บทที่ 2

ทฤษฎีแนวคิดและงานวิจัยที่เกี่ยวข้อง

ผู้วิจัยได้ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้อง โดยจะเสนอเนื้อหาเรียงตามลำดับดังนี้

1. ทฤษฎีการตอบข้อสอบ
   - 1.1 ประวัติความเป็นมาของทฤษฎีการตอบข้อสอบ
   - 1.2 หลักการและข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อสอบ
   - 1.3 ลักษณะเด่นของทฤษฎีการตอบข้อสอบ
   - 1.4 นอร์มัล โอไจฟ์ โมเดล
   - 1.5 ลอจิสติก โมเดล
   - 1.6 เปรียบเทียบ นอร์มัล โอไจฟ์ โมเดล กับ ลอจิสติก โมเดล
   - 1.7 การประมาณค่าต่างๆในโมเดลของทฤษฎีการตอบข้อสอบ
2. ราชค์โมเดล
   - 2.1 ประวัติความเป็นมาของราชค์โมเดล
   - 2.2 แนวคิดของราชค์โมเดล
   - 2.3 พัฒนาการของสูตรของราชค์โมเดล
   - 2.4 ข้อตกลงเบื้องต้นของราชค์โมเดล
   - 2.5 วิธีการประมาณค่าพารามิเตอร์ในราชค์โมเดล
   - 2.6 การทดสอบความสอดคล้องของข้อสอบและผู้สอบกับราชค์โมเดล
   - 2.7 การนำราชค์โมเดลไปใช้ประโยชน์
   - 2.8 ความแกร่งของราชค์โมเดล
   - 2.9 งานวิจัยที่เกี่ยวกับราชค์โมเดล

1. <u>ทฤษฎีการตอบข้อสอบ</u>

1.1 <u>ประวัติความเป็นมาของทฤษฎีการตอบข้อสอบ</u>

ทฤษฎีการตอบข้อสอบเป็นส่วนหนึ่งของทฤษฎีคุณลักษณะแฝง (Latent Trait Theory) ทฤษฎีคุณลักษณะแฝงนี้ เฟอร์กูสัน (Ferguson , 1942) และลอเล่ย์ (Lawley , 1943) เป็นผู้ริเริ่มขึ้น (บนัดดา วัยวัฒนะ 2528 หน้า 14)

ในปี ค.ศ. 1952 เฟรดเดอริค เอ็ม ลอร์ด (Frederic M. Lord) ได้เสนอวิทยานิพนธ์ในระดับปริญญาเอก เกี่ยวกับทฤษฎีการตอบข้อสอบ ซึ่งเขาเรียกว่า ทฤษฎีโค้งลักษณะรายข้อ (Item Characteristic Curve Theory) โดยลอร์ดกล่าวว่า โค้งลักษณะรายข้อ (Item Characteristic Curve : ICC) ของข้อสอบแต่ละข้อมีลักษณะเป็นนอร์มัล โอไจฟ์ (Normal Ogive) ต่อมาเรียกว่า นอร์มัล โอไจฟ์โมเดล (Normal Ogive Model) แต่เนื่องจากความสลับซับซ้อนทางคณิตศาสตร์ที่เป็นพื้นฐานของทฤษฎี การขาดแคลนโปรแกรมคอมพิวเตอร์ที่จำเป็นต้องใช้วิเคราะห์ข้อมูลตามทฤษฎี และความสงสัยในเรื่องของประโยชน์ที่จะได้รับเพิ่มขึ้นจากการวิจัยในแนวนี้ ทำให้ลอร์ดพัฒนาทฤษฎีได้ช้าอยู่ระยะหนึ่ง

ในปี ค.ศ. 1960 ยอร์จ ราชค์ (Georg Rasch) ได้เสนอแนวคิดของทฤษฎีนี้ใหม่ โดยใช้พารามิเตอร์เพียงตัวเดียวคือ ค่าความยากของข้อสอบ (b) ซึ่งเรียกกันโดยทั่วไปว่า ราชค์โมเดล (Rasch Model) ซึ่งเป็นสาเหตุหนึ่งที่ทำให้ลอร์ดกลับมาสนใจศึกษาทฤษฎีนี้อีกครั้งหนึ่งในปี ค.ศ. 1965 (Warm , 1979 : 19 อ้างถึงใน บนัดดา วัยวัฒนะ 2528 หน้า 15)

