2 आंकड़ों का प्रक्रमण

आप पिछले अध्याय में पढ़ चुके हैं कि आंकड़ों का संगठन तथा प्रस्तुतीकरण उन्हें बोधगम्य बनाता है। इससे आंकड़ों का प्रक्रमण सरल हो जाता हैं। आंकड़ों के विश्लेषण के लिए अनेक विधियों को उपयोग किया जाता है। इस अध्याय में आप निम्नलिखित सांख्यिकीय विधियाँ सीखेंगे :

1. केंद्रीय प्रवृत्ति के माप
2. प्रकीर्णन के माप
3. संबंध के माप

जहाँ केंद्रीय प्रवृत्ति के माप पर्यवेक्षणों के समूह का आदर्श प्रतिनिधिकारी मूल्य प्रस्तुत करते हैं, वहीं प्रकीर्णन के माप आंकड़ों की आंतरिक विषमताओं का ब्यौरा देते हैं, जो अक्सर केंद्रीय प्रवृत्ति के माप के संदर्भ में होते हैं। दूसरी ओर संबंध के माप दो या दो से अधिक घटनाओं जैसे वर्षा तथा बाढ़ की घटना अथवा उर्वरकों का उपभोग तथा फ़सलों की उपज के मध्य साहचर्य की गहनता प्रस्तुत करते हैं।

केंद्रीय प्रवृत्ति के माप

मापनीय विशेषताएँ जैसे वर्षा, ऊँचाई, जनसंख्या का घनत्व, उपलब्धियों के स्तर अथवा आयु वर्ग में विभिन्नताएँ पाई जाती हैं। यदि हमें उनको समझना है, तो हमें क्या करना होगा? उसके लिए हमें कदाचित एक मूल्य या मान की आवश्यकता होगी जो पर्यवेक्षणों के समूह का सर्वोत्तम प्रतिनिधित्व करता हो। यह एकल मान सामान्यत: वितरण के किसी भी छोर पर होने की बजाय उसके केंद्र के निकट स्थित होता है। वितरण का केंद्र ज्ञात करने वाली सांख्यिकीय विधियों को **केंद्रीय प्रवृत्ति के माप** के नाम से जाना जाता है। केंद्रीय प्रवृत्ति की द्योतक संख्या सारे आंकड़ों के समूह की प्रतिनिधि संख्या होती है क्योंकि यह उस बिंदु की प्रतीक होती है जिसके निकट इकाइयों के समूहन की प्रवृत्ति होती है।

केंद्रीय प्रवृत्ति के मापों को सांख्यिकीय औसत के नाम से भी जाना जाता है। केंद्रीय प्रवृत्ति के कई माप हैं जिनमें **माध्य, माध्यिका** तथा **बहुलक** सबसे महत्वपूर्ण हैं।

माध्य

माध्य वह मान है जो सभी मूल्यों के योग को कुल प्रेक्षणों की संख्या से विभाजित करने पर प्राप्त होता है।

माध्यिका

माध्यिका उस कोटि का मान होता है जो व्यवस्थित श्रेणी को दो बराबर संख्याओं में विभाजित करता है। यह मान वास्तविक मूल्यों से स्वतंत्र होता है। आंकड़ों को बढ़ते अथवा घटते क्रम में व्यवस्थित करना माध्यम की गणना में सबसे अधिक महत्वपूर्ण हैं। सम संख्याए होने पर दो मध्यस्थ कोटि मानों का औसत माध्यिका होगा।

बहुलक

किसी बिंदु या मान की अधिकतम पुनरावृत्ति अथवा आवृत्ति बहुलक होती है। आपने देखा होगा कि इनमें से प्रत्येक भिन्न-भिन्न प्रकार के आंकड़ों के समूह के लिए उपयुक्त एकल प्रतिनिधि संख्या निर्धारित करने की अलग विधि है।

माध्य

किसी चर के विभिन्न मूल्यों का साधारण अंकगणितीय औसत माध्य कहलाता है। अवर्गीकृत तथा वर्गीकृत आंकड़ों के लिए माध्य ज्ञात करने की विधियाँ निश्चित ही भिन्न हैं। वर्गीकृत व अवर्गीकृत दोनों प्रकार के आंकड़ों के लिए माध्य प्रत्यक्ष व अप्रत्यक्ष विधियों के द्वारा ज्ञात किया जा सकता है।

अवर्गीकृत आंकड़ों से माध्य की गणना

प्रत्यक्ष विधि

अवर्गीकृत आंकड़ों से प्रत्यक्ष विधि द्वारा माध्य की गणना करने के लिए पर्यवेक्षण के सभी मूल्यों को जोड़ कर घटनाओं/पदों की कुल संख्या से भाग देते हैं। इस प्रकार माध्य की गणना निम्नांकित सूत्र के उपयोग द्वारा की जाती है।

$$\overline{X} = \frac{\Sigma x}{N}$$

जिसमें

$\overline{X}$ = माध्य

Σ = मापों के सभी मूल्यों का योग

x = मापों की किसी श्रेणी में एक अपरिष्कृत समंक

Σx = मापों की किसी श्रेणी में एक अपरिष्कृत समंक

N = श्रेणी के पदों की संख्या

उदाहरण 2.1 : मध्य प्रदेश में मालवा पठार के विभिन्न ज़िलों की, *तालिका–2.1* में दी गई वर्षा के आधार पर उस क्षेत्र की माध्य वर्षा की गणना कीजिए।

सारणी 2.1 : माध्य वर्षा की गणना

मालवा के पठार के जिले	*सामान्य वर्षा (मि.मी. में)*	*अप्रत्यक्ष विधि*
	प्रत्यक्ष विधि x	***d= x-800****
इंदौर	979	179
देवास	1083	283
धार	833	33
रतलाम	896	96
उज्जैन	891	91
मंदसौर	825	25
शाजापुर	977	177
Σx and Σd	**6484**	**884**
$\Sigma \frac{x}{N}$ and $\Sigma \frac{d}{N}$	**926.29**	**126.29**

* जिसमें 800 कल्पित माध्य है;
d कल्पित माध्य से विचलन है।

तालिका 2.1 में दिए आंकड़ों के लिए माध्य की गणना निम्न विधि से की जाएगी–

$$\overline{X} = \frac{\sum x}{N}$$

$$= \frac{6,484}{7}$$

$$= 926.29$$

माध्य की गणना से यह समझा जा सकता है कि वर्षा के अपरिष्कृत आंकड़ों का सीधा योग कर लिया गया है तथा उस योग को कुल पदों की संख्या अर्थात् (ज़िलों की संख्या) से विभाजित किया गया है। अत: इसे **प्रत्यक्ष विधि** कहते हैं।

अप्रत्यक्ष विधि

श्रेणी में जहाँ प्रेक्षणों की संख्याएँ बहुत अधिक होती हैं, वहाँ सामान्यत: अप्रत्यक्ष विधि से माध्य की गणना की जाती है। इस विधि में एक स्थिरांक को सभी मूल्यों से घटाने पर प्रेक्षणों की संख्या विस्तार कम हो जाती है। उदाहरण के लिए जैसा *तालिका 2.1* में दर्शाया गया है, वर्षा के मान 800 से 1100 मिलीमीटर तक हैं। एक 'कल्पित माध्य' मानकर हम इन संख्याओं के विस्तार को कम कर सकते हैं। इस उदाहरण में हमने कल्पित माध्य 800 माना है। इस क्रिया को **'कूट पद्धति'** कहते हैं। इसके पश्चात् घटाए हुए मूल्यों के आधार पर माध्य की गणना कर ली जाती है (*तालिका–2.1* में स्तंभ–3)।

