การวิเคราะห์หาความลำเอียงของข้อทดสอบ 3 วิธี

สุพัฒน์ สุกมลสันต์

การวิเคราะห์ข้อทดสอบรายข้อ เพื่อพัฒนาแบบทดสอบให้มีคุณภาพดียิ่งขึ้นนั้นสามารถทำได้หลายรูปแบบ เช่น ใช้รูปแบบประเพณีนิยม (classical model) รูปแบบราสช์ (Rasch Model) และรูปแบบโลจิส (Logistic Model) แต่วิธีการใหม่ที่เพิ่งได้รับความสนใจอย่างจริงจังเมื่อประมาณ ปี ค.ศ. 1970 คือการวิเคราะห์หาความลำเอียง (item bias) ของข้อทดสอบ (Berk, 1982 : 1) ปัจจุบันนี้นักทดสอบได้นำวิธีการทางสถิติหลายวิธีมาใช้เพื่อการวิเคราะห์หาความลำเอียงของข้อทดสอบ แต่วิธีที่ได้รับความสนใจมากมี 3 วิธีคือ

1. วิธีกำหนดจุดเดลต้า (Delta-Plot Method)
2. วิธีทดสอบความแตกต่างของสัดส่วนระหว่างกลุ่มด้วยไคสแควร์ (Chi-square Method)
3. วิธีการตอบสนองของข้อทดสอบ ชนิด 3 พารามิเตอร์ (Three-parameter Logistic Method)

เพื่อให้ผู้อ่านที่สนใจสามารถทำการวิเคราะห์หาความลำเอียงของข้อทดสอบได้โดยง่าย ผู้เขียนใคร่ขอแนะนำวิธีการสำหรับการวิเคราะห์ทั้ง 3 ดังกล่าวแล้ว ดังต่อไปนี้

1. วิธีกำหนดจุดเดลต้า (Delta-Plot Method)

ก. แนวคิด

การวิเคราะห์วิธีนี้มีความเชื่อว่าความลำเอียงของข้อทดสอบหมายถึงข้อทดสอบที่มีค่าความยากง่ายหรือเดลต้า (Δ) ของผู้สอบแต่ละกลุ่มเรียงลำดับแตกต่างกัน

ข. เกณฑ์การตัดสินความลำเอียง

การจะทราบว่าข้อทดสอบใดมีความลำเอียงหรือไม่ ตัดสินจากระยะตั้งฉากของคู่อันดับเดลต้า (Δ) ไปยังเส้นแกนหลัก (d) ว่ามากกว่าเกณฑ์ที่กำหนดไว้หรือไม่ ในทางปฏิบัติยังไม่มีเกณฑ์ที่แน่นอน นักทดสอบบางท่านใช้เกณฑ์ $d = \pm 2S_d$ หรือ $d = \pm 1.964S_d$ สำหรับความเชื่อมั่น (reliability) ร้อยละ 95 และ $d = \pm 3S_d$ หรือ $d = \pm 2.64S_d$ สำหรับความเชื่อมั่นร้อยละ 99 รวมทั้งบางท่านใช้ $d = \pm 0.75$ z-score unit หรือ $d = \pm 1.5$ z-score unit เป็นต้น (Augoff, 1982 : Osterlind, 1963 : 35)

ค. ขั้นตอนการวิเคราะห์

การวิเคราะห์หาความลำเอียงวิธีนี้เป็นวิธีที่ง่ายที่สุด เสียเวลาน้อยที่สุด แต่เป็นวิธีที่มีลักษณะอนุรักษ์มากที่สุด (most conservative method) การวิเคราะห์ทำได้ดังนี้

1. แบ่งกระดาษคำตอบแบบทดสอบ (ปรนัย) ออกเป็น 2 กลุ่ม ตามที่ผู้ศึกษาต้องการศึกษาความลำเอียง เช่น ถ้าต้องการศึกษาความลำเอียงของข้อทดสอบต่อเพศ ก็แบ่งกระดาษคำตอบของผู้สอบออกเป็น 2 กลุ่มคือ กระดาษคำตอบของผู้สอบเพศชายและเพศหญิง

2. นำคำตอบของผู้สอบแต่ละกลุ่มไปวิเคราะห์ข้อทดสอบตามรูปแบบประเพณีนิยม (classical model) เพื่อหาค่าความยากง่ายของข้อทดสอบ คือค่า p แล้วจึงแปลงค่า p เป็นค่าเดลต้า (Δ) ซึ่งเป็นค่าความยากง่ายอีกชนิดหนึ่ง โดยอาจใช้โปรแกรมคอมพิวเตอร์ชื่อ ITEMX ถ้าใช้เครื่องคอมพิวเตอร์ขนาดใหญ่ หรือใช้โปรแกรม CTIA ถ้าใช้เครื่องไมโครคอมพิวเตอร์ (ผู้สนใจอาจศึกษาจาก สุพัฒน์ สุกมลสันต์, 2533; 2534 a)

3. บันทึกค่า Δ ของข้อทดสอบแต่ละข้อและของแต่ละกลุ่มลงในแผ่น diskette เพื่อใช้กับเครื่องไมโครคอมพิวเตอร์

4. ใช้โปรแกรมคอมพิวเตอร์ชื่อ BIAS ตามที่ผู้เขียนได้เขียนขึ้น ทำการคำนวณหาระยะห่างของคู่อันดับเดลต้าไปยังแกนหลัก (d) โดยอาศัยสูตรดังนี้ (Osterlind, 1983 : 34) คือ

$$d_i = \frac{bx_i + a - Y_i}{\sqrt{b^2+1}}$$

$$a = M_x - bMy$$

$$b = \frac{(\sigma_y^2 - \sigma_x^2) \pm \sqrt{(\sigma_y^2 - \sigma_x^2)^2 + 4r_{xy}^2\sigma_x^2\sigma_y^2}}{2r_{xy}\sigma_x\sigma_y}$$

