ความเชื่อมั่น 7

วัตถุประสงค์

เมื่อผู้อ่านได้ศึกษาเนื้อหาบทที่ 7 แล้ว ควรจะมีความสามารถดังนี้

1. อธิบายความหมายของความเชื่อมั่นในเชิงคณิตศาสตร์ได้
2. สรุปสาเหตุที่ทำให้เครื่องมือวัดทางการศึกษามีความเชื่อมั่นต่ำได้
3. คำนวณหาความเชื่อมั่นตามวิธีการต่าง ๆ ได้
4. พิสูจน์ที่มาของสูตรในการคำนวณหาความเชื่อมั่นได้
5. อธิบายปัญหาของวิธีการหาความเชื่อมั่นแบบต่าง ๆ ได้
6. สรุปองค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่นได้

เนื้อหา

7.1 ความสำคัญของความเชื่อมั่น

ความเชื่อมั่นนับว่าเป็นคุณลักษณะที่สำคัญของเครื่องมือในการวัดทุกชนิด กล่าวคือ ถ้าเครื่องมือในการวัดขาดความเชื่อมั่นแล้ว ผลที่ได้จากการวัดก็จะไม่มีความหมาย

แบบทดสอบที่มีความเชื่อมั่นสูง หมายถึงแบบทดสอบที่นำไปสอบกับนักเรียนกี่ครั้ง ๆ ผู้เข้าสอบก็จะได้คะแนนคงที่ แน่นอน ไม่เปลี่ยนแปลง หรือเปลี่ยนแปลงไปจากเดิมแต่เพียงเล็กน้อย

ในการวัดทางการศึกษานั้น มักจะพบปัญหาในการวัดมากมาย และปัญหาที่สำคัญอย่างหนึ่งก็คือความเชื่อมั่นในการวัดทางการศึกษาต่ำกว่าความเชื่อมั่นในการวัดทางกายภาพ

ดร.อนันต์ ศรีโสภา ได้สรุปสาเหตุที่ทำให้ค่าความเชื่อมั่นในการวัดทางการศึกษาต่ำไว้ดังนี้

1. การวัดทางกายภาพนั้นสามารถวัดได้โดยตรง ไม่เหมือนกับการวัดทางการศึกษาซึ่งไม่สามารถวัดได้โดยตรง จึงทำให้ผลที่ได้จากการวัดไม่แน่นอน

2. เครื่องมือที่ใช้วัดทางกายภาพมีความถูกต้องชัดเจนกว่าเครื่องมือวัดทางการศึกษา

3. คุณลักษณะของสิ่งที่ถูกวัดทางกายภาพมีความคงที่แน่นอน ส่วนคุณลักษณะของสิ่งที่ถูกวัดทางการศึกษาเปลี่ยนแปลงอยู่ตลอดเวลา

7.2 ทฤษฎีเกี่ยวกับคะแนนจริง

คะแนนที่ได้จากการสอบ หรือคะแนนที่สังเกตได้ (observed score) จะประกอบด้วยคะแนน 2 ส่วน คือ คะแนนจริง (true score) และส่วนที่เป็นคะแนนอันเกิดจากความคลาดเคลื่อน (error score) คะแนนจริงคือส่วนของคะแนนที่วัดได้โดยไม่ถูกอิทธิพลของความคลาดเคลื่อน คะแนนส่วนนี้ถ้าวัดในสิ่งเดียวกันจะมีค่าคงที่เสมอ ไม่ว่าจะวัดกี่ครั้ง ๆ ก็ตาม คะแนนความคลาดเคลื่อนนั้นอาจเนื่องมาจากตัวข้อสอบ การควบคุมการสอบ การให้คะแนน และการเดาเป็นต้น นั่นคือนักเรียนอาจได้คะแนนโดยที่ไม่มีความรู้ในเรื่องนั้นเลยก็ได้ หรือนักเรียนอาจไม่ได้คะแนนทั้ง ๆ ที่มีความรู้ ความคลาดเคลื่อนในการวัดมี 2 ชนิด คือ ความคลาดเคลื่อนอย่างมีระบบ (systematic error) ความคลาดเคลื่อนชนิดนี้จะไม่มีผลกระทบกระเทือนต่อการวัด ความคลาดเคลื่อนชนิดนี้จะเป็นไปในทิศทางเดียว อาจเป็นลบหรือเป็นบวกทางใดทางหนึ่งเท่านั้น เช่น

แถบวัดที่ใช้มานานอาจยืด เมื่อนำแถบวัดนี้ไปวัดสิ่งใดก็ตาม จะทำให้ความยาวของสิ่งนั้นสั้นกว่าความยาวจริง ๆ เช่น การที่แถบวัดบอกความยาวเกินไป 5 ซม. ทุก ๆ หนึ่งเมตร ก็จะไม่ทำให้ค่าความเชื่อมั่นเปลี่ยนแปลงไปแต่อย่างใด ส่วนความคลาดเคลื่อนอย่างสุ่ม (random error) นั้นเป็นไปได้หลายทิศทาง เช่นบางครั้งอาจเพิ่มขึ้น บางครั้งอาจลดลง ไม่แน่นอน ตัวอย่างเช่น คะแนนที่ได้จากการสอบซึ่งบางครั้งอาจได้คะแนนมากหรือน้อยกว่าคะแนนจริง

อย่างไรก็ตาม คะแนนความคลาดเคลื่อนนี้ยังไม่สามารถหาวิธีการตรวจสอบได้ ดังนั้นจึงอาจสรุปได้แต่เพียงว่า คะแนนที่ได้จากการสอบหรือการสังเกตจะไม่เท่ากับคะแนนจริง และสามารถเขียนสมการของคะแนนในการสอบได้ดังนี้

$$x = t + e \qquad \text{.....(7-1)}$$

เมื่อ x แทน คะแนนผลการสอบ (observed score)

t แทน คะแนนจริง (true score)

e แทน คะแนนความคลาดเคลื่อน (error score)

ความคลาดเคลื่อนอย่างสุ่มนี้ เป็นไปได้ทั้งทางบวกและทางลบ ดังนั้น ในการศึกษาเกี่ยวกับความคลาดเคลื่อน จำเป็นต้องมีข้อตกลงเบื้องต้น ดังนี้

1. ค่าเฉลี่ยของคะแนนความคลาดเคลื่อนเท่ากับ 0

$$Me = 0$$

2. ความสัมพันธ์ระหว่างคะแนนจริงกับคะแนนความคลาดเคลื่อนเท่ากับ 0

$$r_{te} = 0$$

3. ความสัมพันธ์ของคะแนนความคลาดเคลื่อนของแบบทดสอบ 2 ฉบับ ที่ขนานกันเท่ากับ 0

$$r_{e_1e_2} = 0$$

จากข้อสมมุติดังกล่าว อาจเขียนสมการเกี่ยวกับความแปรปรวนได้ดังนี้

$$S_x^2 = S_t^2 + S_e^2 + 2r_{te}S_tS_e$$

แต่เนื่องจาก $r_{te} = 0$ จะได้ว่า

$$S_x^2 = S_t^2 + S_e^2 \qquad \text{.....(7-2)}$$

7.3 ความหมายของความเชื่อมั่น

ความเชื่อมั่น หมายถึงการวัดที่ให้ผลคงที่แน่นอนสม่ำเสมอ ไม่เปลี่ยนไปเปลี่ยนมา การวัดครั้งแรกเป็นอย่างไร เมื่อวัดซ้ำอีกกี่ครั้งก็ตาม ผลการวัดยังคงเหมือนเดิม โดยปกติจะนิยามความเชื่อมั่นว่าเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนจริง และความแปรปรวนของคะแนนที่ได้จากการสอบ เขียนเป็นสมการได้ดังนี้

$$r_{tt} = \frac{S_t^2}{S_x^2} \qquad \text{.....(7-3)}$$

เนื่องจาก $S_x^2 = S_t^2 + S_e^2$

ดังนั้น $$r_{tt} = 1 - \frac{S_e^2}{S_x^2} \qquad \text{.(7-4)}$$

จากสมการ (7-4) สามารถสรุปได้ดังนี้

1. r_{tt} จะมีค่าเป็นศูนย์ เมื่อ $S_e^2 = S_x^2$

2. r_{tt} จะมีค่าเป็น 1.00 เมื่อ $S_e^2 = 0$

3. ถ้าคะแนนของแบบทดสอบมีความคงเส้นคงวาเป็นที่เชื่อถือได้แล้ว S_e^2 จะมีค่าน้อยกว่า S_x^2 เสมอ

