উদাহরণ ঃ নিচের ছকের সাহায্যে মধ্যমা গণনা করুন।

নম্বর ঃ	0-10	10-30	30-60	60-70	70-90
ছাত্রসংখ্যা ঃ	15	25	30	4	10

সমাধান ঃ

মার্ক	ছাত্রসংখ্যা	ক্রমযৌগিক পরিসংখ্যা
0-10	15	15 মধ্যমা $= \frac{84}{2}$ তম পদের আকার
10-30	25	40 $= 42$ " " "
30-60	30	70 $\therefore$ মধ্যমা (30-60) শ্রেণীতে অবস্থিত।
60-70	4	74
70-90	10	84 (N)

$\therefore\ l_1 = 30,\ M = 42,\ C = 40,\ fm = 30,\ i = 60 - 30 = 30$

$\therefore$ মধ্যমা $= 30 + \frac{42-40}{30} \times 30 = 30 + 2 = 32$

টীকা ঃ যদি শ্রেণীবিভাগের দৈর্ঘ্য সমান করে বিভাজন তৈরী করা হয় (পরিসংখ্যা দরকার মত পাল্টিয়ে) সেক্ষেত্রেও আমরা মধ্যমার একই মান পাব। নীচের উদাহরণটি লক্ষ্য করুন।

নম্বর	পরিসংখ্যা	ক্রমযৌগিক পরিসংখ্যা
0-10	15	15 মধ্যমা = 42-তম পদের আকার
10-20	12.5	27.5 (30-40) শ্রেণীতে মধ্যমা অবস্থান করছে
20-30	12.5	46
30-40	10	50 $l_1 = 30,\ M = 42,\ C = 40,\ fm = 10,\ i = 10$
40-50	10	60
50-60	10	70 $\therefore$ মধ্যমা $= 30 + \frac{42-40}{10} \times 10$
60-70	4	74
70-80	5	79 $= 30 + 2 = 32$ নম্বর
80-90	5	84 (N)

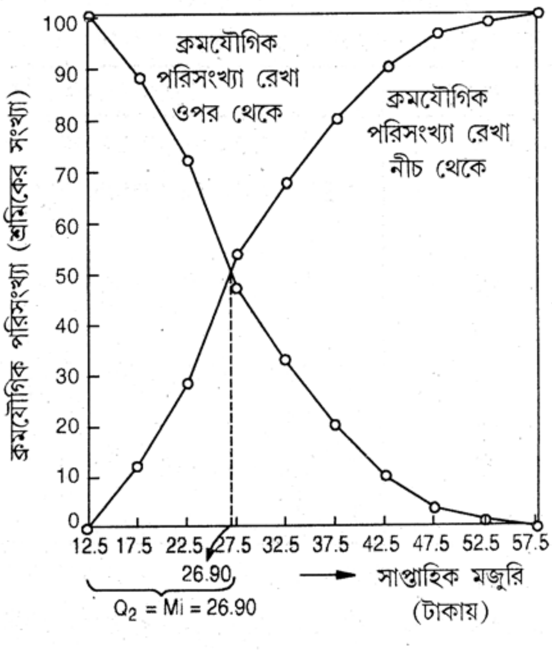
4.4.8 লেখ (Graph)-এর সাহায্যে মধ্যমা নির্ণয় ঃ

দুটি ক্রমযৌগিক পরিসংখ্যা (কম থেকে বেশী less than type বা বেশী থেকে কমের দিকে more than type) রেখা (ogive) অঙ্কণ করেও মধ্যমা নির্ণয় করা যায়। দু'টি রেখার ছেদবিন্দু থেকে x-অক্ষে উলম্ব (Vertical) রেখা আঁকতে হবে। x-অক্ষে ছেদবিন্দু-2 হবে মধ্যমা। নীচের উদাহরণ 3 রেখাচিত্রটি লক্ষ্য করুন।

উদাহরণ ঃ

প্রদত্ত রাশিতথ্যে একটি কারখানার 100-জন শ্রমিকের সাপ্তাহিক মজুরীর পরিমাণ দেওয়া আছে, এই তথ্যরাশি থেকে লেখচিত্রের সাহায্যে মধ্যমা দেখান।

সাপ্তাহিক মজুরী	শ্রমিকের সংখ্যা (f)	ক্রমযৌগিক পরিসংখ্যা (cf)	
		নীচ থেকে	উপর থেকে
12.5-17.5	12	12	100 n
17.5-22.5	16	28	88
22.5-27.5	25	53	72
27.5-32.5	14	67	47
32.5-37.5	13	80	33
37.5-42.5	10	90	20
42.5-47.5	6	96	10
47.5-52.5	3	99	4
52.5-57.5	1	100	1
মোট	$\Sigma fi = n = 100$		



লেখচিত্র : প্রদত্ত পরিসংখ্যা বিভাজনের ক্রমযৌগিক পরিসংখ্যা রেখা।

লেখচিত্রটি লক্ষ্য করুন। এখানে দুটি যৌগিক পরিসংখ্যা (উপর থেকে নীচে ও নীচে থেকে ওপরে) একটি বিন্দুতে মিলছে যেখান থেকে উলম্বভাবে নীচে একটি রেখা টানা হয়েছে (x-অক্ষরেখার দিকে)। এটি x-অক্ষরেখার 26.9 মান ছেদ করেছে। অতএব 26.9 হল মধ্যমা।

সূত্রের সাহায্যেও আমরা একই প্রকার মধ্যমার মান পাই। নীচের সমাধানটি লক্ষ্য করুন।

সমাধান ঃ

$l_1 + \frac{N/2 - C}{fm} \times i$ [এখানে $l_1 = 22.5$; N/2 50, C = 28; i = 5; fm = 25]

$= 22.5 + \frac{50-28}{25} \times 5 = 22.5 + \frac{22 \times 5}{25} = 22.5 + 4.4 = 26.5$।

4.4.9 পরিসংখ্যান গড় হিসেবে মধ্যমা ব্যবহারের সুবিধে ও অসুবিধে (Merits and Demerits of Median as a Statistical Average)

সুবিধে ঃ

(i) যৌগিক গড়ের মত মধ্যমা সহজে নির্ণয় করা যায় ও সহজে বোঝা যায়।

(ii) কোন কোন ক্ষেত্রে মধ্যমা রাশিতথ্যমালা পর্যবেক্ষণ করে নির্ণয় করা যায়।

(iii) শ্রেণীর অন্তিম সংখ্যা দেওয়া না থাকলেও মধ্যমা নির্ণয় করা যায়।

(iv) লেখ (Graph)-র সাহায্যে মধ্যমা নির্দেশ করা যায়।

(v) এর মান রাশিতথ্যমালার সামান্য কয়েকটি অতি উচ্চ বা অতি নিম্নমানের রাশির দ্বারা প্রভাবিত হয় না। এর মান রাশিতথ্যমালার প্রসার অথবা তার মানের নীচে বা ওপরে রাশিতথ্যমালার মানের বিস্তৃতির ওপর নির্ভর করে না।

(vi) যদি শ্রেণীবিভাগের দৈর্ঘ্য পরস্পর সমান না হয় অথবা প্রান্তীয় শ্রেণীবিভাগ মুক্ত (open) থাকে, তবে তার দ্বারা মধ্যমার মান প্রভাবিত হয় না। এক্ষেত্রে যৌগিক গড় অপেক্ষা মধ্যমার ব্যবহার খুব সুবিধেজনক।

(vii) যে সব তথ্য সংখ্যার সাহায্যে প্রকাশ করা যায় না (যেমন বুদ্ধি, সততা ইত্যাধি তাদের বৈশিষ্ট্যের পরিমাণ নিরূপণে মধ্যমার ব্যবহার খুব সুবিধেজনক।

(viii) মধ্যমা রাশিতথ্যমালার সর্বমধ্য পদটির মাঝকে প্রকাশ করে। তাই এর মান রাশিতথ্যমালার কেন্দ্রীয় প্রবণতার সঠিক পরিমাপক বলে মনে করা হয়।

অসুবিধে ঃ

(i) মধ্যমার মান নির্ণয়ে রাশিতথ্যমালার মানসমূহ মানের উর্ধ্বক্রম (বা অধঃক্রম) অনুযায়ী সাজাতে হয়। এটি খুব কষ্টসাধ্য ব্যাপার।

(ii) এর মান নিরূপণে বীজগণিতের নিয়মাবলীর সহজ প্রয়োগ সম্ভব নয়।

(iii) অনেক সময় এর সঠিক মান নির্ণয় করা সম্ভব হয় না; শুধু আনুমানিক মান নির্ণয় করা যায়।

(iv) এর মান নির্ণয়ে রাশিতথ্যমালার প্রত্যেকটি রাশি ব্যবহৃত হয় না।

(v) দুই বা ততোধিক গ্রুপের মিশ্রিত মধ্যমা গণনা করা যায় না।

(vi) রাশিতথ্যের সব মানের ওপর মধ্যমা নির্ভর করে না।

(vii) এর নমুনা বিচ্যুতি অন্যান্য গড়ের তুলনায় বেশী।

4.5 ভগ্নাংশক-চতুর্থক, দশমক এবং শততমক (Partition values-Quartiles, Deciles, Percentiles)

আমরা জানি যে একটি চলরাশির মানগুলোকে যদি ছোট থেকে বড়ো (ঊর্ধ্বক্রমে) বা বড়ো থেকে ছোট (নিম্নক্রমে) পর পর সাজানো যায় তবে মধ্যমা (Median) সেই সারিবদ্ধ শ্রেণী বিভাজনকে দু'টি সমান ভাগে ভাগ করে। সেই রকম আরো কিছু পরিমাপ আছে যারা একটি শ্রেণী বিভাজনকে নির্দিষ্ট কতগুলো সমান অংশে (যেমন 4, 10 বা 100) ভাগ করে। **এগুলিকে সামগ্রিকভাবে ভগ্নাংশক বলে।**

ভগ্নাংশগুলির মধ্যে উল্লেখযোগ্য হল ঃ

(i) **মধ্যমা** (Median), (ii) **চতুর্থক** (Quartiles), (iii) **দশমক** (Decile), (iv) **শততমক** (Percentiles)।

(i) **চতুর্থক** (Quartiles) ঃ একটি চলরাশির মানগুলো ঊর্ধ্বক্রমে সাজালে মধ্যমা (Median) শ্রেণীটিকে সমান দু'টি অংশে বিভক্ত করে। এইভাবে চতুর্থকগুলো (quartiles) একটি চলরাশির সারিবদ্ধ (ছোট থেকে বড়ো) মানগুলোকে সমান চারটি ভাগে ভাগ করে। তিনটি চতুর্থক যথাক্রমে Q_1, Q_2, Q_3 সমগ্র শ্রেণীকে সমান চার অংশে ভাগ করে।

25%	25%	25%	25%
	Q_1	Q_2	Q_3

প্রথম বা নিম্ন চতুর্থক (Q_1)-এর বামদিকে 25% ও ডানদিকে বাকী 75% মান বণ্টিত থাকে। দ্বিতীয় চতুর্থক (Q_2) হল মধ্যমা যার বামদিকে 50% ও ডানদিকে বাকী 50% মান বণ্টিত থাকে এবং তৃতীয় বা ঊর্ধ্ব চতুর্থক (Q_3) হল এমন যার বামদিকে 75% ও ডানদিকে মাত্র 25% মান বণ্টিত থাকে। চতুর্থকের মত রাশিতথ্যমালার মানগুলো ঊর্ধ্বক্রম অনুযায়ী সমান দশটি ভাগে বিভক্ত করে **দশমক** (Decile) এবং একশতটি সমান ভাগে বিভক্ত করে **শততমক** (Percentile) এর সংজ্ঞা দেওয়া হয়।

দশমকে D_1, D_2, D_3 এবং শততমকে P_1, P_2, P_3, P_{50}, P_{55} ইত্যাদি দিয়ে চিহ্নিত করা হয়।

একটি বিষয় মনে রাখা দরকার যে $\mathbf{Q_2 = D_5 = P_{50}}$ **এর প্রত্যেকেই মধ্যমা।**

(2) **চতুর্থক, দশমক ও শততমকের গণনা** (Calculation of quartitles, Deciles and Percentiles) ঃ মধ্যমা যেভাবে গণনা করা হয় ঠিক একই পদ্ধতিতে চতুর্থক, দশমক ও শততমক গণনা করা হয়। সরল রাশিতথ্যমালার (বা সরল পরিসংখ্যা বিভাজনের) ক্ষেত্রে প্রথমে রাশিতথ্যগুলোকে তাদের মান অনুযায়ী ঊর্ধ্বক্রমে সাজানো হয়।

এখন $Q_1 = \frac{n+1}{4}$-তম রাশির মান;

$Q_2 = \frac{2(n+1)}{4} = n + \frac{1}{2}$-তম রাশির মান;

$Q_3 = \frac{3(n+1)}{4}$-তম রাশির মান;

$$D_1 = \frac{(n+1)}{10}\text{-তম রাশির মান;}$$

$$D_2 = \frac{2(n+1)}{10}\text{-তম রাশির মান;}$$

$$\vdots \quad \vdots \quad \vdots$$

$$D_5 = \frac{5(n+1)}{10} = \frac{n+1}{2}\text{-তম রাশির মান;}$$

$$\vdots \quad \vdots \quad \vdots$$

$$D_9 = \frac{9(n+1)}{10}\text{-তম রাশির মান;}$$

$$P_1 = \frac{(n+1)}{100}\text{-তম রাশির মান;}$$

$$P_2 = \frac{2(n+1)}{100}\text{-তম রাশির মান;}$$

$$\vdots \quad \vdots \quad \vdots$$

$$P_{50} = \frac{50(n+1)}{100} = \frac{n+1}{2}\text{-তম রাশির মান;}$$

$$\vdots \quad \vdots \quad \vdots$$

$$P_{99} = \frac{99(n+1)}{100}\text{-তম রাশির মান।}$$

[n = মোট রাশির সংখ্যা]

নীচের উদাহরণটি লক্ষ্য করুন।

উদাহরণ ঃ নিম্নে বর্ণিত ওজন (কিলো) থেকে Q_1, Q_3, D_3, D_4 এবং P_{50} নির্ণয় করুন।

19, 27, 24, 39, 57, 44, 56, 50, 59, 67, 62, 42, 47, 60, 26, 34, 57, 51, 59, 45।

সংখ্যাগুলোকে এইভাবে সাজানো হল—

ক্রমিক সংখ্যা	ওজন (কিলো)	ক্রমিক সংখ্যা	ওজন (কিলো)	ক্রমিক সংখ্যা	ওজন (কিলো)
1	19	8	44	15	57
2	24	9	45	16	59
3	26	10	47	17	59
4	27	11	50	18	60
5	34	12	51	19	63
6	39	13	56	20	67
7	42	14	57		

এখানে $n = 20$

Q_1 (**প্রথম চতুর্থক**) $= \frac{n+1}{4}$-তম পদের মান $= \frac{20+1}{4}$-তম পদের মান

$= 5.25$-তম পদের মান

$= 5$-তম পদের মান $+ \frac{1}{4}$ (6-তম পদের মান – 5-তম পদের মান)

$= 34 + \frac{1}{4}(39 - 34) = 34 + 1.25 = 35.25$ কিলো।

Q_2 (**দ্বিতীয় চতুর্থক**) $= \frac{3(n+1)}{4}$-তম পদের মান $= 3\left(\frac{20+1}{4}\right)$-তম পদের মান

$= 15.75$-তম পদের মান; $= 15$-তম পদের মান $+ \frac{3}{4}$

(16-তম পদের মান – 15-তম পদের মান)

$= 57 + \frac{3}{4}(59 - 57) = 57 + 1.50 = 58.50$ কিলো।

$= 4\left(\frac{n+1}{10}\right)$-তম পদের মান $= 4\left(\frac{20+1}{10}\right)$-তম পদের মান

$= 8.4$-তম পদের মান

$= 8$-তম পদের মান $+ \frac{4}{10}$ (9-তম পদের মান – 8-তম পদের মান)

$= 44 + \frac{4}{10}(45 - 44) = 44 + .4 = 44.4$ কিলো।

$P_{60} = \frac{60(n+1)}{100}$-তম পদের মান $= \frac{60(20+1)}{100}$-তম পদের মান

$= 12.60$-তম পদের মান

$= 12$-তম পদের মান $+ \frac{60}{100}$ (13-তম পদের মান – 12-তম পদের মান)

$= 51 + \frac{60}{100}(56 - 51) = 51 + 3 = 54$ কিলো।

(B) **সরল পরিসংখ্যা বিভাজনের ক্ষেত্রে (For simple frequency distribution) ঃ**

এক্ষেত্রে প্রথমে ক্রমযৌগিক পরিসংখ্যা নির্ণয় করতে হবে, যা নীচের উদাহরণে দেখানো হল।

