इकाई 1 परिकल्पना परीक्षण की प्रक्रिया
(Process of Hypothesis Testing)

1.1 प्रस्तावना (Introduction)

सांख्यिकीय जॉच में , विचाराधीन वस्तुओं की कुलता को समग्र कहा जाता है। एक समग्र जिसमें वस्तुओं/चीजों/इकाईयों की परिमित संख्या होती है उसे परिमित समग्र कहा जाता है। एक समग्र जिसमें वस्तुओं/चीजों/इकाईयों की अपरिमित संख्या हेाती है उसे अपरिमित समग्र कहा जाता हैं । समग्र से संबंधित इकाईयों में विशेषताऐं जैसे ऊँचाई , वजन आदि शामिल है। इस प्रकार , सांख्यिकीय जॉच में हम छात्रों की ऊँचाईयों का जो एक विद्यालय में अध्ययन करते हैं , का हवाला दे सकते हैं। दूसरी परिस्थिति में हम आमों के वजनों का जो एक पेड में व्यस्क हो रहें हैं का हवाला दे सकते हैं। जब समग्र विशाल है , सांख्यिकीय जॉच करते समय, हम समग्र में प्रत्येक इकाई के साथ संपर्क करने में सक्षम नहीं हो सकते है। इसलिए , जॉच नमूने पर आधारित हो सकती है (समग्र का एक प्रतिनिधि भाग) इस परिस्थिति में जॉच को नमूना सर्वेक्षण कहा जाता है। मान लीजिए 'N' आकार के समग्र से n इकाईयों को चयनित करना है आकार 'n' में एक नमूने से ये चयनित इकाईयाँ समग्र के एक नमूने से निकाली जाती है, यदि इकाईयों का चयन निश्चित पूर्व निर्धारित प्रायिकताओं के अनुसार होता है, इस तरह के नमूने को **यादृच्छिक नमूना** कहा जाता है। यदि सभी इकाईयों के लिए प्रायिकताऐं समान हैं, तो चयन सरल **यादृच्छिक नमूनाकरण** कहलाता है।

समग्र में चर राशि के लिए, मान लीजिए हम जैसे माध्य, मानक विचलन आदि स्थिर राशि ज्ञात करते हैं तो इन स्थिर राशियों को समग्र का प्राचल कहा जाता है। दूसरी तरफ, यदि हम नमूने का माध्य, मानक विचलन आदि ज्ञात करते हैं, उन्हें **सांख्यिकीय** कहा जाता है। प्राचलन समग्र की एक सांख्यिकीय स्थिर राशि है। आंकडा नमूनों मानो का एक फलन है।

इस प्रकार, एक विद्यालय के छात्रों की माध्य ऊँचाई एक प्राचल है। जबकि विद्यालय के 50 यादृच्छिक चयनित छात्रों की माध्य ऊँचाई एक ऑकडा है।

समग्र का सांख्यिकीय वितरण एक प्राचल (**पायसन वितरण**) के द्वारा निर्धारित किया जा सकता है या इसका निर्धारण एक से ज्यादा प्राचलों द्वारा किया जा सकता है जैसे सामान्य वितरण में दो प्राचल माध्य एवं मानक विचलन होते हैं, द्विपद वितरण में दो प्राचल n एवं p होते हैं समग्र प्राचल के सभी स्वीकार्य मानो का समूह (समुच्चय) **प्राचल अतंराल** कहलाता है। यदि एक प्राचल अकेले समग्र के सांख्यिकीय वितरण के वर्णन के लिए पर्याप्त है तो प्राचल अंतराल एक आयामी है। दूसरी तरफ , यदि दो प्राचल समग्र का वर्णन करते हैं तो प्राचल अंतराल द्वि आयामी है और इसी तरह।

1.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप-

- ✓ परिकल्पना के परीक्षण के विभिन्न तरीकों को समझ सकेंगे।
- ✓ परिकल्पना के परीक्षण के अनुप्रयोगों के वर्णनों को जानेंगे।

1.3 नमूने का वितरण एवं मानक त्रुटि(Sampling distribution and standard error)

मान लीजिऐ एक समग्र से 'n' आकार का एक नमूना लिया जाता है और नमूना माध्य $\bar{x}$ की गणना की जाती है। समग्र में से समान आकार के इस तरह के बेहत से नमूने लिये जा सकते है। 'n' आकार के सभी नमूनों का समुच्चय जिसे समग्र में से लिया जा सकता है , नमूना अंतराल कहा जाता है। प्रत्येक नमूने के लिए $\bar{x}$ की गणना की जा सकती है। इसलिए $\bar{x}$ के बहुत मान हो सकते हैं मान लीजिए प्राचल अंतराल में, $\bar{x}$ के इन विभिन्न मानो को आवृत्ति वितरण के रूप में सारणीबद्ध किया गया हो , तो परिणामी वितरण को $\bar{x}$ का नमूने का वितरण कहा जाता है। इस नमूने के वितरण का मानक विचलन मानक त्रुटि (S.E) कहलाता है।

समान आकार के विभिन्न नमूनों के लिए ऑकडों के मानों का वितरण आंकडों का नमूना वितरण कहलाता है।

ऑकडों की मानक त्रुटि , ऑकडों के नमूने वितरण का मानक विचलन है। अन्य आंकडों का नमूना वितरण जैसे नमूना माध्यिका भी लिखा जा सकता है।
इन प्रत्येक परिस्थिति में समतुल्य मानक विचलन मानक त्रुटि (S.E) होगा!
एक समग्र जिसका माध्य (μ) एवं मानक विचलन (σ) है का ध्यान करें। चर्चाऐं आरामदायक होंगी , इसलिए यहाँ हम अपने को केवल बडे समग्र तक सीमित करते हैं तब $\bar{x}$ का नमूना वितरण का माध्य μ और मानक त्रुटि $\frac{\sigma}{\sqrt{n}}$ है। जो कि $E(\bar{x}) = \mu$ और $S.E(\bar{x}) = \frac{\sigma}{\sqrt{n}}$ है।
यदि समग्र से आकार n_1 का यादृच्छिक नमूना लिया जाता है जिसका माध्य μ_1 है और मानक विचलन σ_1 है। दूसरे समग्र से यदि n_2 आकार का यादृच्छिक नमूना भी लिया जाता है जिसका माध्य μ_2 और मानक विचलन σ_2 है यदि $\overline{x_1}$ पहले नमूने का माध्य हो और $\overline{x_2}$ दूसरे नमूने का माध्य हो तो

$$E(\overline{x_1} - \overline{x_2}) = (\mu_1 - \mu_2)\ and\ S.E.(\overline{x_1} - \overline{x_2}) = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}$$

मानक त्रुटि या अनुपात
समग्र में, मान लीजिए इकाईयों का द्विपालिक वर्गीकरण (दो वर्गों में वर्गीकरण) उन इकाईयों के रूप में संभव है जो एक गुण रखते हैं और जो गुण नहीं रखते हैं उदाहरण के लिए

(i) गॉव के लोगों का वर्गीकरण जैसे शिक्षित एवं अशिक्षित
(ii) विद्यालय के छात्रों का वर्गीकरणजैस गरीब एवं अमीर
(iii) फलों का वर्गीकरण जैसे पका एवं कच्चा
(iv) कर्मचारियों का वर्गीकरण जैसे संतुष्ट एवं असंतुष्ट
(v) ग्राहकों का वर्गीकरण

समग्र में यदि p इकाईयों का अनुपात है जो विशेषता (गुण) रखता है। इस तरह के समग्र में से, मान लीजिए एक n आकार का यादृच्छिक नमूना लिया जाता है। यदि x इन n इकाईयों के वर्ग से सम्बन्धित है जो इस गुण को रखते हैं।
तब, $p = \frac{x}{n}$ गुण का एक नमूना अनुपात है।
यहाँ $p = \frac{x}{n}$ माध्य $E(p) = p$
और मानक त्रुटि $S.E.(p) = \sqrt{\frac{pq}{n}}$
जहाँ $Q = 1 - p$
यदि एक आकार n_1 का सहज गुण यादृच्छिक नमूना एक समग्र से p_1 अनुपात के साथ लिया जाता है। यदि x_1 इकाईयों में नमूना सहजगुण रखता है। तब , नमूना अनुपात $p_1 = \frac{x_1}{n_1}$ है। यदि एक आकार n_2 का सहजगुण यादृच्छिक नमूना समग्र से p_2 अनुपात के साथ लिया जाता है। यदि x_2 इकाईयो में नमूना सहजगुण रखता है। तब नमूना अनुपात $p_2 = \frac{x_2}{n_2}$ है। यहाँ नमूना अनुपातों के माध्य में अन्तर $E(p_1 - p_2) = (p_1 - p_2)$ और मानक त्रुटि $S.E.(p_1 - p_2) = \sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$ जहाँ $Q_1 = 1 - p_1$ और $Q_2 = 1 - p_2$
यहाँ यदि $P_1 = P_2 = P$
तब मानक त्रुटि $\sqrt{pq(\frac{1}{n_1} + \frac{1}{n_2}}$ है।

इस प्रकार, कुछ ऑकडों के माध्यों एवं मानक त्रुटियॉं निम्नवत है।

आंकड़ा	माध्य	मानक त्रुटि
x (Sample mean)	μ	$\frac{\sigma}{\sqrt{n}}$
$\overline{x_1} - \overline{x_2}$ (Difference of means)	μ_1-μ_2	$\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}$
P (Sample mean)	P	$\sqrt{\frac{PQ}{n}}$
P_1-P_2 (Difference of proportions)	P_1=P_2	$\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}$
P_1-P_2 (when P_1=P_2=P)	0	$\sqrt{PQ[\frac{1}{n_1}+\frac{1}{n_2}]}$

मानक त्रुटि की उपयोगिता

मानक त्रुटि आंकडे की परिवर्तनशीलता की माप है। यह आंकलन एवं परिकल्पनाओं के परीक्षण के लिए उपयोगी है।

(i) आंकलन के सिद्धान्त में , मानक त्रुटि का प्रयोग एक अनुमानक के रूप में ऑकडों की दक्षता एवं तालमेंल को निर्धारित करना है।

(ii) आंकलन अंतराल में, मानक त्रुटि का प्रयोग विश्वसनीयता अंतरालों को लिखना है।

(iii) परिकल्पनाओं के परीक्षण में , परीक्षण की मानक त्रुटि का उपयोग ऑकडे परीक्षण के विवरण को मानकीकृत करने के लिए किया जाता है।

सांख्यिकीय अनुमान - सांख्यिकीय अनुमान सांख्यिकी की वह शाखा है जो समग्र से लिये गये नमूनों का उपयोग करते हुए समग्र के सांख्यिकीय प्रकृति के बारे में निर्णय लेने के सिद्धान्त और तकनीकों से संबंधित है। सांख्यिकीय अनुमानो की दो शाखाऐं होती है, वो हैं (1) **अनुमान के सिद्धान्त** (2) **अनुमानों का परीक्षण**

1.4 अनुमान के सिद्धान्त (Principles of Estimation)

सांख्यिकीय में, हम प्रायः ऐसी स्थिति में आते हैं जहॉं हम समग्र के एक प्राचल के संभावित मूल्य के बारे में बात करेंगें। मान लीजिए एक बागवानी अनुसंधान केंद्र ने केले का संकर किस्म विकसित किया है। केन्द्र हमसे इस तरह के केले के औसत वजन को जानना चाहता है। इस उद्देश्य के लिए, कुल केले यादृच्छिक चुने जाते हैं और उनके माध्य वजन की गणना की जाती है। यह माध्य वजन (नमूना माध्य) संभावित समग्र माध्य के रूप में प्रस्तावित है। इस प्रकार नमूना माध्य $\bar{x}$, समग्र माध्य μ का एक अनुमानक है। एक विशिष्ट नमूने के लिए मान लीजिए नमूना माध्य वजन 84 ग्राम है , $\bar{x}$ के इस विशिष्ट मान समग्र माध्य का आंकलन कहा जाता है। अज्ञात प्राचल का अनुमानक एक आंकडा है जो उस प्राचल के संभावित मान को निर्दिष्ट करता है। अनुमान विशिष्ट नमूने के लिए अनुमानक का एक विशिष्ट मान है। अनुमानक एक आंकडा है , जबकि अनुमान (आंकलन) एक संख्यात्मक मान है।

अनुमान (आंकलन) समग्र से लिये हुए नमूने के सांख्यिकीय का प्रयोग करते हुए समग्र प्राचल के संम्भावित मान के लिए अपनाई गई विधियों और तकनीकों से सम्बन्धित है। आंकडों के दो प्रकार होते हैं वो हैं , (1) बिन्दु आंकलन (2) अंतराल आंकलन

बिन्दु आंकलन (अनुमान):- एक अज्ञात प्राचल का अनुमान लगाते समय, यदि एक एकल मान अनुमान के रूप में प्रस्तावित होता है, तो इस तरह के अनुमान को बिन्दु अनुमान कहते हैं इस प्रकार , नमूना माध्य $\bar{x} = 84$ ग्राम है के आधार पर यदि हम निष्कर्ष निकालें कि समग्र माध्य 84 ग्राम है यह बिन्दु आंकलन है। यहाँ $\bar{x}$ एक समग्र माध्य μ का बिन्दु आंकलन है। निर्दिष्ट मान 84 ग्राम μ का बिन्दु आंकलन है। आम तौर पर n_1 का बिन्दु आंकलन n_1 को ? द्वारा प्रदर्शित किया जाता है। इसलिए ? = $\bar{x}$

अन्तराल आंकलन :- अंतराल आंकलन में, अज्ञात प्राचल के अनुमान के रूप में हम एक एकल मान का प्रस्ताव करते हैं। अधिकांश स्थितियों में प्रस्तावित मूल्य प्राचल के वास्तविक मूल्य की संभावना नहीं होंती है। इसके बजाय, यदि हम बिन्दु अनुमान के आसपास एक छोटे से अंतराल का प्रस्ताव करते हैं , तो प्राचल को शामिल करने की संभावना अंतराल , हमारा प्रस्ताव मजबूत होगा। यह अंतराल जिसमें प्राचल को शामिल करने की संभावना है **अंतराल अनुमान** कहा जाता है। अंतराल आंकलन में , एक अंतराल (T_1 से T_2) जिसमें प्राचल को शामिल करने की संभावना है प्राचल के अनुमानक के रूप में प्रस्तावित है। अंतराल (T_1 से T_2) को विश्वसनीयता अंतराल कहा जाता है। विश्वास अंतराल में प्राचल के सम्मिलित हो ने की संभावना को विश्वास गुणांक कहा जाता है। विश्वास अंतराल की सीमा (T_1 और T_2) को विश्वास सीमा कहा जाता है। ये सीमाऐं सम्बन्धित बिन्दु अनुमानक (सांख्यिकी) के नमूनाकरण वितरण पर आधारित है।

विश्वास अंतराल में प्राचल शामिल होने की संम्भावना को विश्वास गुणांक कहा जाता है। इसे (1-α)है। अन्तराल अनुमान में , हम विभिन्न विश्वास गुणंकों के विश्वास अंतराल लिख सकते हैं , जैसे 9 5%, 99% इत्यादि।

उदाहरण:- हम केले के वनज के बारे में पहले उर्वधृत उदाहरण पर गौर करते हैं। यहाँ , केलों का अज्ञात माध्य वजन (समग्र माध्य μ) का अनुमान नमूना आकार $n = 100$ के प्रयोग से लगाया गया है। जिसके लिए नमूना माध्य $\bar{x} = 84$ ग्राम है। यदि समग्र मानक विचलन $\sigma = 5$ ग्राम है तो केले का वजन सामान्य रूप से वितरित किया जाता है। तब, $\bar{x}$

माध्य 84 ग्राम और मानक विचलन $\frac{\sigma}{\sqrt{n}} = \frac{5}{100}$

1. $\bar{x} = 84$ ग्राम μ का एक बिन्दु अनुमानक है। इसलिए , हम कहते हैं कि केलों का माध्य वजन 84 ग्राम है।
2. μ के लिए 95% विश्वसनीयता अंतराल

   $(\bar{x}- 1.96\frac{\sigma}{\sqrt{n}},\ \bar{x}+1.96\frac{\sigma}{\sqrt{n}})$

   जो $(84 - 2.58\times \frac{5}{\sqrt{100}},\ 84+2.58\times \frac{5}{\sqrt{100}})$ है।

इसलिए 95% विश्वसनीयता के साथ हम कहते हैं कि केलों का माध्य वजन 83.02 एवं 84.98 ग्राम के मध्य है।

μ के लिए 99% विश्वसनीयता $(\bar{x}- 2.58\frac{\sigma}{\sqrt{n}}, \bar{x}+2.58\frac{\sigma}{\sqrt{n}})$

जो $(84 - 2.58\times \frac{5}{\sqrt{100}},\ 84+2.58\times \frac{5}{\sqrt{100}})$

जो (81.62 , 85.29) है। इसलिए, 99% विश्वसनीयता के साथ हम कहते हैं कि कैलों का माध्य वजन 81.61 और 85.29 ग्राम के मध्य है।

टिप्पणी -1 हम जानते हैं कि द्विपद समग्र में माध्य np एक अनुमानक है इसलिए p का अनुमानक

$$\overline{p} = \frac{\overline{x}}{n} = \frac{3.38}{7} = 0.48$$

टिप्पणी -2 पायसन समग्र में माध्य का अनुमानक $\overline{\lambda} = \overline{x} = 0.2$

टिप्पणी -3 पायसन समग्र में माध्य n_1 का अनुमानक $\overline{\lambda} = \overline{x} = 1.2$ है।

1.5 परिकल्पनाओं का परीक्षण (Testing of hypotheses)

अनुमानों का परीक्षण समग्र से लिऐ गये नमूनों का उपयोग करके समग्र के प्राचल के संबंध में अनुमानों की वैधता के सत्यापन के साथ संबंधित हैं मान लीजिए हम मानते हैं कि X शहर के निवासी की औसत आय रू0 50 प्रतिदिन है। इस परिकल्पना का परीक्षण करने के लिए, X शहर के कुछ निवासीयों को यादृच्छिक तरीके से चयनित किया गया और उनकी औसत दैनिक आय ज्ञात की गई। यदि यह नमूना माध्य रू0 50 के करीब है तो हम यह निष्कर्ष निकालते है कि X शहर के निवासी की औसत आय 50 रूपये है इस प्रकार यदि नमूना माध्य रू0 50.60 है, चूकि यह मान रू0 50 के करीब है , हम यह निष्कर्ष निकालते हैं कि X शहर के निवासी की औसत आय रू0 50 प्रतिदिन है। दूसरी और यदि नमूना माध्य रू0 68.60 है चूंकि यह मान रू0 18 से दूर है , तो हम यह निष्कर्ष निकालते हैं, कि X शहर के एक निवासी की औसत आय रू0 50 से भिन्न है।

सांख्यिकीय परिकल्पना:- एक सांख्यिकीय परिकल्पना समग्र के सांख्यिकीय विवतरण के संबंध में एक अभिप्राय है। यह समग्र के मापदंडों के बारे में बयान है।

सांख्यिकीय परिकल्पना को H द्वारा प्रदर्शित किया जाता है।

उदाहरण

1) H: समग्र का माध्य $\mu=25$ ।

2) H: समग्र सामान्य रूप से माध्य $\mu=25$ और मानक विचलन $\sigma=2$ के साथ वितरित है।

एक अनुमान जो पूरी तरह से समग्र के सांख्यिकीय वितरण को निर्दिष्ट करती है, सरल परिकल्पना कहा जाता है यह एक ऐसी अवधारणा है जो समग्र के सभी मापदंडों को निर्दिष्ट करता है। H: समग्र का माध्य $\mu=25$ के साथ सामान्य रूप से वितरण एक सरल परिकल्पना है।

एक परीक्षण प्रक्रिया की शुरूआत करने के लिए , एक परिकल्पना बना दी जाती है , इस परिकल्पना की वैद्यता का परीक्षण किया जाता है यदि परिकल्पना सही पायी जाती है तो इसे स्वीकार किया जाता है। दूसरी ओर, यह गलत पायी जाती है तो इसे अस्वीकार कर दिया जाता है। परिकल्पना संम्भव नकारों की जॉच की जा जिसमें रही है , शून्य परिकल्पना कहा जाता हे। शून्य परिकल्पना को μ_0 से प्रदर्शित किया जाता है यदि शून्य परिकल्पना को गलत पाया जाता है, तो एक और अवधारणा जो शून्य अवधारणा के विपरीत है , जब शून्य परिकल्पना को अस्वीकार किया जाता है , दूसरी परिकल्पना को स्वीकार किया जाता है उसे वैकल्पिक परिकल्पना कहते हैं। वैकल्पिक परिकल्पना को H_1 से प्रदर्शित करते है।

मान लीजिए एक शून्य परिकल्पना H_0: $\mu_1=100$ है(औसत मजदूरी 100 रू0 है)

परिस्थितियों के आधार पर वैकल्पिक परिकल्पना निम्न में से कोई भी हो सकती है।

H_1: ≠ रू0 100 (औसत मजदूरी 100 रूपये से भिन्न हो)

H_2: ≠ रू0 100 (औसत मजदूरी 100 रूपये से ज्यादा हो)

H_3: ≠ रू0 100 (औसत मजदूरी 100 रूपये से कम हो)

परीक्षण प्रक्रिया - एक शून्य परिकल्पना के परीक्षण के एिल सांख्यिकीय प्रक्रिया के चरण इस प्रकार है-

1) शून्य परिकल्पना की स्थापना

2) वैकल्पिक परिकल्पना की स्थापना
3) परीक्षण ऑकडों और शून्य वितरण की पहचान करना
4) Critical क्षेत्र की पहचान करना
5) यादृच्छिक नमूने को लेना और वास्तव में इसके परीक्षण का आयोजन करना।
6) निर्णय लेना (अनुमान देना)

1) शून्य परिकल्पना की स्थापना - संभव अस्वीकृति के लिए परीक्षण किया जा रहा है जो परिकल्पना शून्य अवधारणा है। शून्य परिकल्पना में यह हो सकता है कि

- प्राचल किसी दिए गए मान के बराबर है।
- दो समग्र के एिल प्राचल बराबर है।
- अन्तर नगण्य है। (सार्थक नहीं है)
- वितरण उपयुक्त है।
- विशेषताऐं स्वतन्त्र है, और इसी तरह।

2) वैकल्पिक परिकल्पना की स्थापना-

परिकल्पना को स्वीकार किया जाता है , जब शून्य परिकल्पना अस्वीकार की जाती है , वैकल्पिक परिकल्पना कहते है। यह हो सकता है कि

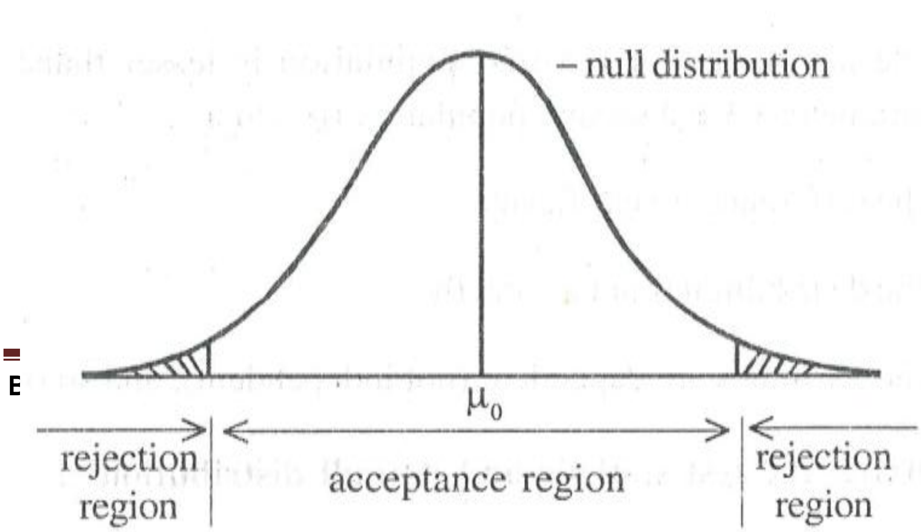
- प्राचल दिए गए मान के बराबर नहीं है।
- प्राचल दिए गए मान से अधिक है।
- दो समग्रो के लिए प्राचल समान नहीं है।
- पहले समग्र का प्राचल दूसरे समग्र के प्राचल से कम है
- अतंर अर्थपूर्ण है।
- वितरण समुचित नहीं है।
- विशेषताऐं निर्भर है। (स्वतन्त्र नहीं)

3) परीक्षण ऑकडों और शून्य वितरण की पहचान करना-

परीक्षण ऑकडे वे ऑकडे हैं जिनके वितरण के आधार पर परीक्षण किया जाता है। H_0 के अंतर्गत परीक्षण ऑकडों के सांख्यिकीय वितरण को शून्य वितरण कहा जाता है। $H_0: \mu=\mu_0$(mean is equal to μ_0) के परीक्षण के लिए, सांख्यिकीय परीक्षण $\bar{x}$ नमूना माध्य है। H_0 के अन्तर्गत $\bar{x}$ का वितरण $N(\mu_0, \frac{\sigma^2}{n})$ है। यह $\bar{x}$ शून्य वितरण है हालाँकि सुविधा के लिए , मानकीकृत रूप $Z = \frac{x-\mu_0}{\sigma}$ समझा जा सकता है जिसमें शून्य N(0,1) वितरण है।

4) Critical (अस्वीकृत) क्षेत्र की पहचान करना-

परीक्षण ऑकडों के उन मानों का समूह जो शून्य अवधारणा के लिए अस्वीकृति का नेतृत्व करता है , को अस्वीकृति क्षेत्र कहा जाता है। परीक्षण ऑकडों के उन मानों का समूह जो शून्य अवधारणा की स्वीकृति का नेतृत्व करता है, स्वीकृति क्षेत्र कहा जाता है। स्वीकृति क्षेत्र की सीमांकन सीमा को स्वीकृति मान कहा जाता है स्वीकृति क्षेत्र ऐसा बना हुआ है कि जब H_0 सही हो , इसकी अस्वीकृति की संभावना



एक छोटा सा पूर्व निर्धारित मान है इस पूर्व निर्धारित मान को महत्व का महत्व का स्तर एक पूर्व निर्धारित ऊँपरी सीमा है जो कि वास्तव में सही है जब शून्य परिकल्पना की अस्वीकृति की संभावना होती है। महत्व का स्तर $\propto$ द्वारा प्रदर्शित किया जाता है। आमतौर पर $\propto$ का पूर्व निर्धारित मान 0.05 या 0.01 होगा। दूसरे शब्दों में , यह 5% या 1% होगा।

परीक्षण की प्रकृति और वैकल्पिक परिकल्पना की प्रकृति के आधार पर Critical क्षेत्र का एक हिस्सा हो सकता हे या इसके दो भाग हो सकते है। यदि Critical क्षेत्र में एक भाग है तो परीक्षण को एकतरफा परीक्षण कहते है। यदि Critical क्षेत्र के दो भाग होते हैं, तो उसे दो तरफा परीक्षण कहते हैं।

5) यादृच्छिक नमूने को लेना और वास्तव में इसके परीक्षण का आयोजन करना- आकार 'n' का एक यादृच्छिक नमूना समग्र से लिया जाता है। ऑकडे परीक्षण के मान की गणना नमूना मान से की जाती है। यदि परिकलित मान (प्रेक्षित मान) Critical क्षेत्र को सम्बन्धित है , H_1 के पक्ष में H_0 अस्वीकार होगी दूसरी ओर यदि परिकलित मान स्वीकृति से सम्बन्धित होता है तो H_0 स्वीकार होगी।

6) निर्णय लेना (अनुमान देना)-

अंतिम निर्णय (अनुमान, निष्कर्ष) की घोषणा की जाती है

निर्णय

(अ) नई दवा पुराने की तुलना में अधिक प्रभावी है।

(ब) नर नवजात शिशुओं के बीच नवजात जन्मोत्तर वृद्धि समान होती है, जैसे कि नर शिशुओं में हो सकता है।

प्रथम एवं द्वितीयक प्रकार की त्रुटियॉं

जब एक वैकल्पिक परिकल्पना के विरूद्ध एक शून्य परिकल्पना का प रीक्षण करते है, निम्न परिस्थितियों में एक सामने आती हैं

क्रम संख्या	**वास्तविक तथ्य**	**नमूने के आधार पर निर्णय**		**त्रुटि**
1	H_0 सत्य है।	स्वीकार H_0	सही निर्णय	....
2	H_0 सत्य है।	अस्वीकार H_0	गलत निर्णय	Type I
3	H_0 असत्य है।	स्वीकार H_0	गलत निर्णय	Type II
4	H_0 असत्य है।	अस्वीकार H_0	सही निर्णय	....

यहॉं परिस्थितियों (2) और (3) में , गलत फैसले आते है। इन गलत निर्णयों को क्रमशः पहली तरह की त्रुटि (I प्रकार त्रुटि) और दूसरी तरह की त्रुटि (II प्रकार त्रुटि) कहा जाता हे। इस प्रकार

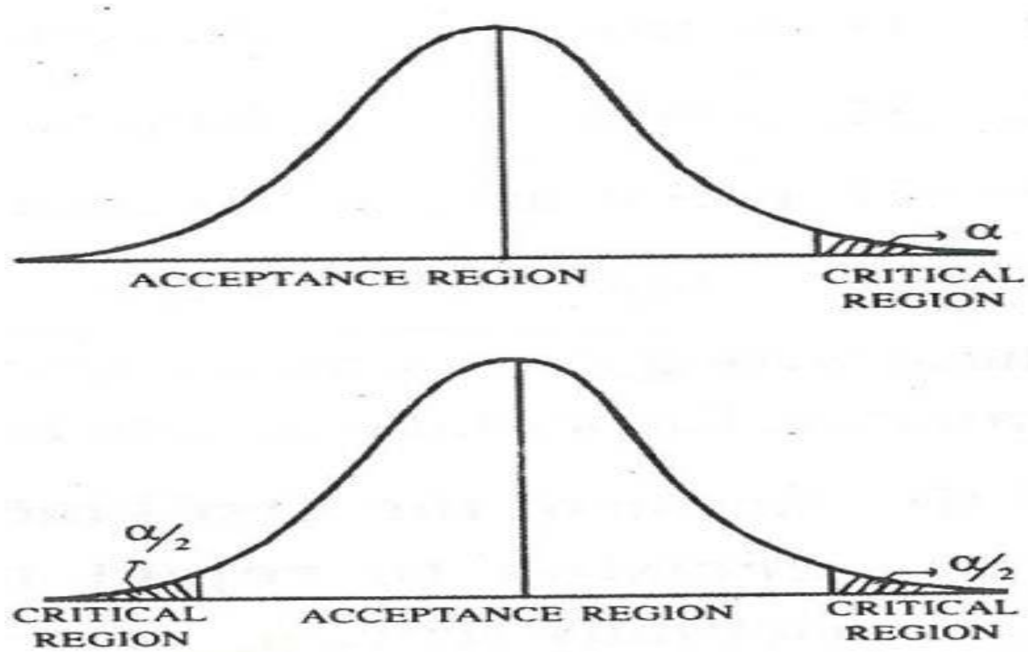
- पहली तरह की त्रुटि वास्तव में सही है जब शून्य अवधारणा के अस्वीकार करने का बेहद गलत निर्णय लिया जाता है।
- द्वितीयक प्रकार की त्रुटि में गलत अनुमानों को स्वीकार करने का गलत निर्णय लिया जाता है , जब यह वास्तव में सत्य है। पहले तरह की त्रुटि होने की संभावना $\propto$ है। यह परीक्षण का एक आकार है।

दूसरी तरह की त्रुटि की संभावना के घटित होने को B द्वारा प्रदर्शित करते हैं

(1-B) के मान को परीक्षण का घात कहा जाता है।

परीक्षण की घात H_0 को अस्वीकार करने की संभावना है जब यह सत्य नहीं है। जबकि परीक्षण, महत्व का स्तर ∝ अग्रिम में तय किया जाता है। फिर महत्वपूर्ण क्षेत्र इस तरह निर्धारित किया जाता है कि (1-B) की अधिकतम हो इस प्रकार Critical मान महत्व के स्तर पर आधारित होता है।

एक पूंछ और दो पूंछ परीक्षण (एक तरफा और दो तरफा परीक्षण)स्तर कहा जाता है।



दूसरी और, यदि Critical क्षेत्र परीक्षण ऑकडों के शून्य वितरण की पूंछ पर माना जाता है, तो परीक्षण दो पुच्छिक परीक्षण (दो तरफा परीक्षण) है। दो पूंछ के मामले में H_0 का मान μ से कम σ से ज्यादा हो सकता है। एक पुच्छीय परीक्षण में $\mu>\mu_0$ या $\mu<\mu_0$ या दिशाऐं दी जाती है।

निम्नलिखित में से कुछ एक पुच्छीय परीक्षण हैं -

- परीक्षण $\mu_0:\mu=\mu_0$ के विपरीत $\mu_1:\mu>\mu_0$
- परीक्षण $\mu_0: \mu_1 = \mu_2$ के विपरीत $\mu_0:\mu_1>\mu_2$
- Goodness of Fit के लिए परीक्षण
- आकस्मिकता तालिका में विशेषताओं की आबादी के लिए परीक्षण

कुछ निम्नलिखित दो पुच्छीय परीक्षण हे।

- परीक्षण $\mu_0: \mu = \mu_0$ के विपरीत $\mu_1:\mu \neq \mu_0$
- परीक्षण $\mu_0: \mu_1 = \mu_2$ के विपरीत $\mu_1:\mu_1 \neq \mu_2$
- परीक्षण $\mu_0\sigma = \sigma_0$ के विपरीत $\mu1:\sigma > \sigma0$

उदाहरण 1 - पिछला अनुभव कहता है कि काजू गिरी में 1.2 ग्राम वजन का मानक विचलन है। 40 माध्य वजन के यादृच्छिक चुने गिरी (दानों) के लिए माध्य त्रुटियाँ ज्ञात कीजिए।

हल- यहाँ, σ = 1.2 ग्राम और n= 40

नमूना माध्य की मानक त्रुटि $S.E(\overline{x}) = \frac{\sigma}{\sqrt{n}}$

$\frac{1240}{\sqrt{}}$= 0.1897 ग्राम है।

उदाहरण 2 - एक कालेज के लडकों के वजन का माध्य एवं मानक विचलन क्रमशः 47 किलोग्राम एवं 31 किलोग्राम है। उसी कालेज के लडकियों के वजन का माध्य एवं मानक विचलन 45 किलोग्राम एवं 2.8 किलोग्राम है। कालेज में से 16 लडके एवं 9 लडकियाँ यादृच्छिक चयनित की जाती है।

I. 16 चयनित लडकों के माध्य वजन का माध्य एवं मानक त्रुटि ज्ञात कीजिए।
II. 9 चयनित लउकियों के माध्य वजन का माध्य एवं मानक त्रुटि ज्ञात कीजिए।
III. चयनित लडकों के औसत वनज एवं चयनित लडकियों के औसत वजन के अन्तर का माध्य एवं मानक त्रुटि ज्ञात कीजिए।

हल:- यहाँ

μ_1=47 किग्रा σ_1= 3.1 किग्रा n_1=16

μ_2=45 किग्रा σ_{12}= 2.8 किग्रा n_2=9

यदि $\overline{x_1}$ चयनित लडको का माध्य वजन है।

$\overline{X}_2$ चयनित लडकियों का माध्य वजन है।

(1) माध्य = E ($\overline{x_1}$) =μ_1=47 किग्रा

S.E ($\overline{x_1}$) $\frac{\sigma_1}{\sqrt{n_1}}$=$\frac{3.116}{\sqrt{}}$= 0.775 किग्रा

(2) माध्य = E($\overline{x_2}$) = 45 किग्रा

S.E ($\overline{x_2}$)= $\frac{\sigma_2}{\sqrt{n_2}}$=$\frac{2.8}{\sqrt{9}}$ =0.993 किग्रा

(3) माध्य = ($\overline{x_1}$- $\overline{x_2}$) = μ_1 –.μ_2

47-45 = 2 किग्रा

$$S..E.(\overline{x}_1 - \overline{x}_2) = \sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} + = \sqrt{\frac{(3.1)^2}{16}} + \frac{(2.8)^2}{9} = 1.213 kgs$$

एक नमूना जॉच , परिणाम पैदा करता है , और इन परिणामों के साथ , समग्र के लिए निर्णय बनाये जाते हैं लेकिन ऐसे निर्णयों में अनिश्चितता का एक तत्व शामिल है, जो गलत फैसले का कारण है। परिकल्पना एक ऐसी धारणा है जो समग्र प्राचल के बारे में सत्य या असत्य भी हो सकती है उदाहरण के लिए , सिक्के को 300 बार उछालने पर कोई 190 चित और 110 बार पट प्राप्त कर सकता है। इस उदाहरण पर , हम यह जांचने में रूचि रखते हैं कि क्या सिक्का निष्पक्ष है या नहीं। इसलिए, हम इसका मूल्यांकन करने के लिए एक परीक्षण आयोजित कर सकते हैं कि क्या अंतर नमूनाकरण के कारण है। एक महत्वपूर्ण परीक्षण के क्रियान्वयन की प्रक्रिया निम्नानुसार है।

अवधारणा का निर्माण - हमारी धारणा को सत्यापित करने के लिए जो नमूना अध्ययन पर आधारित है , हम ऑकडे एकत्र करते हैं और नमूना मान और समग्र मान के बीच के अंतर को पता करते है। यदि इनमें कोई अंतर नहीं है या यदि अंतर बहुत छोटा है , तो हमारा अनुमानित मान सही है , सामान्यतया दो अवधारणाओं का निर्माण किया जाना चाहिए और मगर एक परिकल्पना सही है, तो दूसरे को अस्वीकार करेंगें।

(अ) शून्य परिकल्पना - यह अंतर के महत्व का परीक्षण करने के लिए बहुत उपयोगी उपकरण है। समग्र से सम्बन्धित किसी भी परिकल्पना को सांख्यिकीय अवधारणा कहा जाता है। सांख्यिकीय परीक्षण की प्रक्रिया में , समग्र से प्राप्त नमूने पर आधारित अवधारणा को अस्वीकार या स्वीकार किया जाता है। सांख्यिकीयविद्

अवलोकन के माध्यम से परिकल्पना की जॉच करते हैं और संभावित देते हैं। सरल परिकल्पना से पता चलता है कि नमूना मान और अध्ययन के तहत समग्र मान किसी तरह या अंतर प्रदर्शित नहीं करते है परिकल्पना हमने कल्पित की है, वह शून्य अवधारणा है, इसका अर्थ है कि नमूना के माध्य और समग्र माध्य के बीच वास्त विक अंतर शून्य है कम से कम अंतर होना महत्वहीन है। शून्य परिकल्पना की अस्वीकृति का अर्थ है कि नमूना माध्य और समग्र माध्य के बीच का वास्तविक अंतर काफी महत्वपूर्ण है। शून्य परिकल्पना की अस्वीकृति से पता चलता है कि समय पर वैकल्पिक परिकल्पना।

उदाहरण के लिए -

(1) एक विश्वविद्यालय के विद्यार्थियों की औसत ऊँचाई 155 सेन्टीमीटर है।

(2) एक व्यवसाय की औसत दैनिक बिक्री रू0 1,50,000 है।

(3) किसी विशेष गांव के निवासी की औसत आय 100 रूपये है। इन सभी बयानों को नमूना परीक्षणो के आधार पर सत्यापित करना होगा। सामान्यतया एक परिकल्पना में कहा जाता है कि नमूना माध्य एवं समग्र माध्य के बीच कोई अंतर नहीं हैं या $\mu=\mu_0$ शून्य परिकल्पना को μ_0 द्वारा प्रदर्शित किया जाता है।

(ब) वैकल्पिक परिकल्पना - μ_0 की अस्वीकृति वैकल्पिक परिकल्पना की स्वीकृति की ओर जाता है , जिसे H_1 द्वारा प्रदर्शित किया जाता है जैसे -

$H_0=\mu=155$ (शून्य परिकल्पना)

$H_1=\mu\neq155$ या $\mu>155$ या $\mu<155$ (वैकल्पिक परिकल्पना)

जब दो अवधारणाओं को स्थापित किया जाता है , तो शून्य परिकल्पना की स्वीकृति या अस्वीकृति एक नमूना अध्ययन पर आधारित होती है। इस प्रकार यह दो गलत निष्कर्ष की ओर जाता है। (1) H_0 को अस्वीकृत करते हुए जब H_0 सत्य है (2) H_0 को अस्वीकृत करते हुए जब H_0 असत्य है । इसे निम्नलिखित तालिका में वर्णित किया जा सकता है।

	नमूने में से निर्णय	
	H_0 स्वीकार	H_0 अस्वीकार
H_0 सही	सही	गलत type I त्रुटि
H_0 गलत (H_1 सही)	गलत type II त्रुटि	सही

फिर से लिखते हुए

H_0 अस्वीकृत जब यह सत्य है(Type I त्रुटि) = $\propto$

H_0 स्वीकृत जब यह असत्य है (Type II त्रुटि) = β

H_0 स्वीकृत जब यह सत्य है (सही निर्णय)

H_0 अस्वीकृत जब यह असत्य है (गलत निर्णय)

स्तर का महत्व - टाइप I त्रुटि करने की अधिकतम संभावना , जिसे हमने एक परीक्षण में निर्दिष्ट किया है , को महत्व का स्तर कहा जाता है।

सामान्यतया सांख्यिकीय आंकडों के परीक्षण के लिए 50% महत्व या स्तर निर्धारित किया जाता है। इसका अर्थ है कि हम एक अवधारणा को स्वीकार करने में 95% तक विश्वस्त हो सकतें हैं या हम 5% गलत हो सकते है।

महत्वपूर्ण क्षेत्र - भिन्नता की सीमा में दो क्षेत्र हैं स्वीकृत क्षेत्र और मत्वपूर्ण क्षेत्र या अस्वीकृति क्षेत्र अगर नमूना आंकडा महत्वपूर्ण क्षेत्र में पडते हैं तो हमें शून्य परिकल्पना को अस्वीकार करना पडता है , क्योंकि इससे गलत निर्णय होता है। हम H_1 के लिए जाते हैं, यदि नमूना आंकडों का परिकलित माना अस्वीकृत क्षेत्र में होता है।

एक पुच्छ एवं दो पुच्छ परीक्षण - एक सामान्य चक्र के तहत अस्वीकृत (महत्वपूर्ण) क्षेत्र , जैसा कि पहले कहा गया है दो तरह से विभाजित किया जा सकता है।

(अ) वक्र के नीचे दोनो तरफ

(ब) वक्र के नीचे एक तरफ, या तो दायी पूंछ पर या बांए पूंछ पर

निर्णय या निष्कर्ष निकालना - आखिरकार हम शून्य परिकल्पना को स्वीकार या अस्वीकार करने के लिए निष्कर्ष पर आते हैं, यदि निर्णय परिकलित मान के आधार पर है चाहे वह स्वीकृति क्षेत्र या अस्वीकृत क्षेत्र है।

मानक त्रुटि - नमूना वितरण के मान विचलन को मानक त्रुटि कहते हैं उदाहरण के लिए $\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}$,................इत्यादि सभी नमूनों के माध्यों को समग्र से लिया गया है। इन सभी माध्यों के मानक विचलन माध्य की मानक त्रुटि है। इसके लिए सूत्र $\frac{\sigma}{\sqrt{n}}$ है।

उपयोगिता - (i) यह परिकल्पना के परीक्षण में उपयेागी साधन है। हम परिकल्पना को 5% स्तर के महत्व पर जांच सकते हैं, जिसका अर्थ है कि यदि प्रेक्षित मान एवं अपेक्षित मान के बीच में अंतर ± 1.96 S.E से अधिक है तो शून्य परिकल्पना स्वीकार नहीं की जाती है और किसी को वैकल्पिक परिकल्पना के लिए जाना पडता है। स्तर का महत्व 1% भी हो सकता है। सामान्यतया , परिकल्पना स्वीकार की जाती है यदि अंतर 3 S.E से कम है। 5% स्तर का महत्व प्रचलित है।

(ii) नमूने के विश्वसनीयता की जांच की जा सकती है।

(iii) मापदंडों (प्राचलों) का मान सीमा के साथ निर्धारित किया जा सकता है।

उदाहरण 3 -। शहर में , 32% मतदाताओं ने राजनैतिक पार्टी X के लिए मत दिया। शहर 13 में , 29% मतदाताओं ने X राजनैतिक पार्टी के लिए मत दिया

(i) शहर A में से 70 यादृच्छिक रूप से चुने हुए मतदाताओं में से यदि P_1 उन मतदाताओं का अनुपात है जिन्होंने X राजनैतिक दल के लिए मतदान किया, P_1 के लिए माध्य एवं मानक त्रुटि ज्ञात करें।

(ii) शहर B में से 60 यादृच्छिक रूप से चुने हुए मतदाताओं में से , यदि P_2 एन मतदाताओं का अनुपात है जिन्होने X राजनैतिक दल के लिए मतदान किया, P_2 के लिए माध्य एवं मानक त्रुटि ज्ञात करें।

(iii) $(P_1 - P_2)$ का माध्य एवं मानक त्रुटि ज्ञात करें।

हल:-

यहाँ, $P_1 = \frac{32}{100} = 0.32$ and $P_2 = \frac{29}{100} = 0.29$

$n_1=70$ and $n_2=60$

(i) mean=$E(P_1)=P_1=0.32$

$$S..E.(p_1) = \sqrt{\frac{P_1 Q_1}{n_1}} = \sqrt{\frac{0.32 \times 0.68}{70}} = 0.05575$$

$$(ii)\ Mean = E(p_2) = p_2. = 0.29$$

$$S..E.(p_2) = \sqrt{\frac{P_2Q_2}{n_2}} = \sqrt{\frac{0.29\times 0.71}{60}} = 0.05858$$

$$(iii)\ Mean = E(p_1 - p_2) = p_1 - p_2. = 0.32 - 0.29 = 0.03$$

$$S..E.(p_1 - p_2) = \sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$$

$$= \sqrt{\frac{0.323\times 0.68}{70} + \frac{0.29\times 0.71}{60}}$$

$$= 0.08087$$

उदाहरण 4: एक समाज में महिलाओं का अनुपात 0.48 है। जिसमें से 64 सदस्यों को यादृच्छिक तरीके चयनित किया गया, यदि P_1 महिलाओं का अनुपात है। दूसरे चयन में जिसने 86 सदस्य चयनित किये गये हो उसमें P_2 महिलाओं का अनुपात है।

ज्ञात कीजिए -

(i) P_1 की मानक त्रुटि

(ii) P_2 की मानक त्रुटि

(iii) $(P_1 – P_2)$ के अंतर की मानक त्रुटि

हल -

यहाँ $P_1=P_2=0.48=P$ (माना)

$n_1=64$ और $n_2=86$

$$(i)\ S..E.(p_1) = \sqrt{\frac{PQ}{n_1}} = \sqrt{\frac{0.42\times 0.52}{64}} = 0.06245$$

$$(i)\ S..E.(p_1) = \sqrt{\frac{PQ}{n_1}} = \sqrt{\frac{0.48\times 0.52}{86}} = 0.05387$$

$$S..E.(p_1 - p_2) = \sqrt{PQ\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

$$= \sqrt{0.48\times 0.52\left[\frac{1}{64} + \frac{1}{86}\right]}$$

$$= \sqrt{0.48\times 0.52\left[\frac{86+64}{64\times 86}\right]}$$

$$= 0.08248$$

उदाहरण 5 - विभिन्न बीजों के लिए आवश्यक औसत अंकुरण समय का अनुमान लगाना जरूरी हैं दस यादृच्छिक रूप से चयनित बीज दिखाये जा रहे हैं और अंकुरण का समय नीचे दिया जा रहा है।

समय (दिन) 28,32,27,38,30,31,30,30,27 और 33 नमूना माध्य को समग्र माध्य के अनुमानक के रूप में लेते हुए और सत अंकुरण समय का औसत माध्य बताऐं।

हल - नमूना माध्य $\bar{x} = \frac{\sum x}{n}$

$$\frac{28 + 32 + 27 + 38 + 30 + 31 + 30 + 30 + 27 + 33}{10}$$

$=\frac{306}{10}$ = 31 दिन (लगभग)

इस प्रकार औसत अंकुरण समय का अनुमान 31 दिन है इसलिए हम निष्कर्ष निकालते हैं कि औसतन 31 दिनों में बीज की विविधता अंकुरित होती है।

उदाहरण 6 - पत्नियों की अपेक्षा औसतन , पति कितने लम्बे हैं जानने के लिए 8 जोडों को यादृच्छिक तरीके से लिया जाता है। प्रत्येक मामले में पति और पत्नियों की ऊँचाईयों को मापा जाता है और उनका अंतर निम्नानुसार दर्ज होता है।

जोड़ा	1	2	3	4	5	6	7	8
अंतर(सेमी में)	5	7	12	23	6	-4	10	9

नमूना माध्य x̄ का अनुमानक के रूप प्रयोग करते हुए समग्र माध्य μ को अनुमानित करें।

हल -

अनुमानित $\mu = \bar{x} = \frac{\sum x}{n}$

$= \frac{5+7+12+23+6+4+10+9}{8} = \frac{68}{8} = 8.5$ **सेमी**

इस प्रकार हम निष्कर्ष निकालते हैं कि पति अपनी पत्नियों से औसतन 8.5 सेन्टीमीटर लम्बे हैं।

उदाहरण 7 – 5 एक बेकरी का मालिक शहर में केक की औसत दैनिक मांग का अनुमान लगाना चाहता है। 156 दिनों में, सर्वेक्षण ने लिम्नलिखित मांग का खुलासा किया।

मांग(केक)	120-130	130-140	140-150	150-160	160-170
दिन	13	43	70	27	3

औसत मांग का अनुमान लगाऐं।

हल -

मांग	दिन(f)	मध्यमान (x)	f.x
120-130	13	125	1625
130-140	43	135	5805
140-150	70	145	10150
150-160	27	155	4185
160-170	3	165	495
Total	156	-	22260

x̄ का अनुमान $= \frac{\sum fx}{N} = \frac{22260}{156} = 143$(लगभग)

इस प्रकार औसत मांग को अनुमान **143 केक/प्रतिदिन** है।

निर्णय लेने के लिए परिकल्पना परीक्षण में उचित वितरण का प्रयोग-

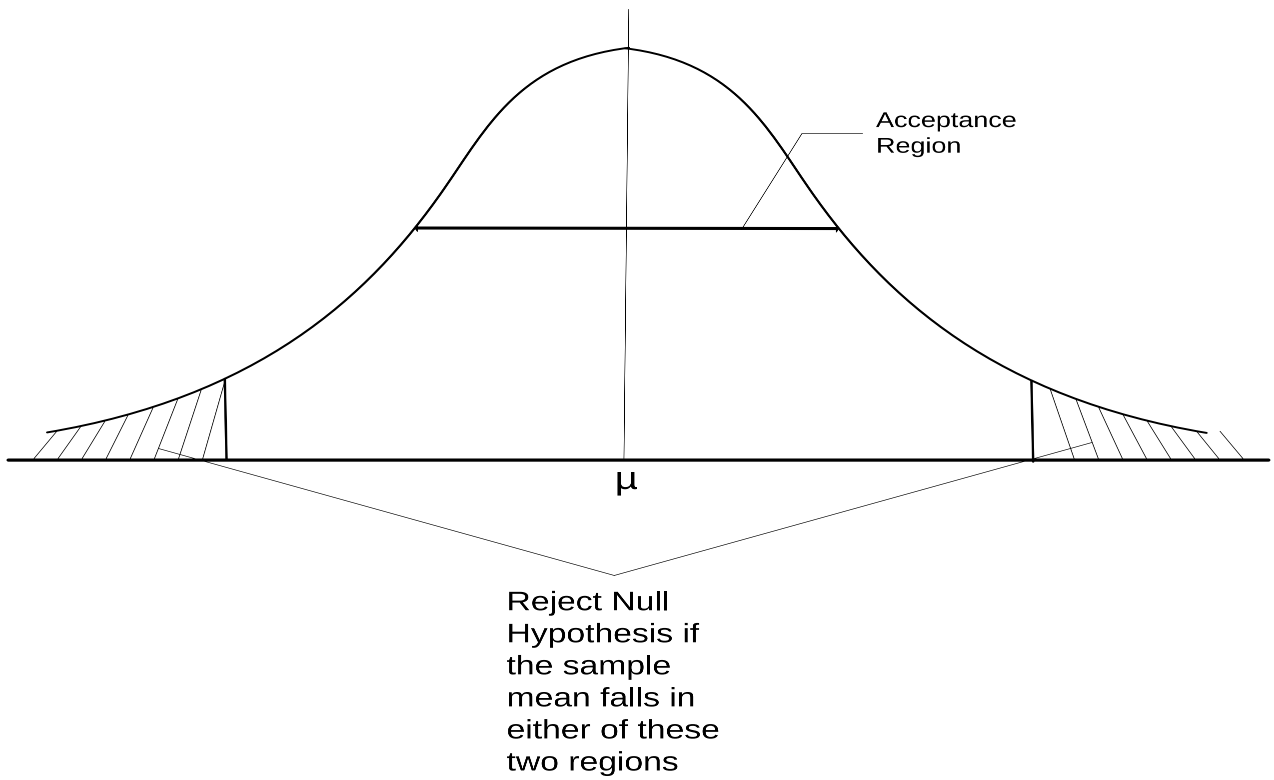
किस स्तर के महत्व का उपयोग करना है , तय करने के लिए , परिकल्पना परीक्षण में हमारा अगला कार्य उचित संभाव्यता वितरण को निर्धारित करना है। हमारे पास सामान्य वितरण और t वितरण के बीच एक विकल्प है। उचित वितरण को चुनने के लिए नियम , उन इकाईयों में समान है जो हमें आकलन के आधार पर मिला था। नीचे दी गई सारणी का सारांश है कि मध्यों को परीक्षण करने में सामान्य और t वितरण का उपयोग किया जाता है। बाद में इस अध्याय, में हम अनुपातों के बारे में परीक्षण परिकल्पना के लिए उपयुक्त वितरण की जांच करेंगे। एक नियम को याद रखे जब एक माध्य के मान पर विचार किया जाए। अनुमान में , जब भी समग्र आकार परिमित है, परिमित जनसंख्या गुणक का उपयोग करें, नमूनाकरण प्रविस्थापन के बिना किया जाता है और नमूनाकरण प्रविस्थापन के बिना किया जाता है और नमूना समग्र से 5% अधिक होता है।

माध्यों के बारे में परीक्षण परिकल्पना में सामान्य औरt वितरण का उपयोग करने के लिए शर्तें		
	जब समग्र मानक विचलन ज्ञात हो	जब समग्र मानक विचलन अज्ञात हो
नमूना आकार द 30 से ज्यादा हो	सामान्य वितरण, Z तालिका	सामान्य वितरण, Z तालिका
नमूना आकार द 30 है या उससे कम और समग्र लगभग सामान्य है	सामान्य वितरण, Z तालिका	t वितरण, t तालिका

दो पुच्छीय परीक्षण और एक पुच्छीय परीक्षण

दो पुच्छीय परीक्षण -

दो पुच्छीय परीक्षण में , शून्य परिकल्पना को खारिज कर दिया जाता है यदि नमूना माध्य अनुमानित समग्र माध्य से अर्थपूर्ण रूप में ज्यादा या कम होता है इस प्रकार दो पुच्छीय परीक्षण में , दो अस्वीकृति क्षेत्र हो ते हैं इसे नीचे दिए गए चित्र में दिखाया गया है -



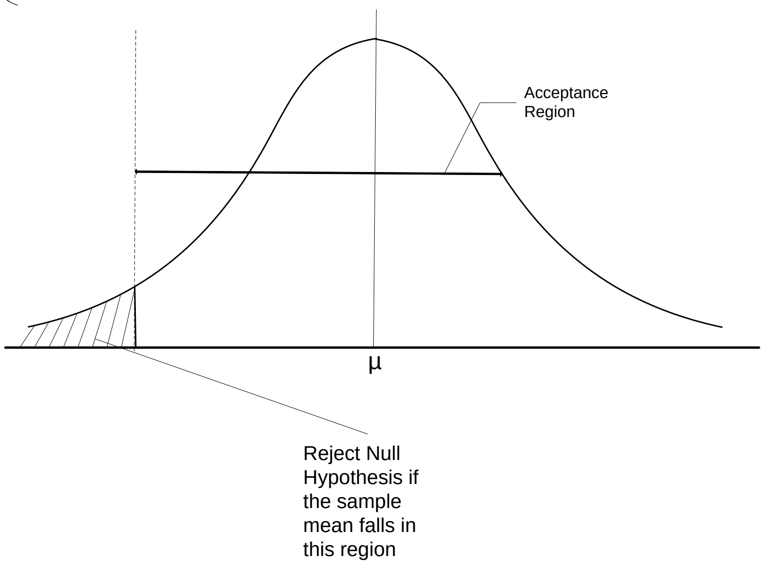
जब शून्य परिकल्पना $\mu = \mu H_o$ है (जहाँ μH_o कोई विनिर्दिष्ट मान है।) और वैकल्पिक परिकल्पना $\mu \neq \mu H_o$ है द्विपुच्छीय परीक्षण उपयुक्त होता है।

मान लें कि प्रकाश बल्ब का निर्माता $\mu = \mu H_o = 1000$ घन्टे के औसत जीवन के साथ बल्ब बनाना चाहता है। यदि जीवनकाल छोटा होता है , वह प्रतिर्स्पधा में अपने ग्राहकों नुकसान। यदि जीवनकाल अधिक हेाता है , तो उसमें उच्च उत्पादन कीमत अधिक लगेगी क्योंकि तन्तु अतिशय प्रगाए कीमत होगी।

क्या उसकी उत्पादन प्रक्रिया सही तरीके से काम कर रही है को देखने के लिए परिकल्पना के परीक्षण $\mu_o \mu = 1000$ के लिए, वह उत्पादन का एक नमूना लेता है। क्योंकि वह किसी भी स्थिति में 1000 घन्टे से आगे नहीं हटना चाहता है , उचित वैकल्पिक परिकल्पना $\mu_o : \mu = 1000$ है और वह द्विपुच्छीय परीक्षण प्रयोग करता है। इसका अर्थ है कि, वह शून्य परिकल्पना को अस्वीकार करता है, यदि नमूने में बल्बों का औसत जीवन या तो 1000 घन्टे से बहुत अधिक हो या 1000 घन्टे से बहुत कम हो।

यद्यपि इन परिस्थितियों में द्विपुच्छीय परीक्षण उपयुक्त नहीं होता है, और हमें एक पुच्छीय परीक्षण का प्रयोग करना चाहिए। थोक व्यापारी मामले में विचार करते हैं कि वह प्रकाश बल्ब के निर्माता के उपरोक्त वर्णन से प्रकाश बल्ब खरीदता है । थोक व्यापारी प्रकाश बल्ब का बढा ढेर खरीदता है बल्ब के ढेरों को तब तक स्वी कार नहीं करना चा हता है जब तक उनका औसत जीवन कम से कम 1000 घन्टा या न्यूनतम 1000 घन्टा हो। प्रत्येक शिपमेंट के आने पर , थोक व्यापारी यह तय करने के लिए एक नमूने का परीक्षण करता है कि क्या शिपमेंट को स्वी कार करना चाहिए। कम्पनी शिपमेंट को तभी अस्वीकार करेगी यदि कम्पनी महसूस करती है कि इनका औसत जीवन 1000 घन्टे से कम हो। यदि कम्पनी महसूस करती है कि बल्ब अपेक्षित की तुलना (औसत जीवन 1000 घन्टे से ज्यादा) में बेहतर है, यह निश्चित रूप से शिपमेंट को अस्वीकार नहीं करेगो अधिक लम्बे जीवन में कोई अतिरिक्त कीमत नहीं हेाती है। इसलिए थोक व्यापारी की परिकल्पनाऐं $\mu_o : \mu = 1000$ घन्टे

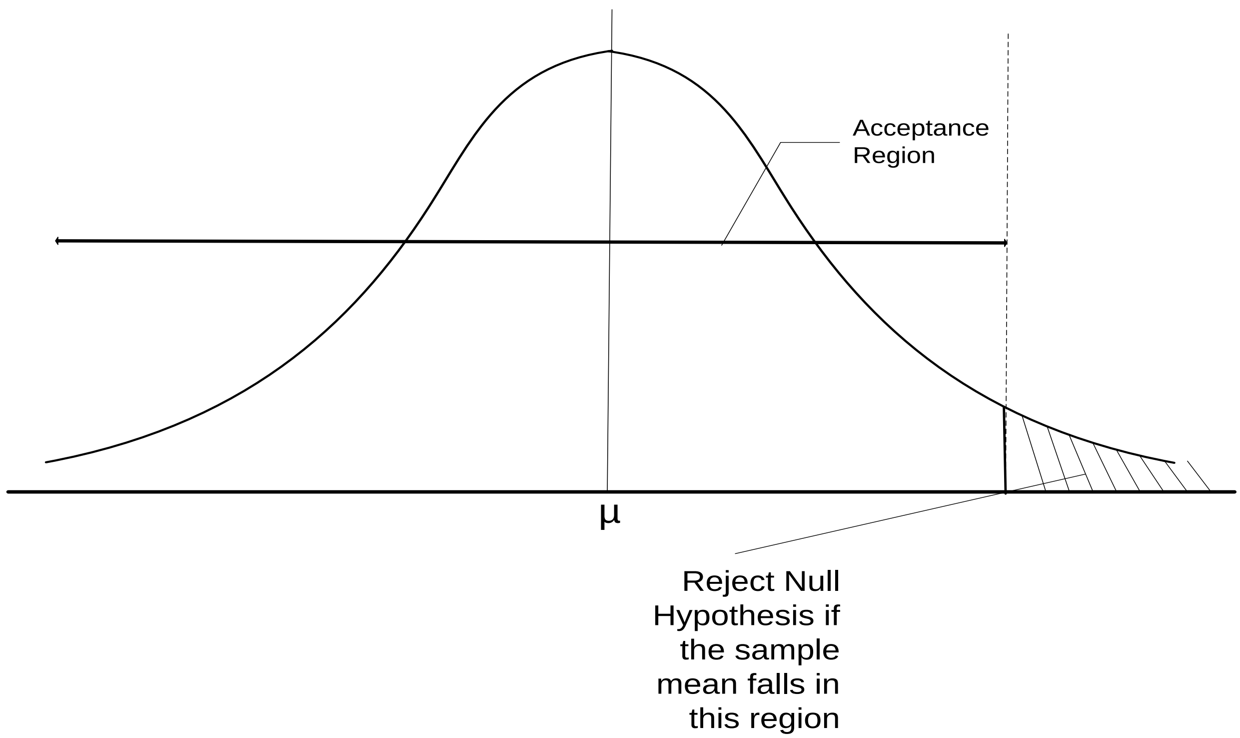
और $\mu_0:\mu<1000$ घन्टे है। यह H_0 को तभी अस्वीकार करती है यदि नमूना बल्बों का औसत जीवन 1000 घन्टों से उल्लेखनीय ढंग से कम होता है। इस परिस्थिति को निम्न चित्र द्वारा स्पष्ट किया जाता है। इस चित्र से , हम देख सकते हैं कि इस परीक्षण को क्यों बायाँ पुच्छीय परीक्षण (या एक निचला पुच्छीय परीक्षण) कहा जाता है।



यदि नमूना माध्य इस क्षेत्र में आता है तो शून्य परिकल्पना अस्वीकार करते हैं।

सामान्य रूप में, यदि परिकल्पनाऐं $\mu_0:\mu=\mu H_0$ और $H_1:\mu<\mu H_0$ हो तो बायें पुच्छीय (निचला पुच्छीय) परीक्षण प्रयोग होता है। इस तरह की परिस्थिति में , यह नमूना साक्ष्य नमूना माध्य के साथ उल्लेखनीय ढंग से अनुमानित समग्र माध्य से कम है जो शून्य परिकल्पना को अस्वीकार करते हुए वैकल्पिक परिकल्पना के पक्ष की और जाता है। दूसरे शब्दों में , अस्वीकृत क्षेत्र नमूना माध्य के वितरण में निचला पुच्छीय (बायीं पुच्छीय) होता है, और इसलिए इसे बायी पुच्दीय परीक्षण कहते हैं।

एक बायी पुच्छीय परीक्षण दो प्रकार के एक पुच्छीय परी क्षणों में से एक हैं सम्भवतः अब तक अनुमान लगाया जाता है , परीक्षण दायीं पुच्छीय परीक्षण (या एक ऊपरी पुच्छीय परीक्षण) है। एक ऊपरी पुच्छीय परीक्षण प्रयोग किया जाता है जब परिकल्पनाऐं $H_0:\mu = \mu H_0$ और $H_1:\mu> \mu H_0$ हों। केवल नमूना माध्य के मानों जो कि अनुमानित समग्र माध्य से अधिक हैं वैकल्पिक परिकल्पना के पक्ष में शून्य परिकल्पना को अस्वीकार करने का कारण होगा। इसे ऊपरी पुच्छीय परीक्षण कहा जाता है क्योंकि अस्वीकृत क्षेत्र नमूना माध्य के वितरण के ऊपरी पुच्छीय है।



यह आपको फिर से याद दिलाना है कि, परिकल्पना परीक्षण के प्रत्येक उदाहरण में, नमूना सूचना के आधार पर जब हम शून्य परिकल्पना को स्वीकर करते हैं, तो हम वास्तव में कहते हैं कि इसमें कोई सांख्यिकी साक्ष्य नहीं है जिससे इसे अस्वीकार कर दिया जाये। हम यह नहीं कह रहे होते हैं कि शून्य परिकल्पना सत्य है। शून्य परिकल्पना को सिद्ध करने के लिए एक ही तरीका समग्र प्राचल को जानना है और जो नमूनाकरण के साथ सम्भव नहीं होता है। इस प्रकार , हम शून्य परिकल्पना स्वीकार करते हैं और व्यवहार करते हैं कि यह सच है क्योंकि हमें इसे अस्वीकार करने के लिए कोई साक्ष्य नहीं मिल सकते हैं।

परिकल्पना परीक्षण में अवधारणाऐं

दो पुच्छीय , ऊपरी पुच्छीय या निचला पुच्छीय परीक्षण के प्रयोग के निर्धारण के लिए नमूना परिणामों का प्रयोग न करें।

किसी भी तरह के आंकडों के संग्रह से पहले , निर्णयकर्ता द्वारा यह मानना है कि उसे क्या पता लगाना है तब परीक्षण के रूप को निर्धारित किया जाता है।

1.6 सारांश (Summary)

परिकल्पना परीक्षण अवधारणा के साथ आरम्भ हेती है, जिसे एक परिकल्पना कहा जाता है, जो समग्र प्राचल के बारे में होती है। तब हम नमूना आंकडा एकत्र करते हैं , जिससे नमूना सांख्यिकी की रचना होती है , और इस जानकारी को निर्धारित करने के लिए कितनी संभावना है कि हमारा अनुमानित समग्र प्राचल सही है। हम मान लेते हैं कि समग्र माध्य के लिए निश्चित मान है। हमारी अवधारणा की वैद्यता के परीक्षण के लिए हम नमूना आंकडा एकत्र करते हैं और परिकल्पित मान एवं नमूना माध्य का वास्तविक मान के मध्य अन्तर निर्धारित करते हैं। तब हम निर्णय लेते है , कि क्या अंतर उल्लेखनीय ढंग से है। छोटा सा अंतर , माध्य के लिए हमारे परिकल्पित मान के सही होने की ज्यादा संभावना हेती है। ज्यादा अंतर होने पर संभावना कम की ओर अग्रसर होती है। दुर्भाग्यवश, परिकल्पित समग्र प्राचल और वास्तविक आंकडे के बीच अंतर प्रायः न तो बहुत बडे हैं जिन्हें हम स्वतः अपनी परिकल्पना में अस्वीकार करते हैं न तो इतने छोटे हेते हैं कि हम उसे शीध्रता से स्वीकार करते हैं। इसलिए परिकल्पना परीक्षण में , सबसे ज्यादा उल्लेखनीय वास्तविक जीवन निर्णयों के रूप

में,स्पष्ट रूप से समाधान अपवाद हैं, न कि नियम। पूर्वानुमान में, जब कभी समग्र का आकार निश्चित हैं, निश्चित समग्र गुणक का प्रयोग करते हैं , नमूनाकरण बिना प्रतिस्थापन के साथ किया जाता है , और नमूना समग्र का 5 प्रतिशत से अधिक हेता है।

1.7 शब्दावली (Glossary)

परिकल्पना : एक शर्त जिसमें से कुछ इस प्रकार है, यह एक भ्रान्ति है।
सरल परिकल्पना : एक परिकल्पना जो सटीक वितरण निर्दिश्ट करता है।

1.8 बोध प्रश्न (Comprehension Question)

1.निम्नलिखित परिस्थितियों में , निर्दिश्ट करें कि कौन सा प्रायिकता वितरण परिकल्पना परीक्षण में प्रयोग होता है -

(अ) $H_0:\mu = 27$, $H1:\mu \neq 27$, $\bar{x} = 33$, नमूना $\sigma = 4$, $n = 25$

(ब) $H_0:\mu = 98.6$, $H1:\mu > 98.6$, $\bar{x} = 99.1$, $\sigma = 1.5$, $n = 50$

(स) $H_0:\mu = 3.5$, $H1:\mu < 3.5$, $\bar{x} = 2.8$, नमूना $\sigma = 0.6$, $n = 18$

(द) $H_0:\mu = 382$, $H1:\mu \neq 382$, $\bar{x} = 363$, $\sigma = 68$, $n = 12$

(घ) $H_0:\mu = 57$, $H1:\mu > 57$, $\bar{x} = 65$, नमूना $\sigma = 12$, $n = 42$

2.भारत में सिनेमाघरों के मालिकों को पता है कि एक हिट फिल्म चलाने के लिए प्रत्येक शहर में औसतन 84 दिन, 10 दिनों के मानक विचलन के साथ फिल्म प्रदर्शित होती थी। एक विशे ष फिल्म वितरक समग्र के साथ अपने क्षेत्र में फिल्म की लोकप्रियता की तुलना में रूचि रखता है। उसने यादृच्छिक रूप से 75 सिनेमाघरों को यादृच्छिक रूप से क्षेत्र में चुना है और पाया कि फिल्म 81.5 दिन तक चली।

I. परीक्षण के लिए उपयुक्त परिकल्पनाएँ लिखें की क्या , समग्र और वितरक के क्षेत्र में सिनेमाघरों के मध्य उल्लेखनीय अन्तर हैं ?
II. 1% महत्व के स्तर पर, इन परिकल्पनाओं का परीक्षण करें ।

1.9 बोध प्रश्नों के उत्तर (Answers to comprehension questions)

इसका अर्थ है अस्वीकृत क्षेत्र दोनों पुच्छों के अन्तर्गत 0.01 है और स्वीकृत क्षेत्र 0.99 है। इसलिए स्वीकृत क्षेत्र का आधा भाग $\frac{0.99}{2} = 0.4950$ है और Z का मान 2.58 इसलिए , स्वीकृत क्षेत्र की सीमाऐं $Z = \pm 2.58$ or

$\bar{x} = \mu H_0 \pm \frac{z\sigma}{\sqrt{n}} = 84 \pm 2.58 \times \frac{1075}{\sqrt{}}$

निचली सीमा 81.02 और ऊपरी सीमा 86.98 के रूप में हैं।

क्योंकि अवलोकित मान स्वीकृत क्षेत्र में है, हम शून्य परिकल्पना H_0 को अस्वीकार नहीं करते हैं। फिल्म के चलने की अवधि दूसरे सिनेमाघरों के समान है। या दूसरे रूप में -

अवलोकित Z मान $\bar{x} - \frac{H_0 SE}{}$ है जहाँ $SE = \frac{\sigma}{\sqrt{n}} = \frac{81.4-84}{(1.155)}$

= - 2.17 स्वीकृत Z क्षेत्र $\pm Z = \pm 2.58$ है।

1.10 स्वपरख प्रश्न (Self-Test Questions)

1.(अ) 24 df के साथ t (स्वतन्त्रता की श्रेणियाँ)
(ब) Z या सामान्य वितरण : दायीं पुच्छीय परीक्षण

(स)17 Df के साथ t

(द) Z या सामान्य वितरण द्वि पुच्छीय परीक्षण

(घ) इस परिस्थिति में वास्तविक 41 d.f के साथ t चूकि 41 df यहाँ नहीं है हम सामान्य वितरण परीक्षण तालिका प्रयोग करते हैं।

2.निम्नलिखित आंकडे दिये गए है :

$\sigma = 10$ दिन, $n = 75$ सिनेमाघर $\bar{x} = 81.5$

$H_0: \mu = 84$ दिन $H1: \mu \neq 84$ दिन $\alpha = 0.01$

3.माइक्रोसाफट ने पिछले साथ अनुमान लगाया था कि 35 प्रतिशत सम्भावित साफटवेयर ग्राहक , नऐ OS विंडोज विस्टा की खरीद की प्रतीक्षा की योजना बना रहे थे , जब तक एक नवीनीकरण जारी न किया गया हो। जनता को आश्वस्त करने के लिए विज्ञापन अभियान के पश्चात माइक्रोसाफट ने 3000 ग्राहकों का सर्वेक्षण किया और पाया कि 950 ग्राहक अभी भी संदेहपूर्ण है। 5 प्रतिशत महत्व के स्तर पर क्या कम्पनी यह निश्कर्श निकाल सकती है कि संदेहपूर्ण लोगों का अनुपात कम हुआ था। (शून्य परिकल्पना अस्वीकार होती है। Z वितरण का प्रयोग करें।)

1.11 सन्दर्भ पुस्तकें (Reference books)

1. मूल सांख्यिकी – गौण, गुप्ता और दासगुप्ता वर्ल्ड प्रेस लिमिटेड-कलकत्ता ।
2. व्यावसायिक आंकडों के बुनियादी सिद्धान्त, संचेती और कपूर ।
3. प्रबंधन में मात्रात्मक विधियाँ - श्रीवास्तव, शोनायँ और गुप्ता ।
4. व्यावसायिक सांख्यिकीय - गुप्ता और गुप्ता ।

इकाई 2 गुणों का साथर्कता परीक्षण
(Validity Testing of Properties)

2.1 प्रस्तावना (Introduction)

2.2 उद्देश्य (Objectives)

2.3 परिकल्पना का परीक्षण (Hypothesis testing)

2.4 मानक त्रुटि (Standard error)

2.5 विशेषताओं के लिए महत्व का परीक्षण (Significance tests for characteristics)

2.6 महत्व सृजन फलन (Importance Generation Function)

2.7 सारांश (Summary)

2.8 शब्दावली (Glossary)

2.9 बोध प्रश्न (Comprehension Question)

2.10 बोध प्रश्नों के उत्तर (Answers to comprehension questions)

2.11 स्वपरख प्रश्न (Self-Test Question)

2.12 सन्दर्भ पुस्तकें (Reference books)

2.1 प्रस्तावना (Introduction)

आंकडों के औद्योगिक अनुप्रयोग प्रायः समग्र और समग्र प्राचलों के बारें में निर्णय लेने से संबंधित हैं। उदाहरण के लिए दो प्रक्रियाओं के लिए क्या बेहतर है के बारे मे निर्णय था किसी विशेष मशीन के उत्पादन को बंद करना है क्योंकि यह एक दोषपूर्ण घटकों की आर्थिक रूप से अस्वीकार्य संख्या पैदा कर रहा है । प्रायः समग्र के निर्धारित माध्य या मानक विचलन पर आधारित होता है , की गणना समग्र में से नमूना आंकडों का प्रयोग किया जाता है। इन निर्णयों तक पहुचनें में , कुछ मान्यताओं का निर्माण किया जाता है , जो सत्य या असत्य भी हो सकता है। बनाई गई धारणाऐं सांख्यिकीय अनुमान या सिर्फ परिकल्पना कहलाती है और आमतौर पर समग्र की संभावना वितरण के बारे में बयान से संबंधित है।

2.2 उद्देश्य (Objectives)

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

- ✓ परिकल्पना के परीक्षण के विभिन्न तरीकों को समझ सकें।
- ✓ विशेषताओं के महत्व के परीक्षण का वर्णन कर सकें।
- ✓ महत्व सृजन फलन को समझ सकें।

2.3 परिकल्पना का परीक्षण (Hypothesis testing)

एक नमूना जांच परिणाम उत्पन्न करता है और इन परिणामों के साथ निर्णय समग्र पर बनाये जाते हैं। लेकिन ऐसे फैसलों में गलत निर्णय लेने के कारण अनिश्चितता का एक तत्व शामिल है , परिकल्पना एक ऐसी धारणा है जो समग्र प्राचल के बारे में सत्य या असत्य भी हो सकती है।

उदाहरण के लिए एक सिक्के को 300 बार उछालने पर कोई 190 चित और 110 पट प्राप्त कर सकता है। इस उदाहरण पर हम यह जांचने में रूचि रखते हैं कि क्या सिक्का निष्पक्ष है या नहीं इसलिए हम इसका मूल्यांकन करने के लिए एक परीक्षण आयोजित कर सकते हैं कि अन्तर नमूनाकरण के कारण है। एक महत्व के परीक्षण करने की प्रक्रिया निम्नानुसार हैः

अवधारणा का निर्माण (Formation of Hypothesis) - हमारी अवधारणा को सत्यापित करने के लिए जो नमूना अध्ययन पर आधारित है , हम नमूना मान और समग्र मान के बीच का अंतर जानने के लिए हम आंकडे एकत्र करत हैं। यदि कोई अंतर नहीं है या यदि अंतर बहुत छोटा है , तो हमारा अनुमानित मान सही है , आमतौर पर दो अनुमानों का निर्माण किया जाना चाहिए और यदि एक परिकल्पना सही है , तो दूसरे को अस्वीकार करेंगें।

(अ) शून्य परिकल्पना (Null Hypothesis) - यह अंतर के महत्व के परीक्षण करने के लिए बहुत उपयोगी उपकरण है। समग्र से संबंधित कोई भी परिकल्पना को एक सांख्यिकीय परिकल्पना कहा जाता है। सांख्यिकीय परीक्षण की प्रक्रिया में समग्र से प्राप्त नमूने के आधार पर अवधारणा को अस्वीकार या स्वीकार किया जाता है। सांख्यिकविद् अवलोकन के माध्यम से परिकल्पना की जांच करते हैं और एक संभावित बयान देते हैं सरल परिकल्पना से पता चलता है कि नमूने का मान और अध्ययन के तहत समग्र के मान में कोई अंतर नहीं दिखता है। हम जिस परिकल्पना को ग्रहण करते हैं उसे शून्य परिकल्पना कहा जाता है । इसका अर्थ है कि नमूने के माध्य और समग्र के माध्य के बीच वास्तविक अंतर शून्य है , न्यूनतम पाया गया अन्तर महत्वहीन है। शून्य परिकल्पनपा की अस्वीकृति का अर्थ है कि नमूना माध्य और समग्र माध्य के बीच वास्तविक अंतर शून्य है। शून्य परिकल्पना की अस्वीकृति प्रदर्शित करती है कि निर्णय सही है।

उदाहरण के लिए -

(1) एक विश्वविद्यालय के छात्रों की औसत ऊँचाई 155 सेमी है।

(2) एक फर्म की औसत दैनिक बिक्री 1500 रूपये हैं।

(3) किसी विशेष गांव की औसत आय 100 रूपये है।

इन सभी बयानों को नमूना परीक्षणों के आधार पर सत्यापित करना होगा। आमतौर पर एक परिकल्पना में कहा जाता है कि नमूना माध्य और समग्र माध्य के बीच कोई अंतर नहीं है। एक सांख्यिकीय अनुमान एक शून्य परिकल्पना है यदि इसे स्वीकार किया जाता है। शून्य परिकल्पना को μ_0 द्वारा प्रदर्शित किया जाता है।

(ब) वैकल्पिक परिकल्पना (Alternative Hypothesis) -

μ_0 की अस्वीकृति वैकल्पिक परिकल्पना की स्वीकृति की ओर जाता है, जिसे H_1 द्वारा दर्शाया गया है। उदाहरण के लिए , $H_0 = \mu = 155$ (शून्य परिकल्पना)

$\mu_1 = \mu \neq 155$ अर्थात $\mu > 155$ या $\mu < 155$ (वैकल्पिक परिकल्पना)

जब दो अवधारणाओं को स्थापित किया जाता है , तो शून्य परिकल्पना की स्वीकृति या अस्वीकृति एक नमूना अध्ययन पर आधारित होती है। इस प्रकार यह दो गलत निष्कर्षों के ओर जाता है अर्थात-

(1) μ_0 अस्वीकृत, जब H_0 सत्य है।

(2) H_0 स्वीकृत, जब H_0 असत्य है। इसे निमनलिखित तालिका में व्यक्त किया जा सकता है।

	नमूने में से निर्णय	
	H_0 स्वीकृत	H_0 अस्वीकृत
H_0 सत्य	सही	गलत (Type I त्रुटि)
H_0 गलत (H_1 सही)	गलत (Type II त्रुटि)	सही

फिर से लिखते हुए -

H_0 अस्वीकार जब यह सत्य है (Type I त्रुटि) $=\alpha$

H_0 स्वीकार जब यह गलत है (Type II त्रुटि)$=\beta$

H_0 स्वीकार जब यह सत्य है (सही निर्णय)

H_0 अस्वीकार जब यह गलत है (सही निर्णय)

महत्व का स्तर

I प्रकार की त्रुटि करने की अधिकतम संभावना, जिसे हमने एक परीक्षण में निर्दिष्ट किया है, को महत्व का स्तर कहा जाता है। सामान्य तथा सांख्यिकीय परीक्षणों में 5 प्रतिशत महत्व का स्तर तय किया जाता है। इसका अर्थ है कि हम एक अवधारणा को स्वीकार करने पर 95 प्रतिशत विश्वस्त हो सकते हैं या हम 5 प्रतिशत गलत हो सकते हैं।

महत्वपूर्ण क्षेत्र (Significance Area) - विविधता की सीमा में दो क्षेत्र हैं - स्वीकृति क्षेत्र और महत्वपूर्ण क्षेत्र या अस्वीकृति क्षेत्र। यदि नमूना आंकडे महत्वपूर्ण क्षेत्र में आते हैं तो हमें परिकल्पना को अस्वीकार करना पडता

है, क्योंकि इससे गलत निर्णय होता है। हम H_1 के लिए जाते हैं , यदि सरल आंकडों का गणित मान अस्वीकृत खेत्र में होता है।

एक पुच्छीय और द्विपुच्छीय परीक्षण(One tailed and two tailed test) -

एक सामान्य वक्र के अन्तर्गत महत्वपूर्ण क्षेत्र , जैसा कि पहले बताया गया है ,दो तरह से विभाजित किया जा सकता है।

(अ) वक्र के नीचे दोनों तरफ

(ब) वक्र के नीचे एक तरफ और दोनों या तो दाई पूंछ पर या बाएं पूंछ पर है।

निर्णय या निष्कर्ष पर पहुंचना (come to a decision or Cconclusion) -

अनंत में हम इस निष्कर्ष मे आते हैं कि या तो शून्य परिकल्पना को स्वीकार किया जाता है या अस्वीकार किया जाता है। यह निर्णय गणना मान के आधार पर है चाहे वह स्वीकृति क्षेत्र में है या अस्वीकृत क्षेत्र में है।

2.4 मानक त्रुटि (Standard error)

नमूनाकरण वितरण के मानक विचलन को मानक त्रुटि कहा जाता है। उदाहरण के लिए इत्यादि समग्र से लिए हुए सभी नमूनों के माध्य है। इन सभी माध्यों का मानक विचलन माध्य की मानक त्रुटि हेाती है। इसके लिए सूत्र $\frac{\sigma}{\sqrt{n}}$. है।

उपयोगिता -

1. यह परिकल्पना के परीक्षण में एक उपयोगी उपकरण है। हम 5 प्रतिशत महत्व के स्तर पर परीक्षण की जांच कर सकते हैं , जिसका अर्थ है , यदि अवकलित और अपेक्षित माध्यों के मध्य अंतर 1.96 S.E ज्यादा है तो परिकलपना स्वीकृत नहीं होती है और उसे वैकल्पिक परिकल्पना की और जाना पडता है। महत्व का स्तर 1 प्रतिशत हो सकता है। सामान्यतः परिकल्पना स्वीकृत होती है यदि अन्तर 3 S.E से कम है, 5 प्रतिशत स्तर लोकप्रिय है।
2. एक नमूने की विश्वसनीयता ज्ञात हो सकती है।
3. प्राचलों के मान सीमाओं के साथ निर्धारित किये जा सकते हैं।

अब हम विभिन्न स्थितियों पर लागू होने वाले महत्व के विभिन्न परीक्षणों पर चर्चा करते हैं । वो है-

1. विशेषताओं के लिए महत्व का परीक्षण
2. चरों के लिए महत्व का परीक्षण

2.5 विशेषताओं के लिए महत्व का परीक्षण (Significance tests for characteristics)

विशेषताओं का नमूनाकरण एक समग्र से नमूनों के चित्रण के रूप में माना जा सकता है जिनके सदस्यों मे एक विशिष्ट विशेषता की उपस्थिति या अनुपस्थिति होती है। उदाहरण के लिए , अंधे (विशेषता) के अध्ययन में, एक नमूना तैयार किया जा सकता है और इसके सदस्यों को अंधे हैं या नही के रूप में वर्गीकृत किया जाता है। विशेषता की उपस्थिति का p द्वारा प्रतिनिधित्व किया जा सकता है और विशेषता की अनुपस्थिति को q द्वारा प्रतिनिधित्व किया जा सकता है। इस प्रकार, 1000 लोगों में, 25 अंधे हैं और शेष अंधे नहीं हैं। दूसरे शब्दों मे $p = \frac{25}{1000}$ या 0.025 और $q = 0.975$ विभिन्न प्रकार के महतव के परीक्षण का अध्ययन निम्नलिखित प्रमुखों के अन्तर्गत किया जा सकता है।

(अ) सफलता की संख्या का परीक्षण(test success number)

यह द्विपद वितरण का अनुसरण करता है। सूत्रः सफलता की संख्या का

$$S.E = \sqrt{npq}$$

n= नमूना आकार

p=प्रत्येक परीक्ष्ज्ञण में सफलता कीसंभावना

q= (1-p) अर्थात विफलता की संभावना

उदाहरण -1 1,00,000 टेनिस के खेप से 400 गेंदों को यादृच्छिक चयनित गेंदों किया गये और जांच की गई। यह पाया गया कि इनमें से 20 दोषपूर्ण थे। कितने दोषपूर्ण गेंदों को आप 95 प्रतिशत विश्वसनीयता के स्तर पर सम्पूर्ण खेप में उचित रूप से उम्मीद कर सकते हैं।

हल: यहाँ

$$p = \frac{20}{400} = 0.05$$

q=0.95

$$\bar{x} = np = 100000 \times 0.05 = 5000$$

$$S.E = \sqrt{npq} = \sqrt{100000 \times .05 \times .95}$$

$$= \sqrt{4750} = 68.9$$

95% विश्वसनीयता सीमाऐं हैं

$5,000 \pm 1.96 \times 68.9 = 5000 \pm 135.044$ (या)

5135 और 4.865

उदाहरण -2 राजस्थान के एक गांव के 500 लोगों के नमूने में, 280 लोग चावल खाने वाले और बाकी गेहूँ खाने वाले पाए गये, क्या हम यह मान सकते हैं कि दोनों खाद्य पदार्थ समान लोकप्रिय है।

हल - हम यह परिकलपना करते हैं कि खाद्य पदार्थ समान रूप से लोकप्रिय है।

तब, गेहूँ खाने वालों और चावल खाने वालों की अपेक्षित आवृत्तिया 250:250 है।

$$S.E. = \sqrt{npq} = \sqrt{500 \times \frac{1}{2} \times \frac{1}{2}} = 11.18$$

वास्तविक और प्रेक्षित के मध्य अंतर =280-250 = 30

$$\frac{\text{Difference}}{\text{S.E.}} = \frac{30}{11.18} = 2.68$$

1% स्तर पर अंतर **2.58 S.E** से ज्यादा है।

यह नमूनाकरण उतार चढाव के कारण नहीं है।

इसलिए हम मान सकते हैं कि दोनों खाद्य पदार्थ समान रूप से लोकप्रिय नहीं है।

उदाहरण 3- एक सिक्के को 400 बार उछाला जाता है और जिसमें 216 बार चिट आता है। इस बात पर चर्चा करें कि क्या सिक्का निष्पक्ष है और इस उद्देश्य के लिए सैद्धांतिक सिद्धान्तों का संक्षेप मे वर्णन करें।

हल - निष्पक्ष सिक्के में चिट आने की संभावना उछालों =$\frac{1}{2}$

400 में अपेक्षित चिटों की संख्या = 200

लेकिन प्रेक्षित चिटों की संख्या = 216

$$S.E. = \sqrt{npq}$$

$$= \sqrt{400 \times \frac{1}{2} \times \frac{1}{2}} = \sqrt{100} = 10$$

वास्तविकता से विचलन = 216-200 = 16

$$Z = \frac{Difference}{S.E.} = \frac{16}{10} = 1.6$$

चूंकि प्रेक्षित विचलन S.E का 1.6 गुना है, जोकि 1.96 S.E., (5% स्तर) से कम है, यह निष्कर्ष निकाला जा सकता है कि शून्य परिकलपना स्वीकार्य है। इसलिए, सिक्का निष्पक्ष है।

(ब) सफलता के अनुपातों का परीक्षण(Testing Success Ratios)

प्रत्येक नमूने में सफलता की संख्या लेने के बजाय, सफलता का एक हिस्सा अर्थात p = $\frac{1}{n}$ अभिलिखित है। हिस्से की स्थिति में मानक त्रुटि की गणना निम्नवत की जाती है।

$$S.E. = \sqrt{\frac{pq}{n}}$$

जहाँ q = 1-p

उदाहरण 4- 500 अनानासों का यादृच्छिक नमूना एक बडे खेप में से लिया गया था और जिसमें 65 खराब पाये गए। दिखाऐं कि खराब अनानासों की इन नमूनों में मानक त्रुटि 0.015 आकार की है,और निष्कर्ष निकाले कि खेप में खराब अनानास का प्रतिशत लगभग 8.5 और 17.5 के बीच में होता है।

हल – यहाँ

$$p = \frac{65}{500} = 0.13,$$

$$q = 1 - 0.13 = 0.87$$

$$S.E. = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.13 \times 0.87}{500}}$$

$$= \sqrt{\frac{0.1131}{500}}$$

$$= \sqrt{.000226} = 0.015$$

खेप में खराब अनानासों की प्रतिशत सीमाऐं हैं-

$$(0.13 \pm 3S.E.) \times 100 = (0.13 \pm 3 \times 0.015)100$$

$$= (.13 \pm .045)\,100$$

$$= (13 \pm 4.5) = 17.5 \text{ और } 8.5$$

टिप्पणी 3S.E सीमाऐं लगभग निश्चित हैं।

उदाहरण 5 - सेब के एक थोक व्यापारी का दावा है कि उसके द्वारा उपलब्ध कराये गऐ सेब में केवल 4 प्रतिशत सेब दोषपूर्ण होते हैं। 600 सेबों का यादृच्छिक नमूने मे 36 सेब दोषपूर्ण थे। थोक व्यापारी के दावे का परीक्षण करें।

हल -

$$S.E = \sqrt{\frac{pq}{n}}$$

$$= \sqrt{\frac{.96 \times .04}{600}} = 0.008$$

95% विश्वसनीयता सीमाऐं =p $\pm$ 1.96 S.E

$$= p \pm 1.96 \times 0.008$$

$$= .96 \pm 0.01568$$

$$= 0.94432 \text{ to } = 0.97568$$

600 सेबों में से, अच्छे सेब 0.94432×600=566.59 से 0.97568×600=585.4 या 567 से 585 के बीच मे हो सकते हैं। इसलिए अपेक्षित दोषपूर्ण सेबों की संख्या 15 से 30 सेबों के बीच संभावित है। उसका दावा है कि दोषपूर्ण सेब 4 प्रतिशत है। लेकिन वास्तविक दोषपूर्ण संख्या 36 है। इसलिए उसका दावा स्वीकार नहीं किया जा सकता है।

उदाहरण 6 - एक बडे शहर से यादृच्छिक चयनित 600 लोगों का एक नमूने आकार दर्शाता है कि नमूने में पुरूषों का प्रतिशत 53 है। यह माना जाता है कि शहर में कुल आबादी के लिए पुरूषों का अनुपात 1:2 है। जांच करें कि इस विश्वास की पुष्टि अवलोकन द्वारा की गई है या नहीं।

हल - शून्य परिकल्पना यह है कि कुल जनसंख्या पुरूषों की संख्या 1:2 या 0.5 है।

प्रेक्षित मान = 0.53

$$S.E. = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\frac{1}{2} \times \frac{1}{2}}{600}} = \sqrt{\frac{\frac{1}{4}}{600}} = \sqrt{\frac{1}{2400}} = 0.02$$

$$S.E. = \frac{0.53 - 0.05}{S.E.} = \frac{(0.53 - 0.5)}{0.02} = 1.5$$

चूंकि z, 1.96 से कम है, 5% विश्वसनीयता के स्तर पर अंतर महत्वपूर्ण नहीं है औश्र नमूनाकरण उतार-चढाव के कारण उत्पन्न हो सकता है। इसलिए शून्य परिकल्पना को अस्वीकार नहीं किया जा सकता है। विश्वसनीयता की पुष्टि है।

(स) अनुपातों मे अंतर का परीक्षण(test of difference in proportions)-

हम विभिन्न समग्रों से दो नमूने लेते हैं और यह सत्यापित करते हैं कि सफलता का अनुपात महत्वपूर्ण है या नहीं।

सूत्र

$$S.E.(p_1 - p_2) = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

यदि S.E $(p_1 - p_2) < 1.96$

अंतर यादृच्छिक नमूनाकरण उतार चढाव के कारण माना जाता है।

उदाहरण 7- एक कारखाने मे एक हजार लेखों की जांच की जाती है और 3 प्रतिशत दोषपूर्ण पाये जाते हैं। दूसरे कारखाने से पन्द्रह सौ समान लेखों की जांच की जाती है केवल 2 प्रतिशत लेख दोषपूर्ण पाये जाते हैं । क्या यह उचित रूप से निष्कर्ष निकाला जा सकता है कि पहले कारखाने का उत्पाद दूसरे से हल्का है।

हल - आइए हम शून्य परिकल्पना तैयार करें -

$H_0: p_1 = p_2$

$$p_1 = \frac{30}{1000} = 0.03$$

$$p_2 = \frac{30}{1500} = 0.02$$

$$S.E.(p_1 - p_2) = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$p = \frac{(1000 \times 0.03) + (1500 \times 0.02)}{1000 + 1500} = \frac{(30 + 30)}{2500} = 0.024$$

$$S.E. = \sqrt{0.024 \times 0.976\left(\frac{1}{1000} + \frac{1}{1500}\right)}$$

$$= 0.006$$

$$Z = \frac{0.03 - 0.020}{.006} = 1.67$$

95% विश्वसनीयता के स्तर पर z=1.96, अंतर महत्वपूर्ण नहीं है। शून्य परिकल्पना जो कि $p_1 = p_2$ स्वीकार्य है।

उदाहरण 8- एक मशीन 500 के एक नमूने मे 16 अपूर्ण वस्तु बनाती है। मशीन की मरम्मत के बाद , यह 100 के एक खेप मे 3 अपूर्ण वस्तु बनाता है। क्या मशीन में सुधार हुआ है।

हल -

$$p_1 = \frac{16}{500} = 0.032 \text{ (पहले नमूने में)}$$

$$p_2 = \frac{3}{100} = 0.03 \quad \text{(दूसरे नमूने में)}$$

कल्पना करते हैं कि मरम्मत के बाद भी मशीन में सुधार नहीं हुआ है या $p_1 = p_2$

$$S.E.(p_1-p_2)=\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$p=\frac{n_1p_1+n_2p_2}{n_1+n_2}$$

$$p=\frac{500\times0.032+100\times0.3}{500+100}$$

$$=\frac{16+3}{600}=0.03$$

q = 1 – 0.03 = 0.97

$$S.E.(p_1-p_2)=\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$=\sqrt{(0.03)(0.97)\left(\frac{1}{500}\right)+\left(\frac{1}{100}\right)}$$

$$=\sqrt{(0.03)(0.97)[0.002+0.01]}$$

= 0.0187

$$Z=\frac{0.032-0.03}{0.0187}=\frac{0.002}{0.0187}=0.106$$

चूंकि (1% स्तर) पर अंतर 2.58 S.E से कम है, प्रयोग का परिणाम परिकल्पना का सर्मथन करता है। इसलिए हम निष्कर्ष निकालते है कि मरम्मत के बाद भी मशीन में सुधार नहीं हुआ है।

उदाहरण 9- A शहर 1000 लोगों के यादृच्छिक नमूनों में , 400 लोग गेहूँ के उपभोक्ता पाए गए। B शहर के 800 लोगों के नमूने में, 400 लोग गेहूँ के उपभोक्ता पाए गए। क्या इन आंकडों से शहर A और शहर B के बीच एक महत्वपूर्ण अंतर का पता चलता है, जहाँ तक गेहूँ उपभोक्ताओं का अनुपात है।

हल - आइए हम इस परिकल्पना को मान लें कि दोनों शहरों में , गेहूँ की खपत के अनुपात के बीच कोई अंतर नहीं हैं।

$H_0: p_1=p_2$

$$p_1=\frac{400}{1000}=0.4,\ p_2=\frac{400}{800}=0.5$$

$$p=\frac{(1000\times0.4)+(800\times0.5)}{1000+800}$$

$$=\frac{4}{9}\qquad \therefore q=\frac{5}{9}$$

$$S.E.(p_1-p_2)=\sqrt{\frac{4}{9}\times\frac{5}{9}\left(\frac{1}{1000}+\frac{1}{800}\right)}$$

$$=\sqrt{\frac{20}{81}\times\frac{9}{4000}}\ =0.024$$

$p_1 - p_2 = 0.4 - 0.5 = 0.1$

$$\frac{\text{Difference}}{\text{S.E.}} = \frac{0.1}{0.024} = 4.17$$

चूंकि अंतर 2.58 S.E. से (1% स्तर) पर ज्यादा है, नमूनाकरण के उतार चढाव के कारण ऐसा नहीं हो सका। इसलिए आंकडे शहर A और शहर B के बीच महत्वपूर्ण अंतर को दर्शाते हैं , जहाँ तक गेहूँ उपभोक्ताओं के अनुपात का सम्बन्ध है।

2.6 महत्व सृजन फलन (Importance Generation Function)

मान लीजिए X एक यादृच्छिक चर है अर्थात X नमूना अंतराल में से वास्तविक संख्याओं का एक फलन है, यादृच्छिक चर X की विभिन्न विशेषताओं की गणना में, अर्थात E(x) या V(x) हम x की प्रायिकता वितरण के साथ सीधे काम करते हैं। ख्संभाव्यता वितरण फलन द्वारा दिया जाता है या तो संभाव्यता वितरण फलन निरंतर स्थिति में , या बिन्दु संभाव्यता $p(x_1) = P(X = x_1)$ असतR स्थिति में। R के उपसमुच्च S के लिए प्रयोगात्मक मानों को लेते हुए X यादृच्छिक चर है। X का महत्व सृजन फलन Mx द्वारा परिभाषित होता है।

Mx(t)= E[exp(tx)] t के लिए R में

टिप्पणी- चूंकि exp(tx) गैर ऋणात्मक यादृच्छिक चर है , Mx(t) किसी t के लिए एक वास्तविक संख्या या धनात्मक अनन्त आस्तित्व में है।

1. दर्शाये कि यदि x एक असत R वितरण सघनता फलन f के साथ है। तो $Mx(t)=\sum_{x\epsilon s} e^{tx} f(x)$
2. दर्शाये कि यदि x सतत वितरण सघनता फलन f के साथ है, तो $Mx(t)=\int_s e^{tx} f(x)dx$

चूंकि चर घांतांकी फलन धनात्मक है , x का महत्व सजृन फलन हमेशा अस्तित्व में होता है या तो वास्तविक संख्या के रूप मे या धनात्मक अनन्त के रूप में।

उदाहरण 10- मान लीजिए x एक समान रूप से अंतराल [a,b] में वितरित है इसलिए

m.g.f $Mx(t)= \int_a^b \frac{e^{tx}}{b-a} dx$

$= \frac{1}{(b-a)t}\left[e^{bt} - e^{at}\right], t \neq 0$

द्वारा किया जाता है।

टिप्पणी – m.g.f स्वतन्त्र चरों की संख्या के योग का गुणनफल उनका m.g.f

$$E\left\{e^{t(x_1+x_2+x_3+\ldots)}\right\} = E\left(e^{tx_1}\right).E\left(e^{tx_2}\right).E\left(e^{tx_3}\right)\ldots$$

द्विपद वितरण का m.g.f

हम जानते हैं कि द्विपद वितरण की स्थिति ${}^nC_x p^x q^{n-x}$ में x सफलता की तुलनात्मक आवृत्ति है। इसलिए m.g.f मूल के बारे दिया जायेगा।

$$M_0(t)= \sum_{x=0}^{n} e^{tx}\, {}^nC_x p^{n-x}$$

$$= \sum_{x=0}^{n} {}^nC_x \left(pe^t\right)^x q^{nx}$$

$=(q + pe^t)^n$

$$=\left[q+p\left(1+t+\frac{t^2}{2!}+\dots\right)\right]^n=\left[1+pt+\frac{pt^2}{2!}+\dots\right]^n$$

$$M_0(t)=\left[1+pt+\frac{pt^2}{2!}+\dots\right]^n$$

हम यह भी जाते हैं कि इस वितरण में माध्य m=np द्वारा निकाला जाता है और

$M_a(t)=e^{-at}M_0(t)$

m.g.f माध्य के बारे निकाला जाता है।

$M_m(t)=e^{-mt}M_0(t)$ जहाँ m=np

$= e^{-npt}(q+pe^t)^n$

$= [qe^{-pt} + pe^{qt}]$

पायसन वितरण का m.g.f

हम जानते हैं कि पायसन वितरण की स्थिति में x की सफलता की प्रायिकता वितरण $e^{-m}\frac{m^x}{x!}$ के द्वारा दी जाती है।

m.g.f. मूल के बारे में द्वारा निकाला जाता है।

$$M_0(t)=\sum_{x=0}^{\infty}e^{tx}\left(\frac{e^{-m}m^x}{x!}\right)$$

$$M_0(t)=e^{-m}\sum_{x=0}^{\infty}\frac{(me^t)^x}{x!}=e^{-m}e^{me^t}$$

$$e^t=\mu'_2=\sum_{x=0}^{n}\{x(x-1)+x\}^n C_x p^x q^{n-x}$$

$$M_0(t)=e^{m(e^t-1)}$$

हम यह भी जानते हैं कि इस वितरण में माध्य m होता है और जहाँ $M_a(t)=e^{-at}M_0(t)$

माध्य के बारे में m.g.f. निकाला जायेगा

$M_m(t)=e^{-mt}M_0(t)$

$$M_m(t)=e^{m\left(e^t-1-t\right)}$$

उदाहरण 11- दर्शायें कि यदि x_1 और x_2 पायसन वितरण के साथ m_1 और m_2 क्रमशः प्राचल के साथी दो स्वतन्त्र यादृच्छिक चर हैं , तब $x_1 + x_2$ का योग पायसन वितरण के साथ प्राचल $m_1 + m_2$ के साथ एक यादृच्छिक चर है।

हल- यदि $M_1(t)$ और $M_2(t)$ X1 और X2 के m.g.f.हैं, तब

$M_1(t)=e^{m_1(e^t-1)}$ और $M_1(t)=e^{m_2(e^t-1)}$

हम यह भी जानते हैं कि स्वतन्त्र संख्या के चरों का योग m.g.f. का गुणनफल है।

(X1+X2) का m.g.f. जहाँ X1, X2 स्वतन्त्र चर है।

=X1 और X2 के गुणनफल का m.g.f.

$=M_1(t) \times M_2(t)$

$= e^{m_1 (e^t - 1)} \times e^{m_2 (e^t - 1)}$

$= e^{(m_1 + m_2)(e^t - 1)}$

= पायसन वितरण का (m_1+m_2) प्राचल के साथ m.g.f.

अतः सिद्ध

सामान्य वितरण का m.g.f.

सामान्य वितरण की स्थिति में हम जानते हैं कि मूल के माध्य के साथ प्रायिकता फलन निकाला जाता है।

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x^2}{2\sigma^2}\right)}$$

$$M_0(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx} e^{-\left(\frac{x^2}{2\sigma^2}\right)}$$

$M_0(t) = e^{mt + (1/6)t^2\sigma^2}$

माध्य m के सन्दर्भ में m.g.f.

$Mm(t) = e^{-mt} M_0(t)$

$= e^{-mt}(e^{mt + (1/2)t^2\sigma^2})$

$Mm(t) = e^{\frac{t^2\sigma^2}{2}}$

उदाहरण 12- सिद्ध करें कि यदि x_1 और x_2 माध्य m_1 और m_2 के साथ और विचलन क्रमशः के साथ स्वतन्त्र सामान्य चर हैं, तब चर $X_1 + X_2$ भी सामान्य चर माध्य m_1+m_2 और के साथ है।

हल - यदि $M_1(t)$ और $M_2(t)$ मूल से X_1 और X2 के m.g.f. है।

$M_1(t) = e^{m_1 t + \frac{1}{2}t^2\sigma_1^2}$ और $M_2(t) = e^{m_2 t + \frac{1}{2}t^2\sigma_2^2}$

और हम यह भी जानते है कि स्वतत्र संख्या के चरों का योग m.g.f.का गुणनफल होता है।

X_1+X_2 का m.g.f.जहा X_1 एवं X_2 स्वतन्त्र चर है।

= m.g.f. के X_1 और X_2 गुणनफल

= ड1;जद्ध ड2;जद्ध

$= M_1(t) \times M_2(t)$

$$= e^{m_1 t + \frac{1}{2}t^2\sigma_1^2} \times e^{m_2 t + \frac{1}{2}t^2\sigma_2^2} = e^{(m_1+m_2)(t) + \frac{1}{2}t^2\left(\sigma_1^2 + \sigma_2^2\right)}$$

सामान्य चर का m.g.f. माध्य (m_1+m_2) और विचलन $\sigma_1^2 + \sigma_2^2$. के साथ

यादृच्छिक चरों का फलन (Function of random variables)

यादृच्छिक चर स्थिति का कार्य प्रायः व्यवस्था विश्लेषण मे उत्पन्न होता है जहां व्यवस्था की कुछ विशेषताओं का ज्ञान, इनपुट के ज्ञान के साथ, आउटपुट में व्यवहार के कुछ अनुमान की अनुमति देगा। उदाहरण के लिए इनपुट यादृच्छिक चर x और इसकी सघनता f(x) ज्ञात है और इनपुट आउटपुट व्यवहार की विशेषता $Y = \phi(x)$ द्वारा दी जाती है।

हम यादृच्छिक चर y की सघनता की गणना करने में रूचि रखते हैं। ध्यान दें किसी दिए गए यादृच्छिक चर x और फलन ϕ के लिए, y एक यादृच्दिक चर की परिभाषा को संतुष्ट नहीं कर सकता। लेकिन यदि हम मानते हैं कि सतत है, तब $Y = \phi(x)$ यादृच्छिक चर होगा।

उदाहरण 13- यदि $Y=\phi(x) = X^2$ उदाहरण के रूप में निश्चित शारीरिक प्रयोग मे मापन त्रुटि को प्रदर्शित करेंगा और तब Y त्रुटि का वर्ग होगा ध्यान दें कि $F_Y(y) = 0$ $y \le 0$ के लिए $y > 0$ के लिए

$$F_Y(y) = P(Y \le y)$$
$$= P(X^2 \le y)$$
$$= P(-\sqrt{y} \le X \le \sqrt{y})$$
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

Y की सधनता के अवकलन द्वारा है

$$F_Y(y) = \frac{1}{2\sqrt{y}}\left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right] \quad ,y>0$$

$= 0$

उदाहरण 14 - यदि x(0,1) में एक समान रूप से वितरित है। हम दर्शाते हैं कि $Y= -\lambda^{-1}\ln(1-X)$ के पास $\lambda>0$ प्राचल के साथ चरघातांकी वितरण है। अवलोकन है कि Y एक गैर ऋणात्मक यादृच्छिक चर है जो दर्शाता है $F_Y(y) = 0$, $y\le 0$ के लिए $y>0$ के लिए, हमारे पास

$$F_y(y) = P(Y \le y) = P[-\lambda^{-1}\ln(1 - X) \le y]$$
$$= P[\ln(1 - X) \ge -\lambda y]$$
$$= P[(1 - X) \ge e^{-\lambda y}]$$

चूंकि e^x x का बढता हुआ फलन है

$$= P(X \le 1 - e^{-\lambda y})$$
$$= F_x(1 - e^{-\lambda y})$$

लेकिन चूंकि x(0,1) पर एकसमान है, $F_x(x) = x$, $0 \le x \le 1$इस प्रकार है-

$$F_y(y) = 1 - e^{-\lambda y}$$

इसलिए Y प्राचल λ के साथ चर घातांकी वितरण है।

नमूनाकरण सिद्धान्त (sampling principle)- यदि आंकडे केवल समग्र के एक हिस्से से एकत्र किये जाते हैं अर्थात समग्र की कुछ ही इकाईयों से , इसे नमूनाकरण के विशिष्ट समग्र (लोगों , विर्निमित वस्तुतओं आदि) जिसके बारे मे हम हर एक वस्तु पर ध्यान दिए बिना कुछ अनुमान बनाना चाहते हैं। इस प्रकार हम नमूने में अर्थात हम कुछ विशिष्ट वस्तुओं पर विचार करने की कोशिश करते हैं , जिससे हम पूरे समग्र की विशेषता जिसकी कुछ समझ है , कुछ जानकारी निकालने की आशा करते हैं। मान लीजिए कि हम एक संख्या के साथ परिमित आबादी के प्रत्येक सदस्य को लगातार वर्गीकृत करते हैं , ताकि सामान्यता के नुकसान के बिना एक

समग्र जो N वस्तुओं को शामिल करता है जिसे 1,2,....,N से प्रदर्शित किया जा सकता है। अब नीचे वर्णित किये गए वस्तुओं में n वस्तु चुनते है। निम्नलिखित यादृच्छिक चरों को परिभाषित करें।

X_1 = प्राप्त समग्र मान जब i^{th} चुना जाता है i = 1,2,....,n

यादृच्छिक चरों $X_1, X_2,....,X_n$ की संभावना वितरण स्पष्ट हम नमूनाकरण के बारे मे कैसे जानते हैं पर निर्भर करता है। यदि हम प्रतिस्थापना के साथ नमूना लेते हैं , प्रत्येक समय में वस्तु को यादृच्छिक चुनते हुए , यादृच्छिक चर $X_1,X_2,...,X_n$ स्वतन्त्र है और एक समान रूप से वितरित होते हैं। जो कि प्रत्येक X_i, i= 1,2,....,n के लिए हमारे पास है।

$P(X_i=j)=1/N$, j=1,2,...,N

इसके सिवाय उनकी संयुक्त संभावना वितरण द्वारा दिया जाता है

$$P[X_i=j_1,\dots,X_n=j_n]= \frac{1}{N(N-1)(N-2)..(N-n+1)}$$

जहां $j_1,....\ j_n$ (1,2,...,N) में से कोई n मान है।

नमूना निकालने की विधियां (Sampling methods)

नमूना निकालने की कुछ विधियां निम्नलिखित हैं-

1. सरल यादृच्दिक नमूनाकरण
2. स्तरीय यादृच्छिक नमूनाकरण

बिन्दु अनुमान (point estimate)

यदि x प्रायिकता वितरण f(x) के साथ एक यादृच्दिक चर है , जिसे अज्ञात प्राचल द्वारा अवगत कराया गया है और यदि $X_1,X_2,...,X_n$ में से n आकार के यादृच्छिक नमूने हैं , आंकडे $\hat{\theta} = h(X_1,X_2,\dots X_n)$ को θ का बिन्दु अनुमानक कहा जाता है। घ्यान दें कि θ यादृच्दिक चर है क्योंकि यह यादृच्छिक चर का एक फलन है। नमूना चयन होने के पश्चात , $\hat{\theta}$ एक विशेष संख्यात्मक मान लेता है को θ का बिन्दु अनुमानक कहते हैं। सामान्यतयाः कुछ समग्र प्राचल θ का बिन्दु अनुमानक $\hat{\theta}$ आंकडे θ का एक संख्यात्मक मान है।

आंकडा $\hat{\theta}$ को बिन्दु अनुमानक कहा जाता है।

निष्पक्ष अनुमानक (unbiased estimator)

बिन्दु अनुमानक $\hat{\theta}$ प्राचल θ के लिए एक निष्पक्ष अनुमानक है यदि $E(\hat{\theta})=\theta$

यदि अनुमानक निष्पक्ष नहीं है, तब अन्तर $E(\hat{\theta})=\theta$ को अनुमानक $\hat{\theta}$ का पक्षपाती कहा जाता है।

जब अनुमानक निष्पक्ष है, तब पक्षपाती शून्य है अर्थात $E(\hat{\theta}) - \theta = 0$

उदाहरण 15- मान लीजिए X माध्य μ और विचलन σ^2 के साथ एक यादृच्छिक चर है और $X_1,X_2,...,X_n$ समग्र में से n आकार के यादृच्छिक नमूना X द्वारा प्रदर्शित है। दर्शायें कि नमूना माध्य $\overline{X}$ और नमूना विचलन S^2 क्रमशः μ और σ^2 के निष्पक्ष अनुमानक हैं।

हल - सबसे पहले नमूना माध्य समझे हम जानते हैं कि $E(\overline{X})=\mu$ इसलिए, नमूना माध्य $\overline{X}$ समग्र माध्य μ का एक निष्पक्ष अनुमानक है। अब नमूना विचलन समझे। हमारे पास

$$E(S^2) = E\left[\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}\right] = \frac{1}{n-1}E\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

$$= \frac{1}{n-1}E\sum_{i=1}^{n}\left(X_i^2 + \overline{X^2} - 2\overline{X}X_i\right)$$

$$= \frac{1}{n-1}E\left(\sum_{i=1}^{n}X_i^2 - n\overline{X^2}\right)$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}E\left(X_i^2\right) - nE\left(\overline{X^2}\right)\right]$$

चूकि $E(X_i^2) = \mu^2 + \sigma^2$ और $E(\overline{X}^2) = \mu^2 + \frac{\sigma^2}{n}$, हमारे पास

$$E(S^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}\left(\mu^2 + \sigma^2\right) - n\left(\mu^2 + \sigma^2/n\right)\right]$$

$$= \frac{1}{n-1}\left(n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2\right)$$

$$= \sigma^2$$

इसलिए, नमूना विचलन S^2 समग्र विचलन σ^2 का एक निष्पक्ष अनुमानक है।

बिन्दु अनुमानक का विचलन (Deviation of point estimator)- यदि हम समझे θ के सभी निष्पक्ष अनुमानक, जिसका सबसे छोटा विचलन का मान (MUVE) न्यूनतम विचरण निष्पक्ष आकलनकर्ता कहलाता है। यदि $X_1, X_2, \ldots, X_n$ माध्य μ और विचलन σ^2 के साथ सामान्य वितरण में से आकार n के यादृच्छिक नमूने हैं, नमूना माध्य $\overline{X}$ μ के लिए MUVE है।

2.7 सारांश (Summary)

इस इकाई मे हमने अध्ययन किया है कि एक सांख्यिकीय परिकल्पना परीक्षण आंकडों का उपयोग करके निर्णय लेने की एक विधि है, चाहें वह नियंत्रित प्रयोग से हो या अवलोकन अध्ययन (नियंत्रित प्रयोग से न हो)। सांख्यिकीय में , एक परिणाम सांख्यिकीय रूप से महत्वपूर्ण कहलाता है , अगर इसके एक पूर्व निर्धारित सीमा संभावना के अनुसार अकेले मौके में होने की संभावना नहीं है, महत्व स्तर वाक्यांश ''महत्व का परीक्षण'' **रोनाल्ड फिशर** द्वारा दिया गया था। ''इस तरह के गंभीर परीक्षण को महत्व का परीक्षण कहा जा सकता है , और जब ऐसे परीक्षण उपलब्ध होते हैं तो हमें पता चल जायेगा कि कोई दूसरा नमूना है या जो पहले से काफी अलग नहीं है।'' अन्वेषण संबंधी आंकडा विश्लेषण के विपरीत, कभी कभी परिकल्पना परीक्षण को पुष्टिक आंकडा विशलेषण कहा जाता है। आवृत्ति की संभावना में , इन निर्णयों को लगभग शून्य परिकल्पना वाले परीक्षणों का उपयोग करके लगभग हमेशा बनाया जाता है। ये ऐसे परीक्षण हैं जो इस प्रश्न का उत्तर देते हैं कि शून्य परिकल्पना सही है, परीक्षण आंकडों के लिए एक मान का निरीक्षण करने की संभावना क्या है जो वास्तव मे प्रेक्षित मान के रूप मे कम से कम चरम हो ? अधिक औपचारिक रूप में , वे इस प्रश्न के उत्तर का प्रतिनिधित्व

करते हैं, जो एक प्रयोग करने से पहले सामने आते हैं , प्रयोग के परिणामों से गलत अस्वीकृति की पूर्व निर्दिष्ट संभावना के लिए शून्य परिकल्पना की अस्वीकृति हो सकती है। परिकल्पना परीक्षण का एक उपयोग यह तय करना है कि पारंपरिक ज्ञान पर संदेह डालने के लिए प्रायोगिक परिणामों के पर्याप्त जानकारी है या नहीं। बायिसियन परिकल्पना के परीक्षण के लिए दृष्टिकोण पिछली संभावना पर परिकल्पना की अस्वीकृति का आधार है। आंकडों के आधार पर निर्णय लेने के लिए अन्य दृष्टि कोण निर्णय सि द्धांत और अनुकूल निर्णय के माध्यम उपलब्ध है। महत्वपूर्ण क्षेत्र एक परिकलपना परीक्षण के सभी परिणामों का समुच्चय है , जिसे वैकल्पिक परिकल्पना के पक्ष में खारिज कर दिया जाता है। महत्वपूर्ण क्षेत्र को आमतौर पर शब्द c द्वारा प्रदर्शित किया जाता है।

2.8 शब्दावली (Glossary)

अनुमानक (estimator)- गुणों के नमूनों का दिया हुआ मान जो आवश्यक अर्थ खोजने के लिए उपयोग किया जाता है।

ऊपरी सीमा (upper limit)- दिये गए समग्र मे ऊपरी मान ।

2.9 बोध प्रश्न (Comprehension Question)

1. थाम्पसन प्रेस की परिकल्पना है कि इसकी नवीनतम बेब आफसेट प्रेस का औसत जीवन 14,500 घंटे है। वो जानते हैं कि प्रेस का S.D 2,100 घंटे है। 25 प्रेसों के एक नमूने मे से , कम्पनी ने 13,000 घंटे नमूना माध्य ज्ञात किया। 0.01 महत्व के स्तर पर , क्या कम्पनी ने निष्कर्ष निकालना चाहिए कि प्रेस का औसत जीवन 14,500 घंटे की तुलना में कम है।
2. फिल्म की जांच होने पर भारत में रंगमंच मालिकों को पता है कि एक हिट फिल्म प्रत्येक शहर मे 10 दिनों के मानक विचलन के साथ औसतन 84 दिनों तक चलती है। एक विशेष फिल्म वितरक को जनसंख्या के साथ अपने क्षेत्र में फिल्म की लोकप्रियता की तुलना करने में दिलचस्पी थी। उसने उस क्षेत्र के 75 थियेटरों को यादृच्छिक तरीके से चुना और पाया कि एक लोकप्रिय फिल्म 81.5 दिनों तक चलती है।

   (अ) परीक्षण के लिए उचित परिकलपनाओं का संज्ञान लेते हुए बताऐं कि क्या वितरकों के क्षेत्र मे थिएटर और आबादी के बीच महत्वपूर्ण अन्तर है।

   (ब) 1% महत्व के स्तर पर, इन अनुमानों का परीक्षण करें।

2.10 बोध प्रश्नों के उत्तर (Answers to comprehension questions)

यहां परिकल्पनाऐं निम्नानुसार लिखी जा सकती है।

H_0:μ =14500 और H_1: ? < 14500 और महत्वपूर्ण स्तर α = 0.01

α = 0.01 के लिए z क्षेत्र में निचलास्वीकृत क्षेत्र z= -2.33 या

(तालिका से z का मान 2.33 होने की संभावना 0.4901 है, स्वीकृति बाऐं पुच्छीय क्षेत्र के)

$$\bar{x} = \frac{\mu H\circ - 2\sigma}{\sqrt{n}} = \frac{145000 - 2.33(2100)}{5}$$

=13521.4 hours

चूंकि नमूना माध्य $\overline{X}$ अनुमानित मान से कहीं कम है, शून्य अवधारणा को अस्वीकार कर दिया जाता है।
निम्नलिखित आंकडों को देखते हुए σ=10 दिन n=75 सिनेमाघर $\overline{X}$ = 81.5 H_0: μ=84 दिन H_1: $\mu \neq 84$ दिन α =0.01 इसका अर्थ है कि दोनों पूंछों के तहत अस्वीकृत क्षेत्र 0.01 है और स्वीकृत क्षेत्र 0.99 है।
इसमें स्वीकृत क्षेत्र का हिस्सा आधा $\frac{0.99}{2} = .4950$ है, Z का मान 2.58 है। इसलिए स्वीकृत क्षेत्र की सीमाऐं हैं।

$$z = \pm 2.58 \text{ or } \bar{x} = \mu_{H_0} \pm \frac{z\sigma}{\sqrt{n}} = 84 \pm 2.58 \times \frac{1075}{\sqrt{}}$$

= 81.02 निचली सीमा और 86.98 ऊपरी सीमा

क्योंकि पर्यवेक्षक का मान स्वीकृति क्षेत्र में है , इसलिए हम रिक्त परिकल्पना H_0 को अस्वीकार नहीं करते हैं। फिल्म के चलने की लम्बाई अन्य सिनेमाघर के समान है।
दूसरे शब्दों में-

Z का प्रेक्षित मान है $\bar{x} - \frac{\mu_{H_0}}{SE}$

जहाँ S.E= $\frac{\sigma}{\sqrt{n}}$

$= \frac{81.4-84}{(1.155)}$

Z का स्वीकृत क्षेत्र है $\pm z = \pm 2.58$

2.11 स्वपरख प्रश्न (Self-Test Question)

1. अनुपात में अन्तर के परीक्षण से आप क्या समझते हैं?
2. महत्व सृजन फलन पर टिप्पणी लिखें?
3. मानक त्रुटि से आप क्या समझते हैं?

2.12 सन्दर्भ पुस्तकें (Reference books)

1. मूल सांख्यिकीय – गौण, गुप्ता और दास गुप्ता वल्ड प्रेस लिमिटेड -कलकत्ता
2. व्यावसायिक सांख्यिकी की बुनियादी बातें - संचेती और कापोर
3. प्रबन्धन मे मात्रात्मक तरीके - श्रीवास्तव, शेनाय और गुप्ता
4. व्यावसायि सांख्यिकी - गुप्ता और गुप्ता

इकाई 3 चरों का साथर्कता परीक्षण (बड़े प्रतिदर्श)

(Testing the Significance of Variables (Large Samples))

3.1 प्रस्तावना (Introduction)

परिकल्पना परीक्षण, आंकडे परीक्षण के नमूना वितरण पर आधारित होती है और इसलिए महत्वपूर्ण (Critical) क्षेत्र को परिभाषित करने के लिए नमूना वितरण को जानना आवश्यक है। यहां , नमूना वितरण उनमें से एक हो सकता है , जिनका पहले के अध्यायों में चर्चा की गई है या उनमें से भिन्न हो सकता है। बडें नमूनों के लिए $n > 30$ परीक्षण संबंधित नमूना वितरण पर आधारित होगा। यद्यपि, बडे नमूनों $n > 30$ के लिए अधिकांश नमूना वितरण सामान्य वितरण की ओर अग्रसर होता है, और इसलिए, परीक्षण सामान्य वितरण पर आधारित हो सकता है। आइए हम कुछ नमूना परीक्षणों पर विचार करते हैं जो सामान्य वितरण पर आधारित है।

3.2 उद्देश्य (Objectives)

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

- ✓ बडे नमूनों के लिए महत्व के परीक्षण को प्रयोग में ला सकें।
- ✓ महत्वपूर्ण परीक्षण के आधार पर एक दिये हुए आंकडे का विश्लेषण कर सकें।

3.3 माध्य के लिए परीक्षण (Test for Mean)

मान लें कि समग्र के लिए माध्य (μ) अज्ञात है। हम परीक्षण करना चाहते हैं कि दिये हुए मान का माध्य μ_0 है।

शून्य परिकल्पना $H_0: H = \mu_0$ है।

एक बडे यादृच्छिक नमूने के आकार के लिए $z = \frac{\bar{x}-\mu_0}{\mu/\sqrt{n}}$N(0,1) है और इसलिए, आंकडा परीक्षण $z = \frac{\bar{x}-\mu_0}{\mu/\sqrt{n}}$ है।

वैकल्पिक परिकलपना निम्न में से कोई भी हो सकती है।

1. $\mu_0: \mu = \mu_0$ यहां, परीक्षण दो पुच्छीय है।
2. $H_1: \mu > \mu_0$ यहां, परीक्षण एक पुच्छीय है ऊपरी पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।
3. $H_1: \mu < \mu_0$ यहां, परीक्षण एक पुच्छीय, निचले पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।
4. दो माध्यों के बीच अन्तर के लिए परीक्षण $\mu_0: \mu_1 < \mu_2$ वैकल्पिक परिकल्पना निम्न में से एक हो सकती है।

1) $H_1: \mu_1 \neq \mu_2$ यहां, परीक्षण दो पुच्छीय है।

2) $H_1: \mu_1 > \mu_2$ यहां, परीक्षण एक पुच्छीय, ऊपरी पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।

3) $H_1: \mu_1 < \mu_2$ यहां, परीक्षण एक पुच्छीय, निचले पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।

दो पुच्छीय परीक्षण की स्थिति में यदि $\propto$ महत्व का स्तर है , महत्वपूर्ण मान की स्थिति में , महत्वपूर्ण $-k_{\propto/2}$ और $k_{\propto/2}$ है। ऊपरी पूंछ परीक्षण की स्थिति में, महत्वपूर्ण मान $k_\propto$ है। निचले पुच्छीय परीक्षण में, यह मान $k_\propto$ है।

ध्यान दें – यहां, यदि μ_1 और μ_2 अज्ञात हो, आंकडा परीक्षण

$Z = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{s_1^2}{n_1}}+\sqrt{\frac{s_2^2}{n_2}}}$ जहां s_1, s_2 नमूना मानक विचलन है।

अनुपात के लिए परीक्षण - मान लीजिए कि समग्र में किसी विशेषता का अनुपात ज्ञात नहीं है , हम जानना चाहते है कि क्या अनुपात का दिया हुआ मान p_0 है।

शून्य परिकल्पना $\mu_0: p = p_0$ (समग्र p_0) है।
वैकल्पिक परिकल्पना निम्न में से कोई एक हो सकती है।
1) $H_1: p \neq p_0$ यहां, परीक्षण दो पुच्छीय है।
2) $H_1: p > p_0$ यहां, परीक्षण एक पुच्छीय, ऊपरी पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।
3) $H_1: p < p_0$ यहां, परीक्षण एक पुच्छीय निचले पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।
समग्र से n आकार के बडे यादृच्छिक नमूने में, यदि x इकाई विशेषता रखते हैं तब, नमूना अनुपात $p = \frac{x}{n}$ है।
और इसलिए, μ_0 के अन्तर्गत $Z = \frac{p-p_0}{\sqrt{\frac{p_0 q_0}{n}}}$ N(0,1) है और यही आंकडा परीक्षण है।
दो पुच्छीय परीक्षण की स्थिति में , यदि $\propto$ महत्व का स्तर है , महत्वपूर्ण मान $-k_{\propto/2}$ और $k_{\propto/2}$ है। निचले पुच्छीय परीक्षण की स्थिति में यह . $-k_{\propto}$ है।

अनुपातों की समानता के लिए परीक्षण

मान लीजिए निश्चित विशेषता के लिए दो समग्र अज्ञात अनुपातों p_1 और p_2 के साथ है। हम परीक्षण करना चाहते हैं कि अनुपात समान है। शून्य परिकल्पना $\mu_0: p_1 = p_2$ (समग्र अनुपात समान है)
वैकल्पिक परिकल्पना निम्न में से कोई एक हो सकती है।

1. $H_1: p_1 \neq p_2$ यहां, परीक्षण दो पुच्छीय है।
2. $H_1: p_1 > p_2$ यहां, परीक्षण एक पुच्छीय, ऊपरी पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।
3. $H_1: p_1 < p_2$ यहां, परीक्षण एक पुच्छीय, निचले पूंछ पर महत्वपूर्ण क्षेत्र के साथ है।

μ_0 के अंतर्गत यदि p सामान्य अनुपात है। यदि पहले समग्र में से n_1 आकार का बडा यादृच्छिक नमूना लिया जाता है n_1 की इन इकाईयों के बीच, यदि x_1 इकाईयां विशेषता रखती है तो नमूना अनुपात $p_1 = \frac{\overline{x_1}}{n_1}$ हैं ।
यदि दूसरे समग्र से n_2 आकार का बडा यादृच्छिक नमूना भी लिया जाता है। n_2 की इन इकाईयों के बीच, यदि x_2 इकाईयां विशेषता रखती है तो नमूना अनुपात $Z = \frac{p_1 - p_2}{pQ\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}$ N(0,1) है। और यही आंकडा परीक्षण है।
सामान्यता, सार्वजनिक अनुपात p ज्ञात नहीं होगा और इसलिए, इसका अनुमान नमूनों में से किया जाता है।
अनुमान $\hat{P} = \frac{x_1+x_2}{n_1+n_2} = \frac{n_1 p_1 + n_2 p_2}{n_1+n_2}$ है।
इसलिए, आंकडा परीक्षण $Z = \frac{P_1 - P_2}{\sqrt{\hat{P}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ है।
दो पुच्छीय परीक्षण की स्थिति में , महत्वपूर्ण मान $-K_{\alpha/2}$ और $K_{\alpha/2}$ है। ऊपरी पुच्छीय परीक्षण की स्थिति में महत्वपूर्ण मान $k_{\propto}$ है। निचले पुच्छीय परीक्षण की स्थिति में यह मान $-k_{\propto}$ है।
उदाहरण 1-'मंदाकिनी मिल्क' 500 मिलीलीटर के प्रत्येक पाउच में दूध बेचता है। दूध को एक मशीन द्वारा पाउच में भरा जाता है जिसके लिए भरने का मानक विचलन 5 मिलीलीटर है। तीन अलग अलग संभावित स्थितियां हैं -

(अ) यह सत्यापित करना आवश्यक है कि क्या मशीन 500 मिलीलीटर दूध औसतन भर रही है । इसका अर्थ है कि मशीन ठीक से व्यवस्थित है।

(ब) ग्राहकों से शिकायत है कि पाउच में 500 मिलीलीटर से कम दूध की मात्रा है।

(स) प्रबंध चाहता है कि औसतन दूध 500 मिलीलीटर से ज्यादा नहीं होना चाहिए।

(द) मान लीजिए कि ऊपर की स्थितियों में से कोई एक स्थिति घटित हेाती है और हमें किसी निष्कर्ष में पहुंचने कि आवश्यकता है। भरे हुए पाउचों के बीच 72 यादृच्छिक तरीके से उठाऐं जाते हैं और उनकी मात्रा को मापा जाता है। इन मापों का माध्य 501.1 मिलीलीटर पाया गया। आपका निष्कर्ष क्या है ? (5% महत्व के स्तर पर परीक्षण करें)

हल- यहां $\mu_0 = 500$ मिलीलीटर, $\sigma = 5$ मिलीलीटर, n=72, $\bar{x} = 501.1$ मिलीलीटर $\propto = 0.05$ मिलीलीटर

(अ) यह सत्यापित करना आवश्यक है कि क्या समग्र माध्य $\mu = 500$ मिलीलीटर है या नहीं और इसलिए, शून्य परकिल्पना $\mu_0 = \mu \neq 500$ मिलीलीटर है। (मशीन ठीक से व्यवस्थित है)

वैकल्पिक परिकल्पना

$\mu_0 = \mu \neq 500$ (मशीन ठीक से व्यवस्थित नहीं है।),

μ_0 के अन्तर्गत, आंकडा परीक्षण $Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1)

यहां, परीक्षण दो पुच्छीय है, क्योंकि $\mu_0 : \mu \neq 500$ मिलीलीटर

दिये हुए नमूने के लिए 5 प्रतिशत महत्व के स्तर पर, महत्वपूर्ण मान $-K_{\alpha/2} = -1.96$ और $K_{\alpha/2} = 1.96$

Z के सही मान के $Z_{obs} = \frac{\bar{x}\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{501.1-500}{\frac{5}{\sqrt{72}}} = 1.87$

क्योंकि $Z_{obs} = 1.87$ एक मान (-1.96, 1.96) के अन्तराल में है, μ_0 स्वीकार्य है।

निष्कर्ष- माध्य 500 मिलीलीटर है इसका अर्थ है कि मशीन ठीक से व्यवस्थित है।

(ब) यह सत्यापित करना आवश्यक है कि क्या समग्र माध्य 500 मिलीलीटर से कम है , और इसलिए, शून्य परिकल्पना $\mu_0 : \mu = 500$ (माध्य 500 मिलीलीटर है)

वैकल्पिक परिकल्पना

$\mu_0 : \mu < 500$ मिलीलीटर है। (माध्य 500 मिलीलीटर से कम है)

μ_0 के अंतर्गत, आंकडा परीक्षण $Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1) है।

यहां, परीक्षण निचला पुच्छ है क्योंकि $\mu_1 : \mu > 500$ मिलीलीटर

5 प्रतिशत महत्व के स्तर पर , महत्वपूर्ण मान $k_\propto = 1.64$ है और महत्वपूर्ण क्षेत्र $Z > 1.645$ है।

$Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{501.1-500}{\frac{5}{\sqrt{72}}} = 1.87$

$Z_{obs} = 1.87 > 1.645$, यह 500 मिलीलीटर से कम नहीं है।

(स) यह सत्यापित करना आवश्यक है कि क्या समग्र माध्य 500 मिलीलीटर से ज्यादा है और इसलिए , शून्य परिकलपना $\mu_0 : \mu = 500$ मिलीलीटर (माध्य 500 मिलीलीटर से ज्यादा है)।

μ_0 के अन्तर्गत, आंकडा परीक्षण $Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1) है।

यहां, परीक्षण ऊपरी पुच्छ है क्योंकि $\mu_1 : \mu > 500$ मिलीलीटर।

5 प्रतिशत महत्व के स्तर पर महत्वपूर्ण मान $k_\propto = 1.645$ है और महत्वपूर्ण क्षेत्र $Z > 1.645$ है।

$Z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{501.1-500}{\frac{5}{\sqrt{72}}} = 1.87$

चूंकि $Z_{obs} + 1.87 > 1645$, μ_0अस्वीकार्य हैं।

निष्कर्ष- माध्य 500 मिलीलीटर से ज्यादा है। इसलिए औसत भरान 500 मिलीलीटर से ज्यादा है।

उदाहरण 2- एक फर्म प्रतिरोधक बनाती है। उने प्रतिरोधों का मानक विचलन 0.02 ओम ज्ञात किया गया। क्या उनका औसत प्रतिरोधक 1.39 ओम है का परीक्षण आवश्यक है। 64 प्रतिरोधकों का एक यादृच्छिक नमूना जिसका माध्य 1.39 ओम है लिया जाता है। इस नमूने के आधार पर , क्या हम निष्कर्ष निकाल सकते हैं कि सम्पूर्ण समूह का औसत प्रतिरोधक 1.4 ओम है।

हल- यहां, $\mu_0 = 1.4$ ओम , $\sigma = 0.02$ ओम, n=64 , $\bar{x} = 1.39$ ओम

महत्व का स्तर अंकित नहीं किया गया है इन परिस्थितियों में हम $\propto = 5\%$ मान सकते हैं। शून्य परिकल्पना $\mu_0 : \mu = 1.4$ ओम (माध्य प्रतिरोधक 1.4 ओम है)

वैकल्पिक परिकलपना $\mu_0 : \mu \neq 1.4$ ओम (माध्य प्रतिरोधक 1.4 ओम के समान नहीं है)

μ_0 के अन्तर्गत आंकडा परीक्षण $Z = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1) है।

यहां, परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $k{\propto}/_2 = -1.96$ है।

$Zobs = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1.39-1.4}{\frac{0.02}{64}} = -4$ चूंकि $Zobs = 4$ का मान अन्तराल (-1.96, 1.96)से बाहर है, μ_0 अस्वीकार्य है।

निष्कर्ष- प्रतिरोधकों का औसत प्रतिरोधक 1.4 ओम के बराबर नहीं है।

उदाहरण 3- इस परिकल्पना की जांच करना आवश्यक है कि औसतन, 1 पंजाबी 180 सेन्टीमीटर लम्बे होते हैं। इसके लिए, 50 पंजाबी यदृच्छया चयनित किये और उनकी ऊँचाई मापी गई। यदि औसत ऊँचाई 181.1 सेमी0 है और मानक विचलन 3.3 सेमी0 है 1 आपका निष्कर्ष है? (महत्व का स्तर 1% उपयोग करें।)

हल- यहां, $\mu_0 = 180$ सेमी., $n = 50$, $\bar{x} = 181.1$सेमी. , $\sigma = 3.3$ सेमी. और $\propto = 0.01$

शून्य परिकल्पना $\mu_0 : \mu = 180$ सेमी. (पंजाबियों की औसत ऊँचाई 180 सेमी. है)

वैकल्पिक परिकल्पना $\mu_0 : \mu > 180$ सेमी. (औसत ऊँचाई 180 सेमी. से ज्यादा है)

μ_0 के अन्तर्गत, आंकडा परीक्षण $Z = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1) है।

यहां, परीक्षण ऊपरी पुच्छ है।

1% महत्व के स्तर पर, महत्वपूर्ण मान $k_\propto = 2.33$ है और महत्वपूर्ण क्षेत्र

Z > 2.33 है।

$Zobs = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{181.1-180}{\frac{3.3}{\sqrt{50}}} = 2.36$ चूंकि $Zobs =$ 2.36, 2.33 से ज्यादा है, μ_0 अस्वीकार्य है।

निष्कर्ष- औसतन पंजाबी 180 सेमी से लम्बे हैं।

उदाहरण 4- समग्र में से 836 माध्य के साथ, 225 प्रेक्षणों वाला यादृच्छिक नमूना लिया गया है, नमूने के लिए माध्य एवं मानक विचलन क्रमशः 840.5 और 45 है।

1% महत्व के स्तर पर , परीक्षण करें कि क्या नमूना माध्य, समग्र माध्य से उल्लेलखनीय ढंग से भिन्न है?