4. ความเชื่อมั่นของแบบทดสอบ (r_{tt}) ไม่เคยมีค่าติดลบ ทั้งนี้เพราะ S_e^2 ไม่เคยมีค่ามากกว่า S_x^2

7.4 การคำนวณหาค่าความเชื่อมั่น

วิธีการหาความเชื่อมั่นของเครื่องมือในการวัด อาจทำได้หลายวิธีดังนี้

1. **วิธีสอบซ้ำ** (Test-retest method) เป็นการหาความเชื่อมั่นของเครื่องมือวัด โดยการนำเอาเครื่องมือนั้นไปสอบวัดกับนักเรียนกลุ่มใดกลุ่มหนึ่ง 2 ครั้ง ในเวลาที่ต่างกัน โดยเว้นระยะเวลาในการสอบทั้ง 2 ครั้ง ให้ห่างกันพอสมควร โดยมีจุดมุ่งหมายที่จะตรวจสอบดูว่าคะแนนหรือผลการวัดทั้ง 2 ครั้งนั้น สอดคล้องหรือสัมพันธ์กันเพียงใด แล้วนำคะแนนที่ได้จากการสอบทั้ง 2 ครั้ง มาคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ โดยใช้วิธี Pearson Product Moment ค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้นี้ก็คือค่าความเชื่อมั่นของเครื่องมือวัดฉบับนั้นนั่นเอง ค่าสัมประสิทธิ์สหสัมพันธ์นี้คำนวณได้จากสูตร

$$r_{xy} = \frac{N \Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[N \Sigma X^2 - (\Sigma X)^2][N \Sigma Y^2 - (\Sigma Y)^2]}} \quad \text{.....(7-5)}$$

ตัวอย่าง 7-1 การหาค่าความเชื่อมั่นของแบบทดสอบฉบับหนึ่ง ซึ่งใช้ทดสอบกับนักเรียนกลุ่มเดียวกัน 2 ครั้ง ปรากฏผลของคะแนนดังนี้

คนที่	สอบครั้งที่ 1 X	สอบครั้งที่ 2 Y	X^2	Y^2	XY
1	1	2	1	4	2
2	2	3	4	9	6
3	2	2	4	4	4
4	3	5	9	25	15
5	4	5	16	25	20
6	5	7	25	49	35
1	6	5	36	25	30
8	8	6	64	36	48
9	9	7	81	49	63
10	IQ	8	loo	64	80
N = 10	$\Sigma X = 50$	$\Sigma Y = 50$	$\Sigma X^2 = 340$	$\Sigma Y^2 = 290$	$\Sigma XY = 303$

แทนค่าจากสูตร r_{xy}

$$= \frac{10(303) - (50)(50)}{\sqrt{[10(340) - (50)^2][10(290) - (50)^2]}}$$

$$= \frac{530}{\sqrt{(900)(400)}}$$

$$= \frac{530}{600}$$

$$= 0.883$$

ดังนั้น แบบทดสอบฉบับนี้มีค่าความเชื่อมั่นเท่ากับ 0.88

ปัญหาที่สำคัญของการหาความเชื่อมั่นโดยวิธีนี้ คือ

1. ช่วงเวลาในการสอบทั้ง 2 ครั้ง ถ้าระยะเวลาระหว่างการสอบนาน ก็จะทำให้ค่า r ลดลง ค่า r ที่ลดลงนี้อาจเนื่องจากเกิดการเปลี่ยนแปลงในตัวผู้สอบ เช่น ผู้สอบอาจไปศึกษาค้นคว้าเพิ่มเติม ซึ่งจะทำให้คะแนนของนักเรียนแต่ละคนเปลี่ยนไป และอาจเปลี่ยนในระดับที่ต่างกัน มีผลทำให้ค่าความเชื่อมั่นของข้อสอบต่ำ และหากเว้นระยะเวลาระหว่างการสอบครั้งแรกกับครั้งหลังสั้นเกินไป จะทำให้เกิด Test wise คือนักเรียนหลายคนในชั้นจะยังจดจำคำตอบบางข้อได้ ทำให้การสอบทั้ง 2 ครั้ง ไม่เป็นอิสระต่อกัน นั่นคือครั้งแรกนักเรียนตอบว่าอย่างไร ครั้งหลังก็จะตอบอย่างนั้น ทำให้ค่าสหสัมพันธ์ที่หาได้ผิดไปจากความจริง และค่าความเชื่อมั่นก็ผิดไปจากความจริงด้วย

2. ความเปลี่ยนแปลงในตัวนักเรียนด้านต่าง ๆ เช่น อารมณ์ ความรู้สึก สภาพทางร่างกายในการสอบทั้ง 2 ครั้ง แตกต่างกัน เช่น อาจเกิดความวิตกกังวล เกิดการเจ็บป่วย ซึ่งสิ่งต่าง ๆ เหล่านี้นับว่ามีอิทธิพลต่อการสอบของนักเรียน

3. สภาพการสอบทั้ง 2 ครั้ง เช่น อากาศ แสง เสียง หรือเรื่องที่อาจเกิดขึ้นโดยนึกไม่ถึงขณะสอบ ซึ่งจะส่งผลต่อการทำข้อสอบของนักเรียน และเป็นสิ่งที่ควบคุมได้ยาก

4. การขาดหายของกลุ่มตัวอย่างในการสอบทั้ง 2 ครั้ง

จะเห็นได้ว่าการหาความเชื่อมั่นโดยวิธีนี้ จำเป็นต้องพิจารณาอย่างรอบคอบ โดยเฉพาะระยะเวลาในการสอบทั้ง 2 ครั้งนั้น ควรเว้นระยะเวลาในการสอบให้ห่างกันพอสมควร

2. วิธีใช้แบบทดสอบคู่ขนาน (Paralell test method) วิธีนี้โดยหลักการแล้ว เป็นวิธีเดียวกับการสอบซ้ำ แต่พยายามป้องกันการจำข้อสอบได้ จึงได้ใช้แบบทดสอบ 2 ฉบับ ที่คู่ขนานกันไปทดสอบแก่นักเรียนกลุ่มหนึ่ง โดยที่ข้อสอบคู่ขนานทั้ง 2 ฉบับนี้ต้องเป็นอิสระต่อกัน คือไม่ใช่ทำถูกฉบับหนึ่ง แล้วต้องไปทำถูกอีกฉบับหนึ่งด้วย ตามทฤษฎีแล้ว แบบทดสอบ 2 ฉบับ จะคู่ขนานกันก็ต่อเมื่อมีจำนวนข้อเท่ากัน ถามเนื้อหาเดียวกัน มีความยากง่ายเท่ากัน นั่นคือคะแนนเฉลี่ยและความแปรปรวนของแบบทดสอบทั้ง 2 ฉบับต้องเท่ากัน เมื่อนำแบบทดสอบคู่ขนานทั้ง 2 ฉบับ ไปสอบนักเรียนแล้ว ก็นำเอาคะแนนที่ได้จากการสอบทั้ง 2 ครั้ง มาหาค่าสัมประสิทธิ์สหสัมพันธ์แบบ Pearson Product Moment เช่นเดียวกับการหาความเชื่อมั่นโดยวิธีสอบซ้ำ