ওজন (কিলো)	পরিসংখ্যা	ক্রমযৌগিক পরিসংখ্যা
40	2	2
42	6	8
45	8	16
50	10	26
51	6	32
54	14	46
56	12	58
59	8	66
60	14	80
62	12	92
64	6	98 (= N)

$Q_1 = \frac{N+1}{4}$ -তম পদের মান (N = মোট পরিসংখ্যা)

$= \frac{98+1}{4}$ " " " $= \frac{4 \times 84}{10}$

$= 24.75$ " " " $= 50$ কিলো।

$Q_3 = 3\frac{(n+1)}{4}$ -তম পদের মান

$= 3\frac{(98+1)}{4}$ " " "

$= 74.25$ " " " $= 60$ কিলো

$D_4 = 4\frac{(N+1)}{10}$ " " "

$= 4\frac{(98+1)}{10}$ " " "

$= 39.6$ " " " $= 54$ কিলো

$P_{60} = \frac{60(N+1)}{10}$ তম পদের মান

$= \frac{60(98+1)}{10}$ " " "

$= 59.4$ " " " $= 59$ কিলো

(C) **গ্রুপ পরিসংখ্যা বিভাজনের ক্ষেত্রে (For grouped frequency distribution) ঃ**

উদাহরণ ঃ

ওজন (কিলো)	পরিসংখ্যা (f)	ক্রমযৌগিক পরিসংখ্যা
20-24	2	2
24-28	3	5
28-32	5	10
32-36	10	20
36-40	8	28
40-44	6	34
44-48	16	50
48-52	12	62
52-56	10	72
56-60	7	79
60-64	5	84 (= N)

$Q_1 = \frac{N}{4}$-তম পদের মান $= \frac{84}{4}$-তম পদের মান $= 21$-তম পদের মান।

Q_1 (36-40) এই শ্রেণীতে অবস্থান করছে।

অন্তঃমান সূত্রের (Interpretation formula) **সাহায্যে** [মধ্যমার সূত্রের অনুরূপ],

$Q_1 = l_1 + \frac{l_2 - l_1}{f_1}(Q - C)$ এখানে $l_1 = 36$, $l_2 = 40$, $f_1 = 8$, $Q = 21$, $C = 20$

$$= 36 + \frac{40-36}{8}(21-20)$$

$$= 36 + \frac{4}{8} \times 1 = 36 + 5 = 36.5$$ কেজি

$Q_3 = \frac{3N}{4}$-তম পদের মান $\frac{3 \times 84}{4}$-তম পদের মান $= 63$-তম পদের মান

$\therefore$ Q_3 (56-52) এই শ্রেণীতে অবস্থান করে। অন্তঃমান সূত্রের সাহায্যে,

$$Q_3 = 52 + \frac{56-52}{10}(63-62) = 52 + \frac{4}{10} = 52 + .4 = 52.4$$ কেজি

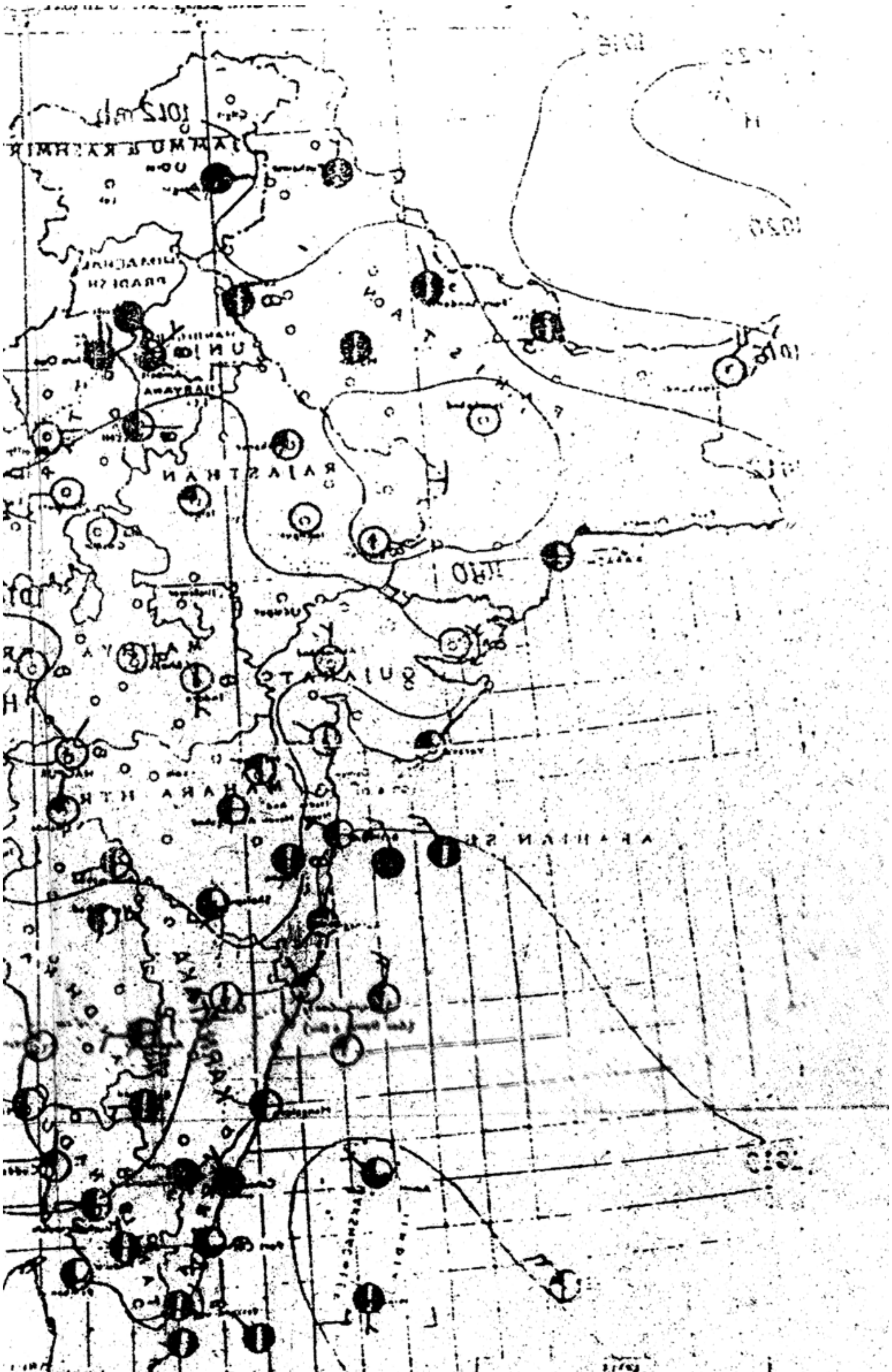
$D_4 = \frac{4N}{10}$-তম পদের মান $= \frac{4 \times 84}{10}$-তম পদের মান $= 33.6$-তম পদের মান

$\therefore$ $D_4 = $ (40-44) এই শ্রেণীতে অবস্থান করছে।

INDIAN DAILY

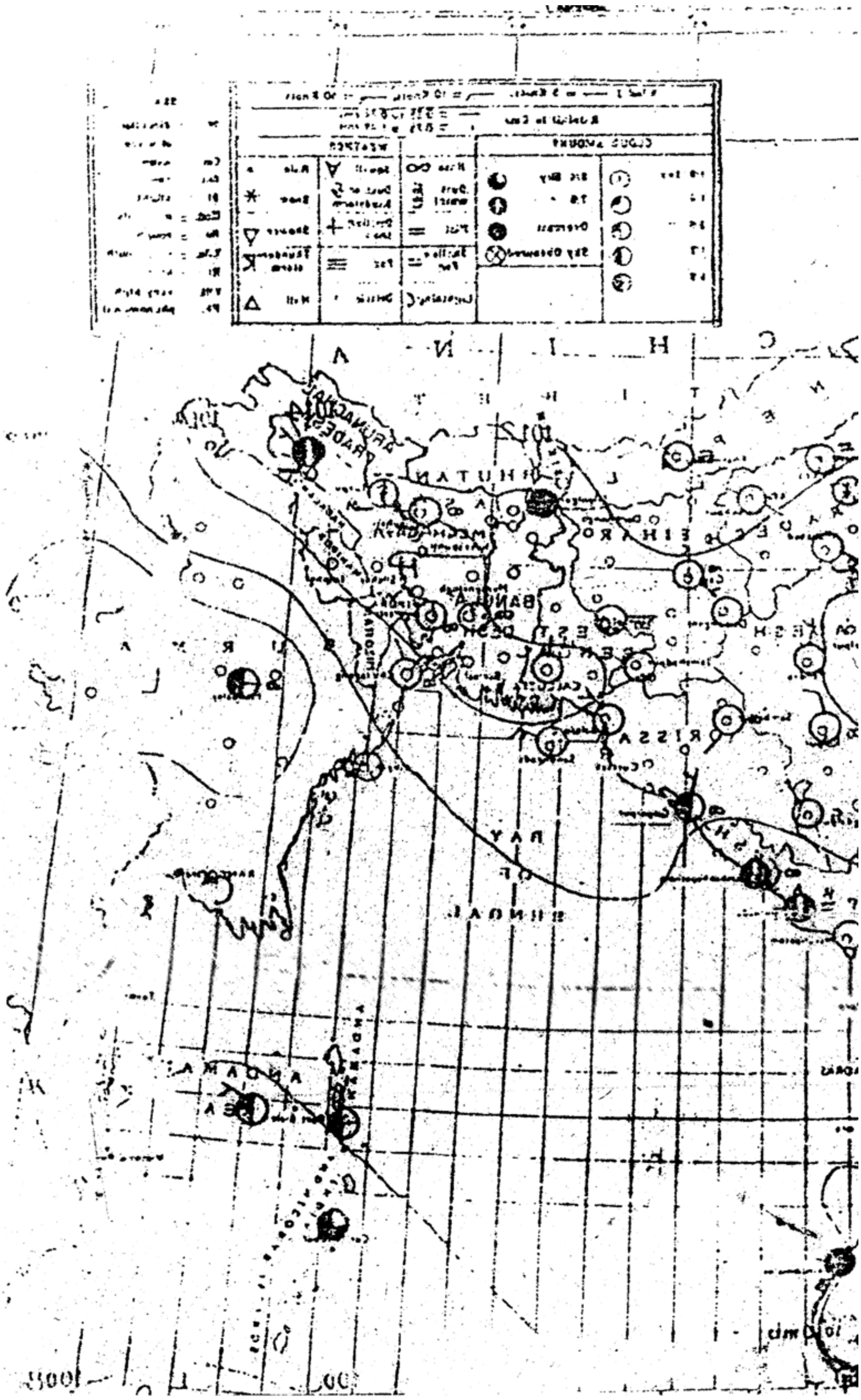
WEATHER MAP AT 0830

SUNDAY 12 JULY 1973



WEATHER REPORT

HRS I.S.T. (0300 HRS. G.M.T.)

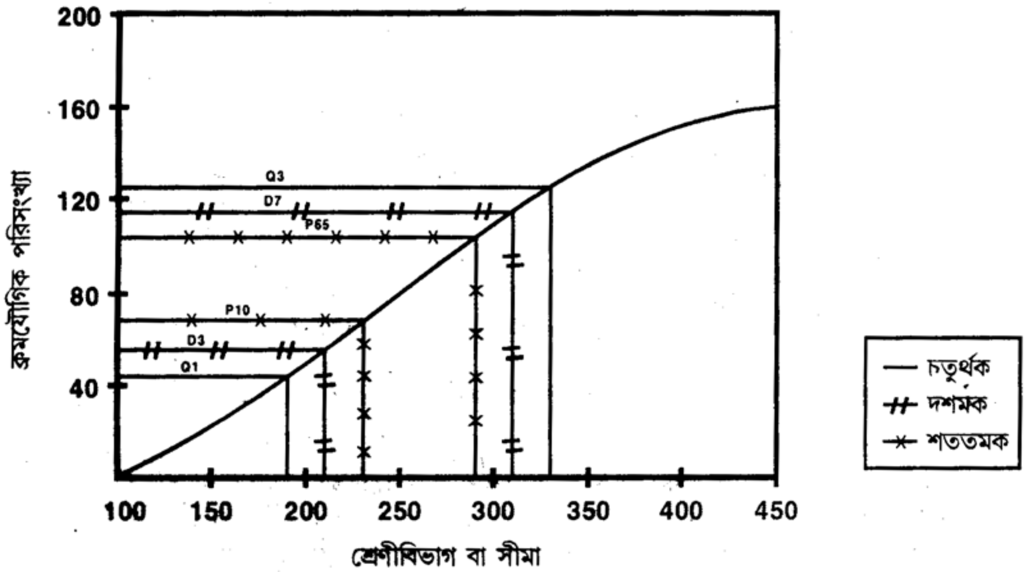


$$D_4 = 40 + \frac{44-40}{6}(33.6-28) = 40 + \frac{4}{6} \times 5.6 = 40 \times 3.7 = 43.7 \text{ কেজি}$$

এইরূপ D_7, P_{17} ইত্যাদিও গণনা করা যায়।

ক্রমযৌগিক পরিসংখ্যারেখা থেকে চতুর্থক, দশম ও শততম নির্ণয় (Determination of Quartile, Decile and Percentile from Cumulative Frequency Curve) ঃ

সরাসরি **ক্রমযৌগিক** পরিসংখ্যারেখা (নীচ থেকে) থেকেও **চতুর্থক, দশমক** ও **শততমক** নির্ণয় করা যায়। চিত্রে **প্রথম** ও তৃতীয় চতুর্থক, তৃতীয় ও **সপ্তম দশমক** এবং চল্লিশ ও পঁয়ষট্টি **শততমক** দেখানো হয়েছে। নীচের **উদাহরণ** ও চিত্র লক্ষ্য করুন।



লেখচিত্র—**ক্রমযৌগিক পরিসংখ্যা লেখচিত্র (নীচ থেকে)** চতুর্থক, দশমক ও শততমক দেখানো।

শ্রেণীবিভাগ	পরিসংখ্যা	ক্রমযৌগিক পরিসংখ্যা (নীচ থেকে)	ক্রমযৌগিক পরিসংখ্যা (ওপর থেকে)
100-150	20	20	170
150-200	25	45	150
200-250	30	75	125
250-300	40	115	95
300-350	28	143	55
350-400	18	161	27
400-500	9	170	9
মোট N = 170			

প্রথম চতুর্থক নির্ণয় ঃ

$$Q_1 = \frac{N}{4} = \frac{170}{4} = 42.5\text{-তম পরিসংখ্যা}$$

$$\alpha_1\,(150) + \frac{42.5\left(\frac{N}{4}\right) - 20(C) \times 50(i)}{25(f)}$$

$$150 + \frac{22.5}{25} \times 50 = 150 + 0.9 \times 50 = 150 + 45 = 195$$

তৃতীয় চতুর্থক নির্ণয় ঃ

$$Q_3 = 3\frac{N}{4} = 170 \times 4 \div 4 = 127.5\text{-তম পরিসংখ্যা}$$

এখানে α_1 হল 300, C হল 115

$$\therefore 300 + \frac{127.5 - 115}{28} \times 50$$

$$= 300 + \frac{12.5}{28} \times 50 = 300 + 0.446 \times 50 = 300 + 22.32 = 322.32$$

অনুরূপভাবে, তৃতীয় দশক ও সপ্তম **দশমক** হবে যথাক্রমে 210 এবং 307.14 আর 10 এবং 65 **শততমক** 20 যথাক্রমে 68 এবং 294.37।

উপরোক্ত সব ফলাফলকেই লেখচিত্রের সাহায্যে প্রকাশ করা হয়েছে।

4.6 সংখ্যাগুরু মান বা ভূয়িষ্ঠক (Mode)

সংগৃহীত রাশিতথ্যমালার মধ্যে যে মানটি সবচেয়ে বেশী ব্যবহৃত হয়, তাকে **সংখ্যাগুরুমান** বলে।

Croxton and Cowden-র ভাষায় "The mode of a distribution is the value at the point around which the items to be most heavily concentrated. It may be regarded as the most typical of a series of values" যখন কোন ব্যক্তির গড় বেতন, গড় উচ্চতার কথা বলে থাকি, তখন এতে করে আমরা সাধারণত সংখ্যাগুরুমানকে বোঝাই। যখন আমরা কোন কারখানার কর্মীদের সংখ্যাগুরু মজুরী মান 70 টাকা বলি তার দ্বারা আমরা বোঝাই যে ঐ কারখানার অধিকাংশ কর্মী 70 টাকা মজুরী পাচ্ছেন। যদি কোন পরিসংখ্যা বিভাজনের একটি সংখ্যাগুরু মান থাকে তবে তাকে **এক সংখ্যা গুরুমান সম্বলিত** (unimodal) **পরিসংখ্যা বিভাজন** বলি। আবার যদি কোন পরিসংখ্যা বিভাজনের দুটি বা তার বেশী সংখ্যক সংখ্যাগুরুমান থাকে, তবে তাকে যথাক্রমে **দুই** বা (Bimodal) **বহু সংখ্যাগুরু** (multimodal) মান সম্বলিত পরিসংখ্যা বিভাজন বলি।

