

# บทที่ 4

## ความเชื่อมั่น

### จุดมุ่งหมายเชิงพฤติกรรม

1. สามารถบอกนิยามของสัมประสิทธิ์ความเชื่อมั่นได้
2. สามารถประมาณค่าความเชื่อมั่นด้วยวิธีสอบซ้ำได้
3. สามารถประมาณค่าความเชื่อมั่นด้วยแบบทดสอบคู่ขนานหรือแบบทดสอบทางเลือกได้
4. สามารถประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายในแบบแบ่งครึ่งได้
5. สามารถประมาณค่าความเชื่อมั่นด้วยวิธีความสอดคล้องภายในได้
6. สามารถประมาณค่าความเชื่อมั่นด้วยสูตรของสเปียร์แมนบราวน์ได้
7. สามารถเปรียบเทียบการประมาณค่าความเชื่อมั่นด้วยวิธีต่าง ๆ ได้
8. สามารถประมาณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ได้
9. สามารถอธิบายอิทธิพลต่าง ๆ ที่อาจมีผลต่อความเชื่อมั่นได้
10. สามารถเปรียบเทียบความแตกต่างของความคลาดเคลื่อนมาตรฐานได้

### เนื้อหา

1. นิยามสัมประสิทธิ์ความเชื่อมั่น
2. การประมาณค่าความเชื่อมั่นด้วยวิธีการสอบซ้ำ
3. การประมาณค่าความเชื่อมั่นด้วยแบบทดสอบคู่ขนานหรือแบบทดสอบทางเลือก
4. การประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายใน : แบบแบ่งครึ่ง
5. ความเชื่อมั่นแบบความสอดคล้องภายใน : กรณีทั่วไป
6. สูตรสเปียร์แมนบราวน์ : กรณีทั่วไป
7. เปรียบเทียบวิธีการประมาณค่าความเชื่อมั่น
8. ความเชื่อมั่นของแบบทดสอบอิงเกณฑ์
9. อิทธิพลที่มีต่อความเชื่อมั่น
10. ความคลาดเคลื่อนมาตรฐานของการวัดและการประมาณค่า

### กิจกรรมการเรียนรู้การสอน

1. บรรยาย
2. ชักถามระหว่างบรรยาย
3. แบบฝึกหัดจากใบงาน

### สื่อการเรียนรู้การสอน

1. แผ่นใส
2. เอกสารประกอบการสอน
3. ใบงาน

### การวัดผล

1. สังเกตความตั้งใจขณะบรรยาย
2. สังเกตการตอบคำถามของผู้เรียน
3. ตรวจสอบผลงานแบบฝึกหัด

### นิยามสัมประสิทธิ์ความเชื่อมั่น

มีหลากหลายวิธีในการนิยามและแปลความหมายความเชื่อมั่นของแบบทดสอบ ตัวอย่างเช่น แบบทดสอบจะเชื่อมั่นได้ถ้าคะแนนสังเกตและคะแนนจริงมีความสัมพันธ์กันสูง นั่นคือ คะแนนที่สังเกตได้และคะแนนจริงที่ได้มาจากผู้สอบทุก ๆ คนที่สอบแบบทดสอบฉบับหนึ่ง แล้วนำคะแนนสังเกตได้และคะแนนจริงมาหาความสัมพันธ์และยกกำลังสอง ( $\rho_{XT}^2$ ) จะเรียกว่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ หรือความเชื่อมั่นสามารถแสดงได้ด้วยสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสังเกตของแบบทดสอบที่คู่ขนานแบบพาราเรล (Parallel) กันสองฉบับ ถ้าแบบทดสอบที่คู่ขนานกันแบบพาราเรลสองฉบับนั้นใช้สอบกับประชากรผู้สอบและผลของคะแนนสังเกตของแบบทดสอบสองฉบับที่คู่ขนานนั้นนำมาหาความสัมพันธ์กัน ค่าสหสัมพันธ์นี้ (ใช้สัญลักษณ์  $\rho_{XX'}$  เมื่อ X และ X' คือคะแนนสังเกตของแบบทดสอบสองฉบับที่คู่ขนานกันแบบพาราเรล) ก็คือค่าสัมประสิทธิ์ความเชื่อมั่น โดยปกติแล้วเราไม่สามารถทราบคะแนนจริงได้ และเป็นไปไม่ได้ที่จะใช้แบบทดสอบที่คู่ขนานกันแบบพาราเรล ดังนั้นความเชื่อมั่นต้องประมาณค่าด้วยวิธีอื่น ๆ ซึ่งมีอยู่ 6 วิธีสำหรับนิยามหรือแปลความหมายของสัมประสิทธิ์ความเชื่อมั่น ดังนี้

1.  $\rho_{XX'}$  = สหสัมพันธ์ระหว่างคะแนนที่สังเกตได้จากแบบทดสอบคู่ขนาน
2.  $\rho_{XX'}^2$  = สัดส่วนของความแปรปรวนใน X ที่อธิบายด้วยความสัมพันธ์เชิงเส้นกับ

X'

3.  $\rho_{XX'} = \sigma_T^2 / \sigma_X^2$
4.  $\rho_{XX'} = \rho_{XT}^2$
5.  $\rho_{XX'} = 1 - \rho_{XE}^2$
6.  $\rho_{XX'} = 1 - \sigma_E^2 / \sigma_X^2$

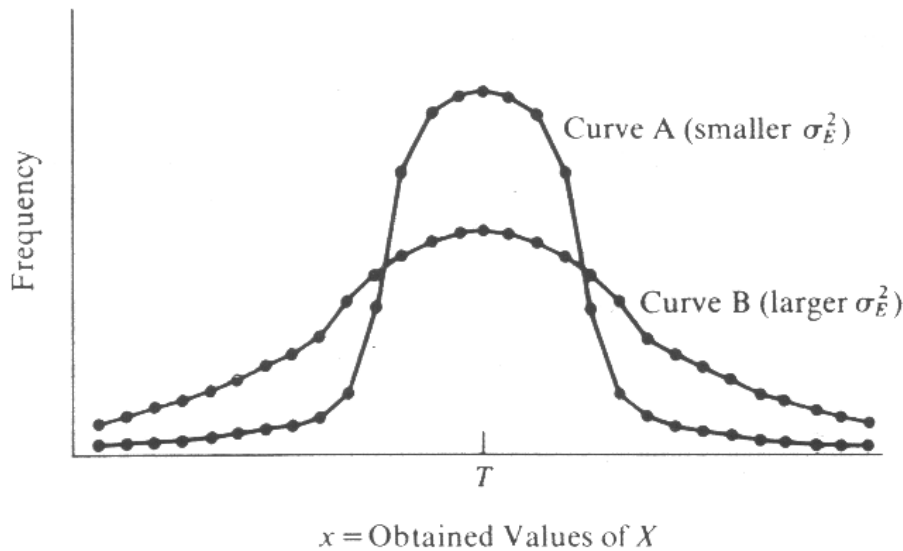
การแปลความหมายในนิยามที่ 1 คือความเชื่อมั่นของแบบทดสอบจะเท่ากับสหสัมพันธ์ของคะแนนสังเกตที่ได้จากแบบทดสอบฉบับหนึ่งกับคะแนนสังเกตที่ได้จากแบบทดสอบอีกฉบับหนึ่งที่คู่ขนานกัน ผู้สอบจะได้คะแนนสังเกตเท่ากันก็ต่อเมื่อแบบทดสอบมีความคู่ขนานและมีความแปรปรวนของคะแนนสังเกตในการสอบแต่ละครั้งเท่ากัน ดังนั้นแบบทดสอบจะมีความเชื่อมั่นที่สมบูรณ์ ( $r_{XX'} = 1$ ) แต่ถ้าผู้สอบมีคะแนนสังเกตในแบบทดสอบฉบับหนึ่งไม่สัมพันธ์กับคะแนนสังเกตอีกฉบับหนึ่งที่คู่ขนานกัน ( $r_{XX'} = 0$ ) แบบทดสอบย่อมเชื่อมั่นไม่ได้

การแปลความหมายในนิยามที่ 2 คือการแปลความหมายโดยใช้สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน กำลังสองของสัมประสิทธิ์สหสัมพันธ์จะแปลความหมายได้ว่า สัดส่วนของความแปรปรวนของตัวแปรหนึ่งสามารถอธิบายได้ด้วยความสัมพันธ์เชิงเส้นกับอีกตัวแปรหนึ่ง ดังนั้น  $\rho_{XX'}^2$  จะเท่ากับสัดส่วนของความแปรปรวนของคะแนนแบบทดสอบฉบับหนึ่งที่อธิบายได้ด้วยความสัมพันธ์เชิงเส้นกับคะแนนในแบบทดสอบอีกฉบับหนึ่งที่คู่ขนานกัน

การแปลความหมายในนิยามที่ 3  $r_{XX'} = \sigma_T^2 / \sigma_X^2$  นั่นคือสัมประสิทธิ์ความเชื่อมั่นเป็นอัตราส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สังเกตได้ สำหรับแบบทดสอบที่มีความเชื่อมั่นอย่างสมบูรณ์  $r_{XX'} = 1$  ดังนั้น  $\sigma_T^2 / \sigma_X^2 = 1$  และความแปรปรวนของคะแนนที่สังเกตได้ทั้งหมดสะท้อนให้เห็นความแปรปรวนของคะแนนจริง ถ้า  $r_{XX'} = 1$  ความแตกต่างระหว่างคะแนนที่สังเกตได้มาจากความแตกต่างระหว่างคะแนนจริง ถ้า  $\sigma_X^2 = \sigma_T^2$  แล้ว  $\sigma_E^2$  ต้องเป็น 0 ดังนั้น  $\varepsilon(E) = 0$ , ความคลาดเคลื่อนทั้งหมดจะเท่ากับ 0 เมื่อ  $\sigma_E^2 = 0$  แล้วการวัดจะต้องปราศจากความคลาดเคลื่อน ถ้า  $r_{XX'} < 1$  แล้วความคลาดเคลื่อนในการวัดยังคงมีอยู่ ถ้า  $r_{XX'} = 0$  แล้ว  $\sigma_X^2 = \sigma_E^2$  ซึ่งหมายความว่าคะแนนที่สังเกตได้ทั้งหมดจะสะท้อนให้เห็นเฉพาะความคลาดเคลื่อนเท่านั้น

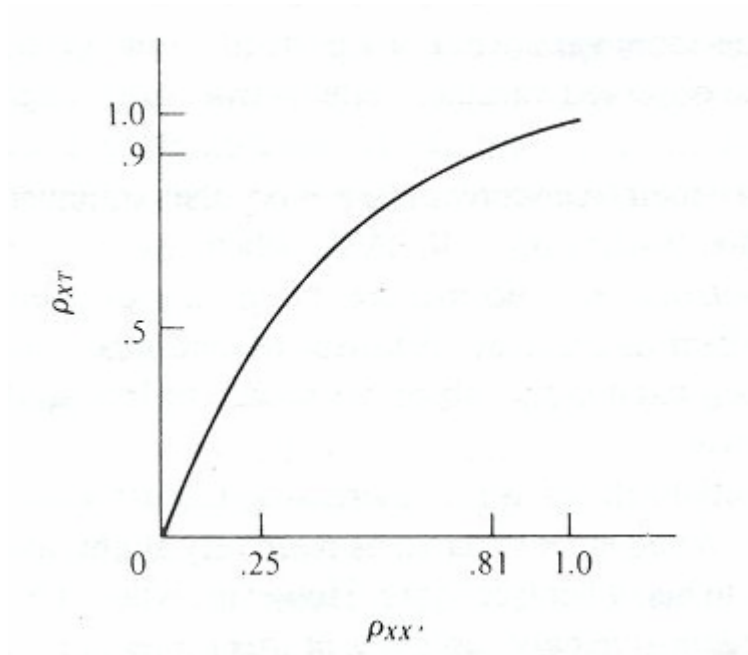
ขณะที่ความเชื่อมั่นของแบบทดสอบเพิ่มขึ้น ความแปรปรวนของคะแนนความคลาดเคลื่อนจะน้อยลง เมื่อความแปรปรวนของความคลาดเคลื่อนมีน้อยแล้ว คะแนนที่สังเกตของผู้สอบจะมีค่าเข้าใกล้คะแนนจริงของผู้สอบมาก อย่างไรก็ตามเมื่อความแปรปรวนของความคลาดเคลื่อนมีมาก คะแนนที่สังเกตได้จะให้การประมาณค่าคะแนนจริงที่ไม่ดี ภาพประกอบ 4.1 จะแสดงความสัมพันธ์นี้ โค้งจะแสดงการแจกแจงเชิงทฤษฎีของคะแนนสังเกตเมื่อคะแนนจริงค่าหนึ่งคงที่ นั่นคือการแจกแจงของคะแนนสังเกตสำหรับผู้สอบคนหนึ่ง คะแนนจริงของผู้สอบจะแสดงด้วยสัญลักษณ์ T ในรูปภาพ เมื่อคะแนนจริงถูกกำหนดให้คงที่  $\sigma_T^2 = 0$  และความ

แปรปรวนของคะแนนสังเกตจะเท่ากับความแปรปรวนของคะแนนความคลาดเคลื่อน ภายใต้โค้ง A ซึ่งมีความแปรปรวนของความคลาดเคลื่อนน้อย คะแนนที่สังเกตได้จะมีค่าเข้าใกล้คะแนนจริง T มาก ภายใต้โค้ง B ซึ่งมีความแปรปรวนของความคลาดเคลื่อนมาก คะแนนที่สังเกตได้จะอยู่ไกลจากคะแนนจริง T

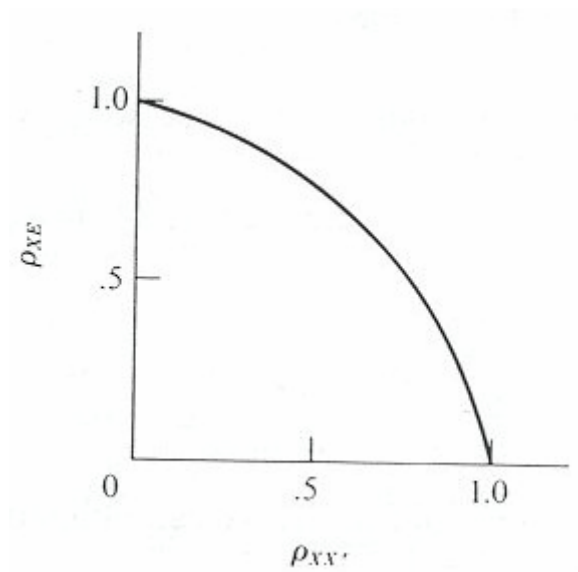


ภาพประกอบ 4.1 แสดงอิทธิพลของความแปรปรวนของความคลาดเคลื่อนบนคะแนนจริงต่อค่าที่สังเกตได้ของ X (Allen and Yen. 1979)

