सहसम्बन्ध
[CORRELATION]

सांख्यिकीय माध्यों द्वारा हमें औसत मूल्य या प्रतिनिधि इकाई का पता चलता है, जबकि अपकिरण के मापों द्वारा समंकों के विचलन एवं श्रेणियों की आकृति के बारे में पता चलता है। सहसम्बन्ध का प्रयोग दो चरों, समूहों अथवा श्रेणियों में सम्बन्ध ज्ञात करने के लिए किया जाता है। यदि हम अपने सामान्य जीवन में विभिन्न चरों, समूहों तथा घटनाओं को देखें तो अनेक दो चरों में, समूहों में अथवा घटनाओं में किसी-न-किसी प्रकार के सम्बन्ध का आभास होता है। उदाहरण के लिए—यदि किसी वस्तु की माँग बढ़ जाए तो उसके मूल्यों में भी वृद्धि हो जाती है अथवा यदि वस्तु की पूर्ति बढ़ जाए तो उसके मूल्यों में कमी हो जाती है।

इसका यह अर्थ हुआ कि वस्तु की माँग या वस्तु की पूर्ति और इनके मूल्यों में सहसम्बन्ध पाया है। केवल आर्थिक पहलुओं में ही नहीं अपितु अनेक सामाजिक घटनाओं एवं तथ्यों में भी इसी प्रकार के सहसम्बन्ध पाए जाते हैं। उदाहरण के लिए—**दुर्खीम** ने आत्महत्या के अध्ययन में यह ज्ञात किया कि आत्महत्या की दर सामाजिक एकीकरण की मात्रा से जुड़ी हुई है अर्थात् यदि समूह में एकता अधिक है तो कम आत्महत्याएँ होती हैं और यदि समूह में एकीकरण कम है तो आत्महत्याएँ अधिक होती हैं। इसका सीधा अर्थ यह हुआ कि आत्महत्या की दर तथा सामाजिक समूह की एकता में सहसम्बन्ध पाया जाता है। इसी प्रकार, हम कमजोर मस्तिष्कों को अपराध से, मानसिक बीमारियों को आत्महत्या से, गरीब गृह परिस्थितियों को अस्वस्थता दर से, बाल अपराध को टूटे परिवारों से तथा तलाक को शिक्षा से सहसम्बन्धित बताने का प्रयास करते हैं।

सहसम्बन्ध का अर्थ एवं परिभाषाएँ

जब दो चरों (Variables) में से एक चर के बढ़ने से दूसरे चर में वृद्धि या कमी हो अथवा एक चर में कमी से दूसरे चर में वृद्धि या कमी हो तो इसका अर्थ हुआ कि दोनों चरों में किसी-न-किसी प्रकार का सहसम्बन्ध पाया जाता है। इसे निम्नांकित प्रकार से परिभाषित किया गया है—

(1) **बोडिंगटन** (Boddington) के अनुसार—''जब दो या दो से अधिक समूहों, वर्गों या आँकड़ों की श्रेणियों के मध्य एक निश्चित सम्बन्ध होता है तो उसे सहसम्बन्ध कहते हैं।''

(2) **कॉनर** (Connor) के अनुसार—''जब दो या अधिक राशियाँ सहानुभूति से परिवर्तित होती हैं, जिसमें एक में होने वाले परिवर्तनों के परिणामस्वरूप दूसरी राशि में भी परिवर्तित होने की प्रवृत्ति पाई जाती है तो वे राशियाँ सहसम्बन्धित कहलाती हैं।''

(3) **क्राक्सटन तथा काउडन** (Croxton and Cowden) के अनुसार—''जब सम्बन्धों की संख्यात्मक प्रकृति होती है, तो उसे ज्ञात करने, मापने एवं उसे एक सूत्र में स्पष्ट करने के उचित सांख्यिकीय यन्त्र को सहसम्बन्ध कहते हैं।''

(4) **किंग** (King) के अनुसार—''यदि यह सत्य सिद्ध होता है कि अधिकांश उदाहरणों में दो चर सदैव एक ही दिशा में या विपरीत दिशाओं में बढ़ने या घटने की प्रवृत्ति रखते हैं, तो हम यह मानते हैं कि उनमें एक सम्बन्ध पाया जाता है, जो सहसम्बन्ध कहलाता है।''