4.6.1 সরল, বিচ্ছিন্ন, অবিচ্ছিন্ন, ক্রমযৌগিক পরিসংখ্যা এবং দুই বা বহুসংখ্যক গুরুমান সম্বলিত পরিসংখ্যা বিভাজন থেকে মধ্যমা নির্ণয়ের বিভিন্ন পন্থা ঃ

(A) সরল পরিসংখ্যা বিভাজন বা বিচ্ছিন্ন শ্রেণী থেকে মধ্যমা নির্ণয় ঃ

নীচের উদাহরণটি লক্ষ্য করুন ঃ

উচ্চতা (ইঞ্চি)	57	59	61	62	63	64	65	66	67	69
লোকের সংখ্যা	3	5	7	10	20	22	24	5	2	2

গ্রুপ টেবিল

উচ্চতা (ইঞ্চি)	পরিসংখ্যা					
	(1)	(2)	(3)	(4)	(5)	(6)
57	3	8		15		
59	5		12		22	
61	7	17				37
62	10		30	52		
63	20	42			65	
64	22		46			51
65	24	29		31		
66	5		7		9	
67	2	4				
69	2					

2 এবং 3 নম্বর স্তম্ভে পরিসংখ্যানগুলো দুটি সংখ্যা একসাথে করে যোগ করা হয়েছে (ছকে দেখান হয়েছে)। আবার 4, 5 এবং 6 নং স্তম্ভে তিনটি সংখ্যা একসাথে করে যোগফল বার করা হয়েছে। প্রতিটি স্তম্ভে সর্বাধিক পরিসংখ্যা মোটা অক্ষরে দেখান হয়েছে। নিচের ছকে বিভিন্ন সর্বাধিক পরিসংখ্যা মান সাজিয়ে দেখান হয়েছে।

বিশেষণ ছক (Analysis table)

স্তম্ভ (column)	সার্বিক পরিসংখ্যাসহ পদের মান .				
1				65	
2		63	64		
3			64	65	
4	62	63	64		
5		63	64	65	
6			64	65	66
মোট সংখ্যা	1	3	5	4	1

উপরের ছক থেকে আমরা পাই যে পদের মান সর্বাধিক ঘটেছে তা হল 64। সুতরাং সংখ্যাগুরু মান 64-এ নির্দেশ করা হয়েছে।

টীকা ঃ নং স্তম্ভ থেকে দেখা যায় যে 65-এর সর্বাধিক পরিসংখ্যা (24)। অতএব সংখ্যাগুরুমান 65। গ্রুপের সাহায্যে এই ভুল ধারণা শুদ্ধ করা হয়েছে। কাজেই শুধুমাত্র পরিদর্শন দেখে এই মান নির্ণয় করা ঠিক নয়।

(B) গ্রুপ পরিসংখ্যা বিভাজন বা অবিচ্ছিন্ন শ্রেণী থেকে মধ্যমা নির্ণয় ঃ

এখানে গ্রুপ ছকের মাধ্যমে সংখ্যাগুরু মান শ্রেণী (modal class) নির্ণয় করতে হবে। তারপর নিচের সূত্রের সাহায্যে আসল মান নির্ণয় করতে হবে।

(C) বিচ্ছিন্ন পরিসংখ্যা বিভাজন থেকে সংখ্যাগুরু মান নির্ণয় ঃ

গ্রুপ পরিসংখ্যা বিভাজনে সংখ্যাগুরু মান (mode) নির্ণয়ে (মধ্যমার মত) শ্রেণী বিভাগসমূহকে অবিচ্ছিন্ন (continuous) হতে হবে। শ্রেণীবিভাগসমূহ বিচ্ছিন্ন (discrete) অবস্থান থাকলে তাদের-কে প্রথমে শ্রেণী সীমাতে রূপান্তরিত করে প্রয়োজনীয় সূত্র ব্যবহারর করতে হবে। নীচের উদাহরণটি লক্ষ্য করুন।

উদাহরণ ঃ

শ্রেণীবিভাগ ঃ	50-59	60-69	70-79	80-89	90-99	100-109
পরিসংখ্যা ঃ	5	20	40	50	30	6

সমাধান ঃ

ওপরের ছকটি লক্ষ্য করুন। এখানে শ্রেণীবিভাগসমূহ বিচ্ছিন্ন অবস্থায় রয়েছে। তাই প্রথমে সেগুলোকে শ্রেণী সীমা-তে রূপান্তরিত করতে হবে।

শ্রেণী সীমা	পরিসংখ্যা
49.5-59.5	5
59.5-69.5	20
69.5-79.5	$40(f_0)$
79.5-89.5	$50(f_1)$
89.5-99.5	$30(f_2)$
99.5-109.5	6

এখানে সর্বাধিক পরিসংখ্যা 50। তাই সংখ্যাগুরু মান শ্রেণী (79.5-89.5)

$l = 79.5,\ f_0 = 40,\ f_1 = 50,\ f_2 = 30,\ i = 10\ (= 89.5\text{-}79.5)$

$\therefore$ সংখ্যাগুরুমান

$$= l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i = 79.5 + \frac{50-40}{2\times 50 - 40 - 30} \times 10$$

$$= 79.5 + \frac{10}{30} \times 10 = 79.5 + 3.33 = 82.83$$

(D) ক্রমযৌগিক পরিসংখ্যা বিভাজনের ক্ষেত্রে সংখ্যাগুরু মান নির্ণয় ঃ

ক্রমযৌগিক পরিসংখ্যা বিভাজনকে প্রথমে সরল গ্রুপ পরিসংখ্যা বিভাজনে পরিবর্তন করে নির্ণেয় মান (সূত্র ব্যবহার করে) বের করতে হবে।

উদাহরণ ঃ নিম্নলিখিত ছক থেকে সংখ্যাগুরুমান (Mode) নির্ণয় করুন।

মার্ক	ছাত্রের সংখ্যা	মার্ক	ছাত্রের সংখ্যা
10-এর বেশী	59	50-এর বেশী	18
20 ” ”	54	60 ” ”	8
30 ” ”	46	70 ” ”	0
40 ” ”	34		

লক্ষ্য করুন যে এখানে কোন শ্রেণীবিভাগ দেওয়া নেই। তাই প্রথমে আমাদের ক্রমযৌগিক পরিসংখ্যা বিভাজনকে সরল পরিসংখ্যা বিভাজনে পরিবর্তন করতে হবে।

শ্রেণী সীমা	ছাত্রের সংখ্যা
10-20	5
20-30	8
30-40	12
40-50	16
50-60	10
60-70	8

প্রমাণ সংখ্যাগুরু মানের সূত্র প্রয়োগ করে আমরা পাই ঃ

$$40(l_1)+\frac{16(f_1)-12(f_0)}{32(2f_1)-12(f_0)-10(f_2)}\times 10(i)$$

$$=40+\frac{4}{10}\times 40=40+4=44 \text{ মার্ক}$$

উদাহরণ ঃ

নিম্নলিখিত ছকের সংখ্যাগুরু মান (Mode) নির্ণয় করুন।

গ্রুপ বিভাগে পরিবর্তন করিয়া পাই

নম্বর	ছাত্রসংখ্যা	নম্বর	ছাত্রসংখ্যা (পরিসংখ্যা)
10-এর নীচে	3	0-10	3
20 " "	8	10-20	$5(f_0)$
30 " "	17	20-30	$9(f_1)$
40 " "	20	30-40	$3(f_2)$
50 " "	22	40-50	2

এক্ষেত্রে সর্বাধিক পরিসংখ্যা 9, সুতরাং সংখ্যাগুরু মানশ্রেণী (20-30)

$l = 20,\ f_0 = 5,\ f_1 = 9,\ f_2 = 3,\ i = 10$

$$\text{সংখ্যাগুরুমান} = 20+\frac{9-5}{2\times 9-5-3}\times 10=20+\frac{4}{10}\times 10=20+4=24 \text{ নম্বর।}$$

প্রশ্ন ঃ নিম্নোক্ত 122 জনের ওজনের পরিসংখ্যা থেকে সংখ্যাগুরু ওজন নির্ণয় করুন।

সমাধান ঃ চোখে দেখে সংখ্যাগুরু ওজন নির্ণয় করা শক্ত ব্যাপার।

ওজন (পাউণ্ড)	মানুষের সংখ্যা	ওজন (পাউণ্ড)	মানুষের সংখ্যা
100-110	4	140-150	33
110-120	6	150-160	17
120-130	20	160-170	8
130-140	32	170-180	2

সমাধান ঃ

চোখে দেখে বলা কঠিন যে কোনটি সংখ্যাগুরুমান শ্রেণী। তাই আমরা প্রথমে একটি গ্রুপ ছক প্রস্তুত করেছি ও পরে বিশ্লেষণ ছক।

গ্রুপ ছক

মানুষের সংখ্যা

ওজন (পাউণ্ড)	I	II	III	IV	V	VI
100-110	9					
		10				
110-120	6					
120-130	20					
		52				
130-140	32					
140-150	33					
		50				
150-160	17					
160-170	8					
		10				
170-180	2					

যে শ্রেণীতে সংখ্যাগুরুমান অবস্থান করছে বলে আশা করা যায়।

বিশ্লেষণ ছক

ক্রমিক নং	120-130	130-140	140-150
1			1
2	1	1	
3		1	1
4		1	1
5	1	1	1
6	1	1	1
	Total 3	5	5

এইটি একটি দ্বি-মাত্রিক সংখ্যাগুরুমান সিরিজ। সুতরাং, নিম্নোক্ত সূত্র প্রয়োগ করে সংখ্যাগুরুমান নির্ণয় করতে হবে সংখ্যাগুরুমান = 3, মধ্যমা = 2 গড়

		মানুষের সংখ্যা		(m − 135)/10	
ওজন (পাউণ্ডে)	m	f	e.f.	d	fd
100-110	105	4	4	− 3	− 12
110-120	115	6	10	− 2	− 12
120-130	125	20	30	− 1	− 20
130-140	135	32	62	0	0
140-150	145	33	96	+ 1	+ 33
150-160	155	17	112	+ 3	+ 34
160-170	165	8	120	+ 3	+ 24
170-180	175	2	122	+ 4	+ 8
	N = 122				Σfd = 55

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

A = 135, Σfd = 55, N = 122, i = 10

$$\overline{X} = 135 + \frac{55}{122} \times 10 = 135 + 4.51 = 139.51$$

মধ্যমা = $\frac{N}{2}$-তম আয়তন, 25 কিনা $\frac{122}{2} = 61$-তম মান।

সুতরাং, 130-140 এই শ্রেণীতে মধ্যমা অবস্থান করছে।

$$\text{মধ্যমা} = L + \frac{\frac{N}{2} - c.f.}{f} \times i$$

L = 130, N/2 = 61, e.f. = 30, f = 32, i = 10

$$\text{মধ্যমা} = 130 + \frac{(61-30)}{32} \times 10 = 130 + \frac{310}{32} = 139.69$$

সংখ্যাগুরুমান = 3 Median − 2 Mean

সংখ্যাগুরুমান = (3 × 139.69) − (2 × 139.51) = 419.07 − 279.02 = 140.05।

সুতরাং, সংখ্যাগুরুমান ওজন হল 140.05 পাউণ্ড।

বিশ্লেষণ ছক প্রস্তুত করার সময় সমস্ত শ্রেণীকে লেখার প্রয়োজন নেই। শুধমাত্র সেইসব শ্রেণীকে গণ্য করতে হবে। যেখানে সংখ্যাগুরুমান অবস্থান করছে বলে আশা করা যায়। এইজন্য আমরা 100-110 বা 150-160 শ্রেণীগুলোকে ধরিনি।

লেখচিত্রের সাহায্যে সংখ্যাগুরু মান নির্ণয় ঃ

পরিসংখ্যা বিভাজনে সংখ্যাগুরু মানকে লেখচিত্রের সাহায্যে দেখানো যেতে পারে। এক্ষেত্রে ধাপগুলো হল নিম্নরূপ ঃ

(a) প্রদত্ত রাশিতথ্য থেকে একটি আয়তলেখচিত্র (Histogram) আঁকুন।

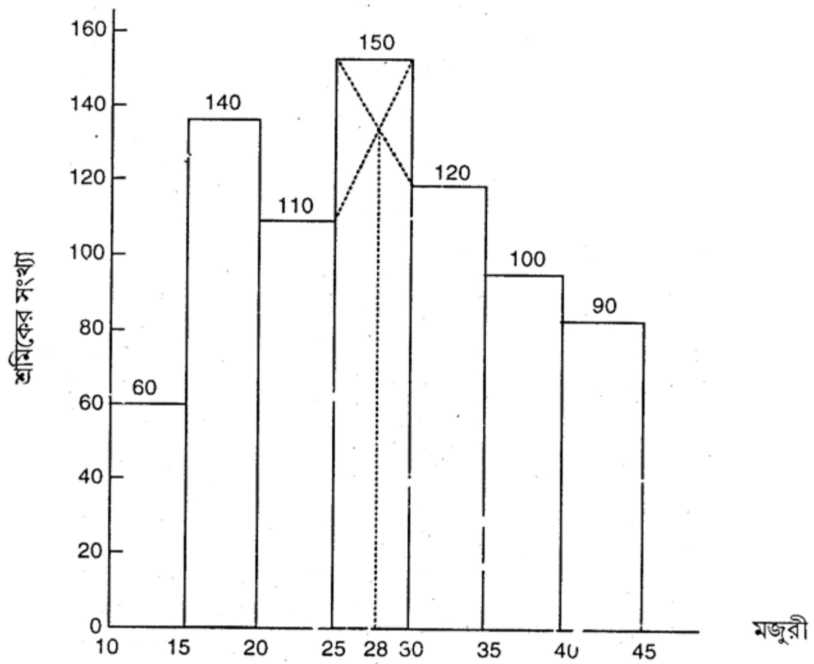
(b) সংখ্যাগুরু মানের স্তম্ভের ভেতরে আড়াআড়িভাবে দুটি রেখা টানা হল। এই দুটি রেখা যেখানে মিলল, সেখান থেকে x-অক্ষরেখার দিকে একটি উলম্বরেখা নীচের দিকে টানা হল। এটি x-অক্ষরেখায় যেখানে মিলল, সেই ছেদবিন্দুর মান হল সংখ্যাগুরু মান।

উদাহরণ ঃ

নীচের শ্রমিকের সংখ্যা ও তাদের মজুরী দেওয়া হল। শ্রমিকের মজুরীর সংখ্যাগুরুমান নির্ণয় করুন। মানটি গণনার মাধ্যমেও ঠিক আছে কিনা দেখান।

মজুরী (টাকায়) ঃ	10-15	15-20	20-25	25-30	30-35	35-40	40-45
শ্রমিকের সংখ্যা ঃ	60	140	110	150	120	100	90

উপরোক্ত পরিসংখ্যান থেকে আয়তলেখচিত্র আঁকা হল।



সংখ্যা = 28
গুরুমান
মজুরী (টাকায়)

4.6.2 সংখ্যাগুরু মানের সুবিধে ও অসুবিধে ঃ

সুবিধে (Advantage) ঃ

(1) চোখে দেখে কখনও কখনও সংখ্যাগুরুমান নির্দেশ করা যায়।

(2) খুব বেশী বা কম সংখ্যা এই গড়ে প্রভাব বিস্তার করে না।

(3) এটি একটি সহজ ও সরল গড়। অবিচ্ছিন্ন রাশি ছাড়া অন্যান্য ক্ষেত্রে এই গড় শ্রেণীর কোন আসল মানকেই বোঝায়।

(4) সাধারণ গড় অর্থে এর ব্যবহার ব্যাপক। অধিকাংশ বেশীরভাগ ইত্যাদি অর্থ সংখ্যাগুরুমানকে সূচিত করে।

(5) বেশীরভাগ ক্ষেত্রে সংখ্যাগুরুমান চলরাশির প্রদত্ত কোন না কোন মাপের সাথে মিশে যায়।

(6) কোন চলরাশির সংখ্যাগুরুমান বের করার জন্য চলরাশির সব মান জানার দরকার নেই। মুক্তপ্রাপ্ত (open end) বিশিষ্ট শ্রেণীবিভাগযুক্ত পরিসংখ্যা বিভাজনের ক্ষেত্রেও সংখ্যাগুরুমান নির্ণয় করা যায়।

