

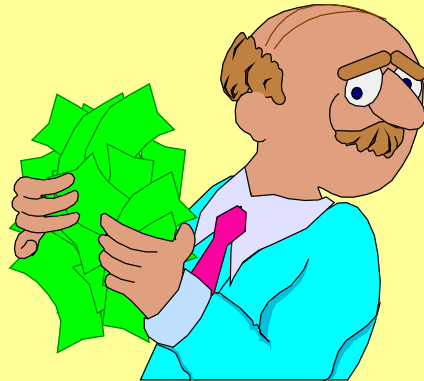
บทที่ 5

การวิเคราะห์ข้อสอบ



อาจารย์ ดร.พัชราภรณ์ พิลาสมบัติ

ความรู้เบื้องต้น
ในการวิเคราะห์ข้อสอบ



วัตถุประสงค์ของการวิเคราะห์ข้อสอบ

1. เพื่อตรวจสอบดูว่าข้อสอบใดที่เป็นข้อสอบที่ดี หรือมีจุดบกพร่อง
2. เพื่อพิจารณาหาความยากง่ายของคำถามแต่ละข้อ
3. เพื่อพิจารณาหาค่าอำนาจจำแนกของข้อสอบแต่ละข้อ
4. เพื่อพิจารณาหาความสัมพันธ์ระหว่างข้อสอบแต่ละข้อกับคะแนนรวม
5. เพื่อจัดหาข้อมูลที่เป็นประโยชน์ต่อการแปลความหมาย หรือการอภิปรายเกี่ยวกับผลการทดสอบ

ประโยชน์ของการวิเคราะห์ข้อสอบ

1. ทำให้รู้ว่าข้อสอบแต่ละข้อมีคุณภาพอย่างไร
2. ช่วยเป็นแนวทางให้ครูเขียนข้อสอบเก่งขึ้น
3. ประหยัดเวลาและแรงงานในการออกข้อสอบระยะยาว
4. เป็นจุดเริ่มต้นของการสร้างข้อสอบคู่ขนาน
5. เป็นจุดเริ่มต้นของการสร้างธนาคารข้อสอบ
6. เป็นจุดเริ่มต้นของการสร้างข้อสอบมาตรฐาน

C ราชบัณฑิต)

การวิเคราะห์ข้อสอบ (Item Analysis)

เป็นการตรวจสอบคุณภาพของข้อสอบเป็นรายข้อ โดยนำผลที่ได้จากการสอบของผู้เรียนมาวิเคราะห์ คำนึงบ่งชี้ของคุณภาพของข้อสอบ ดังนี้

1. ค่าความยาก (Difficulty index)

2. ค่าอำนาจจำแนก (Discriminant index)

1. ความยาก (Difficulty) ตามทฤษฎีการวัดผลแบบอิงกลุ่ม ข้อสอบที่ดีคือข้อสอบที่ไม่ยากเกินไป หรือไม่ง่ายเกินไป เพราะข้อสอบนั้น ๆ จะสามารถ จำแนกผู้สอบได้ว่าใครเก่งใครอ่อน ข้อสอบที่ยากหรือง่ายเกินไปจึงไม่มีคุณค่าในการจำแนก ส่วนทฤษฎีการวัดผลแบบอิงเกณฑ์นั้น ถือว่าการสร้างข้อสอบก็เพื่อที่จะวัดว่า ผู้เรียนได้บรรลุจุดประสงค์หรือไม่ การที่ทุกคนทำข้อสอบ ได้ถูก แสดงว่าเขาบรรลุจุดประสงค์ตามต้องการจะว่า ข้อสอบไม่ดีก็ไม่ได้ ดังนั้นเรื่องความยากง่ายจึงไม่ใช่เรื่อง สำคัญ สิ่งสำคัญอยู่ที่ข้อสอบนั้นได้วัดในจุดประสงค์ที่ ต้องการวัดว่าวัดได้จริงหรือไม่ ถ้าวัดได้จริงก็นับว่าเป็น ข้อสอบที่ดีได้ แม้ว่าจะจะเป็นข้อสอบที่ง่ายก็ตาม

