

วิธีวิเคราะห์ข้อสอบแบบเลือกตอบ (ITEM ANALYSIS)

เดชา ร่มไทรย์*
ส่องสี กุลพงษ์*

1. บทนำ

วิธีวิเคราะห์ข้อสอบแบบเลือกตอบ (multiple choice question) รายข้อนี้เขียนขึ้น เพื่อเป็นแนวทางแก่อาจารย์ผู้สอนที่จะใช้วิธีวิเคราะห์ข้อสอบแบบเลือกตอบ ข้อสอบชนิดนี้เป็นที่นิยมใช้กันมากในคณะเทคนิคการแพทย์ มหาวิทยาลัยเชียงใหม่ เพราะเป็นข้อสอบที่มีประสิทธิภาพในการวัดดีกว่าข้อสอบแบบอื่นๆ ทั้งยังเป็นข้อสอบที่สามารถตรวจให้คะแนนง่าย ถูกต้อง และรวดเร็ว

วิธีวิเคราะห์ที่กล่าวถึงในที่นี้ เป็นวิธีวิเคราะห์แบบง่ายๆ เฉพาะที่คิดว่ามีความจำเป็น เป็นวิธีที่อาจารย์ผู้สอนแต่ละท่านสามารถที่จะนำไปใช้ปฏิบัติได้ เพื่อประโยชน์ในการประเมินผลการเรียน การสอน ให้มีประสิทธิภาพดียิ่งขึ้น

2. วัตถุประสงค์ของการวิเคราะห์ข้อสอบ

การวิเคราะห์ข้อสอบรายข้อนี้ เพื่อต้องการทราบ

1. ระดับความยาก และ
2. อำนาจการจำแนก

2.1 ระดับความยาก (difficulty) ใช้อักษรย่อ P คือเปอร์เซ็นต์ของนักเรียนที่ตอบข้อนั้นๆ ถูก ต่อจำนวนนักเรียนที่เข้าสอบทั้งหมด ดังนั้น ระดับความยาก 20-30 % ก็หมายความว่าข้อสอบข้อนั้น ยากมาก ระดับความยากที่ใกล้เคียง 50% คือข้อสอบที่มีความยากปานกลาง และระดับความยากที่เกิน 70% ก็คือ ข้อสอบที่ง่าย

2.2 อำนาจการจำแนก (discrimination) คือความสามารถของข้อสอบที่จะจำแนกนักเรียนออกได้เป็นนักเรียนที่เรียนเก่ง และนักเรียนที่เรียนไม่เก่งคำนวณได้จากการนำจำนวนคนตอบถูกในกลุ่มนักเรียนที่ได้คะแนนน้อย ไปลบออกจากจำนวนคนตอบถูกในกลุ่มนักเรียนที่ได้คะแนนมาก (การแบ่งกลุ่มนักเรียนจะได้กล่าวถึงต่อไป) แล้วหารด้วยจำนวนนักเรียนทั้งหมด ในกลุ่มใดกลุ่มหนึ่ง (จำนวนนักเรียนในทั้ง 2 กลุ่มเท่ากัน)

ดังนั้นค่าอำนาจการจำแนก (ใช้อักษรย่อ r) จึงมีค่าตั้งแต่ -1 ถึง +1 ถ้า r มีค่าเป็นลบแสดงว่าจำนวนนักเรียนที่ไม่เก่งตอบคำถามนั้นได้ถูกมากกว่านักเรียนที่เรียนเก่ง หรือถ้า r มีค่าเป็น 0 แสดงว่านักเรียนเก่งและไม่เก่งทำข้อสอบนั้นได้เท่ากัน ในทางตรงกันข้ามถ้าข้อสอบนั้นมีค่าเป็นบวก แสดงว่าข้อสอบนั้นนักเรียนเก่งจึงตอบได้ ส่วนนักเรียนที่ไม่เก่งจะตอบไม่ได้ ถ้าข้อสอบนั้นมีค่าที่ใกล้ +1 มากเท่าใด ก็แสดงว่าข้อสอบนั้นมีอำนาจการจำแนกสูง ดังนี้ เป็นต้น