(7) লেখ-র সাহায্যে এই গড় নির্ণয় করা সম্ভব।

অসুবিধে (Disadvantage) ঃ

(1) রাশিতথ্যমালার সংখ্যা কম হলে এই গড় নির্ণয়ে অসুবিধে হয়। কিন্তু এক্ষেত্রে যৌগিক গড় ও মধ্যমা গণনা করা যায়।

(2) শ্রেণীর প্রতিটি রাশির উপর সংখ্যাগুরু মানের মাপ নির্ভর করে না।

(3) বীজগাণিতিক প্রয়োগে এই গড় উপযুক্ত নয়।

(4) এর নমুনা বিচ্যুতি (Sampling fluctuation)-ও বেশী।

4.6.3 কেন্দ্রীয় প্রবণতার বিভিন্ন পরিমাপকগুলোর মধ্যে তুলনা (Comparison of different measures of central tendency) ঃ

একটি উৎকৃষ্ট মানের কেন্দ্রীয় প্রবণতা পরিমাপকের কতগুলো ধর্ম বা বৈশিষ্ট্য থাকা প্রয়োজন। এই বৈশিষ্ট্যগুলো হল ঃ

(a) এর সংজ্ঞা দ্ব্যর্থহীন হওয়া দরকার।

(b) এটি যেন সহজেই গণনা করা যায়।

(c) এটি চলরাশির সব মানের উপর নির্ভরশীল হওয়া উচিত।

(d) এটি যেন সহজ ব্যাখ্যাযোগ্য হয়।

(e) বীজগণিতের সূত্রগুলো প্রয়োগের পক্ষে এটি যেন উপযুক্ত হয়।

(f) চলরাশির চরম মান দ্বারা এটি যেন প্রভাবিত না হয়।

(g) এর নমুনা বিচ্যুতি যেন যথাসম্ভব কম হয়।

আমরা এখন কেন্দ্রীয় প্রবণতার বিভিন্ন পরিমাপকগুলোকে উপরোক্ত গুণাবলীর আলোকে বিচার করতে পারি। বাস্তবক্ষেত্রে কোন পরিমাপকেই উপরোক্ত সব গুণসম্পন্ন হয় না। এখন কেন্দ্রীয় প্রবণতা পরিমাপকের মধ্যে যে পরিমাপকটি সবচেয়ে বেশী গুণসম্পন্ন তাকেই আমরা কেন্দ্রীয় প্রবণতা পরিমাপকের মধ্যে শ্রেষ্ঠ পরিমাপক হিসেবে চিহ্নিত করতে পারি।

আমরা এখন উপরোক্ত গুণাবলীর ভিত্তিতে গাণিতিক গড়, মধ্যমা এবং সংখ্যাগুরুমান-কে বিচার করতে পারি।

গাণিতিক গড় (A.M.) ঃ

(i) এর সংজ্ঞা দ্ব্যর্থহীন।

(ii) ইহা সহজে গণনা করা যায়।

(iii) ইহা চলরাশির সব মাপের ওপর নির্ভর করে।

(iv) বীজগণিতের সূত্রগুলি প্রয়োগের পক্ষে উপযুক্ত এবং প্রমাণ থেকে অনেক গুরুত্বপূর্ণ সূত্র তৈরী করা যায়।

(v) অন্যান্য পরিমাপকের তুলনায় এর নমুনা বিচ্যুতি অনেক কম।

(vi) এটি সহজে ব্যাখ্যা করা যায়।

গাণিতিক গড়ের ত্রুটি হল ইহা চলরাশির চরমমান দ্বারা বিশেষভাবে প্রভাবিত হয়।

মধ্যমা (Median) ঃ

(i) কোন চলরাশির যুগ্মসংখ্যক মান না থাকলে এর সংজ্ঞা দ্ব্যর্থহীন।

(ii) মধ্যমা গণনা সহজ।

(iii) মধ্যমা সহজে ব্যাখ্যা করা যায়।

(iv) চলরাশির চরম মানের দ্বারা এটি প্রভাবিত হয় না।

তবে মধ্যমার **অসুবিধে** হল ঃ

(i) ইহা চলরাশির সব মানের উপর নির্ভর করে না।

(ii) বীজগণিতের সূত্রগুলো প্রয়োগ করার পক্ষে মধ্যমা উপযুক্ত নয়।

(iii) মধ্যমার নমুনা বিচ্যুতিও বেশী।

সংখ্যাগুরু মান (Mode) ঃ

(i) এটি সহজে ব্যাখ্যা করা যায়।

(ii) ইহা চলরাশির চরম মান দ্বারা প্রভাবিত হয় না।

সংখ্যাগুরু মানের অসুবিধে হল ঃ

(i) এর সংজ্ঞা দ্ব্যর্থহীন নয়। কেননা অনেকক্ষেত্রে একাধিক সংখ্যাগুরু মান পাওয়া যায়।

(দ্বি বা বহুসংখ্যাগুরু মান)

(ii) এর গণনা সহজ নয়।

(iii) ইহা চলরাশির সব মানের ওপর নির্ভর করে না।

(iv) বীজগণিতের সূত্রগুলো প্রয়োগের পক্ষে এটি উপযুক্ত নয়।

(v) এর নমুনা বিচ্যুতিও বেশী।

উপরোক্ত আলোচনার পরিপ্রেক্ষিতে আমরা বলতে পারি যে মধ্যমা এবং সংখ্যাগুরু মানের তুলনায় গাণিতিক গড় কেন্দ্রীয় প্রবণতার পরিমাপকের গুণাবলীর বেশীরভাগই মেনে চলে। এই বিচারে গাণিতিক গড়কে আমরা কেন্দ্রীয় প্রবণতার পরিমাপকগুলোর মধ্যে সবচেয়ে উন্নত মানের বলতে পারি।

যদিও গাণিতিক গড়কে সবচেয়ে উৎকৃষ্ট কেন্দ্রীয় প্রবণতার পরিমাপক হিসেবে ধরা হয়, তবুও দু'টি বিশেষ পরিস্থিতিতে গাণিতিক গড়ের তুলনায় মধ্যমা এবং সংখ্যাগুরু মান বেশী উপযুক্ত।

(a) যখন একটি দলে দু'-একটি চরম মান থাকে।

(b) শ্রেণীবিভাগ যুক্ত পরিসংখ্যা বিভাজনের ক্ষেত্রে যখন মুক্তপ্রান্ত বিশিষ্ট শ্রেণী থাকে। উভয় ক্ষেত্রে গাণিতিক গড় ব্যবহার করা যায় না।

যুক্তি হল ঃ প্রথম ক্ষেত্রে গাণিতিক গড় চরম মান দ্বারা বিশেষভাবে প্রভাবিত হয়।

দ্বিতীয় ক্ষেত্রে মুক্তপ্রান্ত বিশিষ্ট শ্রেণীর মধ্যমান নির্ণয় করা যায় না। কিন্তু মধ্যমা বা সংখ্যাগুরু মান নির্ণয়ের ক্ষেত্রে তা কোন সমস্যার সৃষ্টি করে না।

4.7 প্রশ্নোত্তর পর্ব

প্রশ্ন ঃ কোন স্থানের দশটি পরিবারের প্রাত্যহিক আয় দেওয়া আছে। এ থেকে গুণোত্তর গড় নির্ণয় করুন।

85 70 15 75 500 8 45 250 40 36

উত্তর ঃ

গুণোত্তর গড় নির্ণয়			
x	লগ	x	লগ
85	1.9294	8	0.9031
70	1.8451	45	1.6532
15	1.1761	250	2.3979
75	1.8751	40	1.6021
500	2.6990	36	1.5563
			Σ লগ x = 17.6373

গুণোত্তর গড় = এ্যান্টিলগ $\left(\frac{\sum \text{লগ } x}{N}\right)$ = এ্যান্টিলগ $\frac{17.6373}{10}$ = এ্যান্টিলগ 1.7337।

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যান থেকে গুণোত্তর গড় নির্ণয় করুন।

125 1462 38 7 022 0.08 12.75 0.5

উত্তর ঃ

গুণোত্তর গড় নির্ণয়

x	লগ x
125	2.0969
1462	3.1650
38	1.5798
7	0.8451
0.22	.3424
0.08	.9031
12.75	1.1055
0.5	.6990
	Σ লগ x = 6.7360

গুণোত্তর গড় = এ্যান্টিলগ $\left(\frac{\sum \text{লগ } x}{N}\right)$ = এ্যান্টিলগ $\frac{17.6373}{10}$ = এ্যান্টিলগ 1.7637 = 58.03

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যান থেকে গুণোত্তর গড় নির্ণয় করুন।

125 1462 38 7 0.22 0.08 12.75 05.

উত্তর ঃ

গুণোত্তর গড় নির্ণয়

x	লগ x
125	2.0969
1462	3.1650
38	1.5798
7	0.8451
0.22	0.3424
0.08	0.9031
12.75	1.1055
0.5	0.6990
	Σ লগ x = 6.7360

গুণোত্তর গড় = এ্যান্টিলগ $\left(\frac{\sum \text{লগ } x}{N}\right)$ = এ্যান্টিলগ $\left(\frac{6.7368}{8}\right)$ = এ্যান্টিলগ (0.8421 = 6.952)

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যান থেকে গুণোত্তর গড় নির্ণয় করুন।

নম্বর	পরিসংখ্যা	নম্বর	পরিসংখ্যা
4-8	6	24-28	12
8-12	10	28-32	10
12-16	18	32-36	6
16-20	30	36-40	2
20-24	15		

উত্তর ঃ

গুণোত্তর গড় নির্ণয়

নম্বর	মধ্যবর্তী নম্বর	পরিসংখ্যা (f)	লগ m	f × লগ m
4-8	6	6	0.7782	4.6692
8-12	10	10	1.0000	10.0000
12-16	14	18	1.1461	20.6298
16-20	18	30	1.2553	37.6590
20-24	22	15	1.3424	20.1360
24-28	26	12	1.4150	16.9800
28-32	30	10	1.4771	14.7710
32-36	34	6	1.5315	9.1890
36-40	38	2	1.5798	3.1596
		মোট পরিসংখ্যা = 109	$\Sigma f \times$ লগ m = 137.1936	

$$\text{গুণোত্তর গড়} = \text{এ্যান্টিলগ}\left(\frac{\sum f\,\text{লগ}\,x}{N}\right) = \text{এ্যান্টিলগ}\left(\frac{137.1936}{109}\right) = \text{এ্যান্টিলগ} = 1.2587 = 18.14.$$

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যান থেকে জনসংখ্যার গড় বৃদ্ধি নির্ণয় করুন। প্রথম দশকে এই বৃদ্ধি হয়েছে 20 শতাংশ, দ্বিতীয় দশকে 30 শতাংশ এবং তৃতীয় দশকে 40 শতাংশ।

উত্তর ঃ

দশক	% বৃদ্ধি	পূর্ববর্তী দশকের জনসংখ্যাকে 100 ধরে বর্তমান দশকের শেষে x জনসংখ্যা	লগ
প্রথম	20	120	2.0792
দ্বিতীয়	30	130	2.1139
তৃতীয়	40	140	2.1461
			Σ লগ x = 6.3292

গুণোত্তর গড় = এ্যান্টিলগ $\left(\frac{\sum \text{লগ } x}{N}\right)$ = এ্যান্টিলগ $\left(\frac{6.3392}{3}\right)$ = এ্যান্টিলগ = 129.7

অতএব জনসংখ্যার গড় বৃদ্ধির হার হল (129.7 − 100) = 29.7 শতাংশ প্রতি দশকে।

প্রশ্ন ঃ নীচের পরিসংখ্যান থেকে বিবর্ত যৌগিক গড় নির্ণয় করুন।

2574 475 75 5 0.8 0.08 0.005 0.0009

উত্তর ঃ

বিবর্ত যৌগিক গড় গণনা

x	(1/x)	x	(1/x)
2574	0.0004	0.8	1.2500
475	0.0021	0.08	12.5000
75	0.0133	0.005	200.0000
5	0.2000	0.0009	1111.1111
			Σ(1/x) = 1325.0769

বিবর্ত যৌগিক গড় $= \frac{N}{\sum(1/x)} = \frac{8}{1325.0769} = 0.006.$

প্রশ্ন ঃ নিম্নোক্ত অবিচ্ছিন্ন পরিসংখ্যান থেকে বিবর্ত যৌগিক গড় নির্ণয় করুন।

শ্রেণী-সীমা ঃ	10-20	20-30	30-40	40-50	50-60
পরিসংখ্যা ঃ	4	6	10	7	3

উত্তর ঃ

বিবর্ত যৌগিক গড় নির্ণয়

শ্রেণী-সীমা	মধ্যবিন্দু (m)	পরিসংখ্যা (f)	f/m
10-20	15	4	0.267
20-30	25	6	0.240
30-40	35	10	0.286
40-50	45	7	0.156
50-60	55	3	0.055
		N = 30	Σ(f/m) = 1.004

বিবর্ত যৌগিক গড় $(H.M.) = \frac{N}{\sum(1/x)} = \frac{30}{1.004} = 29.88.$

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যা বণ্টন থেকে গাণিতিক গড়, মধ্যমা এবং সংখ্যাগুরু মান নির্ণয় করুন।

শ্রেণী-সীমা	পরিসংখ্যা	শ্রেণী-সীমা	পরিসংখ্যা
10-13	8	25-28	54
13-16	15	28-21	36
16-19	27	31-34	18
19-22	51	34-37	9
22-25	75	37-40	7

উত্তর ঃ

শ্রেণী-সীমা	মধ্যবিন্দু	f	(m – 23.5/3d	fd	c.f.
10-13	11.5	8	– 4	– 32	8
13-16	14.5	15	– 3	– 45	23
16-19	17.5	27	– 2	– 54	50
19-22	20.5	51	– 1	– 51	101
22-25	23.5	75	0	0	176
25-28	26.5	54	+ 1	+ 54	230
28-31	29.5	36	+ 2	+ 72	266
31-34	32.5	18	+ 3	+ 54	284
34-37	35.5	9	+ 4	+ 36	293
37-40	38.5	7	+ 5	+ 35	300
		N = 300		Σfd = 69	

$$\overline{X} = A + \frac{\sum fd}{N} \times i = 23.5 + \frac{69}{300} \times 3 = 24.19$$

মধ্যমা = মধ্যমার আয়তন N/2 বা 150-তম মান

মধ্যমা 22-25 এই শ্রেণীর মধ্যে রয়েছে।

মধ্যমা $= L + \frac{N/22 - c.f.}{f} \times i = 22 + \frac{150-101}{75} \times 3 = 22 + 1.96 = 23.96$

সংখ্যাগুরু মান = এক নজরে আমরা দেখেছি যে 22-25 এই শ্রেণীবিভাগের মধ্যে রয়েছে।

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যা থেকে কোন শহরের 50-টি স্কুলের সমস্ত ছাত্রের গড় মার্কস নির্ণয় করুন।

প্রাপ্ত মার্কস	স্কুলের সংখ্যা	স্কুলে গড় সংখ্যক ছাত্র
35-র বেশী	7	200
30-35	10	250
25-30	15	300
20-25	9	200
15-20	5	150
15-র কম	4	100

উত্তর ঃ প্রথমে উর্দ্ধসারিতে প্রদত্ত পরিসংখ্যাকে সাজাতে হবে এবং এরপর গড় নির্ণয় করতে হবে।

গড় মার্কস নির্ণয়

মার্কস	স্কুলের সংখ্যা	গড় ছাত্রের সংখ্যা	মোট ছাত্রের সংখ্যা (f)	M.P.M.	(M − 27.5)/5 d	fd
(1)	(2)	(3)	(f)	m	d	fd
10-15	4	100	400	12.5	− 3	− 1,200
15-20	5	150	750	17.5	− 2	− 1,500
20-25	9	200	1,800	22.5	− 1	− 1,800
25-30	15	300	4,500	27.5	0	0
30-35	10	250	2,500	32.5	+ 1	+ 2,500
35-40	7	200	1,400	37.5	+ 2	+ 2,800
			N = 11,350			Σfd = + 800

$$\overline{X} = A + \frac{\sum fd}{N} \times i = 27.5 + \frac{800}{11,350} \times 5 = 27.5 + 0.35 = 27.85$$

প্রশ্ন ঃ নিম্নোক্ত পরিসংখ্যান থেকে শ্রেণীবিন্যাস পদ্ধতিকে মধ্যমা ও সংখ্যাগুরু মান নির্ণয় করুন।

মধ্যবর্তী আয়তন ঃ	15	25	35	45	55	65	75	85
পরিসংখ্যা ঃ	5	9	13	21	20	15	8	3

উত্তর ঃ যেহেতু আমাদের মধ্যবর্তী মান দেওয়া আছে, তাই প্রথমে আমরা শ্রেণীবিন্যাসের নিম্ন ও উচ্চসীমা নির্ণয় করব। শ্রেণীর ব্যবধান হল 10। তাই প্রথম শ্রেণী-সীমা হবে 10-20 (যেহেতু মধ্যবিন্দু হল 15)।