हल- यहां, $\mu_0 = 836, n = 225, \bar{x} = 840.5, \sigma = 45$ और $\propto = 0.01$

शून्य परिकलपना $\mu_0: \mu = 836$ (नमूना माध्य समग्र माध्य से उल्लेखनीय ढंग से भिन्न नहीं है।)

वैकल्पिक परिकल्पना $\mu_0: \mu \neq 836$ (नमूना माध्य समग्र माध्य से उल्लेखनीय ढंग से भिन्न है)

μ_0 के अन्तर्गत, आंकडा परीक्षण $Z = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}}$ N(0,1) है।

यहां, परीक्षण दो पुच्छीय हैं

10% महत्व के स्तर पर, महत्वपूर्ण मान $-k\propto/_2 = -2.58$ और $k\propto/_2 = 2.58$

$Zobs = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{840.5-836}{\frac{45}{\sqrt{50}}} = \frac{4.5}{6.33} = 0.1579$ चूंकि $Zobs =$ 1.5 मान अंतराल (- 2.58, 2.58) से बाहर है, μ_0 स्वीकार्य है।

निष्कर्ष - नमूना माध्य समग्र माध्य से उल्लेखनीय ढंग से भिन्न नहीं है।

उदाहरण 5- यह ज्ञात है कि लडकों के IQ का मानक विचलन 10 और लडकियों के IQ का मानक विचलन 12 है। 200 यादृच्छिक चयनित लडकों का माध्य 99 और 300 यादृच्छिक चयनित लडकियों का माध्य 97 है।

I. क्या हम यह निष्कर्ष निकाल सकते हैं कि औसतन लडकों एवं लडकियों का IQ एक समान हो सकता है?

II. क्या हम यह निष्कर्ष निकाल सकते हैं कि औसतन लडकों का IQ लडकियों के IQ से, ज्यादा हो सकता है?

हल- यहां, $\sigma_1 = 10, \sigma_2 = 12, n_1 = 200, \ n_2 = 300, \overline{x_1} = 99$

$\overline{x_2} = 97$

चूंकि $\propto$ अंकित नहीं किया गया है, हम $\propto = 5\% = 0.05$ ले सकते हैं

(i) शून्य परिकल्पना $\mu_0: \mu = \mu_2$ (लडकों एवं लडकियों का IQ एक समान है)

(ii) वैकल्पिक परिकल्पना $\mu_1: \mu_1 = \mu_2$ (लडकों एवं लडकियों का IQ एक समान नहीं है।)

μ_0 के अंतर्गत, आंकडा परीक्षण $Z = \frac{\overline{x_1} - x_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}} isN(0,1)$

$$Z_{obs} = \frac{\overline{x_1} - x_2}{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}} = \frac{99-97}{\sqrt{\frac{10^2}{200} + \frac{12^2}{300}}} = 2.02$$

चूॅंकि $Zobs = 2.02$ का मान अन्तराल (-1.96, 1.96) से बाहर है,

μ_0 अस्वीकार्य है।

निष्कर्ष- लडकें एवं लडकियों का IQ एक समान है।

(ii) यहां वैकल्पिक परिकल्पना

$\mu_1 : \mu_1 > \mu_2$ (लडकों का IQ लडकियों से ज्यादा है।) यहां, परीक्षण ऊपरी पुच्छ है।

5% महत्व के स्तर पर , महत्वपूर्ण मान $k_\propto = 1.645$ है। चूंकि $Zobs = 2.02$, 1.645 से ज्यादा है, μ_0 अस्वीकार्य है।

निष्कर्ष- लडकों का IQ लडकियों के IQ से ज्यादा है।

उदाहरण 6- एक बाग से 1000 सेवों को यादृच्छिक नमूना जिसका माध्य वजन 187 ग्राम एवं मानक विचलन 8 ग्राम है। दूसरे बाग से 800 सेवों का यादृच्छिक नमूना जिसका माध्य वजन 188.4 ग्राम एवं मानक विचलन 10 ग्राम है । परिकल्पना का परीक्षण करें कि दोनों बागों के सेवों का औसत वजन एक समान है।

यहां $n_1 = 1000, \quad \overline{x_1} = 187$ग्राम $\quad \sigma_1 = 8$ ग्राम

$n_2 = 800, \quad \overline{x_2} = 188.4$ग्राम $\quad \sigma_2 = 10$ ग्राम

शून्य परिकल्पना $\mu_0 = \mu_1 = \mu_2$ (माध्य एक समान है)

वैकल्पिक परिकल्पना $\mu_1 \neq \mu_1 \neq \mu_2$ (माध्य एक समान नहीं है)

μ_0 के अन्तर्गत आंकडा परीक्षणर् $Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}}$ N(0,1) है।

परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $-k\propto_{/2} = 1.96$ और $k\propto_{/2} = 1.96$ है।

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}} = \frac{187 - 188.4}{\sqrt{\frac{8^2}{1000} + \frac{10^2}{800}}} = -3.22$$

चूंकि $Zobs = 3.22$ मान अंतराल (-1.96, 1.96) के बाहर है, μ_0 अस्वीकार्य है।

निष्कर्ष- दो बागों के सेवों का औसत माध्य वजन एक समान नहीं है ।

उदाहरण 7- किसी बीमारी से ग्रसित यादृच्छिक चयनित मरीजों के समूह का प्र कुंचक रक्तचाप अध्ययन में एवं दूसरे समूह में 36 व्यक्तियों जो कि किसी बीमारी से ग्रसित नहीं है, निम्नलिखित परिणाम निकाले

	ग्रसित	गैर ग्रसित
नमूना आकार	36	36
माध्य प्रकुंचक चाप	178	141
मानक विचलन	24	12

परीक्षण करें कि क्या बीमारी से ग्रसित मरीजों का औसत प्रकुंचक , गैर ग्रसित व्यक्ति यों से ज्यादा है। परीक्षण 1% महत्व के स्तर पर करें।

हल- यहां $n_1 = 36 \quad \overline{x_1} = 178 \quad \sigma_1 = 24$

$n_2 = 36 \quad \overline{x_2} = 141 \quad \sigma_2 = 12$ और $\propto = 0.01$

शून्य परिकल्पना $H_0 = \mu_1 = \mu_2$ (औसत प्रकुंचक रक्तचाप दोनों समूहों का एक समान है)
वैकल्पिक परिकल्पना $H_1 = \mu_1 > \mu_2$ (औसत प्रकुंचक रक्तचाप बीमारी से ग्रसित मरीजों का गैर ग्रसित व्यक्तियों से ज्यादा है)

H_0 के अंतर्गत,आंकडा परीक्षणर् $Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$ N(0,1) है।

यहां परीक्षण ऊपरी पुच्छ है।

1% महत्व के स्तर पर , महत्वपूर्ण मान $k_\propto = 2.33$है।

$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}} = \frac{178-141}{\sqrt{\frac{24^2}{36} + \frac{12^2}{36}}} = -8.27$

चूंकि $Zobs = 8.27$, 2.33 से ज्यादा है,H_0 अस्वीकार्य है।

निष्कर्ष- औसत प्रंकुचक रक्तचाप बीमारी से ग्रसित मरीजों का, गैर ग्रसित व्यक्तियों से ज्यादा है।

उदाहरण 8- लडके एवं लडकियों के दो समूहों के बुद्धि परीक्षण में निम्नलिखित परिणाम निकले-

	माध्य अंक	मानक विचलन	नमूना आकार
लडके	70	20	250
लडकियां	75	15	150

क्या हम 1% महत्व के स्तर पर यह निष्कर्ष निकाल सकते हैं कि लडकियों के औसत अंक लडकों से ज्यादा है।

हल- यहां $n_1 = 250$ $\overline{x_1} = 70$ $\sigma_1 = 20$
$n_2 = 150$ $\overline{x_2} = 75$ $\sigma_2 = 15$ और $\propto = 1\% = 0.01$

शून्य परिकल्पना $H_0 = \mu_1 = \mu_2$ (माध्य एक समान है) है।

वैकल्पिक परिकल्पना $H_0 = \mu_1 < \mu_2$ (लडको के औसत अंक, लडकियों के औसत अंक से है) है।

H_0 के अन्तगत आंकडा परीक्षण $Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$ N(0,1) है।

परीक्षण एक पुच्छीय है - लडकों के औसत अंक लडकियों से कम है।

1% महत्व के स्तर पर, महत्वपूर्ण मान $-k_{\propto/2} = -2.33$ है।

$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}} = \frac{70-75}{\sqrt{\frac{20^2}{250} + \frac{15^2}{150}}} = -2.84$

चूंकि $Zobs = 2.84$, -2.33 से कम है, H_0 अस्वीकार्य है।

निष्कर्ष- लडकों के औसत अंक लडकियों से कम है या लडकियों के औसत अंक, लडकों से ज्यादा है।

उदाहरण 9- एनाजेसिक (दर्द निवारक) की कुछ खुराक जब 32 महिला रोगियों को दी जाती है तो दर्द निवारण की औसत अवधि 3.5 घन्टे थी। वहीं खुराक जब 36 पुरूष रोगियों को दी जाती है तो दर्द निवारण की औसत अवधि 4 घंटे थी। पिछले अनुभवों से, यह ज्ञात है कि दर्द राहत की अवधि का मानक विचलन 0.5 घंटे हैं। यह परीक्षण करें कि औसतन, पुरूषों और महिलाओं में दर्द निवारण की अवधि एक समान है।

हल- यहां , $\sigma_1 = \sigma_2 = 0.5$ घंटे $n_1 = 32$ $\overline{x_1} = 3.5$घंटे $n_2 = 36$ और $\overline{x_2} = 4$ घंटे चूंकि $\propto$ अंकित नहीं किया गया है, हम $\propto = 5\% = 0.05$ समझ लेते हैं।

शून्य परिकल्पना

$H_0 = \mu_1 = \mu_2$ (औसत दर्द निवारण की अवधि पुरूषो एंव महिलाओं में एक समान है) है।

वैकल्पिक परिकल्पना

$H_1 = \mu_1 \neq \mu_2$ (औसत दर्द निवारण की अवधि पुरूषों एवं महिलाओं में एक समान नहीं है) है।

H_0 के अन्तगत आंकडा परीक्षण $Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$ N(0,1) है।

यहां, परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $-k\propto/_2 = -1.96$ और $k\propto/_2 = 1.96$

$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}} = \frac{3.5-4}{\sqrt{\frac{(0.5)^2}{32} + \frac{(0.5)^2}{36}}} = -33.88$

चूंकि $Zobs = 33.88$ मान अन्तराल (-1.96, 1.96) के मान से ज्यादा है, H_0 अस्वीकार्य है।

निष्कर्ष- औसत दर्द निवारण की अवधि पुरूषों एवं महिलाओं में एक समान नहीं है।

उदाहरण 10- बैंड R कलम का निर्माण करने वाले कहना है कि बैंगलौर के कालेज के छात्रों का अनुपात जो R कलम का प्रयोग करते हैं , 0.3 से अधिक है इस विवाद का परीक्षण करने के लिए , 40 छात्रों को यादृच्छिक तरीके से चुना गया और इस संबंध में पूछताछ की गई। इन 40 छात्रों में से 10 ब्रांड R कलम का उपयोग करने के लिए पाए गए। 0.05 महत्व के स्तर पर, जांच करें कि निर्माताओं का तर्क स्वीकार्य है या नहीं ।

हल- यहां $p_0 = 0.3,\ n = 40,\ x = 10$ और $\propto = 0.05$

और इसलिए, नमूना अनुपात $p = \frac{x}{n} = \frac{10}{40} = 0.25$

शून्य परिकल्पना

$H_0 = p = 0.3$ (ब्रान्ड R के कलम का प्रयोग करने वालों का अनुपात 0.3 है) है।

वैकल्पिक परिकल्पना

$H_0 = p > 0.3$ (ब्रांड R के कलम का प्रयोग न करने वालों का अनुपात 0.3 है) है।

H_0 के अंतर्गत आंकडा परीक्षण $Z = \frac{p - p_0}{\sqrt{\frac{p_0 Q_0}{n}}} = \frac{0.25 - 0.3}{\sqrt{\frac{0.3*0.7}{40}}} = -0.69$

$Zobs = -0.69$, 1.645 से कम है, H_0 स्वीकार्य है।

निष्कर्ष- ब्रान्ड R के कलम का प्रयोग करने वालों का अनुपात 0.3 है और यह 0.3 ज्यादा नहीं है।

उदाहरण 11- इस पुस्तक के लेखक का मानना है कि कर्नाटक के पीयूसी के 90% से ज्यादा छात्र , राजमोहन द्वारा रचित सांख्यिकी पुस्तक का उल्लेख करते हैं पूरे कर्नाटक से 225 सांख्यिकी छात्रों के यादृच्छिक निरीक्षण के लिए लिया जाता है, उनमें से 93.1% उवन्त पुस्तक को दर्शाते हैं। क्या हम यह निष्कर्ष निकाल सकते हैं कि 1% महत्व के स्तर पर लेखक की राय मान्य है।

हल- यहां $P_0 = \frac{90}{100} = 0.9,\ n = 225, \quad p = \frac{93.1}{100} = 0.931$

$\propto = 0.01$

शून्य परिकल्पना

$\mu_0 : p = 0.9$ (यह पुस्तक 90% छात्रों द्वारा उल्लेखित है) है।

$H_1 : P > 0.9$(यह पुस्तक 90% से ज्यादा छात्रों द्वारा उल्लेखित है) है।

μ_0 के अन्तर्गत आंकडा परीक्षण $Z = \frac{p-p_0}{\sqrt{\frac{p_0 Q_0}{n}}}$ N(0,1) है।

यहां, परीक्षण ऊपरी पुच्छ है।

1% महत्व के स्तर पर, महत्वपूर्ण मान $k_{\propto} = 2.33$ है।

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0 Q_0}{n}}} = \frac{0.931 - 0.9}{\sqrt{\frac{0.9 * 0.1}{225}}} = -1.55$$

$Zobs = -1.55, 2.33$ से कम है, H_0 स्वीकार्य है।

निष्कर्ष- पुस्तक को 90% छात्रों द्वारा नहीं अपितु 90% से अधिक छात्रों द्वारा संदर्भित किया जाता है। इस प्रकार, लेखक की राय मान्य नहीं है।

उदाहरण 12- यह सत्यापित करना आवश्यक है कि क्या एक सिक्का पक्षपाती है। सिक्के को 32 बार उछाला जाता है और परिणाम लि खे जाते हैं। 32 में से 19 परिणामों में चिट घटित हेाता है। क्या हम निष्कर्ष निकाल सकते हैं कि सिक्का पक्षपाती है।

हल- हम जानते हैं कि निष्पक्ष सिक्के के लिए चिट आने की प्रायिकता 0.5 है।

इसलिए, हम परीक्षण करते हैं कि क्या $p = 0.5$

इस प्रकार $p_0 = 0.5, \quad n = 32, \quad x = 19$

इसलिए, $p = \frac{x}{n} = \frac{19}{32} = 05938$

शून्य परिकल्पना

$\mu_0 : p = 0.5$ (सिक्का निष्पक्ष है) है।

वैकल्पिक परिकल्पना

$H_0 : p \neq 0.5$ (सिक्का पक्षपाती है)

H_0 के अन्तर्गत आंकडा परीक्षण $Z = \frac{p-p_0}{\sqrt{\frac{p_0 Q_0}{n}}}$ N(0,1) है।

यहां परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $-k_{\propto/2} = -1.96$ और $k_{\propto/2} = 1.96$

है।

$Zobs = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}} = \frac{0.5938-0.5}{\sqrt{\frac{0.5*0.5}{32}}} =$ 1.06

$Zobs =$ 1.06 मान, अन्तराल (-1.96, 1.96) में है, हम μ_0 स्वीकार्य करते हैं

निष्कर्ष- सिक्का निष्पक्ष है।

उदाहरण 13 - प्रायिकता सिद्धान्त के अनुसार , एक ऐसे परिवार जिसमें दो बच्चे हैं और दोनों ही लडके हैं कि प्रायिकता 0.25 है और इलाके में जहां 136 परिवारों में प्रत्येक में दो बच्चे हैं। जिनमें से 46 परिवारों में 2 लडके हैं। क्या यह जानकारी सिद्धान्त की पुष्टि करती है। (1% महत्व के स्तर पर परीक्षण करें)

हल- यहां, $p_0 = 0.25, \quad n = 136, \quad x = 46\,, \quad \propto = 0.01$

इसलिए $p = \frac{x}{n} = \frac{46}{136} = 0.3382$

शून्य परिकल्पना

$H_0 : p = 0.25$ (दो लडकों के साथ परविारों का अनुपात 0.25 है) है।

वैकल्पिक परिकल्पना

$H_1 : p \neq 0.25$ (अनुपात 0.25 में से भिन्न है) है।

H_0 के अंतर्गत, आंकडा परीक्षण $Z = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}}$ N(0,1) है।

यहां, परीक्षण दो पुच्छीय है।

1% महत्व के स्तर पर, महत्वपूर्ण मान . $-k_{\propto/2} = -2.58$और $k_{\propto/2} = 2.58$ है।

$Zobs = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}} = \frac{0.3382-0.25}{\sqrt{\frac{0.25*0.75}{136}}} =$ 2.375

चूंकि $Zobs =$ 2.375 मान, अन्तराल (-2.58, 2.58) में है, हम H_0 स्वीकार्य करते हैं

निष्कर्ष - 2 लडके के साथ परिवारों का अनुपात 0.25 है।

उदाहरण 14- यह परीक्षण करना आवश्यक है कि क्या छात्रों के बीच सिगरेट के घुएं का अनुपात व्याख्याताओं के बीच में कम है। 60 यादृच्छिक चुने गए छात्रों में 2 धूम्रपान करने वाले थे। 17 यादृच्छिक चुने गए व्याख्याताओं में से 5 धूम्रपान करने वाले थे। आपका निष्कर्ष क्या होगा?

हल- यहां $n_1 = 60, \quad x_1 = 2, \quad p_1 = \frac{x_1}{n_1} = \frac{2}{60} = 0.0333$

$$n_2 = 17, \quad x_2 = 5, \quad p_2 = \frac{x_2}{n_2} = \frac{5}{17} = 0.2941$$

शून्य परिकल्पना

$H_0 = p_1 = p_2$ (छात्रों के बीच धूम्रपान करने वालों एवं व्याख्याताओं के बीच धूम्रपान करने वालों का अनुपात समान है) है।

वैकल्पिक परिकल्पना

$H_1 = p_1 < p_2$ है (छात्रों के बीच धूम्रपान करने वालों का अनुपात , व्याख्याताओं के बीच धूम्रपान करने वालों के अनुपात से कम है।)

H_0 के अन्तगत आंकडा परीक्षणर् $Z = \frac{p_1 - p_2}{\sqrt{\hat{p}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ N(0,1) है।

यहां, परीक्षण निचला पुच्छीय है।

5% महत्व के स्तर पर , महत्वपूर्ण मान $-k_{\propto} = -2.33$ है चूंकि सार्वजनिक अनुपात p अज्ञात है , इसका आंकलन दिये हुए आंकडे से करते हैं। आंकलन है-

$$\mathrm{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{2+5}{60+17} = 0.0909$$

$$Zobs = \frac{p_1 - p_2}{\sqrt{\hat{p}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}} = \frac{0.0333 - 0.2941}{\sqrt{0.0909 * 0.9091\left[\frac{1}{60}+\frac{1}{17}\right]}} = \frac{-0.2608}{\sqrt{\frac{0.0909 * 0.9091}{60 * 17}}}$$
$$= -3.302$$

चूंकि $Zobs = -3.302$ मान -2.33 से कम है, H_0 अस्वीकार्य है।

निष्कर्ष- छात्रों के बीच धूम्रपान करने वालों का अनुपात व्याख्याताओं के बीच धूम्रपान करने वालों के अनुपात से कम है।

उदाहरण 15- एक घंटे की अवधि के दौरान बैंगलौर में मैसूर बैंक जंक्शन को पार कर चुके 326 स्कूटरों में से 143 ब्रांड बी स्कूटर थे। एक घंटे की अवधि के दौरान पुणे में ंशिवाजी मूर्ति जंक्शन को पर कर चुके 213 स्कूटरों में, 137 ब्रांड की स्कूटर थे। बैंगलौर की सडकों पर ब्रांड बी स्कूटर का अनुपात पुणे के अनुपात से भिन्न है।

हल- यहां, $n_1 = 326, \quad x_1 = 143, \quad p_1 \frac{x_1}{n_1} = \frac{143}{326} = 0.4387$

$$n_2 = 213, \; x_2 = 137\,, \quad p_2 = \frac{x_2}{n_2} = \frac{137}{213} = 0.6432$$

शून्य परिकल्पना

$H_0 = p_1 = p_2$ है(बैंगलौर एवं पुणे की सडकों पर ब्रांड बी स्कूटर का अनुपात एक समान है)

वैकल्पिक परिकल्पना

$H_1 = p_1 < p_2$ है (बैंगलौर एवं पुणे की सडकों पर ब्राड बी स्कूटर का अनुपात भिन्न है)

H_0 के अंतर्गत, आंकडा परीक्षणर् $Z = \frac{p_1 - p_2}{\sqrt{\hat{p}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$

$$= \frac{0.4387 - 0.6332}{\sqrt{0.5195 * 0.4805\left[\frac{1}{326}+\frac{1}{213}\right]}} = -4.646$$

$Zobs = -4.646$ मान अन्तराल (-1.96, 1.96) से बाहर है। H_0 अस्वीकार्य है।

निष्कर्ष- बैंगलौर की सडकों में ब्रांड B स्कूटरों का अनुपात, पूणे की सडकों में ब्रांड B स्कूटरों के अनुपात से भिन्न है।

उदाहरण 16- निम्नलिखित आंकडों से, परीक्षण करें क्या दो नमूनों के अनुपातों के बीच अन्तर उल्लेखनीय है।

	आकार	अनुपात
नमूना-I	1,000	0.02
नमूना-II	1,200	0.01

हल- यहां $n_1 = 1000, \quad n_2 = 1200, \quad p_1 = 0.02, \quad p_2 = 0.01$

इसलिए, $\hat{p} = \frac{n_1p_1+n_2p_2}{n_1+n_2} = \frac{100*0.02+1200*0.01}{1000+1200} = 0.0146$

शून्य परिकल्पना

$H_0 = p_1 = p_2$ है (अनुपात एक समान है)

वैकल्पिक परिकल्पना

$H_1 = p_1 \neq p_2$ (अनुपात भिन्न है)

H_0 के अंतर्गत, आंकडा परीक्षणर् $Z = \frac{p_1-p_2}{\sqrt{\hat{P}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$ N(0,1) है।

यहां परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $-k\propto/_2 = -1.96$ और $k\propto/_2 = 1.96$

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}} = \frac{0.02 - 0.01}{\sqrt{0.0146 * 0.9854\left[\frac{1}{1000}+\frac{1}{1200}\right]}} = 1.9471$$

चूंकि $Zobs = 1.9471$ मान अन्तर (-1.96, 1.96) के बीच है। H_0 स्वीकार्य हैं।

निष्कर्ष- अनुपात एक समान हैं।

उदाहरण 17- यह परीक्षण करना आवश्यक है कि क्या एक सिक्का पक्षपाती है।

(अ) मान लीजिए सिक्का 40 बार उछाला जाता है और चिट के परिणामस्वरूप सिक्के का अनुपात 0.4 है। निष्कर्ष क्या है।

(ब) मान लीजिए सिक्के को 100 बार उछाला जाता है और चिट के परिणामस्वरूप सिक्के का अनुपात 0.4 है। निष्कर्ष क्या है।

हल- दोनों परिस्थितियों में, हमारे पास $H_0: P = 0.5$ (सिक्का निष्पक्ष है)

$H_1: P \neq 0.5$ (सिक्का पक्षपाती है)

H_0 के अन्तगत आंकडा परीक्षणर् $Z = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}}$ N(0,1) है।

यहां, परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, महत्वपूर्ण मान $-k\propto/_2 = -1.96$ और $k\propto/_2 = 1.96$

(अ) यहां, n=40 और p=0.4, $Zobs = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}} = \frac{0.4-0.5}{\sqrt{\frac{0.5*0.5}{40}}} = -1.2649.$

चूंकि $Zobs = -1.2649$ मान अन्तराल (-1.96, 1.96) के बीच है। H_o स्वीकार्य है।

निष्कर्ष- सिक्का निष्पक्ष है।

(ब) यहां, n=100, p=0.4 अतः $Zobs = \frac{p-p_0}{\sqrt{\frac{p_0Q_0}{n}}} = \frac{0.04-0.5}{\sqrt{\frac{0.5*0.5}{100}}} = -2$

चूंकि $Zobs = -2$ मान अन्तराल (-1.96, 1.96) के बाहर है। H_o अस्वीकार्य है।

निष्कर्ष- सिक्का पक्षपाती है।

ध्यान दें- परिस्थिति (अ) और परिस्थिति (ब) में अनुपात (0,4) एक समान है। इसके बावजूद निर्णय एक समान नहीं है। क्योंकि बडे n के लिए, SE छोटा है।

3.4 बडे नमूने (Large Samples)

बडे एवं छोटे नमूनों में भेद करना बहुत कठिन है। यदि नमूना आकार 30 से बडा है या $n > 30$, तब उन नमूनों को बडा नमूना कहा जा सकता है।

बडे एवं छोटे नमूने के बीच अन्तर महत्व के परीक्षण के प्रयोग से है , क्योंकि हम दो नमूनों के लिए जो धारणाऐं बनाते हैं वो एक समान नहीं होती है। बडे नमूनों के लिए अवधारणाऐं है-

1. आंकडों का यादृच्छिक नमूना वितरण लगभग सामान्य हो ।
2. नमूना मान पर्याप्त रूप से समग्र के करीब हो और इसका प्रयोग आंकलन की मानक त्रुटि की गणना में किया जा सके। बडे नमूनों की स्थिति में , जब हम आंकडों के महत्व का परीक्षण कर रहे हों तो मानक त्रुटि की संकल्पना प्रयोग हेाती है। विभिन्न आंकडों के लिए मानक त्रुटि ज्ञात करने के लिए सूत्र निम्नवत है।

माध्य की मानक त्रुटि - यह केवल नमूना त्रुटियों को मापता है। नमूना त्रुटि सिवाय समग्र के सभी भावात्मक जानकारी नमूने में से एक समग्र प्राचल के आंकलन में शामिल है।

(i) जब समग्र का मानक विचलन ज्ञात है, $S.E\bar{x} = \frac{\sigma_p}{\sqrt{n}}$ सूत्र है।

$S.E\bar{x} =$ माध्य की मानक त्रुटि

$\sigma_p =$ समग्र का मानक विचलन

$n =$ नमूनों में प्रेक्षणों की संख्या

(ii) जब समग्र का मानक विचलन अज्ञात है , हमें माध्य की मानक त्रुटि के लिए नमूने के मान विचलन सूत्र का प्रयोग करना है।

$S.E\bar{x} = \frac{\sigma}{\sqrt{n}}$, σ नमूने का मानक विचलन

यदि नमूने एवं समग्र का मानक विचलन दिया हुआ हो , तो माध्य की मानक त्रुटि की गणना के लिए हमें समग्र के मानक विचलन का प्रयोग करना चाहिए।

उदाहरण 1- निम्नलिखित आंकडे से माध्य के मानक त्रुटि की गणना करें , कोलकाता में दुर्गा पूजा के अवसर पर 100 फर्मों द्वारा देय राशि दिखाई गई है।

मध्य मान रू)0)	39	49	59	69	79	89	99
फर्म की संख्या	2	3	11	20	32	25	7

हल $S.E\bar{x} = \frac{\sigma}{\sqrt{n}}$

मध्य मान (m)	f	$\frac{m-69}{10}=d^1$	fd^1	fd^{1^2}
69	2	-3	-6	18
49	3	-2	-6	12
59	11	-1	-11	11
69	20	0	0	0
79	32	+1	+32	32
89	25	+2	+52	100
99	7	+3	+21	63
	100		$\sum fd^1=80$	$\sum fd^{1^2} = 236$

$$\sigma=\sqrt{\frac{\Sigma fd'^2}{N}-\left(\frac{\Sigma fd'}{N}\right)^2}\times C$$

$$=\sqrt{\frac{236}{100}-\left(\frac{80}{100}\right)^2}\times 10$$

$$=\sqrt{2.36-0.64}\times 10=\sqrt{1.72}\times 10$$

$$= 1.311 \times 10 = 13.11$$

$$S.E.\overline{X}=\frac{13.11}{\sqrt{100}}=\frac{13.11}{10}=1.311$$

अनुपातों का परिकल्पना परीक्षण-

अवधारणाऐं- अनुपात को शामिल करते हुए परिकल्पना परीक्षण में , हम नमूना वितरण के रूप में द्विपदीय वितरण का उपयोग करते है। हम द्विपदीय वितरण के सन्निकट के रूप में सामान्य वितरण का उपयोग कर सकते हैं जब तक np और nq कम से कम 5 हो।

उदाहरण 2- श्री X के पास एक हार्डवेयर स्टोर है और वह एक विशेष ब्रांड की कैंची बेचता है। वह उन सभी की तुलना करना चाहता है जो पूरे देश में बेचे गए हैं। वह अनुभव से जानता है कि पूरे देश में बेची गई कैंची के 15% को पहले वर्ष में मरम्मत की आवश्यकता होती है। उसने 120 ग्राहकों का नमूना लिया और पाया कि उनमें से केवल 22 को उन्हें खरीदने के पहले वर्ष में मरम्मत की आवश्यकता है। महत्व के 2% स्तर पर, क्या पर्याप्त प्रमाण हैं कि उसकी कैंची पूरी दुनिया में बेची गई उन लोगों की विश्वसनीयता में अलग होती है।

हल:- $n = 120, \quad H_0: p = 0.15, \qquad H_1: p \neq 0.15$माध्य

$p = \frac{22}{120} = 0.1833, \qquad \propto= 2\%$ या 0.02

दोनों आधे हिस्से के अंतर्गत स्वीकृत क्षेत्र 0.98 हैं और इसलिए एक आधे हिस्से के अंतर्गत स्वीकृत क्षेत्र सामान्य क्षेत्र का 0.4950 है।

यह Z का 2.33 (महत्वपूर्ण मान) देता है। इसलिए स्वीकृत क्षेत्र की सीमाऐं

$$\pm Z = \pm \frac{p - p_{H_0}}{\sqrt{\frac{p_{H_0} q_{H_0}}{n}}} = \frac{0.1833 - 0.15}{\sqrt{\frac{0.15 * 0.85}{120}}} = \pm 1.02$$

जैसा कि प्रेषित मान महत्वपूर्ण मान 2.33 से कम है। शून्य अवधारणा स्वीकार्य है। इसका अर्थ यह है कि श्री X की दुकान पर बेची जाने वाली कैंची पूरे देश में बेची जाने वाली की तुलना में विश्वसनीय नहीं है।

3.5 सारांश (Summary)

संख्यिकी में, परिणाम को सांख्यिकी सार्थक कहा जाता है यदि मौके पर ऐसा होने की संभावना न घटित हो। सांख्यिकी सार्थक मुहावरे का आविष्कार रोनार्ड फिशर ने किया था। जैसा कि सांख्यिकी में प्रयोग किया जाता है कि, सार्थक का अर्थ महत्वपूर्ण या अर्थपूर्ण नहीं है , जैसा कि रोजमर्रा की बातों में होता है। अनुसंधान विश्लेषकों का ध्यान केवल महत्वपूर्ण परिणामों पर केन्द्रित होता है जो महत्वपूर्ण प्रतिक्रिया तरीकों को दिखा सकते हैं। जो अलग अलग महत्व के परीक्षण के लिए थ्रेसहोल्ड व्यवस्थित कर सकते हैं। कई शोधकर्ताओं ने आग्रह किया कि महत्व के परीक्षणों को हमेंशा प्रभाव के आकार के आंकडों के साथ होना चाहिए, जो आकार का अनुमान लगाते हैं और इस प्रकार का अंतर व्यवहारिक महत्व है।

सक्ष्यों की मात्रा को स्वीकार करने के लिए यह आवश्यक है कि एक घटना अविश्वसनीय संयोग द्वारा उत्पन्न हुई है जिसे महत्व का स्तर या महत्वपूर्ण P मान कहते हैं। पारंपरिक सांख्यिकीय परिकल्पना परीक्षण में , P मान की प्रायिकता कम से कम चरम रूप में देखे गए आंकडों को देखते हुए कि शून्य परिकल्पना सही है। यदि ज्ञात किया हुआ P मान छोटा होता है तो यह कहा जा सकता है शून्य परिकल्पना या तो गलत है या एक असामान्य घटना घटित हुई है। P मान किसी दोहराए जाने वाले नमूनाकरण की व्याख्या नहीं करते हैं।

एक वैकल्पिक (फिर भी सम्बन्धित) सांख्यिकीय परिकल्पना परीक्षण संरचना नाइमन पीयरसन वारकृटिस्ट स्कूल है जिसमें दोनो परिकल्पनाऐं शून्य एवं वैकल्पिक परिभाषित करना चाहिए और जांचना चाहिए कि प्रक्रिया का नमूनाकरण गुण दोहराएं अर्थात् एक शून्य अवधारणा को अस्वीकार करने का निर्णय तब किया जायेगा जब यह वास्तव में सत्य हो और इसे अस्वीकार नहीं किया जाना चाहिए था (इसे गलत सकारात्मक या I प्रकार की त्रुटि कहा जाता है) और संभावना होती है कि कोई निर्णय शून्य परिकल्पना के स्वीकार के लिए किया जायेगा t यह वास्तव में गलत है (II प्रकार की त्रुटि) फिशरियन P मान दार्शनिक रूप से नेमन पियरसन के I प्रकार की त्रुटियों से भिन्न है।

3.6 शब्दावली (Glossary)

महत्व स्तर ()- साक्ष्यों की मात्रा को स्वीकार करने के लिए यह आवश्यक है कि एक घटना अविश्वसनीय संयोग द्वारा उत्पन्न हुई है।

नमूनाकरण त्रुटियां ()- नमूने से समग्र प्राचल आंकलन में नमूनाकरण त्रुटिया◌ॅ◌ं शामिल होती है।

3.7 बोध प्रश्न (Comprehension Questions)

1. एक केचप निर्माता यह तय करने की प्रक्रिया में है कि क्या केचप के एक नए अतिरिक्त मसालेदार ब्रांड का उत्पादन किया जाय। कम्पनी के बाजार शोध दल में पाया कि 6000 परिवारों में किये गये सर्वेक्षण में से 355 परिवार अतिरिक्त मसालेदार ब्रान्ड खरीदेंगें। दो साल पहले किये गये अधिक व्यापक

अध्ययन से पता चला है कि कब 5 प्रतिशत परिवार ब्राड खरीदेंगे। 2 प्रतिशत महत्व के स्तर पर, क्या कम्पनी को निष्कर्ष निकालना चाहिए कि कोई एक अतिरिक्त मसालेदार की रूचि में बढोतरी हुई है?

2. मुम्बई विश्वविद्यालय के 1000 छात्रों का नमूना लिया गया और उनका औसत वजन 112 lbs, 20 lbs के मानक विचलन के साथ पाया गया। क्या समग्र में छात्रों का औसत वजन 120 पाउन्ड हो सकता है।
3. विद्युत प्रकाश बल्बों का निर्माण करने वाली कम्पनी का दावा है कि उनके बल्बों का औसत जीवन 1600 घंटे हैं। 100 के इन बल्बों के यादृच्छिक नमूने का औसत जीवन और मानक विचलन क्रमशः 1570 घन्टे और 120 घन्टे था। क्या हमें कम्पनी के दावे को स्वीकार करना चाहिए?

3.8 बोध प्रश्नों के उत्तर (Answers to Comprehension Questions)

हल- n= 6000 p= 335/6000=0.05583

H_0: P= 0.05 H_1: P> 0.05 α=0.02

β=0.4800 के लिए स्वीकृत क्षेत्र की ऊपरी सीमा 0.4800 है इसके लिए Z का मान Z तालिका 2.05 है।

$$p = p_{Ho} + z\sqrt{\frac{p_{Ho}\,q_{Ho}}{n}} = 0.05 + 2.05\sqrt{\frac{0.05x0.95}{6000}} = 0.05577$$

क्योंकि अवलोकित मान P(0.05577) is > than P(0.05) की तुलना है हम नाममात्र μ_0 को अस्वीकार करते हैं।

3.9 स्वपरख प्रश्न (Self-Assessment Questions)

1. माध्य के लिए परीक्षण की व्याख्या कीजिए।
2. बडे नमूने से आप क्या समझते है?

3.10 सन्दर्भ पुस्तकें (Reference Books)

- बुनियादी सांख्यिकी, गौण, गुप्ता और दास गुप्ता - वर्ल्ड प्रेस लिमिटेड-कलकत्ता।
- व्यावसायिक सांख्यिकी के बुनियादी सिद्धान्त सांचेथी औश्र कपूर।
- प्रबन्ध में मात्रात्मक विधियां श्रीवास्तव, शेनाय और गुप्ता।
- व्यावसायिक सांख्यिकी - गुप्ता और गुप्ता।

इकाई 4 चरों का साथर्कता परीक्षण (छोटे प्रतिदर्श)
(Significance Test of Variables (Small Samples))

4.1 प्रस्तावना (Introduction)

4.2 उद्देश्य (Objectives)

4.3 स्टूडेन्ट का t वितरण (Student's t Distribution)

4.4 t परीक्षण (t Test)

4.5 काई -वर्ग परीक्षण (Chi-Square Test)

4.6 सारांश (Summary)

4.7 शब्दावली (Glossary)

4.8 बोध प्रश्न (Comprehension Questions)

4.9 बोध प्रश्नों के उत्तर (Answers to Comprehension Questions)

4.10 स्वपरख प्रश्न (Self-Assessment Questions)

4.11 सन्दर्भ पुस्तकें (Reference Books)

4.1 प्रस्तावना (Introduction)

यदि नमूना आकार 30 से कम है या n<30, तो उन नमूनों को छोटा नमूना समझा जा सकता है। यथाविधि, छोटे नमूनों के तरीके और छोटे नमूनों के सिद्धान्त बडे नमूनों पर लागू होते हैं , लेकिन बडे नमूनों के तरीके और बडे नमूनों के सिद्धान्त छोटे नमूनों पर लागू नहीं होते हैं। छोटे नमूनों का उपयोग एक अनुमानित अवधारणा को जांचने के लिए किया जाता है , जो प्रेक्षित मानों ज्ञात करने के लिए होता है , जो अग्रिम में दिए गए कुछ मानों के उतार चढाव नमूनाकरण के कारण उत्पन्न होता है। उदाहरण के लिए 12 के नमूने में यदि सहसम्बन्ध गुणांक +0.5 है, हम परीक्षण कर सकते हैं कि क्या सहसम्बन्ध का मान मूल समग्र में महत्वपूर्ण है।

छोटे नमूनों में , जांचकर्ताओं का अनुमान नमूनों से नमूनों तक व्यापक रूप से भिन्न होगा। छोटे नमूने परणिाम से निकला हुआ अनुमान, बडे नमूने परिणाम से कम सटीक होता है।

4.2 उद्देश्य (Objectives)

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

- छोटे नमूनों के लिए महत्व परीक्षण को प्रयोग कर सकें।
- महत्वपूर्ण परीक्षणों के द्वारा छोटे नमूनों के बीच दिए गए आंकडों का विश्लेषण कर सकें।
- t परीक्षण एवं काई वर्ग परीक्षण का वर्णन कर सकें।

4.3 स्टूडेन्ट का t वितरण (Student's t Distribution)

छोटे नमूनों के सिद्धान्त में सबसे बडा योगदान सर गोस्सेट और आर .ए. फिशर द्वारा किया गया था। गोस्सेट ने अपनी खोज 1905 में स्टूडेन्टेस उपनाम के तहत प्रकाशित की और इसे सामान्यतया t परीक्षण या स्टूडेन्ट्स का t वितरण या स्टूडेन्टस वितरण कहा जाता है।

जब नमूना आकार 30 या 30 से कम है और समग्र का मानक विचलन अज्ञात है , हम t वितरण का प्रयोग कर सकते हैं।

$t = (\frac{\bar{x}-\mu}{\sigma}) \times \sqrt{n}$ सूत्र है।

जहां $\sigma = \sqrt{\frac{\varepsilon(x-\bar{x})^2}{n-1}}$

सामान्य समग्र वितरण की अवधारणा के अंतर्गत t वितरण को गणितीय रूप में सिद्ध किया गया है या

$f(t) = c\ (1 + \frac{t^2}{v})^{-\frac{v+1}{2}}$

जहां $t = \frac{\bar{x}-\mu}{\sigma} \times \sqrt{n}$

c=एक स्थिरता को समानता के बराबर व्क के अंतर्गत क्षेत्र बनाने की आवश्यकता होती है।

$v = n - 1$, स्वतन्त्रता के दर्जे की संख्या उदा हरण - निम्नलिखित परिणाम विस्किट के 10 डिब्बों के नमूनों में से प्राप्त किये गये-

निहित वस्तु का माध्य वजन= 490 ग्राम

बजन का मानक विचलन = 9 ग्राम

क्या नमूना आबादी में से आ सकता है जिसका माध्य 500 ग्राम है।

परिकल्पना लें कि $\mu = 500$ ग्राम

$$t=\frac{\overline{X}-\mu}{\sigma}\sqrt{n}$$

$\overline{X}=490;\ \mu=500;\ \sigma=9;\ n=10$

$$t=\frac{490-500}{9}\sqrt{10}$$

df=10-1= 9

$$=\frac{10}{9}\sqrt{10}$$

$$=\frac{10}{9}\times 3.16=\frac{31.6}{9}=3.51$$

df= 9, $t_{0.01}$= 3.25

$3.51 > 3.25$, हमारी परिकल्पना अस्वीकार्य है।

उदाहरण - एक आपरेटर दाबा करता है कि वह एक घंटे में 40 लेख तैयार करता है। 10 यादृच्छिक घंटों का नमूना दिखाता है कि 43, 45, 38, 37, 41, 42, 44, 39, 43 और 38 लेख तैयार होते है। क्या आपरेटर का दावा 5 प्रतिशत महत्व के स्तर पर उचित है। आपरेटर के घंटे में तैयार लेख का वितरण सामान्य है और 9 degree of freedom के लिए 5 प्रतिशत महत्वपूर्ण स्तर पर एक पुच्छीय परीक्षण का मान 1.833 लें।

हल- शून्य परिकल्पनाः आपरेटर का औसत निर्माण 40 है। शून्य परिकल्पना के विरूद्ध वैकल्पिक परिकल्पना है औसत निर्माण 40 से कम है।

$$\overline{X}=\frac{1}{10}(43+45+38+37+41+42+44+39+43+38)=41$$

$$\delta^2=\frac{1}{9}(43-41)^2+(45-41)^2+(38-41)^2+(37-41)^2+(41-41)^2+$$

$$(42-41)^2+(44-41)^2+(39-41)^2+(43-41)^2+(38-41)^2$$

$$=\frac{1}{9}\times 72=8$$

$$\delta=\sqrt{8}=2.8284$$

$$t=\frac{\sqrt{X}-\mu}{\delta}\sqrt{n}=\frac{41-40}{2.8284}\times\sqrt{10}$$

= 1.118 <1.833

इसलिए अन्तर महत्वपूर्ण नहीं हैं इसलिए यह संयोग होने के कारण उत्पन्न हो सका। अवधारण को अस्वीकार्य नहीं किया जा सकता है। आपरेटर की मांग तर्क संगत है।

10.4 t- परीक्षण (t- Test)

एक परिकल्पना परीक्षण आंकडों के नमूने वितरण पर आधारित होता है। और इसलिए एक महत्वपूर्ण खेत्र के परिभाषित करने के लिए नमूना वितरण का ज्ञात होना आवश्यक है। सामान्य वितरण $N(\mu,\sigma^2)$ से

यादृच्छिक नमूना आकार n के लिए, नमूना माध्य $\bar{x}$ माध्य μ मानक विचलन $\frac{\sigma}{\sqrt{n}}$ के साथ सामान्यता वितरित है।

इसलिए $Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ $N(0,1)$ हैं और $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}$ स्टूडेन्ट $t(n-1)$ के साथ चर है।

$N(\mu,\sigma^2)$ से दो क्रमशः लिये गए यादृच्छिक नमूनों n_1 और n_2 आकार के और

$N(\mu,\sigma^2)$ समग्रों $\frac{\bar{x}_1-\bar{x}_2}{\sqrt{\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}\left[\frac{n_1+n_2}{n_1 n_2}\right]}}$

स्टूडेन्ट t चर $(n_1+n_2-2)d.f$ के साथ है।

जब समग्र मानक विचलन अज्ञात है, उपरोक्त t चर को परिकल्पना परीक्षण के लिए प्रयोग करते हैं यद्यपि बडे नमूनों के लिए, t लगभग सामान्य है।

मध्य के लिए छोटे नमूने

मान लें कि $N(\mu,\sigma^2)$ समग्र में, μ और σ दोनों अज्ञात हैं

हम परीक्षण करना चाहते हैं कि क्या माध्य दिया हुआ मान μऋ0 है।

शून्य परिकल्पना $\mu_0 = \mu = \mu_0$ (समग्र माध्य μ_0 है)

आकार n के यादृच्छिक नमूने के लिए , μ_0के अंतर्गत, आंकडा परीक्षण $t = \frac{x-\mu_0}{\frac{\sigma}{\sqrt{n-1}}}$ एक स्टूडेन्ट का t चर (n-1)d.f के साथ है।

निम्न में से कोई एक वैकल्पिक परिकल्पना हो सकती है।

1. $\mu_1 : \mu \neq \mu_0$यहां t परीक्षण दो पुच्छीय है।
2. $\mu_1 : \mu > \mu_0$ यहां परीक्षण महत्वपूर्ण के साथ ऊपरी एक पुच्छीय है।
3. $\mu_1 : \mu > \mu_0$ यहां परीक्षण महत्वपूर्ण के साथ निचला पुच्छीय है।

द्विपुच्छीय परीक्षण की स्थिति में, यदि $\propto$ महत्व का स्तर है, महत्वपूर्ण मान $-t_{\propto/2}$और $t_{\propto/2}$है।

ऊपरी पुच्छीय परीक्षण में, महत्वपूर्ण मान $t_{\propto}$ है।

निचले पुच्छीय परीक्षण में, महत्वपूर्ण मान $-t_{\propto}$ है।

विभिन्न $\propto$ के महत्वपूर्ण मानों के लिए और विभिन्न degree of freedom (स्वतन्त्रता के विभिन्न स्तरों के लिए) t तालिका मान पुस्तक के अंत में प्रदान की गई t वितरण से प्राप्त की जाती है।

ध्यान दें- यह परीक्षण इस धारणा पर आधारित है कि समग्र सामान्य है।

माध्यों की समानता के लिए छोटा नमूना परीक्षण

मान लें कि दो समग्र $N(\mu_1,\sigma^2)$ और $N(\mu_2,\sigma^2)$ अज्ञात μ_1 और μ_2 और σ^2 के साथ है।

हम परीक्षण करना चाहते हैं कि μ_1 और μ_2 समान है।

शून्य परिकल्पना $H_0 = \mu_1 = \mu_2$ (समग्र माध्य समान है।)

इन समग्रों में से, यादृच्छिक नमूना आकार n_1 और n_2 के लिए, आंकडा परीक्षण

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}}$$

$(n_1 + n_2 - 2)d.f$ के साथ स्टूडेन्ट t परीक्षण है।
वैकल्पिक परिकलपना निम्न में से कोई एक हो सकती है।

1. $\mu_1 : \mu \neq \mu_0$ यहां ज परीक्षण दो पुच्छीय है।
2. $\mu_1 : \mu_1 > \mu_2$ यहां परीक्षण एक पुच्छीय महत्वपूर्ण क्षेत्र के साथ ऊपरी पुच्छीय है।
3. $\mu_1 : \mu_1 > \mu_2$ यहां परीक्षण एक पुच्छीय महत्वपूर्ण क्षेत्र के साथ निचला पुच्छीय है।

दो पुच्छीय परीक्षण की स्थिति में, यदि $\propto$ महत्व का स्तर है, महत्वपूर्ण क्षेत्र $-t{\propto}/_2$ और $t{\propto}/_2$है।

ऊपरी पुच्छीय परीक्षण की स्थिति में, महत्वपूर्ण मान $t_{\propto}$ है।

निचले पुच्छीय परीक्षण की स्थिति में, महत्वपूर्ण मान $-t_{\propto}$है।

विभिन्न $\propto$ के मानों के लिए और विभिन्न स्वतंत्रता के स्तरों के लिए t वितरण का मान तालिका से प्राप्त किया जाता है।

ध्यान दें- 1. यह परीक्षण इन अवधारणों पर आधारित है

(अ) समग्र सामान्य है।

(ब) समग्र मानक विचलन समान है (अज्ञात)

ध्यान दें- 2 युगल अवलोकन की स्थिति में n यादृच्छिक युग्मों के साथ आंकडा परीक्षण $t = \frac{\bar{d}}{\frac{Sd}{\sqrt{n-1}}}$ स्टूडेन्ट t चर (n-1)d.f के साथ है।

यहां, d युग्मों के अवलोकनों के बीच अन्तर हैं माध्यों की समानता के लिए परीक्षण जब प्रेक्षण युग्मित है (पेयर टी परीक्षण, अश्रित नमूने) दो चरों के सामान्य समग्र में, यदि समग्र में दो चर इकाई विशेषताऐं x और y जो कि $N(\mu_2, \sigma^2)$, और $N(\mu_1, {\sigma_1}^2)$ क्रमशः है। उदारहण के लिए,

1.उच्च रक्तचाप, के लिए योग उपचार से गुजरने वाले मरीजों में रक्तचाप के दो माप होत हैं - पहला-उपचार से पहले (x) और दूसरा उपचार के बाद किसी परीक्षा में।

2.कोंचिंग कक्षाओं में किसी परीक्षा में अच्छे अंक प्राप्त करने के लिए भाग लेने वाले छात्रों का कोंचिग से पहले (x) कोंचिग के बाद (y)

इस तरह की परिस्थितियों में , मान लें कि हम परीक्षण करना चाहते हैं कि क्या माध्य μ_1 और μ_2 एकसमान है। शून्य परिकल्पना

$H_0 = \mu_1 = \mu_2$ है (माध्य एक समान है)

n यादृच्छिक युग्म प्रेक्षणों के लिए

$(x_1 y_1), (x_2 y_2) \dots \dots \dots . . (x_n y_n)$, यदि $d_1 = x_i - y_i$ होगा। यदि $\bar{d}$ नमूना माध्य और Sd उन विचलनों का नमूना माध्य विचलन है। तो, H_0 के अंतर्गत आंकडा t परीक्षण $t = \frac{\bar{d}}{\frac{Sd}{\sqrt{n-1}}}$ स्टूडेन्ट t चर (n-1)d.f के साथ है।

वैकल्पिक परिकल्पना निम्न में से कोई एक हो सकती है।

1. $H_1 : \mu_1 \neq \mu_2$यहां परीक्षण दो पुच्छीय है।

2. $H_1: \mu_1 > \mu_2$ यहां परीक्षण एक पुच्छीय महत्वपूर्ण क्षेत्र के साथ ऊपरी पुच्छीय है।
3. $H_1: \mu_1 < \mu_2$ यहां परीक्षण एक पुच्छीय महत्वपूर्ण क्षेत्र के साथ निचला पुच्छीय है।

द्विपुच्छीय परीक्षण की स्थिति में , यदि ∝ महत्व का स्तर है , महत्वपूर्ण मान $-t_{\propto/2}$ और $t_{\propto/2}$ है। ऊपरी पुच्छीय परीक्षण की स्थिति में महत्वपूर्ण $t_{\propto}$ है।

निचले पुच्छीय परीक्षण की स्थिति में, महत्वपूर्ण मान $-t_{\propto}$ है। विभिन्न ∝ के महत्वपूर्ण और विभिन्न स्वतंत्रता के स्तरों के लिए t वितरण का मान तालिका से प्राप्त किया जाता है।

उदाहरण- आठ यादृच्छिक दिनों पर , कालेज तक पहुंचने के लिए शहर की बस से लिया गया समय नीचे दिखाया गया है। इस परिकल्पना का परिक्षण करें कि बस से कालेज पहुंचने के लिए औसत समय 30 मिनट है।

दिन	:	1	2	3	4	5	6	7	8
समय (मिनट)	:	27	34	30	35	31	30	29	32

हल- $\mu_0 = 30$ मिनट और n=8 नमूना छोटा और ∝ अज्ञात हैं सामान्य वितरण मानते हुए हम स्टूडेन्ट t परीक्षण का प्रयोग करते हैं वैकल्पिक परकिल्पना $\mu_0 : \mu = 30$ है (माध्य समय 30 मिनट है)

H_0 के अंतर्गत t परीक्षण $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n-1}}}$ स्टूडेन्ट का t परीक्षण (चर) (n-1) के साथ हैं = 8-1=7D.f यहां परीक्षण द्विपुच्छीय है।

5% महत्व के स्तर पर 7D.f के लिए महत्वपूर्ण मान $-t_{\propto/2} = -2.37$ और $t_{\propto/2} = 2.37$ है।

$Time(x)$	x^2
27	729
34	1156
30	900
35	1225
31	961
30	900
29	841
32	1024
248	7736

$$\bar{x} = \frac{\sum x}{n} = \frac{248}{8} = 31$$

$$s = \sqrt{\frac{\sum x^2}{n} - \left[\frac{\sum x}{n}\right]^2} = \sqrt{\frac{7736}{8} - \left[\frac{248}{8}\right]^2} = 2.4495$$

$$t_{0bs} = \frac{\bar{x} - \mu_0}{s / \sqrt{n-1}} = \frac{31-30}{2.4495/\sqrt{8-1}} = 1.08$$

क्योंकि $t_{obs} = 1.08$ अंतराल $(-2.37, 2.37)$ में है, H_0 स्वीकार्य है।

निष्कर्ष- बस से कालेज पहुंचने का माध्य समय 30 मिनट है।

उदाहरण- एक फैक्ट्री का प्रबंधन इस बात का तर्क करता है कि कारखाने में औसत ध्वनि की तीव्रता 120 डेसिबल से कम है। 23 यादृच्छिक मापों में ध्वनि की तीव्रता 117 डेसिबल और मानक विचलन 8 डेसीबल था। प्रबंधन का तर्क स्वीकार है या नहीं, इसके लिए 1% स्तर के महत्व पर परीक्षण करें।

हल: यहां $\mu_0 = 120$ डेसीबल और $n = 23$ $\bar{x} = 117$डेसीबल

$s = 8$ डेसीबल और $\propto = 1\%$ 0.01

चूंकि नमूना छोटा है और σ अज्ञात है, ध्वनि की तीव्रता को सामान्य वितरण मानते हुए हम स्टूडेन्ट के t परीक्षण का उपयोग करते है। **शून्य परिकल्पना** $\mu_0 : \mu = 120$ है (औसत ध्वनि तीव्रता 120 डेसीबल से कम हैं)

μ_0 के अंतर्गत, आंकडा परीक्षण $t = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n-1}}}$

एक स्टूडेन्ट t के (n-1)=23-1= 22 d.f है।

वैकल्पिक परिकल्पना

$\mu_1 : \mu < 120$ है (औसत ध्वनि तीव्रता 120 डेसीबल से कम नहीं हैं)

परीक्षण निचला पुच्छीय है।

यहां परीक्षण एक पुच्छीय है।

22 d.f के लिए, 1% महत्व के स्तर पर महत्वपूर्ण मान $-t_{\propto} = -2.51$ है।

$$t_{obs} = \frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n-1}}} = \frac{117-120}{\frac{8}{\sqrt{23-1}}} = -1.76$$

चूंकि -1.76 -2.51 से कम नहीं है, H_0 स्वीकार्य है।

निष्कर्ष- औसत ध्वनि तीव्रता 120 डेसीबल है और इससे कम नहीं हैं

उदाहरण- 5 मादा बिल्लियों और 8 नर बिल्लियों के दिलों का वजन ग्राम में नीचे दिया गया है-

मादा बिल्ली	:	7.5	7.3	7.1	9.0	7.6			
नर बिल्ली	:	12.7	15.6	9.1	12.8	8.3	11.2	9.4	8.2

1% महत्व के स्तर पर परीक्षण करें कि नर बिल्लियों का वनज मादा बिल्लियों से अधिक है।

हल-यहां $n_1 = 5, \quad n_2 = 8, \quad \propto = 0.01$

शून्य परिकल्पना $\mu_0 : \mu_1 = \mu_2$ है (माध्य वनज एक समान है)

H_0 के अंतर्गत आंकडा परीक्षण

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\left[\frac{n_1+n_2}{n_1 n_2}\right]}}$$

स्टूडेन्ट t का चर $n_1 + n_2 - 2 = (5+8-2) = 11 d.f$ के साथ है।

वैकल्पिक परिकल्पना $H_1: \mu_1 < \mu_2$ है।
(नर बिल्ली के दिल का वजन मादा बिल्ली के वजन से ज्यादा है)
1% महत्व के स्तर पर, 11d.f के लिए, महत्वपूर्ण मान $-t_{\propto/2} = -2.72$ है।

मादा बिल्ली		नर बिल्ली	
x_1	x_1^2	x_2	x_2^2
7.5	56.25	12.7	161.29
7.3	53.29	15.6	243.36
7.1	50.41	9.1	82.81
9.0	81.00	12.8	163.84
7.6	57.76	8.3	68.89
		11.2	125.44
		9.4	88.36
		8.2	67.24
38.5	298.71	87.3	1001.23

मादा बिल्लीयां:

$$\bar{x}_1 = \frac{\sum_1 x}{n_1} = \frac{38.5}{5} = 7.7$$

$$s_1^2 = \frac{\sum_1 x^2}{n_1} - \left[\frac{\sum_1 x}{n_1}\right]^2 = \frac{298.71}{5} - \left[\frac{38.5}{5}\right]^2 = 0.452$$

नर बिल्लीयां

$$\bar{x}_2 = \frac{\sum_2 x}{n_2} = \frac{87.3}{8} = 10.91$$

$$s_1^2 = \frac{\sum_2 x^2}{n_2} - \left[\frac{\sum_2 x}{n_2}\right]^2 = \frac{1001.23}{8} - \left[\frac{87.3}{8}\right]^2 = 6.12$$

$$t_{0bs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}} = \frac{7.7 - 10.1}{\sqrt{\frac{5 \times 0.452 + 8 \times 6.12}{5 + 8 - 2}\left[\frac{5+8}{5 \times 8}\right]}} = -2.61$$

चूंकि $t_{obs} = -2.61, -2.72$ से कम नहीं है, H_0 स्वीकार्य है।

निष्कर्ष- नर बिल्लियों के दिलों का वजन , मादा बिल्लियों के वजन के एक समान है। साक्ष्य यह निष्कर्ष निकालने के लिए प्र्याप्त है कि पुरूष बिल्लियों के दिलों का वजन महिला (मादा) बिल्लीयों की तुलना में अधिक है।

उदाहरण- एसएसएलसी की कक्षा में बेतरतीब ढंग से चुने गए लडके व लडकियों की ऊँचाईयों के बारे में निम्नलिखित आंकडे बताते हैं कि एसएसएलसी के लडके औसतन एसएसएलसी लडकियों से लम्बे हैं।

	लडके	लडकियां

नमूना आकार	9	12
औसत ऊँचाई (सेमी)	171	169
मानक विचलन (सेमी)	3	2

हल- यहां $n_1 = 9, n_2 = 12 , \bar{x}_1 = 171, \bar{x}_2 = 169, s_1 = 3 , s_2 = 2$

शून्य परिकल्पना $H_0 = \mu_1 = \mu_2$ है (लडके एवं लडकियों की औसत ऊँचाई एक समान है)

H_0 के अंतर्गत आंकडा परीक्षण

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}}$$

स्टूडेन्ट का t चर

$n_1 + n_2 - 2 = (9 + 12 - 2) = 19 d.f$ के साथ है।

वैकल्पिक परिकल्पना $H_1 : \mu_1 > \mu_2$ (औसतन लडके लडकियों की तुलना में लम्बे हैं)

परीक्षण ऊपरी पुच्छीय है।

19 d.f के लिए 5% महत्व के स्तर पर महत्वपूर्ण मान $t_\propto = 1.73$ है।

$$t_{0bs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}}$$

$$\frac{171 - 169}{\sqrt{\frac{9 \times 3^2 + 12 \times 2^2}{9 + 12 - 2}\left[\frac{9 + 12}{9 \times 12}\right]}} = 1.74$$

चूंकि $t_{obs} = -1.74$, 1.73 से बडा है, H_0 अस्वीकार्य है।

निष्कर्ष- लडके, औसतन, लडकियेां की तुलना में लम्बे हैं।

उदाहरण- एक मवेशी चारा के निर्माता ढावा करते हैं कि गायों के ज्यादा चारा देने से अधिक दूध मिलता है। आने ढावे को सही साबित करने के लिए निम्न गाय प्रयोगों का आयोजन किया गया।

(अ) 6 गायों को सामान्य चारा खिलाया गया और 8 गायों को निर्माता का चारा खिलाया गया था। 6 गायों की माध्य उपज 9.7 लीटर और मानक विचलन 1.3 लीटर थी। 8 गायों की माध्य उपज 10.5 लीटर औ र मानक विचलन 2.7 लीटर थी।

(ब) 8 गायों को निर्माता का चारा खिलाया गया था। दूध का उत्पादन लीटर में तब होता है जब वे हमेंशा के तरह चारा लेते थे और नीचे दिये हुए निर्माता के बारे मं अन्तर चारा भी लेते थे।

सामान्य चारा	:	6.3	7.4	9.7	12.4	11.1	10.4	9.6	7.1
निर्माता का चारा	:	7.4	7.2	14.6	13.6	10.5	11.6	10.4	8.7

उपरोक्त सभी मामलों, सत्यापित करें कि क्या निर्माता का ढावा सही है।

हल- मामले (अ) में , गायों के दो स्वतन्त्र नमूने हैं जिसमें गायों को सामान्य चारा और निर्माता का चारा खिलाया जाता है। इसलिए, स्वतन्त्र नमूने के लिए t परीक्षण का प्रयोग किया जाता है। मामले (ब) में , गायों के

एकसमान समुच्चय के अंतर्गत सामान्य चारा निर्माता का चारा को प्रेक्षित किया जाता है। यहां आंकडे व्यक्तिगत मामलों में उपज में परिवर्तन दर्शाते हैं। इसलिए t युग्मित परीक्षण प्रयोग किया गया है।

यहां $n_1 = 6, n_2 = 8, \bar{x}_1 = 9.7\ , \bar{x}_2 = 10.5\ , s_1 = 1.3,\ \ s_2 = 2.7$

शून्य परिकल्पना $H_0: \mu_1: \mu_2$ हैं (माध्य एकसमान H_0 के अंतर्गत आंकडा परीक्षण)

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}}$$

स्टूडेन्ट का t चर $n_1 + n_2 - 2 = (6 + 8 - 2) = 12 d.f$ हैं (निर्माण का चारा ज्यादा अच्छा है) परीक्षण निचला पुच्छीय है।

12 d.f के लिए 5% महत्व के स्तर पर, महत्वपूर्ण मान $-t_\propto = 1.78$ है।

$$t_{0bs} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left[\frac{n_1 + n_2}{n_1 n_2}\right]}}$$

$$\frac{9.7 - 10.5}{\sqrt{\frac{6 \times (13.3)^2 + 8 \times (2.7)^2}{6 + 8 - 2}\left[\frac{6+8}{6 \times 8}\right]}} = 0.62$$

चूंकि $t_{obs} = -0.62\ , -1.78$ से कम नहीं है, H_0 स्वीकार्य है।

निष्कर्ष- माध्य एक समान है। (कोई साक्ष्य नहीं है कि निर्माता का चारा बेहतर उपज है)

(ब) यहां, n=8, मान लें d=x-y है जहां x सामान्य चरों में उपज है। और y निर्माता के चारें में उपज है।

H_0 के अंतर्गत $t = \frac{\bar{d}}{\frac{sd}{\sqrt{n-1}}}$ स्टूडेन्ट का चर

(n-1)= (8-1)= 7 के साथ है।

(अ) के मामले में शून्य एवं वैकल्पिक परिकल्पनाऐं एकसमान हैं, परीक्षण निचला पुच्छीय है।

5% महत्व के स्तर पर, 7 d.f के लिए, महत्वपूर्ण मान $-t_\propto = -1.90$ है।

x	y	$d = x - y$	d^2
6.3	7.4	-1.1	1.21
7.4	7.2	0.2	0.04
9.7	14.6	-4.9	24.01
12.4	13.6	-1.2	1.44
11.1	10.5	0.6	0.36
10.4	11.6	-1.2	1.44
9.6	10.4	-0.8	0.64
7.1	8.7	-1.6	2.56
-	-	-10.0	31.70

$$\bar{d} = \frac{\sum d}{8} = \frac{-100}{8} = -1.25$$

$$s_d = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2} = \sqrt{\frac{31.70}{8} - \left[\frac{-10}{8}\right]^2} = 1.55$$

$$t_{0bs} = \frac{\bar{d}}{s_d / \sqrt{n-1}} = \frac{-1.25}{1.55 / \sqrt{8-1}} = 2.13$$

चूंकि $t_{obs} = -2.13 , -1.90$ से कम नहीं है, H_0 अस्वीकार्य है।

निष्कर्ष- निर्माता के दावे के लिए समर्थन दिया गया है कि गायों को अपने चारे से अधिक देने पर अधिक दूध मिलता है।

उदाहरण- CET के लिए एक कोचिंग क्लास है। 10 यादृच्छिक रूप से चयनित छात्रों को कोचिंग से पहले एक परीक्षा दिलाई गई थी और उनहें कोचिंग के बाद भी एक परीक्षा दिलाई गयी थी। परीक्षण स्कोर निम्नानुसार है।

कोंचिग से पहले	:	35	39	47	53	27	19	36	46	08	17
कोचिंग के बाद	:	41	37	45	56	31	21	47	41	05	12

क्या हम निष्कर्ष निकाल सकते हैं कि कोचिंग प्रभावी है।

हल- यहां, कोचिंग से पहले एवं कोचिंग के बाद के अंकों का युग्म बनाया जा सकता है और इसलिए युग्मित t परीक्षण प्रयोग किया जायेगा।

मान लें x कोचिंग से पहले अंकों को प्रदर्शित करते हैं और y कोचिंग के बाद अंकों को प्रदर्शित करते है। **शून्य परिकल्पना** $\mu_0 : \mu_1 = \mu_2$ है।

(माध्य एक समान है)

H_0 के अंतर्गत t आंकडा परीक्षण $t = \frac{\bar{d}}{\frac{sd}{\sqrt{n-1}}}$ स्टूडेन्ट का t परीक्षण

$(n-1) = 10 - 1 = 9\ d.f$ के साथ है।

वैकल्पिक परिकल्पना $H_1 : \mu_1 < \mu_2$ है। (माध्य बढा हुआ है कोचिंग प्रभावशाली है।)

परीक्षण निचला पुच्छीय है।

9 d.f के लिए 5% महत्व के स्तर पर, महत्वपूर्ण मान $-t_\propto = -1.83$ है।

x	y	$d = x - y$	d^2
35	41	-6	36
39	37	2	4
47	45	2	4
53	56	-3	9
27	31	-4	16
19	21	-2	4
36	47	-11	121
46	4	5	25

08 17	05 12	3 5	9 25
-	-	-09	253

$$\bar{d} = \frac{\sum d}{n} = \frac{-9}{10} = -0.9$$

$$s_d = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2} = \sqrt{\frac{253}{10} - \left[\frac{-9}{10}\right]^2} = 4.95$$

$$t_{0bs} = \frac{\bar{d}}{s_d / \sqrt{n-1}} = 4.95 \frac{-0.9}{1.55 / \sqrt{10-1}} = 0.55$$

चूंकि $t_{obs} = -0.55, -1.83$ से कम नहीं है, H_0 स्वीकार्य है।

निष्कर्ष - कोचिंग प्रभावशाली नहीं है।

उदाहरण- 7 पतियों और उनकी पत्नियों के वजन किलोग्राम में निम्नवत हैं।

जोडे	:	1	2	3	4	5	6	7
पति	:	62	56	59	73	49	54	67
पत्नी	:	55	61	62	68	52	51	62

परिकल्पना का परीक्षण करें कि पतियों का माध्य वजन एवं पत्नीयों का माध्य वजन एक समाज है।

हल- यहां हमारें पास n=7 प्रक्षेणों का युग्म और इसलिए, हम t युग्म परीक्षण का प्रयोग करते हैं मान लें x पति का वजन एवं y पत्नी का वजन है। मान लें $d = x - y$ विचलन है।

शून्य परिकल्पना $H_0 = \mu_1 = \mu_2$ है (माध्य वजन एक समान है।)

H_0 के अंतर्गत, t आंकडा परीक्षण $t = \frac{\bar{d}}{\frac{sd}{\sqrt{n-1}}}$ स्टूडेन्ट का t चर $(n-1) = 7 - 1 = 6\ d.f$ के साथ है।

वैकल्पिक परिकल्पना $H_1 : \mu_1 \neq \mu_2$ है।

(माध्य एक समान है)

परीक्षण दो पुच्छीय है।

5% महत्व के स्तर पर, 6d.f के लिए, महत्वपूर्ण मान $-t_{\propto/2} = -2.45$ और $t_{\propto/2} = 2.45$ है।

x	y	$d = x - y$	d^2
62 56 59 73 49 54 67	55 61 62 68 52 51 62	7 -5 -3 5 -3 3 5	49 25 9 25 9 9 25
-	-	9	151

$$\bar{d} = \frac{\sum d}{n} = \frac{9}{17} = 1.29$$

$$s_d = \sqrt{\frac{\sum d^2}{n} - \left[\frac{\sum d}{n}\right]^2} = \sqrt{\frac{151}{7} - \left[\frac{9}{7}\right]^2} = 4.46$$

$$t_{0bs} = \frac{\bar{d}}{s_d / \sqrt{n-1}} = \frac{+1.29}{4.46/\sqrt{7-1}} = 0.71$$

चूंकि 0.71 अंतराल (-2.45 , 2.45) के बीच है, H_0 स्वीकार्य है।

निष्कर्ष- पतियों का माध्य वजन एवं पत्नियों का माध्य वजन एक समान है।

4.5 काई -वर्ग परीक्षण (x^2)(Chi-Square Test)

अनुमानित आंकडों के वितरण को सत्यापित करने के लिए काई वर्ग का प्रयोग सांख्यिकी में गुणों के तर्क संगत के लिए किया जाता है इसलिए, वास्तविक और अपेक्षित आवृत्तियों के विचलन का अध्ययन करने के लिए यह एक उपाय है।

नमूने के अध्ययन में विशेष रूप से आंकडों में इसका और अपेक्षित आवृत्तियों के बीच दो गुना संयोग और नमूने में उतार-चढाव के कारण अंतर को अनदेखा किया जा सकता है। यदि वास्तविक और अपेक्षित आवृत्तियों के बीच कोई अंतर नहीं है x^2 शून्य होता है। इस प्रकार , काई वर्ग परीक्षण सिद्धान्त और अवलोकन के बीच विसंगति का वर्णन करती है।

x^2 परीक्षण की विशेषताऐं

1. परीक्षण आवृत्तियों की घटनाओं पर आधारित है , जबकि सैद्धान्तिक वितरण में , परीक्षण माध्य और मानक विचलन पर आधारित है।
2. निष्कर्ष निकालने के लिए, यह परीक्षण विशेष रूप से परिकल्पना परीक्षण के लिए प्रयोग किया जाता है लेकिन अनुमान के लिए उपयोगी नहीं है।
3. परीक्षण अवलोकन की पूर्ण स्थित एवं अपेक्षित
4. आवृत्तियों के बीच प्रयोग किया जा सकता है।
5. स्वतन्त्रता की श्रेण संख्या में हर वृद्धि के लिए, एक नया x^2 वितरण का गठन होता है।
6. यह एक सामान्य प्रयोजन परीक्षण है और जैसा कि अनुसंधान में बेहद उपयोगी है।

मान्यताऐं-

1. सभी प्रेक्षण स्वतन्त्र होने चाहिए।
2. सभी घटनाऐं परस्पर अनन्य होने चाहिए ।
3. बडे अवलोकन (प्रेक्षण) होने चाहिए।
4. तुलना प्रयोजनों के लिए आंकडे मूल इकाईयों में होने चाहिए।

स्वतन्त्रता की श्रेणी- जब हम x^2 की गणना मान की तुलना तालिका मान के साथ करते हैं। स्वतन्त्रता की श्रेणी होती है। स्वतन्त्रता की श्रेणी का अर्थ है वर्गों की संख्या , जिनके मानों को बिना प्रतिबंध के बिना बंदिशों में सौंपा जा सकता है। उदारहण के लिए हम कोई भी चार अंक चुनते हैं , जिनका योग 50 है। यहां हमारे पास कोई भी तीन अंक के चयन करने का विकल्प है 10 , 15 , 20 और चौथा अंक $[50 - (10 + 15 + 20)]$। इस प्रकार, हमारी आजादी के श्रेणी की पसंद इस शर्त पर एक करके कम हो जाती है कि योग 50 हो। इसलिए

स्वतन्त्रता पर लगा प्रतिबंध एक और स्वतन्त्रता की श्रेणी तीन है। जैसे ही प्रतिबंध बढता है स्वतन्त्रता की श्रेणी कम हो जाती है।

इस प्रकार $v = n - k$

v : (न्यू) स्वतन्त्रता की श्रेणी

k : स्वतन्त्र प्रतिबंधों की संख्या

n : आवृत्ति वर्गोें की संख्या

2×2 आकास्मिक तालिका के लिए, स्वतन्त्रता की श्रेणी

$$v = (c-1)(r-1)$$

$= (2-1)(2-1) = 1$ है।

उपयोग -

1. **गुणों के तर्क संगत का x^2 परीक्षण-** परीक्षण के माध्यम से हम प्रेक्षित मानों और अपेक्षित मानों के बीच के विचलन का पता लगा सकते है। यहां हम प्राचलों से संबन्धित नहीं है। लेकिन वितरण के रूप में सम्बन्धित है। कार्ल पियर्सन ने सैद्धान्तिक मान (परिकलपना) और प्रेक्षित मान के बीच अन्तर का परीक्षण करने के लिए एक विधि विकसित की है। परीक्षण गणना मान का x^2 के तालिका मान के वांक्षित स्वतन्त्रता श्रेणी के साथ तुलना करने से किया जाता है । x^2 ग्रीक शब्द का प्रयोग तथ्य औश्र सिद्धान्त के बीच के अंतर के परिणाम का वर्णन करने के लिए किया जाता है।

x^2 को इस तरीके से परिभाषित किया जाता है।

$$\chi^2 = \Sigma\left\{\frac{(O-E)^2}{E}\right\}$$

O= प्रेक्षित आवृत्तियां

E= अपेक्षित अवृत्तियां

चरण-

1. एक परिकल्पना महत्व के स्तर के साथ स्थापित की गई हैं ।
2. प्रेक्षित मान और अपेक्षित मान के बीच विचलनों की गणना (O-E)।
3. गणित विचलनों का वर्ग $(O-E)^2$
4. $(O-E)^2$ को इसकी अपेक्षित आवृत्तियों से विभाजित करें।
5. चरण 4 से प्राप्त मानों का योग करें।
6. x^2 तालिका में से एक निश्चित महत्व के स्तर पर , सामान्यतया 5% महत्व के स्तर पर ग ‘2 का मान ज्ञात करें।

यदि x^2का परिकलित मान x^2 के तालिका मान से एक निश्चित महत्व के स्तर पर ज्यादा है , हम परिकल्पना को अस्वीकार्य करते हैं यदि x^2 का परिकलित मान शून्य है तब प्रेक्षित मान और अपेक्षित मान पूर्णतया मेंल खाते हैं। यदि x^2 का परिकलित मान पूर्णतया मेंल खाते हैं यदि x^2 का परिकलित मान तालिका मान से एक निश्चित महत्व के स्तर पर कम है, तो यह महत्वपूर्ण नहीं है। इसका अर्थ है कि प्रेक्षित और अपेक्षित आवृत्तियों के बीच नमूनाकरण में उतार चढाव के कारण विसंगतियां हो सकती है।

उदाहरण- 4 सिक्के 160 बार उछाले गये थे और निम्नलिखित परिणाम प्राप्त किये गये थे।

चिट्टों की संख्या	:	0	1	2	3	4
प्रेक्षित आवृत्तियां	:	17	52	54	31	6

0, 1, 2, 3 या 4 सिक्कों के होने की अपेक्षित आवृत्तियां ज्ञात करें और गुणों के तर्क संगत परिकल्पना क्या सिक्का निष्पक्ष है का परीक्षण देखें

हल:- अपेक्षित आवृत्ति $= N.n_{c_4}p^2q^{n-2} = 160.4_{c_2}(0.5)^2(0.5)^{4-2}$

X	अपेक्षित आवृत्ति
	160^4cx $(.5)^4$ = E
0	160 X^4 C0 $(.5)^4$ = 10
1	160 X^4 C1 $(.5)^4$ = 40
2	160 X^4 C2 $(.5)^4$ = 60
3	160 X^4 C3 $(.5)^4$ = 40
4	160 X^4 C4 $(.5)^4$ = 10

जब x^2 प्रयोग करते हैं।

चिट्टों की संख्या	O	E	(O-E)	$(O-E)^2$	$\frac{(O-E)^2}{E}$
0	17	10	7	49	4.9
1	52	40	12	144	3.6
2	54	60	-6	36	0.6
3	31	40	-9	81	2.025
4	6	10	-4	36	1.6

$\Sigma\frac{(O-E)^2}{E}=12.725$

d.f=5-1=4, x^2 (0.05)= 9.488

x^2का परिकलित मतान 12.725 है जो तालिका मान 9.488 से अधिक है, गलत तर्क है।

2. **स्वतन्त्रता के परीक्षण के रूप में-** x^2 परीक्षण का उपयोग यह पता लगाने के लिए किया जा सकता है। उदाहरण के लिए कोंचिंग कक्षा और सफल उम्मीदवार विवाह और विफलता आदि , हम यह पता कर सकते हैं कि क्या वे सम्बन्धित है या स्वतन्त्र। हम एक अवधारणा लेते हैं कि गुण स्वतन्त्र हैं। यदि x^2 का परिकलित मान एक निश्चित महत्व के स्तर पर तालिका मान से कम है, तो अनुमान सही है अन्यथा विपरीत।

उदाहरण - एक गाँव में 120 लोगों के नमूनों में से , इन्फलुएंजा को रोकने के लिए 76 लोगों को एक नई दवा दी गई थी और उनमें से 24 व्यक्ति इन्फलुएंजा द्वारा ग्रसित थे। जिन व्यक्तियों में नई दवा का प्रबंध नहीं किया गया उनमें से 12 इन्फलूएंजा से प्रभावित नहीं थे।

(अ) वास्तविक एवं अपेक्षित आवृत्तियों को दर्शाते हुए 2×2 तालिका तैयार करें।

(ब) काई-वर्ग का प्रयोग करते हुए ज्ञात करें कि क्या नई दवा प्रभावी है या नही ।

(5% महत्व के स्तर पर एक स्वतन्त्रता की श्रेणी के लिए काई वर्ग 3.84 है)

हल-

2×2 तालिका

	A	α	
B	24	32	56 (B)
β	52	12	64
	76	44	120
	(A)		N

मान लें इंफलुएंजा और नई दवा स्वतन्त्र हैं।

अपेक्षित आवृत्तियां

$\frac{76 \times 56}{120} = 35.5$ 56

$\frac{56 \times 44}{120} = 20.5$ $\frac{76 \times 64}{120} = 40.5$ $\frac{64 \times 44}{120} = 23.5$ 64

76 44 120

O	E	O-E	(O-E)²	$\frac{(O-E)^2}{E}$
24	35.5	-11.5	132.25	3.725
52	40.5	11.5	132.25	3.265
32	20.5	11.5	132.25	6.451
12	23.5	11.5	132.25	5.627

$d.f = (2-1)(2-1) = 1, d.f$ के लिए $x^2 = 3.84$

x^2 का परिकलित मान 19.068 है जो कि तालिका मान की तुलना में बहुत ज्यादा है। इसलिए , परिकल्पना अस्वीकार्य है। इसलिए हम यह निष्कर्ष निकालते हैं कि इन्फलूएंजा को नियंत्रित करने में दवा निस्संदेह प्रभावी है।

उदाहरण- 2000 परिवा रों के एक निश्चित नमूने में 1400 परिवार चाय के उपभोक्ता है। 1800 हिन्दू परिवारों में से 1236 परिवार चाय का सेवन करते हैं। x^2 परीक्षण का उपयोग करें और बताएं कि क्या हिंदू और हिन्दू परिवारों के बीच चाय की खपत के बीच कोई महत्वपूर्ण अंतर है।

हल- 2×2 आकस्मिता तालिका में जानकारी के सारणीकरण पर, हम प्राप्त करते हैं:-

	हिन्दू	गैर हिन्दू	योग
उपभोग चाय	1236	164	1400
गैर उपभोग चाय	564	36	600
योग	1800	200	2000

परिकल्पना गुण स्वतन्त्र है।

अपेक्षित आवृत्तियां

$\frac{1800 \times 1400}{2000} = 1260$, $\frac{1800 \times 600}{2000} = 540$

$\frac{200 \times 1400}{2000} = 140,$ $\frac{200 \times 600}{2000} = 60$

x^2 की गणना

O	E	O-E	(O-E)²	$\frac{(O-E)^2}{E}$
1236	1260	-24	576	0.457
564	540	+24	576	1.068
164	140	+24	576	4.114
36	60	-24	576	9.600

$$\sum \frac{(O-E)^2}{E} = 15.239$$

$d.f$ 1 है, 1 $d.f$ के लिए x^2 का तालिका मान = 3.841

x^2 का परिकलित मान 15.239 तालिका मान (3.841) से बहुत अधिक है। इसलिए शून्य परिकल्पना अस्वीकार्य है। इसलिए चाय के उपभोग के संबंध में दो समुदायों में काफी अंतर है।

उदाहरण- निम्नलिखित परिणामों के साथ एक पासा 120 उछाला जाता है।

ऊपर की संख्या	:	1	2	3	4	5	6	Total
आवृत्ति	:	30	25	18	10	22	15	120

इस परिकल्पना का परीक्षण करें कि पासा निष्पक्ष है।

हल- परिकल्पना है कि पासा निष्पक्ष है।

अपेक्षित आवृत्ति $\left[120 \times \frac{1}{6}\right] = 20$ है।

x^2 परीक्षण का प्रयोग करते हुए

O	E	O-E	(O-E)²	$\frac{(O-E)^2}{E}$
30	20	10	100	5.00
25	20	5	25	1.25
18	20	-2	4	0.20
10	20	-10	100	5.00
22	20	2	4	0.20
15	20	-5	25	1.25

$$\sum \frac{(O-E)^2}{E} = 12.90$$

d.f = n -1 = 6 -1 = 5

5% महत्व के स्तर पर 5 d.f के लिए तालिका मान 11.07 है जो कि x^2 के परिकालित मान 12.90 से कम है परिकल्पना जो कि पासा निष्पक्ष है, 5% महत्व के स्तर पर अस्वीकार्य है।

उदाहरण 5- टाइफाइड के खिलाफ इसकी प्रभावशीलता का परीक्षण करने के लिए एक निश्चित इलाके में कुल 720 में से 456 पुरूषों को एक निश्चित दवा दी गई थी। टाइफाइड की घटनाओं को नीचे दिखाया गया है। रोग के खिलाफ दवा की प्रभावशीलता का पता लगाएं।
(1 d.f के लिए 5% महत्व के स्तर पर x^2 का तालिका मान 3.84 है)

हल- 2×2 आकस्मिकता तालिका

	संक्रमण	गैर संक्रमण	योग
दवा	144	312	456
बिना दवा के	192	72	264
योग	336	384	720

परिकल्पना है कि दवा स्वतन्त्र है। अपेक्षित आवृत्तियां

$$\frac{336 \times 456}{720} = 212.80$$
$$\frac{336 \times 264}{720} = 123.2$$
$$\frac{384 \times 456}{720} = 243.2$$

$\frac{384 \times 264}{720} = 140.8$ है।

O	E	O-E	(O-E)²	$\frac{(O-E)^2}{E}$
144	212.8	-68.8	4733.44	22.24
192	123.2	+68.8	4733.44	38.42
312	243.2	+68.8	4733.44	19.46
72	140.8	-68.8	4733.44	33.62

$$\sum \frac{(O-E)^2}{E} = 113.74$$

x^2 का परिकलित मान =113.74 जो कि 1 d.f में 5% महत्व के स्तर पर तालिका मान से बहुत ज्यादा है। इसलिए यह अत्यन्त महत्वपूर्ण है। शून्य परिकल्पना गलत है। इसलिए टायफाइड को नियंत्रित करने में दवा निश्चित रूप से प्रभावी है।

उदाहरण - एक शहर में 8000 स्नातकों में से 800 महिलाएं हैं 1600 स्नातक कर्मचारियों में से 120 महिलाएं हैं x^2 का प्रयोग करके ज्ञात करें कि लिंग के आधारपर नियुक्ति में कोई विभेद है। 1 स्वतन्त्रता की श्रेणी के लिए 5% स्तर पर x^2 का मान 3.84 है।

हल- प्रश्न में दी गई जानकारी को 2×2 तालिका में सारणीबद्ध किया जा सकता है।

	कार्यरत	बेरोजगार	योग
पुरूष	1480	5720	7200
महिला	120	680	800
योग	1600	6400	8000

हम इस अवधारणा को मानते हैं कि लिंग के आधार पर नियुक्ति में कोई अंतर नहीं है।

$\frac{7200\times1600}{8000} = 1440$

$$\frac{7200\times6400}{8000} = 5760$$

$$\frac{1600\times800}{8000} = 160$$

$\frac{6400\times800}{8000} = 640$ है।

O	E	O-E	(O-E)²	$\frac{(O-E)^2}{E}$
1480	1440	40	1600	1.111
120	160	-40	1600	10.00
5720	5760	-40	1600	0.278
680	640	40	1600	2.500

$$\sum \frac{(O-E)^2}{E} = 13.889$$

$$\chi^2 = \sum \frac{(O-E)^2}{E} = 13.889$$

d.f = (r - 1) (c - 1)

$= 1 \times 1 = 1$

1 d.f के लिए $x^2_{0.05} = 3.84$ है जो कि तालिका मान 3.84 से बहुत अधिक है इसलिए , परिकल्पना अस्वीकार्य है। इसका अर्थ है कि लिंग के आधार पर नियुक्तियां हुई है।

4.6 सारांश (Summary)

महत्व का स्तर आमतौर पर यूनानी प्रतीक (लोअरकेस अल्फा) द्वारा निर्दिष्ट किया जाता है। महत्व के लोकप्रिय 10%(0.1), 5%(0.05), 1%(0.01), 0.5%(0.005), 0.1%(0.001) है। यदि महत्व का एक परीक्षण महत्व

स्तर से कम p मान देता है , तो शून्य परिकल्पना को अस्वीकार कर दिया जाता है। ऐसे परिणामों को अनौपचारिक रूप से सांख्यिकी में महत्वपूर्ण कहा जाता हैं उदाहरण के लिए , यदि कोई तर्क करता है कि "एक हजार में केवल एक ही मौका संयोग से होता है।, "सांख्यिकीय महत्व का 0.001 एक स्तर निहित है। कम महत्व के स्तर के लिए मजबूत साक्ष्य की आवश्यकता होती हैं महत्व का स्तर चुनना काफी हद तक मनमाना कार्य है , लेकिन कई अनुप्रयोगों के लिए 5% का स्तर चुना जाता है, इसके कोई बेहतर कारण नहीं है, यह परंपरागत है। कुछ स्थितियों में यह $1-\propto$ के रूप में सांख्यिकीय महत्व व्यक्त करने के लिए सुविधाजनक है। सामान्य तौर पर, जब एक घोषित महत्व की व्याख्या करते हैं , तो सावधानी वरतनी चाहिए कि , वास्तव में सांख्यिकीय तरीके से परीक्षण किया जा रहा है। बराबर प्रभावों को बंद करने के विभिन्न स्तरों में महत्व के निर्धारण विश्वसनीयता बढाने के छोटे स्तर लेकिन गलत शून्य परिकल्पना (II प्रकार की त्रुटियां "गलत नकारात्मक दृढ संकल्प) को अस्वीकार करने में विफल रहने का अधिक जोखिम चलाते हैं , और इसलिए कम सांख्यिकीय शक्ति होती है। इस स्तर का चयन अनिवार्यतः महत्व एवं शक्ति के बीच एक समझौता है , और इसका परिणाम टाइप I त्रुटि और टाइप II त्रुटि के बीच होता है।

4.7 शब्दावली (Glossary)

स्वतन्त्रता की श्रेणी (degree of freedom) - जिसका अर्थ है कि वर्गों की संख्या जिनाक मान बिना किसी प्रतिबंध के उल्लंघन के मुताविक सौंपा जा सकता है।

काई वर्ग परीक्षण (chi square test) - गुणों के तर्क संगत के लिए सांख्यिकीय में प्रयोग किया जाता है।

4.8 बोध प्रश्न (Comprehension Questions)

1. नमूना आकार $n = 10, \mu = 5$ किग्रा

दिए हुए नमूने आंकडों में से सबसे पहले $\bar{x}$ और s की गणना करते हैं।

								Total
x:	4.7	4.9	5.0	5.1	5.2	4.6	4.7	**49.3**
x^2:	22.09	24.01	25.00	26.01	27.04	21.16	22.09	**243.73**

$$\bar{x} = \frac{49.3}{10} = 4.93$$

$$S = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{243.73}{10} - (4.93)^2}$$

$$= \sqrt{2.4373 - 24.30}$$

$$= \sqrt{0.073} = 0.27$$

$H_0 = \mu = 5$ किग्रा

$H_1 = \mu \neq 5$ किग्रा

$$\text{t आंकडा परीक्षण} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} = \frac{4.93-5}{\frac{.27}{\sqrt{9}}} = \frac{-0.07\times 3}{.27} = -0.78$$

$d.f = n-1 = 10-1 = 9$

$9. d. f.$ के लिए 5% महत्व के स्तर पर तालिका मान $= 2.262$

2.$\mu = 2.00$ सेमी, $n = 10$ ट्यूब $\bar{x} = 2.01$ सेमी

$\sigma = \sqrt{0.004}$ सेमी

चूंकि $n < 30$, नमूना, छोटा नमूना है। इसलिए माध्य के परीक्षण के लिए t परीक्षण का प्रयोग करते हुए

$H_0: \mu = 2.00$ सेमी

$H_1: \mu \neq 2.00$ सेमी

आंकडा परीक्षण $t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}}$ है।

$$= \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n-1}}} = \frac{2.01-2.00}{\sqrt{\frac{0.004}{9}}} = \frac{.01\times 3}{\sqrt{.004}} = \frac{0.03}{0.0632} = 0.475$$

$d.f$ ि(स्वतन्त्रता की श्रेणी की संख्या) $= 9$

$9. d. f$ के लिए 5% स्तर पर तालिका मान $= 2.262$

3.$p_1 = \frac{16}{500} = 0.032$ (पहले नमूने में)

$p_2 = \frac{3}{100} = 0.0.$ (दूसरे नमूने में)

मान लें कि मशीन मरम्मत के बाद नहीं सुधरी हुई है।

$$S.E.(p_1 - p_2) = \sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$p = \frac{500\times 0.032 + 100\times 0.3}{500+100}$$

$$= \frac{16+3}{600} = 0.03$$

q = 1 – 0.03 = 0.97

$$S.E.(p_1 - p_2) = \sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$= \sqrt{(0.03)(0.97)\left(\frac{1}{500}\right)+\left(\frac{1}{100}\right)}$$

$$== \sqrt{(0.03)(0.97)(0.002+0.01)}$$

=0.0187

$$z=\frac{0.032-0.03}{0.0187}=\frac{0.002}{0.0187}=0.106$$

4.माना x_1 और x_2 क्रमशः घंटों में कारखानों (A) और कारखाना (B) के मजदूरों की मजदूरी (रू . में) को प्रदर्शित करते हैं तो हमें दिया गया है।

$$n_1 = 150 \quad \bar{x}_1 = 2.56 \quad s_1 = 1.08 \quad = \sigma_1$$
$$n_2 = 200 \quad \bar{x}_2 = 2.87 \quad s_2 = 1.28 \quad = \sigma_2$$

शून्य परिकल्पना H_0

$\mu_1 = \mu_2$ कारखाना A एवं कारखाना B में मजदूरों के मजदूरी के औसत स्तर के बीच कोई महत्वपूर्ण अंतर नहीं है।

वैकल्पिक परिकल्पना H_1

$\mu_2 > \mu_1$ या $\mu_1 < \mu_2$ (आंये पुच्छीय परीक्षण)

आंकडा परीक्षण

H_0 के अंतर्गत, आंकडा परीक्षण (बडे नमूनों के लिए)

$$z=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{\bar{X}_1-\bar{X}_2}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}$$

$$\therefore z=\frac{2.56-2.87}{\sqrt{\frac{(1.08)^2}{150}+\frac{(1.28)^2}{200}}}$$

$$z=\frac{-0.31}{\sqrt{0.016}}=\frac{-0.31}{0.126}$$

= - 2.46

महत्वपूर्ण क्षेत्र

एक पुच्छीय परीक्षण के लिए, 5% महत्व के स्तर पर Z का महत्वपूर्ण मान 1.645 है। बाएं पुच्छीय परीक्षण के लिए महत्वपूर्ण क्षेत्र में Z के सभी मान शामिल है।

4.9 बोध प्रश्नों के उत्तर (Answers to Comprehension Questions)

1.**निष्कर्ष-** 5% स्तर पर शून्य परिकल्पना (H_0) स्वीकार हैं इसलिए मशीन उचित ढंग से काम कर रही है।

2.**निष्कर्ष-** 5% स्तर पर H_0 स्वीकार्य है चूंकि t का परिकलित मान , t के तालिका मान से कम है। इसलिए समग्र माध्य और नमूना माध्य के बीच का अंतर महत्वपूर्ण नहीं है।

3.चूंकि (1% स्तर पर) अंतर 2.58 से कम है, प्रयोगों के परिणाम परिकल्पना को प्रमाणित करते हैं। इसलिए , हम निष्कर्ष निकालते हैं कि मशीन मरम्मत के बाद भी नहीं सुधरी है।

4.**निष्कर्ष-** चूंकि 5% महत्व के स्तर पर Z का परिकलित मान (-2.46) महत्वपूर्ण मान (1.645) की तुलना में कम है। इसलिए शून्य परिकल्पना 5% महत्व के स्तर पर अस्वीकार्य है और हम यह निष्कर्ष निकालते हैं कि कारखाना B द्वारा प्रदत्त औसतन प्रति घंटा मजदूरी कारखाने A द्वारा भुगतान की तुलना में निश्चित रूप से अधिक है।

4.10 स्वपरख प्रश्न (Self-Assessment Questions)

1. एक भरने की मशीन से 5 किलो पाउडर भरने की उम्मीद है। 10 थैलों के नमूने ने 4.7, 4.9, 5.0, 5.1, 5.4, 5.2, 4.6, 5.1, 4.6 और 4.7 के वजन दिए। जांच करें कि मशीन ठीक से काम कर रही है।

2. एक कम्पनी 2.00 सेमी की औसत व्यास के स्टील ट्यूब का निर्माण क र रही है। 10 ट्यूबों का एक नमूना 2.01 सेमी का एक आंतरिक व्यास और 0.004 सेमी^{2} का विचलन देता है। क्या माध्य के मान में अंतर महत्वपूर्ण है। 5% के स्तर पर 9 d.f के लिए t का मान = 2.262

3. एक मशीन ने 500 के नमूने में से 16 अपूर्ण लेखों को मशीन में लिया , मशीन की मरम्मत के बाद यह 100 के खेप में से 3 अपूर्ण लेखों को लेती हैं क्या मशीन में सुधार हुआ है।

4. कारखाना A में 150 श्रमिकों के एक नमूने की औसत प्रतिघंटा मजदूरी रू .2.56 , रू.1.08 के मानक विचलन के साथ थी। कारखाना B में 200 श्रमिकों के एक नमूने की औसत प्रतिघंटा मजदूरी रू.2.87, रू.1.28 के मानक विचलन के साथ थी। क्या एक सुरक्षित रूप में अनुमान लगाया जा सकता है कि कारखाना A द्वारा भुगतान किए गए प्रति घंटे की मजदूरी कारखाना B से अधिक है।

4.11 सन्दर्भ पुस्तकें (Reference Books)

1. मूल सांख्यिकीय – गौण, गुप्ता और दासगुप्ता वर्ल्ड प्रेस लिमिटेड -कलकत्ता
2. व्यावसायिक सांख्यिकीय की बुनियादी बातें संचेती और कपूर
3. प्रबंधन में मात्रात्मक तरीके - श्रीवास्तव, शेनाय और गुप्ता
4. व्यावसायिक सांख्यिकीय - गुप्ता और गुप्ता

इकाई 5 सहसम्बन्ध विश्लेषण
(Correlation Analysis)

5.1 प्रस्तावना (Introduction)

5.2 उद्देश्य (Objective)

5.3 परिभाषा (Definition)

5.4 उपयोगिता / महत्त्व (Importance)

5.5 सहसम्बन्ध के प्रकार (Types of Correlation)

5.6 सहसम्बन्ध गुणांक और उसका विस्तार (Correlation coefficient and its expansion)

5.7 सहसम्बन्ध ज्ञात करने की रीतियाँ (Methods of finding correlation)

5.7.1 बिन्दु रेखीय रीति (Point Linear Method)

5.7.2 विक्षेप चित्र या बिन्दु चित्र (Deflection diagram or point diagram)

5.7.3 कार्ल पियर्सन का सहसम्बन्ध गुणांक (Karl Pearson's correlation coefficient)

5.7.4 स्पियरमैन की कोटि-अन्तर विधि (Spearman's rank-difference method)

5.8 सारांश (Summary)

5.9 अभ्यास प्रश्न (Practice Questions)

5.10 अभ्यास प्रश्नों के उत्तर (Answers to Practice Questions)

5.11 संदर्भ ग्रन्थ सूची (Bibliography)

5.1 प्रस्तावना (Introduction)

सांख्यिकी में 'सह-सम्बन्ध का सिद्धान्त' अति महत्वपूर्ण है। सहसम्बन्ध के अन्तर्गत हम दो चर मूल्यों (variable) के बीच परस्पर आश्रितता का अध्ययन करते हैं। इसके मूल-तत्वों का प्रतिपादन सर्वप्रथम फ्रांस के खगोल-शास्त्री ब्रावे (Bravais A) ने सन् 1846 के लगभग किया था, परन्तु इस सिद्धान्त को विकसित करने व आधुनिक रूप देने का श्रेय प्रसिद्ध प्राणिशास्त्री फ्रांसिस गाल्टन (Francis Galton) तथा कार्ल पियर्सन (Karl Pearson) को प्राप्त है। इन प्रसिद्ध वैज्ञानिकों ने प्राणिशास्त्र (Biology) तथा जनन-विद्या (Genetics) के क्षेत्र में सहसम्बन्ध के सिद्धान्त के आधार पर अनेक समस्याओं का वैज्ञानिक विश्लेषण किया है। सहसम्बन्ध विश्लेषण से हमें यह ज्ञात होता है कि दो सम्बन्धित चर मूल्यों में कितना और किस प्रकार का सम्बन्ध है। इस सिद्धान्त के आधार पर ही प्रत्येक क्षेत्र में दो अथवा दो से अधिक घटनाओं के परस्पर सम्बन्धों का स्पष्टीकरण होता है।

नित्य के अनुभव से ऐसे कई उदाहरण हमारे सामने आते हैं जहाँ विभिन्न चर मूल्यों के बीच एक सम्बन्ध स्थापित होता है। उदाहरण के लिए, जैसे-जैसे बच्चों की ऊँचाई बढ़ती जाती है उनका भार भी बढ़ता है। एक समंक श्रेणी में परिवर्तित होने से दूसरी सम्बन्धित समंक-श्रेणी में भी परिवर्तन आता है। सामान्य अनुभव के आधार पर हम जानते हैं कि हमारे देश में कृषि उत्पादन का स्तर मानसून वर्षा के ऊपर निर्भर करता है। अच्छी वर्षा वाले वर्षों में कृषि उत्पादन का स्तर अधिक होता है , परन्तु मानसून प्रतिकूल हो जाने पर कृषि उत्पादन का स्तर कम हो जाता है। इसी प्रकार हम जानते हैं कि जिन कृषि जोतों (agricultural holdings) में सिंचाई की व्यवस्था अच्छी होती है उनमें कृषि उत्पादन की प्रति हैक्टेयर उपज अधिक होती है , परन्तु असिंचित जोतों में प्रति हैक्टेयर उत्पादकता का स्तर निम्न होता है। इसी प्रकार चर-राशियों के मध्य अन्तर्सम्बन्धों के बहुत से उदाहरण दिये जा सकते है। अर्थशास्त्र के विद्यार्थी भली-भाँति जानते हैं कि उपभोग व्यय व्यक्ति की आय के ऊपर निर्भर करता है , उत्पादन की कुल लागत उत्पादन स्तर के ऊपर निर्भर करती है , वस्तु की पूर्ति बढ़ने से उसकी कीमत गिर जाती है लेकिन वस्तु की माँग बढ़ने पर उसकी कीमत बढ़ जाती है तथा देश में मुद्रा की पूर्ति की मात्रा बढ़ने से सामान्य कीमत स्तर में वृद्धि होगी। गाल्टन ने अध्ययन किया कि लम्बे पिताओं के पुत्र भी सामान्य रूप से लम्बे होते हैं। जब दो चर-मूल्यों में कार्य कारण-सम्बन्ध (cause-effect relationship) हो तो वे सहसम्बन्धित कहलाते है।

जब कभी दो चर मूल्यों में ऐसा सम्बन्ध हो कि एक में परिवर्तन होने से दूसरे में भी परिवर्तन हो - एक में वृद्धि होने पर दूसरे में वृद्धि या कमी हो अथवा एक में कमी होने पर दूसरे में कमी या वृद्धि हो तो ये चर-मूल्य सह-सम्बन्धित कहलाते है। इस गुण को सह-सम्बन्ध (correlation) कहते हैं। एक चर मूल्य में अधिक परिवर्तन होने पर यदि दूसरे चर-मूल्य में भी अधिक परिवर्तन हो तो सह-सम्बन्ध की मात्रा अधिक होगी।

उपर्युक्त उदाहरणों में हमने दो चरों के मध्य अन्तर्सम्बन्धों की चर्चा की। प्रायः सम्बन्ध तीन अथवा उससे अधिक चरों में भी हो सकते है - जैसे कृषि उत्पादन के क्षेत्र में किसी फसल की प्रति हैक्टेयर उत्पादकता सिँचाई सुविधाओं के अतिरिक्त उर्वरकों की मात्रा, श्रम एवं पूँजी की मात्रा, बीजों की किस्म तथा कीटनाशकों के प्रयोग के ऊपर निर्भर करती है। किसी वस्तु की माँग-मात्रा वस्तु की कीमत के अतिरिक्त उपभोक्ता की आय , अन्य वस्तुओं की कीमतें , अभिरुचियों इत्यादि पर निर्भर करती है। इसी प्रकार किसी परिवार का वार्षिक उपभोग व्यय, वार्षिक आय के अतिरिक्त परिवार के आकार , परिवार के सदस्यों की अभिरुचियाँ , परिवार की सामाजिक प्रतिष्ठा इत्यादि पर निर्भर करेगा। इस प्रकार के सांख्यिकीय विश्लेषण को , जिसमें चरों की संख्या दो होती है द्विचरीय विश्लेषण (bivariate analysis) भी कहा जाता है तथा जिसमें चरों की संख्या तीन अथवा इससे अधिक होती है, उसे बहुचरीय विश्लेषण (multivariate analysis) कहा जाता है।

5.2 उद्देश्य (Objective)

इस अध्याय को पड़ने के बाद आप-

- ✓ सहसम्बन्ध की परिभाषा को जानेंगे।
- ✓ प्रकीर्ण आरेख , रेखाचित्र, कार्ल पियर्सन का सहसम्बन्ध गुणांक , कोटि-अन्तर से सहसम्बन्ध गुणांक , संगामी विचलन गुणांक इत्यादि का विवेचन करेंगे।

5.3 परिभाषा (Definition)

सह-सम्बन्ध विश्लेषण की कुछ परिभाषाएँ-

(1) *''यदि दो या दो से अधिक राशियाँ सहानुभूति में परिवर्तित होती है जिससे एक में होने वाले परिवर्तनों के फलस्वरूप दूसरी राशि में भी परिवर्तन होने की प्रवृत्ति पायी जाती है , तो वे राशियाँ सह-सम्बन्धित कहलाती है।*''- **एल. आर. कार्नोर**

(2) *''यदि यह सत्य सिद्ध हो जाता है कि अधिकांश उदाहरणों में चर सदा एक दिशा में या विपरीत दिशा में उच्चावचन की प्रवृत्ति रखते है , तो ऐसी स्थितियों में हम यह समझते हैं कि उनमें एक सम्बन्ध पाया जाता है। यह सम्बन्ध ही सहसम्बन्ध कहलाता है।*'- **डब्ल्यू. आई. किंग**

(3) *''सह-सम्बन्ध का सम्पूर्ण विषय पृथक विशेषताओं के बीच पाये जाने वाले उस पारस्परिक सम्बन्ध की ओर संकेत करता है जिसके अनुसार वे कुछ अंशों में साथ-साथ परिवर्तन होने की प्रवृत्ति रखते है।*'- **डेवनपोर्ट**

(4) *''जब सम्बन्ध मात्रात्मक प्रकृति का होता है , तो उस सम्बन्ध को खोजने तथा मापन करने और उसे एक संक्षेप सूत्र में अभिव्यक्त करने का उपयुक्त सांख्यिकीय उपकरण सहसम्बन्ध के रूप में जाना जाता है।*''- **क्राक्स्टन एवं काउडेन**

(5) "*सहसम्बन्ध विश्लेषण चरों के बीच 'सम्बन्ध की मात्रा"को निर्धारित करने का प्रयास करता है।* ''- **या लुन चाऊ**

(6) *''जब कभी आँकड़ों के दो या अधिक समूहों, वर्गों या श्रेणियों में कुछ निश्चित सम्बन्ध पाया जाता है तो वह सहसम्बन्ध कहलाता है।*''- **बार्डिंगटन**

उपर्युक्त परिभाषाओं से यह स्पष्ट है कि पद 'सहसम्बन्ध' दो या दो से अधिक चरों के बीच सम्बन्ध का अध्ययन करने का संकेत करता है।

5.4 उपयोगिता / महत्त्व (Utility/Importance)

सहसम्बन्ध के अध्ययन की उपयोगिता भौतिक विज्ञान तथा सामाजिक विज्ञान , दोनों में ही पर्याप्त है , तथापि हम यहाँ केवल सामाजिक विज्ञान में सहसम्बन्ध अध्ययन की उपयोगिता की ही व्याख्या करेंगे।

(1) सहसम्बन्ध का अध्ययन निर्णयन लेने से सम्बन्धित अनिश्चितता के परास को कम करता है। सामाजिक विज्ञान, विशिष्टतया व्यावसायिक जगत में , पूर्वानुमान एक महत्त्वपूर्ण तत्व या परिघटना और सहसम्बन्ध अध्ययन सापेक्षतः अधिक विश्वसनीय पूर्वानुमानों के करने में हमारी मदद करता हैं।

(2) सहसम्बन्ध विश्लेषण आर्थिक व्यवहार को समझने में सहायक होता है। यह हमें ऐसे चरों को निर्धारित करने में सहायता करता है जिन पर अन्य चर निर्भर रहते हैं। यह उन घटकों या कारकों के अध्ययन करने में सहायक

होता है जिससे आर्थिक घटनाएँ प्रभावित होती है। उदाहरणार्थ, हम मूल्य वृद्धि अथवा निम्न उत्पादकता के उत्तरदायी कारकों को जान सकते है।

(3) सहसम्बन्ध अध्ययन हमें ऐसे कारकों की पहचान करने में मदद करता है जो एक बाधाग्रस्त आर्थिक स्थिति का स्थायीकरण कर सकता है।

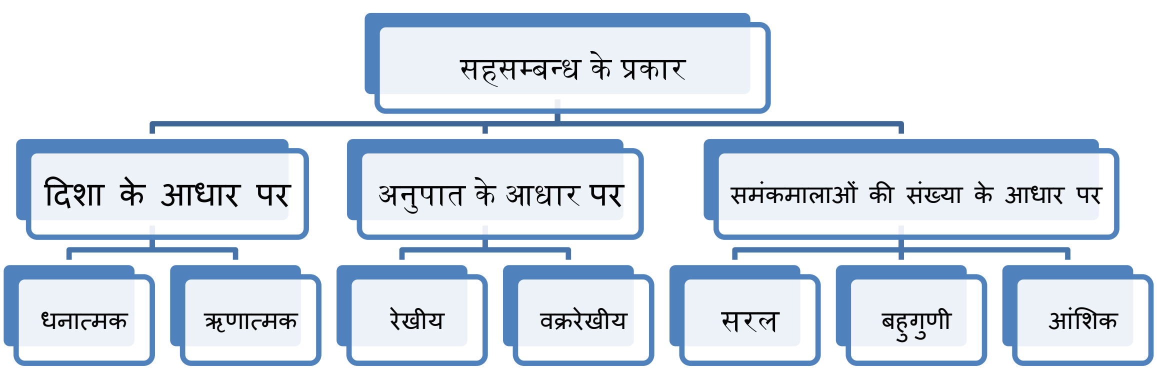
(4) सहसम्बन्ध अध्ययन एक चर में संभाव्य परिवर्तन का सम्बंद्ध दूसरे चर में परिवर्तन की विशिष्ट राशि के साथ आकलन करने में हमारी मदद करता है। उदाहरणार्थ , सहसम्बन्ध अध्ययन कीमत में एक निश्चित राशि के परिवर्तन से माँग में परिवर्तन जानने में मदद कर सकता है।इस दशा में हम समाश्रयण विश्लेषण (regression analysis) की सहायता लेते हैं।

(5) विभिन्न चरों के बीच अन्तर-सम्बन्ध अध्ययन अनुसंधान संवर्द्धन करने तथा ज्ञान के नये क्षेत्र खोलने में बहुत ही सहायक उपकरण होते हैं।

इस प्रकार सहसम्बन्ध अध्ययनों का विभिन्न उद्देश्यों के लिए व्यापक रूप से उपयोग किया जाता है और उन्हें दो या अधिक चरों से सम्बन्धित सांख्यिकीय आँकड़ों के विस्तृत विश्लेषण और निर्वाचन के लिए बुनियादी उपकरण समझा जाता है।

5.5 सहसम्बन्ध के प्रकार (Types of Correlation)

विभिन्न आधारों को लेकर हम सहसम्बन्ध का वर्गीकरण निम्न प्रकार कर सकते है-



- **धनात्मक और ऋणात्मक सहसम्बन्ध (Positive and Negative Correlation)**

सहसम्बन्ध धनात्मक अथवा ऋणात्मक हो सकते है। जब दो चरों में एक ही दिशा में परिवर्तन होता है अर्थात् एक में वृद्धि (या कमी) होने से दूसरे चर के मूल्यों में भी वृद्धि (या कमी) होती है तो ऐसा सहसम्बन्ध प्रत्यक्ष (direct) अथवा धनात्मक (positive) कहलाता है। इसके विपरीत , जब एक चर के मूल्यों में एक दिशा में परिवर्तन होने से दूसरे सम्बद्ध चर के मूल्यों में विपरीत दिशा में परिवर्तन होते हैं तो उनका सहसम्बन्ध ऋणात्मक (Negative), अप्रत्यक्ष या विलोम (inverse) कहलाता है। कुछ ऐसे आँकड़े होते हैं जिनमें

सहसम्बन्ध सामान्यतः धनात्मक और कुछ ऐसे जिसमें सामान्यतः ऋणात्मक होता है , जैसे अन्य बाते सामान्य रहे तो मूल्य और पूर्ति में साधारण तौर पर धनात्मक सहसम्बन्ध होता है। जब मूल्य बढ़ता है तो पूर्ति भी बढ़ती है और जब मूल्य घटता है तो पूर्ति भी घटती है। मूल्य और माँग में सहसम्बन्ध साधारणतः ऋणात्मक होता है। मूल्य में वृद्धि के साथ माँग घटती है और मूल्य के घटने के साथ माँग में साधारणतः वृद्धि होती है।

धनात्मक सहसम्बन्ध (Positive Correlation)

कीमत (Price)	10	15	20	25
पूर्ति (Supply)	100	110	115	130

अथवा

मूल्य या कीमत (Price)	40	30	20	10
पूर्ति (Supply)	150	140	115	100

ऋणात्मक सहसम्बन्ध (Negative Correlation)

मूल्य या कीमत (Price)	10	15	20	25
माँग (Demand)	100	90	80	60

अथवा

मूल्य या कीमत (Price)	25	20	15	10
माँग (Demand)	60	80	90	100

- **रेखीय तथा वक्ररेखीय सहसम्बन्ध (Linear and Curvilinear Correlation)**

परिवर्तनों के अनुपात के आधार पर सहसम्बन्ध रेखीय अथवा वक्र-रेखीय हो सकता है। यदि दो चर-मूल्यों के परिवर्तनों का अनुपात स्थायी (constant ratio) होता है तो उनका सहसम्बन्ध रेखीय (linear) कहलाता है, अर्थात्, यदि प्रत्येक बार मूल्य में 10 प्रतिशत की वृद्धि हो तो पूर्ति में 20 प्रतिशत वृद्धि, रेखीय सम्बन्ध का प्रमाण देगी। उनमें सम्बन्ध $y = a + bx$ के रूप में होगा। यह एक सीधी रेखा का समीकरण होता है। रेखीय सहसम्बन्ध वाले चर-मूल्यों को बिन्दुरेख पर प्रांकित करने से एक सरल रेखा बन जाती है। इस प्रकार का सह-सम्बन्ध भौतिक व पूर्ण विज्ञानों में पाया जाता है। आर्थिक व सामाजिक क्षेत्र में अधिकतर वक्ररेखीय सहसम्बन्ध पाया जाता है। जब दो चर-मूल्यों के परिवर्तनों का अनुपात अस्थिर (variable ratio) या परिवर्तनशील होता है तो उनका सहसम्बन्ध वक्र-रेखीय (Curvilinear) होता है। यदि मुद्रा की मात्रा में 10 प्रतिशत वृद्धि होने से कभी सामान्य कीमत स्तर में 5 प्रतिशत वृद्धि हो जाती है , कभी 6 प्रतिशत, कभी 9 प्रतिशत तो मुद्रा की मात्रा और सामान्य कीमत स्तर का सह-सम्बन्ध वक्ररेखीय कहलाएगा। ऐसी स्थिति में रेखाचित्र पर चर-मूल्यों को प्रांकित करने से एक वक्र रेखा बनेगी।

रेखीय या रैखिक सहसम्बन्ध (inear correlation)

x	2	4	6	8	10

y	5	10	15	20	25

अरेखीय या अरैखिक सहसम्बन्ध (Non-linear correlation)

x	2	4	6	8	10
y	5	8	12	15	25

- **सरल, बहुगुणी एवं आंशिक सह-सम्बन्ध (Simple, Multiple and Partial Correlation)**

स्वतंत्र तथा आश्रित चर-मूल्यों (variable) की संख्या के आधार पर सह-सम्बन्ध सरल , बहुगुणी या आंशिक हो सकता है। दो चर-मूल्यों के सह-सम्बन्ध को सरल सहसम्बन्ध (simple correlation) कहते हैं। इन चर-मूल्यों में से अनाश्रित या प्रधान चर-मूल्य (independent variable) को प्रमाप या आधार श्रेणी (subject series) कहा जाता है तथा दूसरा समंक-समूह आश्रित चर-मूल्य (dependent variable) या सम्बद्ध माला (relative series) कहलाता है।

जब तीन या अधिक कारकों , जैसे उत्पादन, वर्षा और खाद के उपयोग के बीच सम्बन्ध का साथ-साथ अध्ययन किया जाता है , तो इसे 'बहु सहसम्बन्ध (multiple Correlation) कहते हैं। आंशिक सह-सम्बन्ध (partial Correlation) के अन्तर्गत दो से अधिक चर-मूल्यों का अध्ययन किया जाता है परन्तु अन्य चर-मूल्यों के प्रभाव को स्थिर रखकर केवल दो चर-मूल्यों का पारस्परिक सम्बन्ध निकाला जाता है। उदाहरणार्थ , यदि वर्षा की मात्रा और तापक्रम दोनांेें के गैंहू की उपज पर सामूहिक प्रभाव का गणितीय अध्ययन किया जाए तो वह बहुगुणी सहसम्बन्ध कहलाएगा। इसके विपरीत यदि एक स्थिर तापक्रम में वर्षा की मात्रा और गैंहू की उपज के सम्बन्ध का विवेचन किया जाये तो यह आंशिक सहसम्बन्ध कहलाएगा।

5.6 सहसम्बन्ध गुणांक और उसका विस्तार (Correlation coefficient and its Magnitude)

गैरेट के अनुसार, ''सहसम्बन्ध गुणांक दो चलराशियों में पाये जाने वाला ऐसा अनुपात है जिससे यह ज्ञात होता है कि एक चर में होने वाले परिवर्तन ज्ञात दूसरे चर पर किस मात्रा में प्रभाव डालते हैं अथवा किस मात्रा में उसका अनुसरण करते हैं। "अतः स्पष्ट है कि सहसम्बन्ध गुणांक दो या अधिक प्रवृत्तियों के परिमाणात्मक (quantitative) सम्बन्ध को स्पष्ट करता है। वास्तव में यह एक प्रकार का सूचकांक (index) है।

सहसम्बन्ध की मात्रा +1 से -1 तक होती है अर्थात् सहसम्बन्ध कभी भी 1 से अधिक नहीं होता है चाहे यह धनात्मक हो या ऋणात्मक। जब सहसम्बन्ध की मात्रा +1 आती है तो पूर्ण धनात्मक सहसम्बन्ध (perfect positive Correlation) होता है और जब सहसम्बन्ध की मात्रा - 1 होती है तो इसे पूर्ण ऋणात्मक सहसम्बन्ध (perfect negative Correlation) कहते हैं। लेकिन समाज विज्ञानों (social science) से सम्बन्धित चल राशियों में पूर्ण ऋणात्मक अथवा धनात्मक सहसम्बन्ध नहीं आता है। सहसम्बन्ध की मात्रा को निम्न प्रकार से भी प्रदर्शित किया जा सकता है:

-1, -.9, -.8, -.7, -.6, -.5, -.4, -.3, -.2, -.1, 0, .1, .2, .3, .4, .5, .6, .7, .8, .9

सहसम्बन्ध की व्याख्या (interpretation of Correlation) सहसम्बन्ध की मात्रा से पहले यदि (+) चिन्ह आता है तो हम कहेंगे कि सहसम्बन्ध धनात्मक है और यदि सहसम्बन्ध की मात्रा से पहले (-) चिन्ह आता है तो हम कहेंगे कि सहसम्बन्ध ऋणात्मक है। गिलफोर्ड ने सहसम्बन्ध की मात्रा का वर्गीकरण निम्न प्रकार से किया है-

सहसम्बन्ध गुणांक की मात्रा Quantity of Coefficient of Correlation	सम्बन्ध Relationship
$.00 \rightarrow \pm .20$	नगण्य (Negligible)
$\pm .21 \rightarrow \pm .40$	निम्न (Low)
$\pm .41 \rightarrow \pm .60$	साधारण (मध्यम) (Moderate)
$\pm .61 \rightarrow \pm .80$	उच्च (High)
$\pm .81 \rightarrow \pm .99$	अति उच्च (Very High)
± 1	पूर्ण सहसम्बन्ध (Perfect Correlation)

ऊपर दी हुई तालिका के आधार पर सहसम्बन्ध की व्याख्या की जा सकती है। उदाहरण के लिए , यदि सहसम्बन्ध की मात्रा +.85 है तो यहा कहा जाएगा कि दी हुई चलराशियों में धनात्मक और बहुत उच्च सहसम्बन्ध है। धनात्मक सहसम्बन्ध में चलराशियाँ किस प्रकार से एक दूसरे से प्रभावित होती है , यह भी एक रोचक तथ्य है।

5.7 सहसम्बन्ध ज्ञात करने की रीतियाँ (Methods of finding correlation)

सहसम्बन्ध ज्ञात करने की निम्नलिखित प्रमुख रीतियाँ हैं-

5.7.1 बिन्दु रेखीय रीति (Point Linear Method)

5.7.2 विक्षेप चित्र या बिन्दु चित्र (Deflection diagram or point diagram)

5.7.3 कार्ल पियर्सन का सहसम्बन्ध गुणांक (Karl Pearson's correlation coefficient)

5.7.4 स्पियरमैन की कोटि-अन्तर विधि (Spearman's rank-difference method)

5.7.1 बिन्दु रेखीय रीति (Point Linear Method)

इस रीति के अनुसार हम सहसम्बन्ध का अनुमान समय , स्थान, क्रम संख्या आदि को X-axis पर और दोनों आश्रित समंकमालाओं को Y-axis पर अंकित करते हैं। इस विधि से सहसम्बन्ध की मात्रा का ज्ञान नहीं होता बल्कि इसकी दिशा और मात्रा का अनुमान किया जाता है। दोनांेे श्रेणियों के बिन्दुरेख विपरीत दिशाओं में उतार-चढ़ाव को प्रदर्शित करें तो ऋणात्मक सहसम्बन्ध होता है। यदि दोनांेे श्रेणियों के परिवर्तनों की प्रवृत्ति उसी दिशा या विपरीत दिशाओं में न दिखाई दे तो कोई सहसम्बन्ध नहीं होगा।

उदाहरण (Illustration)- 1

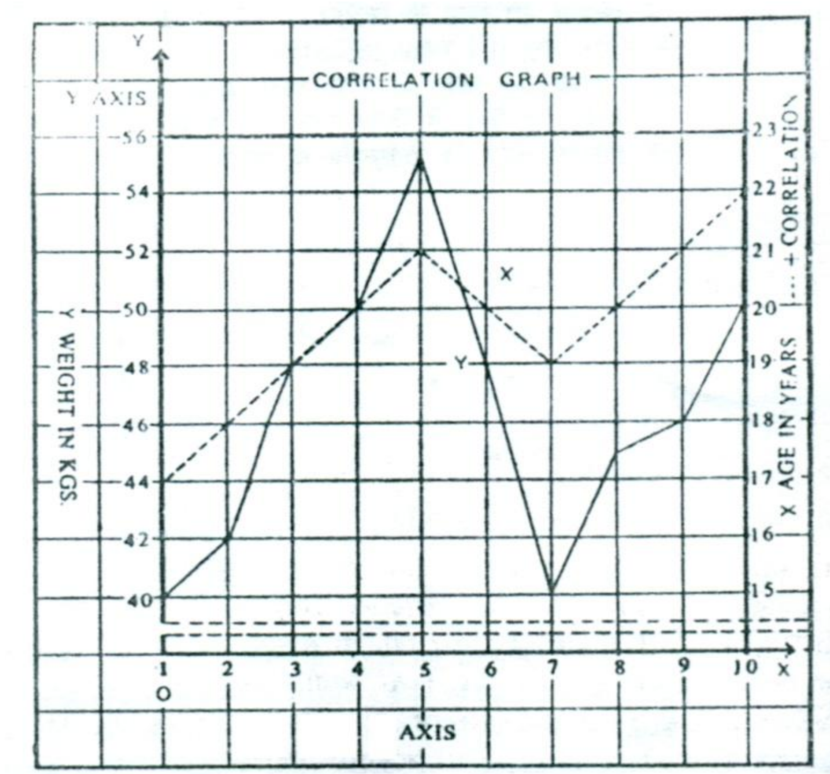
निम्न आँकड़ों से एक सहसम्बन्ध बिन्दु रेखाचित्र बनाइए-

Draw a Correlation graph from the following

उम्र (वर्षों में) Age in Years	17	18	19	20	21	20	19	20	21	22
वजन (कि. ग्रा. में) Weight in kgs	40	42	48	50	55	48	40	45	46	50

क्या उम्र एवं वनज में कोई सहसम्बन्ध है?

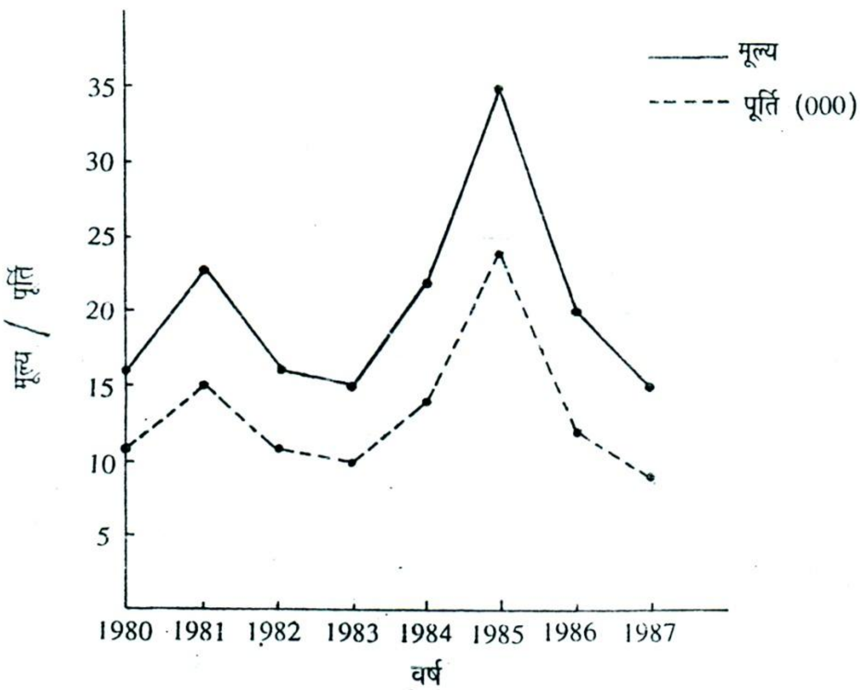
Is there any correlation in age and weight?



उदाहरण (Illustration): 2

मूल्य तथा वस्तु की पूर्ति के सम्बन्ध में नीचे दिये गये आँकड़ों के आधार पर ग्राफिक विधि से मूल्य तथा पूर्ति के बीच सहसम्बन्ध पर प्रकाश डालिए।

वर्ष	1980	1981	1982	1983	1984	1985	1986	1987
मूल्य (प्रति क्विंटल)	16	23	16	15	22	35	20	15
पूर्ति (क्विंटल)	11000	15000	11000	10000	14000	24000	12000	9000

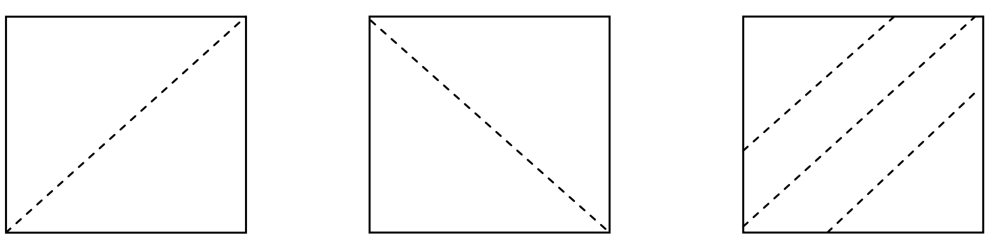


स्पष्ट है कि दोनों समंकमालाओं के बीच सहसम्बन्ध है।

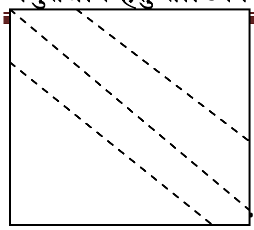
5.7.2 विक्षेप चित्र या बिन्दु चित्र (Scatter diagram or Art diagram

इस विधि से भी सह-सम्बन्ध की मात्रा का ज्ञान नहीं होता , बल्कि इसकी दिशा और मात्रा का अनुमान किया जाता है। इस रीति में स्वतंत्र चर मूल्यों (x) को X-axis पर तथा आश्रित चर मूल्यों (y) को Y-axis पर अंकित किया जाता है। इस प्रकार x तथा y दोनों समंकमालाओं के जितने पदयुग्म (pair of values) होते हैं उतने ही बिन्दु रेखाचित्र पर अंकित कर दिए जाते है। इस प्रकार के चित्र को ही विक्षेप चित्र या बिन्दु चित्र कहते हैं। निम्न पाँच चित्रों की सहायता से हम सह-सम्बन्ध की दिशा और मात्रा का अनुमान लगा सकते है-

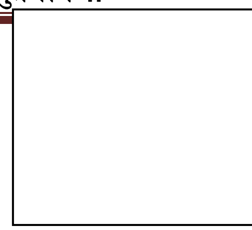
(i) (ii) (iii)



r = +1 **r = -1** Limited Positive



Limited Negative



No Correlation

उदाहरण (Illustration)- 3

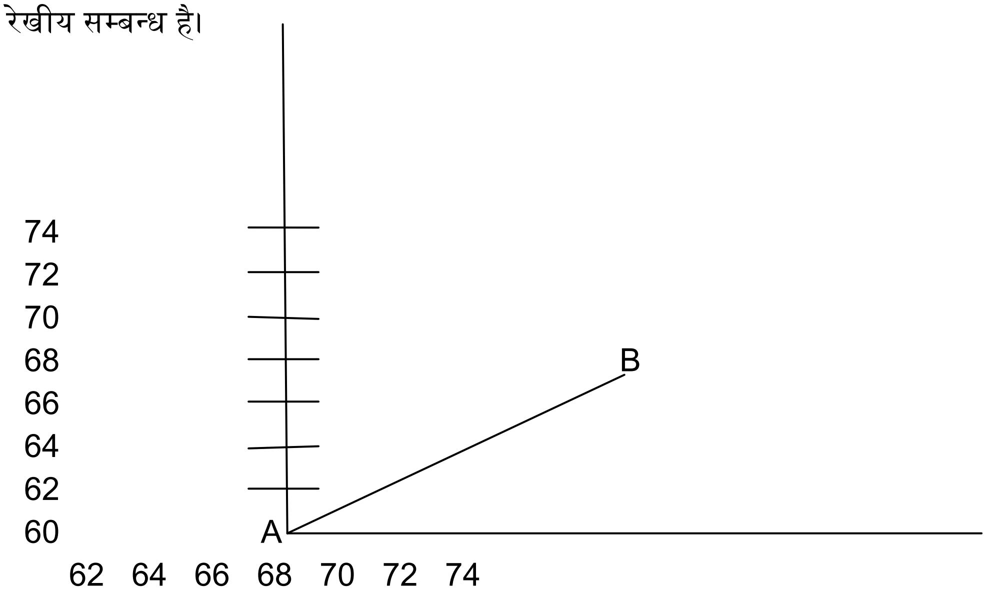
निम्नलिखित सारिणी में 5 पिता तथा उनके अग्रज पुत्रों के भार सम्बन्धी आँकड़े प्रदर्शित हैं -

पिता का भार (kg)	65	63	67	64	68	62	70	66	68	67	69	71
पुत्र का भार (kg)	68	66	68	65	69	66	68	65	71	67	68	70

आँकड़ों को विक्षेप चित्र के द्वारा प्रदर्शित कीजिए।

दिए हुए आँकड़ों में प्रथम , पिता तथा पुत्र के भार क्रमशः 65 तथा 68 कि.ग्रा. है। इन्हें ग्राफ पर बिन्दु के रूप में अंकित किया जाता है। तत्पश्चात् द्वितीय पिता तथा पुत्र के भारों को ग्राफ पर बिन्दु के रूप में प्रदर्शित किया जाता है। इसी प्रकार अन्य पिता तथा पुत्रों के भारों को ग्राफ पर विभिन्न बिन्दुओं के रूप में अंकित कर लिया जाता है। इन बिन्दुओं का प्रवृत्ति पथ AB (trend path) दोनों चर राशियों के मध्य सम्बन्ध को प्रदर्शित करता है।

नीचे चित्र में AB अंकित बिन्दुओं का प्रवृत्ति पथ है। अन्य शब्दों में सम्बन्धित चर राशियों के बीच रेखीय सम्बन्ध है।



विक्षेप चित्र की भाषा में सहसम्बन्ध प्रवृत्ति पथ से विक्षेप बिन्दुओं (scatter points) की निकटता की माप करता है। यदि सभी विक्षेप बिन्दु प्रवृत्ति पथ पर स्थित हैं तो ऐसी स्थिति में चरराशियों के मध्य पूर्ण सहसम्बन्ध (perfect Correlation) होगा तथा फलनात्मक सम्बन्ध (प्रस्तुत उदाहरण में सरल रेखा AB) दिये हुए समंक को पूर्ण रूप से प्रदर्शित करेगा। परन्तु यदि विक्षेप बिन्दु प्रवृत्ति पथ के दो नों ओर बिखरे हुए हैं तो

ऐसी स्थिति में चर राशियों के मध्य अपूर्ण सहसम्बन्ध (imperfect Correlation) होगा अर्थात् प्रवृत्ति पथ 'AB' चर राशियों के मध्य सम्बन्ध को पूर्ण रूप से प्रदर्शित नहीं करेगा। विक्षेप बिन्दुओं के प्रवृत्ति पथ के सन्निकट होने पर यह सहसम्बन्ध प्रबल (Strong Correlation) होगा तथा यदि विक्षेप बिन्दु प्रवृत्ति पथ के दोनांेें ओर दूर-दूर तक फैले हुए हैं, तो ऐसी स्थिति में सहसम्बन्ध निर्बल (Weak Correlation) होगा।

5.7.3 कार्ल पियर्सन का सहसम्बन्ध गुणांक (Karl Pearson's correlation coefficient)

सह-सम्बन्ध ज्ञात करने की यह सर्वश्रेष्ठ विधि है क्योंकि इससे सहसम्बन्ध का संख्यात्मक माप भी प्राप्त होता है। समान्तर माध्य और प्रमाप विचलन पर आधारित इस रीति में गणितीय दृष्टि से पूर्ण शुद्धता है। इस रीति का का प्रतिपादन कार्ल पियर्सन ने सन् 1890 में प्राणिशास्त्र की समस्याओं का अध्ययन करने के लिए किया था। कार्ल **पियर्सन का सह-सम्बन्ध गुणक निम्न मात्राओं पर आधारित है**

(I) दोनों श्रेणियों में रेखीय सम्बन्ध है।

(II) समंक माला को प्रभावित करने वाले स्वतंत्र कारणों में परस्पर कारण व प्रभाव का सम्बन्ध होता है।

(III) सह-सम्बन्धित श्रेणियों पर अनेक कारणों से सामानता आ जाती है।

कार्ल पियर्सन के सह-सम्बन्ध गुणांक की प्रमुख विशेषताएँ निम्न है

(a) यह गुणांक श्रेणी के सभी पदों पर आधारित है।

(b) इससे सह-सम्बन्ध की दिशा ज्ञात हो जाती है।

(c) चूँकि यह गुणांक समान्तर माध्य और प्रमाप विचलन पर आधारित है , इसलिए अनेक बीजगणितीय गुणयुक्त है।

(d) इसको ज्ञात करने के लिए दो नों श्रेणियों के समान्तर माध्य निकाल कर विचलनों की गणना की जाती है और इसके बाद इनका गुणनफल निकाल कर उसके जोड़ में मूल्यों की संख्या से भाग दिया जाता है। इसे सह-विचलन (Co-variance) कहते हैं। इस विधि में प्रयोग किया जाने वाला सूत्र इस प्रकार है -

$$r = \frac{\sum dx.dy}{N\sigma_x.\sigma y}$$

जहाँ - r = सहसम्बन्ध गुणांक

dx = x के मानों का उसके माध्य ($\overline{X}$) से विचलन

dy = y के मानों का उसके माध्य ($\overline{Y}$) से विचलन

σx = x समंक माला का प्रमाप विचलन

σy = y समंक माला का प्रमाप विचलन

N = पदों की संख्या

सूत्र से स्पष्ट है कि दो समंक मालाओं के मध्य सहसम्बन्ध गुणांक ज्ञात करने के लिए सर्वप्रथम प्रत्येक समंक माला का माध्य ($\overline{X}$ एवं $\overline{Y}$) ज्ञात करते हैं। इसके बाद प्रत्येक समंक माला के सभी पदों का उनके माध्य से विचलन ज्ञात कर लेते हैं , जिन्हें dx एवं dy द्वारा व्यक्त किया जाता है। फिर प्रमाप विचलन ज्ञात करने के लिए विचलनों dx तथा dy का वर्ग (dx^2 एवं dy^2) करके उनका अलग- अलग योग ($\sum dx^2$ एवं $\sum dy^2$ ज्ञात कर लेते हैं। इसके अतिरिक्त विचलनों dx एवं dy का गुणनफल करके उनका योग Σ(dx.dy) निकाल लिया जाता है। उपरोक्त सूत्र में -

$\frac{\sum(dx.dy)}{N}$ को सह-विचरण (co-variance) कहते हैं।

उदाहरण (Illustration)- 4

सन् 2008 की परीक्षा में दस विद्यार्थियों द्वारा अर्थशास्त्र एवं सांख्यिकी में पाये गये प्राप्तांकों का विवरण इस प्रकार है-

विद्यार्थी	A	B	C	D	E	F	G	H	I	J
प्राप्तांक (अर्थशास्त्र)	47	57	58	60	62	67	70	71	76	82
प्राप्तांक (सांख्यिकी)	56	50	47	60	62	64	65	70	74	82

हल- निम्न सारणी में अर्थशास्त्र के प्राप्तांकों को 'x' एवं सांख्यिकी के प्राप्तांकों को 'y' के द्वारा प्रदर्शित किया गया है।

विद्यार्थी	प्राप्तांक (x)	अर्थ0 के प्राप्तांकों का माध्य (65) से विचलन (dx)	dx^2	प्राप्तांक (dy)	सांख्य0 प्राप्तांकों का माध्य (63) से विचलन dy	(dy^2)	dx.dy
A	47	-18	324	56	-7	49	126
B	57	-8	64	50	-13	169	104
C	58	-7	49	47	-16	256	112
D	60	-5	25	60	-3	9	15
E	62	-3	9	62	-1	1	3
F	67	2	4	64	1	1	2
G	70	5	25	65	2	4	10
H	71	6	36	70	7	49	42
I	76	11	121	74	11	121	121
J	82	17	289	82	19	361	323
N=10	650		946	630		1020	858

अर्थशास्त्र के प्राप्तांकों का औसत -

$$\overline{X} = \frac{\sum x}{N} = \frac{650}{10} = 65$$

$$\sigma_x = \sqrt{\frac{\sum dx^2}{N}} = \sqrt{\frac{946}{10}} = \sqrt{94.6} \cong 9.7$$ (लगभग 9.7)

सांख्यिकी के प्राप्तांकों का औसत -

$$\overline{Y} = \frac{\sum y}{N} = \frac{630}{10} = 63$$

$$\sigma_y = \sqrt{\frac{\sum dy^2}{N}} = \sqrt{\frac{1020}{10}} = \sqrt{102} \cong 10.1$$

अब सहसम्बन्ध गुणांक $r = \dfrac{\sum(dx.dy)}{N\sigma_x.\sigma_y}$

$$= \frac{858}{10(9.7)(10.1)}$$

$$= \frac{858}{979.7} = 0.88 \text{ लगभग}$$

अतः अर्थशास्त्र एवं सांख्यिकी के प्राप्तांकों के मध्य उच्च धनात्मक सह-सम्बन्ध है।
अन्य शब्दों में जिन विद्यार्थियों ने सांख्यिकी में उच्च अंक प्राप्त किए है उनके अर्थशास्त्र में भी उच्च अंक है।

कार्ल पियर्सन के उपरोक्त सूत्र को ध्यानपूर्वक देखने से हम पाते हैं कि यदि इस सूत्र को एक अन्य रूप में लिखा जाय तो प्रत्येक समंक माला का प्रमाप विचलन निकालने की आवश्यकता नहीं पड़ती है एवं सहसम्बन्ध गुणांक की गणना पहले की अपेक्षा सरलता से हो जाती है।

कार्ल पियर्सन के सूत्र का सरलीकृत रूप -

$$r = \frac{\sum(dx.dy)}{N\sigma_x.\sigma_y}$$

$$= \frac{\sum(dx.dy)}{N \times \sqrt{\frac{\sum dx^2}{N}} \times \sqrt{\frac{\sum dy^2}{N}}} = \frac{\sum(dx.dy)}{N \times \sqrt{\frac{\sum dx^2}{N} \times \frac{\sum dy^2}{N}}}$$

$$= \frac{\sum(dx.dy)}{\frac{N}{N}\sqrt{\sum dx^2 \times \sum dy^2}} = \frac{\sum(dx.dy)}{\sqrt{\sum dx^2 \times \sum dy^2}}$$

अब इस सूत्र में केवल $\sum(dx.dy)$, $\sum dx^2$ एवं $\sum dy^2$ का मान रखकर सहसमबन्ध गुणांक ज्ञात किया जा सकता है। जैसे उदाहरण 5 के लिए -

$$r = \frac{\sum(dx.dy)}{\sqrt{\sum dx^2 \times \sum dy^2}}$$

$$= \frac{858}{\sqrt{(946)(1020)}} = \frac{858}{\sqrt{964920}} = \frac{858}{982} = 0.88 \text{ (लगभग)}$$

उपर्युक्त सूत्रों का दोष यह है कि यदि 'x' तथा 'y' श्रृंखलाओं के माध्यों के मान दशमलव में आते हैं तो इनके द्वारा सहसम्बन्ध की गणना की क्रिया अत्यंत जटिल हो जाती है।

अतः X तथा Y के मूल्यों के बीच सहसम्बन्ध गुणांक ज्ञात करने के लिए निम्नलिखित सूत्र का प्रयोग अधिक व्यावहारिक होता है -

$$r = \frac{N\sum xy - \sum x.\sum y}{\sqrt{\{N\sum x^2 - (\sum x)^2\}\{N\sum y^2 - (\sum y)^2\}}}$$

यह सूत्र गणना की दृष्टि से काफी सरल है। इसका कारण यह है कि इस सूत्र के अन्तर्गत सहसम्बन्ध गुणांक ज्ञात करने के लिए न तो हमें x एवं y के माध्यों को ज्ञात करना पड़ता है और न ही माध्य से विचलनों (dx एवं dy) अथवा प्रमाप विचलनों (x व y) की गणना करनी पड़ती है।

प्रस्तुत उदाहरण में उपर्युक्त सूत्र के द्वारा सहसम्बन्ध गुणांक की गणना निम्न सारिणी के माध्यम से की गयी है।

5.7.4 स्पियरमैन की कोटि-अन्तर विधि (Spearman's rank-difference method)

कार्ल पियर्सन ने दो चर मूल्यों के बीच पाये जाने वाले सम्बन्ध को स्पष्ट करने के लिए जो सूत्र दिया , वह हम स्पष्ट कर चुके हैं। बुद्धिमता , सुन्दरता, स्वास्थ्य आदि ऐसे तथ्य हैं जिन्हें प्रत्यक्ष रूप से अंकों में व्यक्त नहीं किया जा सकता। दूसरे शब्दों में, हम कह सकते हंै कि गुणात्मक तथ्यों के बीच सम्बन्ध जानने के लिए कार्ल पियर्सन द्वारा प्रतिपादित सूत्र नहीं लगाया जा सकता। इन गुणात्मक तथ्यों के बीच सहसम्बन्ध ज्ञात करने के लिए प्रसिद्ध सांख्यिक चार्ल्स एडवर्ड स्पियरमैन (Charles Edward Spearman) ने एक विधि का प्रतिपादन सन् 1904 में किया। उन्हीं के नाम पर इस विधि को स्पियरमैन की कोटि अन्तर विधि कहते हैं। एक सौन्दर्य प्रतियोगिता में माना 10 प्रतियोगी भाग लेते हैं और तीन निर्णायक हैं। विभिन्न प्रतियोगिताओं को गुण की अधिकता के आधार पर ये तीनों निर्णायक अपने ढंग से पहला, दूसरा, तीसरा . . . इत्यादि क्रम प्रदान करते हैं। इन क्रमों के आधार पर ही हम सहसम्बन्ध गुणांक निकालते हैं। माना हम यह जानना चाहते हैं कि इन तीन निर्णायकों में से ऐसे कौन से दो निर्णायक हैं जिनका सौन्दर्य-निर्णय लगभग समान है। यह समस्या कार्ल पियर्सन के सूत्र से हल नहीं हो सकती। हम जानते हैं कि विद्यार्थियों की योग्यता के जाँच के लिए परीक्षा पद्धति बनाई गई है, जिसमें विद्यार्थी प्रश्न-पत्र के उत्तर लिखते हैं और इन उत्तर-पुस्तिकाओं को परीक्षकों के पास भेज दिया जाता है। परीक्षक निर्धारित अधिकतम अंकों में से प्रत्येक उत्तर-पुस्तिका पर अंक देते हैं। अंक देने के लिए कोई निश्चित मापदण्ड नहीं होता यद्यपि मोटे तौर पर कुछ निर्देशों का पालन अवश्य करना होता है। इसी कारण हम सुनते है कि 'मैने कुछ नहीं लिखा और बहुत अच्छे अंक मिले' तथा 'मैने बहुत अच्छा लिखा और पता नहीं बहुत कम अंक मिले।' यह पद्धति दोषपूर्ण होने के कारण अब ग्रेड प्रणाली (grade system) को लाने पर बल दिया जा रहा है। योग्यता की जाँच भी कार्ल पियर्सन द्वारा प्रतिपादित सूत्र से ठीक प्रकार नहीं हो सकती , इसके लिए भी इसी विधि को अपनाना चाहिए। दो परीक्षकों की योग्यता जाँच की समानता देखने के लिए हम इसी कोटि अन्तर विधि द्वारा सह-सम्बन्ध गुणांक निकालते हैं। यहाँ श्रेणियों के पद-मूल्य ज्ञात न हों और उनका क्रम पता हो तो भी यह सहसम्बन्ध गुणांक ज्ञात किया जा सकता है।

इस विधि में सबसे पहले x तथा y दोनों श्रेणियों के पद-मूल्यों को अलग- अलग कोटि-क्रम (rank) प्रदान किए जाते है। इसके बाद कोटि-क्रम अन्तर ज्ञात करके उसका वर्ग निकालते हैं और जोड़ लेते हैं। निम्न सूत्र का उपयोग किया जाता है-

$$\rho = 1 - \frac{6\sum D^2}{N^3 - N}$$

जहाँ, ρ(rho) = Rank Correlation

D = Rank difference

N = Number of pairs

जब किसी श्रेणी में दो या दो से अधिक पद मूल्य बराबर आकार के हों तो उनके मूल्य क्रम निकालकर औसत निकाला जाता है तथा यही औसत क्रम प्रत्येक पद मूल्य के आगे रख दिया जाता हैं। ऐसा करने से गलती की सम्भावना रहती है। इसे समाप्त करने के लिए निम्न सूत्र का प्रयोग किया जाता है-

$$\rho = 1 - \frac{6[\sum D^2 + \frac{1}{12}(m^3 - m)]}{N^3 - N}$$

जहाँ m उस कोटि अथवा कोटियों की बारम्बारता है जो एक से अधिक बार घटित होती है।

रेखीय सहसम्बन्ध गुणांक की भाँति , कोटि अन्तर सहसम्बन्ध गुणांक का मान भी -1 से +1 के बीच स्थिर होता है। r का मान ऋणात्मक अथवा धनात्मक होने पर चरों के बीच का सम्बन्ध भी ऋणात्मक अथवा धनात्मक होता है। r का मान जितना ही +1 अथवा -1 के निकट होगा उतना ही चरों के बीच का सहसम्बन्ध प्रबल (strong) होगा तथा r का मान यदि शून्य है अथवा शून्य के निकट है , तो चरों का सहसम्बन्ध नगण्य होगा अर्थात् सम्बन्ध चर एक दूसरे से स्वतंत्र (independent) होंगे।

उदाहरण (Illustration)- 7

दो अध्यापकों द्वारा 8 विद्यार्थियों का मूल्यांकन नीचे दिया गया है -

विद्यार्थी	1	2	3	4	5	6	7	8
पहला अध्यापक	8	7	3	6	4	1	5	2
दूसरा अध्यापक	5	8	1	6	4	2	7	3

जहाँ तक 8 विद्यार्थियों के मूल्यांकन का प्रश्न है, दोनों अध्यापक किस हद तक एक दूसरे से सहमत है?