अप्रत्यक्ष विधि से माध्य की गणना निम्न सूत्र से की जाती है–

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

जिसमें,

A = घटाया हुआ स्थिरांक

$\sum d$ = स्थिरांक घटाए हुए मूल्यों का योग

N = उक्त श्रेणी में एकल प्रेक्षणों की संख्या

तालिका–2.1 में दिए गए आंकड़ों के लिए अप्रत्यक्ष विधि द्वारा माध्य की गणना निम्नविधि से की जा सकती है–

$$\overline{X} = 800 + \frac{884}{7}$$

$$= 800 + \frac{884}{7}$$

$$\overline{X} = 926.29 \text{ मि.मी.}$$

यहाँ यह ध्यान देने योग्य तथ्य है कि चाहे किसी भी विधि से माध्य की गणना की गई हो, उसका मान समान ही आता है।

वर्गीकृत आंकड़ों से माध्य की गणना

वर्गीकृत आंकड़ों से भी प्रत्यक्ष या अप्रत्यक्ष विधियों से माध्य की गणना की जाती है।

प्रत्यक्ष विधि

जब आवृत्ति वितरण के रूप में आँकड़े वर्गीकृत हों तो उसमें एकाकी मूल्य अपनी पहचान खो देते हैं। इन

सभी मूल्यों का प्रतिनिधित्व वर्ग अंतराल के मध्य बिंदुओं द्वारा होता है, जहाँ वे स्थित हैं। प्रत्यक्ष विधि से वर्गीकृत आंकड़ों के लिए माध्य की गणना करते समय प्रत्येक वर्ग के मध्य बिंदुओं से संबंधित आवृत्ति (f); को गुणा किया जाता है; fx (इसमें X मध्य बिंदु है) के सभी मानों को जोड़कर प्राप्त $\sum fx$ में पदों की संख्या (N) से भाग दिया जाता है। अत: निम्नलिखित सूत्र द्वारा माध्य ज्ञात किया जाता है–

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

जिसमें,

$\overline{X}$ = माध्य

f = आवृत्ति

x = वर्ग अंतराल के मध्य बिंदु

N = पदों की संख्या (इसको $\sum f$ भी कहा जाता है)

उदाहरण 2.2 : *तालिका–2.2* में दिए गए आंकड़ों के प्रयोग से फैक्ट्री में काम करने वालों की माध्य मजदूरी दर की गणना कीजिए

तालिका 2.2 : फैक्ट्री श्रमिकों की मजदूरी दर

मजदूरी (रु./दिन)	***श्रमिकों की संख्या (f)***
वर्ग	***f***
50-70	10
70-90	20
90-110	25
110-130	35
130-150	9

तालिका 2.3 : माध्य की गणना

वर्ग	*आवृत्ति (f)*	*मध्य-बिंदु (x)*	*fx*	*d=x–100*	*fd*	*U = (x–100)/20*	*fu*
50-70	10	60	600	-40	-400	-2	-20
70-90	20	80	1,600	-20	-400	-1	-20
90-110	**25**	**100**	2,500	**0**	0	0	0
110-130	35	120	4,200	20	700	1	35
130-150	9	140	1,260	40	360	2	18
$\sum fx$ तथा $\sum fx$	$\sum f$ =99		$\sum fx$ = 10,160		$\sum fd$ = 260		$\sum fu$ = 13

जिसमें, N = $\sum f$ = 99

तालिका–2.3 में वर्गीकृत आंकड़ों के लिए माध्य की गणना करने की विधि दी गई है। दिए हुए आवृत्ति वितरण में 99 मजदूरों को पारिश्रमिक दर के पाँच वर्गों में बाँटा गया है। इन वर्ग विस्तारों के मध्य बिंदु तृतीय स्तंभ में दिए गए हैं। माध्य ज्ञात करने के लिए प्रत्येक मध्य बिंदु (x) को उससे संबंधित आवृत्ति (f) से गुणा करके (fx) गुणनफल के योग को ($\sum fx$) पदों की संख्या (N) से विभाजित किया गया है। इस प्रकार माध्य की गणना निम्न सूत्र के द्वारा ज्ञात की जा सकती है।

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

$$= \frac{10,160}{99}$$

$$= 102.6$$

अप्रत्यक्ष विधि

वर्गीकृत आंकड़ों से अप्रत्यक्ष विधि द्वारा निम्न सूत्र से माध्य ज्ञात किया जा सकता है। इस विधि से माध्य की गणना के सिद्धांत वही हैं जो अवर्गीकृत आंकड़ों के लिए अप्रत्यक्ष विधि द्वारा माध्य की गणना में दिए गए थे। इसे निम्न प्रकार से अभिव्यक्त किया जाता है–

$$\overline{x} = A \pm \frac{\sum fd}{N}$$

जिसमें,

A = कल्पित माध्य वाले वर्ग का मध्य बिंदु
(*तालिका–2.3* में 90-110 कल्पित माध्य वाला वर्ग माना गया है, जिसका मध्य 100 है।)

f = आवृत्ति

d = कल्पित माध्य वाले वर्ग (A) से विचलन

N = कुल पदों की संख्या अथवा $\sum f$

i = वर्ग अंतराल (इस उदाहरण में यह 20 है)

तालिका–2.3 में अप्रत्यक्ष विधि द्वारा माध्य की गणना करने से संबंधित निम्नलिखित चरण स्पष्ट हैं–

(i) कल्पित माध्य 90–110 वाले वर्ग में माना गया है। कल्पित माध्य जहाँ तक संभव हो, वितरण श्रेणी के मध्य में माना जाता है। इस प्रक्रिया से गणना का परिमाण न्यूनतम होता है। *तालिका 2.3* में A (कल्पित माध्य) 100 है, जो कि 90–110 वाले वर्ग का मध्य बिंदु है।

(ii) पाँचवें स्तंभ (u) में प्रत्येक वर्ग के मध्य बिंदुओं का कल्पित माध्य वाले (90 – 110) के मध्य बिंदु से विचलन दिया गया है।

(iii) छठे स्तंभ में fd प्राप्त करने के लिए प्रत्येक आवृत्ति (f) को उससे संबंधित d के मान से गुणा किया गया है। तत्पश्चात् fd के धनात्मक व ऋणात्मक मानों को अलग–अलग जोड़कर उनका निरपेक्ष अंतर ($\sum fd$) ज्ञात कर लिया जाता है। यहाँ यह ध्यान देने योग्य है कि $\sum fd$ से संलग्न चिह्न को सूत्र में A, के बाद दिए गए चिह्न ± के स्थान पर उपयोग करते हुए माध्य की गणना निम्नानुसार की जाती है :

$$\overline{x} = A \pm \frac{\sum fd}{N}$$

$$= 100 + \frac{260}{99}$$

$$= 100 + 2.6$$

$$= 102.6$$

टिप्पणी : अप्रत्यक्ष विधि समान व असमान दोनों ही वर्ग अंतरालों वाले वितरणों के लिए प्रभावी होती है।

माध्यिका

माध्यिका **स्थितिक औसत** है। इसे "वितरण में ऐसे बिंदु जिसके दोनों ओर बराबर संख्या में पदीय मान हों" के रूप में परिभाषित किया जा सकता है। माध्यिका को प्रतीक M के द्वारा अभिव्यक्त किया जाता है।

अवर्गीकृत आंकड़ों के लिए माध्यिका की गणना

आँकड़े अवर्गीकृत होने पर उन्हें बढ़ते या घटते क्रम में व्यवस्थित कर लिया जाता है। इस व्यवस्थित श्रेणी में मध्यवर्ती पद के मान की स्थिति ज्ञात करके माध्यिका प्राप्त की जा सकती है। बढ़ते या घटते क्रम में व्यवस्थित श्रेणी के किसी भी सिरे से मध्यवर्ती मान की स्थिति निर्धारित की जा सकती है। माध्यिका की गणना करने के लिए निम्नलिखित सूत्र का उपयोग किया जाता है–

$\left(\frac{N+1}{2}\right)$ वाले पद का मान

उदाहरण 2.3 : निम्नांकित ऊँचाईयों का उपयोग करते हुए हिमालय की पर्वतीय-चोटियों की माध्यिका ऊँचाई की गणना कीजिए–

8,126 मी., 8,611मी., 7,817 मी., 8,172 मी., 8,076 मी., 8,848 मी., 8,598 मी.

