Dr. C.V. Raman University
Kargi Road, Kota, BILASPUR, (C. G.),
Ph. : +07753-253801, +07753-253872
E-mail : info@cvru.ac.in | Website : www.cvru.ac.in

Advanced Statistical Analysis

Institute of Open and Distance Education

Faculty of Commerce

Advanced Statistical Analysis

2MCOM3

DR. C.V. RAMAN UNIVERSITY

Chhattisgarh, Bilaspur

A STATUTORY UNIVERSITY UNDER SECTION 2(F) OF THE UGC ACT

2MCOM3

उच्च सांख्यिकीय विश्लेषण

2MCOM3, Advanced Statistical Analysis

Edition: March 2024

Compiled, reviewed and edited by Subject Expert team of University

1. Dr. Vivek Bajpai

(Professor, Dr. C. V. Raman University)

2. Dr. Preeti Shukla

(Associate Professor, Dr. C. V. Raman University)

3. Dr. Aashima Frinklin

(Assistant Professor, Dr. C. V. Raman University)

Published by:

Dr. C.V. Raman University

Kargi Road, Kota, Bilaspur, (C. G.),

Ph. +07753-253801, 07753-253872

E-mail: info@cvru.ac.in

Website: www.cvru.ac.in

अध्याय-1 सम्भावना सिद्धांत (THEORY OF PROBABILITY)

NOTES

इकाई की रूपरेखा

1.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि-

1. संभावना सिद्धांत के अर्थ एवं परिभाषा को समझ सकेंगे।
2. व्यवसाय व सामाजिक जीवन में संभावना सिद्धांत के महत्व एवं उपयोगिता को समझ सकेंगे।
3. संभाविता के नियम एवं प्रमेय का ज्ञान होगा।

1.1 प्रस्तावना

पूर्व में ऐसे समंकों के विश्लेषण का अध्ययन किया गया है जिन्हें उन साधनों से प्राप्त किया जाता है जो हमारे नियंत्रण से बाहर हैं, जैसे, आय के वितरण और मूल्यों में उच्चावचन। इस अध्याय में हम उन काल्पनिक (artificial) समंकों के विश्लेषण का अध्ययन करेंगे जिन्हें हम ऐसी रीतियों से प्राप्त करते हैं जो अधिकांशतः हमारे नियंत्रण में होती हैं। ऐसे समंकों के विश्लेषण से हमें कुछ निष्कर्ष प्राप्त होते हैं जो सांख्यिकीय अध्ययन के लिए काफी महत्वपूर्ण होते हैं। इन निष्कर्षों का अध्ययन संभावना सिद्धांत के अन्तर्गत किया जाता है। अतः हम संभावना सिद्धांत के विकास के विषय में कुछ जानना चाहेंगे।

1.2 संभावना सिद्धांत का विकास (Development of the Theory of Probability)

इस सिद्धांत के विकास का श्रेय 17वीं और 18वीं शताब्दी के सटोरियों (gamblers) को है। प्रारम्भ में ये सटोरिये भाग्य की देवी (Goddess of Fortune) पर ही विश्वास करते थे जो केवल अन्धविश्वास ही था। बाद में इनमें से कुछ लोग अपनी सट्टे सम्बन्धी समस्याओं को हल करने के लिए उस समय के प्रसिद्ध गणितवेत्ताओं की सहायता

NOTES

चाहने लगे। इस प्रकार उस समय के प्रसिद्ध गणितवेत्ता श्री गैलीलियो, पैस्कल, फरमेन्ट, आदि ने उन समस्याओं के समाधान की ओर अपनी मानसिक शक्तियों का उपयोग किया। इन्हीं के प्रयासों से सम्भाविता सिद्धांत का प्रादुर्भाव हुआ। इन विद्वानों ने अवसर (chance) सम्बन्धी कई समस्याओं पर बड़ी गहनता से विचार किया। धीरे-धीरे इस सिद्धांत के विकास को गति मिली। प्रसिद्ध विद्वान लेपलेस, गॉस, डेनियल बरनौली, आदि ने इस सिद्धांत को और विकसित किया और कई क्षेत्रों में इसका प्रयोग सफलतापूर्वक किया। इटली के श्री कारडेनो जो स्वयं एक सटोरिये थे, ने भी इस सिद्धांत के विकास में अपना योगदान एक बीजगणितीय समस्या के हल के रूप में प्रस्तुत किया। कुछ विद्वानों ने इस सिद्धांत का प्रयोग वित्तीय, सैनिक, प्रशासनिक और राजनीतिक क्षेत्र में भी सफलतापूर्वक किया। आधुनिक समय में सम्भाविता सिद्धांत पर आधारित निदर्शन सिद्धांत का प्रतिपादन श्री आर.ए. फिशर, पीयरसन और जे.नीमन आदि द्वारा किया गया है।

1.3 सम्भावना सिद्धांत का महत्व (Importance of the Theory of Probability)

श्री एमाइल बोरेल के शब्दों में इस सिद्धांत का महत्व इस प्रकार व्यक्त किया जा सकता है-

"Probability theory is of interest, not only to card and dice players, who were its godfathers, but also to all men of action, heads of industries or heads of armies, whose success depends on decisions, which in turn depend on two sorts of factors the one known or calculable, the other uncertain and problematical."

इस सिद्धांत का प्रयोग प्रायः सामान्य जीवन में होता है। मनुष्य का परिस्थितियों पर पूर्ण रूप से अधिकार नहीं होता है। इसी कारण अनिश्चितता रहती है और हमें सम्भावना का सहारा लेना पड़ता है। श्री बॉडिंगटन ने इसीलिए कहा है कि "सांख्यिकी अनुमानों और सम्भावनाओं का विज्ञान है" (Statistics is a science of estimates and probabilities)। यह सिद्धांत गणित पर आधारित होने से इस पर विश्वास किया जा सकता है और अनावश्यक परेशानी से बचा जा सकता है।

1.4 सम्भावना सिद्धांत का अर्थ (Meaning of the Probability Theory)

सम्भावना शब्द का अर्थ साधारण व्यवहार में होता है। जैसे, उसके जीतने की सम्भावना कम है, युद्ध होने की सम्भावना है या नहीं, आप सम्भवतः सही हैं, कल वर्षा होने की सम्भावना है, आदि। इन सभी वाक्यों में अनिश्चितता के विचार को मान्यता दी गई है। सांख्यिकीय व्यवहार में सम्भावना शब्द का अर्थ भिन्न होता है। साधारण व्यवहार में किसी विषय की सम्भावना व्यक्ति-विशेष के स्वयं के हित सम्बन्ध पर निर्भर करती है तथा समय और परिस्थिति के अनुसार उसमें परिवर्तन होता है। सांख्यिकीय व्यवहार में किसी की सम्भावना का अर्थ यह होता है कि यदि किसी विशिष्ट घटना से सम्बन्धित प्रयोग बार-बार किया जाये तो उस विशिष्ट घटना के घटित होने की सम्भावना क्या है? उदाहरणार्थ, यदि एक सिक्का ऊपर फेंका जाये तो उसके चित्त (head) या पट (tail) गिरने की सम्भावना बराबर-बराबर (1/2) यदि यह सिक्का संसार के किसी भी कोने में ले जाकर प्रयोग किया जाये तो सम्भावना 1/2 ही रहेगी। इसी प्रकार ताश की गड्डी में जिसमें 52 पत्ते होते हैं और जिनमें पान के 13 पत्ते होते हैं, अनायास खींचे जायें तो पान के पत्ते के खींचे जाने की सम्भावना 13/52 या 1/4 रहेगी।

इसलिए सांख्यिकी में सम्भावना का अर्थ यह है कि किये गये प्रयोग के फलस्वरूप किसी विशिष्ट घटना के होने की सम्भावना क्या है? विभिन्न सांख्यिकवेत्ताओं ने इसका अर्थ इस प्रकार दिया है-

(1) **लेप्लेस** (Laplace)- "सम्भावना, अनुकूल घटनाओं का होने वाली समस्त घटनाओं के साथ अनुपात है।"

"Probability is the ratio of favourable events to the total number number of equally likely events."

(2) **कॉनर** (Connor)- "सम्भावना, अनिश्चित घटनाओं के विषय में मस्तिष्क की एक स्थिति है।"

"Probability is an attitude of mind towards uncertain events."

इस प्रकार, सरल शब्दों में, सम्भावना सिद्धांत को इस प्रकार परिभाषित किया जाता है कि "यदि कोई समान ढंगों से घटित होने वाली घटनाएँ घटित हो सकती हैं, तो उनमें से किसी एक घटना के घटित होने की सम्भावना, अनुकूल ढंग से घटित होने वाली घटनाओं की संख्या का समस्त सम्भव घटनाओं से अनुपात होता है।"

"If there are several equally likely events that may happen, the probability that any one of these events will happen, is the ratio of the number of cases favourable to its happening to the total number of possible cases.

परिभाषाएँ
(Definition)

(1) **गणितीय परिभाषा** (Mathematical Definition)- "कोई विशिष्ट घटना m प्रकार से हो सकती है और यदि वह घटना n प्रकार से नहीं हो सकती हो तो उस विशिष्ट घटना के होने की सम्भावना या 'p' = $\frac{m}{m+n}$ है और उस घटना के न होने की सम्भावना या 'q' = $\frac{n}{m+n}$ है।"

उदाहरणार्थ एक थैले, में 12 गेंदे हैं, 4 का रंग हरा और 8 का रंग लाल है। सभी गेंदों का आकार समान है। यदि थैले में से बिना देखे एक गेंद निकाली जाये तो हरे रंग की गेंद के निकलने की क्या सम्भावना है? यहाँ हरे रंग की गेंद चार हैं अतः हरे रंग की गेंद के निकलने की घटना चार प्रकार से हो सकतीहै। लाल गेंद आठ हैं, अतः हरे रंग की गेंद न निकलने की घटना आठ प्रकार से हो सकती है।

इस प्रकार यह घटना 4 प्रकार से हो सकती है और 8 प्रकार से नहीं हो सकती। अतः $m=4$ और $n=8$ है। उपरोक्त सूत्र के अनुसार-

p (घटना होने की सम्भावना) $= \frac{m}{m+n} = \frac{4}{4+8} = \frac{1}{3}$

q (घटना न होने की सम्भावना) $= \frac{n}{m+n} = \frac{8}{4+8} = \frac{2}{3}$,

कोई भी घटना या तो हो सकती है या नहीं हो सकती है। अतः उस घटना के होने या न होने का योग '1' है। दूसरे शब्दों में $(p + q) = 1$। अतः $q = 1 - p$ और $p = 1 - q$ इस प्रकार किसी भी घटना के होने या न होने की सम्भावना 0 से 1 के बीच होगी। किसी घटना की सम्भावना यदि शून्य (0) है तो उसका होना असम्भव है और यदि सम्भावना '1' है तो उस घटना का होना निश्चित है। हमें 'विशिष्ट घटना का घटित होना' वाक्यांश का महत्व भली-भाँति जान लेना चाहिये। जैसे एक नगर में 20 से 25 वर्ष की आयु के लोगों में से आधे पुरुष और आधे स्त्रियाँ हैं। एक व्यक्ति जिसकी आयु 20 वर्ष है का नाम 'क' है। उसके स्त्री या पुरुष होने की सम्भावना क्या है? सम्भाविता सिद्धांत के अनुसार वह 1/2 है। किन्तु यह वास्तव में सम्भव नहीं है। वह या तो पुरुष या स्त्री हो सकता है। इससे यह निष्कर्ष निकाला जा सकता है कि जिन विषयों के सम्बन्ध में निश्चित रूप से निर्णय लिये जा सकते हैं उन विषयों में सम्भाविता सिद्धांत का प्रयोग अवांछनीय है।

सम्भावना की गणितीय परिभाषा की सीमायें- (1) समान रूप से घटित होने वाली स्थितियाँ सर्वत्र सम्भव नहीं होतीं।

(2) इस परिभाषा की उपयोगिता उस समय समाप्त हो जाती है जबकि सम्भव परिणामों की कुल संख्या अनिश्चित हो।

(3) यह भी सम्भव है कि समान रूप से घटित होने वाली स्थितियों की गणना सम्भव न हो।

अपनी प्रगति की जाँच करें
Test your Progress

(2) **सम्भावना की सांख्यिकीय परिभाषा** (Statistical Definition of Probability)- 'यदि परीक्षणों को सार रूप से समान स्थितियों में बहुत अधिक बार दोहराया जाये तो सम्पूर्ण परीक्षणों की अनन्त रूप से बढ़ती हुई संख्या की तुलना में किसी एक घटना के घटने की अनुपात सीमा उस घटना के घटित होने की सम्भावना कहलाती है। यह कल्पना कर ली जाती है कि यह अनुपात एक निश्चित और अपूर्व सीमा को प्राप्त करता है।'

श्री एम. वैद्यनाथन ने सम्भावना की सांख्यिकीय परिभाषा इस प्रकार दी है : "If a large number of trails be made under constant conditions, then the ratio of the number of trails, in which a certain event happens to the total number of trails (N) would approach a limit as N is indefinitely increased and this limit is the probability of the event happening."

सांख्यिकीय परिभाषा की सीमाएँ

इस परिभाषा की सबसे बड़ी त्रुटि यह है कि-

(i) यह कैसे जाना जावे कि जाँचों (परीक्षणों) की परिस्थितियाँ आवश्यक रूप से एक समान रही हैं या नहीं। वैज्ञानिक प्रयोगों के क्षेत्र में निरन्तर अनुसंधान होते रहते हैं जिसके परिणामस्वरूप परीक्षणों की परिस्थितियाँ बदलती रहती हैं। विशेषतः चिकित्सा सम्बन्धी विज्ञानों में परिस्थितियाँ एक समान नहीं रहती हैं।

(ii) सभी परीक्षणों की अनन्त संख्या, वास्तव में, एक सैद्धांतिक एवं अव्यवहारिक धारणा है।

(iii) सांख्यिकीय सम्भावना का माप पूर्णतः शुद्ध नहीं होता। यह अनुभव एवं प्रयोग पर आधारित होता है, इसलिए यह अनुमानतः ही सत्य होता है। संभावना की यह परिभाषा इसकी गणितीय परिभाषा का ही संशोधित रूप है।

1.5 सम्भावना-मापदण्ड (Probability Scale)

जैसा कि बतलाया गया है, किसी घटना के घटित होने की सम्भावना का माप (p) 0 से लेकर 1 तक हो सकता है। जब सम्भावना का माप शून्य (p = 0) होता है तो सम्बद्ध घटना असम्भव होती है। किसी व्यक्ति अटलांटिक महासागर तैर कर पार करने की सम्भावना '0' है। सम्भावना-मापदण्ड के दूसरी ओर पूर्ण-निश्चितता की स्थिति होती है, जैसे प्रत्येक मनुष्य के मरने की सम्भावना '1' है अर्थात् p = 1। एक अनभिनत सिक्का उछाले जाने पर उसके सिर (head) आने की सम्भावना 1/2 या 0.5 या 50% है। ताश के पत्तों में से किसी 'इक्के' के निकाले जाने की सम्भावना 4/52 या 1/13 या 0.077 या 7.7% है। लाल पान के किसी एक पत्ते को निकाले जाने की सम्भावना 13/52 या 1/4 या 25% या 0.25 होगी। इस प्रकार सम्भावना का माप भिन्न के रूप में या, दशमलव में या प्रतिशत के रूप में प्रस्तुत किया जा सकता है।

Illustration 1. What is the chance of drawing king in a draw from a pack of 52 cards.

ताश के बावन पत्तों की गड्डी में से एक बादशाह के छांटे जाने की प्रायिकता क्या है ?

Solution-

Total No. of cases that can happen = 52 (n)

No. of favourable cases i.e., total No. of king in the pack of cards = 4 (p)

$$\therefore \text{ The probability } = \frac{P}{n} = \frac{4}{52} = \frac{1}{13}$$

Illustration 2. A bag contains 5 black and 10 white balls. What is the probability of drawing?

(i) a black ball, and (ii) a white ball.

एक थैले में 5 काली तथा 10 सफेद गेंदें हैं। (i) एक काली तथा (ii) एक सफेद गेंद छांटे जाने की संभावना (प्रायिकता) क्या है ?

Solution-

$$\text{(i)} \quad p(\text{black}) = \frac{\text{No of black balls}}{\text{Total No. of balls}} = \frac{5}{15} = \frac{1}{3}$$

$$\text{(ii)} \quad p(\text{white}) = \frac{\text{No of white balls}}{\text{Total No. of balls}} = \frac{10}{15} = \frac{2}{3}$$

Illustration 3. The Registrar of vital statistics in a country reported 6,858 sons and 6,543 daughters for a specific region. Treating this as a fair sample from the general population, estimate the probability that the child to be born will be a boy.

Solution- The event can happen in 6,858 ways and can fail in 6,543 ways. Hence,

$$P = \frac{6858}{6858 + 6543} = \frac{6858}{13401}$$

Illustration 4. What is the probability that a vowel selected at random in any English book is an 'e'? इस बात की संभावना क्या है कि अंग्रेजी की एक पुस्तक से यादृच्छिक रीति से छांटा गया स्वर 'e' होगा ?

Solution- Since there are five vowels, total no. of equally likely events = 5 No. of favourable events = 1

$$P = \frac{\text{No. of favourable events}}{\text{Total No. of equally likely events}} = \frac{1}{5}$$

Illustration 5. What is the probability of an ace in a throw of one dice? (एक पांसे के फेंकने पर इक्का अंकित होने की संभावना क्या है ?

Solution- First of all, all the possible cases with one dice will be six- the number of its sides.

Secondly, the number of favourable cases is one because only one side can come upward which is the same as the extent, in question. Hence, the probability we are looking for is the fraction whose numerator is one and whose denominator is six i.e., it is $\frac{1}{6}$.

Illustration 5 (A) : Out of a pack of cards what is the probability of drawing (a)a king; (b) a king of hearts, (c) any card of hearts.

ताश की एक गड्डी से निम्नलिखित के खींचे जाने की क्या संभावना है ?

(अ) एक बादशाह (ब) लाल का बादशाह (स) लाल का कोई पत्ता।

Solution : (a) Since there are 4 kings in 52 cards, Hence

$$p = \frac{4}{52} = \frac{1}{13}.$$

(b) Since there is only 1 king of hearts, $p = \frac{1}{52}$

(c) Since there are 13 cards of hearts, $p = \frac{13}{52} = \frac{1}{4}$

Illustration 5(B) – (i) What is the probability of throwing a number greater than 3 with an ordinary dice whose faces are numbered from 1 to 6?

एक साधारण पासा जिस पर 1 से 6 तक बिन्दु अंकित हैं फेकने पर 3 से अधिक अंकित बिन्दु प्राप्त होने की संभावना क्या हैं ?

(ii) If two dice are thrown together, what is the probability of throwing a total of 7?

यदि दो पाँसे एक साथ फैंके जाते हैं तो इस बात की संभावना क्या है कि उनसे 7 का योग प्राप्त होगा ?

(iii) What is the probability that a digit selected at random from the logarithmic table is (a) 4, (b) 5 or 8?

इस बात की क्या संभावना है कि यादृच्छित रीति (दैव निदर्शन रीति) से छेदा सारणी से छांटा गया कोई अंक (अ) 4, (ब) 5 या 8 होगा ?

Solution –

(i) Total No. of cases = 6

No. of favourable cases (4, 5 & 6) = 3

$$\therefore \quad p = \frac{3}{6} = \frac{1}{2}$$

(ii) Total No. of cases = 6 × 6 = 36

NOTES

No. of favourable cases

[(1 + 6), (2 + 5), (3 + 4), (4 + 3), (5 + 2), (6 + 1) = 6

$\therefore \quad p = \frac{6}{36} = \frac{1}{6}$

(iii) Total no. of digits from 0 to 9 = 10

(a) No. of favourable cases = 1, $p = \frac{1}{10}$

(b) No. of favourable cases = 2, $p = \frac{2}{10}$ or $\frac{1}{5}$

Illustration 5 (C) : The following table gives distribution of marks :

निम्न सारणी में अंकों का वितरण इस प्रकार दिया हुआ है :

Marks (प्राप्तांक) :	5	10	15	20	25
No. of Students (छात्र संख्या) :	45	98	28	16	13

A student is taken at random from the above group. Find the probability that (i) his marks are under 15, (ii) his marks are 15 or more.

उपरोक्त छात्र समूह से एक छात्र दैव निदर्शन रीति से लिया जाता है। संभावना ज्ञात कीजिए कि (i) उसके प्राप्तांक 15 से कम हैं (ii) उसके प्राप्तांक 15 या अधिक हैं।

Solution –

In this group the total No. of cases ie. students are

$$45 + 98 + 28 + 16 + 13 = 200$$

(i) No. of Students getting marks below 15 are 45 + 98 = 143.

$\therefore$ Required Probability $= \frac{143}{200}$

(ii) No. of Students getting marks 15 or above are

28 + 16 + 13 = 57

$\therefore$ Required Probability $= \frac{57}{200}$

1.6 घटनाएँ (Events)

(1) **समान रूप से घटित होने वाली घटनाएं** (Equally likely events)- सम्भाविता सिद्धांत का वर्णन करते समय इस वाक्यांश का प्रयोग किया गया है। ये वह घटनाएं होती हैं जिनमें से कोई भी घटना घटित हो सकती है। सिक्का उछालते समय वह या तो चित्त पड़ेगा या पट। इस प्रकार समान रूप से घटित होने वाली घटनाएं दो हैं। यह निम्न उदाहरण से स्पष्ट हो जायेगा-

'अ' और 'ब' दो खिलाड़ी इन नियमों के अनुसार एक सिक्का उछाल रहे हैं कि वह दो बार उछाला जायेगा। यदि किसी भी उछाल में सिक्का चित्त पड़ता है तो 'अ' विजयी होगा और यदि चित नहीं पड़ता है तो 'ब' विजयी घोषित किया जायेगा।

इस उदाहरण में समान रूप से घटित होने वाली परिस्थितियाँ निम्न होंगी-

प्रथम उछाल	**द्वितीय उछाल**	
चित	चित	'अ' की विजय
चित	पट	'अ' की विजय
पट	चित	'अ' की विजय
पट	पट	'ब' की विजय

अतः 'अ' के विजयी होने की सम्भावना $\frac{3}{4}$ है।

(2) **परस्पर अपवर्जी घटनाएं** (Mutually Exclusive Events)- यह ऐसी घटनाएं होती हैं जिसके घटित होने से दूसरे की सम्भावना समाप्त हो जाती है। उदाहरणार्थ, सिक्का यदि उछाला जाये तो उसके चित पड़ने से पट नहीं पड़ सकेगा और पट पड़ने से वह चित नहीं पड़ सकेगा।

(3) **स्वतंत्र घटनाएं** (Independent Events)- घटनाएँ स्वतंत्र तब मानी जाती हैं जबकि एक घटना के घटित होने से दूसरी घटनाओं के घटित होने पर कोई प्रभाव नहीं पड़ता। दो घटनाएँ तब स्वतंत्र मानी जाती हैं जबकि उनका एक दूसरे पर कोई प्रभाव नहीं होता, जैसे कि सिक्के प्रथम उछाल के परिणाम का दूसरे उछाल पर कोई प्रभाव नहीं होता है।

(4) **आश्रित घटनाएं** (Dependent Events)- यदि एक घटना दूसरी को प्रभावित करती है तो उन्हें आश्रित घटनाएँ कहते हैं। जैसे 52 पत्तों में से 'राजा' के आने की संभावना 1/13 होती है। यदि एक बार राजा निकल जाये और उसे पुनः पत्तों में न मिलाया जाये तो दुबारा 'राजा' निकलने की सम्भावना 3/51 होगी। इस प्रकार प्रथम घटना ने यहाँ द्वितीय घटना को प्रभावित किया है अतः ये आश्रित घटनाएँ हैं।

(5) **सरल एवं संयुक्त घटनाएँ** (Simple and Compound Events)- सरल घटनाओं के लिए किसी घटना के घटित होने या घटित नहीं होने की संभावना पर विचार किया जाता है। ऊपर उदाहरण 1 तथा 2 सरल घटनाओं के ही उदाहरण हैं। 52 पत्तों की एक गड्डी में से बादशाह के खींचे जाने की संभावना $\left(\frac{4}{52}\right)$ सरल घटना का एक उदाहरण है इसी प्रकार एक थैले में 5 लाल और 3 काली गेंदें हों तो एक काली गेंद निकालने की संभावना ज्ञात करना भी सरल घटना ही है।

संयुक्त घटना में दो या दो से अधिक घटनाओं के संयुक्त रूप से घटने की संभावना का अध्ययन किया जाता है। दो पाँसों को एक साथ फैंकना या तीन सिक्कों को एक साथ उछालना और फिर कुछ संभावनाएं निकालना-संयुक्त घटनाओं के उदाहरण हैं।

बोध प्रश्न

1. संभावना सिद्धान्त से आप क्या समझते हैं।

....

....

....

2. संभावना सिद्धान्त की गणितीय परिभाषा की सीमाएँ लिखिए।

....

....

....

3. परस्पर अपवर्जी घटनाओं से आप क्या समझते हैं।

....

....

....

1.7 संभाविता के नियम या प्रमेय (Probability Rules or Theorems)

(1) **संभाविता के योग का नियम** (Rule of Addition of Probability)- कभी-कभी यह जानना आवश्यक होता है कि किन्हीं दो या दो से अधिक घटनाओं में से किन्हीं दो या दो से अधिक घटनाओं के होने की संभावना क्या है? तब हमें इस नियम का उपयोग करना पड़ता है। ये दो या दो से अधिक घटनाएँ अपवर्जी (Mutually Exclusive) होती हैं। घटनाएँ अपवर्जी तभी कहलायेंगी जबकि उनका एक साथ घटित होना असम्भव हो अर्थात् उनमें से एक बार

केवल एक ही घटना घटित हो सकती है। ऐसी परिस्थिति में उन दो या दो से अधिक घटनाओं की अलग-अलग सम्भावना निकाल कर उनका योग करना पड़ता है। इसे सूत्र के रूप में इस प्रकार प्रस्तुत किया जा सकता है-

$$P(A_1 + A_2 \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$$

where,

$P(A_1 + A_2 \ldots + A_n)$ = दो या दो से अधिक घटनाओं की सम्मिलित सम्भावना।

$P(A_1) + P(A_2) \ldots\ldots + P(A_n)$ = घटनाओं की अलग-अलग सम्भावनाएँ।

Illustration 6. If dice having six sides and marked 1 to 6 is thrown, what is the probability of getting 1 or 2?

Solution - We shall have to calculate the probability of getting 1 and probability of getting 2 seperately and then add them-

Probability of getting 1 is $\frac{1}{6}$; Probability of getting 2 is $\frac{1}{6}$

Substituting the above values in the formula,

$$P(A_1 + A_2) = P(A_1) + P(A_2)$$

$$P(A_1 + A_2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

$\therefore$ The probability of getting 1 or 2 $= \frac{1}{3}$

Illustration 7. A bag contains 8 white, 4 black, 6 yellow and 6 red balls. What is the probability of getting a white or red ball at random in a single draw of one?

एक थैले में 8 सफेद, 4 काली, 6 पीली और 6 लाल गेंदें हैं। एक बार निकालने पर एक सफेद या एक लाल गेंद के निकलने की क्या संभावना है ?

Solution-

The probability of getting one white ball $= \frac{8}{24} = \frac{1}{3}$

" " " red " $= \frac{6}{24} = \frac{1}{4}$

Hence " " " white or red ball $= \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$

Illustration 8 (A). There are three events. X, Y and Z of which only one can happen. The odds against X are 4 to 1 against Y are 5 to 3. Find the odds against Z.

Solution-

Probability of X's happening is $= \frac{1}{5}$ Probability of Y's happening is $= \frac{3}{8}$

Probability of Z's happening is $= 1 - \left(\frac{1}{5} + \frac{3}{8}\right) = \frac{23}{40}$

Therefore the odds against Z are 17 to 23.

Illustration 8 (B) : From a pack of 52 cards one card is drawn at random. What is the probability that it will be a queen of clubs or a king of diamonds?

Solution : Drawing a queen of Clubs or a king of diamonds are mutually exclusive events.

$\therefore$ Probability of drawing a queen of Clubs $= \frac{1}{52}$

and Probability of drawing a king of diamonds $= \frac{1}{52}$

$\therefore$ The Probability of drawing a queen of clubs or a king of diamonds is $= \frac{1}{52} + \frac{1}{52} = \frac{2}{52} = \frac{1}{26}$

Illustration 8 (C) : (i) What is the Probability of getting at least total of 8 by throwing two dice once?

दो पाँसों को एक बार फैंकने पर कम से कम 8 का योग आने की संभावना क्या है ?

दो पाँसों को एव बार फेकने पर कम से कम 8 का योग आने की संभावना क्या है ?

(ii) Find out the probability of getting the total of 10 or 12 by throwing two dice.

दो पाँसों के फैंकने पर 10 या 12 के योग आने की संभावना बतलाइये।

Solution- दो पाँसों के एक साथ फैंकने पर कुल $6 \times 6 = 36$ नतीजे प्राप्त हो सकते हैं जैसा कि निम्नलिखित तालिका से स्पष्ट है :

दूसरे पाँसे का परिणाम	पहले पाँसे का परिणाम						योग
	1	2	3	4	5	6	
1	1, 1	1, 2	1, 3	1, 4	1, 5	1, 6	6
2	2, 1	2, 2	2, 3	2, 4	2, 5	2, 6	6
3	3, 1	3, 2	3, 3	3, 4	3, 5	3, 6	6
4	4, 1	4, 2	4, 3	4, 4	4, 5	4, 6	6
5	5, 1	5, 2	5, 3	5, 4	5, 5	5, 6	6
6	6, 1	6, 2	6, 3	6, 4	6, 5	6, 6	6
योग	6	6	6	6	6	6	36

(i) कम से कम 8 का योग प्राप्त करने का अर्थ हुआ - 8 या 9 या 10 या 11 या 12 का योग प्राप्त करना।

8 का योग प्राप्त करने की संभावना (2, 6) (3, 5) (4, 4) (5, 3) (6, 2) $= \frac{5}{36}$

9 का योग प्राप्त करने की संभावना (3, 6) (4, 5) (5,4) (6, 3) $= \frac{4}{36}$

10 का योग प्राप्त करने की संभावना (4, 6) (5, 5) (6, 4) $= \frac{3}{36}$

11 का योग प्राप्त करने की संभावना (5, 6) (6, 5) $= \frac{2}{36}$

12 का योग करने की संभावना (6, 6) $= \frac{1}{36}$

$\therefore$ कम से कम 8 का योग प्राप्त करने की संभावना होगी

$$= \frac{5}{36} + \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{15}{36} = \frac{5}{12}$$

(ii) दो पाँसों के फैंकने पर-

10 का योग प्राप्त करने की संभावना (4, 6) (5, 5) (6, 4) $= \frac{3}{36}$

12 का योग प्राप्त करने की संभावना (6, 6) $= \frac{1}{36}$

10 या 12 का योग प्राप्त होने की संभावना होगी $= \frac{3}{36} + \frac{1}{36} = \frac{4}{36}$ या $\frac{1}{9}$

NOTES

Illustration 8 (D) - (i) In an examination the chance of A's Success is $\frac{1}{8}$, B's Success $\frac{7}{50}$ and C's Success is $\frac{1}{6}$. Find the chance that one of them will succeed. एक परीक्षा में 'अ' की सफलता की संभावना $\frac{1}{8}$ 'ब' की $\frac{7}{50}$ तथा 'स' की $\frac{1}{6}$ हो तो इस बात की संभावना बतलाइये कि उनमें से एक सफल हो जावेगा।

(ii) A bag contains 20 balls numbered from 1 to 20. Find the probability that a ball drawn is a multiple of 3 or 5. एक थैले में 20 गेंद हैं जिन पर 1 से 20 तक नंबर अंकित हैं। यह संभावना बतलाइये कि एक गेंद के निकालने पर उसका नंबर 3 या 5 का गुणक होगा।

(iii) A bag contains 50 balls numbered from 1 to 50. One ball is drawn at random. Find the probability that the number is a multiple of (a) 10 or 12, (b) 10 or 14. एक थैले में 50 गेंदे हैं जिन पर 1 से 50 तक नंबर अंकित हैं। दैव निदर्शन रीति से एक गेंद निकाली जाती है। इस बात की संभावना बतलाइये कि गेंद का नंबर (अ) 10 या 12 (ब) 10 या 14 का गुणक है।

(iv) A bag contains 50 balls numbered from 1 to 50. One ball is drawn at random. Find the probability that the number of the drawn ball is a multiple of 2 or 3 or 10. एक थैले में 50 गेंदें हैं जिन पर 1 से 50 तक नंबर अंकित हैं। दैव निदर्शन रीति से एक गेंद निकाली जाती है। इसकी संभावना बतलाइये कि निकाली गई गेंद की संख्या 2 या 3 या 10 का गुणक अंक है।

(v) What is the probability of drawing a card of diamond or an ace from a pack of cards? ताश की एक गड्डी से एक ईट का पत्ता या एक इक्का खींचने की संभावना क्या है ?

(vi) A person is said to hit the target in 3 out of 4 shots, whereas another person is said to hit at all when they both try. एक व्यक्ति के विषय में कहा जाता है कि वह 4 में से 3 लक्ष्य भेद सकता है जबकि एक दूसरा व्यक्ति 3 में से 2 लक्ष्य भेद सकता है। यदि वे दोनों प्रयास करें तो लक्ष्य भेदे जाने की संभावना बतलाए।

Solution-

(i) These events are mutually exclusive, the probability that one of them will succeed is

$$\frac{1}{8}+\frac{7}{50}+\frac{1}{6}=\frac{75+84+100}{600}=\frac{259}{600}$$

(ii) The probability that the number is a multiple of

$$3=(3, 6, 9, 12, 15, 18)=\frac{6}{20} \quad \text{or} = \frac{3}{10}$$

The Probability that the number is a multiple of

$$5 = (5, 10, 15 \,\&\, 20)=\frac{4}{20} \text{ or } \frac{1}{5}$$

$\therefore$ The Probability is a multiple of 3 or 5 $=\frac{3}{10}+\frac{1}{5}=\frac{5}{10}\ or\ \frac{1}{2}$

But this is not correct because number 15 is not mutually exclusive. It is a multiple of 3 as well as of 5. It is included in both.

Hence, the correct probability that the number is a multiple of 10

$$= (10, 20, 30, 40 \,\&\, 50) = \frac{5}{50}$$

The Probability that the number is a multiple of 12

$$= (12, 24, 36, \& 48) = \frac{4}{50}$$

As the events are mutually exclusive, the probability that the number will be a multiple of 10 or 12

$$= \frac{5}{50} + \frac{4}{50} = \frac{9}{50}$$

(b) The probability that the number is a multiple of 10

$$= (10, 20, 30, 40 \& 50) = \frac{5}{50}$$

The probability that the number is a multiple of 14 $= (14, 28, 42) = \frac{3}{50}$

The events are mutually exclusive, hence, the probability that number will be a multiple of 10 or 14

$$= \frac{5}{50} + \frac{3}{50} = \frac{8}{50}$$

(iv) The probability of multiple of 2

$$= (2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, \text{and } 50) = \frac{25}{50}$$

The probability of multiple of 3

$$= (3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48) = \frac{16}{50}$$

The probability of multiples of 10 $= (10, 20, 30, 40 \& 50) = \frac{5}{50}$

The common multiple of 2 and 3 are :

$$(6, 12, 18, 24, 30, 36, 42, 48) = \frac{8}{50}$$

The common multiple of 2 and 10 are : $(10, 20, 30, 40, 50) = \frac{5}{50}$

The common multiples of 3 & 10 are $= \frac{1}{50}$

The common multiples of 3 and 10 are 30 $= \frac{1}{50}$

Hence, the probability of the number being a multiple of 2, 3 or 10

$$= \frac{25}{50} + \frac{16}{50} + \frac{5}{50} - \frac{8}{50} - \frac{5}{50} - \frac{1}{50} + \frac{1}{50} = \frac{33}{50}$$

Note : जब दो घटनाओं A तथा B में से A अथवा B या A और B दोनों घट सकती हों तो घटनाएं A अथवा B पूर्णतः परस्पर अपवर्जी नहीं कहलाएँगी। अतः योग के नियम को संशोधित रूप में निम्नलिखित प्रकार से प्रयुक्त किया जावेगा। जैसा ऊपर (ii) में किया गया है-

$$p\,(A \text{ or } B) = p\,(A) + p\,(B) - p\,(A = B)$$

इसी प्रकार तीन घटनाओं की दशा में सूत्र इस प्रकार होगा जैसा कि ऊपर (iv) में किया गया है-

$$p\,(A \text{ or } B \text{ or } C) = p(A) + p(B) + p\,(C) - p(A+B) - p\,(A+C) - p\,(B+C) + p\,(A+B+C)$$

NOTES

(v) Probability of drawing a card of Diamond $= \frac{13}{52}$ *or* $\frac{1}{4}$

Probability of Drawing a card of ace $= \frac{4}{52}$ *or* $\frac{1}{13}$

But in 4 aces, one ace of diamond is included which was also included in 13 cards of diamond. Hence, the probability of drawing an ace of diamond which is included in both $= \frac{1}{52}$

The probability of drawing a card of diamond or an ace

$$= \frac{1}{4} + \frac{1}{13} - \frac{1}{52} = \frac{13 + 4 - 1}{52} = \frac{16}{52} \text{ or } \frac{4}{13}$$

(vi) The probability of first person hitting the target $= \frac{3}{4}$

The probability of second person hitting the target $= \frac{2}{3}$

As both them may hit the target, the events are not mutually exclusive. Hence, the probability of hitting the target.

$$= \frac{3}{4} + \frac{2}{3} - \left(\frac{3}{4} \times \frac{2}{3}\right) = \frac{11}{12}$$

(2) सम्भावित के गुणाकार का नियम (Rule of Multiplication of Probability)

इस नियम का प्रयोग सम्भाविता के योग के नियम की विपरीत परिस्थिति में होता है। जब हम यह जानना चाहते हैं कि दो या दो से अधिक घटनाओं की एक साथ या एक के बाद एक निश्चित क्रम के होने की सम्भावना क्या है, तो इस नियम का प्रयोग करते हैं। यह घटनाएँ स्वभाव से ही आपस में स्वतंत्र होती हैं। वे एक दूसरे पर किसी प्रकार निर्भर नहीं रहतीं। जैसे- एक थैली में दो रंग की कुछ गेंद हैं। उसमें से किसी एक रंग के दो या दो से अधिक गेंद निकालने की क्या सम्भावना है? इस परिस्थिति में हर घटना की सम्भावना ज्ञात कर उन सब सम्भावनाओं का गुणाकार किया जाता है। गुणाकार नियम इस प्रकार होता है-

$$P(A_1 A_2 A_3 \ldots A_n) = P(A_1) \times P(A_2) \times P(A_3) \times \ldots P(A_n)$$

Where, P $(A_1 A_2 A_3 \ldots A_n)$ = विभिन्न घटनाओं की एक साथ सम्भावना

$P(A_1) \times (A_2) \times (P(A_3) \times \ldots . P(A_n)$ = विभिन्न घटनाओं की अलग-अलग सम्भावना।

यदि घटनाओं की संख्या कम हो तो इस रीति से सम्भावना ज्ञात करना सरल होता है। यदि उनकी संख्या अधिक हो तो सम्भावना ज्ञात करना बहुत कठिन होता है।

Illustration 9. What is the probability of getting all the tails upward in four throws of a coin?

एक सिक्के को चार बार उछालने पर सभी बार चित पढ़ने की क्या संभावना है?

Solution-

The probability of getting tail in the Ist throw $= \frac{1}{2}$

The probability of getting tail in the IInd throw $= \frac{1}{2}$

The probability of getting tail in the IIIrd throw $= \frac{1}{2}$

The probability of getting tail in the IVth throw $- \frac{1}{2}$

The probability of getting tails upward in all the four throws.

$$=\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{16}$$

Illustration 10. What is the probability of throwing three fours in three throws of a dice?

पाँसा तीन बार फेंकने पर तीनों बार चार आने की संभावना क्या है।

Solution-

The probability of a 'four' in Ist throw $=\frac{1}{6}$

The probability of a 'four' in IInd throw $=\frac{1}{6}$

The probability of a 'four' in IIIrd throw $=\frac{1}{6}$

The probability of three 'four' $=\frac{1}{6}\times\frac{1}{6}\times\frac{1}{6}=\frac{1}{216}$

Illustration 11. In a bag there are 10 green and 15 red balls. If they are drawn at a time, and the ball is replaced after each draw, what is the probability of getting the first ball green, second red and the third green.

एक थैले में 10 हरी तथा 15 लाल गेंदें हैं। यदि उन्हें एक बार निकालने पर वापस अंदर रख दिया जाता है तो प्रथम गेंद के हरी, दूसरी के लाल और तीसरे के हरी होने की संभावना क्या है?

Solution- As the balls are replaced after each draw, so each event is independent.

The probability that the first ball will be green is $=\frac{10}{10+15}=\frac{2}{5}$

The probability that the second ball will be red is $=\frac{15}{10+15}=\frac{3}{5}$

The probability that the third ball will be green is $=\frac{10}{10+15}=\frac{2}{5}$

Hence, desired probability is $\frac{2}{5}\times\frac{3}{5}\times\frac{2}{5}=\frac{12}{125}$

Illustration 12. In a bag, there are 10 black and 10 green balls. Calculate the probability of drawing the black balls in succession if the ball drawn first is not replaced. (एक थैले में 10 काली और 10 हरी गेंदें दोनों बार लगातार काली गेंद निकलने की संभावना क्या है, यदि प्रथम बार निकाली गेंद वापस थैले में नहीं रखी जाती है?)

Solution- Here, the second event is dependent and the first is independent. The probability of getting black ball in the first attempt.

$$=\frac{10}{10+10}=\frac{1}{2}$$

The probability of black ball in the second attempt. $=\frac{9}{10+9}=\frac{9}{19}$

The probability of getting two black ball in succession $=\frac{1}{2}\times\frac{9}{19}=\frac{9}{38}$

NOTES

Illustration 13. From a pack of cards, three cards are drawn without being replaced. What is the probability that they will have the same value? (ताश की गड्डी से बिना वापस रखे तीन पत्ते निकाले जाते हैं। क्या संभावना है कि वे एक ही अंक वाले होंगे ?)

Solution- As the first card may be of any value, the probability of getting any card in the first draw is certain, i.e., $\frac{52}{52}$

In the second draw the probability of getting a card of the same value $\frac{2}{51}$ In the third draw, the probability of getting a card of the same value 1/50. So, the probability that all the three cards will be of the same value.

$$= \frac{52}{52} \times \frac{2}{51} \times \frac{1}{50} = \frac{2}{2550} = \frac{1}{1275}$$

Illustration 14. A problem in statistics is given to three students A, B and C whose chances of solving it are 1/2, 1/3, 1/4 respectively. What is the probability that the problem will be solved?

सांख्यिकी का एक प्रश्न तीन छात्र-अ, ब, स को हल करने के लिए दिया गया है जिनके हल करने की संभावना क्रमशः $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ है। प्रश्न के हल होने की संभावना क्या है ?

Solution- Probability that student 'A' will fail to solve the problem

$$= 1 - \frac{1}{2} = \frac{1}{2}$$

Probability that student B will fail to solve the problem $= 1 - \frac{1}{3} = \frac{2}{3}$

Probability that student C will fail to solve the problem $= 1 - \frac{1}{4} = \frac{3}{4}$

Since the events are independent, the probability that all the three students will fail to solve the problem.

$$= \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = \frac{1}{4}$$

The probability that the problem will be solved $= 1 - \frac{1}{4} = \frac{3}{4}$

Illustration 15. (i) 'A' can solve 75% of the problems is STATISTICS and 'B' can solve 70% what is the probability that either 'A' or 'B' can solve a problem chosen at random.

'अ' सांख्यिकी में 75% प्रश्न हल कर सकता है और 'ब' 70%। क्या संभावना है कि दैव निदर्शन द्वारा चुना गया एक प्रश्न दोनों में से किसी के द्वारा हल हो जावेगा।

(ii) The chance to get a contract for Road Construction for a contractor is 4/9 and that for a water tank is 5/7. What is the probability that at least he will get one contract of the two?

एक ठेकेदार को सड़क निर्माण का ठेका मिलने की संभावना 4/9 है और उसे पानी की टंकी का ठेका मिलने की संभावना 5/7 है तो इन ठेकों में से उसे कम से कम एक ठेका मिलने की संभावना क्या है ?

Solution-

(i) Probability that 'A' will not solve the problem $= 1 - \frac{75}{100} = \frac{25}{100}$

Probability that 'B' will not solve the problem $= 1 - \frac{70}{100} = \frac{30}{100}$

Since the events are independent, the probability that both will not solve the problem is

$$\frac{25}{100} \times \frac{30}{100} = \frac{3}{40}$$

The probability that the problem will be solved is

$$1 - \frac{3}{40} = \frac{37}{40} = 925 \times 100 = 92.5\%$$

(ii) Probability of not getting Road Contract $= 1 - \frac{4}{9} = \frac{5}{9}$

Probability of not getting water tank contract $= 1 - \frac{5}{7} = \frac{2}{7}$

Probability of not getting both the contracts $= \frac{5}{9} \times \frac{2}{7} = \frac{10}{63}$

Probability of getting at least one contract $= 1 - \frac{10}{63} = \frac{53}{63}$

Illustration 16 : (i) Find the probabilty of drawing an ace, a King, a queen and a knave in that order from a pack of cards in four consecutive draws, the cards drawn are not being replaced.

ताश की एक गड्डी में से चार बार लगातार चार पत्ते निकालने पर इस बात की संभावना मालूम कीजिए कि वे क्रमशः इक्का, बादशाह, बेगम तथा गुलाम होंगे। निकाले हुए पत्ते गड्डी में नहीं रखे जाते हैं।

(ii) Four cards are drawn without replacement. What is the probability that (a) they are all aces, (b) they are all of different suits.

बिना वापस रखे चार पत्ते निकाले जाते हैं क्या संभावना है कि (अ) वे सब इक्के हैं (ब) वे सब अलग अलग रंग के हैं।

(iii) The odds against a certain event are 5 to 3 and the odds in favour of another (independent) event are 8 to 5. Find the chance that at least one of the events will happen.

एक घटना के घटित होने के विपक्ष में संयोग अनुपात 5:3 है और दूसरी घटना (स्वतंत्र) के घटित होने के पक्ष में संयोग अनुपात 8:5 है। इस बात की संभावना बतलाइये कि कम से कम एक घटना घटित होगी।

(iv) The captain of the West Indies Cricket team is reported to have observed the rule of calling 'heads' Every time the toss was made during the five matches of the last test series with the Indian team. What is the probability of his winning the toss in all the five matches? How will the probability be affected if he had made a rule of tossing a coil privately to decide whether to call 'heads' or 'tails' on such occasion?

कहा जाता है कि भारतीय टीम के साथ गत टेस्ट श्रंखला में खेले गये 5 मेचों में वेस्ट इंडीज क्रिकेट टीम के कप्तान ने हर बार जब भी टॉस किया गया 'हैड' माँगने के नियम का पालन किया। पाँचों मैचों में उसके जीतने की संभावना क्या है? यह संभावना किस प्रकार प्रभावित होती यदि वह अलग से सिक्के को उछालकर हर बार 'हैड' या 'टेल' माँगने का निर्णय करता?

(v) 'A' can hit a target 3 itmes in 5 shots; 'B' 2 times in 5 shots; 'C' 3 itmes in 4 shots. They fire a volley.What is the probability that 2 shots hit the target?

'अ' 5 में से 3 बार एक लक्ष्य को भेद सकता है; 'ब' 5 में से 2 बार और 'स' 4 में से तीन बार। वे एक राउण्ड निशाना लगाते हैं। इसकी क्या संभावना है कि 2 निशाने लक्ष्य को भेद देंगे?

NOTES

(vi) Three groups of children contain respectively 3 girls and 1 boy; 2 girls and 2 boys; 1 girl and 3 boys one child is selected at random from each group. Show that the chance, that the three selected consist of 1 girl and 2 boys, is 13/32

बच्चों के तीन समूहों में क्रमशः 3 लड़कियाँ और 1 लड़का; 2 लड़कियाँ और 2 लड़के; और 1 लड़की और 3 लड़के हैं। दैव निदर्शन रीति से प्रत्येक समूह में से एक-एक बच्चे का चयन किया जाता है। यह सिद्ध कीजिए कि चयन किये गए 3 बच्चों में 1 लड़की और 2 लड़के होने की संभावना 13/32 है।

(vi) From a pack of 52 cards, two are drawn at random. Find the chance that one is a king and the other a queen.

52 पत्तों की ताश की एक गड्डी में से 2 पत्ते दैव निदर्शन रीति से खींचे जाते हैं। इस बात की संभावना बतलाइये कि उनमें से एक बादशाह और दूसरा बेगम।

Solution :

(i) (a) The probability of drawing an ace $= \frac{4}{52}$

(b) The probability of drawing a king $= \frac{4}{51}$

(c) The probability of drawing a queen $= \frac{4}{50}$

(d) The probability of drawing a knave $= \frac{4}{49}$

The events are dependent, hence the required probability of the compound event is:

$$\frac{4}{52} \times \frac{4}{51} \times \frac{4}{50} \times \frac{4}{49} = \frac{256}{64{,}94{,}400}$$

or $$= \frac{32}{8{,}12{,}175}$$

(ii) (a) The events are dependent. Hence the required probability

$$= \frac{52}{52} \times \frac{39}{51} \times \frac{26}{50} \times \frac{13}{49} = \frac{2{,}197}{20{,}825}$$

(iii) Thc probability that the first event will not happen $= \frac{5}{8}$

The probability that the second event will not happen $= \frac{5}{13}$

The probability that at least one event will not happen $= \frac{5}{8} \times \frac{5}{13} = \frac{25}{104}$

The probability that at least one event will happen $= 1 - \frac{25}{104} = \frac{79}{109}$

(iv) As the chance of winning the toss in one test is $\frac{1}{2}$

Since the events are independent, the probability of his winning the toss in all the five matches is

$$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{32}$$

NOTES

Since the events are independent, the calculated probability would not be affected by his decision.

(v) Firing a volley means that all of them try once to hit the target. Two shots can hit the target in one of the following ways :-

(a) A and B and C fails to hit.

(b) A and C hit and B fails to hit.

(c) B and C hit and A fails to hit.

The Probability of hitting by $A = \frac{3}{5}$ and of not hitting $= 1-\frac{2}{5}=\frac{3}{5}$

" " " " $B = \frac{2}{5}$ " " $= 1-\frac{2}{5}=\frac{3}{5}$

" " " " $C = \frac{3}{4}$ " " $= 1-\frac{3}{4}=\frac{1}{4}$

Hence-

The probability of A and B hitting and C not hitting $= \frac{3}{5}\times\frac{2}{5}\times\frac{1}{4}=\frac{6}{100}$

" " A and C " B " $= \frac{3}{5}\times\frac{3}{4}\times\frac{3}{5}=\frac{27}{100}$

" " B and C " A " $= \frac{2}{5}\times\frac{3}{4}\times\frac{2}{5}=\frac{12}{100}$

Since the events are mutually exclusive, the probability that two shots hit the target is-

$$\frac{60}{100}+\frac{27}{100}+\frac{12}{100} = \frac{45}{100} \; or \; 45\%$$

(vi) 1 girl and 2 boys may be selected in the following ways :

(a) Girl from the first group, boy from the second and boy from the third group:

The probability for it is $= \frac{3}{4}\times\frac{2}{4}\times\frac{3}{4}=\frac{18}{64}$

(b) Boy from the first group, girl from the second group and boy from the third group :

The probability is $= \frac{1}{4}\times\frac{2}{4}\times\frac{3}{4}=\frac{2}{64}$

(c) Boy from the first group, boy from the second group and the girl from the third group :

The probability is $= \frac{1}{4}\times\frac{2}{4}\times\frac{1}{4}=\frac{2}{64}$

All these events are mutually exclusive. Hence the probability that any of these events happen

$$= \frac{10}{4}+\frac{6}{64}+\frac{2}{64}=\frac{26}{64} \; or \; \frac{13}{32}$$

(vii) Two cards-a king and queen may be drawn in one of the following ways :

(a) First card a king and second a queen

(b) First card a queen and second a king.

$$\text{Probability of drawn as in (a)} = \frac{4}{52} \times \frac{4}{51} = \frac{16}{2652}$$

$$\text{Probability of drawn as in (b)} = \frac{4}{52} \times \frac{4}{51} = \frac{16}{2652}$$

(a) and (b) are mutually exclusive, Hence, required probability,

$$= \frac{16}{2652} + \frac{16}{2652} = \frac{32}{2652} \text{ or } \frac{8}{663}$$

1.8 सम्भाविता सिद्धांत में क्रमचय तथा संयोग
(Permutation and Combinations in the Theory of Probability)

ऊपर यह बताया जा चुका है कि सामान्यतः घटनाओं के घटने की अनुकूल और प्रतिकूल परिस्थितियाँ सुगमता से मालूम हो जाती हैं। किन्तु कभी-कभी ऐसा सम्भव नहीं हो पाता है। ऐसी परिस्थिति में क्रमचय और संयोग के नियमों की सहायता लेना आवश्यक हो जाता है। क्रमचय शब्द का अर्थ होता है किसी विषय की रचना किस प्रकार की जा सकती है। यह उन समस्त विन्यासों (arrangements) को प्रकट करता है जो हमें दी हुई वस्तुओं में से कुछ या सभी को एक साथ लेकर प्राप्त होते हैं। जैसे, यदि a,b,c में से दो एक साथ लेने हों तो इसके क्रमचय निम्नलिखित होंगे-

ab,	ac,	bc
ba,	ca,	cb

इस प्रकार यदि दो वस्तुयें एक साथ ली जायें तो इनकी रचना 6 प्रकार से की जा सकती है। इसे सूत्र के रूप में निम्न प्रकार प्रस्तुत किया जा सकता है-

$$^nP_r = \frac{n!}{(n-r)!}$$

where ^nP_r = Permutations

n = number of all the items

n = number of all the items

r = number of items taken together

! = factorial which means n (n–1) (n–2) (n–3)..........

उपरोक्त उदाहरण में इससूत्र द्वारा क्रमचय ज्ञात किया जा सकता है, अर्थात् a, b, c इन तीन शब्दों की रचना दो शब्दों को साथ लेकर किस प्रकार की जा सकती है-

n = 3 और r = 2

अतः $$^nP_r = \frac{n!}{(n-r)!} = \frac{3!}{(3-2)!} = \frac{3 \times 2 \times 1}{1} = 6$$

संयोग
(Combinations)

संयोग का अर्थ है- दिये गये विषय में संयोग या वर्ग किस प्रकार किये जा सकते हैं। दी गई सामग्री के ऐसे वर्ग बनाये जाते हैं जो उन वस्तुओं में से कुछ या सभी वस्तुओं को एक साथ लेने पर प्राप्त होते हैं। जैसे, a,b और c शब्दों का संयोग दो शब्दों के साथ में लेकर कितनी प्रकार से किया जा सकता है। यह संयोग निम्न प्रकार किया जा सकता है-

ab, ac, bc या ba, ac, cb

अर्थात् वह संयोग तीन प्रकार से किया जा सकता है। यह सूत्र के रूप में निम्न प्रकार से प्रस्तुत किया जा सकता है-

NOTES

$$^nC_r = \frac{n!}{(n-r)! \times r!} \text{ or } ^nC_r = \frac{^nP_r}{r!}$$

where ^nC_r = Combinations

n = number of all the items.

r = items taken together

! = factorial

उपरोक्त उदाहरण में

n=3, r=2

अतः $^nC_r = \frac{3 \times 2 \times 1}{(3-2)! \times 2!} = \frac{3 \times 2 \times 1}{1 \times 2 \times 1} = 3$

कुछ एकक समान हों तो क्रमचय (Permutations) **का उपयोग-** क्रमचय का उपयोग कैसे किया जाता है, यह ऊपर बताया जा चुका है। कभी-कभी ऐसा होता है कि जिन एककों की रचना (arrangement) करनी हो वे भिन्न-भिन्न न रहकर उसमें से कुछ एकक एक समान हों। जैसे, एक थैली में 12 गेंद हैं, वे भिन्न-भिन्न रंग की न होकर उसमें से 4 लाल रंग की और 8 हरे रंग की हैं। इस स्थिति में एक ही प्रकार के 4 एकक और दूसरे प्रकार के 8 समान एकक हैं। इस परिस्थिति में निम्न सूत्र का उपयोग किया जायेगा-

$$\text{Desired probability} = \frac{n_1!\, n_2! \ldots n_k!}{n!}$$

where $n_1, n_2 \ldots n_k$ = विभिन्न प्रकार के समान एकक

n = एककों का पूर्ण योग।

Illustration 17. If man has the choice of travelling between Bhopal and Delhi by 10 trains, in how many possible ways he can complete the return journey using a different train in each direction? यदि एक व्यक्ति को भोपाल एवं दिल्ली के बीच दस रेलगाड़ियों से यात्रा करने का विकल्प हो तो प्रत्येक दिशा में अलग गाड़ी का उपयोग करते हुए वह अपनी वापसी की यात्रा कितने प्रकार से पूरी कर सकता है ?

Solution - For the outward journey he has the choice of using all the 10 trains. Having completed the outward journey, he will be left with only 9 trains to complete the return journey. Thus the total number of ways in which he can complete the journey are $10 \times (10-1) = 90$

$$\text{Perm.} = n(n-1) = 10(10-1) = 90.$$

Illustration 18. In how many different ways 4 books at a time can be arranged out of 6 different books on a shelf? (एक अलमारी में 6 भिन्न-भिन्न पुस्तकों में से एक बार में चार पुस्तकें कितने प्रकार से विन्यासित की जा सकती है। :

Solution- $^nP_r = {}^6P_4$

$$= \frac{6!}{(6-4)!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 360$$

Illustration 19. In how many ways 5 seats can be occupied by 4 students? (चार छात्रों द्वारा 5 सीटें कितने प्रकार से ग्रहण की जा सकती है ?

Solution- Here n=5, r=4

Applying the formula,

$$^nP_r = \frac{n!}{(n-r)!} \text{ or } {}^5P_4$$

$$= \frac{5 \times 4 \times 3 \times 2 \times 1}{(5-4)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{1} = 120$$

Illustration 20. There are 6 doors in a room. 5 persons have to enter it. In how many ways they can enter from different doors?

एक कमरे में 6 दरवाजे हैं पाँच व्यक्तियों को इनमें प्रविष्ट होना है। भिन्न-भिन्न दरवाजों से वे कितने प्रकार से प्रविष्ठ हो सकते हैं?

Solution- Perm. = n (n-1) (n-2) (n-3) (n-4)

= 6 (6-1) (6-2) (6-3) (6-4) = 720

The formula can be written as under :-

$$^nP_r = {}^6P_5 = \frac{6\,!}{(6-5)\,!}$$

$$= \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{1} = 720$$

Illustration 21. In how many ways can 12 seats be occupied by 8 students? आठ छात्रों द्वारा बारह सीटें कितने प्रकार से ग्रहण की जा सकती हैं?

Solution- $^nP_r \; = {}^{12}P_8$

$$= \frac{12\,!}{(12-8)}$$

$$= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 2 \times 1}$$

= 1,99,58,400

Illustration 22. In how many ways can 12 students be allotted to three tutorial groups of 2, 4 and 6 respectively? (12 छात्रों को क्रमशः 2, 4 और 6 के तीन ट्यूटोरियल समूह में कितने प्रकार से बाँटा जा सकता है?

Solution- If m = 2, n = r, and p = 6

Then, $$^nP_r = \frac{(m+n+p)\,!}{m\,! \times n\,! \times p\,!} = \frac{(2+4+6)\,!}{2\,! \times 4\,! \times 6\,!}$$

$$= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1) \times (4 \times 3 \times 2 \times 1) \times (6 \times 5 \times 4 \times 3 \times 2 \times 1)}$$

$$= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7}{(2 \times 1) \times (4 \times 3 \times 2 \times 1)} = \frac{665280}{480} = 13860$$

Illustration 23. In how many ways can the letters of the word "BETTER" be arranged? Better शब्द के अक्षरों को कितने प्रकार से व्यवस्थित किया जा सकता है?

Solution - In this letter 'E' comes twice, 'T' comes twice and 'B' and 'R' come only once.

Hence, $$^nP_r = \frac{(m+n+p+q)\,!}{m\,! \times n\,! \times p\,! \times q\,!}$$

$$= \frac{(2+2+1+1)\,!}{(2)\,! \times (2)\,! \times (1)\,! \times (1)\,!}$$

$$= \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1) \times (2 \times 1) \times (1) \times (1)}$$

$$= \frac{6 \times 5 \times 4 \times 3}{(2 \times 1) \times (1) \times (1)} = \frac{360}{2} = 180$$

In this illustration, it has been assumed that

'm'	stands	for	letter	'E'
'n'	stands	for	letter	'T'
'p'	stands	for	letter	'B'
'q'	stands	for	letter	'R'

Illustration 24. In how many ways 3 books can be chosen out of ten different books? (भिन्न-भिन्न दस पुस्तकों में से कोई तीन पुस्तकों को चुना जा सकता है ?)

Solution the formula, we have.

$$^nC_r = \frac{n!}{(n-r)!\,r!} = \frac{10!}{(10-3)!\,3!}$$

$$= \frac{10\times9\times8\times7\times6\times5\times4\times3\times2\times1}{(7\times6\times5\times4\times3\times2\times1)\times(3\times2\times1)} = \frac{10\times9\times9}{(3\times2\times1)} = 120$$

Illustration 25. A bag contains 4 black and 3 white balls. Two draws of 2 balls are made. The balls are replaced after each draw. What is the probability that in the first draw all of them will be black and in the second white? What will be the probability if the balls are not replace?

Solution. The number of ways in which two balls can be drawn out of 7 balls is $= {}^7C_2$

The number of ways in which two black balls can be drawn out of 4 black balls is $= {}^4C_2$

The number of ways in which two white ball can be drawn out of 3 white balls is $= {}^3C_2$

$\therefore$ The probability of getting two black balls in first draw $= \dfrac{{}^4C_2}{{}^7C_2}$

Similarly, the probability of getting two white balls in the second draw $= \dfrac{{}^3C_2}{{}^7C_2}$

The desired probability

$$= \frac{{}^4C_2}{{}^3C_2} \times \frac{{}^3C_2}{{}^7C_2} = \frac{\dfrac{n!}{(n-4)!}}{\dfrac{n!}{(n-r)!\,r!}} \times \frac{\dfrac{n!}{(n-r)!\,r!}}{\dfrac{n!}{(n-r)!\,r!}}$$

$$= \frac{\dfrac{4\times3\times2\times1}{(4-2!)\,2!}}{\dfrac{7\times6\times5\times4\times3\times2\times1}{(7-2)!\,2!}} \times \frac{\dfrac{3\times2\times1}{(3-2!)\,2!}}{\dfrac{7!}{(7-2)!\,2!}}$$

$$= \frac{\dfrac{24}{(2\times1)\times(2\times1)}}{\dfrac{7\times6\times5\times4\times3\times2\times1}{(5\times4\times3\times2\times1)\times(2\times1)}} \times \frac{\dfrac{6}{(1\times2\times1)}}{\dfrac{7\times6\times5\times4\times3\times2\times1}{(5\times4\times3\times2\times1)\times(2\times1)}}$$

$$= \frac{\dfrac{24}{4}}{\dfrac{42}{2}} \times \frac{\dfrac{6}{2}}{\dfrac{42}{2}} = \frac{6}{21} \times \frac{3}{21} = \frac{2}{49}$$

If the balls are not replaced :

The probability of the first event $= \frac{^4C_2}{^7C_2}$

NOTES

However, the probability of the second event will be different if after first draw there will remain only 2 black and 3 white balls in the bag.

The number of ways in which 2 balls can be drawn out of 5 $= {}^5C_2$

The number of ways in which two white balls can be drawn out of 3 white balls $= {}^3C_2$

The probability of drawing 2 white balls is $= \frac{^3C_2}{^5C_2}$

$$= \frac{\frac{n!}{(n-r)!r!}}{\frac{n!}{(n-r)!r!}} \times \frac{\frac{3 \times 2 \times 1}{(1!2!)}}{\frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \times (2 \times 1)}}$$

$$= \frac{\frac{3 \times 2 \times 1}{(1 \times 2 \times 1)}}{\frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1) \times (2 \times 1)}} = \frac{3}{10}$$

Hence, the probability of the compound event is

$$= \frac{^4C_2}{^7C_2} \times \frac{^3C_2}{5C_2} = \frac{6}{21} \times \frac{3}{10} = \frac{3}{35}$$

Illustration 26. In how many ways can the word MANAGEMENT be arranged.

Solution- There are in all 10 letters consisting of

2 Ms, 2 As, 2 Es, 2 Ns, 1 G and 1 T.

The total permutations are

$$\frac{10!}{2!2!2!2!1!1!} = \frac{10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 2 \times 2 \times 2}$$

$$\frac{36{,}28{,}800}{16} = 2{,}26{,}800$$

Illustration 27. Out of 4 officers and 10 clerks in a business firm, a committee consisting of 2 officers and 3 clerks is to be formed. In how many ways can this be done if (a) any officer and any clerk can be included, (b) one particular clerk must be on the committee, (c) two particular officers cannot be on the committee.

Solution.

(a) 2 officers out of 4 can be selected in ${}^4C_2$ ways.

3 clerks out of 10 can be selected in ${}^{10}C_3$ ways.

Total No. of possible selections $= {}^4C_2 \times {}^{10}C_3 = 720$

(b) 2 Officers out of 4 can be selected in ${}^4C_2$ ways.

2 additional clerks out of 9 can be selected in ${}^9C_2$ ways.

Total No. of possible selection $= {}^4C_2 \times {}^9C_2 = 216$

(c) 2 officers out of 2 can be selected in $^2C_2$ ways.

3 clerks out of 10 can be selected in $^{10}C_3$ ways.

Total No. of possible selections $= {}^2C_2 \times {}^{10}C_3 = 120$

Illustration 28. A bag contains 5 green and 7 red balls. Two balls are drawn. What is the probability that one is green and the other red.

Solution- Total No. of balls $= 5 + 7 = 12$.

Now, out of 12 balls, 2 can be drawn in $^2C_2$ ways.

Exhaustive No. of cases $= {}^{12}C_2 = \dfrac{12 \times 11}{2} = 66$

Out of 5 green balls, 1 green ball can be drawn in $^5C_1$ ways and out of 7 red balls, one red ball can be drawn in $^7C_1$ ways. Since each of the former cases can be associated with each of the latter cases, the total number of favourable cases is

$$^5C_1 \times {}^7C_1 = 5 \times 7 = 35$$

Required probability $= \dfrac{35}{66}$

Illustration 29. A bag contains 7 red, 12 white and 4 green balls. What is the probability that :

(i) 3 balls drawn are all white, and

(ii) 3 balls drawn are one of each colour

Solution- Here,

n = Total No. of possible ways of drawing 3 balls from the bag containing 23 balls, irrespective of the colour of balls drawn.

$= {}^{23}C_3$

(i) m = No. of possible ways of drawing 3 white balls out of 12 white balls.

$\therefore$ Probability that all three balls are white

$$\frac{^{12}C_3}{^{23}C_3} = \frac{220}{1771} = 0.1242.$$

(ii) It may be noted that 1 red ball can be drawn out of 7 in $^7C_1$ ways. Similarly, the No. of ways of drawing 1 white ball out of 12 white ball is $^{12}C_1$ and one green ball can be chosen out of 4 green in ways. Since any of 7 groups of 1 red ball may be associated with any of the 12 groups of 1 white ball and 4 groups of 1 green ball, the number of ways of forming a group of 3 with 1 red ball, 1 white ball and 1 green ball is.

$$m = {}^7C_1 \times {}^{12}C_1 \times {}^4C_1$$

Probability that three balls are of different colours

$$= \frac{^7C_1 \times {}^{12}C_1 \times {}^4C_1}{^{23}C_3} = \frac{7 \times 12 \times 4}{1771} = \frac{336}{1771} = 0.1897$$

Miscellaneous Problems on Probability

Illustration 30. If a card is drawn from a well shuffled pack, find

(a) the chance that the card drawn is the ace of spade,

(b) the chance that the card is a spade,

(c) the chance that the card is an ace.

NOTES

Solution. (a) Since there are 52 cards in a pack, and there is only ace of spade, the chance of drawing the ace of spade in 1/52.

(b) Since there are 13 cards of spade, required chance is $= \frac{13}{52} = \frac{1}{4}$

(c) Similarly, there are four aces in a pack of cards. Therefore, the required chance or probability is

$$= \frac{4}{52} = \frac{1}{13}$$

Illustration 31. If, from a pack, four cards are drawn, find the chance that

(a) they will be the four honours of the same suit,

(b) they will be one from each suit.

Solution - The are 52 cards in a pack, and the total number of drawing 4 cards is $^{52}C_4$ or $\frac{52 \times 51 \times 50 \times 49}{4 \times 3 \times 2 \times 1} = 2{,}70{,}725$

(a) There are 4 suits (each containing four honours) in a card and therefore the number of favourable cases is 4.

Therefore, the required chance $= \frac{4}{2{,}70{,}725}$

(b) One card from a suit (containing 13 cards) can be drawn in 13 ways. There are four suits and each suit has 13 cards. Therefore one card from each suit be drawn $\frac{13 \times 13 \times 13 \times 13}{270725} = \frac{2197}{20825}$.

Illustration 32. In a hand of a game of bridge, what is the probability that the 4 kings are held by a specified player?

Solution - The number of ways in which the hand can be made up is $^{52}C_{12}$. The number of favourable ways is the same as the number of ways in which the 9 other cards forming the hand can be chosen from the remaining 41 cards which is $^{48}C_9$

Therefore, the required probability $= \frac{^{48}C_9}{^{52}C_{13}} = \frac{11}{4165}$

Illustration 33. Suppose it is 9 to 7 against a person A who is now 35 years of age living till he is 65, and 3 to 2 against a person B now 45 living till he is 75, find the chance that one at least of these persons will be alive 30 years hence.

Solution- The chance that A will die within 30 years is $\frac{9}{16}$ and chance that B will die within 30 years is $\frac{3}{5}$. The events are independent, therefore, the chance that both will die is $\frac{9}{16} \times \frac{3}{5}$ *or* $\frac{27}{80}$

Therefore, the chance that both will not be dead, i.e., least one will be alive is

$$1 - \frac{27}{80} = \frac{53}{80}$$

Illustration 34. What is the chance that a leap year, selected at random, will contain 53 Sundays?

Solution - A leap year consists of 52 complete weeks with a balance of two days. These two days can be any of the following seven possibilities-

Monday and Tuesday,

Tuesday and Wednesday,

Wednesday and Thursday,

Thursday and Friday

Friday and Saturday,

Saturday and Sunday,

Sunday and Monday.

Out of these seven combinations, only two viz. the the last two are favourable combinations. Hence the required chance is $\frac{2}{7}$.

Illustration 35. A professor claims that he could often tell while his students were still is their first year whether they will obtain first, second, third or fourth (failure) class in their final examination. To substantiate his claim he forecasts the fates of 8 students. Find the probability of his being correct in 5 cases.

Solution - Assuming that the professor has no special gift, the chance of naming a student's class is $\frac{1}{4}$ and therefore the chance of not being correct in naming student's class is $\frac{8}{4}$. Hence the probability of the professor being correct in 5 cases is

$$^8C_5 = \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^3$$

$$^8C_5 = 56 \times \frac{27}{4^8} = \frac{189}{8192}$$

बरनौली प्रमेय (Bernouli Theorem)- इस प्रमेय के जन्मदाता श्री जेम्स बरनौली थे। उन्हीं के नाम से यह नियम 'बरनौली प्रमेय' कहलाता है। किसी घटना के घटित होने की दो दशाएँ हो सकती हैं- (i) अनुकूल, (ii) प्रतिकूल। हम सामान्यतः अनुकूल घटनाओं के 'p' और प्रतिकूल घटनाओं को 'q' संकेतों के रूप में प्रकट करते हैं। इस प्रमेय की निम्नलिखित विशेषताएँ होती हैं-

यदि (1) सभी अन्य परिस्थितियों सामान्य रहें।

(2) घटना को अनेक बार दुहराया जाये।

तो सफलता की एक निश्चित संख्या प्राप्त करने की सम्भावना ज्ञात की जा सकती है-

इसके लिये निम्न सूत्र का प्रयोग किया जाता है-

$$P = {}^nC_r p^r q^{n-r}$$

p = probability of happening of the event, i.e., success,

q = probability of not happening,

n = total no. of trails,

r = no. of Successes in n trails,

^nC_r = no. of combinations of 'n' things taken 'r' at a time.

Illustration 36. (i) What is the probability of exactly four heads while tossing five coins.

(ii) A coin is tossed three times. Find out the probability of (a) exactly two heads; and (b) at least two heads.

Solution-

(i) No. of coins (trails) - n = 5; Success (H) frequency or r = 4; Failure (T) or n - r = 5 - 4 = 1.

$$p = \frac{1}{2};\ q = \frac{1}{2}$$

$$^{n}C_{r} q^{n-r} = {}^{5}C_{4} \left(\frac{1}{2}\right)^{4} \left(\frac{1}{2}\right)^{1} = \frac{5!}{1!\,4!} \times \frac{1}{32} = \frac{5}{32}$$

(ii) (a) Probability of exactly two heads and one tail in tossing a coin three times will be-

n = 3; r = 2; n-r = 3-2 = 1.

p (Probability of head in one toss) = 1/2

q („ „ tail „ „) = 1/2

$p^{r} q^{n-r}$ (Combined probability of 2 heads and 1 tail)

$$= \left(\frac{1}{2}\right)^{2} \left(\frac{1}{2}\right)^{3-2} = \frac{1}{8}$$

But this situation (2H, IT) can be in the following three combinations-

$$^{3}C_{2} = \frac{3!}{(3-2)!\,2!} = 3$$

Trails		I	II	III	Probability
Desired results	(A)	H	H	T	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$
	(B)	H	T	H	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$
	(C)	T	H	H	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

∴ Desired probability

$$= 3 \times \frac{1}{8} = \frac{3}{8} \left[^{n}C_{r} p^{r} q^{n-r} = {}^{3}C_{2} - \left(\frac{1}{2}\right)^{2} \left(\frac{1}{2}\right)^{1} = \frac{3}{8} \right]$$

(b) There can be two situations of getting at least two heads in three tosses-

(i) Exactly 2 Heads and Tail; (ii) 3 Heads.

Bernouli Theorem will be applied twice-

Probability of 2 Heads and 1 Tail $= {}^{3}C_{2} p^{2} q^{1}$

" " 3 Heads and 0 " $= {}^{3}C_{3} p^{3} q^{0}$

Desired probability-

$$^{3}C_{2} p^{2} q^{1} + {}^{3}C_{3} q^{0} = (3 \times \frac{1}{8}) + (1 \times \frac{1}{8}) = \frac{1}{2}$$

Illustration 37. (i) Eight coins are tossed together. Find out the probabilities of the following-

(i) (a) At least six heads (b) No head, (c) All heads.

(ii) If one candidate out of ten fails on an average in an examination, what is the probability that 4 out of 5 will pass?

Solution- (i) (a) At least 6 heads means-6 heads, 7 heads or 8 heads. Therefore, probabilities of all the three situations will be found first and then their total will be arrived at-

$$n = 8;\ r = 6, 7, 8;\ p = \frac{1}{2} = q = \frac{1}{2}$$

$$p\,(r = 6, 7, 8) = {}^8C_6 \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^2 + {}^8C_7 \left(\frac{1}{2}\right)^7 \left(\frac{1}{2}\right)^1 + {}^8C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^0$$

$$= 28 \times \frac{1}{256} + 8 \times \frac{1}{256} + 1 \times \frac{1}{256} = \frac{37}{256}$$

(b) Probability of no head (0 head) in tossing 8 coins is

$$= {}^8C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^8 = 1 \times \frac{1}{256} = \frac{1}{256}$$

(c) Probability of all heads (8 heads) is

$$= {}^8C_8 \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^0 = 1 \times \frac{1}{256} = \frac{1}{256}$$

(ii) Probability of failing $= \frac{1}{10}$;

Probability of Success $= 1 - \frac{1}{10} = \frac{9}{10}$

Probability of success of at least 4 out of 5 candidates is

= Probability of success of 5 + Probability of success of 4.

$$= {}^5C_5 \left(\frac{9}{10}\right)^5 \left(\frac{1}{10}\right)^0 + {}^5C_4 \left(\frac{9}{10}\right)^4 \left(\frac{1}{10}\right)^1$$

$$= \frac{59049}{1{,}00{,}000} + \frac{5 \times 6561}{1{,}00{,}000} = \frac{91854}{1{,}00{,}000} \text{ or } \frac{45927}{50{,}000}$$

1.9 प्रतिबन्धित प्रायिकता (Conditional Probability)

किसी एक घटना, A, के घटने के बाद दूसरी घटना, B के घटने की प्रायिकता उसकी प्रतिबन्धित प्रायिकता कहलाती है, जबकि यह ज्ञात हो कि पहली घटना, A, हो चुकी है। संकेत रूप में,

B की प्रतिबन्धित प्रायिकता जबकि A हो चुकी हो

$$= P(B/A) = \frac{P(A \cap B)}{P(A)},\quad P(A) \neq 0$$

$$P(A \cap B) = P(A)\, P(B/A)$$

Illustration 38. ताश की एक गड्डी को फेंटने के बाद एक पता निकाला जो बादशाह पाया गया। उसको वापस किये बिना दूसरा पत्ता निकाला गया। दूसरे पत्ते के बादशाह होने की प्रायिकता ज्ञात कीजिए। दोनों पत्तों के बादशाह होने की प्रायिकता भी बताइये।

Out of a well shuffled pack of cards a card is drawn, which is found a king. Another card is drawn with out replacing it. Find the probability of second card being a king. Also find the probability of both cards being kings.

हल (Solution)

पहली बार में बादशाह निकालने (घटना A) की प्रायिकता, B (A) $= \frac{4}{52}$

अब गड्डी में 51 ताश रह गये जिनमें 3 बादशाह हैं, अतः दूसरी बार में बादशाह निकालने (घटना B) की प्रायिकता

$$= P(B/A) = \frac{32}{51}$$

दोनों ताशों के बादशाह होने की प्रायिकता $= p(A \cap B) = P(A)\, P(B/A)$

$$= \frac{4}{52} \times \frac{3}{51} = \frac{1}{221}$$

यदि घटना B के होने की प्रायिकता घटना A के होने से प्रभावित होती है तो घटना B को घटना A पर आश्रित कहा जाता है। यदि घटना B के होने की प्रायिकता घटना A के होने से प्रभावित नहीं होती, घटना B को घटना A से स्वतंत्र कहा जाता है। संकेत रूप में,

यदि $P(B/A) = P(B)$ तो B,A से स्वतंत्र है।

यदि $A(A/B) = P(A)$ तो A,B से स्वतंत्र है।

इस प्रकार घटना A तथा B स्वतंत्र घटना कहलाती है यदि और केवल यदि

$P(A \cap B) = P(A)\, P(B)$

उदाहरणार्थ- (i) एक सिक्का दो बार उछाला जाता है। A = पहली बार H आना, B = दूसरी बार T आना, स्वतंत्र हैं।

(ii) उपर्युक्त उदाहरण 28 में, यदि ताश को वापस कर दिया जाये तो $P(B/A) = P(B) = \frac{4}{52}$ होगी। तब A और B स्वतंत्र घटनाएं हैं।

संयोगानुपात (Odds)- कभी-कभी यह न कहकर कि A के होने की प्रायिकता $\frac{a}{a+b}$ है यह भी कहा जाता है कि घटना A के अनुकूल संयोगानुपात (odds in favour of A) $P(A) : P(A^c)$ अर्थात् $\frac{a}{a+b} : \frac{b}{a+b}$ अथवा $a+b$ है या A के होने के प्रतिकूल संयोगानुपात (odd against A), $P(A^c) : P(A)$ अर्थात् $b : a$ है। उदाहरणार्थ, यदि एक थैले में 7 सफेद और 9 काली गेंदें हैं तो उनमें से एक गेंद निकालने पर उसके सफेद होने की प्रायिकता $\frac{1}{7+9} = \frac{1}{16}$ है। इसके अनुकूल संयोगानुपात 7:9 अथवा $\frac{7}{9}$ है तथा इसके प्रतिकूल संयोगानुपात 9:7 अथवा $\frac{9}{7}$ है। यदि किसी घटना के होने की प्रायिकता 0.4 हो तो संयोगानुपात निकालने के लिए P(A) को भिन्न के रूप में लिख लेते हैं : $P(A) = 0.4 = \frac{4}{10} = \frac{2}{5} = \frac{2}{2+3}$ अतः अनुकूल संयोगानुपात = 2:3 होंगे।

गणितीय अनुशंसा
(Mathematical Expectation)

यदि किसी घटना के घटने पर एक व्यक्ति को एक निश्चित धनराशि मिलती हो, तो उस घटना की सम्भावना और सफलता पर उसे मिलने वाली धनराशि का गुणनफल उस व्यक्ति की गणितीय अनुशंसा कहलाती है।

सूत्रानुसार, E = p.M.

E = Mathematical Expectation.

p = Probability of success of an event.

M = Money to be received on the happening of the event.

उदाहरण के लिए यदि किसी घटना के घटने की सम्भावना (p) = 1/4 हो और उसके होने पर मिलने वाली राशि (M) 200 रुपये हो तो गणितीय अनुशंसा $\frac{1}{4} \times 200 = 50$ रुपये होगी।

प्रतिलोम सम्भावना
(Inverse Probability)

प्रतिलोम सम्भावना में हम विपरीत प्रकार की समस्याओं का अध्ययन करते हैं। हमें यह ज्ञात है कि अमुक घटना अनेक कारणों में से किसी एक निश्चित कारण का प्रभाव है। यह मालूम करना कि घटना के उस निश्चित कारण का प्रभाव होने की क्या सम्भावना है- प्रतिलोम सम्भावना कहलाती है। यदि किसी थैले में 3 काली और 2 सफेद गेंदें हों, दूसरे थैले में 4 काली और 3 सफेद गेंदे हों और किसी एक थैले में से एक काली गेंद निकाली गई हो तो यह मालूम करना कि उस गेंद के पहले थैले में से निकाले जाने की क्या सम्भावना होगी, प्रतिलोम सम्भावना का उदाहरण होगा।

नियम- कोई घटना 'N' परस्पर अपवर्जी कारणों में से किसी एक के द्वारा उत्पन्न हो सकती है। जिन कारणों की अलग-अलग सम्भावनाएँ $p_1, p_2, p_3 \ldots p_n$ हैं और ये क्रमशः उन विभिन्न कारणों में से प्रत्येक की अलग-अलग ऐसी संभावनाएँ हैं कि जब 'N' कारणों में से कोई एक कारण होता है तो घटना उत्पन्न होती है। ऐसी स्थिति में उस घटना के mवें कारण से उत्पन्न होने की सम्भावना निम्नलिखित सूत्र द्वारा ज्ञात की जा सकती है-

$$\frac{P_m P_m}{P_1 p_1 + P_2 p_2 + \ldots P_n p_n}$$

1.10 बेज प्रमेय (Bayes Theorem)

कथन (Statement)

कोई घटना A, n परस्पर अपवर्जी तथा निःशेषी कारणों (घटनाओं) $B_1, B_2 \ldots, B_n$ में से किसी भी एक कारण से प्रभावित हो सकती है। यदि उन कारणों की स्वयं सिद्ध प्रायिकताएं $P(B_1) \ldots\ldots, P(B_n)$ ज्ञात हों तथा विभिन्न कारणों से घटना के घटित होने की प्रतिबन्धी प्रायिकताएं $P(A/B_1) \ldots P(A/B_n)$ भी ज्ञात हों तब अनुभव सिद्ध प्रायिकता $P(B_m/A)$ को निम्न प्रकार व्यक्त किया जा सकता है :

$$P(B_m/A) = \frac{P(B_m)\, P(A/B_m)}{\sum_{m=1}^{n} P(B_m)\, P(A/B_m)}$$

जहां $P(B_m/A)$ का अर्थ है, "घटना A के B_m से प्रभावित होने की प्रायिकता जबकि यह ज्ञात हो कि घटना A हो चुकी है।"

उपपत्ति (Proof)

माना $B_1, B_2, \ldots\ldots B_n$ किसी यादृच्छिक प्रयोग की प्रतिदर्श समष्टि की n परस्पर अपवर्जी तथा निःशेषी घटनाएं हैं तथा A कोई अन्य घटना है।

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_n)$$

तथा
$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \ldots + P(A \cap B_n)$$
$$= \sum_{m=1}^{n} P(A \cap B_m) \qquad -(1)$$

पुनः गुणन प्रमेय से,

$$P(A \cap B_m) = P(A)\, P(B_m/A) \qquad -(2)$$

तथा
$$P(A \cap B_m) = P(B_m)\, P(A/B_m) \qquad -(3)$$

उपर्युक्त (2) तथा (3) से,

$$P(A)\, P(B_m/A) = P(B_m)\, P(A/B_m)$$

NOTES

$$P(B_m/A) = \frac{P(B_m)\,P(A/B_m)}{P(A)}$$

$$= \frac{P(B_m)\,P(A/B_m)}{\sum_{m=1}^{n} P(A \cap B_m)} \qquad [(1) \text{ से }]$$

$$= \frac{P(B_m)\,P(A/B_m)}{\sum_{m=1}^{n} P(B_m)\,P(A/B_m)}$$

बोध प्रश्न

1. संभाविता के योग नियम का प्रयोग कब किया जाता है।

....

....

....

2. संभाविता के गुणकार का नियम क्या है ?

....

....

....

3. साँख्यिकी में संभावना सिद्धान्त के महत्व को समझाइए।

....

....

....

Illustration : 39 एक थैले में 4 काली तथा 1 सफेद गेंद हैं। दूसरे थैले में 5 काली व 4 सफेद गेंदें हैं। किसी एक थैले से एक काली गेंद निकाली जाती है तो क्या प्रायिकता है कि पहले थैले से ही निकली है ?

A bag contains 4 black and 1 white ball. Another bag contains 5 black and 4 white balls. One ball has been taken out from one of the bags and found black. What is the probability that it came from the first bag?

हल (Solution)

माना B_1 = पहले थैले को चुनना

B_2 = दूसरे थैले को चुनना

A = काली गेंद निकालना

तब $P(B_1) = P_1 = \frac{1}{2}$

$P(B_2) = P_2 = \frac{1}{2}$

पहले थैले से एक काली गेंद निकालने की प्रायिकता

$$= P(A/B_1) = p_1 = \frac{4}{5}$$

दूसरे थैले से एक काली गेंद निकालना $= P(A/B_2) = p_2 = \frac{5}{9}$

अभीष्ट प्रायिकता $= p(B_1/A) = \dfrac{P(B_1)\,P(A/B_1)}{P(B_1)P(A/B_1)+P(B_2)P(A/B_2)}$

अथवा काली गेंद पहले थैले से निकाली गयी की प्रायिकता

NOTES

$$= \frac{P_1 p_1}{P_1 p_1 + P_2 p_2}$$

$$= \frac{\frac{1}{2} \times \frac{4}{5}}{\frac{1}{2} \times \frac{4}{5} + \frac{1}{2} \times \frac{5}{9}} = \frac{\frac{4}{5}}{\frac{4}{5} + \frac{5}{9}} = \frac{\frac{4}{5}}{\frac{36+25}{45}} = \frac{4}{5} \times \frac{45}{61} = \frac{36}{61}$$

Illustration 40 एक पर्स में एक रुपये के तीन सिक्के तथा 50 पैसे के चार सिक्के हैं। दूसरे पर्स में एक रुपये के चार सिक्के तथा 50 पैसे के पांच सिक्के हैं। किसी एक पर्स में से एक रुपये का सिक्का निकाला गया। यह सिक्का पहले पर्स से आया, इसकी प्रायिकता निकालिए।

A purse contains three one rupees coins and four 50 paise coins. Another purse contains four one rupee coins and five 50 paise coins. A one rupee coin has been taken out from one of the purses. Find the probability that it came from the first purse.

हल (Solution)

माना B_1 = पहले पर्स से सिक्का निकालना

B_2 = दूसरे पर्स से सिक्का निकालना

तब $P(B_1) = P(B_2) = \frac{1}{2}$

माना R = एक रुपये का सिक्का निकालना $= (B_1 \cap R) \cup (B_2 \cap R)$

पहले पर्स से एक रुपये का सिक्का निकालने की प्रायिकता, $P(R/B_1) = \frac{3}{7}$

दूसरे पर्स से एक रुपये का सिक्का निकालने की प्रायिकता, $P(R/B_2) = \frac{4}{9}$

अभीष्ट प्रायिकता $= P(B_1/R) = \frac{P(B_1 \cap R)}{(P(R)}$

$$= \frac{P(B_1)\, P(R/B_1)}{P(B_1)\, P(R/B_1) + P(B_2)\, P(R/B_2)}$$

$$= \frac{\frac{1}{2} \times \frac{3}{7}}{\left(\frac{1}{2} \times \frac{3}{7}\right) + \left(\frac{1}{2} + \frac{4}{9}\right)} = \frac{\frac{3}{14}}{\frac{3}{14} + \frac{2}{9}} = \frac{3}{14} \cdot \frac{14 \times 9}{27 + 28} = \frac{27}{55} = 0.49$$

Illustration 41 एक थैले में 3 काली और 4 सफेद गेंदें हैं। यादृच्छिक विधि से 2 गेंदे बिना वापस किये एक-एक करके खींची गयी हैं। यदि यह ज्ञात हो कि दूसरी गेंद लाल है तो पहली गेंद के काली होने की प्रायिकता ज्ञात कीजिए।

There are 3 black and 4 white balls in a bag. Two balls are drawn one by one by one without replacement. If it known that the second ball is red, then find the probability of the first ball being black.

हल (Solution)

माना B_1 = पहली गेंद काली होना

B_2 = पहली गेंद लाल होना

तथा A = दूसरी गेंद लाल होना

तथा $A = (B_1 \cap A) \cup (B_2 \cap A)$

$$P(A) = P(B_1 \cap A) + P(B_2 \cap A)$$

$$= P(B_1)\, P(A/B_2)\, P(A/B_2)$$

NOTES

या $$P(A) = \frac{3}{7} \times \frac{4}{6} + \frac{4}{7} \times \frac{3}{6} = \frac{12}{42} + \frac{12}{42} = \frac{24}{42} = \frac{4}{7} \quad \ldots\ldots(1)$$

अभीष्ट प्रायिकता $= P(B_1/A) = \dfrac{P(A \cap B_1)}{P(A)}$

$$= \frac{P(A \cap B_1)}{P(A \cap B_1) + P(A \cap B_2)}$$

$$= \frac{P(B_1)\, P(A/B_1)}{P(B_1)\, P(A/B_1) + P(B_2)\, P(A/B_2)}$$

$$= \frac{\frac{3}{7} \times \frac{4}{6}}{\frac{4}{7}} = \frac{2}{7} \times \frac{7}{4} = \frac{1}{2}$$

Illustration 42 फाउण्टेन पेन का निर्माण करने वाले एक कारखाने में यंत्र A, B तथा C फाउण्टेन पेन के कुल उत्पादन का क्रमशः 30%, 30% तथा 40% उत्पादन करते हैं। उनके उत्पादन में क्रमशः 4%, 5% तथा 10% फाउण्टेन पेन दूषित पाये जाते हैं। यदि दैव रीति से एक फाउण्टेन पेन चुना जाता है और वह दूषित पाया जाता है, तो इसकी क्या प्रायिकता है कि उसका उत्पादन C यंत्र द्वारा किया गया था?

In a factory manufacturing fountain pens, machine A, B and C manufacture 30%, 30% and 40% of the total production of fountain pens, respectively. Of their output, 4%, 5% and 10% of the fountain pens are defective. If one fountain pen is selected at random, and if it is found to be defective, what is the probability that it is manufactured by machine C?

हल (Solution)

चुने गये फाउण्टेन पेन के

यंत्र A का उत्पादन होने की प्रायिकता, $P_1 = \dfrac{30}{100}$

यंत्र B का उत्पादन होने की प्रायिकता, $P_2 = \dfrac{30}{100}$

यंत्र C का उत्पादन होने की प्रायिकता, $P_3 = \dfrac{40}{100}$

$p_1 =$ यंत्र A का उत्पादन दूषित होने की प्रायिकता $= \dfrac{4}{100}$

$p_2 =$ यंत्र B का उत्पादन दूषित होने की प्रायिकता $= \dfrac{5}{100}$

$p_3 =$ यंत्र C का उत्पादन दूषित होने की प्रायिकता $= \dfrac{10}{100}$

बेज प्रमेय अनुसार,

अभीष्ट प्रायिकता $= \dfrac{P_3p_3}{P_1p_1 + P_2p_2 + P_3p_3}$

$$= \frac{\frac{40}{100} \times \frac{10}{100}}{\frac{30}{100} \times \frac{4}{100} + \frac{30}{100} \times \frac{5}{100} + \frac{40}{100} \times \frac{10}{100}}$$

$$= \frac{40 \times 10}{(30 \times 4) + (30 \times 5) + (40 \times 10)}$$

$$= \frac{400}{120 + 150 + 400} = \frac{400}{670} = \frac{40}{67}$$

1.11 सारांश

सांख्यिकीय में संभावना सिद्धान्त एक महत्वपूर्ण सिद्धान्त है जो कि घटना के घटित होने संभावना को व्यक्त करता हो आर्थिक एवं व्यापारिक जटिलतम समस्याओं के संख्यात्मक विश्लेषण में संभावना सिद्धान्त अनिवार्य होता जा रहा है, जिसके माध्यम से श्रेष्ठतम् निर्णय का चयन किया जाता है। संभावना सिद्धान्त का महत्व उन क्षेत्रों में अधिक किया जाता है जहाँ अनिश्चितता के वातावरण में भावी अनुमान लगाये जाते हों साँख्यिकीय में संभावना को व्यक्त करने के लिए कुछ महत्वपूर्ण सिद्धान्तों व शब्दावली जैसे- घटना, क्रमचय, संयोग आदि का प्रयोग किया जाता है।

1.12 शब्द कुंजी

यादृच्छिक प्रयोग, दैव प्रयोग, प्रतिदर्श समष्टि, पारस्परिक अपवर्जी घटनाएँ

1.13 अभ्यास प्रश्न

दीर्घउत्तरीय प्रश्न (Long Answer Type Questions)

1. प्रायिकता की परिभाषा दीजिए।
   Define probability.
2. प्रायिकता का उपयुक्त उदाहरण देकर प्रायिकता के योग प्रमेय को समझाइये।
   Explain the addition theorem of probability giving suitable examples.
3. प्रायिकता में स्वतंत्र एवं परस्पर अपवर्जी घटनाओं की अवधारणाओं की व्याख्या कीजिए।
   Explain the concepts of independent and mutually exclusive events in probability.
4. दो घटनाओं के (i) स्वतंत्र व (ii) आश्रित होने की स्थितियों में प्रायिकता के गणन प्रमेय को उपयुक्त उदाहरण देकर समझाइये।
   Explain the multiplication theorem of probability for two events which are (i) independent and (ii) dependent giving suitable example.
5. 'क्रमचय' तथा 'संचय' शब्दों की उदाहरण सहित व्याख्या कीजिए।
   Explain the terms 'permutation' and 'combination' with example.
6. क्रमचय एवं संयोग तथा योग एवं गुणन का नियम सम्भावना सिद्धांत में क्या सहायता प्रदान करते हैं? दर्शाइये।
   What role is played by permutation and combination and rule of addition and multiplication in probability theory? Explain.
7. सांख्यिकी में प्रायिकता की अवधारणा की महत्ता की व्याख्या कीजिए।
   Discuss the importance of the concept of probability in statistics.
8. सांख्यिकी में इस सिद्धांत के महत्व को स्पष्ट कीजिए।
   Explain the importance of probability theory in statistics.
9. सम्भावना की सांख्यिकीय परिभाषा दीजिए।
   Give statistical definition of probability.
10. सम्भाविता की गणितीय और सांख्यिकीय परिभाषा दीजिए और उनमें निहित सम्बन्ध, यदि कोई हो तो, का वर्णन कीजिए।
    Give the classical and statistical definition of probability and state the relationship, if any, between the two definitions.
11. निम्नलिखित धारणाओं का आलोचनात्मक अन्तर स्पष्ट कीजिए :
    Distinguish between the following concepts critically :
    (i) सरल एवं संयुक्त घटनाएं (Simple and compound events)

(ii) स्वतंत्र एवं आश्रित घटनाएं (Independent and dependent events)

(iii) परस्पर अपवर्जी एवं स्वतंत्र घटनाएं (Mutually exclusive and Independent events)

(iv) क्रमचय एवं संयोग (Permutation and Combination) ।

12. प्रायिकता के जोड़ एवं गुणा के प्रमेयों (सूत्रों) को परिभाषित कीजिए तथा उपयुक्त उदाहरणों की सहायता से समझाइये।
Define the theorems (formulae) of addition and multiplication of probabilities and illustrate with suitable examples.

13. पूर्ण एवं मिश्रित प्रायिकता की प्रमेयों का वर्णन कीजिए।
State the theorems of total and compound probability.

14. मान्यताएं एवं सीमाएं बताते हुए प्रायिकता की अवधारणा के दोनों आधारों-सम-सम्भाव्य तथा सापेक्ष आवृत्ति को समझाइये।
Stating the assumptions and limitations exaplain the two bases-Equally likely and relative frequency of concept of probability.

15. प्रायिकता के विभिन्न प्रमेयों को परिभाषित कीजिए।
Define the various theorems of probability.

16. उदाहरण सहित, सम्भाव्यता के गुणन प्रमेय की व्याख्या कीजिए।
Discuss the multiplication theorem of probability with example.

17. सामान्य सम्भावना तथा सीमान्त में अन्तर को उदाहरण सहित समझाइये।
Explain with example the difference between simple probability and marginal probability.

18. समझाइए 'प्रायिकता' शब्द से आप क्या समझते हैं? प्रायिकता की योग एवं गुणन प्रमेयों को लिखिए तथा सिद्ध कीजिए।
Explain what you understand by the term probability? State and prove addition and multiplication theorems of probability.

19. दो घटनाओं (i) परस्पर अपवर्जी होने और (ii) परस्पर अपवर्जी न होने की स्थितियों में प्रायिकता के योग प्रमेय को बताइये।
State the addition theorem of probability in case of (i) mutually exclusive events and (ii) not mutually exclusive events.

20. निम्नलिखित पर संक्षिप्त टिप्पणी लिखिए :

(i) सम्भावना सिद्धांत की उपयोगिता (Utility of probability theory)

(ii) क्रमचय तथा संयोग (Permutation and combination)

(iii) व्युत्क्रम प्रायिकता (Inverse probability)

(iv) सप्रतिबन्ध प्रायिकता (Conditional probability)

(v) बर्नोली प्रमेय (Bernoulli's theorem)

(vi) गणितीय प्रत्याशा (Mathematical expectation)

(vii) बेज प्रमेय (Bayes Theorem)

22. प्रायिकता पर बेज प्रमेय को लिखिए तथा सिद्ध कीजिए।
State and prove Bayes theorem on probability.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. प्रायिकता की परिभाषा दीजिये।
Define probability.

2. सम्भावना की गणितीय परिभाषा दीजिये।
   Give the mathematical definition of probability.
3. सम्भावना के नियम लिखिये।
   Write the Rules of probability.
4. सांख्यिकी में सम्भावना की महत्ता की व्याख्या कीजिये।
   Discuss the importance of probability in statistics.
5. प्रायिकता के योग प्रमेय को समझाइये।
   Explain the addition theorem of probability.
6. प्रायिकता के गणन प्रमेय को समझाइये।
   Explain the multiplication theorem of probability.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. दो पांसे फेंकने पर 8 का जोड़ आने की प्रायिकता-
   (अ) 1/12 (ब) 3/36 (स) 6/36 (द) 5/36
2. अंग्रेजी की पुस्तक में चुना गया एक स्वर A होगा इसकी प्रायिकता होगी-
   (अ) 1/3 (ब) 1/5 (स) 1/6 (द) सभी

एक थैले में तीन सफेद एवं चार काली गेंदे हैं। एक गेंद निकाली गयी यह गेंद सफेद होने की प्रायिकता-
   (अ) 1/2 (ब) 2/7 (स) 3/7 (द) 4/7

4. संभावना अनिश्चित घटनाओं के बारे में मस्तिष्क की एक स्थिति है। यह कथन है-
   (अ) काँनर (ब) लाप्लास (स) कैने (द) कैले

उत्तर- 1. (द), 2. (ब), 3. (स), 4. (अ)

1.14 व्यावहारिक प्रश्न

1. एक थैले में 8 सफेद, 6 काली, 4 हरी और 2 पीली गेंदें हैं। यदि यादृच्छिक रूप से एक गेंद निकाली जाय तो उसके काले या हरे या पीले रंग की होने की क्या सम्भावना है?
   A bag contains 8 white, 6 black, 4 green and 2 yellow balls. If a balls is drawn at random, what is the probability of its being of black or green or yellow colour?
   [उत्तर $\frac{6}{20}+\frac{4}{20}+\frac{2}{20}=\frac{3}{5}$]
2. टिकटों पर 1 से 30 तक संख्याएं लिखी हैं। एक टिकिट यादृच्छिक रूप से निकाला जाता है। इसके 5 या 7 का अपवर्त्य होने की प्रायिकता ज्ञात कीजिए।
   From 30 tickets marked with the first 30 numerals, one is drawn at random. Find the chance that it is a multiple of 5 or 7.
   [उत्तर : $\frac{6}{30}+\frac{4}{30}=\frac{10}{30}=\frac{1}{3}$]
3. दो पासों को एक बार फेंकने पर 10 या 12 का **योग** आने की प्रायिकता ज्ञात कीजिए।
   Find the probability of getting a sum of 10 or 12 when two dice are thrown once.
   [उत्तर : $\frac{3}{36}+\frac{1}{36}=\frac{4}{36}=\frac{1}{9}$]
4. 52 पत्तों की ताश की गड्डी में से एक पत्ता अनायास रूप से निकाला गया। इसकी क्या प्रायिकता होगी कि वह हुक्म का पत्ता या कोई बादशाह होगा?
   From a pack of 52 cards one card is drawn at random. What is the probability that it will be a spade or a king?
   [उत्तर : $\frac{13}{52}+\frac{4}{52}-\frac{1}{52}=\frac{4}{13}$]

NOTES

5. दो पासों को एक बार फेंकने पर कम-से-कम 8 का योग आने की क्या सम्भावना है ?
What is the probability of getting at least eight with two dice?

[उत्तर : $\frac{5}{36}+\frac{4}{36}+\frac{3}{36}+\frac{2}{36}+\frac{1}{36}=\frac{5}{12}$]

6. प्राकृतिक संख्याएं 1 से 30 तक अंकित 30 टिकटों में से एक टिकट यादृच्छिक रूप से निकाला गया। इसकी क्या सम्भावना है कि वह उस टिकट का नम्बर 3 या 5 का गुणक है।
From 30 tickets marked with the first 30 natural numbers, one is drawn at random. What is the chance that number on it is a multiple of 3 or 5.

[उत्तर : $\frac{10}{30}+\frac{6}{30}-\frac{2}{30}=\frac{14}{30}=\frac{7}{15}$]

7. एक थैले में 25 गेंदें रखी हुई हैं जिनमें 1 से 25 अंक तक नम्बर डाले गये हैं। दैव निदर्शन पद्धति से एक गेंद निकाला जाता है। (i) 5 या 6 के गुणित अंक में गेंद निकालने की क्या सम्भावना है ? (ii) 3 या 5 के गणित अंक में गेंद निकालने की क्या सम्भावना है ?
A bag contains 25 balls, numbered from 1 to 25. One ball is drawn at random, (i) what is the probability that a number is a multiple of 5 or 6? (ii) What is the probability that the number is a multiple of 3 or 5?

[उत्तर : (i) $\frac{5}{25}+\frac{4}{25}=\frac{9}{25}$ (*ii*) $\frac{8}{25}+\frac{5}{25}-\frac{1}{25}=\frac{12}{25}$]

8. यादृच्छिक रूप से चुने गये एक लीप वर्ष में 53 बुधवार या 52 गुरुवार होने की क्या प्रायिकता है ?
What is the probability that a leap year selected at random will contain 53 Wednesday or 53 Thursdays?

[उत्तर : $\frac{3}{7}$]

9. खेलने के पत्तों के साधारण पैक में से एक काला पत्ता या एक बादशाह खींचने की प्रायिकता क्या है ?
What is the probability of drawing a black card or a king from a pack of ordinary playing cards?

[उत्तर : $\frac{26}{52}+\frac{4}{52}-\frac{2}{52}=\frac{7}{13}$]

10. घटना A के अनुकूल संयोगानुपात 3 : 4 हैं तथा घटना B के प्रतिकूल संयोगानुपात 5:7 हैं। यदि $P(A \cap B)=\frac{5}{84}$ हो तो कम-से-कम एक घटना होने की प्रायिकता ज्ञात कीजिए।
The odds in favour of an event A are 3:4 and the odds against B are 5:7. Find the probability of the occurrence of at least one of them if $P(A \cap B) = 5/84$.

[उत्तर : $P(A \cup B)=\frac{3}{7}+\frac{7}{12}-\frac{5}{84}=\frac{80}{84}=\frac{20}{21}$]

11. किसी नगर में एक परिवार के पास टेलीविजन सेट होने की प्रायिकता 0.83 है; कपड़े धोने की मशीन होने की प्रायिकता 0.54, है, दोनों टेलीविजन सेट तथा कपड़े धोने की मशीन होने की प्रायिकता 0.45 है। परिवार के पास टेलीविजन सेट या कपड़े धोने की मशीन होने की क्या प्रायिकता है ?
In a city, the probability that a family has a television set is 0.83, a washing machine is 0.54 and both a television set and a washing machine is 0.45. What is the probability that a family has either a television set or a washing machine?

[उत्तर : $P(A \cup B) = 0.83 + 0.54 - 0.45 = 0.92$]

12. ताश की एक गड्डी में से एक पत्ता यादृच्छया खींचा जाता है। इस पत्ते के हुक्म का पत्ता या दरबारी पत्ता या दुक्की होने की प्रायिकता ज्ञात कीजिए।

From a pack of cards one card is drawn at random. Find the probability that the card is a spade or an honour (a court) or a deuce?

[उत्तर : $\frac{25}{52}$]

13. 17 कार्डों की एक गड्डी से जिनमें प्रत्येक पर 1,2,3....., 16,17 संख्याओं में से एक संख्या लिखी हुई है, यादृच्छया एक कार्ड खींचा जाता है। प्रायिकता ज्ञात करो कि इस कार्ड की संख्या 3 या 5 या 7 से विभाज्य हो।

From a set of 17 cards numbered 1,2,3......., 16, 17 one is drawn at random. What is the probability that the card drawn bears a number which is divisible by 3 or 5 or 7?

[उत्तर : $\frac{9}{17}$]

14. 1 से 30 तक संख्या लिखे हुए 30 टिकटों में एक टिकट यादृच्छिक रूप से चुना जाता है। उसको पुनः वापस कर दिया जाता है तथा दुबारा एक टिकट निकाला जाता है। (i) पहली बार में 5 या 7 का अपवर्त्य होने तथा (ii) दूसरी बार में 3 या 11 का अपवर्त्य होने की प्रायिकता ज्ञात कीजिए।

From 30 tickets marked with the first 30 numerals one is drawn at random. It is then replaced and a second drawn is made. Find the chance that in the first drawn (i) it is a multiple of 5 or 7, and (ii) in the second it is a multiple of 3 or 11.

[उत्तर : (i) $\frac{6}{30}+\frac{4}{30}=\frac{1}{3}$ (ii) $\frac{10}{30}+\frac{2}{30}=\frac{2}{5}$]

15. 50 टिकटों पर क्रमानुसार 1 से 50 तक अंक डाले गये हैं. उनसे एक अनायास निकाला जाता है। प्रायिकता बताइये कि निकाला गया टिकट 3 या 4 का गुणित होगा।

Digits from 1 to 50 are written on 50 tickets serially. One of them is drawn at random. Find the probability that the ticket drawn is a multiple of 3 or 4.

[उत्तर : $\frac{16}{50}+\frac{12}{50}-\frac{4}{50}=\frac{12}{25}$]

16. एक जोड़ा साधारण पासे का फेंका जाता है। योग 8 होने की प्रायिकता क्या है यदि यह ज्ञात हो कि योग सम संख्या है?

Two dice are thrown. What is the probability that the sum is 8 if it is known that this sum is an even number?

[उत्तर : $\frac{5/36}{18/36}=\frac{5}{18}$]

17. एक थैले में 5 लाल एवं 3 सफेद गेंदें हैं। एक दूसरे थैले में 4 लाल एवं 6 सफेद गेंदें हैं। यदि दैव निदर्शन से एक गेंद प्रत्येक थैले से निकाली जाती है, दोनों के लाल रंग के होने की क्या सम्भावना है?

A bag contains 5 red and 2 white balls. Another bag contains 4 red and 6 white balls. If one ball is drawn at random from each bag, what is the probability of their being red?

[उत्तर : $\frac{5}{8}\times\frac{4}{10}=\frac{1}{4}$]

18. एक ताश की गड्डी में से एक पत्ता यादृच्छिक रूप से निकाला जाता है। प्रायिकता क्या है कि वह एक (i) गुलाम होगा, (ii) लाल रंग का होगा, (iii) ईट का बादशाह होगा?

A card is drawn at random from a pack of playing cards. What is the probability that the card is (i) a knave, (ii) of red colour, (iii) the king of diamond.

[उत्तर : (i) $\frac{4}{52}=\frac{1}{13}$ (ii) $\frac{26}{52}=\frac{1}{2}$ (iii) $\frac{1}{52}$]

NOTES

19. लघुगणक सारणी में से एक अंक यादृच्छिक रूप से चुना जाता है। 'यह अंक 4 या 8 हो' की प्रायिकता होगी ?

What is the probability that a digit selected at random from the logarithmic table is 4 or 8?

[उत्तर : $\frac{1}{5}$]

20. एक पांसा फेंका जाता है। 4 या 6 का अंक आने की प्रायिकता ज्ञात कीजिए।

A dice is thrown. Find the probability of getting a 4 or 6.