हल- उपर्युक्त प्रश्न में दो अध्यापकों द्वारा निर्धारित '8' विद्यार्थियों की कोटियों (ranks) को प्रदर्शित किया जाता है। दो नों अध्यापकों के मूल्यांकन में समानता का परीक्षण करने के लिए हम कोटि-अन्तर सहसम्बन्ध गुणांक का मान ज्ञात करेंगे। इसका सूत्र निम्न प्रकार है-

$$r = 1 - \frac{6\sum D^2}{N^3 - N}$$

इसकी गणना को निम्न सारिणी की सहायता से दर्शाया गया है -

विद्यार्थी क्रम संख्या	निर्धारित कोटि		कोटि अन्तर	
	पहला अध्यापक	दूसरा अध्यापक	***D***	***D²***
1	8	5	3	9
2	7	8	&1	1
3	3	1	2	4
4	6	6	0	0
5	4	4	0	0
6	1	2	&1	1
7	5	7	&2	4
8	2	3	&1	1
N=8				20

सारिणी से - $N = 8, \quad \sum D^2 = 20$

सूत्र में रखने पर- $r = 1 - \frac{6(20)}{(8^3 - 8)} = 1 - \frac{120}{(512 - 8)} = 1 - \frac{120}{504}$

$= 1 - \frac{5}{21} = \frac{16}{21} = 0.76$

उदाहरण (Illustration)- 8

निम्न समंकों से कोटि सहसम्बन्ध गुणांक निकालिए -

X	115	109	112	87	98	98	120	100	98	118
Y	75	73	85	70	76	65	82	73	68	80

हल-

कोटि सहसम्बन्ध गुणांक का परिगणन

A		B		कोटि अन्तर D	**कोटि अन्तरों के वर्ग D^2**
X	कोटि X	Y	कोटि Y		
115	3	75	5	-2	4
109	5	73	6.5	-1.5	2.25
112	4	85	1	+3	9
87	10	70	8	+2	4
98	8	76	4	+4	16
98	8	65	10	-2	4
120	1	82	2	-1	1
100	6	73	6.5	-0.5	0.25
98	8	68	9	-1	1
118	2	80	3	-1	1
N=10				$\sum D = 0$	$\sum D^2 = 42.50$

Series A में 98 तीन बार आया है तथा तीन समान क्रमों के लिए सूत्र में में जोड़ना होगा। इसी प्रकार Series B में 73 दो बार आया है अतः दो नों समान क्रमों के लिए $\frac{1}{12}(2^3 - 2)$ के बराबर संख्या $\sum D^2$ में जोड़नी पड़ेगी। सूत्रानुसार -

$$P = 1 - \frac{6[\sum D^2 + \frac{1}{12}(m^3 - m)]}{N^3 - N}$$

$$= 1 - \frac{6[42.5 + \frac{1}{12}(3^3 - 3) + \frac{1}{12}(2^3 - 2)}{10^3 - 10}$$

$$= 1 - \frac{6[42.5 + 2 + 0.5]}{990}$$

$$= 1 - \frac{6 \times 45}{990} = 1 - \frac{270}{990} = \frac{990 - 270}{990} = \frac{720}{990} = +.73$$

X और Y में सामान्य रूप से अधिक मात्रा का धनात्मक कोटि सह-सम्बन्ध (moderately high degree of positive rank Correlation) है।

5.8 सारांश (Summary)

जब दो चर-मूल्यों में इस प्रकार का सम्बन्ध हो कि एक में कमी या वृद्धि होने से दूसरे में भी उसी दिशा में या विपरीत दिशा में परिवर्तन होते हों तो वे दो नों सह-सम्बन्धित कहलाते हैं। इससे यह स्पष्ट हो जाता है कि दो सम्बद्ध समंक श्रेणियों में साथ-साथ परिवर्तन होने की प्रवृत्ति को ही सहसम्बन्ध या सह-विचरण (co-variation) कहते हैं।

इसे $\frac{Cov(X,Y)}{\sqrt{Var(X).Var(Y)}}$ द्वारा परिभाषित किया जाता है।

इसे ज्ञात करने के लिए 'गुणन-परिघात सहसम्बन्ध गुणांक ' रीति या 'कार्ल पियर्सन का सहसम्बन्ध गुणांक' की रीति को सर्वोत्तम मानी जाती है क्योंकि इससे सहसम्बन्ध की दिशा और मात्रा का संतोषजनक संख्यात्मक माप ज्ञात हो जाता है।

सम्बद्ध समंकमालाओं में चर-मूल्यों के परिवर्तनों की दिशा , अनुपात तथा मालाओं की संख्या के आधार पर सहसम्बन्ध को धनात्मक तथा ऋणात्मक बताया गया है।

अर्थात् हम यह कह सकते है कि सहसम्बन्ध का सामान्य अर्थ है , दो समंक-श्रेणियो में कारण और परिणाम के आधार पर परस्पर सम्बन्ध का पाया जाना। इस दृष्टि से सहसम्बन्ध दो समंकमालाओं के पारस्परिक सम्बन्ध की दिशा व मात्रा (direction and degree) का विश्लेषण तो करता है लेकिन सहसम्बन्ध की उपस्थिति मात्र से यह निष्कर्ष नहीं निकाल लेना चाहिए कि दो नों सम्बद्ध श्रेणियों में आवश्यक रूप से प्रत्यक्ष कार्य-कारण सम्बन्ध भी है।

अतः निष्कर्ष के रूप में यह कहा जा सकता है कि सहसम्बन्ध की वास्तविक जानकारी केवल उसकी उपस्थिति मात्र से नहीं की जा सकती , जब तक कि दो नों सम्बद्ध मात्राओं में प्रत्यक्ष कार्य-कारण सम्बन्ध की जानकारी न प्राप्त कर ली जाय। प्रो . बाडिंगटन का भी कहना है कि यदि सभी प्रमाण यह संकेत करते हैं कि दोनों सम्बद्ध श्रेणियों में सहसम्बन्ध है अथवा हो सकता है तो भी उन प्रमाणों की अत्यन्त सतक्रतापूर्वक जाँच की जानी चाहिए ताकि निष्कर्ष गलत न हो सकें।

5.9 अभ्यास प्रश्न(Practice Questions)

I. वस्तुनिष्ठ प्रश्न-

A. निम्नलिखित में से कौन सा सही है-

(1) सहसम्बन्ध गुणांक

(i) सदा धनात्मक होता है। (ii) सदा ऋणात्मक होता है।

(iii) या तो धनात्मक या ऋणात्मक हो सकता है।

(iv) इनमें से कोई नहीं।

(2) कार्ल पियर्सन का सहसम्बन्ध गुणांक का सूत्र है।

(i) $r = \frac{\sum xy}{\sigma_x \sigma_y}$ (ii) $r = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}}$

(iii) $r = \frac{\sum xy}{N\sigma_x}$ (iv) $r = \frac{\sum xy}{\sigma_x{}^2 \sigma_y{}^2}$

(3) कोटि सहसम्बन्ध-गुणांक सूत्र द्वारा ज्ञात किया जा सकता है:

(i) $r_s = 1 + \frac{6\sum D^2}{N^3 - N}$ (ii) $r_s = 1 - \frac{\sum D^2}{N^3 + N}$

(iii) $r_s = 1 - \frac{6\sum D^2}{N^3 - N}$ (iv) $r_s = 1 - \frac{6\sum D^3}{N^3 - N}$

(4) सहसम्बन्ध गुणांक का धनात्मक चिन्ह होगा जब-

(i) x के मान बढ़ रहे हो और y के मान घट रहे हो।

(ii) x और y दोनों के मान बढ़ रहे हों।

(iii) x के मान घट रहे हो और y के मान बढ़ रहे हो।

(iv) x और y के मान में कोई परिवर्तन न हो।

II. रिक्त स्थानों को भरिए-

(1) जब दो चरों के मान एक ही दिशा में संचलित होते हैं तो सहसम्बन्ध ----- कहा जाता है।

(2) जहाँ संख्यात्मक मापन कठिन होता है, सहसम्बन्ध गुणांक ----- से परिकलित किया जाता है।

(3) + 1 ----- सहसम्बन्ध है।

(4) धन चिन्ह संकेतिक करते हैं कि सहसम्बन्ध ------ है।

III. लघु उत्तरात्मक प्रश्न-

(1) सहसम्बन्ध से आप क्या समझते हैं? उसके मापन करने की प्रमुख विधियों के नाम लिखिए।

(2) कार्ल पियर्सन के सहसम्बन्ध गुणांक को समझाइए।

(3) संगामी विचलन गुणांक को समझाइए।

(4) क्या दो चरों के बीच सहसम्बन्ध कारण-प्रभाव का सम्बन्ध प्रकट करता है?

(5) कोटि सहसम्बन्ध को परिभाषित कीजिए। कोटि सहसम्बन्ध गुणांक (rs) के लिए स्पियरमैन का सूत्र लिखिए।

IV. निबन्धात्मक प्रश्न-

(1) सहसम्बन्ध की अवधारणा का अर्थ एवं महत्व स्पष्ट कीजिए। सहसम्बन्ध गुणांक के मान का निर्वचन आप किस प्रकार करेंगे?

(2) कार्ल पियर्सन के सहसम्बन्ध-गुणांक की परिभाषा दीजिए। यह किस बात को मापने का आशय करता है ? एक सहसम्बन्ध गुणांक चिन्ह और परिमाण का निर्वचन आप किस प्रकार करेंगे?

V. संख्यात्मक प्रश्न-

(1) निम्नलिखित आँकड़ों के लिए प्रकीर्ण आरेख बनाइए-

X :	8	10	12	11	9	7	13	14	15	17	16
Y :	5	7	9	8	6	4	10	11	12	14	13

x और y के बीच सम्बन्ध का वर्णन भी कीजिए।

(2) निम्न आँकड़ों से एक सहसम्बन्ध लेखाचित्र की रचना कीजिए और पूर्ति तथा मूल्य सूचकांकों के बीच सहसम्बन्ध पर टिप्पणी कीजिए।

वर्ष :	1980	1981	1982	1983	1984	1985
पूर्ति सूचकांक :	166	170	186	154	136	154
मूल्य सूचकांक :	216	200	196	208	214	204

(3) निम्नलिखित आँकड़ों से x और y के बीच कार्ल पियर्सन का सहसम्बन्ध परिकलित कीजिए-

$N = 13, \sum x = 117, \sum x^2 = 1313, \sum y = 260, \sum y^2 = 6580, \sum xy = 2827$

(4) निम्न आँकड़ों से कार्ल पियर्सन का सहसम्बन्ध गुणांक परिकलित कीजिए-

X :	6	8	12	15	18	20	24	28	31
Y :	10	12	15	15	18	25	22	26	28

(5) निम्नलिखित आँकड़ों से संगामी विचलन गुणांक की गणना कीजिए-

मूल्य :	368	284	385	361	347	384	395	403	400	385
आयात :	22	21	24	20	22	26	24	29	28	27

5.10 अभ्यास प्रश्नों के उत्तर (Answers to Practice Questions)

I. वस्तुनिष्ठ प्रश्न-

(1) (iii) (2) (ii) (3) (iii) (4) (ii)

II. रिक्त स्थानों को भरिए-

(1) धनात्मक (2) कोटि-अन्तर (3) पूर्ण (4) धनात्मक
(5) संगामी विचलन

V. संख्यात्मक प्रश्न-

(3) r = 0.81 (4) = +0.96 (5) rc = + 0.333

5.11 संदर्भ ग्रन्थ सूची (Bibliography)

1.बंसल, डॉ. एस. एन., एवं अग्रवाल, डॉ. डी. आर., (1978) सांख्यिकी के मूल तत्त्व, शिवलाल अग्रवाल एण्ड कम्पनी, आगरा।

2.सिंह, एस. पी. , (1997) सांख्यिकी-सिद्धान्त एवं व्यवहार , एस. चन्द एण्ड कम्पनी लिमिटेड, नई दिल्ली।

3.अवस्थी, जी. डी. एवं निगम, सुधीर कुमार, (2007) सांख्यिकीय विश्लेषण, भारत बुक सेन्टर, लखनऊ।

4.नागर, कैलाश नाथ, (2005) सांख्यिकी के मूल तत्व, मिनाक्षी प्रकाशन, मेरठ।

5.Goon, Gupta and Dasgupta, A Fundamental of Statistics, Volume-I, The World Press Private Limited.

इकाई - 6 प्रतीपगमन विश्लेषण

(Regression Analysis)

6.1 प्रस्तावना (Introduction)

सहसम्बन्ध में हमें दो चर मूल्यों के बीच आश्रितता का संख्यामक ज्ञान होता है। मोटे तौर पर हम यह कह सकते हैं कि सहसम्बन्ध गुणांक दो श्रेणियों के बीच सहसम्बन्ध की मात्रा को तो बताता है परन्तु एक श्रेणी के निश्चित चर-मूल्य के आधार पर दूसरी आश्रित श्रेणी के सम्बन्धित चर मूल्य का अनुमान नहीं बताता। प्रतीपगमन विश्लेषण द्वारा हम एक निश्चित चर मूल्य के सापेक्ष आश्रित श्रेणी के चर मूल्य का अनुमान लगा सकते हैं। यदि एक चर मूल्य का पता हो तो दूसरे चर मूल्य का पता जिस सांख्यिकीय रीति से हम लगाते हैं उसे प्रतीपगमन (Regression) कहते हैं।

अपने सामान्य अनुभव के आधार पर हम जानते हैं कि अधिक ऊँचाई वाले व्यक्तियों का वजन सामान्यतया अधिक तथा कम ऊँचाई वाले व्यक्तियों का वजन सामान्यतया कम होता है। अर्थात् व्यक्तियों की ऊँचाई तथा उनके वजनों के बीच सम्बन्ध होता है। इसी प्रकार सामान्यतया यह देखा जाता है कि पति तथा पत्नियों की आयु के मध्य भी सम्बन्ध होता है , अर्थात् अधिक आयु वाले पतियों की पत्नियाँ अधिक आयु वाली होती है तथा कम आयु वाले पतियों की पत्नियाँ कम आयु वाली होती है। इसी प्रकार चरों में सम्बन्धों के अनेक उदाहरण दिए जा सकते हैं। आर्थिक सिद्धान्त हमें बतलाता है कि व्यक्तियों की आय तथा उपभोग के स्तर के मध्य तथा कुल उत्पादन के स्तर और लागत स्तर के मध्य भी सम्बन्ध होता है।

यह सभी दो चर सम्बन्धों (Two variable relationship) के उदाहरण है। दो चर-मूल्यों के बीच कारण-परिणाम सम्बन्ध का सहसम्बन्ध की अपेक्षा प्रतीपगमन विश्लेषण से अधिक स्पष्टीकरण होता है। मुद्रा पूर्ति और सामान्य कीमत-स्तर में सहसम्बन्ध है। परन्तु हम सहसम्बन्ध के आधार पर यह नहीं कह सकते हैं कि एक निश्चित अवधि में मुद्रा पूर्ति कारण है और कीमत स्तर परिणाम। हो सकता है कि सामान्य कीमत स्तर कारण हो और मुद्रा पूर्ति परिणाम। प्रतीपगमन विश्लेषण में जो चर मूल्य दिया होता है उसे सदैव X (Independent variable) मानते हैं तथा जो चर मूल्य ज्ञात करना हो उसे Y (dependent variable) मानते हैं। अतः हम कह सकते हैं कि प्रतीपगमन विश्लेषण से कारण और परिणाम सम्बन्ध स्पष्ट रूप से पता लगता है। इस प्रकार के सम्बन्ध दो से अधिक चरों के मध्य भी हो सकते हैं , जैसे - प्रति हैक्टेयर कृषि उपज की मात्रा सिँचाई की मात्रा पर निर्भर करती है , साथ ही यह उर्वरकों की मात्रा कीटनाशकों के प्रयोग , बीजों के किस्म इत्यादि पर भी निर्भर करती है। इसी प्रकार किसी वस्तु की माँग वस्तु के मूल्य , व्यक्तियों की आय , अभिरुचियों तथा अन्य वस्तुओं के मूल्यों पर निर्भर करती है। एक परिवार का उपभोग व्यय , परिवार की आय के अतिरिक्त परिवार के आकार , परिवार के सदस्यों की अभिरुचियाँ , परिवार का सामाजिक स्तर आदि पर निर्भर करता है।

सांख्यिकीय विधियों द्वारा हम इस सम्बन्धों की कोटि (degree) तथा इनके स्वरूप को ज्ञात कर सकते हैं एवं आनुभाविक रूप से इस तथ्य पर भी प्रकाश डा ल सकते हैं कि इन चरों में सम्बन्ध कारणात्मक है अथवा नहीं।

प्रतीपगमन सांख्यिकीय विश्लेषण की वह विधि है जिसके द्वारा एक चर के किसी ज्ञात मूल्य से सम्बन्धित दूसरे चर का सम्भाव्य मूल्य प्रतीपगमन समीकरण की सहायता से अनुमानित किया जा सकता है। सांख्यिकी के आंग्ल भाषा के 'रिग्रेसन' (Regression) शब्द के लिए हिन्दी भाषा में 'समाश्रयण' शब्द का प्रयोग किया जाता है, यद्यपि कुछ लेखकों ने 'समाश्रयण' शब्द के स्थान पर 'प्रतीपगमन' शब्द प्रयोग किया है। जीव-विज्ञान और भू-विज्ञान में 'रिग्रेशन' शब्द के लिए 'प्रतिक्रमण' शब्द प्रयोग किया जाता है। प्रतीपगमन (या समाश्रयण) शब्द का अर्थ है , वापस लौटना या पीछे की ओर मुड़ना या घूमना (Stepping back or going back)। सांख्यिकी में इस शब्द का प्रयोग सर्वप्रथम सन् 1877 में सर फ्रांसिस गाल्टन (Sir Francis Galton)

नामक प्रसिद्ध वैज्ञानिक ने अपने शोध लेख - "पैतृक ऊँचाई में मध्यमता की ओर प्रतीपगमन " (Regression towards Mediocrity in Hereditary Stature) में किया था। उक्त शोध-लेख में उन्होंने लगभग एक हजार पिताओं और उनके पुत्रों की ऊँचाई या कद में सम्बन्ध का अध्ययन किया और कुछ बहुत ही रोचक निष्कर्ष निकाला। ये निष्कर्ष है-

(i) लम्बे पिताओं लम्बे और नाटे पिताओं के नाटे पुत्र होते हैं।

(ii) लम्बे पिताओं के पुत्रों की माध्य लम्बाई उनके पिताओं की माध्य लम्बाई की अपेक्षा कम होती है।

(iii) नाटे पिताओं के पुत्रों की माध्य लम्बाई उनके पिताओं की माध्य लम्बाई की अपेक्षा अधिक होती है।

(iv) गाल्टन ने यह पाया कि 'जाति (race) की माध्य लम्बाई से पिताओं की माध्य लम्बाई में विचलन की अपेक्षा जाति की माध्य लम्बाई से पुत्रों की माध्य लम्बाई में विचलन कम होता है। जब पिता माध्य लम्बाई से अधिक या कम लम्बे होते हैं तो पुत्रों की लम्बाई माध्य की ओर समाश्रयित (regress) या पीछे की ओर मुड़ जाती है।

इस प्रकार पुत्रों की ऊँचाई के सामान्य माध्य के निकट वापस जाने की इस प्रवृत्ति को ही फ्रांसिस गाल्टन ने 'मध्यमता की ओर प्रतीपगमन ' कहा था। गाल्टन ने इस प्रवृत्ति का प्रयोग एक ज्ञात चर (पिता की ऊँचाई) के तत्संवादी आश्रित चर (पुत्र की ऊँचाई) का सर्वोत्तम अनुमान लगाने के लिए किया था।

आधुनिक समय में प्रतीपगमन का उपयोग केवल पितृगत विशेषताओं के अध्ययन तक ही सीमित नहीं है बल्कि इसकी सामाजिक आर्थिक व व्यावसायिक क्षेत्रों में व्यावहारिक उपयोगिता है। दो या दो से अधिक श्रेणियों के पद मूल्यों में सामान्य माध्य की ओर वापस जाने की प्रवृत्ति होती है- यही प्रतीपगमन है। प्रतीपगमन की सहायता से हम एक चर मूल्य पर आधारित दूसरा चर मूल्य बड़ी सरलता से ज्ञात कर सकते हैं। दो सम्बन्धित श्रेणियों में प्रतीपगमन का अध्ययन बिन्दु-रेखीय ढंग से किया जाता है। विक्षेप चित्र (Scatter diagram) पर सर्वोपयुक्त रेखाएँ (Lines of best fit) खींची जाती हैं। इन्हें प्रतीपगमन रेखाएँ (Regression Lines) कहते हैं।

6.2 उद्देश्य (Objectives)

इस अध्याय को पड़ने के बाद आप-

- ✓ प्रतीपगमन को समझेंगे तथा प्रतीपगमन रेखाएँ को प्राप्त करेंगे।
- ✓ प्रतीपगमन गुणांक के अभिलक्षण या विशेषताएँ एवं उनके उपयोग को जानेंगे।

6.3 परिभाषा (Definition)

प्रतीपगमन की कुछ महत्वपूर्ण परिभाषाएँ निम्न है-

1. "आँकड़ों की मूल इकाइयों के रूप में, दो या अधिक चरों के बीच माध्य सम्बन्ध का माप समाश्रयण कहलाता है।" - **मारिस मेयर्स ब्लेयर**
2. "प्रायः यह ज्ञान करना अधिक महत्त्वपूर्ण होता है कि (दो या अधिक घटनाओं में) वास्तविक सम्बन्ध क्या है जिससे एक चर-मान (स्वतंत्र चर-मान) के ज्ञान के आधार पर दूसरे चर-मान (आश्रित चर-मान) का आकलन किया जा सके और इस प्रकार की दशा में प्रयोग की जाने वाली उपयुक्त सांख्यिकीय प्रविधि समाश्रयण-विश्लेषण कहलाती है।" - **वालिस एवं रॉबर्टस**

6.4 उपयोगिता / महत्त्व (Importance)

उपर्युक्त परिभाषाओं से यह स्पष्ट है कि प्रतीपगमन विश्लेषण एक चर के अज्ञात मान का दूसरे चर के ज्ञात मान से आक लन (estimate) या पूर्वकथन (predicting) के लिए किया जाता है। यह एक बहुत ही उपयोग सांख्यिकीय उपकरण (statistical tool) है जिसका प्रयोग प्राकृतिक और सामाजिक दो नों विज्ञानों में किया जाता है।

आर्थिक व व्यावसायिक जगत में प्रतीपगमन की अत्यधिक व्यावहारिक उपयोगिता है। प्रबन्ध-अधिकारियों द्वारा व्यवसाय के नियंत्रण-उपकरण (control tool) के रूप में प्रतीपगमन विश्लेषण का प्रयोग किया जाता है। इस प्रविधि के आधार पर उचित व्यावसायिक निर्णय लेना सर ल हो जाता है तथा उस निर्णय को व्यवहारिकता की कसौटी पर परखा जा सकता है। उदाहरणार्थ , इसके द्वारा यह अनुमान लगाया जा सकता है कि यदि किसी वस्तु के उत्पादन या उसकी पूर्ति में निश्चित मात्रा में वृद्धि या कमी हो जाए तो उसके मूल्य में संभावित परिवर्तन कितनी मात्रा में होगा। इसी प्रकार यह भी ज्ञात किया जा सकता है कि सामान्य मूल्य-स्तर में निश्चित परिवर्तन कितनी मात्रा में होगा। इसी प्रकार यह भी ज्ञात किया जा सकता है कि सामान्य मूल्य-स्तर में निश्चित वृद्धि होने पर जीवन-निर्वाह व्यय कितना बढ़ जाएगा। मूल्यों के आधार पर माँग का, वर्षा की मात्रा, बीज, खाद आदि के आधार पर कृषि उपज का तथा पूँजी के आधार पर लाभ आदि का अनुमान लगाने में प्रतीपगमन विश्लेषण बहुत सहायक सिद्ध होता है। व्यवसाय की सफ लता के लिए इस प्रकार के अनुमान अनिवार्य होते हैं परन्तु ये अनुमान तभी अधिक यथार्थ होते हैं जब दो नों श्रेणियों में परस्पर घनिष्ठ सहसम्बन्ध हो। प्रतीपगमन विश्लेषण की सहायता से चर-मूल्यों में सह-सम्बन्ध की मात्रा व दिशा का माप भी किया जा सकता है।

समाजशास्त्रीय अध्ययन में तथा आर्थिक आयोजन के क्षेत्र में जनसंख्या पुर्वानुमान (projection of population), जन्म-दरों (birth rates), मृत्यु-दरों (death rates) और इसी प्रकार के अन्य चरों के पुर्वानुमान बड़े उपयोगी होते हैं।

6.5 प्रतीपगमन के प्रकार (Types of Regression)

प्रतीपगमन को नापने की विधियाँ मुख्यतः तीन प्रकार की होती हैं-

I. सरल एवं बहुगुणी प्रतीपगमन

सरल प्रतीपगमन में एक चर स्वतंत्र तथा एक चर आश्रित होता है जैसे मूल्य तथा माँग के मध्य सम्बन्ध में मूल्य स्वतंत्र चर है तथा माँग आश्रित। बहुगुणी प्रतीपगमन में एक से अधिक स्वतंत्र चर के सापेक्ष केव ल एक ही आश्रित चर होता है जैसे किसी वस्तु की कीमत , उपभोक्ता की आय तथा उसकी रुचि इन तीनों स्वतंत्र चरों का प्रभाव उस वस्तु की माँग पर पड़ेगा अतः यहाँ वस्तु की माँग एक आश्रित चर है।

II. कुल एवं आंशिक प्रतीपगमन

कुल प्रतीपगमन में आश्रित चर पर प्रभाव जानने के लिए सभी स्वतंत्र चरों को विचार में लिया जाता है जबकि आंशिक प्रतीपगमन में एक या दो स्वतंत्र चरों पर ही विचार किया जाता है तथा शेष को छोड़ दिया जाता है।

III. रेखीय एवं अरेखीय प्रतीपगमन

जब स्वतंत्र चर तथा आश्रित चर में परस्पर सम्बन्धों को प्रदर्शित करने हेतु सीधी रेखा का प्रयोग किया जाता है तब यह 'रेखीय प्रतीपगमन' कहलाता है तथा जब यह प्रदर्शन वक्रीय रेखा द्वारा किया जाता है तो यह अरेखीय प्रतीपगमन कहलाता है।

सामान्यतः प्रतीपगमन की विधियों में सरल रेखीय प्रतीपगमन सर्वाधिक उपयुक्त मानी जाती है जिसमें प्रतीपगमन रेखाओं द्वारा स्वतंत्र एवं आश्रित चरों का परस्पर सम्बन्ध दर्शाया जाता है।

6.6 रेखीय प्रतीपगमन (Linear Regression)

दो परस्पर सम्बन्धित समंक श्रेणियों में प्रतीपगमन-विश्लेषण का कार्य अधिकतर बिन्दु-रेखीय रीति द्वारा ही किया जाता है। दो सम्बन्धित श्रेणियों के चर-मूल्यों को बिन्दुरेखा पर अंकित करने से जो विक्षेप-चित्र (scatter diagram) या बिन्दु चित्र (dot diagram) तैयार होता है तथा इस चित्र पर अंकित बिन्दुओं के मध्य से गुजरने वाली जो दो 'सर्वोपयुक्त रेखाएँ (Line of Best Fit) निर्मित होती है वास्तव में ये रेखाएँ ही प्रतीपगमन रेखाएँ कहलाती है। स्मरण रहे प्रतीपगमन रेखीय हो सकता है अथवा वक्ररेखीय। उपर्युक्त रेखाओं के सरल (straight) होने पर प्रतीपगमन रेखीय (linear) माना जाता है और अगर यह रेखाएँ सरलित वक्र (smooth curve) के रूप में हो तो प्रतीपगमन , वक्र-रेखीय (curvilinear) माना जाएगा। सर ल प्रतीपगमन रेखाओं के समीकरण एक-घातीय (equation of the first order) होते हैं अर्थात् x पर y की प्रतीपगमन रेखा का समीकरण $x = a + by$ और इसी प्रकार y पर x की प्रतीपगमन रेखा का समीकरण $y = a + xy$ होता है।

6.7 प्रतीपगमन रेखाएँ (Regression Lines)

अर्थ (Meaning) - जैसा कि ऊपर स्पष्ट किया जा चुका है कि दो समंक श्रेणियों के पारस्परिक औसत-सम्बन्ध को दर्शाने वाली सर्वोपयुक्त रेखाओं को **'प्रतीपगमन रेखाएँ'** कहते हैं। ये रेखाएँ वास्तव में किसी एक श्रेणी के मध्यम-मूल्य से सम्बन्धित दूसरी श्रेणी के सर्वोत्तम मध्यम-मूल्यों को व्यक्त करती है।

6.8 प्रतीपगमन रेखाओं के कार्य (Functions of Regression Lines)

प्रतीपगमन रेखाओं के दो महत्वपूर्ण कार्य होते हैं-

(1) सर्वोपयुक्त अनुमान - जैसा कि स्पष्ट किया जा चुका है इन रेखाओं की सहायता से एक श्रेणी के दिये हुए मूल्य के आधार पर दूसरी श्रेणी के तत्संवादी सर्वोपयुक्त औसत मूल्य का सांख्यिकीय अनुमान लगाया जा सकता है। X का Y पर (of X on Y) प्रतीपगमन रेखा से X का तथा Y की X पर (of Y on X) प्रतीपगमन रेखा द्वारा Y का सर्वोत्तम अनुमान लगाया जाता है।

(2) सहसम्बन्ध की मात्रा व दिशा का ज्ञान - प्रतीपगमन रेखाओं की सहायता से निम्नलिखित नियमों के आधार पर यह भी ज्ञात किया जा सकता है कि दोनों श्रेणियों में सहसम्बन्ध कितना और कैसा है-

(i) धनात्मक - जब दो नों प्रतीपगमन रेखाएँ रेखाचित्र पर बाएँ निचले कोने से दाहिने ऊपर के कोने की ओर (ऊर्ध्वगामी) बढ़ती है तो X और Y में धनात्मक सहसम्बन्ध होता है।

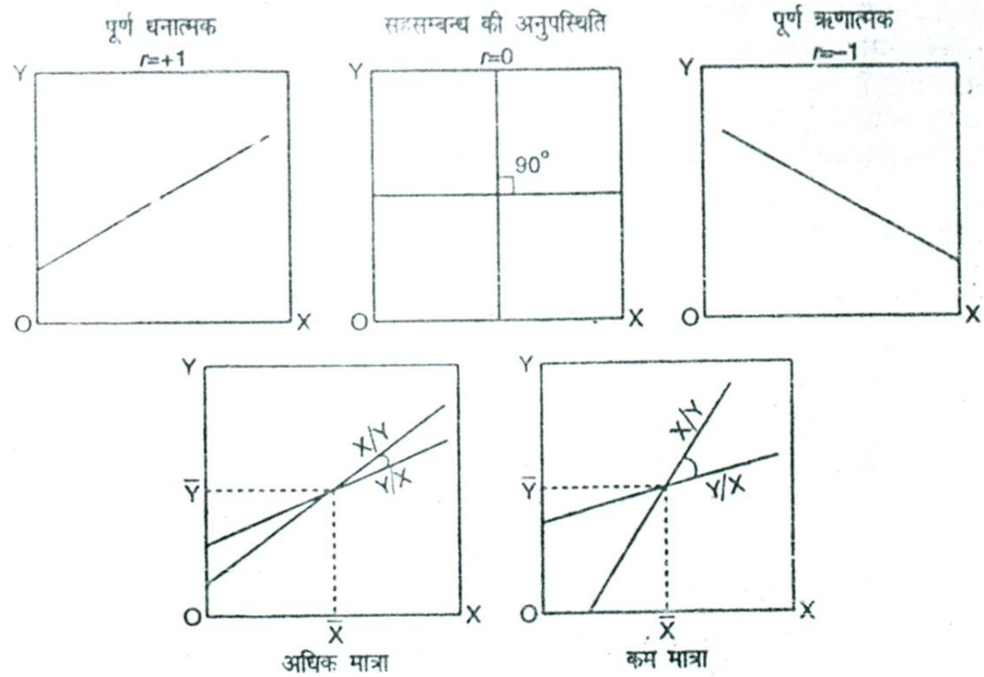
(ii) ऋणात्मक - इसके विपरीत जब ये रेखाएँ ऊपर से नीचे की ओर (अधोगामी) जाती हैं तो सहसम्बन्ध ऋणात्मक होता है।

(iii) पूर्ण सहसम्बन्ध एक रेखा - जब विक्षेप चित्र पर प्रांकित विभिन्न बिन्दु एक ही सीधी रेखा के रूप में हो तो दोनों रेखाएँ एक-दूसरें को पूरी तरह से ढक लेती है। ऐसी स्थिति में श्रेणियों में पूर्ण सहसम्बन्ध होता है। दूसरे शब्दों में X और Y में पूर्ण सहसम्बन्ध होने पर एक ही प्रतीपगमन रेखा बनती है।

(iv) सहसम्बन्ध का अभाव - यदि दो नों रेखाएँ एक दूसरे को समकोण (right angle) अर्थात् 90^o के कोण पर काटती हों तो X और Y में बिल्कुल सहसम्बन्ध नहीं पाया जाता। इस स्थिति में विक्षेप-चित्र पर विभिन्न बिन्दु चारों ओर बिखरे होते हैं तथा उनमें कोई सुनिश्चित प्रवृत्ति स्पष्ट नहीं होती।

(v)सीमित सहसम्बन्ध- दो नों प्रतीपगमन रेखाएँ एक-दूसरे के जितनी निकट होंगी , X और Y में उतना ही अधिक सहसम्बन्ध होगा। इसके विपरीत ये रेखाएँ एक-दूसरे से जितनी दूर होती जाएंगी सहसम्बन्ध की मात्रा उतनी ही कम होती जाएगी। ये रेखाएँ दो नों श्रेणियों के समान्तर माध्य के संयोग से प्रांकित बिन्दु पर एक-दूसरे को काटती है। अतः इनके सर्वनिष्ठ बिन्दु (point of intersection) से दो नों अक्षों पर डाले जाने वाले लम्ब (perpendicular) X तथा Y के समान्तर माध्य-मूल्यों को व्यक्त करते हैं।

निम्न चित्र से प्रतीपगमन रेखाओं से सम्बन्धित उपर्युक्त नियम स्पष्ट हो जाते हैं–



प्रतीपगमन रेखाओं की रचना दो रीतियों द्वारा की जा सकती है-

(क) मुक्त हस्त रीति द्वारा (By freehand method), तथा

(ख) प्रतीपगमन समीकरणों द्वारा (By regression equations)

प्रथम रीति का प्रयोग सामान्यतः नहीं किया जाता , क्योंकि इसके आधार पर विभिन्न व्यक्तियों द्वारा रेखा भिन्न-भिन्न प्रकार से खींची जा सकती है। अतः प्रतीपगमन समीकरण के आधार पर ही इन रेखाओं की रचना की जाती है।

(क) **मुक्त-हस्त वक्र विधि** - मुक्त हस्त वक्र विधि में हम सर्वप्रथम X और Y के मानों के युग्मों को प्रकीर्ण आरेख के रूप में आलेखित करते हैं। मानों के एक युग्म के लिए एक बिन्दु आलेखित किया जाता है। इसके बाद हम दो मुक्त हस्त रेखाएँ खींचते हैं। इन रेखाओं से एक रेखा इस रीति से खींची जाती है कि Y श्रेणी के उसके माध्य से धनात्मक विचलन ऋणात्मक विचलों से निरस्त हो जाते हैं। इस रेखा के एक ओर के विचलनों का योग उसके दूसरी ओर के विच लनों के योग के बराबर होता है। यह Y का X पर प्रतीपगमन रेखा (Regression line of Y on X) होगी। दूसरी प्रतीपगमन रेखा इस रीति से खींची जाएगी कि X श्रेणी के उसके माध्य से धनात्मक विच लन ऋणात्मक विचलों को निरस्त कर देंगे। इस रेखा के ओर के विच लनों का योग भी उसके दूसरी ओर के विच लनों के योग के बराबर होगा। यह प्रतीपगमन रेखा X का Y पर प्रतीपगमन रेखा (Regression Line of X on Y) कहलाएगी। दोनों प्रतीपगमन रेखाएँ दोनों श्रेणियों के माध्यों (Means) के बिन्दु पर एक दूसरे को काटेगी। यदि दोनों चरों के बीच परिपूर्ण धनात्मक या ऋणात्मक सम्बन्ध है, तो केवल एक प्रतीपगमन रेखा होगी।

मुक्तहस्त वक्र विधि से प्रतीपगमन रेखाओं का खींचना बहुत कठिन कार्य है। प्रायः प्रकीर्ण आरेख में बार-बार एक धागा इस रीति से समायोजित किया जाता है कि धनात्मक तथा ऋणात्मक विच लन एक दूसरे को निरस्त कर देते हैं। एक बार जब ये रेखाएँ खींच ली जाती हैं तो हम Y का X पर प्रतीपगमन रेखा से Y के मानों को

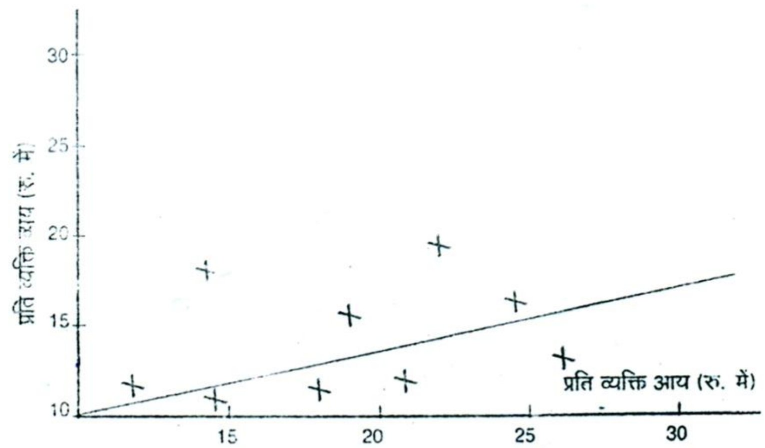
पूर्वकथित या आकलित कर सकते हैं और इसी प्रकार X का Y पर प्रतीपगमन रेखा से Y के मानों को पूर्वानुमानित या आकलित कर सकते हैं।

निम्नलिखित उदाहरण द्वारा इस विधि की व्याख्या दी जा सकती है-

उदाहरण 1- नीचे दिये गये आँकड़े 10 व्यक्तियों की प्रतिदिन की आय तथा व्यय से सम्बन्धित है , प्रतीपगमन विश्लेषण द्वारा यह ज्ञात करना है कि आय घटने या बढ़ने से व्यय किस प्रकार बढ़ता या घटता है।

प्रति व्यक्ति आय (रु0 में)	15	22	28	20	25	30	25	30	22	18
प्रति व्यक्ति व्यय (रु0 में)	12	15	20	18	15	18	12	15	15	12

व्यक्तियों की आय को स्वतंत्र मानते हुए 'X' अक्ष पर लेकर तथा व्यय को आश्रित चर मानते हुए 'Y' अक्ष पर लेकर निम्नलिखित चित्र बनाते हैं–



प्रस्तुत विक्षेप चित्र में वितरित बिन्दुओं के मध्य एक ऐसी प्रवृत्त रेखा खींची गयी है जिसके ऊपर तथा नीचे बिन्दुओं का वितरण लगभग समान है।

अब प्रत्येक 'X' चर के मान के लिए 'Y' के मान को विक्षेप चित्र द्वारा ज्ञात किया जा सकता है। परन्तु इस विधि की सीमा यह है कि भिन्न-भिन्न व्यक्ति इस चित्र में भिन्न-भिन्न प्रवृत्त रेखाएँ खींच सकते हैं तथा इस तरह 'X' चर के लिए 'Y' चर के मानों के अनुमान भी भिन्न हो सकते हैं।

6.9 प्रतीपगमन समीकरण (Regression Equation)

प्रतीपगमन समीकरण जिन्हें प्रायः Estimating Equations भी कहते हैं , प्रतीपगमन रेखाओं के बीजगणितीय स्वरूप है। प्रतीपगमन रेखाओं की भांति ये समीकरण भी दो होते हैं -

(i) X का Y पर प्रतीपगमन समीकरण (Regression equation of X upon Y)

इसकी सहायता से Y (स्वतंत्र चर-मूल्य) के दिए हुए मूल्य के तत्संवादी X (आश्रित चर-मूल्य) का सर्वोत्तम माध्य मूल्य ज्ञात किया जाता है तथा रेखाचित्र पर इस समीकरण के मूल्यों को प्रांकित करने से X की Y पर प्रतीपगमन रेखा प्राप्त हो जाती है।

(ii) Y का X पर प्रतीपगमन समीकरण (Regression equation of Y upon X)

इसके आधार पर X (स्वतंत्र चर-मूल्य) के तत्संवादी Y (आश्रित मूल्य) के सर्वोपयुक्त मूल्य का अनुमान लगाया जाता है और Y की X पर प्रतीपगमन रेखा खीची जाती है।

रेखीय प्रतीपगमन के समीकरण , सरल रेखा के समीकरण (equation of the straight line) पर आधारित है। मूल रूप में ये निम्न प्रकार है-

a. X का Y पर -- $X = a + bY$

b. Y का X पर -- $Y = a + bX$

इस समीकरणों में a और b के मान स्थिरांक (constant) है जो प्रतीपगमन रेखाओं की स्थितियों को निर्धारित करते हैं। प्राचल (parameter) 'a' जिसे अन्तः खण्ड (intercepts) भी कहते हैं, प्रतीपगमन रेखा के स्तर को प्रकट करता है अर्थात् मू ल-बिन्दु (point of origin) से (कोटि-अक्ष पर) ऊपर या नीचे रेखा की दूरी। दूसरे शब्दों में, लेखाचित्र पर मूल-बिन्दु से कोटि - अक्ष पर प्रतीपगमन रेखा के स्पर्श-बिन्दु का अन्तर ही प्राच ल 'a' का मान होता है। जब 'a' का मान धनात्मक (+) होता है तो रेखा Y - अक्ष को मू ल-बिन्दु 'o' से ऊपर की ओर स्पर्श करती है और जब 'a' का मान ऋणात्मक (-) होता है तो रेखा Y - अक्ष पर स्पर्श 'o' से नीचा होता है। यदि 'a' मान शून्य हो तो रेखा मूल-बिन्दु से ही आरम्भ होती है।

अन्तः खण्ड का बीजगणितीय माप-

प्रथम समीकरण ($X = a + bY$) में $a = \bar{X} - b\bar{Y}$

द्वितीय समीकरण ($Y = a + bX$) में $a = \bar{Y} - b\bar{X}$

$\bar{X}$ तथा $\bar{Y}$ समान्तर माध्यों के लिए प्रयुक्त किये जाते हैं।

प्राचल 'b' रेखा का ढलान या ढा ल (slope of the line) को निर्धारित करता है। प्राच ल 'b' को प्रतीपगमन गुणांक (regression coefficient) भी कहते हैं। इससे यह ज्ञात होता है कि X में इकाई का परिवर्तन (unit change) होने से Y में कितना परिवर्तन होगा और इसके विपरीत। यदि 'b' का मान धनात्मक हो तो रेखा का ढलान बाएँ से दाएँ ऊपर की ओर होगा। 'b' का मान ऋणात्मक होने पर रेखा का ढलान नीचे की ओर होगा। बीजगणितीय दृष्टि से 'b' के मान को सहसमबन्ध-गुणांक, मानक विचलन व समान्तर माध्यों के रूप में निम्न प्रकार प्रकट किया जा सकता है-

प्रथम समीकरण X का Y पर -- $b_{xy} = r\frac{\sigma_x}{\sigma_y}$

द्वितीय समीकरण Y का X पर -- $b_{yx} = r\frac{\sigma_y}{\sigma_x}$

σx व σy क्रमशः X और Y श्रेणियों के प्रमाप विच लन (Standard Deviations) है तथा r दोनों श्रेणियों का सहसम्बन्ध गुणांक है। इस विश्लेषण के आधार पर प्रतीपगमन रेखाओं को निम्न रूप में प्रस्तुत किया जा सकता है-

(i) X का Y पर प्रतीपगमन समीकरण

$$X = a + bY$$

$$X = (\bar{X} - b\bar{Y}) + bY$$

$$X - \bar{X} = bY - b\bar{Y}$$

(ii) Y का X पर प्रतीपगमन समीकरण

$$Y = a + bX$$

$$Y = (\bar{Y} - b\bar{X}) + bX$$

$$Y - \bar{Y} = bX - b\bar{X}$$

$$\left(X - \overline{X}\right) = b_{xy}\left(Y - \overline{Y}\right) \qquad \left(Y - \overline{Y}\right) = b_{yx}\left(X - \overline{X}\right)$$

$$X - \overline{X} = r\frac{\sigma_x}{\sigma_y}\left(Y - \overline{Y}\right) \qquad Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}\left(X - \overline{X}\right)$$

प्रयोग- Y से सम्बद्ध X का सर्वोपयुक्त मूल्य अनुमानित करने के लिए प्रथम समीकरण (of X upon Y) का प्रयोग किया जाता है। और X के तत्संवादी Y का सर्वोत्तम मूल्य ज्ञात करने के लिए द्वितीय समीकरण (of Y upon X) का प्रयोग किया जाता है। उपर्युक्त रूप में समीकरणों का प्रयोग तभी करना चाहिए जब प्रश्न में $\overline{X}$ और $\overline{Y}$,σx और σy तथा r के मान दिये गये हों।

यहाँ, $\overline{X}$ = X श्रेणी का समान्तर माध्य , σx = X श्रेणी का प्रमाप विचलन,

$\overline{Y}$ = Y श्रेणी का समान्तर माध्य, σy = Y श्रेणी का प्रमाप विचलन,

r = सहसम्बन्ध गुणांक

उदाहरण 2

निम्नलिखित समंकों से (i) वस्तु की ग्वालियर में सम्भाव्य कीमत ज्ञात कीजिए यदि इन्दौर में कीमत 70 रु0 हो और (ii) इन्दौर में सम्भाव्य कीमत ज्ञात कीजिए यदि ग्वालियर में कीमत 90 रु0 हो।

	माध्य	प्रमाप विचलन
इन्दौर	65	2.5
ग्वालियर	67	3.5

सहसम्बन्ध गुणांक = +0.8

हल- इन्दौर में कीमत को X तथा ग्वालियर में कीमत को Y मानने पर ज्ञात मूल्य इस प्रकार है-

$\overline{X}$ = 65, $\overline{Y}$ = 67, σx = 2.5 σy = 3.5, r = +0.8

ग्वालियर में सम्भाव्य कीमत इन्दौर में सम्भाव्य कीमत

$$Y - \overline{Y} = r\frac{\sigma_y}{\sigma_x}(X - \overline{X}) \qquad X - \overline{X} = r\frac{\sigma_x}{\sigma_y}(Y - \overline{Y})$$

$$Y - 67 = 0.8\frac{3.5}{2.5}(70-65) \qquad X - 65 = 0.8\frac{2.5}{3.5}(90-67)$$

$$Y - 67 = 1.12\ (70-65) \qquad X - 65 = 0.57(90-67)$$

$$Y = (1.12 \times 5) + 67 \qquad X = (0.57 \times 23) + 65$$

$$Y = 72.60 \qquad X = 78.11$$

अतः इन्दौर में (i) 70 रु0 होने पर उस वस्तु की ग्वालियर में सम्भाव्य कीमत 72.60 है और (ii) 75 रु0 होने पर ग्वालियर में कीमत 78.11 रु0 है।

6.10 प्रतीपगमन गुणांक (Regression Coefficients)

प्रतीपगमन समीकरण में 'b' प्रतीपगमन गुणांक का प्रतीक है , जो स्वतंत्र चर मूल्य में परिवर्तन के कारण आश्रित चर-मूल्य में होने वाले परिवर्तन की 'मात्रा' तथा 'दिशा' को बतलाता है। दूसरे शब्दों में, यह इस बात को स्पष्ट करता है कि एक श्रेणी के चर-मूल्यों में 1 का परिवर्तन (unit change) होने से दूसरी श्रेणी के चर-मूल्यों में औसतन कितना परिवर्तन होगा। इस प्रकार , यह प्रतीपगमन रेखा के ढलान (slope of regression line) का बीजगणितीय माप है। प्रतीपगमन रेखाओं ओर समीकरणों की भाँति , प्रतीपगमन गुणांक भी दो होते हैं-

(I) X का Y पर प्रतीपगमन गुणांक (II) Y का X पर प्रतीपगमन गुणांक

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} \qquad\qquad b_{yx} = r\frac{\sigma_y}{\sigma_x}$$

महत्वपूर्ण नोट- उपरोक्त गुणांकों के रूप में हम X तथा Y की दोनों प्रतीपगमन समीकरणों को निम्न रूप में भी व्यक्त कर सकते हैं-

$$X - \overline{X} = b_{xy}\ (Y - \overline{Y}) \quad \text{तथा} \quad Y - \overline{Y} = b_{yx}\ (X - \overline{X})$$

उदाहरण 3 -

(i) यदि दोनों प्रतीपगमन गुणांकों (regression coefficients) के मूल्य 0.64 और 0.81 है, तो सहसम्बन्ध गुणांक (coefficient of correlation) का मान बताइए।

(ii) निम्न आँकड़ों से - (अ) Y का प्रमाप विचलन (σy) और (ब) X और Y के मध्य सहसम्बन्ध गुणांक (rxy) ज्ञात कीजिए।

(iii). एक विद्यार्थी ने Y के X पर (Y on X) और X के Y पर (X on Y) प्रतीपगमन गुणांकों के मान क्रमशः 1.2 और 0.9 ज्ञात किये। कारण सहित बतलाइए कि क्या उसके द्वारा किया गया परिगणन सही है?

(iv)निम्न प्रदत्त सूचना से 'r' का मूल्य ज्ञात कीजिए -

X का प्रसरण (variance of X) = 2.25, σy = 4, तथा X का Y पर प्रतीपगमन समीकरण X = −0.3 + 1.8 है।

हल- (i) $r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{0.64 \times 0.81} = \sqrt{0.5184} = +0.72$

(ii) X का Y पर प्रतीपगमन Y का X पर प्रतीपगमन

X = 0.85Y Y = 0.89X

यदि Y का मूल्य 1 है तो X = 0.85 होगा यदि X = 1 तो Y का मूल्य 0.89 होगा

अतः bxy = 0.85 byx = 0.89

$r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{0.85 \times 0.89} = 0.87$

$b_{xy} = r\frac{\sigma_x}{\sigma_y}$ प्रदत्त मूल्यों को आदिष्ट करने पर-

$$0.85 = 0.87 \times \frac{3}{\sigma_y} \qquad 0.85\ \sigma_y = 2.61$$

$\sigma_y = \frac{2.61}{0.85} = 3.07 \qquad r = 0.87 \; ; \sigma_y = 3.07$

(iii) विद्यार्थी द्वारा प्राप्त परिणाम इस प्रकार हैं-

byx = 1.2, bxy = 0.9 इन दोनों गुणांकों की गुणा (r^2) $1.2 \times 0.9 = 1.08$ है जो 1 से अधिक है, इसका वर्गमूल (r) भी 1 से अधिक होगा, परन्तु यह सहसम्बन्ध गुणांक है जो कि 1 से अधिक नहीं हो सकता। अतः विद्यार्थी ने प्रतीपगमन गुणांकों की गणना में गलती की है।

(iv) X का प्रसरण- $\sigma_x{}^2 = 2.25$

$$\sigma_x = \sqrt{2.25} = 1.5; \quad \sigma_y = 4$$

$$X = -0.3Y + 1.8$$

उक्त समीकरण X का Y पर प्रतीपगमन प्रकट करता है। इसमें अन्तःखण्ड

(a) 1.8 है और X की Y पर सर्वोपयुक्त रेखा का ढाल (b) -0.3 है,

यदि प्रतीपगमन गुणांक है अर्थात् bxy = -0.3

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} \text{ या } -0.3 = r \times \frac{1.5}{4} \text{ या } -1.2 = 1.5 \times r$$

$$r = \frac{-1.2}{1.5} = -0.8$$

6.11 प्रतीपगमन गुणांकों का परिकलन (Calculation of Regression Coefficients)

यदि दो सम्बद्ध श्रेणियों के अ लग-अलग चर मूल्य दिये हुए हों तो उनके आधार पर ऊपर बताए गए सूत्रों द्वारा प्रतीपगमन गुणांक की गणना करना एक अत्यंत कठिन समस्या है क्योंकि r, σx व σy का निर्धारण करने में गणना-क्रिया अनावश्यक रूप से लम्बी हो जाती है। अतः समय व परिश्रम की बचत करने के लिए निम्न दो रीतियों (सूत्रों) का प्रयोग किया जा सकता है-

(अ) प्रत्यक्ष रीति (Product Moment Method)

दो प्रसामान्य समीकरणों पर आधारित इस रीति का सूत्र इस प्रकार है -

X का Y पर प्रतीपगमन गुणांक

$$b_{xy} = \frac{N(\sum xy) - (\sum x)(\sum y)}{N(\sum y^2) - (\sum y)^2}$$

Y का X पर प्रतीपगमन गुणांक

$$b_{yx} = \frac{N(\sum xy) - (\sum x)(\sum y)}{N(\sum x^2) - (\sum x)^2}$$

क्रिया-विधि - ध्यान रहे, इस रीति में 'विचलन' नहीं लिए जाते हैं। सर्वप्रथम, x तथा y श्रेणी के पदों का योग क्रमशः Σx तथा Σy कर लिया जाता है। फिर, अलग-अलग दोनों श्रेणियों के पदों के वर्गों का योग क्रमशः Σx^2 तथा Σy^2 निकाला जाता है। इसके बाद x तथा y श्रेणियों के तत्संवादी पदों की परस्पर गुणा करके उनका योग Σxy प्राप्त कर लिया जाता है।

विद्यार्थियों के लिए नोट - जब दोनों श्रेणियों के पदों का आकार अपेक्षाकृत छोटा हो तो इस रीति का प्रयोग काफी सुविधाजनक रहता है।

उदाहरण 4- दिया हुआ है (Given)

N =12, Σx = 120, Σy = 432, Σxy = 4992, $\Sigma x^2 = 1392$, $\Sigma y^2 = 18252$

ज्ञात कीजिए - (i) दोनों प्रतीपगमन समीकरण (ii) प्रतीपगमन गुणांक, (iii) x और y के मध्य सहसम्बन्ध गुणांक (r)।

हल- $\overline{X} = \frac{\sum x}{N} = \frac{120}{12} = 10$; $\overline{Y} = \frac{\sum y}{N} = \frac{432}{12} = 36$

प्रतीपगमन गुणांक (Regression coefficient)

x का y पर

$$b_{xy} = \frac{N(\sum xy) - (\sum x)(\sum y)}{N\sum y^2 - (\sum y)^2}$$

$$= \frac{12 \times 4992 - (120 \times 432)}{12 \times 18252 - (432)^2}$$

$$= \frac{59904 - 51840}{219024 - 186624}$$

$$= \frac{8064}{32400} = 0.249$$

y का x पर

$$b_{yx} = \frac{N(\sum xy) - (\sum x)(\sum y)}{N\sum x^2 - (\sum x)^2}$$

$$= \frac{12 \times 4992 - (120 \times 432)}{12 \times 1392 - (120)^2}$$

$$= \frac{8064}{16704 - 14400}$$

$$= \frac{8064}{2304} = 3.5$$

प्रतीपगमन समीकरण (Regression equations)

$$x - \overline{x} = b_{xy}(y - \overline{y})$$

$$x - 10 = 0.249\,(y - 36)$$

$$x = 0.249y - 8.964 + 10$$

$$x = 0.249y + 1.036$$

$$y - \overline{y} = b_{yx}(x - \overline{x})$$

$$y - 36 = 3.5(x - 10)$$

$$y = 3.5x - 35 + 36$$

$$y = 3.5x + 1$$

सहसम्बन्ध गुणांक (Coefficient Correlation)

$$r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{0.249 \times 3.5} = \sqrt{0.8715} = 0.93$$

6.12 सारांश (Summary)

प्रतीपगमन वह सांख्यिकीय तकनीक है जो दो या अधिक चरों के मध्य औसत सम्बन्ध को प्रदर्शित करता है तथा इसके द्वारा एक चर के ज्ञात मूल्य के आधार पर दूसरे चर के लिए संभावित मूल्य का अनुमान लगाया जा सकता है।

सहसम्बन्ध से दो चरों के मध्य कारणात्मक सम्बन्ध ज्ञात नहीं किया जा सकता है , किन्तु प्रतीपगमन द्वारा यह ज्ञात करना सर ल है कि कौन सा चर 'कारण' है तथा कौन सा 'प्रभाव' है। अर्थात् प्रतीपगमन उस कार्यमूलक सम्बन्ध को बताता है जिसके द्वारा एक चर के अनुमान दूसरे चर से निकाले जा सकते हैं।

प्रतीपगमन का विश्लेषण दो चरों के मध्य सहसम्बन्ध को प्रस्तुत करने वाले विक्षेप चित्र द्वारा किया जाता है जिसकी मुख्यतः दो विधियाँ प्रचलित है- मुक्त हस्त आरेख तथा न्यूनतम वर्ग विधि।

दो श्रेणियों के पारस्परिक माध्य सम्बन्ध को प्रकट करने वाली सर्वोपयुक्त रेखाओं को प्रतीपगमन रेखाएँ कहा जाता है। ये रेखाएँ एक श्रेणी के माध्य मूल्यों से सम्बन्धित दूसरे सर्वोत्तम माध्य मूल्यों को व्यक्त करती हैं। दो सम्बन्धित श्रेणियों के लिए दो प्रतीपगमन रेखाएँ होती हैं।

प्रतीपगमन गुणांक वह अनुपात है जो यह दर्शाता है कि स्वतंत्र चर की श्रेणी में इकाई परिवर्तन होने पर आश्रित चर के मूल्यों में औसत परिवर्तन दर क्या होगी। वस्तुतः प्रतीपगमन गुणांक , प्रतीपगमन रेखा के ढाल द्वारा प्रदर्शित होता है।

6.13 अभ्यास प्रश्न (Practice Questions)

(A) वस्तुनिष्ठ प्रश्न-

(I)निम्नलिखित में से कौन सा सही है?

(1).प्रतीपगमन विश्लेषण मापन करता है-

(i) x और y श्रेणियों के बीच सहविचरणता के परिमाण का

(ii) y श्रेणी के विचरण का

(iii) x श्रेणी के विचरण का

(iv) x और y श्रेणियों के बीच फलनिक सम्बन्ध का

(2).प्रतीपगमन रेखाएँ एक दूसरे को-

(i) x और y के माध्य (ii) केवल x के माध्य

(iii) केवल y के माध्य

(iv) x और y की माध्यिका के बिन्दु पर काटता है।

(3) यदि उस बिन्दु से, जहाँ दोनों प्रतीपगमन रेखाएँ एक-दूसरे को काटती है X - अक्ष पर लम्ब खींचा जाय तो हमें प्राप्त होगा-

(i) x का समान्तर माध्य (ii) y का समान्तर माध्य

(iii) x का बहुलक मान (iv) x का माध्यिका मान

(4) दो चरों की दशा में केवल एक प्रतीपगमन रेखा होगी यदि-

(i) r या तो +1 हो या -1 (ii) r = + ही हो

(iii) r = 0 हो (iv) r = -1 ही हो

(5) जब एक प्रतीपगमन गुणांक धनात्मक होता है तो दूसरा होगा-

(i) ऋणात्मक (ii) शून्य (iii) धनात्मक (iv) इनमें से कोई नहीं

(II).निम्नलिखित कथनों में से कौन सा सत्य/असत्य है

(i) यदि bxy धनात्मक है तो byx भी धनात्मक होगा। (स / अ)

(ii) यदि दोनों प्रतीपगमन गुणांक ऋणात्मक है तो सहसम्बन्ध गुणांक धनात्मक होगा। (स / अ)

(iii) दोनों प्रतीपगमन रेखाएँ एक दूसरे को ढक लेती हैं जब चरों के मध्य सहसम्बन्ध या तो पूर्ण धनात्मक अथवा पूर्ण ऋणात्मक होता है। (स / अ)

(iv) प्रतीपगमन विश्लेषण दो चरों के बीच कारण-प्रभाव का सम्बन्ध नहीं प्रकट करता है। (स / अ)

(v) स्थिरांक ‘b’ प्रतीपगमन रेखा के ढलान को दर्शाता है। (स /अ)

(III).निम्नलिखित को पूर्ण कीजिए-

a) सहसम्बन्ध गुणांक ----- के गुणनफल का वर्गमूल होता है।

b) प्रतीपगमन रेखाएँ ----- की कल्पना पर खींची जाती है।

c) दोनों ----- गुणांकों के एक समान चिन्ह होने चाहिए।

d) यदि x के y पर प्रतीपगमन गुणांक का मान 1.4 है, तो y के x पर प्रतीपगमन गुणांक का मान ----- नहीं होगा।

e) प्रतीपगमन विश्लेषण दो चरों के बीच ----- का मापन करता है।

(IV) लघु-उत्तरात्मक प्रश्न-

1) प्रतीपगमन क्या है? आर्थिक विश्लेषण में इसकी उपयोगिता की व्याख्या कीजिए।
2) प्रतीपगमन गुणांकों की क्या विशेषताएँ है?
3) सहसम्बन्ध और प्रतीपगमन विश्लेषण में अन्तर स्पष्ट कीजिए।
4) प्रतीपगमन समीकरणों में अचरांक ‘b’ के अर्थ की उदाहरण सहित व्याख्या कीजिए।
5) ’विचरण अनुपात की अवधारणा‘ की सोदाहरण व्याख्या कीजिए।

(V) निबन्धात्मक प्रश्न-

1) प्रतीपगमन अवधारणा की व्याख्या कीजिए। यह सहसम्बन्ध से किस प्रकार भिन्न है? प्रतीपगमन रेखाएँ दो क्यों होती है? किन परिस्थितियों में केवल एक ही प्रतीपगमन रेखा हो सकती है।
2) प्रतीपगमन रेखा किसे कहते हैं? इसे मापने की विधि स्पष्ट कीजिए।
3) प्रतीपगमन विश्लेषण से आप क्या समझते हैं ? प्रतीपगमन विश्लेषण की व्यावसायिक निर्णयों में उपयोगिता की उदाहरण सहित व्याख्या कीजिए।
4) सहसम्बन्ध तथा प्रतीपगमन का अर्थ तथा आर्थिक विश्लेषण में इनकी उपयोगिता बताइए। प्रतीपगमन समीकरण किस प्रकार निकाले जा सकते हैं? उदाहरण देकर समझाइए।
5) यदि r सहसम्बन्ध गुणांक है तो r^2 आश्रित चर में कु ल विचरण का अनुपात है जिसका स्पष्टीकरण प्रतीपगमन विश्लेषण से होता है।

(VI) संख्यात्मक प्रश्न-

(1) निम्न समंकों से y का अनुमानित मूल्य निकालिए यदि ग त्र 70 और x का अनुमानित मूल्य निकालिए यदि y = 90

	x	y
औसत मूल्य	18	100
प्रमाप विचलन	14	20

x और y में सहसम्बन्ध गुणांक त्र ़0.08

(2) दो प्रतीपगमन समीकरण ज्ञात कीजिए तथा y का सम्भावित मूल्य ज्ञात कीजिए जबकि ग त्र 55 है।
दिया हुआ है -

(3) निम्न आँकड़ों से दोनों प्रतीपगमन गुणांक ज्ञात कीजिए -

X रू	8	6	4	7	5
Y रू	9	8	5	6	2

(4) नीचे पतियों तथा पत्नियों की आयु दी गयी है। ज्ञात कीजिए -
(अ) दो प्रतीपगमन समीकरण, (ब) सहसम्बन्ध गुणांक, तथा

(स) पति की सम्भावित आयु जबकि पत्नी की आयु 25 वर्ष है-

पतियों की आयु: 22 23 23 24 26 27 27 28 30 30

पत्नियों की आयु: 18 20 21 20 21 22 23 24 25 26

(5) निम्नलिखित समंकों से बिक्री और लाभ में विचरण-अनुपात ज्ञात कीजिए।

वर्ष:	1987	88	89	90	91	92	93	94	95	96	97	98
बिक्री (करोड़ में):	36	42	33	30	24	21	27	31.5	25.5	28.5	34.5	27
लाभ (करोड़ में):	21	26	24	23	15	14	18	19	17	21	22	20

6.14 अभ्यास प्रश्नों के उत्तर (Answers to Practice Questions)

(A) वस्तुनिष्ठ प्रश्न-

(I) – (iv) (2) – (i) (3) – (i) (4) – (i) (5) – (iii)

(II) (1) – (iii) (2) – (1) (3) – (iii) (4) – (i) (5) – (iii)

(III) (1) दो प्रतीपगमन गुणांकों (2) - न्यूनतम वर्गों ;3द्ध .

(3) प्रतीपगमन (4) - 0.714 से अधिक (5) - फलनिक सम्बन्ध

(VI) संख्यात्मक प्रश्नों के उत्तर

(1) [Y_{70} = 105.94, X_{90} = 17.44, X =0.056 + 12.4, Y = 0.1143 + 97.94]

(2) [x = 0.64y + 12.8, y = x + 7 , Y_{55} = 62]

(3) [bxy = 0.4, byx = 1.2]

(4) [x = त्र 1.107y + 1.646 , y = 0.816x + 0.784, r = +0.95, X = 29.32]

(5) [R.v. = 0.34]

6.15 संदर्भ ग्रन्थ सूची (Bibliography)

1) बंसल, एस. एन., एवं अग्रवाल, डी. आर., (1978), सांख्यिकी के मूल तत्व, शिवलाल अग्रवाल एण्ड कम्पनी, आगरा - 31

2) नागर, कैलाश नाथ, (2005), सांख्यिकी के मूल तत्व, मिनाक्षी प्रकाशन, मेरठ।

3) लाल. एस. एन., चतुर्वेदी, एस., सांख्यिकी, प्रकाशन, इलाहाबाद।

4) Gupta, S.P. (2005), Statistical Methods, S. Chand, New Delhi.

5) Goon, Gupta and Dasgupta, A Fundamental of Statistics, Vol. I, The World Press Private Limited.

इकाई 7 आन्तरगणन एवं बाह्यगणन
(Interpolation counting and Extrapolation counting)

7.1 प्रस्तावना (Introduction)

सामान्यतया जो सांख्यिकीय आँकड़े मिलते हैं वे विभिन्न प्रकार के होते हैं समंक श्रेणी पूर्ण नहीं होती है। सांख्यिकीय विश्लेषण करते समय कभी-कभी यह देखने में आता है कि प्रस्तुत समंक श्रेणी पूर्ण न होकर अपूर्ण होती है अर्थात् श्रेणी के कुछ मूल्य किन्हीं कारणों से अज्ञात बने रहते हैं। ऐसा हो सकता है कि उस अवधि के लिए समंक उपलब्ध ही न हों, ऐसा भी हो सकता है कि आँकड़ों का इकट्ठा करना इतना खर्चीला तथा जटिल हो, जैसे जनगणना को ही लीजिए भारतवर्ष में जनगणना (Census) का कार्य 10 वर्षों के अन्तराल पर किया जाता है, इसका कारण यह है कि देशव्यापी स्तर पर जनगणना का कार्य अत्यधिक खर्चीला है- इसके लिए विशाल मात्रा में संगणकों (investigators), विशेषज्ञों, संसाधनों की आवश्यकता होती है। इसके अतिरिक्त जनगणना से हमें इतनी बड़ी संख्या में वृहद् प्रकार के आँकड़े प्राप्त होते हैं कि इनका विश्लेषण करने में (कम्प्यूटरों की सहायता लेने पर भी) प्रचुर समय लगता है। उपर्युक्त व्याख्या से स्पष्ट हो जाता है कि यह कार्य निश्चित ही प्रत्येक वर्ष करना संभव नहीं है। ऐसी स्थिति में विभिन्न दस-वर्षीय समयावधियों के अन्दर किसी वर्ष की जनसंख्या अनुमानित करने की आवश्यकता पड़ने पर हम अन्तरगणन (interpolation) की सहायता लेते हैं।

अर्थात् कुछ सुनिश्चित मान्यताओं एवं सीमाओं के अन्तर्गत ज्ञात समंको के आधार पर समंक-श्रेणी के बीच किसी अज्ञात मूल्य का सर्वोत्तम सम्भाव्य अनुमान लगाने की क्रिया को आन्तरगणन कहते हैं। उदाहरण के लिए, यदि हमें 1971, 1981, 1991, 2001, 2011 की जनसंख्या दी हो और 2006 की जनसंख्या का अनुमान लगाना हो तो यह समस्या आन्तरगणन की समस्या होगी। इस प्रकार सांख्यिकीय विश्लेषण करते समय बहुत सी समंकमालाएँ पूर्ण नहीं होती और उनके कुछ मूल्य अज्ञात रह जाते हैं। ऐसा भी हो सकता है कि आँकड़े तो इकट्ठे किए गए हों पर किसी कारण से अपठनीय या अनुपलब्ध हो। इन अज्ञात समंकों को जानने की भी कभी-कभी आवश्यकता महसूस होती है। उदाहरण के लिए मान लीजिए हमें 1951, 1961, 1971, 1981 तथा 1991 के भारतीय जनसंख्या सम्बन्धी आँकड़े उपलब्ध है , पर हमें 1985 में प्रतिव्यक्ति आय ज्ञात करने के लिए जनसंख्या की आवश्यकता है, इतना ही नहीं ऐसा भी हो सकता है कि हमें 1991 अब के बाद पाँच वर्षों के लिए जनकल्याण नीति का निर्धारण करना है , इसलिए बाद की अवधि के लिए विभिन्न वर्षों में हमें जनसंख्या के भावी अनुमान की आवश्यकता हो सकती है। उपलब्ध ज्ञात समंकों के आधार पर समंक माला के इन अज्ञात राशियों के सांख्यिकीय अनुमान ज्ञात करने की क्रिया को हम आन्तरगणन तथा बाह्यगणन कहते हैं। जब हम कुछ निश्चित परिकल्पनाओं तथा मान्यताओं के अन्तर्गत समंक माला के ज्ञात समंकों के आधार पर समंक श्रेणी के भीतर के किसी अज्ञात समंक का सम्माक सर्वोत्तम अनुमान करते हैं तो इस क्रिया को बाह्यगणन कहते हैं , वस्तुतः दोनों की क्रियाएं ज्ञात समंकों के आधार पर समंकमाला के अज्ञात समंकों के अनुमान की क्रियाएं हैं और सांख्यिकीय दृष्टिकोण से दोनों क्रियाओं में विशेष अन्तर नहीं होता , जैसा हम आगे देखेंगे , दोनों के सम्बन्ध में एक ही सांख्यिकीय विधि का प्रयोग किया जाता है। इस प्रकार आन्तरगणन और बाह्यबणन में मौलिक अन्तर यह है कि पहले हम चर-मूल्य की दी हुई सीमाओं के अन्तर्गत अज्ञात मूल्य की गणना करते हैं और बाह्यगणन में इन सीमाओं के बाहर किसी मूल्य की गणना की जाती है। निम्न उदाहरण से इन दोनों क्रियाओं का अन्तर स्पष्ट हो जाएगा -

भारत की जनसंख्या

जनगणना वर्ष	:	1941	1951	1961	1971	1981	1991

जनसंख्या (करोड़ों में) : 31.9 36.1 43.9 54.8 68.3 84.6

उपर्युक्त सारणी में दिये गए जनसंख्या में समंकों के आधार पर कुछ मान्यताओं के अन्तर्गत यदि हमें 1941 और 1991 के बीच के किसी वर्ष जैसे 1947, 1975 या 1986 में भारत की जनसंख्या का सर्वोत्तम अनुमान प्राप्त करना हो तो सम्बन्धित क्रिया आन्तरगणन कहलाएगी। इसके विपरीत , उपलब्ध आँकड़ों के आधार पर 1939 (1941 से पहले) या 1999 या 2011 (1991 के बाद के किसी वर्ष) के लिए जनसंख्या का सर्वोपयुक्त अनुमान लगाने की क्रिया को बाह्यगणन कहा जाएगा। सांख्यिकीय दृष्टिकोण से आन्तरगणन व बाह्यगणन का अन्तर को ई विशेष महत्त्व नहीं रखता क्योंकि दोनों क्रियाओं के लिए एक सी रीतियों का ही प्रयोग किया जाता है।

यदि हमें दो चर-मूल्य (variable) x और y दिए हो तथा $y = f(x)$ जहाँ x स्वतंत्र चर-मूल्य (independent variable) और y आश्रित चर-मूल्य (dependent variable) है। किसी निश्चित अन्तराल में x के कुछ मूल्यों के सापेक्ष y के मूल्य दिया हों और यदि हम x के किसी मूल्य जो इसी अन्तराल में हो के सापेक्ष y का मूल्य ज्ञात करना चाहें तो यह क्रिया पूर्व निर्धारित मान्यताओं के अन्तर्गत गणितीय सूत्रों की सहायता से की जाएगी। इसी क्रिया को आन्तरगणन कहते हैं। यदि हम x के किसी मूल्य , जो अन्तराल के बाहर हो , के सापेक्ष y का मूल्य ज्ञात करें तो यह क्रिया बाह्यगणन कहलाती है। आन्तरगणन एवं बाह्यगणन के द्वारा हम किसी चर का वास्तविक मूल्य ज्ञात नहीं करते बल्कि इसके सन्निकट मान (best estimate) का अनुमान लगाते हैं। यह आकलन (estimation) की प्रक्रिया (process) सांख्यिकीय विश्लेषण में बहुत महत्वपूर्ण है और अग्रलिखित मान्यताओं पर आधारित है।

7.2 उद्देश्य (Objective)

कुछ सुनिश्चित परिकल्पनाओं के अन्तर्गत, ज्ञात समंकों के आधार पर समंक-श्रेणी के बीच किसी अज्ञात मूल्य का सर्वोत्तम सम्भाव्य अनुमान लगाना या उपलब्ध सांख्यिकीय तथ्यों के आधार पर, विशेष परिकल्पनाओं के अधीन किसी भावी समंक के पूर्वानुमान प्राप्त करना।

7.3 परिभाषा (Definition)

(i) "एक सांख्यिकीय अनुमान , अच्छा हो या बुरा , ठी हो या गलत , परन्तु प्रायः प्रत्येक दशा में वह एक आकस्मिक प्रेक्षक के अनुमान से अधिक ठीक होगा।" **-डा. ए. एल. वाउले**

(ii) "किन्हीं निश्चित मान्यताओं के अन्तर्गत मात्राओं के सर्वाधिक सम्भाव्य अनुमान लगाने की तकनीक को आन्तरगणन कहते हैं।" **–प्रो. डी. एन. एल्हांस**

(iii) "दो अन्त बिन्दुओं के बीच के स्थित मूल्यों को ज्ञात करने की क्रिया आन्तरगणन तथा इन दोनों बिन्दुओं के बाहर के मूल्यों को ज्ञात करने की क्रिया को बाह्यगणन कहते हैं।" **-डब्ल्यू. एम. हाप्रर**

7.4 मान्यतायें (Assumptions)

ऊपर दी गयी परिभाषा से स्पष्ट है कि इनकी क्रिया कुछ मान्यताओं व परिकल्पनाओं पर निर्भर करती है जिनके अभाव में अज्ञात मूल्यों का अनुमान लगाना सम्भव नहीं हो पाता। यह मान्यताएँ निम्नलिखित है-

(अ) आकस्मिक उतार-चढ़ाव न होना (No Sudden or Violent Fluctuations)

आन्तरगणन व बाह्यगणन की पहली मान्यता यह है कि विचारणीय अवधि के विभिन्न समंकों में कोई अप्रत्याशित परिवर्तन अर्थात् अत्यधिक वृद्धि या अत्यधिक कमी नहीं हुई है। सरल शब्दों में, विचाराधीन अवधि

एक सामान्य अवधि है और इस अवधि में समंकों की प्रवृत्ति नियमित और निरन्तर है अर्थात् इस अवधि में किसी प्रचण्ड उथल-पुथल या परिवर्तन का अनुभव नहीं होता। उदाहरण के लिए , यदि हमें 1961, 1971, 1981 और 1991 के किसी नगर में दिए हुए जनसंख्या-समंकों के आधार पर उसकी 1988 की जनसंख्या का आन्तरगणन करना हो, या 2001 के लिए पूर्वानुमान लगाना हो तो यह मानना पड़ेगा कि उक्त वर्ष प्रसामान्य (normal) थे और बाढ़ , युद्ध, अकाल, शरणार्थियों का भारी संख्या में आगमन आदि कारणों से उन वर्षों की जनसंख्या में एकदम कोई बहुत अधिक कमी या वृद्धि नहीं हुइ थी।

(ब) परिवर्तनों में एकरूपता या नियमितता का पाया जाना (Uniformity or regularity in changes)

दूसरी मान्यता यह है कि समंकों में होने वाले परिवर्तन प्रत्येक अवधि में नियमित रूप से तथा लगभग समान दर से होते हों अर्थात् इस अवधि में जो परिवर्तन होते हैं वे समान (uniform) है। उपर्युक्त उदाहरण में हमारी यह भी मान्यता रहेगी कि 1988 से पहले के तथा बाद के वर्षों में जनसंख्या लगभग एक ही समान गति से लगातार बढ़ रही है।

(स) पद-श्रेणियों में पारस्परिक सम्बन्ध (Mutual Inter-dependence)

यह भी आवश्यक है कि दोनों पद-श्रेणियाँ परस्पर सम्बन्धित हो जिसमें एक स्वतंत्र श्रेणी हो तो दूसरी उस पर आश्रित हो।

7.5 उपयोगिता / महत्व (Importance)

किसी समंक माला की अज्ञात राशियों के आन्तरगणन व बाह्यगणन या पूर्वानुमान का महत्त्व अनेक विषयों में दिखायी पड़ता है पर अर्थशास्त्र , व्यापार, तथा व्यवसाय तथा जनांकिकी के क्षेत्र में इनका विशेष महत्त्व है। अज्ञात राशियों के आन्तरगणन या बाह्यगणन की आवश्यकता हमें निम्न परिस्थितियों में होती हैं-

(i) केन्द्रीय प्रवृत्ति की माप- जब सांख्यिकीय आँकड़े वर्गान्तर तथा वर्ग आवृत्ति के रूप में उपलब्ध हों तो माध्यिका तथा भूयिष्ठक की गणना के लिए आन्तरगणन की विधि का प्रयोग आवश्यक हो जाता हैं। इस प्रकार की क्रिया कुछ निश्चित परिकल्पनाओं के अन्तर्गत की जाती है।

(ii) मध्यवर्ती वर्षों के लिए अनुमान-आन्तरगणन विधि का प्रयोग मध्यवर्ती वर्षों अर्थात् एकत्रित समंकों के बीच की किसी अवधि से सम्बद्ध समंकों का अनुमान लगाने के लिए किया जाता है। उदाहरणार्थ, भारत में जनगणना प्रत्येक दशक (10 वर्षों) में एक बार की जाती है। चूंकि अत्यधिक व्यय के कारण जनगणना का कार्य प्रतिवर्ष नहीं किया जा सकता , अतः जनगणनाओं के उपलब्ध समंकों के आधार पर विभिन्न मध्यवर्ती वर्षों (intercensal year) की जनसंख्या का अनुमान आन्तरगणन द्वारा लगा दिया जाता है।

(iii) समंकों का नष्ट होना या खो जाना - कभी-कभी एकत्रित समंकों में से कुछ आवश्यक समंक खो जाते हैं या नष्ट हो जाते हैं और उनका दुबारा संकलन करना या तो अत्यधिक कठिन होता है या असम्भव। ऐसी दशाओं में , उपलब्ध शेष समंकों के आधार पर रिक्त स्थानों की पूर्ति आन्तरगणन द्वारा की जा सकती है।

(iv) समंकों का अभाव या अपर्याप्तता - कुछ दशाओं में भूतकालीन समंक या तो एकत्र ही नहीं किए जाते या यदि एकत्र भी किये गए हो तो वे सही परिणाम निकालने के लिए सर्वथा अपर्याप्त होते हैं। इस अभाव (gap in coverage) या अपर्याप्तता (inadequacy) की पूर्ति आन्तरगणन द्वारा सर्वोपयुक्त अनुमान लगाकर की जाती है।

(v) भावी अनुमान - समय-समय पर आर्थिक , व्यावसायिक एवं राजकीय क्षेत्रों में विभिन्न उद्देश्यों के लिए भूतकालीन व वर्तमान उपलब्ध सामग्री के आधार पर बाह्यगणन की रीति द्वारा भविष्यकालीन समकों के

पूर्वानुमान लगाने पड़ते हैं। विशेष रूप से आर्थिक नियोजन में बाह्यगणन की रीति का काफी प्रयोग किया जाता है।

(vi) तुलनात्मक अध्ययन हेतु - जब कभी कुछ समस्याओं से सम्बन्धित विभिन्न देशों के समंक अलग-अलग कालों के लिए उपलब्ध हों तो उनका तुलनात्मक अध्ययन करना सम्भव नहीं हो पाता है। अतः ऐसी स्थिति में समंकों को तुलनायोग्य बनाने के लिए आन्तरगणन व बाह्यगणन का सहारा लेना पड़ता है। उदाहरण के लिए अमेरिका में जनगणना 1980 में और भारत में 1981 में की गयी। चूंकि दोनों देशों के जनगणना समंकों की अवधि अलग-अलग है इसलिए तुलना करने के लिए या तो भारत की 1980 की जनसंख्या का आन्तरगणन करना होगा अथवा अमेरिका की 1981 की जनसंख्या का बाह्यगणन करना होगा।

(vii) स्थान सम्बन्धी माध्यों का निर्धारण - एक अविच्छिन्न श्रेणी , भूयिष्ठक, मध्यका आदि स्थानिक माध्यों (averages of position) के मूल्यों का निर्धारण करने के लिए भी आन्तरगणन रीति का प्रयोग किया जाता है।

निम्नांकित परिस्थितियों में भी अज्ञात राशियों को ज्ञात करने के लिए आन्तरगणन तथा बाह्यगणन तकनीक की आवश्यकता पड़ती है-

(अ) हो सकता है कि हम दो देशों की प्रगति का विश्लेषण कर रहे हों पर दोनों के एक समयावधि से समबन्धित आँकड़े न उपलब्ध हों , ऐसी स्थिति में तुलनात्मक अध्ययन के लिए यह आवश्यक है कि दोनों के सम्बन्ध में आँकड़े एक ही समायावधि से सम्बन्धित प्राप्त किए जाएँ। मान लीजिए हम भारत के औद्योगिक उत्पादन की तुलना जापान के साथ करना चाहते है पर भारत में उपलब्ध आँकड़ा 1985 का है , पर जापान का आँकड़ा 1987 का है। ऐसी स्थिति में तुलनात्मक अध्ययन के लिए यह आवश्यक है कि हम भारतीय आँकड़ों के आधार पर 1987 के आँकड़े का आन्तरगणन करें।

(ब) देश की 11वीं योजनाओं की रूपरेखा तैयार करने के लिए यह आवश्यक है कि योजना बनाने से सम्बन्धी भावी आँकड़े ज्ञात हों, इन अज्ञात समंकों के ज्ञान के बिना योजनाएं नहीं बनायी जा सकती।

आन्तरगणन व बाह्यगणन की क्रियाओं का सभी क्षेत्रों में अत्यधिक महत्त्व है। इन विधियों द्वारा प्राप्त आकलनों का प्रशासकों , व्यापारियों, समाजशास्त्रियों, अर्थशास्त्री, नियोजन-विशेषज्ञ, राजनीतिज्ञ, शासक, समाज-सुधारक तथा वैज्ञानिकों के लिए बड़ा व्यावहारिक महत्त्व है। उद्योग एवं व्यापार अनुमानों पर आधारित होते हैं। विश्वसनीय अनुमान लगाने के लिए उद्योगपतियों एवं व्यापारियों द्वारा आन्तरगणन वं बाह्यगणन का प्रयोग किया जाता है। एक वित्तमंत्री अपने बजट सुझावों तथा अनुमानों को इन्हीं आकलनों के आधार पर बनाता है। इसी प्रकार बाह्यगणन का भी व्यापारिक पूर्वानुमान (business forecasting) में बहुत अधिक महत्त्व है।

7.6 आन्तरगणन एवं बाह्यगणन की परिशुद्धता (Accuracy of interpolation and extrapolation)

आन्तरगणन व बाह्यगणन की क्रियाएं उपर्युक्त दो महत्त्वपूर्ण मान्यताओं के आधार पर की जाती है। अतः उनके द्वारा ज्ञात अनुमान यथोचित रूप से ही परिशुद्ध होते हैं। परन्तु यह ध्यान रखना चाहिए कि वे अनुमान-मात्र है। अतः वे वास्तविक समंकों की भाँति परिशुद्ध नहीं हो सकते। यदि आधारभूत मान्यताएँ पूरी नहीं होती तो आन्तरगणन व बाह्यगणन द्वारा प्राप्त सम्भाव्य अनुमान भी भ्रमात्मक और अशुद्ध होते हैं।

डा. बाउले के अनुसार आन्तरगणन की परिशुद्धता निम्न दो बातों पर निर्भर है-

(i) समंकों के सम्भाव्य उच्चावचनों का ज्ञान - दिए हुए समंकों से होने वाले उतार-चढ़ाव के सम्बन्ध में जितनी अधिक जानकारी होगी, आन्तरगणित मूल्यों में उतना अधिक यथार्थता व विश्वसनीयता का अंश होगा। यदि ज्ञात समंकों में लगभग नियमित रूप से उच्चावचन होते हैं तो अज्ञात मूल्य का अनुमान भी यथासम्भव परिशुद्ध होता है।

(ii) समंकों से सम्बन्धित घटनाओं का ज्ञान - यदि सांख्यिकी को उपलब्ध समंकों पर प्रभाव डालने वाली महत्त्वपूर्ण घटनाओं का भी यथेष्ट ज्ञात है, तो वह सभी तथ्यों को ध्यान में रखते हुए आन्तरगणित मूल्यों में आवश्यक संशोधन करके उन्हें अधिक शुद्ध बना सकता है। उदाहरणार्थ, 1947 में भारत की जनसंख्या का आन्तरगणन करते समय देश के विभाजन के कारण उत्पन्न घटनाओं (जैसे शरणार्थियों का भारी संख्या में आना, साम्प्रदायिक दंगे आदि) के आधार पर अनुमानित संख्या में आवश्यक संशोधन कर देने से उसकी शुद्धता अधिक हो जाएगी।

उपर्युक्त दो बातों के अतिरिक्त आन्तरगणित मूल्यों की यथार्थता बहुत कुछ उपयुक्त रीति के प्रयोग पर भी निर्भर करती है। अतः उपयुक्त रीति का चुनाव बहुत महत्त्वपूर्ण है।

7.7 आन्तरगणन एवं बाह्यगणन की विधियाँ (Methods of interpolation and extrapolation)

आन्तरगणन एवं बाह्यगणन की विधियों को मुख्य रूप से दो भागों में बाँटा जा सकता है-

(i) बिन्दु रेखीय या ग्रैफिक विधि (Graphic method)

(ii) बीजगणितीय विधियाँ (Algebraic method)

7.7.1 बिन्दु रेखीय या गैफिक विधि (Graphical Method)

आन्तरगणन एवं बाह्यगणन की यह सबसे सरल रीति है और सब प्रकार के समंकों पर लागू होती है। स्वतंत्र चर मूल्य (independent variable x) को X अक्ष पर प्रदर्शित किया जाता है। रेखा चित्र पर बिन्दु (plot) कर लिये जाते हैं अर्थात् x के सापेक्ष दिए हुए y के मूल्यों को प्रांकित कर लिया जाता है और इन बिन्दुओं को मिला दिया जाता है। x के जिस मूल्य के सापेक्ष y का मूल्य ज्ञात करना हो, वहाँ से एक लम्ब उस वक्र पर डाला जाता है जो बिन्दुओं के मिलाने से प्राप्त हुआ है। यह लम्ब वक्र को जिस बिन्दु पर काटे वहाँ एक लम्ब Y अक्ष पर डाला जाता है और इस मूल्य की गणना कर ली जाती है। यही अभीष्ट आकलन है।

बाह्यगणन करते समय वक्र को आगे बढ़ाया जाता है और फिर x के जिस मूल्य के सापेक्ष y का मूल्य ज्ञात करना हो, वहाँ से इस वक्र पर लम्ब डाला जाता है। इस विधि में बाह्यगणन की अपेक्षा आन्तरगणन अधिक शुद्धता से प्राप्त किया जाता है।

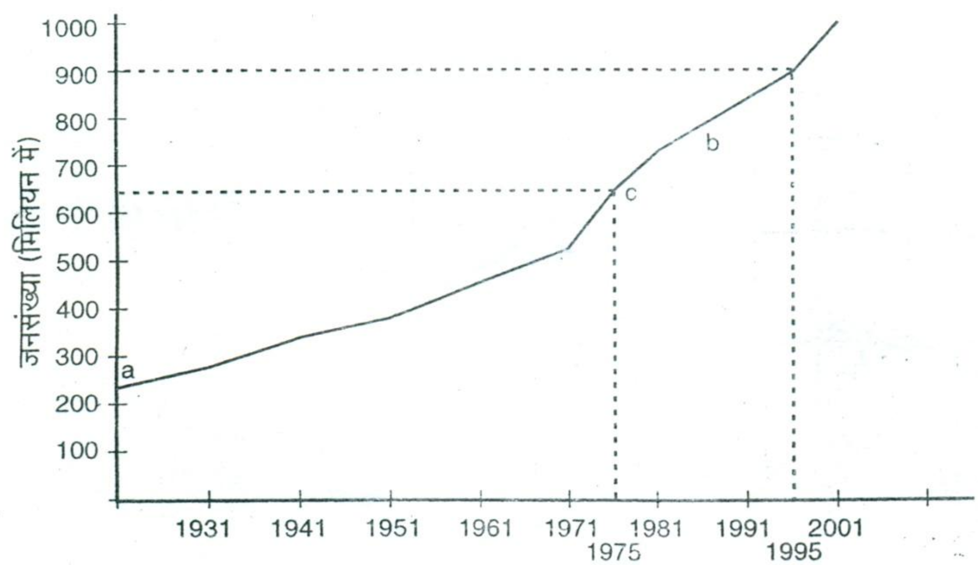
नीचे दिये गये एक उदाहरण के द्वारा इसे और स्पष्ट किया गया है।

उदाहरण (Illustration): 1

नीचे दी गयी सारिणी में भारत में जनगणना के परिणाम दिए हुए हैं जिसके आधार पर 1975 की जनसंख्या का आन्तरगणन तथा 1995 की जनसंख्या का बाह्यगणन ज्ञात कीजिए।

जनगणना वर्ष	1931	1941	1951	1961	1971	1981	1991
जनसंख्या (मिलियन में)	279	319	361	439	548	863	844

हल- ऊपर दी गयी सारिणी में विभिन्न जनगणना से सम्बन्धित वर्षों से सम्बन्धित जनसंख्या दी गयी है , जिसके आधार पर हमें 1975 वर्ष के लिए जनसंख्या का आन्तरगणन तथा 1995 वर्ष के लिए जनसंख्या का बाह्यगणन करना है। सारिणी में 'वर्ष' स्वतंत्र चर तथा उससे सम्बन्धित जनसंख्या आश्रित चर है। स्वतंत्र चर या वर्षों को X अक्ष पर प्रदर्शित किया गया तथा आश्रित चरों को Y अक्ष पर प्रदर्शित किया गया है। X से सम्बन्धित चरों को ग्राफ पर अंकित करके ab वक्र प्राप्त की गयी है जो विभिन्न वर्षों से सम्बन्धित जनसंख्या प्रदर्शित कर रही है , जैसा –



अब प्रश्न के अनुसार हमें 1975 से सम्बन्धित जनसंख्या का आन्तरगणन करना है। सबसे पहले हम X अक्ष पर 1975 वर्ष ज्ञात करेंगे , उसके बाद 1975 से सम्बन्धित बिन्दु से ऊपर लम्ब अक्ष के समानान्तरण एक सीधी रेखा खीचेंगे जो ab को c बिन्दु पर काटती है। c बिन्दु ही 1975 से सम्बन्धित बिन्दु है जिससे होकर ab गुजरती है। अब यदि हम c बिन्दु से Y अक्ष पर लम्ब डालें तो हमें 1975 से सम्बन्धित ज्ञात हो जाएगी, जो ---- है।

इसी प्रकार यदि हमें 1995 के लिए जनसंख्या बाह्यगणन करना हो तो हम सबसे पहले X अक्ष पर 1995 ज्ञात करेंगे जहाँ से एक सीधी रेखा लम्बवत् ऊपर खीचेंगे तथा ab को बढ़ाने में जो उस लम्बवत् रेखा को d पर काटती है। d से Y अक्ष पर लम्ब खींचकर 1995 की जनसंख्या ज्ञात कर लेंगे जो ----- है।

यहाँ एक बात और उल्लेखनीय है कि यदि आँकड़े अविच्छिन्न श्रेणी या वर्गान्तर के रूप में हो तो मध्य बिन्दुओं को X अक्ष पर तथा आवृत्तियों को Y अक्ष पर अंकित करेंगे। शेष प्रक्रिया पहले की ही तरह होगी , और इस स्थिति में भी हम आन्तरगणन तथा बाह्यगणन क्रिया कर लेंगे।

7.7.2 आन्तरगणन तथा बाह्यगणन की बीजगणितीय विधियाँ (Algebraic methods of interpolation and extrapolation)

बीजगणितीय विधि के अन्तर्गत आन्तरगणन तथा बाह्यगणन की अनेक विधियाँ प्रयोग में लायी जाती है जिन्हें हम मोटे तौर पर दो भागों में विभक्त कर सकते हैं-

जब समंक माला के चर बराबर अन्तर (equal intervals) से बढ़े तथा जब समंक माला के चर असमान अन्तर (unequal intervals) से बढ़े। इस स्थिति में निम्नांकित सूत्र प्रयुक्त होते हैं-

(i) प्रत्यक्ष द्विपद-विस्तार विधि (Direct Binomial Expansion Method)

(ii) न्यूटन की प्रगामी-अन्तर विधि (Newton's method of Advancing Differences)

(iii) लाग्रेंज विधि (Lagrange's Method)

(iv) परवलयिक वक्र विधि (Parabolic Curve Method)

(अ) अन्य रीतियाँ (Other Method)

(क) न्यूटन-गॉस (अग्रगामी) विधि (Newton-Gauss (forward) method)

(ख) न्यूटन-गॉस (पृष्ठगामी) विधि (Newton-Gauss (backward) method)

(ग) स्टर्लिंग का सूत्र विधि (Sterling's formula method)

(घ) न्यूटन की विभाजित अन्तर विधि (Newton's divided difference method)

(i) प्रत्यक्ष द्विपद-विस्तार विधि (Direct Binomial Expansion Method)

प्रयोग- यह विधि द्विपद-प्रमेय (Binomial Theorem) पर आधारित है। इसका प्रयोग तब किया जाता है जब निम्न दो शर्तें पूरी होती है - (क) स्वतंत्र चर (x) के पद बराबर अन्तर से बढ़ते हैं , जैसे 1989, 1991, 1993, 1995, 1997, 1999 या 1961, 1971, 1981, 1991, 2001... (ख) इन बराबर अन्तर वाले (equidistant) पदों में से ही किसी एक y मूल्य के आश्रित पद ल का मूल्य अनुमानित करना होता है। उदाहरणार्थ, यदि 1961, 1971, 1981 और 1991 जनगणना वर्षों में से किसी नगर की 1961, 1971 और 1991 की जनसंख्या ज्ञात हो और 1981 की जनसंख्या अनुमानित करनी हो तो द्विपद-विस्तार विधि द्वारा आन्तरगणन किया जाएगा क्योंकि 1961, 1971, 1981 और 1991 के अन्तर समान (10) हैं। इस प्रकार यदि 1961, 1971, 1981 और 1991 के भारत की जनसंख्या के आँकड़े ज्ञात हैं और उनकी सहायता से 2001 के लिए जनसंख्या का बाह्यगणन करना हो तो भी यही विधि अपनायी जाएगी।

प्रक्रिया- इस विधि की निम्नांकित प्रक्रियाएँ हैं-

(i) स्वतंत्र चर-मूल्य (x) के पदों को क्रमानुसार $x_0, x_1, x_2, x_3,\dots$ तथा $y_0, y_1, y_2, y_3,\dots$ आदि संकेताक्षरों द्वारा व्यक्त किया जाता है। उनका प्रमुख अन्तर (leading difference) सदैव शून्य माना जाता हैं। उदाहरणार्थ , मान लीजिए y श्रेणी के ज्ञात मूल्य 5 है तो पाचवाँ प्रमुख अन्तर (5 leading difference) शून्य होगा $\Delta_0^5 = 0$

सूत्र की दृष्टि से- $\Delta_0^n = 0; \quad n = y$ श्रेणी के ज्ञात मूल्यों की संख्या

(ii) y के जितने मूल्य ज्ञात होते, ज्ञात मूल्यो की संख्या $\Delta_0^n = (y-1)^n$

$$= y^n - y^{n-1} + \frac{n(n-1)}{1\times 2} y^{n-2} - \frac{n(n-1)(n-2)}{1\times 2\times 3} y^{n-3} + \dots = 0$$

यदि y के ज्ञात मूल्यो की संख्या (n) 5 हो, तो -

$$\Delta_0^5 = (y-1)^5 = y^5 - \frac{5y^{5-1}}{1} + \frac{5(5-1)}{1\times 2} y^{5-2} - \frac{5(5-1)(5-2)}{1\times 2\times 3} y^{5-3}$$

$$+\frac{5(5-1)(5-2)(5-3)}{1\times2\times3\times4}\,y^{5-4}-\frac{5(5-1)(5-2)(5-3)(5-4)}{1\times2\times3\times4\times5}\,y^{5-5}=0$$

इसको हल करने पर निम्न समीकरण प्राप्त होती है -

$$=y_5-5y_4+10_{Y3}-10y_2+5y_1-y_0=0$$

सूत्र निकालने की एक व्यावहारिक व सरल विधि

द्विपद विस्तार का उपरोक्त ढंग बहुत जटिल है। इसलिए इसकी एक सरल विधि नीचे दी जा रही है-

(क) जिस प्रमुख अन्तर के लिए द्विपद-विस्तार करना हो पहले उसे उस क्रम के y को लिखा जाएगा। फिर अवरोही क्रम में y का घात, अधोलिखित संकेत के रूप में एक-एक कम करते जाएंगे जिससे अन्त में y_0 आ जाए। जैसे यदि 5 मूल्य ज्ञात हों तो पाचवाँ प्रमुखान्तर शून्य होगा और y को निम्न क्रमानुसार लिखा जाएगा-

y_5 y_4 y_3 y_2 y_1 y_0

(ख) प्रथम y धनात्मक होगा, अगला y ॠणात्मक, फिर उससे अगला y धनात्मक और इसी प्रकार अन्त तक चिन्ह एकान्तर (alternate) रूप में लिखे जाएंगे। जैसे $+y_5$ $-y_4$ $+y_3$ $-y_2$ $+y_1$ $-y_0$

(ग) विभिन्न y's के अंकात्मक गुणक (numerical coefficients) निकालने की विधि इस प्रकार होगी। पहले लिखे जाने वाले y का गुणक 1 होगा। इससे आगे के ले ys अंकात्मक गुणक निम्न सूत्रानुसार प्राप्त होंगे-

पिछले y का गुणक × पिछले ल का अधोसंकेत

पिछले y की क्रम-स्थिति

उक्त उदाहरण में,

$$1y_5-\frac{1\times5}{1}y_4+\frac{5\times4}{2}y_3-\frac{10\times3}{3}y_2+\frac{10\times2}{4}y_1-\frac{5\times1}{5}y_0=0$$

$$y_5-5y_4+10y_3-10y_2+5y_1-y_0=0$$

द्विपद विस्तार में y के गुणांक पास्कल त्रिभुज (Pascal's Triangle) से भी ज्ञात किये जा सकते हैं।

पास्कल त्रिभुज (Pascal's Triangle)

n	**अंकात्मक गुणांक (Numerical coefficients)**	**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

कुछ द्विपद-विस्तार-

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	$(y-1)^4 = 0$	$y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$
5	$(y-1)^5 = 0$	$y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0 = 0$
6	$(y-1)^6 = 0$	$y_6 - 6y_5 + 15y_4 - 20y_3 + 15y_2 - 6y_1 + y_0 = 0$

उदाहरण (Illustration): 2

निम्न मूल्यों के आधार पर किसी भी बीजगणितीय विधि का प्रयोग करते हुए y का मूल्य ज्ञात कीजिए जब $x = 3$ हो-

x	:	1	2	3	4	5
y	:	216000	226981	?	250047	262144

हल- द्विपद विस्तार विधि का प्रयोग किया जाएगा क्योंकि इसकी दोनों शर्तें पूरी हो रही है। प्रथम , x श्रेणी के क्रमिक पदों में समान अन्तर है। दूसरा , आन्तरगणित किया जाने वाला मूल्य x_3, x श्रेणी के समान अन्तर वाले पदों में से ही एक पद है।

		x_0	x_1	x_2	x_3	x_4
x	:	1	2	3	4	5
y	:	216000	226981	?	250047	262144
		y_0	y_1	y_2	y_3	y_4

y - श्रेणी के 4 पद ज्ञात हैं अतः $\Delta_0^4 = 0$ मानकर उसका द्विपद विस्तार लिखा जाएगा-

$$\Delta_0^4 = y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$$

ज्ञात मूल्यों को समीकरण में आदिष्ट करने पर–

$$262144 - 4 \times 250047 + 6y_2 - 4 \times 226981 + 216000 = 0$$

$$262144 - 1000188 + 6y_2 - 907924 + 216000 = 0$$

$6y_2 = 1429968 \quad y_2 = 238328$

अतः **x = 3** के लिए **y** का अनुमानित मूल्य **238328** है।

दो अज्ञात मूल्य (Two Missing Values)

जब स्वतंत्र चर-मूल्यों (x's) के अन्तर समान हों और दो अज्ञात मूल्यों (y's) का आन्तरगणन करना हो तो दो समीकरणों की आवश्यकता होती है। प्रथम , ज्ञात मूल्यों की संख्या के बराबर प्रमुख अन्तर को शून्य मानकर द्विपद विस्तार लिखा जाता है। दूसरे, उक्त द्विपद विस्तार को फिर से लिखकर प्रत्येक y के अधोलिखित संकेत (subscript) में 1 की वृद्धि कर देते हैं जिससे , अन्त में y_0 के स्थान पर y_1 प्राप्त हो जाता है। तत्पश्चात् ज्ञात मूल्यों को दोनों समीकरणों में आदिष्ट करके , उनके हल द्वारा अज्ञात मूल्य अनुमानित कर लिए जाते हैं। उदाहरणार्थ, यदि 7 मूल्य ज्ञात हों और 2 अज्ञात मूल्यों का आन्तरगणन करना हो , तो निम्न दो समीकरण बनाए जाएंगे-

$$\Delta_0^7 = y_7 - 7y_6 + 21y_5 - 35y_4 + 35y_3 - 21y_2 + 7y_1 - y_0 = 0 \quad \text{... (1)}$$

$$\Delta_1^7 = y_8 - 7y_7 + 21y_6 - 35y_5 + 35y_4 - 21y_3 + 7y_2 - y_1 = 0 \quad \text{... (2)}$$

इन दोनों द्विपद समीकरणों की सहायता से दो अज्ञात मूल्यों के सम्भाव्य अनुमान लगा लिए जाएंगे।

उदाहरण (Illustration): 3

निम्न सारिणी की सहायता से 1980 और 1990 के लिए उत्पादन का अनुमान लगाइए-

वर्ष :	1965	1970	1975	1980	1985	1990	1995
उत्पादन (000 टनों में) :	200	220	260	?	350	?	430

हल:

	x_0	x_1	x_2	x_3	x_4	x_5	x_6
वर्ष (x) :	1965	1970	1975	1980	1985	1990	1995
उत्पादन (y):	200	220	260	?	350	?	430
	y_0	y_1	y_2	y_3	y_4	y_5	y_6

x's का अन्तर समान होने के कारण प्रत्यक्ष द्विपद-विस्तार विधि प्रयुक्त की जाएगी। y के 5 मूल्य ज्ञात हैं और 2 अज्ञात। इसलिए पाँचवें प्रमुख-अन्तर से सम्बन्धित द्विपद-विस्तार का प्रयोग दो बार निम्न प्रकार किया जाएगा-

$$y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0 = 0 \quad \text{....(1)}$$

$$y_6 - 5y_5 + 10y_4 - 10y_3 + 5y_2 - y_1 = 0 \quad \text{....(2)}$$

ज्ञात मूल्य आदिष्ट करने पर-

$$y_5 - 5 \times 350 + 10y_3 - 10 \times 260 + 5 \times 220 - 200 = 0$$

$$430 - 5y_5 + 10 \times 350 + 10y_3 - 5 \times 260 + 220 = 0$$

$$y_5 + 10y_3 = 3450 \quad \text{...... (3)}$$

$$-5y_5 - 10y_3 = -5010 \quad \text{..... (4)}$$

दोनों समीकरणों को जोड़ने पर निम्न समीकरण उपलब्ध होता है-

$-5y_5 = -1560$ **$y_5=390$**

y_5 के मूल्य को समीकरण (3) में आदिष्ट करने पर y_3 का मूल्य निम्न प्रकार निकाला जाएगा-

$$390 + 10y_3 = 3450$$

$$y_3 = \frac{3450-390}{10} = 306$$

1985 और 1990 में उत्पादन की अनुमानित मात्रा के समंक क्रमशः 306 और 390 हजार टन हैं।

जब मूल्यों में असाधारण उच्चावचन हो

कभी-कभी ऐसा भी देखने में आता है कि दिये गये प्रश्न में एक-आध मूल्य असाधारण रूप से उच्चावचन लिए हुए होता है। ऐसी परिस्थिति में, वांछित मूल्य का आन्तरगणन करने से पूर्व, अनियमित मूल्य को नियमित व सामान्य कर लेना चाहिए अन्यथा अनुमानित किये जाने वाला मूल्य गलत होगा।

उदाहरण (Illustration): 4

एक जिले की विभिन्न वर्षों की जनसंख्या नीचे दी गयी है। वर्ष 1915 की जनसंख्या अनुमानित कीजिए-

वर्ष	:	1910	1911	1912	1913	1914
जनसंख्या (मिलियन में)	:	7	9	36	14	16

हल- नोट: इस प्रश्न में 1915 के वर्ष के लिए जनसंख्या अनुमानित करने से पहले 1912 के लिए जनसंख्या का अनुमान लगाना होगा , क्योंकि 1912 पर जनसंख्या में अत्यधिक उतार-चढ़ाव (jump) हुआ है जो कि आन्तरगणन की मान्यताओं की दृष्टि से सही नहीं है।

	x_0	x_1	x_2	x_3	x_4
वर्ष () :	1910	1911	1912	1913	1914
जनसंख्या () :	7	9	36	14	16
	y_0	y_1	y_2	y_3	y_4

चूंकि ज्ञात मूल्य 4 है इसलिए $\Delta_0^4 = 0$

$$\Delta_0^4 = y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$$

$$16 - (4\times14) + 6y_2 - (4\times9) + 7 = 0$$

$$16 - 56 + 6y_2 - 36 + 7 = 0$$

$$6y_2 = -23 + 92$$

$6y_2 = 69$ मि0 $y_2 = 11.5$ मिलियन

अब 1915 के वर्ष के लिए जनसंख्या अनुमानित की जाएगी-

		x_0	x_1	x_2	x_3	x_4	x_5
वर्ष	:	1910	1911	1912	1913	1914	1915
जनसंख्या (मि0)	:	7	9	11.5	14	16	?
		y_0	y_1	y_2	y_3	y_4	y_5

चूंकि ज्ञात मूल्य 5 है अतएव $\Delta_0^5 = 0$

$$\Delta_0^5 = y_5 - 5y_4 + 10y_3 - 10y_2 + 5y_1 - y_0 = 0$$

or $y_5 - (5\times16) + (10\times14) - (10\times11.5) + (5\times9) - 7 = 0$

$$y_5 - 80 + 140 - 115 + 45 - 7 = 0$$

$y_5 = 202 - 185$ $y_5 = 17$ मिलियन

अतः 1915 वर्ष के लिए जनसंख्या 17 मि0 है।

(ii) न्यूटन की प्रगामी-अन्तर विधि (Newton's method of Advancing Differences)

न्यूटन की प्रगामी अन्तर विधि द्विपद-प्रमेय पर ही आधारित है। इस विधि का प्रयोग उस परिस्थिति में करना चाहिए जिसमें स्वतंत्र श्रेणी (x) के दिए हुए पदों के अन्तर समान हों परन्तु जिस पद (x) के लिए आश्रित चर के पद (yx) का आन्तरगणन करना हो वह दिए हुए स्वतंत्र चर मूल्यों से सर्वथा भिन्न हो अर्थात् वह समान अन्तर वाले x's से भिन्न अन्य कोई मूल्य हो। उदाहरणार्थ यदि राष्ट्रीय आय प्रति 5 वर्ष के अन्तर से दी हुई हो- 1920, 1925, 1930, 1935, 1940 तो ऐसी हालत में अगर 1922 या 1938 के वर्ष के लिए राष्ट्रीय आय ज्ञात करनी हो तो इस न्यूटन प्रगामी अन्तर विधि का प्रयोग किया जाएगा। इस विधि का प्रयोग बाह्यगणन के

लिए भी किया जा सकता है परन्तु यह समंक माला के पूर्वार्द्ध में किसी 'x' के आश्रित मूल्य (yx) का आन्तरगणन करने के लिए अधिक उपयुक्त होती है।

प्रयोग करने की दृष्टि से "प्रगामी अन्तर विधि" तथा "द्विपद विस्तर रीति" में अन्तर-

दोनों, प्रगामी तथा द्विपद विस्तार रीतियों में x श्रेणी या स्वतंत्र चरों का समान रूप से बढ़ना आवश्यक है, परन्तु x के जिस मूल्य आ आन्तरगणन करना होता है वह प्रगामी अन्तर रीति में x श्रेणी का ही एक नियमित मूल्य न होकर उनके बीच का कोई अज्ञात व अनियमित मूल्य हो सकता है जबकि द्विपद विस्तार रीति में आन्तरगणित किये जाने वाला मूल्य x श्रेणी का स्वतः एक नियमित मूल्य होता है।

उदाहरणार्थ -

प्रगामी अन्तर विधि (1)		**द्विपद विस्तार विधि (2)**	
x	y	x	xy
10	132	10	132
20	140	20	140
30	155	30	155
40	172	40	?
50	190	50	190

उपर्युक्त उदाहरण (1) में 10, 20, 30, 40, 50 के मूल्यों को छोड़कर अन्य किसी भी बीच में आने वाले मूल्य का अनुमान प्रगामी अन्तर विधि द्वारा किया जाएगा, जैसे- 12, 18, 28, 35, 42 इत्यादि। जबकि उदाहरण (2) के ज्ञात किये जाने वाला मूल्य x श्रेणी में से ही कोई एक होगा जैसे 40 पर y का मूल्य ज्ञात करना।

क्रिया विधि - न्यूटन प्रगामी अन्तर विधि में निम्न क्रियाएँ अपनायी जाती है-

(1) संकेताक्षर- (i) सबसे पहले x श्रेणी के मूल्यों को क्रमानुसार $x_0, x_1, x_2, x_3, \ldots, x_n$ आदि संकेताक्षरों द्वारा तथा y श्रेणी के मूल्यों को क्रमानुसार $y_0, y_1, y_2, \ldots y_n$ आदि संकेताक्षरों द्वारा दिखाया जाता है। (ii) x श्रेणी के जिस पद का मूल्य आन्तरगणन करना होता है उसे 'x' संकेताक्षर द्वारा प्रकट किया जाता है और (iii) x पर आश्रित y श्रेणी के जिस मूल्य का आन्तरगणन करना हो उसे 'yx' द्वारा प्रकट किया जाता है।

(2) अन्तर सारणी की रचना - y के प्रमुखान्तरों को ज्ञात करने के लिए परिमितान्तरों की सारणी (Table of Finite Differences) बनायी जाती है जिसमें स्वतंत्र व आश्रित चरों के अतिरिक्त y के ज्ञात मूल्यों की संख्या से एक कम संख्या में अन्तरों के खाने (columns) होते हैं। प्रत्येक कॉलम के प्रथम अन्तर को 'प्रमुखान्तर' (leading difference) कहते हैं तथा इसको Δ चिन्ह द्वारा प्रकट किया जाता है- प्रथम, द्वितीय, तृतीय आदि प्रमुखान्तरों के लिए कॉलम के ऊपर $\Delta, \Delta^2, \Delta^3$ आदि चिन्ह रखे जाते हैं। अन्तर निकालनें के लिए y के प्रत्येक मूल्य में से पिछला मूल्य घटाया जाता है, जैसे प्रथम अन्तरों के खाने में $y_1 - y_0 = \Delta_0^1$, $y_2 - y_1 = \Delta_1^1$, $y_3 - y_2 = \Delta_2^1$ आदि। दूसरे खाने के अन्तरों को पहले खाने के अन्तरों की सहायता से इसी प्रकार निकाला जाएगा अर्थात् $\Delta_1^1 - \Delta_0^1 = \Delta_0^2$; $\Delta_2^1 - \Delta_1^1 = \Delta_1^2$; $\Delta_3^1 \ \Delta_2^1 = \Delta_2^2 \ldots$, इसी प्रकार अन्त तक अन्तर निकाले जाएंगे। अन्तरों की संख्या कम होती जाएगी और अन्तिम खाने में एक मात्र अन्तर रहेगा।

अन्तर सारणी (Table of difference)

स्वतंत्र चर (*x*)	आश्रित चर (*y*)	अन्तर							
		प्रथम		द्वितीय		तृतीय		चतुर्थ	
x_0	y_0								
		$y_1 - y_0$	Δ_0^1						
x_1	y_1			$\Delta_1^1 - \Delta_0^1$	Δ_0^2				
		$y_2 - y_1$	Δ_1^1			$\Delta_1^2 - \Delta_0^2$	Δ_0^3		
x_2	y_2			$\Delta_2^1 - \Delta_1^1$	Δ_1^2			$\Delta_1^3 - \Delta_0^3$	Δ_0^4
		$y_3 - y_2$	Δ_2^1			$\Delta_2^2 - \Delta_1^2$	Δ_1^3		
x_3	y_3			$\Delta_3^1 - \Delta_2^1$	Δ_2^2				
		$y_4 - y_3$	Δ_3^1						
x_4	y_4								

नोट - अन्तर सारणी में विभिन्न स्तरों पर अन्तर लेते समय बीजगणितीय चिन्हों (+ व -) का विशेष ध्यान रचना चाहिए क्योंकि किसी एक अन्तर के अशुद्ध होने पर अन्तरों की पूरी श्रृंखला अशुद्ध हो जाती है।

सारणी देखने से स्पष्ट हो जाता है कि यदि सभी प्रमुखान्तर ज्ञात हों तो उनकी सहायता से बाकी सभी अन्तर और y के मूल्य आसानी से ज्ञात किये जा सकते हैं-

सूत्रानुसार-

$$y_1 = y_0 + \Delta_0^1$$

$$y_2 = y_1 + \Delta_1^1 = y_0 + \Delta_0^1 + \Delta_0^2 + \Delta_0^1 = y_0 + 2\Delta_0^1 + \Delta_0^2$$

$$y_3 = y_2 + \Delta_2^1 = (y_0 + 2\Delta_0^1 + \Delta_0^2) + (\Delta_1^1 + \Delta_1^2)$$

$$= y_0 + 2\Delta_0^1 + \Delta_0^2 + (\Delta_0^2 + \Delta_0^1) + (\Delta_0^2 + \Delta_0^3)$$

$$= y_0 + 3\Delta_0^1 + 3\Delta_0^2 + \Delta_0^3$$

प्रमुखान्तर और द्विपद-विस्तार- प्रमुखान्तरों को यदि ज्ञात y's के रूप में व्यक्त किया जाये तो द्विपद-विस्तार प्राप्त हो जाते हैं, उदाहरणार्थ-

$$\Delta_0^1 = y_1 - y_0$$

$$\Delta_0^2 = (y_2 - y_1) - (y_1 - y_0) = y_2 - 2y_1 + y_0$$

$$\Delta_0^3 = \Delta_1^2 - \Delta_0^2 = (\Delta_2^1 - \Delta_1^1) - \Delta_0^2 = (y_3 - y_2) - (y_2 - y_1) - (y_2 - 2y_1 + y_0)$$

$$= y_3 - 3y_2 + 3y_1 - y_0$$

$$\Delta_0^4 = y_4 - 4y_3 + 6y_2 - 4y_1 + y_0$$

(3) स्वतंत्र चर मूल्यों के अन्तर - प्रमुखान्तर निकालने के बाद निम्न सूत्र द्वारा xx और x_0 के अन्तर का x'sके समान अन्तरों पर अनुपात निकाला जाता है -

$$x = \frac{\text{आन्तरगणन पद – मूल पदनिकटवर्ती}}{\text{पदों का अन्तर}}$$

$$x = \frac{\text{Item of Interpolation – Item of Origindifference}}{\text{between adjoining item}}$$

$$x = \frac{x_x - x_0}{x_1 - x_0}$$

(4) सूत्र - अन्त में निम्न सूत्र जिसे ' **न्यूटन ग्रेगोरी सूत्र** ' भी कहते हैं, का प्रयोग किया जाता है-

$$y_x = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1\times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1\times 2\times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1\times 2\times 3\times 4}\Delta_0^4 + \ldots$$

महत्वपूर्ण संकेत- सूत्र का आकार प्रमुखान्तरों व पदों की संख्या पर निर्भर करता है। सूत्र का उपर्युक्त रूप 4 प्रमुखान्तरों वाले प्रश्न के लिए उदाहरण स्वरूप है। अगर मान लीजिए 5वाँ प्रमुखान्तर (leading difference) भी निकलता, तो सूत्र में निम्न विस्तार और करना पड़ता- $+ \frac{x(x-1)(x-2)(x-3)(x-4)}{1\times 2\times 3\times 4\times 5}\Delta_0^5$

उदाहरण (Illustration): 5

निम्न समंकों के आधार पर 16 वर्ष की आयु पर जीवन प्रत्याशा का अनुमान लगाइए-

आयु (वर्ष)	:	15	20	25	30	35
जीवन-प्रत्याशा (वर्ष)	:	32.2	29.1	26.0	23.1	20.4

हल- अन्तर सारणी (जंइसम व िक्पिमतमदबम)

आयु (वर्ष) *x*		जीवन प्रत्याशा (वर्ष) *y*		अन्तर (difference)							
				प्रथम (First)		द्वितीय (Second)		तृतीय (Third)		चतुर्थ (Fourth)	
15	x_0	32.2	y_0								
20	x_1	29.1	y_1	29.1–32.2	−3.1 Δ_0^1	−3.1–(−3.1)	0 Δ_0^2				
25	x_2	26.0	y_2	26.0–29.1	−3.1			0.2–0	0.2 Δ_0^3		
30	x_3	23.1	y_3	23.1–26.0	−2.9	−2.9–(3.1)	0.2			0–0.2	−0.2 Δ_0^4
35	x_4	20.4	y_4	20.4–23.1	−2.7	−2.7– (2.9)	0.2	0.2–0.2	0		
22	x_5	?	y_5								

$$x = \frac{\text{आन्तरगणन – मूल वर्षनिकटवर्ती}}{\text{वर्षों का अन्तर}} = \frac{x_x - x_0}{x_1 - x_0} = \frac{22-1520}{-15} = 1.4$$

ज्ञात मूल्यों की संख्या 5 है अतएव न्यूटन का सूत्र चौथे प्रमुखान्तर (Δ_0^4) तक लिखा जाएगा-

$$y_x = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1\times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1\times 2\times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1\times 2\times 3\times 4}\Delta_0^*$$

$$y_x = 32.2 + 1.4\times(-3.1) + \frac{1.4\times 0.4}{2}\times 0 + \frac{1.4\times 0.4\times(-0.6)}{2\times 3}\times 0.2$$

$$+ \frac{1.4\times 0.4\times(-0.6)\times(-1.4)}{2\times 3\times 4}\times 0.2$$

$$y_x = 32.2 - 4.34 + 0 - 0.0112 - 0.004 = 32.2 - 4.35 = 27.8$$

अतः 22 वर्ष की आयु के लिए जीवन-प्रत्याशा **27.8** वर्ष है।

आवृत्ति वितरण में आन्तरगणन (Interpolation in Frequency Series)

(i) आवृत्ति वितरण में आवृत्तियों को संचयी (cumulative) बनाकर आन्तरगणन किया जाता है। शेष क्रिया पूर्ववत् ही बनी रहती है।

(ii) कभी-कभी अनुमानित मूल्य किसी निश्चित समय या मूल्य के लिए न निकलवा कर दो सीमाओं के बीच के लिए निकलवाया जाता है जिसका एक उपयुक्त उदाहरण हम नीचे दे रहे हैं। ऐसे प्रश्नों के लिए भी न्यूटन की विधि ही उपयुक्त समझी जाती है।

उदाहरण (Illustration): 6

निम्न सारणी से (i) 45 से कम (ii) 55 से कम , तथा (iii) 45-55 के बीच अंक प्राप्त करने वाले विद्यार्थियों की संख्या ज्ञात कीजिए-

प्राप्तांक (Marks)	:	30-40	40-50	50-60	60-70	70-80
विद्यार्थियों की संख्या (No. of Students)	:	31	42	51	35	31

हल- पहले, संचयी आवृत्ति वितरण के रूप बदलकर अन्तर-सारणी बनाई जाएगी-

अन्तर-सारणी (Table of difference)

अंक (x)		विद्यार्थियों की संख्या (y)		अन्तर (difference)							
				प्रथम		द्वितीय		तृतीय		चतुर्थ	
40 से कम	x_0	31	y_0	+42	Δ_0^1	+9	Δ_0^2	−25	Δ_0^3	+37	Δ_0^4
50 से कम	x_1	73	y_1	+51		−16		+12			
60 से कम	x_2	124	y_2	+35		−4					
70 से कम	x_3	159	y_3	+31							

80 से कम	x_4	190	y_4								
45 से कम	x_x	?	y_x								

(i) 45 से कम प्राप्तांकों के लिए आन्तरगणन-

$$x = \frac{x_x - x_0}{x_1 - x_0} = \frac{45-40}{50-40} = \frac{5}{10} = 0.5$$

चौथे प्रमुखान्तर तक न्यूटन का प्रगामी-अन्तर सूत्र लिखकर उसमें ज्ञात मूल्यों को रखा जाएगा-

$$y_x = y_0 + x\Delta_0^1 + \frac{x(x-1)}{1\times 2}\Delta_0^2 + \frac{x(x-1)(x-2)}{1\times 2\times 3}\Delta_0^3 + \frac{x(x-1)(x-2)(x-3)}{1\times 2\times 3\times 4}\Delta_0^4$$

$$y_x = 31 + (0.5\times 42) + \frac{0.5(0.5-1)}{1\times 2}\times 9 + \frac{0.5(0.5-1)(0.5-2)}{1\times 2\times 3}\times(-25) +$$

$$\frac{0.5(0.5-1)(0.5-2)(0.5-3)}{1\times 2\times 3\times 4}\times 37$$

$= 31 + 21 - 1.125 - 1.5625 - 1.4453$

$= 47.8672$ या 48 approx.

