7.(iv) শতাংশ বিন্দু এবং শতাংশ র‍্যাঙ্ক (Percentiles and Percentile Ranks)

আলোচ্য বিষয়বস্তু ঃ শতাংশ বিন্দু ● শতাংশ র‍্যাঙ্ক

❑ শতাংশ বিন্দু (Percentiles) ঃ

পরিমাপক স্কেলের যে বিন্দুর নীচে মোট পরিসংখ্যার (N) নির্দিষ্ট শতাংশ স্কোর থাকে তাকে শতাংশ বিন্দু (Percentile point or simply percentile) বলে (A percentile point is a score point below which a specified percent of the scores in the distribution fall.)।

যেমন কোনো পরিসংখ্যা বিভাজনের মধ্যমমান (Median) বলতে সেই বিন্দুকে বুঝি যার নীচে 50% স্কোর রয়েছে। একে Q_2 বা P_{50} দ্বারা চিহ্নিত করা হয়। আবার Q_1 বা P_{25} এবং Q_3 বা P_{75} বলতে পরিসংখ্যা ক্ষেত্রের সেই বিন্দুকে বুঝি যার নীচে 25% বা 75% স্কোর রয়েছে। অনুরূপভাবে P_{10}, P_{78}, P_{40} ইত্যাদি বলতে পরিমাপক স্কেলের সেই বিন্দুকে বুঝি যার নীচে 10% 78%, 40% স্কোর রয়েছে।

Median (মধ্যমমান) নির্ণেয় সূত্রটি হল ঃ

$$Q_2 \text{ বা } P_{50} = L + \frac{\frac{N}{2} - F}{fm} \times i$$

L = যে শ্রেণিতে মধ্যমমান আছে তার নিম্নসীমা।

$\frac{N}{2}$ = মোট পরিসংখ্যার অর্ধেক।

F = যে শ্রেণিতে মধ্যমমান আছে তার নীচের শ্রেণি পর্যন্ত পরিসংখ্যাগুলির যোগফল।

fm = যে শ্রেণিতে মধ্যমমান আছে সেই শ্রেণির পরিসংখ্যা।

i = শ্রেণির প্রসার।

একইভাবে,

$$Q_1 \text{ বা } P_{25} = L_1 + \frac{\frac{N}{4} - F}{f_p}$$

যেখানে L_1= যে শ্রেণিতে P_{25} রয়েছে তার নিম্নসীমা।

$\frac{N}{4}$ = মোট পরিসংখ্যার $\frac{1}{4}$

F = যে শ্রেণিতে $\frac{N}{4}$ আছে তার নীচের শ্রেণি পর্যন্ত পরিসংখ্যাগুলির যোগফল।

f_p যে শ্রেণিতে P_{25} আছে তার পরিসংখ্যা।

i = শ্রেণি প্রসার।

সুতরাং মধ্যমমান P_{50} বা P_{25} প্রভৃতি সূত্রের মতো যে সাধারণ সূত্র ব্যবহার করে যে কোনো শতাংশ বিন্দু (percentile) নির্ণয় করতে পারি সেটি হল :

$$P_p = L + \frac{pN - F}{f_p}$$

P_p = নির্ণেয় শতাংশ বিন্দু (যেমন 10%, 33% ইত্যাদি)।

L = যে শ্রেণিতে শতাংশ বিন্দু (P_p) আছে তার নিম্নসীমা (exact lower limit of the distribution upon which P_p lies)।

pN = যে শতাংশ বিন্দু নির্ণয় করতে হবে তা N এর শতকরা কত ভাগ (part of N to be counted off in order to reach P_p)।

F = L যে শ্রেণিতে রয়েছে তার নীচের স্কোরসমূহের যোগফল (Sum of all scores upon intervals below (L)।

f_p = যে শ্রেণিতে শতাংশ বিন্দু আছে তার পরিসংখ্যা (frequency within the interval upon which P_F falls)।

i = শ্রেণি প্রসার (length of class interval) P_0 হল প্রথম শ্রেণি প্রসারের নিম্নসীমা (exact lower limit of 1st interval) এবং P_{100} হল শেষ শ্রেণি প্রসারের উর্ধ্বসীমা (exact upper limit of last interval) এই মান দুটি পারসেন্টাইল স্কোরের সীমাবদ্ধতা নির্দেশ করে।

শতাংশগুলিকে অন্য সংকেতের মাধ্যমে লেখা যায় :

$P_{10} = D_1$ = প্রথম দশাংশ; $P_{20} = D_2$ = দ্বিতীয় দশাংশ; $P_{25} = Q_1$ = প্রথম চতুর্থাংশ; $P_{30} = D_3$ = তৃতীয় দশাংশ; $P_{40} = D_4$ = চতুর্থ দশাংশ; $P_{50} = Q_2$ = দ্বিতীয় চতুর্থাংশ; $P_{75} = Q_3$ = তৃতীয় চতুর্থাংশ; $P_{100} = D_{10}$ = দশম দশাংশ;

❒ শতাংশ র‍্যাঙ্ক (Percentile Rank) :

বণ্টনের কোনো স্কোরকে শতকরা হারে যে অবস্থানে স্থাপন করা যায়, তাকে ওই স্কোরের শতাংশ র‍্যাঙ্ক (Percentile Rank) বলে। (Percentile Rank is the position which can be attributed to a score of a distribution in a scale of 100).

