บทที่ 3

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้เป็นการวิเคราะห์แบบแผนการตอบข้อสอบ (Item Response Pattern) ด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคล (Person-fit index) ซึ่งมีวิธีประมาณค่าที่แตกต่างกันภายใต้ข้อตกลงเบื้องต้นของทฤษฎีการตอบสนองข้อสอบ ซึ่งมีวิธีดำเนินการวิจัยตามขั้นตอน ดังต่อไปนี้

1. ศึกษาค้นคว้าเกี่ยวกับ ทฤษฎี และหลักการ ของดัชนีบ่งชี้ความเหมาะสมของบุคคล จากเอกสาร ตำราทางวิชาการ วารสาร และรายงานการวิจัยทางด้านการวัดผลการศึกษา

2. สร้างโปรแกรมสุ่มข้อมูล (Random Program) และโปรแกรมวิเคราะห์ความเหมาะสมของบุคคลกับโมเดลการตอบข้อสอบซึ่งประมาณค่าด้วยฟังก์ชันไลค์ลิฮูด (L_z index) โดยเขียนด้วยภาษา FOXPRO

3. จำลองข้อมูลการตอบข้อสอบของผู้สอบตามตัวแปรที่กำหนด

4. ตรวจสอบความเป็นเอกมิติของแบบสอบที่จำลองด้วยโปรแกรม IRTDATA โดยการวิเคราะห์องค์ประกอบ (Factor Analysis)

5. วิเคราะห์ข้อมูลเบื้องต้นเพื่อให้ทราบลักษณะของข้อมูลในแต่ละเงื่อนไขที่ศึกษา โดยใช้ค่าสถิติพื้นฐาน ได้แก่ ค่ามัชฌิมเลขคณิต ส่วนเบี่ยงเบนมาตรฐาน ความเบ้ ความโด่ง และตรวจสอบคุณภาพในด้านความเที่ยง ด้วยโปรแกรม BILOG

6. วิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคล โดยวิธีประมาณค่าดัชนีด้วยฟังก์ชันไลค์ลิฮูด (The Likelihood-based Approach) ด้วยโปรแกรมที่ผู้วิจัยสร้างขึ้น

7. วิเคราะห์แบบแผนการตอบด้วยดัชนีบ่งชี้ความเหมาะสมของบุคคล โดยวิธีประมาณค่าดัชนีจากคะแนนส่วนที่เหลือ (The Score Residual Approach) ด้วยโปรแกรม BIGSTEPS

8. หาค่าสารสนเทศของแบบสอบหลังการตรวจสอบผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ด้วยโปรแกรม BILOG

9. เปรียบเทียบค่าสารสนเทศของแบบสอบโดยการคำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบสอบก่อนและหลังการตรวจสอบความเหมาะสมของผู้สอบ

10. หาความเที่ยง (reliability) ของแบบสอบ ที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบโดยใช้สูตร KR-20

11. หาความตรงเชิงทฤษฎี (Construct validity) โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) ด้วยโปรแกรม LISREL

12. ทดสอบความแตกต่างของค่าความเที่ยงของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบทั้ง 2 วิธี ด้วยค่าสถิติฟิชเชอร์ - ซี (Fisher-Z)

