การประมาณค่าความเชื่อมั่น (Reliability) ตามทฤษฎีมาตรฐานเดิม

ฉัตรศิริ ปิยะพิมลสิทธิ์

บทนำ

มีหลากหลายวิธีในการนิยามและแปลความหมายความเชื่อมั่นของแบบทดสอบ ตัวอย่างเช่น แบบทดสอบจะเชื่อมั่นได้ถ้าคะแนนสังเกตและคะแนนจริงมีความสัมพันธ์กันสูง นั่นคือ คะแนนที่สังเกตได้และคะแนนจริงที่ได้มาจากผู้สอบทุก ๆ คนที่สอบแบบทดสอบฉบับหนึ่ง กำลังสองของสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้และคะแนนจริง (ρ^2_{XT}) จะเรียกว่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ หรือความเชื่อมั่นสามารถแสดงได้ด้วยสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนสังเกตของแบบทดสอบที่คู่ขนานแบบพาราเรล (Parellel) กันสองฉบับ ถ้าแบบทดสอบที่คู่ขนานกันสองฉบับนั้นใช้สอบกับประชากรผู้สอบและผลของคะแนนสังเกตของแบบทดสอบสองฉบับที่คู่ขนานนั้นนำมาหาความสัมพันธ์กัน ค่าสหสัมพันธ์นี้ (ใช้สัญลักษณ์ $\rho_{XX'}$ เมื่อ X และ X' คือคะแนนสังเกตของแบบทดสอบสองฉบับที่คู่ขนานกันแบบพาราเรล) ก็คือค่าสัมประสิทธิ์ความเชื่อมั่น โดยปกติแล้วเราไม่สามารถทราบคะแนนจริงได้ และเป็นไปไม่ได้ที่จะใช้แบบทดสอบที่คู่ขนานกันแบบพาราเรล ดังนั้นความเชื่อมั่นต้องประมาณค่าด้วยวิธีอื่น ๆ แต่หลังจากที่ได้ตรวจสอบวิธีการโดยทั่วไปในการประมาณค่าความเชื่อมั่นแล้ว มี 6 วิธีสำหรับนิยามหรือแปลความหมายของสัมประสิทธิ์ความเชื่อมั่น ดังนี้

การนิยามและแปลความหมายสัมประสิทธิ์ความเชื่อมั่น

มี 6 วิธีต่อไปนี้สำหรับนิยามสัมประสิทธิ์ความเชื่อมั่นไว้ คือ

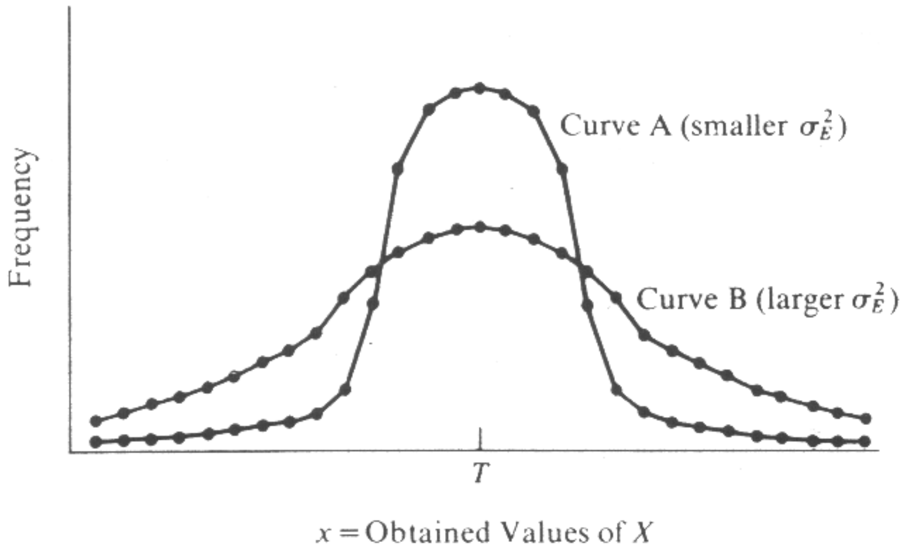
1. $r_{XX'}$ = สหสัมพันธ์ระหว่างคะแนนที่สังเกตได้จากแบบทดสอบคู่ขนาน
2. $\rho^2_{XX'}$ = สัดส่วนของความแปรปรวนใน X ที่อธิบายโดยสหสัมพันธ์เชิงเส้นกับ X′
3. $r_{XX'} = \sigma^2_T / \sigma^2_X$
4. $r_{XX'} = \rho^2_{XT}$
5. $r_{XX'} = 1 - \rho^2_{XE}$
6. $r_{XX'} = 1 - \sigma^2_E / \sigma^2_X$

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

การแปลความหมายในนิยามที่ 2 คือการแปลผลแบบมาตรฐานของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน กำลังสองของสหสัมพันธ์จะแปลความหมายได้ว่าเป็นสัดส่วนของความแปรปรวนของตัวแปรหนึ่งสามารถอธิบายได้ด้วยสหสัมพันธ์เชิงเส้นกับอีกตัวแปรหนึ่ง ดังนั้น $\rho^2_{XX'}$ จะเท่ากับสัดส่วนของความแปรปรวนของคะแนนแบบทดสอบฉบับหนึ่งที่อธิบายได้ด้วยสหสัมพันธ์เชิงเส้นกับคะแนนในแบบทดสอบอีกฉบับหนึ่งที่คู่ขนานกัน

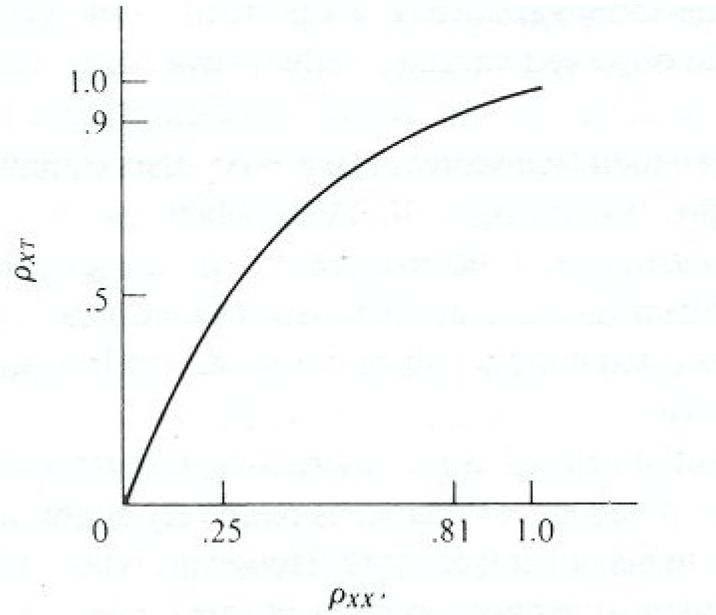
การแปลความหมายในนิยามที่ 3 $r_{XX'} = \sigma^2_T / \sigma^2_X$ คือสัมประสิทธิ์ความเชื่อมั่นเป็นอัตราส่วนของความแปรปรวนของคะแนนจริงกับความแปรปรวนของคะแนนที่สังเกตได้ สำหรับแบบทดสอบที่มีความเชื่อมั่นอย่างสมบูรณ์ $r_{XX'} = 1$ ดังนั้น $\sigma^2_T / \sigma^2_X = 1$ และความแปรปรวนของคะแนนที่สังเกตได้ทั้งหมดสะท้อนให้เห็นความแปรปรวนของคะแนนจริงที่รวมความแปรปรวนของคะแนนคลาดเคลื่อน ถ้า $r_{XX'} = 1$ ความแตกต่างระหว่างคะแนนที่สังเกตได้มาจากความแตกต่างระหว่างคะแนนจริง ถ้า $\sigma^2_X = \sigma^2_T$ แล้ว σ^2_E ต้องเป็น 0 ดังนั้น $\varepsilon(E) = 0$, ความคลาดเคลื่อนทั้งหมดจะเท่ากับ 0 เมื่อ $\sigma^2_E = 0$ ดังนั้นเมื่อ $r_{XX'} = 1$ แล้วการวัดจะต้องปราศจากความคลาดเคลื่อน เมื่อ $r_{XX'} < 1$ แล้วความคลาดเคลื่อนในการวัดยังคงมีอยู่ เมื่อ $r_{XX'} = 0$ แล้ว $\sigma^2_X = \sigma^2_E$ ซึ่งหมายความว่าคะแนนที่สังเกตได้ทั้งหมดจะสะท้อนให้เห็นเฉพาะความคลาดเคลื่อนเท่านั้น

ขณะที่ความเชื่อมั่นของแบบทดสอบเพิ่มขึ้น ความแปรปรวนของคะแนนความคลาดเคลื่อนจะน้อยลง เมื่อความแปรปรวนของความคลาดเคลื่อนมีน้อยแล้ว คะแนนที่สังเกตของผู้สอบจะมีค่าเข้าใกล้คะแนนจริงของเขามาก อย่างไรก็ตามเมื่อความแปรปรวนของความคลาดเคลื่อนมีมาก คะแนนที่สังเกตได้จะให้การประมาณค่าคะแนนจริงที่ไม่ดี ภาพประกอบ 1 จะแสดงความสัมพันธ์นี้ โค้งจะแสดงการแจกแจงเชิงทฤษฎีของคะแนนสังเกตเมื่อคะแนนจริงค่าหนึ่งคงที่ นั่นคือการแจกแจงของคะแนนสังเกตสำหรับผู้สอบคนหนึ่ง คะแนนจริงของผู้สอบจะแสดงด้วยสัญลักษณ์ T ในรูปภาพ จำไว้ว่าเมื่อคะแนนจริงถูกกำหนดให้คงที่ $\sigma^2_T = 0$ และความแปรปรวนของคะแนนสังเกตจะเท่ากับความแปรปรวนของคะแนนความคลาดเคลื่อน ภายใต้โค้ง A ซึ่งมีความแปรปรวนของความคลาดเคลื่อนน้อย คะแนนที่สังเกตได้จะมีค่าเข้าใกล้คะแนนจริง T มาก ภายใต้โค้ง B ซึ่งมีความแปรปรวนของความคลาดเคลื่อนมาก คะแนนที่สังเกตได้จะอยู่ไกลออกจากคะแนนจริง T

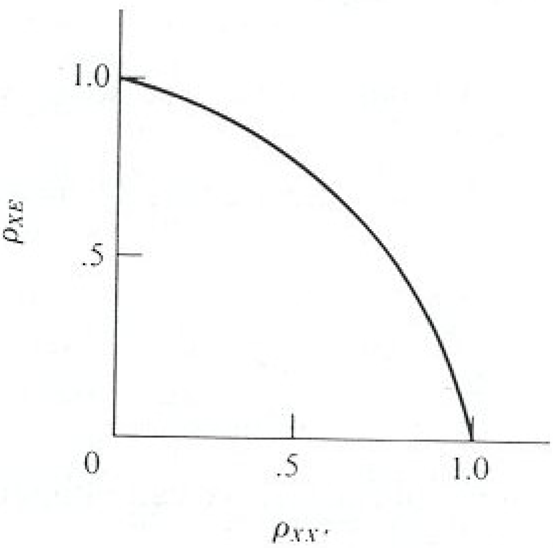


ภาพประกอบ 1 แสดงอิทธิพลของความแปรปรวนของความคลาดเคลื่อนกับคะแนนจริง

การแปลความหมายในนิยามที่ 4 $r_{XX'} = \rho^2_{XT}$ สัมประสิทธิ์ความเชื่อมั่นเท่ากับกำลังสองของสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนจริง ตัวอย่างเช่น ถ้า $r_{xx'} = .81$ ดังนั้น $r_{XT} = .9$ ถ้า $r_{XX'} = .25$ ดังนั้น $r_{XT} = .5$ ความสัมพันธ์นี้จะแสดงในภาพประกอบ 2 เมื่อ $0 < r_{XX'} > 1$ เราจะได้ว่า $r_{XT} > r_{XX'}$ คะแนนสังเกตจะสัมพันธ์กันสูงกับคะแนนจริงและสูงกว่าคะแนนที่สังเกตได้จากแบบทดสอบคู่ขนาน ในความเป็นจริง คะแนนจากแบบทดสอบจะไม่สัมพันธ์กันสูงกับแบบทดสอบอื่น ๆ มากไปกว่าคะแนนจริงของตัวมันเอง สหสัมพันธ์ที่สูงมากระหว่างคะแนนที่สังเกตได้กับแบบทดสอบอื่น ๆ คือ $\sqrt{\rho_{XX'}} = \rho_{XT}$ ถ้าแบบทดสอบ X ใช้สำหรับทำนายแบบทดสอบเกณฑ์ Y แล้ว r_{XY} จะเรียกว่า สัมประสิทธิ์ความเที่ยงตรง เพราะว่า r_{XY} ไม่สามารถจะมีค่าสูงไปกว่า r_{XT} และ r_{XY} ก็จะไม่สูงไปกว่า $\sqrt{\rho_{XX'}}$ ดังนั้นความไม่เชื่อมั่นมีผลต่อความเที่ยงตรง แม้ว่าสัมประสิทธิ์ความเที่ยงตรงไม่สามารถมีค่าสูงไปกว่ากำลังสองของรากที่สองของสัมประสิทธิ์ความเชื่อมั่น ตัวอย่างเช่น ถ้า $r_{XX'} = .49$ แล้ว $r_{XT} = .7$



