

Biserial and Point Biserial Correlation

สุทธินิพนธ์ พิศัยศักดิ์โสภณ*

บทนำ

เราเคยทราบกันมาแล้วว่าในการหาความสัมพันธ์ระหว่างตัวแปรสองตัว โดยที่ตัวแปรแต่ละตัวเป็นตัวแปรต่อเนื่อง และมีการแจกแจงปกติ เราใช้สหสัมพันธ์แบบเพียร์สัน (Pearson Product Moment Correlation) หรือที่นิยมใช้สัญลักษณ์ r_{xy} แต่ถ้าตัวแปรทั้งสองไม่เป็นตัวแปรต่อเนื่องทั้งคู่ เราก็สามารถหาความสัมพันธ์ระหว่างตัวแปรทั้งสองได้เช่นกัน ดังที่จะได้กล่าวต่อไปนี้

* อาจารย์ ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

Point Biserial Correlation

สหสัมพันธ์แบบพอยท์ไบซีเรียล เป็นการหาความสัมพันธ์ระหว่างตัวแปรทั้งสอง โดยที่ตัวแปรหนึ่งเป็นตัวแปรต่อเนื่องที่มีการแจกแจงปกติ และอีกตัวแปรหนึ่งเป็นแบบ Dichotomous แท้หรือตัวแปรที่จำแนกได้เป็นสองพวก เช่น เพศ แบ่งเป็น ชาย-หญิง เราใช้สัญลักษณ์ 1 แทนคนในกลุ่มหนึ่งหรือพวกหนึ่ง(ชาย) และใช้ 0 แทน คนในอีกกลุ่มหนึ่งหรือพวกหนึ่ง(หญิง)เป็นต้น

สูตรที่ใช้ในการคำนวณ

หรือ
$$r_{pbi} = \frac{\bar{x}_p - \bar{x}_q}{s_x} \sqrt{pq} \quad \text{----- (1)}$$

เมื่อ
$$r_{pbi} = \frac{\bar{x}_p - \bar{x}}{s_x} \sqrt{\frac{p}{q}} \quad \text{----- (2)}$$

s_x แทน ส่วนเบี่ยงเบนมาตรฐานของคะแนนในตัวแปรต่อเนื่อง

p แทน สัดส่วนของจำนวนคนในกลุ่มหนึ่งกับจำนวนคนทั้งหมด

q แทน สัดส่วนของจำนวนคนในอีกกลุ่มหนึ่งกับจำนวนคนทั้งหมด

$\bar{x}_p$ แทน ค่าเฉลี่ยของตัวแปรต่อเนื่องเฉพาะในกลุ่มที่กำหนดเป็นพวก 1

$\bar{x}_q$ แทน ค่าเฉลี่ยของตัวแปรต่อเนื่องเฉพาะในกลุ่มที่กำหนดเป็นพวก 0

$\bar{x}$ แทน ค่าเฉลี่ยของตัวแปรต่อเนื่องของจำนวนคนทั้งหมด

ตัวอย่างที่ 1 จงหาความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนกับเพศ จากข้อมูลต่อไปนี้

นักเรียนคนที่	ผลสัมฤทธิ์ทางการเรียน	เพศ
1	6	0
2	8	1
3	8	0
4	11	0
5	16	1
6	25	0
7	27	0
8	31	0
9	31	1
10	39	0
11	44	0
12	50	1
13	56	1
14	68	1

ผลสัมฤทธิ์ทางการเรียน	เป็นตัวแปรต่อเนื่อง	กำหนดให้เป็น x
เพศ	เป็นตัวแปร Dichotomous แท้	กำหนดให้ 1 แทน ชาย และ 0 แทน หญิง