নির্দিষ্ট কোনো শতাংশ বিন্দু থেকে স্কোরের শতাংশ র‍্যাঙ্ক (P.R.) নির্ণয় করা যায়। যদি 30 তম শতাংশ বিন্দু P_{30}=45 হয় তবে 45 স্কোরের শতাংশ র‍্যাঙ্ক হল =30। একটি

বণ্টনের নিম্ন ও ঊর্ধ্বসীমা যদি যথাক্রমে 15 ও 150 হয় তাহলে শতাংশ বিন্দু (Percentile Point বা P_p) এর মান 15 থেকে 150 এর মধ্যে হবে কিন্তু শতাংশ ক্রম (P.R.) এর মান কখনও 100-এর বেশি হবে না। সুতরাং শতাংশ র‍্যাঙ্ক দ্বারা আমরা যে কোনো বণ্টনের নির্দিষ্ট স্কোরের অবস্থান নির্ণয় করতে পারি।

যে সাধারণ সূত্র ব্যবহার করে (P.R.) নির্ণয় করা হয় সেই সূত্রটি হল :

$$P.R. = \left\{F + fp \times \frac{S-L}{i}\right\} \times \frac{100}{N}$$

যেখানে P.R.= নির্ণয় শতাংশ র‍্যাঙ্ক (P.R. of the distribution wanted) F = যে শ্রেণিতে শতাংশ র‍্যাঙ্ক আছে তার নীচের স্কোরসমূহের যোগফল (Sum of all scores below the classes in which the P.R. is present.)।

L = যে শ্রেণিতে শতাংশ র‍্যাঙ্ক আছে তার নিম্নসীমা (exact lower limit of the class interval upon which P.R. lies)।

S = যে স্কোরের শতাংশ র‍্যাঙ্ক নির্ণয় করতে হবে (The score whose P.R. is to be determind) fp = যে শ্রেণিতে শতাংশ র‍্যাঙ্ক আছে তার পরিসংখ্যা

i = শ্রেণি প্রসার (Class interval)।

N = মোট স্কোর সংখ্যা (Total No. of scores)।

আবার অনেক ক্ষেত্রে দেখা যায় বিভিন্ন ব্যক্তি বা বস্তুকে তাদের কোনো বৈশিষ্ট্য অনুযায়ী 1, 2, 3 ইত্যাদি ক্রমকে (order of merit) P. R. দ্বারা প্রকাশ করা যায় নিম্নোক্ত সূত্রের সাহায্যে।

সূত্রটি হল :

$$P.R. = 100 - \left(\frac{100R - 50}{N}\right)$$

যেখানে R = র‍্যাঙ্ক অবস্থান (Rank position)

N = মোট স্কোর সংখ্যা (Total number of score)

Ex : 1 Determine P_{40}, P_{75} and PR of scores 25 and 37 from the following distribution.

Scores	18-20	21-23	24-26	27-29	30-32	33-35	36-38	39-41
freq.	3	4	5	10	10	8	7	3

Scores	Class boundary	f	Cmf
18-20	17.5-20.5	3	3
21-23	20.5-23.5	4	7
24-26	23.5-26.5	5	12

Scores	Class boundary	f	Cmf
27-29	26.5-29.5	10	22
30-32	29.5-32.5	10	32
33-35	32.5-35.5	8	40
36-38	35.5-38.5	7	47
39-41	38.5-41.5	3	50
		N = 50	

Ans.

$$P_{40} = L_1 + \frac{\text{N এর } 40\% - F_1}{f_1} \times i$$

$$\text{N এর } 40\% = 50 \times \frac{40}{100} = 20$$

L_1 = যে শ্রেণিতে N এর 40% = 50 এর $\frac{40}{100} = 20$ Cmf আছে তার নিম্নসীমা = 26·5

F = যে শ্রেণিতে 20 Cmf আছে সেই শ্রেণির আগের পর্যন্ত শ্রেণির Cmf.

f_1 = যে শ্রেণিতে 20 Cmf আছে তার ফ্রিকোয়েন্সি ; i = 20.5 − 17.5 = 3.0

$$\therefore P_{40} = 26 \cdot 5 + \frac{20-12}{10} \times 3$$

$$= 26 \cdot 5 + \frac{24}{10} = 26 \cdot 5 + 2 \cdot 4 = 28 \cdot 9$$

$$P_{75} = Q_3 = L_3 + \frac{N \quad 75\% - F_3}{f_3} i$$

$$\text{N এর } 75\% = 50 \text{ এর } 75\% = 50 \times \frac{3}{4} = \frac{150}{4} = 37 \cdot 5$$

L_3 = যে শ্রেণিতে 37·5 Cmf রয়েছে তার নিম্নসীমা = 32·5

F_3 = যে শ্রেণিতে 37·5 Cmf রয়েছে তার আগে পর্যন্ত ফ্রিকোয়েন্সি সমূহের যোগফল = 32

f_3 = যে শ্রেণিতে 37·5 Cmf রয়েছে তার ফ্রিকোয়েন্সি = 32·5 ; i = 3

$$\therefore P_{75} = 32 \cdot 5 + \frac{37 \cdot 5 - 32}{8} \times 3$$

$$= 32 \cdot 5 + \frac{5 \cdot 5 \times 3}{8} = 32 \cdot 5 + 2 \cdot 06 = 34 \cdot 56$$

PR of 25

$$P.R = \left\{F + f_p \times \frac{S-L}{i}\right\} \times \frac{100}{N}$$

F = যে শ্রেণিতে শতাং র‍্যাঙ্ক আছে তার নীচের শ্রেণি পর্যন্ত Cmf = 7

L = যে শ্রেণিতে শতাংশ র‍্যাঙ্ক আছে তার নিম্নসীমা = 23·5

S = শতাংশ র‍্যাঙ্কের স্কোর = 25

i = 3 ; N = 50

f_p = যে শ্রেণিতে শতাংশ র‍্যাঙ্ক আছে তার ফ্রিকোয়েন্সি = 5

$$\text{P.R.} = \left\{7 + 5 \times \frac{25 - 23\cdot5}{3}\right\} \times \frac{100}{50}$$

$$= \left\{7 + 5 \times \frac{1\cdot5}{3}\right\}$$

$$= \{7 + 5 \times 0\cdot5\} \times 2 = 19$$

P.R. of 37

$$\text{P.R.} = \left\{40 + 7 \times \frac{37 - 35\cdot5}{3}\right\} \frac{100}{50}$$

$$= \{40 + 3\cdot5\} \times 2$$

$$= 43\cdot5 \times 2$$

$$= 87$$

Ex. 2 Calculate directly the percentile rank of a person who scores 65 in the following distribution.