अतः 45 से कम अंक प्राप्त करने वाले विद्यार्थियों की संख्या 48 है।

(ii 55 से कम प्राप्तांकों के लिए आन्तरगणन-

$$x = \frac{\text{आन्तरगणन – मूल वर्षनिकटतम्}}{\text{वर्षों का अन्तर}} = \frac{55-40}{60-50} = \frac{15}{10} = 1.5$$

$$y_x = 31 + (1.5\times 42) + \frac{1.5(1.5-1)}{1\times 2}\times 9 + \frac{1.5(1.5-1)(1.5-2)}{1\times 2\times 3}\times(-25)$$

$$+ \frac{1.5(1.5-1)(1.5-2)(1.5-3)}{1\times 2\times 3\times 4}\times 37$$

$= 31 + 63 + 3.375 + 1.5625 + 0.8672$

= **99.8047** or **100 approx.**

अतः 55 से कम अंक प्राप्त करने वाले विद्यार्थियों की संख्या 100 है।

(iii) 45 से 55 के बीच प्राप्तांकों के लिए आन्तरगणन-

45 से कम अंक प्राप्त करने वाले छात्रों की संख्या = 48

55 से कम अंक प्राप्त करने वाले छात्रों की संख्या = 100

45 से 55 के बीच अंक प्राप्त करने वाले छात्रों की संख्या = 100 - 48 = 52

उदाहरण (Illustration): 7

न्यूटन विधि द्वारा अधिकतम 350 सेन्टीग्रेड से सम्बद्ध न्यूनतम सम्भावित तापमान ज्ञात कीजिए-

अधिकतम तापमान	:	36	34	32	30	28
न्यूनतम तापमान	:	21	19	16	12	11

हल- $x = \frac{x_x - x_0}{x_1 - x_0} = \frac{35-36}{34-36} = \frac{-1}{-2} = 0.5$

अन्तर-सारणी (Table of difference)

तापमान x		तापमान y		प्रमुखान्तर							
				प्रथम Δ^1		द्वितीय Δ^2		तृतीय Δ^3		चतुर्थ Δ^4	
36	x_0	21	y_0	-2	Δ_0^1						
34	x_1	19	y_1	-3		-1	Δ_0^2	0	Δ_0^3		
32	x_2	16	y_2	-4		-1		$+4$		$+4$	Δ_0^4
30	x_3	12	y_3	-1		$+3$					
28	x_4	11	y_4								

$$y_x = 21 + (0.5 \times -2) + \frac{0.5(0.5-1)}{1 \times 2} \times (-1) + \frac{0.5(0.5-1)(0.5-2)}{1 \times 2 \times 3} \times 0 + \frac{0.5(0.5-1)(0.5-2)(0.5-3)}{1 \times 2 \times 3 \times 4} \times 4$$

$$= 21 - 1 + 0.125 + 0 - 0.156$$

$$= 21.125 - 1.156 = 19.97 \quad y_x = 20$$

टिप्पणी- यदि इस प्रश्न को 'अवरोही क्रम' के स्थान पर आरोही क्रम में रखकर हल किया जाये तो ऐसी दशा में प्रमुखान्तर क्रमशः +1, +3, -4, +4 होंगे और x का मान 3.5 आयेगा, किन्तु आन्तरगणन मूल्य एक-समान अर्थात् **19.97** ही निकलकर आयेगा।

(iii) लाग्रैंज की विधि (Lagrange's Method)

प्रयोग- फ्रांस के प्रसिद्ध गणितज्ञ लाग्रैंज द्वारा प्रतिपादित रीति आन्तरगणन एवं बाह्यगणन की सार्वभौमिक रीति (universal method) है। सैद्धान्तिक दृष्टि से लाग्रैंज के सूत्र द्वारा किसी भी प्रकार की परिस्थिति में (चाहे स्वतंत्र चर मूल्यों के अन्तर समान हों या असमान हों) आन्तरगणन व बाह्यगणन किया जा सकता है। परन्तु व्यवहार में इस रीति का प्रयोग वहाँ किया जाता है जहाँ द्विपद-विस्तार रीति तथा न्यूटन की प्रगामी अन्तर-विधि प्रयुक्त न की जा सके अर्थात् जहाँ x's के अन्तर अनियमित या असमान (irregular or unequal intervals) हों। उदाहरणार्थ, यदि किसी नगर की 1981, 1985, 1990, 1991 और 1993 की जनसंख्या ज्ञात हो और 1989 की जनसंख्या का आन्तरगणन करना हो या सन् 2002 ई. की जनसंख्या का बाह्यगणन करना हो तो लाग्रैंज विधि ही अपनायी जाएगी क्योंकि वर्षों के अन्तर असमान है और इस स्थिति में द्विपद-विस्तार विधि या न्यूटन-विधि प्रयोग नहीं की जा सकती।

क्रिया विधि- (1) सर्वप्रथम x श्रेणी को x_0, x_1, x_2 .... आदि संकेताक्षरों तथा y श्रेणी को y_0, y_1, y_2 .... आदि संकेताक्षरों द्वारा प्रदर्शित किया जाता है। स्वतंत्र पदमाला (x - series) के जिस मूल्य के लिए आन्तरगणन करना होता है उसे x द्वारा सम्बोधित किया जाता है और आश्रित श्रेणी के आन्तरगणित किये जाने वाले मूल्य को y_x कहते हैं।

(2) सूत्र के निर्माण करने हेतु नीचे एक काल्पनिक उदाहरण लिया जा रहा है। मान लीजिए x व y दो श्रेणी है और x = 6 पर y का मूल्य निकालना है।

x		y	
4	x_0	120	y_0
7	x_1	140	y_1
8	x_2	165	y_2
12	x_3	183	y_3

(i) अब सूत्र निर्माण के लिए सर्वप्रथम x's से x श्रेणी के सभी मूल्य घटाए जाएंगे परन्तु प्रथम मूल्य (y_0) का तत्सम्बन्धी (corresponding) x_0 मूल्य, x में से नहीं घटाया जाएगा। उदाहरणार्थ-

$$y_0 \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} + \ldots$$

(ii) सूत्र की इस प्रथम पंक्ति के 'हर' (denominator) मूल्सरें के लिए सदैव एक बात याद रखनी चाहिए कि x में से जो मूल्य नहीं घटाया गया था अर्थात् x_0 , अब x श्रेणी के सभी मूल्य उसी (x_0) में से घटाये जायेंगे।

(iii) द्वितीय पंक्ति - $y_1 \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} + \ldots$

स्पष्ट है कि 'अंश' (numerator) के अन्तरों के लिए (अर्थात् पंक्ति के ऊपरी हिस्से में) x में से y_1 का तत्संवादी मूल्य x_1 नहीं घटाया गया है जबकि 'हर' में उसी x_1 में से ही x श्रेणी के सभी मूल्य घटाये गये हैं।

(iv) यह क्रम इसी प्रकार चलता रहेगा।

(v) सूत्र के विस्तार की सीमा, ज्ञात पदों की संख्या अर्थात् x_n पर निर्भर करती है। लाग्रेंज का सूत्र इस प्रकार है-

$$y_x = y_0 \frac{(x-x_1)(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)\ldots(x_0-x_n)} + y_1 \frac{(x-x_0)(x-x_2)(x-x_3)\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)\ldots(x_1-x_n)}$$
$$+ y_2 \frac{(x-x_0)(x-x_1)(x-x_3)\ldots(x-x_n)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)\ldots(x_2-x_n)} + \ldots + y_n \frac{(x-x_0)(x-x_1)(x-x_2)\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)(x_n-x_3)\ldots(x_n-x_{n-1})}$$

एक बार पुनः स्मरण रहे- सांख्यिकी की सभी किताबें यह बताती हैं कि लाग्रेंज का सूत्र मान्यता रहित है अतएव इसका प्रयोग प्रत्येक प्रकार के प्रश्न के लिए किया जा सकता है। लेकिन व्यवहार में परीक्षक ऐसा नहीं मानते। यदि कोई प्रश्न न्यूटन तथा द्विपद-विस्तार रीति से निकालने योग्य है तो भले ही वह लाग्रेंज से भी हल किया जा सकता हो, लेकिन उसे लाग्रेंज से हल नहीं करना चाहिए।

उदाहरण (Illustration): 8

निम्न तालिका में जीवन के प्रथम 6 माह के शिशु का सामान्य भार दिया हुआ है। 4 माह की आयु पर शिशु के भार का अनुमान लगाइए।

आयु (माह में) :	0	2	3	5	6
भार (पौण्ड में) :	5	7	8	10	12

हल- स्वतंत्र चर के अन्तर असमान है अतः लाग्रेंज के सूत्र का प्रयोग किया जाएगा-

	x_0	x_1	x_2	x_3	x_4	x
आयु (age) x :	0	2	3	5	6	4

भार (weight) y :	5	7	8	10	12	?
	y_0	y_1	y_2	y_3	y_4	y_x

$$y_x = y_0 \frac{(x-x_1)(x-x_2)(x-x_3)(x-x_4)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)(x_0-x_4)} + y_1 \frac{(x-x_0)(x-x_2)(x-x_3)(x-x_4)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)(x_1-x_4)}$$
$$+ y_2 \frac{(x-x_0)(x-x_1)(x-x_3)(x-x_4)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)(x_2-x_4)} + y_3 \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_4)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)(x_3-x_4)}$$
$$+ y_4 \frac{(x-x_0)(x-x_1)(x-x_2)(x-x_3)}{(x_4-x_0)(x_4-x_1)(x_4-x_2)(x_4-x_3)}$$

$$y_x = \frac{5\times 2\times 1\times(-1)\times(-2)}{(-2)\times(-3)\times(-5)\times(-6)} + \frac{7\times 4\times 1\times(-1)\times(-2)}{2\times(-1)\times(-3)\times(-4)} + \frac{8\times 4\times 2\times(-1)\times(-2)}{3\times 1\times(-2)\times(-3)}$$
$$+ \frac{10\times 4\times 2\times 1\times(-2)}{5\times 3\times 2\times(-1)} + \frac{12\times 4\times 2\times 1\times(-1)}{6\times 4\times 3\times 1}$$

$y_x = 12.5555 - 3.6666 = 8.8888 \text{ or } 8.9$ पौण्ड

अतः 4 महीने की आयु वाले शिशु का अनुमानित भार 8.9 पौण्ड है।

(iv) परवलयिक वक्र विधि (Parabolic Curve Method)

प्रयोग- लाग्रैंज की विधि की भाँति परवलय-वक्र विधि भी सार्वभौमिक रीति (universal method) है जिस की सहायता से भी किसी प्रकार की आन्तरगणन व बाह्यगणन की समस्या का हल किया जा सकता है परन्तु गणन-क्रिया जटिल होने के कारण व्यवहार में इसका प्रयोग तब किया जाता है जबकि पदों की संख्या कम (3 या 4) हो और स्वतंत्र चर-मूल्यों में अधिकतम समान व थोड़ा अन्तर हो।

इस विधि में x का अज्ञात मूल्य ज्ञात करते समय हम यह मानते है कि y और x में परस्पर गणितीय सम्बन्ध है। दिए हुए समंकों के आधार पर निम्न समीकरण असंगित (fit) किया जाता है-

$y = a + bx + cx^2 + dx^3 + \cdots$

इस समीकरण में x तथा y के मूल्य इस प्रकार रखे जाते हैं कि जितने अचर पद (constants) a, b, c, ... की संख्या है ,, उतने ही समीकरण प्राप्त हो जाएं। उसके बाद इन समीक रणों को हल कर लिया जाता है। इससे a, b, c,... आदि के मान प्राप्त हो जाते हैं और इन्हें उपरोक्त समीकरण में रख दिया जाता है। अब x के किसी भी मूल्य के सापेक्ष y का मूल्य ज्ञात किया जा सकता है क्योंकि हम इस समीकरण में x का वह मान लगा देते हैं।

समीकरण ज्ञात करने के लिए हम सबसे पहले ज्ञात मूल्यों की संख्या देखते हैं। यदि 4 मूल्य ज्ञात हों तो समीकरण $y = a + bx + cx^2 + dx^3 + ex^4$ होगा। जैसे-जैसे पदों की संख्या बढ़ती जाती है , समीकरण के पद भी बढ़ते जाते हैं। n पद ज्ञात होने पर समीकरण $y = a + bx + cx^2 + dx^3 + \cdots + nx^{n-1}$ निम्नांकित सारणी से इस नियम का सरल स्पष्टीकरण हो जाता है-

परवलय-वक्र समीकरण (Parabolic Curve Equation)

ज्ञात मूल्यों की संख्या No. of Known	परवलय-वक्र का घात Power of the Parabola (n-1)	समीकरण **Equation**

Values (n)		
2	1	$y = a+bx$ सरल रेखा समीकरण
3	2	$y = a+bx + cx^2$
4	3	$y = a+bx + cx^2 + dx^3$
5	4	$y = a+bx + cx^2 + dx^3 + ex^4$
n	*n*–1	$y = a+bx + cx^2 + dx^3 + ex^4 + \ldots + nX^{n-1}$

इसके बाद गणनाओं को सरल बनाने के लिए हम आन्तरगणन-पद को शून्य मानकर इसके सापेक्ष सभी स्वतंत्र चर मूल्यों (independent variable) के विचलन ज्ञात कर लेते हैं और इन विचलनों में से समावर्त्तक गुणक (common factor) निकाल देते हैं। x के इन मूल्यों और y के मूल्यों के आधार युगपत समीकरण (simultaneous equation) प्राप्त करके अचर पदों का मान ज्ञात कर लेते हैं।

उदाहरण (Illustration): 9

निम्न सारणी किसी फर्म की गत वर्षों की बिक्री प्रस्तुत करती है। परवलयिक-वक्र विधि (Parabolic curve method) द्वारा उसकी 1991 की बिक्री आन्तरगणित कीजिए-

वर्ष	:	1981	1985	1989	1993
बिक्री (लाख रु0)	:	100	112	136	180

हल-

वर्ष	1985	1989	1991	1993	1997
विचलन : x's	–6 –3	–2 –1	0 0	+2 +1	+6 +3
बिक्री : y's	100	112	y	136	180

ज्ञात मूल्यों की संख्या 4 है। इसलिए तीसरे घात के परवलय वक्र का समीकरण प्रयुक्त किया जाएगा-

$y = a+bx + cx^2 + dx^3$

उक्त समीकरण में ज्ञात मूल्य आदिष्ट करने पर निम्न 5 युगपत् समीकरणों की रचना की जाएगी-

$100 = a + (bx - 3) + (cx - 3^2) + (dx - 3^3)$

$100 = a – 3b + 9c – 27d$....... (i)

112 = a – b + c – d...... (ii) ण्ण्ण् ;पपद्ध

y = a...... (iii) ण्ण्ण् ;पपपद्ध

136 = a + b + c + d....... (iv) ण्ण्ण् ;पअद्ध

180 = a + 3b + 9c + 27d......... (v) ण्ण्ण् ;अद्ध

समीकरण (iii) के अनुसार y का मूल्य a के बराबर है इसलिए बाकी समीकरणों की सहायता से a का मूल्य निकाला जाएगा-

(ii) व (iv) को जोड़ने पर निम्न परिणाम प्राप्त होता है-

112 = a – b + c – d

136 = a + b + c + d

280 = 2a + 2c........ (vi)

इसी प्रकार (i) व (v) समीकरणों को जोड़ देने से निम्न समीकरण प्राप्त होता है-

100 = a – 3b + 9c – 27d

180 = a + 3b + 9c + 27d

280 = 2a + 18c......... (vii)

(vi) को 9 से गुणा करके उसमें से (vii) घटाकर निम्नलिखित परिणाम निकलता है-

2232 = 18a + 18c

280 = 2a + 18c

द्द्द्द्द्द्द्द्द्द्द्द्द्द्द्द्द्द्

1952 = 16a

$$a = \frac{1952}{16} = 122$$

क्योंकि a का मूल्य y के बराबर है इसलिए y = 122

1991 में उस संस्था की बिक्री का आन्तरगणित मूल्य 122 लाख रुपये ही आएगा। यदि इस प्रश्न में परवलय वक्र विधि द्वारा आन्तरगणन करने का निर्देश न हो तो इसे न्यूटन की विधि द्वारा करना ही उपयुक्त होगा।

(अ) अन्य रीतियाँ (other methods)

आन्तरगणन एवं बाह्यगणन की चार प्रमुख रीतियों के अतिरिक्त अन्य रीतियों का भी विशिष्ट परिस्थितियों में प्रयोग किया जा सकता है। इन रीतियों में से अधिकांश न्यूटन के प्रगामी अन्तर सूत्र के ही रूपान्तर है। यहाँ पर निम्न चार अन्य रीतियों का संक्षिप्त वर्णन किया गया है।

(क) न्यूटन-गॉस अग्रगामी विधि (Newton Gauss Forward Method)

रीति का प्रयोग कब किया जाए? यह रीति न्यूटन की प्रगामी अन्तर-विधि का ही एक संशोधित रूप है। इस रीति का प्रयोग उस दशा में किया जाता है जब (i) x श्रेणी समान अन्तर से बढ़ती हो तथा (ii) दिए हुए मूल्यों के अलावा किसी ऐसे x के लिए y_x का आन्तरगणन हो जो श्रेणी के मध्य में आता हो। ध्यान रहे , न्यूटन की प्रगामी अन्तर रीति और इस रीति द्वारा उत्तर एक समान निकलकर आएगा।

क्रिया विधि - (i) संकेताक्षर - स्वतंत्र चर माला अर्थात् x श्रेणी के जिस पद का आन्तरगणन करना होता है उससे ठीक पिछले पद को x_0 और उससे पिछले पदों को क्रमानुसार x_{-1}, x_{-2}, x_{-3} आदि तथा x_0 से अगले पदों को x_1, x_2, x_3 आदि संकेताक्षरों द्वारा प्रकट किया जाता है। ठीक यही क्रिया y श्रेणी के साथ दोहराई जाती है। मान लीजिए एक श्रेणी में छः पद है और x = 32 पद के लिए आन्तरगणन करना है तो न्यूटन की 'प्रगामी अन्तर रीति', 'न्यूटन-गॉस अग्रगामी' और 'न्यूटन गॉस पृष्ठगामी' रीति में संकेताक्षरों का प्रयोग अग्र तालिका के अनुसार किया जाएगा-

न्यूटन प्रगामी		न्यूटन-गॉस अग्रगामी		पृष्ठगामी रीति	
$x = 32$		$x = 32$		$x = 45$	
10 x_0	40 y_0	10 x_{-2}	40 y_{-2}	10 x_{-4}	40
20 x_1	52 y_1	20 x_{-1}	52 y_{-1}	20 x_{-3}	52
30 x_2	68 y_2	30 x_0	68 y_0	30 x_{-2}	68
40 x_3	72 y_3	40 x_1	72 y_1	40 x_{-1}	72
50 x_4	80 y_4	50 x_2	80 y_2	50 x_0	80
60 x_5	85 y_5	60 x_3	85 y_3	60 x_1	85

(ii) अन्तर सारणी - न्यूटन की प्रगामी रीति की भाँति इसमें भी अन्तर-सारणी की रचना की जाती है। अन्तरों के संकेत चिन्ह y's के चिन्हों के अनुकूल होते हैं- जैसे $\Delta^1_{y_0}, \Delta^2_{y_{-1}}, \Delta^3_{y_{-1}}$ आदि-

(iii) x का निर्धारण - x के अन्तर का निर्धारण निम्न सूत्र द्वारा किया जाता है-

$$x = \frac{\text{आन्तरगणन पद} - \text{पिछला पदनिकटतम}}{\text{पदों का अन्तर}}$$

(iv) न्यूटन गॉस अग्रगामी सूत्र इस प्रकार है-

$$y_x = y_0 + x\Delta^1_{y_0} + \frac{x(x-1)}{1\times 2}\Delta^2_{y_{-1}} + \frac{(x+1)\,x\,(x-1)}{1\times 2\times 3}\Delta^3_{y_{-1}} \ldots$$

उदाहरण (Illustration): 10

निम्न आँकड़ों की सहायता से न्यूटन-गॉस विधि द्वारा x = 25 के तत्संवादी y का मूल्य आन्तरगणित कीजिए।

स्वतंत्र चर (x) :	10	20	30	40
आश्रित चर (y) :	25	28	34	45

हल- **अन्तर-सारणी (न्यूटन-गॉस अग्रगामी विधि)**

x's	y's	अन्तर		
		प्रथम	द्वितीय	तृतीय

x_{-1}	25	y_{-1}	3	$\Delta^1_{y_{-1}}$				
x_0	28	y_0	6	$\Delta^1_{y_0}$	3	$\Delta^2_{y_{-1}}$	2	$\Delta^3_{y_{-1}}$
x_1	34	y_1	11	$\Delta^1_{y_1}$	5	$\Delta^2_{y_0}$		
x_2	45	y_2						
x	?	yx						

$$x = \frac{\text{आन्तरगणन पद − पिछला पदनिकटतम}}{\text{पदों का अन्तर}} = \frac{25-2030}{-20} = \frac{5}{10} = 0.5$$

$$y_x = y_0 + x\Delta^1_{y_0} + \frac{x(x-1)}{2}\Delta^2_{y_{-1}} + \frac{(x+1)\,x(x-1)}{2\times 3}\Delta^3_{y_{-1}}$$

$$y_x = 28 + 0.5\times 6 + \frac{0.5\times(-0.5)\times 3}{2} + \frac{1.5\times 0.5\times(-0.5)\times 2}{2\times 3}$$

$$= 28 + 3 - 0.375 - 0.125 \quad or \quad 31 - 0.5 = 30.5$$

न्यूटन की प्रगामी अन्तर-रीति द्वारा भी यही उत्तर आता है।

(ख) न्यूटन-गॉस पृष्ठगामी रीति (Newton Gauss Backward Method)

(i) रीति का प्रयोग- इस रीति का प्रयोग तब किया जाता है जब आन्तरगणन की जाने वाली संख्या श्रेणी के अन्तिम भाग में पड़ती हो। इस रीति की शेष विशेषताएं न्यूटन की सामान्य प्रगामी अन्तर विधि के ही समान है।

(ii) संकेतना एवं क्रिया विधि- संकेतना की दृष्टि से यह रीति उपर्युक्त रीति से थोड़ी भिन्नता लिए हुए है। इस रीति के अन्तर्गत आन्तरगणन पद (x) से अगले (succeeding) पद को x_0 माना जाता है और इस मूल-बिन्दु से पूर्व-पदों को अग्रगामी रीति की ही भाँति x_{-1}, x_{-2}, आदि और x_0 के बाद वाले पदों को x_1, x_2 आदि संकेताक्षरों द्वारा प्रकट किया जाता है। y श्रेणी के लिए भी ठीक इसी तरह से संकेताक्षरों का प्रयोग किया जाता है। तत्पश्चात् अन्तर-सारणी द्वारा अन्तर प्राप्त किए जाते हैं और x का मूल्य निकालने के लिए निम्न सूत्र का प्रयोग किया जाता है-

$$x = \frac{\text{आन्तरगणन पद − पिछला पदनिकटतम}}{\text{पदों का अन्तर}}$$

(iii) सूत्र: अन्त में निम्नलिखित सूत्र द्वारा आन्तरगणन किया जाता है-

$$y_x = y_0 - x\Delta^1_{y_{-1}} + \frac{(x+1)x}{1\times 2}\Delta^2_{y_{-1}} - \frac{(x+1)x(x-1)}{1\times 2\times 3}\Delta^3_{y_{-2}} + \frac{(x+1)x(x-1)(x-2)}{1\times 2\times 3\times 4}\Delta^4_{y_{-2}} \ldots$$

(ग) स्टर्लिंग का सूत्र (Sterling's Formula)

यह सूत्र न्यूटन-गॉस अग्रगामी व पृष्ठगामी दोनों सूत्रों का समान्तर माध्य है और श्रेणी के मध्यवर्ती पद के आश्रित मूल्य का आन्तरगणन करने के लिए उपयुक्त है। अग्रगामी विधि की भाँति इस रीति में भी आन्तरगणन पद से पहले के पद को ही मूल-बिन्दु (x_0, y_0) माना जाता है। सूत्र इस प्रकार है-

$$y_x = y_0 + x\left[\frac{\Delta^1_{y_0} + \Delta^1_{y_{-1}}}{2}\right] + \frac{x^2}{2}\Delta^2_{y_{-1}} + \frac{x(x^2-1)}{2\times 3}\left[\frac{\Delta^3_{y_{-1}} + \Delta^3_{y_{-2}}}{2}\right] + \ldots$$

उदाहरण 10 को स्टर्लिंग सूत्र के प्रयोग द्वारा भी हल किया जा सकता है-

$$y_x = 28 + 0.5\left[\frac{6+3}{2}\right] + \frac{(0.5)^2}{2}\times 3 + \frac{0.5(0.5^2-1)}{2\times 3}\times 2$$

$$= 28 + 0.5\times 4.5 + 0.375 + \frac{0.5\times(-0.75)\times 2}{2\times 3}$$

$$= 28 + 2.25 + 0.375 - 0.125 \ = \ 30.5$$

उपर्युक्त तीनों रीतियों का प्रयोग बहुत कम किया जाता है क्योंकि इनमें अन्तर निकालने और उपयुक्त सूत्र का प्रयोग करने में अधिकतर भ्रान्ति की स्थिति उत्पन्न हो जाती है। आन्तरगणन पद चाहे श्रेणी के आरम्भ , मध्य या अन्त में हो न्यूटन का मूल सूत्र (Newton's advancing differences formula) ही अधिकतर प्रयोग किया जाता है।

(घ) न्यूटन की विभाजित अन्तर-रीति (Newton's Method for divided differences)

रीति का प्रयोग - इस रीति का प्रयोग उस समय किया जाता है जब x श्रेणी के पदों का अन्तर असमान (unequal) हो।

क्रियाविधि - इस रीति के अनुसार पहले विभाजित अन्तर-सारणी (Table of divided differences) बनाई जाती है जिसमें निकटवर्ती y's के अन्तरों से भाग देकर विभाजित अन्तर निकाले जाते हैं। यदि y के चार मूल्य ज्ञात हो तो तीन प्रमुख विभाजित अन्तरों (Leading Divided Differences) का प्रयोग किया जाएगा। अन्तर के प्रथम खाने में तीन विभाजित अन्तर उपलब्ध होंगे जिनमें से पहला , प्रथम प्रमुख विभाजित अन्तर होगा। प्रत्येक y में से पिछले y को घटाकर उनके तत्संवादी x's के अन्तर से भाग कर दिया जाएगा। यही सम्बद्ध विभाजित अन्तर होगा। दूसरे खाने में प्रथम खाने के तीन विभाजितान्तरों की सहायता से इसी प्रकार दो विभाजित अन्तर निकाल लिए जाएंगे जिनमें से पहला, द्वितीय प्रमुख विभाजित अन्तर कहलाएगा। तीसरे खाने में दूसरे कॉलम के दो अन्तरों के आधार पर एकमात्र विभाजित अन्तर प्राप्त किया जाएगा जो तृतीय प्रमुख विभाजितान्तर कहलाएगा। विभाजित अन्तर निकालने की यह प्रक्रिया निम्नांकित सारणी में स्पष्ट की गयी है-

विभाजितान्तर सारणी (Table of Divided Differences)

x's	***y's***	विभाजित-अन्तर **(Divided Differences)**					
		प्रथम First		द्वितीय Second		तृतीय Third	
x_0	y_0	$\frac{y_1-y_0}{x_1-x_0}$	Δ^1_0				
x_1	y_1			$\frac{\Delta^1_1-\Delta^1_0}{x_2-x_0}$	Δ^2_0		
		$\frac{y_2-y_1}{x_2-x_1}$	Δ^1_1			$\frac{\Delta^2_1-\Delta^2_0}{x_3-x_0}$	Δ^3_0
x_2	y_2			$\frac{\Delta^1_2-\Delta^1_1}{x_3-x_1}$	Δ^2_1		
x_3	y_3	$\frac{y_2-y_1}{x_2-x_1}$	Δ^1_2				

x	y

इस विधि द्वारा आन्तरगणन करने के लिए निम्न सूत्र का प्रयोग किया जाएगा-

$$y_x = y_0 + (x - x_0)\Delta_0^1 + (x - x_0)(x - x_1)\Delta_0^2 + (x - x_0)(x - x_1)(x - x_2)\Delta_0^3 + \ldots$$

$\Delta_0^1, \Delta_0^2, \Delta_0^3$ क्रमशः प्रथम, द्वितीय, एवं तृतीय प्रमुख विभाजितान्तर हैं।

उदाहरण (Illustration): 11

निम्न सारणी में सम्पदा कर लगने वाली सम्पदाओं की संख्या दी गयी है -

सम्पदा वर्ग (रु0) :	50,000-75,000	75,000-1,00,000	1,00,000-1,50,000
संख्या :	870	540	450

न्यूटन की विभाजित अन्तर रीति द्वारा 75,000 रु. से 80,000 रु. के बीच की सम्पदाओं की संख्या का आन्तरगणन कीजिए।

हल- **विभाजित अन्तर सारणी**

से कम 000रु. x		संख्या y		विभाजित अन्तर			
				प्रथम Δ^1		द्वितीय Δ^2	
75	x_0	870	y_0	$\frac{1410-870}{100-75}$	Δ_0^1 21.6	$\frac{9-21.6}{150-75}$	Δ_0^2 -0.168
100	x_1	1410	y_1				
150	x_2	1860	y_2	$\frac{1860-1410}{150-100}$	Δ_1^1 9		
80	x	?	y				

सूत्र: $y_x = y_0 + (x - x_0)\,\Delta_0^1 + (x - x_0)(x - x_1)\Delta_0^2$

$$y_{80} = 870 + (80 - 75)\,21.6 + (80 - 75)(80 - 100)(-0.168)$$

$$= 870 + 108 + 16.8 = 994.8 \text{ or } 995$$

75 हजार रु. से कम वाली सम्पदाओं की संख्या = 870

80 हजार रु. से कम वाली सम्पदाओं की संख्या = 995

अतः 75 से 80 हजार रु. के बीच सम्पदाओं की संख्या = 995-870 = 125

7.8 सारांश (Summary)

सांख्यिकीय विश्लेषण करते समय कभी-कभी यह देखने में आता है कि प्रस्तुत समंक श्रेणी पूर्ण (complete) न होकर अपूर्ण (incomplete) होती है अर्थात् श्रेणी के कुछ मूल्य किन्हीं कारणों से अज्ञात बने रहते हैं। चूँकि एक सही निष्कर्ष पर पहुँचने के लिए श्रेणी के सभी मूल्यों की जानकारी का होना अत्यावश्यक है , इसलिए उपलब्ध समंकों के आधार पर उन अज्ञात मूल्यों का अनुमान लगाने के लिए जिन सांख्यिकीय रीतियों का प्रयोग किया जाता है उन्हें आन्तरगणन तथा बाह्यगणन कहते हैं। अर्थात् आन्तरगणन का तात्पर्य दिए गए समंकों से कुछ विशेष मान्यताओं के आधार पर किसी पद का सर्वाधिक सम्भावित मूल्य (most likely value) ज्ञात करने की विधि से है तथा बाह्यगणन से अभिप्राय किन्हीं विशेष मान्यताओं के आधार पर किसी भावी तिथि के भावी समंक का पूर्वानुमान करना होता है।

आन्तरगणन अथवा बाह्यगणन करते समय हम यह कल्पना कर लेते हैं कि समंकों में परिवर्तन की दर सर्वदा समान है एवं समंकों की तिथियों के बीच कोई अचानक घटना नहीं घटी है , अर्थात् समंकों में एक प्रकार की continuity है।

यदि ज्ञात समंकों में लगभग नियमित रूप से उच्चावचन होते हैं तो अज्ञात मूल्य का अनुमान भी यथासम्भव परिशुद्ध होता है।

7.9 अभ्यास प्रश्न (Practice Questions)

वस्तुनिष्ठ प्रश्न

(I) रिक्त स्थानों को भरिए-

(i) भारत में जनगणना प्रत्येक ----- में एक बार आयोजित की जाती है।

(ii) अभाव या अपर्याप्तता की पूर्ति आन्तरगणन द्वारा ----- लगाकर की जाती है।

(iii) आन्तरगणन व बाह्यगणन करते समय यह मान लिया जाता है कि दी हुई अवधि के समंकों में एकदम कोई ----- नहीं हुई हैं।

(iv) यदि ज्ञात समंकों में लगभग नियमित रूप से ----- होते हैं तो अज्ञात मूल्य का अनुमान भी यथासम्भव परिशुद्ध होता है।

(v) प्रत्यक्ष-द्विपद-विस्तार रीति ----- पर आधारित है।

(II) निम्न कथनों में से कौन सी सत्य है और कौन सी असत्य है

(i) आन्तरगणन का उद्देश्य समंक श्रेणी के बीच की रिक्तियों को भरना होता है। (सत्य / असत्य)

(ii) बाह्यगणन से अभिप्राय किन्हीं विशेष मान्यताओं के आधार पर किसी भावी तिथि के भावी समंक का पूर्वानुमान करना होता है। (सत्य / असत्य)

(iii) न्यूटन की प्रगामी-अन्तर विधि द्विपद-प्रमेय पर आधारित नहीं है। (सत्य / असत्य)

(iv) लाग्रैंज द्वारा प्रतिपादित लांग्रैज विधि आन्तरगणन एवं बाह्यगणन की सार्वभौमिक रीति है। (सत्य / असत्य)

(v) 'n' वें घात के परवलयिक वक्र का समीकरण इस प्रकार है-

$y = ax + bx^2 + cx^3 + dx^4$ ... (सत्य / असत्य)

(III) निम्नलिखित में कौन सा विकल्प सही है

(i) किन्हीं निश्चित मान्यताओं के अन्तर्गत मात्राओं के सर्वाधिक सम्भाव्य अनुमान लगाने की तकनीक को कहते हैं।

(अ)बाह्यगणन (ब) सहसम्बन्ध (स) प्रतीपगमन (द) आन्तरगणन

(ii) निम्नांकित रीतियों में कौन सी रीति सार्वभौमिक रीति है-

(अ) बिन्दु-रेखीय रीति (ब) प्रत्यक्ष द्विपद-विस्तार रीति

(स) न्यूटन की प्रगामी अन्तर रीति (द) न्यूटन-गॉस अग्रगामी विधि

(iv) स्टर्लिंग का सूत्र है-

(अ) $y_x = y_0 + x\Delta^1_{y_0} + \frac{x(x-1)}{1\times 2}\Delta^2_{y_{-1}} + \frac{(x+1)x(x-1)}{1\times 2\times 3}\Delta^3_{y_{-1}} \ldots$

(ब) $y_x = y_0 - x\Delta^1_{y_1} + \frac{(x-1)x}{1\times 2}\Delta^2_{y_{-1}} + \frac{(x+1)x(x-1)}{1\times 2\times 3}\Delta^3_{y_{-2}} \ldots$

(स) $y_x = y_0 + x\left[\frac{\Delta^1_{y_0} + \Delta^1_{y_{-1}}}{2}\right] + \frac{x^2}{2}\Delta^2_{y_{-1}} + \frac{x(x^2-1)}{2\times 3}\left[\frac{\Delta^3_{y_{-1}} + \Delta^3_{y_{-2}}}{2}\right] + \ldots$

(द) $y_x = y_0 + (x-x_0)\Delta^1_0 + (x-x_0)(x-x_1)\Delta^2_0 + (x-x_0)(x-x_1)(x-x_2)\Delta^3_0 + \ldots$

(v) आन्तरगणन तथा बाह्यगणन की रीतियों को बाँटा जा सकता है-

(अ) दो श्रेणियों में (ब) तीन श्रेणियों में

(स) चार श्रेणियों में (द) पाँच श्रेणियों में

(IV) लघु उत्तरात्मक प्रश्न

(i) आन्तरगणन एवं बाह्यगणन की दो अन्तर्निहित मान्यताएँ लिखिए।

(ii) आन्तरगणन एवं बाह्यगणन की परिशुद्धता किन दो बातों पर निर्भर करती है?

(iii) $(y - 1)^8$ का द्विपद विस्तार लिखिए।

(iv) न्यूटन प्रगामी विधि के अनुसार तृतीय प्रमुखान्तर (Δ^3_0) तक का अन्तर सारणी बनाइए।

(v) n = 8 तक एक पास्कल त्रिभुज की रचना कीजिए।

(V) निबन्धात्मक प्रश्न

(i) सांख्यिकी में आन्तरगणन एवं बाह्यगणन की उपयोगिता की व्याख्या कीजिए।

(ii) आन्तरगणन से आप क्या समझते हैं? किन मान्यताओं के अन्तर्गत मूल्य की अन्तर्गणना की जाती है?

(iii) आन्तरगणन एवं बाह्यगणन में अन्तर स्पष्ट कीजिए। सांख्यिकीय अध्ययन में उनकी आवश्यकता का संक्षिप्त विवेचन कीजिए।

(iv)आन्तरगणन एवं बाह्यगणन के लिए प्रयोग होने वाली प्रमुख रीतियों का वर्णन कीजिए। वे अवस्थाएँ भी बताइए जिनमें प्रत्येक का प्रयोग उचित रहेगा।

(v) जनगणना के वर्षों के मध्य के वर्षों के परिवर्तनों की गणना आप कैसे करेंगे ? क्या आप 1941, 1951, 1961, 1971 की जनगणना के अंकों के आधार पर 1976 की जनगणना का अनुमान कर सकते हैं?

(VI) संख्यात्मक प्रश्न

(i) निम्न सारणी एक फर्म के 2005 से 2010 तक के लाभ (लाख रु.) के समंक प्रस्तुत करती है। 2009 के लाभ की राशि अज्ञात है। बिन्दु रेखीय रीति द्वारा उसका आन्तरगणन कीजिए।

वर्ष	:	2005	2006	2007	2008	2009	2010

लाभ	:	108	113	111	110	?	114

(ii) 30 वर्ष की आयु के लिए अज्ञात मूल्य का आन्तरगणन कीजिए-

आयु (वर्षों में)	:	10	15	20	25	30	35
मूल्य	:	20	35	40	43	?	55

(iii) नीचे दी गयी सारणी की सहायता से 1970 तथा 1980 के उत्पादन की गणना कीजिए -

वर्ष	:	1955	1960	1965	1970	1975	1980	1985
उत्पादन	:	200	230	270	?	380	?	460

(iv) निम्न समंकों से , आन्तरगणन की न्यूटन विधि द्वारा 25 वर्ष की आयु पर वार्षिक शुद्ध प्रीमियम ज्ञात कीजिए-

आयु (वर्षों में)	:	20	24	28	32
प्रीमियम दर (रु0 में)	:	0.01427	0.01581	0.01772	0.01996

(v) लाग्रेंज का सूत्र प्रयोग करके 6 और 7 हजार के बीच आय कमाने वाले व्यक्तियों की संख्या अनुमानित कीजिए।

आय (हजारों में)	:	0-2	2-3	3-6	6-8
व्यक्तियों की संख्या	:	10	15	35	20

7.10 अभ्यास प्रश्नों के उत्तर (Answer for Practice Questions)

(I) (i) - दशक (ii) - सर्वोपयुक्त अनुमान

(iii) प्रचण्ड वृद्धि (iv) - उच्चावचन (v) - द्विपद-प्रमेय

(II) (i) – स (ii) - स (iii) - अ (iv) - स (v) - अ

(III) (i) - ड (ii) - ड (iii) - अ (iv) - स (v) - अ

(i) – (112 लाख रु.), (ii) – (48)

(iii) – (1970 = 322, 1980 = 433)

(iv)- (0.01625)

(v) – [y_7= 70, 70 -60 = 10]

7.11 संदर्भ ग्रन्थ सूची/उपयोगी पाठ्य सामग्री (Bibliography)

1) बंसल, एस0 एन0, एवं अग्रवाल, डी0 आर0, (1978), सांख्यिकी के मूल तत्व , शिवलाल अग्रवाल एण्ड कम्पनी, आगरा - 31

2) नागर, कैलाश नाथ, (2005), सांख्यिकी के मूल तत्व, मिनाक्षी प्रकाशन, मेरठ।

3) लाल, एस0 एन0, चतुर्वेदी, एस0, सांख्यिकी, प्रकाशन, इलाहाबाद।

4) सिंह, एस0 पी0, (1997) सांख्यिकी-सिद्धान्त एवं व्यवहार, एस0 चन्द एण्ड कम्पनी लिमिटेड, नई दिल्ली।

5) अवस्थी, जी0 डी0 एवं निगम, सुधीर कुमार, (2007) सांख्यिकीय विश्लेषण, भारत बुक सेन्टर, लखनऊ।

6) Gupta, S.P., (2005), Statistical Methods, S. Chand, New Delhi.

7) Goon, Gupta and Dasgupta, A Fundamental of Statistics, Vol. I, The World Press Private Limited.

इकाई-8 सूचकांक
(Index)

8.1 प्रस्तावना (Introduction)

प्रस्तुत इकाई में सूचकांक या निर्देशांक के विषय का विस्तृत रूप से अध्ययन किया गया है। परिवर्तन प्रकृति का ही एक नियम है। आर्थिक क्षेत्र में यह नियम पूर्ण रूप से लागू होता है। आर्थिक क्षेत्र के विभिन्न पहलू जैसे-मूल्य , उत्पादन, व्यापार, जीवन लागत इत्यादि निरन्तर परिवर्तित होते रहते हैं। इनमें सतत उतार-चढ़ाव होता रहता है। इन्हीं परिवर्तनों का अध्ययन करने और इनके प्रभावों को स्पष्ट करने के लिए जिस सांख्यिकीय तकनीक को विकसित किया गया है उसी तकनीक को सूचकांक अथवा निर्देषांक कहते हैं।

सूचकांकों का निर्माण सर्वप्रथम इटली के सांख्यिक कर्प्ली ने किया था। इन्होंने 1764 में मू ल्यो में होने वाले परिवर्तनों की माप करने हेतु सन् 1500 ई. को आधार वर्ष मानकर सन् 1750 के लिए मूल्य सूचकांक का निर्माण किया। प्रारम्भ में सूचकांक को केवल मूल्यस्तर तथा मुद्रा की क्रय शक्ति का माप करने हेतु प्रयोग किया जाता था परन्तु आज के समय में इसका प्रयोग विस्तृत हो गया है। प्रत्येक पहलू में सूचकांक का प्रयोग किया जाता है। जैसे - उत्पादन, उपभोग, निर्यात, आयात, राष्ट्रीय आय, जीवन निर्वाह व्यय, सड़क दुर्घटनाओं जैसे संख्यात्मक तथ्य एवं निर्धनता , स्वास्थ्य, मानवीय विकास, कार्यक्षमता, बुद्धिमता आदि के अध्ययन में भी सूचकांकों का प्रयोग किया जाता है।

सूचकांको के विकास में प्रो. जेवन्स, डॉ. मार्शल, वाल्श, एजवर्थ आदि उल्लेखनीय नाम हैं परन्तु 100 सूत्रों का प्रतिपादन करके नया आयाम देने वाले फिषर का नाम उल्लेखनीय है।

8.2 उद्देश्य (Objectives)

प्रस्तुत इकाई का अध्ययन करने से आप-

- ✓ सूचकांक का आर्थिक एवं व्यावसायिक अध्ययन के महत्व को समझेंगे।
- ✓ सूचकांकों की उपयोगिता एवं सीमाओं को जानेंगे।
- ✓ सूचकांकों के निर्माण को जानेंगे।
- ✓ विभिन्न प्रकार के सूचकांकों का अध्ययन करेंगे।

8.3 सूचकांक का अर्थ एवं परिभाषा (Meaning and Definition of Index)

सूचकांक एक विशेष प्रकार का माध्य है जिनके द्वारा समय , स्थान या अन्य किसी विशेषता के आधार पर सम्बन्धित चर मूल्यों में होने वाले सापेक्ष परिवर्तनों का मापन किया जाता है।

(क) **क्रॉक्सटन एवं काउडेन** के अनुसार, "सूचकांक सम्बन्धित चर मूल्यों के आकार में होने वाले अन्तरों का मापन करने के साधन या उपाय हैं।"

(ख) **डॉ. एल. बाउले** के अनुसार, "सूचकांकों का प्रयोग किसी मात्रा में होने वाले ऐसे परिवर्तनों का माप करने के लिए किया जाता है जिनका हम प्रत्यक्ष रूप से अवलोकन नहीं कर सकते।"

(ग) **सेक्राइस्ट** के अनुसार, "निर्देशांक अंकों की एक ऐसी श्रेणी है जिसके द्वारा किसी तथ्य के परिमाण में होने वाले परिवर्तनों का समय या स्थान के आधार पर मापन किया जाता है।"

(घ) **डॉ. एम. टटिल** के शब्दों में, "निर्देशांक एक अकेले अनुपात के रूप में दो विभिन्न समयों , स्थानों अथवा परिस्थितियों में विभिन्न चरों में परिवर्तन को सामूहिक रूप से मापता है।"

निष्कर्ष रूप में कहा जा सकता है कि सूचकांक प्रतिशत के रूप में व्यक्त किया जाने वाला एक विशेष प्रकार का माध्य है जिसके आधार पर विभिन्न समयों , स्थानों या अन्य समंक समूहों में होने वाले सापेक्षिक परिवर्तनों की सामान्य प्रकृति को मापा जाता है।

जब हम केवल किसी अकेले चर का अध्ययन करते हैं तो वह एक चरीय सूचकांक कहलाता है जबकि कई चरों में होने वाले औसत परिवर्तनों का एक साथ अध्ययन किया जाये तो इसे मिश्रित सूचकांक कहते हैं। जैसे कृषि उत्पादन का सूचकांक मिश्रित सूचकांक का उदाहरण है।

8.4 सूचकांको की विशेषताएँ (Characteristics of Index)

(क) तुलना का आधार (basis of comparison) - सूचकांकों द्वारा परिवर्तनों की तुलना समय अथवा स्थान के आधार पर की जाती है , जिस वर्ष के सूचकांक ज्ञात करने हो उसे चालू वर्ष या प्रचलित वर्ष तथा जिस निश्चित वर्ष से तुलना करनी हो उसे आधार वर्ष कहते हैं।

(ख) विशिष्ट प्रकार के माध्य (special type of mean) - माध्यों द्वारा असमान इकाईयों वाली श्रेणी की तुलना नहीं की जा सकती, परन्तु सूचकांकों द्वारा असमान इकाईयों वाली अनेक श्रेणियों में होने वाले परिवर्तनों का सापेक्ष अध्ययन सरलता से किया जा सकता है।

(ग) प्रत्यक्ष मापन न होने वाले परिवर्तनों का माप (measuring changes that are not directly measurable) - सामान्यतः सूचकांक की तकनीकी का प्रयोग ऐसे मिश्रित एवं जटिल परिवर्तनों के माप के लिए किया जाता है जिनको प्रत्यक्ष रूप से मापा नहीं जा सकता। जैसे मूल्य स्तर , जीवन लागत अथवा आर्थिक क्रियाओं में परिवर्तन। यहाँ सूचकांक की सहायता से सापेक्ष परिवर्तनों का अध्ययन कर लिया जाता है।

(घ) तुलनात्मक माप (comparative measurement) - सूचकांक एक तुलनात्मक अथवा सापेक्ष माप है। उदाहरण के लिए यदि यह कहा जाये कि सन् 1990 की तुलना में सन् 1998 में मूल्य निर्देशांक 160 है तो इसका अर्थ यह है कि 1990 की तुलना में 1998 में मूल्यों में 60 प्रतिशत वृद्धि हो गयी है।

(ङ) सार्वभौमिक उपयोग (universal use) - इस तकनीक का प्रारम्भ मूल्यों में परिवर्तन के मापन के लिए हुआ था, लेकिन वर्तमान समय में इस तकनीक का प्रयोग सर्वव्यापी बन गया है। इसके द्वारा विभिन्न क्षेत्रों में परिवर्तन का मापन किया जाता है।
वास्तव में ऐसा कोई क्षेत्र नहीं है जिसमें संख्यात्मक को मापने के लिए सूचकांकों का प्रयोग न होता हो।

8.5 सूचकांको का महत्त्व एवं उपयोग (Importance and Use of Index)

आर्थिक और व्यावसायिक क्षेत्र में परिवर्तनों के मापन और विश्लेषण की दृष्टि से सूचकांक एक महत्त्वपूर्ण एवं उपयोगी उपकरण बन चुका है और इस कारण इसके "आर्थिक वायुमापक यंत्र" कहा जाने लगा है। सूचकांक आर्थिक जगत का प्राण है क्योंकि उत्पादन , उपभोग, मुद्रा मूल्य, मांग, पूर्ति, मजदूरी, आयात-निर्यात, मूल्य स्तर जैसी प्रमुख समस्याओं का समाधान सूचकांकों के प्रयोग द्वारा ही किया जाता है। संक्षेप में सूचकांक के प्रयोग एवं महत्त्व निम्न प्रकार से हैं-

(क) तुलनात्मक अध्ययन को सम्भव बनाना (Making comparative studies possible) - सूचकांको की सापेक्ष माप के द्वारा किसी समय या स्थान के आधार पर सरलता से तुलना की जा सकती है क्योंकि सूचकांकों के द्वारा विभिनन प्रकृति की इकाईयों को एक अर्थपूर्ण एवं सरल संख्यात्मक मूल्य में परिवर्तित कर लिया जाता है।

(ख) भावी प्रवृत्तियों के संकेतक (Indicators of future trends) - सूचकांक भूतकाल के सन्दर्भ में वर्तमान की व्याख्या करते हैं और भविष्य के लिए पुर्वानुमान लगाने में सहायक सिद्ध होते हैं।

(ग) आर्थिक नीतियों के निर्माण में सहायक (Helpful in formulating economic policies) - सूचकांको द्वारा सरकार मूल्यों में स्थिरता, न्यूनतम मजदूरी, मंहगाई भत्ता आदि सुनिष्चित कर लेती है। साथ ही विभिन्न आर्थिक नीतियों के निर्माण में भी सहायता मिलती है।

(घ) जटिल तथ्यों को सरल बनाना (Simplifying complex facts) - सूचकांकों की सहायता मे जटिल तथ्यों एवं उनके परिवर्तनों के अध्ययन को सरल बनाया जा सकता है। उदाहरण के लिए किसी देष में व्यावसायिक क्रियाओं में परिवर्तन एक जटिल तथ्य है जिसमें उद्योग , व्यापार, बैंकिंग, परिवहन आदि विभिन्न क्षेत्रों में होने वाले परिवर्तन शामिल होते हैं लेकिन सूचकांक द्वारा इसका अध्ययन सरलता से किया जा सकता है।

(ङ) विभिन्न मूल्यों की अवस्फीति में सहायक (Helpful in deflation of various prices) - सूचकांक हमें सिद्धान्त से व्यवहार की ओर ले जाते हैं और वास्तविक स्थिति की जानकारी देते हैं। सूचकांकों से अवस्फीतिकरण होने के कारण वास्तविक जानकारी प्राप्त होती है।

8.6 सूचकांको के प्रकार (Types of Index)

सूचकांक को निम्न रूपों में वर्गीकृत किया जाता है-

1. कीमत या मूल्य सूचकांक (Price or Value Index)
2. मात्रा सूचकांक (Quantity Index)
3. कुल मूल्य सूचकांक या वैल्यू सूचकांक (Total Price Index or Value Index)
4. उद्देश्य विशेष सूचकांक (Purpose Specific Index)
5. वस्तुओं की संख्या के आधार पर (Based on number of commodities)

1. मूल्य सूचकांक (Price or Value Index) - इन सूचकांको के माध्यम से मूल्यों में या मूल्य स्तर में परिवर्तन को मापा जा सकता है। इन्हें दो उपवर्गों में बांटा जा सकता है- थोक मूल्य सूचकांक व निर्वाह व्यय सूचकांक।

2. मात्रा सूचकांक (Quantity Index) - इस प्रकार के सूचकांक भौतिक मात्रा में कमी या वृद्धि को मापने के लिए तैयार किये जाते हैं। जैसे- कृषि उत्पादन सूचकांक, औद्योगिक उत्पान सूचकांक आदि।

3. कुल मूल्य सूचकांक या वैल्यू सूचकांक (Total Price Index or Value Index) - इन सूचकांक का उद्देश्य आधार वर्ष के कुल मूल्य (मात्रा x कीमत) की तुलना में चालू वर्ष के कुल मूल्य में परिवर्तन का अध्ययन किया जाना है। **जैसे** - विक्रय राशि का सूचकांक।

4. उद्देश्य विशेष सूचकांक (Purpose Specific Index) - आर्थिक एवं व्यावसायिक क्षेत्र में किसी विशिष्ट उद्देश्य के लिए भी सूचकांक तैयार किये जा सकते हैं। जैसे-राष्ट्रीय आय सूचकांक , विकास दर सूचकांक , उत्पादकता सूचकांक आदि।

5. वस्तुओं की संख्या के आधार पर सूचकांक (Based on number of commodities) - यदि किसी एक वस्तु के मूल्य के आधार पर सूचकांक तैयार किया जाता है , तो उसे सरल सूचकांक कहते हैं। यदि वस्तुओं के समूह के लिए सूचकांक बनाया जाता है तो उसे संयोजित या 'सकल' सूचकांक कहा जाता है।

8.7 सूचकांक रचना सम्बन्धी सूचनाएँ (Information Related to Index Construction)

(1) सूचकांको का उद्देश्य - सूचकांक बनाते समय सबसे पहले उसके उद्देश्य को निश्चित कर लेना जरूरी है। क्योंकि वस्तुओं का चुनाव, मूल्य उद्धरण, भारों का निर्धारण, सूचकांक की किस रीति का प्रयोग करना है जैसी महत्त्वपूर्ण बातें सही अर्थों में, सूचकांक के उद्देश्य पर ही निर्भर करती है।

(2) वस्तुओं का चयन - दूसरा चरण वस्तुओं के चुनाव का होता है क्योंकि सभी वस्तुओं को शामिल करना आवश्यक नहीं होता। कौन सी वस्तुएँ कितनी संख्या में चुनी जाये उनकी किस्म क्या हो एवं उन्हें कैसे वर्गीकृत करें, ये सब ध्यान देने योग्य है।

(A) वस्तुएँ -

निम्न विशेषताओं वाली वस्तुओं का चुनाव करनी चाहिए-

(i) प्रतिनिधि एवं लोकप्रिय - वस्तुएँ ऐसी होनी चाहिए जो सम्बन्धित वर्ग या क्षेत्र के लोगों में लोकप्रिय हो एवं उनकी आदतें, रीति-रिवाजों व आवश्यकताओं का प्रतिनिधित्व करें।

(ii) पहचानने योग्य - वस्तुएँ ऐसी होनी चाहिए जो आसानी से पहचानी जायें व उनका स्पष्ट वर्णन किया जा सकें।

(iii) प्रमाणित एवं सजातीय - चुनी जाने वाली वस्तुएँ श्रेणीबद्ध व प्रमाणित होनी चाहिए उनकी किस्म में एकरूपता होनी चाहिए।

(B) वस्तुओं की संख्या - इसके सन्दर्भ में कोई दृढ़ व निश्चित नियम नहीं है लेकिन वस्तुओं की संख्या का निर्धारण उपलब्ध समय व धन , शुद्धता तथा उद्देश्य व परिस्थितियों को ध्यान में रखकर करना चाहिए।

(C) किस्म - सूचकांक में ऐसी वस्तुओं को शामिल करना चाहिए , जो सबसे अधिक प्रचलित व प्रमाणित हों। साथ ही उसके गुणों में भी स्थिरता होनी चाहिए।

(D) वर्गीकरण - चुनी गयी वस्तुओं को सजातीयता के आधार पर कुछ निश्चित वर्गों व उपवर्गों में विभाजित कर देते हैं, जिससे सामान्य सूचकांक के साथ-साथ समूह सूचकांक भी ज्ञात किया जा सके।

(3) मूल्य उद्धरण प्राप्त करना - वस्तुओं के चुनाव के बाद मूल्य उद्धरणों का चरण होता है। इसमें निम्न बातों पर विचार होता है-

(i) थोक या फुटकर मूल्य - सूचकांकों के उद्देश्य के आधार पर मूल्य थोक या फुटकर मूल्य हो सकते हैं। परन्तु अधिकांशतः थोक मूल्य ही लिये जाते हैं। क्योंकि वे फुटकर मूल्यो की अपेक्षा अधिक स्थिर होते हैं।

(ii) मूल्य व्यक्त करने का रूप - मूल्य दो प्रकार से व्यक्त किये जा सकते हैं-

a) मुद्रा मूल्य जैसे-1000 रूपये प्रति कुन्तल।
b) मात्रा मूल्य या प्रति लोग मूल्य- जैसे 2 किलो प्रति रूपये।

सूचकांक रचना में सदैव मुद्रा मूल्य का ही प्रयोग किया जाना चाहिए।

(iii) मूल्य उद्धरणों की आवृत्ति या संख्या - वस्तु के महत्त्व के अनुसार यह तय करते हैं कि मूल्य उद्धरण यह कितनी बार व किस अन्तराल से लिये जायें अर्थात् साप्ताहिक या मासिक आधार पर। मूल्य उद्धरणों की संख्या जितनी अधिक होगी, शुद्धता भी उतनी अधिक होगी, परन्तु जटिलता भी बढ़ जायेगी। मूल्य उद्धरणों की आवृत्ति सूचकांक के उद्देश्य, अवधि, उपलब्ध साधन व शुद्धता के स्तर पर निर्भर होती है।

(iv) मूल्य उद्धरण प्राप्ति के स्थान - सूचकांकों के लिए वस्तुओं के मूल्य उद्धरण उन बाजार से प्राप्त करना चाहिए जहाँ वस्तुओ का क्रय-विक्रय वृहद स्तर पर होता है। मूल्य उद्धरण के स्त्रोत स्वतंत्र , निष्पक्ष, विश्वसनीय व उपयुक्त होने चाहिए। विभिन्न पत्र-पत्रिकाओं , रेडियो, दूरदर्शन, कम्प्यूटर, इंटरनेट व अन्य सरकारी व अर्धसरकारी सूत्रों से भी मूल्य सूचना प्राप्त की जा सकती हैं।

(v) मूल्य उद्धरण का औसत निकालना- मूल्य उद्धरणों के बारे में अंतिम चरण औसत निकालना है।

(4) आधार वर्ष का चुनाव तथा मूल्यानुपातों का परिकलन - मूल्य सूचकांक एक प्रमाप वर्ष के आधार पर प्रचलित वर्ष के मूल्य स्तर को व्यक्त करते हैं। पिछला प्रमाप वर्ष जो आगामी वर्षों के तुलनात्मक अध्ययन का आधार होता है, आधार वर्ष कहलाता है।

एक अच्छे आधार वर्ष के सम्बन्ध में यह बाते महत्त्वपूर्ण हैं -

(i) वर्ष सामान्य हो (ii) वास्तविक हो

(iii) उस काल की समस्त सूचनायें उपलब्ध हो (iv) वर्ष अधिक पुराना न हो

आधार वर्ष ज्ञात करने की दो रीतियां हैं-

(i) स्थिर आधार रीति

(a) एक वर्षीय आधार (b) औसत अवधि आधार

(ii) श्रृंखला आधार रीति

(a) एक वर्षीय स्थायी आधार - इस रीति में किसी एक सामान्य वर्ष को आधार वर्ष के रूप में चुन लिया जाता है।

स्थिर आधार के मूल्य को 100 मानकर निकाला गया प्रचलित वर्ष का प्रतिशत ही मूल्यानुपात कहलाता है। आधार वर्ष के मूल्य को P_0 एवं चालू वर्ष के मूल्य के P_1 द्वारा व्यक्त किया जाता है।

$$\text{मूल्यानुपात (R)} = \frac{\text{Current year's price}(P_1)}{\text{Base year's price }(P_0)} \times 100$$

$$(R) = \frac{P_1}{P_0} \times 100$$

(b) औसत मूल्य आधार (Average Price Base)-

$$\text{मूल्यानुपात (R)} = \frac{\text{Current year's price}(P_1)}{\text{Average price }(P_0)} \times 100 \qquad (R) = \frac{P_1}{P_0} \times 100$$

यदि एक ही वस्तु के विभिन्न वर्षों के मूल्य दिये हैं तो इनके मुल्यानुपात ही अभीष्ट सूचकांक हैं। इसके विपरीत प्रत्येक वर्ष के कई वस्तुओं के मूल्य दिये हो तो विभिन्न वस्तुओं के मूल्यानुपात का सामान्तर माध्य ही सम्बन्धित प्रचलित वर्ष का सूचकांक होता है

अर्थात् **चालू वर्ष का सूचकांक** $\frac{\sum R}{N} = \frac{\text{मूल्यानुपातों का योग}}{\text{वस्तुओं की संख्या}}$

उदाहरण -1 - सन् 1998 को आधार मानकर विभिन्न वर्षों के सूचकांक तैयार कीजिए-

Year	1998	1999	2000	2001	2002
Price	44	48	46	52	50

हल:-

Year	Price	$\frac{P_1}{P_0} \times 100$	Index No.
1998	40	--	100
1999	48	$\frac{48 \times 100}{40}$	120

2000	46	$\frac{46 \times 100}{40}$	115
2001	52	$\frac{52 \times 100}{40}$	130

उदाहरण 2 - निम्न समंकों से 1998 से 2000 तक के औसत मूल्य को आधार मानकर मूल्यानुपात की गणना कीजिए।

Year	1998	1999	2000	2001	2002
Price	44	49	57	55	58

हल - $P_0 = \frac{Price\ \ from\ 1998\ to\ 2000}{3} = \frac{44 + 49 + 57}{3} = \frac{150}{3}$ = 50

Year	Price	$\frac{P_1}{P_0} \times 100$	Index No. (PR)
1998	44	$\frac{44 \times 100}{40}$	88
1999	49+	$\frac{49 \times 100}{40}$	98
2000	57	$\frac{57 \times 100}{40}$	114
2001	55	$\frac{55 \times 100}{40}$	110
2002	58	$\frac{58 \times 100}{40}$	110

(ii) श्रृंखला आधार रीति - इस रीति को "**चल आधार रीति**" भी कहा जाता है और इसके आधार पर बना सूचकांक श्रृंखला आधार सूचकांक कहा जाता है। इस रीति में प्रत्येक चालू वर्ष के लि ए हैं। उसका पिछला वर्ष आधार माना जाता है। उदाहरण के लिए 1998 से 2002 तक के वर्षों के सूचकांक तैयार करने हैं तो सन् 1998 के लिए 1997, सन् 1999 के लिए 1998, सन् 2000 के लिए 1999 और ऐसे ही आगे आधार वर्ष माने जायेंगे।

उदाहरण 3 - निम्न समंकों से श्रृंखला आधार सूचकांक बनाइये-

Year	1998	1999	2000	2001	2002
Price	80	120	132	264	396

हल -

गुण एवं दोष - इस आधार का प्रमुख गुण यह है कि इससे तात्कालिक परिवर्तनों का पता चल जाता है। प्रत्येक वर्ष में होने वाले परिवर्तनों की तुलना पिछले वर्ष के परिवर्तनों से की जा सकती है। यह तुलना व्यापारी व

उद्योगपति के लिए बहुत उपयोगी होती है। दूसरे श्रृंखला आधार वाले सूचकांक में आवष्यकतानुसार पुरानी वस्तुओं को हटाकर उनके स्थान पर नई वस्तुओं का समावेश किया जा सकता है। परन्तु श्रृंखला रीति के अनुसार बनाये गये सूचकांक दीर्घकालीन प्रवृत्ति स्पष्ट नहीं करते। इन सूचकांकों की रचना तुलनात्मक रूप से कठिन होती है। यदि किसी एक स्थान पर अशुद्धि हो जाये तो आगे सभी गणनाओं पर उसका प्रभाव पड़ेगा।

श्रृंखला आधार सूचकांकों के निर्माण की क्रिया विधि:-

इसका सूत्र निम्नलिखित है-

$$\text{श्रृंखला मूल्यानुपात} = \frac{\text{चालू वर्ष का मूल्य}}{\text{पिछले वर्ष का मूल्य}} \times 100$$

$$\text{L.R (Link Relative)} = \frac{\text{Current year's price}}{\text{Previous year's price}} \times 100$$

श्रृंखला मूल्यानुपातों को किसी एक ही स्थिर वर्ष पर आधारित करना-

श्रृंखला मूल्यानुपातों द्वारा प्रत्येक वर्ष की पिछली वर्ष से तुलना करते हैं। इस प्रकार दो निकटवर्ती वर्षों में कड़ियाँ स्थापित हो जाती हैं। इन कड़ियों से एक श्रृंखला बन जाती है जिससे सभी वर्षो के परिवर्तन एक निश्चित वर्ष से श्रृंखला हो जाये। इस प्रकार से श्रृंखलित सूचकांक कहते हैं। इसे ज्ञात करने का निम्न सूत्र है-

$$\textbf{चालू वर्ष का सूचकांक} = \frac{\text{गत वर्ष का श्रृंखलित सूचकांक} \times \text{चालू वर्ष का औसत श्रृंखला मूल्यानुपात}}{100}$$

Chain Index for current year =

$$\frac{\text{Previous Year's chain index} \times \text{Current year's average link Relatives}}{100}$$

स्थिर आधार एवं श्रृंखला आधार का अन्तर -

(i) स्थिर आधार में आधार वर्ष स्थिर रहता है और आगे के वर्षों की तुलना इसी आधार वर्ष से की जाती है जबकि श्रृंखला आधार में आधार प्रति वर्ष बदलता रहता है और प्रत्येक वर्ष की तुलना पिछले वर्ष से करते हैं।

(ii) स्थिर आधार सूचकांकों की सहायता से दीर्घकालीन प्रवृत्ति का पता चलता है जबकि श्रृंखला आधार सूचकांक वर्ष प्रतिवर्ष के परिवर्तनों को प्रकट करते हैं।

(iii) स्थिर आधार सूचकांक में शामिल वस्तुओं में परिवर्तन नहीं किये जा सकते , जबकि श्रृंखला सूचकांक में प्रतिवर्ष वस्तु या पद में परिवर्तन किये जा सकते हैं।

(iv) स्थिर आधार सूचकांक की रचना मुल्यानुपातों के आधार पर की जाती है परन्तु श्रृंखला आधार सूचकांकों के निर्माण में श्रृंखला आधार सूचकांकों के निर्माण में श्रृंखला मूल्यानुपातों का उपयोग किया जाता है।

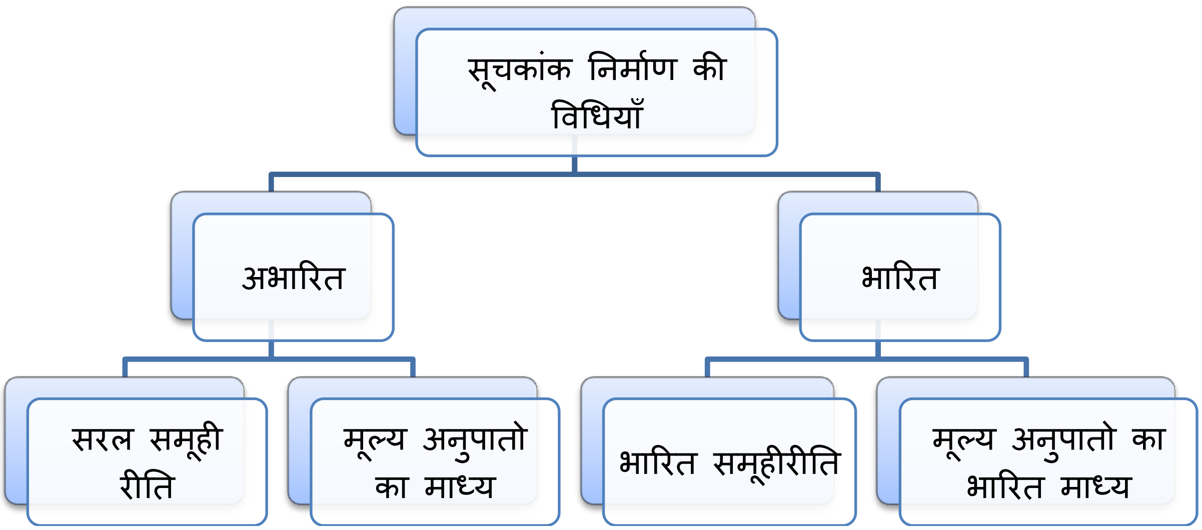
उदाहरण 4- निम्न तालिका से सन् 1998 से 2002 तक के तीन वस्तुओं के औसत थोक मूल्य दिये गये हैं। श्रृंखला आधार रीति से सूचकांकों की रचना कीजिए।

Commodities	Average Wholesale Price				
	1998	1999	2000	2001	2002
I	5	6	8	8	10
II	8	10	12	15	18
III	10	12	15	18	20

P = Price LR = Link Relatives

8.8 सूचकांको की सीमाएँ (Limitations of Index)

विभिन्न विधियों को एक चार्ट के माध्यम से स्पष्ट किया जा सकता है:



अभारित सूचकांक - इसमें मूल्यों को कोई भार प्रदान नहीं किया जाता और यह मान लिया जाता है कि सभी मदों का भार या सापेक्षिक महत्व समान है। रचना तकनीक के आधार पर अभारित सूच कांक निम्न दो प्रकार से तैयार किये जा सकते हैं।

(a) सरल समूही रीति - सबसे सरल रीति है। आधार वर्ष के मूल्यों के योग एवं चालू वर्ष के मूल्यो के योग कहा जाता है। चालू वर्ष के योग में आधार वर्ष के योग का भाग देकर 100 से गुणा कर देते है।

सूत्र $$\text{Index No}\left(P_{01}\right) = \frac{\sum P_1}{\sum P_0} x100$$

सीमाएँ -

1. विभिन्न वस्तुओं के सापेक्षिक महत्व पर ध्यान नहीं दिया जाता।
2. सूचकांक पर मूल्य के विस्तार का प्रभाव पड़ता है।
3. मूल्य जिस इकाई (लीटर, मीटर आदि) में दिये गये है।
4. उनमें परिवर्तन करके सूचकांक का दुरूपयोग किया जा सकता है।

(b) मूल्य अनुपातों की माध्य विधि - इस विधि में सबसे पहले प्रत्येक वस्तु के मूल्य अनुपात निकाले जाते हैं। इसके लिये प्रत्येक वस्तु के चालू वर्ष के मूल्य में आधार वर्ष के मूल्य का भाग देकर 100 का गुणा $\frac{P_1}{P_0}$ x 100 किया जाता है। मूल्यानुपातो के योगमें वस्तुओं की संख्या का भाग देकर सूचकांक ज्ञात कर लिया जाता है। सूत्र-

$$\left(P_{01}\right) = \frac{\sum\left[\frac{P_1}{P_0} x100\right]}{N} \quad \text{or} \quad \frac{\sum[P.R]}{N}$$

इस रीति के कई लाभ है-

- सूचंकाक पर इसका कोई प्रभाव नहीं पड़ता कि मूल्य किसी इकाई में है क्योंकि वे सब मूल्य अनुपातों में बदल जाते हैं।
- सूचकांक पर मूल्य के निरपेक्ष मान का भी कोई प्रभाव नहीं पड़ता। क्योंकि वे प्रतिशत में परिवर्तित हो जाते हैं।

सीमा - यह अभारित होने के कारण विभिन्न वस्तुओं को अका सापेक्षिक महत्व प्राप्त नहीं हो पाता।

उदाहरण 9 - निम्न संमको से 2002 के मूल्यों के आधार पर 2007 के लिये सूचकांक ज्ञात कीजिए-

वस्तु	A	B	C	D	E
2002 मे मूल्य	12	25	10	5	6
2007 में मूल्य	15	20	12	10	15

हल - सूचकांक की गणना

वस्तु	सरल समूही रीति		मूल्यानुपात रीति			
	2002 (Base Price) (P_0)	2007 (Base Price) (P_1)	2007 (Base Price) (P_1)	Relative (R)	2007 (Current) (P_1))	(R)
A	12	15	12	100	15	125
B	25	20	25	100	20	80
C	10	12	10	100	12	120
D	5	10	5	100	10	200
E	6	15	6	100	15	250
N = 6	$\sum P_0 = 58$	$\sum P_1 = 72$				$\sum R = 775$

सरल समूही रीति द्वारा सूचकांक 2007 or P_{01}

$$= \frac{\sum P_1}{\sum P_0} \text{ x } 100 = \frac{72}{58} \text{ x } 100 = 124.14$$

मूल्यानुपात रीति द्वारा सूचकांक 2007 or P_{01}

$$= \frac{\sum R}{N} = \frac{775}{5} = 155$$

भारित सूचकांक (Weighted Index) -

इससे आशय ऐसे सूचकांको से है जिनकी गणना में विभिन्न वस्तुओं को उनका तुलनात्मक या सापेक्षिक महत्व प्रदान किया जाता है। इसलिये इनकी अधिक तर्कपूर्ण नापा जाता है। ये दो प्रकार के होते हैं-

A. भारित समूही रीति - इस सूचकांक में शामिल सभी वस्तुओं को भार आवंटित किये जाते हैं। इसके निमार्ण की अनेक रीतियां है-

1. लास्पेयर रीति (Laspeyre's Method) - इस रीति में आधार वर्ष की मात्रा (q_0) द्वारा भार प्रदान किये जाते हैं।

अर्थात् - $$P_{01} = \frac{\sum P_1 q_0}{\sum P_0 q_0} x100$$

(जहाँ P_1 = चालू वर्ष मूल्य, q_0= आधार वर्ष का मात्रा, q_0 = आधार वर्ष के मूल्य)

निर्माण विधि -

I. इस रीति का प्रतिपादन लोस्पेयर द्वारा 1871 में किया गया था।

II. इस रीति में आधार वर्ष की मात्राओं को भार माना गया है।

III. चालू वर्ष के मूल्य और आधार वर्ष के भार का गुणा करके उनका योग निकाल लेते है। $\left(\sum p_1 q_0\right)$

IV. आधार वर्ष के मूल्य व आधार वर्ष के भारों के गुणनफल का योग निकालते है। $\left(\sum p_0 q_0\right)$

V. अन्त में $\left(\sum p_1 q_0\right)$ को $\left(\sum p_0 q_0\right)$ से विभाजित करके भागफल को 100 से गुणा कर देते है।

इस रीति में यह मान लिया गया है किस आधार वर्ष में वस्तुओं की जो मात्रा थी, वही चालू वर्ष में रही होगी।

2. पाशे रीति (Paasches's Method)- जर्मन के सांख्यिकीविद् पाशे ने अपनी रीति का प्रतिपादन 1874 में किया। इन्होंने चालू वर्ष की मात्रा को भार माना (q_1)

$$P_{01} = \frac{\sum P_1 q_1}{\sum P_0 q_1} x100$$

लास्पेयर और पाशे के सूत्रों की तुलना के संदर्भ मे यह महत्त्वपूर्ण है कि भार में भिन्नता के कारण समान आंकड़ो के आधार पर भी दोनों सूत्रों से उत्तर में भिन्नता आती है।

मूल्य निर्देशांक में फिशर के आदर्श सूचकांक का स्थान है। यह सूचकांक फिशर ने 134 विभिन्न सूत्रो के गहन अध्ययन के पश्चात विकसित किया था। यह भारित सूचकांक का ही रूप है। इस सूत्र में परिवर्तन भारों का प्रयोग किया जाता है।

3. फिशर आदर्श का सूचकांक (Fisher Ideal Index Number) - यह सूचकांक लास्पेयर तथा पाशे सूचकांको का गुणोत्तर माध्य है। फिशर सूचकांक में ये दोनों अभिनति संतुलित हो जाती है।

अतः

$$P_{01} = \sqrt{\frac{\sum P_1q_0 x \sum P_1q_1}{\sum P_0q_0 x \sum P_0q_1}} x100$$

$$P_{01} = \sqrt{\text{Laspeyre Index } x \text{ Paasche Index}}$$

फिशर सूत्र के आदर्श होने के आधार-

1. यह आधार वर्ष व चालू वर्ष दोनों की ही मात्रा व मूल्य का प्रयोग करता है।
2. यह समय उत्क्राम्यता परीक्षण व तत्व उत्क्राम्यता परीक्षण दोनों को ही संतुष्ट करता है।
3. यह स्थिर व परिवर्तनशील भारों दोनों पर आधारित है।

उदाहरण 10 - निम्न से वर्ष 2005 को आधार मानकर 2006 के लिये लास्पेयर, पाशे व फिशर सूचकांक ज्ञात कीजिए।

Article	A		B		C		D		E	
	P	Q	P	Q	P	Q	P	Q	P	Q
Year 2005	4	9	6	12	5	15	4	10	3	14
Year 2006	6	9	8	8	6	11	5	10	2	7

हल- सूचकांको की गणना

वस्तु	आधार 2005		चालू 2006		भारित समूह			
	P_0	q_0	P_1	q_1	P_1q_0	P_0q_0	P_1q_1	P_0q_1
A	4	9	6	9	54	36	54	36
B	6	12	8	8	96	72	64	48
C	5	15	6	11	90	75	66	55
D	4	10	5	10	50	40	50	40
E	3	14	2	7	28	42	14	21
					$\sum P_1q_0$ =318	$\sum P_0q_0$ =265	$\sum P_1q_1$ =248	$\sum P_0q_1$ =200

लास्पेयर सूचकांक $P_{01} = \frac{\sum P_1q_0}{\sum P_0q_0} x100 = \frac{318}{265} x100 = 120$

पाशे सूचकांक $P_{01} = \frac{\sum P_1q_1}{\sum P_0q_1} x100 = \frac{248}{200} x100 = 124$

फिशर सूचकांक $P_{01} = \sqrt{\dfrac{\sum P_1q_0 x \sum P_1q_1}{\sum P_0q_0 x \sum P_0q_1}} x100 = \sqrt{L x P}$

$P_{01} = \sqrt{120 x 124} = \sqrt{148}$

4. मार्शल एजवर्थ रीति (Marshall Edgeworth Method) - इस रीति में आधार वर्ष और चालू वर्ष दोनों की मात्राओं के औसत का भार दिया जाता है, अर्थात्

$$P_{01} = \frac{\Sigma\ (q_o + q_1)P_1}{\Sigma\ (q_o + q_1)P_0} x100$$

$$P_{01} = \left[\frac{\sum P_1q_0 + \sum P_1q_1}{\sum P_0q_0 + \sum P_0q_1}\right] x100$$

5. डोरविश एवं बाउले रीति - यह रीति लास्पेयर तथा पाशे की रीति का मिश्रण है और यह इन दोनों सूचकांको का समान्तर माध्य होता है।

$$P_{01} = \frac{L + P}{2} = \frac{1}{2}\left[\frac{\sum P_1q_0 +}{\sum P_0q_0 +} + \frac{\sum P_1q_1}{\sum P_0q_1}\right] x100$$

6. कैली रीति - इस सूत्र मे आवश्यकतानुसार आधार वर्ष या चालू वर्ष किसी को भी प्रमाणित मानकर उसकी मात्रा या दोनों की मात्रा के औसत भार दिये जाते हैं। इसलिये सूत्र मे के साथ व या 1 का प्रयोग नहीं किया जाता।

$$P_{01} = \frac{\sum P_1q}{\sum P_0q} x100$$

B. मूल्यानुपातो की भारित माध्य रीति- इस रीति में सूचकांक बनाने के लिये सर्वप्रथम प्रत्येक वस्तु का आधार वर्ष के मूल्य के आधार पर चालू वर्ष के लिये मूल्य अनुपात निकाल लेते है। जिसके लिये $\left[\dfrac{P_1 x 100}{P_0}\right]$ सूत्र का प्रयोग करते हैं।

यदि प्रश्न में भार स्पष्ट रूप से दिया हो तो उसका प्रयोग करते हैं लेकिन यदि आधार वर्ष की मात्रा (q_0) दी हो तो प्रत्येक वस्तु की आधार वर्ष का मात्रा और मूल्य (P_0) का गुणा $(P_0 q_0)$ करके मूल्य भार ज्ञात करते हैं। भार (w) का मूल्य अनुपात (PR) में गुणा करके और उनका योग (P_0) लगाकर $\sum$WPR निकालते है और इसमें भार के योग $\sum$W का भाग दिया जाता है।

सूत्र - $\textbf{Weighted Index No.} = \dfrac{\sum \textbf{WPR}}{\sum \textbf{W}}$

मूल्य भार को w के स्थान पर v से भी दर्शाया जा सकता है।

उदाहरण 11 - एक औसत कर्मचारी वर्ग के परिवार के बजट के समूह सूचकांक और समूह भार है। दिये हुये भारो को प्रदान करते हुये सूचकांको की रचना कीजिये।

क्रस.	समूह	सूचकांक	भार
1	भोजन	350	50
2	ईंधन	240	10
3	वस्त्र	230	10
4	किराया	160	14
5	विविध	180	16

हल-

समूह	सूचकांक (PR)	भार (W)	WPR
भोजन	350	50	17500
ईधन	240	10	2400
वस्त्र	230	10	2300
किराया	160	14	2240
विविध	180	16	2880
		$\sum W = 100$	$\sum WPR = 27,320$

$$\text{Index No} = \frac{\sum WPR}{\sum W} = \frac{27,320}{100} = 273.20$$

मूल्य सूचकांक -

ऊपर दिये गये सभी सूचकांक कीमत सूचकांक तथा मात्रा सूचकांक को वर्णित करते हैं। मूल्य कीमत तथा मात्रा का गुणनफल होता है। अर्थात् मूल्य = कीमत × मात्रा ($v = p \times q$)

मूल्य सूचकांक ज्ञात करने के लिये चालू वर्षो के मूल्यों के योग को आधार वर्ष के मूल्यों के योग $\sum(P_0 q_0)$ से विभाजित करके उसे 100 से गुणा कर दिया जाता है अतः सूत्रानुसार-

$$\text{Value Index No. or V} = \frac{\sum P_1 q_1}{\sum P_0 q_0} x100$$

$$= \frac{V_1}{V_0} x100$$

इनका प्रयोग कम होता है।

सूत्रो की उपयुक्ता के मापदण्ड

एक उपयुक्त सूत्र के चुनाव की कसौटी हेतु कुछ मापदण्ड या परीक्षण सुझाये गये है , जो कि निम्नवत् है। **इकाई मापदण्ड-** इस मापदण्ड के अनुसार मूल्य और मात्राएँ किसी भी इकाई में व्यक्त की जा सकती है , सरल समूहो सूचकांक को छोडकर शेष सभी सूत्र इस मापदण्ड को सन्तुष्ट करते हैं। **समय उत्क्राम्यता परीक्षण (Time Reversal Test)** - इस परीक्षण से यह स्पष्ट है कि आधार वर्ष के आधार पर चालू वर्ष का सूचकांक निकाला जाय और फिर चालू वर्ष के आधार पर आधार वर्ष का सूचकांक ज्ञात किया जाए तो दोनों एक दूसरे के व्युत्क्रम होने चाहिए अर्थात् दोनों का गुणनफल 1 होना चाहिए।

$$P_{01} = \frac{1}{P_{10}} \qquad \text{or } P_{01} \text{ x } P_{10} = 1$$

उदाहरण के लिये यदि 1990 के आधार पर 200 के मूल्य 4 गुने हो जाये तो यदि 2000 को आधार मानकर 1990 का सूचकांक बनाया जाय तो वह एक चौथाई होना चाहिये जिससे $4 \text{ x } \frac{1}{4} = 1$ हो सके।

फिशर का सूत्र इस परीक्षण का पूरा करता है, क्योंकि

$$P_{01} = \sqrt{\frac{\sum P_1q_0 x \sum P_1q_1}{\sum P_0q_0 x \sum P_0q_1}}; P_{10} = \sqrt{\frac{\sum P_0q_0 x \sum P_0q_1}{\sum P_1q_0 x \sum P_1q_1}}$$

$$P_{01} \text{ x } P_{10} = \sqrt{\frac{\sum P_1q_0 x \sum P_1q_1}{\sum P_0q_0 x \sum P_0q_1} \text{ x } \frac{\sum P_0q_0 x \sum P_0q_1}{\sum P_1q_0 x \sum P_1q_1}} = 1$$

तत्व उत्क्राम्यता परीक्षण (Factor Reversal Test) - यह परीक्षण यह स्पष्ट करता है कि मूल्य के स्थान पर मूल्य रखकर सूचकांक (q_{01}) तैयार किया जाय तो इसका और मूल्य सूचकांक (P_{01}) का गुणनफल चालू वर्ष के कुल मूल्य $\sum(P_1q_1)$ और आधार वर्ष के कुल मूल्य $\sum(P_0q_0)$ के अनुपात के बराबर होना चाहिए। अर्थात्

$$P_{01} \text{ x } q_{01} = \frac{\sum P_1q_1}{\sum P_0q_0}$$

चक्रीय परीक्षण (Circular Test) - यह परीक्षण समय उत्क्राम्यता परीक्षण का ही विस्तार है इसके अनुसार यदि 2009 का सूचकांक 1999 के आधार पर बनाया जाये और 1999 का सूचकांक 1989 के आधार पर बनाया जाये तो 1989 के आधार पर प्रत्यक्ष रूप से निकाला गया 2009 का सूचकांक असंगत नहीं होना चाहिए। इसमें सूचकांक चक्र के रूप मे तैयार किये जाते हैं। और उन सब का गुणनफल 1 होना चाहिए।

अतः सूत्रानुसार = $\dfrac{\text{चालू वर्ष का नया सूचकांक नये आधार वष का पुराना सूचकांक}}{100}$

उदाहरण 12- निम्नलिखित आंकड़ो से फिशर का आदर्श सूचकांक की गणना कीजियें। समय उत्क्राम्यता और तत्व उत्क्राम्यता परीक्षणों की जाँच भी कीजिए।

Commodity	2000	2005

Rice	Price (Qty)	Rs. 4 (50 kg)	Rs.10 (40 kg)
Wheat	Price (Qty)	Rs. 3 (10 kg)	Rs.8 (8 kg)
Gram	Price (Qty)	Rs. 2 (5 kg)	Rs. 4 (4 kg)

हलः-

Item	2000		2005		P_1q_0	P_1q_1	P_0q_0	P_0q_1
	P_0	q_0	P_1	q_1				
Rice	4	50	10	40	500	400	200	160
Wheat	3	10	8	8	80	64	30	24
Gram	2	5	4	4	20	16	10	8
					600	**480**	**240**	**192**

Fisher's Ideal Index No- $100\sqrt{\frac{\sum P_1q_0}{\sum P_0q_0} x \frac{\sum P_1q_0}{\sum P_0q_0}} = 100\sqrt{\frac{600}{240} x \frac{480}{192}}$

$$= 100\sqrt{2.5 x 2.5} = 250$$

Time Reversal Test- $P_{01} = \sqrt{\frac{\sum P_1q_0}{\sum P_0q_0} x \frac{\sum P_1q_0}{\sum P_0q_0}} = \sqrt{\frac{600}{240} x \frac{480}{192}}$

$$P_{10} = \sqrt{\frac{\sum P_0q_0}{\sum P_1q_0} x \frac{\sum P_0q_1}{\sum P_1q1}} = \sqrt{\frac{240}{600} x \frac{192}{480}}$$

$$P_{01} x P_{10} = \sqrt{\frac{600}{240} x \frac{480}{192} x \frac{240}{600} x \frac{192}{480}} = 1$$

Factor Reversal Test- परीक्षण के अनुसार $P_{01} x q_{01} = \frac{\sum P_1q_1}{\sum P_0q_0}$ होना चाहिए

$$P_{01} = \sqrt{\frac{\sum P_1 q_0}{\sum P_0 q_0} x \frac{\sum P_1 q_1}{\sum P_0 q_1}} = \sqrt{\frac{600}{240} x \frac{480}{192}}$$

$$Q_{01} = \sqrt{\frac{\sum P_0 q_1}{\sum P_0 q_0} x \frac{\sum P_1 q_1}{\sum P_1 q_0}} = \sqrt{\frac{192}{240} x \frac{480}{600}}$$

$$P_{01} \, x \, Q_{01} = \sqrt{\frac{600}{240} x \frac{480}{192} x \frac{192}{240} x \frac{480}{600}} = 1$$

अर्थात् $\frac{\sum P_1 q_1}{\sum P_0 q_0}$

इस प्रकार यह सूत्र तत्त्व उत्क्राम्यता परीक्षण पर सही सिद्ध होता है।

8.8 सूचकांको की सीमाएँ (Limitations of Index)

यह माना गया है कि परिवर्तनो की तुलनात्मक या सापेक्ष मापन की दृष्टि से सूचकंाक एक महत्वपूर्ण सांख्यिकीय उपकरण है लेकिन व्यवहार में इसकी कुछ सीमाएँ है। यह परिसीमाएँ इस प्रकार से है।

1. **न्यादर्श पर आधारित** - सूचकांक की गणना में प्रत्येक मद को शामिल करना अत्यन्त कठिन कार्य है। यदि न्यादर्श में शामिल की गई मदे समग्र का उचित प्रतिनिधित्व नहीं करती , तो सही स्थिति प्रकट नहीं हो पायेगी।
2. **औसत का संकेत** - सूचकांक द्वारा परिवर्तन से औसत का ही संकेत मिलता है। इसी के आधार पर इनका परिवर्तन किया जाना चाहिए।
3. **रचना सम्बन्धी सीमाएँ** - सूचकांको की रचना में असावधानियों का या भ्रम उत्पन्न हो सकता है। जैसे आधार वर्ष का चुनाव, भार का निर्धारण, औसत का प्रयोग आदि।
4. **विशिष्ट उद्देश्यो का प्रभाव** - किसी एक उद्देश्य से बनाया गया सूचकांक का प्रयोग दूसरे उद्देश्य के लिये नहीं हो सकता।
5. **गुणात्मक तथ्यो के परिवर्तन की उपेक्षा** - सूचकांक के माध्यम से संख्यात्मक परिवर्तन का मापन सरलता से हो जाता है, लेकिन यदि सम्बन्धित तथ्यो में गुणात्मक परिवर्तन भी हुआ हो तो उसका सही प्रकटीकरण नहीं हो पाता।

8.9 सारांश (Summary)

आर्थिक क्षेत्र में निरन्तर परिवर्तित होते रहते हैं। इन्हीं परिवर्तनों का अध्ययन करने और इनके प्रभावों को स्पष्ट करने के लिए जिस सांख्यिकीय तकनीक को विकसित किया गया है उसी तकनीक को सूचकांक अथवा निर्देशांक कहते हैं। प्रारम्भ में सूचकांक को केवल मूल्यस्तर तथा मुद्रा की क्रयशक्ति का माप करने हेतु प्रयोग किया जाता था, परन्तु आज के समय में इसका प्रयोग विस्तृत हो गया है। सूचकांको के विकास में प्रो. जेवन्स , डॉ. मार्शल, एजवर्थ फिशर का नाम उल्लेखनीय है। सूचकांक एक विशेष प्रकार का माध्य है, जिनके द्वारा समय, स्थान या अन्य किसी विशेषता के आधार पर सम्बन्धित चर मूल्यों में होने वाले सापेक्ष परिवर्तनों का मापन किया जाता है।

सूचकांक को कीमत या मूल्य सूचकांक , मात्रा सूचकांक, कुल मूल्य सूचकांक या वैल्यू सूचकांक , उद्देश्य विशेष सूचकांक वस्तुओं की संख्या के आधार पर वर्गीकृत किया जाता है। मूल्य सूचकांक एक प्रमाप वर्ष के आधार पर प्रचलित वर्ष के मूल्य स्तर को व्यक्त करते हैं। आधार वर्ष ज्ञात करने की दो रीतियां हैं- स्थिर आधार रीति एवं श्रृंखला आधार। आधार परिवर्तन दो प्रकार के होते हैं - स्थिर आधार से श्रृखंला आधार में एवं श्रृखंला आधार से स्थिर आधार में। आधार वर्ष परिवर्तन आधार परिवर्तन से भिन्न होता है। आधार वर्ष परिवर्तन का आशय है- एक सूचकांक के दिये हुये (पुराने) आधार वर्ष को बदलकर उसके स्थान पर किसी नये आधार वर्ष पर आधारित करके एक नई सूचकांक श्रृखंला की पुर्नरचना करना। आधार वर्ष परिवर्तन की दो रीतियाँ है- प्रत्यक्ष या पुर्ननिर्माण रीति एवं अप्रत्यक्ष अथवा परोक्ष या संक्षिप्त रीति। सूचकांक रचना में किस माध्य का प्रयोग किया जाये यह तय करना जरूरी होता है। व्यवहार में माध्यका , समान्तर माध्य या गुणोत्तर माध्य में से किसी एक ही का प्रयोग करना चाहिए। जब विभिन्न वस्तुओं से सम्बंधित भारों को ध्यान में रखकर सूचकांक बनाया जाता है। तो उसे भारित सूचकांक कहते हैं। भार देने की दो रीतियाँ है प्रत्यक्ष तथा परोक्ष भारांकन स्थिर तथा परिवर्तनशील भार। सूचकांक निर्माण की दो विधियाँ है- अभारित एवं भारित। भारित समूही रीति में शामिल सभी वस्तुओं को भार आवंटित किये जाते हैं। इसके निमार्ण की अनेक रीतियां है - लास्पेयर रीति , पाशे रीति, फिशर आदर्श का सूचकांक, मार्शल एजवर्थ रीति, डोरविश एवं बाउले रीति व कैली रीति। एक उपयुक्त सूत्र के चुनाव की कसौटी हेतु कुछ मापदण्ड या परीक्षण सुक्षाये गये है। इकाई मापदण्ड , समय उत्क्राम्यता परीक्षण , तत्व उत्क्राम्यता परीक्षण, चक्रीय परीक्षण। शिरोबन्धन का अर्थ दो सूचकांको मालाओं के शिरो को बांधने से है अर्थात् शिरो बन्धन का अर्थ दो या अधिक अधिव्याप्त सूचकांको की मालाओं को किसी एक सामान्य आधार पर एक सूचकांक माला में परिवर्तित करने से है। उपभोक्ता मूल्य सूचकांक या निर्वाह लागत सूचकांक किसी स्थान विशेष पर वर्ग विशेष के व्यक्ति के निर्वाह व्यय में होने वाले परिवर्तनो की दशा और मात्रा को प्रकट करते हैं। सूचकांक एक तुलनात्मक अथवा सापेक्ष माप है।वास्तव में ऐसा कोई क्षेत्र नहीं है जिसमें संख्यात्मक को मापने के लिए सूचकांकों का प्रयोग न होता हो। तुलनात्मक अध्ययन को सम्भव बनाना , भावी प्रवृत्तियों के संकेतक ,आर्थिक नीतियों के निर्माण में सहायक , जटिल तथ्यों को सरल बनाना , विभिन्न मूल्यों की अवस्फीति में सहायक सूचकांक की उपयोगिता को दर्शाता है। निष्कर्ष रूप में कहा जा सकता है कि सूचकांक प्रतिशत के रूप में व्यक्त किया जाने वाला एक विशेष प्रकार का माध्य है जिसके आधार पर विभिन्न समयों , स्थानों या अन्य समंक समूहों में होने वाले सापेक्षिक परिवर्तनों की सामान्य प्रकृति को मापा जाता है।

8.10 शब्दावली (Glossary)

- तुलनात्मक अध्ययन (comparative study)
- सापेक्ष मापन (relative measurement)
- प्रतिनिधि मूल्य (Representative Price)
- स्थिर आधार रीति (Fixed base method) - जब आधार मूल्य स्थिर रहते है।
- बहुवर्षीय माध्य आधार (Multi-year mean basis) - कुछ वर्षो के माध्य को आधार मान लेते है।
- चल आधार रीति (Moving base method) - चालू वर्ष के लिये पिछला वर्ष आधार वर्ष मान लेते है
- शिरोबन्ध (Headbandh) - दो या अधिक सूचकांक मालाओ को किसी एक सूचकांक माला में परिवर्तित करने से है।
- मूल्य अनुपात Price Relative (P.R.)
- उद्धरण (Quotation)
- भारांकन (Weighting)
- सजातीय (Homogenous)

8.11 लघु उत्तरीय प्रश्न (Short Answer Questions)

1. फिशर के आदर्श सूचकांक के निर्धारण में सामान्यता कितने खाने होते हैं।

2. पाशे का सूचकांक आधारित है।

(1) आधार वर्ष की मात्रा पर (2) चालू वर्ष की मात्रा पर

(3) दोनों के औसत पर (4) इनमे से कोई नहीं।

3. एक अध्ययन सूचकांक वह है जो संतुष्ट करता है।

(1) इकाई परीक्षण (2) समय उत्क्राम्यता परीक्षण

(3) तत्व उत्क्राम्यता परीक्षण (4) चक्रीय परीक्षण

4. निम्न में से आदर्श सूचकांक है

(1) पाशे का सूत्र (2) फिशर का सूत्र

(3) लास्पेयर का सूत्र (4) वाश का सूत्र

5. सूचकांको की रचना के लिये सर्वोत्तम माध्य है।

(1) मध्यक (2) माध्यिका

(3) बहुलक (4) गुणोत्तर माध्य

6. सूचकांक होते हैं आर्थिक

(1) लेक्टोमीटर (2) स्पाइरामीटर

(3) बैरो मीटर (4) कैलोरीमीटर

उत्तर- (1) 1, (2) 2, (3) 3, (4) 2, (5) 4, (6) 3

8.12 सदंर्भ सहित ग्रन्थ (Books with References)

- डा0 एस सचदेवा - परिमाणात्मक विधियाँ ,लक्ष्मी नारायण अग्रवाल, आगरा।
- डा0 के0 एल0 गुप्ता एवं डा0 हरिओम गुप्ता - परिमाणात्मक तकनीकें, नवयुग साहित्य भवन, आगरा।
- डा0 के 0 एल 0 गुप्ता , रवि कान्त - अर्थशास्त्र की आधारभूत परिमाणात्मक विधियाँ , नवनीत पब्लिकेशन्स, आगरा।
- एस0पी0 सिंह - साख्यिकीः सिद्धान्त एवं व्यवहार, एस0 चन्द पब्लिकेशन्स, नई दिल्ली।

8.13 कुछ उपयोगी पुस्तकें (Some Useful Books)

- Kumar, Anil, (2008) Statistical Research Methodology, Aifa Publishing House.
- Singh, S.P.(2010) Principles of Statistics, S & Chand Publishing House.
- Bhardwaj,R.S. (2000). Mathematics for Economics and Business, Excel Books.
- Bose,D.,(2003), An introduction to Mathematical Economics, Himalaya Publishing House.

8.14 निबन्धात्मक प्रश्न (Essay Type Questions)

1. सूचकांक क्या है? इसका निर्माण कैसे किया जाता है ? फिशर का सूत्र आदर्श सूचकांक क्यो कहलाता है?
2. सूचकांक की परिभाषा दीजिये ? सूचकांक बनाने की स्थिर आधार विधि व श्रृखंला आधार विधि में अंतर स्पष्ट कीजिये व उनके तुलनात्मक गुणो का वर्णन कीजिये।
3. वर्ष 2004 को आधार मानकर नये सूचकांक ज्ञात कीजिये।

वर्ष	2001	2002	2003	2004	2005	2006
सूचकांक	100	108	120	150	210	225

4. निम्न में से फिशर आदर्श मूल्य सूचकांक ज्ञात कीजिये-

वस्तुएँ	**आधार वर्ष**		**प्रचलित वर्ष**	
	मूल्यइकाई/	**कुल व्यय**	**मूल्यइकाई/**	**कुल व्यय**
A	2	40	5	75
B	4	16	8	40
C	1	10	2	24
D	5	25	10	60

इकाई - 9 काल श्रेणी विश्लेषण
(Time Series Analysis)

9.1 प्रस्तावना (Introduction)

प्रस्तुत इकाई में काल श्रेणी विश्लेषण पर प्रकाश डाला गया हैं। आधुनिक एवं व्यावसयिक क्षेत्रो में समय के साथ-साथ निरन्तर रीति से अनेक प्रकार के परिवर्तन होते रहते हैं। किसी भी चल के मूल्यों में समय में परिवर्तन के साथ साथ परिवर्तन होते रहते हैं। इस परिवर्तन के कई कारण हो सकते हैं जैसे- जनसँख्या में परिवर्तन, मूल्य में परिवर्तन, उपभोक्ताओं की रुचियों में परिवर्तन, उनकी आयों में परिवर्तन, आर्थिक स्थिति में परिवर्तन, आदि। काल की गति के साथ मूल्यों मे होने वाले विभिन्न दीर्घकाल एवं अल्पकालीन उच्चवचनों का विधिवत् विश्लेषण किसान, उपभोक्ता, व्यापारी, प्रशासक आदि सभी वर्गो के व्यक्तियों के लिये आवश्यक और उपयोगी होता है।

यदि किसी वस्तु के विक्रय सम्बन्धी पिछले 50 वर्षों के समंकों का अध्ययन किया जाए तो यह मालूम होगा कि वार्षिक विक्रय में प्रत्येक वर्ष परिवर्तन होता रहता है । यदि हम श्रेणी के अद्ध्ययन के आधार पर अगले वर्षों में होने वाले विक्रय का पूर्वानुमान करना चाहते हैं तो हमें विक्रय को प्रभावित करने वाले सभी तत्वों के प्रभावों का अलग अलग अध्यन करना पडेगा।

अतः काल श्रेणी के विश्लेषण में हम किसी चल के मूल्यों को प्रभावित करने वालें विभिन्न तत्वों के प्रभावों का अध्ययन करते हैं। साथ ही काल श्रेणी के महत्त्व का भी अवलोकन करेंगे और भविष्य के बारे में पूर्वानुमान एवं पिछले अनुभवों से होने वाले लाभ में काल श्रेणी की भूमिका का भी अध्यन करेंगे।

9.2 उद्देश्य (Objectives)

प्रस्तुत इकाई का अध्ययन करने के बाद आप-

- ✓ काल श्रेणी विश्लेषण को समझेंगे।
- ✓ काल श्रेणी के संघटक को जानेंगे।
- ✓ दीर्घकालीन प्रवृत्ति एवं अल्पकालीन उच्चावचन के बीच अंतर को जानेंगे।
- ✓ अल्पकालीन एवं दीर्घकालीन उच्चावचनों के मापन की विधियों को जानेंगे।
- ✓ काल श्रेणी विश्लेषण के महत्त्व को समझेंगे।

9.3 काल श्रेणी विश्लेषण (Time Series Analysis)

एडवर्ड-डे लेविस के अनुसार,"अर्थशास्त्री के लिए यह जानने के प्रयासों मे कि आर्थिक व्यवस्था कैसे कार्य करता है, काल श्रेणी का अध्ध्यन सम्भवना, सूचना का सबसे महत्वपूर्ण स्रोत है।"काल की गति के साथ मूल्यों में होने वाला दीर्घकालीन एवं अल्पकालीन उच्चावनों का विश्लेषण किसान उपभोक्ता , व्यापारी, प्रशासक आदि सभी वर्गों के व्यक्तियों के लिये आवश्यक और उपयोगी होता है।

9.3.1 काल श्रेणी का अर्थ एवं परिभाषा (Meaning and Definition of Time Series)

काल श्रेणी का आशय ऐसी श्रेणी या समंकमाला से है, जिसमें 'काल' अर्थात् 'समय' के आधार पर संमक प्रस्तुत किय जाते हैं।

बर्नर हर्श के अनुसार, समय के क्रमिक बिन्दुओं के तत्संवादी उसी चर के मूल्यों का व्यवस्थित अनुक्रम ही काल श्रेणी कहलाता है। या **लुन चाऊ** के अनुसार , एक काल श्रेणी को विभिन्न समय अवधियों में किसी आर्थिक चर या चरों के मिश्रण से सम्बन्धित संख्याओं के संकलन के रूप में परिभाषित किया जा सकता है।

"A time series may be defines as a collection of reading belonging to different time periods, of economic variable or composite of variables. – **Prof. Ya-Lun-Chou**

तकनीकी दृष्टि से काल श्रेणी विश्लेषण में स्वतन्त्र चर मूल्य एवं समंक आश्रित चर होते हैं। यह संमक समय के साथ-साथ होने वाले परिवर्तनों को स्पष्ट करते हैं।

उदाहरण -

भारत में जनसंख्या-

वर्ष	जनसंख्या (करोड़)
1957	36.20
1961	43.90
1971	54.00
1981	68.40

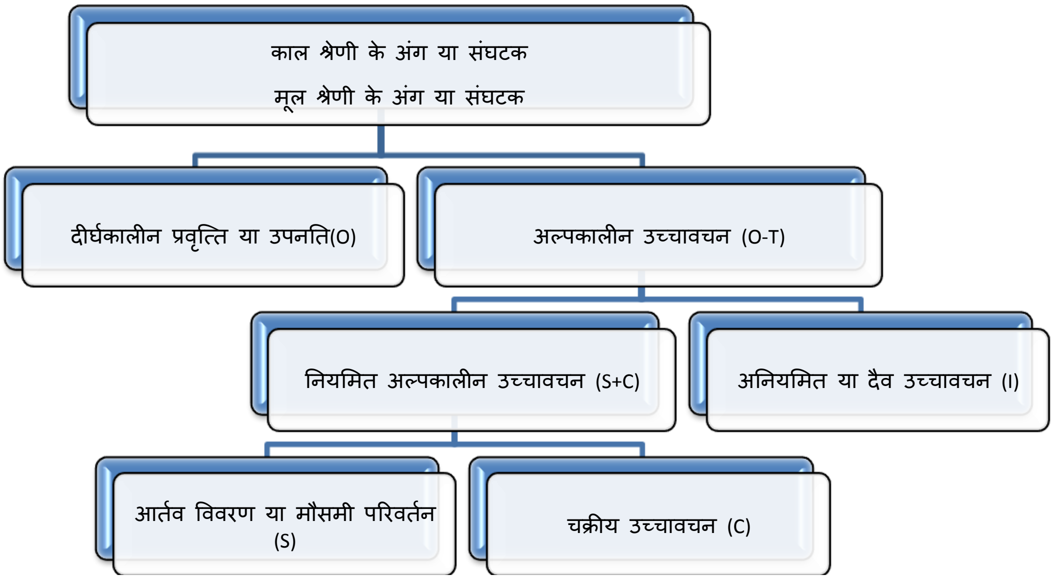
निष्कर्ष रूप में यह कहा जा सकता है कि काल श्रेणी का आशय समय क्रम में सांख्यिकीय संमको की व्यवस्था से है। यह श्रेणी समय परिवर्तन के साथ ही तथ्य विशेष के संमकों में होने वाले परिवर्तनों को स्पष्ट करती है। काल श्रेणी में होने वाले दीर्घकालीन एवं अल्पकालीन उच्चावचनों का अध्ययन न सिर्फ व्यापारी वरन् अर्थशास्त्री के लिए भी बड़ा महत्व रखता है। भूतकाल के परिवर्तन के विश्लेषण करके वे पिछले अनुभव के आधार पर भविष्य की नीतियाँ निर्धारित कर सकते हैं। और अपनी क्रिया ओं पर नियंत्रण करके भविष्य के जोखिमों से अपने व्यापार की सुरक्षा कर सकते हैं। अतः यह कह सकते हैं कि विभिन्न वर्ग चाहे वो अर्थशास्त्री हो या उपभोक्ता, योजनाकार, किसान, राजनीतिक आदि सभी के लिये काल श्रेणी में से वाले परिवर्तनों का विश्लेषण विशेष रूप से उपयोगी होता है। एक विवेकपूर्ण विश्लेषण तथा संकेतको का वैज्ञानिक विवेचन काल श्रेणी की महत्ता में वृद्धि करता है।

9.4 काल श्रेणी के अंग या संघटक (Components of Time Series)

अनेक प्रकार के घटक या परिवर्तन काल श्रेणी पर अपना प्रभाव डालते है। इन परिवर्तनों को कुछ वर्गों में बाँट सकते हैं और वर्ग ही काल श्रेणी के संघटक कहे जाते हैं।

मूल संमकों को 'O' से दर्शाया जाता है, इसके चार संघटक है।

a. दीर्घकालीन प्रवृत्ति (T)
b. मौसमी विचरण (S)
c. चक्रीय उच्चारण (C)
d. अनियमित उच्चावचन (I)



अनेक प्रकार के घटक या परिवर्तन काल श्रेणी पर अपना प्रभाव डालते है। इन परिवर्तनों को कुछ वर्गों में बाँट सकते हैं और वर्ग ही काल श्रेणी के संघटक कहे जाते हैं।

9.4.1 दीर्घकालीन प्रवृति या उपनति (Secular Trend or long term movement or trend)

इसे 'T' से सम्ब®धित किया जाता है।जब दीर्घकाल में परिर्वतन की सामान्य दिशा का अध्ययन होता है तो उस प्रवृत्ति को दीर्घकालीन प्रवृति कहते हैं। इस प्रकार Secular Trend किसी समंक में दीर्घकाल में घटने अथवा बढ़ने की प्रवृत्ति की ओर संकेत करती है न कि उनमे होने वाले अल्पकालीन उच्चावचनों की ओर।

प्रो0 सिम्पसन और काफ्का के अनुसार,"उपनति जिसे दीर्घकालीन प्रवृत्ति भी कहते हैं, किसी समयावधि मे बढ़ने या घटने की आधारभूत प्रवृत्ति होती है। उपनति की धारणा में अप्लकालीन परिवर्तन शामिल नहीं होते, वरन् दीर्घकालीन में हुये स्थिर परिवर्तन शामिल होते हैं।"

सरल शब्दों में यह कह सकते हैं कि अल्पकाल में समय-समय पर कई उतार चढ़ाव होते हैं पर दीर्घकाल में इन्हीं उतार-चढ़ाव में एक अर्न्तविहीन प्रवृत्ति देखने को मिलती है। इसी प्रवृत्ति को दीर्घकालीन प्रवृत्ति कहते हैं। उदाहरण के तौर पर यदि देश में मूल्यों की बात करें तो हम पायेंगे कि उनमें समय-समय पर कई उतार-चढ़ाव हुये परन्तु दीर्घकाल में उनकी प्रवृत्ति बढ़ने की ही है। उत्पादन , उपभोग, विक्रय, निर्यात, विनियोग, कीमत, रोज़गार आदि के काल श्रेणियों में दीर्घकाल में बढ़ने अथवा घटने की प्रवृत्ति प्रतिलक्षित होती है। यह प्रवृत्ति जनसँख्या में वृद्धि , तकनीकी परिवर्तन, व्यावसायिक संगठन और विधि में परिवर्तन , प्राकृतिक साधनों की खोज, सरकारी नीति आदि ऐसे अनेक महत्वपूर्ण तथ्यों के प्रभावों के परिणामस्वरूप उत्पन्न होती है जिनमे परिवर्तन धीरे धीरे होता है। ऐसे तत्वों का आर्थिक समंको पर प्रभाव भी नियमित एवं मंद होता है।

"Trend, also called Secular trend or long term trend, is the basic tendency (of series) to grow or decline over a long period of time. The concept of trend does not include short term oscillations but rather study movements over a long period of time." - ***Simpson and Kafka***

दीर्घकालीन प्रवृत्ति को मापने के उद्देश्य

इस प्रवृत्ति को मापने के दो प्रमुख उद्देश्य है।

(1) अन्य संघटको की जानकारी। जैसे- अल्पकालीन उच्चावचन, मौसमी विचरण, चक्रीय परिवर्तन आदि।

(2) भविष्य का अनुमान।

प्रमुख विशेषताएँ-

(1) तीन पहलू - (i) **वृद्धि प्रवृत्ति** - देश में मूल्यों की स्थिति

(ii) **कमी प्रवृत्ति** - जनसंख्या की मृत्यु दर

(iii) **स्थिर प्रवृत्ति** - स्थान विशेष का तापमान

(2) विभिन्न समयों में विभिन्न प्रवृत्ति - यह भी मुमकिन है कि एक दीर्घकालीन प्रवृत्ति के अन्दर एक समय में एक प्रवृति, दूसरे समय मे दूसरों प्रवृत्ति देखने को मिले।

(3) तुलनात्मक धारणा - क्योंकि दीर्घकाल एक तुलनात्मक धारण है अतः यह श्रेणी विशेष की विशेषताओं से प्रभावित होता है। मृतकों की संख्या किसी विशेष परिस्थिति के दौरान एक माह से कुछ माह में ही दीर्घकालीन अवधि के अन्तर्गत आ सकती है जबकि मूल्यों के उतार-चढ़ाव कई सालों की अवधि को इस श्रेणी में लाया जाता है।

9.4.2 नियमित अल्पकालीन उच्चावचन (Regular Short Term Oscillation)

कई बार देखा गया है कि कुछ ऐसी शक्तियों काल श्रेणी को प्रभावित करती है, जिसकी समय-समय पर पुनरावृत्ति होती है। क्योंकि यह पुनरावृति नियमित रूप से होती है। अतः उन्हें नियमित अप्लकालीन उच्चावचन कहा जाता है। विभिन्न रीतियों से निकाले गए प्रवृत्ति मूल्यों (trend values) को काल श्रेणी के मूल समंकों में से यदि घटा दिया जाए, तो जो शेष बचता है उसे अल्पकालीन उच्चावचन (Short Term Oscillation) कहते हैं। इन्हें दो भागों में बांट सकते हैं।

(A) आर्त्तव विचरण या मौसमी परिवर्तन (Seasonal Variations) -जो परिवर्तन एक वर्ष से कम की अवधि में ही नियमितता और लगभग एकरूप प्रवृत्ति के रूप में होते रहते है , उन्हें मौसमी परिवर्तन कहते हैं-जैसे दिन, सप्ताह, माह, छमाहि आदि। इसके मुख्य कारण है।

a) जलवायु
b) रीतिरिवाज, परम्परा और स्वभाव
c) समय विशेष की परिस्थिति- जैसे अप्रैल आते ही स्कूल की ड्रेस और किताबों की मांगो में उछाल।

आर्थिक समंकों के दृष्टिकोण से प्रतिवर्ष घटित होने वाली periodic movement अधिक महत्वपूर्ण होती हैं। यद्यपि मौसमी परिवर्तन का विस्तार परिवर्तनीय हो सकता है परन्तु उनकी अवधि सर्वदा निश्चित होती है। यही कारण है कि वार्षिक समंकों में मौसमी परिवर्तन दृष्टिगोचर नही होते।

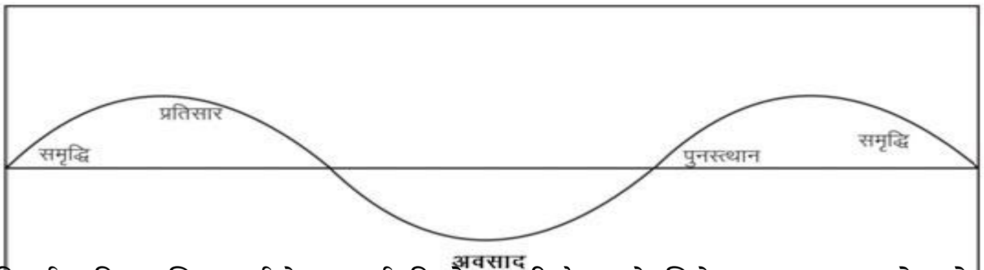
मौसमी परिवर्तन की विशेषताएँ-

(i) नियमित परिवर्तन
(ii) दोनों दशाओं में परिवर्तन-अर्थात् उतार भी और चढ़ाव भी।
(iii) पूर्वानुमान सम्भव-उपभोक्ता, उत्पादक विक्रेता आदि अपने निर्णय लेते समय परिवर्तनों का विशेष ध्यान रखते है और भविष्य पुर्वानुमान के लिए कर पाते है।

(B) चक्रीय उच्चावचन (Cyclical fluctuations) - यह प्रायः देखा गया है कि आर्थिक काल श्रेणियों में तरंग सरीखे (wave like) उतार चढ़ाव दिखलाई पड़ते हैं। इस प्रकार के परिवर्तन के कारण उत्पन्न होते हैं। यह उचावचन भी नियमित होने वाले होते हैं किन्तु इनकी पुनरावृत्ति एक वर्ष से अधिक की होती है। इन्हें चक्रीय

इसलिये कहा जाता है क्योंकि इनका क्रम चक्रीय स्वभाव का होता है। जैसे व्यापार का चक्र जिसमें सामान्यतः चार अवस्थायें देखने को मिलता है-

(i) समृद्धि (Prosperity)
(ii) प्रतिसार (Recession)
(iii) अवसाद (Depression)
(iv) पुनस्त्थान (Recovery)



यहाँ परिवर्तन की अवधि 3 वर्ष से 10 वर्ष की हो सकती है। इसके लिये 'C' शब्द का प्रयोग होता है।

चक्रीय उच्चावचन मौसमी परिवर्तन से भिन्न होते हैं। इनकी अवधि नियमित नहीं होती। फिर भी cyclical variation का परिवर्तन क्रम (sequence of change) लगभग नियमित होता है। इनकी अवधि मौसमी विचरण की अवधि से अधिक होती है।

9.4.3 अनियमित या दैव उच्चारण (Irregular or Random Variation)

जब अकस्मात कोई घटना या परिस्थिति से उच्चावचन होते हैं तो उन्हें अनियमित उच्चावचन कहा जाता है। जैसे किसी कारखाने में आग लगना, जिसके कारण लाभ कम हो जाना, हड़तालों के कारण उत्पादन प्रभावित होना आदि। इन परिवर्तनों की दिशा या मात्रा के बारें में अनुमान लगाना असंभव होता है। उनका अध्ययन करने के लिए पहले दीर्घकालीन प्रवृत्ति और फिर अल्पकालीन उच्चावचन के प्रभावों को दूर करना आवश्यक है। परन्तु ऐसा करने पर भी अनियमित होने के कारण इनका वैज्ञानिक अध्ययन संभव नहीं। अनेक ऐसे तत्त्व है जहाँ पर अनियमित उच्चावचन देखने को मिलता है। जैसे युद्ध , बाढ़, भोचाल, हड़ताल, राजनीतिक परिवर्तन आदि जो कि अवधि एवं तीव्रता की दृष्टि से अनियमित होते हैं। ये न किसी काल क्रम से प्रभावित होते हैं न किसी अन्य प्रकार से नियमित होते हैं।

विशेषताएँ-

(i) पूर्वानुमान नहीं-क्योंकि यहाँ जो शक्तियों क्रियाशील होता है उनके बारे में पहले से जानकारी मिलना सम्भव नहीं होता। अतः इसका पूर्वानुमान भी संभव नहीं हो पाता।
(ii) निश्चित प्रारूप न होना-नहीं ऐसे उच्चावचना का कोई निश्चित प्रारूप होता है न ही इनके पुनः होने की निश्चित अवधि होती है।
(iii) अल्पकालिक-ऐसे उच्चावचन प्रायः अल्पकालिक होते हैं पर इनका प्रभाव कभी- कभी अत्यन्त गहरे होते हैं।
(iv) अनियमित परिवर्तन-इसके अन्तर्गत उन सभी परिवर्तना को शामिल किया जाता है जो न दीर्घकालिक प्रवृति और न मौसमी परिवर्तन की श्रेणी में आते है।

9.5 काल श्रेणी का विश्लेषण-आशय एवं मॉडल (Analysis of Time Series – Purpose and Models)

काल श्रेणी निदर्श-काल श्रेणी के चार संघटकों का मापन निम्न दो मॉडलों पर आधारित है।

9.5.1 योज्य मॉडल (Additive Model)

यह इस मान्यता पर आधारित है कि मूल समंक चारों संघटक अंगो का योग होता है।

$$O = T + S + C + I$$

दीर्घकालीन उतपत्ति ;ज्द्ध को मूल संमक में से घटाकर अल्पकालीन उच्चावचनों का पृथक्क्करण किया जाता है।

$$O - T - S - C = I$$

अल्पकालीन उच्चावचनों (O-T) में से मौसमी विचरणों (S) को घटाकर चक्रीय व अनियमित परिवर्तन ज्ञात किया जा सकता है।

$$O - T - S = C + I$$

यदि अल्पकालीन उच्चावचनो (O-T) में से मौसमी और चक्रीय उच्चावचनों (S+C) को घटाकर अनियमित परिवर्तन ज्ञात किया जा सकता है।

$O - T - (S + C) = I$

$$O - T - S - C = I$$

9.5.2 गुणात्मक मॉडल (Multiplicative Model)

इस में मूल संमक चारों सघंटकों को गुणनफल होता है।

$$O = T \times S \times C \times I$$

अल्पकालीन विचरण को मापने के लिये, इन्हें अलग-अलग ढंग से प्रयोग किया जा सकता है।

$$\frac{O}{T} = S \text{ x } C \text{ x } I$$

$$\frac{O}{T \text{ x } C} = S \text{ x } I$$

$$\frac{O}{T \text{ x } S \text{ x } C} = I$$

इस मॉडल में दीर्घकालीन प्रवृत्ति को मूल संमको की इकाई के रूप में व्यक्त किया जाता है।
व्यवहार में योज्य एवं गुणात्मक दोनों मॉडलों का मिश्रण भी अपनाये जा सकते हैं।

$$O = TSC + I$$

$$O = TC + SI$$

$$O = T + SCI$$

$$O = T + S + CI$$

9.6 दीर्घकालीन प्रवृत्ति का मापन (Measurement of long term trend)

इस प्रवृति को मापने के लिये चार प्रमुख रीतियाँ निम्न प्रकार है -

1. मुक्त हस्त रीति
2. अर्द्ध मध्यक रीति
3. चल माध्य रीति
4. न्यूनतम वर्ग रीति

1. मुक्त हस्त वक्र रीति (Free hand curve method)

इस रीति में मूल काल श्रेणी को बिन्दु रेखीय पत्र पर अंकित करके एक चित्र बनाया जाता है। तथा उसके पश्चात् आकड़ो के उतार-चढाव को ध्यान में रखके उच्चावचनों के लगभग गुजरता हुआ एक सरलित वक्र खीचां जाता है। यही वक्र मुक्त हस्त रीति द्वारा दीर्घकालिन प्रवृत्ति या उपनति का प्रदर्शित करता है।
इसे बिन्दुरेखीय रीति भी कहते हैं अथवा निरीक्षण द्वारा वक्र अन्वायोजन की रीति भी कहते हैं। यह एक सरलतम रीति है , क्योंकि इसमें जटिल गणितीय क्रिया ओं का प्रयोग नहीं होता है। परन्तु इस रीति के दोष निम्नलिखित है -

(1) विषयगत रीति - सरलित वक्र खीचने में व्यक्ति के पक्षपात और पूर्वाग्रहो का प्रभाव पड़ सकता है।

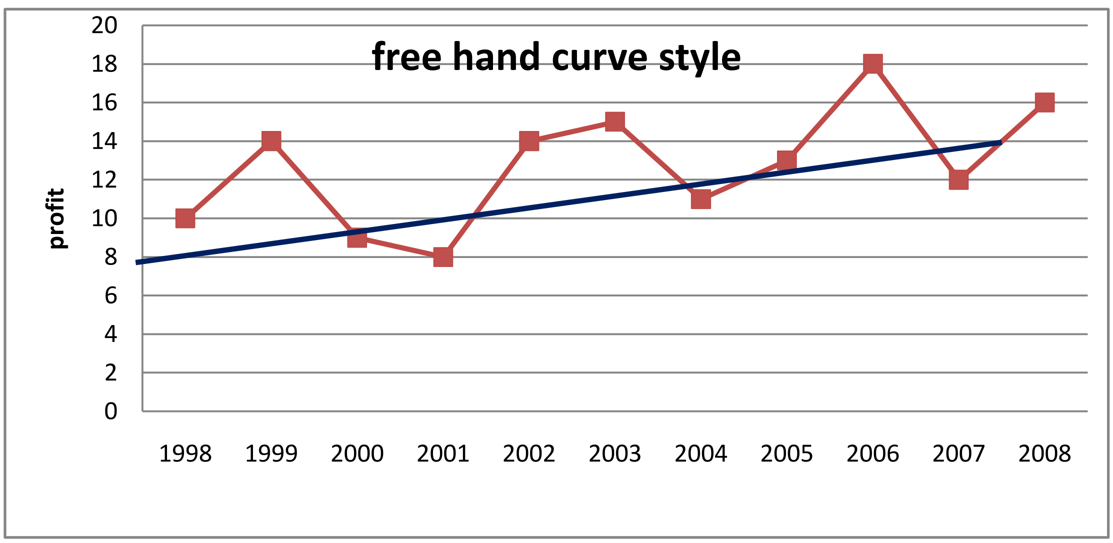
(2) इस रीति में परिशुद्धता का अभाव होता है।

(3) पूर्वानुमान में खतरा

उदाहरण -

वर्ष	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
लाभ	10	14	9	8	14	15	11	13	18	12

हल - वर्षों की संख्या के आधारपर एक बिन्दुरेखीय ग्राफ अंकित किया जाता है तथा मुक्त हस्त द्वारा एक सीधी रेखा अंकित की जाती है।



2. अर्द्ध मध्यक रीति (Semi Average Method)

इस रीति का अर्थ है- श्रेणी के प्रत्येक आधे भाग (पूर्वाद्व तथा उत्तरार्द्व) के मूल्यों का समान्तर माध्य इस रीति के द्वारा दीर्घकालीन प्रवृत्ति ज्ञात करने की प्रक्रिया निम्न प्रकार से है-

(1) काल श्रेणी का दो समान भागों में विभाजन- ऐसा करने के पश्चात् प्रत्येक भाग का माध्य निकालकर उस भाग के मध्य के समय बिन्दु के सामने रखा जाता है।

(2) दो माध्यों की गणना-दोनों समान भागों का अलग- अलग समान्तर माध्य ज्ञात कर लेते है। इस माध्यों को ही अर्द्धमध्यक कहते हैं।

(3) ग्राफ पेपर पर मूल बिन्दुओं का अंकन- पहले अर्द्धमध्यक का बिन्दु पहले भाग के समय के माध्यका बिन्दु के ऊपर और दूसरे अर्द्धमध्यक का बिन्दु के ऊपर लगाया जाता है।

(4) प्रवृत्ति सेवा- उपलब्ध सरल रेखा ही अर्द्धमध्यक रीति द्वारा प्राप्त प्रवृत्ति रेखा है।

यदि मूल्यों की संख्या विषम हो तो बिल्कुल बीच के संमक को छोड़ दिया जाता है शेष क्रिया पूर्वरत रहता है।

उदाहरण 2- निम्न संमकों से अर्द्धमध्यक रीति का प्रयोग करते हुए दीर्घकालीन प्रवृत्ति निर्धारित कीजिए तथा 2002 के मूल्य का अनुमान कीजिये-

वर्ष	1995	1996	1997	1998	1999	2000
उत्पादन	40	48	44	60	56	64

हल - यहाँ कुछ 6 वर्षो के मूल्य दिये गये है इन्हें दो बराबर के भाग 3-3 वर्षों के होंगे और उनके माध्यम निकालकर बिन्दुओं पर दीर्घकालीन प्रवृत्ति रेखा खींची जायेगी।

Year	Production	3 Year Semi Total	Semi Average
1995	40		
1996	48	132	$\frac{132}{3} = 44$
1997	44		
1998	60		
1999	56	180	$\frac{180}{3} = 60$
2000	64		

वर्ष 1999 के लिये अर्द्ध माध्य = 60

वर्ष 1996 के लिये अर्द्ध माध्य = 44

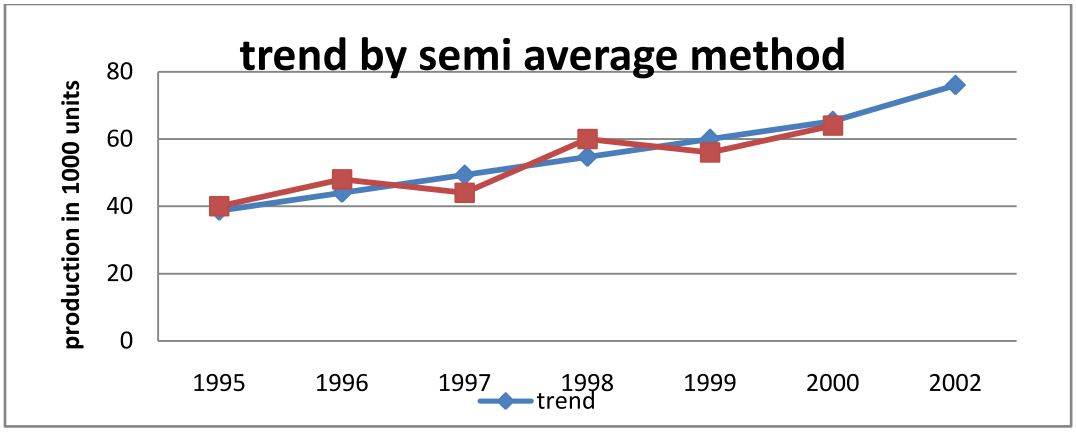
वार्षिक वृद्धि $\frac{16}{3} = 5.33$

प्रवृत्ति मूल्यों की गणना

वर्ष		उत्पादन
1995	44-5.33	38.67
1996		44
1997	44+5.33	49.33
1998	60-5.33	54.67
1999		60
2000	60+5.33	65.33

उपर्युक्त गणना के आधार पर 2002 का मूल्य

= 1999 का अर्द्ध मूल्य + 5.33 x 3 = 60 + 16 = 76



अर्द्ध मध्यक रीति के गुण-

1. सरलता

2. वतुनिष्ठता एवं निश्चितता

3. पूर्व या भावी अनुमान

दोष -

(1) रेखीय प्रवृत्ति - यह रीति तभी प्रयोग हो सकती है जब दीर्घकालीन प्रवृत्ति लगभग रेखीय है।

(2) चरम मूल्यों का प्रभाव - मूल्य बहुत बड़े या छोटे होने पर अद्ध मध्यकों पर प्रभाव पड़ता है। और प्रवृत्ति रेखा उचित प्रतिनिधित्व नहीं कर पाती।

2. अर्द्ध मध्यक रीति (Semi-median method)

यह एक लोचपूर्ण रीति है। जिसके अन्तर्गत दीर्घकालीन प्रवृत्ति को सरलता एवं प्रर्याप्त शुद्धता से ज्ञात किया जा सकता है।

यह रीति एक समान्तर माध्यों की श्रृख्लां है इसमें काल श्रेणी के निरन्तर अगले अतिव्यापी भाग के लिये समान्त माध्यों की गणना की जाती है।

यदि a, b, c, d, e, f छः वर्ष है और इनमें तीन वर्षीय चल माध्यांे की गणना करनी है तो यह गणना इस प्रकार की जायेगी।

$$\frac{a+b+c}{3}, \frac{b+c+d}{3}, \frac{c+d+c}{3}, \frac{d+e+f}{3}$$

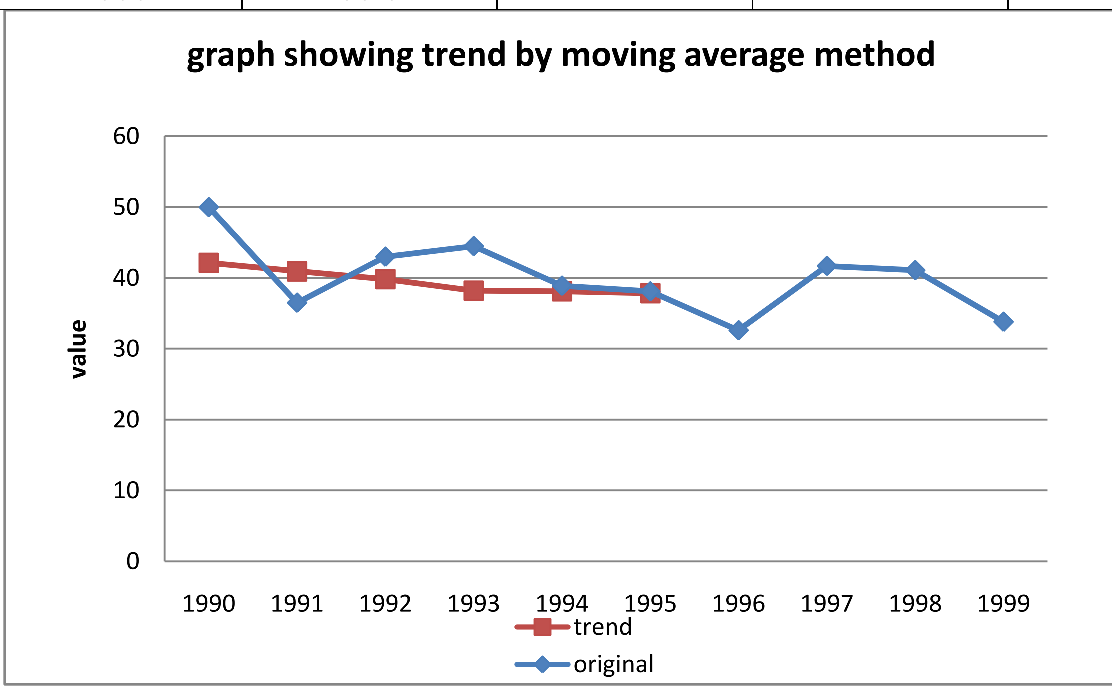
मूल प्रश्न यह उठता है कि कितने वर्षों का चल माध्य निकाला जाये-जैसे तीन वर्षीय , चार वर्षीय इत्यादि। चल माध्य की गणना की दृष्टि से प्रश्नों को दो भागों में बांटा जा सकता है-विषम अवधि चल माध्य एवं सम अवधि चल माध्य।

उदाहरण 3- निम्न समंको से 4 वर्षीय चलमाध्य की गणना कीजिये और प्रवृत्ति को बिन्दुओं पत्र पर प्रोकित कीजिए।

वर्ष	मूल्य	वर्ष	मूल्य
1990	30.0	1995	38.10
1991	36.5	1996	32.60
1992	43.0	1997	41.70
1993	44.5	1998	41.10
1994	38.9	1999	33.80

हल- Calculation of trend values by 4 yearly moving average-

year	value	4 yearly moving tables	2 yearly moving tables	Moving Average
1990	50.0			
1991	36.5	174.00		
1992	43.0	162.90	336.9	42.11
1993	44.5	164.50	327.4	40.93
1994	38.9	154.10	318.6	39.83
1995	38.1	151.30	305.4	38.18
1996	32.6	153.50	304.8	38.10
1997	41.7	149.20	302.7	37.84
1998	41.7			
1999	33.8			



चल माध्य की अवधि -

जितनी अधिक अवधि का चल माध्य होगा, उतनी ही अनियमित उच्चावचनों की गहनता उतनी ही कम होती जायेगी अतः यदि अनियमित उच्चावचनों को कम करना हो तो लम्बी अवधि का चल माध्य लेना चाहिये।

परन्तु ऐसा करने मे एक दोष उत्पन्न होता है वह यह कि लम्बी अवधि का चल माध्य लेने पर प्रवृत्ति मूल्य वास्तविकता मूल्यों से उतनी ही दूर होते चले जायेंगे। अतः माध्य की अनुकूलतम अवधि वह होती है जो काल श्रेणी में विद्यमान चक्रीय अवधि के बराबर या गुणांक में है। ऐसा करने से चक्रीय विचरण अनियमित उच्चारण लगभग कम हो जाते हैं। और प्रवृति मूल्य का श्रेष्ठ सम्भावित मान मिल जाता है।

चल माध्य रीति के गुण -

1. सरल
2. वस्तुनिष्ठता एवं निश्चिचता

3. लोचदार ने मूल्य बढ़ने पर सभी गणनाएँ पुनः नहीं करनी होती वरन् कुछ अतिरिक्त माध्य बढ़ जाते हैं।
4. चक्रीय उच्चावचनों का उन्मूलन यह तब संभव है जब चल माध्य का अवधि काल श्रेणी के चक्र की अवधि को ध्यान में रखकर निर्धारित कर ली जाये।
5. चरम मूल्यों का प्रभाव जिसके कारण प्रवृत्ति मूल्यों को उचित प्रकार ज्ञात नहीं किया जा सकता।

यदि काल श्रेणी में उच्चावचनों नियमित हो तो यह रीति सर्वश्रेष्ठ मानी जाती है।

3. न्यूनतम वर्ग रीति (Least square method)

यह रीति दीर्घकालीन प्रवृत्ति को ज्ञात करने की सर्वश्रेष्ठ रीति माना जाता है। इसके अन्तर्गत गणितीय समीकरणें के प्रयोग द्वारा न्यूनतम वर्ग मान्यता के आधार पर श्रेणी के लिये सर्वाधिक उपयुक्त रेखा खींची जाती है। यह रेखा सरल या परवलयिक वक्र का रूप ले सकती है।

इस रीतिको न्यूनतम वर्ग रीति इसलिये कहा जाता है क्योंकि इस रीति के आधार पर खींजी गयी प्रवृत्ति रेखा से मूल संमको के बिन्दुओं के विचलनों के वर्गों का योग अन्य किसी भी रेखा की तुलना में न्यूनतम होता है।

इस रीति के आधार पर प्रवृत्ति निर्धारण को तीन वर्गों में बाटा जा सकता है -

1. सरल रेखा प्रवृत्ति अन्वायोजन।
2. परवलय वक्रीय अथवा अरेखीय अप्रवृत्ति अन्यायोजन।
3. अर्द्ध लघुगणकीय या घातांकीय वक्र।

(A) सरल रेखीय प्रवृत्ति अन्वायोजन (Fitting a Straight line trend)-

इसके अन्तर्गत निम्न आधारभूत समीकरण का प्रयोग किया जाता है।

$y_c = a + bx$ y_c = अभष्टि उपनत्ति मूल्य

X = समय की इकाई

अचर मूल्य a, b के इस प्रकार की जाती है

(1) दीर्घ रीति द्वारा (2) लघु रीति द्वारा

दीर्घ रीति - समय बिन्दुओं के लिये आरम्भ से क्रम संख्याएँ (1, 2, 3, ...आदि) प्रयुक्त की जाती है। ये क्रम संख्याएं x द्वारा व्यक्त की जाती है और इनका योग $\sum x$ कर लिया जाता है।

1. क्रम संख्याओ के वर्गों का योग $\sum x^2$ निकाला जाता है।
2. x और मूल समंको y के मूल्यों की गुणा करके उनका जोड़ $\sum xy$ प्राप्त किया जाता है।
3. y मूल्यो का जोड़ $\sum x$ प्राप्त किया जाता है।
4. $\sum x, \sum x^2, \sum xy, \sum y$, निकालने के बाद निम्न समीकरणों के द्वारा और के मूल्य निकाले जाते हैं।

$\sum y = Na + b\sum x$

$$\sum xy = a\sum x + b\sum x^2$$

a और b प्राप्त करके सरल रेखा के आधारभूत समीकरण को प्रयोग करके प्रवत्ति मूल्य निकाला जाता है।

लघु रीति - प्रवृत्ति निकालने के लिये यदि मध्यका वर्ष को मूलबिन्दु (O) माना जाये तो गणन क्रिया अत्यन्त सरल हो जाता है। $\sum xy$ शून्य हो जाता है, अतः

$$\sum y = Na$$
$$\sum xy = b\sum x^2$$

अतः $a = \frac{\sum y}{N}$ $\quad b = \frac{\sum xy}{\sum x^2}$

उदाहरण 4- काल श्रेणी के निम्न संमकों से न्यूनतक वर्ग रीति द्वारा प्रवृत्ति ज्ञात कीजिए।

Year	2001	2002	2003	2004	2005	2006
Sales	5	7	9	10	12	17

-

Year	y	x	xy	x^2	y_c
2001	5	1	5	1	2.4+2.17x = 4.47
2002	7	2	14	4	2.4+2.17x2 = 6.74
2003	9	3	27	9	2.4+2.17x3 = 8.91
2004	10	4	40	16	2.4+2.17x4 = 11.08
2005	12	5	60	25	2.4+2.17x5 = 13.25
2006	17	6	102	35	2.4+2.17x6 = 15.42

N = 6, $\sum y = 60$, $\sum x = 21$, $\sum xy = 248, \sum x^2 = 91$

$\sum y = Na + b\sum x$

60 = 6a + 21b

$\sum xy = a\sum x + b\sum x^2$

248 = 21a + 91b

दोनों समीकरणों को हल करने पर-

420 = 42a + 147b

<u>496 = 42a + 182b</u> (2 से गुणा करने पर)

-76 = -35b

का मान (1) पर रखने पर- 60 = 6a + 21 x 2.17

$a = \frac{14.43}{6} = 2.4$

$y_c = 2.4 + 2.17x$

उदाहरण 5 - न्यूनतम वर्ग रीति द्वारा निम्नलिखित संमको की सरल रेखीय प्रवृत्ति का अन्वायोजन कीजिए।

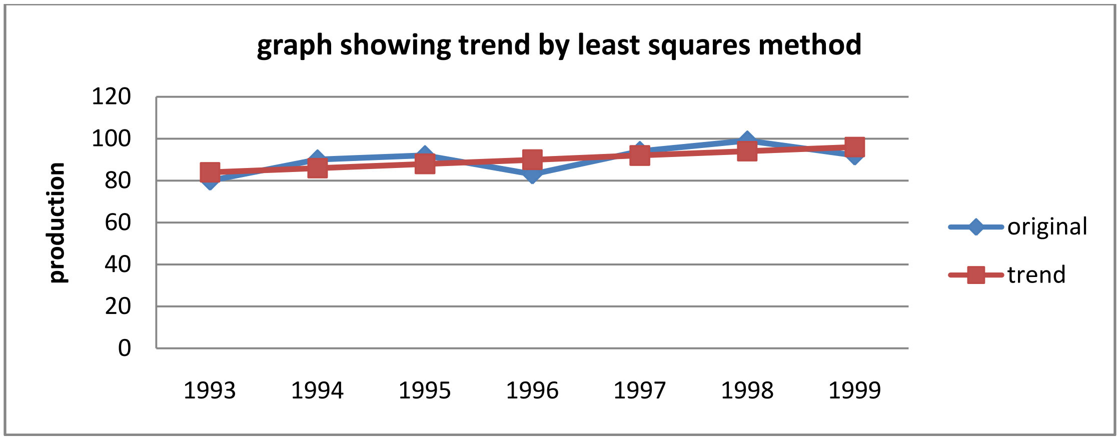
Year	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**
Sales	80	90	92	83	94	99	92

-

Year	Prood (y)	Deviation from 1996 (x)	Square (x^2)	xy	TREND VALUE $a + bx = y_0$
1993	80	-3	9	-240	90 + 2x-3 = 84
1994	90	-2	4	-180	90 + 2x-2 = 86
1995	92	-1	1	-92	90 + 2x-1 = 98
1996	83	0	0	0	90 + 2x+ = 90
1997	94	1	1	94	90 + 2x1 = 92
1998	99	2	4	198	90 + 2x2 = 96
1999	92	3	9	276	90 + 2x3 = 96
N = 7	$\sum y = 630$	$\sum x = 0$	$\sum x^2 = 28$	$\sum xy = 56$	$\sum y_c = 630$

$$a = \frac{\sum y}{N} = \frac{630}{7} = 90 \quad , \quad b = \frac{\sum xy}{\sum x^2} = \frac{56}{28} = 2$$

$$y_c = 90 + 2x$$



B. परवलय-वक्रीय अथवा अरेखीय प्रवृत्ति अन्वायोजन (Fitting a parabolic Non linear trend)

कभी कभार ऐसी स्थिति होती है जहाँ सरल रेखा दीर्घकालीन प्रवृत्ति का यथार्थ रूप में प्रस्तुतिकरण नहीं कर पाती। ऐसी स्थिति में प्रवृति निकालने के लिये निश्चित घात का परवलयिक वक्र या ऐकन्द्रित वक्र खीचंना पड़ता है। उदाहरण के लिये द्वितीय घात के परवलयिक वक्र (Second degree or parabolic curve of the second degree) को स्पष्ट करना चाहे तो इसका मूल समीकरण निम्न प्रकार से है-

$y = a + bx + cx^2$

यहाँ a, b, c अचल मूल्य है, जिन्हें ज्ञात करने के लिये निम्न समीकरणों का प्रयोग किया जाता है।

$$\sum y = Na + b\sum x + c\sum x^2$$

$$\sum xy = a\sum x + bx\sum x^2 + cx\sum x^2$$

$$\sum x^2 y = ax\sum x^2 + + bx\sum x^3 + c\sum x3$$

यदि विचलन काल श्रेणी के ठीक माध्य से लिया हो तो $\sum x = 0$ का मान शून्य हो जायेगा और उपर्युक्त समीकरण सरल रूप से निम्न हो जायेंगे।

$$\sum y = Na + c\sum x^2$$

$$\sum xy = b\sum x^2$$

$$\sum x^2 y = a\sum x^2 + c\sum x4$$

यहाँ $\sum x^3$ इसलिये समाप्त हो गया है, क्योंकि $\sum x = 0$ होगा तो $\sum x^3$ कामान भी शून्य हो जायेगा।

उदाहरण - निम्न आंकड़ो के लिये द्वितीय कोटि का परवलयिक वक्र अन्वायोजित कीजिए-

Year	1996	1997	1998	1999	2000
Value	10	12	113	10	8

-

YEAR	y	x	x^2	x^3	x_y	xy	x^2y
1996	10	-2	4	-8	16	-20	40
1997	12	-2	1	-1	1	-12	12
1998	13	0	1	0	0	0	0
1999	10	1	1	-1	1	10	10
2000	8	2	4	-8	16	16	32
N = 5	$\sum y = 53$	$\sum x = 0$	$\sum x^2 = 10$	$\sum x^3 = 0$	$\sum xy = 34$	$\sum xy = -6$	$\sum x^4 y = 94$

यहाँ पर $\sum x$ और $\sum x^3$ का योग 0 है।

$\sum y = Na + c\sum x^2$ or $53 = 5a + 10c$

$\sum xy = b\sum x^2$ or $-6 = 10b$

$\sum x^2 y = a\sum x^2 + c\sum x4$ or $94 = 10a + 34c$

समी. (2) में **10b = - 6 or b = -0.6**

समी. (1) में 2 से गुणा करने पर और समी (3) से घटाने पर-

$$94 = 10a + 34c$$

$$\frac{106 = 10a + 20c}{-12 = 14c}, \quad c = \frac{-12}{14} = 0.85$$

समी. (1) में c का मान रखने पर-

$$53 = 5a + 10x - 0.857$$

$$53 + 8.57 = 5a \text{ or } \quad 5a = 61.57$$

$$a = 12.314$$

$$y = a + bx + cx^2$$

$$= 12.314 + (-0.6)\ x\ (-0.85)x^2$$

$$= 12.314 - 0.6x - 0.857x^2$$

वर्ष	X	गणना	प्रवृत्ति मूल्य (y_c)
1996	-2	= 12.314 – 0.6x – 0.857 - 2^2	= 10.08
1997	-1	= 12.314 – 0.6 x – 1 – 0.857 x 1^2	= 12.05
1998	0	= 12.314	= 12.314
1999	1	= 12.314 – 0.6 x 1 -0.857 x1^2	= 10.85
2000	2	= 12314 -0.6 x 2 – 0.857 x 2^2	= 7.68

C- अर्द्ध-लघुगणकीय या घातांकीय चक्र(Semi-logarithm or exponential curve)-

यदि काल श्रेणी में डाक स्थिर प्रतिशत की दर से वृद्धि या कमी होता है तो अर्द्ध लघुगणकीय अथवा घाताकीय वक्र का प्रयोग उचित रहता है।

समी.

a और b के मान की गणना के लिये निम्न समी. का प्रयोग किया जाता है-

$$\sum(\log y) = N \log a + \log b \text{ x } \sum x$$

$$\sum(\log y) = \log a \text{ x } \sum x + \log b \text{ x } \sum x^2$$

यदि मूल बिन्दु मध्यका से लिये जाते हैं। तो उप्रयुक्त सूत्र की निम्न रूप से संक्षिप्तीकृत हो जाते हैं-

$$\sum(\log y) = N \log a \qquad \text{or} \qquad \log = \frac{\sum \log y}{N}$$

$$\sum(x \log y) = \log b \sum x^2 \qquad \text{or} \qquad \log b = \frac{\sum(x \log y)}{\sum x^2}$$

यह वक्र सरल रेखा के रूप में ही बनता है यदि इसे अर्द्ध लघुगणकीय ग्राफ पर अंकित किया जाये लेकिन सामान्य ग्राफ पर यह वक्र अरेखीय हो जाता है।

न्यूनतम वर्ग रीति के लाभ-

1. पूर्णता वस्तुनिष्ठ
2. पूर्वानुमान की सुविधा
3. पूरी अवधि के लिये प्रवृत्ति ज्ञात हो जाती है।
4. सर्वोपयुक्त रेखा
5. परिवर्तन दर की जानकारी

सीमाएँ-

1. कठिन एवं जटिल रीति
2. लोच का अभाव

3. समीकरण का गलत चुनाव
4. भावी पूर्वानुमान की सीमाएँ क्योंकि इसमें मौसमी चक्रीय आदि उच्चावचनो को ध्यान में नहीं रखा जाता।

9.7 अल्पकालीन उच्चावचन का मापन (Measurement of short term variation)

काल श्रेणी पर दीर्घकालीन प्रवृत्ति और अप्लकालीन उच्चवचनों दो नों का ही सामूहिक प्रभाव पड़ता है। अतः चल माध्य या न्यूनतम वर्ग रीति द्वारा निकाल गये प्रवृत्ति सहायक को मूल श्रेणी में से कर दिया जाये तो अल्पकालीन उच्चावचन शेष रह जाता है। इनको मापने की प्रमुख रीतियां निम्न प्रकार से है-

1. सरल माध्य या आर्त्तव माध्य या आर्त्तव विचरण (Simple average or seasonal Average or seasonal variation index method)
2. चल माध्य द्वारा आर्त्तव विचरण (Seasonal variation through moving average)
3. श्रृखंला मूल्यानुपात रीति (trend relative method)
4. प्रवृत्ति अनुमान रीति (Ratio to trend method)
5. चल माध्य अनुपात रीति (Ratio to moving average method)

1. सरल माध्य या आर्त्तव माध्य या आर्त्तव विचरण (Simple average or seasonal Average or seasonal variation index method)

आत्तर्व विचरण निकालने की यह सबसे सरल रीति है। इसका प्रयोग अधिकतर 12 मास आंकड़ो से ऋतुनिष्ठता का माप करने के लिये किया जाता है। यह रीति उस परिस्थिति में उपयुक्त है जहाँ आँकडो में कोई सुनिश्चित दीघकालीन प्रवृति स्पष्ट रूप से दृष्टिगोचर न हो।

$$\text{ऋतुनिष्ठ विचरण सूचकांक} = \frac{\text{ऋतुकालिक माध्य}}{\text{सामान्य माध्य}} \times 100$$

उदाहरण निम्न संमको से ऋतुनिष्ठ सूचकाकों की गणना कीजिए-

Year	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
2004	15	16	18	23	23	23	20	28	29	33	33	38
2005	23	22	28	31	31	28	22	28	32	37	34	44
2006	25	25	35	36	36	30	30	24	38	48	41	53

Calculation of seasonal variation indices by monthly average

Month	1998	1999	2000	Total	Monthly Avg	Seasonal India No.
Jan	15	23	25	63	21	70
Feb	16	22	25	63	21	70

Mar	18	28	35	81	27	90
Apr	18	27	36	81	27	90
May	23	21	36	90	30	100
Jun	23	28	30	81	27	90
July	20	22	30	72	24	80
Aug	28	28	34	90	30	100
Sept	29	32	38	99	33	110
Oct	33	37	47	117	39	130
Nov	33	34	41	108	36	120
Dec	38	44	53	135	45	150
Total				1080	360	1200
Average				90	30	100

आर्त्तव विचरण निदेशांक = $\frac{\text{मासिक माध्य}}{\text{सामान्य माध्य}}$ X 100.