วิธีทำ 1. จากโจทย์จะได้ว่า $N = 14$

$$p = \frac{6}{14} = .43$$

$$q = \frac{8}{14} = .57$$

2. คำนวณค่า S_x , $\bar{X}_p$, $\bar{X}_q$ และ $\bar{X}$

$$\begin{aligned}\bar{X} &= \frac{(6+8+8+11+16+25+27+31+31+39+44+50+56+68)}{14} \\ &= 30\end{aligned}$$

$$S_x = \sqrt{\frac{\sum (X - \bar{X})^2}{N}} = 18.19$$

$$\bar{X}_D = \frac{8+16+31+50+56+68}{6} = 38.17$$

$$\bar{X}_q = \frac{6+8+11+25+27+31+39+44}{8} = 23.88$$

3. แทนค่าลงในสูตร (1) หรือ สูตร (2)

* สูตร (1)

$$\begin{aligned}r_{pbi} &= \frac{38.17 - 23.88}{18.19} \sqrt{(.43)(.57)} \\ &= 0.39\end{aligned}$$

* สูตร (2)

$$\begin{aligned}r_{pbi} &= \frac{38.17 - 30}{18.19} \sqrt{\frac{.43}{.57}} \\ &= 0.39\end{aligned}$$

เพราะฉะนั้น ความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนและเพศ เท่ากับ 0.39

การทดสอบนัยสำคัญ

เมื่อคำนวณค่า r_{pbi} เราสามารถทดสอบว่า r_{pbi} แตกต่างจาก 0 หรือไม่ โดยตั้งสมมติฐานดังนี้

$$H_0 : r_{pbi} = 0$$

$$H_1 : r_{pbi} \neq 0$$

สูตรที่ใช้ทดสอบนัยสำคัญคือ

$$t = r_{pbi} \sqrt{\frac{N-2}{1-r_{pbi}^2}} \quad \text{-----} \quad (3) \quad df = N-2$$

เมื่อ t แทน ค่า t -distribution

N แทน จำนวนคนทั้งหมด

ถ้า t จำนวน มากกว่าหรือเท่ากับ t ตาราง ณ ระดับนัยสำคัญที่กำหนด แสดงว่าตัวแปรทั้งสอง มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ

จากตัวอย่างที่ 1 นำมาทดสอบนัยสำคัญทางสถิติได้ดังนี้

$$t = .39 \sqrt{\frac{14-2}{1-(.39)^2}} = 1.47$$

ค่า t จำนวน เท่ากับ 1.47 เมื่อเปิดตารางการแจกแจงของ t ที่ $df = 12$ และ $= .05$ ได้ค่าเท่ากับ 2.179 มีค่ามากกว่า t จำนวน ดังนั้นผลสัมฤทธิ์ทางการเรียนกับเพศ มีความสัมพันธ์กันอย่างไม่มีนัยสำคัญทางสถิติ

Biserial Correlation

สหสัมพันธ์แบบไบซีเรียล เป็นการหาความสัมพันธ์ระหว่างตัวแปรสองตัว โดยที่ตัวแปรตัวหนึ่งเป็นตัวแปรต่อเนื่อง และอีกตัวแปรตัวหนึ่งถูกทำให้เป็นตัวแปรแบบ Dichotomous และมีการแจกแจงแบบปกติ เช่น การหาความสัมพันธ์ระหว่างสติปัญญากับการปรับตัวทางสังคม ซึ่งตัวแปรทั้งสอง เป็นตัวแปรต่อเนื่องทั้งคู่ แต่ตัวแปรตัวหนึ่งถูกทำให้เป็นตัวแปร Dichotomous โดยที่การปรับตัวทางสังคมแบ่งเป็น กลุ่มที่ปรับตัวทางสังคมได้ และกลุ่มที่ไม่สามารถปรับตัวทางสังคมได้ หรือความสนใจ แบ่งเป็น พวกสนใจ และพวกไม่สนใจ พวกที่เล่นกีฬา และพวกไม่เล่นกีฬา เป็นต้น

สูตรที่ใช้คำนวณ

$$r_{bi} = \frac{\bar{x}_p - \bar{x}_q}{s_x} \cdot \frac{pq}{y} \quad \text{-----} \quad (4)$$