Scores	90-99	80-89	70-79	60-69	50-59	40-49	30-39
freq.	2	12	22	20	14	4	1

Ans :-

Scores	Classes inverval scores	f	cmf
30-39	29.5-39.5	1	1
40-49	39.5-49.5	4	5
50-59	49.5-59.5	14	19
60-69	59.5-69.5	20	39
70-79	69.5-79.5	22	61
80-89	79.5-89.5	12	73
90-99	89.5-99.5	2	75

$$\text{P.R.} = \left\{F + fp \times \frac{S-L}{i}\right\} \times \frac{100}{N}$$

F = 19, fp 20, S = 65, i = 10, N = 75, L = 59.5

$$\text{P.R} = \left\{19 + 20 \times \frac{65 - 59\cdot5}{10}\right\} \frac{100}{75}$$

$$= \{19 + 2 \times 5{\cdot}5\} \times \frac{100}{75}$$

$$= (19 + 11) \times \frac{100}{75} = 30 \times \frac{100}{75} = 40$$

$\therefore$ P.R. = 40

Ex. 3 : On an achievement test on Phy. Sc. on 25 students of class IX, they are arranged in order (rank) of their performances. Niladri scores 6th position; what will be his percentile Rank?

Ans :

$$\text{P.R} = 100 - \left(\frac{100R - 50}{N}\right)$$

$$\text{P.R of Niladri} = 100 - \left(\frac{100 \times 6 - 50}{N}\right) = 100 - \frac{550}{25}$$

$$= 100 - 22 = 78$$

Ex. 4 : একটি পারদর্শিতার পরীক্ষায় 20 জন শিক্ষার্থীর স্কোরগুলি হল—

12, 20, 25, 15, 8, 32, 28, 35, 22, 44,

36, 17, 29, 13, 9, 37, 40, 21, 10, 42

17 স্কোরের Percentile Rank বের করো।

Ans : স্কোরগুলিকে নিম্নক্রম অনুযায়ী সাজানো হল। 17 স্কোরের Rank order = 14

Scores	Rank order	Scores	Rank order
44	1	22	11
42	2	21	12
40	3	20	13
37	4	17	14
36	5	15	15
35	6	13	16
32	7	12	17
29	8	10	18
28	9	9	19
25	10	8	20

$$\therefore \text{P.R} = 100 - \frac{100 \times 14 - 50}{20}$$

$$= 100 - \frac{1350}{20} = 100 - 67{\cdot}5 = 32{\cdot}5$$

$\therefore$ 17 স্কোরটির PR = 32 (পূর্ণসংখ্যা ধরে)

7.(vi) সহ সম্বন্ধ

(Corelation & Co-efficient of Correlation)

আলোচ্য বিষয়বস্তু ঃ সহ সম্বন্ধ—ধনাত্মক সহগতি—ঋণাত্মক সহগতি—শূন্য সহগতি ● প্রোডাক্ট মোমেন্ট পদ্ধতি—অবিন্যস্ত স্কোরের সহগাঙ্ক—সহগতি সহগাঙ্ক 'r' এর তাৎপর্য—অসুবিধা ● র‍্যাঙ্ক পার্থক্য পদ্ধতি—সুবিধা—ত্রুটি প্রোডাক্ট মোমেন্ট সহগতি ও র‍্যাঙ্ক পার্থক্য সহগতি ● শিক্ষাক্ষেত্রে সহগতি সহগাঙ্কের ব্যবহার—উদাহরণ।

❑ সহ সম্বন্ধ (Corelation & Co-efficient of Correlation) ঃ

সহ সম্বন্ধ হল কোন দুই চলের (Variable) মধ্যে এমন একটি সম্বন্ধ যার মাধ্যমে চল দুটির পারস্পরিক নির্ভরতার মাত্রা নির্ণয় করা যেতে পারে। Guilford-এর মতে 'যখন দুটি শ্রেণি বা ঘটনার মধ্যে একটি পরিমানগত বা গুণগত সম্বন্ধ থাকে, তখন সেই সম্বন্ধ নির্ধারণ, পরিমাপ ও সঠিকভাবে প্রকাশ করার উপযুক্ত পরিসংখ্যানমূলক সূত্র হল সহ-সম্বন্ধ।

(When there is a relationship of a quantitative or qualitative nature between two sets of phenomena, the appropriate statistical test for discovering and measuring the relationship and expresing it in a precise way is known as correlation).

যদি দুটি চলরাশি (Variable) পরস্পরের সাথে এমন ভাবে সম্পর্কিত হয় যে একটির পরিবর্তন ঘটলে অপরটিও পরিবর্তিত হয় তখন বলা হয় দুটি রাশির মধ্যে সহগতি রয়েছে। যেমন ঃ- একটি বৃত্তের ব্যাসের কমা বাড়ার সঙ্গে সঙ্গে তার পরিধি কমে বাড়ে। অর্থাৎ ব্যাস ও পরিধি পরস্পরের সঙ্গে সম্পর্কিত। শিক্ষাক্ষেত্রে বুদ্ধি ও গণিতে পারদর্শিতা পরস্পরের মধ্যে সম্পর্কিত।

যে গাণিতিক সূচক দ্বারা দুটি চলের মধ্যকার সহগতির পরিমাপ ও গুণগত বৈশিষ্ট্য প্রকাশ করা হয় তাকেই বলে সহগতির সহগাঙ্ক (Co-efficient of Correlation) সহগাঙ্ককে সাধারণত r দ্বারা প্রকাশ করা হয়। r এর মান –1 থেকে +1 পর্যন্ত হতে পারে।

Karl Pearson দুটি চলের মধ্যে সহসম্বন্ধ গুণাঙ্ক নির্ণয়ের পদ্ধতি প্রবর্তন করেন। পিয়ারসন পদ্ধতিতে যখন সহসম্বন্ধ নির্ণয় করা হয় তখন সর্বদা ধরে নেওয়া হয় ঐ দুটি চলের মধ্যে সরলরৈখিক সম্পর্ক আছে। এই সম্বন্ধকে Product moment coefficient বলে। এর মান '–'1 থেকে '+'1 পর্যন্ত বিস্তৃত হতে পারে। যখন দুটি চলের মধ্যে একটির মানের বৃদ্ধি বা হ্রাস অপরটির মান যথাক্রমে সমপরিমাণ বৃদ্ধি বা হ্রাস পায় অর্থাৎ চলরাশিদ্বয়ের মানের পরিবর্তন সমমুখী হয় তখন সহগতি গুণাঙ্ককে +1 হয়। যেমন একটি শ্রেণির ছাত্রদের I.Q ও পরীক্ষার ফলাফল একই হারে পরিবর্তিত হয়