ภาพประกอบ 2 ความสัมพันธ์ระหว่าง $r_{XX'}$ และ r_{XT}



ภาพประกอบ 3 ความสัมพันธ์ระหว่าง $r_{XX'}$ และ r_{XE}

การแปลความหมายในนิยามที่ 5 $r_{XX'} = 1-\rho^2_{XE}$ สัมประสิทธิ์ความเชื่อมั่นมีค่าเท่ากับ 1 ลบด้วยกำลังสองของสหสัมพันธ์ระหว่างคะแนนที่สังเกตได้กับคะแนนความคลาดเคลื่อน ในแนวคิดนี้ r_{XE} ควรจะมีค่า 0 แต่ $r_{XE} = 0$ เมื่อ $r_{XX'} = 1.0$ ความสัมพันธ์ระหว่าง r_{XE} และ $r_{XX'}$ จะแสดงในภาพประกอบ 3

การแปลความหมายในนิยามที่ 6 $r_{XX'} = 1-\sigma_E^2/\sigma_X^2$ ความเชื่อมั่นเป็นความสัมพันธ์ของความแปรปรวนของคะแนนความคลาดเคลื่อนและความแปรปรวนของคะแนนที่สังเกตได้ ดังที่อธิบายไปแล้วในตอนต้น เมื่อ $r_{XX'} = 1$, $\sigma_E^2 = 0$ และเมื่อ $r_{XX'} = 0$, $\sigma_E^2 = \sigma_X^2$ ระดับของความแปรปรวนที่เป็นวิวิธพันธ์ของคะแนนสังเกตได้มาจากกลุ่มของผู้สอบซึ่งมีผลกระทบต่อความเชื่อมั่นเป็นสำคัญ ถ้าแบบทดสอบใช้กับกลุ่มประชากรที่ทำแบบทดสอบได้พิสัยของคะแนนสังเกตน้อย (เช่น ใช้แบบทดสอบ IQ กับกลุ่มประชากรที่มีสมองช้า) σ_X^2 จะมีค่าน้อยลง ถ้าความแปรปรวนของความคลาดเคลื่อนในกลุ่มที่มีช่วงคะแนนน้อยเท่ากับกลุ่มที่เป็นวิวิธพันธ์แล้วความเชื่อมั่นของกลุ่มที่มีพิสัยน้อยจะมีค่าน้อยกว่า กล่าวอีกอย่างว่า ความเชื่อมั่นที่ประมาณได้จากกลุ่มที่เป็นวิวิธพันธ์จะมีแนวโน้มสูงกว่าการประมาณจากกลุ่มที่เป็นเอกพันธ์

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
7) ความเชื่อมั่นก็คือสัดส่วนของความแปรปรวนของคะแนนสังเกตได้ในส่วนที่เป็นความแปรปรวนของคะแนนจริง ($r_{XX'} = \sigma_T^2 / \sigma_X^2$) และ
8) ค่า $r_{XX'}$ ที่สูงจะช่วยก่อให้เกิดความเชื่อมั่นในการประมาณค่า T จาก X เพราะความแปรปรวนของความคลาดเคลื่อนมีค่าน้อยลง

การประมาณค่าความเชื่อมั่นด้วยวิธีการสอบซ้ำ

ความเชื่อมั่นของการสอบซ้ำ $r_{XX'} = \hat{\rho}_{XX'}$ จะนิยามอยู่บนพื้นฐานของการใช้กลุ่มผู้สอบกลุ่มเดียวกันที่ใช้แบบทดสอบฉบับเดียวกันซ้ำสองครั้งแล้วมานำหาความสัมพันธ์กัน ถ้าผู้สอบแต่ละคนได้คะแนนสอบเหมือนกันในการสอบซ้ำสองครั้ง และมีความแปรปรวนของคะแนนสังเกตเท่ากันแล้ว สหสัมพันธ์จะได้เท่ากับ 1.0 บ่งชี้ถึงความสัมพันธ์กันอย่างสมบูรณ์ หรือในอีกกรณีหนึ่ง ถ้าคะแนนสังเกตสำหรับผู้สอบทุกคนอ้างอิงมาจากการทดสอบฉบับที่หนึ่งที่มีความสัมพันธ์ของคะแนนสังเกตเป็นเชิงเส้นตรงอย่างสมบูรณ์กับแบบทดสอบฉบับที่สองแล้ว การประมาณค่าความเชื่อมั่นจะเท่ากับ 1.0 แต่ถ้าชุดของคะแนนจากแบบทดสอบฉบับแรกไม่มีความสัมพันธ์กับชุดของคะแนนจากแบบทดสอบฉบับที่สอง การประมาณค่าความเชื่อมั่นจะได้ 0.0 วิธีการสอบซ้ำดูเหมือนจะมีเหตุมีผลมากในการประมาณค่าความเชื่อมั่นของแบบทดสอบ แต่มันเป็นวิธีการที่ยุ่งยาก

ปัญหาที่สำคัญมากกับการประมาณค่าความเชื่อมั่นแบบสอบซ้ำคือผลของการเกิด carry-over effect ระหว่างการทดสอบ การทดสอบครั้งแรกอาจจะมีอิทธิพลต่อการสอบครั้งที่สองเกิดความคลาดเคลื่อนของคะแนนสำหรับการสอบ ผู้สอบเมื่อสอบครั้งที่สองอาจจะคำตอบได้ที่ตอบไปในครั้งแรกได้ และก็ง่ายที่ตอบคำตอบเดิมซ้ำ ถ้าผู้สอบส่วนใหญ่จำคำตอบได้เช่นนี้ $r_{XX'}$ จะประมาณค่าได้สูงกว่า $\rho_{XX'}$ หรือในบางการทดสอบ carry-over effects อาจจะเนื่องมาจากการฝึกฝน เช่นผู้สอบส่วนใหญ่มีแนวโน้มที่ได้รับการฝึกฝนจากการทดสอบซ้ำทำให้มีความคล่องแคล่วในการทำแบบทดสอบและในบางความสามารถที่แบบทดสอบวัด ถ้ามีผู้สอบบางคนที่ได้รับการฝึกฝนมากกว่าผู้สอบคนอื่น สหสัมพันธ์ของคะแนนที่สังเกตได้ในการทดสอบทั้งสองครั้งจะประมาณค่าได้ต่ำกว่า $\rho_{XX'}$ เว้นแต่ว่าระดับของการฝึกฝนนั้นจะมีความสัมพันธ์กันสูงกับคะแนนในการทดสอบในครั้งใดครั้งหนึ่งจากการทดสอบทั้งสองครั้ง

การเปลี่ยนเจตคติของผู้สอบหรือระดับของข้อสอบสามารถมีสาเหตุให้เกิด carry-over effects ได้ การไม่ให้ความร่วมมือของผู้สอบ ผู้สอบอาจจะไม่ยอมทำแบบทดสอบในครั้งที่สองและจงใจจะเดาหรือตอบผิดในการสอบครั้งที่สอง ผลนี้จะทำให้สัมประสิทธิ์ความเชื่อมั่นต่ำระหว่างการสอบทั้งสองครั้ง หรือหลังจากการสอบครั้งแรก ผู้สอบบางคนอาจจะไปค้นหาข้อมูลเพิ่มเติมเพื่อช่วยให้คะแนนของตนเองเพิ่มขึ้น ดังนั้นสัมประสิทธิ์สหสัมพันธ์ระหว่างการทดสอบ

ซ้ำสองครั้งควรจะมีแนวโน้มต่ำ carry-over effects สามารถมีผลต่อการประมาณค่าความเชื่อมั่น ช่วยให้การประมาณค่าความเชื่อมั่นแบบสอบซ้ำต่ำหรือสูงกว่าความเป็นจริง

ปัญหาประการที่สองกับการประมาณค่าความเชื่อมั่นแบบสอบซ้ำจะเกี่ยวข้องกับระยะเวลาที่เว้นช่วงห่างระหว่างการสอบทั้งสองครั้ง การเว้นช่วงห่างที่สั้นมากจะมีผลให้เกิด carry-over effects อันเนื่องมาจากความจำข้อสอบได้ การฝึกฝน หรืออารมณ์ การเว้นช่วงห่างที่ยาวนานอาจจะมีผลเนื่องมาจากการเปลี่ยนแปลงของความรู้หรืออารมณ์ ถ้าคุณลักษณะที่แบบทดสอบวัดมีการเปลี่ยนแปลงไปตามระยะเวลา เช่นความสามารถทางสมองของเด็ก การเว้นช่วงเวลาระหว่างการสอบยาวนานเกินไปจะมีแนวโน้มการประมาณค่าความเชื่อมั่นต่ำ ความแตกต่างของการเว้นช่วงระยะเวลาจะมีอิทธิพลต่อการประมาณค่าความเชื่อมั่น ในบางครั้งจะมีผลการประมาณค่าความเชื่อมั่นต่ำหรือสูงเกินไป

การประมาณค่าความเชื่อมั่นของการสอบซ้ำอยู่บนพื้นฐานของการออกแบบที่ตรงไปตรงมา สัมพันธ์กันง่ายกับผลของการใช้แบบทดสอบซ้ำ carry-over effects และการเว้นระยะเวลาสอบซ้ำมีอิทธิพลต่อขนาดของการประมาณค่าความเชื่อมั่นแบบสอบซ้ำ การประมาณค่าความเชื่อมั่นของการสอบซ้ำจะมีความเหมาะสมมากสำหรับแบบทดสอบที่วัดคุณลักษณะที่ไม่อ่อนไหวต่อ carry-over effects และมีความคงที่เมื่อเว้นช่วงห่างของการสอบซ้ำ เช่น กระบวนการสอบซ้ำที่เหมาะสมในการประมาณค่าความเชื่อมั่นกับแบบทดสอบที่ใช้โสตประสาทสัมผัสทั้งห้า (เช่นแบบทดสอบวัดการมองเห็น หรือการฟัง)

การประมาณค่าด้วยการใช้แบบทดสอบคู่ขนานและแบบทดสอบทางเลือก

ความเชื่อมั่นของการใช้แบบทดสอบคู่ขนานสามารถคำนวณได้ด้วยสูตร $r_{XX'}$ เป็นสหสัมพันธ์ระหว่างคะแนนที่สังเกตของแบบทดสอบสองฉบับที่คู่ขนานกันแบบพาราเรล (parelle) ในทางปฏิบัติเป็นไปไม่ได้ที่จะมีแบบทดสอบสอบฉบับที่คู่ขนานกัน และแบบทดสอบทางเลือกที่มักนำมาใช้แทนที่แบบทดสอบคู่ขนานเสมอ แบบทดสอบทางเลือกคือแบบทดสอบอีกฉบับหนึ่งที่มีโครงสร้างที่มีผลต่อแบบทดสอบคู่ขนาน แบบทดสอบคู่ขนานและแบบทดสอบทางเลือกนี้จะมีความเท่ากันในค่าเฉลี่ยของคะแนนสังเกต ความแปรปรวนของคะแนนสังเกต และสหสัมพันธ์กับแบบทดสอบอื่น อย่างไรก็ตาม ซึ่งยากที่จะมีคุณลักษณะแบบนี้ ไม่มีการพิสูจน์ให้เห็นถึงประโยชน์ในแบบทดสอบทางเลือก สหสัมพันธ์ระหว่างคะแนนสังเกตบนแบบทดสอบทางเลือกคือ r_{XZ} เป็นการประมาณค่าความเชื่อมั่นของแบบทดสอบฉบับหนึ่งกับแบบทดสอบทางเลือก สหสัมพันธ์นี้จะมีอิทธิพลต่อความเชื่อมั่นของแบบทดสอบ หรืออีกกรณีหนึ่งก็คือมันคู่ขนานกัน ดังนั้นแบบทดสอบอื่นจะมีแนวโน้มในการประมาณค่าความเชื่อมั่นที่แตกต่างไปจากการทดสอบแบบสอบซ้ำ หรือการประมาณค่าความเชื่อมั่นแบบคู่ขนานจะไม่มีผลต่อช่วงระยะเวลา อย่างไรก็ตามในการใช้แบบทดสอบทางเลือกหรือแบบทดสอบคู่ขนานจะไม่สามารถขจัด carry-over effect ให้หมดไปได้ ซึ่งยังคงมีผลต่อรูปแบบการตอบ อารมณ์หรือเจตคติ carry-over effects ยังคงมีผลต่อการประมาณค่า $\rho_{XX'}$ หรือ $\rho_{ZZ'}$ ให้สูงกว่าหรือต่ำกว่าความเป็นจริง ช่วงเวลายังคงมีปัญหา

การเว้นช่วงเวลาที่สั้นไปประหว่างการสอบสองฉบับจะมีผลเนื่องมาจากความจำ การฝึกฝนและอารมณ์ การเว้นช่วงเวลาที่นานเกินไปจะไม่เหมาะกับแบบทดสอบที่วัดคุณลักษณะที่เปลี่ยนแปลงไปตามเวลา

เมื่อแบบทดสอบทางเลือก X และ Z ไม่มีความคู่ขนานกันแล้ว 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 $\sigma^2_{E_X} > \sigma^2_{E_Z}$ 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 แม้ว่าสหสัมพันธ์ระหว่างคะแนนสังเกตจะเท่ากับกับความสัมพันธ์ของแบบทดสอบคู่ขนาน 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 $\sigma^2_{E_X} \neq \sigma^2_{E_Z}$) $\rho_{XZ} = \rho_{XX'}$ เมื่อ Z คือฟังก์ชันเชิงเส้นของ X’ สหสัมพันธ์ของ X และ Z จะเท่ากับสหสัมพันธ์ของ X กับ X’ ผลนี้มีโอกาสเกิดขึ้นเพราะว่า สหสัมพันธ์จะไม่มีผลกับการแปลงเชิงเส้นในแต่ละตัวแปรที่นำเข้าหาสหสัมพันธ์