ปัญหาที่สำคัญของการหาความเชื่อมั่นโดยวิธีนี้ก็คือ เป็นการยากมากที่จะสร้างแบบทดสอบทั้ง 2 ฉบับ ให้คู่ขนานกัน เมื่อไม่สามารถสร้างแบบทดสอบทั้ง 2 ฉบับ ให้ขนานกันได้จริง ๆ ก็ย่อมทำให้เกิดความคลาดเคลื่อนในการหาความเชื่อมั่นได้

3. วิธีหาความคงที่ภายใน (Internal consistency method) เนื่องจากการหาความเชื่อมั่น 2 วิธีที่กล่าวมานั้น มีปัญหาบางประการ เช่น ต้องใช้การสอบถึง 2 ครั้ง ทำให้เกิดความไม่สะดวก ดังนั้นเพื่อขจัดปัญหานี้ จึงหลีกเลี่ยงด้วยการสอบเพียงครั้งเดียว แล้วนำผลที่ได้จากการสอบมาคำนวณหาค่าความเชื่อมั่น แต่ทั้งนี้ต้องเป็นไปตามเงื่อนไขของแต่ละวิธี วิธีการคำนวณหาค่าความคงที่ภายใน แบ่งออกเป็นดังนี้

3.1 วิธีแบ่งครึ่งข้อสอบ (Split - half method) การหาความเชื่อมั่นโดยวิธีนี้ คล้ายกับการหาความเชื่อมั่นโดยใช้แบบทดสอบคู่ขนาน คือแทนที่จะนำแบบทดสอบไปสอบกับนักเรียน 2 ครั้ง ก็เปลี่ยนเป็นนำแบบทดสอบนั้นไปสอบกับนักเรียนเพียงครั้งเดียว แล้วแบ่งข้อสอบออกเป็น 2 ชุดย่อย โดยพยายามทำให้แบบทดสอบทั้ง 2 ชุดนั้น มีลักษณะคล้ายคลึงกันมากที่สุดเท่าที่จะทำได้ ทั้งนี้อาจแบ่งข้อสอบออกเป็นข้อคู่ข้อคี่ หรือแบ่งเป็น 2 ตอน คือ ตอนแรก-ตอนหลัง เมื่อได้คะแนน 2 ชุดแล้ว ก็นำมาหาค่าสัมประสิทธิ์สหสัมพันธ์แบบ Pearson Product Moment ค่าสัมประสิทธิ์สหสัมพันธ์ที่ได้นี้จะเป็นค่าความเชื่อมั่นของแบบทดสอบเพียงครึ่งฉบับ ฉะนั้นถ้าต้องการหาค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ ก็ต้องนำไปคำนวณขยายโดยใช้สูตรของสเปียร์แมน - บราวน์ (Spearman - Brown) ดังนี้

$$r_{tt} = \frac{nr_{xy}}{1 + (n - 1)r_{xy}} \qquad \text{.....(7-6)}$$

เมื่อ r_{tt} แทน ความเชื่อมั่นของแบบทดสอบที่มีความยาว n เท่า ของแบบทดสอบที่ให้ค่า r_{xy}

ถ้าแบ่งข้อสอบฉบับหนึ่งออกเป็น 2 ชุดย่อย และต้องการหาความเชื่อมั่นของข้อสอบเต็มฉบับ ก็จะต้องขยายความเชื่อมั่นเป็น 2 เท่าของฉบับเดิม สูตรขยายความเชื่อมั่นจากครึ่งฉบับเป็นเต็มฉบับ จะเป็นดังนี้

$$r_{tt} = \frac{2r_{xy}}{1 + (2 - 1)r_{xy}} \qquad \text{.....(7-7)}$$

$$= \frac{2r_{xy}}{1 + r_{xy}}$$

ตาราง 7-2 แสดงการคำนวณหาความเชื่อมั่นโดยวิธีแบ่งครึ่ง

นักเรียน	ข้อคู่ a	ข้อคี่ b	a	b^2	ab
1	1	0	1	0	0
2	3	3	9	9	9
3	2	0	4	0	0
4	4	3	16	9	12
5	2	2	4	4	4
6	1	1	1	1	1
7	3	2	9	4	6
8	2	1	4	1	2
9	4	2	16	4	8
10	2	2	4	4	4
11	3	4	9	16	12
12	3	2	9	4	6
13	3	3	9	9	9
14	2	2	4	4	4
1s	2	1	4	1	2
16	3	1	9	1	3
17	1	2	1	4	2
18	4	4	16	16	16
19	2	3	4	9	6
20	3	2	9	4	6
Σ	so	40	142	104	112

$$r_{ab} = \frac{112 - (SO \times 40)/20}{\sqrt{(142 - 50^2/20)\ (104 - 40^2/20)}} = 0.594$$

$$r_{tt} = \frac{2 \times 0.594}{1 + 0.594} = \frac{1.188}{1.594} = 0.745$$

ปัญหาที่สำคัญของการหาความเชื่อมั่นโดยวิธีนี้คือ ข้อสอบแต่ละข้อจะต้องมีความยากง่ายใกล้เคียงกันจึงจะสามารถแบ่งเป็นข้อคู่-ข้อคี่ หรือ ตอนแรก-ตอนหลัง ได้ การแบ่งข้อสอบเป็น ตอนแรก-ตอนหลัง จะมีปัญหาเรื่องเวลา เช่น นักเรียนทำข้อสอบไม่ทัน จะทำให้ผลที่ได้จากการหาความเชื่อมั่นนั้นไม่แน่นอน การหาความเชื่อมั่นโดยวิธีแบ่งครึ่งนี้ ไม่เหมาะกับข้อสอบที่เป็น Speed test เพราะถ้าให้เวลาน้อย นักเรียนทุกคนก็จะไม่ได้ทำข้อหลัง ๆ เหมือนกันหมด ทำให้ค่าความเชื่อมั่นที่ได้ผิดไปจากความจริง ในการหาค่าความเชื่อมั่นของวิธีแบ่งครึ่งอาจใช้สูตรของ Rulon ดังนี้

$$r_{tt} = 1 - S_d^2/S_x^2 \qquad \text{..... (7-8)}$$

เมื่อ r_{tt} แทน ค่าความเชื่อมั่นทั้งฉบับ

S_d^2 แทน ความแปรปรวนของคะแนนความแตกต่างของนักเรียนแต่ละคน

S_x^2 แทน ความแปรปรวนของข้อสอบทั้งฉบับ

สูตรของ Rulon นี้ Guttman ได้แปลงให้อยู่ในรูปที่ง่ายต่อการคำนวณ ดังนี้

$$r_{tt} = 2\left[1 - (S_a^2 + S_b^2)/S_x^2\right] \qquad \text{..... (7-9)}$$

เมื่อ a, b แทนข้อสอบครึ่งแรกและครึ่งหลัง ตามลำดับ

เนื่องจาก $d = a - b$ และ $x = a + b$

ดังนั้น $S_d^2 = S_a^2 + S_b^2 - 2r_{ab}S_aS_b$

$S_x^2 = S_a^2 + S_b^2 + 2r_{ab}S_aS_b$

สมการ (7-8) สามารถเขียนใหม่ได้ดังนี้

$$r_{tt} = 1 - \frac{S_a^2 + S_b^2 - 2r_{ab}S_aS_b}{S_a^2 + S_b^2 + 2r_{ab}S_aS_b}$$

$$r_{tt} = \frac{S_a^2 + S_b^2 + 2r_{ab}S_aS_b - S_a^2 - S_b^2 + 2r_{ab}S_aS_b}{S_a^2 + S_b^2 + 2r_{ab}S_aS_b}$$