ในปี ค.ศ. 1965 ลอร์ด (Lord) ได้กลับมาศึกษาอีกครั้งและได้ข้อสรุปอันหนึ่งจากการใช้กลุ่มตัวอย่างกว่าแสนคน ผลการศึกษาแสดงให้เห็นว่า ปัญหาที่จะทำให้การศึกษาของเขาต้องชงักไปเป็นเวลานานนั้นไม่ได้เป็นปัญหาแต่อย่างใด และทฤษฎีการตอบข้อสอบเหมาะสมกับข้อสอบแบบเลือกตอบ เขาจึงเริ่มศึกษาทฤษฎีการตอบข้อสอบอีกครั้ง

ในปี ค.ศ. 1968 เบิร์นบอม (Birnbaum) ได้เสนอแนวคิดของทฤษฎีนี้ใหม่ ในรูปของลอจิสติกโมเดล (Logistic Model) โดยใช้พารามิเตอร์สองตัวคือ ค่าความยากของข้อสอบ (b) และค่าอำนาจจำแนกของข้อสอบ (a) ซึ่งเป็นรูปที่ง่ายกว่าที่ลอร์คได้เสนอไว้ในปี ค.ศ. 1952 และในเวลาต่อมา ได้มีการพัฒนาขึ้นเรื่อยๆ จนภายหลังได้พัฒนาใช้กับพารามิเตอร์ตัวเดียวคือ ค่าความยากของข้อสอบ (b) และพารามิเตอร์สามตัวคือ ค่าความยากของข้อสอบ (b) ค่าอำนาจจำแนกของข้อสอบ (a) และค่าการเดา (c) (Warm , 1979 : 21 อ้างถึงใน นันคดา วัยวัฒนะ 2528 หน้า 15)

1.2 <u>หลักการและข้อตกลงเบื้องต้นของทฤษฎีการตอบข้อสอบ</u>

หลักการสำคัญของทฤษฎีการตอบข้อสอบคือ ผลการสอบของผู้สอบจากแบบทดสอบใดๆนั้น ขึ้นอยู่กับความสามารถของผู้สอบ หรือกล่าวอีกอย่างหนึ่งได้ว่า คะแนนของผู้สอบจากแบบทดสอบใดๆนั้น สามารถอธิบายและทำนายได้จากความสามารถของคนนั้นๆ

ให้ผลการสอบหรือคะแนน (Performance or Score) เป็นตัวแปรที่สามารถสังเกตและบันทึกได้โดยตรง ส่วนความสามารถ (Trait or Ability) เป็นตัวแปรที่มองเห็นไม่ได้และวัดไม่ได้โดยตรง

ถ้าให้ P แทน ผลการสอบ
θ แทน ความสามารถ
f แทน ฟังก์ชั่น

แล้วจะได้ความสัมพันธ์ในรูปทั่วไปดังนี้

$$P = f(\theta)$$

จากความสัมพันธ์ในรูปทั่วๆ ไปดังกล่าวข้างต้น ต้องอาศัยข้อตกลงเบื้องต้นดังนี้

1.2.1 <u>ความเป็นมิติเดียว</u> (Unidimensionality)

ความเป็นมิติเดียว หมายถึง ข้อสอบวัดเนื้อหาความรู้หรือความสามารถเดียวเท่านั้น (Warm , 1978 : อ้างถึงใน ประเวศ เวชชะ 2520 หน้า 21) หรือข้อสอบที่ข้อคำถามแต่ละข้อเป็นเอกพันธ์ (Homogeneous) กล่าวคือข้อคำถามเหล่านี้วัดความสามารถเดียวกัน (A Single Ability or Latent Trait) (ผจงจิต อินทสุวรรณ 2525 หน้า 55) ได้มีผู้กล่าวเกี่ยวกับแบบทดสอบที่มีลักษณะความเป็นมิติเดียวหลายท่าน ดังนี้