หรือ

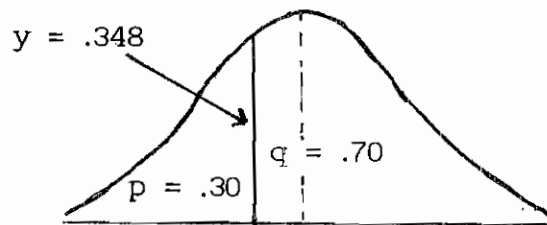
$$r_{bi} = \frac{\bar{x}_p - \bar{x}}{s_x} \cdot \frac{p}{y} \quad \text{-----} \quad (5)$$

เมื่อ y แทน ความสูงของ ordinate ของโค้งปกติ ณ จุดที่แบ่ง p และ q

$\bar{x}_p$, $\bar{x}_q$, $\bar{x}$, s_x , p และ q ทำนองเดียวกับสูตร r_{pbi}

ตัวอย่างการหาค่า y

สมมติว่า $p = .30$ และ $q = .70$ จากตารางการแจกแจงปกติ จะได้ความสูงของ ordinate = .348



ตัวอย่างที่ 2 จงหาความสัมพันธ์ระหว่างการวัดความซาบซึ้งทางดนตรีระหว่างนักเรียนที่ได้รับการฝึก และไม่ได้รับการฝึกทางดนตรี จากข้อมูลต่อไปนี้

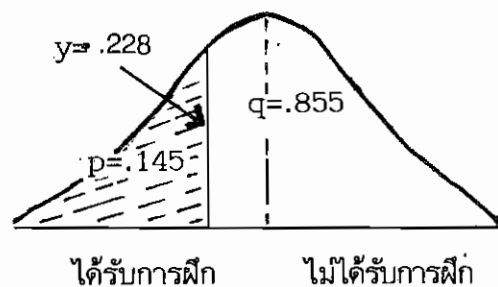
x	กลุ่มที่ได้รับการฝึก	กลุ่มที่ไม่ได้รับการฝึก	รวม
87	5	6	11
82	2	16	18
77	6	19	25
72	6	27	33
67	1	19	20
62	0	21	21
57	1	16	17
$N_1 = 21$		$N_2 = 124$	$N = 145$

วิธีทำ 1. จากโจทย์จะได้ว่า $N = 145$
 $N_1 = 21$
 $N_2 = 124$

$$P = \frac{21}{145} = .145$$

$$q = \frac{124}{145} = .855$$

$$y = .228$$



2. คำนวณค่า S_x , $\bar{x}_p$, $\bar{x}_q$, $\bar{x}$

S_x แทน ส่วนเบี่ยงเบนมาตรฐานของคะแนน x จากคนทั้งหมด (145 คน)
มีค่าเท่ากับ 8.80

$\bar{x}_p$ แทน ค่าเฉลี่ยของคะแนนจากกลุ่ม N_1 (21 คน) ซึ่งเท่ากับ 77.00

$\bar{x}_q$ แทน ค่าเฉลี่ยของคะแนนจากกลุ่ม N_2 (124 คน) ซึ่งเท่ากับ 70.39

$\bar{x}$ แทน ค่าเฉลี่ยของคะแนนจากคนทั้งหมด (145 คน) ซึ่งเท่ากับ 71.35

3. แทนค่าลงในสูตร (4) หรือสูตร (5) ก็ได้

* สูตร (4)

$$\begin{aligned} r_{bi} &= \frac{77.00 - 70.39}{8.80} \cdot \frac{(.145)(.855)}{.228} \\ &= .41 \end{aligned}$$