* ภาควิชาสถิติศาสตร์ไม่โครสโคปี คณะเทคนิคการแพทย์ มหาวิทยาลัยเชียงใหม่

2.3 การแปลความหมายของค่า P และ r

โดยทั่วไปแล้วค่าความยากและค่าอำนาจการจำแนกนั้นจะแปลความหมายได้ดังนี้.-

2.3.1 ค่าความยาก (P)

85 - 100	เป็นข้อสอบที่ง่ายมาก
60 - 84	เป็นข้อสอบง่าย
51 - 59	เป็นข้อสอบค่อนข้างง่าย
50	เป็นข้อสอบยากง่ายพอเหมาะ
40 - 49	เป็นข้อสอบค่อนข้างยาก
15 - 39	เป็นข้อสอบยาก
0 - 14	เป็นข้อสอบยากมาก

2.3.2 ค่าอำนาจการจำแนก (N)

1.00	เป็นข้อสอบที่มีอำนาจจำแนกดีเลิศ
0.80 - 0.99	เป็นข้อสอบที่มีอำนาจจำแนก ดีมาก
0.60 - 0.79	เป็นข้อสอบที่มีอำนาจจำแนก ดี
0.40 - 0.59	เป็นข้อสอบที่มีอำนาจจำแนก ปานกลาง
0.20 - 0.39	เป็นข้อสอบที่มีอำนาจจำแนก น้อย
0.01 - 0.19	เป็นข้อสอบที่มีอำนาจจำแนก บ้าง
0	เป็นข้อสอบที่ไม่มีอำนาจจำแนกเลย

ในการเลือกข้อสอบไปใช้นั้นควรเลือกเฉพาะข้อสอบที่มีค่า r สูงไว้เป็นหลักก่อน ส่วนค่า P นั้น จะต้องพิจารณาตามวัตถุประสงค์ ตัวอย่างเช่น

ก. การเลือกข้อสอบเพื่อใช้เป็นแบบทดสอบวัดผลสัมฤทธิ์ (classification test) คือการสอบเพื่อดูว่านักเรียนแต่ละคนมีความสามารถระดับใด ได้แก่ แบบทดสอบที่ใช้ในการสอบซ่อม หรือสอบไล่ นั้นเอง ซึ่งควรเลือกข้อสอบที่มีค่า P ดังนี้ คือ

ค่า P ตั้งแต่ 75 ขึ้นไป	ประมาณ 25%
ค่า P ตั้งแต่ 25-74	ประมาณ 50%
ค่า P ตั้งแต่ 24 ลงมา	ประมาณ 25%

หรือถ้าจะให้ค่า P ตั้งแต่ 40-60 ทุกข้อก็จะดี ทั้งนี้ย่อมแล้วแต่การพิจารณาของผู้สอนเป็นเกณฑ์ด้วย แต่พึงระลึกไว้เสมอว่า ถึงแม้ข้อสอบค่อนข้างง่าย แต่ถ้ามีอำนาจจำแนกสูงแล้ว ย่อมดีกว่าข้อสอบที่ยาก แต่ไม่มีอำนาจจำแนก

ข. การเลือกข้อสอบเพื่อใช้เป็นแบบทดสอบวัดผลเพื่อคัดเลือกบุคคล (Selection test) เช่น การสอบเพื่อคัดเลือกนักเรียนเข้าเรียนในมหาวิทยาลัย เป็นต้น ควรเลือกข้อสอบที่ค่อนข้างยาก คือ มีค่า P ดังนี้

ปีที่ 15 ฉบับที่ 1 มกราคม 2525

ค่า P ตั้งแต่ 40-49 ประมาณ 25 %

ค่า P ตั้งแต่ 15-30 ประมาณ 50 %

ค่า P ตั้งแต่ 5-14 ประมาณ 25 %

3. ลำดับขั้นของการวิเคราะห์ข้อสอบ

3.1 ตรวจให้คะแนนแบบทดสอบทั้งชุดก่อน

3.2 เรียงกระดาษคำตอบจากนักเรียนที่ได้คะแนนสูงที่สุดมาต่ำสุด

3.3 แบ่งกระดาษคำตอบเป็นกลุ่มดังนี้

3.3.1 ถ้านักเรียนไม่เกิน 50 คน ให้แบ่งเป็น 2 กลุ่ม โดยแบ่งกลุ่มที่ได้คะแนนสูง (H) และกลุ่มที่ได้คะแนนต่ำ (L) ให้มีจำนวนเท่ากัน เช่น นักเรียน 40 คน ก็แบ่งกลุ่มละ 20 คน แต่ถ้านักเรียน 41 คน ก็คัดกระดาษคำตอบของคนที่อยู่กลางออก 1 คน แล้วจึงแบ่งเป็นกลุ่มละ 20 คน