[उत्तर : $\frac{1}{3}$]

21. एक साधारण पांसे को फेंकने पर 2 से अधिक संख्या आने की प्रायिकता बताइये।

Find the probability of getting more than 2 with a ordinary dice.

[उत्तर : $\frac{2}{3}$]

22. एक पिनोकिल गड्डी से बादशाह छांटने की क्या प्रायिकता है ? (पिनोकिल गड्डी में 2 इक्के, 2 बादशाह, 2 बेगम, 2 दहले तथा 2 नहले प्रत्येक सूट के होते हैं इनके अलावा कोई पत्ता नहीं होता है।)

What is the probability of selecting a king from a pinochle deck? (A pinochle deck consists of 1 ace, 2 kings, 2 queens, 2 jacks, 2 tens and 2 nines of each suit. There are no other cards.)

[उत्तर : $\frac{8}{48} = \frac{1}{6}$]

23. दो पासे एक साथ फेंके जाते हैं। प्रायिकता ज्ञात कीजिए :

Two dice are thrown simultaneously. Find the probability :

(i) पहले पासे पर विषम अंक प्राप्त होने की

of getting odd digit on first die

(ii) योग 9 होने की

of getting a sum of 9

(iii) योग 11 होने की

of getting a sum of 11

(iv) योग 9 नहीं होने की

of not throwing a total of 9.

[उत्तर : (i) $\frac{1}{2}$, (ii) $\frac{1}{9}$, (iii) $\frac{1}{18}$, (iv) $1 - \frac{1}{9} = \frac{8}{9}$]

24. 120 रेडियो के प्रतिदर्श में निम्न सूचना प्राप्त है :

In a sample of 120 radios the following information is available :

दोषों की संख्या (No. of defects) :	0	1	2	3
रेडियो की संख्या (No. of Radios) :	15	80	20	5

यादृच्छिक रूप से चयनित एक रेडियो दोष रहित होगा, इसकी क्या प्रायिकता है ?

What is the probability that a radio selected at random will have no defect?

[उत्तर : $\frac{15}{120} = \frac{1}{8}$]

25. एक थैले में 5 लाल, 8 सफेद और 10 काली गेंदें हैं। एक गेंद खींची जाती है। इस बात की क्या सम्भावना है कि खींची गई गेंद सफेद या काली होगी ?

A bag contains 5 red balls, 8 white balls and 10 black balls. A ball is drawn. What is the probability that the ball drawn will be either white or black?

[उत्तर : $\frac{18}{23}$]

26. एक थैले में 5 सफेद, 4 काली, 7 पीली तथा 6 लाल गेंदें हैं। एक गेंद यादृच्छिक रूप से निकालने पर उसके काली या लाल होने की क्या सम्भावना है ?

27. 'अ' 75 प्रतिशत अवस्थाओं में सच बोलता है और 'ब' 80 प्रतिशत अवस्थाओं में सच बोलता है। बताइये कि एक ही तथ्य का वर्णन करते समय वे कितनी प्रतिशत अवस्थाओं में एक-दूसरे के विरुद्ध बोलेंगे ?
A speaks truth in 75% cases and B in 80% cases. In what percentage of cases will they contradict each other in stating a same fact?
[उत्तर : $75 \times 2 + 25 \times .8 = 0.35$ अर्थात् 35%]

28. एक ठेकेदार को सड़क निर्माण का ठेका मिलने की प्रायिकता 4/9 है और उसे जल की टंकी बनाने का ठेका मिलने की प्रायिकता 5/7 है तो इस ठेकेदार को कम-से-कम एक ठेका मिलने की प्रायिकता बताइये।
The probability that a contractor will get a contract for road construction is 4/9 and the probability that he will get contract for the construction of a water tank is 5/7. What is the probability of getting at least one of the contract?
[उत्तर : $1 - \frac{5}{9} \times \frac{2}{7} = 1 - \frac{10}{63} = \frac{53}{63}$]

29. किसी परीक्षा में सांख्यिकी के दो पेपर I एवं पेपर II हैं। कोई विद्यार्थी पेपर I में पास हो जाता है, उसकी सम्भावना 40% है और वह पेपर II में पास होता है, उसकी सम्भावना 60% है। वह प्रायिकता क्या होगी जिसमें कि कोई विद्यार्थी (i) दोनों पेपरों में पास होगा? (ii) कम-से-कम एक पेपर में पास होगा ?
These are two papers (paper I and paper (ii) of statistics in a examination. The chance that a student will pass in paper I is 40% and he will pass in paper II is 60%. What is the probability that a student will pass in (i) both papers? (ii) will pass in at least one of the papers?
[उत्तर : (i) $\frac{40}{100} \times \frac{60}{100} = 0.24$, (*ii*) $1 - \frac{60}{100} \times \frac{40}{100} = 0.76$]

30. यह सम्भावना कि A एक प्रश्न हल कर लेगा, 2/3 है तथा यह सम्भावना कि B उसे हल कर लेगा, 5/12 है। यदि वे दोनों प्रयत्न करें तो (i) प्रश्न के हल किये जाने की सम्भावना ज्ञात कीजिए;
(ii) केवल एक प्रश्न को हल करे इसकी क्या प्रायिकता है ?
The probability of A question is 2/3 and that B solving the same question is 5/12. (i) Find the probability of solving the question if A and B together attempt the same question; (ii) what is the probability that only one of them solves the problems.
[उत्तर : (i) $1 - \frac{1}{3} \times \frac{7}{12} = \frac{29}{36}$, (ii) $\frac{2}{3} \times \frac{7}{12} + \frac{1}{3} \times \frac{5}{12} = \frac{19}{36}$]

31. शैलेष के पास एक लॉटरी योजना के 10 में से 5 टिकट हैं और दूसरी लॉटरी योजना के 10 में से 8 टिकट हैं। उसके इनाम आने की प्रायिकता क्या है ?
Sailesh has 5 tickets out of 10 tickets of a lottery scheme and 8 tickets out of 10 tickets of a second lottery scheme. What is the probability of his getting a prize?
[उत्तर : $1 - \frac{5}{10} \times \frac{2}{10} = \frac{9}{10}$]

32. सांख्यिकी का एक प्रश्न चार विद्यार्थियों अ, ब, स तथा द को दिया जाता है। उनके प्रश्न को हल करने की क्रमशः $\frac{1}{2}, \frac{3}{4}, \frac{1}{2}$ तथा $\frac{1}{3}$ सम्भावनाएं हैं। प्रश्न के हल हो जाने की क्या सम्भावना है ?
A problem in statistics is given to four students A, B, C and D. Their chances of solving it are $\frac{1}{4}, \frac{3}{4}, \frac{1}{2}$ and $\frac{1}{3}$ respectively. What is the probability that it will be solved?
[उत्तर : $1 - \frac{3}{4} \times \frac{1}{4} \times \frac{1}{2} \times \frac{2}{3} = 1 - \frac{1}{16} = \frac{15}{16}$]

33. तीन विद्यार्थियों A, B तथा C को एक प्रश्न हल करने को दिया जाता है। हल करने की उनकी सम्भावनाएं क्रमशः $\frac{1}{2}, \frac{1}{3}$ और $\frac{1}{4}$ हैं। प्रश्न को हल किये जाने की प्रायिकता ज्ञात कीजिए।

Three students A, B and C are given a problem to solve. Their chances of solving it are $\frac{1}{2}, \frac{1}{3}$ and $\frac{1}{4}$ respectively. Find out the probability that the problem will be solved.

[उत्तर : $1 - \left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{4}\right) = 1 - \frac{1}{2} \times \frac{2}{3} \times \frac{3}{4} = \frac{3}{4}$]

34. A 5 में से 4 बार निशाना लगा सकता है, B 4 में से 3 बार, C 3 में से 2 बार निशाना लगा सकता है। ये तीनों एक-एक गोली चलाते हैं। कम-से-कम दो का निशाना ठीक लगाने की प्रायिकता ज्ञात कीजिए।

35. बिना प्रतिस्थापन के अंकों 2, 3, 4, 5, 6, 7 से 4 अंकों की कितनी विभिन्न संख्याएं बन सकती हैं ?

How many different numbers of four digits can be formed with the digits 2, 3, 4, 5, 6, 7; none of the digits being repeated in any of the numbers so formed.

[उत्तर : ${}^6P_4 = 360$]

36. PRODUCT शब्द के अक्षरों से तीन-तीन अक्षरों के कितने विभिन्न शब्द बन सकते हैं ?

How many different words of 3 letters can be formed out of the letters of the word PRODUCT?

[उत्तर : ${}^7P_3 = 210$]

37. PROBLEM शब्द से 4 अक्षरों के कितने शब्द बन सकते हैं ?

How many different words of letters can be formed out of the letters of the word PROBLEM?

[उत्तर : ${}^7P_4 = 840$]

38. BANKER शब्द को कितनी प्रकार से क्रम में रखा जा सकता है ?

In how many ways can the word BANKER be rearranged?

[उत्तर : ${}^6P_6 = 6! = 720$]

39. COSTING शब्द के अक्षरों को कितने प्रकार से व्यवस्थित किया जा सकता है ?

In how many ways can the alphabets of the word COSTING be rearranged?

[उत्तर : ${}^7P_7 = 7! = 5{,}040$]

40. एक कमरे में 6 दरवाजे हैं। चार व्यक्ति इसमें प्रवेश करते हैं। यदि चारों व्यक्ति विभिन्न दरवाजों से प्रवेश करते हैं तो कुल कितने प्रकार से अन्दर आ सकते हैं ?

There are six doors in a room. Four persons have to enter it. In how many ways can they enter from different rooms?

[उत्तर : ${}^6P_4 = 360$]

41. एक कमरे में 8 कुर्सियाँ हैं। पांच विद्यार्थी उन पर कितनी तरह से बैठ सकते हैं ?

There are eight chairs in a room. In how many ways 5 students can sit on them?

[उत्तर : ${}^8P_5 = 6{,}820$]

42. मिस तरुलता, सुनीता, श्वेता और वीना श्रीमाया होटल पहुंचती हैं। वहां पहुंचकर देखती हैं कि होटल में 5 प्रवेश द्वार हैं। किस प्रकार चारों बहिनें विभिन्न रास्तों से प्रवेश कर सकती हैं ?

43. दो शहर A और B पांच सड़कों से जुड़े हुए हैं। यदि एक व्यक्ति शहर A से किसी एक सड़क से B शहर को जाये और वापस आये तो उसके उसी सड़क से वापस न आने की प्रायिकता क्या है ?

Two cities A and B are connected by five roads. If a person goes from city A to B by a road and returns, what is the probability that he does not return with the same road?

[उत्तर : $\frac{5 \times 4}{5 \times 5} = \frac{4}{5}$]

44. किसी नगर में 5 होटल हैं। यदि चार व्यक्ति इस नगर में आते हैं तो इसकी प्रायिकता ज्ञात कीजिए कि वे अलग-अलग होटलों में ठहरेंगे।

There are 7 hotels in a city. If 4 persons come to the city, find the probability that they will stay in different hotels.

[उत्तर : $\frac{5 \times 4 \times 3 \times 2}{5 \times 5 \times 5 \times 5} = \frac{24}{125}$]

45. एक थैले में 5 काली और 3 लाल गेंदें हैं। 2 गेंद एक साथ निकाली जाती हैं। दोनों गेंदों के एक ही रंग की होने की प्रायिकता ज्ञात करो।

A bag contains 5 black and 3 red balls. 2 balls are drawn at a time. Find the probability that both are of the same colour.

[उत्तर : $\frac{{}^5C_2}{{}^8C_2} + \frac{{}^3C_2}{{}^8C_2} = \frac{13}{28}$

वैकल्पिक विधि : P (दोनों अलग-अलग रंग की हों)

$= \frac{{}^5C_1 \times {}^3C_1}{{}^8C_2} = \frac{5 \times 3}{28} = \frac{15}{28}$

अभीष्ट प्रायिकता $= 1 - \frac{15}{28} = \frac{13}{28}$]

46. एक थैले में 13 लाल और 16 सफेद गेंदें हैं। 2 गेंद निकाली जाती हैं। दोनों गेंदों में एक ही रंग की होने की प्रायिकता ज्ञात कीजिए।

A bag contains 13 red and 16 white balls. 2 balls are drawn at random. Find the probability that both are of the same colour.

[उत्तर : $\frac{{}^{13}C_2}{{}^{29}C_2} + \frac{{}^{16}C_2}{{}^{29}C_2} = \frac{99}{203}$]

47. एक थैले में 8 लाल तथा 7 सफेद गेंदें हैं। एक बार में दो गेंदें निकाली जाती हैं। क्या सम्भावना है कि (i) दोनों गेंदें लाल होंगी, (ii) दोनों गेंदें सफेद होंगी, तथा (iii) एक गेंद लाल तथा दूसरी सफेद होगी?

A bag contains 8 red and 7 white balls. Two balls are drawn at random. What is the probability that (i) they are both red, (ii) they are both white, and (iii) one as red and the other white?

[उत्तर : (i) $\frac{28}{205}$, (ii) $\frac{21}{105}$, (iii) $\frac{56}{105}$]

48. एक थैले में दस गेंदों में तीन लाल रंग की हैं। यदि थैले में से दो गेंदें एक साथ निकाली जायें तो क्या सम्भावना है कि उनमें से (i) कम-से-कम एक गेंद लाल रंग की होगी? (ii) दोनों लाल रंग की होंगी?

In a bag 3 balls are red out of 10 balls. If two balls are drawn at a time from the bag, what is the probability that (i) at least one ball is of red colour and (ii) both are of red colour?

[उत्तर : (i) $\frac{{}^3C_1 \times {}^7C_1 \times {}^3C_2}{{}^{10}C_2} = \frac{8}{15}$, (ii) $\frac{{}^3C_2}{{}^{10}C_2} = \frac{1}{15}$]

49. एक बर्फ विक्रेता मौसम गरम होने पर 200 रु. प्रति दिन कमा सकता है, परन्तु मौसम ठण्डा होने पर उसे 75 रु. की हानि उठानी पड़ती है। बर्फ विक्रेता की प्रति दिन प्रत्याशित आमदनी बताइये। मौसम गरम होने की प्रायिकता 0.6 है।

An ice-seller earns Rs. 200 per day in hot season but he looses Rs. 75 in the cold season. Find his expectation per day. The probability of hot season in 0.6.

[उत्तर : 90 रु.]

NOTES

50. एक कम्पनी के विभिन्न लाभों की प्रायिकताएं निम्न हैं :

The probability of the profits of a company are as follows :

लाभ रुपये (Profit Rs.) : 10.000 20.000 25.000 40.000

प्रायिकता (Probability) : 0.3 0.2 0.4 0.1

कम्पनी का प्रत्याशित लाभ ज्ञात कीजिए।

Find expected profit of the company.

[उत्तर : 21,000 रु.]

51. एक थैले में 3 लाल और 4 हरी गेंदें हैं। एक व्यक्ति को थैले में से 1 गेंद निकालने को कहा गया है। शर्त यह है कि यदि वह लाल गेंद निकाले तो उसे 30 रु. मिलेंगे किन्तु हरी गेंद निकालने पर उसे 20 रु. की हानि होगी। उसकी आकांक्षा क्या हो सकती है?

A bag contains 3 red and 4 green balls. A man is asked to draw one balls from the bag and is reminded that if the ball be drawn is red he will win Rs. 30 but if it is green he will lose Rs. 20. Find his expectation.

[उत्तर : शून्य]

52. यदि किसी दिन वर्षा हो जाती है तो एक रेनकोट बेचने वाला 500 रु. कमा सकता है, परन्तु वर्षा न होने पर उसकी आय केवल 50 रु. होगी। यदि किसी दिन वर्षा होने की प्रायिकता 0.25 हो तो उसकी प्रत्याशा क्या होगी?

If it rains, a raincoat dealer can earn Rs. 500 per day, but on a fair day his earning will be Rs. 50 only. What is his expectation if the probability of rain is 0.25.

[उत्तर : 387.50 रु.]

53. एक व्यापारी सेबों की एक खेप को 10,000 रु., 8.500 रु., 7,500 रु. तथा 7,000 रु. में बेच सकेगा। इसकी संम्भाविताएं क्रमशः 0.1, 0.2, 0.3 तथा 0.4 हैं। वह एक खेप 8,000 रु. में खरीदता है। इस सौदे में उसके लाभ का प्रत्याशित मूल्य निकालिए।

A trader can sell a consignment of apples in Rs. 10,000, Rs. 8,500, Rs. 7,500 and Rs. 7,000. Its probabilities are 0.1, 0.2, 0.3 and 0.4 respectively. He purchase a consignment in Rs. 8,000. Find the expected value of his profit in this transaction.

[उत्तर : प्रत्याशित लाभ = प्रत्याशित विक्रय मूल्य - क्रय मूल्य = 7,750– 8,000 = – 250 रु.]

54. माना X किसी व्यक्ति के व्यापार में लाभ को व्यक्त करता है। उसे 2,800 रु. का लाभ होने की प्रायिकता 0.5 है, 5,500 रु. की हानि होने की प्रायिकता 0.3 है तथा न लाभ और न हानि होने की प्रायिकता 0.2 है। X की गणितीय प्रत्याशा ज्ञात कीजिए।

Let X denote the profit that a man makes in a business. He may earn Rs. 2,800 with a probability 0.5, he may lose Rs. 5,500 with a probability 0.3 and he may neither earn not lose with probability 0.2. Calculate the mathematical expectation of X.

[उत्तर : $E(X) = -250$ रु.]

55. रेफ्रीजरेटरों का एक विक्रेता अपने पुराने अनुभव से एक दिन में X रेफ्रीजरेटर बेचने की प्रायिकता अनुमानित करता है। ये प्रायिकताएं निम्न प्रकार हैं :

A dealer in refrigerators estimates from his past experience the probability of his selling X refrigerators in a day. These are as follows :

$X = x$	0	1	2	3	4	5	6
$P(X = x)$	0.03	0.20	0.23	0.15	0.12	0.10	0.07

Find the expected number of refrigerators sold in a day.

[उत्तर : $E(X) = 2.81$]

56. एक थैले में 3 लाल 4 हरी गेंदें हैं। एक व्यक्ति उसमें से दो गेंदें निकालता है। यदि प्रत्येक लाल गेंद निकालने के 35 रुपये मिलें और प्रत्येक हरी गेंद निकालने के लिए 14 रुपये मिलें तो उसकी प्रत्याशा क्या होगी ?

A bag contains 3 red and 4 green balls. A man draws 2 balls from this bag. If he is to receive Rs. 35 for every red ball which he draws and Rs. 14 for every green ball, what is his expectation?

[उत्तर : 46 रु.]

57. एक कारखाने की दो मशीनों में से प्रथम मशीन 45% तथा द्वितीय मशीन 60% उत्पादन करती है। प्रथम मशीन का 10% तथा द्वितीय मशीन का 8% उत्पादन खराब है। यदि एक दिन के उत्पादन में से चुनी गयी एक वस्तु खराब पायी जाती है तो उसके प्रथम मशीन द्वारा उत्पादित होने की प्रायिकता क्या है ?

Out of two machines in a factory first machine produces 40% and second machine 60%. 10% of machine first and 8% of machine second is defective. An item is selected from a day's output and is found defective, what is the probability of its being produced by machine first?

[उत्तर : $\frac{5}{11}$]

58. बोल्ट के एक कारखाने में वोल्ट के उत्पादन की तीन मशीनें A, B और C क्रमशः 25%, 35% तथा 40% उत्पादन करती है। उनके उत्पादन में क्रमशः 5%, 4% व 2% बोल्ट दूषित पाये जाते हैं। दैव रीति से एक वोल्ट चुना जाता है और यह दूषित पाया जाता है। इसकी क्या सम्भावना है कि उसका उत्पादन मशीन C द्वारा किया गया था ?

In a bolt factory, three machines A, B and C manufacture 25, 35 and 40 percent of the total output. Of their output 5, 4 and 2 percent are defective respectively. A bolt is drawn at random and found to be defective. What is the probability that it was manufactured by machine C.

[उत्तर : $\dfrac{\frac{40}{100}\times\frac{2}{100}}{\frac{25}{100}\times\frac{5}{100}+\frac{35}{100}\times\frac{4}{100}+\frac{40}{100}\times\frac{2}{100}}=\frac{16}{69}$]

59. एक कारखाने में मशीन M_1, M_2 तथा M_3 क्रमशः कुल उत्पादन का 30%, 30% तथा 40% बनाती है। उनके उत्पादन में क्रमशः 1%, 3% तथा 2% वस्तुएं खराब होती हैं। एक दिन के उत्पादन में से एक वस्तु ली जाती है जो खराब पायी गई। इसकी क्या प्रायिकता है कि :

In a factory, machine M_1, M_2 and M_3 manufacture respectively 30. 30 and 40 percent of the total output. of their output 1, 3 and 2 percent are defective items. An item is drawn from a day's output and is found defective. What is the probability that :

(i) वह वस्तु मशीन M_1 द्वारा, मशीन M_2 द्वारा मशीन M_3 द्वारा बनायी गयी ?
it was manufactured by M_1 , by M_2, by M_3?

(ii) वह वस्तु मशीन M_1 या M_3 द्वारा बनायी गई ?
it was manufactured by M_1 *or* M_3?

[उत्तर : (i) $\frac{3}{20}, \frac{9}{20}, \frac{8}{20}$, *(ii)* $\frac{11}{20}$]

60. A 6 दशाओं में से 5 में सच बोलता है और वह यह कहता है कि 9 काली और 1 सफेद गेंद वाले एक थैले में से एक सफेद गेंद निकाली गई। क्या सम्भावना है कि वास्तव में सफेद गेंद ही निकाली गयी थी ?

A speaks truth 5 times out of 6 and he says that a white ball is drawn from a bag consisting of 9 black balls and 1 white ball. What is the probability that the ball drawn was really white?

$$[\text{उत्तर : } \frac{\frac{5}{6}\times\frac{1}{10}}{\frac{5}{6}\times\frac{1}{10}+\frac{1}{6}\times\frac{9}{10}} = \frac{5}{5+9} = \frac{5}{14}]$$

NOTES

61. एक बॉक्स में 3 खराब और 7 अच्छी वस्तुएं हैं तथा दूसरे बॉक्स में 1 खराब तथा 9 अच्छी वस्तुएं हैं। किसी एक बॉक्स को यादृच्छिक रूप से चुना जाता है तथा फिर उसमें से एक वस्तु निकाली जाती है ?

A box contains 3 defective and 7 non-defective items and another box contains 1 defective and 9 non-defective items. A box is selected at random and an item is drawn from it.

(i) खराब वस्तु निकालने की क्या प्रायिकता है ?

What is the probability of drawing a defective item?

(ii) यदि खराब वस्तु चुनी गयी हो तो उसके पहले बॉक्स में से आने की प्रायिकता क्या है ?

What is the probability that first box was chosen, given a defective item is drawn.

$$[\text{उत्तर : } (i)\ \frac{1}{2}\times\frac{3}{10}+\frac{1}{2}\times\frac{1}{10} = \frac{4}{20}, (ii)\ \frac{\frac{1}{2}\times\frac{3}{10}}{\frac{1}{2}\times\frac{3}{10}+\frac{1}{2}\times\frac{1}{10}} = \frac{3}{4}]$$

62. आपने अनुभव किया है कि आप जितनी बार अपने अधिकारी के पास गये हैं, उसमें से 60% अवसरों पर वे प्रसन्न दिखाई पड़े ? अतः आपने उनके पास जाने पर प्रसन्न दिखाई पड़ने की संभावना 0.6 या $\frac{6}{10}$ मान ली। आपने यह भी अनुभव किया है कि यदि वे प्रसन्न दिखाई पड़ते हैं तो इस बात की सम्भावना कि वे आपकी प्रार्थना को स्वीकार कर लेंगे 0.4 या $\frac{4}{10}$ है, और यदि प्रसन्न नहीं हैं तो आपकी प्रार्थना स्वीकार करने की सम्भावना 0.1 या $\frac{1}{10}$ है। आप एक दिन उनके यहां गये और उन्होंने आपकी प्रार्थना स्वीकार कर ली। बताइये कि उस दिन उनके प्रसन्न रहने की संभावना क्या है ?

You note that your officer is happy on 60 percent of your calls, so you assign a probability of his being happy on your visit as 0.6 or $\frac{6}{10}$. You have noticed also that if he is happy, he accedes to your request with a probability of 0.4 or $\frac{4}{10}$, whereas if he is not happy, he accedes to your request with a probability of 0.1 or $\frac{1}{10}$. You call one day, and he accedes to your request. What is the probability of his being happy on that day?

$$[\text{उत्तर : } \frac{\frac{6}{10}\times\frac{4}{10}}{\frac{6}{10}\times\frac{4}{10}+\frac{4}{10}\times\frac{1}{10}} = \frac{24}{28} = \frac{6}{7}]$$

63. कलश A_1 में 8 काली तथा 2 सफेद गोलियां हैं, कलश A_2 में 3 काली तथा 7 सफेद गोलियां हैं और कलश A_3 में 5 काली गोलियां हैं। एक साधारण पासा फेंका जाता है। यदि पासे पर 1,2 या 3 आता है तो कलश A_1 में से एक गोली चुनी जाती है। यदि पासे पर 4 या 5 आता है तो कलश A_2 में से एक गोली चुनी जाती है। अन्तिम रूप से यदि पासे पर 6 आता है तो एक गोली कलश A_3 में से चुनी जाती है। यदि चुनी गयी गोली काली हो तो इसके A_2 में से आने की क्या प्रायिकता है ?

Urn A_1 contains 8 black and 2 white marbles, Urn A_2 contains 3 black and 7 white marbles and urn A_3 contains 5 white and 5 black marbles. A fair die is to be cast. If the die turns up 1, 2 or 3, then a marble will be selected from A_1 . If the die turns up 4 or 5, a marble will be selected from A_2 . Finally a marble will be selected from A_3, if the die turns up 6. Given that the marble selected is black, what is the probability that the marble was from urn A_2 ?

[उत्तर : $\dfrac{\frac{2}{6}\times\frac{3}{10}}{\frac{3}{6}\times\frac{8}{10}+\frac{2}{6}\times\frac{3}{10}+\frac{1}{6}\times\frac{5}{10}} = \dfrac{6}{35}$]

64. एक थैले में 5 सफेद और 4 काली गेंदें हैं। इस थैले में से एक गेंद निकाली जाती है और वापस रख दी जाती है तब दूसरे ड्रा से एक गेंद और निकाली जाती है। इसकी क्या प्रायिकता है कि निकाली गयी गेंदें अलग-अलग रंग की थीं?

A bag contains 5 white and 4 black balls. A ball is drawn from this bag and replaced and then a second draw of a ball is made. What is the probability that the balls drawn were of different colours?

[उत्तर : $\frac{5}{9}\times\frac{4}{9}+\frac{4}{9}\times\frac{5}{9}=\frac{40}{81}$]

65. एक थैले में 5 सफेद और 7 लाल गेंदें हैं। एक गेंद निकाली जाती है और वापस रख दी जाती है। क्या प्रायिकता है कि एक सफेद और एक लाल गेंद इसी क्रम में निकाली जाती है?

There are 5 white and 7 red balls in a bag. A ball is drawn and then replaced. What is the probability that a white and a red ball are drawn in that order?

[उत्तर : $\frac{5}{12}\times\frac{7}{12}=\frac{35}{144}$]

66. एक थैले में 6 सफेद और 12 काली गेंदें हैं। दैव निदर्शन से एक साथ तीन गेंदें निकालने पर सभी सफेद आने की सम्भावना क्या है?

A bag contains 6 white and 12 black balls. What is the probability of getting all white balls if three balls are drawn at random?

[उत्तर : $\frac{^{6}C_3}{^{18}C_3}=\frac{6\times5\times4}{18\times17\times16}=\frac{5}{204}$]

67. एक थैले में 3 काली, 3 सफेद और 4 लाल गेंदें हैं। तीन गेंद यादृच्छिक रूप से निकाली जाती हैं। इनमें एक काली, एक सफेद और एक लाल होने की क्या प्रायिकता है?

A bag contains 3 black, 3 white and 3 red balls. Three balls are drawn at random. What is the probability of drawing a black, a white and a red ball?

[उत्तर : $\frac{3}{10}$]

68. एक थैले में 6 सफेद और 4 काली गेंदें हैं। एक दूसरे थैले में 4 सफेद और 8 काली गेंदें हैं। दोनों में से एक थैला यादृच्छिक रूप से चुना जाता है और इसमें से 2 गेंदें निकाली जाती हैं। प्रायिकता ज्ञात कीजिए कि इनमें एक सफेद और एक काली गेंद है।

A bag contains 6 white and 4 black balls and a second one 4 white and 8 black balls. One of the bags in chosen at random and draw of 2 balls is made from it. Find the probability that one is white and the other is black.

[उत्तर : $\frac{1}{2}\times\frac{^{6}C_1\times{}^{4}C_1}{^{10}C_2}+\frac{1}{2}\times\frac{^{4}C_1\times{}^{8}C_1}{^{12}C_2}=\frac{4}{15}+\frac{8}{33}=\frac{84}{165}=0.51$]

69. एक थैले में से जिसमें 4 काली और 5 लाल गेंदें हैं, 3 गेंद निकाली जाती हैं। वे सब काली होंगी इसकी क्या प्रायिकता है?

From a bag containing 4 black and 5 red balls a draw of 3 balls is made. What is the probability that all of them would be black?

[उत्तर : $\frac{^{4}C_3}{^{9}C_3}=\frac{4}{84}=\frac{1}{21}$]

NOTES

70. ताश के एक खेल में एक विशिष्ट खिलाड़ी को चारों बादशाह मिलने की प्रायिकता क्या है ?
What is the chance that a particular player gets all the four kings?

[उत्तर : $\frac{{}^{4}C_4 \times {}^{48}C_9}{{}^{52}C_{13}} = \frac{11}{4,165}$]

71. एक थैले में 3 सफेद, 5 काली तथा 7 लाल गेंदें हैं। दो गेंद यादृच्छिक रूप से चुनी जाती हैं। वे दोनों लाल होंगी इसकी क्या प्रायिकता है ?
A bag contains 3 white, 5 black and 7 red balls. Two balls are drawn at random. Find the probability that they will both be red?

[उत्तर : $\frac{{}^{7}C_2}{{}^{15}C_2} = \frac{21}{105} = \frac{1}{5}$]

72. एक थैले में 4 सफेद, 5 लाल और 6 हरी गेंदें हैं। तीन गेंद यादृच्छिक रूप से निकाली जाती हैं। एक . सफेद, एक लाल तथा एक हरी गेंद निकालने की प्रायिकता क्या है ?
A bag contains 4 white, 5 red and 6 green balls. Three balls are drawn at random. What is the probability that a white, a red and a green ball are drawn?

[उत्तर : $\frac{{}^{4}C_1 \times {}^{5}C_1 \times {}^{6}C_1}{{}^{52}C_3} = \frac{24}{91}$]

73. ताश के 52 पत्तों की गड्डी से 2 पत्ते यादृच्छिक रूप से निकाले जाते हैं। इनमें एक बादशाह तथा एक बेगम होने की प्रायिकता ज्ञात कीजिए।
From a pack of 52 cards, 2 are drawn at random. Find the chance that one is king and the other is queen.

[उत्तर : $\frac{{}^{4}C_1 \times 4C_1}{{}^{52}C_2} = \frac{8}{663}$]

74. यदि 52 पत्तों की गड्डी में से 4 पत्ते निकाले जाते हैं तो उनके विभिन्न सूटों के होने की प्रायिकता क्या है ? वे सब इक्के हों इसकी क्या प्रायिकता है ? इनमें 2 लाल तथा 2 काले होने की क्या प्रायिकता है ?
If 4 cards are drawn from a pack of 52 cards, what is the probability that they belong to different suits? What is the probability that all are aces. What is the probability that two are black and two are red?

[उत्तर : $\frac{{}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1 \times {}^{13}C_1}{{}^{52}C_4} = \frac{2,197}{20,825}, \frac{{}^{4}C_4}{{}^{52}C_4} = \frac{325}{833}$]

75. 10 बिजली के बल्वों में से 4 खराब हैं। यादृच्छिक रूप से तीन बल्वों का चयन किया जाता है। प्रायिकता ज्ञात कीजिए- (i) एक भी बल्व खराब न हो तथा (ii) एक बल्व खराब हो।
Out of 10 electric bulbs 4 are defective. Three balls are selected at random. Find the probability that (i) no bulb is defective and (ii) one bulb is defective.

[उत्तर : (i) $\frac{{}^{6}C_3}{{}^{10}C_3} = \frac{1}{6}$, (ii) $\frac{{}^{4}C_1 \times {}^{6}C_2}{{}^{10}C_3} = \frac{2}{3}$]

76. एक पासे को तीन बार उछालने पर कम-से-कम एक बार 6 आने की क्या प्रायिकता है ?
What is the probability of throwing 6 with a die at least once in 3 attempts?

[उत्तर : $1- \left(1 - \frac{1}{6}\right)^3 = 1 - \left(\frac{5}{6}\right)^3 = 1 - \frac{125}{216} = \frac{91}{216}$]

77. ताश की एक गड्डी को फेंटते समय अकस्मात तीन पत्ते गिर जाते हैं। इस बात की सम्भावना ज्ञात कीजिए कि ये तीनों पत्ते अलग-अलग सूट के हैं?

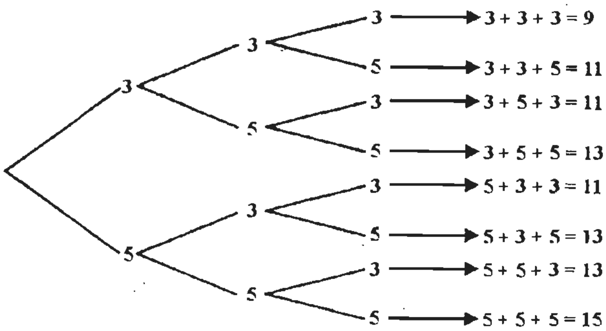
Three cards are dropped accidently while shuffling a pack of playing cards. What is the probability that they belong to different suits?

[उत्तर : $\frac{{}^{4}C_3 \times 13^3}{{}^{52}C_3}$]

78. एक सिक्के को जिसके एक ओर 3 तथा दूसरी ओर 5 लिखा है, तीन बार उछाला जाता है। योग (i) 9 होने की (ii) 11 होने की (iii) 13 होने की (iv) 15 होने की (v) 11 से अधिक होने की (vi) 13 से अधिक न होने की प्रायिकता ज्ञात कीजिए।

A coin whose faces are marked 3 and 5 is tossed three times, Find the probability of getting a sum of (i) 9 (ii) 11 (iii) 13 (iv) 15 (v) more than 11 (vi) note more than 13.

[संकेत- वृक्ष चित्र की सहायता से



कुल तरीके $= 2 \times 2 \times 2 = 8$

(i) $\frac{1}{8}$ (ii) $\frac{3}{8}$ (iii) $\frac{3}{8}$ (iv) $\frac{1}{8}$ (v) $\frac{3}{8} + \frac{1}{8} = \frac{1}{2}$ (vi) $\frac{1}{8} + \frac{3}{8} = \frac{1}{2}$]

79. 7 लड़के और 5 लड़कियां एक पंक्ति में खड़े होते हैं। किन्हीं भी दो लड़कियों के साथ-साथ खड़ी न होने की क्या प्रायिकता है?

7 boys and 5 girls stand in a row. What is the probability that no two girls are adjacent to each other?

[उत्तर : $\frac{7! \times {}^{8}P_3}{12!} = \frac{7}{99}$]

80. ताश के 52 पत्तों की गड्डी में से चार ताश बिना पुनःस्थापन के खींचे जाते हैं। प्रायिकता ज्ञात करो कि :

Out of 52 cards four are drawn at random without replacement. What is the probability that :

(i) वे भिन्न-भिन्न नामों के हों,
they belong to different denominations.

(ii) वे भिन्न-भिन्न जाति (सूट) एवं भिन्न नामों के हों.
they belong to different suits and different denominations,

(iii) वे दरबारी ताशों में से हों।
they are court cards.

[उत्तर : (i) $\frac{{}^{13}C_4 \times 4^4}{{}^{52}C_4} = \frac{52 \times 48 \times 44 \times 40}{52 \times 51 \times 50 \times 49} = \frac{2816}{4,165}$ (ii) $\frac{52 \times 36 \times 22 \times 10}{52 \times 51 \times 50 \times 49}$

या $\frac{{}^{13}C_1 \times {}^{12}C_1 \times {}^{11}C_2 \times {}^{10}C_1}{{}^{52}C_4} = \frac{264}{4,165}$ (iii) $\frac{{}^{12}C_4}{{}^{52}C_4} = \frac{99}{54,745}$]

NOTES

81. एक कमरे में तीन बल्व होल्डर हैं। 10 बल्वों में से जिनमें 6 बल्व खराब हैं, एक व्यक्ति यादृच्छया 3 बल्व चुनता है और बल्व होल्डरों में लगा देता है। उसे प्रकाश मिल जाय इसकी क्या प्रायिकता है?

A room has three bulb holders. From a collection of 10 light bulbs of which 6 are not good, a person selects 3 at random and puts them in the sockets. What is the probability that he will have light?

[उत्तर : $1 - \frac{1}{6} = \frac{5}{6}$]

82. 12 व्यक्तियों की एक पार्टी में 5 औरतें हैं। 2 व्यक्ति यादृच्छिक रूप से चुने जाते हैं। दोनों में आदमी होने की क्या प्रायिकता है?

In a part of 12 persons, 7 are men and 5 are ladies. Two persons are selected at random. What is the probability of both being men?

[उत्तर : $\frac{{}^{7}C_2}{{}^{12}C_2} = \frac{7}{22}$]

83. एक लॉटरी में 3 इनाम की और 6 खाली टिकटें हैं, उसमें A के 3 हिस्से हैं। दूसरी लॉटरी में। इनाम की और 2 खाली टिकटें हैं, उसमें B का एक हिस्सा है। दिखाइये कि A के इनाम जीतने की प्रायिकता B के जीतने की प्रायिकता से अधिक है जिनका अनुपात 16:7 है।

A has 3 share in a lottery in which these are three prizes and 6 blanks. B has one share in another, where these is but I prize and 2 blanks. Show that A has a better chance of winning a prize than B in the ratio of 16 to 7.

[संकेत- P(A के इनाम जीतने की) $= \frac{{}^{6}C_2}{{}^{9}C_3} = 1 - \frac{5}{21} = \frac{16}{21}$

P (B के इनाम जीतने की) $= 1 - \frac{{}^{2}C_1}{{}^{3}C_1} = \frac{1}{3}$

अभीष्ट अनुपात $\frac{16}{21} : \frac{1}{3}$ अर्थात् 116 : 7]

84. ताश की एक गड्डी में से 4 पत्ते एक साथ निकाले जाते हैं। यदि यह क्रिया पत्ते वापस रखते हुए तीन बार दोहराई जाय तो प्रत्येक बार चारों बादशाह आने की प्रायिकता बताइये।

4 cards are drawn at a time from a pack of playing cards. If this act is repeated thrice after replacing the cards every time, find the probability that they are all kings every time.

[उत्तर : $\left(\frac{{}^{4}C_4}{{}^{52}C_4}\right)^3 = \left(\frac{1}{2,78,725}\right)^3$]

85. 20 व्यक्तियों के एक समूह में 5 महिलाएं हैं। यदि 20 में से 3 व्यक्ति दैव आधार पर चुने जाते हैं तो (i) सभी महिला होने की क्या प्रायिकता है तथा (ii) कम-से-कम एक महिला होने की क्या प्रायिकता है?

In a group of 20 persons 5 are females. If 3 persons are picked out of 20 at random (i) what is the probability that all are females and (ii) what is the probability of at least one being female?

[उत्तर : (i) $\frac{{}^{5}C_3}{{}^{20}C_3} = \frac{1}{114}$, $(ii)\ 1 - \frac{{}^{15}C_3}{{}^{20}C_3} = \frac{137}{228}$]

86. उत्पादन विभाग के 3 ऑफिसर, क्रय विभाग के 4 ऑफिसर, बिक्री विभाग के 2 ऑफिसर तथा 1 चार्टर्ड एकाउण्टेण्ड से 4 व्यक्तियों की एक समिति बनानी है। निम्न प्रकार समिति बनाने की प्रायिकता ज्ञात कीजिए:

A committee of 4 people is to be appointed from 3 officers of the production department, 4 officers of the purchase department, 2 officers of the sales department and 1 chartered accountant. Find the probability of forming the committee in the following manner :

(i) प्रत्येक कैटेगरी में से एक अवश्य हो।
There must be one from each category.

(ii) क्रय विभाग से कम-से-कम एक व्यक्ति अवश्य हो।
It should have at least one from the purchase

(iii) समिति में चार्टर्ड एकाउण्टैण्ड अवश्य होना चाहिए।
The chartered accountant must be in the committee.

[उत्तर : (i) $\frac{4}{35}$, *(ii)* $\frac{13}{14}$, *(iii)* $\frac{2}{5}$]

87. किसी बस में 10 स्थानों पर चार व्यक्ति कितनी तरह से बैठ सकते हैं ?
In how many ways can ten seats in a Bus be occupied by four passengers?
[उत्तर : $^{10}P_4 = 10 \times 9 \times 8 \times 7 = 5{,}040$]

88. एक प्रतीक्षालय में चार कुर्सियां हैं। तीन व्यक्तियों को कितने प्रकार से बैठाया जा सकता है, यदि प्रत्येक व्यक्ति एक पृथक कुर्सी पर ही बैठे ?
There are four chairs in a waiting room. In how many ways three persons can sit if each persons sit on a separate chair.
[उत्तर : $^{4}P_3 = 4 \times 3 \times 2 = 24$]

89. 10 प्रतियोगियों को प्रथम, द्वितीय व तृतीय पुरुस्कार कितने प्रकार से वितरित किये जा सकते हैं ?
In how many ways first, second and third prizes can be distributed to three of ten competitors?
[उत्तर : $^{10}P_3 = 10 \times 9 \times 8 = 720$]

90. चार यात्री एक बस में सवार होते हैं जिसमें 6 सीट खाली हैं। वे यात्री कितने विभिन्न प्रकार से बैठ सकते हैं ?
Four passengers board a bus in which there are 6 empty seats. In how many different ways can they be seated?
[उत्तर : $^{6}P_4 = \frac{6!}{(6-4)!} = \frac{6 \times 5 \times 4 \times 3 \times 2 \times 1}{2 \times 1} = 360$]

91. एक कमरे में 8 कुर्सियां हैं। यदि 8 व्यक्ति कमरे में प्रवेश करते हैं तो वे कितने प्रकार से कुर्सियों पर बैठ सकते हैं ?
There are eight chairs in a room. If 8 persons enter the room, in how many ways they can sit on the chairs?
[उत्तर : $8! = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40{,}320$]

92. 6 पारितोषिक 3 व्यक्तियों को कितने प्रकार से दिये जा सकते हैं, यदि एक व्यक्ति को कितने भी पारितोषिक मिलें ?
In how many ways 6 prizes be given to 3 persons, if a person may get any number of prizes?
[उत्तर : $3^6 = 729$]

93. 3 काली, 5 सफेद तथा 7 लाल गेंदों में से कितने प्रकार से गेंदें चयनित की जा सकती है ताकि प्रत्येक रंग की कम-से-कम एक गेंद हो ?
In how many ways can a selection be made from 3 black, 5 white and 7 red balls so as to have at least one of each colour?
[उत्तर : $(2^3 - 1)(2^5 - 1)(2^7 - 1) = 27{,}559$]

94. DEED शब्द के अक्षरों से कितने विन्यास बनाये जा सकते हैं ?

How many different arrangements can be formed out of the word DEED?

[उत्तर : $\frac{4!}{2!2!} = 6$]

95. BANANA शब्द के अक्षरों को कितने प्रकार से विन्यसित किया जा सकता है ?

In how many ways can the letters of the word 'BANANA' be arranged?

[उत्तर : $\frac{6!}{3!2!} = 60$]

96. 'ELEMENT' शब्द के अक्षरों को कितने ढंग से अनुविन्यसित किया जा सकता है ?

In how many ways the letters of the word 'ELEMENT' be arranged?

[उत्तर : $\frac{7!}{3!} = 840$]

97. 'EXAMINATION' शब्द के अक्षरों को कितने ढंग से अनुविन्यसित किया जा सकता है ?

In how many ways the letters of the word 'EXAMINATION' be arranged?

[उत्तर : $\frac{11!}{2!2!2!2!}$]

98. निम्न शब्दों के अक्षरों को एक साथ लेकर कितने क्रमचय अलग-अलग बनाये जायेंगे ?

How many permutations separately can be formed from the letters of the following words?

(i) MAC MILLAN and (ii) STATISTICALLY

[उत्तर : (i) $\frac{9!}{2!2!2!} = 15{,}120$, (*ii*) $\frac{13!}{3!2!2!2!2!}$ 6,48,64,800]

99. COMMITTEE शब्द को कितने बार क्रम में रखा जा सकता है ?

In how many ways can the word 'COMMITTEE' be arranged?

[उत्तर : $\frac{9!}{2!2!2!}$]

100. 'ASSESSMENT' शब्द को कितने ढंग से अनुविन्यसित किया जा सकता है ?

In how many ways can the word 'ASSESSMENT' be arranged?

[उत्तर $\frac{10!}{4!2!}$]

101. 'INDEPENDENCE' शब्द के अक्षरों को कितनी प्रकार से व्यवस्थित किया जा सकता है ?

In how many ways can the word INDEPENDENCE be arranged?

[उत्तर : $\frac{12!}{3!2!4!}$]

102. PROBABILITY शब्द के अक्षरों से कितने अन्य क्रमचय बनाये जा सकते हैं ?

How many other permutations can be made from the letters of the word 'PROBABILITY'?

[उत्तर : $\frac{11!}{2!2!} - 1$]

103. कितनी अन्य रीतियों से 'SIMPLETON' शब्द के अक्षरों को विन्यसित किया जा सकता है ?

By how many other methods, the letters of the word 'SIMPLETON' can be arranged?

[उत्तर : $9! - 1$]

104. 'CALCULATION' शब्द में काम आने वाले अक्षरों से कितने विन्यास बनाये जा सकते हैं ?
In how many ways can the alphabets of the word 'CALCULATION' be arranged?

[उत्तर : $\frac{11!}{2!2!2!}$]

105. 'MINIMUM' शब्द को कितने विभिन्न क्रम में लिखा जा सकता है ?
In how many different ways can the letters of the word MINIMUM be arranged?

[उत्तर : $\frac{7!}{3!2!} = 420$]

106. 'REGRESSION' शब्द को कितने प्रकार से क्रमबद्ध किया जा सकता है ?
In how many ways can the word 'REGRESSION' be arranged?

[उत्तर : $\frac{10!}{2!2!2!}$]

107. MANGO शब्द के अक्षरों के कितने क्रमचय होंगे यदि पहला अक्षर M सर्वदा प्रारंभ में रहे ?
In how many ways the letters of the word 'MANGO' can be arranged so that M is always in the beginning?

[उत्तर : $1 \times {}^4P_4 = 4! = 24$]

108. 'STRANGE' शब्द के अक्षरों को कुल कितनी विधियों से व्यवस्थित किया जा सकता है ताकि स्वर (A व E) विषम स्थानों पर ही आयें ?
In how many ways can the letters of the word 'STRANGE' be arranged so that the the vowels (A and E) may appear in the odd places?

[उत्तर : ${}^4P_2 \times 5! = 1{,}440$]

109. LAHORE शब्द के अक्षरों से कितने अक्षर बन सकते हैं (i) जिनके प्रारम्भ में L तथा अन्त में E हो तथा (ii) जिनके शीर्षों पर L तथा E हों।
How many words can be formed by the letters of the word 'LAHORE' so that (i) they begin with L and end with E, and (ii) L and E occupy end place?

[उत्तर : (i) 4!, (ii) $4! \times 2$]

110. PERSONAL शब्द के चार विभिन्न अक्षरों से कितने शब्द बनाये जा सकते हैं जबकि प्रारम्भ में तथा अन्त में स्वर हों ?
How many different words, each containing four different letters can be formed from the letters of the word PERSONAL, if each word is to begin and end with a vowel?

[उत्तर : ${}^3P_2 \times {}^6P_2 = 180$]

111. 'INDIA' शब्द के अक्षरों से कुल कितने शब्द बन सकते हैं ? उनमें से कितने ऐसे हैं जिनमें दो I इकट्ठी न हों और इनमें कितने I से शुरू तथा A पर समाप्त होते हैं ?
How many different words can be formed with the letters of the word 'INDIA'? In how many of them 2 I's are never together and how many of them begin with I and end with A.

[उत्तर : 60:60.- 24 = 36;6]

112. JAIPUR शब्द में प्रयुक्त सभी अक्षरों को सम्मिलित करते हुए विभिन्न शब्दों की संख्या बताइये यदि प्रत्येक शब्द (i) J से प्रारम्भ हो तथा (ii) J से प्रारम्भ न हो ?
Find the number of words that can be formed out of the letters of the word JAIPUR if each word (i) begins with J, and (ii) does not begin with J.
[उत्तर : 5! = 120, (ii) 6! - 5! = 600]

NOTES

113. एक प्रश्न-पत्र में 10 प्रश्न हैं जिनमें से 5 हल करने हैं। पांच प्रश्नों के कितने प्रकार के संचय हो सकते हैं?
A test paper contains 10 questions out of which 5 are to be attempted. How many combinations of 5 questions are there?

[उत्तर : ${}^{10}C_5 = \frac{10!}{5!5!} = 252$]

114. 15 सदस्यीय क्रिकेट टीम में से एक क्रिकेट एकादश कितने प्रकार से छांटी जा सकती है यदि, (i) एक विशिष्ट सदस्य को अवश्य छांटना हो, (ii) उस सदस्य को एकादश में सम्मिलित ही न करना हो।
In how many ways a cricket eleven can be selected from a 15 members cricket team if (i) one particular member must be, (ii) that member never be selected in the cricket eleven.

[उत्तर : (i) $1 \times {}^{14}C_{10}$, (ii) ${}^{14}C_{11}$]

115. 3 लड़कियों एवं 5 लड़कों में से पांच की एक समिति का चुनाव करना है। समिति में 2 लड़कियां एवं 3 लड़के रखे जाते हैं। कितने प्रकार से समिति सम्भव है?
A committee of 5 is to be elected from 3 girls and 5 boys, 2 girls and 3 boys are kept in the committee. In how many ways the committee is possible?

[उत्तर : ${}^{3}C_2 \times {}^{5}C_3 = 30$]

116. एक क्रिकेट क्लब के 18 सदस्य हैं जिनमें से 2 विकेट कीपर, 5 गेंद फेंकने वाले तथा शेष बल्लेबाज हैं। इनमें से 11 खिलाड़ियों की टीम कितने तरीकों से बनाई जा सकती है यदि उसमें एक विकेटकीपर तथा कम-से-कम 3 गेंदबाज सम्मिलित हों।
There are 18 members of a cricket club, in which 2 are wicketkeepers, 5 are bowlers and the remaining are batsmen. In how many ways a team of 11 players can be formed of there is one wicketkeeper and at least 3 bowlers.

[उत्तर : ${}^{2}C_1 \times {}^{5}C_3 \times {}^{11}C_7 + {}^{2}C_1 \times {}^{5}C_4 \times {}^{11}C_6 + {}^{2}C_1 \times {}^{5}C_5 \times {}^{11}C_4$
$= 6{,}600 + 4{,}620 + 924 = 12{,}144$]

117. 10 अच्छे व 5 फ्यूज बल्बों में 7 बल्ब कितने प्रकार से चुने जा सकते हैं जबकि प्रत्येक समूह में कम-से-कम 5 अच्छे हों?
In how many different ways 7 bulbs can be chosen from 10 good and 5 fuse bulbs, if each selection must have at least 5 good bulbs?

[उत्तर : 2520 + 1,050 + 120 = 3,690]

118. 7 पुरुष और 8 महिला सदस्यों में से एक समिति बनानी है। समिति में कुल 6 सदस्य होंगे, जिनमें 2 महिलाएं अनिवार्य रूप से सम्मिलित की जायेंगी। बताइये कितने प्रकार से इस समिति का गठन किया जा सकता है?
A committee is to be formed from 7 men and 8 women. There will 6 members in the committee, when two women are compulsorily included. In how manys this committee can be formed?

[उत्तर : ${}^{8}C_2 \times {}^{7}C_4 + {}^{8}C_3 \times {}^{7}C_3 + {}^{8}C_4 \times {}^{7}C_2 + {}^{8}C_5 \times {}^{7}C_1 + {}^{8}C_6 \times {}^{7}C_0$]

119. 6 भारतीय और 5 जापानियों में सात सदस्यों की कितनी समितियां बनायी जा सकती हैं जिसमें (i) कम-से-कम 2 भारतीय रहें? तथा (ii) कम से कम 3 भारतीय हों?
How many 7 men committee can be formed from 6 Indians and 5 Japanese, if it consists of (i) at least two Indians? (ii) at least three Indians?

[उत्तर : (i) 15 + 100 + 150 + 60 + 5 = 330 (ii) 100 + 150 + 60 + 5 = 315]

120. 6 छात्रों और 5 छात्राओं में से 5 सदस्यों वाली कितनी कमेटियां बनायी जा सकती हैं जिसमें (i) कम-से-कम 2 छात्र रहें; (ii) कम-से-कम 3 छात्राएं रहें ?

How many 5 members committee can be formed from 6 male students and 5 female student if committee consists if (i) at least two male students, (ii) at least three female student?

[उत्तर : (i) 431, (ii) 181]

121. 7 हिन्दू और 4 मुसलमानों के समूह में से 6 सदस्यों की एक समिति कितने प्रकार से बनाई जा सकती है, जबकि उसमें कम-से-कम दो मुसलमान अवश्य हों ?

In how many ways a committee of 6 persons be formed from 7 Hindus and 4 Muslims so that at least 2 Muslims are there in the committee?

[उत्तर : $210 + 140 + 21 = 371$]

122. 6 सज्जनों तथा 4 महिलाओं में से एक 5 की समिति की रचना करनी है। यदि कम-से-कम एक महिला अवश्य ही रहे तो यह कितने प्रकार से हो सकता है ?

From 6 gentlement and 4 ladies, a committee of 5 is to be formed. In how many ways can this be done so as to include at least one lady.

[उत्तर : 246]

123. 6 भारतीय एवं 4 अमेरिकियों में से 5 सदस्यों की एक कमेटी बनायी जाती है। यह कितने प्रकार से बनाई जा सकती है यदि (i) कमेटी में ठीक दो भारतीय हों ? (ii) कम-से-कम दो भारतीय हों ?

Out of 6 Indians and 4 Americans, in how many ways a committee of 5 can be consistituted provided it must have (i) exactly two Indians? (ii) at least two Indians?

[उत्तर : (i) ${}^6C_2 \times {}^4C_3 = 60$
(ii) ${}^6C_2 \times {}^4C_3 + {}^6C_3 \times {}^4C_2 + {}^6C_4 \times {}^4C_1 = 240$]

124. 7 भारतीय एवं 4 पाकिस्तानियों के एक समूह में से 5 व्यक्तियों की एक समिति कितने प्रकार बनायी जा सकती है, अगर उस समिति में,

In how many ways a committee can be formed out of 7 Indians and 4 Pakistanis when the committee contains :

(अ) सभी भारतीय हों।
all Indians.

(ब) कम-से-कम दो पाकिस्तानी ।
at least two Pakistani?

(स) कम-से-कम दो भारतीय तथा दो पाकिस्तानी हों।
at least two Indians and two Pakistanis.

(द) दो से अधिक पाकिस्तानी न हों।
not more than two Pakistanis.

[उत्तर : (a) ${}^7C_5 = 21$, (b) ${}^4C_2 \times {}^7C_2 + {}^4C_3 \times {}^7C_2 + {}^4C_4 \times {}^7C_1 = 301$
(c) ${}^7C_2 \times {}^4C_3 + {}^7C_3 \times {}^4C_2 = 294$,
(d) ${}^4C_0 \times {}^7C_5 + {}^4C_1 \times {}^7C_4 + {}^4C_2 \times {}^7C_3 = 371$]

126. 8 अमेरिकन तथा 5 भारतीयों में से 6 सदस्यों की समिति कितने प्रकार से बनायी जा सकती है जबकि उसमें (अ) 3 अमेरिकन और 3 भारतीय हों, (ब) कम-से-कम 3 भारतीय हों, (स) ठीक 2 अमेरिकन हों।

In how many ways a committee of 6 members can be formed from 8 Americans and 5 Indians when the committee consists of : (a) 3 Americans and 3 Indians, (b) at least 3 Indians, (c) exactly two Americans.

NOTES

[उत्तर : (a) ${}^{8}C_3 \times {}^{5}C_3 = 560$, (b) ${}^{8}C_3 \times {}^{5}C_3 + {}^{8}C_2 \times {}^{5}C_4 + {}^{8}C_1 \times {}^{5}C_5 = 708$, (c) ${}^{8}C_2 \times {}^{5}C_4 = 140$]

127. 6 अध्यापकों एवं 4 छात्रों में 8 सदस्यों की एक समिति बनानी है। यह समिति कितने प्रकार से बनाई जा सकती है? यदि समिति में 4 से कम अध्यापक न हों?

From 6 teachers and four students, a committee of 8 is to be formed. In how many ways can this be done when these should not be less than four teachers in the committee?

[उत्तर : $15+24+6=45$]

128. 15 खिलाड़ियों में 3 अध्यापक हैं। कितने ढंग से 11 खिलाड़ी चुने जा सकते हैं यदि कम-से-कम एक अध्यापक अवश्य खेले?

There are three teachers in 15 players. In how many ways selection of 11 players be made so that in each case at least one teacher should play?

[उत्तर : ${}^{15}C_{11} - {}^{12}C_{11}$]

129. EXAMINATION शब्द के अक्षरों में से एक बार में 4 अक्षरों में से एक बार में 4 अक्षरों को लेने पर कितने शब्द बन सकते हैं?

How many different permutations can be formed from the letters of the word EXAMINATION taken four at time?

[उत्तर : $1,680+756+18=2,454$]

अपनी प्रगति की जाँच करें
Test your Progress

अध्ययाय–2 सैद्धांतिक आवृत्ति वितरण

(THEORETICAL FREQUENCY DISTRIBUTION)

इकाई की रूपरेखा

2.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. सैद्धान्तिक आवृत्ति वितरण के अर्थ को समझ सकेंगे।
2. सैद्धान्तिक आवृत्ति वितरण की उपयोगिता को समझ सकेंगे।
3. सैद्धान्तिक आवृत्ति वितरण के विभिन्न प्रकारों का ज्ञान होगा।

2.1 प्रस्तावना

कुछ घटनाओं के आवृत्ति वितरण गणितीय रीति से ज्ञात किये जा सकते हैं। ऐसे वितरण, प्रायः "सैद्धांतिक वितरण" कहलाते हैं। आवृत्ति-वितरण का अध्ययन दो प्रकार से किया जा सकता है-

(1) **एकत्रित समंकों के आधार पर-** अपकिरण और विषमता द्वारा एकत्रित समंकों का विश्लेषण करके वितरण के विषय में विस्तृत विवेचन किया जाता है। इसका अध्ययन पिछले अध्यायों में किया जा चुका है।

(2) **सम्भावना सिद्धांत के आधार पर-** इस सिद्धांत के आधार पर एक अनुसंधान से पूर्व ही आवृत्ति वितरण के विषय में अनुमान लगाया जा सकता है। इससे यह पूर्व अनुमान हो जाता है कि एकत्रित होने के बाद आवृत्ति वितरण किस प्रकार का होगा।

अब तक हमने ऐसे आवृत्ति-वितरणों का अध्ययन किया है जो सांख्यिकीय अनुसंधानों से उपलब्ध वास्तविक या अवलोकित (observed) समंकों के आधार पर तैयार किये जाते हैं। इन्हें **अवलोकित आवृत्ति-वितरण** (Observed Frequency Distribution) कहते हैं। किसी कक्षा के 100 विद्यार्थियों की ऊँचाई का माप करके और परीक्षा में उनके प्राप्तांक मालूम करके निम्नलिखित दो अवलोकित आवृत्ति वितरण बनाये जा सकते हैं-

NOTES

अवलोकित आवृत्ति वितरण के उदाहरण

ऊंचाई (सेमी)	विद्यार्थियों की संख्या (आवृत्ति)	प्राप्तांक	आवृत्ति
150–155	4	1–20	10
155–160	12	21–40	16
160–165	56	41–60	54
165–170	20	61–80	14
170–175	6	81–100	6
175–180	2		
	100		100

2.2 सैद्धांतिक आवृत्ति वितरण का अर्थ

कुछ क्षेत्रों में किसी सुनिश्चित एवं पूर्व निर्धारित कल्पना या मान्यता से प्रारम्भ करके, गणितीय आधार पर प्रत्याशित आवृत्तियों का अनुमान लगाया जा सकता है। उदाहरण के लिए यदि 3 के सिक्के 64 बार उछाले जायें, चित्त आने को सफलता माने, तो 'द्विपद सम्भावना' सिद्धांत के आधार पर निम्नलिखित प्रत्याशित आवृत्तियों की कल्पना की जाती है-

सफलता की संख्या	संभावना	प्रत्याशित आवृत्ति
0	$\frac{1}{8}$	8
1	$\frac{3}{8}$	24
2	$\frac{3}{8}$	24
3	$\frac{1}{8}$	8
योग	1	64

सैद्धांतिक आवृत्ति वितरण की परिभाषा इस प्रकार की जा सकती है-

ऐसे आवृत्ति वितरण जिन्हें वास्तविक अवलोकनों या प्रयोगों द्वारा प्राप्त न करके कुछ सुनिश्चित पूर्वकल्पनाओं के आधार पर गणितीय रूप में अनुमानित किया जाता है, सैद्धांतिक आवृत्ति वितरण कहलाते हैं। इन्हें प्रत्याशित या आदर्श आवृत्ति वितरण भी कहते हैं।

"Distributions which are not obtained by actual observations or experiments, but are deduced mathematically under certain hypotheses or assumptions, are called Theoretical Frequency Distribution."

2.3 सैद्धांतिक आवृत्ति वितरण की उपयोगिता

सैद्धांतिक आवृत्ति वितरण आधुनिक सांख्यिकी के आधार पर माने गये हैं। इनकी उपयोगिता निम्नलिखित बातों से स्पष्ट है-

(1) इनमें से हमें यह पता चलता है कि निश्चित मान्यताओं के अन्तर्गत समंकों की प्रवृत्ति क्या होगी?

(2) इनके द्वारा प्राप्त प्रत्याशित समंक विवेकपूर्ण निर्णय लेने का आधार प्रस्तुत करते हैं।

(3) प्रत्याशित समंकों की सहायता से भावी पूर्वानुमान लगाये जा सकते हैं।

(4) ये वितरण अवलोकित वितरण के स्थानापन्न होते हैं, विशेषतः उन परिस्थितियों में, जबकि वास्तविक अनुसंधानों में व्यय बहुत अधिक होता हो।

(5) इनसे प्राप्त प्रत्याशित आवृत्तियों की वास्तविक अनुसंधान द्वारा प्राप्त अवलोकित आवृत्तियों से तुलना की जा सकती है और यह निष्कर्ष निकाला जा सकता है कि दोनों का अन्तर न्यादर्श के उच्चावचनों के कारण उत्पन्न हुआ है या किसी अन्य कारण द्वारा। जैसे, एक सिक्के को 200 बार उछालने पर चित (Head) आने की प्रत्याशा (Expectation) 100 है। यदि वास्तव में सिक्का उछालने पर 90 बार वह चित आवे और 110 बार पट तो अर्थपूर्णता जाँच की सहायता से यह जाँच की जा सकती है कि यह अन्तर (90-100) दैव के कारण है या सिक्के की सुडौलता में कमी होने के कारण।

(6) इनसे वास्तविक अनुसंधान में खर्च होने वाले धन, श्रम व समय का पूर्वाभास हो जाता है। इससे यह निर्णय किया जा सकता है कि सम्बन्धित अनुसंधान होगा या नहीं।

(7) विभिन्न प्रकार की प्रत्याशित आवृत्ति वितरणों का ज्ञान दैनिक जीवन की कई समस्याओं के समाधान में सहायक होता है। उदाहरणार्थ, व्यवसाय में सफलता प्राप्त करने के लिए एक रेडीमेड कपड़ों के निर्माता को यह जानकारी होनी चाहिए कि किस आकार के कॉलर की कितनी प्रत्याशित या अनुमानित माँग होगी? यह अनुमान प्रसामान्य-वितरण के आधार पर लगाया जा सकता है। गुण-नियंत्रण विशेषज्ञ, प्वॉयसन-वितरण के आधार पर यह निश्चित करते हैं कि अमुक उत्पादन-विधि सुचारू रूप से चल रही है या नहीं। विपणन-विशेषज्ञ, काई-वर्ग वितरण का उपयोग कर यह देखते हैं कि वस्तु के स्वरूप में थोड़ा परिवर्तन करने से उपभोक्ताओं की प्रतिक्रियाओं में क्या महत्वपूर्ण परिवर्तन होते हैं?

इस प्रकार विभिन्न क्षेत्रों में सैद्धांतिक आवृत्ति वितरणों का अध्ययन विशेष रूप से उपयोगी होता है। मेरिल एवं फॉक्स के अनुसार, ये सांख्यिकीय सिद्धांत में कई महत्वपूर्ण भूमिकाओं का निर्वाह करते हैं।

सैद्धांतिक आवृत्ति वितरण के प्रकार

सैद्धांतिक दृष्टि से आवृत्ति वितरण कई प्रकार के हो सकते हैं, किन्तु उनमें से मुख्य तीन निम्न हैं, जिनका सांख्यिकी विश्लेषण में काफी महत्व है-

(1) द्विपद वितरण (Binomial Distribution)

(2) सामान्य वितरण (Normal Distribution)

(3) प्वॉयसन वितरण (Poisson Distribution)

2.4 द्विपद वितरण (Binomial Distribution)

इस वितरण के निर्माण की खोज श्री जेम्स बरनौली (स्विस गणितज्ञ) द्वारा 1654-1705 में की गई। यद्यपि इसका प्रकाशन उनकी मृत्यु के आठ वर्ष बाद हुआ। इस वितरण का आधार बरनौली प्रमेय (Bernoulli Theorem) ही है। जैसा कि इस वितरण के नाम से ही संकेत मिलता है, यह दो वर्गों में बंटा हुआ होता है। एक वर्ग में वे सब वस्तुएं शामिल की जाती हैं, जिनमें एक निश्चित विशेष गुण होता है और दूसरे वर्ग में वे वस्तुएं शामिल होती हैं जिनमें वह विशेष गुण नहीं होता। इस प्रमेय की सहायता से सफलता या विफलता की सम्भावनाओं को ज्ञात कर लिया जाता है। उनमें कुल संख्याओं से गुणा करके पदों की आवृत्तियाँ मालूम हो जाती हैं-

अपनी प्रगति की जाँच करें
Test your Progress

जैसे, यदि दो सिक्के 'A' और 'B' एक साथ उछाले जायें तो सम्भाव्य परिणाम ये होंगे-

A	B
H	H
H	T
T	H
T	T

Here, 'H' stands for Head and T for Tail.

इस प्रकार चार में से एक साथ दो सिरों (heads) के आने की सम्भावना 1 अर्थात् $\frac{1}{4}$ है। इसे सम्भावित प्रमेय द्वारा निम्न प्रकार व्यक्त किया जा सकता है-

NOTES

प्रथम उछाल में सिर के आने की सम्भावना $= \frac{1}{2}$

द्वितीय उछाल में सिर के आने की सम्भावना $= \frac{1}{2}$

अतः एक साथ (simultaneously) सिर के आने की सम्भावना होगी $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ होगी। यही बात दुम (Tail) के विषय में लागू होगी।

यदि उपरोक्त उदाहरण में सफलता की सम्भावना (H) के लिए 'p' और विफलता की सम्भावना (T) के लिए 'q' संकेत का प्रयोग किया जाये तो चार परिणामों को निम्न प्रकार व्यक्त किया जायेगा-

	HH	HT	TH	TT
	pp	pq	qp	qq
or	p^2		$2pq$	q^2

$p^2 + 2pq + q^2$ का विस्तार $(p + q)^2$ से होता है। अतः दो स्वतंत्र घटनाओं की संयुक्त सम्भावना को निम्न द्विपद विस्तार से प्रकट किया जा सकता है-

$(p + q)^2 a$

The probability of coming Head or p $= \frac{1}{2}$

The probability of coming Tail or q $= \frac{1}{2}$

Then $(p = q)^2 = (\frac{1}{2} + \frac{1}{2})^2 = \frac{1}{4} + \frac{1}{2} + \frac{1}{4}$

यही ऊपर व्यक्त किया गया है। यह विस्तार आगे भी निम्न सूत्र की सहायता से किया जा सकता है-

$(p + q)^n$

यदि हम प्रत्येक परिस्थिति में कुल आवृत्ति ज्ञात करना चाहते हैं तो वह इस प्रकार ज्ञात होगी-

N (p + q)

where 'N' = Number of trails

'n' = Number of independent events.

Illustration 1. Find out the probable frequencies if there are 100 trials and 2 independent events.

Solution - The formula applicable in this case is

$N(p + q)^n = 100\,(p + q)^2$

$= 100\,(p^2 + 2pq + q^2)$, if $p = \frac{1}{2}, q = \frac{1}{2}$

or $= 100\left(\frac{1}{4}\right) + 100\left(\frac{1}{2}\right) + 100\left(\frac{1}{4}\right)$

or $= 25 + 50 + 25$

= 25 for two successes.

= 50 for one success.

= 25 for no success.

Ilustration 2. (Comparison of Actual and Theoretical frequencies.)

Six dice were thrown 74 times. Each 4, 5, or 6 spot appearing was considered to be success and each 1, 2 or 3 spot a failure. The results were :

NOTES

No. of Successes	0	1	2	3	4	5	6
Actual Frequencies	0	5	13	22	17	7	0

Find out the Expected frequencies from the above, and compare them.

Solution - Expected frequencies,

$$N(p+q)6 = 64\,(p^6 + 6p^5q + 15p^4q^2 + 20p^3q^3 + 15p^2q^4 + 6pq^5 + p^6)$$

$$= 64\left(\frac{1}{64}\right) + 64\left(\frac{6}{64}\right) 64 + \left(\frac{15}{64}\right) 64 + \left(\frac{20}{64}\right) 64 \left(\frac{15}{64}\right)$$

$$+ 64\left(\frac{6}{64}\right) + 64\left(\frac{1}{64}\right)$$

$$= 1 + 6 + 15 + 20 + 15 + 6 + 1 = 64$$

Hence the Expected frequencies for different successes are as under-

No. of Successes	0	1	2	3	4	5	6
Expected frequencies	1	6	15	20	15	6	1

For comparison of Actual and Expected frequencies we will calculate the Mean and Standard deviation of the two as under-

m	Actual Frequencies					Expected Frequencies				
	F	mf	dx(3)	fdx	fd^2x	F	mf	dx(3)	fdx	fd^2x
0	0	0	−3	0	0	1	0	−3	−3	2
1	5	5	−2	−10	20	6	6	−2	−12	24
2	13	26	−1	−13	13	15	30	−1	−15	15
3	22	66	0	0	0	20	60	0	0	0
4	17	68	+1	+17	17	15	60	+1	+15	15
5	7	35	+2	+14	28	6	30	+2	+12	24
6	0	0	+3	0	0	1	6	+3	+3	9
	64	200		+8	78	64	192		0	96

$$a = \frac{\Sigma mf}{n} = \frac{200}{64} = 3.125$$

$$\text{S.D.} = \sqrt{\left[\frac{\Sigma fd^2x}{n} - \left(\frac{\Sigma fdx}{n}\right)^2\right]}$$

$$= \sqrt{\left[\frac{78}{64} - \left(\frac{8}{64}\right)^2\right]}$$

$$= 1.1$$

$$a = \frac{\Sigma mf}{n} = \frac{196}{64} = 3$$

$$\text{S.D.} = \sqrt{\left[\frac{\Sigma fd^2x}{n} - \left(\frac{\Sigma fdx}{n}\right)^2\right]}$$

$$= \sqrt{\left[\frac{96}{64} - \left(\frac{0}{64}\right)^2\right]}$$

$$= 1.22$$

द्विपद वितरण का सामान्य रूप - उपरोक्त उदाहरणों से यह स्पष्ट है कि यदि 'n' घटनाएं हों तो द्विपद वितरण के विस्तार का स्वरूप निम्न प्रकार होगा-

$$(p+q)^n = p^n + np^{n-1}q^1 + \frac{n(n-1)}{2\times 1}p^{n-1}q^2 + \frac{n\,(n-)\,(n-2)}{3\times 2\times 1}$$

$$p^{n-3}q^3 + \frac{n\,(n-1)\,(n-2)\,(n-3)}{4\times 3\times 2\times 1}p^{n-4}q^4 + \ldots q^n$$

यदि 'n' = 5 हो तो $(p+q)^n = (p+q)^5$

$$\therefore\ (p+q)^5 = p^5 + 5y^{5-1}q^1 + \frac{5\,(5-1)}{2}p^{5-2}q^2 + \frac{5\,(5-1)\,(5-2)}{3\times 2}p^{5-3}q^3 +$$

$$\frac{5\,(5-1)\,(5-2)\,(5-3)}{4\times 3\times 2}p^{5-4}q^4 +$$

NOTES

$$\frac{5(5-1)(5-2)(5-3)(5-4)}{5\times4\times3\times2}p^{5-5}q^{5}$$

$$= p^5 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5$$

पास्कल त्रिभुज (Pascal-Triangle)- द्विपद वितरण के विभिन्न पदों को संख्यात्मक गुणांक पास्कल के त्रिभुज से प्रत्यक्ष रूप में देखे जा सकते हैं।

पास्कल त्रिभुज

$(p+q)^n$ के विस्तार के गुणांक

घात Power	द्विपद गुणांक Binomial Coefficient	योग 2^n
1	1 1	2
2	1 2 1	4
3	1 3 3 1	8
4	1 4 6 4 1	16
5	1 5 10 10 5 1	32
6	1 6 15 20 15 6 1	64
7	1 7 21 35 35 21 7 1	128
8	1 8 28 56 70 56 28 8 1	256
9	1 9 36 84 126 84 36 9 1	512
10	1 10 45 120 210 252 210 45 10 1	1024

पास्कल त्रिभुज में प्रत्येक गुणांक उससे पिछली पंक्ति के दोनों ओर के दो गुणांकों का योग है- 6 = 3 + 3; 28 = 7 + 21; 126 = 70 + 56; 10 = 9+1 आदि।

यह नीचे कुछ उदाहरणों द्वारा स्पष्ट हो जावेगा।

Illustration 3. Assuming that half of the population is vegetarian so that choice of an individual being a vegetarian is 1/2. Assuming that 100 investigators can take sample of 10 individuals each to see whether they are vegetarians. how many investigators would you expect to report that three people or less were vegetarians.

Solution - Probability of a person being vegetarian,

$$(P) = \frac{1}{2}$$

$$q = 1 - p = 1 - \frac{1}{2} = \frac{1}{2}$$

Out of 10, 3 or less vegetarians

Veg.	Non-veg.	Probability
3	7	$p = \frac{10!}{3!(10-3)!}(.5)^3(.5)^7 = .1171875$
2	8	$p = \frac{10!}{2!(10-8)!}(.5)^2(.5)^8 = .0439453125$
1	9	$p = \frac{10!}{1!(10-1)!}(.5)^1(.5)^9 = .009365625$
0	10	$p = \frac{10!}{0!(10-0)!}(.5)^0(.5)^{10} = \frac{.0009765625}{.1718750000}$

Regd. No. of investigators

$$= 100 \times 0.171875 = 17.1875 = 17$$

Illustration 4. Four coins are tossed simultaneously. What is the probability of getting (i) 2 heads and 2 tails, (ii) at least 2 heads, and (iii) at least one head.

Solution - Here the random experiment consists in tossing 4 coins and observing the number of heads. Let us call the occurance of heads as 'success', then

$$p = P\text{ (head with single coin)} = \frac{1}{2}; q = \frac{1}{2}; n = 4$$

Since the value of 'p' is constant fro each coin and the trails are independent, using formula for binomial distribution, the probability of x successes is-

$$p(x) = {}^4C_x \left(\frac{1}{2}\right)^{4-x} = {}^4C_x \left(\frac{1}{2}\right)^4$$

(i) Putting $x = 2$, we get

$$P\text{ (2 heads and 2 tails)} = p(2) = {}^4C_2 \left(\frac{1}{2}\right)^4 = \frac{3}{8}$$

(ii) Since 'at least 2 heads' implies '2 or 3 or 4 heads', the probability of at least 2 successes is given by the sum of probabilities p (2) + p (3) + p (4).

$$\therefore \quad P\text{ (at least 2 heads)} = {}^4C_x \left(\frac{1}{2}\right)^4 + {}^4C_3 \left(\frac{1}{3}\right)^4 + {}^4C_x \left(\frac{1}{2}\right)^4$$

$$= \frac{3}{8} + \frac{1}{4} + \frac{1}{18} = \frac{11}{18}$$

$$\text{(iii)} \quad P\text{ (at least 2 heads)} = 1 - p\text{ (no head)} = 1 \quad {}^4C_0 \left(\frac{1}{2}\right)^4 = \frac{15}{18}$$

द्विपद-वितरण के अचर-मूल्य (Constants of the Binomial Distribution)- द्विपद-वितरण के माध्य और प्रमाप विचलन ज्ञात करने के लिए निम्नलिखित सूत्रों का उपयोग किया जाता है-

(i) समान्तर माध्य - $\bar{x} = np$

(ii) प्रमाप विचलन - $\sigma = \sqrt{npq}$

Illustration 5. (i) एक फैक्टरी में औसत रूप से 40% स्क्रू दोषपूर्ण बनाये जाते हैं। 10 स्क्रू में से दोषपूर्ण स्क्रुओं के सैद्धांतिक वितरण के लिये समान्तर माध्य एवं प्रमाप विचलन मालूम कीजिए। परिघातों पर आधारित विषमता एवं पृथुशीर्षतत्वक गुणांक भी ज्ञात कीजिये।

(ii) ऐसा सम्पूर्ण द्विपद वितरण लिखिए जिसका समान्तर माध्य 3 है और प्रसरण (Variance) 2 है।

(iii) किसी द्विपद वितरण के लिए एक छात्र ने समान्तर माध्य का माप 5 और प्रसरण (Variance) का मान 9 परिकलित किया है। क्या उसका परिकलन सही है ?

Solution - दोषपूर्ण स्क्रू का अनुपात $(p) = \frac{40}{100} = \frac{2}{5}$

दोषरहित स्क्रू की सम्भावना $(q) = 1 - p = 1 - \frac{2}{5} = \frac{3}{5}$;

$$n = 10$$

द्विपद वितरण में $\bar{x} = np$ और $\sigma = \sqrt{npq}$

$$\therefore \quad \bar{x} = 10 \times \frac{2}{5} = 4, \sigma = \sqrt{10 \times \frac{2}{5} \times \frac{3}{5}} = \sqrt{2.4} = 1.55$$

परिघातों पर आधारित विषमता गुणांक

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}} = \frac{q - p}{\sqrt{npq}} = \frac{.6 - .4}{\sqrt{2.4}} = 1.3$$

NOTES

पृथु शीर्षस्थ का माप-

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{1-6pq}{npq} = 3 = \frac{1-6\times.4\times.6}{2.4} = 3 - .183 = 2.817$$

ii. एक द्विपद वितरण में समान्तर माध्य $\bar{x} = 3$,

प्रसरण $= 2, \sigma = \sqrt{2} = 1.414$

$$\bar{x} = np = 3; \sigma^2 = npq = 2; \frac{npq}{np} = 1 = \frac{2}{3}$$

$\therefore \quad p = 1 - \frac{2}{3} = \frac{1}{3} \quad \Rightarrow \quad npq = n \times \frac{1}{3} \times \frac{2}{3} = 2$

या $\quad \frac{2}{9}n = 2 \quad \therefore \quad n = 9$

$n = 9; p = \frac{1}{3}, q = \frac{2}{3}$, द्विपद वितरण $= (q+p)^n = \left(\frac{2}{3}+\frac{1}{3}\right)^9$

$$\left(\frac{2}{3}+\frac{1}{3}\right)^9 = \left(\frac{2}{3}\right)^9 + 9\left(\frac{2}{3}\right)^8\left(\frac{1}{3}\right)^1 + 36\left(\frac{2}{3}\right)^7\left(\frac{1}{3}\right)^2 + 84\left(\frac{2}{3}\right)^6\left(\frac{1}{3}\right)^3$$
$$+ 126\left(\frac{2}{3}\right)^5\left(\frac{1}{3}\right)^4 + 126\left(\frac{2}{3}\right)^4\left(\frac{1}{3}\right)^5 + 84\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^6$$
$$+ 36\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)^7 + 9\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^8 + \left(\frac{1}{3}\right)^9$$

$$= \frac{1}{19683}[(512\times1) + (9\times256) + (36\times128)$$
$$+ (84\times64) + (126\times32) + (126\times16) + (84\times8)$$
$$+ (36\times4) + (9\times2) + 1]$$

सफलता की संख्या (x)	सम्भावना p (x)
0	$\frac{512}{19683}$
1	$\frac{2304}{19683}$
2	$\frac{4608}{19683}$
3	$\frac{5376}{19683}$
4	$\frac{4032}{19683}$
5	$\frac{2016}{19683}$
6	$\frac{672}{19683}$
7	$\frac{144}{19683}$
8	$\frac{18}{19683}$
9	$\frac{1}{19683}$

(iii) समान्तर माध्य $\bar{x} = 5$, प्रसरण $\sigma^2 = 9$;

$$\bar{x} = np = 5\,;\, npq = 9, q = \frac{npq}{np} = \frac{9}{5}$$

किन्तु q या p, 1 से अधिक कदापि नहीं हो सकता, वरन् $P+q=1$ होता है। यहाँ पर $q=1\frac{4}{5}$ है जो सम्भव है अतः गणन क्रिया में उसका परिकलन सही नहीं है।

वास्तविक और प्रत्याशित आवृत्तियों की तुलना
(Comparison of actual and expected frequencies)

यदि 5 सिक्के 250 बार उछाले जायें और 0, 1, 2, 3, 4, 5 सिरों (heads) की संख्या ज्ञात की जावे तो वास्तविक आवृत्तियाँ उपलब्ध हो जावेंगी। द्विपद वितरण के आधार पर सम्भावना निकालकर प्रत्याशित आवृत्तियाँ भी ज्ञात की जा सकती हैं। यदि सिक्के सुडौल (दोष रहित) हों और जाँच (trials) की संख्या अत्यधिक हो तो वास्तविक तथा प्रत्याशित आवृत्तियों में अन्तर बहुत कम हो जाता है।

वास्तविक तथा प्रत्याशित के अन्तर का माप निम्नलिखित दो रीतियों द्वारा किया जा सकता है-

(i) विन्दुरेखीय रीति।

(ii) काई-वर्ग जाँच (X^2 – test) रीति।

प्रथम रीति में, दोनों ही आवृत्तियाँ रेखाचित्र पर अंकित की जाती हैं। यदि दोनों वक्र एक-दूसरे के बहुत निकट हों तो अन्तर निरर्थक माना जाता है और यह कहा जाता है कि वक्र-आसंजन उत्कृष्ट है (The fit is good)। यदि दोनों की दूरी अधिक होती है तो वास्तविक तथा प्रत्याशित आवृत्तियों में अधिक अन्तर होता है तो वक्र आसंजन उत्कृष्ट नहीं होता है। (The fit is not good)।

द्वितीय रीति में, वास्तविक (f_a) तथा प्रत्याशित (f_e) आवृत्तियों के अन्तर के आधार पर काई-वर्ग की गणना कर ली जाती है-

$$X^2 \Sigma \left\{\frac{(f_a - f_e)^2}{f_e}\right\} \quad \text{या } \Sigma \left\{\frac{(O - E)^2}{E}\right\}$$

काई-वर्ग निकालने के बाद, स्वातंत्र्यांश या स्वातंत्र्य संख्या (Degree of Freedom) ज्ञात कर ली जाती है। द्विपद-वितरण में सह सफलताओं की संख्या से 1 से कम होती है। इसके बाद सम्बन्धित स्वातंत्र्य संख्या के 5% सार्थकता स्तर पर x^2 - तालिका में से काई-वर्ग का मूल्य देख लिया जाता है। यदि परिकलित मूल्य इस सारणी मूल्य से अधिक होता है तो आसंजन उत्कृष्ट नहीं होता और वास्तविक तथा प्रत्याशित आवृत्ति में अन्तर सार्थक (Significant) है। इसके विपरीत, परिकलित मूल्य सारणी-मूल्य से कम होने पर आसंजन उत्तम माना जाता है अर्थात् अन्तर अर्थहीन (Not Significant) होता है।

Illustration 6. छः भुजा वाले 12 पाँसे 4096 बार फेंके गये और 4, 5 तथा 6 को सफलता माना गया। सफलता की आवृत्तियाँ निम्नांकित हैं-

सफलता	0	1	2	3	4	5	6	7	8	9	10	11	12
आवृत्ति	-	7	60	198	430	731	948	847	536	257	72	11	-

वास्तविक माध्य और प्रमाप विचलन ज्ञात कीजिए और प्रत्याशित वितरण के लिए भी इन मापों की गणना कीजिए। अपनी गणना को स्वच्छ सारणी रूप में दर्शाइये।

Solution - एक पाँसे में 6 परिणाम होते हैं। 4, 5 या 6 और 3 का फेंका जाना सफलता माना जावेगा। अतः $p = q = \frac{1}{2}$

पाँसों की संख्या, $n = 12$, परीक्षणों की संख्या, $N = 4096$

द्विपद - वितरण-

$$n\,(q+p)^n = 4096\,(\frac{1}{2}+\frac{1}{2})^{12}$$

NOTES

इसका विस्तार करने से- 0, 1, 2,...., 12 सफलताओं की प्रत्याशित आवृत्ति मालूम हो जावेगी-

$$= 4096 \left\{1 \left(\tfrac{1}{2}\right)^{12}\right\} + \left\{12 \left(\tfrac{1}{2}\right)^{11} \left(\tfrac{1}{2}\right)^{1}\right\} + \left\{66 \left(\tfrac{1}{2}\right)^{10} \left(\tfrac{1}{2}\right)^{2}\right\}$$

$$+ \left\{220 \left(\tfrac{1}{2}\right)^{9} \left(\tfrac{1}{2}\right)^{3}\right\} + \left\{495 \left(\tfrac{1}{2}\right)^{8} \left(\tfrac{1}{2}\right)^{4}\right\} + \left\{792 \left(\tfrac{1}{2}\right)^{7} \left(\tfrac{1}{2}\right)^{6}\right\}$$

$$+ \left\{924 \left(\tfrac{1}{2}\right)^{6} \left(\tfrac{1}{2}\right)^{6}\right\} + \left\{792 \left(\tfrac{1}{2}\right)^{5} \left(\tfrac{1}{2}\right)^{7}\right\} + \left\{495 \left(\tfrac{1}{2}\right)^{4} \left(\tfrac{1}{2}\right)^{8}\right\} +$$

$$\left\{220 \left(\tfrac{1}{2}\right)^{3} \left(\tfrac{1}{2}\right)^{9}\right\} + \left\{66 \left(\tfrac{1}{2}\right)^{2} \left(\tfrac{1}{2}\right)^{2} \left(\tfrac{1}{2}\right)^{10}\right\} + \left\{12 \left(\tfrac{1}{2}\right)^{1} \left(\tfrac{1}{2}\right)^{11}\right\} + \left\{1 \left(\tfrac{1}{2}\right)^{12}\right\}$$

$$= \frac{4096 \times 1}{4096} (1 + 12 + 66 + 220 + 495 + 792 + 924 +$$

$$792 + 495 + 220 + 66 + 12 + 1)$$

सफलता	0	1	2	3	4	5	6	7	8	9	10	11	12
आवृत्ति	1	12	66	220	495	792	924	792	495	220	66	12	1

समान्तर माध्य एवं प्रमाप विचलन की गणना

वास्तविक (Actual)					प्रत्याशित (Expected)			
सफलता	आवृत्ति f_n	dx (6)	fdx	fd^2x	f_e (6)	dx	fdx	fd^2x
0	0	–6	0	0	1	–6	–6	36
1	7	–5	–35	175	12	–5	–60	300
2	60	–4	–240	960	66	–4	–264	1056
3	198	–3	–594	1782	220	–3	–660	1980
4	430	–2	–860	1720	495	–2	–990	1980
5	731	–1	–731	731	792	–1	–792	792
6	948	0	0	0	924	0	0	0
7	847	+1	+847	847	792	+1	+792	792
8	536	+2	+1072	2144	495	+2	+990	1980
9	257	+3	+771	2313	220	+3	+660	1980
10	71	+4	+284	1136	66	+4	+264	1056
11	11	+5	+55	275	12	+5	+60	300
12	0	+6	0	0	1	+6	+6	36
योग	4,096	0	569	12,083	4,096	0	0	12,288

वास्तविक (Actual) f_a	प्रत्याशित (Expected) f_e
$\bar{x} = A + \frac{\Sigma fdx}{n}$	$\bar{x} = A + \frac{\Sigma fdx}{n}$
$= 6 + \frac{569}{4096} = 6.14$	$= 6 + \frac{0}{4096} = 6$
S.D. $= \sqrt{\left[\frac{\Sigma fd^2x}{n} - \left(\frac{\Sigma fdx}{n}\right)^2\right]}$	$= \sqrt{\left[\frac{\Sigma fd^2x}{n} - \left(\frac{\Sigma fdx}{n}\right)^2\right]}$
$= \sqrt{\left[\frac{12083}{4096} - \left(\frac{596}{4096}\right)^2\right]}$	$= \sqrt{\left[\frac{12288}{4096} - \left(\frac{0}{4096}\right)^2\right]}$
$= \sqrt{2.95 - (.14)^2}$	$= \sqrt{3 - 0}$
$= 1.71$	$= 1.732$

प्रत्याशित वितरण के अचर मूल्य निम्नलिखित सूत्रों द्वारा भी ज्ञात किये जा सकते हैं-

$$\bar{x} = np = 12 \times \frac{1}{2} = 6\ ; \sigma = \sqrt{npq} = \sqrt{12 \times \frac{1}{2} \times \frac{1}{2}}$$

$$= \sqrt{13} = 1.732$$

वास्तविक एवं प्रत्याशित आवृत्ति विवरणों में विशेष अन्तर नहीं है। यदि इन दोनों के वक्र खींचे जावें तो वे लगभग एक दूसरे के समरूप ही होंगे। अतः वक्र-आसंजन उत्कृष्ट है।

Illustration 7. 192 परिवारों में जिनके लिये (albinos) बच्चे के उत्पन्न होने की सम्भावना 0.25 समझी जाती है, प्रथम तीन बच्चों में सूरजमुखी बच्चों का वितरण इस प्रकार था-

सूरजमुखी बच्चों की संख्या;	0	1	2	3
परिवारों की संख्या :	77	90	20	5

यह मानते हुए कि द्विपद नियम लागू होता है, सैद्धांतिक आवृत्तियों को मालूम कीजिये तथा X^2 – परीक्षण का प्रयोग करके आसंजन-सौष्ठव (Goodness of fit) की जाँच कीजिये।

Solution - सूरजमुखी बच्चे के पैदा होने की सम्भावना-

$$p = 0.25 = \frac{1}{4}\ ; q = 1 - p = \frac{3}{4}\ ; n = 3\ ; N = 192$$

0, 1, 2, 3 सूरजमुखी के जन्म की प्रत्याशित आवृत्तियाँ मालूम करने के लिये निम्नलिखित द्विपद-वितरण के सभी पद निकाले जावेंगे-

$$N(q+p)^n = 192\left(\frac{3}{4} + \frac{1}{4}\right)^3$$

$$= 192\left[\left\{\left(\frac{3}{4}\right)^3\right\} + \left\{3\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)\right\} + \left\{3\left(\frac{3}{4}\right)^1\left(\frac{1}{4}\right)^2\right\} + \left\{\left(\frac{1}{4}\right)^3\right\}\right]$$

$$= \frac{192 \times 1}{64}(27 + 27 + 9 + 1) = 81 + 81 + 27 + 3$$

आसंजन-सौष्ठव (Goodness of fit) की जाँच-

x	f_a	f_e	$f_a - f_o$	$(f_a - f_e)^2$	$\frac{(f_a - f_e)^2}{f_e}$
0	77	81	–4	16	$\frac{16}{81} = 0.198$
1	90	81	+9	81	$\frac{81}{81} = 1.000$
2	20	27	–7	49	$\frac{49}{27} = 1.815$
3	5	3	+2	4	$\frac{4}{3} = 1.333$
N = 192		192			$x^2 = 4.346$

स्वतंत्र्य संख्या = 4 – 1 = 3

3 d.f. के लिये 5% सार्थकता स्तर पर काई वर्ग का सारणी - मूल्य 7.815 है। ज्ञात किया गया मूल्य 4.346 सारणी मूल्य से कम है। अतः वास्तविक एवं प्रत्याशित आवृत्तियों का अन्तर सार्थक नहीं है, अर्थात् वक्र आसंजन उत्तम है।

Illustration 8. पाँच सिक्कों का एक समुच्चय (set) 3200 बार उछाला जाता है और प्रत्येक बार आने वाले सिरों (heads) की संख्या लिख ली जाती है। इस प्रकार प्राप्त परिणाम निम्नांकित है-

सिरों की संख्या :	0	1	2	3	4	5
आवृत्ति :	80	570	1100	900	500	50

इस परिकल्पना की जाँच कीजिये कि सिक्के सुडौल हैं। X^2 जाँच का प्रयोग कीजिये।

Solution- सिक्के सुडौल होने की स्थिति में सिर (head) व दुम (tail) की सम्भावनायें $p = q = \frac{1}{2}$ होंगी।

यहाँ पर सिक्कों की संख्या, $n = 5$; जाँच (परीक्षण) की संख्या, N = 3,200 है।

अतः सैद्धांतिक आवृत्ति वितरण-

$N(q+p)^n = 3200\left(\frac{1}{2}+\frac{1}{2}\right)^5$ के विस्तार के द्वारा ज्ञात किया जायेगा।

$$3200\left(\frac{1}{2}+\frac{1}{2}\right)^5 = 3200\left(\frac{1}{2}\right)^5 [1 + 5 + 10 + 10 + 5 + 1]$$

$$= 3200 \times \frac{1}{82}(1 + 5 + 10 + 10 + 5 + 1)$$

$$= 100 + 500 + 1000 + 1000 + 500 + 100$$

x	f_a	f_e	$f_a - f_e$	$(f_a - f_e)^2$	$\frac{(f_a - f_e)^2}{f_e}$
0	80	100	–20	400	4.00
1	570	500	+70	4900	9.80
2	1100	1000	+100	10000	10.00
3	900	1000	–100	10000	10.00
4	500	500	0	0	0.00
5	50	100	–50	2500	25.00
N =	3,200	3,200	0	27,800	$x^2 = 58.80$

∴ स्वातंत्र्यांश (d.f.) 5 के लिये 5% स्तर पर x^2 का सारणी मूल्य 11.07 है। परिकलित मूल्य इससे बहुत अधिक है, अतः वक्र आसंजन उत्तम नहीं है, अर्थात् सिक्का सुडौल नहीं है।

द्विपद वितरण की विशेषतायें (Characteristics of Binomial distribution)- उपरोक्त उदाहरणों के आधार पर द्विपद-वितरण की निम्नलिखित विशेषतायें होती हैं-

(1) **सैद्धांतिक वितरण** (Theoretical Distribution)- द्विपद-वितरण एक सैद्धान्तिक आवृत्ति वितरण है जो बरनौली-प्रमेय पर आधारित है। इस सम्भावना वितरण को कुल संख्या (N) से गुणा करके प्रत्याशित आवृत्तियाँ मालूम की जाती हैं।

(2) **खण्डित-वितरण-** यह एक खण्डित आवृत्ति वितरण है जिसमें सफलताओं की संख्या 0,1,2,3...n होती है और उनकी आवृत्तियाँ द्विपद-विस्तार द्वारा निकाली जाती हैं।

(3) **आवृत्ति-बहुभुज-** द्विपद-वितरण को रेखाचित्र पर आवृत्ति-बहुभुज (Frequency Polygon) के रूप में प्रस्तुत किया जा सकता है।

(4) **स्वरूप-** इसका स्वरूप 'p' और 'q' के माप और घातांक 'n' के मूल्य पर आधारित है।

(5) **अचर-मूल्य** (Constant) -- इसमें अचर-मूल्य विभिन्न सूत्रों द्वारा उपलब्ध किये जाते हैं, जैसे- समान्तर माध्य $(\bar{x}) \equiv np$; प्रमाप विचरण $(\sigma) = \sqrt{npq_1}$ प्रथम परिघात $(\mu_1) = 0$; द्वितीय परिघात $(N_2) = npq$; तृतीय परिघात $(\mu_3) = npq\,(q - p)$।

(6) **'p' और 'q' का क्रम** - n, p और q ज्ञात होने पर पूर्ण द्विपद वितरण लिखा जा सकता है। यदि सफलता की संख्या आरोही क्रम में- 0, 1, 2, 3....n लिखनी है तो द्विपद सम्भावना वितरण (q+p)n का विस्तार होगा। इसके विपरीत, सफलता की संख्या अवरोही क्रम- n......., 3, 2, 1, 0 - में होने पर (p+q) n का विस्तार होगा। विभिन्न पदों के गुणांक पास्कल त्रिभुज से ज्ञात किये जा सकते हैं। कुल परीक्षणों की संख्या अर्थात् 'N' से इस सम्भावना वितरण को गुणा करके प्रत्याशित आवृत्ति वितरण प्राप्त किया जा सकता है-

$N(q+p)^n$ या $N(p+q)^n$

(7) **उपयोग-** इसका उपयोग वहाँ उपयुक्त होता है जहाँ घटनाओं की सफलता-असफलता के आधार पर द्वन्द्वभाजन वर्गीकरण किया जा सकता है, जैसे सिक्का उछालने पर चित या पट गिरना, जनसंख्या का स्त्री-पुरुष में वितरण आदि। निर्णय लेने तथा भावी अनुमान के लिये यह एक आवश्यक उपकरण है।

बोध प्रश्न

1. सैद्धांतिक आवृत्ति वितरण का क्या अर्थ है।

....

....

....

2. सैद्धान्तिक आवृत्ति वितरण के महत्व की विवेचना कीजिए।

....

....

....

3. द्विपद वितरण की विशेषताएँ लिखिए।

....

....

....

2.5 सामान्य वितरण (Normal Distribution)

सामान्य वितरण के सर्वप्रथम अन्वेषण का श्रेय प्रसिद्ध अंग्रेजी-गणितज्ञ डॉ. ए.डी. मॉयर (Dr. A.D. Moiver) को है। बाद में विज्ञान में इसका प्रयोग फ्रांसीसी गणितज्ञ लेप्लेस (Laplace) द्वारा किया गया।

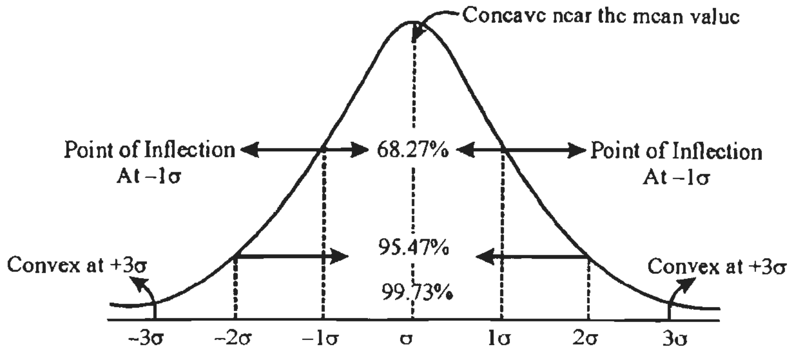
जर्मन-गणितज्ञ गॉस (Gauss) ने भी काफी विकास किया। कालान्तर में इसका प्रयोग क्वेलेट (बेलजियम) तथा गॉल्टन (इंगलैंड) ने भी किया तथा इसके क्षेत्र को विकसित किया। सामान्य वितरण को प्रसामान्य वितरण, सामान्य वक्र (Normal Curve), सामान्य आवृत्ति वक्र (Normal frequency distribution), सामान्य सम्भावना वक्र (Normal Probability Curve), सामान्य विभ्रम वक्र (Normal Curve of error) और गॉस-वक्र (Gauss-Curve) भी कहते हैं।

महत्व- सांख्यिकीय सिद्धांत में यह सबसे महत्वपूर्ण सम्भावना वितरण है। अधिकांश न्यादर्श (sample) इसी के आधार पर लिए जाते हैं। इसकी सहायता से यह अनुमान लगाया जाता है कि विशाल समग्र में से लिये गये न्यादर्श समग्र का प्रतिनिधित्व ठीक प्रकार से करते हैं या नहीं। इसकी सहायता से विभिन्न न्यादर्शों का विश्लेषण और तुलनात्मक अध्ययन सम्भव होता है। एक प्रसिद्ध सांख्यिकी श्री काउडन के शब्दों में इसका महत्व इस प्रकार स्पष्ट किया जा सकता है-

"The normal curve of error stands out in the experience of mankind as one of the broadest generalisation of natural philosophy. It serves as the guiding instrument in researches in the physical and social sciences and in medicine, agriculture and engineering. It is an indispensable tool for the analysis and the interpretation of the basic data obtained by observation and experiment."

सामान्य वक्र का आकार - यह घण्टी के आकार का होता है। यह पूर्णतया समभित (Symmetrical) होता है यद्यपि इसका वास्तविक आकार प्रमाप विचलन द्वारा निर्धारित होता है इस अभाव में न्यादर्श के सिद्धांत का विकास सम्भव नहीं था। इस वक्र का आकार निम्न प्रकार का होता है-

NOTES



इस वक्र को निम्न समीकरण के रूप में भी प्रस्तुत किया जा सकता है-

$$y = \frac{N}{\sigma\sqrt{(2\pi^{0})}} \quad \frac{-(x-\bar{x})^2}{2\sigma^2}$$

where, y = Ordinate of the curve at a point x.

N = Number of items in the frequency distributions.

σ = Standard deviation of distribution.