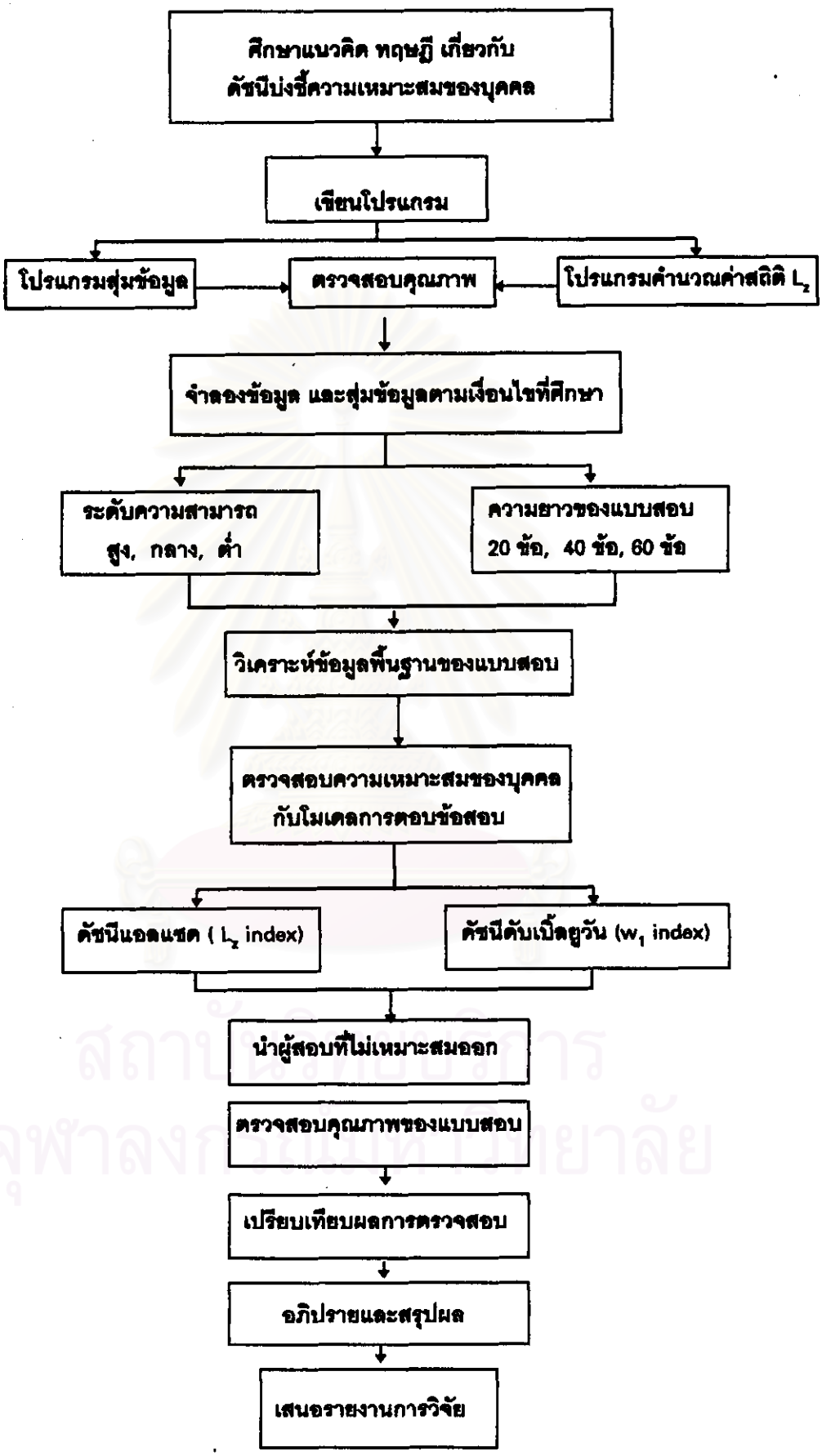
13. วิเคราะห์ความแตกต่างของค่าความตรงของแบบสอบที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบทั้ง 2 วิธี ด้วยค่าสถิติไค - สแควร์สัมพัทธ์ (Relatived Chi-square)

14. อภิปราย และสรุปผลการวิจัย

15. เสนอรายงานการวิจัย

สามารถแสดงวิธีดำเนินการวิจัยได้ดังแผนภาพที่ 1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย


ศึกษาแนวคิด ทฤษฎี เกี่ยวกับ
ดัชนีบ่งชี้ความเหมาะสมของบุคคล
เขียนโปรแกรม
โปรแกรมสุ่มข้อมูล
ตรวจสอบคุณภาพ
โปรแกรมคำนวณค่าสถิติ L_z
จำลองข้อมูล และสุ่มข้อมูลตามเงื่อนไขที่ศึกษา
ระดับความสามารถ
สูง, กลาง, ต่ำ
ความยาวของแบบสอบ
20 ข้อ, 40 ข้อ, 60 ข้อ
วิเคราะห์ข้อมูลพื้นฐานของแบบสอบ
ตรวจสอบความเหมาะสมของบุคคล
กับโมเดลการตอบข้อสอบ
ดัชนีลอลแชด (l_z index)
ดัชนีดับเบิ้ลยูวัน (w_1 index)
นำผู้สอบที่ไม่เหมาะสมออก
ตรวจสอบคุณภาพของแบบสอบ
เปรียบเทียบผลการตรวจสอบ
อภิปรายและสรุปผล
เสนอรายงานการวิจัย