सहसम्बन्ध के प्रकार

मुख्य रूप से सहसम्बन्धों में दो प्रकार से भेद किया जाता है—

(1) धनात्मक और ऋणात्मक सहसम्बन्ध (Positive and negative correlation)— जब चरों, समूहों अथवा श्रेणियों में परिवर्तन की दिशा एक ही हो (अर्थात् यदि एक में वृद्धि या कमी होती है तो दूसरे में भी वृद्धि या कमी होती है) तो उनके बीच पाए जाने वाले सम्बन्धों को धनात्मक सम्बन्ध अथवा प्रत्यक्ष सम्बन्ध कहा जाता है। इसके विपरीत, जब दो चरों, समूहों अथवा श्रेणियों के बीच परिवर्तन की दिशा विलोम हो (अर्थात् एक में परिवर्तन होने पर उससे सम्बद्ध दूसरे में विपरीत दिशा में परिवर्तन हो) तो उसे ऋणात्मक सहसम्बन्ध अथवा अप्रत्यक्ष सहसम्बन्ध कहा जा सकता है।

(2) रेखीय तथा वक्रीय सहसम्बन्ध (Linear and curvilinear correlation)—जब एक चर, समूह या श्रेणी में परिवर्तन होने पर दूसरे चर, समूह या श्रेणी में भी उसी अनुपात में परिवर्तन हो, तो ऐसे सहसम्बन्ध को रेखीय सहसम्बन्ध कहा जाता है। इसके विपरीत, जब दो चरों, समूहों या श्रेणियों में परिवर्तन

एक समान अनुपात में न हो (अर्थात् एक में परिवर्तन होने पर दूसरे में परिवर्तन उसी अनुपात में न हो) तो उसे वक्रीय सहसम्बन्ध कहा जाता है।

सहसम्बन्धों की सीमा अथवा परिमाण

सहसम्बन्ध गुणांक (Coefficient of correlation) के माध्यम से हम सहसम्बन्धों की सीमा को ज्ञात करते हैं। सहसम्बन्ध का माप सदैव 1 के मध्य होता है। साधारणत: दो चरों में कभी भी पूर्ण धनात्मक सहसम्बन्ध (1) अथवा पूर्ण ऋणात्मक सहसम्बन्ध (1) नहीं होता है। धनात्मक तथा ऋणात्मक सहसम्बन्ध की निम्नलिखित सीमाएँ हो सकती हैं—

(1) पूर्ण सहसम्बन्ध (Perfect correlation)—जब दो चरों या श्रेणियों में समान दिशा तथा समान अनुपात में परिवर्तन होता है तो ऐसा सहसम्बन्ध **पूर्ण धनात्मक सहसम्बन्ध** कहलाता है। यह सहसम्बन्ध गुणांक **+ 1** होता है। इसके विपरीत, जब दो चरों अथवा श्रेणियों में परिवर्तन समान अनुपात परन्तु विपरीत दिशा में होता है तो इसे **पूर्ण ऋणात्मक सहसम्बन्ध** कहा जाता है। ऐसी स्थिति में सहसम्बन्ध गुणांक **- 1** होता है।

(2) सहसम्बन्ध की अनुपस्थिति (Absence of correlation)—यदि दो चरों में अथवा श्रेणियों में परस्पर कोई आश्रितता तथा सहानुभूति नहीं है (अर्थात् एक में परिवर्तन का दूसरे पर कोई प्रभाव नहीं पड़ता) तो ऐसी स्थिति में सहसम्बन्ध की अनुपस्थिति पाई जाती है। इस स्थिति में सहसम्बन्ध गुणांक **शून्य** (0) होता है।

(3) सहसम्बन्ध के सीमित परिमाण (Limited degrees of correlation)—यह वह स्थिति है जिसमें दो चरों या श्रेणियों में न ही पूर्ण सहसम्बन्ध पाया जाता है और न ही सहसम्बन्ध की अनुपस्थिति। इसे **आंशिक सहसम्बन्ध** भी कहा जाता है तथा यह धनात्मक एवं ऋणात्मक दोनों प्रकार का हो सकता है। सामान्यत: इसकी तीन दशाएँ हो सकती हैं—

