

สหสัมพันธ์

ความหมาย

สหสัมพันธ์ (Correlation) หมายถึง ปริมาณความเกี่ยวข้องระหว่างลักษณะของบุคคล หรือสิ่งของ 2 ลักษณะขึ้นไป เช่น ทศนคติต่อสภาพแวดล้อม และระดับการศึกษาเป็นลักษณะที่อาจเกี่ยวข้องกันหรือมีความสัมพันธ์กัน ซึ่งเรียกว่า สหสัมพันธ์อย่างง่าย (simple correlation) หรือ หมายถึงปริมาณความเกี่ยวข้องระหว่างลักษณะมากกว่า 2 ลักษณะขึ้นไป เรียกว่า สหสัมพันธ์พหุคูณ (multiple correlation) หรืออาจหมายถึง ปริมาณความสัมพันธ์ระหว่างลักษณะ 2 ลักษณะ เมื่อต้องการควบคุมตัวแปรอื่นๆ ที่อาจเข้ามาเกี่ยวข้องกับลักษณะทั้งสองนี้ ซึ่งเรียกว่าสหสัมพันธ์พาร์เชียล (partial correlation)

ในทางสถิติ สหสัมพันธ์ หมายถึง ปริมาณความสัมพันธ์เชิงเส้น (linear relationship) ระหว่างตัวแปรสุ่ม (random variables) ตั้งแต่สองตัวขึ้นไป โดยที่ตัวแปรสุ่มเหล่านี้มีการแจกแจงหลายตัวแปรแบบปกติ (bivariate normal distribution หรือ multivariate normal distribution)

ความเป็นมา

การศึกษาเรื่องสหสัมพันธ์นี้เกิดขึ้นหลังจากที่ เซอร์ ฟรานซิส แกลตัน (Sir Francis Galton) นักมานุษยวิทยาชาวอังกฤษ ตีพิมพ์ การค้นพบใน ค.ศ. 1885 เกี่ยวกับ การทำนาย ลักษณะทางกายภาพของบุตร เมื่อรู้ลักษณะทางกายภาพของบิดามารดาแล้ว ซึ่งแกลตันพบว่าใน

กรณีที่บิดา มีความสูงมาก โดยเฉลี่ยแล้วมีบุตรที่สูงน้อยกว่าตนเอง ในขณะที่บิดาที่มีความสูงน้อย โดยเฉลี่ยแล้วมีบุตรที่สูงกว่าตนเอง ปรากฏการณ์นี้เขาใช้คำว่า การถดถอย (regression)

คาร์ล เพียร์สัน (Karl Pearson) นักวัดผลทางชีววิทยา (biometrician) ชาวอังกฤษ เป็นบุคคลแรกที่นำความคิดของ แกลตัน มาขยายและพัฒนาวิธีการขึ้นมาใหม่ที่เรียกว่า การศึกษาสหสัมพันธ์

ดัชนีสหสัมพันธ์

ดัชนีสหสัมพันธ์ เป็นค่าทางสถิติที่บอกให้ทราบถึงปริมาณของความสัมพันธ์ระหว่างตัวแปรโดยทั่วไป ดัชนีสหสัมพันธ์หรือสัมประสิทธิ์สหสัมพันธ์ จะมีค่าตั้งแต่ -1 ถึง +1 ซึ่งหมายความว่าถ้าตัวแปรคู่ใดมีค่าดัชนีสหสัมพันธ์เป็น -1 หรือ +1 แสดงว่าตัวแปรคู่นั้นมีความสัมพันธ์กันสูงที่สุด นั่นคือเมื่อค่าของตัวแปรหนึ่งเพิ่มขึ้น ค่าของตัวแปรอีกตัวหนึ่งเพิ่มขึ้นด้วย (ดัชนีสหสัมพันธ์ = +1) หรือค่าของตัวแปรหนึ่งเพิ่มขึ้นขณะที่ค่าของตัวแปรอีกตัวหนึ่งลดลง (ดัชนีสหสัมพันธ์ = -1) ถ้าดัชนีสหสัมพันธ์เป็น 0 แสดงว่าตัวแปรคู่นั้นไม่มีความสัมพันธ์กันเลย

ประเภทของดัชนีสหสัมพันธ์

1. กรณีที่ศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว

- 1.1 สัมประสิทธิ์สหสัมพันธ์เพียร์สัน ใช้สัญลักษณ์ r ใช้เมื่อตัวแปรทั้งสองตัว มีการ

วัดในระดับ อินตรภาค (interval)* ขึ้นไป เป็น
ดัชนีที่มีผู้นิยม ใช้มากที่สุด อีกทั้งมีค่าความค
ลาดเคลื่อน (standard error) ต่ำที่สุดเมื่อ
เปรียบเทียบกับดัชนีตัวอื่นๆ

1.2 สัมประสิทธิ์สหสัมพันธ์ไบซีเรียล
(biserial correlation coefficient) ใช้สัญลักษณ์
 r_{bis} ใช้เมื่อตัวแปรหนึ่งมีการวัดในระดับอินตรภาค
ขึ้นไป และตัวแปรอีกตัวหนึ่งถูกทำให้มีค่า 2 ค่า
แบบเทียม (artificial dichotomy) นิยมใช้ใน
การวิเคราะห์รายข้อซึ่งเป็นการหาความสัมพันธ์
ระหว่างคะแนนเป็นรายข้อกับคะแนนรวมทั้งฉบับ
ค่าที่คำนวณได้อาจมีค่าเกิน 1 ดัชนีความสัมพันธ์
ชนิดนี้ความคลาดเคลื่อนสูงกว่า r

1.3 สัมประสิทธิ์สหสัมพันธ์พอยท์
ไบซีเรียล (point-biserial correlation coeffi-
cient) ใช้สัญลักษณ์ r_{pbis} ใช้เมื่อตัวแปรหนึ่งมี
การวัดในระดับอินตรภาคขึ้นไป และตัวแปรอีกตัว
หนึ่งมีค่า 2 ค่า (dichotomous variable) ค่า
ที่คำนวณได้มีค่าต่ำกว่า r_{bis}

1.4 สัมประสิทธิ์สหสัมพันธ์
เตตราโคริก (tetrachoric correlation coeffi-
cient) ใช้สัญลักษณ์ r_t ใช้เมื่อตัวแปรทั้งสองตัว
มีค่า 2 ค่า แบบเทียม

1.5 สัมประสิทธิ์คอนทินเจนซี
(contingency coefficient) ใช้สัญลักษณ์ c
ใช้เมื่อตัวแปรทั้งสองตัวมีระดับของการวัดเป็น
นามบัญญัติ (nominal scale) ดัชนีความ
สัมพันธ์ตัวนี้มีความเกี่ยวข้องกับค่าสถิติไคสแควร์
(chi-square)

1.6 สัมประสิทธิ์สหสัมพันธ์โร
(rank difference correlation coefficient)
ใช้สัญลักษณ์ ρ ใช้เมื่อตัวแปรทั้งสองมีการวัด
เป็นมาตราอันดับ (ordinal scale) การคำนวณ
ใช้สูตรที่ดัดแปลงมาจากสูตรของ r

1.7 สัมประสิทธิ์สหสัมพันธ์พาร์เชียล
(partial correlation coefficient) ใช้เมื่อ
ต้องการศึกษาความสัมพันธ์ระหว่างตัวแปรสองตัว
และมีการดึงความสัมพันธ์ของตัวแปรตัวที่สามที่
มีต่อตัวแปร 2 ตัวแรกออกมาแล้ว

2. กรณีที่ศึกษาความสัมพันธ์ของ
ตัวแปรมากกว่า 2 ตัว

สัมประสิทธิ์สหสัมพันธ์พหุคูณใช้
สัญลักษณ์ R ใช้เมื่อต้องการศึกษาความสัมพันธ์
ระหว่างตัวแปร 1 ตัว ที่เรียกว่าตัวแปรเกณฑ์
(criterion variable) กับตัวแปรอีกกลุ่มหนึ่ง
ซึ่งมีตั้งแต่ 2 ตัวแปรขึ้นไปที่เรียกรวมกันว่า
ตัวแปรทำนาย (predictor variables)

*ระดับของการวัดแบ่งออกเป็น 3 มาตรา คือ

(1) มาณานามบัญญัติ (nominal scale) หมายถึง การใช้ตัวเลข หรือสัญลักษณ์ในการจัดกลุ่มคนหรือ
สิ่งของออกเป็นประเภทหรือชนิดต่างๆ เช่น เลข 1 หมายถึง เพศชาย เลข 2 หมายถึง เพศหญิง ซึ่งตัวเลขเหล่านี้มิ
ได้มีความหมายในเชิงปริมาณ