ন তাদের মধ্যে সহসম্বন্ধ গুণাঙ্ক হল +1। আবার যদি একটি চলের মানের হ্রাস বা তে অপরটির মান যথাক্রমে সম পরিমাণ বৃদ্ধি বা হ্রাস পায় অর্থাৎ চলরাশিদ্বয়ের নর পরিবর্তন বিপরীতমুখী হয় তখন সহসম্বন্ধ গুণাঙ্ক –1 ধরা হয়। ধর, যেমন নো একটি শ্রেণিতে ছাত্ররা বাংলাতে ভাল নম্বর পেয়েছে এবং আনুপাতিক হারে খারাপ নম্বর পেয়েছে। এক্ষেত্রে উভয় বিষয়ের মধ্যে সহসম্বন্ধ গুণাঙ্ক হল –1.

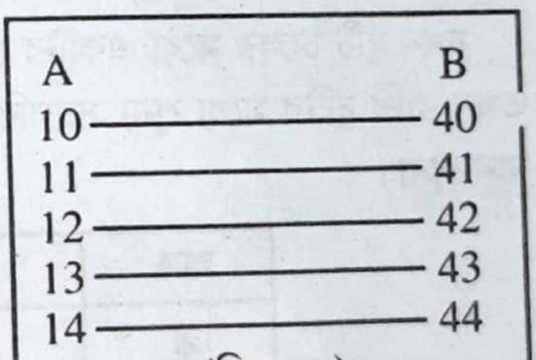
আবার দুটি চলরাশির মধ্যে এমন সম্পর্ক থাকে যে একটির মানের পরিবর্তনের র অপরটির মান নির্ভরশীল নয়। তাহলে সেরূপ ক্ষেত্রে রাশি দুটি পরস্পর সম্বন্ধহীন ncorrelated) হবে। এক্ষেত্রে সহসম্বন্ধ গুণাঙ্ক হল 0. যেমন ছাত্রদের উচ্চতার সঙ্গে দের অঙ্কের ফলাফলের কোনো সম্পর্ক নেই।

তাই সহসম্বন্ধের গুণাঙ্কের মান –1 থেকে +1 পর্যন্ত্য যে কোন মান হতে পারে।

সহগতি তিন ধরনের হত পারে—(1) ধনাত্মক সহগতি (Positive Correlation)) ঋণাত্মক সহগতি (Negative Correlation) (3) শূন্য সহগতি (Zero Correlation)

) ধনাত্মক সহগতি ঃ

যদি দুটি চলরাশির একটি বৃদ্ধি পেলে অপরটি বৃদ্ধি পায় আবার একটি হ্রাস পেলে পরটিও হ্রাস পায় তখন চলরাশি দুটির মধ্যে ধনাত্মক সহগতি আছে বলা হয়। যখন টি চলের মধ্যে পরিপূর্ণ ধনাত্মক সহগতি থাকে র্থাৎ একটি চল রাশির যে কোনো পরিমাণ রিবর্তনে অপরটির সমপ্রকৃতির সমপরিমাণ রিবর্তন হয় তখন চলরাশি দুটির মধ্যে সহগতির হগাঙ্ক (r) হয় +1 ধরা যাক কোনো বিশেষ রিস্থিতিতে কোনো একটি শ্রেণির ছাত্রদের বুদ্ধ্যাঙ্ক হারে বাড়ছে তাদের প্রাপ্ত পরীক্ষার নম্বরও সেই রে বাড়ছে এক্ষেত্রে বুদ্ধ্যাঙ্ক ও পরীক্ষার ফলাফলের ধ্যে সহগতির সহগাঙ্ক +1. তাপমাত্রা ও থার্মোমিটারে পারদ স্তম্ভের উচ্চতার মধ্যে হগতির সহগাঙ্ক +1.



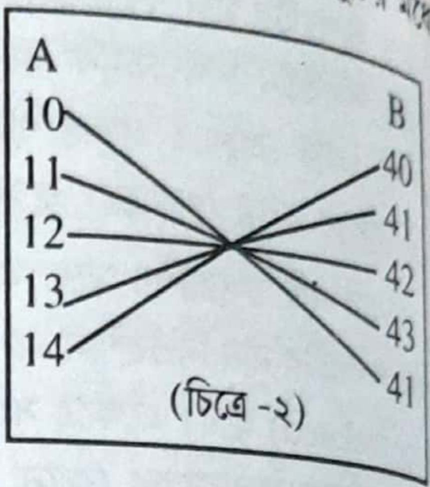
(চিত্রে -১)

চিত্রে -১ এতে A স্তম্ভের ও B স্তম্ভের একদল ছাত্রের দুটি বিষয়ে নম্বর আছে। চিত্রে খা যাচ্ছে যারা A স্তম্ভে যারা বেশি পেয়েছে তারা B তেও বেশি পেয়েছে। এক্ষেত্রে পরিবর্তনের সঙ্গে B এর সমপরিমাণ পরিবর্তন হয়েছে। অর্থাৎ এক্ষেত্রে $r = +1$

) ঋণাত্মক সহগতি ঃ

যখন দুটি চলরাশির মধ্যে বিপরীত সম্পর্ক থাকে অর্থাৎ একটি চলরাশির বৃদ্ধিতে পরটি হ্রাস পায় তখন রাশি দুটির মধ্যে ঋণাত্মক সহগতি আছে বলা হয়।

যখন দুটি চলের মধ্যে সম্পূর্ণ বিপরীত সম্পর্ক থাকে তখন চলরাশির দুটির মধ্যে সহগতির সহগাঙ্ক r এর মান –1 হয়।



(চিত্রে -২)

যেমন ঃ কোনো বিশেষ পরিস্থিতিতে দেখা গেল, যে সব ছাত্ররা ইংরাজিতে ভালো নম্বর পেয়েছে তারা আনুপাতিক হারে অঙ্কে খারাপ নম্বর পেয়েছে বা যারা অঙ্কে ভালো নম্বর পেয়েছে, আনুপাতিক হারে তারা ইংরাজিতে খারাপ নম্বর পেয়েছে তখন ইংরাজি ও অঙ্কের মধ্যে সহগতির সহগাঙ্ক –1.