สำเริง บุญเรืองรัตน์ (2529 หน้า 44) กล่าวว่า สมมุติว่าแบบทดสอบฉบับหนึ่งมี n ข้อ และใช้กลุ่มตัวอย่าง r กลุ่ม และสมมุติอีกว่าแบบทดสอบฉบับนี้วัดกลุ่มบุคคลที่มีความสามารถเท่ากับ θ ถ้านำคะแนนจากการสอบทั้ง r กลุ่มนั้นมาแจกแจงความถี่ แล้วคะแนนมีลักษณะการแจกแจงเป็นอย่างเดียวกัน ก็ถือได้ว่าแบบทดสอบนั้นมีลักษณะความเป็นมิติเดียว แต่ถ้าการกระจายของคะแนนมีลักษณะแตกต่างกันไป ก็แสดงว่าแบบทดสอบนั้นคงวัดสิ่งอื่นอีกด้วย ไม่ได้วัดความสามารถเดียวเท่านั้น เทคนิคที่ใช้ตรวจสอบข้อตกลงว่าแบบทดสอบนั้นมีลักษณะเป็นมิติเดียวหรือไม่ ก็ควรทำการวิเคราะห์องค์ประกอบ ถ้าพบว่ามีองค์ประกอบที่สำคัญมากกว่าหนึ่งองค์ประกอบ ก็ถือว่าแบบทดสอบนั้นไม่เป็นมิติเดียว ในทางปฏิบัติถ้าต้องการสร้างแบบทดสอบให้มีลักษณะเป็นมิติเดียว ก็อาจทำได้ด้วยการสร้างคำถามเพื่อวัดความสามารถสิ่งเดียวกัน แล้วก็ทำการวิเคราะห์องค์ประกอบเพื่อรวมคำถามที่มีองค์ประกอบเดียวกันไว้ด้วยกัน ก็จะได้ข้อสอบที่มีลักษณะเป็นแบบทดสอบที่วัดความสามารถเดียวกันได้

ผจงจิต อินทสุวรรณ (2528 หน้า 26) กล่าวว่า เราสามารถจะมีวิธีคร่าวๆในการดูว่า คำถามชุดนั้นเป็นมิติเดียวหรือไม่คือ

ก. หาแมทริกซ์ของสหสัมพันธ์เทคระคอริคระหว่างข้อ

ข. ประมาณค่าคอมมิวนอลลิตี้ (Communalities) แล้วนำไปแทนลงในแนวเส้นทะแยงมุมของแมทริกซ์

ค. คำนวณค่าเลเท้นรู้ท (Latent Root) ของแมทริกซ์ของสหสัมพันธ์เทคระคอริคระหว่างข้อ ที่มีค่าประมาณของคอมมิวนอลลิตี้อยู่ใน

แนวเส้นทะแยงมุมนั้น ถ้าผลที่ได้มีลักษณะดังนี้ (1) เลเท้นรู้ทตัวแรกมีค่าสูง เมื่อเปรียบเทียบกับค่าตัวที่สอง และ (2) ค่าของเลเท้นรูทตัวที่สองไม่สูงกว่าค่าของเลเท้นรูทตัวอื่นๆมากนัก เราจึงสรุปได้ว่า ข้อคำถามเหล่านี้มีมิติเดียวโดยประมาณ วิธีนี้ยังคงใช้ได้แม้ในกรณีที่สหสัมพันธ์เทตระคอริคไม่ค่อยเหมาะสมนัก (ข้อสังเกตคือ การวิเคราะห์องค์ประกอบโดยวิธีไล้ค์ลี่ฮู้คสูงสุดของโจเรสโขก (Jores Kog) และการทดสอบนัยสำคัญที่ใช้ควบคู่กันนั้นไม่สามารถใช้กับแมทริกซ์ของสหสัมพันธ์เทตระคอริคได้อย่างแท้จริง)