মধ্যমা ও সংখ্যাগুরু মান নির্ণয়

শ্রেণী-সীমা	মধ্যবিন্দু	f	(m – 55)/10d	fd	c.f.
10-20	15	5	– 4	– 20	5
20-30	25	9	– 3	– 27	14
30-40	35	13	– 2	– 26	27
40-50	45	21	– 1	– 21	48
50-60	55	20	0	0	68
60-70	65	15	+ 1	+ 15	83
70-80	75	8	+ 2	+ 16	91
80-90	85	3	+ 3	+ 9	94
		N = 94		Σfd = – 54	

মধ্যমা নির্ণয় ঃ মধ্যমা = মোট পরিসংখ্যা N ÷ 2 অর্থাৎ $\frac{94}{2}$ । 47-তম মান। মধ্যমা 40-50 শ্রেণী-সীমাতে অবস্থান করছে।

$$\text{মধ্যমা} = L + \frac{N/2 - c.f.}{f} \times i = 40 + \frac{47-27}{21} \times 10$$

$$= 40 + \frac{47-27}{21} \times 10$$

$$= 40 + 9.524 = 49.524$$

শ্রেণী-সীমা	I	II	III	IV	V	VI
10-20	5					
		14		27		
20-30	9		22		43	
30-40	13	34				
40-50	21		41			54
50-60	20			56		
		35			43	
60-70	15		23			
70-80	8					26
		11				
80-90	3					

যে শ্রেণীতে সংখ্যাগুরু মান অবস্থান করছে বলে আশা করা যায়

ক্রমিক সংখ্যা	40-50	50-60	60-70
I	1		
II		1	1
III		1	1
IV	1	1	1
V	1	1	
V	1	1	1
	5	5	3

এইটি হল দ্বিমাত্রিক (bimodal) সংখ্যাগুরু মান শ্রেণী এবং আমরা পরোক্ষ পদ্ধতিতে সংখ্যাগুরু মান নির্ণয় করব।

সংখ্যাগুরু মান = 3 মধ্যমা (Median) – 2 গড় (Mean)

গড় নির্ণয় $\overline{X} = A + \frac{\sum fd}{N} \times i = 55 - \frac{54}{94} \times 10 = 55 - 5.75 = 49.255$

সংখ্যাগুরু মান = (49.524) – 2(49.255) = 148.572 – 98.51 - 50.062

প্রশ্ন ° নিম্নলিখিত পরিসংখ্যান থেকে সংখ্যাগুরু মান নির্ণয় করুন।

নম্বর	ছাত্রসংখ্যা	নম্বর	ছাত্রসংখ্যা
0-র বেশী	80	60-র বেশী	28
10-র "	77	70-র "	16
20-র "	72	80-র "	10
30-র "	65	90-র "	8
40-র "	55	100-র "	0
50-র "	43		

উত্তর ঃ যেহেতু এটি যৌগিক পরিসংখ্যা বিভাজনকে নির্দেশ করছে, তাই প্রথমে আমরা এটিকে সরল পরিসংখ্যা বিভাজনে পরিবর্তন করছি।

নম্বর	ছাত্রসংখ্যা	নম্বর	ছাত্রসংখ্যা
0-10	3	50-60	15
10-20	5	60-70	12
20-30	7	70-80	6
30-40	10	80-90	2
40-50	12	90-100	8

এক নজরে আমরা লক্ষ্য করছি যে সংখ্যাগুরুমান শ্রেণী হল 50-60

$$\text{সংখ্যাগুরুমান} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

এখানে $L = 50$; $\Delta_1 = (15 - 12) = 3$; $\Delta_2 = (15 - 12) = 3$; $i = 10$

$$\therefore \text{ সংখ্যাগুরু মান হল } = 50 + \frac{3}{3+3} \times 10 = 50 + 5 = 55 \text{।}$$

প্রশ্ন ঃ নীচে একটি পরিসংখ্যা দেওয়া আছে। এ থেকে সংখ্যাগুরু মান নির্ণয় করুন।

ওজন (কেজি)	ছাত্রসংখ্যা	ওজন (কেজি)	ছাত্রসংখ্যা
93-97	2	113-117	14
98-102	5	118-122	6
103-107	12	123-127	3
108-112	7	128-132	1

উত্তর ঃ লক্ষ্য করলে দেখতে পাচ্ছি যে সংখ্যাগুরু মানের শ্রেণী-সীমা হল 108-112। কিন্তু এর প্রকৃত সীমা হল 107.5-112.5।

$$\text{সংখ্যাগুরু মান} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

এখন $L' = 107.5$, $\Delta_1 = f_1 - f_0 = (17 - 12) = 5$, $\Delta_2 = f_1 - f_2$
$= (17 - 14) = 3$, $i = 5$

$$\text{সংখ্যাগুরু মান হল } = 107.5 + \frac{5}{5+3} \times 5 = 107.5 + 3.125 = 110.625 \text{ কেজি।}$$

4.7 গ্রন্থ নির্দেশিকা

চট্টোপাধ্যায়, কৃষ্ণদাস, 2000 ঃ **সামাজিক গবেষণা ঃ পদ্ধতি ও প্রক্রিয়া,** আরামবাগ বুক হাউস, কলকাতা।

দে, সৌরেন্দ্রনাথ, 2002 ঃ **ব্যবসায় গণিত ও পরিসংখ্যান,** ছায়া প্রকাশনী, কলকাতা।

নাগ, এন. কে., 1988 ঃ **ব্যবসায় গণিত ও পরিসংখ্যান,** ভারতী বুক ষ্টল, কলকাতা।

বাগচী, কনক কান্তি ও অনিল ভুইমালী, 2000 ঃ অর্থনৈতিক উন্নয়ন ও রাশিবিজ্ঞান, বি. সরকার এন্ড কোং, কলকাতা।

সরখেল, জয়দেব ও সন্তোষ কুমার দত্ত, 2002 ঃ রাশিবিজ্ঞানের ভূমিকা, বুক সিন্ডিকেট, কলকাতা।

Bhat, L.S. and Aslam Mohmood, 1977 : **Field work and Laboratory Techniques in Geography,** Oxford & IBH Pub. Co., New Delhi.

Gupta, S. P., 2000 : **Statistical Methods,** S. Chand, New Delhi.

Sarkar, Ashis, 1997 : **Practical Geography : A Systematic Approach,** Orient Longman, Kolkata.

একক 5 ☐ বিস্তৃতির বিভিন্ন পরিমাপ চতুর্থক পার্থক্য, গড় পার্থক্য, সমক পার্থক্য

গঠন

5.1 প্রস্তাবনা

পূর্ববর্তী খণ্ডে আপনারা পড়েছেন যে পরিসংখ্যানীয় গড় (বা কেবল মাত্র গড়) রাশিতথ্যের ধরন সম্বন্ধে একটি ভাল ধারণা দিতে পারে। কিন্তু গড় সবরকম বৈশিষ্ট্য প্রকাশ করতে পারে না। যেমন গড়ের সাহায্যে চলকের মানগুলির গড়ের চতুর্দিকে কিভাবে বিস্তৃত (scattered or dispersed) তা বোঝা যায় না। সম সংখ্যক মানের দ্বারা গঠিত দুটি শ্রেণীর একই গড় থাকতে পারে। কিন্তু একটির মানগুলি গড় থেকে বহুদূর বিস্তৃত এবং অপরটির মানগুলি কাছাকাছি হতে পারে। সুতরাং মানগুলির গড় থেকে তাদের বিস্তার (scattered or dispersed) জানবার প্রয়োজন আছে। এই এককে চলকের মানের বিস্তৃতির বিভিন্ন পরিমাপ আলোচনা করা হচ্ছে।

5.2 উদ্দেশ্য

এই এককটি পাঠ করে আমরা বিস্তৃতির বিভিন্ন পরিমাপ সম্বন্ধে ধারণা ও তাদের সুবিধা ও অসুবিধা বিষয়ে জানতে পারি।

5.3 বিস্তৃতি (Dispersion)

চলকের বিভিন্ন মানের তাদের গড় থেকে ভেদ বা পার্থক্যকে বিস্তৃতি বলা হয়।

বিস্তৃতির পরিমাপ

পরম পরিমাপ	আপেক্ষিক পরিমাপ
(ক) প্রসার	(ক) ভেদাঙ্ক
(খ) চতুর্থক পার্থক্য	(খ) চতুর্থক পার্বক্যাঙ্ক
(গ) গড় পার্থক্য	(গ) গড় পার্বক্যাঙ্ক
(ঘ) সমক পার্থক্য	

এখানে আমরা 'পরম পরিমাপ' নিয়ে আলোচনা করব।

প্রসার : বিস্তৃতির সর্বাপেক্ষা সরল পরম পরিমাপ হল প্রসার। এটি চলকের বৃহত্তম এবং ক্ষুদ্রতম মান দুটির অন্তরকলের সমান। অর্থাৎ প্রসার হল কোন একটি বিস্তারের (distribution) বৃহত্তম ও ক্ষুদ্রতম মানের পার্থক্য। যেমন কোন একটি স্থানের (suppose A) সর্বোচ্চ তাপমাত্রা হল 44°C (in June) এবং সর্বনিম্ন হল 0.4°C (in January) এক্ষেত্রে বার্ষিক তাপমাত্রার প্রসার হল (44°C – .4°) = 43.6°C এভাবে কোন Class (শ্রেণীর) ছাত্রদের সর্বোচ্চ ও সর্বনিম্ন নম্বরের পার্থক্য, কোন স্থানের জনগোষ্ঠীর বয়সের পার্থক্য, আয়ের পার্থক্য প্রসারের (Range) মাধ্যমে প্রকাশ করা যায়।

উদাহরণ : Rs. 8, Rs. 5, Rs. 10, Rs. 7, Rs. 12, Rs. 6, মানগুলির প্রসার নির্ণয় করুন।

(Rs. 12 – Rs. 7) = Rs. 5 Answer

উদাহরণ : 3, 5, 8, –1, 4 এই মানগুলির প্রসার নির্ণয় করুন।

সর্বোচ্চমান = 8

সর্বনিম্ন মান = –1

প্রসার = (8) – (–1) = 9 Answer

শ্রেণীবন্ধ পরিসংখ্যার বিভাজনের ক্ষেত্রে প্রসারের পরিমাপ হল উচ্চতম শ্রেণীর উর্ধ্বসীমা এবং নিম্নতম শ্রেণীর নিম্নসীমার অন্তরকলের সমান।

$$\text{প্রসারাঙ্ক (Co-efficient of Range)} = \frac{\text{প্রসার}}{\text{উচ্চতম ও নিম্নতম মানের যোগফল}}$$

5.4 সুবিধা ও অসুবিধা

বোঝবার পক্ষে সরল। গণনার পক্ষে সহজ। কিন্তু অসুবিধাও অনেক আছে। খুব বেশী উচ্চ মান ও খুব বেশী নিম্ন মানের উপস্থিতি দ্বারা এটি প্রভাবিত হয়। চলকের সব মানের উপর নির্ভরশীল নয়। মুক্ত পান্তের পরিসংখ্যা বিভাজন (open class) থেকে প্রসার গণনা করা যায় না।

5.5 চতুর্থক পার্থক্য (Quartile Deviation)

বিস্তৃতির পরিমাপ প্রকাশের দ্বিতীয় পদ্ধতিটি হল চতুর্থক পার্থক্য (Quartile Deviation বা Semi Interquartile deviation) চলকের এক প্রস্থ মানগুলিকে উর্ধ্বক্রমে সাজালে ঠিক মধ্যম অবস্থানের মানটি সমগ্র শ্রেণীটিকে দুই অংশে বিভক্ত করে এবং একে বলা হয় মধ্যমা। এইভাবে এইরূপ তিনটি মান Q_1, Q_2, Q_3 নেওয়া হয় যারা সমগ্র শ্রেণীকে চারটি অংশে বিভক্ত করে। এই তিনটি মান Q_1, Q_2, Q_3 কে যথাক্রমে প্রথম (বা নিম্ন), চতুর্থক, দ্বিতীয় চতুর্থক এবং তৃতীয় (বা উর্ধ্ব) চতুর্থক বলা হয়। অর্ধ আন্ত চতুর্থক প্রসার (Semi Interquartile Range) বা চতুর্থক পার্থক্য (Quartile deviation) হল প্রথম (বা নিম্ন) এবং তৃতীয় (বা উর্ধ্ব) চতুর্থকের তফাতের অর্ধাংশ অর্থাৎ

$$\text{Q.D.} = \frac{Q_3 - Q_1}{2}$$

Q_1 = প্রথম (বা নিম্ন) চতুর্থক

Q_3 = তৃতীয় (বা উর্ধ্ব) চতুর্থক

এটি বিস্তৃতির একটি পরম পরিমাপ। যে রাশিতথ্যে বিস্তৃতি স্বাভাবিক (Normal distribution) সেখানে আন্ত চতুর্থক প্রসার এর (Interquartile Range) অর্ধভাগ ও মধ্যমা ও যে কোন চতুর্থকের (any Quartile) পার্থক্য সমান হয়। অর্থাৎ কোন রাশিতথ্যের বিস্তারের 50% দৃষ্টান্ত Semi interquartile প্রসার ± মধ্যমার মধ্যে থাকে।

উদাহরণ : নিম্নে প্রদত্ত 7 জন ব্যক্তির দৈনিক মজুরী (In Rs) চতুর্থক পার্থক্য নির্ণয় করুন।

12, 7, 15, 10, 19, 17, 25

সমা : রাশিগুলিকে মানের উর্ধ্বক্রমে লিখলে পাওয়া যায়।

7, 10, 12, 15, 17, 19, 25.

এখানে $n = 7 \; \frac{n+1}{4} = \frac{7+1}{4} = 2, \; \frac{3(n+1)}{4} = \frac{3(7+1)}{4} = 6$

$\therefore \; Q_1 = \left(\frac{n+1}{4}\right)$ তম মান = দ্বিতীয় মান = 10.

এবং $Q_3 = \frac{3(n+1)}{4}$ তম মান = ষষ্ঠ মান = 19.

সুতরাং চতুর্থক পার্থক্য $= \frac{Q_3 - Q_1}{2} = \frac{19-10}{2}$ – Rs. 4.5 Answer.