ในเมื่อ d_i = ระยะห่างตั้งฉากจากคู่อันดับเดลต้าของข้อทดสอบข้อที่ i ไปยังแกนหลัก

x, y = ค่าเดลต้าของข้อทดสอบข้อที่ i จากกลุ่มที่ 1 และ 2

a = ค่าที่ตัดแกน y

b = ค่าความชันของ slope

M_x, M_y = ค่าเฉลี่ยของค่าเดลต้าของกลุ่มที่ 1 และ 2

σ_X, σ_Y = ค่าเบี่ยงเบนมาตรฐานของค่าเดลต้าของกลุ่มที่ 1 และ 2

r_{xy} = ค่าสัมประสิทธิ์สหสัมพันธ์ของค่าเดลต้าของกลุ่มที่ 1 และ 2

5. คำนวณหาค่าเบี่ยงเบนมาตรฐานของระยะห่างตั้งฉาก จากสูตร

$$S_d = \sqrt{1 - r_{xy}}$$

6. กำหนดเกณฑ์การตัดสินความลำเอียง เช่น ในกรณีความเชื่อมั่นในการตัดสินร้อยละ 95 อาจใช้เกณฑ์ = $d = 1.96\ S_d$ เป็นต้น

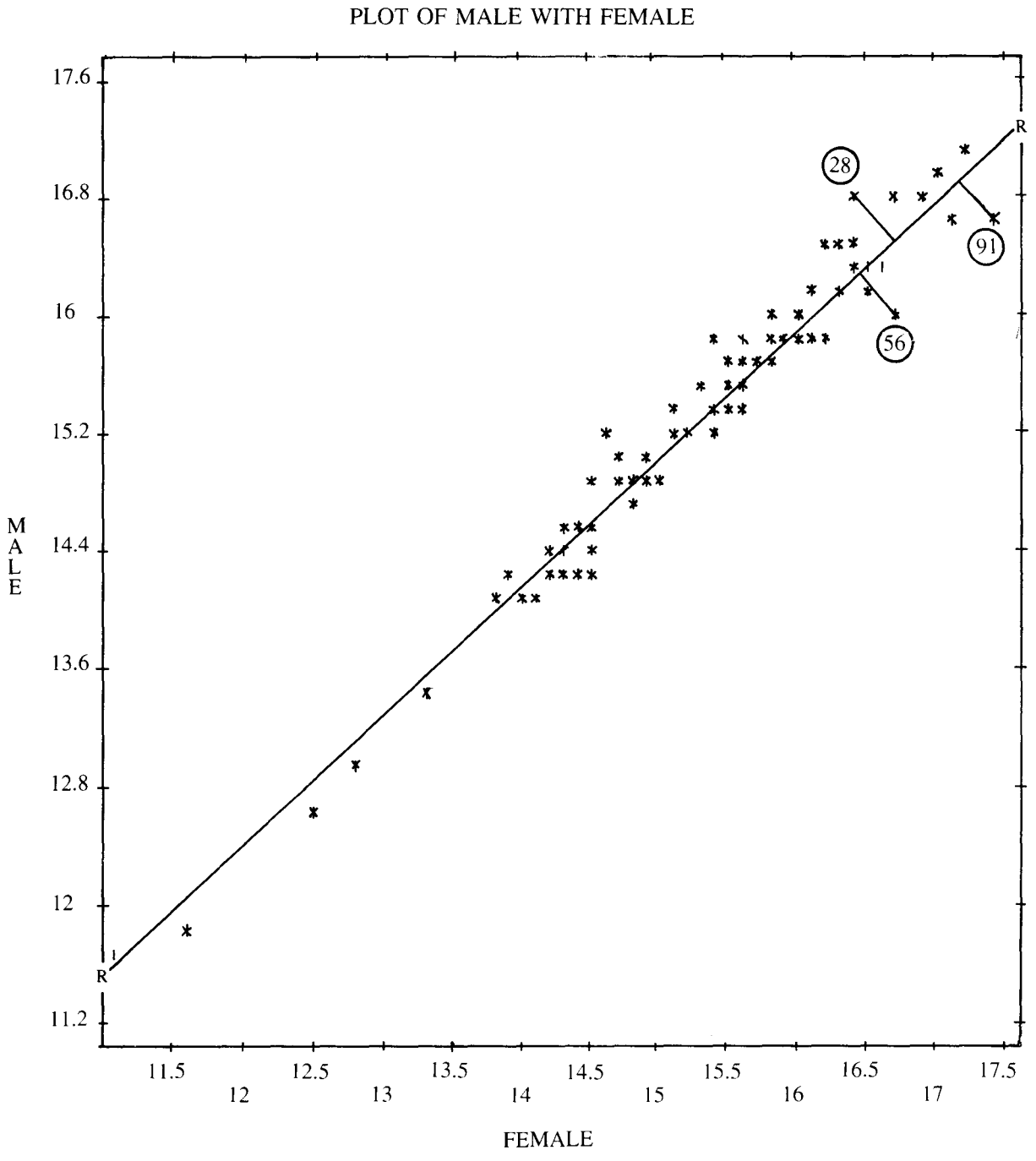
7. เปรียบเทียบค่า d ที่คำนวณได้จากข้อ 4 กับเกณฑ์ที่กำหนดขึ้นในข้อ 6 ถ้าข้อทดสอบข้อใดมีค่า d > เกณฑ์ แสดงว่าเป็นข้อที่มีความลำเอียง ข้อทดสอบข้อใดที่ลำเอียงและมีค่า d เป็นค่า + แสดงว่า ข้อทดสอบนั้นลำเอียงต่อผู้สอบในกลุ่มแนวตั้ง (ผู้สอบในกลุ่มแนวตั้งเสียเปรียบ) แต่ถ้าค่า d เป็น - แสดงว่าข้อทดสอบนั้นลำเอียงต่อผู้สอบในกลุ่มแนวนอน (ผู้สอบในกลุ่มแนวนอนเสียเปรียบ)

8. ผู้ที่ทำการศึกษาอาจใช้โปรแกรมคอมพิวเตอร์ ชื่อ SPSS หรือ SPSS/PC ทำการ Plot กราฟ ค่าเดลต้าของแบบทดสอบทั้งฉบับ เพื่อศึกษาการกระจายของค่าเดลต้าว่ามีลักษณะอย่างไร และจะพบว่าข้อที่ลำเอียงคือข้อที่อยู่ห่างจากแกนหลักกว่าข้ออื่นๆ (อาจจะ plot ค่าเดลต้าด้วยมือเองก็ได้)