สรุป สหสัมพันธ์ระหว่างคะแนนสังเกตบนแบบทดสอบทางเลือกให้การประมาณค่าความเชื่อมั่นที่ดีถ้าแบบทดสอบทางเลือกมีความคู่ขนานกันหรือคะแนนมีความสัมพันธ์กันเชิงเส้นตรง และถ้า carry-over effect และการเปลี่ยนแปลงของคะแนนที่ขึ้นอยู่กับช่วงเวลาห่างไม่มีผลต่อสหสัมพันธ์

การประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายใน : แบบแบ่งครึ่ง

ความเชื่อมั่นแบบสอดคล้องภายในจะถูกประมาณค่าด้วยการใช้แบบทดสอบเพียงฉบับเดียวสอบเพียงครั้งเดียว ดังนั้นจึงเป็นการหลีกเลี่ยงปัญหาที่เกิดขึ้นจากวิธีสอบซ้ำ วิธีนี้เป็นที่นิยมใช้กันอย่างแพร่หลาย แบบทดสอบจะถูกแบ่งครึ่งออกเป็นสองส่วน ซึ่งแต่ละส่วนจะคู่ขนานกันแบบพาราเรล (parelle) การประมาณค่าความเชื่อมั่นของแบบทดสอบจะใช้สูตรสเปียร์แมนบราวน์ (Spearman-Brown formula) แต่ถ้าทั้งสองส่วนนั้นคู่ขนานกันแบบทอ (essentially τ–equivalent) จะใช้สัมประสิทธิ์แอลฟา (α-coefficient) ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ

การใช้สูตรสเปียร์แมนบราวน์ คะแนนจากแบบทดสอบแบ่งครึ่ง (เรียกว่า Y และ Y') จะนำมาหาสหสัมพันธ์กัน ผลที่ได้คือ $\rho_{YY'}$ สหสัมพันธ์นี้ควรจะเป็นเหตุผลในการวัดความเชื่อมั่นของแบบทดสอบที่แบ่งครึ่ง ความเชื่อมั่นของแบบทดสอบทั้งฉบับ X = Y + Y' ควรจะมีค่ามากกว่าความเชื่อมั่นของแบบทดสอบเพียงแค่ครึ่งฉบับ สูตรสเปียร์แมนบราวน์จะให้ค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ คือ

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

ตาราง 1 จะแสดงค่าความเชื่อมั่น โดยปกติ $\rho_{XX'}$ จะมีค่าสูงกว่า $\rho_{YY'}$ เพราะ $\rho_{XX'}$ เป็นค่าความเชื่อมั่นของแบบทดสอบทั้งฉบับ และ $\rho_{YY'}$ เป็นค่าความเชื่อมั่นของแบบทดสอบเพียงครึ่งฉบับ

ตาราง 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) แต่ถ้าคะแนนแบ่งครึ่งนั้นไม่มีความเท่ากันในความแปรปรวนหรืออื่น ๆ ที่บ่งชี้ว่าไม่คู่ขนานกันแบบพาราเรล (parallel) แล้ว ก็จะใช้สัมประสิทธิ์แอลฟา ในการประมาณค่าความเชื่อมั่น ถ้าสองส่วนที่แบ่งครึ่งนั้นมีความคู่ขนานกันแบบทอ (essentially τ–equivalent) แต่ถ้าทั้งสองส่วนที่แบ่งครึ่งไม่มีความคู่ขนานกันแบบทอแล้ว ค่าสัมประสิทธิ์แอลฟาก็จะประมาณค่าความเชื่อมั่นได้ต่ำ (นั่นคือความเชื่อมั่นของแบบทดสอบต้องสูงกว่าหรือเท่ากับผลที่ได้จากสูตรสัมประสิทธิ์แอลฟา) ถ้าผลของสัมประสิทธิ์แอลฟามีค่าสูงแน่นอนว่าค่าความเชื่อมั่นของแบบทดสอบต้องมีค่าสูง ถ้าสัมประสิทธิ์แอลฟามีค่าต่ำ คุณไม่รู้ว่าความเชื่อมั่นของแบบทดสอบต่ำหรือแบบทดสอบที่แบ่งครึ่งไม่คู่ขนานกันแบบทอ สูตรสำหรับสัมประสิทธิ์แอลฟาแบ่งครึ่งคือ

$$\rho_{XX'} \geq \alpha = \frac{2[\sigma_X^2 - (\sigma_{Y_1}^2 + \sigma_{Y_2}^2)]}{\sigma_X^2}$$

เมื่อ $\sigma^2_{Y_1}$ และ $\sigma^2_{Y_2}$ คือความแปรปรวนของคะแนนบนแบบทดสอบแบ่งครึ่ง และ σ^2_X คือความแปรปรวนของคะแนนบนแบบทดสอบทั้งฉบับ กับ $X = Y_1 + Y_2$

การตรวจสอบสูตรนี้ให้ผลในการนำไปใช้ ความแปรปรวนของ X เท่ากับผลรวมของความแปรปรวนของ Y_1 และ Y_2 บวกกับสองเท่าของความแปรปรวนของ Y_1 และ Y_2 ดังนั้นตัวเศษของสัมประสิทธิ์แอลฟาคือฟังก์ชันของความแปรปรวนร่วมระหว่างสองส่วนที่แบ่งครึ่งของแบบทดสอบ และสัมประสิทธิ์แอลฟาสามารถแสดงได้ด้วยสูตร

$$\frac{2[2 \times \mathrm{cov\,ariance.between.halves}]}{\mathrm{total.test.score.\,var\,iance}}$$

เมื่อความแปรปรวนหรือสหสัมพันธ์ระหว่างสองส่วนที่แบ่งครึ่งเพิ่มขึ้นแล้วสัมประสิทธิ์แอลฟาจะเพิ่มขึ้น

สมการคำนวณสัมประสิทธิ์แอลฟาและสูตรสเปียร์แมนบราวน์จะมีค่ามากถ้าแบบทดสอบที่แบ่งครึ่งมีความสัมพันธ์กันสูงและจะมีค่าต่ำเมื่อแบบทดสอบไม่มีความสัมพันธ์กัน แบบทดสอบที่แบ่งครึ่งจะมีความสัมพันธ์กันสูงโดยเฉพาะเมื่อแบบทดสอบวัดคุณลักษณะเดียวกัน ดังนั้นความเชื่อมั่นแบบสเปียร์แมนบราวน์และสัมประสิทธิ์แอลฟาจะบ่งชี้ถึงแบบทดสอบที่มีความสอดคล้องภายในหรือเป็นเอกพันธ์กัน

ถ้าความแปรปรวนของคะแนนสังเกตของแบบทดสอบที่แบ่งครึ่งเท่ากัน สูตรสเปียร์แมนบราวน์และสูตรสัมประสิทธิ์แอลฟาจะมีค่าเท่ากัน ถ้าความแปรปรวนของคะแนนสังเกตของแบบทดสอบที่แบ่งครึ่งเท่ากันแต่ครึ่งนั้นไม่คู่ขนานกันแบบทอ ทั้งสูตรสเปียร์แมนบราวน์และสัมประสิทธิ์แอลฟาจะมีค่าความเชื่อมั่นต่ำกว่าความจริง ถ้าความแปรปรวนของคะแนนสังเกตของแบบทดสอบแบ่งครึ่งเท่ากันและสองส่วนนั้นคู่ขนานกันแบบทอแล้ว สูตรสเปียร์แมนบราวน์และสัมประสิทธิ์แอลฟาจะได้ค่าความเชื่อมั่นเท่ากัน

การใช้การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งสามารถแสดงได้ด้วยตัวอย่างดังนี้

สมมติว่าสหสัมพันธ์ระหว่างแบบทดสอบครึ่งฉบับเป็น 0.5 ความแปรปรวนของคะแนนคือ 7 และ 5 และความแปรปรวนของคะแนนรวมคือ 17.9 ใช้สูตรสเปียร์แมนบราวน์คำนวณค่าความเชื่อมั่นโดยรวมทั้งฉบับได้

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

ความเชื่อมั่นของแบบทดสอบทั้งฉบับประมาณค่าด้วยสัมประสิทธิ์แอลฟาได้ดังนี้

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

ตัวอย่างนี้สัมประสิทธิ์แอลฟาประมาณ่าความเชื่อมั่นได้ต่ำกว่าสูตรสเปียร์แมนบราวน์เพียงเล็กน้อย

ประโยชน์หลักของการประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายในคือใช้แบบทดสอบเพียงฉบับเดียวสอบเพียงครั้งเดียว อย่างไรก็ตามวิธีหาความสอดคล้องภายในไม่เหมาะสมเมื่อแบบทดสอบไม่สามารถแบ่งออกเป็นส่วน ๆ ที่คู่ขนานกันแบบพาราเรล

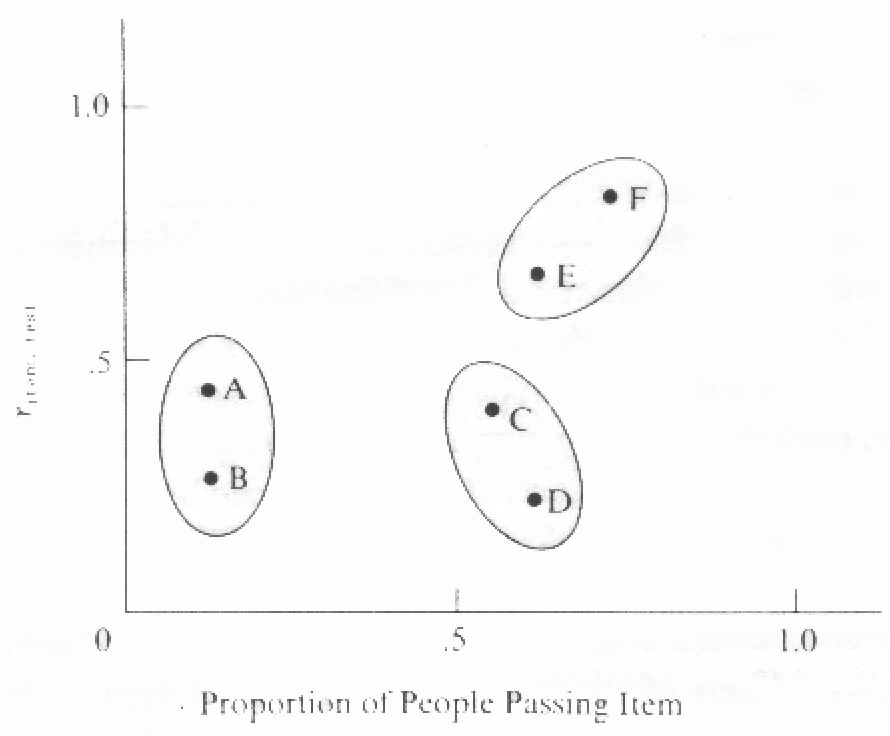
(parallel)หรือแบบทอได้ หรือเมื่อแบบทดสอบไม่มีข้อสอบที่เป็นอิสระจากกันทำให้ไม่สามารถแบ่งออกเป็นส่วน ๆ ได้ ตัวอย่างเช่น ในบางแบบทดสอบผู้สอบต้องจัดการกับวัตถุในช่วงเวลาที่กำหนดไม่สามารถจะแยกออกเป็นส่วน ๆ ได้ เพราะว่าการจัดการกับวัตถุในแต่ละชิ้นขึ้นอยู่กับเวลาและความคลาดเคลื่อนในขณะทำงานกับวัตถุชิ้นอื่น ๆ ในสถานการณ์นี้ การประมาณค่าความเชื่อมั่นแบบสอบซ้ำกับการใช้แบบทดสอบอื่นจะเหมาะสมมากกว่า

มีสามวิธีในการแบ่งครึ่งแบบทดสอบที่นิยมใช้กัน วิธีแรกจะเรียกว่า วิธีแบ่งข้อคู่ข้อคี่ (odd/even method) เป็นการจัดแบ่งข้อสอบออกเป็นสองกลุ่มโดยอาศัยตัวเลขข้อเป็นหลักในการแบ่ง ผู้สอบแต่ละคนจะมีคะแนนของแบบทดสอบในกลุ่มข้อคู่และข้อคี่ วิธีที่สองคือแบบเรียงอันดับ (order method) คือจะแบ่งข้อสอบออกเป็นครึ่งแรกกับครึ่งหลัง ผู้สอบแต่ละคนจะมีคะแนนของแบบทดสอบในครึ่งแรกและครึ่งหลังของแบบทดสอบ โดยทั่วไปการแบ่งครึ่งแรกกับครึ่งหลังจะมีความเหมาะสมน้อยกว่าแบบแบ่งข้อคู่และข้อคี่ เพราะว่าผู้สอบบางคนอาจจะได้รับการฝึกฝนมา (มีอิทธิพลกับแบบทดสอบครึ่งหลังให้คะแนนสูงกว่าปกติ) และผู้สอบบางคนทำแบบทดสอบไม่เสร็จ (มีอิทธิพลกับแบบทดสอบครึ่งหลังให้คะแนนต่ำกว่าปกติ) อย่างไรก็ตาม ปัญหาของผู้สอบบางคนที่ทำไม่เสร็จในครึ่งหลังของแบบทดสอบสามารถแก้ไขได้ด้วยการแบ่งครึ่งเวลา นั่นคือผู้สอบทำแบบทดสอบไปจนเสร็จครึ่งแรกและเมื่อเวลาหมด ผู้สอบทั้งหมดจึงค่อยลงมือทำแบบทดสอบครึ่งที่สอง จะช่วยให้ผู้สอบได้ทำแบบทดสอบสมบูรณ์ในทั้งครึ่งแรกและครึ่งหลัง การแบ่งครึ่งชนิดนี้จะมีความเท่าเทียมกับการใช้แบบทดสอบทางเลือกฉบับสั้น 2 ฉบับ