2. อำนาจจำแนก (Discrimination) หมายถึง

ความสามารถในการจำแนกผู้สอบ ข้อสอบที่ดีจะต้องมีอำนาจจำแนกสูงตามทฤษฎีการวัดผลแบบอิงกลุ่ม (Norm Reference Measurement) อำนาจจำแนกของข้อสอบหมายถึง ความสามารถของข้อสอบในการจำแนกผู้สอบออกเป็นกลุ่มสูง/กลุ่มต่ำ ส่วนทฤษฎีการวัดผลแบบอิงเกณฑ์ (Criterion Reference Measurement) หมายถึงความสามารถของข้อสอบนั้นในการจำแนกผู้สอบออกเป็น 2 กลุ่มคือกลุ่มรอบรู้ และกลุ่มยังไม่รอบรู้

วิธีการวิเคราะห์ข้อสอบรายข้อ

1. ตรวจสอบคำตอบแล้วรวมคะแนนของกระดาษคำตอบแต่ละคน
2. เรียงกระดาษคำตอบจากคะแนนมากไปน้อย ชำกั้นเอาไว้ด้วยกัน
3. แบ่งกระดาษเป็น 2 กลุ่ม (กลุ่มสูง H และกลุ่มต่ำ L) ถ้ามีน้อยให้กลุ่มละ 50% แต่ถ้ามีมาก กลุ่มละ 27% ที่เหลืออีก 46% ไม่ต้องนำมาคิด

การกำหนดกลุ่มเก่งและกลุ่มอ่อน

สมมติว่าใช้เทคนิค 27%

และขนาดตัวอย่างเท่ากับ 200 คน

คะแนนเรียงลำดับแล้ว

และมีความถี่เป็นดังนี้

| คะแนน | จำนวน (ความถี่) |
|-------|-----------------|
| 100 | 1 |
| 99 | 3 |
| 98 | 2 |
| 97 | 4 |
| 96 | 7 |
| 95 | 10 |
| ... | |
| ... | |
| ... | |
| ... | |
| 5 | 4 |
| 4 | 3 |
| 3 | 6 |
| 2 | 2 |
| 1 | 0 |
| 0 | 1 |
| รวม | 200 |

$$100 - 27$$

$$200 \times 27 = 54$$

กลุ่มเก่ง (H) 27% = 54 คน ~~X~~

~~กลุ่มปานกลาง (M) 46% = 92 คน~~

กลุ่มอ่อน (L) 27% = 54 คน ~~X~~

รวม 100% = 200 คน

4. นับจำนวนนักเรียนที่ตอบถูกแต่ละข้อทั้งกลุ่มสูงและกลุ่มต่ำ ดังตาราง

| ข้อที่ | กลุ่มสูง (H) | | | | | กลุ่มต่ำ (L) | | | | |
|--------|--------------|---|----|---|-----|--------------|---|---|---|-----|
| | ก | ข | ค | ง | รวม | ก | ข | ค | ง | รวม |
| 1 | 0 | 0 | 20 | 0 | 20 | 4 | 8 | 2 | 6 | 20 |
| 2 | | | | | | | | | | |
| 3 | | | | | | | | | | |
| . | | | | | | | | | | |
| . | | | | | | | | | | |

5. รวมจำนวนนักเรียนที่ตอบถูกของแต่ละข้อของกลุ่มสูง และกลุ่มต่ำ หาค่าด้วยจำนวนกระดาษคำตอบของกลุ่มสูง และกลุ่มต่ำรวมกัน ค่าที่ได้จะเป็นดัชนีความยากของข้อสอบ (Index of Item Difficulty)

6. เอาจำนวนนักเรียนที่ตอบถูกในกลุ่มสูงลบด้วยจำนวนนักเรียนที่ตอบถูกในกลุ่มต่ำ และหารด้วยจำนวนกระดาษคำตอบในกลุ่มสูงหรือต่ำ ค่าที่ได้เป็นอำนาจการจำแนก (Discrimination Power) ของข้อสอบนั้น

R

สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อสอบ

$$P = \frac{H + L}{2n}$$

ค่า
0 - 1
0.2-0.8
0.4-0.6
ง่าย

P = ตัวบ่งชี้ความยากของข้อสอบ (difficulty index)

H = จำนวนผู้ตอบถูกในกลุ่มสูง (high)