ตัวแปรที่ศึกษา

1. ตัวแปรอิสระ (Independent Variable) มี 3 ตัว คือ
   1.1 ดัชนีบ่งชี้ความเหมาะสมของบุคคล มี 2 วิธี คือ
      1.1.1 ดัชนีแอลแซด
      1.1.2 ดัชนีดับเบิ้ลยูวัน แบ่งออกเป็น 2 กรณี คือ
         1) การใช้ค่าสถิติ infit เป็นเกณฑ์ในการพิจารณา
         2) การใช้ค่าสถิติ outfit เป็นเกณฑ์ในการพิจารณา
   1.2 ความยาวของแบบสอบ มี 3 ขนาด คือ 20, 40 และ 60 ข้อ
   1.3 ระดับความสามารถของผู้สอบ มี 3 ระดับ คือ
      1.3.1 กลุ่มผู้สอบที่มีความสามารถสูง ($\bar{\theta}$= 1.5)
      1.3.2 กลุ่มผู้สอบที่มีความสามารถปานกลาง ($\bar{\theta}$= 0.0)
      1.3.3 กลุ่มผู้สอบที่มีความสามารถต่ำ ($\bar{\theta}$= -1.5)
2. ตัวแปรตาม (Dependent Variable) เป็นผลของวิธีตรวจสอบความเหมาะสมบุคคล มี 3 ตัว คือ
   2.1 ค่าสารสนเทศของแบบสอบ (Test Information Function)
   2.2 ความเที่ยงของแบบสอบ (Reliability)
   2.3 ความตรงตามทฤษฎี (Construct Validity)

พลวิจัย

พลวิจัยที่ใช้ในการวิจัยครั้งนี้เป็นการจำลองเมตริกซ์คำตอบจากโปรแกรม IRTDATA ซึ่งพัฒนา โดย ยอร์ช เอ โจแฮนสัน (Johanson,1992) มหาวิทยาลัยโอไฮโอ สหรัฐอเมริกา ลักษณะของโปรแกรมเป็นดังนี้

1. IRTDATA เป็นโปรแกรมสร้างข้อมูลการตอบแบบสอบของผู้สอบ ที่ให้คะแนนเป็น 0,1 โดยสร้างขึ้นตามทฤษฎีการตอบสนองข้อสอบ สามารถกำหนดให้เป็น โมเดลการตอบสนองข้อสอบแบบ 1,2 และ 3 พารามิเตอร์

2. โปรแกรมมีสมรรถนะในการจำลองข้อมูลตามลักษณะและสถานการณ์ของการสอบ โดยมีจำนวนผู้สอบและข้อสอบไม่เกิน 3,000 คน และ 200 ข้อ ตามลำดับ

3. การสร้างเมตริกซ์คำตอบ ในการจำลองข้อมูลสามารถกำหนดลักษณะการแจกแจงของข้อมูล ค่าพารามิเตอร์ของผู้สอบและค่าพารามิเตอร์ของข้อสอบได้ตามลักษณะ ดังนี้

3.1 ข้อมูลมีการแจกแจงเป็นแบบสม่ำเสมอ (Uniform distribution) หรือ การแจกแจงเป็นแบบปกติ (Normal distribution)

3.2 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าพารามิเตอร์ของผู้สอบ (ability of examinees)

3.3 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของค่าพารามิเตอร์ของข้อสอบ (a, b และ c)

3.4 จำนวนผู้สอบ

3.5 จำนวนข้อสอบ

4. ข้อมูลที่ได้จากการจำลอง จะเป็นข้อมูล 3 ลักษณะคือ

4.1 ค่าพารามิเตอร์ของข้อสอบ (a, b และ c)

4.2 ค่าพารามิเตอร์ของผู้สอบ เป็นคะแนนดิบ (จำนวนข้อที่ตอบถูก) และ คะแนนจริง (true ability)