(i) उच्च सहसम्बन्ध (High degree of correlation)—जब दो चरों या श्रेणियों का सहसम्बन्ध गुणांक 0 75 और 1 के मध्य होता है तो उसे उच्च सहसम्बन्ध कहते हैं। अधिकतर यह 0 9 के पास ही बना रहता है।

(ii) मध्यम सहसम्बन्ध (Moderate degree of correlation)—जब दो चरों में सहसम्बन्ध गुणांक 0 25 और 0 75 के मध्य होता है तो उसे मध्यम सहसम्बन्ध माना जाता है।

(iii) निम्न सहसम्बन्ध (Low degree of correlation)—जब दो चरों में सहसम्बन्ध गुणांक 0 और 0 25 के बीच होता है तो इसे निम्न सहसम्बन्ध माना जाता है।

निम्नांकित तालिका सहसम्बन्ध गुणांक द्वारा सहसम्बन्ध की सीमाओं को स्पष्ट करती है—

सहसम्बन्ध की सीमाएँ

सीमा (Degree)	**धनात्मक ( )**	**ऋणात्मक ( )**
पूर्ण सहसम्बन्ध	1	-1
उच्च सहसम्बन्ध	0 75 से 1 के बीच	0 75 से 1 के बीच
मध्यम सहसम्बन्ध	0 25 से 0 75 के बीच	0 25 से 0 75 के बीच
निम्न सहसम्बन्ध	0 से 0 25 के बीच	0 से 0 25 के बीच
सहसम्बन्ध की अनुपस्थिति	शून्य (0)	शून्य (0)

सहसम्बन्ध ज्ञात करने की विधियाँ

सहसम्बन्ध ज्ञात करने की प्रमुख विधियाँ निम्नलिखित हैं—

(क) अगणितीय विधियाँ (Non-mathematical methods) :

(1) विशेष चित्र विधि (Scatter or dat diagram method)

(2) बिन्दुरेखीय विधि (Graphic method)।

(ख) सहसम्बन्ध सारणी (Correlation table)।

(ग) गणितीय विधियाँ (Mathematical methods) :

(1) कार्ल पियर्सन का सहसम्बन्ध गुणांक (Karl Pearson's coefficient of correlation)

(2) संगामी विचलन गुणांक (Coefficient of concurrent deviation)

(3) स्पियरमैन की श्रेणी अन्तर विधि (Spearman's Ranking Method)

(4) न्यूनतम वर्ग विधि (Least square method)।

कार्ल पियर्सन का सहसम्बन्ध गुणांक

कार्ल पियर्सन का सहसम्बन्ध गुणांक बीजगणितीय दृष्टि से उत्तम है क्योंकि यह श्रेणी के सभी पदों अथवा मूल्यों पर आधारित है। इसमें सहसम्बन्ध की दिशा तथा मात्रा का अंकात्मक माप भी प्राप्त होता है। समान्तर माध्य एवं प्रमाप विचलन पर आधारित होने के कारण इसमें पूर्ण शुद्धता पाई जाती है।

इसे निम्नांकित सूत्रों द्वारा व्यक्त किया जा सकता है—

(अ) व्यक्तिगत श्रेणी में सहसम्बन्ध गुणांक

इसे प्रत्यक्ष विधि एवं लघु विधि द्वारा निकाला जा सकता है—

(1) प्रत्यक्ष विधि (Direct method) : सूत्र,

(i) $$\boxed{r \quad \frac{x \quad y}{n\sigma x.\sigma y}}$$ अथवा

(ii) $$\boxed{r \quad \frac{dxdy}{\sqrt{\frac{dx^2}{a}}\quad\sqrt{\frac{dy^2}{n}}}}$$ अथवा

(iii) $$\boxed{r \quad \frac{dxdy}{\sqrt{dx^2 \quad dy^2}}}$$

जहाँ, r सहसम्बन्ध का गुणांक (Coefficient of correlation),

dxdy दोनों श्रेणियों के माध्यों से निकाले गए सम्बन्धित विचलनों के गुणनफलों का योग (Sum of products of deviation of *x* and *y* series from their respective deviations)

n पदों की संख्या (Number of items),

x *y* *x* व *y* श्रेणियों का प्रमाप विचलन (Standard deviation of *x* and *y* series),

dx^2 श्रेणी के विचलनों के वर्ग का योग (Sum of the square of deviations of x series),

dy^2 y श्रेणी के विचलनों के वर्ग का योग (Sum of the square of deviations of y series)।