การแปลความหมายในนิยามที่ 4  $\rho_{XX'} = \rho_{XT}^2$  นั่นคือ สัมประสิทธิ์ความเชื่อมั่นเท่ากับกำลังสองของสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนจริง ตัวอย่างเช่น ถ้า  $r_{XX'} = 0.81$  แล้ว  $r_{XT} = 0.9$  หรือ  $r_{XX'} = 0.25$  แล้ว  $r_{XT} = 0.5$  ความสัมพันธ์นี้จะแสดงในภาพประกอบ 4.2 เมื่อ  $0 < r_{XX'} < 1$  เราจะได้ว่า  $r_{XT} > r_{XX'}$  คะแนนสังเกตและคะแนนจริงจะมีความสัมพันธ์กันสูงกว่าความสัมพันธ์ระหว่างคะแนนที่สังเกตได้จากแบบทดสอบคู่ขนาน ในความเป็นจริง คะแนนจากแบบทดสอบจะไม่สัมพันธ์กันสูงกับแบบทดสอบอื่น ๆ มากไปกว่าคะแนนจริงของตัวเอง สหสัมพันธ์ที่สูงมากระหว่างคะแนนที่สังเกตได้กับแบบทดสอบอื่น ๆ คือ  $\sqrt{\rho_{XX'}} = \rho_{XT}$  ถ้า Y เป็นแบบทดสอบฉบับที่เป็นเกณฑ์ แล้วนำมาหาความสัมพันธ์กับแบบทดสอบ X แล้ว  $r_{XY}$  จะเรียกว่า สัมประสิทธิ์ความเที่ยงตรง เพราะว่า  $r_{XY}$  จะมีค่าไม่สูงไปกว่า  $r_{XT}$  และ  $r_{XY}$  ก็จะมีค่าไม่สูงไปกว่า  $\sqrt{\rho_{XX'}}$  ดังนั้นความเชื่อมั่นจึงมีผลต่อความเที่ยงตรง แม้ว่าสัมประสิทธิ์ความเที่ยงตรงมีค่าไม่สูงไปกว่ารากที่สองของสัมประสิทธิ์ความเชื่อมั่น ตัวอย่างเช่น ถ้า  $r_{XX'} = 0.49$  แล้ว  $r_{XT} = 0.7$



ภาพประกอบ 4.2 ความสัมพันธ์ระหว่าง  $\rho_{XX'}$  และ  $\rho_{XY}$  (Allen and Yen. 1979)



ภาพประกอบ 4.3 ความสัมพันธ์ระหว่าง  $\rho_{XX'}$  และ  $\rho_{XE}$  (Allen and Yen. 1979)

การแปลความหมายในนิยามที่ 5  $\rho_{XX'} = 1 - \rho_{XE}^2$  นั่นคือสัมประสิทธิ์ความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยกำลังสองของสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนความคลาดเคลื่อน ในแนวคิดนี้  $r_{XE}$  ควรจะมีค่า 0 แต่  $r_{XE} = 0$  เมื่อ  $r_{XX'} = 1.0$  ความสัมพันธ์ระหว่าง  $r_{XE}$  และ  $r_{XX'}$  ดังแสดงในภาพประกอบ 4.3 นั่นคือเมื่อสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนความคลาดเคลื่อนลดลงแล้วค่าสัมประสิทธิ์ความเชื่อมั่นจะมีค่าเพิ่มมากขึ้น

การแปลความหมายในนิยามที่ 6  $\rho_{XX'} = 1 - \sigma_E^2 / \sigma_X^2$  นั่นคือ ความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยสัดส่วนของความแปรปรวนของคะแนนความคลาดเคลื่อนกับความแปรปรวนของคะแนนที่สังเกตได้ เมื่อ  $r_{XX'} = 1$ ,  $\sigma_E^2 = 0$  และเมื่อ  $r_{XX'} = 0$ ,  $\sigma_E^2 = \sigma_X^2$  ค่าความเชื่อมั่นขึ้นอยู่กับกลุ่มผู้สอบด้วย นั่นคือถ้ากลุ่มผู้สอบมีความเป็นวิวิธพันธ์ คือมีความแตกต่างของความสามารถมากค่าความเชื่อมั่นมีแนวโน้มจะสูงกว่ากลุ่มผู้สอบที่มีความเป็นเอกพันธ์ เช่น ใช้แบบทดสอบ IQ กับกลุ่มผู้สอบที่มีสมองช้า กับกลุ่มผู้สอบปกติ ค่าความเชื่อมั่นที่คำนวณจากกลุ่มผู้สอบที่มีสมองช้ามีแนวโน้มจะมีค่าต่ำกว่าค่าความเชื่อมั่นที่คำนวณจากกลุ่มผู้สอบปกติ

โดยสรุปแล้ว เมื่อ  $r_{XX'} = 1$  มีสาเหตุจาก

- 1) การวัดปราศจากความคลาดเคลื่อน ( $E = 0$ )
- 2)  $X = T$  ในผู้สอบทุก ๆ คน
- 3) ความแปรปรวนของคะแนนสังเกตทั้งหมดสะท้อนให้เห็นความแปรปรวนของคะแนนจริง ( $\sigma_X^2 = \sigma_T^2$ )
- 4) ความแตกต่างทั้งหมดระหว่างคะแนนสังเกตสะท้อนถึงความแตกต่างของคะแนนจริง
- 5) สหสัมพันธ์ระหว่างคะแนนสังเกตกับคะแนนจริงมีค่า 1 ( $r_{XT} = 1$ ) และ
- 6) สหสัมพันธ์ระหว่างคะแนนสังเกตกับคะแนนคลาดเคลื่อนมีค่า 0 ( $r_{XE} = 0$ )

เมื่อ  $r_{XX'} = 0$  มีสาเหตุจาก

- 1) เกิดเฉพาะความคลาดเคลื่อนอย่างสุ่มในการวัด
- 2)  $X = E$  สำหรับผู้สอบทุก ๆ คน
- 3) ความแปรปรวนของคะแนนสังเกตทั้งหมดสะท้อนถึงความแปรปรวนของคะแนนคลาดเคลื่อน ( $\sigma_X^2 = \sigma_E^2$ )
- 4) ความแตกต่างทั้งหมดระหว่างคะแนนสะท้อนถึงความคลาดเคลื่อนในการวัด
- 5) สหสัมพันธ์ระหว่างคะแนนสังเกตกับคะแนนจริงมีค่า 0 ( $r_{XT} = 0$ ) และ
- 6) สหสัมพันธ์ระหว่างคะแนนสังเกตกับคะแนนคลาดเคลื่อนมีค่า 1 ( $r_{XE} = 1$ )

เมื่อ  $0 \leq r_{XX'} \leq 1$  มีสาเหตุจาก

- 1) การวัดจะรวมความคลาดเคลื่อนบางอย่างเข้าไว้ด้วย
- 2)  $X = T + E$

- 3) ความแปรปรวนของคะแนนสังเกตประกอบด้วยความแปรปรวนของคะแนนจริงบางส่วนรวมกับความแปรปรวนของคะแนนคลาดเคลื่อนบางอย่าง  
( $\sigma_X^2 = \sigma_T^2 + \sigma_E^2$ )
- 4) ความแตกต่างระหว่างคะแนนสามารถสะท้อนให้เห็นความคลาดเคลื่อนของการวัด เช่นเดียวกับความแตกต่างของคะแนนจริง
- 5) สหสัมพันธ์ระหว่างคะแนนสังเกตและคะแนนจริง  $r_{XT}$  เท่ากับ  $\sqrt{\rho_{XX'}}$
- 6) สหสัมพันธ์ระหว่างคะแนนสังเกตและคะแนนความคลาดเคลื่อน  $r_{XE}$  คือ  $\sqrt{1 - \rho_{XX'}}$
- 7) ความเชื่อมั่นก็คือสัดส่วนของความแปรปรวนของคะแนนสังเกตได้ในส่วนที่เป็นความแปรปรวนของคะแนนจริง ( $r_{XX'} = \sigma_T^2 / \sigma_X^2$ ) และ
- 8) ค่า  $r_{XX'}$  ที่สูงจะช่วยให้เกิดความเชื่อมั่นในการประมาณค่า T จาก X เพราะความแปรปรวนของความคลาดเคลื่อนมีค่าน้อยลง

### การประมาณค่าความเชื่อมั่นด้วยวิธีการสอบซ้ำ

ความเชื่อมั่นของการสอบซ้ำ จะนิยามอยู่บนพื้นฐานของการใช้กลุ่มผู้สอบกลุ่มเดียวกันที่ใช้แบบทดสอบฉบับเดียวกันซ้ำสองครั้งแล้วนำมาหาความสัมพันธ์กัน ถ้าผู้สอบแต่ละคนได้คะแนนสอบเหมือนกันในการสอบซ้ำสองครั้ง และมีความแปรปรวนของคะแนนสังเกตเท่ากันแล้ว สหสัมพันธ์จะได้เท่ากับ 1.0 บ่งชี้ถึงความสัมพันธ์กันอย่างสมบูรณ์ หรือในอีกกรณีหนึ่ง ถ้าคะแนนสังเกตสำหรับผู้สอบทุกคนอ้างอิงมาจากการทดสอบครั้งที่หนึ่งที่มีความสัมพันธ์ของคะแนนสังเกตเป็นเชิงเส้นตรงอย่างสมบูรณ์กับการทดสอบครั้งที่สองแล้ว การประมาณค่าความเชื่อมั่นจะเท่ากับ 1.0 แต่ถ้าคะแนนจากการทดสอบครั้งแรกไม่มีความสัมพันธ์กับคะแนนจากการทดสอบครั้งที่สอง การประมาณค่าความเชื่อมั่นจะได้ 0.0

ปัญหาที่สำคัญในการประมาณค่าความเชื่อมั่นแบบสอบซ้ำ ประการแรกคือผลของการเกิด carry-over effect ระหว่างการทดสอบ การทดสอบครั้งแรกอาจจะมีอิทธิพลต่อการสอบครั้งที่สอง เกิดความคลาดเคลื่อนของคะแนนสำหรับการสอบ ผู้สอบเมื่อสอบครั้งที่สองอาจจะจำคำตอบที่ได้ตอบไปในครั้งแรก และเป็นไปได้ที่ตอบคำตอบเดิมซ้ำ หรือในบางการทดสอบ carry-over effects อาจจะมีเนื่องมาจากการฝึกฝน เช่นผู้สอบส่วนใหญ่มีแนวโน้มที่ได้รับการฝึกฝนจากการทดสอบซ้ำทำให้มีความคล่องแคล่วในการทำแบบทดสอบและในบางความสามารถที่แบบทดสอบวัด ถ้ามีผู้สอบบางคนที่ได้รับการฝึกฝนมากกว่าผู้สอบคนอื่น

การเปลี่ยนแปลงคติของผู้สอบหรือระดับความสามารถของผู้สอบอาจมีสาเหตุก่อให้เกิด carry-over effects ได้ การไม่ให้ความร่วมมือของผู้สอบ ผู้สอบอาจจะไม่ยอมทำแบบทดสอบในครั้งที่สองและจงใจจะเดาหรือตอบผิดในการสอบครั้งที่สอง ผลนี้จะทำให้สัมพันธ์กับความเชื่อมั่น

ต่ำระหว่างการสอบทั้งสองครั้ง หรือหลังจากการสอบครั้งแรก ผู้สอบบางคนอาจจะไปค้นหาข้อมูลเพิ่มเติมเพื่อช่วยให้คะแนนของตนเองเพิ่มขึ้น ดังนั้นสัมประสิทธิ์สหสัมพันธ์ระหว่างการทดสอบซ้ำสองครั้งควรมีแนวโน้มต่ำ ดังนั้น carry-over effects สามารถมีผลต่อการประมาณค่าความเชื่อมั่น ช่วยให้การประมาณค่าความเชื่อมั่นแบบสอบซ้ำต่ำหรือสูงกว่าความเป็นจริง

ปัญหาประการที่สองในการประมาณค่าความเชื่อมั่นแบบสอบซ้ำจะเกี่ยวข้องกับระยะเวลาที่เว้นช่วงห่างระหว่างการสอบทั้งสองครั้ง การเว้นช่วงห่างที่สั้นมากจะมีผลให้เกิด carry-over effects อันเนื่องมาจากความจำข้อสอบได้ การฝึกฝน หรืออารมณ์ การเว้นช่วงห่างที่ยาวนานอาจจะมีผลเนื่องมาจากการเปลี่ยนแปลงของความรู้หรืออารมณ์ ถ้าคุณลักษณะที่แบบทดสอบวัดมีการเปลี่ยนแปลงไปตามระยะเวลา เช่นความสามารถทางสมองของเด็ก การเว้นช่วงระยะห่างระหว่างการสอบยาวนานเกินไปจะมีแนวโน้มการประมาณค่าความเชื่อมั่นต่ำ ความแตกต่างของการเว้นช่วงระยะเวลาจะมีอิทธิพลต่อการประมาณค่าความเชื่อมั่น ในบางครั้งจะมีผลการประมาณค่าความเชื่อมั่นต่ำหรือสูงกว่าความเป็นจริง

การประมาณค่าความเชื่อมั่นของการสอบซ้ำอยู่บนพื้นฐานของการออกแบบในการสอบซ้ำ ซึ่งการประมาณค่าความเชื่อมั่นแบบสอบซ้ำอาจจะมีอิทธิพลมาจากผลของ carry-over effects และการเว้นระยะเวลาสอบซ้ำ การประมาณค่าความเชื่อมั่นของการสอบซ้ำมีความเหมาะสมมากสำหรับแบบทดสอบที่วัดคุณลักษณะที่ไม่อ่อนไหวต่อ carry-over effects และมีความคงที่เมื่อเว้นช่วงห่างของการสอบซ้ำ เช่น ใช้การประมาณค่าความเชื่อมั่นแบบสอบซ้ำกับแบบทดสอบที่ใช้สไตประสาธสัมพันธ์สิ่งทำ (เช่นแบบทดสอบวัดการมองเห็น หรือการฟัง)