4.3 เมตริกซ์ของคำตอบซึ่งเป็นคะแนนดิบของการตอบข้อสอบของของผู้สอบแต่ละคน

5. ข้อมูลที่สร้างขึ้นสามารถนำไปใช้เป็นข้อมูลนำเข้า (Input data) สำหรับการวิเคราะห์โดยใช้โปรแกรมอื่นได้ ซึ่งผู้วิจัยได้สร้างโปรแกรมสุ่ม (Random Program) เพื่อสุ่มข้อมูลที่จำลองขึ้นไปใช้ศึกษา ตามตัวแปรและเงื่อนไขที่กำหนด

วิธีการจำลองข้อมูล

ผู้วิจัยจำลองข้อมูลการตอบข้อสอบของผู้สอบตามตัวแปรที่กำหนดโดยดำเนินการ ดังนี้

1. จำลองเมตริกซ์คำตอบตามสมรรถนะสูงสุดของโปรแกรม IRTDATA คือ ผู้สอบจำนวน 3,000 คน ข้อสอบ 200 ข้อ จำนวน 3 เมตริกซ์ โดยกำหนดให้ความสามารถ

เฉลี่ยของผู้สอบในแต่ละเมตริกซ์ (mean of ability) เป็น -1.5, 0 และ +1.5 ลำดับ ค่าพารามิเตอร์ของข้อสอบ (a, b และ c) มีค่าเฉลี่ยเท่ากันในทุกเมตริกซ์ และมีการแจกแจงเป็นโค้งปกติ

2 ทำการสุ่มข้อมูลจากเมตริกซ์ ทั้ง 3 เมตริกซ์ ตามเงื่อนไขที่ต้องการศึกษา ด้วยโปรแกรมสุ่มข้อมูลซึ่งผู้วิจัยสร้างขึ้น ดังนี้

2.1 การสุ่มครั้งที่ 1 ทำการสุ่มผู้สอบซึ่งมีระดับความสามารถสูง ($\bar{\theta}$ =1.5) จากเมตริกซ์คำตอบซึ่งผู้สอบมีค่าความสามารถเฉลี่ย เท่ากับ +1.5 จำนวน 1,000 คน และสุ่มข้อสอบที่มีความยาวแตกต่างกัน 3 ชุด คือ 20, 40, และ 60 ข้อ

2.2 การสุ่มครั้งที่ 2 ทำการสุ่มผู้สอบซึ่งมี ระดับความสามารถปานกลาง ($\bar{\theta}$ = 0)) จากเมตริกซ์คำตอบซึ่งผู้สอบมีระดับความสามารถเฉลี่ย เท่ากับ 0 จำนวน 1,000 คน และสุ่มข้อสอบที่มีความยาวแตกต่างกัน 3 ชุด คือ 20, 40, และ 60 ข้อ

2.3 การสุ่มครั้งที่ 3 ทำการสุ่มผู้สอบซึ่งมี ระดับความสามารถต่ำ ($\bar{\theta}$ = -1.5) จากเมตริกซ์คำตอบซึ่งผู้สอบมีค่าความสามารถเฉลี่ย เท่ากับ -1.5 จำนวน 1,000 คน และสุ่มข้อสอบที่มีความยาวแตกต่างกัน 3 ชุด คือ 20, 40, และ 60 ข้อ

3. ข้อมูลที่ใช้ศึกษา เป็นเมตริกซ์คำตอบ ซึ่งมีความยาวของแบบสอบและระดับความสามารถของผู้สอบแตกต่างกัน รวมทั้งสิ้น 18 เงื่อนไข ดังตารางที่ 3

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3 ข้อมูลที่ใช้ศึกษา

ดัชนีบ่งชี้ความเหมาะสมของบุคคล	ความยาวของแบบสอบ	ระดับความสามารถเฉลี่ย
ดัชนีแอลแซด	20	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5
	40	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5
	60	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5
ดัชนีดับเบิ้ลยูวัน	20	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5
	40	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5
	60	$\bar{\theta}$ = +1.5
		$\bar{\theta}$ = 0.0
		$\bar{\theta}$ = -1.5