जैसे जनवरी = $\frac{21x100}{30} = 70$ आदि

2. चल माध्य द्वारा मौसमी विचरण (Seasonal variation by moving mean)

यह रीति प्रथम रीति से अधिक श्रेष्ठ है। इसके द्वारा चक्रीय उच्चावचनों को छोड़कर सभी प्रकार के वित्रानी प्रवृत्ति (trend) अल्पकालीन परिवर्तन (short time oscillations) मौसमी, तथा दैव उच्चावचन (random fluctuations) आदि का विशलेषण हो जाता है। यदि काल श्रेणी के मूल संमको पर उपनति का भी प्रभाव हो तो चल माध्यों का प्रयोग करके मॉसमी विच रणों का मापन किया जा सकता है। इस रीति का यह विशेषज्ञ लाभ है कि इसके द्वारा लगभग सभी प्रकार के विचर णों, प्रवृत्ति अल्पकालिक परिवर्तन तथा ऋतुनिष्ठ एवं अनियमित या दैव उच्चारण का विश्लेषण हो जाता है। यह रीति काल श्रेणी विश्लेषण के योगशील निदर्श पर आधारित है।

स्पष्तः यदि मौसमी परिवर्तनों को कुल उच्चावाचनों में से घटा दिया जाए तो शेष अनियमित उच्चावचन होंगे।

3. श्रृंखला मूल्यानुपात विधि (Series Ratio Method)

मौसमी विचरण का विश्लेषण करने की यह एक सन्तोषजनक रीति है। इसके अनुसार पहले मौसम के श्रृखलानुपात परिगणित किये जाते हैं तथा फिर उनमें से अवशिष्ट प्रवृत्ति निकाल ली जाती है। इसकी क्रिया विधि इस प्रकार है-

1. प्रत्येक मौसम का निम्न सूत्र द्वारा श्रृंखला मूल्यानुपात ज्ञात किया जायेग।

$$\text{श्रृंखला मूल्यानुपात} = \frac{\text{प्रचलित ऋतु मूल्य}}{\text{पिछला ऋतु मूल्य}}$$

2. प्रत्येक अवधि के श्रृंख्ला मूल्यानुपातो को समान्तर माध्य निकाला जायेगा।
3. उक्त श्रृंखला मूल्यानुपात माध्यों का प्रथम कालाविधि अधिक पर श्रृंख्ला सूचंकाकों में बदला जायेगा।

प्रचलित ऋतु का श्रृंख्ला सूचकांक =.
CR = Chain Relative (श्रृंखला सूचकांक)
ALR = Link Relation (श्रृखं्ला मूल्यानुपात का माध्य)
अन्तिम अवधि को आधार मानकर प्रथम अवधि का श्रंख्ला सूचकांक निकाला जायेगा।
प्रथम ऋतु का संगणित (R) = अन्तिम ऋतु का LR प्रथम ऋतु का ALR / 100

4. प्रवृत्ति अनुपात विधि (Trend Ratio Method)

यह रीति गुणनात्मक निदर्श पर आधारित है। प्रवृति को अधिक महत्व देती है और गणना क्रिया जटिल होने के कारण इसका प्रयोग भी कम किया जाता है। मूल समंकों को वर्ष के अनुसार अनुविन्यसित करके उनका मसमी माध्य ज्ञात कर लिया जाता है। न्यूनतम वर्ग रीति से मूल समंकों के लिए दीर्घकालीन प्रवृत्ति (Secular trend) मालूम कर ली जाती है जो वार्षिक वृद्धि अथवा हास की दर प्रदर्शित करती है। इसे मासिक समंको की स्थिति में 12 तथा त्रैमासिक समंकों की स्थिति में 4 से भाग देकर मौसमी वृद्धि अथवा ह्मास दर मालूम कर ली जाती है। फिर मालूम किये गए मौसम माध्यों का माध्य मालूम कर लिया जाता है । यह माध्य श्रेणी के मध्य बिंदु का प्रवृत्ति मूल्य है। इस प्रकार मालूम किये गए प्रवुत्ति मूल्यों को 100 मानकर तत्संबंधी माध्य मूल्यों को प्रतिशत में बदल दिया जाता है । प्राप्त परिणाम ही मौसमी परिवर्तन के निर्देशांक होते हैं ।

9.8 काल श्रेणी का महत्व (Importance of Time Series)

(1) भूतकाल के व्यवहार का विश्लेषण - इस आधार पर व्यवहारों को नियन्त्रण करने की सुव्यवस्था हो सकती है।

(2) भविष्य के विषय में अनुमान- इसके बारे में बर्नर हिर्श का कहना है, काल श्रेणी का विश्लेषण करने का मुख्य उद्देश्य भावी घटनाओं की गतिविधि का वास्तविक अनुमान लगाने के लिये आर्थिक तथ्यों में होने वाली परिवर्तनो को समझना, समझाना एवं मूल्याकिंत करना है।

(3) तुलनात्मक अध्ययन - दो या दो से अधिक सम्बन्धित समय अवधियों के संमकों का तुल्नात्मक अध्ययन करना सम्भव हो जाता है।

(4) व्यापार चक्रों का अनुमान - चक्रीय उच्चावचनों के आधार पर व्यापार चक्रों का अनुमान लगाया जा सकता है। व्यवसायों अपनो क्रियाओं को नियोजित कर सकता है।

अन्य साख्यिंकीय उपकरणों की भॉति काल श्रेणी विश्लेषण भी अनुमानित एंव सामान्य परिणाम तथा संकेत प्रदान करता है।

9.9 सारांश (Summary)

काल की गति के साथ मूल्यों मे होने वाले विभिन्न दीर्घकाल एवं अल्पकालीन उ च्चावचनों का विधिवत् विश्लेषण किसान, उपभोक्ता, व्यापारी, प्रशासक आदि सभी वर्गो के व्यक्तियों के लिये आवश्यक और उपयोगी होता है। निष्कर्ष रूप में यह कहा जा सकता है कि काल श्रेणी का आशय समय क्रम में सांख्यिकीय संमको की व्यवस्था से

है। यह श्रेणी समय परिवर्तन के साथ ही तथ्य विशेष के संमकों में होने वाले परिवर्तनों को स्पष्ट करती है। इन परिवर्तनों को कुछ वर्गों में बाँट सकते हैं और वर्ग ही काल श्रेणी के संघटक कहे जाते हैं।मूल संमकों को 'σ' से दर्शाया जाता है, इसके चार संघटक है।

दीर्घकालीन प्रवृत्ति (T) मौसमी विचरण (S) चक्रीय उच्चारण (C) अनियमित उच्चावचन (I)

दीर्घकालीन प्रवृति को मापने के लिये चार प्रमुख रीतियाँ निम्न प्रकार है - मुक्त हस्त रीति, अर्द्ध मध्यक रीति, चल माध्य रीति, एवं न्यूनतम वर्ग रीति। अल्पकालीन उच्चावचन को मापने की प्रमुख रीतियां सरल माध्य या आर्त्तव माध्य या आर्त्तव विचरण , चल माध्य द्वारा आर्त्तव विचरण , श्रृखंला मूल्यानुपात रीति , चल माध्य अनुपात रीति एवं प्रवृत्ति अनुपात रीति है । काल श्रेणी में होने वाले दीर्घकालीन एवं अल्पकालीन उच्चावचनों का अध्ययन न सिर्फ व्यापारी वरन् अर्थशास्त्री के लिए भी बड़ा महत्व रखता है। भूतकाल के परिवर्तन के विश्लेषण करके वे पिछले अनुभव के आधार पर भविष्य की नीतियाँ निर्धारित कर सकते हैं। और अपनी क्रिया ओं पर नियंत्रण करके भविष्य के जोखिमों से अपने व्यापार की सुरक्षा कर सकते हैं। अतः यह कह सकते हैं कि विभिन्न वर्ग चाहे वो अर्थशास्त्री हो या उपभोक्ता , योजनाकार, किसान, राजनीतिक आदि सभी के लिये काल श्रेणी में से वाले परिवर्तनों का विश्लेषण विशेष रूप से उपयोगी होता है। एक विवेकपूर्ण विश्लेषण तथा संकेतको का वैज्ञानिक विवेचन काल श्रेणी की महत्ता में वृद्धि करता है।

9.10 शब्दावली (Glossary)

- **समयानुसार (Chronological series)**–
- **संघटक (Components)** –
- **मौसमी परिवर्तन ()** – ऐसे परिवर्तन जो नियमित और आवर्त्तक होते हैं।
- **पुनरावृत्ति (Repetition)** -
- **दैव उच्चावचन ()** – अनियमित उतार-चढाव
- **सरलित वक्र (Smoothed Curve)** -
- **मध्यका बिन्दु (Mediain pt)** -
- **अगले अतिव्यापी (Successive overlapping)**-
- **सर्वाधिक उपयुक्त रेखा (Line of best fit)** -
- **उपनति (Trend)** -

9.11 लघु उत्तरीय प्रश्न (Short Answer Questions)

(1) नियमित अल्पकालीन उच्चावचन को विभाजित किया जाता है।

(a) ------------------------ (b) ----------------------------

(2) काल माला के गुणन मॉडल में γ =

(3) O - T - S = --------------------------

(4) किसी काल माला के समंको में बढ़ने या घटने की दीर्घकालीन प्रवृत्ति को .... कहते हैं।

(5) मौसमी विचरण .... अवधि के अल्पकालीन उच्चावचन है।

उत्तर-

1) मौसमी, चक्रीय
2) γ = T x S x C x I
3) C + T

4) सुदीर्घकाल उपनति
5) चक्रीय

9.12 बहुविकल्पीय प्रश्न (Multiple Choice Questions)

(1) सामान्त माध्य की प्रमाप त्रुटि का सूत्र-

(i) $\frac{\sigma\, p}{n}$ (ii) $\frac{\sqrt{\sigma p}}{n}$

(iii) $\sqrt{\frac{\sigma\, p}{n}}$ (iv) $\frac{\sigma\, p}{\sqrt{n}}$

(2) माध्य विचलन को प्रमाप त्रुटि होती है-

(i) $0.78672\ \sigma/\sqrt{n}$ (ii) $0.6028\ \sigma/\sqrt{n}$

(iii) $\sigma/\sqrt{n}$ (iv) $1.36\ \sigma/\sqrt{n}$

9.13 सदंभ सहित ग्रन्थ (Books with References)

- **डा0 एस सचदेवा** - परिमाणात्मक विधियाँ ,लक्ष्मी नारायण अग्रवाल, आगरा
- **डा0 के0 एल0 गुप्ता एवं डा0 हरिओम गुप्ता** - परिमाणात्मक तकनीकें ,नवयुग साहित्य भवन, आगरा।
- **डा0 के 0 एल 0 गुप्ता , रवि कान्त** - अर्थशास्त्र की आधारभूत परिमाणात्मक विधियाँ ,नवनीत पब्लिकेशन्स, आगरा
- **एस0पी0 सिंह** – सांख्यिकी: सिद्धान्त एवं व्यवहार, एस0 चन्द पब्लिकेशन्स नई दिल्ली।

9.14 कुछ उपयोगी पुस्तकें (Some Useful Books)

- Kumar, Anil, (2008) Statistical Research Methodology, Aifa Publishing House.
- Singh, S.P.(2010) Principles of Statistics, S & Chand Publishing House.
- Bhardwaj,R.S. (2000). Mathematics for Economics and Business, Excel Books.
- Bose,D.,(2003), An introduction to Mathematical Economics, Himalaya Publishing House.

9.15 निबन्धात्मक प्रश्न (Essay Type Question)

(1) काल श्रेणी के विश्लेषण से आप क्या समझते है? उपनति मापन की विधियों का संक्षिप्त वर्णन कीजिये।

(2) काल श्रेणी क्या है? दीर्घकालीन प्रवृत्ति, मौसमी परिवर्तनों तथा चक्रीय उच्चावचनों में अन्तर स्पष्ट कीजिये? किन्हीं दिये गये संमकों में दीर्घकालीन प्रवृत्ति की माप किस प्रकार करेंगे।

इकाई 10 काई-वर्ग परीक्षण
(Chi-Square Test)

10.1 प्रस्तावना (Introduction)

जैसा कि आप जानते हैं कि विभिन्न सांख्यिकीय उपकरणों का प्रयोग परिकल्पनाओं के परीक्षण के लिए किया जाता है। इन सांख्यिकीय परीक्षणों को दो अलग अलग समूहों में वर्गीकृत किया जा सकता है। प्राचल परीक्षण और गैर प्राचल परीक्षण। प्राचल परीक्षण वे होते हैं जो समग्र के मापदंडों पर आधारित होते हैं। प्राचल परीक्षणों को प्रयोग करने में , समग्र वितरण की निश्चित मान्यताओं को पूरा करना आवश्यक है। यह इन परीक्षणों के कार्यान्वयन के लिए प्रतिबंध बनता है। इन प्रतिबंधों को रोकने के लिए हम दूसरे वर्ग के परीक्षणों का प्रयोग कर सकते हैं जिन्हें गैर प्राचल परीक्षणों के रूप में जाना जाता है। इस ढांचे में , आप विभिन्न गैर प्राचल परीक्षणों के बारे में अध्ययन करेंगें। इन सभी परीक्षणों में , काई वर्ग परीक्षण सबसे लोकप्रिय गैर प्राचल परीक्षण है लेकिन इसे प्राचल परीक्षण के रूप में भी प्रयोग किया जा सकता है। इस अध्याय में आप सैद्धान्तिक अवधारणा और काई वर्ग परीक्षण के विभिन्न उपयोगों के बारे में विस्तार से अध्ययन करेंगें।

10.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप-

- ✓ काई-वर्ग वितरण की अवधारणा की व्याख्या को जानेंगे।
- ✓ काई-वर्ग वितरण के प्रयोगों को सीखेंगे।

10.3 काई-वर्ग परीक्षण का अर्थ एवं प्रयोग (Meaning and Use of Chi-Square Test)

संख्यिकीविदों द्वारा विकसित किये गये कई परीक्षणों के बीच काई-वर्ग परीक्षण एक महत्वपूर्ण परीक्षण है। 1900 में प्रोफेसर **कार्ल पियरसन** द्वारा इस परीक्षण का प्रस्ताव किया था। काई वर्ग परीक्षण को x^2 प्रतीक द्वारा प्रदर्शित किया जाता है। इसकी उत्पत्ति ग्रीक अक्षर 'ची' से होती है। काई-वर्ग परीक्षण का उपयोग समग्र के भिन्नता की तुलना करने के लिए प्राचलन के साथ ही गैर प्राचल परीक्षण के रूप में किया जा सकता है . जैसे स्वतन्त्र के परीक्षण के रूप में या Goodness of fit के परीक्षण के रूप में। सैद्धांतिक विचरण को भिन्नता के लिए नमूना विश्लेषण के संदर्भ में यह परीक्षण मुख्य रूप से प्रयोग किया जाता है। नील आर उल्लमन के अनुसार,"गैर प्राचलन परीक्षण के रूप में , यह निर्धारित करने के लिए प्रयोग किये जा सकते हैं कि क्या आंकडे स्पष्ट निर्भरता दिखाते हैं या दो वर्ग स्वतन्त्र है। जब वर्ग प्रयोग किये जाते हैं, इसका प्रयोग सैद्धान्तिक समग्र एवं वास्तविक आंकडों के बीच तुलना करने में किया जा सकता है। "x^2 की यात्रा सैद्धान्तिक एवं अवलोकित विसंगति के परिमाण का वर्णन करती है , अर्थात् x^2 की सहायता से हम जान सकते हैं कि क्या सिद्धा न्त एवं अवलोकन के बीच दी हुई विसंगति संयोगवश विशेषता से हुई है या क्या यह सिद्धान्त की अपर्याप्तता से बनाये अवलोकित तथ्यों को जोडने के लिए है।

x^2 परीक्षण सांख्यिकीय कार्य में सबसे आसान और सर्वाधिक व्यापक रूप से प्रयोग किया जाने वाला गैर प्राचल परीक्षणों में से एक है , जिसे कई स्थितियों में प्रयोग किया जा सकता है। इस तकनीक को निम्नलिखित प्रयोजनों के लिए प्रयोग किया जाता है-

- Goodness of fit के परीक्षण के लिए
- दो विशेषताओं के बीच सम्बन्ध के बीच महत्व के परीक्षण के लिए
- समरूपता या समग्र में भिन्नता के महत्व को जानने के लिए

10.4 x^2 वितरण (x^2 Distribution)

काई-वर्ग चर के प्रायिकता वितरण को काई वर्ग वितरण कहते हैं। य दि $Z_1, Z_2, \ldots\ldots, Z_K$ स्वतन्त्र यादृच्छिक चर हों, प्रत्येक के पास एक मानक सामान्य वितरण N(0,1) हो, तब $\chi_k^2 = Z_1^2 + Z_2^2 + \ldots\ldots\ldots + Z_k^2$ के वितरण को x^2 **वितरण** कहा जाता है।

वैचारिक रूप में ,x^2 अवलोकित एवं अपेक्षित आवृत्तियों के अस्तित्व के बीच विसंगति की माप है जिसका मान मुख्य रूप से स्वतन्त्रता की मात्रा पर निर्भर करता है। स्वतन्त्रता की मात्रा मानो की वह संख्या है जिसे हम स्वतन्त्र रूप से चुन सकते हैं , अर्थात वे अन्य पूर्व निर्धारित प्राचलों द्वारा तय नहीं किया जाता है। इसका अर्थ है कि काई वर्ग परीक्षण में केवल एक प्राचल होता है अर्थात स्वतन्त्रता की मात्रा की संख्या । उदाहरण के लिए, यदि यह दिया गया है कि तीन चरों का योग 50 के बराबर हो, तो हम कोई भी दो चरों का चयन करने के लिए स्वतन्त्र हैं लेकिन तीसरा चर 50 (दो चरों का योग) के बराबर होना चाहिए, क्योंकि केवल तभी तीन चरों का योग 50 के बराबर होगा। यहाँ स्वतन्त्रता की मात्रा 2 है। स्वतन्त्रता की मात्रा को सामान्यतया r या df द्वारा प्रदर्शित किया जाता है। x^2 परीक्षण के माध्यम से हम सिद्धान्त या अपेक्षित मान और अवलोकित या वास्तविक मान के बीच अंतर की सीमा निर्धारित करने में सक्षम है।

गणितीय रूप में इसे निम्नवत परिभाषित किया जाता है-

$$x^2 = \sum \frac{(O - E)^2}{E}$$

जहाँ O = अवलोकित आवृत्तियाँ

E = अपेक्षित आवृत्तियाँ

काई वर्ग परीक्षण विशेष रूप से संज्ञात्मक आंकडों से जुडे परीक्षणों में उपयोगी है , लेकिन इसका उपयोग उच्च मानदण्डों के लिए भी किया जा सकता है। संज्ञात्मक आंकडा माप का सबसे प्रारम्भिक रूप है , जो विभाजन को एक श्रेणी में व्यवस्थित करता है जो कि पारस्परिक रूप से अनन्य और सामूहिक रूप से संपूर्ण है। उदाहरण के लिए, समग्र को दो श्रेणियों में वर्गीकृत किया जा सकता है। पुरूष और महिलाएं।

10.4.1 x^2 वितरण के गुण (Properties of C'2 Distribution)

x^2 वितरण में कई गुण है, वितरण के कुछ महत्वपूर्ण गुण निम्नानुसार है-

1. x^2 वितरण एक निरंतर संभाव्यता वितरण है जिसका निचली सीमा में मान शून्य है और धनात्मक दिशा में अनंनता तक फैली हुई है। x^2 का नकारात्मक मान सम्भव नहीं है क्योंकि अवलोकित और अपेक्षित आवृत्तियों के बीच के अंतर की हमेंशा वर्ग होता है , इसलिए x^2 का मान कभी नकारात्मक नहीं हो सकता है।

2. वितरण का सही आकार स्वतन्त्रता की मात्रा की संख्या (v) पर निर्भर करता है। सामान्य रूप में जब अ छोटा है, वक्र का आकार दायीं ओर तिरछा है और जैसे ही अ बडा हो जाता है , वितरण अधिक से अधिक सममित हो जाता है और सामान्य वितरण द्वारा अनुमानित किया जा सकता है।

3. x^2 वितरण का माध्य स्वतन्त्रता की मात्रा द्वारा दी जाती है और भिन्नता स्वतन्त्रता की मात्रा की दो गुनी है। इसे निम्नानुसार व्यक्त किया जा सकता है:

E(x^2) = μ = v

V(x^2) = σ^2 = 2v

4. x^2 एक नमूना आंकडा है जिसके पास कोई संगत प्राचल नहीं है। यह x^2 वितरण गैर प्राचल वितरण को बनाता है।

5. x^2 वितरण के लिए योगात्मक गुण अच्छा निर्णय देता है। इसका अर्थ है कि स्वतन्त्र x^2 चरों का योग भी x^2 चर होता है। इस प्रकार यदि $x_1{}^2$ v_1 d.f के साथ x^2 चर है और $x_2{}^2$ v_2 d.f के साथ x^2 का दूसरा चर है। $x_1{}^2$ स्वतन्त्र है, तब उनका योग $x_1{}^2 + x_2{}^2$ भी एक $v_1 + v_2$ के साथ x^2 का चर है।

10.4.2 x^2 परीक्षण के प्रयोगों के लिए शर्तें (Conditions for Using x^2 Test)

एक काई -वर्ग परीक्षण का उपयोग किया जाना चाहिए, यदि निम्न स्थितियाँ संतुष्ट होती है-

1. सैद्धान्तिक रूप से सही वितरण और अध्ययन के नमूनाकरण वितरण के बीच समानता के एक माडेम के दायित्व के लिए ‘N’ की कुल संख्या बहुत बडी होनी चाहिए। N का सामान्यतया स्वीकार किया गया मान 50 होता है। इसलिए नमूने कम से कम 50 अवलोकन होने चाहिए।
2. नमूना आंकडा लक्षित समग्र से यादृच्छिक तरीके से लिया जाना चाहिए ताकि पक्षपाती (पूर्वाग्रह) कोई तत्व न हो ।
3. प्रयोगिक आंकडा या नमूना अवलोकन, अर्पित सभी वस्तुऐं या नमूने में अवलोकलन एक दूसरे से अलग होने चाहिए।
4. तुलनात्मकता की सुविधा के लिए आंकडों को मूल इकाईयों (पूर्ण रूप) में व्यक्त किया जाना चाहिए न कि तुलनात्मक रूप में जैसे कि प्रतिशत या अनुपात या समानुपात इत्यादि।
5. किसी भी कक्ष में पाँच अवलोकन से कम नहीं होना चाहिए। (प्रत्येक आंकडे प्रविष्टि को एक कक्ष के रूप में जाना जाता है।) यदि किसी समूह में स्वीकार स्तर के नीचे आवृत्तियाँ होती हैं तो आवृत्तियों का एकत्रीकरण किया जाता है जिससे कम आवृत्तियों को पूर्ववर्ती या बाद की आवृत्तियों में जोडा जाता है ताकि परिणामस्वरूप मान स्वीकार्य स्तर से अधिक हो कुछ सांख्यिकीयविद 5 के बजाय न्यूनतम स्वीकार्य स्तर के लिए 10 अंक को बेहतर मानते हैं।
6. प्रतिबंधों को रैखिक होना चाहिए अर्थात् प्रतिबंधों की परिभाषित करने वाले सभीकरणों में आवृत्तियों का कोई वर्ग या उच्च घातें नहीं होनी चाहिए।

10.5 विचरण की तुलना करने के लिए काई-वर्ग एक परीक्षण के रूप में (Chi-Square as a Test for Comparing Variance)

काई-वर्ग मान प्रायः समग्र विचरण के महत्व के आकलन के प्रयेाग के लिए किया जाता है। इसका अर्थ यह है कि जब सामन्य वक्र जिसका माध्य μ और विचलन σ_p^2 है में से एक यादृच्छिक नमूना लिया जाता है तो x^2 परीक्षण का प्रयोग किया जा सकता है। समग्र विचलन के महत्व के आंकलन के लिए, यह माना जाता है कि नमूना विचलन समग्र विचलन के बराबर हो। इस प्रकार शून्य परिकल्पना को निम्न प्रकार से लिया जाता है-

$$H_0 = \sigma_s^2 = \sigma_p^2$$

यह स्पष्ट है कि x^2 परीक्षण x^2 वितरण पर आधारित होता है जो कि मानों के संग्रह से संबंधित होती है और जिसमें वर्गों का योग शामिल हेाता है। जब हमें काई वर्ग का प्रयोग समग्र विचलन के परीक्षण के रूप में करना होता है, तो हमें शून्य परिकल्पना का परीक्षण करने के लिए निम्नलिखित सूत्र द्वारा x^2 का मान निकलना पडेगा।

$$x^2 = \frac{\sigma_s^2}{\sigma_p^2}(n - 1)$$

जहाँ

σ_s^2 = नमूने का विचलन

σ_p^2 = समग्र का विचलन

(n-1) = स्वतन्त्रता का अंश

n = नमूने में वस्तुओं की संख्या

स्वतन्त्रता के विभिन्न अंशों एवं स्तरों के लिए x^2 के महत्वपूर्ण मान तालिका के रूप में उपलब्ध हैं। आप इस तालिका को सांख्यिकीय की किसी भी अच्छी किताबों के परिशिष्टों में पा सकते हैं। किसी निर्णय में पहुॅचने के लिए उपरोक्त सूत्र द्वारा x^2 के मान की गणना की जाती है, गणना किये हुए मान को (n-1) स्वतंत्रता के अंश के साथ किसी विशेष स्तर पर तालिका मान से तुलना करते हैं। यदि x^2 का गणितीय मान तालिका के मान से कम है तो शून्य परिकल्पना को स्वीकृत किया जाता है। इसके विपरीत , यदि x^2 का गणना किया हुआ मान , तालिका मान के बराबर या अधिक हो तब परिकल्पना को अस्वीकार किया जाता है।

इस सम्बन्ध में याद रखने के लिए एक महत्वपूर्ण बात यह है कि काई वर्ग वितरण सममित नहीं है और सभी मान सकारात्मक है। भिन्नता की तुलना करने के लिए x^2 परीक्षण का उपयोग समग्र के सामान्य वितरण की धारणा पर आधारित होती है।

उदाहरण 1- 10 छात्रों का वजन इस प्रकार है-

क्रम. संख्या	1	2	3	4	5	6	7	8	9	10
वजन (किग्रा)	38	40	45	53	47	43	55	48	52	49

क्या हम कह सकते हैं कि 10 छात्रों के ऊपर दिए गए नमूने से सभी छात्रों के वजन का विचलन 20 किलोग्राम के बराबर है? 5 प्रतिशत एवं प्रतिशत महत्व के स्तर पर परीक्षण करें।

हल- सबसे पहले हमें नमूना आंकडा ज्ञात करना चाहिए, अर्थात σ^2 जिसकी गणना निम्नानुसर की जाती है:-

क्रम.सं.	X_i (वनज किग्रा में)	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
1	38	-9	81
2	40	- 7	49
3	45	- 2	04
4	53	+ 6	36
5	47	0	00
6	43	- 4	16
7	55	+8	64
8	48	+1	01
9	52	+5	25
10	49	+2	04
n = 10	$\sum X_i$ = 470		$\sum(X_i - \bar{X})^2$ = 280

$$\bar{X} = \frac{\sum X_i}{n} = \frac{470}{10} = 47 \text{ kgs}$$

$$\therefore \quad \sigma_s = \sqrt{\frac{(X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{280}{10-1}} = \sqrt{31.11}$$

या $$\sigma_s^2 = 31.11$$

$$\mathrm{H}_0: \sigma_s^2 = \sigma_p^2$$

इस रिक्त परिकल्पना के परीक्षण के लिए, हमें x^2 के मान पर निम्नानुसार कार्य करना पडेगा।

$x^2 = \frac{\sigma_s^2}{\sigma_p^2}(\mathrm{n}-1)$

$= \frac{31.11}{20}(10-1)$

स्वतन्त्रता का अंश = (n – 1) or 10 -1 = 9

5% महत्व के स्तर पर, x^2 का तालिका मान 16.92 है और 1% महत्व के स्तर पर यह 9 d.f. के लिए 21.67 है और ये दोनों मान x^2 के गणना किये हुए मान जो कि 13.999 हे से ज्यादा है। इसलिए , हम शून्य परिकल्पना को स्वीकार करते हैं और यह निष्कर्ष निकालते हैं कि दिये हुए वितरण का विचलन 5% एवं 1% महत्व के स्तर पर 20 किलोग्राम लिया जा सकता है। दूसरे शब्दों में, नमूने को समग्र से 20 किलोग्राम वजन के साथ लिया जा सकता है।

उदाहरण 2- 15 बोतलों का एक नमूना यादृच्छिक तरीके से एक निश्चित समग्र से लिया जाता है। दिये हुए नमूने के माध्य से विचलन के वर्ग का योग 55 है। क्या इस नमूने का 6 के विचलन के साथ समग्र से लिया गया है?

हल- n = 15, $\sum(\mathrm{X_i} - \bar{\mathrm{X}}) = 55$, $\sigma_p^2 = 6$ दिया गया है-

शून्य परिकल्पना को निम्नानुसार लिया जा सकता है- भ्0रू $\mathrm{H_0}: \sigma_S^2 = \sigma_p^2$

सबसे पहले हमें नमूना विचलन की गणना

$\sigma_s^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{55}{15-1} = 3.93$ के अन्तर्गत करते हैं

अब हमें x^2 के मान की गणना $x^2 = \frac{\sigma_s^2}{\sigma_p^2}(\mathrm{n}-1)$ के अन्तर्गत करनी है।

$x^2 = \frac{\sigma_s^2}{\sigma_p^2}(\mathrm{n}-1) = \frac{3.93}{6}(15-1) = 9.17$

$(\mathrm{n}-1)$ के लिए अर्थात् (15-1) या 14 स्वतन्त्रता के लिए 5% महत्व के स्तर पर x^2 का परिकलित मान x^2 के तालिका मान से कम है , इसलिए शून्य परिकल्पना स्वीकार्य है , अर्थात् नमूना विचलन और समग्र विचलन के बीच कोई महत्वपूर्ण अंतर नहीं है और नमूने को 6 के विचलन के साथ समग्र से लिया गया है।

10.6 काई वर्ग एक गैर प्राचलिक परीक्षण के रूप में (Chi-Square as a Non-Parametric Test)

काई-वर्ग का उपयोग प्राचल के साथ ही गैर-प्राचल परीक्षण के रूप में किया जा सकता है। पहले के खंड में, आपने पढा है कि काई-वर्ग का प्रयोग नमूना विचलन की तुलना , समग्र विचलन के साथ की जा सकती है। इस स्थिति में, इसे प्राचल परीक्षण के रूप में प्रयोग किया जाता है क्योंकि यह समग्र प्राचल पर आधारित है। लेकिन यह ज्यादातर गैर-प्राचल परीक्षण के रूप में प्रयोग किया जाता है। वास्तव में, काई वर्ग सबसे महत्वपूर्ण और लोकप्रिय गैर प्राचल परीक्षण में से एक है जो उन सम्बन्ध में पूरा करना है। यह परीक्षण विशेष रूप से संज्ञात्मक आंकडों के सम्बन्ध में उपयोगी है। लेकिन इसे उच्च स्तरों (क्रमिक , अंतराल, अनुपात) के लिए भी

प्रयोग किया जा सकता है। गैर प्राचलन परीक्षण के रूप में , निम्न परिस्थितियों में काई वर्ग प्रयोग किया जा सकता है।

10.6.1 काई वर्ग स्वतंत्र परीक्षण के रूप में (Chi-square as a test of independence)

गुणों के समूह के क्षेत्र में काई वर्ग परीक्षण का प्रयोग बहुत उपयोगी है। विशेषताओं के समूह से , हमारा उद्देश्य यह पता लगाना है कि क्या दो विशेषताऐं स्वतन्त्र है या उनके बीच कोई सम्बन्ध है। इस प्रयोजन के लिए, हम शून्य परिकल्पना बनाते हैं कि दो विशेषताओं के बीच कोई समबन्ध नहीं है अर्थात दो विशेषताऐं स्वतन्त्र है। उसके बाद x^2 के मान की गणना की जाती है और इसकी तुलना x^2 के तालिका मान से की जाती है यदि x^2 का परिकलित मान उसके तालिका मान से कम या उसके बराबर है तो शून्य अवधारणा को स्वीकार किया जाता है और दो विशेषतओं को स्वतन्त्र माना जाता है , अन्यथा अगर परिकलित मान उसके तालिका मान से अधिक है तो शून्य परिकलित मान उसके तालिका मान से अधिक है तो शून्य परिकल्पना अस्वीकार कर दी जाती है और यह माना जाता है कि दो विशेषताओं के बीच सम्बन्ध है। इसलिए दो विशेषताओं के बीच परस्पर स्वतन्त्रता के परीक्षण के लिए काई वर्ग परीक्षण का उपयोग किया जाता है।
उदाहरण के लिए, मान लिजिए हमें यह जानने में रूचि है कि क्या एक नई दवा बुखार को नियंत्रित करने में प्रभावी है या नहीं। हम यह निर्णय लेने में काई वर्ग परीक्षण की सहायता ले सकते हैं। सबसे पहले हम शून्य अवधारणा लेंगें कि दो विशेषताओं अर्थात् नई दवा और बुखार का नियंत्रण स्वतन्त्र है जिसका अर्थ है कि नई दवा बहुखर को नियंत्रित करने में प्रभावी नहीं है। इस आधार पर , हम पहले अपेक्षित आवृत्तियों की गणना करते हैं और फिर x^2 का मान ज्ञात करते हैं। यदि x^2 का परिकलित मान स्वतन्त्रता के दिये गए अंश के लिए एक निश्चित स्तर पर तालिका मूलय से कम है, तो हम निष्कर्ष निकालते हैं कि शून्य परिकल्पना सत्य है जिसका अर्थ है कि दो विशेषताऐं स्वतन्त्र हैं या सम्बन्धित नहीं है। (अर्थात नई दवा बुखार को नियंत्रित करने में प्रभावी नहीं है) लेकिन यदि x^2 का परिकलित मान उसके तालिका मान से अधिक है तो हमारा अनुमान गलत हेागा जिसका अर्थ है कि दो विशेषताऐं सम्बन्धित है और सम्बन्ध संयोगवश नहीं है अपितु यह वास्तविकता में मौजूद है (अर्थात नई दवा बुखार को नियंत्रित करने में प्रभावी है और जैसा कि निर्धारित किया जा सकता है)। यहाँ पर घ्यान देने योग्य एक महत्वपूर्ण बात यह है x^2 सम्बन्ध के अंश की माप नहीं है या दो विशेषता के बीच सम्बन्ध के रूप में उपाय नहीं है , लेकिन यह केवल इस तरह के सहयोग के महत्व को पहचानने की एक तकनीक या दो विशेषताओं के संबंध को पहचानने की एक तकनीक है।

10.6.2 काई वर्ग Goodness of fit के रूप में (Chi-square as equation and factor)

काई-वर्ग परीक्षण का प्रयोग यह पता लगाने के लिए भी किया जाता है कि सैद्धान्तिक आवृत्ति वितरण के अनुरूप अपेक्षित वितरण में कितना अंतर है अर्थात Goodness of fit का परीक्षण किया जाता है। इसका अर्थ यह है कि x^2 का परीक्षण Goodness of fit के रूप में प्रयोग यह निर्धारित करने के लिए प्रयोग किया जाता है कि अवलेाकित आंकडे सैद्धान्तिक वितरण में कितने अच्छे हैं। कई बार , अपेक्षित आवृत्तियों को गणितीय तकनीकों जैसे द्विपद, सामान्य और पायसन वितरण आदि की सहायता से ज्ञात किया जाता है। कभी कभी यह जानना हमारे लिए महत्वपूर्ण हो जाता है कि कितनी वास्तविक आवृत्तियॉं , अपेक्षित आवृत्तियों से मिलती है। यह आवश्यकता मुख्य रूप से नमूना अध्ययन के मामले में उत्पन्न हेाती है। ऐसी परिस्थितियों में, हमें यह देखना होगा कि नमूना अध्ययन से प्राप्त वास्तविक आवृत्तियों का सैद्धांतिक या गणितीय वितरण से प्राप्त अपेक्षित आवृत्तियों के साथ मिलान है और उनके बीच का अंतर महत्वपूर्ण है या नहीं। इसे Goodness of fit कहा जाता है।

अस्वीकृति यह बताती है कि Goodness of fit खराब है। इस सम्बन्ध में Goodness of fit एक अन्य उल्लेखनीय बात यह है कि अपेक्षित और वास्तविक आवृत्तियों के वक्र एक समान होते है। दूसरी ओर , यदि Goodness of fit नहीं है तब दोनों वक्र एक समान नहीं होते है। जैसा कि आप जानते हैं कि Goodness of fit के रूप में , काई वर्ग मुख्य रूप से नमूने अध्ययनों में उपयोग किया जाता है। लेकिन अलग अलग तरह के नमूनों में काई वर्ग का उपयोग करते समय एक बात का ध्यान रखा जाना चाहिए कि Goodness of fit भी इसकी विश्वसनीयता पर प्रश्न चिन्ह उठाती है। **चाउ** के अनुसार ,"यह ध्यान में रखा जाना चाहिए खराब Goodness of fit के रूप में एक समान है। जब गणना किये गए काई वर्ग का मान शून्य के निकट है तो हमें इस सम्भावना पर संदेह करना चाहिए कि दो आवृत्तियों के आवंटन को उनके साथ सहमत होने के लिए बाध्य किया गया है और इसलिए हमारे प्रयोग के प्रारूप की पूर्ण रूप से जांच करनी चाहिए।"

10.6.3 काई वर्ग एकरूपता के रूप में (Chi-square as a test of uniformity)

विभिन्न नमूनों के बीच एकरूपता परीक्षण के लिए काई वर्ग परीक्षण का भी प्रयोग किया जाता है। एकरूपता के परीक्षण के लिए, यह पता लगाया जाता है कि क्या दो विभिन्न नमूने एक ही समग्र से लिये गये हैं या नहीं। दूसरे शब्दों में , काई वर्ग परीक्षण का प्रयोग दो अलग नमूनों के मानों के बीच अंतर के महत्व के परीक्षण करने के लिए किया जाता है। इसलिए , यह दो विशेषताओं के बीच स्वतन्त्र परीक्षण के समान है। लेकिन एक ही समय में , यह दो बिन्दुओं के स्वतन्त्र परीक्षण से भिन्न है। सबसे पहले स्वतन्त्र परीक्षण यह पता करने की कोशिश करता है कि एक विशेषता दूसरे से स्वतन्त्र है, और समरूपता का परीक्षण यह पता लगाने की कोशिश करता है कि यादृच्छिक नमूने समग्र से लिये गये है। दूसरा स्वतन्त्र परीक्षण एक नमूने का प्रयोग करता है जबकि समरूपता की जांच में दो या अधिक नमूनों का प्रयोग होता है।

10.7 काई वर्ग परीक्षण के प्रयोग में चरण (Steps in using the Chi-square test)

x^2 परीक्षण प्रारम्भ करने के लिए निम्नलिखित चरणों का पलन किया जाता है-

- काई वर्ग x^2 परीक्षण की प्रक्रिया शून्य परिकलपना की अवधारणा के साथ प्रारम्भ होती है। यह माना जाता है कि अपेक्षित और वास्तविक आवृत्तियों के बीच कोई अंतर नहीं है। स्वभाविक रूप से , अपेक्षित और वास्तविक आवृत्तियों के बीच अंतर के लिए वैकल्पिक अनुमान भी शून्य परिकल्पना से तैयार किया जाता है। इसे निम्नानुसार व्यक्त किया जा सकता है।

$H_0: O_i = E_i$
$H_0: O_i \neq E_i$
जहॉं $O_i =$ अवलोकित आवृत्ति और
$E_i =$ अपेक्षित आवृत्ति

कुछ सांख्यिकीयविदों द्वारा अवलोकित और अपेक्षित आवृत्तियों को दर्शाने के लिए क्रमशः f_0 और f_e या O और E प्रतीक चिन्ह प्रयोग किये जाते हैं।

- x^2 के मान की गणना करने के लिए , हमारे पास वास्तविक और अपेक्षित आवृत्तियॉं हमारे पास पहले से उपलब्ध है। दूसरे चरण में , इन वास्तविक आवृत्तियों के आधार पर , परिस्थितियों के आधार पर अपेक्षित आवृत्तियों की गणना करते हैं। अपेक्षित आवृत्तियों को यह मानकर विकसित किया जाता है कि संबंधित सांख्यिकीय समग्र के लिए एक विशेष संभावना वितरण उचित है। सामान्यता, 2X2 किसी भी आकस्मिकता तालिका के लिए किसी भी कक्ष में अपेक्षित आवृत्ति को निम्नानुसार ज्ञात किया जाता है।

किसी भी कक्ष में अपेक्षित आवृत्ति =

(उस कक्ष की पंक्ति के लिए पंक्ति योग) X (उस कक्ष के स्तम्भ के लिए स्तम्भ योग)/ कुल योग

- अगले चरण में, अवलोकित और अपेक्षित आवृत्तियों के बीच अन्तर को देखते हैं, अर्थात (O-E)
- इसके पश्चात्, अन्तर का वर्ग करते हैं। अर्थात् $(O-E)^2$
- अगले चरण में, इस अंतर के वर्ग को इसकी अपेक्षित आवृत्तियों से विभाजित करते हैं, अर्थात $\frac{(O-E)^2}{\mathrm{E}}$। यह प्रत्येक कक्षा की आवृत्तियों के लिए दोहराया जाना चाहिए।
- उसके बाद $\frac{(O-E)^2}{\mathrm{E}}$ का कुल योग ज्ञात करते हैं। यही आवश्यक x^2 मान है। इसे निम्नवत व्यक्त किया जा सकता है।$x^2 = \sum \frac{(O-E)^2}{E}$
- उपर्युक्त विधि द्वारा x^2 के मान की गणना करने के बाद , शून्य परिकल्पना की स्वीकृति या अस्वीकृति के सम्बन्ध में निर्णय लेने के लिए इसकी तुलना तालिका मान से की जाती है, हमें दो पहलुओं के बारे में निणय लेना चाहिए- महत्व का स्तर और स्वतन्त्रता का अंश।
- महत्व के स्तर का अर्थ है कि यादृच्छिक नमूने के उतार चढाव के कारण गलत होने वाले अधिकतम संभावित प्रतिशत। उदाहरण के लिए 1% महत्व का अर्थ है कि यादृच्छिक नमूने में उतार चढाव के कारण अधिकतम 1% संख्या गलत हो सकती है। इसी तरह, 5% महत्व का अर्थ है कि यादृच्छिक नमूने में उतार चढाव के कारण अधिकतम 5% संख्या गलत हो सकती है। x^2 के तालिका मान ज्ञात करने के लिए शोधकर्ता द्वारा उनकी सुविधा एवं अध्ययन के उद्देश्य के अनुसार महत्व के स्तर का निर्णय लेना होता है। अभ्यास में 5% महत्व का स्तर ज्यादा प्रचलित है।
- x^2 के तालिका मान को जानने एक और मापदंड तय किया जाता है वह स्वतन्त्रता का अंश है। स्वतन्त्रता के अंश का अर्थ चयन की स्वतन्त्रता की सीमा से है। किसी भी परिस्थिति में , जब विभिन्न आवृत्ति अंतराल दिये हों, पहले के तरह ही योग निकालते हैं, आवृत्ति बदलने के लिए उपलब्ध विकल्पों की संख्या को वांछित स्वतंन्त्रता के रूप में जाना जाता है। स्वतन्त्रता के वांछित अंश को कुल संख्या में से एक को घटाकर प्राप्त किया जाता है। इसे निम्न प्रकार से व्यक्त किया जा सकता है। d.f = n-1 यदि वर्ग आवृत्तियाँ को पंक्तियों और स्तम्भों में व्यवस्थित किया जाता है , तो स्वतन्त्रता के अंश जानने के लिए, एक पंक्ति की संख्या और स्तम्भ की संख्या दोनों से घटाया जाता है क्योंकि पहले की तरह पंक्तियों और स्तम्भों का योग ज्ञात करना आवष्यक होता है। इन स्थितियों में , स्वतन्त्रता के अंश का सूत्र निम्नवत दिया जाता है-

  d.f. = (C – 1)(R – 1)
  जहाँ C = स्तम्भों की संख्या
  R = पंक्तियों की संख्या
- महत्व के स्तर और स्वतन्त्रता का अंश तय करने के पश्चात x^2 के तालिका मान को विशिष्ट स्तर और अंश को देखा जाता है। काई वर्ग का तालिका मान अधिकतम सीमा तक , परिकलित मान है जो नमूनाकरण के उतार चढाव के कारण उत्पन्न हुआ माना जाता है।

अंतिम चरण में , x^2 के गणना मान और तालिका मान के बीच तुलना की जाती है। यदि परिकलित मान , तालिका मान से कम है तो शून्य परिकल्पना स्वीकार होती है जो इंगित करता है कि अपेक्षित ओर वास्तविक आवृत्तियों के बीच अंतर महत्वपूर्ण नहीं समझा जाता है। इसके विपरीत, यदि परिकलित मान, तालिका मान से ज्यादा है तो शून्य परिकल्पना अस्वीकार होती है औश्रअपेक्षित एवं वास्तविक आवृत्तियों के बीच अंतर महत्वपूर्ण समझा जाता है। इस सम्बन्ध में एक और उल्लेखनीय तथ्य यह है कि जब स्वतन्त्रता का अंश 30 से

अधिक हो तो $\sqrt{2\chi^2}$ का वितरण सामान्य वितरण के अनुरूप होता है , जिसमें $\sqrt{2\chi^2}$ वितरण का माध्य $\sqrt{2\, d.f.-1}$ है और मानक विचलन = 1। फल स्वरुप जब स्वतन्त्रता का अंश 30 से ज्यादा हो , $\left[\sqrt{2\chi^2} - \sqrt{2d.f.-1}\right]$ की मात्रा का प्रयोग इकाई के विचलन के साथ सामान्य चर के लिए किया जा सकता है, अर्थात् $Z\alpha = \left[\sqrt{2\chi^2} - \sqrt{2d.f.-1}\right]$

उदाहरण 3 - नीचे दी गई तालिका में हैजा के महामारी के दौरान प्राप्त आंकडों को दर्शाया गया है-

	आक्रमण	गैर आक्रमण	योग
टीकाकरण	31	469	500
गैर टीकाकरण	185	1315	1500
योग	216	1784	2000

हैजा के हमले को रोकने में टीकाकरण की प्रभावशीलता का परीक्षण करें। 5% महत्व के स्तर पर x^2 का मान 1 स्वतन्त्रता की श्रेणी के लिए 3.84 है।

हल - हम परिकल्पना को निम्नानुसार लेंगें-

शून्य परिकल्पना - टीकाकरण एवं हैजे के हमले की रोकथाम के बीच कोई सम्बन्ध नहीं है , अर्थात दो विशेषताऐं स्वतन्त्र हैं। x^2 के मान की गणना करने के लिए हमारे पास वासतविक एवं अपेक्षित आवृत्तियॉं होनी चाहिए। वास्तविक या अवलोकित आवृत्तियॉं प्रश्न में दी गई हैं (हैजे का आक्रमण A द्वारा प्रदर्शित है और टीकाकरण B द्वारा) जो निम्नवत है।

अवलोकित आवृत्ति

	आक्रमण (A)	**गैर आक्रमण (α)**	**Total**
टीकाकरण (B)	31	469	**500**
गैर टीकाकरण (β)	185	1315	**1500**
Total	**216**	**1784**	**2000**

इन वास्तविक आवृत्तियॉं के आधार पर, हम निम्नलिखित तरीके से अपेक्षित आवृत्तियॉं ज्ञात करते हैं।

$(AB) = \frac{500 \times 216}{2000} = 54 \quad \therefore (\alpha B) = (B) - (AB)$ or $500 - 54 = 446$

$(A\beta) = (A) - (AB)$ or $216 - 54 = 162$

$(\alpha\beta) = (\beta) - (A\beta)$ or $1500 - 162 = 1338$

इसे निम्नलिखित तरीके से प्रदर्शित किया जा सकता है:

अपेक्षित आवृत्ति

	आक्रमण (A)	**गैर आक्रमण (α)**	**Total**

टीकाकरण (B)	$\frac{500 \times 216}{2000}$ = 54 (AB)	500-54 = 446 (αB)	**500**
गैर टीकाकरण (β)	216 – 54 = 162 (Aβ)	1500-162 = 1338 ($\alpha\beta$)	**1500**
Total	**216**	**1784**	**2000**

$$x^2 = \sum \frac{(O-E)^2}{E}$$

इन मानों को प्रतिस्थापित करने पर-

$$x^2 = \frac{(31-54)^2}{54} + \frac{(469-446)^2}{446} + \frac{(185-162)^2}{162} + \frac{(1315-1338)^2}{1338}$$

$$= \frac{(-23)^2}{54} + \frac{(23)^2}{446} + \frac{(23)^2}{162} + \frac{(-23)^2}{1338}$$

$$= 9.8 + 1.19 + 3.27 + 0.40 = 14.66$$

इस प्रकार, x^2 का परिकलित मान 14.66 है।

स्वतन्त्रता के अंश की संख्या = (c – 1)(r – 1)

5% महत्व के स्तर पर 1 d.f. के साथ =1 x^2 का तालिका मान = 3.84 है, जबकि x^2 का परिकलित मान 14.66 है जो कि तालिका मान से बहुत अधिक है इसलिए शून्य परिकल्पना अस्वीकार है जिसका अर्थ है टीकाकरण एवं हैजे के हमले की रोकथाम स्वतन्त्र नहीं है। इस प्रकार , हम कह सकते हैं कि टीका करण हैजे की रोकथाम के लिए प्रभावशाली है।

x^2 की गणना के लिए वैकल्पिक विधि

हम 2X2 अकस्मिकता तालिका के मामले में x^2 के मान की गणना के लिए अन्य विधि का प्रयोग कर सकते है। हम मान सकते हैं कि अवलोकित आवृत्तियों को निम्नलिखित तरीकों से व्यवस्थित किया गया है-

a	b	(a+b)
c	d	(c+d)
(a+c)	(b+d)	N

तब x^2 की गणना निम्न सूत्र के आधार पर की जा सकती है-

$$x^2 = \frac{(ad - bc)^2 \times N}{(a+b)(c+d)(a+c)(b+d)}$$

इस सूत्र को इस प्रश्न में प्रयोग किया जा सकता है ताकि x^2 के मान की गणना के बिना अपेक्षित आवृत्तियों को भी ज्ञात किया जा सके। ऐसी स्थिति में , हमें केवल वास्तविक आवृत्तियों की आवश्यकता होती है जो प्रश्न में पहले से ही दिए गए है।

$$x^2 = \frac{[(31 \times 1315) - (185 \times 469)]^2 \times 2000}{(31+469)(185+1315)(31+185)(469+1315)}$$

$$= \frac{(40675-86765)^2 \times 2000}{(500)(1500)(216)(1784)}$$

$$= \frac{4232000000000}{28900800} = 14.6$$

उदाहरण 4- 1000 छात्रों पर जो डिस्कवरी चैनल देखते हैं और उनके बुद्धि स्तर में किये गए सर्वेक्षण में निम्नलिखित जानकारियाँ सामने आई है।

	डिस्कवरी देखने वाले	**गैर डिस्कवरी योग देखने वाले**	**योग**
उच्च (IQ)	415	185	**600**
निम्न (IQ)	65	335	**400**
योग	**480**	**520**	**1000**

5% महत्व के स्तर पर परीक्षण करें कि डिस्कवरी चैलन देख रहे छात्रों में उच्च IQ होता है।

हल- आइऐं हम शून्य परिकल्पना बनाते हैं कि डिस्कवरी चैनल देखने और IQ स्तर में कोई भी सम्बन्ध नहीं है।
अब वास्तविक आवृत्तियों के आधार पर अपेक्षित आवृत्तियों की गणना करते हैं

$\frac{480 \times 600}{1000} = 288$	600-288=312	600
480-288=192	400-192=208	400
480	**520**	**1000**

अवलोकित अपेक्षित

आवृत्ति (O)	**आवृत्ति (E)**	$(\mathbf{O-E})$	$(\mathbf{O-E})^2$	$\frac{(\mathbf{O-E})^2}{\mathbf{E}}$
415	288	127	16129	56
185	312	-127	16129	51.7
65	192	-127	16129	84
335	208	127	16129	77.54
			Total	**269.24**

$x^2 = \frac{(O-E)^2}{E} = 269.24$

स्वतन्त्रता के श्रेणीयों की संख्या = (c – 1)r(– 1) = (2 – 1)(2 – 1) = 1

1 d.f. के साथ 5% महत्व के स्तर पर x^2 का तालिका मान 3.841 है। चूंकि, परिकलित मान (269.24) तालिका मान (3.841) से बहुत अधिक है। इसलिए , शून्य परिकल्पना अस्वीकृत की जाती है और हम यह निष्कर्ष निकालते हैं कि डिस्कवरी चैनल देखने वाले छात्रों का IQ ऊँचा है।

उदाहरण 5- एक पासे को 150 बार उछालने पर निम्नलिखित परिणाम आये

अंकों की संख्या	1	2	3	4	5	6
आवृत्ति	19	23	28	17	32	31

परिकल्पना का परीक्षण करें कि पासा निष्पक्ष है।

हल- पासे को उछालने पर हम यह परिकल्पना करते हैं कि अवलोकित और अपेक्षित आवृत्तियों के बीच कोई सम्बन्ध नहीं है , अर्थात् पासा निष्पक्ष है। एक पासे में 6 मुख होते हैं और प्रत्येक मुख के घटित होने की प्रायिकता एक समान है, इसलिए 150 उछालों के लिए प्रत्येक मुख की अपेक्षित आवृत्ति 150/6 = 25 होगी।

O	E	$(\mathbf{O-E})$	$(\mathbf{O-E})^2$	$\frac{(\mathbf{O-E})^2}{\mathbf{E}}$
19	25	-6	36	1.44
23	25	-2	4	0.16
28	25	3	9	0.36
17	25	-8	64	2.56
32	25	7	49	1.96
31	25	6	36	1.44
			Total	**7.92**

$$x^2 = \frac{(O-E)^2}{E} = 7.92$$

स्वतन्त्रता के श्रेणीयों की संख्या = n – 1 = 6 – 1 = 5

5 स्वतन्त्रता की श्रेणी के साथ , 5% महत्व के स्तर पर x^2 का तालिका मान 11.07 है। चूंकि x^2 का परिकलित मान (7.92) तालिका मान (11.07) से कम है। इसलिए शून्य परिकल्पना स्वीकार है। इस प्रकार हम निष्कर्ष निकाल सकते हैं कि पासा निष्पक्ष है।

उदाहरण 6- निम्नलिखित आकस्मिकता तालिका 300 लोगों के आंखों के रंग और बालों के रंग का विश्लेषण प्रदर्शित करती है। जांच के लिए x^2 परीक्षण का प्रयोग करें, क्या आंखों के रंग और बालों के रंग के बीच कोई सम्बन्ध है।

हल-

आंख का रंग		**बाल का रंग**		**योग**
	काला	**सुन्दर**	**भूरा**	
भूरा	30	10	40	**80**
नीला	40	20	40	**100**
धुंधला	50	30	40	**120**
योग	**120**	**60**	**120**	**300**

शून्य परिकल्पना- आंखों के रंग और बालों के रंग के बीच कोई सम्बन्ध नहीं है, अर्थात् दोनों स्वतन्त्र हैं। प्रश्न में दिये गए वास्तविक आवृत्तियों के आधार अपेक्षित आवृत्तियों की गणना करते हैं।

अपेक्षित आवृत्ति

आंख का रंग	बाल का रंग			योग
	काला	सुन्दर	भूरा	
भूरा	$\frac{80 \times 120}{300}$ = 32	$\frac{80 \times 60}{300}$ = 16	80 – (32 +16) = 32	**80**
नीला	$\frac{100 \times 120}{300}$ = 40	$\frac{100 \times 60}{300}$ = 20	100 – (40 + 20) = 40	**100**
धुंधला	120 – (32 + 40) = 48	60 – (16 +20) = 24	120 – (32 + 40) = 48	**120**
योग	**120**	**60**	**120**	**300**

O	E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
30	32	-2	4	0.125
10	16	-6	36	2.250
40	32	8	64	2.00
40	40	0	0	0
20	20	0	0	0
40	40	0	0	0
50	48	2	4	0.08
30	24	6	36	1.50
40	48	-8	64	1.33
			Total	**7.285**

$$x^2 = \frac{(O-E)^2}{E} = 7.285$$

स्वतन्त्रता के श्रेणीयों की संख्या $= (c-1)(r-1) = (3-1)(3-1) = 4$

4 स्वतन्त्रता की श्रेणी के लिए 5% महत्व के स्तर पर x^2 का तालिका मान 9.488 है। x^2 का परिकलित मान 7.285 है जो कि तालिका मान से कम है। इसलिए, शून्य परिकल्पना स्वीकार है। इसका अर्थ है कि आंखों के रंग एवं बालों के रंग के बीच कोई सम्बन्ध नहीं है।

उदाहरण 7- प्रत्येक 5 बच्चों वाले 320 परिवारों के एक सर्वेक्षण से निम्नलिखित वितरण का पता चला-

लडकों की संख्या	5	4	3	2	1	0
लडकियों की संख्या	0	1	2	3	4	5

परिवारों की संख्या	14	56	110	88	40	12

क्या हम परिणाम कल्पना के अनुरूप हैं कि पुरूष और महिला जन्म समान रूप से संभावित है।

हल- आइए हम शून्य परिकल्पना लेते हैं कि पुरूष और महिला का जन्म समान रूप से संभावित है , अर्थात् पुरूष और महिला के जन्म की संभावना के बीच कोई अंतर नहीं है।

इस प्रश्न में, अपेक्षित आवृत्तियों को जानने के लिए द्विपद वितरण प्रयोग किया जा सकता है क्योंकि यह द्विपद वितरण की सभी शर्तों को पूरा करता है।

पुरूष के जन्म की संभावना $P = \frac{1}{2}$

महिला के जन्म की संभावना $q = 1 - \frac{1}{2} = \frac{1}{2}$

यदि प्रत्येक 5 बच्चों के 320 परिवारों पर एक सर्वेक्षण किया जाता है, तो द्विपद विस्तार निम्नानुसार होगा-

$320(p+q)^5$

or $320\left(\frac{1}{2}+\frac{1}{2}\right)^5$

$= 320[p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5]$

$$= 320[(\frac{1}{2})^5 + 5\left(\frac{1}{2}\right)^4(\frac{1}{2}) + 10\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 + 10\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 + 5(\frac{1}{2})(\frac{1}{2})^4 + (\frac{1}{2})^5]$$

$$= 320\left(\frac{1}{32}\right) + 320(5)\left(\frac{1}{16}\right)\left(\frac{1}{2}\right) + 320(10)\left(\frac{1}{8}\right)\left(\frac{1}{4}\right) + 320(10)\left(\frac{1}{4}\right)\left(\frac{1}{8}\right) + 320(5)\left(\frac{1}{2}\right)\left(\frac{1}{16}\right) + 320(\frac{1}{32})$$

$$= 10 + 50 + 100 + 100 + 50 + 10$$

उपरोक्त द्विपदीय विस्तार की विभिन्न पदों से अपेक्षित आवृत्तियों का पता चलता है।

O	E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
14	10	4	16	1.60
56	50	6	36	0.72
110	100	10	100	1.00
88	100	-12	144	1.44
40	50	-10	100	2.00
12	10	2	4	0.40
			Total	**7.16**

$$x^2 = \frac{(O-E)^2}{E} = \mathbf{7.16}$$

स्वतन्त्रता के श्रेणीयों की संख्या = n – 1 = 6 – 1 = 5

5 स्वतन्त्रता के श्रेणी के साथ 5% महत्व के स्तर पर x^2 का तालिका मान 11.07 है। x^2 का परिकलित मान तालिका मान से कम है। इसलिए , शून्य परिकल्पना स्वीकार्य है और यह निष्कर्ष निकाला जा सकता है कि पुरूष और महिला जन्म एक समान रूप से संभावित है।

उदाहरण 8- 200 एमबीए के परीक्षा परिणामों का एक विश्लेषण किया गया था। यह पाया गया कि 46 छात्र असफल हुए हैं, 68 ने तृतीय श्रेणी प्राप्त की है , 62 ने द्वितीय श्रेणी और शेष ने प्रथम श्रेणी प्राप्त की है। क्या ये आंकडे सामान्य परीक्षा परिणाम के अनुरूप हैं जो क्रमशः विभिन्न श्रेणीयों के लिए 2:3:3:2 के अनुपात में है?

हल- आइए हम शून्य परिकल्पना बनाते हैं कि अवलोकित और अपेक्षित परिणामों में कोई अंतर नहीं है। अपेक्षित आवृत्तियों को 2:3:3:2 के सामान्य परीक्षा परिणाम अनुपातों के आधार पर ज्ञात किया जा सकता है। इन अनुपातों के आधार पर , विद्यार्थियों के असफल , तृतीय श्रेणी , द्वितीय श्रेणी एवं प्रथम श्रेणी प्राप्त की अपेक्षित आवृत्तियॉ क्रमश: $\frac{200 \times 2}{10} = 40$, $\frac{200 \times 3}{10} = 60$, $\frac{200 \times 3}{10} = 60$ and $\frac{200 \times 2}{10} = 40$ होना चाहिए।

O	E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
46	40	6	36	0.900
68	60	8	64	1.067
62	60	2	4	0.067
24	40	-16	256	6.400
			Total	**8.434**

$$x^2 = \frac{(O-E)^2}{E} = \mathbf{8.434}$$

स्वतन्त्रता के श्रेणी की संख्या = n – 1 = 4 – 1 = 3

3 स्वतन्त्रता की श्रेणी के साथ 5% महत्व के स्तर पर x^2 का तालिका मान 7.81 है। x^2 का परिकलित मान **8.434** है जो कि तालिका मान से अधिक है। इसलिए , शून्य परिकल्पना अस्वीकार है और हम यह निष्कर्ष निकालते हैं कि दिए गए परिणाम सामान्य परीक्षा परिणाम के अनुरूप नहीं है।

10.8 येट के सुधार (Yates's correction)

1934 में, येट ने एक (2 X 2) तालिका के सम्बन्ध में गणना की गई x^2 मान निरंतरता के लिए सुधार का सुझाव दिया है, विशेषकर जब कक्ष में छोटी आवृत्तियॉं होती हैं और x^2 केवल महत्व के स्तर पर होता है। x^2 विश्लेषण का उपयोग करते समय , यह महत्वपूर्ण है कि कक्ष में (कम से कम 5) कम से कम 80 प्रतिशत अपेक्षित या सैद्धान्तिक आवृत्तियॉं हों और किसी भी कक्ष में अपेक्षित आवृत्ति 5 से कम हो, तो 2 X 2 आकस्मिकता तालिका में 1 d.f. के लिए हम येट का सुधार का प्रयोग करते हैं। इस सुधार के अनुसार , अवलोकित आवृत्ति जो कि 5 से कम है में 0.5 तक वृद्धि की जाती है और दूसरी आवृत्तियों को भी व्यवस्थित किया जाता है (0.5 जोडकर और 0.5 घटाकर) कि कुल पंक्ति और कुल स्तम्भ एक समान हों।

आवृत्तियों का वर्गीकरण

यदि छोटी सैद्धान्तिक आवृत्तियॉं घटित होती है , तो आमतौर पर दो या दो से अधिक कक्षाएं एक साथ जोडकर इस समस्या को दूर करना सम्भव हो। दूसरे शब्दों में, 5 से कम सैद्धान्तिक आवृत्तियों को एक या अधिक कक्षों के मिश्रण में से लेकर एक वर्ग में किया जा सकता है, जो अवलोकित और अपेक्षित आवृत्तियों के बीच अंतर की गणना करने से पहले हो सकता है। पुर्नगठन के बाद आजादी के श्रेणी की संख्या को निर्धारित किया जायेगा।

येट का सुधार (2×2) आकस्मिकता तालिका में उपयोग किया जाता है। आवृत्तियों का वर्गीकरण n x n ($m > 2, n > 2$) आकस्मिक तालिकाओं में उपयोग किया जाता है जहॉं 5 से कम अपेक्षित अवृत्तियों को आसन्न आवृत्ति में जोडा जाता है।

उदाहरण 9- 50 छोटी सामान्य दुकानों के एक नमूने में निम्नलिखित जानकारी प्राप्त की गई थी। क्या यह कहा जा सकता है कि शहर की तुलना में गॉवों में अपेक्षाकृत अधिक महिला दुकानदार हैं ? x^2 परीक्षण का प्रयोग करें-

	दुकानें		
	शहर में	गॉवों में	योग
पुरूष दुकानदार	17	18	35
महिला दुकानदार	3	12	15
योग	20	30	50

1 d.f. के साथ 5% महत्व के स्तर पर x^2 का मान 3.841 है।

हल- शून्य परिकल्पना महिला दुकानदारों की संख्या शहर और गॉव में एक समान है। कक्ष आवृत्ति 3 में 0.5 जोडना और अन्य कक्ष आवृत्तियों को समायोजित करना है ताकि पंक्ति का योग समान ही रहे , हमारे पास निम्नलिखित आकस्मिकता तालिका है-

अवलोकित आवृत्ति तालिका

	A	α	योग
B	16.5	18.5	**35**
β	3.5	11.5	**15**
योग	**20**	**30**	**50**

अपेक्षित आवृत्ति तालिका

	A	α	योग
B	$\frac{35 \times 20}{50} = 14$	$\frac{35 \times 30}{50} = 21$	**35**
β	$\frac{15 \times 20}{50} = 6$	$\frac{15 \times 30}{50} = 9$	**15**
योग	**20**	**30**	**50**

O	E	$(O-E)$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
16.5	14	+ 2.5	6.25	0.45
18.5	21	- 2.5	6.25	0.30
3.5	6	- 2.5	6.25	1.04
11.5	9	+ 2.5	6.25	0.69
			Total	**2.48**

$$x^2 = \frac{(O-E)^2}{E} = \mathbf{2.48}$$

1 d.f. के साथ 5% महत्व के स्तर पर x^2 का तालिका मान 3.841 है। x^2 का परिकलित मान 2.48 है जो कि तालिका मान से कम है। इसलिए , शून्य परिकल्पना स्वीकार है। इसका अर्थ है कि महिला दुकानदारों की संख्या शहर और गॉव में एक समान है।

10.9 काई-वर्ग परीक्षण का महत्वपूर्ण अव लोकन (Important observations of the Chi-square test)

आपने उपरोक्त खण्डों में देखा है कि काई-वर्ग परीक्षण का प्रयोग कई क्षेत्रो में किया जाता है जैसे कि नमूना विचरण और समग्र विचरण के बीच तुलना करने के लिए , दो विशेषताओं के बीच स्वतन्त्रता या सम्बन्ध के लिए या विभिन्न नमूनों के बीच समरूपता को जानने के लिए। यह सैद्धान्तिक वितरण के सम्बन्ध में Goodness of fit के निर्णय के लिए भी प्रयोग किया जाता है। इस प्रकार , x^2 परीक्षण के महत्व के बारे में कोई संदेह नहीं हैं।

x^2 एक बहुत ही लोकप्रिय परीक्षण है और अपने गुणों के कारण अक्सर (प्रायः) प्रयोग किया जाता है।

x^2 परीक्षण की मुख्य विशेषता यह है कि यह मूल वितरण या इसके प्राचल के रूप में कोई अवधारणा नहीं होती है। चूंकि x^2 अवलोकित और अपेक्षित आवृत्तियों पर आधारित होता है , न कि प्राचलों में जैसे माध्य एवं मानक विचलन, इसलिए मूल वितरण के सन्दर्भ में इसे कोई अवधारणा बनाने की आवश्यकता नहीं होती है। मुक्त परीक्षण वितरण होने के नाते , इसका प्रयोग किसी भी प्रकार के समग्र वितरण में किया जा सकता है जो x^2 परीक्षण के दायरे को बढाता है। इस परीक्षण की लोकप्रियता का एक कारण यह है कि प्राचल परीक्षणों t परीक्षण, Z परीक्षण और f परीक्षण की तुलना में , x^2 परीक्षण की गणना और व्याख्या की प्रक्रिया आसान (सरल) है। फिर भी x^2 परीक्षण का एक अन्य लाभ इसका योगात्मक गुण है , जिसके कारण स्वतन्त्र से

सम्बन्धित नमूनों के परिणामों को जोडना सम्भव होता है। इन सभी गुणों के कारण , काई वर्ग परीक्षण का प्रयोग प्रायः व्यापारिक समस्याओं और सामाजिक विज्ञान के क्षेत्र में किया जाता है।

इन गुणों के बावजूद, काई वर्ग परीक्षण की निश्चित सीमाऐं या दोष भी है। जैसा कि आप जानते हैं, x^2 परीक्षण का प्रयोग करने के लिए निश्चित शर्तों को पूरा किया जाता है। इन शर्तों को पूर्ण करना ही इस परीक्षण की एक सबसे बडी सीमा है। इस सम्बन्ध में एक उल्लेखनीय बात यह है कि प्राचल परीक्षणों की तरह x^2 परीक्षण विश्वसनीय नहीं है। इस प्रकार, किसी स्थिति में यदि x^2 एवं प्राचल दोनों तरह के परीक्षण प्रयोग होते हों तो उस स्थिति में वरीयता प्राचल परीक्षणों को दी जानी चाहिए। इस प्रकार इस परीक्षण का प्रयोग केवल परिकल्पना के परीक्षण के लिए किया जा सकता है। यह आंकलन के लिए उपयुक्त नहीं हैं इस परीक्षण की दूसरी सीमा यह है कि, घटनाओं के घटित होने के साथ घटनाओं के घटित न होने के सम्बन्ध में आंकडे आवश्यक हैं। फिर भी काई वर्ग परीक्षण का एक और दोष यह है कि x^2 का मान परिकलित नहीं किया जा सकता है , यदि एक ही तालिका में समान या मिलान वाले समूहों की दोहराई गई माप दर्शायी जाती है। लेकिन इन सभी सीमाओं के पश्चात भी x^2 परीक्षण का महत्व या लोकप्रियता को कम नहीं किया जा सकता है। इसके महत्व या लोकप्रियता के बारे में कोई संदेह नहीं है लेकिन इसका सही अनुप्रयोग भी एक महत्वपूर्ण और कठिन कार्य है। x^2 परीक्षण का उपयोग करते समय हमेंशा याद रखना चाहिए कि परीक्षण केवल तभी प्रयोग किया जा सकता है जब नमूने के व्यक्तिगत अवलोकन स्वतन्त्र हों। इसका अर्थ है कि व्यक्तिगत पद या घटना या अवलोकन की घटना का विचाराधीन नमूना के घटित अवलोकन का कोई प्रभाव नहीं पडता है। छोटे सैद्धान्तिक आवृत्तियों वाला नमूना विशेष तरीके से समझा जाना चाहिए। इस परीक्षण के अनुचित प्रयोग या दुरूप्रयोग से सम्बंधित गैर घटनाओं के आवृत्तियों के बारे में लापरवाही , परिकलित मान, अवलोकित मानों का योग और अपेक्षित मानों का योग और गलत गणना आदि संभावित कारण के गलत निर्धारण से विफलता हो सकती है। इस प्रकार शोधकर्ता को इन सभी सावधानियों को ध्यान में रखना चाहिए जब वे x^2 परीक्षण का प्रयोग कर रहे हैं और परिकल्पना के सम्बन्ध में निष्कर्ष निकालते हैं।

10.10 सारांश (Summary)

1900 में कार्य पीयरसन द्वारा काई वर्ग परीक्षण का प्रतिपादन किया गया था। x^2 परीक्षण प्राचल के साथ साथ एक गैर प्राचल परीक्षण है। लेकिन यह मुख्यतः गैर प्राचल परीक्षण के रूप में प्रयोग किया जाता है। x^2 परीक्षण से हमें, सैद्धान्तिक या अपेक्षित मान और अवलोकित या वास्तविक मान के बीच अंतर की सीमा निर्धारण में सहायता मिलती है। कुछ परिस्थितियाँ ऐसी होती हैं जहाँ x^2 परीक्षण प्रयोग करने के लिए कुछ पदों की संख्या कम से कम 50 होनी चाहिए। और किसी भी कक्षा में आवृत्ति 5 से कम नहीं होनी चाहिए। नमूना यादृच्छिक तरीके से चयनित होना चाहिए और आंकडों को पूर्ण रूप में दिखाया जाना चाहिए। यदि यादृच्छिक नमूना, सामान्य वितरण में माध्य μ और विचलन σ_p^2 के साथ लिया गया हो तो परीक्षण के लिए x^2 का प्रयोग किया जा सकता है।

गैर प्राचल परीक्षण के रूप में, x^2 का प्रयोग स्वतन्त्र परीक्षण के रूप में Goodness of fit और समरूपता के रूप में किया जा सकता है। x^2 स्वतन्त्र परीक्षण के रूप में स्थापित किया जा सकता है यदि दो या अधिक विशेषताऐं एक दूसरे से सम्बन्धित हैं या स्वतन्त्र हैं। x^2 Goodness of fit के रूप में अवलोकित आंकडों में सैद्धान्तिक वितरण के निर्धारण के लिए प्रयोग किया जाता है। x^2 समरूपता के एक परीक्षण के रूप में यह ज्ञात करने के लिए प्रयोग किया जाता है कि दो या अधिक यादृच्दिक रूप से चयनित स्वतन्त्र नमूने समान समग्र से लिये गऐ हैं या नहीं। सूत्र $\sum\frac{(O-E)^2}{E}$ का प्रयोग x^2 की गणना के लिए किया जाता है। x^2 के परिकलित मान की तुलना निश्चित स्वतंत्रता की श्रेणी और महत्व के स्तर पर तालिका मान के साथ की जाती है। यदि

परिकलित मान, तालिका मान से कम है तो शून्य परिकल्पना स्वीकार होती है अन्यथा अस्वीकार। स्वतन्त्रता की श्रेणी वर्गों की संख्या को संदर्भित करते हैं , जिसमें मान को स्वतन्त्र रूप से निर्धारित किये जा सकता है जो सीमाओं के बिना बडे (n – 1) या (c – 1)(n – 1) के प्रयोग से स्वतन्त्रता की श्रेणी के संख्या के निर्धारण के लिए होती है। यदि किसी कक्ष में आवृत्ति 5 से कम है तो येट का सुधार प्रयोग होता है जिसके द्वारा 0.5 को उस आवृत्ति और अन्य आवृत्तियों में इस तरह से समायोजित किया जाता है कि पंक्ति और स्तम्भ का योग एक समान रहें।

10.11 शब्दावली (Glossary)

- **स्वतन्त्रता की श्रेणी (Degree of freedom):** आंकडों के एक समूह में स्वतन्त्र बाधाओं की संख्या ।
- **महत्व का स्तर (Significance level):** यादृच्छिक नमूनाकरण के उतार चढाव के कारण संख्या में होने वाले अधिकतम संभावित प्रतिशत ।
- **Goodness of fit:** सिद्धान्तिक वितरण का अवलोकित वितरण से सुमेंल।

10.12 बोध प्रश्न (Comprehension Questions)

(अ) रिक्त स्थानों की पूर्ति

1. जब अवलोकित एवं अपेक्षित आवृत्तियॉं पूर्णतया सुमेंलित होती हैं , तो x^2 का मान ------ होगा।
2. प्रोफेसर -----ने x^2 परीक्षण का प्रतिपादन किया।
3. x^2 की मात्रा सिद्धान्त एवं ----के बीच विसंगति के परिमाण का वर्णन करता है।
4. x^2 वितरण एक -----प्रायिकता वितरण है।
5. येट का सुधार -----आकस्मिकता तालिका में प्रयोग होता है।

(ब) सही या गलत

1. x^2 की गणना के लिए सूत्र $\sum \frac{(O-E)^2}{O}$ है।
2. x^2 का परिकलित मान सकारात्मक या नकारात्मक है।
3. स्वतन्त्रता की श्रेणी अ द्वारा प्रदर्शित की जाती है।
4. (c-2)(r-2) सूत्र का प्रयोग आकस्मिकता तालिका में स्वतन्त्रता की श्रेणी निर्धारित करने के लिए किया जाता है।
5. यदि x^2 का परिकलित मान उसके तालिका मान से कम होता है तो शून्य परिकल्पना को अस्वीकार कर दिया जाता है।

10.13 बोध प्रश्नों के उत्तर (Answers to Comprehension Questions)

(अ) रिक्त स्थानों की पूर्ति

1. शून्य 2. कार्ल पियर्सन 3. अवलोकन 4. सतत् (निरंतर) और 5. 2ग2

(ब) सही या गलत

1. असत्य, 2. असत्य 3. सत्य 4. असत्य 5. असत्य

10.14 स्वपरख प्रश्न (Self-Assessment Questions)

1. x^2 परीक्षण के विभिन्न उपयोगों को बताऐं।
2. स्वतन्त्रता की श्रेणी से आप क्या समझते हैं घ्
3. किस स्थिति में येट का सुधार प्रयोग किया जाता है।
4. काई वर्ग परीक्षण क्या है घ् काई वर्ग परीक्षण के अनुप्रयोग में शामिल विभिन्न चरणों का उल्लेख करें।
5. काई वर्ग परीक्षण के अनुप्रयोगों का समीक्षात्मक विश्लेषण करें।
6. एक कक्षा के सात छात्रों का नमूना निम्नानुसार दिया गया है

क्रम सं0	1	2	3	4	5	6	7
अंक(प्रतिशत में)	52	50	56	61	45	54	39

यह निर्धारित करने के लिए x^2 परीक्षा का उपयोग करें कि क्या उपरोक्त नमूना एक छात्र समग्र से लिया गया है जिसका वि चरण 25 है। 5% महत्व के स्तर पर परीक्षण करें। स्वतन्त्रता की 6 श्रेणीयों के साथ x^2 का तालिका मान 14.1 है।

[12.64, H_0 स्वीकार]

7. 10 का एक नमूना यादृच्छिक तरीके से एक निश्चित समग्र से लिया गया है। दिये गये नमूने के माध्य से विचलनों का योग 50 है। इस परिकलपना का परीक्षण करें कि 5 प्रतिशत महत्व के स्तर पर समग्र का विचरण 5 है। x^2 का तालिका मान 9 स्वतन्त्रता की श्रेणीयों के साथ 16.92 है।

[10, H_0स्वीकार]

8. नीचे दिये गए आंकडे मलेरिया से हुए महामारी के दौरान लिये गए आंकडे दिखाता है

	ग्रसित	**गैर ग्रसित**	**योग**
टीकाकरण	120	240	360
गैर टीकाकरण	280	360	640
योग	400	600	1000

मलेरिया के आक्रमण को रोकने में टीकाकरण के प्रभाव का परीक्षण करें। स्वतन्त्रता की श्रेणी के साथ x^2 का तालिका मान 3.841 है।

[10.41, H_0 अस्वीकार]

9. एक निश्चित शहर में पुलिस दस्तावेज , जनवरी 2012 के पहले सप्ताह के दौरान हुए दुर्घटनाओं की संख्या से संबंधित निम्न आंकडे दिखाते हैं-

दिन	रवि	सोम	मंगल	बुध	गुरू	शुक्र	शनि	योग
दुर्घटनाओं की संख्या	20	12	13	17	19	20	18	119

आप को यह ज्ञात करना है कि दुर्घटनाऐं सप्ताह में एक समान रूप से हुई हैं। 5% महत्व के स्तर पर स्वतन्त्रता की 6 श्रेणीयों के साथ x^2 का मान 12.59 हैं।

[3.77, H_0 स्वीकार]

10. तालिकाओं के एक समूह में से 200 अंकों को यादृच्छिक रूप से चुना गया था। अंकों की आवृत्तियाँ थी-

अंक	0	1	2	3	4	5	6	7	8	9
आवृत्ति	18	19	23	21	16	25	22	20	21	15

अभिकल्पना की शुद्धता के आंकलन के लिए x^2 परीक्षण का प्रयोग करें, जो उन तालिकाओं में अंकों की समान संख्या में वितरित किए गए थे जिनसे इन्हें चयनित किया गया था। 5% महत्व के स्तर पर x^2 का मान 9 d.f. के लिए 16.919 है।

[4.3, H_0 स्वीकार]

11. नमूना अध्ययन के आधारपर आय वर्गों में कुछ लोगों को वर्गीकृत करने के लिए दो शोध अध्ययन किए गए। उनके परिणाम इस प्रकार थे-

अनुसंधान अध्ययन	**आय वर्ग**			
	गरीब	**मध्य**	**अमीर**	**योग**
अ	160	30	10	200
ब	140	120	40	300
योग	300	150	50	500

परीक्षण करें कि आय वर्गीकरण के लिए दोनों अध्ययन एक समान परिणाम देते हैं। 5% महत्व के स्तर पर 9 d.f के लिए x^2 का मान 16.919 है।

[55.54, H_0 अस्वीकार]

12. पांच सिक्कों को 3200 बार उछाला जाता है और हर बार प्रदर्शित होने वाले चिटों की संख्या का उल्लेख किया जाता है। अन्त में, निम्नलिखित परिणाम प्राप्त किए गए-

चिटों की संख्या	0	1	2	3	4	5
आवृत्ति	80	570	1100	900	500	50

यह निर्धारित करने के लिए कि क्या सिक्का निष्पक्ष है या नहीं , Goodness of fit के लिए काई वर्ग परीक्षण का प्रयोग करें। 5 d.f के लिए 5% महत्व के स्तर पर x^2 का मान 11.07 है।

[58.8, H_0 अस्वीकार]

13. एंथ्रेक्स के बकरियों के प्रतिरक्षण पर एक प्रयोग के बाद निम्नलिखित परिणाम प्राप्त किए गए-

	मरे हुए	**जीवित**	**कुल**
टीकाकरण	2	10	12
गैर टीकाकरण	6	6	12
योग	8	16	24

येट के सुधार से x^2 की गणना करें और वैक्सीन की प्रभावकारिता पर अपना अनुमान दें।

10.15 सन्दर्भ पुस्तकें (Reference Books)

- रॉय रामेंड, 'सांख्यिकीय के सिद्धान्त' प्रयाग पुस्तक भवन
- गुप्ता एस.पी. और गुप्ता एम.पी., 'व्यावसायिक सांख्यिकी' सुल्तान चंद एंड संस, नई दिल्ली
- शुक्ला एस.एम. और सहाय एस.पी. 'उन्नत सांख्यिकी' साहित्य भवन प्रकाशन, आगरा।

इकाई 11 साइन व माध्यिका परीक्षण

(Sign and Median Test)

11.1 प्रस्तावना (Introduction)

पिछली इकाई में, आप काई-वर्ग के बारे में जो कि सबसे प्रचलित गैर प्राचल परीक्षण है , का अध्ययन कर चुके हैं। काई वर्ग परीक्षण के अतिरिक्त हाल के वर्षो के दौरान कई अन्य गैर प्राचल परीक्षणों को विकसित किया गया है। जैसा कि आप जानते हैं कि गैर - प्राचल परीक्षण वितरण मुक्त परीक्षण होते हैं जिसमें कल्पना नहीं की जाती है कि कोई विशेष वितरण लागू हुआ है या कोई निश्चित मान समग्र के प्राचल से जुडा हुआ है। इस प्रकार इनका प्रयोग आसान होता है। यही कारण है कि जहाँ भी इन परीक्षणों में समतुल्य प्राचल परीक्षण की विधियाँ एकसमान होती हैं, गैर प्राचल संस्करणों को अधिक यथार्थवादी होने के लिए पसंद किया जाता है , इसलिए परिकल्पित समग्र को सामान्य या सामान्य के निकट होने की आवश्यकता नहीं होती है। कई गैर-प्राचल परीक्षणों जो हमारे लिए उपलब्ध है के मध्य उन वास्तविक जीवन परिस्थितियों की संख्याओं पर सोचेगें जो प्रायः प्रयोग किये जाते हैं। इस इकाई में आप प्रायः प्रयोग होने वाले एवं प्रचलित गैर प्राचल परीक्षणों जैसे साइन परीक्षण माध्यिका परीक्षण , फिशर इरविन परीक्षण , मन व्हाइटने U परीक्षण मैक नेयर परीक्षण , लिकोक्सन परीक्षण, एक नमूना रन्स परीक्षण आदि का अध्ययन करेगें।

11.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप-

- ✓ विभिन्न प्रकार के गैर प्राचल परीक्षण का वर्णन करेंगे।
- ✓ गैर प्राचल परीक्षणों की उपयुक्तता को समझ सकेंगे।
- ✓ गैर प्राचल परीक्षणों के लाभ एवं सीमाओं का वर्णन कर पाएंगे।

11.3 साइन परीक्षण (Sign test)

साइन परीक्षण सबसे पहले में से एक और सबसे सरलतम गैर प्राचल है। इसका नाम इस तथ्य से आता है कि यह अवलोकन के एक जोडे की दिशा (धनात्मक या ऋणात्मक) पर आधारित होता है न कि उनकी संख्यात्मक परिमाण पर। इस प्रकार इस परीक्षण में पूर्वानुमानित मानों और अवलोकित मानों के मध्य परिमाणों का अन्तर आवश्यक नहीं होता है, बल्कि दिशा क अन्तर का, अर्थात + या - चिन्ह प्रासंगिक होता है। यह दो प्रकार के तरीकों की प्रभावशीलता का मूल्यांकन करने के लिए उपयोगी होता है जिनके प्रभावों को मापा नहीं जा सकता है। लेकिन इसका केवल उत्कृष्ट/निकृष्ट या अच्छा/बुरा या/बेहतर/बेहतर नही के रूप में अनुमान लगाया जा सकता है। उदाहरण के लिए , छात्रों के एक समूह को दो विभिन्न प्रकार के शिक्षण विधियों का मूल्यांकन करने के लिए कहा जाता है। दो विधियों के मूल्यांकन करने के लिए कहा जाता है। दो विधियों के मूल्यांकन को चिन्हों में परिवर्तित किया जाता है। धनात्मक चिन्ह का अर्थ पहली विधि के लिए प्राथमिकता है , ऋणात्मक चिन्ह का अर्थ दूसरी विधि की प्राथमिकता है और शून्य बराबरी को प्रदर्शित करता है। अर्थात कोई प्राथमिकता नहीं है। हम + चिन्हों और - चिन्हों की गणना करते हैं और बराबर के मूलयांकनों को छोड देते हैं। इस परीक्षण का उपयोग करने के लिए आवश्यक एकमात्र आवश्यकता यह है कि समग्र वितरण लगभग माध्य μ_0 के सममित हो। साइन परीक्षण दो प्राकर का होता है

- एकल प्रतिदर्श साइन परीक्षण
- द्वि प्रतिदर्श साइन परीक्षण

11.3.1 एक प्रतिदर्श साइन परीक्षण (One Sample Sign Test)

एकल प्रतिदर्श साइन परीक्षण बहुत सरल गैर-प्राचल परीक्षण होता है, जब हम एक निरंतर सममित समग्र का प्रतिदर्श लेते हैं , तो उस स्थिति में प्रतिदर्श मान कम होने की संभावना $\frac{1}{2}$ होती है और अपेक्षा से

अधिक प्रतिदर्श मान प्राप्त करने की संभावना $\frac{1}{2}$ होती है। साइन परीक्षण आयोजित करने की प्रक्रिया निम्नानुसार होती है-

- सबसे पहले समग्र में से n आकार का यादृच्छिक प्रतिदर्श चयनित करते है और हम शून्य परिकलपना लेते हैं कि समग्र माध्य परिकल्पित माध्य के बराबर है अर्थात $\mu_0 : \mu = \mu_0$
- n प्रतिदर्श के प्रत्येक मानों को अवलोकित कर यह पता लगाया जाता है कि यह मान μ_0 से अधिक है या इससे कम है। μ_0 से ज्यादा प्रतिदर्श मानों को + चिन्ह निर्दिष्ट किया जाता है और इसे सफल के रूप में निर्दिष्ट करते हैं और जो μ_0 से कम होते हैं उन्हें - चिन्ह निर्दिष्ट करते हैं और उस असफल के रूप में निर्दिष्ट करते हैं।
- यदि कोई ऐसा एकांश है जिसका मान माध्य के बराबर है , तो उसे शून्य निर्दिष्ट करते हैं और इनके मानों को केवल त्याग दिया जाता है। यह प्रतिदर्श आकार को छोटा (कम) कर देता है।
- कुल चिन्हों की संख्या को n द्वारा प्रदर्शित करते हैं और प्रायः कम चिन्हों की संख्या को s द्वारा प्रदर्शित करते हैं।
- इसके उपरान्त 5 प्रतिशत महत्व के स्तर पर दो तरफा विकल्प के लिए गणना करते हैं और इसे 'k' प्रदर्शित करते हैं। k के मान की गणना के लिए निम्न सूत्र का प्रयोग करते हैं-
  $k = \frac{n-1}{2} - (0.98)\sqrt{n}$
- अन्तिम चरण में, s और k के मान के बीच तुलना की जाती है। यदि s का मान k के मान से अधिक होता है, तब शून्य परिकल्पना स्वीकार्य होती है। यदि फिर भी s का मान k के मान से कम या इसके समान होता है तब शून्य परिकलपना अस्वीकार होती है। एक प्रतिदर्श चिन्ह परीक्षण तब प्रयोग करते हैं जब प्रतिदर्श छोटा होता है , हम द्विपद प्रायिकताओं की सारणीयों का प्रयोग कर सकते हैं। सामान्य वितरण के z महत्वपूर्ण मानों का प्रयोग करते हुए , शून्य परिकलपना का परीक्षण किया जाता है। निम्नलिखित सूत्र का प्रयोग करते हैं-

$$z = \frac{p \pm 0.50 - {}^{n}/_{2}}{\sqrt{\frac{n}{4}}}$$

जहाँ p = धन चिन्हों की संख्या

n = कुल धन एवं ऋण चिन्हों की संख्या (निकाले हुए शून्य चिन्ह)

जब $p < \frac{n}{2}$ $(+0.5)$ प्रयोग होता है और जब $p > \frac{n}{2}$ (-0.5) सूत्र में प्रयोग होता है ।

यदि परिकल्पित z मान सारणी मान से अधिक होता है , शून्य परिकल्पना अस्वीकार्य होती है अन्यथा शून्य परिकल्पना स्वीकार्य होती है। जब प्रतिदर्श बडा होता है, हम द्विपद वितरण के लिए सामान्य सादृश्य का प्रयोग करते हैं। z के लिए निम्नलिखित सूत्र का प्रयोग करते हैं-

$z = \frac{s-np}{\sqrt{np\,(1-p)}}$ या $z = \frac{s-np}{\sqrt{npq}}$

जहाँ p = सफलता का अनुपात और

q = 1 – p

s = धनात्मक चिन्हों की संख्या

यदि परिकलित z मान, महत्वपूर्ण मान से कम होता है तब शून्य परिकल्पना स्वीकार्य होती है और यदि परिकलित z मान सारणी मान से अधिक होता है तब यह अस्वीकार्य होता है।

यहाँ इस बिन्दु में ध्यान देना आवश्यक है कि यह अपेक्षित है कि समग्र एक सतत् द्विपद समग्र नहीं है तब शून्य परिकलपना को माध्य के बदले माध्यिका के रूप में अभिव्यक्त किया जा सकता है।

उदाहरण 1- एक विक्रेता ने अपने क्षेत्र के बिक्री प्रबंधक के लिए 12 यात्राओं का भुगतान किया और कहा कि उसने कार्यालय के लिए क्रम से 10, 15, 20, 17, 11, 25, 30, 27, 36, 40, 5 और 26 मिनट का प्रतीक्षा करना पडता है। क्षेत्रीय प्रबंधक का दावा है कि उनसे मिलने वाले विक्रेता को 20 मिनट से अधिक समय तक प्रतीक्षा नहीं करनी पडती। साइन परीक्षण का प्रयोग करते हुए 0.05 महत्व के स्तर पर प्रमाणित करें कि क्षेत्रीय बिक्री प्रबन्धक का दावा सही है।

हल-

$\mu_0 : \mu = 20$ मिनट

$\mu_0 : \mu > 20$ मिनट

अब हम प्रतिदर्श मानो के आधार पर धनात्मक एवं ऋणात्मक चिन्ह देंगें कि ये मान 20 से अधिक है या कम है।

समय (मिनट में):	10	15	20	17	11	25	30	27	36	40	5	26
चिन्ह :	-	-	0	-	-	+	+	+	+	+	-	+

चिन्हों की संख्या या n = 11

प्रायः कम चिन्हों की कुल संख्या, अर्थात (-) या

s = 5

$$k = \frac{n-1}{2} - (0.98)\sqrt{n}$$

$$k = \frac{11-1}{2}(0.98)\sqrt{11}$$

$$k = 5 - 3.25 = 1.75$$

यदि s(5) का मान k(1.75) के मान से अधिक है, इसलिए शून्य परिकल्पना स्वीकार्य होती है, इसका अर्थ है कि क्षेत्रीय बिक्री प्रबंधक द्वारा प्रस्तुत दावा प्रमाणित है।

उदाहरण 2- मान लें कि शहर के क्लब में गोल्फ के चार छोरों को खेल रहै 11 पेशेवर खिलाडियों ने 280, 282, 290, 273, 283, 283, 275, 284, 282, 279 और 281 की संख्या बनाई। 5 प्रतिशत महत्व के स्तर पर साइन परीक्षण का प्रयोग करें कि वैकल्पिक परिकल्पना $\mathrm{H}_0 < 284$ के विपरीत शून्य परिकल्पना चार दोरों के लिए पेशेवर गोल्फरों का औसत $\mu_0 = 284$ है।

हल- $\mathrm{H}_0 : \mu_0 = 284$

$\mathrm{H}_1 : \mu_0 < 284$

अब हम प्रतिदर्श मानों के आधार पर धनात्मक एवं ऋणात्मक मान देंगें कि ये मान 284 से अधिक है या 284 से कम है।

अंक	चिन्ह
280	-
282	-
290	+

273	-
283	-
283	-
275	-
284	0
282	-
279	-
281	-

शून्य को अलग करते हुए कुल चिन्हों की संख्या, अर्थात् 'n' = 10

प्रायः कम चिन्हों की कुल संख्या (+) अर्थात s = 1 इस प्रश्न को विभिन्न वैकल्पिक विधियों द्वारा हल किया जा सकता है जो कि निम्नवत स्पष्ट है-

1. **प्रायोगिक विधि के अन्तर्गत-**

$$k = \frac{n-1}{2} - 0.98\sqrt{n}$$
$$= \frac{10-1}{2} - 0.98\sqrt{10}$$
$$= 4.5 - 3.099 = 1.4$$

चूंकि s (1) का मान k(1.4) के मान से कम है इसलिए शून्य परिकल्पना अस्वीकार है। इसका अर्थ है कि गोल्फ के चार दौरों में गोल्फरों का औसत 284 से कम है।

2. **द्विपद प्रायिकता विधि के अन्तर्गत-** जब प्रतिदर्श आकार छोटा होता है, साइन परीक्षण का प्रयोग करने के लिए हम द्विपद प्रायिकता वितरण का प्रयोग कर सकते है।

n =10, $p = \frac{1}{2}$, $q = \frac{1}{2}$

प्रायः कम चिन्हों की संख्या (+1) है। इसलिए, n =10 और $p = \frac{1}{2}$ के साथ एक या कम सफलता की संभावना की प्रक्रिया निम्नवत है।

$$P(1) = {}^{10}C_1 p^1 q^9 + {}^{10}C_0 p^0 q^{10}$$
$$= 10\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^9 + 1\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{10}$$
$$= 0.010 + 0.001 = 0.011$$

चूंकि p(s), अर्थात 0.011 मान 0.5 (अर्थात वांछित महत्व स्तर) से कम है इसलिए शून्य परिकलपना को अस्वीकार किया जाना चाहिए। इसका अर्थ है कि गोल्फ के चार दौरों के गोल्फरों का औसत 284 से कम है।

आप को याद रखना चाहिए कि शून्य परिकल्पना स्वीकार होती है यदि $p(s) > \propto$ और शून्य परिकल्पना अस्वीकार होती है यदि $p(s) < \propto$

3. **सामान्य वक्र विधि के अन्तर्गत-** $n = 10$, $p = \frac{1}{2}$, $q = \frac{1}{2}$

चिन्हों के आधार पर, अवलोकित सफलता अनुपात $\hat{p} = \frac{1}{10} = 0.1$

शून्य परिकल्पना $p = \frac{1}{2}$ मानते हुए अनुपात की मानक त्रुटि निम्नवत है-

$$\text{S.E.}_{\text{prop}} = \sqrt{\frac{p.q}{n}} = \sqrt{\frac{\frac{1}{2}\times\frac{1}{2}}{10}} = 0.1581$$

शून्य परिकल्पना के परीक्षण के लिए, अर्थात

$p = \frac{1}{2}$ वैकल्पिक परिकलपना के विपरीत

$p < \frac{1}{2}$ एक पुच्छीय परीक्षण (बायीं पुच्छीय) उपयुक्त है।

चूंकि महत्व स्तर 5 प्रतिशत है इसलिए स्वीकृत क्षेत्र 0.5-0.05 = 0.45 क्षेत्र है। सामान्य वक्र के अन्तर्गत क्षेत्र सारणी के प्रयोग द्वारा हम निर्धारित करते हैं कि 0.45 क्षेत्र के समतुल्य z मान के लिए क्षेत्र 1.64 है। अब, हम स्वीकृत क्षेत्र की सीमा को p से अनुपात के मानक त्रुटि को घटाके ज्ञात करेंगें (क्योंकि यह बायी पुच्छीय परीक्षण हैं, इस प्रकार $\text{S.E.}_{\text{prop}}$ कम होना चाहिए न कि जुडना

स्वीकृत क्षेत्र की सीमा =p-z. $\text{S.E.}_{\text{prop}}$

$$= \frac{1}{2} - (1.64)(0.1581)$$

= 0.5 - 0.2593 = 0.2407

जैसा कि सफलता का अवलोकित अनुपात केवल 0.1 है जो कि अस्वीकृत क्षेत्र में आता है (क्योकि यह 0.247 क्षेत्र के अन्तर्गत आता है) , इसलिए शून्य परिकल्पना अस्वीकृत होती है और परिणामस्वरूप वैकल्पिक परिकल्पना स्वीकृत होती है जिसका अर्थ है कि गोल्फ के चार दोरों के लिए पेशेवर गोल्फर का औसत 284 से कम है।

11.3.2 दो प्रतिदर्श साइन परीक्षण (Two Sample Sign Test)

साइन परीक्षण में समस्याओं के महत्वपूर्ण अनुप्रयोग होते हैं, जहाँ हम युग्मित आंकडों का वर्णन करते हैं यह दो सममित समग्रों से संबंधित n युग्मित अवलोकनों पर प्रयोग किया जा सकता है। इसलिए इसे युग्मित आंकडों के लिए साइन परीक्षण के रूप में जाना जाता है। परीक्षण आंकडें एवं निर्णय नियम एक प्रतिदर्श साइन परीक्षण के समान होते है। अन्तर केवल धन चिन्ह एवं ऋण चिन्ह कैसे निर्दिष्ट होते हैं , पर निर्भर करता है। प्रत्येक वस्तु या एकांश के लिए पहले अंक को दूसरे अंक के साथ तुलना की जाती है। यदि अन्तर धनात्मक होता है अर्थात पहला अंक दूसरे अंक से अधिक है , हम धन चिन्ह निर्दिष्ट करते हैं , यदि अंतर ऋणात्मक होता है अर्थात पहला अंक दूसरे से छोटा है तब ऋण चिन्ह निर्दिष्ट करते है। वस्तुओं के लिए , जहाँ दो मान या अंक एक समान हैं, उनको 0 निर्दिष्ट किया जाता है और इन युग्मों को अलग कर देते हैं। जिस परिस्थिति में दो प्रतिदर्श समान आकार में नहीं होते हैं तब बडे प्रतिदर्श के कुछ मानों को जिनका कोई युग्म नहीं होता है , अलग कर देते हैं। दो प्रतिदर्श साइन परीक्षण का मुख्यतः प्रयोग दो पुनरावृत्ति मापों के परीक्षण के लिए किया जाता है यदि दो प्रतिदर्श के माध्यों के बीच कोई अर्थपूर्ण अन्तर होता है।

उदाहरण 3- 22 मरीजों की स्पंद दर एक औषधि के प्रबन्ध के पहले और बाद मापी जाती है जो निम्नवत है-

मरीज	औषधि लेने से पहले स्पंद दर	औषधि लेने के बाद स्पंद दर	मरीज	औषधि लेने से पहले स्पंद दर	औषधि लेने के बाद स्पंद दर
1	73	75	12	70	72
2	71	73	13	70	69
3	69	70	14	67	70
4	68	69	15	74	75

5	74	73	16	72	74
6	72	73	17	71	71
7	73	73	18	73	75
8	71	72	19	71	69
9	70	68	20	70	72
10	69	74	21	73	75
11	73	70	22	74	75

5 प्रतिशत महत्व के स्तर पर साइन परीक्षण का प्रयोग करते हुए "औषधि का स्पंद दर में कोई प्रभाव नहीं है" परिकल्पना का परीक्षण करें।

हल-

मरीज	औषधि लेने से पहले स्पद दर	औषधि लेने के बाद स्पंद दर	चिन्ह
1	73	75	-
2	71	73	-
3	69	70	-
4	68	69	-
5	74	73	+
6	72	73	-
7	73	73	0
8	71	72	-
9	70	68	+
10	69	74	-
11	73	70	+
12	70	72	-
13	70	69	+
14	67	70	-
15	74	75	-
16	72	74	-
17	71	71	0
18	73	75	-
19	71	69	+
20	70	72	-
21	73	75	-
22	74	75	-

धन चिन्हों की कुल संख्या = 5

ऋण चिन्हों की कुल संख्या = 15

प्रतिदर्श आकार या n = 20

जो ‘0’ चिन्ह के साथ निर्दिष्ट है हम दो युग्मों को अलग करेंगें। इसलिए प्रतिदर्श आकार 22 – 2 = 20 होगा। र्

$$Z = \frac{S - np}{\sqrt{np(1-p)}}$$

$$= \frac{5 - \left(20 \times \frac{1}{2}\right)}{\sqrt{20 \times \frac{1}{2}\left(1 - \frac{1}{2}\right)}}$$

$$= \frac{5 - 10}{\sqrt{20 \times \frac{1}{2} \times \frac{1}{2}}} = \frac{-5}{\sqrt{5}}$$

$$= -2.23$$

चूंकि z (2.23) का परिकल्पित मान z के सारणी मान 5 प्रतिशत महत्व के स्तर पर कम है , इसलिए, शून्य परिकल्पना अस्वीकार्य है। इसका अर्थ है कि औषधि का स्पंद दर पर कोई प्रभाव नहीं है।

11.4 माध्यिका परीक्षण (Median test)

पिछली इकाई में, आपने काई वर्ग परीक्षण के बारे में जो कि संज्ञात्मक स्तर पर दो या अधिक स्वतन्त्र प्रतिदर्श मापों के लिए प्रयोग होता है का अध्ययन कर चुके हैं लेकिन यदि प्रतिदर्शों को क्रमवार स्तर पर मापा जाता है तब हम माध्यिका परीक्षण का प्रयोग कर सकते हैं। इसी तरह पिछले खण्ड में , अपने साइन परीक्षण के बारे में अध्ययन किया जो मुख्यतः उन परिस्थितियों में प्रयोग होता है जहाँ हम युग्म अवलोकनों के n समुच्चयों के साथ वर्णन करते हैं। लेकिन साइन परीक्षण के अनुप्रयोग के लिए एक आवश्यक शर्त यह है कि एक ही आकार के दो प्रतिदर्शों को लिया जाना चा हिए यदि प्रतिदर्श परिणामी आंकडा युग्मित परिणाम हो। व्यवहारिक परिस्थितियों में, हमें उन समस्याओ के साथ वर्णन करना पडता है जहाँ दो स्वतन्त्र प्रतिदर्शों का चयन आवश्यक होता है। विभिन्न समग्रों से समान आकार का प्रतिदर्श आवश्यक नहीं है। इन परिस्थितियों में यह प्रमाणित करना आवश्यक होता है कि समग्रों से लिये गये प्रतिदर्श उनके माध्य मानों से भिन्न है। माध्यों की तुलना द्वारा , माध्यिका परीक्षण का प्रयोग यह निर्धारित करने के लिए किया जाता है कि समगों से लिये गये प्रतिदर्शों की माध्यिका एकसमान है। यह दो या अधिक यादृच्छिक प्रतिदर्शों के माध्यिकाओं के मध्य अर्थपूर्ण अन्तर को निर्धारित करती हैं माध्यिका परीक्षण के संचालन की प्रक्रिया निम्नवत है।

- पहले चरण में, मिश्रित प्रतिदर्शों के माध्यिका की गणना की जाती है। इस चरण में , प्रतिदर्शों के चयन के पश्चात प्रत्येक अवलोकनों में वर्णित प्रतिदर्शों की मिश्रित किया जाता हैं , प्रतिदर्शों को परिमाणों के आधार पर व्यवस्थित करते हैं और माध्यिका ज्ञात की जाती है। माध्यिका मध्य का अवलोकन होता है जब n एक विषय संख्या होती है और जब n एक सम संख्या होती है तो यह दो मध्य के अवलोकनों का माध्य होती हैं
- अगले चरण में, पहले प्रतिदर्श n_1 के सभी अवलोकनों की तुलना माध्यिका मान के साथ की जाती है और इन्हें दो वर्गों में वर्गीकृत किया जाता है।
    1. माध्यिका से अधिक a_1 और
    2. माध्यिका से कम b_1

इसी तरह, माध्यिका के साथ तुलना करने के पश्चात दूसरे प्रतिदर्श n_2 के सभी अवलोकनों को दो वर्गों में वर्गीकृत किया जाता है। माध्यिका से अधिक a_2 और माध्यिका से कम b_2

- तदपश्चात परिणामी आंकडों को 2 x 2 आकस्मिकता तालिका के रूप में प्रदर्शित करते हैं।

माध्यिका परीक्षण के लिए प्रतिदर्श आंकडे का वर्गीकरण

	माध्यिका से अधिक	माध्यिका से कम	प्रतिदर्श आकार
प्रतिदर्श I	a_1	b_1	$n_1 = a_1 + b_1$
प्रतिदर्श II	a_2	b_2	$n_2 = a_2 + b_2$
कुल	$a_1 + a_2$	$b_1 + b_2$	n = $n_1 + n_2$

- जब प्रतिदर्श आंकडा का वर्गीकरण 2 x 2 आकस्मिकता तालिका में करते हैं , यदि कोई अवलोकन माध्यिका के मान के बराबर पाया जाता है तब या तो इसे प्रतिदर्श में से अपमार्जित किया जा सकता है या इसे माध्यिका वर्ग से ऊपर सम्मिलित किया जा सकता हैं। यदि प्रतिदर्श का आकार पर्याप्त रूप से बडा होता है तब माध्यिका मान के बराबर अवलोकनों को प्रतिदर्श में से अपमार्जित किया जा सकता है लेकिन यदि प्रतिदर्श का आकार छोटा होता है तब इसे माध्यिका वर्ग से ऊपर समिमलित किया जाना चाहिए।
- यदि प्रतिदर्श आकार बडा है (न्यूनतम 30 या 30 से अधिक) तब हम शून्य परिकल्पना के परीक्षण के लिए काई वर्ग परीक्षण का प्रयोग करेंगें। इस परिस्थिति में, 2 x 2 आकस्मिकता तालिका के आधार पर अपेक्षित आवृत्तियाँ ज्ञात करने के पश्चात हम x^2 की गणना करेंगें और इसकी तालिका मान के साथ तुलना पूर्व निर्धारित महत्व के स्तर के साथ 1 d.f. में करेंगें। यदि परिकल्पित मान तालिका मान से कम होता है तो शून्य परिकलपना स्वीकार्य होती है और यदि परिकल्पित मान तालिका मान से अधिक होता है तो शून्य परिकल्पना अस्वीकार्य होती है।
- यदि प्रतिदर्श आकार छोटा होता है तब हम शून्य परिकल्पना के परीक्षण के लिए फिशर के यथार्थ प्रायिकता परीक्षण का प्रयोग करेंगें। इस परिस्थिति में हम आवृत्तियों के युग्म के लिए व्यवस्थित 2 x 2 आकस्मिकता तालिका से फिशर की यथार्थ प्रायिकता का निर्धारण निम्नलिखित विधि से हाइपर ज्यामितीय वितरण के प्रयोग द्वारा करेंगें।

$$P_{(a_1\, a_2)} = \frac{\left({n_1}_{C_{a_1}}\right)\left({n_2}_{C_{a_2}}\right)}{\left(n_1 + n_{2\; C_{\,a_1 + a_2}}\right)}$$

शून्य परिकल्पना स्वीकार्य है या अस्वीकार्य का निर्णय p के परिकल्पित मान के साथ महत्व के स्तर α की तुलना द्वारा लिया जाता है। यदि p का परिकल्पित मान महत्व के स्तर की तुलना में कम होता है तो शून्य परिकल्पना अस्वीकार्य होता है और यदि p का मान महत्व के स्तर की तुलना में अधिक होता है तो शून्य परिकल्पना स्वीकार्य होती है।

उदाहरण 4- टीवी स्वरसमंजक के s_1 और s_2 के रूप में चिन्हित दो बडे नौवहन एक आयातक द्वारा प्राप्त किये जाते हैं। वह दो नौवहन से दो प्रतिदर्शों को जिसमें 25 स्वरसमंजक शामिल हैं का चयन करता है। दोषपूर्ण टुकडों की संख्या को जांचने के लिए , वह दोषपूर्ण स्वरसमंजक की संख्या के लिए निम्नलिखित आंकडे प्रदान करता है-

प्रतिदर्श I (s_1)	:	2	1	0	4	2	3	6	5	3	1
प्रतिदर्श II (s_2)	:	3	2	0	6	3	4	8	6	5	2

माध्यिका परीक्षण प्रयोग द्वारा 0.05 महत्व के स्तर पर प्रमाणित करें कि दो नौवहन में दोषपूर्ण वस्तुओं की संख्या माध्यिका संख्या के एकसमान है।

हल- शून्य परिकल्पना μ_0 दो नौवहन में दोषपूर्ण वस्तुओं की संख्या माध्यिका संख्या के एक समान है। वैकल्पिक परिकल्पना H_0 दो नौवहन में दोषपूर्ण वस्तुओं की संख्या माध्यिका संख्या के समान नहीं है। अब , हम

माध्यिका के मान की गणना करने के लिए पहले एवं दूसरे प्रतिदर्श के अवलोकनों को बढते हुए क्रम में व्यवस्थित करेगें।

S.No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Values	0	0	1	1	2	2	2	2	3	3	3	3	4	4	5	5	6	6	6	8

माध्यिका $= \frac{n+1}{2}$ वें पद का मान

$= \frac{20+1}{2}$ वें पद का मान

= 10.5 वें पद का मान

= 10 वें और 11 वें पद का मान

$= \frac{3+3}{2} = 3$

अब हम पहले एवं दूसरे प्रतिदर्श के मानों को माध्यिका 3 के मान के साथ तुलना करेंगें और उनको माध्यिका से अधिक और माध्यिका से कम वर्गों में वर्गीकृत करेगें। इसे निम्नलिखित 2 x 2 आकस्मिकता तालिका में अभिव्यक्त किया जा सकता है-

	M_d से अधिक	M_d से कम	प्रतिदर्श आकार
S_1	5	5	10
S_2	7	3	10
कुल	12	8	20

चूंकि प्रतिदर्श (20) का आकार छोटा है (30 से कम) है इस प्रकार हम फिशर का यथार्थ प्रायिकता परीक्षण प्रयोग करेंगें।

$$P_{(a_1\, a_2)} = \frac{\left(n_{1_{C_{a_1}}}\right)\left(n_{2_{C_{a_2}}}\right)}{\left(n_1 + n_{2\ C\ a_1 + a_2}\right)}$$

$$P_{(5,7)} = \frac{\left(10_{C_5}\right)\left(10_{C_7}\right)}{\left(20_{\ C_{12}}\right)} = \frac{252\ \times 120}{125970} = 0.24$$

चूँकि p(5,7) 0.24 है जो महत्व के स्वर (0.05) से अधिक है, इसलिए शून्य परिकलपना स्वीकार्य है , इसका अर्थ है कि दो पोतलदानों में दोषपूर्ण वस्तुओं की संख्या माध्यिका संख्या के समान है।

11.5 बिल्कसन मिलान युग्म परीक्षण (Wilkerson Matched Pair Test)

यह परीक्षण मिलान युग्मित आंकडों के लिए उपयुक्त होता है। अर्थात सम्बन्धित प्रतिदर्शों के साथ एक द्विभाजित चर और निरंतर चर के महत्व के परीक्षण करने के लिए किया जाता है। इस परीक्षण को बिल्कसन चिन्हित श्रेणी भी जाना जाता है क्योंकि चिन्ह परीक्षण में हम केवल मानों के मध्य अन्तर का चिन्हों के साथ वर्णन करते हैं जबकि यह परीक्षण न केवल दिशा परीक्षण करता है , अपितु मिलान युग्मों के मध्य के अन्तर के परिमाणों का भी परीक्षण करता है। यह परीक्षण विशेष रूप से मिलान युग्मों के पहले और बाद के प्रयोग प्रकार के लिए उपयुक्त होता है। यह मुख्यतया मिलान युग्मों की परिस्थिति में प्रयोग होता है जैसे एक अध्ययन ज हाँ पति और पत्नी में मिलान हाता है या जब हम दो एक ही तरह के मशीनों के परिणामों की तुलना करते हैं या पहले बाद के प्रयोग के परिणाम इस प्रकार यह एक महत्वपूर्ण गैर प्राचल परीक्षण जो मिलान युग्मों के बीच

अन्तर की दिशा एवं परिमाण का संज्ञान लेता है। दो एक ही तरह के के मशीनों के परिणामों की तुलना करते हैं या पहले बाद के प्रयोग के परिणाम। इस प्रकार यह एक महत्वपूर्ण गैर प्राचल परीक्षण जो मिलान युग्मों के बीच अन्तर की दिशा एवं परिमाण का संज्ञान लेता है। विल्कसन चिन्हित श्रेणी परीक्षण के संचालन की प्रक्रिया निम्नवत है-

- सबसे पहले, समस्या के गुण के आधार पर, विचार के अर्न्तगत दो श्रेणीयों के मध्य कोई अन्तर नहीं है , शून्य परिकल्पना बनाई जाती है।
- प्रत्येक युग्म को अंकों या मानों के बीच अन्तर पर कार्य किया जाता है।
- चिन्ह के बगैर छोटे से बडे के बीच श्रेणीयाँ निर्दिष्ट की जाती है। इसका अर्थ है कि श्रेणी 1 सबसे छोटे अन्तर को निर्दिष्ट करते हैं , 2 इसी क्रम में अगले और इसी तरह। जब श्रेणीयाँ निर्दिष्ट की जाती है , अन्तर के चिन्ह को संज्ञान में नहीं लिया जाता है।
- जब इस परीक्षण को प्रयोग करते है, दो प्रकार की बराबर की परिस्थितियाँ समझ में आ सकती है।
- पहली स्थिति दृष्टिगोचर होती है जब कुछ मिलान युग्मों के दो मान समान होते हैं अर्थात मानों के मध्य अन्तर शून्य होता है। इन युग्मों को गणना से अलग कर देते हैं। इस प्रकार उन युग्मों को जिनका अन्तर समान या शून्य होता है उन्हें बाद की गणनाओं से अलग कर देते हैं।
- दूसरी बराबर की स्थिति दृष्टिगोचर होती है जब दो या अधिक युग्मों का अन्तर एकसमान होता है। इस स्थिति में हम सम्बन्धित स्थिति सदिश के औसत श्रेणी पर कार्य करते हैं और उन्हें औसत श्रेणी निर्दिष्ट करते हैं। उदारहण के लिए मान लें 1, 2 और 3 श्रेणी निर्दिष्ट करने के बाद हमारे पास दो समान अन्तर के मान हैं। यदि उनके अन्तर एक समान नहीं होगें तो उनको श्रेणी 4 और 5 के साथ निर्दिष्ट किया जायेगा। इस प्रकार, हम दोनों को 4 और 5 वीं श्रेणी को औरसत निर्दिष्ट करेंगें, अर्थात ;4़5द्धध्2त्र4.5 अगले अन्तर मान को 6वीं श्रेणी निर्दिष्ट करेंगें।
- एक बार अन्तरों को श्रेणीबद्ध कर लिया, प्रत्येक को तब वास्तविक अंतर के चिन्ह के साथ निर्दिष्ट करते हैं।
- अगले चरण में , परीक्षण आंकडा ;जद्ध की गणना की जाती है जो दो योगों से छोटा घटित होता है अर्थात ऋणात्मक श्रेणीयों का योग और घनात्मक श्रेणीयों का योग।
- मिलान युग्मों की कुल संख्या, संज्ञान के पश्चात छोडे हुये युग्मों की संख्या 25 के बराबर या कम होती है, ज् के महत्वपूर्ण मान की तालिका का प्रयोग शून्य परिकल्पना के स्वीकृत या अस्वीकृत के प्रयोग के लिए की जाती है। शून्य परिकलपना स्वीकार हाती है यदि ज् आंकडे का परिकल्पित मान तालिका मान से अधिक होता है यदि ज् तालिका मान से कम या बराबर होता है शून्य परिकल्पना अस्वीकार्य होती है।
- जब मिलान युग्म 25 से अधिक होते हैं तब शून्य परिकल्पना के स्वीकृत या अस्वीकृत सम्बन्धित निर्णय के लिए z के मान की गणना के लिए विधि निम्नवत है- र्

त्र $Z = \frac{T - U_T}{\sigma_T}$

जहाँ U_T = माध्य

σ_T = मानक विचलन

T = छोटे चिन्हित वर्ग के श्रेणीयों का योग

माध्य या $U_T = \frac{n(n+1)}{4}$

मानक विचलन या $\sigma_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$

जहाँ- n = अलग किये गये युग्मों को छोडकर मिलान युग्मों की संख्या

उदाहरण 5- गणित के अभ्यास के लिए पॉचवी कक्ष के छात्रों को के दो सेट दिये गए थे। अभ्यास पुस्तिकाओं के नीचे दिये गये आंकडे क्रमशः कार्य पुस्तिका A और B से अभ्यास करने वाले छात्रों द्वारा प्राप्त अंकों को दर्शातें हैं। शोधकर्ता इस बात को जानने में रूचि रखते हैं कि क्या अंकों में स्पष्ट अंतर है जिसमें अभ्यास पुस्तिका के प्रकार के उपयोग के लिए बच्चों को जिम्मेदार ठहराया जा सकता है।

Student No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	73	43	47	53	58	47	52	58	38	61	56	56	34	55	65	75
B	51	41	43	41	47	32	24	58	43	53	52	57	44	57	40	68

हल- इस समस्या के लिए शून्य एवं वैकल्पिक परिकल्पना निम्नवत की जा सकती है।

$\mu_0 =$ विद्यार्थियों के दो वर्गो के अंकों के बीच कोई अंतर नहीं है।

$\mu_a =$ दो वर्गों के अंकों के बची अंतर।

विल्कसन मिलान युग्मों परीक्षण का प्रयोग करते हुए , हम T आंकडा परीक्षण के मान के लिए निम्न के अन्तर्गत कार्य करते हैं-

युग्म	कार्य पुस्तिका A	कार्य पुस्तिका B	अन्तर (di)	अन्तर की श्रेणी |di	चिन्ह (+ / -)	श्रेणीयॉं
1	73	51	+ 22	13	+13	....
2	43	41	+ 2	2.5	+2.5	....
3	47	43	+ 4	4.5	+4.5	....
4	53	41	+ 12	11	+11	....
5	58	47	+ 11	10	+10	....
6	47	32	+ 15	12	+12	....
7	52	24	+ 28	15	+15	....
8	58	58	0	....	....	....
9	38	43	-5	6	....	-6
10	61	53	+8	8	+8	....
11	56	52	+4	4.5	+4.5	....
12	56	57	-1	1	....	-1
13	34	44	-10	9	....	-9
14	55	57	-2	2.5	....	-2.5
15	65	40	+25	14	+14	....
16	75	68	+7	7	+7	....
				Total	**+101.5**	**-18.5**

हम युग्म संख्या 8 को अलग करेंगें क्योंकि इसका अन्तर मान शून्य है।

n = 16.1 =15 और +,- छोटे साइन वर्ग का योग =18.5

जब n =15 5% महत्व के स्तर पर T का तालिका मान 25 है (द्विपुच्छीय परीक्षण का प्रयोग करते हुए क्योंकि हमारी वैकल्पिक परिकल्पना यह है कि दो वर्गों के मध्य अन्तर है) T का परिकल्पित मान 18.5 है जो तालिका मान 25 से कम है। उसी रूप में हम शून्य परिकलपना को अस्वीकार करते हैं और निष्कर्ष निकालते हैं कि दो वर्गों के मध्य अंतर है।

11.6 बिल्कसन-मन-ब्हिटनी परीक्षण (U परीक्षण) Wilkerson-Mann-Whitney test (U test))

श्रेणी योग परीक्षणों के मय U परीक्षण सबसे अधिक लोकप्रिय परीक्षण है। इसे सामान्यतया विल्कसन मन व्हाइटने परीक्षण के रूप में भी जाना जाता है। पूर्ववर्ती परीक्षणों के समान , U परीक्षण भी दो स्वतन्त्र प्रतिदर्शों से प्राप्त आंकडों पर आधारित होता है। इसका प्रयोग यह निर्णारित करने में किया जाता है कि लिये गये प्रतिदर्श समान समग्र में से है या समान वितरण के दो विभिन्न समग्रों से है इसे साइन परीक्षण के ऊपर और फिशर ईरविन समझा जाता है। क्योंकि धन और ऋण चिन्ह के बदले इसमें श्रेणी सूचना का प्रयोग होता है। मन -व्हाइटने U और विल्कसन मिलान युग्म सामान्यतः समान होते हैं जिसमें उनको दो माध्यिकाओं के बीच तुलना करने की सलाह दो प्रतिदर्श समान समग्र से लिये गये हैं या नहीं। यदि आपके दोनों प्रतिदर्श पूर्ण रूप से एक दूसरे से स्वतन्त्र नहीं है और कुछ कारक एक समान है , अर्थात भौगोलिक परिस्थिति या पहले / बाद प्रतिपादन में विल्कसन मिलान युग्म परीक्षण का प्रयोग किया जा सकता है। यदि आपके पास दो प्रतिदर्श जो कि स्वतन्त्र हैं, आपको मन -व्हाइटने U परीक्षण का प्रयोग करना चाहिए। अत्यधिक परिवर्तनशील गैर प्राचल परीक्षण के रूप में, इसे उन परिस्थितियों में प्रयोग किया जाता है जहाँ प्रतिदर्श छोटे और समान आकार के नहीं होते हैं। यह परीक्षण बहुत सामान्य शर्तों के अन्तर्गत लागू होता है और केवल समग्र प्रतिदर्श सतत होने की आवश्यकता होती है।

U परीक्षण के संचालन की प्रक्रिया निम्नवत है-

- एकल समग्र में से या दो विभिन्न समग्रों से दो स्वतन्त्र प्रतिदर्श सामान्यतः विभिन्न आकार के लिये जाते है। छोटे आकार के प्रतिदर्श जिसमें n_1 अवलोकन निहित है और बडे आकार के प्रतिदर्श जिसमें n_2 अवलोकन समाहित होते है।
- शून्य और वैकल्पिक परिकल्पनाऐं ली जाती है। शून्य परिकल्पना व्यक्त करती है कि अंकों के दो समूहों में कोई अन्तर नहीं है जबकि वैकल्पिक परिकल्पना व्यक्त करती है कि अंकों के दो समूहों में व्यवस्थित ढंग से अंतर होता है। यह एकपुच्छीय या द्विपुच्छीय हो सकती है।
- दो प्रतिदर्शों को मिश्रित किया जाता है और सभी $n = (n_1 + n_2)$ अवलोकनों को छोटे से लेकर बडे तक बढते क्रम में व्यवस्थित करते हैं। तदपश्चात श्रेणीयाँ निर्दिष्ट की जाती है। प्रतिदर्शों को ध्यान दिये बिना मिश्रित प्रतिदर्शों n_1 और n_2 के मानों को सबसे कम से लेकर सबसे अधिक तक श्रेणीबद्ध किया जाता है, सबसे छोटे अंक को श्रेणी 1, अगले को श्रेणी 2 और इसी क्रम में प्रत्येक प्रतिदर्श की पहचान इंगित करता है। पुनरावृत्ति मानेां को अनेक प्रारम्भिक मानों के औसत के साथ श्रेणीबद्ध किया जाता है।
- तब पहले प्रतिदर्श के श्रेणी के योग को प्राप्त करते हैं और इसे R_1 के रूप में प्रदर्शित करते हैं और तब दूसरे प्रतिदर्श के श्रेणी के योग को प्राप्त करते हैं और इसे R_2 के रूप में प्रदर्शित करते हैं।
- अगले चरण में, हम आंकडे परीक्षण के मान पर कार्य करते हैं अर्थात U जो श्रेणीबद्ध अवलोकन के दो प्रतिदर्शों के अनतर के मध्य की माप के अन्तर्गत इस रूप में होती है- $U = n_1 \times n_2 + \frac{n_1(n_1+1))}{2} - R_1$

- तब U तालिका से n_1 और n_2 के लिए U का महत्वपूर्ण मान लिया जाता है। यदि U तालिका उपलब्ध नहीं है और प्रतिदर्श आकार (n_1 और $n_2 > 8$) बडा है तब U आंकडे को Z आंकडों में परिवर्तित किया जाता है। $n_1 + n_2$ अवलोकनों की शून्य परिकल्पना एक जैसी समग्र के लिए सत्य है , तब U आंकडे का प्रतिदर्श वितरण माध्य या $U = \frac{n_1 \times n_2}{2}$ और मानक विचलन या $\sigma_U = \sqrt{\frac{n_1.n_2(n_1+n_2+1)}{12}}$ के साथ होता है।

इसलिए Z आंकडे का निम्नलिखित सूत्र से गणना की जा सकती है।

$$Z = \frac{U - (n_1\ n_2)/2}{\sqrt{\frac{n_1\ .\ n_2(n_1+n_2+1)}{12}}}$$

- यदि Z का परिकल्पित मान, महत्वपूर्ण मान से छोटा या कम होता है , तब शून्य परिकल्पना स्वीकार्य होती है। दूसरी ओर, यदि Z का परिकलिपत मान महत्वपूर्ण मान से अधिक होता है, शून्य परिकल्पना अस्वीकार्य होती है।
- प्रतिदर्श आकार के छोटे होने की स्थिति में , अर्थात n_1 या $n_2 < 8$ तक हम वैकल्पिक विधि प्रयोग कर सकते हैं। U को w_s से न्यूनतम w_s का घटाकर ज्ञात किया जाता है, जहाँ w_s R_1 या R_2 से छोटा है और s प्रतिदर्श में तत्वों की संख्या छोटे योग के साथ है। तब हम इस परिकल्पित मान को U के महत्वपूर्ण मान की बिल्कसन तालिका से तुलना करते हैं और शून्य परिकलपना के स्वीकृत या अस्वीकृत के सम्बन्ध में निर्णय लेते हैं।

उदाहरण 6- एक परीक्षा में लडकों एवं लड़कियां द्वारा प्राप्त अंकों से सम्बिन्धत आंकडे निम्नवत दिये गये हैं-

लड़के (B)	44	56	32	36	52	48	40	44	56	52	36	32
लड़कियां (G)	40	48	44	36	44	24	32	16	36	44	28	30

U परीक्षण का प्रयोग 10 प्रतिशत महत्व के स्तर पर यह परीक्षण करने के लिए करें कि दोनों लड़के एवं लड़कियां समान माध्य के साथ एक समग्र से ली गई है।

हल:- H_0: लड़के एवं लड़कियां का प्रतिदर्श समान माध्य के साथ एक समग्र से लिया गया है।

H_1: लड़के एवं लड़कियां का प्रतिदर्श समान माध्य के साथ भिन्न समग्र से लिया गया है।

अब हम सभी अवलोकनों को बढते क्रम में व्यवस्थित करेंगे और उन्हें श्रेणी निर्दिष्ट करेंगे।

प्रतिदर्श मान	श्रेणी	लडकों की श्रेणी	लड़कियां की श्रेणी
16 (G)	1	-	1
24 (G)	2	-	2
28 (G)	3	-	3
30 (G)	4	-	4
32 (B)	6	6	-
32 (B)	6	6	-
32 (G)	6	-	6
36 (B)	9.5	9.5	-
36 (B)	9.5	9.5	-

36 (G)	9.5	-	9.5
36 (G)	9.5	-	9.5
40 (B)	12.5	12.5	-
40 (G)	12.5	-	12.5
44 (B)	16	16	-
44 (B)	16	16	-
44 (G)	16	-	16
44 (G)	16	-	16
44 (G)	16	-	16
48 (B)	19.5	19.5	-
48 (G)	19.5	-	19.5
52 (B)	21.5	21.5	-
52 (B)	21.5	21.5	-
56 (B)	23.5	23.5	-
56 (B)	23.5	23.5	-
Total		**R_1=185**	**R_2 = 115**

अब, इसके अर्न्तगत U आंकडे का मान ज्ञात करेंगें-

$$U = n_1 \times n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= 12 \times 12 + \frac{12\ (12+1)}{2} - 185$$

$$= 144 + 78 - 185 = 37$$

or $$U = n_1 \times n_2 + \frac{n_2\ (n_2+1)}{2} - R_2$$

$$= (12 \times 12) + \frac{12\ (12+1)}{2} - 115$$

$$= 144 + 78 - 115 = 107$$

चूंकि $n_1 = n_2$ n_1 = 12 और n_2 = 12 (दोनों ही 8 से अधिक है), इसलिए, U का प्रतिदर्श वितरण लगभग सामान्य वक्र की ओर होता है। निम्नलिखित सूत्र U को Z आकड़े में परिवर्तित किया जाता है-

$$Z = \frac{U - (n_1\ n_2)/2}{\sqrt{\frac{n_1\ .\ n_2(n_1+n_2+1)}{12}}}$$

$$\frac{37-(12 \times 12)/2}{\sqrt{\frac{12\times 12(12+12+1)}{12}}}$$

$$= \frac{37-72}{17.32} = -2.02$$

त्र ;37.72द्धध्17.32 त्र दृ 2.02

या $Z = \frac{107 - (12 \times 12)/2}{\sqrt{\frac{12\times12(12+12+1)}{12}}}$

$= \frac{107-72}{17.32}$ = + 2.02

चूंकि यह एक द्विपुच्छीय परीक्षण है , 10 प्रतिशत महत्व के स्तर पर्र का महत्वपूर्ण मार्न ं 1.64 हैर्। का परिकल्पित मान +2.02 महत्वपूर्ण मान से अधिक है। इसलिए शून्य परिकल्पना अस्वीकार्य होती है और हम निष्कर्ष निकालते हैं कि लड़के एवं लडकियां समान माध्य के साथ समग्र से लिये गए हैं।

उदाहरण 7- दो प्रतिदर्श पहली स्थिति में 90, 94, 36 और 44 मानों के साथ ओर दूसरी स्थिति में 53, 39, 6, 24 और 33 मानों के साथ दिये गए हैं। U परीक्षण का प्रयोग 10 प्रतिशत महत्व के स्तर पर करें कि समान माध्य के साथ प्रतिदर्श समग्र से लिये गये हैं।

हल- H_0: समान माध्य के साथ दो प्रतिदर्श समग्र से लिये गये हैं।

H_1: विभिन्न माध्यों के साथ दो प्रतिदर्श समग्र से लिये गए हैं। अब , हम अवलोकनों को बढते हुए क्रम में व्यवस्थित करेंगें और उन्हें श्रेणीबद्ध करेंगें।

प्रतिदर्श मान	श्रेणी	पहले प्रतिदर्श की श्रेणी	दूसरे प्रतिदर्श की श्रेणी
6 (II)	1	....	1
24 (II)	2	....	2
33 (II)	3	....	3
36 (I)	4	4	....
39 (II)	5	....	5
44 (I)	6	6	....
53 (II)	7	....	7
90 (I)	8	8	....
94 (I)	9	9	....
	कुल	**R_1 = 27**	**R_2 = 18**

जैसा कि दो प्रतिदर्शों में पदों की संख्या 8 से कम है $(n_1 = 4$ और $n_2 = 5)$ हम सामान्य वक्र करीब (निकट) तकनीकी प्रयोग नहीं कर सकते हैं। हम दिये हुए विल्कसन के (अयुग्मित) वितरण तालिका का प्रयोग करेंगें।

$w_s =$ दो योगों का छोटा = 18

S = छोटे योग के साथ प्रतिदर्श में पदों की संख्या = 5

$w_l =$ दो योगों का बडा = 27

L = बडे योग के साथ प्रतिदर्श में पदों की संख्या = 4

w_s का न्यूनतक मान =1+2+3+4+5 =15 (जब S = 5)

w_l का अधिकतम मान = 6+7+8+9=30 (जब L = 4)

$U = w_s - w_s$ न्यूनतम =18 -15 =3 या $U = w_l$ अधिकतम $w_l = 30 - 27 = 3$

U का प्रायिकता मान विल्कसन तालिका के अनुसार स्तम्भ 3 के कक्ष से, S = 5 और L = 4 टुकडे द्वारा 0.056 है। यह 3 से छोटी या छोटी से छोटी मान प्राप्त करने की आवश्यक प्रायिकता है हमें इसकी 10 प्रतिशत महत्व के स्तर के साथ तुलना करनी चाहिए चूँकि वैकल्पिक परिकल्पना यह है कि दो प्रतिदर्श विभिन्न माध्यों के साथ समग्र से लिये गये हैं, एक द्विपुच्छीय परीक्षण उपयुक्त है और 10 प्रतिशत महत्व स्तर का अर्थ 5 प्रतिशत बायॉं पुच्छीय और 5 प्रतिशत दायॉं पुच्छीय अनुसार है। दूसरे शब्दों में , हमें परिकल्पित प्रायिकता की 0.05 प्रायिकता के साथ तुलना करनी चाहिए। चूँकि परिकल्पित प्रायिकता (0.056) 0.05 से अधिक है , इसलिए, शून्य परिकल्पना स्वीकार्य है और हम निष्कर्ष निकालते हैं कि दो प्रतिदर्श समान माध्य से साथ समग्र से लिये गए हैं।

11.7 मैकनेयर परीक्षण (McNemar's test)

मैकनेयर परीक्षण महत्वपूर्ण गैर प्राचल परीक्षणों में से एक है जो प्रायः संज्ञात्मक आंकडों और दो सम्बन्धित प्रतिदर्शों से सम्बन्ध में प्रयोग होता है। इसका प्रयोग यह निर्धारित करने में किया जाता है कि दो सम्बन्धित प्रतिदर्शों के अनुपातों के मध्य अंतर प्रमाणित है। एकल प्रतिदर्श गिनती आंकडे पर आधारित, यह पूर्व निर्णय और बाद में निर्णय, प्रतिक्रिया परिणामों की तुलना करता है। दूसरे शब्दों में मैकनेमर परीक्षण स्थितियों में दो सम्बन्धित प्रतिदर्श के लिए प्रयोग होता है जहाँ लोगों की प्रवृत्ति का यदि कोई विचा रों में परिवर्तन के महत्व का परीक्षण पहले और बाद के प्रतिपादन के मूल्यांकन से किया जाता है।

इस परीक्षण के उपयोग के लिए प्रयोग को इस तरीके से अभिकल्पित किया जाता है कि तत्रं के बारे में विषयों की आरम्भ में उनके अनुकूल ओर प्रतिकूल दृष्टिकोणों को समान वर्ग में विभाजित किया जाता है। कुछ प्रतिपादन के पश्चात समान संख्या के विषयकों से दिये हुए तन्त्र के बारे में अपने दृष्टिकोण व्यक्त करने को कहा जाता है कि क्या वे इसके पक्ष में है या नहीं। समान विषयक पहले और बाद की प्रतिक्रियाओं के प्रबंधकीय प्रतिपादन से निम्नलिखित 2×2 आकस्मिकता तालिका के रूप में प्रदर्शित किया जा सकता है।

परिवर्तन के महत्व के परीक्षण के लिए प्रतिक्रिया तालिका

पहले प्रतिपादन	बाद का प्रतिपादन	
	पक्ष	विपक्ष
पक्ष	A	B
विपक्ष	C	D

इस तालिका में A प्रतिक्रियादाताओं के दृष्टिकोण हमेंशा धनात्मक होते हैं पहले और बाद के प्रतिपादन से कोई परिवर्तन नहीं होता है के रूप को प्रदर्शित करता है। इसी तरह D भी प्रतिक्रियादाताअें के दृष्टिकोण में पहले और बाद के प्रतिपादन में कोई अन्तर नहीं होता है और वे हमेंशा ऋणात्मक होते हैं को इंगित करता है। लेकिन B और C प्रतिपादन के प्रभाव के कारण प्रतिक्रियादाताओं के दृष्टिकोण में परिवर्तन को दर्शाता है। B प्रतिक्रियादाताओं की उस संख्या को प्रदर्शित करता है जो प्रतिपादन से पहले धनात्मक थे और बाद में ऋणात्मक थे। इसी तरह , C उन प्रतिक्रियादाताओं को इंगित करता है। जो प्रतिपादन से पहले ऋणात्मक थे और बाद में धनात्मक थे। इस प्रकार , B और C केवल दो प्रासंगिक निर्णय की आवृत्ति कक्ष है जो महत्व में परिवर्तन के होने या न होने को दर्शाता है। चूँकि (B+C) प्रतिक्रियादाताअें के दृष्टिकोण के कुल परिवर्तन को इंगित करती है, इस प्रकार शून्य परिकल्पना के अन्तर्गत अपेक्षा यह है कि (B+C) स्थिति में परिवर्तन एक दिशा में और समान अनुपात में दूसरी दिशा में होता है। मैकनेमर परीक्षण आंकडा एक x^2 रूपान्तर परीक्षण प्रारूप निम्नवत रूप में है-

$x^2 = \frac{(|B-C|-1)^2}{(B+C)}$ (with 1 d.f.)

असतत् वितरण से सतत् वितरण बनाने के लिए उपरोक्त वर्णित x^2 सूत्र को -1 से सुधार किया जाता है। अन्तिम चरण में , पूर्व निर्धारित महत्व के स्तर पर 1 स्वतन्त्रता के अंश के साथ x^2 के परिकल्पित मान की तालिका मान से तुलना की जाती है। यदि परिकल्पित मान तालिका मान से कम होता है महत्वपूर्ण परिवर्तन नहीं है की शून्य परिकलपना स्वीकृत होती है और यदि परिकल्पित मान तालिका मान से अधिक होता है तो शून्य परिकल्पना अस्वीकार होती है। मैकनेमर परीक्षण का x^2 के ऊपर लाभ यह है कि इस परीक्षण में मिलान युग्मों को संज्ञान में लिया जाता है जबकि x^2 परीक्षण में इन्हें संज्ञान में नहीं लेते हैं।

उदाहरण 8- एक कंपनी एक नई ब्रांडिंग रणनीति पर कार्य कर रही है, जो सोचती है कि वह अधिक प्रभावी है। इसकी स्वीकृत होने पर , प्रबन्धक जानना चाहता है कि नई रणनीति अपेक्षा से अधिक प्रभावी है। 50 प्रतिक्रियादाताओं का प्रतिदर्श नई रणनीति के अंगीकरण के दोनों पहले और बाद की प्रतिक्रिया जानने के लिए चयनित किया जाता है। प्रतिदर्श प्रतिक्रिया आंकडे का विश्लेषण निम्नलिखित परिणाम देता है-

अंगीकरण के पहले	**अंगीकरण के बाद**	
	धनात्मक	**ऋणात्मक**
धनात्मक	10	16
ऋणात्मक	12	11

5 प्रतिशत महत्व के स्तर पर प्रमाणित करें कि क्या नई ब्रांडिग रणनीति वास्तविकता में अधिक प्रभावशाली है।

हल- H_0: नई ब्रांडिग रणनीति अधिक प्रभावशाली नहीं है।

H_1: नई ब्रांडिग रणनीति अर्थपूर्ण ढंग से अधिक प्रभावशाली है।

$x^2 = \frac{(|B-C|-1)^2}{(B+C)} = \frac{(|16-12|-1)^2}{(16+12)} = \frac{9}{28} = 0.32$

1 d.f. के साथ 5 प्रतिशत महत्व के स्तर पर x^2 का तालिका मान 3.84 है। चूँकि परिकल्पित मान 0.32 तालिका मान 3.84 से कम है , इसलिए शून्य परिकल्पना स्वीकार होती है। इसका अर्थ है कि नई ब्रांडिग रणनीति अधिक प्रभावशाली है।

11.8 एक प्रतिदर्श रन्स परीक्षण (One Sample Runs Test)

एक प्रतिदर्श रन परीक्षण एक परीक्षण होता है जिसका उपयोग एक प्रतिदर्श की यादृच्छिकता के क्रम के आधार पर किया जाता है जिससे अवलोकन लिये जाते हैं। कई बाद हमें उन स्थितियों से समझौता करना पडता है जहाँ हमारा आंकडे पदो के चयन में कोई नियंत्रण नहीं होता है। इन परिस्थितयों में , निर्णय लेना बहुत कठिन होता है कि चयनित प्रतिदर्श यादृच्छिक है या नहीं। इन परिस्थितियों में , हमें प्रतिदर्श में यादृच्छिक परीक्षण के लिए रन परीक्षण प्रयोग करना चाहिए। यहाँ एक महत्वपूर्ण बिन्दु ध्यान में रखना चाहिए कि यादृच्छिक प्रतिचयन के लिय यह आवश्यक है लेकिन पर्याप्त परीक्षण नहीं है। यदि कल्पित यादृच्छिक प्रतिदर्श रन परीक्षण को असफल करता है , यह प्रकट करता है कि ये यादृच्छिक प्रतिचयन के साथ प्रतिदर्श विसंगत के क्रम में ये असामान्य है , न यादृच्छिक काल चक्र में है। एक रन परीक्षण समान अक्षरों या किसी अन्य प्रकार के प्रतीकों का अनुक्रम होता है जो विभिन्न अक्षरों या किसी भी अक्षर का नहीं द्वारा अनुसरण और पूर्ववर्तिता करता है। स्थिति की संख्या , पद या रन में प्रतीकों की संख्या को रन की लम्बाई के रूप में जाना जाता है।

यादृच्छिक आंकडे समूह में $(l+1)^{th}$ मान की प्रायिकता lth मान से अधिक या कम हो तो द्विपद वितरण का अनुसरण होता है जो रन परीक्षण का आधार बनता है।
उदाहरण के लिए, XX YYYYY XXX ZZZZ XX एक रन को प्रदर्शित करते हैं।
हम निम्न प्रकार से उपरनों में एक उप भोग को विभाजित करने के लिए रेखांकन के माध्यम से लगातार समान अक्षर समूह बना सकते हैं।
XX YYYYY XXX ZZZZ XX इस स्थिति में, हमारे पास 5 उप रन (r) है, x की $7(n_1)$ घटनाऐं हैं, y की $5(n_2)$ घटनाऐं हैं और z की $3(n_3)$ घटनाऐं हैं। इस प्रकार, रन की लम्बाई या कुल अवलोकनों (N) की संख्या $(7+5+3)=15$ है।

यदि किसी भी प्रकार के अव लोकनों के आकार 10 से कम होता है (अर्थात n_1 या n_2 या $n_3 < 10$) तब r के परिकल्पित मान की तुलना रन तालिका में से प्राप्त r के तालिका मान के साथ से की जाती है। लेकिन जब सभी प्रकार से अवलोकनों का आकार 10 या इससे अधिक होता है तब निम्नलिखित तरीके से r में आधारित सांख्यिकी की गणना करते हैं-

$$Z = \frac{r - \mu_r}{\sigma_r}$$

जहाँ: रू $\mu_r = \frac{2n_1 n_2}{n_1+n_2} + 1$ & $\sigma_r = \sqrt{\frac{2n_1 n_2\,(2n_1\, n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$.

