) - सामान्य वितरण में, चतुर्थांश विचलन प्रमाप विचलन का 2/3 है अर्थात्

$$Q.D. = 2/3\ S.D.$$

11. मध्य विचलन(M.D.)- सामान्य वितरण में, माध्य विचलन प्रमाप विचलन का 4/5 है अर्थात M.D.= 4/5 S.D.

12. संक्रमण का बिन्दु (Point of inflexion)- सामान्य वितरण में दो Point of inflexion है (अर्थात् वह बिन्दु जहाँ पर वक्र अपनी वक्रता बदलता है) और वह बिन्दु $\bar{X}$-1σ तथा $\bar{X}$+1σ अन्य शब्दों में inflexion बिन्दु $\bar{X} \pm 1\ \sigma$ पर होता है। यह निम्न चित्र से साफ है।



13. सतत प्रायिकता वितरण- सामान्य वितरण सतत चरों का वितरण है। अतः इसे सतत प्रायिकता वितरण कहते हैं।

14. धारित (constant)- सामान्य वितरण के धारित निम्न चिह्नों से प्रदर्शित करते हैं। माध्य = $\bar{X}$ या μ या m

विषमता गुणांक संवेग=$\sqrt{\beta_1}$ = 0

प्रमाप विचलन = σ

प्रथुशीर्षत्व गुणांक संवेग = $\sqrt{\beta_2}$ =3

प्रसरण = σ

15. प्रमुख प्राचाल- सामान्य वितरण के दो प्राचल हैं माध्य ($\bar{X}$)तथा प्रमाप विचलन (σ) इन दो प्राचालों की सहायता से हम पूरा वितरण ज्ञात कर सकते हैं।

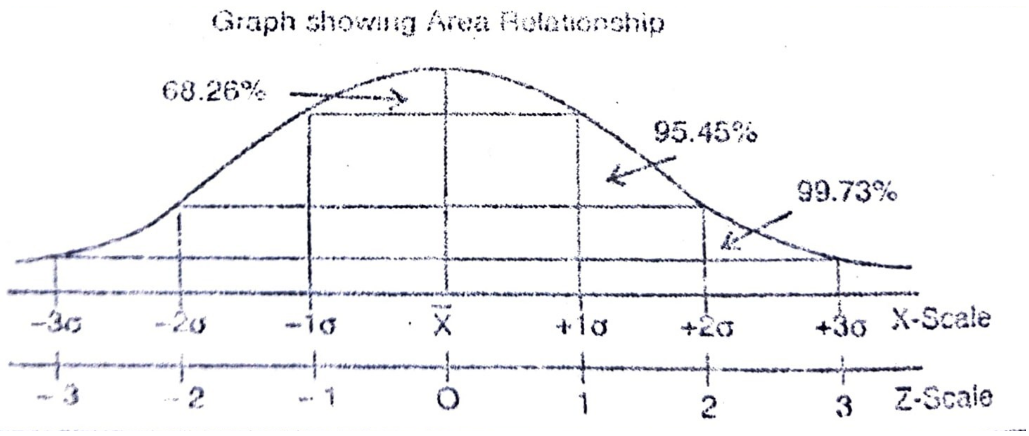
16. क्षेत्रफल विशेषता- सामान्य वक्र की सबसे प्रमुख विशेषताओं में से एक विशेषता इसकी क्षेत्रफल विशेषता है। वक्र के नीचे का पूरा क्षेत्रफल 1 होता है। यह पाया जाता है किः

(अ) $\bar{X}$-1σ तथा $\bar{X}$+1σ के मध्य सामान्य वक्र का क्षेत्रफल 0.6826 है अर्थात माध्य ± 1σ वक्र के कुल क्षेत्रफल का 68.26% क्षेत्रफल धारण करता है।

(ब) $\bar{X}$-2σ तथा $\bar{X}$+2σ के मध्य सामान्य वक्र का क्षेत्रफल 0.95450.9973 अर्थात माध्य ± 2σ वक्र के कुल क्षेत्रफल का 95.45% भाग रहता है।

(स) $\bar{X}$-3σ तथा $\bar{X}$+3σ के मध्य सामान्य वक्र का क्षेत्रफल 0.9973 है अर्थात माध्य ± 3σ वक्र के कुल क्षेत्रफल का 99.73% भाग रहता है।

निम्नलिखित चित्र क्षेत्रफल विशेषता को दर्शाता है-



10.4.4 सामान्य वितरण का महत्व (Importance of Normal Distribution)-

संख्यिकीय समीक्षा में सामान्य वितरण का बहुत महत्व है। यह नवीन सांख्यिकी का आधार है। निम्न सामान्य वितरण के उपयोग तथा महत्व को चिह्नांकित करते हैंः

1. **प्राकृतिक घटनाओं का अध्ययन-** सभी प्राकृतिक घटनाएं सामान्य वितरण की चारित्रिक विशेषताएंधारण करते हैं जैसे एक पेड़ की पत्तियों की लम्बाई,जन्म दर तथा मृत्यु दर इत्यादि। सामान्य वितरण प्राकृतिक घटनाओं के अध्ययन में बहुतायत उपयोग होता है।
2. **निदर्शन सिद्धान्त का आधार-** सामान्य वितरण का निदर्शन सिद्धान्त में भी बहुत महत्वपूर्ण है। सामान्य वितरण की सहायता से हम यह परीक्षण कर सकते हैं कि ब्रह्मांड न्दपअमतेम से चुने गए निदर्श ब्रह्मांड को पूर्णतः प्रदर्शित करते हैं या नहीं।
3. **सांख्यिकीय गुणवत्ता नियं=ण-** यह टालरेंस या स्पेस्फिक जिसके मध्य उत्पाद की गुणवत्ता होती है, जानने में मदद करता है। उत्पाद की गुणवता ॅचमबपपिबंजपवद सपउपजे के मध्य ही स्वीकार्य है।
4. **वृहद प्रतिदर्श परीक्षणः-** सामान्य वितरण का प्रयोग वृहद प्रतिदर्श परीक्षण में भी होता है। वृहद प्रतिदर्श परीक्षण सामान्य वितरण की विशेषताओं पर आधारित है।
5. **द्विपद तथा प्वायसनवितरण का सन्निकटीकरणः**

सामान्य वितरण कई सैद्धांतिक वितरणों जैसे द्विपद, प्वायसन इत्यादि का अच्छा सन्निकटीकरण है। जैसे अवलोकनों की संख्या बढ़ती है प्वायसन, द्विपद इत्यादि से संबंधित समस्याओं को हल करने में सामान्य वितरण का महत्व बढ़ता है।

6. प्रो योडेन ने सामान्य वक्र के आकार के आधार पर सामान्य वितरण का महत्व बताया है जो कि नीचे दर्शाया गया हैः

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY. IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE, AGRICULTURAL AND ENGINEERING
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT.

10.4.5 द्विपद (BD), प्वायसन (PD) तथा सामान्य वितरण (ND) में संबंध-

द्विपद प्वायसन तथा सामान्य वितरण एक दूसरे से संबंधित हैं। संबंध नीचे दर्शाया गया हैः

अ. द्विपद तथा सामान्य वितरण के मध्य संबंध

द्विपद वितरण निम्न परिस्थितियों में सामान्य वितरण की ओर उन्मुख होता हैः

1. n अनंत की ओर उन्मुख हो अर्थात्
2. p तथा q में से कोई भी बहुत छोटा न हो।

ब. प्वायसन और सामान्य वितरण के मध्य संबंध

जब प्वायसन वितरण का प्राचाल ‘m’ बहुत बड़ा हो जाता है अर्थात् m -- ∞ तो प्वायसन वितरण सामान्य वितरण की ओर उन्मुख हो जाता है।

10.4.6 सामान्य तथा द्विपद वितरण में अंतर (Difference Between Normal and Binomial Distribution)

सामान्य तथा द्विपद वितरण में निम्नलिखित अंतर होते हैंः-

1. **प्रकृति-** द्विपद वितरण असतत प्रायिकता है जबकि सामान्य वितरण सतत प्रायिक वितरण है।
2. **प्रायिकता फलन-** द्विपद वितरण का प्रायिकता फलन निम्न प्रकार से है-

$$P(X = x) = {}^nC_x\, q^{n-x}\, p^x$$

सामान्य वितरण का प्रायिकता फलन

$$P\ (X = x) = \ 1/\sigma\sqrt{2\pi}\ .\ e^{-1/2(x-\bar{x}/\sigma)2}$$

3. **N का मान-** द्विपद वितरण में परखों की संख्या निश्चित होती है जबकि सामान्य वितरण में अनंत के निकट होता है अर्थात n-- ∞
4. **प्राचल-** द्विपद वितरण के दो प्राचल हैं, n तथा p जबकि सामान्य वितरण के भी दो प्राचल हैं $\bar{X}$ तथाσ
5. **आकार-** सामान्य वितरण सममित तथा असममित हो सकता है। वह तथा के मान पर निर्भर करता है जबकि दूसरी तरफ सामान्य वितरण हमेंशा सममित होता है।

10.5 सामान्य वक्र के नीचे के क्षेत्रफल का मापन (Measurement of the Area Under the Normal Curve)

सामान्य वक्र के नीचे के क्षेत्रफल को मापने के लिए निम्न कदम उठाए जाते हैं

1. सर्वप्रथम दिए गए सामान्य चर को standard normal variate में बदलते हैं। बदलीकरण का सूत्र निम्न हैः

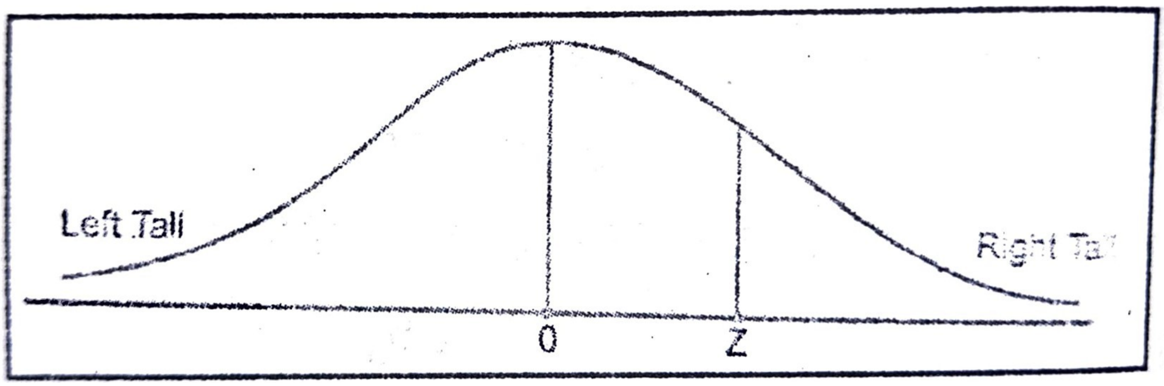
$$Z =$$

उदाहरण के लिए, यदि $\bar{X}$=30, σ = 5 तथा X 35 हो तो 35 के सापेक्ष standard normal variate होगाः $Z = (35.30)/5 = 1$

अतः 35 का बदलीकरण 1 होगा

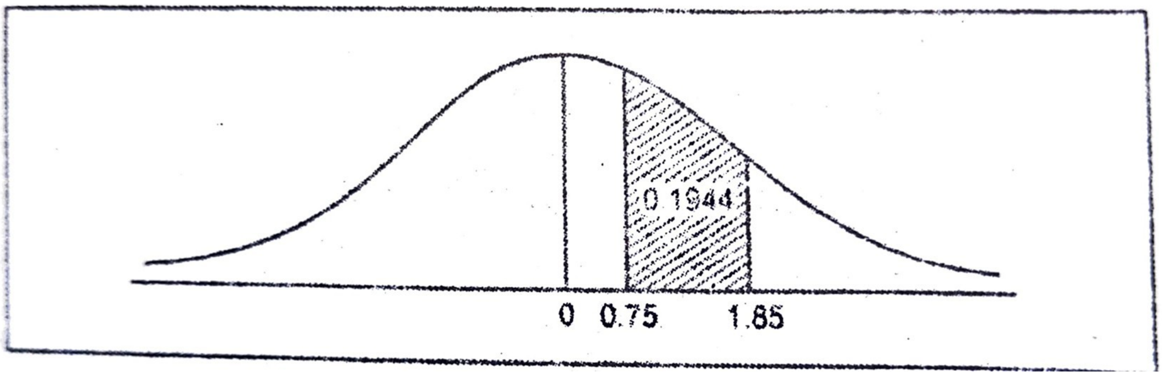
2) तब किसी निश्चित मान के लिए क्षेत्रफल अध्याय के अंत में दिए गए क्षेत्रफल तालिका से प्राप्त किया जा सकता है।

अध्याय के अंत में दी गयी तालिका 0 से Z के मध्य का क्षेत्रफल दर्शाती है जो कि निम्न चित्र से दर्शाया गया हैः



उदाहरण 10.1 Z = 0.75 तथा Z = 1.85 के मध्य सामान्य वक्र के नीचे का क्षेत्रफल ज्ञात करो।

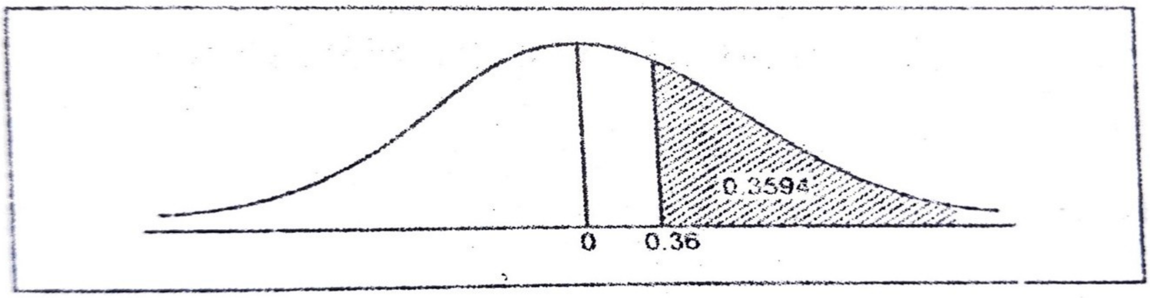
हलः



वांछित क्षेत्रफल = (Z= 0 से Z= 1.85 के मध्य क्षेत्रफल). (Z= 0 से Z=0.75 के मध्य क्षेत्रफल)
= 0.4678 - 0.2734 = 0.1944

उदाहरण 10.2 : Z = +0.36 के दायें का क्षेत्रफल ज्ञात करो।

हलः



वांछित क्षेत्रफल
=(Z = 0 के मध्य क्षेत्रफल) - (Z= 0 से Z= 0.36 के मध्य क्षेत्रफल)
= 0.5000 . 0.1406
= 0.3594

उदाहरण 10.3 : Z = -1.25 के दायीं ओर या Z =-1.25 से अधिक का क्षेत्रफल ज्ञात करो

हल

वांछित क्षेत्रफल = (Z=-1.25 तथा Z= 0 के मध्य क्षेत्रफल के दायीं ओर का क्षेत्रफल)
= 0.3944 + 0.5000
= 0.8944

10.6 सामान्य वितरण का उपयोग (Use of Normal Distribution)

अब आप सामान्य वितरण के उपयोंगों का अध्ययन करेंगे।

10.10.1 सामान्य चर के Xbar तथा σ का मान ज्ञात होने/ दिए जाने पर क्षेत्रफल ज्ञात करना

सामान्य वितरण के नीचे का क्षेत्रफल ज्ञात करने के लिए सर्वप्रथम सामान्य चर को Z चर में बदलना पड़ता है। उदाहरण के लिए यदि $\bar{X}$=30, σ=5 तथा X =35 हो तो Standard normal variate इस प्रकार बदला जाएगाः

Z= (35.30)/5 =1

जहाँ Z= (X-$\bar{X}$)/σ

अतः X=35 के लिए Standard normal variate =1 है।

Z बदलीकरण के पश्चात सामान्य वक्र के नीचे के क्षेत्रफल वाले तालिका लेते हैं।

उदाहरण 10.4: 1000 परीक्षार्थियों पर किए गए अभिक्षमता परीक्षण करने पर पता चलता है कि औसत स्कोर 42 तथा स्कोर का प्रमाप विचलन 24 है। स्कोर के लिए सामान्य वितरण मानते हुए ज्ञात कीजिए 1) 60 से अधिक स्कोर पाने वाले परीक्षार्थियों की संख्या तथा 2) 30 से 66 के मध्य स्कोर करने वाले परीक्षार्थियों की संख्या

हलः दिया है $\bar{X}$=42, σ =24, N=1000

1) 60 से अधिक

Z= 0 के सापेक्ष Standard normal variate (SNV)

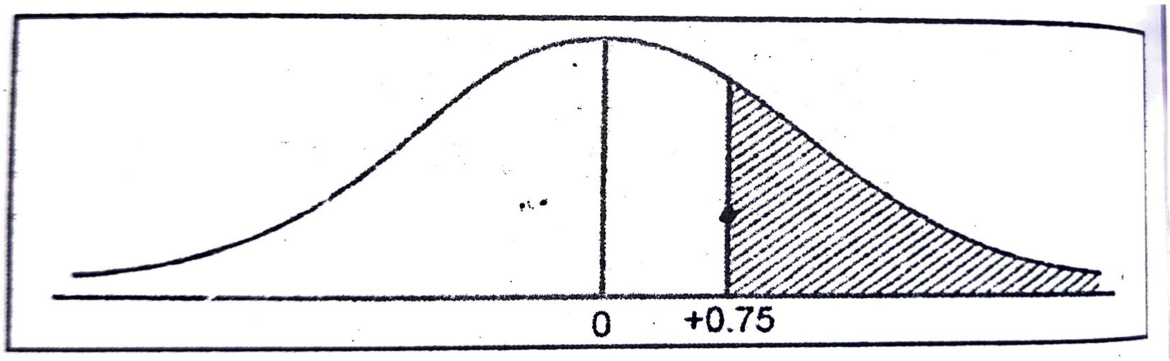
= $(X-\bar{X})/\sigma$

=(60.42)/24

=18/24

=3/4

=0.75



वांछित प्रायिकता =

(Z= 0 के दायीं ओर क्षेत्रफल) - (Z= 0 से Z= 0.75 के मध्य क्षेत्रफल)

= 0.5000. 0.2734 = 0.2266

अर्थात् 60 से अधिक स्कोर करने वाले परीक्षार्थियों की संख्या =1000 x 0.2266 =222.6 या 227

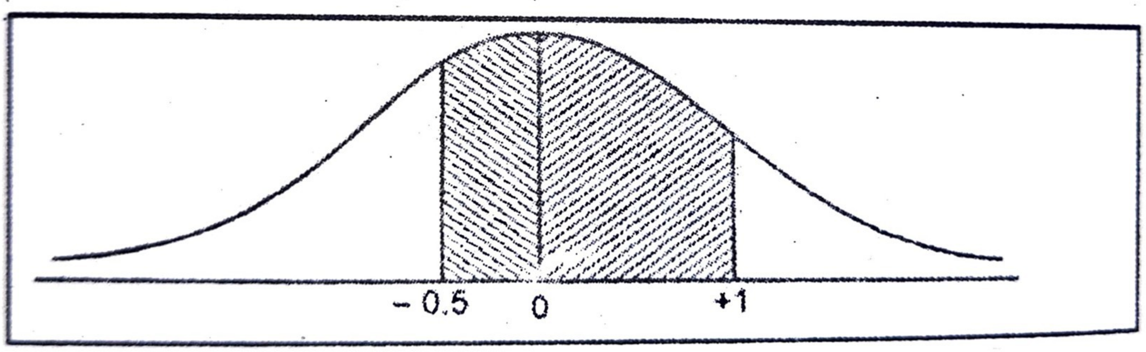
2) 30 तथा 66 के मध्य

Z_1= 30 के सापेक्ष Standard normal variate (SNV)

= $(X-\bar{X})/\sigma$ = (30.42)/24 = -12/24 = -0.5

Z_2 = 66 के सापेक्ष Standard normal variate (SNV)

= $(X-\bar{X})/\sigma$ = (66.42)/24 = 24/24 = +1



वांछित क्षेत्रफल

=(Z=-0.5 से Z= 0 के मध्य क्षेत्रफल)+(Z= 0 से Z=1 के मध्य क्षेत्रफल)

=0.1915 + 0.3413 = 0.5328

अतः 30 तथा 66 के मध्य स्कोर करने वाले परीक्षार्थियों की संख्या=1000 x 0.5328 = 532.8 या 533

उदाहरण 10.5 : सैनिकों की औसत लम्बाई 68.22 इंच तथा प्रसरण 10.8 है। 1000 सैनिकों वाली टुकड़ी में 6 फीट से अधिक लम्बे सैनिकों की संख्या बताओ।

हलः दिया है $\bar{X}$= 68.22, σ^2= 10.8 या $\sigma = \sqrt{10.8}$ = 3.28

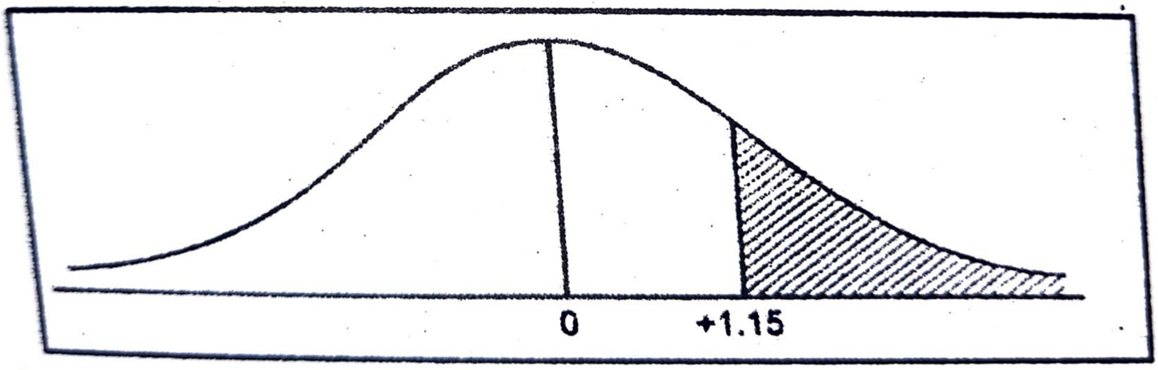
6 फीट से अधिक (अर्थात्72 इंच से अधिक)

X=72 के लिए Z = (X –$\bar{X}$) / σ = (72.68.22)/3.28 = 3.78/3.28 = 1.15

वांछित क्षेत्रफल=

(Z=0 के दायीं ओर क्षेत्रफल) - (Z= 0 तथा Z= 1.15 के मध्य क्षेत्रफल)

= 0.5000. 0.3749 = 0.12251



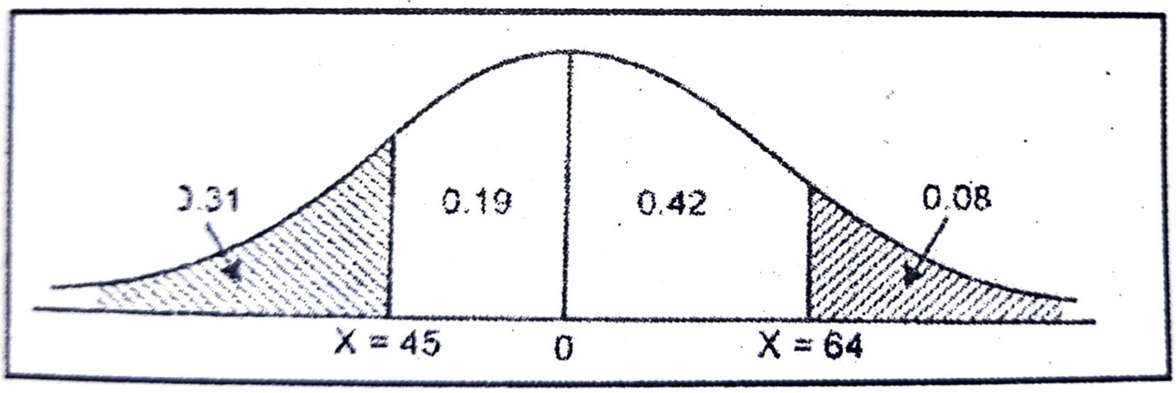
अतः 6 फीट अधिक लम्बाई वाले सैनिकों की प्रत्याशित संख्या= 1000x 0.1251 = 125 ≈125

10.6.2 सामान्य वक्र के नीचे के क्षेत्रफल ज्ञात होने पर माध्य तथा प्रमाप विचलन ज्ञात करना

जब सामान्य वक्र के नीचे का क्षेत्रफल ज्ञात हो तो माध्य तथा प्रमाप विचलन का मान ज्ञात कर सकते हैं।

उदाहरण 10.6 : एक साममन्य वितरण में 31% मद के नीचे तथा 8% मद 64 से ऊपर हैं। $\bar{X}$ तथा σ का मान ज्ञात करो।

हलः



Z= (X –$\bar{X}$) / σ

तालिका से Z=0.5 -0.3 के सापेक्ष Z का मान

= 0.19 क्षेत्रफल

= -0.5

∴ -0.5 = (45- $\bar{X}$)/σ

या, -0.5σ = 45 – $\bar{X}$

$-0.5\sigma = 45$ (i)

$Z=0.5-0.08$ के सापेक्ष Z का मान

$= 0.42$ क्षेत्रफल $= +1.41$ (तालिका स)

$\therefore 1.41 = (64- \bar{X})/ \sigma$

या $1.41\ \sigma = 64.$

या $\bar{X}+ 1.41\sigma = 64$ (तालिका से)

दोनों समीकरणों को हल करने पर

$\bar{X} - 0.5\ \sigma = 45$

$\bar{X} +1.41\ \sigma = 64$

. . .