गणना : माध्यिका M की गणना निम्न चरणों में की जा सकती है–

(i) दिए हुए आंकड़ों को बढ़ते अथवा घटते क्रम में व्यवस्थित कीजिए।

(ii) श्रेणी में मध्यवर्ती मूल्य का मान जानने के लिए सूत्र का उपयोग कीजिए। इस प्रकार–

$\left(\frac{N+1}{2}\right)$ वाले पद का मान

$=\left(\frac{7+1}{2}\right)$ वाले पद का मान

$=\left(\frac{8}{2}\right)$ वाले पद का मान

अर्थात् व्यवस्थित श्रेणी में चौथे पद का मान माध्यिका होगी।

आंकड़ों का बढ़ते क्रम में व्यवस्थापन–

7,817; 8,076; 8,126; 8,172; 8,598; 8,611; 8,848

चौथे पद का मान

अत:

M = 8,172 मीटर

वर्गीकृत आंकड़ों से माध्यिका की गणना

आँकड़े वर्गीकृत होने पर हमें उस बिंदु का मान ज्ञात करना होता है, जहाँ कोई व्यक्ति प्रेक्षण किसी वर्ग के माध्य में स्थित होता है। इसकी गणना निम्न सूत्र से की जा सकती है–

$$M = l + \frac{i}{f}\left(\frac{N}{2} - c\right)$$

जिसमें,

M	=	वर्गीकृत आंकड़ों के लिए माध्यिका
l	=	माध्यिका वर्ग की निम्न सीमा
i	=	वर्ग अंतराल
f	=	माध्यिका वर्ग की आवृत्ति
N	=	आवृत्ति का कुल योग अथवा प्रेक्षणों की संख्या
c	=	माध्यिका वर्ग से पहले वाले वर्ग की संचयी आवृत्ति।

उदाहरण–2.4 **:** निम्न वितरण के लिए माध्यिका की गणना कीजिए

वर्ग	50-60	60-70	70-80	80-90	90-100	100-110
f	3	7	11	16	8	5

तालिका–2.4 : माध्यिका की गणना

वर्ग	*आवृत्ति (f)*	*संचयी आवृत्ति (F)*	*माध्यिका वर्ग की गणना*
50-60	3	3	
60-70	7	10	
70-80	11	*21c*	
80-90 (माध्यिका वर्ग)	***16 f***	**37**	$M=\frac{N}{2}$
90-100	8	45	
100-110	5	50	$=\frac{50}{2}$
	$\sum f$ **या N= 50**		= 25

नीचे दिए गए चरणों के अनुसार माध्यिका की गणना की जाती है–

(i) *तालिका–2.4* की भाँति आवृत्तियों के लिए सारणी बना ली जाती है।

(ii) *तालिका–2.4* के स्तंभ 3 में दिए अनुसार प्रत्येक अगली साधारण आवृत्ति को जोड़कर संचयी आवृत्तियों **(F)** प्राप्त की जाती है।

(iii) $\frac{N}{2}$ के द्वारा माध्यिका संख्या ज्ञात की जाती है, जो कि इस उदाहरण में $\frac{50}{2}$ = **25** है। इसकी गणना *तालिका–2.4* के चौथे स्तंभ में दर्शाई गई है।

(iv) $\frac{\mathbf{N}}{\mathbf{2}}$ से अधिक मान प्राप्त होने तक संचयी आवृत्ति के वितरण (**F**) में ऊपर से नीचे की ओर गणना कीजिए। इस उदाहरण में $\frac{N}{2}$ = 25 है, जो कि 40–44 वाले वर्ग में सम्मिलित है। अत: इसे माध्यिका वर्ग कहते हैं। इस वर्ग की संचयी आवृत्ति 37, साधारण आवृत्ति 16 तथा माध्यिका वर्ग से पहले वाले वर्ग की संचयी आवृत्ति 21 है।

(v) चौथे चरण में निर्धारित इस सभी मानों को निम्न सूत्र में प्रतिस्थापित करके माध्यिका की गणना की जाती है–

$$M = l + \frac{i}{f}(m - c)$$

$$= 80 + \frac{10}{16}(25 - 21)$$

$$= 80 + \frac{5}{8} \times 4$$

$$= 80 + \frac{5}{2}$$

$$= 80 + 2.5$$

$$M = 82.5$$

बहुलक

किसी श्रेणी में जिस मान की सर्वाधिक पुनरावृत्ति होती है। वह मान **बहुलक** कहलाता है इसके संकेताक्षर **Z** अथवा $\mathbf{M_O}$ हैं। माध्य तथा माध्यिका की तुलना में बहुलक का उपयोग कम प्रचलित है। किसी श्रेणी में एक से अधिक बहुलक भी हो सकते हैं।

अवर्गीकृत आंकड़ों के लिए बहुलक की गणना

दिए हुए आंकड़ों के समूह से बहुलक की गणना करने के लिए पहले सभी मापों को बढ़ते या घटते क्रम में व्यवस्थित कर लिया जाता है। इससे सर्वाधिक पुनरावृत्ति वाले मान की पहचान करने में आसानी रहती है।

उदाहरण 2.5 : निम्नांकित दस विद्यार्थियों के भूगोल की परीक्षा में प्राप्तांकों के लिए बहुलक की गणना कीजिए।

61, 10, 88, 37, 61, 72, 55, 61, 46, 22

गणना : बहुलक ज्ञात करने के लिए निम्नानुसार सभी प्राप्तांकों को बढ़ते क्रम में व्यवस्थित कर लिया जाता है–

10, 22, 37, 46, 55, **61, 61, 61**, 72, 88

दिए हुए आंकड़ों में तीन बार की पुनरावृत्ति वाला मान 61, दी हुई श्रेणी का बहुलक है। चूँकि इस श्रेणी में अन्य किसी संख्या के मान में ऐसी विशेषता नहीं है, अतः यह, इस श्रेणी में **एक-बहुलक** है।

उदाहरण 2.6 : दस विद्यार्थियों के एक अन्य प्रतिदर्श के लिए निम्नांकित प्राप्तांकों के आधार पर बहुलक ज्ञात कीजिए–

82, 11, 57, 82, 08, 11, 82, 95, 41, 11

गणना : निम्नानुसार सभी दिए गए प्राप्तांको को बढ़ते क्रम में व्यवस्थित कीजिए–

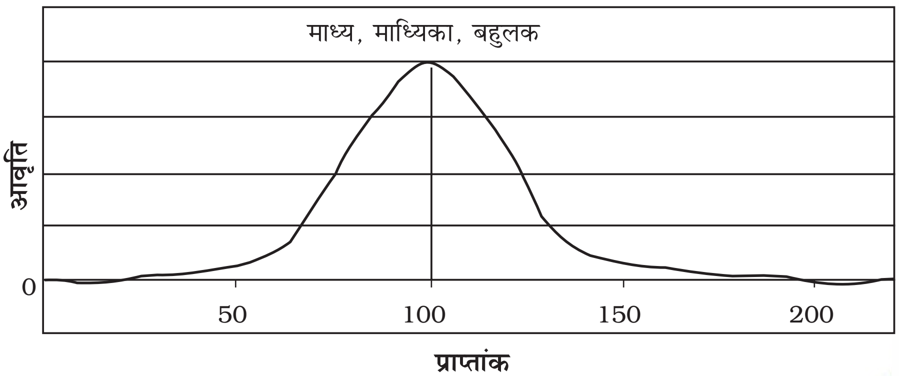
08, 11, 11, 11, 41, 57, 82, 82, 82, 95

उपरोक्त व्यवस्थित श्रेणी में आसानी से देखा जा सकता है कि 11 तथा 82, दोनों मानों के वितरण में तीन बार पुनरावृत्ति हुई है। अतः आंकड़ों के इस समूह का स्वरूप **द्वि-बहुलक** है। यदि किसी श्रेणी में तीन मानों की पुनरावृत्ति समान तथा सबसे अधिक बार होती है तो उस श्रेणी को **त्रि-बहुलक** श्रेणी कहते हैं। ऐसे ही कई मानों की समान बार पुनरावृत्ति होने पर **बहु-बहुलक** श्रेणी बन जाती है तथापि किसी श्रेणी में एक भी मान की पुनरावृत्ति न होने पर वह **बहुलक-रहित** श्रेणी कहलाती है।