การศึกษาผลการตรวจสอบความเหมาะสมของบุคคลในครั้งนี้ ผู้วิจัยทำการสุ่มข้อมูลเป็นผลการตอบข้อสอบของผู้สอบเงื่อนไขละ 1,000 คน โดยกำหนดให้ระดับความสามารถของผู้สอบ และระดับความยาวของแบบสอบ แตกต่างกัน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

การวิเคราะห์ข้อมูล

การวิเคราะห์ข้อมูลในครั้งนี้ใช้ดัชนีบ่งชี้ความเหมาะสมของบุคคล 2 วิธี แต่เนื่องจากการประมาณค่าดัชนีบ่งชี้ความเหมาะสมของบุคคลจากคะแนนส่วนที่เหลือ มีค่าสถิติซึ่งเป็นเกณฑ์ในการพิจารณาความเหมาะสมกับโมเดล 2 ค่า คือ ค่าสถิติ infit และ ค่าสถิติ outfit ซึ่งมีสูตรในการคำนวณแตกต่างกัน ผู้วิจัยจึงศึกษาดัชนี W_i ใน 2 กรณี คือ W_i(infit) และ W_i(outfit) โดยใช้ค่าสถิติ infit และ outfit เป็นเกณฑ์ในการวิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ตามลำดับ ผู้วิจัยดำเนินการวิเคราะห์ข้อมูลตามลำดับต่อไปนี้

1. คำนวณค่าสถิติพื้นฐาน เพื่อบรรยายลักษณะของข้อมูลตามเงื่อนไขที่ศึกษา โดยใช้โปรแกรมวิเคราะห์ข้อสอบ BILOG ด้วยโปรแกรมย่อย IS11 ซึ่งเป็นการวิเคราะห์ข้อสอบตามทฤษฎีการสอบแบบดั้งเดิม (Classical Test Theory)

2. วิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีแอลแซด (L_z index) ตามขั้นตอน ดังนี้

   2.1 วิเคราะห์ข้อมูลในแต่ละเงื่อนไขเพื่อเตรียมค่าพารามิเตอร์ของข้อสอบและค่าพารามิเตอร์ของผู้สอบ ด้วยโปรแกรม BILOG

   2.2 วิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ (person-fit) ซึ่งมีวิธีการประมาณค่าดัชนี L_z ดังนี้

      2.2.1 ประมาณค่าความน่าจะเป็นที่ผู้สอบคนที่ i จะตอบข้อสอบข้อที่ j ได้ถูกต้อง โดยใช้ค่าพารามิเตอร์ที่ประมาณค่าได้จากข้อ 2.1

$$P_i(\theta) = C_i + (1 - C_i)\frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

เมื่อ

$P_i(\theta)$ คือ โอกาสที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูก

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

b_i คือ ระดับความยากของข้อสอบข้อที่ i

c_i คือ ค่าโอกาสการเดาถูก

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

D คือ scaling factor ของฟังก์ชั่นโลจิสติกมีค่าเท่ากับ 1.7

2.2.2 ประมาณค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics)

$$l_0 = \sum_{i=1}^{n} \{ U_i[\ln P_i(\theta)] + (1-U_i) * \langle \ln[1-P_i(\theta)] \rangle \}$$

2.2.3 แปลงค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics) เป็นค่าดัชนีมาตรฐาน (standardized person-fit indexes)

$$l_z = \frac{l_0 - E(l_0)}{[Var(l_0)]^{1/2}}$$

เมื่อ $E(L_0)$ คือ ค่าที่คาดหวังของ L_0

$Var(L_0)$ คือ ความแปรปรวนของ L_0

โดย $$E(l_0) = \sum_{i=1}^{n} \left\{ P_i(\hat{\theta}) InP_i(\hat{\theta}) + \left[1 - P_i(\hat{\theta})\right] \ln\left[1 - P_i(\hat{\theta})\right] \right\}$$

โดย $$Var(L_0) = \sum_{i=1}^{n} P_i(\theta)[1-P_i(\theta)] \left\{ \ln\left[\frac{P_i(\theta)}{1-P_i(\theta)}\right] \right\}^2$$

1 1-2เมื่อ $P_i(\theta)$ คือ โอกาสที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

