

การวิเคราะห์ข้อสอบอิงเกณฑ์

ฉัตรศิริ ปิยะพิมลสิทธิ์*

การวิเคราะห์อิงเกณฑ์คล้ายกับการวิเคราะห์อิงกลุ่ม ก็คือมีการวิเคราะห์ข้อสอบเป็นรายข้อ โดยตามแนวคิดของอิงเกณฑ์นั้นจะต้องจำแนกผู้สอบออกเป็นกลุ่มผู้รอบรู้และไม่รอบรู้ในจุดประสงค์ต่าง ๆ ได้อย่างถูกต้อง ดังนั้นการวิเคราะห์รายข้อจึงเน้นที่อำนาจจำแนก ส่วนการวิเคราะห์ทั้งฉบับนั้น ในด้านของความเที่ยงตรง จะต้องสามารถวัดเนื้อหาและพฤติกรรมได้ถูกต้องตามจุดประสงค์ที่ตั้งไว้ นั่นคือจะต้องมีความเที่ยงตรงตามเนื้อหาและแบบทดสอบสามารถจำแนกผู้สอบออกเป็นผู้รอบรู้และไม่รอบรู้ได้ถูกต้องตามความสามารถที่แท้จริงเพียงไร นั่นคือจะต้องมีความเที่ยงตรงตามสภาพ ส่วนความเชื่อมั่นนั้นก็แน่นอนว่าจะต้องหาเหมือนแบบทดสอบอิงกลุ่ม นั่นคือแบบทดสอบสามารถสอบวัดได้โดยปราศจากความคลาดเคลื่อนมากนักน้อยเพียงไร

การวิเคราะห์คุณภาพรายข้อ

อำนาจจำแนกของข้อสอบอิงเกณฑ์ คือประสิทธิภาพของข้อสอบแต่ละข้อที่สามารถจำแนกกลุ่มผู้เรียนออกเป็นกลุ่มรอบรู้และไม่รอบรู้ วิธีการคำนวณหาอำนาจจำแนกมีอยู่หลายวิธี ดังจะได้นำเสนอต่อไปนี้

1. ดัชนี B ของเบรนนอน (Brennan)

เบรนนอน ได้พัฒนาสูตรอำนาจจำแนกเรียกว่า Discrimination Index B มีสูตรดังนี้

$$B = \frac{H}{N_H} - \frac{L}{N_L}$$

เมื่อ B คือดัชนีค่าอำนาจจำแนก

H คือจำนวนนักเรียนในกลุ่มผู้รอบรู้ที่ตอบข้อสอบข้อนั้นถูก

L คือจำนวนนักเรียนในกลุ่มไม่รอบรู้ที่ตอบข้อสอบข้อนั้นถูก

N_H คือจำนวนนักเรียนในกลุ่มรอบรู้ทั้งหมด

N_L คือจำนวนนักเรียนในกลุ่มไม่รอบรู้ทั้งหมด

* กศ.ม. (การวัดผลการศึกษา) <http://www.watpon.com/>

การแบ่งกลุ่มผู้เรียนออกเป็นกลุ่มรอบรู้และไม่รอบรู้ให้ยึดคะแนนจุดตัดเป็นหลัก และแน่นอนว่าจำนวนนักเรียนในกลุ่มรอบรู้และไม่รอบรู้อาจไม่เท่ากันได้

2. วิธีของคอกซ์ และวากัส (Cox and Vargas)

เป็นการหาอำนาจจำแนกโดยอาศัยการสอบก่อนสอนและหลังสอน โดยการคำนวณหาความแตกต่างระหว่างสัดส่วนของนักเรียนที่ตอบถูกก่อนสอนและตอบถูกหลังสอน โดยเรียกว่าดัชนี Dpp มีสูตรดังนี้

$$Dpp = p_{.1} - p_{1.}$$

เมื่อ Dpp คือดัชนีค่าอำนาจจำแนก
 $p_{.1}$ คือสัดส่วนของผู้ตอบถูกหลังสอน
 $p_{1.}$ คือสัดส่วนของผู้ตอบถูกก่อนสอน