วิธีที่สามสำหรับการแบ่งครึ่งแบบทดสอบให้เท่าเทียมกันนี้เป็นวิธีที่ใหม่กว่าสองวิธีแรก วิธีนี้เรียกว่าการจับคู่แบบทดสอบย่อยอย่างสุ่ม (matched random subsets) ซึ่งมีอยู่หลายขั้นตอน ดังนี้

1. ต้องคำนวณสถิติของข้อสอบสองตัวคือ
    1.1 สัดส่วนของผู้สอบที่ทำข้อสอบนั้นถูก (ความยากง่ายของข้อสอบ)
    1.2 สหสัมพันธ์ไบซีเรียบหรือพอยท์ไบซีเรียลระหว่างคะแนนแบบทดสอบกับคะแนนรวม

ในแต่ละข้อพล็อตกราฟโดยใช้สถิติสองตัวนี้ ข้อสอบจะถูกจับคู่กันบนกราฟ โดยสองจุดใด ๆ ที่อยู่ใกล้กันให้จับทั้งคู่สุ่มไปใส่ในกลุ่มครึ่งฉบับ ตัวอย่างในภาพประกอบ 4 แสดงข้อสอบ 6 ข้อที่ถูกพล็อตลงบนกราฟ และจับกลุ่มเป็นคู่ ถ้าข้อ A ถูกเลือกเข้ากลุ่มครึ่งแรกแล้ว ข้อสอบ B ก็จะปรากฏอยู่ในอีกครึ่งหนึ่ง ความเป็นไปได้ของแบบทดสอบที่จะถูกสุ่มเข้ากลุ่มเป็นดังนี้ ACE และ BDF, ADE และ BCF, ACF และ BDE และอื่น ๆ วิธีนี้จะช่วยให้แน่ใจว่าสองส่วนนั้นมีความยากง่ายเหมือนกันและการวัดนั้นก็ประมาณค่าในสิ่งเดียวกัน (ดังนั้นคะแนนจริงจึงเท่ากัน)



ภาพประกอบ 4 การเลือกแบบทดสอบย่อยโดยใช้วิธีจับคู่สุ่มด้วยกราฟ

แบบทดสอบที่พิจารณาถึงมิติของความเร็ว (speed test) ตั้งแต่เริ่มทำจนเสร็จและแบบทดสอบที่ใช้เวลาในการคิดนาน (power test) แบบทดสอบที่ใช้ความเร็ว (speed test) สอดคล้องกับข้อสอบที่ผู้สอบทุก ๆ คนสามารถตอบได้ถูกหมดในเวลาที่พอเพียง แต่แบบทดสอบที่ให้เวลาน้อยเกินไปผู้สอบจะต้องพยายามตอบข้อสอบให้ได้โดยเร็ว ตัวอย่างเช่น แบบทดสอบที่ให้คำมาเป็นคู่ จำนวน 100 ข้อแล้วบอกถึงความแตกต่างควรจะทำให้เสร็จภายในเวลา 60 นาที หรืออีกแบบหนึ่งเป็นแบบทดสอบที่ต้องใช้ความสามารถมาก (power test) ซึ่งข้อสอบจะมีความยากให้เวลาไม่จำกัดในการสอบ การตอบได้หรือไม่ได้จึงขึ้นอยู่กับความสามารถของผู้สอบสามารถตอบคำถามได้ถูกต้องเฉพาะข้อที่แน่ใจ การจำกัดเวลาในการสอบโดยทั่วไปต้องให้แน่ใจว่าผู้สอบแต่ละคนจะสามารถทำข้อสอบแต่ละข้อได้เสร็จ การทดสอบความสามารถหรือผลสัมฤทธิ์โดยมากมักจะใช้ทั้ง speed test และ power test

การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งไม่ควรใช้กับ speed test เพราะว่าผู้สอบโดยมากจะต้องพยายามตอบให้ถูกต้องภายในเวลาที่จำกัด ถ้าข้อสอบมี 30 ข้อการแบ่งข้อคู่ข้อคี่โดยปกติก็คือ 15 ข้อ และทั้งสองส่วนนี้ควรจะคู่ขนานกัน จะทำให้การประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งมีค่าเข้าใกล้ 1 และถ้าการประมาณค่าความเชื่อมั่นแบบแบ่งครึ่งอยู่บนพื้นฐานของคะแนนที่สัมพันธ์กันระหว่างครึ่งแรกและครึ่งสองของแบบทดสอบ speed test การประมาณค่าความเชื่อมั่นจะเข้าใกล้ 0 ผู้สอบส่วนใหญ่ควรจะทำได้คะแนนดีมากในครึ่งแรก และได้คะแนนไม่ดีในครึ่งหลัง ในกรณีนี้ สหสัมพันธ์ระหว่างคะแนนบนแบบทดสอบที่แบ่งครึ่งควรจะเป็นผลสะท้อนให้เห็นความสมพันธ์ระหว่างความคลาดเคลื่อนของแบบทดสอบครึ่งแรกและความเร็วใน

การทำแบบทดสอบครึ่งหลัง การประมาณค่าความเชื่อมั่นด้วยวิธีวิธีการจับคู่ข้อสอบก็ไม่เหมาะสมกับ speed test เพราะว่าความยากของข้อสอบและความสัมพันธ์ของข้อสอบกับคะแนนรวมควรจะทำหน้าที่ในการบ่งบอกตำแหน่งของข้อสอบในแบบทดสอบมากกว่าจะบอกคุณลักษณะของข้อสอบ

วิธีการแบ่งครึ่งที่อยู่บนพื้นฐานของการแบ่งครึ่งแบบทดสอบ ลองจินตนาการต่อไปว่าถ้าแบบทดสอบถูกแบ่งออกเป็นสามส่วน สี่ส่วน ห้าส่วน และอื่น ๆ แล้ว การตรวจสอบความสัมพันธ์ระหว่างองค์ประกอบของแบบทดสอบ ถ้ากระบวนการนี้จำกัดว่า ในแต่ละข้อเป็นตัวบ่งชี้ถึงองค์ประกอบของแบบทดสอบ และความสัมพันธ์ระหว่างข้อสอบควรจะใช้เป็นตัวบ่งชี้ความสอดคล้องภายในของแบบทดสอบ ถ้าทุกองค์ประกอบของแบบทดสอบมีความสอดคล้องภายในสูง แบบทดสอบควรจะประมาณค่าความเชื่อมั่นได้สูง บ่งชี้ถึงความสอดคล้องภายในระหว่างองค์ประกอบของแบบทดสอบ

นอกจากนี้ยังมีสูตรเฉพาะสำหรับประมาณค่าความเชื่อมั่นกรณีแบบทดสอบถูกแบ่งออกเป็น 2 ส่วน และทั้งสองส่วนคู่ขนานกันแบบทอ สูตรที่ใช้ในการประมาณค่าความเชื่อมั่นได้แก่

$$F\ r_{XX'} = \frac{4\sigma_{X_1X_2}}{\sigma_X^2}$$

สูตรนี้ (มาจาก Flanagan, Rulon. 1939) อาจจะแสดงในรูปการเท่ากันทางพีชคณิตได้หลายแบบอย่างเช่น (Guttman. 1945; Rulon. 1939)

$$GR\ r_{XX'} = 2\left(1-\frac{[\sigma_{X_1}^2+\sigma_{X_2}^2]}{\sigma_X^2}\right) = 1-\frac{\sigma_{(X_1-X_2)}^2}{\sigma_X^2}$$

สำหรับในบางกรณี อาจจะเป็นไปไม่ได้ที่จะแบ่งเครื่องมือออกเป็นสองส่วนที่เป็นลักษณะ essentially tau-equivalent อย่างเช่นแบบทดสอบการอ่านที่แบ่งออกเป็นข้อคู่ข้อคี่ ซึ่งจะไม่สามารถทำได้ถ้าข้อสอบมีการให้สถานการณ์มาให้อ่านและตอบคำถามเป็นจำนวนหลาย ๆ ข้อ หรือการแบ่งความยาวของทั้งสองส่วนอาจจะไม่เท่ากัน ในกรณีนี้เราจะใช้ส่วนในลักษณะของ Congeneric ความแปรปรวนของคะแนนสังเกตจะมีผลเกิดความไม่เท่าเทียมกันในค่า σ_T^2 และ σ_E^2 สูตรสามารถหาได้โดยใช้ Horst (1951), Angoff (1953), Raju (1970), Kristof (1971), และ Feldt (1975) สำหรับในสถานการณ์นี้

สูตรของ Raju จะต้องรู้ความยาวของแต่ละส่วน นั่นก็คือจะต้องรู้จำนวนข้อของแต่ละส่วน ถ้า k_1 และ k_2 แทนความยาวของของแบบทดสอบ สูตรที่เหมาะสมคือ $\lambda_1 = k_1/(k_1 + k_2)$ และ $\lambda_2 = k_2/(k_1 + k_2)$ แล้วคะแนนสังเกตสำหรับส่วนที่ทดสอบอาจจะแสดงได้ว่า

$$X_1 = \lambda_1 T + C_1 + E_1 \text{ และ } \quad X_2 = \lambda_2 T + C_2 + E_2$$

ในที่นี้ T คือคะแนนจริงของคะแนนรวมทั้งแบบทดสอบ C_1 และ C_2 คือค่าคงที่ซึ่งเป็นความแตกต่างของค่าเฉลี่ยในส่วนที่ทดสอบ และ E_1 และ E_2 คือความคลาดเคลื่อนในการวัด ความแปรปรวนร่วมระหว่างส่วนที่เท่ากับคือ $\sigma_{X_1X_2} = \lambda_1\lambda_2\sigma_T^2$ และ

$$\sigma_T^2 = \frac{\sigma_{X_1X_2}}{\lambda_1\lambda_2}$$

ความเชื่อมั่นของแบบทดสอบรวมทั้งฉบับเท่ากับ

$$Rr_{XX'} = \frac{\sigma_{X_1X_2}}{\lambda_1\lambda_2\sigma_X^2} \quad (17)$$

ถ้าในแต่ละส่วนเป็น Congeneric และความยาวไม่เท่ากันและไม่รู้ความยาว ในกรณีนี้เพราะว่าเราไม่รู้ 4 ส่วนคือ σ_T^2, $\sigma_{E_1}^2$, $\sigma_{E_2}^2$ และ λ_1 (สังเกตว่า $\lambda_2 = 1 - \lambda_1$) แต่มีเฉพาะการประมาณค่าพารามิเตอร์สามตัวเท่านั้นที่เกี่ยวข้องคือ $\sigma_{X_1}^2$, $\sigma_{X_2}^2$ และ $\sigma_{X_1X_2}$ อย่างไรก็ตาม ปัญหากลายมาเป็นความสามารถในการจัดการถ้าสมมติประการหนึ่งว่าความแปรปรวนของความคลาดเคลื่อนมีความแตกต่างกันตามข้อตกลงเบื้องต้นแล้วละก็

$$\frac{\sigma_{E_1}^2}{\sigma_{E_2}^2} = \frac{\lambda_1}{\lambda_2} = \frac{\lambda_1}{1-\lambda_1}$$

$$\sigma_{E_2}^2 = \frac{(1-\lambda_1)}{\lambda_1}\sigma_{E_2}^2$$

และจำนวนของค่าที่ไม่รู้ลดเหลือ 3 ตัว ภายใต้กระบวนการของ Angoff และ Feldt สุดท้ายจะได้สูตรว่า

$$AFr_{XX'} = \frac{4\sigma_{X_1X_2}}{\sigma_X^2 - \left(\dfrac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X}\right)^2} \quad (18)$$

สัมประสิทธิ์นี้เท่ากับ $Rr_{XX'}$ กับ $\lambda_1 = (\sigma_{X_1}^2 + \sigma_{X_1X_2})/\sigma_X^2$ และ $\lambda_2 = 1 - \lambda_1$

กับข้อมูลในตัวอย่างก่อนหน้าที่ สูตรประมาณค่าข้างต้นนี้จะได้ $r_{XX'}$ = .7505 แน่นอนว่าค่าที่ได้ไม่ได้มีความหมายแตกต่างไปจาก .750 หรือ .748 ความแตกต่างที่มีมากระหว่างความแปรปรวนของแต่ละส่วนของแบบทดสอบควรจะต้องทำให้มีความแตกต่างกันมากระหว่างสัมประสิทธิ์ทั้ง 3 ตัว กระบวนการ Kristof (1971) สำหรับปัญหานี้เก๋มาก แต่การคำนวณจะลำบากมาก มันปรากฏผลว่ามีความสอดคล้องกันสูงกับสูตรข้างต้น

ความเชื่อมั่นแบบความสอดคล้องภายใน : กรณีทั่วไป

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

ในหัวข้อนี้จะสมมติว่าแบบทดสอบถูกแบ่งออกเป็น N ส่วน ความแปรปรวนของคะแนนในแต่ละส่วนและความแปรปรวนของคะแนนรวมของแบบทดสอบจะใช้ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ ถ้าในแต่ละส่วน (ข้อสอบ หรือชุดของข้อสอบ) มีความคู่ขนานกันแบบทอ (τ-equivalent) สูตรที่นำเสนอในหัวข้อนี้จะให้ค่าความเชื่อมั่นของแบบทดสอบ ถ้าแต่ละส่วน