ลอร์ด (Lord , 1980 : 20) กล่าวว่า แบบทดสอบที่น่าจะเป็นมิติเดียวโดยประมาณนั้นคือ แบบทดสอบการสะกดคำ คำศัพท์ อ่านเอาเรื่อง เหตุผลเชิงคณิตศาสตร์ อุปมาอุปไมย อนุกรมตัวเลข และแบบทดสอบที่เกี่ยวกับมิติสัมพันธ์ ตัวอย่างของแบบทดสอบที่ไม่เป็นมิติเดียวคือ แบบทดสอบวัดผลสัมฤทธิ์วิชาเคมี ซึ่งบางส่วนต้องอาศัยทักษะคณิตศาสตร์ และบางส่วนไม่ต้องอาศัยความรู้ทางคณิตศาสตร์

1.2.2 <u>ความเป็นอิสระที่ระดับความสามารถเดียวกัน</u> (Local Independence)

หมายความว่า โอกาสของผู้สอบคนหนึ่ง จะตอบข้อสอบได้ถูกนั้นไม่ได้มีผลมาจากคำตอบของข้อสอบข้ออื่นในแบบทดสอบชุดเดียวกัน รวมทั้งหมายความถึงข้อสอบไม่มีความสัมพันธ์กับความแตกต่างของผู้สอบที่มีระดับความสามารถเดียวกัน

วอร์ม (Warm , 1978 : 101 อ้างถึงใน ประเวศ เวชชะ 2530 หน้า 21) กล่าวว่า การทดสอบความเป็นอิสระที่ระดับความสามารถเดียวกันมีความยุ่งยากมาก ดังนั้นโดยปกติแล้วจึงไม่มีการทดสอบกัน

1.2.3 <u>โค้งลักษณะของข้อสอบ</u> (Item Characteristic Curve)

โอกาสที่ผู้สอบจะตอบข้อสอบข้อหนึ่งถูกนั้น ขึ้นอยู่กับโค้งลักษณะข้อสอบของแต่ละโมเดลที่ใช้ ไม่ขึ้นกับการแจกแจงความสามารถของกลุ่มตัวอย่าง ดังนั้นลักษณะเด่นของคุณสมบัติข้อคำถามจึงไม่ขึ้นอยู่กับ

ลักษณะของกลุ่มผู้สอบและข้อคำถามจึงมีคุณสมบัติคงที่ตลอดเวลา

1.3 ลักษณะเด่นของทฤษฎีการตอบข้อสอบ

สงบ ลักษณะ (2525 หน้า 49) กล่าวว่า ถ้าข้อตกลงเบื้องต้นของการสอบตรงตามเงื่อนไขของทฤษฎี เราสามารถคำนวณค่าพารามิเตอร์ได้เหมาะสมกับโมเดล ผลดีที่ได้รับคือ

1.3.1 ค่าพารามิเตอร์ของผู้สอบคือ ค่าความยาก ค่าอำนาจจำแนก และค่าสัมประสิทธิ์การเดา จะเป็นค่าที่ไม่เปลี่ยนแปลงไปตามกลุ่มผู้สอบ ไม่ว่าจะนำไปสอบกับผู้ใดก็ตาม

1.3.2 เมื่อทราบลักษณะข้อสอบแต่ละข้อของผู้สอบคนใด เราสามารถคำนวณหาค่าความสามารถที่แท้จริงของบุคคลนั้นได้ ค่าความสามารถที่แท้จริงนี้มีความสัมพันธ์โดยตรงกับคะแนนจริง การคำนวณหาค่าความสามารถที่แท้จริงอาจใช้ข้อใดก็ได้ที่วัดในสิ่งเดียวกัน ลักษณะเช่นนี้ถือว่าเป็นลักษณะของความเป็นอิสระของข้อสอบ

1.4 นอร์มัล โอไจฟ์ โมเดล (Normal Ogive Model)

ลอร์ด (lord , 1952) เป็นผู้เสนอรูปแบบของ นอร์มัล โอไจฟ์ โมเดล ซึ่งมี ICC เป็น นอร์มัล โอไจฟ์ ฟังก์ชั่น (Normal Ogive Function) ของความสามารถ โดยโค้งนี้อธิบายได้ด้วยพารามิเตอร์สองตัวคือ ความยากของข้อสอบและอำนาจจำแนก ซึ่งเราเขียนแสดงเป็นรูปสมการได้ดังนี้

$$P_i(\theta) = \phi\,[L_i(\theta)] \quad ; \quad i = 1,2,3,\ldots,n \quad \text{-----} (1)$$