3.3.2 ถ้านักเรียนเกิน 50 คน ให้แบ่งเป็น 3 กลุ่ม คือกลุ่มที่มีคะแนนสูง (H) 27 % ของจำนวนกระดาษคำตอบ และกลุ่มที่มีคะแนนต่ำ (L) 27 % ของจำนวนกระดาษคำตอบ ดังนั้น จำนวนกระดาษคำตอบกลุ่ม H และ L จะเท่ากัน ส่วนกลุ่มที่มีคะแนนกลางนั้น ไม่ต้องนำมาใช้วิเคราะห์

3.4 นำกระดาษคำตอบของนักเรียนในกลุ่ม H และ L มาบันทึกในตาราง ดังตัวอย่างที่ 1

	a	b	c	d	e	0
กลุ่ม H (20 คน)	////	//// ///		//		
กลุ่ม L (20 คน)	////	///	////	//// //		//

ตัวอย่างที่ 1 การบันทึกคำตอบของนักเรียนในกลุ่มที่ได้คะแนนสูง และกลุ่มที่ได้คะแนนต่ำ : ตัวอักษร a, b, c, d, e คือ ตัวเลือกของข้อสอบข้อนี้ 0 คือ จำนวนนักเรียนที่ไม่ได้ตอบข้อสอบข้อนั้น ตัวเลือกที่ถูก คือ ข้อ b

3.5 นำผลที่ได้ไปวิเคราะห์ ดังนี้

3.5.1 ค่าความยาก (P) ใช้สูตรคำนวณดังนี้

$$P = \frac{(H + L)}{N} \times 100$$

P คือ ค่าความยาก

H คือ จำนวนคนในกลุ่ม H ที่ตอบถูก

L คือ จำนวนคนในกลุ่ม L ที่ตอบถูก

N คือ จำนวนกระดาษคำตอบที่นำมาวิเคราะห์ทั้งหมด

จากตัวอย่างที่ 1 จะคำนวณค่า P ได้ดังนี้

$$P = (13 + 3) \frac{100}{40} \\ = 40 \%$$

3.5.2 ค่าอำนาจจำแนก (r) ใช้สูตรคำนวณดังนี้

$$r = \frac{2(H - L)}{N}$$

r คืออำนาจการจำแนก

H คือจำนวนคนในกลุ่ม H ที่ตอบถูก

L คือจำนวนคนในกลุ่ม L ที่ตอบถูก

N คือจำนวนกระดาษคำตอบที่นำมาวิเคราะห์ทั้งหมด

จากตัวอย่างที่ 1 จะคำนวณค่า r ได้ดังนี้

$$r = \frac{2(13 - 3)}{40} \\ = + 0.5$$

3.5.3 การพิจารณาตัวलग

เราอาจพิจารณาตัวलगได้จากตารางวิเคราะห์ข้อสอบ เช่น จากตัวอย่างที่ 1 นั้น จะเห็นได้ว่าตัวलगข้อ a นั้นเป็นตัวलगที่ไม่ดี เนื่องจากว่าตัวलगข้อ a นี้ लगนักเรียนเก่งได้ถึง 5 คน แต่กลับलगนักเรียนไม่เก่งได้เพียง 4 คนเท่านั้น ตัวलगข้อ c นั้นดี เนื่องจากलगนักเรียนเก่งไม่ได้เลย แต่लगนักเรียนไม่เก่งได้ถึง 4 คน นอกจากนั้นตัวलगข้อ e ก็ไม่ดีนัก เนื่องจากलगนักเรียนไม่ได้เลย ดังนี้เป็นต้น

ก. ค่า P ของตัวलगอาจใช้สูตรคำนวณ เช่นเดียวกับการคำนวณค่า P ของตัวเลือก (ข้อที่ถูก) ค่า P ของตัวलगที่ดีจะอยู่ในช่วง 5%-50% จากตัวอย่างที่ 1 ค่า P ของตัวलगข้อ a คือ 22.50% ข้อ c คือ 10% ข้อ d คือ 22.50 % และ ข้อ e คือ 0 จะเห็นได้ว่าถ้าพิจารณาค่า P ของตัวलगในข้อนี้แล้ว ข้อที่ต้องปรับปรุง คือ ข้อ e.