ง. ตัวอย่างผลการวิเคราะห์

ความลำเอียงของข้อทดสอบภาษาอังกฤษเข้ามหาวิทยาลัยชุด กข. ปี พ.ศ. 2531 ซึ่งวิเคราะห์จากข้อมูลจริง มีลักษณะดังนี้ (สุพัฒน์ สุกมลสันต์, 2536 ; ภาคผนวก ง)

ตัวอย่างที่ 1 : ลักษณะของกราฟและระยะห่างตั้งฉากจากแกนหลัก (d_1)
กรณีของแบบทดสอบภาษาอังกฤษชุด กข. ปี 2531



2. วิธีทดสอบความแตกต่างของสัดส่วนระหว่างกลุ่มด้วยไคสแควร์ (Chi-square Method)

ก. แนวคิด

การวิเคราะห์วิธีนี้มีความเชื่อมั่น ความลำเอียงของข้อทดสอบ หมายถึง ข้อทดสอบซึ่งผู้สอบที่ได้คะแนนในช่วงระดับคะแนนเดียวกัน (ผู้สอบมีความสามารถในระดับเดียวกัน) มีสัดส่วนของการตอบข้อทดสอบถูกหรือผิดแตกต่างกันเนื่องมาจากกลุ่มที่แตกต่างกัน (Shepard, 1982 : 23)

ข. เกณฑ์การตัดสินความลำเอียง

ใช้ระดับความมีนัยสำคัญของความแตกต่างกันของสัดส่วนของการตอบถูกหรือผิดข้อทดสอบแต่ละข้อซึ่งอาจเป็น α = 0.05 หรือ 0.01 แล้วแต่ความต้องการของผู้ต้องการศึกษา

อนึ่ง เกณฑ์การตัดสินดังกล่าวอาจมีความเข้มงวดแตกต่างกันตามจำนวนกลุ่มระดับความสามารถของผู้สอบแต่ละกลุ่มที่ผู้ศึกษาต้องการหาความลำเอียงของข้อทดสอบ กล่าวคือหากแบ่งผู้สอบเพศชายออกเป็น 3 กลุ่มตามระดับความสามารถ เกณฑ์การตัดสินย่อมมีความเข้มงวดน้อยกว่าการแบ่งผู้สอบนั้นออกเป็น 5 กลุ่ม ทั้งนี้เพราะระดับความมีนัยสำคัญรวมของ χ^2 - test เมื่อปรับค่าความคลาดเคลื่อนชนิดที่ 1 แล้วคือ

$$\alpha = \frac{1}{j}\alpha_t$$

ในเมื่อ α_t = ระดับนัยสำคัญรวม

α = ระดับนัยสำคัญที่ปรับค่า type I error แล้ว

ดังนั้น หากว่าเราแบ่งผู้สอบแต่ละกลุ่มออกเป็น 3 กลุ่มย่อยตามระดับความสามารถ เกณฑ์การตัดสินความลำเอียงก็คือ $\chi^2_{1,0.017}$ = 5.7242

ค. ขั้นตอนการวิเคราะห์

การวิเคราะห์หาความลำเอียงวิธีนี้ ยุ่งยากกว่าวิธีที่ 1 เพราะมีการคำนวณหลายขั้นตอนมากกว่า แต่เป็นวิธีที่มีความไว (sensitive) มากกว่าวิธีที่ 1 และที่สำคัญมากก็คือ เป็นวิธีที่มีเกณฑ์ในการตัดสินแน่นอนมากกว่า การวิเคราะห์ทำได้ดังนี้

1. แบ่งกระดาษคำตอบแบบทดสอบออกเป็น 2 กลุ่ม ในลักษณะเดียวกับข้อที่ 1 ของวิธีที่ 1
2. นำกระดาษคำตอบไปตรวจให้คะแนน แล้วพิจารณาแบ่งผู้สอบออกเป็นกลุ่มย่อย โดยอาศัยคะแนนรวมเป็นเกณฑ์ การจะแบ่งผู้สอบออกเป็นกี่กลุ่มย่อยอาจพิจารณาเช่นเดียวกับการกำหนดระดับคะแนน ("ตัดเกรด") ข้อควรคำนึงก็คือ แต่ละกลุ่มย่อยควรมีผู้ตอบข้อทดสอบแต่ละข้อไม่น้อยกว่า 5 คน (ซึ่งเป็นข้อตกลงเบื้องต้นอย่างหนึ่งของ X^2 - test) หรือไม่ก็ต้องใช้วิธีการปรับแก้ค่า χ^2
3. ใช้เกณฑ์เดียวกันกับข้อ 2 แบ่งผู้สอบอีกกลุ่มหนึ่งออกเป็นกลุ่มย่อย ซึ่งจะได้จำนวนกลุ่มย่อยเท่ากัน
4. บันทึกผลการตอบข้อทดสอบแต่ละข้อของแต่ละกลุ่มลงแผ่น diskette หรือ computer tape
5. นำข้อมูลที่บันทึกแล้วไปวิเคราะห์ โดยอาจใช้โปรแกรมย่อย ชื่อ Crosstabs ของโปรแกรม SPSS หรือ SPSS/PC การวิเคราะห์ให้ใช้การหาค่า χ^2 ระหว่างกลุ่ม (1, 2) × กลุ่มย่อย i (ถูก, ผิด) ตามจำนวนครั้งของกลุ่มย่อย
6. เปรียบเทียบค่า χ^2 ที่คำนวณได้จากข้อ 5 กับเกณฑ์ที่กำหนดขึ้น เช่นในกรณีที่แบ่งผู้สอบแต่ละกลุ่มเป็น 3 กลุ่มย่อย เกณฑ์ก็คือ ค่า 5.7242 หากค่า χ^2 ของกลุ่มใด > เกณฑ์ แสดงว่าข้อทดสอบข้อนั้น มีความลำเอียงต่อผู้สอบกลุ่มนั้น