$-1.91\ \sigma = -19$

$\sigma = 10$

σ का मान समीकरण 1) पर रखने पर

$\bar{X} - 0.5\ (10) = 45$

$\bar{X} - 5= 45$

या, $\bar{X}=50$

$\therefore \bar{X}= 50, \sigma =10$

उदाहरण 10.7 : एक वितरण पूर्णतया 7% मद 35 से कम हैं तथा 89% मद 63 से कम/नीचे हैं। वितरण का माध्य तथा प्रमाप विचलन ज्ञात करो।

हलः 0.43 क्षेत्रफल के सापेक्ष Z का मान =1.48

$\therefore -1.48 = (35 - \bar{X})/\sigma$

या $-1.48\ \sigma = 35- \bar{X}$

या $\bar{X}-1.48\ \sigma = 35$ (i)

0.39 क्षेत्रफल के सापेक्ष Z का मान = +1.23

$1.23 = (63- \bar{X})/ \sigma$

$1.23\ \sigma = 63 - \bar{X}$

$\bar{X}+ 1.23\ \sigma = 63$ (पप)

दोनों समीकरण हल करने पर

$\bar{X} -1.48\ \sigma = 35$

$\bar{X} + 1.23\ \sigma = 63$

$-2.71\ \sigma = -28$

$\sigma = 28/2.71$

$= 10.33$

σ का मान समीकरण 1) पर रखने पर

$\bar{X} - 1.48(10.33) = 35$

$\bar{X} -15.3 = 35$

$\bar{X} = 50.3$

अतः $\bar{X}= 50.3, \sigma =10.33$

10.6.3 उच्चतम तथा न्यूनतम वर्ग में उच्चतम तथा न्यूनतम स्कोर ज्ञात करना।

जब $\bar{X}$, σ तथा उच्चतम तथा न्यूनतम वर्ग का समानुपात ज्ञात हो तब आप उच्चतम तथा न्यूनतम वर्ग में उच्चतम तथा न्यूनतम स्कोर ज्ञात कर सकते हैं।

उदाहरण 10.8 : 5000 कर्मचारियों का भाड़ा सामान्य रूप से वितरित है जिसका माध्य रू 2000 तथा प्रमाप विचलन रू 120 है। ऊपर से 500 सबसे धनी कर्मचारियों के लिए न्यूनतम भाड़ा क्या है?

हलः दिया है N=5000, $\bar{X}$= 2000, σ =120

सबसे धनी कर्मचारियों का अनुपात = 500/5000 = 1/10 = 0.10

0.40 क्षेत्रफल के सापेक्ष का मान = 1.29

हम जानते हैं Z= (X –$\bar{X}$) / σ

$\bar{X}$ - 2000 = 154.8

$\bar{X}$ = 2154.8

अतः सबसे धनी 500 कर्मचारियों के लिए न्यूनतम भाड़ा रू 2154.8 है।

उदाहरण 10.9 एक परीक्षण के अंक/ मार्क्स सामान्य रूप से वितरित हैं जिसका माध्य 75 तथा प्रमाप विचलन 5 है। यदि ऊपर से 5 प्रतिशत छात्र ग्रेड ए प्राप्त करें तथा नीचे 25 प्रतिशत ग्रेड एफ तो ए का न्यूनतम तथा एफ का अधिकतम अंक ज्ञात करें।

हलः दिया है $\bar{X}$= 75, σ =5

Z= (X –$\bar{X}$) / σ

0.5 - 0.25 के सापेक्ष Z का मान = 0.25 क्षेत्रफल

= .0.68 (तलिका से)

$\therefore$ -0.68 = (X-75)/5

या, -0.68x5= X-75

X = 71.6 या 72

अतः नीचे से 25 प्रतिशत छात्रों का उच्चम अंक 72 होगा।

(0.05 - 0.05)= 0.45 क्षेत्रफल के सापेक्ष Z का मान

= 1.65 (तालिका से)

$\therefore$ 1.65 = (X-75)/5

या,1.65x5= X-75

या,8.25 = X-75

या,X= 83.25 या 83

अतः ऊपर से 5 प्रतिशत छात्रों का न्युनतम अंक 83 है।

10.7 सामान्य वक्र को फिट करना (Fitting the Normal Curve)

सामान्य वक्र को फिट करने के दो तरीके हैंः

10.7.1 कोटि विधि तरीका

इस विधि में SNV के कोटि का उपयोग करते हैं। इस विधि में निम्नलिखित चरण होते हैंः

1. सर्वप्रथम दिए गए वितरण का माध्य तथा प्रमाप विचलन का मान ज्ञात करो।
2. प्रत्येक वर्ग अंतराल का मध्य बिन्दु ज्ञात करो तथा इसे से प्रदशित करो।
3. प्रत्येक ग्के लिए Z= (X –$\bar{X}$) / σ ज्ञात करो।
4. कोटि तालिका से Z के प्रत्येक मान के लिए कोटि का मान ज्ञात करो।
5. इन प्रत्येक मानों को N x i/σ से गुणा करो तथा प्रत्याशित आवृत्ति ज्ञात करो।

यहाँ पर N = मदों की संख्या

I =वर्ग अंतराल की लम्बाई

σ =प्रमाप विचलन

उदाहरण 10.10 निम्नलिखित डाटा पर कोटि विधि से सामान्य वक्र फिट करो।

चर	0.10	10.20	20.30	30.40	40.50
f:	3	5	8	3	1

हलः सामान्य वक्र फिट करने के लिए $\bar{X}$ तथा σ ज्ञात करो

चर	**F**	**MV (X)**	**d**	**d' = d/i**	**fd'**	**fd'²**
0-10	3	5	-20	-2	-6	12
10-20	5	15	-10	-1	-5	5
20-30	8	25	0	0	0	
30-40	3	35	+10	+1	+3	3
40-50	1	45	+20	+2	+2	4
	N = 20				∑ fd'= -6	∑ fd'²= 24

$\bar{X}$ = A + ∑ fd'/N x i = 25+ -6/20 x 10 = 22

Σ = √ [∑ fd'²/N] – [∑ fd'/N]² x i = √24/20 –[-6/20]² x10

= 10.53

$\bar{X}$ तथा σ का मान ज्ञात करने के पश्चात निम्न प्रक्रिया अपनाओ-

चर	f	MV (X)	Z= (X – $\bar{X}$)/σ	कोटि तालिका से कोटि का मान	fe = कोटि x N x i/σ
0-10	3	5	-1.61	0.1092	2.07=2
10-20	5	15	-0.66	0.3209	6.09=6
20-30	8	25	0.28	0.3836	7.28=7
30-40	3	35	1.28	0.1872	3.55=4
40-50	1	45	2.18	0.0371	0.7046=1
					N= 20

10.7.2 क्षेत्रफल विधि (Area Method)

इस विधि में SNV के अन्तर्गत क्षेत्रफल तालिका का उपयोग करते हैं। इस विधि में निम्न चरण होते हैं-

1. सर्वप्रथम दिए गए वितरण का $\bar{X}$ तथा σ ज्ञात करो।
2. प्रत्येक वर्ग अन्तराल की न्यून सीमा को लिखो तथा इसे X से प्रदर्शित करो।
3. प्रत्येक न्यून वर्ग सीमाए X ए के लिए Z= (X –$\bar{X}$) / σ ज्ञात करो।
4. Z के प्रत्येक मान के लिए क्षेत्रफल ज्ञात कीजिए। (क्षेत्रफल तालिका से)
5. इसके बाद प्रत्येक दो लगातार मानों के मध्य क्षेत्रफल ज्ञात करो। इनके चिन्ह समान हैं और जब Z में भिन्न चिन्ह हों तो उनको जोड़ो।
6. इन प्रत्येक मानों में N से गुणा करके प्रत्याशित आवृत्ति ज्ञात करो।

उदाहरण 10.11 : निम्न डाटा पर सामान्य वक्र फिट करो।

चर	0-10	10-20	20-30	30-40	40-50
f:	3	5	8	3	1

हल : ऊपर के उदाहरण से हम प्राप्त करते हैं कि $\bar{X}$ =22, σ =10.53 तथा N=20

$\bar{X}$ तथा σ का मान ज्ञात करने के पश्चात हम निम्न विधि अपनाते हैंः

चर	**न्यून वर्ग सीमा**	**Z= (X –$\bar{X}$) / σ**	**0 से Z का क्षेत्रफल**	**प्रत्येक वर्ग अंतराल का क्षेत्रफल**	**fe=क्षेत्रफल x N**
0-10	0	-2.09	0.4817	0.1088	2.17=2

10-20	10	-1.04	0.3729	0.2975	5.95=6
20-30	20	0.19	0.0753	0.3518	7.036=7
30-40	30	0.76	0.2764	0.1800	3.6=4
40-50	40	1.71	0.4564	0.0397	0.794=1
50-60	50	2.66	0.4961		
					N=20

10.8 सारांश (Summary)

प्वायसन प्रक्रिया में दो घटनाओं के मध्य समय को घातांकी वितरण से वर्णन करते हैं। अर्थात एक प्रक्रिया जिसमें घटनाएं औसत दर से सतत तथा स्वतंत्र रूप से घटित होती हैं। यह गुणोत्तर वितरण का सतत रूप है। बीटा वितरण भी सतत प्रायिकता वितरण है जो कि अंतराल (0,1) पर परिभाषित है तथा इसके दो आकार प्राचल हैं तथा। बीटा वितरण का प्रयोग उन अनुपातों की सांख्यिकय माडलिंग में करते हैं जहाँ पर अनुपात 0 या 1 न हो। सामान्य वितरण एक बहुत महत्वपूर्ण तथा बहुतायत उपयोग होने वाला वितरण है इसका उपयोग मुख्यतः सतत दैव चरों जैसे वजन, लम्बाई, तथा छात्रों के समूह का बुद्धिमत्ता इत्यादि के व्यवहार के अध्ययन में होता है।

10.9 शब्दावली

- **घातांकी वितरणः** एक ऐसी प्रक्रिया जिसमें घटनाएं स्वतंत्र तथा सतत रूप से एक निश्चित औसत दर से घटित होती है।

10.10 अभ्यास प्रश्न (Practice Questions)

1.द्विपद वितरण असतत प्रायिकता है जबकि सामान्य वितरण ....प्रायिकता वितरण है।

2. सामान्य प्रायिकता वितरण की खोज अंग्रेजी गणितज्ञ ....ने 1733 में की थी।

3. सामान्य वितरण का प्रसार दोनों ही तरफ .... तक है ।

4. $\bar{X}$-2σ तथा $\bar{X}$+2σ के मध्य सामान्य वक्र का क्षेत्रफल ....है ।

10.11 अभ्यास प्रश्नों के उत्तर (Answers of Practice Questions)

1. सतत 2. अब्राहम डिमोरे 3. अनन्त 4. 0.9973

10.12 स्वपरख प्रश्न (Self Test Questions)

1. सामान्य वक्र के अंतर्गत तालिका का उपयोग करते हुए निम्नलिखित स्थितियों में क्षेत्रफल ज्ञात करोः

i) Z=0 तथा Z=1.3 के मध्य।

ii) Z=0.75 तथा Z=0 के मध्य।

iii) Z=-0.56 तथा Z=2.45 के मध्य।

iv) Z=0.85 तथा Z=1.96 के मध्य।

2. 1000 मजदूरों के प्रतिदर्श में औसत वजन 45 किग्रा तथा प्रताप विचलन 15 किग्रा है। सामान्य वितरण मानते हुए 40 तथा 60 किग्रा के मध्य वजन वाले मजदूरों की संख्या बताइए।

3. सामान्य वितरण जिसका माध्य 5 तथा SD = 3 हो से एक मद दैव निदर्शन द्वारा चुने जाने पर उसके 2.57 तथा 4.34 के मध्य होने की प्रायिकता ज्ञात करो।

4. एक सामान्य वितरण का माध्य 12 तथा प्रमाप विचलन 2 है। X_1= 9.6 तथा X_2= 13.8 के मध्य का क्षेत्रफल ज्ञात करो।

5. एक सामान्य वितरण जिसका माध्य 12 तथा प्रमाप विचलन 2 हो X_1=6 तथा X_2= 18 के मध्य वक्र के अंतर्गत क्षेत्रफल ज्ञात करो।

6. एक परीक्षा में परीक्षार्थियों द्वारा प्राप्त अंक सामान्य वितरित है। यदि 10 प्रतिशत ने 40 से कम अंक प्राप्त किए हों तथा 15 प्रतिशत ने 80 से अधिक अंक प्राप्त किए हों तो अंको का माध्य तथा प्रमाप विचलन ज्ञात करो।

7. एक परीक्षा में 15 प्रतिशत उम्मीदवार विशिष्टता के साथ उत्तीर्ण हुए जबकि 25 प्रतिशत फेल हो गए। हमें यह ज्ञात है कि 100 में से 40 से कम अंक प्राप्त करने पर उम्मीदवार फेल हो जाता है तथा विशिष्टता प्राप्त करने के लिए 75 से अधिक अंक प्राप्त करना होता है। अंकों को सामान्य वितरित मानते हुए उनके वितरण का माध्य तथा प्रमाप विचलन ज्ञात करो।

8. दिया गया है कि 84 प्रतिशत आदमियों की लम्बाई 65.2 इंच से कम है तथा 68 प्रतिशत आदमियों की लम्बाई 65.2 इंच 62.8 के मध्य है। आदमियों के समूह के लम्बाई को सामान्य वितरित मानते हुए उसका माध्य तथा प्रमाप विचलन ज्ञात करो।

9. एक परीक्षा में उत्तीर्ण तथा विशिष्टता का प्रतिशत क्रमशः 46 तथा 9 है। उम्मीदवारों द्वारा प्राप्त औसत अंक तथा उनका प्रमाप विचलन ज्ञात करो। जब न्यूनतम उत्तीर्ण अंक तथा विशिष्टता अंक क्रमशः 40 तथा 75 है। (मानो कि अंकों का वितरण सामान्य है।)

10. 500 मजदूरों की मासिक आय सामान्य वितरित है जिसका माध्य रू 2000 तथा प्रमाप विचलन रू 200 है। सबसे अमीर मजदूरों की न्यूनतम आय बताओ।

10.13 स्वपरख प्रश्नों के उत्तर (Answer to Self Test Questions)

1. i) 0.4032 ii) 0.2734 iii) 0.7052 iv) 0.1727
2. 471
3. 0.2039
4. 70.08 %
5. 99.74 %
6. [=62.15, =17.16]
7. [=53.79. =20.46]
8. [=64, =1.2]
9. [=37.18, =28.22]
10. 2,134rs

10.14 सन्दर्भ पुस्तकें (References)

1. Roy Ramendu, ***'Principles of Statistics'***, Prayag Pustak Bhandar, Allahabad.
2. Gupta, S.P. & Gupta, M.P., ***"Business Statistics"*** , Sultan Chand & Sons, New Delhi.
3. Shukla S.M., & Sahai S.P., ***"Advanced Statistics",*** Sahitya Bhawan Publication, Agra.
4. Goon, Gupta and Dasgupta, ***"Basic Statistics"*** World Press Limited, Calcutta.

5***. “Fundamentals of Business Statistics”*** – Sanchethi and Kapoor.
6. Srivastava, Shenoy and Guptha, **“Quantitative Methods in Management”.**

इकाई 11 केन्द्रीय प्रवृत्ति की माप
(Measures of Central Tendency)

11.1 प्रस्तावना (Introduction)

प्रस्तुत इकाई मे पाठकों को केन्द्रीय प्रवृत्ति की प्रमापें की जानकारी दी जाएगी। इसके अंतर्गत समंको को सांख्यिकीय माध्य द्वारा प्रकट करने की विभिन्न विधियों की व्याख्या की जाएगी। सांख्यिकीय विष्लेषण के अन्तर्गत वर्गीकरण और सारणीयन द्वारा समकों के विषल समूह को आवृत्ति बंटन के रूप में प्रस्तुत करके सरल और समझने योग्य बनाया जाता है। परन्तु इससे समंको की महत्वपूर्ण विशेषताओं का पता नहीं चलता। सारांश रूप में समंको को सांख्यिकीय माध्य द्वारा प्रकट किया जा सकता है। यह एक ऐसा मूल्य है जिसके आस-पास अन्य सकंको के केन्द्रित होने की प्रवृत्ति पायी जाती है तथा जो समंक लगभग श्रेणी के मध्य में होता है तथा केन्द्रीय प्रवृत्ति की माप है। माध्य सांख्यिकीय विष्लेषण की महत्वपूर्ण माप है। यह मूल्य चिड़िया की आँख जैसी दृष्टि प्रदान करता है। माध्यों को स्थान सम्बन्धीया प्रतिरूपी मूल्य कहते हैं।

सारांश रूप में समंकों को प्रस्तुत करने के लिये एक संख्यात्मक माप की आवश्यकता होती है। एक उदाहरण में एक बाग में 360 पेड़ हैं, जिनकी ऊँचाई का आवृत्ति बंटन निम्न हैः

ऊँचाई (फीट में)	0.7	7.14	14.21	21.28	28.35	35.42
आवृत्ति	26	31	35	49	82	71

11.2 उद्देश्य (Objective)

इस इकाई का अध्ययन करने के उपरान्त आप

- ✓ माध्यों के विभिन्न प्रकारों को जानने में सक्षम होंगे।
- ✓ गणना की विधियाँ एवम् सूत्रों के बारे में जान सकेंगे।
- ✓ इन माध्यों की विभिन्न उपयोगों की जानकारी प्राप्त करेंगें।

11.3 सांख्यिकीय माध्यों के उद्देश्य व कार्य (Objectives and Functions of Statistical Means)

माध्य आसानी से व्यक्त न होने वाले जटिल-

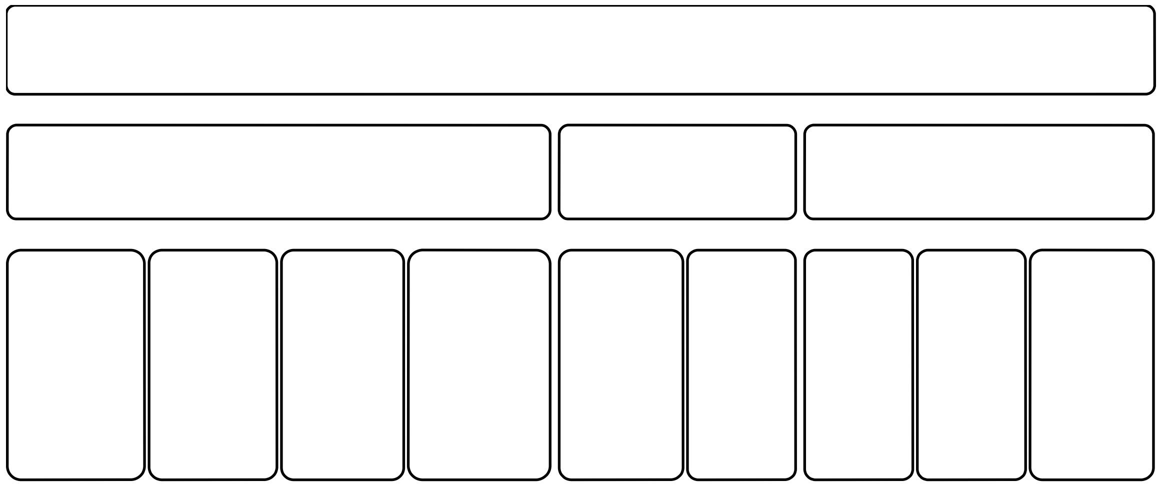
i) समंकों का संक्षिप्त चित्र प्रस्तुत करता है।
ii) माध्यों की एक महत्वपूर्ण उपयोगिता विभिन्न समंक समूहों के बीच तुलना की सुविधा है।
iii) अन्य सांख्यिकीय विवेचन जैसे, उपकिरण, विषमता और प्रथुषीर्षत्व की गणना में माध्यों का उपयोग होता है।
iv) माध्य पथ प्रदर्शन का कार्य करते हैं।

11.4 आदर्श माध्य के आवश्यक तत्व (Essential Elements of Ideal Mean)

प्रो0 यूल (Prof. Yule) के द्वारा आदर्श माध्य के आवश्यक तत्वों का वर्गीकरण निम्न प्रकार किया गया है।

i) इन्हें सांख्यिक (Statistician) के अनुमान पर आधारित नहीं होना चाहिये।
ii) इन्हें सभी मूल्यों पर आधारित (Based on all observations) होना चाहिए।
iii) यह आसानी से समझ मं आने और रेखांकित (Simple to understand and locate) होने योग्य होना चाहिये।
iv) इनमें निर्धारण की सरलता होनी चाहिये।
v) इनमें अधिकतम अथवा न्यूनतम चरों के मूल्य का अधिक प्रभाव नहीं पड़ना चाहिये।
vi) इनकी कुछ सरल और स्पष्ट गुण होने चाहिये।
vii) माध्यों पर प्रतिचयन के परिवर्तन (effect of sampling) का न्यूनतम प्रभाव हो। एक ही समग्र के विभिन्न प्रतिचयन/प्रतिदर्श के विभिन्न माध्य नहीं होने चाहिये।
viii) इनका अग्रिम गणितीय विश्लेषण में उपयोग होना चाहिये।

11.5 सांख्यिकीय माध्यों के प्रकार (Kinds of Statistical Average)



उपर्युक्त विभिन्न माध्यों में समान्तर माध्य, मध्यका और बहुलक महत्वपूर्ण हैं।

11.5.1 गणितीय माध्य (Mathematical Average)

समान्तर माध्य - यह सबसे लोकप्रिय माध्य है। समान्तर माध्य वह मूल्य है जो किसी समंकमाला के सभी पदों के मूल्यों के योग में उन पदों की संख्या से भाग देने पर प्राप्त होता है। समान्तर माध्य को सामान्य रूप में x लिखते हैं। यदि पद $x_1, x_2, x_3, \ldots\ldots, x_n$ हों तो उनका समान्तर माध्य

$$\bar{X} = =$$

यहाँ सिग्मा योग की अभिव्यक्ति है और i उपसंकेत के लिये प्रयुक्त है यदि पद $x_1, x_2, x_3, \ldots\ldots, x_n$ और हों तो इनका योग $\sum$ के रूप में व्यक्त होगा और

$$\bar{X} = =$$

सामान्तर माध्य की उपर्युक्त गणना व्यक्तिगत श्रेणी में होती है। यदि पदों की आवृतियाँ दी गई हों अर्थात् यदि खण्डित श्रेणी (discrete series) में सामान्तर माध्य ज्ञात करना हो तो-

$$\bar{X} =$$

$$\bar{X} = f_1X_1 + f_2X_2 + f_3X_3 + \ldots\ldots + f_nX_n \,/\, f_1 + f_2 + f_3 + \ldots f_n$$

अविछिन्न श्रेणी (continuous series) में समान्तर माध्य ज्ञात करने के लिए वर्गों के मध्य-मूल्य (mid value) ज्ञात कर लेते हैं फिर खण्डित श्रेणी की भांति ही समान्तर माध्य की गणना करते है।

समान्तर माध्य की गणना दो प्रमुख विधियों द्वारा की जाती है-

i) प्रत्यक्ष रीति (Direct Method)
ii) अप्रत्यक्ष रीति (Indirect method)

दोनों रीतियों से निकाला गया समान्तर माध्य बराबर ही आता है। अप्रत्यक्ष रीति अथवा लघु रीति का प्रयोग गणना की क्रिया को सरल बनाने के लिए किया जाता है। इसके अन्तर्गत इन मूल पदों को, एक ऐच्छिक मूल्य जिसे कल्पित माध्य (Assumed Mean) कहा जाता है तथा जिसे A से लिखते हैं, द्वारा एक नये पद (dx तथा dx) में परिवर्तित करते हैं और फिर उचित सूत्र द्वारा समानान्तर माध्य की गणना कर लेते हैं।

व्यक्तिगत श्रेणी (Individual Series)

हमने देखा कि व्यक्तिगत श्रेणी में समान्तर माध्य हैं।

$\bar{X}$ =

यह प्रत्यक्ष रीति है।

लघुरीति (Short-cut Method)

इससे नया पद इस प्रकार का होता है कि dx= X-A अर्थात् मूल पदों में से एक स्थिर पद घटाते हैं। A का मान श्रेणी मे से ही मध्य का रखा जाता है।
अब प्रत्येक (X-A) का जोड़कर ∑dx प्राप्त करते हैं तथा मूल श्रेणी का समान्तर माध्य का मान ज्ञात करते हैं-

$$\bar{X} = A +$$

इसे निम्न सूत्र की उत्पत्ति द्वारा भी समझा जा सकता है-

या $dx = X-A$

$X = A + dx$

दोनो तरफ योग करने पर -

$X = \sum A + dx$

या = = (दोनो तरफ N से भाग देने पर)

$\bar{X} = A +$

इसी को इस रूप में भी लिख सकते हैं-

$\bar{X} = A + \bar{d}x$ क्योंकि $\bar{d}x$ =

उदाहरण 1. **:** निम्नलिखित आंकड़ों का माध्य ज्ञात कीजिए

विद्यार्थी	A	B	C	D	E	F	G	H	I	J
लम्बाई	158	150	165	163	162	166	164	180	155	167

हलः प्रत्यक्ष तथा लघु रीति द्वारा समान्तर माध्य

$$\bar{X} = A +$$

विद्यार्थी	लम्बाई	dx= X-A (A=160)
A	158	158-160=-2
B	150	150-160=-10
C	165	165-160=+5
D	163	163-160=+3
E	162	162-160=+2
F	166	166-160=+6
G	164	164-160=+4
H	180	180-160=+20
I	155	155-160=-5
J	167	167-160=+7
	∑X = 1630	∑dx = +37

i) प्रत्यक्ष रीति द्वारा माध्य = $\bar{X}$ = 1630/10 = 163

ii) लघु रीति द्वारा माध्य = 160+37/10 = 163.7 जो लगभग एक समान है।

खण्डित श्रेणी में माध्य की गणना (Calculating the Mean in a Discrete Series):