### การประมาณค่าความเชื่อมั่นด้วยการใช้แบบทดสอบคู่ขนานและแบบทดสอบทางเลือก

ความเชื่อมั่นของการใช้แบบทดสอบคู่ขนานสามารถคำนวณได้ด้วยสูตร  $r_{xx'}$  เป็นสหสัมพันธ์ระหว่างคะแนนที่สังเกตของแบบทดสอบสองฉบับที่คู่ขนานกันแบบพาราเรล (parallel) ในทางปฏิบัติเป็นไปได้ที่จะมีแบบทดสอบสองฉบับที่คู่ขนานกันแบบพาราเรล และแบบทดสอบทางเลือกมักถูกนำมาใช้แทนที่แบบทดสอบคู่ขนานเสมอ แบบทดสอบทางเลือกคือแบบทดสอบอีกฉบับหนึ่งที่มีโครงสร้างทำนองเดียวกับแบบทดสอบคู่ขนาน แบบทดสอบคู่ขนานและแบบทดสอบทางเลือกนี้จะมีความเท่ากันในค่าเฉลี่ยของคะแนนสังเกต ความแปรปรวนของคะแนนสังเกต และสหสัมพันธ์กับแบบทดสอบอื่น อย่างไรก็ตาม ซึ่งยากที่จะหาแบบทดสอบสองฉบับที่มีคุณลักษณะแบบนี้ สหสัมพันธ์ระหว่างคะแนนสังเกตกับแบบทดสอบทางเลือกคือ  $r_{xz}$  เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบฉบับหนึ่งกับแบบทดสอบทางเลือก สหสัมพันธ์นี้จะมีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบ หรืออีกกรณีหนึ่งก็คือแบบทดสอบสองฉบับนี้คู่ขนานกัน ดังนั้นการใช้แบบทดสอบทางเลือกจะเป็นการประมาณค่าความเชื่อมั่นที่แตกต่างไปจากแบบสอบซ้ำ อย่างไรก็ตามในการใช้แบบทดสอบทางเลือกหรือแบบทดสอบคู่ขนานจะไม่



สามารถจัด carry-over effect ให้หมดไปได้ ซึ่งยังคงมีผลต่อรูปแบบการตอบ อารมณ์หรือเจตคติ carry-over effects ยังคงมีผลต่อการประมาณค่า  $\rho_{XX'}$  หรือ  $\rho_{ZZ'}$  ให้สูงกว่าหรือต่ำกว่าความเป็นจริง ช่วงเวลายังคงมีปัญหา การเว้นช่วงเวลาที่สั้นไประหว่างการสอบสองฉบับจะมีผลเนื่องมาจากความจำ การฝึกฝนและอารมณ์ การเว้นช่วงเวลาที่นานเกินไปจะไม่เหมาะกับแบบทดสอบที่วัดคุณลักษณะที่เปลี่ยนแปลงไปตามเวลา

เมื่อแบบทดสอบฉบับหนึ่ง X และแบบทดสอบทางเลือก Z ไม่มีความคู่ขนานกันแล้วความสัมพันธ์ระหว่างคะแนนสังเกตของแบบทดสอบทั้งสองฉบับนี้  $r_{XZ}$  โดยทั่วไปจะเป็นตัวประมาณค่าที่ไม่ถูกต้องของ  $\rho_{XX'}$  หรือ  $\rho_{ZZ'}$  ตัวอย่างเช่น ให้  $X = T_X + E_X$  และ  $Z = T_Z + E_Z$  ถ้า  $T_X = T_Z$  แต่  $\sigma_{E_X}^2 > \sigma_{E_Z}^2$  แล้ว X จะมีความเชื่อมั่นน้อยกว่า Z สหสัมพันธ์  $r_{XZ}$  จะมีแนวโน้มประมาณค่าได้สูงกว่า  $\rho_{XX'}$  และประมาณค่าได้ต่ำกว่า  $\rho_{ZZ'}$  ถ้า  $T_X \neq T_Z$  เป็นไปได้ว่าแบบทดสอบสองฉบับนี้จะวัดคุณลักษณะที่แตกต่างกัน และ  $r_{XZ}$  จะมีแนวโน้มประมาณค่าได้ต่ำกว่า  $\rho_{XX'}$  และ  $\rho_{ZZ'}$  ตัวอย่างเช่น ถ้า X คือคะแนนของแบบทดสอบคณิตศาสตร์คำนวณ และ Z คือคะแนนของแบบทดสอบคณิตศาสตร์เหตุผล  $r_{XZ}$  คือสหสัมพันธ์ระหว่างคะแนนคณิตศาสตร์คำนวณและคณิตศาสตร์เหตุผล และไม่จำเป็นว่าจะต้องเป็นตัวประมาณค่าที่ดีของความเชื่อมั่นในแบบทดสอบทั้งสองฉบับ

เป็นไปได้ที่แบบทดสอบทางเลือกจะมีความไม่เท่ากันของคะแนนจริงและความแปรปรวนของคะแนนคลาดเคลื่อน แม้ว่าสหสัมพันธ์ระหว่างคะแนนสังเกตจะเท่ากับความสัมพันธ์ของแบบทดสอบคู่ขนาน ตัวอย่างเช่น ให้  $X = T_X + E_X$  และ  $X' = T_{X'} + E_{X'}$  เมื่อ X และ X' คือคะแนนของแบบทดสอบคู่ขนาน ให้  $Z = aX' + b$  เมื่อ a และ b คือค่าคงที่ และ  $a > 0$  นั่นคือ Z เป็นฟังก์ชันเชิงเส้นของ X' แม้ว่า Z และ X จะไม่ใช่แบบทดสอบคู่ขนาน ( $T_Z \neq T_X$  และ  $\sigma_{E_X}^2 \neq \sigma_{E_Z}^2$ ) แต่  $\rho_{XZ} = \rho_{XX'}$  เมื่อ Z คือฟังก์ชันเชิงเส้นของ X' สหสัมพันธ์ของ X และ Z จะเท่ากับสหสัมพันธ์ของ X กับ X'

สรุป สหสัมพันธ์ระหว่างคะแนนสังเกตบนแบบทดสอบทางเลือกให้การประมาณค่าความเชื่อมั่นที่ดีถ้าแบบทดสอบทางเลือกมีความคู่ขนานกันหรือคะแนนมีความสัมพันธ์กันเชิงเส้นตรง และถ้า carry-over effect และการเว้นช่วงระยะเวลาของการสอบไม่มีผลต่อสหสัมพันธ์

### การประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายใน : แบบแบ่งครึ่ง

ความเชื่อมั่นแบบสอดคล้องภายในจะถูกประมาณค่าด้วยการใช้แบบทดสอบเพียงฉบับเดียวสอบเพียงครั้งเดียว ดังนั้นจึงเป็นการหลีกเลี่ยงปัญหาที่เกิดขึ้นจากวิธีสอบซ้ำ วิธีนี้เป็นที่นิยมใช้กันอย่างแพร่หลาย แบบทดสอบจะถูกแบ่งครึ่งออกเป็นสองส่วน ซึ่งแต่ละส่วนจะคู่ขนานกันแบบพาราเรล (parallel) และการประมาณค่าความเชื่อมั่นของแบบทดสอบจะใช้สูตรของสเปียร์แมนบราวน์ (Spearman-Brown formula) แต่ถ้าทั้งสองส่วนนั้นคู่ขนานกันแบบทอ (essentially

$\tau$ -equivalent) จะใช้สูตรสัมประสิทธิ์แอลฟา ( $\alpha$ -coefficient) ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ

การใช้สูตรสเปียร์แมนบราวน์ คำนวณจากแบบทดสอบแบ่งครึ่ง (เรียกว่า  $Y$  และ  $Y'$ ) จะนำมาหาสหสัมพันธ์กัน ผลที่ได้คือ  $\rho_{yy'}$  สหสัมพันธ์นี้จะเป็นการวัดความเชื่อมั่นของแบบทดสอบเพียงครึ่งฉบับ ความเชื่อมั่นของแบบทดสอบทั้งฉบับ  $X = Y + Y'$  ควรจะมีค่ามากกว่าความเชื่อมั่นของแบบทดสอบเพียงแค่ครึ่งฉบับ สูตรสเปียร์แมนบราวน์จะให้ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ คือ

$$\rho_{xx'} = \frac{2\rho_{yy'}}{1 + \rho_{yy'}}$$

ตาราง 4.1 จะแสดงค่าความเชื่อมั่น โดยปกติ  $\rho_{xx'}$  จะมีค่าสูงกว่า  $\rho_{yy'}$  เพราะ  $\rho_{xx'}$  เป็นค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ และ  $\rho_{yy'}$  เป็นค่าความเชื่อมั่นของแบบทดสอบเพียงครึ่งฉบับ

ตาราง 4.1 สหสัมพันธ์ระหว่างแบบทดสอบที่แบ่งครึ่งฉบับ ( $\rho_{yy'}$ ) และความเชื่อมั่นของแบบทดสอบทั้งฉบับ ( $\rho_{xx'}$ )

$\rho_{xx'}$	$\rho_{yy'}$
0.00	0.00
0.33	0.20
0.57	0.40
0.75	0.60
0.89	0.80
1.00	1.00

สูตรสเปียร์แมนบราวน์สามารถใช้หาความเชื่อมั่นของแบบทดสอบที่แบ่งครึ่งฉบับแล้วมีความคู่ขนานกันแบบพาราเรล (parallel) แต่ถ้าคะแนนแบ่งครึ่งนั้นมีความไม่เท่ากันในความแปรปรวนหรือไม่เท่ากันในกรณีอื่น ๆ ที่บ่งชี้ว่าไม่คู่ขนานกันแบบพาราเรล (parallel) แล้วก็ต้องไปใช้สัมประสิทธิ์แอลฟาในการประมาณค่าความเชื่อมั่น ถ้าสองส่วนที่แบ่งครึ่งนั้นมีความคู่ขนานกันแบบทอ (essentially  $\tau$ -equivalent) แต่ถ้าทั้งสองส่วนที่แบ่งครึ่งไม่มีความคู่ขนานกันแบบทอแล้ว ค่าสัมประสิทธิ์แอลฟาก็จะประมาณค่าความเชื่อมั่นได้ต่ำ (นั่นคือความเชื่อมั่นของแบบทดสอบต้องสูงกว่าหรือเท่ากับผลที่ได้จากสูตรสัมประสิทธิ์แอลฟา) ถ้าผลของสัมประสิทธิ์แอลฟามีค่าสูงแน่นอนว่าค่าความเชื่อมั่นของแบบทดสอบต้องมีค่าสูง ถ้าสัมประสิทธิ์แอลฟามีค่า

ต่ำ ก็ไม่อาจรู้ได้ว่าความเชื่อมั่นของแบบทดสอบต่ำหรือแบบทดสอบที่แบ่งครึ่งไม่คู่ขนานกันแบบ  
ทอ สูตรสำหรับสัมประสิทธิ์แอลฟาแบ่งครึ่งคือ

$$\alpha = \frac{2[\sigma_x^2 - (\sigma_{Y_1}^2 + \sigma_{Y_2}^2)]}{\sigma_x^2}$$

เมื่อ  $\sigma_{Y_1}^2$  และ  $\sigma_{Y_2}^2$  คือความแปรปรวนของคะแนนแบบทดสอบที่แบ่งครึ่งฉบับเป็น ส่วนที่ 1 และ  
ส่วนที่ 2 และ  $\sigma_x^2$  คือความแปรปรวนของคะแนนแบบทดสอบทั้งฉบับ

สมการคำนวณสัมประสิทธิ์แอลฟาและสูตรสเปียร์แมนบราวน์จะมีค่ามากถ้า  
แบบทดสอบที่แบ่งครึ่งมีความสัมพันธ์กันสูงและจะมีค่าต่ำเมื่อแบบทดสอบที่แบ่งครึ่งไม่มี  
ความสัมพันธ์กัน แบบทดสอบที่แบ่งครึ่งจะมีความสัมพันธ์กันสูงเมื่อแบบทดสอบวัดคุณลักษณะ  
เดียวกัน ดังนั้นความเชื่อมั่นแบบสเปียร์แมนบราวน์และสัมประสิทธิ์แอลฟาจะบ่งชี้ถึงแบบทดสอบ  
ที่มีความสอดคล้องภายในหรือเป็นเอกพันธ์กัน

ถ้าความแปรปรวนของคะแนนสังเกตของแบบทดสอบที่แบ่งครึ่งเท่ากัน สูตรสเปียร์  
แมนบราวน์และสูตรสัมประสิทธิ์แอลฟาจะมีค่าเท่ากัน ถ้าความแปรปรวนของคะแนนสังเกตของ  
แบบทดสอบที่แบ่งครึ่งเท่ากันแต่ครั้งนั้นไม่คู่ขนานกันแบบทอ ทั้งสูตรสเปียร์แมนบราวน์และ  
สัมประสิทธิ์แอลฟาจะมีค่าความเชื่อมั่นต่ำกว่าความจริง

การใช้การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งสามารถแสดงได้ด้วยตัวอย่างดังนี้

สมมติว่าสหสัมพันธ์ระหว่างแบบทดสอบครึ่งฉบับเป็น 0.5 ความแปรปรวนของคะแนน  
คือ 7 และ 5 และความแปรปรวนของคะแนนรวมคือ 17.9 ใช้สูตรสเปียร์แมนบราวน์คำนวณค่า  
ความเชื่อมั่นโดยรวมทั้งฉบับได้

$$r_{XX'} = \frac{2(0.5)}{1 + 0.5} = 0.67$$

ความเชื่อมั่นของแบบทดสอบทั้งฉบับประมาณค่าด้วยสัมประสิทธิ์แอลฟาได้ดังนี้

$$r_{XX'} = \frac{2[17.9 - (7 + 5)]}{17.9} = 0.66$$

ตัวอย่างนี้สัมประสิทธิ์แอลฟาประมาณค่าความเชื่อมั่นได้ต่ำกว่าสูตรสเปียร์แมนบราวน์  
เพียงเล็กน้อย