L = จำนวนผู้ตอบถูกในกลุ่มต่ำ (low)

n = จำนวนผู้ตอบในกลุ่มสูงหรือกลุ่มต่ำ

P ย่อมาจากคำ Proportion หรือค่าร้อยละ (Percentage) ของการตอบถูก

P จึงมีค่าเริ่มต้นจาก 0 (0%) ถึง 1 (หรือ 100%) ดังนั้น ค่า P ยิ่งน้อยยิ่ง

ยาก ยิ่งมากยิ่งง่าย โดยค่า P ที่เหมาะสมคือ 20% ถึง 80%

สัญลักษณ์ที่ใช้ในการวิเคราะห์ข้อสอบ

$$r = \frac{H - L}{n}$$

- 1 ถึง 1
↓
ทำในกระดาษ
ทำแบบฝึก

r = อำนาจจำแนกของข้อสอบ (discrimination power)

H = จำนวนผู้ตอบถูกในกลุ่มสูง (high)

L = จำนวนผู้ตอบถูกในกลุ่มต่ำ (low)

n = จำนวนผู้ตอบในกลุ่มสูงหรือกลุ่มต่ำ

R มีค่าจาก -1 ถึง 1 โดยค่า R ยิ่งมากยิ่งดี และ
มาตรฐานขั้นต่ำของ R ที่ยอมรับได้คือ 0.20 (คือจำแนกได้ 20%
ขึ้นไป)

ตัวอย่าง การสอบวิชาภาษาไทยข้อสอบเป็นแบบ 4 ตัวเลือก

จำนวน 5 ข้อ มีผู้สอบจำนวน 50 คน ตรวจให้คะแนนและเรียงคะแนนจากสูงไปต่ำเรียบร้อยแล้ว และนำมาแบ่งเป็นกลุ่มสูง 27% (ประมาณ 14 คน) และกลุ่มต่ำ 27% (ประมาณ 14 คน) ผลการวิเคราะห์จำนวนคนที่ตอบถูกในกลุ่มสูงและกลุ่มต่ำปรากฏดังตารางต่อไปนี้

| ข้อที่ | H | L | H+L | H-L | P | r |
|---------------|-----------|-----------|------------|------------|-------------|--------------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |
| 3 | 14 | 14 | 28 | 0 | 1.00 | 0.00 |
| 4 | 6 | 10 | 16 | -4 | 0.57 | -0.28 |
| 5 | 12 | 5 | 17 | 7 | 0.61 | 0.50 |

การหาค่า P และ r เมื่อ n = 14

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|---|---|-----|-----|------|------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |

$$P = \frac{H + L}{2n}$$

$$P = \frac{8 + 8}{2 \times 14}$$

$$P = 0.57$$

$$r = \frac{H - L}{n}$$

$$r = \frac{8 - 8}{14}$$

$$r = 0.00$$

การหาค่า P และ r เมื่อ n = 14

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|---|-----|-----|------|------|
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |

$$P = \frac{H + L}{2n}$$

$$P = \frac{14 + 0}{2 \times 14}$$

$$P = 0.50$$

$$r = \frac{H - L}{n}$$

$$r = \frac{14 - 0}{14}$$

$$r = 1.00$$

0-1 -1 จำนวน +1

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|---|-----|-----|------|------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |
| 3 | | | | | | |
| 4 | | | | | | |
| 5 | | | | | | |

ข้อ 1 ทั้ง 2 กลุ่มตอบถูกเท่ากัน แปลว่า ข้อสอบข้อนี้ไม่สามารถจำแนกนักเรียนทั้งสองกลุ่มได้ แม้ว่าจะมีความยากดีมาก (0.57) เป็นข้อสอบที่ต้องปรับปรุง

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|----|-----|-----|------|------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |
| 3 | 14 | 14 | 28 | 0 | 1.00 | 0.00 |
| 4 | | | | | | 0.28 |
| 5 | | | | | | 0 |

ข้อ 2 ผู้สอบกลุ่มสูงตอบถูกหมด และกลุ่มต่ำ
 ตอบผิดหมด มีค่าความยากเท่ากับ 0.50
 แปลว่ามีค่าความยาก ~~ยาก~~ และค่าอำนาจจำแนก
 เท่ากับ 1 เป็นข้อสอบที่ดีมาก

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|----|-----|-----|---------------------|-------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |
| 3 | 14 | 14 | 28 | 0 | 1.00 | 0.00 |
| 4 | 6 | 10 | 16 | -4 | 0.57 | -0.28 |
| 5 | | | | | $p = \frac{H+L}{n}$ | |