$$r_{tt} = 4r_{ab}S_aS_b / S_x^2 \qquad \textbf{(7-10)}$$

ตาราง 7-3 แสดงการคำนวณหาความเชื่อมั่น โดยใช้สูตรของ Rulon

นักเรียน	ข้อคู่ a	ข้อคี่ b	d_j	d_j^2	x_j	x_j^2
1	1	0	1	1	1	1
2	3	3	0	0	6	36
3	2	0	2	4	2	4
4	4	3	1	1	7	49
5	2	2	0	0	4	16
6	1	1	0	0	2	4
7	3	2	1	1	5	25
8	2	1	1	1	3	9
9	4	2	2	4	6	36
10	2	2	0	0	4	16
11	3	4	-1	1	7	49
12	3	2	1	1	5	25
13	3	3	0	0	6	36
14	2	2	0	0	4	16
IS	2	1	1	1	3	9
16	3	1	2	4	4	16
17	1	2	-1	1	3	9
18	4	4	0	0	8	64
19	2	3	-1	1	5	25
20	3	2	1	1	5	25
Σ	50	40	10	22	90	470

$$S_d^2 = \frac{22}{20} - \left(\frac{10}{20}\right)^2 = 1.10 \quad 0.25 = 0.85$$

$$S_x^2 = \frac{470}{20} - \left(\frac{90}{20}\right)^2 = 23.50 - 20.25 = 3.25$$

$$r_{tt} = 1 - \frac{0.85}{3.25} = 1 - 0.262 = 0.738$$

3.2 วิธีคูเดอร์ ริชาร์ดสัน (Kuder - Richardson method) เป็นการดูความคงที่ภายในของแบบทดสอบ นั่นคือจะใช้ข้อสอบฉบับเดียวทดสอบกับนักเรียนเพียงครั้งเดียว แล้วนำผลที่ได้มาคำนวณหาค่าความคงที่ภายใน

สมการความแปรปรวนของข้อสอบ ซึ่งประกอบด้วยข้อสอบ n ข้อ

$$S_x^2 = \Sigma S_i^2 + 2\,\Sigma r_{ik} S_i S_k,\ i < k$$

จากสมการนี้ เทอมสุดท้ายจะมีความแปรปรวนร่วม n(n − 1) เทอม ถ้าข้อสอบแต่ละข้อมีโอกาสที่จะตอบถูกเท่ากัน และสหสัมพันธ์ระหว่างข้อสอบแต่ละข้อเท่ากัน สามารถแทน S_iS_k ด้วย S_i^2 และผลบวกของเทอมที่เป็นความแปรปรวนร่วม สามารถเขียนแทนด้วย $n(n-1)\ \bar{r}_{ik}\ S_i^2$

$$S_x^2 = \Sigma S_i^2 + n(n-1)\,\bar{r}_{ik}\,S_i^2 \qquad \text{..... (7-11)}$$

เนื่องจาก $nS_i^2 = \Sigma S_i^2$

ดังนั้น

$$\bar{r}_{ik} = \frac{S_x^2 - \Sigma S_i^2}{(n-1)\,\Sigma S_i^2} \qquad \text{..... (7-12)}$$

สมการ 7-12 เป็นค่าความเชื่อมั่นของข้อสอบข้อเดียว ดังนั้นถ้าต้องการหาค่าความเชื่อมั่นของข้อสอบที่ประกอบด้วย n ข้อ ก็ต้องขยายโดยใช้สูตรของสเปียร์แมน-บราวน์ ดังนี้

$$r_{tt} = \frac{n\bar{r}_{ik}}{1 + (n-1)\bar{r}_{ik}}$$

$$= n\frac{S_x^2 - \Sigma S_i^2}{(n-1)\,\Sigma S_i^2}\ \frac{1}{1 + (n-1)\left[(S_x^2 - \Sigma S_i^2)/(n-1)\,\Sigma S_i^2\right]}$$

$$= \frac{n}{n-1}\ \frac{S_x^2 - \Sigma S_i^2}{\Sigma S_i^2}\ \frac{1}{1 + (S_x^2/\Sigma S_i^2) - 1}$$

$$= \frac{n}{n-1}\ \frac{S_x^2 - \Sigma S_i^2}{\Sigma S_i^2}\ \frac{\Sigma S_i^2}{S_x^2}$$

$$= \frac{n}{n-1}\ \frac{S_x^2 - \Sigma S_i^2}{S_x^2} \qquad \text{..... (7-13)}$$

สมการ 7-13 เป็นการหาความเชื่อมั่นของข้อสอบที่มี n ข้อ เรียกการหาความเชื่อมั่นตามวิธีการนี้ว่า Coefficient Alpha (α)

สมการ 7-13 เป็นสูตรทั่วไปในการคำนวณหาค่าความเชื่อมั่นของแบบทดสอบ นั่นคือ

ไม่ว่าการให้คะแนนจะเป็นแบบใดก็ตาม สามารถคำนวณหาความเชื่อมั่นของแบบทดสอบฉบับนั้นด้วย Coefficient Alpha ได้เสมอ

ถ้าเป็นข้อสอบแบบ 0, 1 คือตอบผิดให้ 0 คะแนน ตอบถูกให้ 1 คะแนน จะได้ว่า

$$S_i^2 = p_i q_i$$

ดังนั้น $$r_{tt} = \frac{n}{n-1} \frac{S_x^2 - \Sigma pq}{S_x^2} \qquad (7\text{-}14)$$

หรือ $$r_{tt} = \frac{n}{n-1} \left(1 - \frac{\Sigma pq}{S_x^2}\right) \qquad (7\text{-}15)$$

สมการ 7-14 และ 7-15 นี้ ก็คือสูตรการหาความเชื่อมั่นตามวิธีของคูเดอร์-ริชาร์ดสัน สูตรที่ 20 นั่นเอง สูตรนี้มีข้อจำกัดว่าการให้คะแนนจะต้องเป็นแบบ 0, 1

ถ้าให้ข้อสอบแต่ละข้อมีระดับความยากเท่ากัน หรือค่า p คงที่ จะได้ว่า

$$\Sigma pq = n\bar{p}\bar{q}$$

นำไปแทนในสมการ (7-15) จะได้ว่า

$$r_{tt} = \frac{n}{n-1} \left(1 - \frac{n\bar{p}\bar{q}}{S_x^2}\right) \qquad \text{..... (7-16)}$$

$$= \frac{n}{n-1} \left(1 - \frac{\frac{n\Sigma p}{n} \cdot \frac{\Sigma(1-p)}{n}}{S_x^2}\right)$$

$$= \frac{n}{n-1} \left(1 - \frac{\Sigma p \cdot \frac{\Sigma 1 - \Sigma p}{n}}{S_x^2}\right)$$

$$= \frac{n}{n-1} \left(1 - \frac{\frac{\bar{X}(n-\bar{X})}{n}}{S_x^2}\right)$$

$$= \frac{n}{n-1} \left(1 - \frac{\bar{X}(n-\bar{X})}{nS_x^2}\right) \qquad (7\text{-}17)$$

สมการ 7-16 และ 7-17 เป็นการหาความเชื่อมั่นตามวิธีการของคูเดอร์-ริชาร์ดสัน สูตรที่ 21 สูตรนี้นับว่ามีความสะดวกในการคำนวณมาก เพราะเพียงแต่หาค่าเฉลี่ยและความแปรปรวนของแบบทดสอบก็สามารถคำนวณค่าความเชื่อมั่นออกมาได้ โดยนำค่าที่คำนวณ