চিত্র ২নং দেখা যাচ্ছে যারা A তে বেশি নম্বর পেয়েছে তারা B তে কম নম্বর ও যারা A তে কম নম্বর পেয়েছে তারা B তে বেশি নম্বর পেয়েছে। অর্থাৎ এক্ষেত্রে A এর পরিবর্তনের সঙ্গে B এর পরিবর্তন সম্পূর্ণ বিপরীত। তাই এক্ষেত্রে $r = -1$

সুতরাং পূর্ণ ধনাত্মক সহগতির সহগাঙ্ক (r) এর মান +1 এবং পূর্ণ ঋণাত্মক সহগতির সহগাঙ্ক (r) এর মান –1. এই দুই চরম প্রান্তের মধ্যে (+1 ও –1) এর মধ্যে r এর মান বর্তমান থাকে।

○ শূন্য সহগতি ঃ

যখন দুটি চলের মধ্যে একটির কোনো পরিবর্তন অপরকে প্রভাবিত করতে পারেনা তখন রাশি দুটির মধ্যে শূন্য সহগতি আছে বলা হয়। এক্ষেত্রে সহগতির সহগাঙ্ক r এর মান শূন্য।

ছাত্র	বয়স	উপস্থিতির দিন
ক	16	150
খ	15	160
গ	14	135
ঘ	13	170
ঙ	12	145

(চিত্রে -৩)

চিত্র ৩-এ দেখা যাচ্ছে ছাত্রদের বয়সের সঙ্গে উপস্থিতি দিনের কোনো সম্পর্ক নেই। এক্ষেত্রে সহগতির সহগাঙ্ক প্রায় 0 এর কাছাকাছি।

ছাত্রের উচ্চতার সঙ্গে অঙ্কে নম্বরের কোন সম্পর্ক নেই। তাই এক্ষেত্রে সহগাঙ্ক $r = 0$।

সহসম্বন্ধ রৈখিক (linear) বা অরৈখিক (non linear) হতে পারে। দুটি চলের মানের পরিবর্তনের অনুপাত ধ্রুবক হলে তাদের সহসম্বন্ধ রৈখিক কিন্তু চল দুটির মানের অনুপাত ধ্রুবক না হলে তাদের সহ সম্বন্ধের মানকে অরৈখিক বলে। যেমন মানুষের বয়স ও শারিরীক শক্তির মধ্যে সহ সম্বন্ধ অরৈখিক।

সহগতির সহগাঙ্ক নির্ণয়ের বিভিন্ন পদ্ধতি আছে। এর মধ্যে রৈখিক সম্পর্ক যুক্ত দুটি পদ্ধতি রয়েছে। যেমন—(1) প্রোডাক্ট মোমেন্ট পদ্ধতি (Product Moment Method) এবং (2) র‍্যাঙ্ক পার্থক্য পদ্ধতি (Rank Difference Method)।

● **(1) প্রোডাক্ট মোমেন্ট পদ্ধতি (Product Moment Method) ঃ**

রৈখিক সম্পর্ক যুক্ত দুটি চলের মধ্যে সহগতির সহগাঙ্ক নির্ণয়ের জন্য যে পদ্ধতি রয়েছে তাকে প্রোডাক্ট মোমেন্ট পদ্ধতি বলে। এই পদ্ধতিতে সহগতির সহগাঙ্ককে 'r' দিয়ে চিহ্নিত করা হয়।

পিয়ারসনের মতে ঃ সহসম্বন্ধের গুণফল গুণাংক (Co-effrent currelation) হল এমন একটি অনুপাত যা একটি চলের পরিবর্তনের সাথে অপর চলের যে পরিবর্তন হচ্ছে তার সীমা প্রকাশ করে (Product noment coefficient of correlation is a kind of ratio that express the extent to which change in the variable is accompanied by change in another variable). এই পদ্ধতিতে মান নির্ণয়ের বিভিন্ন পদ্ধতি আছে তার মধ্যে কয়েকটি হল ঃ

অবিন্যস্ত স্কোরের সহগাঙ্ক ঃ X ও Y যদি দুটি চল শ্রেণি হয় এবং M_1 ও M_2 যদি ওই দুই শ্রেণির গড় হয়। এবং শ্রেণি দুটিতে স্কোর সংখ্যা যদি N হয় তবে x হল X শ্রেণির প্রতিটি রাশিকে M_1 দ্বারা বিয়োগ করে যে চ্যুতি পাওয়া যায় এবং অনুরূপে y হল Y শ্রেণির প্রতি রাশিকে M_2 দ্বারা বিয়োগ করে যে চ্যুতি পাওয়া যায় চ্যুতি দ্বয়ের গুণফল xy। এইভাবে সব xy যোগ করে পাওয়া যায় Σxy এরপর X শ্রেণির চল রাশির S.D হল σ_x ও Y শ্রেণির চলরাশির S.D হয় σ_y প্রোডাক্ট মোমেন্ট পদ্ধতি অনুযায়ী সহগতির সহগাঙ্ক

$$(r) = \frac{\Sigma xy}{N\sigma_x\sigma_y} \quad [x = X - M_1;\ y = Y - M_2]$$

X = X চল শ্রেণির যে কোনো স্কোর

Y = Y চল শ্রেণির যে কোনো স্কোর

প্রত্যক্ষভাবে স্কোর থেকে ঃ

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2} \times \sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

যেখানে N = স্কোর সংখ্যা

ΣXY = X শ্রেণির স্কোর ও Y শ্রেণির স্কোরগুলির গুণফল এর যোগফল

ΣX = X শ্রেণির স্কোর সমূহের যোগফল ও ΣY = Y শ্রেণির স্কোর সমূহের যোগফল

ΣX^2 ও ΣY^2 হল যথাক্রমে X ও Y শ্রেণির স্কোরগুলির বর্গের সমষ্টি

কল্পিত গড় পদ্ধতিতে সহগাঙ্ক নির্ণয় ঃ

$$r_{xy} = \frac{\frac{\sum x'y'}{N} - c_x c_y}{\sigma'x\,\sigma'y}$$ [দুটি পরিসংখ্যা বিভাজনের স্কোর (X ও Y) সমূহের থেকে কল্পিত গড়ের বিয়োগ ফল যথাক্রমে x′, y′

$$\sigma'_x = \sqrt{\frac{\sum fx'^2}{N} - \left(\frac{\sum fx'}{N}\right)^2}$$ [σ′x , σ′y হলে ওই দুটি পরিসংখ্যা বিভাজনের SD.]