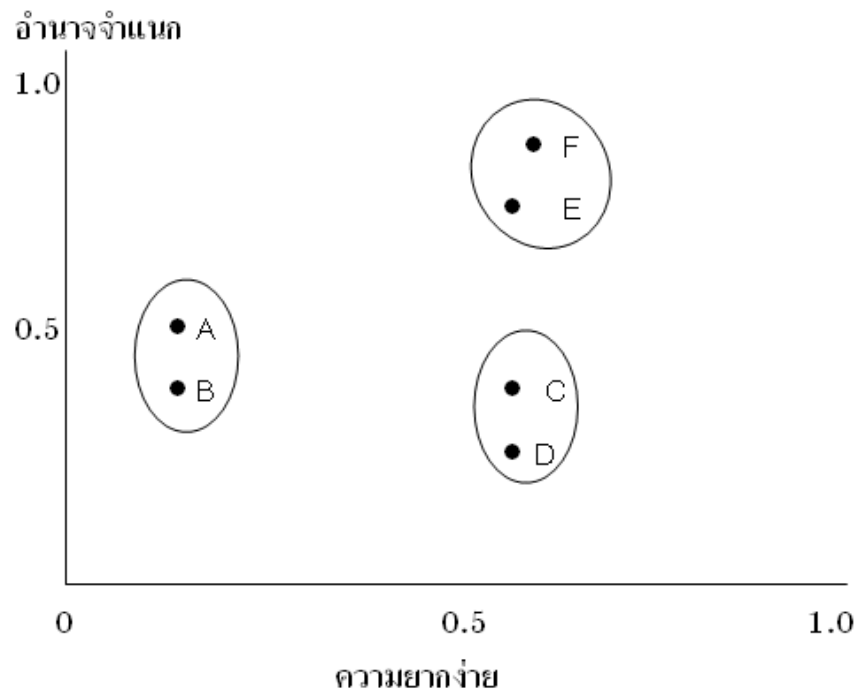
ประโยชน์หลักของการประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายในคือใช้  
แบบทดสอบเพียงฉบับเดียวสอบเพียงครั้งเดียว อย่างไรก็ตามวิธีหาความสอดคล้องภายในจะไม่  
เหมาะสมเมื่อแบบทดสอบไม่สามารถแบ่งออกเป็น ส่วน ๆ ที่คู่ขนานกันแบบพาราเรล (parallel)  
หรือแบบทอ (essentially  $\tau$ -equivalent) ได้ หรือเมื่อแบบทดสอบไม่มีข้อสอบที่เป็นอิสระจาก  
กันทำให้ไม่สามารถแบ่งออกเป็น ส่วน ๆ ได้ ตัวอย่างเช่น ในบางแบบทดสอบผู้สอบต้องจัดการกับ  
วัตถุในช่วงเวลาที่กำหนดไม่สามารถจะแยกออกเป็น ส่วน ๆ ได้ เพราะว่าการจัดการกับวัตถุในแต่ละ  
ระดับขึ้นอยู่กับเวลาและความคลาดเคลื่อนในขณะที่ทำงานกับวัตถุชิ้นอื่น ๆ ในสถานการณ์นี้ การ  
ประมาณค่าความเชื่อมั่นแบบสอบซ้ำกับการใช้แบบทดสอบทางเลือกจะเหมาะสมมากกว่า

มีสามวิธีในการแบ่งครึ่งแบบทดสอบที่นิยมใช้กัน วิธีแรกจะเรียกว่า วิธีแบ่งข้อคู่ข้อคี่ (odd/even method) เป็นการจัดแบ่งข้อสอบออกเป็นสองกลุ่มโดยอาศัยตัวเลขข้อเป็นหลักในการแบ่ง ผู้สอบแต่ละคนจะมีคะแนนของแบบทดสอบในกลุ่มข้อคู่และข้อคี่ วิธีที่สองคือแบบเรียงอันดับ (order method) คือจะแบ่งข้อสอบออกเป็นครึ่งแรกกับครึ่งหลัง ผู้สอบแต่ละคนจะมีคะแนนของแบบทดสอบในครึ่งแรกและครึ่งหลังของแบบทดสอบ โดยทั่วไปการแบ่งครึ่งแรกกับครึ่งหลังจะมีความเหมาะสมน้อยกว่าแบบแบ่งข้อคู่และข้อคี่ เพราะว่ามีผู้สอบบางคนอาจจะได้รับอิทธิพลของการฝึกฝนจากข้อสอบที่เพิ่งทำผ่านมา (มีอิทธิพลกับแบบทดสอบครึ่งหลังได้คะแนนสูงกว่าปกติ) และผู้สอบบางคนทำแบบทดสอบไม่เสร็จ (มีอิทธิพลกับแบบทดสอบครึ่งหลังได้คะแนนต่ำกว่าปกติ) อย่างไรก็ตาม ปัญหาของผู้สอบบางคนที่ไม่เสร็จในครึ่งหลังของแบบทดสอบสามารถแก้ไขได้ด้วยการแบ่งครึ่งเวลา นั่นคือผู้สอบทำแบบทดสอบไปจนเสร็จครึ่งแรกและเมื่อเวลาหมด ผู้สอบทั้งหมดจึงค่อยลงมือทำแบบทดสอบครึ่งที่สอง จะช่วยให้ผู้สอบได้ทำแบบทดสอบสมบูรณ์ในทั้งครึ่งแรกและครึ่งหลัง การแบ่งครึ่งชนิดนี้จะมีความเท่าเทียมกับการใช้แบบทดสอบทางเลือกฉบับสั้น 2 ฉบับ

วิธีที่สามสำหรับการแบ่งครึ่งแบบทดสอบให้เท่าเทียมกันนี้เป็นวิธีที่ใหม่กว่าสองวิธีแรก วิธีนี้เรียกว่าการจับคู่แบบทดสอบย่อยอย่างสุ่ม (matched random subsets) ซึ่งมีขั้นตอนดังต่อไปนี้

1. ต้องคำนวณสถิติของข้อสอบสองตัวคือ
  - 1.1 สัดส่วนของผู้สอบที่ทำข้อสอบนั้นถูก (ความยากง่ายของข้อสอบ)
  - 1.2 สหสัมพันธ์ไบซีเรียลหรือพอยท์ไบซีเรียลระหว่างคะแนนแบบทดสอบกับคะแนนรวม (อำนาจจำแนกของข้อสอบ)

2. ในข้อสอบแต่ละข้อพล็อตกราฟโดยใช้สถิติสองตัวนี้ ข้อสอบจะถูกจับคู่กันบนกราฟโดยสองจุดใด ๆ ที่อยู่ใกล้กันให้จับทั้งคู่สุ่มไปใส่ในกลุ่มครึ่งฉบับ ตัวอย่างในภาพประกอบ 4 แสดงข้อสอบ 6 ข้อที่ถูกพล็อตลงบนกราฟ และจับกลุ่มเป็นคู่ ถ้าข้อ A ถูกเลือกเข้ากลุ่มครึ่งแรกแล้ว ข้อสอบ B ก็จะไปปรากฏอยู่ในอีกครึ่งหนึ่ง ความเป็นไปได้ของแบบทดสอบที่จะถูกสุ่มเข้ากลุ่มเป็นดังนี้ ACE และ BDF, ADE และ BCF, ACF และ BDE และอื่น ๆ วิธีนี้จะช่วยให้แน่ใจว่าสองส่วนนั้นมีความยากง่ายและอำนาจจำแนกเหมือนกันและการวัดนั้นก็วัดในสิ่งเดียวกัน (ดังนั้นคะแนนจึงจึงเท่ากัน)



ภาพประกอบ 4.4 การเลือกแบบทดสอบย่อยโดยใช้วิธีจับคู่สุ่มด้วยกราฟ (Allen and Yen. 1979)

สำหรับแบบทดสอบใช้ความเร็ว (speed test) และแบบทดสอบที่ใช้เวลาในการคิดนาน (power test) นั้น แบบทดสอบที่ใช้ความเร็ว (speed test) ควรจะเป็นข้อสอบที่ผู้สอบทุก ๆ คน สามารถตอบได้หมดในเวลาที่มีพอเพียง แต่แบบทดสอบที่ให้เวลาน้อยเกินไปผู้สอบจะต้องพยายามตอบข้อสอบให้ได้โดยเร็ว ตัวอย่างเช่น แบบทดสอบที่ให้คำมาเป็นคู่ จำนวน 100 ข้อแล้วบอกถึงความแตกต่างควรจะทำให้เสร็จภายในเวลา 60 นาที และอีกแบบหนึ่งเป็นแบบทดสอบที่ต้องใช้ความสามารถมาก (power test) ซึ่งข้อสอบจะมีความยาก ให้เวลาไม่จำกัดในการสอบ การตอบได้หรือไม่ได้จึงขึ้นอยู่กับความสามารถของผู้สอบ สามารถตอบคำถามได้ถูกต้องเฉพาะข้อที่แน่ใจ การจำกัดเวลาในการสอบโดยทั่วไปต้องให้แน่ใจว่าผู้สอบแต่ละคนจะสามารถทำข้อสอบแต่ละข้อได้เสร็จ การทดสอบความสามารถหรือผลสัมฤทธิ์โดยมากมักจะใช้ทั้ง speed test และ power test

การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งไม่ควรใช้กับ speed test เพราะว่าผู้สอบโดยมากจะต้องพยายามตอบให้ถูกต้องภายในเวลาที่จำกัด ถ้าข้อสอบมี 30 ข้อการแบ่งข้อคู่ข้อคี่โดยปกติก็คือ 15 ข้อ และทั้งสองส่วนนี้ควรจะคู่ขนานกัน จะทำให้การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งมีค่าเข้าใกล้ 1 และถ้าการประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งอยู่บนพื้นฐานของการแบ่งเป็นครั้งแรกและครั้งหลังของแบบทดสอบ speed test การประมาณค่าความเชื่อมั่นจะเข้าใกล้ 0 ผู้สอบส่วนใหญ่ควรจะได้คะแนนดีมากในครั้งแรก และได้คะแนนไม่ดีในครั้งหลัง

การประมาณค่าความเชื่อมั่นด้วยวิธีการจับคู่ข้อสอบก็ไม่เหมาะสมกับ speed test เพราะว่าการยากของข้อสอบและความสัมพันธ์ของข้อสอบกับคะแนนรวมควรจะทำหน้าที่ในการบ่งบอกตำแหน่งของข้อสอบในแบบทดสอบมากกว่าจะบอกคุณลักษณะของข้อสอบ

### ความเชื่อมั่นแบบความสอดคล้องภายใน : กรณีทั่วไป

เทคนิคการแบ่งครึ่งข้อสอบ (แบ่งข้อคู่ข้อคี่ แบบเรียงอันดับ และแบบจับคู่อย่างสุ่ม) สามารถทำให้อยู่ในรูปทั่วไปโดยการแบ่งแบบทดสอบออกมากกว่าสองส่วน เช่น วิธีแบ่งข้อคู่ข้อคี่สามารถปรับใช้โดยการแบ่งออกเป็นสามส่วน สำหรับแบบทดสอบที่มี 9 ข้อ โดยอาจจะให้ข้อหนึ่งสี่ และเจ็ด เป็นส่วนแรก ข้อสอง ห้า และแปด เป็นส่วนที่สอง และข้อสาม หก และเก้า เป็นส่วนที่สาม วิธีการจับคู่อย่างสุ่มอาจจะใช้สามส่วน โดยการเลือกสามข้อที่อยู่ใกล้กันแล้วสุ่มแบ่งออกเป็นสามส่วน

ในหัวข้อนี้จะสมมติว่าแบบทดสอบถูกแบ่งออกเป็น  $N$  ส่วน ความแปรปรวนของคะแนนในแต่ละส่วนและความแปรปรวนของคะแนนรวมของแบบทดสอบจะใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ ถ้าในแต่ละส่วน (ข้อสอบ หรือชุดของข้อสอบ) มีความคู่ขนานกันแบบทอ ( $\tau$ -equivalent) สูตรที่นำเสนอในหัวข้อนี้จะให้ค่าความเชื่อมั่นของแบบทดสอบตามความเป็นจริง ถ้าแต่ละส่วนไม่คู่ขนานกันแบบทอ ( $\tau$ -equivalent) สูตรในหัวข้อนี้จะประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง นอกจากนี้ สูตรจะประมาณค่าความเชื่อมั่นได้ดีเมื่อแบบทดสอบวัดคุณลักษณะเดียว นั่นคือเมื่อแบบทดสอบมีเนื้อหาเป็นเอกพันธ์ (homogeneous) แต่แบบทดสอบวัดเชาวน์ปัญญาซึ่งวัดความสามารถทางภาษา มิติสัมพันธ์ และอื่น ๆ ควรจะเป็นวิวิธพันธ์ (heterogeneous) นั่นคือการวัดความเชื่อมั่นแบบความสอดคล้องภายในไม่เหมาะที่จะใช้กับแบบทดสอบที่เป็นวิวิธพันธ์

สูตรสำหรับความเชื่อมั่นแบบความสอดคล้องภายในกรณีทั่วไปคือสัมประสิทธิ์แอลฟา ( $\alpha$ -coefficient)

$$\alpha = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - \sum_{i=1}^N \sigma_{Y_i}^2}{\sigma_X^2} \right]$$

เมื่อ  $X$  คือ คะแนนรวมของแบบทดสอบที่รวมกัน  $N$  ส่วน ( $X = \sum_{i=1}^N Y_i$ )

$\sigma_X^2$  คือ ความแปรปรวนของแบบทดสอบที่รวมกัน  $N$  ส่วน

$\sigma_{Y_i}^2$  คือ ความแปรปรวนของส่วนที่  $i$ ,  $Y_i$

$N$  คือ จำนวนส่วนที่รวมกันเป็นคะแนน  $X$  เช่น ถ้า  $N = 3$  คะแนนของแบบทดสอบ  $X$  ก็จะมาจากผลรวมของคะแนนในสามส่วน

สัมประสิทธิ์แอลฟาโดยทั่วไป นิยมเขียนเป็นสมการว่า

$$\alpha = \left[ \frac{N}{N-1} \right] \left[ 1 - \frac{\sum_{i=1}^N \sigma_{Y_i}^2}{\sigma_X^2} \right]$$

ถ้าในแต่ละส่วนเป็นข้อสอบแบบ 0, 1 (dichotomous) สมการข้างต้นสามารถเขียนในรูปแบบเฉพาะได้ว่า

$$KR20 = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - \sum_{i=1}^N p_i(1-p_i)}{\sigma_X^2} \right]$$

เมื่อ  $p_i$  คือสัดส่วนของผู้สอบที่ตอบข้อสอบข้อที่  $i$  ได้ถูกต้องหรือก็คือความยากง่ายนั่นเอง สมการ KR20 ข้างต้นสะท้อนให้เห็นความแปรปรวนของคะแนนในข้อที่  $i$  เมื่อคะแนนของข้อสอบให้คะแนนเป็น 0, 1 เท่ากับ  $p_i(1-p_i)$  เมื่อ  $p_i$  คือสัดส่วนของผู้สอบที่ได้ 1 คะแนนในข้อ  $i$  (นั่นคือสอบผ่านในข้อนั้น) สมการ KR20 ข้างต้นก็คือสูตร Kuder-Richardson formula 20 เขียนย่อว่า KR20 เพราะว่า Kuder-Richardson นำเสนอสูตรนี้เป็นสูตรที่ 20 อีกชื่อหนึ่งของสูตรนี้ก็คือ coefficient  $\alpha$ -20 สูตร KR-20 โดยทั่วไปนิยมเขียนเป็นสมการว่า

$$KR20 = \left[ \frac{N}{N-1} \right] \left[ 1 - \frac{\sum_{i=1}^N p_i(1-p_i)}{\sigma_X^2} \right]$$