เมื่อ $\phi\,(x)$ เป็น Normal Cumulative Distribution Function

$$\phi\,(x) = \int_{-\infty}^{\infty} (t)\,dt$$

เมื่อ $\psi(t)$ เป็น Normal Frequency Function

และ $L_i(\theta) = a_i(\theta - b_i)$

สมการ (1) เราจะได้รูปแบบของ Normal Ogive Function ดังนี้

$$P_i(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt \quad ; i = 1,2,3,\ldots,n \quad \text{--(2)}$$

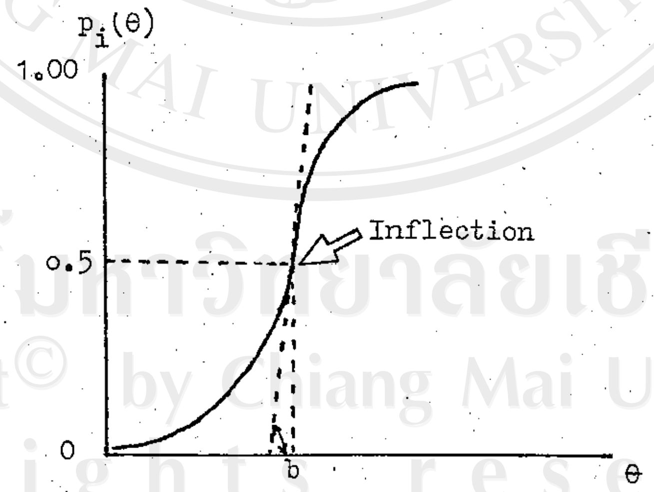
เมื่อ $P_i(\theta)$ คือ ความน่าจะเป็นที่ผู้สอบคนหนึ่งซึ่งมีความสามารถ θ จะตอบข้อสอบข้อที่ i ถูกต้อง

b_i คือ ค่าความยากของข้อสอบข้อที่ i

a_i คือ ค่าอำนาจจำแนกของข้อสอบข้อที่ i

θ คือ ระดับความสามารถของผู้สอบ

ภาพ 1 แสดงความสัมพันธ์ระหว่าง $P_i(\theta)$ และ θ ของนอร์มัล โอไจฟ์ โมเดล ที่มี 2 พารามิเตอร์



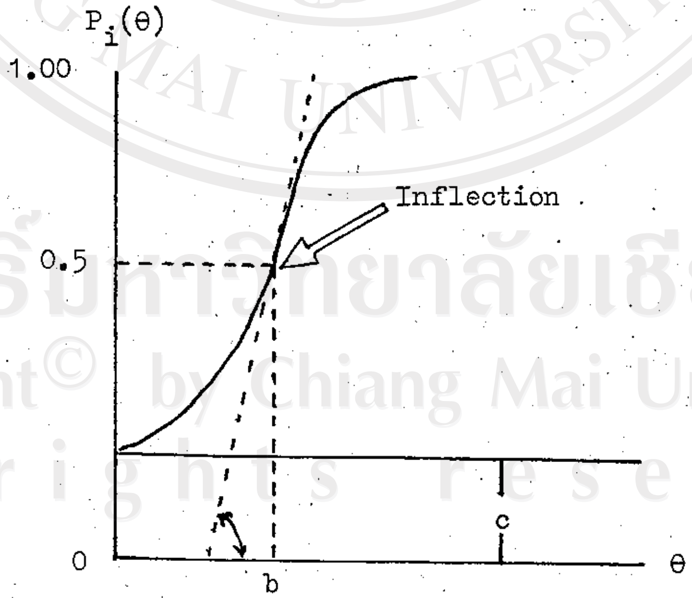
จากภาพที่ 1 ค่าความยาก (b) ได้มาจากค่าความสามารถที่ตรงกับจุดเปลี่ยนโค้ง (Inflection Point) ความสามารถระดับนี้จำเป็นสำหรับผู้สอบคนหนึ่งที่จะตอบข้อสอบข้อนั้นได้ถูกต้องด้วยความน่าจะเป็นเท่ากับ 0.5 โดยถือว่าคำตอบที่ถูกไม่ใช่ได้จากการเดา ($c=0$) ค่าอำนาจจำแนก (a) เป็นสัดส่วนของความชัน (Slope) ของ $P_i(\theta)$ ที่จุดเปลี่ยนโค้งหรือที่จุด $\theta = b$