$$\sigma'_y = \sqrt{\frac{\sum fy'^2}{N} - \left(\frac{\sum fy'}{N}\right)^2} \quad \left[c_x = \frac{\sum fx'}{N}, c_y = \frac{\sum fy'}{N}\right]$$

N = মোট স্কোর সংখ্যা

সহগতি সহগাঙ্ক ‘r’ এর তাৎপর্য্য ঃ

গিলফোর্ড সহগতি সহগাঙ্ক গুণাঙ্কের মাত্রার উপর ভিত্তি করে সহ সম্বন্ধকে নিম্ন লিখিত শ্রেণিগুলিতে ভাগ করেন ঃ

সহগাঙ্ক (r এর মান)	সম্বন্ধ
(i) .00 থেকে ±0.20	খুবই কম
(ii) ±0.21 থেকে ±0.40	কম
(iii) ±0.41 থেকে ±0.60	সাধারণ
(iv) ±0.61 থেকে ±0.80	অধিক
(v) ±0.81 থেকে ±0.99	খুবই বেশি
(vi) ±1.00	পরিপূর্ণ সহসম্বন্ধ

বিশেষ ক্ষেত্রে কোন্‌টি গ্রহণ করা হবে তা বিচার করা হয় (a) চল গুলির প্রকৃতি (b) যে উদ্দেশ্যে সহগতি নির্ণয় করা হয় তার পরিপ্রেক্ষিতে।

শিক্ষাতত্ত্ব, মনোবিজ্ঞান প্রভৃতি বিষয়গুলির তাৎপর্য্য নির্ণয়ের ক্ষেত্রে উক্ত মানগুলি অত্যন্ত গুরুত্বপূর্ণ।

পিয়ারসনের প্রোডাক্ট মোমেন্ট পদ্ধতির অসুবিধা ঃ

(1) সহগাঙ্ক নির্ণয়ের এই পদ্ধতি দীর্ঘ ও শ্রমসাধ্য।

(2) শিক্ষামূলক পরিমাপের ক্ষেত্রে আমরা যে তথ্য পাই তা সব সময় সাংখ্য মানের নাও হতে পারে সেক্ষেত্রে এই পদ্ধতি অসুবিধাজনক।

(3) শিক্ষার্থীদের যোগ্যতাকে গুণগত দিকে প্রকাশ করলে (প্রথম, দ্বিতীয়, তৃতীয় ইত্যাদি) এই পদ্ধতিতে সহগাঙ্ক নির্ণয় করা যায় না।

(4) 'r' এর মান রৈখিক সম্পর্কের ক্ষেত্রে প্রযোজ্য। রৈখিক সম্পর্ক না থাকলে এই মান সহগতির প্রকৃত মান ব্যক্ত করে না।

(2) র‍্যাঙ্ক পার্থক্য পদ্ধতি (Rank Difference Method) ঃ

যে সব চল রাশির মানসমূহকে সংখ্যা দ্বারা প্রকাশ করা যায় না যেমন বুদ্ধি, সৌন্দর্য্য প্রভৃতি গুণগত বৈশিষ্ট্যসমূহের ক্ষেত্রে দুটি চলের পারস্পরিক সম্পর্কের প্রকৃতি কে 1, 2, 3 ... ইত্যাদি সংখ্যা ব্যবহার করে বিভিন্ন পদসমূহের নির্দিষ্ট বৈশিষ্ট্যের সাপেক্ষে ক্রম প্রকাশ করা হয়। এই ভাবে পদসমূহকে ক্রম অনুযায়ী বিন্যাস করণকে Ranking বলে এবং যে সংখ্যা দিয়ে কোন্ পদের ক্রম স্থির করা হয় তা ঐ পদের সারিতে অবস্থানকে চিহ্নিত করে। এই ভাবে মানুযায়ী দুটি চলের দুটি শ্রেণি তৈরি করা হয়। এর পর চল দুটির র‍্যাঙ্কের পার্থক্য নির্ণয় করাহয়। এই পার্থক্য নির্ণয়কে কাজে লাগিয়ে র‍্যাঙ্ক পার্থক্য পদ্ধতিতে সহগতি সহগাঙ্ক (Rank co-efficent of cerretion by differences network) নির্ণয় করা যায়। একে Rank difference method বলে। Spearman এক্ষেত্রে সহগতিকে ρ দ্বারা চিহ্নিত করেন। এক্ষেত্রে যে সূত্রটি প্রযোজ্য সেটি হল ঃ

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

$D = R_1 \sim R_2$ (দুটি র‍্যাঙ্কের পার্থক্য)

N = মোট স্কোর সংখ্যা।

ρ এর মান –1 থেকে +1 এর মধ্যে থাকে। যখন এর মান হয় +1 তখন দুটি সারিতে ক্রম অনুযায়ী প্রত্যেক পদের Rank একই থাকে এবং –1 হবে বলে ক্রম অনুযায়ী প্রত্যেক পদের অবস্থান একই ক্রমে বিপরীত মুখী হবে।

র‍্যাঙ্ক পার্থক্য পদ্ধতির সুবিধা ঃ

(1) গুণগত ও পরিমানগত উভয় স্কোর এর ক্ষেত্রে সহগতির সহগাঙ্ক নির্ণয়ে ব্যবহার করা যায়।

(2) এই পদ্ধতিতে ρ (সহগতির সহগাঙ্ক) অপেক্ষাকৃত সহজে নির্ণয় করা যায়। এই কারণে মনোবিজ্ঞান, শিক্ষাতত্ত্ব ও সমাজশিক্ষায় এর অধিক ব্যবহার হয়।