एकल प्रतिदर्श रन्स परीक्षण केवल श्रेणी के सहजगुण के यादृच्छिकता सीमाबद्ध नहीं होता है। यहाँ तक कि माध्यिका से अधिक और माध्यिका से कम वर्गों या रन्स के अन्दर मानों के पृथक्करण द्वारा संख्यात्मक मानों में शामिल प्रतिदर्श हैतु प्रयोग किया जा सकता है। यह आर्थिक आंकडे से संबंधित प्रवृत्तियों या चक्रि य तरीके के पीरक्षण के लिए मुख्य रूप से उपयोगी होता है।

उदाहरण 9- 26 लोगों के प्रतिदर्श जिसमें 16 महिलाऐं (W) और 10 पुरूषों (M) का साक्षात्कार लिया गया। इनका साक्षात्कार निम्नलिखित क्रम में था।
M WWWW MMM WW M WWW MM WWW MMM WWWW
5 प्रतिशत के स्तर पर इस प्रतिदर्श के यादृच्छिक परीक्षण के लिए रन परीक्षण का प्रयोग करें।

हल- H_0: प्रतिदर्श यादृच्छिक है।
H_1: प्रतिदर्श यादृच्छिक नहीं है।
रन्स की संख्या = r = 10

महिला के घटित होने की संख्या = $n_1 = 4 + 2 + 3 + 3 + 4 = 16$

पुरूष के घटित होने की संख्या = $n_2 = 1 + 3 + 1 + 2 + 3 = 10$

कुल अवलोकनों की संख्या = N = 16 + 10 = 26

चूँकि दोनों n_1 और $n_2 \geq 10$, इसलिए, आंकडे की गणना निम्नवत करेंगें-

$$\mu_r = \frac{2n_1 n_2}{n_1+n_2} + 1 = \frac{2(16)(10)}{16+10} + 1 = 13.3$$

$$\sigma_r = \sqrt{\frac{2n_1 n_2\,(2n_1\, n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}} = \sqrt{\frac{2(16)(10)(2\times16\times10-16-10)}{(16+10)^2(16+10-1)}}$$

$$= \sqrt{\frac{94080}{16900}} = 2.359$$

$Z = \frac{r - \mu_r}{\sigma_r} = \frac{10 - 13.3}{2.359} = -1.398$

5 प्रतिशत महत्व के स्तर पर , द्विपुच्छीय परीक्षण के लिए Z का महत्वपूर्ण मा न $\pm$1.96 होता है। इसलिए परिकल्पित मान तालिका मान से छोटा है , इसलिए शून्य परिकल्पना स्वीकार्य है और हम निष्कर्ष निकाल सकते हैं कि प्रतिदर्श यादृच्छिक है।

11.9 गैर प्राचल परीक्षणों का समीक्षात्मक मूल्यांकन (Critical evaluation of non-parametric tests)

परिकल्पनाओं के परीक्षण के लिए हमारे पास दो प्रकार के परीक्षण होते हैं प्राचल परीक्षण और गैर प्राचल परीक्षण। हमें प्राचल परीक्षण के निश्चित रूप से चयन करना चाहिए यदि हम विश्वस्त होते हैं कि समग्र से प्रतिदर्शित किया हुआ आंकडा सामान्य वितरण का अनुसरण करता है। लेकिन कई बार , हमें उन परिस्थितियों के साथ समझौता करना पडता है जहाँ महत्व के मानक परीक्षणों के लिए विभिन्न अवधारणाऐं आवश्यक होती है जैसे समग्र सामान्य है, प्रतिदर्श स्वतन्त्र है, मानक विचलन ज्ञात है इत्यादि को पूरा नहीं किया जा सकता है तब हम गैर प्राचल विधियों का प्रयोग कर सकते हैं। निम्नलिखित तीन परिस्थितियों में हमें निश्चित रूप से गैर प्राचल परीक्षण का चयन करना चाहिए-

- परिणाम एक श्रेणी या एक अंक है और समग्र पूर्णतया सामान्य नहीं है।
- कुछ मान पैमाने से बाहर होते है अर्थात माप में बहुत बडे या बहुत छोटे। यहॉ तक की समग्र सामान्य है, प्राचल परीक्षण के साथ इन आंकडों का विश्लेषण असम्भव होता है चूँकि हम स भी मानों को नहीं जानते है। इन आंकडों के साथ गैर प्राचल परीक्षण का प्रयोग आसान होता है।
- आंकडे को क्रम संख्या पैमाने पर मापा जाता है। और समग्र का वितरण गौसियन तरीके से नहीं हुआ है।

गैर-प्राचल परीक्षणों में प्राचल परीक्षणो के ऊपर बहुत लाभ होते हैं। गैर -प्राचल परीक्षण का सबसे बढा लाभ इसकी अस्थिरता होती है। इन परीक्षणों का प्रयोग सभी प्रकार के आंकडो के लिए किया जा सकता है चाहै समग्र सामान्य है या असामान्य , परिमाणात्मक है या गुणात्मक है। यह श्रेणीबद्ध आंकडे के लिए सबसे अधिक उपयोगी होता है। जब हम इस तरह के आंकडों के साथ समझौता करते हैं जिन्हें प्रतिक्रियादा तों की पसंद के अनुसार श्रेणीबद्ध किया जा सकता है। लेकिन उनका सटीक परिमाणीकरण सम्भव नही होता है, तब हमारे पास एक ही विकल्प गैर -प्राचल परीक्षण होता है। इसी प्रकार , यह निर्णयात्मक या संज्ञात्मक आंकडे के साथ सबसे अच्छे विकल्प का भी समझौता है। कभी कभी हम इस तरह के आंकडे के साथ कार्य करते हैं जिसे विभिन्न समग्रों से सम्बन्धित प्रतिदर्शों के माध्यम से प्राप्त करते हैं। इस तरह की परिस्थितियों में , हमें कुछ अयथार्थवादी अवधारणाओं को प्राचल परीक्षणों के प्रयोग के लिए बनाना पडता है। लेकिन गैर प्राचल परीक्षणों के अनुप्रयोग से इन परिस्थितियों में कोई समस्या नहीं होती है। जब प्रतिदर्श आकार छोटा है या केवल कुछ अवलोकन उपलब्ध होते हैं तब भी केवल गैर-प्राचल परीक्षण का प्रयोग किया जाना चाहिए।

गैर-प्राचल परीक्षण के लोकप्रियता का मुख्य कारण उनकी प्राचल परीक्षण की तुलना में आसान गणना का होना है, चूँकि गैर प्राचल परीक्षण समझ में आसान गणना में सरल , सभी प्रकार के आंकडों के लिए अनुकूल और कम समय लेने वाले होते हैं, इसलिए ये अनुसंधानकर्ताओं द्वारा पसंद किये जाते हैं ।

यद्यपि गैर प्राचल परीक्षणों के बहुत लाभ है लेकिन इसकी कुछ हानियॉं भी होती है इस कारण से पहली पसंद हमेंशा प्राचल परीक्षणों को दी जाती है। गैर-प्राचल परीक्षण , प्राचल परीक्षणों से कम शक्तिशाली होते हैं क्योंकि वो बहुत सी अवधारणाओं पर आधारित नहीं होती है। अवधारणाओं का अभाव निर्णय लेने के क्षेत्र को सीमाबद्ध करता है। इस प्रकार , प्रतिदर्शी आंकडा सभी वांछित अवधारणाओं को पूर्ण करता है , या आंकडों को अन्तराल में या अनुपात पैमाने में मापा जाता है , तब इसे हमेंशा प्राचल परीक्षण का प्रयोग , गैर-प्राचल परीक्षण की तुलना में अच्छा समझा जाता है। इसी तरह, प्रतिदर्श का आकार बडा होता है तब गैर-

प्राचल परीक्षणों में शामिल गणनाऐ ज्यादा लम्बी होती है। इस प्रकार , बडे प्रतिदर्शों की स्थिति में, गैर-प्राचल परीक्षणों को नकार देना चाहिए। गैर-प्राचल परिक्षण के क्रियान्वयन के साथ दूसरी समस्या तालिका समीक्षात्मक मान की उपलब्धता होती है। अर्थपूर्ण निर्णयों के पहुँच के क्रम में, समीक्षात्मक मान आवश्यक होते हैं। यद्यपि, इनमें से कुछ मानों को प्रासंगिक तालिकाओं में संकलित नहीं किया गया होता है और प्रचलित तालिका हमेंशा ही उपलब्ध नहीं होती है, ये आसानी से उपलब्ध नहीं है।

11.10 सारांश (Summary)

गैर-प्राचल परीक्षण वे परीक्षण होते हैं जो समग्र के प्राचलें पर आधारित नहीं होते हैं, ये वितरण मुफ्त परीक्षण होते हैं । इस इकाई में आप कुछ लोकप्रिय और प्रयाः गैर-प्राचल परीक्षणों , काई वर्ग परीक्षण के अतिरिक्त के बारे में जिसे पिछले इकाई में अध्ययन किया गया है।

गैर प्राचल परीक्षणों में साइन परीक्षण सबसे महत्वपूर्ण है जिसका परीक्षण दिशा के अन्तरों के लिए प्रयोग होता है । यदि समग्र माध्य परिकलित माध्य के बराबर होता है। दो प्रकार के साइन परीक्षण होते है एक प्रतिदर्श साइन परीक्षण और द्वि प्रतिदर्श साइन परीक्षण ।

माध्यिका परीक्षण का प्रयोग यह निर्धारित करने में होता है कि समान माध्यिका के साथ समग्र में से प्रतिदर्श लिये गये है। यह दो या अधिक यादृच्छिक प्रतिदर्शों के माध्यिकाओं के बीच महत्वपूर्ण अन्तर को निर्धारित करता है।

दूसरा महत्वपूर्ण गैर प्राचल परीक्षण विल्कसन मिलान युग्म परीक्षण है जो पहले आर बाद के प्रयोग प्रकार का मिलान युग्मों के लिए उपयुक्त होता है। इस परीक्षण में, दिशाओं के साथ साथ अन्तर के परिमाणों को संज्ञान में लिया जाता है। तब भी दूसरा गैर-प्राचल परीक्षण विल्कसन मन व्हाइटले परीक्षण है जो U परीक्षण के रूप में जाना जाता है। यह दो स्वतन्त्र प्रतिदर्शों के बीच अलग के अंशों की माप करता है।
इसका प्रयोग यह निर्धारित करने में किया जाता हैं कि दो स्वतन्त्र प्रतिदर्श एक ही समग्र से लिये गये हैं या समान वितरण के दो या अधिक समग्रों से लिये गये हैं।

मैकनेयर परीक्षण का प्रयोग उन परिस्थितियों में जहाँ दो सम्बन्धित प्रतिदर्श स्थितियों में लोगों की प्रवृत्ति को परीक्षण के पहले और बाद के प्रतिपादन में विचार में यदि कोई अर्थपूर्ण परिवर्तन होता है वहाँ एकल प्रतिदर्श रन परीक्षण का प्रयोग प्रतिदर्श के यादृच्छिक परीक्षण के लिये किया जाता है। श्रेणी आंकडे के प्रयोज्य संज्ञात्मक और छोटे आकार प्रतिदर्श में गैर प्राचल परी क्षणों के अस्थिरता होने के लाभ होते हैं। इसकी कुछ हानियाँ भी हैं, जैसे ये परीक्षण बडे प्रतिदर्श के लिए कम शक्तिशाली एवं उपयुक्त नहीं है। इस प्रकार , प्राचल परीक्षण गैर-प्राचल परीक्षणों से अधिक पसंदनीय होते हैं।

11.11 शब्दावली (Glossary)

- **साइन परीक्षण (Sign Test)**- यह अवलोकन के एक जोडे की दिशा (धनात्मक या ऋणात्मक) पर आधारित होता है न कि उनकी संख्यात्मक परिमाण पर ।

11.12 बोध प्रश्न (Comprehension Question)

(अ) रिक्त स्थानों की पूर्ति

1. विल्कसन मिलान युग्म परीक्षण को विल्कसन -------परीक्षण के रूप में भी जाना जाता है।
2. एक रन में घटनाओं की संख्या को रन के ---- रूप में जाना जाता है ।
3. मैकनेयर परीक्षण आंकड़ा प्रयोग एक परिवर्तित ------परीक्षण प्रारूप है।
4. ------परीक्षण श्रेणी योग परीक्षणों के मध्य बहुत प्रसिद्ध है।

5. साइन परीक्षण में, प्रतिदर्श मान μ_0 की तुलना में अधिक ------निर्दिष्ट किये जाते हैं।

(ब) सत्य या असत्य

1. जब समग्र का वितरण स्पष्टतया सामान्य होता है , हमें निश्चित रूप से गैर प्राचल परीक्षणों का चयन करना चाहिए। (सत्य/असत्य)
2. यदि प्रतिदर्शों को क्रमवार मापा जाता है , तब हम माध्यिका परीक्षण का उपयोग कर सकते हैं। (सत्य/असत्य)
3. U परीक्षण में जब n_1 एवं $n_2 < 8$ हो तब U को z में परिवर्तित करना चाहिए। (सत्य/असत्य)
4. माध्यिका परीक्षण में, हम z के परिकलित मान की तुलना z के महत्वपूर्ण मान से शून्य परिकल्पना के परीक्षण के लिए करते हैं। (सत्य/असत्य)
5. गैर प्राचल परीक्षण छोटे प्रतिदर्शों के लिए अधिक उपयुक्त होते हैं। (सत्य/असत्य)

11.13 बोध प्रश्नों के उत्तर (Answers to comprehension questions)

(अ) रिक्त स्थानों की पूर्ति

1. साइन की श्रेणी (Signed Rank) 2. लम्बाई (Length) 3. काई - वर्ग (Chi - Square) 4. नए 5 धन (Plus)

(ब) सत्य या असत्य

1. असत्य 2. सत्य 3. असत्य 4. असत्य 5. सत्य

11.14 स्वपरख प्रश्न (Self-Test Question)

1. साइन परीक्षण के उपयोग का उद्देश्य क्या होता है घ्
2. माध्यिका परीक्षण में परीक्षण से सन्दर्भित शून्य परिकल्पना को लिखें घ्
3. गैर प्राचल परीक्षणों के तीन लाभों का वर्णन करें घ्
4. दो गैर प्राचल परीक्षणों के महत्व को समझाते हुए इनकी संक्षिप्त विवेचना करें।
5. गैर प्राचल परीक्षणों के लाभों एवं (नुकसान) हानियों का वर्णन करें।
6. 5 प्रतिशत महत्व के स्तर पर साइन परीक्षण का प्रयोग करते हुए विद्यालय के सभी विद्यार्थियों ने औसतन 80 प्रतिशत अंक प्राप्त किये हैं यह सत्य है या नहीं-

क्र. सं.	1	2	3	4	5	6	7	8	9	10	11	12
अंक	81	70	93	94	82	80	76	78	83	95	75	89

(μ_0 स्वीकार)

7. 30 दिनों में एक प्राचीन चट्टान पर दो पुरातात्विकों द्वारा खोदें गए कलाकृतियों की संख्या निम्नलिखित है-

X द्वारा	1	0	2	3	1	0	2	2	3	0	1	1	4	1	2	1	3	5	2	1	3	2	4	1	3	2	0	2	4	2
Y	0	0	1	0	2	0	0	1	1	2	0	1	2	1	1	0	2	2	6	0	2	3	0	2	1	0	1	0	1	0

द्वारा																													

1 प्रतिशत महत्व के स्तर पर साइन परीक्षण का प्रयोग करते हुए शून्य परिकल्पना का परीक्षण करें कि दोनों पुरातात्विक x एवं y कलाकृतियों ज्ञात करने में वैकल्पिक परिकल्पना कि x अच्छा है के विरूद्ध एक समान है। (μ_0 अस्वीकार)

8. एक शारीरिक प्रशिक्षण दावा करता है कि एक विशेष व्यायाम को जब न दिनों के लिए लगातार किया जाता है, शरीर का वजन न्यूनतम 3.5 किग्रा कम हो जाता है। 5 अधिक वनज वाली लड़कियां में 7 दिनों के लिए व्यायाम किया और उनके शरीर का वजन निम्नवत पाया गया था-

लड़कियां	व्यायाम से पहले वजन	व्यायाम के बाद वजन
1	70	66
2	72	70
3	75	72
4	71	66
5	78	72

साइन परीक्षण का प्रयोग करते हुए $\propto = 0.05$ पर दावे को प्रमाणित करें कि व्यायाम वजन को न्यूनतम 3.5 किग्रा करता है। (μ_0 स्वीकार)

9. एम0बी0ए0 विद्यार्थियों के दो वर्गो को लागत लेखांकन के समान पाठ्यक्रम को अध्यापन की दो भिन्न विधियों द्वारा बताया गया था। जिन्हें T_1 तथा T_2 द्वारा जानें। 6 विद्यार्थियों के एक प्रतिदर्श को दो वर्गों के प्रत्येक में से 20 अंकों का एक कक्ष परीक्षा दिया गया था। अंक निम्नवत पाये गये थे।

प्रतिदर्श (T_1)	15	10	11	12	18	15
प्रतिदर्श (T_2)	12	17	14	11	09	15

0.05 महत्व के स्तर पर माध्यिका का परीक्षण का प्रयोग करते हुए परीक्षण करें कि एम 0बी0ए0 विद्यार्थियों द्वारा प्राप्त अंक दोनों वर्गों में माध्यिका अंक के रूप में एक समान है। (μ_0 स्वीकार)

10. नीचे दिये गये आंकडों के मिलान युग्म दो मशीनों A एवं B के उत्पादन क्षमता से सम्बन्धित हैं-

मशीन A	142	140	144	144	142	146	149	150	142	148
मशीन B	138	136	147	139	143	141	143	145	136	136

विल्कोवन्सोन साइन श्रेणी का प्रयोग करते हुए शून्य परिकल्पना को 1 प्रतिशत महत्व के स्तर पर प्रमाणित करें कि दो मशीनों के मध्य उत्पादन क्षमता में कोई अन्तर नहीं हैं। (μ_0 अस्वीकार)

11. एक पहले एवं बाद के प्रयोग में 300 प्रतिक्रियादाताओं से प्राप्त प्रतिक्रियाओं को निम्नवत वर्गीकृत किया गया था-

पहले उपचार	बाद में उपचार	
	प्रतिकूल	अनुकूल
अनुकूल	60	90

प्रतिकूल	120	30

5 प्रतिशत महत्व के स्तर पर परीक्षण करें , मकनर परीक्षण का प्रयोग करते हुए प्रमाणित करें कि उपचार के पश्चात लोगों की राय में कोई महत्वपूर्ण अन्तर है। (μ_0 स्वीकार)

13. बहुत वर्ष पहले एक सडक के पास पास 30 आम के पेड स्थापित किये गये थे। एक शोधकर्ता ने पेडों को निरोगी (H) एवं रोगी (D) क्रम में निम्नवत पाया-

HH	DD	HHHHH	DDD	HHHH	DDDDD	HHHHHHHHH

1 प्रतिशत महत्व के स्तर पर रन्स परीक्षण का प्रयोग करते हुए इस प्रतिदर्श के लिए यादृच्छिकता का परीक्षण करें।

11.15 संदर्भ पुस्तकें (Reference books)

- होडा आर.पी., व्यवसाय एवं अर्थशास्त्र के लिए सांख्यिकीय, मैक मिल्लन व्यवसाय पुस्तकें, नई दिल्ली।
- राय रमनद्यू एवं बैनर्जी सुमोजित, अनुसंधान प्रणाली के मूल किताब महल इलाहाबाद।
- शुक्ला एस.एम. एण्ड शशि एस.पी., उन्नत सांख्यिकीय साहित्य भवन प्रकाशन आगरा।

इकाई 12 F -परीक्षण और प्रसरण का विश्लेषण
(F-Test and Analysis of Variance (ANOVA))

12.1 प्रस्तावना (Introduction)

पिछले दो इकाईयों में आपने काई-वर्ग परीक्षण और दूसरे गैर प्राचल परीक्षणो का अध्ययन किया है। आपको ज्ञात होना चाहिए कि प्राचल परीक्षण गैर प्राचल परीक्षणो की तुलना में अधिक प्रभावशाली होते हैं। इस प्रकार, आपको परिकल्पना परीक्षणों या निष्कर्ष निकालने के लिए प्राचल परीक्षणों में अधिक निभ्रर रहना चाहिए।

पिछले इकाईयों में आप पहले ही कुछ महत्वपूर्ण प्राचल परीक्षणों जैसे t परीक्षण z परीक्षण इत्यादि के बारे में अध्ययन कर चुके हैं। दो प्रतिदर्शों के माध्यों के बीच के महत्वपूर्ण अन्तर को या तो z परीक्षण या t परीक्षण द्वारा अपनिर्णीत किया जा सकता है। लेकिन जब हम एक ही समय में दो प्रतिदर्श माध्यों से अधिक अन्तर के महत्व का अध्ययन कर रहे होते हैं, ये दोनो परीक्षण उपयोगी नहीं होते हैं और हमें परिवर्तनशीलता के विश्लेषण का प्रयोग करना पडता है। दूसरा महत्वपूर्ण प्राचल परीक्षण F परीक्षा होता है जो स्वतन्त्र अनुमानों के लिए समग्र परिवर्तनशीलता के महत्व परीक्षण में प्रयोग किया जाता है।

इस इकाई में, आप परीक्षण और परिवर्तनशीलता के विश्लेषण के बारे में अध्ययन करेंगें जो दो प्रतिदर्शों के बीच से अधिक परिवर्तनशीलता के महत्व के निर्णय में अपकी सहायता करेंगें।

12.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के पश्चात आप-

- ✓ F- परीक्षण की अवधारणा एवं उसके अनुप्रयोगों को जानेंगे।
- ✓ ANOVA की अवधारणा को जानेंगे।
- ✓ प्रसारण के विश्लेषण की तकनीकी को समझेंगे।

12.3 F – परीक्षण (F-Test)

प्राचल परीक्षणों के क्षेत्र में परिवर्तनशीलता अनुपात परीक्षण या F परीक्षण एक महत्वपूर्ण परीक्षण होता है। इसे सामान्यतया F परीक्षण के रूप में जाना जाता है क्योंकि आर .ए.फिशर महान साख्यिकीयविद् ने पहली बाद परिवर्तनशीलता शब्द का प्रयोग और परीक्षण को विकसित किया था। F परीक्षण सामान्यतया उपयोगी होता है जब बहु प्रतिदर्श स्थितियॉ शामिल होती है और आंकडे के अन्तराल या अनुपात पैमाने में मापा जाता है। F परीक्षण का उद्देश्य यह निर्धारित करने में किया जाता है क्योंकि दो स्वतन्त्र अनुमानों की समग्र परिवर्तनशीलता में अर्थपूर्ण अन्तर है या क्या दो प्रतिदर्श सामान्य समग्रों से जिनकी परिवर्तनशीलता समान है से लिये जा सकते है। F परीक्षण एक बहुत उपयोगी परीक्षण है जिसे दो सामान्य समग्रों के समानता की परिवर्तनशीलता के परीक्षण में प्रयोग किया जा सकता है। दो स्वतन्त्र प्रतिदर्शों से अधिक के लिए यह परिवर्तनशीलता का विश्लेषण कर सकता है। इसका प्रयोग सह प्रसरण के विशलेषण के लिये किया जा सकता है। इस प्रकार, यह एक महत्वपूर्ण, लोकप्रिय और उपयोगी प्राचल परीक्षण है जिसे सभी क्षेत्रों में जैसे अर्थशास्त्र , व्यवसाय, शिक्षा कृषि इत्यादि में प्रयोग किया जा सकता है।

12.3.1 F - परीक्षण की अवधारणाऐं (Concepts of F-Test)

F परीक्षण निश्चित अवधारणाओं पर आधारित होता है जिसे इसके अनुप्रयोग के लिए पूर्ण होना चाहिए ये अवधारणाऐं निम्नवत हैं-

- पहली अवधारणा समग्र की सामान्य स्थिति होती है। इसका अर्थ है कि प्रत्येक वर्ग में मान सामान्य रूप से वितरित हुए हैं।

- दूसरी अवधारणा वर्गों की एकरूपता होती है। इसका अर्थ है कि प्रत्येक वर्ग के अर्न्तगत प्रसरण सभी वर्गों के लिए एकसमान होना चा हिए। यह अवधारणा मिश्रित या संघीय वर्गों के अर्न्तगत प्रसरणों के एकल भीतर वर्ग प्रसरण के स्रोत के क्रम के अंदर लिए आवश्यक होता है।
- तीसरी अवधारणा त्रुटि की स्वतन्त्रता होती है। इसका अर्थ है कि प्रत्येक मान के लिए अपने स्वयं के समूह के चारों ओर एक मान के विचरण का मध्य स्वतन्त्र होना चाहिए।
- अन्तिम अवधारणा या-च्छिकता की होती है। इसका अर्थ है कि प्रतिदर्श पदों को या-च्छिक तरीकों से समग्र से लिया जाना चाहिए।

12.3.2 F - परीक्षण की तकनीकें (Techniques of F-Test)

F परीक्षण दो प्रसरणों के अनुपात पर आधारित होता है। इसी कारण इसे , प्रसरण अनुपात परीक्षण कहा जाता है। दो प्रसरणों का अनुपात F वितरण का अनुसरण करता है जो उपरोक्त वर्णित अवधारणाओं पर आधारित होता है। इस परीक्षण में , सबसे पहले, एक शून्य परिकल्पना ली जाती है जिसका कथन यह होता है कि दो समग्रों के प्रसरण के बीच कोई अन्तर नहीं हैं। इस परिकल्पना के परीक्षण के लिए , हमें F (प्रसरणों का अनुपात) के मान के लिए कार्य करना पडता है। F की गणना निम्नवत की जाती है।

$F = \frac{S_1^2}{S_2^2}$

जहाँ $S_1^2 = \frac{\sum(X_1-\bar{X}_1)^2}{n_1-1}$ & $S_2^2 = \frac{\sum(X_2-\bar{X}_2)^2}{n_2-1}$

यहाँ यह बात ध्यान रखनी चाहिए कि अंशगणक हमेंशा ही ज्यादा प्रसरण का होता है। इसका अर्थ है कि S_1^2 हमेंशा की प्रसरण के बडे अनुमान का होता है। (अर्थात $S_1^2 > S_2^2$) इसे निम्नलिखित सूत्र के रूप में व्यक्त किया जा सकता है।

F = प्रसरण का बडा अनुमान / प्रसरण का छोटा अनुमान

v_1 के लिए $v_1 = n_1 - 1$ = बडे प्रसरण के प्रतिदर्श के लिए स्वतन्त्रता का अंश

$v_2 = n_2 - 1$ = छोटे प्रसरण के प्रतिदर्श के लिए स्वतन्त्रता का अंश

F के मान की गणना के बाद, इसे वांछित महत्व के स्तर (5% या 1%) पर v_1 और v_2 (बडे एवं छोटे प्रसरणों की स्वतन्त्रता के अंश के लिए) F के तालिका मान से तुलना करते हैं। यदि F अनुपात का परिकल्पित मान F के तालिका मान से कम होता है , तब F अनुपात अर्थपूर्ण नहीं होता है और शून्य परिकल्पना स्वीकार होती है। तब यह अनुमान लगाया जा सकता दोनों प्रतिदर्श समान प्रसरण के समग्र में से लिये गये हैं। दूसरी तरफ, यदि F का परिकल्पित मान F के तालिका मान से अधिक होता है, तब F अनुपात अर्थपूर्ण समझा जाता है और शून्य परिकल्पना अस्वीकार होती है।

उदाहरण 1- दो या-च्छिक नमूने दो सामान्य समगों में से लिये गये थे और उनके मान निम्नरूप में हैं-

A	66	67	75	76	82	84	88	90	92	------	------
B	64	66	74	78	82	85	87	92	93	95	97

5% महत्व के स्तर पर परीक्षण करें कि क्या दोनों समगों में समान प्रसरण है (संकेत: 5% स्तर पर $v_1 = 10$ और $v_2 = 8$ के लिएF = 3.36)

हल- हम शून्य परिकल्पना लेते हैं कि दोनों समग्रों में समान प्रसरण हैं

A (X_1)	($X_1-\overline{X_1}$)	$(X_1-\overline{X_1})^2$	B (X_2)	($X_2-\overline{X_2}$)	$(X_2-\overline{X_2})^2$
66	-14	196	64	-19	361
67	-13	169	66	-17	289
75	-5	25	74	-9	81
76	-4	16	78	-5	25
82	+2	4	82	-1	1
84	+4	16	85	+2	4
88	+8	64	87	+4	16
90	+10	100	92	+9	81
92	+12	144	93	+10	100
			95	+12	144
			97	+14	196
720	**0**	**734**	**913**	**0**	**1298**

$\bar{X}_1 = \frac{\sum X_1}{n_1} = \frac{720}{9} = 80$ $\bar{X}_2 = \frac{\sum X_2}{n_2} = \frac{913}{11} = 83$

$S_1^2 = \frac{\sum(X_1-\bar{X}_1)^2}{n_1-1} = \frac{734}{9-1} = 91.75$

$S_2^2 = \frac{\sum(X_2-\bar{X}_2)^2}{n_2-1} = \frac{1298}{11-1} = 129.8$

$F = \frac{S_1^2}{S_2^2} = \frac{129.8}{91.75} = 1.4$ (दूसरे प्रतिदर्श के प्रसरण को अंश गुणक बनाया गया क्योंकि दूसरे प्रतिदर्श का प्रसरण पहले की तुलता में अधिक है) 5% महत्व के स्तर पर $v_1 = 10$ और $v_2 = 8$ के लिए F का तालिका मान 3.36 है। चूँकि F (1.4) का परिकल्पित मान तालिका मान (3.36) से कम है , इसलिए शून्य परिकल्पना स्वीकार्य है। इस प्रकार, यह निष्कर्ष निकाला जा सकता है कि दो समग्रों में एक समान प्रसरण है।

उदाहरण 2- दो कृषि भूखंडों के समान क्षेत्र के सार्वजनिक 10 उपखण्डों के एक प्रतिदर्श में गेहूँ की उत्पादकता अवलोकित की गई। यह देखा गया था कि वर्ग विचलनों का योग माध्य से क्रमशः 0.92 और 0.26 था। 5% महत्व के स्तर पर परीक्षण करें कि क्या दो या-च्छिक समग्रों से लिए गऐ प्रतिदर्शों का प्रसरण समान है।

हल- हम शून्य परिकल्पना लेते है। कि दो समग्रों के प्रसरण के बीच कोई अन्तर नहीं हैं
दिया है-

$n_1 = 10$, $n_2 = 10$, $\sum(X_1 - \bar{X}_1)^2 = 0.92$, $\sum(X_2 - \bar{X}_2)^2 = 0.26$

$S_1^2 = \frac{\sum(X_1-\bar{X}_1)^2}{n_1-1} = \frac{0.92}{10-1} = 0.102$

$S_1^2 = \frac{\sum(X_2-\bar{X}_2)^2}{n_2-1} = \frac{0.26}{10-1} = 0.028$

$F = \frac{S_1^2}{S_2^2} = \frac{0.102}{0.028} = 3.64$

$v_1 = n_1 - 1 = 10 - 1 = 9$, $v_2 = n_2 - 1 = 10 - 1 = 9$

5% महत्व के स्तर पर $v_1 = 9$ और $v_2 = 9$ के लिए F का तालिका मान 3.18 है। चूँकि F (3.64) का परिकल्पित मान तालिका मान (3.18) से अधिक है। इसलिए, शून्य परिकल्पना अस्वीकार्य है। इसका अर्थ है कि समग्र से लिये गये प्रतिदर्शों में का प्रसरण भिन्न है।

12.4 प्रसरण का विश्लेषण (Analysis of Variance)

प्रसरण के विशलेषण को प्रायः ANOVA के रूप में उल्लेखित किया जाता है। विचरण में वर्गों की वजह से प्रसरण के पृथककरण जो कि अन्य वर्गों की वजह से होते हैं को सांख्यिकीय तकनीक के रूप में परिभाषित किया जाता है। अर्थशास्त्र , जीवविज्ञान, शिक्षा, समाजशास्त्र, मनोविज्ञान, व्यवसाय या उद्योग के क्षेत्रों और विभिन्न दूसरे शिक्षणों में यह शोध सम्बन्धी अत्यधिक उपयोगी तकनीक होती है। ANOVA तकनीक का आरम्भ में कृषि सम्बन्धी शोध में प्रयोग किया गया था और अब इसे सक्रिय रूप से प्रायोगिक प्रारूप पर आधारित शोधों के लिए प्रयोग किया जाता है , जैसे कि प्राकृतिक विज्ञान या सामाजिक विज्ञान। यह तकनीक तब प्रयोग की जाती है जब विविध प्रतिदर्श घटनाऐं शामिल होती है। ANOVA को यह परीक्षण करने के लिए कि क्या दो से अधिक परिमाणात्मक समग्रों के माध्य समान होते हैं , विशेषरूप से प्रारूपित किया जाता है। इसमें आंकडों का वर्गीकरण और क्रास वर्गीकरण शामिल होता है तब परीक्षण होता है यदि निर्दिष्ट वर्ग के माध्य में अर्थ पूर्ण अन्तर है।

डोनाल्ड एल हरनैट और जेम्स एल मर्फी के अनुसार “ANOVA का सार यह है कि आंकडों के समूह में प्रसरण की कुल मात्रा को दो हिस्सों में बांटा जाता है जैसे कि मात्रा जो संयोगवश सहजगुण के कारण हो सकती है और मात्रा जो कि निर्दिष्ट घटनाओं के सहजगुण के कारण हो सकती है।“ प्रतिदर्शों के बीच प्रसरण हो सकता है और प्रतिदर्श पदो के बीच भी हो सकता है। ANOVA में विश्लेषणात्मक उद्देश्यों के लिए प्रसरण विखंडन शामिल होते है। आप जानते है कि ज परीक्षण का प्रयोग जहाँ दो समग्र माध्य समान होते है के परीक्षण के लिए किया जाता है जबकि ANOVA का प्रयोग विविध समगों के माध्यों के बीच समानता के परीक्षण के लिए किया जाता है। इस प्रकार, ANOVA को ज परीक्षण के विस्तार के रूप में सुविचारित किया जा सकता है।

ANOVA एक तकनीक है जिसे विभिन्न क्षेत्रों में प्रयोग किया जा सकता है। उदाहरण के लिए , यह तकनीक हमें वर्णन करने में सहायता करती है कि बीजों की विभिन्न प्रजातियाॅ या रासायनिक उर्वरको या मिट्टीयो में अर्थपूर्ण अन्तर है जिसकी विजह से कृषि शोधों के क्षेत्र में नीति निर्णय तद्नुसार लिये जा सकते हैं इसी तरह, इस तकनीक के अनुप्रयोग के माध्य से , जानवर के विशेष वर्ग के लिए तैयार संभरण में अन्तर या निर्दिष्ट बिमारी के संसाधन के लिए विभिन्न प्रकार की औषधि प्रौद्योगिक में अन्तर का अध्यययन किया जा सकता है या अन्तर अर्थपूर्ण है या नहीं का निर्णय लिया जा सकात है। इसका प्रयोग व्यवसाय से सम्बन्धित नीति निर्णय क्षेत्रों में भी प्रयोग किया जा सकता है। उदाहरण के लिए , एक बडे कारोबार का एक प्रबन्धक अपनी देखरेख में आने वाले विभिन्न विक्रेताओं के कार्यो के प्रदर्शन का विश्लेषण कर सकता है और उनके प्रदर्शन को अर्थपूर्ण अन्तर में जानने के क्रम में नियंत्रित कर सकता है। इसी तरह, विभिन्न मशीनों के परिणामों के माध्य गुणों में अर्थपूर्ण अन्तर को निर्धारित किया जा सकता है। इस तरह का अध्ययन निर्धारित करेगा कि परिणामों के गुणों में एकरूपता को संचालन की मानकीकरण प्रक्रिया द्वारा बढाया जा सकता है या इसे मशीनों के मानकीकरण द्वारा बढाया जा सकता है। इस तरह से ANOVA व्यवसाय से सम्बन्धित नीति निर्णयों के लिए एक बहुत महत्वपूर्ण सांख्यिकीय तकनीक सि) हो सकता है। आपको हमेंशा याद रखना चाहिए कि प्रसरण का विशलेषण परीक्षण दो प्रतिदर्श प्रसरणों के मध्य अर्थपूर्ण अन्तर के परीक्षण के सर्वश्रेष्ठ उद्देश्य के लिए अभीष्ट

नहीं है, बल्कि इसका उद्देश्य प्रतिदर्श माध्यों के मध्य अन्तर के अर्थपूर्ण उद्देश्य के परीक्षण के लिए होता है। दो प्रसरणों के मध्य अर्थपूर्ण अन्तर के लिए इसे F परीक्षण की प्रक्रिया माध्यम से सम्पन्न किया जाता है , लेकिन परीक्षण को इस तरीके से प्रारूपित किया जाता है कि तुलना किये जा रहे प्रसरण भिन्न होते हैं केवल यदि संज्ञान के अन्तर्गत माध्य स भांगी नहीं होते हैं इस तरह से , F का अर्थपूर्ण मान निर्दिष्ट करता है कि माध्य एक दूसरे से अर्थपूर्ण तरीके से भिन्न है।

12.4.1 प्रसरण के स्रोत (Sources of Variance)

प्रसरण का विश्लेषण उन सभी परिस्थितियों के सन्दर्भ में महत्वपूर्ण तकनीक होती है जब शोधकर्ता दो से अधिक समग्रों की तुलना करना चाहता है। विभिन्न समग्रों के मध्य अन्तर के विश्लेषण के लिए हमें निर्णय करना पडता है कि प्रतिदर्श माध्यों के मध्य अन्तर केवल घटना के कारण होता है या क्या अन्तर विभिन्न समग्रों के माध्य के कारण जो लिये गये वास्तविक प्रतिदर्शों में से घटित होता है। आंकडों में दो प्रकार के प्रसरण हो सकते है। और ANOVA तकनीक आंकडों में हमें इन दो प्रकार के प्रसरण के अध्ययन में सहायता करता है। पहला "विभिन्न प्रतिदर्शों के बीच" और दूसरा "प्रतिदर्श के भीतर"

यदि प्रसरण प्रतिदर्श के भीतर और प्रतिदर्श के बीच एक दूसरे से अर्थपूर्ण भिन्न नही होते हैं , तब प्रतिदर्श केवल प्रसरण के भीतर एकसमान होते हैं । यदि प्रतिदर्श के मध्य विचरण, प्रतिदर्श के भीतर विचरण से बहुत अधिक होता है, इसका अर्थ है कि प्रतिदर्श समग्र के विभिन्न प्रकार से लिये गये हैं अन्यथा प्रतिदर्शों के मध्य और प्रतिदर्शो। के भीतर कोई अर्थपूर्ण अन्तर नहीं रहेगा। इसलिए , प्रसरण के विश्लेषण में, हम प्रतिदर्शों के बीच और प्रतिदर्शों के भीतर सम्बन्ध ज्ञात करते हैं। यदि समग्रों के सभी माध्य समान होते हैं , तब प्रतिदर्शों के मध्य परिवर्तनशीलता केवल घटना परिणाम होगा और इसलिए प्रतिदर्शे के भीतर उत्पन्न हुए परिवर्तनशीलता के समान होगा। दूसरी ओर, यदि समग्र माध्य एक समान नहीं होते हैं, प्रतिदर्शों के मध्य परिवर्तनशीलता प्रतिदर्शों के भीतर परिवर्तनशीलता से अधिक होगी।

परितर्वनशीलता को प्रसरण के विश्लेषण में मापने को 'माध्य वर्ग' कहा जाता है जिसकी गणना निम्नलिखित सूत्र से की जाती है।

माध्य वर्ग = माध्य से विचलनों के वर्ग का योग / स्वतन्त्रता का अंश

प्रतिदर्शों के भीतर परिवर्तनशीलता मापने के लिए, विचलनों को विशेष प्रतिदर्श माध्यों से और विचलनों के वर्ग के योग को स्वतन्त्रता के अश से विभाजित (प्रतिदर्श की संख्या को कुल प्रतिदर्श आकार से घटाकर) कर लिया जाता है जिसे प्रतिदर्शों के भीतर माध्य वर्ग कहा जाता है। यह माध्य वर्ग परिवर्तनशीलता की माप को जोकि घटना या प्रयोगात्मक त्रुटि के कारण हुई है को प्रदर्शित करता है। प्रतिदर्शों के मध्य परिवर्तनशीलता मापने के लिए प्रतिदर्श माध्य के विचलनों को सभी अवलोकनों के सर्वोच्च माध्य से लिया जाता है और विचलनों के वर्ग के योग को स्वतन्त्रता के अंश द्वारा (प्रतिदर्शों की संख्या को एक से घटाकर) विभाजित किया जाता है जिसे प्रतिदर्शों के मध्य माध्य वर्ग द्वारा जाना जाता है। यह माध्य वर्ग वर्ग प्रभाव को या प्रतिदर्शों के मध्य संभावित अंतर को प्रदर्शित करता है।

यदि सभी समग्रों का माध्य समान होता है. उनमें कोई वर्ग प्रभाव नहीं होता है और प्रतिदर्शों का माध्य वर्ग भी अकेले घटना के कारण परिवर्तनशीलता को प्रदर्शित करेगा। इसलिए , जब समग्र में प्रतिदर्श माध्य एकसमान होते है , प्रतिदर्शों के भीतर माध्य वर्ग और प्रतिदर्शों के बीच माध्य वर्ग में बहुत अधिक अंतर नहीं होना चाहिए और उनका अनुपात एक के करीब होना चाहिए। असमान्यतः बडे अनुपात इंगित करेंगे कि समग्र में प्रतिदर्श माध्य एक समान नहीं होते हैं।

12.4.2 ANOVA (एनोवा) का औचित्य (Rationale of ANOVA)

ANOVA के पीछे वैचारिक औचित्य यह है कि आंकडों के समूह में प्रसरण की मात्रा की दो तरह से विशेषता हो सकती है अर्थात घटना और निर्दिष्ट कारणों से और ANOVA के प्रयोग से हम विश्लेषणात्मक उद्देश्य के लिए इस प्रसरण को विभाजित कर सकते है। ANOVA कारकों के किसी संख्या के अनुसंधान के लिए स्वीकृति देता है जो आश्रित चर के प्रभाव के लिए अनुमानित हों। ANOVA का मूलभूत नियम प्रतिदर्शों भीतर प्रसरण की मात्रा का समग्रों के माध्य अन्तर के लिए परीक्षण द्वारा जांच पडताल करना और प्रतिदर्शों के मध्य सम्बन्धित प्रसरण की मात्रा का परीक्षण करना होता है। जबकि ANOVA का प्रयोग करते हुए, हम मानते है, कि प्रत्येक प्रतिदर्श को सामान्य समग्र से लिया गया है और प्रत्येक समग्र का प्रसरण एक समान है। यह भी कल्पना की जाती है कि एक या अधिक कारको के अतिरिक्त किये जा रहे परीक्षण प्रभावशाली तरीके से नियत्रण में है।
तद्पश्चात प्रत्येक आंकडे वर्ग के लिए स्वतन्त्र या-च्छिक प्रतिदर्शों को चयनित किया जाता है , प्रतिदर्शो के बीच प्रसरण की मात्रा और प्रतिदर्शो के भीतर प्रसरण की मात्रा के अनुपात पर कार्य किया जाता है , इसे F अनुपात के रूप में जाना जाता है। इसे निम्नलिखित सूत्र के आकार में वर्णित किया जा सकता है।

F = प्रतिदर्शो के मध्य प्रसरण पर आधारित समग्र प्रसरण का अनुमान / प्रतिदर्शों के भीतर प्रसरण पर आधारित समग्र प्रसरण का अनुमान

सामान्यतया प्रतिदर्शे के मध्य प्रसरण , प्रतिदर्शो के भीतर प्रसरण की तुलना में अधिक होगा। यदि परिस्थिति विपरीत होती है , अर्थात प्रतिदर्शो के बीच प्रसरण , प्रतिदर्शों के भीतर प्रसरण की तुलना में कम होता है तो अंश एवं हर की स्थितियों को परिवर्तित करना चाहिए और तद्नुसार निष्कर्ष निकाले जाने चाहिए लेकिन यह बहुत कदाचित होगा। F मान के गणना के पश्चात इसे दिये हुए अंश की स्वतन्त्रता के लिए तालिका मान के साथ तुलना की जाती है। यदि F का परिकल्पित मान तालिका मान के समान या अधिक होता है तब प्रतिदर्श माध्यों के बीच अर्थपूर्ण अंतर न होने की शून्य परिकल्पना अस्वीकार होती है। यह याद रखा जाना चाहिए कि ANOVA परीक्षण हमेंशा एक पुच्छीय परीक्षण होता है , चूँकि प्रतिदर्श आंकडों में से F का छोटा परिकल्पित मान का अर्थ होगा कि शून्य परिकल्पना के लिए समग्र माध्य बहुत लायक है।
ANOVA परीक्षण का अनुप्रयोग कुछ अवधारणाओं पर आधारित है जो निम्न रूप में है-

- समग्र की प्रसामान्यता
- प्रसरण की समरूपता
- या-च्छिकरण
- त्रुटि की स्वतन्त्रता

आप देख सकते है। कि ANOVA परीक्षण एवं F परीक्षण की अवधारणाऐं समान होती है। इन अवधारणाओं की पूर्ति प्रत्यक्ष रूप से इस परीक्षण की विश्वसनीयता को बढायेगी लेकिन यदि समग्र एक रूपात्मक हों और प्रतिदर्श आकर लगभग समान हों तब समग्र के प्रसामान्यता की अवधारणा का उल्लंधन परीक्षण की उपयुक्तता को प्रभावित नहीं करेगा।

12.4.3 ANOVA (एनोवा) तकनीक (ANOVA Technique)

प्रसरण के विश्लेषण के माध्यम से, शोधकर्ता कारकों की किसी संख्या जिसे अनुमानित किया जाता है की छानबीन कर सकते हैं। यदि शोधकर्ता केवल एक कारक को लेता है और इसके विभिन्न वर्गों के मध्य अन्तरों की छानबीन करता है जिसके बहुत संभावित मान है , तब शोधकर्ता एकतरफा ANOVA का प्रयोग करते हैं और उस स्थिति में जहाँ वह दो कारकों की छानबीन एक साथ करता है , तब उसके द्वारा दो तरफा ANOVA का

प्रयोग होता है। अच्छे निर्णय निर्धारण के लिए दो स्वतन्त्र चरों का एक आश्रित चर को प्रभावित करने का अध्ययन किया जाता है। आंकडों के वर्गीकरण के आधार पर या कारकों की सहभागिता ANOVA तकनीक को विभिन्न वर्गों जैसे एक तरफा ANOVA, दो तरफा ANOVA, ANOVA लैटिन वर्ग प्रारूप इत्यादि में विभाजित किया जा सकता है। विभिन्न स्थितियों में विभिन्न विधियों का प्रयोग किया जा सकता है जिसका सार निम्नलिखित रूप में है-

1. एक तरफा ANOVA
    - I. प्रत्यक्ष विधि
    - II. सरल मार्ग विधि
    - III. सांकेतिक विधि
2. दो तरफा ANOVA
    - I. अपुनरावृत्ति मानों के साथ
    - II. पुनरावृत्ति मानों के साथ
    - III. रेखांचित्रिय विधि

1. एकतरफा ANOVA

एकतरफा या एकल कारक ANOVA की स्थिति में , केवल एक कारक सुविचारित होता है और यह अवलोकित किया जाता है कि एकल कारक प्रतिदर्शो के भीतर परिवर्तनशीलता और प्रतिदर्शों के बीच परिवर्तनशीलता का अध्ययन महत्वपूर्ण होता है। यदि कारक के भीतर अन्तर होते है तो हमें निरीक्षण करना पडता हैं एकतरफा वर्गीकरण में , आंकडों को केवल एक मानदण्ड के अनुसार वर्गीकृत किया जाता है और शून्य परिकल्पना की जाती है जिसका कथन हातेा है कि समग्रों के समान्तर माध्यों का जिसमें इसके k प्रतिदर्श या-च्छिक लिये गये थे वे एक दूसरे के बराबर हैं। इसे निम्नवत वर्णित किया जा सकता है।

H_0: $\mu_1 = \mu_2 = \mu_3 .... = \mu_k$

इस शून्य परिकल्पना के परीक्षण के लिए हम विभिन्न वैकल्पिक विधियों का प्रयोग कर सकते हैं जिनका नीचे वर्णन किया गया है।

(अ) प्रत्यक्ष विधि- एकतरफा ANOVA परीक्षण में प्रत्यक्ष विधि के अन्तर्गत निम्नलिखित चरण शामिल हैं-

- सबसे पहले, प्रत्येक प्रतिदर्श का माध्य परिकल्पित किया जाता है-

$\overline{X}_1, \overline{X}_2, \overline{X}_3 .... \overline{X}_k$ (जब k प्रतिदर्श हैं)

- तद्पश्चात प्रतिदर्श माध्यों का माध्य निम्नलिखित तरीके से परिकल्पित किया जाता है-

$$\bar{\bar{X}} = \frac{n1 * \overline{X}_1 + n2 * \overline{X}_2 + n3 * \overline{X}_3 + \cdots .... + nk * \overline{X}_k}{n1 + n2 + n3 + \cdots ....nk}$$

- अगले चरण में, प्रतिदर्श माध्यों के विचलनों को प्रतिदर्श माध्यों के माध्य से परिकल्पित किया जाता है।
- तद्पश्चात इन विचलनों का वर्ग किया जाता है और समरूपी प्रतिदर्श में इन्हें पदों की संख्या द्वारा गुणा किया जाता है और उनका संकलन प्राप्त किया जाता है इसे प्रतिदर्श के बीच प्रसरण या SS बीच के लिए वर्गों का योग कहा जाता है। इसे निम्नलिखित रूप में वर्णित किया जा सकता है।

$$\text{SS बीच} = n_1\left(\bar{X}_1 - \bar{\bar{X}}\right)^2 + n_2\left(\bar{X}_2 - \bar{\bar{X}}\right)^2 + ... + n_k\left(\bar{X}_k - \bar{\bar{X}}\right)^2$$

- तब प्रतिदर्शो के बीच प्रसरण के लिए वर्गों के योगों का प्रतिदर्शो। के बीच स्वतंत्रता के अंश द्वरा विभाजित किया जाता है जो माध्य वर्ग मध्य प्रदान करता है। लाक्षणिक रूप से

MS मध्य = $\frac{SS\ Between}{(k-1)}$

जहॉ (k - 1) प्रतिदर्शो के मध्य स्वतंत्रता का अंश ।

- अगले चरण में , SS भीतर परिकल्पित किया जाता है। इसके लिए , सभी प्रतिदर्शों के लिए समरूपी प्रतिदर्श माध्य में से प्रतिदर्श पदो के विचलन का मान परिकल्पित करते हैं , इन विचलनों का वर्ग किया जाता है और उनका संकलन प्राप्त करते हैं। इसे प्रतिदर्शों के भीतर प्रसरण के लिए वर्गों के योग के रूप में जाना जाता है। इसे निम्नलिखित रूप में व्यक्त किया जा सकता है।

SS भीतर =$\sum(X_{1i} - \overline{X}_1)^2 + \sum(X_{2i} - \overline{X}_2)^2 + ... + \sum(X_{ki} - \overline{X}_k)^2$

with i =1, 2, 3..............के साथ

- तद्पश्चात "प्रतिदर्श के भीतर माध्य वर्ग " की गणना प्रतिदर्श के भीतर स्वतन्त्रता के अंश के साथ प्रतिदर्श के भीतर प्रसरण के लिए वर्गों के योग द्वारा विभाजन से की जाती है।

MS भीतर = ss भीतर / (n - k)

जहाँ n सभी प्रतिदर्शों के कुल पदों की संख्या अर्थात $n_1 + n_2 + ...+ n_k$

k = प्रतिदर्शों की कुल संख्या

इस प्रकार (n-k) प्रतिदर्शों के भीतर स्वतन्त्रता के अंशों को i = 1, 2, 3........ के साथ प्रदर्शित करता है।

- अन्तिम चरण में, F अनुपात की गणना निम्नलिखित सूत्र द्वारा की जाती है।

F- ratio = $\frac{MS\ Between}{MS\ Within}$

तद्पश्चात F के परिकल्पित मान की तुलना F के तालिका मान के साथ निर्दिष्ट महत्व के स्तर पर दिये हुए स्वतन्त्रता के अंश के लिए किया जाता है। यदि F अनुपात का परिकल्पित मान तालिका मान से छोटा होता है तब शून्य परिकल्पना स्वीकार होती है और यदि परिकल्पित मान तालिका मान से अधिक होता है तब शून्य परिकल्पना अस्वीकार होती है। इस अनुपात का प्रयोग यह निर्णय करने में किया जाता है कि क्या विविध प्रतिदर्श माध्यों के मध्य अन्तर अर्थपूर्ण है या यह केवल एक प्रतिचयन अस्थिरता का विषय है।

ANOVA तकनीक का योगात्मक गुण- कुल प्रसरण के लिए विचलन के वर्ग का योग प्रतिदर्शों के भीतर प्रसरण के लिए वर्ग के योग के जोड और प्रतिदर्शों के बीच प्रसरण के लिए वर्ग के योग द्वारा प्राप्त किया जा सकता है।

लाक्षणिक रूप से कुल प्रसरण के लिए SS = SS माध्य + SS भीतर कुल प्रसरण के लिए वर्ग के इस योग को एक वैकल्पिक विधि द्वारा भी ज्ञात किया जा सकता है इस प्रक्रिया में विचलनों के वर्गों का योग शामिल होता है जब सभी प्रतिदर्शों में एकल पदो के लिए विचलनों को प्रतिदर्श लाक्षणिक रूप से,

कुल प्रसरण के लिए SS = $\sum\left(X_{ij} - \overline{\overline{X}}\right)^2$

i = 1, 2, 3,..........

j = 1, 2, 3,.........

कुल प्रसरण के लिए अंश की स्वतन्त्रता = (n - 1) = (k - 1) + (n - k)

इसका अर्थ है कि कुल प्रसरण के लिए अंशों की स्वतंत्रता सभी प्रतिदर्शों में पदों की संख्या ऋण एक के बराबर होगी। इसे प्रतिदर्शों के बीच के लिए अंश की स्वतन्त्रता का योग और प्रतिदर्शो के भीतर अंश की स्वतन्त्रता के योग द्वारा भी ज्ञात किया जा सकता है। यही कारण ANOVA तकनीक के योगात्मक गुण का है।

एकतरफा या एकल कारक ANOVA तकनीक में शामिल विभिन्न चरणों के आधार पर , उनकी गणना निम्नलिखित प्रसरण के विश्लेषण तालिका के रूप में संक्षिप्त की जा सकती है।

एकतरफा ANOVA के लिए प्रसरण की विश्लेषण तालिका

विचरण के स्रोत	वर्गों का योग (SS)	अंश की स्वतन्त्रता (d.f.)	माध्य वर्ग (MS)	F-अनुपात
(i) प्रतिदर्शो के बीच	$n_1\left(\bar{X}_1-\bar{\bar{X}}\right)^2 + n_2\left(\bar{X}_2-\bar{\bar{X}}\right)^2 + \ldots\ldots + n_k\left(\bar{X}_k-\bar{\bar{X}}\right)^2$	(k – 1) k = प्रतिदर्शों की संख्या	$\frac{SS\ Between}{k-1}$	$\frac{MS\ Between}{MS\ Within}$
(ii) प्रतिदर्शों के भीतर	$\sum(X_{1i}-\bar{X}_1)^2 + \sum(X_{2i}-\bar{X}_2)^2 + \ldots\ldots + \sum(X_{ki}-\bar{X}_k)^2$ i =1,2,3……………. के साथ	(n – k) n = कुल पदों की संख्या	$\frac{SS\ Within}{(n-k)}$	
(iii) कुल	$\sum\left(X_{ij}-\bar{\bar{X}}\right)^2$ i = 1, 2, 3,…………….. j = 1, 2, 3, ………………	(n – 1)		

उदाहरण 3- एक शहर के कोनवेन्ट स्कूलों के मध्य किसी परीक्षा में प्रदर्शन के सम्भव विचरण के महत्व के आंकलन में एक सार्वजनिक परीक्षा चार स्कूलों से सम्बन्धित प्रत्येक अपर पांचवीं कक्षा से या-च्छिक रूप से लिए गये छात्रों को दिलाई गयी थी। परिणाम नीचे दिये गये हैं। आंकडों के प्रसरण के विश्लेषण का सृजन करें।

स्कूल			
A	B	C	D
8	12	18	13
10	11	12	9
12	9	16	12
8	14	6	16
7	4	8	15

हल- हम शून्य परिकल्पना लेते है। कि $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

प्रत्येक प्रतिदर्श का माध्य-

$\overline{X_1} = \frac{8+10+12+8+7}{5} = \frac{45}{5} = 9$

$\overline{X}_2 = \frac{12+11+9+14+4}{5} = \frac{50}{5} = 10$

$\overline{X}_3 = \frac{18+12+16+6+8}{5} = \frac{60}{5} = 12$

$\overline{X}_4 = \frac{13+9+12+16+15}{5} = \frac{65}{5} = 13$

प्रतिदर्श माध्यों का माध्य

$\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3 + \overline{X}_4}{k}$

$= \frac{9+10+12+13}{4} = 11$

SS मध्य

SS Between = $n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + n_3(\bar{X}_3 - \bar{\bar{X}})^2 + n_4(\bar{X}_4 - \bar{\bar{X}})^2$

$= 5(9-11)^2 + 5(10-11)^2 + 5(12-11)^2 + 5(13-11)^2$

$= 20 + 5 + 5 + 20$

$= 50$

MS मध्य

MS Between = $\frac{SS\ Between}{(k-1)}$

$= \frac{50}{4-1} = 16.7$ (there are 4 samples)

SS भीतर

SS भीतर = $\sum(X_{1i} - \bar{X}_1)^2 + \sum(X_{2i} - \bar{X}_2)^2 + \sum(X_{3i} - \bar{X}_3)^2 + \sum(X_{4i} - \bar{X}_4)^2$

$= \{8-9)^2 + (10-9)^2 + (12-9)^2 + (8-9)^2 + (7-9)^2\}$

$+ \{(12-10)^2 + (11-10)^2 + (9-10)^2 + (14-10)^2 + (4-10)^2\}$

$+ \{(18-12)2 + (12-12)2 + (16-12)2 + (6-12)2 + (8-12)^2\}$

$+ \{(13-13)2 + (9-13)2 + (12-13)2 + (16-13)2 + (15-13)^2\}$

$= \{1+1+9+1+4\} + \{4+1+1+16+36\} + \{36+0+16+36+16\} + \{0+16+1+9+4\}$

$= 16 + 58 + 104 + 30 = 208$

MS भीतर

MS Within = $\frac{SS\ Within}{(n-k)}$

$= \frac{208}{20-4} = \frac{208}{16} = 13$

F-अनुपात

F-अनुपात = $\frac{MS\ Between}{MS\ Within}$

$= \frac{16.7}{13} = 1.285$

उपरोक्त वर्णित गणनाओं को निम्नलिखित तालिका के रूप में संक्षिप्त किया जा सकता है-

विचरण के श्रोत	SS	d.f.	MS	F- अनुपात	5% F – सीमा
मध्य प्रतिदर्श	50	4-1=3	16.7	1.285	F (3, 16)
भीतर प्रतिदर्श	208	20-4=16	13		3.24
कुल	258				

F (1.285) का परिकल्पित मान तालिका मान (3.24) से कम है , इसलिए शून्य परिकल्पना स्वीकार होती है और हम यह निष्कर्ष निकाल सकते हैं कि प्रतिदर्श समान समग्रों से लिये गये हैं।

(ब) सरल मार्ग विधि - उपरोक्त वर्णित ANOVA के एकतरफा तकनीक की प्रत्यक्ष विधि बहुत विस्तृत और अधिक समय लेने वाली होती है। प्रत्यक्ष विधि के स्थान पर , सरल मार्ग विधि को एकतरफा ANOVA से सम्बन्धित समस्याओं के लिए प्रयुक्त किया जा सकता है और हम समान परिणाम प्राप्त करेंगे। वास्तव में , सरल मार्ग विधि प्रत्यक्ष विधि की तुलना में अधिक लोकप्रिय है और सामान्यतया इसे एकतरफा ANOVA के अभ्यास के लिए प्रयोग किया जाता है क्योकि यह कम समय लेने वाली , आसान और यह विशेष रूप से संगणनात्मक कार्य को कम करता है। सरल मार्ग विधि में निम्नलिखित चरण शामिल होते हैं-

- सबसे पहले, सभी प्रतिदर्शों में एकल पदों का योग ज्ञात किया जाता है और इसे 'T' के रूप में जाना जाता है। लाक्षणिक रूप से,

$T = \sum X_{ij}$ जहॉं $i = 1, 2, 3....$ $j = 1, 2, 3,....$

- तद्पश्चात "संशोधन कारक" के अर्न्तगत कार्य करते हैं" संशोधन कारक = $\frac{(T)^2}{n}$
- अगले चरण में, कुल विचरण के वर्गो के योग के लिए हम सभी पद मानों के वर्ग द्वारा और इसका योग लेते हुए और संशोधन कारक को इसमें से घटाते हैं।

योग SS = $\sum X^2_{ij} - \frac{(T)^2}{n}$

- तब हम प्रतिदर्शों के बीच प्रसरण के लिए वर्गों का योग ज्ञात करते हैं। इस मान को प्राप्त करने के लिए , प्रत्येक प्रतिदर्श वर्ग $(T_j)^2$ के वर्ग को प्रतिदर्श में सम्बन्धित पदों की संख्या द्वारा विभाजित किया जाता है, उनका संकलन ज्ञात किया जाता है और संशोधन कारक को इस संकलन से घटाया जाता है।

SS मध्य = $\sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$ जहॉं $j = 1, 2, 3,....$

- अगले चरण में, प्रतिदर्शों के बीच के लिए वर्ग के योग को कुल प्रसरण के लिए वर्गों के योग में घटाया जाता है और परिणामित मान प्रतिदर्शों के भीतर के लिए वर्गो के योग को निर्दिष्ट करता है। लाक्षणिक रूप से,

$$\text{SS भीतर} = \left\{\sum X_{ij}^2 - \frac{(T)^2}{n}\right\} - \left\{\sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}\right\}$$

$$= \sum X_{ij}^2 - \sum \frac{(T_j)^2}{n_j}$$

तद्पश्चात एक प्रत्यक्ष विधि के प्रयोग के लिए समान तरीके से ANOVA तालिका को निर्मित किया जाता है।

उदाहरण 4- तीन प्रजातियों के गेहूँ 4 भूखंडों में विकास के लिए , प्रसरण के विश्लेषण को स्थापित करें इसके लिए प्रति उत्पादन आंकड़ा निम्नवत दिया है और कथन की पुष्टि करें कि प्रजातियों में अंतर महत्वपूर्ण है ।

प्रति एकड़ उत्पादन आंकड़ा

जमीन का भूखंड	**गेहूँ की प्रजाति**		
	A	B	C
1	6	5	5
2	7	5	4
3	3	3	3
4	8	7	4

हल-

हम इस समस्या को सरल मार्ग विधि द्वारा हल करेंगें

$H_0 : \mu_1 = \mu_2 = \mu_3$

मान लें कि गेहूँ की तीन प्रजातियों में कोई अर्थपूर्ण अन्तर नहीं है, शून्य परिकल्पना है।

$T = \sum X_{ij}$

$= 6+7+3+8+5+5+3+7+5+4+3+4 = 60$

संशोधन कारक $= \frac{(T)^2}{n} = \frac{(60)^2}{12} = 300$

कुल SS $= \sum X^2_{ij} - \frac{(T)^2}{n}$

$= (6)^2 + (7)^2 + (3)^2 + (8)^2 + (5)^2 + (5)^2 + (3)^2 + (7)^2 + (5)^2 + (4)^2 + (3)^2 + (4)^2 - \frac{(60)^2}{12}$

$= 332 - 300 = 32$

SS मध्य $= \sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$

$= \frac{(24)^2}{4} + \frac{(20)^2}{4} + \frac{(16)^2}{4} - \frac{(60)^2}{12}$

$= 144 + 100 + 64 - 300 = 308 - 300 = 8$

SS भीतर $= \sum X_{ij}^2 - \sum \frac{(T_j)^2}{n_j}$

$= 332 - 308 = 24$

ANOVA तालिका निम्नवत है-

प्रसरण के श्रोत	**SS**	**d.f.**	**MS**	**F- ratio**	**5% F-सीमा**
प्रतिदर्श के बीच	8	3-1=2	$\frac{8}{2} = 4.00$	4.00/2.67 = 1.5	F (2, 9) = 4.26

प्रतिदर्श के भीतर	24	12-3=9	$\frac{24}{9}$ = 2.67		
योग	32				

F (1.5) का परिकल्पित मान तालिका मान (4.26) से कम है, इसलिए शून्य परिकल्पना स्वीकृत होती है और हम निष्कर्ष निकाल सकते हैं कि गेहूँ की पैदावार में अन्तर प्रजातियों की वजह से अर्थपूर्ण नहीं है और यह केवल घटना के तरीके से हुआ है।

(स) सांकेतिक विधि- कभी कभी हमें बडे मानों के साथ समझौता करना पडता है। यह गणना की प्रक्रिया को बहुत जटिल बनाता है। इन स्थितियों में , हम सांकेतिक विधि की सहायता ले सकते है। यह एक सरल मार्ग विधि का विस्तार है। इस प्रकार , सांकेतिक विधि का प्रयोग समस्याओं के सरलीकरण के लिये किया जाता है जिसमें बडे मान शामिल होते हैं। बीजांक एक स्थायी द्वारा जोड , घटाना, गुणा या भाग का उल्लेख करता है। यदि सभी n पदों को एक सार्वाजनिक कारक जिसे स्थायी कहा जाता है द्वारा या तो गुणा किया जाता है या भाग दिया जाता है या एक अपरिवर्तनशील को प्रत्येक n पदों में जोडा या घटाया जाता है , तब भी F अनुपात का मान प्रभावित नहीं होता है। इसका अर्थ है कि वास्तविक माप की गणना को परिणामों के अनुगामी सामजस्यों के आवश्यकता बिना सरलीकरण किया जा सकता है। एक बार दिये हुए मान कुछ सार्वजनिक मान के साथ परिवर्तित किये जाते हैं, तब सरल मार्ग विधि के सभी चरणों को F अनुपात को ज्ञात और व्याख्या करने के लिए अंगीकृत किया जा सकता है।

उदाहरण 5- 5% मोटर कार टायरों के यादृच्छिक प्रतिदर्शों को तीन कम्पनीयों द्वारा निर्मित प्रत्येक के 3 ब्रान्डों से लिया गया है। इन टायरों (मीलवार दौड द्वारा मापने के रूप में) नीचे दिखाया गया है। आंकडों के आधार पर, परीक्षण करें कि 3 ब्रान्ड के टयरों का औसत जीवनकाल समान है या नहीं।

टायरों का जीवनकाल (000 मील)

ब्रान्ड

A	B	C
35	32	34
34	32	33
34	31	32
33	28	32
34	29	33

हल- $H_0 : \mu_1 = \mu_2 = \mu_3$

शून्य परिकल्पना मानते है कि तीन ब्रान्ड के टायरों के मध्य कोई अन्तर नहीं है।

गणनाओं के सरलीकरण के क्रम में, प्रत्येक अवलोकन को 30 द्वारा कम किया जाता है।

बीजांक आंकड़ा है-

A	B	C
5	2	4
4	2	3
4	1	2
3	-2	2

4	-1	3

T = $\sum X_{ij}$

= 5 + 4 + 4 + 3 + 4 + 2 + 2 +1 – 2 - 1 + 4 + 3 + 2 + 2 + 3 = 36

संशोधन कारक = $\frac{(T)^2}{n} = \frac{(36)^2}{15}$ = 86.4

कुल SS = $\sum X^2_{ij} - \frac{(T)^2}{n}$

= $(5)^2+(4)^2+(4)^2+(3)^2+(4)^2+(2)^2+(2)^2+(1)^2+(-2)^2+(-1)^2+(4)^2+(3)^2+(2)^2+(2)^2+(3)^2-86.4$

= 138 - 86.4 = 51.6

SS मध्य = $\sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$

= $\frac{(20)^2}{5} + \frac{(2)^2}{5} + \frac{(14)^2}{5}$ - 86.4

= 80 + 0.8 + 39.2 - 86.4

= 120 - 86.4 = 33.6

SS भीतर = $\sum X_{ij}^2 - \sum \frac{(T_j)^2}{n_j}$

= 138 -120 = 18

ANOVA तालिका निम्नवत है-

प्रसरण के स्रोत	**SS**	**d.f.**	**MS**	**F-ratio**	**5% F-सीमा**
मध्य प्रतिदर्श **भीतर प्रतिदर्श**	33.6 18.0	3 – 1= 2 15– 3 = 12	$\frac{33.6}{2}$ = 16.8 $\frac{18}{12}$ = 1.5	$\frac{16.8}{1.5}$ = 11.2	F (2,12) = 3.89
योग	51.6				

F (11.2) का परिकल्पित मान तालिका मान (3.89) से अधिक है, इसलिए शून्य परिकल्पना अस्वीकार है और हम यह निष्कर्ष निकाल सकते हैं कि टायरों के 3 ब्राण्डों का औसत जीवनकाल एकसमान नहीं है।

2. दो तरफा ANOVA

एकतरफा ANOVA में आपको ध्यान देना चाहिए कि एकल कारक के विभिन्न स्तरों के निरूपण संगठन जो कि प्रयोग में नियंत्रित रहता है। लेकिन वास्तविक जीवन स्थितियों में , एक शोधकर्ता एक ही समय में एक से अधिक कारकों के प्रभाव को जानने में रूचि रख सकते हैं या हम बहुत सी उन स्थितियों का सामना कर सकते हैं जिसमें रूचि का प्रतिक्रिया चर एक से अधिक कारकों द्वारा प्रभावित हो सकता है। उदाहरण के लिए, कृषि सम्बन्धी परिणाम उर्वरक के प्रकार एवं बीज की प्रजाति द्वारा प्रभावित किया जा सकता है , उत्पादन को मशीनों की विभिन्न प्रजातियों एवं मजदूरों के विभिन्न वर्गो द्वारा प्रभावित किया जा सकता है , उत्पाद विक्री को विज्ञापन स्तरें एवं कीमत स्तरों द्वारा प्रभावित किया जा सकता है। इन स्थितियों में , हम दो

तरफा ANOVA का प्रयोग करेगें। इस प्रकार, दो तरफा ANOVA तकनीक प्रयोग की जाती है जब आंकडों का वर्गीकरण दो कारकों के आधार पर होता है। हम परीक्षण की रचना इस तरीके से कर सकते है। कि एक ही समय में दो कारकों के प्रभाव का परीक्षण प्रसरण के विश्लेषण से किया जा सके। दो तरफा ANOVA के साथ, एक ही समय, एक से आंकडे के साथ हम परिकल्पना के दो समूहों का परीक्षण कर सकते हैं। इस तकनीक का सबसे बडा लाभ यह है कि यह शोधकर्ता को कारकों के मध्य पारस्परिक प्रभाव के निरीक्षण के लिए समर्थ बनाता है। दो तरफा प्रारूप के प्रत्येक कारक में पुनरावृत्ति मापें हो सकती है या पुनरावृत्ति मापें नहीं हो सकती है।

(अ) अपुनरावृत्ति मानों के साथ- दो तरफा वर्गीकरण में प्रसरण के विश्लेषण के लिए प्रक्रिया एकतरफा वर्गीकरण की स्थिति से थेाडी सी भिन्न होती है। जब हमारे पास पुनरावृत्ति मान नहीं होते हैं , प्रतिदर्शों के भीतर वर्गों के योगों की गणना सीधे नहीं की जा सकती हैं इस अवशेष या विचरण त्रुटि की गणना एक निरूपण के मध्य प्रजातियों के प्रसरण के लिए वर्गो के योग द्वारा एवं दूसरे निरूपण के मध्य प्रजातियों के प्रसरण के लिए वर्गों के योग को कुल प्रसरण के लिए वर्गों के योग में से घटाकर की जाती है। संगणना प्रक्रिया में निम्नलिखित चरण शामिल होते हैं-

- जब दोतरफा ANOVA की गणना करते हैं, यदि मान जटिल होते हैं, तब बीजांक प्रारम्भ में किया जा सकता है और तद्पश्चात अगले चरणों का अनुसरण करते हैं।
- सभी प्रतिदर्शो में एकल पदो के मानों या बीजांक मानों का योग किया जाता है और इस संगकलन को ज् के रूप में जाना जाता है। सांकेतिक रूप में,

$$T = \sum X_{ij}$$

- तद्पश्चात निम्नलिखित तरीके से संशोधनकारक ज्ञात किया जाता है

$$\text{संशोधन कारक} = \frac{(T)^2}{n}$$

- अगले चरण में संशोधन कारक को एकल पदों के वर्गों के योग में से घटाया जाता है और परिणामी मान कुल प्रसरण के लिए वर्गों के विचलनों का कुल योग होता है।
  सांकेतिक रूप में,

$$\text{कुल SS} = \sum X^2_{ij} - \frac{(T)^2}{n}$$

- अब स्तम्भों के मध्य प्रसरण के लिए हमें विचलनों के वर्गो का योग ज्ञात करना पडता है। इस मान की गणना के लिए, विभिन्न स्तम्भों का योग किया जाता है और प्रत्येक स्तम्भ के योग के वर्ग को सम्बन्धित स्तम्भ में पदों की संख्या द्वारा विभाजित किया जाता है और इन मानों को संकलित किया जाता है और तद्पश्चात इस संकलन में से संशोधन कारक को घटाया जाता है, सांकेतिक रूप में,

$$\text{स्तम्भों के मध्य SS} = \sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$$

- अब स्तंभों के मध्य प्रसरण के लिए ह में विचलनों के वर्गों का योग ज्ञात करना पडता है। इस मान की गणना के लिए, विभिन्न स्तम्भों का योग किया जाता है और प्रत्येक स्तम्भ के योग के वर्ग को सम्बन्धित स्तम्भ में पदों की संख्या द्वारा विभाजित किया जाता है और इन मानों को संकलित किया जाता है और तद्पश्चात इस संकलन में से संशोधन कारक को घटाया जाता है सांकेतिक रूप में,

$$\text{स्तम्भों के मध्य SS} = \sum \frac{(T_i)^2}{n_i} - \frac{(T)^2}{n}$$

- स्तम्भों के मध्य गणना के पश्चात हमें पंक्तियों के मध्य ss की गणना करनी पडती है। पंक्तियों के मध्य प्रसरण के लिए विचलनों के वर्गों के योग की गणना कुल पंक्ति के वर्ग के योग को सम्बन्धित पंक्ति में पदों की संख्या द्वारा विभाजन में से संशोधन कारक को घटाकर की जाती है। सांकेतिक रूप से

$$\text{पंक्तियों के मध्य SS} = \sum \frac{(T_i)^2}{n_i} - \frac{(T)^2}{n}$$

- अगले चरण में, स्तम्भों के मध्य प्रसरण के लिए विचलनों के वर्गों का योग और पंक्तियों के मध्य प्रसरण के लिए विचलनों के वर्गों का योग को कुल प्रसरण के लिए विचलनों के वर्गों के योग में से घटाकर किया जाता है और परिणामी मान अवशेष या त्रुटि प्रसरण के लिए विचलनों के वर्गों के योग को दर्शाता है। इसे निम्नवत वर्णित किया जा सकता है-

अवशेष SS = कुल SS - (स्तम्भों के मध्य SS = पंक्तियों के मध्य SS)

- F अनुपात के मान को प्राप्त करने के लिए , हमें भिन्न वर्गों के योग के लिए अंशों की स्वतन्त्रता ज्ञात होनी चाहिए जिसके लिए निम्न के अन्तर्गत कार्य करना पड सकता है:-

कुल प्रसरण के लिए d.f. = (c . r - 1)

स्तम्भों के मध्य प्रसरण के लिए d.f. = (c - 1)

पंक्तियों के मध्य प्रसरण के लिए d.f. = (r - 1)

अवशेष प्रसरण के लिए d.f. = (c - 1) (r -1)

जहाँ c = स्तम्भों की संख्या r = पंक्तियों की संख्या

- तब एक दो तरफा ANOVA तालिका निम्नलिखित तरीके से निर्मित की जाती है-

दो तरफा ANOVA के लिए प्रसरण के विश्लेषण की तालिका

विचरण के स्रोत	वर्गों का योग SS	स्वतंत्रता का अंश (d.f.)	माध्य वर्ग (MS)	F-अनुपात
स्तम्भों के मध्य	$\sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$	(c – 1)	$\frac{\text{SS between Columns}}{(c-1)}$	$\frac{\text{MS between columns}}{(\text{MS residual})}$
पंक्तियों के मध्य	$\sum \frac{(T_i)^2}{n_i} - \frac{(T)^2}{n}$	(r – 1)	$\frac{\text{SS between rows}}{(r-1)}$	$\frac{MS\ between\ rows}{(MS\ residual)}$
अवशेष या त्रुटि	कुल ss – (स्तंभों का मध्य SS + पंक्तियों के मध्य SS)	(c–1)(r–1)	$\frac{\text{SS residual}}{(c-1)(r-1)}$	
कुल	$\sum X^2_{ij} - \frac{(T)^2}{n}$	(c.r – 1)		

आपकों यह स्पष्ट होना चाहिए कि अपुनरावृत्ति मानों के साथ दो तरफा ANOVA में अवशेष प्रसरण का आधार F अनुपात होता है। अवशेष प्रसरण के घटित होने के लिए कारण प्रतिचयन की अस्थिरता होती है। महत्व के निर्दिष्ट स्तर पर दिये हुए अंश की स्वतन्त्रता के लिए दोनों F अनुपातों की तुलना उनके समरूपी तालिका मानों के साथ की जाती है। F के आधार पर शून्य परिकल्पनाओं के लिए स्वीकृत एवं अस्वीकृत मानदण्ड समान रहते हैं।

उदाहरण 6- निम्नलिखित आंकडे इकाईयों की संख्या के प्रतिदिन उत्पादन को प्रदर्शित करते है। जो 3 भिन्न मजदूरों द्वारा 4 भिन्न प्रकार की मशीनों से निकले है। नीचे दिये हुए आंकडों में दो तरफा ANOVA को प्रदर्शित करें।

मजदूर	**मशीन के प्रकार**			
	A	**B**	**C**	**D**
I	38	40	41	39
II	45	42	49	36
III	40	38	42	42

(सांकेतिक विधि का प्रयोग दिये हुए संख्याओं को 40 से घटाकर करें)

हल- हम शून्य परिकल्पना लेते है कि उत्पादकता माध्य में मशीन प्रकार एवं विभिन्न मजदूरों के सन्दर्भ में कोई अर्थपूर्ण अन्तर नहीं है। प्रत्येक मान को 40 में से घटाकर कर, हम प्राप्त करते हैं-

मजदूर	**मशीन प्रकार**				**कुल**
	A	B	C	D	
I -	2	0	+1	-1	**-2**
II +	5	+2	+9	-4	**+12**
III	0	-2	+2	+2	**+2**
कुल +	**3**	**0**	**+12**	**-3**	**+12**

उपरोक्त तालिका से 'T' स्पष्ट है कि या $\sum X_{ij} = 12$

संशोधन कारक $= \frac{(T)^2}{n} = \frac{(12)^2}{12} = 12$

कुल SS $= \sum X^2_{ij} - \frac{(T)^2}{n}$

$= (-2)^2 + (5)^2 + (0)^2 + (0)^2 + (2)^2 + (-2)^2 + (1)^2 + (9)^2 + (2)^2 + (-1)^2 + (-4)^2 + (2)^2 - 12$

$= 4 + 25 + 0 + 0 + 4 + 4 + 1 + 81 + 4 + 1 + 16 + 4 - 12 = 144 - 12 = 132$

मशीनों के मध्य वर्गों का योग $= \sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$

$= \frac{(3)^2}{3} + \frac{(0)^2}{3} + \frac{(12)^2}{3} + \frac{(-3)^2}{3} - 12$

$= 3 + 0 + 48 + 3 - 12 = 42$

मजदूरों के मध्य वर्गों का योग = $\sum \frac{(T_i)^2}{n_i} - \frac{(T)^2}{n}$

$$= \frac{(-2)^2}{4} + \frac{(12)^2}{4} + \frac{(2)^2}{4} - 12$$

$$= 1 + 36 + 1 - 12 = 26$$

SS अवशेष = कुल SS - (मशीनों के मध्य SS + मजदूरों के मध्य SS)

= 132 - (42 + 26) = 132 - 68 = 64

ANOVA तालिका निम्नवत है-

विचरण के स्रोत	SS	d.f.	MS	F- अनुपात	5% F- सीमा
मशीनों के मध्य	42	4– 1= 3	$\frac{42}{3} = 14$	$\frac{14}{10.67} = 1.31$	F (3,6) = 4.76
मजदूरों के मध्य	26	3– 1 =2	$\frac{26}{2} = 13$	$\frac{13}{10.67} = 1.22$	F(2,6) = 5.14
अवशेष	64	(4 –1)(3–1) =6	$\frac{64}{6} = 10.67$		
कुल	132	(4×3 – 1) =11			

चूँकि F अनुपातों (1.31, 1.22) दोनों परिकल्पित मान उनके तालिका मानों (4.76ए 5.14) से कम है इसलिए, दोनों शून्य परिकल्पनाऐं स्वीकार है और हम निष्कर्ष निकाल सकते हैं कि उत्पादकता माध्यों में मशीन प्रकार के साथ साथ मजदूरों के सन्दर्भ में कोई अर्थपूर्ण अन्तर नहीं है।

(ब) पुनरावृत्ति मानों के साथ- कदाचित हमें दो तरफा प्रारूप की कुछ स्थितियों में सामना करना पढ सकता है जहाँ सभी वर्गों के लिए पुनरावृत्ति मापें होती है। दो तरफा प्ारूप के अपुनरावृत्ति मानों के साथ और दो तरफा पुनरावृत्ति मानों की संगणना प्रक्रिया में केवल एक अन्तर होता है । कुल ss स्तम्भों के मध्य ss और पंक्तियों के मध्य ss की गणना समान तरीके से की जाती है। पुनरावृत्ति मानों की स्थिति में , हमें अन्योन्यक्रिया विचरण की गणना करनी पडती है। दो तरफा विश्लेषण में अन्योन्यक्रिया का तात्पर्य यह है कि दो निरूपण तन्त्र नहीं है और एक कारक का विशेष निरूपण का प्रभाव दूसरे कारक के स्तर पर आश्रित रहता है और विपरीत क्रम में प्रतिदर्शों के भीतर प्रसरण के लिए वर्गों के योग की गणना एक तरफा ANOVA की स्थिति के रूप में समान तरीके से की जाती है। अन्योन्य क्रिया विचरण की गणना शेष बचे वर्गो के योग के पर एवं बचे शेष स्वतन्त्रता के अंशों के आधार पर की जाती है।

एक अर्थपूर्ण अन्योन्यक्रिया प्रभाव इंगित करता है कि एक कारक के लिए निरूपण का प्रभाव दूसरे कारक द्वारा दृढ़ता से प्रभावित हुआ है।

ANOVA तालिका को सामान्य तरीके से तैयार किया जाता है।

उदाहरण 7- क्या अन्योन्य क्रिया विचरण निम्नलिखित सूचना सम्बन्धित मील संख्या आधारित विभिन्न ब्रान्डों की गैसोलीन एवं कारों की स्थिति में अर्थपूर्ण है-

	गैसोलीन के ब्रान्ड		
कार	X	Y	Z
A	12	10	9
	12	9	11
B	12	7	10
	11	8	11
C	10	11	8
	11	11	7

हल- H_0 कारों एवं गैसोलीन ब्रान्डों के मध्य कोई अर्थपूर्ण अन्योन्यक्रिया नहीं है।

$T = \sum X_{ij}$

$= (12 + 12 + 10 + 9 + 9 + 11) + (12 + 11 + 7 + 8 + 10 + 11) + (10 + 11 + 11 + 11 + 8 + 7)$

$= 63 + 59 + 58 = 180$

संशोधन कारक $= \frac{(T)^2}{n} = \frac{(180)^2}{18} = 1800$

कुल SS $= \sum X^2_{ij} - \frac{(T)^2}{n}$

$= (12)^2 + (12)^2 + (10)^2 + (9)^2 + (9)^2 + (11)^2 + (12)^2 + (11)^2 + (7)^2 + (8)^2 + (10)^2 + (11)^2 + (10)^2 + (11)^2 + (11)^2 + (11)^2 + (8)^2 + (7)^2 - 1800$

$= 1846 - 1800 = 46$

SS स्तम्भों के मध्य $= \sum \frac{(T_j)^2}{n_j} - \frac{(T)^2}{n}$

$$= \left\{\left(\frac{68 \times 68}{6}\right) + \left(\frac{56 \times 56}{6}\right) + \left(\frac{56 \times 56}{6}\right)\right\} - 1800$$

$= 770.67 + 522.67 + 522.67 - 1800 = 1816.01 - 1800 = 16.01$

SS पंक्तियों के मध्य $= \sum \frac{(T_i)^2}{n_i} - \frac{(T)^2}{n}$

$$= \left\{\left(\frac{63 \times 63}{6}\right) + \left(\frac{59 \times 59}{6}\right) + \left(\frac{58 \times 58}{6}\right)\right\} - 1800$$

$$= 661.5 + 580.17 + 560.67 - 1800 = 1802.34 - 1800 = 2.34$$

प्रतिदर्शों के भीतर ss की गणना वर्ग के भीतर इसके माध्य के साथ प्रत्येक पद द्वारा घटाकर करते हैं। [उदाहरण के लिए A & X →(12 +12)/2 = 12, A & Y →(10 + 9)/2 = 9.5 और इसी प्रकार]

SS भीतर =

$(12 -12)^2 + (12 - 12)^2 + (12 -11.5)^2 + (11-11.5)^2 + (10-10.5)^2 + (11-10.5)^2 + (10 -9.5)^2 + (9 -9.5)^2 + (7-7.5)^2 + (8 - 7.5)^2 + (11 - 11)^2 + (11 - 11)^2 + (9 -10)^2 + (11 - 10)^2 + (10 - 10.5)^2 + (11-10.5)^2 + (8 - 7.5)^2 + (7 - 7.5)^2$

$= 0 + 0 + 0.25 + 0.25 + 0.25 + 0.25 + 0.25 + 0.25 +0.25 +0.25 + 0 + 0 + 1 + 1 + 0.25 + 0.25 +0.25 +0.25 = 5$

SS अन्योन्यक्रिया = कुल SS - (SS स्तम्भ + SS पंक्ति + SS भीतर)

$= 46 - (16.01 + 2.34 + 5) = 46 - 23.35 = 22.65$

ANOVA तालिका निम्नवत है-

विचरण के श्रोत	SS	d.f.	MS	F-अनुपात	5% F-सीमा
स्तम्भों के मध्य	16.01	3– 1= 2	$\frac{16.01}{2} = 8$	$\frac{8}{0.56} = 14.28$	F (2,9) = 4.26
पंक्तियों के मध्य	2.34	3– 1 =2	$\frac{2.34}{2} = 1.17$	$\frac{1.17}{0.56} = 2.09$	F (2,9) = 4.26
प्रतिदर्श के भीतर	5.00	18 – 9 = 9	$\frac{5}{9} = 0.56$		
अन्योन्यक्रिया	22.65	17–(2+2+9) =4	$\frac{22.65}{4} = 5.66$	$\frac{5.66}{0.56} = 10.1$	F(4,9) = 3.63
कुल	46	18-1=17			

अन्योन्यक्रिया (10.1) के लिए F अनुपात का परिकल्पित मान इसके तालिका मान (3.63) से अधिक है , इसलिए शून्य परिकल्पना अस्वीकार्य है। इसका अभिप्राय है कि कारों एवं गैसोलीन के ब्रान्डों के मध्य अर्थपूर्ण अन्योन्यक्रिया है, इसलिए स्तम्भ प्रभाव एवं पंक्ति प्रभाव के परिणामों का कोई उपयोग नहीं हैं।

(स) रेखाचित्रीय विधि- यदि आपको दो तरफा ANOVA के पुनरावृत्ति मानों के साथ आवश्यक समस्याओं के साथ व्यवहार करना पडता है, तब आपके पास रेखाचित्रीय विधि का भी विकल्प होता है। इस प्रकार, दो तरफा

प्रारूप में, रेखाचित्रीय विधि विभिन्न कारकों के मध्य अन्योन्यक्रिया का भी अध्ययन किया जा सकता है। रेखाचित्रीय विधि में, एक कारक को x अक्ष में रेखांकित और दूसरे कारक को y अक्ष में रेखांकित किया जाता है। सभी प्रतिदर्शों के लिए माध्यों को बिदुरेख में रेखांकित किया जाता है और पृथक रेखाओं द्वारा तर्कसंगत किया जाता है। यदि प्रत्येक प्रतिदर्श पदों को जोडने वाली रेखाऐं एक दूसरे के वि रुद्ध नहीं होती है, तब इसका सूचक अन्योन्यक्रिया का नहीं होता है, जबकि रेखाऐं एक दूसरे विपरीत (वि रुद्ध) होती है इसका अभिप्राय कारकों के मध्य एक अन्योन्यक्रिया का होना है। प्रत्येक प्रतिदर्श पदों से सम्बन्धित रेखाओं के रेखाचित्रीय निरूपण अन्योन्यक्रिया के प्रकार के बारे में इंगित करता है। उदाहरण के लिए , अन्योन्यक्रिया क्रमवाचक प्रकार की हो सकती है जहॉ एक कारक से सम्बन्धित प्रभाव का श्रेणी क्रम एक समान रहता है। दूसरे कारक के परिवर्तन से सम्बन्धित प्रभावों का श्रेणी क्रम हो तो अन्योन्यक्रिया क्रमवार प्रकार की नहीं हेागी। यह गैर क्रमवार अन्योन्यक्रिया गैर विपरीत शेष या विपरीत शेष प्रकार की हो सकती है।

12.5 सारांश (Summary)

इस इकाई में, आपने F अनुपात और ANOVA के बारे में अध्ययन किया। F परीक्षण दो प्रसरणों के अनुपात पर आधारित होता है। इसका प्रयोग यह निर्धारण करने में किया जाता है कि क्या दो सम्बन्धित प्रतिदर्शों को समान प्रसरण के सामान्य समग्रों से लिया गया है। F परीक्षण सामान्य स्थिति , समरूपता, यादृच्छिकता, त्रुटि की स्वतन्त्रता की अवधारणा पर आधारित होता है। ANOVA कारणों के वर्ग की वजह से विचरण के पृथककरण में से दूसरे वर्गो की वजह से विचरण की एक सांख्यिकीय तकनीक के लिए हैं विशेष रूप से इसे उस परीक्षण के लिए प्रारूपित किया जाता है जहॉ दो से अधिक मात्रात्मक (परिमाणात्मक) समग्रों के माध्य समान होते हैं। यदि आंकडे को केवल एक कारक के अनुसार वर्गीकृत किया जाता है तब एक तरफा ANOVA प्रयोग होता है जबकि यदि आंकडे को दो करकों के अनुसार वर्गीकृत किया जाता है तब दो तरफा ANOVA प्रयुक्त होता है। एक तरफा ANOVA में, F अनुपात की गणना माध्य वर्ग के माध्य और माध्य वर्ग के भीतर अनुपात के रूप में की जाती है। शून्य परिकल्पना की स्वीकार्यता और अस्वीकारिता के सम्बन्ध में निर्णय F के परिकल्पित एक तालिका मान की तुलना के आधार पर लिया जाता है। दो तरफा ANOVA में एक साथ एक कारक से अधिक के प्रभाव का अध्ययन किया जाता है। अपुनरावृत्ति मानों की दशा में , दो F अनुपातों की गणना की जाती है। पहली स्तम्भों के मध्य और दूसरी पंक्तियों के मध्य पुनरावृत्ति मानों की दशा में , उपरोक्त वर्णित दो F अनुपातों के अतिरिक्त, दूसरा F अनुपात भी अन्योन्यक्रिया विचरण की गणना के लिए लिया जाता है।

एकतरफा और दो तरफा ANOVA के लिए स्वीकृत और अस्वीकृत मानदण्ड एक समान रहते हैं। जटिल आंकडों की दशा में, आंकडों के सरलीकरण के लिए सांकेत विधि का प्रयोग किया जा सकता है।

12.6 शब्दावली (Glossary)

- **ANOVA**: प्रसारण का विश्लेषण।
- **SS**: प्रसरण के लिए विचलनों के वर्गो का योग।
- **MS**: औसत (माध्य) वर्ग।
- **अन्योन्यक्रिया प्रभाव**: एक कारक दूसरे कारक के लिए उपचार का प्रभाव।

12.7 बोध प्रश्न (Comprehension Questions)

(अ) रिक्त स्थानों की पूर्ति

1. ‘‘प्रसरण‘‘ शब्द सबसे पहले सांख्यिकीविद् -------द्वारा प्रयोग किया गया था।
2. ANOVA तकनीक प्रारम्भ में ------अनुसंधान में प्रयोग की गई थी।
3. प्रतिदर्श के भीतर माध्य वर्ग गणना के लिए स्वतन्त्रता के अंश में सम्मिलित -----घटाने पर ------।
4. -------विधि ANOVA तकनीक में गणनात्मक कार्य सरलीकरण के लिए प्रयोग होती है।
5. आलेखी विधि के अन्तर्गत ------कारकों के मध्य रेखाओं को पार करने द्वारा इंगित किया जाता है।

(ब) सत्य या असत्य

1. F- परीक्षण के अनुप्रयोग के लिए वर्गों की समरूपता एक आवश्यक अवधारणा होती है। (सत्य/असत्य)
2. F- परीक्षण दो मानक विचलनों के अनुपात पर आधारित है। (सत्य/असत्य)
3. प्रसरण के विश्लेषण परीक्षण का उद्देश्य दो प्रतिदर्श प्रसरणों के मध्य अन्तर के महत्व का परीक्षण करना होता है। (सत्य/असत्य)
4. द्विमार्गी ANOVA पुनरावृत्ति मानों के साथ न होने पर की स्थिति में अन्योन्यक्रिया चिरण की मणना की जाती है। (सत्य/असत्य)
5. (n-k) अवशेष प्रसारण के लिए स्वतन्त्रता के अंश को इंगित करता है। (सत्य/असत्य)

12.8 बोध प्रश्नों के उत्तर (Answers to Comprehension Questions)

(अ) रिक्त स्थानों की पूर्ति

1. आर.ए. फिशर 2. भूमि विषयक 3. कुल प्रतिदर्श आकार प्रतिदर्शों की संख्या

4. सांकेतिक शब्दों में बदलना 5. अन्योन्यक्रिया

(ब) सत्य या असत्य

1. सत्य 2. असत्य 3. असत्य 4. असत्य 5. असत्य

12.9 स्वपरख प्रश्न (Self-Assessment Questions)

1. F- परीक्षण का उद्देश्य क्या होता है?
2. शब्द माध्य वर्ग को परिभाषित करें?
3. संशोधन कारक का सूत्र लिखें?
4. F- परीक्षण की अवधारणाऐं एवं तकनीक का वर्णन करें?
5. प्रसरण के विश्लेषण का अर्थ समझाइयें। द्विमार्गी वर्गीकरण के लिए ANOVAकी तकनीक का संक्षेप में वर्णन करें?
6. दो यादृच्छिक प्रतिदर्शों को दो सामान्य समग्रों में से लिया गया है:

प्रतिदर्श 1	75	68	65	70	84	66	55
प्रतिदर्श 2	42	44	56	52	46	....	....

5 प्रतिशत महत्व के स्तर में प्रसरण अनुपात का प्रयोग करते हुए परीक्षण करें कि दो समग्रों में समान प्रसरण है।

(F = 2.37, H_0: स्वीकार्य)

7. 8 अवलोकनों के एक प्रतिदर्श में, माध्य में से पदो के विचलनों के वर्गों का 84.4 था। 10 अवलोकनों के एक दूसरे प्रतिदर्श में ये मान 102.6 पाया गया था। 5 प्रतिशत स्तर पर परीक्षण करें कि अन्तर महत्वपूर्ण है। आपको 5 प्रतिशत का स्तर दिया गया है , $v_1 = 7$ एवं $v_2 = 9$ स्वतन्त्रता की श्रेणीयों के लिए F का समीक्षात्मक मान 3.29 है और $v_1 = 8$ एवं $v_2 = 10$ स्वतन्त्रता श्रेणीयों के लिए यह मान 3.07 है। (F=1.06, H_0 : स्वीकार्य)

8. सामान्य समग्र में से समान प्रसरणों के साथ नीचे तीन प्रतिदर्श प्राप्त किये गये। परिकल्पना परीक्षण करें कि प्रतिदर्श माध्य एक समान है-

8	7	12
10	5	9
7	10	13
14	9	12
11	9	14

$v_1 = 2$ एवं $v_2 = 12$ के लिए 5 प्रतिशत महत्व के स्तर पर F का तालिका मान 3.88 है।

(F =4, H_0 : अस्वीकार्य)

9. एक कम्पनी यह जानने के लिए इच्छुक है कि क्या तीन बिक्रीकर्ता एक समान प्रर्दशन कर रहे हैं। तीन बिक्रीकर्ताओं का साप्ताहिक बिक्री अभिलेख है।

A (Rs)	B (Rs)	C (Rs)
300	600	700
400	300	300
300	300	400
500	400	600
0	....	500

10. तीन प्रयोग शक्ति के प्रतिदर्श के संतुष्टि आर्द्रता का निर्धारण करते हैं, प्रत्येक व्यक्ति प्रत्येक 4 प्रेषणों में से एक प्रतिदर्श लेता है। परिणाम निम्नवत है।

प्रयोग	प्रेषण			
	I	II	III	IV
A	9	10	9	10
B	12	11	9	11
C	11	12	10	12

इन आकडों से प्रसरण का विश्लेषण निर्धारित करें और वर्णन करें कि प्रेषणों के मध्य या प्रयोगों के मध्य कोई महत्वपूर्ण अन्तर है प्रेषणों के मध्य (F = 4.02, H_0: स्वीकृत), प्रयोगों के मध्य F=6.91, H_0 अस्वीकृत

11. निम्नलिखित तीन ड्रगों के से सम्बन्धित सूचना के आधार पर ANOVA तालिका का निर्माण करें और यह परीक्षण करें कि ड्रगों की प्रभावशीलता तीन विभिन्न वर्गों के व्यक्तियों रक्तचाप को कम करती है।

व्यक्ति का वर्ग	ड्रग		
	X	Y	Z
A	14	10	11
	15	9	11
B	12	7	10
	11	8	11
C	10	11	8
	11	11	7

a) क्या ड्रग्स भिन्न प्रकार से प्रतिक्रिया करते हैं?

b) क्या भिन्न वर्ग के व्यक्ति भिन्न प्रकार से प्रभावित हैं?

c) क्या अन्योन्यक्रिया पद महत्वपूर्ण है? (F=36.9, 19.1, 18.78) सभी तीन H_0 अस्वीकार है।

12.10 संदर्भ पुस्तकें (Reference Books)

- गुप्ता एस. पी., "सांख्यिकीय विधियाँ सुल्तान चन्द एवं सन्स, नई दिल्ली
- दास एन.जी. सांख्यिकीय विधियाँ टाटा मेंग्रो हिल, नई दिल्ली
- बाजपेई नवल व्यवसाय सांख्यिकीय पियरसन

इकाई 13 कंप्यूटर के आधार - इतिहास, भाषा और घटक (Basics of Computer - History, Language and Components)

13.1 परिचय (Introduction)

13.2 उद्देश्य (Objectives)

13.3 कंप्यूटर का इतिहास (History of Computers)

(कंप्यूटर क्या है, प्राचीन गणना पद्धतियां, अबेकस और अन्य गणना प्रणालियां, कंप्यूटर का विकास, कंप्यूटर की पीढ़ियां, कंप्यूटर के प्रकार, पर्सनल कंप्यूटर का विकास, कंप्यूटर के गुण-उपयोग)

13.4 कंप्यूटर के बुनियादी अवयव (Basic Components)

(प्रोसेसिंग, इनपुट, आउटपुट, मेमोरी, प्रोग्राम)

13.5 कंप्यूटर की कार्यपद्धति (Working Process)

(डाटा, सूचना, हार्डवेयर, सॉफ्टवेयर)

13.6 कंप्यूटर लैंग्वेज (Computer Languages)

(प्रोग्रामिंग भाषाएं, बाइट, बाइनरी संख्याएं, कोडिंग सिस्टम, कम्पाइलर)

13.7 उपसंहार (The Conclusion)

13.8 अभ्यास प्रश्न (Exercise)

13.9 निबंधात्मक प्रश्न (Theoretical Question)

13.1 परिचय (Introduction)

समाजशास्त्र और कंप्यूटर पहली नजर में ये दोनों शब्द एक दूसरे से जुड़े हुए प्रतीत नहीं होते हैं। किन्तु वैज्ञानिक प्रगति के दौर में कंप्यूटर जिस तरह मानव जीवन का अभिन्न अंग बनकर रह गया है, उससे सामाजिक अभिरचना को जानने समझने में भी कई मायनों में मदद मिली है। शिक्षा हो , स्वास्थ्य होए सुरक्षा हो , बैंकिंग हो या कोई भी अन्य क्षेत्र , मानव जीवन का शायद ही कोई पहलू आज कंप्यूटर से अछूता रह गया हो। वस्तुतः कंप्यूटर आज मानव जीवन के दैनिक कार्यों की सबसे बड़ा सहायक मशीन बन गया है। यही वजह है कि सामाजिक शोध कार्यों में भी कंप्यूटर और कंप्यूटर एप्लीकेशन का इस्तेमाल आज आवश्यक है।

13.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के उपरान्त आप-

- ✓ मानव जीवन के विकास के साथ किस तरह गणना के उपकरण विकसित हुए, को जानेंगे।
- ✓ कंप्यूटर क्या है, किस तरह इस बहुउपयोगी मशीनों के विकास हुआ।
- ✓ कंप्यूटर किस तरह काम करता है और इसके प्रमुख अवयव क्या हैं।
- ✓ कंप्यूटर एप्लीकेशन क्या हैं और इनका महत्व क्या है।

13.3 कंप्यूटर का इतिहास (History of Computers)

विकास के लंबे अनुक्रम में मनुष्य ने जीवन के नये पहलुओं की खोज अपने अनुभवों के आधार पर की। इन्हीं खोजों में शामिल थी गणनाएं। यूं तो मानव मस्तिष्क स्वयं में सूचनाओं को सुरक्षित रखने का अथाह भंडार है, किन्तु जब प्रश्न गणनाओं और गणनाओं में भी त्वरित गणनाओं का आता है तो मस्तिष्क कुछ पीछे रह जाता है। शायद यही वह कारण रहा होगा , जिसने प्राचीन काल से ही मनुष्य को ऐसे तरीके ईजाद करने के लिए प्रेरित किया हो , जो गणनाओं को चुटकियों में हल कर सकें। प्राचीन गणना पद्धतियों अबेकस से लेकर कैलकुलेटर और फिर कंप्यूटर तक की यात्रा भी इसी प्रेरणा का परिणाम है।

- **कंप्यूटर क्या है (What is Computer)**

कंप्यूटर शब्द की उत्पत्ति अंग्रेजी शब्द कंप्यूट (Compute) से हुई है, जिसका अर्थ गणना करना है। यही वजह है कि हिन्दी में इस उपकरण को गणक या संगणक भी कहा जाता है। अपने विकास की शुरूआत में कंप्यूटर का इस्तेमाल मुख्यतः जटिल गणनाओं में ही किया जाता रहा , लेकिन कालान्तर में ज्यों-ज्यों मानवीय आवश्यकताएं बढ़ती गईं , कंप्यूटर का स्वरूप भी बहुआयामी (Multitasking) होता चला गया। आज हम कंप्यूटर पर गाने सुन सकते हैंए वीडियो देख सकते हैं, इसके जरिये इंटरनेट पर दुनियाभर की खबरें एक चुटकी में हासिल कर सकते हैं, चिकित्सकीय सुविधाएं हासिल कर सकते हैं, शिक्षा प्राप्त कर सकते हैं और हर वो काम कर सकते हैं जो हम चाहते हैं। यानी कंप्यूटर वह मशीन है , जो वर्तमान दौर में हमारे जीवन को सरल और अधिक सक्षम बनाती है।

- **कंप्यूटर का विकास (Development of Computer)**

यद्यपि मानव सभ्यता के विकास के साथ ही गणनाओं के भी प्रमाण मिलते रहे हैं। हजारों वर्ष पहले अंगुलियों की मदद से गणनाओं की जानकारी मध्यपूर्व एशिया , यूरोप की कई सभ्यताओं में मिलती है , लेकिन उपकरणों की मदद से कंप्यूटर के विकास की यात्रा को जानने.समझने के लिए हमें करीब तीन हजार साल पीछे लौटना

होगा। मानव जीवन में गणनाओं का विशेष महत्व रहा है , लेकिन यह पहले ही स्पष्ट हो चुका है कि मानव मस्तिष्क जटिल गणनाओं का त्वरित हल निकाल पाने में सक्षम नहीं है। ऐसे में गणनाओं के लिए किसी उपकरण की आवश्यकता महसूस की जाने लगी। ऐतिहासिक साक्ष्यों के अनुसार चीनी वैज्ञानिकों ने करीब तीन हजार साल पूर्व पहला ऐसा उपकरण बनाया, जो गणनाओं को मानव के लिए सुगम और सरल बनाने में सफल रहा। यह उपकरण था अबेकस (Abacus) इसे हम निम्न चित्र के जरिये जान सकेंगे। अबेकस में लकड़ी या लोहे के फ्रेम में कुछ लोहे की छड़ें होती हैंए जिनमें लकड़ी की बनी गोलियां लगाई जाती थीं। इन गोलियों को इसे इस्तेमाल करने वाला व्यक्ति उपर.नीचे करके आसानी से गणनाएं कर सकता था। आज भी नन्हे स्कूली बच्चों को गणनाओं का प्रारंभिक पाठ पढ़ाने में अबेकस की मदद ली जाती है। हालांकिए इसकी मदद से सिर्फ छोटी गणनाएं ही कर पाना संभव है। फिर भी यही वह उपकरण थाए जो मौजूदा कंप्यूटर के आविष्कार की बुनियाद बना। इस लिहाज से अबेकस को पहला कंप्यूटर का दर्जा दिया जाता है।

(प्राचीन अबेकस, जिसकी मदद से गणनाएं की जाती थीं)

अबेकस के बाद गणनाओं के लिए एक नया उपकरण ईजाद हुआ सन 1617 में। स्कॉटलैंड के गणितज्ञ नेपियर ने एक गणितीय उपकरण बनाया, जो दिखने में अबेकस की तरह ही था। अंतर सिर्फ यह था कि इसमें गोलियों के बजाय छड़ें ही फ्रेम में लगी होती थीं। खासियत यह थी कि इन छड़ों पर अंक लिखे होते थे , जिनकी मदद से गणनाएं की जा सकती थीं। इसके कुछ ही समय बाद 1642 में एक और नये उपकरण का आविष्कार अपने दौर के महान फ्रांसीसी गणितज्ञ ब्लेज पास्कल ने किया। इस उपकरण का नाम पास्कल के नाम पर ही पास्कलाइन (Pascaline) रखा गया। यह अबेकस और नेपियर बोन से अधिक तेजी से गणना करने में सक्षम था। हालांकिए अब भी गुणा और भाग की गणनाएं करना संभव नहीं हो सका था। ऐसे में सन 1671 में जर्मन वैज्ञानिक गडॉफ्रिट लेन्ज ने पास्कलाइन को ही परिष्कृत (Modified) किया, जिसका परिणाम लेन्ज कैल्कुलेटर के रूप में

सामने आया। इसकी खासियत यह थी कि इसमें जोड़ और घटाने के अलावा गुणा.भाग जैसी जटिल गणनाएं भी आसानी से कर पाना संभव हुआ।

हालांकि, समय के साथ जिस तेजी से मानव सभ्यताएं विकसित होती गईं और हर रोज नई खोजों के लिए जटिलतम गणनाएं सामने आती रहीं, अबेकस की तरह पास्कलाइन भी अनुपयोगी लगने लगा। ऐसे में सन सर चार्ल्स बैबेज एनालिटिकल इंजन (Analytical Engine) नाम का उपकरण सामने लाए। यह कहीं अधिक तेजी से और त्रुटिरहित गणनाएं करने में सक्षम था।

सबसे बड़ी बात यह थी कि इस मशीन में गणनाओं को सुरक्षित भी रखा जा सकता था। स्टोरेज के लिए इसमें पंचकार्ड का इस्तेमाल किया जाता था। यह 25 हजार छोटे पुर्जों से बना करीब 15 टन वजनी और आठ फीट ऊँचा उपकरण था। भारी भरकम स्वरूप की वजह से हर किसी के लिए इसका इस्तेमाल करना न तो सरल थाए न ही संभवए लेकिन एनालिटिकल इंजन ही वह रास्ता बनाए जो आगे चलकर कंप्यूटर पर खत्म हुआ। यही कारण है कि **सर चार्ल्स बैबेज** को ही कंप्यूटर के जनक के तौर पर जाना जाता है। कालान्तर में सर बैबेज के ही डिजाइन किए उपकरण में निरन्तर सुधार किए जाते रहे और आज का कंप्यूटर विकसित होता गया। अब भी कंप्यूटर की दुनिया में लगातार खोज और सुधार जारी हैं।

- **कंप्यूटर की पीढ़ियां (Generations of Computer)**

सर चार्ल्स बैबेज ने जो एनालिटिकल इंजन पेश किया था , वह गणनाओं में खासा सहायक साबित हुआ , लेकिन चूंकि समय के साथ परिवर्तन आवश्यक है , निरन्तर गणनाओं का दायरा और सूचनाओं को सुरक्षित रखने की जरूरत महसूस की जाने लगी। सर बैबेज के एनालिटिकल इंजन से आधुनिक कंप्यूटर के विकास का सफर शुरू हुआ। इस लिहाज से सामान्यतः कंप्यूटर के विकास को पीढ़ियों में भी बांटकर देखा जाता है। पहली पीढ़ी से लेकर आज के दौर के कंप्यूटर यानी पांचवीं पीढ़ी तक।

i. पहली पीढ़ी (First Generation)

सन 1946 में दुनिया का पहला इलेक्ट्रॉनिक कंप्यूटर अस्तित्व में आया। दो वैज्ञानिकों जे.पी एकर्ट और जे.डब्ल्यू. मॉशी इस कंप्यूटर के आविष्कारक थे। दोनों ने अपने इस कंप्यूटर को नाम दिया ENIAC (Electronic Numerical Integrated and Calculator) लेकिन यह कंप्यूटर बहुत अधिक भारी था। उस वक्त इस कंप्यूटर का वनज करीब 30 टन था। दोनों वैज्ञानिकों ने इस कंप्यूटर में आंकड़ों के संग्रहण के लिए वैक्यूम ट्यूबों का इस्तेमाल कियाए लेकिन कमी यह थी कि वैक्यूम ट्यूब की कार्यक्षमता बहुत अधिक नहीं थी। इसके अलावा इस कंप्यूटर को ठंडा रखने के लिए काफी बड़े कूलिंग सिस्टम (Cooling System) की भी जरूरत पड़ती थी। पहली पीढ़ी के कंप्यूटर के कालखंड को 1946 से 1959 तक बांटकर देखा जा सकता है।

(पहली पीढ़ी का कंप्यूटर ENIAC)

ii. दूसरी पीढ़ी (Second Generation)

समय के साथ आते गए बदलावों के फलस्वरूप दूसरी पीढ़ी के कंप्यूटर अस्तित्व में आए। इस पीढ़ी के कंप्यूटरों का कालखंड 1959 से 1964 रहा। इस पीढ़ी के कंप्यूटरों की खासियत यह थी कि इसमें आंकड़ों के संग्रहण के लिए भारी भरकम वैक्यूम ट्यूबों के स्थान पर ट्रांजिस्टर (Transistors) का उपयोग किया गया। ट्रांजिस्टर वैक्यूम ट्यूब के मुकाबले आकार में भी काफी छोटे थेए लिहाजा कंप्यूटर का स्वरूप और वजन पूर्ववर्ती पीढ़ी के सापेक्ष काफी कम हो गया। दूसरी ओर, ट्रांजिस्टर की गणनात्मक कार्यक्षमता और आंकड़ों को सुरक्षित रखने की क्षमता भी एनिआक के मुकाबले काफी बेहतर थी।

iii. तीसरी पीढ़ी (Third Generation)

सन 1964 में तीसरी पीढ़ी के कंप्यूटरों की खोज हुई। इस पीढ़ी के कंप्यूटरों की विशेषता यह थी कि इसमें इंटीग्रेटेड सर्किट (Integrated Circuit : IC) का इस्तेमाल कंप्यूटर के प्रमुख इलेक्ट्रॉनिक घटक के रूप में किया गया था। आईसी की खोज और कंप्यूटर में इसका इस्तेमाल आगे चलकर माइक्रो इलेक्ट्रॉनिक्स (Micro Electronics) का जरिया बना। वैज्ञानिक टीएस बिल्की की खोज आईसी की सबसे बड़ी खासियत इसका बेहद छोटा आकार, लेकिन संग्रहण की अकूत क्षमता थी। इसके अलावा इसमें पहले के मुकाबले कई गुना अधिक और कहीं ज्यादा तेजी से गणनाएं करने की क्षमता भी थी। तीसरी पीढ़ी के कंप्यूटरों का कालखंड (Time Period) 1965 से 1971 रहा।

iv. चौथी पीढ़ी (Fourth Generation)

चौथी पीढ़ी के कंप्यूटर वह हैं , जिनका इस्तेमाल हम आज करते हैं। इस पीढ़ी के कंप्यूटरों की खासियत इनमें इस्तेमाल किया जाने वाला माइक्रो प्रोसेसर (Micro Processor) है। 1971 में अमेरिका के वैज्ञानिक टेड हॉफ (Tedd Hoff) को माइक्रो प्रोसेसर की ईजाद का श्रेय जाता है। टेड तब कंप्यूटर निर्माता कंपनी इनटेल में काम करते थे और उन्होंने अपने माइक्रोप्रोसेसर को इनटेल- 4004 नाम दिया। माइक्रोप्रोसेसर दरअसल एक सिंगल चिप हैए जिसमें आंकड़ों को सुरक्षित रखा जा सकता है। इसके इस्तेमाल से कंप्यूटरों का न सिर्फ आकार छोटा हुआए बल्कि इनकी कार्यक्षमता भी बढ़ी। इस पीढ़ी का कालखंड 1971 से 1980 रहा।

v. पांचवीं पीढ़ी (Fifth Generation)

1980 से आज के दौर तक इस्तेमाल किए जाने वाले कंप्यूटरों को पांचवीं पीढ़ी में शामिल किया जाता रहा है। कुछ विद्वान आज के कंप्यूटरों को भी चौथी पीढ़ी का ही कंप्यूटर मानते हैं तो कुछ ने इन्हें पांचवीं पीढ़ी में रखा है। कंप्यूटरों को चौथी पीढ़ी का ही मानने की बड़ी वजह यह है कि मौजूदा कंप्यूटरों का मूलाधार माइक्रोप्रोसेसर ही है, लेकिन इन्हें पांचवीं पीढ़ी में रखने वाले यह मानते हैं कि माइक्रोप्रोसेसर की क्षमताओं और आकार में भी लगातार बदलाव आते रहे हैं।

इसके अलावा प्रोसेसर अब सिर्फ कंप्यूटर तक ही सीमित नहीं रह गया है , बल्कि मोबाइल स्मार्टफोन के जरिये ये मनुष्य के हाथों में समाहित हो जाने वाला उपकरण बन चुका है। कंप्यूटर की भावी पीढ़ी की बात करें तो वैज्ञानिक इस तरह के कंप्यूटर बनाने की दिशा में प्रयास कर रहे हैं जो कृत्रिम बुद्धि (Artificial Intelligence) से लेस हो। इस दिशा में निरन्तर शोध किए जा रहे हैं। रोबोट को कुछ हद तक इस श्रेणी में रखा जा सकता है , लेकिन वह भी उतने ही काम करता हैए जितने का उसे निर्देश दिया जाता है।

वैज्ञानिकों की सोच यह है कि ऐसे कंप्यूटर बनाए जाएं जो आवश्यकता के अनुरूप स्वतः निर्णय ले सके और आंकड़ों.सूचनाओं का इस्तेमाल कर खुद ही अपेक्षित परिणाम दे सके। हालांकि, यह बिन्दु इस लिहाज से विवाद का विषय भी बनता रहा है कि यदि कंप्यूटर स्वतः बुद्धि-विवेक से काम करने लगेगा तो मनुष्य उस पर नियंत्रण कैसे रख सकेगा। और यदि अनहोनीवश कृत्रिम बुद्धि-विवेकयुक्त कंप्यूटर नकारात्मक दिशा में चलने लगा तो यह विनाशकारी साबित हो सकता है।

• कंप्यूटर के प्रकार (Types of Computer)

कंप्यूटर का मुख्य कार्य उन आंकड़ों को सुरक्षित रखना है , जो इसे इस्तेमाल करने वाला व्यक्ति (user) कंप्यूटर को उपलब्ध कराता है। कंप्यूटर उपयोगकर्ता के निर्देशों के आधार पर इन आंकड़ों का उपयोग कर परिणाम देता है। कार्यक्षमता के आधार पर कंप्यूटर को इन श्रेणियों में बांटा गया हैः सुपर कंप्यूटर , मेनफ्रेम कंप्यूटर, मिनी कंप्यूटर और माइक्रो कंप्यूटर (Micro Computer)। इन सभी श्रेणियों पर नजर डालें तो सुपर कंप्यूटर सर्वोच्च श्रेणी का माना जाता हैए जबकि माइक्रो कंप्यूटर सबसे छोटी। आइए अब हर श्रेणी को कुछ विस्तार से समझते हैं।

1) सुपर कंप्यूटर (Super Computers)

सुपर कंप्यूटर, कंप्यूटरों की लंबी श्रृंखला में सबसे तेज गति से काम करने वाले कंप्यूटर हैं। कल्पनातीत डाटा को यह न्यूनतम समय में सूचनाओं में बदलने में सक्षम हैं। इनका इस्तेमाल सामान्यतः बेहद बड़ी गणनाओं में ही किया जाता है। कंप्यूटर का प्रयोग मौसम की भविष्यवाणी , मिसाइलों के डिजाइन जैसे जटिल कार्यों में किया जाता है। सुपर कंप्यूटरों में कई माइक्रो प्रोसेसर (Micro Processor) लगे होते हैं। यह एक प्रकार की बेहद छोटी मशीन है जो कम्प्यूटिंग यानी गणना के कार्य को बेहद कम समय में कर पाने में सक्षम है। भारत में

विकसित सुपर कंप्यूटर का नाम परम है। निम्नवत चित्र से समझा जा सकता है कि सुपर कंप्यूटर दरअसल , कई सारे प्रोसेसर का एक सामूहिक स्वरूप है।

यहां यह सवाल उठना लाजिमी है कि प्रोसेसर किस तरह गणना में मदद करते हैं। दरअसल , किसी जटिल गणना को कम समय में पूरा करने के लिये बहुत से प्रोसेसर एक साथ काम करते हैं। इस प्रक्रिया को समान्तर प्रोसेसिंग (Parallel Processing) कहा जाता है। इसके तहत कंप्यूटर को मिलने वाले डाटा अलग.अलग काम के लिए अलग.अलग प्रोसेसर को बांट दिए जाते हैं। हर प्रोसेसर अपने हिस्से की गणना करने के बाद कंप्यूटर को सूचना उपलब्ध कराता है और कंप्यूटर सभी प्रोसेसर से मिलने वाली सूचनाओं को एकत्र कर लेने के बाद सटीक अंतिम परिणाम उपलब्ध करा देता है।

(सुपर कंप्यूटर)

2) मेनफ्रेम कंप्यूटर (Mainframe Computers)

मेनफ्रेम कंप्यूटर कार्यक्षमता के लिहाज से सुपर कंप्यूटर से कुछ कमतर , लेकिन फिर भी काफी अधिक क्षमतावान होते हैं। इसकी कार्यक्षमता का अंदाजा इसी से लगाया जा सकता है कि मेनफ्रेम कंप्यूटरों पर एक ही समय में 250 से अधिक लोग एकसाथ काम कर सकते हैं। इन कंप्यूटरों का इस्तेमाल बल्क डाटा (Bulk Data) की प्रोसेसिंग में किया जाता है। यानी ऐसी जगहों पर ये कंप्यूटर प्रयुक्त होते हैंए जहां एक ही समय में भारी मात्रा में और निरन्तर गणनाओं की जरूरत होती है। मुख्यतः इस तरह के कंप्यूटर बड़ी कंपनियों में उपभोक्ताओं की जानकारी सुरक्षित रखने में , जनगणना और इसी तरह के अन्य ऐसे कार्यों में इस्तेमाल किए जाते हैं , जहां भारी डाटा आता है।

3) मिनी और माइक्रो कंप्यूटर (Mini, Micro Computers)

मिनी कंप्यूटर, मेनफ्रेम कंप्यूटरों से छोटे लेकिन माइक्रो कंप्यूटरों से बड़े होते हैं। माइक्रो कंप्यूटरों को पर्सनल कंप्यूटर (Personal Computers, PC) भी कहा जाता है। पर्सनल कंप्यूटर कंप्यूटरों की श्रृंखला में आकार के लिहाज से सबसे छोटे होते हैं। पर्सनल कंप्यूटर का विकास सबसे पहले 1981 में हुआ था। आगे हम इसे विस्तार से समझेंगे। माइक्रो या पर्सनल कंप्यूटर के अन्य प्रकारों को इस तरह समझ सकते हैं।

डेस्कटॉपः वह कंप्यूटर जिसे मेज पर रखकर काम किया जा सके

लैपटॉपः ऐसा कंप्यूटरए जिसे उपयोगकर्ता गोद में रखकर काम करे

पामटॉपः वह कंप्यू टर जो उपयोगकर्ता की हथेली में समा सके , इस श्रेणी में स्मार्टफोन (smart phones) म्यूजिक प्लेयर, वीडियो प्लेयर, टैबलेट रखे जा सकते हैं

• पर्सनल कंप्यूटर का विकास (Development of PCs)

कंप्यूटर की शुरूआत के साथ ही इनका आकार बेहद बड़ा था, जो कालान्तर में जरूरत के हिसाब से छोटा होता गया। समय के साथ कंप्यूटर में आते गए इन बदलावों ने कंप्यूटर को सिर्फ गणनाएं करने वाली मशीन के बजाय एक समय में एकसाथ कई काम करने वाला उपकरण बना दिया। इससे यह मनुष्य के दैनिक जीवन के लिए लगातार उपयोगी बनता गयाए लेकिन सबसे बड़ी समस्या यह थी कि आम आदमी कैसे करोड़ों का सुपर कंप्यूटर इस्तेमाल करे। इस जवाब के तलाश में 1970 में माइक्रो प्रोसेसर का आविष्कार हुआ। यही माइक्रो प्रोसेसर आगे चलकर माइक्रो कंप्यूटरों की खोज का जरिया बने। कंप्यूटर निर्माता कंपनी आईबीएम ने वर्ष 1981 में पहला पर्सनल कंप्यूटर बनाने की घोषणा की , जिसे आई.बी.एम.पी.सी नाम दिया गया। यह कंप्यूटर प्रारंभिक रूप से मुख्यतः शौकिया बनाया गया थाए लेकिन यह इस कदर लोकप्रिय हुआ कि बाद में सभी कंप्यूटर निर्माता कंपनियों का ध्यान पीसी की ओर गया।

(पर्सनल कंप्यूटर या माइक्रो कंप्यूटर)

खास बात यह है कि अब दुनियाभर में सैकड़ों कंपनियों के पर्सनल कंप्यूटर बाजार में हैंए लेकिन वे सभी आई.बी.एम.पीसी कंपेटीबल (IBM-PC Compatible) ही होते हैं। इसका अर्थ यह है कि ये सभी पर्सनल कंप्यूटर आकार, संरचना, हार्डवेयर आदि में आईबीएम.पीसी के समान ही होते हैं। इस तरह आईबीएम.पीसी स्वतः कंप्यूटर निर्माता कंपनियों के लिए एक मानक (Standard) बन गया है। समय के साथ पर्सनल कंप्यूटरों की क्षमताओं में भी लगातार बदलाव होते आए हैं। 1981 में पहले पर्सनल कंप्यूटर के जन्म के बाद से अब तक पीसी की कई पीढ़ियां सामने आ चुकी हैं। इनमें पीसी.पेंटियम , पीसी.कोर 2, इंटेल आई सीरीज प्रमुख हैं। सभी कंप्यूटर सामान्यतः एकसमान होते हैं , लेकिन हर श्रेणी और पीढ़ी में अंतर सिर्फ इसकी संग्रहण क्षमता (Storage Power) और प्रोसेसर (Processor) का होता है।

कंप्यूटर के गुण-उपयोग(Qualities-Uses of Computers)

कंप्यूटर आज के प्रतिस्पर्धी और वैज्ञानिक युग में सिर्फ गणनाओं को चुटकी में हल कर देने भर का साधन नहीं रह गया है वरन यह आज मनोरंजन , शिक्षा, चिकित्सा, सुरक्षा का भी बड़ा माध्यम बन चुका है। कंप्यूटर के गुणों की बात करें तो यह किसी भी काम को बहुत तेज गति से करने वालाए उपयोगकर्ता की ओर से मिलने वाले निर्देशों का अपेक्षित पालन करने वालाए जितना निर्देश दिया जाए , उतना ही काम करने वाला, हर काम को त्रुटिरहित करने वाला , आंकड़ों के आंकड़ों के असीमित भंडार को कम से कम जगह में संग्रह करके रखने वाला और जरूरत पड़ने पर अभीष्ट आंकड़ों को तुरंत उपलब्ध कराने वाला उपकरण है। इस लिहाज से यह मौजूदा मानव जीवनशैली में मानव का सबसे बड़ा सहायक उपकरण बन जाता है। दूसरी ओर , यदि कंप्यूटर के उपयोगों की बात की जाए तो इस लिहाज से भी यह अपने पूर्ववर्ती उपकरणों से कहीं आगे निकल चुका है। इसके दैनिक के कार्यों में होने वाले उपयोग निम्नवत हैं-

ईमेल (Email) - ईमेल इलेक्ट्रॉनिक मेल (Electronic Mail) का संक्षिप्त रूप है। ईमेल का तात्पर्य उस मेल यानी पत्र से है, जिसे हम कंप्यूटर पर लिखकर इंटरनेट के माध्यम से किसी को भेजते हैं। सामान्य डाक प्रक्रिया से इतर यह पूरी प्रक्रिया चंद सेकंडों की होती है। इसके लिए उपयोगकर्ता को एक ईमेल पते की आवश्यकता होती है जो उपयोगकर्ता (user) और मेल सुविधा देने वाली कंपनी के डोमेन नेम (Domain name) का संयुक्त स्वरूप होता है। उदाहरण के लिए- xyzsharma@gmail.com

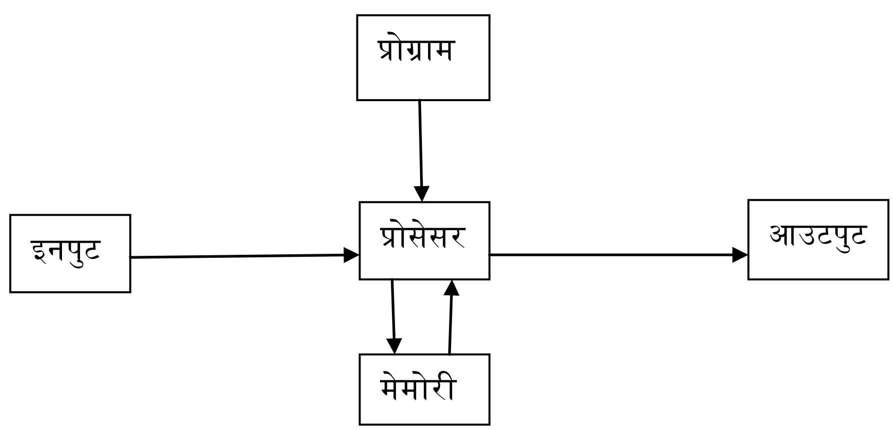
जानकारी संजोना एवं सहयोगः कंप्यूटर उपयोगकर्ता (user) के लिए सहयोगी की तरह काम करता है। वह उपयोगकर्ता की ओर से मिलने वाले निर्देशों का पालन करने के साथ ही जरूरत के अनुरूप जानकारी, सूचनाएं, आंकड़े उपलब्ध कराता है। इस तरह यह उपयोगकर्ता के लिए एक चुटकी में दुनियाभर की जानकारी देने का जरिया बन जाता है।

शिक्षा एवं संचार सुविधाः शिक्षा के क्षेत्र में कंप्यूटर आज के दौर में अति आवश्यक तत्व बन गया है। स्कूल से लेकर विश्वविद्यालयी शिक्षा तक शायद ही शिक्षा का कोई हिस्सा हो, जहां कंप्यूटर का इस्तेमाल नहीं होता हो। दूरस्थ शिक्षा के क्षेत्र में तो कंप्यूटर के सहयोग से क्रान्ति आई है। दुनिया के किसी भी कोने में बैठा शिक्षक आज इंटरनेट के जरिये छात्रों को पढ़ाने में सक्षम हो सका है। दूसरी ओर , संचार सुविधाएं भी कंप्यूटर की मदद से तेजी से विकसित हुई और बढ़ी हैं। वह चाहे ईमेल हो या स्मार्टफोन , सबका विकास कंप्यूटर सिस्टम के जरिये ही हो पाना संभव हो सका है। इससे कुछ पीछे जाएं तो टेलीफोन के दौर में एसटीडी और आईएसडी कॉल की शुरूआत का श्रेय भी कंप्यूटर क्रान्ति को ही जाता है।

शोध, स्वास्थ्यः कंप्यूटर शोधार्थियों के लिए अहम उपकरण है। वस्तुतः शोध कार्यों में एकत्र होने वाले डाटा , आंकड़ों को संग्रहित कर सूचनाओं का संकलन करने में यह शोधार्थी का सबसे बड़ा सहायक बन जाता है। दूसरी ओर, स्वास्थ्य सुविधाओं के क्षेत्र में भी कंप्यूटर मददगार साबित हुआ है। सीटी स्कैन हो या अल्ट्रासाउंड या एमआरआई चिकित्सा क्षेत्र में निरन्तर नये बदलावों के जरिये कंप्यूटर मानव जीवन को स्वस्थ बनाने में सहायक बना है। और अब तो टेलीमेडिसिन चिकित्सा विधा की समग्र शाखा के तौर पर सामने आई है। इसके तहत डॉक्टर दुनिया के किसी भी कोने में रहकर मरीज का इलाज कर पाने में सक्षम हुए हैं।
सुरक्षा एवं अन्य सुविधाएं- कंप्यूटर मनुष्य जीवन के अहम बिन्दु सुरक्षा के लिहाज से खासे मददगार साबित हुए हैं। आम जनजीवन में क्लोज सर्किट कैमरे (Close Circuit Cameras) हों या सैन्य जीवन में अत्याधुनिक उपकरण, रडार और स्वचालित हथियार , सभी कुछ कंप्यूटरीकृत तकनीक पर आधारित हैं। इसके अलावा सड़कों पर यातायात व्यवस्था को सुगम-सुचारू बनाए रखने वाली ट्रैफिक लाइटें हों या एक कॉल पर घायलों को अस्पताल पहुंचाने वाली 108 एंबुलेंस या फिर आपराधिक वारदातों की त्वरित सूचनाएं पुलिस तक पहुंचाने वाला 100 नंबर, सभी जगह कंप्यूटर ही मूल तकनीकी बुनियाद के तौर पर नजर आता है।

13.4 कंप्यूटर के बुनियादी अवयव (Basic Components)

कंप्यूटर चाहे सुपर हो या माइक्रो यानी पर्सनल, हर कोई पांच प्रमुख भागों से मिलकर तैयार होता है, इन भागों को हम कंप्यूटर के बुनियादी अवयव भी कह सकते हैं। ये पांचों हैं- इनपुट (Input), आउटपुट (Output), प्रोसेसर (Processor), मेमोरी (Memory) और प्रोग्राम (Programme) कंप्यूटर की संरचना में इन पांचों का विशेष महत्व है। निम्नवत ग्राफ की मदद से हम इनके कार्य को समझ सकते हैं-



1. प्रोसेसर (Processor)

जैसा कि नाम से ही स्पष्ट हो रहा है कि प्रोसेसर कंप्यूटर का वह हिस्सा होगा , जहां प्रोसेसिंग (Processing) यानी पूरी प्रक्रिया चलती होगी। इस लिहाज से प्रोसेसर को कंप्यूटर का सर्वाधिक महत्वपूर्ण भाग माना जा सकता है, या इसे यूं भी कहा जा सकता है कि प्रोसेसर ही दरअसल असल कंप्यूटर है , बाकि के सभी भाग तो प्रोसेसर की ओर से किए जा रहे कार्यों को सफलतापूर्वक पूर्ण करने में सहायक हैं। ग्राफ से भी यह आसानी से समझ में आता है कि कंप्यूटर के सभी भाग सीधे तौर पर प्रोसेसर से ही जुड़े हुए हैं। इसे यूं भी कहा जा सकता है कि प्रोसेसर कंप्यूटर का दिमाग है , जिस तरह मनुष्य का दिमाग उसे सोचने.समझने , तर्क करने या किसी समस्या का हल निकालने की क्षमता प्रदान करता है , ठीक उसी तरह प्रोसेसर भी कंप्यूटर को मिलने वाले

निर्देशों का सही हल निकालने का काम करता है। इस लिहाज से यह साफ है कि प्रोसेसर कंप्यूटर का वह हिस्सा है जो उपयोगकर्ता (user) की ओर से दिए जाने वाले आदेशों को ठीक से समझकर उनका ठीक से पालन करने, गणितीय क्रियाएं करने, किसी विशेष लक्ष्य या कार्य की जांच आदि करने का काम करता है।

कंप्यूटर के प्रोसेसर वाले हिस्से को सेंट्रल प्रोसेसिंग यूनिट (Central Processing Unit) कहा जाता है, जिसे आमतौर पर संक्षिप्त रूप में हम सीपीयू भी कह लेते हैं। अब सीपीयू के भी तीन अहम हिस्से होते हैं , जिनके जुड़ने से प्रोसेसिंग यूनिट अपना सही आकार लेती है और ठीक से कार्य कर पाती है। ये तीन भाग हैं- मेमोरी (Memory) अर्थमेटिक लॉजिक यूनिट (Arithmatic Logic Unit) यानी एएलयू और कंट्रोल (control) प्रोसेसर के इन तीनों हिस्सों के जिम्मे अलग.अलग तरह के निर्देशों का ठीक से पालन करना और परिणामों को बिल्कुल सही प्राप्त करना है। सबसे पहले बात करते हैं अर्थमेटिक लॉजिक यूनिट की। अर्थमेटिक का हिन्दी अर्थ ही अंक गणित है, यानी इस यूनिट के जिम्मे सभी तरह की गणनाएं और तुलनाएं हैं। अब लॉजिक पर आएं तो इसके तहत गणितीय प्रक्रियाओं से इतर मिलने वाले सभी तरह के निर्देश शामिल हैं। कंट्रोल यूनिट का काम कंप्यूटर के सभी भागों की निगरानी करना और उपयोगकर्ता की ओर से मिलने वाले निर्देशों को अभीष्ट यूनिट तक पहुंचाना होता है। तीसरी और सबसे अहम यूनिट है मेमोरीए चूंकि यह वृहद् विषय है , इसे हम आगे विस्तार से जानेंगे।

2. मेमोरी (Memory)

मेमोरी, यानी याददाश्त। हम पहले ही जान चुके हैं कि मानव विकास के अनुक्रम में जिस तेजी से गणितीय गणनाएं लगातार बढ़ती गईं, उसी तेजी से यह जरूरत भी बढ़ती चली गई कि हम जो भी गणना कर रहे हैं , उनके परिणाम स्मृति में लंबे समय तक संजोकर रखें। अबेकस से लेकर कंप्यूटर तक के विकास की सैकड़ों सालों की यात्रा का परिणाम है मेमोरी। कंप्यूटर पर उपयोगकर्ता जो भी जानकारी, सूचना, आंकड़ा, परिणाम बाद के इस्तेमाल के लिए सुरक्षित रखना चाहता है, वह मेमोरी में ही जाकर संग्रहीत (Stored) होता है।

मानव मस्तिष्क में जिस तरह चेतन और अवचेतन मस्तिष्क की अवधारणा है और अब तो यह विभिन्न शोधों से पता भी चला है कि मस्तिष्क के अलग-अलग हिस्से अलग.अलग तरह की सूचनाओं को संग्रहीत कर स्मृति में बनाए रखते हैं, ठीक उसी तरह कंप्यूटर की मेमोरी भी काम करती है। कंप्यूटर की मेमोरी भी मानव मस्तिष्क के अलग.अलग हिस्सों की तरह कई छोटे टुकड़ों (Blocks) में बंटी होती है। इन ब्लॉक को सामान्यतः बाइट (Byte) कहा जाता है। कंप्यूटर मेमोरी में हर ब्लॉक की अपनी एक खास लोकेशन (Location) होती है, जो मनुष्य की पहचान के लिए दिए जाने वाले नामों की तरह इन पर दर्ज नंबरों से तय मानी जाती है। इन नंबरों को बाइट या ब्लॉक का पता (Address) माना जा सकता है।

हर बाइट अपने से भी छोटी इकाई बिट (Bit) से बनती है। बिट को कंप्यूटर मेमोरी का सबसे छोटा हिस्सा माना जा सकता है और हर आठ बिट की श्रृंखला (Chain) मिलकर एक बाइट का निर्माण करती है। बिट किस तरह काम करती है, इसे हम 'हाँ' या 'ना' के उदाहरण से समझते हैं। हमें कुछ काम करना है तो हमारे उसे करने या नहीं करने की दो ही स्थितियां हो सकती हैं , हां या ना। या इसे किसी स्विच के ऑन या ऑफ होने से भी समझ सकते हैं। यानी किसी बाइट में मौजूद बिटों की शृंखला में कुछ बिट हां या ऑन हैं तो कुछ ना या ऑफ। इस आधार पर ऑन बिट को 0 और ऑफ को 1 माना जाता है। कंप्यूटर पर हम जो भी काम करते हैं या

सूचनाएं संग्रहीत रखते हैं, वह सब 0 और 1 के रूप में ही दर्ज होता है, इन्हें बाइनरी संख्या कहा जाता है, जिसे हम इसी यूनिट के अगले हिस्से में जानेंगे। किसी भी कंप्यूटर की संग्रहण क्षमता यानी उसकी मेमोरी को बाइट में ही मापा जाता है। जिस कंप्यूटर की बाइट जितनी अधिक होगी, वह आंकड़ों के संग्रहण, गणनाओं और सूचनाओं तथा परिणाम के निष्पादन में उतना ही सक्षम होगा। बाइट से लेकर गीगा बाइट और इससे भी कहीं आगे एक्साबाइट तक मेमोरी की क्षमता की यह श्रृंखला जाती है। इस लिहाज से जितनी अधिक बाइट वाला कंप्यूटर होगा, उसकी मेमोरी उतनी ही अधिक होगी। इसे हम निम्न सारिणी से समझ पाएंगे-

8 बिट	1 बाइट
1024 बाइट	1 किलोबाइट
1024 किलोबाइट	1 मेगाबाइट
1024 मेगाबाइट	1 गीगाबाइट

• इनटर्नल मेमोरी (Internal memory)

कंप्यूटर की मेमोरी दो तरह की होती है , भीतरी और बाहरी। भीतरी यानी इनटर्नल मेमोरी को कंप्यूटर की मुख्य मेमोरी (Main Memory) माना जाता है। कंप्यूटर की इनटर्नल यानी मेन मेमोरी को भी दो भागों में बांटा जा सकता है। पहला है रैम (RAM) और दूसरा रॉम (ROM) ये दोनों मेमोरी सेंट्रल प्रोसेसिंग यूनिट में ही मौजूद होती हैं, लेकिन दोनों के काम करने का तरीका अलग होता है जो कंप्यूटर को आंकड़ों को संग्रहीत करके रखने में मददगार बनता है।

रैम (RAM)- पहले बात करते हैं रैम की। रैम का पूरा नाम है रैंडम एक्सेस मेमोरी (Random Access Memory) यानी मेमोरी का वह हिस्सा या वह प्रकार , जिसे उपयोगकर्ता अपनी इच्छा के अनुसार इस्तेमाल कर सकता है। इस मेमोरी में कोई भी जानकारी, आंकड़ा या सूचना कम समय के लिए ही संग्रहीत हो सकती है। कोई नया या दूसरा डाटा आने की स्थिति में पिछला डाटा सुरक्षित नहीं रह पाता है।

रॉम (ROM)- रॉम यानी रीड ओनली मेमोरी (Read Only Memory) जैसा कि नाम से ही स्पष्ट हो रहा है कि इसमें संग्रहीत आंकड़ों को उपयोगकर्ता पढ़ यानी इस्तेमाल तो कर सकता है , लेकिन इसमें बदलाव नहीं किया जा सकता। रॉम कंप्यूटर निर्माता कंपनी की ओर से उपलब्ध ऐसा डाटा है , जिनकी उपयोगकर्ता को निरन्तर आवश्यकता होती है। इसमें संग्रहीत डाटा कभी मिटता या खत्म नहीं होता है।

कैश मेमोरी (Cache Memory)- कैश भी रैंडम एक्सेस मेमोरी के समान है , लेकिन इन दोनों में मुख्य अंतर यह है कि रैम जहां कंप्यूटर सिस्टम में स्टोर रहती है , कैश मेमोरी गतिशील होती है और इसे सर्वर में स्टोर किया जाता है। दोनों का उपयोग और कार्यशैली समान ही होते हैं , लेकिन कंप्यूटर इस मेमोरी का उपयोग अधिकतर हाल में देखे गए वेब पेजों को याद रखने में करता है।

• बाहरी मेमोरी (External Memory)

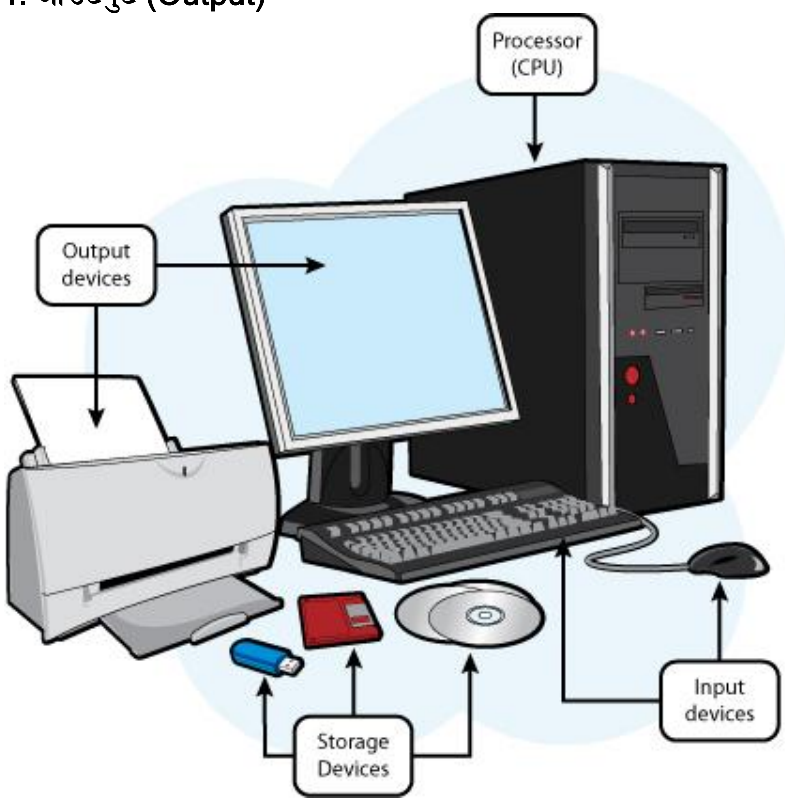
कंप्यूटर की भीतरी या मुख्य मेमोरी की अपनी कुछ सीमाएं होती हैं। हर कंप्यूटर को अलग मेमोरी क्षमता से डिजाइन किया जाता है। लेकिन अकसर यह होता है कि डाटा या आंकड़े इतने अधिक हो जाते हैं कि उन्हें कंप्यूटर में ही संग्रहीत रख पाना संभव नहीं हो पाता। या कई बार जरूरत यह होती है कि कंप्यूटर में दर्ज परिणामों का इस्तेमाल कहीं और करना होता है। ऐसे में बाहरी मेमोरी मददगार साबित होती है। शायद यही