$1\text{-}P_i$ คือ โอกาสที่ผู้สอบจะตอบข้อสอบข้อที่ i ผิด

U_i คือ พฤติกรรมการตอบที่สังเกตได้ (0,1)

θ คือ ค่าประมาณความสามารถของผู้สอบ

2.2.4 ทดสอบสมมติฐานโดยใช้สถิติทดสอบ ซี (z-test) แบบสองทาง

2.2.4.1 บุคคลที่ไม่เหมาะกับโมเดลการตอบ (person-misfit) คือ ผู้สอบที่มี $|L_z|$ มากกว่า 1.96 : $\alpha = .05$

2.2.5 นำผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ (person- misfit) ออกจากการวิเคราะห์

2.3 วิเคราะห์คุณภาพของแบบสอบในแต่ละเงื่อนไข โดยการคำนวณค่าสารสนเทศของแบบสอบ (TIF) และค่าความเที่ยงเชิงความสอดคล้องภายใน ด้วยสูตร KR-20 ของคูเดอร์ ริชาร์ดสัน (Kuder-Richardson) ดังนี้

2.3.1 คำนวณค่าสารสนเทศของแบบสอบ ตามสูตรการคำนวน ดังนี้

$$I(\theta) = \sum_{i=1}^{k} I_i(\theta)$$

เมื่อ $I_i(\theta)$ คือ ค่าสารสนเทศของข้อสอบข้อที่ I

2.3.2 คำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบ ดังนี้

$$\text{RE (Y,X)} = \frac{I_Y(\theta)}{I_X(\theta)}$$

เมื่อ RE (Y,X) คือ ประสิทธิภาพสัมพัทธ์ของแบบสอบ Y เมื่อเทียบกับ X

$I_Y(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบสอบ Y เมื่อผู้สอบมีระดับความสามารถ θ

$I_X(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบสอบ X เมื่อผู้สอบมีระดับความสามารถ θ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.3.2 คำนวณค่าความเที่ยง ด้วยสูตร KR 20 ดังนี้

$$r_{tt} = \left[\frac{n}{n-1}\right]\left[1-\frac{\sum pq}{s_x^2}\right]$$

เมื่อ r_{tt} คือ ค่าความเที่ยง

n คือ จำนวนข้อสอบ

Σpg คือ ผลรวมของความแปรปรวนรายข้อ

S^2_x คือ ความแปรปรวนของคะแนนรวม

P คือ สัดส่วนของผู้ที่ตอบถูก

Q คือ สัดส่วนของผู้ที่ตอบผิด

3. วิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีดับเบิ้ลยูวัน ($W_i$index) ตามขั้นตอนดังนี้

3.1 วิเคราะห์ผู้สอบที่เหมาะสมกับโมเดลการตอบข้อสอบ (person-fit) ด้วยโปรแกรม BIGSTEPS ซึ่งประมาณค่าด้วยวิธี PROX และ UCON ตามขั้นตอนต่อไปนี้

3.1.1 ประมาณค่าพารามิเตอร์ของข้อสอบและความสามารถของผู้สอบ โดยโมเดลโลจิสติกแบบหนึ่งพารามิเตอร์

3.1.2 ประมาณค่าความน่าจะเป็นที่ผู้สอบคนที่ i จะตอบข้อสอบข้อที่ j ได้ถูกต้อง โดยใช้ค่าพารามิเตอร์ที่ประมาณได้จากข้อ 1

$$P_i(\theta) = \frac{e^{D(\theta-b_i)}}{1+e^{D(\theta-b_i)}}$$

เมื่อ

$P_i(\theta)$ คือ โอกาสที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

$P_i(\theta)$ คือ โอกาสที่ผู้สอบที่มีความสามารถ θ จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

θ คือ ระดับความสามารถที่แท้จริงของผู้สอบ

b_i คือ ระดับความยากของข้อสอบข้อที่ i

e คือ ค่าคงที่มีค่าเท่ากับ 2.7182818

D คือ scaling factor เพื่อปรับเปลี่ยนค่าของฟังก์ชั่นโลจิสติก มีค่าเท่ากับ 1.7

3.1.3 ประมาณค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics) Unweighted total mean square : MNSQ - outfit

$$W_i = \frac{1}{N}\sum_{j=1}^{N}\frac{(U_{ij} - P_{ij})^2}{P_{ij}(1-P_{ij})}$$

เมื่อ N คือ จำนวนผู้เข้าสอบ

บุคคลที่มีค่า Unweighted total mean square มากกว่า 1.20 เป็นผู้สอบที่ไม่เหมาะกับโมเดลการตอบ