เพราะฉะนั้น การฝึกดนตรีและความชอบซึ่งในดนตรี มีความสัมพันธ์กัน เท่ากับ .41

การทดสอบนัยสำคัญ

การทดสอบนัยสำคัญของ r_{bi} ทำได้โดยใช้สูตร

$$t = \frac{r_{bi}}{S_{rbi}} \quad \text{-----} \quad (6)$$

เมื่อ

$$S_{rbi} = \frac{1}{Y} \sqrt{\frac{pq}{N}}$$

จากตัวอย่างที่ 2 นำค่า r_{bi} มาทดสอบนัยสำคัญได้ดังนี้

$$t = \frac{.41}{\frac{1}{.288} \sqrt{\frac{(.145)(.855)}{145}}} = 3.20$$

แสดงว่า r_{bi} เท่ากับ .41 มีนัยสำคัญทางสถิติที่ระดับ .01

นั่นคือ การฝึกดนตรีมีความสัมพันธ์กับความชาญซึ่งในดนตรี อย่างมีนัยสำคัญทางสถิติที่ระดับ .01

ประโยชน์ของ Point Biserial และ Biserial Correlation ทางการศึกษา

ประโยชน์จากการใช้ Point Biserial และ Biserial Correlation คือนำไปหาคุณภาพของข้อคำถาม ได้แก่ ค่าอำนาจจำแนก (Discrimination) ซึ่งจะเป็นตัวบ่งชี้ว่าข้อคำถามนั้นสามารถจำแนกความสามารถของผู้ตอบได้หรือไม่ หรือจำแนกคนเก่งออกจากคนอ่อนได้ โดยการหาความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมทั้งฉบับ ซึ่งคะแนน

รายข้ออาจเป็น 0-1 หรือไม่เป็น 0-1 ก็ได้

ค่าสหสัมพันธ์แบบพอยท์ไบซีเรียล จะให้ค่าสูงสุดเมื่อ $p=q=0.50$ ดังนั้น r_{pbi} เหมาะสำหรับข้อสอบที่มีความยากปานกลาง แต่ถ้าข้อสอบง่ายเกินไปหรือยากเกินไป r_{pbi} จะให้ค่าที่น้อยกว่าความเป็นจริง (Underestimate) จึงควรใช้สหสัมพันธ์ไบซีเรียลจะดีกว่า ซึ่งจะแก้ข้อด้อยของสหสัมพันธ์พอยท์ไบซีเรียลได้

$$r_{pbi} \text{ และ } r_{bi} \text{ มีความสัมพันธ์กันดังนี้} \quad r_{bi} = r_{pbi} \frac{\sqrt{pq}}{y}$$

บทสรุป

r_{pbi} และ r_{bi} เป็นการหาความสัมพันธ์ระหว่างตัวแปรสองตัว ซึ่งตัวแปรหนึ่งเป็นแบบ Dichotomous และอีกตัวแปรหนึ่งเป็นแบบต่อเนื่อง มีประโยชน์สำหรับนำไปใช้หาคุณภาพของเครื่องมือได้ แต่การที่จะเลือกใช้วิธีใดนั้น ควรดูจุดมุ่งหมายและลักษณะของข้อมูลด้วย ถ้าหากข้อมูลมีจำนวนมาก การคำนวณด้วยวิธีนี้อาจยุ่งยากและเสียเวลา และเกิดข้อผิดพลาดได้ จึงควรใช้คอมพิวเตอร์ช่วยในการคำนวณ ซึ่งจะทำให้ข้อมูลเชื่อถือได้มากขึ้น

เอกสารอ้างอิง

- บุษกร เพชรวิวรรธน์. สหสัมพันธ์ : ประเภทและวิธีการ. กรุงเทพฯ : สำนักทดสอบทางการศึกษาและจิตวิทยา มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร, 2523.
- สุณี รักษาเกียรติศักดิ์. "การตรวจสอบคุณภาพของเครื่องมือด้วย SPSS", วารสารวัดผลการศึกษา. 16(46) : พฤษภาคม-สิงหาคม 2537 ; 58-84.
- Ferguson, George A. Statistical Analysis in Psychology and Education. 4th ed. New York : McGraw-Hill, 1976.
- Garrett, Henry E. and R.S. Woodworth. Statistics in Psychology and Education. 4th ed. New York : David Mckay, 1967.