ต่อมาภายหลังได้มีการดัดแปลง นอร์มัล โอไจฟ์ โมเดล จากที่ใช้กับพารามิเตอร์ 2 ตัว มาใช้กับพารามิเตอร์ 3 ตัว ซึ่งมีรูปสมการดังนี้

$$P_i(\theta) = c_i + (1-c_i) \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt \text{ ------ } (3)$$

เมื่อ $P_i(\theta)$, a_i , b_i , และ θ มีความหมายเหมือนข้างต้น ส่วน c_i คือ ค่าการเดาข้อที่ i (Guessing Parameter)

ภาพ 2 แสดงความสัมพันธ์ระหว่าง $P_i(\theta)$ และ θ ของนอร์มัล โอไจฟ์ โมเดล ที่มี 3 พารามิเตอร์



1.5 ลอจิสติกโมเดล (Logistic Model)

สำหรับข้อสอบแบบเลือกตอบ (Multiple Choice Item Test) นั้น แต่ละข้อคำถามซึ่งมีข้อเลือกตอบ A ข้อ ผู้สอบคนหนึ่งอาจเลือกตอบอย่างสุ่มหรือเดาคำตอบถูกได้ด้วยความน่าจะเป็นเท่ากับ 1/A ซึ่งแสดงว่าผู้สอบที่มีระดับความสามารถต่ำ ไม่จำเป็นว่าเขาจะมีความน่าจะเป็นในการตอบถูกเท่ากับศูนย์ หรือกล่าวอีกนัยหนึ่งได้ว่าเมื่อมีการเดาเกิดขึ้น โลเวอร์เอสซิมโทต์ (Lower Asymptote) ของ Pi (θ) ไม่เป็นศูนย์ พารามิเตอร์ของการเดา (Guessing Parameter) ต้องนำมาพิจารณา นอร์มัล โอไจฟ์ โมเดล (Normal Ogive Model) จึงไม่เหมาะสำหรับเหตุการณ์เช่นนี้ แม้ว่าจะสามารถดัดแปลงใช้กับพารามิเตอร์ 3 ตัวได้ แต่ไม่สะดวกเท่ากับการใช้ ลอจิสติก โมเดล

1.5.1 ลอจิสติกโมเดลที่ใช้พารามิเตอร์ 2 ตัว (Two-parameter Logistic Model)

เบิร์นบอม (Birnbaum,1968) เป็นผู้เสนอรูปแบบนี้ ซึ่งมีพารามิเตอร์ 2 ตัวคือ ความยากของข้อสอบ (b) และอำนาจจำแนก (a) และยังสามารถดัดแปลงใช้กับพารามิเตอร์ 3 ตัว หรือ 1 ตัว ได้ สามารถเขียนอยู่ในรูปของสมการได้ดังนี้

$$P_i(\theta) = \psi[DL_i(\theta)] \quad ; i = 1,2,3,\ldots,n \quad \text{--} (4)$$

เมื่อ $\psi(x)$ เป็น The Logistic Cumulation Distribution Function

$\psi(x) = [1+e^{-x}]^{-1}$; $L(\theta) = a_i(\theta - b_i)$

และ D เป็น Scaling factor ซึ่งเท่ากับ 1.7

e 2.71828...