(3) যেখানে সঠিক তথ্য জানা থাকে না, তথ্য ক্রম অনুযায়ী সারিবদ্ধ অবস্থায় থাকে সেখানে Rank difference পদ্ধতিতে ρ নির্ণয় করা সুবিধা জনক।

র‍্যাঙ্ক পার্থক্য পদ্ধতির ত্রুটি ঃ

(1) স্কোরের উপর কোনো গুরুত্ব দেওয়া হয় না। শুধুমাত্র স্কোরের অবস্থানের উপর গুরুত্ব দেওয়া হয়।

(2) স্কোরগুলিকে সারিবদ্ধ করার সময় সারির পুনরাবৃত্তি ঘটলে ত্রুটি বেড়ে যায়।

(3) পদ সংখ্যা 30 এর চেয়ে বেশি হলে ρ এর মান নির্ণয় কষ্ট সাধ্য হয়।

প্রোডাক্ট মোমেন্ট সহগতি (r) ও র‍্যাঙ্ক পার্থক্য সহগতি (ρ) ঃ

(1) এই দুটি পদ্ধতিতে সহ সম্বন্ধ দুটি রৈখিক সম্পর্কের ক্ষেত্রে প্রযোজ্য।

(2) যখন স্কোর সংখ্যা কম হয় তখন r ও ρ এর মান প্রায় সমান।

(3) স্কোরের পুনরাবৃত্তি ঘটলে r ও ρ এর মান সমান হবে না।

(4) তবে এই দুটি পদ্ধতিতে নির্ণীত r ও ρ এর মান প্রায় সমান থাকে না কারণ ρ নির্ণয়ে বেশ কিছু ত্রুটি আছে যেমন Score মানের উপর ρ নির্ণয়ের কোনো গুরুত্ব দেওয়া হয় না, কেবল মাত্র স্কোরের অবস্থানের উপর গুরুত্ব আরোপ করা হয়। আর প্রোডাক্ট মোমেন্ট পদ্ধতিতে স্কোরের মান ও অবস্থান উভয়ের ওপর গুরুত্ব দেওয়া হয়। এই কারণে Product mement পদ্ধতিতে নির্ণীত সহগাঙ্ক অপেক্ষাকৃত ত্রুটি কম থাকে।

(5) স্কোরের পুনরাবৃত্তি ঘটলে র‍্যাঙ্ক পদ্ধতিতে নির্ণীত মানে ত্রুটি বাড়ে।

(6) Product mement পদ্ধতিতে সহগাঙ্ক নির্ণয় অপেক্ষা Rank difference পদ্ধতিতে সহগতির সহগাঙ্ক নির্ণয় অপেক্ষাকৃত সহজ তাই শিক্ষাতত্ত্ব, মনোবৈজ্ঞানিক ও সমাজতত্ত্বে এই পদ্ধতি ব্যবহৃত হয়।

(7) র‍্যাঙ্কগুলিকে যদি কোনো আদর্শমানের থেকে চ্যুতি হিসেবে বিবেচনা করে Product moment সহগাঙ্ক নির্ণয় করা যায় তবে সেক্ষেত্রে r ও ρ এর মান সমান হয়।

শিক্ষাক্ষেত্রে সহগতি সহগাঙ্কের ব্যবহার (Use of co-efficient of co-relation in Education) ঃ

সহসম্বন্ধ গুণাঙ্ক মান (r) একটি গাণিতিকসূচক। শিক্ষামূলক পরিমাপের ক্ষেত্রে এর অনেক উপযোগিতা লক্ষ করা যায়। এই গুণাঙ্কের সাহায্যে শিক্ষা বিষয়ক বিভিন্ন প্রকার সিদ্ধান্তও গ্রহণ করা হয়।

(i) এই গুণাঙ্কের ভিত্তিতে শিক্ষা বিষয়ক অভীক্ষার বৈধতা (validity) নির্ণয় করা হয়। এক্ষেত্রে r কে Validity coefficient বলে। অভীক্ষা তৈরি করার সময় যখন Predictive validity বা Concurrent validity নির্ধারণ করা হয়, তখন অভীক্ষার প্রাপ্ত স্কোরের সঙ্গে সমতুল্য অভীক্ষায় প্রাপ্ত স্কোরের r নির্ণয় করে অভীক্ষার বৈধতা স্থির করা হয়।

(ii) শিক্ষা বিষয়ক অভীক্ষাসমূহের নির্ভর যোগ্যতা (reliability) নির্ধারণ করার জন্য সহসম্বন্ধ গুণাঙ্ককে সূচক হিসেবে গণনা করা হয়। এই সূচককে reliability coefficient বলে।

(iii) অভীক্ষার পদগুলির বিশ্লেষণে (item analysis), (পদগুলির কাঠিন্য মান (Difficulty value) ও তাদের পার্থক্য নির্ণায়ক ক্ষমতা নির্ণয়ে (Discriminating power)) সহসম্বন্ধ গুণাঙ্কের প্রয়োজন হয়।

(iv) সহ পরিবর্তনের মান নির্ণয়ের মাধ্যমে ভবিষ্যৎ গণনা করা সম্ভব। যেমন শিশু বুদ্ধির অভীক্ষায় যদি উন্নত ফল দেখায় তবে তা থেকে ভবিষ্যৎ গণনা করা যায় সে বিদ্যালয় পরীক্ষাতেও ভালো ফল করবে।

(v) বিদ্যালয়ের শিক্ষার্থী পরীক্ষায় যে সব নির্বাচনমূলক পদ্ধতি আজকাল ব্যবহৃত হয় সেগুলিতে r এর মান ব্যবহার হয়ে থাকে।

(vi) শিক্ষা বিজ্ঞানের ক্ষেত্রে অনেক সময় বিশেষ মানসিক বৈশিষ্ট্যের মূলে যে উপাদানগুলি আছে তা বিশ্লেষণ করার জন্য factor analysis করা যায়। এই প্রক্রিয়ার প্রাথমিক পর্যায় r এর মান নির্ণয় করা হয়। সুতরাং শিক্ষা বিষয়ক গবেষণার ক্ষেত্রেও সহ সম্বন্ধ গুণাঙ্কের গুরুত্ব যথেষ্ট রয়েছে।

Example 1 :

Use the rank difference method to compute the correlation co-efficient of the following distribution and comment on it.

x	10	12	16	18	22	17	19	20	19	18	15	21
y	12	10	19	15	19	15	17	21	22	13	11	18

Ans :

X	Y	Rx	Ry	D=Rx~Ry	D^2
10	12	12	10	2.0	4.00
12	10	11	12	1.0	1.00
16	19	9	3.5	5.5	30.25
18	15	6.5	7.5	1.0	1.00
22	19	1	3.5	2.5	6.25
17	15	8	7.5	0.5	0.25

Condt...