ง. ตัวอย่างผลการวิเคราะห์

การวิเคราะห์ข้อทดสอบเพื่อหาความลำเอียงโดย Chi-square Method อาจทำได้หลายวิธี แต่วิธีที่ได้กล่าวมาแล้วเรียกว่า Marascuilo Method One ซึ่งผลการวิเคราะห์หาความลำเอียงของข้อทดสอบภาษาอังกฤษเข้ามหาวิทยาลัย ชุด กข. ปี พ.ศ. 2531 ข้อ 4 จากข้อมูลจริงมีลักษณะดังนี้ (สุพัฒน์ สุกมลสันต์ , 2534b : ภาคผนวก ง)

ตัวอย่างที่ 2 การคำนวณหาความลำเอียงของข้อทดสอบ โดยวิธี Chi-square Method (ข้อที่ 4 ชุด กข. ปี 2531)

กลุ่มคะแนนต่ำ

	ชาย	หญิง	รวม
ผิด	705 5401 52.4	597 45.9 47.4	1302 50.0
ถูก	641 49.2 47.6	662 50.8 52.6	1303 50.0
รวม	1346 51.7	1259 48.2	26.5 100.0

χ^2 = 6.2017

กลุ่มคะแนนปานกลาง

	ชาย	หญิง	รวม
ผิด	601 49.6 56.1	611 50.4 55.3	1212 55.7
ถูก	470 48.8 43.9	494 51.2 44.7	964 44.3
รวม	1071 49.2	1105 50.8	2176 100.0

χ^2 = 0.1173

กลุ่มคะแนนสูง

	ชาย	หญิง	รวม
ผิด	228 46.0 39.1	268 54.0 42.1	496 40.7
ถูก	355 49.1 60.9	368 50.9 57.9	723 59.3
รวม	583 47.8	636 52.2	1219 100.0

χ^2 = 1.0352

การตัดสิน

α = 0.05

เกณฑ์การตัดสิน = $\frac{\alpha}{k} = \frac{0.05}{3}$ 0.0167

และ $\chi^2_{1,.00167}$ = 5.7246

ดังนั้น ข้อทดสอบข้อที่ 4 มีความลำเอียงต่อผู้สมัครสอบเพศชาย (เนื่องจากอัตราส่วนของผู้ที่ตอบถูกเป็นผู้สมัครสอบเพศชายน้อยกว่าเพศหญิง ทั้ง ๆ ที่เป็นผู้ที่มีคะแนนรวมอยู่ในกลุ่มคะแนนรวมเดียวกัน)

3. วิธีการตอบสนองของข้อทดสอบชนิด 3 พารามิเตอร์ (Three-parameter Logistic Method)

ก. แนวคิด

การวิเคราะห์วิธีนี้มีความเชื่อว่าความลำเอียงของข้อทดสอบหมายถึงข้อทดสอบซึ่งผู้สอบที่มีระดับความสามารถเท่ากันมีโอกาสตอบข้อทดสอบได้ถูกต้องแตกต่างกัน เนื่องจากมาจากกลุ่มที่แตกต่างกัน

ข. เกณฑ์การตัดสินความลำเอียง

เกณฑ์การตัดสินของการวิเคราะห์วิธีนี้ยังไม่แน่นอน แต่จากงานวิจัยที่ค้นพบจะใช้ค่าดัชนีแสดงพื้นที่ระหว่างโค้งลักษณะของข้อทดสอบ (Item Characteristic Curve : ICC) ของผู้ตอบจากกลุ่มที่แตกต่างกันตามเกณฑ์ดังนี้ (Rudner. 1977 อ้างจาก ทัศนีย์ พีรมนตรี, 2530 : 9) คือ

1. $\phi < 0.40$ แสดงว่าข้อทดสอบมีความลำเอียงน้อย
2. $0.40 < \phi < 0.70$ แสดงว่าข้อทดสอบมีความลำเอียงปานกลาง
3. $\phi > 0.70$ แสดงว่าข้อทดสอบมีความลำเอียงมาก

แต่อย่างไรก็ตาม ยังไม่ปรากฏว่ามีงานวิจัยใด ที่กล่าวว่าความลำเอียงระดับใดจึงถือว่ามีนัยสำคัญ

ค. ขั้นตอนการวิเคราะห์

การวิเคราะห์หาความลำเอียงวิธีนี้เป็นวิธีที่ซับซ้อนที่สุด เสียเวลามากที่สุด เสียค่าใช้จ่ายมากที่สุด และผู้วิเคราะห์ต้องอาศัยพื้นความรู้ทางสถิติชั้นสูง แต่ผลการวิเคราะห์มีความละเอียดและถูกต้องตามหลักการวัดผลมากที่สุด จึงถือกันว่าเป็นวิธีที่มีความไวมากที่สุด (most sensitive mothod) การวิเคราะห์ทำได้ดังนี้

1. นำผลการสอบเป็นรายข้อของแบบทดสอบทั้งฉบับไปทดสอบก่อนว่าแบบทดสอบนี้วัดความสามารถ เพียงมิติเดียว (unidimensionality) หรือไม่ ซึ่งอาจทำได้หลายวิธี เช่น วิเคราะห์องค์ประกอบ (Factor Analysis) ของแบบทดสอบด้วยวิธี Principal Component Analysis แล้วหมุนแกนด้วยวิธี Varimax หากว่า eigenvalue ขององค์ประกอบที่ 1 มีขนาดแตกต่างจาก eigenvalue ขององค์ประกอบที่ 2 มากและ eigenvalue ขององค์ประกอบที่ 2 แตกต่างจาก eigenvalue ขององค์ประกอบอื่น ๆ ที่เหลือเล็กน้อย แสดงว่าแบบทดสอบนั้นวัดเพียงมิติเดียว (Lord, 1980 : 21) หรือคำนวณหาค่าสัมประสิทธิ์สัมพันธ์ระหว่างค่า biserial correlation ของข้อทดสอบแต่ละข้อกับค่า factor loading หากค่า r_{xy} ดังกล่าว > 0.80 ก็แสดงว่าแบบทดสอบนั้นวัดเพียงมิติเดียว (Warm, 1978 : 104)

