अध्याय – 2
आँकड़ें का प्रक्रमण
(Data Processing)

आँकड़ो के विश्लेषण के लिए अनेक विधियों का उपयोग किया जाता है। इनमें से निम्नलिखित विधियों का अध्ययन इस अध्याय में करेंगे –

1. केन्द्रीय प्रवृत्ति के माप
2. प्रकीर्णन के माप
3. सम्बन्ध के माप

केन्द्रीय प्रवृत्ति के माप

सांख्यिकी में समंक श्रेणी की केन्द्रीय प्रवृत्ति को सरल व सारांश रूप में अभिव्यक्त करने वाला प्रतिनिधि मूल्य **केन्द्रीय प्रवृत्ति का माप** कहलाता है। यह एक ऐसा संक्षिप्त तथा सरल माप है जो समंक श्रेणी के प्रमुख अभिलक्षणों पर प्रकाश डालता है। केन्द्रीय प्रवृत्ति के माप वैसे तो कई प्रकार के होते हैं, किन्तु सामान्यतः अधिक प्रचलित माप **समान्तर माध्य (Mean)**, **माध्यिका (Median)** व **बहुलक (Mode)** हैं।

माध्य

किसी चर के विभिन्न मूल्यों का साधारण अंकगणितीय औसत माध्य कहलाता है। अवर्गीकृत तथा वर्गीकृत आँकड़ों के लिए माध्य

ज्ञात करने की विधियाँ निश्चित ही भिन्न हैं। वर्गीकृत व अवर्गीकृत दोनों प्रकार के आँकड़ों के लिए माध्य प्रत्यक्ष व अप्रत्यक्ष विधियों के द्वारा ज्ञात किया जा सकता है।

अवर्गीकृत आँकड़ों से माध्य की गणना

प्रत्यक्ष विधि

अवर्गीकृत आँकड़ां 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।

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

जिसमें,

$\overline{X}$ = माध्य

$\sum$ = मापों के सभी मूल्यों का योग

x = मापों की किसी श्रेणी में एक अपरिष्कृत समंक

$\sum x$ = मापों की किसी श्रेणी में अपरिष्कृत समंकों का योग

N = श्रेणी के पदों की संख्या

उदाहरण : किसी कक्षा में 10 विद्यार्थियों के एक प्रश्न–पत्र में प्राप्तांक 50, 80, 90, 70, 60, 35, 15, 75, 45, 80 हैं। इनका माध्य प्रत्यक्ष विधि से ज्ञात कीजिए।

हल : माध्य की गणना के लिए सूत्र

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

जिसमें

$\overline{X}$ = माध्य

$\sum x$ = मापों के सभी मूल्यों का योग

N = पदों की संख्या

$$\therefore \overline{X} = \frac{50+80+90+70+60+35+15+75+45+80}{10}$$

$$= \frac{600}{100} = 60$$

$$\therefore \overline{X} = 60$$

अभ्यास

1. एक विद्यार्थी के पाँच प्रश्न–पत्रों के प्राप्तांक 30, 35, 54, 91, 50 हैं तो उनका माध्य प्रत्यक्ष विधि से ज्ञात कीजिए।

अप्रत्यक्ष विधि

श्रेणी में जहाँ प्रेक्षणों की संख्याएँ बहुत अधिक होती हैं, वहाँ सामान्यतः 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।

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

उदाहरण : किसी कक्षा में 10 विद्यार्थियों के एक प्रश्न–पत्र में प्राप्तांक 50, 80, 90, 70, 60, 35, 15, 75, 45, 80 हैं तो उनका माध्य अप्रत्यक्ष विधि से ज्ञात कीजिए।

हल :

प्राप्तांक **(x)**	**d = x – A (A=50)**
50	0
80	–30
90	40
70	20
60	10
35	–15
15	–35
75	25
45	–5
80	30
	$\sum d = 40$

माध्य की गणना के लिए सूत्र–

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

जहाँ

$A = 50$

$\sum d = 40$

$N = 10$

$\therefore \overline{X} = 50 + \frac{40}{10}$

$= 50 + 4$

$\overline{X} = 54$

अभ्यास

2. एक विद्यार्थी के छः प्रश्न–पत्रों के प्राप्तांक 50, 60, 75, 80, 55, 40 है तो उनका माध्य अप्रत्यक्ष विधि से ज्ञात कीजिए।

वर्गीकृत आँकड़ों से माध्य की गणना

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

प्रत्यक्ष विधि

जब आवृत्ति वितरण के रूप में आँकड़ें वर्गीकृत हो तो उसमें एकाकी मूल्य अपनी पहचान खो देते हैं। इन सभी मूल्यों का प्रतिनिधित्व वर्ग अंतराल के मध्य बिन्दुओं द्वारा होता है, जहाँ वे स्थित हैं। प्रत्यक्ष विधि से वर्गीकृत आँकड़ों के लिए माध्य की गणना करते समय प्रत्येक वर्ग के मध्य बिन्दुओं से सम्बन्धित आवृत्ति (f) को गुणा किया जाता है, fx (इसमें X मध्य बिन्दु है) के सभी मानों को जोड़कर प्राप्त $\sum fx$ में पदों की संख्या (N) से भाग दिया जाता है। अतः 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 बिन्दु

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

उदाहरण : निम्नलिखित आँकड़ों में एक परीक्षा के प्राप्तांक व विद्यार्थियों की संख्या दी गई है। प्राप्तांकों का माध्य प्रत्यक्ष विधि से ज्ञात कीजिए।

वर्ग (प्राप्तांक)	आवृत्ति (f) (विद्यार्थियों की संख्या)
20–30	05
30–40	09
40–50	16
50–60	12
60–70	08

हल :

वर्ग	आवृत्ति (f)	मध्यबिन्दु (x)	fx
20–30	05	25	125
30–40	09	35	315
40–50	16	45	720
50–60	12	55	660
60–70	08	65	520
	$\sum f = 50$		$\sum fx = 2340$

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

जहाँ $\sum fx = 2340$

$N = 50$

$$\therefore \overline{X} = \frac{2340}{50}$$

$= 46.8$

$\therefore \overline{X} = 46.8$

अभ्यास

3. निम्नांकित वर्गीकृत समंकों में कुछ श्रमिकों की दैनिक आय (Daily Wage) दशाई गई है। उनकी माध्य आय प्रत्यक्ष विधि से ज्ञात कीजिए।

वर्ग	(f)
40—49	16
50—59	20
60—69	32
70—79	12
80—89	20

अप्रत्यक्ष विधि

वर्गीकृत आँकड़ों 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 आँकड़ों के लिए अप्रत्यक्ष विधि द्वारा माध्य की गणना में दिए गए थे। इसे निम्न प्रकार से अभिव्यक्त किया जा सकता है–

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

जिसमें

A = कल्पित माध्य वाले वर्ग का मध्य बिन्दु

f = आवृत्ति

d = कल्पित माध्य वाले वर्ग (A) से विचलन

N = कुल पदों की संख्या अथवा $\sum f$

उदाहरण : निम्नलिखित आँकड़ों से प्राप्तांकों का माध्य अप्रत्यक्ष विधि से ज्ञात कीजिए।

प्राप्तांक	20–30	30–40	40–50	50–60	60–70
विद्यार्थियों की संख्या	05	09	16	12	08

हल :

वर्ग (प्राप्तांक)	आवृत्ति (विद्यार्थियों की संख्या)	मध्य बिन्दु (x)	$d = x - A$ $(A = 45)$	fd
20–30	05	25	–20	–100
30–40	09	35	–10	–90
40–50	16	45	0	0
50–60	12	55	10	120
60–70	08	65	20	160
	$\sum f = 50$			$\sum f = 90$

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

जहाँ

$$A = 45$$

$$\sum fd = 90$$

$$N = \sum f = 50$$

$$\therefore \overline{X} = 45 + \frac{90}{50}$$

$$= 45 + 1.8$$

$$= 46.8$$

$$\therefore \overline{X} = 46.8$$

अभ्यास

4. पचास विद्यार्थियों की कक्षा में आयोजित एक टैस्ट में प्राप्तांकों का श्रेणीवार विभाजन निम्नांकित है। इन वर्गीकृत समंकों के लिये प्राप्तांकों का अप्रत्यक्ष विधि से माध्य ज्ञात कीजिए।

वर्ग	आवृत्ति
120—129	07
130—139	10
140—149	14
150—159	09
160—169	10

माध्यिका

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

अवर्गीकृत आँकड़ों 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 : निम्नांकि अवर्गीकृत आँकड़ों से माध्यिका ज्ञात कीजिए–

99, 23, 28, 88, 19, 24, 58, 51, 37, 45, 14, 22, 103

हल : इन समंकों को निम्नानुसार क्रमबद्ध कर लेते हैं –

क्रम संख्या	**आँकड़े**
1	14
2	19
3	22
4	23
5	24
6	28
7	**37**
8	45
9	51
10	58
11	88
12	99
13	103

$M = \left(\frac{N+1}{2}\right)$ का मान

$= \frac{13+1}{2}$ का मान

$= \frac{14}{2}$ का मान

= **7वें पद का मान,** जो कि उपरोक्त क्रमबद्ध श्रेणी में **37** है

अतः **माध्यिका 37** है।

उदाहरण : निम्नांकित अवर्गीकृत आँकड़ों से माध्यिका ज्ञात कीजिए।

10, 7, 25, 52, 14, 1, 19, 39, 27, 9, 47, 66

हल : इन आँकड़ों को निम्नानुसार क्रमबद्ध कर लेते हैं –

क्रम संख्या	आँकड़ें
1	1
2	7
3	9
4	10
5	14
6	**19**
7	**25**
8	27
9	39
10	47
11	52
12	66

$M = \left(\frac{N+1}{2}\right)$ का मान

$= \frac{12+1}{2}$ का मान

$= \frac{13}{2}$ का मान

= **6.5वें पद का मान,**

छठे व सातवें पद के समंकों को जोड़कर दो का भाग देने पर 6.5वें पद का मान ज्ञात किया जाता है, अतः

$$\frac{19+25}{2}=\frac{44}{2}=22$$

इस प्रकार माध्यिका 22 है।

अभ्यास

5. निम्नांकित समंको के लिये माध्यिका ज्ञात कीजिए –

(i) 41, 62, 15, 34, 80, 16, 40, 54

(ii) 40, 70, 38, 14, 84, 62, 32, 12

वर्गीकृत आँकड़ों से माध्यिका की गणना

आँकड़ें वर्गीकृत होने पर हमें उस बिन्दु 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 आँकड़ों 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 प्रवृत्ति

उदाहरण : निम्नांकित दिये गये वर्गीकृत आँकड़ो से माध्यिका ज्ञात कीजिए।

वर्ग	**आवृत्ति** (f)
10–15	04
15–20	05
20–25	09
25–30	10
30–35	10
35–40	07
40–45	05

हल :

वर्ग	आवृत्ति (f)	F (संचयी आवृत्ति)
10—15	04	04
15—20	05	09
20—25	09	18 c
25—30	10 (f)	28
30—35	10	38
35—40	07	45
40—45	05	50
	$\sum f = N =$ 50	

$$\therefore m = \frac{N}{2} = \frac{50}{2} = 25$$

इसका अर्थ यह है कि माध्यिका ज्ञात करने के लिये 25वें पद का मान निकालना है—

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

$$= 25 + \frac{5}{10} \times (25 - 18)$$

$$= 25 + \frac{1}{2} \times 7$$

$$= 25 + 3.5$$

$$M = 28.5$$

उदाहरण : निम्नांकित वर्गीकृत आँकड़ों से माध्यिका ज्ञात कीजिए।

वर्ग	आवृत्ति (f)
30—39	08
40—49	18
50—59	30
60—69	24
70—79	20

हल :

वर्ग	आवृत्ति (f)	संचयी आवृत्ति (F)
30—39	08	08
40—49	18	26 c
50—59	30 f	56
60—69	24	80
70—79	20	100
	$\sum f = N =$ 100	

$\therefore m = \frac{N}{2} = \frac{100}{2}$ = 50वें पद का मान ज्ञात करना है।

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

$$= 50 + \frac{9}{30}(50 - 26)$$

$$= 50 + \frac{9}{30} \times 24$$

$$= 50 + 7.2$$

$$= 57.2$$

$$\therefore M = 57.2$$

अभ्यास

6. निम्नांकित आँकड़ो से माध्यिका ज्ञात कीजिए।

वर्ग	आवृत्ति (f)
10—20	06
20—30	10
30—40	14
40—50	16
50—60	12
60—70	09
70—80	06

7. निम्न वर्गीकृत आँकड़ों के लिये माध्यिका ज्ञात कीजिए।

वर्ग	आवृत्ति (f)
150—159	03
160—169	05
170—179	08
180—189	12
190—199	22
200—209	18
210—219	16
220—229	10
230—239	06

बहुलक

किसी श्रेणी में जिस मान की सर्वाधिक पुनरावृति होती है। वह मान **बहुलक** कहलाता है। इसके संकेताक्षर Z अथवा M_0 है। माध्य तथा माध्यिका की तुलना में बहुलक का उपयोग कम प्रचलित है। किसी श्रेणी में एक से अधिक बहुलक भी हो सकते हैं।

अवर्गीकृत आँकड़ों के लिए बहुलक की गणना

दिए हुए आँकड़ों 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 : एक फैक्टरी में कुछ श्रमिकों की दैनिक मजदूरी 15, 44, 17, 29, 35, 44, 15, 67, 35, 44, 36, 44, 21, 38 रूपये है। बहुलक मजदूरी ज्ञात कीजिए।

हल : इन आँकड़ो को क्रमबद्ध जमाने पर

$\underline{15, 15,}$ 17, 21, 29, $\underline{35, 35,}$ 36, 38, $\underline{44, 44, 44, 44,}$ 67

इससे हम देखते हैं कि 44 मूल्य सबसे अधिक बार आया है, अतः यही बहुलक है।

विशेष : यदि समंकों में दो मूल्यों की आवृत्तियाँ समान हो तो इन्हें **द्वि–बहुलक (Bi-modal)**, तीन मूल्यों की आवृत्तियाँ समान होने पर **त्रि–बहुलक (Tri-modal)** अथवा अनेक मूल्यों की आवृत्तियाँ समान होने पर **विविध–बहुलक (Multi-modal)** श्रेणियाँ कहते हैं। यदि दी गई श्रेणी के किसी भी मूल्य में पुरावृत्ति नहीं होती तो वह **बहुलक –रहित (Without-modal)** श्रेणी होती है।

अभ्यास

8. निम्नांकित आँकड़ो से बहुलक ज्ञात कीजिए–

40, 42, 60, 14, 04, 12, 25, 20, 40, 25, 18, 10, 25

प्रकीर्णन

प्रकीर्णन से तात्पर्य केन्द्रीय प्रवृत्ति के माप से इकाईयों 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 केन्द्रीय मान से विभिन्न मूल्यों के बिखराव अथवा विषमता की मात्रा है।

प्रकीर्णन के मापन की निम्नलिखित विधियाँ हैं –

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

इनमें से प्रत्येक विधि के अपने विशेष गुण एवं सीमाएँ हैं। अतः इनमें से किसी भी विधि का प्रयोग 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 है। 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 आँकड़ों के लिए विस्तार की गणना

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

रू. 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' क्रमशः अधिकतम तथा न्यूनतम मान के प्रतीक है।

अतः

$$R = L - S$$

$= \quad 100 - 40 = 60$

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

अभ्यास

9. निम्नांकित मूल्यों के लिए विस्तार की गणना कीजिए :

   50, 30, 70, 90, 25, 20, 100

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।

अवर्गीकृत आँकड़ों के लिये मानक विचलन की गणना

अवर्गीकृत आँकड़ो के लिये मानक विचलन की गणना करना आसान है। इसकी गणना विधि के निम्न चरण हैं –

(i) सबसे पहले समंक श्रेणी का माध्य ज्ञात किया जाता है।

(ii) उसके पश्चात् प्रत्येक मूल्य में से माध्य घटाकर विचलन ज्ञात किये जाते हैं। माध्य $(\overline{X})$ से विचलन की सांख्यिकी में x प्रदर्शित किया जाता है, इसलिये सूत्र के रूप में इसकी अभिव्यक्ति इस प्रकार की जाती है–

$$x = (X - \overline{X})$$

(iii) फिर सभी विचलनों के वर्ग अर्थात् x^2 ज्ञात किये जाते हैं।

(iv) इसके बाद सभी विचलनों के वर्गों को जोड़ लिया जाता है, अर्थात $\sum x^2$ ज्ञात किया जाता है।

(v) विचलन के वर्गों के योग $\left(\sum x^2\right)$ में कुल पदों (N) का भाग दिया जाता है, अर्थात $\frac{\sum f}{x}$ ज्ञात किया जाता है।

(vi) भाग देने से मूल्य का वर्गमूल निकाला जाता है, अर्थात $\sqrt{\frac{\sum x^2}{N}}$ ज्ञात किया जाता है। यही मानक विचलन होता है।

अतः मानक विचलन की गणना निम्न सूत्र के आधार पर की जाती है।

$$S.D. \text{ या } \sigma = \sqrt{\frac{\sum x^2}{N}}$$

जिसमें

$S.D.$ या σ = मानक विचलन

$\sum x^2$ = माध्य से विचलनों के वर्गों का योग, जो कि ऊपर बिन्दु संख्या $(ii), (iii)$ तथा (iv) में समझाया गया है।

N = पदों की संख्या

उदाहरण : दस विद्यार्थियों के भूगोल के टैस्ट में निम्नलिखित आँकड़ों के लिये मानक विचलन ज्ञात कीजिए–

3, 5, 8, 12, 16, 13, 8, 4, 21, 10

हल :

X	$x=(X-10)$	x^2
03	− 7	49
05	− 5	25
08	− 2	4
12	2	4
16	6 − 22	36
13	3	9
08	− 2	4
04	− 6 + 22	36
21	11	121
10	0	0
$\sum X = 100$		$f\,x'^2$

$N = 10$

$\therefore \bar{X} = 10$

मानक विचलन

$$\sigma = \sqrt{\frac{\sum x^2}{N}}$$

$$= \sqrt{\frac{288}{10}}$$

$$= \sqrt{28.8}$$

$= 5.37$ या 5.4

अभ्यास

10. दस श्रमिकों के पारिश्रमिक निम्नानुसार हैं। इनका मानक विचलन ज्ञात कीजिए।

42, 38, 36, 45, 50, 58, 52, 55, 65, 59

वर्गीकृत समंकों के लिये मानक विचलन की गणना

वर्गीकृत आँकड़ों से मानक विचलन ज्ञात करने की अनेक विधियाँ हैं। इनमें से केवल एक सबसे आसान विधि को यहाँ समझाया गया है। इसकी गणना करने का सूत्र निम्नानुसार है–

$$\sigma = i \times \sqrt{\frac{\sum fx'^2}{N} - \left(\frac{\sum fx'}{N}\right)^2}$$

जिसमें

σ = मानक विचलन

i = वर्ग विस्तार

f = आवृत्ति

N = आवृत्तियों का योग अर्थात $\sum f$

x' = कल्पित माध्य से पद विचलन

उदाहरण : निम्नांकित वर्गीकृत आँकड़ों के लिये मानक विचलन ज्ञात कीजिए।

वर्ग	110–120	120–130	130–140	140–150	150–160
आवृत्ति	2	3	7	5	3

हल :

वर्ग	आवृत्ति f_1	x' (पद विचलन)	fx' (f व x' का गुणा)	($f\,x'^2$ व fx' का गुणा)
110—120	2	- 2	- 4	8
120—130	3	- 1	- 3	3
130—140	7	0	0	0
140—150	5	1	5	5
150—160	3	2	6	12

$\sum f = N = 20$ $\sum fx' =$ $\sum fx'^2 = 28$

मानक विचलन का सूत्र

$$\sigma = i \times \sqrt{\frac{\sum fx'^2}{N} - \left(\frac{\sum fx'}{N}\right)^2}$$

सूत्र में उपरोक्त मान रखने पर

$$\sigma = 10 \times \sqrt{\frac{28}{20} - \left(\frac{4}{20}\right)^2}$$

$$= 10 \times \sqrt{1.4 - (0.2)^2}$$

$$= 10 \times \sqrt{1.4 - 0.04}$$

$$= 10 \times \sqrt{1.36}$$

$$= 10 \times 1.166$$

$= 11.66$ अथवा 11.7

अभ्यास

11. निम्नांकित वर्गीकृत आँकड़ो के लिये मानक विचलन की गणना कीजिए–

वर्ग	आवृत्ति (f)
05—09	02
10—14	04
15—19	14
20—24	18
25—29	22
30—34	20
35—39	12
40—44	08

12. निम्नांकित आँकड़ो के लिये मानक विचलन ज्ञात कीजिए–

वर्ग	आवृत्ति (f)
10—20	09
20—30	12
30—40	14
40—50	18
50—60	16
60—70	12
70—80	08

विचरण गुणांक (CV)

यदि आँकड़ें माप की अलग–अलग इकाईयों 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 माध्य को माध्य के प्रतिशत के रूप में अभिव्यक्त करता है। इसका निर्धारण निम्नांकित सूत्र के उपयोग द्वारा होता है:

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

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

उदाहरण : यदि किसी श्रेणी में मानक विचलन $(\sigma) = 2.83$ एवं माध्य $(\overline{X}) = 5$ हो तो विचरण गुणांक (CV) ज्ञात कीजिए।

हल :

$$\text{विचरण गुणांक } (CV) = \frac{\sigma}{\overline{X}} \times 100$$

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

$$CV = 56\%$$

अभ्यास

13. यदि किसी श्रेणी में मानक विचलन (σ) का मान 5.37 एवं माध्य $(\overline{X})$ का मान 10 हो तो उस श्रेणी का विचरण गुणांक (CV) ज्ञात कीजिए।

स्पीयरमैन का कोटि आकार सह–सम्बन्ध (Spearman's Rank Order Correlation)

सह–सम्बन्ध (Correlation)

जब कोई दो चर मूल्य (Variables) परस्पर एक दिशा में या विपरीत दिशा में घटने–बढ़ने की प्रवृत्ति रखते हैं, तो ऐसी स्थिति में उनके मध्य एक विशेष प्रकार का सम्बन्ध पाया जाता है। इस सम्बन्ध को ही **सह–सम्बन्ध** कहते हैं।

सह–सम्बन्ध के पहलू

सह–सम्बन्ध के दो पहलू बहुत महत्वपूर्ण है– (1) सह–सम्बन्ध की दिशा (Direction of Correlation) तथा (2) सह–सम्बन्ध की गहनता (Degree of Correlation)।

1. **सह–सम्बन्ध की दिशा** : सह–सम्बन्ध दो प्रकार का होता है – **धनात्मक तथा ऋणात्मक**। जब दो समंक चर परस्पर एक ही दिशा में विचरण करते हैं अर्थात् एक चर में वृद्धि (या कमी) होने पर दूसरे चर में भी वृद्धि (या कमी) हो तो उसे **धनात्मक सह–सम्बन्ध (Positive Correlation)** कहते हैं।

किन्तु यदि एक चर में वृद्धि होने पर दूसरे चर में कमी हो अथवा एक चर में कमी होने पर दूसरे चर में वृद्धि हो तो दोनों चरों के परस्पर विपरीत गमन के कारण यह **ऋणात्मक सह–सम्बन्ध (Negative Correlation)** कहलाता है। इस प्रकार सह–सम्बन्ध की दो दिशाएँ होती है– धनात्मक व ऋणात्मक। अतः सह–सम्बन्ध की गणना करने पर प्राप्त उत्तर में दिशा का प्रतीक चिन्ह (+ या –) लगाना अत्यन्त आवश्यक होता है। यदि किसी संख्या से पूर्व कोई चिन्ह लगा हुआ नहीं हो तो उसे धनात्मक मान लिया जाता है। इसलिये ऋणात्मक सह–सम्बन्ध प्राप्त होने पर **ऋणात्मक चिन्ह (–) लगाने का विशेष ध्यान रखना चाहिये**। वस्तुतः सह–सम्बन्ध की मात्रा पर प्रतीकात्मक चिन्ह, विशेष रूप से ऋणात्मक चिन्ह, नहीं लगाने पर उत्तर त्रुटिपूर्ण हो जाता है।

2. **सह–सम्बन्ध की गहनता** : सह–सम्बन्ध की गहनता के मापक का विस्तार 0 से 1 होता है। मात्रात्मक दृष्टि से सह–सम्बन्ध का अभाव (No Correlation) 0 से प्रदर्शित होता है तथा घनिष्ठतम सह–सम्बन्ध (Strongest Correlation) 1 से प्रदर्शित होता है। घनिष्ठतम सह–सम्बन्ध को पूर्ण सह–सम्बन्ध (Perfect Correlation) कहते हैं। इस प्रकार सह–सम्बन्ध की गहनता 0 से 1 के मध्य रहती है। यदि इस गहनता में सह–सम्बन्ध की दिशा का पहलू जोड़ दे तो सह–सम्बन्ध की गहनता के मापक विस्तार 1 से शून्य पर गुजरते हुए –1 तक रहता है जिसमें –1 **पूर्ण ऋणात्मक सह–सम्बन्ध (Perfect Negative Correlation)**, 0 सह–सम्बन्ध

के अभाव **(No Correlation)** तथा +1 **पूर्ण धनात्मक सह–सम्बन्ध (Perfect Positive Correlation)** का प्रतीक होता है।

सावधानियाँ (Cautions) : सह–सम्बन्ध की गणना करने में तथा उत्तर लिखने में विशेष रूप से दो सावधानियाँ रखना अति–आवश्यक है–

(i) सह–सम्बन्ध पर दिशा सूचक चिन्ह (+ या –) लगाना विशेष रूप से यदि सह–सम्बन्ध ऋणात्मक (Negative) हो।

(ii) सह–सम्बन्ध **किसी भी परिस्थिति में एक से अधिक नहीं हो सकता।** यदि ऐसा हो रहा है तो निश्चित रूप से गणना करने में किसी न किसी प्रकार की त्रुटि हुई है। इस त्रुटि को ढूंढ कर दूर कर लेना चाहिये।

गणना की विधि (Method of Calculation) :

गणना : सह–सम्बन्ध कई प्रकार के होते हैं। किन्तु इस अध्याय में केवल **स्पीयरमैन के कोटि आकार सह–सम्बन्ध** को ही सम्मिलित किया गया है। इसकी गणना करना बहुत आसान है। इस विधि का प्रतिपादन स्पीयरमैन ने किया था, इसीलिये इस विधि को **स्पीयरमैन का सह–सम्बन्ध** कहा जाता है। इसे सांख्यिकी में p या *Rho* से व्यक्त किया जाता है। इसकी गणना को निम्न चरणों में समझा जा सकता है–

(i) सह–सम्बन्ध की गणना दो चरों (Variables) के मध्य की जाती है। इनहे सांख्यिकी में X तथा Y से सम्बोधित किया जाता है। इनमें X **स्वतंत्र चर** (Independent Variable) तथा Y **निर्भर चर** (Dependent Variable) होता है। उदाहरण के लिये विभिन्न स्थानों का समुद्रतल से ऊँचाई तथा उन स्थानों के तापमान सह–सम्बन्ध की गणना करने के लिये दिये हुए है, तो चूंकि तापमान समुद्रतल से ऊँचाई पर निर्भर करता है, इसलिये ऊँचाई (X) तथा तापमान (Y) चर होंगे। इसमें समुद्रतल से ऊँचाई स्वतंत्र चर तथा तापमान निर्भर चर है।

(ii) X तथा Y दोनों चरों के समंकों की कोटियाँ (Ranks) निर्धारित की जाती है। X चर में जिसका मान सर्वाधिक होगा, उसके कोटि 1 (Ranks 1) तथा घटते मान के साथ क्रमशः 2, 3, 4. .... कोटि अंक प्रदान करते हैं। ये सभी X की कोटियाँ (Ranks) अर्थात् X_R है। इसी प्रकार Y_R के मान भी निर्धारित किये जाते हैं।

महत्वपूर्ण : कभी–कभी किसी श्रेणी में एक से मान एक से अधिक हो सकते हैं। उदाहरण के लिये तालिका में X का मान देखिये। इसमें सर्वाधिक मान 20 है, अतः उसे कोटि 1, उसके बाद 16 को कोटि 2 देंगे। लेकिन उसके बाद 15 का मान **दो बार** आया है, जिसकी कोटियाँ 3 व 4 है, अतः इन्हें जोड़कर **दो** से भाग देने पर

प्राप्त मान 3.5 $\left(\frac{3+4}{2}\right)$ को इनकी कोटियाँ माना जायेगा। इसके बाद के मान 12 की कोटि 5 होगी क्योंकि दो बार घटित 15 को 3 व 4 कोटियाँ दी जा चुकी हैं।

तालिका 1

X	X_R
8	6
15	3.5
2	8
12	5
16	2
20	1
7	7
15	3.5

तालिका 2

Y	Y_R
10	6
11	3.5
6	8
10	6
14	2
16	1
10	6
11	3.5

अब तालिका संख्या 2 में Y के मान देखिये। इसमें 11 की कोटियों को 3.5 देने के बाद अगला मान 10 भी तीन बार आया है। इन सभी की कोटियाँ $\left(\frac{5+6+7}{3}\right)6-6$ होंगी तथा अगले मान 6 की कोटि 8 होगी क्योंकि 7 तक कोटियाँ आवंटित की जा चुकी है।

(iii) इस प्रकार दोनों चरो की कोटियाँ निर्धारित करके उनका अन्तर (Difference) ज्ञात किया जाता है। इसे D से सम्बोधित किया जाता है।

(iv) बाद में D का वर्ग (Square) ज्ञात किया जाता है जिसे D^2 द्वारा सम्बोधित किया जाता है।

(v) इसके पश्चात् निम्न सूत्र से सह–सम्बन्ध की गणना की जाती है। इसका अवकलन (Calculation) बहुत आसान है।

$$Rho / p = 1 - \frac{6\sum D^2}{N\left(N^2-1\right)}$$

जिसमें

$D^2 =$ दोनों चरों की कोटियों के अन्तर का वर्ग (Square of the difference of the two ranks)

$N =$ युग्मों की संख्या (Number of Pairs)

उदाहरण : निम्नांकित चरो का कोटि आकार सह–सम्बन्ध ज्ञात कीजिए–

X - 12, 27, 9, 35, 18, 16, 19, 31, 6, 22

Y - 6, 14, 5, 28, 17, 9, 25, 21, 2, 11

हल :

X	Y	X_R	Y_R	D	D^2
12	6	8	8	0	0
27	14	3	5	2	4
9	5	9	9	0	0
35	28	1	1	0	0
18	17	6	4	2	4
16	9	7	7	0	0
19	25	5	2	3	9
31	21	2	3	1	1
6	2	10	10	0	0
22	11	4	6	2	4
$N=10$					$\sum D^2 = 22$

गणना –

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

$$= 1 - \frac{6 \times 22}{10(10^2 - 1)}$$

$$= 1 - \frac{132}{10(100 - 1)}$$

$$= 1 - \frac{132}{10(99)}$$

$$= 1 - \frac{132}{990}$$

$$= 1 - 0.13$$

$$= 0.87$$

अर्थात दोनों चरों के मध्य अत्यन्त **घनिष्ठ धनात्मक सह–सम्बन्ध है।**

उदाहरण : निम्नांकित चरों का कोटि आकार सह–सम्बन्ध ज्ञात कीजए –

X – 32, 25, 28, 8, 19, 25, 3, 25

Y – 14, 5, 2, 16, 14, 9, 18, 11

हल :

X	Y	X_R	Y_R	D	D^2
32	14	1	3.5	2.5	6.25
25	5	4	7.0	3.0	9.00
28	2	2	8.0	6.0	36.00
8	16	7	2.0	5.0	25.00
19	14	6	3.5	2.5	6.25
25	9	4	6.0	2.0	4.00
3	18	8	1.0	7.0	49.00
25	11	4	5.0	1.0	1.00
	$N = 8$				$\sum D^2 = 136.5$

गणना

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

$$= 1 - \frac{6 \times 136.5}{8(8^2 - 1)}$$

$$= 1 - \frac{819}{8(64 - 1)}$$

$$= 1 - \frac{819}{8(63)}$$

$$= 1 - \frac{819}{504}$$

$$= 1 - 1.625$$

-0.625 अथवा 0.63

अर्थात दोनों चरों के मध्य **मध्यम ऋणात्मक सह–सम्बन्ध** (Moderate Negative Correlation) है।

उदाहरण : निम्नांकित चरों का कोटि आकार सह–सम्बन्ध ज्ञात कीजिये –

X – 110, 57, 63, 98, 48, 57, 84, 57

Y – 117, 196, 183, 145, 201, 199, 167, 201

हल :

X	**Y**	**X_R**	**Y_R**	**D**	**D^2**
110	117	1	8.0	7.0	49.00
57	196	6	4.0	2.0	4.00
63	183	4	5.0	1.0	1.00
98	145	2	7.0	5.0	25.00
48	201	8	1.5	6.5	42.25
57	199	6	3.0	3.0	9.00
84	167	3	6.0	3.9	9.00
57	201	6	1.5	4.3	20.25
	$N = 8$				$\sum D^2 = 159.5$

गणना

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

$$= 1 - \frac{6 \times 159 - 5}{8(8^2 - 1)}$$

$$= 1 - \frac{957}{8(64-1)}$$

$$= 1 - \frac{957}{8(63)}$$

$$= 1 - \frac{957}{504}$$

$$= 1 - 1.9$$

$= -0.9$ उत्तर

अर्थात दोनों चरों के मध्य **अत्यन्त घनिष्ठ ऋणात्मक सह–सम्बन्ध (Very Strong Negative Correlation)** है।

अभ्यास

14. निम्नांकित चरों के आधार पर कोटि आकार सह–सम्बन्ध ज्ञात कीजिए।

X	08	10	40	80	12	64	14	20
Y	06	14	48	16	10	54	25	11

15. निम्नांकित चरों के आधार पर कोटि आकार सह–सम्बन्ध ज्ञात कीजिए।

X	64	72	10	24	32	40	90	25	14	50
Y	03	06	10	12	18	20	24	30	32	40

□□□