จากสมการ (4) เราสามารถเขียนรูปเต็มของฟังก์ชันลอจิสติกที่ใช้พารามิเตอร์ 2 ตัว (Two - parameter Logistic Function) ได้ดังนี้

$$P_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \quad ; i = 1,2,3,...,n \text{--} (5)$$

1.5.2 <u>ลอจิสติกโมเดลที่ใช้พารามิเตอร์ 3 ตัว</u> (Three - parameter Logistic Model)

เมื่อมีเหตุผลสมควรที่จะรวมเอาพารามิเตอร์ของการเดา (Guessing Parameter : c) เข้าไปในโมเดล เราสามารถดัดแปลงโมเดลเล็กน้อย ดังนั้นรูปสมการคือ

$$P_i(\theta) = c_i + (1-c_i) \quad [Dl_i(\theta)] \quad \text{-----} (6)$$

เมื่อ $\psi(x)$ และ $L_i(\theta)$ มีความหมายดังกล่าวมาแล้ว

ดังนั้น จากสมการ (6) เราสามารถเขียนรูปเต็มของฟังก์ชั่นลอจิสติกที่ใช้พารามิเตอร์ 3 ตัว (Three-parameter Logistic Function) ได้ดังนี้

$$P_i(\theta) = c_i + (1-c_i) \frac{e^{Da_i(\theta-b_i)}}{1+e^{Da_i(\theta-b_i)}} \quad ;i=1,2,3,..,n\text{-} (7)$$

1.5.3 <u>ลอจิสติกโมเดลที่ใช้พารามิเตอร์ 1 ตัว</u> (One - parameter Logistic Model)

โดยมิได้เกี่ยวข้องกันเลย ราชค์ (Rasch,1960, 1966) ได้พัฒนาโมเดลคุณลักษณะแฝง (Latent Trait Model) ที่ฟังก์ชั่น (Function) สามารถอธิบายได้ด้วยพารามิเตอร์ของข้อสอบเพียงตัวเดียวคือ ค่าความยากของข้อสอบ (b) ปรากฏว่าโมเดลที่พัฒนาขึ้นนี้กลายเป็นลอจิสติกโมเดลที่ใช้พารามิเตอร์ 1 ตัว (One-parameter Logistic

Model) ที่พัฒนาโดย เบร์นบอม (Birnbaum,1968) (Wright & Panchapakeson,1969;Wright,1977อ้างถึงใน ผจงจิต อินทสุวรรณ 2525 หน้า 61) โมเดลนี้ถือว่าไม่มีการเดา และอำนาจจำแนกของข้อสอบคงที่ตลอดทั้งฉบับ สามารถเขียนเป็นสมการได้ดังนี้

$$P_i(\theta) = \Psi[DL_i(\theta)] \quad ; i = 1,2,3,\ldots,n \text{ ------ } (8)$$

เมื่อ $\Psi(x)$ มีความหมายเหมือนที่กล่าวมาแล้ว

และ $L_i(\theta) = a(\theta - b_i)$

ดังนั้นจากสมการ (8) เราสามารถเขียนรูปเต็มของฟังก์ชั่นลอจิสติกที่ใช้พารามิเตอร์ 1 ตัว (One-parameter Logistic Function) ได้ดังนี้

$$P_i(\theta) = \frac{e^{Da(\theta - b_i)}}{1+e^{Da(\theta - b_i)}} \quad ; i = 1,2,3,\ldots,n \text{ ------ } (9)$$

เมื่อ a เป็นอำนาจจำแนกของทุกข้อซึ่งมีค่าเท่ากัน

ถ้าเราสมมุติให้ $a_i = 1$ และกำหนดให้ $D = 1$ แล้วจะได้รูปสมการที่ง่ายขึ้นดังนี้

$$P_i(\theta) = \frac{e^{(\theta - b_i)}}{1+e^{(\theta - b_i)}} \quad ; i = 1,2,3,\ldots,n \text{ ------ } (10)$$