อนึ่ง การวิเคราะห์องค์ประกอบและการหาค่า factor loading รายข้อ อาจใช้โปรแกมคอมพิวเตอร์ชื่อ SPSS หรือ SPSS/PC โดยใช้โปรแกรมย่อยชื่อ FACTOR ส่วนการหาค่า biserial correlation รายข้อนั้น อาจใช้ โปรแกรม ITEMX ดังกล่าวมาแล้ว

สำหรับการทดสอบความมิติเดียวของแบบทดสอบนี้ เป็นข้อกำหนดเบื้องต้นของการวิเคราะห์ข้อทดสอบตามทฤษฎีการตอบสนองของข้อทดสอบ (Item Response Theory) หรือที่รู้จักกันทั่วไป คือ Logistic Model หากปรากฏว่าแบบทดสอบไม่ได้วัดเพียงมิติเดียวแล้ว จะวิเคราะห์ข้อทดสอบด้วยวิธีการดังกล่าวไม่ได้เว้นแต่จะได้รับการแก้ไขก่อน ซึ่งอาจทำได้หลายวิธี (ผู้สนใจอาจศึกษาได้จาก Warm, 1978)

2. แบ่งกระดาษคำตอบแบบทดสอบออกเป็น 2 กลุ่มในลักษณะเดียวกับข้อที่ 1 ของวิธีที่ 1 (หลังจากทราบว่าแบบทดสอบวัดเพียงมิติเดียวแล้ว)

3. นำกระดาษคำตอบของแต่ละกลุ่มไปวิเคราะห์ข้อทดสอบรายข้อ โดยใช้โปรแกรมคอมพิวเตอร์ชื่อ LOGIST การวิเคราะห์ขั้นนี้จะต้องให้เครื่องคอมพิวเตอร์วิเคราะห์อย่างน้อยจนถึงระยะที่ 25 (stage 25) เพื่อให้ได้ค่า พารามิเตอร์ที่คงที่ทั้ง 3 ค่า คือ

a = ค่าอำนาจจำแนก
b = ค่าความยากง่าย และ
c = ค่าการเดาถูก

ในขณะที่ทำการวิเคราะห์ข้อทดสอบนี้ หากข้อทดสอบใด "ไม่ดี" ผู้วิเคราะห์ควรตัดข้อทดสอบนี้ออกจากการวิจัย และทำการวิเคราะห์ใหม่จนกว่าการวิเคราะห์จะได้พารามิเตอร์ที่คงที่ ในกรณีที่ไม่สามารถทำให้ค่าดังกล่าวคงที่ได้ควรต้องขยายระยะการวิเคราะห์ให้มากกว่า 25 ระยะ (ผู้สนใจอาจศึกษาวิธีการวิเคราะห์ข้อทดสอบด้วย LOGIST ได้จาก สุพัฒน์ สุกมลสันต์, 2533)

4. ใช้ค่าพารามิเตอร์ของผู้สอบกลุ่มใดกลุ่มหนึ่งเป็นหลักแล้วปรับเทียบ (equate) ค่าพารามิเตอร์ของอีกกลุ่มหนึ่ง เพื่อปรับให้ค่าเหล่านี้อยู่ในมาตรวัด (scale) เดียวกัน ดังนี้ (Osterlind, 1983 : 49) :

a^*_{12} = $(1/A)\, a_{12}$

b^*_{12} = $Ab_{12} + B$

A = $\sigma^2_{bi1} / \sigma^2_{bi2}$

B = $b_{.1} - Ab_{.2}$

ในเมื่อ A = ค่าอำนาจจำแนก (a) ของข้อทดสอบก่อนปรับเทียบ

B = ค่าความยากง่าย (b) ของข้อทดสอบก่อนปรับเทียบ

δ_{bi1} = ค่าเบี่ยงเบนมาตรฐานของค่า b ของกลุ่มหลัก

δ_{bi2} = ค่าเบี่ยงเบนมาตรฐานของค่า b ของกลุ่มปรับเทียบ

$b_{.1}$ = ค่าเฉลี่ยของค่า b ของกลุ่มหลัก

$b_{.2}$ = ค่าเฉลี่ยของค่า b ของกลุ่มที่ปรับเทียบ

a^*, b^* = ค่า a และ b ที่ปรับเทียบแล้ว

5. คำนวณหาพื้นที่ของโค้งลักษณะของข้อทดสอบของผู้สอบแต่ละกลุ่มเมื่อผู้สอบมีความสามารถที่แท้จริง(true ability) ระหว่าง $\theta = \pm 3$ หรือ ± 5 จากสูตรต่อไปนี้ (Hambleton and Cook, 1977 : 82) :

$$\int_{-5}^{5} P(\theta) = c_i + (1 - c_I) \frac{e^{Da(\theta - b_i)}}{1 + e^{Da(\theta - b_i)}}$$

ในเมื่อ D = ค่า scale factor มีค่า 1.7

6. หาความแตกต่างระหว่างพื้นที่ของโค้งลักษณะของข้อทดสอบ (ICC) แต่ละข้อของผู้สอบแต่ละกลุ่ม คือ

$$\phi_i = \int_{-5}^{5} P_1(\theta)\ \Delta\theta - \int_{-5}^{5} P_2(\theta)\ \Delta\theta$$

7. เปรียบเทียบค่า ϕ ของข้อทดสอบแต่ละข้อ กับเกณฑ์ที่กำหนดไว้ ถ้าหากว่าค่า $\phi >$ เกณฑ์ แสดงว่าข้อทดสอบ นั้นมีความลำเอียง

8. นำ ICC ของข้อทดสอบที่ลำเอียงมา plot ในภาพเดียวกัน หากว่าเส้นกราฟของกลุ่มใดอยู่เหนือกว่าเส้นกราฟของอีกกลุ่มหนึ่ง แสดงว่าข้อทดสอบนั้นมีความลำเอียงต่อกลุ่มที่เส้นกราฟอยู่ใต้กว่า นอกจากนี้เส้นโค้งลักษณะของข้อทดสอบ ข้อหนึ่งอาจตัดกัน 1-2 ครั้ง ดังนั้นการตีความลำเอียงต้องสังเกตดู ICC แต่ละจุดก่อนและหลังตัดกันด้วย เช่น ข้อทดสอบอาจลำเอียงต่อผู้สอบกลุ่มที่ 1 ในบางความสามารถแต่ลำเอียงต่อกลุ่มที่ 2 ในบางความสามารถก็ได้