वजह है कि इस मेमोरी को सहायक मेमोरी (Auxilliary Memory) भी कहा जाता है। हम सभी लोग इस तरह की मेमोरी का अकसर दैनिक जीवन में उपयोग करते हैं। फ्लॉपी , पेनड्राइव, सीडी, डीवीडी, हार्ड डिस्क आदि कंप्यूटर की सहायक मेमोरी ही हैं। इनमें सैकड़ों.हजारों गीगाबाइट तक आंकड़े , सूचनाएं, गणनाएं, परिणाम आदि संग्रहीत कर रखे जा सकते हैं।

3. इनपुट (Input)

यह तो हम स्पष्ट रूप से जानते हैं कि कंप्यूटर कोई भी कार्य उपयोगकर्ता की ओर से दिए जाने वाले निर्देशों के पालन के अनुक्रम में करता है। ऐसे में इनपुट कंप्यूटर की वह इकाई है , जिसकी मदद से उपयोगकर्ता सेंट्रल प्रोसेसिंग यूनिट यानी सीपीयू तक अभीष्ट निर्देश पहुंचा पाता है। उपयोगकर्ता की ओर से कंप्यूटर को निर्देश देने की इस प्रक्रिया को ही इनपुट कहा जाता है। कंप्यूटर को इनपुट देने के लिए उपयोगकर्ता कुछ उपकरणों (Devices) का इस्तेमाल करता है , जिन्हें इनपुट डिवाइस भी कहा जाता है। मसलन , हम जब कंप्यूटर पर टाइपिंग करते हैं तो हम उसके लिए की.बोर्ड (Keyboard) पर टाइप करते हैं। इस तरह की.बोर्ड कंप्यूटर के लिए एक इनपुट डिवाइस है , क्योंकि यह उपयोगकर्ता की ओर से टाइप किए जाने वाले अक्षर-अंक की जानकारी कंप्यूटर के सीपीयू को पहुंचाता है। की.बोर्ड के अलावा माउस, जॉयस्टिक, लाइट पेन, माइक, स्कैनर आदि भी इनपुट डिवाइस हैं।

4. आउटपुट (Output)



उपयोगकर्ता जो भी इनपुट कंप्यूटर को देता है वह सीपीयू में जाकर प्रोसेस किया जाता है। जो परिणाम कंप्यूटर उपयोगकर्ता तक पहुंचाता है, उसे आउटपुट कहा जाता है। आउटपुट पाने में कुछ मशीनें या उपकरण कंप्यूटर के

सहायक होते हैं। इन मशीनों या उपकरणों पर उपयोगकर्ता अपनी ओर से दिए गए निर्देशों के परिणाम कंप्यूटर के स्तर पर की जाने वाली डाटा प्रोसेसिंग के बाद हासिल कर पाता है। इनमें सबसे अधिक महत्वपूर्ण और सर्वाधिक इस्तेमाल की जाने वाली डिवाइस है मॉनीटर (Monitor), मॉनीटर पर ही हम हर परिणाम देख-सुन सकते हैं। इसके अलावा प्रिंटर , स्पीकर आदि भी आउटपुट डिवाइस हैं। इनपुट-आउटपुट डिवाइस और कंप्यूटर अन्य प्रमुख घटक यानी सिस्टम यूनिट को हम उपरोक्त चित्र की मदद से आसानी से समझ सकते हैं।

4. प्रोग्राम (Program)

दैनिक जीवन में हम जो भी काम करते हैं, उनके लिए निश्चित और पूर्वनियत प्रक्रियाओं के एक समूह से गुजरते हैं। मसलन हमें नहाना है तो यह निश्चित है कि हम सबसे पहले बाथरूम तक पहुंचेंगे , नल खोलेंगे, बाल्टी लगाकर पानी भरेंगे और फिर नहाना शुरू करेंगे। ठीक इसी तरह कंप्यूटर भी उपयोगकर्ता के लिए जो भी काम करता है, वह दरअसल आदेशों का एक ऐसे समूह के जरिये तय हो पाता है , जो पहले से कंप्यूटर के सीपीयू में दर्ज हैं।

कंप्यूटर पर हर कार्य के लिए अलग आदेश समूह व्यवस्थित रहता है। उदाहरण के लिए हम जब भी कंप्यूटर पर कुछ काम करते हैं तो देखने में तो वह माउस के एक क्लिक पर चुटकी में हो जाता है, लेकिन दरअसल, प्रोसेसर तक माउस की वह एक क्लिक अभीष्ट काम से जुड़े आदेशों का समूह पहुंचाती है। ये आदेश चरणबद्ध तरीके से कंप्यूटर की भीतरी मेमोरी में दर्ज रहते हैं और प्रोसेसिंग यूनिट उस पर बेहद तेजी से काम (Execution) करती है, जिससे सेकंड से भी कम समय के भीतर जरूरी परिणाम हमारे सामने आउटपुट डिवाइस यानी मॉनीटर या प्रिंटर पर उपलब्ध हो जाता है। किसी अभीष्ट कार्य को सफलतापूर्वक निष्पादित करने के लिए जरूरी आदेशों के समूह को कंप्यूटर के लिए प्रोग्राम कहा जाता है।

13.5 कंप्यूटर की कार्यपद्धति (Working Process)

कंप्यूटर के सभी भागों के बारे में जानकारी मिल जाने के बाद यह जानना जरूरी लगता है कि कंप्यूटर इन सबकी मदद से काम करता कैसे है। इससे पहले हम यह जान लेते हैं कि कंप्यूटर की कार्यपद्धति में किन तत्वों की सबसे अधिक आवश्यकता होती है। ये तत्व हैं- डाटा (Data), सूचना (Information), हार्डवेयर (Hardware) और सॉफ्टवेयर (Software)

हम जानते हैं कि कंप्यूटर पर उपयोगकर्ता की ओर से कुछ निर्देश दिए जाते हैं , ये निर्देश सूचनात्मक होते हैं , यानी हम कंप्यूटर के सेंट्रल प्रोसेसिंग यूनिट अर्थात् सीपीयू को कुछ डाटा उपलब्ध कराते हैं , जिसके आधार पर वह हार्डवेयर और सॉफ्टवेयर की मदद से परिणाम हासिल करता है। सामान्यतः दैनिक जीवन में भी हम कई तरह के डाटा का इस्तेमाल कर किसी परिणाम पर पहुंचते हैं , इस प्रक्रिया को डाटा प्रोसेसिंग (Data Processing) कहते हैं। कंप्यूटर पर यही कार्य इलेक्ट्रॉनिक डाटा प्रोसेसिंग (Electronic Data Processing) बन जाता है , क्योंकि कंप्यूटर एक इलेक्ट्रॉनिक मशीन है। आइए अब हम डाटा प्रोसेसिंग के प्रमुख तत्वों को समझते हैं-

• डाटा क्या है (What is Data)

सामान्य शब्दों में कहा जाए तो डाटा दरअसल जानकारी है। इसे इस उदाहरण से समझते हैं , मान लीजिए कि हम क्रिकेट खेल रहे हैं। अब क्रिकेट में किन खेल उपकरणों का इस्तेमाल होता है , क्रिकेट को खेलने का सही

तरीका क्या है, क्रिकेट के मैच कितने तरह के होते हैं , क्रिकेट के एक मैच में कितनी टीमें खेलती हैं , क्रिकेट की एक टीम में कितने खिलाड़ी होते हैं , इस तरह सवालों की एक लंबी शृंखला जो बनेगी , वह क्रिकेट को लेकर अलग-अलग तरह का डाटा बन जाएगा।
अब इसे कंप्यूटर की भाषा में समझें तो डाटा दो तरह का होता है। पहला संख्यात्मक (Numeric) और दूसरा चिह्नात्मक (Alpha Numeric) संख्यात्मक जैसा कि नाम से ही स्पष्ट हो रहा है कि यह डाटा अंकों से संबंधित है, जिनका उपयोग जोड़, घटाना, गुणा-भाग या अन्य तरह की गणनाओं में किया जा सकता है। दूसरी ओर चिह्नात्मक डाटा का मतलब ऐसी जानकारियों से है , जिन्हें अंकीय स्वरूप में दर्ज नहीं किया जा सकता। जैसे- किसी व्यक्ति का नाम, किसी किताब का नाम आदि। इस तरह के डाटा के साथ गणितीय प्रक्रिया संपन्न नहीं की जा सकती है , लेकिन इनके जरिये तुलनात्मक परिणाम (Comparative Results) जरूर हासिल किए जा सकते हैं।

• सूचना (Information)

किसी भी काम के संबंध में हमारे पास जो भी डाटा यानी जानकारी उपलब्ध होती है , वह अव्यवस्थित (Unarranged) होती है। इसकी वजह से कई बार यह डाटा इसलिए उपयोगी साबित नहीं हो पाता , क्योंकि इसके व्यवस्थानुक्रम में नहीं होने के कारण अभीष्ट परिणाम प्राप्त करना असंभव होता है। इसके लिए जरूरी है कि हम डाटा के अकूत भंडार में से सिर्फ उसी डाटा को अपने लिए चुनें , जो समय विशेष पर हमारे लिए उपयोगी है। उदाहरण के लिए, यदि हमें प्यास लगी है तो हम जानते हैं कि प्यास पानी से बुझेगी लेकिन पानी के संबंध में हमारे मस्तिष्क में कई तरह का डाटा उपलब्ध है। पानी नदी से आता है , पानी बारिश से भी आता है, तालाब में भी पानी भरा रहता है और नल में भी पानी आता है, लेकिन इन सब जानकारियों में से महज एक जानकारी हमें इच्छित परिणाम तक पहुंचा सकती है और वह है नल से पानी आना। इसी तरह कंप्यूटर में कोई इच्छित परिणाम प्राप्त करने के लिए उपयोगकर्ता डाटा के भंडार में से जरूरी और उपयोगी डाटा का चयन करता है। इस चयनित डाटा को ही सूचना कहा जाता है।

• हार्डवेयर (Hardware)

कंप्यूटर पर हम जो भी काम करते हैं , वह दो चीजों की मदद के बिना असंभव है। ये चीजें हैं हार्डवेयर और सॉफ्टवेयर। पहले बात करते हैं हार्डवेयर की। हार्डवेयर कंप्यूटर से जुड़े वे कल-पुर्जे (Spare Parts) या उपकरण हैं, जिन्हें उपयोगकर्ता आंखों से देख सकता है या छूकर महसूस कर सकता है। सीपीयू , मॉनीटर, माउस, की.बोर्ड, प्रिंटर, पैन ड्राइव आदि कंप्यूटर के हार्डवेयर हैं।

• सॉफ्टवेयर (Software)

कंप्यूटर के सफल तरीके से कार्य करने (Execution) में सॉफ्टवेयर की अहम भूमिका है। सॉफ्टवेयर दरअसल प्रोग्रामों का समूह है। यानी उपयोगकर्ता जो भी काम कंप्यूटर पर करना चाहता है या निर्देश कंप्यूटर को देना चाहता है, उसके सफल निष्पादन के लिए जिन आदेशों की आवश्यकता कंप्यूटर को पड़ती है , उस प्रोग्राम को सॉफ्टवेयर कहा जाता है। बिना सॉफ्टवेयर के कंप्यूटर पर कोई भी काम कर पाना असंभव सा है , क्योंकि यदि सॉफ्टवेयर नहीं होगा तो इसका सीधा तात्पर्य यह है कि संबंधित कंप्यूटर के पास उपयोगकर्ता के इच्छित आदेशों का पालन करवाने वाले आदेशों का समूह उपलब्ध नहीं है। ऐसी स्थिति में कंप्यूटर के लिए अभीष्ट

परिणाम देना संभव नहीं हो सकेगा। कार्यक्षमता के लिहाज से सॉफ्टवेयर को भी दो भागों में बांटा जा सकता है। पहला है सिस्टम सॉफ्टवेयर और दूसरा है एप्लीकेशन सॉफ्टवेयर।

सिस्टम सॉफ्टवेयर वे प्रोग्राम हैं , जिनका काम सिस्टम यानी कंप्यूटर को चलाते रहना है। ऑपरेटिंग सिस्टम (Operating System), कंपाइलर (Compiler), यूटिलिटी प्रोग्राम (Utility Program) ऐसे ही सॉफ्टवेयर हैं। यह भी कहा जा सकता है कि यही वे प्रोग्राम हैं, जिनकी वजह से कंप्यूटर चलता है, यानी ये कंप्यूटर के प्राण हैं। कंप्यूटर पर जो भी उपयोगकर्ता काम करता है , उसे इन्हीं सॉफ्टवेयर पर काम करना होगा। दूसरी ओर , एप्लीकेशन सॉफ्टवेयर वे प्रोग्राम हैं , जिन्हें उपयोगकर्ता की जरूरत के हिसाब से डिजाइन किया गया है। उदाहरण के लिए जो उपयोगकर्ता पेंटिंग करना चाहता है , उसके लिए पेंटिंग के प्रोग्राम हैं , किसी को वेतन का रिकॉर्ड दर्ज करना है तो उसके लिए अलग एप्लीकेशन हैं। ये प्रोग्राम कंप्यूटर में पहले से उपलब्ध नहीं होते हैं , इन्हें अलग से इंस्टॉल (Install) करना होता है। सिस्टम और एप्लीकेशन सॉफ्टवेयरों की मदद से ही कंप्यूटर पूरा होता है। इन दोनों के संयुक्त स्वरूप को सॉफ्टवेयर पैकेज भी कहा जाता है।

13.6 कंप्यूटर लैंग्वेज (Computer Languages)

हम यह भली-भांति जानते हैं कि कंप्यूटर मानव जीवन के लिए बहुधा उपयोगी मशीन है जो गणनाओं के जरिये मानव जीवन को सरल-सुगम बना रही है। लेकिन यह भी उतना ही सत्य है कि कंप्यूटर स्वयं कोई परिणाम मनुष्य को नहीं देता , बल्कि यह उपयोगकर्ता के निर्देशों के पालन के अनुक्रम में ही काम करता है। कंप्यूटर पर किस निर्देश के आधार पर डाटा प्रोसेसिंग का क्या परिणाम निकलेगा , यह तय करते हैं प्रोग्राम और ये प्रोग्राम आदेशों का एक समूह होते हैं, यह हम पहले ही जान चुके हैं।

इस लिहाज से कंप्यूटर के लिए हर काम के लिए आदेशों का एक ऐसा समूह यानी प्रोग्राम तैयार किया जाता है , जिसे कंप्यूटर समझ सके। कंप्यूटर के लिए प्रोग्राम जिन भाषाओं में लिखे जाते हैंए उन्हें कंप्यूटर प्रोग्रामिंग लैंग्वेज कहा जाता है। कंप्यूटर बस इतना करता है कि जो भी प्रोग्राम उसके सीपीयू में इंस्टॉल हो जाए , उसके आदेशों के क्रम को वह मेमोरी में सेव कर लेता है। इसके बाद जब भी कभी उपयोगकर्ता को आवश्यकता होती है, कंप्यूटर का प्रोसेसर मेमोरी से अभीष्ट आदेशों के प्रोग्राम का चयन कर लेता है और इसके आधार पर परिणाम उपयोगकर्ता को उपलब्ध करा देता है। कंप्यूटर के लिए प्रोग्राम बनाने वाली भाषाओं में मुख्यतः अंग्रेजी के कुछ शब्द और चिह्न प्रयुक्त किए जाते हैं।

हर प्रोग्रामिंग भाषा का अपना एक अलग व्याकरण (Grammar or Syntax) होता है। ऐसे में यह जरूरी होता है कि जिस भाषा में प्रोग्राम तैयार किया जा रहा हो , उसके व्याकरण का पूरा पालन किया जाए , ऐसा नहीं करने पर कंप्यूटर प्रोग्राम को ठीक से समझ नहीं सकेगा और आदेशों का ठीक पालन नहीं कर पाने से वह परिणाम नहीं दे सकेगा।

कंप्यूटरों के लिए प्रयुक्त होने वाली प्रमुख भाषाएं निम्नवत हैं-

- बेसिक (BASIC)
- सी (C)
- सी++ (C++)
- जावा (JAVA)

• डॉटनेट (DOTNET)

- **बाइनरी संख्या प्रणाली (Binary Number System)**

यहां यह उल्लेखनीय है कि कंप्यूटर बाइनरी भाषा ही समझते हैं, यानी कंप्यूटर का सारा काम सिर्फ दो अंकों 0 और 1 पर चलता है। यह हम पहले ही जान चुके हैं कि कंप्यूटर मेमोरी की सबसे छोटी इकाई बिट है जो 0 और 1 से ही मिलकर बनती है। हम सामान्य जीवन में दशमलव संख्या प्रणाली का इस्तेमाल करते हैं , यानी एक से नौ तक के अंक, लेकिन कंप्यूटर सिर्फ 0 और 1 का ही प्रयोग करता है।

मान लीजिए कि हमें 9 लिखना है तो हम सीधे 9 लिखेंगे , लेकिन यदि कंप्यूटर को 9 लिखना है तो प्रोससिंग यूनिट इसे बाइनरी नंबरों में तोड़कर समझेगा। इसे सामान्य शब्दों में इस तरह समझ सकते हैं कि हम कंप्यूटर पर जो भी काम करते हैं, वह हमारे लिए भले ही सीधा समझ में आता हो, लेकिन कंप्यूटर उसे अपनी भाषा में समझता है। हालांकि, आउटपुट पर कंप्यूटर जो परिणाम उपलब्ध कराता है , वह उसी रूप में होता हैए जो हमारा अभीष्ट है।

दशमलव संख्या	बाइनरी संख्या	दशमलव संख्या	बाइनरी संख्या
0	0	8	1000
1	1	9	1001
2	10	10	1010
3	11	11	1011
4	100	12	1100
5	101	13	1101
6	110	14	1110
7	111	15	1111

यहां यह उल्लेखनीय है कि दशमलव संख्या प्रणाली में हम 10 को आधार मानते हैं , क्योंकि इस संख्या प्रणाली में हम 0 से लेकर 9 तक कुल 10 अंकों की मदद से किसी भी बड़ी से बड़ी संख्या को लिख सकते हैं , लेकिन बाइनरी संख्या प्रणाली में 2 ही हर संख्या का आधार है , क्योंकि इस प्रणाली में सिर्फ दो अंकों 0 और 1 का ही इस्तेमाल किया जाता है। विशेष पहलू यह है कि किसी भी दशमलव संख्या को बाइनरी संख्या में बदला जा सकता है और किसी भी बाइनरी संख्या को दशमलव संख्या में परिवर्तित किया जा सकता है। लेकिन चूंकि यह बिन्दु हमारे विषय के लिए बहुत उपयोगी नहीं है , लिहाजा हमने उपर दिए गए ग्राफ में सिर्फ समझने के लिए हम प्रथम 16 अंकों के बाइनरी नंबर लिए हैं।

यहां यह तथ्य उल्लेखनीय है कि हमें पहले ही मालूम है कि एक बाइट का अर्थ आठ बिट से होता है और हर बिट के दो मान हो सकते हैं, 0 या 1। अब गणितीय सिद्धांतों के अनुसार हर एक बाइट में बिटों की संख्या 2 का आठ गुना यानी दो को आठ बार गुणा करने से मिलने वाला मान अर्थात् 256 हो सकती है। यानी सीधे शब्दों में कहें तो हर एक बाइट में 0 से लेकर 255 तक यानी कुल 256 संख्याएं दिखाई जा सकती हैं। इसी तरह किसी एक बाइट में कुल 256 चिह्न दर्ज किए जा सकते हैं।

- **कोडिंग सिस्टम (Coding System)**

कंप्यूटर पर बाइनरी भाषा के चलते अक्षरों और चिह्नों को बाइनरी संख्या प्रणाली के हिसाब से ही लिखना जरूरी है। ऐसे में यह स्पष्ट है कि हम जो भी संख्या या चिह्न या डाटा लिखना चाहते हैं उसका कोई बाइनरी

कोड होना आवश्यक है। तभी कंप्यूटर समझ सकेगा कि हम क्या लिखना चाहते हैं। इस लिहाज से कोडिंग सिस्टम कंप्यूटर और उपयोगकर्ता के बीच परस्पर बातों को समझाने का जरिया बन जाता है। कोडिंग के जरिये अक्षरों और चिह्नों को बाइटों में सुरक्षित कर लिया जाता है। इसके लिए मुख्यतः दो प्रकार के कोड प्रयोग में लाए जाते हैं , ये हैं आस्की कोड (American Standard Code for Information) और एब्सडिक (Extended Binary Coded Decimal Interchange Code) हालांकि, माइक्रो यानी पर्सनल कंप्यूटरों में मुख्यतः आस्की कोड का ही इस्तेमाल किया जाता है। अंग्रेजी अक्षरों को आस्की और एब्सडिक कोड में कैसे लिखा जाता है, यह निम्न सारिणी से समझा जा सकता है-

अक्षर	**आस्की**	**एब्सडिक**
A	01000001	11000001
B	01000010	11000010
C	01000011	11000011
D	01000101	11000101

- **कम्पाइलर (Compiler)**

हमें मालूम है कि कंप्यूटर के प्रोग्राम ऐसी भाषा में होने जरूरी हैं , जिसे कंप्यूटर समझ सके और यह भाषा है बाइनरी संख्या प्रणाली आधारित। कंप्यूटर की इस भाषा को मशीन लैंग्वेज (Machine Language) कहा जाता है। इसे सामान्य तौर पर निम्न स्तरीय भाषा (Low Level Language) भी कहा जाता है। इसलिए यह जरूरी हो जाता है कि कंप्यूटर के लिए जो प्रोग्राम तैयार किए जाएं , वे मशीनी भाषा में ही हों , लेकिन ऐसा करना संभव नहीं हो पाता , क्योंकि हर अंक, चिह्न को बाइनरी संख्या प्रणाली में 0 और 1 के रूप में लिख पाना बेहद लंबा और दुष्कर कार्य है। ऐसे में मददगार साबित होता है कंपाइलर। दरअसलए कंपाइलर एक ऐसा प्रोग्राम है जो उच्चस्तरीय भाषा में लिखे गए किसी भी प्रोग्राम को मशीनी भाषा में बदल देता है , ताकि कंप्यूटर उसे आसानी से समझ सके। इसे निम्न ग्राफ की मदद से समझ सकते हैं-

प्रोग्राम → कम्पाइलर → बाइनरी

कंपाइलर किसी कंप्यूटर के सिस्टम सॉफ्टवेयर का ही हिस्सा होता है। कंपाइलर दो भाग में काम करता है। पहला यह कि कंपाइलर उपयोगकर्ता की ओर से दिए जाने वाले आदेश की अभीष्ट प्रोग्राम के व्याकरण के आधार पर पूरी जांच करता है। पता लगाता है कि आदेश प्रोग्राम के व्याकरण के अनुरूप है कि नहीं। अगर कोई गलती है तो कंपाइलर रूक जाता है , जिसके बाद उपयोगकर्ता को दोबारा ठीक से आदेश देना होता है। कंपाइलर आदेश को प्रोग्राम के व्याकरण सम्मत पाता है तो इसे तत्काल मशीनी भाषा यानी बाइनरी कोड में बदल देता है। उपयोगकर्ता के एक आदेश को पूरा पढ़ने के बाद कंपाइलर उस एक आदेश को बाइनरी कोड के हिसाब से कई छोटे आदेशों में भी बदल सकता है। ये आदेश सीपीयू में जाते हैं , जहां मेमोरी, प्रोसेसर, एएलयू आदेशों के अनुरूप काम करती हैं।

13.7 उपसंहार (The Conclusion)

हम न सिर्फ कंप्यूटर के विकास और इसके इतिहास से रूबरू हुए हैं , बल्कि यह भी समझ पाने में सक्षम रहे हैं कि किस तरह मानवीय सभ्यताओं के विकास के साथ कंप्यूटर भी आगे बढ़ा। प्राचीन काल में सामान्य गणनाओं से लेकर आज के वैज्ञानिक युग में मगल ग्रह तक मनुष्य के सफर को कामयाब बनाने में कंप्यूटर का किसी न किसी रूप में योगदान रहा। इस लिहाज से कह सकते हैं कि कंप्यूटर मानव समाज का अभिन्न अंग बन चुका है।

13.8 अभ्यास प्रश्न (Exercise)

1. गणनाओं के लिए सर्वप्रथम उपयोग किया गया ज्ञात उपकरण है-

a) पास्कलाइन

b) एनियाक

c) अबेकस

d) सुपर कंप्यूटर

2. कंप्यूटर के विकास की मूल अवधारणा इनमें से क्या थी-

a) गणनाएं

b) मनोरंजन

c) खेल

d) उपरोक्त में से कोई नहीं

3. इनमें से किसे कंप्यूटर का जनक माना जाता है-

a) ब्लेज पास्कल

b) सर चार्ल्स बैबेज

c) जेपी एकर्ट

d) जेडब्ल्यू मॉशी

4. इंटीग्रेटेड सर्किट यानी आईसी की खोज किसने की-

a) सर चार्ल्स बैबेज

b) जेपी एकर्ट

c) टीएस बिल्की

d) बिल गेट्स

5. कम्पाइलर इनमें से क्या है-

a) एक प्रोग्राम

b) एक प्रोग्रामिंग भाषा

c) इनपुट डिवाइस

d) आउटपुट डिवाइस

6. एक बाइट का मान होता है-

a) 1 बिट

b) 8 बिट

c) 1024 बिट

d) उपरोक्त में से कोई नहीं

7. बाइनरी संख्या प्रणाली में आधार अंक हैं-

a) 0 से 9 तक

b) 0 और 1

c) 2 और 10

d) कोई आधार अंक नहीं है

8. हर एक बाइट में चिह्नों या अंकों की संख्या हो सकती है-

a) 200

b) 512

c) 1024

d) 256

9. पैन ड्राइव इनमें से किस मेमोरी का उदाहरण है-

a) बाहरी मेमोरी

b) रैम

c) रॉम

d) कैश मेमोरी

10. अर्थमेटिक लॉजिक यूनिट हिस्सा है-

a) सीपीयू का

b) एक विशेष प्रोग्राम का

c) कम्पाइलर का

d) कंप्यूटर उपकरणों का

11. पर्सनल या माइक्रो कंप्यूटर अस्तित्व में आए-

a) 1970 में

b) 1942 में

c) 1981 में

d) 1990 में

12. अमेरिकी वैज्ञानिक टेड हॉफ ने खोज की थी-

a) माइक्रो प्रोसेसर की

b) बाइनरी संख्या प्रणाली की

c) एनालिटिकल कंप्यूटर की

d) पास्कलाइन की

13. दूसरी पीढ़ी के कंप्यूटरों में इस्तेमाल किया जाता था-

a) वैक्यूम ट्यूब

b) इंटीग्रेटेड सर्किट

c) माइक्रोप्रोसेसर

d) ट्रांजिस्टर

14. पहले माइक्रो प्रोसेसर का नाम था-

a) इनटेल.4004

b) एनियाक

c) परम

d) इनमें से कोई नहीं

15. अल्फा न्यूमेरिक डाटा का तात्पर्य है-

a) अंकों में प्रदर्शित किए जाने वाले डाटा से

b) ऐसे डाटा सेए जिसे अंकों में नहीं दिखाया जा सकता

c) उपरोक्त में से दोनों

d) उपरोक्त में से कोई नहीं

16. निम्नलिखित में से कौन आउटपुट डिवाइस नहीं है-

a) प्रिंटर

b) मॉनीटर

c) स्कैनर

d) सभी आउटपुट डिवाइस हैं

17. ब और ब़ क्या हैं

a) कंप्यूटर एप्लीकेशन

b) कंप्यूटर प्रोग्राम

c) कंप्यूटर प्रोग्रामिग भाषाएं

d) कंप्यूटर कम्पाइलर

18. कंप्यूटर प्रोग्राम का तात्पर्य है-

a) खास परिणाम के लिए तय आदेशों का क्रम

b) खास परिणाम पाने के लिए जरूरी आउटपुट

c) कंप्यूटर पर इस्तेमाल किए जाने वाले इनपुट उपकरण

d) उपरोक्त में से कोई नहीं

19. इस मेमोरी में दर्ज सूचनाएं बदली नहीं जा सकतीं-

a) रैम

b) रॉम

c) कैश मेमोरी

d) उपरोक्त सभी

20. भारत में विकसित सुपर कंप्यूटर का नाम है-

a) आईबीएम

b) एनियाक

c) लेन्ज कैल्कुलेटर

d) परम

13.9 निबंधात्मक प्रश्न (Theoretical Question)

1. कंप्यूटर क्या हैए अबेकस से लेकर कंप्यूटर तक की विकास यात्रा का विस्तृत वर्णन के साथ इसकी आवश्यकता भी समझाएं।
2. कंप्यूटर को कितनी पीढ़ियों में बांटा जा सकता हैए इनका क्रमवार वर्णन करने के साथ हर पीढ़ी में आए अंतर का विश्लेषण करें।
3. कंप्यूटर के प्रमुख अवयव क्या हैं , हर अवयव कंप्यूटर प्रणाली के लिए किस तरह महत्वपूर्ण है और ये किस तरह काम करते हैं?
4. कंप्यूटर मेमोरी क्या है, यह कितने प्रकार की होती है?
5. कंप्यूटर किस तरह काम करता है, डाटा-सूचना क्या हैं, इनमें प्रमुख अंतर क्या है, सॉफ्टवेयर और हार्डवेयर में क्या अंतर है?
6. बाइनरी संख्या प्रणाली क्या है, यह मानव जीवन में प्रयुक्त की जाने वाली दशमलव संख्या प्रणाली से किस तरह भिन्न हैए कंप्यूटर में इस संख्या प्रणाली का उपयोग क्यों किया जाता है, कोडिंग सिस्टम का भी वर्णन करें।
7. कंप्यूटर के विकास अनुक्रम को संक्षिप्त में समझाते हुए मानव समाज की प्रगति में इसके योगदान का विश्लेषण करें।

इकाई 14 कंप्यूटर ऑपरेटिंग सिस्टम और इंटरनेट के अनुप्रयोग
(Applications of Computer Operating System and Internet)

14.1 प्रस्तावना (Introduction)

हम इस तथ्य से भली -भांति परिचित हैं कि कंप्यूटर आज मानव जीवन की अभिन्न आवश्यकता बन चुका है। जीवन का शायद ही कोई पहलू आज ऐसा बचा रह गया हो , जिसमें छोटे या बड़े रूप में कंप्यूटर का इस्तेमाल नहीं किया जाता हो। अब जिस तरह मानवीय सामाजिक व्यवस्था अलग -अलग घटकों में बंटी हुई है, उसी तरह कंप्यूटर की पूरी कार्य व्यवस्था भी कई अंगों का एक सामूहिक स्वरूप है। पिछली इकाई में हमने कंप्यूटर के इतिहास से लेकर इसके विकास के अनुक्रम को विस्तार से समझा है। अब इस इकाई में हम जानेंगे कि एप्लीकेशन (Application Softwares) और ऑपरेटिंग सिस्टम (Operating Systems) किस तरह कंप्यूटर के सफल कार्य निष्पादन में सहयोगी हैं।

14.2 उद्देश्य (Objectives)

इस इकाई के अध्ययन के बाद आप-

- ✓ कंप्यूटर ऑपरेटिंग सिस्टम और इसका कंप्यूटर की कार्यप्रणाली में उपयोग को जानेंगे।
- ✓ कंप्यूटर ऑपरेटिंग सिस्टम के विकास को जानेंगे।
- ✓ नेटवर्किंग और इसका कंप्यूटर की कार्य व्यवस्था में महत्व को जानेंगे।
- ✓ इंटरनेट, इसका इस्तेमाल एवं इसने मानव जीवन को सरल-सुगम बनाने में क्या योगदान किया है
- ✓ इंटरनेट का इस्तेमाल करने में सुरक्षा का ध्यान रखना क्यों जरूरी है , इंटरनेट का इस्तेमाल करने के दौरान किस तरह की सावधानियां बरती जानी चाहिए

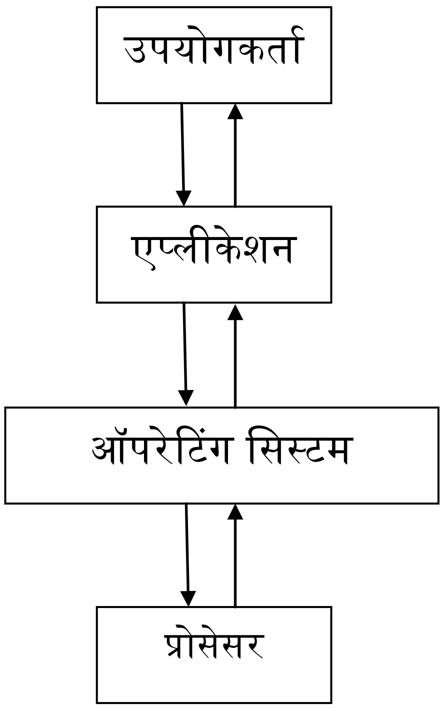
14.3 सिस्टम सॉफ्टवेयर (System Softwares)

पिछली इकाई में हम जान चुके हैं कि सॉफ्टवेयर क्या होते हैं। सॉफ्टवेयर मुख्यतः दो प्रकार के होते हैं , सिस्टम सॉफ्टवेयर (System Softwares) और एप्लीकेशन सॉफ्टवेयर (Application Software) सिस्टम सॉफ्टवेयर वे प्रोग्राम हैं, जिनका काम कंप्यूटर को चलाना होता है। हम इस इकाई में जिस ऑपरेटिंग सिस्टम के बारे में जानने वाले हैं, वह भी मूलतः सिस्टम सॉफ्टवेयर ही है। इसके अलावा एप्लीकेशन सॉफ्टवेयर वे प्रोग्राम हैं, जो किसी खास काम को करने और अभीष्ट परिणाम हासिल करने में उपयोगकर्ता की मदद करते हैं।

14.4 ऑपरेटिंग सिस्टम (Operating System)

हम जान चुके हैं कि ऑपरेटिंग सिस्टम दरअसल सिस्टम सॉफ्टवेयर है , यानी इसकी मदद से ही कोई कंप्यूटर काम कर सकता है। इस लिहाज से कोई भी ऑपरेटिंग सिस्टम वह माध्यम है , जो उपयोगकर्ता और कंप्यूटर के बीच की महत्वपूर्ण कड़ी का काम करता है।

यहां यह बिन्दु अति महत्वपूर्ण है कि ऑपरेटिंग सिस्टम के बिना किसी उपयोगकर्ता के लिए कंप्यूटर से अभीष्ट कार्य करा पाना असंभव तो नहीं है , लेकिन बेहद कठिन जरूर है। ऑपरेटिंग सिस्टम का महत्व इससे समझा जा सकता है कि यह उपयोगकर्ता (user) के आदेशों, एप्लीकेशन सॉफ्टवेयर पर दर्ज निर्देशों को कंप्यूटर तक पहुंचाने और प्रोसेसिंग के बाद मिलने वाले परिणाम और अन्य सूचनाओं को वापस उपयोगकर्ता तक पहुंचाने का काम करता है। इसके अलावा किसी खास कार्य के निष्पादन या मनचाहे परिणाम प्राप्त करने के लिए उपयोगकर्ता ने जो प्रोग्राम तैयार किए हैं , उन्हें शुरू कराने से लेकर पूरी प्रक्रिया के बाद खत्म कराने तक की जिम्मेदारी भी ऑपरेटिंग सिस्टम पर होती है। हार्डवेयर के सभी संसाधनों को जरूरत पड़ने पर प्रोग्राम के लिए उपलब्ध कराना और उपयोगकर्ता के लिए उपयोगी डाटा को सुरक्षित रखने का काम भी ऑपरेटिंग सिस्टम की मदद से ही संभव हो पाता है। ऑपरेटिंग सिस्टम , एप्लीकेशन सॉफ्टवेयर, उपयोगकर्ता और कंप्यूटर के बीच संबंध को निम्न ग्राफ से समझा जा सकता है-



इतिहास और विकास (History and Development)

हम जानते हैं कि कंप्यूटर के विकास की शुरूआत एकल उद्देश्य की पूर्ति के लिए हुई थी , जिसे गणना कहा जाता है। इस तरह के कंप्यूटर मुख्यतः कैल्कुलेटर ही थेए लेकिन जिस तरह गणनाएं और जरूरतें बढ़ती गईं , एक से अधिक कार्य कंप्यूटर की मदद से किए जाने लगे। वर्ष 1950 में अस्तित्व में आए प्रारंभिक कंप्यूटर में कोई ऑपरेटिंग सिस्टम तो नहीं थाए लेकिन इनमें रेजीडेंट मॉनीटर (Resident Monitor) नाम का खास फंक्शन मौजूद था, इसकी वजह से कंप्यूटर की कार्यक्षमता , एक्यूरेसी (Accuracy) और गति (Speed) में भी खासी बढ़ोतरी हुई। इसके बाद ऑपरेटिंग सिस्टम के विकास पर कंप्यूटर अनुसंधानकर्ताओं का ध्यान गया। वर्ष 1960 तक कंप्यूटरों में बैच प्रोसेसिंग , इनपुट-आउटपुट इंटरप्ट, बफरिंग जैसे कार्य करना संभव हो गया था। हालांकि , अब भी यह सिंगल टास्किंग मशीन (Single Tasking Machine) ही थी, यानी कंप्यूटर पर एक समय में एक ही काम कर पाना संभव हो सकता था।

मेनफ्रेम ऑपरेटिंग सिस्टम (Mainframe OS)

हम जानते हैं कि वर्ष 1980 में पर्सनल कंप्यूटर के विकास से पहले सुपर और मेनफ्रेम कंप्यूटर ही अस्तित्व में थे। चूंकि सुपर कंप्यूटर बेहद महंगे थे , लिहाजा मेनफ्रेम कंप्यूटर ही अधिकतर प्रयोग किए जाते थे और ऑपरेटिंग सिस्टम भी मेनफ्रेम कंप्यूटरों के लिए ही विकसित हुए। मेनफ्रेम कंप्यूटरों के लिए कब-कैसे ऑपरेटिंग सिस्टम का विकास हुआ, यह निम्नवत समझा जा सकता है-

- **वर्ष 1950** - रेजीडेंट मॉनीटर फंक्शन
- **वर्ष 1959** - आईबीएम ने अपने मेनफ्रेम कंप्यूटर आईबीएम-704 के लिए शेयर (SHARE) ऑपरेटिंग सिस्टम तैयार किया। आईबीएम के आईबीएम -709 और आईबीएम -7090 मेनफ्रेम कंप्यूटरों में भी यही ऑपरेटिंग सिस्टम प्रयोग किया गया , हालांकि जल्द ही कंपनी ने एक और नया ऑपरेटिंग सिस्टम विकसित कर लियाए जिसे आईबीएम -709, 7090 और 7094 मेनफ्रेम कंप्यूटरों पर इस्तेमाल किया गया। इस ऑपरेटिंग सिस्टम का नाम था आईबीसिस या आईबीजॉब (IBSYS / IBJOB)

वर्ष 1960 - आईबीएम कंपनी ने हर तरह के काम के लिए एक सिंगल ऑपरेटिंग सिस्टम (Single Operating System) तैयार किया , जिसे नाम दिया गया ओएस -360 (OS-360) आईबीएम का यह ऑपरेटिंग सिस्टम

आज के दौर के सभी ऑपरेटिंग सिस्टम का मूलाधार है। खास बात यह है कि उस वक्त इस ऑपरेटिंग सिस्टम के लिए प्रोग्राम इस तरह लिखे गए थे कि यह सिस्टम आज के दौर के कंप्यूटरों पर भी आसानी से चलाया जा सकता है।

ओएस-360 की खासियत यह थी कि यह पहला ऐसा सिस्टम था जो उपयोगकर्ता की जरूरत के मुताबिक संसाधनों को उपलब्ध कराने के अलावा डाटा को मेन और सहायक मेमोरी में सेव करने में मदद करता था। यह पहला सिस्टम था, जिसके जरिये फाइल लॉकिंग (File Locking) का काम संभव हो सका।

कालान्तर में आईबीएम के दूसरे जितने भी ऑपरेटिंग सिस्टम विकसित हुए , वे सभी दरअस ओएस -360 में ही कुछ सुधार कर तैयार किए जाते रहे। दूसरी ओर 1960 में ही कंट्रोल डाटा ऑपरेशन्स (Control Data Operation) और मिनोसेटा यूनिवर्सिटी के संयुक्त प्रयासों से बैच प्रोसेसिंग के मकसद से एक ऑपरेटिंग सिस्टम विकसित किया गया। इसका नाम था स्कोप (SCOPE)

- **वर्ष 1961**- बरॉज कॉरपोरेशन ने बी-5000 नाम से नया मेनफ्रेम कंप्यूटर पेश किया जो मास्टर कंट्रोल प्रोग्राम (MCP) नाम के ऑपरेटिंग सिस्टम से सुसज्जित था। यह दुनिया का पहला ऐसा ऑपरेटिंग सिस्टम था, जिसके लिए पहली बार हाई लेवल लैंग्वेज (High Level Language) ESPOL में प्रोग्राम लिखे गए थे। यही नहीं, इस मशीन में पहली बार वर्चुअल मेमोरी का भी इस्तेमाल किया गया था। यह अपने दौर का बेहद क्रान्तिकारी कदम था। शायद यही वजह थी कि उस दौर की सबसे बड़ी कंप्यूटर निर्माता कंपनी ने अपने हार्डवेयर प्रोजेक्ट एएस400 (AS400) के लिए बरॉज कॉरपोरेशन से इस ऑपरेटिंग सिस्टम के इस्तेमाल की इजाजत मांगी, लेकिन कंपनी ने इनकार कर दिया। एमसीपी का इस्तेमाल आज भी यूनिसिस क्लियरपाथ कंप्यूटरों में किया जा रहा है। दूसरी ओरए बाद में आईबीएम ने सीपी-67 नाम से अपने सिस्टम पर काम कियाए जो वर्चुअल मेमोरी (Virtual Memory) पर फोकस था।
- **वर्ष 1970**- 1961 से 1970 तक ऑपरेटिंग सिस्टम के विकास को लेकर लगातार शोध होते रहे। हर पुराने सिस्टम में कुछ संशोधन कर जल्द ही नया सिस्टम तैयार कर लिया जाता था। इस कड़ी में यह साल भी शामिल रहा। इस वर्ष कंट्रोल डाटा कॉरपोरेशन और मिनोसेटा यूनिवर्सिटी ने क्रोनोर और एनओएस सिस्टम पेश किए। इनकी खासियत टाइमशेयरिंग और एक साथ कई काम किया जाना थी। कंट्रोल डाटा ने ही बाद में यूनिवर्सिटी ऑफ इलियोनिस के साथ मिलकर प्लेटो (PLATO) नाम से ऑपरेटिंग सिस्टम तैयार कियाए जिसकी मदद से पहली बार रियल टाइम चैटिंग और मल्टी यूजर ग्राफिकल गेम्स जैसे फीचरों का सफल निष्पादन संभव हो सका। प्लेटो अपने दौर का सबसे आधुनिक ऑपरेटिंग सिस्टम बन गया।

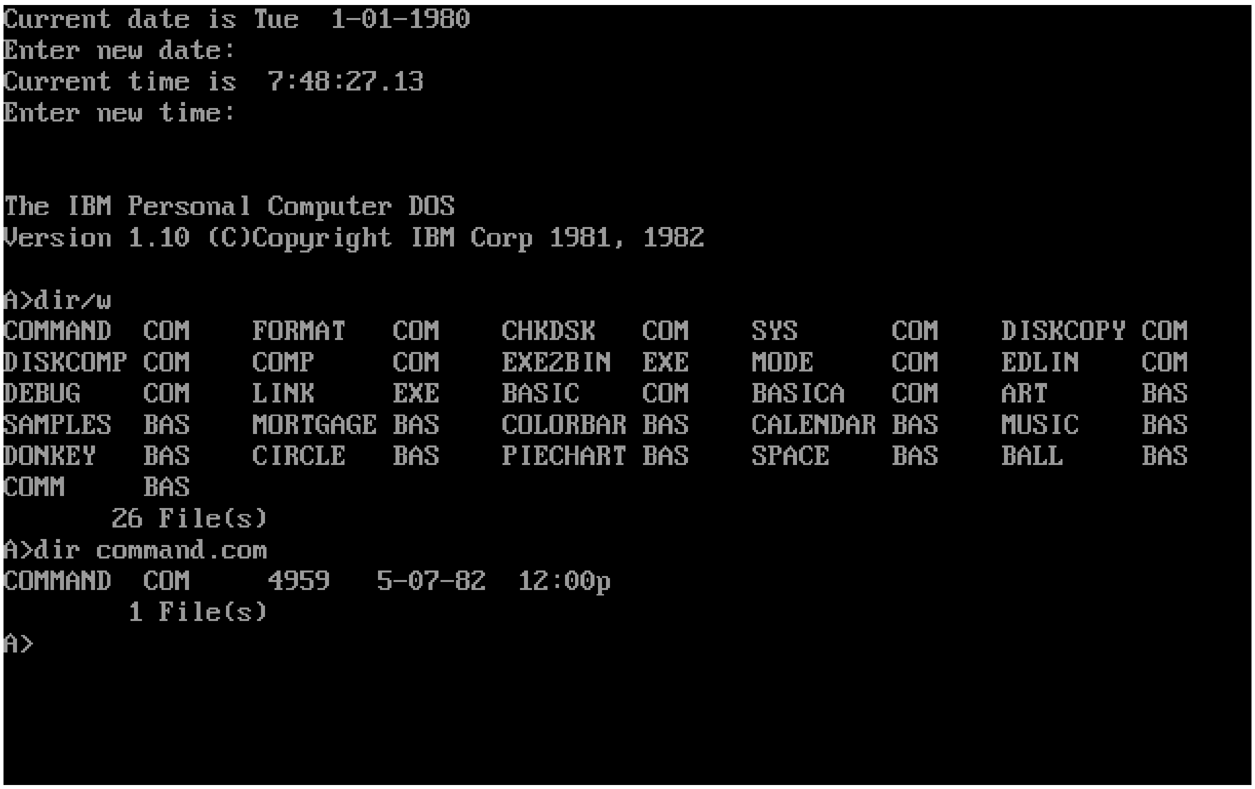
इसी साल पहली कॉमर्शियल कंप्यूटर निर्माता कंपनी यूनिवैक (UNIVAC) ने एक्जेक (EXEC) नाम से ऑपरेटिंग सिस्टम की एक सीरीज पेश की जो रियल टाइम बेस्ड (Real Time Based) थी। इसी तरह जनरल इलेक्ट्रिक और एमआईटी ने जनरल कांप्रेहेन्सिव ऑपरेटिंग सिस्टम (GCOS) तैयार किया। वहीं , डिजिटल इक्विपमेंट कॉरपोरेशन ने टॉप्स -10 (TOPS-10) और टॉप्स-20 (TOPS-20) जैसे ऑपरेटिंग सिस्टम तैयार किए जो मुख्यतः विश्वविद्यालयों के लिए खासे उपयोगी साबित हुए। इनके अलावा भी कई अन्य ऑपरेटिंग सिस्टम लगातार विकसित किए जाते रहे।

माइक्रो कंप्यूटर सिस्टम (Micro Computer OS)

हम जानते हैं कि पहले माइक्रो कंप्यूटर या पर्सनल कंप्यूटर का विकास आईबीएम कंपनी ने 1980 में किया था। उस वक्त यह सिर्फ प्रयोग के तौर पर तैयार किए गए थे। शुरूआती दौर में पर्सनल कंप्यूटरों में अधिक क्षमता भी नहीं थीए लिहाजा इनके लिए अलग से ऑपरेटिंग सिस्टम की जरूरत महसूस नहीं की गईए क्योंकि तब कंप्यूटरों का दैनन्दिन जीवन में कोई विशेष उपयोग नहीं किया जाता था। हालांकिए तब भी इन कंप्यूटरों में

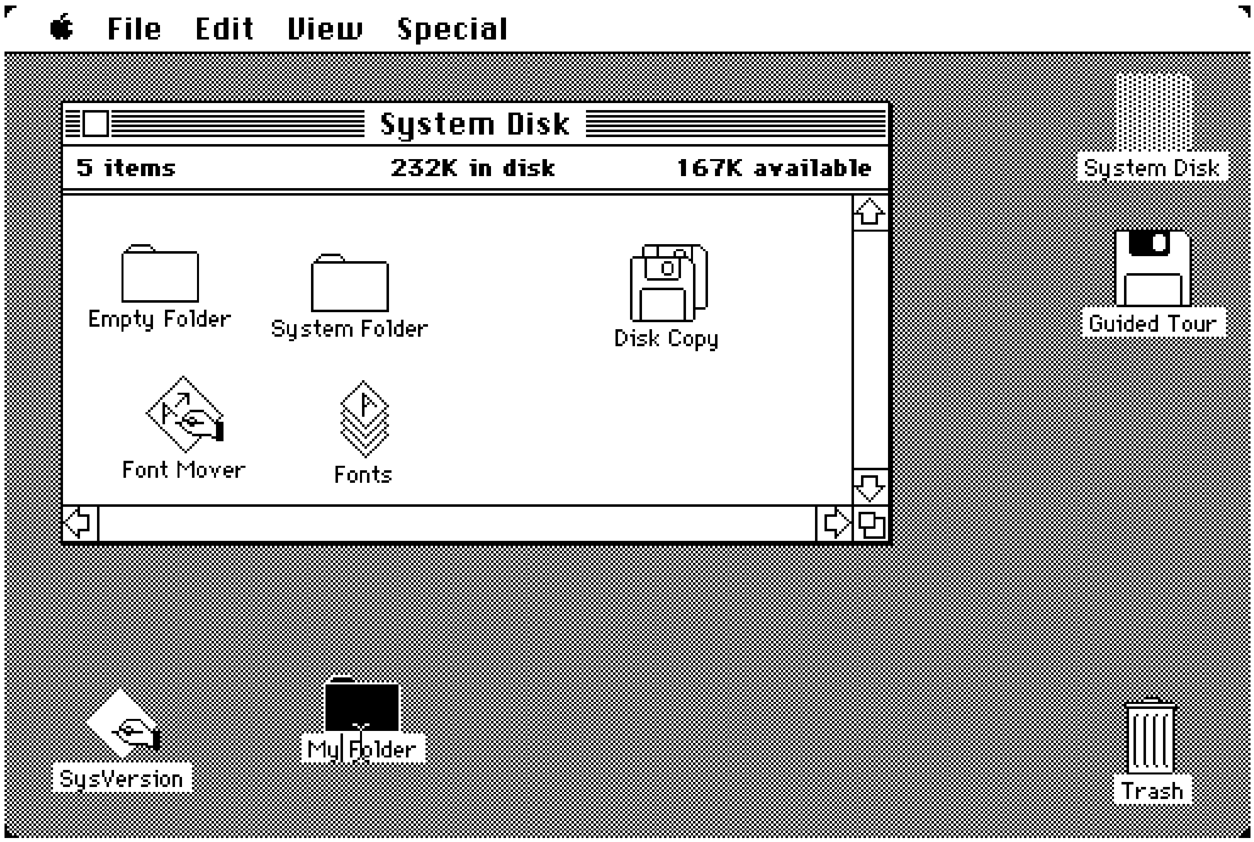
रॉम यानी मेमोरी उपलब्ध रहती थी। उस दौर में इन कंप्यूटरों को मॉनीटर (IBM-DOS or PC-DOS) कहा जाता था।

पर्सनल कंप्यूटर के प्रारंभिक दौर में पहला ऑपरेटिंग सिस्टम था सीपी -एम (CP-M) जो डिस्क ऑपरेटिंग सिस्टम (Disk Operating System) था। लेकिन जल्दी ही माइक्रोसॉफ्ट ने अपना ऑपरेटिंग सिस्टम एमएस-डॉस (MS-DOS) पेश किया। लांचिंग के साथ ही यह सिस्टम सबसे अधिक लोकप्रिय हो गया। इसकी एक बड़ी वजह यह भी थी कि आईबीएम कंपनी ने अपने माइक्रो कंप्यूटरों के लिए इसी ऑपरेटिंग सिस्टम का चयन किया, जिसे तब आईबीएम डॉस या पीसी डॉस (IBM-DOS or PC-DOS) भी कहा जाता था।



(आईबीएम कंप्यूटर में इस्तेमाल किया जाने वाला पीसी डॉस)

दूसरी ओर, कंप्यूटर निर्माता दूसरी बड़ी कंपनी एप्पल (Apple Inc) ने भी लगभग आईबीएम के समानांतर एप्पल मेकिनटोश (Apple Macintosh) नाम से अपना माइक्रो कंप्यूटर पेश किया। इस कंप्यूटर की खासियत थी इसका ऑपरेटिंग सिस्टम मैकिन्टोश परेटिंग सिस्टमए जिसे मैक ओएस (MAC OS) भी कहा जाता है। इस ऑपरेटिंग सिस्टम की मदद से एप्पल कंपनी अपने पर्सनल कंप्यूटर में ग्राफिकल यूजर इंटरफेस (Graphical User Interface - GUI) देने में सफल रही। इस इंटरफेस का तात्पर्य ऐसी व्यवस्था से है , जिसके तहत उपयोगकर्ता मशीन पर चल रहे प्रोग्राम को आइकन (Icons) की मदद से पहचान सके, ताकि उसे काम करने में आसानी हो। सबसे खास बात यह थी कि अपने इस नये प्रोजेक्ट को आगे बढ़ाने के लिए एप्पल कंपनी ने अपने शुरूआती पर्सनल कंप्यूटर प्रोजेक्ट एप्पल 2 (Apple II) को बंद कर दिया था।



(एप्पल मैकिनटोश में इस तरह आइकन बने नजर आते थे)

वर्ष 1985 में 32 बिट आर्किटेक्चर और पेजिंग क्षमता वाली इनटेल 80386 सीपीयू चिप ने पर्सनल कंप्यूटर के विकास में नयी क्रान्ति पैदा की। दरअसल , इस चिप के इस्तेमाल के बाद ही पर्सनल कंप्यूटर मेनफ्रेम और मिनी कंप्यूटरों की तरह मल्टी टास्किंग (Multi Tasking) ऑपरेटिंग सिस्टम को चलाने लायक बन सका। इस बिन्दु को ध्यान में रखते हुए माइक्रोसॉफ्ट कंपनी ने वीएमएस (VMS) ऑपरेटिंग सिस्टम बनाने वाले डेविड कटलर को अपने साथ जोड़ लिया और उन्हें माइक्रोसॉफ्ट के पुराने ऑपरेटिंग सिस्टम (DOS) को आगे बढ़ाते हुए विंडोज ऑपरेटिंग सिस्टम (Windows Operating System) को तैयार करने की कमान सौंप दी गई। दूसरी तरफ एप्पल कंपनी के सह संस्थापक स्टीव जॉब्स का ध्यान भी इनटेल 80386 चिप ने खींचा। स्टीव ने नेक्स्ट कंप्यूटर नाम से अपनी अलग कंपनी बनाई और इसके तहत नेक्स्टस्टेप (NEXTSTEP) ऑपरेटिंग सिस्टम तैयार किया। कालान्तर में एप्पल ने यह सिस्टम खरीद लिया और मैकिनटोश के साथ इसका उपयोग किया। मौजूदा पर्सनल कंप्यूटरों में अधिकतर इस्तेमाल होने वाले ऑपरेटिंग सिस्टम विंडोज और मैकिनटोश ही हैं। हालांकि, लाइनक्स (LINUX) यूनिक्स (UNIX) भी ऑपरेटिंग सिस्टम हैं , लेकिन पर्सनल कंप्यूटरों में इनका बहुत अधिक इस्तेमाल नहीं किया जाता है।

1. ऑपरेटिंग सिस्टम के प्रकार (Types of Operating System)

ऑपरेटिंग सिस्टम को इनकी कार्यक्षमता और इनकी कार्यशैली के आधार पर दो अलग तरह से बांटा जा सकता है। यूनिट के इस हिस्से में दोनों तरीकों से ऑपरेटिंग सिस्टमों को आसानी से जान सकेंगे। कर्यक्षमता के आधार पर ऑपरेटिंग सिस्टम को छह प्रमुख भागों में बांटा जा सकता है। ये हैं एकल एवं बहुल कार्य (Single and Multi Tasking) एकल एवं बहुल उपयोगकर्ता (Single and Multi Users), वितरित सिस्टम (Distributed), टेंपलेटेड (Templated), एंबेडेड (Embedded), लाइब्रेरी (Library) और रियल टाइम (Real Time)

- **एकल और बहुल कार्य (Single and Multi Tasking) -** शुरूआत करते हैं एकल एवं बहुल कार्य सिस्टम से। जैसा कि नाम से ही स्पष्ट है कि सिंगल ऑपरेटिंग सिस्टम वे सिस्टम हैं, जो एक समय में एक ही काम करने में सक्षम हैं, दूसरी ओर मल्टी टास्किंग ऑपरेटिंग सिस्टम उपयोगकर्ता को एक ही वक्त में एक से अधिक काम करने की क्षमता प्रदान करते हैं। मल्टी टास्किंग ऑपरेटिंग सिस्टम टाइम अचीविंग (Time Achieving) के जरिये ऐसा कर पाते हैं। लेकिन इसमें भी ऑपरेटिंग सिस्टम दो तरह से काम करते हैं। पहला है प्रीएंप्टिव और दूसरा को-ऑपरेटिव। प्रीएंप्टिव (Preemptive) ऑपरेटिंग सिस्टम के तहत प्रोसेसर में हर प्रोग्राम के लिए टाइम शेयर (Time Share) कर लिया जाता है , जिससे प्रोसेसर तय समय में एक के बाद एक हर प्रोग्राम पर काम करता है। लेकिन , इसमें परेशानी यह होती है कि एक प्रोग्राम की प्रोसेसिंग पूरी होने के बाद ही दूसरा शुरू हो सकता है। यानी उपयोगकर्ता को दूसरे प्रोग्राम पर जाने के लिए इंतजार करना होता है। इस तरह के ऑपरेटिंग सिस्टम हैं लाइनक्स (LINUX) यूनिक्स (UNIX), सोलरिस (Soalaris), अमीगा (Amiga) दूसरी ओर , को -ऑपरेटिव ऑपरेटिंग सिस्टम में हर काम को इस तरीके से प्रोसेस किया जाता है कि प्रोसेसर में हर काम के लिए अलग टाइम स्लॉट तय करने के साथ प्रोसेसिंग को भी बांट दिया जाता है। इससे एक ही समय में एक साथ अलग - अलग काम करना संभव हो पाता है। विंडोज 16-बिट ऐसा ही ऑपरेटिंग सिस्टम है। हालांकिए विंडोज का 32-बिट ऑपरेटिंग सिस्टम और विंडोज 9x ऑपरेटिंग सिस्टम प्रीएंप्टिव सिस्टम थे।
- **एकल एवं बहुल उपयोगकर्ता (Single and Multi Users) -** एकल यूजर ऑपरेटिंग सिस्टम एकल उपयोगकर्ता के लिए ही उपयोगी होता है। हालांकि , यह भी बहुत अधिक सुविधाएं प्रदान नहीं करता , फिर भी इतनी सहूलियत जरूर होती है कि इसमें एक साथ कुछ प्रोग्राम चलाए जा सकते हैं। दूसरी ओर, मल्टीटास्किंग यूजर ऑपरेटिंग सिस्टम एक से अधिक उपयोगकर्ताओं को डिस्क स्पेस (Disk Space) सुविधा के जरिये कंप्यूटर (Interact) करने की सुविधा प्रदान करता है। इससे इस तरह के ऑपरेटिंग सिस्टम पर एक साथ कई उपयोगकर्ता एक ही समय पर काम करने में सक्षम होते हैं। इसी तरह टाइम शेयरिंग ऑपरेटिंग सिस्टम भी खास तरीकों से प्रोसेसर टाइमिंग , प्रिंटर, मास स्टोरेज और अन्य संसाधनों का एलॉकेशन (Allocation) करता है , जिससे एक ही समय पर अलग -अलग उपयोगकर्ता अलग-अलग संसाधन का उपयोग अपने अभीष्ट परिणाम प्राप्त करने में कर सकें।
- **वितरित सिस्टम (Distributed)-** इस तरह के ऑपरेटिंग सिस्टम को सिंगल एंड मल्टीटास्किंग -यूजर का वृहद और विस्तृत स्वरूप माना जा सकता है। दरअसल , इस सिस्टम के जरिये ऐसे कई कंप्यूटरों को साथ जोड़ा जा सकता है , जो दरअसल भौतिक रूप से एक.दूसरे से दूर हों। यह काम नेटवर्किंग (Networking) के जरिये किया जाता है , जिसकी प्रक्रिया को हम आगे जानेंगे। वस्तुतः इस तरह के सिस्टम का विकास ही नेटवर्किंग की अवधारणा के बाद हुआ। इसके जरिये एक ही समय में एक साथ कई सारे कंप्यूटरों को ऑपरेट किया जाना संभव हो सका। कंप्यूटरों पर को -ऑपरेशन (Co-operation) के तहत होने वाले काम को ही वितरित सिस्टम कहा जाता है।
- **टेंपलेटेड (Templated)-** टेंपलेट का शाब्दिक अर्थ होता है खास पैटर्न (Pattern) यानी किसी खास मकसद की पूर्ति के लिए तैयार किया जाने वाला ऑपरेटिंग सिस्टम। यह ऑपरेटिंग सिस्टम वितरित सिस्टम का और अधिक परिष्कृत स्वरूप है। मुख्यतः इस तरह के ऑपरेटिंग सिस्टम का प्रयोग इंटरनेट बेस्ड (Internet Based), क्लाउड कंप्यूटिंग (Cloud Computing) में किया जाता है। इस तरह के ऑपरेटिंग सिस्टम में किसी भी डाटा , सूचना को वर्चुअलाइज (Virtualize) कर लिया जाता है। इसके बाद यह डाटा या सूचना सर्वर (Server) तक पहुंचा दिया जाता है , जहां वह स्टोर रहता है। अब भविष्य में जब भी किसी उपयोगकर्ता को किसी खास डाटा की आवश्यकता होती है तो टेंपलेट ऑपरेटिंग सिस्टम की मदद से वह आसानी से उसे सर्वर से हासिल कर लेता है।

- **एंबेडेड सिस्टम (Embeded System)-** एंबेडेड ऑपरेटिंग सिस्टम एंबेडेड कंप्यूटरों के लिए बनाए जाते हैं। एंबेडेड कंप्यूटरों का अर्थ उन कंप्यूटरों से है , जिनका निर्माण कुछ खास मकसद से किया जाता है, जो कम आकार , कम स्पेस और कम संसाधनों के बावजूद सुरक्षित और विश्वसनीय तरीके से उपयोगकर्ता के निर्देशों का पालन कर सकें।

(पीडीए आधारित एक मोबाइल डिवाइस)

उदाहरण के लिए इसे पीडीए (Personal Digital Assistant) से समझा जा सकता है। पीडीए दरअसल एक मोबाइल डिवाइस है, जो इंटरनेट से जुड़ सकती है , डाटा और सूचनाएं संग्रहीत कर सकती है और उपयोगकर्ता की जरूरत के मुताबिक जानकारी उपलब्ध करा सकती है। यही वजह है कि पीडीए को हैंडहोल्ड पीसी (Handhold PC) भी कहा जाता था। हालांकि वर्ष 2010 के बाद स्मार्टफोन के विकास और आईफोन ऑपरेटिंग सिस्टम (i-OS) और एंड्रॉयड (Android) के विकास के बाद पीडीए का उपयोग काफी कम , लगभग नगण्य, रह गया।

- **रियल टाइम सिस्टम (Real Time Operating System)-** इस तरह के ऑपरेटिंग सिस्टम यह सुनिश्चित करते हैं कि उपयोगकर्ता जो काम करना चाहता है या जो डाटा इस्तेमाल करना चाहता है , वह निश्चित समयावधि में परिणाम के रूप में उसके सामने उपलब्ध हो। ये ऑपरेटिंग सिस्टम सिंगल टास्किंग भी हो सकते हैं और मल्टी टास्किंग भी। अंतर सिर्फ यह होता है कि मल्टी टास्किंग होने की स्थिति में ये ऑपरेटिंग सिस्टम निर्धारित कलन विधियों (Scheduled Algorithms) की मदद से लक्ष्य हासिल करता है। कलन विधियां, गणितीय शब्द है।

ये दरअसल किसी एप्लीकेशन प्रोग्राम के वे स्टेप हैं , जिनपर चलकर प्रोसेसर उपयोगकर्ता को अभीष्ट परिणाम उपलब्ध कराता है। एल्गोरिथम में भी ये ऑपरेटिंग सिस्टम दो तरह से काम करते हैं। पहला है इवेंट ड्राइवन सिस्टम (Event Driven System) इसके तहत ऑपरेटिंग सिस्टम उपयोगकर्ता की ओर से मिले आदेशों को प्राथमिकता (Priority) के क्रम में तय करता है और एक के बाद एक तय समय में इन्हें पूरा करता है। वहीं , टाइम शेयरिंग सिस्टम (Time Sharing System) में कार्यों के लिए समय निर्धारण किया जाता है।

- **लाइब्रेरी (Library)-** लाइब्रेरी ऑपरेटिंग सिस्टम भी मुख्यतः कंप्यूटर नेटवर्किंग से जुड़ा हुआ है। यह सिस्टम दरअसल किसी खास तरह की नेटवर्किंग में इस्तेमाल किए जाने वाले सभी ऑपरेटिंग सिस्टमों का एक समूह है, जो लाइब्रेरी के स्वरूप में उपलब्ध रहता है।

प्रमुख ऑपरेटिंग सिस्टम (Some Operating Systems)

कंप्यूटर के विकास के अनुक्रम में ही ऑपरेटिंग सिस्टमों का भी विकास तेजी से हुआ। जिस हिसाब से जरूरतें बढ़ती गईं, उसी हिसाब से लगातार शोध और अनुसंधानों की मदद से ऑपरेटिंग सिस्टमों की ईजाद कर समस्याओं का हल निकाला जाता रहा। कुछ प्रमुख ऑपरेटिंग सिस्टमों के बारे में हम यहां जानने का प्रयास करेंगे-

- **यूनिक्स (Unix)**- यूनिक्स मल्टीटास्किंग, मल्टीयूजर कंप्यूटर ऑपरेटिंग सिस्टम है , जिसे 1970 में अमेरिका की अमेरिकन टेलीफोन एंड टेलीग्राफ कंपनी (AT&T) की बेल रिसर्च लैब (Bell Lab) में केन थॉमसन, डेनिस रिची की टीम ने तैयार किया था। टीम ने यूनिक्स सिस्टम बनाने का प्रोजेक्ट 1968 में शुरू किया था। शुरूआत में यह ऑपरेटिंग सिस्टम असेंबलिंग लैंग्वेज (Assembling Language) में लिखा गया था, जो उस समय प्रोग्रामिंग की प्रचलित भाषा थी। प्रारंभ में यह ऑपरेटिंग सिस्टम सिर्फ बेल लैब के ही कार्यों के निष्पादन के लिए तैयार किया गया था। बाद में एटीएंडटी ने यह ऑपरेटिंग सिस्टम अन्य संस्थाओं को भी देना शुरू किया। इसके लिए यूनिक्स के एकेडमिक और कॉमर्शियल दो वर्जन तैयार किए गए। इसके शुरूआती उपयोगकर्ताओं में यूनिवर्सिटी ऑफ कैलीफोर्निया, माइक्रोसॉफ्ट, बर्कले, आईबीएम, सन माइक्रोसिस्टम्स जैसी कंपनियां रहीं। यूनिक्स अपनी खास पद्धति पर काम करता है, जिसे अकसर कंप्यूटर विशेषज्ञ यूनिक्स फिलॉसफी (Unix Pholosophy) भी कहते हैं। यह सिस्टम उपयोगकर्ता को ऐसे टूल्स (Tools) का समूह उपलब्ध कराता है , जिनमें से हरेक एक खास फंक्शन (Function) को पूरा करते हैं। इसके अलावा यह इन सभी टूल्स की मदद से संयुक्त यूनिफाइड फाइल सिस्टम और शेल (Shell) कमांड सिस्टम भी विकसित करता है, जिससे वर्कफ्लो (Workflow) में मदद मिलती है। बेहतरीन कार्यक्षमता और उपयोगकर्ता के लिए खासा मददगार साबित हुआ यूनिक्स पहला पोर्टेबल ऑपरेटिंग सिस्टम (Portable Operating System) माना जाता है। यही वजह है कि इसके बाद विकसित हुए अधिकतर ऑपरेटिंग सिस्टमों का मूल आधार यूनिक्स ही रहा। यही नहीं, समय के साथ जैसे -जैसे प्रोग्रामिक भाषाएं विकसित होती रहीं , वैसे -वैसे हर भाषा में यूनिक्स को हर बार नये स्वरूप में तैयार किया गया।
- **यूनिक्स लाइक फैमिली (Unix Like Family)**- यूनिक्स कंप्यूटर के विकास का बड़ा आविष्कार था। मेनफ्रेम और मिनी कंप्यूटरों के लिहाज से यह बेहद उपयोगी था , जहां बल्क डाटा (Bulk data) आता था। एटीएंडटी-बेल रिसर्च लैब में विकास के बाद यूनिक्स के टेडमार्क द ओपन ग्रुप ने हासिल कर लिए , जिसने एचपी, आईबीएम, एप्पल और सन माइक्रोसिस्टम्स को यूनिक्स ऑपरेटिंग सिस्टम को अपने कंप्यूटरों में प्रयोग करने को ही अधिकृत किया है। ऐसे में यूनिक्स से मिलते -जुलते ऑपरेटिंग सिस्टम तैयार करने शुरू किए गए। यूनिक्स के समकक्ष कई नये ऑपरेटिंग सिस्टम उभरकर सामने आएए जिन्हें यूनिक्स लाइक फैमिली कहा जाता है। इनमें लाइनक्स (Linux), वी सिस्टम (V System), बीएसडी (BSD) शामिल हैं। इनमें से अधिकतर का उपयोग एकेडमिक संस्थाओं, इंजीनियरिंग कंपनियों के सर्वर में किया जाता है।
- **लाइनक्स (Linux)**- यह ऑपरेटिंग सिस्टम फिनलैंड के एक इंजीनियरिंग छात्र लाइनस टोर्वेल्ड्स ने तैयार किया। पढ़ाई के दौरान एक प्रोजेक्ट पर काम करते हुए लाइनस ने अपने इस ऑपरेटिंग सिस्टम के बारे में एक अखबार में जानकारी प्रकाशित की। हालांकि , तब तक यह पूरी तरह तैयार नहीं हुआ था, लेकिन अखबार में प्रकाशन के बाद कई विशेषज्ञ, इंजीनियरिंग छात्रों ने लाइनस को इस प्रोजेक्ट में मदद की, अपेक्षित सुधार किए, जिसके बाद लाइनक्स सिस्टम वजूद में आया।

(मौजूदा दौर में सबसे अधिक इस्तेमाल किए जाने वाले एंड्रॉयड सिस्टम वाले मोबाइल फोन का ऑपरेटिंग सिस्टम लाइनक्स ही है)

लाइनक्स को यूनिक्स लाइक ऑपरेटिंग सिस्टम माना जाता है, लेकिन अपनी तरह के दूसरे सिस्टम से लाइनक्स इस लिहाज में अलग है कि इसे बनाने में यूनिक्स कोड का इस्तेमाल नहीं किया गया है। ओपन लाइसेंस मोड होने के कारण लाइनक्स कोड अध्ययन और सुधारीकरण के लिए भी खुला है। अपनी इसी खूबी के कारण लाइनक्स सुपर कंप्यूटरों से लेकर स्मार्टवॉच तक का ऑपरेटिंग सिस्टम बन गया। मल्टीटास्किंग , मल्टीयूजर सर्वर से लेकर मोबाइल फोन जैसे एंबेडेड कंप्यूटरों में भी लाइनक्स का पूरा इस्तेमाल किया जाता है। गूगल क्रोम और क्रोम ब्राउजर भी लाइनक्स आधारित हैं।

- **मैक ओएस (Mac-OS)-** मैकिन्टोश ऑपरेटिंग सिस्टम (Macintosh Operating System) एप्पल कंपनी की ओर से तैयार किया गया ऑपरेटिंग सिस्टम है। ग्राफिकल यूजर इंटरफेस (GUI) आधारित यह पहला ऑपरेटिंग सिस्टम नहीं था, लेकिन जीयूआई का पहला सबसे अधिक लोकप्रिय सिस्टम बना।

मैक से पहले 1980 में जेरॉक्स कॉरपोरेशन (Xerox Corporation) ने सबसे पहले जीयूआई पर शोध किया। इस शोध से सिद्ध हुआ कि हाथ में पकड़े जा सकने वाले किसी साधन (Tool) की मदद से कंप्यूटर को निर्देश समझाना अधिक आसान और सुगम है। कंपनी ने अपने इस शोध के आधार पर अपना खुद का कंप्यूटर जेरॉक्स स्टार (Xerox Star) भी लांच किया , लेकिन इसमें जीयूआई सिस्टम यानी ग्राफिकल यूजर इंटरफेस की परिकल्पना पूरी तरह सफल नहीं हो सकी थी। दूसरी ओर , एप्पल भी इसी विषय पर शोध कर रहा था और उसने संपूर्ण जीयूआई आधारित ऑपरेटिंग सिस्टम यानी मैक तैयार कर बाजी मार ली। एप्पल ने वर्ष 1984 में अपना पहला मैक ऑपरेटिंग सिस्टम पेश किया था, जिसे बाद में परिष्कृत किया जाता रहा।

- **माइक्रोसॉफ्ट विंडोज (Microsoft Windows)-** पर्सनल या माइक्रो कंप्यूटर आज डेस्कटॉप (Desktop) या लैपटॉप (laptop) के रूप में लगभग हर घर में इस्तेमाल हो रहा है। और जब भी हम

अपना डेस्कटॉप या लैपटॉप खोलते हैं तो उसमें हमें विंडोज 7 , 8 या एक्सपी ही बतौर ऑपरेटिंग सिस्टम नजर आती है। इसकी वजह यह है कि दुनियाभर के कुल वेब कनेक्टेड कंप्यूटरों में से 88.9 प्रतिशत में विंडोज ऑपरेटिंग सिस्टम इस्तेमाल किया जाता है। इसके बारे में हम आगे विस्तार से जानेंगे।

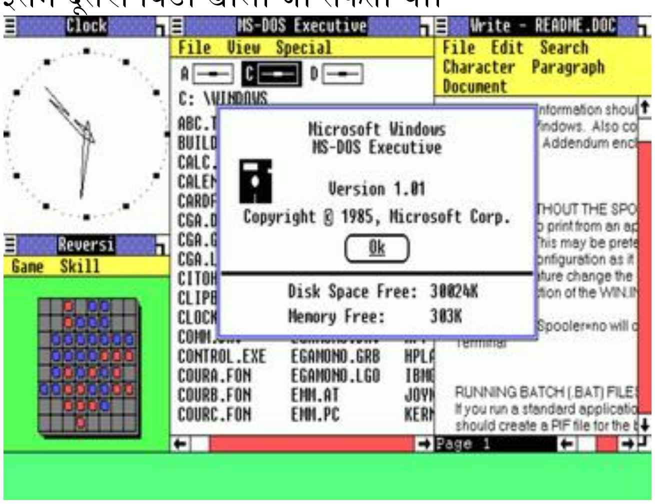
- **माइक्रोसॉफ्ट विंडोज (Microsoft Windows)-**

वर्ष 1985 में माइक्रोसॉफ्ट कॉरपोरेशन ने पहली बार विंडोज 1.0 ऑपरेटिंग सिस्टम को लांच किया था। पूरी तरह ग्राफिकल यूजर इंटरफेस आधारित यह सिस्टम जल्द ही बेहद लोकप्रिय हो गया। यही वजह थी कि आईबीएम ने अपने कंप्यूटरों के लिए इस ऑपरेटिंग सिस्टम को आधिकारिक रूप से स्वीकृत और उपयोग किया। आईबीएम के अलावा भी अन्य कंप्यूटर निर्माता कंपनियों ने अपने पर्सनल कंप्यूटरों में विंडोज का ही इस्तेमाल किया है।

माइक्रोसॉफ्टकॉरपोरेशन 1985 से लेकर 2015 तक अभी तक विंडोज 1.0 से लेकर विंडोज 10 तक ऑपरेटिंग सिस्टम के अलग.अलग वर्जन लांच कर चुका है। विंडोज ऑपरेटिंग सिस्टम का हर नया वर्जन यानी संस्करण पिछले वाले संस्करण में रह गई कमियों को दूर करके बनाया जाता रहा , जिसकी वजह से हर नया विंडोज सिस्टम उपयोगकर्ताओं के लिए और अधिक उपयोगी और लाभकारी बनता चला गया। दुनियाभर के अधिकतर कंप्यूटरों में विंडोज सिस्टम इस्तेमाल किए जाने के पीछे शायद यही वजह है। यहां यह उल्लेखनीय है कि विंडोज 7.0 सबसे अधिक लोकप्रिय और सर्वाधिक इस्तेमाल किया जाने वाला ऑपरेटिंग सिस्टम है। विंडोज में समय के साथ आए बदलावों को हम निम्नवत समझ सकते हैं-

विंडोज 1.0- माइक्रोसॉफ्टकी ओर से वर्ष 1985 में यह सबसे पहला जीयूआई आधारित ऑपरेटिंग सिस्टम लांच किया गया था, इसकी सबसे बड़ी खासियत उपयोगकर्ता की जरूरत के हिसाब से मल्टीटास्किंग करना भी थी। 32X32 पिक्सल (Pixels) के आइकन और कलर स्कीम इस ऑपरेटिंग सिस्टम की विशेषताएं रहीं।

विंडोज 1.2- विंडोज 1.0 की कामयाबी के दो साल बाद यानी वर्ष 1987 में माइक्रोसॉफ्टने अपने ऑपरेटिंग सिस्टम का यह परिष्कृत स्वरूप पेश किया। इस ऑपरेटिंग सिस्टम की विशेषता यह थी कि इसमें विंडोज की ओवरलैपिंग (Overlapping) की सुविधा उपलब्ध थी। ओवरलैपिंग का मतलब यह है कि एक विंडो के उपर इसमें दूसरी विंडो खोली जा सकती थी।



(विंडोज का पहला जीयूआई ऑपरेटिंग सिस्टम विंडोज 1.0)

विंडोज 2.10- वर्ष 1987 में ही माइक्रोसॉफ्ट कंपनी ने अपना अगला ऑपरेटिंग सिस्टम विंडोज 2.10 ऑपरेटिंग सिस्टम के नाम से लांच किया। इस ऑपरेटिंग सिस्टम की खासियत रही आभासी मशीन (Virtual Machines) इस मशीन का तात्पर्य ऐसे सिस्टम से है जो मुख्य कंप्यूटर से जुड़कर एक ऐसी व्यवस्था बनाता है, जो पूरे ऑपरेटिंग सिस्टम पर निगरानी रखते हुए जरूरत के हिसाब से किसी काम को करने के लिए

हार्डवेयर को इस तरह नियंत्रित करते हैं कि वे एक ही कंप्यूटर में अवस्थित होने के बावजूद अलग -अलग काम करने में सक्षम हों।

विंडोज 3.0- माइक्रोसॉफ्ट ने वर्ष 1990 में यह ऑपरेटिंग सिस्टम जारी किया। ग्राफिकल यूजर इंटरफेस (GUI) इंटरफेस प्लेटफॉर्म पर यह विंडोज का सबसे सफल ऑपरेटिंग सिस्टम रहा। उस दौर के जीयूआई आधारित मैक और अमीगा ऑपरेटिंग सिस्टम के मुकाबले यह सिस्टम उतारा गया था , जो काफी हद तक उपयोगकर्ताओं को लुभाने में कामयाब भी रहा। इस ऑपरेटिंग सिस्टम में पहली बार आइकन आधारित प्रोग्राम मैनेजर (Program Manager) और फाइल मैनेजर (File Manager) की व्यवस्था दी गई।

इससे पहले माइक्रोसॉफ्टकंपनी के सभी पुराने ऑपरेटिंग सिस्टम में डॉस (DOS) आधारित फाइल और प्रोग्राम मैनेजर दिया जाता था। लेकिन विंडोज ऑपरेटिंग सिस्टम में ऐसी नयी सुविधाएं दी गई , जिनकी मदद से सिस्टम को केंद्रीयकृत (Centralised) करना आसान हो गया। इनके अलावा विंडोज का बेहद लोकप्रिय गेम सालिटेयर पहली बार इसी सिस्टम में लांच हुआ। यही नहीं , आज टाइपिंग के लिए सर्वाधिक प्रयोग किया जाने वाला नोटपैड, कैल्कुलेट और कलरबार तथा विशेष मेनु के साथ पेंटब्रश भी परिष्कृत स्वरूप में इसी ऑपरेटिंग सिस्टम में लांच किए गए।

विंडोज 3.1- विंडोज का यह नया परिष्कृत ऑपरेटिंग सिस्टम वर्ष 1992 में पेश किया गया। इस सिस्टम की खासियत थी मल्टीमीडिया और नेटवर्किंग की क्षमता। खास बात यह थी कि इस सिस्टम में पहली बार माइक्रोसॉफ्टमेल (Microsoft Mail) की सुविधा उपयोगकर्ताओं को मिली। इस सिस्टम में माइक्रोसॉफ्ट ने नोटपैड के लिए तीन फॉन्ट का इस्तेमाल किया , ये थे - Times New Roman, Arial और Courior New इनके अलावा चिह्नों (Symbols) को भी शामिल किया गया।

विंडोज 3.11- यह ऑपरेटिंग सिस्टम वर्ष 1993 में लांच किया गया था। इसकी खासियत यह थी कि इसमें 32 बिट नेटवर्किंग और 32 बिट फाइल सिस्टम की सुविधा उपलब्ध थी। इसके जरिये यह ऑपरेटिंग सिस्टम मल्टीटास्किंग के साथ मल्टीयूजर भी बन गया। इससे एक ही ऑपरेटिंग सिस्टम से एक साथ 20 से अधिक कंप्यूटरों को जोड़ना संभव हो सका। माइक्रोसॉफ्ट की ओर से यह ऑपरेटिंग सिस्टम इस तरह तैयार किया गया था कि यह पर्सनल कंप्यूटरों के अलावा नेटवर्किंग उपयोगकर्ताओं और ऑफिस में उपयोग के लिए तैयार किया गया था।

विंडोज 95- वर्ष 1995 में माइक्रोसॉफ्ट ने विंडोज 95 ऑपरेटिंग सिस्टम लांच किया। यह पूर्णतः 32 बिट ऑपरेटिंग सिस्टम था , जिसकी मदद से मल्टीटास्किंग और नेटवर्किंग का काम और अधिक आसान होता गया। सबसे बड़ी खासियत यह थी कि इस सिस्टम में माइक्रोसॉफ्ट ने अपने शुरूआती सिस्टम डॉस (DOS) और विंडोज 3.1 के फीचर्स को संयुक्त करने में कामयाबी हासिल की। प्लग एंड प्ले फीचर इस विंडोज के सबसे बड़े साधन (Tools) थे। आज भी हम कंप्यूटर पर जो स्टार्ट बटन देखते हैं (जिस पर क्लिक करने के बाद कंप्यूटर पर मौजूद सभी प्रोग्राम , फाइल मैनेजर आदि की सारिणी खुल जाती है) वह सबसे पहले इसी विंडोज ऑपरेटिंग सिस्टम में पेश किया गया था। यही नहीं , जब भी हम कोई प्रोग्राम बंद करना चाहते हैं तो उसके लिए हमें लंबी प्रोसेस के बजाय सीधे क्लोज (Close) बटन पर क्लिक करना होता है। यह क्लोज बटन भी सबसे पहले विंडोज 95 में ही शामिल किया गया था।

(विंडोज 95 ऑपरेटिंग सिस्टम का होमपेज)

विंडोज 98- वर्ष 1998 में माइक्रोसॉफ्ट ने यह ऑपरेटिंग सिस्टम लांच किया। इस ऑपरेटिंग सिस्टम की सबसे बड़ी खासियत यह थी कि इसकी मदद से इंटरनेट का इस्तेमाल कर पाना संभव और सुगम हो सका।
पहली बार इस ऑपरेटिंग सिस्टम में माइक्रोसॉफ्ट ने इंटरनेट एक्सप्लोरर (Internet Explorer) 4.01 दिया, इसके अलावा इंटरनेट पर इस्तेमाल की जा सकने वाली अन्य एप्लीकेशन जैसे आउटलुक एक्सप्रेस , विंडोज एक्सप्रेस बुक, फ्रंटपेज एक्सप्रेस , माइक्रोसॉफ्ट चैट, पर्सनल वेब सर्वर , वेब पब्लिशिंग विजार्ड , नेट मीटिंग भी इस ऑपरेटिंग सिस्टम में शामिल की गईं।
विंडोज ने वर्ष 1999 में इस ऑपरेटिंग सिस्टम में कुछ और सुधार करते हुए विंडोज 98 सेकंड एडिशन (SE) लांच किया। इस सिस्टम में इंटरनेट एक्सप्लोरर को और अधिक परिष्कृत करते हुए 5.0 वर्जन पेश किया गया। इसके अलावा पिछले सिस्टम में शामिल नेट शो प्लेयर की जगह विंडोज मीडिया प्लेयर भी डाला गया।
विंडोज 2000 एमई- इस ऑपरेटिंग सिस्टम का मूल आधार भी विंडोज 98 ही था। वर्ष 2000 में लांच किया गया यह ऑपरेटिंग सिस्टम इंटरनेट के बढ़ते स्कोप को ध्यान में रखते हुए विकसित किया गया था। इसमें अधिकतर फीचर्स विंडोज 98 वाले ही थे , लेकिन इसमें यह सुविधा दी गई थी कि इसकी मदद से इंटरनेट पर नेटवर्किंग का काम आसान हो सके। यही वजह थी कि इसे विंडोज एनटी भी कहा जाता है। माइक्रोसॉफ्ट ने इस सिस्टम के चार वर्जन प्रोफेशनल , सर्वर, एडवांस्ड सर्वर और डाटा सर्वर लांच किए। इससे यह सिंगल यूजर से लेकर मल्टी यूजर तक के लिए उपयोगी ऑपरेटिंग सिस्टम बन सका।

- **विंडोज XP-** विंडोज एनटी फैमिली की अगली कड़ी के तौर पर वर्ष 2004 में यह ऑपरेटिंग सिस्टम लांच किया गया। शुरूआत में यह सिस्टम व्यावसायिक उपयोग के लिए ही तैयार किया जा रहा था , लेकिन पर्सनल कंप्यूटरों की बढ़ती मांग को देखते हुए इसे पर्सनल और व्यावसायिक दोनों उपयोग के लिए बनाया गया। इस सिस्टम की खासियत इसका बेहतर जीयूआई , सुधारीकृत हार्डवेयर सपोर्ट , विस्तृत मल्टीमीडिया श्रृंखला रहीं। विंडोज एक्सपी इस कदर लोकप्रिय हुआ कि लांचिंग के महज पांच साल के भीतर चार लाख कंप्यूटरों पर यह ऑपरेटिंग सिस्टम इंस्टॉल कर लिया गया था। वर्ष 2014 में

पूरी तरह बंद होने तक यह ऑपरेटिंग सिस्टम प्रयोग करने वाले कंप्यूटर उपयोगकर्ताओं की संख्या दुनियाभर में दस लाख से भी अधिक हो चुकी थी।

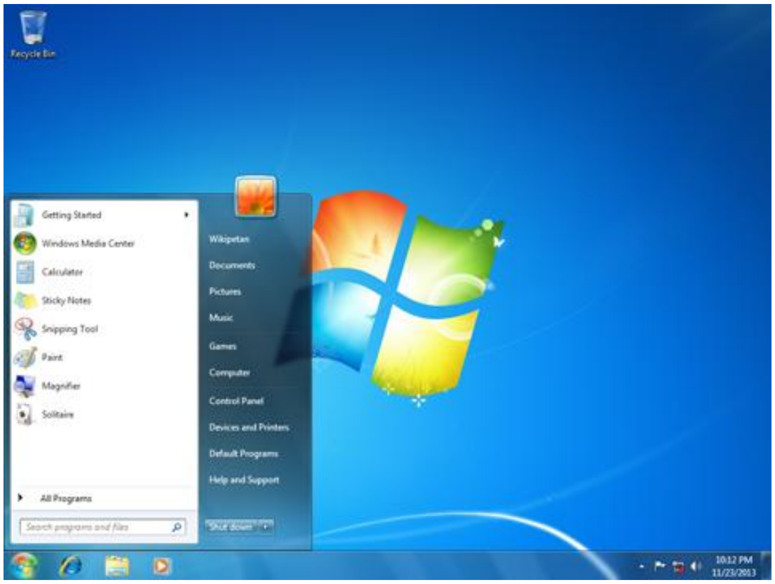
विंडोज विस्टा- वर्ष 2007 में विंडोज का यह ऑपरेटिंग सिस्टम लांच किया गया। इसमें नेटवर्किंग की बेहतर सुविधाओं के साथ प्रिंट, ऑडियो प्ले, विंडोज डीवीडी मेकर जैसे नये फीचर्स भी शामिल किए गए। इस सिस्टम में सबसे अहम खासियत थी इसका ऐरो ग्लास लुक (Aero Glass Look) इसके तहत विंडोज के पिछले ऑपरेटिंग सिस्टम में चले आ रहे ग्राफिक यूजर इंटरफेस (Graphical User Interface) को रि-डिजाइन करने के साथ आकर्षक स्वरूप दिया गया। इसके तहत लेआउट में बदलाव के साथ एप्लीकेशन में भी उपयोगकर्ता के लिए उपयोगी परिवर्तन किए गए। संचार (Communication) के स्तर पर यह विंडो प्रोग्राम लिखने वाले विशेषज्ञों के लिए खासी मददगार साबित हुई। इसके अलावा इस विंडोज में नेटवर्किंग पर खासा ध्यान दिया गया था। इसके तहत इस ऑपरेटिंग सिस्टम की मदद से अलग -अलग कंप्यूटरों पर मल्टीमीडिया , फाइलों का आदान-प्रदान कर पाना संभव हो सका। लेकिन परेशानी यह थी कि इस विंडोज को चलाने के लिए सिस्टम में काफी हैवी हार्डवेयर की जरूरत होती थी।

इसके अलावा इसकी लाइसेंसिंग प्रक्रिया भी काफी जटिल थी। सुरक्षा के पहलू पर भी इसकी गुणवत्ता को लेकर सवाल उठते रहे। इसके बावजूद वर्ष 2009 में विंडोज के नये वर्जन विंडोज 7 की लांचिंग तक दुनियाभर में चार लाख से अधिक इंटरनेट यूजर्स विस्टा का प्रयोग करने लगे थे। हालांकि, यह संख्या विंडोज एक्सपी से काफी कम थी।

(विंडोज विस्टा का ऐरो ग्लास लुक)

विंडोज 7 - माइक्रोसॉफ्टने वर्ष 2009 में सिर्फ पर्सनल कंप्यूटर आधारित अपना पहला ऑपरेटिंग सिस्टम विंडोज 7 लांच किया। आलोचकों ने विंडोज विस्टा की जिन कमियों को उजागर (Point Out) किया था , कंपनी ने विंडोज 7 में उन्हें दूर करने पर फोकस किया।

विंडोज ऐरो में लगातार सुधार के साथ इस सिस्टम में कुछ नये फीचर्स जोड़े गएए जिनमें इंटरनेट एक्सप्लोरर 8ए विंडोज मीडिया प्लेयरए विंडोज मीडिया सेंटरए सुरक्षात्मक प्रक्रियाओं के लिए एक्शन सेंटरए नया रिडिजाइन्ड टास्कबार और लाइब्रेरी शामिल हैं। इस सिस्टम को इस तरह तैयार किया गया कि यह कंप्यूटर के हार्डवेयर और सॉफ्टवेयरके बीच बेहतर सामंजस्य स्थापित करने का जरिया बन सके। सबसे बड़ी बात यह थी कि जिन आलोचकों ने विंडोज विस्टा पर सवालिया निशान खड़े किए थे , उन्होंने ही विंडोज 7 को अब तक का बेहतरीन ऑपरेटिंग सिस्टम करार दिया।



(विंडोज 7 की होमस्क्रीन)

विंडोज 7 माइक्रोसॉफ्ट कंपनी के लिए बेहतरीन वरदान साबित हुआ। विशेष पहलू यह है कि कंपनी ने ऑनलाइन रिटेल कंपनी Amazon.com पर अपने इस उत्पाद की बिक्री शुरू की थी और महज छह महीने के भीतर ही एक लाख से अधिक ग्राहकों ने यह ऑपरेटिंग सिस्टम खरीद लिया जो 2012 तक करीब साढ़े साठ लाख हो गए। ताजा आंकड़ों पर नजर डालें तो विंडोज 7 डेस्कटॉप ऑपरेटिंग सिस्टम के मार्केट में 47.77 प्रतिशत हिस्सेदारी रखता है। यह माइक्रोसॉफ्ट का सबसे अधिक उपयोग किया जाने वाला ऑपरेटिंग सिस्टम है।

विंडोज 8- माइक्रोसॉफ्ट ने वर्ष 2012 में विंडोज 8 नाम से नया पर्सनल कंप्यूटर ऑपरेटिंग सिस्टम लांच किया। इसे हम निम्न चित्र से आसानी से समझ सकेंगे-

(विंडोज 8 ऑपरेटिंग सिस्टम की होमस्क्रीन)

यह सिस्टम दरअसल, इस तरीके से डिजाइन किया गया है कि यह टैबलेट का इस्तेमाल करने वाले उपभोक्ताओं के लिए मददगार साबित हो सके। मोबाइल फोन की दुनिया में इस समय तक एंड्रॉयड (Android) आईफोन ऑपरेटिंग सिस्टम (i-OS) विंडोज से काफी आगे निकल चुके थे। मूलतः विंडोज 8 का स्वरूप इस तरह रखा गया है कि इसे मेट्रो डिजाइन (Metro Design) कहा जाता है। इसकी होम स्क्रीन पर प्रोग्राम और एप्लीकेशन पिछली विंडोज की तरह सारिणी में दिखने के बजाय ग्रिड में नजर आते हैं, ठीक वैसे ही जैसे हमें अपने मोबाइल फोन में दिखते हैं। माइक्रोसॉफ्ट ने इस ऑपरेटिंग सिस्टम को इस तरह तैयार किया है कि यह माउस के साथ अंगुलियों से छूकर भी परफॉर्म (Perform) करे, यानी यह ऑपरेटिंग सिस्टम टचस्क्रीन (Touchscreen) प्रक्रिया पर काम करता है। इसके अलावा सुरक्षा की दृष्टि से इस ऑपरेटिंग सिस्टम में इन.बिल्ट ;प्द ठनपसजद्ध एंटीवायरस (Antivirus) उपलब्ध है, साथ ही यह माइक्रोसॉफ्ट स्मार्ट स्क्रीन फिशिंग फिल्टरिंग (Microsoft Smart Screen Phishing Filtering) सिस्टम से भी ऑनलाइन जुड़ सकता है, जो वायरस से इस सिस्टम की रक्षा करता है। जुलाई 2015 में माइक्रोसॉफ्ट ने अपना नवीनतम ऑपरेटिंग सिस्टम विंडोज 10 लांच किया है।

ऑपरेटिंग सिस्टम का बाजार(Market Share of OSs)

कंप्यूटर और मोबाइल फोन के बढ़ते इस्तेमाल ने दुनिया को ग्लोबल विलेज (Global Village) की शक्ल दे दी है। टनों वजनी मशीन से प्रारंभ हुई कंप्यूटर की विकास यात्रा आज महज 100-150 ग्राम वजनी मोबाइल फोन तक आ चुकी है। इसके पीछे जहां वैज्ञानिक शोधों -अनुसंधानों का परिणाम है, वहीं इसके पीछे लगातार परिष्कृत होते गए ऑपरेटिंग सिस्टम भी महत्वपूर्ण हैं। इन दिनों दुनियाभर में कंप्यूटरों और मोबाइल फोन में इस्तेमाल किए जा रहे ऑपरेटिंग सिस्टम के कितने उपभोक्ता हैं और बाजार में कौन सा ऑपरेटिंग सिस्टम कितना शेयर रखता है, यह हम निम्न सारिणी से समझ सकते हैं-

ऑपरेटिंग सिस्टम	उपभोक्ता
एंड्रॉयड	878
विंडोज	328
मैक और आईफोन	267
ब्लैकबेरी	24
अन्य	803
कुल	2300

(नोटः यह आंकड़े वर्ष 2013 के हैं, स्रोतः गूगल)

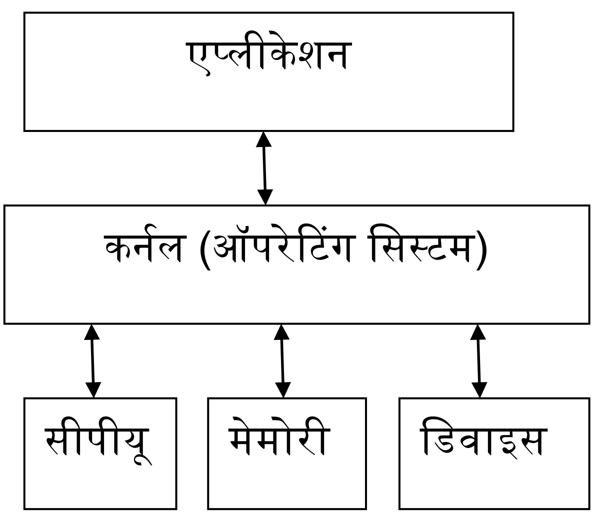
14.5 ऑपरेटिंग सिस्टम के घटक (Operating System Components)

विकीपीडिया (Wikipedia) पर नजर डालें तो उसमें ऑपरेटिंग सिस्टम की परिभाषा कुछ यूं दी गई है , 'ऑपरेटिंग सिस्टम किसी कंप्यूटर का मेरूदंड होता है , जो इसके सॉफ्टवेयर और हार्डवेयर को नियंत्रित रखता है। यह हार्डवेयर और सॉफ्टवेयर के बीच सेतु का काम करता है और सॉफ्टवेयर घटक होता है। इसकी मदद से एक से अधिक सीपीयू में भी प्रोग्राम चलाए जा सकते हैं।

इस यूनिट के अध्ययन के जरिये अब तक हम ऑपरेटिंग सिस्टम के कार्य , उसके प्रकार , ऑपरेटिंग सिस्टम के महत्व आदि से अच्छी तरह वाकिफ हो चुके हैं। लेकिन यह जानना भी हमारे लिए बहुत आवश्यक है कि ऑपरेटिंग सिस्टम के मुख्य घटक (Components) क्या हैं। यूनिट के इस हिस्से में हम इन्हीं घटकों के बारे में विस्तार से जानेंगे। ये घटक ऑपरेटिंग सिस्टम के वे हिस्से हैं , जिनकी बदौलत ऑपरेटिंग सिस्टम कंप्यूटर के सॉफ्टवेयर यानी एप्लीकेशन प्रोग्राम और हार्डवेयर यानी सीपीयू के बीच बेहतर सामंजस्य स्थापित करने में सफल हो पाता है।

कर्नल (Kernel)

कर्नल किसी ऑपरेटिंग सिस्टम का सबसे अहम और केंद्रीय (Central) भाग है। यानी ऑपरेटिंग सिस्टम की जो भी गतिविधियां होती हैं, वे सब कर्नल के ही इर्द -गिर्द होती हैं या यूं भी कह सकते हैं कि कर्नल की वजह से ही ऑपरेटिंग सिस्टम ठीक से काम कर पाता है। हालांकि , इस सबके बावजूद कंप्यूटर उपयोगकर्ता कभी भी न तो कर्नल को देख पाता है , न ही इसे महसूस कर सकता है , क्योंकि यह नेपथ्य (Behind The Scene) रहकर काम करता है। किसी ऑपरेटिंग सिस्टम में कर्नल किस तरह काम करता है , यह हम निम्नवत ग्राफ से समझ सकते हैं-



सामान्य शब्दों में यह भी कहा जा सकता है कि कर्नल ही किसी ऑपरेटिंग सिस्टम की बुनियाद है। इसकी मदद से ही ऑपरेटिंग सिस्टम यह तय कर पाता है कि किसी एप्लीकेशन सॉफ्टवेयरके लिए कब.किस वक्त पर कौन सा हार्डवेयर समुचित परिणाम प्राप्त करने के लिए इस्तेमाल करने की जरूरत होगी। हम जब भी कंप्यूटर ऑन करते हैं, वह कर्नल ही है जो सिस्टम को रिबूट (Reboot) करता है, मेमोरी को चेक करता है, किसी प्रोग्राम के लिए मेमोरी लोकेट (Locate) करना या नहीं करना यह सब कर्नल की मदद से ही संभव हो पाता है। अब कर्नल (ऑपरेटिंग सिस्टम) किसी एप्लीकेशन सॉफ्टवेयर की मदद में किस तरह पूरी प्रक्रिया करता है , इसे हम निम्न बिंदुओं में समझ सकेंगे-

- **प्रोग्राम एक्जीक्यूशन (Program Execution)-** एक्जीक्यूशन का हिन्दी अर्थ होता है निष्पादन या प्रक्रिया यानी किसी काम को संपन्न करना या करने का तरीका। अब हम यह भली -भांति जानते हैं कि ऑपरेटिंग सिस्टम का मुख्य काम एप्लीकेशन प्रोग्राम और कंप्यूटर हार्डवेयर के बीच सामंजस्य बनाना

है। इसमें कर्नल सबसे अधिक मददगार साबित होता है। दरअसल , कर्नल ही ऑपरेटिंग सिस्टम का वह हिस्सा है जो प्रोग्राम के लिए मेमोरी स्पेस तय करता है, इसके लिए जरूरी हार्डवेयर उपलब्ध करवाता है, मल्टीटास्किंग सिस्टम में एक से अधिक एप्लीकेशन के लिए टाइम शेयरिग (Time Sharing) करता है, उपयोगकर्ता की ओर से मिलने वाले निर्देशों को बाइनरी कोड में तब्दील कर हार्डवेयर , सीपीयू तक पहुंचाता है और फिर परिणाम हासिल कर उन्हें दोबारा हाई लेवल लैंग्वेज में बदलकर उपयोगकर्ता के समझने लायक बनाता है।

- **व्यवधान (Interrupt)-** प्रोग्राम एक्जीक्यूशन के दौरान कई बार हार्डवेयर और एप्लीकेशन प्रोग्राम के बीच बाधाएं या व्यवधान उत्पन्न होते हैं। दरअसल, ये व्यवधान सिग्नल (Signal) के रूप में होते हैं, जो हार्डवेयर से ऑपरेटिंग सिस्टम को या एप्लीकेशन प्रोग्राम से ऑपरेटिंग सिस्टम को मिलते हैं। ये सिग्नल असल में तब आते हैं, जब किसी प्रोग्राम को चलाने के लिए किसी खास हार्डवेयर की जरूरत होती है या एक्जीक्यूशन के दौरान हार्डवेयर प्रोग्राम एप्लीकेशन के किसी खास हिस्से को और बेहतर समझना चाहता है। ऐसी स्थिति में यह कर्नल की जिम्मेदारी है कि वह तुरंत प्रक्रिया जहां तक पहुंची है , वहीं रोक दे, लेकिन जितनी प्रोसेसिंग हो चुकी है, उसे सुरक्षित भी रखे। इसके बाद कर्नल हार्डवेयर के लिए जरूरी प्रोग्राम या प्रोग्राम के लिए जरूरी हार्डवेयर को तलाशकर प्रक्रिया को दोबारा वहीं से शुरू करवाता है, जहां वह रुकी थी।

मोड (Modes)- आधुनिक सीपीयू (Central Processing Unit) कई मोड पर काम करती हैं। इनमें यूजर मोड (User Mode) और सुपरवाइजर मोड (Supervisor Mode) प्रमुख हैं। सुपरवाइजर मोड में कर्नल खुद ही सभी प्रोग्राम के एक्जीक्यूशन के लिए जरूरी निर्णय लेता है और हार्डवेयर को निर्देश प्रदान करता है। दूसरी ओर कुछ प्रोग्राम एप्लीकेशन ऐसे होते हैं , जो कर्नल की मदद के बिना खुद ही सीधे ऑपरेटिंग सिस्टम की लाइब्रेरी और अन्य संसाधनों का उपयोग करते हैं। अब ऐसे किसी प्रोग्राम के संचालन की स्थिति में कंप्यूटर सिस्टम भ्रमित न हो और वह क्रैश (Crash) न हो जाए, यह सुनिश्चित करता है कर्नल। कर्नल यूजर मोड और सुपरवाइजर मोड के बीच एक लक्ष्मणरेखा सी खींच देता है , जिससे किसी एप्लीकेशन के स्वतंत्र रूप से काम करने के दौरान कर्नल के स्तर पर कोई बाधा उत्पन्न न हो।

- **मेमोरी प्रबंधन (Memory Management)-** हम जानते हैं कि मल्टीटास्किंग सिस्टम का तात्पर्य एक ऐसे सिस्टम है, जिस पर एक ही समय में एकसाथ एक से अधिक प्रोग्राम संचालित किए जा सकें। अब यदि इसे मानवीय उदाहरण के जरिये समझने की कोशिश करें तो हम जानेंगे कि जब कभी हम एक ही समय में एक साथ दो या दो से अधिक काम करने लगते हैं तो आशंका इस बात की अधिक रहती है कि हमारा कोई काम या तो अधूरा रह जाएगा या पूरा ध्यान नहीं दे पाने के कारण प्रारंभ ही नहीं होगा। कंप्यूटर सिस्टम में ऐसी स्थिति से ऑपरेटिंग सिस्टम को बचाता है कर्नल। कर्नल एक से अधिक प्रोग्राम चलने पर यह सुनिश्चित करता है कि सिस्टम की पूरी मेमोरी का सही उपयोग हो। इसके लिए वह हर प्रोग्राम को जरूरत के हिसाब से मेमोरी उपलब्ध कराता है। यही नहीं , यह भी तय करता है कि जिस वक्त किसी खास मेमोरी लोकेशन पर कोई एक प्रोग्राम एक्जीक्यूट (Execute) हो रहा है, उसी मेमोरी लोकेशन पर दूसरा प्रोग्राम न जा सके।
- **मल्टीटास्किंग (Multitasking)-** हम जब भी किसी कंप्यूटर पर एक साथ एक से अधिक प्रोग्राम चलाते हैं, तो हमें भले ही यह अनुभूति (Experience) होती है कि दोनों प्रोग्राम एकसाथ एक ही समय पर चल रहे हैं , लेकिन दरअसल दोनों अलग -अलग समय पर चलते हैं। होता यह है कि यह सब एक्जीक्यूशन इतनी तेजी से और इतने व्यवस्थित ढंग से होता है कि समय का यह अंतर बेहद नगण्य होता है। प्रोग्राम संचालन की यह समय व्यवस्था भी कर्नल की बदौलत संभव हो पाती है। कर्नल ही तय करता है कि किस प्रोग्राम के एक्जीक्यूशन के लिए कितना समय लगने वाला है। वह एक से अधिक प्रोग्राम के संचालन के लिए समय निर्धारण (Time Scheduling) करता है , जिससे मल्टीटास्किंग

संभव होती है। एक प्रोग्राम पहले से चल रहे दूसरे प्रोग्राम के संपन्न होने तक कतार (Queue) में रहता है।

- **डिस्क एक्सेस-फाइल सिस्टम (Disk Access-File System)-** हम जब भी कंप्यूटर खोलते हैं तो उसमें सी.डी.ई.एफ. सीडी के आइकन नजर आते हैं। इनमें से सी .डी.ई.एफ... आदि उस हार्डडिस्क ड्राइव के भाग हैं, जो कंप्यूटर की मेमोरी और सीपीयू का हिस्सा है। सीडी या डीवीडी ड्राइव कंप्यूटर की एक्सटर्नल मेमोरी का हिस्सा है। अब ऑपरेटिंग सिस्टम का यह काम है कि किसी डाटा को फाइल की शक्ल में इन मेमोरी में सुरक्षित रखे।

इस पूरी प्रक्रिया को फाइल सिस्टम (File System) कहा जाता है, जिसमें उपयोगकर्ता को यह सहूलियत मिलती है कि वह किसी डाटा, सूचना, परिणाम को फाइल की शक्ल में सुरक्षित रखे और इसे नाम या चिह्नों की मदद से पहचान दे। अब जब भी कोई प्रोग्राम कंप्यूटर पर चलता है तो कर्नल तय करता है कि प्रोग्राम के लिए कौन सा डाटा उपयुक्त है और यह फाइल सिस्टम में कहां उपलब्ध है। इसके बाद हार्डवेयर और एप्लीकेशन आसानी से संबंधित जानकारी तक पहुंच सकते हैं।

- **डिवाइस ड्राइवर्स (Device Drivers)-** डिवाइस ड्राइवर ऑपरेटिंग सिस्टम का अहम हिस्सा हैं। ये भी दरअसल कुछ खास तरह के प्रोग्राम हैं, जो किसी खास एप्लीकेशन की मदद में हार्डवेयर के लिए तैयार किए जाते हैं। किसी प्रोग्राम को चलाने के दौरान कौन सा हार्डवेयर किस तरह काम करेगा , यह इन ड्राइवर के जरिये तय किया जाता है। यही वजह है कि अकसर एप्लीकेशन प्रोग्राम के लिए अलग डिवाइस ड्राइवर कंप्यूटर में इंस्टॉल करने की जरूरत पड़ती है। कंप्यूटर ऑपरेटिंग सिस्टम के पुराने वर्जनों में अकसर यह होता था कि डिवाइस ड्राइवर किसी प्रोग्राम के चलने पर खुद ही एक्जीक्यूशन शुरू कर देते थे। लेकिन विंडोज के विस्टा वर्जन की लांचिंग के बाद से ऑपरेटिंग सिस्टम में बदलाव किया गया है। इसके तहत अब डिवाइस ड्राइवर प्रोग्राम के चलने पर कर्नल की मदद लेते हैं। कर्नल एक बार एक्जीक्यूशन शुरू हो जाने के बाद खुद को प्रक्रिया से अलग कर लेता है और प्रक्रिया पूरी हो जाने या प्रक्रिया के बीच कोई अगला निर्देश नहीं मिलने तक डिवाइस ड्राइवर को अपने स्तर पर ही काम करने की स्वतंत्रता प्रदान करता है।

नेटवर्किंग (Networking)

हम जानते हैं कि विकास के अनुक्रम में कंप्यूटर मल्टीयूजर , मल्टीटास्किंग मशीन बन चुका है। इसी का एक स्वरूप है नेटवर्किंग। नेटवर्किंग का तात्पर्य उस व्यवस्था से है जो एक से अधिक कंप्यूटरों को एक.दूसरे के बीच डाटा एक्सचेंज की सुविधा प्रदान कर सके। ये कंप्यूटर या तो तारों के जाल के जरिये एक.दूसरे से जुड़े हो सकते हैं या फिर वायरलेस (Wireless) नेटवर्क की मदद सेए जिसे नेटवर्क नोड (Network Nodes) कहा जाता है। इस प्रक्रिया में पर्सनल कंप्यूटर , सर्वर, फोन आदि कुछ भी जोड़ा जा सकता है। आज के दौर में नेटवर्क का सर्वाधिक प्रचलित स्वरूप है इंटरनेट, जिसके बारे में हम आगे विस्तार से जानेंगे।

- **सुरक्षा (Security)**

कंप्यूटर की बढ़ती जरूरतों और दैनिक मानव जीवन में उपयोग की वजह से आधुनिक दौर के ऑपरेटिंग सिस्टम ऐसे असंख्य संसाधनों को कंप्यूटर पर मौजूद एप्लीकेशन को चलाने की आजादी प्रदान करते हैं। लेकिन इस पूरी प्रक्रिया के बीच ऑपरेटिंग सिस्टम यह भी तय करते हैं कि प्रोग्राम संचालन के लिए उन्हें नेटवर्क के जरिये जो भी निर्देश मिलते हैं , वे अनुमतियोग्य हैं भी या नहीं। यही वजह है कि आज के दौर के अधिकतर ऑपरेटिंग सिस्टम ऐसे सुरक्षा फीचर्स से लेस हैं, जो कंप्यूटर सिस्टम पर किसी भी तरह के अनधिकृत गतिविधियों को रोक सकती हैं। कंप्यूटर पर यूजर एकाउंट, मोबाइल फोन पर फिंगरप्रिंट एक्सेस, ईमेल पर ईमेल-पासवर्ड आदि ऐसे ही फीचर्स हैं, जिनसे गुजरने के बाद ही कोई उपयोगकर्ता कंप्यूटर एप्लीकेशन और हार्डवेयर का इस्तेमाल कर पाता है।

विंडोज ऑपरेटिंग सिस्टम की बात करें तो इनमें एन्टी फिशिंग फिल्टर (Anti Phishing Filter) इंटरनेट एक्सप्लोरर में पहले से मौजूद होता है। फिशिंग वह प्रक्रिया है , जिसके जरिये उपयोगकर्ता की व्यक्तिगत जानकारी, जैसे- डेबिट कार्ड का पिनकोड , ई-मेल के पासवर्ड आदि , निकालने का प्रयास किया जाता है। एन्टी फिशिंग फिल्टर की मदद से इंटरनेट एक्सप्लोरर इस तरह की गतिविधियों को पहचान कर उन्हें नुकसान पहुंचाने से पहले ही रोक देता है। इसके अलावा विंडोज सिस्टम फायरवॉल (Firewall) से सुसज्जित होता है , जिसकी मदद से वायरस से बचा जा सकता है। हालांकि , फायरवॉल कंप्यूटर में पहले से स्थापित और इंटरनेट तक सूचनाएं पहुंचाने में सक्षम प्रोग्रामों को नियंत्रित नहीं करता है। ऐसे में अधिकतर उपयोगकर्ता बाह्य फायरवॉल को कंप्यूटर इंस्टॉल करते हैं। विंडोज डिफेंडर भी विंडोज ऑपरेटिंग सिस्टम का हिस्सा है। यह स्वयं काम करता है और किसी भी तरह की अनधिकृत प्रक्रिया की सूचना उपयोगकर्ता तक पहुंचा देता है। इसके अलावा एक्शन सेटिंग (Action Settings) के जरिये उपयोगकर्ता को यह सुविधा मिलती है कि वह ऑपरेटिंग सिस्टम में प्रदत्त सुरक्षा व्यवस्था को अपनी सहूलियत के अनुरूप शुरू या बंद कर सके।

- **यूजर इंटरफेस (User Interface)**

हम जानते हैं कि कंप्यूटर हमारी यानी मानवों की भाषा नहीं समझ सकता , न ही हम कंप्यूटर की बाइनरी भाषा को समझ सकते हैं। ऐसे में यह जरूरी हो जाता है कि कंप्यूटर और इसे उपयोग करने वाले के बीच कुछ ऐसा अंतराफलक (Interface) हो , जो एक -दूसरे को समझ नहीं पाने के बावजूद दोनों के बीच बेहतर समझदारी विकसित कर सके। यही प्रक्रिया यूजर इंटरफेस कहलाती है और कंप्यूटर.मानव संबंध में यही ऑपरेटिंग सिस्टम की बड़ी जिम्मेदारी है।

कंप्यूटर के विकास के क्रम में इसकी शुरुआत चिह्नों , संकेतकों, अक्षरों से हुई थी और मौजूदा दौर के अधिकतर ऑपरेटिंग सिस्टम ग्राफिकल यूजर इंटरफेस (Graphical User Interface) का इस्तेमाल करते हैं , जिसमें आइकन के जरिये प्रोग्राम एप्लीकेशन और उपयोगकर्ता के बीच बेहतर संबंध बन पाता है। यूजर मॉनीटर पर नजर आने वाले आइकन के जरिये किसी फाइल, प्रोग्राम या डाटा को आसानी से पहचान सकता है और उस पर क्लिक कर अभीष्ट परिणाम हासिल करता है।

14.6 इंटरनेट (Internet)

कंप्यूटर के मानव जीवन का अभिन्न अंग बन जाने के बाद आज के दौर में शायद ही ऐसा कोई व्यक्ति होगा जो इंटरनेट से परिचित न हो या जिसने इस सुविधा का कभी इस्तेमाल नहीं किया हो। हम जानते हैं कि नेटवर्किंग वह व्यवस्था है , जिसमें एक से अधिक कई कंप्यूटरों को डाटा एक्सचेंज के लिए आपस में जोड़ा जा सकता है, लेकिन इस तरह की नेटवर्किंग की सीमाएं तय हैं। इस तरह का नेटवर्क किसी एक संस्थान में , ऑफिस में, शिक्षण संस्थान में संभव है , जहां सभी कंप्यूटर एक-दूसरे से जुड़े हुए हों। अब इंटरनेट शब्द भी नेटवर्किंग से ही जुड़ा हुआ है , लेकिन इसका तात्पर्य किसी निश्चित या सीमित दायरे में कंप्यूटरों का एक -दूसरे से जुड़ना ही नहीं है। बल्कि यह नेटवर्कों का एक ऐसा नेटवर्क (Network of Networks) है, जो असीमित है। इसमें आम आदमी से लेकर निजी संस्थाओं, शैक्षणिक संस्थानों, कंपनियों, व्यापार, स्वास्थ्य, खेल-मनोरंजन समेत जीवन के हर आयाम की जानकारियों का स्थानीय से लेकर वैश्विक (Global) पहुंच का जाल है जो इंटरनेट प्रोटोकॉल (Internet Protocol), वर्ल्ड वाइड वेब (World Wide Web) इलेक्ट्रॉनिक मेल (E-Mail) टेलीफोन के जरिये दुनियाभर के कंप्यूटरों से जुड़ा हुआ है। कंप्यूटरों के बीच यह जुड़ाव वायरलेस , इलेक्ट्रिक और ऑप्टिकल तकनीक के माध्यम से संपन्न होता है।

- **इंटरनेट का संक्षिप्त इतिहास (Internet's Brief History)**

इंटरनेट की शुरुआत कबए कैसे हुई यह हम निम्न बिन्दुओं से समझेंगे-

- वर्ष 1969 में अमेरिकी रक्षा विभाग ने एडवांस रिसर्च प्रोजेक्ट एजेंसी (Advanced Research Project Agency- ARPA) नाम से एक नेटवर्क लांच किया , यह नेटवर्क युद्धकाल में गोपनीय सूचनाओं के त्वरित आदान-प्रदान के उद्देश्य से तैयार किया गया था।
- एआरपीए की कामयाबी के बाद इसे रक्षा मामलों से इतर सामान्य जनजीवन के लिए उपयोग करने लायक बनाने का प्रोजेक्ट प्रारंभ किया गया। तब इसे नाम दिया गया एडवांस रिसर्च प्रोजेक्ट एजेंसी नेटवर्क (ARPANET) अमेरिकी वैज्ञानिक लियोनॉर्ड क्लिनरॉक और पॉल बैरन तथा ब्रिटिश वैज्ञानिक डोनाल्ड डेविस और लॉरेंस रॉबट्स ने इस सिस्टम का कांसेप्ट डिजाइन किया था।
- ARPANET में कार्यरत अमेरिकी वैज्ञानिक रेमंड सैमुअल टॉम्लिनसन या रे टॉम्लिनसन ने नेटवर्क के लिए पहला फाइल ट्रांसफर प्रोग्राम (FTP) सीपीवाईनेट (CPYNET) तैयार किया। इसके जरिये ARPANET से जुड़े कंप्यूटरों पर सूचनाओं का आदान -प्रदान संभव हो सका। टॉम्लिनसन ने ही सबसे पहले 1972 में ई -मेल की शुरुआत की। हालांकि , प्रारंभ में इस तरह की ई -मेल उसी उपयोगकर्ता को भेजी जा सकती थी , जो उसी कंप्यूटर को प्रयोग करता हो , जिससे ई -मेल भेजी गई है। यानी ई -मेल भेजने के बाद उसे खोलने के लिए उसी कंप्यूटर काम करना जरूरी था। इस दिक्कत से निजात के लिए टॉम्लिनसन ने / की ईजाद की। इसके बाद ई-मेल को एक से दूसरे कंप्यूटर और बाद में एक से दूसरे देश तक भेजना सरल हो गया।
- 1979 में ब्रिटिश डाकघर इंटरनेट तकनीक का इस्तेमाल करने वाला पहला संस्थान बना।
- 1984 तक 1000 से अधिक कंप्यूटर इंटरनेट तकनीक से जोड़े जा चुके थे , धीरे-धीरे यह तकनीक तेजी से बढ़ने लगी और लोग इससे जुड़ने लगे।
- 1985 में अमेरिका ने नेशनल साइंस फाउंडेशन नेटवर्क (NSFNET) प्रोजेक्ट शुरु किया। इसके बाद इंटरनेट तकनीक का तेजी से विकास हुआ और यह दुनियाभर में फैलती चली गई।
- हमारे देश भारत में वर्ष 1980 में इंटरनेट की शुरुआत हुई , जब एजुकेशन एंड रिसर्च नेटवर्क (ERNET) प्रोजेक्ट प्रारंभ हुआ। इस प्रोजेक्ट को भारत सरकार और संयुक्त राष्ट्र के विकास कार्यक्रम (UNDP) की मदद से प्रारंभ किया गया।
- 15 अगस्त 1995 को विदेश संचार निगम लिमिटेड (VSNL) ने गेटवे सिस्टम शुरू कर इंटरनेट सुविधा भारत में आम उपयोग के लिए उपलब्ध कराई। इसके बाद से देश में इंटरनेट सुविधा लगातार बढ़ती गई। आज भारत संचार निगम लिमिटेड समेत कई मोबाइल कंपनियां , ब्रॉडबैंड कंपनियां इंटरनेट सुविधा दे रही हैं, जिनसे 13 करोड़ से अधिक लोग जुड़ चुके हैं। उल्लेखनीय पहलू यह है कि दुनिभर के देशों में इंटरनेट इस्तेमाल करने वाले लोगों की संख्या के मामले में भारत का हिस्सा 13.5 प्रतिशत है। आम आदमी तक इंटरनेट की पहुंच के हिसाब से अमेरिका दुनिया का सबसे बड़ा देश है। वहां की कुल आबादी 31 करोड़ से कुछ अधिक है, जबकि इंटरनेट सुविधा से 24 करोड़ से अधिक लोग जुड़े हुए हैं।

इंटरनेट कनेक्शन के प्रकार (Types of Internet Connection)

इंटरनेट की मदद से हम घर बैठे अपने कंप्यूटर पर दुनियाभर की सूचनाएं पलक झपकते ही हासिल कर सकते हैं। लेकिन कंप्यूटर पर इंटरनेट सुविधा प्राप्त करने के लिए हमें इंटरनेट कनेक्शन की आवश्यकता होती है। आधुनिक दौर में डेस्कटॉप से लेकर लैपटॉप , गेमिंग कन्सोल, टैबलेट्स, मोबाइल फोन तक में इंटरनेट कनेक्शन का इस्तेमाल किया जाता है। यह उपयोगकर्ता पर निर्भर करता है कि वह किस तरह के इंटरनेट कनेक्शन से जुड़ना चाहता है। कुछ प्रमुख कनेक्शन निम्नवत हैं-

- **डायलअप कनेक्शन (Dial Up Connection)-** इस प्रक्रिया में उपभोक्ता का कंप्यूटर फोन लाइन के जरिये जोड़ा जाता है। इस तरह के कनेक्शन को एनालॉग (Analog) कनेक्शन कहा जाता है। इस कनेक्शन के जोड़ने के बाद फोन का इस्तेमाल करना संभव नहीं होता। हालांकि , गति धीमी होने के कारण अब इस कनेक्शन का प्रचलन लगभग खत्म हो चुका है।

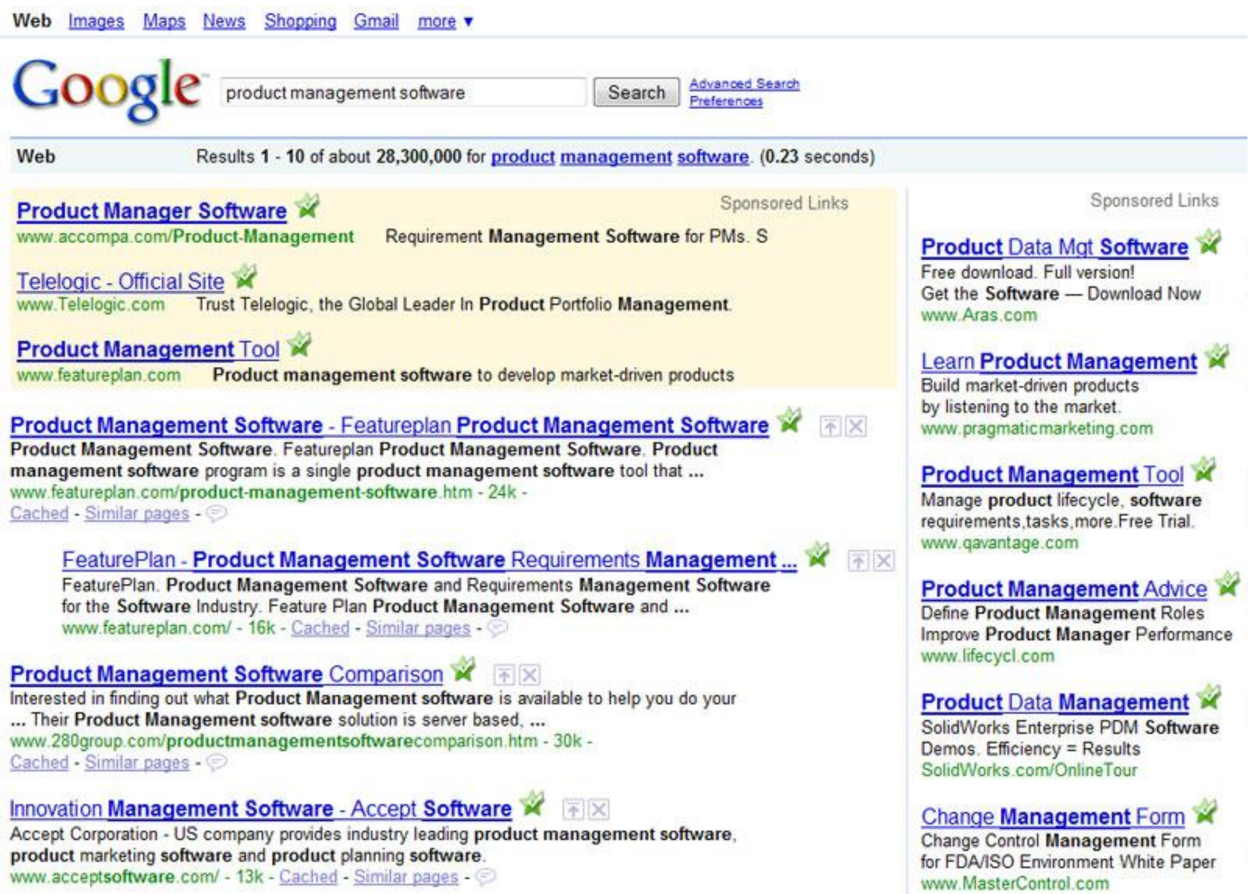
- **ब्रॉडबैंड कनेक्शन (Broadband Connection)-** ब्रॉडबैंड कनेक्शन सबसे ज्यादा तीव्र गति वाला इंटरनेट कनेक्शन है। इसमें भारी मात्रा में सूचनाएं भेजने के लिए एक से अधिक डाटा चैनलों का इस्तेमाल किया जाता है। ब्रॉडबैंड ब्रॉड बैंडविथ (Broad Bandwidth) का संक्षिप्त रूप है। केबल और टेलीफोन कंपनियां ब्रॉडबैंड सेवाएं उपलब्ध कराती हैं।
- **डीएसएल कनेक्शन (DSL Connection)-** डीएसएल कनेक्शन की फुलफॉर्म है, डिजिटल सब्सक्राइबर लाइन (Digital Subscriber Line) इस कनेक्शन में उपभोक्ता के घर में उपलब्ध दो तारों वाली टेलीफोन लाइन का इस्तेमाल किया जाता है। इससे यह सुविधा लैंडलाइन कनेक्शन के साथ ही उपलब्ध हो जाती है। डायल अप कनेक्शन से इतर इस व्यवस्था में इंटरनेट के इस्तेमाल के दौरान उपभोक्ता लैंडलाइन फोन का भी प्रयोग कर सकता है।
- **वायरलेस कनेक्शन (Wireless Connection)-** जैसा कि नाम से ही स्पष्ट है कि इस तरह के कनेक्शन में तारों की मदद नहीं ली जाती है। इसमें केबल या टेलीफोन नेटवर्क के बजाय रेडियो तरंगों (Radio Frequency) का प्रयोग किया जाता है। इस कनेक्शन की सबसे बड़ी सुविधा यह है कि इसमें कनेक्शन हमेशा ऑन रहता है।
- **मोबाइल कनेक्शन (Mobile Connection)-** संचार क्रांति के दौर में अब इंटरनेट हर उपयोगकर्ता के हाथों तक आसान पहुंच बना चुका है। इसका जरिया बना है मोबाइल फोन। जीएसएम (GSM), 3-जी, 4-जी जैसी नयी तकनीकों की मदद से अब हम मोबाइल , टैबलेट पर आसानी से इंटरनेट सुविधा हासिल कर सकते हैं।

इंटरनेट के साधन (Tools of Internet)

इकाई के इस हिस्से हम उन साधनों को जानने का प्रयास करेंगे जो इंटरनेट सुविधा को सफल बनाने का काम करते हैं। इनमें किसी कंप्यूटर को इंटरनेट से जोड़ने वाले उपकरण भी शामिल हैं तो कंप्यूटर पर इंटरनेट और संचार सुविधा संचालित करने वाले एप्लीकेशन प्रोग्राम भी। आइए, इनका संक्षिप्त परिचय लेते हैं-

- **मोडेम (Modem) -** मोडेम का विस्तारित शब्द है मॉडुलेटर डि मॉडुलेटर (Modulator-De-Modulator) यह एक ऐसा उपकरण है जो कंप्यूटर में मौजूद डिजिटल डाटा को एनालॉग सिग्नलों (Analog Signal) में बदलता है। एनालॉग सिग्नल वे सिग्नल होते हैं , जो टेलीफोन लाइन या अन्य संचार माध्यम के जरिये एक से दूसरे स्थान तक भेजे जा सकते हैं। इसी तरह वह एनालॉग सिग्नल को डिजिटल डाटा में बदल देता है, ताकि कंप्यूटर सिग्नल को समझ सके।
- **वेब ब्राउजर (Web Browser)-** वेब ब्राउजर दरअसल एक तरह के सॉफ्टवेयर प्रोग्राम हैं, जो कंप्यूटर में ही स्थापित रहते हैं। इनकी मदद से उपभोक्ता इंटरनेट का इस्तेमाल सूचनाएं , डाटा की तलाश करने में कर पाता है। उदाहरण - इंटरनेट एक्स्प्लोरर, गूगल का गूगल क्रोम ब्राउजर , मोजिला फायर फॉक्स, एप्पल सफारी आदि।
- **वर्ल्ड वाइड वेब (WWW)-** हम जानते हैं कि वेब का अर्थ जाले से होता है। वर्ल्ड वाइड वेब का अर्थ सूचनाओं या डाटा के एक ऐसे जाल से है जो पूरी दुनिया में विस्तृत हो और कोई भी इंटरनेट उपयोगकर्ता इस डाटाबेस से अपनी जरूरत के मुताबिक सूचना हासिल कर सकता है। यह मूलतः डाटाबेस के अलग-अलग पेजों का एक समूह है जो शीर्षकों (Titles) में बंटे रहते हैं और जिन्हें वेबसाइट कहा जाता है।
- **वेबसाइट (Website)-** इंटरनेट पर कोई भी जानकारी डाटाबेस संबंधित पेजों के रूप में उपलब्ध रहती है, जिन्हें वेबसाइट कहा जाता है। ब्राउजर के जरिये उपयोगकर्ता इन वेबसाइट तक पहुंच सकता है। वेबसाइट जीवन के हर आयाम , पहलू पर आधारित होती हैं। खेल , मनोरंजन, विज्ञान अलग -अलग विषय की हजारों -लाखों वेबसाइट यानी पेज इंटरनेट पर उपलब्ध रहते हैं। शोधकार्यों के लिए ये वेबसाइट शोधार्थियों (Research Fellows) की खासी मददगार साबित होती हैं।

- वेब पेज और एचटीएमएल (Webpage and HTML)- एचटीएमएल एक उच्चस्तरीय प्रोग्रामिंग लैंग्वेज है, जो वेबपेज तैयार करने में काम आती है। वेबपेज क्या है, यह हम पहले ही जान चुके हैं। कोई भी वेबसाइट कई वेबपेजों का एक समूह हो सकता है। एचटीएमएल का विस्तृत शब्दरूप है हाइपर टेक्स्ट मार्कअप लैंग्वेज (Hypertext Markup Language)
- **एचटीटीपी (HTTP/http)-** एचटीटीपी का विस्तृत शब्दरूप है हाइपर टेक्स्ट ट्रांसफर प्रोटोकॉल (Hypertext Transfer Protocol) यह प्रोटोकॉल दरअसल वर्ल्ड वाइड वेब में मौजूद डाटाबेस की बुनियाद है, हम जब भी ब्राउजर पर किसी वेबसाइट को सर्च करने के लिए किसी वेबसाइट का नाम लिखते हैं तो उसके आगे http:// लिखा जाता है। इसका तात्पर्य यह है कि उपयोगकर्ता वेब पर वह फाइल तलाशना चाहता है, जो एचटीएमएल भाषा में उपलब्ध हो। एचटीटीपी को वर्ल्ड वाइड वेब की आचार संहिता भी माना जाता है।
- **डोमेन नेम (Domain Name)-** इंटरनेट पर एक ही विषय से जुड़ी हजारों -लाखों वेबसाइट उपलब्ध होती हैं, ऐसे में इनमें से उपयोगकर्ता के वास्तविक उपयोग वाली वेबसाइट तलाशना लंबा समय और उर्जा खाने वाला काम बन जाता है। ऐसे में हर वेबसाइट को जो नाम दिया जाता है वह डोमेन नेम कहलाता है। वास्तव में डोमेन नेम इंटरनेट पर किसी वेबसाइट का पता होता है। ब्राउजर पर जब भी उपयोगकर्ता किसी वेबसाइट का नाम लिखता है तो ब्राउजर तुरंत लाखों वेबपेज में से संबंधित वेबपेज को आसानी से तलाश लेता है।
- अब जब भी हम ब्राउजर पर किसी वेबसाइट को तलाश करते हैं तो उसे इस तरह पूरा लिखा जाता है - www.facebook.com इसमें शुरूआती तीन अक्षर www बताते हैं कि हम जिस पेज की तलाश कर रहे हैं, वह वर्ल्ड वाइड वेब पर उपलब्ध है , जबकि बाकी के दो शब्द यानी facebook.com इस वेबपेज का डोमेन नेम है। किन्हीं भी दो वेबसाइट का डोमेन नेम कभी भी एकसमान नहीं हो सकता है। यही वजह है कि ब्राउजर पर वेबसाइट का पूरा नाम लिखते ही अभीष्ट वेबपेज तुरंत खुल जाता है।
- **यूआरएल (URL)-** यूआरएल यानी यूनिफॉर्म रिसोर्स लोकेटर (Uniform Resource Locator) किसी वेबसाइट का पूरा नाम यानी वर्ल्ड वाइड वेब पर उस वेबसाइट या वेबपेज का पूरा पता है। इसे हम इस उदाहरण से समझ सकते हैं। यदि हमें अपने विश्वविद्यालय यानी उत्तराखंड मुक्त विवि (Uttarakhand Open University) की वेबसाइट खोलनी है तो हम वेब ब्राउजर पर इस तरह लिखते हैं- http://www.uou.ac.in अब हम जानते हैं कि इस नाम के आखिरी तीन शब्द डोमेन नेम , पहले शब्द और चिह्न हाइपर टेक्स्ट प्रोटोकॉल और www वर्ल्ड वाइड वेब के परिचायक हैं। इन सभी से मिलकर वेबसाइट का जो पूरा पता बना है, वह यूआरएल कहलाता है।
- **सर्च इंजन (Search Enginers)-** कई बार होता यह है कि उपयोगकर्ता को उस विषय की तो जानकारी रहती है, जिसके लिए उसे डाटा या सूचनाओं की आवश्यकता है, लेकिन उसे यह मालूम नहीं होता कि कौन सी वेबसाइट उसके लिए उपयोगी होगी। कई बार उसे अभीष्ट वेबसाइट का नाम भी मालूम नहीं होता है। ऐसे में सर्च इंजन इंटरनेट उपयोगकर्ता के खासे मददगार साबित होते हैं।



(गूगल सर्च इंजन पर इस तरह की वर्ड की मदद से साइट ढूंढी जाती हैं)

- दरअसल, सर्च इंजन पर उपयोगकर्ता को वेबसाइट का पूरा नाम लिखने के बजाय सिर्फ कुछ कीवर्ड (Keywords) ही लिखने की जरूरत होती है। उदाहरण के लिए अगर उपयोगकर्ता समाज में बढ़ते अपराधों के विषय पर डाटा -सूचनाएं और जानकारी जुटाना चाहता है , लेकिन उसे नहीं मालूम है कि वह किस वेबसाइट पर जाए तो वह ब्राउजर पर क्राइम (Crime) या समाज (Society) या समाज में अपराध पद (Crime in Society) जैसे शब्द ही लिख सकता है। सर्च इंजन तुरंत इन शब्दों के आधार पर एक साथ कई वेबपेज की सूची उपलब्ध करा देता है , जिन पर क्लिक कर उपयोगकर्ता अभीष्ट जानकारी हासिल कर पाता है। गूगल, याहू, बिंग आदि ऐसे ही सर्च इंजन हैं।
- **ईमेल (E-mail)-** ई -मेल, जैसा कि नाम से ही स्पष्ट है कि यह एक ऐसी चिट्ठी या संदेश है जिसका स्वरूप इलेक्ट्रॉनिक यानी डिजिटल है। वास्तव में ई -मेल भी एक तरह का सॉफ्टवेयर है, जो उपयोगकर्ता को कोई संदेश दूसरे उपयोगकर्ता तक पहुंचाने की सुविधा देता है। ई -मेल दो प्रकार की होती हैं। पहली ब्राउजर आधारित , इस तरह की मेल में उपयोगकर्ता को इंटरनेट पर मौजूद ई -मेल सुविधा देने वाली कंपनी से जुड़ना होता है। इसके लिए उपयोगकर्ता को संबंधित कंपनी में अपना विशेष खाता बनाना होता है। जीमेल , याहूमेल, रेडिफमेल, हॉटमेल ऐसी ही कंपनियां हैं जो ई -मेल सुविधा देती हैं। यह प्रक्रिया निःशुल्क होती है। इनसे जुड़ा उपयोगकर्ता इस तरह अपना ई -मेल खाता या ई-मेल आईडी बनाता है- xyz@gmail.com दूसरी ई-मेल होती हैं उपभोक्ता आधारित। इस तरह की मेल कंप्यूटर में इंस्टॉल सॉफ्टवेयरपर ही उपलब्ध होती हैं। मसलन माइक्रोसॉफ्ट ऑफिस सॉफ्टवेयर पर आउटलुकए आउटलुक एक्सप्रेस आदि।

- **इंटरनेट का सामाजिक प्रभाव (Socio Impact of Internet)**

किसी भी सुविधा के दो पहलू होते हैं। हर काम, हर संसाधन के साथ सकारात्मक और नकारात्मक परिणाम जुड़े होते हैं, यह हम पर निर्भर करता है कि हम किस पहलू को अधिक तवज्जो देते हैं। इंटरनेट सुविधा भी इस सार्वभौमिक सत्य का अपवाद नहीं है।

पहले चर्चा करते हैं इंटरनेट के सकारात्मक पहलू की। इंटरनेट मानव समुदाय को संचार क्रांति का सबसे बड़ा उपहार है। आज आधुनिक दौर में यह एक ऐसा हथियार बन गया है जो दुनियाभर के मानव समुदाय में समभाव (Equalty) का जरिया है, चाहे वह जाति-धर्म के आधार पर हो या फिर अमीरी -गरीबी के आधार पर। इंटरनेट न तो छुआछूत देखता है न सामाजिक स्थिति के हिसाब से किसी व्यक्ति का आकलन करता है। उपयोगकर्ता के स्टेटस (Status) का ध्यान रखे बगैर यह हर उस व्यक्ति को दुनिया -जहान की हर जानकारी लैपटॉप, डेस्कटॉप या मोबाइल फोन पर उपलब्ध कराता है, जो इसका उपयोग कर पाने में सक्षम है या उपयोग करना चाहता है।

इंटरनेट आज न सिर्फ आम से आम आदमी तक दुनियाभर की जानकारियों के अकूत भंडार के तौर पर सहज-सुलभ है, बल्कि फेसबुक, ट्वीटर, इंस्टाग्राम और इन जैसी तमाम सोशल साइट्स के जरिये यह उस व्यक्ति को भी अपनी बात पूरी दुनिया के सामने रखने की छूट और आजादी प्रदान कर रहा है, जो कभी जाति तो कभी स्टेटस के भेद के कारण खुलकर कहने -सुनने में खुद को सक्षम नहीं पाता था। इस लिहाज से यदि यह कहें कि इंटरनेट सामाजिक, वैचारिक परिवर्तन का भी एक माध्यम है तो शायद यह अतिश्योक्ति नहीं होगी।

अब बात नकारात्मक पहलू की। इंटरनेट पर बीते कुछ वर्षों में सोशल साइट्स का प्रचलन बहुत तेजी से बढ़ा है। फेसबुक आज हर उस शख्स के जीवन का अभिन्न अंग बन गया है जो कंप्यूटर चलाता है तो व्हाट्सएप हर उस व्यक्ति की जरूरत जो स्मार्टफोन इस्तेमाल कर रहा है। लेकिन यहीं इंटरनेट पर सवाल खड़े होने लगते हैं। दरअसल, पिछले कुछ समय में जिस तेजी से संचार क्रांति बढ़ी है , उसके साथ ही यह चिंता भी बढ़ती चली गई है कि यह सुविधा मानव समुदाय की सामाजिक संरचना को नुकसान पहुंचाने की वजह बनती जा रही है। दरअसल, हुआ यह है कि आधुनिक दौर में लोग सदियों से चली आ रही सामाजिक संरचनाओं से विमुख होते जा रहे है। कई शोध रिपोर्ट ये बताती हैं कि आज का मनुष्य परिवार , समाज से कहीं अधिक वक्त लैपटॉप , टैबलेट या स्मार्टफोन पर ही बिता रहा है। परिजन , रिश्तेदार या समाज क्या कह , कर रहा है , इससे कहीं अधिक अहम उसके लिए सोशल साइट्स होती जा रही हैं। इसकी वजह से सामाजिक ढांचा सामूहिक से एकल की ओर बढ़ने लगा है। इंटरनेट सूचनाओं के आदान -प्रदान का सबसे तेज जरिया बन गया है। कोई घटना हो , दुर्घटना हो या सांस्कृतिक इवेंट हो, इसकी जानकारी सैकड़ों.हजारों मील दूर बैठे दूसरे शख्स तक चंद सेकंडों में पहुंच जाती है। शायद यही वजह है कि इंटरनेट और सोशल साइट्स ने राजनेताओं और राजनीतिक दलों का भी ध्यान तेजी से खींचा है। राजनीतिक परिदृश्य में भी अब यह माना जाने लगा है कि मतदाताओं तक कम समय में पैठ बनाने और अपना संदेश पहुंचाने के लिए सोशल साइट्स ही सबसे उपयुक्त माध्यम हैं। भारत समेत दुनिया के कई देशों में इंटरनेट और सोशल साइट्स चुनाव प्रचार का बड़ा हथियार बन गई हैं। दूसरी तरफ , टेलीमेडिसिन, ऑनलाइन एजुकेशन , रोजगार जैसी कई ऐसी सुविधाएं भी हैं , जिनके जरिये इंटरनेट ने आज मानव समुदाय के जीवन को और अधिक सरलीकृत किया है। शिक्षा , स्वास्थ्य की पहुंच इसके जरिये उन क्षेत्रों और लोगों तक भी बढ़ी है जो वर्षों तक इन बुनियादी सुविधाओं से वंचित रहे।

- **सुरक्षा (Security)**

इंटरनेट का जिस तेजी से इस्तेमाल बढ़ा है , उसी तेजी से इसके नकारात्मक पहलू भी लगातार सामने आए हैं। दरअसल, इंटरनेट पर जहां उपयोगकर्ता की मदद के लिए कई तरह के वेबपेज, प्रोग्राम उपलब्ध हैं, वहीं कई ऐसे प्रोग्राम और सॉफ्टवेयर भी इस पर मौजूद रहते हैं जो उपयोगकर्ता को नुकसान पहुंचा सकते हैं। यूनिट के इस हिस्से में हम ऐसे ही कुछ प्रोग्राम को जानेंगे जो हानिकारक हैं-

- **वायरस (Virus)-** वायरस एक प्रोग्राम या कंप्यूटर कोड होता है जो इंटरनेट पर कंप्यूटर के जुड़ते ही कंप्यूटर में प्रवेश कर जाता है। जिस तरह मानव शरीर में वायरस घुसता है तो संक्रमण फैलाता है। उसी तरह वायरस कंप्यूटर में घुसकर इसके सिस्टम को नुकसान पहुंचाता है। कंप्यूटर से महत्वपूर्ण फाइलें डिलीट करने के साथ यह हार्डडिस्क को भी करप्ट कर देता है। वायरस के कंप्यूटर पर आने की बड़ी वजह उपयोगकर्ता के संक्रमित फाइलें या ई-मेल अटैचमेंट खोलना होती हैं। इंटरनेट पर संदिग्ध वेबपेज खोलने पर भी अकसर वायरस आ जाते हैं।
- **वर्म (Worm)-** वर्म वह कंप्यूटर प्रोग्राम है , जो कंप्यूटर में प्रवेश करने के बाद ऑटोमेटिक तरीके से अपनी ही कई प्रतियां बना लेता है। इसके बाद यह कंप्यूटर की प्रक्रियाओं को बाधित कर देता है। वायरस से इतर यह खुद को कंप्यूटर की फाइलों या प्रोग्रामों से नहीं जोड़ता , बल्कि खुद ही प्रोग्राम बनकर प्रक्रिया रोकने लगता है। अगर कंप्यूटर नेटवर्किंग से जुड़े हुए हों तो यह संक्रमित कंप्यूटर से जुड़े दूसरे कंप्यूटरों में भी पहुंच जाता है। यह रैंडम एक्सेस मेमोरी को प्रभावित कर कंप्यूटर की प्रोसेसिंग को बेहद धीमा कर देता है।
- **ट्रोजन हॉर्स (Trojan Horse)-** ट्रोजन हॉर्स वे प्रोग्राम हैं जो उपयोगकर्ता के सामने लाभदायक प्रोग्राम के रूप में आते हैं, लेकिन प्रयोग की कोशिश करते ही ये कंप्यूटर में घुसकर उसमें वायरस डाल देते हैं। वायरस और वर्म की तरह ट्रोजन हॉर्स अपनी कई प्रतियां नहीं बनाते , बल्कि यह कंप्यूटर मेमोरी में मौजूद संवेदनशील डाटा, जानकारी, फाइलें और व्यक्तिगत जानकारियां तलाशते हैं। मूलतः आपराधिक किस्म के लोग इसका इस्तेमाल करते हैं , जिससे वे किसी व्यक्ति की गोपनीय जानकारी हासिल कर सकें। ऑनलाइन ठगी के अधिकतर मामलों को इस श्रेणी में रखा जा सकता है।

बचाव के तरीके (Prevention)

1.कंप्यूटर पर एंटी वायरस (Anti Virus)- प्रोग्राम स्थापित किया जाना चाहिए, एंटी वायरस प्रोग्राम इंटरनेट पर जुड़ने के दौरान हर प्रोग्राम , फाइल, वेबपेज को स्कैन करते हैं और इनमें किसी भी तरह का संदेह होने की स्थिति में उपयोगकर्ता को संबंधित फाइल या वेबपेज से नहीं जुड़ने का संदेश देते हैं, इसके अलावा ये अनधिकृत प्रोग्रामों को कंप्यूटर में प्रवेश करने से भी रोकते हैं

2. उपयोगकर्ता को कंप्यूटर पर फायरवॉल का इस्तेमाल करना चाहिए , यह एक खास तरह का प्रोग्राम है , जो विंडोज ऑपरेटिंग सिस्टम में पहले से उपलब्ध रहता है , उपयोगकर्ता को करना सिर्फ यह होता है कि सेटिंग में जाकर इसे ऑन करना होता है। इसके बाद यह किसी भी बाहरी साधन को उपयोगकर्ता के कंप्यूटर तक पहुंचने से रोकने का काम करता है।

3. इंटरनेट पर कोई भी संदिग्ध फाइल , वेबपेज और ई -मेल पर कोई भी ऐसा संदेश कभी नहीं खोलें , जिसे भेजने वाला संदिग्ध हो, ई-मेल पर आने वाले अटैचमेंट को खोलने से पहले स्कैन जरूर करें।

4. ई-मेल अटैचमेंट के फाइल एक्सटेंशन को ध्यान से जरूर देखें , यदि फाइल का एक्सटेंशन exe, pif, bat, bas, cmd, com, cml, inf, js, lnk, msi, scr, vbs हो तो इन्हें खोलने से पहले एंटी वायरस प्रोग्राम की मदद से स्कैन जरूर करें-

5. ई-मेल और सोशल साइट्स पर कई बार ऐसे लिंक आते हैं , जो उपयोगकर्ता को लालच देकर फांसते हैं। इस तरह के लिंक में कई बार उपयोगकर्ता की लॉटरी लगने की जानकारी दी जाती है तो कभी कोई दूसरा ऐसा संदेश भेजा जाता है, जिसे पढ़ते ही उपयोगकर्ता उस पर क्लिक करे, लेकिन इससे हमेशा बचना चाहिए।

6. इंटरनेट उपयोगकर्ताओं को वायरस से बचाने के लिए माइक्रोफ्ट पैच ट्यूजडे सेवा चलाता है। इसके जरिये माइक्रोसॉफ्ट हर महीने के दूसरे मंगलवार को उन सभी प्रोग्रामों की सूची तैयार करता है जो कंप्यूटर को नुकसान पहुंचा सकते हैं। इसमें इस तरह के प्रोग्रामों से बचने के तरीके भी सुझाए जाते हैं, जिन्हें पैच कहा जाता है। इन पैचों का प्रयोग कर उपयोगकर्ता इंटरनेट का सुरक्षित उपयोग कर सकते हैं।

7. सोशल साइट्स पर फोन नंबर, बैंक खाते, पासवर्ड, एटीएम पिन कोड जैसी गोपनीय जानकारी कभी भी दर्ज नहीं करें। फेसबुक ट्विटर जैसी सोशल साइटों पर उपयोगकर्ता सेटिंग के जरिये अपनी जानकारियों को छिपा भी सकते हैं या यह तय कर सकते हैं कि कौन लोग इन जानकारियों को देख सकते हैं।
8. ऑनलाइन नेटबैंकिंग के इस्तेमाल के दौरान हमेशा बैंक या कंपनी के अधिकृत वेबसाइट का ही इस्तेमाल करें। कई बार अधिकृत वेबसाइट के बजाय होस्ट वेबसाइट पर कंपनी का लिंक दिया जाता है , इस तरह की होस्ट साइट पर क्लिक करने से उपयोगकर्ता का गोपनीय डाटा चोरी होने की आशंका रहती है।

14.7 उपसंहार (Conclusion)

यूनिट के अध्ययन के बाद हम इस परिभाषा के तात्पर्य को समझने के साथ यह जान चुके हैं कि कंप्यूटर के संचालन में ऑपरेटिंग सिस्टम की कितनी बड़ी भूमिका है। इसके अलावा हम यह भी जाने हैं कि इंटरनेट किस तरह आज मानव के सामाजिक जीवन का अभिन्न अंग बन गया है और किस तरह इंटरनेट के सकारात्मक और नकारात्मक पहलू मानव के दैनिक जीवन पर असर डाल रहे हैं। हालांकि , इस सबके बीच यह जरूर माना जा सकता है कि ऑपरेटिंग सिस्टम के विकास और इसके जरिये इंटरनेट के अविर्भाव ने मानव जीवन को सरल जरूर बनाया है।

14.8 अभ्यास प्रश्न (Practice Question)

1. हाईलेवल लैंग्वेज में तैयार पहला ऑपरेटिंग सिस्टम था-
   a) बी-5000
   b) ओएस-360
   c) आईबीएम-709
   d) विंडोज 1.0
2. ऑपरेटिंग सिस्टम का बुनियादी घटक है-
   a) सेंट्रल प्रोसेसिंग यूनिट
   b) की-बोर्ड
   c) प्रोग्राम
   d) कर्नल
3. माइक्रोसॉफ्टका पहला सिर्फ पर्सनल कंप्यूटर आधारित ऑपरेटिंग सिस्टम था-
   a) विंडोज 2.0
   b) विंडोज 10
   c) विंडोज 7
   d) विंडोज विस्टा
4. वर्तमान दौर में सर्वाधिक प्रचलित ऑपरेटिंग सिस्टम है-
   a) एंड्रॉयड
   b) आईफोन
   c) ब्लैकबेरी
   d) विंडोज
5. इनमें से कौन सा ऑपरेटिंग सिस्टम टचस्क्रीन सपोर्ट करता है-
   a) मैक

b) विंडोज 8
c) विंडोज विस्टा
d) विंडोज 1.0

6. विंडोज से पहले आईबीएम में प्रयुक्त ऑपरेटिंग सिस्टम था-
a) डॉस
b) मैक
c) एंड्रॉयड
d) उपरोक्त में से कोई नहीं

7. स्टार्ट और क्लोज बटन सबसे पहले इस सिस्टम में लाए गए-
a) विंडोज 8
b) विंडोज 3.1
c) विंडोज 95
d) विंडोज 7

8. विंडोज का पहला ग्राफिकल यूजर इंटरफेस ऑपरेटिंग सिस्टम था-
a) विंडोज 1.0
b) विंडोज 8
c) विंडोज 3.1
d) इनमें से कोई नहीं

9. विंडोज 98 में पहली बार यह लांच किया गया-
a) पेंट ब्रश
b) ग्राफिकल यूजर इंटरफेस
c) एरो ग्लास लुक
d) इंटरनेट एक्सप्लोरर 4.0

10. मोबाइल फोन पर इस्तेमाल होने वाले एंड्रॉयड का मूल आधार यह ऑपरेटिंग सिस्टम है-
a) लाइनक्स
b) यूनिक्स
c) ब्लैकबेरी
d) विंडोज

11. पहली पीढ़ी के कंप्यूटरों में ऑपरेटिंग सिस्टम की तरह काम करते थे-
a) लाइनक्स
b) एंड्रॉयड
c) रेजीडेंट मॉनीटर
d) इनमें से कोई नहीं

12. इनमें से कौन ऑपरेटिंग सिस्टम ओपन लाइसेंस मोड है-
a) यूनिक्स फैमिली

b) मैक ओएस
c) विंडोज
d) लाइनक्स

13. मैक ओएस की निर्माता कंपनी है-
a) माइक्रोसॉफ्ट
b) डाटा कॉरपोरेशन
c) एप्पल
d) जेरॉक्स कॉरपोरेशन

14. गूगल और गूगल क्रोम का मूलाधार ऑपरेटिंग सिस्टम है-
a) विंडोज 1.0
b) लाइनक्स
c) यूनिक्स
d) विंडोज 7

15. पर्सनल डिजिटल असिस्टेंट (पीडीए) में इस्तेमाल ऑपरेटिंग सिस्टम है-
a) एंबेडेड
b) मल्टी यूजर
c) मल्टी टास्किंग
d) इनमें से कोई नही

16. इनमें से किसे इंटरनेट की आचार संहिता माना जाता है-
a) http://
b) html
c) www
d) इनमें से कोई नही

17. एनालॉग सिग्नल को डिजिटल में और डिजिटल को एनालॉग में बदलने वाला उपकरण है-
a) गूगल
b) वर्ल्ड वाइड वेब
c) मोडेम
d) उपरोक्त में से सभी

18. इनमें से कौन सर्च इंजन है-
a) गूगल
b) मोडेम
c) आउटलुक
d) इनमें से कोई नहीं

19. वेबपेज इस भाषा में तैयार किए जाते हैं-
a) http

b) html

c) ब़c++

d) java

20. वेब ब्राउजर इनमें से क्या है-

a) सॉफ्टवेयर प्रोग्राम

b) हार्डवेयर

c) प्रोग्रामिंग लैंग्वेज

d) उपरोक्त में से सभी

14.9 निबंधात्मक प्रश्न (Essay Type Question)

1. ऑपरेटिंग सिस्टम क्या है, कंप्यूटर के सफल संचालन में महत्व को समझाते हुए ऑपरेटिंग सिस्टम के विकास की यात्रा का वर्णन करें।

2. ऑपरेटिंग सिस्टम कितने प्रकार का होता है ? उदाहरण समेत विस्तार से बताएं , अलग -अलग तरह के ऑपरेटिंग सिस्टम की जरूरत क्यों महसूस हुई, इसकी जानकारी भी दें।

3. कुछ प्रमुख ऑपरेटिंग सिस्टम, इनकी विशेषताओं का उल्लेख करें।

4. विंडोज ऑपरेटिंग सिस्टम के विकास और हर वर्जन में आने वाले बदलाव की विशेषताए इसकी जरूरत आदि का विष्लेशण करें।

5. ऑपरेटिंग सिस्टम के प्रमुख घटक क्या हैं? इनके कार्यों और जरूरतों का उल्लेख करें।

6. किसी ऑपरेटिंग सिस्टम में कर्नल क्या होता है ? यह किस तरह एप्लीकेशन प्रोग्राम और हार्डवेयर के बीच संतुलन बनाता है?