उदाहरण 2 : एक कपड़ा मिल में कार्य करने वाले श्रमिकों की दैनिक मजदूरी (रु0) निम्न है, इनकी औसत मजदूरी ज्ञात कीजिए।

दैनिक मजदूरी (X)	100	120	140	160	180	200	220
श्रमिकों की संख्या (f_i)	3	6	10	15	24	42	75

हल:

	X	**f_i**	**d_i = X -A**	**f_id_i**
	100	3	-60	-180
	120	6	-40	-240
	140	10	-20	-200
A →	160	15	0	0
	180	24	+20	480
	200	42	+40	1680
	220	75	+60	4500
		$\sum f_i = 175$		$\sum f_id_i = 6040$

नोट → यहाँ अवश्य जाँच लें इससे ज्ञात होता है कि मानक माध्य (Assumed Mean) का मान सही है। fidi

$\bar{X}$ = A + = 160 +6040/175 = 194.5

नोट यह गणना लघु रीति द्वारा की गयी है, पाठक स्वयं प्रत्यक्ष रीति द्वारा माध्य की गणना कर जाँच करें।

उदाहरण 3: निम्न समंको से माध्य (Arithmetic Average) की गणना कीजिए

	X	**f_i**	**d_i = X_i -A**	**f_id_i**
	5	3	-5	-15
	6	4	-4	-16
	7	2	-3	-6
	9	4	-1	-4
A →	10	2	0	0
	12	5	2	10
	13	3	3	9
	15	2	5	10
		$\sum f_i = 25 = N$		$\sum f_id_i = -12$

यहाँ $\bar{X}$ = A + = 10 – 12/25 = 9.52 (hrs)

उदाहरण 4 : निम्न समंको द्वारा माध्य की गणना कीजिए

	X	**f_i**	**d_i = X_i -A**	**f_id_i**
	1	5	-3	-15
	2	9	-2	-18
	3	12	-1	-12
A →	4	17	0	0

	5	14	+1	14
	6	10	+2	20
	7	6	+3	18
		$\sum f_i = 73 = N$		$\sum f_i d_i = -7$

$\bar{X} = A + \quad = 4 + (-7)/73 = 4 – 7/73 = 3.904$

अविच्छिन्न श्रेणी के लिये माध्य (Mean for Continuous Series)

अविछिन्न श्रेणी में समान्तर माध्य की गणना करने के लिये वर्ग अन्तराल (class interval) के मध्य बिन्दु ज्ञात कर लिये जाते हैं। अविच्छिन्न श्रेणी में यह मान लिया जाता है कि आवृत्तियाँ मध्य बिन्दुओं पर केन्द्रित हैं।

माध्य निकालने की रीति

अविच्छिन्न श्रेणी में माध्य की गणना -

प्रत्यक्ष रीति

अप्रत्यक्ष रीति तथा पद विचलन रीति द्वारा की जा सकती है।

उदाहरण 5: निम्न समंको से माध्य की गणना प्रत्यक्ष एवं अप्रत्यक्ष रीति द्वारा कीजिए ।

वर्ग अन्तराल	**मध्य बिन्दु (X)**	**बारम्बारता (fi)**	**fx**	**(X-A) di**	**fidi**
0-10	5	10	50	-20	-200
10-20	15	12	180	-10	-120
20-30	25	25	625	0	0
30-40	35	18	630	10	180
40-50	45	15	675	20	300
		$\sum f_i = 80$	$\sum fx = 2160$		$\sum f_i d_i = 160$

यहाँ X मध्य बिन्दु की गणना $L_1+L_2 / 2$ द्वारा करते हैं।

$$\bar{X} = \frac{\sum fx}{\sum f} = \frac{2160}{80} = 27$$ (प्रत्यक्ष रीति द्वारा)

अप्रत्यक्ष रीति द्वारा

$$\bar{X} = A + \frac{\sum fdx}{\sum f} = 25 + \frac{160}{80} = 27$$

दोनों रीति से उत्तर समान है।

पद विचलन रीति- यदि अविच्छिन्न श्रेणी में वर्ग-विस्तार ($L_{2-} L_1$) समान हो अर्थात् प्रत्येक वर्गों में वर्ग अन्तराल एक निश्चित संख्या के बराबर हो तो हम समानान्तर माध्य की गणना करने के लिये एक ऐसा पद बनाते है जो X-A/ N के बराबर होता है।

इसमें दो बातें ध्यान देने योग्य हैं-

1. **मूल का परिवर्तन (Change of Origin)** -लघु रीति में प्रत्येक पद (X) में से स्थिर राशि मानक माध्य (A) को घटाते हैं।

2. **पैमाने का परिवर्तन (Change of Scale)-** इसे dx’ लिखते हैं तथा X-A / N जहाँ N वर्ग अन्तराल का मान है।

इस रीति द्वारा माध्य की गणना निम्न सूत्र द्वारा की जाती हैं-

$$\bar{X} = A + N$$

उदाहरण 6:

वर्ग अन्तराल	**मध्य बिन्दु (X)**	**बारम्बारता (f)**	**dx' = (X-A)/N**	**fdx'**
0-10	5	10	-2	-20
10-20	15	12	-1	-12
20-30	**25** (A)	25	0	0
30-40	35	18	1	18
40-50	45	15	2	30
		∑ f = 80		∑ fdx' =+ 32

यहाँ सभी वर्गों का अन्तराल 10 है अतः समान्तर माध्य

$$\bar{X} = A + N$$

$$= 25+10 \text{ x } 25+ 2 = 27$$

स्मरणीय तथ्य

1. इसमें कुछ बाते स्मरणीय है जैसे यदि वर्गान्तरों को इस रूप में दिया जाए की आवृतियाँ संचयीं हों (वर्ग अन्तराल कम अथवा अधिक रुप में दिये हों) तो वितरण को सामान्य वितरण में बदला जा सकता है तथा समान्तर माध्य की गणना की जा सकती है।

जैसे -

वर्ग अन्तराल	**आवृत्ति**
10 से कम	10
20 से कम	18
30 से कम	30
40 से कम	37

सामान्य वितरण

वर्ग अन्तराल	**आवृत्ति**
0-10	10
10-20	18
20-30	30
30-40	37

2. असमान वर्ग वितरण पर भी समान्तर माध्य की गणना उसी रुप में की जा सकती हैं।

3. खुले हुए वर्ग अन्तराल (Open end classes) में माध्य की गणना करने के लिये खुले वर्ग को बन्द कर लेना आवश्यक होता है।

समान्तर माध्य के गणितीय गुण (Mathematical Properties of Arithmetic Mean)-

विशेषता 1: - विभिन्न मूल्यों (X) के विचलनों (dx) का बीजगणितीय योग शून्य होता है।

$$\sum (X-\bar{X}) = 0$$

विशेषता 2: यदि $\bar{X}$ = में $\sum X$,N, $\bar{X}$ में से कोई दो की माप ज्ञात हो तो तीसरे की गणना की जा सकती है। इसी विषेषता के आधार पर हम अज्ञात आवृत्तियों का निर्धारण तथा गलत माध्य को सही माध्य में परिवर्तित कर सकते हैं।

उदाहरण 8- विश्व कप फुटबाल मैचों में एक सप्ताह में सोमवार से शनिवार तक औसत रुप में प्रति मैच 3 गोल हुए । रविवार को एक संघर्षपूर्ण मैच में अत्यधिक गोल होने पर पूरे सप्ताह में गोल का औसत 4 हो गया। बताइये रविवार को कुल कितने गोल हुए ?

हल - सोमवार से शनिवार तक औसत गोल = 3

6 दिनों में कुल गोल = 3 x 6 = 18

सोमवार से रविवार तक गोल का औसत = 4

अतः सात दिनों में कुल गोल = 7 x 4 =28

सातवें दिन (रवि0) को कुल गोल = 28 – 18 = 10

उदाहरण 9: यदि बारम्बारता ज्ञात न हो –

यदि $205 + fi = \sum f$

$$722 + 5fi = \sum fx$$

$\bar{X} = 3.763$ दी हुई हो तो fi ज्ञात कीजिए

$$3.763 = \frac{722 + 5f_1}{205 + f_1}$$

$$= 771.415 + 3.763 f_1 = 5f_1$$

$$49.415 = 1.123f_1$$

$$f_1 = \frac{49415}{1123} = 44.0$$

अतः अज्ञात आवृत्ति है fi = 44

विशेषता 3: सही माध्य का निर्धारण (Determination of True Mean)

यदि N पदों के समान्तर माध्य X ज्ञात किया और बाद में यह पता चला कि उसमें एक या दो से अधिक मूल्य गलत जोड़े गये, तो सही मूल्य ज्ञात होने की स्थिति में हम निम्न प्रक्रिया अपनाते हैं-

1. पहले हम गलत योग $\sum X = \bar{X}X(N)$ ज्ञात करते हैं
2. इस गलत योग $\sum X$ में से गलत पदों को घटाकर सही पदों को जोड़ देते हैं।
3. यही योग $\sum X$ में N का भाग कर सही $\bar{X}$ माध्य प्राप्त होता है।

विशेषता 4: सामूहिक माध्य (Combined Mean)

यदि किसी समूह के दो या दो से अधिक उपसमूहों के समान्तर माध्य और पदों की संख्या ज्ञात होतो सामूहिक माध्य की गणना निम्न सूत्र द्वारा की जा सकती है-

$\bar{X}$ =

सामान्य रुप में

$\bar{X}$ =

विशेषता 5: यदि किसी श्रेणी के सभी मूल्यों में निश्चित अचर राशि जोड़ दी जाए घटा दी जाए, गुणा कर दी जाए तो उस श्रेणी के समान्तर माध्य में वह अचर राशि क्रमशः जुड़ जाती है गुणा हो जाती है।

विशेषता 6: दो श्रेणियों के तत्संवादी मूल्यों (Coressponding Values) के जोड़ों और अन्तरों का समान्तर माध्य उन दोनों श्रेणियों के समान्तर माध्य के याग या अन्तर के बराबर होता है।

समान्तर माध्य के गुण (Properties of Arithmetic Mean) :

एक आदर्श माध्य के रुप में समान्तर माध्य की निम्न विशेषताएँ पायी जाती हैं - यह

1. सदैव निष्चयात्मक होता है
2. गणना करने में सरल तथा सामान्य होता है
3. यह सभी मूल्यों पर आधारित होता है
4. दूसरे माध्यों की तुलना में प्रतिचयन के प्रभावों से कम प्रभावित होता है।

समान्तर माध्य की सीमाएँ (Limitations of Arithmetic Mean):

समान्तर माध्य दोषमुक्त नही है इसकी निम्न सीमाएँ हैं-

1. यह चरम मूल्यों से प्रभावित होता है
2. इसका निरीक्षण अथवा बिन्दुरेखीय निर्धारण सम्भव नही है।
3. अनुपात, दर, प्रतिशत की गणना में उपयुक्त नही है
4. असमित वितरण की दषा में समान्तर माध्य वितरण का उचित प्रतिनिधित्व नही करता।

11.5.2 भारित समान्तर माध्य (Weighted Arithmetic Mean)

व्यवहारिक जीवन में अनेकों ऐसी स्थितियाँ होती हैं जिनमें सभी पद समान महत्व के नही होते। ऐसी स्थिति में समान्तर माध्य निकालते समय मूल्यों के सापेक्षिक महत्व को ध्यान में रखना आवश्यक होता है, इस प्रकार निकाले जाने वाले माध्य को समानान्तर माध्य कहते हैं।

यदि मूल्य $\rightarrow \quad X_1, \; X_2, \; X_3, \; X_4, \; \ldots\ldots, X_n$

सम्बद्ध भार $\rightarrow \quad W_1, W_2, W_3, \; \ldots\ldots, W_n$ हैं तो -

$$\bar{X}_w = \frac{W_1X_1 + W_2x_2 + W_3X_2 + \ldots\ldots + W_nX_n}{W_1 + W_2 + W_3 + \ldots\ldots + W_n} = \frac{\sum WX}{\sum W}$$

आवृत्ति वितरण में -

$$\bar{X}_w = \frac{W_1(f_1X_1) + W_2(f_2X_2) + \ldots\ldots + W_n(f_nX_n)}{W_1 + W_2 + \ldots . W_n} = \frac{\sum W(fX)}{\sum W}$$

भारित माध्य (Weighted Mean) की गणना प्रत्यक्ष और अप्रत्यक्ष रीति दोनो प्रकार से की जा सकती है। लेकिन इसे प्रत्यक्ष रीति द्वारा ही निकालते है।

उदाहरण 10: दो विद्यालयों का परीक्षाफल दिया गया है, कौन सा विद्यालय बेहतर है ?

प्रतिशत परीक्षाफल

	विद्यालय A	विद्यालय B

हाइस्कूल	70% (200 विद्यार्थी)	80% (150 विद्यार्थी)
बी0 ए 0	60% (150 विद्यार्थी)	80% (50 विद्यार्थी)

हल : विद्यालय A का भारित माध्य -

$$\bar{X}W_1 = \frac{200\times70 + 150\times60 + 100\times80}{200+150+100} = \frac{31000}{450} = 68.9$$

स्पष्ट है कि यहाँ विद्यार्थियों की संख्या को भार (W) तथा प्रतिशत परीक्षाफल को पद मूल्य (X) माना गया है ।

विद्यालय **B** का भारित समानान्तर माध्य

$$\bar{X}W_2 = \frac{150\times80 + 100\times60 + 50\times80}{150+100+50} = \frac{22000}{300} = 73.3$$

अतः विद्यालय B का परीक्षाफल बेहतर है।

11.5.3 गुणोत्तर माध्य (Geometric Mean)

परिभाषा: किसी श्रेणी के n पदों के गुणनफल को n वाँ मूल (root √) ही उसका गुणोत्तर माध्य है। पदों की संख्या से ही मूल का मान ज्ञात होता है। अतः 2,3 संख्याओं में गुणोत्तर माध्य की गणना आसानी से की जा सकती है जैसे -

पद $\rightarrow$ X, Y $\qquad$ G.M $\rightarrow \sqrt[2]{X.Y}$

पद $\rightarrow$ X, Y, Z $\qquad$ G.M $\rightarrow \sqrt[3]{X.Y.Z}$

जब दो या तीने से अधिक पद होते हैं तो गुणोत्तर माध्य

$$G.M. = \text{Anti log}\left[\frac{\log X_1 + \log X_2 + \ldots\ldots + \log X_n}{n}\right]$$ (∽ =n)

$$= \text{Antilog}\left[\frac{\sum \log X}{n}\right]$$

यदि खण्डित अविच्छिन्न श्रेणी हो तो गुणोत्तर माध्य

$$G.M = \text{Antilog}\left[1/N\sum f \log X\right] \quad \text{जहाँ} \quad N = \sum f$$

उपयोग: गुणोत्तर माध्य का प्रयोग प्रतिशत वृद्धि-दरों जैसे जनसंख्या वृद्धि दर, चक्रवृद्धि ब्याज, मूल्यों में होने वाले प्रतिशत परिवर्तन आदि की औसत दरें ज्ञात करने के लिये होता है। इसका संबंधित सूत्र निम्न है-

i) $$P_n = P_0(1+r)^n$$ (∽ =n)

ii) $$r = \left[\sqrt[n]{\frac{P_n}{P_0}} - 1\right]$$

जहाँ

P_n or P_n = निश्चित अवधि के उपरान्त पर मूल्य की राशि

P_0 = अवधि के आरम्भ में चर मूल्य की राशि

$\backsim$ or n= वर्षों की संख्या

r = प्रति इकाई परिवर्तन की दर

इसे चक्रवृद्धि ब्याज सूत्र Compound Interest Formula) कहते हैं।

उदाहरण 11:

दशक	**वृद्धि दर**	**जनसँख्या दशक के अंत में जब पूर्व दशक में 100 मान लें**	$\log_{10} x$
1	5	105	2.0212
2	8	108	2.0334
3	12	112	2.0492

$$\text{G.M.} = \text{Antilog}\left\{\frac{1}{\backsim}\sum \log x\right\} = \text{Antilog}\left\{\frac{1}{3}(6.1038\right\}$$

$$= \text{Antilog}\,(2.0346) = 108.2$$

सामूहिक गुणोत्तर माध्य

इसी प्रकार सामूहिक गुणोत्तर माध्य की गणना की जा सकती हैं

सामूहिक गुणोत्तर माध्य $= \text{Antilog}\left[\frac{N_1 \log G_1 + N_2 \log G_2}{N_1 + N_2}\right]$

जहाँ $G_1 \rightarrow$ एक भाग का गुणोत्तर माध्य

$N_1 \rightarrow$ एक भाग के पदों की संख्या

$G_2 \rightarrow$ दूसरे भाग का गुणोत्तर माध्य

$N_2 \rightarrow$ दूसरे भाग के पदों की संख्या

11.5.4 भारित गुणोत्तर माध्य (Weighted Geometric Mean)

यदि विभिन्न मूल्यों का सापेक्षिक महत्व अलग-अलग हो तो भारित समान्तर माध्य की ही तर भारित गुणोत्तर माध्य ज्ञात किया जा सकता है।

भारित गुणोत्तर माध्य $\text{W.G.M} = \text{Antilog}\left[\frac{\sum(\log X.W)}{\sum W}\right]$

11.5.5 हरात्मक माध्य (Harmonic Mean)

परिभाषा - किसी समंक श्रेणी के पदों की संख्या को पदों के व्युत्क्रमों (reciprocal) के योग से भाग देने पर जो मूल्य प्राप्त होता है उसे उस श्रेणी का हरात्मक माध्य कहतें हैं -

$$\text{H.M.} = \frac{N}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots + \frac{1}{x_\backsim}}$$

$$= \frac{N}{\sum \frac{1}{x}}$$

हरात्मक माध्य का व्यावहारिक उपयोग सीमित है। औसत गति, चलन वेग तथा प्रति रुपये वस्तु की मात्रा का प्रयोग किया जाता है। कभी-कभी भारित हरात्मक माध्य की गणना भी करनी पड़ती है। भारित हरात्मक माध्य का सूत्र निम्न है।

भारित हरात्मक माध्य

$$(\text{W.H.M.}) = \frac{N}{\sum \frac{W_i}{X_i}}$$

या

$$\frac{1}{\frac{w_1}{x_1} + \frac{w_2}{x_2} + \frac{w_3}{x_3} + \ldots + \frac{w_n}{x_n}}$$

समान्तर माध्य, गुणोत्तर माध्य तथा हरात्मक माध्य के संबंध में कुछ स्मरणी बातें –

1. किसी श्रेणी के गैर ऋणात्मक मूल्यों के लिए

$$\text{A.M.} \geq \text{G.M.} \geq \text{H.M}$$

यदि गैर ऋणात्मक मूल्य बराबर नही हैं-

$$\text{A.M.} > \text{G.M.} > \text{H.M}$$

यदि सभी मूल्य बराबर हैं -

$$\text{A.M.} \geq \text{G.M.} \geq \text{H.M}$$

2. यदि केवल दो संख्याएँ होती है तो -

$$(\text{G.M.})^2 = \text{A.M.} \times \text{H.M.}$$

यदि दो संख्याएँ $a > 0$ तथा $b > 0$ हों तो हम जानते हैं कि इनका

$$\text{A.M.} = \frac{a+b}{2},$$

$$\text{H.M.} = \frac{2}{\frac{1}{a} + \frac{1}{b}} = \frac{2ab}{a+b}$$

$$\text{G.M.} = \sqrt{ab}$$

तथा

$$\text{A.M.} \times \text{H.M.} = \frac{a+b}{2} \times \frac{2ab}{a+b}$$

$$= ab$$

$$= \left(\sqrt{ab}\right)^2$$

अतः $$(\text{A.M}) \times (\text{H.M}) = (\text{G.M})^2$$

11.5.6 स्थिति - संबंधी माध्य (Positional Averages)

मध्यका (Median):

मध्यका किसी समंक श्रेणी में मध्यका वह मूल्य होता है जो पुरे श्रेणी को दो बराबर भागों में विभाजित करता है, जिसमें श्रेणी को घटते से बढ़ते क्रम में व्यवस्थित करते हैं तो उसका मध्य मूल्य मध्यका होता है, जिसे M या Md से व्यक्त करतें हैं।

नोटः यहाँ श्रेणी को घटते से बढ़ते अथवा बढ़ते से घटते क्रम में व्यवस्थित करना आवश्यक हैं।

मध्यका $=\left(\frac{n+1}{2}\right)$ वाँ पद का मान

यदि पदों की संख्या सम (even) हो तो $\left(\frac{n+1}{2}\right)$ वाँ पद दशमलव में आता है अतः

4.5 वाँ पद का मान = चौथा पद का मान + पाँचवे पद का मान / 2

उदाहरण 12: $X \rightarrow 7,\ 11,\ 12,\ 15.$ का मध्यका $\left(\frac{4+1}{2}\right)$ वाँ पद

अतः मध्यका $=\frac{11+12}{2}=11.5$

उदाहरण 13: - निम्न श्रेणी का मध्यका मूल्य ज्ञात कीजिए।

$25,\ 13,\ 23,\ 40,\ 27,\ 25,\ 23,\ 24,\ 22,\ 30$

हल -

क्रम सं0	1	2	3	4	5	6	7	8	9	10
पद मूल्य	13	22	23	23	24	25	25	27	30	40

मध्यका $=\frac{10+1}{2}=5.5$ वाँ पद

अतः 5 वाँ पद मान + 6 वाँ पद मान /2

खण्डित श्रेणी में मध्यका (Median in Discrete Series)

निम्न विधि द्वारा ज्ञात की जाती हैं:

1. पहले संचयी आवृत्ती ज्ञात करते हैं।
2. $(N+1/2)$ वाँ पद निकालते हैं।
3. संचयी आवृति के जिस पद में $(N+1/2)$ वाँ पद शामिल हो उसी के सामने वाले मूल्य को मध्यका मान कहते है।

उदाहरण 14: निम्न श्रेणी में मध्यका मूल्य ज्ञात कीजिए ।

x	→	0	1	2	3	4	5	6	7	8
f	→	1	9	26	59	72	52	29	7	1

हल:

x	→	0	1	2	3	4	5	6	7	8
f	→	1	9	26	59	72	52	29	7	1
c.f	→	1	10	36	95	167	219	248	255	256

$$M = \left(\frac{N+1}{2}\right)^{th} \text{ item} = \frac{256+1}{2} = 128.5^{th} \text{ item}$$ यहाँ 128.5 वाँ संचयी आवृत्ति 167 में पहली बार आ रहा है, अतः मध्यका का मान 4 हैं।

अविच्छिन्न श्रेणी में मध्यका (Median in Continuous Series)

अविच्छिन्न श्रेणी में मध्यका का पद मान ज्ञात करने के लिये निम्न सूत्र का प्रयोग करते हैं।

$$Md = \frac{N}{2}$$ वाँ पद

तथा मध्यका की गणना निम्न सूत्र द्वारा करते हैं।

$$M = L_1 + \left(\frac{\frac{N}{2} - C}{f}\right) \times i$$

जहाँ M = मध्यका

L_1 = मध्यका वर्ग की निचली सीमा

c = मध्यका वर्ग के ठीक पहले वाली संचयी आवृत्ति

f = मध्यका वर्ग की आवृत्ति

i = मध्यका वर्ग का विस्तार $(L_2 - L_1)$

पहले खण्डित, श्रेणी के समान ही संचयी आवृत्ति ज्ञात करते है-

$\frac{N}{2}$ द्वारा मध्यका वर्ग ज्ञात करतें हैं-

जिस संचयी आवृत्ति में प्रथम बार $\frac{N}{2}$ शामिल हो उसी के सामने वाला वर्ग मध्यका वर्ग होगा।

उदाहरण 15: निम्न वितरण से मध्यका मान ज्ञात कीजिए

मजदूरी (रु0)	→	20-30	30-40	40-50	50-60	60-70
श्रमिकों की संख्या	→	3	5	20	10	5

हल:

	→	मजदूरी	श्रमिकों की संख्या	संचयी आवृत्ति
			f	c.f
		20-30	3	3
		30-40	5	8
Md	→	40-50	20	28
		50-60	10	38
		60-70	5	43 = N

यहाँ $N/2 = 21.5$ तथा संबंधित संचयी बारम्बारता 28 अतः 40-50 मध्यका वर्ग अन्तराल सूत्र का प्रयोग करते हुएः

मध्यका $= 40 + \frac{10}{20}(21.5 - 8) = 40+6.75=46.75$

अतः मजदूरी की मध्यका = 46.5

स्मरणीय बिन्दुः

1. यदि श्रेणी को संचयी आवृत्ति वितरण के रुप में प्रस्तुत किया गया है तो इसे पहले सामान्य श्रेणी बना लेना चाहिए।

2. यदि वर्ग अवरोधी क्रम में है तो मध्यका का सूत्र होगा-

$$M = L_2 - \frac{\left(\frac{N}{2} - c\right)}{f} \times i$$

यहाँ L = मध्यका वर्ग की ऊपरी सीमा

C= संचयी आवृत्ति मध्यका वर्ग से पहले वाले वर्ग की होती है।

उदाहरण 16 - निम्न वितरण का मध्यका मूल्य ज्ञात कीजिए ।

	प्राप्तांक	आवृत्ति	संचयी आवृत्ति
		(f)	(c.f)
	70-80	10	100
	60-70	10	90
Md →	50-60	20	80
C.I.	30-40	15	30 = c
	20-30	15	15