π = The ratio of the circumference of a circle to its diameter approximately 3.1416;

$\sqrt{(2\pi)}$ = 2.5066

e = 2.718 (Base of the Napierian Logarithm)

सामान्य वितरण के उपयोग (Uses of Normal Distribution)- आधुनिक सांख्यिकी में सामान्य-वितरण का केन्द्रीय स्थान है। इसके प्रमुख उपयोग इस प्रकार हैं-

(1) कई क्षेत्रों में समंकों के विचरण का स्वरूप बहुत कुछ सामान्य वितरण के अनुरूप होता है। प्राकृतिक विज्ञानों, प्राणिशास्त्र, मनोविज्ञान, समाजशास्त्र और आंशिक रूप से आर्थिक और औद्योगिक क्षेत्र में सामान्य-वितरण के कई उदाहरण मिलते हैं। एक पेड़ के पत्तों की लम्बाई, मनुष्यों की ऊंचाई, जन्म-मृत्यु और विवाह दरें, छात्रों के बौद्धिक स्तर का माप, मापन और अवलोकन की त्रुटियां, इस्पात के तारों की तनाव शक्ति, एक नस्ल के प्राणियों का आयु-वितरण आदि घटनाओं पर सामान्य वितरण के नियम लागू होते हैं। अतः इस वितरण का व्यापक रूप से उपयोग होता है।

(2) न्यादर्श सिद्धांत के आधार पर सांख्यिकीय निर्वचन में सामान्य वितरण का महत्व सर्वोपरि है। यदि प्रतिदर्श (Sample) बड़े और दैव-निदर्शन के आधार पर चुने गए हों तो प्रतिदर्श-मापों का वितरण सामान्य होता है भले ही समग्र के समंकों का वितरण सामान्य हो या नहीं हो। उदाहरण के लिये यदि हम किसी शहर के विश्वविद्यालय स्तर के 5,000 छात्रों के समग्र में से दैव-निदर्शन आधार पर 50-50 के 50 प्रतिदर्श चुनें, उनकी ऊंचाई का माप करें और सभी प्रतिदर्शों के समान्तर माध्य निकालें तो इन माध्य-ऊंचाइयों का वितरण सामान्य वितरण होगा चाहे पूरे समग्र के 5,000 छात्रों की ऊंचाई का वितरण सामान्य हो या न हो। न्यादर्श-वितरणों का यह गुण पूरे समग्र के माप का अनुमान लगाने में, विभिन्न प्रतिदर्श मापों की शुद्धता की जाँच करने में और न्यादर्श के उच्चावचनों की सीमाएँ निर्धारित करने में बहुत उपयोगी होता है। इसीलिये सामान्य वितरण को सांख्यिकीय निर्वचन तथा न्यादर्श-सिद्धांत का मूल-आधार माना जाता है।

(3) चर-मूल्यों का रूपान्तर करके अनेक अ-सामान्य वितरणों को सामान्य वितरण में बदला जा सकता है।

(4) सामान्य वितरण में कई गणितीय गुण होते हैं जिनके कारण इसकी उपयोगिता और भी बढ़ जाती है।

प्रसिद्ध सांख्यिकी यूडेन ने तो सामान्य वितरण की सार्वभौमिक उपयोगिता से प्रभावित होकर इसके उपयोग को सम्मित आकृति के रूप में ही प्रकट किया है- "विभ्रमों का सामान्य नियम मानव जाति के अनुभव में प्राकृतिक दर्शन के व्यापक सामान्यीकरण का महत्वपूर्ण सिद्धांत है। भौतिक तथा सामाजिक विज्ञानों, चिकित्सा, कृषि और इंजीनियरिंग में, शोधकार्य में यह निर्देशक उपकरण का कार्य करता है। अवलोकन तथा प्रयोग से प्राप्त मूल समंकों के विश्लेषण एवं निर्वचन में यह अपरिहार्य शास्त्र है।"

सामान्य-वितरण की मान्यताएँ (Assumptions of Normal Distribution)

(1) **स्वतंत्र कारण-** प्रत्येक अवलोकन तथा स्वतंत्र घटना पर प्रभाव डालने वाली कारण- सम्बन्धी शक्तियाँ परस्पर स्वतंत्र होना चाहिये।

(2) **विविध कारण-** व्यक्तिगत घटनाओं पर प्रभाव डालने वाली कारण- सम्बन्धी शक्तियाँ अनेक होना चाहिये और वे लगभग सभी समान भार या महत्व की होना चाहिए।

(3) **समिति** - सम्बन्धी मान्यता- सभी कारण संबंधी शक्तियों का ऐसा प्रभाव होना चाहिये कि समान्तर माध्य के दोनों ओर के विचलन आकार और संख्या में संतुलित हो जावें।

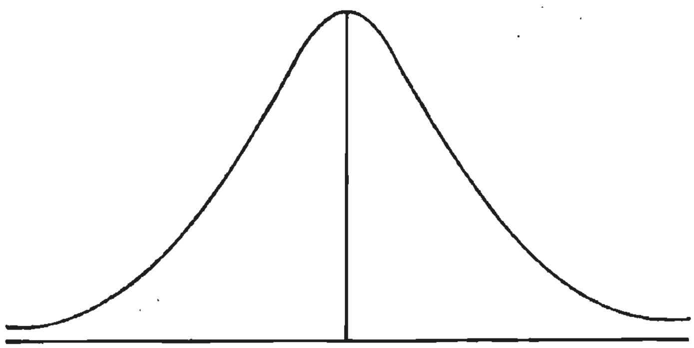
(4) **समरूपता की मान्यता-** जिस समग्र में से प्रतिदर्श का दैव-चुनाव किया जावे, उस समग्र के सम्बन्ध में ये शक्तियाँ एक समान होना चाहिए भले ही उनका प्रभाव भिन्न-भिन्न घटनाओं पर भिन्न-भिन्न पड़े।

स्पष्ट है कि उपरोक्त मान्यताएँ बहुत कठोर हैं जिसके कारण यह वितरण केवल एक आदर्श-वितरण मात्र रह जाता है। व्यावहारिक जीवन में, प्राकृतिक विज्ञानों, भौतिकी, जीवशास्त्र, मनोविज्ञान, कृषि आदि क्षेत्रों में उन सभी घटनाओं पर यह वितरण लागू होता है जो दैव पर आधारित होती हैं। आर्थिक क्षेत्र में सामान्य वितरण का प्रयोग सीमित होता है।

प्रसामान्य वक्र की विशेषताएं (Characteristics or Properties of the Normal Curve)

सामान्य वक्र में निम्नलिखित विशेषताएं पायी जाती हैं :

(1) **घण्टाकार आकृति** (Bell Shaped)- प्रसामान्य वक्र पूर्ण रूप से सममित होता है। यह वक्र घण्टी के आकार का होता है। यदि शीर्ष बिन्दु से आधार रेखा पर लम्ब डाला जाय तो यह लम्ब नीचे के क्षेत्रफल को दो बराबर भागों में विभाजित करेगा। यदि इसके एक ओर के आधे भाग को मोड़ा जाय तो यह दूसरे भाग को पूर्णतः ढंक लेगा। निम्नांकित चित्र से स्पष्ट होता है।



(2) **सतत वितरण** (Continuous Distribution)- प्रसामान्य वितरण एक सतत या अखण्डित वितरण है जबकि द्विपद एवं प्वॉयसन वितरण खण्डित है।

(3) **केन्द्रीय मापों की समानता** (Equality of Central Value)- पूर्ण सममित वितरण होने के कारण सभी केन्द्रीय माप समान होते हैं अर्थात् $\overline{X} = Z = M$।

(4) **एक भूयिष्ठक** (Uni-modal)- प्रसामान्य वितरण में अधिकतम ऊंचाई एक ही स्थान पर होती है अतः यह एक ही भूयिष्ठक वाला वितरण होता है।

(5) **मध्यका से चतुर्थकों की समान दूरी** (Equal distance of Quartiles from Median)- इस वितरण में प्रथम एवं तृतीय चतुर्थक मध्यका से समान दूरी पर होते हैं अर्थात् $Q_3 - M = M - Q_1$

(6) **चतुर्थक विचलन एवं प्रमाप विचलन का सम्बन्ध** (Relationship between Q.D. and S.D.)- चतुर्थक विचलन (Q.D.) प्रमाप प्रचलन (S.D.) का 0.6745 अर्थात् $\frac{2}{3}$ होता है।

(7) **माध्य विचलन एवं प्रमाप विचलन में सम्बन्ध** (Relationship between M.D. and S.D.)- ऐसे वितरण में माध्य विचलन (M.D.) प्रमाप विचलन (S.D.) का .7979 अर्थात् $\frac{4}{5}$ होता है।

NOTES

(8) **चतुर्थक विचलन, माध्य विचलन एवं सम्भाव्य विभ्रम** (Q.D., M.D. and P.E.)- प्रसामान्य वितरण में चतुर्थक विचलन तथा सम्भाव्य विभ्रम समान होते हैं अर्थात् Q.D. = P.E. सम्भाव्य विभ्रम माध्य विचलन का .845 गुना होता है, अर्थात्,

$$P.E. = .845\,M.D. \quad \text{अथवा} \quad P.E. = \frac{5}{6}\delta$$

(9) **अनन्तस्पर्शी** (Asymptotic to the base line)- इस वक्र के दोनों सिरे एक ही प्रकार से दोनों ओर बढ़ते हैं। वे आधार रेखा के निकटतम आते-जाते हैं पर उसका स्पर्श नहीं करते हैं। सैद्धांतिक दृष्टि से वक्र अनन्त (Infinity) की ओर अग्रसर दोनों दिशाओं में होता है।

(10) **प्रमाप विचलन एवं परिवर्तित बिन्दु** (S.D. and Changed Points)- वक्र के जिन बिन्दु पर झुकाव की दिशा में परिवर्तन होता है यदि वहाँ से आधार रेखा पर लम्ब डाले जाएं तो समान्तर माध्य से उन लम्बों तक की दूरी प्रमाप विचलन की विभिन्न इकाइयों को प्रकट करती है। यथा $\overline{X} + \sigma; \overline{X} + 2\sigma; \overline{X} + 3\sigma;$

(11) **प्रसामान्य वितरण के अचर मूल्य** (Constant Values of Normal Distribution)- प्रसामान्य वितरण के अचर मूल्य निम्नलिखित हैं :

समान्तर माध्य $= X$ अथवा μ, प्रमाप विचलन $= \sigma$

परिघात- प्रथम द्वितीय तृतीय चतुर्थ

$$\mu_1 = 0, \quad \mu_2 = \sigma^2, \quad \mu_3 = 0, \quad \mu_4 = 3\,\mu_2^2 = 3\sigma^4$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0, \beta_2 = \frac{\mu_4}{\mu_3^2} = \frac{3\mu_2^2}{\mu_2^2} = 3$$

प्रसामान्य वितरण सममित होने के कारण इसके विषम (odd) परिघातों का मूल्य सदैव शून्य होता है और इसलिए विषमता भी शून्य होगी, परन्तु पृथुशीर्षत्व (Kurtosis) जो मध्य शीर्ष वाले वितरण का प्रतीक है, का मान 3 होता है।

(12) **वितरण के प्राचल** (Parameters of Distribution)- प्रसामान्य वितरण के दो प्रमुख प्राचल हैं : समान्तर माध्य (X) एवं प्रमाप विचलन (σ)। इन दोनों की सहायता से सम्पूर्ण वितरण लिखा जा सकता है। μ (म्यू) संकेताक्षर सम्पूर्ण समग्र के समान्तर माध्य के लिए प्रयोग होता है।

(13) **कोटि-सम्बन्ध** (Ordinate Relationship)- प्रसामान्य वक्र के समान्तर माध्य पर कोटि अक्ष की अधिकतम ऊंचाई होती है। कोटि अक्ष पर माध्य से एक बार प्रमाप विचलन की दूरी माध्य पर कोटि अक्ष की ऊंचाई की 60.653 प्रतिशत होती है। इस प्रकार अन्य दूरियों पर भी यह ऊंचाई का अनुपात निश्चित होता है। ये अनुपात कोटि अक्ष की ऊंचाई के अनुपात सम्बन्धित सारणी (Table of Ordinates of the Normal Curve) से देखे जा सकते हैं।

(14) **क्षेत्र-सम्बन्ध** (Area Relationship)- प्रसामान्य वक्र की सबसे महत्वपूर्ण विशेषता क्षेत्रफल सम्बन्धी है। आधार रेखा और प्रसामान्य वक्र के मध्य का क्षेत्र प्रसामान्य वक्र के अधीनस्थ क्षेत्र (area under the normal curve) कहलाता है। माध्य कोटि अक्ष के साथ प्रमाप विचलन (σ) की माध्य से निश्चित दूरियों पर कोटि अक्ष का सदैव एक ही सम्बन्ध होता है। इसका अर्थ यह है कि माध्य कोटि अक्ष और माध्य से किसी निश्चित प्रमाप विचलन की दूरी पर कोटि अक्ष के मध्य वक्र का क्षेत्रफल वक्र सम्पूर्ण क्षेत्रफल का सदैव एक निश्चित अनुपात होता है।

माध्य एवं प्रमाप विचलन के आधार पर प्रसामान्य वक्र का क्षेत्रफल निम्न प्रकार होता है:

(i) माध्य + 1 σ के अन्तर्गत 68.268% क्षेत्रफल, 34.134% क्षेत्रफल माध्य के दोनों ओर होता है।

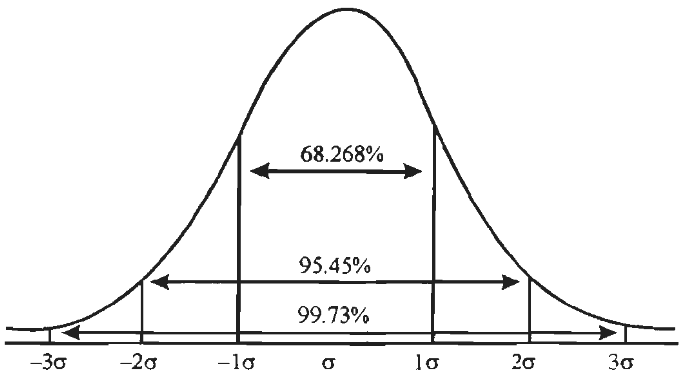
(ii) माध्य + 2 σ के अन्तर्गत 95.45% क्षेत्रफल, 47.725% क्षेत्रफल माध्य के दोनों ओर होता है।

(iii) माध्य + 3 σ के अन्तर्गत 99.73% क्षेत्रफल, 49.865% क्षेत्रफल माध्य के दोनों ओर होता है।

केवल .27% क्षेत्रफल वक्र से बाहर रह जाता है।

निम्न चित्र क्षेत्रफल सम्बन्ध को स्पष्ट प्रकट करता है :

प्रसामान्य वक्र में क्षेत्रफल सम्बन्ध



पमाणित प्रसामान्य वितरण (Standard Normal Distribution)- प्रसामान्य वक्र के अधीनस्थ आने वाले विविध क्षेत्रफल को सारणी से देखा जा सकता है, परन्तु जब $\overline{X}$ और σ का मान अलग-अलग होता है तो कठिनाई उत्पन्न हो जाती है क्योंकि ऐसी स्थिति में अलग-अलग सारणियों का प्रयोग करना होता है। इस कठिनाई को दूर करने के लिए अलग-अलग माध्य ($\overline{X}$) और प्रमाप विचलन (σ) वाले प्रसामान्य वितरण को 'प्रमापित प्रसामान्य वितरण' में बदल लिया जाता है। इसे **'इकाई प्रसामान्य वितरण'** भी कहते हैं। इसके अन्तर्गत माध्य का मूल्य शून्य ($\overline{X} = 0$) और प्रमाप विचलन एक इकाई ($\sigma = 1$) होता है। इस प्रक्रिया को Z परिवर्तन विधि (Z = transformation technique) कहते हैं।

सूत्र : $Z = \dfrac{X - \overline{X}}{\sigma}$ अथवा $Z = \dfrac{X}{\sigma}$ $\quad [\because X = X - \overline{X}]$

जहां Z = z-परिवर्तन विधि

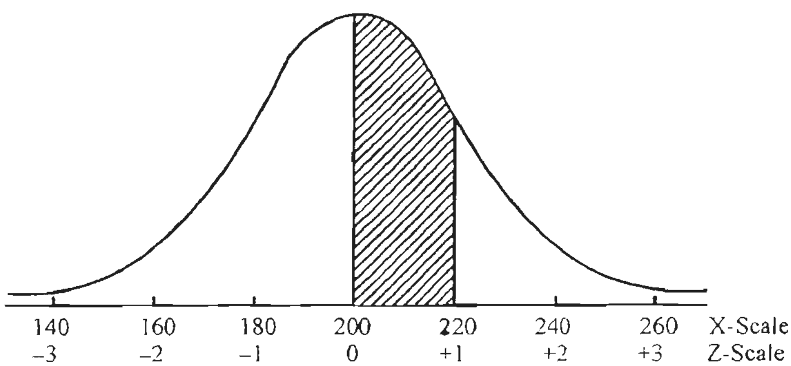
x = अवलोकित मान

$\overline{X}$ = वितरण का माध्य

σ = वितरण का प्रमाप विचलन

माना कि एक वितरण में माध्य ($\overline{X}$) = 200 एवं प्रमाप विचलन (σ) = 20 है अतः यदि X का मूल्य 200 है तो Z का मूल्य होगा $= \dfrac{200 - 200}{20} = 0$। इसी प्रकार यदि X के मूल्य हैं 180 एवं 220 तो Z के मूल्य होंगे क्रमशः $= \dfrac{180 - 200}{20} = \dfrac{-20}{20} = -1; \dfrac{220 - 200}{20} = \dfrac{20}{20} = 1$

इसी प्रकार अन्य Z के मूल्य ज्ञात किए जा सकते हैं। प्रसामान्य वक्र को Y और Z के रूप में निम्न चित्र में प्रदर्शित किया गया है :



क्षेत्रफल देखने की विधि (The Method of obtaining Area)

Z-मूल्य का क्षेत्रफल देखने के लिए प्रसामान्य वक्र के अन्तर्गत क्षेत्रफल (Area under the Normal Curve) नामक सारणी में पहले खाना Z अर्थात् $\frac{x}{\sigma}$ का होता है। इसमें 0.0 से 3.0 तक के मूल्य लिखे रहते हैं।

NOTES

इससे आगे समानान्तर .00 से .09 तक के 10 खाने होते हैं जिनमें Z-मूल्य के क्षेत्रफल अनुपात दिये होते हैं। यदि Z का मूल्य 1.45 है तो क्षेत्रफल 1.4 की सीध में .05 वाले खाने में देखेंगे जहां .4265 दिया हुआ है अर्थात् 42.65%। Z का मूल्य धनात्मक होने पर क्षेत्रफल माध्य कोटि से दाहिनी ओर माना जायेगा और यदि Z का मूल्य ऋणात्मक है तो क्षेत्रफल माध्य कोटि से बायीं ओर माना जाता है। प्रारम्भ से Z का ऋणात्मक मूल्य निकालने के लिए क्षेत्रफल के आधे अर्थात् .5000 में से सारणी मूल्य को घटा दिया जाता है यथा (.5000 – .4265 = .0735), यदि Z की दो सीमाओं अर्तात् Z_1 एवं Z_2 के मध्य का क्षेत्रफल ज्ञात करना हो तो दोनों Z के सारणी मूल्यों को जोड़ दिया जाता है जबकि Z_1 का मान ऋणात्मक तथा Z_2 का मान धनात्मक है।

$Z_1 = -1.45$ एवं $Z_2 = +1.05$ के मध्य का क्षेत्रफल = – 1.45 का क्षेत्रफल = .4265 + (1.05 का क्षेत्रफल) = .4265 + .3531 = .7796 होगा।

यदि $Z_1 = -1.45$ एवं $Z_2 = -1.05$ के मध्य का क्षेत्रफल ज्ञात करना है तो .4265 में से .3531 घटाने पर वांछित क्षेत्रफल = .4265 - .3531 = .0734 होगा।

Illustration 9. 100 प्रयासों में 5 सफलताएं प्राप्त करने की आवृत्ति को प्रसामान्य बंटन की सहायता से किस प्रकार ज्ञात करोगे यदि प्रत्येक प्रयास में सफलता की प्रायिकता p = 0.1 हो?

How would you use normal distribution to find approximately the frequency of exactly 5 successes in 100 trails, the probability of success in each trail being p = 0.1?

हल (Solution)

माना प्रयासों की संख्या = n, तब $n=100$, $p=0.1$ तथा $q=1-q=1-.1=0.9$

द्विपद वितरण की अवस्था में माध्य $= np = 100 \times 0.1 = 10$

तथा $\sigma = \sqrt{npq} = \sqrt{100 \times 0.1 \times 0.9} = \sqrt{9} = 3$

जब n बड़ी संख्या होती है तो द्विपद वितरण सामान्य प्रकार का होता है जो अविच्छिन्न वितरण होता है। इसलिए सामान्य वितरण में 5 सफलता उस आवृत्ति के समान होगी जो 4.5-5.5 वर्ग में स्थित होगा और द्विपद समीकरण का माध्य और प्रमाप ही सामान्य वितरण का माध्य और प्रमाप विचलन होगा।

$$X = 4.5 \text{ के लिए } Z_1 = \frac{4.5 - 10}{3} = -1.83$$

$$X = 5.5 \text{ के लिए } Z_2 = \frac{55 - 10}{3} = -1.5$$

अब सारणी से 1.83σ के लिए क्षेत्रफल = 0.4664

-1.5σ के लिए क्षेत्रफल = 0.4332

-1.83 और -1.5 के बीच क्षेत्रफल = 0.4664 – 0.4332 = 0.0332

जब N कुल आवृत्ति हो तो 4.5 और 5.5 वर्ग की आवृत्ति = .0332 N

यदि N = 100 मान लिया जाय तो आवृत्ति = 3.32%

Illustration 10. You are incharge of rationing in a State affected by goods shortage. Following reports were received from the investigators-

Daily calorific value of food available per adult during current period-

Area	Mean	S.D.
A	2000	350
B	1750	100

The estimated requirement of an adult is taken at 2500 calories daily and the absolute minimum of 1000. Comment on the reported figures and determine which area in your opinion need more urgent attention.

Solution-

Area A	Area B
Mean $\pm$ 3σ	Mean $\pm$ 3σ
2000 $\pm$ (3 $\times$ 350)	1750 $\pm$ (3 $\times$ 100)
Between 3050 and 950 calories	Between 2050 and 1450 calories

Hence, area A needs urgent attention because there are some people to whom only 950 calories are available when the absolute minimum is 100 calories.

Illustration 11. (Normal curve by formula) Average sales of Bata Shoe Company's Multiple shop is Rs. 12,500 per month with a standard deviation of Rs. 4050. Find out what proportion of all shops sold between Rs. 12,500 and Rs. 16,550.

Solution.

$$\frac{x-\bar{x}}{\sigma} = \frac{16550 - 12500}{4050} = \frac{4050}{4050} = 1.00$$

By referring to table, we find the area covered by the normal curve = .34135 = .34134
or 34134% shops sold between Rs. 12500 and Rs. 16550.

Illustration 12 (i) एक कम्पनी द्वारा उत्पादित इस्पात की छड़ों का समान्तर माध्य 10 मीटर और प्रमाप विचलन 20 सेमी है। एक कांट्रेक्टर ने 5,000 छड़ें खरीदी हैं। बताइये इनमें से कितनी छड़ों के 9.75 मीटर से छोटे होने की सम्भावना है। अपने उत्तर के लिये यह मान लीजिए कि इस्पात की छड़ों की लम्बाई सामान्य रूप से फैली हुई है और सामान्य-वक्र में माध्य और माध्य से दूरी $\frac{(x-x)}{\sigma}$ के बीच क्षेत्रफल बनाने वाली सारणी के निम्नलिखित उद्धरण का उपयोग कीजिए-

$\frac{(X-\bar{X})}{\sigma}$	:	1.10	1.15	1.20	1.25	1.30
क्षेत्रफल	:	3643	.3749	.3849	.3944	.4032

(ii) एक परीक्षा में बैठने वाले 1,000 छात्रों के माध्य-प्राप्तांक 34.4 तथा प्रमाप-विचलन 16.6 हैं। सामान्य वितरण मानते हुए यह बतलाइये कि 30 तथा 60 के बीच अंक पाने वाले कितने छात्र होंगे। केन्द्रीय 70% छात्रों के प्राप्तांकों की सीमाएं भी निर्धारित कीजिए।

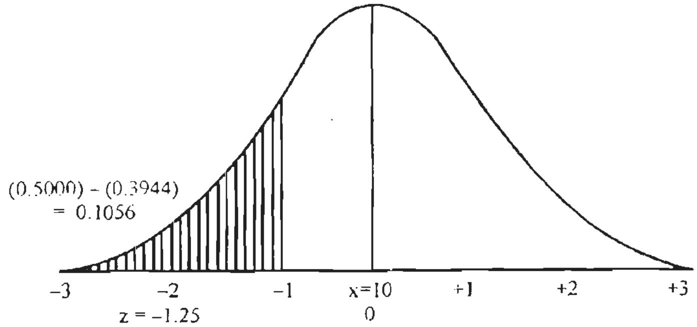
Solution - (i) $\bar{x}$ = 10 मीटर; $\sigma = 0.2$ मीटर

$$Z = \frac{(x-\bar{x})}{\sigma}; x = 9.75,$$

तो, $$Z = \frac{(9.75-10)}{0.2} = \frac{-0.25}{0.2} = -1.25$$

माध्य कोटि से $-$ 1.25 तक का क्षेत्रफल = .3944

अत: प्रारम्भ से Z = $-$ 1.25 तक का क्षेत्रफल = .5000 $-$ .3944 = .1056



NOTES

इस प्रकार 9.755 मीटर से कम लम्बाई की छड़ों का अनुपात 0.1056 है और उनकी संख्या $500 \times .1056 = 528$ है।

(ii) $\bar{x} = 34.4;\ \sigma = 16.6;\ N = 1000$

30 और 60 के मध्य प्राप्तांक पाने वाले छात्रों की संख्या-

$$Z_1 = \frac{30 - 34.4}{16.6} = -0.265;$$ मध्य से $Z = -.265$ तक का क्षेत्रफल $= 0.1045$

$$Z_2 = \frac{60 - 34.4}{16.6} = +1.54;$$ मध्य से $Z = +1.54$ तक का क्षेत्रफल $= 0.4382$

$\therefore$ 30 और 60 के मध्य प्राप्तांक पाने वाले छात्रों की संख्या -

$$= 0.5427 \times 1000 = 543$$

मध्य (केन्द्र) में स्थित 70% छात्रों के प्राप्तांकों की सीमा-केन्द्रीय कोटि-अक्ष से बाई ओर 35% (–.3500) और दाहिनी ओर 35% (+.3500) से सम्बद्ध Z ज्ञात किये जावेंगे। सारणी में .3500 क्षेत्रफल से सम्बद्ध Z का मूल्य 1.037 है। अतः सीमाएँ $\bar{x} + 1.037\,\sigma$ होगी।

न्यूनतम सीमा $\quad \bar{x} - 1.037\,\sigma = 34.4 - 1.037 \times 16.6 = 17.2$

अधिकतम सीमा $\quad \bar{x} - 1.037\,\sigma = 34.4 + 1.037 \times 16.6 = 51.6$

अतः माध्य के 70% छात्रों के प्राप्तांक 17.2 (17) और 51.6 (52) के बीच में होंगे।

Illustration 13. 1000 छात्रों पर किये गए एक बौद्धिक परीक्षण में माध्य और प्रमाप विचलन क्रमशः 42 तथा 24 थे। उन छात्रों की संख्या मालूम कीजिए जिनके अंक (i) 50 से अधिक हों (ii) 30 और 54 के बीच हों, तथा (iii) सर्वोच्च 100 छात्रों द्वारा प्राप्त न्यूनतम अंक मालूम कीजिए।

Solution- (i) $\bar{x} = 42;\ \sigma = 24;\ x = 50$

यदि $x = 50$ है तो $Z = \dfrac{50 - 42}{24} = +0.33;$ क्षेत्रफल .1304 हैं।

$Z = 50$ से अधिक वाला क्षेत्रफल $= .5000 - .1304 = .3696$

50 से अधिक अंक प्राप्त करने वाले छात्रों की संख्या $= .3696 \times 1000 = 370$

(ii) $$Z_1 = \frac{30 - 42}{24} = -0.5;$$ क्षेत्रफल 0.1915

$$Z_2 = \frac{54 - 42}{24} = +0.5;$$ क्षेत्रफल 0.1915

Z_1 और Z_2 के बीच का क्षेत्रफल $= 0.1915 + 0.1915 = 0.3830$

अतः 30 और 54 के बीच अंक पाने वालों की संख्या $= 0.3830 \times 1000 = 383$

(iii) सर्वोच्च अंकों वाले 100 छात्रों की सम्भावना $= \dfrac{100}{1000} = 0.1$

वक्र में दाहिनी ओर अन्तिम 0.1 क्षेत्रफल = केन्द्र से $.5 - .1 = .4$ से अधिक वाला क्षेत्र।
.4000 क्षेत्रफल के लिये Z का मूल्य 1.281 है।

$$Z = \frac{x - \bar{x}}{\sigma};\ 1.281 = \frac{x - 42}{24};\ x - 42 = 30.744$$

$$x = 42 + 30.744 = 72.744$$

अतः सर्वोच्च 100 छात्रों के न्यूनतम प्राप्तांक 72.7 हैं।

2.6 प्वायसन वितरण (Poisson Distribution) :

प्वायसन वितरण (Poisson Distribution) का प्रतिपादन सन् 1837 ई. में फ्रेंच गणितज्ञ साइमन डेनिस प्वॉयसन ने द्विपद-विस्तार की सहायता से किया। इसे प्वॉयसन संभाविता वितरण भी कहते हैं। जब घटनाएं असामान्य

एवं दुर्लभ होती हैं तथा उनके घटित होने की संभावना बहुत कम होती है तब सही स्थिति निरूपित करने के लिए प्वॉयसन वितरण का प्रयोग होता है तथा, एक देश में भयंकर बाढ़ संख्या, अणु परीक्षण से प्रभावित नहीं होने की संभावना, किसी नगर में एक निर्दिष्ट अवधि में हृदयरोग से मरने वालों की संख्या, पानी की एक बूंद में कीटाणुओं की संख्या।

प्वॉयसन वितरण का स्वरूप (Form of Poisson Distribution) :

यह द्विपद वितरण की भांति खण्डित होता है। अतः सफलताएँ धनात्मक एवं पूर्णांक होती हैं। 1,2,3.......... सफलताओं की संभावनाएँ निम्न प्रकार के विभिन्न पदों से ज्ञात हो जाती हैं :-

$$p(r) = \frac{C^{-m} m^r}{r}$$

where r = Number for which probable frequencies are to be calculated.

C = Base for the natural Logarithms. (C = 2.7183)

m = Arithmetic average or mean.

r = r!

प्वॉयसन वितरण, द्विपद वितरण की सीमान्त रूप है जिसमें 'n' बहुत बड़ा 'p' बहुत छोटा तथा 'm' एक परिमित धनात्मक संख्या होती है।

उपरोक्त गणन क्रिया में C^{-m} का मान निकालने में विशेष सावधानी की आवश्यकता है। इसका मान इस प्रकार निकाला जाता है-

$$C^{-m} = \frac{1}{m} = \frac{1}{(2.7183)^m}$$

जहाँ तक (2.7183) m का संबंध है। यह लघुगणक की सहायता से निकाला जाता है।

माना कि $(2.7183)^m = Z$

अतः

$$\text{Log } Z = m \log (2.7183)$$

$$Z = \text{Antilog } m \log (2.7183)$$

$$= \text{Antilog } m = \times 0.4343$$

$$= \text{Antilog } 0.4343$$

$$Z = (2.7183)^m = \mathit{Antilog}\ .4343^m$$

then, $$C^{-m} = \frac{1}{(2.7183)^m} = \frac{1}{\mathit{Antilog}\ .4343^m}$$

$$= \text{Reciprocal } (\text{A. log . } 4343^m)$$

यहाँ 'm' का कोई भी मान रखकर C^{-m} ज्ञात कर लिया जाता है।

प्रत्याशित आवृत्तियाँ (Expected Frequencies) :

0,1,2,3......... सफलताओं की प्वॉयसन संभावनाएँ क्रमानुसार लिख दी जावें तो प्वॉयसन वितरण प्राप्त हो जाता है। यदि परीक्षणों की संख्या N हो तो प्रत्याशित आवृत्तियाँ मालूम करने के लिए प्रत्येक पद की संभावना को परीक्षणों की संख्या (N) से गुणा कर दिया जाता है।

$$\text{Expected frequencies N.p (r)} = \frac{NC^{-m} m^r}{r}$$

प्वॉयसन वितरण के स्थिरांक (Constants of Poisson Distribution) :

समान्तर माध्य, प्रमाप विचलन, चारों परिघात, विषमता तथा पृथुशीर्षत्व के माप प्वॉयसन वितरण के स्थिरांक कहलाते हैं। इस वितरण के लिए इनके मान इस प्रकार हैं:-

NOTES

Mean	$= m$
Standard Deviation	$= \sqrt{m}$
First Moment (μ^1)	$= 0$
Second Moment (μ^2)	$= m$
Third Moment (μ^3)	$= m$
Fourth Moment (μ^4)	$= m + 3m^2$
Skewness (β_1)	$= \dfrac{1}{m}$
Kurtosis (β_2)	$= 3 + \dfrac{1}{m}$

$m = 0.0$ से $m = 0.99$ तक के लिए e^{-m} के मानों की तालिका

m	0	1	2	3	4	5	6	7	8	9
0.0	1.0000	.9900	.9802	.9704	.9608	.9512	.9418	.9324	.9231	.9139
0.1	.9048	.8869	.8869	.8681	.8694	.8607	.8521	.8437	.8353	.8207
0.2	.8187	.8025	.8025	.7945	.7866	.7788	.7711	.7634	.7558	.7483
0.3	.7408	.7261	.7261	.7189	.7118	.7047	.6977	.6907	.6839	.6771
0.4	.6703	.6570	.6570	.6505	.6440	.6376	.6313	.6250	.6188	.6126
0.5	.6065	.5945	.5945	.5886	.5827	.5770	.5712	.5655	.5599	.5543
0.6	.5488	.5379	.5379	.5326	.5273	.5220	.5169	.4117	.5066	.5016
0.7	.4966	.4868	.4868	.4819	.1771	.4724	.4677	.4630	.4548	.4538
0.8	.4493	.4404	.4404	.4360	.4317	.4274	.4234	.4190	.4148	.4107
0.9	.4066	.3985	.3946	.3946	.3906	.3867	.3829	.3791	.3753	.3716

$m = 1$ से $m = 10$ तक के लिए e^{-m} के मानों की तालिका

m	1	2	3	4	5
e^{-m}	0.36788	0.13534	0.04979	0.01832	0.006738
m	6	7	8	9	10
e^{-m}	0.002479	0.000912	0.000335	0.000123	0.006738

उपर्युक्त तालिका से m = 0.51 e^{-m} का मान देखने के लिए पहले स्तम्भ में 0.5 लें इसके सामने तथा स्तम्भ 1 के नीचे लिखा मान पढ़ें। इस प्रकार

$$e^{-0.51} = 0.6005$$

इसी प्रकार $e^{-0.61} = 0.5434$ तथा

$$e^{-2.51} = e^{-2.0 - 0.51}$$

$$= e^{-20} \times e^{-0.51} = 0.13534 \times 0.605 = .08127$$

Illustration 14. किसी उत्पादन का निरीक्षण करने पर प्रति इकाई 2 दोष पाए जाते हैं। प्वॉयसन वितरण का उपयोग करते हुए बिना किसी दोष, 3 दोष या 4 दोष वाली इकाइयों के पाए जाने की प्रायिकता ज्ञात कीजिए। (दिया है $e^{-2} = 0.135$)

Suppose that a manufactured product has 2 defects per unit if product is inspected. Using Poisson distribution calculate the probabilities of finding a product without any defect, 3 defects or 4 defects. (Given $e^{-2} = 0.135$).

हल (Solution)

माना x = प्रति इकाई दोषों की संख्या

तथा $p(x) = P(X = x) = e^{-m}\dfrac{m^x}{x\,!}, x = 0, 1, ..., \infty$

यहां m = 2

$$p(0) = e^{-2} = 0.135 \text{ (दिया है)}$$
$$p(1) = p(0) \times m = 1.35 \times 2 = 0.27$$
$$p(2) = p(1) \times \frac{m}{2} = .27 \times \frac{2}{2} = .27$$
$$p(3) = p(2) \times \frac{m}{3} = .27 \times \frac{2}{3} = 0.18$$
$$p(4) = p(3) \times \frac{m}{4} = .27 \times \frac{2}{4} = .9$$

इस प्रकार,

कोई भी दोष न पाये जाने की प्रायिकता $= p(0) = 0.135$

तीन दोष पाये जाने की प्रायिकता $= p(3) = 0.18$

चार दोष पाये जाने की प्रायिकता $= p(4) = 0.9$

Illustration 15 : The mean of the Poisson Distribution is 3.24. Find other contants of the distribution.

प्वॉयसन वितरण का माध्य 3.24 है। अन्य स्थिरांक मालूम कीजिए।

Solution :

$$m = 3.24 \text{ (given)}$$
$$\text{S.D.} = \sqrt{m} = \sqrt{3.24} = 1.8$$
$$\mu_1 = 0 \qquad \mu_2 = m = 3.24$$
$$\mu_3 = m = 3.24 \qquad \mu_4 = m + 3m^2$$
$$= 3.24 + 3\,(3.24)^2 = 34.732$$
$$\beta_1 = \frac{1}{m} = \frac{1}{3.24} = 0.38$$
$$\beta_2 = 3 + \frac{1}{m} = 3 + \frac{1}{3.24} = 3 + 0.38 = 3.38$$

Illustration 16 : The average number of defective articles in a certain manufacturing process is 3%. These articles are supplied in packets of 50. Use Poisson distribution to calculate the approximate number of packets containing none-defective, one defective, two defective and three defective articles in a consignment of 10,000 packets. (given $C^{-15} = .2231$)

किसी निर्माण प्रक्रिया में दोषपूर्ण सामान का माध्य (औसत) 3% है। यह सामान 50 के पैकेटों में रखा जाता है। प्वॉयसन वितरण का प्रयोग करते हुए 10,000 पैकेटों में ऐसे पैकेटों की संख्या बतलाइये जिनमें कोई सामान दोषपूर्ण न हो, एक दोषपूर्ण हो तथा तीन दोषपूर्ण हो।

Solution

$$N = 10{,}000$$
$$P = \frac{3}{100}$$
$$m = 50 \times \frac{3}{100} = 1.5$$
$$C^{-m} = C^{-1.5} = 0.2231$$
$$p\,(r) = \frac{NC^{-m}m^r}{\lfloor r}$$

NOTES

Packets without any defective article (p_0)

$$= \frac{NC^{-m} m^r}{\lfloor r}$$

$$= \frac{NC^{-m} m^0}{\lfloor 0} = 10000\ C^{-1.5} = 10000 \times .2231 = 2231$$

Packets with one defective article (p_1)

$$= \frac{NC^{-m} m^1}{\lfloor 1} = 10000 \times .2231 \times 1.5 = 3346.5$$

Packets with two defective articles (p_2)

$$= \frac{NC^{-m}\ m^2}{\lfloor 2}$$

$$= \frac{10000 \times .2231\ (1.5)^2}{\lfloor 2} = 2510$$

Packets with three defective articles (p)

$$= \frac{NC^{-m}\ m^3}{\lfloor 3} = \frac{10000 \times .2231\,(1.5)^3}{\lfloor 3} = 1255$$

Illustration 17 : One fifth percent of the blades manufactured by a factory turn out to be defective. The blades are supplied in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing no defective, one defective and two defective blades respectively in a consignment of 1,00,000 packets. (given $C^{-.02} = .9802$)

Solution.

Here $\quad p = \frac{1}{500}$; $n = 10$

$$m = np = \frac{1}{500} \times 10 = 0.02$$

$$N = 1{,}00{,}000$$

$$C^{-m} = C^{-.02} = 0.9802$$

$$p(r) = \frac{N\,C^{-m}\, m^r}{r\,!}$$

Packets with no defective blades :

$$= \frac{100{,}000 \times 0.9802 \times (.02)^0}{0\,!}$$

Packets with one defective blade :

$$= \frac{100{,}000 \times 0.9802 \times (.02)^1}{1\,!} = 98020 \times \frac{2}{100} = 1960$$

Packets with two defective blades :

$$= \frac{100{,}000 \times 0.9802 \times (.02)^2}{2\,!} = 19.6 \text{ or } 20$$

Illustration 18 : The distribution of Typing mistakes committed by a typist is given below. Assuming a Poisson Model, find out the expected frequencies :

Mistakes per page :	0	1	2	3	4	5
No. of pages :	142	156	69	27	5	1

Solution :

Mistakes per page (x)	No. of Pages (f)	xf
0	142	0
1	156	156
2	69	138
3	27	81
4	5	20
5	1	5
	400	400

$$m = \frac{\Sigma xf}{\Sigma f} = \frac{400}{400} = 1$$

Taking the mean of the given distribution as the mean of the Poisson distribution, we want to fit, we get m = 1 and

$$p(r) = \frac{C^{-m} m^r}{r^1}; r = 0, 1, 2, ...,$$

Also $\quad P(0) = C^{-m} = C^{-1} = 0.3679$ (By Log tables)

Computation of Expected Frequencies

x	Expected frequencies	Observed frequency
0	$400 \times 0.3679 = 147$	142
1	$400 \times 0.3679 \times 1 = 147$	156
2	$\frac{400 \times 0.3679 \times (1)^2}{2!} = 74$	69
3	$\frac{400 \times 0.3679 \times (1)^3}{3!} = 25$	27
4	$\frac{400 \times 0.3679 \times (1)^4}{4!} = 6$	5
5	$\frac{400 \times 0.3679 \times (1)^5}{5!} = 1$	1
Total	400	400

प्वॉयसन वितरण की विशेषताएँ (Characteristics of Poisson Distribution)-

(1) **खण्डित वितरण** – द्विपद-वितरण की भाँति यह भी एक खण्डित वितरण है जिसमें सफलताओं की संख्या 0, 1, 2, 3........ x आदि पूर्णांकों के रूप में होती है। इनकी सम्भावना प्वॉयसन-प्रमेय के अनुसार ज्ञात की जाती है।

(2) **'p' और 'q' की मात्रा** – यह उन स्थितियों में प्रयुक्त होता है जिनमें घटनाओं के घटने की सम्भावना 'p' बहुत कम होती है और न घटने की सम्भावना 'q' अत्यधिक (लगभग 1 के बराबर) होती है और 'n' भी अधिक होता है।

(3) **प्रमुख प्राचल** (Parameter)– समान्तर माध्य (m = np) इस वितरण का प्रमुख प्राचल है और उसका मान ज्ञात होने से पूरा वितरण लिखा जा सकता है।

(4) **अचर मूल्य** (Constants)- प्वॉयसन वितरण के निम्नलिखित अचर मूल्य हैं –

समान्तर माध्य m = np; प्रमाप विचलन $\sigma = \sqrt{m} = \sqrt{np}$

परिघात $\mu_1 = 0; \mu_2 = m;$ आदि।

NOTES

(5) **स्वरूप** - यह वितरण असममित होता है। जैसे-जैसे समान्तर माध्य (m) का मूल्य बढ़ाता जाता है, यह वितरण दाहिनी ओर को प्रवृत्त होता जाता है और विषमता की मात्रा में कुछ कमी होती जाती है।

(6) **मान्यता** - इस वितरण की यह आधारभूत मान्यता है कि घटना की प्रत्याशा स्थिर होनी चाहिये। उस परिस्थिति में जबकि प्रमुख-प्राचल में ही परिवर्तन होता रहता है, इस वितरण का प्रयोग नहीं किया जाना चाहिये। इसके लिये कुछ सांख्यिकों ने इस वितरण का एक संशोधित रूप प्रस्तुत किया है।

(7) **उपयोग** - उन सभी घटनाओं में, जिनकी सफलता की सम्भावना (p) बहुत कम हो और (n) अधिक हो, यह वितरण उपयोगी सिद्ध होता है। प्राणीशास्त्र, जनांकिकी, यातायात-नियंत्रण, दूरभाष-संदेशवाहन, उद्योग तथा सांख्यिकीय गुण-नियंत्रण की विभिन्न समस्याओं का विश्लेषण प्वॉयसन वितरण के आधार पर होता है।

बोध प्रश्न

1. सांख्यिकी में सामान्य वितरण की क्या महत्ता है?

....

....

....

2. प्रसामान्य वक्र की विशेषताओं को समझाइए।

....

....

....

3. प्यॉगसन वितरण का क्या अर्थ है? इसकी विशेषताएँ लिखिए।

....

....

....

2.8 सारांश

सैद्धांतिक आवृत्ति वितरण आधुनिक सांख्यिकी के आधार माने गये हैं। जिनके माध्यम से निश्चित मान्यताओं के आधार पर समंक की प्रवृत्ति का ज्ञान होता है एवं विवेकपूर्ण निर्णय लेने के आधार मिलते है। इसके आधार पर भावी पूर्वनुमान लगाये जाते हैं। इन्हें प्रत्याशित या आदर्श आवृत्ति वितरण भी कहा जाता है। सैद्धांतिक दृष्टि से आवृत्ति वितरण कई प्रकार के हो सकते हैं, किन्तु उनमें से मुख्य तीन का सांख्यिकीय विश्लेषण में सर्वाधिक महत्व है, जो निम्न है– 1. द्विपद वितरण, 2. सामान्य वितरण, 3. प्वॉयसन वितरण।

2.9 शब्द कुंजी

(1) सैद्धांतिक वितरण, (2) खंडित वितरण, (3) आवृत्ति बहुभुज, (4) अचर मूल्य, (5) घंटाकार आवृत्ति, (6) अनन्तस्पर्शी।

2.10 अभ्यास प्रश्न

दीर्घउत्तरीय प्रश्न (Long Answer type Questions)

1. सैद्धांतिक वितरण का क्या अर्थ है?
   What is meant by Theoretical Distribution?
2. द्विपद बंटन का वर्णन कीजिए।
   Describe Binomial distribution.

3. द्विपद बंटन की विशेषताओं का वर्णन कीजिए।
   Give the characteristics of Binomial distribution.
4. सैद्धांतिक आवृत्ति का क्या अर्थ है ?
   What is meant by theoretical frequency?
5. द्विपद वितरण पर उन दशाओं जिनमें वह उत्पन्न होता है तथा उसकी प्रमुख विशेषताओं का वर्णन करते हुए एक टिप्पणी लिखिए।
   Write a note on the Binomial distribution describing the situation in which it arises and its chief characteristics.
6. उन परिस्थितियों की उदाहरण सहित व्याख्या कीजिए जिनके अन्तर्गत द्विपद बंटन प्राप्त होता है। इस बंटन का समान्तर माध्य और प्रसरण ज्ञात कीजिए।
   Explain the conditions under which binomial distribution is obtained, giving examples. Obtain the mean and variance of this distribution.
7. प्वॉयसन बंटन की परिभाषा दीजिए।
   Define Poisson distribution.
8. प्वॉयसन बंटन के प्रमुख लक्षणों को स्पष्ट कीजिए।
   Explain the main characteristics of Poisson distribution.
9. द्विपद आवृत्ति वितरण कब प्वॉयसन आवृत्ति वितरण की ओर प्रवृत्त होता है ? आप किन परिस्थितियों में द्विपद वितरण के स्थान पर प्वॉयसन वितरण का प्रयोग करेंगे ?
   When does a Binomial distribution tend to a Poisson distribution? In which case would you apply a Poisson distribution in place of Binomial distribution?
10. किन परिस्थितियों में प्वॉयसन आवृत्ति वितरण का प्रयोग किया जा सकता है ? प्वॉयसन वितरण का सामान्य पद बताइये और इसका माध्य और प्रसरण निकालिए।
    In what circumstances may a Poisson distribution be used? Give the general terms of Poisson distribution and derive its mean and variance.
11. द्विपद वितरण तथा प्वॉयसन वितरण किस प्रकार भिन्न हैं ?
    How do Binomial distribution and Poisson distribution differ?
12. द्विपद तथा प्वॉयसन वितरणों में माध्य तथा प्रमाप विचलन के मूल्यों की गणना किस प्रकार की जाती है ?
    How mean and standard deviation are computed in Binomial and Poisson distributions?
13. प्वॉयसन वितरण का क्या अर्थ है ? प्रति दिन के जीवन में जहां यह वितरण लागू होता है उनमें से कुछ के उदाहरण दीजिए।
    What is meant by Poisson distribution? Give some examples of its occurence in every day life.
14. प्वॉयसन वितरण की विशेषताओं की विवेचना कीजिए।
    Discuss the characteristics of Poisson distribution.
15. सैद्धांतिक आवृत्ति वितरण का क्या अर्थ है ? द्विपद, प्रसामान्य एवं प्वॉयसन वितरणों के मुख्य लक्षणों की विवेचना कीजिए।
    What is meant by theoretical frequency distribution. Discuss the salient features of the Binomial, Normal and Poisson distribution.
16. सैद्धांतिक आवृत्ति वितरणों से आप क्या समझते हैं ? इन वितरणों के महत्व की विवेचना कीजिए।
    What do you understand by theoretical frequency distribution? Discuss the importance of these distributions.

17. द्विपद आवृत्ति वितरण कब प्वॉयसन आवृत्ति वितरण की आवृत्ति प्राप्त करता है? आप किन परिस्थितियों में द्विपद वितरण के स्थान पर प्वॉयसन वितरण का प्रयोग करेंगे?
    When does a binomial distribution tend to become a Poisson distribution? In which case would you apply a Poisson distribution in place of binomial distribution?
18. प्रसामान्य आवृत्ति वितरण की विशेषताएं बताइये। सांख्यिकी में इस वितरण को प्रमुख स्थान क्यों प्राप्त है?
    List the chief properties of the normal distribution. AWhy is this distribution given a central place in statistics?
19. प्रसामान्य एवं द्विपद बंटनों में भेद कीजिए तथा सांख्यिकीय निष्कर्ष के लिए प्रसामान्य बंटन के महत्व की विवेचना कीजिए।
    Distinguish between normal and binomial distribution and discuss the importance of normal distribution for statistical inference.
20. प्रसामान्य वितरण द्विपद वितरण से किस प्रकार भिन्न है? प्रसामान्य वितरण की प्रमुख विशेषताएं क्या हैं? दैव निदर्शन अनुसन्धान में वे किस प्रकार उपयोगी हैं?
    How does a normal distribution differ from a binomial distribution? What are the important properties of a normal distribution? How are they useful in random sampling investigation?
21. सांख्यिकी में प्रसामान्य वितरण की क्या महत्ता है? आर्थिक विश्लेषणों में भी इसकी उपयोगिता बताइये।
    What is the significance of normal distribution in statistics? Give also its utility in economic analysis.
22. प्रसामान्य एवं द्विपद वितरण की महत्वपूर्ण विशेषताओं की तुलना कीजिए और उनका निदर्शन सिद्धांत में महत्व स्पष्ट कीजिए।
    Compare the important properties of the normal and binomial distributions and point out their significance in sampling theory.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. सैद्धान्तिक वितरण क्या है?
   What is Theoretical Distribution.
2. द्विपद वितरण क्या है?
   What is Binomial distribution.
3. प्वॉयसन वितरण की परिभाषा दीजिये।
   Define Poission distribution.
4. प्रसामान्य आवृत्ति वितरण की विशेषतायें बताइये।
   List the chief properties of the normal distribution.
5. प्वॉयसन वितरण के प्रमुख लक्षणों को स्पष्ट कीजिये।
   Explain the main characteristics of poission distribution.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. प्वॉयसन विधि सिमोन डेनिस द्वारा कब बनाई गयी-
   (अ) 1820 (ब) 1825 (स) 1836 (द) 1837
2. प्रसामान्य वितरण विधि में औसत से दायें का क्षेत्र क्या है-
   (अ) +.5 (ब) –.5 (स) 1 (द) कोई नहीं
3. सैद्धान्तिक आवृत्ति वितरण के कितने प्रकार हैं-
   (अ) 2 (ब) 4 (स) 3 (द) 5

4. द्विपद वितरण में सभी पदों का योग होता है-

(अ) 0 (ब) 1 (स) 100 (द) 50

उत्तर- 1. (द), 2. (अ), 3. (स), 4. (ब)

2.11 व्यावहारिक प्रश्न

1. किसी उद्योग में व्यावसायिक बीमारी के किसी कर्मचारी को लगाने की सम्भावना 20% है। 6 कर्मचारियों में से 4 अथवा अधिक से इस बीमारी से ग्रसित होने की प्रायिकता क्या है?

The incidence of occupational disease in an industry is such that the workman have a 20% chances of suffering from it. What is the probability that out of six workmen 4 or more will contact disease?

[उत्तर : $\frac{240 + 24 + 1}{15,625} = 0.1696$]

2. पशुओं में किसी विशेष बीमारी के फैलने की सामान्य दर 40% है। 6 पशुओं को नया इन्जेक्शन लगने पर उनमें से किसी पर भी बीमारी का असर नहीं हुआ। अवलोकित परिणाम की प्रायिकता ज्ञात कीजिए।

The normal rate of infection of a certain disease in animal is known to be 40%. In an experiment with new injection 6 animals, none caught infection. Find the probability of the observed result.

[उत्तर : $\left(\frac{3}{5}\right)^6 = \frac{729}{15,625} = .0467$]

3. एक बम के निशाने पर गिरने की प्रायिकता $\frac{1}{5}$ है। एक पुल को नष्ट करने के लिए दो बम पर्याप्त हैं। यदि पुल पर 6 बम डाले जायें तो उसके नष्ट होने की प्रायिकता ज्ञात कीजिए।

The probability of a bomb hitting a target is $\frac{1}{5}$. Two bombs are enough to destroy a bridge. If six bombs are aimed at the bridge, find the probability that the bridge is destroyed.

[उत्तर : $1 - \left\{\left(\frac{5}{5}\right)^6 + 6\left(\frac{5}{6}\right)^5 \left(\frac{1}{6}\right)\right\} = 1 - (.2621 + .3952) = .3447$]

4. छः पासे 729 बार फेंके गए। आप कितनी बार कम-से-कम तीन पासों द्वारा पांच छः दिखाने की उम्मीद करते हैं?

Six dice are thrown 729 times. How many times do you expect at least three dice to show a five or six?

[उत्तर : 223]

5. आठ सिक्के उछाले जाते हैं और यह प्रयोग 100 बार दोहराया जाता है। द्विपद बंटन ज्ञात कीजिए और उसका माध्य तथा प्रमाप विचलन निकालिए।

Eight coins are tossed and this experiment is repeated 100 times. Find binomial distribution and calculate its mean and standard deviation.

[उत्तर : $p(x) = 100.\ {}^{10}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x}$, $x = 0, 1, 2, \ldots\ldots, 10.$

माध्य $= 10 \times \frac{1}{2} = 5$, प्रमाप विचलन $= \sqrt{10 \times \frac{1}{2} \times \frac{1}{2}} = 1.6$]

NOTES

6. यह मानकर कि किसी नगर के 50% व्यक्ति अखबार पढ़ते हैं और यह भी मानकर कि 1,024 खोजकर्ताओं द्वारा 10 व्यक्तियों के विषय में जानकारी की जाती है कि वह अखबार पढ़ते हैं कि नहीं। कितने अन्वेषक प्रत्याशित हैं जो यह सूचना देंगे कि 3 या उससे कम व्यक्ति अखबार पढ़ते हैं?

   Assuming that 50% of a population in a town reads newspapers and further assuming that 1,024 investigators each take 10 individuals to find out if they read newspapers, how many investigators would you expect to report that three or less people read newspapers.

   [उत्तर : 176 अन्वेषक]

7. यह मानते हुए कि आधी जनसंख्या शाकाहारी होने की प्रायिकता $\frac{1}{2}$ है और यह भी मानते हुए कि 100 अन्वेषकों में से प्रत्येक को 10 व्यक्तियों का प्रतिदर्श लेकर उनसे यह पूछना है कि वे शाकाहारी या नहीं। आप कितने अन्वेषक अनुमानित करते हैं जो यह रिपोर्ट करेंगे कि तीन या इससे कम व्यक्ति शाकाहारी हैं?

   Assuming that half the population is vegetarian so that chance of an individual being a vegetarian is $\frac{1}{2}$ and assuming that out of 100 investigators each takes a sample of 10 individuals to see whether they are vegetarian. How many investigators would you expect to report that three or less people are vegetarian?

   [उत्तर : 17]

8. किसी सेना की टुकड़ी में 3/5 जवान विवाहित हैं और शेष 2/5 अविवाहित। एक पंक्ति में यदि 5 जवान हों तो उनमें 0, 1, 2, 3, 4, 5 विवाहित जवान होने की प्रायिकता क्या है?

   In an army battalion 3/5 of the soldiers are married and the remainder 2/5 unmarried. What is the probability of being 0,1,2,3,4,5 married soldiers in a row of soldiers?

   [उत्तर : $\left(\frac{2}{5}\right)^5, {}^5C_1 \left(\frac{2}{5}\right)^4 \left(\frac{3}{5}\right), {}^5C_2 \left(\frac{2}{5}\right)^3 \left(\frac{3}{5}\right)^2, {}^5C_3 \left(\frac{2}{5}\right)^2 \left(\frac{3}{5}\right)^3 {}^5C_4 \left(\frac{2}{5}\right) \left(\frac{3}{5}\right)^4, \left(\frac{2}{5}\right)^5$]

9. सैद्धांतिक द्विपद बंटन $128\left(\frac{1}{2}+\frac{1}{2}\right)^5$ की सभी प्रत्याशित आवृत्तियों को ज्ञात कीजिए।

   Find all the expected frequencies of a theoritical binomial distribution $128\left(\frac{1}{2}+\frac{1}{2}\right)^5$.

   [उत्तर : 4, 20, 40, 40, 20, 4]

10. 4 पासे एक साथ 64 बार फेंके गये। बिन्दु 3 या 4 को सफलता तथा बिन्दु 1 या 2 को असफलता माना गया। बिन्दु 5 या 6 आने पर प्रयास नहीं माना गया। परिणाम निम्न प्रकार पाये गये :

    Four dice were thrown 64 times, Spot 3 or 4 was considered a success and spot 1 or 2 a failure. The outcome of a 5 to 6, was not considered a trail. The results were :

सफलताओं की संख्या **(Number of successes)**	:	0	1	2	3	4
आवृत्ति **(Frequency)**	:	0	5	13	22	24

    अवलोकित आवृत्ति वितरण तथा सैद्धांतिक द्विपद वितरण के माध्य और प्रमाप विचलन ज्ञात कीजिए।

    Find the mean and standard deviation of the observed frequency distribution and the theoretical binomial distribution.

    [उत्तर : अवलोकित आवृत्ति वितरण के लिए $\bar{x} = 3.02, \sigma = .94$ द्विपद वितरण के लिए,
    $\bar{x} = np = 4 \times \frac{1}{2} = 2$ तथा $\sigma = \sqrt{npq} = \sqrt{4 \times \frac{1}{2} \times \frac{1}{2}} = 1$]

11. 8 सिक्कों के एक समूह को 256 बार उछाला गया और प्रत्येक बार आने वाले शीर्षों की संख्या को लिखा गया। प्राप्त परिणाम निम्न तालिका में दिये गये हैं। शीर्षों की माध्य संख्या अनुमानित कीजिए तथा इस अनुमान का प्रयोग करते हुए प्रत्याशित आवृत्तियां परिकलित कीजिए यदि द्विपद नियम लागू होता है।

A set of 8 coins is thrown 256 times and the number of heads appearing in each throw is recorded. The results obtained are given in the following table. Estimate the mean number of heads and then calculate the expected frequencies using this estimate if the binomial law holds :

शीर्षों की संख्या (No. of heads)	0	1	2	3	4	5	6	7	8
उछालों की संख्या (No. of throws)	2	6	38	52	59	56	32	10	6

[उत्तर : $\bar{x} = \frac{\Sigma fx}{fx} = \frac{1{,}024}{256} = 4; p = \frac{\bar{x}}{n} = \frac{4}{8} = \frac{1}{2}$

प्रत्याशित आवृत्तियां, f_e : 1 8 28 56 70 56 28 8 1]

12. एक बगीचे में 64 क्यारियां हैं एवं प्रत्येक क्यारी में एक विशेष प्रकार के फूल के 3 बीज बोये जाते हैं। एक नीले फूल की प्रायिकता 1/4 है। 3,2,1 एवं 0 नीले फूलों वाली क्यारियों की संख्या ज्ञात कीजिए।

There are 64 beds in a garden and 3 seeds of particular type of flower are sown in each bed. The probability of a flower being blue is 1/4. Find the number of beds with 3, 2, 1 and 0 blue flowers.

[उत्तर : 27, 27, 9 तथा 1]

13. 4 चूहों के 104 समूहों में 0, 1, 2, 3, 4 मादा चूहों की संख्या निम्न प्रकार है :

मादा चूहों की संख्या,	x :	0	1	2	3	4
समूहों की संख्या,	f :	8	28	34	24	10

किसी चूहे के मादा चूहा होने की प्रायिकता अनुमानित कीजिए तथा प्रत्याशित आवृत्तियां ज्ञात कीजिए।

In 104 litters of 4 mice, the number of letters which contain 0, 1, 2, 3, 4 female mice is as follows :

No. of female mice, x :	0	1	2	3	4
No. of litters, f :	8	28	34	24	10

Estimate the probability of a mouse being a female and find the expected frequencies.

[उत्तर : $p = \frac{1}{2}$, f_e : 6.5, 26, 39, 26, 6.5]

14. निम्नलिखित के लिए द्विपद वितरण का माध्य तथा बहुलक बताइये यदि।

Obtain the mean and mode for the following binomial distribution if,

(i) n = 99, p = 0.6 (ii) n = 60, p = 1/6 (iii) n = 8, p = 1/5

[उत्तर : (i) माध्य = 59.4, बहुलक = 60 तथा 59, (ii) माध्य = 10 = बहुलक

(iii) माध्य = $\frac{8}{5}$, बहुलक = 1]

15. एक द्विपद बंटन में समान्तर माध्य 3 एवं प्रसरण 2 है तो शेष अचर मूल्य $(\mu_1, \mu_2, \mu_3, \mu_4, \beta_1, \beta_2)$ ज्ञात कीजिए।

In a binomial distribution mean is 3 and variance is 2, find the remaining constants $(\mu_1, \mu_2, \mu_3, \mu_4, \beta_1, \beta_2)$.

[उत्तर : $\frac{npq}{np} = \frac{2}{3}, p = 1 - q = \frac{1}{3}, n = \frac{np}{p} = 9, \mu_1 = 0, \mu = \sigma^2$

$= 2, \mu_3 = \frac{2}{3}, \mu_4 = \frac{34}{3}, \beta_1 = \frac{1}{18}, \beta_2 = \frac{17}{6}$]

NOTES

16. 4 सिक्के 160 बार उछाले गये और निम्नांकित अवलोकित आवृत्ति वितरण प्राप्त हुआ :

शीर्षों की संख्या :	0	1	2	3	4
4 सिक्कों की उछालों की संख्या :	17	32	54	31	6

यह मानते हुए कि सिक्के सुडौल हैं शीर्षों की संख्याओं की प्रत्याशित आवृत्तियां ज्ञात कीजिए और आसंजन-सौष्ठव का परीक्षण भी कीजिए।

4 coins are tossed 160 times and following observed frequency distribution :

No. of heads :	0	1	2	3	4
No. of tosses :	17	32	54	31	6

Assuming that the coins are unbaised calculate the expected frequencies for the number of heads and test the goodness of fit.

[उत्तर : प्रत्याशित आवृत्तियां : 10, 40, 60, 40, 10, $x^2 = 12.7$;5% सार्थकता स्तर पर आसंजन सही है]

17. एकल प्रयास में 5 सिक्कों के एक समूह को 96, पर उछाला गया। निम्न तालिका में चित्त की संख्या का अवलोकित आवृत्ति वितरण दिया हुआ है। इस मान्यता पर कि सिक्के अनभिनत थे, आशानुरुप आवृत्ति परिकलित करें। इस मान्यता की सत्यता का परीक्षण करने के लिए x^2 परीक्षण का उपयोग करें (5 सिक्कों को एक साथ उछालने पर सिक्कों की संख्या जिन पर चित्त आता है = x, तथा उछालों की संख्या = f):

A set of 5 coins was tossed 96 times. The following table gives the observed frequency distribution of the number of heads in a single trail. Calculate the expected frequencies on the assumption that the coins were unbaised. Use X^2 test to examine the correctness of this assumption (x = number of coins that show head in a single trail, f = No. of throws).

x :	5	4	3	2	1	0
f :	7	19	35	24	8	3

[उत्तर : प्रत्याशित आवृत्तियां, f_e = 3,15,30,30,15, 3 x^2 परिकलित 11.62, $x2_{(.05,5)} = 11.07$;
मान्यता सही नहीं है]

18. 7 सिक्कों को 128 बार उछालने पर उनके चित्त गिरने का अवलोकित आवृत्ति वितरण निम्न प्रकार है। यह मानते हुए कि सिक्के सुडौल हैं द्विपद बंटन का आसंजन कीजिए। आसंजित बंटन के माध्य तथा प्रमाप विचलन क्या हैं?

The observed frequency distribution of 128 throws of 7 coins, according to the number of heads. Fit a binomial distribution under the assumption that the coins are unbiased. What are the mean and standard deviation of the fitted distribution?

चित्तों की संख्या (No. of heads) :	0	1	2	3	4	5	6	7
उछालों की संख्या (No. of throws) :	7	6	19	35	30	23	7	1

[उत्तर : प्रत्याशित आवृत्ति : 1, 7, 21, 35, 35, 21, 7, 1 : माध्य = 3.5, प्रमाप विचलन = 32]

19. निम्नलिखित पर संक्षिप्त टिप्पणी लिखिए-

Write short note on the following :

(i) द्विपद बंटन (Binomial Distribution)

(ii) द्विपद बंटन की विशेषताएं (Characteristics of Binomial Distribution)

(iii) द्विपद बंटन के अचर (Constants of Binomial distribution)

(iv) द्विपद बंटन के आसंजन की श्रेष्ठता का परीक्षण
(Test of goodness of fit of a Binomial distribution)

20. निम्नलिखित कथन की व्याख्या कीजिए-

किसी द्विपद वितरण का माध्य 8 तथा प्रमाप 3 है।

Criticise the following statement :

The mean of a binomial distribution 8 and standard deviation is 3.

21. यदि एक ढेर में दोषपूर्ण वस्तु होने का अनुपात 4% हो तो 10 के एक प्रतिदर्श में 2 से अधिक दोषपूर्ण न होने की प्रायिकता क्या है ? (दिया है $e^{-4} = .6703$)

If the proportion of defective items in a bulk is 4 percent, find the probability of not more that two defective in a sample of 10. (It is known that $e^{-.4} = 6703$)

[उत्तर : $P_{(0)} + p_{(1)} + p_{(2)} = 0.6703 + 0.2681 + 0.536 = 0.992$]

22. पिछला अनुभव यह बताता है कि 1 माह में औसतन 4 औद्योगिक दुर्घटनाएं होती हैं। किसी विशेष माह में 4 से कम दुर्घटनाएं होने की प्रायिकता ज्ञात करें (दिया है : $e^{-4} = .018$)

The past experience shows that there occur on an average 4 industrial accidents per month. Find the probability of the occurence of less than 4 industrial accidents in a certain month. (given : $e^{-4} = 0.18$)

[उत्तर : $p_{(0)} + p_{(1)} + p_{(2)} + p_{(4)} = \sum_{x=0}^{x=3} e^{-4} 4^x$]

23. निर्मित वस्तुओं का निरीक्षण करके एक प्रति इकाई 2 त्रुटियां दिखाई देती हैं। किसी निर्मित वस्तु में (i) कोई दोष न हो तथा (ii) दोष हों की प्रायिकता ज्ञात कीजिए। (दिया है : $e^{-2} = .135$)

In an inspection of manufactured articles 2 defects per unit are found. Calculate the probability finding an article (i) without any defect and (ii) with 4 defects.

[उत्तर : (i) 0.135, (ii) .09]

24. किसी दुकान पर औसत रूप से प्रति मिनट 5 ग्राहक आते हैं। प्वॉयसन वितरण की सहायता से किसी विशिष्ट एक मिनट में 6 ग्राहक आने की प्रायिकता ज्ञात कीजिए।

The average number of customers per minute arriving at a shop is 5. Using Poisson distribution find the probability that during one particular minute exactly 6 customers will arrive.

[उत्तर : 0.1454]

25. किसी पुस्तक के 50 पृष्ठों के संशोधन के पश्चात् यह ज्ञात हुआ कि सामान्यतः 5 पृष्ठों में 3 गलतियां हैं। प्वॉयसन बंटन के आधार पर यह अनुमान लगायें कि 1,000 पृष्ठों की पुस्तक में 0,1, 2, 3, गलतियां वाले पृष्ठों की संख्या क्या है ? (दिया है : $e^{-6} = .5488$)

After correcting the errors of the first 50 pages of a book, it is found that on the average there are 3 errors per 5 pages. Use Poisson distribution to estimate the number of pages with 0, 1, 2, 3 errors in the whole book of 1,000 pages (Given $e^{-6} = .5488$)

[उत्तर : 549, 329, 99, 20]

26. एक उत्पादक इस बात की घोषणा करता है कि उसके उत्पादनों का 5% दोषपूर्ण है। वह 100-100 उत्पादनों की पेटियों में विक्रय करता है और गारण्टी देता है कि पेटी में 3 से अधिक दोषपूर्ण उत्पादन नहीं होगा। इसकी क्या सम्भावना है कि एक पेटी के उत्पादन उसकी गारण्टी की पूर्ति नहीं कर सकेंगे ?

A producer claims that 5% of his product is defective. He sells his product in boxes of 100 and gurantees that not more than 34 items will be defective. What is the probability that a box will fail to meet the gurantee?

27. टेलीफोन कॉलों की संख्या जो एक टेलीफोन ऑपरेटर 8 बजे से 8 बजकर 5 मिनट के बीच में प्राप्त करता है, m = 3 के प्वॉयसन बंटन का पालन करती है। इस बात की क्या संभावना है कि आगामी काल उसी अवधि में उसे कोई फोन कॉल नहीं मिलेगी?

The number of telephone calls an operator receives from 8 to 5 minutes past 8 follows a Poisson distribution with m=3. What is the probability that the operator will not receive a telephone call in the same time interval tomorrow?

[उत्तर : .04979]

28. लेन्स बनाने वाली किसी फैक्टरी में दोषयुक्त लेन्स बनाने की सम्भावना $\frac{1}{100}$ है। प्रति पैकिट 10 लेन्स रखे जाते हैं। प्वॉयसन वितरण का उपयोग करते हुए यह ज्ञात कीजिए कि 20,000 पैकिटों में से कितने पैकिटों में (i) एक भी दोषयुक्त लेन्स नहीं होगा, (ii) एक दोषयुक्त लेन्स होगा, (iii) 2 दोषयुक्त लेन्स होंगे, (iii) 2 दोषयुक्त लेन्स होंगे, (iv) 3 दोषयुक्त लेन्स होंगे तथा (v) 4 दोयुक्त लेन्स होंगे? (दिया है $-e^{.02} = .9802$]

In a certain factory turning out optical lenses there is a small chance $\frac{1}{100}$ for any lens to be defective. The lenses are supplied in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing no defective, one defective, two defective, three defective and four defective lenses respectively in a consignment of 20,000 packets. (Given $e^{-.02} = .9802$).

[उत्तर : (i) 19,604, (ii) 392, (iii) 4, (iv) 0, (v) 0]

29. फाउण्टेन पेन बनाने के कारखाने में जहां 0.5 प्रतिशत दोषपूर्ण पेन बनते हैं, सौ-सौ पेन डिब्बों में रखे जाते हैं। ऐसे डिब्बों का क्या प्रतिशत होगा जिनमें (i) एक भी दोषपूर्ण पेन न हो, (ii) कम-से-कम एक दोषपूर्ण पेन हो और 2 या 2 से अधिक दोषपूर्ण पेन हों? (दिया है $e^{-5} = .6065$)

In factory manufacturing fountain pen, the chance having a defective pen is 0.5 percent, 100 pens are kept in a box. What is the percentage of boxes in which (i) there is no defective pen, (ii) there is at least one defective pen and (iii) there are 2 or more than 2 defective pens? (Given $e^{-5} = .6065$)

[उत्तर : (i) 60.65%, (ii) 39.35%, (iii) 9.025%]

30. एक कार्यालय प्रबंधक अपने टाइपिस्ट का कार्य तभी स्वीकार करता है जबकि उसमें कोई अशुद्धि न हो। टाइपिस्ट को प्रति दिन औसतन 20 पत्र टाइप करने पड़ते हैं जिनमें से प्रत्येक में लगभग 200 शब्द होते हैं। प्वॉयसन वितरण का प्रयोग करके उसके द्वारा अशुद्धि किए जाने की प्रायिकता ज्ञात कीजिए यदि (i) उसके द्वारा प्रस्तुत पत्रों में से 1% से कम अस्वीकृत कर दिये जाते हैं। (ii) 90% दिनों में उसके द्वारा प्रस्तुत पूरा कार्य स्वीकार कर लिया जाता है।

A manager accepts the work submitted by his typist only when there is no mistake in the work. The typist has to type on an average 20 letters per day of about 200 words each. Use Poisson distribution to find the chance of his making a mistake if (ii) less than 1% of the letters submitted by him are rejected and (ii) on 90% of days all the work submitted by him is accepted.

[उत्तर : (i) $p_{(0)} = 0.99 \Rightarrow m = .01014$ तथा $p = .00005$ (ii) $p_{(0)} = 0.9 = m = .1054$ तथा p = .00026]

31. एक पुस्तक में प्रति पृष्ठ निम्न त्रुटियां पायी गयीं। इन आंकड़ों के आधार पर एक प्वॉयसन वितरण का आसंजन करें :

प्रति पृष्ठ त्रुटियां :	0	1	2	3	4
पृष्ठों की संख्या :	200	90	20	10	0

(दिया है : $e^{-5} = 6065$)

The following mistakes per page were observed in a book. Fit a Poisson distribution to the data :

No. of mistakes per page :	0	1	2	3	4
No. of pages :	200	90	20	10	0

(Given : $e^{-5} = .6065$)

32. एक टाइपिस्ट के द्वारा 100 पृष्ठों की टंकण में निम्न प्रकार गलतियां कीं। प्वॉयसन वितरण का आसंजन करो तथा प्रत्याशित आवृत्तियां ज्ञात कीजिए :

प्रति पृष्ठ त्रुटियां :	0	1	2	3	4	5
पृष्ठों की संख्या :	42	35	14	6	4	1

(दिया है : $e^{-.1} = .3679$)

A typist commits the following number of mistakes per page in typing 100 pages. Fit a Poisson distribution and calculate theoretical frequencies :

Mistakes per page : 0 1 2 3 4 5

No. of pages : 42 35 14 6 4 1

(Given : $e^{-.1} = .3679$)

[उत्तर : प्रत्याशित आवृत्तियां : 36.8, 36.8, 18.4, 6.1, 1.5, 0.4]

33. निम्न आंकड़े किसी हाई कोर्ट में पिछले 100 वर्षों में हुए जजों के रिक्त स्थानों को व्यक्त करते हैं। प्वॉयसन वितरण का आसंजन कीजिए :

Below are given the number of vacancies of Judges occuring in a high court over a period of 100 years. Fit a Poisson distribution :

रिक्तियां (Vacancies) :	0	1	2	3
आवृत्ति (Frequency) :	56	32	8	4

[उत्तर : m = .5488, प्रत्याशित आवृत्तियां : 54.88, 32.93, 9.88, 9.31]

34. निम्न अवलोकनों से प्वॉयसन वंटन का अन्वायोजन कीजिए और सैद्धांतिक आवृत्तियों की गणना कीजिए : ($e^{-5} = .6065$)

Fut a Poisson distribution to the following data and also calculate theoretical frequencies :

मृत्यु (Death) :	0	1	2	3	4
आवृत्ति (Frequency) :	122	60	15	2	1

[उत्तर : m = 0.5, सैद्धांतिक आवृत्तियां : 121.2, 60.65, 15.16, 2.53, 0.32]

35. प्वॉयसन वितरण के माध्य एवं प्रसरण ज्ञात कीजिए।

Find the mean and variance of Poisson distribution.

36. दिखाइये कि निम्न कथन में कोई असंगति है : "प्वॉयसन वितरण का समान्तर माध्य 16 है और उसका प्रमाप विचलन 9 है।"

Show that there is a inconsistency in the following statement : "The mean of a Poisson distribution is 16 and the standard deviation is 9."

37. निम्न कथन की विवेचना कीजिए :

Comment on the following statement :

प्वॉयसन वितरण के लिए माध्य 8 और प्रसरण 7 है।

"For a Poisson distribution mean is 8 and variance is 7."

NOTES

38. एक नये कारखाने में 1,000 बल्ब लगाये जाते हैं जिनका औसत जीवन 120 दिन है। उनका जीवन प्रसामान्य बंटन के अनुरूप है जिसका प्रमाप विचलन 20 दिन है।

(i) इनमें से कितने 90 दिन से आगे समाप्त हो जाएंगे?

(ii) यदि सभी के स्थान पर नये बल्ब लगाने की योजना हो तो उनके बदलने से पहले 10% से अधिक कितने बदले जा सकेंगे?

1,000 light bulbs with a mean life of 120 days are installed in a new factory. Their length of life is normally distributed with standard deviation 20 days. (i) How many will expire in less than 90 days? (ii) It is decided to replace all the bulbs together what interval should be allowed between replacement if not more than 10% should expire before replacement.

[उत्तर : (i) 67, (ii) 94 दिन]

39. एक प्रसामान्य बंटन में 7% वस्तुएं 35 से कम है। 89% वस्तुएं 63 से कम हैं। इस बंटन के $\bar{X}$ तथा σ क्या हैं?

In a distribution exactly normal 7% of the items are under 35 and 89% are under 63. What are the mean and S.D. of the distribution?

[उत्तर : $\bar{X} = 50.3$ and $S.D. = 10.33$]

40. एक स्टील कारखाने में कार्यरत 1,000 श्रमिकों की औसत लम्बाई 67 इंच है। इसका प्रमाप विचलन, σ 5 इंच है। उस कारखाने में 72 इंच से अधिक लम्बे श्रमिकों की संख्या क्या होगी?

दिया है : 0 और 1σ के बीच क्षेत्रफल = .34134 तथा 0 और 1.5σ के बीच = .43319

The mean height of the 1,000 workers in a steel plant is 67 inches with a standard deviation of 5 inches. How many workers are expected to be above 72 inches in that plant?

Given : The area under one side of the normal curve for 1σ is .34134 and for 1.5σ is .43319.

[उत्तर : 159]

41. यदि एक प्रसामान्य वितरण में 500 छात्रों की औसत ऊंचाई 65 इंच तथा प्रमाप विचलन, σ 5 इंच हो, तो कितने छात्रों की ऊंचाई

(i) 70 इंच से अधिक होगी। (ii) 60 और 70 इंच के मध्य होगी।

If the heights of 500 students are normally distribution with mean 65 inches and standard deviation 5 inches. How many students have heights.

(i) greater than 70 inches. (ii) between 60 and 70 inches.

[उत्तर : (i) 11, (ii) 68]

42. मुम्बई म्युसिपल कॉरपोरेशन ने सड़कों पर 2,000 नये बल्ब लगाए। यदि इनके चलन की औसत आयु 1,000 घण्टे हो और σ = २०० घण्टे, तो प्रथम 700 जलने वाले घण्टों में कितने बल्ब बेकार हो जाएंगे? निम्न तालिका में सामान्य वक्र-रेखा सम्बन्धी जानकारी दी गई है:

$\frac{X-\bar{X}}{\sigma}$:	1	1.25	1.50
प्रायिकात :	0.159	0.106	0.067

The Mumbai Municipal Corporation installed 2,000 bulbs in the streets of Mumbai. If these bulbs have an average life of one thousand burning hours with a S.D. of 200 hours, what number of bulbs might be expected to fail in the first 700 burning hours? the table of the area of normal curve at selected values is as follows :

$\frac{X-\bar{X}}{\sigma}$:	1	1.25	1.50
Probability :	0.159	0.106	0.067

[उत्तर : 134]

43. यदि जवानों की औसत ऊंचाई 68.22 इंच हो तथा प्रसरण $\sigma^2 = 1.08$ इंच हो तो एक रेजीमेण्ट के 1,000 जवानों में 6 फीट से अधिक ऊंचाई होने की संभावना क्या है ?

(प्रसामान्य वितरण में 1.15 पर कोटि के दायीं ओर का क्षेत्रफल .6221 है।)

If the mean height of soldiers is 68.22 inches with a variance of 1.08 inches, how many soldiers in a regiment of 1,000 can be expected to be over 6 feet all? [In a normal distribution are to the right of an ordinate at 1.15 = 6221.]

[उत्तर : 125]

44. 10,000 व्यक्तियों के समूह की प्रति माह आय का वितरण प्रसामान्य है जिसमें माध्य 750 रु. प्रतिमाह और प्रमाप विचलन 50 रु. है। दिखाइये कि इस समूह के 95% व्यक्तियों की आय 668 रु. से अधिक तथा 5% व्यक्तियों की आय 832 रु. से अधिक होगी। सर्वाधिक धनी 100 व्यक्तियों ममें सबसे कम आय कितनी होगी ?

The income per month of a group of 10,000 persons was found to be normally distributed with mean Rs. 75 with a standard deviation of Rs. 50. Show that of this group 95% had income mean Rs. 750 with a standard deviation of Rs. 50. Show that of this group 95% had income exceeding Rs. 668 and only 5% had income exceeding Rs. 832. What was the lowest income among the richest 100?

[उत्तर : 866.5 रु.]

45. एक शहर में 1,000 लघु व्यावसायिक फर्मों के सांख्यिकीय अनुसंधान में यह पाया गया कि उनकी मासिक औसत बिक्री 8,000 रु. थी तथा प्रमाप विचलन 2,000 रु. था। यह कल्पना करते हुए कि बिक्री प्रसामान्य रूप से वितरित है, उन फर्मों की संख्या ज्ञात कीजिए :

(i) जिनकी औसत मासिक बिक्री 6,000 रु. से कम थी, तथा

(ii) जिनकी औसत मासिक बिक्री 7,000 रु. से 9,000 रु. के बीच थी।

In a statistical survey of 1,000 small business firms in a city, it was found that their monthly average sales amounted to Rs. 8,000 with a standard deviation of Rs. 2,000. Assuming that the sales are normally distributed, estimate :

(i) the number of firms whose monthly average sales were less than Rs. 6,000 and
(ii) the number of firms whose monthly average sales were between Rs. 7,000 and Rs. 9,000.

प्रसामान्य वक्र के नीचे का क्षेत्रफल
(Area Under the Normal Curve)

$Z = \frac{x-\mu}{\sigma}$:	0.1	0.5	1.0	1.5	2.0
A :	0.0398	0.1915	0.3413	0.4332	0.4772

[उत्तर : (i) 158.7, (ii) 383]

46. एक बिक्री-कर अधिकारी ने सूचित किया है कि एक वर्ष में उसको जिन 500 व्यवसायों से निपटना पड़ता है उनकी औसत बिक्री 3,60,000 रु. वार्षिक है तथा उनका प्रमाप विचलन 1,00,000 रु. है। यह मानकर कि इन व्यवसायों की बिक्री प्रसामान्य वितरण के अनुसार है, ज्ञात कीजिए :

(i) उन व्यवसायों की संख्या जिनकी बिक्री 4,00,000 रु. से अधिक है।

(ii) उन व्यवसायों की संख्या जिनकी बिक्री 30,0000 रु. से 4,00,000 रु. से अधिक होने की सम्भावना है।

NOTES

(iii) दैव रूप से चुने गये किसी व्यवसाय की बिक्री 3,00,000 रु. से अधिक होने की सम्भावना।

A Sales-tax officer has reported that the average sales of the 500 business that he has to deal a Sales-tax officer has reported that the average sales of the 500 business that he has to deal with during a year amount to Rs. 3,60,000 with a standard deviation of Rs. 1,00,000. Assuming that the sales in these business are normally distributed, find :

(i) the number of business the sales of which are over Rs. 4,00,000.

(ii) The percentage of business the sales of which are over Rs. 4,00,000.

(iii) The probability that the sales of business selected at random will be over Rs, 3,00,000.

$\frac{x-\mu}{\sigma}$:	0.25	0.40	0.50	0.60
A :	0.0987	0.1554	0.1915	0.2257

[उत्तर : (i) 158.7, (ii) 383]

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय-3 निदर्शन के सिद्धान्त एवं सार्थकता परीक्षण
(THEORY OF SAMPLING AND TEST OF SIGNIFICATION)

इकाई की रूपरेखा

3.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि-

1. निदर्शन एवं समग्र प्रणाली के अर्थ को समझ सकेंगे।
2. निदर्शन चुनने की विभिन्न पद्धतियों का ज्ञान होगा।
3. छोटे न्यादर्श की सार्थकता की जाँच कर सकेंगे।

3.1 प्रस्तावना

किसी विषय के सम्बन्ध में दो प्रकार से अनुसंधान किया जा सकता है-

(1) समग्र या संगणना प्रणाली (Census Method)

(2) निदर्शन या न्यादर्श प्रणाली (Sample Method)

दोनों प्रणालियों का विस्तृत विवेचन करने से पूर्व यह बतला देना उपयुक्त होगा कि सांख्यिकी में समस्त अनुसंधान के क्षेत्र को समग्र (Population or Universe) कहते हैं। किसी भी समग्र का अनुसंधान करने के लिये या तो समग्र के प्रत्येक मद (item) के सम्बन्ध में सूचना प्राप्त की जाती है जिसे संगणना प्रणाली कहते हैं या समग्र के सभी मदों में से किसी पद्धति (जिसका विवेचन किया जायेगा) से प्रतिनिधि मद चुन लिये जाते हैं जिसे निदर्शन या न्यादर्श प्रणाली कहते हैं। अब हम दोनों प्रणालियों का विस्तृत विवेचन करेंगे-

3.2 समग्र या संगणना प्रणाली (Census Method)

ऊपर बताया जा चुका है कि जब अनुसंधान के विषय में सम्बन्धित समग्र (Universe) की प्रत्येक इकाई की जाँच की जाए तो यह समग्र या संगणना अनुसंधान कहलाता है। इस प्रणाली में अनुसंधानकर्ता द्वारा सामग्री संकलन के समय समस्त समूह की जाँच की जाती है और सभी इकाईयों का सविस्तार अध्ययन किया जाता है। इस प्रकार के अनुसंधान की आवश्यकता उस समय पड़ती है जब किसी विषय के सम्बन्ध में गहन अध्ययन करना अथवा उच्चस्तरीय शुद्धता एवं विश्वसनीयता प्राप्त करना हो। इस प्रणाली का उपयोग विशेषतः जनगणना, उत्पादन तथा आयात-निर्यात, आदि की गणना के लिए किया जाता है। यदि हम किसी महाविद्यालय के सभी विद्यार्थियों की आर्थिक स्थिति का सर्वेक्षण करने के लिए सभी के आय-व्यय की जाँच करें तो यह समग्र तथा संगणना अनुसंधान ही कहलाता है। इस प्रकार का अनुसंधान शासन अथवा ऐसी संस्था ही कर सकती है जिसकी आर्थिक स्थिति सुदृढ़ हो। अधिकांशतः निदर्शन प्रणाली (Sample Method) का ही प्रयोग किया जाता है।

संक्षेप में हम उन परिस्थितियों का वर्णन करना भी उचित समझते हैं जिनमें यह प्रणाली उपयुक्त है और उसका प्रयोग करना सर्वथा अनुकूल है। इन परिस्थितियों को इस प्रकार वर्गीकृत किया जा सकता है-

(1) **अनुसंधान का सीमित क्षेत्र** (Limited field of Investigation)- यह प्रणाली अनुसंधान का क्षेत्र सीमित होने पर उपयोग में लाई जा सकती है। यदि क्षेत्र विस्तृत हो तो इस प्रणाली का उपयोग करने से अनुसंधान में काफी धन, समय और व्यक्तियों की आवश्यकता पड़ेगी।

(2) **अनुसंधान की कम परिवर्तनशील प्रकृति** (Less Dynamic Nature and Investigation)- इस प्रणाली का उपयोग करते समय इस बात पर विचार कर लेना भी उचित होगा कि हम जिस समस्या के विषय में अनुसंधान करना चाहते हैं उससे संबंधित आँकड़ों की प्रकृति परिवर्तनशील तो नहीं है। यदि ऐसा है तो आँकड़ों में परिवर्तन के बाद संकलित आँकड़ों का कोई महत्व शेष नहीं रहेगा और संकलन में किया गया व्यय व्यर्थ हो जायेगा।

(3) **गहन अध्ययन की आवश्यकता** (Need of Intensive Study)- इस प्रणाली का उपयोग उन परिस्थितियों में भी किया जा सकता है जबकि हम अनुसंधान के विषय के संबंध में गहन अध्ययन करना चाहते हों।

(4) **शुद्धता एवं विश्वसनीयता** (Accuracy and Reliability)- ऐसी परिस्थिति में भी यह प्रणाली आवश्यक हो जाती है। जब हम अपने अनुसंधान के निष्कर्षों में अधिक से अधिक शुद्धता और विश्वसनीयता प्राप्त करना चाहते हैं। यदि हमें यह संदेह है कि लेशमात्र भी ढिलाई के कारण अनुसंधान का उद्देश्य व्यर्थ हो जायेगा तो हमें इसी प्रणाली को अपनाना चाहिए।

समग्र प्रणाली के लाभ (Merits of Census Method)- प्रणाली जैसा कि ऊपर बतलाया गया है, कुछ परिस्थितियों में आवश्यकता होती हैं। फिर भी इस प्रणाली के अपने गुण-दोष हैं। हम पहले इसके गुणों का वर्णन करेंगे:-

(1) समग्र के प्रत्येक मद से सूचना प्राप्त करने के कारण निकाले गये निष्कर्षों में अधिक शुद्धता एवं यथार्थता रहती है।

(2) प्रत्येक क्षेत्र की प्रत्येक इकाई के विषय में विस्तृत सूचना प्राप्त हो जाती है। इस सूचना का उपयोग बाद में कई अन्य समस्याओं का अध्ययन करने में किया जा सकता है।

(3) इस प्रणाली के कारण विषय का गहन अध्ययन हो जाता है और अनुसंधानकर्ता को उस विषय के बारे में पूर्ण ज्ञान हो जाता है। इस प्रणाली के द्वारा ऐसी बारीक बातों का भी पता चल जाता है जो अन्यथा ध्यान में न आयें।

(4) यह प्रणाली ऐसी समस्याओं के अध्ययन के लिए भी उपयोगी है जिनका अध्ययन प्रत्यक्ष रूप से सांख्यिकी के अन्तर्गत नहीं हो सकता है, जैसे दरिद्रता, बेकारी, ईमानदारी आदि।

(5) जहाँ इकाइयाँ एक दूसरे से काफी भिन्न हों और न्यादर्श प्रणाली द्वारा सफलतापूर्वक अनुसंधान करना कठिन हो वहाँ इस प्रणाली का उपयोग लाभप्रद रहता है।

समग्र प्रणाली के दोष (Demerits of Census Method)- इस प्रणाली का उपयोग व्यापक क्षेत्र में विस्तृत रूप से किया जाता है अतः इसमें कई कठिनाइयाँ एवं बाधायें उपस्थित हो जाती हैं-

इस प्रणाली में अधिक धन, अधिक समय, अधिक गणक और अधिक शक्ति (energy) लगानी पड़ती है जिससे यह बहुत असुविधाजनक हो जाती है।

(2) इस प्रणाली के प्रयोग में व्यवस्था सम्बन्धी कठिनाइयाँ भी बहुत बढ़ जाती हैं। इसके लिए अलग से पूरा एक बड़ा विभाग बनाना पड़ता है। यह सरकार अथवा एक बड़ी संस्था के लिए ही सम्भव है।

(3) समग्र अनुसंधान द्वारा निकाले गये निष्कर्षों और न्यादर्श प्रणाली द्वारा निकाले गये निष्कर्षों में प्रायः विशेष अन्तर नहीं पड़ता है फिर क्यों अनावश्यक रूप से इस प्रणाली का उपयोग किया जाये।

3.3 निदर्शन अथवा न्यादर्श प्रणाली (Sample Method)

ऊपर हमने समग्र अनुसंधान के गुण-दोषों का विवेचन किया है। इससे यह स्पष्ट हो जाता है कि इस प्रकार का अनुसंधान व्यक्ति विशेष की सामर्थ्य से बाहर होता है। व्यक्ति विशेष ही नहीं, वे संस्थायें भी इस प्रणाली को नहीं अपना सकती हैं जिनके आर्थिक साधन सीमित होते हैं। इस कारण ही अधिकांशतः निदर्शन अथवा न्यादर्श प्रणाली के द्वारा अनुसंधान किए जाते हैं। इस प्रणाली में समग्र का अध्ययन नहीं किया जाता है। केवल समग्र (Universe) में से कुछ पद (Units) अथवा इकाइयाँ चुन ली जाती हैं जो समग्र का प्रतिनिधित्व करती हों। इनका अध्ययन कर सामग्री का संकलन किया जाता है। इसे ही निदर्शन अनुसंधान कहते हैं। उदाहरण के लिए एक महाविद्यालय में 500 विद्यार्थी हैं। हमें उनका आर्थिक एवं सामाजिक सर्वेक्षण करना है। यदि इस सर्वेक्षण में सभी से जानकारी प्राप्त कर हम अपना निष्कर्ष निकालें तो यह समग्र अनुसंधान कहलायेगा। यदि इन 500 विद्यार्थियों में से केवल 50 विद्यार्थी चुन लिए जायें और उनसे ही जानकारी प्राप्त कर इसके आधार पर अपने निष्कर्ष निकाल लिए जायें तो यह निदर्शन अनुसंधान कहलायेगा। दोनों पद्धतियों द्वारा जो परिणाम अथवा निष्कर्ष निकलेंगे वे लगभग एक से ही होंगे। इसलिये सदैव समग्र-प्रणाली को हर परिस्थिति में अपनाने से कोई विशेष प्रयोजन हल नहीं होता है और अनावश्यक रूप से धन, समय व शक्ति का व्यय होता है और मानव ज्ञान का भी बहुत धीमी गति से विकास होता है। इस प्रणाली के महत्व को प्रसिद्ध सांख्यिक स्नेडेकोर के शब्दों में इस प्रकार व्यक्ति किया जा सकता है- "केवल कुछ पौंड कोयले जाँच के आधार पर एक गाड़ी कोयला स्वीकृत या अस्वीकृत कर दिया जाता है। केवल एक बूँद खून की जाँच करके एक रोगी के खून के विषय में चिकित्सक निष्कर्ष निकालता है। न्यादर्श ऐसी युक्तियाँ (devices) हैं जिनके द्वारा केवल कुछ इकाइयों का निरीक्षण करके बड़ी मात्राओं के विषय में जाना जाता है।"

यहाँ यह बतला देना उपयुक्त होगा कि इस प्रकार के अनुसंधान में प्रतिनिधि इकाई का चुनाव एक महत्वपूर्ण कार्य है। यदि चुनी गई इकाइयाँ या पद समग्र का प्रतिनिधित्व करते हैं तो उनसे निकाले गए निष्कर्ष सही होंगे, अन्यथा गलत। अतः चुनी गई इकाइगों (निदर्शन) में निम्न गुण होने चाहिए-

(1) निदर्शन (Samples) ऐसे होने चाहिए जो समग्र का पूर्णरूपेण प्रतिनिधित्व करते हों।

(2) निदर्शन की संख्या अधिक होनी चाहिए जिससे निष्कर्ष अधिक सही निकलेंगे। किन्तु यह भी स्पष्ट है कि निदर्शन की संख्या जितनी अधिक होगी उतना ही धन, समय व श्रम अधिक लगेगा और कम संख्या होने की दशा में निष्कर्ष कम सही होंगे। अतः मध्य मार्ग ही उत्तम होगा कि निदर्शन की संख्या न बहुत अधिक हो और न बहुत कम।

निदर्शन का आधार (Basis of Sampling)

संगणना प्रणाली के स्थान पर निदर्शन प्रणाली का प्रयोग करने के लिए क्या आधार है? निदर्शन का आधार समग्र (Universe) की इकाइयों में पाई जाने वाली एकरूपता है। सभी इकाइयाँ एक दूसरे से सभी बातों में मिलती-जुलती हों, ऐसा तो प्रायः असम्भव ही होता है। किन्तु इकाइयों में बाह्य भिन्नता होते हुए भी कुछ गुण ऐसे होते हैं जो सभी में समान रूप से पाये जाते हैं। बाह्य रूप से प्रत्येक श्रमिक एक दूसरे से भिन्न होता है। चतुरता, स्वास्थ्य, चरित्र, कार्य-क्षमता तथा मानसिक प्रवृत्तियों में एक श्रमिक और दूसरे श्रमिक में बड़ा अन्तर होता है। किन्तु उनमें कुछ बातों में समानता भी होती है जो सभी स्थानों के श्रमिकों पर उन परिस्थितियों में लागू होती है। जैसे अहमदाबाद के वस्त्र-उद्योग में लगे हुए 10% श्रमिकों का अनुसंधान किया गया, जिससे यह पता लगा कि उनमें से 20% श्रमिक मद्यपान करते हैं। यदि सभी श्रमिकों का अनुसंधान भी किया जाये तो भी मद्यपान करने वालों का प्रतिशत लगभग इतना ही निकलेगा। स्पष्ट है कि जो बात निदर्शन के विषय में सत्य है, वह समग्र के विषय में भी सत्य होगी।

NOTES

निदर्शन प्रणाली की आवश्यकता

वैसे तो उपरोक्त वर्णन से यह स्पष्ट हो ही जाता है कि इस प्रणाली के उपयोग की आवश्यकता कब और क्यों पड़ती है। फिर भी हम इस प्रणाली की आवश्यकता के कारणों का संक्षेप में नीचे वर्णन करते हैं-

(1) कभी-कभी एकमात्र यही प्रणाली सम्भव होती है। संगणना प्रणाली का उपयोग हो ही नहीं सकता है। दो परिस्थितियों के उत्पन्न होने से ऐसा हो सकता है। प्रथम, समग्र असीमित हो जैसे किसी औद्योगिक पदार्थ का उत्पादन। उत्पादन की कोई सीमा नहीं है, अतः उसकी किस्म की जाँच संगणना प्रणाली से नहीं हो सकती है। द्वितीय, ऐसा सम्भव है कि किस्म अथवा टिकाऊपन की जाँच करते समय वस्तु स्वयं नष्ट हो जाये। ऐसी परिस्थिति में संगणना प्रणाली द्वारा जाँच करने पर सभी वस्तुएँ नष्ट हो जायेंगी। अतः निदर्शन प्रणाली ही एकमात्र प्रणाली होती है जिसका उपयोग किया जाता है। इसी प्रकार यदि हमें किसी डिब्बे के घी की जाँच करनी है तो निदर्शन प्रणाली अपनाकर नमूने के तौर उसकी जाँच करनी होगी, अन्यथा सभी घी जाँच ही में समाप्त हो जाएगा।

(2) सामान्यतः यही प्रणाली व्यावहारिक होती है। यदि समग्र का क्षेत्र बहुत बड़ा हो तो संगणना प्रणाली के उपयोग में धन, समय तथा शक्ति बहुत व्यय होंगे। जैसे एक ब्लेड बनाने वाली कम्पनी अपने ब्लेड के सम्बन्ध में लोगों की राय जानने के लिए ब्लेड का उपयोग करने वाले प्रत्येक व्यक्ति से सम्पर्क स्थापित नहीं कर सकती क्योंकि उनकी संख्या भी अधिक होगी और वे काफी दूर-दूर रहते होंगे। ऐसी परिस्थिति में निदर्शन प्रणाली ही सर्वोत्तम रहती है।

(3) इस प्रणाली से प्राप्त होने वाले निष्कर्ष प्रायः संगणना प्रणाली के निष्कर्षों के समान ही ठीक होते हैं। फिर यही प्रणाली क्यों नहीं अपनाई जाये। प्रो. नीस्वैंगर के अनुसार, "वास्तव में निदर्शन के लिए निकाली गई अपेक्षाकृत थोड़ी सी इकाइयाँ अधिक शुद्धता से संकलित की जा सकती हैं तथा समग्र की पूर्ण संगणना की अपेक्षा अधिक उत्तम फल प्रदान कर सकती हैं।"

(4) ऐसी परिस्थितियाँ भी कभी-कभी उत्पन्न हो जाती हैं जबकि सम्पूर्ण सामग्री का मिलना असम्भव हो जाता है तब निदर्शन प्रणाली ही अपनानी पड़ती है। किसी गोदाम में आग लगने से सब रुई की गाठें जल गई और केवल पाँच गाँठें शेष रहीं तो निदर्शन प्रणाली के आधार पर ही रिपोर्ट बीमा कम्पनी को देनी होगी।

(5) जब वस्तुओं की प्रकृति शीघ्र बदलने वाली हो तब संगणना प्रणाली के प्रयोग के कारण अधिक समय लगने से निष्कर्ष असंतोषजनक रहेगा और ऐसी परिस्थिति में निदर्शन प्रणाली का प्रयोग ही उत्तम रहता है।

अन्त में निदर्शन प्रणाली की आवश्यकता को **प्रो. नीस्वैंगर** के शब्दों में इस प्रकार व्यक्त किया जा सकता है-

"निदर्शन प्रणाली का उपयोग आर्थिक व व्यापारिक अनुसंधान में विस्तृत रूप से होता है क्योंकि सामूहिक समंकों के अध्ययन में कभी-कभी यह एकमात्र सम्भव प्रणाली, प्रायः सबसे अधिक व्यावहारिक और साधारणतया सबसे उत्तम प्रणाली होती है।"

3.4 निदर्शन चुनने की पद्धतियाँ (Methods of Selecting Samples)

निदर्शन चुनने की निम्नलिखित पद्धतियाँ प्रचलित हैं-

(1) विस्तृत या व्यापक निदर्शन (Extensive Sampling)

(2) सविचार निदर्शन (Deliberate, Intentional, Purposive, Conscious or Representative Sampling)

(3) दैव प्रवरण या आकस्मिक निदर्शन (Random Sampling or Chance Selection)-

- i लॉटरी पद्धति (Lottery System)
- ii ढोल घुमाकर (Rotating the Drum)
- iii आँख बन्द करके चुनना (Blind fold Selection)
- iv 'स' वाँ नम्बर या नियमानुसार दैव निदर्शन (n^{th} number or Systematic random sampling)
- v पदों को किसी रीति से सजाकर (Arrangement of items in some order)

vi टिपेट की संख्याओं द्वारा (Tippet's tables)

(4) मिश्रित या स्तरित निदर्शन (Mixed or Stratified Sampling)

(5) बहुस्तरीय निदर्शन (Multi-stage sampling)

(6) अन्य निदर्शन प्रणालियाँ (Other Sampling methods)

i बहुचरण निदर्शन प्रणाली (Multi-phase Sampling)

ii अभ्यंश निदर्शन (Quota Sampling)

iii अनुक्रमिक निदर्शन (Sequential Sampling)

iv सुविधानुसार निदर्शन (Convenience Sampling)

v संतुलित निदर्शन प्रणाली (Balanced Sampling)

vi अजातीय पद समूह निदर्शन (Cluster Sampling)

अब हम उपरोक्त पद्धतियों का अलग-अलग विवेचन करेंगे:-

(1) विस्तृत या व्यापक निदर्शन (Extensive Sampling)

संगणना प्रणाली के समान ही विस्तृत निदर्शन की प्रणाली होती है। इस निदर्शन में न्यादर्श अधिक मात्रा में लिया जाता है। कभी-कभी तो सभी उपलब्ध इकाइयों का अध्ययन इसके अन्तर्गत किया जाता है। इसलिए यदि कहा जाये तो अतिशयोक्ति नहीं होगी कि विस्तृत निदर्शन संगणना प्रणाली का ही एक रूप है। फिर भी दोनों में अन्तर है, यद्यपि यह अन्तर विशेष महत्वपूर्ण नहीं है। संगणना प्रणाली में प्रत्येक पद का निश्चित रूप से अध्ययन किया जाता है। इसके प्रतिकूल विस्तृत निदर्शन में केवल उन समस्त पदों का अध्ययन किया जाता है जो सुगमता से उपलब्ध हो जाते हैं।

इस पद्धति के गुण (Merits)- (1) संगणना प्रणाली की अपेक्षा यह प्रणाली सुगम होती है। जो पद सुगमता से प्राप्त हो जाते हैं, उन्हें अध्ययन में सम्मिलित कर लिया जाता है और जिनकी प्राप्ति दुर्लभ अथवा कठिन होती है, उन्हें छोड़ दिया जाता है।

(2) इस प्रणाली द्वारा किया गया अनुसंधान विश्वसनीय होता है क्योंकि इसमें भी संगणना प्रणाली की भाँति ही प्रायः सभी पदों की जाँच की जाती है।

(3) इसके द्वारा विषय का गहन अध्ययन सम्भव होता है तथा परिणाम बहुत अंशों तक शुद्ध होते हैं।

दोष (Demerits)- (1) संगणना प्रणाली की भाँति ही इस प्रणाली द्वारा अनुसंधान में बहुत अधिक धन, समय तथा परिश्रम की आवश्यकता होती है। अतः इस प्रणाली का उपयोग भी सीमित रूप से ही हो सकता है।

(2) ऐसी परिस्थिति में इस प्रणाली का उपयोग दोषपूर्ण हो जाता है जबकि अनुसंधानकर्ता पक्षपातपूर्ण नीति का अनुकरण कर रहा हो। इस बात का न्यादर्श पर प्रतिकूल प्रभाव पड़ता है और अनुसंधान के सच्चे लक्ष्य की प्राप्ति में बाधा उपस्थित हो जाती है।

(3) सुगमता से उपलब्ध पदों को ही लेने के कारण यह भय रहता है कि सम्भवतः अधिक महत्वपूर्ण पदों की जाँच न हो पाये जिनको प्राप्त करना कठिन है। इसके कारण निष्कर्ष अशुद्ध होने का भय रहता है।

(2) सविचार निदर्शन
(Deliberate, Intentional, Purposive, Conscious or Representative Sampling)

जैसा कि इस प्रणाली के नाम से ही स्पष्ट होता है, इसमें अनुसंधानकर्ता न्यादर्श का चुनाव विचारपूर्वक करता है। वह चुनाव करते समय इस बात का प्रयत्न करता है कि सम्पूर्ण की सब विशेषतायें न्यादर्श में आ जायें। इस उद्देश्य की पूर्ति हेतु वह अपने न्यादर्श में ऐसे पदों को सम्मिलित करता है जो समग्र की सभी विशेषताओं को प्रकट करने वाले हों। ऐसे पदों के चुनाव में सामान्यतः कोई प्रमाप निश्चित कर लिया जाता है। उदाहरण के लिए, यदि आर्थिक पर्यवेक्षण के लिए मध्यप्रदेश के गाँवों का एक न्यादर्श चुनना हो तो केवल औसत श्रेणी के गाँव चुने जायेंगे। अधिक उन्नतिशील अथवा पिछड़े हुए गाँवों को छोड़ दिया जायेगा।

इस प्रणाली का उपयोग उस अनुसंधान में किया जाता है, जिसमें कुछ इकाइयाँ इतनी महत्वपूर्ण होती हैं कि उनका सम्मिलित करना अनिवार्य रहता है। जैसे भारतवर्ष में लौह-इस्पात उद्योग का पर्यवेक्षण टाटा आयरन एण्ड स्टील

कम्पनी को सम्मिलित किए बिना अधूरा ही रह जायेगा। इस प्रणाली की सफलता बहुत कुछ अनुसंधानकर्ता की ईमानदारी, निष्पक्षता एवं ज्ञान पर आधारित है।

इस प्रकार सविचार निदर्शन की तीन मुख्य प्रणालियाँ होती हैं जिनका उल्लेख इस प्रकार किया जा सकता है-

(1) **औसत गुण की इकाइयों का चुनाव** (Selection of the unit of average quality)- जैसा कि ऊपर बतलाया गया है केवल औसत गुण वाली इकाइयों का चुनाव के आधार पर निष्कर्ष निकाले जा सकते हैं। इस प्रकार के चुनाव से निकाले हुए निष्कर्षों से समग्र को प्रकट किया जा सकता है। यदि अधिक पिछड़ी हुई इकाइयों को छोड़ दिया जाये तो बहुमत पर बुरा प्रभाव पड़ेगा।

(2) **उद्देश्यानुसार सविचार न्यादर्श का चुनाव** (Intentional selection according to the object)- उद्देश्य के अनुसार जान-बूझ कर न्यादर्श का इस प्रकार चुनाव किया जाता है ताकि कोई महत्वपूर्ण इकाई न छूटने पावे।

(3) **आनुपातिक चुनाव** (Proportional Selection)- इसके अनुसार प्रत्येक समूह को उसी अनुपात में न्यादर्श में सम्मिलित किया जाता है जिस अनुपात में वे अनुसंधान के क्षेत्र में होते हैं। इस प्रणाली में अनुसंधानकर्ता की भावना का चुनाव पर अधिक प्रभाव पड़ता है। इसलिए इस प्रकार के निकाले गए निष्कर्ष विश्वसनीय नहीं माने जा सकते हैं। उदाहरणार्थ एक साम्यवादी विचारधारा वाला अनुसंधानकर्ता ऐसे श्रमिकों की मजदूरी को प्रतिनिधि चुनेगा जिनको न्यूनतम मजदूरी मिलती हो। इसके बिल्कुल विपरीत पूँजीपति अनुसंधानकर्ता उन मजदूरों का चुना करेगा जिसकी अधिकाधिक मजदूरी होगी। इस प्रकार दोनों द्वारा एकत्रित तथ्य दोषपूर्ण होंगे। पहले के अनुसार औसत मजदूरी बहुत कम और दूसरे के अनुसार बहुत अधिक सिद्ध होगी।

पद्धति के गुण (Merits)- (1) यह पद्धति सरल है।

(2) यदि योजना ठीक ढंग से बना ली जाती है और प्रमाप निश्चित कर लिया जाता है तो न्यादर्श का चुनाव (Selection of Sample) ठीक होने की सम्भावना रहती है, और न्यादर्श का चुनाव ठीक होता है तो निष्कर्ष भी विश्वसनीय होते हैं।

(3) इस पद्धति का उपयोग उस अनुसंधान के लिए तो आवश्यक ही है जहाँ कुछ इकाइयाँ इतनी अधिक महत्वपूर्ण हों कि उन्हें न्यादर्श में सम्मिलित करना अनिवार्य ही हो- जैसा कि ऊपर एक उदाहरण में बतलाया गया है कि लौह-इस्पात उद्योग का अनुसंधान टाटा आयरन एण्ड स्टील कम्पनी को न्यादर्श के रूप में लिये बिना उपयोगी ही नहीं होगा।

पद्धति के दोष (Demerits)- ऊपर हमने सविचार निदर्शन के गुणों की व्याख्या की है किन्तु उस पद्धति में कुछ दोष भी हैं जिनका उल्लेख हम इस प्रकार कर सकते हैं-

(1) इस पद्धति की सफलता न्यादर्श के चुनावकर्ता की ईमानदारी, पक्षपात विहीनता तथा ज्ञान पर निर्भर करती है। चुनाव करने वाले की पूर्व विचारधारा या धारणा का प्रभाव न्यादर्श के चुनाव पर अवश्य पड़ता है जो निष्कर्ष को अशुद्ध बना देता है।

(2) जैसा कि बतलाया गया है कि न्यादर्श के चुनावकर्ता में उचित ज्ञान का होना आवश्यक है ताकि वह समग्र के प्रत्येक अंग की विशेषता को ठीक प्रकार समझ सके। अतः इस पद्धति का प्रयोग सीमित हो जाता है।

(3) दैव प्रवरण या आकस्मिक निदर्शन (Random Sampling or Chance Selection)

इस पद्धति में सविचार पद्धति के विपरीत अनायास ही कुछ इकाइयाँ चुन ली जाती हैं। चुनावकर्ता इसके लिये अपनी बुद्धि का उपयोग नहीं करता है कि कौनसी इकाइयों का चुनाव किया जाये। चुनाव का कार्य संयोग पर छोड़ दिया जाता है। कोई भी इकाई चुनाव में आ सकती है। इस प्रकार प्रत्येक इकाई के चुने जाने का समान अवसर रहता है। उदाहरणार्थ, यदि हम मध्यप्रदेश के आदिवासी क्षेत्र का पर्यवेक्षण करना चाहते हैं तो इस पद्धति द्वारा आकस्मिक रूप से कुछ गाँव चुन लिये जायेंगे। यह असम्भव है कि इस प्रकार चुनाव करने से सभी आदिवासी क्षेत्र के प्रगतिशील या पिछड़े हुये गाँव की इकाइयों का समावेश हमारे चुनाव में हो जाये किन्तु ऐसी परिस्थिति प्रायः नहीं होती है। देखा यह गया है कि जब चुनाव इस आधार पर होता है तथा चुनावकर्ता अपनी बुद्धि का उपयोग नहीं करता है तो, अच्छी, बुरी तथा सामान्य सभी प्रकार की इकाइयाँ उसमें आ जाती हैं तथा न्यादर्श समग्र का पूर्णरूपेण प्रतिनिधित्व कर सकता है।

इस पद्धति द्वारा चुनाव करने की कई विधियाँ हैं जिनका संक्षेप में उल्लेख इस प्रकार किया जा सकता है-

(1) **लॉटरी पद्धति** (Lottery System)- इस पद्धति में समस्त इकाइयों के नाम अथवा क्रम-संख्या अलग-अलग कागजों पर लिख लिए जाते हैं और उन्हें एक स्थान पर रख दिया जाता है या एक बॉक्स में डाल दिया जाता है और उनमें से कुछ उठा लेते हैं या निकाल लेते हैं।

(2) **ढोल घुमाकर** (Rotating the Drum)- एक ढोल में एक सी आकार की लकड़ी या अन्य धातु के टुकड़े रखे जाते हैं जिन पर अलग-अलग संख्याएँ लिखी जाती हैं। ढोल को हाथ से या बिजली की शक्ति से घुमाया जाता है ताकि सब टुकड़े ऊपर-नीचे (Randomise) हो जायें। तदुपरान्त किसी भी व्यक्ति द्वारा या मशीन की सहायता से एक-एक टुकड़ा निकाला जाता है जिस पर अंकित संख्या को देख लिया जाता है। इस प्रकार न्यादर्श में जितने मदों की आवश्यकता होती है, उतने ही टुकड़े निकालकर उनकी अंकित संख्या लिख दी जाती है। उन सब संख्या वाले मदों का एक न्यादर्श बन जाता है। भारत सरकार द्वारा निर्गमित इनामी बॉण्डों के त्रैमासिक इनाम मालूम करने के लिये लगभग यही पद्धति काम में लाई जाती थी।

(3) **आँख बंद करके चुनना** (Blind fold selection)- इस पद्धति में चुनावकर्त्ता विभिन्न मदों के लिये अंकित संख्याओं में आँख बंद करके कुछ को उठा लेता है और वे ही संख्यायें न्यादर्श में सम्मिलित की जाती हैं। इस पद्धति का प्रयोग एक अन्य प्रकार से भी हो सकता है। सामने दीवार पर एक बहुत बड़ा चौकोर नक्शा होता है जिस पर एक से लेकर 100, 200, 500 या 1000, तक समान लम्बाई-चौड़ाई के खाने (Square) या वृत्त (Circles) बने होते हैं। उन पर बिना किसी क्रम के संख्यायें लिखी जाती हैं। किसी निश्चित दूरी से कोई भी व्यक्ति उस कागज पर (नक्शे पर) आँख बन्द करके या आँख पर पट्टी बाँध कर तीर फेंकता है। तीर जिस संख्या पर जाकर लगता है उसी संख्या को नोट कर लिया जाता है। इस प्रकार न्यादर्श में जितने मदों की आवश्यकता हो उतनी ही बार तीर मार कर संख्यायें नोट कर ली जाती हैं। इन सब संख्याओं वाले मदों का एक न्यादर्श बन जाता है। यदि कोई तीर दो संख्याओं के बीच या एक ही संख्या पर एक से अधिक बार गिरे तो उसे रद्द कर दिया जाता है और उसके लिये तीर फिर से फेंकना पड़ता है।

(4) **'स' वाँ नम्बर या नियमानुसार दैव निदर्शन** (nth Number or systematic Random Sampling)- समग्र के समस्त मदों को किसी भी क्रम में- भौगोलिक (geographical), वर्णनात्मक (alphabertical) या संख्यात्मक (numerical) जमा दिया जाता है। फिर 'स' वाँ (nth) नम्बर मालूम कर लिये जाते हैं। जैसे हम स = 10 मानते हैं तो दस-दस के अन्तर से नम्बर नोट कर लिये जाते हैं। यदि हम आरम्भ में तीसरे नम्बर से आरम्भ करते हैं तो दस-दास के अन्तर पर तेरहवाँ, तेईसवाँ, तैंतीसवाँ, आदि नम्बर नोट कर लिये जाते हैं। इस प्रकार से मदों की संख्या ज्ञात करके न्यादर्श बना लिया जाता है। इस पद्धति को व्यवस्थित दैवनिदर्शन भी कहते हैं। इस पद्धति में दैव निदर्शन की अपेक्षा व्यक्तिगत नियंत्रण की मात्रा अधिक रहती है। प्रो. नोस्वैंगर के अनुसार "व्यवस्थित दैव निदर्शन में सूची की इकाइयों के क्रम पर नियंत्रण के साथ-साथ अनायास चुनाव का फल प्राप्त होता है जिससे इकाइयों का प्रादेशिक विभाजन प्राप्त हो सकता है जो कि अनियंत्रित दैव निदर्शन में सम्भव नहीं रहता।" इस पद्धति का उपयोग उस समय उचित नहीं रहता है जब इकाइयों के मूल्य में एक निश्चित गति के अनुसार परिवर्तन होता है।

(5) **पदों को किसी रीति से सजाकर** (Arrangement of items in some order)- यह पद्धति उपरोक्त पद्धति से ही मिलती-जुलती है। इसमें पहले पदों को किसी ढंग से सजा लेते हैं और उनमें से आकस्मिक ढंग से कुछ पदों को चुन लेते हैं।

(6) **टिपेट की संख्याओं द्वारा** (Tippet's Tables)- टिपेट प्रसिद्ध सांख्यिक हुये हैं। इन्होंने 41,600 इकाइयों के चार-चार अंकों की 10,400 संख्याओं की एक सारणी (table) तैयार की है। इस सारणी की सहायता से न्यादर्श का चुनना सरल होता है। सबसे पूर्व सभी पदों के लिये संख्यायें निश्चित कर लेते हैं और फिर बाद में सारणी की सहायता से किन्हीं 10, 20 या 50 या अन्य संख्याओं को चुन लेते हैं। ये संख्यायें जिन पदों को प्रकट करती हैं उन्हें न्यादर्श में सम्मिलित कर लिया जाता है। बड़े-बड़े सर्वेक्षणों में टिपेट की सारणी का ही प्रयोग किया जाता है।

दैव निदर्शन प्रणाली के गुण (Merits)- (1) इस पद्धति से चुनाव करने में चुनावकर्त्ता की इच्छा का कोई प्रभाव नहीं पड़ता है। इसलिये पक्षपात की गुंजाइश नहीं रहती है। सभी पदों के चुने जाने का समान अवसर रहता है।

(2) चुनावकर्त्ता को कोई बुद्धि नहीं लगानी पड़ती है। वह अनायास चुनाव करता है।

(3) निदर्शन-भूल आसानी से ज्ञात की जा सकती है क्योंकि यह पद्धति सम्भाविता सिद्धान्त (Theory of Probability) पर ही आधारित है।

(4) चुनाव के लिये किसी विस्तृत योजना के बनाने की आवश्यकता नहीं पड़ती है।

(5) इस पद्धति द्वारा धन, श्रम व समय भी अधिक नहीं लगता है।

(6) इस पद्धति में न्यादर्श की शुद्धता की जाँच अन्य न्यादर्श लेकर भी की जा सकती है।

दोष (Demerits)- (1) इस पद्धति के अपनाने से कभी-कभी ऐसे पदों का चुनाव हो जाता है जो समग्र का बिल्कुल ही प्रतिनिधित्व नहीं करते।

(2) इस पद्धति का उपयोग उस अनुसंधान के लिये नहीं किया जा सकता जहाँ कुछ इकाइयाँ इतनी महत्वपूर्ण हों कि उन्हें न्यादर्श में सम्मिलित करना आवश्यक हो।

इन दोषों को देखते हुये दैव निदर्शन पद्धति का प्रयोग तभी सम्भव है जबकि चुनावकर्त्ता में पक्षपात की भावना तनिक भी न हो तथा प्रत्येक पद या अंश को चुनाव में आ जाने का समान अवसर प्राप्त हो।

(4) मिश्रित या स्तरित निदर्शन (Mixed or Stratified Sampling)

मिश्रित निदर्शन जिसे स्तरित निदर्शन प्रणाली भी कहते हैं, का प्रयोग- सविचार निदर्शन और दैव निदर्शन प्रणाली दोनों के लाभों को एक साथ प्राप्त करने के लिये किया जाता है। इस प्रणाली में सविचार तथा दैव निदर्शन दोनों ही प्रणालियों के दोष दूर हो जाते हैं।

प्राय: समग्र में विभिन्नता अधिक होने के कारण एक साथ समग्र एकत्रित करना कठिन होता है। ऐसी दशा में समग्र को सर्वप्रथम कई भागों में बाँट दिया जाता है। यह बंटवारा या विभाजन इस प्रकार किया जाता है कि प्रत्येक भाग में एक ही प्रकार के लक्षण वाले समंक प्राप्त हो सकें। इसके उपरान्त प्रत्येक भाग में से दैव निदर्शन प्रणाली द्वारा समंकों को एकत्रित कर लिया जाता है। इस प्रकार सभी लक्षणों वाले सब भागों के प्रतिनिधि अंक प्राप्त कर लिये जाते हैं। इसे उदाहरण द्वारा इस प्रकार स्पष्ट किया जा सकता है- यदि हम किसी कक्षा के विद्यार्थियों के विषय में कोई जानकारी प्राप्त करना चाहते हैं तो सर्वप्रथम हम विद्यार्थियों को कई भागों में विभागों के अनुसार बाँट देंगे जैसे विज्ञान, वाणिज्य कला, शिक्षण, विधि, इत्यादि। इसके उपरान्त प्रत्येक विभाग से एक निश्चित मात्रा में दैव निदर्शन पद्धति द्वारा इकाइयों का चुनाव कर लेंगे। यदि भिन्न-भिन्न विभागों में विद्यार्थियों की संख्या में काफी अन्तर होगा तो चुनी जाने वाली इकाइयों की संख्या भी उसी अनुपात में कम या ज्यादा कर ली जायेगी।

यह पद्धति दैव निदर्शन तथा सविचार निदर्शन पद्धतियों की तुलना में अधिक उत्तम है क्योंकि इसमें इन दोनों के गुणों का समावेश है। क्राक्सटन तथा काउडेन के कथनानुसार, "मिश्रित प्रणाली में विभागों या वर्गों की उपस्थिति की मान्यता न्यादर्श (Sample) के चुनाव में और अधिक नियंत्रण उपस्थित करती है तथा हमें प्रतिनिधित्व का और अधिक विश्वास दिलाती है जो विभागों की संख्या के अनुसार ही बढ़ता है।" यही कारण है कि यह पद्धति आजकल अधिक प्रयोग में लाई जाती है। इस पद्धति का सबसे बड़ा दोष यह है कि यदि विभाग या वर्ग बनाने में कोई त्रुटि रह जाये तो इस पद्धति द्वारा निकाले गये निष्कर्ष संतोषजनक नहीं होते हैं।

(5) बहुस्तरीय निदर्शन (Multi-stage Sampling)

इस पद्धति का उपयोग उस समय किया जाता है जब किसी बड़े शहर में से क्षेत्रों के आधार पर जनसंख्या का नमूना लिया जाता है। इसके अनुसार नमूने का चुनाव कई स्तरों में होता है। सबसे पहले सम्पूर्ण शहर को भिन्न-भिन्न क्षेत्रों में बाँटा जाता है। क्षेत्रों में बाँटते समय यह ध्यान रखा जाता है कि एक क्षेत्र में लगभग एक ही प्रकार के लोग रहते हों ताकि क्षेत्र विभाजन में एकरूपता लाई जा सके। इसके बाद प्रत्येक क्षेत्र से दैव निदर्शन पद्धति के अनुसार एक गृह-समूह (Block-cluster) का चुनाव किया जाता है। फिर प्रत्येक गृह-समूह से कुछ गृहों को लिया जाता है। इन गृहों में से कुछ व्यक्तियों को चुनकर उनकी जाँच की जाती है।

इस पद्धति से यह यह स्पष्ट है कि न्यादर्श (नमूने) का चुनाव एक साथ न होकर कई स्तरों (stages) में होता है। प्रत्येक स्तर पर चुनाव करते समय दैव निदर्शन प्रणाली का उपयोग होता है जिसके कारण प्रत्येक इकाई के चुने जाने का समान अवसर रहता है। इस प्रणाली की सबसे मुख्य विशेषता यही है कि जनसंख्या का विभाजन किसी आधार पर न होकर क्षेत्रीय आधार पर होता है।

यद्यपि यह प्रणाली दैव निदर्शन के सभी लाभों से युक्त होती है, किन्तु क्षेत्रों में जनता की एकरूपता प्राप्त होना कठिन होता है। इसी से क्षेत्रों का विभाजन, जैसा इस प्रणाली के अनुसार होना चाहिए, वैसा नहीं हो पाता है। इस प्रणाली में वे दोष भी होते हैं जो दैव निदर्शन प्रणाली में पाये जाते हैं। बड़े समग्र में आजकल इस प्रणाली का प्रयोग अधिक होने लगा है।

(6) अन्य निदर्शन प्रणालियाँ (Other Sampling Methods)

उपरोक्त महत्वपूर्ण प्रणालियों के अतिरिक्त निदर्शन की कुछ अन्य प्रणालियाँ भी हैं जिनका संक्षेप में वर्णन नीचे करेंगे-

(1) **बहुचरण निदर्शन प्रणाली** (Mutli-phase Sampling)- हम कभी-कभी एक ही समग्र में एक साथ कई समस्याओं के सम्बन्ध में सूचना प्राप्त करना चाहते हैं। ऐसी परिस्थिति में अगर प्रत्येक समस्या के लिए अलग-अलग तथ्य एकत्रित किए जायें तो व्यय अधिक होगा। इस कारण एक बड़ा न्यादर्श (Sample) चुन लिया जाता है और उसी में से प्रत्येक समस्या के लिए एक उप-न्यादर्श (Sub-sample) चुन लिया जाता है। इस तरह पहला उप-न्यादर्श प्रथम चरण न्यादर्श तथा दूसरा उप-न्यादर्श द्वितीय चरण न्यादर्श (Second Phase Sample) आदि कहलाता है।

(2) **अभ्यंश निदर्शन** (Quota Sampling)- इस प्रणाली में पहले समग्र (universe) को कई भागों में बाँट दिया जाता है। यह बंटवारा (विभाजन) इस तरह किया जाता है कि प्रत्येक भाग में समंकों की समरूपता हो जाए। इसके बाद गणकों (Enumerators) को यह बतला दिया जाता है कि किस भाग में से कितनी इकाइयों का चुनाव करना है। इस तरह इकाइयों का अभ्यंश (Quota) निश्चित कर दिया जाता है। अभ्यंश निश्चित कर देने के बाद गणक को यह अधिकार होता है कि वह प्रत्येक भाग में से अभ्यंश इकाइयों का चुनाव स्वयं अपनी इच्छानुसार कर ले।

इस प्रणाली में हम गणकों पर निर्भर रहते हैं। उन्हें इकाइयाँ छाँटने का अधिकार होता है। इस अधिकार-प्राप्ति के कारण वे अपना कार्य बड़ी ईमानदारी से करने का प्रयास करेंगे। इस कारण उनका कार्य बड़ा शुद्ध एवं विश्वसनीय होगा। किन्तु गणक यदि लापरवाही से कार्य करें तो सारा कार्य दोषपूर्ण एवं अशुद्ध हो जायेगा। अतः इस प्रणाली का प्रयोग उसी समय करना चाहिए जब गणक पूर्णरूप से प्रशिक्षित एवं ईमानदार हो।

(3) **अनुक्रमिक निदर्शन** (Sequential Sampling)- साधारणतया सभी निदर्शन प्रणालियों में सर्वप्रथम न्यादर्श का चुनाव करते हैं और फिर उनमें होने वाली भूलों (errors) का अनुमान लगाया जाता है। प्रणाली में इसके विपरीत होता है। इसमें पहले निदर्शन भूल (Sampling error) मालूम करके उसके आधार पर न्यादर्श का आकार (Size) निश्चित करते हैं। इस प्रकार इस प्रणाली में परिशुद्धता की अपेक्षित मात्रा (expected standard of accuracy) का अनुमान वास्तविक निदर्शन से पूर्व ही कर लिया जाता है। इसके बाद चुनाव दैव निदर्शन प्रणाली के द्वारा किया जाता है। इस प्रणाली के प्रतिपादक प्रो. ए. वाल्ड हैं।

(4) **सुविधानुसार निदर्शन** (Convenience Sampling)- इस प्रणाली में अनुसंधानकर्त्ता को जो भी तरीका सुविधाजनक मालूम पड़ता है उसी के अनुसार न्यादर्श का चुनाव कर उसकी जाँच करता है। यदि हमें विश्वविद्यालय के वाणिज्य के प्राध्यापकों में से न्यादर्श लेना हो तो हम कालेजों के प्रविवरण (Prospectus) का उपयोग कर सकते हैं। यह पद्धति अवैज्ञानिक है तथा इसके अन्तर्गत प्राप्त न्यादर्शों के सच्चा प्रतिनिधित्व करने की सम्भावना संदिग्ध रहती है। इस प्रणाली का प्रयोग बहुत कम किया जाता है।

(5) **संतुलित निदर्शन प्रणाली** (Balanced Sampling)- यदि किसी न्यादर्श के मदों का इस प्रकार प्रविरण किया जाए कि उनके अध्ययन से निकाला गया परिणाम ठीक वही हो जो कि समग्र के सभी मदों का परिणाम है तो ऐसे न्यादर्श को संतुलित न्यादर्श कहा जाता है। उदाहरणार्थ, किसी महाविद्यालय में 1,000 विद्यार्थियों की ऊंचाई नापने के लिये 100 विद्यार्थियों का एक न्यादर्श चुना जाता है। उनकी औसत ऊंचाई 5'-2" आती है। यदि 1,000 विद्यार्थियों में प्रत्येक की ऊंचाई नापने के बाद भी औसत ऊंचाई 5'-2" ही आये तो हम ऐसे न्यादर्श को संतुलित न्यादर्श कहेंगे। वास्तव में ऐसे न्यायदर्श की आशा करना स्वप्न मात्र है क्योंकि व्यावहारिक दृष्टि से ऐसा न्यादर्श चुना जाना कठिन है। यदि इस प्रणाली में पूर्ण सफलता मिल जाए तो संगणना प्रणाली की कोई आवश्यकता ही नहीं रहे।

(6) **'अजातीय-पद-समूह' निदर्शन प्रणाली** (Cluster Sampling)- यह प्रणाली स्तरित (Stratified) प्रणाली के बिल्कुल विपरीत है। इसमें ऐसे मदों को चुनने का प्रयत्न किया जाता है जो बिल्कुल भिन्न या अजातीय (Heterogeneous) हो ताकि प्रत्येक न्यादर्श एक छोटे समग्र का रूप धारण कर ले।

NOTES

निदर्शन चुनाव की विभिन्न पद्धतियों का विवेचन करने के बाद हम सम्भावना सिद्धान्त और निदर्शन अनुसंधान (Theory of Probability and Sample Investigation) के विषय में विचार करेंगे। प्रकृति में एक प्रकार की एकरूपता है और यही कारण है कि निदर्शन प्रणाली द्वारा निष्कर्ष बहुत कुछ ठीक निकलते हैं। यदि प्रकृति में इस प्रकार की एकरूपता (uniformity) न होती तो पूरी जाँच के अभाव में संतोषजनक और शुद्ध परिणाम पर पहुंचना कठिन हो जाता। अब हम सम्भावना सिद्धान्त पर विस्तृत रूप से विचार करेंगे।

निदर्शन के सिद्धान्त एवं सार्थकता परीक्षण

1. सांख्यिकीय विचार में समग्र से आप क्या समझते हैं? समग्र प्रणाली किन परिस्थितियों में उपयुक्त है?