উদাহরণ : নিম্নে প্রদত্ত সারণী থেকে চতুর্থক পার্থক্য নির্ণয় কর।

Monthly expenditure		Fe	Cu fe
<Rs. 1000	–	30	30
Rs. 1000 – 1250	–	45	75
Rs. 1250 – 1500	–	70	145
Rs. 1500 – 1750	–	82	227
Rs. 1750 – 2000	–	66	293
Rs. 2000 – 2250	–	57	350
Rs. 2250 – 2500	–	28	378
Rs. 2500 – 2750	–	22	400
Rs. 2750 – 3000	–	18	418
>Rs. 300		12	430
		$\Sigma fe = 430$	

$$n = 430 \qquad \frac{n}{4} = 107.5 \qquad \frac{3n}{4} = 322.5$$

ক্রমযৌগিক পরিসংখ্যান থেকে বোঝা যাচ্ছে যে 1250 – 1500, এবং 2000 – 2250 শ্রেণী প্রসারের (Class interval) যথাক্রমে Q_1 এবং Q_3 বর্তমান।

তাই $$Q_1 = 1250 + \left(\frac{107.5 - 75}{70}\right) \times 250$$

$$= 1250 + \left(\frac{8125}{70}\right)$$

$$= 1250 + 116.07 = \text{Rs. } 1366.07$$

এবং $$Q_3 = 2000 + \left(\frac{322.5 - 293}{57}\right) \times 250$$

$$= 2000 + \left(\frac{322.5 - 293}{57}\right) \times 250$$

$$= 2000 + \left(\frac{7375}{57}\right)$$

$$= 2000 + 129.38 \quad = \text{Rs. } 2129.38$$

$$Q_D = \frac{Q_3 - Q_1}{2} \qquad = \frac{\text{Rs. } 2129.38 - \text{Rs. } 1366.07}{2}$$

$$= \frac{763.31}{2} = \text{Rs. } 381.65 \quad \text{Answer}$$

5.6 সুবিধা ও অসুবিধা

চতুর্থক পার্থক্য গণনা সহজ। এই গণনা প্রথম ও তৃতীয় চতুর্থকের উপর নির্ভরশীল। কেবল Q_1 এবং Q_3-র উপর নির্ভরশীল হওয়ার সমুদায় মানের পরিবর্তনশীলতা-এর বিবিচনাধীন নয় এবং সেইজন্য বাস্তবক্ষেত্রে এর বিবেচনাধীন নয় এবং সেইজন্য বাস্তবক্ষেত্রে এর ব্যবহার খুব বেশী নয়। তবে মুক্ত প্রান্ত বিশিষ্ট শ্রেণীযুক্ত শ্রেণীবদ্ধ পরিসংখ্যা বিভাজন থেকে এটি গণনা করা যায়।

5.6 গড় পার্থক্য (Mean Deviation)

কোন চলকের মানগুলির কোন শ্রেণীর গড় পার্থক্য হল ঐ শ্রেণীর যে কোন একটি গড় (গড়, মধ্যমা) থেকে সমুদয় পরম পার্থক্যগুলির যৌগিক গড়। এটি বিস্তৃতির পরম পরিমাপ। একপ্রস্ত n সংখ্যক মান $x_1, x_2 \ldots x_n$-এর A.M. সাপেক্ষে গড় পার্থক্যের সংজ্ঞা হল গড় পার্থক্য = M.D. $= \dfrac{\sum_{j=1} |x_j - \bar{x}|}{N} = \dfrac{\sum |x - \bar{x}|}{N} = \left|\overline{x - x}\right|$

যেখানে $\bar{x}$ হল সংখ্যা সমূহের মধ্যমা বা A.M. (Arithmetic Mean) এবং $|x_j - \bar{x}|$ হল $\bar{x}$-র থেকে পার্থক্যের পরম মান।

উদাহরণ (১) : নিম্নলিখিত অশ্রেণীবদ্ধ (ungrouped) রাশিতথ্যের গড় পার্থক্য বা Mean Deviation নির্ণয় করুন।

2, 3, 6, 8, 11

Arithmetic Mean (A.M.) বা মধ্যমা

$$\bar{x} = \frac{2+3+6+8+11}{5} = 6$$

Mean Deviation = M.D. = গড় পার্থক্য

$$\frac{|2-6|+|3-6|+|6-6|+|8-6|+|11-6|}{5}$$

$$= \frac{|-4|+|-3|+|0|+|2|+|5|}{5}$$

$$= \frac{4+3+0+2+5}{5}$$

$$= 2.8 \text{ Answer}$$

পরিসংখ্যা বিভাজন অর্থাৎ শ্রেণীবস্তু রাশিতথ্য (Grouped Data)-র ক্ষেত্রে

$$M.D. = \frac{\sum f|x - \bar{x}|}{N}$$

f = শ্রেণীর পরিসংখ্যা

$\bar{x}$ = গাণিতিক গড়

N = মোট পরিসংখ্যান

N.B. : রাশি কে মড্ । d । এইভাবে পড়া হয় এবং এটি d-এর চিহ্ন নিরপেক্ষ সংখ্যা বা পরম মানকে নির্দেশ করে যেমন

$$|-3| = 3, |+4| = |4|, |-0.67| = .67$$

পার্থক্যগুলির বীজগাণিতিক মান না নিয়ে কেবল পরম মান নেবার কারণ হল গড় থেকে মানগুলির পার্থক্যসমূহের বীজগাণিতিক যোগফল শূন্য।

উদাহরণ (২) : ১০০ জন ছাত্রের ভরের (mass) গড় পার্থক্য বার কর।

$\bar{x}$ = 67.45 kg.

Mass (kg)	class mark x	$x - \bar{x}$	f	$f(x - \bar{x})$
60-62	61	6.45	5	32.25
63-65	62	3.45	18	62.10
66-68	67	0.45	42	18.90
69-71	70	2.55	27	68.85
72-74	71	5.55	08	44.40
			$N = \sum f = 100$	$\sum f|x - \bar{x}| = 226.50$

গড় পার্থক্য $= M.D. = \frac{\sum f|x - \bar{x}|}{N} = \frac{226.50}{100}$

= 2.26 kg. Answer.

উদাহরণ (৩) : কোন একটি জেলার (district) C-D. Block-এর No. of Household (পরিবারের সংখ্যা) দেওয়া হল। এই সংখ্যাগুলির যৌগিক গড় সাপেক্ষে গড় পার্থক্য নির্ণয় কর।

C.D. Block	No. of household
1	31
2	35
3	29
4	63
5	55
6	72
7	37

যৌগিক গড় $\bar{x} = \frac{31+35+29+63+55+72+37}{7}$

$= \frac{322}{7} = 46$

মান (x)	গড় থেকে পার্থক্য $d = x - \bar{x}$ $(x - 46)$	পরম পার্থক্য $\lvert d \rvert$
31	– 15	15
35	– 11	11
29	– 17	17
63	17	17
55	9	9
72	26	26
37	– 9	9
Total $\sum x = 322$		$\sum \lvert d \rvert = 104$

∴ গড় সাপেক্ষে নির্ণেয় গড় পার্থক্য

$$= \frac{\sum \lvert d \rvert}{n} = \frac{104}{7} = 14.86 \text{ Household}$$

5.8 সুবিধা ও অসুবিধা

গড় পার্থক্য চলকের সমুদয় মানের উপর নির্ভরশীল এবং কখনও কখনও এর থেকে বিস্তৃতির পরিমাপ হিসেবে মোটামুটি ভাল ফল পাওয়া যায়। কিন্তু গড় পার্থক্য গণনার জন্য চিহ্ন বর্জিত পরম পার্থক্যগুলি নেওয়ার রীতি অযৌক্তিক বলে মনে হয় এবং এর ফলে বীজগণিত সূত্র প্রয়োগ কঠিন হয়ে পড়ে। যদিও পূর্ববর্তী পরিমাপগুলিতে চরম মানগুলি নেওয়া হয়, এখানে সেইরূপ হয় না, রাশিতত্ত্বের বিস্তারিত আলোচনার জন্য এই পদ্ধতি খুব একটা উপযোগী নয়।

5.9 সমক পার্থক্য (Standard Deviation)

সমক পার্থক্য বিস্তৃতির সর্বাপেক্ষা গুরুত্বপূর্ণ পরিমাপ। সমক পার্থক্য হল গড় থেকে পার্থক্যগুলির বর্গসমূহের গড়ের বর্গমূল। সমক পার্থক্য কে সাধারণতঃ গ্রীক অক্ষর δ (সিগমা) বলে পড়তে হয়। সমক পার্থক্যের আর একটি সংজ্ঞা বলা যেতে পারে যে কোন চলকের এক প্রসুমানের সমক পার্থক্য (S.D.) হল মানগুলির যৌগিক গড় থেকে,

তাদের সকল পার্থক্যের বর্গগুলির যৌগিক গড়ের ধনাত্মক বর্গমূল। এখানে নঞর্থক পার্থক্য (negative deviation) বাদ দিয়ে দাওয়া হয় বলে গাণিতিক ক্ষেত্রে এর বহুল ব্যবহার দেখা যায় ও অত্যন্ত জনপ্রিয় এই বিস্তৃতি পরিমাপ।

যদি কোন চলকের মানের একটি শ্রেণী $x_1, x_2, x_3 \ldots x_n$ হয় এবং যদি তাদের A.M. হয় তবে S.D. (δ) এর সংজ্ঞা অনুসারে

$$\delta = \sqrt{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + (x_n - \bar{x})^2}$$

অথবা $\sqrt{\dfrac{\sum (x - \bar{x})^2}{N}}$

পরিসংখ্যা বিভাজন এর (Grouped frequency)

ক্ষেত্রে $\delta = \sqrt{\dfrac{\sum f(x - \bar{x})^2}{N}}$

উদাহরণ (১) : নিম্নলিখিত রাশিসমূহের S.D. (δ) নির্ণয় কর।

12, 6, 7, 3, 15, 10, 18, 5

গড় বা $\bar{x} = \dfrac{\sum x}{n} = \dfrac{12+6+7+3+15+10+5}{8}$

$$\delta = \sqrt{\frac{(12-95)^2 + (6-9.5)^2 + (7.95)^2 + (3-9.5)^2 + (15-95)^2 + (10-9.5)^2 + (18.95)^2 + (5.95)^2}{8}}$$

$= \sqrt{23.75} = 4.87$ Answer.

উদাহরণ (২) : 100 জন ছাত্রের ভরের S.D. বা সমক পার্থক্য নির্ণয় কর : $\bar{x} = 67.45$

Mass (kg)	Class mark (x)	$x - \bar{x}$ (n – 67.45)	$(x - \bar{x})^2$	F	$F(x - \bar{x})^2$
60 – 62	61	– 6.45	41.60	5	208.0
63 – 65	64	– 3.45	11.90	18	214.25
66 – 68	67	– C.45	0.20	42	8.50
69 – 71	70	2.55	6.50	27	175.57
72 – 74	73	5.55	30.80	08	246.42
	$\bar{x} = 67.45$			$N = \sum f$ = 100	$\sum F(x - \bar{x})^2$ = 852.75

$$\delta = \sqrt{\frac{\sum F(x-\bar{x})^2}{N}} = \sqrt{\frac{852.75}{100}} = \sqrt{8.5275} = 2.92 \text{ Kg} \quad (\text{Answer})$$

সমক পার্থক্য আমরা ভেদমান (variance) নির্ণয় করেও বার করে থাকি।

Variance : সমক পার্থক্যের বর্গ ভেদমান (variance) নামে পরিচিত। অর্ক্য ভেদমান = $(\delta)^2 = (S.D.)^2$ ভেদমানের ধনাত্মক বর্গমূলকে S.D.-র সংজ্ঞা হিসেবে ধরা হয়।

অতএব "It is the average of sum of squares of deviation of individual sizes of the variable from Arithmetic mean, multiplied by corresponding frequency.

আমরা প্রথমে নিয়ে থাকি :

1) (Variables) চলকের গড় থেকে বিস্তৃতির (size)

2) বিস্তৃতির বর্গ (square)

3) বর্গকে উল্লেখিত frequency দিয়ে গুণ করতে হবে

4) প্রতিটি মাপের বর্গের বিস্তৃতিকে যোগ করতে হবে

5) (Δ) যোগফলকে মোট frequency দিয়ে ভাগ করতে হবে।

$S.D.^2$ = বর্গমূল নিয়ে S.D. পাওয়া যাবে।

উদাহরণ : নীচের সারণি থেকে সমপার্থক্য (S.D.) নির্ণয় কর।

x	f
1	6
2	8
3	10
4	12
Total	$\sum f = 36$

$$\delta^2 = \frac{\sum(x_1 - \bar{x})^2}{N}$$

$$\bar{x} \text{ or A.M.} = \frac{6\times1 + 8\times2 + 10\times3 + 12\times4}{36}$$

$$= 2.8$$

$$\delta^2 = \frac{6(1-2.8)^2 + 8(2-2.8)^2 + 10(3-2.28)^2 + 12(4-2.28)^2}{36}$$

$$= 1.17$$

$$\delta = 1.08$$

Method (L) $\delta^2 = \left\{ \frac{fd^2}{N} - \left(\frac{fd}{N} \right)^2 \right\}$

ইচ্ছামত নির্বাচিত মূল বিন্দু থেকে পার্থক্য নিয়ে S.D.^{2}/S.D. নির্ণয়।

A = কোন নির্বাচিত মূলবিন্দু, ধরা যাক x = 3 এবং 3 এর থেকে বিস্তৃতি নেওয়া হল এবং fd table তৈরী করা হল। d = (x – A)

x		f	fx	fd^2
1	(–2)	6	–12	24
2	(–1)	8	–8	8
3	(0)	10	0	0
4	(1)	12	12	12
Total		$\sum f = 36$	$\sum fd = -8$	$\sum fd^2 = 44$

$$\delta^2 = \frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N} \right)^2$$

$$-\frac{44}{36} - \left(\frac{-8}{36} \right)^2 = 1.17$$

$\delta = 1.08$ Answer.

পরিসংখ্যা বিভাজনের ক্ষেত্রে সূত্রটি হবে

$$\delta^2 = \left\{ \frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N} \right)^2 \right\} \times C^2.$$

উদাহরণ : $$\delta = \sqrt{\left\{ \frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N} \right)^2 \right\} \times C^2}$$

Class Interval	x	f	d	fd	d^2	fd^2
0 – 10	5	6	– 2	– 12	4	24
10 – 20	15	8	– 1	– 8	1	8
20 – 30	25	10	0	0	0	0
30 – 40	35	12	1	12	1	12
	$\sum f$ or $N = 36$		$\sum fd = -8$			$\sum fd^2 = 44$

$$\varepsilon^2 = \left\{ \frac{44}{36} - \left(\frac{-6}{36} \right)^2 \right\} \times C^2 = 1.17 \times (10)^2$$

$$\delta = \sqrt{1.17} \times 10 = 10.8 \text{ Answer}$$

উদাহরণ : নিম্নের রাশিগুলি কোন স্থানের গড় বৃষ্টিপাতের পরিমাণ নির্দেশ করছে। এই রাশিতথ্যগুলি থেকে সমক পার্থক্য নির্ণয় করুন।

প্রথম পদ্ধতি :

মান (x) (in cm)	গড় 56 থেকে পার্থক্য $d = (x - \bar{x})$	d^2
49	−7	49
63	7	49
46	−10	100
59	3	৯
65	9	81
52	−4	16
60	4	16
54	−2	4
Total $\sum x = 448$		$\sum d^2 = 324$

গড় $\bar{x} = \frac{\sum x}{N} = \frac{448}{8}$

$= 56$

সমক পার্থক্য $= \sqrt{\frac{\sum d^2}{N}} = \sqrt{\frac{324}{8}}$

$= \sqrt{40.5} = 6.36$ cm.

দ্বিতীয় পদ্ধতি :

মান (x) (in cm)	ইচ্ছামত নির্বাচিত মূল মূল বিন্দু A থেকে পার্থক্য $d = (x - A) = (x - 54)$	d^2
49	−5	25
63	৯	81
46	−৪	64

মান (x) (in cm)	ইচ্ছামত নির্বাচিত মূল মূল: বিন্দু A থেকে পার্থক্য $d = (x - A) = (x - 54)$	d^2
59	5	25
65	11	121
52	−2	04
60	6	36
54 = A	0	0
	$\sum d = 16$	$\sum d^2 = 356$

$n = 8$

$$\delta = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2}$$

$$= \sqrt{\frac{356}{8} - \left(\frac{16}{8}\right)^2} = \sqrt{44.5 - 4}$$

$$= \sqrt{40.5} = 6.36 \text{ cm}$$

উদাহরণ : নিম্নের ছকে কোন একটি বিদ্যালয়ে 540 ছাত্রদের (marks) নাম্বারের বিভাজন দেওয়া হল। এর থেকে সমক পার্থক্য নির্দেশ করুন।

প্রাপ্ত নাম্বার (marks obtained)	30	40	50	60	70
ছাত্রদের সংখ্যা (no. of students)	64	132	153	140	51

প্রাপ্ত নম্বর x	A (= 50 থেকে পার্থক্য) $d = (x - 50)$	ছাত্র সংখ্যা f	fd	fd^2
30	−20	64	−1280	25600
40	−10	132	−1320	13200
50 = A	0	153	0	0

প্রাপ্ত নম্বর x	A (= 50 থেকে পার্থক্য) d = (x − 50)	ছাত্র সংখ্যা f	fd	fd^2
60	10	140	1400	14000
70	20	51	1020	29400
		$\sum f$ or N = 540	$\sum fd = -2600$ 2420 = − 180	$\sum fd^2 = -73200$

$$\therefore \text{ সমক পার্থক্য } \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{n}\right)^2}$$

$$= \sqrt{\frac{73200}{540} - \left(\frac{180}{540}\right)^2}$$

$$= \sqrt{135.555 - .1111}$$

$$= \sqrt{135.4444}$$

= 11.64 নম্বর Answer

5.10 সুবিধা ও অসুবিধা

১) সমক পার্থক্য হল বিস্তৃতির পরিমাপগুলির মধ্যে সর্বপেক্ষা গুরুত্বপূর্ণ ও ব্যাপক ভাবে ব্যবহৃত পরিমাপ।

২) বিস্তৃতির সুপরিমাপের জন্য সব গুণ এখানে বর্তমান।

৩) এর সংজ্ঞা সম্পূর্ণ, দ্ব্যর্থহীন, চলকের সব মানের উপর নির্ভরশীল।

৪) বীজগণিতের সূত্রগুলি প্রয়োগের পক্ষে এটি উপযুক্ত। বিস্তৃতির অপর যে কোন পরম পরিমাপের তুলনায় S.D. নমুনা বিচ্যুতির দ্বারা সর্বাপেক্ষা কম পরিমাণে প্রভাবিত হয়।