อีกสูตรหนึ่งของ Kuder-Richardson ก็คือสูตร KR21

$$KR21 = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - N\bar{p}(1-\bar{p})}{\sigma_X^2} \right]$$

เมื่อ  $\bar{p}$  คือค่าเฉลี่ยของความยากข้อสอบ เพราะว่า  $\bar{p}$  สามารถคำนวณได้โดยใช้  $\bar{p} = \Sigma(X)/N$  สมการ KR21 สามารถคำนวณโดยใช้ค่าเฉลี่ยและความแปรปรวนของข้อสอบ  $N$  ข้อในแบบทดสอบ ซึ่งเป็นกรณีเฉพาะของสูตร KR20 และสูตรนี้ นิยมเขียนเป็นสมการว่า

$$KR21 = \left[ \frac{N}{N-1} \right] \left[ 1 - \frac{\bar{X}(N-\bar{X})}{N\sigma_X^2} \right]$$

โดยปกติสูตร KR20 และ KR21 จะเกี่ยวข้องกันโดยที่  $KR20 \geq KR21$

ทั้งสองสูตรนี้จะมีค่าเท่ากันเมื่อความยากง่ายของข้อสอบเท่ากันทุกข้อ ถ้าข้อสอบมีค่าความยากง่ายไม่เท่ากันแล้ว KR21 จะประมาณค่าได้ต่ำกว่า KR20 และเป็นการประมาณค่าความเชื่อมั่นที่ต่ำกว่าความเป็นจริง

ผลต่างของสูตร KR20 และ KR21 เสนอโดย Tucker (1949) ดังนี้

$$KR20 - KR21 = \frac{n\sum(p_i - \bar{p})^2}{(n-1)\sigma_X^2} = \frac{n^2\sigma_{p_i}^2}{(n-1)\sigma_X^2}$$

ผลจากสูตร  $\alpha$  และ KR20 จะประมาณค่าได้ต่ำกว่าหรือเท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบ โดยจะประมาณค่าได้เท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบเมื่อในแต่ละองค์ประกอบ ( $Y_i$ ) มีความคู่ขนานกันแบบทอ ( $\tau$ -equivalent) (นั่นคือจำเป็นที่คะแนนจริงของแต่ละองค์ประกอบจะต้องเท่ากัน) ส่วน KR21 จะเท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบถ้าข้อสอบมีความยากง่ายเท่ากันและคู่ขนานกันแบบทอ ( $\tau$ -equivalent) และทั้งสามสูตรที่นำเสนอข้างต้นนี้จะให้ค่าความเชื่อมั่นสูงถ้าคะแนนในแต่ละส่วนกับคะแนนรวมมีความสัมพันธ์กันสูง และจะให้ค่าความเชื่อมั่นต่ำถ้าคะแนนในแต่ละส่วนกับคะแนนรวมมีความสัมพันธ์กันต่ำ ในแต่ละส่วนจะมีความสัมพันธ์กันสูงถ้าแบบทดสอบนั้นวัดคุณลักษณะเดียวกัน ดังนั้นสูตรที่นำเสนอในหัวข้อนี้จะเป็นตัวบ่งชี้ถึงความสอดคล้องภายในของแบบทดสอบหรือความเป็นเอกพันธ์ของแบบทดสอบ

#### ตัวอย่างคำนวณ

ต่อไปนี้จะเป็นอย่างในการใช้สูตรที่นำเสนอในหัวข้อนี้มาประมาณค่าความเชื่อมั่นแบบความสอดคล้องภายใน

ตาราง 4.2 ข้อมูลสำหรับประมาณค่าความเชื่อมั่น

ผู้สอบ	ข้อสอบ						รวม
	1	2	3	4	5	6	
1	0	0	0	0	0	0	0
2	0	0	0	0	1	0	1
3	1	0	1	1	1	0	4
4	1	1	1	1	1	1	6
5	1	1	1	1	1	1	6
6	0	0	1	0	0	0	1
7	0	0	1	1	1	0	3
8	0	0	0	1	0	0	1
9	1	0	1	1	1	0	4
10	0	1	0	1	0	1	3
ค่าเฉลี่ย							2.9
ส่วนเบี่ยงเบนมาตรฐาน							2.02



1. กรณีแบบทดสอบถูกแบ่งครั้งออกเป็น 2 ส่วน โดยส่วนแรกประกอบด้วยข้อ 1 - 3 และส่วนที่สองประกอบด้วยข้อ 4 - 6 และสมมติทั้ง 2 ส่วนคู่ขนานกันแบบพาราเรล (Parallel) สัมประสิทธิ์สหสัมพันธ์ระหว่างสองส่วนเท่ากับ 0.82 ประเมินค่าความเชื่อมั่นด้วยสูตรของ สเปียร์แมนบราวน์

$$SB\rho_{XX'} = \frac{2\rho_{yy'}}{1+\rho_{yy'}} = \frac{2(0.82)}{1+0.82} = 0.90$$

2. กรณีแบบทดสอบถูกแบ่งออก 6 ส่วน โดยแต่ละข้อถือเป็น 1 ส่วน และทุกข้อ คู่ขนานกันแบบทอ ประเมินค่าสัมประสิทธิ์แอลฟา KR20 และ KR21

$$\alpha = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - \sum_{i=1}^N \sigma_{Y_i}^2}{\sigma_X^2} \right] = \left[ \frac{6}{6-1} \right] \left[ \frac{4.54 - 1.5}{4.54} \right] = 0.80$$

$$KR20 = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - \sum_{i=1}^N p_i(1-p_i)}{\sigma_X^2} \right] = \left[ \frac{6}{6-1} \right] \left[ \frac{4.54 - 1.35}{4.54} \right] = 0.84$$

$$KR21 = \left[ \frac{N}{N-1} \right] \left[ \frac{\sigma_X^2 - N\bar{p}(1-\bar{p})}{\sigma_X^2} \right] = \left[ \frac{6}{6-1} \right] \left[ \frac{4.54 - 6(0.48)(0.52)}{4.54} \right] = 0.80$$

ผลต่างของ KR20 และ KR21 คำนวณได้ด้วยสูตรของทักเกอร์ (Tucker) ดังนี้

$$KR20 - KR21 = \frac{n\sum(p_i - \bar{p})^2}{(n-1)\sigma_X^2} = \frac{10(0.148)}{(10-1)4.54} = 0.04$$

### สูตรสเปียร์แมนบราวน์ : กรณีทั่วไป

อีกวิธีการหนึ่งในการประเมินค่าความเชื่อมั่นของแบบทดสอบก็คือสูตรสเปียร์แมนบราวน์ สูตรสเปียร์แมนบราวน์จะใช้ในการทำนายอิทธิพลที่เปลี่ยนแปลงไปของความยาวของแบบทดสอบที่จะมีต่อค่าความเชื่อมั่น สูตรนี้จะอ้างอิงกับการประเมินค่าความเชื่อมั่นแบบแบ่งครึ่ง สูตรสเปียร์แมนบราวน์โดยทั่วไปคือ

$$\rho_{XX'} = \frac{N\rho_{YY'}}{1 + (N-1)\rho_{YY'}}$$

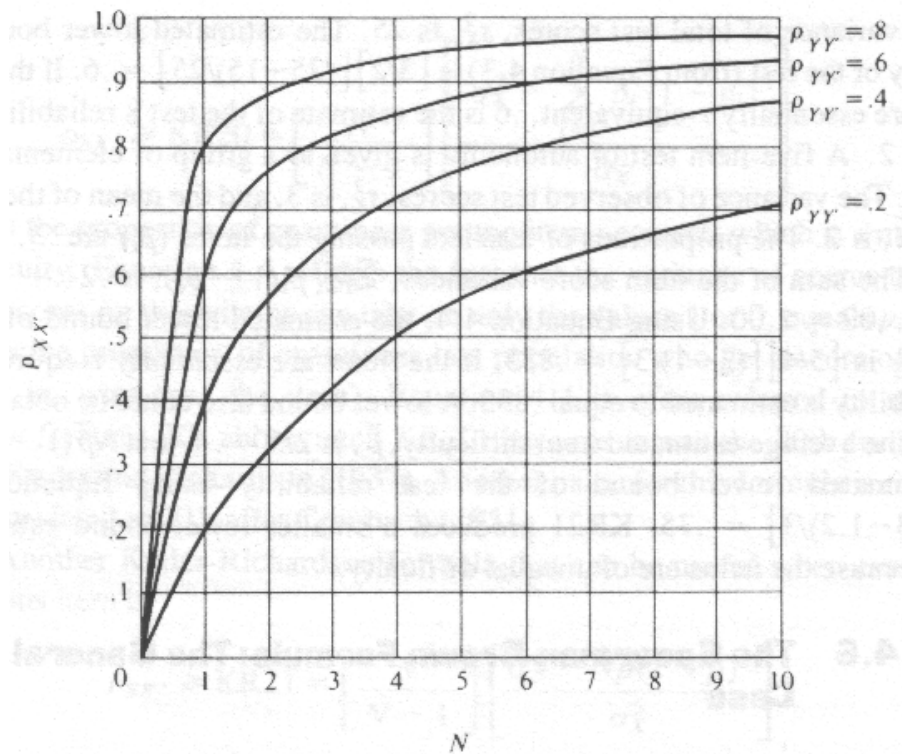
เมื่อ  $Y_i$  คือคะแนนของแบบทดสอบฉบับดั้งเดิม

$\rho_{XX'}$  คือความเชื่อมั่นของแบบทดสอบฉบับใหม่ (X)

$\rho_{YY'}$  คือความเชื่อมั่นของแบบทดสอบฉบับดั้งเดิม ( $Y_i$ ) และ

N คือจำนวนเท่าของข้อสอบที่เพิ่มเข้าไปโดยมีความคู่ขนานกันแบบ

พาราเรลกับข้อสอบเก่า และรวมกันเป็นฉบับ X



ภาพประกอบ 4.5 ความสัมพันธ์ระหว่างความยาวของแบบทดสอบและความเชื่อมั่น  
(Allen and Yen. 1979)

สูตรสเปียร์แมนบราวน์จะแสดงค่าความเชื่อมั่นของแบบทดสอบ  $\rho_{XX'}$  ในเทอมของความเชื่อมั่นในแต่ละข้อที่คู่ขนานกันของแบบทดสอบ สังเกตว่าในสูตรนี้  $\rho_{XX'}$  จะมีค่ามากกว่าหรือเท่ากับ  $\rho_{YY'}$  เสมอ ค่าความเชื่อมั่นของแบบทดสอบที่เกิดจากการรวมข้อสอบหรือองค์ประกอบที่คู่ขนานกันต้องมีค่ามากกว่าหรือเท่ากับค่าความเชื่อมั่นขององค์ประกอบใดองค์ประกอบหนึ่ง ซึ่งค่าความเชื่อมั่น  $\rho_{XX'}$  ในบางครั้งเรียกว่า Stepped-up reliability เพราะคือการปรับแก้ค่าความเชื่อมั่นให้สูงขึ้นจากความเชื่อมั่นในฉบับที่สั้นกว่า

ภาพประกอบ 4.5 จะแสดงอิทธิพลโดยทั่วไปของการเปลี่ยนแปลงความยาวแบบทดสอบที่จะมีผลกับ  $\rho_{XX'}$  สำหรับแบบทดสอบที่แต่ละองค์ประกอบมีค่าความเชื่อมั่น  $\rho_{YY'}$  เป็น 0.2, 0.4, 0.6 และ 0.8 เมื่อเราทราบค่าของสองจำนวนจากสามจำนวนคือ  $N$ ,  $\rho_{YY'}$  และ  $\rho_{XX'}$  แล้ว เราสามารถหาจำนวนที่สามได้ ถ้า  $\rho_{YY'} = 0.4$  และ  $\rho_{XX'} = 0.8$  แบบทดสอบนี้จะยาวกว่าแบบทดสอบเดิม 6 เท่า ( $N = 6$ ) ภาพประกอบ 4.5 จะแสดงการเพิ่มขึ้นของความยาวแบบทดสอบที่สัมพันธ์กับความเชื่อมั่นของแบบทดสอบ เมื่อ  $N \rightarrow \infty$  แล้ว ความเชื่อมั่นของแบบทดสอบที่มี  $N$  ข้อหรือองค์ประกอบจะมีค่าถึง 1.0 โดยที่  $\rho_{YY'} \neq 0$

สมการข้างต้นจะใช้เมื่อเรารู้  $N$  และ  $\rho_{YY'}$  และต้องการหาค่า  $\rho_{XX'}$  เมื่อเรารู้ความเชื่อมั่นของแบบทดสอบ  $\rho_{XX'}$  และต้องการกำหนดความเชื่อมั่นของแบบทดสอบในฉบับที่สั้นกว่า ( $\rho_{YY'}$ ) และมีความยาวเป็น  $1/N$  เท่าของฉบับเต็ม สูตรสเปียร์แมนบราวน์จะเขียนได้ใหม่เป็น

$$\rho_{YY'} = \frac{\frac{1}{N}\rho_{XX'}}{1 + \left(\frac{1}{N} - 1\right)\rho_{XX'}}$$

ในกรณีที่รู้ค่า  $\rho_{YY'}$  และ  $\rho_{XX'}$  เราสามารถแก้สมการของสเปียร์แมนบราวน์ใหม่เป็นสูตรในการคำนวณหาจำนวนเท่าของข้อสอบที่เพิ่มเข้าไป ได้ว่า

$$N = \frac{\rho_{XX'}(1 - \rho_{YY'})}{\rho_{YY'}(1 - \rho_{XX'})}$$

ต่อไปนี้จะแสดงตัวอย่างของการประยุกต์ใช้สูตรสเปียร์แมนบราวน์ในการประมาณค่าความเชื่อมั่น ( $r_{XX'}$  และ  $r_{YY'}$ ) แทนค่าความเชื่อมั่นของประชากร ( $\rho_{XX'}$  และ  $\rho_{YY'}$ ) การประมาณค่าความเชื่อมั่นนี้ สามารถใช้ควบคู่ไปกับการประมาณค่าความเชื่อมั่นแบบสอบซ้ำ แบบคู่ขนาน แบบทดสอบทางเลือก หรือแบบความสอดคล้องภายใน