....

....

....

2. न्यादर्श प्रणाली क्या है? इसकी आवश्यकता बताइए?

....

....

....

3. सविचार निदर्शन क्या है? इसके गुणों को बताइए?

....

....

....

3.5 सम्भावना सिद्धान्त (Theory of Probability)

यह गणित का एक मुख्य सिद्धान्त है। सभी सांख्यिकीय नियम इसी सिद्धान्त पर आधारित हैं। लाप्लाज के अनुसार यह सिद्धान्त अनुकूल घटनाओं का समस्त होने वाली घटनाओं के आधार के साथ एक अनुपात है। (Probability is the ratio of favourable events of the total number of equally likely events) **कॉनर** के अनुसार, "किसी अनिश्चित घटना के बारे में मस्तिष्क में होने वाली प्रतिक्रिया की ही सम्भावित कहते हैं।" (Probability is an attitude of mind towards uncertain events.)

यह नियम बतलाता है कि यदि एक सिक्का हवा में उछाला जाये तो यह आशा की जाती है कि वह आधी वार चित्र की ओर गिरेगा और आधी बार पीठ की ओर। यदि एक थैले में 5 लाल व 7 सफेद गेंद हों तो एक लाल गेंद की उस थेले में से निकलने की सम्भावना 5/12 है और सफेद गेंद के निकाले जाने की सम्भावना 7/12 है। इस सिद्धान्त का मानव-जीवन में काफी महत्व है। इसके आधार पर कई सिद्धान्त बनाए गए हैं। निदर्शन पद्धति इसी सिद्धान्त की देन है। सट्टा और बीमा व्यवसाय करने वाले लोग इसी सिद्धान्त को आधार मानकर व्यवहार करते हैं। इस सिद्धान्त की महत्वपूर्ण मान्यता है कि प्रयोग के लिए समूह बड़ा हो। इसी सिद्धान्त पर सांख्यिकी में निम्नलिखित महत्वपूर्ण नियम आधारित हैं-

(1) सांख्यिकीय नियमितता नियम (Law of Statistical Regularity)

(2) महांक जड़ता नियम (Inertia of Large Numbers)

अब हम इन सिद्धान्तों के विषय में विवेचन करेंगे-

3.6 सांख्यिकीय नियमितता नियम (Law of Statistical Regularity)

यह नियम गणित के सम्भावना सिद्धान्त पर आधारित है। इस नियम का अर्थ यह है कि यदि किसी विशाल समूह से अनायास ही पर्याप्त मात्रा में इकाइयाँ चुन ली जायें तो इस प्रकार चुनी गयी इकाइयों में समूह की सब विशेषतायें

पाई जायेंगी। समूह को हम समग्र (universe) कहते हैं और चुनी गई इकाइयों को न्यादर्श (Sample)। प्रसिद्ध अर्थशास्त्री किंग ने इस नियम की परिभाषा इस प्रकार की है-

"गणित के सम्भावना सिद्धान्त के आधार पर बना सांख्यिकीय नियमितता नियम बतलाता है कि यदि किसी विशाल समूह में से दैव निदर्शन द्वारा पर्याप्त बड़ी संख्या में पदों को चुन लिया जाये तो यह लगभग निश्चित है कि इन पदों में औसत रूप से विशाल समूह के गुण होंगे।"

एक उदाहरण द्वारा हम इस नियम को इस प्रकार स्पष्ट कर सकते हैं। यदि हमें किसी महाविद्यालय के विद्यार्थियों की आर्थिक स्थिति के विषय में जानकारी प्राप्त करनी है तो यह आवश्यक है कि सभी विद्यार्थियों से सम्पर्क स्थापित किया जाये। यदि थोड़े-थोड़े विद्यार्थी अलग-अलग कक्षाओं में चुन लिए जायें और उनसे उनकी आर्थिक स्थिति के विषय में जानकारी प्राप्त की जाये तो उससे प्राप्त होने वाले परिणाम कॉलेज के समस्त विद्यार्थियों पर लागू होंगे। कहने का तात्पर्य यह है कि चुने हुए विद्यार्थियों का समूह समस्त विद्यार्थियों का प्रतिनिधित्व कर सकता है।

विशेषतायें- इस नियम में निम्नलिखित विशेषतायें होती हैं-

(1) चुना हुआ अंश समग्र का प्रतिनिधि हो सकता है अर्थात् केवल कुछ इकाइयों के परिणाम समस्त इकाइयों पर लागू किए जा सकते हैं। ऐसा इसलिये होता है क्योंकि यद्यपि बाहरी रूप से एक समूह की प्रत्येक इकाई एक दूसरे से भिन्न दिखाई पड़ती है फिर भी कुछ सामूहिक गुण सभी इकाइयों में समान रूप से पाये जाते हैं।

(2) इस नियम के अनुसार यदि अनायास कुछ इकाइयाँ चुनी जायें तो अच्छी, बुरी, तथा औसत सभी प्रकार की इकाइयों के चुने जाने की समान सम्भावना रहती है। यदि हम एक सिक्के को 100 बार उछालें तो यह सम्भावना रहती है कि लगभग आधी बार चित्र की तरफ तथा आधी बार दूसरी ओर गिरेगा।

(3) जो परिणाम किसी विशेष तथ्य के सम्बन्ध में किसी समय तथा किसी स्थान में लागू होते हैं, देश-काल के अनुसार ही उनमें थोड़ा-बहुत परिवर्तन करके उन्हें सभी समयों तथा सभी स्थानों पर लागू किया जा सकता है। उदाहरणार्थ यदि इन्दौर शहर में जनसंख्या में वृद्धि की दर पिछले कई वर्षों की औसत के आधार पर 10% आई हो तो बहुत सम्भव है कि आगे आने वाले समय में भी औसत दर यही रहे। इस प्रकार भविष्य के वर्षों के लिए भी इन्दौर शहर की जनसंख्या का अनुमान लगाया जा सकता है। परिस्थितियाँ समान होने पर इस औसत वृद्धि को अन्य शहरों पर भी लागू किया जा सकता है।

नियम की सीमायें- (1) न्यादर्श पर्याप्त मात्रा में होना चाहिए। चुनी हुई इकाइयों की संख्या समूह के आकार के अनुसार ही होनी चाहिए। 100 विद्यार्थियों की आर्थिक दशा की जानकारी प्राप्त करने के लिए केवल 10 विद्यार्थियों की जाँच पर्याप्त नहीं होगी।

(2) इकाइयों का चुनाव अनायास ही होना चाहिये और वह पक्षपात-रहित होनी चाहिये।

(3) इकाइयों में चुनाव किया जा रहा है उनमें कोई विशेष भिन्नता नहीं होनी चाहिये।

(4) प्राप्त परिणाम समूह की अन्य इकाइयों पर केवल सामूहिक रूप से ही लागू होंगे। विद्यार्थियों के एक न्यादर्श के परिणाम हर एक विद्यार्थी पर पूर्णरूपेण लागू नहीं किये जा सकते हैं। वे सब पर औसत रूप से ही लागू किये जा सकते हैं।

नियम की उपयोगिता- इस नियम का, जैसा कि ऊपर बतलाया गया है, सांख्यिकीय अनुसंधान में व्यापक रूप से प्रयोग होता है। किन्तु मुख्य रूप से इस नियम का प्रयोग निदर्शन प्रणाली तथा आंतर्गणन और बाह्यगणन (Interpolation and Extrapolation) में किया जाता है। निदर्शन प्रणाली में कुछ इकाइयाँ न्यादर्श के रूप में चुन ली जाती हैं तथा उनके परिणाम समस्त समूह पर लागू किये जाते हैं। आंतर्गणन एवं बाह्यगणन में प्राप्त समंकों का अध्ययन करके उसके आधार पर बीच के लुप्त (अप्राप्त) मूल्यों को मालूम किया जाता है तथा भविष्य में सम्भावित समंकों का अनुमान किया जाता है। बीमा व्यवसायी मृत्यु दर के आधार पर प्रीमियम की राशि निर्धारित करते हैं। इसी तरह अन्य सांख्यिकीय अनुसंधानों में इसका प्रयोग होता है।

3.7 महांक जड़ता नियम (Law of Inertia of Large Number)

यह नियम सांख्यिकीय नियमितता नियम पर आधारित है। प्रकृति में बाह्य रूप में भिन्नता दृष्टिगोचर होती है किन्तु आन्तरिक रूप में उसमें एकरूपता होती है इसलिये सम्पूर्ण में से थोड़ा अंश जाँचकर ही हम सम्पूर्ण के विषय में

निर्णय दे देते हैं। यह तो हमें विदित ही है कि न्यादर्श की संख्या जितनी अधिक होगी, परिणाम प्राप्त करने में शुद्धता की आशा भी उतनी ही अधिक होगी। इसका मुख्य कारण यह है कि यदि अनुसंधान के क्षेत्र में विभिन्न आकार वाले अंक हों तो यह सम्भावना रहेगी कि कुछ अंक यदि एक दिशा में परिवर्तन करते हैं तो कुछ अन्य अंक विपरीत दिशा में। परिणामस्वरूप परिवर्तन का प्रभाव औसतन साधारण या नाममात्र का रहेगा। ज्यों-ज्यों न्यादर्श की संख्या बढ़ती जायेगी अशुद्धि उतनी ही कम होती जायेगी। अतः यह नियम बतलाता है कि अधिक या बड़े-बड़े अंकों में एक जड़ता होती है अर्थात् उनमें परिवर्तन बहुत धीमा और लगभग नहीं के बराबर होता है। इसके विपरीत छोटे अंकों में परिवर्तन काफी होता है।

डॉ. बॉउले ने यह लिखा है कि "बड़ी-बड़ी संख्याओं तथा उनके माध्यों में स्थिरता होने के कारण ही माप सम्भव हो सकता है।" उन्होंने यह भी बतलाया है कि "एक समूह की व्यक्तिगत इकाइयाँ लगातार बदलती हैं किन्तु समस्त समूह बहुत धीरे-धीरे परिवर्तित होता है।"

सारांश यही है कि बड़ी संख्यायें अपेक्षाकृत अधिक स्थिर तथा अपरिवर्तनशील होती हैं। इसे हम कुछ उदाहरणों द्वारा स्पष्ट कर सकते हैं। एक व्यक्ति की आय में परिवर्तन एक निश्चित अवधि में बहुत अधिक हो सकता है। किन्तु जब हम एक परिवार की आय लेंगे तो उसमें व्यक्ति की आय की अपेक्षा कम परिवर्तन दृष्टिगोचर होगा। जब हम किसी राज्य या देश की आय के परिवर्तन की दर की ओर देखेंगे तो यह अपरिवर्तित या स्थिर ही रहेगी। इसी प्रकार जनसंख्या की वृद्धि के विषय में भी कहा जा सकता है। एक परिवार में संख्या की वृद्धि बहुत अस्थिर और परिवर्तनशील होती है। किन्तु जब हम गाँव या शहर की वृद्धि दर को देखेंगे तो यह अपेक्षाकृत अधिक स्थिर होगी और सम्पूर्ण देश में वृद्धि दर लगभग अपरिवर्तनशील होती है। एक व्यापारी का प्रतिदिन का विक्रय बदलता रहता है किन्तु वर्ष भर में कुल व औसत विक्रय में स्थिरता रहती है।

इस नियम के कार्यान्वित होने की सबसे बड़ी शर्त यही है कि समूह बड़ा होना चाहिये। समूह जितना बड़ा होगा, नियम उतना ही अधिक प्रभावशील होगा। किन्तु यह नियम सदैव ही लागू होगा, ऐसी बात नहीं है। अधिक समयान्तर पर पर्याप्त परिवर्तन दिखाई दे सकता है। उपरोक्त उदाहरणों में यदि हम जनसंख्या में वृद्धि की दर या विक्रय में परिवर्तन की दर को 10-15 वर्षों बाद देखें तो उसमें परिवर्तन स्वाभाविक रूप से हो सकता है।

नियम की उपयोगिता (Utility of the Law)- इस नियम का सांख्यिकीय अनुसन्धान में अत्यन्त महत्व है। सांख्यिकीय अनुसन्धान में जो निष्कर्ष निकाले जाते हैं उन्हें भविष्य की संख्याओं पर भी लागू किया जाता है। यदि मूल्यों और उत्पादन में स्थिरता न हो तो सांख्यिकीय अनुसन्धान से कोई निष्कर्ष ही नहीं निकलें। यदि जनसंख्या में वृद्धि की दर हर वर्ष बदलती रहे तो उसकी गणना से लाभ ही क्या होगा?

इससे यह स्पष्ट है कि भविष्य के लिये अनुमानों का आधार यही नियम है और इसी से इस नियम की उपयोगिता स्पष्ट है।

निदर्शन का महत्व (Importance of Sampling)

निदर्शन का महत्व हमें जीवन के हर क्षेत्र में देखने को मिलता है। हमारे जीवन की कई क्रियायें इस प्रणाली पर आधारित हैं। जब हम कोई वस्तु बाजार से अधिक मात्रा में खरीदना चाहते हैं तो पहले हम उस वस्तु का थोड़ा–सा भाग नमूने के रूप में लेकर यह निर्णय करते हैं कि वह वस्तु खरीदने योग्य है अथवा नहीं। किसी नगर या राज्य के कुछ भागों की विशेषताओं को देखकर उस नगर या राज्य के विषय में हम अपने विस्तृत विचार बना सकते हैं। परीक्षा में कुछ प्रश्न पूछकर, परीक्षक, विद्यार्थी के ज्ञान के विषय में अपना निर्णय दे देता है। किसी व्यक्ति के साथ थोड़ा–सा सम्पर्क स्थापित होने से उसके सम्पूर्ण व्यक्तित्व और चरित्र के विषय में अपनी राय कायम की जाती है। इस प्रकार, निदर्शन प्रणाली हमारे जीवन के विभिन्न क्षेत्रों में बड़ी ही उपयोगी सिद्ध होती है। इसी कारण इसका व्यापक प्रयोग किया जाता है। यहाँ हम केवल इसके विभ्रम के विषय में विस्तृत अध्ययन करेंगे।

3.8 निदर्शन के विभ्रम (Errors of Sampling)

न्यादर्श (Sample) समग्र (Universe) का प्रतिनिधि होता है। जो सूचना इसके द्वारा प्राप्त होती है उसमें समग्र की विशेषतायें उपलब्ध होती हैं। किन्तु हमें न्यादर्श मध्यक को समग्र मध्यक के बिल्कुल समान नहीं समझ लेना चाहिये। न्यादर्श लेने की रीति चाहे कितनी ही उपयुक्त क्यों न हो, समग्र व न्यादर्श में थोड़ा बहुत अन्तर अवश्य रहता है। ऐसा कम सम्भव है कि न्यादर्श समग्र का पूर्ण रूप से प्रतिनिधित्व करेगा। एक ही समग्र में से यदि कई

न्यादर्श अलग- अलग किये जायें तो उनमें भी आपस में अन्तर होता है। **न्यादर्श और समग्र के मध्य इन कमियों को ही निदर्शन के विभ्रम कहते हैं** । गणित की सहायता से इसका अनुमान लगाया जा सकता है । प्रत्येक निदर्शन विधि की सापेक्ष शुद्धता का भी अनुमान लगाया जा सकता है । प्रत्येक निदर्शन विभ्रम के ज्ञान के कारण निदर्शन की रीति अधिक विश्वसनीय हो गई है और इसकी उपयोगिता का क्षेत्र बहुत व्यापक हो गया है ।

न्यादर्श के विभ्रम के दो कारण हो सकते हैं-

(1) न्यादर्श के चुनाव की पद्धति में व्यक्तिगत झुकाव (Bias) का प्रभाव ।

(2) सम्भावना (Chance), क्योंकि किसी भी न्यादर्श में समग्र की सभी विशेषतायें नहीं होतीं ।

प्रथम दोष को न्यादर्श विधि के उपयुक्त प्रयोग द्वारा दूर किया जा सकता है ।

3.9 प्रमाप विभ्रम (Standard Error)

यदि किसी बड़े समग्र में से एक निश्चित मात्रा के कई न्यादर्श दैव-निदर्शन (Random sampling) की रीति से लिये जायें और इनमें से प्रत्येक का समान्तर माध्य निकाला जाये, तो इन माध्यों के आधार पर जिस आवृत्ति वितरण (Frequency distribution) की रचना होती है, वह निदर्शन वितरण (Sampling Distribution) कहलाता है । अब, यदि हम इन सभी मध्यकों का समान्तर माध्य निकालें और उसकी सहायता से मध्यकों को मूल्य मानते हुये प्रमाप विचलन निकालें तो यह प्रमाप विचलन ही मध्यकों का प्रमाप विभ्रम (Statndard Error of the mean) कहलायेगा । यह उस सीमा का निर्धारण करता है जिसके अन्दर ही दैव-निदर्शन से प्राप्त न्यादर्श समग्र से दूर रहता है । प्रमाप विभ्रम समग्र की विचरणता (Variability) और न्यादर्श के आकार पर निर्भर होता है । इसकी सीमा ज्ञात करने में सामान्य वक्र काफी सहायक हैं । जिस अनुमान में प्रमाप विभ्रम जितना कम होता है वह अनुमान उतना ही शुद्ध और विश्वसनीय होता है ।

बड़े न्यादर्शों में प्रमाप विभ्रम
(Standard Error in large Samples)

ऊपर यह बताया जा चुका है कि बड़े न्यादर्शों के सामान्य वक्र की सहायता से भी प्रमाप विभ्रम का अनुमान लगाया जा सकता है । सम्भावित सिद्धान्त में सामान्य वक्र का वर्णन करते समय यह बतलाया जा चुका है कि माध्य के ± 1 प्रमाप विचलन के क्षेत्र में सम्पूर्ण क्षेत्र का 68.27% आता है। इसी भाँति ± 2 प्रमाप विचलन और ± 3 प्रमाप विचलन में क्रमशः 95.45% और 99.73% क्षेत्र सम्मिलित होता है। इस तरह $\pm 3\,\sigma$ में लगभग सभी पदों का समावेश हो जाता है ।

समान्तर माध्य का प्रमाप विभ्रम (Standard Error of the Mean) - समान्तर माध्य या मध्यक का प्रमाप विभ्रम निकालने के लिये समग्र का प्रमाप विचलन निकालना पड़ता है । यदि यह सम्भव न हो तो न्यादर्श के प्रमाप विचलन से भी सूत्र में कुछ संशोधन करके काम चलाया जा सकता है । इसके सूत्र इस प्रकार हैं-

(अ) जब समग्र का प्रमाप विचलन ज्ञात हो-

$$\sigma_a \text{ (Standard Error of the Mean)} = \frac{\sigma \text{ (universe)}}{(\sqrt{n})}$$

(ब) जब न्यादर्श का प्रमाप विचलन ज्ञात नहीं हो-

$$\sigma_s = \frac{\sigma_s}{(\sqrt{N-1})} \quad \text{or} \quad \sigma_a = \frac{\sigma_a}{(\sqrt{N})}$$

where

σ_a = Standard Error of the mean

σ_s = Standard Error of the sample

N = Number of the items

समान्तर माध्य के प्रमाप विभ्रम की सहायता से यह जाना जाता है कि समग्र के माध्य की किसी सीमा के भीतर होने की सम्भावना है । हमें यह मालूम है कि सामान्य वक्र का 95% क्षेत्रफल $M \pm 2\sigma$ के भीतर आता है । अतः इसकी सहायता से हम निम्न निष्कर्ष निकाल सकते हैं-

NOTES

यदि किसी निदर्शन का समान्तर माध्य 10 हो और प्रमाप विभ्रम .002 हो तो यह सम्भावना है कि यदि हम इस प्रकार के 100 निदर्शनों की माध्य और उनकी प्रमाप विभ्रम निकालें तो 95 प्रयत्नों में समग्र की माध्य निदर्शन की माध्य $\pm 2\sigma$ सीमा के भीतर ही रहे ।

इस प्रमाप विभ्रम का उपयोग उदाहरणों में एक उपकल्पना (Hypothesis) देकर उसकी जाँच करने के लिये किया जाता है । उपकल्पना यह रहती है कि क्या निदर्शन उसी समग्र में से निकाला गया है ? इस उपकल्पना की जाँच का सूत्र निम्न है–

$$T = \frac{M - M_a}{\sigma}$$

where, T = Deviation of the mean of the sample from the mean of the hypothesis.

M = Mean of the sample.

M_h = hypothetical Mean of the Universe.

σ_m = Standard error of the mean.

Illustration 1. The average weight of 900 students is 50 kg. Find the standard error of the mean if the Standard Deviation of the universe is 9 kilograms.

Solution– Standard Error of the Mean, or $\sigma_a = \frac{\sigma u}{\sqrt{(n)}}$

$$= \frac{9}{\sqrt{(900)}} = \frac{9}{30} = .3 \text{ kilograms.}$$

Illustration 2. The average income of 100 persons is Rs. 50. Calculate the Standard Error of the Mean if the Standard Deviation of the sample is Rs. 5.

Solution- Standard Error of the Mean of $\sigma_s = \frac{\sigma_3}{\sqrt{(N-1)}}$

$$= \frac{5}{\sqrt{(100-1)}} = \frac{5}{\sqrt{99.}} = .5 \text{ Rs.}$$

Illustration 3. To Know the Mean weight of all the 10 year old boys in the State of Madhya Pradesh, a sample of 625 is taken. The mean weight of this sample is found to be 67 pounds with a standard deviation of 15 pounds. Can you draw any inference from it about the Mean weight of the universe ?

Solution– Standard Error of the sample Mean weight

$$= \frac{\sigma_s}{\sqrt{(n)}} = \frac{15}{\sqrt{(625)}} = \frac{15}{25} = \frac{3}{5} = .6$$

Assuming that the conditions of simple sampling hold good, the Mean Weight of the universe in all probability would be lying within the limit set by 3 times this Standard Error to the Mean of the sample.

So, the Mean weight of all 10 year old boys in the State of Madhya Pradesh lies between £ (67 ± 3 × .6) or (67 ± 1.8) i.e. £ 68.8 and £ 65.2

Illustration 4. Suppose that it has been determined that the average pulse rate of males in the 20–25 year age–group is 72 beats per minute and that the S.D. is 9.5 beats per minute. If a group of 55 distance–runners all in the given age– group, examined and found to have an average pulse rate of 65, should this be regarded as a significant deviation from the general average ?

Solution- $\sigma_a = \frac{\sigma_u}{\sqrt{(N)}} = \frac{9.5}{\sqrt{(55)}} = \frac{9.5}{7.4} = 1.28$ beats per minute.

The difference between the two average pulse rate is 7 beats per minute, which is 5.47 times this standard error. Hence, the deviation of the average pulse rate of distance runners from the general average is significant.

Illustration 5. Suppose that for tyres of a certain factory, the mean mileage is 15190 and the standard deviation is 1250 miles.

Suppose further that a sample of 900 tyres shows a mean mileage of 15242 miles with a standard deviation of 1212 miles. Ascertain if the Sample Mean represents a significant divergence from the mean of the universe.

Solution–

$$\sigma_a = \frac{\sigma_p}{\sqrt{(n)}} \text{ or } \frac{\sigma^a}{\sqrt{(n)}} = \frac{1250}{\sqrt{(900)}} = \frac{1250}{30} = 41.67 \text{ Miles.}$$

The observed difference between the two Means of mileage is 15242–15190 or 52 miles. This is only 1.2 times this standard error and so could have arisen through the fluctuations of simple sampling. Hence it can be said that the divergence of the sample mean from the mean of the universe is not significant.

Note– *When Standard Deviation of the universe is given, then the S.D. of the sample should not be used.*

मध्यका, चतुर्थक आदि का प्रमाप विभ्रम
(Standard Error of the Median, Quartiles etc.)

निदर्शन की मध्यकाओं में उच्चावचन अधिक होते हैं और इनके अपकिरण माध्य के अपकिरण से 25% अधिक रहते हैं। इनके सूत्र इस प्रकार हैं–

(i) Standard Error of the Median $= 1.25331 \frac{\sigma}{\sqrt{(N)}}$

(ii) " " " the Quartiles $= 1.36263 \frac{\sigma}{\sqrt{(N)}}$

(iii) " " " Mean Deviation $= .6028 \frac{\sigma}{\sqrt{(N)}}$

(iv) " " " Quartile Deviation $= .78672 \frac{\sigma}{\sqrt{(N)}}$

(v) " " " Variance $= \sigma^2 \sqrt{\left(\frac{2}{N}\right)}$

(vi) " " " Coefficient of Correlation $= \frac{1 - r^2}{\sqrt{(N)}}$

(vii) " " " Regression coefficient $= \frac{\sigma_x \sqrt{(1 - r^2)}}{\sigma_y \sqrt{(N)}}$

(viii) Standard Error of the Regression estimate =
(y on x) $= \sigma_y \sqrt{(1 - r^2)}$

(ix) Standard Error Regression estimate $=$ (x on y) $= \sigma_x \sqrt{(1 - r^2)}$

(x) Standard Error Coefficient of Association

$$\frac{(1 - Q^2)}{2} \sqrt{\left(\frac{1}{(AB)} + \frac{1}{(Ab)} + \frac{1}{(aB)} + \frac{1}{(ab)}\right)}$$

Illustration 6. Compute standard errors of Median, Quartiles, Mean Deviation, Variance and Quartile Deviation, if Standard Deviation, is 5 and number of cases included in a sample is 100 presuming that the sample has been drawn from a normal population.

NOTES

Solution–

(a) Standard Error of Median $= 1.25331 \frac{\sigma}{\sqrt{(N)}} = 1.25331 \frac{5}{\sqrt{100}} = .626655$

(b) Standard Error of Q1 and Q2 $= 1.36263 \frac{\sigma}{\sqrt{(N)}} = 1.36263 \frac{5}{\sqrt{(N)}} = .681315$

(c) Standard Error of M.D. $= .6028 \frac{5}{\sqrt{(10)}} = .3014$

(d) Standard Error of Variance $= \sigma^2 \sqrt{\left(\frac{2}{N}\right)}$

$$= (5)^2 \sqrt{\left(\frac{2}{100}\right)} = 25 \sqrt{\left(\frac{1}{50}\right)}$$

$$= 25 \times .14 = 3.50$$

(e) Standard Error of Q.D. $= .78672 \frac{\sigma}{\sqrt{(N)}} = .78672 \frac{5}{\sqrt{(100)}} = .39336$

Illustration 7. To study the correlation between the stature of the father and the stature of the son, a sample of 625 is taken from the universe of fathers and sons. The sample study gives the correlation between the two to be + .67. Within what limits does it hold true for the universe ?

Solution– Standard Error of the Coefficient of Correlation

$$= \frac{2 - r^2}{\sqrt{(N)}} = \frac{1 - (.67)^2}{\sqrt{(625)}} = \frac{1 - .4489}{25} = \frac{.5511}{25} = .022$$

Hence, the correlation in the universe most probably lies within $.67 \pm 3 \times .022 = .67 \pm .066$, i.e., .60 and .74 .

Illustration 8. Find out the regression Coefficient of Hapur prices over Karachi prices from the following data and also calculate its standard error.

Assuming 'n' to be 900	**Hapur Rs.** (x)	**Karachi Rs.** (y)
Average price per md. of wheat	18	12
Standard Deviation	3	2

' r' between price at Hapur and Karachi = .67

Solution– Regression Coefficient of Hapur price (x) Karachi prices (y)

$$bxy = r \frac{\sigma_x}{\sigma_y} = .67 \frac{3}{2} = 1.005$$

The standard error of this regression Coefficient is

$$= \frac{\sigma_x \sqrt{(1 - r^2)}}{\sigma_y \sqrt{(N)}} = \frac{3\sqrt{1 - (.67)^2}}{2 \times \sqrt{(900)}} = .028$$

Illustration 9. Two groups of children, one belonging to the professional class 125 in number and the other belonging to the labour classes 124 in number are compared and the following results are obtained–

	Poor Children %	**Well-to-do Children %**
Below normal weight	55	13
Above normal weight	11	48

Find the coefficient of Association between the weight and the social status. Also calculate the standard error of this Coefficient of Association.

Solution–Denotion poor children by 'A' and well–to–do by 'a' normal weight by 'B' and above normal weight by 'b', we get

(AB) = (55) aB = (13) Ab = (11) and ab = (48)

Substituting these values in the formula, we get

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)}$$

$$= \frac{(55)\times(48) - (11)\times(13)}{(55)\times(48) + (11)\times(13)} = \frac{2640 - 143}{2640 + 143} = \frac{2497}{2783} = +.89$$

The standard error of the coefficient of Association

$$= \frac{1 - Q^2}{2}\sqrt{\left(\frac{1}{AB} + \frac{1}{Ab} + \frac{1}{aB} + \frac{1}{ab}\right)}$$

$$= \frac{1 - (.89)^2}{2}\sqrt{\left(\frac{1}{55} + \frac{1}{11} + \frac{1}{13} + \frac{1}{48}\right)} = 0.1477$$

Illustration 10. (A) The following table shows frequency distribution of yield of wheat (in quintals) per acre in 2,000 irrigated fields selected at random. Calculate Sample error of the mean :–

निम्नांकित सारणी में दैव निदर्शन से चुने गए 2,000 सिंचित खेतों की प्रति एकड़ गेहूं की उपज (क्विटल) का आवृत्ति वितरण दर्शाया गया है। समान्तर माध्य का न्यादर्श विभ्रम निकालिए :–

Limit (सीमा) Qtls.	0–2	2–4	4–6	6–8	8–10	10–12	12–14	14–16	16–18
No. of Fields (खेतों की संख्या) :	90	368	562	456	310	158	44	10	2

Solution :

Limit IAj 9|I Qlts.	M.V.	No of fields (f)	(dx 9)	fdx	fd^2
0-2	1	90	–8	–720	5760
2-4	3	368	–6	–2208	13248
4-6	5	562	–4	–2248	8992
6-8	7	456	–2	912	1824
8-10	9	310	0	0	0
10-12	11	158	+2	316	632
12-14	13	44	+4	176	704
14-16	15	10	+6	60	360
16-18	17	2	+8	16	128
		2,000		-5520	31648

$$S.D.(s) = \sqrt{\frac{\Sigma fd^2x}{n} - \left(\frac{\Sigma fdx}{n}\right)^2}$$

$$= \sqrt{\frac{31648}{2000} - \left(\frac{-5520}{2000}\right)^2}$$

$$= \sqrt{15.82 - (2.76)^2}$$

$$= \sqrt{15.82 - 7.62}$$

$$= \sqrt{8.2}$$

$$= 2.86$$

Sample error of the mean

$$= \frac{S.D.(s)}{\sqrt{n-1}}$$

$$= \frac{2.86}{\sqrt{2000-1}}$$

$$= \frac{2.86}{44.7}$$

$$= 0/062$$

NOTES

न्यादर्श माध्यों के अन्तर का प्रमाप विभ्रम
(Standard Error of the Difference of Sample means)

यदि हमें दो न्यादर्शों से दो माध्य–मूल्य प्राप्त हुये हों और हमें यह ज्ञात करना हो कि इन दो मूल्यों में महत्वपूर्ण अन्तर हैं या नहीं अथवा क्या वे दोनों एक ही समग्र अथवा ऐसे समग्रों से प्राप्त हुये हैं जिनका समान्तर माध्य और प्रमाप विचलन समान ही हैं। यहाँ हम दोनों न्यादर्श माध्यों के अन्तर के प्रमाप विभ्रम की गणना करेंगे और फिर यह देखेंगे कि क्या उन दोनों के माध्यों का अन्तर इस प्रमाप विभ्रम के तिगुने से अधिक है अथवा कम। यदि वास्तविक अन्तर इस सीमा से अधिक है तो इसे महत्वपूर्ण माना जायेगा अन्यथा अन्तर न्यादर्श के उच्चावचन के कारण से माना जायेगा।

दो न्यादर्श माध्यों के अन्तर का प्रमाप विभ्रम ज्ञात करने के लिए निम्नलिखित सूत्रों का प्रयोग किया जाता है–

(i) जब समग्र का प्रमाप विचलन ज्ञात न हो, $\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)}$

(ii) जब समग्र का प्रमाप विचलन ज्ञात हो, $\sqrt{\left[\sigma_u^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\right]}$

(iii) जब एक न्यादर्श के माध्य की तुलना दोनों न्यादर्शों के मिश्रित माध्य से करनी हो –

$$\sqrt{\left(\sigma_u^2\,\frac{n_2}{n_1(n_1+n_2)}\right)}$$

इन अवस्थाओं में हम न्यादर्श की साधारण मान्यताओं को स्वीकार करते हैं।

यदि दोनों न्यादर्श (Samples) ऐसे समग्र से लिये गये हों जिसके माध्य सह–सम्बन्धित हैं तो न्यादर्श माध्यों के अन्तर का प्रमाप विभ्रम निम्न सूत्र द्वारा ज्ञात होगा–

$$\sqrt{\left(\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}-2r\,\frac{\sigma_1\times\sigma_2}{n_1\times n_2}\right)}$$

सूत्रों का प्रयोग निम्न उदाहरणों द्वारा स्पष्ट हो जायेगा–

[I] जब समग्र का प्रमाप विचलन न दिया गया हो

Illustration 11. A random sample of 200 villages was taken from Indore district and the average population per village was found to be 485 with a standard deviation of 50. Another random sample of 200 village from the same district gave an average population of 510 per village with a S.D. of 40. Is the difference between the average of the two samples statistically significant? Give reasons.

Solution–Standard Error of the difference of the two means, or

$$\sigma m_1 - m_2 = \sqrt{\left[\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right]} = \sqrt{\left[\frac{(50)^2}{200}+\frac{(40)^2}{200}\right]}$$

$$= \sqrt{(20.5)} = 4.53$$

The observed difference of the two average populations per village is (510–485) or 25 which is thus 5.5 times this standard error and so could not have arisen through the fluctuations of simple sampling. Hence, the difference between the average of the two samples is statistically significant.

Illustration 12. A random sample of 1000 farms in a certain year gives an average yield of wheat of 2000 lbs. per acre with a standard deviation of 192 lbs. A random sample of 1000 farms in the following year gives an average yield of 2100 lbs. per acre with a standard deviation of 224 lbs. Show that these data are inconsistent with the hypothesis that the average yields in the country as a whole were the same in two years.

Solution-

$$\sigma m_1 - m_2 = \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)} = \sqrt{\left[\frac{(192)^2}{1000} + \frac{(224)^2}{1000}\right]}$$

$$= \sqrt{(36.864 + 50.176)} = \sqrt{(87.04)} = 9.3$$

The actual difference between the two averages is of 100 lbs. which is nearly 10.5 times of this standard error. Hence, it is inconsistent.

[II] जब समग्र का प्रमाप विचलन ज्ञात हो

Illustration 13. The mean produce of wheat of a sample of 100 fields comes to 200 lbs. per acre with a standard deviation of 10 lbs. Another sample of 150 fields gives the mean at 220 lbs. with a standard deviation of 12 lbs. Assuming the standard deviation of the yield at 11 lbs. for the universe; find out if there is significant difference between the mean yields of the two samples.

Solution- Supposing the samples are independent and come from the same universe, the standard error of the difference of the mean yield would be–

$$\sigma m_1 - m_2 = \sqrt{\left[\sigma_u^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]} = \sqrt{(11)^2\left[\left(\frac{1}{100} + \frac{1}{150}\right)\right]} = 1.42$$

The observed difference between the two means is 20 lbs. which is more than three times the standard error of the difference of the mean. Hence, the difference is significant and could not have arisen due to fluctuations of sampling.

[III] जब प्रथम न्यादर्श (Sample) के माध्य की तुलना दोनों न्यादर्शों के मिश्रित माध्य से करनी हो

Illustration 14. A random sample of 100 villages in a district gives the mean population of 500 persons per village. Another sample of 150 villages from the same district gives the mean at 504 . If the standard deviation of the mean population of village in that district is 20. Find out, if the mean of the first sample is significantly different from the combined of the two samples taken together.

Solution. The combined mean of the two samples.

$$\frac{(100\times500) + (150\times504)}{100 + 150} = 502.4$$

The difference between the first sample mean and the combined mean is thus (502.4–500) = 2.4.

The standard error of the difference of the first sample mean and the combined mean is.

$$= \sqrt{\left(\sigma_u^2 \frac{n^2}{n_1(n_1+n_2)}\right)} = \sqrt{\left[(20)^2\left(\frac{150}{100\,(100 + 150)}\right)\right]} = \sqrt{(2.4)} = 1.549$$

The observed difference is less than three times the standard error of the difference and as such it could have arisen due to fluctuations of sampling, the difference is not significant.

[IV] सह-सम्बन्ध वाले न्यादर्शों में प्रमाप विभ्रम

Illustration 15. In an intelligence test administered to 60 fathers and their 100 children; the following results were obtained–

NOTES

Father's Mean Score	114;	S.D. = 13
Son's ,, ,,	110;	S.D. = 11

Assuming the 'r' between the two to be = 75, calculate the standard error of the difference of the two means and state whether the difference is significant ?

Solution- Since the two samples are related, the standard error of the difference of their Mean Scores, is

$$\sigma_{m_1 - m-2} = \sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} - 2r \times \frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}\right)}$$

$$= \sqrt{\left(\frac{(13)^2}{60} + \frac{(11)^2}{100} - 2 \times 7.5 \times \frac{13}{60} + \frac{11}{100}\right)} = \sqrt{(3.9793)} = 1.99.$$

Thus, the standard error of the difference of the two Mean Scores is 1.99. The observed difference is (114–110) or 4 which is only twice this standard error, and so could have arisen through the fluctuation of the simple sampling. Hence, the difference is not significant.

गुण समंक का सरल निदर्शन
(Simple Sampling of Attributes)

निदर्शन दो प्रकार के हो सकते हैं– (i) सरल (ii) जो सरल न हो । यदि किसी घटना के होने की सम्भावना स्वतंत्र हो तो निदर्शन सरल होता है और यदि घटना के होने की सम्भावना निर्भर (Dependent) है तो निदर्शन सरल नहीं कहा जाता ।

सार्थकता की परीक्षा (Test of significance)

सामान्यतः हर न्यादर्श की सहायता से ज्ञात किये गये सांख्यिकीय माप भिन्न–भिन्न होते हैं । इस भिन्नता के निश्चित कारण बतलाना बहुत कठिन है । हर न्यादर्श की आवृत्ति अलग–अलग होती है । अतः उसकी सहायता से निकाले गये समान्तर माध्य, प्रमाप विचलन आदि भी अलग–अलग होते हैं । यह भिन्नता उच्चावचन कहलाती है । यह उच्चावचन सार्थक (significant) है अथवा नहीं, इसकी जांच निदर्शन के सिद्धान्त, जो कि सम्भावना के सिद्धान्तों पर आधारित है, की सहायता से करनी होगी ।

गुणसमंक के सरल निदर्शन का समान्तर माध्य और प्रमाप–विचलन

दिये गये समग्र में से यदि कई न्यादर्श दिये गये जिनमें कुछ पद हों, और न्यादर्श में एक घटना के होने की सम्भावना (p) निकाली जाये तो उस घटना के न होने की सम्भावना (q)= 1–p होगी । इसका समान्तर माध्य होगा

$$M \text{ (or a)} = np$$

और पदमाला का प्रमाप विचलन होगा

$$\sigma \text{ (S.D.)} = \sqrt{(npq)}$$

यदि हम हर न्यादर्श में घटना होने की सम्भावना की संख्या न लिखकर उस घटना के होने का प्रमाप या अनुपात (Proportion) लिखें और उनका प्रमाप विचलन निकालें तो उसका सूत्र इस प्रकार होगा -

$$\text{S.D.} = \sqrt{\left(\frac{pq}{n}\right)}$$

यदि दो न्यादर्श अलग-अलग समग्रों से लिये जायें तो उनके प्रमाप (Proportion) को (p_1) और (p_2) कहा जायेगा । इसके अन्तर की सार्थकता जानने के लिए निम्न सूत्र का प्रयोग किया जायेगा–

$$\text{Standard Error in one sample} = \sqrt{\left(\frac{p_0 q_0}{n_1}\right)}$$

$$\text{Standard Error in the second sample} = \sqrt{\left(\frac{p_0 q_0}{n_2}\right)}$$

Standard Error of the difference $= \sqrt{\left\{ p_0 q_0 \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right\}}$

Illustration 16. Balls are drawn from a bag containing equal number of black and white balls, each ball being returned before drawing another. In 2250 drawings, 1018 black and 1232 white balls have been drawn. Do you suspect some bias on the part of the drawer ?

Solution– The expectation of drawing a white ball in one draw is 1/2. Since the bag contains equal number of black and white balls–

In 2250 drawings, the expected number of white balls is 1125.

But the number of white balls drawn is 1232.

Thus the difference from the expected result is 107.

the Standard Deviation of Simple Sampling is

$$\sigma = \sqrt{(npq)} = \sqrt{\left(\frac{1}{2} \times \frac{1}{2} \times 2250 \right)} = 32.7$$

The divergence from the expectation is thus about 4.5 times of this and hence it is not probable that it arose due to fluctuations of sampling. Explanation of this deviation must be sought somewhere else and it seems reasonable to suspect that the drawer was biassed.

Illustration 17. A coin is tossed 400 times and it turns up head 216 times. Discuss whether the coin may be an unbiased one.

Solution–If the coin is unbiased, the chance that it will turn up head in one toss is 1/2

Therefore, the expected number of heads in 400 tosses is 200. The observed number is 216 heads, and thus it is in excess of the expected number by 16.

The S.D. of Simple Sampling is

$$= -(npq) = \sqrt{\left(\frac{1}{2} \times \frac{1}{2} \times 400 \right)} = 10$$

The deviation observed is thus only 1.6 times this S.D. and it is therefore probable that it arose as a Sampling fluctuation. Therefore, the coin can be said to be an unbiased one.

Illustration 18. A random sample of 500 apples was taken from a large consignment and 65 were found to be bad. Estimate the proportion of the bad apples in the consignment as well as the Standard error of the estimate. Deduce that the percentage of bad apples in the consignment almost certainly lies between 8.5 and 17.5.

Solution–Proportion of bad apples in the sample or (p)

$$= \frac{65}{500}$$

and so (q) $= 1 - \frac{65}{500} = \frac{435}{500}$

Standard Error of the proportion of bad apples

$$\sqrt{\left(\frac{pq}{n} \right)} = \sqrt{\left(\frac{65}{500} \times \frac{435}{500} \times \frac{1}{500} \right)} = 0.015 \text{ or } 1.5\%$$

The percentage of bad apples in the sample is

$$\left(\frac{65}{500} \times 100\right) = 13\%$$

So, the limits of bad apples in the consignment would be $13 \pm (3 \times 1.5) = 8.5$ and 17.5.

Illustration 19. In a locality containing 18,000 families, a sample of 840 families was selected at random. Of these 840, 206 families were found to have a monthly income of Rs. 50 or less. It is desired to estimate how many out of 18,000 families have a monthly income of Rs. 50 or less. Within what limits would you place your estimate ?

Solution-Proportion of families having a monthly income of Rs. 50 or less i.e.,

$$(p) = \frac{206}{840} = \frac{103}{420} \text{ and so } (q) = \frac{317}{420}$$

Standard error of the proportion of the families having a monthly income of Rs. 50 or less.

$$= \sqrt{\left(\frac{pq}{n}\right)} = \sqrt{\left(\frac{103}{420} \times \frac{317}{420} \times \frac{1}{420}\right)}$$

$$= 0.015 \text{ or } 1.5\%$$

Whatever be the percentage of families who have a monthly income of Rs. 50 or less, a simple sample should give a percentage within three times standard error.

Hence, taking $\left(\frac{103}{420} \times 100\right)$ or 24.5% to be the estimate of such families we have the limits as $24.5 \pm (3 \times 1.5) = 20\%$ and 29% approximately.

छोटे न्यादर्शों में प्रमाप विभ्रम
(Standard Error in Small Samples)

बड़े आकार के न्यादर्शों के विषय में जो विश्लेषण ऊपर प्रस्तुत किया गया है, वह छोटे न्यादर्शों पर लागू नहीं होता है। जहाँ पदों की संख्या 30 (कुछ विद्वानों ने इसे 50 माना है) से भी कम होती है तो उन्हें छोटे न्यादर्श माना जाता है। न्यादर्शों का आकार जब छोटा होता है तो समग्र का प्रतिनिधित्व ठीक प्रकार से नहीं होता है। जिन मान्यताओं पर हम बड़े न्यादर्शों के समान्तर माध्य, माध्य विचलन आदि का विश्लेषण करते हैं वे मान्यताएँ छोटे न्यादर्शों में नहीं पाई जाती हैं। बड़े न्यादर्शों में यह मान लिया जाता है कि दैव निदर्शन से प्राप्त वितरणों में सामान्य वितरण (Normal distribution) की विशेषताएँ रहती हैं और न्यादर्शों के अनुसंधान से जो सांख्यिकीय माप (समान्तर माध्य आदि) प्राप्त होते हैं, वे समग्र के माप के निकट होते हैं और उन्हीं का प्रयोग प्रमाप-विभ्रम की गणना में किया जा सकता है। लेकिन छोटे न्यादर्श में इसे विश्वसनीय नहीं समझा जा सकता है। संभावना के सामान्य नियम भी छोटे न्यादर्शों में खरे नहीं उतरते हैं क्योंकि महांक जड़ता नियम का अभाव उनमें बाधक होता है। अत: छोटे न्यादर्शों के लिए नई प्रक्रियाएँ तथा विधियाँ अपनाई जाती हैं।

यह कहा जा सकता है कि छोटे न्यादर्शों की अपेक्षा बड़े न्यादर्श ही क्यों नहीं ले लिए जावें ताकि अलग प्रक्रिया अपनाना ही नहीं पड़े। आर्थिक, राजनैतिक एवं सामाजिक क्षेत्रों में तो न्यादर्शों का आकार बड़ा होता ही है लेकिन जहाँ समंक प्रयोगशाला द्वारा संग्रह करने पड़ते हैं वहाँ व्यय तथा समय अधिक लगता है, इसलिए बड़े आकार के न्यादर्श प्राप्त करना उचित नहीं समझा जाता है। तब यह जरूरी हो जाता है कि कोई विश्वसनीय रीति ऐसी निकाली जावे कि छोटे न्यादर्शों का विश्लेषण संभव हो सके। अभी-अभी ऐसी रीतियाँ विकसित हुई हैं जिनसे काफी शुद्धता से निष्कर्ष प्राप्त किये जा सकते हैं। छोटे न्यादर्शों का विश्लेषण करते समय हमारा उद्देश्य न्यादर्शों के अवलोकित मूल्य तथा समग्र के प्रचलित मूल्यों में तुलना कर यह देखना होता है कि इनका अंतर सार्थक (Significant) है या नहीं अर्थात् दोनों में अन्तर निदर्शन उच्चावचनों के परिणाम स्वरूप उत्पन्न हुआ है या किसी अन्य कारण से।

छोटे न्यादर्श की विशेषताएँ (Characteristics of Small Samples) :

(i) न्यादर्श जितना छोटा होता है उनके माध्यों में अन्तर उतना ही विस्तृत होता है क्योंकि थोड़ी संख्या में पदों के अध्ययन में क्षतिपूरक प्रभाव की कमी रहती है।

(ii) समग्र के प्रमाप विचलन की तुलना में छोटे न्यादर्श के प्रमाप विचलन के कम अनुमानित किये जाने की संभावना रहती है ।

(iii) छोटे न्यादर्शों के प्रमाप विचलन समग्र के प्रमाप विचलन की अपेक्षा कम होते हैं, इसलिए उनके आधार पर गणना किये गये प्रमाप विभ्रम भी कम आंके जावेंगे । इनमें संशोधन के लिए बेस्सेल का संशोधन सूत्र लगाया जाता है ।

बोध प्रश्न

1. साख्यिकीय नियमितता नियम की विशेषताएँ लिखिए।

....

....

....

2. महांक जड़ता नियम को समझाइए।

....

....

....

3. आर्थिक विश्लेषण में निदर्शन का क्या योगदान है।

....

....

....

3.10 सार्थकता परीक्षण (Test of Significance) :

छोटे न्यादर्शों में सार्थकता परीक्षण करने से पहले यह मानना जरूरी है कि हमारा अध्ययन इस परिकल्पना पर आधारित है कि वह समग्र (जिससे यह न्यादर्श लिया गया है) प्रसामान्यता की विशेषता रखता है । अनुसंधानों के अनुसार यदि मूल समग्र में सामान्य बंटन से कुछ भिन्नता भी हो तो सार्थकता परीक्षण का प्रयोग किया जा सकता है किन्तु इसमें अधिक असंमितता हो तो इन परीक्षण विधियों की विश्वसनीयता समाप्त हो जाती है ।

छोटे न्यादर्शों में सार्थकता परीक्षण विशेष प्रकार से होता है । छोटे न्यादर्शों की जाँच मुख्यतः निम्नलिखित तीन परीक्षणों पर आधारित होती है :-

(i) स्टुडेन्ट का टी-परीक्षण (Student's T-Test)

(ii) फिशर का जैड-परीक्षण (Fisher's Z-Test)

(iii) एफ-परीक्षण (F-Test)

(i) टी-परीक्षण पर आधारित सार्थकता-परीक्षण:-

विलियम गोस्सेट (William Gosset) ने इस परीक्षण का प्रतिपादन किया है । छोटे न्यादर्शों के समान्तर माध्य के महत्व की जाँच करने के लिए निम्न सूत्र की सहायता से 't' का मूल्य निकालते हैं :-

$$t = \frac{\bar{x} - m}{\sigma} \sqrt{n}$$

Where $\bar{x}$ = Mean of the Sample

m = Mean of the universe

NOTES

σ = Standard deviation

n = No. of items.

गोस्सेट के इस सिद्धात पर 't' सारणी बनाई गई है । इस सारणी में स्वतंन्त्रता के विभिन्न अंकों के लिए 10% , 5% ,1% आदि स्तरों पर 't' के मूल्य दिये हुए होते हैं । प्राप्त मूल्यों और सारणी मूल्यों की तुलना की जाती है। यह पूर्व कल्पना कर ली जाती है कि न्यादर्श और समग्र के माध्यों में कोई अन्तर नहीं है अर्थात् अन्तर शून्य है ।

इस प्रकार, यदि निकाला गया 't' का मूल्य उसके सारणी मूल्य से अधिक है तो अन्तर सार्थक होगा और हमारी शून्य की कल्पना असत्य होगी । यदि यह मूल्य सारणी मूल्य से कम है तो अन्तर निरर्थक होगा और शून्य- कल्पना सत्य होगी । परीक्षण 1% सार्थकता स्तर पर भी किया जा सकता है किन्तु अधिकतर यह 5% सार्थकता स्तर पर किया जाता है ।

Illustration 20. The ten items of a sample had the following values–

19., 22, 21, 25, 24, 27, 26, 31, 30 and 35.

In the light of these data discuss the suggestion that the mean value in the universe is 24.

Solution-

S.No.	Values (x)	Deviation from average 26 (d)	Square of deviations (d^2)
1	19	–7	49
2	22	–4	16
3	21	–5	25
4	25	–1	1
5	24	–2	4
6	27	+1	1
7	26	0	0
8	31	+5	25
9	30	+4	16
10	35	+9	81
	$\Sigma m = 260$	0	$\Sigma d^2 = 218$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{260}{10} = 26.$$

$$\text{S.D. of the sample} = \sqrt{\left(\frac{\Sigma d^2}{n-1}\right)} = \sqrt{\left(\frac{218}{10-1}\right)} = \sqrt{\left(\frac{218}{9}\right)} = \sqrt{(14.22)} = 4.9$$

$$= t = \frac{\bar{x} - m}{\sigma}\sqrt{(n)} = \frac{26-24}{4.9}\sqrt{(10)} = \frac{2}{4.9} \times 3.16 = 1.3$$

The number of degrees of freedom is n–1 = 10–1 = 9. The value of 't' for 9 degrees of freedom at 5% level of significance is 2.262. The calculated value of 't' is 1.38 which is less than 2.262, hence difference is not significant. This error could have arisen due to fluctuations of sampling and can be said that the mean value in the universe is 24.

Illustration 21 (A)

Ten individuals are chosen at random from a population and their incomes are found to be Rs. 63, 63, 64, 65, 66, 69, 69, 70, 70, and 71. Discuss the suggestion that the mean–Income in the universe is Rs. 65. Given that for a degree of freedom, the values of Students t–test at 5% level of Significance is 2.262.

Solution : यह मान लिया गया है कि न्यादर्श माध्य आय तथा समग्र माध्य आय में कोई सार्थक अंतर नहीं है ।

Calculation of Mean and S.D. or the Sample

NOTES

Income (x) Rs.	**dx (67)**	**d^2**
63	–4	16
63	–4	16
64	–3	9
65	–2	4
66	–1	1
69	+2	4
69	+2	4
70	+3	9
70	+3	9
71	+4	16
670		88

$$\bar{x} \text{ (Mean of the Sample)} = \frac{\Sigma x}{n} = \frac{670}{10} = 67 \text{ Rs.}$$

m (Mean of the Universe) = 65 Rs. (given)

$$\sigma \text{ (S.D. of the Sample)} = \sqrt{\frac{\Sigma d^2}{n-1}} = \sqrt{\frac{88}{n-1}} = \sqrt{\frac{88}{9}} = 3.1$$

$$t = \frac{\bar{x} - m}{\sigma}\sqrt{n} = \frac{67-65}{3.1}\sqrt{10} = \frac{2 \times 3.16}{3.1} = 2.03$$

t (calculated) 2.03 < (is less than) t (tabulated) 2.262 . Hence, Null Hypothesis is true and the difference of two averages is not significant.

दो लघु–न्यादर्शों के माध्यों के अंतर का 't' परीक्षण (t–test of the difference of means of two small samples)

(i) दो न्यादर्शों से प्राप्त समान्तर माध्यों के बीच अंतर की सार्थकता के परीक्षण के लिए 't' का मूल्य या मान निम्नलिखित सूत्र से निकाला जाता है –

$$t = \frac{\bar{x} - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

उपरोक्त सूत्र में $\bar{x}_1$ तथा $\bar{x}_2$ दोनों न्यादर्शों के औसत हैं । S का मान निकालने के लिए निम्न सूत्र का प्रयोग होगा :

$$S = \sqrt{\frac{\Sigma d_1^2 + \Sigma d_2^2}{n_1 + n_2 - 2}}$$

यदि दोनों न्यादर्शों के प्रमाप विचलन प्रश्न में दिये हों तो बेस्सेल संशोधन के आधार पर संशोधित 's' का सूत्र होगा–

$$S = \sqrt{\frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2 - 2}}$$

उपरोक्त सूत्रों में $(n_1 + n_2 - 2)$ = स्वातंत्र्य संख्या (Degree of freedom)

NOTES

निष्कर्ष-

यदि t (calculated) > t (Tabulated) = Difference Significant

„ t („) < t („) = „ not Significant

Illustration 22. The means of two random samples of sides 9 and 7 respectively are 196.42 and 198.82 .The sums of the square of the deviation from the mean are 26.94 and 18.93 respectively. Can the samples be considered to have been drawn from the same normal population ?

d.f.	5% Value of 't'	1% Value of 't'
13	2.160	3.012
14	2.145	2.977
15	2.131	2.947
16	2.120	2.921

Solution :

The Standard error (S) of the difference of the means of the two samples is–

$$S = \sqrt{\frac{\Sigma d_1^2 + \Sigma d_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{26.94 + 18.93}{9 + 7 - 2}}$$

$$= \sqrt{\frac{45.87}{14}} = \sqrt{3.28} = 1.8$$

Applying to the given data,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{196.42 - 198.82}{1.8\sqrt{\frac{1}{9} + \frac{1}{7}}} = \frac{-2.4}{1.8 \times 0.25} = \frac{-2.4}{0.9} = 2.67$$

The number of degrees of freedom is $n_1 + n_2 - 2$ i.e. $9 + 7 - 2 = 14$. The value of 't' for 14 degrees of freedom at 5% level of significance is 2.145 . The calculated value of 't' is more than its tabulated value. Hence, the difference between the two means is significant and the samples cannot be considered to have been drawn from the same normal population.

Illustration 23 : Two types of batteries are tested for their length of file and following results are obtained :-

	No. in Sample	Mean	Variance (σ^2)
Battery A	10	500 hrs.	100
Battery B	10	560 hrs.	121

Is there a significant difference in the two means :

Solution : On the hypothesis that the two samples are taken from populations with similar means and standard deviations, the value of combined standard error will be-

$$S = \sqrt{\frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(10 \times 100) + (10 \times 121)}{10 + 10 - 2}} = \sqrt{\frac{1000 + 1210}{18}}$$

$$= \sqrt{\frac{2210}{18}} = \sqrt{122.8} = 11.08$$

Applying to the given data-

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{500 - 560}{11.08\sqrt{\frac{1}{10} + \frac{1}{n_{10}}}} = \frac{-60}{11.08 \times 0.45} = 12.03$$

For 18 degrees of freedom at 5% level of significance the value of 't' according to the table is 2.1. The calculated value is. 12.03 being much higher than this, the difference between the two means is highly significant.

Illustration 24 : The annual salary of professors in the colleges averages Rs. 30,000 and has standard deviation or Rs. 1,000 . The salary of doctors averages Rs. 25,000 with standard deviation Rs. 1500 . These data relate to samples of size of 25 for each group chosen at random. Test at 5% level of significance, whether there is significant difference in the mean of the salaries of the two groups.

Solution : Standard error (s) of the difference of the two means

$$S = \sqrt{\frac{n_1 \sigma_1^2 + n_2 \sigma_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(25 \times 1000^2) + (25 \times 1500^2)}{25 + 25 - 2}} = \sqrt{\frac{2,50,00,000 + 5,62,50,000}{48}}$$

$$= \sqrt{\frac{8,12,50,000}{48}} = \sqrt{16,92,708} = 1301$$

Applying to the given data,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{30,000 - 25,000}{1301\sqrt{\frac{1}{25} + \frac{1}{25}}} = \frac{5000}{1301 \times 0.28} = \frac{5000}{364.28} = 13.7$$

For 48 degrees of freedom at 5% level of significance the value of 't' according to the table is 2.01 which is less than the calculated values. Hence, there is significant difference between mean salaries of the two groups.

Illustration 25 : Two kinds of manures are applied to 16 plots of one acre each, other conditions remaining the same, the yield (in quintals) is given below :

Manure I	18,	20,	36,	50,	49,	36,	34,	49,	41,	(9 plots)
Manure II	29,	28,	26,	35,	30,	44,	46,	(7 plots)		

Examine the significance of the difference between the mean yields due to the application of different kinds of manures.

NOTES

Solution :

Manure I			Manure II		ई
yield x_1	d 1 (37)	$d^2{}_1$	yield x_2	d_2 (34)	d_2^2
18	−19	361	29	−5	25
20	−17	289	28	−6	36
36	−1	1	26	−8	64
50	+3	169	35	+1	1
49	+12	144	30	−4	16
36	−1	1	44	+10	100
34	−3	9	46	+12	144
49	+12	144	-		
41	+4	16	-		
333		1134	238		386

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{333}{9} = 37$$

$$\bar{x}_2 = \frac{\Sigma x_2}{n_2} = \frac{238}{7} = 34$$

$$S = \sqrt{\frac{\Sigma d_1^2 + \Sigma d_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{1123 + 386}{9 + 7 - 2}} = \sqrt{\frac{1520}{14}} = \sqrt{108.57} = 10.42$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{N_1 N_2}{N_1 + N_2}}$$

$$= \frac{37 - 34}{10.42}\sqrt{\frac{9 \times 7}{9 + 7}} = \frac{3}{10.42} \times \frac{1}{2} = 0.57$$

For 14 degrees of freedom at 5% level of significance, the value of 't' according to the table, is 2.14. The Calculated Value of 't' is 0.57 which is less than the tabled value. Hence, the mean yields do not differ significantly due to the applications of the kinds of manures.

Illustration 26 : Two laboratories carry out independent estimates of fat content for ice-cream made by a certain firm. A sample is taken from each batch halved, and the separate halves sent to the two laboratories . They obtain the following results :-

percentage fat content in Ice-cream.

Batch No.	:	1	2	3	4	5	6	7	8	9	10
Lab A	:	7	8	7	3	8	6	9	4	7	8
Lab B	:	9	8	8	4	7	7	9	6	6	6

Is the testing reliable ?

Solution : If the testing is reliable, the mean of the difference in results should not significantly differ from zero.

We shall calculate the mean and the standard deviation of the difference of two results :

NOTES

Batch No.	Lab A	Lab B	Difference of results (Lab B- Lab A)	d (0.3)	d^2
1	7	9	+2	+1.7	2.89
2	8	8	0	–0.3	0.09
3	7	8	+1	+0.7	0.49
4	3	4	+1	+0.7	0.49
5	8	7	–1	–1.3	1.69
6	6	7	+1	+0.7	0.49
7	9	9	0	–0.3	0.09
8	4	6	+2	+1.7	2.89
9	7	6	–1	–1.3	1.69
10	8	6	–2	–2.3	5.29
			+3		16.10

Mean of the difference or $\bar{x} = \frac{3}{10} = 0.3$

Standard deviation (σ) of the differences

$$= \sqrt{\frac{\Sigma d^2}{n-1}} = \sqrt{\frac{16.1}{9}} = 1.34$$

Applying the given information,

$$t = \frac{x - m}{\sigma}\sqrt{n}$$

$$= \frac{0.3 - 0}{1.34}\sqrt{10} = \frac{0.3}{1.34} \times \frac{3.16}{1} = 0.71$$

The value or 't' for 9 degrees of freedom at 5% level is 2.57. The calculated value is much less than this, hence, the difference is insignificant and the tests have been reliable.

छोटे न्यादर्शों में सहसंबंध गुणक की सार्थकता का 't' परीक्षण

(Testing the significance of Coefficient or Correlation in small samples– 't'–test):–

छोटे न्यादर्शों में सहसंबंध गुणक के प्रमाप विभ्रम की गणना भी उसी आधार पर की जाती है जिस पर बड़े न्यादर्शों के लिए । किन्तु इसमें थोड़ा अन्तर हो जाता है । उस सूत्र के अंश में $1 - r^2$ के स्थान पर $\sqrt{1 - r^2}$ और हर में $\sqrt{n}$ के स्थान पर $\sqrt{n-2}$ का प्रयोग होता है । 'n' में से 2 कम करने का कारण यह है कि 'r' की गणना में 2 स्वतंत्रता की सीमाएँ कम हो जाती हैं । इस प्रकार सहसंबंध गुणक के प्रमाप विभ्रम को ज्ञात करने के लिए निम्नलिखित सूत्र का प्रयोग किया जाता है :–

$$S\,E\,(r) = \frac{\sqrt{1 - r^2}}{\sqrt{n - 2}}$$

't' का मूल्य (मान) ज्ञात करने के लिए निम्नलिखित सूत्र का प्रयोग करते हैं ।

$$t = \frac{r}{\sqrt{1 - r^2}} \times \sqrt{n - 2}$$

Illustration 27 : Find the least value of 'r' in a sample of 27 pairs from a bivariate normal population significant at 5% level.

Solution : n = 27

Degree of freedom (d.f.) = 27–2 = 25

NOTES

$$t = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} = \frac{r}{\sqrt{1-r^2}} \sqrt{25}$$

$$t = \frac{5r}{\sqrt{1-r^2}}$$ (Value of 't' 25 degrees of freedom of 5% level is 2.06)

For significance of 'r' the value of 't' must be more than 2.06. Therefore,

$$\frac{5r}{\sqrt{1-r^2}} > 2.06$$

$$5r > 2.06 \sqrt{1-r^2}$$

$$r\ 25\ r^2 > 4.2436\ (1-r^2)$$

$$r\ 25\ r^2 > 4.2436 - 4.2436\ r^2$$

$$r\ 25\ r^2 + 4.2436\ r^2 > 4.2436,$$

$$29.2436\ r^2 > 4.2436$$

$$r > \sqrt{\frac{4.2436}{29.2436}}, \qquad r > .381$$

Therefore, at 5% level of significance, the value of 'r' must be atleast 0.381

Illustration 28 : A random sample of 18 pairs of observations from a normal population gives a coefficient Correlation of 0.52. Is it likely that the variables in the population are uncorrelated ?

Solution :

$$t = \frac{r}{\sqrt{1-r^2}} \times \sqrt{n-2} = \frac{0.52}{\sqrt{1-(0.52)^2}} \times \sqrt{18-2}$$

$$= \frac{0.52}{\sqrt{1-0.2704}} \times \sqrt{16} = \frac{0.52}{\sqrt{1-0.7296}} \times \sqrt{16}$$

$$= \frac{0.52}{0.85} \times 4 = 2.44$$

For 16 degrees of freedom, according to the table, the value of 't' at 5% level of significance is 2.12 and at 2% is 2.58.

The observed value of 't' is significant at 5% level but not at 2% level. This gives reasons to suspect the hypothesis that the variables are uncorrelated. They may be correlated.

[II] फिशर का 'Z'–परीक्षण (Fisher's 'Z'–test):

प्रोफेसर फिशर ने सहसंबंध गुणक की सार्थकता की जाँच करने के लिए न्यादर्श 'Z' का प्रतिपादन किया है । इस परीक्षण में 'r'को 'Z' में परिवर्तित कर दिया जाता है । इसी आधार पर इसको फिशर का 'Z' परीक्षण या 'Z'– रूपान्तरण कहते हैं ।

'Z'-परीक्षण का उपयोग (Uses of Z-test) :

(i) न्यादर्श में 'r' के मान तथा समग्र के 'r' के मान में अन्तर सार्थक है अथवा उसमें संगति है;

(ii) दो न्यादर्शों के सहसंबंध गुणक का अन्तर सार्थक है अथवा अर्थहीन ।

(i) **न्यादर्श के 'r' तथा समग्र के 'r' के अन्तर की सार्थकता :**

यहाँ दोनों 'r' का 'Z' रूपान्तरण करने के लिए निम्नलिखित सूत्र प्रयोग में लाये जाते हैं :–

$$Z(s) = \frac{1}{2} \log_e\left(\frac{1+r}{1-r}\right) = 1.1513 \log_{10}\left(\frac{1+r}{1-r}\right)$$

$$[\log_e = \log_{10} 2.3026]$$

$$Z(p) = 1.1513 \log_{10}\left(\frac{1+r(p)}{1-r(p)}\right)$$

$$S.E(z) = \frac{1}{\sqrt{n-3}}$$

निष्कर्ष:

यदि $Z(s) - Z(p) > 2.58\,\sigma_z$ तो अन्तर 1% सार्थकता स्तर पर सार्थक है अन्यथा अर्थहीन।

" $Z(s) - Z(p) > 1.96\,\sigma_z$ तो अन्तर 5% सार्थकता स्तर पर सार्थक है अन्यथा अर्थहीन।

" $Z(s) - Z(p) > 3\,\sigma_z$ तो अन्तर निश्चित रूप से सार्थक है अन्यथा अर्थहीन।

(ii) दो न्यादर्शों के सहसंबंध गुणक की सार्थकता जाँच :

दो न्यादर्शों के सहसंबंध गुणक (r_1 तथा r_2) के अन्तर की सार्थकता की जाँच के लिए उपरोक्त प्रकार से ही Z_1 तथा Z_2 रूपान्तरण प्राप्त किये जाते हैं तथा अन्तर के प्रमाप विभ्रम हेतु निम्नलिखित सूत्र का प्रयोग किया जाता है :–

$$S.E.(Z_1 - Z_2) = \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}$$

$$Z_1 = 1.1513 \log_{10}\left(\frac{1+r_1}{1-r_1}\right) \qquad Z_2 = 1.1513 \log_{10}\left(\frac{1+r_2}{1-r_2}\right)$$

(n_1 तथा n_2 न्यादर्शों की इकाई संख्या है)

निष्कर्ष :

(i) $(Z_1 - Z_2) > 2.58\,\sigma(Z_1 - Z_2)$ तो 1% सार्थकता स्तर पर सहसंबंध गुणक अंतर सार्थक है, अन्यथा अर्थहीन।

(ii) $(Z_1 - Z_2) > 1.96\,\sigma(Z_1 - Z_2)$ तो 5% सार्थकता स्तर पर सहसंबंध गुणक अंतर सार्थक है, अन्यथा अर्थहीन।

(iii) $(Z_1 - Z_2) > 3\,\sigma(Z_1 - Z_2)$ तो अंतर निश्चित रूप से सार्थक है अन्यथा अर्थहीन।

Illustration 29 : Coefficient of Correlation of 28 pairs of observations of a sample is 0.7 and Corresponding population value is 0.4. Is the difference significant?

$n = 28$; $r(s) = 0.7$; $r(p) = 0.4$

Applying Z–Transformation.

$$Z(s) = 1.1515 \log_{10}\left(\frac{1+0.7}{1-0.7}\right)$$

$$= 1.1513 \log_{10}\left(\frac{1.7}{0.3}\right) = 1.1513 \log_{10}\left(\frac{17}{3}\right)$$

$$= 1.1513\,(\log_{10}{}^{17} - \log_{10}{}^{3})$$

$$= 1.1513\,(1.2304 - 0.4771)$$

$$= 1.1513 \times 0.7533 = 0.87$$

$$Z(p) = 1.1513 \log_{10}\left(\frac{1+0.4}{1-0.4}\right)$$

$$= 1.1513 \log_{10}\left(\frac{1.4}{0.6}\right) = 1.1513 \log_{10}\left(\frac{7}{3}\right)$$

NOTES

$$= 1.1513\,[\log_{10}{}^{7} - \log_{10}{}^{3}] = 1.1513\,[.8451\text{–}4771]$$

$$= 1.1513 \times 0.368 \quad = 0.423$$

$$Z_{(s)} - Z_{(p)} = 0.87 - 0.423 \quad = 0.447$$

$$S.E\ (z) = \frac{1}{\sqrt{n-3}}$$

$$\sigma_{(z)} = \frac{1}{\sqrt{28-3}} \times \frac{1}{\sqrt{25}} = \frac{1}{5} = 0.2$$

At 1% level of significance

$$Z_{(s)} - Z_{(p)} < 2.58\ \sigma Z$$

$$< 2.58 \times 0.2$$

$$0.447 < 0.516$$

Hence, the difference between sample and universe Coefficient of correlation is not significant.

Illustration 30 : Two groups of students are given an intelligence and an arithmetic test.

$n_1 = 45$ $\qquad r_1 = 0.45$

$n_2 = 39$ $\qquad r_2 = 0.38$

Is the difference between values of two 'r' significant

Solution :

First sample (Intelligence test)

$$Z_1 = \frac{1}{2}\log_{10}\frac{1-r}{1+r}$$

$$= 1.1513\ \log_{10}\frac{1.45}{0.55}$$

$$= 1.513\ \log_{10}\ 2.6364$$

$$= 1.1513 \times 0.4207$$

$$= 0.4844$$

Second Sample (Arithmetic test)

$$Z_2 = \frac{1}{2}\log_{10}\frac{1-r}{1+r}$$

$$= 1.1513\ \log_{10}\frac{1.38}{0.62}$$

$$= 1.1513\ \log_{10}\ 2.2258$$

$$= 1.513 \times 0.3476$$

$$= 0.4002$$

If the two samples are random samples from the same normal population , $Z_1 - Z_2$ is distributed normally with mean –zero and standard error.

$$S.e = \sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}} = \sqrt{\frac{1}{45-3} + \frac{1}{39-3}}$$

$$= \sqrt{\frac{1}{42} + \frac{1}{36}} = \sqrt{\frac{13}{252}} = \sqrt{0.0516} \qquad - 0.227$$

$$\therefore \frac{Z_1 - Z_2}{Se} = \frac{0.4844 - 0.4002}{0.227} = \frac{0.084}{0.227} = 0.37$$

At 5% level of Significance, value or normal deviate 1.96. The observed value is much less than this. Therefore, the difference between the values of two 'r' is not significant.

[III] 'F'–परीक्षण या प्रसरण अनुपात परीक्षण (F–Test or Variance Ratio Test) :–

इस परीक्षण के प्रणेता प्रो. फिशर एवं जार्ज डब्ल्यू. स्नेडिकॉर माने जाते हैं। फिशर के नाम से जुड़ा होने के कारण ही इसे 'F–Test' कहते हैं। इसमें प्रसरणों के अन्तर की सार्थकता की जाँच की जाती है। इस जाँच का आधार यह है कि समग्र प्रसरण के दो स्वतंत्र अनुमान सार्थक रूप से भिन्न हैं या दोनों ही एक समग्र– प्रसरण वाले प्रसामान्य वितरण से लिये गये हैं या नहीं। यहाँ प्रसरणों के अनुपात का प्रयोग होने के कारण इस परीक्षण को प्रसरण् अनुपात परीक्षण (Variance Ratio test) भी कहते हैं। इसे निम्न प्रकार से व्यक्त किया जा सकता है :–

$$F \text{ (प्रसरण अनुपात)} = \frac{\text{Larger estimate of variance}}{\text{Smaller estimate of variance}}$$

$$= \frac{S_1^2}{S_2^2} \qquad \text{यहाँ } (S_1^2 > S_2^2)$$

$$S_1^2 = \frac{\Sigma d^2}{n_1 - 1} \qquad S_2^2 = \frac{\Sigma d^2}{n_2 - 1}$$

'n_1' तथा n_2 क्रमशः प्रथम एवं द्वितीय न्यादर्शों की इकाई संख्या है। यहाँ बड़े प्रसरण वाले न्यादर्श की स्वातन्त्र्य संख्या (d.f.) = n–1 (v_1) कम प्रसरण वाले न्यादर्श की स्वातन्त्र्य संख्या = n–2 (v_2).

प्रसरण अनुपात की सार्थकता देखने के लिए 'F' के लिए परिकल्पित मूल्य तथा F.05 सारणी मूल्य से तुलना की जाती है। यदि F > F.05 है तो प्रसरण अनुपात को सार्थक माना जाता है अर्थात् यह माना जावेगा कि दोनों न्यादर्श एक ही समग्र से नहीं लिये गये हैं। यदि F < F.05 है तो प्रसरण अनुपात निरर्थक होगा। यह माना जावेगा कि दोनों न्यादर्श एक ही समग्र से लिये गये हैं।

Illustration 31 : Two Samples of size 10 and 9 give the sum of squares of deviations from their respective means as 135 and 80 square centimeters. Can they be regarded as drawn from the same normal population ?

Solution : $S_1^2 = \frac{\Sigma d^2}{n_1 - 1}$ $\qquad S_2^2 = \frac{\Sigma d^2}{n^2 - 1}$

$$= \frac{135}{10-1} \qquad = \frac{80}{9-1}$$

$$= \frac{135}{9} \qquad = \frac{80}{8}$$

$$= 15 \qquad = 10$$

$$F = \frac{S_1^2}{S_2^2} = \frac{15}{10} = 1.5$$

At 5% level of Significance , the value of F.05 , according to the table is 3.39 F –Calculated 1.5 < F.05 tabulated 3.39 . Hence, Variance ratio is not significant, i.e. the two samples belong to the same population (universe).

Illustration 32 : For a random sample of 10 pigs, fed on diet A, the increase in weight in a certain period was 10, 6, 16, 17, 13, 12, 8, 14, 15 and 9 lbs.

For an other random sample of 12 pigs, fed on diet B, the increase was 7, 13, 22, 15, 12, 14, 18, 8, 21, 23, 10, and 17 lbs.

Show that the estimate or the population variance from the two samples are not significantly different.

For V_1 = 11, V_2 = 9, the 5% Value of F is 3.112)

NOTES

Solution :

Calculation of Variances

Sample 1			Sample-2		
x_1	d_1 (12)	d_1^2	x_2	d_2 (15)	d_2^2
10	–2	4	7	–8	64
6	–6	36	13	–2	4
16	+4	16	22	+7	49
17	+5	25	15	0	0
13	+1	1	12	–3	9
12	0	0	14	–1	1
8	–4	16	18	+3	9
14	+2	4	8	–7	49
15	+3	9	21	+6	36
9	–3	9	23	+8	64
			10	–5	25
			17	+2	4
120	0	120	180	0	314

$$\bar{x}_1 = \frac{\Sigma x_1}{n} = \frac{120}{10} = 12 \qquad \bar{x}_2 = \frac{\Sigma x_2}{n} = \frac{180}{12} = 15$$

$$S_1^2 = \frac{\Sigma d_1^2}{n_1 - 1} = \frac{120}{10 - 1} = \frac{120}{9} = 13.33 \qquad S_2^2 = \frac{\Sigma d_2^2}{n_2 - 1} = \frac{314}{12 - 1} = \frac{314}{11} = 28.55$$

$$F = \frac{\text{Larger estimate}}{\text{Smaller estimate}} = \frac{S_2^2}{S_1^2} = \frac{28.55}{13.33} = 2.14$$

$\therefore$ F < F.05; Hence, population variance from the two samples are not significantly different.

काई–वर्ग जाँच (Chi–Square Test)

इस परीक्षा का उपयोग ज्ञात किये गये अंक और सम्भावित अंकों के अन्तर का अध्ययन करने के लिए किया जाता है । इन दोनों अंकों के एक होने की सम्भावना नगण्य रहती है । ऐसी स्थिति में इन दो में जो अन्तर है, उसे यह मानकर कि वह निदर्शन उच्चावचन के कारण है, कहाँ तक ध्यान में नहीं रखा जाये । काई–वर्ग इस अन्तर का मापक है । यदि दोनों का अन्तर शून्य हो तो काई–वर्ग भी शून्य आता है । यदि दोनों में अन्तर है तो वह शून्य से अधिक होगा । यह अन्तर शून्य से अधिक निदर्शन के उच्चावचन के कारण भी आ सकता है । अत: यदि काई–वर्ग का मूल्य कुछ अधिक हो तो उसे छोड़ देना चाहिए, वह सार्थक (Significant) नहीं है । विभिन्न परिस्थिति में किस सीमा तक का अन्तर सार्थक नहीं है, यह मालूम करने के लिये काई–वर्ग की सारिणी मिलती है जिसकी सहायता से यह सीमा जानी जा सकती है । यदि काई–वर्ग का प्राप्त किया हुआ मूल्य सारिणी में दिये हुए मूल्य से अधिक हो तो प्राप्त अंक और सम्भावित अंक का अन्तर सार्थक होता है । काई–वर्ग ज्ञात करने का वर्ग इस प्रकार है–

$$\text{Chi-Square of } X^2 = \Sigma \left\{\frac{(f_0 - f)^2}{f}\right\}$$

where,

X^2 = Chi–Square

f_0 = Observed frequency

f = Corresponding theoretical of probable frequency.

3.11 बोध प्रश्न

1. टी टेस्ट क्या है ?

....

....

....

2. एफ टेस्ट क्या है ?

....

....

....

3.12 सारांश

समग्र में से जब किसी पद्धति के द्वारा प्रतिनिधि अंश चुन लिया जाता है तो उसे निदर्शन कहते हैं। निदर्शन ज्ञात करने के प्रमुख प्रचलित विधियाँ निम्न हैं– व्यापक निदर्शन, सविचार निदर्शन, दैव निदर्शन व बहुस्तरीय निदर्शन आदि। निदर्शन का आधार समग्र की इकाइयों में पाई जाने वाली एकरूपता है अतः निदर्शन में समग्र की विशेषताएँ उपलब्ध रहती है। किन्तु न्यादर्श मध्यक को समग्र मध्यक के बिल्कुल समान नहीं समझना चाहिए क्योंकि इसमें अंतर होता है। न्यादर्श एवं समग्र मध्य के अंतर को न्यादर्श विभ्रम कहते हैं। छोटे न्यादर्श के सार्थकता परीक्षण मुख्यतः तीन प्रकार के परीक्षणों पर आधारित होता है– स्टुडेन्ट का टी परीक्षण, फिशर का जैड परीक्षण, एफ परीक्षण आदि।

3.13 अभ्यास प्रश्न

दीर्घउत्तरीय प्रश्न (Long Answer Type Questions)

1. संगणना प्रणाली और निदर्शन प्रणाली का अन्तर बतलाइए। इनके तुलनात्मक गुण-दोषों की विवेचना कीजिए। उन परिस्थितियों को स्पष्ट कीजिए, जिनके अन्तर्गत इनमें से प्रत्येक रीति का प्रयोग लाभदायक हो सकता है।
   Distinguish between a census and sample enquiry. Discuss their relative advantages and disadvantages. Explain the conditions under which each of these methods may be used with advantage.
2. "एक अच्छा न्यादर्श दैव-निदर्शन पर आधारित होना चाहिए।"विवेचन कीजिए।
   a good sample must be based on random selection. discuss.
3. "एक न्यादर्श बड़ा होते हुए भी व्यर्थ हो सकता है क्योंकि वह दैव-निदर्शन पर आधारित नहीं अथवा दैव-निदर्शन पर आधारित होते हुए भी अविश्वसनीय हो सकता है, क्योंकि वह छोटा है।" इस कथन की समीक्षा कीजिए और दैनिक जीवन में निदर्शन के महत्व को समझाइए।
   "A sample may be large yet worthless, because it is not random.or it may be random but underliable, because it is small." Comment upon this statemant and explain the importance of sampling inaur daily life.
4. "कुछ निश्चित परिस्थितियों में निदर्शन एक आवश्यकता है।" इस कथन को सउदाहरण समझाइए। निदर्शन की अधिक प्रचलित विधियाँ कौन-कौनसी हैं ? दैव-निदर्शन की तुलना में स्तरित न्यादर्श के कुछ लाभों की व्याख्या कीजिये।
   "Sampling is a necessary under certain conditions." Illustrate it by a suitable example. What are the well known methods of sampling? Indicate some of the advantages of stratified sampling over random sampling.

5. "किसी न्यादर्श सर्वेक्षण में विभ्रम के अनेक स्रोत होते हैं। एक पूर्णरूप से यथार्थ सर्वेक्षण कल्पना मात्र है।" इस कथन का विवेचन कीजिए।
"In any sample survey, there are many sources of error. A perfect survey is a myth." Discuss the statement.

6. सांख्यिकीय नियमितता नियम और महांक जड़ता नियम को स्पष्ट रूप से समझाइए।
State and explain the law of statistical regularity and the law of inertia of large numbers.

7. आर्थिक विश्लेषण में निदर्शन का क्या योगदान है, विवेचना कीजिए।
Discuss the role of sampling in economic analysis.

8. न्यादर्श चुनने की विभिन्न विधियों का वर्णन कीजिए। उदाहरण देते हुए प्रत्येक के गुण और दोष बतलाइए।
Describe the various methods of selecting a sample. State the merits and demerits of each, giving examples.

9. न्यादर्श की विभिन्न तकनीकें बतलाइए।
Give an account of the different techniques of sampling.

10. निम्नलिखित पर संक्षिप्त टिप्पणियाँ लिखिए-
    (क) सम्भावना समंक (Theory of Probability)
    (ख) प्रतिनिधि समंक (Representative numbers)
    (ग) टिप्पेट के दैव अंक (Tippett's Random numbers)
    (घ) बहुस्तरीय दैव-निदर्शन (Multi-stage Area Random Sampling)
    (ङ) निदर्शन (प्रतिचयन) के उद्देश्य (Objects of Sampling)

11. Distinguish clearly between any two of the following giving, suitable example wherever necessary-
    (i) Primary and Secondary data,
    (ii) Census and sample enquiry,
    (iii) Classification and Tabulation.

12. समंकों के संकलन की संगणना एवं निदर्शन रीतियों में अन्तर बताइये और संक्षेप में उनके तुलनात्मक लाभों का विवेचन कीजिये।
Distinguish between the census and sample methods of collecting statistics and discuss briefly their comparative advantages.

13. दैव निदर्शन तथा स्तरित निदर्शन में अंतर बतलाइए।
Differentiate between Random Sampling and Statified Sampling.

14. टी–टेस्ट क्या है ?
What is t–Test ?

15. निदर्शन सांख्यिकीय अन्वेषणों में क्यों आवश्यक है ? निदर्शन की सामान्य रीतियों का वर्णन कीजिए। प्रतिचयन सिद्धांत में सार्थकता की धारणा को स्पष्ट कीजिए।
Why is sampling necessary in statistical–investigations? Describe the general methods of sampling. Clarify the concept of significance in the theory of sampling.

16. निदर्शन का महत्व बतलाइए। निदर्शन-विभ्रम के कारण हैं ?
Depict the importance of sampling. What are the causes of errors of sampling.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. निदर्शन क्या है ?
   What is Sample.
2. निदर्शन की विभिन्न विधियों का संक्षिप्त विवरण दीजिये।
   Discuss briefly the different methods of sampling.
3. निदर्शन का महत्व बतलाइये।
   Explain the importance of sampling.
4. निदर्शन विभ्रम के कारण क्या हैं ?
   What are the causes of errors of sampling.
5. दैव निदर्शन की व्याख्या कीजिये।
   Discuss random sampling.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. निदर्शन अनुसंधान का गुण है-
   (अ) मितव्ययिता (ब) शुद्धता का सत्यापन (स) व्यावहारिकता (द) गहन अध्ययन
2. निदर्शन आधारित है-
   (अ) कलात्मक सिद्धान्त (ब) अर्थशास्त्रीय सिद्धान्त (स) गणितीय सिद्धान्त (द) वैज्ञानिक सिद्धान्त
3. युवक युवती जीवन साथी चुनने में भी निदर्शन का प्रयोग करते हैं-
   (अ) किंग (ब) ब्लेयर (स) डॉल्टन (द) हैमवर्ग

उत्तर- 1. (ब), 2. (द), 3. (ब).

3.14 व्यावहारिक प्रश्न

1. A coin is tossed 10,000 times and head turns up 5125 times . Is it reasonable to think that the coin is unbiassed ?
   [S.E. 50, Actual difference is greater than 3 times the S.E.]
2. A Coin is tossed 400 times and it turns head 216 times. Discuss whether the coin may be an unbiassed one and explain briefly the theoretical principles you would use for this purpose.

   [S.E. 10, coin is unbiassed.]
3. A random sample of 500 pineapples was taken from a large consignment and 65 were found to be bad. Estimate the proportion of the pineapples in the consignment as well as the standard error of the estimate. Deduce that the percentage of bad pineapples in the consignment almost certainly lies between 8.5 and 17.5.

   [S.E. of proportion = .015 –Proportion of bad pineapples lies between
   $(13 \pm 3 \times .015)$
   $(t = -1.5)$]
4. Calculate the standard error of the mean from the following data collected in one of the many random sample inquiries conducted to find out average earnings of a particular class

Earning p.m. (Rs)	Number of persons
upto 10	50
„ 20	150
„ 30	300
„ 40	500
„ 50	700
„ 60	800
„ 70	900
„ 80	1000

[S.E. .61]

5. What is meant by the standard error and what are its practical uses ? Intelligence tests on two groups of boys and girls, give the following results. Examine if the diffference is significant.

	Mean	S.D.	No.
Girls	84	10	121
Boys	81	12	81

[S.E. 1.61–Actual difference is less than three times the standard error.]

6. A random sample of 500 villages was taken from district Kanpur and the average of population per village was found to be 480 with standard deviation of 48 . Another random sample of 400 villages from the same district gave an average population of 500 per village with S.D. of 54. Is the difference between the averages statisticallly significant ? Give reasons.

[S.E.= .61]

7. An astrologer assured 400 candidates of their success in the examination. 320 of them passed. Do you admit his claim to power of knowing future events?

[S.E.=10]

8. A random sample of 1000 farms in a certain year gives an average of wheat 2000 lbs. per acre with a standard deviation of 192 lbs. A random sample of 1000 farms in a following year gives an average yield of 2100 lbs. per acre with a standard deviation of 224 lbs. Show these data are inconsistent with the hypothesis that the average yields in the country as a whole were the same in the years.

[S.E. 1–2 = 9.32, t = 10.7, hence difference is significant.]

9. In a locality containing 18,000 families a sample of 840 families was selected at random. Of these 840 families. 206 families were found to have a monthly income of Rs. 50 or less. It is desired to estimate how many out of 18,000 families have a monthly income of Rs. 50 or less. Within what limits would you place your estimates?

[S.E. =0.015]

10. In a random sample of 500 men from a particular district of U.P. 300 are found to be smokers. In one of 1000 men from another district 550 are smokers. Do the data indicate that the two districts are significantly different with respect to the prevalence of smoking among men?
(P.C.S.)
[S.E. =0.027]
11. A random sample of 200 villages was taken from a district and average population per village was found to be 485 with a standard deviation of 50. Another random sample of 200 villages from the same district gave an average population of 510 per village with standard deviation of 40 . Is the difference between the two samples statistically significant ? Give reasons.
[S.E. = 4.54]
12. Explain what do you understand by Standard Error or the mean of random sample. The data concerning height measurement of a random sample of individuals from a given population are as follows–
Mean = 172; S.D. = 12; n = 65
If a larger number of samples of the same size were selected at random from the given population what would be the limits of 20% confidence interval for the true mean.
[S.E. = 1.5]
13. The means of simple samples of 1,000 and 2,000 are 67.5 and 68.0 inches respectively. Can the samples be regarded as drawn from the same population of S.D. 2.5.
[S.E. =0.097]
14. Ten individuals are chosen at random from a proportion and their heights are found to be in inches 63, 63, 64, 65, 66, 69, 69, 70, 70, 71. Discuss the suggestions that the mean height in the universe is 65 inches, given that a degree of freedom the value of students at 5% level of significance is 2.262,
[t = 2.02]
15. The median height of 100 M.Com. students is 66 inches with a standard deviation of 3 inches and the Median height of 121 M.A. students is 64 inches with a standard deviation of 4 inches. Are M.Com. students taller than M.A. students ?
[S.E. Median 37599, .45575, S.E. of difference =5.6]
16. Suppose a sample of 197 families is randomly chosen out of 8000 families residing in an area, and their indebtedness position is analysed. The following table describes the position.

Indebtedness	No. of Families	Indebtedness	No. of families
1–20	1	100–120	43
20–40	3	120–140	31
40–60	14	140–160	6
60–80	29	160–180	2
80–100	27	180–200	1
			197

NOTES

What is the range within which the average indebtedness of the 2800 families is likely to be?

(Rs. 96.96 to 98.73)

17. A random sample of 1000 men from Northern India given their mean wage to be Rs. 2.50, per day with a standard deviation of Rs. 1.50. A sample of 1500 men from Southern India gives a mean wage of Rs. 2–69 per day with a standard deviation of Rs. 200. Discuss whether the mean rate of wages varies as between the two regions.

18. What is the chance that a leap year, selected at random, will contain 53 Sundays.

$\left[P = \frac{2}{7}\right]$

19. A and B stand in a ring with 10 other persons. If the arrangement of 12 persons is at random, find the chance that there are exactly three persons between A and B.

$\left[P = \frac{2}{11}\right]$

20. A, B, and C in order toss a coin. The first to throw a head wins. What are their respective chances of winning?

[A = 4/7, B = 2/7 and C = 1/7]

21. An experiment succeeds twice as often as it fails. Find the chance that in the next six trials there will be at least four successes.

[Ans. $= \frac{2}{3}, q = \frac{1}{3} \left(\frac{2}{3} + \frac{1}{3}\right)^6 = \frac{496}{726}$]

22. Find the chance of drawing a king, a queen and a knave in the order from a pack of 52 cards in three consecutive draws; the cards drawn not being replaced.

[Ans. $\frac{4 \times 4 \times 4}{52 \times 51 \times 50}$]

23. Three groups of children contain respectively.

3 girls and 1 boy.

2 girls and 2 boys,

1 girl and 3 boys

One child is selected at random from each group. Show that the chance that the three selected

consists of 1 girl and two boys is $\frac{13}{12}$.

p = girl from the first, boy from the second., boy from the third

$$= \frac{3}{4} \times \frac{1}{2} \times \frac{3}{4} = \frac{9}{32}$$

p = boy from the first, girl from the second, boy from the third

$$= \frac{1}{4} \times \frac{1}{2} \times \frac{3}{4} = \frac{3}{32}$$

p = boy from the first, boy from the second, girl from the third

$$= \frac{1}{4} \times \frac{1}{2} \times \frac{1}{4} = \frac{1}{32}$$

Joint $\quad p = \frac{9}{32} + \frac{3}{32} + \frac{1}{32} = \frac{13}{32}$

24. One purse contain 1 sovereign and 4 rupees, a second one contains 2 sovereigns and 3 rupees and a third contains 3 sovereigns and 2 rupees. One purse is selected at random out of these and a coin is taken from it. What is the probability that it is a sovereign?
(Ans. 6/15)

25. Peter and Paul play a game with two dice. Peter plays by throwing the dice together. If the total number of points is a prime number other than two, he wins outright. If it is even he throws again under the same conditions, in other cases the throw is passed to Paul who thrown under the same conditions. What is the probability of Peter's winning?
[Ans. 18/22].

26. What is meant by theoretical frequency distributions? Discuss the salient features of the Binomial, Normal and Poisson distributions.

27. What do you understand by Theoretical frequency distribution.'? Point out the properties of Normal and Poisson distributions.

28. Explain the conditions under which we get a binomial distribution, giving examples of its occurrence. Obtain the mean and the variance of this distribution.

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय–4 प्रसरण विश्लेषण (ANALYSIS OF VARIANCE)

NOTES

इकाई की रूपरेखा

4.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. प्रसरण विश्लेषण के अर्थ को समझ सकेंगे।
2. प्रसरण विश्लेषण की विभिन्न तकनीकों का ज्ञान होगा।
3. दो से अधिक न्यादर्शों के अंतरों की सार्थकता का अध्ययन कर सकेंगे।

4.1 प्रस्तावना

निदर्शन–सिद्धांत (Sampling Theory) में एक न्यादर्श माध्य की सार्थकता, दो न्यादर्शों के माध्यों के अंतर की सार्थकता का परीक्षण करने वाली विभित्र विधियों एवं तकनीकों का अध्ययन किया गया है। इनमें 'टी–परीक्षण' तथा 'काई–वर्ग परीक्षण' विशेष महत्वपूर्ण हैं। जब दो से अधिक न्यादर्शों के अन्तरों की सार्थकता का अध्ययन करना होता है तो ये विधियाँ उपयोगी नहीं होती हैं। यहाँ प्रसरण–विश्लेषण तकनीक का प्रयोग होता है। इस तकनीक का प्रतिपादन सर रॉनल्ड ए. फिशर ने किया है तथा श्री स्नेडेकॉर आदि ने इसके विकास में महत्वपूर्ण योगदान दिया है। यह तकनीक प्रसरणों के बीच के अन्तर का परीक्षण करने वाली एक तकनीक है जिसके द्वारा माध्यों की सजातीयता को जाँचा जाता है। इसके द्वारा प्रकृति में विचरण का अध्ययन किया जाता है।

4.2 प्रसरण विश्लेषण का अर्थ (Meaning of Analysis of Variance)

किसी भी समंक–समूह के पूर्ण विचलन को विभित्र अऋणात्मक भागों के योग के रूप में प्रदर्शित करने की गणितीय–प्रक्रिया को ही प्रसरण विश्लेषण कहते हैं। प्रत्येक भाग विचलन के स्वतंत्र कारण को प्रदर्शित करता है। **श्री फिशर** के अनुसार "कारणों के एक समूह द्वारा अभिनिर्धारित प्रसरण का अन्य वर्गों से उत्पत्र प्रसरण से पृथक्करण

प्रसरण कहलाता है।" (The separation of the variance ascribable to one group of cases from the variance ascribable to other groups is called variance)

श्री ऑवेनडेविस के अनुसार, "प्रसरण विश्लेषण आवश्यक रूप से, विचरण के अभिनिर्धारित स्रोतों के अनुरूप विभिन्न संघटकों में प्रसरण का विश्लेषण करने की एक प्रविधि है।"

इस प्रकार प्रसरण विश्लेषण में कुल प्रसरण को विभिन्न प्रसरण संघटकों में विभाजित करके न्यादर्श माध्यों की सजातीयता का परीक्षण किया जाता है। कुल प्रसरण को दो संघटकों में बाँटा जाता है:-

कुल प्रसरण = न्यादर्शों के मध्य प्रसरण + न्यादर्शों के अंतर्गत प्रसरण Total variance = variance between samples + variance within samples.