ข้อ 3 ผู้สอบกลุ่มสูงและกลุ่มต่ำตอบถูกหมด แสดงว่า
ข้อสอบง่ายมาก และไม่สามารถจำแนกนักเรียนได้เป็น
ข้อสอบที่ต้องตัดทิ้ง

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|----|-----|-----|------|-------|
| 1 | 8 | 8 | 16 | 0 | 0.57 | 0.00 |
| 2 | 14 | 0 | 14 | 14 | 0.50 | 1.00 |
| 3 | 14 | 14 | 28 | 0 | 1.00 | 0.00 |
| 4 | 6 | 10 | 16 | -4 | 0.57 | -0.28 |
| 5 | | | | | | |

ข้อ 4 ข้อสอบข้อนี้มีความยากเท่ากับ 0.57 น่าจะ
ใช้ได้ แต่ค่าอำนาจจำแนกติดลบ นั่นคือ คนกลุ่มสูง
ตอบไม่บ่อยได้ คนกลุ่มต่ำตอบได้มากกว่า เป็น
ข้อสอบที่ต้องปรับปรุง

| ข้อที่ | H | L | H+L | H-L | P | r |
|--------|----|---|-----|-----|------|-------|
| 1 | | | | | | |
| 2 | | | | | | |
| 3 | | | | | | |
| 4 | 6 | | 16 | -4 | 0.57 | -0.28 |
| 5 | 12 | 5 | 17 | 7 | 0.61 | 0.50 |

ข้อ 5 ความยากเท่ากับ 0.61 ค่าอำนาจจำแนกเท่ากับ 0.50 มีค่าความยากและอำนาจจำแนกอยู่ในเกณฑ์ดี เป็นข้อสอบที่ดี

ตัวบ่งชี้ของข้อสอบที่ดีคืออะไร?

คำตอบคือ **อำนาจจำแนก**

ข้อสอบที่ง่ายเกินไปก็ไม่ดี ยากเกินไปก็ไม่ดี
และความยากที่เหมาะสมก็ใช้ว่าจะดีเสมอไป

ข้อสอบที่ดีคือ **ข้อสอบที่มีอำนาจจำแนกได้มาตรฐาน**

ข้อสอบที่มีอำนาจจำแนกได้มาตรฐาน คือ

ตัวถูก H มากกว่า L ไม่น้อยกว่า 20%

ตัวลวงทุกตัว L มากกว่า H ไม่น้อยกว่า 5%

๔ ข้อสอบที่ยากที่สุด (สร้างได้ง่าย)

ตัวอย่างโจทย์

“เส้นผมของนักศึกษามีกี่เส้น ”

หรือ

“จงหาค่า $10\sqrt{2}$ โดยใช้ทศนิยม 5 ตำแหน่ง”

และมีเวลาทำข้อสอบข้อละ 2-3 นาที
(ไม่อนุญาตให้ใช้เครื่องคิดเลข)

การแปลความค่า P

ตัวบ่งชี้ความยากของข้อสอบ

| | P | ความหมาย |
|---|-------------|--------------|
| 1 | 0.00 - 0.19 | ยาก |
| 2 | 0.20 - 0.39 | ค่อนข้างยาก |
| 3 | 0.40 - 0.59 | ปานกลาง |
| 4 | 0.60 - 0.80 | ค่อนข้างง่าย |
| 5 | 0.81 - 1.00 | ง่าย |

ความยากของข้อสอบที่เหมาะสมคือ P มีค่าอยู่ระหว่าง 0.20 ถึง 0.80

การแปลความค่า R

อำนาจจำแนกของข้อสอบ

| | R | ความหมาย |
|---|---------------|-------------------------------|
| 1 | น้อยกว่า 0.00 | จำแนกผิดพลาด |
| 2 | 0.00 ถึง 0.19 | จำแนกได้น้อย หรือ จำแนกไม่ได้ |
| 3 | 0.20 ถึง 0.29 | พอใช้ |
| 4 | 0.30 ถึง 0.39 | ดี |
| 5 | 0.40 ถึง 1.00 | ดีมาก |

อำนาจจำแนกที่ดีคือ

r มีค่าไม่น้อยกว่า 0.20 สำหรับตัว ถูก (เครื่องหมาย บวก)

และมีค่าไม่น้อยกว่า 0.05 สำหรับตัว ลวง (เครื่องหมาย ลบ)