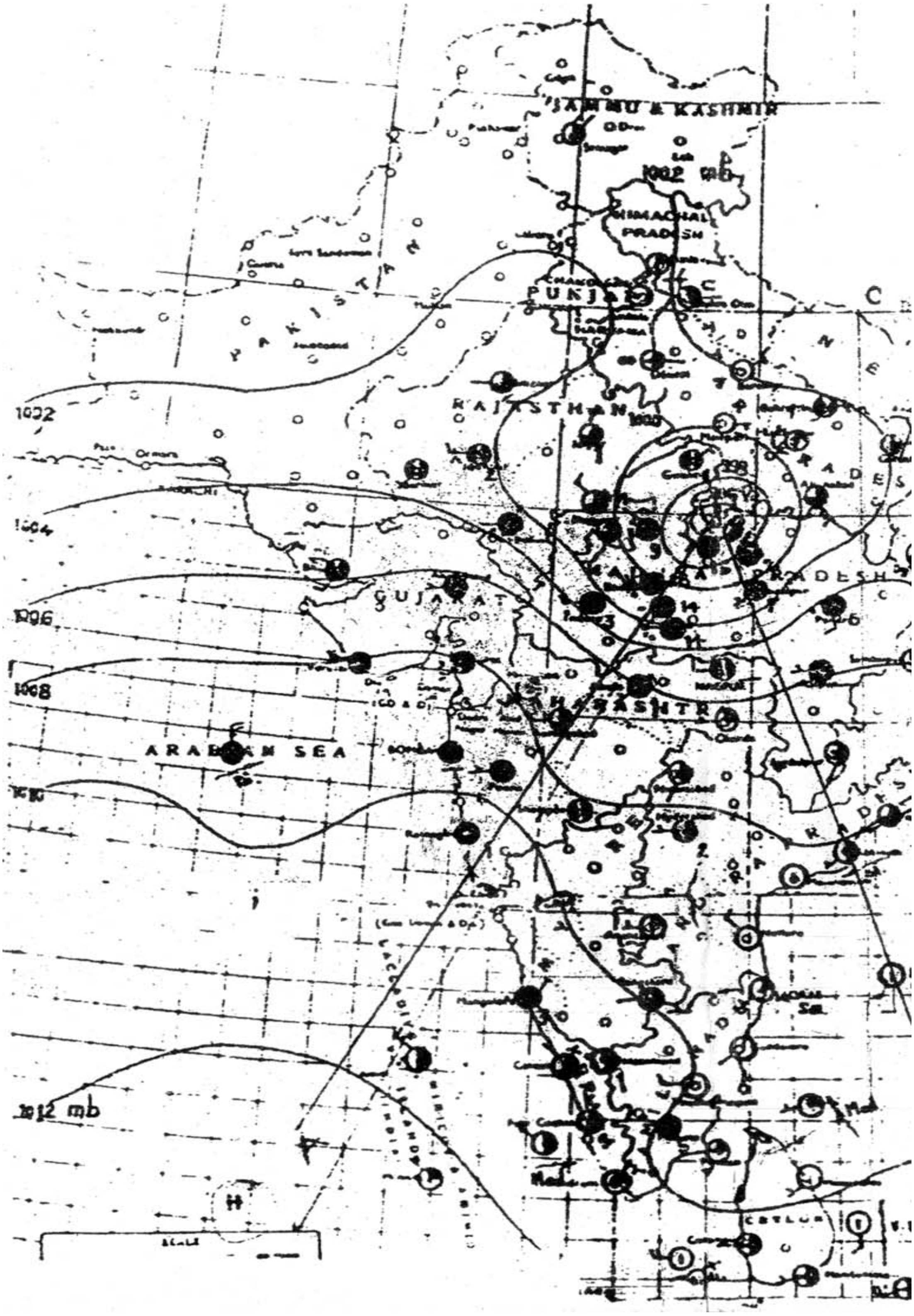
তবে S.D. সব সময় সহজবোধ্য নাও হতে পারে। A.M. থেকে পার্থক্যগুলির বর্গ করে এই বর্গগুলির A.M. এর বর্গমূল নেওয়া একটি জটিল প্রক্রিয়া। মূলবিন্দু ও স্কেলের সুবিধামত পরিবর্তন দ্বারা S.D. গণনা সহজতর করা যেতে পারে। বিভিন্ন এককে প্রদত্ত দুই বা ততোধিক বিভাজনের পরিবর্তনশীলতার তুলনাতে একে ব্যবহার করা যায় না।

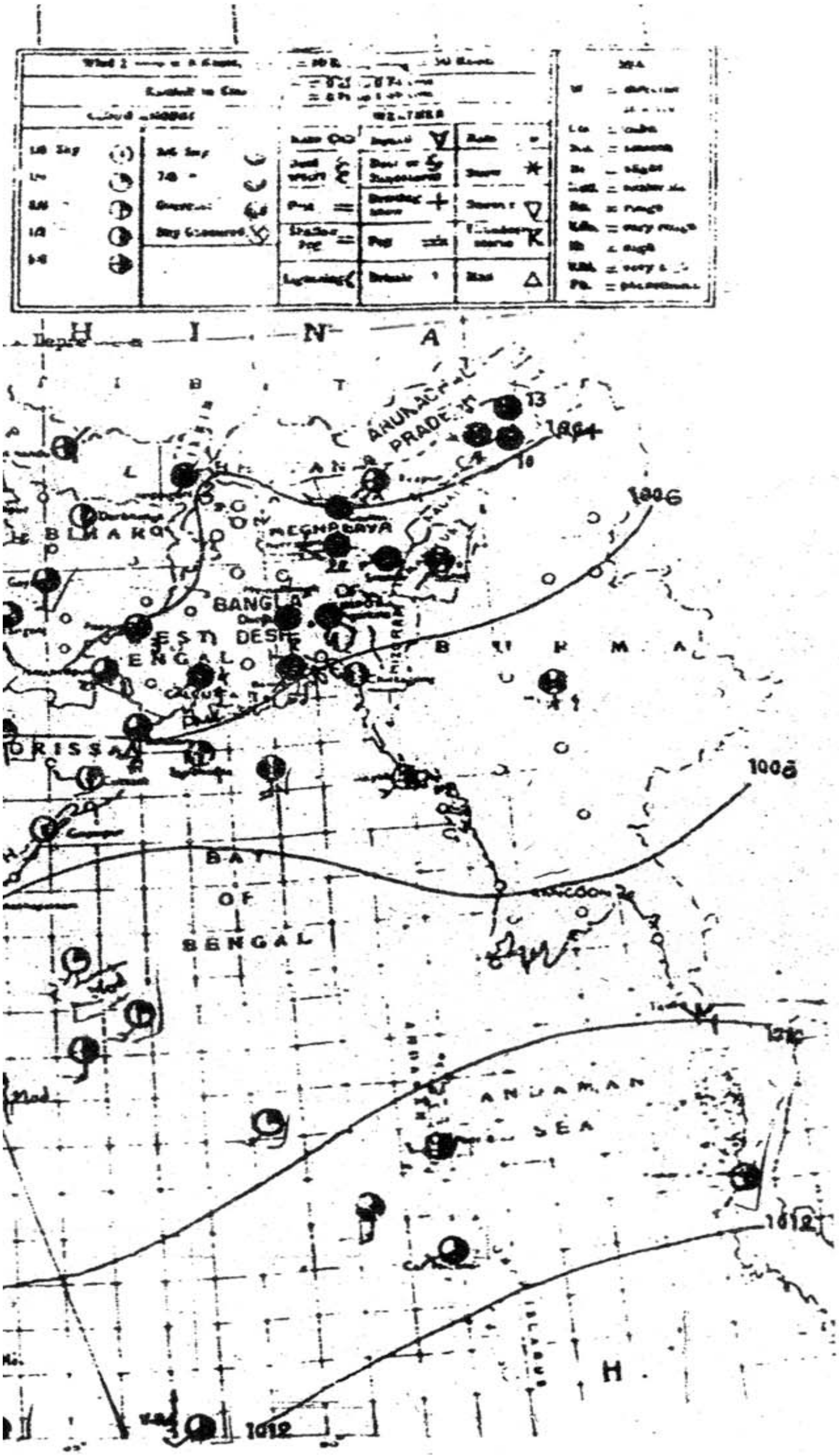
INDIAN DAILY WEATHER REPORT

WEATHER MAP AT 0830 HRS. I. S. T. (0300 HRS. G M. T.)

12 July

Sunday ~~[illegible]~~ 1973 ([illegible] Asadha 1895 Saka)





ARUNACHAL PRADESH
MEGHALAYA
BANGLA DESH
WEST BENGAL
CALCUTTA
ORISSA
BURMA
BAY OF BENGAL
ANDAMAN SEA
1006
1008
1012
H

5.11 অনুশীলনী

1. **জন ছাত্রের প্রাপ্ত নম্বরের প্রসার নির্ণয় কর :**

9, 7, 25, 18, 38, 12, 30, 35.

উঃ 31 টা

2. **নিম্নের পরিসংখ্যা বিভাজনের গড় পার্থক্য নির্ণয় কর :**

Size of family (X)	–	5	7	9	11	13.
No. of family (F)	–	4	10	22	10	4.

উঃ 1.44.

3. **200 টি গাছের ফুলের সংখ্যা থেকে সমক পার্থক্য নির্ণয় কর :**

ফুলের সংখ্যা	–	50	55	60	65	70.
গাছের সংখ্যা	–	30	40	65	50	15.

উঃ 5.79.

4. **নীচের সারণী থেকে সমক পার্থক্য ও গড় পার্থক্য নির্ণয় কর।**

Station	(Alipur)	(rainfall in mm)
J	–	2.2
F	–	6.3
M	–	126.3
A	–	188.1
M	–	228.1
J	–	478.5
J	–	915.1
A	–	866.4
S	–	726.4
O	–	410.4
N	–	34.3
D	–	1.6

উঃ S.D. = 328.52 mm

M.D. = .989456 mm.

5. **নীচের সারণী থেকে সমক পার্থক্য ও গড় পার্থক্য নির্ণয় কর।**

Station x		(rainfall in mm)
J	–	0.3
F	–	1.1
M	–	22.6

A	–	34.3
M	–	32.5
J	–	317.7
J	–	307.9
A	–	475.1
S	–	403.4
O	–	55.7
N	–	[illegible]4
D	–	1.6

উঃ S.D. = 172.77 mm

M.D. = 1.2425 mm.

6. **নীচের Frequency Table থেকে Quartile deviation চতুর্থক পার্থক্য নির্ণয় কর।**

Class boundary			–	frequency
0.5	–	400.5	–	05
400.5	–	800.5	–	11
800.5	–	1200.5	–	09
1200.5	–	1600.5	–	03
1600.5	–	2000.5	–	02

উঃ 299 (approx)

7. **নীচের সারণী থেকে গড় পার্থক্য ও সম্যক পার্থক্য নির্ণয় কর।**

Class boundary			–	frequency
0	–	50	–	8
50	–	100	-	10
100	–	150	–	16
150	–	200	–	9
200	–	250	–	4
250	–	300	–	3

উঃ গড় পার্থক্য (M.D.) = 52

S.D. = 68.55

8. নীচের সারণী থেকে S.D. এবং M.D. নির্ণয় কর।

District	–	(yield of Awan (kg/hec) (2001)
Burdwan	–	2029
Birbhum	–	2327
Bankura	–	2435
Midnapore	–	2117
Howrah	–	1413
24 Parganas (N)	–	1844
24 Parganas (S)	–	1873
Nadia	–	1595
Purulia	–	1818
Hooghli	–	2217

উঃ S.D. = 304.78 (kg/hec)

M.D. = 258.24 (kg/hec)

9. নীচের পরিসংখ্যা বিভাজন থেকে চতুর্থক পার্থক্য নির্ণয় কর।

Class Group	–	ব্যক্তির সংখ্যা
40 – 45	–	10
45 – 50	–	22
50 – 55	–	28
55 – 60	–	20
60 – 65	–	12
65 – 70	–	08

উঃ Q.D. = 5.17

10. নীচের রাশিতথ্য থেকে S.D. সমক পার্থক্য গণনা কর :

উচ্চতা (মঞ্চ)	–	59–61	61–63	63–65	65–67	67–69
ছাত্র সংখ্যা	–	4	30	45	15	6

উঃ S.D. = 1.83 মঞ্চ

একক 6 □ সহজ দ্বিচলকের প্রতিগমান ও সহপরিবর্তন, কালীন শ্রেণী বিশ্লেষণ

গঠন

6.1 প্রস্তাবনা

পূর্ববর্তী এককগুলিতে আমরা একক চলকের বিভাজন বৈশিষ্ট্য বৃষ্টিপাত, আয়, বয়স, ওজন প্রভৃতির বিভাজন সমূহের মধ্যমা গণনা, বিস্তৃতির পরিমাপ ইত্যাদি আলোচনা করেছি। এখন দুইটি চলকের সম্বন্ধে একই সঙ্গে

আলোচনা করা হবে এবং তাদের পরিমাণগত সম্বন্ধটি নির্ণয়ের চেষ্টা করা হবে। যেমন আমরা Census থেকে জনসংখ্যার বিবিধ তথ্য অর্থাৎ জনগণের বয়স, কর্ম, শিক্ষা, আয় ইত্যাদি রাশিতথ্য সংগ্রহ তাদের মধ্যে (cause & effect) কার্যকারণ সম্পর্ক স্থাপন করতে পারি। (১) যেমন মানুষের উচ্চতা ও 3 জন একটি অপরটির উপর নির্ভরশীল (২) কোন বস্তুর চাহিদা ও দামের সম্পর্ক (৩) বৃষ্টিপাত ও কৃষি উৎপাদনতা (৪) উৎপাদন ও দ্রব্যমূল্য (৫) জন সংখ্যার শিক্ষা ও দারিদ্রতা ইত্যাদি।

6.2 দ্বিচলকের প্রতিগমান ও সহপরিবর্তন [Bivariate (Linear) Regression with Correlation)

6.2.1 সহপরিবর্তন (Correlation) :

দুইটি চলকের রাশি যখন একই সঙ্গে নথিভুক্ত করা হয় তখন তাকে দ্বিচলক বলে। যেমন (১) কিছু পরিবারের আয় ও ব্যায় (২) কোন রাজ্যের পর পর কিছু বছরের বৃষ্টিপাতের পরিমাণ ও ধানের ফলন ইত্যাদি। দুটি চলরাশির মধ্যে সম্বন্ধের মাত্রা নির্ণয়ের জন্য পরিসংখ্যানীয় পদ্ধতি হল 'সহপরিবর্তন'। 'সম্বন্ধ' বলতে আমরা এখানে বলতে পারি 'চলক' দুটি একই সঙ্গে পরিবর্তনের প্রবণতা। চলক দুটিকে 'x' ও 'y' বলা হয়। স্বাধীন (Independent) বা ব্যাখ্যাকৃত (explained) চলকে 'x' এবং অধীন (Dependent) বা ব্যাখ্যাকারী (explaining) চলকে 'y' বলা হয়। যদি 'চলক) অর্থাৎ 'x' এবং 'y' এইরূপ সম্বন্ধযুক্ত হয় যে একটির গতির (বা পরিবর্তনের) সঙ্গে সঙ্গে অপর চলক এর অনুরূপ গতির (বা পরিবর্তনের) প্রবণতা থাকে, অর্থাৎ 'x' এর গতির সঙ্গে 'y' এর অনুরূপ গতির প্রবণতা থাকে তাকে সহপরিবর্তন ধর্মের অধীন (correlated) বলা হয়। গতিগুলি একই অভিমুখে হতে পারে (অর্থাৎ x ও y দুটির বৃদ্ধি পেতে পারে।) যেমন কোন অঞ্চলে বৃষ্টিপাত বাড়লে সেই অঞ্চলে আমন ধানের উৎপাদন বৃদ্ধি পায়। গতিগুলি দুটিই হ্রাস পেতে পারে। যেমন বৃষ্টিপাত কমলে পাট উৎপাদন কম হবে। গতিগুলি পরস্পর বিপরীত অভিমুখে হতে পারে যেমন চিকিৎসা শাস্ত্রের উন্নতি ও মৃত্যুহার হ্রাস। চলক দুটির গতি একই দিকে হলে তাকে ধনাত্মক সহপরিবর্তন (Positive Correlation) বা বিপরীত দিকে হলে ঋণাত্মক সহ পরিবর্তন (negative correlation) বলা হয়। যদি x এর কোন পরিবর্তনের জন্য y অপরিবর্তিত থাকে বা স্থির থাকে তাকে বলা হয় পরস্পর সম্বন্ধহীন বা (uncorrelated) যেমন কোন অঞ্চলের 'উষ্ণতা' ও 'শিক্ষার হার'।

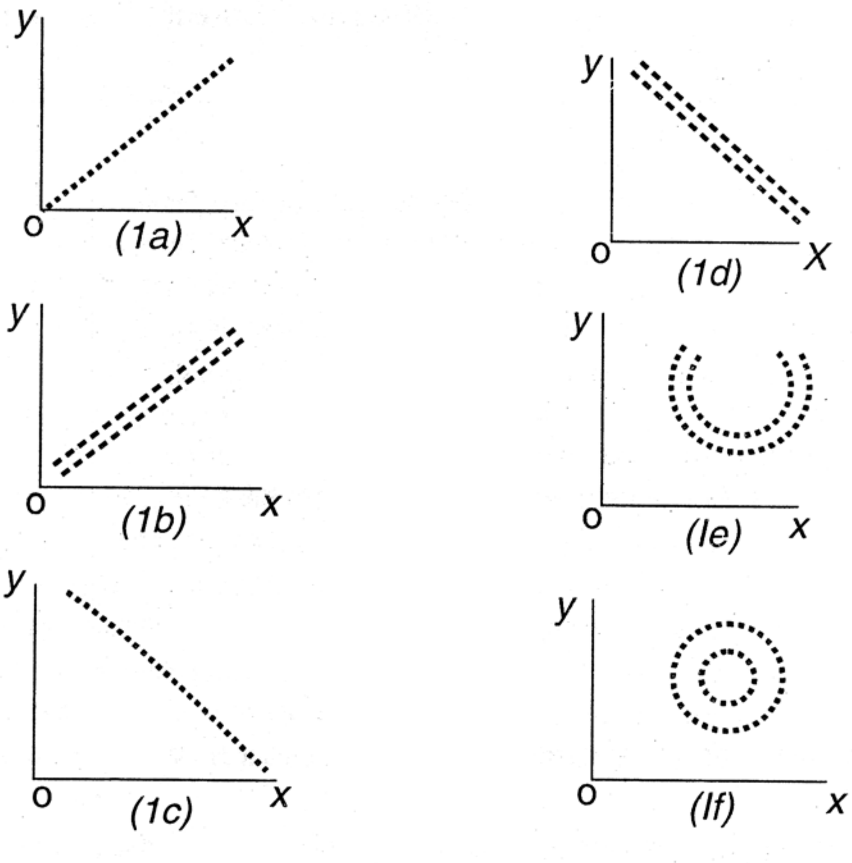
L. R. 'Conner'-এর কথায় 'যদি দুই বা ততোধিক রাশির পরিবর্তন ঐকতানিক হয়, যাহাতে একটির গতির সাথে সাথে অপরটি বা অপরগুলিও অনুরূপ গতির অধীন হয়, তবে উহাদের সহপরিবর্তন ধর্মবিশিষ্ট বলা হয়।" "If two or more quantities very in sympathy so that movements in the one tend to be accompanied by corresponding movements in the other than they are said to be correlated".