4.3 मान्यताएँ (Assumptions)

यह तकनीक निम्नलिखित मान्यताओं पर आधारित है :-

(i) **समग्र का प्रसामान्य वितरण**:- न्यादर्श जिस समग्र से लिये गये हैं, उनका वितरण प्रसामान्य (Normal) होना चाहिए तथा न्यादर्श इकाइयों (पदों) का चयन भी दैव-निदर्शन रीति से होना चाहिये।

(ii) **न्यादर्श स्वतंत्र होना चाहिए**:- समग्र से लिये गए न्यादर्श स्वतंत्र होने चाहिये। न्यादर्श स्वतंत्र होने पर ही प्रसरण अनुपात की गणना स्वतंत्र होगी। इसी स्थिति में F-अनुपात उपयोगी सिद्ध हो सकता है।

(iii) **संयोज्यता** (Additive Property) :- विभिन्न संघटक प्रसरणों का योग कुल प्रसरण के बराबर होना चाहिए। प्रसरण विश्लेषण की उपयोगिता उपर्युक्त मान्यताओं के पूर्ण होने पर ही है। इनके अभाव में प्रसरण विश्लेषण द्वारा निकाले गये निष्कर्ष भ्रमात्मक होंगे और महत्वहीन भी। ऐसी स्थिति में प्रसरण-विश्लेषण के निष्कर्षों का निर्वचन पूरी सावधानी से किया जाना चाहिए।

4.4 प्रसरण की गणना-रीति (Method of Calculating variance)

न्यादर्शों के 'बीच' तथा उनके 'अंतर्गत' प्रसरण की गणना इस प्रकार की जाती है :-

(i) **न्यादर्शों के बीच प्रसरण** :- विभिन्न न्यादर्शों के अलग-अलग माध्य निकालिए। सभी माध्यों का सामूहिक माध्य निकालिए। विभिन्न न्यादर्शों के अलग अलग माध्यों से सामूहिक माध्य निकालिए तथा विचलनों का वर्ग $(\bar{x} - \bar{\bar{x}})^2$ निकालिए। विचलनों के वर्गों के योग में न्यादर्शों से सम्बद्ध स्वतंत्र संख्या (k) में से एक घटाकर (k-1) प्राप्त संख्या से भाग दीजिए। यह ही न्यादर्शों के बीच प्रसरण होता है।

(ii) **न्यादर्शों के अंतर्गत प्रसरण** :- न्यादर्शों के 'बीच' प्रसरण मालूम करने के बाद न्यादर्शों के अंतर्गत प्रसरण निकाला जाता है। इस हेतु अलग-अलग न्यादर्श इकाइयों से उनसे संबंधित न्यादर्श माध्य के विचलन $(x - \bar{x}_1)$, $(x - \bar{x}_2)$ आदि मालूम कीजिए। फिर प्रत्येक के वर्ग $(x - \bar{x}_1)^2$, $(x - \bar{x}_2)^2$ निकालिए। वर्गों के योग में कुल मूल्यों से संबंधित स्वातंत्र्य संख्या $d.f._2$ या $V_2 = N-K$ या (n-1) से भाग देकर न्यादर्शों के अंतर्गत प्रसरण मालूम हो जाता है।

(iii) **प्रसरण अनुपात की गणना** (Variance Ratio or 'F'-Calculation) :- उपरोक्त दोनों प्रसरणों के आधार पर निम्न लिखित सूत्र द्वारा प्रसरण अनुपात (F) की गणना की जाती है :-

$$F = \frac{\text{Variance between samples (न्यादर्शों के बीच प्रसरण)}}{\text{Variance within samples (न्यादर्शों के अंतर्गत प्रसरण)}}$$

या

$$= \frac{\text{Greater Variance (बड़ा प्रसरण)}}{\text{Smaller Variance (छोटा प्रसरण)}}$$

(दोनों में जो प्रसरण बड़ा होता है, वह अंश के रूप में और छोटा हर के रूप में रखा जाता है)

अपनी प्रगति की जाँच करें
Test your Progress

NOTES

4.5 प्रसरण विश्लेषण सारणी (Analysis of Variance Table)

उपरोक्त गणन क्रिया की शुद्धता की जाँच के लिये कुल प्रसरण की गणना कीजिए । इस हेतु प्रत्येक मूल्य का बड़े माध्य ($\bar{\bar{x}}$) से विचलन निकाला जाता है, विचलनों के वर्ग किये जाते हैं और इनका योग निकाला जाता है । यह योग न्यादर्श के बीच वाले विचलन वर्गों तथा न्यादर्शों के अंतर्गत वाले विचलन वर्गों के योग के बराबर होता है ।

जैसे कि :

$$\Sigma (x_1 - \bar{\bar{x}})^2 + \ldots\ldots\ldots\ldots (x_k - \bar{x})^2 = \Sigma \left[nk\ (\bar{x}_k - \bar{\bar{x}})^2\right] + \left[\Sigma(x - \bar{x}_k)^2\right]$$

कुल प्रसरण से संबंधित स्वातंत्र्य संख्या (d.f. for total variance) :

सभी न्यादर्श इकाइयों की कुल संख्या में से 1 घटाकर स्वातन्त्र्य संख्या निकाली जाती है (n–1)

प्रसरण विश्लेषण सारणी का प्रारूप
(Proforma of Analysis of Variance Table)

प्रसरण स्रोत (Variance Source)	वर्गों का योग (Sum of Squares)	स्वातन्त्र्य संख्या (dx)	प्रसरण (Mean Squares)	प्रसरण अनुपात (variance Ratio) (F-Ratio)
न्यादर्शों के बीच (Between Sample)	$\Sigma[nk\ (\bar{x}_k - \bar{\bar{x}})^2]$	(K–1)	$\dfrac{\Sigma\ [nk\ (\bar{x}_k - \bar{\bar{x}})^2]}{(k-1)}$	Explained Variance
न्यादर्शों के अंतर्गत (Within sample)	$\Sigma\ (x - \bar{x}_k)^2$	(n – k)	$\dfrac{\Sigma\ (x - \bar{x}_k)^2}{(n-k)}$	Unexplained variance $F = \dfrac{\textit{Greater Variance}}{\textit{Smaller Variance}}$
कुल प्रसरण (Total variance)	$\Sigma\ (x - \bar{\bar{x}})^2$	(n – 1)	$\dfrac{\Sigma\ (x - \bar{\bar{x}})^2}{(n-1)}$	Total Variance

निर्वचन (Interpretation) :

F–सारणी में 5% सार्थकता स्तर पर V_1 तथा V_2 वाली पंक्ति में प्राप्त संख्या 'F' का सारणी मान है। निर्वचन इस प्रकार किया जाता है –

F_c (परिकलित मान) $> F_{.05}$ है तो न्यादर्शों में अंतर सार्थक है ।अतः शून्य परिकल्पना असत्य है ।

F_c (परिकलित मान) $< F_{.05}$ है तो न्यादर्शों में अंतर सार्थक नहीं है । अतः शून्य परिकल्पना सत्य है ।

4.6 प्रसरण विश्लेषण तकनीक–एकमार्गीय वर्गीकरण (Analysis of Variance Technique–One–way classification)

समंक–समूहों का वर्गीकरण दो आधार– एक मार्गीय या द्विमार्गीय पर किया जा सकता है । इसे एक कारक (साधन) अथवा द्वि–कारक (साधन) वर्गीकरण भी कहते हैं । प्रसरण विश्लेषण का प्रयोग दोनों प्रकार के वर्गीकरण में हो सकता है, लेकिन दोनों के लिए इसकी अलग–अलग तकनीक अपनाई जाती है ।

जब समंकों का वर्गीकरण किसी एक कारक के आधार पर किया जाता है तो वह एक–मार्गीय वर्गीकरण कहलाता है । जैसे, औद्योगिक उत्पादन पर श्रमिकों के प्रकार का प्रभाव, खाद की चार किस्मों का पाँच-पाँच खेतों की न्यादर्श इकाइयों के उत्पादन पर प्रभाव, या चार महाविद्यालयों के एम.ए. (अर्थशास्त्र) के 10–10 छात्रों के सार्वजनिक राजस्व विषय में प्राप्तांकों का विश्लेषण देखना ।

जब वर्गीकरण एक ही कारक के आधार पर किया गया है तो प्रसरण विश्लेषण के लिए निम्नलिखित तीन रीतियों का प्रयोग किया जाता है :–

(अ) प्रत्यक्ष रीति (Direct Method) ,

(ब) ऋजु रीति (Short cut method) ,

(स) मूल बिन्दु परिवर्तन रीति (Coding Method) ।

(अ) प्रत्यक्ष रीति :-

इस रीति के अंतर्गत न्यादर्शों के 'बीच' तथा न्यादर्शों के 'अन्तर्गत ' प्रसरण ज्ञात करते हैं, प्रसरण विश्लेषण सारणी तैयार करते हैं, प्रसरण अनुपात का आकलन करते हैं और उसकी तुलना F–सारणी में 5% सार्थकता स्तर पर तत्संबंधी स्वातन्त्र्य संख्याओं (V_1 तथा V_2) के लिए सारणी में दिये गए मूल्य से की जाती है और अंत में न्यादर्शों के बीच माध्य अंतर की सार्थकता का निर्वचन किया जाता है ।

Illustration 1 : The following table gives the yields of four plots each of four varieties of wheat . Find out, if the variety–differences are significant :–

Yield of Different varieties

A	B	C	D
16	20	32	28
20	22	24	20
20	16	28	24
16	22	12	32

Solution :– We assume as hypothesis that the four varieties of wheat are not different from each other.

Table

Variety–A x_1	Variety–B x_2	Variety–C x_3	Variety–D x_4
16	20	32	28
20	22	24	20
20	16	28	24
16	22	12	32
$\Sigma x_1 = 72$	$\Sigma x_2 = 80$	$\Sigma x_3 = 96$	$\Sigma x_4 = 104$
$\bar{x}_1 = 72 \div 4 = 18$	$\bar{x}_2 = 80 \div 4 = 20$	$\bar{x}_3 = 96 \div 4 = 24$	$\bar{x}_4 = 104 \div 4 = 26$

$$\text{Combined Mean } (\bar{\bar{x}}) = \frac{\Sigma x_1 + \Sigma x_2 + \Sigma x_3 + \Sigma x_4}{n}$$

$$= \frac{72 + 80 + 96 + 104}{16} = \frac{352}{16} = 22$$

न्यादर्शों के बीच प्रसरण (Variance between samples) :

विचलन वर्गों का योग (SSC) :

$$= n_1 (\bar{x}_1 - \bar{\bar{x}})^2 + n_2 (\bar{x}_2 - \bar{\bar{x}})^2 + n_3 (\bar{x}_3 - \bar{\bar{x}})^2 + n_4 (\bar{x}_4 - \bar{\bar{x}})^2$$

$$= 4 (18 - 22)^2 + 4 (20 - 22)^2 + 4 (24 - 22)^2 + 4 (26 - 22)^2$$

$$= 64 + 16 + 16 + 64 = 160$$

स्वातन्त्र्य संख्या = 4–1 = 3 (K–1)

न्यादर्शों के बीच प्रसरण (MSC) $= \frac{ssc}{k - 1} = \frac{160}{3} = 53.33.$

NOTES

न्यादर्शों के अन्दर प्रसरण (Variance within samples) :

Table

Variety – A			Variety – B			Variety – C			Variety – D		
x_1	$(x_1-\bar{x}_1)$	$(x_1-\bar{x}_1)^2$	x_2	$(x_2-\bar{x}_2)$	$(x_2-\bar{x}_2)^2$	x_3	$(x_3-\bar{x}_3)$	$(x_3-\bar{x}_3)^2$	x_4	$(x_4-\bar{x}_4)$	$(x_4-\bar{x}_4)^2$
16	–2	4	20	0	0	32	+8	64	28	+2	4
20	+2	4	22	+2	4	24	0	0	20	–6	36
20	+2	4	16	–4	16	28	+4	16	24	–2	4
16	–2	4	22	+2	4	12	–12	144	32	+6	36
		16			24			224			80

विचलन वर्गों का योग (SSE) $= 16 + 24 + 224 + 80 = 344$

स्वातन्त्र्य संख्या $= N-K = 16-4 = 12$

न्यादर्शों के अन्दर प्रसरण (MSE) $= \frac{SSE}{N-K} = \frac{344}{12} = 28.67$

प्रसरण–विश्लेषण सारणी
(Analysis of Variance Table)

Sources	Sum of Squares (S.S.)	Degrees of Freedom (D.f)	Mean Squares (ms)	F–ratio
Between samples	160 (SSC)	(K–1) = 3	53.33 (MSC)	$\frac{53.33\ (MSC)}{28.67\ (MSE)} = 1.86$
Within Samples	344 (SSE)	(n–k) = 12	28.67 (MSE)	
Total Variance	504 (SST)	15 (N – 1)		

$V_1 = 3$ तथा $V_2 = 12$ के लिए 5% सार्थकता स्तर पर 'F' का तालिका मान 5.95 है । परिकलित मान केवल 1.86 है जो सारणी मान से कम है । अत: हमारी परिकल्पना सही है कि गेहूँ की चारों किस्मों में कोई अंतर नहीं है ।

(ब) ऋजु रीति (Short cut Method) :

प्रत्यक्ष रीति में प्रसरण के लिए न्यादर्शों के बीच एवं अंदर वर्गों के योग निकालने में अधिक समय लगता है तथा गणनाएँ भी जटिल होती हैं । इसलिए प्रसरण के लिए न्यादर्शों के बीच तथा अन्दर वर्गों का योग मालूम करने के लिए ऋजु रीति (लघुरीति) का प्रयोग किया जा सकता है । शेष प्रक्रिया यथावत रहती है ।

(i) न्यादर्श इकाइयों का योग और उनके वर्गों का योग ज्ञात करते हैं :–

न्यादर्श इकाइयों का योग $= \Sigma x_1, \Sigma x_2, \Sigma x_3, \ldots\ldots\ldots\Sigma x_k$

न्यादर्श इकाइयों के वर्गों का योग $= \Sigma x_1^2, \Sigma x_2^2, \Sigma x_3^2 \ldots\ldots\ldots\Sigma x_k^2$

(ii) **संशोधन कारक** (Correction factor) ज्ञात किया जाता है । इसके लिए सभी न्यादर्श इकाइयों के कुल योग के वर्ग में न्यादर्श इकाइयों की कुल संख्या का भाग दिया जाता है । इसका सूत्र है–

$$C.F. = \frac{T^2}{n}$$

CF = Correction factor

T^2 = Square of the Total units of the samples.

n = Total no. of units of samples.

(iii) वर्गों का कुल योग (SST) निकाला जाता है ।

$$SST = [\Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \ldots\ldots\ldots\ldots\ \Sigma x_k^2] - \frac{T^2}{n} \text{ or C F}$$

(iv) न्यादर्शों के बीच वर्गों का जोड़ निकाला जाता है।

$$SSC = \left[\frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} + \ldots\ldots \frac{(\Sigma x_k)^2}{n_k} \right] - \frac{T^2}{n} \text{ or } CF$$

(v) न्यादर्शों के अंतर्गत वर्गों का जोड़ निकालिए।

(vi) शेष प्रक्रिया प्रत्यक्ष रीति जैसी ही होती है (प्रसरण अनुपात निकालने तथा निर्वचन के लिए)।

Illustration 2 : The following table gives the results of experiments of four varieties of wheat grown in 24 plots :–

		Plot		
	A	B	C	D
Yield	2	5	20	5
	4	9	9	12
	8	8	14	9
	5	12	18	15
	6	12	18	17
		14	11	10
				16

Is there any significant difference in production of these varieties ?

Solution : Hypothesis :–There is no difference in the yield of four varieties of wheat i.e. variance is insignificant

(i) Totals of sample–units and their squares :–

A		B		C		D	
x_1	x_1^2	x_2	x_2^2	x_3	x_3^2	x_4	x_4^2
2	4	5	25	20	400	5	25
4	16	9	81	9	81	12	144
8	64	8	64	14	196	9	81
5	25	12	144	18	324	15	225
6	36	12	144	18	324	17	289
		14	196	11	121	10	100
						16	256
25	145	60	654	90	1446	84	1120

(ii) Correction factorr $= \dfrac{(\Sigma x_1 + \Sigma x_2 + \Sigma x_3 + \Sigma x_4)^2}{n}$

$$(CF) = \frac{(25 + 60 + 90 + 84)^2}{24} = \frac{(259)^2}{24} = 2795.94$$

(iii) SST $= \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 - CF$

$= 145 + 654 + 1446 + 1120 - 2795.04$

$= 3365 - 2795.04 = 569.96$

(iv) SSC $= \left[\dfrac{(\Sigma x_1)^2}{n_1} + \dfrac{(\Sigma x_2)^2}{n_2} + \dfrac{(\Sigma x_3)^2}{n_3} + \dfrac{(\Sigma x_k)^2}{n_4} \right] - CF$

$$= \left[\frac{(25)^2}{5} + \frac{(60)^2}{6} + \frac{(90)^2}{6} + \frac{(84)^2}{7} \right] - 2795.04$$

NOTES

$= [\ 125 + 600 + 1350 + 1008\ ] - 2795.04$

$= 3083 - 2795.04 = 287.96$

(v) SSE $= SST - SSC$

$= 569.96 - 287.96 = 282$

(vi) Analysis of variance Table for interpretation.

Source	S.S.	d.f.	M.S.	F–Ratio
Between Samples	287.96 (SSC)	$4 - 1 = 3$	$\frac{SSC}{K-1} = \frac{287.96}{3} = 95.99$	$\frac{95.99\ (MSC)}{14.10\ (MSE)} = 6.807$
Withn Samples	282 (SSE)	$24 - 4 = 20$	$\frac{SSE}{n-K} = \frac{283}{20} = 14.10$	
Total Variance	569.96 (SST)	23		

(vii) Interpretation :–

The calculated value of F (Fc) > 3.10 (F .05).

Hence, our hypothesis is not true. Hence, the difference in the yield of varieties of wheat is significant.

(स) मूल–बिन्दु परिवर्तन रीति (Coding Method) :–

प्रसरण अनुपात (F) की एक महत्वपूर्ण विशेषता यह है कि सभी न्यादर्शों के सभी पदमूल्यों में किसी सामान्य कारक (Common factor) या स्थिरांक को जोड़ने; घटाने, गुणा करने या भाग देने पर इस अनुपात का मान यथावत रहता है । इस रीति के प्रयोग का मूल उद्देश्य प्रसरण विश्लेषण की जटिल प्रक्रिया को सुगम बनाना है । इसके लिए दिये हुए मूल अकों में से कोई स्थिर संख्या घटा या जोड़ दी जाती है अथवा उससे गुणा या भाग कर दिया जाता है ताकि मूल अंक छोटे हो जाते हैं । इससे प्रसरण अनुपात निकालने की गणन क्रिया सरल हो जाती है और निष्कर्ष भी प्रभावित नहीं होता है । यह रीति उन न्यादर्शों के लिए उपयुक्त होती है जिनके पद मूल्य बड़े होते हैं ।

इस रीति की गणन–क्रिया लघुरीति की भाँति ही होती है (केवल मूल बिन्दु परिवर्तन क्रिया को छोड़कर)

Illustration 3 :- Prepare a table of Analysis of Variance for the following data and test if the varieties are different :–

Plots	Varieties			
	A	B	C	D
1	200	230	250	300
2	190	270	300	270
3	240	150	145	180
	630	650	695	750

Solution : - This illustration is of Two–way classification but only one factor i.e. difference of variety for the yield is significant or not, has been asked. Hence, we will not give any thought towards plots.

Hypothesis = The difference in the varieties for the yield is insignificant.

Codified Table after deducting 200 from the original data

NOTES

A		B		C		D	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2	X_4	X_4^2
0	0	+30	900	+50	2500	+100	10,000
–10	100	+70	4900	+100	10000	+70	4900
+40	1600	–50	2500	–55	3025	–20	400
$\Sigma x_1 = 30$	$\Sigma X_1^2 =$ 1700	$\Sigma X_2 = 50$	$\Sigma X_2^2 =$ 8300	$\Sigma X_3 = 95$	$\Sigma X_3^2 = 1552$ 5	$\Sigma X_4 =$ 150	$\Sigma X_4^2 = 15,3$ 00

T $= \Sigma x_1 + \Sigma x_2 + \Sigma x_3 + \Sigma x_4$

$= 30 + 50 + 95 + 150 = 325$

C.F. $= \dfrac{T_2}{N} = \dfrac{(325)^2}{12} = \dfrac{105625}{12} = 8,802.08$

SST $= [\Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2]$ –CF

$= [1,700+8,300+15,525+15,300]-8,802.08$

$= 40,825 - 8,802.08 = 32022.92$

SSC $= \left[\dfrac{(\Sigma x_1)^2}{n_1} + \dfrac{(\Sigma x_2)^2}{n_2} + \dfrac{(\Sigma x_3)^2}{n_3} + \dfrac{(\Sigma x_k)^2}{n_k}\right] - CF$

$= \left[\dfrac{(30)^2}{3} + \dfrac{(50)^2}{3} + \dfrac{(95)^2}{3} + \dfrac{(150)^2}{3}\right] - 8,802.92$

$= [300 + 833.3 + 3,008.3 + 7,500] - 8,802.08$

$- 11,608.06 - 8,802.08 = 2,806.52$

SSE $=$ SST – SSC

$= 32,022.92 - 2,806.52 = 29,216.40$

d.f. (Between samples) : $V_1 = k-1 = 4-1 = 3$

(Inside samples) : $V_2 = N-K = 12-4 = 8$

Total d.f. $= 11$

Table of Analysis of Variance

Source	S.S.	d.f.	MS	F–Ratio
Between Samples	2,806.52	3	935,51	$F = \dfrac{3652.05}{935.51}$ $= 3.90$
Within Samples	29,216.40	8	3,652.05	
Total Variance	32,022.92	11		

Interpretation : The calculated Value of Fc 3.90 < 4.07 (F.05)

Hence, Zero hypothesis is true , i.e. there is no difference in the varieties from the point of view of the yield.

Illustration 4 : The following data relate to the results obtained by four investigators investing a common problem. Each of the 4 investigators had taken a sample of 6 items. Do these results vary significantly from each other.

NOTES

Samples

A	B	C	D
126	102	114	138
141	126	150	114
120	90	120	120
111	120	141	102
120	96	120	141
150	108	111	132

Solution : Zero hypothesis that the results of the four investigators are not different from each other, using Coding method, deducting 120 from each investigation and then divide the balance by 3, we get :–

$$\frac{X - 120}{3}$$

Sample A		Sample B		Sample C		Sample D	
(X_1)	$(X_1)^2$	(X_2)	$(X_2)^2$	(X_3)	$(X_3)^2$	(X_4)	$(X_4)^2$
+2	4	–6	36	–2	4	+6	36
+7	49	+2	4	+10	100	–2	4
0	0	–10	100	0	0	0	0
–3	9	0	0	+7	49	–6	36
0	0	–8	64	0	0	+7	49
10	100	–4	16	–3	9	+4	16
$\Sigma x_1=16$	$\Sigma x_1^2=162$	$\Sigma x_2=-26$	$\Sigma x_2^2=220$	$\Sigma X_3=12$	$\Sigma x_3^2=162$	$\Sigma x_4=9$	$\Sigma x_4^2=141$

(i) T $= \Sigma x_1 + \Sigma x_2 + \Sigma x_3 + \Sigma x_4 = (16) + (-26) + (12) + (9) = 11$

(ii) CF $= \frac{T_2}{N} = \frac{(11)^2}{24} = 5.04$

(iii) d.f $=$ n–K $= 24\text{-}4 = 20$; K–1 $= 4–1 = 3$

(iv) SST $= \Sigma x_1^2 + \Sigma x_2^2 + \Sigma x_3^2 + \Sigma x_4^2 - CF$

$= 162 + 220 + 162 + 141 = 685 - 5.04 = 679.96$

(v) SSC $= \left[\frac{(\Sigma x_1)^2}{n_1} + \frac{(\Sigma x_2)^2}{n_2} + \frac{(\Sigma x_3)^2}{n_3} + \frac{(\Sigma x_4)^2}{n_4}\right] - CF$

$= \left[\frac{(16)^2}{6} + \frac{(-26)^2}{6} + \frac{(12)^2}{6} + \frac{(9)^2}{6}\right] - 5.04$

$= 192.83 - 5.04 = 187.79$

(vi) SSE $=$ SST – SSC

$= 679.96 - 187.79 = 492.17$

Table of Analysis of Variance

Source of Variation	d.f.	S.S.	M.S.	F	F.05 (Table value)
Between Samples	$V_1=3$	187.79 (SSC)	$\frac{187.79}{3} = 62.6$ (MSC)	$\frac{MSC}{MSE}$	3.10
Within Samples	$V_2=20$	492.17 (SSE)	$\frac{492.17}{20} = 24.61$ (MSE)	$\frac{62.6}{24.61} = 2.54$	
Total of Variance	23	679.96 (SST)			

$\because F_c < F_{.05}$ (t)

Hence, there is no difference between the results obtained by the four investigators.

NOTES

Illustration 5 : (Analysis of grouped data) :–

Two samples of 40 and 50 Students Score the following marks in a test carrying maximum marks 15 : –

Marks	:	10	11	12	13	14	15	
f_1	:	6	14	12	6	2	0	= 40
f_2	:	2	6	10	16	14	2	= 50

Does the average level of knowledge of the students differ ?

Solution :

Table

		Sample – 1			Sample – 2		
Marks X	x^2	f_1	$f_1 x$	$f_1 x^2$	f_2	$f_2 x$	$f_2 x^2$
10	100	6	60	600	2	20	200
11	121	14	154	1,694	6	66	726
12	144	12	144	1,728	10	120	1,440
13	169	6	78	1,014	16	208	2,704
14	196	2	28	392	14	196	2,744
15	225	0	0	0	2	30	450
		$\Sigma f_1 = 40$	$\Sigma f_1 x = 464$	$\Sigma f_1 x^2 = 5,428$	$\Sigma f_2 = 50$	$\Sigma f_2 x = 640$	$\Sigma f_2 x^2 = 8264$

(i) T $= \Sigma f_1 x + \Sigma f_2 x = 464 + 640 = 1104$

(ii) CF $= \frac{T^2}{N} = \frac{(1104)^2}{40 + 50} = \frac{12,18,816}{90} = 13,542.4$

(iii) SST $= \Sigma f_1 x^2 = \Sigma f_2 x^2 - CF$

$= 5,428 + 8,264 - 13,542.4$

$= 13,962 - 13,542.4 = 149.6$

(iv) SSC (Between Samples) $= \left[\frac{(\Sigma f_1 x)^2}{n_1} + \frac{(\Sigma f_2 x)^2}{n_2}\right] - CF$

$= \left[\frac{(464)^2}{40} + \frac{(640)^2}{50}\right] - 13,542.4$

$= \left[\frac{(215296)}{40} + \frac{(409600)}{50}\right] - 13,542.4$

$= 5,382.4 + 8,192 - 13,542.4$

$= 13,574.4 - 13,542.4 = 32$

(v) SSE (Within Samples) $= SST - SSC$

$= 149.6 - 32 = 117.6$

(vi) d.f. (Degrees of freedom)

(a) Between samples (v_1) $= k-1 = 2-1 = 1$

(b) Within Samples (v_2) $= N - k = 90-2 = 88$

NOTES

Analysis of variance–Table

Source of Variation	S.S.	df	MS	F–Ratio	$F_{.05}$
Between Samples Within Samples	32 (SSC) 117.6 (SSE)	1 88	32 (MSC) 1.33 (MSE)	$\frac{32\ (MSC)}{1.33\ (MSE)}$ = 24.06	9,28
Total	149.6 (SST)	89			

$\because \quad Fc > F_{.05}$ (t)

Hence, the average level of knowledge of the students differs significantly.

प्रसरण विश्लेषण

1. प्रसरण विश्लेषण किसे कहते हैं?

....

....

....

2. प्रसरण विश्लेषण की मान्यता लिखिए।

....

....

....

3. प्रसरण विश्लेषण की विधि को समझाइए।

....

....

....

4.7 प्रसरण–विश्लेषण तकनीक–द्वि–मार्गीय वर्गीकरण (Analysis of Variance Technique–Two way chassification)

एक मार्गीय वर्गीकरण में एक कारक के अनुसार ही हम अपने प्रेक्षणों (Investigations) को वर्गों में बाँटते हैं जबकि द्वि–मार्गीय वर्गीकरण में इन्हें दो कारकों के अनुसार अलग-अलग वर्गों में बाँटते हैं। उदाहरण के लिए, एक प्लॉट की उपज का विश्लेषण दो कारकों–खाद की किस्म और बीज की किस्म के आधार पर किया जाता है। इसी प्रकार औद्योगिक उत्पादन पर मशीन एवं श्रमिकों के प्रकारों का भिन्न-भिन्न प्रभाव देखा जाता है।

इसमें वर्गीकरण द्वि–मार्गीय होता है। स्तम्भ (Column) एक प्रकार के कारक को व्यक्त करता है जबकि पंक्तियाँ (Rows) दूसरे प्रकार के कारक को। इस वर्गीकरण में प्रसरण विश्लेषण करते समय वर्गों के कुल योग (Total sum of squares) को तीन भागों में बाँटते हैं :-

(i) अन्तःस्तम्भ वर्ग (Sum of Squares between Columns–SSC)

(ii) अन्तः पंक्ति वर्ग (Sum of Squares between rows–SSR)

(iii) अवशिष्ट वर्ग या त्रुटिवर्ग (Sum of Squares due to error–SSE)

द्वि–मार्गीय विश्लेषण की गणना विधि :

(i) गणनाओं की जटिलता से बचने के लिए मूल बिन्दु परिवर्तित किया जाता है। (Coding Method)

(ii) संशोधन कारक (Common factor) निकाला जाता है। सूत्रः

$$C.F. = \frac{T^2}{N}$$

(iii) कुल वर्ग योग (SST) निकाला जाता है । सूत्र :

$$SST = SSC + SSR + SSE - CF$$

(iv) अन्तः स्तम्भ वर्ग योग (SSC) ज्ञात किया जाता है । सूत्रः

$$SSC = \Sigma \left[\frac{(\Sigma xc)}{n_c} \right] - CF$$

Σx_c = प्रत्येक स्तम्भ के मूल्यों के योग का वर्ग

n_c = प्रत्येक स्तम्भ में इकाई संख्या

(v) अन्तः पंक्ति वर्ग योग (SSR) की गणना की जाती है । सूत्र :

$$SSR = \Sigma \left[\frac{(\Sigma xr)_2}{n_r} \right] - CF$$

Σx_r = प्रत्येक पंक्ति के मूल्यों के योग का वर्ग

n_r = प्रत्येक पंक्ति में इकाई संख्या

(vi) त्रुटिवर्ग योग (SSE) की गणना की जाती है । सूत्र :

$$SSE = SST - (SSC + SSR)$$

(vii) स्वातन्त्र्य संख्या (d.f.) निकालना :

(अ) d.f. for columns = $c-1$

(ब) d.f. for rows = $r-1$

(स) d.f. for residual = $(c-1)(r-1)$

(viii) उपरोक्त प्राप्त सूचनाओं को प्रसरण-विश्लेषण सारणी के रूप में प्रदर्शित करना ।

(ix) **निर्वचन** :- न्यादर्श स्तम्भ एवं न्यादर्श पंक्ति दोनों के लिए परिगणित प्रसरण अनुपातों (F-Ratio) को उनके सारणी-मानों से तुलना की जाती है । यदि परिकलित मान, सारणीमान से अधिक है तो अन्तर सार्थक है, अतः शून्य परिकल्पना असत्य है । यदि परकलित मान, सारणीमान से कम है तो शून्य परिकल्पना सत्य है ।

Illustration 6 : Three varieties- A, B and C of a crop are tested in a randomized block design with four replications (पुनरावृत्तियाँ) , the layout being given in the diagram below. The plot yields (in pounds) are also indicated therein. Analyse the experimented yield and state your conclusions.

A–6	C–5	A–8	B–9
C–8	A–4	B–6	C–9
B–7	B–6	C–10	A–6

Solution- Zero Hypothesis :- There is no difference in the Block-mean and the kind mean.

Deducting 6 from all experiments, we get

Varieties	Blocks				Total
	1	2	3	4	
A	0	–2	+2	0	0
B	+1	0	0	+3	4
C	+2	–1	+4	+3	8
Total	3	–3	6	+6	12

(i) Grand Total (T) = 12

NOTES

(ii) $CF = \frac{T^2}{N} = \frac{(12)^2}{12} = 12$

(iii) $SST = [(0)^2 + (1)^2 + (2)^2 + (-2)^2 + (0)^2 + (-1)^2 + (2)^2 + (0)^2 + (4)^2 + (0)^2 + (3)^2 (3)^2] - 12$

$= 0 + 1 + 4 + 4 + 0 + 1 + 4 + 0 + 16 + 0 + 9 + 9 - 12 = 36$

d.f. $= C_r - 1 = 4 \times 3 - 1 = 11$

(iv) $SSC = \frac{(3)^2}{3} + \frac{(-3)^2}{3} + \frac{(6)^2}{3} + \frac{(6)^2}{3} - 12$

$= 3 + 3 + 12 + 12 - 12 = 18$

df. $= (c - 1) = (4 - 1) = 3$

(v) $SSR = \frac{(0)^2}{4} + \frac{(4)^2}{4} + \frac{(8)^2}{4} - 12$

$= 0 + 4 + 16 - 12 = 8$

df. $= (r - 1) = (3 - 1) = 2$

(vi) $SSE = SST - (SSC + SSR)$

$= 36 - (18 + 8) = 10$

df. $= (c - 1)(r - 1) = (4 - 1)(3 - 1) = 3 \times 2 = 6$

Analysis of Variance - Table

Sources of Variation	S.S.	d.f.	M.S.	(Fc)	F.05(t)
Between Blocks (Columns)	18	3	6	$F2 = \frac{6}{1.67} = 3.58$	4.76
Between Varieties (Rows)	8	2	4	$F1 = \frac{4}{1.67}$	5.14
Residual	10	6	1.67		
Total	36	11			

Both the Calculated Values of F are less than the tabulated Values.
Hence, the difference between the experiments of Blocks and kinds is insignificant.

4.8 लेटिन वर्ग में प्रसरण विश्लेषण (Analysis of Variance in Latin Square)

यह द्विमार्गीय विश्लेषण का ही व्यापक रूप है। इसका प्रयोग अधिकतर कृषि अनुसंधान के क्षेत्र में होता है। इस तकनीक का आधार है कि एक मानक के विभिन्न प्रकार एक से नहीं होते हैं। अत: विभिन्न कारकों को समान प्रकारों में बाँटकर जाँच करना चाहिए। जैसे कि, खाद की विभिन्न किस्मों का प्रभाव कृषि उपज पर देखना है तो इसके लिए एक खेत को उतने खण्डों में बाँटना होगा जितनी कि खाद की किस्में हैं और फिर प्रत्येक खण्ड को उतने ही उपखण्डों में बाँटना होगा जिनके किस्म की खादों का प्रयोग करना है। ऐसा करना इसलिए आवश्यक है कि एक खेत की सभी मिट्टी समान रूप से उपजाऊ नहीं होती है।

लैटिन वर्ग में प्रसरण-विश्लेषण प्रायोगिक अभिकल्पना पर आधारित है। अभिकल्पना का आयोजन इस प्रकार किया जाता है कि प्रत्येक प्रकार की खाद का प्रयोग उपखण्ड पर हो। उपखण्डों के लिए अंग्रेजी वर्णमाला A, B, C तथा D का प्रयोग होता है।

NOTES

लैटिन-वर्ग के मानक प्रारूप इस प्रकार के हो सकते हैं :-

(4 ×4)

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

(5 × 5)

A	B	C	D	E
B	C	D	E	A
C	D	E	A	B
D	E	A	B	C
E	A	B	C	D

उपरोक्त वर्ग-विन्यास से पता चलता है कि प्रत्येक वर्णाक्षर का प्रयोग स्तम्भ या पंक्ति में एक बार ही हुआ है चाहे क्रम कुछ भी रहा हो।

प्रसरण-विश्लेषण प्रक्रिया :

यहाँ कुल वर्ग योग चार संघटक में बँट जाता है :-

(i) अन्तः स्तम्भ वर्ग योग

(ii) अन्तः पंक्ति वर्ग योग

(iii) अन्तः अभिक्रिया वर्ग योग (Between Treatment S.S.)

(iv) अवशिष्ट वर्ग योग

ऊपर वर्णित योगों का अध्ययन पूर्व में किया जा चुका है, केवल अभिक्रिया वर्ग योग के। इसकी गणना निम्नलिखित सूत्र से की जाती है :-

$$\text{Between Treatment S.S.} = \frac{\Sigma(\Sigma x_1)^2}{n} - CF$$

प्रत्येक वर्ग से संबंधित मूल्यों का योग (Σx_1) मालूम करके सभी वर्ग योगों के वर्गों का कुल योग $(\Sigma x_1)^2$ मालूम कर लिया जाता है।

स्वातंत्र्य संख्याएँ प्रथम तीन वर्गों के लिए $(n-1)$ के द्वारा निकाली जाती है क्योंकि यहाँ $n = c = r$ है। अवशिष्ट वर्ग के लिए $d.f. = (n-1)(n-2)$ और कुल वर्ग के लिए $d.f. = n^2 - 1$ होता है।

Illustration 7 : A latin square design experiments were conducted on three plots of Sugarcane. The design and yield per plot are given below. Analyse the data and interpret the results.

A 20	B 18	C 16
B 15	C 20	A 25
C 25	A 15	B 20

Solution :

यह 3 × 3 लेटिन वर्ग है। प्रत्येक अंक में से 20 कम करने पर सारणी इस प्रकार होगी।

Basic point x = 20 Square

	1	2	3	ΣXr	1	2	3	Squares Total
1	0	–2	–4	–6	0	4	16	20
2	–5	0	+5	0	25	0	25	50
3	+5	–5	0	0	25	25	0	50
Σx_c	0	–7	+1	–6 (T)	50	29	41	120

NOTES

(i) $T = -6$

(ii) $C.F. = \frac{T^2}{N} = \frac{(-6)^2}{3 \times 3} = \frac{36}{9} = 4$

(iii) $TSS = \Sigma(x_1^2 + x_2^2 + \ldots . x_k^2) - C.F.$

$= 120 - 4 = 116$

$d.f. (t) = n^2 - 1 = (3)^2 - 1 = 9 - 1 = 8$

(iv) SSC (Between Columns S.S.) $= \frac{\Sigma (\Sigma x_c)^2}{n} - C.F.$

$= \frac{\Sigma [(0)^2 + (-7)^2 + (1)^2]}{3} - 4$

$= 16.67 - 4 = 12.67$

$d.f.(c) = (n - 1) = (3) - (1) = 2$

(v) SSR (Between Rows S.S.) $= \frac{\Sigma (\Sigma x_r)}{n} - C.F.$

$= \frac{\Sigma[(-6)^2 + (0)^2 + (0)^2]}{3} - 4$

$= \frac{36}{3} - 4 = 8$

$d.f. (r) = (n - 1) = 3 - 1 = 2$

(vi) Between Treatment SS :– इसके लिए सारणी को पुनः व्यवस्थित करना होगा। इसमें प्रत्येक वर्ग से संबंधित मूल्य इसी में रखना होगा।

इस प्रकार सारणी वर्गानुसार विन्यासित होगी –

	A	B	C	Total
1	0	–2	–4	–6
2	+5	–5	0	0
3	–5	0	+5	0
$\Sigma x_{(1)}$	0	–7	+1	–6

Now, Between Treatment S.S. $= \frac{\Sigma \left[(xt)^2 \right]}{n} - CF$

$= \frac{\Sigma \left[(0)^2 + (-7)^2 + (1)^2 \right]}{3} - 4$

$= \frac{50}{3} - 4 = 16.67 - 4 = 12.67$

$d.f. (t) = n - 1 = 3 - 1 = 2$

(vii) Residual S.S. $= TSS - (SSC + SSR + SST)$

$= 116 - (12.677 + 8 + 12.67)$

$= 116 - 33.34 = 82.66$

$d. f. = (n-1)(n-2)$

$= (3-1)(3-2) = 2$

Analysis or Variance - Table

Source of Variation	S.S.	d.f.	M.S.	Fc	F.05(1)
Between Column	12.67	2	6.34	$F_1 = \frac{6.34}{41.33} = 0.15$	19
Between Rows	8.00	2	4.00		
Between Treatment	12.67	2	6.34	$F_2 = \frac{4}{41.33} = 0.09$	19
Residual	82.66	2	41.33	$F_3 = \frac{6.34}{41.33} = 0.15$	19
Total	116.00	8			

(viii) **Interpretation :** As the Calculated Value (Fc) $< eF_{.05}$ (t) in all cases hence, the difference is insignificant.

4.9 प्रसरण विश्लेषण की उपयोगिता (Utility of Analysis of Variance)

प्रारंभ में इस तकनीक का उपयोग कृषि अनुसंधान तक सीमित रहा, लेकिन आजकल इसका विज्ञान के कई क्षेत्रों में प्रयोग होता है जैसे अर्थशास्त्र, समाजशास्त्र, शिक्षा, मनोविज्ञान आदि ।

प्रयोगात्मक अनुसंधान के लिए तो यह अत्यधिक उपयोगी है। इसकी सहायता से निष्कर्षों की परिशुद्धता एवं संक्षिप्तता को प्राप्त किया जाता है । इससे यह जानने में सहायता मिलती है कि विभिन्न प्रयोगों के निष्कर्ष में अर्थपूर्ण (Significant) अंतर है या नहीं । इस तकनीक का प्रयोग प्रमुख रूप से निम्नलिखित क्षेत्रों में किया जाता है–

(1) **न्यादर्श (दो से अधिक) माध्यों के अन्तर की जाँच :-**

दो से अधिक न्यादर्शों के माध्यों के बीच अंतर की जाँच अन्य किसी परीक्षण द्वारा एक बार में करना संभव नहीं है ।इस तकनीक द्वारा अनेक माध्यों के अन्तर की सार्थकता को एक ही बार में जाँचा जा सकता है ।

(2) **प्रसरणों के अंतर की सार्थकता की जाँच :-**

प्रसरण अनुपात (F–Ratio) द्वारा विभिन्न न्यादर्शों के प्रसरणों के अन्तर की सार्थकता को जाँचा जाता है ।

(3) **द्विमार्गीय वर्गीकरण में सजातीयता का अध्ययन :-**

इस तकनीक द्वारा समंक–समूह पर दो प्रकार के व्यवहारों के भिन्न -भिन्न प्रभाव को मालूम किया जा सकता है। इस रीति द्वारा द्वि–मार्गीय वर्गीकरण में कई वर्गों के अंतर्गत सजातीयता का अध्ययन किया जा सकता है।

(4) **सह संबंध एवं प्रतीपगमन का परीक्षण :-**

इस तकनीक का प्रयोग सह संबंध अनुपात एवं बहुगुणी सह संबंध गुणक की सार्थकता की जाँच में भी किया जाता है । इसकी सहायता से प्रतीपगमन रेखा की प्रकृति की जाँच भी की जा सकती है।

इस प्रकार प्रसरण– विश्लेषण तकनीक प्रयोगों के आधुनिक युग में बहुत उपयोगी एवं महत्वपूर्ण है।

4.10 बोध प्रश्न

1. द्वि-मार्गीय विश्लेषण की गणना विधि समझाइए।

....

....

....

....

....

NOTES

2. प्रसरण विश्लेषण की उपयोगिता लिखो।

....

....

....

4.11 सारांश

जब दो से अधिक न्यादर्शों के अन्तरो की सार्थकता का अध्ययन करना होता है तो वहाँ प्रसरण विश्लेषण विधि का प्रयोग किया जाता है। प्रसरण विश्लेषण निम्न मान्यताओं समग्र का वितरण, न्यादर्श स्वतंत्र होना व संयोज्यता पर आधारित होता है। प्रसरण विश्लेषण, एक मार्गीय, द्विमार्गीय वर्गीकरण तकनीक का प्रयोग किया जाता है। यह तकनीक वहाँ महत्वपूर्ण है जहाँ न्यादर्श (दो से अधिक) माध्यों के अंतर की जाँच, प्रसरणों के अंतर की सार्थकता की जाँच एवं सह संबंध एवं प्रतीपगमन का परीक्षण करना होता है।

4.12 शब्द कोष

प्रसरण, संयोज्यता, ऋजुरीति

4.13 अभ्यास प्रश्न

दीर्घ उत्तरीय प्रश्न (Long Answer Type Questions)

1. प्रसरण- विश्लेषण क्या है ? प्रसरण–विश्लेषण विधि को समझाइये जबकि प्रेक्षण एकमार्गी वर्गीकृत हो ।
   What is Analysis of Variance ? Explain clearly the technique of Analysis of Variance for data with one–way classification.
2. प्रसरण–विश्लेषण की उदाहरण सहित व्याख्या कीजिये ।
   Explain with illustrations the Analysis of Variance technique.
3. अन्तर्निहित परिकल्पनाओं के विशेष सन्दर्भ में प्रसरण–विश्लेषण के मूल सिद्धान्तों की विवेचना कीजिये ।
   Discuss the fundamental principles of Analysis of Variance with special reference to the assumptions made therein.
4. प्रसरण–विश्लेषण क्या है और इसके अध्ययन का क्या महत्व है ? इसके अध्ययन में काम आने वाले F गुणांक का अर्थ समझाइये ।
   What is Analysis of Variance ? What is the significance of its study ? Also explain the meaning of F Coefficient used in this study.
5. द्वि–साधन वर्गीकरण में प्रसरण–विश्लेषण की प्रविधि की स्पष्ट व्याख्या कीजिये। प्रसरण–विश्लेषण की उपयोगिता को भी समझाइये ।
   Explain clearly the technique of Analysis of variance for the data with two–way classification. Also give the use of Analysis of variance.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. प्रसरण विश्लेषण क्या है ?
   What is Analysis of Variance?
2. प्रसरण विश्लेषण विधि को समझाइये।
   Explain clearly the technique of Analysis of Variance.
3. F गुणांक का अर्थ समझाइये।
   Explain the meaning of F Coefficient.

4. प्रसरण विश्लेषण के मूल सिद्धान्तों की विवेचना कीजिये।
Discuss the fundamental principles of Analysis of Variance.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. प्रसरण विश्लेषण में किसकी जाँच की जाती है-
(अ) बहुलक, (ब) माध्य, (स) गुणक, (द) प्रसरण।
2. प्रसरण विश्लेषण का प्रतिपादन किसने किया है-
(अ) फिशर, (ब) गौसेट, (स) कार्ल पियर्सन, (द) जार्ज।
3. कुल प्रसरण को कितने घटकों में बाँटा जाता है-
(अ) 3, (ब) 2, (स) 4, (द) 5।
4. प्रसरण विश्लेषण के लिये निम्न रीतियाँ हैं-
(अ) प्रत्यक्ष, (ब) ऋजु, (स) मूल बिन्दु, (द) सभी।

[उत्तर- 1. (ब), 2. (अ), 3. (ब), 4. (द) .]

व्यावहारिक प्रश्न

1. निम्न समंकों से एक प्रसरण विश्लेषण तालिका की रचना करिये :
Set up a table of Analysis of Variance for the following data.

	Varieties			
Plots	a	b	c	d
1	200	230	250	300
2	190	270	300	270
3	240	150	145	180

[F = 3.85 < F.05, Not Significant]

2. एक नगर के विभिन्न Grammar Schools में योग्यता सम्बन्धी परीक्षण के परिणामों के सम्भावित अन्तर का मूल्यांकन करने के उद्देश्य से यादृच्छिक रीति से चारों सम्बन्धित स्कूलों के पाँचवीं कक्षा के कुछ चुने हुए छात्रों के लिये एक सामान्य परीक्षा आयोजित की गई। उसके परिणाम निम्न प्रकार से थे। समंकों का विश्लेषण करिये।
To assess the significance of possible variation in performance in a certain test between the grammar schools of a city, a common test was given to a number of students taken at random from the senior fifth class of each of the four schools concerned. The results are given below. Make an analysis of data.

Schools			
A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15

[F Cal. = 1.285 F Cal. < F Tab. (Not Significant)]

3. नीचे यादृच्छिक रूप से चुने हुए तीन खेतों में बोये गये दो किस्म के बीजों की प्रति एकड़ उपज दी हुई है।उपज की दृष्टि से किस्मों के मध्य अन्तर की सार्थकता के परीक्षण के लिये प्रसरण विश्लेषण का प्रयोग कीजिये।
(सम्बद्ध स्वातन्त्र्यांशों के लिये F का 5% पर मान 7.71 है।)

NOTES

Given below is the per acre yield of two strains of wheat sown in three randomly selected farms. Use analysis of variance to test whether the strains are significantly different is regard to yield.

(For relevent degrees of freedom the value of F at 5% 7.71)

	Qnt.	Qnt.	Qnt.
A	30	32	22
B	20	18	16

[F Cal. 9.4; diff. Significant]

4. गेहूँ की तीन किस्में–A, B और C खेतों के 12 खण्डों में बोयी गई जिनकी उपज के आँकड़े (कि.ग्रा) निम्न प्रकार से थे :–

The following figures relate to production in Kg. of three varieties A,B,C of wheat sown in 12 plots.

A	14	16	18		
B	14	13	15	22	
C	18	16	19	19	20

क्या तीनों किस्मों के उत्पादन का अन्तर महत्वपूर्ण है ?

Is there any significant difference in the production of three varieties.

[F Cal. = 1.157 F Cal. < F Tab. (Not Significant)]

5. खेत के प्रत्येक 4–4 खण्डों पर पाँच प्रकार के खादों का प्रयोग किया गया । अग्र तालिका में प्रत्येक खण्डों के लिए आलुओं की उपज के आँकड़े दिए हुए हैं । यह परीक्षण करिए कि क्या भिन्न–भिन्न प्रकार की खादों के कारण उपज का अन्तर महत्वपूर्ण है ?

Five fertilizers were applied to four pots each. The yields of potatoes on each of these plots are given in the following table. Examine whether the effects of these fertilizers on the yields are significantly different. [F 5% = 3.06 given]

	Fertilizers				
Plots	(1)	(2)	(3)	(4)	(5)
a	19	22	26	18	21
b	25	19	23	26	22
c	17	19	22	20	21
d	21	18	25	23	24

[F Cal.= 1.79 F Cal.< F Tab.(Not Significant)]

6. तीन प्रकार की gasoline पर किये गये अनेक परीक्षणों से प्राप्त औसत निम्न प्रकार से थे (Km./100 litres):

Several tests with three types of gasoline revealed the following averages (Kilometers/100 litres):

Gasolines		
A	B	C
519	523	520
521	520	517
520	522	521
518	520	519
512	524	520
521	523	517

एक प्रसरण विश्लेषण तालिका की रचना करिये । इसकी गणना करिये तथा 5% स्तर पर तीन प्रकार की gasoline के अन्तर की सार्थकता की जाँच करिये ।

Set up an analysis of variance table. Calculate F & test the significance of differences between the three types of gaxoline at 5% level.

(F Cal. = 5.83 F Cal. > F Tab. Difference Significant]

7. मई, जून एवं जुलाई के तीन महीनों में 4 सेल्समैनों द्वारा बेचे गए रेफ्रीजरेटरों की संख्या निम्न तालिका में दी हुई है :

Following table gives the number of refrigerators sold by 4 salesmen in three months–May, June & July.

Month	A	B	C	D
May	50	40	48	39
June	46	48	50	45
July	39	44	40	39

क्या चारों सेल्समैनों द्वारा दी गई बिक्री के मध्य अन्तर महत्वपूर्ण हैं ?

Is there a significant difference in the sales made by the four salesmen ?

[F Cal. = 1.55 F Tab. > F Cal. Not Significant]

8. एक कारखाने का प्रबन्धक, जो उत्पादन की किसी प्रक्रिया में प्रयोग की जाने वाली मशीनें खरीदना चाहता है, ऐसी मशीनों का उत्पादन करने वाली चार कम्पनियों में से प्रत्येक से एक–एक मशीन खरीदता है । वह तीन व्यक्तियों को जो इन चारों मशीनों पर एक–एक दिन काम करते हैं यादृच्छिक ढंग से नियुक्त करता है । निम्न तालिका से उत्पादित इकाइयाँ दी गई हैं :–

विभिन्न प्रकार की मशीनों एवं व्यक्तियों के मध्य उत्पादित इकाइयों के अन्तर की सार्थकता की जाँच कीजिये ।

(1 प्रतिशत महत्ता स्तर पर F का मान 2 एवं 6 d.f. के लिये 10.92 तथा 3 और 6 d.f. के लिये 9.78 है ।)

A Factory Manager, wishing to buy machines for a certain operation in a production process, obtains one machine from each of the four companies making such machines, and puts three men each of whom works one day on each of the four machine in a random order. The resulting units are given in a table given below.

Discuss the significance of Variation of production among the different types of machines and also among the workers.

(One percent value of F for 2, 6 and 3, 6 degrees of freedom is 10.92 and 9.78)

	Workers		
Machines	W_1	W_2	W_3
M_1	62	63	64
M_2	64	66	68
M_3	67	67	70
M_4	68	69	69

Among Workers– F Cal. (9.45) < $F_{.01}$ Tab. 10.92 (Diff. Significant)

Among machines–F Cal. (29) F .01 Tab. 9.78 (Diff. Significant)

9. तीन प्रकार के विभिन्न क्षेत्रों में चार सेल्समैनों द्वारा की गई औसत मासिक बिक्री (हजार रुपयों में) के आँकड़े निम्न तालिका में दिए हुए हैं :–

The following table gives the average monthly sale (in thousands of rupees) of four salesmen in three different types of territories.

NOTES

Salesmen

Territory	A	B	C	D	Teritory Totals
X	5	4	4	7	20
Y	7	8	5	4	24
Z	9	6	6	7	28
	21	18	15	18	72

उपर्युक्त तालिका से एक प्रसरण विश्लेषण तालिका की रचना कीजिए।

Set up a table of analysis of variance for the table given above.

[F Ca. = .75 Between Salesmen.] [F Cal. = 1.50 „ „ Territories]

10. एक संयंत्र की विभिन्न निर्माणशालाओं द्वारा उत्पादित इकाइयाँ निम्न प्रकार से हैं। उक्त समंकों पर प्रसरण–विश्लेषण कीजिये :–

Units produced by different foundry–shops of a plant are as under. Perform analysis of variance of these data.

Through–put obtained

	A	84	60	40	47	34				
Foundry	B	67	92	95	40	98	60	59	108	86
	C	46	93	100						

[F Cal. = 2.2 $< F_{.05}$, Difference not significant]

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय-5 आकस्मिकता एवं काई वर्ग परीक्षण
(CONTINGENCY AND CHI SQUARE TEST)

इकाई की रूपरेखा

5.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. आकस्मिकता गुणक की आवश्यकता को समझ सकेंगे।
2. काई वर्ग की जाँच के उपयोग का ज्ञान होगा।
3. काई वर्ग महत्ता के स्तर का ज्ञान होगा।

5.1 प्रस्तावना

गुण-सम्बन्ध का अध्ययन करते समय हमने केवल द्वन्द्वात्मक वर्गीकरण के आधार पर ही दी गई समंक सामग्री का अध्ययन किया है। हमारी यह मान्यता रही है कि समंकों में किसी विशेष गुण की उपस्थिति अथवा अनुपस्थिति होती है, लेकिन कभी-कभी तीसरी और चौथी स्थितियाँ भी उत्पन्न हो सकती हैं और व्यवहारिक जीवन में ऐसा होता है। द्वन्द्वात्मक वर्गीकरण के अनुसार यदि शिक्षा का वर्गीकरण करना है तो एक वर्ग शिक्षितों (A) का तथा दूसरा अशिक्षितों (a) का होगा। किन्तु वास्तविक स्थिति बीच की भी हो सकती है जैसे कुछ लोग न तो शिक्षितों के वर्ग में आते हैं और न ही अशिक्षितों के। ऐसी स्थिति में बहुमुखी वर्गीकरण की आवश्यकता होती है और जो सारणी (Table) इस प्रकार से बनायी जाती है, वह आकस्मिकता सारणी (Contingency table) कहलाती है।

इसको निम्न उदाहरण द्वारा स्पष्ट किया जा सकता है—

Contingency Table showing the Temperament of Brothers and Sisters

Brother (A)	Sisters (B)			Total
	Quick (B_1)	Good Natured (B_2)	Sullen (B_3)	
Quick (A_1)	($A_1 B_1$)	($A_1 B_2$)	($A_1 B_3$)	(A_1)
GoodNatured (A_2)	($A_2 B_1$)	($A_2 B_2$)	($A_2 B_3$)	(A_2)
Sullen (A_3)	($A_3 B_1$)	($A_3 B_2$)	($A_3 B_3$)	(A_3)
Total	**(B_1)**	**(B_2)**	**(B_3)**	**N**

NOTES

ऐसी सारणियों का गुण-सम्बन्ध ज्ञात करने के लिये सबसे सरल उपाय यह है कि सारणियों को 2×2 में परिवर्तित कर दिया जाये। ऐसा करने के लिये गुणों के कई वर्गों को केवल दो वर्गों में कर दिया जाता है। ऐसा सदैव सम्भव नहीं होता है।

Illustration 1 : From the table given below, find out the Coefficient of Association between intelligence in fathers and in sons (निम्नलिखित सारणी से पिता और पुत्रों की बुद्धिमत्ता में गुण संबंध गुणक निकालिए) : —

Fathers (A)	Sons (B)				Total
	Very Intelligent	Intelligent	**Ordinary**	**Dull**	
Very Intelligent	20	16	10	2	48
Intelligent	10	14	8	4	36
Ordinary	14	12	18	6	50
Dull	4	6	6	24	40
Total	48	**48**	42	**36**	**174**

Solution– The above table will be converted into 2×2 table by merging very intelligent and intelligent, and ordinary and dull. The changed table will be as follows :-

Fathers (A)	Sons (B)		Total
	Intelligent (B)	Dull (b)	
Intelligent (A)	60 (AB)	24 (Ab)	A = 84
Dull (a)	36 (aB)	54 (ab)	a = 90
Total	(B) = 96	**(b) = 78**	**N = 174**

Coefficient of Association or

$$Q = \frac{(AB)\,(ab) - (Ab)\,(aB)}{(AB)\,(ab) + (Ab)\,(aB)} = \frac{(60 \times 54) - (24 \times 36)}{(60 \times 54) + (24 \times 36)}$$

$$= \frac{3240 - 864}{3240 + 864} = \frac{2376}{4104} = 0.58$$

5.2 आकस्मिकता गुणक (Coefficient of Contingency)

किसी गुण से उपस्थिति या अनुपस्थिति के आधार पर दो श्रेणियों में सम्बन्ध मालूम करने के लिए कार्ल पियर्सन द्वारा प्रतिपादित माध्य वर्ग आकस्मिकता गुणक (Coefficient of Mean Square Contingency) निकाला जाता है। इसे संक्षेप में आकस्मिकता का गुणक (Coefficient of Contingency) कहते हैं। इसे प्राप्त करने के लिए पहले काई-वर्ग (Chi-Squarc) निकालना पड़ता है। काई-वर्ग निम्न सूत्र द्वारा निकाला जाता है–

$$x^2 = \Sigma \left\{ \frac{(f - fx)^2}{fx} \right\} \quad \text{or}$$

$$= \Sigma \left[\frac{(\text{Difference of actual and expected frequencies})^2}{\text{Expected frequencies}} \right]$$

where, f = actual frequency

fx = expected frequency

x^2 = Chi–Square or Square contingency

काई-वर्ग निकालने की रीति– (1) सर्वप्रथम आकस्मिकता सारणी तैयार कीजिये।

(2) प्रत्येक स्तम्भ (Cell or Column) की सम्भावित आवृत्ति इस सूत्र की सहायता से निकालिये—

$$A_1 B_1 = \frac{(A_1) \times (B_1)}{N} \qquad A_2 B_2 = \frac{(A_2) \times (B_2)}{N}$$

और इसी प्रकार जितने स्तम्भ हों उनकी सम्भावित आवृत्तियाँ निकालिये।

(3) वास्तविक आवृत्ति में से सम्भावित आवृत्ति घटाकर उसका वर्ग निकालिये।

(4) इस वर्ग में सम्भावित आवृत्ति का भाग दीजिये ।

(5) प्राप्त भजनफलों का योग काई–वर्ग होगा ।

इस प्रकार काई–वर्ग आकस्मिकता (Mean Square Contingency) या फाई–वर्ग (Phi Square) कहलाता है । इसे सूत्र के रूप में इस प्रकार व्यक्त कर सकते हैं–

$$\text{Mean Square Contingency or} \quad \Phi = \frac{x^2}{N}$$

कार्ल पियर्सन ने माध्य वर्ग आकस्मिकता अथवा आकस्मिकता गुणक ज्ञात करने के लिए निम्न सूत्र का प्रयोग किया है –

Mean Square Contingency or Coefficient of Contingency

$$\text{or} \quad C = \sqrt{\left(\frac{x^2}{N + X^2}\right)} \quad \text{or} \quad \sqrt{\frac{\Phi^2}{1 + \Phi^2}} \; \Phi^2 = \frac{X^2}{N}$$

इस गुणक का एक दोष है कि यह कभी भी '1' की सीमा तक नहीं पहुँचता है ।

(Φ संकेत फाई–वर्ग के लिए प्रयोग किया जाता है ।)

Illustration 2 : Discuss the resemblance of height of parent and offspring from the following (निम्नलिखित से माता पिता एवं उनकी संतान की ऊंचाई में समानता की व्याख्या कीजिए) –

Offspring	**Parent**				**Total**
	Very Tall	**Tall**	**Medium**	**Short**	
Very Tall	20	30	20	2	72
Tall	14	125	85	12	236
Medium	3	140	165	125	433
Short	3	37	68	151	259
Total	**40**	**332**	**338**	**290**	**1000**

Solution :–

Actual Frequencies

Offspring	**Parent B**				**Total**
	Very Tall (B_1)	**Tall (B_2)**	**Medium (B_3)**	**Short (B_4)**	
Very Tall $(A_1B_1)(A_1)$	(A_1B_2)	(A_1B_3)20	(A_1B_4)30	20	(A_1) = 72
Tall $(A_2B_1)(A_2)$	(A_2B_2)	(A_2B_3)14	(A_2B_4)125	85	(A_2) = 236
Medium $(A_3B_1)(A_3)$	(A_3B_2)	(A_3B_3) 3	(A_3B_4)140	165	(A_3) = 1433
Short $(A_4B_1)(A_4)$	(A_4B_2)	(A_4B_3) 3	(A_4B_4)37	68	(A_4) = 1259
Total	(B_1) = 40	(B_2) = 332	(B_3) = 338	(B_4) = 290	N = 1000

Expected Frequencies (fx)

Offspring	**Parent**				**Total**
	Very Tall	**Tall**	**Medium**	**Short**	
Vary Tall	2.9	23.9	24.3	20.9	72
Tall	9.4	78.4	79.8	68.4	236
Medium	17.3	143.8	146.4	125.5	433
Short	10.4	85.9	87.5	75.5	259
Total	40	332	338	290	1000

Expected Frequencies will be calculated as follows (assuming that the two attributes are not associated, *i.e.*, they are independent) –

$$(A_1 B_1) = \frac{(A_1) \times (B_1)}{N} = \frac{72 \times 40}{1000} = 2.9$$

$$(A_1 B_2) = \frac{(A_1) \times (B_2)}{N} = \frac{72 \times 332}{1000} = 23.9$$

NOTES

$$(A_1 B_3) = \frac{(A_1) \times (B_3)}{N} = \frac{72 \times 338}{1000} = 24.3$$

$$(A_1 B_4) = \frac{(A_1) \times (B_4)}{N} = \frac{72 \times 290}{1000} = 20.9$$

$$(A_2 B_1) = \frac{(A_2) \times (B_1)}{N} = \frac{236 \times 40}{1000} = 29.4$$

$$(A_2 B_2) = \frac{(A_2) \times (B_2)}{N} = \frac{236 \times 332}{1000} = 78.4$$

$$(A_2 B_3) = \frac{(A_2) \times (B_3)}{N} = \frac{236 \times 338}{1000} = 79.8$$

$$(A_2 B_4) = \frac{(A_2) \times (B_4)}{N} = \frac{236 \times 290}{1000} = 68.4$$

$$(A_3 B_1) = \frac{(A_3) \times (B_1)}{N} = \frac{433 \times 40}{1000} = 17.3$$

$$(A_3 B_2) = \frac{(A_3) \times (B_2)}{N} = \frac{433 \times 332}{1000} = 143.8$$

$$(A_3 B_3) = \frac{(A_3) \times (B_3)}{N} = \frac{433 \times 338}{1000} = 146.4$$

$$(A_3 B_4) = \frac{(A_3) \times (B_4)}{N} = \frac{433 \times 290}{1000} = 125.5$$

$$(A_4 B_1) = \frac{(A_4) \times (B_1)}{N} = \frac{259 \times 40}{1000} = 10.4$$

$$(A_4 B_2) = \frac{(A_4) \times (B_2)}{N} = \frac{259 \times 332}{1000} = 85.9$$

$$(A_4 B_3) = \frac{(A_4) \times (B_3)}{N} = \frac{259 \times 338}{1000} = 87.5$$

$$(A_4 B_4) = \frac{(A_4) \times (B_4)}{N} = \frac{259 \times 290}{1000} = 75.2$$

	f	fx	(f − fx)	$(f - fx)^2$	$\frac{(f-fx)^2}{fx}$
(A_1B_1)	20	2.9	17.1	292.41	100.83
(A_1B_2)	30	23.9	6.1	37.21	1.55
(A_1B_3)	20	24.3	−4.3	18.49	0.76
(A_1B_4)	2	20.9	−18.9	375.21	17.09
(A_2B_1)	14	9.4	4.6	21.16	2.25
(A_2B_2)	125	78.4	46.6	2171.56	27.70
(A_2B_3)	85	79.8	5.2	27.04	0.34
(A_2B_4)	12	68.4	−56.4	3180.96	46.50
(A_3B_1)	3	17.3	−14.3	204.49	11.82
(A_3B_2)	140	143.8	−3.8	14.44	0.10
(A_3B_3)	165	146.4	18.6	345.96	2.36
(A_3B_4)	125	125.5	−0.5	00.25	0.00
(A_4B_1)	3	10.4	−7.4	54.76	5.26
(A_4B_2)	37	85.9	−48.9	2391.21	27.83
(A_4B_3)	68	87.5	−19.5	380.25	4.34
(A_4B_4)	151	75.2	75.8	5745.64	76.40
					Σ or X^2 = 325.132

$$C=\sqrt{\left(\frac{X^2}{N+X^2}\right)} = \sqrt{\left(\frac{325.132}{100+325.132}\right)} = .495$$

This indicates that the association between the height of parent and the offspring is significant. From the inspection of the table, the contingency is positive which means that there is resemblance between the height of the two.

5.3 काई-वर्ग और स्वतन्त्रता की मात्रा (Chi-Square and Degree of Freedom)

काई-वर्ग के आधार पर दिये हुए दो गुणों में सम्बन्ध की जाँच की जा सकती है । इसके लिए विशेष प्रकार की सारणियाँ बनी होती हैं जिनमें सामान्यतः 1% से 99% तक के विभिन्न स्तरों के लिए काई-वर्ग (X^2) के मूल्य दिये होते हैं । इस तरह सबसे पूर्व हम स्वतन्त्रता की जाँच (Test of freedom) करते हैं । इसके लिए निम्न सूत्र का प्रयोग किया जाता है-

Degree of Freedom = (R–1) (C–1) where, R = No. of Rows

C = No. of Columns.

इस सूत्र से स्वतंत्रता की मात्रा मालूम हो जाती है। इस मात्रा के दिये हुए प्रतिशत के स्तर पर सारणी की सहायता से 'काई-वर्ग' प्राप्त कर लिया जाता है । अब सूत्र द्वारा निकाले गये काई-वर्ग और सारणी द्वारा प्राप्त काई-वर्ग में तुलना की जाती है । यदि वास्तविक निकाला गया काई-वर्ग सारणी द्वारा प्राप्त काई-वर्ग से अधिक है तो दोनों गुणों में अर्थपूर्ण सम्बन्ध (Significant Association) माना जायेगा । यदि यह कम है तो अर्थहीन (Insignificant) सम्बन्ध माना जायेगा । यदि दोनों मूल्य लगभग समान हों तो दोनों का अन्तर निदर्शन-उच्चावचन के कारण माना जायेगा ।

Illustration 3– Two samples of polls of Votes for two candidates A and B for a public office are taken, one from among residents of urban areas, and the other from residents of rural areas. the results are given below. Examine whether the nature of the area is related to voting preference in this election.

Votes for/Area	A	B	Total
Rural	620	380	1000
Urban	550	450	1000
Total	1170	830	2000

Solution– Setting up the hypothesis that the nature of area is independent of voting preference, the expected frequencies would be —

Votes for/Area	A	B	Total
Rural	$\frac{1000 \times 1170}{2000} = 585$	$\frac{1000 \times 830}{2000} = 415$	1000
Urban	$\frac{1000 \times 1170}{2000} = 585$	$\frac{1000 \times 830}{2000} = 415$	1000
Total	**1170**	**830**	**2000**

$$\therefore \quad X^2 = \frac{(620-585)^2}{585} + \frac{(380-415)^2}{415} + \frac{(550-585)^2}{585} + \frac{(450-415)^2}{415}$$

$$= \frac{(35)^2}{585} + \frac{(-35)^2}{415} + \frac{(-35)^2}{585} + \frac{(35)^2}{415} = 2.1 + 2.9 + 2.1 + 2.9$$

$= 10$ Approx.

The number of degrees of freedom $= (2-1)(2-1) = 1$

NOTES

The calculated value of X^2 which is 10 is much greater than the table value of X^2 for 1 d.f. at 5% level of significance. Therefore our hypothesis is wrong and the nature of area is related to voting preference.

Illustration 4. From the following table, showing the number of plants having certain characters, test the hypothesis that the flower colour is independent of flatness of leaf–

	Flat Leaves	Cur led Leaves	Total
White Flowers	99	36	135
Red Flowers	20	5	25
Total	**119**	**41**	**160**

You may use the following table giving the value of X^2 for one degree of freedom for different values of P :-

P	.99	.96	.90	.50	.10	.05	.01
X^2	.000157	.09393	.0158	.455	2.706	3.841	6.635

Solution . Let the flower colour and flatness of leaf are not associated. Expected frequency for white flowers and flat leaves

$$= \frac{135}{160} \times 119 = 100.4$$

In this way, the following expected frequency table can be obtained —

	Flat Leaves	Curled Leaves	Total
White flowers	100.4	34.6	145
Red flowers	18.6	6.4	25
Total	**119**	**41**	**160**

$$= \frac{(99-100.4)^2}{100.4} + \frac{(36-34.6)^2}{34.6} + \frac{(20-18.6)^2}{18.6} + \frac{(5-6.4)^2}{6.4}$$

$$= \frac{(-1.4)^2}{100.4} + \frac{(1.4)^2}{34.6} + \frac{(1.4)^2}{18.6} + \frac{(-1.4)^2}{6.4}$$

$$= \frac{1.96}{100.4} + \frac{1.96}{34.6} + \frac{1.96}{18.6} + \frac{1.96}{6.4} \quad = .0195 + .0566 + .1053 + .3062$$

$$= .4876$$

The number of degrees of freedom

$= (2-1)\ (2-1) = 1$

The value for degree of freedom at 5% level of significance is 3.841 which is greater than calculated value X^2. Thus, the hypothesis is correct and the flower of colour is independent of the leaf.

शुपरो का गुणक (Tschuprow's Coefficient) — कार्ल पियर्सन के आकस्मिकता गुणक का दोष यह है कि यह अधिकतम सीमा '1' तक नहीं पहुँचता । शूपरो ने निम्न गुणक का सुझाव दिया है–

$$T = \sqrt{\frac{C^2}{(1-c^2)\sqrt{(r-1)(c-1)}}}$$

यहाँ T = Tschuprow's Coefficient (शुपरो का गुणक)

C = Coefficient of Contingency (आसंग गुणक)

r = No. of rows (कतारों की संख्या)

c = No. of columns (खानों की संख्या)

5.4 येट का संशोधन (Yate's Correction)

काई–वर्ग बंटन एक सतत् बंटन है जिसके प्रयोग की एक आवश्यक शर्त यह है कि कोई भी कोष्ठ–आवृत्ति (Cell frequency) अर्थात् अन्तिम आवृत्ति (ultimate class frequency) 5 से कम नहीं होनी चाहिए अन्यथा काई–वर्ग (X^2) का मान अनावश्यक रूप से बढ़ा रहेगा और निष्कर्ष भी भ्रमपूर्ण होंगे। अतः काई–वर्ग बंटन की सततता (अविच्छिन्नता) बनाये रखने तथा सही परिणाम निकालने के लिए 2×2 –सारणी में अविच्छिन्नता के लिए येट का संशोधन (Yate's Correction for continuity) करना आवश्यक है।

यह संशोधन तब किया जाता है जबकि किसी कोष्ठ की आवृत्ति 5 से कम हो। ऐसी स्थिति में सारणी में दी हुई सबसे छोटी आवृत्ति में 1/2 या 0.5 जोड़ दिया जाता है और शेष 3 आवृत्तियों को इस प्रकार समायोजित किया जाता है कि सीमांत जोड़ (marginal total) पूर्ववत् ही रहे। इस संशोधन का प्रभाव यह होता है कि प्रत्येक वास्तविक और अनुमानित आवृत्ति का अन्तर 0.5 कम हो जाता है जिससे काई–वर्ग का मूल्य भी वास्तविकता के अधिक निकट हो जाता है और निष्कर्ष भी ठीक निकलते हैं।

5.5 आवृत्तियों का समूहन (Pooling of individual frequencies)

जब अनुमानित आवृत्तियाँ कम होती हैं (विशेष रूप से 5 से कम), तो वास्तविक और प्रत्याशित आवृत्तियों को अन्तर मालूम करने से पहले दो या दो से अधिक ऐसी आवृत्तियों को जोड़ लिया जाता है। स्वातन्त्र्यांशों की संख्या (No. of degrees of freedom) का निर्धारण 5 से कम वाली निकटतम आवृत्तियों को जोड़ने के बाद प्राप्त वर्गों की संख्या के आधार पर किया जाता है। उदाहरण के लिए, 8 वर्गों में से 3 की सैद्धान्तिक आवृत्तियाँ कम (<s) हों जिन्हें एक वर्ग में समूहित कर दिया जाये तो कुल पुनः व्यवस्थित वर्गों की संख्या 6 हो जायेगी। ऐसी स्थिति में स्वातन्त्र्य संख्या 8—1 = 7 न होकर 6 – 1 = 5 होगी।

Illustration 5 – एक भयंकर रोग से बकरियों के प्रतिरक्षण (Immunisation) से सम्बन्धित प्रयोगों से निम्नलिखित परिणाम प्राप्त हुए —

	रोग से मरी (b)	जीवित (B)	कुल (N)
टीका लगा (A)	2	10	12
टीका नहीं लगा (a)	6	6	12
कुल (N)	8	16	24

(i) येट संशोधन के बिना तथा (ii) येट संशोधन सहित X^2 (काई–वर्ग) की गणना कीजिए और निष्कर्ष की व्याख्या कीजिए।

Solution –

(i) Calculation of X^2 without Yate's correction–

Expected frequencies (fx) =

$$(Ab) = \frac{(A)\times(b)}{N} = \frac{12\times 8}{24} = 4 \qquad (AB) = \frac{(A)\times(B)}{N} = \frac{12\times 16}{24} = 8$$

$$(ab) = \frac{(a)\times(b)}{N} = \frac{12\times 8}{24} = 4 \qquad (aB) = \frac{(a)\times(B)}{N} = \frac{12\times 16}{24} = 8$$

Now,

NOTES

f	fx	(f–fx)	$(f–fx)^2$	$\frac{(f-fx)^2}{fx}$
2	4	–2	4	1.0
10	8	+2	4	0.5
6	4	+2	4	1.0
6	8	–2	4	0.5
				X^2 = 3.0

The X^2 arrived at is 3 which is less than the Table–Value 1 df (3.841). Therefore, it is true that innoculation and death are independent.

ii. Calculation of X^2 with Yate's correction–

Expected Frequencies		
4	8	12
4	8	12
8	10	24

Actual Frequencies		
2.5	9.5	12
5.5	6.5	12
8	16	24

Now,

f	fx	(f–fx) – 0.5	$[(f–fx)–0.5]^2$	$\frac{[(f-fx)-0.5]^2}{fx}$
2	4	(2–4)–0.5 = –1.5	2.25	2.25 ÷ 4 = 0.5625
10	8	(10–8)–0.5 = +1.5	2.25	2.25 ÷ 8 = 0.2812
6	4	(6–4)–0.5 = +1.5	2.25	2.25 ÷ 4 = 0.5625
6	8	(6–8)–0.5 = –1.5	2.25	2.25 ÷ 8 = 0.2813
				X^2 = 1.6875

Here also, the calculated value of X^2 is 1.6875 which is less than the table–value; hence the result remains unchanged.

5.6 अन्वायोजन की उत्तमता की जाँच (Test of Goodness of fit)

काई–वर्ग जाँच का उपयोग अन्वायोजन की उत्तमता की जाँच के लिए किया जाता है । इससे हमें सिद्धान्त (Theory or expectation) तथा तथ्य (Fact or Observation) के अन्तर की सार्थकता (significance) का ज्ञान होता है । वास्तव में, अन्वायोजन–उत्तमता जाँच सैद्धान्तिक और न्यादर्श बंटन की एकरूपता या संगति की जाँच है । यदि X^2 निकाला गया मूल्य X^2 –सारणी में दिये गए मूल्य से कम होता है तो अन्वायोजन उत्तम समझा जाता है अर्थात् वास्तविक और अनुमानित आवृत्तियों के वक्र लगभग एक दूसरे के अनुरूप हैं ; उनका अन्तर केवल दैव के कारण है, अत: वह अर्थहीन (insignificant) है । इसके विपरीत, यदि X^2 का निकाला गया मूल्य उसके सारणी–मूल्य से अधिक है तो अन्वायोजन उत्तम नहीं कहा जावेगा । वास्तविक और अनुमानित आवृत्तियों के वक्रों में काफी दूरी है अर्थात् उनका अन्तर अर्थपूर्ण (significant) है । यह दैव के कारण नहीं है ।

Illustration 6. ग्रेगोर मेण्डल ने मटर–प्रजनन (Pea– Breeding) पर किये गए प्रयोगों से बीजों की निम्नलिखित आवृत्तियाँ प्राप्त कीं —

गोल व पीले (Round and yellow) = 315
झुर्रीदार व पीले (Wrinkled and yellow) = 101
गोल व हरे (Round and green) = 108

झुर्रीदार व हरे (Wrinkled and green) = 32

कुल (Total) = 556

सिद्धान्त के अनुसार आवृत्तियाँ 9 : 3 : 3 : 1 के अनुपात में होना चाहिए।

सिद्धान्त और प्रयोग के सामंजस्य की जाँच कीजिए।

Solution- Calculation of X^2

Seed mixture	**Actual freq. (f)**	**fx**	**f–fx**	**(f–fx)^2**	$\frac{(f-fx)^2}{fx}$
Round and yellow	315	$\frac{556 \times 9}{16} = 313$	+2	4	0128
Wrinkled and yellow	101	$\frac{556 \times 3}{16} = 104$	–3	9	.0865
Round and green	108	$\frac{556 \times 3}{16} = 104$	+4	16	.1538
Wrinkled and green,	32	$\frac{556 \times 1}{16} = 35$	–3	9	.2571
Total	**556**	**556**	**0**	**38**	**X^2 = .5102**

d.f. (derived frequencies) = (n–1) =(4–1) =3

The value of X^2 in the Table at 5% significant level for 3 independent frequencies is 7.815 . The calculated value of X^2 is 0.5102 which is very less than its Table– value. Hence, the fit is good or there is good is coordination between theory and experiment.

5.7 काई–वर्ग के विशेष गुण (Special Properties of X^2)

(1) **संयोगात्मक या संचयात्मक गुण** (Additive property) — यदि किसी समग्र से कई यादृच्छिद न्यादर्श का चुनाव कर उनका अध्ययन किया जाये तो न्यादर्शों के अलग– अलग काई– वर्गों के मान जोड़कर पूरे समग्र के विषय में अधिक विश्वसनीय निष्कर्ष निकाले जा सकते हैं।

(2) **X^2 बंटन का स्वरूप एवं स्वातन्त्र्यांश** (Nature and degree of freedom of X^2–distribution)– X^2 बंटन का स्वरूप स्वातन्त्र्य संख्याओं पर निर्भर होता है। प्रत्येक स्वातंत्र्य संख्या के लिए एक काई–वर्ग वक्र बनता है। बहुत कम स्वातन्त्र्य संख्याओं के लिए काई– वर्ग बंटन धनात्मक विषमता वाले दाहिनी ओर को असंमित वक्र के रूप में होता है। स्वतन्त्रता की संख्या बढ़ने के साथ-साथ, वक्र की असंमिति कम होती जाती है और वक्र संमिति की ओर बढ़ती जाती है। जब स्वातन्त्र्य संख्या 30 से अधिक हो जाती है तो $\sqrt{2X^2}$ का बंटन प्रसामान्य बंटन (Normal distribution) के अनुरूप हो जाता है। $\sqrt{2X}$ के बंटन का माध्य $\sqrt{2d.f. - 1}$ और प्रमाप-विचलन 1 होता है। यह गुण बहुत उपयोगी बनता है क्योंकि 30 से अधिक d.f. वाले x^2 का निर्वचन करने के लिए प्रसामान्य बंटन के क्षेत्रफल सारणी का प्रयोग किया जा सकता है।

काई–वर्ग जाँच के उपयोग (Uses of X^2 –test)

काई– वर्ग जाँच का उपयोग आधुनिक सांख्यिकी में बहुत अधिक होता है। यह शोधकार्य के लिए एक महत्वपूर्ण उपकरण है। इसका उपयोग निम्नलिखित जाँचों में किया जाता है —

(1) **स्वतन्त्रता की जाँच** (Test of Independence)– इसके द्वारा दो गुणों में गुण सम्बन्ध की जाँच की जाती है, जैसे इस तकनीक द्वारा यह जाँच की जा सकती है कि टीका चेचक से बचने के लिए प्रभावोत्पादक है या नहीं; साक्षरता और रोजगार में कोई सम्बन्ध है या नहीं; सिगरेट पीने तथा केन्सर रोग में होने में कोई गुण सम्बन्ध है या वे स्वतन्त्र हैं; पिता और पुत्रों की बुद्धिमता में कोई गुण सम्बन्ध है या नहीं, आदि। इस जाँच के लिए पहले दोनों गुण स्वतन्त्र मान लिए जाते हैं, फिर अनुमानित आवृत्तियाँ निकाली जाती हैं जिनका वास्तविक आवृत्तियों से अन्तर मालूम करके काई वर्ग के माप की गणना की जाती है। अन्त में, एक निश्चित सार्थकता स्तर (5%) पर सम्बन्धित

NOTES

स्वातन्त्र्य संख्या के अनुरूप X^2 का सारणी–मूल्य देख लिया जाता है । यदि X^2 का परिगणित मूल्य सारणी मूल्य से अधिक है तो दोनों गुण स्वतन्त्र नहीं होते वरन् उनमें गुण सम्बन्ध होता है ।

(2) **अन्वायोजन–उत्तमता की जाँच** (Test of Goodness of Fit)– X^2 का प्रयोग सैद्धान्तिक आवृत्ति–बंटन (द्विपद, प्वॉयसन या प्रसामान्य) और वास्तविक बंटन में अन्तर की जाँच करने के लिए किया जाता है । इस जाँच से यह पता चलता है कि वास्तविक आवृत्ति बंटन कहाँ तक सैद्धान्तिक आवृत्ति बंटन के अनुरूप है या दोनों में अन्तर सार्थक अथवा अर्थहीन है । यदि X^2 का परिगणित मूल्य उसके सारणी मूल्य से अधिक होता है तो अन्वायोजन उत्तम नहीं होता । यदि यह कम होता है तो अनुमानित और वास्तविक आवृत्तियों का अन्तर अर्थहीन होता है अर्थात् वह दैव के कारण होता है; अन्य किसी कारण से नहीं ।

(3) **सजातीयता की जाँच** (Test of Homegeneity)– यह स्वातन्त्र्य–जाँच का ही विस्तृत स्वरूप है । काई वर्ग के प्रयोग द्वारा इस तथ्य की भी जाँच की जाती है विभिन्न न्यादर्श एक ही समग्र से लिये गये हैं या नहीं ।

काई–वर्ग के प्रयोग की मुख्य शर्तें (Main conditions of the application of X^2–test) :

(1) समग्र की इकाइयों की संख्या यथोचित रूप से अधिक होनी चाहिए अन्यथा वास्तविक व अनुमानित आवृत्तियों के अन्तरों का बंटन प्रसामान्य नहीं होगा । इकाइयों की संख्या, व्यवहार में 50 से अधिक ही होना चाहिए ।

(2) कोई भी अनुमानित कोष्ठ–आवृत्ति 5 से कम नहीं होना चाहिए । यदि कोई आवृत्ति 5 से कम है तो उसे निकटवर्ती आवृत्तियों के साथ मिलकर X^2 निकालना चाहिए ।

(3) न्यादर्श दैव आधार पर चुना हुआ होना चाहिए ।

(4) कोष्ठ आवृत्तियों के अवरोध रेखीय (Linear) होने चाहिए ।

5.8 काई वर्ग तथा महत्ता का स्तर (Chi-Square and Standard of Significance)

सूत्र की सहायता से ज्ञात किए गए काई-वर्ग की उचित स्वतंत्रता अंश के एक निश्चित महत्ता-स्तर से तुलना करते हैं। इस प्रकार काई-वर्ग तथा सारणी से प्राप्त काई-वर्ग में तुलना की जाती है। यदि वास्तविक काई वर्ग सारणी द्वारा प्राप्त काई-वर्ग से अधिक है तो दोनों गुणों में महत्वपूर्ण गुण-सम्बन्ध माना जाता है और हमारी शून्य-मान्यता गलत हो जाती है। इसके विपरीत ज्ञात किया गया काई-वर्ग तालिका से ज्ञात मूल्य से कम है तो उसे महत्वहीन (insignificant) मानते हैं, अर्थात् शून्य मान्यता सही मानी जाती है और उनमें कोई गुण-सम्बन्ध नहीं होता। यदि दोनों मूल्य समान हों तो यह माना जाएगा कि दोनों का अन्तर निदर्शन उच्चावचन के कारण है। सामान्यतया 5% तक महत्ता स्तर को विचार में लिया जाता है।

Illustration 7. निम्न सारणी से जिसमें प्लांट का कुछ विशेषता के आधार पर संख्या दी गयी है, मान्यता के आधार पर परीक्षण कीजिए, कि फूल का रंग पत्ती के समतल से स्वतंत्र है-

	Flat leaves	Curled leaves	Total
White Flowers	99	36	185
Red Flowers	20	5	25
Total	**119**	**41**	**160**

You may use the following table giving the value of x^2 for one degree of freedom for different values of.

p :	.99	.96	.90	.50	.10	.05	.01
x^2:	.000157	.00393	.0158	.455	2.706	3.841	6.635

हल- माना Flower colour and flatness of leaf में गुण-सम्बन्ध नहीं है-

White flowers and flat leaves में सम्भावित आवृत्ति-

$$= \frac{135}{160} \times 119 = 100.4$$

इसी प्रकार अन्य सम्भावित आवृत्तियाँ निम्न प्रकार हैं-

	Flat leaves	Curled leaves	Total
White Flowers	100.4	34.6	135
Red Flowers	18.6	6.4	25
Total	**119**	**41**	**160**

$$x^2 = \frac{(99-100.4)^2}{100.4} + \frac{(36-34.6)^2}{34.6} + \frac{(20-18.6)^2}{18.6} + \frac{(5-6.4)^2}{6.4}$$

$$= \frac{(-1.4)^2}{100.4} + \frac{(1.4)^2}{36.4} + \frac{(1.4)^2}{18.6} + \frac{(-1.4)^2}{6.4}$$

or $\frac{1.96}{100.4} + \frac{1.96}{36.4} + \frac{1.96}{18.6} + \frac{1.96}{6.4}$

or .019 + .0566 + .1053 + .3062 or .4876

∴ The number of degree of freedom-

$$= (2-1)\ (2-1) = 1$$

अत: 5% महत्ता स्तर पर 1 अंश स्वतंत्रता का मूल्य 3841 है जो गणना किए गए x^2 मान से अधिक है। अत: विचार कल्पना सही है और flower colour, leaf से स्वतंत्र है।

Illustration 8. मतदान से दो प्रतिदर्श दो व्यक्तियों अ एवं ब का लिया गया है- उनमें से एक शहरी क्षेत्र एवं दूसरा ग्रामीण क्षेत्र से है, जिनका परिणाम निम्न प्रकार दिया है। परीक्षण कीजिए कि क्षेत्र की प्रकृति मतदान के प्रायिकता से सम्बन्धित है-

Area \ Votes for	A	B	Total
Rural	620	380	1000
Urban	550	450	1000
Total	**1170**	**830**	**2000**

हल– माना कि nature of area, अपने voting preference से स्वतंत्र है। The expected frequencies are –

Area \ Vote for	A	B	Total
Rural	$\frac{1000\times1170}{2000} = 585$	$\frac{1000\times830}{2000} = 415$	1000
Urban	$\frac{1000\times1170}{2000} = 585$	$\frac{1000\times830}{2000} = 415$	1000
Total	**1170**	**830**	**2000**

NOTES

$$x^2 = \frac{(620-585)^2}{585} + \frac{(380-415)^2}{415} + \frac{(550-585)^2}{585} + \frac{(450-415)^2}{415}$$

$$= \frac{(35)^2}{585} + \frac{(-35)^2}{415} + \frac{(-35)^2}{585} + \frac{(35)^2}{415}$$

$$= 2.1 + 2.9 + 2.1 + 2.9 = 10$$

The number of degree of freedom

$= (2 - 1)\ (2 - 1) = 1$

काई वर्ग 10 है जो कि सारणी के काई वर्ग 1 d.f. से 5% महत्ता स्तर से बहुत अधिक है, अत: हमारी कल्पना गलत है और nature of area का Voting preference से सम्बन्ध है।

Illustration 9. निम्न समंकों से पुरुष व स्त्री के मध्य साक्षरता एवं बेरोजगारी के मध्य गुण-सम्बन्ध ज्ञात कीजिए-

	Males	Females
Total	250	200
Literate	100	40
Unemployed	50	12
Literate and Unemployed	30	4

[The 5% value of Chi square for one degree of freedom = 3.841]

हल- माना A = Literacy

a = Not literate

B = Unemployed

b = Not unemployed

Given date -	Males	Females
	N = 250	N = 200
	(A) = 100	(A) = 40
	(B) = 50	(B) = 12
	(AB) = 30	(AB) = 4

	A	a	
A	(AB) 30	(aB) 20	(B) 50
b	(Ab) 70	(ab) 130	(b) 200
	(A) 100	(a) 150	(N) 250

	A	a	
B	(AB) 4	(aB) 8	(B) 12
b	(Ab) 36	(ab) 152	(b) 188
	(A) 40	(a) 160	(N) 200

Males : Expected value of (AB) $\frac{(A)\times(B)}{N} = \frac{100\times 50}{250} = 20$

Females : Expected value of (AB) $= \frac{(A)\times(B)}{N} = \frac{40\times 12}{200} = 2.4$

Other frequencies will be -

20	30	50
80	120	200
100	150	250

2.4	9.6	12
37.6	150.4	188
40	160	200

NOTES

Calculation of x^2

Males						Females					
	F	F_1	(F − F₁)	$(F-F_1)^2$	$\frac{(F-F_1)^2}{F_1}$		F	F_1	(F − F₁)	$(F-F_1)^2$	$\frac{(F-F_1)^2}{F_1}$
AB	30	20	10	100	5.00	AB	4	2.4	1.6	2.56	1.07
Ab	70	80	-10	100	1.25	Ab	36	37.6	-1.6	2.56	.07
aB	20	30	-10	100	3.33	aB	8	9.6	-1.6	2.56	.27
ab	130	120	10	100	0.83	ab	152	150.4	1.6	2.56	.02
					x^2 = 10.41						x^2 = 1.43

Association of Literacy and unemployment is more significant in Males.

महत्वपूर्ण सूत्र

आसंग

(1) काई वर्ग $x^2 = \Sigma \left[\frac{(f_0 - f_e)^2}{f_e}\right]$ Ùee $\Sigma \left[\frac{(0 - E)^2}{E}\right]$

(2) आसंग गुणांक $C = \sqrt{\frac{x^2}{N + x^2}}$

काई-वर्ग जाँच

(1) **स्वातन्त्रय जाँच: विधि-**

(i) H_n = शून्य परिकल्पना (ii) x^2 का मान

(iii) स्वातन्त्रय संख्या d.f. = (C - 1) (r - 1)

(iv) 5% सार्थकता स्तर पर d.f. सारणी मूल्य

(v) परिकलन x^2 (ii) > सारणी x^2 H_0 असत्य A और B में साहचर्य

(2) **आसंजन-उत्कृष्टता: विधि की जाँच-**

(i) $H_0 - f_0$ और fe में शून्य अन्तर है।

(ii) Z_2 का परिकलित मूल्य

(iii) स्वातन्त्रय संख्या: d.f. = (n - k) (n - 1)

(iv) 5% सार्थकता स्तर पर d.f. से मूल्य

(v) परिकलित मूल्य : > सारणी मूल्य अन्वायोजन उत्कृष्ट नहीं।

बोध प्रश्न

1. काई वर्ग परीक्षण को समझाइये।

....

....

....

2. काई वर्ग परीक्षण की प्रक्रिया को लिखिए।

....

NOTES

....

....

3. 'येट का संशोधन' पर टिप्पणी कीजिए।

....

....

....

5.10 शब्दावली या शब्दकुंजी

5.10 अभ्यास प्रश्न

दीर्घ उत्तरीय प्रश्न (Long Answer Type Questions)

1. काई वर्ग जाँच के विभिन्न उपयोगों का वर्णन कीजिये।
   Describe the various uses of the chi square test.
2. काई वर्ग बंटन के विशेष गुणों की उदाहरण सहित व्याख्या कीजिये। काई वर्ग जाँच के प्रयोग की आवश्यक शर्तों का वर्णन कीजिये।
   Explain with illustrations the special properties of chi square distribution. Also state the necessary conditions for applying the chi square test.
3. काई वर्ग परीक्षण क्या है ? सैद्धान्तिक और अवलोकित आवृत्तियों के अन्तर के महत्व को ज्ञात करने में यह कैसे सहायक होता है ?
   What is chi square test? How does it help in finding out the significance of difference between theory and observations?
4. X^2 परीक्षण क्या है ? इसकी सीमाओं का वर्णन कीजिये।
   What is X^2 test? State the limitations of chi square test.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. काई वर्ग परीक्षण क्या है ?
   What is Chi Square test.
2. काई वर्ग परीक्षण की सीमाओं का वर्णन कीजिये।
   Explain the limitations of Chi Square test.
3. काई वर्ग परीक्षण की प्रक्रिया क्या है ?
   What is process of Chi Square test.?
4. आकस्मिकता गुणांक क्या है ?
   What is coefficient of contingency?
5. आकस्मिकता का अर्थ स्पष्ट कीजिये।
   Explain the meaning of contingency.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. येट का संशोधन सामान्यत: स्वातन्त्रय कोटि होने पर किया जाता है-
   (अ) 2, (ब) 1, (स) 3, (द) 4।

2. आसंग तालिका की एक 5 × 4 में स्वातन्त्र कोटियाँ है-

(अ) 25, (ब) 16, (स) 12, (द) 15 ।

3. काई वर्ग परीक्षण का चिन्ह प्रयोग किया जाता है-

(अ) X^2, (ब) T, (स) F, (द) कोई नहीं ।

4. काई वर्ग के विशेष गुण होते हैं-

(अ) सयोगात्मक (ब) स्वातन्त्रयाश (स) संचयात्मक (द) उपरोक्त सभी ।

[**उत्तर**- 1. (ब), 2. (स), 3. (अ), 4. (द)]

5.11 व्यावहारिक प्रश्न-

1. निम्न सारणी 1000 अपराधियों में उनके भार और मनोवृत्ति के मध्य सम्बन्ध प्रदर्शित करती है। दोनों गुणों में आसंग-गुणांक परिकलित कीजिए-

	भार (पौण्ड में)					
मनोवृत्ति	90-120	120-130	130-140	140-150	150 से अधिक	योग
सामान्य	50	102	198	210	240	800
दुर्बल	30	38	72	30	30	200
योग	80	140	270	240	270	1000

[C = .216, y_2 = 49.05]

2. निम्न सारणी से आसंग गुणांक परिकलित कीजिए और अपने निष्कर्ष निकालिए। उक्त सारणी दो स्वतन्त्र अनुसन्धानकर्त्ताओं द्वारा किए गए ऐसे अनुसंधान से सम्बन्धित है, जिसमें एक ही बस्ती में रहने वाले व्यक्तियों को विभिन्न आय-वर्गों में प्रस्तुत किया गया है-

		आय-वर्ग	
अनुसन्धानकर्ता-	दरिद्र	मध्य-वर्ग	सम्पन्न
अ	140	100	15
ब	140	50	20

[C = .165]

3. निम्नलिखित सारणी चेचक के प्रारम्भ होने पर प्राप्त सामग्री दिखाती है-

	आक्रान्त हुए	आक्रान्त नहीं हुए
टीका लगे हुए	31	469
टीका नहीं लगे हुए	185	1315

चेचक को रोकने में टीके के प्रभाव का परीक्षण कीजिए। आप अपने परिणाम का 5% सार्थकता स्तर पर x^2 के द्वारा परीक्षण कीजिए। 5% सार्थकता स्तर पर स्वातन्त्र संख्या (d.f.) 1 के लए x^2 का मूल्य 3.841 है।

[x^2 = 14.65 टीका प्रभावशाली है]

4. पशुओं की क्षय रोग से मुक्ति से सम्बन्ध एक प्रयोग में निम्न परिणाम प्राप्त हुए-

	मृत या प्रभावित	अप्रभावित	योग
टीका लगा	12	24	36
टीका नहीं लगा	16	8	24
योग	28	32	60

क्षय रोग की ग्रहणशीलता पर नियंत्रण करने में टीके के प्रभाव का परीक्षण कीजिए।

[x^2 = 6.43 टीका प्रभावशाली है]

NOTES

5. (i) क्षय रोग से पशुओं के प्रतिरक्षण के एक प्रयोग से निम्न परिणाम प्राप्त हुए-

	रोग ग्रस्त	निरोग
टीका लगे हुए	12	28
टीका न लगे हुए	13	7

बीमारी को नियंत्रित करने में टीके के प्रभाव की परीक्षा कीजिए।

(ii) एक नगर में खाद्य सामग्री सर्वेक्षण में निम्न परिणाम प्राप्त हुए-

	हिन्दू	मुस्लिम
चाय पीने वाले परिवार	1236	164
चाय न पीने वाले परिवार	564	36

इस तथ्य का परीक्षण कीजिए कि क्या चाय पीने की आदत के सम्बन्ध में दोनों धर्मावलम्बियों में कोई सार्थक अन्तर है ?

[(i) $x^2 = 7.8$ टीका प्रभावशाली है। (ii) $x^2 = 15.2$ सार्थक अन्तर है।]

6. प्रतिदर्श अध्ययन के आधार पर दो अन्वेषकों ने कुछ व्यक्तियों को आय के अनुसार वर्गीकृत किया। परिणाम निम्न प्रकार रहा-

अन्वेषक	दरिद्र	मध्यवर्गीय	समृद्ध	योग
अ	160	30	10	200
ब	140	120	40	300
योग	300	150	50	500

सिद्ध कीजिए कि कम से कम एक अन्वेषक द्वारा प्रयुक्त प्रतिचयन प्रविधि दोषपूर्ण है। (2df के लिए x^2 0.5 = 5.991; $x^2_{03} = 9.21$)

[$x^2 = 55.54$ दोषपूर्ण]

7. सारणियों के एक समूह में से 200 यादृच्छिक रूप से चुने गए। उनकी आवृत्तियाँ निम्नांकित हैं-

अंक :	0	1	2	3	4	5	6	7	8	9	योग
आवृत्ति :	18	19	23	21	16	25	22	20	21	15	200

x^2 के प्रयोग द्वारा इस परिकल्पना की यथार्थता का मूल्यांकन कीजिए कि उन सारणियों में अंक समान रूप से वितरित थे।

[$x^2 = 4.3 < 1692$ यथार्थ]

8. एक टेलीफोन निदर्शिका से यादृच्छिया चुनी गई संख्याओं में अंकों का वितरण निम्न प्रकार है-

अंक :	0	1	2	3	4	5	6	7	8	9	योग
आवृत्ति :	1026	1107	997	966	1075	933	1107	972	964	853	10000

इस तथ्य की जाँच कीजिए कि क्या निर्देशिका में उपयुक्त अंक बिल्कुल समान आवृत्ति से पाए जाते हैं।

[$x^2 = 58.5 > 16.92$ नहीं]

9. एक प्रजनन सिद्धान्त के अनुसार जिन बच्चों के माता-पिता में से एक M रक्त वर्ग का और दूसरा N रक्त वर्ग का होगा, वे बच्चे सदा तीर रक्त-वर्गों M, MN व N में से एक वर्ग के होंगे तथा इन वर्गों में बच्चों की औसत संख्या क्रमश: 1:2:1 के सम्भावित अनुपात में होगी। इस प्रयोग की रिपोर्ट निम्न है-

162 ऐसे बच्चों में से जिनके माता-पिता में से M वर्ग वाला तथा दूसरा N वर्ग वाला है। 28.4% M वर्ग के थे, 42% MN वर्ग तथा शेष N वर्ग के थे।

क्या रिपोर्ट में प्रस्तुत तथ्य प्रत्याशित प्रजनन अनुपात 1:2:1 के अनुरूप हैं।

[$x^2 = 4.22$ हाँ अनुरूप है]

10. चार बड़े नगरों के प्रौढ़ पुरुषों में से निम्न आकार के यादृच्छिक प्रतिदर्श चुने गए और विवाहित तथा अविवाहित व्यक्तियों की संख्याएँ अंकित कर ली गईं। इन आँकड़ों से क्या पुरुषों की विवाह करने की प्रवृत्ति में इन नगरों के बीच कोई विशेष विभिन्नता दृष्टिगोचर होती है ?

नगर	अ	ब	स	द	योग
विवाहित	137	164	152	147	600
अविवाहित	32	57	56	35	180
योग	169	221	208	182	780

(स्वातन्त्र्य संख्या 3 के लिए x^2 का 5% का मान 7.815 है)

[x^2 = 5.8 कोई विभिन्नता नहीं]

11. निम्न सारणी उन बाजारों में मूल्य वृद्धि व मूल्य ह्रास का विवरण प्रस्तुत करती है, जिनमें साख नियंत्रण प्रचलित है तथा जहाँ वह लागू नहीं है-

	कीमत ह्रास	कीमत वृद्धि	योग
प्रचलित	862	10	872
अप्रचलित	582	18	600
योग	1444	28	1472

इस तथ्य की जाँच कीजिए कि क्या साख नियंत्रण कीमतों में वृद्धि को रोकने में प्रभावशाली रहा है ?

[x^3 = 6.6 neB meneÙekeâ nw]

12. छोटे आकार की 50 साधारण दुकानों के विषय में निम्न सूचना प्राप्त हुई-

	दुकानें		
	कस्बों में	गाँवों में	योग
पुरुषों द्वारा संचालित	17	18	35
स्त्रियों द्वारा संचालित	3	12	15
योग	20	30	50

क्या यह कहा जा सकता है कि कस्बों की अपेक्षा गाँवों में स्त्री संचालक अपेक्षाकृत अधिक है ? काई-वर्ग परीक्षण का प्रयोग कीजिए।

[x^2 = 3.57 मेट संशोधन सहित 2.48 नहीं H_0 कोई अन्तर नहीं]

13. निम्नलिखित सारणी चेचक के आरम्भ होने पर प्राप्त सामग्री दिखाती है-

	आक्रान्त हुए	आक्रान्त नहीं हुए
टीका लगे हुए	31	469
टीका नहीं लगे हुए	185	1315

चेचक को रोकने में टीके के प्रभाव का परीक्षण कीजिए। आप अपने परिणाम का 5% सार्थकता स्तर पर स्वातन्त्र्य संख्या (d.f.) के लिए x^2 का मूल्य 3.841 है।

[x^2 = 14.65 टीका प्रभावशाली है।]

14. आसंग-गुणांक परिकलित कीजिए-

	A_1	A_2	A_3	A_4
B_1	11	6	2	1
B_2	5	12	15	8
B_3	-	2	3	15

[C = .59]

15. In the following table, the number of student according to age and the regular players among them is given. Assuming that the majority is achieved at the of 18. find out coefficient of association between Majority and playing habit.

Age (yrs.)	15	16	17	18	19	20
No. of students	250	200	150	120	100	80
Regular Players	200	150	90	48	30	12

[M.Com. Indore. 1993]

[Q = – 0.73; No. of minors = 250 + 200 + 150 = 600. Minor players = 200 + 150 + 90 = 440; No. of Majors = 120 + 100 + 80 = 300; Major players = 48 + 30 + 12 = 90]

16. In a study of the stature of husband and wife, the following information is given-

	Tall husband%	**Short-husband%**
Tall wives %	56	13
Short wives %	11	48

Find out coefficient of association in the stature of husband & wife.

[Q = +0.9]

17. The table given below shows the data obtained during an epidemic of cholera-

	Attacked	**Not attacked**	**Total**
Inoculated	31	469	500
Not inoculated	185	1315	1500
	216	1784	2000

Test the effectiveness if inoculation in preventing the attack of cholera.

[Five percent value of x^2 0 for one degree of freedom is 3.84]

[Ans. 0 = 14.64 Not justified]

18. The fortune magazine of U.S.A. published the following results of a sample survey of public opinion regarding election of Roosevelt as the President of U.S.A.-

Attitude towards election	Rich	Poor	Total
Favourable	508	1559	2067
Unfavourable	905	1114	2019
	1413	2673	4086

Is attitude towards election issue guided by the economic status of the voters? What test would you apply? The following table 1% values of Chi-Square is reproduced for your use-

Degree of Freedom :	1	2	3	4
1% value of Chi-Square :	6.635	9.210	11.341	13.277

[Ans. x^2 = 187.1, Yes]

19. In the course of anti-malarial work in Birnagore in the third quarter of 1932, quinine was administered to 606 adults out of a total population of 3,540. The incidence of malarial fever is shown below. Discuss the preventive value of quinine.

	Fever	**No. Fever**	**Total**
Quinine	19	587	606
No. quinine	193	2741	2934
Total	212	3328	3540

You may use the 5% value of Chi-Square for 'n' degrees of freedom, equal to 1, the value being 3.841.

[Ans. x^2 = 10.59- Not independent]

NOTES

20. The following table gives the results of series of controlled experiments. Discuss whether the treatment may be considered to have any positive effect-

	Positive	No effect	Negative
Treatment	9	2	1
Control	3	6	3
	12	8	4 = 24

[Ans. x^2 = 9.367 associated]

21. Find the value of Chi-Square for the following table-

Class	A	B	C	D	E
Observed Frequency	8	29	44	15	4
Theoretical Frequency	7	24	38	24	7

22. In the contingency table, given below, use x^2 test to test for independence of hair colour and eye colour of persons-

Hair Colour Eye Colour	Light	Dark	Total
Blue	26	9	35
Brown	7	18	25
Total	33	27	60

[Ans. The calculated value of x^2 = 15 which is much greater than table value which is 3.84, so that our hypothesis is wrong. We therefore conclude that hair colour and eye colour are associated.]

23. Calculate coefficient of contingency among the weight and mentality of 1,000 criminals from the following table-

Mentality	Weight (Lbs)					Total
	90-120	120-130	130-140	140-150	150 & above	
Normal	50	102	198	210	240	800
Weak	30	38	72	30	30	200
Total	80	140	270	240	270	1,000

[C = 0.216; x^2 = 49.05]

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय-6 आन्तरगणन एवं बाह्यगणन (INTERPOLATION AND EXTRAPOLATION)

इकाई की रूपरेखा

6.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. आन्तरगणन एवं बाह्यगणन के अर्थ को समझ सकेंगे।
2. आन्तरगणन एवं बाह्यगणन की आवश्यकता को समझ सकेंगे।
3. आन्तरगणन एवं बाह्यगणन की विभिन्न रीतियों का ज्ञान होगा।

6.1 प्रस्तावना

केन्द्रीय प्रवृत्ति के विभिन्न मापों (Measures of central tendency) का अध्ययन पिछले अध्यायों में किया गया है। जब अखंडित श्रेणी में मध्यका (Median) अथवा भूयिष्ठक (Mode) ज्ञात करना हो तो हमें सूत्र के माध्यम से वर्गान्तर (Class-interval) में इनका मूल्य ज्ञात करने के लिए आन्तरगणन (Interpolation) का उपयोग करना पड़ता है। यहां आन्तरगणन और बाह्यगणन के संबंध में हम विस्तृत रूप से विवेचन करेंगे।

सांख्यिकीय अनुसंधान करते समय प्राय: यह कठिनाई उपस्थित होती है कि हमें भूतकालीन कई वर्षों के सांख्यिकी आंकड़े प्राप्त तो हो जाते हैं किन्तु किसी–किसी वर्ष के आंकड़े या तो संकलित ही नहीं किये जाते या किन्हीं कारणों से उपलब्ध नहीं होते। ऐसी परिस्थिति में उपलब्ध आँकड़ों द्वारा आन्तरगणन विधि से बीच के वर्षों के आँकड़े प्राप्त किये जा सकते हैं। उपलब्ध आँकड़ों से किसी भविष्य के समय के लिए भी आँकड़ों का अनुमान लगाया जा

सकता है। हमारे देश में जनगणना प्रति दस वर्ष में होती है। सन् 1981 से सन् 1991 तक के जनसंख्या के आँकड़े हमें शासकीय रिकार्ड से प्राप्त हो सकते हैं किन्तु यदि हम सन् 1955 या 1987 के लिए जनसंख्या के आँकड़ें प्राप्त करना चाहते हों तो यह क्रमशः आन्तरगणन और बाह्यगणन द्वारा किया जा सकता है। सांख्यिकीय क्षेत्र में इनका काफी महत्व है।

6.2 आन्तरगणन और बाह्यगणन का अर्थ (Meaning of Interpolation and Extrapolation)

जब दिये हुये आंकड़ों के आधार पर किसी बीच के समय के लिए, किसी विशेष रीति द्वारा, अनुमानित आंकड़े निकाले जाते हैं तो वह आन्तरगणन कहलाता है। यदि यही आँकड़े किसी भावी समय के लिए निकाले जायें तो वह बाह्यगणन होता है। इस प्रकार अनुमानित आँकड़े कुछ कल्पनाओं या मान्यताओं (assumptions) के आधार पर निकाले जाते हैं। अतः यह कहा जा सकता है कि आन्तरगणन और बाह्यगणन ऐसी विधियां हैं जिनसे कुछ मान्यताओं के आधार पर अधिकतम सम्भाव्य अनुमानित आँकड़े प्राप्त किये जाते हैं। आन्तरगणन की परिभाषा इस प्रकार की जा सकती है–

Interpolation and extrapolation are the techniques of obtaining the most likely estimates of certain quantity under certain assumptions.