1. คุณมีแบบทดสอบที่ใช้เวลาในการสอบเพียง 5 นาทีและประมาณค่าความเชื่อมั่นได้ 0.6 ถ้าคุณต้องการเพิ่มแบบทดสอบอีก 3 เท่าที่คู่ขนาน ความเชื่อมั่นของแบบทดสอบที่ยาวกว่าคืออะไร ในที่นี้  $N = 3$ ,  $r_{YY'} = 0.6$  ใช้สมการสเปียร์แมนบราวน์คำนวณได้

$$r_{XX'} = \frac{3(0.6)}{1 + (2)(0.6)} = 0.82$$

ความเชื่อมั่นนี้อ้างอิงได้จากภาพประกอบ 4.5

2. คุณมีข้อสอบ 50 ข้อที่ประมาณค่าความเชื่อมั่นได้ 0.9 ถ้าคุณเอาข้อสอบออกมา 10 ข้อแล้ว ความเชื่อมั่นของแบบทดสอบ 10 ข้อนั้นคือเท่าไร ในที่นี้  $N = 5$ ,  $r_{XX'} = 0.9$  และเราจะคำนวณ  $r_{YY'}$  ได้ดังนี้

$$r_{YY'} = \frac{\frac{1}{5}(0.9)}{1 + \left(\frac{1}{5} - 1\right)(0.9)} = 0.64$$

3. คุณมีแบบทดสอบฉบับสั้น 10 ข้อที่มีค่าความเชื่อมั่น 0.8 แบบทดสอบควรจะยาวเท่าไรจึงจะมีค่าความเชื่อมั่น 0.9 ในที่นี้  $r_{YY'} = 0.8$ ,  $r_{XX'} = 0.9$  และเราจะคำนวณหาความยาวของแบบทดสอบได้ดังนี้

$$N = \frac{(0.9)(1 - 0.8)}{(0.8)(1 - 0.9)} = 2.25$$

แบบทดสอบใหม่ควรจะมีความยาวเป็น 2.25 เท่าของแบบทดสอบเดิมหรือก็คือ 23 ข้อ

สูตรสเปียร์แมนบรวานี่จะอยู่บนพื้นฐานของข้อสอบหรือองค์ประกอบในแบบทดสอบ ต้องคู่ขนานกันแบบพาราเรล รวมทั้งชุดของข้อสอบหรือองค์ประกอบที่เพิ่มเข้าไปในแบบทดสอบด้วย ถ้าเพิ่มอย่างระมัดระวังข้อสอบที่เพิ่มมีความคู่ขนานกับข้อสอบในฉบับเดิม ความเชื่อมั่นควรจะสูงขึ้น แต่ถ้าเพิ่มอย่างไม่ระวัง ความเชื่อมั่นอาจจะลดต่ำลง อย่างไรก็ตาม แบบทดสอบที่ยาวกว่าย่อมมีความเชื่อมั่นสูงกว่า เพราะว่าข้อตกลงเบื้องต้นของทฤษฎีคะแนนจริงมาตรฐานเดิม เมื่อ  $N$  เพิ่มขึ้น ความแปรปรวนของคะแนนจริงจะเพิ่มมากกว่าความแปรปรวนของความคลาดเคลื่อน

ถ้าข้อสอบหรือองค์ประกอบของแบบทดสอบไม่คู่ขนานกันแล้ว สูตรสเปียร์แมนบรวานี่จะประมาณค่าได้ต่ำกว่าหรือสูงกว่าความเป็นจริง ตัวอย่างเช่น แบบทดสอบ 10 ข้อคำนวณค่าความเชื่อมั่นได้ 0.6 เมื่อเพิ่มข้อสอบที่คู่ขนานกับข้อสอบเดิมอีกเท่าตัว จะได้ค่าความเชื่อมั่น  $[2(0.6)]/[1+(0.6)] = 0.75$  อย่างไรก็ตาม ถ้าข้อสอบที่เพิ่มเข้าไปไม่คู่ขนานกับข้อสอบเดิม โดยข้อสอบ 10 ข้อใหม่ที่เพิ่มเข้าป็นั้นไม่มีความแปรปรวน จึงไม่มีผลต่อคะแนนของผู้สอบ จึงไม่ช่วยเพิ่มค่าความเชื่อมั่นให้สูงขึ้น ในกรณีนี้ข้อสอบ 20 ข้อจะได้ค่าความเชื่อมั่นเท่ากับ 0.6 (เท่ากับแบบทดสอบฉบับเดิม)

ในอีกสถานการณ์หนึ่ง สูตรสเปียร์แมนบรวานี่สามารถประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง ตัวอย่างเช่น สมมติว่าข้อสอบ 10 ข้อมีความเชื่อมั่น 0.0 สูตรสเปียร์แมนบรวานี่คำนวณค่าความเชื่อมั่นเมื่อเพิ่มข้อสอบอีกเท่าตัวที่คู่ขนานกับข้อสอบเดิมได้  $[2(0.0)]/[1+(0.0)] = 0.0$  อย่างไรก็ตาม ถ้าข้อสอบที่เพิ่มเข้าไปไม่มีความคู่ขนานกับแบบทดสอบเดิมแล้ว ความเชื่อมั่นใหม่ที่คำนวณได้จะต่ำกว่า 0.7 ความเชื่อมั่นของข้อสอบฉบับใหม่ 20 ข้อจะประมาณค่าได้มากกว่า 0.0 ในกรณีนี้การใช้สูตรสเปียร์แมนบรวานี่ที่ไม่เหมาะสมจะทำให้ประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง ผลของความเชื่อมั่นที่ใช้สูตรสเปียร์แมนบรวานี่จะมีความถูกต้องเมื่อข้อสอบหรือองค์ประกอบที่เพิ่มเข้าไปมีความคู่ขนานกัน

สามารถประยุกต์ใช้สูตรสเปียร์แมนบรวานี่ได้ เมื่อต้องการเปรียบเทียบความเชื่อมั่นของแบบทดสอบสองฉบับ แบบทดสอบที่ยาวมากกว่าดูเหมือนจะมีค่าความเชื่อมั่นสูงกว่า การประยุกต์ใช้สูตรสเปียร์แมนบรวานี่จะช่วยให้เราประมาณค่าความเชื่อมั่นของแบบทดสอบ ถ้าแบบทดสอบสองฉบับนั้นมีความยาวเท่ากัน สำหรับแบบทดสอบที่สั้นมากมีแนวโน้มจะมีค่าความเชื่อมั่นที่ต่ำกว่า ซึ่งควรจะระมัดระวังเมื่อมีการเปรียบเทียบความเชื่อมั่นของแบบทดสอบฉบับที่สั้นกว่า

### เปรียบเทียบวิธีการประมาณค่าความเชื่อมั่น

เราได้อธิบายความแตกต่างของวิธีการประมาณค่าความเชื่อมั่นของแบบทดสอบและวิธีการที่แตกต่างกันนี้จะคำนวณค่าความเชื่อมั่นได้ต่างกัน ดังนั้นการนำไปใช้มีข้อสังเกตดังนี้

1. การประมาณค่าความเชื่อมั่นของแบบทดสอบ speed test นั้น ควรจะใช้การประมาณค่าความเชื่อมั่นแบบสอบซ้ำ แบบทดสอบทางเลือกหรือแบบทดสอบคู่ขนาน การใช้สัมประสิทธิ์แอลฟาและสูตรคูเดอริชาร์ดสันจะให้ค่าความเชื่อมั่นขั้นต่ำ ซึ่งค่าความเชื่อมั่นขั้นต่ำนี้จะเท่ากับค่าความเชื่อมั่นของแบบทดสอบ

2. ถ้าข้อสอบหรือแต่ละส่วนในแบบทดสอบมีความคู่ขนานกันแบบทอ (essentially  $t$ -equivalent) สัมประสิทธิ์แอลฟาและสูตรคูเดอริชาร์ดสันก็จะเหมาะสม แต่ข้อสอบในแบบทดสอบควรจะมีความเป็นเอกพันธ์กัน ถ้าแบบทดสอบไม่มีความเป็นเอกพันธ์คือมีการวัดในหลายคุณลักษณะแล้ว สูตรสัมประสิทธิ์แอลฟาและคูเดอริชาร์ดสันก็จะไม่เหมาะสม

3. สูตรสเปียร์แมนบราวน์สามารถประมาณค่าได้ต่ำกว่าหรือสูงกว่าความเป็นจริงถ้าข้อสอบหรือองค์ประกอบในแบบทดสอบไม่คู่ขนานกันแบบพาราเรล เมื่อข้อสอบหรือองค์ประกอบในแบบทดสอบมีความคู่ขนานกันแบบพาราเรลแล้ว สูตรสเปียร์แมนบราวน์จะมีประโยชน์ในการประมาณค่าความเชื่อมั่นของแบบทดสอบและประมาณค่าความเชื่อมั่นเมื่อความยาวของแบบทดสอบเปลี่ยนไป

### ความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์

ผลการวัดควรจะต้องมีความถูกต้องแม่นยำ ปราศจากความคลาดเคลื่อน มีผู้เสนอวิธีการคำนวณหาความเชื่อมั่นของข้อสอบอิงเกณฑ์หลายวิธีดังนี้

1. วิธีของคาร์เวอร์ (Carver) ใช้แบบทดสอบคู่ขนาน โดยให้ข้อสอบคล้ายกันข้อต่อข้อ แล้วนำไปสอบนักเรียน แล้วนำข้อมูลที่ได้มาสร้างตารางดังนี้

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

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

**ตัวอย่างคำนวณ**

นักเรียน 8 คน ทำแบบทดสอบอิงเกณฑ์ 2 ฉบับ ฉบับละ 10 ข้อ คะแนนจุดตัดคือ 5 คะแนน มีผลการสอบดังนี้

คนที่	1	2	3	4	5	6	7	8
ฉบับที่ 1	7	8	9	4	5	3	4	5
ฉบับที่ 2	6	8	7	5	6	4	4	3

คำนวณค่าความเชื่อมั่นของแบบทดสอบอิงเกณฑ์ ดังตาราง

	ฉบับ 2	
ฉบับ 1	สอบไม่ผ่าน	สอบผ่าน
สอบผ่าน	1	4
สอบไม่ผ่าน	2	1

พบว่าผู้สอบผ่านทั้ง 2 ฉบับจำนวน 4 คน และสอบไม่ผ่านทั้ง 2 ฉบับจำนวน 2 คน แทนค่าคำนวณค่าความเชื่อมั่นดังนี้

$$\text{ค่าความเชื่อมั่น} = \frac{4+2}{8} = 0.75$$

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)

$\sigma^2$  คือความแปรปรวนของคะแนนสอบทั้งฉบับ

$\bar{X}$  คือคะแนนเฉลี่ยของคะแนนสอบทั้งฉบับ

$c$  คือคะแนนจุดตัด

วิธีนี้ใช้ได้ผลดีเมื่อคะแนนมีการกระจายแบบฐานนิยมเดียว

**ตัวอย่างคำนวณ**

ผลการสอบข้อสอบจำนวน 10 ข้อ มีคะแนนเฉลี่ย 6.4 และความแปรปรวนคือ 1.8 มีคะแนนจุดตัดคือ 5 มีค่าความเชื่อมั่นที่คำนวณจากสูตร KR-20 คือ 0.86 คำนวณค่าความเชื่อมั่นแบบอิงเกณฑ์ได้ดังนี้

$$\begin{aligned}
 r_{cc} &= \frac{0.86(1.8) + (6.4 - 5)^2}{1.8 + (6.4 - 5)^2} \\
 &= \frac{1.548 + 1.96}{1.8 + 1.96} \\
 &= 3.508 / 3.76 \\
 &= 0.9330
 \end{aligned}$$

3. วิธีของสวามินาทานและคณะ (Swaminathan) วิธีนี้ใช้แบบทดสอบฉบับเดียวแต่สอบซ้ำ 2 ครั้งหลังจากสิ้นสุดการสอน เพื่อดูความคงเส้นคงวา

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

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

เมื่อ

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

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

**ตัวอย่างคำนวณ**

จากข้อมูลในตัวอย่างคำนวณค่าความเชื่อมั่นด้วยวิธีของคาร์เวอร์

	ฉบับ 2	
ฉบับ 1	สอบผ่าน	สอบไม่ผ่าน
สอบผ่าน	4	1
สอบไม่ผ่าน	1	2

คำนวณค่าความเชื่อมั่นด้วยวิธีสอบสวามินาธานและคณะ ได้ดังนี้

$$P_o = \frac{4+2}{8} = 0.75$$

$$P_e = \frac{(4+1)(4+1)+(1+2)(1+2)}{(4+1+1+2)^2} = \frac{25+9}{64} = 0.53$$

$$\begin{aligned} \text{ค่าความเชื่อมั่น (K)} &= \frac{0.75-0.53}{1-0.53} \\ &= \frac{0.22}{0.47} = 0.468 \end{aligned}$$

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

**ตัวอย่างคำนวณ**

แบบทดสอบอิงเกณฑ์ฉบับหนึ่งมี 10 ข้อ มีคะแนนจุดตัดที่ 5 คะแนน ไปสอบกับนักเรียนจำนวน 8 คน ปรากฏผลดังนี้

คนที่	1	2	3	4	5	6	7	8	ผลรวม
คะแนน (X)	7	8	9	4	5	3	4	5	45
$X^2$	49	64	81	16	25	9	16	25	285
$(X - c)$	2	3	4	-1	0	-2	-1	0	5
$(X - c)^2$	4	9	16	1	0	4	1	0	35

แทนค่าในสูตรคำนวณค่าความเชื่อมั่นของโลเวทท์ ได้ค่าดังนี้

$$\begin{aligned} r_{cc} &= 1 - \frac{10(45) - (285)}{(10-1)(35)} \\ &= 1 - \frac{165}{315} \\ &= 1 - 0.5238 \\ &= 0.4762 \end{aligned}$$



## อิทธิพลที่มีต่อความเชื่อมั่น

### ความเป็นเอกพันธ์ของกลุ่มตัวอย่าง

ขนาดของสัมประสิทธิ์ความเชื่อมั่นขึ้นอยู่กับความแปรปรวนของคะแนนจริงและคะแนนคลาดเคลื่อนของแต่ละบุคคล ดังนั้น ความเป็นเอกพันธ์ของกลุ่มผู้สอบมีความสำคัญในการพัฒนาแบบทดสอบและเลือกใช้แบบทดสอบ สมมติว่าแบบทดสอบที่พัฒนาขึ้นวัดความวิตกกังวลในวิชาคณิตศาสตร์ ถ้าแบบทดสอบนี้ใช้กับกลุ่มตัวอย่างกลุ่มหนึ่ง que เลือกเรียนคณิตศาสตร์แล้วแน่นอนว่านักเรียนแต่ละคนจะต้องได้คะแนนความวิตกกังวลในระดับต่ำ ดังนั้นความแปรปรวนของคะแนนจริงจะต่ำและความเชื่อมั่นก็จะต่ำด้วย ถ้าใช้แบบทดสอบเดียวกันนี้กับอีกกลุ่มตัวอย่างหนึ่งที่มีลักษณะแตกต่างกันออกไป คะแนนจริงจะมีความแปรปรวนสูงมาก สมมติว่าควบคุมความแปรปรวนของความคลาดเคลื่อนให้คงที่ และขนาดของกลุ่มตัวอย่างเท่ากัน สัมประสิทธิ์ความเชื่อมั่นของกลุ่มที่สองย่อมสูงกว่ากลุ่มแรก ตาราง 4.3 แสดงสถานการณ์ในตัวอย่างนี้ ในการใช้แบบทดสอบที่มีประสิทธิภาพนั้น จะต้องแสดงวิธีการหาความเชื่อมั่นในคู่มือสอบด้วยและแสดงค่าสถิติต่าง ๆ ของกลุ่มที่ใช้ ถ้าแบบทดสอบที่ผลิตออกมาได้ใช้กลุ่มผู้สอบที่มีความเป็นวิวิธพันธ์กันมาก ความเชื่อมั่นก็จะมีค่าสูง และความเชื่อมั่นจะมีค่าลดลงเมื่อนำแบบทดสอบไปใช้กับกลุ่มตัวอย่างที่เป็นเอกพันธ์มาก