ข. ค่า r ของตัวलगใช้สูตรดังนี้

$$r = \frac{2(L - H)}{N}$$

(ความหมายของอักษรย่อต่างที่ได้กล่าวแล้วในสูตรคำนวณค่า r ของตัวलगข้อ 3.5.2)

ตัวลวงที่ดีต้องเป็นบวก ตัวลวงที่มีค่าเป็น 0 หรือเป็นลบ แสดงว่าเป็นตัวลวงที่ต้องปรับปรุงจากตัวอย่างที่ 1 ค่า r ของตัวลวง ข้อ a คือ -0.05 ข้อ c คือ +0.2 ข้อ d คือ +0.25 ข้อ e คือ 0 จะเห็นได้ว่า ถ้าพิจารณาค่า r ของตัวลวงข้อนี้แล้ว ข้อที่ต้องแก้ไขปรับปรุงคือข้อ a และข้อ e

4. การบันทึกผลการวิเคราะห์ข้อสอบ

4.1 การบันทึกผลการวิเคราะห์ข้อสอบรายข้อ

จากตัวอย่างที่ 1 ดังกล่าวมาแล้ว เมื่อนำมาคำนวณค่า P และ r แล้วก็นำมาบันทึกไว้เพื่อพิจารณาปรับปรุง ดังตัวอย่างที่ 2

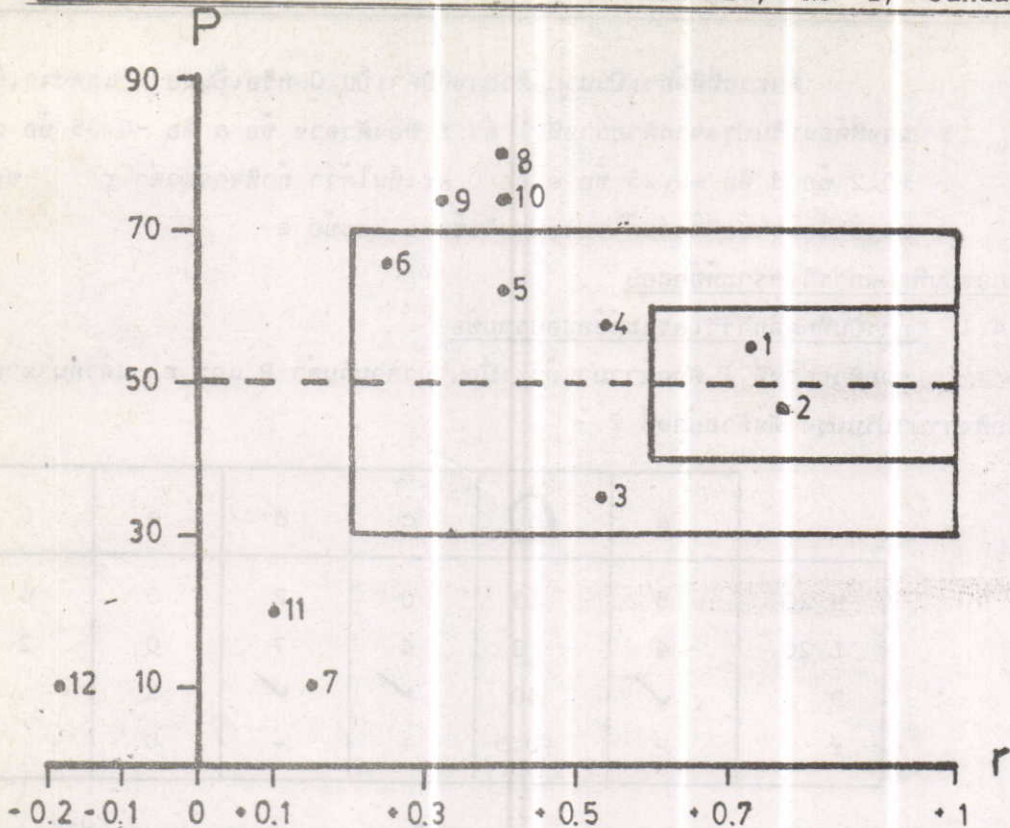
	a	b	c	d	e	0
H 20	5	13	0	2	0	0
L 20	4	3	4	7	0	2
P	✓	40	✓	✓	X	
r	-	+0.5	+	+	0	