इसमें स्वतंत्र चर (Independent Variable) को 'x' और आश्रित चर (Dependent Variable) को 'y' से संबोधित किया जाता है।

6.3 मान्यतायें (Assumptions)

यह ऊपर बताया जा चुका है कि आन्तरगणन और बाह्यगणन कुछ मान्यताओं पर आधारित है। ये मान्यतायें निम्नलिखित हैं–

(1) **पद श्रेणियों का पारस्परिक संबंध** – आंतरगणन की यह मान्यता है कि पद श्रेणियां आपस में संबंधित हैं। पहली पद श्रेणी स्वतंत्र और दूसरी उस पर आश्रित है।

(2) **पद श्रेणी के मूल्यों में एकदम उतार–चढ़ाव नहीं** – यह भी माना जाता है कि दो विभिन्न समयों में समंकों में एकदम उतार–चढ़ाव नहीं होता है। जैसे यदि हम 1931, 1941, 1951 और 1961 की जनगणना के आधार पर सन् 1956 की जनसंख्या ज्ञात करना चाहें तो हमारी यह मान्यता होगी कि जनसंख्या के इन आंकड़ों में अधिक उतार–चढ़ाव नहीं हुए हैं।

(3) **परिवर्तन की दर** (Rate of change) **में एकरूपता** – हमारी आन्तरगणन के समय यह भी मान्यता होती है कि दी हुई समयावधि में संमकों में होने वाले परिवर्तन नियमित (regular), एकरूप (uniform) और सामान्य (general) हैं।

(4) **समान परिस्थितियां** – यदि हमें जो समंक उपलब्ध हैं वे विभिन्न वर्गों से संबंधित हैं तो यह माना जाता है। कि उन वर्गों की परिस्थितियां एक समान या लगभग समान हैं।

हमें यह भी नहीं भूलना चाहिए कि आन्तरगणन से प्राप्त समंक विशेष मान्यताओं पर आधारित अधिकतम सम्भाव्य अनुमान मात्र हैं।

अपनी प्रगति की जाँच करें
Test your Progress

6.4 आन्तरगणन और बाह्यगणन की आवश्यकता (Necessity of Interpolation and Extrapolation)

सांख्यिकीय अनुसंधान में इन विधियों की आवश्यकता बहुत अधिक होती है। इसके निम्नलिखित कारण हैं–

(1) **समंकों का अभाव या अपर्याप्तता** (Inadequacy of data) — यदि हम किसी निश्चित और विश्वसनीय परिणाम पर पहुंचना चाहते हों तो यह आवश्यक है कि पर्याप्त आंकड़े उपलब्ध हों। किन्तु व्यवहार में यह देखा जाता है कि आंकड़े या तो उपलब्ध ही नहीं होते और यदि होते भी हैं तो वे अपर्याप्त होते हैं तब आन्तरगणन या बाह्यगणन आवश्यक हो जाता है।

(2) समंकों का नष्ट हो जाना — कभी-कभी किसी विशेष अनुसंधान से संबंधित आंकड़े खो जाते हैं, जल जाते हैं या किसी कारण से नष्ट हो जाते हैं तब भी इन विधियों द्वारा उनका अनुमान लगाया जा सकता है। जैसे किसी व्यापारी के बहीखाते या किसी विद्यार्थी के प्राप्तांक खो जायें या नष्ट हो जायें तो भी इस रीति का प्रयोग होता है।

(3) किसी बीच के समय के लिए समंक प्राप्त करना — आन्तरगणन और बाह्यगणन का प्रयोग उस समय भी करना पड़ता है जबकि समंक एकत्रित किये जाने वाले समय के बीच के किसी समय के लिए समंक प्राप्त करना आवश्यक हो। हमारे देश में, जैसा कि ऊपर बतलाया गया है, जनगणना दसवें वर्ष होती है। यदि शासन, उद्योगपति या अर्थशास्त्री को इन दिये गये वर्षों की जनगणना के अलावा किसी वर्ष की जनगणना मालूम करना हो तो इन विधियों का प्रयोग आवश्यक हो जायेगा।

(4) योजना और वित्तीय प्रबंध में उपयोगी —देश के आर्थिक विकास के लिये योजनाओं का निर्माण किया जाता है और वित्त-व्यवस्था को सुदृढ़ किया जाता है। दोनों के ही लिए उत्पादन और आय आदि का अनुमान लगाना आवश्यक होता है। यह इन्हीं विधियों द्वारा किया जाता है।

(5) व्यापारी और उद्योगपति लाभान्वित — उत्पादन और वस्तु की मांग का भावी अनुमान लगाना प्रत्येक व्यापारी और उद्योगपति के लिए बहुत आवश्यक है। व्यावसायिक पूर्वानुमान उनके लिए अतिमहत्वपूर्ण होता है। व्यापारी तथा उद्योगपति उपलब्ध आंकड़ों के आधार पर आन्तरगणन द्वारा महत्वपूर्ण निष्कर्ष निकालते हैं।

(6) तुलना — आन्तरगणन और बाह्यगणन के द्वारा दो विभिन्न स्थानों के विषय में दिये गये आंकड़ों में तुलना सरल हो जाती है।

आन्तरगणित मूल्य पर किस सीमा तक विश्वास किया जा सकता है ?

आन्तरगणन तथा बाह्यगणन दोनों ही केवल मात्र अनुमान हैं। इनके पूर्णरूपेण शुद्ध होने की आशा करना उचित नहीं है। ऊपर यह बतलाया जा चुका है कि कुछ मान्यताओं पर आधारित ये केवल अधिकतम सम्भाव्य अनुमान हैं। ऊपर जिन मान्यताओं का उल्लेख किया गया है, यदि वे दी गई जानकारी के संबंध में सही हैं तो आन्तरगणित मूल्य पर अधिकतम विश्वास किया जा सकता है। किन्तु प्रश्न यह है कि क्या ये मान्यताएँ सही हैं ? इसके विषय में वही व्यक्ति सही दृष्टिकोण प्रस्तुत कर सकेगा, जिसे पद श्रेणी की पूर्ण जानकारी होगी। यदि दो पद श्रेणियां असंबंधित हैं तो आन्तरगणित मूल्य अविश्वसनीय ही होगा। इस प्रकार यह कहा जा सकता है कि आन्तरगणन और बाह्यगणन की शुद्धता मूल्यों के संभव उच्चावचनों, संबंधित घटनाओं और गणना की उचित रीति के पूर्ण ज्ञान पर निर्भर हैं।

बोध प्रश्न

1. आन्तरगणन से आप क्या समझते हैं।

....

....

....

....

2. बाह्यगणन क्या है।

....

....

....

....

3. आन्तरगणन एवं बाह्यगणन की आवश्यकता समझाइए।

....

....

....

....

NOTES

आन्तरगणन और बाह्यगणन की रीतियां
(Method of Interpolation and Extrapolation)

ये रीतियां मुख्य रूप से दो हैं—

(I) **बिन्दुरेखीय रीति** (Graphic Methods) (II) **बीजगणितीय रीतियां** (Algebraic Method)

(1) एकेन्द्र वक्र अन्वायोजन रीति (Method of Fitting a Parabolic Curve)

(2) परिमितान्तर या द्विपद विस्तार रीति
(Finite Differences Method or Binomical Expansion Method)

(3) न्यूटन- गॉस (अग्रगामी) रीति (Newton–Gauss Forward Method)

(4) न्यूटन- गॉस (विलोमगामी) रीति (Newton–Gauss Backward Method)

(5) स्टर्लिग का सूत्र (Sterling's Formula)

(6) लैग्रेंज की रीति (Lagrange's Method)

अब उपरोक्त रीतियों का सविस्तार विवेचन किया जायेगा।

6.5 बिन्दुरेखीय रीति (Graphic Method)

यह बहुत सरल रीति है। इसमें समय (Time) या मूल्य को क्षैतिज (Horizontal) या भुजाक्ष पर तथा मूल्य या आवृत्ति को उदग्र (Vertical) या कोटिअक्ष पर प्रदर्शित किया जाता है। इसके उपरान्त पैमाना निश्चित करके बिन्दुओं को प्रांकित किया जाता है। इस प्रकार मूल्य या आवृत्तियों का एक वक्र बनता है। अब जिस समय का मूल्य या जिस मूल्य की आवृत्ति ज्ञात करनी हो वहां से एक लम्ब वक्र की ओर खींचा जाता है। जहां यह लम्ब वक्र को काटता है, उस बिन्दु से एक और लम्ब क्षैतिज रेखा के समानान्तर खींचा जाता है। यह उदग्र माप रेखा को किसी एक बिन्दु पर काटेगा। उस बिन्दु का मूल्य या आवृत्ति ही अभीष्ट मूल्य या आवृत्ति होगी। अग्र उदाहरण द्वारा इसे स्पष्ट किया जायेगा।

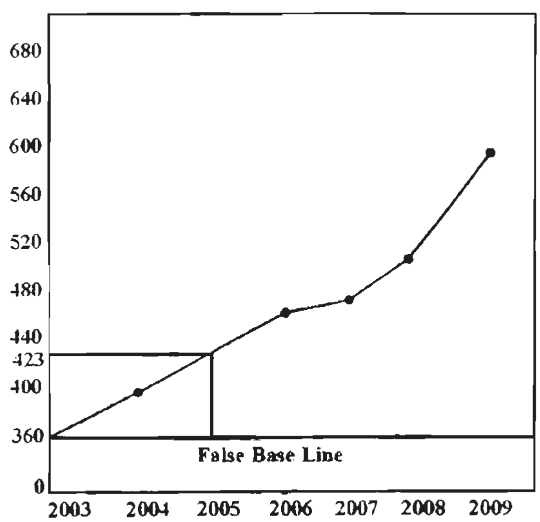
Illustration 1. : The following table gives the population of undivided Bengal during the years 2003 and 2009 in which the figures for 2005 are missing. (अविभाजित बंगाल की 2003 से 2009 तक 2005 को छोड़कर की जनसंख्या नीचे दी गई है)

Year	2003	2004	2006	2007	2008	2009
Population (in lakhs)	363	391	455	467	501	603

With the help of a graph, find the population for the year 2005.

(बिन्दुरेख की सहायता से 2005 की जनसंख्या निकालिए)

Solution –



उपरोक्त बिन्दुरेख पत्र में 2005 से एक लम्ब रेखा खींची गई जो जनसंख्या वक्र को 'P' बिन्दु पर स्पर्श करती है, यहीं से कोटि अक्ष पर एक लम्ब खींचा गया जो 423 (लाख) पर कटता है, अतः 2005 में जनसंख्या का अनुमान 423 लाख व्यक्ति रहा।

6.6 बीजगणितीय रीतियां (Algebraic Method)

(1) **एकेन्द्र वक्र अन्वायोजन रीति** (Method of Fitting a Parabolic Curve) यह रीति इस मान्यता पर आधारित है कि दो पद श्रेणियों में परस्पर संबंध है और जब हमें एक पद श्रेणी के मूल्य ज्ञात हों तो उससे संबंधित दूसरी पद श्रेणी के मूल्य भी ज्ञात किये जा सकते हैं।

आन्तरगणन और बाह्यगणन के लिए इस रीति का प्रयोग उन अवस्थाओं में किया जा सकता है जबकि मूल्य समान समयान्तर से दिये गये हों।

इस वक्र के लिए निम्न सूत्र का प्रयोग होता है–

$$y = a + bx + cx^2 + dx^3 + ex^4 ....nx^n$$

इस सूत्र में 'x' को स्वतंत्र चल (Independent Variable) और 'y' को 'x' पर आश्रित अथवा उसका कार्य (Function) माना गया है। 'a', 'b', 'c', 'd' और 'e' अचल पद (Constants) हैं जिन्हें ज्ञात करना है। एकेन्द्र वक्र (Parabolic Curve) इसी समीकरण की सहायता से खींचा जाता है। इस सूत्र का प्रयोग करते समय यह ध्यान रखना है कि दिये गये प्रश्न में जितने पद ज्ञात हों उससे एक वक्र कम (n–1) का एकेन्द्र वक्र वाला सूत्र प्रयोग करना चाहिए। यदि प्रश्न में ज्ञात पद 5 हों तो 4 घात (Power) वाले सूत्र का प्रयोग करना होगा।

Illustration 2: The following table gives the cube of the items. Find out cube of 2.5 (निम्न सारणी में पदों का घनमूल्य दिया है। 2.5 का घनमूल्य ज्ञात कीजिए)–

Size of Item (पद)	1	2	3	4
Cube Value (घन मूल्य)	1	8	27	64

Solution – इस उदाहरण में ज्ञात मूल्य 4 हैं, अतः 3 एकेन्द्र वक्र वाले सूत्र का प्रयोग होगा जो इस प्रकार है – $y = a + bx + cx^2 + dx^3$

Size of Item (x)	Deviation from 2.5	Common factor	Cube Value (y)
1	–1.5	–3	1
2	–0.5	–1	8
2.5	0	0	y_0
3	+0.5	+1	27
4	+1.5	+3	64

'x' और 'y' के मूल्यों को सूत्र में लिखने से निम्न समीकरण प्राप्त होते हैं–

$$1 = a + b(-3) + c(-3)^2 + d(-3)^3$$

or $\quad 1 = a - 3b + 9c - 27d \quad ..... \quad ...... \quad (i)$

$$8 = a + b(-1) + c(-1)^2 + d(-1)^3$$

or $\quad 8 = a - b + c - d \quad ..... \quad ..... \quad (ii)$

$$y0 = a + b(0) + c(0)^2 + d(0)^3$$

or $\quad y0 = a \quad ..... \quad ..... \quad (iii)$

$$27 = a + b(1) + c(1)^2 + d(1)^3$$

or $27 = a + b + c + d$ .... (iv)

$64 = a + b\,(3) + c\,(3)^2 + d\,(3)^3$

or $64 = a + 3b + 9c + 27\,b$ .... (v)

Adding Equation (i) and (iv) we have,

$65 = 2a + 18\,c$ .... (vi)

Adding Equation (ii) and (iv) we have ,

$35 = 2a + 2c$ .... (vii)

Multiplying Equation (vii) by 9, we have

$315 = 18\,a + 18\,c$ .... (viii)

Subtracting Equation No. (vi) from (viii), we have

$250 = 16a$ .... (ix)

or $a = \frac{250}{16} = 15.625$

Substituting the above value of 'a' in equation (iii) we have

$y0 = a$ or $y0 = 15.625$

Hence cube value of 2.5 is 15.625.

Illustration 3. The following table gives the census of population of an Indian State in 2006, 2007, 2008, and 2009.

(निम्न सारणी में भारत राज्य की जनगणना दी है)

Estimate the population of the state in 2009, aking your method clear– (2009 में राज्य की जनसंख्या का अनुमान लगाइये)

Year	2006	2007	2008	2010
Population	2,797	2,935	3,047	3,354
(In thousands)				

Solution– Estimate of the census of population in 2009 by fitting a parabolic curve –

Year	2006	2007	2008	2009	2010
Deviation from 2009 (x)	–23	–13	–3	0	+7
Population (In thousands)	2,797	2,935	3,047	y_o	3354

Since the known values are four, we will take parabola of the 3rd order, i.e.,

$$y = a + bx + cx^2 + dx^3$$

Substituting the different values of x and y, we get the following equations–

$2797 = a - 23b + 529\,c - 12167d$ .... (i)

$2935 = a - 13b + 169\,c - 2197d$ .... (ii)

$3047 = a - 3b + 9c - 27d$ .... (iii)

$y0 = a$ .... (iv)

$3354 = a + 7b + 49\,c + 343\,d$ .... (v)

NOTES

To find out the values of 'y_0' which will give the estimated population of the state in 2009, we proceed as follows–

Multiplying (i), (ii) and (v) by 3 and the (iii) by 23, 13 and 7, we get

$$8391 = 3a - 69b + 1587c - 36501d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(vi)}$$

$$8805 = 3a - 39b + 507c - 6591d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(vii)}$$

$$10062 = 3a + 21b + 147c + 1029d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(viii)}$$

$$70081 = 23a - 69b + 207c - 621d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(ix)}$$

$$39611 = 13a - 39b + 117c - 351d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(x)}$$

$$21329 = 7a - 21b + 63c - 189d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xi)}$$

Substracting (vi) from (ix) and (vii) from (x),

$$61690 = 20a - 1380c + 35880d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xii)}$$

$$30806 = 10a - 390c + 6240d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xiii)}$$

Adding (viii) and (xi) ,

$$31391 = 10a + 210c + 840d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xiv)}$$

Multiplying (xii) and (xiii) by 7 and (xiv) by 46 and 13,

$$431830 = 140a - 9660c + 251160d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xv)}$$

$$215642 = 70a - 2730c + 43680d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xvi)}$$

$$1443986 = 460a + 9660c + 38640d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xvii)}$$

$$408083 = 130a + 2730c + 10920d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xviii)}$$

Adding (xv) and (xvii), and (xvi) and (xviii)

$$1875816 = 600a + 289800d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xix)}$$

$$623725 = 200a + 53600d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xx)}$$

Multiplying (xix) by 13 and (xx) by 69,

$$24385608 = 7800a + 3767400d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xxi)}$$

$$43037025 = 13800a + 3767400d \quad \ldots\ldots \quad \ldots\ldots \quad \text{(xxii)}$$

Substracting (xxi) from (xxii)

$$18651417 = 6000a$$

or $$a = \frac{18651417}{6000} = 3108.5595$$

or $$y_0 = 3108.5695$$

(Population for 2009 in thousands.)

Illustration 4. : Estimate annual sales of pencils for 2005 from the following records of wholesale merchants (थोक व्यापारियों के निम्न रिकार्ड से 2005 के लिए वार्षिक विक्रय का अनुमान कीजिए)

Year	2002	2003	2004	2006	2007
Sale of Pencils (in lakhs of dozens)	25	30	40	55	60

Solution : Since the known values are five, an equation of the parabola of the 4th order will be taken, as

$$y = a + bx + cx^2 + dx^3 + ex^4$$

NOTES

Now –

Year	2002	2003	2004	2005	2006	2007
Deviations from 2005 (x)	–10 or – 5	–6 –3	–2 –1	0 0	+2 +1	+6 +3
y (Sale of Pencils	25	30	40	y_0	55	60

Substituting the given values of x and y, we get the following equations–

$$25 = a - 5b + 25c - 125d + 625e \quad \text{.......} \quad \text{(i)}$$

$$30 = a - 3b + 9c - 27d + 81e \quad \text{.......} \quad \text{(ii)}$$

$$40 = a - b + c - d + e \quad \text{.......} \quad \text{(iii)}$$

$$y0 = a \quad \text{.......} \quad \text{(iv)}$$

$$55 = a + b + c + d + e \quad \text{.......} \quad \text{(v)}$$

$$60 = a + 3b + 9c + 27d + 81e \quad \text{.......} \quad \text{(vi)}$$

Multiplying (v) by 5 and 3,

$$275 = 5a + 5b + 5c + 5d + 5e \quad \text{.......} \quad \text{(vii)}$$

$$165 = 3a + 3b + 3c + 3d + 3e \quad \text{.......} \quad \text{(viii)}$$

Adding (i) with (vii) with (ii) with (viii), we get

$$300 = 6a + 30c - 120d + 630e \quad \text{.......} \quad \text{(ix)}$$

$$195 = 4a + 12\ c - 24d + 84e \quad \text{.......} \quad \text{(x)}$$

Multiplying (x) by 5,

$$975 = 20a + 60c - 120d + 420e \quad \text{.......} \quad \text{(xi)}$$

Subtracting (ix) from (xi),

$$675 = 14\ a + 30\ c - 210\ e \quad \text{........} \quad \text{(xii)}$$

Adding (ii) with (vi) and (iii) with (v),

$$90 = 2a + 18c + 162e \quad \text{.........} \quad \text{(xiii)}$$

$$95 = 2a + 2c + 2e \quad \text{........} \quad \text{(xiv)}$$

Multiplying (xiv) by 9 and 15,

$$855 = 18a + 18c + 18e \quad \text{........} \quad \text{(xv)}$$

$$1425 = 30a + 30c - 30e \quad \text{........} \quad \text{(xvi)}$$

Subtracting (xiii) from (xv) and (xii) by (xvi)

$$765 = 16a + 144\ e \quad \text{........} \quad \text{(xvii)}$$

$$750 = 16a + 240\ e \quad \text{........} \quad \text{(xviii)}$$

Multiplying (xvii) by 5 and (xviii) by 3,

$$3825 = 80\ a - 720\ e \quad \text{........} \quad \text{(xix)}$$

$$2250 = 48\ a + 720\ e \quad \text{........} \quad \text{(xx)}$$

Adding (xix) and (xx)

$$6075 = 128\ a$$

NOTES

or $a = \frac{6075}{128} = 47.46$

or $y0 = 47.46$ (Lakhs dozens of Pencils)

इस पद्धति का सबसे बड़ा दोष यह है कि जब पदों की संख्या अधिक होती है तो समीकरण बहुत बन जाते हैं जिन्हें एक साथ हल करना पड़ता है। समीकरणों को इस प्रकार हल करना बहुत कठिन है। अतः इस रीति का प्रयोग उन्हीं अवस्थाओं में करना चाहिए जबकि पदों की संख्या कम हो। गणितीय दृष्टि से तो यह रीति सर्वोत्तम है क्योंकि यह हर प्रकार की सतत श्रेणियों में प्रयुक्त की जा सकती है।

6.7 परिमितान्तर या द्विपद विस्तार रीति

(Finite Differences Method or Binomial Expansion Method)

इस सूत्र का प्रतिपादन न्यूटन द्वारा परिमित (finite) अन्तर के आधार पर किया गया है। यह द्विपद प्रमेय (Binomial Theorem) पर आधारित रीति है।

इस रीति का प्रयोग निम्न अवस्थाओं में ही हो सकता है —

i. श्रेणी समान वर्ग-विस्तार (Equal class-interval) वाली हो।

ii. वर्ग-विस्तार में ही आने वाले किसी पद मूल्य को ज्ञात करना हो।

यह निम्न उदाहरण से स्पष्ट हो जायेगा —

मान लीजिये हमें निम्न श्रेणी दी हुई है —

x	10	20	30	40	50
y	4	5	6	?	8

इसमें 'x' श्रेणी समान वर्गान्तर (10 इकाइयों) से बढ़ती है और 'x' का मूल्य (40) जिसके लिए हमें 'y' का मूल्य ज्ञात करना है, वह भी वर्ग-विस्तार का ही एक भाग है। इस प्रकार यह रीति वहीं अपनाई जा सकती है जहाँ प्रत्येक पद के बीच का अन्तर बराबर हो और उन्हीं में से किसी एक पद का आन्तरगणन या बाह्यगणन करना हो।

इस रीति का आधार इस मान्यता पर निर्भर है कि 'N' वाँ प्रमुख अन्तर (Leading difference) शून्य होगा। इसे समीकरण द्वारा निम्न प्रकार स्पष्ट किया जा सकता है-

$$(y-1)^n = 0$$

If $n = 4$; $(y-1)^4 = y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$

If $n = 5$; $(y-1)^5 = y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0 = 0$

इसी प्रकार इसे आगे बढ़ाया जाता है। इनमें प्रथम मूल्य धनात्मक, दूसरा ऋणात्मक, तीसरा धनात्मक, चौथा ऋणात्मक. पाँचवा धनात्मक और इसी प्रकार होता जाता है। इस रीति की विस्तृत प्रक्रिया निम्न प्रकार है-

i. स्वतन्त्र चर-मूल्य (x) के पदों को क्रमानुसार x_0, x_1, x_2, x_3 .... आदि संकेताक्षरों द्वारा प्रदर्शित किया जाता है। इसी प्रकार, आश्रित श्रेणी (y) के मूल्यों के y_0, y_1, y_2, y_3 ...आदि संकेतों का प्रयोग किया जाता है। इन्हीं संकेताक्षरों में से एक अज्ञात होता है।

ii. y के जितने ज्ञात मूल्य होते हैं उतने क्रम के प्रमुख अन्तर (Leading difference) को शून्य (0) माना जाता है। यदि 5 पद ज्ञात हों तो पाँचवा प्रमुख अन्तर शून्य होगा। इसका सूत्र इस प्रकार है —

$$\Delta^n{}_0 = 0$$

'n' का अर्थ ज्ञात मूल्यों की संख्या होता है।

iii. ज्ञात-मूल्यों की संख्या के अनुसार प्रमुख-अन्तर को शून्य मानते हुए उसका द्विपद-विस्तार निम्नलिखित सूत्र के रूप में लिखा जावेगा –

$$\Delta^n{}_0 = (y-1)^n = 0$$

$$= y^n - ny^{n-1} + \frac{n(n-1)}{1 \times 2} y^{n-2} - \frac{n(n-1)(n-2)}{1 \times 2 \times 3} y^{n-3} + \ldots\ldots\ldots\ldots\ldots = 0$$

यदि y के ज्ञात मूल्य 5 हों तो —

$$\Delta^n_0 = (y-1)^5 = 0$$

$$= y^5 - ny^{5-1} = \frac{5(5-1)}{1\times 2}y^{5-2} - \frac{5(5-1)(5-2)}{1\times 2\times 3}y^{5-3}$$

$$+ \frac{5(5-1)(5-2)(5-3)}{1\times 2\times 3\times 4}y^{5-4}$$

$$- \frac{5(5-1)(5-2)(5-3)(5-4)}{1\times 2\times 3\times 4\times 5}y^{5-5} = 0$$

$$= y^5 - 5y^4 + 10y^3 - 10y^2 + 5y^1 - y^0 = 0$$

इस सूत्र का उपयोग करते समय घातों (Powers) का उपयोग उपसंकेत (subscript) के रूप में करते हैं।

iv. द्विपद विस्तार का ऊपर–वर्णित स्वरूप जटिल है। इसे निकालने की सरल विधि इस प्रकार है–

(अ) जिस प्रमुख अन्तर के लिए द्विपद विस्तार करना हो, पहले उस क्रम के y को लिखा जावेगा फिर अवरोही क्रम में y का घात एक–एक कम करते जावेंगे ताकि अन्त में y^0 आ जावे। जैसे,

$$y^5,\ y^4,\ y^3,\ y^2,\ y^1,\ y^0$$

(ब) प्रथम y धनात्मक होगा, अगला ऋणात्मक, फिर उससे अगला धनात्मक और इसी प्रकार अन्त तक चिन्ह लिखे जावेंगे –

$$+y^5,\ -y^4,\ +y^3,\ -y^2,\ +y^1,\ -y^0$$

(स) विभिन्न y's के अंकात्मक गुणक (numerical coefficients) इस प्रकार निकाले जावेंगे –

सबसे पहले लिखे जाने वाले y का गुणक –1 होगा। इसके बाद के y's अंकात्मक गुणक प्राप्त करने के लिए इस सूत्र का उपयोग किया जावेगा –

$$= \frac{(\text{पिछले } y \text{ का गुणक}) \ \text{X} \ (\text{पिछले } y \text{ का उपसंकेत या घात})}{\text{पिछले } y \text{ की क्रम स्थिति}}$$

उपर्युक्त उदाहरण में अंकात्मक–गुणक इस प्रकार निकलेंगे–

$$1y^5 - \frac{1\times 5}{1}y^4 - \frac{5\times 4}{2}y^3 - \frac{10\times 3}{3}y^2 - \frac{10\times 2}{4}y^1 - \frac{5\times 1}{5}y^0$$

$$= y^5 - 5y^4 + 10y^3 - 10y^2 + 5y^1 - y^0 = 0$$

कुछ द्विपद-विस्तार इस प्रकार ज्ञात होंगे –

ज्ञात मूल्यों की संख्या	मूल–सूत्र	द्विपद–विस्तार
2	$(y-1)^2 = 0$	$y^2 - 2y^1 + y^0 = 0$
3	$(y-1)^3 = 0$	$y^3 - 3y^2 + 3y^1 - y^0 = 0$
4	$(y-1)^4 = 0$	$y^4 - 4y^3 + 6y^2 - 4y^1 + y^0 \equiv 0$
5	$(y-1)^5 = 0$	$y^5 - 5y^4 + 10y^3 - 10y^2 + 5y^1 - y^0 = 0$
6	$(y-1)^6 = 0$	$y^6 - 6y^5 + 15y^4 - 20y^3 + 15y^2 - 6y^1 + y^0 = 0$
7	$(y-1)^7 = 0$	$y^7 - 7y^6 + 21y^5 - 35y^4 + 35y^3 - 21y^2 + 7y^1 - y^0 = 0$

अग्रलिखित उदाहरणों द्वारा यह रीति स्पष्ट हो जावेगी–

Illustration 5. Using any interpolation method, find out the likely Index Number for 2007 from the following table (2007 के लिए संभावित निर्देशांक ज्ञात कीजिए।)–

Year	2005	2006	2007	2008	2009
Index No.	100	107	y_2	157	212

NOTES

Year(x)	2005	2006	2007	2008	2009
Index No. (y)	100 y_0	107 y_1	....	157y_3	212 y_4

Hence, the equation will be

$(y-1)^4 = 0$

इस समीकरण का विस्तार इस प्रकार होगा —

$$y_4 - 4y_{4-1} + \frac{4(4-1)}{1\times 2}y_{4-2} - \frac{4(4-1)(4-2)}{1\times 2\times 3}y_{4-3} + \frac{4(4-1)(4-2)(4-3)}{1\times 2\times 3\times 4}y_{4-4} = 0$$

or $\quad y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$

By substituting values, we get

$212 - (4 \times 157) + 6y_2 - (4 \times 107) + 100 = 0$

or $\quad 212 - 628 + 6y_2 - 428 + 100 = 0$

or $\quad 6y_2 = -212 + 628 - 100 + 428$

or $\quad 6y_2 = 1056 - 312 \qquad$ or $\quad 6\ y_2 = 744$

or $\quad y_2 = 124$

$\therefore$ Index Number for 2007 will be 124.

Illustration 6. Obtain the missing figure in the following table —(निम्न सारणी में लुप्त अंक प्राप्त कीजिए) –

Value of Chi–Square at 1% Level of Significance.

Degrees of freedom	%Chi–Square	
1	6.64	y_0
2	9.21	y_1
3	11.34	y_2
4	13.28	y_3
5		y_4
6	16.81	y_5
7	18.48	y_6
8	20.07	y_7
9	21.67	y_8

Solution — Since the known quantities are eight, the eighth leading difference will be zero.

$(y - 1)^8 = 0$

or $\quad y_8 - 8y_7 + 28y_6 - 56y_4 + 70y_4 - 56y_3 + 28y_2 - 8y_1 + y_0 = 0$

Substituting the given values, we get

$$21.67 - (8 \times 20.07) + (28 \times 18.48) - (56 \times 16.81) + 70y_4 - (56 \times 13.28) + (28 \times 11.34) - (8 \times 9.21) + 6.64 = 0$$

or $\quad 21.67 - 160.56 + 517.44 - 941.36 + 70\ y_4 - 743.68 + 317.52 - 73.68 + 6.64 = 0$

or $\quad 70y_4 = 160.56 - 21.67 - 517.44 + 941.36 + 743.68 - 317.52 + 73.68 - 6.64$

or $\quad 70y_4 = 1056.01 \quad$ or $\quad y_4 = 15.09$

Thus the interpolated value of Chi–Square at 1% level of significance for 5 degrees of freedom is 15.09.

दो अज्ञात मूल्य (Two unknown values) — जब स्वतंत्र चर–मूल्यों के अन्तर समान हों और आश्रित मूल्यों में दो अज्ञात मूल्यों का आन्तरगणन करना हो तो दो समीकरण बनाये जाते हैं। प्रथम, ज्ञात मूल्यों की संख्या के बराबर प्रमुख अन्तर को शून्य मानकर द्विपद–विस्तार लिखा जाता है। दूसरे, इस द्विपद विस्तार को फिर से लिखकर प्रत्येक y के उपसंकेत (Subscript or power) में 1 की वृद्धि कर देते हैं। जिससे अन्त में y^0 के स्थान पर y^1 प्राप्त हो जाता है। या, दूसरा समीकरण ऐसे तैयार किया जाता है, मान लो पहले समीकरण से एक मूल्य ज्ञात हो गया है और कुछ मूल्यों की संख्या एक से बढ़ गई है। दोनों समीकरणों में ज्ञात मूल्यों को आविष्ट करके अज्ञात मूल्य निकाले जाते हैं। यह निम्नलिखित उदाहरण द्वारा स्पष्ट किया जायेगा–

Illustration 7. Interpolate the missing figures in the following table– (निम्नलिखित सारणी में लुप्त संख्याएँ आन्तरगणित कीजिए)

Year	Acres in Millions
2000	76.6
2001	78.7
2002	.....
2003	77.7
2004	78.7
2005	....
2006	80.6
2007	77.6
2008	78.7

Solution– इस प्रश्न में दो मूल्यों का आन्तरगणन करना है अत: एक समीकरण में अन्तिम पद को छोड़ दिया जायेगा ताकि दोनों मूल्यों की गणना की जा सके।

Year	Acres in Millions	
2000	76.6	y_0
2001	78.7	y_1
2002	....	y_2
2003	77.7	y_3
2004	78.7	y_4
2005	....	y_5
2006	80.6	y_6
2007	77.6	y_7
2008	78.7	y_8

इस प्रश्न में निम्न दो समीकरण इस प्रकार बनेंगे –

(i) $\quad (y-1)^8 = y_8 - 8y_7 + 28y_6 - 56\,y_5 + 70y_4 - 56y_3 + 28y_2 - 8y_1 + y_0 = 0$

(ii) $\quad (y-1)^7 = y_7 - 7y_6 + 21y_5 - 35y_4 + 35y_3 - 21y_2 + 7y_1 - y_0 = 0$

NOTES

By substituting values, we have

$$78.7 - (8 \times 77.6) + (28 \times 80.6) - 56y_5 + (70 \times 78.7)$$
$$-(56 \times 77.7) + 28y_2 - (8 \times 78.7) + 76.6 = 0 \quad \text{(i)}$$

$$77.6 - (7 \times 80.6) + 21y_5 - (35 \times 78.7) + (35 \times 77.7)$$
$$- 21y_2 + (7 \times 78.7) - 76.6 = 0 \quad \text{(ii)}$$

or
$$- 56y_5 + 28y_2 + 7921.1 - 5601.6 = 0 \quad \text{(iii)}$$
$$21y_5 - 21y_2 + 3348 - 3395.3 = 0 \quad \text{(iv)}$$

or
$$- 56y_5 + 28y_2 = -2319.5 \quad \text{(v)}$$
$$21y^5 - 21y^2 = +47.3 \quad \text{(vi)}$$

Multiplying (v) by 3 and (vi) by 4, we get

$$- 168y_5 + 84y_2 = - 6958.5 \quad \text{(vii)}$$
$$84y_5 - 84y_2 = 189.2 \quad \text{(viii)}$$

Adding (vii) and (viii) we get

$$- 84\ y_5 = - 6769.3 \quad \text{or} \quad y_5 = \frac{6769.3}{84} = 80.6$$

Substituting the value of 'y_5' in equation (v) we get,

$$(- 56 \times 80.6) + 28y_2 = - 2319.5$$

or $$- 4513.6 + 28y_2 = - 2319.5$$

or $$28y_2 = 4531.6 - 2319.5$$

or $$28y_2 = 2194.1$$

or $$y_2 = \frac{194.1}{28} = 78.3$$

Hence, the Acres in Millions in 2002 = 78.3 in 1990 = 80.6

Illustration 8. The following table gives the population of Indore at the time of the last six censuses — (पिछले 6 जनगणनाओं में इंदौर की जनगणना दी हुई है) :–

Years	1951,	1961,	1971,	1981,	1991,	2001
Population	75401,	82984,	86686,	44947,	93091,	1,27,327

Estimate the population for 2011. [2011 की जनसंख्या का अनुमान कीजिए]

Solution– उपरोक्त आँकड़ों में जनसंख्या सन् 1951 से 1971 तक निरन्तर बढ़ रही है जबकि 1981 में यह एकदम कम हो जाती है और फिर 1991 में पुनः एक दम बढ़ी है। अतः 2011 की जनसंख्या का अनुमान लगाने से पूर्व 1981 की जनसंख्या का आन्तरगणन आवश्यक है क्योंकि इसकी मान्यताओं में यह बतलाया गया है कि श्रेणी में परिवर्तन एकरूप होना चाहिए।

1981 के लिए आन्तरगणन (Interpolation for 1981)

Year	**Population**
1951	75,401 y_0
1961	82,984 y_1
1971	86,686 y_2
1981	...? y_3
1991	93,091 y_4
2001	1,27,327 y_5

Equation will be

$$(y-1)^5 \text{ or } y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0 = 0$$

Substituting the values, we have

$127327 - (5 \times 93091) + 10y_3 - (10 \times 86686) + (5 \times 82984) - 75401 = 0$

or $\quad 127327 - 465455 + 10y_3 - 866860 + 414920 - 75401 = 0$

or $\quad 10y_3 = -127327 + 465455 + 866860 - 414920 + 75401$

or $\quad 10y_3 = 865469 \quad$ or $\quad y_3 = 86547.$

इस प्रकार सन् 1981 की जनसंख्या 86,547 मानकर सन् 2011 की जनसंख्या का अनुमान निम्न प्रकार से लगाया जायेगा —

Year	Population
1951	75,401 y_0
1961	82,984 y_1
1971	86,686 y_2
1981	86,547 y_3
1991	93,091 y_4
2001	1,27,327 y_5
2011	...? y_6

Since the known figures are six, the equation will be

$$(y-1)^6 = 0 \text{ or } \quad y_6 - 6y_5 + 15y_4 - 20y_3 + 15y_2 - 6y_1 + 6y_0 = 0$$

Substituting the values, we have

$$y_6 - (6 \times 127327) + (15 \times 93091) - (20 \times 86547)$$
$$+ (15 \times 86686) - (6 \times 82984) + 75401 = 0$$

or $\quad y_6 - 763962 + 1396365 - 1730940 + 1300290 - 497904 + 75401 = 0$

or $\quad y_6 = 763962 - 1396365 + 1730940 - 1300290 + 497904 - 75401 = 0$

or $\quad y_6 = 2992806 - 2772056 \quad$ or $\quad y_6 = 220750.$

Hence, the population estimate for 2011 will be 220750.

नोट– यहाँ भी सन् 2001 में जनसंख्या में यकायक वृद्धि ने सन् 2011 की जनसंख्या को भी बहुत बढ़ा दिया है। इस प्रश्न को इसलिए, सन् 2001 की जनसंख्या को ज्ञात करके भी किया जा सकता है।

6.8 न्यूटन की प्रगामी अन्तर रीति (Newton's Method of Advancing Differences)

इस रीति का प्रयोग तभी होता है जबकि स्वतन्त्र चर (Independent variable) समान अन्तर से पड़ता है और इससे श्रेणी के प्रारम्भ में आन्तरगणन का सर्वोत्तम अनुमान प्राप्त होता है। गणना की दृष्टि से यह सरल रीति है। इस रीति की निम्नलिखित विशेषताएँ हैं —

(1) यह रीति द्विपद प्रमेय (Binomial Theorem) पर आधारित है।

(2) इस रीति से आन्तरगणन और बाह्यगणन का अनुमान लगाया जा सकता है।

(3) इस रीति का प्रयोग समान अन्तर के पदों वाली श्रेणी में किया जाता है।

(4) इस रीति का प्रयोग सामान्यत: श्रेणी के प्रारम्भिक मूल्यों को ज्ञात करने के लिए किया जाता है।

(5) इस रीति में 'y' के मूल्यों का आपसी अन्तर निकाला जाता है और इसके बाद इन अन्तरों का भी अन्तर उस सीमा तक निकाला जाता है जब तक कि अन्तर केवल एक पद नहीं रह जाये।

NOTES

(6) अन्तर निकालने के लिए हमेशा अगले मूल्य से पिछले मूल्य को घटाते हैं। अन्तर निकालते समय धन (+) और ऋण (–) चिन्हों का प्रयोग किया जाता है —

(7) इस रीति में आन्तरगणन और बाह्यगणन के लिए निम्न सूत्र का प्रयोग किया जाता है —

$$yx = y_0 + x\,\Delta_0^1 + \frac{x(x-1)}{1\times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1\times 2\times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1\times 2\times 3\times 4}\Delta_0^4 \;.....\text{ and so on.}$$

where, 'yx' = Figure to be interpolated. y_o = Value of the year of origin.

"Δ' = Differences between the adjoining years

$$'x' = \frac{(\text{Item of interpolation}) - (\text{Item of origin})}{\text{Distance between two adjoining items}}$$

Table showing finie or Advancing Differences

X	Y	Finiter of Advancing Differences							
		First Differences Δ^1		Second Differences Δ^2		Third Differences Δ^3		Fourth Differences Δ^4	
x_o	y_o								
x_1	y_1	$y_1 - y_o$	Δ_o^1						
x_2	y_2	$y_2 - y_1$	Δ_1^1	$\Delta_1^1 - \Delta_o^1$	Δ_o^2				
x_3	y_3	$y_3 - y_2$	Δ_2^1	$\Delta_2^1 - \Delta_1^1$	Δ_1^2	$\Delta_1^2 - \Delta_o^2$	Δ_o^3		
x_4	y_4	$y_4 - y_3$	Δ_3^1	$\Delta_3^1 - \Delta_2^1$	Δ_2^2	$\Delta_2^2 - \Delta_1^2$	Δ_1^3	$\Delta_1^3 - \Delta_o^3$	Δ^4

अन्तरों को बड़ी सावधानी से निकालना चाहिए। यदि एक अन्तर की गणना में त्रुटि हो गई तो बाद के सब अन्तरों में त्रुटि हो जायेगी।

Illustration 9. The following table shows the expectation of life at different ages. You are required to find out the expectation of life at the age of 16. (विभिन्न आयु पर जीवन अनुशंसा दी हुई है। 16 वर्ष की आयु पर जीवन-अनुशंसा ज्ञात कीजिए) : —

Age in years	Expectation of life
10	35
15	33
20	29
25	27
30	22
35	20

Solution — $x = \dfrac{\text{Age of interpolation} - \text{Age of origin}}{\text{Age distance between adjoining ages}} = \dfrac{16-10}{5} = \dfrac{6}{5} = 1.2$

Age in Year 5	Expectation in life	First Diff. Δ^1	Second Diff Δ^2	Third Diff. Δ^3	Fourth Diff. Δ^4	Fifth Diff. Δ^5
10 x_0	35 y_0					
15 x_1	33 y_1	$-2\,\Delta_0^1$				
20 x_2	29 y_2	$-4\,\Delta_1^1$	$-2\,\Delta_0^2$	$+4\,\Delta_0^3$	$-9\,\Delta_0^4$	$+20\,\Delta_0^5$
25 x_3	27 y_3	$-2\,\Delta_2^1$	$+2\,\Delta_1^2$	$-5\,\Delta_1^3$	$+11\,\Delta_1^4$	
30 x_4	22 y_4	$-5\,\Delta_3^1$	$-3\,\Delta_2^2$	$+6\,\Delta_2^3$		
35 x_5	20 y_5	$-2\,\Delta_4^1$	$+3\,\Delta_3^2$			

Application of formula —

$$yx = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1 \times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1 \times 2 \times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1 \times 2 \times 3 \times 4}\Delta_0^4 + \frac{x(x-1)(x-2)(x-3)(x-4)}{1 \times 2 \times 3 \times 4 \times 5}\Delta_0^5$$

where, 'yx' = The figure to be interpolated Δs = Differences

$$x = \frac{\text{Age of interpolation} - \text{Age of origin}}{\text{Age distance between adjoining ages}} = 1.2$$

$$yx = 35 + (1.2 \times -2) + \frac{1.2(1.2-1)}{1 \times 2} \times -2 + \frac{1.2(1.2-1)(1.2-2)}{1 \times 2 \times 3} \times 4$$

$$+ \frac{1.2(1.2-1)(1.2-2)(1.2-3)}{1 \times 2 \times 3 \times 4} \times -9$$

$$+ \frac{1.2(1.2-1)(1.2-2)(1.2-3)(1.2-4)}{1 \times 2 \times 3 \times 4 \times 5} \times 20$$

$$= 35 + (-2.4) + \left(\frac{.24}{2} \times -2\right) + \left(\frac{-.192}{6} \times 4\right)$$

$$+ \left(\frac{.3456}{34} \times -9\right) + \left(\frac{-.96768}{120} \times 120\right)$$

$$= 35 - 2.4 - .24 - .128 - .1296 - .1613 = 31.94 .$$

Illustration 10. : (Interpolation in a frequency Distribution)

From the following data, estimate the number of persons earning wages between 60 and 70 rupees (60 से 70 रु. के बीच मजदूरी पाने वालों की संख्या निकालिए) :–

Wages (in Rs.)	No. of Persons (in thousands)
Below 40	250
40–60	120
60–80	100
80–100	70
100–120	50

Solution— अविच्छिन्न श्रेणी में सर्वप्रथम आवृत्ति को संचयी आवृत्ति में निम्न प्रकार बदला जायेगा–

Wages (Rs.)	No. of Persons (Thousands)
Below 40	250
" 60	370
" 80	470
" 100	540
" 120	590

Wages in Rs.		No. of Persons in Thousands		First Diff. Δ^1	Second Diff. Δ^2	Third Diff. Δ^3	Fourth Diff. Δ^4
Below 40	x_0	250	y_0	$+120\ \Delta_0^1$			
" 60	x_1	370	y_1	$+100\ \Delta_1^1$	$-20\ \Delta_0^2$		
" 80	x_2	470	y_2	$+70\ \Delta_2^1$	$-30\ \Delta_1^2$	$-10\ \Delta_0^3$	
" 100	x_3	540	y_3	$+50\ \Delta_3^1$	$-20\ \Delta_2^2$	$+10\ \Delta_1^3$	$+20\ \Delta_0^4$
" 120	x_4	590	y_4				

NOTES

$$x = \frac{\text{Wages below 70} - \text{Wages below 40}}{\text{Wage distance 20}} = \frac{70 - 40}{20} = 1.5$$

Newton's Formula –

$$yx = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1 \times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1 \times 2 \times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1 \times 2 \times 3 \times 4}\Delta_0^4$$

$$= 250 + (1.5 \times 120) + \frac{1.5(1.5-1)}{1 \times 2} \times -20 + \frac{1.5(1.5-1)(1.5-2)}{1 \times 2 \times 3} \times -10 + \frac{1.5(1.5-1)(1.5-2)(1.5-3)}{1 \times 2 \times 3 \times 4} \times 20$$

$$= 250 + 180 - 7.5 + .625 + .469 = 423.6$$

∴ The number of persons getting wages below seventy is 423.6 thousands.

Hence, the number of persons getting wages between Rs. 60 and Rs. 70 is 423.6 –370 = 53.6 thousands.

Illustration 11. If L_x represents the numbers living at age x in a life table, find as accurately as the data will permit, L_x for values of x = 35, 42 and 47, give $L_{20} = 512$; $L_{30} = 439$; $L_{40} = 346$; $L_{50} = 243$.

Solution –

Age in years	Nos. of living	Δ^1	Δ^2	Δ^3
20 x_o	512 y_o	–73 Δ_o^1		
30 x_1	439 y_1	–93 Δ_1^1	–20 Δ_o^2	
40 x_2	346 y_2	–103 Δ_2^1	–10 Δ_1^2	+10 Δ_o^3
50 x_3	243 y_3			

$$x = \frac{\text{Age of Interpolation} - \text{Age of origin}}{\text{Age distance between adjoining ages}}$$

$$= \text{(i) } \frac{35-20}{10} = 1.5 \qquad \text{(ii) } \frac{42-20}{10} = 2.2 \qquad \text{(iii) } \frac{47-20}{10} = 2.7$$

Newton's Formulae –

$$y_x = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1 \times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1 \times 2 \times 3}\Delta_0^3$$

(i) $$y_x = 512 + 1.5(-73) + \frac{1.5(1.5-1)}{1 \times 2} \times -20 + \frac{1.5(1.5-1)(1.5-2)}{1 \times 2 \times 3} \times 10$$

$$= 512 - 109.5 - 7.50 - .625 = 394 \ (L_{35})$$

(ii) $$L_{42} = 512 + 2.2(-73) + \frac{2.2(2.2-1)}{1 \times 2} \times -20 + \frac{2.2(2.2-1)(2.2-2)}{1 \times 2 \times 3} \times 10$$

$$= 512 - 160.6 - 26.4 + .88 = 326$$

(iii) $$L_{47} = 512 + 2.7(-73) + \frac{2.7(2.7-1)}{1 \times 2} \times -20 + \frac{2.7(2.7-1)(2.7-2)}{1 \times 2 \times 3} \times 10$$

$$= 512 - 197.1 - 45.9 + 5.355 = 374$$

Illustration 12. Extrapolate the population of a town for 1986 from the following data about its population during the previous four censuses–(1986 की जनसंख्या की बाह्यगणना कीजिए)।

Census year	:	1951, 1961, 1971, 1981
Population in thousands	:	473, 468, 454, 484

Solution –

Year		Population		Δ^1	Δ^2	Δ^3
1951	x_0	473	y_0	$-5\ \Delta_0^1$	$-9\ \Delta_0^2$	$+53\ \Delta_0^3$
1961	x_1	468	y_1	$-14\ \Delta_1^1$	$+44\ \Delta_1^2$	
1971	x_2	454	y_2	$+30\ \Delta_2^1$		
1981	x_3	484	y_3			

$$x = \frac{\text{Year of Extrapolation} - \text{Year of Origin}}{\text{Time distance}} = \frac{1986-1951}{10} = \frac{35}{10} = 3\cdot5$$

Newton's formula —

$$y_x = y_0 + x\,\Delta_0^1 + \frac{x(x-1)}{1\times2} + \Delta_0^2 + \frac{x(x-1)(x-2)}{1\times2\times3}\Delta_0^3$$

$$= 437 + (3.5\times-5) + \frac{3.5\,(3.5-1)}{1\times2}\times-9 + \frac{3.5\,(3.5-1)\,(3.5-2)}{1\times2\times3}\times53$$

$$= 473 - 17.5 - 39.375 + 115.937 \qquad = 532.062\,.$$

Illustration 13. (Newton's formula for having a group)

Given the population of a town in three ten-year age-groups. It is required to estimate the number of people in 20–25 and 25–30 age-groups separately.

Age group	:	10–20	20–30	30–40
Population	:	14,500 f_0,	15420 f_1 ,	13,500 f_2

Solution– ऐसे प्रश्नों को हल करने को लिये न्यूटन ने एक सूत्र का प्रतिपादन किया है। इस सूत्र के लिये उस दस वर्षीय आयु वर्ग में जनसंख्या मालूम होनी चाहिए जो उस वर्ग के पहिले और बाद में पड़ते हैं जिसे दो पाँच वर्षीय आयु वर्गों में विभाजित करता हो (f_1a और f_1b)। यह सूत्र इस प्रकार है –

$$f_1\,a - \frac{1}{2}\left\{ f_1 + \frac{1}{8}(f_0 - f_2) \right\} \qquad f_1\,b = f_1 - f_1\,a$$

where, f_0 = number in the preceding ten year age group.

f_1 = number in the ten year age group to be broken into two equal groups of f_1a and f_1b.

f_2 = number in the succeeding ten year age group.

Hence, $$f_1a\ (20-25) = \frac{1}{2}\left\{ 15420 + \frac{1}{8}(14500 - 13500) \right\}$$

$$= \frac{1}{2}(15420 + 125) = 7773 \qquad f_1b\ (25-30) = f_1 - f_1a$$

$$= 15420 - 7773 = 7647\,.$$

6.9 न्यूटन गॉस (अग्रगामी) रीति (Newton-Gauss Forward Method)

इस रीति का प्रयोग वहाँ किया जाता है जहाँ आंतरगणन की जाने वाली संख्या श्रेणी के मध्य में हो और श्रेणी का आकार मूल्य का अन्तर या वर्ग–विस्तार समान हो। इसका सूत्र है —

$$y_x = y_0 + x\,\Delta^1 y_0 + \frac{x(x-1)}{1\times2}\Delta^2 y_{-1} + \frac{(x+1)(x-1)}{1\times2\times3}\Delta^3 y_{-1} + \frac{(x+1)(x-1)(x-2)}{1\times2\times3\times4}\Delta^4 y_{-2}\ldots\ldots$$

NOTES

इस सूत्र में 'y_0' सामान्यतः उस पद का मूल्य होता है जो आन्तरगणन के पद के पूर्व होता है। 'y_0' के पहले आने वाली संख्याओं के सामने क्रमशः 'y_{-1}' y_2' आदि लिखा जाता है और 'y_0' के बाद वाली संख्याओं के सामने 'y_1, y_2, y_3' आदि लिखा जाता है। इसमें

$$x = \frac{\text{(आंतरगणन वाली मद) – (आंतरगणन वाली मद के बिल्कुल पहले वाली मद)}}{\text{वर्गान्तर}}$$

उपरोक्त सूत्र को इस प्रकार भी लिखा जा सकता है —

$$y_x = y_0 + x\Delta^1 y_0 + \frac{x(-1)}{1\times 2}\Delta^2 y - 1 + \frac{(x+1)x(x-1)}{1\times 2\times 3}\Delta^3 y_{-1} + \frac{(x+1)x(x-1)(x-2)}{1\times 2\times 3\times 4}\Delta^4 y_{-2}\ldots\ldots$$

Illustration 14. Estimate the value of 'y' if x is 3.75 from the following table–(निम्न सारणी से 'y' का मूल्य निकालिए यदि x = 3.75 है)

x	:	2.5	3.0	3.5	4.0	4.5	5.0,
y	:	24.145,	22.043	20.225	18.644	17.262	16.047

Solution –

x	y		$\Delta^1 y$		$\Delta^2 y$		$\Delta^3 y$		$\Delta^4 y$	ह
2.5	24.145	y_{-2}	–2.102	$\Delta^1 y_{-2}$						
					+.248	$\Delta^2 y_{-2}$				
3.0	22.043	y_{-1}	–1.818	$\Delta^1 y_{-1}$						
				$\Delta^1 y_0$	+.237	$\Delta^2 y_{-1}$	–.047	$\Delta^3 y_{-2}$		
3.5	20.225	y_0	–1.581	$\Delta^1 y_1$		$\Delta^2 y_0$				
					+.199		–.038	$\Delta^3 y_{-1}$	+.009	$\Delta^4 y_{-2}$
4.0	18.644	y_1	–1.381	$\Delta^1 y_2$		$\Delta^2 y_1$				
					+.167		–.032	$\Delta^3 y_0$	+.006	$\Delta^4 y_{-1}$
4.5	17.262	y_2	–1.215							
5.0	16.017	y_3								

$$x = \frac{3.75 - 3.5}{4.5 - 3.5} = \frac{.25}{.5} = .5$$

The Formula is : $y_x = y_0 + x\Delta^1 y_0 + \frac{x(x-1)}{1\times 2}\Delta^2 y_{-1} + \frac{(x+1)x(x-1)}{1\times 2\times 3}\Delta^3 y_{-1} + \frac{(x+1)x(x-1)(x-2)}{1\times 2\times 3\times 4}\Delta^4 y_{-2}$

Substituting the values, we get

$$y_x = 20.225 + (.5\times -1.58) + \left\{\frac{.5(.5-1)}{1\times 2}\times .237\right\} + \left\{\frac{(.5+1).5(.5-1)}{1\times 2\times 3}\right\}\times(-.038) + \left\{\frac{(.5+1).5(.5-1)(.5-2)}{1\times 2\times 3\times 4}\times .009\right\}$$

$$= 20.225 - .7905 - .029625 + .002375 + .0001206 = 19.407 \text{ Approx.}$$

Hence the value of y when x is 3.75 = 19.407.

6.10 न्यूटन–गॉस विलोमगामी रीति (Newton–Gauss Backward Method)

इस रीति प्रयोग उस समय किया जाता है जबकि आन्तरगणन की जाने वाली संख्या श्रेणी में अन्तिम भाग में पड़ती है। इस सूत्र का प्रयोग तभी किया जाना उचित होता है जबकि मूल्यों का अन्तर या वर्ग विस्तार समान हो। इसके लिए निम्न सूत्र का प्रयोग होता है—

$$y_x = y_o - x\,\Delta^1 y_{-1} + \frac{x\,(x+1)}{1\times 2}\,\Delta^2 y_{-1} - \frac{x\,(x+1)\,x\,(x-1)}{1\times 2\times 3}\,\Delta^3 y_{-2}$$
$$+ \frac{x\,(x+1)\,x\,(x-1)\,(x-2)}{1\times 2\times 3\times 4}\,\Delta^4 y_{-3}$$

इस सूत्र में 'y_0' आंतरगणन वाली संख्या के बाद की संख्या का मूल्य होता है। 'y_0' के पहले वाली संख्याओं के सामने क्रमशः y_{-1}, y_{-2}' आदि लिखा जाता है और 'y_0' के बाद वाली संख्या के सामने 'y_1' आदि (यदि आवश्यक हो तो)। इसमें,

$$x = \frac{\text{(आंतरगणन वाली मद के बिल्कुल बाद वाली मद) – (आंतरगणन वाली मद)}}{\text{वर्गान्तर}}$$

Illustration 15. Estimate the value of y when x = 22 from the following data (निम्न लिखित सूचना से y का मूल्य निकालिए यदि x = 22 है)

x –	5	10	15	20	25	30
y –	25	32	40	47	55	64

Solution –

x		y		Δ		Δ^2		Δ^3		Δ^4	
5	x_{-4}	25	y_{-4}	7	$\Delta^1 y_{-4}$						
10	x_{-3}	32	y_{-2}	8	$\Delta^1 y_{-3}$	+1	$\Delta^2 y_{-4}$	–2	$\Delta^3 y_{-4}$		$\Delta^4 y_{-4}$
15	x_{-2}	40	y_{-2}	7	$\Delta^1 y_{-2}$	–1	$\Delta^2 y_{-3}$	+2	$\Delta^3 y_{-3}$	4	$\Delta^4 y_{-3}$
20	x_{-1}	47	y_{-1}	8	$\Delta^1 y_{-1}$	+1	$\Delta^2 y_{-2}$	0	$\Delta^3 y_{-2}$	–2	
25	x_o	55	y_o	9	$\Delta^1 y_o$	+1	$\Delta^2 y_{-1}$				
30	x_{+1}	64	y_{+1}								

$$x = \frac{25-23}{25-20} = \frac{2}{5} = .4$$

$$y_x = y_o - x\,\Delta^1 y^{-1} - \frac{x\,(x+1)}{1\times 2}\,\Delta^2 y^{-1} - \frac{x\,(x+1)\,x\,(x-1)}{1\times 2\times 3}\,\Delta^3 y^{-2}$$

$$= 55 - (.4 \times 8) - \frac{.4\,(.4+1)}{2} \times 1 - \frac{(.4+1)\,(.4)\,(.4-1)}{6}$$

$$= 55 - 3.2 + .28 - 0 = 52.08$$

6.11 स्टर्लिंग का सूत्र (Sterling's Formula)

न्यूटन गॉस का सूत्र की भाँति, यह सूत्र भी उस समय उपयुक्त होता है जबकि आन्तरगणन वाली संख्या श्रेणी के मध्य में हो और वर्ग विस्तार समान हो। इसका सूत्र निम्न प्रकार है –

$$y_x = y_o + x\,\frac{\Delta_1 y_{-1} + \Delta_1 y_o}{2} + \frac{x_2}{2}\,\Delta^2 y_{-1} + \frac{x\,(x_2 - 1_2)}{6}$$
$$\times \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \frac{x_2}{24}\,(x^2 - 1^2)\,\Delta^4 y_{-2} \ldots\ldots$$

Illustration 16. – Use the sterling's formula, interpolate the value of y when x = 15, (स्टर्लिंग सूत्र का प्रयोग कर y का मूल्य निकालिए यदि x = 15 है) :

x –	10	12	14	16	18	20
y –	50	60	75	95	120	150

NOTES

Solution –

x		y		Δ^1		Δ^2		Δ^3	
10	x_{-2}	50	y_2	10	$\Delta^1 y_{-2}$	5	$\Delta^2 y_{-2}$	0	$\Delta^3 y_{-3}$
12	x_{-1}	60	y_1	15	$\Delta^1 y_{-1}$	5	$\Delta^2 y_{-1}$	0	$\Delta^3 y_{-1}$
14	x_o	75	y_o	20	$\Delta^1 y_o$	5	$\Delta^2 y_o$	0	$\Delta^3 y_o$
16	x_{+1}	95	y_{+1}	25	$\Delta^1 y_{+1}$	5	$\Delta^2 y_{+1}$		
18	x_{+2}	120	y_{+2}	30	$\Delta^1 y_{+2}$				
20	x_{+3}	150	y_{+3}						

$$x = \frac{15-14}{2} = \frac{1}{2} = .5$$

$$y_x = y_o + \frac{x\,\Delta^1 y_{-1} + \Delta^1 y_o}{2} + \frac{x^2}{2}\Delta^2 y_{-1} + \frac{x\,(x^2-1^2)}{6} \times \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2}$$

$$= 75 + .5\,\frac{15+20}{2} + \frac{(5)^2}{2} \times 5 + \frac{.5\,(.5^2-1^2)}{2} \times \frac{0+0}{2}$$

$$= 75 + 8.75 + .625 + 0 = 84.375\,.$$

6.12 लैग्रेंज की रीति (Lagrange's Method)

इस रीति का प्रतिपादन फ्रांस के प्रसिद्ध गणितज्ञ श्री लैग्रेंज महोदय ने किया। यह रीति सर्वश्रेष्ठ समझी जाती है। इसके द्वारा किसी श्रेणी में आन्तरगणन और ब्राह्यगणन सम्भव है। यह आवश्यक नहीं है कि श्रेणी समान वर्ग विस्तार की ही हो। ऊपर जिन रीतियों का उल्लेख किया गया है उनकी अपनी सीमायें हैं किन्तु इस रीति के लिये ऐसी कोई सीमा नहीं है। इसका सूत्र निम्न प्रकार का है –

$$y_x = y_o \frac{(x-x_1)(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_o-x_1)(x_o-x_2)(x_o-x_3)(x_o-x_n)}$$

$$+ y_1 \frac{(x-x_o)(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_1\quad x_o)(x_1\quad x_2)(x_1\quad x_3)\ldots(x_1\quad x_n)}$$

$$+ y_2 \frac{(x-x_o)(x-x_1)(x-x_3)\ldots(x-x_n)}{(x_2-x_o)(x_2-x_1)(x_2-x_3)\ldots(x_2-x_n)}$$

$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$+ y_n \frac{(x-x_o)(x-x_1)(x-x_2)\ldots(x-x_n-1)}{(x_n-x_o)(x_n-x_1)(x_n-x_2)\ldots(x_n-x_n-1)}$$

where, x = the figure for which interpolation is to be done x_o, x_1, x_2, x_3 etc. and y_o, y_1, y_2, y_3 etc are the variables of x and y series respectively.

Illustration 17 – In the following table 'h' is the height above sea-level and 'p' the barometric pressure

Calculate 'p' when h = 5280

h = 0,	4763,	6942,	10593
p = 27,	25,	23,	20

Solution –

Height above Sea level		Barometric Pressure	
0	x_o	27	y_o
4763	x_1	25	y_1
6942	x_2	23	y_2
10593	x_3	20	y_3

x = 5280, the height above sea level for which the barometric pressure has to be interpolated.

Applying the Lagrange's Formula, we get

$$y_x = 27\frac{(5280-4763)(5280-6942)(5280-10593)}{(0-4763)(0-6942)(0-10593)} +$$

$$25\frac{(5280-0)(5280-6942)(5280-10593)}{(4763-0)(4763-6942)(4763-10593)} +$$

$$23\frac{(5280-0)(5280-4763)(5280-10593)}{(6942-0)(6942-4763)(6942-10593)} +$$

$$20\frac{(5280-0)(5280-4763)(5280-6942)}{(10593-0)(10593-4763)(10593-6942)}$$

$$\text{or } y_x = 27\frac{(517)(-1662)(-5313)}{(-4763)(-6942)(-10593)} + 25\frac{(5280)(-1662)(-5313)}{(4763)(-2179)(-5830)} +$$

$$23\frac{(5280)(517)(-5313)}{(6942)(2179)(-3651)} + 20\frac{(5280)(517)(-1662)}{(10593)(5830)(3651)} +$$

$$\text{or } y_x = -.353 + 19.23 + 6.34 - .402 = 24.8.$$

∴ The estimated Barometric pressure when the height above sea level is 5280 = 24.8

Illustration 18 – The observed values of a function are respectively 168, 120 72 and 63 at four positions 3, 7, 9 and 10 of the independent variable. What is the best estimate you can give of the value of the function at the position 6 of the independent variable.

Solution –

Independent variable		Function	
3	x_o	168	y_o
6	x	?	
7	x_1	120	y_1
9	x_2	72	y_2
10	x_3	63	y_3

Applying Lagrange's Formula, we have

$$y_x = 1168\frac{(6-7)(6-9)(6-10)}{(3-7)(3-9)(3-10)} +$$

$$120\frac{(6-3)(6-9)(6-10)}{(7-3)(7-9)(7-10)} + 72\frac{(6-3)(6-7)(6-10)}{(9-3)(9-7)(10-9)}$$

$$+ 63\frac{(6-3)(6-7)(6-9)}{(10-3)(10-7)(10-9)}$$

$$\text{or} \quad y_x = 168\frac{(-1)(-3)(-4)}{(-4)(-6)(-7)} + 120\frac{(3)(-3)(-4)}{(4)(-2)(-3)}$$

$$+ 72\frac{(3)(-1)(-4)}{(6)(2)(-1)} + 63\frac{(3)(-1)(-3)}{(7)(3)(1)}$$

$$\text{or} \quad y_x = 12 + 180 - 72 + 27 \quad \text{or} \quad y_x = 147$$

Illustration 19 – The following are the marks obtained by 65 students in Statistics in 1991 M.Com (Final) examination of a certain university (एक विश्वविद्यालय की एम.कॉम. परीक्षा में सांख्यिकी विषय में 65 छात्रों के प्राप्तांक इस प्रकार हैं) :–

Marks out of 100 More than	25	36	45	55	70
No. of Students	65	63	40	18	7

NOTES

Find out the number of students who secured first class marks in Statistics. प्रथम श्रेणी पाने वालों की संख्या बतलाइये।

Solution – first class marks means 60% or more. So we have to interpolate the number of students who secured 60 marks or more than 60 marks. As the marks are not given in fractions, we may interpolate from the given data, the number of students who secured more than 59 marks.

Marks out of 100	**Number of Students**
Marks more than 25% x_o	65 y_o
Marks more than 36% x_1	63 y_1
Marks more than 45% x_2	40 y_2
Marks more than 55% x_3	18 y_3
Marks more than 70% x_4	7 y_4

x = 59, corresponding to which the value of y has to be interpolated.

Applying Lagrange's Formula, we get.

$$y_x = 65\frac{(59-36)(59-45)(59-55)(59-70)}{(25-36)(25-45)(25-55)(25-70)}$$
$$+\ 63\frac{(59-25)(59-45)(59-55)(59-70)}{(36-25)(36-45)(36-55)(36-70)}$$
$$+\ 40\frac{(59-25)(59-36)(59-55)(59-70)}{(45-25)(45-36)(45-55)(45-70)}$$
$$+\ 18\frac{(59-25)(59-36)(59-45)(59-70)}{(55-25)(55-36)(55-45)(55-70)}$$
$$+\ 7\frac{(59-25)(59-36)(59-45)(59-55)}{(70-25)(70-36)(70-45)(70-55)}$$

or $$y_x = 65\frac{(23)(14)(4)(-11)}{(-11)(-20)(-30)(-45)} + 63\frac{(34)(14)(4)(-11)}{(11)(-9)(-19)(-34)}$$
$$+\ 40\frac{(34)(23)(4)(-11)}{(20)(9)(-10)(-25)} + 18\frac{(34)(23)(14)(-11)}{(30)(19)(10)(-15)}$$
$$+\ 18\frac{(34)(23)(14)(-11)}{(30)(19)(10)(-15)} + 7\frac{(34)(23)(14)(4)}{(45)(34)(25)(15)}$$

or $y_x = -3.1 + 20.6 - 30.6 + 25.4 + .5$ or $y_x = 12.8$ or 13.

Therefore, the estimated number of students who secured First Class-marks in Statistics in 13.

Illustration 20 – Estimate the population for 1997 from the following table. (निम्न सारणी से 1997 की जनसंख्या का अनुमान लगाइये) :–

Year	1971 x_o	1981 x_1	1986 x_2	1991 x_3
Population in lakhs	22 y_o	31 y_1	37 y_2	44 y_3

x = 1947

Applying Lagrange's Formula, we get

$$y_x = 22\frac{(1997-1981)(1997-1986)(1997-1991)}{(1997-1981)(1971-1986)(1971-1991)}$$
$$+\ 31\frac{(1997-1971)(1997-1986)(1997-1991)}{(1981-1971)(1981-1986)(1981-1991)}$$

$$+ 37\frac{(1997-1971)(1997-1981)(1997-1991)}{(1986-1971)(1986-1981)(1986-1991)}$$

$$+ 44\frac{(1997-1971)(1997-1981)(1997-1986)}{(1991-1971)(1991-1981)(1991-1986)}$$

or $$y_x = 22\frac{(16)(11)(6)}{(-10)(-15)(-20)} + 31\frac{(26)(11)(6)}{(10)(-5)(-10)}$$

$$+ 37\frac{(26)(16)(6)}{(15)(5)(-5)} + 44\frac{(26)(16)(11)}{(20)(10)(5)}$$

or $$y_x = -7.74 + 106.39 - 246.27 + 201.34 \qquad y_x = 53.72 \text{ Lakhs}$$

Therefore, the population for 1997 will be 53.69 lakhs.

Comparative Study of the Methods of Interpolation

Method	Cases in which applicable	Remarks
(i) Graphic Method	All (General assumptions of no abnormality and similar rate of change.)	(i) Gives idea of the function. (ii) Easy in application. (iii) Requires graphical skill. Different persons may find different results.
(ii) Fitting the parabolic Curve	All	May require lengthy calculations.
(iii) Lagrange's Formula	All	Though easier to apply but sometimes may prove cumbersome.
(iv) Binominal Expansion Method	Values to be found, together with those given are placed at equal intervals of the independent variable	Perhaps the easiest to apply
(v) Newton's Formula	Given values are place at equal intervals. Value to be found at unequal Intervel.	Fairly easy.

बोध प्रश्न

1. बिन्दुरेखीय रीति पर टिप्पणी कीजिए।

....

....

....

....

2. द्विपद रीति की अवस्थाएँ लिखिए।

....

....

....

....

NOTES

6.13 सारांश

जब कई वर्षों के भूतकालीन आँकड़े प्राप्त होते हैं एवं किसी एक वर्ष के आँकड़े उपलब्ध न हो तो ऐसी परिस्थिति में उपलब्ध आँकड़ों द्वारा आन्तरगणन विधि के द्वारा बीच के आँकड़े प्राप्त किए जा सकते हैं। उपलब्ध आँकड़ों के माध्य से भविष्य के समय के लिए भी आँकड़ों का अनुमान लगाया जा सकता है।

6.14 शब्दावली या शब्दकुंजी

6.15 अभ्यास प्रश्न

दीर्घ उत्तरीय प्रश्न (Important Questions)

1. 'आन्तरगणन' एवं 'बाह्यगणन' में अन्तर स्पष्ट कीजिए। सांख्यिकीय अध्ययन में उनकी आवश्यकता और उपयोगिता का संक्षिप्त विवेचन कीजिए।
   Explain clearly the difference between 'Interpolation' and 'Extrapolation'. Discuss briefly their necessity and usefulness in statistical studies.
2. आन्तरगणन करने में किन मान्यताओं का ध्यान रखना पड़ता है ?
   What are the assumptions underlying interpolation ?
3. एक व्यापारी के लिए आन्तरगणन और बाह्यगणन की उपयोगिता का विवेचन कीजिए। आन्तरगणन की विभिन्न रीतियाँ कौन-कौन सी हैं ?
   Describe the usefulness of Interpolation and Extrapolation for a businessman. What are the various methods of Interpolation ?
4. आन्तरगणन क्या है ? यह पूर्वानुभाव से किस प्रकार भिन्न है ? उदाहरण देकर स्पष्ट कीजिए कि ये वाणिज्य और उद्योग में किस प्रकार उपयोगी होते हैं ?
   What is interpolation ? How does it differ from forecasting ? Explain giving examples, how these are useful in commerce and industry ?
5. एक श्रेणी में अज्ञात समंक को अनुमानित करने की किन्हीं तीन रीतियों का संक्षिप्त वर्णन कीजिए और उन परिस्थितियों व मान्यताओं का भी उल्लेख कीजिए जिनमें प्रत्येक रीति उचित रूप से प्रयुक्त की जा सकती है ?
   Describe briefly any three methods of estimating a missing figure in a series giving the circumstances in which each of them can be most suitably used and the assumptions made therein.

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. अन्तरगणन से आपका क्या आशय है ?
   What do you mean by interpolation.
2. बाह्यगणन से आप क्या समझते हैं ?
   What do you understand by Extrapolation.
3. अन्तरगणन की कौन-कौन सी विधियाँ हैं ?
   What are the different methods for interpolation.
4. अन्तरगणन एवं बाह्यगणन में अन्तर स्पष्ट कीजिये।
   Explain clearly the difference between interpolation and extrapolation.
5. अन्तरगणन की आवश्यकता का संक्षिप्त विवेचन कीजिये।
   Discuss briefly the interpolation of necessity.

NOTES

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. अन्तरगणन एवं बाह्यगणन की रीतियाँ होती हैं-

   (अ) दो (ब) तीन (स) चार (द) पाँच

2. अन्तरगणन की रीति कौन सी नहीं है-

   (अ) कार्ल पियर्सन (ब) द्विपद विस्तार (स) न्यूटन (द) लाग्रेंज

3. जब दो अज्ञात मूल्यों का अन्तरगणन करना हो तो समीकरण होते हैं-

   (अ) तीन (ब) चार (स) दो (द) एक

4. लैग्रेंज महोदय किस देश के निवासी थे-

   (अ) फ्रांस (ब) अमेरिका (स) रूस (द) ब्रिटेन

उत्तर- 1. (अ), 2. (अ), 3. (स), 4. (अ)

6.16 व्यावहारिक प्रश्न

1. Interpolate the missing figure by the graphic method.

Year :	2000	2001	2002	2003	2004	2005	2006	2007	2008
No. of Factories :	27	29	32	57	112	130	?	140	146

2. A Life Assurance Company advertises the following immediate life annuities per 100 pounds paid :

Age in Year :	50	60	65	70
Annuity in (lbs) :	6/5	8/6	9/18	12/2

By graphic mean (or otherwise) estimate the corresponding values for age 62 and 67 years.

[Convert Annuities in shillings and then plot the figures – (i) 8/16, (ii) 10/12]

3. The following gives the sales of a concern for the following years :

Year :	1990	1995	2000	2005	2010
Sales of Cloth (Lakhs of metres) :	125	163	204	238	282

Assuming the conditions to be the same, estimate the sales for the year 1940.
[380 Lakhs of yds.]

4. Interpolate the missing figure in the following table with the help of a suitable formula :

x :	2000	2001	2002	2003	2004	2005	2006
y :	1331	1728	2197	?	3375	4096	4913

[Binomial Method – 2744]

5. Estimate the missing in the following table :

x :	1	2	3	4	5	6	7
y :	2	4	8	?	32	64	128

[Binomial Method – 16]

6. Discuss briefly the nature and suitability of the chief methods of interpolation. Estimate the annual sales of cloth for 2008 from the following data :

Year :	2005	2006	2007	2008	2009
Sales of Cloth :	250	285	328	?	444

[Binomial Method – 380.5]

NOTES

7. The following values are given in a table :

x	:	1	2	3	4	5
y	:	2,16,000	2,26,981	?	2,50,047	2,62,144

Using any suitable algebraic method, find the value of y for x = 3

[y = 2,38,328]

8. State carefully any two methods of interpolation and use one of them to find the value of y when x = 5, from the following set of values

x	:	2	3	4	6	7
y	:	1	5	13	61	125

[Binomial Method – 28.6 or 29]

9. Obtain by graphical method the missing figure in the following table of 5% values of chi–square and check the result by an algebraic method :

Degrees of freedom	:	2	3	4	5	6	7
Chi–square (5%)	:	5.99	7.82	9.49	?	12.59	14.07

(Binomial Method – 11.07)

10. The following table gives the age–composition of post–graduate students on an Evening College:

Age (years)	:	15–17	17–19	19–21	21–23	23–25	25–27
No. of Students	:	18	25	33	?	56	59

Interpolate the most likely number of students of 21–23 age–group.

[Binomial Method– 44.4 or 44]

11. Find out the missing values in the following :

x	:	5	10	15	20	25	30
y	:	7	?	13	15	?	25

(Binomial Method –10.87; 18.47)

12. Estimate the production for the years 1995 and 2005 with the help of the following table :

Year	:	1980	1985	1990	1995	2000	2005	2010
Production (Lakh tons)	:	200	220	260	?	350	?	430

[Binomial Method–306 and 390 Lakh tons]

13. The table below gives the expectations of life at different ages:

Ages	40	45	60
Expectations	30	36	48

Find the expectation at the age of 55.

[Ans. 41.25 Parabolic Method]

14. From the following table, estimate by using Newton's formula the premium payable at the age of 22 years.

Premium Table to secure Rs. 100

Age Years	Premium
20	25
25	28
30	32
35	37
40	43.50
45	52.25

[Ans. 26 Approx.]

15. From the following data, find the number of students who obtained less than 45 marks:

Marks	No. of Students
30–40	31
40–50	42
50–60	51
60–70	35
70–80	31

(Ans. 48, Newton's Method)

16. Following are the marks obtained by 492 candidates in a certain examination:

	Candidates
Not more than 40 marks	212
" " 45 "	296
" " 50 "	368
" " 55 "	429
" " 60 "	460
" " 65 "	481
" " 70 "	490
" " 75 "	492

Find out the number of candidates who secured more than 42 but not more than 45 marks.

[Ans. 40, Newton's Method]

17. Estimate by Newton's method of interpolation the expectation of life at age 22 from the following data stating the assumption underlying the formula used by you :

Age in years	10	15	20	25	30	35
Expectation of life (years)	35.4	32.2	29.1	26.0	23.1	20.4

[Ans. 27.88 yrs.]

18. Using Newton's formula for interpolation estimate the population of Agra for the years 1976 :

Year	1961	1971	1981	1991	2001
Population	98,754	1,32,285	1,68,076	1,95,690	2,46,050

[Ans. 1, 51, 519]

19. Estimate the number of persons having incomes between 1,000 and 1,500 in the table given below in the groups A and B.

Income (in Rs)	No.of persons (Group A)	No. of persons (Group B)
Below 500	6,000	5,000
500–1000	4,250	4,500
1000–2000	3,600	4,800
2000–3000	1,500	2,200
3000–4000	650	1,500

[Ans. Group A = 2141, Group B = 2844, Newton's Method]

NOTES

20. Find out from the following data the number of workers earning Rs. 24 or more but less than Rs. 25.

Earning less than	20	25	30	35	40
Number of workers	296	599	804	918	966

[Ans. 53, Newton's Method]

21. State Newton's formula for interpolation for equal intervals and the assumption underlying it? Use it to find the annual net premium at the age of 25 from the table given gelow :

Age	20	24	28	32
Annual Net Premium	0.01427	0.01581	0.01772	0.01996

[Ans. 0.01625]

22. The following table relates to weekly earnings of a number of workers in a certain manufacturing concern :

Earnings (in Rs) Up to Rs.	10	20	30	40	50	60
No. of workers	50	150	300	500	700	800

Interpolate the number of workers earning between Rs. 25 and Rs. 35.

(178, Newton's Method)

23. The following figures showing the relationship between the amount of manure used and the yield per acre of rice were supplied by a research institute after several years of experimentation:

Amount of manure per acre (Tons)	0	10	20	30
Yield of rice per acre (Mds.)	10	15	17	18

Obtain by using a suitable interpolation formula, the yield of rice per acre corresponding to 5 and 15 tons of manure per acre.

[Ans. 13 and 16.25 mds., Newton's Method]

24. Interpolate the probable number of persons earning between Rs. 20 and Rs. 25 from the following figures :

Income in Rs. Less then	10	10–20	20–30	30–40	40–50
No. of persons	150	170	200	250	180

[Ans. 91, Newton' s Method]

25. Use some appropriate interpolation method and reconstruct the following frequency table with the intervals halved :

X	0–2	2–4	4–6
Frequency	35	52	84

[Hint –Convert into cumulative frequency table and apply Newton's Method. Ans. = Frequency for 0–1, 2–3 and 5–6 would be 21, 22 and 38 respectively.]

26. The length of the day was 12 hours on March 19th, 14hrs. on April 18th and 15hrs. 40 minutes on May 18th required an approximate value of (a) the length of the day on May 3rd (b) the mean length of the day during the period March 19th May 18th.

[Ans. (a) 14 Hrs. 53 minutes (b) 13.9 Hrs.]

27. Given $\log_{10} 654 = 2.8156$; $\log_{10} 658 = 2.8182$

$\log_{10} 659 = 2.8189$; $\log_{10} 661 = 2.8202$

Find Log $_{10}$ 656

[Ans. 2.8168]

28. The following table gives the population of Indore City at the time of last six censuses–

1901	99,880
1911	57,285
1921	1,07,948
1931	1,47,100
1941	2,03,695
1951	3,10,859

Estimate the population for 1961.

[Ans. 531. 859 after interpolation for 1911 as well, because there is sudden fall in that year.]

29. Extrapolate the population of a town for 1966 from following data about its population during the previous four censues;

Census year	Population in thousands
1931	473
1941	468
1951	454
1961	484

[Ans. 532. App. Lagrange's Method]

30. Explain the methods used in forecasting the growth the population. The population of a certain town is given below in the years mentioned. Estimate it for the year 1957.

Year	1921	1931	1941	1947	1951
Population	22,000	27,000	34,000	39,000	42,000

[Ans. 44560, Lagrange' s Method]

31. In the following table 'h' is the height above sea level and p the barometric pressure. Calculate p when h = 5280.

h = 0.	4763,	6942,	10,593
p = 27.	25,	23,	20.

[Ans. 24.8 Lagrange' s formula]

32. The observed value of function are respectively 168, 120, 72 & 63 at the four positions 3, 7, 9 & and 10 of the independent variable. What is the best estimate you can given of the function at the position 6 of the independent variable ?

[Ans 147. Lagrange' s formula method]

33. Determine by Lagrange's formula the percentage number of criminals under years 35 :–

Age under 25 years	25	30	40	50
% number of criminals.	52	67.3	84.1	94.4

[Ans. 77.4%] or

Determine the percentage number for criminals under 35 Yrs.

Age	0–25	25–30	30–40	40–50
% No. of Criminals	52.0	15.3	16.8	10.3

[Ans. 77.4%]

34. Estimate the probable number of persons earning between Rs. 30 and Rs. 40 from the following:

Income (Rs.)	15–20	20–30	30–45	45–55	55–70
No. of persons	73	97	110	180	140

[Ans. 53, Lagrange' s Method]

35. What do you understand by Interpolation and Extrapolation? What are their use? The following table gives the normal weights of babies during the first twelve months of life:

Age (in months)	0	2	5	8	10	12
Weights (in Ibs.)	15/2	41/4	15	16	18	21

Estimate the weight of the baby at the age of 7 months.

[Ans. 15.66 Ibs., Lagrange's Method]

36. Estimate the population of 'City of Taj' for the year 1939 from the following census figures:

Year	1901	1911	1921	1931	1941	1951
Population ('000)	12	15	20	27	39	52

[Ans. 33.125, Newton Gauss Backward Method]

37. Interpolate the missing figures :

Year	1951	1952	1953	1954	1955	1956	1957	1958
Sales in ('000 Rs.)	10%	18	25	%	35	37	45	

[1952 = 11.3; 1955 = 31.4]

38. The age of mothers and number of average children born per mother are given in the following table, Find out the unknown value :-

Age of months (yrs.)	15-19	20-24	25-29	30-34	35-39	40-44
No. of children born	0-7	2.1	3.5	?	5.7	5.8

[4.8]

39. Estimate the expectation of life at the age of 22 from the following data-

Age (yrs)	15	20	25	30	35
Expectation of life (yr)	32.2	29.1	26.0	23.1	20.4

[27.84 yrs] [M.A. Agra, 1984]

40. Find out the value of $\sin 52^0$, using a suitable method of interpolation, if-

$\sin 45^\circ = 0.7071$; $\sin 55^\circ = 0.8192$; $\sin 50^\circ = 0.8660$ $\sin 52^\circ = 0.7880$.

41. Find out the missing value-

x	0	3	6	9
y	30	?	80	120

[50]

अध्याय–7 गुण सम्बन्ध (ASSOCIATION OF ATTRIBUTES)

इकाई की रूपरेखा

7.0 उद्देश्य

इस इकाई का अध्ययन करने के पश्चात आप इस योग्य हो सकेंगे कि–

1. गुणों के आधार पर वर्गीकरण के प्रकारों का ज्ञान होगा।
2. गुण संबंध में प्रयुक्त विभिन्न वर्ग चिन्हों व संकेत अक्षरों का अर्थ जान सकेंगे।
3. आंशिक गुण संबंध एवं भ्रमात्मक गुण संबंध का ज्ञान होगा।

7.1 प्रस्तावना

वर्गीकरण और सारणीयन (Classification and Tabulation) के अध्याय में यह बतलाया जा चुका है कि वर्गीकरण की दो मुख्य रीतियाँ होती हैं —

(1) **वर्गान्तरों के अनुसार वर्गीकरण** (Classification according to class–intervals)– इसका विस्तृत विवेचन पीछे किया जा चुका है।

(2) **गुणों के अनुसार वर्गीकरण** (Classification according to qualities or attributes)

इस वर्गीकरण में गुणों के अनुसार वर्ग (classes) बनाये जाते हैं अर्थात् आँकड़ों के गुणों को प्रधानता दी जाती है। संकलित आँकड़ों को निश्चित गुणों के आधार पर बाँट दिया जाता है। गुण कई प्रकार के होते हैं, जैसे– शिक्षा, लिंग, स्वास्थ्य, वजन, आदत आदि।

गुणों के आधार पर वर्गीकरण भी दो प्रकार का हो सकता है।

NOTES

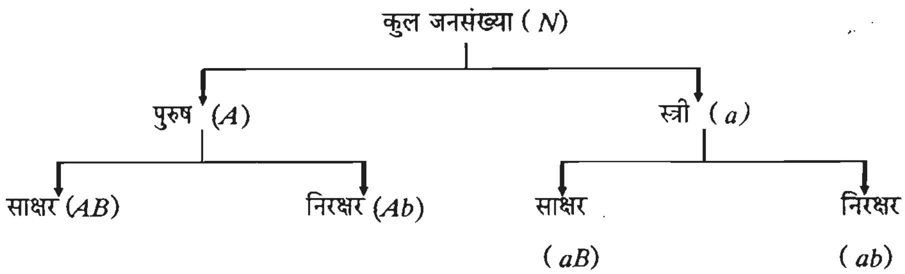
7.2 द्वंद्व-भाजन वर्गीकरण (Classification according to Dichotomy)

यह साधारण वर्गीकरण (Simple Classification) भी कहलाता है । इसमें आँकड़ों को पूर्व-निश्चित गुण की उपस्थिति या अनुपस्थिति के अनुसार वर्गीकृत किया जाता है । इसमें केवल दो वर्ग (classes) बनते हैं-

(1) निश्चित गुण की उपस्थिति वाले पद ।

(2) निश्चित गुण की अनुपस्थिति वाले पद ।

दोनों वर्गों का योग सम्पूर्ण के बराबर होता है । लिखने और बोलने की सुविधा की दृष्टि से किसी निश्चित गुण की उपस्थिति या अनुपस्थिति के आधार पर किये गये वर्गों को सरल संकेत अक्षरों (Simple notations) में व्यक्त करना अधिक उपयुक्त रहता है । सामान्यत: गुणों की उपस्थिति के लिए बड़े अक्षरों (Capital Letters) A, B, C, आदि और उनकी अनुपस्थिति के लिये छोटे अक्षरों (Small Letters) a, b, c, आदि का प्रयोग किया जाता है । यदि किसी स्थान की जनसंख्या को लिंग और शिक्षा के आधार पर विभाजित करना हो तो उसे निम्न प्रकार से व्यक्त किया जायेगा-



7.3 संकेत अक्षरों और वर्ग चिन्हों के अर्थ (Meaning of Notations and Class-symbols)

(1) **धनात्मक गुण** (Positive Attribute)— पूर्व-निश्चित गुण की उपस्थिति को संकेत अक्षर 'A', 'B', 'C' आदि से व्यक्त किया जाता है । यदि वर्ग में इस गुण की उपस्थिति होती है तो उसे धनात्मक या अनुलोम वर्ग (Positive Class) कहते हैं । ये वर्ग बड़े अक्षरों से बनते हैं, जैसे-A, AB, ABC आदि ।

(2) **ऋणात्मक गुण** (Negative Attribute)— पूर्व-निश्चित गुण की अनुपस्थिति को संकेत अक्षर 'a', 'b', 'c', आदि से व्यक्त किया जाता है । यदि वर्ग में विलोम या ऋणात्मक गुण हो तो वह ऋणात्मक या विलोम वर्ग (Negative Class) कहलाता है । ये वर्ग छोटे अक्षरों द्वारा बनाये जाते हैं जैसे –a, ab, abc आदि ।

(3) **गुणों का संयोग** (Combination of Attributes)—विभिन्न गुण जब एक साथ मिलते हैं तो उसे गुणों का संयोग कहते हैं । ऊपर सारणी में दिये उदाहरण में पुरुष (A) और साक्षर (B) के लिये (AB), पुरुष निरक्षर के लिये (Ab); स्त्री-साक्षर के लिये (aB) और स्त्री निरक्षर के लिये (ab) का प्रयोग किया गया है । यह गुणों का संयोग है । जो वर्ग बड़े और छोटे अक्षरों से मिले-जुले बनते हैं, वे विपरीत वर्ग (Contrary class) कहलाते हैं ।

(4) **वर्ग-आवृत्ति** (Class-Frequency)— किसी वर्ग में कितने पद हैं उसे पद-संख्या या वर्ग-आवृत्ति कहते हैं और यह बतलाने के लिये उस अक्षर (Letter) को कोष्ठक में लिखते हैं । इसका अर्थ उस गुण को रखने वालों की संख्या होता है । अत: जो गुण या गुणों के संयोग कोष्ठकों के भीतर प्रस्तुत किये जाते हैं, वे वर्ग की आवृत्ति को प्रकट करते हैं ।

(5) **वर्गों की संख्या** (Number of Classes) — वर्गों की संख्या गुणों की संख्या पर निर्भर होती है । एक गुण के आधार पर वर्गीकरण करने से 2 वर्ग दो गुणों के आधार पर 4 वर्ग और तीन गुणों के आधार पर 8 वर्ग बनते हैं । यह निम्न तरीके से स्पष्ट किया जा सकता है-

(A) एक गुण के वर्गीकरण के आधार पर प्रथम क्रम की आवृत्ति (Frequency of the first order)-

(i) (A) (ii) (a)

(B) दो गुणों के वर्गीकरण के आधार पर द्वितीय क्रम की आवृत्ति (Frequency of the second order)

(i) (AB) (iii) (aB)

(ii) (Ab) (iv) (ab)

(C) तीन गुणों के वर्गीकरण के आधार पर तृतीय क्रम की आवृत्ति (Frequency of the third order)—

(i) (ABC) (v) (Abc)

(ii) (ABc) (vi) (abC)

(iii) (AbC) (vii) (aBc)

(iv) (aBC) (viii) (abc)

(6) समग्र (Universe) — पदों की कुल संख्या समग्र कहलाती है । इसको व्यक्त करने के लिये संकेत अक्षर '*N*' का प्रयोग किया जाता है । विभिन्न गुणों की संख्या के आधार पर इसे निम्न प्रकार से ज्ञात किया जाता है–

(i) **एक गुण के आधार पर वर्गीकरण में –**

$$N = (A) + (a)$$

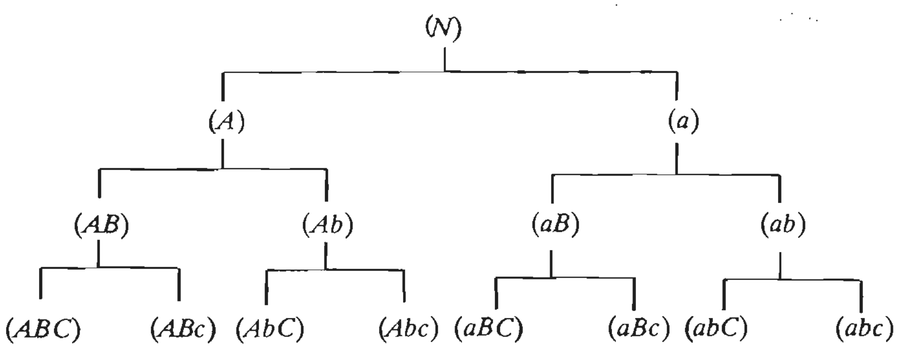
(ii) **दो गुणों के आधार पर वर्गीकरण में–**

$$N = (AB) + (Ab) + (aB) + (ab)$$

(iii) **तीन गुणों के आधार पर वर्गीकरण में–**

$$N = (ABC) + (AbC) + (ABc) + (Abc) + (aBC) + (aBc) + (abC) + (abc)$$

इसे चार्ट द्वारा निम्न रूप में प्रदर्शित किया जा सकता है–



अन्तिम वर्ग आवृत्तियों (Ultimate class frequencies) को उदाहरण द्वारा निम्न प्रकार समझाया जा सकता है –

Order 0	*N*		
Order 1	(A)	(B)	(C)
	(a)	(b)	(c)
Order 2	(AB)	(BC)	(AC)
	(Ab)	(Bc)	(aC)
	(aB)	(bC)	(Ac)
	(ab)	(bc)	(ac)
Order 3	(ABC)	(aBC)	
	(ABc)	(aBc)	
	(Abc)	(abC)	
	(AbC)	(abc)	

NOTES

दो गुणों के आधार पर किये गये वर्गों की आवृत्ति ज्ञात करने के लिए निम्न सारणी का प्रयोग भी किया जाता है–

A	*a*	*N*
(AB),	(*aB*),	(B)
(Ab),	(ab),	(b)
(A),	(a),	(N)

इस सारणी से यह स्पष्ट हो जाता है कि–

$$(B) = (AB) + (aB)$$
$$(A) = (AB) + (Ab)$$
$$(b) = (Ab) + (ab)$$
$$(a) = (aB) + (ab)$$
$$(N) = (A) + (a) \text{ or } (B) + (b)$$

इस प्रकार इस सारणी की सहायता से किन्हीं चार आवृत्तियों के ज्ञात होने से पाँचवीं आवृत्ति को जोड़कर या घटाकर मालूम किया जा सकता है—

$$(AB) = (A) - (Ab) \text{ or } (B) - (aB)$$
$$(Ab) = (A) - (AB) \text{ or } (b) - (ab)$$
$$(aB) = (B) - (AB) \text{ or } (a) - (ab)$$
$$(ab) = (a) - (aB) \text{ or } (b) - (Ab)$$
$$(N) = (AB) + (Ab) + (aB) + (ab)$$

इसी तरह तीन गुणों के आधार पर वर्गीकरण की दशा में भी आवृत्तियाँ ज्ञात की जा सकती हैं किन्तु इसके लिये ऊपर जैसी नौ खानों वाली सारणी नहीं बन सकती है। नीचे उदाहरणों द्वारा इसे ज्ञात किया जायेगा:-

7.4 अज्ञात वर्ग–आवृत्तियाँ निकालने के नियम (Rules to determine unknown class-frequencies)

एक समग्र (universe) में यदि कुछ वर्ग– आवृत्तियाँ मालूम हों तो शेष अज्ञात वर्ग– आवृत्तियाँ जोड़–घटाकर निर्धारित की जा सकती हैं। उदाहरण के लिए, यदि एक समग्र (N) का वर्गीकरण लिंग और साक्षरता के अनुसार किया जावे तो कुल संख्या (N) = पुरुषों (A) तथा स्त्रियों (a) को जोड़कर मालूम की जा सकती है, अर्थात् $N = (A) + (a)$। कुल संख्या (N) में से पुरुषों की संख्या (A) घटाकर स्त्रियों की संख्या (a) निकाली जा सकती है, जैसे $-(a) = N - (A)$, इसी प्रकार पुरुषों की संख्या होगी $(A) = N - (a)$

अब साक्षरता के आधार पर वर्गीकरण करने से कुल संख्या (N) – साक्षरों (B) तथा निरक्षरों (b) की संख्या को जोड़कर मालूम की जा सकती है। यह निम्नलिखित संकेत अक्षरों के रूप में व्यक्त किये जा सकते हैं–

$N = (B) + (b) =$ (कुल संख्या)

$(B) = N - (b) =$ (साक्षरों की संख्या)

$(b) = N - (B) =$ (निरक्षरों की संख्या)

इसी भाँति पुरुषों (A) की संख्या – साक्षर पुरुषों (AB) तथा निरक्षर पुरुषों (Ab) के योग से मालूम की जा सकती है। अर्थात्–

$(A) = (AB) + (Ab)$	$(B) = (AB) + (aB)$
$\therefore (AB) = (A) - (Ab)$	$\therefore (AB) = (B) - (aB)$
$\therefore (Ab) = (A) - (AB)$	$\therefore (aB) = (B) - (AB)$

	(a)	$= (aB) + (ab)$		(b)	$= (Ab) + (ab)$
$\therefore$	(aB)	$= (a) - (ab)$	$\therefore$	(ab)	$= (b) - (ab)$
$\therefore$	(ab)	$= (a) - (aB)$	$\therefore$	(ab)	$= (b) - (Ab)$

साक्षर पुरुष व निरक्षर पुरुष तथा साक्षर स्त्रियों तथा निरक्षर स्त्रियों की संख्याओं का योग कुल समग्र (N) के बराबर होगा, जैसे–

$$N = (AB) + (Ab) + (aB) + (ab)$$

इसी प्रकार तीन गुणों वाले वर्गीकरण में भी जोड़–घटाकर अज्ञात वर्ग–आवृत्तियाँ ज्ञात की जा सकती हैं। इसके लिए प्रमुख रूप से निम्नलिखित नियमों का पालन किया जाता है–

(1) किसी एक क्रम की वर्ग–आवृत्ति, उससे ऊँचे क्रम की सुनिश्चित वर्ग–आवृत्तियों को जोड़कर मालूम की जा सकती है। इस प्रकार प्रथम क्रम (order) की कोई वर्ग–आवृत्ति (A), द्वितीय क्रम (second order) की ऐसी दो आवृत्तियों को जोड़कर मालूम की जा सकती है जिनमें वह संकेताक्षर (A) विद्यमान है, अर्थात् (A) = (AB) + (Ab)। इसके लिए यह ध्यान रखना है कि जिस संकेताक्षर की आवृत्ति ज्ञात करनी है वह दोनों बार एक सा लिखा जायेगा तथा दूसरा संकेताक्षर (B) या (C) एक बार धनात्मक तथा दूसरी बार ऋणात्मक गुण के रूप में लिखा जावेगा। इस प्रकार —

$(A) = (AB) + (Ab) \quad \therefore \quad (AB) = (A) - (Ab); \quad (Ab) = (A) - (AB)$

$(a) = (aB) + (ab) \quad \therefore \quad (aB) = (a) - (ab); \quad (ab) = (a) - (aB)$

$(B) = (AB) + (aB) \quad \therefore \quad (AB) = (B) - (aB); \quad (aB) = (B) - (AB)$

$(b) = (Ab) + (ab) \quad \therefore \quad (Ab) = (b) - (ab); \quad (ab) = (b) - (Ab)$

यह नियम तीन गुणों वाली वर्ग–आवृत्तियों की स्थिति में भी लागू किया जा सकता है, जैसे—

यह नियम तीन गुणों वाली वर्ग-आवृत्तियों की स्थिति में भी लागू किया जा सकता है, जैसे–

$(AB) = (ABC) + (ABc); \quad \therefore \quad (ABC) = (AB) - (ABc);$
$(ABc) = (AB) - (ABC)$

$(Ab) = (AbC) + (Abc); \quad \therefore \quad (AbC) = (Ab) - (Abc);$
$(Abc) = (Ab) - (AbC)$

$(aB) = (aBC) + (aBc); \quad \therefore \quad (aBC) = (aB) - (aBc);$
$(aBc) = (aB) - (aBC)$

$(ab) = (abC) + (abc); \quad \therefore \quad (abC) = (ab) - (abc);$
$(abc) = (ab) - (abC)$

(2) तीन गुणों वाले समग्र में, प्रथम क्रम (First order) की एक वर्ग आवृत्ति (A) तीसरे क्रम की उन चार वर्ग आवृत्तियों का योग होती है जिनमें संकेताक्षर (A) पाया जाता है, जैसे–

$$(A) = (ABC) + (ABc) + (AbC) + (Abc)$$

$$(B) = (ABC) + (ABc) + (aBC) + (aBc)$$

$$(C) = (ABC) + (aBC) + (AbC) + (abC)$$

यह नियम प्रथम नियम का ही उपप्रमेय (Corollary) है।

(3) यदि सभी अंतिम वर्ग–आवृत्तियाँ मालूम हों तो शेष सभी कोटियों की वर्ग आवृत्तियाँ मालूम की जा सकती हैं और सभी अंतिम वर्ग–आवृत्तियों को जोड़कर कुल समग्र (N) की संख्या ज्ञात की जा सकती है।

दो गुण (Two attributes) —

$N = (AB) + (Ab) + (aB) + (ab)$

$(A) = (AB) + (Ab); \quad (a) = (aB) + (ab)$

$(B) = (AB) + (aB); \quad (b) = (Ab) + (ab)$

NOTES

तीन गुण (Three attributes)

N = (ABC) + (ABc) + (AbC) + (aBC) + (Abc) + (abC) (aBc) + (abc)

(AB) = (ABC) + (ABc) ; (BC) = (ABC) + (aBC)

(AC) = (ABC) + (AbC)

(Ab) = (AbC) + (Abc); (bC) = (AbC) + (aBC)

(aC) = (aBc) + (Abc)

(aB) = (aBC) + (aBc); (bC) = (AbC) + (abC);

(aC) = (aBC) + (abC)

(ab) = (abC) + (abc); (bc) = (Abc) + (abc);

(ac) = (aBc) + (abc)

(A) = (AB) + (Ab) (B) = (BC) + (Bc);

(C) = (AC) + (aC)

या (A) = (AC) + (Ac) ; (B) = (AB) + (aB);

(C) = (BC) + (bC)

Illustration 1. Given the following ultimate class–frequencies, find the frequencies of positive and negative classes and the total number of observations– (अंतिम वर्ग आवृत्तियों से **धनात्मक तथा ऋणात्मक वर्गों की** आवृत्तियाँ तथा कुल संख्या ज्ञात कीजिए):–

(AB) = 125 (Ab) = 60

(aB) = 100 (ab) = 35

Solution :–

(AB)	(aB)	(B)
(Ab)	(ab)	(b)
(A)	(a)	(N)

N = (A) + (a)

= (AB) + (Ab) + (aB) + (ab)

= 125 + 60 + 100 + 35

= 320

(b) = (Ab) + (ab)

= 60 + 35

= 95

(A) = (AB) + (Ab)

= 125 + 60

= 185

(B) = (AB) + (aB)

= 125 + 100

= 225

(a) = (aB) + (ab)

= 100 + 35

= 135

Illustration 2. Given the following table, calculate the frequencies of the remaining classes (निम्नलिखित सारणी से शेष वर्गों की आवृत्तियाँ निकालिए) :–

(A) = 25 (AB) = 15

(B) = 20 (N) = 50

Solution– इस उदाहरण में हमें निम्न गुणों के मूल्य ज्ञात करने हैं–

(a) ; (b) ; (Ab) ; (aB) ; (ab)

Now

$$\begin{aligned}(a) &= (N) - (A) \\ &= 50 - 25 \\ &= 25\end{aligned} \qquad \begin{aligned}(b) &= (N) - (B) \\ &= 50 - 20 \\ &= 30\end{aligned}$$

$$\begin{aligned}(Ab) &= (A) - (AB) \\ &= 25 - 15 \\ &= 10\end{aligned} \qquad \begin{aligned}(aB) &= (B) - (AB) \\ &= 20 - 15 \\ &= 5\end{aligned}$$

$$\begin{aligned}(ab) &= (a) - (aB) \\ &= (N) - (A) - \{(B) - (AB)\} \\ &= 50 - 25 - \{20 - 15\} \\ &= 50 - 25 - 5 = 20\end{aligned}$$

Illustration 3. A number of labourers in a factory were examined for the presence or absence of certain defects of which three chief descriptions were noted–

A – Physical weakness

B – Nervous signs.

C – Mental dullness.

Given the following ultimate frequencies, find the frequencies of the positive classes including the whole number of observations– N – (निम्नलिखित अंतिम वर्ग आवृत्तियों से धनात्मक वर्गों की आवृत्तियाँ तथा कुल संख्या निकालिए) :–

(ABC) = 75	(aBC) = 98
(ABc) = 310	(aBc) = 702
(AbC) = 106	(abC) = 74
(Abc) = 489	(abc) = 8415

Solution — The positive classes including the whole number of observations are –

(N), (A), (B), (C), (AB), (AC), (BC), (ABC)

Now

$$\begin{aligned}(N) &= (ABC) + (ABc) + (AbC) + (aBC) + (abC) + (aBc) + (Abc) + (abc) \\ &= 75 + 310 + 106 + 98 + 74 + 702 + 489 + 8415 = 10{,}269.\end{aligned}$$

$$\begin{aligned}(A) &= (ABC) + (ABc) + (AbC) + (Abc) \\ &= 75 + 310 + 106 + 489 = 980.\end{aligned}$$

$$\begin{aligned}(B) &= (ABC) + (ABc) + (aBC) + (aBc) \\ &= 75 + 310 + 98 + 702 = 1185.\end{aligned}$$

$$\begin{aligned}(C) &= (ABC) + (AbC) + (aBC) + (abC) \\ &= 75 + 106 + 98 + 74 = 353.\end{aligned}$$

$$\begin{aligned}(AB) &= (ABC) + (ABc) \\ &= 75 + 310 = 385.\end{aligned}$$

$$\begin{aligned}(AC) &= (ABC) + (AbC) \\ &= 75 + 106 \\ &= 181.\end{aligned} \qquad \begin{aligned}(BC) &= (ABC) + (aBC) \\ &= + 75 + 98 \\ &= 173.\end{aligned}$$

Note : The frequency of any first order class is given by the total of second order frequencies, the class–symbols for which contain the same letter; similarly the frequency of any second order class can be given by the total of third order frequencies and so on.

Illustration 4. Given the following frequencies of the positive classes, find out all the class–frequencies (धनात्मक वर्गों की निम्न आवृत्तियों से शेष सभी वर्ग की आवृत्तियाँ निकालिए) :—

$$N = 12{,}000 \qquad (AB) = 453$$
$$(A) = 977 \qquad (AC) = 284$$
$$(B) = 1{,}185 \qquad (BC) = 250$$
$$(C) = 596 \qquad (ABC) = 127$$

Solution– The frequencies for the following classes are to be found out–

(a), (b), (c), (Ab), (aB), (ab), (Ac), (aC), (ac), (Bc), (bC), (bc), (Abc), (ABc), (aBC), (AbC), (abC), (aBc), (abC), (abc).

Now,

$$(a) = N - (A) = 12{,}000 - 977 = 11{,}023. \qquad (b) = N - (B) = 12{,}000 - 1{,}185 = 10{,}815. \qquad (c) = N - (C) = 12{,}000 - 596 = 11{,}404.$$

$$(ABc) = (AB) - (ABC) = 453 - 127 = 326. \qquad (AbC) = (AC) - (ABC) = 284 - 127 = 157.$$

$$(aBC) = (BC) - (ABC) = 250 - 127 = 123.$$

$$(Abc) = (Ab) - (AbC) = (A) - (AB) - (AbC) = 977 - 453 - 157 = 367.$$

$$(aBc) = (aB) - (aBC) = (B) - (AB) - (aBC) = 1{,}185 - 453 - 123 = 609.$$

$$(abC) = (bC) - (AbC) = (C) - (BC) - (AbC) = 596 - 250 - 157 = 189$$

$$(Ab) = (A) - (AB) = 977 - 453 = 524. \qquad (aB) = (B) - (AB) = 1{,}185 - 453 = 732.$$

$$(Ac) = (A) - (AC) = 977 - 284 = 693. \qquad (aC) = (C) - (AC) = 596 - 284 = 321.$$

$$(Bc) = (B) - (BC) = 1{,}185 - 250 = 935. \qquad (bC) = (C) - (BC) = 596 - 250 = 346.$$

$$(abc) = (ab) - (abC)$$
$$= (b) - (Ab) - (abC)$$
$$= N - (B) - \{(A) - (AB)\} - (abC)$$

$= 12{,}000 - 1{,}185 - \{977 - 453\} - (189)$
$= 12{,}000 - 1{,}185 - 524 - 189$
$= 12{,}000 - 1{,}898$
$= 10{,}102.$

(ab) $= (abC) + (abc)$
$= 189 + 10{,}102$
$= 10{,}291.$

(ac) $= (aBc) + (abc)$
$= 609 + 10{,}102$
$= 10{,}711$

(bc) $= (Abc) + (abc)$
$= 367 + 10{,}102$
$= 10{,}469$

Illustration 5— Arrange the following figures in a nine–square table and find out the unknown class–frequencies. (निम्न समंकों को 9 खाने वाली सारणी में जमाकर अज्ञात वर्ग आवृत्तियाँ निकालिए) :

Intelligent fathers and their intelligent sons	=	248
Dull fathers and their intelligent sons	=	92
Intelligent fathers and their dull sons	=	81
Dull fathers and their dull sons	=	579

Solution – Taking, Intelligence in fathers = A, and intelligence in sons = B, we have—

(AB) = 248; (aB) = 92; (Ab) = 81; (ab) = 579.