อนึ่ง ท่านที่สนใจอาจใช้โปรแกรมคอมพิวเตอร์ที่ผู้เขียนได้เขียนขึ้นชื่อ ICC.TEST ทำการวิเคราะห์ในขั้นที่ 4 -6 ได้ ส่วนในขั้นที่ 8 นั้นอาจใช้โปรแกรมคอมพิวเตอร์สำหรับรูปชื่อ Math Cad ช่วยในการคำนวนหาพื้นที่ และ plot กราฟได้

8. ตัวอย่างผลการวิเคราะห์

ความลำเอียงของข้อทดสอบภาษาอังกฤษเข้ามหาวิทยาลัย ชุด กข. ปี พ.ศ. 2531 ข้อ 41 ซึ่งวิเคราะห์จากข้อมูลจริง มีลักษณะดังนี้ (สุพัฒน์ สุกมลสันต์, 2534b : ภาค ผนวก ง)

ตัวอย่างที่ 3 : ผลการวิเคราะห์หาความลำเอียงด้วย Three-parameter Logistic Method

a : = 0.5413

b : = -.1482 **2531-ENGLISH EXAM FORM AB, ITEM NO. 41**

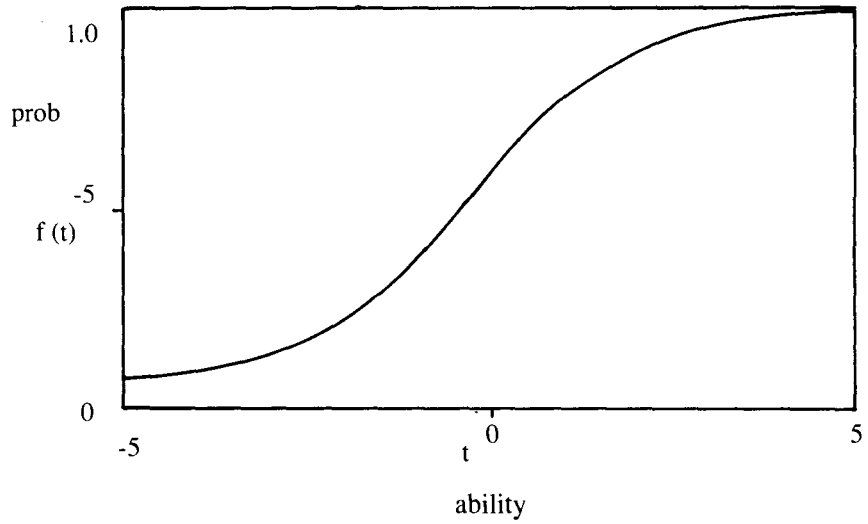
c : = 0.0767 **ITEM BIASES AGAINST SEXES**

t : = -5, -4.95 . . 5

d : = 1.7

f(t) : = $c + (1 - c) \frac{e \times p(d.a.(t-b))}{1 + e \times p(d.a.(t-b))}$ $\int_{-5}^{5} f(t)\ dt$ = 5.518

MALE



transformed

parameters

x : = 0.8412

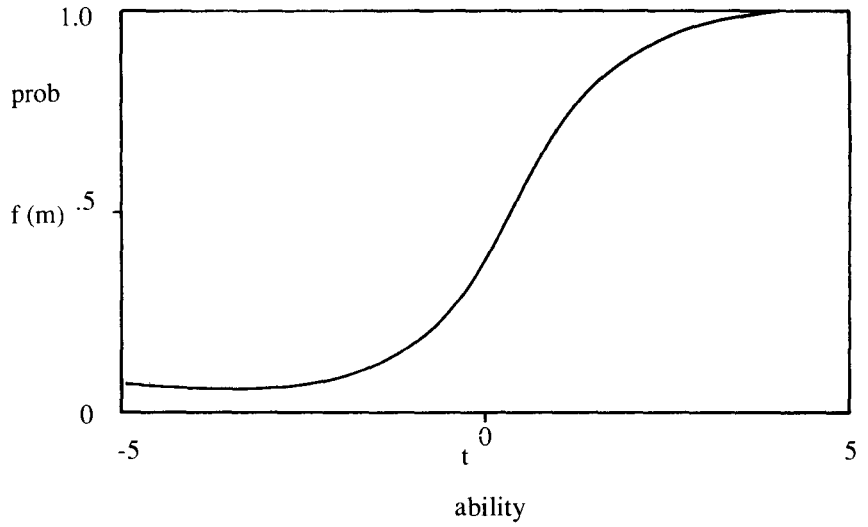
y : = 0.6094

z : = 0.0758

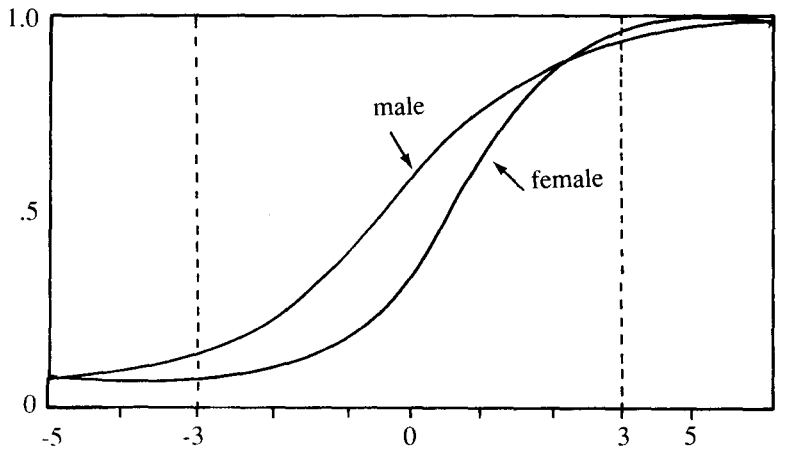
m : = -5, -4.95 ..5

$$f(m) := z + (1 - z)\,\frac{\exp(d \cdot x \cdot (m - y))}{1 + \exp(d \cdot x \cdot (m - y))} \qquad \int_{-5}^{5} f(m)\,dm = 4.817$$

