

ความเชื่อมั่นของผู้ประเมิน

ผศ. ดร. สุทธิวรรณ นิจศักดิ์โสภณ *

ความนำ

จากสาระสำคัญในหมวด 4 ตามพระราชบัญญัติแห่งชาติ พ.ศ. 2542 มาตรา 26 ว่าด้วยเรื่องของการประเมินผลการเรียน ทำให้ครูผู้สอนมีความตื่นตัวในเรื่องวิธีการประเมินว่าจะประเมินอย่างไร จึงจะสอดคล้องกับสภาพที่แท้จริงของผู้เรียน และการประเมินนั้นถ้ามีผู้ประเมินหลายคนจะมีความเชื่อถือได้เพียงไรยังคงเป็นคำถามที่ต้องการคำตอบอยู่

สำหรับการประเมินผู้เรียนหรือประเมินกลุ่มเป้าหมายที่มีผู้ประเมินตั้งแต่สองคนขึ้นไป เราสามารถหาความเชื่อถือได้ระหว่างผู้ประเมินเหล่านั้น โดยใช้วิธีทางสถิติที่เป็นความสัมพันธ์ระหว่างผู้ประเมิน (Interrater Correlation) ทั้งนี้ก็ต้องขึ้นอยู่กับลักษณะของเครื่องมือที่ใช้ประเมิน คะแนนที่ประเมิน วิธีที่ใช้คำนวณนั้นมีอยู่หลายวิธี แต่ละวิธีต่างก็มีข้อตกลงของการใช้ต่างกัน จะเลือกใช้วิธีใด ก็ควรให้เหมาะสมกับลักษณะของข้อมูลที่ได้มา

สำหรับบทความนี้จะขอกล่าวถึงการหาความเชื่อมั่นระหว่างผู้ประเมินสองคน และในคราวต่อไปนี้จะนำเสนอการหาความเชื่อมั่นกรณีที่มีผู้ประเมินหลายคน

การหาความเชื่อมั่นระหว่างผู้ประเมินสองคน : กรณีที่ให้คะแนนแบบทวิ

ในการหาความสัมพันธ์ระหว่างผู้ประเมินสองคนเพื่อดูว่าผู้ประเมินทั้งสองคนนี้ให้คะแนนนักเรียนได้สอดคล้องกันหรือไม่ เราสามารถหาได้โดยใช้สัมประสิทธิ์สหสัมพันธ์แบบ Pearson r product-moment ถ้าคะแนนที่ประเมินเป็นแบบต่อเนื่องหรืออยู่ในมาตราวัดระดับอัตราภาค (interval) ขึ้นไป แต่ถ้าคะแนนที่ประเมินเป็นแบบจัดกลุ่มออกมาเป็น 2 กลุ่มหรือที่เรียกว่าให้คะแนนแบบทวิ (ใช่ - ไม่ใช่, ถูก - ผิด หรือ 0 - 1) เราคงไม่สามารถหาความสอดคล้องระหว่างผู้ประเมินสองคนโดยใช้สัมประสิทธิ์แบบเพียร์สันได้ แต่เราก็สามารถหาได้โดยวิธีอื่น

* ผู้ช่วยศาสตราจารย์ประจำภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์
มหาวิทยาลัยมหาสารคาม

26 การวัดผลการศึกษามมส

ลักษณะของข้อมูลที่จะประเมินความสอดคล้องระหว่างผู้ประเมินสองคนเมื่อมีการให้คะแนนแบบ 0 - 1 สามารถจัดรูปข้อมูลได้ดังนี้

	ผู้ประเมินคนที่ 1 (1)	ผู้ประเมินคนที่ 1 (0)
ผู้ประเมินคนที่ 2 (1)	A	B
ผู้ประเมินคนที่ 2 (0)	C	D

จากรูปแบบลักษณะของข้อมูล เราสามารถคำนวณหาความสอดคล้องระหว่างผู้ประเมินสองคน ได้ดังวิธีต่อไปนี้

$$(1) \quad \text{Phi} = \frac{(AD - BC)}{\sqrt{(A + B)(C + D)(A + C)(B + D)}}$$

$$(2) \quad \text{Yule's Q (หรือ gamma)} = \frac{(AD - BC)}{(AD + BC)}$$

$$(3) \quad \text{The Jaccard ; J} = \frac{A}{(A + B + C)}$$

$$(4) \quad \text{The G-Index ; G} = \frac{[(A + D) - (B + C)]}{N}$$

$$(5) \quad \text{Bennett's B index ; B} = \frac{AD - x_{kon}^2}{(A + x_{kon}^2)(D + x_{kon}^2)}$$

$$\text{เมื่อ } x_{kon} = (B+C)/2$$

วิธีการทั้ง 5 ที่กล่าวมานี้สามารถดาวน์โหลดโปรแกรม DICHOT3.0 ซึ่งเป็นโปรแกรมที่ใช้ประมวลค่าความเชื่อมั่นของผู้ประเมินสองคนที่มีการให้คะแนนแบบ 0 - 1 ได้จาก <http://www.liv.ac.uk/~pbarrett/programs.htm>

ตัวอย่างการคำนวณ

สมมติผู้ประเมินสองคนประเมินพฤติกรรมของผู้เรียน 64 คน มีความเห็น ดังตารางต่อไปนี้

	ผู้ประเมินคนที่ 1 (1)	ผู้ประเมินคนที่ 1 (0)
ผู้ประเมินคนที่ 2 (1)	33	13
ผู้ประเมินคนที่ 2 (0)	10	8

จากตารางจะเห็นว่าทั้งผู้ประเมินคนที่ 1 และคนที่ 2 มีความเห็นสอดคล้องกันในการประเมินนักเรียนว่าให้ 1 คะแนน จำนวน 33 คน และมีความเห็นสอดคล้องกันว่าให้ 0 คะแนน จำนวน 8 คน เมื่อคำนวณโดยใช้วิธีทั้ง 5 จะได้ผลดังนี้

วิธี	ค่าที่ได้
Phi	0.1550
Yule's Q (Gamma)	0.3401
Jaccard	0.5893
G-Index	0.2813
Bennett's B-Index	0.1498

ตัวอย่างการคำนวณเมื่อใช้สัมประสิทธิ์ Phi มีดังนี้

$$\begin{aligned} \text{Phi} &= \frac{(33)(8) - (13)(10)}{\sqrt{(33+13)(10+8)(33+10)(13+8)}} \\ &= \frac{134}{\sqrt{747684}} \\ &= 0.1550 \end{aligned}$$

สำหรับวิธีอื่น ๆ ก็คำนวณง่าย ๆ โดยแทนค่าลงในแต่ละสูตร

การหาความเชื่อมั่นระหว่างผู้ประเมินสองคน : กรณีมีการให้คะแนนมากกว่า 2 กลุ่ม

การหาความเชื่อมั่นระหว่างผู้ประเมินสองคน เมื่อมีการให้คะแนนแบบมากกว่า 2 กลุ่ม สามารถใช้สัมประสิทธิ์แคปป่าของโคเฮน (Cohen's Kappa) จากสูตรต่อไปนี้

$$(6) \quad K = \frac{\sum f_o - \sum f_e}{N - \sum f_e}$$

เมื่อ K เป็น ค่าสัมประสิทธิ์ความเชื่อมั่นระหว่างผู้ประเมินสองคน

$\sum f_o$ เป็น ความถี่ที่สังเกตได้ในแนวทแยง

$\sum f_e$ เป็น ความถี่ที่คาดหวังในแนวทแยง

N เป็น จำนวนคนทั้งหมด

สำหรับการคำนวณหาความถี่ที่คาดหวังเหมือนกับการหาความถี่ที่คาดหวังในสถิติทดสอบไคสแควร์นั่นเอง ค่าสัมประสิทธิ์แคปป่าจะมีค่า 0 ถึง 1

28 การวัดผลการศึกษามมส

ลักษณะของข้อมูลเมื่อใช้สัมประสิทธิ์แคปป่านี้จะทำนองเดียวกับกรณีที่เป็นการให้คะแนนแบบทวิ เพียงแต่มีจำนวนคอลัมน์และแถวเพิ่มขึ้น ดังตัวอย่างการหาค่าสัมประสิทธิ์แคปป่าเมื่อมีผู้ประเมินสองคนให้คะแนนเป็น 3 กลุ่ม ดังนี้

		ผู้ประเมินคนที่ 1		
		สูง	กลาง	ต่ำ
ผู้ประเมินคนที่ 2	สูง	5 (2.4)	3	4
	กลาง	0	7 (4)	3
	ต่ำ	0	0	3 (1.2)

จากตัวอย่างความถี่ที่คาดหวังในแต่ละเลขแยงมุมคำนวณเหมือนกับการหาความถี่ที่คาดหวังในสถิติทดสอบไคสแควร์ ได้ดังนี้

$$f_{e11} = \frac{12 \times 5}{25} = 2.4$$

$$f_{e22} = \frac{10 \times 10}{25} = 4$$

$$f_{e33} = \frac{3 \times 10}{25} = 1.2$$

เมื่อกำหนดจากสูตร (6) จะได้ดังนี้

$$\begin{aligned} K &= \frac{(5 + 7 + 3) - (2.4 + 4 + 1.2)}{25 - (2.4 + 4 + 1.2)} \\ &= \frac{15 - 7.6}{25 - 7.6} \\ &= 0.43 \end{aligned}$$

นั่นคือ มีความสอดคล้องกันระหว่างผู้ประเมิน 0.43

ข้อดีของสัมประสิทธิ์แคปป่าก็คือ สามารถแปลความหมายเป็นร้อยละของความสอดคล้องกันระหว่างผู้ประเมินได้ทำนองเดียวกับสัมประสิทธิ์ Jaccard

สรุป

จากการคำนวณหาความเชื่อมั่นหรือความสอดคล้องระหว่างผู้ประเมินสองคนเมื่อมีการให้คะแนนแบบทวิจะมีอยู่หลายวิธี แต่ละวิธีต่างก็ให้ค่าความเชื่อมั่นที่แตกต่างกัน บางวิธีให้ค่าสูงจากค่าอื่น ๆ จะเลือกใช้วิธีใดก็พิจารณาให้เหมาะสมกับลักษณะของข้อมูลด้วย

เอกสารอ้างอิง

Barret, Paul. (1999, November). Assessing the Reliability of Rating Data. (online).
Available : <http://www.liv.ac.uk/~pbarrett/programs.htm>