यहाँ संचयी आवृत्ति नीचे से बनाई गई है, क्योंकि अवरोही वर्गान्तर घटते क्रम में है अतः मूल सूत्र का प्रयोग करते हुए - $N/2 = 50$ स 0 जो, 40-50 वर्ग अन्तराल में स्थित है।

$$M = L_1 + \left(\frac{\frac{N}{2} - C}{f}\right) i$$

$$= \frac{40 + \frac{100}{2} - 30}{30} \times 10$$

$$= 40 + \frac{50-30}{30}$$

$$= 40+6.67 = 46.67$$

यदि वर्ग अन्तराल बढ़ते क्रम का होता जैसे 70 से अधिक तो संशोधित सूत्र का प्रयोग किया जाएगा।

3. आसमान वर्गान्तरों की स्थिति में मध्यका ज्ञात करने के लिये उन्हें यथा सम्भव समान वर्गान्तरों में बदल लेना चाहिए, श्रेणी के अधिकतम वर्ग विस्तार के आधार पर पुर्नगठन करना चाहिए।
4. प्रथम वर्ग ही मध्यका वर्ग हो तो ब का मान शून्य होगा, शेष क्रियाएँ यथावत होगी।
5. खुले वर्ग अन्तराल से मध्यका प्रभावित नही होती

मध्यका की विशेषताएँ (Features of Median)-

गुण (Merits)

1. इसकी गणना सरल होती है।
2. चरम मूल्यों से यह प्रभावित नहीं होती।
3. नए पदों से मध्यका पर प्रभाव न्यूनतम होता है।
4. इसका बिन्दुरेखीय निरुपण भी सम्भव है- संचयी आवृत्ति वक्र (ogive curve) द्वारा मध्यका का निर्धारण कर सकते हैं।
5. मध्यका वर्ग की जानकारी से ही मध्यका की गणना की जा सकती है।

दोष (Demerits)

1. यह बीजगणितीय विवेचन के लिये अनुपयुक्त है-
2. सीमान्त मूल्यों को भार देने में यह अनुपयुक्त है
3. अविच्छिन्न श्रेणी में मध्यका का सूत्र इस मान्यता पर आधारित है कि प्रत्येक वर्ब में आवृत्तियाँ समान रुप से वितरित हैं। लेकिन यह अव्यवहारिक है।

विभाजन मूल्य -

मध्यका पूरे वितरण को दो समान भागों में विभाजित करती है, परन्तु मध्यका के सिद्धान्त पर पूरे वितरण चार, दस या सौ भागों में बाँटा जा सकता है यह निम्न है-

	खण्डित श्रेणी	अविच्छिन्न श्रेणी
a) चतुर्थक	- $Q_1 = \frac{N+1}{4}$ वाँ पद	$\frac{N}{4}$ वाँ पद
(चार हिस्से)	- $Q_3 = \frac{3(N+1)}{4}$ वाँ पद	$\frac{3N}{4}$ वाँ पद
b) दशमक	- $D_1 \rightarrow \frac{N+1}{10}$;	$\frac{N}{10}$ वाँ पद का मान
(10 हिस्से)	- $D_2 \rightarrow \frac{2(N+1)}{10}$;	$\frac{2N}{10}$ वाँ पद का मान
c) शतमक	$\rightarrow P_1 \rightarrow \frac{N+1}{100}$;	$\frac{N}{100}$ वाँ पद का मान
(100 हिस्से)	$P_{80} \rightarrow \frac{80(N+1)}{100}$;	$\frac{80N}{100}$ वाँ पद का मान

शेष गणना क्रिया एवं सूत्र मध्यका का ही होता है।

बहुलक (Mode)

परिभाषा : वह मूल्य है जो श्रेणी में सबसे अधिक बार आता हो या जिसकी आवृत्ति सबसे अधिक हो। यह मूल्यों के अधिकतम संकेद्रण का बिन्दु कहलाता है।

व्यक्तिगत श्रेणी में बहुलक

उदाहरण 17:

x	→	46	47	48	49	50	51	52
f	→	2	4	11	23	10	3	2

हल: यहाँ आवृत्तियाँ नियमित है अतः अधिकतम आवृत्ति (23) के समान वाला मूल्य अर्थात् 49 बहुलक है

खण्डित श्रेणी में बहुलक -

समूहन रीति द्वारा - निरीक्षण द्वारा हम बहुलक का निर्धारण करते हैं जब आवृत्तियाँ नियमित हों, परन्तु जब आवृत्तियाँ अनियमित हों तथा अधिकतम आवृत्ति केन्द्र में न होकर अन्त/आरम्भ में हो तो समूहन रीति अपनायी जाती है।

इसमें 6 स्तम्भ बनाये जाते हैं जिनमें

प्रथम स्तम्भ → आवृत्तियाँ होती हैं।

द्वितीय स्तम्भ → आरम्भ से दो-दो आवृत्तियों को जोड़ लिया जाता है।

तृतीय स्तम्भ → स्तम्भ 1 की पहली आवृत्ति छोड़कर दो-दो आवृत्तियों का जोड़ होता है।

चतुर्थ स्तम्भ → स्तम्भ 1 की तीन तीन आवृत्तियों को जोड़ होता है।

पंचम स्तम्भ→ स्तम्भ 1 की पहली आवृत्ति छोड़कर तीन-तीन आवृत्तियों को जोड़ होता है।

षष्ठम स्तम्भ → स्तम्भ 1 की पहली आवृत्ति छोड़कर तीन-तीन का जोड़ होता है।

अविच्छिन्न श्रेणी में बहुलक -

$$Z = L_1 + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

जहाँ Z= बहुलक

L_1 = बहुलक वर्ग की न्यूनतम वर्ग

f_0 = बहुलक वर्ग से पहले की आवृत्ति

f_2 = बहुलक वर्ग के बाद की आवृत्ति

f_1 = बहुलक वर्ग की आवृत्ति

i = वर्ग विस्तार

इस सूत्र की यह मान्यता है, कि बहुलक का मूल्य बहुलक वर्ग के समीप के वर्गों की आवृत्तियों से प्रभावित होता है। इसे निम्न उदाहरण से स्पष्ट किया जा सकता है-

उदाहरण 18: निम्न विवरण का बहुलक ज्ञात कीजिए-

x	-	1	2	3	4	5	6	7	8	9	10	11	12
f	-	3	8	15	23	35	40	32	28	20	45	14	6

X	(i)	(ii)	(iii)	(iv)	(v)	(vi)
1	3	11				
2	8		23	26		

3	15	38			46	
4	23		58			73
5	35	75		98		
6	40		72			
7	32	60				
8	28		48	80		
9	20	65			93	
10	45		59			79
11	14	20		65		
12	6					

LrEHk & la[;k	vf/kdre vko`fRr	x dk eku
(i)	45	10
(ii)	75	5,6
(iii)	72	6,7
(iv)	98	4,5,6
(v)	107	5,6,7
(vi)	100	6,7,8

यहाँ x के 6 वाँ का मान सर्वाधिक बार आया है अतः बहुलक का मान 6 होगा ।

बहुलक की गणना में महत्वपूर्ण तथ्य (Important Facts in Calculating Mode)-

1. यदि वर्गान्तर समावेषी है, तो इसे अपवर्जी में बदल देना चाहिए
2. यदि वर्गान्तर अवरोही क्रम में हो तो बहुलक की गणना ऊपरी सीमा से करनी चाहिए।

$$Z = \left(L_2 - \frac{f_1 - f_2}{2f - f_0 - f_2} \right) i$$

समान्तर माध्य और मध्यका द्वारा बहुलक की गणना

बहुलक = 3 मध्यका - 2 माध्य

बहुलक की विशेषताएँ (Characteristics of Mode) -

गुण (Merits):

1. यह सरल एवं लोकप्रिय है
2. निरीक्षण मात्र से निर्धारण सम्भव
3. चरम मूल्यों का न्यूनतम प्रभाव
4. विवरण का प्रतिरुपी मूल्य होता है
5. सर्वोत्तम प्रतिनिधि

दोष (Demerits)

1. यह बहुत अनिश्चित एवं अस्पष्ट है
2. चरम मूल्यों का महत्व नही
3. पदों की संख्या कम हो तो सार्थक नही
4. वर्ग विस्तार में परिवर्तन से मूल्य परिवर्तन

माध्य के चुनाव

1. उद्देश्य जिसके लिये माध्य का उपयोग करना है
2. समको की प्रकृति एवं विशेषताएँ तथा माध्य का कार्य

उपयोग -

1. समान्तर माध्य

→ इसका प्रयोग सार्वभौमिक होता है

→ आदर्श माध्य

2. गुणोत्तर माध्य

→ सापेक्ष परिवर्तन जैसे जनसंख्या वृद्धि, अनुपात, चक्रवात दरों के अध्ययन में सहायक

3. हरात्मक माध्य -

→ काल श्रेणी के अध्ययन में विषेष महत्व है

→ गति, मात्रा के रुप में प्रदत्त भाव में उपयोगी

4. मध्यका

→ गुणात्मक समंको का अध्ययन में सहायक

5. बहुलक

→ अपूर्ण समंको में गणना संभव। प्रति व्यक्ति उत्पादन, जूतों के माडल साइज मेंयह उपयोगी है।

माध्य की सीमाएँ (Limits of Mean)-

विभिन्न माध्य केवल केन्द्रीय प्रवतृत्ति को मापते हैं, समंको की प्रवृत्ति के घटते बढ़ने की जानकारी संभव नही होती। माध्य विषमता की माप नही कर सकते अतः विषमता का अध्ययन विवरण में आवश्यक होता है।

सांख्यिकीय माध्य कुछ विद्वानों का मत-

Croxten and Cowden के अनुसार,***"माध्य, समंकों के विस्तार के अन्तर्गत स्थिर ऐसा मूल्य है जिसका प्रयोग श्रेणी के सभी मूल्यों का प्रतिनिधित्व करने के लिये किया जाता है। समंकमाला के विस्तार के मध्य मे स्थित होने के कारण माध्य को केन्द्रीय मूल्य का माप भी कहा जाता है। "***

A. L. Bowley के अनुसार *" **माध्य वे सांख्यिकीय अचर हैं जो हमें सम्पूर्ण की सार्थकता का एक ही प्रयास में समझने की योग्यता प्रदान करते हैं।**"*

11.6 सारांश (Summary)

समान्तर माध्य वह मूल्य है जो किसी समंकमाला के सभी पदों के मूल्यों के योग में उन पदों की संख्या से भाग देने पर प्राप्त होता है। समान्तर माध्य की गणना दो प्रमुख विधियों द्वारा की जाती है- i) प्रत्यक्ष रीति (Direct Method) ii) अप्रत्यक्ष रीति (Indirect method) दोनों रीतियों से निकाला गया समान्तर माध्य बराबर ही आता है। अप्रत्यक्ष रीति अथवा लघु रीति का प्रयोग गणना की क्रिया को सरल बनाने के लिए किया जाता है। यदि $\bar{X} = (\sum X)/N$ में $\sum X, N, \bar{X}$ में से कोई दो की माप ज्ञात हो तो तीसरे की गणना की जा

सकती है। एक आदर्श माध्य के रुप में समान्तर माध्य की निम्न विशेषताएँ पायी जाती हैं - यह 1. सदैव निष्चयात्मक होता है। 2. गणना करने में सरल तथा सामान्य होता है। 3. यह सभी मूल्यों पर आधारित होता है। गुणोत्तर माध्य- किसी श्रेणी के n पदों के गुणनफल को n वाँ मूल (root √) ही उसका गुणोत्तर माध्य है। बहुलक वह मूल्य है जो श्रेणी में सबसे अधिक बार आता हो या जिसकी आवृत्ति सबसे अधिक हो। यह मूल्यों के अधिकतम संकेद्रण का बिन्दु कहलाता है। मध्यका किसी समंक श्रेणी में मध्यका वह मूल्य होता है जो पुरे श्रेणी को दो बराबर भागों में विभाजित करता है, जिसमें श्रेणी को घटते से बढ़ते क्रम में व्यवस्थित करते हैं तो उसका मध्य मूल्य मध्यका होता है।

11.7 अभ्यास के लिये प्रश्न (Practice Questions)

बहुविकल्पीय प्रश्न :

(1) निम्नलिखित में से कौन स्थिति सम्बन्धी माध्य हैं?

i) मध्यका ii) समान्तर माध्य

iii) गुणोत्तर माध्य iv) कोई भी नहीं

(2) निम्न में से कौन से सम्बन्ध सही हैं?

(i) A.M. = $\sqrt{G.M. \times H.M.}$

(ii) H. M. = $\sqrt{A.M. \times H.M.}$

(iii) G.M. = $\sqrt{A.M. \times H.M.}$

(iv) G.M. = $\frac{A.M. + H.M.}{2}$

(3) बहुलक ज्ञात करने की विधियों में से एक है -

(i) Z = 3M + $2\bar{X}$ (ii) Z = 2M – $3\bar{X}$

(iii) Z = 3M – $2\bar{X}$ (iv) Z = 3M – $3\bar{X}$

(4) निम्न से कौन सबसे अनिश्चित माध्य है?

i) बहुलक ii) मध्यका

iii) गुणोत्तर माध्य iv) हरात्मक माध्य

(5) निम्न में से कौन सा सम्बन्ध सही है?

(i) ($\bar{X}$ – Z) = 2/3 ($\bar{X}$ – Z)

(ii) ($\bar{X}$ – Z) = 1/3 ($\bar{X}$ - Z)

(iii) ($\bar{X}$ – Z) = 2/3 ($\bar{X}$ – M)

(iv) ($\bar{X}$ – M) = 1/3 ($\bar{X}$ – z)

उत्तर : 1. i) 2. iii) 3. iii) 4. i) 5. iv)

सही अथवा गलत चिन्हित करें :

(1) समान्तर माध्य से विभिन्न मूल्यों के विचलनों का बीजगणितीय योग शून्य होता है।
(2) बहुलक का बीजगणितीय विवेचन सम्भव है।
(3) एक आदर्श माध्य पर प्रतिचयन के परिवर्तनों का प्रभाव अधिकतम होना चाहिए।
(4) मध्यका चरम मूल्यों द्वारा प्रभावित होती है।
(5) बहुलक का बिन्दुरेखिय निर्धारण सम्भव है।
(6) एक आवृत्ति बंटन के मुख्य विशेषताओं को स्पष्ट करने में माध्य अकेले ही सक्षम है।

उत्तर : 1.सही 2.गलत 3. गलत 4. गलत 5. सही 6.गलत

11.9 सहायक / उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

1. सुदामा सिंह, ओ 0 पी0 सिंह, वाई 0 के0 सिंह (2002) – ***अर्थशास्त्रीय गणित एवं प्रारम्भिक सांख्यिकी*** - राधा पब्लिकेषन्स, नई दिल्ली।
2. एस 0 एन 0 लाल, एल 0 के0 चतुर्वेदी (2010) - **परिमाणात्मक विष्लेषण**, शिव पब्लिशिंग हाउस, इलाहाबाद।
3. J.K. Sharma (2008) – ***Business Statistics***, Dorling Kinderseley (India) Pvt. Ltd. (Pearson Education), Delhi.

11.10 निबंधात्मक प्रश्न (Essay Type Questions)

1. केन्द्रीय प्रवृत्ति की माप से आप क्या समझते हैं? एक अच्छे माध्य की विशेषताएँ क्या हैं? किस माध्य को आप आदर्श माध्य मानेंगे और क्यों ?
2. विभिन्न प्रकार के माध्यों का वर्णन कीजिए और उनकी सापेक्षिक विशेषताओं की व्याख्या कीजिए।
3. विभिन्न माध्यों के उपयोगों की चर्चा कीजिए।
4. यदि किसी समंक श्रेणी के सभी पदों का मूल बराबर हो तो सिद्ध कीजिए कि -

$$\bar{X} = G.M. = H.M$$

इकाई -12 अपकिरण की माप

(Measurement of Dispersion)

12.1 प्रस्तावना (Introduction)

केन्द्रीय प्रवृत्ति की माप, किसी श्रेणी के लक्षणों को सारांष-रूप में एक संख्या में व्यक्त करता है, परन्तु विभिन्न पदों के आपसी अन्तर या विभिन्न पदों के केन्द्रीय मूल्य से अन्तर को केन्द्रीय प्रवृत्ति की माप व्यक्त नहीं कर सकती समंकमाला की बनावट का पता माध्य द्वारा नही चल पाता। माध्य की शक्तिशाली माप बनाने के लिये हम समर्थक माप ज्ञात करतें हैं, जो श्रेणी में विचरण की माप करता है, इसे श्रेणी में विचरण की माप करता है, इसे अपकिरण (Dispersion) कहते हैं। केन्द्रीय प्रवृत्ति की माप को प्रथम क्रम का माध्य (Averages of first Order) कहते हैं। अपकिरण की माप को द्वितीय क्रम का माध्य कहते हैं।

अपकिरण की माप वास्तव में उस श्रेणी में संगतता या सिमरूपता के अभाव की सीमाओं को या परिमाण को मापता है। यदि पदों में अन्तर अधिक होगा तो अपकिरण का मान भी अधिक होगा। अपकिरण की माप की वही विशेषताएँ है जो आदर्श माध्य की होती हैं।

12.2 उद्देश्य (Objective)

प्रस्तुत इकाई का अध्ययन करने के पश्चात आप

- ✓ अपकिरण की विभिन्न मापों के बारे में जान सकेंगे।
- ✓ अपकिरण की विभिन्न मापों की विशेषताओं को जानने में सक्षम होंगे।
- ✓ अपकिरण को मापने की बिन्दुरेखीय विधि को समझ सकेंगे।

निरपेक्ष एवं सापेक्ष अपकिरण -

निरपेक्ष अपकिरण (Absolute dispersion) - जब अपकिरण की माप को मूल इकाइयों में ही व्यक्त किया जाए तो उसे निरपेक्ष अपकिरण कहा जाता है।

सापेक्ष अपकिरण (Relative Dispersion)- जब अपकिरण की माप को अनुपात या प्रतिशत में व्यक्त किया जाता है। यह विभिन्न प्रवृत्ति के समंकों की तुलना करने में अधिक उपयोगी है।

दो श्रेणियों की तुलना में निरपेक्ष अपकिरण नहीं प्रयुक्त किया जाता है। तुलना के लिये सदैव सापेक्ष अपकिरण का प्रयोग किया जाता है।

12.3 अपकिरण की माप एवं रीति (Measurement and Method of Dispersion)

क्रम संख्या	माप (Measure)		रीति (Method)
1.	विस्तार	Range	
2.	अन्तर - चतुर्थक विस्तार	Interquartile Range	सीमा रीति
3.	शतमक विस्तार	Percentile Range	
4.	चतुर्थक विचलन	Quartile Deviation	विचलन माध्य रीति
5.	माध्य विचलन	Mean Deviation	
6.	प्रमाप विचलन	Standard Deviation	

7.	लारेंज वक्र	Lorenz Curve	बिंदुरेखीय विधि

हम एक-एक कर सभी मापों की गणना रीति की चर्चा उदाहरणों सहित करेंगें।

12.3.1 विस्तार (Range)

किसी श्रेणी के अधिकतम और न्यूनतम मूल्यों के अन्तर को विस्तार (Range) कहते हैं। यह उस सीमा को व्यक्त करता है। जिनके, मध्य मूल्यों का अस्तित्व होता है। विस्तार को R से व्यक्त करते हैं। तथा इसका सूत्र निम्न है -

$$R = X_{max} - X_{min}$$ या $R = L - S$

वर्गीकृत श्रेणी में विस्तार की मान

R = उच्चवर्ग की ऊपरी सीमा - निम्न वर्गी की निचली सीमा

यह अपकिरण का निरपेक्ष माप है, सापेक्ष विस्तार या विस्तार गुणांक के लिए निम्न सूत्र का प्रयोग करना होता है

विस्तार गुणांक (Coefficient of Range) = $=\frac{L - S}{L + S}$

उदाहरण 1 निम्न श्रेणी की सापेक्ष एवं निरपेक्ष माप ज्ञात कीजिए

वर्गान्तर	5-10	10-15	15-20	20-25	25-30
आवृत्ति	2	5	12	8	3

हल .1 यहाँ उच्चतम वर्गान्तर = 25.30 जिसकी उच्चतम सीमा 30 है

न्यूनतम वर्गान्तर = 5 - 10 जिसकी न्यूनतम सीमा 5 है

निरपेक्ष विस्तार = L – S = 30 - 5 = 25

सापेक्ष विस्तार/विस्तार गुणांक $=\frac{L - S}{L + S} = \frac{30 - 5}{30 + 5} = \frac{25}{35} = \frac{5}{7}$

नोटः यदि वर्गान्तर का स्वरूप समावेषी हो तो विस्तार की गणना करते समय उसे उपवर्जी (exclusive) में बदल लेना चाहिए।

विस्तार के गुण एवं उपयोग

गुण (Merits)

1. यह सरलता से मापित किया जा सकता है।
2. इससे समय की बचत होती है।
3. यह उद्योगों में गुण नियंत्रण, विनिमय दरों में परिवर्तनशीलता की माप या मौसम भविष्यवाणी में प्रयुक्त होता है।

दोष (Demerits)

1. विस्तार केवल चरम मूल्यों पर आधारित होता है। यह चरम मूल्यों के बीच के मूल्य की जानकारी नहीं देता है।
2. इसमें स्थिरता नहीं होती, चरम मूल्य के एक पद में अन्तर होने से विस्तार प्रभावित हो जाता है।
3. आवृत्ति बंटनों के लिये विस्तार उपयुक्त नहीं है।
4. बीजगणितीय विवेचना के दृष्टिकोण से विस्तार अनुपयुक्त है।

अन्तर चतुर्थक विस्तार (Interquartile Range)

विस्तार (Range) आंकड़ों में विचरणशीलता की एक महत्वपूर्ण माप प्रस्तुत करता है, परन्तु इस माप का प्रमुख दोष यह है कि यह चरम मानों के द्वारा अत्यधिक प्रभावित होता है। चरम मानों के प्रभाव को निरस्त करने के लिये हम एक अन्य विधि अपनाते हैं, जिसे हम अर्द्ध अन्तर-चतुर्थक विस्तार (Interquartile Range) कहते हैं। इस विधि के अन्तर्गत हम तृतीय एवं प्रथम चतुर्थकों के आधे अन्तर को आँकड़ों की विचरणशीलता की माप के रूप में स्वीकार करते हैं। इस माप को इस प्रकार परिभाषित किया जाता है -

$$Q = \frac{Q_3 - Q_1}{2}$$

यद्यपि इस विधि के अन्तर्गत चरम मानों का प्रभाव निरस्त हो जाता है, परन्तु विस्तार की भाँति यह भी आँकड़ों के सभी मानों पर आधारित नहीं है। पूर्णतया समान आँकड़ों (Perfectly Symmetric Data) में मध्यिका का मान तृतीय चतुर्थक (Q_3) तथा प्रथम चतुर्थक (Q_1) के ठीक बीच में स्थित होता है। अत: ऐसी स्थिति में

$$Q_3 = M_d + Q$$
$$Q_1 = M_d + Q$$

अन्य शब्दों में पूर्णतया सममित वितरण में Q_3 तथा Q_1 माध्यिका से चतुर्थक विचलन की दूरी पर स्थित होंगे। परन्तु असममित वितरणों (Non-symmetrical Distribution) में Q_3 तथा Q_1 माध्यिका से चतुर्थक विचलन, दूरी पर स्थित नहीं होते। यह स्थिति बहुत सारे आर्थिक आँकड़ों में दृष्टिगोचर होती है- जो कि मूलतः असमामित होते हैं - जैसे आय का वितरण, भू-जलों का वितरण अथवा परिसम्पत्तियों का वितरण।

शतमक विस्तार (Percentile Range)

कभी-कभी अपकिरण ज्ञात करने के लिए शतमक विस्तार का भी प्रयोग किया जाता है। शतमक विस्तार P_{90} तथा P_{10} के अन्तर को व्यक्त करता है।

अपकिरण की यह रीति विस्तार (Range) तथा I.R. (Inter Quartile Range) से अपेक्षाकृत बेहतर है। क्योंकि एक तरफ तो यह चरम मूल्यों से प्रभावित नहीं होती और दूसरी तरफ श्रेणी के मध्य के 80 प्रतिशत मूल्यों पर आधारित है। इस रीति के दोष भी विस्तार तथा अन्तर चतुर्थक विस्तार के दोषों जैसे ही हैं।

12.3.2 चतुर्थक विचलन (Quartile Deviation)

यह विचलन एक प्रकार का विस्तार है जो चतुर्थकों पर आधारित है। हमने देखा है कि अन्तर चतुर्थक विस्तार उच्च चतुर्थक (Q_3) और निम्न चतुर्थक पर निर्भर होता है।

$$I.R. = Q_3 - Q_1$$

यदि हम को दो से भाग दे दें तो हमें चतुर्थक विचलन प्राप्त होता है।

$$Q.D = \frac{Q_3 - Q_1}{2}$$

यह अपकिरण का निरपेक्ष माप है। विभिन्न श्रेणियों में चतुर्थक विचलन की तुलना करने के लिए हम सापेक्ष माप ज्ञात करते हैं और इसे चतुर्थक विचलन गुणांक कहते हैं।

$$\text{चतुर्थक विचलन गुणांक} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