7. नेटवर्किंग और इंटरनेट क्या हैं ? इंटरनेट पर काम जितना सुविधाजनक है , उतना ही असुरक्षित भी , इस कथन का विष्लेशण करते हुए जरूरी सावधानियों का भी जिक्र करें।

इकाई 15 माइक्रोसॉफ्ट ऑफिस, एमएस वर्ड और एमएस एक्सेल (Microsoft Office, MS Word and MS Excel)

15.1 परिचय / प्रस्तावना (Introduction)
15.2 उद्देश्य (Objectives)
15.2 कंप्यूटर के प्रमुख साधन (Important Tools of Computer)
(नोटपैड, वर्डपैड, पेंट, कैल्कुलेटर, विंडोज मीडिया प्लेयर)
15.3 एमएस ऑफिस (MS Office)
(एमएस ऑफिस क्या है , ऑपरेटिंग सिस्टम के विकास के हिसाब से ऑफिस का विकास , एमएस ऑफिस के घटक)
15.4 एमएस वर्ड (MS Word)
(एमएस वर्ड की जरूरत, कार्यशैली, कमांड, उपयोग)
15.5 एमएस एक्सेल (MS Excel)
(एमएस एक्सेल का उपयोग, कमांड)
15.6 उपसंहार (Conclusion)
15.7 अभ्यास प्रश्न (Practice Question)
15.8 निबंधात्मक प्रश्न (Essay Type Question)

15.1 परिचय / प्रस्तावना (Introduction)

हम इस तथ्य से भली -भांति परिचित हैं कि कंप्यूटर आज मानव समाज का अभिन्न अंग और दैनन्दिन उपयोग का सबसे बड़ा उपकरण बन गया है। हम यह भी जानते हैं कि कंप्यूटर का विकास जटिल गणनाओं (Calculations) के समाधान के तौर पर होता चला गया। लेकिन समय के साथ हुए विकास में कंप्यूटर महज गणनाओं तक सीमित नहीं रह गया है। यह जीवन के हर आयाम को छूता है और मानव के लिए सर्वाधिक उपयोगी मशीन बना है। टाइपिंग हो या ऑडियो विजुअल फीचर मानव जीवन का हर काम अब कंप्यूटर पर संभव है। इस यूनिट में कंप्यूटर के ऐसे ही फीचर या साधनों (Tools) का अध्ययन करेंगे, जिनकी मदद से हम अपने दैनिक उपयोग के कार्य आसानी से निष्पादित कर पाते हैं।

15.2 उद्देश्य (Objectives)

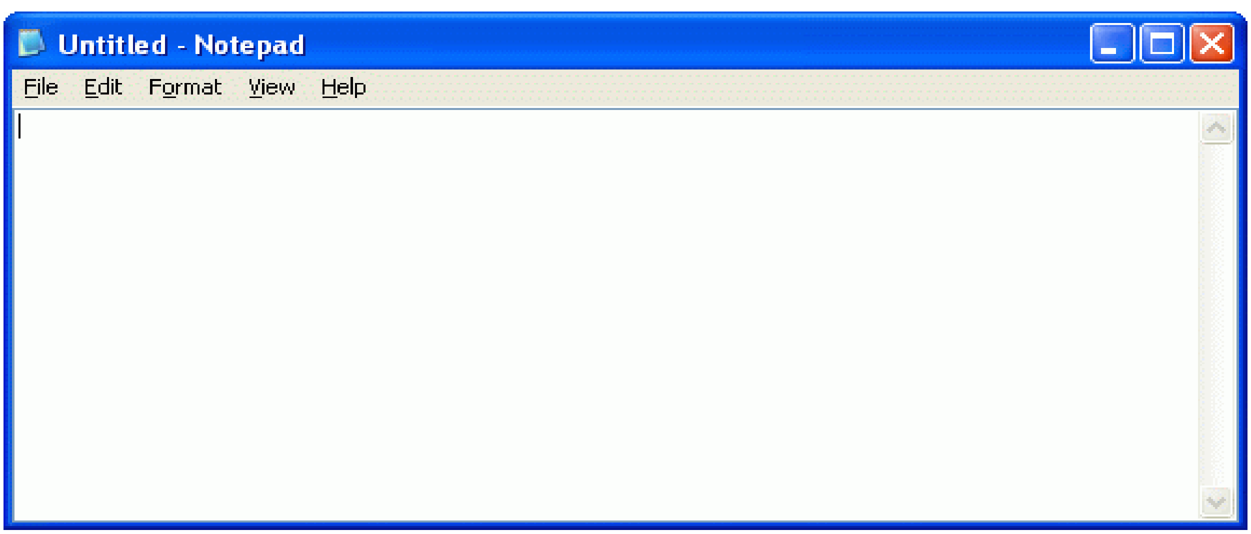
इस इकाई के अध्ययन के उपरान्त आप-

- ✓ कंप्यूटर के कुछ विशेष साधनों का उपयोग जानेंगे।
- ✓ माइक्रोसॉफ्ट ऑफिस एवं इसके उपयोगों को जानेंगे।
- ✓ एमएस-वर्ड, एक्सेल, पॉवरप्वाइंट एवं इनके उपयोगों को जानेंगे।

15.3 कंप्यूटर के प्रमुख साधन (Important Tools of Computer)

कंप्यूटर का विकास ही इस उद्देश्य के साथ हुआ कि जिन गणनाओं को हल करने में मानव को लंबा समय लगता थाए उनका समाधान चुटकियों में प्राप्त किया जा सके। कालांतर में गणनाओं से इतर कई अन्य कार्य भी इस श्रेणी में जुड़ते चले गए। इसी अनुक्रम में ऑपरेटिंग सिस्टम का विकास हुआ और सिस्टम में ही कुछ ऐसे उपयोगी प्रोग्राम जोड़े जाते गए, जो मानवोपयोगी थे। इनकी मदद से उपयोगकर्ता को ऐसे काम मिनटों में कर पाने की सहूलियत मिली , जिन्हें किसी अन्य साधन या विधि से करने में लंबा समय लगता। ऐसे ही कुछ साधनों के बारे में हम यहां जानने वाले हैं।

1. **नोटपैड (Notepad)-**



नोटपैड वह प्रोग्राम है , जिसका उपयोग हम अकसर करते हैं। विंडोज में यह प्रोग्राम प्री -इंस्टॉल (Pre Installed) रहता है। नोटपैड एक टेक्स्ट एडिटर (Text Editor) प्रोग्राम है। टेक्स्ट एडिटर प्रोग्राम का तात्पर्य उन प्रोग्राम से है, जिनमें उपयोगकर्ता अपनी जरूरत की टेक्स्ट फाइल (Text Files) तैयार कर सकता है। आम

जीवन में हम डायरी , कॉपी, कागज पर सूचनाएं दर्ज करते रहते हैं। कंप्यूटर पर यही काम नोटपैड पर किया जाता है। यह भी कहा जा सकता है कि नोटपैड वह प्रोग्राम है जो उपयोगकर्ता को सूचनाओं या डाटा के डॉक्यूमेंटेशन (Documentation) में मदद करता है। हालांकि, नोटपैड में हम किसी फाइल को आकर्षक स्वरूप नहीं दे सकते हैं। इस प्रोग्राम में सारा पाठ्य (Text) एक ही फॉन्ट में दिखाया जा सकता है। नोटपैड के संक्षिप्त इतिहास की चर्चा करें तो वर्ष 1983 में **रिचर्ड ब्रॉडी** ने माइक्रोसॉफ्ट के लिए डिस्क ऑपरेटिंग सिस्टम (DOS) के लिए पहला नोटपैड तैयार किया था , जिसे माउस की मदद से भी ऑपरेट किया जा सकता था। कालान्तर में माइक्रोसॉफ्ट विंडोज के हर वर्जन में नोटपैड में अपेक्षित सुधार किए जाते रहे। मौजूदा दौर में विंडोज के सभी प्रचलित ऑपरेटिंग सिस्टम का यह अभिन्न अंग है।

- **नोटपैड के घटक (Components of Notepad)**

सामान्यतः डेस्कटॉप या लैपटॉप की होमस्क्रीन के टूलबार में नोटपैड का शॉर्टकट की (टूलबार में नोटपैड का आइकन) रहती है। इस पर क्लिक करने से नोटपैड प्रोग्राम खुल जाता है। यदि यह शॉर्टकट न हो तो कंप्यूटर के स्टार्ट बटन पर क्लिक करने के बाद ऑल प्रोग्राम्स (All Programs) पर क्लिक करना होता है। यहां खुलने वाली सूची में एसेसरीज (Accessories) पर क्लिक करते ही नोटपैड का ऑप्शन सामने आता है। इस पर क्लिक करके नोटपैड प्रारंभ हो जाएगा। नोटपैड की विंडो में टाइटल बार (Title Bar), पाठ्यक्षेत्र (Text Area) मेन्यू बार (Menu Bar) और कर्सर (Cursor) होते हैं। नोटपैड की सबसे उपर की पंक्ति टाइटल बार है। इसमें उस दस्तावेज या फाइल का नाम नजर आता है, जो उपयोगकर्ता लिख रहा हो और उसे सेव कर कुछ नाम दिया हो। अगर फाइल सेव नहीं है तो यहां अनटाइटल्ड (Untitled) लिखा दिखता है। टाइटल बार के ठीक नीचे मेन्यू बार होता है। इसमें एडिट, फाइल, फॉरमेट, व्यू, हेल्प जैसे ऑप्शन होते हैं। हर मेन्यू में कई विकल्प होते हैं, जिनका प्रयोग उपयोगकर्ता अपनी जरूरत के हिसाब से कर सकता है।

पाठ्यक्षेत्र नोटपैड विंडो का वह हिस्सा है , जहां उपयोगकर्ता टाइपिंग की मदद से अपनी फाइलए सूचना, दस्तावेज तैयार करता है। नोटपैड खोलने पर यह हिस्सा खाली नजर आता है और टाइप करते जाने पर यह भरता जाता है। पाठ्यक्षेत्र में एक खड़ी लकीर (।) नजर आती है , जिसे कर्सर कहा जाता है। कर्सर जिस जगह पर हो, वहां से टाइपिंग प्रारंभ की जा सकती है। अगर किसी शब्द में सुधार (Correction) करना हो तो उपयोगकर्ता माउस की मदद से कर्सर को संबंधित शब्द पर ले जाकर टाइप कर सकता है।

- **वर्डपैड (Wordpad)**

वर्डपैड भी नोटपैड की तरह टेक्स्ट एडिटर (Text Editor) है, लेकिन यह इस तरह का सॉफ्टवेयर या प्रोग्राम है, जिसकी मदद से फाइल के निर्माण के अलावा उसका संपादन और छपाई (Printing) भी की जा सकती है। सामान्यतः इस तरह के प्रोग्रामों को वर्ड प्रोसेसर (Word Processor) कहा जाता है। विंडोज ऑपरेटिंग सिस्टम में नोटपैड की तरह ही वर्डपैड प्रोग्राम भी प्री -इंस्टॉल रहता है। नोटपैड से इतर इस प्रोग्राम में उपयोगकर्ता को टाइपिंग के जरिये टेक्स्ट फाइल तैयार करने के अलावा ग्राफिक (Graphics) यानी चित्र और संकेत या चिह्न भी शामिल करने की सहूलियत मिलती है। इसके लिए इसमें टाइटल बार और मेन्यू बार के अलावा टूल बार भी शामिल किया गया है। इसे हम निम्न चित्र से समझ सकते हैं।

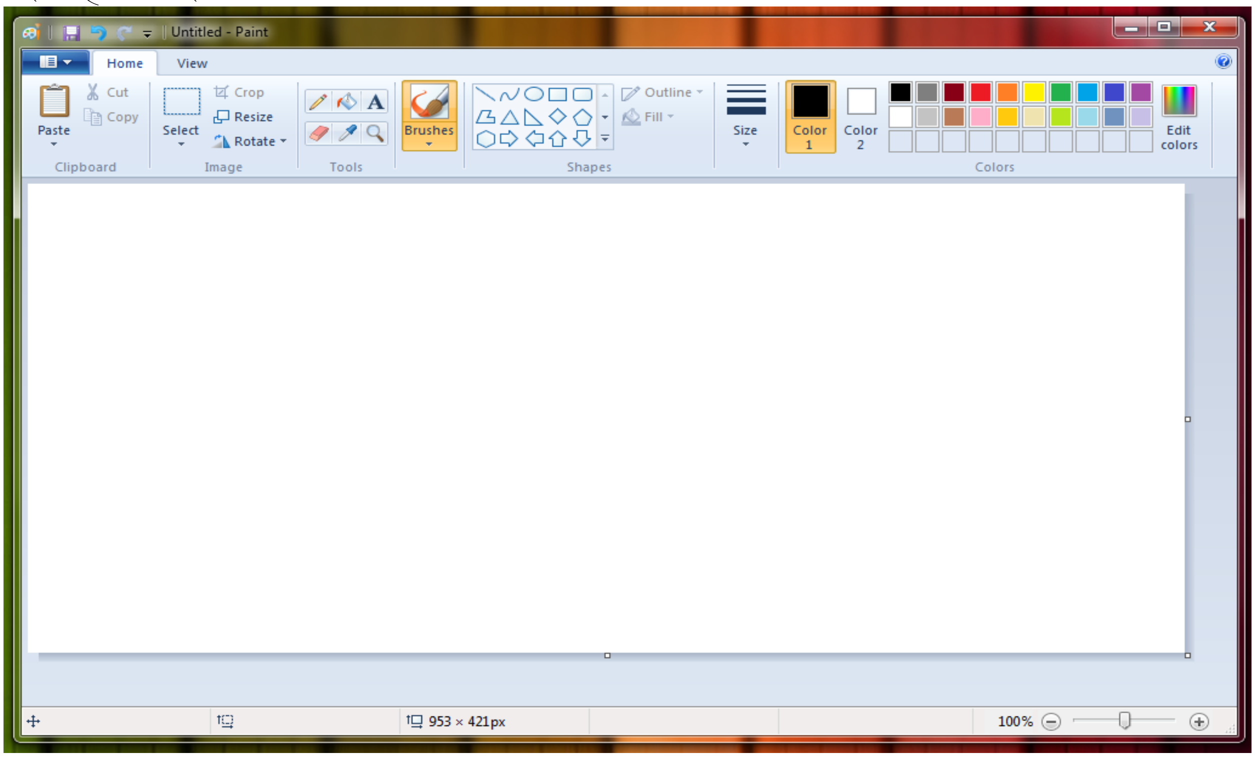
(माइक्रोसॉफ्ट वर्डपैड की विंडो)

वर्डपैड में बाकी सभी चीजें मूलतः नोटपैड जैसी ही हैं , लेकिन इसमें जोड़े गए टूल की मदद से उपयोगकर्ता के लिए टेक्स्ट फाइल को आकर्षक बनाने में मदद मिलती है। इसमें बोल्ड , इटेलिक, अंडरलाइन जैसे ऑप्शन के अलावा पेज (Page) को जूम या अनजूम करने की भी सुविधा उपलब्ध है। साथ ही इन्सर्ट (Insert) विकल्प के जरिये वर्डपैड में टेक्स्ट फाइल के साथ फोटो भी जोड़ी जा सकती है। फॉन्ट की विस्तृत श्रृंखला से उपयोगकर्ता मनचाहा फॉन्ट चुन सकता है, जबकि फॉन्ट का साइज भी वह अपने हिसाब से तय कर सकता है। माइक्रोसॉफ्ट ने सबसे पहले वर्ष 1995 में अपने ऑपरेटिंग सिस्टम विंडोज 95 में वर्डपैड को लांच किया था। इसके जरिये वर्ड फाइल को रिच टेक्स्ट फॉरमेट (Rich Text Format- RTF) में सुरक्षित कर पाना संभव हो सका। आरटीएफ दरअसल, माइक्रोसॉफ्ट फाइल सिस्टम का खास अधिकृत स्वरूप है , जिसमें किसी दस्तावेज , डाटा को वर्ड फाइल में इस तरह सुरक्षित किया जाता है कि माइक्रोसॉफ्ट के सभी प्रोग्राम इस फाइल को आसानी से पढ़-समझ सकें। माइक्रोसॉफ्ट के लिए नोटपैड प्रोग्राम तैयार करने वाले रिचर्ड ब्रॉडी ने अपने साथियों चार्ल्स सिमोनी और डेविड ल्यूबर्ट के साथ मिलकर आरटीएफ फॉरमेट का तरीका तैयार किया था। इसकी मदद से ही वर्डपैड को और अधिक परिष्कृत कर पाना संभव हुआ। माइक्रोसॉफ्ट ने वर्ष 2008 में आरटीएफ फॉरमेट की मेंटेनेंस का काम बंद कर दिया है, लेकिन अब भी यह विंडोज ऑपरेटिंग सिस्टम में उपयोग किया जाता है।

- **पेंट (Paint)**

विंडोज ऑपरेटिंग सिस्टम के सभी संस्करणों की एसेसरीज में यह प्रोग्राम भी उपलब्ध रहता है। पेंट दरअसल , ग्राफिकल प्रोग्राम है यानी इसकी मदद से उपयोगकर्ता रंग -बिरंगे चित्र बना सकता है। इसमें कई ऐसे साधन (Tools) मौजूद हैं, जिनकी मदद से उपयोगकर्ता अभीष्ट आकार, रंग, रेखाएं आदि के जरिये मनचाहा चित्र बना सकता है। इस प्रोग्राम की सबसे बड़ी खासियत यह है कि यह विंडोज के प्रारंभिक प्रोग्रामों में से एक है। वर्ष 1985 में विंडोज ने अपना पहला ऑपरेटिंग सिस्टम विंडोज 1.0 लांच किया था तो उसमें भी पेंट प्रोग्राम शामिल था।

हालांकि, उस वक्त इसे माइक्रोसॉफ्ट पेंट की जगह पेंटब्रश (Paintbrush) कहा जाता था। तब इसे जेडसॉफ्ट कॉरपोरेशन ने पीसी पेंटब्रश के नाम से तैयार किया था। इस लिहाज से इस प्रोग्राम का अधिकृत लाइसेंस इसी कंपनी के पास था। शुरुआती दौर में यह प्रोग्राम सिर्फ एक बिट मोनोक्रोम ग्राफिक्स को ही सपोर्ट करता था। माइक्रोसॉफ्ट ने जब अपना ऑपरेटिंग सिस्टम विंडोज 3.0 वर्ष 1990 में लांच किया तो पेंटब्रश को नये ग्राफिकल यूजर इंटरफेस (Graphical User Interface) के हिसाब से दोबारा तैयार किया गया। इससे यह कई तरह की फाइल को सपोर्ट करने लगा।



अक्टूबर 2016 में सोशल नेटवर्किंग साइट ट्विटर पर किसी व्यक्ति ने विंडोज पेंट का नया वर्जन विंडोज पेंट 3 -डी (Paint 3-D) का एक ट्यूटोरियल वीडियो डाला। उस समय तक माइक्रोसॉफ्ट ने ऐसा कोई आधिकारिक प्रोग्राम लांच नहीं किया था। लेकिन , यह वीडियो लीक होने के बाद 28 अक्टूबर 2016 को माइक्रोसॉफ्ट ने अपना डमी एप न्यूकैसल (Newcastle) जारी किया , ताकि उपयोगकर्ता लीकेज वीडियो से लिंक के बजाय कंपनी का अधिकृत सॉफ्टवेयर ही इस्तेमाल करें।

माइक्रोसॉफ्ट की ओर से यह भी जानकारी दी गई कि यह प्रोग्राम जुलाई 2015 में लांच हुए विंडोज के नवीनतम ऑपरेटिंग सिस्टम विंडोज 10 के लिए तैयार किया गया है। माइक्रोसॉफ्ट ने विंडोज 10 उपयोगकर्ताओं के लिए विशेष वेबसाइट भी बनाई है , जिसमें इस प्रोग्राम के संचालन के तरीके बताए गए हैं। इस प्रोग्राम की खासियत इस पर 3 -डी पेंटिंग करना है। इसके अलावा इस प्रोग्राम में पारदर्शी 2 -डी पेंटिंग भी संभव है।

- **कैल्कुलेटर (Calculator)**

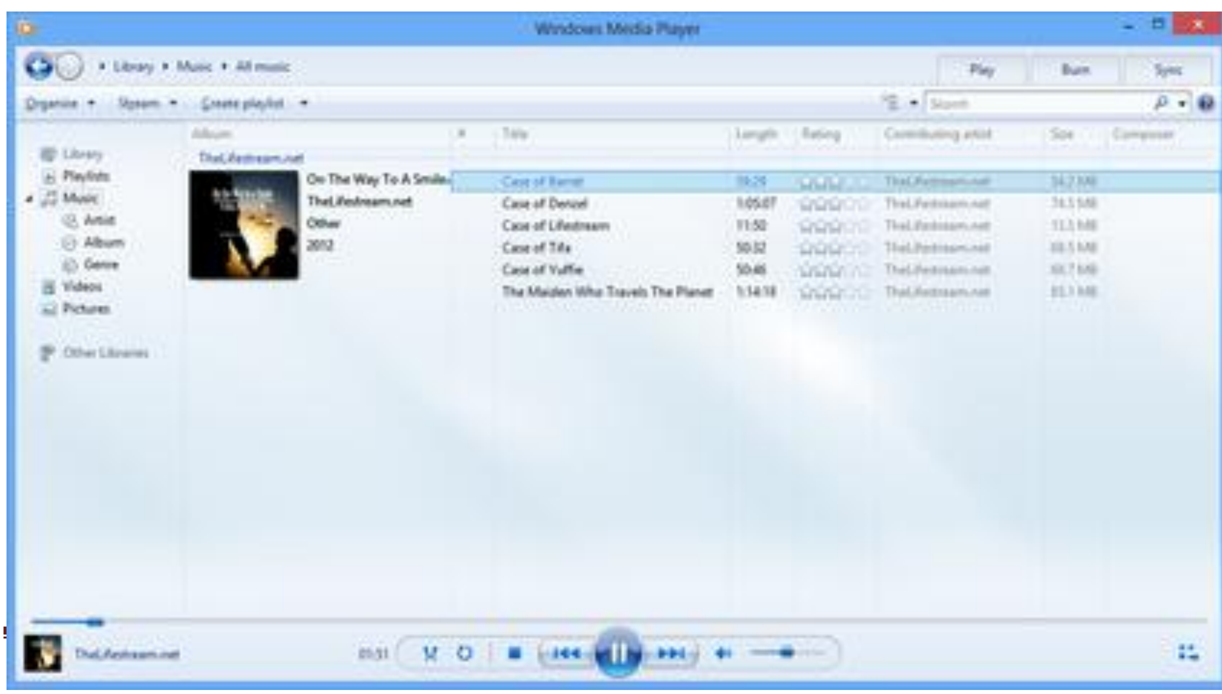
हम जानते हैं कि कंप्यूटर की खोज और विकास का मूल आधार गणनाएं (Calculators) थीं। समय के साथ कंप्यूटर पर अन्य आयाम जुड़ते चले गए , लेकिन गणनाएं आज भी कंप्यूटर का मूल उद्देश्य है। यही वजह है कि विंडोज के हर ऑपरेटिंग सिस्टम पर कैल्कुलेटर (Calculator) प्रोग्राम अनिवार्य रूप से उपलब्ध रहता है। कंप्यूटर पर काम करते वक्त किसी भी तरह की गणना करने में यह उपयोगकर्ता की मदद करता है। कैल्कुलेटर प्रोग्राम भी दो तरह का होता है , पहला सामान्य (Standard) और दूसरा वैज्ञानिक (Scientific) सामान्य

कैल्कुलेटर में गुणा-भाग, जोडना-घटाना, प्रतिशत मान निकालना जैसी समान्य गणितीय प्रक्रियाओं को संपन्न करने की सुविधा उपलब्ध होती है। कैल्कुलेटर प्रोग्राम पर लिखे अंकों और गुणा-भाग, जोड़-घटाव के चिह्नों की मदद से उपयोगकर्ता आसानी से अभीष्ट परिणाम हासिल कर सकता है, लेकिन यदि उपयोगकर्ता को और जटिल गणनाएं करनी हों तो वह साइंटिफिक कैल्कुलेटर का इस्तेमाल कर सकता है। इस कैल्कुलेटर में सामान्य गणितीय प्रक्रियाओं के अलावा वर्ग (Square) घन (Cube) त्रिकोणमितीय मान (Trigonometrical Values) समेत सांख्यिकीय (Statistical) गणनाएं करने की भी सुविधा उपलब्ध रहती है।

- **विंडोज मीडिया प्लेयर (Windows Media Player)**

विंडोज मीडिया प्लेयर (WMP) भी माइक्रोसॉफ्ट द्वारा निर्मित एक ऐसा प्रोग्राम है, जो पर्सनल कंप्यूटरों पर इंस्टॉल विंडोज ऑपरेटिंग सिस्टम के साथ उपलब्ध रहता है। हम लगातार इस बात को दोहरा रहे हैं कि कंप्यूटर के विकास के अनुक्रम में समय के साथ गणनाएं ही एकमात्र उद्देश्य नहीं रह गया था। कंप्यूटर मनोरंजन का भी बड़ा साधन बनते चले गए और विंडोज मीडिया प्लेयर ऐसा ही एक साधन है, जो कंप्यूटर के जरिये उपयोगकर्ता को ऑडियो सुनने तथा वीडियो और फोटो देखने की सुविधा प्रदान करता है। पर्सनल कंप्यूटर के अलावा विंडोज मीडिया प्लेयर उन पॉकेट पीसीए टैबलेट और मोबाइल फोन पर भी उपलब्ध रहता है, जो विंडोज ऑपरेटिंग सिस्टम पर चलते हैं।

विंडोज ने सबसे पहले वर्ष 1991 में मीडिया प्लेयर लांच किया था, जब विंडोज ऑपरेटिंग सिस्टम का वर्जन विंडोज 3.0 जारी किया गया था। उस वक्त मीडिया प्लेयर एनीमेशन फाइलों को ही देखने में उपयोग किया जा सकता था, लेकिन जैसे-जैसे विंडोज ऑपरेटिंग सिस्टम के नये वर्जन लांच होते गए, मीडिया प्लेयर में भी सुधार आता गया। विंडोज के लगभग हर वर्जन के साथ मीडिया प्लेयर का भी सुधारीकृत वर्जन लांच किया जाता रहा। विंडोज 3.1 के साथ पहली बार मीडिया प्लेयर में वीडियो चलाने की भी सुविधा उपलब्ध कराई गई। विंडोज 95 में मीडिया प्लेयर की वीडियो चलाने की क्षमताओं में और अधिक सुधार किया गया। विंडोज मीडिया प्लेयर का आखिरी वर्जन विंडोज मीडिया प्लेयर 12 वर्ष 2009 में विंडोज 7 ऑपरेटिंग सिस्टम के साथ लांच किया था। मीडिया प्लेयर का यह वर्जन विंडोज 7 और इसके बाद अब तक जारी हुए विंडोज के सभी ऑपरेटिंगसिस्टमों में संचालित किया जाता है। विंडोज मीडिया प्लेयर 12 को हम निम्न चित्र से समझ सकते हैं- विंडोज मीडिया प्लेयर उपयोगकर्ता को सिर्फ गाने सुनने, वीडियो देखने की ही सुविधा प्रदान नहीं करता, बल्कि इसकी मदद से उपयोगकर्ता ऑडियो सीडी, एमपी3 सीडी तैयार करने या सीडी से ऑडियो -वीडियो को



कंप्यूटर पर सुरक्षित रखने का काम भी कर सकता है।
यही नहीं, विंडोज मीडिया प्लेयर इंटरनेट के जरिये उपयोगकर्ता को ऑनलाइन म्यूजिक स्टोर (Online Music Store) या विंडोज की ऑनलाइन मीडिया लाइब्रेरी (Windows Media Library) से भी जोड़ देता है। यहां यह भी उल्लेखनीय है कि विंडोज मीडिया प्लेयर को विंडोज ऑपरेटिंगसिस्टम के अलावा मैक ऑपरेटिंगसिस्टम पर भी उपयोग किया जाता था।

15.4 एमएस ऑफिस (MS Office)

माइक्रोसॉफ्ट ऑफिस (MS-Office) एप्लीकेशन प्रोग्रामों, सर्वर और सुविधाओं का एक ऐसा ऑफिस सुइट (Office Suite) है, जिसकी मदद से किसी कार्यालय (Office) के दैनिक कार्यों को कम समय में और प्रामाणिकता (Authenticity) और शुद्धता (Accuracy) के साथ संपन्न किया जा सके। ऑफिस सुइट का तात्पर्य उन सुविधाओं से है , जो ऑफिशियल कार्यों में मददगार हों। इनमें दस्तावेजों का संरक्षण , निर्माण , बिल , प्रजेंटेशन, हिसाब-किताब आदि गतिविधियां शामिल हैं।

बिल गेट्स ने माइक्रोसॉफ्ट ऑफिस सुइट तैयार करने की घोषणा सबसे पहले 1 अगस्त 1988 को अमेरिका के लास वेगास में आयोजित कॉमडेक्स (COMDEX) में की थी। कॉमडेक्स का अर्थ है कंप्यूटर डीलर्स एक्जीबिशन (Computer Dealer’s Exhibition) यह प्रदर्शनी जैसा कि नाम से ही स्पष्ट है , नवीनतम कंप्यूटर, कंप्यूटर तकनीक और कंप्यूटर उत्पादों की जानकारी लोगों को देने के मकसद से आयोजित होती थी। वर्ष 1979 से 2003 तक हर साल आयोजित होती रही इस प्रदर्शनी में दुनियाभर से कंप्यूटर निर्माता कंपनियों के प्रतिनिधि, कंप्यूटर उपकरण बेचने वाले डीलर और कंप्यूटर विशेषज्ञ-वैज्ञानिक शामिल होते थे।
वर्ष 1990 में लांच हुए विंडोज ऑपरेटिंगसिस्टम विंडोज 3.0 में माइक्रोसॉफ्ट ऑफिस सुइट लांच की गई , जिसमें ऑफिस के तीन प्रमुख प्रोग्राम माइक्रोसॉफ्ट वर्डए माइक्रोसॉफ्ट एक्सेल और माइक्रोसॉफ्ट पॉवर प्वाइंट शामिल थे। माइक्रोसॉफ्ट ने समय के साथ ऑफिस सुइट में कुछ और नये फीचर्स भी जोड़े , हालांकि, ये तीनों प्रोग्राम हमेशा सुइट के मूल आधार बने रहे। खास बात यह है कि एमएस -ऑफिस के तीनों प्रोग्रामों के ग्राफिकल यूजर इंटरफेस में भी उसी तरह समान बदलाव किए गए , जैसे संबंधित ऑपरेटिंगसिस्टम के ग्राफिकल यूजर इंटरफेस में किए जाते थे। इससे विंडोज के हर संस्करण के साथ एमएस-ऑफिस भी लगातार अपडेट होते रहे। दुनियाभर में प्रचलित ऑपरेटिंगसिस्टम, प्रोग्राम, ऑनलाइन गेम्स और कंप्यूटर से जुड़े अन्य उत्पादों की बिक्री , डिमांड जैसे पहलुओं पर नजर रखने वाली वेबसाइट सॉफ्टपीडिया (Softpedia) के अनुसार वर्ष 2012 तक विश्वभर में माइक्रोसॉफ्ट ऑफिस इस्तेमाल करने वाले उपयोगकर्ताओं की संख्या एक अरब से भी अधिक है।

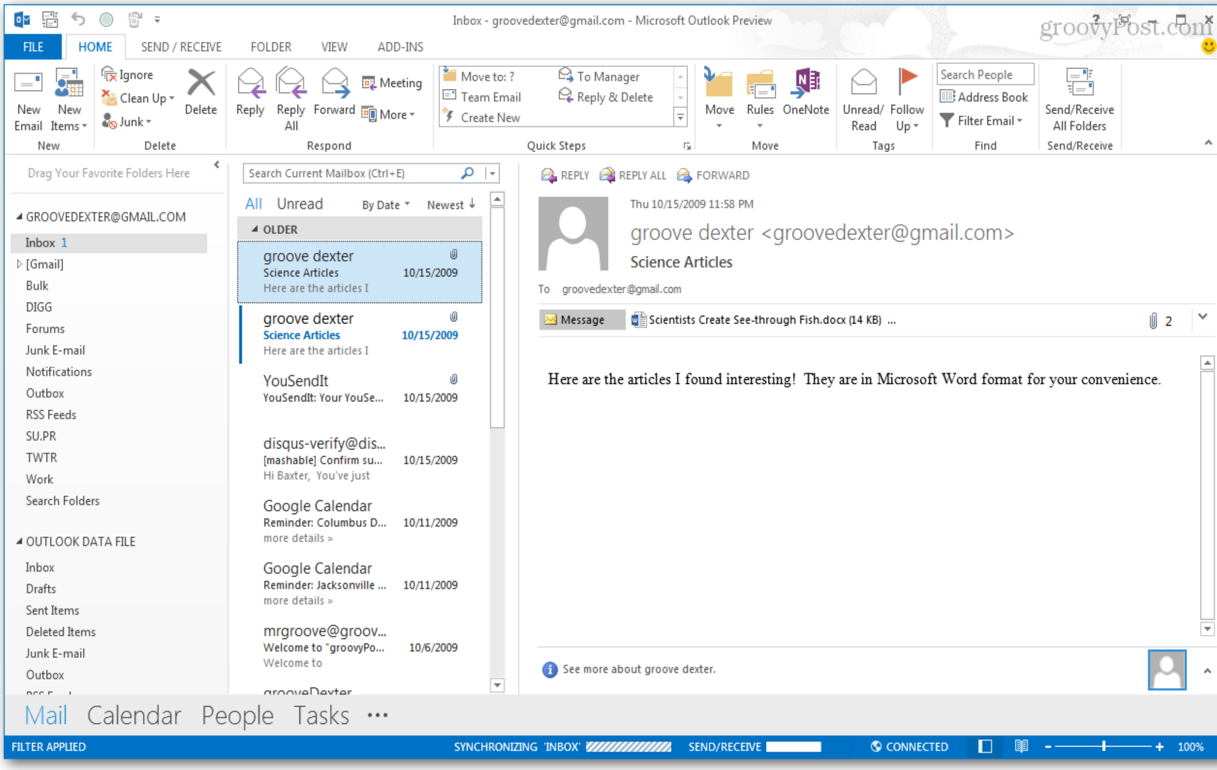
एमएस-ऑफिस के घटक (Components of MS-Office)

माइक्रोसॉफ्ट ऑफिस को समय की मांग के अनुरूप लगातार सुधारा और परिष्कृत किया जाता रहा है। ऐसे में एमएस-ऑफिस में तीनों बुनियादी प्रोग्रामों के अलावा कई नये प्रोग्राम भी लगातार जुड़ते गए हैं। इकाई के इस हिस्से में हम एमएस-ऑफिस के ऐसे ही कुछ प्रोग्राम के बारे में जानकारी हासिल करेंगे-

- **एमएस-वर्ड (MS-Word)-** माइक्रोसॉफ्ट वर्ड एक वर्ड प्रोसेसर हैए यानी एक ऐसा प्रोग्रामए जिसमें उपयोगकर्ता टेक्स्ट फाइल तैयार करने के साथ उसमें ग्राफिकल सुधार भी कर सकता है। यह विंडोज और मैक ऑपरेटिंगसिस्टम पर समान रूप से काम करता है। एमएस.वर्ड के बारे में हम यूनिट के अगले हिस्से में विस्तार से जानेंगे।
- **एमएस-एक्सेल (MS-Excel)-** माइक्रोसॉफ्ट एक्सेल मूलतः स्प्रेडशीट (Spreadsheet) आधारित प्रोग्राम है। यहां दिलचस्प पहलू यह है कि एमएस -एक्सेल को सबसे पहले वर्ष 1985 में मैक ऑपरेटिंगसिस्टम के लिए तैयार किया गया था, दो साल बाद वर्ष 1987 में यह प्रोग्राम विंडोज के साथ कुछ सुधारीकरण के बाद शामिल किया गया। एमएस -एक्सेल के बारे में भी हम यूनिट के अगले हिस्से में विस्तार से जानेंगे।

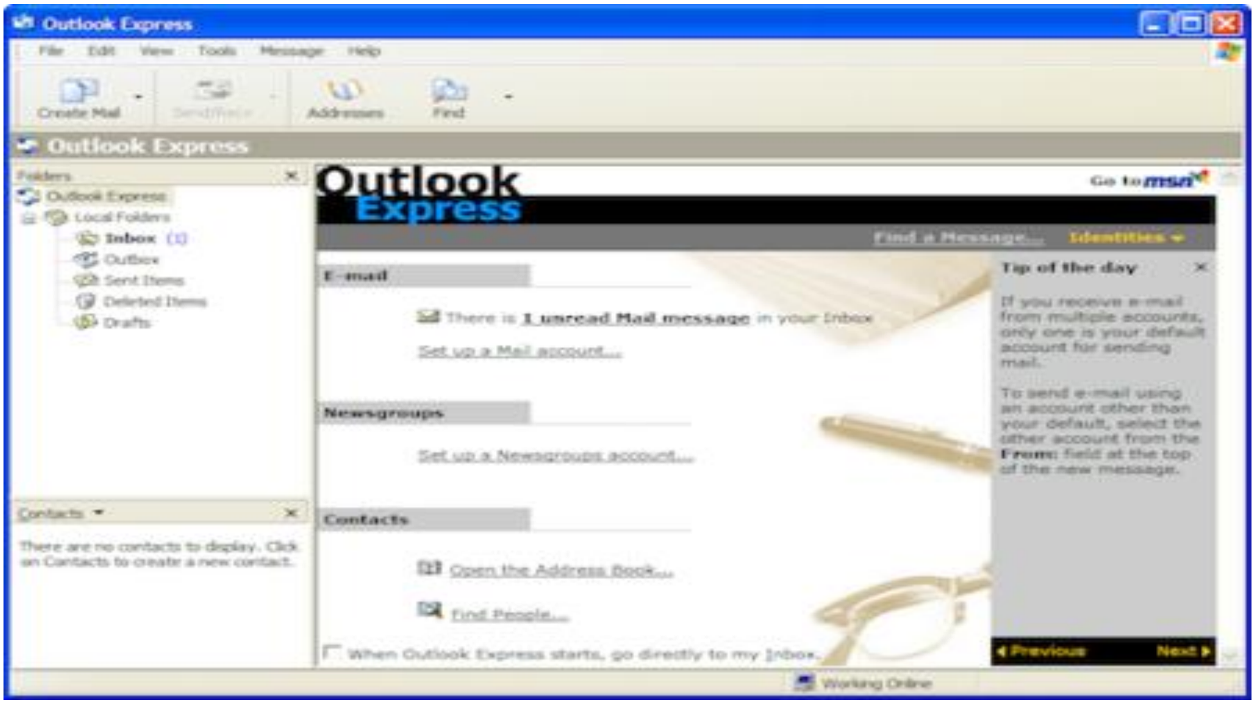
- **एमएस-पॉवरप्वाइंट (MS-Powerpoint)-** माइक्रोसॉफ्ट पॉवरप्वाइंट विंडोज और मैक ऑपरेटिंगसिस्टम का प्रजेंटेशन प्रोग्राम (Presentation Program) है। इसकी मदद से उपयोगकर्ता टेक्स्ट, ग्राफिक्स आदि की मदद से स्लाइड शो (Slide Show) तैयार कर सकता है, जिन्हें प्रिंट किया जा सकता है या प्रोजेक्टर की मदद से प्रजेंट करना संभव हो पाता है।
- **एमएस-एक्सेस (MS-Access)-** माइक्रोसॉफ्ट एक्सेस मूलतः डाटाबेस मैनेजमेंट सिस्टम (Database Management) है, जो एमएस-ऑफिस सुइट का हिस्सा है। हालांकि, पर्सनल कंप्यूटरों से इतर इसका उपयोग प्रोफेशनल (Proffessional) कंप्यूटरों पर ही किया जाता है। डाटाबेस का अर्थ प्रोग्राम स्कीम, क्वेरी (Queries), टेबल, रिपोर्ट और उन अन्य जरूरी डाटा का सामूहिक स्वरूप है जो किसी प्रोग्राम या ऑपरेटिंगसिस्टम के सफल संचालन के लिए जरूरी होता है। डाटाबेस मैनेजमेंट सिस्टम दरअसल एक तरह का सॉफ्टवेयर प्रोग्राम होता है जो उपयोगकर्ता , कंप्यूटर पर मौजूद दूसरी एप्लीकेशनों और अपने ही भीतर मौजूद डाटा का परीक्षण (Analyzation) करता है, जिससे प्रोग्राम की कार्यशैली में सुधार आता है। किसी तरह की दिक्कत आने पर यह ऑनलाइन जुड़ने पर जेट डाटाबेस इंजन (Jet Database Engine) से संपर्क करता है। जेट डाटाबेस इंजन , उन सभी प्रोग्रामों , एप्लीकेशनों के डाटाबेसों का विस्तृत खजाना है, जो माइक्रोसॉफ्ट ने तैयार किए हैं।
- **माइक्रोसॉफ्ट आउटलुक (MS-Outlook)-** माइक्रोसॉफ्ट आउटलुक मूलतः पर्सनल इंफॉर्मेशन मैनेजर (Personal Information Manager) है। विंडोज ऑपरेटिंगसिस्टम के साथ यह प्रोग्राम इसलिए शमिल किया गया था कि उपयोगकर्ता इसकी मदद से अपनी व्यक्तिगत जानकारियां , फोटो, डाटा, वीडियो और कोई भी अपेक्षित सूचना इस प्रोग्राम में सुरक्षित कर सकता था।

विंडोज ने वर्ष 1997 में लांच विंडोज ऑपरेटिंगसिस्टम 97 के साथ पहली बार आउटलुक को शामिल किया था। आउटलुक की खासियत यह है कि इसे ऑपरेटिंगसिस्टम के सहयोगी प्रोग्राम के तौर पर भी इस्तेमाल किया जा सकता है और एक स्वतंत्र एप्लीकेशन या प्रोग्राम के तौर पर भी। हालांकि , यह विशेष तौर पर स्वतंत्र प्रोग्राम के तौर पर ही उपयोग किया जाता है। इसकी वजह इसमें ई -मेल, कैलेंडर, टास्क मैनेजर (Task Manager) फंक्शन की उपलब्धता है। अपने फीचर की वजह से यह कार्यालयी (Official) कार्यों के लिए उपयुक्त प्रोग्राम बन गया। माइक्रोसॉफ्ट एक्सचेंज सर्वर (Microsoft Exchange Server) और माइक्रोसॉफ्ट शेयरप्वाइंट सर्वर (Microsoft Sharepoint Server) से जुड़कर एमएस -आउटलुक साझा मेल बॉक्स (Mailbox), एक्सचेंज पब्लिक फोल्डर, शेयर प्वाइंट सूची और उपयोगकर्ता के अन्य ई -मेल कंपनियों पर बने ई-मेल खातों (E-mail Accounts) पर मिलने वाली ई-मेल को आयात करने जैसे कार्यालयी उपयोगी कार्यों में खासा मददगार होता है। यही वजह है कि आज अधिकतर कंपनियों में आउटलुक को ही कर्मचारियों की अधिकृत ई-मेल आईडी के तौर पर प्रयोग किया जाता है। जनवरी 2015 में माइक्रोसॉफ्ट ने एंड्रॉयड और आईफोन ऑपरेटिंगसिस्टम पर चलने वाले स्मार्टफोन (Smartphones) और टैबलेट (Tablets) के लिए ऑफिस 365 के साथ ई-मेल, कैलेंडर और कांटेक्ट फीचर वाला आउटलुक वर्जन जारी किया।



(माइक्रोसॉफ्ट आउटलुक एप्लीकेशन)

- **आउटलुक एक्सप्रेस (Outlook Express)-** माइक्रोसॉफ्ट के इस एप्लीकेशन का नाम भी पिछले एप्लीकेशन या प्रोग्राम के समान है , लेकिन यह आउटलुक से भिन्न है। दरअसल यह प्रोग्राम मूलतः ई - मेल और माइक्रोसॉफ्ट इंटरनेट एक्सप्लोरर से जुड़ा हुआ था। वर्ष 1996 में माइक्रोसॉफ्ट ने माइक्रोसॉफ्ट इंटरनेट मेल एंड न्यूज (MS-Internet Mail and News) नाम से नया फीचर विंडोज 95 में शामिल किया। इसके साथ ही इंटरनेट एक्सप्लोरर 3 वर्जन भी लांच किया गया था। हालांकि , तब यह मेल तब सिर्फ प्लेन टेक्स्ट (Plain Text) या रिच टेक्स्ट फॉरमेट (RTF) को ही सपोर्ट करती थी, हाइपर टेक्स्ट मार्कअप लैंग्वेज (HTML) को नहीं। वर्ष 1997 में माइक्रोसॉफ्ट इंटरनेट मेल एंड न्यूज को परिष्कृत कर इंटरनेट एक्सप्लोरर 4.0 के साथ आउटलुक एक्स्प्रेस लांच किया गया।



(आउटलुक एक्सप्रेस)

आउटलुक एक्सप्रेस की बड़ी खामी यह थी कि इसमें सुरक्षा उपकरणों का अभाव था। हालांकि , माइक्रोसॉफ्ट ने इसमें कुछ फीचर जोड़ने का प्रयास तो किया, लेकिन वे नाकाफी साबित हुए। आखिर वर्ष 2005 में विंडोज मेल (Windows Mail) की लांचिंग के साथ आउटलुक एक्सप्रेस को रिप्लेस (Replace) कर दिया गया।

- **एमएस-वन नोट (MS-One Note)-** माइक्रोसॉफ्ट वन नोट भी एमएस -ऑफिस सुइट का की (Key) कंपोनेंट है। यह प्रोग्राम मूलतः नेटवर्किंग पर आधारित है। वर्ष 2003 में रिलीज हुआ यह प्रोग्राम एमएस-ऑफिस के ऑनलाइन संस्करण के जरिये यह किसी वन नोट उपयोगकर्ता को दूसरे वन नोट उपयोगकर्ता तक सूचनाएं भेजने की सुविधा प्रदान करता है। इसके अलावा यह उपयोगकर्ता के हर तरह के डाटा को संग्रहीत (Collect) करने की भी सहूलियत देता है , चाहे वे हस्तलिखित (Handwritten) हों, ड्राइंग (Drawing) के रूप में हों या ऑडियो (Audio) के रूप में। विंडोज 10 समेत यह फीचर इस तरह के ऑपरेटिंगसिस्टमों में ज्यादा कारगर है, जो टचस्क्रीन हैं या जिनमें कीबोर्ड के बजाय पेन (Digital Pen) का इस्तेमाल किया जाता है। यही वजह है कि माइक्रोसॉफ्ट ने इस प्रोग्राम का स्टैंडअलोन (Standalone) संस्करण (Version), विंडोज फोन, आईफोन और एंड्रॉयड के लिए लांच किया है।
- **एमएस-स्वे (MS-Sway)-** माइक्रोसॉफ्ट स्वे एमएस -ऑफिस का नवीनतम प्रोग्राम है , जो विंडोज 10 के साथ रिलीज किया गया है। 39 भाषाओं में उपलब्ध स्वे मूलतः पॉवरप्वाइंट की तरह प्रजेंटेशन पर आधारित है, लेकिन यह प्रोग्राम उपयोगकर्ता को यह सुविधा प्रदान करता है कि वह टेक्स्ट , फोटो, ऑडियो आदि को जोड़कर प्रजेंटेशन लायक वेबसाइट (Website) बना सके। हालांकि , इसके लिए उपयोगकर्ता का माइक्रोसॉफ्ट एकाउंट होना जरूरी है। एकाउंट बनने के बाद उपयोगकर्ता अपने डेस्कटॉप, लैपटॉप पर संग्रहीत डाटा को स्वे के जरिये माइक्रोसॉफ्ट सर्वर पर सुरक्षित रख सकता है। यही नहीं, स्वे उपयोगकर्ता को यह भी सहूलियत प्रदान करता है कि वह फेसबुक जैसी सोशल नेटवर्किंग साइट, यू-ट्यूब आदि से सीधे कोई फोटो, ऑडियो, लिंक या कोई अन्य डाटा-जानकारी भी स्वे के जरिये अपनी वेबसाइट में सीधे जोड़ सकता है। इन वेबसाइट को किसी भी ऐसी वेब एप (Web Applications) की मदद से देखा या संपादित किया जा सकता है , जो एमएस-ऑफिस ऑनलाइन एप

पर संचालित होने में सक्षम हों। विंडोज 10 के अलावा आईफोन भी ऐसे ऑपरेटिंगसिस्टम हैं , जिन पर स्वे को चलाया जा सकता है। एंड्रॉयड और विंडोज फोन पर यह सुविधा फिलहाल उपलब्घ नहीं है , लेकिन माइक्रोसॉफ्ट की ओर से इस दिशा में काम किया जा रहा है।

- **एमएस-डेस्कटॉप पब्लिशिंग (Desktop Publishing)-** डेस्कटॉप पब्लिशिंग एमएस -ऑफिस सुइट में शामिल वह प्रोग्राम है, जो उपयोगकर्ता को पर्सनल कंप्यूटर पर ऐसी सामग्री तैयार करने की सहूलियत प्रदान करता है , जो छापी (Print) जानी हो। इस प्रोग्राम में पेज लेआउट (Page Layout) टेक्स्ट, इमेज आदि की लंबी श्रृंखला रहती है। इनकी मदद से ग्रीटिंग कार्ड , कैलेंडर , पत्रिकाएं , ब्रोशर (Brouchures) लेबल, स्टिकर, बिजनेस कार्ड, पोस्टकार्ड, वेबसाइट आदि डिजाइन किए जाते हैं। इस एप्लीकेशन का इस्तेमाल मूलतः पब्लिशिंग कारोबारी ही करते हैं।
- **एमएस-सर्वर और वेब सर्विस (MS-Server & Web Services)-** माइक्रोसॉफ्ट अपने सभी ऑपरेटिंगसिस्टम, प्रोग्राम और ए मएस-ऑफिस सुइट को साझा सर्वर के जरिये संयुक्त रखता है। माइक्रोसॉफ्ट शेयर प्वाइंट ऐसा ही सर्वर है। इसकी वजह से माइक्रोसॉफ्ट के सभी एप्लीकेशन का इस्तेमाल ऑनलाइन करना भी संभव हो पाता है। एक्सेल सर्वर , एमएस-प्रोजेक्ट सर्वर, एमएस-सर्च सर्वर, इंफोपाथ फार्म सर्विस और माइक्रोसॉफ्ट लिंक सर्वर ऐसे ही कुछ सर्वर हैं।

माइक्रोसॉफ्ट वेब सर्विसेज की बात करें तो कंपनी अपने लगभग सभी उत्पादों की ऑनलाइन सेवा भी उपलब्ध कराती है। वर्ड ऑनलाइन , एक्सेल ऑनलाइन , पॉवरप्वाइंट ऑनलाइन , वननोट ऑनलाइन , आउटलुकण्कॉमए पीपुल (आउटलुक.कॉम से संबद्ध एड्रेस बुक), कैलेंडर , डॉक्स.कॉम Docs.com। एमएस-ऑफिस उपयोगकर्ता इसकी मदद से अपने प्रोफाइल में अपनी वर्डफाइल , पीडीएफ, पॉवरप्वाइंट प्रजेंटेशन आदि सुरक्षित रख सकता है, वन ड्राइव, स्वे, प्लानर, वीडियो आदि एमएस-ऑनलाइन सर्विस हैं।

- **एमएस-ऑफिस के वर्जन (Versions of MS-Office)**

एमएस-ऑफिस के कई वर्जन व्यावसायिक उपयोग और पर्सनल कंप्यूटरों के लिए उपलब्ध हैं , लेकिन इनमें सबसे अधिक प्रचलन वाला वर्जन एमएस-ऑफिस डेस्कटॉप वर्जन है। यह वर्जन मूलतः पर्सनल कंप्यूटरों के लिए डिजाइन किया गया था, जिन पर विंडोज ऑपरेटिंगसिस्टम या मैक ऑपरेटिंगसिस्टम चलाए जाते हैं। हम जान चुके हैं कि एमएस-ऑफिस का पहला वर्जन वर्ष 1990 में लांच हुआ था। एमएस-ऑफिस का नवीनतम संस्करण (Version) ऑफिस 2016 (Office 2016) है, जो विंडोज और मैक दोनों के लिए क्रमशः 22 सितंबर 2015 और 9 जुलाई 2015 को लांच किया गया था। विंडोज और मैक के लिए एमएस -ऑफिस सुइट के वर्जन , उनमें उपलब्ध सुविधाएं और लांचिंग का वर्ष निम्नवत हैं-

- **विंडोज (Windows)-** विंडोज ऑपरेटिंगसिस्टम के लिए माइक्रोसॉफ्ट ने पहला एमएस -ऑफिस सुइट वर्ष 1990 में लांच किया था। इसका नाम एमएस -ऑफिस 4.0 था। इसके दो साल बाद एमएस - ऑफिस 3 या 92 वर्ष 1992 में रिलीज हुआ। वर्ष 1993 में एमएस -ऑफिस 4 .X. वर्ष 1995 में एमएस-ऑफिस 95 और वर्ष 1997 में एमएस -ऑफिस 8.0 या 97 लांच किए गए। विंडोज के लिए एमएस-ऑफिस सुइट की श्रृंखला यहीं नहीं थमी। वर्ष 2000 में माइक्रोसॉफ्ट ने एमएस -ऑफिस 9.0 या 2000 लांच किया। इसके दो साल बाद यानी वर्ष 2002 में एमएस -ऑफिस XP रिलीज किया गया। वर्ष 2003 में एमएस-ऑफिस 11.0, 2007 में एमएस-ऑफिस 12.0 और वर्ष 2010 में एमएस-ऑफिस 14.0 माइक्रोसॉफ्ट ने लांच किए। वर्ष 2013 में एमएस -ऑफिस 15 या 2013 के बाद माइक्रोसॉफ्ट का अब तक का अंतिम एमएस-ऑफिस सुइट 2016 वर्ष 2016 में लांच हुआ।

मैक (Mac)- मैक ऑपरेटिंगसिस्टम के लिए माइक्रोसॉफ्ट ने वर्ष 1984 में पहली बार एमएस -वर्ड इंट्रोड्यूस (Introduce) किया था। इसका वर्जन एमएस -वर्ड 1.0 था। एक साल बाद 1985 में मैक के साथ एक्सेल 1.0 और इसके दो साल बाद यानी वर्ष 1987 में पॉवरप्वाइंट 1.0 भी मैक ऑपरेटिंगसिस्टम में शामिल किए गए। वर्ष 1989 में माइक्रोसॉफ्ट ने पहली बार एमएस -ऑफिस सुइट के तीनों बुनियादी प्रोग्रामों वर्ड , एक्सेल और

पॉवरप्वाइंट को संयुक्त कर ऑफिस मैक नाम से मैक ऑपरेटिंगसिस्टम के लिए लांच किया। वर्ष 1991 में ऑफिस मैक 1.5 , 1992 में ऑफिस 3.0 और 1994 में ऑफिस 4.2 जारी किया गया। वर्ष 1998 में माइक्रोसॉफ्ट ने मैक ऑपरेटिंगसिस्टम के लिए अपना ऑफिस सुइट 98 लांच किया। वर्ष 2000 में ऑफिस मैक 2001, वर्ष 2001 में ऑफिस V.X और वर्ष 2004 में ऑफिस 2004 जारी हुए। इसके बाद 2008 में ऑफिस 2008, 2010 में ऑफिस 2011 जारी किए गए। मैक ऑपरेटिंगसिस्टम के साथ वन नोट और आउटलुक की सुविधाएं माइक्रोसॉफ्ट की ओर से वर्ष 2014 में जोड़ी गईं। वर्ष 2015 में माइक्रोसॉफ्ट ने मैक ऑपरेटिंगसिस्टम के लिए अपना अब तक का अंतिम ऑफिस सुइट ऑफिस मैक 2016 लांच किया।

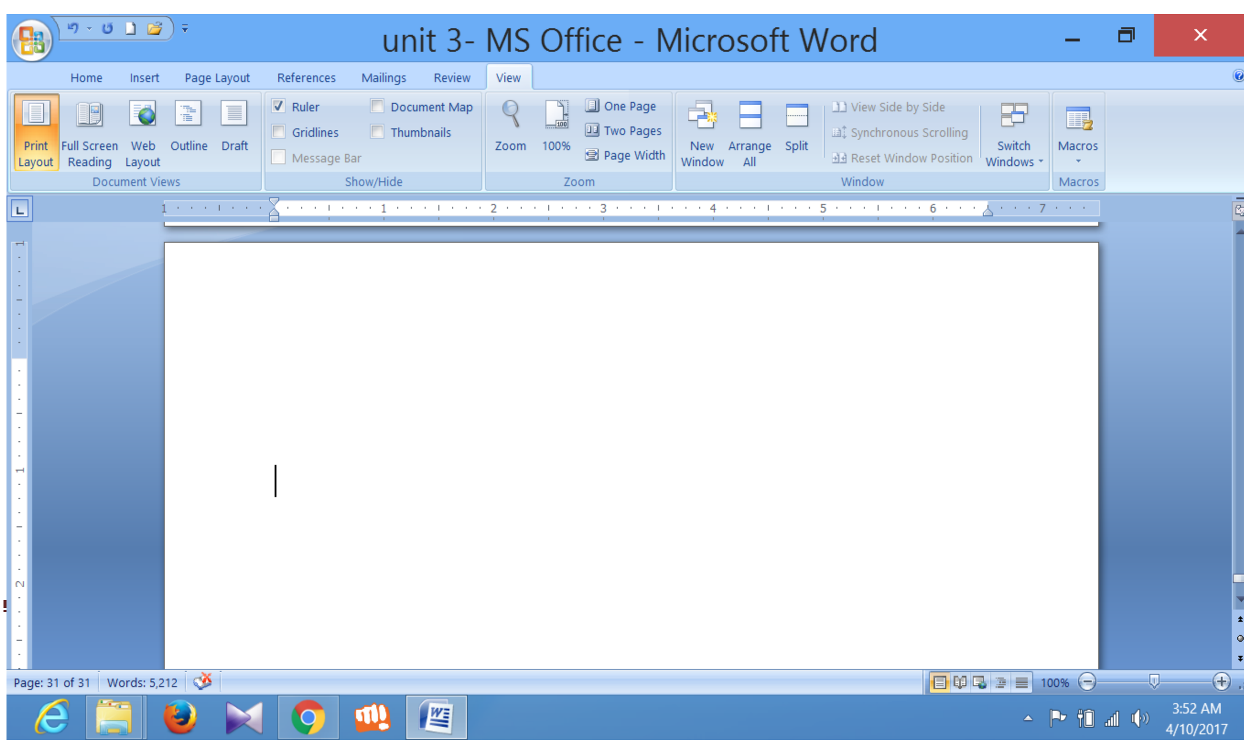
15.5 एमएस वर्ड (MS Word)

माइक्रोसॉफ्ट वर्ड मूलतः एक वर्ड प्रोसेसर (Word Processor) है। माइक्रोसॉफ्ट ने वर्ष 1983 में यूनिक्स (Unix) आधारित जेनिक्स (Xenix) ऑपरेटिंगसिस्टम पर चलने वाले माइक्रोकंप्यूटर के लिए अपना पहला वर्ड प्रोसेसिंग प्रोग्राम लांच किया था, तब इसका नाम मल्टी टूल वर्ड (Multi Tool Word) था। इसी वर्ष इसे आईबीएम के डिस्क ऑपरेटिंगसिस्टम (DOS) आधारित कंप्यूटरों में भी शामिल किया गया। कालान्तर में वर्ड के वर्जन लांच होते गए। वर्ष 1989 में एमएस-वर्ड को विंडोज ऑपरेटिंगसिस्टम में शामिल किया गया और वर्ष 1990 में यह एमएस-ऑफिस सुइट का अभिन्न अंग बन गया।

वर्ष 2010 में एमएस-वर्ड का नवीनतम संस्करण (Version) लांच किया गया है। हालांकि, एमएस-वर्ड 2007 सर्वाधिक प्रचलित एमएस-वर्ड संस्करण है। अधिकतर पर्सनल कंप्यूटरों पर एमएस-ऑफिस सुइट में वर्ड के इसी संस्करण का उपयोग किया जाता है। इकाई के इस भाग में हम भी एमएस-वर्ड के बारे में ही विस्तार से अध्ययन करेंगे। यहां यह उल्लेखनीय है कि माइक्रोसॉफ्ट विंडोज ऑपरेटिंगसिस्टम में वर्डपैड (Wordpad) नाम से एक स्वतंत्र वर्ड प्रोसेसर भी उपलब्ध रहता है , लेकिन वर्डपैड में उपयोगकर्ता को मिलने वाली सुविधाएं काफी सीमित रहती हैं। एमएस -वर्ड में वर्डपैड के मुकाबले कहीं अधिक साधन (Tools) उपलब्ध रहते हैं , जिससे उपयोगकर्ता के लिए अपने दस्तावेजों को तैयार करने और इन्हें बेहतर साज-सज्जा के साथ आकर्षक स्वरूप देना अधिक सरलीकृत और सुगम हो जाता है। इस लिहाज से हम यह भी मान सकते हैं कि एमएस-वर्ड वर्ड प्रोसेसिंग प्रोग्रामों के लिहाज से वर्डपैड की अगली और परिष्कृत कड़ी है।

- **एमएस-वर्ड के साधन (Tools of MS-Word)**

उपरोक्त चित्र में एमएस -वर्ड प्रोग्राम के प्रमुख हिस्से दर्शाए गए हैं , जिन्हें हम इस प्रोग्राम के साधन (Tools)



मान सकते हैं। यहां साधन का तात्पर्य प्रोग्राम में मौजूद उन सुविधाओं से है , जो एमएस-वर्ड को उपयोगकर्ता के प्रयोग के लिए सरल बनाते हैं और अभीष्ट परिणाम उपलब्ध कराने में मददगार होते हैं। इन सभी साधनों को हम निम्नवत क्रमवार विस्तार से जान लेते हैं-

- **टाइटल बार (Title Bar)-** जैसा कि नाम से ही स्पष्ट हो रहा है , प्रोग्राम के इस हिस्से में शीर्षक यानी टाइटल नजर आता है। उपयोगकर्ता जो भी फाइल , दस्तावेज बना रहा है यदि उसने इसे कुछ नाम दिया है तो वह नाम टाइटल बार में नजर आता है। यदि कोई नाम नहीं दिया है तो टाइटल बार में अनटाइटल्ड डॉक्यूमेंट (Untitled Document) लिखा दिखता है।
- **ऑफिस बटन (Office Button)-** एमएस-वर्ड के पुराने वर्जनों में इसके स्थान पर फाइल मेन्यू रहता था। लेकिन वर्ड 2007 में इसे बदलकर ऑफिस बटन कर दिया गया है। टाइटल बार के सबसे उपरी हिस्से में बायीं ओर विंडोज के लोगो वाला गोल बटन नजर आता है , वही ऑफिस बटन है। इस बटन पर क्लिक करते ही कई विकल्प (Option) सामने खुल जाते हैं। इनमें न्यू (New) बटन पर क्लिक करते ही नयी फाइल खुल जाती है , जिस पर उपयोगकर्ता अपना नया दस्तावेज तैयार कर सकता है। ओपन (Open) पर क्लिक करने से कंप्यूटर पर पहले से सुरक्षित दस्तावेज या फाइल खोली जा सकती है। सेव (Save) पर क्लिक करने से उपयोगकर्ता अपनी फाइल को सुरक्षित कर सकता है। सेव एज (Save as) उपयोगकर्ता को यह सुविधा प्रदान करता है कि वह अपने दस्तावेज को वर्ड डॉक्यूमेंट फाइल या अन्य किसी फॉरमेट में सुरक्षित कर सके। प्रिंट (Print) बटन पर क्लिक करने से उपयोगकर्ता अपने दस्तावेज का प्रिंट हासिल कर सकता है। इसका अगला बटन है प्रीपेयर (Prepare) की मदद से उपयोगकर्ता अपने दस्तावेज में कई ऐसे फीचर जोड़ सकता है , जो दस्तावेज को किसी अन्य उपयोगकर्ता के प्रयोग करने की स्थिति में इसे सुरक्षित रखें और अपरिवर्तनीय बना सकें। सेंड (Send) बटन दस्तावेज को ई-मेल के जरिये किसी दूसरे कंप्यूटर या उपयोगकर्ता तक भेजने की सुविधा देता है। पब्लिश (Publish) की मदद से उपयोगकर्ता अपने दस्तावेज को ब्लॉग बना सकता है या माइक्रोसॉफ्ट के डाटा मैनेजमेंट सर्वर में सुरक्षित रख सकता है। आखिरी बटन क्लोज (Close) है, जिसका अर्थ है बंद करना, यानी इसकी मदद से उपयोगकर्ता अपनी फाइल बंद कर सकता है।
- **क्विक एक्सेस टूलबार (Quick Access Toolbar)-** ऑफिस बटन के ठीक उपर दायीं ओर क्विक एक्सेस टूलबार होता है। इसमें वे कमांड (Commands) शामिल हैं , जिनका इस्तेमाल उपयोगकर्ता एमएस-वर्ड पर काम करते वक्त बार -बार करता है। ऐसे में उसे लंबी प्रक्रिया से गुजरने के बजाय सीधे इन पर क्लिक कर काम करने की सुविधा मिलती है। इनमें सेव (Save) अनडू (Undo) रिपीट (Repeat), ओपन (Open) जैसी कमांड शामिल हैं। उपयोगकर्ता अपनी जरूरत के हिसाब से टूलबार में कमांड जोड़ या हटा सकता है।
- **रिबन (Ribbon)-** एमएस-वर्ड में अलग-अलग सात टैब (Tab) हैं। हर टैब एक खास तरह के कमांड का समूह है और ये सातों टैब जहां अवस्थित रहती हैं, उस हिस्से को रिबन कहा जाता है। यह टाइटल बार से नीचे अगली लंबी पट्टी होती है। ये सात टैब हैं होम , इन्सर्ट, पेज लेआउट, रेफरेंसेज, मेलिंग्स, रिव्यू और व्यू। आइए हम संक्षेप में इन सातों टैब के बारे में जानते हैं-

1.होम टैब (Home Tab)- होम टैब की मदद से मुख्यतः पांच कार्य किए जा सकते हैं। यह भी मान सकते हैं कि किसी उपयोगकर्ता के लिए अपनी डॉक्यूमेंट फाइल तैयार करते वक्त सबसे अधिक जिन कमांड की जरूरत होती है। इनमें क्लिपबोर्ड , फॉन्ट , पैराग्राफ , स्टाइल और एडिटिंग शामिल हैं। इन सभी कमांडों की मदद से उपयोगकर्ता इच्छानुसार काम कर सकता है।

2. इन्सर्ट टैब (Insert)- इन्सर्ट का हिन्दी अर्थ है जोड़ना या घुसाना। उपयोगकर्ता अपनी फाइल में जब भी कोई ग्राफिक, चित्र, कोई खास चिह्न, टेबल आदि जोड़ना चाहता है तो इन्सर्ट टैब मददगार साबित होती है। इस टैब के भी सात हिस्से हैं , पेजेस (Pages), टेबल्स (Tables), इलस्ट्रेशन (Illustrations), लिंक्स (Links), हेडर

एंड फुटर (Header and Footer) टेक्स्ट (Text) और सिंबल्स (Symbals) पेजेस की मदद से उपयोगकर्ता मनचाहा पेज इस्तेमाल कर सकता है , जिस पर वह अपनी फाइल तैयार करना चाहता है। टेबल की मदद से उपयोगकर्ता सारिणी तैयार कर सकता है, जो कई दस्तावेजों में चीजों को समझाने में खासी उपयोगी होती हैं। इलस्ट्रेशन टैब उपयोगकर्ता को पिक्चर यानी फोटो , क्लिपआर्ट यानी चित्र , लोगो लगाने की सुविधा देती हैं। शेप्स की मदद से उपयोगकर्ता कई तरह की आकृतियों का इस्तेमाल कर सकता है। स्मार्टआर्ट में कई खास तरह के ग्राफिक्स का समूह पहले से उपलब्ध रहता है , जिनमें से उपयोगकर्ता मनचाहा चुन सकता है। चार्ट की मदद से दस्तावेज में ग्राफ (Graph) बनाए जा सकते हैं जो तुलनात्मक अध्ययन में मददगार होते हैं। लिंक्स की मदद से उपयोगकर्ता अपने दस्तावेज को वेबपेज की तरह तैयार कर सकता है , बुकमार्क बना सकता है। हेडर एंड फुटर किसी दस्तावेज का हेडर, यानी शीर्षक या परिचय और फुटर यानी निष्कर्ष को विशेष तरह से लिखने की सुविधा देता है। इसी टैब में पेज नंबर का भी विकल्प मौजूद होता है , इसकी मदद से उपयोगकर्ता अपने दस्तावेज में पन्नों के नंबर तय कर सकता है। टेक्स्ट टैब से उपयोगकर्ता को अपने पाठ्य को आकर्षक बनाने में मदद मिलती है। सिंबल्स की मदद से कुछ ऐसे संकेत, चिह्न पाठ्य में जोड़े जा सकते हैं, जिन्हें सामान्य तौर पर लिखना या टाइप कर पाना संभव नहीं हो पाता। इसमें कई गणितीय संकेत भी शामिल हैं।

3. पेज लेआउट (Page Layout)- इस टैब में पांच कार्यसमूह होते हैं। थीम्स (Themes) पेज सेटअप (Page Setup), पेज बैकग्राउंड (Page Background), पैराग्राफ (Paragraph), अरेंज (Arrange) इन सभी की मदद से उपयोगकर्ता यह तय कर पाता है कि उसे अपना दस्तावेज किस तरह तैयार करना है। उसका स्वरूप कैसा होगा, पन्ने पर हाशिये (Margins) कितनी चौड़ाई के होंगे, पन्ने का आकार (Size) कितना होगा, कॉलम कैसे होंगे, पैराग्राफ के बीच दूरी कितनी रहेगी आदि।

4. रेफरेंसेज (References)- उपयोगकर्ता जब कोई दस्तावेज तैयार करता है तो कई बार पाठ्य के साथ कुछ अन्य टिप्पणियां जोड़ने की भी जरूरत महसूस होती है। उदाहरण के तौर पर विषयसूची , शीर्षक, संदर्भ, फुटनोट, एंडनोट आदि। इस टैब में उपलब्ध कमांड टेबल ऑफ कंटेन्ट्स (Table of Contents), फुटनोट्स, साइटेशन एंड बिबिलोग्राफी (Citations and Bibilography), कैप्शन (Caption) इन्डेक्स और टेबल ऑफ अथॉरिटीज (Table of Authorities) से यह सब कर पाना संभव होता है।

5. मेलिंग्स (Mailings)- इस टैब के पांच भाग हैं , क्रिएट (Create) स्टार्ट मेल मर्ज, राइट एंड इन्सर्ट फील्ड्स (Write & Insert Fields), प्रीव्यू रिजल्ट्स और फिनिश । यह टैब मुख्यतः तब इस्तेमाल में लाया जाता है , जब उपयोगकर्ता अपने दस्तावेज को ई -मेल के जरिये किसी दूसरे उपयोगकर्ता तक भेजना चाहता हो या वेब संदेश तैयार करना चाहता हो। हालांकि, इसके लिए वर्ड ऑनलाइन से जुड़ा होना आवश्यक है।

6. रिव्यू (Review)- जैसा कि नाम से ही स्पष्ट है , यह टैब उपयोगकर्ता को सुधार का अवसर देता है। इसकी छह सहायक टैब हैं, प्रूफिंग (Proofing), कमेंट्स, ट्रैकिंग, चेंजेस, कंपेयर और प्रोटेक्ट (Protect) इनकी मदद से उपयोगकर्ता अपने दस्तावेज में वर्तनी , व्याकरण की अशुद्धियां दूर कर सकता है , टिप्पणी जोड़ सकता है , दस्तावेज में एक ही बात बार-बार गलती से रिपीट हो रही हो या कॉपी-पेस्ट हो गई हो तो इसे चिह्नित किया जा सकता है। दस्तावेज में कोई सुधारात्मक परिवर्तन करना संभव होता है। यही नहीं , दस्तावेज को इस तरह सुरक्षित किया जा सकता है कि कोई अन्य उपयोगकर्ता इसमें बदलाव नहीं कर सके।

7. व्यू टैब (View Tab)- इस टैब के पांच सहायक टैब हैं , डॉक्यूमेंट व्यूज (Document Views), शो-हाइड जूम, विंडो, और मैक्रो। इनकी मदद से उपयोगकर्ता अपने दस्तावेज का प्रिंट लेआउट तय कर सकता है। पन्ने को जूम करके बड़ा या छोटा कर सकता है , ताकि पाठ्यक्षेत्र पर टाइप करने में आसानी हो। दस्तावेज की विंडो को दो टुकड़ों में बांट सकता है, ताकि पाठ्य के दो अलग-अलग भागों को एक ही समय पर पढ़ सके। मैक्रो की मदद से उन कमांड को एक समूह के रूप में व्यवस्थित किया जा सकता है जो दस्तावेज तैयार करने के दौरान उपयोगकर्ता कई बार इस्तेमाल करता है। इन कमांड को सामूहिक रूप देने के बाद एक बार मैक्रो पर क्लिक करने से सभी कमांड काम करती हैं।

8. अन्य टैब (Other Tab)- रिबन में कुछ टैब नजर नहीं आती हैं। जैसे डेवलपर टैब (Developer Tab) यह टैब मूलतः तकनीकी उपयोग में काम आती है। वे ही लोग इसे अधिक प्रयोग करते हैं , जो वर्ड के लिए एप्लीकेशन (Applications) तैयार करते हैं। सामान्य दैनिक जीवन में इस टैब का अधिक उपयोग नहीं होता है। इसके अलावा रिलीवेंट टैब (Relevant Tab) भी वर्ड का हिस्सा हैं। ये टैब वे हैं, जो वर्ड में कोई खास काम करने के दौरान रिबन पर स्वतः उभर आती हैं। मसलन , जब हम दस्तावेज में कोई चित्र जोड़ते हैं तो पिक्चर टूल्स (Picture Tools) खुदबखुद सामने आ जाते हैं।

- **रूलर (Ruler)-** वर्ड में काम करते वक्त हमें दो रूलर नजर आते हैं। दरअसल, यह वर्ड का मापक साधन है। इसकी मदद से उपयोगकर्ता यह तय कर पाता है कि दस्तावेज के पन्नों पर हाशिये किस तरह व्यवस्थित होंगे। व्यू टैब में जाकर शो या हाइड से रूरल को हटाया भी जा सकता है।
- **पाठ्यक्षेत्र (Text Area)-** हम जानते हैं कि पाठ्यक्षेत्र वर्ड का वह हिस्सा है, जहां उपयोगकर्ता टाइपिंग करता है। सामान्य शब्दों में इसे वर्ड दस्तावेज का एक पन्ना भी मान सकते हैं।
- **कर्सर (Cursor)-** कर्सर पाठ्यक्षेत्र में एक खड़ी लकीर (।) की तरह दिखता है। उपयोगकर्ता पाठ्यक्षेत्र में जो भी लिखता यानी टाइप करता है, वह इस कर्सर से ही प्रारंभ होता है। पाठ्यसामग्री के किसी भाग में यदि उपयोगकर्ता को कोई शब्द बदलना हो तो उसके लिए कर्सर को संबंधित शब्द पर ले जाना अनिवार्य होता है , तभी टाइपिंग संभव हो पाती है। इसी तरह पाठ्यक्षेत्र में एक माउस प्वाइंटर (Mouse Pointer) भी नजर आता है जो दिखने में अंग्रेजी अक्षर कैपिटल आई प् की तरह होता है। इसे स्क्रीन पर माउस की मदद से कहीं भी ले जाया जा सकता है। जहां पर भी इसे माउस के क्लिक से छोड़ा जाता है, कर्सर वहीं आ जाता है।
- **स्टेटस बार (Status Bar)-** यह एमएस-वर्ड के सबसे नीचे की पट्टी है। इसमें जूम स्लाइड के अलावा कुछ अन्य क्विक कमांड मौजूद होती हैं। इसके अलावा बार के बायें हिस्से में दस्तावेज के पन्नों और इसमें टाइप किए गए कुल शब्दों की संख्या दर्शाई जाती है।

15.6 एमएस एक्सेल (MS Excel)

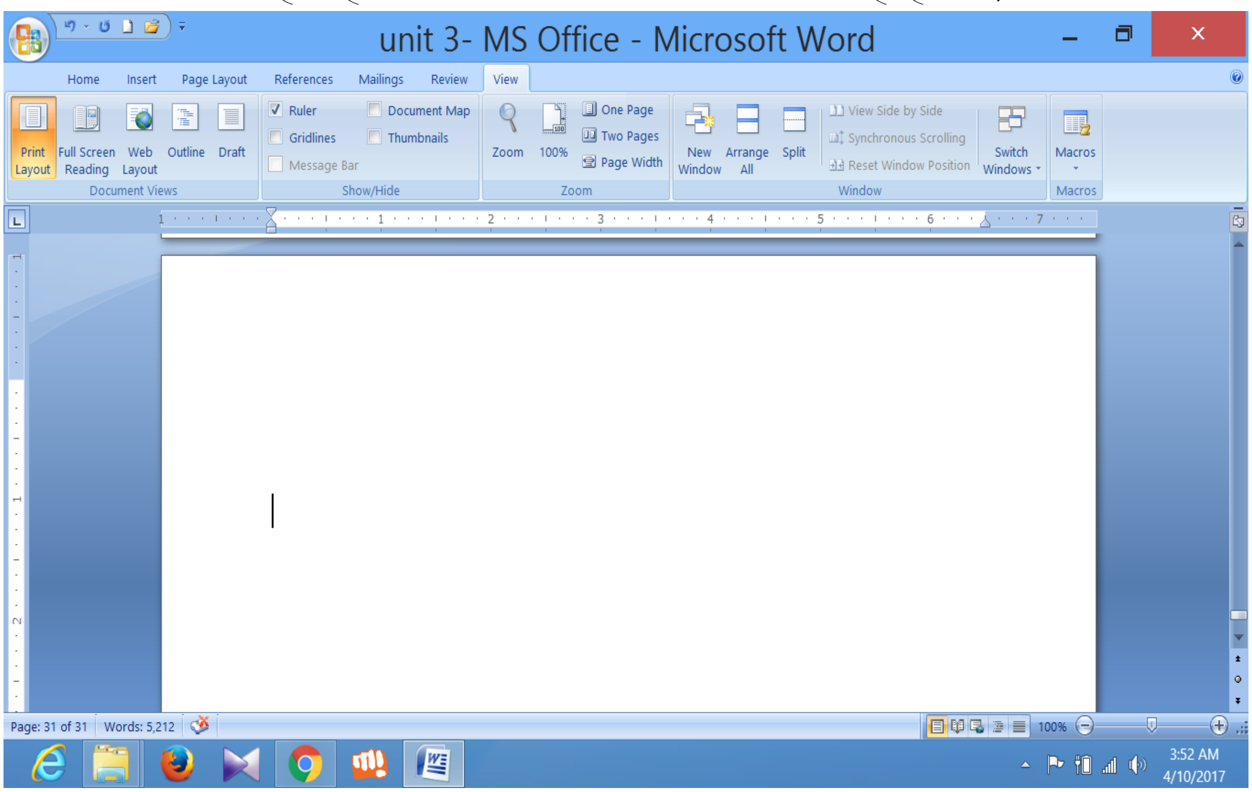
माइक्रोसॉफ्ट एक्सेल मूलतः एक स्प्रेडशीट (Spreadsheet) प्रोग्राम है। स्प्रेडशीट दरअसल डाटा या सूचनाओं का सारिणी रूप (Tabular Form) प्रस्तुतीकरण है , जिसमें डाटा या आंकड़ों को सारिणी के छोटे हिस्सों, जिन्हें सेल (Cells) कहा जाता है, में दर्ज किया जाता है। यह डाटा अंकीय भी हो सकता है और शब्दीय (Text) के रूप में भी। स्प्रेडशीट कुछ विशेष सूत्रों पर काम करती है , जिसके तहत अलग -अलग सेल में दर्ज आंकड़ों या डाटा के आधार पर पूरा परिणाम प्राप्त किया जा सकता है। ये फॉर्मूला स्वतः (Automatically) काम करते हैं, यानी उपयोगकर्ता को स्प्रेडशीट पर तय सेल में आंकड़े सिर्फ भरने होते हैं और फॉर्मूला की मदद से प्रोग्राम खुद ही नतीजा उपलब्ध करा देता है।

माइक्रोसॉफ्ट ने वर्ष 1987 में पहली बार विंडोज के लिए एमएस.एक्सेल को रिलीज किया था। 1990 में एमएस.ऑफिस सुइट के लांच होने पर इसे सुइट में शामिल कर लिया गया। विंडोज के अलावा यह प्रोग्राम मैकए आईफोन और अब एंड्रॉयड ऑपरेटिंगसिस्टम पर भी उपलब्ध है। संक्षिप्त इतिहास की चर्चा करें तो माइक्रोसॉफ्ट ने वर्ष 1982 में अपना पहला स्प्रेडशीट प्रोग्राम मल्टीप्लान (Multiplan) नाम से लांच किया था। लेकिन आईबीएम के स्प्रेडशीट प्रोग्राम लोटस 1 -2-3 (Lotus 1-2-3) के आगे यह टिक नहीं सका। ऐसे में माइक्रोसॉफ्ट ने वर्ष 1985 में मैक ऑपरेटिंगसिस्टम के लिए एमएस-एक्सेल नाम से प्रोग्राम को लांच किया।

- **एमएस-एक्सेल के साधन (Tools of MS-Excel)**

एमएस-एक्सेल की विंडो काफी हद तक एमएस-वर्ड की तरह ही नजर आती है। टाइटल बारए स्टेटस बार इसमें वर्ड के ही समान होता हैए अंतर सिर्फ रिबन में रहता हैए जहां पांच टैब होम , इन्सर्ट मल्टीप्लान, पेज लेआउट, फॉर्मूलाज, डाटा, रिव्यू और व्यू होते हैं। इनमें होम , इन्सर्ट, व्यू, पेज लेआउट और रिव्यू टैब लगभग वही हैं , जिन्हें हम एमएस -वर्ड में जान चुके हैं। इन सभी में फीचर्स का बेहद मामूली अंतर है , जो उपयोग के दौरान समझ में आ जाता है, मसलन एक्सेल में इन्सर्ट टैब की सहायक टैब चार्ट में ग्राफ के कुछ नये फीचर सामने आते हैं। इसी तरह हर टैब में हल्का अंतर है। एक्सेल पर उपयोगकर्ता जब भी काम करता है तो स्क्रीन पर सामने आने वाले पाठ्यक्षेत्र को वर्कशीट (Worksheet) कहा जाता है। वर्कशीट कतारों (Row) और कॉलम (Columns) में बंटा होता है। दिलचस्प पहलू यह है कि एक्सेल में अधिकतम एक लाख 48 हजार 576 रो और 16 हजार 384 कॉलम हो सकते हैं।

हम जानते हैं कि वर्कशीट में क्षैतिज पंक्तियां और ऊध्वाकार कॉलम होते हैं। कॉलमों पर A से लेकर Z और फिर AA, AB, AC, AD,.... दर्ज होता है, पंक्तियों में 1,2,3,4,5....अंक और संख्याएं लिखी होती हैं। हर कॉलम और पंक्ति का छोटा हिस्सा सेल (Cell) होता है। हर सेल की पहचान इसके कॉलम में दर्ज अक्षर और पंक्ति में दर्ज संख्या से होती है। मसलन उपरोक्त चित्र में जो सेल नजर आ रही है , उसका नाम या उसकी



पहचान A_1 है, क्योंकि यह सेल कॉलम A की पहली पंक्ति पर स्थित है। इसी तरह डी15 (D15) का तात्पर्य यह होगा कि संबंधित सेल D कॉलम की 15वीं पंक्ति पर स्थित है। इस पहचान को संबंधित सेल का पता (Address) भी कहा जाता है। सेल में भरे जाने वाले टेक्स्ट या शब्दों को लेबल (Label) कहा जाता है। इसी तरह सेल में भरी जाने वाली संख्याओं को वैल्यू (Value) कहा जाता है। एमएस -वर्ड में जिस तरह हमें पाठ्यक्षेत्र में कर्सर मिलता था , एक्सेल की वर्कशीट में जोड़ (+) का एक मोटा निशान नजर आता है , जिसे सेल प्वाइंटर कहते हैं। किसी सेल पर काम करने के लिए इस सेल प्वाइंटर को संबंधित सेल पर ले जाना जरूरी होता है, इससे वह सेल एक्टिव हो जाती है। इसके बाद हम इसमें लेबल और वैल्यू भर सकते हैं। वैल्यू का अर्थ सिर्फ 0 से 9 तक के अंकों से नहीं है , बल्कि इसमें हम गुणा -भाग, जोड़-घटाने और अन्य निशान भी भर सकते हैं।

एक्सेल पर काम करते हुए उपयोगकर्ता को सेलों का एक आयताकार समूह चुनना पड़ता है। इसके बाद इसी समूह में सारी क्रियाएं की जाती हैं। एक्सेल के इस समूह को रेंज (Range) कहा जाता है।

- **फॉर्मूलाज (Formulas)-** एमएस-एक्सेल में फॉर्मूले का विशेष महत्व है। जब हम कोई गणना करना चाहते हैं, मसलन किसी कॉलम की कुछ सेल को जोड़ना , घटाना या गुणा करना चाहते हैं तो इसके लिए हमें फॉर्मूला बार में इसके लिए सूत्र डालना अनिवार्य होता है। उल्लेखनीय है कि एक्सेल में फॉर्मूले हमेशा = से शुरू होते हैं। अब उदाहरण के लिए मान लें कि सेल ई5 (E5) में हमें जो परिणाम चाहिए, वह ए5, बी5, सी5 का योग और इस योग में से डी5 का अंतर हो तो ई5 सेल पर क्लिक करने के बाद हम फॉर्मूला इस तरह भरेंगे, =A5+B5+C5-D5
- **फॉर्मूला ऑपरेटर (Formula Operator)-** एक्सेल में हम फॉर्मूला तैयार करने में जिन चिह्नों , टेक्स्ट का इस्तेमाल करते हैं, उन्हें ऑपरेटर कहा जाता है। ये ऑपरेटर निम्नवत हैं-
- **अंकगणितीय ऑपरेटर (Arithmatic Operators)-** +, -, ×, /, %, ^ हैं, जिनका अर्थ क्रमशः जोड़ , घटाना, गुणा, भाग, प्रतिशत और घात है। अब यदि हमें डी16 (D16) सेल में बी10 का 45 प्रतिशत मान जानना है तो डी16 सेल को क्लिक करने के बाद फॉर्मूला इस तरह लिखा जाएगा , =b10×45% कोष्ठकों के इस्तेमाल से जटिल गणनाएं करना भी संभव है। इस तरह के फॉर्मूले इस तरह लिख सकते हैं, =d8+(b5×6)-(c3×25).
- **तुलना ऑपरेटर (Comparison Operators)-** से दो मानों की तुलना करना संभव हो पाता है। ये ऑपरेटर इस प्रकार हैं- =, >>, >>=, <<, <<=, <> इनके अर्थ क्रमशः बराबर, बड़ा, बड़ा या बराबर, छोटा, छोटा या बराबर और बराबर नहीं हैं। इन चिह्नों का प्रयोग सामान्यतः तार्किक फंक्शन (Logical Functions) में किया जाता है।
- **टेक्स्ट ऑपरेटर (Text Operator)-** वह ऑपरेटर है, जो किन्हीं दो सेलों में लिखे शब्दों को जोड़ता है। एक्सेल में प्रयुक्त होने वाला एकमात्र टेक्स्ट ऑपरेटर है ,- इसका प्रयोग इस तरह होता है, मान लीजिए कि सेल a3 में books और सेल b6 में pens लिखा है और वर्कशीट के सेल c8 में हम books & pens साथ लेना चाहते हैं तो इसका फॉर्मूला =a3&b6 लिखा जाएगा।
- **सन्दर्भ ऑपरेटर (Reference Operators)-** हम जानते हैं कि एक्सेल पर काम करने के लिए हम जितनी रो और कॉलम का इस्तेमाल करने वाले हैं, उन्हें वर्कशीट पर पहले सेलेक्ट करके रेंज तय करनी होती है। अब इस रेंज को दर्शाने के लिए कोलोन चिह्न (:) का प्रयोग किया जाता है। मसलन यदि उपयोगकर्ता की रेंज a4 से f16 तक है तो इस रेंज को इस तरह प्रदर्शित किया जाएगा, a4:f16.
- **फॉर्मूलों का क्रम (Orders of Formulas)-** जिस तरह सामान्य गणित में किसी जटिल गणना का हल निकालने के लिए हम गणितीय चिह्नों को तय क्रम यानी सबसे पहले कोष्ठक, फिर गुणा, भाग.... करते हैं, उसी तरह एक्सेल में भी फॉर्मूला ऑपरेटर का गणना क्रम तय है, यह इस प्रकार है-

क्रम संख्या	चिह्न	आशय
1	:	रेंज सन्दर्भ
2	-	ऋणात्मक संख्या
3	%	प्रतिशत
4	^	घातांक
5	× या /	गुणा या भाग
6	+ या -	जोड़ या घटाना
7	&	पाठ्य का जोड़

8	=/<>/<=/>=	तुलना

- **नंबर फॉरमेट (Number Format)-** होम टैब में सहायक टैब है नंबर , इसकी मदद से हम नंबर यानी संख्याओं का फॉरमेट तय कर सकते हैं। उल्लेखनीय है कि एक्सेल में कोई संख्या सेल में किस तरह दिखाई देगी, यह सेल के फॉरमेट पर ही निर्भर करता है। इस सहायक टैब में कई तरह के फॉरमेट हैं , लेकिन सामान्य उपयोग में इनमें से मुख्यतः सात -आठ ही इस्तेमाल में आती हैं। जनरल का अर्थ है कि सेल में संख्या को किसी खास फॉरमेट में नहीं दिखाया जाना है , यानी एक्सेल का जो तय फॉरमेट है , उपयोगकर्ता उसे ही इस्तेमाल करना चाहता है। नंबर पर क्लिक करने के बाद दशमलव संख्याओं को सेल में टाइप करना संभव हो पाता है। करेंसी किसी संख्या के आगे मुद्रा का निशान लगाने के लिए यह कमांड उपयोग की जाती है। डेट की मदद से संख्या को तारीख के रूप में प्रदर्शित करना संभव हो पाता है। टाइम कमांड की मदद से संख्याओं को समय के रूप में सेल में दर्शाया जाता है। परसेंटेज यानी संख्या को प्रतिशत रूप में दिखाने के लिए उपयोगी कमांड।
- **फंक्शन (Function)-** एमएस-एक्से में फंक्शन वे सुविधाएं हैं , जिनकी मदद से जटिलतम गणनाएं करना भी आसान हो जाता है। जटिल गणनाओं के फॉर्मूले बनाने के लिए खास गणितीय व्याकरण (Mathematical Syntax) का इस्तेमाल करना होता है। इसमें ये फंक्शन काम आते हैं। एक्सेल में सैकड़ों फंक्शन उपलब्ध होती हैं , जिन्हें वित्तीय तारीख और समय , गणित एवं त्रिकोणमिती , सांख्यिकीय, संदर्भ, डाटाबेस, पाठ्य, तार्किक , सूचना, अभियांत्रिकी (Engineering) और घन (Cube). इन फंक्शन के उपयोग और महत्व के बारे में अधिक जानने के लिए एक्सेल हेल्प की मदद ली जा सकती है। हालांकि, इनका उपयोग सामान्य गणनाओं में बेहद कम किया जाता है।
- **डाटाबेस (Database)-** हम जानते हैं कि डाटा का व्यवस्थित समूह डाटा बेस कहलाता है। एक्सेल में रो और कॉलम में दर्ज आंकड़ों की सामूहिक वर्कशीट या रेंज को डाटाबेस कहा जा सकता है। एक्सेल में डाटाबेस की हर पंक्ति को रिकॉर्ड (Record) कहा जाता है। मसलन किसी वर्कशीट में यदि किसी कक्षा के 50 छात्रों के सात विषयों में प्राप्ताकों का विवरण दर्ज है तो a कॉलम की 1 से 50 तक पंक्तियों में छात्रों के नाम लिखे जाएंगे , अब मान लीजिए कि हमें a5 पर दर्ज छात्र के अंक देखने हैं तो पांच नंबर पंक्ति में b5, c5, d5, e5, f5, g5, h5 पर दर्ज विषयवार अंक संबंधित छात्र का रिकॉर्ड होगा।
- **फील्ड (Field)-** एक्सेल में हर कॉलम को फील्ड कहा जाता है। इस लिहाज से हर एकल सेल को भी फील्ड माना जा सकता है। हर फील्ड में पाठ्य , संख्या, तारीखें, फंक्शन और फॉर्मूले भरे जा सकते हैं। जब किसी फील्ड में कोई फंक्शन या फॉर्मूला भरा जाता है , तो यह फील्ड गणनाकृत फील्ड (Computed Field) कहा जाता है।

15.7 उपसंहार (Conclusion)

इकाई के अध्ययन के बाद हम यह जान पाने में सक्षम रहे हैं कि कंप्यूटर पर मानव जीवन के दैनन्दिन कार्यों को सरल बनाने के लिए कौन-कौन से प्रोग्राम उपलब्ध हैं। एमएस-ऑफिस सुइट किस तरह काम करता है और इसके प्रोग्रामों और उनमें उपलब्ध साधनों की मदद से किस तरह हम डॉक्यूमेंट तैयार करने से लेकर गणितीय हिसाब-किताब भी आसानी से कर सकते हैं। साथ ही ऑफिस सुइट की मदद से हम अपने प्रस्तुतीकरण को बेहतर बना सकते हैं। चूंकि, यह प्रायोगिक विषय है, लिहाजा कंप्यूटर पर इन साधनों के उपयोग से इसे और बेहतर समझा जा सकता है।

कुछ महत्वपूर्ण तथ्य (Some Important Facts)-

हम माइक्रोसॉफ्ट ऑफिस सुइट के बारे में विस्तार से जान चुके हैं। अब यह जानना भी आवश्यक हो जाता है कि उपयोगकर्ता जब जिस प्रोग्राम में काम करता है , उसके अनुरूप जो भी दस्तावेज वह बनाता है, उसे

विशेष नाम से सुरक्षित करता है। किसी फाइल या दस्तावेज का नाम दो हिस्सों में बंटा होता है। पहला तो वह नाम, जो उपयोगकर्ता संबंधित दस्तावेज को देता है और दूसरा एक्सटेंशन (Extension), एक्सटेंशन दरअसल, इस बात का परिचायक है कि कोई दस्तावेज माइक्रोसॉफ्ट ऑफिस सुइट के किस प्रोग्राम को इस्तेमाल करके बनाया गया है। निम्न सारिणी से जानेंगे कि किस प्रोग्राम का दस्तावेज किस एक्सटेंशन से सेव किया जाता है-

फाइल	एक्सटेंशन	प्रकार	प्रोग्राम
xyz.txt	.txt	टेक्स्टफाइल	नोटपैड
abc.rtf	.rtf	टेक्स्टफाइल	वर्डपैड
puneet.jpg	.jpg	फोटो	पेण्ट
uou.doc	.doc	टेक्स्टफाइल	एमएस-वर्ड
123.xls	.xls	स्प्रेडशीट	एमएस-एक्सेल
uou.ppt	.ppt	प्रजेंटेशन	एमएस-पॉवरप्वाइंट

15.8 अभ्यास प्रश्न (Practice Question)

1. इनमें से कौन माइक्रोसॉफ्ट ऑफिस सुइट का हिस्सा नहीं है-
   a) वर्ड
   b) एक्सेल
   c) मीडिया प्लेयर
   d) आउटलुक
2. विंडोज ने मीडिया प्लेयर लांच किया था-
   a) 1991 में
   b) 2001 में
   c) 1985 में
   d) इनमें से कोई नहीं
3. उपयोगकर्ता को वेबसाइट बनाने की सुविधा इनमें से कौन सा प्रोग्राम प्रदान करता है-
   a) स्वे
   b) एक्सेल
   c) डेस्कटॉप पब्लिशिंग
   d) उपरोक्त में से सभी
4. माइक्रोसॉफ्ट ऑफिस 365 लांच किया गया-
   a) 1990 में
   b) 2015 में
   c) 2016 में
   d) 2003 में
5. ऑफिस सुइट से तैयार किसी टेक्स्ट फाइल का एक्सटेंशन निम्न में से कौन सा होता है-
   a) .txt
   b) .doc

c) .ppt
d) इनमें से कोई नहीं

6. माइक्रोसॉफ्ट स्वे प्रोग्राम इनमें से किस विंडोज ऑपरेटिंगसिस्टम के ऑफिस सुइट का हिस्सा है-
a) विंडोज 3.0
b) विंडोज XP
c) विंडोज 8
d) विंडोज 10

7. मैक ऑपरेटिंगसिस्टम में एमएस.ऑफिस का कौन सा प्रोग्राम सबसे पहले जारी किया गया था-
a) वर्ड
b) एक्सेल
c) पॉवरप्वाइंट
d) उपरोक्त में से सभी

8. वर्ष 2005 में आउटलुक एक्सप्रेस को इस प्रोग्राम से रिप्लेस कर दिया गया-
a) विंडोज मेल
b) आउटलुक
c) इंटरनेट एक्सप्लोरर
d) इनमें से कोई नहीं

9. इनमें से कौन मूलतः पर्सनल इंफॉर्मेशन मैनेजर की तरह काम करता है-
a) आउटलुक
b) आउटलुक एक्सप्रेस
c) उपरोक्त दोनों
d) इन दोनों में से कोई नहीं

10. माइक्रोसॉफ्ट एक्सेल निम्न में से किस पर आधारित प्रोग्राम है-
a) वर्ड प्रोसेसर
b) प्रजेंटेशन
c) स्प्रेडशीट
d) उपरोक्त सभी

11. बिल गेट्स ने माइक्रोसॉफ्ट ऑफिस सुइट लांच करने की घोषणा कब की थी-
a) 1988 में
b) 2000 में
c) 1997 में
d) इनमें से कोई नहीं

12. व्यावसायिक प्रिंटिंग में इस्तेमाल किया जाने वाला माइक्रोसॉफ्ट प्रोग्राम है-
a) स्वे
b) वन नोट

c) डेस्कटॉप पब्लिशिंग
d) पॉवरप्वाइंट

13. एमएस.ऑफिस सुइट का नवीनतम वर्जन है-
a) एमएस.ऑफिस 5
b) एमएस.ऑफिस 10
c) एमएस.ऑफिस एक्सपी
d) एमएस.ऑफिस 16

14. इनमें से कौन सा प्रोग्राम एंड्रॉयडए आईफोन और विंडोज फोन पर भी इस्तेमाल किया जा सकता है-
a) वन नोट
b) वर्ड
c) एक्सेल
d) उपरोक्त सभी

15. इस प्रोग्राम के इस्तेमाल के लिए उपयोगकर्ता का माइक्रोसॉफ्ट पर एकाउंट होना आवश्यक है-
a) वर्ड
b) स्वे
c) एक्सेल
d) इनमें से कोई नही

16. किसी वर्ड दस्तावेज में कितने पन्ने और कितने शब्द हैंए यह कहां देखा जा सकता है-
a) टाइटल बार में
b) इन्सर्ट टैब में
c) स्टेटस बार में
d) इनमें से कोई नहीं

17. एक्सले में प्रयोग किया जाने वाला झ्झ किस तरह का ऑपरेटर है-
a) तुलना ऑपरेटर
b) अंकगणितीय ऑपरेटर
c) पाठ्य ऑपरेटर
d) उपरोक्त में से सभी

18. एमएस.एक्सेल में कॉलम को यह भी कहा जाता है-
a) फंक्शन
b) रिकॉड्र
c) फील्ड
d) इनमें से कोई नहीं

19. एक्सेल में कॉलमों की अधिकतम संख्या है-
a) 64
b) 1,048,576

c) 256
d) 16384

20. एमएस.एक्सेल में उपयोगी पाठ्य ऑपरेटर है-

a) &
b) =
c) ×
d) उपरोक्त में से सभी

15.9 निबंधात्मक प्रश्न (Essay Type Question)

1. कंप्यूटर पर उपयोगकर्ता की मदद के लिए मौजूद कुछ प्रमुख साधनों यानी टूल्स के बारे में विस्तार से जानकारी दें।
2. माइक्रोसॉफ्ट ऑफिस और इसके घटकों के बारे में बताएं।
3. माइक्रोसॉफ्ट ईमेल, स्वे और वन नोट किस तरह के एमएस..ऑफिस टूल हैं। इनका इस्तेमाल एमएस-ऑफिस के सामान्य टूल से किस तरह अलग है। इनके क्या लाभ हैं।
4. एमएस.वर्ड क्या है , यह किस तरह काम करता है। एमएस -वर्ड में कोई नया दस्तावेज बनाने के लिए उपयोगकर्ता किन साधनों की मदद लेता हैए इनके बारे में विस्तार से बताएं।
5. एमएस-एक्सेल क्या है। यह किस तरह काम करता है , विस्तार से बताएं। एक्सेल का मानव जीवन में क्या उपयोग है।

इकाई 16 एसपीएसएस
(Statistical Package for Social Sciences)

16.1 प्रस्तावना (Introduction)
16.2 उद्देश्य (Objectives)
16.3 एसपीएसएस में सांख्किीय विधियां (Statistical Methods in SPSS)
16.4 एसपीएसएस में प्रयुक्त प्रोग्राम (Programs used in SPSS)
16.5 डाटा एन्ट्री (Data Entry)
16.6 सारांश (Summary)
16.7 शब्दावली (Glossary)
16.8 स्वमूल्यांकन हेतु प्रश्न (Self Assessment Question)
16.9 संदर्भ ग्रन्थ सूची (Bibliography)
16.10 निबन्धात्मक प्रश्न (Essay Type Question)

16.1 प्रस्तावना (Introduction)

SPSS एक साफ्टवेयर पैकेज है जिसे सामाजिक विज्ञान के शोधों के लिए प्रयुक्त किया जाता है। हालांकि अब यह बाजार अनुसंधानों, स्वास्थ्य शोध, कंपनियों द्वारा किये जाने वाले सर्वेक्षणों, सरकारों, शैक्षिक अनुसंधानों तथा बाजार विश्लेषण के लिए भी किया जाता है।

नील बेट तथा लूंट द्वारा 1990 में (SPSS) का मौलिक मैनुअल जो समाज शास्त्र की सबसे प्रभावशाली पुस्तकों में से एक मानी जाती है के द्वारा साधारण शोधों का भी अच्छी तरह सांख्यिकीय विश्लेषण किया जाता है। इसके द्वारा न केवल सांख्यिकीय विश्लेषण किया जाता है बल्कि इसके द्वारा आंकड़ों का व्यवस्थापन (Management) केस चुनाव , फाइल , रीशंपिंग आदि से किया जाता है तथा आंकड़ो का लेखीकरण (Documentation) भी किया जाता है।

SPSS के द्वारा जहाँ एक तरफ आंकड़ो का साधारण प्रतिशत ज्ञात कर सकते हैं वहीं दूसरी ओर जटिल से जटिल सांख्यिकीय विश्लेषण भी कर सकते हैं।

16.2 उद्देश्य (Objectives)

इस इकाई का अध्ययन करने के बाद आप-

- ✓ एसपीएसएस के सॉफ्टवेयर के बारे में जानेंगे।
- ✓ डाटा इन्ट्री की विधि के बारे में जान पाएंगे।
- ✓ एसपीएसएस में प्रयुक्त सांख्यिकी विधियों के बारे में जानेंगे।

16.3 एसपीएसएस में सांख्किीय विधियां (Statistical Methods in SPSS)

SPSS में सभी प्राथमिक सांख्यिकीय परीक्षण तथा Multivariate Analysis सम्मिलित होते है जैसे-

1. t-tests
2. Chi-Square tests
3. ANOVA
4. Correlation and other Associations measures
5. Regression
6. Non Parametric Tests
7. Factor Analysis
8. Cluster Analysis

16.4 एसपीएसएस में प्रयुक्त प्रोग्राम (Programs used in SPSS)

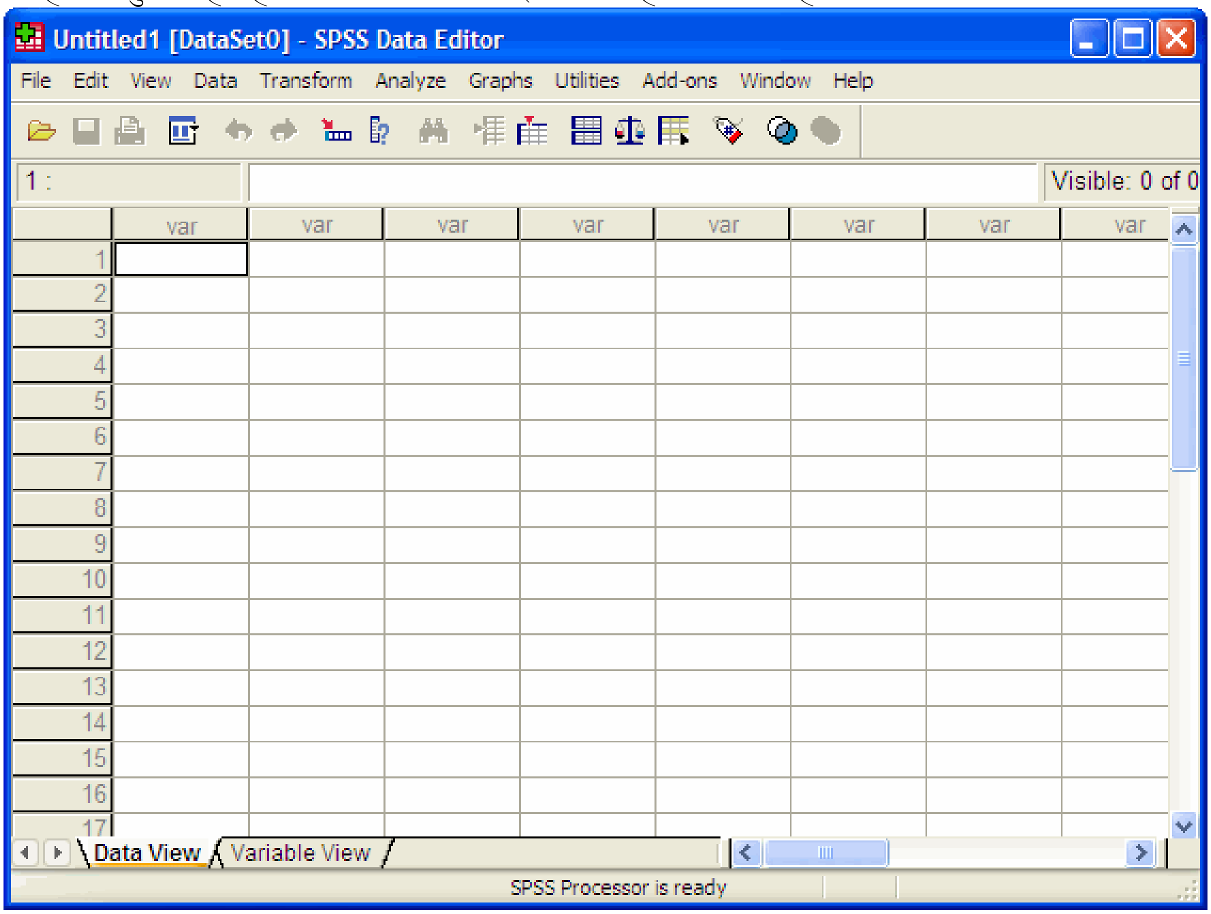
SPSS साफ्टवेयर खोलने का तरीका- हम किस तरह के कम्प्यूटर पर कार्य कर रहे हैं इसको देखते हुये दो तरह से खोला जा सकता है-

1. अगर डेस्कटॉप पर SPSS का शॉटकट हो तो करसर उसपे रखकर बांयी तरफ माउस पर डबल क्लिक करें।
2. माउस को बांयी तरफ से स्कीन के स्टार्ट बटन पर क्लिक करें , फिर कर्सर को All Programs पर ले जायं और माउस पर बायीं ओर क्लिक करें। फिर SPSS 12 की विंडो पर माउस बायीं तरफ से क्लिक करें। SPSS 12 का प्रयोग उदाहरण के लिए है।

इन दोंनो में से किसी भी प्रकार से करने पर SPSS का लेआउट खुलता है व डाटा एडिटर बिंडो में जो नया स्थान होता है जो कि स्कीन की बायी तरफ नीचे से चयनित किया जा सकता है।

SPSS Data Editor Window

SPSS Data Editor Window, SPSS का मुख्य विडो है। यही एक ऐसा विडो है जो SPSS Run कराने पर हमेशा खुलता है यह बाये कोने में लाल आइकॉन से पहचाना जाता है।



फाइल (File)- फाइल में वे सारे विकल्प होते है जो अन्य प्रोग्राम में होते है जैसे ओपन , सेव, एक्जीट इसमें पुरानी फाइल खोली जाती है या कई प्रकार की नयी फाइलों को खोला जाता है।

एडिट (Edit) में कट, कापी, पेस्ट का कमांड होता है इसके द्वारा आंकड़ा व परिणामों के प्रदर्शन के विभिन्न विकल्प मिलते हैं।

ऑप्शन (Option) पर क्लिक कीजिये और आपको बायीं तरफ डायलाग बाक्स दिखेगा। इसका उपयोग आंकड़ो को प्रारूपित करने, परिणाम तथा चार्ट आदि के लिए प्रयुक्त किये जाते हैं।

व्यू (View) के द्वारा फोंटसाइज, ग्रिड लाइन को जोड़ना या घटाना या मूल आंकड़ो को प्रदर्शित करना या न करने या आंकड़ों का लेबल प्रदर्शित किया जाता है।

Data के द्वारा निश्चित केसों को चुनने करने तथा विशिष्ट चरों के आधार पर आंकड़ो को चयनित किया जाता है।

ट्रांसफार्म (Transform) में प्रस्तुत चरों को परिवर्तित करने के लिए कई विकल्प होते हैं। उदाहरण के लि ए, सतत चरों को श्रेणीगत चरों में परिवर्तित किया जा सकता है। प्राप्तांकों को श्रेणियों में परिवर्तित किये जा सकते हैं।

एनेलाइज (Analyze) के द्वारा सांख्यिकी विश्लेषण का कमांड होता है तथा इसके द्वारा विश्लेषित सांख्यिकी का प्रयोग होता है।

ग्राफ (Graph) में बाक्स प्लाट, लाईन ग्राफ, बार चार्ट जैसे कमांड होते हैं।

यूटिलिटी (Utilities) में सभी चरों की लिस्ट बतायी जाती है इसके द्वारा लेवल , वैल्यू, डाटा की लोकेशन तथा प्रकार का प्रयोग होता है।

(Addons) एडआन्स एक प्रोग्राम है जो SPSS पैकेज में जोड़ा जा सकता है।

विंडो (Windows) के द्वारा यह तय किया जाता है कि किस तरह के विंडो को हम देखना चाहते हैं जैसे - डाटा एडिटर, आउटपुट, यूसर या सिनटैक्स

Help में SPSS पैकेज के कई महत्वपूर्ण विकल्प होते हैं।

डाटा एडिटर के बॉये नीचे के कोने में दो टैब होते है- डाटा व्यू व वैरीबल व्यू। डाटा व्यू डाटा एडीटर में दो विंडो होते हैं। डेटा व्यू डिफाल्ट में दिखता है जिससे आंकड़ो को प्रविष्ट किया जाता है। आंकड़ो को डेटा व्यू स्प्रेडशीट में प्रविष्ट किया जाता है। अधिकतम विश्लेषण के लिए SPSS यह मान लेता है कि रो के द्वारा केसों एवं कॉलम के द्वारा चरों का प्रतिनिधित्व होता है।

डाटा व्यू, स्प्रेडशीट ब्यू व टाप डाउन मेन्यू द्वारा नियंत्रित होता है। इसके द्वारा सेल के फोट को बदला जा सकता है पंक्तियों को हटाया जा सकता है तथा वेल्यू लेबल को दर्शनीय बनाया जा सकता है। डाटा व्यू का प्रयोग तब करते है जब हम एसपीएसएस में डाटा की एण्टी करते है। इसके स्तम्भों को चर कहा जाता है । स्तंभ के शीर्ष पर चर का नाम लिखा जाता है। पंक्तियों को केस कहा जाता है। डाटा सेल में ही वैल्यू निहित होता है जिसमें मूल्य निर्धारण होता है ।

वैरीवल ब्यू के द्वारा चरों को परिभाषित किया जाता है। जैसे ही आंकड़ो को डेटा व्यू के अन्तर्गत कॉलम में प्रविष्ट किया जाता है। वैसे ही वैरीवल व्यू में चर कालम का डिफाल्ट नाम एक पंक्ति का रूप ले लेते हैं। वैरीवल व्यू में निम्न विशिष्टता पायी जाती है।

नाम (Name)- इसमें चुने हुये चरों का नाम होता है इसमें केवल आठ अल्फावेट तक के नाम आ सकते है। इसमें **अंडरस्कोर (Underscore)-** तो स्वीकार्य है परंतु हाइफन (-) तथा स्पेस स्वीकार्य नहीं है।

टाइप (Type)- इसमें आंकडो के प्रकार को रखा जाता है इसका एक डिफाल्ट सेल होता है।

चौड़ाई (Width)- इसके द्वारा वास्तविक आंकड़ो की प्रविष्ट आंकड़ो की प्रविष्ट आंकड़ो का फैलाव दिखाया जाता है इसमें आंकिक चरों की डिफाल्ट प्रविष्टि 8 है। तीसरे कॉलम में सेल को हाइलाइट करके चौड़ाई को बढ़ाया या घटाया जा सकता है। ऐसा केवल सेल में नये नंबर को टाइप करके भी प्राप्त किया जा सकता है।

दशमलव (Decimals)- प्रविष्ट आंकड़ो में दशमलव के दायी तरफ आंकड़ो का प्रदर्शन होता है लेकिन यह String data में नहीं प्रयुक्त होता है।

आंकड़ो का संग्रहण तथा आंकड़ो की पुर्नप्राप्ति (Storing and retrieving data files)- आंकड़ो का संग्रहण तथा आंकड़ो की पुर्नप्राप्ति मेन्यूबार में फाइल को सेलेक्ट करने के बाद उपलब्ध हुये टॉप डाउन के द्वारा की जा सकती है।

वैल्यू लेबल (Value Level)

वैल्यू के अन्तर्गत वैल्यू लेवल प्राप्त होता है जिसमें हमारे आकडों का मूल्य प्रदर्शित होता है। वैल्यू लेवल के द्वारा चर के खास मूल्यों को लेवल प्रदान किया जाता है। वैल्यू लेवल ज्यादातर नामित या वर्गीकृत चरों के लिए प्रयुक्त होता है-

1. हिन्दू, 2. मुसलमान, 3. ईसाइ, 4 नास्तिक, 5. अन्य

वैल्यू लेवल का एक अन्य महत्वपूर्ण उपयोग होता है चरों का समूहीकरण करना। जैसे , मान लीजिये हमें एल्कोहल के विभिन्न मात्रा लेने वाले प्रतिभागियों के प्रतिक्रिया समय में अन्तर देखना है। हम ग्रुप 1 वैल्यू

लेवल का इस्तेमाल उनके लिए कर सकते है जिन्होनें एल्कोहल का सेवन नहीं किया , ग्रुप 2 वैल्यू लेवल का इस्तेमाल उनके लिए कर सकते है , जिन्होंने 1 यूनिट एल्कोहल का प्रयोग किया था व ग्रुप 3 वैल्यू लेवल वाले समूह ने 2 यूनिट एल्कोहल का प्रयोग किया था। वैल्यू लेवल को एसपीएसएस में अंतर्निहित कर दिया जाता है, जिससे इन वैल्यू के मतलब का पता चल सके ।

मिसिंग वैल्यू (Missing Value)

कभी कभी हमारे पास आकडों का पूरा सेट उपलब्ध नहीं हो पाता। उदाहरण के लिए , कुछ प्रतिभागी अपने धर्म को नहीं बताते है या कुछ प्रतिभागियों से आकडें उपलब्ध नही हो पाते है । आकडों के इस अन्तर को मिसिंग वैल्यू कहतें है। जब हमारे पास मिसिंग वैल्यू होता है तो यह आवश्यक होता है कि एसपीएसएस को यह बताया जाये कि हमारे पास उस चर पर इस प्रतिभागी का वैध आकडा उपल्बध नहीं है। इसके लिए हम एक ऐसे वैल्यू को चूनते है जो उस चर के लिए सामान्यतः प्रयूक्त नही होते है। जैसे धर्म के लिए हम कोड 9 को ले सकते है एजब उत्तरदाता अपने धर्म को नहीं बताता है। इस प्रकार कोड 9 धर्म के लिए मिसिंग वैल्यू है। मिसिंग वैल्यू सभी चरों के लिए अलग अलग हो सकता है।

16.5 डाटा एन्ट्री (Data Entry)

जब (SPSS) विंडो को खोलते हैं तो एक डिफाल्टर डायलाग बाक्स खुलता है जिसके द्वारा कई विकल्प प्राप्त होते हैं। जब टाइप इन डाटा का चयन होता है तो एक खाली स्प्रेडशीट खुलती है जिसे डाटा एडीटर कहते हैं। स्कीन के ऊपर एक मेन्यू बार होता है तथा नीचे की ओर एक स्टेटस बार होता है। स्टेटस बार द्वारा यह पता चलता है कि कौन सी सुविधायें अभी सक्रिय हैं। सेशन की शुरूआत में साधारणतः यह कहता है SPSS processories ready SPSS के द्वारा एक टूलबार भी प्राप्त होता है जो सामान्य कार्यो को तेजी से आसानी से कर सकता है। प्रत्येक टूल के बारे में संक्षिप्त जानकारी टूलबार पर कर्सर ले जाकर प्राप्त किया जा सकता है।

जब हम नये डाटा सेट का निर्माण कर रहे होते है तो चरों के नाम व अन्य विशेषताओं से शुरूआत करते है फिर प्रत्येक स्वतन्त्र स्रोत के लिए प्रत्येक चर पर विशिष्ट वैल्यू को enter करते है। आकडों के स्वतन्त्र स्रोत के लिए एक पंक्ति तथा प्रत्येक विशेषता (वैरीबल) के लिए एक कॉलम होता है।

यह ध्यान रखना होता है कि एक प्रतिभागी का डेटा एक समय में इन्टर करना चाहिये। उदाहरण के लिए एक प्रतिभागी को लिंग , उम्र व किसी स्वतन्त्र चर पर उसके प्रांप्ताक मदजमत करते है, फिर इसी रो में दूसरे प्रतिभागी का कंजं इंटर करने के पश्चात एक बार पुनः उनका निरीक्षण किया जाना चाहिये ।

डेटा फाइल को सेव करना- इसके लिए मेन्यू आइटम में फाइल पर क्लिक करते है। Save का Option आता है। इसके लिए फाइल का नाम लिखते है फिर सेव करते है। इस फाइल का नाम का प्रयोग करते है ।

SPSS Variable Types-

SPSS वैरीबल टाइप एवं फारमेट का अध्धयन करने से कार्य और तेजी व विश्वसनीय तरीके से किया जाता है।

SPSS में दो वैरीबल टाइप होते है- 1. स्ट्रिंग, 2.न्यूमेरिक। न्यूमेरिक चरों में केवल अंक आते है। स्ट्रिग वैरीबल में अक्षर, संख्या व अन्य विशेषतायें आती है। न्यूमेरिक वैरीबल के साथ गणना की जा सकती है पर स्ट्रिग वैरीबल के साथ नहीं।

Syntax SPSS की तीन महत्वपूर्ण विडो में से दूसरा महत्वपूर्ण विडो है। यह वहां पर उपथित होता है जहां पर से हम फाइल को खोलने का कंमाड, एडिटिग का एरिजल्ट जनरेट का व फाइल सेव करने का कमांड देते है । सिन्टेकस के द्वारा जटिल से जटिल कमान्ड को सेव कर लिया जाता है। सिन्टेकस के द्वारा आकडा विश्लेषण (Data Analysis) का रिकार्ड रखा जाता है।

Output विडो वह विडो है जो सभी आउटपुट जिन्हें हमने विश्लेषण के दौरान उत्पन्न किया है को समाहित करता है इसके द्वारा मुख्यतः टेबल एवं चार्ट का प्रयोग किया जाता है। आउटपुट व्यूवर विडो तब अपने आप उत्पन्न हो जाता है जब हम परिणाम उत्पन्न करते है। यह बैगनी आइकॉन से प्रदर्शित किया जाता है। SPSS आउटपुट फाइल का उपयोग रिर्पोटिग के लिए नहीं किया जा सकता ।

डाटा एनालिसिस के लिए एनालिसिस में जाकर के जो सांख्यिकीय विधि प्रयुक्त करनी होती है उस पर क्लिक कर देते है। इसमें आवृति से लेकर कारक विश्लेषण अर्थात साधारण से लेकर जटिलतम सांख्यिकीय विधि आती है।

अपरिष्कृत आकडों को विश्लेषण करने से पहले कुछ एडिट करने की आवश्यकता होती है। इसमें नये चरों का निर्माण, आकडो का पुर्ननिर्माण किया जाता है ।

सभी चार्ट व तालिकायें SPSS द्वारा आसानी से प्राप्त की जा सकती है। इसके लिए डाटा एनालिसिस में ही जाना होता है लेकिन SPSS में एक कमी यह है कि इसके द्वारा उत्पन्न किया हुआ चार्ट देखने में अच्छा नहीं लगता है। इससे उबरने के लिए SPSS Graph Editor का प्रयोग करना होता है।

16.6 सारांश (Summary)

SPSS एक बहुत ही महत्वपूर्ण सॉफ्टवेयर है। इसके द्वारा हम बहुत बड़े आंकड़ों का सांख्यिकी विश्लेषण आसानी से कर सकते हैं। इस सॉफ्टवेयर का उपयोग वर्तमान जगत में बहुत तेजी से हो रहा है।

SPSS में विभिन्न तरह के एप्लीकेशन जिसमें डेटा बेस मैनेजमेन्ट तथा रिर्पोटिग , सांख्यिकी विश्लेषणव ग्राफिक सम्मिलित है ।

SPSS के द्वारा सभी स्रोतों से प्राप्त आकडों का संसाधन किया जा सकता है। इसमें डेटा एडीटर विडो , सिन्टेकस व आउटपुट विडो तीन महत्वपूर्ण विडो होती है।

डाटा एडीटर विडो में दो टैब होते है- डेटा व्यू एवैरीबल व्यू। डेटा व्यू व वैरीबल व्यू का प्रयोग डेटा एंट्री के लिए किया जाता है।

सिन्टेक्स विडो द्वारा सभी कमाडो का रिकार्ड रखा जाता है।

आउटपुट विडो द्वारा उत्पन्न परिणामों का संग्रहण किया जाता है।

इसके द्वारा आंकड़ों की कोडिंग की जाती है और उनका प्रविष्टन करके विभिन्न सांख्किीय विधियों द्वारा परिणाम पाया जा सकता है।

इसमें साधारण से लेकर जटिल सांख्किीय विधियों का प्रयोग किया जा सकता है।

16.7 शब्दावली (Glossary)

- **एसपीएसएस (SPSS)-** यह एक ऐसा सॉफ्टवेयर है जिसे सामाजिक विज्ञान के शोधों , बाजार अनुसंधानों, स्वास्थ्य शोधों , कम्पनियों द्वारा किये जाने वाले सर्वेक्षणों , सरकारों व शैक्षिक अनुसंधानों द्वारा किया जाता है व बाजार विश्लेषण के लिए भी इसका उपयोग होता है।
- **डाटा इन्ट्री (Data Entry)-** यह आंकड़ों के विश्लेषण के पहले की अत्यन्त महत्वपूर्ण प्रक्रिया है। इसमें आंकड़ों को आंकिक रूप से सॉफ्टवेयर में अंकित किया जाता है।

16.8 स्वमूल्यांकन हेतु प्रश्न (Self Assessment Question)

1.एसपीएसएस द्वारा न केवल आंकड़ों का सांख्किीय विश्लेषण बल्कि उसके द्वारा आंकड़ों का .... भी किया जाता है।

2. एडिट मे.......... का कमांड होता है।

3. एनेलाइज द्वारा .... का कमांड होता है।
4. वैरीबल व्यू में .... को परिभाषित किया जाता है।
उत्तर- 1. व्यवस्थापन 2. कट, कापी, पेस्ट, 3. सांख्किीय विश्लेषण 4. चरों

16.9 संदर्भ ग्रन्थ सूची (Bibliography)

SPSS for Psychologists Nicola Brace, Richard Remp and Rosemery Shelgen Palgrave

16.10 निबन्धात्मक प्रश्न (Essay Type Question)

1. एसपीएसएस में सॉफ्टवेयर का क्या उपयोग है। इसके द्वारा आंकड़ों का विश्लेषण कैसे होता है?
2. एसपीएसएस द्वारा आंकड़ों की एन्ट्री का क्या तरीका है?