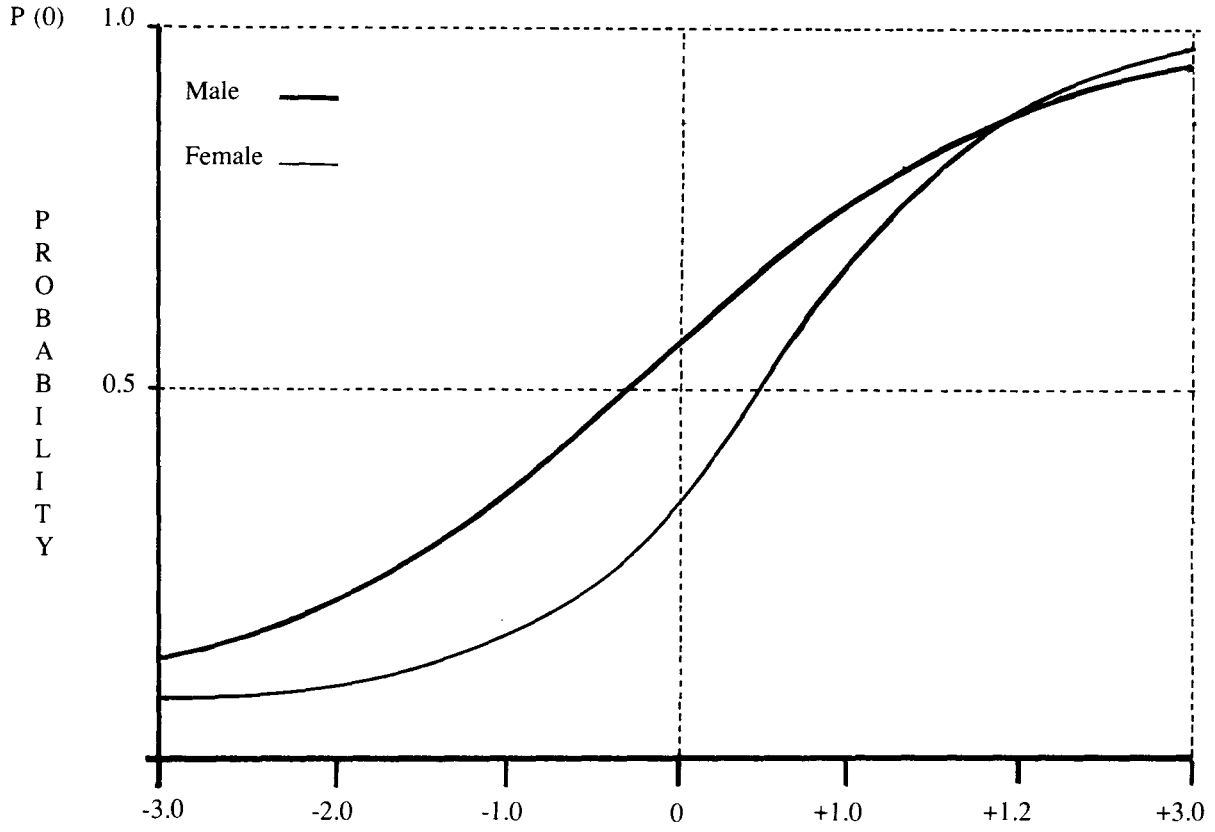
FEMALE



ภาพซ้อน



ภาพขยายระหว่าง $\theta = \pm 3$



DETAIL OF BIAS

I_NO	FROM ABILIT	TO ABILITY	AGAINST	INDEX
1	-3.0000000000E+00	1.9799999994E+00	FEMALE	6.8347039691E-01
2	1.9799999994E+00	2.9999999993E+00	MALE	1.4187977449E-02
			TOTAL INDEX	6.9765837436E-01

ความเหมือนและความแตกต่างของการวิเคราะห์หาความลำเอียง 3 วิธี

การวิเคราะห์หาความลำเอียงของข้อทดสอบ 3 วิธีดังกล่าวแล้ว เนื่องจากใช้แนวคิดและเกณฑ์ตัดสินความลำเอียงต่างกันจึงคงทำให้ได้ผลการวิเคราะห์ต่างกัน เช่นในการวิเคราะห์หาความลำเอียงของข้อทดสอบเข้ามหาวิทยาลัย วิชาภาษาอังกฤษ ปี พ.ศ. 2531-2535 ได้ผลแตกต่างกัน แต่บางวิธีก็ได้ผลเหมือนกันดังต่อไปนี้ (สุพัฒน์ สุกมลสันต์, 2534b : 79)

ตารางที่ 1 เปรียบเทียบจำนวนข้อทดสอบที่มีความลำเอียงต่อเพศของผู้สอบเมื่อวิเคราะห์ด้วย 3 วิธี

ปี พ.ศ.	แบบทดสอบชุด	n ที่ลำเอียงต่อเพศ			ข้อที่ลำเอียงต่อเพศเกิน 1 วิธี	
		DPM	CSM	TPLM	n	ข้อที่
2531	กข	3	16	23	3	7, 13, 92
2532	กข	13	28	54	22	16, 19, 24, 28, 34, 39, 40, 43, 49, 51, 55-57, 63, 71, 76, 77, 81, 87, 96-98,
2533	กข	4	9	8	3	59, 69, 100
$\overline{X}$		6.67	17.67	28.33	9.33	14.86%
2531	กขค	6	22	8	7	16, 25, 31, 76, 90, 91, 96
2532	กขค	4	17	61	15	4, 7, 8, 10, 11, 13, 19, 28, 30, 34, 55, 65, 67, 78, 84
2533	กขค	2	14	55	9	7, 12, 16, 20, 23, 30, 36, 48, 98
$\overline{X}$		4.00	17.67	41.33	10.33	16.80%
$\overline{\overline{X}}$		5.33	17.67	34.83	9.83	15.83%
F-test		**	**	**		
r_{xy}		-	0.84*	0.18*		
			-	0.29*		