ตัวอย่างที่ 2 การบันทึกผลการวิเคราะห์ข้อสอบ : H20 คือ นักเรียนที่ได้คะแนนสูง 20 คน และ L20 ก็คือนักเรียนที่ได้คะแนนต่ำ 20 คน ข้อถูกคือ ข้อ b, ค่า P ของข้อสอบข้อนี้คือ 40, เครื่องหมาย X ในช่อง P แสดงว่าตัวลวงที่ไม่ดี, เครื่องหมาย ✓ ในช่อง P แสดงว่าตัวลวงนั้นๆ มีค่า P ระหว่าง 5%-50% ค่า r ของข้อสอบนี้ คือ +0.5 เครื่องหมาย - หรือ 0 หรือ + แสดงว่าอำนาจการจำแนกของผู้ตอบตัวลวงนั้นมีค่า - หรือ 0 หรือ + ตามลำดับ

จากตัวอย่างที่ 2 นี้ แสดงว่าข้อสอบข้อนี้ค่อนข้างยาก ($P = 40$) อำนาจการจำแนกปานกลาง ($r = 0.5$) ตัวลวง ข้อ a และ e ควรปรับปรุงแก้ไขใหม่,

4.2 การบันทึกผลการวิเคราะห์ข้อสอบทั้งชุด

เราอาจใช้กระดาษกราฟธรรมดา เพื่อบันทึกผลการวิเคราะห์ข้อสอบแต่ละชุดได้ ดังตัวอย่างที่ 3



ตัวอย่างที่ 3 การบันทึกผลการวิเคราะห์ข้อสอบทั้งชุด : ข้อสอบชุดนี้มีอยู่ 12 ข้อ ตัวเลขที่เขียนข้างจุดกราฟ คือหมายเลขข้อสอบ ข้อสอบที่ดีที่สุดในกรอบสี่เหลี่ยมเล็กและข้อสอบที่ใช้ได้อยู่ในกรอบสี่เหลี่ยมใหญ่

ข้อสอบที่มีคุณภาพดีนั้นจะต้องมีค่า P ใกล้เคียง 50 และ r ยิ่งสูงยิ่งดี ดังนั้นจากตัวอย่างที่ 3 จะเห็นว่าข้อสอบที่อยู่ในกรอบสี่เหลี่ยมเล็กนั้นเป็นข้อสอบที่ดีที่สุดได้แก่ ข้อสอบที่ 1 และข้อที่ 2 ข้อสอบที่มีคุณภาพดีรองลงมา คือข้อที่ 3, 4, 5 และ 6 ส่วนข้อที่ 7 นั้นยากมาก และอ่านาจจำแนกน้อย : ข้อสอบ ข้อที่ 8, 9 และ 10 เป็นข้อสอบที่ไม่ดี เพราะเป็นข้อสอบที่ง่าย และอ่านาจการจำแนกน้อย ข้อที่ 11 เป็นข้อสอบที่ยาก และไม่มีอ่านาจการจำแนก ส่วนข้อสอบที่ 12 นั้น เป็นข้อสอบที่ใช้ไม่ได้