माध्य, माध्यिका तथा बहुलक की तुलना

सामान्य वितरण वक्र की सहायता से केंद्रीय प्रवृति के तीनों मापों की तुलना आसानी से की जा सकती है। सामान्य वक्र आवृत्तियों का ऐसा वितरण होता है जिसको प्रदर्शित करने वाला रेखाचित्र **घंटाकार वक्र** कहलाता है। बौद्धिकता, व्यक्तित्व, समंक तथा विद्यार्थियों की उपलब्धि के समंक जैसी अनेक मानवीय विशेषताओं

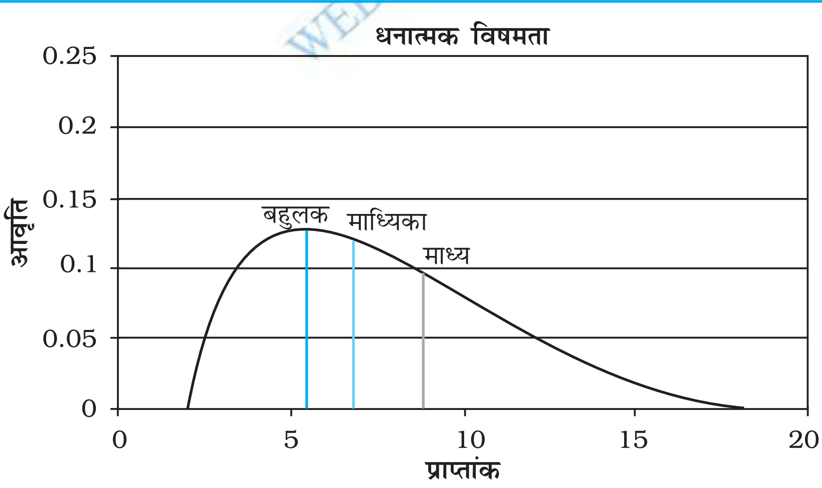
का सामान्य वितरण होता है। सामान्य वक्र की आकृति घंटाकार वक्र जैसी होती है क्योंकि यह वक्र सममित होता है। दूसरे शब्दों में अधिकांश प्रेक्षण श्रेणी के मध्य मान पर अथवा आस-पास एकत्रित होते हैं। जैसे-जैसे दूरस्थ मानों की ओर जाते हैं, वैसे-वैसे पर्यवेक्षित प्रेक्षणों की संख्या सममित रूप से घटती जाती है। सामान्य वक्र में आंकड़ों की परिवर्तनशीलता कम अथवा अधिक हो सकती है। सामान्य वक्र का एक उदाहरण *चित्र-2.3* में दर्शाया गया है।



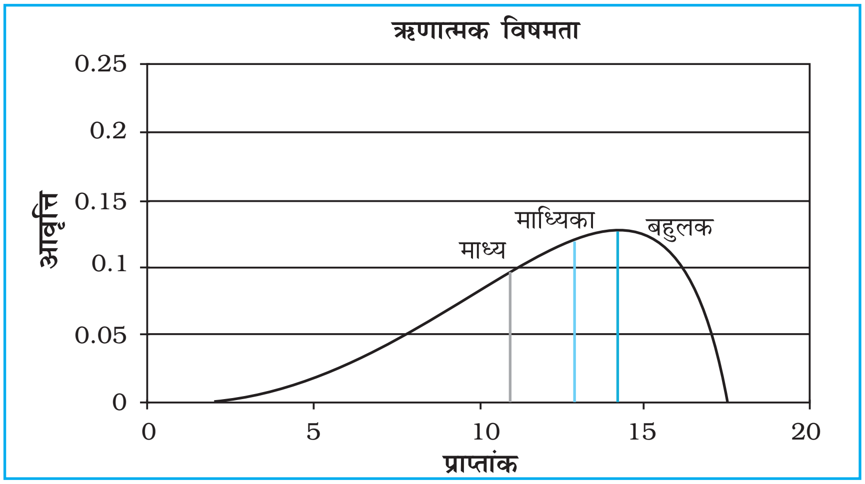
चित्र 2.3 : *सामान्य वितरण वक्र*

सामान्य वितरण की एक विशेषता होती है। इसमें **माध्य, माध्यिका तथा बहुलक का मान समान होता है** *(चित्र-2.3* में यह मान 100 है) क्योंकि सामान्य वितरण सममित होता है। अधिकतम आवृत्ति का मान वितरण के मध्य में होता है तथा इस बिंदु से आधी इकाइयाँ ऊपर तथा आधी नीचे होती हैं। अधिकतर इकाइयाँ वितरण के मध्य में अथवा माध्य के निकट होती हैं। अति उच्च तथा अति निम्न मूल्यों की बारंबारता अधिक नहीं होता, अत: वे विरले ही होते हैं।

यदि आंकड़े किसी प्रकार विषम अथवा विकृत हों तो माध्य, माध्यिका तथा बहुलक संपाती नहीं होंगे तथा विषम आंकड़ों के प्रभाव पर विचार करने की आवश्यकता है *(चित्र-2.4 तथा 2.5)*



चित्र 2.4 : *धनात्मक विषमता*



चित्र 2.5 : ऋणात्मक विषमता

प्रकीर्णन के माप

केवल केंद्रीय प्रवृत्ति माप ही वितरण का समुचित रूप से वर्णन नहीं करते हैं क्योंकि वे केवल वितरण का केंद्र ही चिह्नित करते हैं तथा उसे यह ज्ञात नहीं होता कि विभिन्न मूल्य अथवा माप केंद्र के परिप्रेक्ष्य में किस प्रकार प्रकीर्णित हैं। केंद्रीय प्रवृत्ति के माप की इन सीमाओं को समझने के लिए *तालिका–2.5* तथा 2.6 में दिए गए आंकड़ों का उपयोग करते हैं।

तालिका 2.5 : व्यक्तियों के प्राप्तांक

इकाई	प्राप्तांक
X_1	52
X_2	55
X_3	50
X_4	48
X_5	45

तालिका 2.6 : व्यक्तियों के प्राप्तांक

इकाई	प्राप्तांक
X_1	28
X_2	00
X_3	98
X_4	55
X_5	69

दोनों ही श्रेणियों के लिए $\overline{X} = 50$ है।

स्पष्ट है कि आंकड़ों के दोनों समूहों से प्राप्त किया गया माध्य एक समान अर्थात् 50 है। *तालिका–2.5* में दिए गए आंकड़ों में उच्चतम व निम्नतम मान क्रमशः 55 तथा 45 हैं। *तालिका–2.6* में दिए गए वितरण में ये अधिकतम तथा न्यूनतम मान क्रमशः 98 तथा 00 (शून्य) हैं। यद्यपि दोनों वितरणों का माध्य समान है, तथापि द्वितीय वितरण जो कि अधिक **अस्थिर** अथवा **विषम** है की अपेक्षा प्रथम वितरण **स्थिर** तथा **समरूप** है। इससे यह प्रश्न उत्पन्न होता है कि क्या माध्य वितरण की सभी विशेषताओं का पर्याप्त संकेतक है। ये उदाहरण ठोस प्रमाण देते हैं कि ऐसा नहीं है। अतः वितरण का श्रेष्ठतर प्रतिबिंब प्राप्त करने के लिए हमें **केंद्रीयता की प्रवृत्ति** के माप तथा **प्रकीर्णन** या **विषमता** के माप की भी आवश्यकता होती है।

प्रकीर्णन से तात्पर्य केंद्रीय प्रवृत्ति के माप से, इकाइयों के बिखराव से लगाया जाता है। यह माप औसत मूल्य से किसी इकाई अथवा संख्यात्मक मान की विषमता या बिखराव की प्रवृत्ति का मापन करता है। इस प्रकार प्रकीर्णन केंद्रीय मान से विभिन्न मूल्यों के बिखराव अथवा विषमता की मात्रा है।