These figures will be arranged in the nine–square table as under —

	A	a	N
B	(AB) 248	(aB) 92	(B) 340
b	(Ab) 81	(ab) 579	(b) 660
N	(A) 329	(a) 671	(N) 1,000

Intelligent fathers	(A)	=	329
Dull fathers	(a)	=	671
Intelligent sons	(B)	=	340
Dull sons	(b)	=	660
Total number	(N)	=	1,000

Illustration 6– After an air–raid, survey was made of the local hospital specially reserved for those injured in the raid. These were 1200 beds in all. 600 patients were found to have fractured their arms, legs and skulls as a result of the bombing. There were 50 patients who had no injury over their body but were admitted only to provide speedy recovery from the shock of the raid. Patients with fractured arms has a majority of 192 over those with no arm–injury. The number of patients who had escaped head injury was 270. There were 36 patients with a fractured arm but had fortunately no injury in their legs. Similarly there were 204 patients with a fractured leg but without any head injury. The majority of those with a fractured leg over those without injury to their legs was 620.

Make an analysis of the injury of patients according to the injuries received by them.

NOTES

Solution – Let 'A' stand for injury to arm and 'a' non–injury.

'B' stand for injury to leg and 'b' for non–injury.

'C' stand for injury to skull and 'c' for non injury.

Then the above data can be stated as–

$$
\begin{aligned}
N &= 1200 & (c) &= 270\\
(ABC) &= 600 & (Ab) &= 36\\
(abc) &= 50 & (Bc) &= 204\\
(A)-(a) &= 192 & (B)-(b) &= 620.
\end{aligned}
$$

We have to find out the remaining ultimate class–frequencies (ABc), (AbC), (Abc), (aBC), (aBc), (abC). In order to find them out, we have to first calculate the frequencies of positive classes–

(A), (B), (C), (AB), (AC), (BC).

(i) First order frequencies–

$$
\begin{aligned}
N = (A) + (a) &= 1200\\
(A) - (a) &= 192\\
\hline
2\,(A) &= 1392\\
\text{Or}\quad (A) &= 696 \text{ (Number of persons with fractured arms.)}
\end{aligned}
$$

$$
\begin{aligned}
N = (B) + (b) &= 1200\\
(B) - (b) &= 620\\
\hline
2\,(B) &= 1820\\
\text{Or}\quad (B) &= 910 \text{ (Number of persons with fractured legs.)}\\
(C) &= N - (c)\\
&= 1200 - 270\\
&= 930. \text{ (Number of persons with fractured skulls.)}
\end{aligned}
$$

(ii) Second order frequencies –

$$
\begin{aligned}
(AB) &= (A) - (Ab)\\
&= 696 - 36\\
&= 660 \text{ (No. of persons with fractured arms and legs.)}\\
(BC) &= (B) - (Bc)\\
&= 10 - 204\\
&= 706 \text{ (No. of persons with fractured legs and skulls.)}\\
(AC) &= (C) - (aC)\\
&= (C) - \{(aBC) + (abC)\}\\
&= (C) - [(aBC) + \{(ab) - (abc)\}]\\
&= (C) - [(aBC) + \{(b) - (Ab) - (abc)\}]\\
&= (C) - [(aBC) + \{(N) - (B) - (Ab) - (abc)\}]\\
&= (C) - [\{(BC) - (ABC)\} + \{(N) - (B) - (Ab) - (ac)\}]\\
&= (C) - (BC) + (ABC) - (N) + (B) + (Ab) + (abc)\\
&= 930 - 706 + 600 - 1200 + 910 + 36 + 50\\
&= 620 \text{ (No. of persons with fractured arms and skulls.)}
\end{aligned}
$$

(iii) Ultimate class–frequencies –

$(ABc) = (AB) - (ABC)$
$= 660 - 600 = 60$ (No. of persons with fractured arms and legs but not skull.)

$(AbC) = (AC) - (ABC)$
$= 620 - 600 = 20$ (No. of persons with fractured arms and skull but not legs)

$(aBC) = (BC) - (ABC)$
$= 706 - 600 = 106$ (No. of persons with fractured legs and skull but not arms.)

$(aBc) = (aB) - (aBC)$
$= (B) - (AB) - (aBC)$
$= 910 - 660 - 106$
$= 144$ (No. of persons with fractured legs but not arms and skulls.)

$(Abc) = (Ab) - (AbC)$
$= (A) - (AB) - (AbC)$
$= 696 - 660 - 20$
$= 16$ (No. of persons with fractured arms but not legs and skulls.)

$(abC) = (bC) - (AbC)$
$= (C) - (BC) - (AbC)$
$= 930 - 706 - 20$
$= 204$ (No. of persons with fractured skulls but not arms and legs.)

Illustration 7. In a free vote in the House of Commons 600 members voted, 300 government members representing English Constituencies (including Welsh) voted in favour of the motion. 25 opposition members representing Scottish Constituencies voted against motion. The government majority among those who voted was 96. 135 of the members voting represented Scottish Constituencies; 18 government members voted against the motion. The motion was carried by 310 votes. 102 members of Scottish Constituency voted in favour of the motion.

Analyse the voting according to the Nationality of the Constituencies and party.

Solution: Denoting

(i) The government and opposition parties by 'A' and 'a';

(ii) Voting for and against the motion by 'B' and 'b';

(iii) English and Scottish members by 'C' and 'c'.

The given data are

$N = 600;\ (ABC) = 300;\ (abc) = 25;\ (A) - (a) = 96\ ;$

$(c) = 135;\ (Ab) = 18;\ (Bc) = 102$

$(B) - (b) = 310$

To find out, Ultimate class–frequencies –

(i) $(A) + (a) = 600$ (ii) $(B) + (b) = 600$

$(A) - (a) = 96$ $(B) - (b) = 310$

$\therefore\ 2(A) = 696$ $\therefore\ 2(B) = 910$

$\therefore\ (A) = 348.$ $\therefore\ (B) = 455$

(iii) $(C) = N - (c)$

$= 600 - 135$

$= 465.$

(iv) $(AB) = (A) - (Ab) = 348 - 18 = 330.$

(v) $(BC) = (B) - (Bc) = 455 - 102 = 353.$

(vi) We also have,

$(abc) = (bc) - (Abc) = (c) - (Bc) - [(Ab) - (AbC)]$

$= (c) - (Bc) - (A) + (AB) + (AC) - (ABC)$

$25 = 135 - 102 - 348 + 330 + (AC) - 300$

or $(AC) = 135 - 102 - 348 + 330 - 300 - 25$

or $(AC) = 310$

Now, the fruquencies of the Ultimate classes will be calculated as under–

(i) $(ABC) = 330$ (given)

(ii) $(ABc) = (AB) - (ABC) = 330 - 300 = 30$

(iii) $(AbC) = (AC) - (ABC) = 310 - 300 = 10$

(iv) $(aBC) = (BC) - (ABC) = 353 - 300 = 53$

(v) $(Abc) = (Ab) - (AbC) = 18 - 10 = 8$

(vi) $(aBc) = (aB) - (aBC)$

$= (B) - (AB) - (aBC) = 455 - 330 - 53 = 72$

(vii) $(abC) = (bC) - (AbC)$

$= (C) - (BC) - (AbC)$

$= 465 - 353 - 10 = 102$

(viii) $(abc) = 25$ (given)

Illustration 8. A University result consists of three attributes, rest of which is divisible into two parts. What are the different class frequencies obtainable?

At a competitive examination at which 600 graduates appeared, boys outnumbered girls by 96. Those qualifying for interview exceeded in number those failing to qualify by 310. The number of science graduate boys interviewed was 300, while among the Arts graduate girls there were 25 who failed to qualify for interview. Altogether there were only 135 Arts graduates and 33 among them failed to qualify. Boys who failed to qualify numbered 18. Find (a) the number of boys who qualified for inrerview, (b) the total number of science graduate boys appearing and (c) the number of science graduate girls who qualified.

Solution – Let

'A' stand for Boys and 'a' for girls.

'B' stand for those who qualified for interview and 'b' for those who failed to qualify.

'C' stand for science graduates and 'c' for arts graduates.

The given data are:

N = 600; (A) – (a) = 96; (B) – (b) = 310; (ABC) = 300

= (abc) = 25; (c) = 135; (bc) = 33; (Ab) = 18.

To find – (i) AB (ii) AC (iii) aBC.

Now,

(1) (A) – (a) = 96	(2) (B) – (b) = 310
(A) + (a) = 600	(B) + (b) = 600
∴ 2 (A) = 696	∴ 2 (B) = 910
∴ (A) = 248	∴ (B) = 455

(3) (C) = N – (c) = 600 – 135 = 465.

(i) (AB) = (A) – (Ab)

= 348 – 18 = 330.

(ii) (AC) = (ABC) + (AbC)

= (ABC) + (Ab) – (Abc)

= (ABC) + [(Ab) – {(bc) – (abc) }]

= (ABC) + (Ab) – (bc) + (abc)

= 300 + 18 – 33 + 25 = 310.

(iii) (aBC) = (BC) – (ABC) = (C) – (bC) – (ABC)

= (C) – [(b) – (bc)] – (ABC)

= 465 – [145 – 33] – 300 = 53.

Illustration 9. Measurements are made on a thousand husbands and a thousand wives. If the measurements of the husbands exceed the measurements of the wives in 789 cases for one measurement, in 741 cases for another, and in 690 cases for both measurements. In how many cases will both measurements on the wife exceed the measurements on the husband.

Solution — Let 'A' stand for the cases in which measurements of husbands exceed the measurements by the wives in one measurement and 'B' for those in the other measurement.

N = 1000; (A) = 789; (B) = 741; (AB) = 690.

We have to find out the frequencies of (ab)

(ab) = (a) – (aB) = (N) – (A) – (B) + AB

= 1000 – 789 – 741 + 690 = 160

Thus the number of cases in which both measurements on the wife exceed the measurements on the husbands is 160.

NOTES

Illustration 10. There were 400 students in the B.Com. class of a university. Their results in the various terminal examinations are given below–

I	Terminal		passed	180
II	„	„	passed	140
III	„	„	passed	180

60 passed in all terminals;

80 failed in all the three;

400 passed in the Ist and IInd terminals and failed in the IIIrd;

70 failed in the Ist and IInd terminals and passed in the IIIrd;

Find out how many students passed atleast two examinations.

Solution – Let, Success in Ist Terminal be denoted by 'A' and failure by 'a'.

Success in IInd Terminal be denoted by 'B' and failure by 'b'.

Success in IIIrd Terminal be denoted by 'C' and failure by 'c'.

Then, the given data are :

$$N = 400 \qquad (ABC) = 60$$
$$(A) = 180 \qquad (abc) = 80$$
$$(B) = 140 \qquad (ABc) = 40$$
$$(C) = 180 \qquad (abC) = 70$$

We have to find out the values of,

$$(ABC) + (ABc) + (AbC) + (aBC)$$

Now.
$$\begin{aligned}(aBC) &= (aC) - (abC)\\ &= (C) - (AC) - (abC)\\ &= (C) - \{(ABC) + (AbC)\} - (abC)\\ &= (C) - (ABC) - (AbC) - (abC)\end{aligned}$$

$$(AbC) = (AC) - (ABC) \qquad = (ABC) + (AbC) - (ABC)$$

Hence, $(aBC) + (AbC)$
$$= (C) - (ABC) - (AbC) - (abC) + (ABC) + (AbC\} - (ABC)$$
$$= (C) - (ABC) - (abC) = 180 - 60 - 70 = 50$$

Hence, $(ABC) + (ABc) + (aBC) + (AbC)$
$$= 60 + 40 + 50 \qquad = 150$$

$\therefore$ The number of students passing atleast in two examinations is 150.

This illustration can also be solved in the following manner—

Now, $(aBC) + (ABC) + (AbC) + (abC) = C$

$$(aBC) + (AbC) = (C) - (ABC) - (abC)$$
$$= 180 - 60 - 70 = 50$$

$$(ABC) + (ABc) + (aBC) + (AbC)$$
$$60 + 40 + 50 = 150$$

7.5 समंकों की संगति (Consistency of Data)

ऊपर यह बतलाया जा चुका है कि संकलित सामग्री का वर्गीकरण किसी भी समस्या के अनुरूप गुण के आधार पर किया जाता है। यह सम्भव है कि वर्गीकृत आँकड़ों में अशुद्धि हो जाये और उनकी जाँच आवश्यक हो। ये अशुद्धियाँ दो तरह की हो सकती हैं —

(1) **योग तो ठीक हो किन्तु आवृत्ति में अन्तर पड़ जाये** – कभी-कभी ऐसा होता है कि वर्गीकरण करते समय वर्गों की आवृत्तियों का योग (जोड़) तो ठीक हो किन्तु आवृत्तियों में अन्तर हो। उदाहरणार्थ, किसी गाँव की जनसंख्या 1000 है जिसमें 200 साक्षर और 800 निरक्षर हैं लेकिन वे क्रमशः 240 और 760 लिख दिये जायें तो यहाँ समंकों में संगति (Consistency) है, अतः ऊपर से अशुद्धि का पता नहीं लग पाता है।

(2) **योग में गलती** – योग में गलती होने पर भी अशुद्धि होती है। वर्गों की आवृत्तियों का योग भी गलत हो जाये और उस कारण कुल योग भी–जैसे उपरोक्त उदाहरण में साक्षर 270 और निरक्षर 770 लिख दिये जायें और कुल योग 1040 हो जाये। यह स्पष्टतः अशुद्ध लगता है और यहाँ समंकों की संगति भी नहीं है। यदि समंक असंगत होते हैं तो किसी भी एक वर्ग की आवृत्ति ऋणात्मक हो जाती है।

अतः वर्ग–आवृत्तियों की संगति के लिये एक आवश्यक शर्त यह है कि किसी भी अन्तिम (Ultimate) वर्ग–आवृत्ति का मूल्य ऋणात्मक न हो। अतः यदि हमें दी हुई वर्ग–आवृत्तियों में संगति की जाँच करनी हो तो हमें केवल अन्तिम वर्ग आवृत्तियों को मालूम कर यह देख लेना चाहिये कि उनमें कोई ऋणात्मक तो नहीं है। यदि कोई ऋणात्मक हो तो समंक असंगत होते हैं। यदि ऐसा न हो तो समंकों में संगति होती है।

कई स्थानों पर वर्गों की आवृत्तियाँ धनात्मक होती हैं। ऐसे समय संगति की जाँच करने के लिए निम्न नियमों का प्रयोग किया जा सकता है–

(A) एक गुण के आधार पर वर्गीकरण में–

(i) $(A) \nless 0$, otherwise (A) will be negative.

(ii) $(A) \ngtr N$, otherwise (a) will be negative.

(B) दो गुणों के आधार पर वर्गीकरण में–

(iii) $(AB) \nless 0$, otherwise (AB) will be negative.

(iv) $(AB) \nless (A)$, otherwise (Ab) will be negative, because

$(B) = (AB) + (Ab)$

(v) $(AB) \ngtr (B)$ otherwise (aB) will be negative because

$(B) = (AB) + (aB)$

(vi) $(AB) \ngtr (A) + (B) - N$, otherwise (ab) will be negative because $(ab) = (a) - (aB)$

or $(ab) = N - (A) - (B) + (AB)$

or $-(AB) = N - (A) - (B) - (ab)$

or $(AB) = -N + (A) + (B) + (ab)$

or $(AB) = (A) + (B) - N + (ab)$

(C) तीन गुणों के आधार पर वर्गीकरण में –

(vii) $(ABC) \nless 0$. otherwise (ABC) will be negative.

(viii) $(ABC) \nless (AB) + (AC) - (A)$, otherwise (Abc) will be negative,

because $(Abc) = (Ab) - (AbC)$

or $(Abc) = (A) - (AB) - (AC) + (ABC)$

or $-(ABC) = +(A) - (AB) - (AC) - (Abc)$

$(ABC) = (AB) + (AC) - (A) + (Abc)$

NOTES

(ix) $(ABC) \nless (AB) + (BC) - (B)$, otherwise (aBc) will be negative.

(x) $(ABC) \nless (AC) + (BC) - (C)$, otherwise (abC) will be negative.

(xi) $(ABC) \ngtr (AB)$ otherwise (ABc) will be negative.

(xii) $(ABC) \ngtr (AC)$, otherwise (AbC) will be negative.

(xiii) $(ABC) \ngtr (BC)$ otherwise (aBC) will be negative.

(xiv) $(ABC) \ngtr (AB) + (BC) + (AC) - (A) - (B) - (C) + N$,

otherwise (abc) will be negative.

because $(abc) = (ab) - (abC)$

or $(abc) = (a) - (aB) - \{(bC) - (AbC)\}$

or $(abc) = (a) - (aB) - (bC) + (AbC)$

or $(abc) = N - (A) - (B) + (AB) - (C) + (BC) + (AC) - (ABC)$

or $(ACB) = N - (A) - (B) + (AB) - (C) + (BC) + (AC) - (abc)$

Now if (ABC) is more than the above, it is obvious that (abc) will be negative.

(xv) $(AB) + (AC) + (BC) \nless (A) + (B) + (C) - (N)$

(xvi) $(AB) + (AC) - (BC) \ngtr (A)$

(xvii) $(AB) - (AC) + (BC) \ngtr (B)$

(xviii) $-(AB) + (AC) + (BC) \ngtr (C)$

उपरोक्त में चिन्ह, '$\ngtr$' ≠ should not be greater than (से अधिक नहीं होना चाहिए)

'$\nless$' = should not be less than (से कम नहीं होना चाहिए)

Illustration 11. – The following data are given in a report –

$(N) = 1000, (AB) = 200, (Ab) = 350, (aB) = 500$

Show that there must be a misprint or mistake of some sort.

Solution – There is nothing obviously wrong with the figures. Yet, they are certainly inconsistent.

$$
\begin{aligned}
(ab) &= (a) - (aB) &&= N - (A) - (aB) \\
&= N - \{(AB) + (Ab)\} - (aB) &&= \{200 + 350\} - 500 \\
&= 1000 - 550 - 500 = -50
\end{aligned}
$$

क्योंकि वर्ग आवृत्ति ऋणात्मक नहीं हो सकती है, इसलिए दिये हुए समंकों के किसी प्रकार की अशुद्धि अवश्य होनी चाहिए।

Illustration 12. – If $(A) = 50, (B) = 60, (C) = 80, (AB) = 35, (AC) = 45$ and $(BC) = 42$, find the greatest and least possible value of (ABC)

Solution –

$$
\begin{aligned}
&(ABC) \nless 0 \\
&(ABC) \nless (AB) + (AC) - (A) &&= 35 + 45 - 50 = 30 \\
&(ABC) \nless (AB) + (BC) - (B) &&= 35 + 42 - 60 = 17 \\
&(ABC) \nless (BC) + (AC) - (C) &&= 42 + 45 - 80 = 7
\end{aligned}
$$

Therefore, the least value of $(ABC) = 30$

$(ABC) \ngtr (AB) = 35$

$$(ABC) \ngtr (AC) = 45$$

$$(ABC) \ngtr (BC) = 42$$

Hence the greatest value of (ABC) = 45

Illustration 13. – If in a collection of house actually invaded by small pox, 70 percent of the inhabitants are attacked and 85% have been vaccinated. What is the lowest percentage of the vaccinated that must have been attacked.

Solution – Writing 'A' to denote inhabitants invaded by small pox, 'B' to denote those vaccinated, the data are :

$$(A) = 70,\ (B) = 85,\ N = 100$$

First we find the lowest value of (AB),

$$(AB) \nless 0$$

$$\begin{aligned} (AB) &\nless (A) + (B) - N \\ &\nless 70 + 85 - 100 \\ &\nless 55 \end{aligned}$$

$\therefore$ The lowest value of (AB) = 55

Hence the lowest percentage of inhabitants vaccinated which have been attacked

$$= \frac{(AB)}{(B)} \times 100 \qquad = \frac{55}{85} \times 100$$

$$= 64.7\%$$

Illustration 14. – In a very hotly fought battle 70% atleast of the combatants lost an eye, 75% atleast lost an ear, 80% atleast lost an arm and 85% atleast lost a leg. How many atleast must have lost all four.

Solution – Let (N) = 100; (A), (B), (C) and (D) denote respectively losing an eye, an arm and a leg. Then (A) = 70, (B) = 75, (C) = 80, (D) = 85.

We have to find out (ABCD)

$$\begin{aligned} (ABCD) &\nless (A) + (B) + (C) + (D) - (n - 1)N \\ &\nless (A) + (B) + (C) + (D) - (4 - 1)N \\ &\nless 70 + 75 + 80 + 85 - 3 \times 100 \\ &\nless 310 - 300 \\ &\nless 10 \end{aligned}$$

Hence atleast 10% lost all the four.

Illustration 15. – To investigate the association between eye-colour of husband and eye-colour of wife, the following data are available –

Husbands with light eyes and wives with not light eyes = 414

Husbands with not light eyes and wives with light eyes = 260

Husbands with not light eyes and wives with not light eyes = 238

Husbands with light eyes = 400

Do you find any inconsistency in the data ?

NOTES

Solution – Denoting 'A' husbands with light eyes.

'a' husbands with not light eyes.

'B' wives with light eyes.

'b' wives with not light eyes

The given data are –

$$(Ab) = 414,\ (aB) = 260,\ (ab) = 238 \text{ and } (A) = 400$$

From the above

$$(AB) = (A) - (Ab) = 400 - 414 = -14$$

Hence, the given data are inconsistent.

Illustration 16. – The following is a summary of the statistical features of a census of ration – cards.

Item No.	Category	Total No. of Cards.
1.	The whole census	1000
2.	Permanent residents	510
3.	Males	490
4.	Consumers of rice	427
5.	Permanent male residents	189
6.	Consumers of rice among permanent residents	140
7.	Male consuming rice	97

Show that the entry against item No. 7 is inconsistent with the entries against the previous items viz. 1, 2, 3, 4, 5 and 6 taken together.

Solution – Denote permanent residents by 'A', males by 'B' consumers of rice by 'C'. The given data then are –

$N = 1000$, $(A) = 510$, $(B) = 490$, $(C) = 427$, $(AB) = 189$, $(AC) = 140$, $(BC) = 97$. Now according to the conditions of consistency.

$$AB + AC + BC < (A) + (B) + (C) - N$$

$$189 + 140 + 97 < 510 + 490 + 427 - 1000$$

$$426 < 427, \text{ which is not true.}$$

Hence, there is some inconsistency in the data.

or

$$(BC) \nless (A) + (B) + (C) - N - (AB) - (AC)$$

$$\nless 510 + 490 + 427 - 1000 - 189 - 140$$

$$98.$$

The given value of (BC) is 97 which is less than 98. Hence it can be said that the entry against item No. 7 is inconsistent with the entries against all the previous items, namely 1, 2, 3, 4, 5 and 6 taken together.

Illustration 17. – 1000 persons of London were asked by a B.B.C. investigator to give the nationality of the music they liked. He returns the following data :

570 liked English; 650 liked French; 480 liked German;

440 liked English and French; 360 liked French and German

& 240 liked English and German;

125 liked all the three.

Show that the information as it stands must be incorrect

Solution : Denoting

Persons who liked music of English nationality by 'A'.

Person who liked music of French nationality by 'B'

Persons who liked music of German nationality by 'C'.

The given data are :

N = 1000; (AB) = 440; (A) = 570; (B) = 650; (C) = 480; (AC) = 240; (BC) = 360; (ABC) = 125

According to the conditions of consistency.

$$\begin{aligned}(ABC) &\nless (AB) + (BC) - (B)\\ &\nless 440 + 360 - 650\\ &\nless 150\end{aligned}$$

As per the data given the value of (ABC) is 125 which is less than 150. Hence the information is incorrect.

Illustration 18. –

(i) It is found out from a report on Consumers Test that out of 500 persons who were investigated, 410 liked A–quality, 380 liked B–Quality and 270 liked both the qualities. Is the given data inconsistent ?

(ii) The following are the actual frequencies given in a report. Show that there is a mistake or misprint in it and probably the statement that (BC) = 85 is not true–

N = 1,000; (A) = 510; (B) = 490; (C) = 427; (AB) = 189; (AC) = 140 and (BC) = 85

Solution – (i) Assuming – Liking A – quality = A

Liking B – quality = B

The given class frequencies are :–

N = 500; (A) 410; (B) = 380; (AB) = 270

If any of the ultimate class – frequencies is negative, the given information will be inconsistent.

$$\begin{aligned}(Ab) &= (A) - (AB) = 410 - 270 = 140\\ (aB) &= (B) - (AB) = 380 - 270 = 110\\ (ab) &= (a) - (aB) \text{ or } N - (A) - \{(B) - (AB)\}\\ &= 500 - 410 - \{380 - 270\} = -20\end{aligned}$$

The frequency of (ab) is –20 which is negative; hence the given data is inconsistent

NOTES

(ii) According to the condition of consistency –

$$(AB) + (AC) + (BC) \nless (A) + (B)(C) - N$$
$$189 + 140 + (BC) \nless 510 + 490 + 427 - 1000$$
$$(BC) + 329 \nless 427$$
$$(BC) \nless 427 - 329 \text{ or } 98$$

But, the given frequency of (BC) is 85 which is less than 98. Therefore the given data are inconsistent and the class – frequency of (BC) is not true.

7.6 सम्भावना सिद्धान्त और आशंसा (Theory of Probability and Expectation)

सांख्यिकी में निदर्शन रीति (Sample Method) का विस्तृत विवेचन पिछले अध्यायों में किया जा चुका है। इसमें सम्भावना सिद्धान्त का भी उल्लेख हुआ है। सम्भावना (Probability) का अर्थ होता है किसी भी घटना के घटित होने या न होने के सम्बन्ध में कोई अनुमान लगाना। यदि एक घटना कई प्रकार से घटित हो सकती है और यह निश्चित नहीं है कि वह किस प्रकार घटेगी तो उस घटना के प्रत्येक प्रकार से घटित होने की सम्भावना उसी अनुमान में होती है जितने कि प्रकार हैं। उदाहरणार्थ, यदि किसी समूह में 100 व्यक्ति हैं, जिनमें 60 पुरुष और 40 स्त्रियाँ हैं। यदि उन 100 में से चिट्ठी डालकर (Lottery) इनाम देना है तो प्रत्येक पुरुष के इनाम पाने की सम्भावना $\frac{60}{100}$ और प्रत्येक स्त्री के इनाम पाने की सम्भावना $\frac{40}{100}$ होगी। यदि 20 बार चिट्ठी डालकर इनाम निकाला जाये तो पुरुषों के इनाम पाने की आशंसा (Expectation) $\frac{60}{100} \times 20 = 12$ होगी और स्त्रियों के इनाम पाने की आशंसा $\frac{40}{100} \times 20 = 8$ होगी। इस प्रकार, किसी घटना की आशंसा उसकी सम्भावना में संख्या से गुणा करके ज्ञात की जाती है। गुण-सम्बन्ध का अध्ययन इसी सिद्धान्त पर आधारित है। यदि हम ताश के पत्तों की जोड़ी में से लाल पान का पत्ता निकालना चाहें तो उसकी सम्भावना होगी $\frac{13}{52}$ क्योंकि लाल पान के पत्ते 13 होते हैं और कुल ताश के पत्ते 52 होते हैं। यदि इस प्रकार 100 बार ऐसा किया जाये तो लाल पान की आशंसा $\frac{13}{52} \times 100 = 25$ होगी।

अतः किसी घटना की सम्भावना को इस प्रकार मालूम किया जाता है

$$\frac{\text{Number of favourable cases}}{\text{Number of Cases (Total)}}$$

The Probability of (A) attribute is $\frac{(A)}{N}$

The Probability of (B) attribute is $\frac{(B)}{N}$

The Probability of (A) and (B) $= \frac{(A)}{N} \times \frac{(B)}{N}$

The Expectation of the two attributes will be

$$\frac{(A)}{N} \times \frac{(B)}{N} \times N = \frac{(A) \times (B)}{N}$$

7.7 गुणों की स्वतन्त्रता (Independence of Attributes)

जब किन्हीं दो गुणों में परस्पर सम्बन्ध नहीं होता तो उन्हें गुणों की स्वतंत्रता कहते हैं। इन गुणों में न तो धनात्मक सम्बन्ध होता है और न ही ऋणात्मक। ऐसा उस परिस्थिति में होता है, जबकि किसी वर्ग की वास्तविक आवृत्ति उस वर्ग की आशंसा के बराबर होती है। इसे निम्न सूत्रों द्वारा स्पष्ट किया जा सकता है :–

(A) and (B) are independent if (AB) $= \frac{(A) \times (B)}{N}$

(A) and (b) are independent if (Ab) $= \frac{(A) \times (b)}{N}$

(a) and (B) are independent if (aB) $= \frac{(a) \times (B)}{N}$

(a) and (b) are independent if (ab) $= \frac{(a) \times (b)}{N}$

Illustration 19 – Show whether A and B are independent in the following case–

N = 150, (A) = 50, (B) = 60, (AB) = 20

Solution –

Expectation of (AB) $= \frac{(A) \times (B)}{N} = \frac{50 \times 60}{150} = 20$

The actual observation is equal to the expectation. Therefore, A and B are independent.

Illustration 20 – Let the actual observation be as follows –

People vaccinated but not attacked by Small pox (AB) = 60

People not vaccinated and not attacked by Smallpox (aB) = 204

People vaccinated and attacked by Smallpox (Ab) = 80

People not vaccinated and attacked by Smallpox (ab) = 272

It is required to find whether attributes A and B are independent.

Solution – According to the criterion, (A) and (B) are independent only when,

$$(AB) \times (ab) = Ab \times (aB)$$

$$60 \times 272 = 80 \times 204$$

$$16320 = 16320$$

$\therefore$ A and B are independent.

Illustration 21. – In a population of 200 students the number of married is 80. Out of 60 students who failed, 24 belonged to the married group. It is required to find out whether the attributes of marriage and failure are independent.

Solution : Let, married = (A) = 80

Failure = (B) = 60

(AB) = 24

$\therefore$ Expectation of (AB) $= \frac{(A) \times (B)}{N} = \frac{80 \times 60}{200} = \frac{4800}{200} = 24$

Hence A and B independent.

7.8 गुण–सम्बन्ध (Association of Attributes)

सांख्यिकी में 'संबंध' शब्द का अर्थ इसके सामान्य अर्थ से भिन्न होता है। सामान्यत: हम 'A' और 'B' को सम्बन्धित मानते हैं यदि वे कई बार एक साथ आते है, किन्तु सांख्यिकी में ऐसा नहीं होता। सांख्यिकी में 'A' और 'B' में तभी सम्बन्ध माना जायेगा जबकि कई बार वे आशा से अधिक एक साथ प्रकट हों और यदि वे स्वतंत्र न हों। अत: जब कोई दो गुण सम्भावित इकाइयों से अधिक इकाइयों में एक साथ उपस्थित या अनुपस्थित हों अथवा एक की उपस्थिति तथा दूसरे की अनुपस्थिति एक साथ पाई जाये तो दोनों गुण आपस में सम्बन्धित कहलाते हैं। इस प्रकार गुण–सम्बन्ध दो प्रकार का हो सकता है।

(1) धनात्मक (Positive), (2) ऋणात्मक (Negative)

NOTES

(1) **धनात्मक गुण-सम्बन्ध** (Positive Association) – जब दो गुण साथ-साथ उपस्थित या अनुपस्थित हों तो उनमें धनात्मक गुण-सम्बन्ध होता है। जैसे जलवायु और स्वास्थ्य में धनात्मक गुण-सम्बन्ध होता है क्योंकि अच्छी जलवायु तथा अच्छा स्वास्थ्य और खराब जलवायु तथा खराब स्वास्थ्य साथ-साथ पाये जाते हैं। यदि वास्तविक आवृत्ति वर्ग आशंसा से अधिक हो तो धनात्मक गुण-सम्बन्ध पाया जाता है। इसे सूत्र रूप में निम्न प्रकार व्यक्त कर सकते हैं–

Association is positive, if

$$AB > \frac{(A) \times (B)}{N}$$

(2) **ऋणात्मक गुण-सम्बन्ध** (Negative Association) – जब एक गुण के होने पर दूसरे का अभाव हो तो इन दोनों गुणों में ऋणात्मक गुण-सम्बन्ध होता है, जैसे अच्छा स्वास्थ्यप्रद वातावरण और रोग में ऋणात्मक गुण-सम्बन्ध होता है। स्वास्थ्यप्रद वातावरण में वृद्धि के साथ रोगों में कमी होती है। यदि वास्तविक आवृत्ति वर्ग की आशंसा से कम होती है तो ऋणात्मक गुण-सम्बन्ध होता है इसे 'Dissociation' भी कहते हैं। सूत्र अग्र प्रकार व्यक्त किया जा सकता है –

Association is negative, if

$$(AB) < \frac{(A) \times (B)}{N}$$

Illustration 22.

Given (A) = 40 (B) = 30

(AB) = 20 N = 100

Study the association between 'A' and 'B'; 'a' and 'b'; 'A' and 'b'; 'a' and 'B'.

Solution – We can represent the given data in the shape of a table and obtain the frequencies of missing classes.

(AB) 20	(aB) 10	(B) 30
(Ab) 20	(ab) 50	(b) 70
(A) 40	(a) 60	(N) 100

The above values are observed or actual values. We can now calculate the expected values –

$$\left.\begin{aligned}(AB) &= \frac{(A) \times (B)}{N} = \frac{40 \times 30}{100} = 12 \\ (ab) &= \frac{(a) \times (b)}{N} = \frac{60 \times 70}{100} = 42\end{aligned}\right\} \text{Positively Associated}$$

$$\left.\begin{aligned}(Ab) &= \frac{(A) \times (b)}{N} = \frac{40 \times 70}{100} = 28 \\ (aB) &= \frac{(a) \times (B)}{N} = \frac{60 \times 30}{100} = 18\end{aligned}\right\} \begin{array}{c}\text{Negatively Associated} \\ \text{or Dissociated}\end{array}$$

Illustration 23. – In an anti-malaria campaign in a certain area, quinine was administered to 812 persons out of a total population of 3,248.

The number of fever cases is shown below –

Treatment	**Fever**	**No-fever**
Quinine	20	792
No-quinine	220	2216

Discuss the usefulness of quinine in checking malaria.

Solution – Denoting 'A' for quinine treatment

'a' for No quinine treatment

'B' for No attack of treatment

'b' for attack of fever

Then– (AB) = 79 2 (Ab) = 20

(aB) = 2216 (ab) = 220

To test association,

$$(AB) \times (ab) = (Ab) \times (aB)$$

$$(792) \times (220) = (20) \times (2216)$$

$$174240 > 44320$$

Hence there is positive association between A and B. It means quinine is helpful in checking Malaria.

गुण-सम्बन्ध को निम्न चार्ट द्वारा और स्पष्ट किया जा सकता है –

Attributes Independence	**Positive Association**	**Dissociation or Negative Association**
A and B (AB) $= \frac{(A) \times (B)}{N}$	AB $> \frac{(A) \times (B)}{N}$	AB $< \frac{(A) \times (B)}{N}$
A and b (Ab) $= \frac{(A) \times (b)}{N}$	Ab $> \frac{(A) \times (b)}{N}$	Ab $< \frac{(A) \times (b)}{N}$
a and B (aB) $= \frac{(a) \times (B)}{N}$	aB $> \frac{(a) \times (B)}{N}$	ab $< \frac{(a) \times (B)}{N}$
a and b (ab) $= \frac{(a) \times (b)}{N}$	ab $> \frac{(a) \times (b)}{N}$	ab $< \frac{(a) \times (b)}{N}$

7.9 गुण–सम्बन्ध गुणक (Coefficient of Association)

उपरोक्त रीति से केवल दो गुणों में सम्बन्ध है या नहीं – यही जाना जा सकता है। उनमें किन अंशों तक सम्बन्ध है, यह नहीं जाना जा सकता। गुण–सम्बन्ध किस सीमा तक है, यह ज्ञात करने के लिए प्रो. यूल ने गुण-सम्बन्ध का निम्न सूत्र दिया है –

$$Q = \frac{(AB)\,(ab) - (Ab)\,(aB)}{(AB)\,(ab) + (Ab)\,(aB)}$$

गुण सम्बन्ध गुणक सदैव + 1 और – 1 के बीच में होता है। इसका निर्वचन कार्ल पियर्सन के सह-सम्बन्ध की भाँति ही होता है। यदि गुण-सम्बन्ध गुणक + 1 आये तो पूर्णतया धनात्मक गुण-सम्बन्ध होता है, यदि –1 आये तो ऋणात्मक और यदि '0' आये तो कोई सम्बन्ध नहीं होता है। उपरोक्त सूत्र में 'Q' का अर्थ **गुण सम्बन्ध गुणक** (Coefficient of Association) होता है।

Illustration 24. : Calculate the coefficient of association between extravagance in fathers and sons from the following data–

Extravagant fathers with extravagant sons	327
Extravagant fathers with miserly sons	545
Miserly fathers with extravagant sons	741
Miserly fathers with miserly sons	235

Solution – Let, Extravagant father be denoted as 'A'

Miserly father " " 'a'

Extravagant sons " " 'B'

Miserly sons " " 'b'

The given data are –

(AB) = 327; (Ab) = 545, (aB) = 741; (ab) = 235

Now, applying yule's formula to find Coefficient Association –

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)} = \frac{(327)(235) - (545)(741)}{(327)(235) + (545)(741)}$$

$$= \frac{(76845) - (403845)}{(76845) + (403845)} = \frac{-327000}{480690} = -.68$$

$\therefore$ There is a negative association between extravagance in father and sons.

Illustration 25 – In a case-test in which 135 candidates were examined or proficiency in English and Economics. It was discovered that 75 students failed in English, 90 failed in Economics and 50 failed in both. Find if there is any association between failing in English and also state the magnitude of association.

Solution – If we express the above data by symbols the figures will appear as follows–

Failing in	English	(A)	= 75
" "	Economics	(B)	= 90
" "	both	(AB)	= 50
Total No. of Candidates		(N)	= 135

The magnitude of the association is measured by –

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)}$$

From the given data, we have to find out the ultimate class-frequencies required in the above formula –

(a) = N – (A) = 135 – 75 = 60

(b) = N – (B) = 135 – 90 = 45

(aB) = (B) – (AB) = 90 – 50 = 40

(Ab) = (A) – (AB) = 75 – 50 = 25

(ab) = (a) – (aB) = 60 – 40 = 20

Substituting the above values in the formula, We get,

$$Q = \frac{(50) \times (20) - (25) \times (40)}{(50) \times (20) + (25) \times (40)}$$

$$= \frac{(1000) - (1000)}{(1000) + (1000)} = \frac{0}{2000} = 0$$

Therefore, failure in English and Economics are completely independent of each other.

Illustration 26. – From the table given below, compare the intensity of association between literacy and unemployment among males in urbran areas with that in the rural areas -

	(N)	(A)	(B)	(AB)
Urban (lakhs)	25	10	5	3
Rural (lakhs)	200	40	12	4

Solution : For comparing the intensity of association, coefficient of association should be calculated –

$$Q = \frac{(AB)\,(ab) - (Ab)\,(aB)}{(AB)\,(ab) + (Ab)\,(aB)}$$

Urban Areas – (Ab) = (A) – (AB) = 10 – 3 = 7

(aB) = (B) – (AB) = 5 – 3 = 2

(ab) = (a) – (aB) = N – (A) – (aB)

= 25 – 10 – 2 = 13

$$Q = \frac{(3) \times (13) - (7) \times (2)}{(3) \times (13) + (7) \times (2)} = \frac{39 - 14}{39 + 14}$$

$$= \frac{25}{53} = .472$$

Rural Areas – (Ab) = (A) – (AB) = 40 – 4 = 36

(aB) = (B) – (AB) = 12 – 4 = 8

(ab) = (a) – (aB) = N – (A) – (aB)

= 200 – 40 – 8 = 152

$$Q = \frac{(4) \times (152) - (36) \times (8)}{(4) \times (152) + (36) \times (8)}$$

$$= \frac{608 - 288}{608 + 288} = 0.357$$

Therefore, coefficient of Association between A and B for urban areas is more than that of rural areas.

Illustration 27. – Can vaccination be regarded as a preventive measure for Small pox from the data given below ?

'Of 1482 persons, in a locality exposed to small pox, 368 in all were attacked.'

'Of 1482 persons, 343 had been vaccinated and of these only 35 were attacked.'

Solution – Let,

'A' be the attribute of vaccination and 'a' that of non-vaccination.

'B' be the attribute of exemption from attack of small pox and

'b' that of attack of small pox.

Given data are –

N = 1482; (A) = 343; (b) = 368; (Ab) = 35;

From the above,

(AB) = (A) – (Ab) = 343 – 35 = 308

(a) = N – (A) = 1482 – 343 = 1139

NOTES

$$(aB) = (B) - (AB) \quad = N - (b) - (AB)$$
$$= 1482 - 368 - 308 = 806$$

Now, percentage of vaccinated who were not attacked.

$$\text{i.e. } \frac{(AB)}{(A)} \times 100 = \frac{308}{343} \times 100 = 90\%$$

and percentage of not vaccinated who were not attacked.

$$\text{i.e. } \frac{(aB)}{(a)} \times 100 = \frac{806}{1139} \times 100 = 71\%$$

A comparison of the two percentages reveal that vaccination and exemption from attack are positively associated. Hence vaccination may be regarded preventive measure for small pox.

Illustration 28 – In a diphtheria epidemic, out of 254 children who were not protected against it, 20 suffered and 12 of them died. But out of 294 children to whom triple-antigen was injected, 6 suffered and only one died.

Do you find any association between (a) innoculation and suffering from diphtheria, (b) innoculation and mortality among children who suffered from it ?

Solution : (a) Denoting

Innoculation by (A), absence of it by 'a' infection by 'B' and absence of it by 'b'. The given data are –

$$(A) = 294;\ (a) = 254;\ = (aB) = 20;\ (AB) = 6$$
$$(Ab) = (A) - (AB) = 294 - 6 = 288\ ;$$
$$(ab) = (a) - (aB) = 254 - 20 = 234$$

Substituting values in Yule's formula

$$Q = \frac{(AB)(ab) - (Ab)(aB)}{(AB)(ab) + (Ab)(aB)} = \frac{(6 \times 234) - (288 \times 20)}{(6 \times 234) + (288 \times 20)}$$
$$= \frac{1404 - 5760}{1404 + 5700} = \frac{-4356}{7164} = -0.6$$

Thus, the coefficient of association between inoculation and contracting the disease is – 0.6.

(b) If 'A' and 'a' denote vaccination and its absence respectively, amongst those who contacted the disease and 'B' and 'b' represent death and recovery respectively, the given data are –

$$(A) = 6; \quad (a) = 20$$
$$(AB) = 17; \quad (aB) = 12$$

The values of

$$(Ab) = (A) - (AB) = 6 - 1 = 5$$
$$(ab) = (a) - (aB) = 20 - 12 = 8$$

Substituting the values in Yule's formula of Coefficient of association, we get,

$$Q = \frac{(1) \times (8) - (5) \times (12)}{(1) \times (8) + (5) \times (12)} = \frac{8 - 60}{8 + 60} = \frac{-52}{68} = -0.76$$

Thus, coefficient of association between inoculation and mortality among persons who contacted disease is – 0.76.

अनुपात-तुलना (प्रतिशत) रीति से गुण-सम्बन्ध
(Association by Method of Comparison of Proportions)

इस रीति के अनुसार दो गुणों के आधार पर दो प्रतिशतें निकाली जाती हैं –

एक तो B की A में $\left[\frac{(AB)}{(A)} \times 100\right]$ तथा दूसरी B की a में $\left[\frac{(aB)}{(a)} \times 100\right]$ इन दोनों के तुलनात्मक अध्ययन द्वारा गुण-सम्बन्ध का निर्धारण किया जाता है। यदि दोनों प्रतिशतें बराबर हों तो गुण-स्वातन्त्र्य (Independence) होता है। यदि A में B का अनुपात a में B के अनुपात से अधिक होता है तो A और B में धनात्मक गुण सम्बन्ध होता है। इसकी विपरीत स्थिति हो तो यह ऋणात्मक होता है।

संकेताक्षरों के रूप में इसके विभिन्न सूत्र निम्न प्रकार हैं –

B के A और a में अनुपात	A के B और b में अनुपात	निष्कर्ष
$\frac{(AB)}{(A)} \times 100 = \frac{(aB)}{(a)} \times 100$	$\frac{(AB)}{(B)} \times 100 = \frac{(Ab)}{(b)} \times 100$	गुण-स्वातन्त्र्य
$\frac{(AB)}{(A)} \times 100 > \frac{(aB)}{(a)} \times 100$	$\frac{(AB)}{(B)} \times 100 > \frac{(Ab)}{(b)} \times 100$	धनात्मक (Positive)
$\frac{(AB)}{(A)} \times 100 > \frac{(aB)}{(a)} \times 100$	$\frac{(AB)}{(B)} \times 100 < \frac{(Ab)}{(b)} \times 100$	ऋणात्मक (Negative)

Illustration 29 – In a district of India; out of 70,000 literates 500 were criminals and out of 9,30,000 illiterates, 15,000 were criminals. Do you find any association between illiteracy and criminality from these figures ?

Solution – Representing illiteracy by A and Criminality by B, we find the following class – frequencies–

(a) = 70,000; (aB) = 500; (A) = 9,30,000, (AB) = 15,000

Therefore, the percentage ratio of criminals among illiterates

$$= \frac{(AB)}{(A)} \times 100 = \frac{15,000}{9,30,000} \times 100 = 1.6\%$$

The percentage ratio of criminals among literates

$$= \frac{(aB)}{(a)} \times 100 = \frac{500}{70,000} \times 100 = 0.71\%$$

Hence, there is positive association, between illiteracy and criminality because $\frac{(AB)}{(A)} \times 100 > \frac{(aB)}{(a)} \times 100$

तथ्यानुबन्ध गुणांक (Coefficient of Colligation)

यूल ने गुण-सम्बन्ध के माप के लिए एक अन्य सूत्र का भी प्रयोग किया है। यह सूत्र अग्र प्रकार है–

$$\text{Coefficient of colligation or } Y_{AB} = \frac{1 - \sqrt{\frac{(Ab)\,(aB)}{(AB)\,(ab)}}}{1 + \sqrt{\frac{(Ab)\,(aB)}{(AB)\,(ab)}}}$$

गुण सम्बन्ध गुणांक (Q_{AB}) तथा तथ्यानुबन्ध गुणांक (Y_{AB}) में निम्नलिखित सम्बन्ध है –

$$Q_{AB} = \frac{2Y}{1 + Y^2}$$

व्यवहार में, तथ्यानुबंधन गुणांक का प्रयोग नहीं किया जाता है।

NOTES

7.10 आंशिक गुण-सम्बन्ध (Partial Association)

गुण-सम्बन्ध, मात्रा की दृष्टि से दो प्रकार के होते हैं –

(1) सम्पूर्ण गुण-सम्बन्ध (Total Association)

(2) आंशिक गुण-सम्बन्ध (Partial Association)

सम्पूर्ण गुण-सम्बन्ध तब होता है जबकि यह पूरे समूह में पाया जाता हो और वहाँ किसी तीसरे गुण का प्रभाव नहीं पड़ता हो। आंशिक गुण-सम्बन्ध की अवस्था में गुण-सम्बन्ध अप्रत्यक्ष होता है और उस पर किसी तीसरे गुण का प्रभाव होता है। जैसे, 'अ' और 'ब' में कोई प्रत्यक्ष गुण सम्बन्ध नहीं है किन्तु उनमें यह सम्बन्ध किसी तीसरे गुण 'स' के कारण है जिसका सम्बन्ध 'अ' और 'ब' दोनों से है। उदाहरणार्थ, टीका लगाने (A) और चेचक से मुक्ति (B) में धनात्मक गुण-सम्बन्ध होना स्वाभाविक है। इसलिए यह निष्कर्ष निकालना स्वाभाविक ही है टीका लगाने से चेचक से मुक्ति मिल जाती है। किन्तु यह निष्कर्ष unwarranted भी समझा जा सकता है। यह कहा जा सकता है कि टीका लगाने की प्रथा केवल सम्पन्न और शिक्षित घरों (C) में प्रचलित है और अन्य लोग इसे संदेह से देखते हैं। अतः (A) और (B) में जो गुण-सम्बन्ध हमें दिखाई देता है वह इन दोनों का (C) के साथ गुण-सम्बन्ध होने के कारण है। उप-समग्र (Sub-Universe) 'C' में 'A' और 'B' के गुण-सम्बन्ध को आंशिक् गुण-सम्बन्ध कहते हैं।

'A' और 'B' में यदि आंशिक गुण-सम्बन्ध हो तो इसका उल्लेख स्पष्ट रूप से कर देना चाहिये, क्योंकि जो एक अंश 'C' के लिये सच हो, आवश्यक नहीं है कि वह अन्य अंशों 'D' और 'E' आदि के लिए सच हो।

आंशिक गुण-सम्बन्ध निकालने की भी वही विधि है जो सम्पूर्ण गुण-सम्बन्ध के लिये है। अन्तर केवल इतना है कि यहाँ दो के स्थान पर तीन गुणों का विवेचन होता है। तीसरे गुण की उपस्थिति को 'C' और अनुपस्थिति को 'c' के रूप में प्रदर्शित करते हैं। इसके लिए अग्र सूत्र का प्रयोग होता है –

'C' वर्ग में 'A' और 'B' का गुण सम्बन्ध गुणक –

$$Q(C) = \frac{(ABC)\,(abC) - (AbC)\,(aBC)}{(ABC)\,(abC) + (AbC)\,(aBC)}$$

'c' वर्ग में 'A' और 'B' का गुण-सम्बन्ध गुणक –

$$Q(c) = \frac{(ABc)\,(abc) - (Abc)\,(aBc)}{(ABc)\,(abc) + (Abc)\,(aBc)}$$

7.11 भ्रमात्मक गुण-सम्बन्ध (Illusory Association)

ऐसा गुण-सम्बन्ध जो दो गुणों के मध्य किसी वास्तविक सम्बन्ध के कारण न हो, भ्रमात्मक गुण-सम्बन्ध होता है। ऊपर आंशिक गुण-सम्बन्ध का उल्लेख करते समय यह स्पष्ट किया जा चुका है कि वास्तव में आंशिक गुण-सम्बन्ध होने पर उसे सम्पूर्ण गुण-सम्बन्ध मान लिया जा सकता है। यह निष्कर्ष भ्रमोत्पादक होगा। यह निम्न में से किसी भी ढंग से निकाले हुए हो सकते हैं–

(1) **आंशिक गुण-सम्बन्ध होने पर सम्पूर्ण गुण-सम्बन्ध मान लेना** – आंशिक गुण-सम्बन्ध के होने पर उसे सम्पूर्ण मान लिया जाये और तीसरे गुण के प्रभाव को ध्यान में नहीं रखा जाये।

(2) **दो गुण ठीक प्रकार से परिभाषित नहीं किये गये हों** – यदि गुणों की ठीक-ठीक परिभाषा नहीं दी गई हो तो सामग्री गलत हो जाती है और उनमें स्थापित गुण-सम्बन्ध पूर्णतः भ्रमात्मक हो सकता है।

(3) **धनात्मक को ऋणात्मक और ऋणात्मक को धनात्मक मान लेना** – जब धनात्मक गुण-सम्बन्ध को ऋणात्मक या ऋणात्मक गुण–सम्बन्ध को धनात्मक मान लिया जाये तो भी यह भ्रमात्मक होगा।

(4) **दो समूहों में विपरीत गुण** – सम्बन्ध होने पर अल्प मात्रा का गुण – सम्बन्ध मानना या स्वतंत्र मान लेना।

(5) **दो समूहों में से एक के स्वतंत्र गुणों को और दूसरे के उच्च स्तरीय गुण** – सम्बन्ध को सम्पूर्ण के लिये सामान्य मात्रा का मान लेना।

भ्रमात्मक गुण-सम्बन्ध, अलग-अलग वर्गों का गुण–सम्बन्ध एक साथ निकालने, न्यादर्श चुनने में पक्षपात, गुणन क्रिया में असावधानी, गुणक की अशिक्षा और अनुभवहीनता आदि के कारण हो सकता है।

Illustration 30 – Explain clearly what you understand by association and illusory association ?

In a state with a total population of 70,000 adults; 34,000 are males and out of a total of 6,000 graduates, 700 are females. Out of 1,200 graduate employees of the state 200 are females. Is there any sex bias in education among people ? The state holds that no distinction is made in appointment in respect of sex. How far is their claim substantiated by the data given above ?

Solution – The percentage of males who are graduates is

$$\frac{5300 \times 100}{34,000} = 116\% \text{ approx}$$

The percentage of females who are graduates is

$$\frac{700 \times 100}{36,000} = 2\% \text{ approx}$$

The above percentages clearly show that there is a sex bias in education. A larger percentage of males than females get education.

$$\frac{700 \times 1200}{6,000} = 1400$$

The actual number is 200 which shows that there is a bias in favour of females.

7.12 गुण-सम्बन्ध तथा सह सम्बन्ध (Association of Attributes and Correlation)

गुण-सम्बन्ध और सह-सम्बन्ध में कई बातों में समानता और कई बातों में असमानता है जिसका उल्लेख आगे किया जाता है –

समानता (Similarity)

(1) **दो समूहों में गणना** – दोनों ही मापों की गणना दो समूहों के मध्य की जाती है।

(2) **धनात्मक और ऋणात्मक** – दोनों ही माप या तो धनात्मक या ऋणात्मक निष्कर्ष प्राप्त करते हैं। दोनों मापों की मात्रा की सीमायें भी + 1 और – 1 के मध्य होती हैं।

(3) **सापेक्ष माप** (Relative Measure) – दोनों ही माप सापेक्ष हैं। अत: किसी भी अन्य श्रेणी के माप से उनकी तुलना की जा सकती है।

असमानता (Dissimilarity)

(1) **विभिन्न आधारों पर श्रेणियों का विभाजन** – दोनों मापों की गणना विभिन्न आधारों पर विभाजन करके की जाती है। गुण-सम्बन्ध में विभाजन गुणों के आधार पर होता है और सह-सम्बन्ध में यह संख्या के आधार पर किया जाता है।

(2) **गुणों की उपस्थिति और अनुपस्थिति** – गुण-सम्बन्ध निकालते समय गुणों की उपस्थिति और अनुपस्थिति दोनों पर विचार करना आवश्यक होता है जबकि सह-सम्बन्ध ज्ञात करने के लिये गुणों की उपस्थिति ही पर्याप्त होती है।

(3) **श्रेणियों की संख्या** – गुण-सम्बन्ध सामान्यत: एक ही श्रेणी के विभिन्न समूहों के गुणों में मालूम किया जाता है और सह-सम्बन्ध दो श्रेणियों के बीच निकाला जाता है।

(4) **समूह और श्रेणी** – गुण-सम्बन्ध एक ही श्रेणी के विभिन्न समूहों में निकाला जा सकता है तथा सम्पूर्ण का भी निकाला जा सकता है किन्तु सह-सम्बन्ध दो श्रेणियों के बीच ही निकाला जाता है।

(5) **गणना में सरलता** – गुण-सम्बन्ध गुणक ज्ञात करना सह-सम्बन्ध गुणक की अपेक्षा सरल होता है। और कम समय भी लगता है।

(6) **गुणक ज्ञात करने की रीतियाँ** – गुण-सम्बन्ध गुणक ज्ञात करने की एक ही रीति होती है किन्तु सह-सम्बन्ध गुणक कई रीतियों से निकाला जा सकता है।

NOTES

बोध प्रश्न

1. गुण संबंध किसे कहते हैं ?

....

....

....

2. आंशिक गुण संबंध को समझाइए।

....

....

....

3. गुण संबंध एवं सह-संबंध में समानताएँ बताइए।

....

....

....

4. आर्थिक सांख्यिकी विश्लेषण में गुण-संबंध गुणक के उपयोग पर टिप्पणी कीजिए।

....

....

....

7.13 सारांश

7.14 शब्दावली या शब्दकुंजी

7.15 अभ्यास प्रश्न

दीर्घउत्तरीय प्रश्न (Long answer type question)

1. गुण-सम्बन्ध से आप क्या समझते हैं ? उसकी उपस्थिति अथवा अनुपस्थिति का किस प्रकार निर्धारण किया जाता है ?

   What do you understand by Association of Attributes? How is its existence or non-existence determined?

2. गुण-सम्बन्ध और सह-सम्बन्ध का भेद स्पष्ट कीजिए। आंशिक गुण सम्बन्ध के विचार और उसके महत्व पर भी प्रकाश डालिए।

   Distinguish between Association and Correlation. Also throw light on the concept of Partial-Association and its importance.

3. सह सम्बन्ध, प्रतीपगमन और गुण सम्बन्ध में अन्तर कीजिए। उन परिस्थितियों का वर्णन कीजिए जिनमें इनका प्रयोग किया जाना चाहिये। अपने उत्तर को सउदाहरण स्पष्ट कीजिए

   Distinguish between correlation, regression and association and describe the situations in which each of these should be used. Illustrate you answer with examples.

4. दो गुणों में गुण सम्बन्ध का अध्ययन करने के उद्देश्य से संकलित समंकों की कसौटियों की व्याख्या कीजिए।
State the criteria for consistency of data collected for studying association of two attributes.
5. गुण सम्बन्ध से आप क्या समझते हैं? विभिन्न गुणों के अनुसार वर्गीकृत समंकों की विश्वसनीयता (संगीत) की जाँच आप कैसे करेगें?
What do you understand by association of attributes? How will you examine the consistency of data classified according to different attributes.
6. आर्थिक सांख्यिकी के विश्लेषण में गुण सम्बन्ध गुणक के उपयोग पर टिप्पणी लिखिए।
Write a note on the use of Coefficient of association in analysing economic-statistics.
7. आसंजन-उत्तमता की काई-वर्ग जाँच क्या है? इस जाँच का प्रयोग करते समय कौन-कौन सी सावधानियों रखना आवश्यक है?
What is the Chi-Square test of goodness of fit? What precautions are necessary in using this test?
8. निम्नलिखित पर संक्षिप्त टिप्पणी लिखिए-
Write short-notes on the following-
   i. आंशिक गुण सम्बन्ध (Partial Association)
   ii. भ्रमात्मक गुण सम्बन्ध (Illusory Association)
   iii. समंकों की संगति (Consistency of data)
   iv. अंतिम वर्ग आवृत्तियाँ (Ultimate class-frequencies)
   v. तथ्यानुबंधन गुणक (Coefficient of Colligation)
   vi. माध्य वर्ग आकस्मिकता गुणक (Coefficient of mean Square Contingency)
   vii. शुप्रो-गुणक (Tschuprow's Coefficient)
   viii. येट के संशोधन (Yate's correlation's)
   ix. शून्य परिकल्पना (Null Hypothesis)
   x. स्वतन्त्रयांश (Degrees of freedom)

लघु उत्तरीय प्रश्न (Sort Answer Type Quetions)

1. गुण सम्बन्ध से आप क्या समझते है?
What do you understand by Association of attributes?
2. गुण सम्बन्ध और सह सम्बन्ध में भेद स्पष्ट कीजिये।
Distinguish between Association and corraelation.
3. गुण सम्बन्ध की माप कैसे करेंगे।
How would you calculate Association Attributes.?
4. दो गुणों में गुण सम्बन्ध ज्ञात करने की एक विधि का वर्णन कीजिये।
Explain one method of finding association betweet two attributes.

वस्तुनिष्ठ प्रश्न- (Objective Type Questions)

1. गुण सम्बन्ध गुणांक का प्रतिपादन किसने किया है-
(अ) फिशर, (ब) कार्ल पियर्सन, (स) यूल, (द) केट।
2. दो गुणों में गुण सम्बन्ध के वर्गों की संख्या है-
(अ) 8, (ब) 7, (स) 4, (द) 10।

3. द्वंद्व-भाजन वर्गीकरण में कितने वर्ग होते है-

(अ) 4, (ब) 3, (स) 2, (द) 6 ।

4. विभिन्न गुणों की संख्या के आधार पर निम्न प्रकार ज्ञात किया जाता है-

(अ) एक गुण के आधार पर, (ब) दो गुणों के आधार पर,

(स) तीन गुणों के आधार पर, (द) उपरोक्त सभी।

(उत्तर - (1) ब, (2) अ, (3) स, (4) द)

7.16 व्यावहारिक प्रश्न -

1. The following the number of boys obsered with certain classes of defects amongst a number of school children. 'A' denotes development defects,'B'nerve signs and 'C' low nutrition.

(ABC) = 149;(ABc) = 738;(AbC) = 225;(aBC) = 204;(Abc) = 1,196; (aBc) = 1,762;(abC) = 171;(abc) = 21,842

Find the frequencies of the positive classes.

(N) = 26287,(A) = 2308,(B) 2583, (C) = 749,(AB) = 887,(AC) = 374, and (BC) = 353)

2. Given the following positive class-frequencies, find all the ultimate class-frequencies-

N = 10000; (A) = 877; (B) = 1086; (C) = 286;

(AB) = 338;(AC) = 143; (BC) = 135; (ABC) = 57.

[(ABC) = 57; (ABc) = 281; (AbC) = 86;(aBC) = 78;(Abc) = 453; (aBc) = 670; (abC) = 65; (abc) = 8310]

3. Out of 818 inhabitants of a locality, 279 were inoculated against cholera, 3 of whom were attacked.

The total number of persons exempt from attack of cholera was 749.

Arrange the information in a nine-square table and obtain the following class-frequencies-

(i) Number of persons not inoculated;

(ii) Number of persons attacked;

(iii) Number of persons inoculated and not attacked;

(iv) Number of persons not inoculated but exempt from attack;

(v) Number of persons not inoculated and attacked;

[(i) 539, (ii) 69, (iii) 276, (iv) 473, (v) 66]

4. At an examination at with 600 candidates appeared, boys outnumbered girls by 16%. Also those passing the examination exceed in number those failing by 310. The numbers of successful boys choosing science subjects was 300 while among the girls offering Arts subjects there were 25 failures. Altogether only 135 offered Arts and 33 among them failed. Boys failing in the examination numbered 18. Obtain all the class frequencies.

[(A) = 348, (AB) = 330, (aB) = 125, (ABC) = 330, (Abc) = 8

(a) = 252, (AC) = 310, (aC) = 155, (AbC) = 10, (abc) = 25

(B) = 455, (BC) = 353, (bC) = 112, (aBC) = 53, N = 600

(b) = 145, (Ab) = 18, (ab) = 127, (abC) 102,

(C) = 465, (Ac) = 38, (ac) = 97, (aBc) = 30,

(c) = 135, (BC) = 102, (bc) = 33, (aBc) = 72]

5. 100 Students took three examinations, 40 passed the first, 39 passed the second, and 48 passed the third. 10 passed all three, 21 failed all three, 9 passed the first two and failed the third and 19 failed the first two and passed the third. Find how many children passed at least two examinations.

6. Three aptitude tests A,B,C, were given to 200 apprentice trainees. from amongst them 80 passed test A, 78 passed test B and 96 passed the third test while 20 passed all three tests,42 failed the three, 18 passed A and B but failed C, and 38 failed A and B but passed the third. Determine (i) how many trainees passed atleast two of the three tests and (ii) whether the performances in tests A and B are associated.

((i) 76; (ii) Yes, positively associated as (AB))

7. Given the following positive class frequencies, find all the ultimate class-frequencies:-

N = 23713, (A) = 1618; (B) = 2015; (C) = 770; (AB) = 587; (BC) = 428; (AC) = 335; (ABC) = 156.

[(ABC = 156 given), (ABc) = 431, (AbC) = 179, (aBC) = 272, (Abc) = 582, (aBc) = 1156, (abC) = 163, (abc) = 20504]

8. (a) Given the following frequencies of the positive classes, find the frequencies of the ultimate classes:-

(A) = 40; (B) = 60; (AB) = 30; N = 130.

(b) Examine whether A and B are independent in the following case-

(A) = 490; (AB) = 294; (a) = 560; (aB) = 380;

[(Ab) = 10; (ab) = 60; (aB) = 30]; (b) A and B are not independent.]

9. What do you understand by 'Consistence of Data'? Do you find any inconsistency in the following data made available for investigation the eye colour of brothers and sisters?

Brothers with lights eyes and Sisters with not light eyes = 414

Brothers with not light eyes and Sisters with light eyes = 260

Brothers with light eyes = 400

(AB = -14, Hence data are inconsistent)

10. In a market survey relation to consumer preferences for tea, coffee and milk, the following information about 2000 persons was obtained-1040 liked taa, 620 liked coffee; 940 liked milk; 160 liked tea and coffee; 300 liked tea and milk; 170 liked coffee and milk; 50 liked all the three. Show that the information as it stands must be incorrect.

[Incorrect; (ABC) + (AB) + (AC) + (BC) – (A) – (B) – (C) + N]

NOTES

11. A market investigator returns the following data of 1000 people consulted-400 liked chocolates; 390 liked toffee; 480 liked lemon drops; 380 liked chocolates and toffees; 350 liked chocolates and lemon drops; 370 liked toffee and lemon drops.

Show that the information as it stands must be incorrect.

[(AB) + (BC) – (AC) (B)]

12. In a college, 50% of students are boys; 60% of the students are above 18 years and 80% receive scholarship. 35% of the students are boys above 18 years of age. 45% are boys receving scholership and 42% are above 18 years and recieve scholarship. Determine the limits (lowest and highest), of the proportion of boys above 18 years who are in receipt of scholarships

(limits 30% and 32%)

13. In a very hotly fought battle-

70% at least of the combatants lost an eye; 75% at least an ear, 80% at least a leg and 85% at least an arm. What percentage at least lóst all four?

(10%)

14. (i) If there are 144 A's and 384 B's in 1024 observations. How many AB's will there be A & B being independent?

(ii) If the A's are 60% of the whole number of observations what must be the percentage of AB's in order that we may conclude that A and B are independent?

[(i) 54(ii) 21%]

15. Out of 900 persons 300 were literate and 400 had travelled beyond the limits of their district; 200 of the literates were among those who had travelled. Is there any relation between travelling and literacy?

(Q = + 0.6)

16. In the study of aggregates having different attributes how would you determine whether the attributes are mutually independent or are associated in some way?

1660 candidates appeared for a competitive examination. 422 were successful, 256 had attended a Coaching class and of these 150 came out successful. Estimate the utility of the Coaching class.

(% of success among those attending the class = 89.8%

" " " " " not attending " " " " " " " " ;19.4% Q = +.71

Hence coaching class is useful.)

17. Can vaccination be regarded as a preventive measure of small pox from the data given below? "Of 1482 persons in a locality exposed to small pox, 368 in all were attacked. Of 1482 persons 343 had been vaccinated and of these only 35 were attacked.

(Ans. = + 0.573, percentage = 89.8%; 70.8%)

18 Explain the method of finding association between two attributes. Out of 70,000 literates in a particular district of India, number of criminals was 590. Out of 930 thousand of illiterates in the same district number of criminals was 15,000. On the basis of these figures do you find any assocition between illiteracy and criminality?

(And (i) Illiterate Criminals = 16% (ii) L iterate Crim nals = 0.7%)

19. The male population of a town is 250 lakhs. The number of literate males is 50 lakhs and the total number of male criminals is 36 thousand. The number of literate male criminals is 2 thousand. Do you find any association between literacy and criminality? (Ans. Q = -0.6)

20. 750 candidates appeared and 470 passed in an examination, 365 attended classes and 51 of them failed. Prove the utility of the classes.

    (Ans. Pass percentage of those who attended classes = 86.6: who did not attend classes = 13.4)

21. An investigation was carried out to determine whether there is any association between the eye-colour of parents and the eye-colour of children.

    The eye colours were noted in the case of a random sample of 1,000 fathers and their eldest sons. In 471 cases both fathers and sons hadlight eyes, in 230 cases both had dark eye, in 148 cases the fathers were dark-eyed and sons light-eyed and all remaining cases the sons were dark-eyed and the fathers were light eyed.

    Determine whether eye color in fathers and in sons is associated or independent.(Ans. Associated)

22. In a certain investigation carried on with regard to 500 graduates and 1,500 non-graduates, it was found that the number of employed graduates was 450 while the number of unemployed non-graduates was 300. In the second investigation, 5,000 cases were examined. The number of non-graduates was 3,000 and the number of employed non-graduates was 2,500. the number of graduates who were found to be employed was 1,600.

    Calculate the coefficient of association between graduation and employment in both the investigations.

    (Ans. Q(1) = + 38 and Q(2) = -.11)

23. Calculate the Coefficient of Association between Intelligence in father and sons from the following data:

Intelligent fathers with intelligent sons	248
Intelligent fathers with dull sons	81
Dull fathers with intelligent sons	92
Dull fathers with dull sons	579

    (Ans. Q = + 0.9)

24. Calculate the Coefficient of Association between extravagance in fathers and sons from the following data:

Extravagant fathers with extravagant sons	327
Extravagent fathers with miserfy sons	545
Miserly fathers with extravagant sons	751
Miserly fathers with miserly sons	235

    (Ans. Q = -0.68)

25. Find out the Coefficient of Association between the type pf training and success in teaching from the following table. Also find coefficient of contingency:

Institution	Successful	Unsuccessful	Total
Teacher's College	58	42	100
University	49	51	100
Total	107	93	200

(Ans. Q = +0.185; C = 0.09; X2 = 1.628)

26. What do you understand by 'Contingency'? In an investigation into the health and nutrition of certain children (between the ages of one and five years) two groups of childern were compared, one belonging to the well-to-do class, 125 in number, and the other belonging to the poor class, 124 in number. The following results were obtained:

	poor children (%)	well-to-do children (%)
Below normal weight	75	23
Above normal weight	5	42

Find the Coefficient of Association between the weight of the children and their parents financial condition.

(Ans. Q = + 0.93)

27. The following table shows the distribution of the temper in pairs of sisters in an exhaustive school inquiry:

Second Sister	First Sister		Total
	Good natured	Sullen	
Good natured	1,040	180	1,220
Sullen	160	120	280
Total	1,200	300	1,500

Trace the association, if any, in the distribution of tempers in first and second sisters.

(Ans. Q = + 0.62)

28. A census revealed following figures of the blind and the insane in two age groups in a certain population:

	Age-group (15-25 years)	Age-group (Over 26 years.)
Total population	2,70,000	1,60,200
No.of blinds	1,000	2,000
No. of insanes	6,000	1,000
No.of insanes among the blinds	19	9

(a) Obtain a measure of association between blindness and insanity in each of the two age groups.

(b) Do you consider that blindness and insanity are associated or dissociated with each other in the two age groups or more in one-age proup than in other?

(Q for First = – 0.07 and Q for Second = – 0.16. There is greater degree of dissociation in over 25 years age group)

29. (a) How would you distinguish between 'association' and 'correlation' as the terms are used in statistics?

(b) From the figures given in the following table, compare the association between literacy and unemployment in the rurel and urban areas, and give reasons for the difference, if any:

	Urban	Rural
Total Adult males	25 lakhs	200 lakhs
Literate "	10 "	40 "
Unemployed "	5 "	12 "
Literate and unemployed males	3 "	4 "

(Ans. Urban Q = +.47; Rural Q = +.35)

30. The following table gives the number of literates and criminals in three cities:

	Kanpur	Allahabad	Agra
Total Number (in thousands)	244	184	230
Literates (")	40	47	33
Literate Criminals (in hand)	3	2	2
Illiterate : (")	40	20	24

Compare the degree of association between criminality and illiteracy in each of the three towns.

[Ans. Q. (Kanpur) = +.45; Allahabad = +.55; Agra = +.34]

31. The following table gives the number of persons suffering from certain infirmities in Burma in 2001 :

Sex	Total in Lakhs	Insanes	Deaf Mutes	Deaf Mutes and Insanes
Males	260	12,650	21,301	545
Females	241	9,055	14,136	317

Trace the association between insanity and deaf muteness for males and females separately.

(Ans. Q of Males = + 0.965 and Females = + 0.968)

32. In an anti-malaria campaign in a certain area, quinine was administered to 812 persons out of a total population of 3248. The number of fever cases is as below:

Treatment	Fever	No fever
Quinine	20	792
No Quinine	220	2216

Discuss the usefulness of quinine in checking Malaria.

(Ans. Q = –.59)

33. What is partial association of the attributes of a group of population? Of the 60,000 persons in a town, 10,000 are literate. The total number of unemployed persons is 540 of whom 35 are literate. Is there any association between literacy and employment?
(Ans. Q = + .48)

34. The male population of Orissa is 250 lakhs. The number of literate males is 20 lakhs and the total number of male criminals is 2,000 of whom 100 are literate. Do you find any association between literacy and criminality?
(Ans. Q = + .48)

35. In a state with a total population of 70,000 adults, 34,000 are males and out of a total of 6,000 graduates, 700 are females. Out of 1,200 graduate employees of the state. 200 are females. Is there any sex-bias in education among the people. The state holds that no distinction is made in appointments in respect of sex. How far is their claim substantiated by the data given above?

Perceny of Male graduates	= 15.6%
" Female "	= 1.94%
Graduate males Employed	= 18.86%
" Females "	= 28.57%

36. A study was made about the studying habits of students of a certain university and the following summary is given at one place in the report of the students surveyed, 75% were from well-to-do families, 55% were boys and 69% were irregular in their studies. Out of the irregular ones, 50% were boys and 2/3 were from well-to-do families. The percentage of irregular boys from well-to-do families was 8. Is there any inconsistency in the data?
Give proof for your answer.
[Ans. There is inconsistency]

37. From the following find whether blindness and baldness are associated-

Total Population	1,62,64,000
Number of bald headed	24,441
Number of blind	7,623
Number of bald headed blind	221

[Ans. There is positive association because expected frequencies of (AB) are less than the actual.]

38. The following summary appears in a report on a survey covering 1,000 fields. Find out if the data are consistent.

Manured fields	510
Irrigated fieldst	490
Fields growing improved varieties	427
Fields both irrigated and manured	189
Fields both manured and growing improved varieties	140
Fields both irrigated and growing improved varieties	85

[There is consistency]

NOTES

39. Following information relates to literacy and unemployment among 500 persons. Find out coefficient of association between literacy and unemployment and conclude from it-

Illiterate unemployed	220
Literate employed	20
Illiterate employed	180

[Q = + .532]

40. Trace the association between darkness of eye colour of fathers and sons from the following figures-

Fathers with dark eyes and sons with dark eyes	50
Fathers with dark eyes but sons with not dark eyes	79
Fathers with not dark eyes but sons with dark eyes	89
Fathers with not dark eyes and sons with not dark eyes	782

[Q = + .695]

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय–8 प्रतीपगमन-विश्लेषण (REGRESSION-ANALYSIS)

NOTES

इकाई की रूपरेखा

8.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. प्रतीपगमन के अर्थ को जान सकेंगे।
2. प्रतीपगमन विश्लेषण की विभिन्न विधियों का ज्ञान होगा।
3. प्रतीपगमन एवं सह-संबंध में अंतर को जान सकेंगे।
4. प्रतीपगमन विश्लेषण के आधार पर अनुमान कर सकेंगे।

8.1 प्रस्तावना

दो या दो से अधिक पद श्रेणियाँ जो परस्पर सबन्धित हैं– क्या वे अपने सम्बन्धों पर कुछ प्रकाश डालती हैं? यदि हाँ, तो क्या उसे जाना जा सकता है ? यह प्रश्न हमारे मस्तिष्क में उठ सकते हैं। जैसे, एक वस्तु का कितना उत्पादन बढ़ाने से उसका कितना मूल्य कम हो जायेगा ? पूँजी के विनियोग में कितनी वृद्धि से कितना अधिक लाभ होगा ? कितने मजदूर लगाने से उत्पादन में कितनी वृद्धि होगी ? इन प्रश्नों का उत्तर सह-सम्बन्ध गुणक से नहीं मिलता। सांख्यिकीय अनुसन्धान का विश्लेषण करते समय ये उत्तर तो आवश्यक ही हैं, किन्तु हमें यह भी जानना है कि ये उत्तर कहाँ तक सही हैं ? इनका उत्तर प्रतीपगमन के मापों द्वारा मिलता है।

8.2 प्रतीपगमन का अर्थ व परिभाषा

सांख्यिकी में 'प्रतीपगमन' के प्रयोग का श्रेय श्री गाल्टन को है जिन्होंने प्रतीपगमन का सर्वप्रथम विश्लेषणात्मक अध्ययन किया। जब उन्होंने पिता और पुत्रों की ऊँचाइयों (heights) का अध्ययन किया तो उन्होंने देखा कि सामान्यतः व्यक्तिगत ऊँचाइयों का झुकाव औसत ऊँचाई की ओर है अर्थात् ऊँचे पिता के पुत्र कम ऊँचे और ठिगने पिता के पुत्र कम ठिगने हैं। इस प्रकार, पिता की ऊँचाइयों की अपेक्षा पुत्र की ऊँचाइयाँ औसत के अधिक निकट होती हैं। इसी प्रवृत्ति को प्रतीपगमन या वापस जाना कहते हैं।

जब दो पद श्रेणियाँ दी जाती हैं तब उनमें से एक स्वतन्त्र रूप से बदलती है अथवा हम यह मान लेते हैं कि स्वतन्त्र रूप से बदलती हैं और दूसरी पद श्रेणी इसलिए बदलती है कि प्रथम पद श्रेणी में परिवर्तन होता। यह पद श्रेणी भी या तो वास्तव में प्रथम पद श्रेणी में हुए परिवर्तन के कारण बदलती है अथवा हम यह मान लेते हैं कि दूसरी पद श्रेणी में होने वाले परिवर्तन प्रथम पद श्रेणी में हुए परिवर्तनों का परिणाम है। प्रथम पद श्रेणी स्वतन्त्र चल (Independent Variable) और दूसरी आश्रित चल (Dependent Variable) कहलाती है। जब दो चलों के मध्य कारण और प्रभाव का सम्बन्ध होता है तो प्रतीपगमन समीकरण (Regression Equation) की सहायता से एक मूल्य पर आधारित दूसरा मूल्य बड़ी सरलता से निकाला जा सकता है।

इस प्रकार **मौरिस मेयर्स ब्लेयर** ने प्रतीपगमन की परिभाषा इस प्रकार दी है–

"दो या अधिक पद श्रेणियों में दिये गये एकांकों की दृष्टि से,जो औसत सम्बन्ध हैं उसे ज्ञात करने के माप को प्रतीपगमन (Regression) कहते हैं।"

"Regression is the measure of the average relationship between two or more variables in terms of the original units of the data."

8.3 प्रतीपगमन विश्लेषण की उपयोगिता (Utility of Regression Analysis)

वर्तमान युग में प्रतीपगमन की धारणा केवल पितागत विशेषताओं के अध्ययन तक ही सीमित नहीं है। आजकल इसका प्रयोग उन सभी क्षेत्रों में किया जाता है जिनमें दो या अधिक सम्बन्धित श्रेणियों में विभिन्न पद मूल्यों की सामान्य माध्य की ओर वापस जाने की प्रवृत्ति पाई जाती है। इसके आधार पर आर्थिक, व्यावसायिक एवं सामाजिक क्षेत्र में विभिन्न घटनाओं के माध्य सम्बन्धों का विश्लेषण करके एक स्वतन्त्र पद-मूल्य से सम्बन्धित दूसरे आश्रित मूल्य का अनुमान लगाया जा सकता है। आर्थिक एवं व्यावसायिक जगत में प्रतीपगमन की बहुत अधिक उपयोगिता है। इसका प्रयोग प्रबन्धकों द्वारा व्यवसाय के नियन्त्रण-उपकरण के रूप में किया जाता है। इसके आधार पर उचित व्यावसायिक निर्णय लिए जा सकते हैं और उन्हें व्यावहारिकता की कसौटी पर कसा जा सकता है। इसके द्वारा यह अनुमान लगाया जा सकता है कि किसी वस्तु के उत्पादन या उसकी पूर्ति में निश्चित मात्रा में वृद्धि या कमी होने पर उसके मूल्य में क्या सम्भावित परिवर्तन होगा। मूल्यों के आधार पर माँग का, वर्षा की मात्रा, खाद, बीज आदि के आधार पर कृषि-उत्पादन का और पूँजी के आधार पर लाभ का अनुमान लगाने में प्रतीपगमन विश्लेषण बहुत उपयोगी सिद्ध होता है। व्यवसाय एवं व्यापार की सफलता के लिए ऐसे अनुमान आवश्यक होते हैं।

सह-सम्बन्ध एवं प्रतीपगमन में अन्तर– (1) **सम्बन्ध की मात्रा एवं प्रकृति**– सह-सम्बन्ध के माप से दो या अधिक पद मूल्यों (चरों) में औसत सम्बन्ध की मात्रा ज्ञात होती है जबकि प्रतीपगमन से इस सम्बन्ध की प्रकृति स्पष्ट होती है और यह भी पता चलता है कि एक चर के औसत मूल्य का प्रभाव दूसरे आश्रित चर के औसत मूल्य पर क्या होगा?

(2) **कारण-परिणाम सम्बन्ध**– सह-सम्बन्ध विश्लेषण में दो चर मूल्यों में कारण परिणाम सम्बन्ध अधिक स्पष्ट रूप से व्यक्त किया जाता है। दो चरों में अत्यधिक मात्रा का सह-सम्बन्ध होने से यह निश्चित रूप से नहीं कहा जा सकता कि एक कारण है और दूसरा परिणाम लेकिन प्रतीपगमन में ऐसा सम्भव है, क्योंकि उसमें एक चर (variable) स्वतन्त्र माना जाता है और दूसरा आश्रित जिसका अनुमान लगाया जाता है।

अपनी प्रगति की जाँच करें
Test your Progress

8.4 प्रतीपगमन विश्लेषण के प्रकार (Kind of Regression Analysis)

प्रतीपगमन की तकनीक से स्वतन्त्र चर (Variable) के आधार पर आश्रित चर का अनुमान लगाया जाता है। स्वतन्त्र चर को 'x' तथा आश्रित चर को 'y' माना जाता है।

प्रतीपगमन विश्लेषण दो प्रकार का हो सकता है–

(i) रेखीय एवं वक्रीय प्रतीपगमन (Linear and Curvilinear Regression)

दो परस्पर संबंधित समंक श्रेणियों में प्रतीपगमन का अध्ययन ज्यादातर बिन्दुरेखीय पद्धति से किया जाता है। x तथा y श्रेणियों के चल-मूल्यों के बिन्दुरेख-पत्र पर प्रांकित करने से एक विक्षेप चित्र बन जाता है। विक्षेप चित्र (Scatter Diagram) पर प्रांकित बिन्दुओं के मध्य से गुजरने वाली दो सर्वोपयुक्त रेखाएँ खींची जा सकती हैं। इन्हें

प्रतीपगमन रेखाएँ कहते हैं। यदि ये रेखाएँ सरल होती हैं तो इन्हें रेखीय प्रतीपगमन (Linear Regression) कहते हैं। यदि ये रेखाएँ वक्र (curve) के रूप में होती हैं तो इन्हें वक्रीय प्रतीपगमन कहते हैं।

(ii) सरल एवं बहुगुणी प्रतीपगमन (Simple and Multiple Regression)

जब केवल दो चरों के बीच प्रतीपगमन का अध्ययन किया जाता है तो यह सरल प्रतीपगमन कहलाता है। इन दो चरों में एक स्वतन्त्र चर होता है और दूसरा आश्रित। जब प्रतीपगमन का अध्ययन दो से अधिक चरों में किया जाता है तो वह बहुगुणीय-प्रतीपगमन कहलाता है। इसमें दो या अधिक स्वतन्त्र चर होते हैं और आश्रित चर केवल एक।

प्रतीपगमन रेखाएँ
(Lines of Regression)

मोटे तौर पर प्रतीपगमन रेखा उस रेखा को कहते हैं जो यह बतलाये कि एक चल में यदि एक एकक की घट-बढ़ होती है तो दूसरे चल में कितने एककों की घटबढ़ होगी। यह सम्बन्ध औसत रूप में (on an average) होगा। जैसे, इसमें यह बतलाया जायेगा कि यदि एक इंच ऊँचाई बढ़ती है तो औसत वजन कितना बढ़ेगा ? इसके लिए यह जरूरी है कि दोनों पद श्रेणियों में एक की अपेक्षा अधिक पद हों। प्रतीपगमन माध्यों पर आधारित है। पदों की केवल एक जोड़ी की सहायता से प्रतीपगमन रेखा नहीं निकाली जा सकती। यदि दो पद श्रेणियाँ x और y लें, तो औसत सम्बन्ध प्रकट करने वाली दो प्रतीपगमन रेखाएँ बनेंगी। इन रेखाओं की सहायता 'y' का मूल्य ज्ञात किया जा सकता है। इन रेखाओं को सर्वोत्कृष्ट उपयुक्त रेखाएँ (Lines of the best fit) भी कहते हैं।

इन रेखाओं की सहायता से सह-सम्बन्ध जानने के लिए निम्नलिखित नियम हैं–

(1) यदि दोनों रेखाएँ एक दूसरी को पूर्ण रूप से ढँक लें तो पद श्रेणियों में पूर्ण सह-सम्बन्ध माना जायेगा।

(2) यदि ये रेखाएँ एक दूसरे को समकोण पर काटें तो सह-सम्बध की मात्रा शून्य होगी।

(3) रेखाएँ एक-दूसरे के जितनी निकट होंगी, सह-सम्बन्ध की मात्रा उतनी ही अधिक होगी।

(4) रेखाएँ एक दूसरे से जितनी दूर होंगी, सह-सम्बन्ध की मात्रा उतनी ही कम होगी।

(5) किसी भी चल बिन्दु पर विचरण का अनुपात (Ratio of variation) क्या है ?

प्रतीपगमन रेखाएँ दो क्यों होती हैं
(Why there are two Regression Lines)

प्रतीपगमन के दो समीकरण होते हैं और उन पर आधारित दो प्रतीपगमन रेखाएँ होती हैं। एक 'x' का 'y' पर प्रतीपगमन दर्शाती है और दूसरी 'y' का 'x' पर। अन्य शब्दों में एक रेखा 'x' का सर्वोत्तम संभावित मान बतलाने के लिए जबकि 'y' का मान दिया हुआ है और दूसरी 'y' का मान बताने के लिए जबकि 'x' का औसत मान दिया हुआ है। यह प्रश्न उठता है कि क्या एक ही रेखा से दोनों मान नहीं बतलाये जा सकते हैं जैसाकि साधारण बिन्दु देखा में होता है ? इसका उत्तर है– नहीं। इसका कारण स्पष्ट है कि प्रतीपगमन एक श्रेणी में होने वाले विचलनों की तुलना में दूसरी श्रेणी में होने वाले विचलन बतलाता है। इन दोनों श्रेणियों में से एक को स्वतन्त्र और दूसरी को आश्रित मानते हैं और स्वतन्त्र श्रेणी में होने वाले प्रत्येक इकाई परिवर्तन के लिए आश्रित श्रेणी में होने वाले परिवर्तन की मात्रा मालूम करते हैं। यदि स्वतन्त्र एवं आश्रित श्रेणी सदैव स्थिर हो जैसा कि समय श्रेणियों में होता है तो एक ही प्रतीपगमन रेखा की आवश्यकता होती। हम अपनी आवश्यकतानुसार कभी एक श्रेणी को स्वतन्त्र और कभी दूसरी को स्वतन्त्र मानते रहते हैं। ऐसी स्थिति में दो प्रतीपगमन रेखाएँ ही बनेंगी। उदाहरण के लिए, मूल्य और माँग दो श्रेणी में से यदि हम मूल्य में परिवर्तन के फलस्वरूप माँग में होने वाले परिवर्तन का अध्ययन करना चाहते हैं तो मूल्य की मात्रा स्वतन्त्र तथा माँग की मात्रा आश्रित होगी। इसके विपरीत यदि हम माँग में परिवर्तन के फलस्वरूप मूल्य पर पड़ने वाले प्रभाव का अध्ययन करना चाहते हैं तो पहले तथ्य यह बताने में असमर्थ होंगे। कारण और प्रभाव के तर्क को हमें ध्यान में रखना होगा। अतः यहाँ की माँग की मात्रा स्वतन्त्र होगी और मूल्य की मात्रा आश्रित। स्पष्ट है कि प्रतीपगमन रेखाएँ दो होती हैं।

प्रतीपगमन रेखाओं का खींचना

प्रतीपगमन रेखाएँ दो प्रकार से खींची जा सकती हैं–

(i) **मुक्त हस्त विधि (Free Hand Method)** तथा (ii) **प्रतीपगमन समीकरणों द्वारा (By Regression equations)**

प्रथम विधि उपयुक्त नहीं है, क्योंकि विभिन्न व्यक्तियों द्वारा इनकी रचना भिन्न-भिन्न हो सकती है। अतः द्वितीय विधि-प्रतीपगमन समीकरणों के आधार पर ही इन रेखाओं को खींचा जाता है।

प्रतीपगमन समीकरण (Regression Equation) :

प्रतीपगमन रेखा को बीजगणितीय रूप देने के लिए उसे प्रतीपगमन समीकरण के रूप में लिखा जा सकता है। प्रतीपगमन रेखाएँ दो होती हैं। अतः प्रत्येक रेखा के लिए प्रतीपगमन समीकरण होंगे।

(i) 'x' का 'y' पर प्रतीपगमन समीकरण (Regression equation of x on y) –

$$(x - ax) = r\frac{\sigma_x}{\sigma_y}(y - ay)$$

or

$$(x - \bar{x}) = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

(ii) 'y' का 'x' पर प्रतीपगमन समीकरण (Regression equation of y on x) –

$$(y - ay) = r\frac{\sigma_y}{\sigma_x}(x - ax)$$

or

$$(y - \bar{y}) = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

यदि हम 'x' का मूल्य ज्ञात करना चाहते हों तो प्रथम समीकरण का और यदि 'y' का मूल्य ज्ञात करना हो तो द्वितीय समीकरण का प्रयोग करें। इन समीकरणों में 'x' और 'y' का आशय 'x' और 'y' श्रेणियों के औसत (Average) से होता है।

Illustration 1 – Find the most likely price in Mumbai corresponding to the price of Rs. 70 at Kolkata from the following data –

कोलकाता में माध्य मूल्य (Average Price at Kolkata)	65
मुम्बई में माध्य मूल्य (Average Price at Mumbai)	67
कोलकाता का प्रमाप विचलन (Standard Deviation of Kolkata)	2.5
मुम्बई का प्रमाप विचलन (Standard Deviation of Mumbai)	3.5

Coefficient of Correlation is + .8, between the two prices of the commodity in the two towns. (दोनों नगरों में सह संबंध गुणक 0.8 है)

Solution – Let 'y' stand for Kolkata and 'x' for Mumbai. Hence, we are two calculate the regression of x on y according to the following equation.

$$(x - \bar{x}) = r\frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

or $$(x - 67) = +\ 0.8\,\frac{3.5}{2.5}(y - 65)$$

or $$(x - 67) = 1.12\,(y - 65)$$

or $$(x - 67) = 1.12y - 72.8$$

or $$(x - 67) = (1.12 \times 70) - 72.8$$

or $$x = 78.4 - 72.8 + 67$$

or $$x = 72.6$$

Thus the most likely price at Mumbai is Rs. 72.6 when the corresponding price at Kolkata is Rs. 70.

NOTES

Illustration 2 — Given the following data, calculate the expected value of y when x is 12. (निम्नलिखित समंकों से y का संभावित मूल्य निकालिए यदि x = 12 हो)

		x	y
Average (समान्तर माध्य)	:	8	5
Standard Deviation (प्रमाप विचलन)	:	4	3

Coefficient of Correlation between x and y = + .99

(x तथा y में सह सम्बन्ध गुणक)

Solution —

Regression of y on x

$$(y - \bar{y}) = r\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

or $$(y - 15) = .99 \times \frac{3}{4}(x - 8)$$

or $$(y - 15) = .7425\,(x - 8)$$

or $$(y - 15) = .7425x - 5.94$$

or $$(y - 15) = .7425 \times 12 - 5.94$$

or $$y = 8.91 - 5.94 + 15 = 17.97$$

नोट—x पर y का प्रतीपगमन समीकरण इस प्रकार से निकाला जा सकता है—

जैसे $y = a + bx$ यहाँ 'a' और 'b' सामान्य समीकरणों (Normal Equation) द्वारा ज्ञता किये जाते हैं जो निम्न हैं—

(i) $$\Sigma y = na + b\Sigma x$$

(ii) $$\Sigma xy = a\Sigma x + b\Sigma x^2$$

इस भाँति 'y' पर 'x' का प्रतीपगमन समीकरण भी निकाला जा सकता है।

8.5 प्रतीपगमन गुणक (Coefficient of Regression)

जैसा कि बतलाया जा चुका है, प्रतीपगमन समीकरण (Regression Equation) दो सम्बन्धित पद श्रेणियों के माध्य के रूपों में प्रस्तुत किये जाते हैं या वे माध्य पर आधारित रहते हैं। इनके द्वारा यह प्रकट होता है कि एक पद-श्रेणी में उसके माध्य से होने वाले विचलन की तुलना में दूसरी पद श्रेणी में उसके माध्य से कितना विचलन होता है। प्रतीपगमन गुणक यह बतलाता है कि एक पद श्रेणी में एक पद के परिवर्तन से दूसरी पद श्रेणी में कितना परिवर्तन होगा। जिस प्रकार प्रतीपगमन समीकरण दो होते हैं उसी प्रकार दो प्रतीपगमन गुणक भी होते हैं। इन्हें निम्नलिखित सूत्र द्वारा ज्ञात किया जाता है—

(i) **'x' का 'y' पर प्रतीपगमन गुणक** (Regression Coefficient of x on y)—

$$bxy = r\frac{\sigma_x}{\sigma_y}$$

(ii) **'y' का 'x' पर प्रतीपगमन गुणक** (Regression Coefficient of y on x)—

$$bxy = r\frac{\sigma_y}{\sigma_x}$$

यदि इन दोनों प्रतीपगमन गुणकों का गुणनफल निकालकर उसका वर्गमूल निकाल लिया जाये तो सह-सम्बन्ध गुणक ज्ञात हो जायेगा। इसका सूत्र इस प्रकार होगा—

$$r = \sqrt{\left(r\frac{\sigma_x}{\sigma_y} \times r\frac{\sigma_y}{\sigma_x}\right)}$$

$$r = \sqrt{(bxy \times byx)}$$

Illustration 3 – Given the following data, calculate Regression Coefficient of x on y and of y on x – (निम्नलिखित समंकों से y पर आधारित x तथा x पर आधारित y के प्रतीपगमन गुणक निकालिए)–

	x	y
Average :	8	15
Standard Deviation :	4	3

Coefficient of Correlation between x and y = + .99

Solution–

(i) Regression coefficient of x on y – (ii) Regression Coefficient of y on x

$$bxy = r\frac{\sigma_x}{\sigma_y} \qquad byx = r\frac{\sigma_y}{\sigma_x}$$

$$= .99\,\frac{4}{3} \qquad = .99\,\frac{3}{4}$$

$$= .99 \times 1.33 \qquad = .99 \times .75$$

$$= 1.3167 \qquad = .7425$$

अत: $x = 1.3167y$ और $y = .7425x$

इसका अर्थ यह हुआ कि 'y' में 1 का परिवर्तन होने से 'x' में 1.3167 का और 'x' में एक का परिवर्तन होने से y में .7425 का परिवर्तन होगा।

इन गुणांकों की सहायता से सह-संबंध गुणक इस प्रकार जाना जा सकता है–

$$r = \sqrt{(bxy \times byx)} = \sqrt{(1.3167 \times .7425)} = .99$$

अत: यहाँ यह सिद्ध होता है कि दोनों प्रतीपगमन गुणकों का गुणाकार 1 से अधिक नहीं आना चाहिए। यदि यह 1 से अधिक हुआ तो सह-संबंध गुणक भी एक से अधिक होगा, लेकिन यह असम्भव है।

यदि केवल प्रतीपगमन गुणक निकालना हो तो सह-संबंध गुणक निकालने की आवश्यकता नहीं रहती है। यह निम्न सूत्र की सहायता से निकाला जा सकता है–

(i) 'x' का 'y' पर प्रतीपगमन गुणक (Regression Coefficient of x on y)

$$bxy = \frac{\Sigma xy}{\Sigma y^2}$$

(ii) 'y' का 'x' पर प्रतीपगमन गुणक (Regression Cocfficient of y on x)–

$$byx = \frac{\Sigma xy}{\Sigma x^2}$$

where,

Σxy = x और y पद श्रेणियों की माध्य से प्राप्त विचलनों के गुणनफलों का योग–

Σy^2 = 'y' पद श्रेणी के समान्तर माध्य से ज्ञात विचलनों के वर्गों का योग।

Σx^2 = 'x' पद श्रेणी के समान्तर माध्य से ज्ञात विचलनों के वर्गों का योग।

Illustration 4 – Calculate Regression Coefficients from the following data (निम्नलिखित समंकों से दोनों प्रतीपगमन गुणक ज्ञात कीजिए)–

x	:	1	2	3	4	5
y	:	1	4	9	16	25

NOTES

Solution—

m1	X-series deviations from Mean (3) (x)	x^2	m2	Y-series deviations from Mean (11) (y)	y2	xy
1	–2	4	1	–10	100	20
2	–1	1	4	–7	49	7
3	0	0	9	–2	4	0
4	+1	1	16	+5	25	5
5	+2	4	25	+14	196	28
15		**10**	**55**		**374**	**60**

(i) Regression Coefficient of x on y – (ii) Regression Coefficient of y on x

$$bxy = \frac{\Sigma xy}{\Sigma y^2} = \frac{60}{374} = 0.16 \qquad byx = \frac{\Sigma xy}{\Sigma x_w} = \frac{60}{10} = 6$$

लघुरीति (Shortcut Method) – उपरोक्त उदाहरण में माध्य संख्या पूर्णांक होने से प्रतीपगमन गुणक सरलता से निकाले जा सकते हैं। यदि माध्य संख्या पूर्णांक में न आये तो ऋजुरीति से गुणक निकालने में कठिनाई होगी। ऐसे समय लघुरीति का प्रयोग ही करना चाहिए। लघुरीति में निम्न सूत्र का प्रयोग होता है–

(i) Regression Coefficient of x on y –

$$bxy = \frac{\Sigma\, dx\, dy - \left(\dfrac{\Sigma\, dx \times \Sigma\, dy}{n}\right)}{\Sigma\, d^2 y - \dfrac{(\Sigma\, dy)^2}{n}}$$

(ii) Regression Coefficient of y on x –

$$byx = \frac{\Sigma\, dx\, dy - \left(\dfrac{\Sigma\, dx \times \Sigma\, dy}{n}\right)}{\Sigma\, d^2 x - \dfrac{(\Sigma\, dx)^2}{n}}$$

उदाहरण नं. 4 को लघुरीति से इन सूत्रों द्वारा हल किया जा सकता है।

8.6 अनुमान का प्रमाप विभ्रम (Standard Error of the Estimate)

यह बतलाया जा चुका है कि दी गई पद श्रेणी को सरल रेखा के रूप में प्रस्तुत करने के लिए निर्भर चल (Dependent Variable) के मूल्य अल्पतम वर्ग पद्धति (Method of Least Squares) द्वारा ज्ञात किये जाते हैं। निर्भर चल के इस प्रकार निकाले गये मूल्यों में और वास्तविक मूल्यों में अन्तर रह सकता है। इस प्रकार जो विभ्रम होता है उसे अनुमान का प्रमाप विभ्रम कहते हैं। इसे ज्ञात करने के लिए निम्न सूत्र का प्रयोग किया जाता है–

(i) $$Sx = \sqrt{\left(\frac{(x - x^1)^2}{N}\right)}$$

(ii) $$Sy = \sqrt{\left(\frac{(y - y^1)^2}{N}\right)}$$

where, "Sy" and "Sx" = Standard Error of Estimate

x and y = Actual value of x and y

x^1 and y^1 = Estimated value of x and y

N = Number of items.

Illustration 5 – Compute the Standard Error of Estimate for the Illustration No. 4 (उदाहरण नं. 4 के लिए अनुमान का प्रमाप विभ्रम ज्ञात कीजिए।)

Solution – (i) Calculation of Regression Equation of y on x by the Method of Least Squares–

x	y	xy	x^2	y^2
1	1	1	1	1
2	4	8	4	16
3	9	27	9	81
4	16	64	16	256
5	25	125	25	625
15	**55**	**225**	**55**	**979**

Substituting the above values in the two normal equations–

$$\Sigma y = na + b\Sigma x \qquad .....(1)$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \qquad .....(2)$$

we get,

$$55 = 5(a) + b(15)$$

or $55 = 5a + 15b$ .....(i)

and $225 = a(15) + b(55)$

or $225 = 15a + 55b$ .....(ii)

Multiplying equation (i) by 3, we have

$165 = 15a + 45b$ .....(iii)

Subtracting equation (iii) from equation (ii), we have,

$$60 = 10\,b$$

or $b = 6$

Substituting the value of 'b' in equation (i), we have,

$$55 = 5a + 15 \times 6$$

or $55 = 5a + 90$

or $55 - 90 = 5a$

or $a = -7$

Substituting the above values in the Regression equation of y on x : viz.

$$y = a + bx$$

we get,

$$y = -7 + 6x$$

अब निम्न प्रकार से 'y' के सम्भाव्य मूल्य ज्ञात कर सकते हैं।

If $x = 1$ $y = -1$

$x = 2$ $y = 5$

$x = 3$ $y = 11$

$x = 4$ $y = 17$

$x = 5$ $y = 23$

NOTES

(ii) **Calculation of Standard Error of Estimate—**

x	y	y^1 (Estimated Value of y)	y – y1	$(y-y^1)^2$
1	1	–1	2	4
2	4	5	–1	1
3	9	11	–2	4
4	16	17	–1	1
5	25	23	2	4
N = 5				14

$$Sy = \sqrt{\left(\frac{(y-y^1)^2}{n}\right)} = \sqrt{\left(\frac{14}{5}\right)} = \sqrt{(2.8)} = 1.67$$

Illustration 6 – A regression equation explaining the average relationship between the dividend per share and the price per share in 1990 for 100 corporations was,

$$y = \text{Rs. } 5.49 + 12.14x$$

Estimate the value of a share of stock which pays a dividend of Rs. 5 per share. The Standard error of the estimate.

(Sy) is = Rs. 4.5

Solution– $y = \text{Rs. } 5.49 + 12.14x$

$$= \text{Rs. } 5.49 + 12.14 \times 5 = \text{Rs. } 66.19$$

(i) There is 68.3% chance that its values will be (Rs. 66.19 ± 1 × 4.5 *Sy*) i.e. 70.69 and Rs. 61.69

(ii) There is 95.4% chance that its value will be Rs. 66.19 ± 2 × 4.5 i.e. Rs. 75.19 and Rs. 57.19

(iii) There is 99.7% chance that its value will be Rs. 66.19 ± 3 × 4.5 i.e. Rs. 79.69 and Rs. 52.69.

Illustration 7 – (a) Given

	x-series	y-series
Mean (समान्तर माध्य) :	18	100
Standard Derivation (प्रमाप विचलन) :	14	20

Coefficient of Correlation between x and y = +0.8
(x तथा y के मध्य सह-संबंध गुणक)

Find the most probable of y if x is 70 and most probable value of x if y is 90. (y का संभावित मूल्य निकालिए यदि x = 70 हो तथा x का संभावित मूल्य निकालिए यदि y = 90 हो)

(b) If two regression Coefficient are 0.8 and 0.6, what would be the value of the Coefficient of Correlation. (यदि दो प्रतीपगमन गुणक 0.8 तथा 0.6 हों तो सह-संबंध गुणक क्या होगा ?)

Solution–

(a) (i) Regression of x on y

$$(x-\bar{x}) = r\frac{\sigma_x}{\sigma_y}(y-\bar{y})$$

or $(x-18) = 0.8 \times \frac{14}{20}(y-100)$

or $(x-18) = .56\,(90-100)$

or $(x-18) = .56 \times -10$

(ii) Regression of y on x

$$(y-\bar{y}) = r\frac{\sigma_y}{\sigma_x}(x-\bar{x})$$

or $(y-100) = .8 \times \frac{20}{14}(70-18)$

or $(y-100) = 1.143\,(52)$

or $y = 100 + 1.143\,(52)$

NOTES

$$x = 18 - 5.6$$

$$= 12.4$$

$$= 159.436$$

(b) $r = \sqrt{[(bxy \times byx)]}$

$$r = \sqrt{[(0.8 \times 0.6)]} = 0.69$$

Illustration 8 – In a partially destroyed Laboratory-record of an analysis of correlation data the following results only are legible–

Variance of x= 9

Regression equations are

(i) $8x - 10y = -66$ (ii) $40x - 18y = 214$

What are the (a) Mean Values of x and y, (b) Coefficient of Correlation between x and y (c) the σ of y.

आंशिक रूप से नष्ट हुए एक प्रयोगशाला के रिकार्ड से सह संबंध समंकों के विश्लेषण संबंधी निम्नलिखित निष्कर्ष ही उपलब्ध हैं–

x का विचरण मापांक = 9

प्रतीपगमन समीकरण;

(i) $8x - 10y = -66$ (ii) $40x - 18y = 214$

गणना कीजिए (i) x तथा y का समांतर माध्य

(ii) x तथा y के मध्य सह संबंध गुणक (iii) y का प्रमाप विचलन

Solution – (a) Calculation of Mean Values of x and y –

$$8x - 10y = -66 \quad .....(i)$$

$$40x - 18y = 214 \quad .....(ii)$$

Multiplying equation (i) by 5, we have,

$$40x - 50y = -330 \quad .....(iii)$$

Deducting (iii) from (ii), we have,

$$32y = 544$$

$$y = \frac{544}{32} = 17$$

Substituting the value of y in equation (i), we have,

$$8x - 10 \times 17 = -66$$

or $8x = -66 + 170$

or $8x = 104$ or $x = 13,$

$\therefore$ Mean of x = 13, Mean of y = 17

(b) Coefficient of correlation between x and y –

$$r = \sqrt{(bxy \times byx)}$$

The value of bxy and byx will be found out as under–

From the first equation, we get,

$$10y = 8x + 66$$

$$y = .8x + 6.6 \quad \text{or} \quad byx = .8$$

From the second equation, we get,

$$40x = 18y + 214$$

$x = .45y + 5.35$ or $bxy = .45$

Hence, $r = \sqrt{(bxy \times byx)} = \sqrt{(.45 \times .8)} = \sqrt{(.360)} = 0.6$

(c) Calculation of σ of y–

Variance or $\sigma_x^2 = 9$ or 'σ_x' $= 3$

Regression coefficient of x on y–

$$bxy = r\frac{\sigma_x}{\sigma_y} = .45$$

or $$= .6\frac{3}{\sigma_y} = .45$$

or $$.6 \times 3 = .45 \times \sigma_y$$

or $$1.8 = .45\sigma_y$$

or $$\sigma_y = \frac{1.8}{.45} = 4$$

Illustration 9 – For certain data y = 1.3x and x = 0.7y are the regression lines. Compute the Coefficient of Correlation between x and y. (किन्हीं समंकों के लिए y = 1.3x तथा x = 0.7y दो प्रतीपगमन रेखाएँ हैं। x तथा y के मध्य सह संबंध गुणक की गणना कीजिए)

Solution– $y = 1.3x$ or $byx = 1.3$

$x = 0.7y$ or $bxy = 0.7$

$r = \sqrt{(bxy \times bxy)} = \sqrt{(1.3 \times 0.7)}$

Illustration 10–(General) Write down the two regression equation associated with the following pairs of values–

(x)	:	152	114	138	154	144	153	141	117	136	154
(y)	:	193	300	414	594	676	549	320	483	481	659

उपरोक्त मूल्यों के दो युग्मों से संबंधित दो प्रतीपगमन समीकरण बनाइये।

Solution–Calculation of the Regression Coefficients :

x			y			dxdy
x	dx(130)	d^2x	y	dy(500)	d^2y	
152	+22	484	193	–307	94249	–6754
114	–16	256	300	–200	40000	+3200
138	+8	64	414	–86	7396	–688
154	+24	576	594	+94	8836	+2256
144	+14	196	676	+176	30976	+2464
153	+23	529	549	+49	2401	+1127
141	+11	121	320	–180	32400	–1980
117	–13	169	483	–17	289	+221
136	+6	36	481	–19	361	–114
154	+24	576	659	+159	25281	+2816
	+103	**3007**		**–331**	**2,42,189**	**3548**

Regression Coefficient of x on y, or

$$bxy = r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma\, dx\, dy \left(\frac{\Sigma\, dx \times \Sigma\, dy}{n}\right)}{\Sigma\, d^2 y - \frac{(\Sigma\, dy)^2}{n}}$$

$$= \frac{3548 - \left(\frac{103 \times (-331)}{10}\right)}{242189 - \frac{(-331)^2}{10}}$$

$$= \frac{3548 + 3409.3}{242189 - 10956.1} \qquad = \frac{6957.3}{231232.9} = 0.03$$

Similarly, Regression Coefficient of y on x or

$$byx \quad = r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma\, dx\, dy - \left(\frac{\Sigma\, dx \times \Sigma\, dy}{n}\right)}{\Sigma\, d^2x - \frac{(\Sigma\, dx)^2}{n}}$$

$$= \frac{3548 - \left(103 \times \frac{(-331)}{10}\right)''}{3007 - \frac{(103)^2}{10}} = \frac{6957.3}{1946.1} = 3.575$$

$$\text{Ax or x} \quad = x + \frac{\Sigma\, dx}{n} \qquad \text{Ay} \quad \text{or} \quad \text{y} \quad = y + \frac{\Sigma\, dy}{n}$$

$$= 130 + \frac{103}{10} = 140.3 \quad = 500 + \frac{-331}{10} = 466.9$$

Substituting the above values in the following regression equation of x on y

$$(x - \bar{x}) \quad = bxy\,(y - \bar{y})$$

$$(x - 104.3) \quad = .03\,(y - 466.9)$$

$$x \quad = .03y - 14.007 + 140.3$$

$$x \quad = .03y + 126.293$$

Similarly, the regression equation of y on x is

$$(y - \bar{y}) \quad = byx\,(x - \bar{x})$$

$$(y - 466.9) \quad = 3.575\,(x - 140.3)$$

$$y \quad = 3.575x + 466.9 - 501.57$$

$$y \quad = 3.575x - 34.67$$

Thus, the two regression equation that may be associated with the given pairs of values are—

$$x \quad = .03y + 126.293$$

$$y \quad = 3.575x - 34.67$$

Illustration 11—Given,

$$x = .85y; \quad y = .89x; \quad \sigma_x = 3$$

Find out (i) σ_y (ii) r.