X	Y	Rx	Ry	D=Rx~Ry	D^2
19	17	4.5	6	1.5	2.25
20	21	3	2	1.0	1.00
19	22	4.5	1	3.5	12.25
18	13	6.5	9	2.5	6.25
15	11	10	11	1.0	1.00
21	18	2	5	3.0	9.00
					$\Sigma D^2 = 74.50$

$$r = 1-\frac{6\Sigma D^2}{N(N^2-1)} = 1-\frac{6\times 74\cdot 5}{12(12^2-1)} \quad \text{[এক্ষেত্রে N=12]}$$

$$= 1-\frac{470}{12\times 143} = 1-0.26 = 0.74$$

এক্ষেত্রে সহগতির সহগাঙ্কের মান + 0.74 এবং এই মান ধনাত্মক এবং উচ্চমান সম্পন্ন তাই X এর মান বৃদ্ধিতে Y এর মানও বৃদ্ধি পাবে। আবার X এর মান হ্রাসে Y এর মানও হ্রাস পাবে।

Example 2 :

Compute the correlation coefficient given :

$\Sigma X=1$^ $\Sigma Y=100$ $\Sigma XY=86.2$ $\Sigma X^2=60$ $\Sigma Y^2=428$ $N=25$

Coefficient of correlation

$$r = \frac{N\Sigma XY-(\Sigma X)(\Sigma Y)}{\sqrt{\left[N\Sigma X^2-(\Sigma X)^2\right]\left[N\Sigma Y^2-(\Sigma Y)^2\right]}}$$

এক্ষেত্রে $\Sigma XY=86.2$, $\Sigma X=12$, $\Sigma Y=100$,
$\Sigma X^2=60$, $\Sigma Y^2=428$, $N=25$

$$r = \frac{25\times 86\cdot 2-12\times 100}{\sqrt{\left[25\times 60-(12)^2\right]\left[25\times 428-(100)^2\right]}}$$

$$\frac{955}{\sqrt{1356\times 700}} = \frac{955}{\sqrt{949200}} = \frac{955}{974\cdot 27} = +0\cdot 98$$

এক্ষেত্রে চলরাশি দুটির মধ্যে উচ্চধনাত্মক সহগতি রয়েছে।

Example 3 :

Find out the Correlation Co-efficient form the data given below. Interpret your finding.

Individuals		A	B	C	D	E	F	G	H	I	J
Scores	X	13	12	10	10	8	6	6	5	3	2
	Y	11	14	11	7	9	11	8	7	6	1

Ans :

Individuals	Scores X	Scores Y	R_X	R_Y	D=Rx~Ry	D^2
A	13	11	1	3	+2	4.0
B	12	14	2	1	+1	1.0
C	10	11	3.5	3	+0.5	0.25
D	10	7	3.5	7.5	+4.0	16.00
E	8	9	5	5	0	0.0
F	6	11	6.5	3	+3.5	12.25
G	6	8	6.5	6	+0.5	0.25
H	5	7	8	7.5	+0.5	0.25
I	3	6	9	9	0.0	0.00
J	2	1	10	10	0.0	0.00
						$\Sigma D^2 = 34.00$

এখানে N = 10

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}$$

$$\therefore \rho = 1 - \frac{6 \times 34}{10(100-1)} = 1 - \frac{204}{10 \times 99} = 1 - 0.206 = 0.794$$

Interpretations :

X ও Y স্কোরের মধ্যে ধনাত্মক সহগতি এবং উচ্চমানের Score X এর মান বৃদ্ধি পেলে Y এর মান বৃদ্ধি পাবে, আবার X এর মান হ্রাস পেলে Y এর মানও হ্রাস পাবে।

Example 4 :

Find out the Correlation of Coefficient from the data given below and interpret your finding :

Scores on two	X	2	3	4	5	6	7
Measures of memory span	Y	0	3	4	4	6	11

Scanned by CamScanner

Ans :

X-Score	Y-Score	Xস্কোরের র‍্যাঙ্ক (R^1)	Yস্কোরের র‍্যাঙ্ক (R_2)	$D=R_1 \sim R_y$	D^2
2	0	6	6	0	0
3	3	5	5	0	0
4	4	4	3.5	0.5	0.25
5	4	3	3.5	0.5	0.25
6	6	2	2	0	0
7	11	1	1	0	0

$\Sigma D^2 = 0.50$

র‍্যাঙ্ক পার্থক্যের নিয়ম অনুযায়ী সহগতির সহগাঙ্ক (ρ)- এর মান নির্ণয় করা সম্ভব।

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)}$$

$N = 6, \Sigma D^2 = 0.50$

$$\therefore \rho = 1 - \frac{6 \times 0.50}{6(6^2-1)}$$

$$= 1 - \frac{3}{6 \times 35} = 1 - \frac{1}{70} = \frac{69}{70} = 0.986$$

Interpretation :

দুটি পরিমাপের স্কোরের মধ্যে সহগতি সহগাঙ্কের মান 0.986 অর্থাৎ এদের মধ্যে সম্পর্ক অতি উচ্চমানের ও ধনাত্মক। অর্থাৎ একটির মান পরিবর্তনে অপরটির মানেরও প্রায় সমপরিমাণ ও সমপ্রকৃতির পরিবর্তন পরিলক্ষিত হচ্ছে।

Example 5 :

Compute the value of co-efficient of co-relation between the following two series of scores obtained by a group of 10 students in Physical Science and Mathematics.

Physic	50	65	73	80	65	70	60	67	55	75
Math	70	81	90	85	90	90	65	75	72	92

Ans. (a) Rank difference Method (co-efficient correlation)