नोटः हम जानते हैं कि मध्यका पूरी श्रेणी को दो बराबर भागों में बांटती है। एक सममित विवरण में-

$$Q_3 - M = M - Q_1$$

या

$$2M = Q_1 + Q_3$$

$$M = \frac{Q_1 + Q_3}{2}$$

इस प्रकार

$$Q.D + Q_1 = \frac{Q_3 - Q_1}{2} + Q_1 = \frac{Q_1 + Q_3}{2} = M$$

और

$$Q_3 - Q.D = Q_3 - \frac{Q_3 - Q_1}{2} = \frac{Q_1 + Q_3}{2} = M$$

अत: $Q_1 = M - Q.D.$

और $Q_3 = M + Q.D$

इस प्रकार हम यह निष्कर्ष निकाल सकते हैं कि एक सममित वितरण में $M \pm Q.D$ समंक श्रेणी के 50% मूल्यों को व्यक्त करता है।

कुछ विद्वानों का मत है कि चतुर्थक विचलन वास्तव में एक स्थिति माध्य Positional Average) है और यह माध्य से पदों के विखराव को नहीं मापता। इसे अपकिरण की माप के बजाय विभाजन - माप (Measure of partition) कहना श्रेष्ठ है।

चतुर्थक विचलन को अर्द्ध अंतर- चतुर्थक विस्तार (Semi Inter quartile range) भी कहते हैं।

चतुर्थक विचलन के गुण (Merits of Quartile Deviation)

चतुर्थक विचलन विस्तार के कुछ दोषों के दूर करता है। इसके मुख्य गुण निम्न हैं-

1. इसकी गणना तथा इसे समझना दोनों सरल है।
2. यह श्रेणी के महत्वपूर्ण भाग (मध्य के 50 :) पर विचार करता है।
3. यह चरम मूल्यों (Extreme Value) से प्रभावित नहीं होता ।
4. यह खुले सिरे वाले वर्गान्तरों में भी अपकिरण की माप कर सकता है।

चतुर्थक विचलन के दोष (Demerits of Quartile Deviation)

1. यह श्रेणी के सभी मूल्यों पर आधारित नहीं है।
2. यह प्रतिचयन के परितर्वनों से बहुत प्रभावित होता है।
3. हसमें बीजगणितीय विवेचन का अभाव है।
4. यह अपकिरण का ठीक-ठाक माप नहीं करता केवल सन्निकट माप (Approximate Measures) करता है। वास्तव में यह सीमा रीति (Method of limit) द्वारा अपकिरण- माप के दोषों से मुक्त हो भी नहीं सकता।

विच्छिन्न श्रेणी में चतुर्थक विचलन :

उदाहरण 2 - दिये गये समंक में चतुर्थक विचलन ज्ञात कीजिए ।

चर मूल्य	(x)	:	6	7	8	9	10	11	12
आवृत्ति	(f)	:	3	6	9	13	8	5	4

हल-

(1) प्रस्तुत उदाहरण में एक विच्छिन्न आवृत्ति विवरण प्रदर्शित है। विच्छिन्न आवृत्ति विवरण के चतुर्थक ज्ञात करने के पूर्व, संचयी, आवृत्तियों को ज्ञात किया जाता है। यह प्रक्रिया निम्न सारणि में प्रदर्शित है।

x	f	cf
6	3	3
7	6	9

8	9	18
9	13	31
10	8	39
11	5	44
12	4	48

$Q_1 =$ वितरण के $\left(\frac{N+1}{4}\right)$ वें पद का मान

$=$ वितरण के $\left(\frac{48+1}{4}\right)$ वें पद का मान

$=$ वितरण के $\left(\frac{49}{4}\right)$ वें पद का मान

$=$ 12.25 वें पद का मान $= 8$

$Q_3 =$ वितरण के $\left(\frac{3N+1}{4}\right)$ वें पद का मान

$=$ वितरण के $\left(\frac{3(48+1)}{4}\right)$ वें पद का मान

$=$ वितरण के $\frac{3\times 49}{4}$ वे पद का मान

$= \frac{147}{4}$ वाँ मान $= 36.75$ वाँ मान $= 10$

$Q.D = \frac{Q_3 - Q_1}{2} = 10$

$Q.D = \frac{10-8}{2} = \frac{2}{2} = 1$

चतुर्थक विचलन गुणांक $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$

$= \frac{10-8}{10+8} = \frac{2}{18} = 0.11$

अविच्छिन्न श्रेणी में चतुर्थक विचलन :

उदाहरण 3: नीचे दिये गये समंक में चतुर्थक विचलन एवं चतुर्थक विचलन गुणांक ज्ञात कीजिए?

वर्गान्तर	10-20	20-30	30-40	40-50	50-60	60-70	70-80
आवृत्ति	5	10	10	25	25	15	10
संचयी	5	15	25	50	75	90	100
आवृत्ति	5	15	25	50	75	90	100

$Q_1 =$ वितरण के $\left(\frac{N}{4}\right)$ वें पद का मान

$=$ वितरण के $\left(\frac{100}{4}\right)$ वें पद का मान $=$ 25 वाँ पद

25 वाँ पद वर्गान्तर (30-40) में निहित है।

अतः $Q_1 = L_1 x \frac{i}{f}\left(\frac{N}{4} - C\right)$

$$Q_1 = 30 + \frac{10}{10}(25 - 15)$$

$$Q_1 = 30 + 1(10)$$

$$Q_1 = 40$$

$Q_3 = \frac{3N}{4}$ वें पद का मान

$=$ वितरण के $\frac{3(100)}{4}$ वें पद का मान

$=$ 75 वाँ पद का मान

75 वाँ पद वर्गान्तर (50-60) में निहित है -

अतः $Q_3 = L_1 + \frac{i}{f}\left(\frac{3N}{4} - C\right)$

$$Q_3 = 50 + \frac{10}{25}(75 - 50)$$

$$Q_3 = 50 + \frac{10}{25} \times 25$$

$$Q_3 = 50 + 10$$

$$Q_3 = 60$$

चतुर्थक विचलन $(Q.D.) = \frac{Q_3 - Q_1}{2} = \frac{60 - 40}{2} = \frac{20}{2} = 10$

चतुर्थक विचलन गुणांक $= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{60 - 40}{60 + 40} = \frac{20}{100}$

12.3.3 माध्य विचलन (Mean Deviation)

विचरणशीलता की अब तक चर्चित मापों - विस्तार तथा चतुर्थक विचलन, अथवा इन पर आधारित गुणांकों का प्रमुख दोष यह था कि समंक के केवल दो मूल्यों पर ही आधारित थे। यह समंक के अन्य मूल्यों की उपेक्षा करते हैं। विचरणशीलता की आदर्श माप समंक के सभी मूल्यों पर आधारित होनी चाहिये। इस प्रकार की एक माप है माध्य विचलन। 'माध्य विचलन' विचलनों के औसत को व्यक्त करता है। इन विचलनों की गणना समंक के किसी केन्द्रीय मान (माध्य, मध्यिका अथवा बहुलक) से की जाती है। सामान्य तौर पर माध्य विचलन की गणना के लिये प्रयुक्त विचलनों की गणना हम समंकों के माध्य से ही व्यक्त करते हैं।

माध्य विचलन की गणना की प्रक्रिया में इस प्रकार सर्वप्रथम हम समंकों के मूल्यों का माध्य से विचलन ज्ञात करते हैं - अर्थात् प्रत्येक समंक मूल्य में से माध्य के मान को घटाते हैं। तत्पश्चात विचलनों के चिन्हों की उपेक्षा करते हुए हम विचलनों का औसत ज्ञात करते हैं। विचलनों के बीजगणितीय चिन्हों की उपेक्षा करने का कारण यह हैं कि माध्य से विचलनों का योग सदैव शून्य के बाराबर होता है - क्योंकि विचलनों के योग की प्रक्रिया में धनात्मक तथा ऋणात्मक विचलन एक दूसरे को पूर्णतया निरस्त कर देते हैं।

चरराशि के मूल्यों का माध्य $= X - \overline{X}$ से विचलन

अब हम विचलनों के बीजगणितिय चिन्हों की उपेक्षा करते हुए, इनके निरपेक्ष मानो (Absolute values) अथवा परिमाणों को ज्ञात करते हैं। विचलनों के निरपेक्ष मानों को दो खड़ी लकीरों '||' के बीच रखकर व्यक्त किया जाता है, अर्थात्

विचलनों के निरपेक्ष मान $=\left|X - \overline{X}\right|$ अब $\left|X - \overline{X}\right|$ किसी राशि के परिमान अथवा निरपेक्ष मान को व्यक्त करता है तथा इसे हम Mod of $\left|X - \overline{X}\right|$ पढ़ते हैं।

माध्य विचलन = विचलनों का योग / विचलनों की संख्या $=\frac{\Sigma\left|X - \overline{X}\right|}{N}$

जहाँ $\Sigma\left|X - \overline{X}\right|$ = निरपेक्ष विचलनों का योग

N = आँकड़ों की संख्या।

यदि विचलनों को "$\overline{X}$" के द्वारा प्रदर्शित किया जाये जहाँ $dX = X - \overline{X}$ हो तो

$$MD = \frac{\Sigma\left|dX\right|}{N}$$

वर्गीकृत आँकड़ों के आकृति वितरण के लिये माध्य विचलन का सूत्र निम्न प्रकार होगा -

$$MD = \frac{\Sigma f.\left|d - \overline{X}\right|}{N}$$

जहाँ, $N = \Sigma f$

माध्य विचलन गुणांक (Coefficient of Mean Deviation)

माध्य विचलन अपकिरण का निरपेक्ष माप है जो विभिन्न श्रेणियों के लिये तुलना योग्य नहीं होता। इसके लिये हमें सापेक्ष माध्य विचलन या माध्य विचलन गुणांक का परिकलन करना पड़ता है। माध्य विचलन में जब हम संबंधित माध्य से भाग देते हैं तो हमें माध्य विचलन गुणांक प्राप्त होता है।

$=\frac{\delta\overline{X}}{X}$ या $=\frac{\delta M}{M}$ या $=\frac{\delta Z}{Z}$

(सामान्तर माध्य) (मध्यका) (बहुलक) से लिये गये विचलन गुणांक है।

व्यक्तिगत श्रेणी में माध्य विचलन

उदाहरण 4: श्रेणी 10, 14, 20, 24, 12 का समान्तर माध्य से माध्य विचलन तथा माध्य विचलन का गुणांक ज्ञात कीजिए।

हल - माध्य विचलन $\delta\overline{x} = \frac{\Sigma\left|x - \overline{x}\right|}{n} = \frac{\Sigma\left|d\overline{x}\right|}{n}$

x	$\left|x - \overline{x}\right| = \left|d\text{-}\overline{x}\right|$
10	6
14	2
20	4
24	8

12	4
80	24

$$\bar{x} = \sum x/\sim = 80/5 = 16$$

(∽ = n= 5)

$$\delta\,\bar{x} = |d - \bar{x}|\,/\sim = 24/5 = 4.8$$

माध्य विचलन गुणांक

$$\delta\,\bar{x}/\bar{x} = 4.8/16 = 0.3$$

खण्डित श्रेणी में माध्य विचलन (Mean Deviation in a Discrete Series)

खण्डित श्रेणी में माध्य विचलन $|d\bar{x}|$ या $|d_x|$ या $|d_z|$ ज्ञात करने के लिये पहले वह माध्य ज्ञात करते हैं जिससे विचलन लेना है फिर $d\,|\bar{x}|$ या $|dm|$ या $|d_z|$ ज्ञात करते है। आवृत्ति f से इसमें गुणा करके इसका योग कर लेते हैं। इस योग में आवृत्ति के योग से भाग देने पर माध्य विचलन प्राप्त होता है।

$$\delta\,\bar{x} = \frac{\sum f|d\bar{x}|}{N = \sum f}, \quad \delta_M = \frac{\sum f\,|dM|}{N}, \quad \delta_z = \frac{\sum f(dz)}{N}$$

मध्यका से माध्य विचलन की गणना -

X	→	5	6	7	8	9	10	11
1	→	2	1	3	6	14	3	1

x	f	c.f	मध्यिका 8 से विचलन	f|dMd|
5	2	2	3	6
6	1	3	2	2
7	3	6	1	3
8	6	12	0	0
9	4	16	1	4
10	3	19	2	6
11	1	20	3	3
	N = 20			$\sum f\,\|dMd\| = 24$

$Md = \left(\frac{N+1}{2}\right)$ वे पद का मान $= \frac{20+1}{2}$ वे पद का मान $= \left(\frac{21}{2}\right)$ वे पद का मान $= 10.5$ वें पद का मान.
संचयी आवृत्ति 10.5 व संचयी आवृत्ति 12 में सम्मिलित अर्थात् मध्यिका $Md = 8$

माध्य विचलन, $\delta\,Md = \frac{\sum f\,|Md|}{N}$

$$= \frac{24}{20} = 1.2$$

माध्य विचलन गुणांक $= \frac{\delta Md}{Md} = \frac{1.2}{8} = 0.15$

अविच्छिन्न श्रेणी में माध्य विचलन (Mean Deviation in a Continuous Series)

अविच्छिन्न श्रेणी में माध्य विचलन का खण्डित श्रेणी वाला सूत्र ही प्रयोग में लाया जाता है यहाँ x वर्ग अन्तराल का मध्य बिन्दु ज्ञात करतें हैं।

उदाहरण 5 - निम्न समंको में समान्तर माध्य, माध्यिका एवं बहुलक से माध्य विचलन एवं उनके गुणांक ज्ञात कीजिए -

वर्गान्तर	20-30	30-40	40-50	50-60	60-70	70-80	80-90
आवृत्ति	9	12	15	20	25	10	9

हल:

वर्गान्तर	माध्य बिन्दु	आवृत्ति	A – 65 से विचलन (dx)	dx'	fdx'	माध्य 55.6 से निरपेक्ष विचलन |dx'|	$f\|d\bar{x}\|$
20-30	25	9	-40	-4	-36	30.6	275.4
30-40	35	12	-30	-3	-36	20.6	247.2
40-50	45	15	-20	-2	-30	10.6	159.0
50-60	55	20	-10	-1	-20	0.6	12.0
60-70	65	25	0	0	0	9.4	235.0
70-80	75	10	+10	1	10	19.4	194.0
80-90	85	9	+20	2	18	29.4	264.6
		N = 100			fdx' = 94		$\sum f\|d\bar{x}\|=1387.2$

पहले हम समान्तर माध्य की गणना लघु रीति से करेंगे

$$\bar{X} = A + \frac{\sum fdx'}{N} \times i$$

$$= 65 + \frac{(-94)}{100} \times 10$$

$$= 65 - \frac{94}{10} = 65 - 94 = 55.6.$$

अब समान्तर माध्य 55.6 द्वारा हम माध्य विचलन ज्ञात करेंगे-

$$\delta\bar{X} = \frac{f|dx|}{N} = \frac{1387.2}{100}$$

$$= 13.872 \simeq 13.87$$

माध्य विचलन गुणांक $= \frac{\delta\bar{X}}{\bar{X}} = \frac{13.87}{55.6} = 0.25$ (उत्तर)

मध्यिका से माध्य विचलन (Mean Deviation from the Median) -

यहाँ हम उपरोक्त समंको को सूचना का प्रयोग करेंगे

$\sum$ (संचयी) आवृत्ति $= 100$

$\sum$ मध्यिका '57' से $= 1404$

विचलन $|dMd|$
का में गुणा

अब, $Md = \left(\frac{N}{2}\right)$ वें पद का मान $= \frac{100}{2} = 50$ वें पद का मान 50 वाँ पद $(50 - 60)$ वर्गान्तर में पड़ता है
अतः

$$Md = \ell + \frac{i}{f}\left(\frac{N}{2} - C\right)$$

$$= 50 + \frac{10}{20}(59 - 36)$$

$$Md = 50 + 7$$

$$Md = 57$$

माध्य विचलन मध्यिका से

$$\delta\ Md = \frac{\sum f|dMd|}{N}$$

$$\delta Md = \frac{1404}{100}$$

$$= 14.04$$

माध्य विचलन गुणांक $\frac{\delta Md}{Md} = \frac{14.04}{57} = 0.25$

बहुलक से माध्य विचलन (Mean Deviation from the Mode)

समंकमाला में बहुलक 60-70 वर्गान्तर में होगा क्योंकि इसकी आवृत्ति (25) सबसे अधिक है -

बहुलक

$$M_0 = 1 + i \times \frac{D_1}{D_1 + D_2}$$ जहाँ $D_1 = f_1 - f_0$

$$= 60 + 10\ \frac{5}{5 + 15}$$ $D_2 = 25 - 10$

$$= 62.5$$

माध्यविचलन - $\delta\ M_0 = \frac{\sum f|dM_0|}{N}$

$$= \frac{1470.0}{100} = 14.70$$

माध्य विचलन गुणांक $= \frac{\delta M_0}{M_0} = \frac{14.7}{62.5} = 0.24$ (उत्तर)

नोट: पाठकों से अनुरोध है कि वे स्वयं मध्यिका से विचलन एवं बहुलक से विचलन की सारणी ज्ञात करें।

उदाहरण 6 - निम्न सारणी की दो श्रेणियों में विचरण की तुलना माध्य विचलन गुणांक द्वारा कीजिए ।

A	70	100	50	130	140	150	90	60	110	600
B	1250	1350	1600	1450	1550	1700	1750	1800	1400	1650
A	176	124								
B	-	-								

हल: (1) सर्वप्रथम हम दोनो श्रेणियों के माध्य ज्ञात करते हैं।
(2) यदि माध्य समान हो तो माध्य विचलन की गणना करते हैं।
(3) यदि माध्य भिन्न हों तो माध्य विचलन गुणांक ज्ञात करेंगें।

श्रेणी A		
क्र स0	**आय (x)**	**|dx|**
1	70	80
2	100	50
3	50	100
4	130	20
5	140	10
6	150	0
7	90	60
8	60	90
9	110	40
10	600	450
11	176	26
12	124	26
	1800	952

$$d = \frac{1800}{12} = 150$$

माध्य विचलन गुणांक

(माध्य विचलन) $\delta \bar{x} = \frac{\sum |d\bar{x}|}{N}$

$$= \frac{952}{12} = 79.33$$

माध्य विचलन गुणांक $\frac{\delta \bar{x}}{\bar{x}} = \frac{79.33}{150} = 0.53$

श्रेणी B		
क्र स0	**आय (x)**	**|d $\bar{x}$|**
1	1250	300
2	1350	200
3	1600	50
4	1450	100
5	1550	0
6	1700	150
7	1750	200
8	1800	250
9	1400	150
10	1650	100

	15500	1500

$$d = \frac{15500}{10} = 1550$$

माध्य विचलन $= \delta \bar{x} = \frac{\sum |d\bar{x}|}{N}$

$$= \frac{1500}{10} = 150$$

माध्य विचलन गुणांक $= \frac{\delta \bar{x}}{\bar{x}}$

$$= \frac{150}{1550} = 0.10$$

उत्तर श्रेणी B में माध्य विचलन का गुणांक मान श्रेणी A से कम है अतः श्रेणी B में विवरण श्रेणी A की तुलना में अधिक समान है।

माध्य विचलन के गुण (Merits of Mean Deviation)-

1. माध्य विचलन की परिभाषा स्पष्ट है।
2. माध्य विचलन की गणना सरल है ।
3. यह सभी मूल्यों पर आधारित प्रवृत्ति हैं ।
4. यह चरम मानों से अपेक्षाकृत कम प्रभावित हैं ।
5. यह केन्द्रीय मान से आंकड़ो की औसत दूरी प्रदर्शित करता है।

माध्य विचलन के दोष(Demerits of Mean Deviation) –

1. यह विचलनों के गणितीय चिन्हों की उपेक्षा करता है ।
2. आंकड़ों में किसी अज्ञात पद के होने से इसकी गणना संभव नही है।
3. आंकड़ों की संख्या अधिक होने से इसकी गणना में कठिनाई होती है।
4. माध्य विचलन का महत्व सैद्धान्तिक है तथा व्यावहारिक सांख्यिकीय में विचरणषीलता की माप के लिए मानक विचलन का प्रयोग किया जाता है।

12.3.4 प्रमाप विचलन (Standard Deviation)

मानक विचलन के अंतर्गत, सर्वप्रथम हम समंक मूल्यों के माध्य से विचलन ज्ञात करते हैं। कुछ विचलन धनात्मक एवं कुछ ऋणात्मक होते हैं अतः हम विचलनों के वर्ग ज्ञात करते हैं। इस प्रकार विचलनों के चिन्हों की समस्या समाप्त हो जाती है। इसे सिग्मा (σ) द्वारा व्यक्त करते हैं। मानक विचलन की गणना निम्न प्रकार होती है -

- माध्य से विचलन $dx = x - \bar{x}$
- विचलनों का वर्ग $dx^2 = (x - \bar{x})^2$
- विचलनों वर्गों का योग $= \sum dx^2 = \sum (x-\bar{x})^2$
- विचलन वर्गों का औसत अथवा प्रसरण $= \frac{\sum (x-\bar{x})^2}{N}$
- मानक विचलन $= +\sqrt{\frac{\sum (x - \bar{x})^2}{N}}$

मानक विचलन सदैव धनात्मक होता है। यदि मानक विचलन दिया हो तो प्रसरण का मान निम्न प्रकार व्यक्त कर सकते हैं।

प्रमाप विचलन किसी समंक के मूल्यों का माध्य के दोनो ओर बिखराव का माप है।

प्रमाप विचलन को भिन्न नामों से जाना जाता है जैसे -

- ‘विभ्रम’ (Mean Error)
- माध्य विचलन वर्ग मूल्य (Mean Square Error)
- माध्य विचलन वर्ग मूल (Root Mean Square Deviation)

प्रमाप विचलन का प्रत्यक्ष सूत्र इस प्रकार है -

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{N}}$$

यदि माध्य का मान दशमलव बिन्दुओं में हो तो -

$$\sigma = \sqrt{\frac{\sum x^2}{N} = \left(\frac{\sum x}{N}\right)^2}$$

12.3.5 वितरण गुणांक (Coefficient of Variation)

$$\text{वितरण गुणांक} = \frac{\sigma}{\bar{x}} \times 100$$

अर्थात सापेक्षित विचरणषीलता ज्ञात करने के लिये निरपेक्ष (मानक विचलन) माप को केन्द्रीय मान (माध्य) से भाग देते हैं, यह प्रतिशत के रूप मे व्यक्त होता है, अतः इसे 100 से गुणा कर देते हैं। विचरणषीलता गुणांक का प्रयोग उस स्थिति में भी किया जाता है, जब दोनों श्रेणियों की माप की इकाइयाँ भिन्न भिन्न हों।

मानक माप ज्ञात करने की विधि (Methods of Measuring Standard Deviation)

सामान्य श्रेणी में प्रत्यक्ष रीति $\sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}$

$$\text{विचरण गुणांक} - \frac{\sigma}{\bar{x}}$$

उदाहरण 7: प्रमाप विचलन तथा विचरण गुणांक की गणना कीजिए

श्रमिक	-	1	2	3	4	5	6	7	8	9	10
मजदूरी (रू0)	-	80	85	95	90	100	75	65	105	70	85

हल: $\sigma = \sqrt{\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2}$

विचरण गुणांक $= \dfrac{\sigma}{\bar{x}}$

प्रमाप विचलन तथा विचरण गुणांक की गणना को निम्न सारणी द्वारा दिखाया गया है।

श्रमिक	मजदूरी (रू0)	X^2	माध्य 85 से $dx = (x-\bar{x})$	dx^2
1	80	6400	-5	25
2	85	7225	0	0
3	95	9025	10	100
4	90	8100	5	25
5	100	10000	15	225
6	75	5625	-10	100
7	65	4225	-20	400
8	105	11025	20	400
9	70	4900	-15	225
10	85	7225	0	0
N = 10	$\sum x = 850$	$\sum x^2 = 73750$		$\sum dx^2 = 1500$

$$\sigma = \sqrt{\frac{13750}{10} - \left(\frac{850}{10}\right)^2}$$

$$= \sqrt{7375 - (85)^2}$$

$$\sigma = \sqrt{7375 - 7225} = \sqrt{150} = 12.25$$ (उत्तर)

$$\bar{X} = \frac{\sum X}{N} = \frac{850}{10} = 85$$

$$C.V. = \frac{12.25}{85} = 0.144$$ (उत्तर)

2 अप्रत्यक्ष विधि - $\sigma = \sqrt{\dfrac{\sum d\,x^2}{N}}$

$$\sum dx^2 = 1500, \quad N = 10$$

तो $\sigma = \sqrt{\dfrac{1500}{10}} = 12.25$ (उत्तर)

विचलन गुणांक $= \dfrac{12.25}{85} = 0.144$ (उत्तर)

दोनों रीति से उत्तर समान है ।

विच्छिन्न श्रेणी का प्रमाप विचलन (Standard Deviation Of Discrete Series)

$$\sigma = \sqrt{\frac{\sum fdx^2}{N}} = \sqrt{\frac{\sum f(x-\bar{x})^2}{N}}$$

$\sum fdx^2$ = माध्य से विचलनों के वर्गों तथा संगत आवृत्तियों का गुणनफल

विचलन रीति (Deviation Method)

$$\sigma = \sqrt{\frac{\sum fdx^2}{N} - \left(\frac{\sum fdx}{N}\right)^2}$$

dX = कल्पित माध्य से विचलन

A = सर्वाधिक आवृत्ति वाला चर को कल्पित माध्य मान लेते हैं।

पद विचलन रीति (Step Deviation Method)

$$dX = X - A$$

$$dX' = \frac{dX}{i} = \frac{X - A}{i}$$

$$\sigma = \sqrt{\frac{\sum f.dX'^2}{N} - \left(\frac{\sum f.dx'}{N}\right)^2} \times i$$