प्रकीर्णन निम्नलिखित दो आधारभूत उद्देश्यों की पूर्ति करता है :

(i) इससे हमें वितरण या श्रेणी के संघटन की प्रकृति का ज्ञान होता है तथा

(ii) इसकी सहायता से दिए हुए वितरण की तुलना स्थिरता अथवा समरूपता के आधार पर हो जाती है।

प्रकीर्णन के मापन की विधियाँ

प्रकीर्णन के मापन की निम्नलिखित विधियाँ हैं :

1. विस्तार
2. चतुर्थक विचलन
3. माध्य विचलन
4. मानक विचलन (SD) तथा विचरण गुणांक (CV)
5. लॉरेंज़ वक्र

इनमें से प्रत्येक विधि के अपने विशेष गुण एवं सीमाएँ हैं। अत: इनमें से किसी भी विधि का उपयोग सावधानीपूर्वक करने की आवश्यकता है। विस्तार के साथ-साथ प्रकीर्णन के सापेक्ष माप के रूप में मानक विचलन तथा प्रकीर्णन के सापेक्षिक माप के रूप में विचरण गुणांक (CV), प्रकीर्णन के सबसे अधिक प्रचलित माप हैं। हम इन सभी मापों की गणना विधियों का विवेचन करेंगे।

विस्तार

किसी श्रेणी में अधिकतम व न्यूनतम मूल्यों के अंतर को **विस्तार** (R) कहते हैं। इस प्रकार यह किसी श्रेणी में सबसे छोटे से लेकर सबसे बड़े माप की दूरी है। इसे उच्चतम मान में न्यूनतम मान के घटाए हुए परिणाम के रूप में परिभाषित किया जा सकता है।

अवर्गीकृत आंकड़ों के लिए विस्तार की गणना

उदाहरण 2.7 : निम्नांकित दैनिक मजदूरी के वितरण के लिए विस्तार की गणना कीजिए

रु. 40, 42, 45, 48, 50, 52, 55, 58, 60, 100

विस्तार की गणना

R की गणना निम्नलिखित सूत्र की सहायता से हो सकती है–

$$R = L - S$$

जिसमें,

'R' = विस्तार

'L' तथा 'S' क्रमश: अधिकतम तथा न्यूनतम मान के प्रतीक हैं।

अत:

$$\mathbf{R} = L - S = 100 - 40 = \mathbf{60}$$

यदि हम उपरोक्त वितरण में से दसवें मूल्य को हटा दें तो **R** का मान 20 (60–40) रह जाएगा। श्रेणी में से केवल एक मूल्य को हटा देने पर **R** का मान घटकर केवल एक-तिहाई रह गया है। इससे स्पष्ट है कि प्रकीर्णन के माप के रूप **R** के साथ कठिनाई है कि इसका मान पूर्णत: दो चरम मूल्यों पर आधारित होता है। इस प्रकार प्रकीर्णन के माप के रूप में **R** का क्रियात्मक रूप ठीक वैसा ही जैसा केंद्रीय की प्रवृत्ति के माप में बहुलक का है। दोनों ही माप **अत्यधिक अस्थिर** हैं।

मानक विचलन

प्रकीर्णन के माप के रूप में मानक विचलन (SD) सबसे अधिक प्रचलित माप है। इसे विचलनों के वर्ग के औसत के वर्गमूल के रूप में परिभाषित किया जाता है। इसकी गणना हमेशा माध्य के परिप्रेक्ष्य में की जाती है। मानक विचलन प्रकीर्णन का सर्वाधिक स्थिर माप है जिसका अन्य सांख्यिकीय गणनाओं में उपयोग किया जाता है। ग्रीक अक्षर σ इसका संकेताक्षर है।

SD ज्ञात करने के लिए किसी श्रेणी के माध्य से प्रत्येक मूल्य के विचलन (x) का वर्ग (x^2) किया जाता है। यहाँ यह तथ्य ध्यान देने योग्य है कि इस चरण के कारण विचलनों के सभी ऋणात्मक मान धनात्मक हो जाते हैं। यह प्रक्रिया मानक विचलन को औसत विचलन की एक बड़ी आलोचना से बचा लेते हैं जो *मापांक* x का उपयोग करता है। इसके पश्चात् विचलनों के सभी वर्गों को जोड़ लिया जाता है ($\sum x^2$)। इसमें यह **सावधानी** रखनी होती है कि विचलनों को **पहले जोड़कर फिर वर्ग नहीं** किया जाता। इस वर्ग विचलनों के योग को पदों की संख्या से विभाजित करके उसका वर्गमूल निकाला जाता है। इसलिए **मानक विचलन को वर्गमूल माध्य वर्ग विचलन** के रूप में परिभाषित किया जाता है। दिए हुए आंकड़ों के लिए इसकी गणना निम्न सूत्र के उपयोग के द्वारा की जाती है–

$$s = \sqrt{\frac{\text{å } x^2}{N}}$$

गणना के पदों में वर्गमूल निकालने से पहले एक पारिभाषिक शब्द आता है। इसे **प्रसरण** कहा जाता है। प्रसरण का उपयोग अग्रिम सांख्यिकीय गणनाओं में किया जाता है। इसका वर्गमूल ही मानक विचलन होता है। इसी प्रकार इसका विपरीत भी सत्य है अर्थात् मानक विचलन (SD) का वर्ग ही प्रसरण है।

अवर्गीकृत आंकड़ों के लिए मानक विचलन की गणना

उदाहरण 2.7 : निम्नांकित मूल्यों के लिए मानक विचलन ज्ञात कीजिए

01, 03, 05, 07, 09

$$s = \sqrt{\frac{\text{å } x^2}{N}}$$

$$= \sqrt{\frac{40}{5}}$$

$$= \sqrt{8} = 2.828$$

$$= \mathbf{2.83}$$

तालिका 2.7 : मानक विचलन की गणना

X	$x(X - \overline{X})$	x^2
1	–4	16
3	–2 / –6	4
5	0	0
7	2	4
9	6-Apr	16
å X = 25		
N =5		
∴= 5		

उपरोक्त गणनाओं के पदों के सारांश निम्नानुसार हैं :

(i) सभी मूल्यों को **X** द्वारा चिह्नित स्तंभ में रखा गया है।

(ii) सभी मूल्यों को जोड़कर उसमें कुल पदों का भाग देकर माध्य ज्ञात किया गया है।

(iii) प्रत्येक मूल्य का विचलन **(*x*)** वास्तविक मूल्य से माध्य को घटाकर प्राप्त किया गया है। इसकी शुद्धता की जाँच विचलनों के योग से की जा सकती है, जो कि सदैव शून्य होता है। इस अभ्यास में भी यह तथ्य देखा जा सकता है।

(iv) विचलन **(*x*)** को वर्ग करके उसका योग किया गया है।

(v) सभी वर्ग विचलनों के योग को पदों की संख्या से विभाजित किया गया है। पुनर्स्मरण कीजिए कि इससे प्रसरण ज्ञात हो जाता है।

(vi) इसका वर्गमूल निकालने से मानक विचलन प्राप्त हो जाता है।

वर्गीकृत आंकड़ों के लिए मानक विचलन की गणना

उदाहरण : निम्नलिखित आंकड़ों के लिए मानक विचलन की गणना कीजिए

वर्ग –	120-130	130-140	140-150	150-160	160-170	170-180
f –	2	4	6	12	10	6

निम्नांकित तालिका में वर्गीकृत आंकड़ों के लिए मानक विचलन ज्ञात करने की विधि समझाई गई है। इस तालिका के प्रथम चार स्तंभों तक पद वही है, जो वर्गीकृत आंकड़ों के लिए माध्य की गणना करते समय अपनाए थे। हम गणना का प्रारंभ मध्यांतर वर्ग 150–160 में कल्पित माध्य मानकर करते हैं, अतः इस मध्यांतर वर्ग के सम्मुख तीसरे स्तंभ में अर्थात् कल्पित माध्य से पद–विचलन शून्य अंक दिया गया है। इसके पश्चात् अन्य वर्गों के पद–विचलन अंकित किए जाते हैं। स्तंभ–4 (fx') के मान पिछले दो स्तंभों के गुणनफल हैं। पाँचवे स्तंभ ($fx'2$) के मान तीसरे व चौथे स्तंभों के गुणनफल हैं। उसके बाद सभी स्तंभों के मानों का योग कर लिया जाता है।