ไม่คู่ขนานกันแบบทอ (τ-equivalent) สูตรในหัวข้อนี้จะประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง นอกจากนี้ สูตรจะประมาณค่าความเชื่อมั่นได้ดีเมื่อแบบทดสอบวัดคุณลักษณะเดียว นั่นคือเมื่อแบบทดสอบมีเนื้อหาเป็นเอกพันธ์ (homogeneos) แต่แบบทดสอบวัดเชาวน์ปัญญาซึ่งวัดความสามารถทางภาษา มิติสัมพันธ์ และอื่น ๆ ควรจะเป็นวิวิธพันธ์ (heterogeneos) การวัดความเชื่อมั่นแบบความสอดคล้องภายในไม่เหมาะที่จะใช้กับแบบทดสอบที่เป็นวิวิธพันธ์

สูตรสำหรับความเชื่อมั่นแบบความสอดคล้องภายในกรณีทั่วไปคือสัมประสิทธิ์แอลฟา (α-coefficient)

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

เมื่อ X คือ คะแนนรวมของแบบทดสอบที่รวมกัน N ส่วน

$X = \sum_{i=1}^{N} Y_i$

σ_X^2 คือ ความแปรปรวนของ X

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

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

ถ้าในแต่ละส่วนเป็นข้อสอบแบบ 0, 1 (dichotomous) สมการข้างต้นสามารถเขียนในรูปแบบเฉพาะได้ว่า

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

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 α-20

อีกสูตรหนึ่งของ Kuder-Richardson ก็คือ

$$\rho_{xx'} \geq KR21 = \left[\frac{N}{N-1}\right]\left[\frac{\sigma_X^2 - N\overline{p}(1-\overline{p})}{\sigma_X^2}\right]$$

เมื่อ $\overline{p}$ คือค่าเฉลี่ยของความยากข้อสอบ เพราะว่า $\overline{p}$ สามารถคำนวณได้โดยใช้ $\overline{p}$ = ε(X)/N สมการ KR21 สามารถคำนวณโดยใช้ค่าเฉลี่ยและความแปรปรวนของข้อสอบ N ข้อในแบบทดสอบ ซึ่งเป็นกรณีเฉพาะของสูตร KR20 และโดยปกติจะเกี่ยวข้องกันโดยที่

$$KR20 \geq KR21$$

ทั้งสองสูตรนี้จะมีค่าเท่ากันเมื่อความยากง่ายของข้อสอบเท่ากันทุกข้อ ถ้าข้อสอบมีค่าความยากง่ายไม่เท่ากันแล้ว KR21 จะประมาณค่าได้ต่ำกว่า KR20 และเป็นการประมาณค่าความเชื่อมั่นที่ต่ำกว่าความเป็นจริง

ผลต่างของสูตร KR20 และ KR21 เสนอโดย Tucker (1949) ดังนี้

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

ผลจากสูตร α และ KR20 จะประมาณค่าได้ต่ำกว่าหรือเท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบ จะประมาณค่าได้เท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบเมื่อในแต่ละองค์ประกอบ (Y_i) มีความคู่ขนานกันแบบทอ (τ-equivalent) (นั่นคือจำเป็นที่คะแนนจริงจะต้องเท่ากัน) ส่วน KR21 จะเท่ากับความเชื่อมั่นแท้จริงของแบบทดสอบถ้าข้อสอบมีความยากง่ายเท่ากันและคู่ขนานกันแบบทอ (τ-equivalent) และทั้งสามสูตรที่นำเสนอข้างต้นนี้จะให้ค่าความเชื่อมั่นสูงถ้าคะแนนในแต่ละส่วนกับคะแนนรวมมีความสัมพันธ์กันสูง และจะให้ค่าความเชื่อมั่นต่ำถ้าคะแนนในแต่ละส่วนกับคะแนนรวมมีความสัมพันธ์กันต่ำ ในแต่ละส่วนจะมีความสัมพันธ์กันสูงถ้าแบบทดสอบนั้นวัดคุณลักษณะเดียวกัน ดังนั้นสูตรที่นำเสนอในหัวข้อนี้จะเป็นตัวบ่งชี้ถึงความสอดคล้องภายในของแบบทดสอบหรือความเป็นเอกพันธ์ของแบบทดสอบ

อีกวิธีการหนึ่งในกรณีที่แบ่งออกเป็นหลายส่วน นำเสนอโดย Guttman (1945) ดังนี้

$$Gr_{XX'} = \frac{\Sigma\Sigma\sigma_{X\,X_g} + \sqrt{\frac{n}{n-1}\Sigma\Sigma\sigma_{X_f X_g}^2}}{\sigma_X^2} \qquad (22)$$

ถ้าในแต่ละส่วนคู่ขนานแบบ essential tau-euivalent แล้ว $Gr_{XX'}$ จะเท่ากับ $\alpha r_{XX'}$

ฮอยท์ (Hoyt. 1941) ได้พัฒนากระบวนการประมาณค่าความเชื่อมั่นที่ให้ผลทำนองเดียวกับการใช้สัมประสิทธิ์แอลฟา วิธีการของฮอยท์จะอยู่บนพื้นฐานของการวิเคราะห์ความแปรปรวนโดยจัดให้ผู้สอบและข้อสอบเป็นแหล่งความแปรปรวน สัญลักษณ์ที่ใช้จะเหมือนกับการวิเคราะห์ความแปรปรวน ฮอยท์ได้นิยามความเชื่อมั่นตามวิธีของเขาด้วยสูตร

$$\rho_{XX'} = \frac{MS_{persons} - MS_{residual}}{MS_{persons}}$$

เมื่อ $MS_{persons}$ คือค่าเฉลี่ยกำลังสองของผู้สอบที่ได้จากการวิเคราะห์ความแปรปรวน และ $MS_{residual}$ คือค่าเฉลี่ยกำลังสองของความคลาเคลื่อนจากการวิเคราะห์ความแปรปรวน ตาราง 3 จะสรุปผลการวิเคราะห์ความแปรปรวนจากข้อมูลในตาราง 2 และแสดงผลการวิเคราะห์สัมประสิทธิ์ความเชื่อมั่นตามวิธีของฮอยท์ ฮอยท์ได้อธิบายว่าสูตรของเขามาจากนิยามในเชิง

ทฤษฎีของสัมประสิทธิ์ความเชื่อมั่นโดยใช้สัญลักษณ์ $MS_{persons}$ แทนความแปรปรวนของคะแนนสังเกต และ $MS_{residual}$ แทนความแปรปรวนของคะแนนคลาดเคลื่อนในทฤษฎีความเชื่อมั่นที่ว่า

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

นอกจากนี้ ยังมีสัมประสิทธิ์ความเชื่อมั่นอีกชุดหนึ่ง กรณีแต่ละส่วนในแบบทดสอบคู่ขนานกันแบบ congeneric ดังนี้

ถ้าเราใช้หลักการในการแบ่งกลุ่มของข้อสอบ เราจะให้ข้อสอบแต่ละข้อก็คือส่วน หรือกลุ่มของข้อสอบอาจจะกลายเป็นหนึ่งส่วน และส่วนแต่ละส่วนอาจจะไม่เท่ากัน ในกลุ่ม Congeneric นี้ Raju (1977) ได้มาจากการสรุปอ้างอิงของสัมประสิทธิ์สองส่วน (1970) แล้วปรับให้ใช้ได้กับกรณีมากกว่า 2 ส่วน โดยจะต้องรู้ความยาวของแต่ละส่วน ในเทอมของจำนวนข้อสอบนั้น เราจะให้ $k_1, k_2, \ldots, k_n$ แทนจำนวนข้อของแต่ละส่วนและให้ $\lambda_f = k_f/(\Sigma k_f)$ แล้วสัมประสิทธิ์ของ Raju จะได้ว่า

$$Rr_{XX'} = \frac{\sigma_X^2 - \Sigma\sigma_{X_f}^2}{(1-\Sigma\lambda_f^2)\sigma_X^2} \qquad (23)$$

ถ้าส่วนแต่ละส่วนมีจำนวนข้อเท่ากันแล้ว λ_f จะเท่ากับ 1/n และสัมประสิทธิ์ Raju จะเท่ากับสัมประสิทธิ์แอลฟา นอกจากนี้ยังขยายค่าของสัมประสิทธิ์แอลฟาที่คำนวณจากข้อมูลที่มีแต่ละส่วนเท่ากัน

Kristof (1974) สนใจปัญหาการประมาณค่าความเชื่อมั่นของแบบทดสอบที่ประกอบด้วยแต่ละส่วนที่เป็น Congeneric ในกรณีที่ไม่รู้ความยาวของแต่ละส่วน โดยแบ่งแบบทดสอบออกเป็น 3 ส่วน โมเดลของคะแนนของแต่ละส่วนจะอยู่ในรูปของส่วนที่เป็น congeneric คือ

$X_f = \lambda_f T + C_f + E_f$

สังเกตว่า $\sigma_{X_f X_g} = \lambda_f \lambda_g \sigma_T^2$ และ $\Sigma\lambda_f = 1.0$ (f,g = 1,2,3)

อัตราส่วนของการจับคู่ของความแปรปรวนร่วม จากความจริงที่ว่า $\Sigma\lambda_f = 1.0$ จะได้

$$\lambda_f = \left[\frac{\sigma_{X_g X_h}}{\sigma_{X_f X_g}} + \frac{\sigma_{X_g X_h}}{\sigma_{X_f X_h}} + 1\right]^{-1}$$

คำตอบของ $\lambda_1, \lambda_2, \lambda_3$ อาจจะแทนในการหาสัมประสิทธิ์ของ Raju ผลสุดท้ายจะได้รูปแบบที่ง่ายเรียกว่าสัมประสิทธิ์ของ Kristoff

$$k\rho_{XX'} = \frac{(\sigma_{X_1X_2}\sigma_{X_1X_3} + \sigma_{X_1X_2}\sigma_{X_2X_3} + \sigma_{X_1X_3}\sigma_{X_2X_3})^2}{\sigma_{X_1X_2}\sigma_{X_1X_3}\sigma_{X_2X_3}\sigma_X^2} \qquad (24)$$

สูตรของ Kristoff นี้เป็นสูตรเฉพาะกรณีแบ่งออกเป็น 3 ส่วน แต่ละส่วนคู่ขนานกันแบบ Congeneric

ต่อไปนี้จะเป็นตัวอย่างในการใช้สูตรที่นำเสนอในหัวข้อนี้มาประมาณค่าความเชื่อมั่นแบบความสอดคล้องภายใน

ตาราง 2 ข้อมูลสำหรับประมาณค่าความเชื่อมั่น

	ข้อสอบ						
ผู้สอบ	1	2	3	4	5	6	รวม
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. กรณีแบบทดสอบถูกแบ่งครึ่งออกเป็น 2 ส่วน โดยส่วนแรกประกอบด้วยข้อ 1 – 3 และส่วนที่สองประกอบด้วยข้อ 4 – 6 และสมมติทั้ง 2 ส่วนคู่ขนานกันแบบพาราเรล (Parallel) สัมประสิทธิ์สหสัมพันธ์ระหว่างสองส่วนเท่ากับ 0.82 ประมาณค่าความเชื่อมั่นด้วยสูตรของสเปียร์แมนบราวน์

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

สมมติทั้งสองส่วนคู่ขนานแบบทอ ความแปรปรวนร่วมของทั้งสองส่วนเท่ากับ 0.92 ประมาณค่าความเชื่อมั่นด้วยสูตรของฟลานาแกน

$$\text{F}\,\rho_{xx'} = \frac{4\sigma_{X_1X_2}}{\sigma_X^2} = \frac{4(0.92)}{4.54} = 0.80$$

และสมมติทั้งสองส่วนคู่ขนานแบบ congeneric ประมาณค่าความเชื่อมั่นด้วยสูตรของราชู

$$\text{R}\,\rho_{xx'} = \frac{\sigma_{X_1X_2}}{\lambda_1\lambda_2\sigma_X^2} = \frac{0.92}{(0.5)(0.5)(4.54)} = 0.81$$

และประมาณค่าความเชื่อมั่นด้วยสูตรของแองกอฟและเฟล์ด

$$\text{AFr}_{XX'} = \frac{4\sigma_{X_1X_2}}{\sigma_X^2 - \left(\frac{\sigma_{X_1}^2 - \sigma_{X_2}^2}{\sigma_X}\right)^2} = \frac{4(0.92)}{4.54 - \left(\frac{1.34 - 1.16}{2.13}\right)^2} = 0.81$$

2. กรณีแบบทดสอบถูกแบ่งออก 6 ส่วน โดยแต่ละข้อถือเป็น 1 ส่วน และทุกข้อคู่ขนานกันแบบทอ ประมาณค่าสัมประสิทธิแอลฟา 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 คำนวณได้ด้วยสูตรของทัคเกอร์ (Tucker) ดังนี้

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

ประมาณค่าความเชื่อมั่นด้วยวิธีวิเคราะห์ความแปรปรวนของฮอยท์ ได้ผลการวิเคราะห์ความแปรปรวนดังตาราง 3