ได้ไปแทนในสมการ 7-17 สูตรนี้มีข้อจำกัดเพิ่มเติมจากสูตรที่ 20 ก็คือ ข้อสอบแต่ละข้อต้องมีความยากง่ายเท่ากัน (มีค่า p เท่ากันทุกข้อ) จะเห็นได้ว่าในทางปฏิบัติแล้วสูตรนี้ไม่มีโอกาสได้ใช้เลยเนื่องจากเป็นการยากมากที่จะสร้างข้อสอบทุกข้อให้มีความยากง่ายเท่ากัน

สูตรที่ 21 นี้เป็น under estimate ของสูตรที่ 20 ดังนั้นถ้าคำถามแต่ละข้อมีความยากง่ายต่างกันแล้ว ค่าความเชื่อมั่นที่คำนวณโดยใช้สูตรที่ 21 นี้ จะให้ค่าต่ำกว่าที่คำนวณโดยใช้สูตรที่ 20 เสมอ แต่ถ้าข้อสอบทุกข้อมีความยากง่ายเท่ากัน ไม่ว่าจะคำนวณโดยใช้สูตรที่ 20 หรือสูตรที่ 21 ก็จะให้ค่าความเชื่อมั่นออกมาเท่ากันเสมอ

ตารางที่ 7-4 แสดงการหาค่าความเชื่อมั่นโดยใช้วิธีของคูเดอร์-ริชาร์ดสัน

นักเรียนคนที่	ข้อที่						คะแนนรวม
	1	2	3	4	5	6	
1	1	1	1	0	1	1	5
2	0	1	1	0	0	1	3
3	1	0	1	1	0	0	3
4	1	1	0	1	1	1	5
5	1	1	1	1	0	0	4
6	0	1	0	0	1	0	2
7	1	1	0	1	1	0	4
8	1	1	0	1	0	1	4
9	1	1	1	1	1	1	b
10	1	1	1	1	1	1	6
p	.8	.9	.6	.7	.6	.6	
q	.2	.1	.4	.3	.4	.4	
pq	.16	.09	.24	.21	.24	.24	

$\Sigma pq = 1.18$

$\Sigma X = 42 \quad N = 10 \quad \overline{X} = 4.2 \quad n = 6$

$\Sigma X^2 = 192$

$$S_x^2 = \frac{192}{10} - (4.2)^2 = 19.20 - 17.64 = 1.56$$

$$KR\ 20 : r_{tt} = \frac{n}{n-1}\left[1 - \frac{\Sigma pq}{S_x^2}\right]$$

$$= \frac{6}{5}\left(1 - \frac{1.18}{1.56}\right)$$

$$= 1.2\left(\frac{0.38}{1.56}\right)$$

$$= 0.29$$

$$KR\ 21 : r_{tt} = \frac{n}{n-1}\left[1 - \frac{\bar{X}(n-\bar{X})}{nS_x^2}\right]$$

$St = 1.56 \quad \bar{X} = 4.2 \quad n = 6$

$$r_{tt} = \frac{6}{5}\left[1 - \frac{4.2\ (6-4.2)}{6\ (1.56)}\right]$$

$$= \frac{6}{5}\left[1 - \frac{(0.7)\ (1.8)}{1.56}\right]$$

$$= \frac{6}{5}\ \frac{(1.56 - 1.26)}{1.56}$$

$$= \frac{6}{5}\left(\frac{0.30}{1.56}\right)$$

$$= 0.23$$

7.5 ความคลาดเคลื่อนมาตรฐานของการวัด

ถ้าเครื่องมือในการวัดมีความเชื่อมั่นสูง (r_{tt} = 1.00) ความคลาดเคลื่อนในการวัดก็จะไม่มี แต่ในทางการศึกษาจะไม่มีเครื่องมือใดที่มีความเชื่อมั่นถึง 1.00 และเมื่อใดก็ตามถ้าความเชื่อมั่นของเครื่องมือในการวัดต่ำแล้ว ค่าความคลาดเคลื่อนก็จะสูง จากสมการ 7-3 และ 7-4 จะเห็นว่าความเชื่อมั่นจะสูงขึ้นเมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนลดลง โดยที่ความแปรปรวนของคะแนนที่ได้จากการสอบคงที่ หรือเมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนคงที่ และความแปรปรวนของคะแนนที่ได้จากการสอบเพิ่มขึ้น ความเชื่อมั่นก็จะสูงขึ้นเช่นกัน

เนื่องจาก $r_{tt} = 1 - \frac{S_e^2}{S_x^2}$

$$S_e^2 = S_x^2\ (1 - r_{tt}) \qquad \text{..... (7-17)}$$

หรือ $$S_e = S_x \sqrt{1 - r_{tt}} \qquad \text{..... (7-18)}$$

เมื่อ S_e แทนความคลาดเคลื่อนมาตรฐานในการวัด

S_x แทนความเบี่ยงเบนมาตรฐานในการวัด

r_{tt} แทนความเชื่อมั่น

ตัวอย่าง แบบทดสอบฉบับหนึ่งมีความเชื่อมั่นเท่ากับ 0.84 ความเบี่ยงเบนมาตรฐานในการวัดเท่ากับ 10 จงหาความคลาดเคลื่อนมาตรฐานในการวัด

$$S_e = 10\sqrt{1-0.84}$$
$$= 10\sqrt{0.16}$$
$$= 10\,(0.4)$$
$$= 4.0$$

สมมุติว่าคะแนนที่ได้จากการสอบเป็น 40 คะแนน โอกาสที่จะได้คะแนนจริง อยู่ระหว่าง 40 ± 4 หรือ 36-44 ประมาณ 68 เปอร์เซ็นต์ และโอกาสที่จะได้คะแนนจริง อยู่ระหว่าง 40 ± 1.96 (4.0) หรือระหว่าง 32-48 ประมาณ 95 เปอร์เซ็นต์

7.6 การหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ (Criterion Referenced Test)

การวัดผลแบบอิงเกณฑ์นั้นมีจุดประสงค์เพื่อตัดสินการเรียนรู้ของผู้เรียนว่าผ่านเกณฑ์ที่กำหนดไว้หรือไม่ ในการตัดสินนั้นมีสิ่งที่ต้องคำนึงถึงอยู่ 2 ประการ คือ

1. A false-negative decision เป็นการตัดสินว่าผู้เรียนไม่ผ่านเกณฑ์ ทั้ง ๆ ที่ผู้เรียนมีความสามารถ

2. A false-positive decision เป็นการตัดสินว่าผู้เรียนผ่านเกณฑ์ ทั้ง ๆ ที่จริง ๆ แล้วผู้เรียนมีความสามารถไม่ถึงเกณฑ์

จะเห็นได้ว่าการหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์มุ่งไปในส่วนที่เกี่ยวข้องกับความถูกต้องในการตัดสินใจ (Consistency of decision making) เกี่ยวกับสภาวะการเรียนรู้ของผู้เรียน ส่วนการวัดแบบอิงกลุ่มนั้นสนใจเกี่ยวกับความถูกต้องของคะแนน กล่าวคือ คะแนนผลการสอบควรจำแนกความแตกต่างระหว่างบุคคลในกลุ่มผู้สอบ ดังนั้นการหาความเชื่อมั่นของแบบทดสอบโดยวิธีอิงกลุ่มจึงไม่เหมาะที่จะนำมาใช้หาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ (CRT) ทั้งนี้เพราะคะแนนที่ได้จากแบบทดสอบแบบอิงเกณฑ์อาจมีความแปรปรวนเพียงเล็กน้อย ดังนั้น หากนำคะแนนจากแบบทดสอบแบบอิงเกณฑ์มาหาความเชื่อมั่นโดยวิธีอิงกลุ่มแล้ว ก็จะทำให้ความเชื่อมั่นของแบบทดสอบต่ำ