แต่สูตรนี้จะมีค่าแปรเปลี่ยนไปตามความรู้เดิมของกลุ่มผู้เรียน ถ้าผู้เรียนมีความรู้เดิมมากแน่นอนว่าค่า Dpp ย่อมมีค่าต่ำ แต่ถ้ามีความรู้เดิมน้อย ค่า Dpp ย่อมมีค่าสูง ต่อมา Herbig ได้พัฒนาโดยปรับแก้อิทธิพลของความรู้เดิมที่มีมาก่อนสอน ได้สูตรใหม่ดังนี้

$$Dpp = \frac{P_{01} - P_{10}}{P_{0.} + P_{.0}}$$

เมื่อ Dpp คือดัชนีค่าอำนาจจำแนก
 p_{01} คือสัดส่วนของผู้ตอบผิดก่อนสอนแต่หลังสอนตอบถูก
 p_{10} คือสัดส่วนของผู้ตอบถูกก่อนสอนแต่หลังสอนตอบผิด
 $p_{0.}$ คือสัดส่วนของผู้ตอบผิดก่อนสอน
 $p_{.0}$ คือสัดส่วนของผู้ตอบผิดหลังสอน

3. ดัชนีความไว (Sensitivity Index)

เป็นการวิเคราะห์โดยใช้ความแตกต่างระหว่างการสอบก่อนสอนและการสอบหลังสอน มีสูตรในการคำนวณดังนี้

$$S = \frac{R_{post} - R_{pre}}{N}$$

เมื่อ S คือดัชนีความไว
 R_{post} คือจำนวนผู้ตอบถูกหลังสอน
 R_{pre} คือจำนวนผู้ตอบถูกก่อนสอน
 N คือจำนวนนักเรียนทั้งหมด

4. วิธีของป๊อปแฮม (Popham)

1) Pretest-Posttest Differences

$$D_{ppd} = P_{post} - P_{pre}$$

เมื่อ D_{ppd} คือความแตกต่างของกลุ่มก่อนสอนและกลุ่มหลังสอน
 P_{post} คือสัดส่วนของผู้ตอบถูกหลังสอน
 P_{pre} คือสัดส่วนของผู้ตอบถูกก่อนสอน

2) Uninstructed versus instructed group differences

$$D_{uigd} = P_i - P_u$$

เมื่อ D_{uigd} คืออำนาจจำแนกของกลุ่มได้รับการสอนกับกลุ่มไม่ได้รับการสอน
 P_i คือสัดส่วนการตอบถูกของผู้ได้รับการสอน
 P_u คือสัดส่วนการตอบถูกของผู้ไม่ได้รับการสอน

การแปลความหมายค่าอำนาจจำแนก

ค่าอำนาจจำแนกแบบอิงเกณฑ์จะมีค่าอยู่ระหว่าง -1.00 ถึง 1.00 เหมือนกับค่าอำนาจจำแนกแบบอิงกลุ่ม และมีการแปลความหมายคล้ายคลึงกันดังนี้

ค่าอำนาจจำแนก	การแปลความหมาย
1.00	จำแนกผู้รอบรู้และไม่รอบรู้ได้ถูกต้องทุกคน
0.50 – 0.99	จำแนกผู้รอบรู้และไม่รอบรู้ได้ถูกต้องเป็นส่วนใหญ่
0.20 – 0.49	จำแนกผู้รอบรู้และไม่รอบรู้ได้ถูกต้องบางส่วน
0.00 – 0.19	จำแนกผู้รอบรู้และไม่รอบรู้ได้ถูกต้องน้อยมากหรือไม่จำแนก

ค่าอำนาจจำแนกที่ติดลบ แสดงว่าข้อสอบจำแนกผู้รอบรู้และไม่รอบรู้ได้ในทางกลับกัน อาจหมายถึงข้อสอบง่ายหรือยากจนเกินไป ผู้เรียนรู้อาแล้ว หรือผู้สอนสอนไม่ตรงตามจุดประสงค์ที่ตั้งไว้ หรือข้อสอบไม่ชัดเจน พิมพ์ผิด เกลยผิด ตรวจผิด เป็นต้น

เกณฑ์ของค่าอำนาจจำแนกที่ยอมรับว่าข้อสอบนั้นมีคุณภาพสามารถจำแนกได้ก็คือ 0.20 ขึ้นไป

การหาคุณภาพทั้งฉบับด้านความเที่ยงตรง

แบบทดสอบอิงเกณฑ์จะต้องสามารถวัดเนื้อหาและพฤติกรรมได้ถูกต้องตามที่ระบุไว้ในจุดประสงค์ และสามารถจำแนกผู้รอบรู้และไม่รอบรู้ในจุดประสงค์นั้นได้อย่างถูกต้อง จึงเน้นที่ความเที่ยงตรงเพียง 2 ชนิดคือ ความเที่ยงตรงตามเนื้อหาและความเที่ยงตรงตามสภาพ