$\sum f.dx^2$ = पद विचलनों के वर्गों तथा संगत आवृत्तियों के गुणनफल का योग

= पद विचलनों तथा संगत आवृत्तियों के गुणनफल का योग

अविच्छिन्न श्रेणी में प्रमाप विचलन (Standard Deviation in a Continuous Series)

$$\sigma = \sqrt{\frac{\sum fdx^2}{N} - \left(\frac{\sum fdx}{N}\right)^2}$$

जहाँ $\sum fdx^2$ = माध्य बिन्दु के वर्ग एवं संगत आवृत्तियाँ के गुणनफल का योग

$\left(\frac{\sum fdx}{N}\right)^2$ = मध्य बिन्दु की संगत आवृत्ति से गुणफल के योग को N से भाग करने पर प्राप्त मान का वर्ग

उदाहरण 8: निम्नलिखित समंक का प्रमाप विचलन एवं प्रमाप विचलन गुणांक कीजिए

चर मूल्य	(x)	:	6	7	8	9	10	11	12
आवृत्तियाँ	(f)	:	3	6	9	13	8	5	4

हल:

x	f	fx	dx	dx^2	f dx^2
6	3	18	-3	9	27
7	6	42	-2	4	24

8	9	72	-1	1	9
9	13	117	0	0	0
10	8	80	1	1	8
11	5	55	2	4	20
12	4	48	3	9	36
	N = 48	432			124

$$\bar{X} = \frac{432}{48} = 9$$

$$\sigma = \sqrt{\frac{\sum fdx^2}{N}} = \sqrt{\frac{124}{48}} = \sqrt{2.58} = 1.60$$

विचलन गुणांक $= \frac{\sigma}{\bar{X}} = \frac{1.6}{9} = 0.18$

उदाहरण 9: निम्नलिखित समंको का पद विचलन रीति से मानक विचलन ज्ञात कीजिए ।

वर्गान्तर	मध्य बिन्दु	आवृत्ति	काल्पित माध्य से विचलन		fdx'	f(dx'²)
	(X)	(f)	dx	dx'		
10-20	15	5	-30	-3	-15	45
20-30	25	10	-20	-2	-20	40
30-40	35	10	-10	-1	-10	10
40-50	45	25	0	0	0	0
50-60	55	25	10	1	25	25
60-70	65	15	20	2	30	60
70-80	75	10	30	3	30	90
		N = 100			$\sum fdx' = 40$	$\sum fdx'^2 = 270$

$$\sigma = \sqrt{\sum fdx'^2 - \left(\frac{\sum fdx'}{N}\right)^2} \times i = \sqrt{\frac{270}{100} - \frac{40}{100}} \times 10$$

$$\sigma = \sqrt{2.7 - (0.4)^2} \times 10 = \sqrt{2.7 - 0.16} \times 10$$

$$\sigma = \sqrt{2.54} \times 10 = 1.593 \times 10 = 15.93$$

प्रमाप विचलन गुणांक $= \frac{\sigma}{\bar{X}}$

जहाँ $\bar{X} = A + \frac{\sum dx'}{N} \times i$

$$= 45 + \frac{40}{100} \times 10$$

$$\bar{X} = 45 + 4 = 49$$

तो $\frac{\sigma}{\bar{X}} = \frac{15.93}{49} = 0.33$ उत्तर

उदाहरण 10 दो कारखानों A तथा B में, जो एक ही उद्योग से सम्बन्धित हैं, औसत साप्ताहिक मजदूरी तथा प्रमाप विचलन इस प्रकार दिये हैं-

कारखाना	माध्य (रू0)	प्रमाप विचलन
A	40	12.5
B	36.4	9.1

दोनो कारखानों की मजदूरी में किसमें संगतता (Compatibility) अधिक है।

हल - हम दोनो कारखानों के विचलन गुणांक ज्ञात करेंगें।

$$A = C.V = \frac{\sigma_1}{X_1} \times 100$$

$$\sigma_1 = 12.5$$

$$x_1 = 40$$

अत :
$$C.V. = \frac{12.5}{40} \times 100 = 31.25\%$$

$$B \rightarrow C.V = \frac{\sigma_2}{x_2} \times 100$$

$$\sigma_2 = 9.1$$

$$x_2 = 36.4$$

अतः
$$= \frac{9.1}{36.4} \times 100$$

$$C.V. = 25\%$$

जहाँ B कारखाने का विचलन गुणांक कम है, अतः अधिक मजदूरी संगतता B में हैं।

सामूहिक प्रमाप विचलन (Combined Standard Deviation) -

जिस प्रकार हम एक से अधिक विवरणों के माध्यों के आधार पर सामूहिक माध्य ज्ञात करते है उसी प्रकार हम एक से अधिक विवरणों के माध्यों एवं प्रमाप विचलन के आधार पर सम्मिलित विवरण का प्रमाप विचलन ज्ञात करते हैं - इसकी विधि निम्न है-

1. विवरणों का सामूहिक माध्य ज्ञात करते हैं।
2. प्रत्येक विवरण माध्य का सामूहिक माध्य से विचलन ज्ञात किया जाता है।

   जैसे $D_1 = \bar{X}_1 - \bar{X}$, $D_2 = \bar{X}_2 - \bar{X}$, $D_n = \bar{X}_n - \bar{X}$.
3. अब निम्न सूत्र के द्वारा हम सामूहिक विवरण का प्रमाप विचलन (σ) ज्ञात करते हैं।

$$\sigma = \sqrt{\frac{N_1(\sigma_1^2 + D_1^2) + N_2(\sigma_2^2 + D_2^2) + \ldots\ldots + N_2(\sigma_n^2 + D_n^2)}{N_1 + N_2 + \ldots\ldots + N_n}}$$

उदाहरण 11: दो आवृत्ति वितरणों की कुल आवृत्ति, माध्य एवं प्रमाप विचलनों का वितरण इस प्रकार है -

$N_1 = 200$ $\quad \bar{X}_1 = 25$ $\quad \sigma_1 = 3$

$N_2 = 300$ $\quad \bar{X}_2 = 10$ $\quad \sigma_2 = 4$

हलः वितरणों का सामूहिक माध्य $\overline{X}$ =

$$\overline{X} = \frac{N_1 - \overline{X}_1 + N_2 \overline{X}_2}{N_1 + N_2} = \frac{200(25) + 300(10)}{200 + 300}$$

$$= \frac{8000}{500} = 16$$

सामूहिक प्रमाप विचलन

$$(\sigma) = \frac{\sqrt{N_2\left(\sigma_1^2 + D_1^2 + N_2\left(\sigma_2^2 + D_2^2\right)\right)}}{N_1 + N_2}$$

प्रश्न द्वारा प्रदत्त सूचना के अनुसार -

$N_1 = 200$ $\sigma_1 = 3$ $D_1 = \overline{X}_1 - \overline{X} = 9$

$N_2 = 300$ $\sigma_2 = 4$ $D_2 = \overline{X}_2 - \overline{X} = -6$

सूत्र में मान रखने पर

$$\sigma = \sqrt{\frac{200(3^2 + 9^2) + 300(4^2 + (-6)^2}{200 + 300}}$$

$$\sigma = \sqrt{\frac{200(9 + 81) + 300(16 + 36)}{500}}$$

$$\sigma = \sqrt{\frac{200(90) + 300(52)}{500}}$$

$$= \sqrt{\frac{18000 + 15,600}{500}} = \sqrt{\frac{336}{5}} = \sqrt{67.2} = 8.2$$

सामूहिक विचलन प्रमाप = 8.2

प्रमाप विचलन के गुण एवं दोष (Merits and Demerits of Standard Deviation)

गुण (Merits)

i) इसका मान निश्चित है तथा यह परिभाषित है।
ii) इसका मान समंकमाला के सभी मूल्यों पर आधारित है।
iii) गणितीय क्रियाओं के लिये सर्वथा उपयुक्त है।
iv) यह विचलन के वर्गों पर आधारित होता है।
v) प्रमाप विचलन का मान न्यादर्श परिवर्तन के फलस्वरूप अधिक परिवर्तित नहीं होता।
vi) इसकी संकल्पना सहसम्बन्ध तथा समाश्रयण में भी अत्यन्त उपयोगी है।

दोष (Demerits)

i) यह गणना करने में कठिन है।
ii) यह बड़े विचलनों को छोटे विचलनों से अधिक महत्व देता है।

12.4 बिन्दु रेखीय विधि- लारेंज वक्र (Graphical Method- Lorentz Curve)

बिन्दु रेखीय विधि से अपकिरण मापने का सर्वप्रथम प्रयोग **डॉ0 मैक्य ओ. लारेंज (Max O. Lorenz)** ने किया था। यह समान वितरण रेखा से समंक माला के असमानताओं का अध्ययन करने में सहायक

है। इसका प्रयोग भू जोतों, आय एवं सम्पत्ति तथा परिसम्पत्तियों की असमानता का अध्ययन करने में सहायक होता है। लारेंज वक्र समान वितरण रेखा से वास्तविक विवरण के विचलन का माप है। यह वक्र समान विवरण से जितना दूर होगा असमानताओं उतनी अधिक होंगी।

ग्राफ ????

यहाँ A ग्रुप में वितरण आय की असमानाएँ अधिक हैं।

आय की असमानताओं को परिमाणात्मक रूप में मापने के लिये हम एक गुणांक का प्रयोग करते है जिसे '**संकेन्द्रण अनुपात**' अथवा '**Gini Coefficient'** कहते हैं।

गिनी गुणांक = समान विवरण रेखा लानेंज वक्र के मध्य क्षेत्रफल / समान विवरण रेखा व अक्षों के बीच क्षेत्रफल

इस गुणांक का मान जितना कम होगा आय की असमानताएँ भी उतनी कम होंगी तथा अधिक होने पर असमानताएँ अधिक होंगी।

12.5 सारांश (Summary)

श्रेणी में विचरण की माप ज्ञात करने वाले माध्य को अपकिरण कहतें हैं। अपकिरण को द्वितीय क्रम का माध्य भी कहते हैं। अपकिरण के माप की वहीं विशेषताएँ होती हैं, जो आदर्श माध्य की होती हैं। जब अपकिरण की माप को मूल इकाइयों में ही व्यक्त किया जाता है तो उसे निरपेक्ष अपकिरण कहतें हैं तथा जब इसे अनुपात या प्रतिशत में व्यक्त किया जाता है तो इसे सापेक्ष अपकिरण कहते हैं। दो श्रेणियों की तुलना करने के लिये सापेक्ष अपकिरण का उपयोग किया जाता है। अपकिरण को तीन रीति से मापा जा सकता है- **सीमा रीति**- विस्तार, अन्तर-चतुर्थक विस्तार, शतमक विस्तार, **विचलन माध्य रीति** - चतुर्थक विचलन, माध्य विचलन, प्रमाप विचलन, **बिन्दुरेखीय विधि-** लोरेन्ज वक्र। विस्तार, अपकिरण की निरपेक्ष माप है जो श्रेणी के अधिकतम एवं न्यूनतम मूल्यों के अन्तर से प्राप्त होता है। विस्तार गुणांक द्वारा हमें अपकिरण की सापेक्ष माप प्राप्त होती है। विस्तार का प्रमुख उपयोग गुण नियंत्रण, विनिमय दरों में परिवर्तन, मौसम भविष्यवाणी में प्रयुक्त। चरम मानों से अत्यधिक प्रभावित होने के कारण विस्तार विचलन की सही माप नही दे पाता, अतः चरम मानों के प्रभाव निरस्त करने के लिए हम अर्द्ध अन्तर चतुर्थक विस्तार का प्रयोग करते हैं। पूर्णतया सभावित आँकड़ों में मध्यिका का मान तृतीय चतुर्थक तथा प्रथम चतुर्थक के ठीक बीच में स्थित होता है।

भूजोत, आय तथा सम्पत्ति विवरण में असमित विवरण की माप में अन्तर चतुर्थक विचलन उपयोगी होता है। श्रेणी के चरम मूल्यों से प्रभावित न होकर 8% मूल्यों पर आधारित अपकिरण की बेहतर माप शतमक विस्तार है। चतुर्थक विचलन अन्तर चतुर्थक विचलन को 2 में से भाग देने पर प्राप्त होता है। चतुर्थक विचलन गुणांक का प्रयोग हम विभिन्न श्रेणियों में विचलन ज्ञात करने के लिए करते है चतुर्थक विचलन को स्थिति माध्य कहना ज्यादा उचित है। माध्य विचलन विचलनों के औसत व्यक्त करता है इसकी गणना समंक के किसी केन्द्रीय माध्य (समान्तर माध्य, माधिका बहुलक) द्वारा की जाती है। सापेक्ष अपकिरण के लिए माध्य विचलन गुणांक की गणना की जाती है। प्रमाप विचलन या मानक विचलन, समंकमाला के सभी मूल्यों पर आधारित होता है। मानक विचलन का माप सदैव धनात्मक होता है, इसकी सापेक्ष माप को विचरण गुणांक कहते हैं। समान विवरण रेखा से समंकमालाओं की असमानता का बिन्दुरेखीय विधि से अध्ययन लारेंज वक्र द्वारा किया जाता है। आय की असमानताओं के परिमाणात्मक माप के लिये संक्रेन्द्रण अनुपात (Gini Coefficient) का प्रयोग किया जाता है।

12.6 अभ्यास के लिये प्रश्न (Practice Questions)

प्रश्न: 1 A तथा B दो कम्पनियों के अंशपत्रों (Shares) के मूल्य सम्बन्धी आँकड़ो नीचे दिये गये हैं।

X	55	51	52	53	56	58	52	50	51	49
Y	108	107	105	105	106	107	104	103	104	101

प्रश्नः 2 एक कक्षा के विद्यार्थियों के प्राप्तांक निम्न है -

32, 15, 20, 20, 21, 22, 24, 24, 35, 33, 28, 30, 29

आंकड़ो के आधार पर i) चतुर्थक विचलन ii) चतुर्थक विचलन गुणांक ज्ञात कीजिए ।

प्रश्न: 3 निम्न आंकड़ों के आधार पर मध्यिका विचलन गुणांक, माध्य गुणांक की गणना कीजिए -

वस्तु का आकार	(x)	→	4	6	8	10	12	14	16
आवृत्ति	(f)	→	2	4	5	3	2	1	4

प्रश्न: 4 निम्न आवृत्ति विवरण के i) माध्य ii) प्रमाप विचलन की गणना कीजिए ।

x	10	20	30	40	50	60	70	80
y	15	30	53	75	100	110	115	125

प्रश्न: 5 निम्न समंको का प्रमाप विचलन ज्ञात कीजिए -

x	1	2	3	4	5	6	7	8	9
y	2	60	101	152	205	155	79	40	1

प्रश्न: 6 100 मदों से सम्बद्ध एक विवरण का माध्य '50' तथा प्रमाप विचलन '4' है, इन मदों के वर्गों का योग ज्ञात कीजिए।

प्रश्नः 7 निम्न समंकों द्वारा माध्य और माध्य विचलन ज्ञात कीजिए ।

x	3-4	4-5	5-6	6-7	7-8	8-9	9-10
y	3	7	22	60	85	32	8

प्रश्न 8: निम्न आंकड़ों के माध्य और प्रमाप विचलन ज्ञात कीजिए -

x	10	20	30	40	50	60	70	80
y	15	30	53	75	100	110	115	125

प्रश्नः 9 निम्न आंकड़ों से लारेंज वक्र खीजिए -

आय का विस्तार	परिवार संख्या	कुल आय
51-50	1744	268
251-450	557	186
451-650	302	166
51-850	186	139
851-1050	123	117
1051-1250	90	104

प्रश्नः 10 एक विवरण की निम्न सूचना उपलब्ध है-

$$N = 100 \quad \bar{X} = 40 \quad \sigma = 0.51$$

परन्तु बाद में पता चलात है कि एक मूल्य को के स्थान पर ले लिया था अब इस विवरण की सही माध्य एवं प्रभाव विचलन ज्ञात कीजिए

12.7 अभ्यास प्रश्नों के उत्तर (Answer of Practice Questions)

1. $c.v.(x) = 4.99\% \quad c.v(y) = 1.9\%,$ अतः Y अधिक स्थिर ।
2. $(Q.D. = 4.5; \ \text{Coeff of} \ \ Q.D. = 0.18)$

3. i) 0.405 ii) 0.33

4. i) $\bar{X} = 35$ (ii) $\sigma = 19.76$

5. $\sigma = 1.61$

6. $\sum x^2 = 2, 51, 600$

7. 0.915

8. $\bar{X} = 35.16$, $\sigma = 19.76$

9. $\bar{X} = 35.16$, $\sigma = 19.76$

10. $\bar{X} = 39.90$; $\sigma = 5$

12.8 उपयोगी पाठ्य सामग्री (Useful/Helpful Text)

1. सुदामा सिंह, ओ0 पी0 सिंह, वाई0 के0 सिंह (2002) – ***अर्थशास्त्रीय गणित एवं प्रारम्भिक सांख्यिकी*** - राधा पब्लिकेषन्स, नई दिल्ली।
2. एस0 एन0 लाल, एल0 के0 चतुर्वेदी (2010) - **परिमाणात्मक विष्लेषण**, शिव पब्लिशिंग हाउस, इलाहाबाद।
3. J.K. Sharma (2008) – ***Business Statistics***, Dorling Kinderseley (India) Pvt. Ltd. (Pearson Education), Delhi.

इकाई : 13 विषमता और पृथुशीर्षत्व
(Skewness and Kurtosis)

13. 1 प्रस्तावना (Introduction)

इससे पहले की इकाइयों में आप केन्द्रीय प्रवृत्ति की विभिन्न मापों एवं अपकिरण के बारे में पढ़ चुके हैं। प्रस्तुत इकाइ मे आप को विषमता (Skewness) पृथुशीर्षत्व (Kurtosis) की जानकारी दी जाएगी। प्रस्तुत इकाइ में बाउले का विषमता माप, केली का विषमता माप, विषमता गुणांक की विभिन्न विधियों पृथुशीर्षत्व (Kurtosis) की व्याख्या की जाएगी। अन्त में सारांश एवं अभ्यास के लिए प्रश्न जिससे पाठकों को ईकाई का अध्ययन करने में सुविधा होगी।

13.2 उद्देश्य -

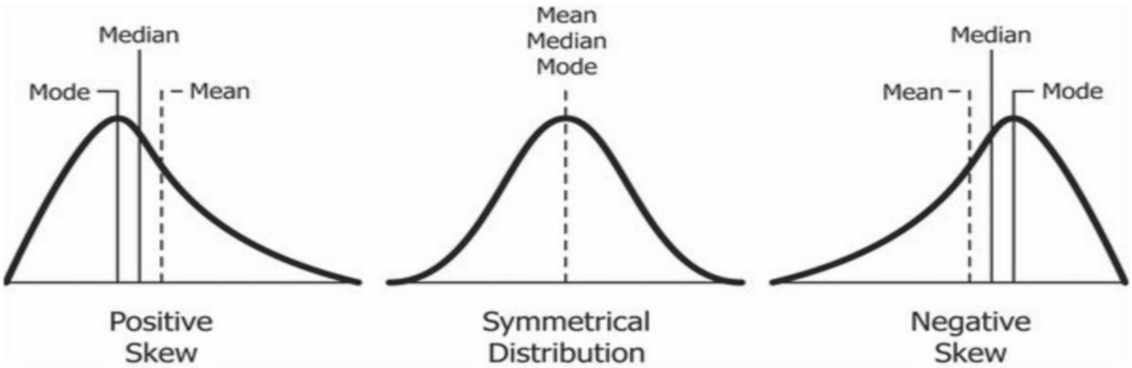
प्रस्तुत इकाई का अध्ययन करने के पश्चात आप

- ✓ विषमता को जानने में सक्षम होंगे
- ✓ विषमता की प्रमुख मापों को जान पाएँगे।
- ✓ विषमता की निरपेक्ष एवं सापेक्ष मापों के बारे में जानकारी प्राप्त कर सकेंगें।

13.3 विषमता (Skewness)

यदि कोई आवृत्ति विवरण केन्द्रीय मान के दोनो ओर सममित है, तथा केन्द्रीय मान के दोनो ओर आवृत्ति वक्र का आकार एक सा है, तो हम आवृत्ति वक्र को पूर्णतया सममित कहेंगे। यदि आवृत्ति वक्र केन्द्रीय मान के दोनो ओर सममित नहीं है अर्थात् वक्र की पुच्छ केन्द्रीय मान के दोनो ओर सममित नहीं है अर्थात वक्र की पुच्छ केन्द्रीय मान के एक ओर अपेक्षाकृत दीर्घ तथा एक ओर अपेक्षाकृत लघु है तो हम ऐसे आवृत्ति वितरण को विषम कहते हैं। विषमता का अर्थ सममिति का अभाव है। यदि किसी विवरण में सममिति से दूर हटने की प्रवृत्ति है तो उस वितरण में विषमता होती है।

सममित बंटन (Symmetrical Distribution) में आवृत्तियाँ एक निश्चित क्रम में बढ़ती हैं तथा फिर उसी क्रम में घटती हैं। इस तरह के वक्र को प्रसामान्य वक्र कहतें हैं। सममित बंटन में समान्तर माध्य, मध्यिका और बहुलक बराबर होते हैं तथा वक्र का आकार केन्द्रीय मान के दोनो ओर एक सा होता है।



जब वक्र का झुकाव दाहिनी ओर अधिक होता है तो बंटन (distribution) में धनात्मक विषमता (Positive Skewness) पायी जाती है - जिसमें

(i) $\bar{X} > M > Z$

(ii) $(Q_3 - M) > (M - Q_1)$

जब असमित वक्र का झुकाव बायीं ओर अधिक होता है तो इसमें ऋणात्मक विषमता (Negative Skewness) पायी जाती है। जिसमें-

(i) $\bar{X} < M < Z$

$$(ii)\ (Q_3 - M) < (M - Q_1)$$

किसी बंटन में विषमता की जाँच आवश्यक होती है - यदि

1) आवृत्ति बंटन का वक्र सममित न हो अर्थात उसका स्वरूप घंटी के आकर का न हो।
2) यदि बंटन में समान्तर माध्य, मध्यिका और बहुलक के बीच जितनी ही अधिक दूरी होगी आवृत्ति बंटन में विषमता उतनी अधिक होगी।
3) मध्यिका से चतुर्थक मूल्यों की दूरी असमान हो।
4) मध्यिका तथा बहुलक से निकाले गये विचलनों का योग शून्य न हो।

विषमता की माप के द्वारा हमें किसी बंटन विषमता या असमिति की मात्रा (अंकात्मक मान) तथा दिशा (धनात्मक या ऋणात्मक) का ज्ञान होता है।

13.3 विषमता की माप (Measures of Skewness)

आवृत्ति बंटन की विषमता माप निरपेक्ष या सापेक्ष हो सकती है, निरपेक्ष माप तुलना योग्य होता है। दो या दो से अधिक श्रेणियों में विषमता की तुलना करने के लिए विषमता की सापेक्ष ज्ञात करते हैं। इसे विषमता गुणांक Coefficient of Skewness) भी कहते हैं। यह गुणांक इकाई विहीन होते हैं तथा इन्हें प्रमाप विचलन से भाग देते है प्रमुख विषमताओं की मापों का विवरण निम्न असममित बंटन में समान्तर माध्य से बहुलक की दूरी जितनी अधिक होगी श्रेणी में विषमता उतनी ही अधिक होगी।

विषमता माप

$$S_k = \bar{X} - Z$$

यह निरपेक्ष माप है, यदि इस निरपेक्ष-माप में हम अपकिरण की माप से भाग दे दें तो हमें सापेक्ष माप या विषमता गुणांक प्राप्त होगा।

कार्ल पियर्सन का विषमता गुणांक - $J = \frac{\bar{X} - Z}{\sigma}$

13.3.1. कार्ल पियर्सन का विषमता माप (Karl Pearson's Skewness Measure)

कभी-कभी बहुलक का निर्धारण कठिन हो जाता है ऐसी स्थिति में हम विषमता माप के वैकल्पिक सूत्र का प्रयोग करते हैं।

कार्ल पियर्सन का विषमता - माप $S_k = 3(\bar{X} - M)$

$$\bar{X} - Z = 3(\bar{X} - M)$$

कार्ल पियर्सन का विषमता - गुणांक

$$J = \frac{(\bar{X} - M)}{\sigma}$$

उपर्युक्त सूत्रों के आधार पर विषमता - माप ज्ञात करने की रीति को विषमता का 'प्रथम माप' कहते है।

विषमता का 'द्वितीय माप' मध्यका से चतुर्थक मूल्यों के अन्तर पर आधारित है इसे बाउले (Bowley) का विषमता माप भी कहतें हैं, इसका प्रथम प्रयोग बाउले ने किया था।

13.3.2 बाउले का विषमता माप (Bowley's Measures of Skewness)

$$S_{KB} = (Q_3 - M) - (M - Q_1) = Q_3 + Q_1 - 2M$$

बाउले का विषमता गुणांक -

$$J_B = \frac{(Q_3 - M) - (M - Q_1)}{(Q_1 - M) + (M - Q_1)} = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

यह रीति बहुत व्यावहारिक नहीं है क्योंकि इसमें श्रेणी के केवल आधे भाग की विषमता का ही माप से पाता है। यह कार्ल पियर्सन के विषमता गुणांक से भिन्न होता है। इसमें दोनों की रीति भिन्न होने से विषमता तुलनात्मक नही होती है।

13.3.3 केली का विषमता माप (Kelly’s Measures of Skewness)