นักวัดผลได้พยายามคิดหาวิธีการหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ไว้หลายวิธีด้วยกัน แต่ละวิธียังมีข้อจำกัดและข้อโต้แย้งมาก และยังหาข้อสรุปไม่ได้ว่า วิธีใดเหมาะสมที่สุด ในที่นี้จึงขอยกตัวอย่างวิธีการหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์เพียงบางวิธีเท่านั้น

1. Livingston (1972) ได้เสนอแนะวิธีหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์โดยทำการปรับปรุงวิธีการหาความเชื่อมั่นของแบบทดสอบแบบอิงกลุ่มดังนี้

$$K^2(X, T_x) = \frac{\rho^2(X, T_x)\,\sigma^2 X + (\mu_x - C_x)^2}{\sigma_x^2 + (\mu_x - C_x)^2} \qquad \text{..... (7-19)}$$

เมื่อ $K^2(X, T_x)$ แทนความเชื่อมั่นของแบบทดสอบอิงเกณฑ์

$\rho^2(X, T_x)$ แทนความเชื่อมั่นของแบบทดสอบอิงกลุ่ม

σ_x^2 แทนความแปรปรวนของคะแนน

μ_x แทนค่าเฉลี่ยของคะแนนจากแบบทดสอบ

C_x แทนเกณฑ์มาตรฐานที่กำหนดไว้สำหรับตัดสินว่า ใครมีความรู้มากพอที่จะให้ผ่านหรือให้สอบได้

จากสูตร (7-19) สามารถสรุปได้ดังนี้

1. ถ้า $\rho^2(X, T_x)$ เพิ่ม ค่า $K^2(X, T_x)$ จะเพิ่มขึ้นด้วย
2. ถ้า $\rho^2(X, T_x) = 1.00$ แล้ว $K^2(X, T_x)$ จะมีค่าเท่ากับ 1.00 ด้วย
3. ถ้า $\mu_x = C_x$ แล้วจะได้ว่า $K^2(X, T_x) = \rho^2(X, T_x)$

จะเห็นได้ว่า ค่าความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์ที่คำนวณโดยวิธีของ Livingston จะมีค่ามากกว่าหรือเท่ากับค่าความเชื่อมั่นของข้อสอบฉบับเดียวกันที่คำนวณโดยใช้สูตรของการวัดแบบอิงกลุ่ม

2. Harris (1972) ได้เสนอวิธีการหาความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์โดยใช้วิธีการวิเคราะห์ความแปรปรวนแบบทางเดียว (One way ANOVA) ซึ่งมีสูตรในการคำนวณดังนี้

$$\mu_c^2 = \frac{SS_B}{SS_B \text{ t } ssw} \qquad \text{. . . (7-20)}$$

เมื่อ μ_c^2 แทน ความเชื่อมั่นของข้อสอบแบบอิงเกณฑ์

SS_B แทน ผลบวกกำลังสองระหว่างกลุ่ม (Sum Square Between)

SS_W แทน ผลบวกกำลังสองภายในกลุ่ม (Sum Square Within)

3. Hambleton (1974) และคณะได้เสนอแนะวิธีการหาความเชื่อมั่นโดยใช้การสอบ 2 ครั้ง แล้วดูสัดส่วนของนักเรียนที่ถูกกำหนดอย่างคงเส้นคงวาในแต่ละสภาพ (ผ่าน-ไม่ผ่าน) ของการเรียนรู้ทั้ง 2 ครั้ง โดยใช้ค่าสัมประสิทธิ์แคปป้า (Kappa-K) ของ Cohen ดังนี้

$$K = (P_o - P_e) / (1 - P_e) \qquad \text{..... (7-21)}$$

เมื่อ K แทน ความเชื่อมั่นของแบบทดสอบแบบอิงเกณฑ์

P_o แทน สัดส่วนของความสอดคล้องที่สังเกตได้จากการสอบซ้ำ

P_e แทน สัดส่วนของความสอดคล้องที่คาดหวัง

สอบครั้งที่ 1 \ สอบครั้งที่ 2	Master	Non-Master	รวม
Master	.66	.14	.80
Non-Master	.04	.16	.20
รวม	.70	.30	

$P_o = .66 + .16$

$= .82$

$P_e = (.80)(.70) + (.20)(.30)$

$= .62$

$K = (.82 - .62) / (1 - .62)$

$= .53$

4. Carver ได้เสนอแนะวิธีหาความเชื่อมั่นของแบบทดสอบ โดยใช้แบบทดสอบอิงเกณฑ์ 2 ฉบับที่คู่ขนานกัน ทดสอบกับนักเรียนกลุ่มเดียวกัน แล้วแบ่งนักเรียนดังตารางข้างล่าง

		แบบทดสอบ B ตก	ได้
แบบทดสอบ A	ได้	a	b
	ตก	c	d

Reliability = (b + c) / N

N = a + b + c + d

7.7 องค์ประกอบที่มีอิทธิพลต่อค่าความเชื่อมั่น

เป็นที่ทราบกันดีอยู่แล้วว่า ความเชื่อมั่นของข้อสอบไม่คงที่อาจเพิ่มมากขึ้นหรือลดน้อยลงก็ได้ วิธีการต่าง ๆ ในการคำนวณหาความเชื่อมั่นก็เป็นสาเหตุหนึ่งที่ทำให้ค่าความเชื่อมั่นเปลี่ยนไป นอกจากนี้แล้วยังมีองค์ประกอบอื่น ๆ ซึ่งมีอิทธิพลต่อค่าความเชื่อมั่นของแบบทดสอบ ซึ่งพอสรุปได้ดังนี้

1. ความยาวของแบบทดสอบ (Test length) ความยาวของแบบทดสอบมีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบ แบบทดสอบที่ยาวจะมีความเชื่อมั่นสูงกว่าแบบทดสอบที่สั้น ทั้งนี้พิจารณาได้จากสูตรของ Spearman-Brown ดังนี้

$$r_{tt} = \frac{nr_{xy}}{1 + (n-1)\,r_{xy}}$$

เมื่อ r_{tt} แทน ความเชื่อมั่นของแบบทดสอบใหม่ที่มีความยาวเป็น n เท่า ของข้อสอบเดิม

r_{xy} แทน ความเชื่อมั่นของข้อสอบเดิม

ตัวอย่าง ถ้าความเชื่อมั่นของแบบทดสอบฉบับหนึ่งซึ่งมีอยู่ 20 ข้อ มีค่าเท่ากับ .60 และถ้าต้องการเพิ่มข้อสอบเป็น 60 ข้อ แบบทดสอบใหม่จะมีค่าความเชื่อมั่นเป็นเท่าไร

$$r_{tt} = \frac{3\,(.60)}{1 + (3-1)\,(.60)}$$

$$= 0.818$$

จากตัวอย่างนี้ จะเห็นว่าข้อสอบยิ่งยาว ความเชื่อมั่นของแบบทดสอบจะมีค่าสูงขึ้น และในทำนองเดียวกัน ถ้าลดจำนวนข้อในแบบทดสอบก็จะทำให้ค่าความเชื่อมั่นของแบบทดสอบนั้นลดลงด้วย