ตาราง 4.3 สัมประสิทธิ์ความเชื่อมั่นของกลุ่มตัวอย่าง 2 กลุ่มที่แตกต่างกันในความแปรปรวนของคะแนนจริง

	กลุ่ม 1	กลุ่ม 2
ความแปรปรวนของคะแนนจริง	20	60
ความแปรปรวนของคะแนนคลาดเคลื่อน	10	10
ความแปรปรวนของคะแนนสังเกต	30	70
สัมประสิทธิ์ความเชื่อมั่น	.67	.85

แมกนีสัน (1967) ได้เสนอสูตรสำหรับการทำนายค่าความเชื่อมั่นที่เปลี่ยนไปเมื่อความแปรปรวนของกลุ่มตัวอย่างเปลี่ยนไป

$$\hat{\rho}_{UU'} = 1 - \frac{\sigma_X^2(1 - \hat{\rho}_{XX'})}{\sigma_U^2}$$

เมื่อ  $\sigma_U^2$  คือความแปรปรวนของกลุ่มใหม่  $\sigma_X^2$  คือความแปรปรวนของกลุ่มเดิม  $\hat{\rho}_{XX'}$  คือความเชื่อมั่นของกลุ่มเดิม และ  $\hat{\rho}_{UU'}$  คือค่าความเชื่อมั่นใหม่ที่ถูกทำนาย สิ่งสำคัญที่ควรสังเกตคือ สูตรนี้ได้มีข้อตกลงว่า ความแปรปรวนของคะแนนความคลาดเคลื่อนเท่ากันในทุก

สองกลุ่ม และความเปลี่ยนแปลงของคะแนนที่สังเกตได้เนื่องมาจากความแตกต่างของคะแนนจริง ผู้ที่ใช้แบบทดสอบสามารถตรวจสอบข้อตกลงได้เมื่อนำแบบทดสอบไปใช้กับกลุ่มตัวอย่างและตรวจสอบเชิงประจักษ์โดยประมาณค่าความเชื่อมั่นของกลุ่มตัวอย่างใหม่

### ความยาวของแบบทดสอบ

คุณลักษณะหนึ่งของแบบทดสอบที่มีอิทธิพลต่อความแปรปรวนของคะแนนจริงและความแปรปรวนของคะแนนที่สังเกตได้คือ ความยาวของแบบทดสอบ ลองพิจารณาในกรณีที่ผู้สอบใช้แบบทดสอบที่มีข้อสอบ 1 ข้อ กับที่ใช้ข้อสอบ 10 ข้อวัดเนื้อหาเดียว แบบทดสอบฉบับไหนจะน่าเชื่อถือได้มากกว่ากัน คำตอบย่อมเป็นแบบทดสอบที่ใช้จำนวนข้อมาก ๆ ความสัมพันธ์ระหว่างความยาวของแบบทดสอบและความเชื่อมั่นแสดงดังสูตร

$$\rho_{YY'} = \frac{N\rho_{XX'}}{1 + (N-1)\rho_{XX'}}$$

เมื่อ  $\rho_{XX'}$  คือความเชื่อมั่นของแบบทดสอบเดิม  $N$  คือจำนวนเท่าของข้อสอบที่เพิ่มเข้าไปในแบบทดสอบ  $\rho_{YY'}$  คือความเชื่อมั่นใหม่ ดังนั้น ถ้าความเชื่อมั่นของแบบทดสอบย่อยฉบับหนึ่งมีค่าเป็น 0.75 ความเชื่อมั่นของแบบทดสอบที่เพิ่มจำนวนข้อสอบเข้าไป 5 เท่า จะคำนวณได้ค่าความเชื่อมั่น

$$\rho_{XX'} = \frac{5(.75)}{1 + (5-1)(.75)} = 0.94$$

หรือถ้าแบบทดสอบย่อย  $X$  มี 50 ข้อแบบทดสอบย่อยใหม่มี 150 ข้อ แบบทดสอบย่อยใหม่มีจำนวนข้อเพิ่มเป็น 3 เท่าของฉบับเดิม หรือ  $N = 3$  นอกจากนี้จำนวนเท่า  $N$  ไม่จำเป็นต้องเป็นจำนวนเต็มเสมอไป และไม่จำเป็นต้องมากกว่า 1.00 เช่น แบบทดสอบ 100 ข้อ และพัฒนาแบบทดสอบต่อไปจนเหลือแค่เพียง 75 ข้อ และ  $N$  คือความยาวของแบบทดสอบที่ต้องคุณเมื่อมีความยาวใหม่  $100(N) = 75$  ข้อ  $N = 0.75$  ดังนั้นถ้า  $\rho_{XX'} = 0.60$  ความเชื่อมั่นสำหรับแบบทดสอบฉบับสั้นกว่าคือ

$$\rho_{XX'} = \frac{0.75(0.60)}{1 + (0.75-1)(0.60)} = 0.53$$

เมื่อความยาวของแบบทดสอบเพิ่มขึ้น ค่า  $N$  ก็จะมีค่าเกิน 1 เสมอ เมื่อความยาวของแบบทดสอบลดลง ค่า  $N$  มีค่าต่ำกว่า 1 เสมอ สังเกตว่าการเพิ่มขึ้นของความเชื่อมั่นมีผลมาจากการเพิ่มขึ้นของความยาวข้อสอบ นั่นคือถ้าเพิ่มความยาวเท่าตัวของแบบทดสอบที่มีความเชื่อมั่น 0.60 แล้ว ความเชื่อมั่นจะเพิ่มเป็น 0.75 ถ้าเพิ่มเป็นสามเท่า ความเชื่อมั่นจะเพิ่มเป็น 0.81 ถ้าเพิ่มถึงห้าเท่า ความเชื่อมั่นจะเพิ่มเป็น 0.88 ดังนั้นการเพิ่มข้อสอบเพื่อให้ได้ความเชื่อมั่นเพิ่มขึ้น บางครั้งก็ต้องพิจารณาถึงต้นทุนในการเขียนและการใช้แบบทดสอบด้วย

### การจำกัดเวลาในการสอบ

เมื่อการทดสอบมีการจำกัดเวลาที่ตายตัวเช่น กลุ่มผู้สอบจำนวนหนึ่งทำข้อสอบเสร็จ แต่คนอื่น ๆ ยังทำไม่เสร็จ เวลาที่ให้ในการทำข้อสอบจะมีอิทธิพลต่อความคลาดเคลื่อนอย่างเป็นระบบต่อการทำข้อสอบของนักเรียนทั้งหมด ดังนั้น ความแปรปรวนจากระยะเวลาที่ผู้สอบได้ทำข้อสอบกลายเป็นส่วนหนึ่งของความแปรปรวนของคะแนนจริง ในบางแบบทดสอบ (เช่น แบบทดสอบวัดความถนัด) เป้าหมายในการทำข้อสอบต้องการประเมินความสามารถในทางปฏิบัติที่วางไว้ในการทำงาน ระยะเวลาในการตอบจะสั้น การให้เวลาในการตอบข้อสอบขึ้นอยู่กับลักษณะที่ถูกวัดด้วย ในแบบทดสอบหลาย ๆ ชนิดที่มีการจำกัดเวลา ซึ่งควรจะให้เวลานานเพียงพอที่เด็กส่วนใหญ่สามารถทำเสร็จได้พอดีเวลา

### คุณลักษณะของข้อสอบ

ความเชื่อมั่นของคะแนนแบบทดสอบที่มีข้อสอบตั้งแต่ 2 ข้อขึ้นไปต้องขึ้นอยู่กับคุณลักษณะบางประการของข้อสอบ ในหัวข้อนี้เราจะพิจารณาตัวบ่งชี้ที่มีผลต่อความเชื่อมั่นของคะแนนแบบทดสอบ ก็คือ ดัชนีความเชื่อมั่นและดัชนีอำนาจจำแนก และดัชนีทางอ้อมที่เกี่ยวข้องกับความเชื่อมั่นก็คือ ดัชนีความยากง่าย ในที่นี้ ดัชนีความยากง่ายก็คือค่าเฉลี่ยหรือค่าคาดหวังของคะแนนที่สังเกตได้ของข้อสอบ ดัชนีอำนาจจำแนกสำหรับข้อสอบเป็นนิยามของสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนที่สังเกตได้ของข้อสอบและคะแนนรวม ดัชนีความเชื่อมั่นของข้อสอบเป็นผลคูณของส่วนเบี่ยงเบนมาตรฐานของคะแนนสังเกตของข้อสอบและดัชนีอำนาจจำแนกของข้อสอบ ซึ่งเราจะพิจารณาแต่ละดัชนีในรายละเอียดต่อไปนี้

### ดัชนีความเชื่อมั่นและดัชนีอำนาจจำแนก

ดัชนีความเชื่อมั่นของข้อสอบแต่ละข้อจะเท่ากับผลคูณของคะแนนเบี่ยงเบนมาตรฐานกับสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม

$$\sigma_{Y_i} \rho_{XY_i}$$

ผลคูณของส่วนเบี่ยงเบนมาตรฐานและสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมเรียกว่า ดัชนีความเชื่อมั่น และผลรวมของดัชนีความเชื่อมั่นในแต่ละข้อจะมีค่าเท่ากับคะแนนเบี่ยงเบนมาตรฐานของคะแนนรวม

$$\sigma_X = \sum \sigma_{Y_i} \rho_{XY_i}$$

ในโครงสร้างของแบบทดสอบ สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมก็คือดัชนีอำนาจจำแนก นั่นคือข้อสอบสามารถจำแนกผู้สอบที่มีคุณลักษณะที่ต้องการวัดมากหรือน้อยออกจากกันได้

ตาราง 4.4 สถิติพื้นฐานของข้อสอบ 36 ข้อ และผู้สอบ 498 คน

ข้อสอบ	ค่าเฉลี่ย	ส่วนเบี่ยงเบน มาตรฐาน	Item-Total Correlation	Index of Reliability
1	0.81	0.40	0.25	0.10
2	0.55	0.50	0.27	0.14
3	0.84	0.36	0.17	0.06
4	0.38	0.49	0.43	0.21
5	0.59	0.49	0.45	0.22
6	0.76	0.43	0.20	0.09
7	0.82	0.38	0.29	0.11
8	0.28	0.45	0.34	0.15
9	0.77	0.42	0.26	0.11
10	0.90	0.31	0.30	0.09
11	0.30	0.46	0.32	0.15
12	0.55	0.50	0.44	0.22
13	0.93	0.26	0.24	0.06
14	0.11	0.32	0.21	0.07
15	0.67	0.47	0.34	0.16
16	0.52	0.50	0.28	0.14
17	0.72	0.45	0.26	0.12
18	0.45	0.50	0.30	0.15
19	0.76	0.43	0.37	0.16
20	0.36	0.48	0.38	0.18
21	0.67	0.47	0.23	0.11
22	0.51	0.50	0.42	0.21
23	0.45	0.50	0.42	0.21
24	0.72	0.45	0.31	0.14
25	0.27	0.44	0.31	0.14
26	0.26	0.44	0.34	0.15
27	0.37	0.48	0.45	0.22
28	0.21	0.41	0.25	0.10
29	0.53	0.50	0.26	0.13

ข้อสอบ	ค่าเฉลี่ย	ส่วนเบี่ยงเบน มาตรฐาน	Item-Total Correlation	Index of Reliability
30	0.34	0.47	0.34	0.16
31	0.19	0.69	0.27	0.11
32	0.21	0.41	0.31	0.13
33	0.50	0.50	0.42	0.21
34	0.41	0.49	0.33	0.16
35	0.24	0.43	0.29	0.12
36	0.06	0.24	0.09	0.02

หมายเหตุ สัมประสิทธิ์ความเชื่อมั่นคือผลคูณระหว่างส่วนเบี่ยงเบนมาตรฐานกับความสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวม

จากข้อมูลในตาราง 4.4 จะแสดงการใช้ดัชนีความเชื่อมั่น ผลของข้อสอบ 36 ข้อ ใช้กับผู้สอบ 498 คน สถิติพื้นฐานแสดงดังตาราง 1 สังเกตว่า ผลรวมกำลังสองของส่วนเบี่ยงเบนมาตรฐานของทั้ง 36 ข้อมีค่า 7.03 ผลรวมดัชนีความเชื่อมั่น 5.01 ซึ่งเท่ากับส่วนเบี่ยงเบนมาตรฐานของคะแนนรวม และค่าความเชื่อมั่นทั้งฉบับมีค่า 0.74

สังเกตว่ามีอยู่ 16 ข้อที่มีค่าสหสัมพันธ์ระหว่างคะแนนรายข้อกับคะแนนรวมน้อยกว่า 0.30 สมมติว่าทั้ง 16 ข้อนี้มีค่า 0.30 ทั้งหมด ในขณะที่ส่วนเบี่ยงเบนมาตรฐานมีค่าคงที่ ดัชนีความเชื่อมั่นจะเพิ่มขึ้น ซึ่งผลรวมของความเชื่อมั่นรายข้อมีค่า 5.36 (จากเดิม 5.01) มีผลทำให้ค่าความเชื่อมั่นทั้งฉบับเพิ่มสูงขึ้นเป็น 0.78 ซึ่งเพิ่มขึ้นจากเดิม 0.04 ถึงแม้ว่าจะไม่มากนักแต่ก็สามารถทำให้ความเชื่อมั่นเพิ่มขึ้นได้โดยไม่ต้องเพิ่มความยาวของแบบทดสอบ ยิ่งกว่านั้น ถ้าลองเพิ่มความยาวของแบบทดสอบเพื่อช่วยให้ความเชื่อมั่นเพิ่มขึ้น เราพบว่าถ้าเพิ่มความยาวของแบบทดสอบ 20% คือเพิ่มข้อสอบอีก 8 ข้อ แน่นอนว่า ความเชื่อมั่นก็จะเพิ่มสูงขึ้น