बाउले के विषमता माप की त्रुटि केली के विषमता माप द्वारा की जाती है, इसमें शतमक या दषमक मूल्यों का प्रयोग किया जाता है।

केली का विषमता माप -

$$KS_K (P_{90} - P_{50}) - (P_{50} - P_{10}) = P_{90} + P_{10} - 2P_{50}$$

केली का विषमता गुणांक :

$$J_K = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{(P_{90} - P_{50}) + (P_{50} - P_{10})} = \frac{P_{90} + P_{10} - 2P_{50}}{P_{90} - P_{10}}$$

कुछ उदाहरणों द्वारा हम ऊपर दी गयी विषमता की मापों का स्पष्ट उल्लेख करेंगें।

- **स्मरणीय बिन्दु-**

पूर्णतया सममित विवरणों के लिये तृतीय चतुर्थक तथा मध्यिका का अन्तर $(Q_3 - Md)$ एवम् मध्यिका एवं प्रथम चतुर्थक का अन्तर $(Md - Q_1)$ समान होंगें। धनात्मक विषमता वाले आवृत्ति वक्रों में (Q_3) का मान $(Md - Q_1)$ से अधिक होगा तथा ऋणात्मक विषमता वाले आवृत्ति वक्रो में $(Q_3 - Md)$ का मान $(Md - Q_1)$ से कम होगा।

अन्य शब्दों में उपर्युक्त परिभाषित चतुर्थक विषमता गुणांक (Quartile Coefficient of Skewness) का मान धनात्मक, शून्य अथवा ऋणात्मक होने पर आवृत्ति वक्र क्रमशः धनात्मक विषमता वाला, पूर्णतया सममित अथवा ऋणात्मक विषमता वाला कहलायेगा।

उदाहरण - 1 निम्न आंकड़ों से कार्ल पिर्यसन का विषमता गुणांक ज्ञात कीजिए -

x	रू	2	4	6	8	10	12	14
f	रू	10	18	30	25	12	3	2

हलः इस प्रश्न को बहुलक या मध्यका, दोनों के आधार पर ज्ञात कर सकते हैं। यहाँ स्पष्ट है, अतः हम इसी का प्रयोग करेंगें।

x	f	fx	fx^2
2	10	20	40
4	10	72	288
6	30	180	1080

8	25	200	1600
10	12	120	1200
12	3	36	432
14	2	28	392
	100	656	5032

समान्तर माध्य - $x = \frac{656}{100} = 656$

प्रमाप विचलन - $\sigma = \sqrt{\frac{1}{N}\sum fx^2 - \left(\frac{\sum fx}{N}\right)^2}$

$$= \sqrt{\frac{5032}{100} - \left(\frac{656}{100}\right)^2}$$

$$= \sqrt{50.32 - 43.0336} = \sqrt{7.2864} = 2.6994$$

चूँकि अधिकतम आवृत्ति 30 हैं, अतः इससे सम्बद्ध मूल्य अर्थात् 6 बहुलक होगा।

कार्ल पियर्सन का विषमता गुणांक -

$$J = \frac{\bar{X} - Z}{\sigma} = \frac{6.56 - 6}{2.6994} = 0.2075.$$

अतः वितरण में धनात्मक विषमता है।

उदाहरण 2: निम्न श्रेणी में कार्ल पियर्सन का विषमता गुणांक ज्ञात कीजिए

प्राप्तांक	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
विद्यार्थी सं0	5	20	10	0	5	20	8	7

यहाँ बहुलक स्पष्ट नहीं हो रहा है अतः वैकल्पिक सूत्र द्वारा विषमता गुणांक ज्ञात करेंगें।

प्राप्तांक	**मध्य बिन्दु**	**आवृत्ति**	**dx^1=X-A/n A=45, n=10**	**fdx^1**	**fdx^2**	**संचयी आवृत्ति**
0-10	5	5	-4	-20	80	5
10-20	15	20	-3	-60	180	25
20-30	25	10	-2	-20	40	35
30-40	35	0	-1	0	0	35
40-50	45	5	0	0	0	40
50-60	55	20	1	20	20	60
60-70	65	8	2	16	32	68
70-80	75	7	3	21	63	75
Total		75		-43	415	

सामान्तर माध्य - $\overline{X} = \frac{A + n\Sigma fdx'}{N}$

$$= 45 + \frac{10 \times (-43)}{75} = 45 - \frac{86}{15} = 45 - 5.7$$

$$= 39.3$$

$$\frac{N}{2} = \frac{75}{2} = 37.5$$

अतः 40-50 वाले वर्ग में मध्यका होगी।

मध्यका $M = \ell_1 + \left(\frac{\frac{N}{2} - c}{f} \right) \times h$

$$= 40 + \frac{37.5 - 35}{5} \times 10 = 40 + \frac{2.5}{5} \times 10 = 45.$$

प्रमाप विचलन

$$\sigma = h \times \sqrt{\frac{1}{N}\Sigma fdx'^2 - \frac{(\Sigma fdx')^2}{N}}$$

$$\sigma = h \times \sqrt{\frac{415}{75} - \left(\frac{-43}{75}\right)^2} = 10 \times \sqrt{5.53 - 0.33}$$

$$= h \times \sqrt{\frac{415}{75} - \left(\frac{-43}{75}\right)^2} = 10 \times \sqrt{5.53 - 0.33}$$

$$= 10 \times 2,28 = 22.8$$

अतः कार्ल पिर्यसन का विषमता गुणांक -

$$J = \frac{3(\overline{X} - M)}{S} = 3\frac{(39.3 - 45.0)}{22.8} = -\frac{17.1}{22.8}$$

$$= -0.75.$$

उदाहरण 3. बंटन A और बंटन B के संबंध में निम्न सूचनाएँ प्राप्त हैं

	बंटन A	बंटन B
समान्तर माध्य	50	45
मध्यका	45	40
प्रमाप विचलन	5	5

निम्न की जाँच कीजिए।

i) बंटन A और बंटन B में समान विचरण है।

ii) दोनों बंटनों में विषमता समान है।

हल - विचरण की तीव्रता के लिये हम विचरण गुणांक ज्ञात करते हैं।

i) बंटन A में $C.V = \sigma / \overline{X} \times 100 = 5/50 \times 100 = 10$

बंटन B में $C.V = \sigma / \overline{X} \times 100 = 5/45 \times 100 = 11.11$

स्पष्ट रूप से बंटन B में विचलन अधिक है अतः i) कथन असत्य है।

ii) कार्ल पियर्सन का विषमता गुणांक -

बंटन A में $= \frac{3(\overline{X} - M)}{\sigma} = \frac{3(50 - 45)}{5} = \frac{3 \times 5}{5} = 3$

बंटन B में $= \frac{3(\overline{X} - M)}{\sigma} = \frac{3(45 - 40)}{5} = \frac{3 \times 5}{5} = 3$

दोनों बंटनों में विषमता समान है तथा दिया गया कथन ii) सत्य है।

उदाहरण 4: निम्न वितरण से कार्ल पिर्यसन का विषमता गुणांक और बाउले का विषमता गुणांक ज्ञात कीजिए -

	श्रेणी A	**श्रेणी B**
समान्तर माध्य	150	140
मध्यका	142	155
प्रमाप विचलन	30	55
तृतीय चतुर्थक	195	260
प्रथम चतुर्थक	62	80

हल -

श्रेणी A में कार्ल पियर्सन का विषमता गुणांक -

$$J = \frac{3(\overline{X} - M)}{\sigma} = \frac{3(150 - 142)}{30} = \frac{3 \times 8}{30} = 0.8$$

बाउले का विषमता गुणांक -

$$J_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = \frac{195 + 62 - 2 \times 142}{195 - 62} = \frac{157 - 284}{133}$$

$$= -\frac{27}{133} = -0.203$$

श्रेणी B में कार्ल पियर्सन का विषमता गुणांक

$$J = \frac{3(\bar{X} - M)}{6} = \frac{3(140 - 155)}{55} = \frac{3 \times (-15)}{55} = -\frac{9}{11} = -0.82$$

तथा बाउले का विषमता गुणांक -

$$J_B = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1} = \frac{260 + 80 - 2 \times 155}{260 - 80} = \frac{340 - 310}{180}$$

$$= \frac{1}{6} = 0.167$$

यहाँ एक ही श्रेणी में कार्ल पियर्सन की रीति से निकाला गया विषमता गुणांक और बाउले की रीति से निकाला गया विषमता गुणांक भिन्न हो सकता है।

अपकिरण एवं विषमता (Dispersion and Skewness)

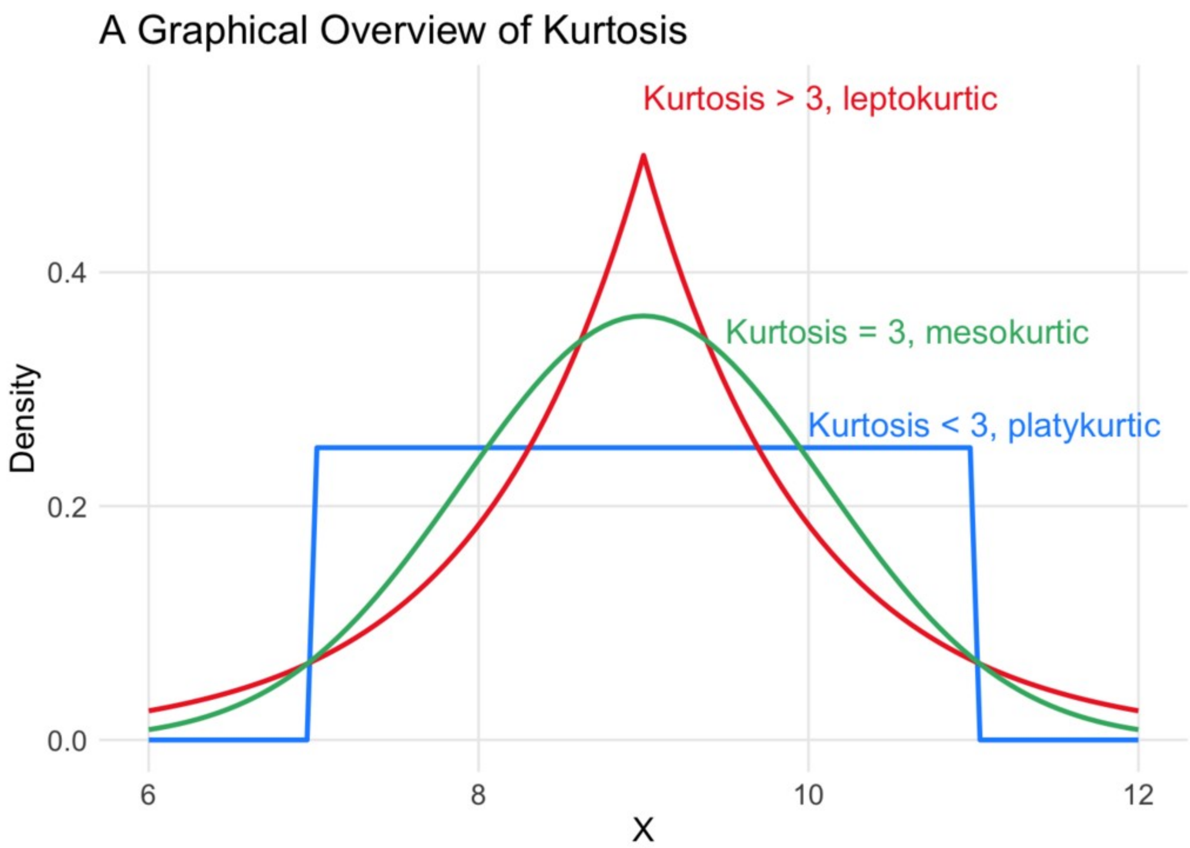
किसी बंटन में अपकिरण एवं विषमता विश्लेषण के प्रमुख अंग है। इस दृष्टि से ये माप एक दूसरे के पूरक हैं। इनमें प्रमुख अन्तर निम्न हैं-

अपकिरण	विषमता
यह पद मूल्यों के बिखराव या श्रेणी की बनावट बताता है।	यह बंटन की सममिति अथवा असमिति बताता हैं।
यह पूरी श्रेणी के बिखराव को मापता है।	यह माध्य के दोनो तरफ के बिखराव की तुलना करता है।
यह प्रथम, द्वितीय एवं तृतीय परिघातों से संबंधित हैं।	यह केवल तृतीय परिघात से संबंधित हैं।

13.4 पृथुशीर्षत्व (Kurtosis)

हमने पूर्व में ही स्पष्ट किया है कि किसी आवृत्ति बंटन की संपूर्ण विषेषताओं की जानकारी करने के लिए चार प्रकार के माप आवश्यक हैं। केन्द्रीय प्रवृत्ति की माप, अपकिरण की माप और विषमता की माप, अत्यन्त महत्वपूर्ण होते हुए भी किसी बंटन की सम्पूर्ण कहानी को नहीं कह पाते। इसके लिए हमें चौथी माप अर्थात् पृथुशीर्षत्व की माप की जानकारी आवश्यक है। पृथुशीर्षत्व को कार्ल पियर्सन वक्र की उत्तलता (Convexity of Curve) कहते हैं। विषमता से हमें यह जानकारी प्राप्त होती है कि बंटन सममित है अथवा असममित है। दूसरे शब्दों में विषमता आवृत्ति वक्र की दायीं अथवा बाँयी पूंछ (Right or Left Tail) की स्थिति की जानकारी प्रस्तुत करती है। पृथुशीर्षत्व हमें वक्र के माध्य अथवा शीर्ष के नुकीलेपन (Peakedness) या चपटेपन (Flatness) की जानकारी प्रदान करता है।

किसी आवृत्ति बंटन का वक्र नुकीला अथवा चपटा हो सकता है। यदि हम प्रसामान्य वक्र (Normal curve) को आधार माने तो पृथुशीर्षत्व हमें प्रसामान्यता से अलगाव (Departure from Normality) की जानकारी देता है। इसकी माप से हमें श्रेणी के मध्य भाग में आवृत्तियों के जमाव का ज्ञात प्राप्त होता है। यदि मध्य भाग में आवृत्तियों का जमाव सामान्य है तो वह आवृत्ति वक्र मध्यम शीर्ष वाला (Mesokurtic) या प्रसामान्य (Normal) कहलाता है। यदि मध्य भाग में आवृत्तियों का जमाव अत्यधिक सघन है तो वक्र नुकीले शीर्ष वाला (Leptokurtic) होता है। इसके विपरीत यदि केन्द्र में आवृत्तियों का जमाव कम है तो वक्र चपटे शीर्ष वाला (Platykurtic) कहलाता है। चित्र में वक्र C चपटे शीर्ष वाला वक्र है। पृथुशीर्षत्व आकृति (नुकीलेपन अथवा चपटेपन) को मापता है।



Source: https://commons.wikimedia.org/wiki/File:14fig-kurtosis.png

13.5 कार्ल पियर्सन का β2 और γ2 गुणांक

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4} \text{ and } \gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\sigma^4} - 3$$

यदि $\beta_2 = 3$ या $\gamma_2 = 0$ तो वक्र मध्यम शीर्ष वाला (Mesokurtic) है।
यदि $\beta_2 > 3$ या $\gamma_2 > 0$ तो वक्र नुकीले शीर्ष वाला (Leptokurtic) है।
यदि $\beta_2 < 3$ या $\gamma_2 < 0$ तो वक्र चपटे शीर्ष वाला (Platykurtic) है।

उदाहरण 1. यदि किसी बंटन में समान्तर माध्य $(\bar{X}) = 1$ हो और μ_2, μ_3 तथा μ_4 का मान क्रमशः 3, 0 और 27 हो तो इनकी सहायता से विषमता तथा पृथुशीर्षत्व की जाँच कीजिए।

हल - कार्ल पियर्सन का विषमता गुणांक -

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = 0 \text{ or } \beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0$$

अतः बंटन सममित है या $\bar{X} = M = Z$

कार्ल पियर्सन का पृथुशीर्षत्व माप $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{27}{(3)^2} = \frac{27}{9} = 3$

या $\gamma_2 = \beta_2 - 3 = 0$

चूँकि $\mu_2 = 0$ अतः दिया हुआ बंटन मध्यम शीर्ष वाला (Mesokurtic) या प्रसामान्य है।

चूँकि $\mu_3 = 0$ और $\mu 2 = 3$ अतः बंटन समान्तर माध्य और प्रमाप विचलन $\sqrt{3} = 1.732$ के साथ प्रसामान्य है।

उदाहरण 2: यदि $\mu_2 = 120$ और $\mu_4 = 36000$ तो बंटन का स्वरूप ज्ञात कीजिए।

हल - बंटन के स्वरूप के लिए हमें पृथुशीर्षत्व की माप करनी होगी।

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{36000}{120 \times 120} = 2.5 \text{ or } \gamma_2 = \beta_2 - 3 = -0.5$$

चूँकि $\beta_2 < 3 \text{ or } \gamma_2 < 0$

अतः बंटन चपटे शीर्ष वाला (Platykurtic) है।

उदाहरण 3: यदि एक सममित बंटन में प्रमाप विचलन हो तो चतुर्थ केन्द्रीय परिघात का क्या मूल्य हो ताकि बंटन

i) नुकीले शीर्ष वाला हों।

ii) मध्यम शीर्ष वाला हो

iii) चपटे शीर्ष वाला हो।

हल - $\sigma = 4$ अतः $\sigma^2 = \mu^2 = (4)^2 = 16$

i) नुकीले शीर्ष वाले बंटन के लिए $\beta_2 > 3$ होना चाहिए।

या $\frac{\mu_4}{\mu_2^2} > 3$ या $\mu_4 > 3\mu_2^2$

या $\mu_4 > 3 \times (16)^2$

$\mu_4 = > 3 \times 256 > 768$

ii) मध्यम शीर्ष वाले बंटन में -

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3$$

या $\mu_4 = 3\mu_2^2$

$$= 3 \times (16)^2$$

$$= 3 \times 256 = 768$$

iii) चपटे शीर्ष वाले बंटन में -

$$\beta_2 = \frac{\mu_4}{\mu_2^2} < 3$$

या $\mu_4 < 3\mu_2^2$

$$\mu_4 < 3 \times 256$$

$$< 768$$

उदाहरण – 4. निम्न वितरण की कुकुदता गुणांक ज्ञात कीजिए।

वर्गान्तर	0-4	4-8	8-12	12-16	16-20
आवृत्ति	1	3	12	8	6

हल:

वर्गान्तर	**मध्य बिन्दु**	**आवृत्ति**	$(X - \bar{X})$	$(X - \bar{X})^2$	$(x - \bar{x})^4$	$f(x - \bar{x})^4$
0-4	2	1	-10	100	10000	10000
4-8	6	3	-6	36	1296	3888
8-12	10	12	-2	4	16	192
12-16	14	8	2	4	16	128
16-20	18	6	6	36	1296	7776

$$M_4 = \frac{\Sigma f(x - \bar{x})^4}{N} = \frac{21984}{30} = 732.80$$

σ का मान = 4.06

इसे कुकुदता के सूत्र में रखने पर

कुकुदता गुणांक $= \frac{M_4}{\sigma^4} = \frac{732.80}{(4.06)^4} = \frac{732.80}{271.71} = 2.70$

कुकुदता का मान चूँकि 3 से कम है अतः दिये हुए वितरण का आवृत्ति वक्र चपटे शीर्ष वाला (Platykurtic) होगा।

13.6 परिघात (Moment)

सांख्यिकी में परिघात का प्रयोग किसी आवृत्ति बंटन के विभिन्न मापों, जैसे - केन्द्रीय प्रवृत्ति, अपकिरण, विषमता प्रथुषीर्षत्व आदि के विश्लेषण में होता है। किसी श्रेणी में परिघातों का परिकलन समान्तर माध्य अथवा किसी कल्पित माध्य से परिघात (Moment About the Arithmetic Mean) कहते हैं। परिघातों की संख्या कई होती है। परिघात या अपूर्ण शब्द का भौतिक विज्ञान या यन्त्रिका विज्ञान (Mechanics) में प्रायः प्रयोग होता है, जहाँ यह किसी बिन्दु के सापेक्ष घुमाव पैदा करने वाले बल को मापता है।

परिघातों की संख्या, कल्पित मूल बिन्दु से परिघात, केन्द्रीय परिघातों के परिकलन, व्यक्तिगत श्रेणी-आवृत्ति श्रेणी, शेपर्ड के संशोधन, चार्लियर की शुद्धता जाँच, परिघातों पर आधारित कार्ल पिर्यसन के गुणांक, परिघातों पर आधारित विषमता गुणांक, पृथुषीर्षत्व, पृथुषीर्षत्व की माप के रूप में कार्ल पियर्सन ने β2 और γ2 गुणांकों का प्रयोग किया है।

13.6.1 उद्देश्य -

परिघात क्या है?

परिघात के उपयोग क्या हैं ?

परिघात को मापने की विधि क्या है?

प्रथुषीर्षत्व क्या है तथा इसकी उपयोगिताएँ क्या हैं?

13.6.2 परिघातों की संख्या (Number of Moment)

परिघातों की संख्या कई होती हैं, सामान्य रूप से किसी आवृत्ति विवरण में r वाँ केन्द्रीय परिघात -

$$\mu_r = \frac{1}{N}\sum f(X - \bar{X})^r;\ r = 0, 1, 2\ \ldots\ldots (1)$$

$$= \frac{1}{N}\left[f_1(X - \bar{X})^r + f_2(X_2 - \bar{X})^r + \ldots\ldots\ldots f_n(X_n - \bar{X})^r\right]$$

समीकरण (1) में r = 0 रखने पर

$$M_0 = \frac{1}{N}\sum f\ (X - X)^0 = \frac{1}{N}\sum f = 1$$

क्योंकि $\sum f = N$ और $(X - \bar{X})^0 = 1$

अतः $\mu_0 = 1.$

पुनः समीकरण (1) में r = 1 रखने पर

$\mu_1 = \frac{1}{N}\sum f\ (X - \bar{X}) = 0$ क्योंकि $\sum(X - \bar{X}) = 0$

याद करें, समान्तर माध्य से निकाले गये विचलनों का योग शून्य होता है।

पुनः r = 2 रखने पर समीकरण (1) द्वारा

$$\mu_2 = \frac{1}{N}\sum f\ (X - \bar{X})^2 = \sigma^2 = \text{variance}$$

इस प्रकार द्वितीय केन्द्रीय परिघात प्रसरण को व्यक्त करता है

पुनः $$\mu_3 = \frac{1}{N} \sum f(X - \bar{X})^3$$

और $$\mu_4 = \frac{1}{N}\sum f(X - \bar{X})^4$$

कल्पित मूल बिन्दु से परिघात (Moments about arbritrary origin)

समान्तर माध्य $\bar{X}$ के बजाय हम किसी कल्पित बिन्दु (कल्पित माध्य) A से परिघातों का परिकलन कर सकते हैं। सामान्य रूप से कल्पित मूल बिन्दु A से r वाँ परिघात

$$\mu_r^1 = \frac{1}{N}\sum f(X - A)^r,\ r = 0, 1, 2, \ldots \qquad (2)$$

$$= \frac{1}{N}\left[f_1(X_1 - A)^r + f_2(X_2 - A)^r + \ldots f_n(X_n - A)^r\right]$$

समीकरण (2) में r = 0 रखने पर -

$$\mu_0^1 = \frac{1}{N}\sum f(X - A)^0$$

$$= \frac{1}{N}\sum f = 1$$

इसी समीकरण में r =1 रखने पर -

$$\mu_1 = \frac{1}{N}\sum f(X - A)$$

$$= \frac{1}{N}\left[\sum fX - \sum fA\right]$$

$$= \frac{1}{N}\sum fX - \frac{A\sum f}{N}$$

$$= \frac{1}{N}\sum fX - A$$

$$= \bar{X} - A$$

अतः $$= \bar{X} = \mu_1 + A \qquad (3)$$

समीकरण (3) बहुत ही महत्वपूर्ण समीकरण है। हम यह देख सकते हैं कि यदि $A = O$ तो $\bar{X} = \mu_1$ अर्थात शून्य मूल बिन्दु से लिया गया परिघात समान्तर माध्य को व्यक्त करता है।

समीकरण (2) में $r = 2, 3, 4, \ldots\ldots$ इत्यादि रखने पर

$$\mu_2 = \frac{1}{N}\sum f(X - A)^2$$

$$\mu_3 = \frac{1}{N}\sum f(X - A)^3$$

$$\mu_4 = \frac{1}{N}\sum f(X - A)^4$$

पाठक अब यह समझ चुके हैं कि समान्तर माध्य से लिए गये परिघात अथवा केन्द्रीय परिघात को μ (म्यू) तथा कल्पित बिन्दु A से लिए गए परिघात को μ_1 से लिखते हैं।

हम जानते हैं कि

$$\mu_r = \frac{1}{N}\sum f(X - \bar{X})^r$$

इसे हम निम्न रूप में लिख सकते हैं।

$$\mu_r = \frac{1}{N}\sum f\,(X - A + A - \bar{X})^r \qquad (4)$$

समीकरण (3) से

$\bar{X} = \mu_1 + A$ या $A - \bar{X} = u_1^{'}$

अतः समीकरण (4) से

$$\mu_r = \frac{1}{N}\sum f\ \left[(X - A) - \mu_1\right] \qquad (5)$$

समीकरण (5) का द्विपद प्रमेय से विस्तार करने पर μ और μ_r में संबंध निकल जाता है। हम यहाँ पर इस विस्तार को पूर्णरूपेण प्रस्तुत करने की कोई आवश्यकता महसूस नही कर रहे हैं। इस विस्तार का अन्ततः सामान्य रूप निम्न होगा-

$$\mu_r = \mu_r - {}^rC_1\mu_{r-1}\ \mu_1 + {}^rC_2\ \mu_{r-1}\ \mu_1^2 \qquad (6)$$