ตาราง 3 ตารางสรุปผลการวิเคราะห์ความแปรปรวนจากข้อมูลในตาราง 2

แหล่งความแปรปรวน	Sums Squares	df	Mean Squares
ผู้สอบ	6.817	9	0.7574
ข้อสอบ	1.483	5	0.2967
ความคลาเคลื่อน	6.683	45	0.1485

$$\rho_{XX'} = \frac{MS_{persons} - MS_{residual}}{MS_{persons}} = \frac{0.7574 - 0.1485}{0.7574} = 0.8039$$

กรณีที่ทั้ง 6 ข้อคู่ขนานกันแบบ congeneric และรู้ความยาวของแบบทดสอบ ประมาณค่าสัมประสิทธิ์ความเชื่อมั่นของราชู

$$Rr_{XX'} = \frac{\sigma_X^2 - \Sigma\sigma_{X_f}^2}{(1-\Sigma\lambda_f^2)\sigma_X^2} = \frac{4.54-(1.5)}{(1-0.17)(4.54)} = 0.80$$

สังเกตว่าสัมประสิทธิ์ความเชื่อมั่นของราชูจะเท่ากับสัมประสิทธิ์แอลฟาเมื่อแต่ละส่วนมีจำนวนข้อเท่ากัน

สูตรเฉพาะของคริสทอฟ (Kristof) กรณีแบบทดสอบถูกแบ่งออกเป็น 3 ส่วน แต่ละส่วนคู่ขนานกันแบบ congeneric และไม่รู้ความยาวของแต่ละส่วน โดยส่วนแรกประกอบด้วยข้อสอบข้อ 1 และ 2 ส่วนที่สองประกอบด้วยข้อสอบข้อ 3 และ 4 และส่วนที่สามประกอบด้วยข้อสอบข้อ 5 และ 6 เมตริกความแปรปรวนและความแปรปรวนร่วมของทั้งสามส่วนแสดงได้ดังตาราง 4

ตาราง 4 เมตริกความแปรปรวนและความแปรปรวนร่วมของทั้ง 3 ส่วน

ส่วนที่	1	2	3
1	0.61	0.39	0.47
2	0.39	0.61	0.33
3	0.47	0.33	0.49

ประมาณค่าความเชื่อมั่นด้วยสูตรของคริสทอฟ

$$k\rho_{XX'} = \frac{(\sigma_{X_1X_2}\sigma_{X_1X_3} + \sigma_{X_1X_2}\sigma_{X_2X_3} + \sigma_{X_1X_3}\sigma_{X_2X_3})^2}{\sigma_{X_1X_2}\sigma_{X_1X_3}\sigma_{X_2X_3}\sigma_X^2}$$

$$= \frac{((0.39)(0.47)+(0.39)(0.33)+(0.47)(0.33))^2}{(0.39)(0.47)(0.33)(4.54)}$$

$$= 0.79$$

สัมประสิทธิ์ความเชื่อมั่นที่ไม่มีข้อตกลงเบื้องต้น

นอกจากนี้ Traub (1994 : 87 - 94) ยังได้กล่าวถึงสัมประสิทธิ์ความเชื่อมั่นที่คิดค้นโดย Louis Guttman (1945) ออกมาเป็น 3 สัมประสิทธิ์คือ L_1, L_2 และ L_3 ซึ่ง L_3 มีรูปแบบเท่าเทียมกับสัมประสิทธิ์แอลฟา Guttman ได้แสดงสัมประสิทธิ์ประมาณค่าความเชื่อมั่นที่ไม่มีข้อตกลงเบื้องต้น (no assumptions) ใด ๆ เกี่ยวกับธรรมชาติของความสัมพันธ์ระหว่างแต่ละส่วนของแบบทดสอบ สัมประสิทธิ์ L_1 อย่างง่าย ๆ คือ

$$L_1 = 1 - \frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_X^2}$$

เมื่อ σ_i^2 คือความแปรปรวนของคะแนนที่สังเกตได้ในแต่ละส่วน และ σ_i^2 คือความแปรปรวนของคะแนนที่สังเกตได้ในฉบับเต็ม

อีกสูตรหนึ่งของความเชื่อมั่นที่ไม่มีข้อตกลงเบื้องต้นคือ L_2 นิยามได้ว่า

$$L_2 \quad = \quad 1-\frac{\sum_{i=1}^{n}\sigma_i^2}{\sigma_X^2}+\frac{\sqrt{\frac{n}{(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}^2}}{\sigma_X^2} \qquad (i \neq j)$$

เมื่อ σ_{ij}^2 คือกำลังสองของความแปรปรวนร่วมระหว่างคะแนนสังเกตในคู่ของแบบทดสอบส่วนที่ i และ j (i, j = 1, 2, …, n ; $i \neq j$) สูตร L_2 สามารถประมาณค่าได้มากกว่าหรือเท่ากับสูตรสัมประสิทธิ์แอลฟา แม้ว่าสูตรของ L_2 จะดูเทอะทะลำบากในการคำนวณ แต่ก็เป็นเรื่องง่ายหากเราคำนวณด้วยเครื่องคอมพิวเตอร์

สูตรสเปียร์แมนบราวน์ : กรณีทั่วไป

อีกวิธีการหนึ่งในการประมาณค่าความเชื่อมั่นของแบบทดสอบก็คือสูตรสเปียร์แมนบราวน์ สูตรสเปียร์แมนบราวน์จะใช้ในการทำนายอิทธิพลที่เปลี่ยนแปลงไปของความยาวของแบบทดสอบที่จะมีต่อค่าความเชื่อมั่น สูตรนี้จะอ้างอิงกับการประมาณค่าความเชื่อมั่นแบบแบ่งครึ่ง สูตรสเปียร์แมนบราวน์โดยทั่วไปคือ

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

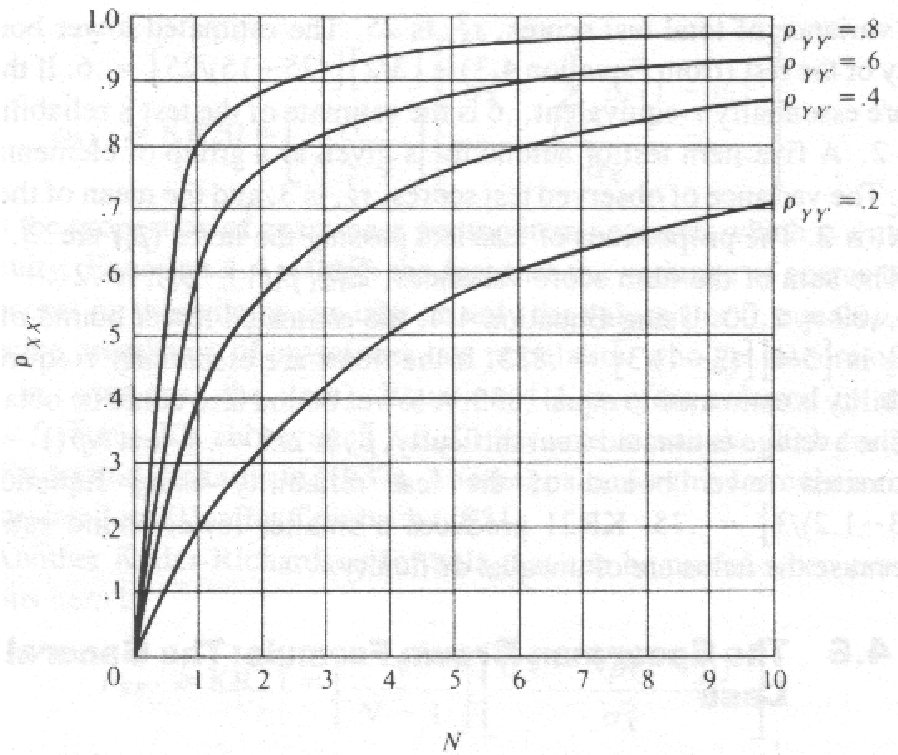
เมื่อ X คือคะแนนรวมของแบบทดสอบที่สังเกตได้จากการรวมคะแนนในแต่ละองค์ประกอบที่คู่ขนานกันแบบพาราเรลของแบบทดสอบม, $X = \sum_{i=1}^{N} Y_i$

Y_i คือคะแนนของแบบทดสอบในแต่ละองค์ประกอบ

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

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

N คือจำนวนขององค์ประกอบที่คู่ขนานกันแบบพาราเรลที่รวมกันเป็นฉบับ X



ภาพประกอบ 5 ความสัมพันธ์ระหว่างความยาวของแบบทดสอบและความเชื่อมั่น

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

ภาพประกอบ 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 5 จะแสดงการเพิ่มขึ้นของความยาวแบบทดสอบที่สัมพันธ์กับความเชื่อมั่นของแบบทดสอบ เมื่อ N -> ∞ แล้ว ความเชื่อมั่นของแบบทดสอบที่มี N องค์ประกอบจะมีค่าถึง 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

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

ในกรณีที่รู้ค่า $\rho_{YY'}$ และ $\rho_{XX'}$ เราสามารถแก้สมการของสเปียร์แมนบราวน์ใหม่ได้ว่า

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

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 5

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

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

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

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

แบบทดสอบใหม่ควรจะมีความยาวเป็น 2.25 เท่าของแบบทดสอบเดิมหรือก็คือ 23 ข้อ

สูตรสเปียร์แมนบราวนี้จะอยู่บนพื้นฐานขององค์ประกอบแต่ละองค์ประกอบในแบบทดสอบต้องคู่ขนานกันแบบพาราเรล รวมทั้งชุดของข้อสอบหรือองค์ประกอบที่เพิ่มเข้าไปในแบบทดสอบด้วย ถ้าเพิ่มอย่างระมัดระวังข้อสอบที่เพิ่มมีความคู่ขนานกับข้อสอบเดิมในฉบับ ความเชื่อมั่นควรจะสูงขึ้น แต่ถ้าเพิ่มอย่างไม่ระวัง ความเชื่อมั่นจะลดต่ำลง อย่างไรก็ตาม แบบทดสอบที่ยาวกว่าย่อมมีความเชื่อมั่นสูงกว่า เพราะว่าข้อตกลงเบื้องต้นของทฤษฎีคะแนนจริงมาตรฐานเดิม เมื่อ N เพิ่มขึ้น ความแปรปรวนของคะแนนจริงจะเพิ่มมากกว่าความแปรปรวนของความคลาดเคลื่อน

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

ในอีกสถานการณ์หนึ่ง สูตรสเปียร์แมนบราวน์สามารถประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง ตัวอย่างเช่น สมมติว่าข้อสอบ 10 ข้อมีความเชื่อมั่น 0.0 สูตรสเปียร์แมนบราวน์คำนวณค่าความเชื่อมั่นเมื่อเพิ่มข้อสอบอีกเท่าตัวที่คู่ขนานกับข้อสอบเดิมได้ [2(0.0)]/[1+(0.0)] = 0.0 อย่างไรก็ตาม ถ้าข้อสอบที่เพิ่มเข้าไปไม่มีความคู่ขนานกับแบบทดสอบเดิมแล้ว ความเชื่อมั่นใหม่ที่คำนวณได้จะได้เท่ากับ 0.7 ความเชื่อมั่นของข้อสอบฉบับใหม่ 20 ข้อจะประมาณค่าได้มากกว่า 0.0 ในกรณีนี้การใช้สูตรสเปียร์แมนบราวน์ที่ไม่เหมาะสมจะทำให้ประมาณค่าความเชื่อมั่นได้ต่ำกว่าความเป็นจริง ผลของความเชื่อมั่นที่ใช้สูตรสเปียร์แมนบราวน์จะมีความถูกต้องเมื่อข้อสอบหรือองค์ประกอบที่เพิ่มเข้าไปมีความคู่ขนานกัน

สามารถประยุกต์ใช้สูตรสเปียร์แมนบราวน์ได้ในอีกสองสถานการณ์คือ สถานการณ์แรกเมื่อต้องการเปรียบเทียบความเชื่อมั่นของแบบทดสอบสองฉบับที่มีความยาวของแบบทดสอบต่างกัน แบบทดสอบที่ยาวมากกว่าดูเหมือนจะมีค่าความเชื่อมั่นสูงกว่า การประยุกต์ใช้สูตรสเปียร์แมนบราวน์จะช่วยให้เราประมาณค่าความเชื่อมั่นของแบบทดสอบถ้าสองฉบับนั้นมีความยาวเท่ากัน สถานการณ์ที่สอง เพราะว่าแบบทดสอบที่สั้นมากมีแนวโน้มจะมีค่าความเชื่อมั่นที่ต่ำกว่าแบบทดสอบที่ยาวกว่า ซึ่งควรจะมีความระมัดระวังเมื่อมีการเปรียบเทียบคะแนนจากแบบทดสอบฉบับสั้น

การเปรียบเทียบวิธีการประมาณค่าความเชื่อมั่น

เราได้อธิบายความแตกต่างของวิธีการประมาณค่าความเชื่อมั่นของแบบทดสอบและวิธีการที่แตกต่างกันนี้จะคำนวณค่าความเชื่อมั่นได้ต่างกัน ในการประมาณค่าความเชื่อมั่นของแบบทดสอบ speed test นั้น ควรจะใช้การประมาณค่าความเชื่อมั่นแบบสอบซ้ำ แบบทดสอบทางเลือกหรือแบบทดสอบคู่ขนาน เพราะว่าถ้าจะใช้การประมาณค่าความเชื่อมั่นด้วยวิธีหาความสอดคล้องภายในจะไม่ถูกต้อง การใช้สัมประสิทธิ์แอลฟาและสูตรคูเดอร์ริชาร์ดสันจะให้ค่าความเชื่อมั่นขั้นต่ำ ซึ่งค่าความเชื่อมั่นขั้นต่ำนี้จะเท่ากับค่าความเชื่อมั่นของแบบทดสอบ ถ้าข้อสอบหรือแต่ละส่วนในแบบทดสอบมีความคู่ขนานกันแบบทอ (essentially τ-equivalent) สัมประสิทธิ์แอลฟาและสูตรคูเดอร์ริชาร์ดสันควรจะใช้เมื่อข้อสอบเป็นเอกพันธ์กัน ถ้าแบบทดสอบที่วัดมีหลายคุณลักษณะ สูตรสัมประสิทธิ์แอลฟาและคูเดอร์ริชาร์ดสันก็จะไม่เหมาะ สูตรสเปียร์แมนบ