সহ পরিবর্তন রৈখিক (linear) অথবা অরৈখিক (non linear) হতে পারে। যদি 'x' এর পরিবর্তনের মান 'y' এর অনুরাগ পরিবর্তনের পরিমাণের সঙ্গে একটি ধ্রুবক অনুপাতে থাকে তবে x ও y এর সহ পরিবর্তনকে সরল, বা রৈখিক বা একমাত্রিক বলা হয়। অন্যরূপ হলে তাকে অরৈখিক সহপরিবর্তন বলা হয়।

দুটি চলকের সহপরিবর্তন নির্ণয় (Determination of correlation Scatter Diagram বা বিক্ষিপ্ত চিত্রের মাধ্যমে করা হয়ে থাকে বিক্ষিপ্ত চিত্র (Scatter Diagram).

দ্বিচালক রাশিতথ্যের লৈখিক প্রকাশ বিক্ষিপ্ত চিত্র নামে পরিচিত। এখানে সহগতি (বা পরিবর্তন) সম্বন্ধীয় পরিসংখ্যানীয় রাশিতথ্যকে বিন্দুর সাহায্যে লৈখিক পদ্ধতিতে প্রকাশ করা যায়। দুটি চলকের মধ্যে স্বাধীন (independent) চলককে 'x' অক্ষ (অনুভূমিক) বরাবর ও অধীন (Dependent) চলককে 'y' (উলম্ব) অক্ষ বরাবর নেওয়া হয়। x এবং y এর যুগলমানকে এইভাবে ছক কাগজে বিন্দু সমূহের মাধ্যমে প্রদর্শন করা হয়।

এই বিন্দুগুলির বিক্ষিপ্ত চিত্র এবং বিক্ষিপ্ততার অভিমুখ চলক দুটির সহপরিবর্তনের প্রকৃতি (Nature) ও মাত্রা (Degree) প্রকাশ করে। কতগুলি উদাহরণ নীচে দেখানো হল।



'1a' এবং '1b'-তে চলক দুটির গতি (বা পরিবর্তন) সমমুখী এবং বিক্ষিপ্ত চিত্রে একটি ঋজুপথ (linear path) দেখা যাচ্ছে। এদের মধ্যে সম্পর্ক সরল (direct) ও ধনাত্মক (positive).

'1c' ও '1d'-তে বিক্ষিপ্ত চিত্রে ঋজুপথে দেখা যাচ্ছে কিন্তু চলক দুটির গতি বিপরীতমুখী এই ক্ষেত্রে সহপরিবর্তন ঋণাত্মক (negative) ও সম্বন্ধ পরোক্ষ (indirect)। চিত্র '1e' এবং '1f'-এ বিন্দুগুলি একটি ঋজুপথের পরিবর্তে একটি বক্র পথ বরাবর অবস্থান করছে (curvilinear) বা একটি ঝাঁক বেঁধে অবস্থান করছে আর ও একটি উদাহরণ দিয়ে দেখানো যেতে পারে।

(a) বৃষ্টিপাত ও ফসল উৎপাদন

(b) একটি শ্রেণীতে ছাত্রদের উচ্চতা ও ওজন।

দ্রষ্টব্য : (a) মানগুলির বিক্ষিপ্ত ভাবে আছে।

(b) মানগুলি একটি রেখা বরাবর থাকার চেষ্টা করছে।

সিদ্ধান্ত : 'a'-এর তুলায় 'b'-এর সহপরিবর্তন বেশী।

6.2.2 রৈখিক সহপরিবর্তনের গণনা (Measurement of Linear Correlation) :

বিক্ষিপ্ত চিত্র আমাদের দুটি চলকের সম্বন্ধ বা সহপরিবর্তন সম্পর্কে একটি ধারণা দিতে পারে। কিন্তু দুটি চলকের সম্বন্ধের শক্তি (মাত্রা) (degree) বা দিক (direction) পরিমাপের পরিমাণগত (Quantitative) বিশ্লেষণ এর জন্য Karl Pearson সূত্রটি গ্রহণ করা হয়।

সূত্রটি হল

$$r=\frac{\sum xy-\frac{\sum x\sum y}{N}}{\sqrt{\sum x^2-\frac{\left(\sum x\right)^2}{N}}\sqrt{\sum y^2-\frac{\left(\sum y\right)^2}{N}}}$$

এই পরিমাপটিকে বলা হয় Product Moment Correlation অথবা Correlation Coefficient বা সহভেদ মান সহ পরিবর্তনের গুণাঙ্ক (Corelation Coefficient) অথবা সহভেদ মান বা Covariance পদ্ধতিতেও করে থাকি। দুটি চলক x এবং y-এর n সংখ্যক যুগ্ম মান (x_1y_1) (x_2y_2) ... (x_ny_n) হয়, তবে x এবং y-এর কোভ্যারিয়েন্স-এর সংজ্ঞা এইরূপ

$$\text{Cov}(x,y)=\sum_{n}^{1}(x-\bar{x})(y-\bar{y})$$

Covariance দুটি চালকের মধ্যে যৌগিক পরিবর্তন নির্দেশ করে। x এবং y-এর সহ পরিবর্তনের গুণাঙ্ক (r)-এর সংজ্ঞা হল

$$r=\frac{\text{Cov}(x,y)}{\delta x\,\delta y}=\frac{\sum(x-\bar{x})(y-\bar{y})}{n\,\delta x\,\delta y}$$

যেখানে $\delta x \delta y$ হল যথাক্রমে x এবং y-এর সমক সহপরিবর্তনের গুণাঙ্ক 'r'-এর সূত্রগুলিকে বিভিন্ন আকারে লেখা যায়

যদি $x=x-\bar{x}=x$-এর মধ্যক থেকে x-এর পার্থক্য এবং

$y=y-\bar{y}=y$-এর মধ্যক থেকে y-এর পার্থক্য তবে

সূত্র ১) হল Karl Pearson-এর সূত্র।

এখন $\delta x=\sqrt{\frac{\sum x^2}{n}-\left(\frac{\sum x}{n}\right)^2}=\sqrt{\frac{\sum x^2}{n}}$ $\qquad\left[\,\therefore\sum x=\sum(x-\bar{x})=0\,\right]$

সেইভাবে $\delta y = \sqrt{\frac{\sum y^2}{n}}$

থেকে $r = \frac{\sum xy}{\sqrt{\frac{\sum x^2}{n}} \times \sqrt{\frac{\sum y^2}{n}}} = \frac{\sum xy}{\sqrt{\sum x^2 \times \sum y^2}}$

6.2.3 সহপরিবর্তন গুণাঙ্কে ব্যবহার ও বৈশিষ্ট্য (Properties and uses of Correlation co-efficient) :

যদি চলকগুলির মান বিক্ষিপ্ত চিত্রে একটি ঋজু রৈখিক পথ নির্দেশ করে, তবে চলকগুলির মধ্যে সম্পর্ক নির্ণয়ের ক্ষেত্রে সহপরিবর্তন গুণাঙ্ককে (Correlation Coefficient) একটি উপযোগী পরিমাপ রূপে বিবেচনা করা হয়। 'r'-এর ধনাত্মক মান নির্দেশ করে যে চলক দুটির গতি সম অভিমুখী, অর্থাৎ x চলকের উচ্চ ও নিম্ন মানের সঙ্গে y চলকের যথাক্রমে উচ্চ ও নিম্ন মানের সংস্রব আছে। r-এর ঋণাত্মক মান নির্দেশ করে যে চলকদ্বয়ের গতি পরস্পরে বিপরীত অভিমুখী, অর্থাৎ একটি চলকের উচ্চ মানগুলির সঙ্গে অপরটির নিম্নমানের সংস্রব আছে। r-এর মান −1 থেকে +1-এর মধ্যে সীমিত থাকে। r-এর মান যদি +1 থেকে −1-এর মধ্যে থাকে তাহলে চলক দুটির মধ্যে সহপরিবর্তন ঋণাত্মক বা ধনাত্মক। সহপরিবর্তন মাত্রা কমতে থাকলে 'r'-এর মান হবে '0'.

উদাহরণ : নিম্নের রাশিতথ্য থেকে সহ পরিবর্তন গুণাঙ্ক নির্ণয় করুন।

x	1	2	3	4	5	6	7
y	6	8	11	9	12	19	14

x	y	x^2	y^2	xy
1	6	1	36	6
2	8	4	64	16
3	11	9	121	33
4	9	16	81	36
5	12	25	144	60
6	10	36	100	60
7	14	49	196	98
$\sum x = 28$	$\sum y = 70$	$\sum x^2 = 140$	$\sum y^2 = 742$	$\sum yx = 309$

$$\therefore \qquad r = \frac{\sum xy - \frac{\sum x \sum y}{N}}{\sqrt{\sum x^2 = \frac{\left(\sum x\right)^2}{N}}\ \sqrt{\sum y^2 = \frac{\left(\sum y\right)^2}{N}}}$$

$$r = \frac{309 - \frac{28 \times 70}{7}}{\sqrt{140 - \frac{784}{N}} \times \sqrt{742 - \frac{4900}{7}}}$$

$$r = \frac{309 - 280}{\sqrt{28} \times \sqrt{42}} = \frac{29}{5.29 \times 6.48}$$

$$= + .84622.$$

r-এর মান থেকে বোঝা যাচ্ছে যে তাদের মধ্যে ধনাত্মক সম্পর্ক (positive relation) বর্তমান।

দ্বিতীয় পদ্ধতি :

x	y	$x = x - \bar{x}$	$y = y - \bar{y}$	x^2	y^2	xy
1	6	−3	−4	9	16	12
2	8	−2	−2	4	4	4
3	11	−2	1	1	1	1
4	9	0	−1	0	1	0
5	12	1	2	1	4	2
6	10	2	0	4	0	0
7	14	3	4	9	16	12
$\sum x = 28$	70	0	0	$\sum x^2 = 28$	$\sum y^2 = 42$	$\sum xy = 29$

$$\bar{x} = \frac{\sum x}{n} = \frac{28}{7} = 4. \qquad \bar{y} = \frac{\sum y}{n} = \frac{70}{7} = 10$$

$$r = \frac{29}{\sqrt{28} \times \sqrt{42}} = + .84622$$

6.3 প্রতিগমন (Regression)

দুটি চলকের মধ্যে গড় সম্বন্ধ যার দ্বারা বোঝানো হয় তাকে প্রতিগমান (Regression) বলা হয়। এম্ এম্ ব্লেয়ারের কথার "প্রতিগমন হল দুই বা ততোধিক চলকের মধ্যে গড় সম্বন্ধের রাশি তথ্যগুলির প্রাথমিক এককের মাধ্যমে প্রকাশিত পরিমা" [Regression in the measure of the average relationship between two or more variables in terms of the original units of the data.]

আমরা বিক্ষিপ্ত চিত্রে দেখেছি যে রাশিতথ্যের দুটি ভিন্ন তথ্যের মধ্যে সম্পর্ক স্থাপন করা হয়। এতে 'x' অক্ষে স্বাধীন চল বা independent variable-কে নেওয়া হয় বা y অক্ষে dependent বা অধীন চলকে নেওয়া হয়। এক : x এবং y-এর পরিপ্রেক্ষিতে যে বিন্দুগুলি চিহ্নিত হয় তাদের একত্রে scatter diagram বা বিক্ষিপ্ত চিত্র বলা হয়। এই সম্পর্ক ঋণাত্মক বা ধনাত্মক হতে বা চলক দুটির মধ্যে কোন সম্পর্ক নাও থাকতে পারে। যদি ধারণা করা যায় যে সম্পর্কটি একটি সরল রেখার মাধ্যমে প্রকাশ করা যায়, তাহলে এই সরল রেখার উপর y বিন্দু অবস্থিত হলে প্রতিটি x বিন্দুর জন্য আদর্শ 'y' বিন্দুর মান নির্ণয় করতে হবে। দুটি চলকের মধ্যে কার্যকারণ সম্পর্ক সুস্পষ্টভাবে প্রকাশ করতে গেলে আমরা গাণিতিক পদ্ধতি (mathematical form) প্রকাশ করলে তার প্রয়োজনীয়তা বৃদ্ধি পায়। একটি সরলরেখার যে সমীকরণ পাওয়া যায় তা হল $y = a + bx$। এভাবে আদর্শ 'y' নির্ণয় করে যে রৈখিক সম্পর্ক চিহ্নিত হয় তাকে, Linear Regression বলা হয়।

6.3.1 প্রতিগমন সমীকরণ (Regression equation) :

এখানে একটি চলকের নির্ধারিত মানগুলির জন্য অন্য চলকটির সর্বাপেক্ষা সম্ভাব্য মানগুলির পূর্বাভাস প্রদান (বা প্রাক্ কলনী হিসাবকরণ) ঐ দুটি চলকের উপযুক্ত সমীকরণের সাহায্যে করা হয়। এই সমীকরণকে প্রতিগমন সমীকরণ বলে পরিচিত।

x-এর উপর y-এর প্রতিগমন সমীকরণ (Regression equation of y or x)

রৈখিক প্রতিগমনের (Linear Regression) ক্ষেত্রে রাশিতথ্য সমূহের সঙ্গে যদি $y = a + bx$ ধরনের কোন সরললেখাকে 'বর্গসমষ্টির' ক্ষুদ্রতমকরণ (least square method) বা লঘিষ্ঠ বর্গসমষ্টি) পদ্ধতিতে খাপ খাওয়ানো যায়, বা ফিট্ করা যায় অর্থাৎ a এবং b-এর মান নির্ণয় করা যায়। তবে x-এ উপর y-এর প্রতিগমন সমীকরণ পাওয়া যায়। ধরা যাক্, (x_1y_1) (x_2y_2) ... (x_ny_n) হল n জোড়া পর্যবেক্ষণলব্ধ মান এবং এই রাশিতথ্যসমূহের সঙ্গে যে সরলরেখাটি ফিট্ করা হবে, সেটি হল

$$y = a + bx \qquad ..... \quad (1)$$

এখন, লঘিষ্ঠ বর্গসমষ্টি পদ্ধতিতে নীচে উল্লেখিত স্বাভাবিক সমীকরণ (normal equation) পাওয়া যায়।

$$\sum y = na + b\sum x \qquad ..... \quad (2)$$

এবং $$\sum xy = a\sum x + b\sum x^2 \qquad ..... \quad (3)$$