นอกจากนี้ยังควรที่จะพิจารณาองค์ประกอบอื่น ๆ ที่ช่วยให้ความเชื่อมั่นสูงขึ้นก็คือ

#### ความเที่ยงตรงของข้อสอบ

โดยปกติไม่ควรจะเลือกข้อสอบเพื่อจัดฉบับเป็นแบบทดสอบโดยใช้ดัชนีความเชื่อมั่นโดยไม่คำนึงถึงความเที่ยงตรงของข้อสอบ ซึ่งสิ่งสำคัญประการแรกของแบบทดสอบคือข้อสอบต้องเที่ยงตรง นั่นคือต้องทดสอบความเที่ยงตรงแบบต่าง ๆ ของข้อสอบเสียก่อน แต่เมื่อมีข้อสอบ 2 ข้อในแบบทดสอบที่มีความเที่ยงตรงและวัดความรู้และทักษะเดียวกันแล้ว ในการเลือกข้อสอบนั้นควรจะเลือกข้อสอบที่มีดัชนีความเชื่อมั่นที่สูงกว่า

### ข้อสอบถูกรวมอยู่ในคะแนนรวมหรือไม่

ดัชนีอำนาจจำแนก เป็นค่าความสัมพันธ์ระหว่างคะแนนของข้อสอบกับคะแนนรวมของคุณลักษณะที่ต้องการวัด ค่าถามที่เกิดขึ้นก็คือ ความสัมพันธ์ระหว่างคะแนนของข้อสอบกับคะแนนรวมควรจะอยู่บนพื้นฐานของคะแนนรวมทุกข้อในองค์ประกอบที่วัดหรือคะแนนรวมทุกข้อที่ยกเว้นข้อที่นำมาคำนวณหาความสัมพันธ์ เมื่อข้อสอบที่นำมาหาความสัมพันธ์ข้อนั้นถูกรวมอยู่ในคะแนนรวม สัมประสิทธิ์สหสัมพันธ์ (หรือดัชนีอำนาจจำแนก) ของข้อสอบข้อนั้นจะลำเอียงสูง (มีค่าเพิ่มขึ้น) เพราะว่าคะแนนของข้อสอบจะสัมพันธ์กันอย่างสมบูรณ์กับคะแนนรวมในองค์ประกอบที่วัด ดังนั้นในแบบทดสอบที่มี  $n$  ข้อ สัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนของข้อสอบข้อนั้นกับคะแนนรวมที่คำนวณบนพื้นฐานของข้อสอบ  $n - 1$  ข้อ หรือสหสัมพันธ์ระหว่างคะแนนของข้อสอบข้อนั้นกับคะแนนรวมที่หักข้อนั้นออก

### ดัชนีความยากง่าย

แบบทดสอบที่ประกอบด้วยข้อสอบที่มีความง่ายมากหรือยากมากอย่างใดอย่างหนึ่งมีแนวโน้มว่าค่าความเชื่อมั่นจะต่ำ ในกรณีนี้ ความแปรปรวนของคะแนนจริงจะต่ำ ความแปรปรวนของคะแนนสังเกตได้จะเกิดความคลาดเคลื่อนในการวัด เมื่อเราต้องการให้ค่าความเชื่อมั่นมีค่าสูงแล้ว แนะนำว่าข้อสอบในแบบทดสอบควรเลือกข้อที่มีค่าความยากง่ายปานกลาง แล้วความแปรปรวนของคะแนนจริงจะสูงย่อมส่งผลให้ความเชื่อมั่นของแบบทดสอบมีค่าสูงด้วย

### ความคลาดเคลื่อนมาตรฐานของการวัด

ความเชื่อมั่นคือสัดส่วนของความแปรปรวนคะแนนจริงกับคะแนนที่สังเกตได้ อย่างไรก็ตาม ผู้ใช้แบบทดสอบแน่ใจว่าความคลาดเคลื่อนของการวัดย่อมมีผลต่อการแปลความหมายคะแนนสอบของผู้เข้าสอบ แต่ว่าเป็นไปไม่ได้ที่จะทราบจำนวนของความคลาดเคลื่อนที่มีอยู่ในคะแนนที่สังเกตได้ ในทฤษฎีมาตรฐานเดิมกำหนดวิธีการสำหรับอธิบายความแปรปรวนที่คาดหวังของคะแนนที่สังเกตของแต่ละบุคคล ซึ่งจะเกี่ยวข้องกับคะแนนจริงของผู้สอบแต่ละบุคคล จำได้ว่าคะแนนจริงนิยามว่าเป็นค่าเฉลี่ยหรือค่าคาดหวังของคะแนนที่สังเกตได้ของผู้สอบจากการสอบซ้ำกันหลาย ๆ ครั้งด้วยข้อสอบเดิม ตามทฤษฎีผู้สอบแต่ละคนก็จะมีผลการแจกแจงที่เป็นไปได้ของคะแนนที่สังเกตได้ซึ่งจะรวมคะแนนจริงเอาไว้ด้วยและก็จะมีส่วนเบี่ยงเบนมาตรฐานของคะแนนเมื่อนำส่วนเบี่ยงเบนมาตรฐานของแต่ละบุคคลมาเฉลี่ยทั้งกลุ่มแล้ว ผลที่ได้ก็คือความคลาดเคลื่อนมาตรฐานของการวัด (Standard error of measurement) ใช้สัญลักษณ์ว่า  $\sigma_E$  ซึ่งการแสดงสูตรการคำนวณความคลาดเคลื่อนมาตรฐานในการวัดสามารถได้มาโดยใช้ความสัมพันธ์ของ

$$\sigma_T^2 + \sigma_E^2 = \sigma_X^2$$

หารทั้งสองข้างของสมการนี้ด้วย  $\sigma_X^2$  จะได้

$$\frac{\sigma_T^2}{\sigma_X^2} + \frac{\sigma_E^2}{\sigma_X^2} = 1$$

สังเกตว่าเทอมแรกทางซ้ายมือสามารถแสดงได้ด้วยนิยามของความเชื่อมั่น  $\rho_{XX'}$  ดังนั้น

$$\rho_{XX'} + \frac{\sigma_E^2}{\sigma_X^2} = 1$$

แก้สมการเพื่อให้ได้  $\sigma_E$  นั่นคือ

$$\frac{\sigma_E^2}{\sigma_X^2} = 1 - \rho_{XX'}$$

และ

$$\begin{aligned} \sigma_E^2 &= \sigma_X^2(1 - \rho_{XX'}) \\ \sigma_E &= \sigma_X \sqrt{1 - \rho_{XX'}} \end{aligned}$$

ดังนั้น ถ้าส่วนเบี่ยงเบนมาตรฐานของคะแนนที่สังเกตได้คือ 10 และสัมประสิทธิ์ความเชื่อมั่นหรือ  $\rho_{XX'} = 0.91$  ค่าความคลาดเคลื่อนมาตรฐานของการวัดควรจะคำนวณได้ค่า

$$\sigma_E = 10 \sqrt{1 - 0.91} = 3.0$$

สมมติว่าความคลาดเคลื่อนของการวัดเป็นไปอย่างสุ่มที่มีการแจกแจงเป็นโค้งปกติ เราควรจะคาดหวังว่าประมาณ 68% ของคะแนนที่สังเกตได้ของผู้สอบนั้นตกอยู่ในช่วง  $T \pm 1 \sigma_E$  และประมาณ 95% ของคะแนนที่สังเกตได้ของผู้สอบตกอยู่ในช่วง  $T \pm 1.96 \sigma_E$  และประมาณ 99% ของคะแนนที่สังเกตได้ของผู้สอบตกอยู่ในช่วง  $T \pm 2.58 \sigma_E$

ในการทดสอบโดยมาก ผู้สอบจะได้รับการทดสอบเพียงครั้งเดียวและมีคะแนนที่สังเกตได้เพียงค่าเดียว ดังนั้นถ้าเรามีการประมาณค่าความคลาดเคลื่อนมาตรฐานของการวัดสำหรับแบบทดสอบเอาไว้ เราไม่สามารถกำหนดช่วงของคะแนนจริงของผู้สอบได้เพราะว่าในทางปฏิบัติ นั้นเราไม่รู้คะแนนจริง ดังนั้น เราจึงใช้การประมาณค่าของความคลาดเคลื่อนมาตรฐานสร้างช่วงความเชื่อมั่นที่อยู่รอบคะแนนที่สังเกตได้จาก  $X \pm 1 \sigma_E$  เราสามารถเชื่อมั่นได้ 68% ว่าคะแนนจริงจะตกอยู่ภายในช่วงนี้ ถ้าเป็นไปได้เราอาจจะใช้คะแนนจริงแทนที่คะแนนที่สังเกตได้ แต่ก็เป็นไปได้ สมมติว่าสุชาติมีคะแนนจริงของมาตรวัดเจตคติเท่ากับ 50 และความคลาดเคลื่อนมาตรฐานของการวัดเท่ากับ 5 ตามทฤษฎีถ้าสุชาติทำข้อสอบ 100 ครั้งแล้ว จะมีประมาณ 68 ครั้งที่คะแนนที่สังเกตได้จะตกอยู่ระหว่าง 45 ถึง 55 ซึ่งคะแนน 32 จะตกอยู่นอกขอบเขต 45 ถึง 55 นี้

นั่นคือช่วงความเชื่อมั่นที่สร้างจากคะแนนสังเกตได้ของสุชาติ 100 ค่า จะมีประมาณ 68 ค่าที่ตกอยู่ในช่วงระหว่าง 45 ถึง 55 และช่วงนี้จะรวมคะแนนจริงของสุชาติเอาไว้ด้วย ในการทดสอบของสุชาติเพียงครั้งเดียว (ในความเป็นจริงย่อมสอบเพียงครั้งเดียว) โดยดึงมาอย่างสุ่มจากคะแนนที่สังเกตได้ 100 ค่า จะมีโอกาส 68% ที่คะแนนจะตกอยู่ในช่วง 45 ถึง 55 ถ้าเราไม่

โซคร้ายดึงได้คะแนน 32 ที่ตกอยู่นอกช่วง 45 ถึง 55 ซึ่งช่วง 5 คะแนนที่อยู่รอบคะแนนจริงนี้ไม่ใช่คะแนนจริง สำหรับเหตุผลนี้มีความสำคัญที่จะต้องจำว่าคะแนนจริงที่สังเกตได้เพียงค่าเดียวอาจจะไม่สามารถประมาณค่าของคะแนนจริงได้ใกล้เคียง ค่าความคลาดเคลื่อนมาตรฐานของการวัดมีประโยชน์สำหรับการจัดเตรียมการประมาณค่าคะแนนจริงที่อาจจะตกอยู่ห่างจากคะแนนที่สังเกตได้ของกลุ่มประชากร แต่ก็ไม่สามารถประกันได้ว่าคะแนนจริงของแต่ละบุคคลจะตกอยู่ในช่วงความเชื่อมั่นที่ได้ นอกจากนี้ควรจะสังเกตว่าค่าของ  $\sigma_E$  สะท้อนให้เห็นถึงค่าเฉลี่ยของความคลาดเคลื่อนมาตรฐานของผู้เข้าสอบ

### ความคลาดเคลื่อนมาตรฐานของการประมาณค่า

เป็นความคลาดเคลื่อนมาตรฐานสำหรับทำนายคะแนนของผู้สอบบนแบบทดสอบ 2 ฉบับที่คู่ขนานกันในกรณีที่รู้คะแนนของแบบทดสอบฉบับหนึ่ง มีสูตรในการคำนวณคือ

$$\sigma_{X_2 \cdot X_1} = \sigma_{X_2} \sqrt{1 - \rho_{X_1 X_2}^2}$$

เช่น มาตรฐานวัดเจตคติที่คู่ขนานกัน 2 ฉบับ มาตรฐานวัดเจตคติฉบับที่ 2 มีคะแนนเบี่ยงเบนมาตรฐาน 10 และสัมประสิทธิ์สหสัมพันธ์ระหว่างแบบทดสอบคู่ขนานทั้งสองฉบับคือ 0.75 ความคลาดเคลื่อนในการประมาณค่าจะเท่ากับ  $10\sqrt{1 - (0.75)^2} = 6.6$  ดังนั้นสุชาติที่ทำคะแนนบนมาตรฐานวัดเจตคติฉบับที่ 1 ได้ 50 คะแนน สามารถทำนายคะแนนของสุชาติบนมาตรฐานวัดเจตคติฉบับที่ 2 อยู่ระหว่าง  $50 \pm \sigma_{X_2 \cdot X_1}$  หรือประมาณ 43.4 ถึง 56.6 คะแนน ที่ระดับความเชื่อมั่น 68%



### หนังสืออ่านประกอบ

- Allen, Mary J. and Yen, Wendy M. (1979). *Introduction to Measurement Theory*. Monterey : Brooks/Cole Publishing Company.
- Crocker, Linda and Algina, James. (1986). *Introduction to Classical and Modern Test Theory*. New York : CBS College Publishing.
- Feldt, Leonard S. and Brennan, Robert L. (1989). "Reliability," in *Educational Measurement*. Linn, Robert L. (Ed.). Third Edition. U.S.A. : Macmillan Publishing Company.
- Gulliksen, Harold. (1950). *Theory of Mental Tests*. U.S.A. John Wiley & Sons, Inc.



Traub, Ross E. (1994). **Reliability for the Social Sciences : Theory and Applications**. Thousand Oaks : SAGE Publications.

Trochin, William M. K. (1999). **Research Methods Knowledge Base**. (Online)  
 Available : <http://trochim.human.cornell.edu/kb/>. Retrieved December, 2000.

**คำถามท้ายบท**

1. สัมประสิทธิ์ความเชื่อมั่นคืออะไร
2. มีอิทธิพลอะไรบ้างและมีอิทธิพลอย่างไรที่มีผลต่อการประมาณค่าความเชื่อมั่น
3. ความคลาดเคลื่อนมาตรฐานแต่ละประเภทแตกต่างกันอย่างไร นำไปใช้ต่างกันอย่างไร
4. จากข้อมูลต่อไปนี้ คำนวณค่าความเชื่อมั่นแบบต่าง ๆ

คนที่	ข้อที่				
	1	2	3	4	5
1	1	1	1	1	1
2	0	0	0	0	1
3	0	0	1	0	0
4	1	1	1	1	1
5	1	1	1	1	1
6	1	1	1	1	1
7	0	1	1	1	0
8	0	1	1	1	1
9	0	1	0	0	0
10	0	1	1	1	1
11	1	1	1	1	1
12	1	1	1	1	0
13	1	0	0	0	0
14	1	1	1	1	1
15	0	1	1	1	1