1. ความเที่ยงตรงตามเนื้อหา (Content Validity)

เป็นการตรวจดูว่าข้อสอบที่สร้างขึ้นวัดได้สอดคล้องกับจุดประสงค์เชิงพฤติกรรมหรือไม่ และจุดประสงค์เชิงพฤติกรรมเหล่านั้นชัดเจน ครอบคลุมและสอดคล้องกับเนื้อหาหรือไม่ คำถามในแบบทดสอบเหล่านั้นสามารถเป็นตัวแทนของพฤติกรรมที่ต้องการวัดได้ดีเพียงไร

วิธีการตรวจสอบความเที่ยงตรงตามเนื้อหาของข้อสอบอิงเกณฑ์ ที่ใช้กันมากคือ การให้ผู้เชี่ยวชาญทางเนื้อหาวิชาหลายคนเป็นผู้พิจารณา แล้วตอบคำถาม 2 ข้อคือ

1. จุดประสงค์เชิงพฤติกรรมหรือขอบเขตของพฤติกรรมเขียนไว้ชัดเจนและครอบคลุมเนื้อหาหรือไม่ อาจทำได้โดยการสร้างตารางกำหนดรายละเอียด แล้วให้ผู้เชี่ยวชาญหลายคนพิจารณาให้คะแนน

+1 แน่ใจว่าจุดประสงค์เชิงพฤติกรรมหรือขอบเขตนั้นชัดเจนและครอบคลุมเนื้อหา

0 ไม่แน่ใจว่าจุดประสงค์เชิงพฤติกรรมหรือขอบเขตนั้นชัดเจนและครอบคลุมเนื้อหา

-1 แน่ใจว่าจุดประสงค์เชิงพฤติกรรมหรือขอบเขตนั้นไม่ชัดเจนและครอบคลุมเนื้อหา

แล้วนำผลที่ได้มาหาคะแนนเฉลี่ยในแต่ละจุดประสงค์ หากได้ 0.5 ขึ้นไป ถือว่าจุดประสงค์เชิงพฤติกรรมหรือขอบเขตเนื้อหาที่กำหนดไว้ชัดเจนและครอบคลุมเนื้อหาดีแล้ว แต่ถ้าไม่ถึง 0.5 ต้องปรับปรุงใหม่

2. ข้อสอบแต่ละข้อวัดได้ตรงตามจุดประสงค์เชิงพฤติกรรมหรือเป็นตัวแทนของขอบเขตเนื้อหานั้นหรือไม่ ซึ่งมีหลายวิธี เช่น

2.1 วิธีหาความสอดคล้องระหว่างข้อสอบกับจุดประสงค์ Rovinelli และ Hambleton ได้เสนอให้นำจุดประสงค์เชิงพฤติกรรมและข้อสอบทั้งหลายที่วัดจุดประสงค์นั้น ไปให้ผู้เชี่ยวชาญทางเนื้อหาหลายคนพิจารณาแล้วให้คะแนน

+1 แน่ใจว่าข้อสอบวัดจุดประสงค์นั้น

0 ไม่แน่ใจว่าข้อสอบวัดจุดประสงค์นั้น

-1 แน่ใจว่าข้อสอบไม่วัดจุดประสงค์นั้น

แล้วหาค่าเฉลี่ยในแต่ละข้อ เรียกค่านั้นว่า “ดัชนีความสอดคล้อง” ถ้าข้อใดมีดัชนีความสอดคล้องมากกว่าหรือเท่ากับ 0.5 ถือว่าข้อนั้นวัดจุดประสงค์นั้นจริง มีความเที่ยงตรงตามเนื้อหา แต่ถ้าต่ำกว่า 0.5 ให้ตัดออกหรือแก้ไขปรับปรุงแล้วนำมาให้ผู้เชี่ยวชาญพิจารณาใหม่

2.2 วิธีจับคู่ระหว่างข้อสอบกับจุดประสงค์ซึ่ง Rovinelli และ Hambleton ได้เสนอไว้โดยให้ผู้เชี่ยวชาญจับคู่ระหว่างจุดประสงค์กับข้อสอบว่าข้อสอบข้อนั้นวัดจุดประสงค์ใด แล้วนำคำตอบของผู้เชี่ยวชาญแต่ละคนมาแจกแจงความถี่