3.1.4 แปลงค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics) เป็นค่าดัชนีมาตรฐาน (standardized mean square infit statistic) : PERSON OUTFIT

$$t_i = (v_i^{1/3} - 3/q_i) + (q_i/3)$$

เมื่อ t_i คือ ค่าดัชนีความเหมาะสมของคนที่ j

v_i คือ ค่า unweighted mean square ของคนที่ j

q_i คือ ส่วนเบี่ยงเบนมาตรฐานของ unweighted mean square

บุคคลที่มี |ค่าสถิติ OUTFIT| มากกว่า 2.00 เป็นผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ

3.1.5 ประมาณค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics) : Weight total mean square : MNSQ - infit

$$W_1 = \frac{\sum_{i=1}^{n} (U_{ij} - P_{ij})^2}{\sum_{j=1}^{N} P_{ij}(1-P_{ij})}$$

(Smith, 1994; Linacre & Wright, 1994)

เมื่อ

U_{ij} คือ พฤติกรรมการตอบข้อสอบที่สังเกตได้ (0 , 1) ของผู้สอบคนที่ j ข้อสอบข้อที่ i

P_{ij} คือ โอกาสที่ผู้สอบคนที่ j จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

บุคคลที่มีค่า MNSQ-infit มากกว่า 1.20 เป็นผู้สอบที่ไม่เหมาะกับโมเดลการตอบข้อสอบ

3.1.6 แปลงค่าสถิติบ่งชี้ความเหมาะสมของบุคคล (person-fit statistics) เป็นค่าดัชนีมาตรฐาน (standardized mean square outfit statistic) : PERSON INFIT

$$t_i = (v_i^{1/3} - 3/q_i) + (q_i/3)$$

เมื่อ t_i คือ ค่าดัชนีความเหมาะสมของคนที่ i

v_i คือ ค่า weighted mean square ของคนที่ i

q_i คือ ส่วนเบี่ยงเบนมาตรฐานของ weighted mean square

ค่าดัชนีตัวนี้ เรียกว่า Person outfit เมื่อวิเคราะห์ความเหมาะสมของบุคคลกับโมเดล

บุคคลที่มี |ค่าสถิติ INFIT| มากกว่า 2.00 เป็นผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ

3.2 นำผู้สอบที่ไม่เหมาะสมกับโมเดลการตอบข้อสอบ (person-misfit) ออกจากการวิเคราะห์

3.2.1 คำนวณค่าสารสนเทศของแบบสอบ ตามสูตรการคำนวณ ดังนี้

เมื่อ $I_i(\theta)$ คือ ค่าสารสนเทศของข้อสอบข้อที่ i

3.2.2 เปรียบเทียบค่าสารสนเทศของแบบสอบโดยการคำนวณค่าประสิทธิภาพสัมพัทธ์ของแบบสอบ ดังนี้

$$RE\ (Y,X) = \frac{I_Y(\theta)}{I_X(\theta)}$$

เมื่อ RE (Y,X) คือ ประสิทธิภาพสัมพัทธ์ของแบบสอบ Y เมื่อเทียบกับ X

$I_Y(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบสอบ Y เมื่อผู้สอบมีระดับความสามารถ θ

$I_X(\theta)$ คือ ค่าฟังก์ชันสารสนเทศของแบบสอบ X เมื่อผู้สอบมีระดับความสามารถ θ

ถ้าค่า RE(Y,X) > 1.00 แสดงว่า แบบสอบชุด Y มีค่าสารสนเทศสูงกว่าแบบสอบชุด X ในระดับความสามารถนั้นๆ