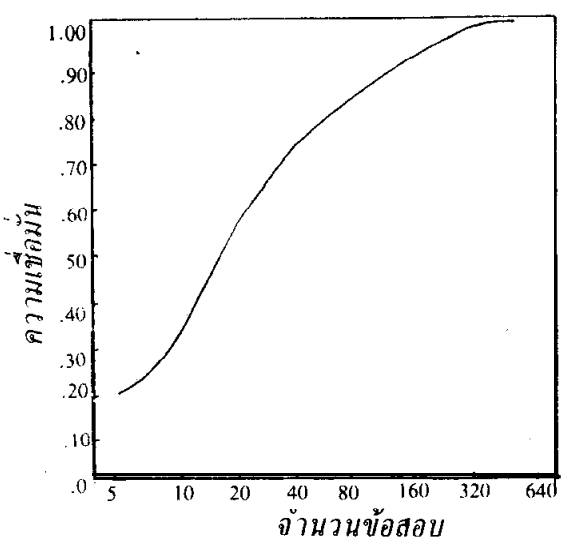
การเพิ่มค่าความเชื่อมั่นของแบบทดสอบโดยใช้สูตรนี้ มีข้อตกลง (Assumption) ว่าข้อสอบที่เพิ่มขึ้นนั้นจะต้องมีคุณสมบัติทางสถิติเหมือนกับข้อสอบเดิม คือข้อสอบที่เพิ่มเข้าไปจะต้องมีค่าความยากง่ายเหมือนเดิม หรือกล่าวง่าย ๆ ว่า ข้อสอบที่เพิ่มเข้าไปนั้นจะต้องขนานกับข้อสอบเดิม แต่ในทางปฏิบัติเมื่อเพิ่มข้อสอบแล้วนำมาหาความเชื่อมั่นใหม่ก็จะไม่ได้เท่ากับความเชื่อมั่นที่เพิ่มขึ้นตามสูตรของ Spearman-Brown ทั้งนี้เพราะข้อสอบที่เพิ่มเข้าไปอาจไม่คู่ขนานกับข้อสอบเดิมในแง่ของความยาก อำนาจจำแนก และความแปรปรวนของข้อสอบแต่ละข้อ นอกจากนี้สภาพการสอบก็เปลี่ยนไป คือเมื่อเพิ่มจำนวนข้อสอบมากขึ้น นักเรียนจะ

ต้องใช้เวลาทำนานขึ้น จะทำให้นักเรียนเกิดความเหนื่อยล้า และเบื่อหน่ายในการทำข้อสอบ ซึ่งนับว่าเป็นสาเหตุสำคัญที่ทำให้สูตรของ Spearman-Brown มีความคลาดเคลื่อน

อย่างไรก็ตาม ค่าความเชื่อมั่นจากสูตรของ Spearman-Brown ก็มิได้แปรผันโดยตรงกับความยาวที่เปลี่ยนไป ดังตัวอย่างแสดงความสัมพันธ์ระหว่างความเชื่อมั่นกับจำนวนข้อในแบบทดสอบ ดังนี้

จำนวนข้อ	ความเชื่อมั่น
5	0.20
10	0.33
20	0.50
40	0.67
80	0.80
160	0.89
320	0.94
640	0.97
∞	1.00

จากตารางจะเห็นว่า เมื่อข้อสอบมีความเชื่อมั่นสูงแล้ว การเพิ่มความยาวของข้อสอบจะทำให้ความเชื่อมั่นของข้อสอบเพิ่มขึ้นเพียงเล็กน้อยเท่านั้น เช่น เพิ่มข้อสอบ 80 ข้อ เป็น 160 ข้อ จะทำให้ความเชื่อมั่นเพิ่มขึ้นเพียง .09 เท่านั้น ดังเห็นได้จากกราฟแสดงความสัมพันธ์ระหว่างจำนวนข้อในแบบทดสอบกับความเชื่อมั่นของแบบทดสอบ ดังนี้



รูป 7-1 แสดงความสัมพันธ์ระหว่างจำนวนข้อในแบบทดสอบกับความเชื่อมั่นของแบบทดสอบ

อนึ่ง ในการเพิ่มจำนวนข้อสอบเพื่อให้แบบทดสอบมีความเชื่อมั่นสูงนั้น ย่อมจะต้องมีขอบเขตจำกัด เช่น ช่วงเวลาที่กำหนดไว้สำหรับการสอบ และยังจะต้องคำนึงถึงประสิทธิภาพด้วย กล่าวคือถ้าใช้ข้อสอบจำนวนน้อยข้อแล้วให้ผลไม่แตกต่างจากข้อสอบที่มีจำนวนมากข้อก็ให้เลือกใช้ข้อสอบที่มีจำนวนน้อยข้อ เพราะเป็นการประหยัดทางเศรษฐกิจ แต่ทั้งนี้จะต้องน้อยพอที่จะมีประสิทธิภาพในทางปฏิบัติ

2. การกระจายของคะแนน จากนิยามความเชื่อมั่นที่ว่า $r_{11} = 1 - \frac{S_e^2}{S_x^2}$ จะเห็นได้ว่า เมื่อความแปรปรวนของคะแนนความคลาดเคลื่อนคงที่ และความแปรปรวนของคะแนนรวมจากแบบทดสอบมีค่าสูงจะทำให้ค่าความเชื่อมั่นสูง แบบทดสอบที่การกระจายของคะแนนมีมาก จะส่งผลให้ความแปรปรวนของคะแนนรวมจากแบบทดสอบมีค่าสูง และส่งผลให้ความเชื่อมั่นของแบบทดสอบมีค่าสูงด้วย ในเรื่องการกระจายของคะแนนนี้เป็นผลเนื่องมาจาก

2.1 ความเป็นเอกพันธ์ของกลุ่มนักเรียน (Group Homogeneity) ถ้านักเรียนที่ไปทดสอบมีความเป็นเอกพันธ์มากจะทำให้การกระจายของคะแนนมีน้อย ซึ่งจะส่งผลให้ความเชื่อมั่นของแบบทดสอบต่ำด้วย ในทำนองกลับกัน หากกลุ่มที่ไปทดสอบมีความแตกต่างกันมากหรือมีความเป็นวิวิธพันธ์ จะทำให้คะแนนกระจายมาก ซึ่งมีผลทำให้ความแปรปรวนของคะแนนสูง อันจะทำให้ความเชื่อมั่นสูงตามไปด้วย

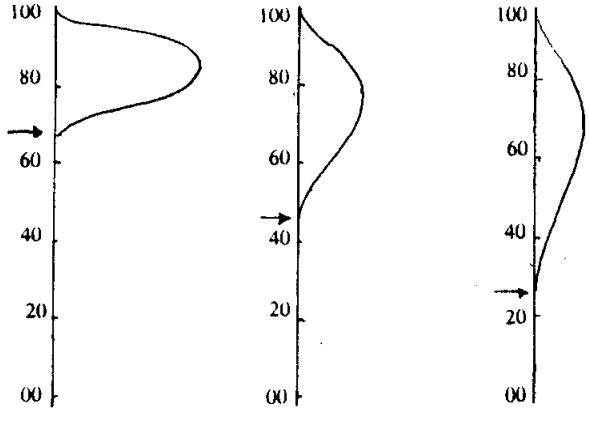
2.2 ความยากง่ายของข้อสอบ ความเชื่อมั่นของแบบทดสอบย่อมขึ้นอยู่กับความแปรปรวนของคะแนนผลการสอบ กล่าวคือถ้าความแปรปรวนระหว่างคะแนนผลการสอบมีน้อย ย่อมจะทำให้ค่าความเชื่อมั่นต่ำกว่าคะแนนผลการสอบที่มีความแปรปรวนมาก แบบทดสอบที่ยากเกินไปนักเรียนทำผิดเกือบทุกคน หรือแบบทดสอบที่ง่ายเกินไปนักเรียนทำถูกเกือบทุกคน จะจำแนกความสามารถของนักเรียนได้น้อย คะแนนจะไปรวมกันที่ใดที่หนึ่ง ทำให้ความแปรปรวนของคะแนนมีน้อย ซึ่งมีผลทำให้ความเชื่อมั่นของแบบทดสอบต่ำ ข้อสอบที่มีค่าความยากเท่ากับ .50 จะมีผลต่อความเชื่อมั่นมาก ทั้งนี้เพราะจะทำให้ความแปรปรวนในแต่ละข้อสูงสุด คือ $pq = .25$

ดังนั้น ในการสร้างข้อสอบเพื่อให้มีความเชื่อมั่นสูง ๆ จึงไม่ควรใช้ข้อสอบที่ยากหรือง่ายเกินไป ควรใช้ข้อสอบที่มีความยากง่ายพอเหมาะ คือข้อสอบที่มีค่า p ใกล้ ๆ .50

2.3 ชนิดของแบบทดสอบ แบบทดสอบต่างชนิดกันมีผลทำให้คะแนนกระจายต่างกันและมีผลทำให้ความเชื่อมั่นของแบบทดสอบแตกต่างกันด้วย ตัวอย่างเช่น ข้อสอบบรรยาย

คะแนนเต็ม 100 คะแนน เวลาให้คะแนนก็มักจะให้ระหว่าง 70 ถึง 100 ซึ่งทำให้พิสัยเท่ากับ 30 สำหรับข้อสอบแบบปรนัยนั้นพิสัยตามทฤษฎีคือผลต่างระหว่างคะแนนเต็มกับคะแนนที่ได้จากการเดาถูก คะแนนที่ได้จากการเดาถูกนั้น หมายถึงคะแนนเต็มคูณด้วยโอกาสที่จะเดาถูกแต่ละข้อ เป็นต้นว่า ข้อสอบแบบถูกผิด โอกาสที่จะเดาถูกเป็น $\frac{1}{2}$ ถ้าเป็นข้อสอบแบบเลือกตอบ 4 ตัวเลือก โอกาสที่จะเดาถูกเป็น $\frac{1}{4}$ ถ้าข้อสอบถูกผิด 100 ข้อ คะแนนเต็ม 100 คะแนน คะแนนที่ได้จากการเดาก็เป็น 50 ดังนั้นพิสัยของคะแนนก็เท่ากับ 100 − 50 = 50 ถ้าเป็นข้อสอบเลือกตอบ 4 ตัวเลือก 100 ข้อ 100 คะแนน คะแนนที่ได้จากการเดาก็จะเป็น 25 ดังนั้นพิสัยของคะแนนก็เท่ากับ 100 − 25 = 75

Robert L. Ebel ได้สร้างแบบของการกระจายของคะแนน ดังรูป



รูป 7-2 แสดงการเปรียบเทียบความเชื่อมั่นของแบบทดสอบทั้งสาม

Ebel ได้สรุปว่า ข้อสอบประเภทที่มีพิสัยของคะแนนมาก จะทำให้ค่าความเชื่อมั่นของข้อสอบสูงขึ้น

3. ความเป็นปรนัยในการให้คะแนน ข้อสอบที่มีความเป็นปรนัยในการให้คะแนนสูงจะมีความเชื่อมั่นสูง ส่วนข้อสอบที่ไม่มีความเป็นปรนัยในการให้คะแนน เช่น ข้อสอบประเภทความเรียง ซึ่งการตรวจให้คะแนนขึ้นอยู่กับอารมณ์ของผู้ตรวจ ไม่มีเกณฑ์การให้คะแนนที่แน่นอนจะมีผลทำให้ค่าความเชื่อมั่นของแบบทดสอบฉบับนั้นต่ำ

ไพศาล หวังพานิช ได้เสนอแนะวิธีการที่จะสร้างเครื่องมือในการวัดให้มีความเชื่อมั่นสูงไว้ดังนี้

1. มีความชัดเจนทั้งตัวคำถาม และคำอธิบายหรือคำชี้แจงในการตอบ

2. เวลาที่ใช้ในการตอบต้องเหมาะสมพอดี ไม่มากหรือน้อยเกินไป

3. ต้องให้คะแนนอย่างมีหลักเกณฑ์แน่นอน

4. ใช้คำถามที่ถามตามหลักเกณฑ์หรือหลักวิชา เช่น ถ้าเป็นข้อสอบก็ต้องถามโดยสามารถหาคำตอบถูกได้โดยใช้หลักวิชา ไม่ใช่ถูกหรือผิดตามความรู้สึกนึกคิดของบุคคล

5. พยายามใช้คำถามมาก ๆ ข้อ

สรุปเนื้อหาบทที่ 7

1. ความเชื่อมั่น หมายถึง ความคงที่สม่ำเสมอของคะแนนจากการวัด หรือเป็นอัตราส่วนระหว่างความแปรปรวนของคะแนนจริง กับความแปรปรวนของคะแนนที่ได้จากการสอบ
2. ความเชื่อมั่นของแบบทดสอบ จะมีค่าอยู่ระหว่าง 0 ถึง 1 และค่าความเชื่อมั่นนี้จะไม่มีค่าติดลบ ทั้งนี้เพราะเป็นไปไม่ได้ที่ S_e^2 จะมีค่ามากกว่า S_x^2
3. วิธีการหาความเชื่อมั่นของแบบทดสอบ อาจทำได้ 3 วิธี คือการใช้วิธีสอบซ้ำ การใช้วิธีแบบทดสอบคู่ขนาน และการหาความคงที่ภายใน
4. สูตร α-coefficient เป็นสูตรทั่วไปในการคำนวณหาความเชื่อมั่น นั่นคือไม่ว่าแบบทดสอบจะมีการให้คะแนนแบบใดก็สามารถคำนวณหาความเชื่อมั่นของแบบทดสอบนั้นด้วย α-coefficient ได้เสมอ แต่ถ้าการให้คะแนนเป็นแบบ 0, 1 การคำนวณด้วยสูตร KR 20 จะสะดวกกว่า และถ้าข้อสอบทุกข้อมีค่าความยากง่ายเท่ากันแล้ว การคำนวณด้วย KR 21 จะเป็นวิธีที่สะดวกที่สุด แต่ในทางปฏิบัติโอกาสที่จะใช้สูตร KR 21 นั้นแทบไม่มีเลยเนื่องจากเป็นการยากมากที่จะสร้างข้อสอบให้มีความยากง่ายเท่ากันทุกข้อ
5. ความเชื่อมั่นของแบบทดสอบจะสูงหรือต่ำขึ้นอยู่กับการกระจายของคะแนน กล่าวคือถ้าคะแนนผลการสอบมีการกระจายมากค่าความเชื่อมั่นของแบบทดสอบก็จะสูง ถ้าคะแนนผลการสอบมีการกระจายน้อยค่าความเชื่อมั่นของแบบทดสอบก็จะต่ำ ในเรื่องการกระจายของคะแนนนั้นขึ้นอยู่กับองค์ประกอบหลายอย่างเช่น กลุ่มนักเรียนที่ทดสอบ ความยากง่ายของข้อสอบ ชนิดของแบบทดสอบ และความยาวของแบบทดสอบ

คำถามท้ายบทที่ 7

1. ถ้ามีนักวัดผลคนหนึ่งกล่าวว่า "การหาความเชื่อมั่นของข้อสอบโดยใช้สูตร KR 21 จะให้ค่าความเชื่อมั่นต่ำกว่าการคำนวณหาความเชื่อมั่น โดยใช้สูตร KR 20 เสมอ" ท่านเห็นด้วยกับคำกล่าวข้างต้นนี้หรือไม่ เพราะเหตุใด
2. จงพิสูจน์ว่า ถ้าข้อสอบทุกข้อมีความยากเท่ากันแล้ว สูตร KR 20 จะกลายมาเป็น KR 21

$$\text{KR 20} : r_{tt} = \frac{n}{n-1}\left(1 - \frac{\Sigma pq}{S_x^2}\right)$$

$$\text{KR 21} : r_{tt} = \frac{n}{n-1}\left(1 - \frac{X(n-X)}{nS_x^2}\right)$$

3. ในการคำนวณค่าความเชื่อมั่นของข้อสอบที่เป็น 0, 1 จะใช้ α-coefficient ได้หรือไม่ เพราะเหตุใด