इसी समीकरण (6) में और r = 2, 3 और 4 रखने पर

$$\mu_2 = \mu_2^{'} - \mu_1^{'2} \qquad (7)$$

$$\mu_3 = \mu_3^{'} - 3\mu_2^{'}\ \mu_1^{'} + 2\mu_1^{'3} \qquad (8)$$

$$\mu_4 = \mu_4' - 4\mu_3' \mu_1' + 6\mu_2' \mu_1'^2 - 3\mu_1'^4 \quad (9)$$

समीकरण (7), (8), (9) बहुत महत्वपूर्ण हैं तथा सदैव याद रखने योग्य है। हम केन्द्रीय परिघातों से कल्पित मूल बिन्दु पर आधारित परिघातों की भी गणना कर सकते हैं। इसके लिए हमें सर्वप्रथम समान्तर माध्य की गणना करनी पड़ती है तथा कल्पित मूल बिन्दु की जानकारी रखनी पड़ती हैं। समीकरण (3) में हम जानते हैं कि

$\bar{X} = \mu_1' + A$ या $\mu_1' = \bar{X} - A$ समीकरण (7), (8), (9) से

$$\mu_2' = \mu_2 + \mu_1'^2 \quad (10)$$

$$\mu_3' = \mu_3 + 3\mu_2\mu_1' + \mu_1'^3 \quad (11)$$

$$\mu_4^1 = \mu_4 + 4\mu_3\mu_1^1 + 6\mu_2\mu_1'^2 + \mu_1'^4 \quad (12)$$

परिघात ज्ञात करने संबंधी प्रश्नों का हल करते समय कुछ बाते ध्यान देने योग्य हैं-

13.6.3 केन्द्रीय परिघातों के परिकलन

केन्द्रीय परिघातों के परिकलन करते समय यदि समान्तर माध्य पूर्ण संख्या है तब तो सीधे सूत्र

$$\mu_r = \frac{1}{N}\sum f\ (X - \bar{X})^r$$

का प्रयोग करना चाहिए। हमें केन्द्रीय परिघातों के सूत्र को यहाँ फिर देख लें।

व्यक्तिगत श्रेणी

$$\mu_1 = \frac{\sum(X - \bar{X})}{N} = 0 = \frac{\sum d}{N}$$

$$\mu_2 = \frac{\sum(X - \bar{X})^2}{N} = \frac{\sum d^2}{N} = \sigma^2$$

$$\mu_3 = \frac{\sum(X - \bar{X})^3}{N} = \frac{\sum d^3}{N}$$

$$\mu_4 = \frac{\sum(X - \bar{X})^4}{N} = \frac{\sum d^4}{N}$$

आवृत्ति श्रेणी

$$\mu_1 = \frac{1}{N}\Sigma f(X - \bar{X}) = \frac{\Sigma fd}{N} = 0$$

$$\mu_2 = \frac{1}{N}\Sigma f(X - \bar{X})^2 = \frac{\Sigma fd^2}{N} = \sigma^2$$

$$\mu_3 = \frac{1}{N}\Sigma f(X - \bar{X})^3 = \frac{\Sigma fd^3}{N}$$

$$\mu_4 = \frac{1}{N}\Sigma f(X - \overline{X})^4 = \frac{\Sigma fd^4}{N}$$

यदि समान्तर माध्य दशमलव में आता है तो सीधे सूत्र का प्रयोग करके केन्द्रीय परिघातों का परिकलन बहुत कठिन हो जाता है। ऐसी स्थिति में हम पहले किसी कल्पित माध्य (A) से परिघात μ_r' ज्ञात करते हैं और फिर इनके आधार पर केन्द्रीय परिघातों की गणना कर ली जाती है। हम जानते हैं कि किसी कल्पित मूल बिन्दु A से परिघात

$$\mu r' = \frac{1}{N}\Sigma f(X - A)^r$$

यदि $dx = X - A$ तो $\mu r' = \frac{1}{N}\Sigma f(dx)^r$

परिघातों के संबंध में निम्न बातें महत्वपूर्ण हैं.

परिघातों के संबंध में निम्न बातें भी काफी महत्वपूर्ण हैं -

(1) $\mu_0 = 1, \mu_1 = 0$ तथा समान्तर माध्य $\overline{X} = A + \mu_1^1$

(2) यदि सममित वितरण है तो हम जानते हैं कि ऐसे वितरण में आवृत्तियाँ जिस क्रम में बढ़ती हैं उसी क्रम में घटती हैं। सममित वितरण में समान्तर माध्य से विचलन लिया जाए और इस विचलन का विषम घात (1, 3, 5, 7..... इत्यादि) किया जाए तो धनात्मक विचलन और ऋणात्मक विचलन एक दूसरे के बराबर होते हैं। इस कारण μ_1, μ_3, μ_5 इत्यादि का मान शून्य होता है अर्थात् -

$$\Sigma f(X - \overline{X}) = \Sigma f(X - \overline{X})^3 = \Sigma f(X - \overline{X})^5 = ... = 0$$

या $\mu_1 = \mu_3 = \mu_5 .... = 0$

या $H_{2r+1} = 0$ जहाँ $r = 0, 1, 2, 3.....$

(3) यदि आवृत्ति वचरण (वर्गीकृत) में हम पैमाने का परिवर्तन (Change of Scale) करें अर्थात् यदि

$$dx' = X - A / h$$

या $X = A + hdX'$

और $\overline{X} = A + hd\overline{X}'$

इस प्रकार

$$\mu_r = \frac{1}{N}\Sigma f\left(X - \overline{X}\right)^r$$

$$= h^r \frac{1}{N}\Sigma f\left(dX' - d\overline{X}'\right)$$

(4) हम जानते हैं कि $\mu_2 = \sigma^2 =$ प्रसरण तथा

$$\mu_r' = \frac{1}{N}\Sigma f(X - A)^2 = \text{ विचलन वर्ग माध्य}$$

$$\mu_2 = \mu_2' - \mu_1'^2$$

चूँकि μ_1' एक वास्तविक संख्या है, अतः इसका वर्ग एक गैर ऋणात्मक राशि होगी। इस प्रकार

$\mu_2 = \mu_2'$ - एक गैर ऋणात्मक राशि

अतः $\mu_2 < \mu_2'$

या प्रसरण < विचलन-वर्ग माध्य

या Variance ≤ Mean Square Deviation

या S.D. ≤ Root Mean Square Deviation

उदाहरण – 1. किसी चर के 10 पदों पर आधारित प्रथम चार परिघात क्रमषः 5, 30, 40 और 50 हैं। समान्तर माध्य, प्रसरण μ_3 तथा μ_4 का परिकलन कीजिए।

हल -

$$A = 10, \mu_1' = 5, \mu_2' = 30, \mu_3' = 40, \mu_4 = 50$$

समान्तर माध्य $\overline{X} = A + \mu_1' = 10 + 5 = 15$

प्रसरण $\mu_2 = \mu_2' - \mu_1'^2 = 30 - 5^2 = 30 - 25 = 5$

$$\mu_3 = \mu_3^1 - 3\mu_2^1\mu_1^1 + 2\mu_1^{1^3}$$

$$= 40 - 3 \times 30 \times 5 + 2(5)^3$$

$$= 40 - 450 + 2 \times 125 = 290 - 450 = 160$$

$$\mu_4 = \mu_4^1 - 4\mu_3^1\mu_1^1 + 6\mu_2^1 + 6\mu_2^1\mu_1^2 - 3\mu_1^{1^4}$$

$$= 50 - 4 \times 40 \times 5 + 6 \times 30 \times 25 - 3 \times 625$$

$$= 50 - 800 + 4500 - 1875$$

$$= 4550 - 2675$$

$$= 1875$$

उदाहरण 2. यदि किसी श्रेणी का समान्तर माध्य 7 और प्रथम चार केन्द्रीय परिघात क्रमषः 0, .16, 64 और 162 हो तो – (1) कल्पित मूल बिन्दु 5 पर आधारित और (2) शून्य पर आधारित परिघातों की गणना कीजिए।

हल -

(1) $\overline{X} = A + \mu_1'$

या $\mu_1' = \overline{X} - A = 7 - 5 = 2$

$$\mu_2' = \mu_2 + \mu_1'^2 = 16 + 4 = 20$$

$$\mu_3' = \mu_3 + 3\mu_2\mu_1' + \mu_1'^3$$

$$= -64 + 3 \times 16 \times 2 + (2)^3$$

$$= -64 + 96 + 8 = 40$$

$$\mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2\mu_1'^2 - \mu_1'^4$$

$$= 162 + 4 \times (-64) \times (2) + 6 \times 16(2)^2 + (2)^4$$

$$= 162 - 512 + 384 + 16$$

$$= 562 - 512 = 50$$

इस प्रकार मूल बिन्दु 5 पर आधारित प्रथम चार परिघात 2, 20, 40 और 50 है।

(2) $\overline{X} = 7, A = 0$

अतः $\mu_1' = \overline{X} - A = 7 - 0 = 7$

$$\mu_2' = \mu_2 + \mu_1'^2$$

$$= 16 + (7)^2 = 16 + 49 = 65$$

$$\mu_3' = \mu_3 + 3\mu_2\mu_1' + \mu_1'^3$$

$$= -64 + 3 \times 16 \times 7 + (7)^3$$

$$= -64 + 336 + 343 = 615$$

$$\mu_4' = \mu_4 + 4\mu_3\mu_1' + 6\mu_2\mu_1'^2 - \mu_1'^4$$

$$= 162 + 4(-64) \times 7 + 6 \times 16 \times (7)^2 + (7)^4$$

$$= 162 - 1792 + 3136 + 2401$$

$$= 5699 - 1792$$

$$= 3907$$

13.6.4 शेपर्ड के संशोधन (Sheppard's Correction for Moments)

वर्गीकृत श्रेणी में परिघातों की गणना करते समय वर्गान्तरों के मध्य बिन्दु को ही चर (X) मानते हैं। यह मान लिया जाता है कि आवृत्तियों का जमाव वर्ग के मध्य बिन्दु पर ही होता है। यह मान्यता सममित वितरण में लगभग सही होता है लेकिन सामान्यतः या असममित वितरण में ऐसा मानना उचित नहीं होता और परिणामस्वरूप कुछ त्रुटि जिसे समूहन त्रूटि (Grouping error) कहते हैं, परिघातों की गणना में विद्यमान हो जाती है। इस विभ्रम या त्रुटि को दूर करने के लिए शेपर्ड (W.F. Sheppard) ने निम्न सूत्रों का प्रयोग किया और इसे ही शेपर्ड के संशोधन कहते हैं -

$$\mu_2 \text{ (संशोधित)} = \mu_2 - \frac{h^2}{12}$$

$$\mu_3 \text{ (संशोधित)} = \mu_3$$

$$\mu_4 \text{ (संशोधित)} = \mu_4 - \frac{1}{2}h^2\mu_2 + \frac{7}{240}h^4$$

हम जानते हैं कि μ_1 सदैव शून्य होता है। μ_1 और μ_3 का संशोधन इसलिए भी आवश्यक नहीं है कि इनमें विचलनों के धनात्मक और ऋणात्मक चिन्ह बने रहते हैं। अतः अशुद्धि क्षतिपूरक प्रकृति के कारण लगभग स्वतः समाप्त हो जाती है।

शेपर्ड का संशोधन सममित या साधारण सममित आवृत्ति बंटन में ही उपयुक्त होता है। J या $\cup$ आकृति वाले आवृत्ति बंटन में यह उपयुक्त नहीं होता। साथ ही यदि आवृत्तियों की संख्या बहुत बड़ी हो (1000 से अधिक) तभी यह संशोधन उपयुक्त होता है।

13.6.5 चार्लियर की शुद्धता जाँच (Charlier's check)

चालिर्ययर की शुद्धता जाँच की चर्चा हमने पूर्व के अध्यायों में की है। हमें यह भी ज्ञात है कि 'चार्लियर-जाँच' के द्वारा गणना क्रिया के शुद्धता की परीक्षा होती है। परिघातों के संबंध में अगर निम्न शर्त पूरी होती है तो गणना-क्रिया में कोई अशुद्धि नहीं है -

परिघात	जाँच-सूत्र
प्रथम	$\Sigma f(dx+1) = \Sigma fdx + N$
द्वितीय	$\Sigma f(dx+1)^2 = \Sigma fdx^2 + 2\Sigma fdx + N$
तृतीय	$\Sigma f(dx+1)^3 + \Sigma fdx^3 + 3\Sigma fdx^2 + 3\Sigma fdx + N$
चतुर्थ	$\Sigma f(dx+1)^4 = \Sigma fdx^4 + 4\Sigma fdx^3 + 6\Sigma fdx^2 + 4\Sigma fdx + N$

उपर्युक्त जाँच-सूत्र, परिघात ज्ञात करने की अप्रत्यक्ष या लघुरीति पर आधारित है। यदि प्रत्यक्ष रीति का प्रयोग कर रहे हैं तो $dx\,(X - A)$ की जगह रखा जाएगा।

13.6.5 परिघातों पर आधारित कार्ल पिर्यसन के गुणांक (Karl Pearson's Coefficients based on moments)

प्रथम चार केन्द्रीय परिघातों पर आधारित चार गुणांकों की चर्चा कार्ल पियर्सन ने किया है। इन गुणांकों का प्रयोग तुलनात्मक अध्ययन में सुविधापूर्वक किया जाता है। व्यवहार में विषमता एवं पृथुशीर्षत्व के माप में ये गुणांक बहुत उपयोगी है। ये बीटा और गामा गुणांक निम्न हैं - $\beta = \frac{\mu_3^2}{\mu_2^3}$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}$$

$$Y_1 = \sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3} \quad \text{क्योंकि} \quad \mu_2 = \sigma^2$$

$$Y_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3$$

कभी-कभी अल्फा (alpha) गुणांक की भी चर्चा की जाती है। अल्फा गुणांक निम्न हैं -

$$\alpha_1 = \frac{\mu_1}{\sigma} = 0, \alpha_2 = \frac{\mu_2}{\sigma^2} = 1$$

$$\alpha_3 = \frac{\mu_3}{\sigma^3} = \sqrt{\beta_1} = \gamma_1, \alpha_4 = \frac{\mu_4}{\sigma^4} = \beta_2$$

13.6.6 परिघातों पर आधारित विषमता गुणांक (Coefficient of Skewness based on Moments)

परिघातों पर आधारित कार्ल पियर्सन का विषमता गुणांक -

$$S_k = \frac{\sqrt{\beta_1}(\beta_2 + 3)}{2(5\beta_2 - 6\beta_1 - 9)}$$

पाठक यहाँ धान दें, पूर्व में हमने कार्ल पियर्सन के विषमता गुणांक के लिए हमने J का प्रयोग किया है। यहाँ पर S_k का प्रयोग इसलिए कर रहे हैं कि दोनों सूत्रों में अन्तर आसानी से हो सके।

यहाँ β_1 और β_2 कार्ल पियर्सन के गुणांक है। इस सूत्र द्वारा यदि $\overline{X} > Z$ धनात्मक विषमता होगी तथा यदि $X < Z$ तो ऋणात्मक विषमता होगी। यदि $S_k = 0$ तो या तो $\beta_1 = 0$ या $\beta_2 + 3 = 0$ या $\beta_2 = -3$

लेकिन $\beta_2 = \mu^4 / \mu_2^2$ जहाँ $\mu_4 = {}^1\!/_N \Sigma f(x - \overline{x})^4 > 0$

और μ_2 प्रसारण है जिसका वर्ग ऋणात्मक नहीं हो सकता अतः -

$$\beta_2 = \frac{\mu_4}{\mu_2^2} > 0$$

इस प्रकार यदि $S_k = 0$ तो $\beta_1 = 0$ या $\mu_3 = 0$ । इस प्रकार सममित वितरण में $\beta_1 = 0$ । अतः को विषमता के माप के रूप में प्रयोग किया जाता है। लेकिन इसकी एक गम्भीर सीमा ;सपउपजंजपवदद्ध है। चूँकि जहाँ तो धनात्मक या ऋणात्मक हो सकता है लेकिन सदैव धनात्मक होगा। इसी प्रकार भी धनात्मक होगा। इस प्रकार , विषमता की दिषा (धनात्मक या ऋणात्मक) को बताने में असमर्थ होता है। इस दोष को दूर करने के लिए कार्ल पियर्सन के गामा गुणांक का प्रयोग किया जाता है।

$$\gamma_1 = +\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3}$$

इस प्रकार विषमता का चिन्ह μ_3 पर निर्भर हो जाता है। यदि μ_3 ऋणात्मक होगा तो विषमता भी ऋणात्मक होगी और यदि μ_3 धनात्मक होगा तो विषमता भी धनात्मक होगी।

13.7 सारांश (Summary)

विषमता का अर्थ सममिति का अभाव है। यदि किसी वितरण में सममिति से दूर हटने की प्रवृत्ति है तो उसे विषय कहते हैं। जब वक्र का झुकाव दाहिनी ओर अधिक होता है तो बंटन में धनात्मक विषमता पायी जाती है। जब असममित वक्र का झुकाव बायीं ओर अधिक होता हैं तो इसमें ऋणात्मक विषमता पायी जाती है। विषमता की माप द्वारा हमें बंटन में असममिति की मात्रा दिशा का ज्ञान होता है। दो या दो से अधिक श्रेणियों में तुलना करने के लिये विषमता गुणांक का प्रयोग करते हैं। यह गुणांक इकाई विहीन होते हैं। विषमता की प्रमुख माप - कार्ल पियर्सन का निरपेक्ष माप, बाउले का विषमता माप और केली का विषमता माप। प्रथुशीर्षत्व की माप द्वारा हमें वक्र के माध्य अथवा शीर्ष के नुकीलेपन या चपटेपन की जानकारी प्राप्त होती है। इसे कार्ल पियर्सन द्वारा वक्र की उत्तलता (Convexity of the Curve) भी कहते हैं। इसकी माप से हमें श्रेणी के मध्य भाग में आवृत्तियों के जमाव का ज्ञात प्राप्त होता है। इसे β_2, γ_2 गुणांकों द्वारा चिन्हित करते हैं। यदि $\beta_2 = 3$ या $\gamma_2 = 0$ तो वक्र मध्यम शीर्ष वाला है। यदि $\beta_2 > 3$ या $\gamma_2 > 0$ तो वक्र नुकीले शीर्ष वाला है। यदि $\beta_2 < 3$ या $\gamma_2 < 0$ तो वक्र चपटे शीर्ष वाला है।

13.8 अभ्यास के लिए प्रश्न (Practice Questions)

वस्तुनिष्ठ प्रश्न :

1. विषमता किसी आवृत्ति बंटन के किस विशेषता को प्रकट करती है-
   i) आवृत्तियों के केन्द्रिकरण को ii) आवृत्तियों के अपकिरण को
   iii) किसी आवृत्ति बंटन के आकार को iv) इन सभी को

2) निम्न से किसने विषमता के माप को विकसित नहीं किया -
   i) क्राक्सटन ii) कार्ल पियर्सन
   iii) बाउले iv) केली

3) बाउले के विषमता गुणांक की क्या सीमा है-
   i) 0 से 1 ii) -1 से +1
   iii) 0 से 3 iv) -3 से +3

4) कार्ल पियर्सन के विषमता गुणांक का सूत्र है-

(i) $\dfrac{X - Z}{S.D.}$

(ii) $\dfrac{X - M}{}$

S.D.

(iii) $\frac{3\ (\bar{X} - Z)}{S.D.}$

(iv) $\frac{2\ (\bar{X} + M)}{S.D.}$

5) एक सममित वितरण में विषमता गुणांक होती है-

i) शून्य ii) धनात्मक
iii) ऋणात्मक iv) इनमें से कोई नहीं

6) वक्र के माध्य अथवा शीर्ष के नुकीलेपन या चपटेपन की जानकारी देता है-

i) पृथुशीर्षत्व ii) परिघात
iii) विषमता iv) समान्तर माध्य

7) यदि β2 > 3 है तो वक्र कैसा होगा

i) नुकीले शीर्ष वाला ii) चपटै शीर्ष वाला
iii) मध्यम शीर्ष वाला iv) इनमें से कोई नहीं

8) पृथुशीर्वत्च को वक्र की उत्तलता किसने कहा-

i) केली ii) बाउले
iii) कार्ल पियर्सन iv) इनमें से कोई नहीं

सही (True) अथवा गलत (False) चिन्हित करें -

(1) विषमता धनात्मक होती है, यदि X < M < Z ।

(2) बाउले का विषमता गुणांक चुतर्थकों और मध्यिका पर आधारित होता है।

(3) खुले सिरे वाले श्रेणियों के सन्दर्भ में बाउले का विषमता गुणांक कार्ल पियर्सन के विषमता गुणांक से बेहतर होता है।

(4) एक सममित वितरण में $Q_3 - M = M - Q_1$ होता है।

(5) सभी वितरणों को धनात्मक अथवा ऋणात्मक विषमता के रूप में विभाजित किया जा सकता है।

(6) मध्यम शीर्ष वाला वक्र प्रसामान्य वक्र कहलाता है।

(7) मध्यम शीर्ष वाले वक्र के सन्दर्भ में β2 < 3 होता है।

(8) द्वितीय केन्द्रिय परिघात प्रसरण को कहते है।

(9) एक सममित वितरण के सभी विषम परिघातों का मूल्य हमेशा शुन्य से अधिक होता है।

(10) चतुर्थ केन्द्रित परिघात पृथुशीर्षत्व की माप करता है।

13.9 अभ्यास प्रश्नों के उत्तर (Answer for Practice Questions)

वस्तुनिष्ठ प्रश्न :

1.iii) 2.i) 3.ii) 4.i) 5.i) 6.i) 7.i) 8.iii)

सही (True) अथवा गलत (False) चिन्हित करें -

1. False 2. True 3 True 4 True 5. False 6 True 7. False 8 True 9. False 10 True

13.10 संदर्भ ग्रन्थ (Bibliography)

- सुदामा सिंह, ओ0पी0 सिंह, वाई0 के0 सिंह (2002) - ***अर्थषास्त्रीय गणित एवं प्रारम्भिक सांख्यिकी,*** राधा पब्लिकेषन्स, नई दिल्ली।
- J. K. Sharma (2008) ***Business Statistics***, Dorling Kinderseley (India) Pvt. Ltd. (Pearson Education), Delhi.
- एस0एन0 लाल, एल0के0 चतुर्वेदी (2010) -***परिमाणात्मक विष्लेषण***, शिव पब्लिशिंग हाउस, इलाहाबाद।

13.11 निबंधात्मक प्रश्न (Essay Type Questions)

प्र0 1: विषमता से क्या अभिप्राय है? इसकी जाँच किस प्रकार से करते हैं?

प्र0 2: निम्न समकों के आधार पर कार्ल पियर्सन का विषमता गुणांक ज्ञात कीजिए।

x	58	59	60	61	62	63	64	65
f	10	18	30	42	35	28	16	8

प्र0 3: निम्न संमकों के आधार पर कार्ल पियर्सन का विषमता गुणांक ज्ञात कीजिए

x	0	10	20	30	40	50	60	70	80
f	150	140	100	80	80	70	30	14	0

प्र0 4: किसी बंटन में कार्ल पियर्सन का विषमता गुणांक 0. 40 है इसका प्रमाप विचलन 8 और माध्य 30 है। बंटन के बहुलक एवं मध्यका की गणना कीजिए।

प्र0 5: किसी बंटन में निम्न परिणाम प्राप्त हुए -

समान्तर माध्य = 45
मध्यका = 48
विषमता गुणांक= -0. 4
बंटन के प्रमाप विचलन की गणना कीजिए

प्र0 6: निम्न समंकों से बाउले का विषमता गुणांक ज्ञात कीजिए।

Class Interval	40-36	36-32	32-28	28-24	24-20	20-16	16-12	12-8	8-4
Frequency	2	6	10	12	15	30	18	10	6

प्र0 7: निम्न समूहों में किसका वितरण अधिक सममित हैं ?

i) माध्य = 22, मध्यिका = 24, मान विचलन = 10,
ii) माध्य = 22, मध्यिका = 21, मान विचलन = 12

प्र0 8: यदि मध्यिका का मान 24 तथा माध्य का मान 26 हो तो समूह की विषमता धनात्मक होगी अथवा ऋणात्मक ?

प्र0 9: निम्नलिखित आंकड़ों की सहायता से विचरण - गुणांक तथा विषमता - गुणांक की गणना कीजिए ।

वर्ष	1910	'11	'12	'13	'14	'15	'16	'17	'18	'19
गेहूँ का मूल्य सूचकांक	83	87	93	109	124	126	130	118	105	104

प्र 0 10: प्रथुशीर्षत्व की व्याख्या कीजिए। विषमता तथा प्रथुशीर्षत्व में अन्तर स्पष्ट कीजिए।

प्र 0 11: कार्ल पियर्सन के बीटा तथा गामा गुणांकों की व्याख्या कीजिए।

प्र 0 12: निम्न आवत्ति वितरण किस प्रकार का है -
i) चपटे शीर्ष
ii) नुकीले शीर्ष

X	0-10	10-20	20-30	30-40
f	1	3	4	2

हल: 2. $[J = 0.228]$ 3. $[J = -0.6622]$

4. $[M_0 = 26.8,\ M_e = 28.93]$ 5. $[\sigma = 22.5]$ 6. $[J_B = 0.188]$

9. $c.v = 4.4\%$ (विषमता गुणांक) 12. ii)