ข้อสอบ	จุดประสงค์					
	1		2		3	
	f	%	f	%	f	%
1						
2						
3						
:						

จากนั้นสังเกตเปอร์เซ็นต์การแจกแจงความถี่ โดยใช้เปอร์เซ็นต์ความถี่ 80% เป็นเกณฑ์ตัดสินยอมรับว่าข้อสอบวัดจุดประสงค์นั้น ส่วนข้อที่ไม่ถึง 80% ให้ตัดทิ้งหรือปรับปรุงใหม่

2.3 วิธีจัดอันดับความสอดคล้องระหว่างข้อสอบกับขอบเขตเนื้อหา ซึ่ง Hambleton และคณะ ได้เสนอให้ใช้ผู้เชี่ยวชาญทางเนื้อหาวิชา 2 คน จัดอันดับความสอดคล้องระหว่างข้อสอบกับขอบเขตเนื้อหา โดยใช้มาตราส่วนประมาณค่า 4 ระดับคือ

- 1 ไม่เกี่ยวข้อง
- 2 เกี่ยวข้องบ้าง
- 3 เกี่ยวข้อง
- 4 เกี่ยวข้องมาก

จากนั้นนำมาจัดทำตารางดังนี้

ผู้จัดอันดับคนที่ 2	ผู้จัดอันดับคนที่ 1		รวม
	1 หรือ 2	3 หรือ 4	
1 หรือ 2	A	B	A + B
3 หรือ 4	C	D	C + D
รวม	A + C	B + D	N

$N = A + B + C + D$

จากนั้นคำนวณหาดัชนีความเที่ยงตรงตามเนื้อหา (Content Validity Index)

$$CVI = \frac{D}{N}$$

CVI จะต้องมีค่ามากกว่า 0.50 ถือว่ามีความเที่ยงตรงตามเนื้อหา

2.4 วิธีของครอนบาค (Cronbach) เสนอวิธีให้ผู้เขียนข้อสอบ 2 กลุ่ม ที่มีความสามารถทัดเทียมกัน เขียนข้อสอบกลุ่มละ 1 ชุด วัดจุดประสงค์เดียวกัน แล้วนำข้อสอบทั้ง 2 ชุด ไปสอบนักเรียนกลุ่มหนึ่งที่เรียนรู้จุดประสงค์นั้นแล้ว แล้วนำผลการสอบมาเปรียบเทียบกัน โดยอาจใช้ t-test แล้วค่าเฉลี่ยไม่แตกต่างกัน ถือว่ามีความเที่ยงตรงตามเนื้อหา

2. ความเที่ยงตรงตามสภาพ (Concurrent Validity)

เป็นการตรวจสอบว่าข้อสอบนั้นสามารถจำแนกผู้สอบออกเป็นผู้เรียนรู้แล้วกับผู้ที่ยังไม่เรียนรู้ได้ถูกต้องเพียงใด หากผู้ที่เรียนรู้แล้วทำคะแนนได้สูงกว่าผู้ที่ยังไม่เรียนรู้ แสดงว่า ข้อสอบชุดนี้มีความเที่ยงตรงตามสภาพ

การตรวจสอบความเที่ยงตรงสภาพมี 2 วิธี ดังนี้

1. วิธีของมิลแมน (Millman) สอบนักเรียนกลุ่มเดียว แต่สอบ 2 ครั้งคือก่อนเรียน และหลังเรียน แล้วนำข้อมูลมาสร้างตารางดังนี้

	ก่อนเรียน	หลังเรียน	รวม
สอบผ่าน	B	A	A + B
สอบไม่ผ่าน	D	C	D + C
รวม	B + D	A + C	N

$$\text{ความเที่ยงตรงตามสภาพ } (\phi) = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

มิลแมน ยังได้เสนออีกวิธีหนึ่งคือ สอบครั้งเดียว แล้วแบ่งเป็นผ่านกับไม่ผ่าน แล้วให้ผู้สอนประเมินนักเรียนกลุ่มนี้ว่ามีใครบ้างที่มีความสามารถจริง ใครบ้างที่ยังไม่มีความสามารถ โดยสังเกตจากพฤติกรรมและผลงานของนักเรียน

	ไม่มีความสามารถ	มีความสามารถ	รวม
สอบผ่าน	B	A	A + B
สอบไม่ผ่าน	D	C	D + C
รวม	B + D	A + C	N