* P > 0.05

** P < 0.05

ผลการวิจัยที่เกี่ยวข้อง

Ironson และ Subkoviak (1979) ได้ทำการศึกษาเปรียบเทียบผลของการวิเคราะห์ความลำเอียงของแบบทดสอบหลายวิธีและพบว่า Three-parameter Logistic Method เป็นวิธีที่ดีที่สุด รองลงมา ได้แก่ Chi-square Method และ Delta-Plot Method ซึ่งในปีเดียวกัน Laksana (1979) ก็ศึกษาพบว่า Three-parameter Logistic Method เป็นวิธีวิเคราะห์ความลำเอียงที่ดีกว่า ANOVA Method แต่ว่าไม่เหมาะสมในทางปฏิบัติ เนื่องจากต้องใช้กลุ่มขนาดใหญ่

ในอีก 5 ปีต่อมา Shepard และคณะ (อ้างจาก Subkoviak, และคณะ : 1984 : 50) ได้ทำการวิเคราะห์ความลำเอียงของข้อทดสอบหลายวิธีและพบว่า Three-parameter Logistic Method เป็นวิธีที่ดีที่สุด รองลงมาได้แก่ Chi-square Method และ Delta-Plot Method เป็นวิธีที่ใช้ได้ดีในทางปฏิบัติ เพราะคำนวณง่าย ซึ่งจะสอดคล้องกับผลการศึกษาของ Subkoviak และคณะ (1984) Rudner และคณะ (1980) และสุพัฒน์ สุกมลสันต์ (2534b)

สรุป

ดังนั้น ในทางปฏิบัติ หากเป็นไปได้ผู้สนใจควรทำการวิเคราะห์ความลำเอียงด้วย Three-parameter Logistic Method เพราะว่าแม้จะเสียเวลามาก เสียค่าใช้จ่ายสูง ใช้กลุ่มตัวอย่างขนาดใหญ่ และต้องอาศัยโปรแกรมคอมพิวเตอร์ แต่ก็จะได้ข้อ

สนเทศ (information) มากกว่าวิธีอื่น หากวิธีดังกล่าวไม่อาจทำได้ ก็ควรเลือกใช้ Chi-square Method โดยเฉพาะวิธี Marascuilo Method One ที่แบ่งผู้สอบแต่ละกลุ่มออกเป็น 5 กลุ่มย่อย ตามระดับความสามารถ (Burrill, 1982 : 162) และหากยังไม่สามารถทำการวิเคราะห์หาความลำเอียงของข้อทดสอบได้ด้วยวิธีที่ 2 ดังกล่าวแล้ว ก็ควรใช้ Detlta-Plot Method เพราะเป็นวิธีที่ง่ายที่สุด แม้ว่าจะเป็นวิธีที่มีลักษณะอนุรักษ์มากที่สุด (most conservative method) ก็ตาม แต่ก็ดีกว่าที่ผู้สร้างแบบทดสอบจะไม่ให้ความสนใจทำการวิเคราะห์หาความลำเอียงของข้อทดสอบที่สร้างขึ้นเลย เพราะผู้สร้างแบบทดสอบอาจทำให้เกิด "บาปบริสุทธ์" แก่ผู้รับการทดสอบได้

เอกสารอ้างอิง

ทัศนีย์ พีรมนตรี. "การวิเคราะห์หาความลำเอียงของแบบสอบวิชาคณิตศาสตร์ โครงการตรวจสอบคุณภาพการศึกษาชั้นมัธยมศึกษาปีที่ 6." *วิทยานิพนธ์ครุศาสตร์มหาบัณฑิต.*

สุพัฒน์ สุกมลสันต์. *คู่มือการใช้โปรแกรม CTIA/GRADING.* กรุงเทพมหานคร : สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2534a.

_____. "การวิเคราะห์ความลำเอียงของข้อทดสอบภาษาอังกฤษเข้ามหาวิทยาลัย ปี พ.ศ. 2531-2533". *รายงานวิจัย.* กรุงเทพมหานคร : สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย. 2534b.

_____. *การวิเคราะห์ข้อทดสอบแนวใหม่ด้วยคอมพิวเตอร์.* กรุงเทพมหานคร : สถาบันภาษา จุฬาลงกรณ์มหาวิทยาลัย, 2533.

Angoff, W.H. "Use of Difficulty and Discrimination Indice for Detecting Item Bias." *In Handbook of Methods for Detecting Test Bias.* R.A. Berk (ed.) Maryland : The Johns Hopkins University Press. 1982.

Berk, R.A. *Handbook of Methods for Detecting Test Bias.* Maryland : The Johns Hopkins University Press, 1982.

Burill, L.E. "Comparative Studies of Item Bias Methods". *In Handbook of Methods for Detecting Test Bias.* R.A. Berk (ed.) Maryland : The Johns Hopkins University Press, 1982.

Hambleton, R.K. and Cook, L.L. "Latent Trait Models and Their Use in the Analysis of Educational Test Data." *Journal of Educational Measurement.* 14 (Summer, 1977) : 75-95.

Ironson, G.H. and Subkoviak, M.J. "A Comparison of Several Methods of Assessing Item Bias." *Journal of Educational Measurement* 16 (1976) : 209-225.

Lord. F.M. *Applications of Item Response Theory to Practical Testing Problems.* New Jersey : Lorence Erlbaum. 1980

Osterlind, S.J. *Test Item Bias,* California : SAGE Publications, Inc., 1983.

Rudner, L.M. and Others. "A Monte Carlo Comparison of Seven Biased Item Detection Techniques." *Journal of Educational Measurement.* 17 (Spring, 1980) : 1-10.

Shepard, L.A. "Definitions of Bias." In *Handbook of Methods for Detecting Test Bias.* R.A. Berk (ed.) Maryland : The Johns Hopkins University Press, 1982.

Subkoviak, M.J. and Others. "Empirical Camparison of Selected Item Bias Detection Procedures with Bias Manipulation." *Journal of Educational Measurement.* 21 (Spring 1984) : 49-58.

Warm, T.A. *A Primer of Item Response Theory.* Oklahoma : US Coast Guard Institute. 1978.