(1) वर्ग	(2) *f*	(3) *x´*	(4) *fx´*	(5) *fx´2*
120 - 130	2	–3	–6	18
130 - 140	4	–2	–8	16
140 - 150	6	–1	- 6 - 20	6
150 - 160	12	0	0	0
160 - 170	10	1	10	10
170 - 180	6	2	12 22	24
	N=40		$\sum fx' = 2$	$\sum fx'2 = 74$

मानक विचलन की गणना के लिए निम्नलिखित सूत्र का उपयोग किया जाता है :

$$SD = i^{2}\sqrt{\sum fx'^{2} - \frac{\sum fx'}{N}}$$

विचरण गुणांक (CV)

यदि आँकड़े माप की अलग–अलग इकाइयों में भिन्न–भिन्न स्थानों अथवा अवधियों के हों तथा उनकी परस्पर तुलना करनी हो तो विचरण गुणांक बहुत उपयोगी सिद्ध होता है। विचरण गुणांक मानक विचलन के माध्यम को माध्य के प्रतिशत के रूप में अभिव्यक्त करता है। इसका निर्धारण निम्नांकित सूत्र के उपयोग द्वारा होता है :

$$CV = \frac{\text{मानक विचलन}}{\text{माध्य}} \quad 100$$

$$CV = \frac{\sigma}{\bar{X}} \times 100$$

इस प्रकार तालिका-2.7 में दिए गए आंकड़ों के लिए CV निम्नानुसार होगा :

$$CV = \frac{\sigma}{\bar{X}} \times 100$$

$$CV = \frac{2.83}{5} \times 100$$

$$CV = \mathbf{56\%}$$

इसी सूत्र से वर्गीकृत आंकड़ों का विचरण गुणांक ज्ञात किया जा सकता है।

कोटि सहसंबंध

अभी तक जितनी सांख्यिकीय विधियों की विवेचना की गई है, उन सभी का संबंध एक ही चर से था। अब हम दो चरों के मध्य साहचर्य के अन्वेषण करने वाली विधियों की व्याख्या करेंगे तथा यह भी देखेंगे कि इस साहचर्य की अभिव्यक्ति संख्यात्मक रूप से कैसे की जा सकती है? दो या दो से अधिक चरों के बारे में चर्चा होने पर यह जिज्ञासा उठती है कि क्या किसी एक चर में परिवर्तन का प्रभाव दूसरे चर में किसी प्रकार के परिवर्तन पर पड़ता है।

बहुधा हमारी रुचि दो या दो से अधिक चरों के मध्य साहचर्य अथवा पारस्परिक निर्भरता की प्रकृति जानने में रहती है। ऐसा समझा जाता है कि **सहसंबंध** इस उद्देश्य से उपयोगी है। आधारभूत रूप से यह दो या दो से अधिक चरों के मध्य साहचर्य का माप है। चूँकि हम इसके अंतर्गत यह अध्ययन करते हैं कि संबंधित घटक एक-दूसरे के साथ किस प्रकार विचरण करते हैं अत: इन्हें **चर** कहा जाता है। इस प्रकार पारिभाषिक शब्दावली के रूप में सहसंबंध से तात्पर्य दो चरों के मध्य अनुरूपता अथवा साहचर्य की प्रकृति एवं गहनता से है। इस परिभाषा में सम्मिलित पारिभाषिक शब्दावली के रूप में **प्रकृति** तथा **गहनता** का आशय **दिशा** एवं **मात्रा** से है, जिसके अनुरूप दो चर परस्पर विचरण करते हैं।

सहसंबंध की दिशा

यह हमारा सामान्य अनुभव है कि कुछ प्राप्ति के लिए निवेश किया जाता है। इससे तीन संभावनाएँ रहती हैं :

1. निवेश में वृद्धि से प्राप्ति में भी वृद्धि हो।
2. निवेश में वृद्धि से प्राप्ति में कमी हो।
3. निवेश की मात्रा में परिवर्तन से प्राप्ति की मात्रा में कोई परिवर्तन न हो।

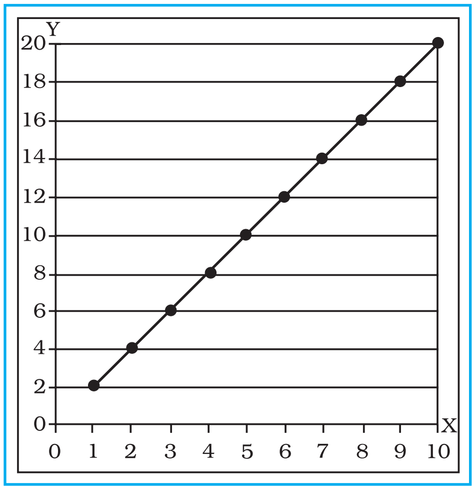
प्रथम स्थिति में निवेश तथा प्राप्ति में साहचर्य की दिशा एक ही है। इस स्थिति में ऐसा कहा जाता है कि दोनों के मध्य **धनात्मक सहसंबंध** है।

द्वितीय स्थिति में निवेश व प्राप्ति के मध्य परिवर्तन की दिशा एक-दूसरे के विपरीत है, अत: कहा जाता है कि दोनों के मध्य **ऋणात्मक सहसंबंध** है।

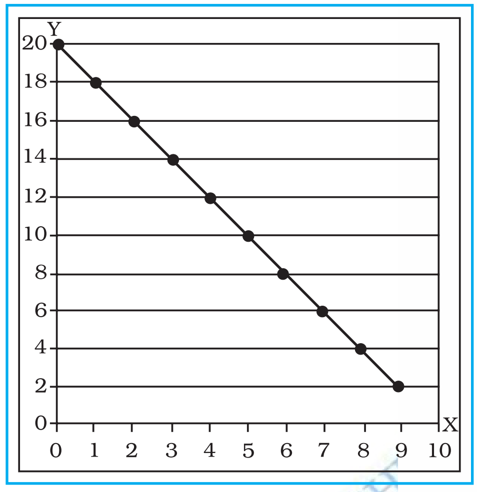
तृतीय स्थिति में निवेश व प्राप्ति के मध्य कोई साहचर्य विद्यमान नहीं है। अत: यह कहा जाता है कि दोनों के मध्य कोई महत्वपूर्ण सांख्यिकीय सहसंबंध नहीं है।

आइए अब हम *चित्र 2.7* देखें जो *चित्र 2.6* से एकदम विपरीत है। उसमें रेखाचित्र पर अंकित मानों की दिशा ऊपर बाएँ से नीचे दाईं ओर है। यह भी ध्यान दीजिए कि X-अक्ष पर प्रत्येक एक इकाई वृद्धि के साथ-साथ Y-अक्ष पर दो इकाइयों की कमी हो जाती है। यह **ऋणात्मक सहसंबंध** का उदाहरण है। इसका अर्थ यह है कि दोनों चरों में एक-दूसरे के विपरीत गति करने की प्रवृत्ति है, अर्थात् यदि एक चर में वृद्धि

हो तो दूसरे में कमी होती है तथा इसका विपरीत भी। इस प्रकार साहचर्य हमें कई भौगोलिक युग्म चरों में मिल सकता है। समुद्र तल से ऊँचाई तथा वायुदाब, तापमान तथा वायुदाब आदि कुछ ऐसे ही उदाहरण हैं। इससे यह भी अर्थ निकलता है कि सहसंबंध की द्योतक संख्या से पहले गणितीय चिह्न होना आवश्यक है (+ या –) अधिक आवश्यक रूप से जबकि सहसंबंध ऋणात्मक हो।