ราวน์สามารถประมาณค่าได้ต่ำกว่าหรือสูงกว่าความเป็นจริงถ้าข้อสอบหรือองค์ประกอบในแบบทดสอบไม่คู่ขนานกันแบบพาราเรล เมื่อข้อสอบหรือองค์ประกอบในแบบทดสอบมีความคู่ขนานกันแบบพาราเรลแล้ว สูตรสเปียร์แมนบราวน์จะมีประโยชน์ในการประมาณค่าความเชื่อมั่นของแบบทดสอบเมื่อความยาวเปลี่ยนไป ถ้าข้อสอบหรือแต่ละส่วนในแบบทดสอบมีความคู่ขนานกันแบบ congeneric แล้ว หากแบ่งออกเป็น 2 ส่วนสูตรของฟลานาแกน (Flanagan) ก็จะเหมาะสม แต่ถ้าแบ่งออกเป็นหลายส่วนและรู้ความยาวของแบบทดสอบ ควรใช้สูตรของราชู (Raju) แต่ถ้าไม่รู้ความยาวสูตรและแบ่งออกเป็น 3 ส่วนจะเหมาะกับสูตรของคริสทอฟ (Kristof) แต่ถ้าแบ่งออกเป็นหลายส่วนควรใช้สูตรของแองกอฟและเฟล์ด (Angoff and Feldt)

ความเชื่อมั่นของคะแนนความแตกต่าง

บางโอกาสในการวิจัย ประเมินและการวินิจฉัย เมื่อมีแบบทดสอบ 2 ฉบับที่ใช้ร่วมกัน แต่ตัวแปรที่สนใจก็คือความแตกต่างระหว่างคะแนนของแบบทดสอบ ตัวอย่างเช่น

1. นักจิตวิทยาคลินิกที่ศึกษาเกี่ยวกับเด็กที่ขาดความสามารถในการเรียนรู้ สนใจความขัดแย้งกันระหว่างการทำข้อสอบ 2 ฉบับที่แตกต่างกัน (เช่นกระบวนการทางภาษา และผลผลิตทางภาษา)
2. ผู้ประเมินต้องการตัดสินผลการทำข้อสอบเกินเวลาโดยใช้แบบทดสอบฉบับเดียวกัน

ในกรณีนี้ตัวแปรที่สนใจศึกษาก็คือความแตกต่างของคะแนน นิยามว่า

$$D = X - Y$$

เมื่อ X คือคะแนนที่ได้จากการสอบวัดครั้งแรกหรือฉบับแรก และ Y คือคะแนนที่ได้จากการสอบวัดครั้งที่สองหรือฉบับที่สอง ถ้าต้องการการตัดสินใจเป็นรายบุคคลบนพื้นฐานของความแตกต่างของคะแนน ก็น่าจะคำนวณหาความเชื่อมั่นของคะแนนความแตกต่าง

ถ้าให้ T_X แทนคะแนนจริงของการทดสอบครั้งแรก และ T_Y แทนคะแนนจริงของการทดสอบครั้งที่สองแล้ว ผลต่างของคะแนนจริงก็คือ

$$T_D = T_X - T_Y$$

ส่วนความแปรปรวนสามารถเขียนได้ดังสมการ

$$\sigma_D^2 = \sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}$$

และ

$$\sigma_{T_D}^2 = \sigma_{T_X}^2 + \sigma_{T_Y}^2 - 2\sigma_{XY}$$

จากนิยามพื้นฐานของความเชื่อมั่น เรารู้ว่า $\sigma_{T_X}^2 = \rho_{XX'}\sigma_X^2$ และ $\sigma_{T_Y}^2 = \rho_{YY'}\sigma_Y^2$ ในทั้ง 3 นิยามนี้ เราสามารถแสดง $\sigma_{T_D}^2$ ด้วยคะแนนที่สังเกตได้ดังสมการ

$$\sigma_{T_D}^2 = \rho_{XX'}\sigma_X^2 + \rho_{YY'}\sigma_Y^2 - 2\sigma_{XY}$$

สามารถคำนวณค่าสัมประสิทธิ์ความเชื่อมั่นของคะแนนความแตกต่างได้ด้วยสมการ

$$\rho_{DD'} = \frac{\sigma^2_{T_D}}{\sigma^2_D}$$

สามารถเขียนใหม่ในอีกรูปหนึ่งโดยการแทนค่า $\sigma^2_{T_D}$ และ σ^2_D และแทนค่า σ_{XY} ด้วย $\rho_{XY}\sigma_X\sigma_Y$ จะได้

$$\rho_{DD'} = \frac{\rho_{XX'}\sigma^2_X + \rho_{YY'}\sigma^2_Y - 2\rho_{XY}\sigma_X\sigma_Y}{\sigma^2_X + \sigma^2_Y - 2\rho_{XY}\sigma_X\sigma_Y}$$

ตาราง คะแนนเฉลี่ย ส่วนเบี่ยงเบนมาตรฐาน และความเชื่อมั่นของเครื่องมือวัดผลผลิตทางภาษา และกระบวนการทางภาษา

สถิติ	เครื่องมือวัดผลผลิตทางภาษา	เครื่องมือวัดกระบวนทางภาษา
คะแนนเฉลี่ย	32	40
คะแนนเบี่ยงเบนมาตรฐาน	6	8
ความเชื่อมั่น	0.75	0.80

สมมติว่ามีผู้วิจัยวางแผนจะศึกษาเด็กที่มีความขัดแย้งระหว่างผลผลิตทางภาษาและกระบวนการทางภาษา ความเชื่อมั่นของเครื่องมือวัดทั้งสองฉบับแสดงในตาราง 1 ถ้าผู้วิจัยพบว่าความสัมพันธ์ระหว่างผลผลิตทางภาษาและกระบวนการทางภาษาคือ 0.70 ความเชื่อมั่นของคะแนนความแตกต่างควรจะเป็น

$$\rho_{DD'} = \frac{(0.75)(36) + (0.80)(64) - 2(0.70)(6)(8)}{36 + 64 - 2(0.70)(6)(8)} = 0.34$$

ถ้าในอีกกรณีหนึ่งที่สหสัมพันธ์ระหว่างเครื่องมือทั้งสองเป็น 0.30 แล้ว ความเชื่อมั่นของคะแนนความแตกต่างจะสูงขึ้นเป็น

$$\rho_{DD'} = \frac{(0.75)(36) + (0.80)(64) - 2(0.30)(6)(8)}{36 + 64 - 2(0.30)(6)(8)} = 0.69$$

จะสังเกตได้ว่าค่าความเชื่อมั่นของคะแนนความแตกต่างจะมีค่าสูง เมื่อเราเลือกใช้เครื่องมือวัดแต่ละฉบับที่มีความเชื่อมั่นสูง แต่มีความสัมพันธ์กันต่ำ แต่เงื่อนไขนี้บางครั้งก็ได้รับการยกเว้น ตัวอย่างงานวิจัยของ Zimmerman และ Williams (1982) ได้แสดงคะแนนความแตกต่างที่สามารถเชื่อมั่นได้เมื่อผลคะแนนเป็นแบบทดสอบฉบับเดียวที่สอบก่อนและสอบหลัง

ความเชื่อมั่นขององค์ประกอบ

องค์ประกอบจะนิยามว่าเท่ากับคะแนนรวมของแบบทดสอบตั้งแต่ 2 ฉบับขึ้นไป พิจารณาการสร้างแบบทดสอบสองฉบับที่คู่ขนานกันโดยฉบับหนึ่งคือ A และอีกฉบับหนึ่งคือ B ถ้าสอบฉบับนี้คู่ขนานกันแบบ parallel สัมประสิทธิ์ความเชื่อมั่นของในแต่ละฉบับจะเท่ากับ ρ_{AB} อย่างไรก็ตาม หากมีการใช้แบบทดสอบทั้งสองฉบับที่คู่ขนานกันนี้ คะแนนรวมของผู้สอบแต่ละคนจะเรียกว่าคะแนนองค์ประกอบนิยามว่า

$$C = A + B$$

ความเชื่อมั่นของคะแนนองค์ประกอบ ($\rho_{CC'}$) จะคำนวณได้อย่างไร ได้มีผู้เสนอแนะว่าให้สร้างแบบทดสอบอีก 2 ฉบับที่คู่ขนานกันแบบ parallel และรวมคะแนนองค์ประกอบกับ 2 ฉบับแรกก่อนหน้านี้ ปัญหาก็จะมีว่า จะหาความเชื่อมั่นของคะแนนองค์ประกอบทั้งหมด 4 ฉบับได้อย่างไร ดังนันจึงมีวิธีการหรือนิยามวามเชื่อมั่นขององค์ประกอบในเทอมของคุณลักษณะทางสถิติในองค์ประกอบภายใน ในหัวข้อนี้เราจะพิจารณา 2 วิธีในการพิจารณาความเชื่อมั่นขององค์ประกอบที่สามารถแสดงในเทอมของคุณสมบัติทางสถิติขององค์ประกอบ วิธีการแรกจะใช้กระบวนการที่เรียกว่า Spearman Brown prophecy ที่ให้เราประมาณค่าความเชื่อมั่นขององค์ประกอบของแบบทดสอบคู่ขนานกันแบบ parallel เมื่อเราทราบความเชื่อม่นของแบบทดสอบฉบับหนึ่ง วิธีที่สองเป็นกระบวนที่เรียกว่า Cronbach’s alpha ที่ให้เราประมาณค่าความเชื่อม่นขององค์ประกอบเมื่อเรารู้ความแปรปรวนของคะแนนองค์ประกอบและความแปรปรวนร่วมระหว่างองค์ประกอบทั้งหมด สำหรับวิธีง่าย ๆ ที่จะอธิบายนี้ เราจะเริ่มต้นกับแบบทดสอบที่มีองค์ประกอบคู่ขนานกันแบบ parallel ต่อมาเราจะพิจารณาผลที่ได้ที่มีอิทธิพลถ้าองค์ประกอบของแบบทดสอบไม่ได้คู่ขนานกันแบบ parallel

The Spearman Brown Prophecy

มีข้อตกลงเบื้องต้นดังนี้

1. ความเชื่อมั่นขององค์ประกอบสามารถนิยามได้ว่าเท่ากับ $\rho_{CC'} = \dfrac{\sigma^2_{T_C}}{\sigma^2_C}$

2. การวัดทั้งหมดคู่ขนานกันแบบ parallel สามารถแสดงให้เห็นถึงความเท่ากันของค่าเฉลี่ย ความเท่ากันของคะแนนเบี่ยงเบนมาตรฐาน และความเท่ากันของความแปรปรวน ยิ่งกว่านั้นเมื่อมีแบบทดสอบ k ฉบับ ความสัมพันธ์ระหว่างคู่ของแบบทดสอบที่คู่ขนานกันจะเท่ากันในทุก ๆ คู่

3. ถ้ามี k ด้านในองค์ประกอบ ความแปรปรวนขององค์ประกอบนี้จะมีผลรวมจะเท่ากับผลรวมของความแปรปรวนจำนวน k ค่า และความแปรปรวนร่วมจำนวน k(k - 1) ค่า

เดี๋ยวนี้เราพร้อมที่ได้สูตรสำหรับความเชื่อมั่นขององค์ประกอบ ให้เรานิยามแรกขององค์ประกอบของแบบทดสอบคู่ขนาน k ฉบับนั้นคือ

$$C = A + B + \dots + K$$

ความแปรปรวนของคะแนนสังเกตของององค์ประกอบจะได้ว่า

$$\sigma_C^2 = \sigma_A^2 + \sigma_B^2 + ... + \sigma_K^2 + \sum_{i \neq j} \rho_{ij}\sigma_i\sigma_j \qquad (1)$$

เมื่อ $\sum_{i \neq j} \rho_{ij}\sigma_i\sigma_j$ b คือผลรวมของ k(k - 1) ในเทอมของความแปรปรวน และ i และ j เป็นสัญลักษณ์ในคู่ของแบบทดสอบฉบับ A ถึง K เพราะว่าแบบทดสอบทั้งหมดคู่ขนานกันแบบ parallel และค่า ρ_{ij} ทั้งหมดจะเท่ากัน ดังนั้น

$$\sigma_A = \sigma_B = ... = \sigma_i = \sigma_j$$

ดังนั้นเราสามารถเขียนสมการ (1) ได้ในรูปใหม่ว่า

$$\sigma_C^2 = k\sigma_i^2 + k(k-1)\rho_{ij}\sigma_i^2$$

ขจัดองค์ประกอบของ $k\sigma_i^2$ ออกจากเทอมต่าง เราจะได้

$$\sigma_C^2 = k\sigma_i^2[1 + (k-1)\rho_{ij}] \qquad (2)$$

ผลสุดท้ายสังเกตว่า เพราะ i และ j ถูกวัดอย่างคู่ขนานกันแบบ parallel, ρ_{ij} สามารถเป็นค่าสัมประสิทธิ์ความเชื่อมั่นของแบบทดสอบ i และสมการ (2) สามารถเขียนใหม่ได้ว่า

$$\sigma_C^2 = k\sigma_i^2[1 + (k-1)\rho_{ii'}] \qquad (3)$$

เราจะใช้ผลการแสดงนี้เป็นตัวหารของความเชื่อมั่นขององค์ประกอบ

ให้เรากลับไปสู่ความแปรปรวนของคะแนนจริงสำหรับองค์ประกอบ C ซึ่ง

$$\sigma_{T_C}^2 = \sigma_{T_A}^2 + \sigma_{T_B}^2 + ... + \sigma_{T_K}^2 + \sum_{i \neq j} \rho_{T_iT_j}\sigma_{T_i}\sigma_{T_j}$$