3.2.3 หาความเที่ยง (reliability) ของแบบสอบ ที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบโดยใช้สูตร KR-20 (Kuder-Richardson) ดังนี้

$$r_{tt} = \left[\frac{n}{n-1}\right]\left[1-\frac{\sum pq}{s_x^2}\right]$$

เมื่อ r_{tt} คือ ค่าความเที่ยง

n คือ จำนวนข้อสอบ

Σpg คือ ผลรวมของความแปรปรวนรายข้อ

S_x^2 คือ ความแปรปรวนของคะแนนรวม

P คือ สัดส่วนของผู้ที่ตอบถูก

Q คือ สัดส่วนของผู้ที่ตอบผิด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Q คือ สัดส่วนของผู้ที่ตอบผิด

4. เปรียบเทียบผลการตรวจสอบความเหมาะสมของบุคคล โดยการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีแอลแซด และดัชนีดับเบิ้ลยูวัน ด้านความเที่ยงของแบบสอบ ในแต่ละเงื่อนไข โดยทดสอบนัยสำคัญของความแตกต่างของค่าความเที่ยงของแบบสอบหลังการวิเคราะห์ด้วยดัชนีบ่งชี้ความเหมาะสมทั้งสองวิธีที่ได้จากการ ด้วยค่าสถิติซี (Z-test)

$$Z = \frac{Z_{rtt1} - Z_{rtt2}}{\sqrt{\frac{1}{n_1 - 3} - \frac{1}{n_2 - 3}}}$$

(ศิริชัย กาญจนวาสี, 2532)

เมื่อ Z_{rtt1} คือ ค่าฟิชเชอร์ - ซี ของ ค่าความเที่ยงของแบบสอบชุดที่ 1

Z_{rtt2} คือ ค่าฟิชเชอร์ - ซี ของ ค่าความเที่ยงของแบบสอบชุดที่ 2

n คือ จำนวนข้อสอบในแต่ละฉบับ

5. วิเคราะห์ข้อสอบที่เหมาะสมกับโมเดลการตอบข้อสอบ (item-fit) โดยใช้ข้อมูลที่ได้จากการวิเคราะห์แบบแผนการตอบข้อสอบด้วยดัชนีทั้งสองวิธี โดยการทดสอบค่าสถิติไค - สแควร์

5.1 ประมาณค่าพารามิเตอร์ของผู้สอบและข้อสอบด้วย โมเดลโลจิสติกแบบ 2 พารามิเตอร์ ด้วยโปรแกรม BILOG

5.2 วิเคราะห์ข้อสอบที่เหมาะสมกับโมเดลการตอบข้อสอบ ด้วยการทดสอบค่าสถิติไค-สแควร์ ดังนี้

$$\chi^2 = \sum_{j=1}^{N} \frac{n_{ij}(O_{ij} - E_{ij})^2}{E_{ij}(1 - E_{ij})}$$

(Rost and Davier,1994)

เมื่อ O_{ij} คือ พฤติกรรมการตอบข้อสอบที่สังเกตได้ (0,1)

E_{ij} คือ โอกาสที่ผู้สอบแต่ละคนในกลุ่ม j จะตอบข้อสอบข้อที่ i ได้ถูกต้อง

n_{ij} คือ จำนวนผู้สอบในกลุ่ม j เมื่อตอบข้อสอบข้อที่ i

j คือ จำนวนกลุ่มของผู้สอบที่อยู่ในช่วงของโอกาสที่จะตอบข้อสอบข้อที่ i ได้ถูกต้องเดียวกัน

ข้อสอบที่ไม่เหมาะกับโมเดลการตอบ (item-misfit) คือ
ข้อสอบที่มีค่า χ^2 มากกว่า ค่า χ^2 ที่มี df เท่ากับ j - 1

6. หาความตรงเชิงทฤษฎี (Construct validity) โดยการวิเคราะห์องค์ประกอบเชิงยืนยัน (Confirmatory Factor Analysis) ด้วยโปรแกรม LISREL

7. เปรียบเทียบคุณภาพด้านความตรงตามทฤษฎีของแบบสอบในแต่ละเงื่อนไข หลังจากการตรวจสอบความเหมาะสมของบุคคลด้วยดัชนีทั้งสองวิธี โดยเปรียบเทียบจากค่าสถิติไค-สแควร์สัมพัทธ์ (Relative chi-square: χ^2/ df)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย