โมเดลราช (Rash Model)

ดร.ณัชชา มหปุญญานนท์**

โมเดลราชเป็นรูปแบบของการทดสอบอย่างหนึ่งของรูปแบบลอจิสต์ ที่มักเรียกกันว่าแบบลอจิสต์ 1 พารามิเตอร์ (one parameter logistic model) แนวคิดนี้ จอร์จ ราช (Georg Rash) นักคณิตศาสตร์ชาวเดนมาร์กได้คิดขึ้นและเสนอแนวคิดนี้ในปี ค.ศ. 1960 แล้วได้ตีพิมพ์บทความในทำนองเดียวกันอีกในปี ค.ศ. 1961 และ 1966 ต่อมาในปี ค.ศ. 1967 เบนจามิน ไรต์ (Benjamin Wright) ได้นำแนวคิดนี้ไปเผยแพร่จนเป็นที่แพร่หลายทั่วไปนับแต่นั้นมา

ประวัติความเป็นมาโดยสังเขป

แนวคิดที่สำคัญของการวิเคราะห์ข้อทดสอบโมเดลราชเชื่อว่าความเป็นปรนัยของการวัดผล (objectivity of measurement) ที่ไม่สามารถจะหาได้จากแบบประเพณีนิยมทั่วๆไป คือ

1. ความเป็นอิสระจากกลุ่มตัวอย่าง (group independent) กล่าวคือ การคำนวณหาค่าพารามิเตอร์ของข้อทดสอบเป็นอิสระจากกลุ่มตัวอย่าง ค่าต่างๆของข้อทดสอบ เช่น ระดับความยากของข้อสอบ (β) จะไม่แปรเปลี่ยนไปตามคุณลักษณะกลุ่มตัวอย่าง
2. ความเป็นอิสระจากข้อทดสอบ (test independent) กล่าวคือ การคำนวณหาค่าพารามิเตอร์ของบุคคลเป็นอิสระจากข้อทดสอบ เช่น ความสามารถของบุคคล (θ) จะไม่แปรเปลี่ยนไปตามลักษณะของการทดสอบ ความสามารถของบุคคลจะคงที่ไม่ว่าจะวัดเมื่อใด เช่น ข้อทดสอบข้อหนึ่งๆที่เคยทดสอบแล้วจะมีค่าคงที่เมื่อวัดกับบุคคลเดิมไม่ว่าข้อทดสอบนั้นจะไปปรากฏอยู่ส่วนใดของข้อทดสอบ

แนวความคิดดังกล่าวนี้นับว่ามีประโยชน์มากในการทดสอบและวัดผลโดยเฉพาะการทดสอบเกี่ยวกับความสามารถทางสมอง(mental measurement) ในปี ค.ศ. 1969 ไรต์และปัญจปากีสาน (Panchapakesan) ได้ทำการทดสอบแนวคิดดังกล่าวแล้วสนับสนุนว่าเป็นไปได้จริง และในปีต่อๆมาได้มีนักทดสอบแนวคิดดังกล่าวอีกหลายครั้ง และยืนยันว่าคุณลักษณะดังกล่าวนี้เป็นไปได้จริง แม้ว่ากลุ่มตัวอย่างจะไม่ได้มาจากการสุ่มก็ตาม และมี

---

* ภาควิชาการประเมินผลและวิจัย คณะศึกษาศาสตร์ มหาวิทยาลัยทักษิณ

ความแกร่ง (robustness) มาก คือสามารถทดสอบเพื่อคำนวณหาความสามารถของบุคคล (θ) และระดับความยากของข้อทดสอบ (β) ได้ ทั้งๆที่กลุ่มตัวอย่างบุคคลและข้อทดสอบแตกต่างจากข้อตกลงเบื้องต้นของการใช้รูปแบบนี้ก็ตาม เช่น ขนาดของกลุ่มตัวอย่างมีเพียง 100 คน และไม่ต้องได้มาจากการสุ่มก็ได้ (Wright 1977: 219)

แนวคิดของการวิเคราะห์ข้อทดสอบตามโมเดลราช

ตามแนวคิดของการวิเคราะห์ข้อทดสอบตามโมเดลราชนั้นโอกาสที่ผู้สอบจะทำข้อทดสอบได้หรือไม่ขึ้นอยู่กับปัจจัย 2 อย่าง คือระดับความสามารถของตนเอง (ability parameter, θ) และระดับความยากของข้อทดสอบ (dificulty parameter, β) เช่นถ้า θ = 0.5 และ β = 0.5 โอกาสที่คนผู้นั้นจะสามารถตอบข้อทดสอบนั้นได้ถูกต้องมี 50% ถ้าหากความสามารถของบุคคล (θ) น้อยกว่าความสามารถของข้อทดสอบ (ระดับความยากหรือ β) แล้ว โอกาสที่จะตอบข้อทดสอบนั้นได้ถูกต้องย่อมน้อยกว่า 50% และในทำนองเดียวกัน ถ้าหากว่าค่า θ มากกว่าค่า β แล้ว โอกาสที่คนผู้นั้นจะตอบข้อทดสอบได้ถูกก็มีมากกว่า 50%

ดังนั้นตามแนวความคิดของโมเดลราชนั้นค่าพารามิเตอร์ที่เกี่ยวข้องมีเพียง 2 ค่าเท่านั้น ไม่มีค่าอำนาจจำแนกหรือโอกาสของการเดาเข้ามาเกี่ยวข้องด้วย เพราะเป็นข้อตกลงเบื้องต้นว่าข้อทดสอบที่นำมาใช้กับโมเดลราชนั้นจะต้องมีอำนาจจำแนกเท่าๆกัน และมีลักษณะที่จะทำให้เกิดการเดาได้น้อยที่สุด แต่ในทางปฏิบัติลักษณะทั้งสองนี้เป็นความแกร่งของรูปแบบซึ่งข้อทดสอบที่ไม่มีลักษณะดังกล่าวนี้ก็อาจใช้การวิเคราะห์ข้อทดสอบตามโมเดลราชได้ด้วย

พัฒนาการของสูตรโมเดลราช

ดังได้กล่าวมาแล้วว่าแนวคิดของโมเดลราชเกี่ยวข้องกับความสัมพันธ์ของความสามารถของบุคคล (θ) และระดับความยากของข้อทดสอบ (β) เท่านั้น ความสัมพันธ์นี้เป็นความสัมพันธ์เชิงคณิตศาสตร์ กล่าวคือ โอกาสที่บุคคล v ที่มีระดับความสามารถ θ (หรือ θ_v) จะทำข้อทดสอบ e ที่มีระดับความยากβ(หรือ β_e) ได้ถูกต้องมากน้อยเพียงใดขึ้นอยู่กับระดับความแตกต่างของ ($\theta_v - \beta_e$) นั่นคือ

โอกาสของความสำเร็จ (odd of success) = $(\theta v - \beta e)$

แต่เนื่องจาก $(\theta v - \beta e)$ นั้นมีค่าระหว่าง $\pm\infty$ แต่ว่าโอกาสของความสำเร็จมีได้ระหว่าง 1 กับ 0 เท่านั้น เพื่อให้ค่า $(\theta v - \beta e)$ เป็นค่าที่มีหน่วยเล็กลงและคงที่เหมาะสมแก่การนำมาใช้ จึงได้ค่าเลขชี้กำลัง (exponent) ของ $(\theta v - \beta e)$ แทนซึ่งจะมีค่าอยู่ระหว่าง 0 กับ $\pm\infty$ เช่น

$$\text{โอกาสของความสำเร็จ} = e^{(\theta v - \beta e)} = \exp(\theta v - \beta e) \qquad \ldots\ldots\ldots\ldots..1$$

หน่วยของการวัดเรียกว่า log-odds unit หรือลอจิตส์ (logits) (ค่า exp หรือ e = 2.71828...) เพื่อทำให้ $\exp(\theta v - \beta e)$ มีค่าเป็นมาตรอันตรภาค (interval scale) ระหว่าง 0 และ 1 สมการที่ 1 อาจเขียนได้ดังนี้

$$\text{โอกาสของความสำเร็จ} = \frac{\exp(\theta v - \beta e)}{1 + \exp(\theta v - \beta e)} \qquad \ldots\ldots\ldots\ldots..2$$

ดังนั้นสูตรที่แสดงว่าโอกาสที่บุคคล v มีความสามารถ θ จะทำข้อทดสอบข้อที่ e ที่มีระดับความยาก β ได้ถูกต้อง (ได้คะแนน = 1 หรือ X_v) คือ

$$P\{X_v = 1 | \theta v, \beta e\} = \frac{\exp(\theta v - \beta e)}{1 + \exp(\theta v - \beta e)} \qquad \ldots\ldots\ldots\ldots..3$$

และนี่คือสูตรของโมเดลราช

ตัวอย่าง เช่น ถ้า (θv - βe) = 2 log-odds units

โอกาสที่บุคคลจะทำข้อทดสอบข้อที่ e ได้ถูกต้อง คือ

$$P\{X_{ve} = 1 | \theta v, \beta e\} = \frac{e^2}{1+e^2} = \frac{2.71828^2}{1+2.71828^2}$$

$$= 0.881 \text{ หรือ } 88.1\ \%$$

ความสัมพันธ์ระหว่างโอกาสที่จะตอบข้อทดสอบได้ถูกต้องกับระดับความยากของข้อทดสอบมีการกระจายเป็นรูปตัวเอส (ogive distribution) ต่อมาเพื่อให้การกระจายของโค้งดังกล่าวเป็นรูปตัวเอสปกติ (normal ogive) จึงนิยมใช้ค่า 1.7 เข้ามาเกี่ยวข้องในการคำนวณด้วยเพื่อปรับปรุงโค้งของความสัมพันธ์ ปกติค่าที่นำมาปรับนี้ทำให้ความแตกต่างของโค้งทั้งสองลักษณะแตกต่างกันน้อยกว่า 1%

อนึ่ง ความสัมพันธ์ระหว่าง θv และ βe ดังสมการที่ 3 นั้นเรียกว่าฟังก์ชันลอจิสต์ (logistic function) และเป็นที่รู้จักกันมานานแล้วในหมู่นักวัดทางชีววิทยา (biometricians) ตั้งแต่ปี ค.ศ. 1920) แต่ จอร์จ ราช เป็นคนแรกที่มองเห็นความสำคัญทางการวัดทางจิตวิทยา (PSYCHOMETRICS) ราชเชื่อว่าไม่มีสูตรทางคณิตศาสตร์ใดๆอีกแล้วที่จะสามารถทำให้เกิดความสัมพันธ์ระหว่างค่าพารามิเตอร์ดังกล่าวที่สามารถจะคำนวณหาค่า θv และ βe อย่างเป็นอิสระจากกันได้ กล่าวคือ ถ้าต้องการจะคำนวณหาค่า θv ไม่จำเป็นต้องอาศัยค่า βe เข้ามาเกี่ยวข้องก็ได้ และในทำนองเดียวกัน ถ้าต้องการหา βe ก็ไม่จำเป็นต้องรู้ค่า θv ก็ได้ การคำนวณดังกล่าวนี้อาจใช้วิธีดำเนินการแบบเป็นไปได้มากที่สุดอย่างมีเงื่อนไข (conditional maximum lielihood estimators) ซึ่งเป็นวิธีใช้กันอย่างแพร่หลาย หรืออาจใช้วิธีการดำเนินการแบบเป็นไปได้มากที่สุดอย่างไร้เงื่อนไข (unconditional maximum likelihood estimators) ก็ได้เมื่อกลุ่มตัวอย่างขนาดเล็ก วิธีแรกเป็นการกำหนดให้ค่า θ ของผู้สอบคงที่เพราะถือว่ามีอยู่จริงและเป็นค่าพารามิเตอร์ แล้วจึงคำนวณหาค่าพารามิเตอร์อื่นๆของข้อทดสอบ แต่วิธีหลังนั้นเป็นการคำนวณหาค่าพารามิเตอร์ของผู้สอบและของข้อทดสอบจากการทดสอบทุกค่า โดยไม่มีการกำหนดค่าใดไว้ล่วงหน้า (Baker 1992: 143)

จากสูตรที่ 1 ในกรณีที่ $X_{ve} = 0$ ก็คือ

$$P\{X_{ve} = 0 \mid \theta v, \beta e\} = \frac{1}{1 + \exp(\theta v - \beta e)} \quad \text{............4}$$

และอาจเขียนเป็นสูตรทั่วๆไปได้ดังต่อไปนี้

$$P\{X_{ve} \mid \theta v, \beta e\} = \exp\left(\frac{X_{ve}(\theta v - \beta e)}{1 + \exp(\theta v - \beta e)}\right) \quad \text{..........5}$$

ดังนั้นสูตรที่ 3 จึงเป็นสูตรทั่วไปของโมเดลราชและในกรณีที่ข้อทดสอบมีทั้งหมด L ข้อ และกลุ่มตัวอย่างมี N คน คะแนนรวมของคนที่มีความสามารถ θv ก็คือ $r_v = \sum X_{ve}$ และจำนวนคนที่ตอบข้อทดสอบแต่ละข้อก็คือ $S_i = \sum X_v$ ดังนั้นถ้าหากกำหนดให้ $[(X_{ve})]$ แทนเมทริกซ์ (matrix) ทั้งหมดของผลการสอบแล้วโอกาสที่เกิดผลการสอบดังกล่าวได้ก็คือ

$$P\{[(X_{ve})] \mid (\theta v), \beta e\} = \frac{[\exp(\sum r_v \theta_v)][\exp(\sum S_e \beta_e)]}{\prod\prod[1 + \exp(\theta_v - \beta_e)]} \quad \text{.............6}$$

สมการที่ 6 มีความสำคัญมาก เพราะแสดงให้เห็นว่า

1. ในการที่จะคำนวณหาความสามารถของแต่ละข้อ (ระดับความยากหรือ β_e) นั้นอาศัยเพียงคะแนนรวมของแต่ละบุคคล (person score คือ r_v) เท่านั้นก็เพียงพอในการกำจัดค่า θ_v ออกจากสมการ ดังนั้นวิธีการคำนวณหาจึงเป็นอิสระจากกลุ่มตัวอย่าง
2. ในการคำนวณหาความสามารถของแต่ละบุคคล อาศัยเพียงคะแนนรวมของแต่ละข้อ (item score หรือ S_e) เท่านั้นก็เพียงพอในการกำจัดค่าβ_e ออกจากสมการ ดังนั้นการคำนวณหาค่าθ_v จึงเป็นอิสระจากกลุ่มข้อทดสอบ

3. การคำนวณหาค่า θ_v และ β_e เป็นอิสระจากกันและกัน วิธีการคำนวณหาค่า θ_v และ β_e โดยวิธีการกำจัดค่าใดค่าหนึ่งออกไปจากสมการใช้วิธีดำเนินการแบบเป็นไปได้มากที่สุดอย่างมีเงื่อนไขซึ่งเป็นวิธีที่ยุ่งยากสับสนในการวิเคราะห์แบบทดสอบที่ยาวเกินกว่า 25 ข้อไม่ได้ ดังนั้น ไรต์และปัญจปากีสาน (wright 1978: 17) จึงได้พัฒนาวิธีคำนวณค่า θ_v และ β_e โดยใช้วิธีดำเนินการแบบเป็นไปได้มากที่สุดอย่างไร้เงื่อนไข (UNCON) ขึ้น ซึ่งสามารถใช้ได้กับการวิเคราะห์แบบทดสอบที่มีขนาดยาวหรือสั้นได้ด้วย แม้ว่าจะมีอคติ (bias) ในการคำนวณเพียงเล็กน้อยแต่ก็แก้ไขได้ วิธีดังกล่าวนี้ในโปรแกรม BICAL (Binomial Calibration) แนะนำให้ใช้เมื่อข้อทดสอบมีขนาดสั้น (ประมาณ 25 ข้อ) และการกระจายของคะแนนมีลักษณะเบ้ แม้ว่าจะมีขนาดของกลุ่มตัวอย่างมากหรือน้อยก็ตาม

ในกรณีที่แบบทดสอบมีขนาดยาว (มากกว่า 25 ข้อขึ้นไป) การวิเคราะห์ข้อทดสอบโดยโมเดลราช นิยมใช้วิธีกำลังสองน้อยที่สุด (least square method) ในการคำนวณหาค่าพารามิเตอร์ (wright 1977: 219) แต่ว่าต่อมาไม่เป็นที่นิยมเพราะต้องอาศัยกลุ่มตัวอย่างขนาดใหญ่มากแลพให้ค่าไม่ถูกต้องนัก

ข้อตกลงเบื้องต้นในการใช้การวิเคราะห์ข้อทดสอบโมเดลราช

ข้อตกลงเบื้องต้นในการใช้การวิเคราะห์ข้อทดสอบโมเดลราชมีดังนี้ (Hambleton 1977: 78; Tinsley 1975: 326)

1. ข้อทดสอบแต่ละข้อจะต้องวัดคุณลักษณะ(trait) เดียวกัน หรือวัดเพียงมิติเดียว (unidimensional latent space) กล่าวคือ ข้อทดสอบจะต้องเป็นเอกพันธ์ ในการที่จะตัดสินว่าข้อทดสอบมีลักษณะดังกล่าวหรือไม่ อาจทำได้หลายวิธี เช่น โดยการอาศัยการวิเคราะห์องค์ประกอบ ถ้าค่าไอเกน (eigenvalue) ที่มากกว่า 1.00 มีหลายจำนวน และค่าอันดับที่ 2 ต่างจากค่าอันดับที่ 1 มาก และค่าอันดับที่ 3 ต่างจากค่าอันดับต่อไปไม่มากนัก ก็ถือได้ว่าแบบทดสอบนั้นวัดมิติเดียว (Lord 1980: 21) หรือโดยการหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายระหว่างค่าน้ำหนักองค์ประกอบรายข้อขององค์ประกอบที่ 1 ของข้อทดสอบทุกข้อกับค่า r_{bis} หากว่าค่า $r_{bis} > 0.80$ แล้วก็อาจแสดงว่าแบบทดสอบนั้นวัดเพียงมิติเดียวได้ด้วย (Warm 1978:

104) แต่ถ้าข้อทดสอบเป็นแบบวิวิธพันธ์จะต้องมีการวิเคราะห์องค์ประกอบด้วยวิธีดังกล่าวเพื่อจัดกลุ่มเอกพันธ์ก่อน

2. ข้อทดสอบจะต้องมีความเป็นอิสระจากตำแหน่ง (local independence) ใน 2 ลักษณะ คือ

2.1 มีความเป็นอิสระทางสถิติ (statistically independent) กล่าวคือข้อทดสอบแต่ละข้อเป็นอิสระไม่เกี่ยวข้องกัน แต่ละข้อวัดความสามารถไม่ซ้ำกันเลย ดังนั้นคำตอบแต่ละข้อของแต่ละคนเป็นอิสระต่อกัน แต่รวมกันแล้วจะวัดคุณลักษณะเดียวเท่านั้น

2.2 มีความเป็นอิสระจากตำแหน่ง (uncorrelated independence) กล่าวคือ ข้อทดสอบแต่ละข้อจะปรากฏอยู่ในตำแหน่งใดของแบบทดสอบก็ได้ จะไม่มีผลต่อการตอบของผู้สอบ

3. คะแนนมีเพียง 2 ค่า (dichotomous) เท่านั้น คือ เป็น 1 หรือ 0

4. ความเร็วในการตอบข้อทดสอบไม่มีผลต่อโอกาสที่จะตอบข้อสอบได้หรือไม่

5. โอกาสที่จะตอบข้อทดสอบได้ถูกต้องหรือไม่ขึ้นอยู่กับความสัมพันธ์ของความสามารถของบุคคล (θ) กับระดับความยากของข้อทดสอบ (β) เท่านั้น

ความแกร่งของการวิเคราะห์ข้อทดสอบโมเดลราช

การใช้การวิเคราะห์ข้อทดสอบโมเดลราชนั้นได้มีการทดลองใช้ในรูปแบบที่แตกต่างกันเพื่อหาความแกร่งของรูปแบบนี้ ปรากฏว่าโมเดลราชได้รับความนิยมมาก เพราะมีความแกร่งมากกล่าวคือ

1. ขนาดของกลุ่มตัวอย่างที่จะใช้โมเดลราชนั้นอาจจะมีขนาดเล็กประมาณ 100 คน เท่านั้นก็ได้ (Wright 1977: 219)

2. การกระจายของความสามารถของกลุ่มตัวอย่างหรือความยากของข้อทดสอบไม่จำเป็นต้องเป็นโค้งปกติ เพียงแต่มีแนวโน้มว่าจะเป็นโค้งปกติเท่านั้นก็ได้และกลุ่มตัวอย่างไม่จำเป็นที่จะต้องได้มาจากการสุ่มก็ได้ (Wright & Stone 1979: 20)

3. ข้อทดสอบไม่จำเป็นต้องมีค่าอำนาจจำแนกเท่ากันและไม่จำเป็นต้องคำนึงถึงค่าการเดาก็ได้ (Wright & Panchapakesan 1969: 25)

4. ข้อทดสอบที่ให้คะแนนรายข้อต่างกัน (multiple – point item) ก็ใช้กับโมเดลราชได้ (Willmott 1980: 195)

ประโยชน์ของโมเดลราชในการทดสอบและวัดผล

ปัจจุบันได้มีการนำเอาการวิเคราะห์ข้อทดสอบแบบความสามารถแฝงเช่นโมเดลราชมาใช้อย่างกว้างขวางในการวัดและทดสอบดังต่อไปนี้

1. ใช้ประโยชน์ในการวิเคราะห์ข้อทดสอบ (item analysis)

เนื่องจากการวิเคราะห์ข้อทดสอบตามแนวคิดเดิม เช่น วิธี 27% นั้นมีจุดบกพร่องอยู่มากดังได้กล่าวมาแล้ว เช่น ค่าสถิติต่างๆขึ้นอยู่กับสภาพของกลุ่มตัวอย่างที่ทดสอบ และค่าอำนาจจำแนกเป็นค่าที่ไม่ถูกต้อง ปัญหาต่างๆเหล่านี้อาจจะแก้ไขได้โดยวิเคราะห์ด้วยวิธีโมเดลราช

2. ใช้ในการสร้างธนาคารข้อสอบ (item bank)

เนื่องจากข้อทดสอบที่วิเคราะห์แล้วค่าพารามิเตอร์มีลักษณะคงที่ (Invariant) ดังนั้นข้อทดสอบเหล่านี้จึงนำมาใช้สร้างแบบทดสอบชุดใหม่ๆซึ่งต้องมีการรวมข้อทดสอบ จำนวนมากเข้าด้วยกันและมีการสุ่มข้อทดสอบมาใช้ตามเกณฑ์ที่ต้องการได้โดยง่าย

3. ใช้สร้างแบบทดสอบที่ให้คะแนนรายข้อต่างกัน (multiple-point test)

โมเดลราชใช้วิเคราะห์ข้อทดสอบที่ให้คะแนนรายข้อต่างกันได้ตามระดับขั้น (degree) ของความถูกต้องของคำตอบได้ และสามารถจะสร้างข้อทดสอบลักษณะดังกล่าวได้จากธนาคารข้อทดสอบที่วิเคราะห์ไว้แล้ว

4. ใช้ในการกำหนดเกณฑ์ของระดับผลสัมฤทธิ์ (master level) ของแบบทดสอบแบบอิงเกณฑ์

ผลของการวิเคราะห์ข้อทดสอบ คะแนนของผู้สอบจะถูกแปลงเป็นคะแนนความสามารถของแต่ละบุคคล (ability score) จึงสามารถเปรียบเทียบกับคะแนนความสามารถทั่วไป ซึ่งเป็นค่าพารามิเตอร์และใช้เป็นเกณฑ์คงที่ได้ ทำให้เราทราบได้ว่าระดับสัมฤทธิ์ต่ำสุด (minimum mastery level) ของแบบทดสอบแต่ละชุดควรเป็นเท่าใด รวมทั้งใช้ในการกำหนดแถบความสามารถ (ability band) ของระดับความสามารถหลายๆกลุ่มจากผลการทดสอบได้

ด้วย โดยอาศัยการจับกลุ่มกันของข้อทดสอบที่ได้จากการวิเคราะห์ระดับความยาก (Wright 1977: 108)

5. ใช้วินิจฉัยความสามารถของผู้ตอบ (diagnostic)

ในกรณีที่โค้งลักษณะเฉพาะของข้อทดสอบไม่แนบสนิท (fit) กับโค้งมาตรฐานของรูปแบบ แสดงว่ามีบางสิ่งบางอย่างผิดปกติในตัวผู้สอบ ซึ่งอาจใช้เป็นข้อมูลเพื่อวินิจฉัยความสามารถของผู้ตอบได้

6. ใช้ในการค้นหาความเป็นอคติของข้อทดสอบ (item bias)

เมื่อเกิดเหตุการณ์ในข้อ 5 แสดงว่ามีความสามารถอื่นแฝงเข้ามาในความสามารถที่ต้องการวัด เราสามารถจะตรวจสอบความเป็นอคติของข้อทดสอบได้โดยการเปรียบเทียบความแตกต่างของพื้นที่โค้งลักษณะเฉพาะของกลุ่มบุคคลที่แตกต่างกันแต่มีระดับความสามารถเท่ากัน แม้แต่คำๆเดียวที่ทำให้เกิดอคติก็สามารถค้นพบสาเหตุได้ (Wright 1977: 12)

7. ใช้ในการสร้างแบบทดสอบวัดความสามารถของแต่ละบุคคล

ผู้ใช้อาจสุ่มข้อทดสอบที่วิเคราะห์แล้วมาจัดเรียงระดับตามความยากและง่ายเป็นรูปพีระมิด เพื่อใช้ทดสอบระดับความสามารถของแต่ละบุคคลได้ แบบทดสอบชนิดนี้อาจเรียกว่า แบบทดสอบแบบปรับเปลี่ยน (adaptive test) เพราะแบบทดสอบสามารถปรับเปลี่ยนความยากง่ายของแต่ละข้อให้เหมาะสมกับระดับความสามารถของแต่ละบุคคล หรืออาจเรียกว่าแบบทดสอบเฉพาะบุคคล หรือแบบทดสอบทรงพีระมิด (pyramidal test) และในปัจจุบันนี้นิยมใช้เครื่องคอมพิวเตอร์ช่วยจึงมักเรียกแบบทดสอบแบบปรับเปลี่ยนโดยอาศัยคอมพิวเตอร์ (computerized adaptive test : CAT) ข้อทดสอบที่จะใช้เพื่อสร้างแบบทดสอบชนิดนี้มักได้รับการวิเคราะห์อย่างน้อยก็ด้วยวิธีโมเดลราช

8. ใช้ในการจัดชั้นเรียน (grade – placement tailoring)

ค่า θ ของผู้สอบที่ได้จากการวิเคราะห์ข้อทดสอบอาจใช้ในการจัดชั้นเรียนให้เหมาะสมกับระดับความสามารถของผู้เรียนเป็นกลุ่มๆได้

9. ใช้ในการเปรียบเทียบคะแนนต่างชุด (equating scores)

ถ้าแบบทดสอบต่างชุดกัน มีข้อทดสอบบางส่วที่เหมือนกัน (anchored items) เราสามารถนำผลการสอบที่วิเคราะห์แล้วด้วยแบบความสามารถแฝงเปรียบเทียบกันได้ เพราะคะแนนแต่ละชุดจะถูกแปลงเป็นคะแนนมาตรฐานที่สามารถใช้เปรียบเทียบกันได้ใช้ศึกษา

ปัญหาการเดาและความบกพร่องในการตอบแบบทดสอบ การศึกษาดังกล่าวอาศัยเส้นแนวโน้มที่ได้จากสมการถดถอยระหว่างค่าคงเหลือจากการตอบสนอง (response residual) ($-e^{(\theta - \beta)/2}$ ถ้า X = 0 และ $-e^{-(\theta - \beta)/2}$ ถ้า X = 1) กับระดับความยากของข้อทดสอบ ถ้าเส้นโค้งขึ้นแสดงว่าผู้สอบตอบข้อทดสอบด้วยการเดา และถ้าเส้นโค้งลงแสดงว่าผู้สอบตอบข้อทดสอบด้วยความสะเพร่า (ตอบข้อสอบที่ง่ายผิด) และถ้าค่าสมการถดถอยระหว่างค่าคงเหลือจากการตอบสนองกับตำแหน่งของข้อทดสอบเป็นบวกแสดงว่าผู้สอบตอบข้อทดสอบได้ไม่เป็นระบบ และถ้าค่าเป็นลบแสดงว่าผู้สอบตอบข้อทดสอบได้ไม่ทันเวลาที่กำหนดไว้ (Wright 1977: 112)

10. ใช้ในการสร้างแบบทดสอบที่ดีที่สุด (best test design)

ผลจากการวิเคราะห์ข้อทดสอบโดยแบบความสามารถแฝงสามารถนำไปใช้ในการสร้างแบบทดสอบที่มีคุณลักษณะต่างๆตามที่ต้องการได้ เช่น ทำให้ระดับความยากของข้อทดสอบมีโค้งตามลักษณะที่ต้องการได้ และถ้าแบบทดสอบมีข้อทดสอบที่มีระดับความยากอยู่ระหว่าง M = 2 S.D, (M = mean ability) และมีความยาวเป็น L = 6/SEM แล้วจะเป็นแบบทดสอบที่ดีที่สุดได้

ข้อเสนอแนะในการใช้โมเดลราช

อาร์ อาร์ เรนตซ์ (R. R. Rentz) และ ซี อาร์ เรนตซ์ (C. R. Rentz) ได้ศึกษาสารสนเทศและงานวิจัยที่เกี่ยวข้องกับการใช้การวิเคราะห์ข้อทดสอบโมเดลราชจำนวนมาก เพื่อศึกษาว่าโมเดลราชนี้สามารถใช้ได้จริงหรือไม่ ปรากฏว่าบุคคลทั้งสองได้สรุปข้อค้นพบและข้อเสนอแนะเกี่ยวกับการใช้การวิเคราะห์ข้อทดสอบโมเดลราชที่น่าสนใจหลายประการ (Rentz and Rentz 1978: 1 – 18) ดังนี้

โมเดลราชกับการสอบและแบบทดสอบ

<u>ข้อค้นพบ</u>

1. รูปโมเดลราชสามารถใช้วิเคราะห์ข้อทดสอบของแบบทดสอบวิชาต่างๆ ได้ดี เช่น คณิตศาสตร์ ภูมิศาสตร์ ชีววิทยา และภาษาอังกฤษ แบบทดสอบอาจเป็นแบบทดสอบวินิจฉัย แบบทดสอบชุด (battery test) แบบทดสอบอิงเกณฑ์ แบบทดสอบ สัมฤทธิผล และ

แบบทดสอบเชาวน์ปัญญา หรือความถนัดเฉพาะด้าน (aptitude test) แต่แบบทดสอบเหล่านี้มักเป็นแบบทดสอบแบบเลือกตอบ

2. รูปโมเดลราชใช้ได้ดีทั้งกับแบบทดสอบอิงเกณฑ์และแบบทดสอบอิงกลุ่มแต่ปกติแล้วจะใช้ได้ดีกับแบบทดสอบที่มีลักษณะเป็นเอกพันธ์และค่อนข้างง่าย

3. ลักษณะของเนื้อหา (test content) และรูปแบบของข้อทดสอบ (item format) ไม่ใช่ลักษณะที่จะกำหนดว่าแบบทดสอบนั้นเหมาะสำหรับที่จะใช้กับโมเดลราชหรือไม่ เช่น ข้อทดสอบปลายเปิด (open-endes item) และแบบสอบถามชนิดลิเคิร์ต (Likert-type) ก็สามารถใช้วิเคราะห์ด้วยโมเดลราชได้

<u>ข้อเสนอแนะ</u>

1. เนื่องจากยังไม่มีวิธีที่ดีที่จะทดสอบความเป็นมิติเดียวของข้อทดสอบแม้แต้วิธีการวิเคราะห์องค์ประกอบก็ยังไม่ดีพอ และแนวคิดนี้ก็ยังไม่ชัดเจนนักในเชิงปฏิบัติ ดังนั้นถ้าต้องการให้ข้อทดสอบสอดคล้องกับรูปแบบของราชแล้วควรจะสร้างข้อทดสอบหรือแบบทดสอบให้มีลักษณะดังนี้ คือ

1.1 ให้มีเนื้อหาในขอบเขตจำกัด เช่น กำหนดปริเขต (domain)

1.2 ให้มีข้อทดสอบที่ค่อนข้างง่ายเพื่อหลีกเลี่ยงการเดา

1.3 เขียนข้อทดสอบให้เป็นข้อทดสอบที่ดี ซึ่งอาจใช้แนวคิดการเขียนข้อทดสอบแบบประเพณีนิยมทั่วไปก็ได้

2. ในกรณีที่สามารถทำได้ควรจะทดสอบความเป็นมิติเดียวของข้อทดสอบก่อน เพื่อทำให้แบบทดสอบมีประสิทธิภาพ (efficiency) มากขึ้น เพราะจะทราบได้ว่าควรจะตัดข้อทดสอบข้อใดออกจากแบบทดสอบบ้าง เป็นต้น

3. ข้อทดสอบที่มีอำนาจจำแนกแตกต่างกันมากมักจะไม่สอดคล้องกับโค้งลักษณะเฉพาะของโมเดลราช ดังนั้นควรระวังการคัดเลือกข้อทดสอบด้วย

โมเดลราชกับการวิเคราะห์ข้อทดสอบ

<u>ข้อค้นพบ</u>

1. ข้อทดสอบที่มีลักษณะไม่ดี (bad item) ซึ่งข้อค้นพบจากการวิเคราะห์ข้อทดสอบโดยวิธีประเพณีนิยม เมื่อวิเคราะห์ด้วยวิธีโมเดลราชก็เป็นข้อทดสอบที่ไม่ดีด้วย ซึ่งมักพบได้จากการที่ข้อทดสอบนั้นไม่สอดคล้องกับโค้งลักษณะเฉพาะของข้อทดสอบ ยกเว้นในกรณีที่ข้อทดสอบมีพิสัยของอำนาจจำแนกแตกต่างกันมาก ข้อทดสอบที่มีอำนาจจำแนกต่ำหรือสูงจะไม่สอดคล้องกับโค้งลักษณะเฉพาะของข้อทดสอบเมื่อวิเคราะห์ด้วยวิธีโมเดลราช และถือว่าเป็นข้อทดสอบที่ไม่ดี แต่การวิเคราะห์ด้วยวิธีประเพณีนิยม ข้อทดสอบที่มีอำนาจจำแนกต่ำเท่านั้นที่ถือว่าเป็นข้อทดสอบที่ไม่ดี จากผลการวิจัยจำนวนมากพบว่าปกติแล้วข้อทดสอบที่วิเคราะห์ด้วยวิธีประเพณีนิยมและคัดเลือกแล้วว่าเป็นข้อทดสอบที่ดีประมาณ 5% - 10% จะไม่สอดคล้องกับโค้งลักษณะเฉพาะของข้อสอบที่ถือว่าเป็นข้อทดสอบที่ไม่ดีเมื่อวิเคราะห์ด้วยวิธีโมเดลราช (Rentz and Rentz 1978: 10)

อนึ่ง การวิเคราะห์ข้อทดสอบด้วยโมเดลราชไม่วิเคราะห์ตัวเลือก ดังนั้นข้อทดสอบที่จะวิเคราะห์ด้วยวิธีนี้ควรวิเคราะห์ด้วยวิธีประเพณีนิยม เพื่อปรับปรุงแก้ไขข้อทดสอบให้ดีขึ้นก่อนแล้วจึงวิเคราะห์ด้วยวิธีโมเดลราช

2. จุดมุ่งหมายของการวิเคราะห์ข้อทดสอบด้วยโมเดลราชและวิธีประเพณีนิยมแตกต่างกัน กล่าวคือ วิธีโมเดลราชมุ่งวิเคราะห์เพื่อประมาณระดับความยากของข้อทดสอบ และเพื่อประเมินความสอดคล้องของข้อทดสอบกับโค้ลลักษณะเฉพาะของข้อทดสอบ เพื่อค้นหาข้อทดสอบที่ไม่ดี แต่วิธีประเพณีนิยมมุ่งวิเคราะห์หาระดับความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์สหสัมพันธ์ของข้อทดสอบกับเกณฑ์เพื่อประเมินคุณภาพของแบบทดสอบ เช่น ค่าความแม่นตรง ค่าความเชื่อมั่น ค่าเฉลี่ย และความแปรปรวนของคะแนน เป็นต้น แต่วิธีโมเดลราชใช้ค่าความคลาดเคลื่อนมาตรฐานในการวัด (SEM) แทนค่าความเชื่อมั่นของแบบทดสอบ

อนึ่ง ปกติแล้วในการวิเคราะห์ข้อทดสอบด้วยวิธีประเพณีนิยมมักนิยมตัดข้อทดสอบที่ไม่ดีออก แล้ววิเคราะห์ข้อทดสอบใหม่ใหม่เพื่อหาคุณภาพของแบบทดสอบว่าดีขึ้น

หรือไม่โดยอาศัยข้อมูลเดิม แต่จากการศึกษางานวิจัยหลายชิ้นพบว่า การกระทำเช่นนี้ไม่มีความจำเป็นสำหรับการวิเคราะห์ข้อทดสอบโดยวิธีโมเดลราช

3. ขนาดของกลุ่มตัวอย่างผู้สอบที่จำเป็นต้องใช้ขึ้นอยู่กับความต้องการความถูกต้องจากการคำนวณว่ามีมากเพียงใด จากการวิจัยหลายชิ้นพบว่ากลุ่มตัวอย่างขนาด 500–1,000 คน เป็นขนาดที่ทำให้ค่าต่างๆที่คำนวณได้ถูกต้องและคงที่มาก (ซึ่งมีความจำเป็นสำหรับธนาคารข้อทดสอบ) กลุ่มตัวอย่างที่มากกว่า 1,000 คน ไม่มีความจำเป็นเพราะทำให้ค่าต่างๆถูกต้องเพิ่มขึ้นเพียงเล็กน้อยเท่านั้น

แต่อย่างไรก็ตาม หากผู้ใช้ไม่ต้องการความถูกต้องมากนัก กลุ่มตัวอย่างขนาด 200 คนก็สามารถให้ค่าที่คำนวณได้มีความคงที่มากพอ และถ้าต้องการค่าต่างๆเพื่อเทียบผลการสอบ (equating) อาจใช้ขนาดของตัวอย่างเพียง 125 คนก็เพียงพอ เพราะจะได้ค่าต่างๆเพื่อการเทียบผลที่คงที่เพียงพอแล้ว และเพื่อการเรียนการสอนหรือเพื่อประโยชน์ทั่วไปที่ไม่ต้องการความถูกต้องมาก กลุ่มตัวอย่างขนาด 100 คนก็อาจใช้ได้

<u>ข้อเสนอแนะ</u>

1. ใช้ค่าสถิติความสอดคล้องกับค่าเฉลี่ยกำลังสองเพื่อการตัดสินความสอดคล้องของข้อทดสอบกับโค้งลักษณะเฉพาะของข้อทดสอบ ค่าสถิตินี้เป็นค่าที่เกี่ยวข้องกับค่าเฉลี่ยและมักเรียกกันว่าความสอดคล้องกับค่าเฉลี่ยกำลังสอง (mean square fit) ค่าเฉลี่ยกำลังสองนี้คำนวณจากค่าเฉลี่ยกำลังสองของข้อทดสอบแต่ละข้อของกลุ่มบุคคลที่มีความสามารถในกลุ่มเดียวกัน (มีคะแนนดิบรวมแล้วเท่ากัน) แล้วทดสอบว่าค่าเฉลี่ยกำลังสองนี้มากกว่าค่าที่กำหนดไว้ในรูปแบบหรือไม่ ค่านี้เป็นค่าที่คาดหวังไว้และปกติจะมีค่าเท่ากับ 1.5 แต่ในโปรแกรม BICAL กำหนดค่าดังกล่าวไว้เท่ากับ 1.0

2. เลือกข้อทดสอบที่ดีไว้แล้วจำกัดข้อทดสอบที่ไม่ดีทิ้งไปหรือนำไปปรับปรุงแก้ไข

อนึ่ง ภายหลังจากการศึกษาจากงานวิจัยและบทความต่างๆที่เกี่ยวข้องกับรูปโมเดลราชจำนวนประมาณ 50 ชิ้น อาร์ อาร์ เรนตซ์ และ ซี อาร์ เรนต์ (Rentz and Rentz 1978: 18) สรุปว่านักพัฒนาข้อทดสอบสามารถใช้การวิเคราะห์ข้อทดสอบโมเดลราชเพื่อการสร้างข้อทดสอบได้อย่างสบายใจ แต่ก็ควรตระหนักถึงข้อควรระวังบางอย่างในการแปลความหมายของการวิเคราะห์ข้อมูลไว้ด้วย

สรุป

เราจะเห็นว่าการวิเคราะห์ข้อทดสอบตามโมเดลราช ซึ่งเป็นรูปแบบหนึ่งของแบบความสามารถแฝงที่มีจำนวนพารามิเตอร์น้อยที่สุดในจำนวนรูปแบบต่างๆที่ได้กล่าวมาแล้วข้างต้น มีประโยชน์มากสำหรับการทดสอบและวัดผล การคำนวณหาค่าพารามิเตอร์ต่างๆมักเสียเวลาและค่าใช้จ่ายน้อยกว่าแบบอื่นๆ ดังนั้นจึงมีผู้นิยมใช้การวิเคราะห์ข้อทดสอบโมเดลราชไปประยุกต์ในการวัดและประเมินผลอย่างกว้างขวาง และได้มีผลงานวิจัยที่ศึกษาเกี่ยวกับความแกร่งของโมเดลนี้มากผู้ที่สนใจต้องการใช้วิธีโมเดลราชจึงควรศึกษาเรื่องความแกร่งเหล่านี้ด้วยนอกเหนือจากข้อตกลงเบื้องต้นทั่วๆไปในการใช้รูปแบบนี้ ทั้งนี้ เพื่อที่จะได้แน่ใจว่าการนำไปใช้จะได้ผลตามความมุ่งหมาย

บรรณานุกรม

สุพัฒน์ สุกมลสันต์. **การวิเคราะห์ข้อทดสอบและตัดเกรดด้วยคอมพิวเตอร์.** กรุงเทพ: บริษัทวิทยพัฒน์ จำกัด, 2542.

Andrich David. **Rasch Model for Measurement**. Sage Plubications, 1988.

Baker, B. B. **Item Response Theory: Parameter Estimation Techniques**. New York: Marcel Dekker, Inc, 1992.

Embretson, S. E. And Steven, P. Reise. **Item Response Theory for Psychologists**. London: Lawrence Erlbaum Associates Publishers, 2000.

Fischer, G. H. And Molennar, I. W. Rasch Models: **Foundation, Recent Developments, and Application**. New York: Springer-Verlag New York, Inc, 1995.

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(1977).

Hambleton, R. K. and others. **Fundamentals of Item Response Theory**. London: Sage Publications, 1991.

Linden, W. J. And Hambleton, R. K. **Handbook of Modern Item Response Theory**. USA: Springer-Verlag New York Inc, 1997.

Hulin, C. L. And others. **Item Response Theory: Application to Psychological Measurement.** DOW JONES-IRWIN, 1983.

Lord, F. M. **Applications of Item Response Theory to Practical Testing Problems**. New Jersy: Lawrence Erlbaum Associates, 1980.

Lore, F. M. And Novick, M. R. " **Statistical Theories of Mental Test Scores".** Addison-Wesley Publishing Company, 1968.

Hambleton, R. K. And Swaminathan, H. Item Response Theory: Principles and Applications.Kliwer-Nijhoff Publishing, 1985.

Rentz R. R. and Rentz C. R. *" Does The Rasch Model Really Work?"* **ERIC Report** 67, New Jersy: Educational Testing Service, 1978.

Tinsley, H.A. and Davis, R. V. " *An Investigation of the Rasch Simple Logistic Model: Sample Free Items and Test Calibration*." **Journal of Educational and Psychological Measurement** 35 (1975).

Willmott, A. S. And Fowles, D. E. **The Objective Interpretation of Test Performance: The Rasch Model Applied.** London: NFER Publishing, 1974.

Willmott, A. S. " *What does Rash promise?*" **Journal of NFER** 22(1980).

Wright, B.D. and Panchapakesan, N. " *A Procedure for Sample-free Item Analysis."* **Journal of Educational and Psychological Measurement**, 1969.

Wright, B. D. And Stone, M. H. **Best Test Design**. Chicaco: Mesa Press, 1979.

Wright, B.D. " *Misunderstanding the Rasch Model*." **Journal of Educational Measurement** 14 (1977). (a)

.... " *Solving Measurement Problems with the Rasch Model.* " **Journal of Educational Measurement** 14(1977). (b)