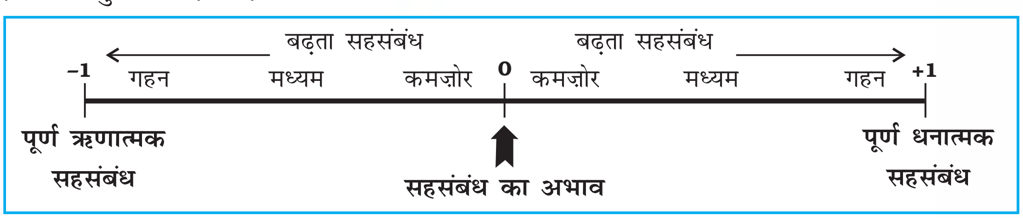
चित्र 2.6 : पूर्ण धनात्मक सहसंबंध



चित्र 2.7 : पूर्ण ऋणात्मक सहसंबंध

सहसंबंध की गहनता

जब सहसंबंध की दिशा ऋणात्मक अथवा धनात्मक के विषय में संदर्भ आ चुका होता है तो स्वाभाविक रूप से यह जानने के लिए जिज्ञासा जागृत होती है कि दोनों चरों में अनुरूपता अथवा साहचर्य की गहनता की मात्रा कितनी है। इस अनुरूपता अथवा साहचर्य की गहनता की मात्रा गणितीय दृष्टि से अधिकतम **1 (एक)** तक होती है। इस मात्रा में सहसंबंध की दिशा का पहलू जोड़ने पर इसका अधिकतम विस्तार –1 **से शून्य की ओर होते हुए** +1 तक होता है। इसका मान **किसी भी परिस्थिति में एक से अधिक नहीं हो सकता।** इस विस्तार का रैखिक वर्णन *चित्र 2.8* में दर्शाया गया है। सहसंबंध पूरा 1 (एक) होने पर (चाहे धनात्मक हो या ऋणात्मक) इसे **पूर्ण सहसंबंध** कहते हैं। इस प्रकार गहनतम सहसंबंध के दो विपरीत सिरों के ठीक मध्य में **शून्य (O) सहसंबंध** स्थित होता है, जिस बिंदु पर चरों के मध्य **सहसंबंध का अभाव** अथवा सहसंबंध अनुपस्थित होता है।



चित्र 2.8 : सहसंबंध की दिशा व गहनता का विस्तार

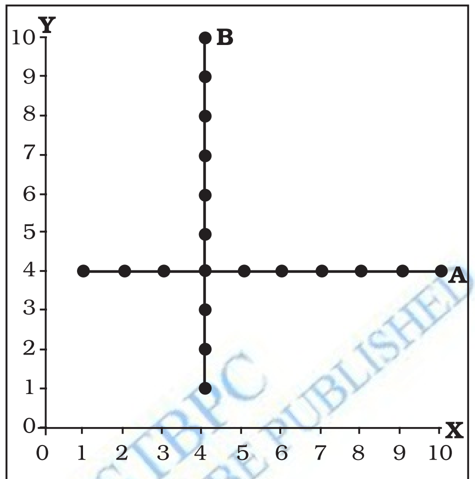
पूर्ण सहसंबंध

चित्र 2.6 तथा *2.7* की रचना दो चरों के मध्य विशिष्ट साहचर्य को दर्शाने के लिए किया गया है। ध्यान दीजिए कि ये रेखाचित्र X तथा Y के मानों का बिखराव अथवा प्रकीर्णन दर्शाते हैं। इसलिए ऐसे रेखाचित्रों को **प्रकीर्ण आरेख** अथवा **प्रकीर्ण अंकन** कहते हैं। *चित्र 2.6* से यह स्पष्ट है कि जब इस प्रकार के युग्म के मानों को अंकित किया जाता है, तो सभी बिंदु एक सरल रेखा पर स्थित होते हैं। जब यह सरल रेखा प्रकीर्ण आरेख के निचले बाएँ से ऊपरी दाएँ भाग की ओर जाती है तो यह **पूर्ण धनात्मक सहसंबंध (1.00)** का उदाहरण होता है। *चित्र 2.7* इसका ठीक विपरीत है। इसमें भी सभी बिंदु एक सरल रेखा पर अंकित

हैं। यह रेखा प्रकीर्ण आरेख के ऊपरी बाएँ भाग से इनके निचले दाएँ भाग की ओर विस्तारित है। यह **पूर्ण ऋणात्मक सहसंबंध** (जिसका मान –1.00 है) का उदाहरण है। **सहसंबंध का अभाव** (अथवा **शून्य सहसंबंध**) तब होता जबकि युग्म के दोनों एक-दूसरे में परिवर्तन का कोई प्रत्युत्तर नहीं देते। इस स्थिति में दोनों चरों के मध्य कोई सहसंबंध नहीं होता, अतः इसे **सहसंबंध के अभाव** अथवा **शून्य सहसंबंध** की स्थिति कहते हैं। इसे *चित्र 2.9* में दर्शाया गया है। X-चर में परिवर्तन का Y-चर द्वारा प्रत्युत्तर नहीं दिए जाने के कारण उत्पन्न शून्य सहसंबंध को **प्रकीर्ण अंकन-A** द्वारा दर्शाया गया है। इसी प्रकार **प्रकीर्ण अंकन – B** में भी शून्य सहसंबंध की स्थिति है, जो कि Y-चर में परिवर्तन पर X-चर द्वारा कोई प्रत्युत्तर नहीं दिए जाने के कारण उत्पन्न हुई है।

A

X	Y
1	4
2	4
3	4
4	4
5	4
6	4
7	4
8	4
9	4
10	4



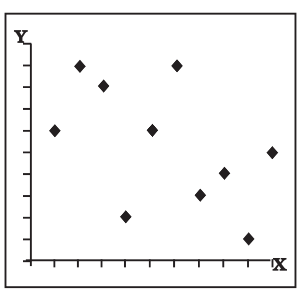
B

X	Y
4	1
4	2
4	3
4	4
4	5
4	6
4	7
4	8
4	9
4	10

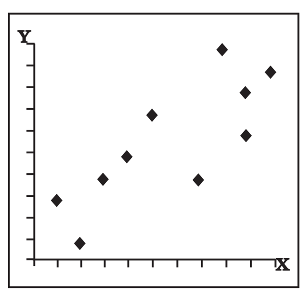
चित्र 2.9: शून्य सहसंबंध को दर्शाने वाला प्रकीर्ण आरेख

अन्य सहसंबंध

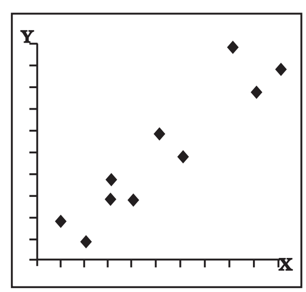
पूर्ण सहसंबंधों (±1) व शून्य सहसंबंध के मध्य में साहचर्य की सामान्य परिस्थितियाँ पाई जाती हैं जिन्हें **कमजोर, मध्यम** तथा **गहन** सहसंबंध के नाम से जाना जाता है। इन तीनों परिस्थितियों को क्रमशः *चित्र 2.10, 2.11* तथा *2.12* में स्पष्ट रूप से दर्शाया गया है। इनमें अंकित बिंदुओं के बिखराव अथवा प्रकीर्णन के स्वरूप तथा उनको दिए गए विशिष्ट नाम, यथा कमज़ोर, मध्यम तथा गहन की ओर ध्यान दीजिए (ये परिस्थितियाँ सामान्य विशेषण हैं, जिनकी कोई विशिष्ट सीमाएँ निर्धारित नहीं हैं) बिखराव जितना अधिक होगा, सहसंबंध उतना ही कमज़ोर होगा। प्रकीर्णन जितना कम होगा, सहसंबंध उतना ही गहन होगा, तथा अंकित बिंदुओं के एक सरल रेखा पर स्थित हो जाने पर पूर्ण सहसंबंध होगा *(चित्र 2.6 तथा 2.7)*।



चित्र 2.10 :
कमज़ोर ऋणात्मक सहसंबंध



चित्र 2.11 :
मध्यम धनात्मक सहसंबंध



चित्र 2.12 :
गहन धनात्मक सहसंबंध

सहसंबंध की गणना करने की विधियाँ

सहसंबंध की गणना करने की अनेक विधियाँ हैं किंतु समय व स्थान की सीमाओं को ध्यान में रखते हुए यहाँ हम केवल स्पीयरमैन के कोटि सहसंबंध की व्याख्या करेंगे।

स्पीयरमैन का कोटि सहसंबंध

स्पीयरमैन ने कोटियों के आधार पर सहसंबंध की गणना विधि की युक्ति प्रदान की। प्रचलित रूप से इसे स्पीयरमैन के कोटि सहसंबंध के नाम से जाना जाता है जिसका सांख्यिकी में संकेताक्षर r (ग्रीक अक्षर जिसका उच्चारण है रो **-rho**) है। इसकी गणना विधि आसान होने के कारण स्पीयरमैन के सहसंबंध का उपयोग अधिक प्रचलित है। इस संबंध की गणना निम्न चरणों में की जाती है :

(i) अभ्यास में दिए गए X तथा Y चरों के आंकड़ों को तालिका के क्रमशः प्रथम व द्वितीय स्तंभों में उतार लिया जाता है।

(ii) दोनों चरों की अलग-अलग कोटियाँ निर्धारित की जाती हैं। X-चर की कोटियों को तृतीय स्तंभ में अंकित किया जाता है जिसका शीर्षक **(XR)** (X-चर की कोटियाँ) है। इसी प्रकार Y-चर की कोटियों **(YR)** चतुर्थ स्तंभ में अंकित किया जाता है। आंकड़ों में उच्चतम मान को कोटि एक, दूसरे सर्वोच्च मान को कोटि दो तथा इसी प्रकार अन्य कोटियों का आवंटन किया जाता है। मान लीजिए X-चर के आँकड़े 4, 8, 2, 10, 1, 9, 7, 3, 0 तथा 5 हैं तो **XR** क्रमशः 6, 3, 8, 1, 9, 2, 4, 7, 10 व 5 होंगी। ध्यान दीजिए कि अंतिम कोटि (इस उदाहरण में 10) श्रेणी में कुल इकाइयों की संख्या के बराबर होती है। इसी प्रकार **YR** का भी निर्धारण किया जाता है।

(iii) **XR** तथा **YR** के निर्धारण के पश्चात् दोनों कोटियों में अंतर ज्ञात किया जाता है (जिसमें धनात्मक या ऋणात्मक चिह्नों का ध्यान नहीं रखते)। इस अंतर का अभिलेखन पाँचवें स्तंभ में लिखा जाता है। चूँकि अगले चरण में इन अंतरों का वर्ग निकाला जाता है, इसलिए इन अंतरों के साथ जुड़े **ऋणात्मक अथवा धनात्मक चिह्नों का कोई महत्त्व नहीं है।**

(iv) प्रत्येक अंतर का वर्ग ज्ञात करके उनका योग कर लिया जाता है। ये मान छठे स्तंभ में लिखे जाते हैं।

(v) इसके पश्चात् कोटि सहसंबंध की गणना निम्न सूत्र के आधार पर की जाती है–

$$r = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

जिसमें,

r = कोटि सहसंबंध

$\sum D^2$ = दोनों कोटियों के अंतर के वर्ग का योग

N = X-Y युग्मों की संख्या

उदाहरण 2.8 : निम्नांकित आंकड़ों के लिए स्पीयरमैन के कोटि सहसंबंध की गणना कीजिए–

अर्थशास्त्र में प्राप्तांक (X)	02 08 00 20 12 16 06 18 09 10
भूगोल में प्राप्तांक (Y)	04 12 06 24 16 18 08 20 09 10

तालिका 2.8 : स्पीयरमैन के कोटि सहसंबंध की गणना

(1) X	(2) Y	(3) XR	(4) YR	(5) D	(6) D^2
2	4	9	10	1	1
8	12	7	5	2	4
0	6	10	9	1	1
20	24	1	1	0	0
12	16	4	4	0	0
16	18	3	3	0	0
6	8	8	8	0	0
18	20	2	2	0	0
9	9	6	7	1	1
10	10	5	6	1	1
N=10					D^2=8

गणना :

जब r कोटि सहसंबंध **D** दोनों चरों X तथा Y की कोटियों का अंतर तथा **N** दोनों चरों अर्थात् x – y युग्मों की संख्या हो तो

$$= 1 - \frac{6 \times 8}{10(10^2 - 1)}$$

$$= 1 - \frac{48}{10(100 - 1)}$$

$$= 1 - \frac{48}{10(99)}$$

$$= 1 - \frac{48}{(990)}$$

$$= 1 - 0.05$$

$$= 0.95$$

जब आंकड़ों के अंतर्गत दी हुई इकाइयों की संख्या कम हो तो अन्य प्रकार के सहसंबंधों की तुलना में 'रो' अधिक उत्तम स्थानापन्न होता है। इकाइयों की संख्या अधिक होने पर यह लगभग अनुपयोगी हो जाता है क्योंकि जब तक सभी युग्मों की कोटियों की गणना की जाती है तब तक अन्य प्रकार के सहसंबंध की गणना की जा सकती है।

अभ्यास

1. निम्नांकित चार विकल्पों में से सही विकल्प चुनिए :
   - (i) केंद्रीय प्रवृत्ति का जो माप चरम मूल्यों से प्रभावित नहीं होता है वह है :
     - (क) माध्य
     - (ख) माध्य तथा बहुलक
     - (ग) बहुलक
     - (घ) माध्यिका
   - (ii) केंद्रीय प्रवृत्ति का वह माप जो किसी वितरण के उभरे भाग से हमेशा संपाती होगा वह है :
     - (क) माध्यिका
     - (ख) माध्य तथा बहुलक
     - (ग) माध्य
     - (घ) बहुलक
   - (iii) ऋणात्मक सहसंबंध वाले प्रकीर्ण अंकन में अंकित मानों के वितरण की दिशा होगी :
     - (क) ऊपर बाएँ से नीचे दाएँ
     - (ख) नीचे बाएँ से ऊपर दाएँ
     - (ग) बाएँ से दाएँ
     - (घ) ऊपर दाएँ से नीचे बाएँ

2. निम्नलिखित प्रश्नों के उत्तर लगभग 30 शब्दों में दीजिए :
   - (i) माध्य को परिभाषित कीजिए।
   - (ii) बहुलक के उपयोग के क्या लाभ हैं?
   - (iii) अपकिरण किसे कहते हैं?
   - (iv) सहसंबंध को परिभाषित कीजिए।
   - (v) पूर्ण सहसंबंध किसे कहते हैं?
   - (vi) सहसंबंध की अधिकतम सीमाएँ क्या हैं?

3. निम्नलिखित प्रश्नों के उत्तर लगभग 125 शब्दों में दीजिए :
   - (i) आरेखों की सहायता से सामान्य तथा विषम वितरणों में माध्य, माध्यिका तथा बहुलक की सापेक्षिक स्थितियों की व्याख्या कीजिए।
   - (ii) माध्य, माध्यिका तथा बहुलक की उपयोगिता पर टिप्पणी कीजिए (संकेत : उनके गुण तथा दोषों से)।
   - (iii) एक काल्पनिक उदाहरण की सहायता से मानक विचलन के गणना की प्रक्रिया समझाइए।
   - (iv) प्रकीर्णन का कौन-सा माप सबसे अधिक अस्थिर है तथा क्यों?
   - (v) सहसंबंध की गहनता पर एक विस्तृत टिप्पणी लिखिए।
   - (vi) कोटि सहसंबंध की गणना के विभिन्न चरण कौन-से हैं?

क्रियाकलाप

1. भौगोलिक विश्लेषण के लिए प्रयुक्त कोई काल्पनिक उदाहरण लीजिए तथा अवर्गीकृत आंकड़ों की गणना करने की प्रत्यक्ष व अप्रत्यक्ष विधियों को समझाइए।
2. विभिन्न प्रकार के पूर्ण सहसंबंध दर्शाने के लिए प्रकीर्ण आरेख बनाइए।