จากข้อมูลในตารางคำนวณได้ 2 วิธีคือ

- คำนวณหาค่าสหสัมพันธ์ Phi

$$\phi = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

- คำนวณหาค่าสหสัมพันธ์ Q

$$Q = \frac{ad - bc}{ad + bc}$$

2. วิธีของคาร์เวอร์ (Carver) เป็นการสอบกับนักเรียน 2 กลุ่ม คือกลุ่มที่ได้รับการสอน กับกลุ่มที่ไม่ได้รับการสอน

	ไม่ได้รับการสอน	ได้รับการสอน
สอบผ่าน	b	a
สอบไม่ผ่าน	c	d

คำนวณค่าความเที่ยงตรงตามสภาพด้วยสูตร

$$\text{ค่าความเที่ยงตรงตามสภาพ} = \frac{a + c}{a + b + c + d}$$

การหาคุณภาพทั้งฉบับด้านความเชื่อมั่น

ผลการวัดควรจะต้องมีความถูกต้องแม่นยำ ปราศจากความคลาดเคลื่อน มีผู้เสนอวิธีการคำนวณหาความเชื่อมั่นของข้อสอบอิงเกณฑ์หลายวิธีดังนี้

1. วิธีของคาร์เวอร์ (Carver) ใช้แบบทดสอบคู่ขนาน โดยให้ข้อสอบคล้ายกันข้อต่อข้อ แล้วนำไปสอบนักเรียน แล้วนำข้อมูลที่ได้มาสร้างตารางดังนี้

	ฉบับ B	
ฉบับ A	สอบไม่ผ่าน	สอบผ่าน
สอบผ่าน	b	a
สอบไม่ผ่าน	c	d

$$\text{ค่าความเชื่อมั่น} = \frac{a + c}{a + b + c + d}$$

2. วิธีของลิวิงสตัน (Livingston) ใช้การสอบครั้งเดียวหลังเรียนจบ แล้วคำนวณด้วยสูตร

$$r_{cc} = \frac{r_{tt}\sigma^2 + (\bar{X} - c)^2}{\sigma^2 + (\bar{X} - c)^2}$$

เมื่อ r_{tt} คือความเชื่อมั่นของแบบทดสอบที่คำนวณด้วยวิธีอิงกลุ่ม (เช่น KR-20, KR-21 ฯลฯ)

σ^2 คือความแปรปรวนของคะแนนสอบทั้งฉบับ

\bar{X} คือคะแนนเฉลี่ยของคะแนนสอบทั้งฉบับ

c คือคะแนนจุดตัด

วิธีนี้ได้ผลดีเมื่อคะแนนมีการกระจายแบบฐานนิยมเดียว

3. วิธีของสวามินาทานและคณะ (Swaminathan) วิธีนี้ใช้แบบทดสอบฉบับเดียวแต่สอบซ้ำ 2 ครั้งหลังจากสิ้นสุดการสอน เพื่อดูความคงเส้นคงวา

	สอบครั้งที่ 1	
สอบครั้งที่ 2	สอบผ่าน	สอบไม่ผ่าน
สอบผ่าน	a	b
สอบไม่ผ่าน	c	d

$$\text{ค่าความเชื่อมั่น (K)} = \frac{P_o - P_e}{1 - P_e}$$

$$\text{เมื่อ } P_o = \frac{a + d}{a + b + c + d}$$

$$P_e = \frac{(a + b)(a + c) + (c + d)(b + d)}{(a + b + c + d)^2}$$

4. วิธีของโลเวทท์ (Lovett) ใช้การสอบครั้งเดียวหลังเรียน มีสูตรดังนี้

$$r_{cc} = 1 - \frac{k \sum x_i - \sum x_i^2}{(k-1) \sum (x_i - c)^2}$$

เมื่อ	x_i	คือคะแนนของแต่ละคน
	k	คือจำนวนข้อสอบทั้งหมด
	c	คือคะแนนจุดตัด



บรรณานุกรม

- เซาวนา ชลิตธำรง. (2538). “การตรวจสอบคุณภาพของเครื่องมือ : อิงเกนท์”, ใน การสร้างเครื่องมือวัดผลที่ใช้ในการวิจัย. กรุงเทพฯ : ภาควิชาการวัดผลและวิจัยการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ,
- บุญเชิด ภิญโญนนท์พงษ์. (2527). การทดสอบแบบอิงเกนท์ : แนวคิดและวิธีการ. กรุงเทพฯ : โอเดียนสโตร์,