(2) लघु विधि (Short-cut method)

इसमें निम्नलिखित सूत्र का प्रयोग किया जाता है—

(i) $$\boxed{r \quad \frac{dxdy - n\,(a_1 - x_1)\,(a_2 - x_2)}{n \quad x \quad y}}$$ अथवा

(ii) $$\boxed{r \quad \frac{dxdy - n \quad \frac{dx}{n} \quad \frac{dy}{n}}{\sqrt{\frac{dx^2}{n} - \quad \frac{dx}{n}^{\;2} \qquad \frac{dy^2}{n} - \quad \frac{dy}{n}^{\;2}}}}$$ अथवा

(iii) $$r \quad \frac{dxdy - \dfrac{dx - \ dy}{n}}{\sqrt{dx^2 - \dfrac{(\ dx)^2}{n}} \quad \sqrt{dy^2 - \dfrac{(\ dy)^2}{n}}}$$ अथवा

(iv) $$r \quad \frac{dxyx \quad n - (\ dx)(\ dy)}{\sqrt{dx^2 \quad n - (\ dx^2)} \quad \sqrt{dy^2 \quad n - (\ dy)^2}}$$

जहाँ, *dxdy* कल्पित माध्यों से दोनों श्रेणियों के विचलनों का योग,

n पदों की संख्या,

a_1 श्रेणी का वास्तविक समान्तर माध्य,

a_2 y श्रेणी का वास्तविक समान्तर माध्य,

x_1 x श्रेणी का कल्पित माध्य,

y_2 y श्रेणी का कल्पित माध्य,

x x श्रेणी का प्रमाप विचलन, तथा

y y श्रेणी का प्रमाप विचलन।

(ब) वर्गीकृत श्रेणियों में सहसम्बन्ध गुणांक

इसमें पहले द्विचर आवृत्ति सारणी (Two-way frequency table) के अन्तर्गत x तथा y चरों के मापों को निम्नांकित सूत्र द्वारा वर्गीकृत किया जाता है—

$$r \quad \frac{fdxdy \quad n \quad (\ fdx)(\ fdy^2)}{\sqrt{[\ fdx^2 \quad n \quad (\ dyx^2)]} \quad \sqrt{[\ fdy^2 \quad n \quad (\ fdy)]}}$$

(स) काल श्रेणियों के अन्तर्गत सहसम्बन्ध गुणांक

काल श्रेणियों में भी सहसम्बन्ध गुणांक का प्रयोग किया जाता है। दीर्घकालीन परिवर्तनों में सहसम्बन्ध ज्ञात करने के लिए पहले तीन या पाँच वर्षीय चल माध्य ज्ञात किए जाते हैं। इनको प्रवृत्ति मूल्य भी कहा जाता है। इसके बाद प्रथम श्रेणी के चल माध्यों को x तथा दूसरी को y मानकर प्रत्यक्ष अथवा लघु रीति द्वारा कार्ल पियर्सन सहसम्बन्ध गुणांक निकाला जाता है।

अल्पकालीन परिवर्तनों के सहसम्बन्ध के लिए पहले x तथा y श्रेणी के चल माध्य ज्ञात किए जाते हैं। फिर प्रत्येक माप में से तत्सम्बन्धी चल माध्य घटाकर अल्पकालीन विचलन ज्ञात किए जाते हैं। फिर प्रत्येक माप में से तत्सम्बन्धी चल माध्य घटाकर अल्पकालीन विचलन ज्ञात किए जाते हैं। इसके पश्चात् निम्नांकित सूत्र द्वारा सहसम्बन्ध गुणांक निकाला जाता है—

$$r \quad \frac{dxdy}{\sqrt{(\ dx^2 \quad dy^2)}}$$

जहाँ, *dx* तथा *dy* क्रमशः श्रेणी *x* तथा श्रेणी *y* के विचलन हैं। ●