ข้อสอบที่มีความยากมากนั้น ถ้าพิจารณาจากการคำนวณค่า P และ r แล้วจะเห็นได้ว่า ถ้าค่า P น้อย(ยาก) จะเป็นไปได้ยากที่ค่า r จะใกล้ +1 เนื่องจากข้อที่ยากมากๆ นั้น นักเรียน กลุ่ม H ก็จะต้องถูกเพียงไม่กี่คน ในขณะที่นักเรียนกลุ่ม L ตอบไม่ได้เลย เช่น ตัวอย่างที่ 4 ในกรณีนี้ค่า r จะน้อย ถ้าแปลความหมายดังที่กล่าวมาแล้วก็จะได้ว่า ข้อสอบข้อนี้มีอ่านาจจำแนกน้อย แต่อย่างไรก็ตามข้อสอบข้อนี้อาจจะดีในกรณีที่มีจุดประสงค์ว่านักเรียนที่เรียนเก่งจริงๆ เท่านั้นจึงตอบได้ ในทางกลับกัน ข้อสอบที่ง่ายก็เป็นไปได้ยากที่จะให้ค่า r ใกล้เคียง 1 ทั้งนี้เพราะข้อสอบที่ง่ายนั้นทั้งนักเรียนเก่งและนักเรียนไม่เก่งก็อาจตอบได้ ข้อสอบชนิดนี้ก็เหมาะสมกับเนื้อหาที่นักเรียนทุกคนจำเป็นต้องรู้ ถ้าไม่รู้เรื่องที่ถามในข้อนี้ แสดงว่าไม่สมควรที่จะผ่านการสอบ ดังนั้น นอกจากหลักการที่ว่าต้องเลือกข้อสอบที่มีค่า r สูงๆ ไว้ก่อนนั้น จึงต้องใช้ดุลพินิจของอาจารย์ ผู้

สอนเป็นเกณฑ์ในการเลือกอีกด้วย ดังนั้นการคำนวณค่า P และ r รวมทั้งการวิเคราะห์ดังกล่าว และการบันทึกผลข้อสอบทั้งชุดนั้น จึงเป็นเพียงเครื่องมือที่ใช้วัดลักษณะของข้อสอบเท่านั้น ส่วนการตัดสินใจว่าดี หรือไม่ดี ควรเลือกไปใช้ หรือไม่นั้น ต้องอยู่ในดุลพินิจของอาจารย์ผู้สอน

	a	b	c	d	e	0
H 25	5	2	5	5	8	0
L 25	6	0	8	6	5	0
P	✓	4	✓	✓	✓	
r	+	0.08	+	+	-	

ตัวอย่างที่ 4 ข้อสอบที่มีความยากและอำนาจการจำแนกน้อย

4.3 ธนาคารข้อสอบ (item bank)

เมื่อบันทึกผลการวิเคราะห์ข้อสอบทั้งชุดแล้ว เราก็สามารถที่จะเลือกข้อสอบที่มีคุณภาพ ดีมาเก็บรวบรวมไว้เพื่อใช้ในการสอบคราวต่อไปได้ โดยนำข้อสอบที่เห็นว่ามีคุณภาพดีตามเกณฑ์ที่ได้กล่าวมาแล้วข้างต้นมาบันทึกรายละเอียดทั้งหมดเก็บไว้เป็นต้นว่า ค่า P และค่า r ของการนำไปใช้แต่ละครั้ง, จำนวนนักเรียนที่สอบ, ชั้น, ปีที่สอบ ฯลฯ.

กล่าวโดยสรุปแล้วการวิเคราะห์ข้อสอบนี้จะเป็นสิ่งที่ช่วยให้อาจารย์ผู้สอนสามารถที่จะปรับปรุงข้อสอบของตนให้มีคุณภาพดียิ่งขึ้น และยังสามารถที่จะสะสมข้อสอบที่ดีไว้ใช้ในการสอบคราวต่อไป ซึ่งจะทำให้เกิดความสะดวกและรวดเร็วในการออกข้อสอบ ทั้งยังเป็นการเพิ่มทักษะในการออกข้อสอบของอาจารย์ผู้สอนอีกด้วย

เอกสารอ้างอิง

1. บุญเชิด ภิญโญนนท์พงษ์, การวัดและประเมินผลการศึกษา, คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ, 2519, หน้า 140-163.
2. อนันต์ ศรีโสภะ, การวัดและประเมินผลการศึกษา, บริษัทสำนักพิมพ์ไทยวัฒนาพานิช จำกัด, 2520, หน้า 148-165.
3. ชุมพร ยงกิตติกุล, การสร้างเครื่องมือประเมินผลการเรียน, ใน "คู่มืออาจารย์ ด้านการเรียนการสอน", โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, บรรณาธิการ ไพฑูรย์ สินลารัตน์, 2520 หน้า 237.