(2) มาณอันดับ (ordinal scale) หมายถึง การใช้ตัวเลขในการแบ่งประเภทของคนหรือสิ่งของ และ
ตัวเลขมีความสัมพันธ์กันในเชิงที่บอกได้ว่า มากกว่าหรือน้อยกว่า เช่น ระดับเศรษฐกิจ 1 หมายถึง ระดับต่ำ 2
หมายถึงระดับกลาง และ 3 หมายถึง ระดับสูง ซึ่งจะเห็นว่า $3 > 2 > 1$

(3) มาณอินตรภาค (interval scale) หมายถึง การใช้ตัวเลข เพื่อแบ่งประเภทของคนหรือสิ่งของ
และตัวเลขนั้นมีคุณสมบัติที่บอก ความสัมพันธ์ในเชิงมากกว่าหรือน้อยกว่า และช่วงระหว่างตัวเลขสองตัวเป็นค่าที่
ประมาณได้ เช่น อุณหภูมิ ใช้การวัดระดับมาตราอินตรภาค

ความสำคัญและการนำมาใช้ประโยชน์

ดัชนีสหสัมพันธ์เหล่านี้มีประโยชน์สองประการ ประการแรก คือ ชี้ให้เห็นว่าตัวแปรมีความเกี่ยวข้องกันหรือไม่ และถ้าเกี่ยวข้องกัน ลักษณะของความเกี่ยวข้องเป็นอย่างไร ถ้าความสัมพันธ์ระหว่างตัวแปรไม่เป็นศูนย์ เราอาจทำนายตัวแปรหนึ่งจากการรู้ค่าของตัวแปรอื่นได้ ยกตัวอย่างเช่น ถ้าคะแนนสอบเข้ามหาวิทยาลัยมีความสัมพันธ์ทางบวก (สมมติ $r = .75$) กับคะแนนสอบภาคปลายของนิสิตชั้นปีที่ 1 ดังนั้นถ้าเรารู้คะแนนสอบเข้ามหาวิทยาลัยของ นาย ก เราก็อาจทำนายคะแนนสอบปลายภาคของเขาได้อย่างค่อนข้างใกล้เคียง ประโยชน์อีกประการหนึ่งคือประโยชน์ทางการวิจัยเพื่อค้นหาตัวแปรที่อาจช่วยในการอธิบายปรากฏการณ์ใหม่ๆ โดยการหาความสัมพันธ์ระหว่างตัวแปรที่สนใจ ผลที่

ได้จากการวิเคราะห์จะเป็นข้อมูลเบื้องต้นในการศึกษาให้ลึกซึ้งต่อไป ข้อที่ควรระมัดระวังในการใช้ดัชนีสหสัมพันธ์ก็คือ เมื่อพบว่าตัวแปรคู่ใด ๆ มีความสัมพันธ์กัน มิได้หมายความว่าตัวแปรหนึ่งเป็นสาเหตุของอีกตัวแปรหนึ่ง เพราะถ้า A และ B มีความสัมพันธ์กันอาจเป็นเพราะ 1) A เป็นสาเหตุให้เกิด B หรือ 2) B เป็นสาเหตุให้เกิด A หรือ 3) C เป็นสาเหตุให้เกิดทั้ง A และ B เช่นงานวิจัยหลายเรื่องพบความสัมพันธ์ระหว่างความพึงพอใจในงานกับประสิทธิภาพในการทำงาน ผลของการวิจัยอาจสรุปได้ว่า เพราะ 1) ความพึงพอใจในงานทำให้คนงานทำงานอย่างมีประสิทธิภาพ 2) เมื่อคนงานมีประสิทธิภาพเขาจึงมีความพึงพอใจ หรือ 3) คนงานได้เงินเดือนสูงมากทำให้เกิดความพึงพอใจและทำงานอย่างมีประสิทธิภาพ เป็นต้น

ดุษฎี โยเหลา

บรรณานุกรม

- Borg, W.R. and M.D. Gall. **Educational Research**. London : Longman, 1989.
- Ferguson, G.A. **Statistical Analysis in Psychology and Education**. New York : McGraw-Hill, 1981.
- Pedhazur, E.J. **Multiple Regression in Behavioral Research**. New York : Holt, Rinehart and Winston, 1982.