Solution—From the given data

Regression of x on y or bxy $= r\frac{\sigma_x}{\sigma_y} = .85$

Regression of y on x or byx $= r\frac{\sigma_y}{\sigma_x} = .89$

Substituting these values in the following formula

$$r = \sqrt{(bxy \times byx)}$$

NOTES

we get, $r - \sqrt{(.85 \times .89)} = \sqrt{(.7565)} = .87$

Similarly, $r = \frac{\sigma_x}{\sigma_y} = .85$

or $.85\,\sigma_y = r \times \sigma_x$

or $\sigma_y = \frac{r \times \sigma_x}{.85}$

$= \frac{.87 \times 3}{.85} = 3.07$

Illustration 12. The following table gives the ages of husbands and wives for 50 newly married couples. Find the two regression lines. Also estimate (a) the age of husband when wife is 20 and (b) age of wife when husband is 30.

निम्नलिखित तालिका में 50 नवविवाहित दम्पत्तियों के आयु संबंधी आँकड़े दिये हुए हैं। दोनों प्रतीपगमन रेखायें प्राप्त कीजिए और ज्ञात कीजिए (i) पति की आयु जबकि पत्नी की आयु 20 वर्ष हो तथा (ii) पत्नी की आयु जबकि पति की आयु 30 वर्ष है।

Age of wives (पत्नियों की आयु)	**Age of husbands (HeefleÙeeW keâer DeeÙeg)**			**Total**
	20-25	**25-30**	**30-35**	
16–20	9	14	–	23
20–24	6	11	3	20
24–28	–	–	7	7
Total	**15**	**25**	**10**	**50**

Solution—

Ages of wives (y)			Age of husbands (x)			Total (f)	fdy	fd^2y	fdxdy
			20–25	25–30	30–35				
			22.5	27.5	32.5				
			–5	0	+5				
16–20	18	–4	20 9 180	0 14 0	–	23	–92	368	180
20–24	22	0	0 6 0	0 11 0	0 3 0	20	0	0	0
24–28	26	+4	–	–	20 7 140	7	28	112	140
Total (f)			15	25	10	50	–64	480	320
fdx			–75	0	50	–25			
fd^2x			375	0	250	625			
fdxdy			180	0	140	320			

(i) **Actual Means :**

X

$a_x = \bar{x} + \frac{\Sigma fdx}{n}$

$= 27.5 + \frac{-25}{50}$

$= 27.5 - 0.5 = 27$

Y

$a_y = \bar{x} + \frac{\Sigma fdy}{n}$

$= 22 + \frac{-64}{50}$

$= 22 - 1.28 = 20.72$

(ii) **Regression Coefficients :**

x on y :

$$bxy = \frac{\Sigma f\,dx\,dy - \left(\frac{\Sigma f\,dx \times \Sigma f\,dy}{n}\right)}{\Sigma f d^2 y - \frac{(\Sigma f\,dy)^2}{n}}$$

$$= \frac{320 - \frac{(-25 \times -64}{50}}{480 - \frac{(-64)^2}{50}}$$

$$= \frac{320 - 32}{480 - 81.92} = \frac{288}{398.08} = 0.7$$

y on x :

$$byx = \frac{\Sigma f\,dx\,dy - \left(\frac{\Sigma f\,dx \times \Sigma f\,dy}{n}\right)}{\Sigma f d^2 x - \frac{(\Sigma f\,dx)^2}{n}}$$

$$= \frac{320 - \frac{(-25 \times -64)}{50}}{625 - \frac{(-25)^2}{50}}$$

$$= \frac{320 - 32}{625 - 12.5} = \frac{288}{612.5} = 0.47$$

(iii) **Coefficient or Correlation (r) :**

$$r = \sqrt{bxy \times byx}$$

$$= \sqrt{0.7 \times 0.47} = \sqrt{0.329} = 0.57$$

(iv) **Regression equation of x on y :**

$$(x - \bar{x}) = bxy\,(y - \bar{y})$$

$$(x - 27) = 0.7\,(y - 20.72)$$

$$(x - 27) = 0.7y - 14.504$$

$$x = 0.7y - 14.504 + 27$$

$$x = 0.7y + 12.496$$

$$x = 0.7 \times 20 + 12.496$$

$$x = 14 + 12.496$$

$$x = 26.496$$

अत: पति की आयु 26.5 होगी जबकि पत्नी की आयु 20 वर्ष है।

(v) **Regression equation of y on x :**

$$(y - \bar{y}) = byx\,(x - \bar{x})$$

$$(y - 20.72) = 0.47\,(x - 27)$$

$$(y - 20.72) = 0.47x - 12.69$$

$$y = 0.47x - 12.69 + 20.72$$

$$y = 0.47x + 8.03$$

$$y = 0.47 \times 30 + 8.03$$

$$y = 14.10 + 8.03$$

$$y = 22.13$$

अत: पत्नी की आयु 22.13 वर्ष होगी जबकि पति की आयु 30 वर्ष है।

Illustration 13 : The following regression equations were obtained in a certain investigation. Find the mean of x and y and coefficient of correlation.

(किसी अनुसंधान में निम्नलिखित समीकरण प्राप्त हुए। x तथा y का समान्तर माध्य तथा सह संबंध गुणक निकालिए।)

$$x = 19.13 - 0.87y$$

NOTES

$$y = 11.64 - 0.50x$$

Solution :

Mean of x and y

$$x + 0.87y = 19.13 \quad .....(i)$$

$$0.5x + y = 11.64 \quad .....(ii)$$

Multiplying (ii) equation by 2, we have

$$x + 2y = 23.28 \quad .....(iii)$$

Substracting equation (iii) from (i) we have

$$1.13y = 4.15$$

$$\therefore \quad y = \frac{4.15}{1.13} = 3.67$$

Now Substituting the value of y in equation (ii)

We get,

$$0.5x + 3.67 = 11.64$$

or $\quad 0.5x = 11.64 - 3.67$

or $\quad 0.5x = 7.97$

or $\quad x = \frac{7.97}{0.5} = 15.94$

Hence Mean of $x = 15.94$ and that of $y = 3.67$

Coefficient of Correlation (r) :

From the two equations given, we can note that

$$bxy = -0.87 \quad \text{and} \quad byx = -0.5$$

Hence, $\quad r = \sqrt{bxy \times byx}$

$$= \sqrt{-0.87 \times -0.5}$$

$$= -\sqrt{.435}$$

$$= -0.66$$

As the two regression coefficients are negative the coefficient of correlation is also negative.

बोध प्रश्न

1. प्रतीपगमन विश्लेषण से आप क्या समझते हैं।

....

....

....

2. प्रतीपगमन एवं सह-संबंध में अंतर लिखिए।

....

....

....

3. प्रतीपगमन रेखाएँ दो क्यों होती है एवं किन परिस्थितियों में प्रतीपगमन रेखा एक हो सकती है ?

....

....

....

8.7 विचरण का अनुपात (Ratio of Variation)

प्रायः हम दो पद-श्रेणियों के मध्य विचरण का अनुपात जानना चाहते हैं। जिससे हमें यह ज्ञात हो सके कि एक श्रेणी के परिवर्तन और दूसरी श्रेणी के परिवर्तन में क्या अनुपात है। यदि एक का परिवर्तन 1% है तो दूसरी का परिवर्तन 1% का कौन-सा भाग है। इस तरह, **विचरण का अनुपात आश्रित श्रेणी** (y) **में प्रधान श्रेणी** (x) **की तुलना में समान्तर माध्य से प्रतिशत विचलनों का माध्य अनुपात है।** [Ratio of Variation is the average ratio of the percentage deviations from the mean in the relative (y) as compared with those in the subject (x)] विचरण का अनुपात सह-संबंध गुणक का उपप्रमेय (Corollary) होता है। यह सम्भव है कि दो चलों में पूर्ण सह-सम्बन्ध हो किन्तु दोनों चलों का आनुपातिक उतार-चढ़ाव समान न हो। विचरण के अनुपात में हम इसी का अध्ययन करते हैं। यह अध्ययन काफी लाभदायक है। उदाहरणार्थ, एक अनाज-व्यापारी यह जानने को उत्सुक होता है कि अनाज के भाव कितने बढ़ेंगे यदि फसल में 10% की कमी है।

विचरण का अनुपात ज्ञात करने की विधियाँ

विचरण का अनुपात ज्ञात करने की दो विधियाँ हैं—

(1) गणितीय विधि (Mathematical Method),

(2) बिन्दु रेखीय विधि (Graphic method)।

(1) गणितीय विधि

इस विधि से विचरण के अनुपात की संगणना करने के लिए निम्न क्रियाएँ करनी पड़ती हैं—

(i) सर्वप्रथम वह निश्चित कीजिए कि कौन-सी श्रेणी को प्रधान और किसे आश्रित माना जाये। उस श्रेणी को प्रधान मानना उचित है जिसमें अधिक माध्य आनुपातिक विचरण हों। इसमें विचरण का अनुपात 1 की तुलना में सदा कम आयेगा।

(ii) प्रधान और आश्रित श्रेणियों का समान्तर माध्य निकालिये।

(iii) इन माध्यों से दोनों श्रेणियों के मूल्यों का विचलन (deviation) अलग-अलग निकालिये।

(iv) दोनों श्रेणियों में प्रत्येक पद का प्रतिशत विचलन निकालिये। इसके लिए निम्न सूत्र का प्रयोग होता है—

$$\frac{d}{a} \times 100$$

(v) आश्रित श्रेणी के प्रत्येक पद के प्रतिशत विचलन में प्रधान श्रेणी के उसके सामने वाले प्रतिशत विचलन का भाग देकर भजनफल निकालिये।

(vi) इन भजनफलों के योग में पदों के जोड़ों की संख्या का भाग दीजिये। प्राप्त भजनफल विचरण का अनुपात होगा।

Illustration 14— From the following data ascertain the ratio variation between the sales and profits.

Year	:	1998	1999	2000	2001	2002	2003
Sales (Lakhs)	:	72	84	66	60	48	42
Profits (Lakhs)	:	42	52	48	46	30	28
Year	:	2004	2005	2006	2007	2008	2009
Sales (Lakhs)	:	54	63	51	57	69	54
Profit (Lakhs)	:	36	38	34	42	44	40

NOTES

Solution —

Year	(Sales) X-series			(Profits) Y-series			Ratio
	Sales x	dx (60)	dx %	Profits y	dy (40)	dy (%)	dy % / dx %
1998	72	12	12/60 × 100 = 20	42	+2	6	0.25
1999	84	+24	24/60 × 100 = 40	52	+12	30	0.75
2000	66	+6	6/60 × 100 = 100	48	+8	20	2.00
2001	60	0	–12/60 × 100 = –20	46	+6	15	0
2002	48	–12	–18/60 × 100 = –30	30	–10	–25	1.25
2003	42	–18	–6/60 × 100 = –30	28	–12	–30	1.00
2004	54	–6	3/60 × 100 = +5	36	–4	–10	1.00
2005	63	+3	–3/60 × 100 = –15	38	–2	–5	–1.00
2006	51	–9	–3/60 × 100 = –5	34	–6	–15	1.00
2007	57	–3	9/60 × 100 = +15	42	+2	+5	–1.00
2008	69	+9	–6/60 × 100 = –10	44	+4	+10	0.67
2009	54	–6		40	0	0	0

$$\text{Ratio of Variation} = \frac{\frac{dy\%}{dx\%}}{N} = \frac{5.92}{12} = 0.494$$

Thus for a deviation of one unit in x-series, there is on an average a deviation of 0.494 in y-series.

(2) **बिन्दुरेखीय विधि** (Graphic Method)

पद-श्रेणियाँ, जो सामाजिक और आर्थिक समस्याओं से सम्बन्धित होती हैं, प्रायः नियमित नहीं होती हैं। ऐसी परिस्थिति में विचरण का अनुपात गणितीय विधि से ज्ञात करना उपयुक्त नहीं होता। यहाँ बिन्दुरेखीय विधि अधिक उपयुक्त है। इस विधि के जन्मदाता श्री फ्रांसिस गाल्टन थे। इस विधि में निम्न क्रियाएँ की जाती हैं—

(1) दोनों श्रेणियों के समान्तर माध्य निकालिए।

(2) दोनों श्रेणियों के समान्तर माध्य को आधार मानकर उसके प्रत्येक मूल्य को प्रतिशत में बदल लीजिये।

(3) प्रधान श्रेणी को उदग्र माप, अर्थात् कोटि-अक्ष (axis of y) पर प्रदर्शित कीजिए।

(4) आश्रित श्रेणी को क्षैतिज माप अर्थात् भुजाक्ष (axis of x) पर प्रदर्शित कीजिए।

(5) प्राप्त निर्देशांक को लिए गये माप के आधार पर बिन्दुरेख पत्र पर अंकित कीजिए।

(6) तब एक सर्वाधिक उपयुक्तता की रेखा खींचिए। इस रेखा को खींचने में यह ध्यान रखा जाये कि जहाँ तक सम्भव हो लगभग आधे अंकित बिन्दुरेखा की ओर और आधे दूसरी ओर हों। रेखा के दोनों ओर के बिन्दु यथासम्भव रेखा के समान दूरी पर हों तथा यह रेखा समान्तर माध्य से होकर गुजरनी चाहिए।

गाल्टन बिन्दुरेख का अध्ययन

(1) यदि दोनों श्रेणियों का विचरण एक ही अनुपात में होता है, तो अंकित बिन्दुरेख एक रेखा के रूप में होंगे। दोनों में पूर्ण सह-संबंध होगा।

(2) यदि रेखा बाईं ओर से दाहिनी ओर ऊपर उठी हो तो धनात्मक सह-संबंध होता है।

(3) यदि रेखा बाईं ओर से दाहिनी ओर नीचे की ओर झुकी हो, तो ऋणात्मक सह-संबंध होगा।

(4) यदि दोनों श्रेणियों में एक ही अनुपात में परिवर्तन होता है, तो यह रेखा आधार रेखा से 45° का कोण बनाती हुई जायेगी। यदि प्रधान श्रेणी में आश्रित श्रेणी की अपेक्षा विचलन की दर अधिक होती है तो यह कोण 45° से अधिक अन्यथा इससे कम होता है।

(5) यह रेखा समान विचरण रेखा से जितनी दूर होगी, सह-संबंध उतना ही कम होगा।

विचरण के अनुपात का माप

कोटि अक्ष (axis of y) पर कोई बिन्दु (A) ले लीजिए। इस बिन्दु से विचरण रेखा तक एक रेखा खींचिये। जहाँ तक रेखा विचरण रेखा को स्पर्श करे (B), वहाँ से भुजाक्ष (axis of x) पर लम्ब (C) खींचिये। अब, AB और BC की दूरी माप लीजिए। AB के माप में BC के माप का भाग दे दीजिये। भजनफल विचरण का अनुपात बतलायेगा।

Illustration 15—Show Ratio of Variation graphically from the following—

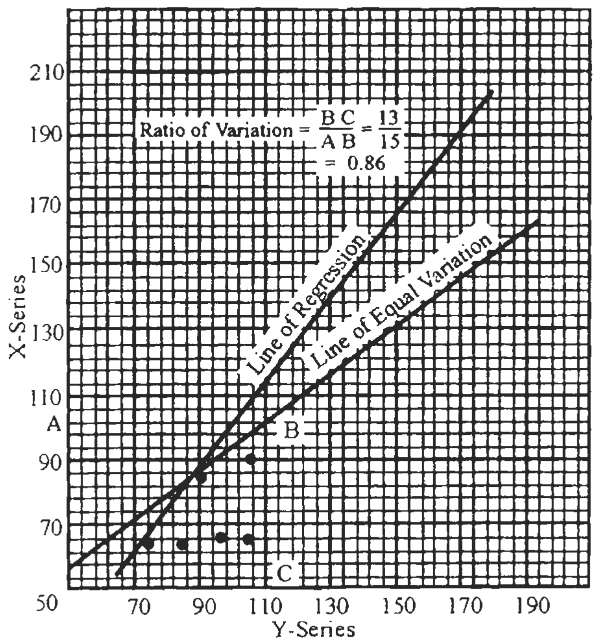
X-Series	12	20	18	25	18	13	12	17	26	39 = 200
Y-Series	140	180	180	190	210	190	160	180	250	320 = 2000

Solution.

x	Index of x (20 = 100)	y	Index of y (200 = 100)
12	60	140	70
20	100	180	90
18	90	180	90
25	125	190	95
18	90	210	105
18	65	190	95
12	60	160	80
17	85	180	90
26	130	250	125
39	195	320	160
200		2000	
a = 200		a = 200	

गाल्टन बिन्दुरेख की उपयोगितायें—(1) यह प्रदर्शित करता है कि दो श्रेणियों के सह-संबंध हैं या नहीं तथा कैसा सह-संबंध है।

(2) इसके द्वारा प्रतीपगमन का माप भी ज्ञात किया जा सकता है।



(3) विचरण के अनुपात का माप भी इसके द्वारा ज्ञात किया जा सकता है।

(4) इसमें प्रधान श्रेणी के किसी भी मूल्य का आश्रित श्रेणी में अनुमानित मूल्य का ज्ञान हो जाता है।

NOTES

8.8 बहुगुणी प्रतीपगमन (Multiple Regression)

जब एक आश्रित चर पर कई स्वतंत्र चरों के सामूहिक प्रभाव का अध्ययन करना होता है तो उसके लिए बहुगुणी प्रतीपगमन एक उचित रीति है। यह सरल प्रतीपगमन का ही विस्तृत रूप है। इसके द्वारा दो या दो से अधिक स्वतंत्र चर मूल्यों के आधार पर आश्रित चर के सर्वोत्तम मूल्य का अनुमान लगाया जाता है। उदाहरण के लिए, 20 हैक्टेयर भूमि में गेहूँ की उपज (x_1) का, वर्षा की मात्रा (x_2) और खाद की मात्रा (x_3) दिये हुए होने पर अनुमान एक निश्चित वर्षा एवं खाद की मात्रा के लिए बहुगुणी प्रतीपगमन विश्लेषण द्वारा लगाया जा सकता है।

सरल प्रतीपगमन में आश्रित चर के अनुमान एक रेखा में आते हैं, किन्तु जब दो स्वतंत्र चर होते हैं तो ये अनुमान एक क्षेत्र (plane) में आवेंगे।

बहुगुणी प्रतीपगमन समीकरण (Multiple Regression Equation)

यह समीकरण विविध चरों के औसत सम्बन्ध को व्यक्त करता है। इस औसत संबंध के आधार पर ही आश्रित चर के लिए सर्वोत्तम अनुमान लगाया जाता है।

यह समीकरण इस प्रकार लिखा जाता है-

$$y_c = a + b_1x_1 + b_2x_2 + b_3x_3 + ...$$

संकेताक्षर:

y_c = आश्रित चर का संगणित मान (Computed Value of dependent Variable)

x_1, x_2, x_3 = स्वतंत्र चर (Independent variables)

b_1, b_2, b_3 = प्रतीपगमन गुणक (Regression Coefficient)

गेहूँ की उपज (x_1) तथा वर्षा एवं खाद की मात्रा (x_2 तथा x_3) के लिए बहुगुणी प्रतीपगमन समीकरण इस प्रकार लिखा जावेगा:-

$$x_{1.23} = a_{1.23} + b_{12.3}x_2 + b_{13.2}x_3$$

यहाँ

x_1 = गेहूँ की उपज

$a_{1.23}$ = स्थिरांक

$b_{12.3}$ = आंशिक प्रतीपगमन गुणक जो x_2 में इकाई परिवर्तन का प्रभाव x_1 पर x_3 को स्थिर मानते हुए मापता है।

$b_{13.2}$ = आंशिक प्रतीपगमन गुणक जो x_3 में इकाई परिवर्तन का प्रभाव x_1 पर x_2 को स्थिर मानते हुए मापता है।

x_2 = वर्षा की मात्रा।

x_3 = खाद की मात्रा।

वास्तविक माध्य से विचलन लेने पर :

जब विचलन चरों के वास्तविक माध्य से लिये जाते हैं तो प्रतीपगमन समीकरण छोटा हो जाता है क्योंकि ऐसी स्थिति में $a_{1.23}$ स्थिरांक का मान शून्य हो जाता है। इस स्थिति में समीकरण होगा :

$$x_1 = b_{12.3}x_2 + b_{13.2}x_3$$

आंशिक प्रतीपगमन गुणकों की गणना के सूत्र :

$$b_{12.3} = \frac{r_{12} - r_{13}r_{23}}{1 - r_{23}^2} \times \frac{\sigma_1}{\sigma_2} \text{ या } = r_{12.3} \times \frac{\sigma 1.23}{\sigma 2.13}$$

$$b_{13.2} = \frac{r_{13} - r_{12}r_{23}}{r - r_{23}^2} \times \frac{\sigma_1}{\sigma_3} \text{ या } = r_{13.2} \times \frac{\sigma 1.23}{\sigma 3.12}$$

आंशिक प्रतीपगमन गुणकों को मालूम करके समीकरण के रूप में रख दिया जाता है और इस प्रकार x_1 का सर्वोत्तम मान निकाल लिया जाता है। इन गुणकों के आधार पर प्रतीपगमन समीकरण इस प्रकार होगा:-

Regression equation of x_1 on x_2 and x_3

$$x_{1.23} = a_{1.23} + b_{12.3}x_2 + b_{13.2}x_3$$

$$a_{1.23} = \bar{x}_1 - b_{12.3}\bar{x}_2 - b_{13.2}\bar{x}_3$$

आंशिक सह संबंध गुणकों तथा आंशिक प्रतीपगमन गुणकों में संबंध (Relation between Partial Correlation Coefficient and Partial Regression Coefficients):-

जिस प्रकार सह संबंध गुणक (r_{12}) दोनों प्रतीपगमन गुणकों के गुणनफल का वर्गमूल होता है, उसी प्रकार आंशिक सहसंबंध गुणक ($r_{12.3}$) दोनों आंशिक प्रतीपगमन गुणकों के गुणनफल का वर्गमूल होता है :-

$$r_{12.3} = \sqrt{b_{12.3} \times b_{21.3}} \quad \text{या} \quad r^2_{12.3} = b_{12.3} \times b_{21.3}$$

Illustration 16: Find the multiple linear Regression equation of x_1 on x_2 and x_3 from the data relating to three Variables given below :-

x_1	:	11	17	26	28	31	35	41	49	63	69
x_2	:	2	4	6	5	8	7	10	11	13	14
x_3	:	2	3	4	5	6	7	9	10	11	13

Solution: The regression equation of x_1 on x_2 and x_3 is

$$x_1 = a_{1.23} + b_{12.3}x_2 + b_{13.2}x_3$$

The Values of the Constants $a_{1.23}$, $b_{12.3}$ and $b_{13.2}$ are obtained by solving the following three normal equations:

$$\Sigma x_1 = Na_{1.23} + b_{12.3}\,\Sigma x_2 + b_{13.2}\,\Sigma x_3$$

$$\Sigma x_1 x_2 = a_{1.23}\,\Sigma x_2 + b_{12.3}\,\Sigma x_2^2 + b_{13.2}\,\Sigma x_2 x_3$$

$$\Sigma x_1 x_3 = a_{1.23}\,\Sigma x_3 + b_{12.3}\,\Sigma x_2 x_3 + b_{13.2}\,\Sigma x_3^2$$

x_1	x_2	x_3	$x_1 x_2$	$x_3 x_1$	$x_2 x_3$	x_2^2	x_3^2	x_1^2
11	2	2	22	22	4	4	4	121
17	4	3	68	51	12	16	9	289
26	6	4	156	104	24	36	16	676
28	5	5	140	140	25	25	25	784
31	8	6	248	186	48	64	36	961
35	7	7	245	225	49	49	49	1225
41	10	9	410	369	90	100	81	1681
49	11	10	539	490	110	121	100	2401
63	13	11	819	693	143	169	121	3969
69	14	13	966	897	182	196	169	4761
370	80	70	3,613	3,197	687	780	610	16,868

Substituting the values in the normal equations :

$$10\,a_{1.23} + 80\,b_{12.3} + 70\,b_{13.2} = 370 \qquad \text{....(i)}$$

$$80\,a_{1.23} + 780\,b_{12.3} + 687\,b_{13.2} = 3{,}613 \qquad \text{....(ii)}$$

$$70\,a_{1.23} + 687\,b_{12.3} + 610\,b_{13.2} = 3{,}197 \qquad \text{....(iii)}$$

Dividing each equation by the coefficient of $b_{12.3}$, we get

$$.125\,a_{1.23} + b_{12.3} + .875\,b_{13.2} = 4.625 \qquad \text{.....(iv)}$$

$$.103\,a_{1.23} + b_{12.3} + .881\,b_{13.2} = 4.632 \qquad \text{.....(v)}$$

NOTES

$$.102\,a_{1.23} + b_{123} + .888\,b_{13.2} = 4.654 \quad \text{.....(vi)}$$

Subtracting equation, (iv) from (v) and (v) from (vi), we have

$$-.022\,a_{1.23} + .006\,b_{13.2} = .007 \quad \text{......(vii)}$$

$$-.001\,a_{1.23} + .007\,b_{13.2} = .022 \quad \text{......(viii)}$$

Multiplying (vii) and (viii) by 1,000 we get

$$-22\,a_{1.23} + 6\,b_{13.2} = 7 \quad \text{.......(ix)}$$

$$-a_{1.23} + 7\,b_{13.2} = 22 \quad \text{.....(x)}$$

Solving these equations, we get

$$a_{1.23} = 0.561;\ b_{123} = 1.735 \text{ and } b_{13.2} = 3.223$$

Hence, the required equation is

$$x_1 = 0.561 + 1.735x_2 + 3.223x_3$$

Illustration 17: The following information about a trivariate population is known :-

$\sigma_1 = 3; \sigma_2 = 4$ and $\sigma_3 = 5 \quad r_{23} = 0.40; r_{31} = 0.60$ and $r_{12} = 0.70$

Determine the regression equation of x_1 on x_2 and x_3

Solution: Regression equation of x_1 on x_2 and x_3 is

$$x_1 = a + b_{123}x_2 + b_{13.2}x_3$$

If the Variables x_1, x_2 and x_3 are measured as deviations from their respective means, 'a' will be zero. So, let us assume x_1, x_2 and x_3 represent deviations from means. Now the regression equation of x_1 on x_2 and x_3 is-

$$x_1 = b_{123}x_2 + b_{13.2}x_3$$

$$b_{123} = \frac{r_{12} - r_{13}r_{23}}{1 - r_{23}^2} \times \frac{\sigma_1}{\sigma_2}$$

$$= \frac{0.70 - 0.60 \times 0.40}{1 - (0.40)_2} \times \frac{3}{4} = 0.41$$

$$b_{13.2} = \frac{r_{13} - r_{12} \times r_{23}}{1 - r_{23}^2} \times \frac{\sigma_1}{\sigma_2}$$

$$= \frac{0.60 - (0.70 \times 0.40)}{1 - (0.40)^2} \times \frac{3}{5} = 0.228$$

∴ Required regression equation is

$$x_1 = 0.41\,x_2 + 0.228\,x_3$$

Illustration 18: Given the following data, Calculate the expected value of x_3 when $x_1 = 58$ and $x_2 = 52.5$;

$$\bar{x}_1 = 55.95; \quad \sigma_1 = 2.26; \quad r_{12} = 0.578$$

$$\bar{x}_2 = 51.48; \quad \sigma_2 = 4{,}39; \quad r_{13} = 0.581$$

$$\bar{x}_3 = 56.03; \quad \sigma_3 = 4.41; \quad r_{23} = 0.974$$

Solution :

The Multiple regression equation of x_3 on x_1 and x_2 is

$$x_{3.12} = a_{3.12} + b_{31.2}x_1 + b_{32.1}x_2 \quad \text{......(i)}$$

Now,

$$b_{31.2} = \frac{r_{13} - r_{23}\,r_{12}}{1 - r_{12}^2} \times \frac{\sigma_3}{\sigma_1} = \frac{0.581 - (0.974 \times 0.578)}{1 - (0.578)^2} \times \frac{4.41}{2.26}$$

$$= \frac{0.581 - 0.563}{1 - 0.334} \times \frac{4.41}{2.26} = \frac{0.018}{0.666} \times 1.95 = 0.053$$

$$b_{32.1} = \frac{r_{23} - r_{13}\,r_{12}}{1 - r_{12}^2} \times \frac{\sigma_3}{\sigma_2} = \frac{0.974 - (0.581 \times 0.578)}{1 - (0.578)^2} \times \frac{4.41}{4.39}$$

$$= \frac{0.974 - 0.336}{1 - 0.334} \times 1.005 = \frac{0.638}{0.666} \times 1.005 = 0.963$$

$$a_{3.12} = \bar{x}_3 - x_1\,\bar{b}_{31.2} - b_{32.1}\,\bar{x}_2$$

$$= 56.03 - (0.053 \times 55.95) - (0.963 - 51.48)$$

$$= 3.48$$

Substituting the values in the equation No. (i), we get

$$x_{3.12} = 3.48 + (0.053 \times 58) + (0.963 \times 52.5) = 57.11$$

बोध प्रश्न

1. विचरण का अनुपात ज्ञात करने की विधि को संक्षिप्त में समझाइए।

....

....

....

2. बहुगुणी प्रतीपगमन पर संक्षिप्त टिप्पणी कीजिए।

....

....

....

3. प्रतीपगमन विचार की व्याख्या कीजिए?

....

....

....

8.10 सारांश

आर्थिक एवं व्यापारिक जगत में प्रतीपगमन की अत्यन्त व्यावहारिक उपयोगिता है। इस तकनीक के प्रयोग से अनेक प्रकार के तथ्यों के पूर्वानुमान किए जा सकते हैं। प्रतीपगमन विश्लेषण प्रकृति तथा सीमा की माप करके हमें भावी अनुमान लगाने की क्षमता प्रदान करता है 1. रेखीय एवं वक्रीय प्रतीपगमन, 2. सरल एवं बहुगुणी प्रतीपगमन। प्रतीपगमन रेखाएँ दो होती हैं एवं प्रतीपगमन रेखाएँ दो प्रकार से खींची जा सकती है– 1. मुक्त हस्त विधि तथा प्रतीपगमन समीकरण के द्वारा। प्रथम विधि उपयुक्त नहीं है क्योंकि तथ्यों की रचना भिन्न हो सकती है अतः द्वितीय विधि प्रतीपगमन समीकरणों के आधार पर ही इन रेखाओं को खींचा जा सकता है।

8.11 शब्दावली या शब्दकुंजी

NOTES

8.12 अभ्यास प्रश्न

दीर्घ उत्तरीय प्रश्न (Long Answer Type Questions)

1. 'प्रतीपगमन' से क्या आशय है ? प्रत्येक द्विचर वितरण के लिए सामान्यतः दो प्रतीपगमन रेखाओं का होना क्यों आवश्यक है ?

   What is meant by 'Regression'. Why should there be in general, two lines of regression for each bivariate-distribution.

2. प्रतीपगमन-विचार की व्याख्या कीजिए। यह सह सम्बन्ध से किस प्रकार भिन्न हैं ?

   Explain the concept of Regression. How does it differ from Correlation?

3. 'प्रतीपगमन' की परिभाषा दीजिए। प्रतीपगमन रेखाएँ दो क्यों होती हैं ? किन परिस्थितियों में केवल एक ही प्रतीपगमन रेखा हो सकती है ?

   Define 'Regression'. Why are there two regression lines? Under what conditions can there be only one regression line?

4. सह सम्बन्ध, प्रतीपगमन तथा विचरण अनुपात की धारणाओं की व्याख्या कीजिए और आर्थिक अनुसंधान के क्षेत्र में उनकी उपयोगिता बतलाइए।

   Explain the concepts of Correlation, regression and ratio of variation and state their utility in the field of economic enquiries.

5. सह सम्बन्ध गुणांक और प्रतीपगमन गुणांक का अर्थ समझाइये। प्रतीपगमन समीकरण दो क्यों होना चाहिये ?

   Explain the meaning of Correlation Coefficient and regression coefficient. Why should there be two equations?

लघु उत्तरीय प्रश्न (Short Answer Type Questions)

1. प्रतीपगमन से क्या आशय है ?

   What is meant by Regression?

2. प्रतीपगमन रेखाएँ दो क्यों होती हैं ?

   Why are there two regression lines?

3. प्रतीपगमन समीकरण दो क्यों होना चाहिये।

   Why should there be two equations?

4. प्रतीपगमन गुणांक का अर्थ समझाइये।

   Explain the meaning of regression coefficient.

5. प्रतीपगमन विचार की व्याख्या कीजिये।

   Explain the concept of Regression.

वस्तुनिष्ठ प्रश्न (Objective Type Questions)

1. प्रतीपगमन का प्रयोग सर्वप्रथम किसने किया-

   (अ) स्टोव, (ब) जेवन्स, (स) फिशर, (द) गाल्टन।

2. प्रतीपगमन गुणांक प्रतीपगमन रेखा का कैसा माप है-

   (अ) गणितीय, (ब) रेखा गणितीय, (स) बीज गणितीय, (द) सामान्य।

3. दोनों प्रतीपगमन गुणांकों का मूल्य धनात्मक होने पर सहसंबंध गुणांक होगा-

   (अ) ऋणात्मक, (ब) धनात्मक, (स) शून्य, (द) अनंत।

4. दो चरों की दशा में केवल एक प्रतीपगमन रेखा होगी जबकि-

   (अ) $r = 0$, (ब) $r = +1$, (स) $r = -1$, (द) $r = +1$ या -1

[उत्तर- 1. (द), 2. (स), 3. (ब), 4. (द)]

8.13 व्यावहारिक प्रश्न

1. You are given the following results for the heights (x) and weights (y) of 1000 workers of a factory.

   Mean Height (ax) = 68.0 inches; and = 2.5 inches

   Mean Weight x(y) = 150.0 lbs; and = 20 lbs.

   Coefficient of Correlation = +0.60

   Estimate from the above data (i) the height of a particular factory worker whose weight is 200 lbs. (ii) The weight of a particular factory worker who is 5 feet tall.

   [71.75 inches and 111.6 lbs]

2. The following data are given for marks in subjects A and B in a certain examination–

   Mean Marks in A = 39.5

   Mean Marks in B = 47.5

   S.D. in A = 10.8

   S.D. in B = 16.3

   Coefficient of Correlation between A and B = +0.42

   Draw the two lines of regression and explain why there are two equations of regression. Also give the expectations of Marks in subject B for candidate who secured 50 marks in subject A.

   [Ans. 54.325, Equation are (i) $y - 47.5 = .63\,(x - 39.5)$ (ii) $x - 39.5 = .27\,(y - 47.5)$]

3. The following table gives the frequency according to age groups of marks obtained by 5 students in an intelligence test–

Test Marks	Age in Years					Total
	19	20	21	22	23	
10-20	4	4	2	1	1	12
20-40	3	5	4	2	2	16
40-60	3	6	8	5	3	25
60-80	-	4	6	8	4	22
Total	10	19	20	16	10	75

   Calculate the Coefficient of Correlation between age and intellegence.

   [r = .35]

4. Calculate Regression equation from the following data–

   Mean height = 50.07 inches

   Mean age = 9.98 yrs.

   S.D. of height = 5.26 inches

   S.D. of age = 2.59 yrs.

   Coefficient of Correlation = +0.898

   [$x = 1.797y + 32.14$; $y = .442x - 12.21$]

NOTES

5. Given the following data. calculate the expected value of y and x is 12.

	X	Y
Average	7.6	14.8
Standard Deviation	3.6	2.5

[r = +0.99]
[Ans. 17.825]

6. The ages of husband and wife in a community were found to have a Correlation Coefficient equal to +0.8; the average ages of husband and wife were 25 and 22 years and their standard deviation 4 and 5 years respectively. Draw the two line of regression and estimate the expected age of husband when the wife's age is 12 years and the expected age of wife when husband's age is 33 years.
[Ans. Husband's age 18.60, wife's age 30 years]

7. Determine the Equation of the straight line which best fits the following data—

X	:	10	12	13	16	17	20	25
Y	:	19	22	24	27	29	33	37

[Ans. x on y = x = .8124 - 6.013; y on x = y = 1.213 x + 7.705]

8. From the following data, find the probable yield when rainfall is 29 inches - Mean rainfall -25 inches, Mean Production -40 units per acre. Coefficient of Correlation between rainfall and production -8; S.D. of rainfall -3 inches; S.D. of production -6 units.
[46.4 units]

9. Given the following data for two tests—

History (x)	English (y)
$\bar{x} = 75$	$\bar{y} = 70$
$\sigma_x = 6$	$\sigma_y = 8$

[r = +0.72]
(a) Work out the two regression equations.
(b) Predict the probably grade in English of a student whose histroy marks are 65.
[y = .96x - 2, x = .54y + 37.2; 60.4]

10. Find out Coefficient of Correlation and regression equation—

x	:	1	2	3	4	5	6	7
y	:	9	8	10	12	11	13	14

Also find estimated value of y when x = 6.2
[r = .9, y =.929x + 7.248; x = -929y - 6.219; y = 13.04]

11. Calculate Coefficient of Correlation and obtain the lines of regression for the following data—

x	:	1	2	3	4	5	6	7	8	9
y	:	9	8	10	12	11	13	14	16	15

Obtain an estimate of y which should correspond on the average to x = 6.2.
[Ans. r = +.95, Regression equation y on x = .95x + 7.25]
[Regression equation x on y = .95y – 6.4; y = 13.14]

12. The following marks have been obtained by a class of students in Statistics (Out of 100).

Paper I	80	45	55	56	58	60	65	68	70	75	85
Paper II	82	56	50	48	60	62	64	65	70	74	90

Compute the Coefficient of Correlation for the above data. Find the lines of regression and examine the relationship.

[r = +.91, y = .99x + .9, x = .85y + 9.55; P.E. = 0.45]

13. Recorded data showing the test scores made by salesman on an intelligence test and their weekly sales is given in the following.

Salesman	A	B	C	D	E	F	G	H	I	J
Test Scores	40	70	50	60	80	50	90	90	60	60
Sales in '000 units	2.5	6.0	4.5	5.0	4.5	2.0	5.5	3.0	4.5	3.0

Calculate the regression line of sales on test score and estimate the most probable weekly sales volume if a salesman makes a score of 70.

What will be the sampling error of your estimate?

[Ans. bxy = .66, y = 5.65]

$$\text{Sampling Error} = \sqrt{\frac{(\Sigma dy^2 - r\frac{\sigma_y}{\sigma_x}\Sigma xy)}{n}} = \pm .88$$

14. Heights of fathers and sons are given in inches—

Height of father	65	66	67	67	68	69	71	73
Height of son	67	68	64	68	72	70	69	70

From the two lines of Regression, calculate the expected average height of the son when the height of the father is 67.5 inches.

[Ans. = 68.18 inches]

15. If the mean and S.D. of three variables are as follows :

	X_1	X_2	X_3
Mean	28.02	4.91	594
S.D.	4.42	1.10	85

and of $r_{12} = .80$, $r_{23} = -\cdot 56$ and $r_{31} = -40$. Find the plane of regression of X 1 on X 2 and X 3 .

(b) Calculate R21.23 in the above example deriving the formula you use. Interpret the result.

16. (a) Distinguish between " Multiple" and "Partial" correlations, in the case of three variables. Write down the formula for them.

(b) X_1, X_2, X_3 are three variables measured from their means with $N = 10$; $\Sigma X_1^2 = 90$; $\Sigma X_2^2 = 160$; $\Sigma X_3^2 = 40$;

$$\Sigma X_1 X_2 = 60; \Sigma X_2 X_3 = 60; \Sigma X_3 X_1 = 40$$

Calculate $r_{12.3}$ and $R_{2.31}$.

NOTES

17. Calculate coefficient of linear multiple correlation of X_3 on X_1 and X_2 from the following data. Also find $r_{12.3}$, $r_{13.2}$ and $r_{23.1}$

X_1	3	5	6	8	12	14
X_2	16	10	7	4	3	2
X_3	90	72	54	42	30	12

[Ans. $r_{12.3}$ = 0.5950, $r_{13.2}$ = − 0.8995]

18. Obtain the best rule for predicting X_1 in terms of X_2 and X_3, given the following Data:

Observation No.	:	1	2	3	4	5	6	7
Variable X_1	:	5	3	2	4	3	1	8
Variable X_2	:	2	4	2	2	3	2	4
Variable X_3	:	21	21	15	17	20	13	32

19. Calculate (a) R3.12 (b) R1.23 (c) R2.13 from the following data:

X_1 = 6.8	X_2 = 7.0	X_3 = 74
s_1 = 1.0	s_2 = .80	s_3 = .90
r_{12} = .60	r_{13} = .70	r_{23} = .65

[Ans. (a) .7567, (b) .7255, (c) . 6810]

अपनी प्रगति की जाँच करें
Test your Progress

अध्याय–9 सांख्यिकीय निर्णयन सिद्धान्त
(STATISTICAL DECISION THEORY)

NOTES

इकाई की रूपरेखा

9.0 उद्देश्य

इस इकाई का अध्ययन करने के बाद आप इस योग्य हो सकेंगे कि–

1. निर्णयन का अर्थ व इसकी अवधारणा का ज्ञान होगा।
2. निर्णयन सिद्धान्त में मूल्यों को सारणीकरण कर सकेंगे।
3. निर्णयन समस्या के तत्वों का ज्ञान होगा।

9.1 प्रस्तावना

निर्णयन प्रबन्ध का हृदय होता है, एक प्रबन्धक अपनी समस्त गतिविधियाँ एवं कार्य निर्णयन के द्वारा ही करता है। वह सतत् निर्णयन प्रक्रिया के साथ संलग्न रहता है। चाहे वह लक्ष्यों का निर्धारण करता है, योजना तैयार करता है, व्यूह रचना जमाता है, नीति निर्धारण एवं प्रक्रिया का निर्धारण करता है अथवा कर्मचारियों की भर्ती, चयन एवं नियुक्ति करता है वह केवल निर्णयन द्वारा ही ये सभी कार्य का पाता है। इस प्रकार निर्णयन समस्त प्रबन्धकीय क्रियाओं का आधार है। प्रसिद्ध प्रबन्ध शास्त्री **पीटर एफ.ड्रकर** ने कहा है कि **"एक प्रबन्धक जो कुछ भी करता है वह निर्णयन के द्वारा ही करता है।"**

निर्णयन का आशय (Meaning of Decision Making)– निर्णयन के विभिन्न पक्षों की व्याख्या करने के पूर्व निर्णय शब्द का अर्थ जानना आवश्यक है। निर्णय (Decision) शब्द की उत्पत्ति लेटिन शब्द 'decides' से हुई है, जिसका शाब्दिक अर्थ काटना होता है। इस प्रकार निर्णय उन विकल्पों जो वांछित होते हैं एवं उन विकल्पों जो अवांछित होते हैं, के मध्य काटने का काम करता है। व्यावहारिक रूप में निर्णय विभिन्न विकल्पों में से सर्वाधिक उपयुक्त विकल्प का चयन करना है। किसी भी व्यवसाय अथवा उद्योग में आरम्भ से लेकर अन्त तक अनेक निर्णय लेने पड़ते हैं, यही कारण है कि कुछ विद्वानों ने प्रबन्ध को निर्णय लेने की प्रक्रिया के रूप में परिभाषित किया है। **ड्रकर** ने कहा है कि **"निर्णय विभिन्न विकल्पों के मध्य चयन करता है।"**

NOTES

9.2 निर्णयन की अवधारणाएँ (Concept of Decision Making)

निर्णयन कोई तथ्य या विषयवस्तु न होकर एक प्रक्रिया के रूप में जाना एवं माना जाता है। अतः इसे प्रक्रिया के रूप में ही सहज मान्यता मिली है, लेकिन प्रबन्ध शास्त्र के नियंताओं ने अपनी पृथक-पृथक राय इसके बारे में कायम की थी जो विभिन्न अवधारणाओं के रूप में परिलक्षित हुई। इनमें से कुछ प्रमुख अवधारणाएँ निम्नलिखित हैं :

1. **मानसिक प्रक्रिया अवधारणा** (Mental Process Concept)– इस विचार के जनक प्रसिद्ध प्रबन्धशास्त्री **प्रो. जी. एल. एस. शेकल** थे। उनके अनुसार 'निर्णयन' एक मानसिक प्रक्रिया है, जिसमें निर्णय लेने वाला अपने मस्तिष्क का उपयोग करके विवेकपूर्ण एवं तर्कसंगत निर्णय लेता है। यह सर्वथा उचित एवं प्रासंगिक विचार है क्योंकि विकल्पों की खोज, चयन एवं त्वरित निष्कर्ष निकालना जैसी समस्त क्रियाएँ मस्तिष्क से ही संचालित की जाती हैं।

2. **चयन अवधारणा** (Selection Concept)– आरम्भिक तौर पर **कूण्ट्ज ओ'डोनेल, मैकफारलैण्ड** आदि विद्वानों ने निर्णयन को केवल चयन प्रक्रिया तक ही सीमित किया था। उनके अनुसार जब निर्णयन प्रक्रिया प्रारम्भ होती है तो उसकी प्रथम एवं मुख्य कड़ी विभिन्न विकल्पों के मध्य तुलना करके श्रेष्ठ विकल्प का चयन करना होती है। अतः समस्याओं के समाधान हेतु प्रबन्धक वर्ग इस चयन के द्वारा ही अपने लक्ष्यों को पूर्ण करते हैं।

3. **समस्या निवारण अवधारणा** (Problem Solving Concept)– यह एक स्वाभाविक सत्य है कि कोई भी निर्णय किया क्यों जाता है, क्योंकि कोई समस्या होती है एवं इसका समाधान करने के लिए निर्णयन प्रक्रिया का उपयोग ही श्रेष्ठ हल हो सकता है। अतः समस्या निवारण अवधारणा भी मुख्य रूप से निर्णयन के लक्ष्यों को उभारती है। यदि कोई समस्या नहीं है तो निर्णयन का प्रश्न ही नहीं उठता है।

4. **सतत् प्रक्रिया अवधारणा** (Continuity Process Concept)– निर्णयन कोई इस प्रकार का कार्य नहीं होता कि काम पूरा हुआ एवं निर्णयन पूरा हुआ, यह तो एक सतत् चलने वाले प्रक्रिया होती है। बहुधा यह समझ लिया जाता है कि श्रेष्ठ विकल्प का चयन करके निर्णयन समाप्त हो जाता है। यह समझना भ्रांति है क्योंकि श्रेष्ठ का चयन करके अपने संगठन को कार्य पर लगा देता है, पूँजी की व्यवस्था करता है, श्रम को अभिप्रेरित करता है, लाभार्जन यद्यपि मुख्य लक्ष्य है फिर भी सामाजिक उत्तरदायित्वों का निर्वहन करता है, मानव संसाधनों का विकास एवं पर्यावरण संरक्षण जैसे मुद्दों पर भी चिन्तन-मनन करना प्रबन्धकीय निर्णयन प्रक्रिया में समाहित किये जाते हैं। अतः निर्णयन एक निरन्तर चलने वाली प्रक्रिया होती है।

5. **उद्देश्यात्मक अवधारणा** (Objective Concept)– कोई व्यक्ति किसी भी कार्य को हमेशा किसी उद्देश्य के लिए ही करता है तो प्रबन्ध तो कोई कार्य निरुद्देश्य कर ही नहीं सकता। निर्णयन प्रक्रिया का मुख्य उद्देश्य लक्ष्य निर्धारण को ध्यान में रखकर उस तक पहुँचने के उपाय ढूँढ़ना, विकल्प प्रस्तुत करना आदि होता है। सामान्यतया निर्णयन का उद्देश्य सकारात्मक होता है, लेकिन कभी-कभी उद्देश्य नकारात्मक भी हो सकते हैं।

6. **चैतन्य मानवीय क्रिया अवधारणा** (Conscious Human Activity Concept)– नियोजन करके लक्ष्य तक पहुँचना है तो एक बाँर उस लक्ष्य प्राप्ति का नियोजन जड़ हो सकता है लेकिन निर्णयन तो व्यक्ति, उद्यमी एवं प्रबन्धक की मानसिक क्रिया से किया जाता है अतः इसमें उन्हें सदैव चैतन्य अवस्था में ही रहना होता है। यदि ये जड़ हुए तो व्यवसाय भी जड़ हो जायेगा एवं निर्णयन दिशाहीन हो जाएगा। अतः निर्णयन एक चैतन्य क्रिया होती है।

7. **वचनबद्धता की अवधारणा** (Commitment Concept)– जब भी प्रबन्ध या प्रशासन कोई निर्णय लेता है तब वह उस निर्णयानुसार कार्य सम्पन्न करने के लिए वचनबद्ध होता है अर्थात् यह वचनबद्धता उसे लक्ष्य पूर्ण होने तक निभानी होती है। निर्णयन के पश्चात् निर्णय लेने वाला वचनबद्ध हो जाता है जिसे वह पूर्ण करने या कराने हेतु बाध्य होता है। यह अवधारणा निर्णयकर्ता को उत्तरदायित्व की भावना से बद्ध करती है।

8. **सर्वव्यापकता की अवधारणा** (Universality Concept)– प्रबन्ध किसी भी स्तर का हो चाहे वह शीर्ष प्रबन्ध हो या मध्य अथवा निम्न प्रबन्ध उसे कोई न निर्णय अवश्य लेना पड़ता है अतः निर्णयन प्रबन्ध की मूल आधारशिला है। यह न केवल उद्योग, प्रबन्ध एवं प्रशासन तक बल्कि प्रत्येक व्यक्ति, समाज एवं हर एक घटक निर्णयन से प्रभावित होता है अतः निर्णयन एक सर्वव्याप्त विचार एवं अवधारणा है।

इसीलिए कई विद्वान निर्णयन एवं प्रबन्ध में कोई भेद करना नहीं चाहते क्योंकि ये दोनों समानार्थक (Synonymous) एवं पूरक माने जाते हैं।

9.3 निर्णयन की परिभाषाएँ (Definitions of Decision Making)

यह निःसन्देह सत्य है कि निर्णयन कोई तथ्य या विषयवस्तु न होकर एक मानसिक प्रक्रिया एवं सतत् गतिविधि होती है इसलिए इसे परिभाषित करना भी अत्यन्त जटिल कार्य है, लेकिन कुछ विद्वानों ने निर्णयन को अपने तरीके से परिभाषित किया है जो निम्न प्रकार से है:

1. **वेल्सटर शब्द** कोष के अनुसार "निर्णयन का अर्थ किसी एक व्यक्ति के मस्तिष्क में किसी सम्मति अथवा कार्यवाही के तरीके के निर्धारण से लगाया जाता है।"

[**टिप्पणी:** यह परिभाषा निर्णयन के केवल एक पक्ष सहमति व तरीके का उल्लेख करती है जो निर्णयन के सही अर्थ को स्पष्ट नहीं करती है।]

2. **कून्ट्ज एवं ओ'डोनेल** के अनुसार, "निर्णयन किसी क्रिया को करने के विभिन्न विकल्पों में से किसी एक का वास्तविक चयन नियोजन का अन्तर्भाग होता है।"

3. **अर्नेस्ट डेल** के अनुसार, "प्रबंधकीय निर्णय वे निर्णय होते हैं जो सदैव सही प्रबन्धकीय क्रियाओं यथा नियोजन, संगठन, कर्मचारियों की भर्ती, निर्देशन, नियन्त्रण, नव-प्रवर्तन और प्रतिनिधित्व में से किसी एक के दौरान लिये जाते हैं।"

[**टिप्पणी** : ये परिभाषाएँ यद्यपि निर्णयन को उसकी सभी विशेषताओं के साथ प्रस्तुत करता है, लेकिन इनमें निर्णयन की निरन्तरता को स्पष्ट नहीं किया गया है।]

4. **जार्ज आर. टेरी** के अनुसार, "निर्णयन किसी मापदण्ड के आधार पर दो अथवा दो से अधिक सम्भावित विकल्पों में से किसी एक का चुनाव करना होता है।"

5. **मैक-फारलैण्ड** के अनुसार, "निर्णयन एक चयन प्रक्रिया है जिसके अन्तर्गत एक अधिशाषी दी हुई परिस्थिति में, क्या किया जाना है, के सम्बन्ध में किसी निष्कर्ष पर पहुँचता है। निर्णय किसी व्यवहार का प्रतिनिधित्व करता है जिसका चयन अनेक सम्भावित विकल्पों में से किया जाता है।"

[**टिप्पणी** : **जॉज टेरी** ने तो निर्णयन हेतु केवल दो विकल्पों की बात करके इसे अत्यन्त सीमित कर दिया है, लेकिन **मैक-फारलैण्ड** ने निर्णयन के उन पहलुओं को छुआ है जो अभी तक स्पष्ट नहीं किये गये थे जैसे यह एक प्रक्रिया है, व्यवहार का प्रतिनिधित्व करने वाला व अनेक विकल्पों में से श्रेष्ठ का चयन करना आदि।]

6. **जी.एल.एस.शेकल** के अनुसार, "निर्णयन रचनात्मक मानसिक क्रिया वा वह केन्द्र बिन्दु है जहाँ कार्यपूर्ति के लिए ज्ञान, विचार, भावना तथा कल्पना का संयोग होता है।"

7. **एलन** के अनुसार, "निर्णयन वह कार्य है जिसे एक प्रबन्धक किसी निष्कर्ष तथा फैसले पर पहुँचने के पूर्व करता है।"

[**टिप्पणी** : इन दोनों परिभाषाओं में एक-एक बिन्दुओं को स्पष्ट किया है- **शेकल** ने निर्णयन को मानसिक क्रिया जिसमें ज्ञान, विचार, भावना व कल्पना को जोड़ा है वहीं **एलन** ने इसे एक ऐसा कार्य निरूपित किया है जो प्रबन्धक निष्कर्ष तक पहुँचने से पहले करता है।]

8. **फेलिक्स लॉपेज** के अनुसार, "निर्णय एक फैसले, विभिन्न विरोधाभासी आवश्यकताओं, संसाधनों एवं उद्देश्यों के पृथक्करण एवं अनिश्चितता, जटिलता अथवा यहाँ तक कि अविवेकपूर्णता की स्थिति में किसी कार्य के प्रति वचनबद्धता को इंगित करता है।"

[**टिप्पणी** : **फेलिक्स लॉपेज** ने निर्णयन को लगभग पूर्णता प्रदान करने की चेष्टा की है। लॉपेज ने उद्देश्यों के प्रथक्करण, कार्य के प्रति वचनबद्धता आदि को अपनी परिभाषा में समेटकर इसे पूर्णता प्रदान की है।]

निर्णयन के अर्थ एवं परिभाषाओं की विवेचना पश्चात् इसकी एक श्रेष्ठ एवं आदर्श परिभाषा इस प्रकार से दी जा सकती है:

9. **आदर्श परिभाषा** – "निर्णयन प्रबन्ध की एक निरन्तर चलने वाली प्रक्रिया है जिसमें किसी भी समस्या के निदान हेतु प्रबन्ध द्वारा उपलब्ध विकल्पों में से श्रेष्ठ का चयन करके अपने लक्ष्य तक पहुँचने तथा उद्यम के सभी घटकों को सक्रिय रखने व उनकी प्रत्येक समस्या के निदान के लिए छोटे-बड़े निर्णय त्वरित लेकर अपनी वचनबद्धता को स्थापित किया जाता है।"

NOTES

9.4 निर्णयन के लक्षण या विशेषताएँ (Essentials or Characteristics of Decision-Making)

विवेकपूर्ण निर्णय लेने के लिये आवश्यक होता है कि निर्णय की विशेषताओं को उनके सही अर्थों एवं सन्दर्भों में जाना जाये। एक निर्णय में मुख्य रूप से निम्न विशेषताएँ पायी जाती हैं:

(1) **निर्णयन एक अन्तिम प्रक्रिया है** (Decision is the Last Process)– निर्णय लेना प्रबन्ध का तर्कपूर्ण एवं गहन विचार-विमर्श के बाद किया जाने वाला कार्य होता है अतः इसे अन्तिम प्रक्रिया माना गया है। यह निर्णय लेने के कार्य का परिणाम होता है। यह परिणाम विभिन्न सम्भावित विकल्पों पर काफी तर्कपूर्ण विचार-विमर्श करने के पश्चात् प्राप्त होता है। इसीलिये निर्णय, जो कि एक अन्तिम प्रक्रिया होती है, वह बौद्धिक विश्लेषण, विचार-विनिमय और विकल्पों के तुलनात्मक व विश्लेषणात्मक अध्ययन का निष्कर्ष होती है।

(2) **निर्णयन सर्वोत्तम चयनित विकल्प है** (Decision is a Best Selected Alternative)– निर्णय लेते समय प्रबन्धक के सामने विभिन्न विकल्प होते हैं। उन विकल्पों का उनकी उपादेयता एवं उनसे उत्पन्न होने वाले प्रभावों के सन्दर्भ में तुलनात्मक विश्लेषण करना होता है और अधिक लाभों तथा सपरिणामों की उपलब्धि कराने वाले या कम से कम नुकसान वाले विकल्प का चयन किया जाता है। अतएव निर्णय चुना गया एक सर्वोत्तम विकल्प होता है जो या तो अधिक लाभ प्रदान करने वाला या कम से कम हानिकारक होता है।

(3) **निर्णयन सकारात्मक एवं नकारात्मक हो सकता है** (Decision might be Positive and Negative)– लिये जाने वाले निर्णयों का सदैव ही धनात्मक अर्थात् सकारात्मक होना जरूरी नहीं है। वे नकारात्मक भी हो सकते हैं। किसी कार्य को करने या योजना को लागू करने का निर्णय सकारात्मक होता है और उसी कार्य को न करने या योजना को लागू न करने का निर्णय नकारात्मक होता है। कई बार नकारात्मक निर्णय भी संस्था की भलाई के लिए लेने पड़ते हैं।

(4) **निर्णयन साधन या साध्य या दोनों से सम्बन्धित होता है** (A Decision May Relate to the End or Means or Both)– निर्णय का सम्बन्ध साधन-साध्य दोनों से हो सकता है। कुछ परिस्थितियों में निर्णय साधन का कार्य करता है और कुछ परिस्थितियों में साध्य का। जिन परिस्थितियों में उद्देश्य दिया होता है और उनकी पूर्ति के तरीकों को मालूम करना होता है तो उनमें निर्णय साधन का काम करता और कई बार निर्णयन का सम्बन्ध साधन एवं साध्य दोनों से ही हो सकता है।

(5) **निर्णयन अधिकार दर्शाता है** (Decision Making Shows Rights)– निर्णय लेना यह दर्शाता है कि निर्णयकर्ता के पास निर्णय लेने के लिए पर्याप्त अधिकार हैं। किसी संगठन के अधिकारों के भारार्पण के आधार पर यह निश्चित किया जाता है कि किस व्यक्ति द्वारा किस प्रकार का निर्णय लिया जा सकता है।

(6) **निर्णयन संकल्प का प्रतीक होता है** (Decision Making is an Indicator of a Commitment)– प्रबन्ध द्वारा किसी भी प्रकार का निर्णय लिया गया हो, वह उसके प्रति दृढ़ संकल्प का द्योतक होता है जिसके क्रियान्वयन हेतु सामूहिक या वैयक्तिक प्रयत्न करने होते हैं।

(7) **विशेषज्ञों का परामर्श** (Advice of Specialist)– यदि प्रबन्ध या प्रशासन किसी जटिल समस्या के निर्णयन में कोई असहजता या कठिनाई पाते हैं तो वे सम्बन्धित विषय के विशेषज्ञों से परामर्श भी ले सकते हैं।

बोध प्रश्न

1. निर्णयन की अवधारणा को समझाइए।

....

....

....

....

....

2. निर्णयन की परिभाषा दीजिए।

....

....

....

3. निर्णयन के लक्षण लिखिए।

....

....

....

....

....

सांख्यिकीय निर्णय विश्लेषण (Statistical Decision Analysis)

सांख्यिकीय निर्णय विश्लेषण वह है जिसके अन्तर्गत सर्वश्रेष्ठ निर्णय क्रिया निर्धारित करने के उद्देश्यों से आर्थिक परिणामों के साथ-साथ प्रायिकता मूल्यों को शामिल किया जाता है। इसे निर्णय सिद्धान्त (Decision Theory) के नाम से भी जाना जाता है।

सांख्यिकीय निर्णय विश्लेषण का आरम्भ सशर्त आर्थिक परिणामों को दर्शाने वाली सारणी को निर्णय सारणी कहते हैं। जब निर्णय क्रियाओं के अनुक्रम की आवश्यकता महसूस होती है तब इस विश्लेषण में निर्णय-वृक्ष की रचना की जाती है। निर्णय-वृक्ष में प्रत्येक विकल्प के लिए एक शाखा का उपयोग किया जाता है और उस विकल्प के चयन से सम्भावित घटनाओं को उपशाखाओं द्वारा प्रस्तुत किया जाता है। इस प्रकार, एक उपशाखा से फिर कई शाखाएँ निकल सकती हैं और रेखाचित्र देखने पर शाखायुक्त वृक्ष की तरह दिखाई देता है।

निर्णय-वृक्ष में आयत अथवा वर्ग उस बिन्दु को प्रदर्शित करता है, जहाँ निर्णय लेना है। इसी वर्ग में से निकली शाखा पर बने वृत्त से निकलने वाली शाखाएँ सम्बन्धित विकल्प के चयन किए जाने पर सम्भावित घटनाओं और परिणामों को प्रदर्शित करती हैं।

निर्णय-सारणी (Decision-Table)

किसी निर्माणी कम्पनी को निर्णय लेना है कि किसी विशिष्ट उत्पाद का निर्माण किया जाए अथवा उसे बाजार से खरीदा जाए। इस प्रकार निर्माण के सम्बन्ध में दो सम्भावित निर्णय क्रिया हैं-

1. निर्माण, 2. अनिर्माण,

हम बाजार स्थिति को माँग के सन्दर्भ में तीन सम्भावित स्तरों में व्यक्त कर सकते हैं-

1. निम्न (Low), 2. सामान्य (Moderate), 3. उच्च (High),

अब हम इनके आधार पर निर्णय-सारणी की निम्न रूपरेखा तैयार कर सकते हैं-

बाजार स्थिति-माँग के तीन सम्भावित स्तर

निर्णय क्रिया			बाजार स्थिति
	S_1 निम्न माँग	S_2 सामान्य माँग	S_3 उच्च माँग
D_1 निर्माण			
D_2 निर्माण नहीं	0	0	0

9.5 सारणी में मूल्य रखना (Keep the Values in Table)

अब हम सारणी में मूल्यों को रखें। ये मूल्य निर्णय क्रिया के सशर्त आर्थिक परिणामों (रुपयों में) के रूप में प्रस्तुत होने चाहिए। जैसे कि यदि उत्पादन का 'निर्माण नहीं' किया जाना है तो आर्थिक परिणाम शून्य (Zero) होगा और बाजार की स्थिति कुछ भी हो, इससे कोई फर्क नहीं पड़ता, क्योंकि निर्माण नहीं किया जाना है।

NOTES

इस प्रकार, जब उत्पाद बाजार में जाएगा ही नहीं तो लाभ या हानि का प्रश्न ही नहीं उठता है। इसलिए हम 'निर्णय नहीं' वाली दूसरी पंक्ति में शून्य मूल्य अंकित कर सकते हैं।

सारणी के प्रथम पंक्ति को पूरा करने के लिए हमें जोखिमपूर्ण दशाओं पर विचार करना होगा। माना कि यदि माँग निम्न है तो कुल पूँजी निवेश पर शुद्ध हानि 3,00,000 रुपये होगी और सामान्य माँग की दशा में शुद्ध लाभ 50,000 रु. और माँग की दशा में शुद्ध लाभ 2,50,000 रु. होगा। अब हम इन मूल्यों को सारणी में रख सकते हैं।

9.6 प्रायिकता का निर्धारण (Determination of Probability)

सांख्यिकीय निर्णय विश्लेषण में प्रायिकता मूल्यों और आर्थिक परिणामों, दोनों को शामिल किया जाता है। इसलिए आर्थिक परिणामों के बाद हम प्रत्येक माँग स्थिति के लिए प्रायिकता का निर्धारण करेंगे। पूर्ववर्ती अनुभवों के आधार पर माना कि विभिन्न माँग की दशाओं में प्रायिकता इस प्रकार है-

निम्न माँग की दशा में - 0.20

सामान्य माँग की दशा में - 0.50

उच्च माँग की दशा में - 0.30

अब इन आर्थिक परिणामों और प्रायिकता को निम्न सारणी में प्रदर्शित करेंगे-

आर्थिक परिणाम और प्रायिकता

निर्णय क्रिया	स्थिति S_1 निम्न माँग (P = 0.20) Rs.	स्थिति S_2 सामान्य माँग (P = 0.50) Rs.	स्थिति S_3 उच्च माँग (P = 0.30) Rs.
D_1 निर्माण	–30,00,000	50,000	2,50,000
D_2 निर्माण नहीं	0	0	0

प्रत्येक सम्भावित निर्णय क्रिया के लिए प्रत्याशित मौद्रिक मूल्य (EMV) निर्धारित करना ताकि सर्वोत्तम निर्णय क्रिया की पहचान की जा सके। इसके लिए हम अंकगणित का सहारा ले सकते हैं-

$$\text{EMV } (D_1) = 0.20\ (-3,00,000) + 0.50\ (50,000) + 0.30\ (2,50,000) = 40,000 \text{ रु.}$$

$$\text{EMV } (D_2) = 0.20\ (0) + 0.50\ (0) + 0.30\ (0) = 0 \text{ रु.}$$

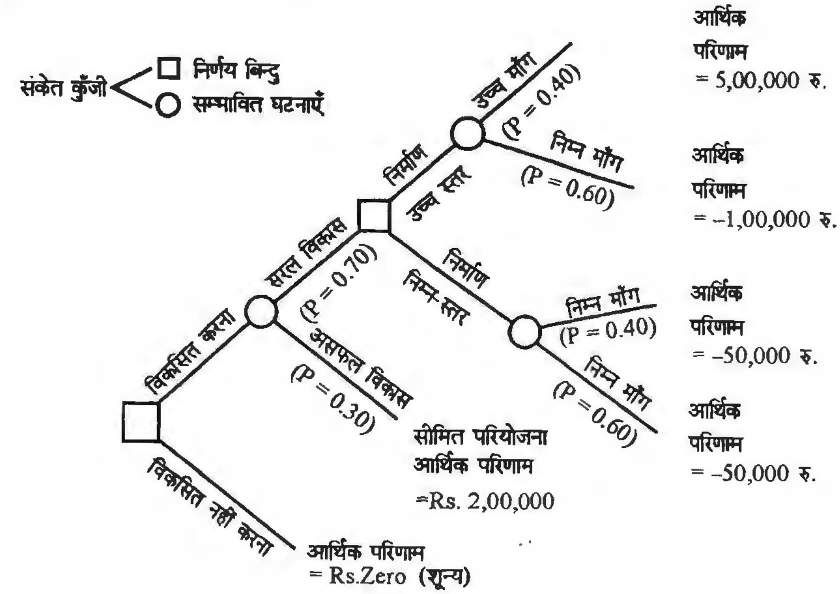
उपलब्ध निर्णय क्रियाओं के लिए प्रत्याशित मौद्रिक मूल्यों के तुलनात्मक अवलोकन से पता चलता है कि सर्वश्रेष्ठ निर्णय D_2 के बजाय D_1 में है, किन्तु यहाँ हमें यह नहीं भूलना चाहिए कि किसी विशिष्ट निर्णय स्थिति के लिए सर्वोत्तम क्रिया में अशुद्धि भी हो सकती है। जैसे कि इसी उदाहरण में यह सम्भव है कि निम्न बाजार माँग की स्थिति में प्रायिकता 0.20 सही नहीं हो। इस प्रकार, सांख्यिकीय निर्णय विश्लेषण के अन्तर्गत सर्वश्रेष्ठ निर्णय क्रिया का चयन यह आश्वासन नहीं देता कि प्रत्येक निर्णय सही ही है, फिर भी इतना अवश्य है कि यह हमें दीर्घकालीन औसत आर्थिक परिणाम की अनुकूलतम दिशा की ओर ले जाता है।

9.7 निर्णय-वृक्ष (Decision-Tree)

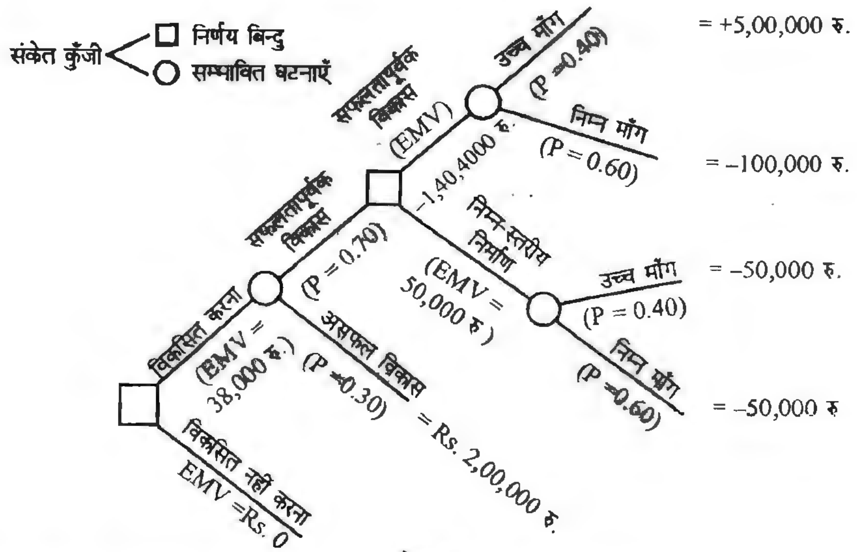
उपर्युक्त निर्णय स्थिति में यदि निर्माण करने अर्थात् पहला निर्णय लिया जाता है तो निर्णय की समस्या जटिल बन जाती है, क्योंकि प्रथम निर्णय लेने पर अब आवश्यक हो जाता है कि उसके बाद वाले सम्भावित परिणामों के आधार पर और भी अनेक निर्णय लिए जाएँ। इस तरह की समस्या को प्रस्तुत रेखाचित्र (1) द्वारा प्रदर्शित किया गया है।

अनुक्रमिक निर्णय समस्या का निर्णय वृक्ष

NOTES



रेखाचित्र 1



रेखाचित्र 2

प्रत्याशित मुद्रा मूल्यों सहित निर्णय वृक्ष

निर्णय वृक्ष उन सभी निर्णय क्रियाओं को और उन स्थितियों को दर्शाते हैं जो अनुक्रमिक रूप से घटित होती हैं। दिए गए चित्र (2) में निर्णय बिन्दुओं को वर्ग संकेत द्वारा एवं सम्भावित घटनाओं को वृत प्रतीक द्वारा दर्शाया गया है।

चित्र (2) में दो निर्णय बिन्दु हैं-

1. उत्पाद को विकसित करना है अथवा नहीं, और
2. यदि उत्पाद को सफलतापूर्वक विकसित करना है तो उसका निर्माण स्तर कैसा होगा?

निर्णय वृक्ष की उपयोगिता यह है कि यह निर्णयकर्ता को प्राथमिक निर्णय के साथ-साथ सम्भावित परिणामों की दिशा में विचार करने की आवश्यकता प्रतिपादित करता है। रेखाचित्र में प्रदर्शित सम्भावित परिणामों की दृष्टि से हम यह कह सकते हैं कि उत्पाद को विकसित करने का निर्णय तभी सही होगा जब 1. उत्पाद को विकसित किया जाए, और 2. इस उत्पाद के लिए अनुक्रमिक माँग उच्च हो।

प्रत्येक निर्णय बिन्दु पर सर्वश्रेष्ठ निर्णय क्रिया सारणी विश्लेषण के समान ही है। रेखाचित्र में प्रदर्शित निर्णय-वृक्ष में सभी निर्णय क्रियाओं के लिए प्रत्याशित मौद्रिक मूल्यों को भी दर्शाया गया है। इसकी गणना पहले की तरह ही की

NOTES

गई है। यहाँ उल्लेखनीय है कि उत्पाद को विकसित किए जाने सम्बन्धी EMV की गणना और उच्च-स्तरीय निर्माण के लिए EMV की गणना इस प्रकार की गई है-

EMV = 0.70 (1,40,000) + 0.30 (-2,00,000) = 38,000 रु.

EMV = 0.40 (5,00,000) + 0.60 (-1,00,000) = 1,40,000 रु.

1. उच्च-स्तर पर निर्माण, जहाँ प्रत्याशित मौद्रिक मूल्य 1,40,000 रु. है, और
2. निम्न-स्तरीय निर्माण, जहाँ प्रत्याशित मौद्रिक मूल्य -50,000 रु. है।

प्रस्तुत दशा में उत्पाद के सफलतापूर्वक विकास के साथ उच्च-स्तरीय निर्माण कार्य सर्वश्रेष्ठ निर्णय क्रिया है। जैसा कि रेखाचित्र में प्रदर्शित है कि विकसित नहीं करने और निम्न-स्तरीय निर्माण को वरीयता नहीं देने के कारण उसके लिए दोहरे धारी (H) का प्रयोग किया गया है।

सारांश में हम कह सकते हैं कि प्रथम निर्णय बिन्दु पर उत्पाद को विकसित करने के चयन में 38,000 रु. प्रत्याशित मौद्रिक मूल्य और विकसित नहीं करने का चयन करने पर शून्य मौद्रिक मूल्य उपलब्ध है। अतएव, सर्वोत्तम निर्णय क्रिया उत्पाद को विकसित करना है। इसके बाद, यदि उत्पाद को असफलतापूर्ण विकसित किया जाए तो 2,00,000 रु. की हानि होगी। इसलिए सर्वश्रेष्ठ निर्णय यह है कि उत्पाद का सफलतापूर्वक विकास किया जाए और उच्च-स्तरीय निर्माण कार्य किया जाए।

9.8 निर्णयन समस्या के तत्व (Ingredients of Decision Problem)

निर्णयन समस्या के तत्व निम्नलिखित हैं-

(1) **उद्देश्य का निर्धारण** (Determination of Objectives)- निर्णय लेने हेतु उद्देश्यों का पूर्व निर्धारण आवश्यक है। निर्णायक के सभी कदम उद्देश्यों पर ही आधारित होते हैं। निर्णय लेना, निर्णय को स्थगित करना, कोई निर्णय न लेना या यथास्थिति कायम रखना भी निर्णय प्रक्रिया के ही अंग हैं, लेकिन यह सभी पूर्व-निर्धारित उद्देश्यों पर निर्भर करते हैं।

(2) **विकल्पों का ज्ञान** (Identification of Alternatives)- निर्णय लेने के लिए समस्या से सम्बन्धित विभिन्न विकल्पों की खोज तथा उन्हें सूचीबद्ध करना आवश्यक है। निर्णय लेने की स्थिति उत्पन्न ही तब होती है, जबकि समस्या से सम्बन्धित एक से अधिक विकल्प उत्पन्न हों। यदि कोई समस्या इस प्रकार की है कि उसके हल के लिए कोई विकल्प उपलब्ध ही नहीं है तो निर्णय नहीं लिया जा सकता है।

(3) **घटनाओं का पूर्वानुमान** (Estimation of Events)- भावी घटनाओं का निर्णयन पर सीधा प्रभाव पड़ता है, क्योंकि भावी घटनायें विभिन्न बाहरी शक्तियों (आर्थिक, सामाजिक, मौसमी, राजनैतिक तथा तकनीकी) के संयोग के कारण प्रकट होती हैं। अतः निर्णायक के लिए यह आवश्यक है कि उसे विभिन्न विकल्पों तथा विकल्पों के संयोगों से सम्बन्धित विभिन्न भावी घटनाओं की जानकारी हो। यदि निर्णायक विभिन्न घटनाओं का पूर्वानुमान नहीं कर सकता है तो उसे व्यवसाय में सफलता की प्राप्ति संदिग्ध रहती है।

(4) **विकल्पों की उपादेयता की माप** (Evaluation of Alternatives)- सामान्य आर्थिक परिस्थितियों में उपादेयता की माप के लिए दो प्रकार की इकाइयों का प्रयोग किया जाता है-

(a) मुद्रा मूल्य (Monetary Value),

(b) उपयोगिता मूल्य (Utility Value)।

निर्णयन के लिए विकल्पों की उपादेयता की माप किया जाना आवश्यक है।

(5) **निर्णय वातावरण** (Decision Environment)- निर्णयकर्ता को निर्णय सम्बन्धी उपलब्ध सूचनाओं के आधार पर निर्णय परिस्थितियों को तीन भागों में बाँटा जा सकता है-

(i) निश्चितता की स्थिति में निर्णय (Decision making under certainty)

(ii) अनिश्चितता की स्थिति में निर्णय (Decision making under uncertainity)

(iii) जोखिम की स्थिति में निर्णय (Decision making under risk)

इसमें प्रयोग होने वाले शब्दों का अर्थ इस प्रकार है-

(1) **कृत्य या विकल्प** (Acts or Alternatives)- निर्णयकर्ता को उपलब्ध विभिन्न विकल्प को ही कृत्य कहते हैं। इन्हीं विकल्पों में से निर्णयकर्ता को अनुकूलतम विकल्प का चुनाव करना होता है। एक निर्णय सम्बन्धी विभिन्न विकल्पों के समुच्चय को निम्न प्रकार से व्यक्त किया जा सकता है-

$$a = (a_1, a_2, a_3, \ldots.. a_n)$$

या

$$A = (A_1, A_2, A_3, \ldots.. A_n)$$

(2) **घटनायें** (Events or States of Nature)- निर्णयकर्ता द्वारा किसी कृत्य के लिए किये गए निर्णय का परिणाम कई तत्वों पर आश्रित होता है जो निर्णयकर्ता की पहुँच से परे होता है। अर्थात् निर्णयकर्ता का इन पर कोई प्रभावी नियंत्रण नहीं होता है। इनको ही घटनाएँ कहते हैं। विभिन्न घटनाओं के समुच्चय को निम्न प्रकार से व्यक्त किया जा सकता है-

$$E = [E_1, E_2, E_3, \ldots.. E_m]$$

या

$$S = (S_1, S_2, S_3, \ldots\ldots.. S_m)$$

(3) **प्राप्तियाँ** (Pay-Off)- विकल्पों तथा घटनाओं के संयोग पर प्राप्त मौद्रिक मान "Pay-Off" कहलाता है।

निर्णायक की प्राप्तियाँ एक याच्छिक चर है जिसका मान निर्णायक द्वारा चुने गये विकल्प एवं भावी घटनाओं के संयोग पर निर्भर करता है। विभिन्न विकल्पों तथा घटनाओं सम्बन्धी प्राप्तियों को सारणी रूप में निम्न प्रकार प्रस्तुत किया जा सकता है।

Events \ Acts	a_1	a_2	a_3	.....	a_n
E_1	O_{11}	O_{12}	O_{13}	.....	O_{1n}
E_2	O_{21}	O_{22}	O_{23}	....	O_{2n}
E_3	O_{31}	O_{32}	O_{33}	.....	O_{3n}
:	:	:	:	......	:
:	:	:	:	......	:
:	:	:	:	....	:
E_m	O_{m1}	O_{m2}	O_{m3}	.....	O_{mn}

(4) **विकल्प त्याग** (Opportunity Loss or Regret)- किसी घटना से सम्वन्धित विभिन्न विकल्पों में से निर्णयकर्त्ता को सर्वोत्तम विकल्प चुनना होता है। लाभ का वह भाग जो निर्णयकर्त्ता को दी हुई परिस्थिति में सर्वोत्तम विकल्प न चुनने के कारण छोड़ना पड़ता है, विकल्प त्याग संबंधी हानि कहलाती है। विकल्प त्याग की लागत की गणना करने के लिए प्रत्येक घटना की सर्वाधिक प्राप्ति से उस घटना से सम्बन्धित विकल्पों की प्राप्तियों को घटा दिया जाता है।

निश्चितता की स्थिति में निर्णय (Decision Making Under Certainty)- निश्चितता की स्थिति में निर्णयकर्त्ता को निर्णय सम्बन्धी सभी सूचनाएँ उपलब्ध होती हैं तथा वह भविष्य की घटनाओं के वारे में भी पूर्णता से पूर्वानुमान कर सकता है। निर्णयकर्त्ता यह मानकर चलता है कि उसके निर्णय के लिए एक ही घटना सम्बन्धित है। निश्चितता की स्थिति में निर्णय लेने की प्रमुख तकनीकें निम्न हैं-

1. आदा-प्रदा विश्लेषण (Input-output Analysis)
2. लागत-लाभ विश्लेषण (Cost Profit Analysis)
3. सामग्री प्रबन्ध तकनीक (Inventory Control Method)
4. रेखीय कार्यक्रम (Linear Programming)

अनिश्चितता की स्थिति में निर्णय (Decision Under Uncertainty)- निर्णयकर्त्ता के पास निर्णय परिस्थिति के बारे में कोई सूचना उपलब्ध नहीं है। इस स्थिति में वह घटनाओं को प्रायिकता भी प्रदान नहीं कर सकता है। भावी घटनाओं का कोई पूर्व इतिहास भी नहीं होता है तथा निर्णयकर्त्ता के पास प्रत्याशित लाभ प्राप्त करने का कोई साधन नहीं होता। निर्णयकर्त्ता इस स्थिति में निम्न निर्णय कसौटियों में किसी भी एक को अपनाता है-

(1) **अधिकतम से अधिकतम कसौटी** (Maximax Criterion)- आशावादिता पर आधारित इस कसौटी के आधार पर निर्णय करने वाला निर्णायक प्रत्येक विकल्प के लिए अधिकतम प्राप्ति का निर्णय करता है या फिर उस विकल्प का चुनाव करता है जिसकी तुलनात्मक दृष्टि से सर्वाधिक प्राप्ति हो। यह कसौटी उस समय उपयुक्त हो सकती है, जबकि निर्णयकर्त्ता विकल्प के चुनाव के परिणामस्वरूप होने वाली हानि को सहन करने की स्थिति में हो।

(2) **न्यूनतम से अधिकतम कसौटी** (Maxi-Min Criterion)- यह कसौटी निराशावादी निर्णयकर्त्ता है जो निकृष्टतम में से श्रेष्ठ चुनने का आकांक्षी है। इस कसौटी के अन्तर्गत निर्णयकर्त्ता प्रत्येक विकल्प से न्यूनतम प्राप्ति का चुनाव करता है तथा इन न्यूनतम प्राप्तियों से उस विकल्प का चुनाव करता है, जिसमें सर्वाधिक प्राप्ति हो।

(3) **अधिकतम से न्यूनतम (विकल्प त्याग लागत कसौटी)** (Mini-Max Regret Criterion)- इस कसौटी के अन्तर्गत विभिन्न विकल्पों सम्बन्धी विकल्प त्याग की लागतों में से निर्णायक सर्वाधिक विकल्प त्याग की लागत का चुनाव करता है तथा उस विकल्प का चुनाव करता है जहाँ विकल्प त्याग लागत न्यूनतम हो।

(4) **आशावादी गुणांक कसौटी** (Coefficient of Optimism Criterion-Hurwitz Criterion)- यह सिद्धान्त आशावादी व निराशावादी सिद्धान्त का समन्वय है और आशावादी गुणांक पर आधारित है। इस सिद्धान्त के प्रतिपादक **हरविज** (Hurwitz) हैं। **इसके** गुणांक का मूल्य 0 तथा 1 के मध्य होता है। पूर्णतः आशावादी निर्णायक के लिए **हरविज गुणांक** (α) का मान एक (1) तथा पूर्णतः निराशावादी निर्णायक के लिए हरविज गुणांक (α) का मान शून्य (0) होता है। सामान्य निर्णायक इन दोनों मानों के मध्य अपना आशावादी गुणांक निर्धारित करता है। गुणांक निर्धारण के बाद निर्णायक उस विकल्प का चुनाव करेगा, जिसके लिए हरविज मान का मूल्य सर्वाधिक हो। हरविज मान का मूल्य (Value Based on Hurwitz Criterion) निम्न प्रकार निकाला जा सकता है।

$$H(a_1) = \text{विकल्प } a_1 \text{ से अधिकतम प्राप्ति} \times (\alpha) + \text{विकल्प } a_1 \text{ से न्यूनतम प्राप्ति} \times (1-\alpha)$$

or

$$(\alpha)\ (\text{Maximum}) + (1 - \alpha)\ (\text{Minimum})$$

(5) **बर्नोली की विधि** (Jacob Bernoulli Method)- इस रीति का प्रयोग उस समय किया जाता है, जबकि यह मानने के लिए कोई कारण नहीं हो कि किसी भी घटना के घटने की प्रत्याशा दूसरों से अधिक हो। इस प्रकार प्रत्येक घटना के घटने के लिए समान प्रायिकता निर्धारित की जाती है। उदाहरण के लिए-

$$\frac{1}{n}[O_{i1} + O_{i2} + \ldots\ O_{in}]$$

जोखिम की स्थिति में निर्णय (Decision Making Under Risk)- अनेक परिस्थितियों में निर्णायक को निर्णय परिस्थितियों के बारे में पूर्ण जानकारी तो नहीं होती है परन्तु वह पूर्व अनुभव के आधार पर भावी घटनाओं को प्रायिकता प्रदान कर सकने में सम्भव होता है। उदाहरणस्वरूप एक खुदरा व्यापारी अगले दिन की बिक्री के लिए आदेश प्रेषित करना चाहता है। वस्तु की प्रकृति शीघ्र नश्वर है। अतः यदि शाम तक नहीं बिक पाती है तो वह नष्ट हो जायेगी या फिर कम मूल्य पर बेचनी पड़ती है। व्यापारी अपने पूर्व अनुभव के आधार पर विभिन्न माँगों को सापेक्ष बारम्बारता (Relative Frequencies) प्रदान कर सकता है। इस प्रकार की समस्याओं का हल जोखिम की स्थिति में निर्णयन की विभिन्न विधियों का प्रयोग कर किया जा सकता है।

अनिश्चय की स्थिति में प्रायिकता की सहायता से निर्णय लेना

(Decision Making Under Uncertainty with the use of Probabilities)

(i) **प्रत्याशित मौद्रिक मान कसौटी** (Expected Monetary Value Criterion or EMV)- निर्णायक को भावी घटनाओं से सम्बन्धित प्रायिकताओं का ज्ञान होने पर प्रत्याशित मौद्रिक मान कसौटी (EMV) के आधार पर निर्णय किया जा सकता है। इस विधि के अन्तर्गत निर्णायक अपने पूर्व अनुभव या परीक्षण द्वारा विभिन्न घटनाओं को

प्रायिकता प्रदान करता है तथा फिर विभिन्न विकल्पों सम्बन्धी प्रत्याशित मौद्रिक मान का परिकलन निम्न सूत्र की सहायता से करता है-

$$EMV = \sum_{j=1}^{n} p_i O_{ij} = P_1 O_{i1} + P_2 O_{i2} + \ldots + P_n O_{in}$$

प्रत्याशित मौद्रिक मान कसौटी ('बे' का सिद्धान्त) के आधार पर निर्णायक उस विकल्प का चुनाव करता है जिसका प्रत्याशित मौद्रिक मान (EMV) सर्वाधिक हो।

(ii) **प्रत्याशित विकल्प त्याग लागत कसौटी** (Expected Opportunity Loss Criterion EOL)- प्रत्याशित विकल्प त्याग लागत कसौटी (EOL) प्रत्याशित मौद्रिक मान कसौटी का वैकल्पिक रूप है। प्रत्येक कृत्य के लिए प्रत्याशित विकल्प त्याग हानि उसी प्रकार निकाली जाती है जिस प्रकार प्रत्याशित मौद्रिक मान ज्ञात किये जाते हैं। सूत्रानुसार-

$$EOL = \sum_{j=1}^{n} p_i R_{ij} = P_1 R_{i1} + P_2 R_{i2} + \ldots + P_n R_{in}$$

प्रत्याशित विकल्प त्याग लागत कसौटी के आधार पर निर्णायक को उस विकल्प का चुनाव करना चाहिए, जिसकी प्रत्याशित विकल्प त्याग हानि न्यूनतम हो।

प्रत्याशित विकल्प त्याग हानि को कहते हैं जो कि अधिकतम लाभ में से वास्तविक लाभ घटाने के बाद प्राप्त होता है।

पूर्ण सूचना का प्रत्याशित मूल्य (Expected Value of Perfect Information EVPI)- पूर्ण सूचना का प्रत्याशित मूल्य (EVPI)वह मान है जो निर्णायक अधिक सूचना एकत्रित करने के लिए व्यय कर सकता है। पूर्ण सूचना के प्रत्याशित मान की गणना निम्न प्रकार से की जा सकती है-

(i) Expected Value With Perfect Information

= [Best outcome for 1st event × Probability of 1st event] +

[Best outcome for 2nd event × Probability of 2nd event] +

[Best outcome for nth event × Probabiltity of nth event]

(ii) EVPI = Expected value with perfect Information - EMV of optimal act.

वैकल्पिक रूप में, अनुकूलतम हल से सम्बन्धित कृत्य के लिए प्रत्याशित विकल्प त्याग लागत भी पूर्ण सूचना प्रत्याशित मूल्य कहलाती है।

'बे' का निर्णय सिद्धान्त (Bayesian Decision Theory)- इसका विस्तृत वर्णन प्रायिकता सिद्धान्त अध्याय में किया जा चुका है। यहाँ तो सिर्फ इतना ही बताना है कि प्रायिकता की गणना निम्न सूत्र द्वारा की जा सकती है-

$$\frac{\text{Frequency of an Outcome}}{\text{Total Number of Events}}$$

सीमान्त एवं संयुक्त प्रायिकता (Marginal and Joint probability)-इसका विस्तृत विवरण भी प्रायिकता सिद्धान्त के अन्तर्गत किया जा चुका है। सीमान्त विश्लेषण विधि के अनुसार अतिरिक्त इकाई का संचय उस समय तक किया जाना चाहिए जब तक कि अतिरिक्त इकाई के बेचे जाने पर प्राप्त होने वाला प्रत्याशित लाभ उस इकाई के न बिकने पर होने वाली प्रत्याशित हानि से अधिक या बराबर हो।

यदि 'P' अतिरिक्त इकाई के बेचे जाने की प्रायिकता है,'IP'उस इकाई के बेचे जाने पर प्राप्त होने वाले लाभांश तथा 'IL' उस इकाई के न बिकने पर होने वाली हानि हो तो

$$P(IP) \geq (I - P)(IL)$$

Or

$$P \geq \frac{IL}{IP + IL}$$

अर्थात् अतिरिक्त इकाई का संचय उस समय तक किया जाना चाहिए जब तक कि उसके बेचे जाने की प्रायिकता (P) अनुपात $\frac{IL}{IP + IL}$ के बराबर या बड़ी हो।

NOTES

Illustration 1. The management of a company is faced with the problem of choosing one of the two products for manufacturing. The probability matrix after market survey for the two products was as follows:

State of Nature → / ↓ Acts	Good	Fair	Poor
Product 'A'	0.75	0.15	0.10
Product 'B'	0.60	0.30	0.10

The profits that the management can make for different levels of market acceptability of the products are as follows :

Profits (in Rs. if market is)

State of Nature→ / ↓ Acts	Good	Fair	Poor
Product 'A'	35,000	15,000	5,000
Product 'B'	50,000	20,000	(–)3,000

You have to calculate expected value from choice of either of the products.

Solution.

Let us put the above information in a pay-off matrix with probabilities associated with the state of nature.

Market Conditions / Strategy		Good	Fair	Poor
Product 'A'	Probability	0.75	0.15	0.10
	Profit (Rs.)	35,000	15,000	5,000
Product 'B'	Probability	0.60	0.30	0.10
	Profit (Rs.)	50,000	20,000	(–)3,000

The expected values of the two stragegies are :

Product A : (35,000) (0.75) + (15,000) (0.15) + (5,000) (0.10) = **Rs. 29,000**

Product B : (50,000) (0.60) + (20,000) (0.30) + (-3,000) (0.10) = **Rs. 35,700**

The expected pay-off for product B is more and therefore it is preferred.

Illustration 2. A boy purchases a news magazine for Rs. 3 and sells it for Rs. 5 per news magazine. He can return the unsold stock for Re. 1.00 per magazine. Past records of sales are as follows :

Sales of Magazine	20	21	22	23	24	25	26
Probability	0.03	0.05	0.17	0.23	0.30	0.18	0.04

Find the ordering quantity for optimal solution.

Solution.

Firstly the conditional pay-off and than expected monetary value (EMV) will be calculated. Calculation of EMV and conditional pay-off is like this :

Pay off Table (Value in Rs.)

Events	P	Acts						
		20	21	22	23	24	25	26
20	0.03	40	38	36	34	32	30	28
21	0.05	40	42	40	38	36	34	32
22	0.17	40	42	44	42	40	38	36
23	0.23	40	42	44	46	44	42	40
24	0.30	40	42	44	46	48	46	44
25	0.18	40	42	44	46	48	50	48
26	0.04	40	42	44	46	48	50	50
EMV		**40**	**41.74**	**43.56**	**44.56**	**44.64**	**43.52**	**41.60**

Calculation of Conditional pay-off

Ordered Units	Demand	Profit on Sale (in (Rs.)	Loss on Not Sold Goods (in Rs.)	Conditional Profit
20	20	40	–	40
21	20	40	2	38
	21	42	–	42
22	20	40	4	36
	21	42	2	40
	22	44	–	44
23	20	40	6	34
	21	42	4	38
	22	44	2	42
	23	46	–	46

Similarly other conditional pay-off may be calculated.

Calculation (EMV) Expected Monetary Value

Act EMV

$20 \rightarrow 40 \times 0.03 + 40 \times 0.05 + 40 \times 0.17 + 40 \times 0.23 + 40 \times 0.30 + 40 \times 0.18 + 40 \times 0.04 = 40$

$21 \rightarrow 38 \times 0.03 + 42 \times 0.05 + 42 \times 0.17 + 42 \times 0.23 + 42 \times 0.30 + 42 \times 0.18 + 42 \times 0.04 = \mathbf{41.74}$

$22 \rightarrow 36 \times 0.03 + 40 \times 0.05 + 44 \times 0.17 + 44 \times 0.23 + 44 \times 0.30 + 44 \times 0.18 + 44 \times 0.04 = 43.56$

$23 \rightarrow 34 \times 0.03 + 38 \times 0.05 + 42 \times 0.17 + 46 \times 0.23 + 46 \times 0.3 + 46 \times 0.18 + 46 \times 0.04 = \mathbf{44.56}$

$44 \rightarrow 32 \times 0.03 + 36 \times 0.05 + 40 \times 0.17 + 44 \times 0.23 + 48 \times 0.3 + 48 \times 0.18 + 48 \times 0.04 = \mathbf{44.64}$

and similar other values will be calculated for Act 25 and 26.

Decision. EMV (44.64) is geatest for the Act 24. Therefore for optimum solution 24 magazines should be ordered.

Illustration 3. Under an employment promotion programme it is proposed to allow sale of newspapers on the buses during off-peak hours.

The vendor can purchase the newspapers at a special concessional rate of 0.25 paisa per copy against the selling price of 40 paise. Any unsold copies are however, a dead loss.

NOTES

A vendor has estimated the following probability distribution for the number of copies demanded :

Number of Copies	15	16	17	18	19	20
Probability	0.04	0.19	0.33	0.26	0.11	0.07

How many copies should be ordered so that his expected profits will be maximum?

Solution.

The vender does not purchase less than 15 copies or more than 20 copies. His profit is determined by the demand (D) and the number of copies purchased (P). Whan the demand is more than the number of copies purchased by him profit will be:

$$\text{Rs. } 0.40 \times P - \text{Rs. } 0.25 \times P \qquad \text{....(i)}$$

When the demand is less or equal to the number of copies purchased by him his profit is:

$$\text{Rs. } 0.40 \times D - \text{Rs. } 0.25 \times P \qquad \text{....(ii)}$$

From above equation (i) and (ii),we can construct the pay-off Table

D \ P	15	16	17	18	19	20
15	2.25	2.00	1.75	1.50	1.25	1.00
16	2.25	2.40	2.15	1.90	1.65	1.40
17	2.25	2.40	2.55	2.30	2.05	1.80
18	2.25	2.40	2.55	2.70	2.45	2.20
19	2.25	2.40	2.55	2.70	2.85	2.60
20	2.25	2.40	2.55	2.70	2.85	3.00
EMV	**2.25**	**2.38**	**2.44**	**2.37**	**2.19**	**1.97**

Calculation of Expected Profits

Act

$15 = (2.25 \times 0.04) + (2.25 \times 0.19) + (2.25 \times 0.33) + (2.25 \times 0.26) + (2.25 \times 0.11) + (2.25 \times 0.7) = 2.25$

$16 = (2 \times 0.04) + (2.4 \times 0.19) + (2.4 \times 0.33) + (2.4 \times 0.26) + (2.4 \times 0.11) + (2.4 \times 0.07) = 2.38$

$17 = (1.75 \times 0.04) + (2.15 \times 0.19) + (2.55 \times 0.33) + (2.55 \times 0.26) + (2.55 \times 0.11) + (2.55 \times 0.07) = 2.44$

$18 = (1.5 \times 0.04) + (1.90 \times 0.19) + (2.30 \times 0.33) + (2.70 \times 0.26) + (2.70 \times 0.11) + (2.70 \times 0.07) = 2.37$

$19 = (1.25 \times 0.04) + (1.65 \times 0.19) + (2.05 \times 0.33) + (2.45 \times 0.26) + (2.85 \times 0.11) + (2.85 \times 0.07) = 2.19$

$20 = (1.00 \times 0.04) + (1.40 \times 0.19) + (1.80 \times 0.33) + (2.20 \times 0.26) + (2.60 \times 0.11) + (3.00 \times 0.07) = 1.97$

Decision. It is clear from the above calculations that the expected profit is maximum in the third action i.e., 17. Hence he should order 17 copies in order to maximise his profits.

Illustration 4. Calculate the expected monetary value (EMV) of the three conditions with prior probabilities and pay-offs indicated therein:

Events	Probability	Acts		
		A_1	A_2	A_3
E_1	0.1	25	–10	–125
E_2	0.7	400	440	400
E_3	0.2	650	740	750

Solution.

Computation of Expected Monetary Values

Events	Probability	Acts		
		A_1	A_2	A_2
E_1	0.1	$25 \times 0.1 = 2.5$	$-10 \times 0.1 = -1.0$	$-125 \times 0.1 = -12.5$
E_2	0.7	$400 \times 0.7 = 280$	$400 \times 0.7 = 308$	$400 \times 0.7 = 280$
E_3	0.2	$650 \times 0.2 = 130$	$740 \times 0.2 = 148$	$750 \times 0.2 = 150$
Total	1.00	412.5	455.0	417.5

The highest expected value of pay-off is from Act A_2

बोध प्रश्न

1. निर्णयन वृक्ष को समझाइए।

....

....

....

2. निर्णयन समस्या के तत्व लिखिए।

....

....

....

3. पे. ऑफ तथा रिग्रेट क्या है ?

....

....

....

9.9 सारांश

निर्णयन व्यावसायिक प्रबंध का आधार होता है जो कि सांख्यिकीय सिद्धान्तों व तकनीकों पर आधारित होता है। प्रबंध अपनी समस्त गतिविधियाँ एवं कार्य निर्णयन के आधार पर ही करता है एवं यह एक सतत् जारी रहने वाली प्रक्रिया है। जिसके अंतर्गत प्रबंधक योजना तैयार करता, रणनीति बनाता है, नीति निर्धारण, कर्मचारियों का चयन, नियुक्ति एवं व्यवसाय से संबंधित अन्य प्रकार के कार्य निर्णयन से ही संभव होते हैं।

9.10 शब्दावली या शब्दकुंजी

NOTES

9.11 अभ्यास प्रश्न

दीर्घ उत्तरीय प्रश्न (Long answer type questions)

1. सांख्यिकीय निर्णय सिद्धान्त क्या है ? निर्णय पर्यावरण की विवेचना कीजिये।
   What is statistical decision theory? Disens decisin environment.
2. सम्भावना का उपयोग न करते हुए अनिश्चितता की अवस्था में निर्णय लेने की विभिन्न विधियों का उल्लेख कीजिये।
   Describe the various methods of decision making under uncertainly without use of probobility.
3. पे ऑफ तालिका की सहायता से रिग्रेट तालिका की प्रविष्टियाँ किस प्रकार प्राप्त की जाती हैं ?
   How can entries in a Regret table be derived from a pay-off table?

लघु उत्तरीय प्रश्न (Short answer type Questions)

1. सांख्यिकीय निर्णय सिद्धान्त क्या है ?
   What is statistical decision Theory?
2. पे ऑफ तथा रिग्रेट क्या है ?
   What are pay-offs and Regrets?
3. निर्णय वृक्ष क्या है ?
   What is Decision tree?
4. EMV को समझाइये।
   Explain EMV.

9.12 व्यावहारिक प्रश्न

1. तीन कृत्य A, B, C तथा तीन प्राकृत अवस्थाएँ P, Q, R की प्राप्तियाँ नीचे दी गयी हैं-
   Pay-offs of three acts A, B, C and states of nature P, Q, R are given below :

प्राप्तियाँ (Pay-offs) (in Rupees)

प्राकृत अवस्था (State of Nature)	क्रियाएँ (Acts)		
	A	B	C
P	–35	120	–100
Q	250	–350	200
R	550	650	700

प्राकृत अवस्थाओं की प्रायिकताएँ क्रमश: 0.5, 0.1 तथा 0.4 हैं। उपर्युक्त आँकड़ों के लिए प्रत्याशित मौद्रिक मूल्य सारणी बनाइए तथा यह भी बताइए कि कौन-सी क्रिया को सर्वोत्तम मान कर चुना जा सकता है ?

The probabilities of the states of nature are 0.5, 0.1 and 0.4 respectively. Tabulate the Expected Monetary Values for the above data and state which can be chosen as the best act?

[उत्तर: प्रत्याशित मौद्रिक मूल्य सारणी

प्राकृत अवस्था	क्रिया		
	A	B	C
P	–17.5	60	–50
Q	25	–35	20
R	220	260	280
EMV	227.5	285	250

न्यूनतम EMV, क्रिया A के लिए है, अत: क्रिया A सर्वोत्तम है।]

2. निम्नलिखित प्राप्ति-आव्यूह पर विचार कीजिए:

प्राकृत अवस्था	प्रायिकता	कृत्य या क्रिया		
		विस्तार न करना	100 इकाइयों का विस्तार	200 इकाइयों का विस्तार
ऊँची माँग	0.2	3,500	4,500	6,000
औसत माँग	0.3	3,500	4,500	3,500
नीची माँग	0.5	3,500	3,500	2,000

ई. एम. वी. कसौटी का उपयोग करते हुए सबसे अच्छा कृत्य चुनिए।

Given the following pay-off matrix :

State of Nature	Probability	Act or Action		
		Do not Expand	Expand 100 units	Expand 200 units
Hight Demand	0.2	3,500	4,500	6,000
Medium Demand	0.3	3,500	4,500	3,500
Low Demand	0.5	3,500	3,500	2,000

Using EMV criterion decide the best act.

[उत्तर: अधिकतम (3,500 12,200 3,250) = 12,200.
100 इकाई का विस्तार करना सबसे अच्छा कृत्य है]

3. निम्न सारणी में तीन क्रियाओं A_1, A_2, A_3 तथा प्राकृत अवस्थाओं E_1, E_2, E_3 के पे-ऑफ दिए गए हैं। प्राकृत अवस्थाओं की प्रायिकताएँ क्रमशः 0.1 , 0.7 तथा 0.2 हैं। सुअवसर-हानि सारणी बनाइए तथा सबसे उत्तम क्रिया बताइए:

पे-ऑफ सारणी

प्राकृत अवस्था	क्रिया		
	A_1	A_2	A_3
E_1	25	–10	–125
E_2	400	440	400
E_3	650	740	750

4. Given is the following pay-off matrix:

State of Nature	Probability	Decision		
		Do not Expand Rs.	Expand 200 units (Rs.)	Expand 400 units (Rs.)
Hight Demand	0.4	2,500	3,500	5,000
Medium Demand	0.4	2,500	3,500	2,500
Low Demand	0.2	2,500	1,500	1,000

What should be the decision if we use (i) EMV criterion,(ii) The minimax criterion, and (iii) the maximax criterion?

5. The probalility of the demand for lorries for hiring on any day in a given district is as follows:

Number of Lorries Demanded	0	1	2	3	4
Probability	0.1	0.2	0.3	0.4	0.2

Lorries have a fixed of Rs.90 each a day to keep and daily hire charge (net of variable costs of running) is 200. If the lorry-hire company owns four lorries, what is its daily expectation? If the lorry-hire company is about to go into business and currently has no lorries, how many lorries should it buy?

NOTES

6. (i) Marketing staff of a certain industrial organization has submitted the following pay-off table, giving profits in million rupees, concerning a certain proposal depending upon the rate of technological advance in the next three years:

Decision		
Technological advance	Accept	Reject
Much	2	3
Little	5	2
None	–1	4

The probabilities are 0.2, 0.5 and 0.3 for Much, Little and none technological advance respectively.What decision should be taken?

(1) It is known that 40% of the students in a certain college are girls and 50% of the students are above the median height. If $\frac{2}{3}$ of the boys are above the median height, what is the probability that a randomly selected student who is below the median height is a girl ? (Answer the problem preparing a joint probability table).

7. A management is faced with the problem of choosing one of the products for manufacturing. The potential demand for each product may turn out to be good, moderate or poor. The probabilities for each of the states of nature were estimated as follows:

Product	Nature of Demand		
	Good	Moderate	Poor
X	0.70	0.20	0.10
Y	0.50	0.30	0.20
Z	0.40	0.50	0.10

The estimated profit or loss under the three states may be taken together as:

	Rs	Rs.	Rs.
X	30,000	20,000	10,000
Y	60,000	30,000	20,000
Z	40,000	10,000	–15,000 (loss)

Prepare the expected value teble, and advise the management about the choice of product.

8. A Newpaper distributor assigns probabilities to the demand for a magazine as follows:

Copies	Probability
1	0.4
2	0.3
3	0.2
4	0.1

A copy of the magazine sells for Rs.7 and costs Rs. 6. What can be the maximum possible expected monetary value (EMV) if the distributor can return unsold copies for Rs.5 each?

[Ans. If the distributor can return back unsold copies for Rs.5 each, the maximum possible EMV is Rs.1.20.]

9. A group of students raises money each year by selling Souvenirs outside the stadium after a criket match between Teams A and B. They can buy any of the three different types of Souvenirs from a supplier. Their conditional pay-off table is as under:

	Types of Souvenir			
		I	II	III
Team A wins	Rs.	1,200	800	300
Team B wins	Rs.	2.50	700	1,100

(i) Construct the opportunity loss table.

(ii) Which type of Souvenir should the students buy if probability of Team A's winning is 0.6?

(iii) Find out the cost of uncertainty.

10. The conditional pay-offs in rupees for each action-event combination are as under:

Event	Action			
	1	2	3	4
	Rupees			
A	4	–2	7	8
B	0	6	3	5
C	--5	9	2	–3
D	3	1	4	5
E	6	6	3	3

(i) Which is the best action in accordance with the maximum criterion?

(ii) Which is the best action in accordance with EMV criterion, presuming all events have equal probabilities of occurrences.

11. A management is faced with the problem of choosing one of the three products for manufacturing. The potential demand for each product may turn out to be good. fair or poor. The probabilities for each type of demand were estimated as follow:

Product	Type of Demand		
	Good	Fair	Poor
A	0.75	0.15	0.10
B	0.60	0.30	0.10
C	0.50	0.30	0.20

The estimated profit of loss under the three states of demand in respect of each product may be taken as:

	Rs.	Rs.	Rs.
A	35,000	15,000	5,000
B	50,000	20,000	loss or 3,000
C	60,000	30,000	20,000

Prepare the expected value table and advice the management about the choice of the product.

12. Pay-offs of three Acts A, B and C the states of nature P.Q.R. are given below:

State of Nature	Pay-offs (Rupees)		
	A	B	C
P	–35	120	–100
Q	250	–350	200
R	550	650	700

The probabilities of the states of nature are 0.5, 0.1 and 0.4 respectively. Tabulate the expected monetory value for the above data and state which can be chosen as the best act.

अपनी प्रगति की जाँच करें
Test your Progress

सुझाव पत्र (विद्यार्थियों के लिये)

नाम	–	कार्यक्रम का नाम	–
नामांकन नं.	–	कोर्स का नाम	–
फोन नं.–		सत्र	–
ई–मेल आईडी –			

प्रिय छात्र–छात्राओं,

विश्वविद्यालय के द्वारा दूरस्थ शिक्षण संस्था में पंजीकृत छात्र–छात्राओं को दी जाने वाली पाठ्यसामग्री को हमेशा बेहतर बनाने का प्रयास रहा है। इस उद्देश्य की पूर्ति हेतु आपके विचार एवं सुझाव आंमत्रित है, कृपया आपको प्रदान की जाने वाली पाठ्य–सामग्री के संबंध में अपने विचार एवं सुझाव 500 शब्दों में लिखकर प्रेषित करें, ताकि उक्त विचार एवं सुझाव का अमल करते हुये हम अपने पाठ्य सामग्री को और अधिक सरल, सहज एवं रोचक बनाया जा सकें।

सुझाव –

....

....

....

....

....

....

....

....

....

छात्र का नाम एवं हस्ताक्षर

सुझाव पत्र (विषय विशेषज्ञ/पाठ्यक्रम समन्वयक/कार्यक्रम समन्वयक के लिये)

नाम — पद —
विभाग/विषय — पता —
फोन नं.— सत्र —
ई–मेल आईडी —

प्रिय विषय विशेषज्ञ/पाठ्यक्रम समन्वयक/कार्यक्रम समन्वयक,

विश्वविद्यालय के द्वारा दूरस्थ शिक्षण संस्था में पंजीकृत छात्र–छात्राओं को दी जाने वाली पाठ्यसामग्री को हमेशा बेहतर बनाने का प्रयास रहा है। इस उद्देश्य की पूर्ति हेतु आपके विचार एवं सुझाव प्रार्थनीय है, कृपया आप इस पाठ्य–सामग्री के संबंध में अपने विचार एवं सुझाव 500 शब्दों में लिखकर प्रेषित करें, ताकि उक्त विचार एवं सुझाव का अमल करते हुये हम अपने पाठ्य सामग्री को और अधिक सरल, सहज एवं रोचक बनाया जा सकें।

सुझाव —

....

....

....

....

....

....

....

....

....

धन्यवाद,

नाम एवं हस्ताक्षर