เพราะว่าคะแนนจริงสำหรับผู้สอบแต่ละคนต้องเท่ากับแบบทดสอบคู่ขนานทั้งฉบับ i และฉบับ j ดังนั้น $\rho_{T_iT_j}$ = 1.00 ในแบบทดสอบทุกฉบับ ยิ่งกว่านั้นเพราะว่าเรามีแบบทดสอบคู่ขนาน

$$\sigma_{T_A} = \sigma_{T_B} = \sigma_{T_i} = \sigma_{T_j}$$

ดังนั้นความแปรปรวน k ตัวและความแปรปรวนร่วม k(k - 1) ในองค์ประกอบของคะแนนจริงนี้ เราจะได้

$$\sigma_{T_C}^2 = k\sigma_{T_i}^2 + k(k-1)\sigma_{T_i}^2$$

สามารถเขียนให้อยู่ในรูปอย่างง่ายได้ว่า

$$\sigma_{T_C}^2 = k^2\sigma_{T_i}^2 \qquad (4)$$

ใช้ผลของ $\sigma_{T_C}^2$ (จากสมการ 4) และ σ_C^2 จากสมการ 3 เราจะเขียนสมการสำหรับประมาณค่าความเชื่อมั่นขององค์ประกอบได้ว่า

$$\rho_{CC'} = \frac{k^2\sigma_{T_i}^2}{k\sigma_i^2[1 + (k-1)\rho_{ii'}]}$$

เพราะว่า $\frac{\sigma_{T_i}^2}{\sigma_i^2} = \rho_{ii'}$ สามารถเขียนในอยู่ในรูปอย่างง่ายได้ว่า

$$\rho_{CC'} = \frac{k\rho_{ii'}}{1+(k-1)\rho_{ii'}} \qquad (5)$$

สมการ 5 คือสูตรทั่วไปสำหรับการประมาณค่าความเชื่อมั่นขององค์ประกอบของ Spearman Brown ความเชื่อมั่นขององค์ประกอบนี้สามารถแสดงเป็นความเชื่อมั่นขององค์ประกอบเดียวก็คือ (สมมติว่าทุกองค์ประกอบคู่ขนานแบบ parallel) ความสำคัญในการใช้สูตรนี้ในการประมาณค่าความเชื่อมั่นและพัฒนาแบบทดสอบจะอธิบายต่อไป

ความเชื่อมั่นขององค์ประกอบกับสัมประสิทธิ์แอลฟา

จุดมุ่งหมายทั้งหมดในหัวข้อนี้เป็นการแสดงความเชื่อมั่นขององค์ประกอบที่สามารถแสดงฟังก์ชันของความแปรปรวนของคะแนนงค์ประกบและความแปรปรวนร่วมของแบบทดสอบที่เกิดจากการรวมองค์ประกอบ สูตรที่ได้มาจะเรียกว่า Coefficient alpha ต่อไปนี้จะแสดงที่มาของสูตรที่จะช่วยให้ผู้อ่านได้เข้าใจ

1. สำหรับทุก ๆ คู่ของแบบทดสอบ i และ j ความแปรปรวนร่วมของแบบทดสอบใช้สัญลักษณ์ว่า σ_{ij} หรือ

$$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$$

2. เมื่อแบบทดสอบ i และ j คู่ขนานกันแบบ parallel ความแปรปรวนของคะแนนจริงของกบบทดสอบ i จะเท่ากับความแปรปรวนร่วมของคะแนนจริงบนแบบทดสอบ i และ j หรือ

$$\sigma^2_{T_i} = \sigma_{T_iT_j}$$

3. สำหรับแบบทดสอบทุก ๆ คู่ i และ j ความแปรปรวนร่วมของคะแนนจริงเท่ากับความแปรปรวนร่วมของคะแนนสังเกต หรือ

$$\sigma_{T_iT_j} = \sigma_{ij}$$

4. นิยามขององค์ประกอบ C จะเท่ากับผลรวมของคะแนนแบบทดสอบย่อยที่คู่ขนานกัน k ฉบับ C = A + B + … + K, องค์ประกอบของคะแนนจริงจะเท่ากับ $T_C = T_A + T_B + \ldots + T_K$ จำได้ว่าความแปรปรวนของคะแนนจริงขององค์ประกอบนี้ก็คือ

$$\sigma^2_{T_c} = \sigma^2_{T_A} + \sigma^2_{T_B} + \ldots + \sigma^2_{T_K} + \sum_{i\neq j}\sigma_{T_iT_j}$$

เมื่อ i และ j ในทุกคู่ของแบบทดสอบคู่ขนาน และ $\sum_{i\neq j}\sigma_{T_iT_j}$ คือผลรวมจำนวน k(k - 1) ตัว เพราะแบบทดสอบคู่ขนาน k ฉบับมีความแปรปรวนเท่ากันและความแปรปรวนร่วมกับแบบทดสอบอื่นเท่ากัน

$$\sigma^2_{T_c} = k\sigma^2_{T_i} + k(k-1)\sigma_{T_iT_j} \qquad (6)$$

ยิ่งกว่านั้น เรารู้ว่า $\sigma^2_{T_i} = \sigma_{T_iT_j}$ ดังนั้นเราจะแทนค่าในสมการ 6 จะได้

$$\sigma^2_{T_c} = k\sigma_{T_iT_j} + k(k-1)\sigma_{T_iT_j}$$

สามารถทำให้อยู่ในรูปอย่างง่ายได้ว่า

$$\sigma^2_{T_c} = k^2\sigma_{T_iT_j}$$

จากจุดนี้ $\sigma_{T_iT_j} = \sigma_{ij}$ ดังนั้นแทนค่าในสมการจะได้

$$\sigma^2_{T_c} = k^2\sigma_{ij}$$

ถ้าเราใช้ตัวอักษรแสดงความแปรปรวนของคะแนนจริงขององค์ประกอบ เราสามารถเขียนแสดงค่าความเชื่อมั่นขององค์ประกอบได้เท่ากับ

$$\rho_{CC'} = \frac{k^2\sigma_{ij}}{\sigma^2_C}$$

เมื่อแบบทดสอบทุกฉบับคู่ขนานกันแบบ parallel

ในสถานการณ์การทดสอบจริง เราไม่แน่ใจว่าแบบทดสอบทั้งหมดในองค์ประกอบคู่ขนานกันแบบ parallel จริงหรือไม่ ในกรณีนี้ มันเป็นไปได้ที่จะใช้ผลรวมของความแปรปรวนร่วมของแบบทดสอบและความแปรปรวนขององค์ประกอบในการประมาณค่าต่ำสุดของความเชื่อมั่นขององค์ประกอบ ค่าความเชื่อมั่นต่ำสุดคือสัมประสิทธิที่มีค่าน้อยกว่าสัมประสิทธิ์ความเชื่อมั่น การแสดงผลนี้เราต้องมีความไม่เท่ากันใน 3 ประการคือ

1. เมื่อแบบทดสอบ k ฉบับในองค์ประกอบอาจจะไม่คู่ขนานกันแบบ parallel มีอย่างน้อย 1 ฉบับ (ฉบับ g) ที่มีความแปรปรวนของคะแนนจริงมากกว่าหรือเท่ากับความแปรปรวนร่วมระหว่างแบบทดสอบฉบับ g กับฉบับอื่น หรือ

$$\sigma^2_{T_g} \geq \sigma_{ig}$$

2. สำหรับแบบทดสอบ 2 ฉบับที่อาจจะไม่คู่ขนานแบบ parallel ผลรวมของความแปรปรวนคะแนนจริงจะมากกว่าหรือเท่ากับความแปรปรวนร่วมของทั้งสองฉบับ หรือ

$$\sigma^2_{T_i} + \sigma^2_{T_j} \geq 2\sigma_{ij}$$

3. ผลรวมของความแปรปรวนของคะแนนจริง k ค่าสำหรับแบบทดสอบที่ไม่คู่ขนานแบบ parallel จะมีค่ามากกว่าหรือเท่ากับผลรวมของความแปรปรวนร่วมระหว่างแบบทดสอบทั้งหมด k(k - 1) ค่า หรือ

$$\Sigma\sigma^2_{T_i} \geq \frac{\sum_{i\neq j}\sigma_{ij}}{k-1}$$

อักษรที่แสดงถึงความไม่เท่ากันเป็นผลของการขยายตรรกของแบบทดสอบ k ฉบับ รายละเอียดทางขั้นตอนของพีชคณิตแสดงไว้ใน Lord and Novick (1968) สำหรับผู้อ่านที่สนใจเพิ่มเติมเกี่ยวกับผลรวมของความแปรปรวนร่วมสำหรับแต่ละข้างที่ไม่เท่ากันเราจะได้

$$\Sigma\sigma^2_{T_i} + \sum_{i\neq j}\sigma_{ij} \geq \frac{\sum_{i\neq j}\sigma_{ij}}{k-1} + \sum_{i\neq j}\sigma_{ij}$$

ผลรวมของความแปรปรวนร่วมทางด้านขวาไม่เท่าเทียมกันสามารถรวมกันได้เป็น

$$\frac{\sum_{i\neq j}\sigma_{ij}}{k-1}+\frac{(k-1)\sum_{i\neq j}\sigma_{ij}}{k-1}=\frac{k\sum_{i\neq j}\sigma_{ij}}{k-1}$$

ยิ่งกว่านั้น ด้านซ้ายของความไม่เท่าเทียมกันแสดงได้ด้วย $\sigma^2_{T_c}$ ดังนั้น

$$\sigma^2_{T_c} \geq \frac{k}{k-1}\sum_{i\neq j}\sigma_{ij} \qquad (7)$$

เมื่อ $\sum_{i\neq j}\sigma_{ij}$ คือผลรวมของความแปรปรวนร่วมของแบบทดสอบที่อาจจะไม่คู่ขนานกันแบบ parallel k(k - 1) จำนวน ถ้าเราหารแต่ละข้างของสมการ 7 ด้วย σ^2_C เราจะได้

$$\frac{\sigma^2_{T_c}}{\sigma^2_C} \geq \frac{k}{k-1}\left(\frac{\sum_{i\neq j}\sigma_{ij}}{\sigma^2_C}\right)$$

ซึ่งจะเหมือนกับ

$$\rho_{CC'} \geq \frac{k}{k-1}\left(1-\frac{\Sigma\sigma^2_i}{\sigma^2_C}\right) \qquad (8)$$

ผลที่แสดงทางข้างขวาของสมการ 8 โดยปกติก็คือสัมประสิทธิ์แอลฟา โดยสรุปแล้ว ในทางทฤษฎีสัมประสิทธิ์ความเชื่อมั่นสามารถแสดงคุณลักษณะของสัมประสิทธิ์การตัดสินใจ (เช่น สหสัมพันธ์ที่อ้างอิงระหว่างแบบทดสอบสองฉบับที่คู่ขนานกันแบบ parallel) เมื่อองค์ประกอบของแบบทดสอบไม่คู่ขนานกันแล้ว เราสามารถประมาณค่าสัมประสิทธิ์ขั้นต่ำของการตัดสินใจโดยใช้สัมประสิทธิ์แอลฟา ผลการคำนวณนี้เราต้องรู้จำนวนของแบบทดสอบ ความแปรปรวนของคะแนนองค์ประกอบ และผลรวมของความแปรปรวนร่วมในแบบทดสอบทั้งหมด ประโยชน์ของความเกี่ยวข้องนี้จะปรากฏมากถ้าเราจำได้ว่าทุก ๆ แบบทดสอบอาจจะนำมารวมเป็นองค์ประกอบและข้อสอบแต่ละข้อรวมกันเป็นแบบทดสอบ ดังนั้น สัมประสิทธิแอลฟาจะเป็นวิธีที่สะดวกในการประมาณค่าขั้นต่ำของสัมประสิทธิ์การตัดสินใจของแบบทดสอบโดยใช้ข้อมูลการตอบข้อสอบที่อ้างอิงจากการใช้แบบทดสอบเพียงครั้งเดียว

บรรณานุกรม

Allen, Mary J. and Yen, Wendy M. (1979). **Introduction to Measurement Theory**. Monterey : Brooks/Cole Publishing Company.

Crocker, Linda and Algina, James. **Introduction to Classical and Modern Test Theory**. New York : CBS College Publishing, 1986.

Feldt, Leonard S. and Brennan, Robert L. "Reliability," in **Educational Measurement**. Linn, Robert L. (Ed.). Third Edition. U.S.A. : Macmillan Publishing Company, 1989.

Gulliksen, Harold. **Theory of Mental Tests**. U.S.A. John Wiley & Sons, Inc., 1950.

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

Trochim, William M.K. **Research Methods Knowledge Base**. 2nd Edition. http://trochim.human.cornell.edu/kb/. 1999.

จัดทำเสร็จเมื่อเดือนมกราคม พ.ศ.2547