1.6 เปรียบเทียบนอร์มัล โอไจฟ์ โมเดล กับ ลอจิสติกโมเดล

พิจารณา นอร์มัล โอไจฟ์ โมเดล จะเห็นว่า ICC เป็น นอร์มัล โอไจฟ์ ฟังก์ชั่น (Normal Ogive Function) ของความสามารถ โดยโค้งดังกล่าวจะอธิบายได้ด้วยพารามิเตอร์สองตัวคือ ความยาก (Item Difficulty : b_i) และอำนาจจำแนก (Item Discrimination :a_i) ดังนั้นในกรณีที่มีแบบทดสอบฉบับหนึ่งซึ่งประกอบด้วยข้อคำถามที่ตรวจให้คะแนนเป็นสองแบบ (Binary Item)คือ ตอบถูก ให้คะแนน 1 คะแนน ตอบผิด ให้คะแนน 0 คะแนน และแบบทดสอบฉบับวัดความสามารถเดียวกัน (Unidimension) การใช้โค้งลักษณะข้อสอบที่เป็น นอร์มัล โอไจฟ์ (Normal Ogive Item Characteristic Curve) ย่อมเป็นรูปแบบที่เหมาะสม (Lord and Novick , 1968) สำหรับในกรณีที่แบบทดสอบประกอบด้วยข้อคำถามที่เป็นแบบเลือกตอบ (Multiple-choice Item Tests) ถ้าแต่ละข้อคำถามมี A ตัวเลือก ความน่าจะเป็นในการตอบถูกอย่างสุ่มจะเป็น 1/A นั่นคือ สำหรับผู้มีความรู้ต่ำก็มิได้หมายความว่าผลการสอบจะเป็นศูนย์ กล่าวอีกนัยหนึ่งได้ว่า เมื่อมีการเดาเกิดขึ้น โลเวอร์ เอสซิมโทต (Lower Asymptote) ของ $P_i(\theta)$ ไม่เป็นศูนย์ พารามิเตอร์ของการเดา (Guessing Parameter : c_i)จึงมีความจำเป็นที่จะต้องนำมาพิจารณา ซึ่งนอร์มัล โอไจฟ์ โมเดล ไม่เหมาะสมที่จะใช้สำหรับกรณีนี้ แม้จะสามารถดัดแปลงโมเดลให้มีพารามิเตอร์เป็นสามตัวได้ (Lord and Novick , 1968) แต่ไม่สะดวกเท่ากับใช้ ลอจิสติกโมเดล ซึ่งมีความคล้าย นอร์มัล โอไจฟ์ มาก โดยสรุปอาจแสดงความคล้ายคลึงกันและแตกต่างกันของทั้งสองโมเดลได้ดังนี้

นอร์มัล โอไจฟ์ โมเดล	ลอจิสติก โมเดล
1. ICC ได้จาก Normal Distribution Function 2. ปกติอธิบายได้ด้วยพารามิเตอร์ 2 ตัว คือ a_i หรือ b_i (อาจดัดแปลงเป็น 1 หรือ 3 พารามิเตอร์ได้) 3. ไม่สะดวกต่อการใช้	1. ICC ได้จาก Logistic Distribution Function 2. อธิบายได้ด้วยพารามิเตอร์ 1 หรือ 2 หรือ 3 ตัว 3. เลือกใช้ได้สะดวกตามความเหมาะสม

1.7 <u>การประมาณค่าต่างๆในโมเดลของทฤษฎีการตอบข้อสอบ</u>

เมื่อสร้างสมการแสดงความสัมพันธ์ระหว่าง P และ θ ได้แล้ว ก็จะคำนวณหาค่า a_i, b_i, c_i และ θ โดยที่

ค่า a_i เป็นค่าอำนาจจำแนกของข้อสอบ ซึ่งเป็นสัดส่วนของความชัน (Slope) ของโค้ง ณ จุดเปลี่ยนโค้ง หรือที่ $\theta = b_i$ แม้ว่าโดยทฤษฎีค่าของ a_i จะนิยามให้อยู่ในช่วง $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติค่า a_i ที่เป็นลบ ก็แสดงว่าข้อสอบไม่ดี ใช้ไม่ได้ จะต้องตัดทิ้งข้อสอบข้อนั้นไป จึงมักพบว่า a_i มีค่าตั้งแต่ 0 ถึง +2 โดยที่ $a_i = 0$ ซึ่ง แสดงว่าข้อสอบข้อนั้นไม่มีอำนาจจำแนกหรืออำนาจจำแนกต่ำ และ $a_i = +2$ แสดงว่าข้อสอบจะมีอำนาจจำแนกสูง

ค่า b_i เป็นค่าความยากของข้อสอบ ซึ่งได้มาจากค่าความสามารถ (θ) ที่ตรงจุดเปลี่ยนโค้ง ความสามารถระดับนี้จำเป็นสำหรับผู้สอบคนหนึ่งที่จะตอบข้อสอบข้อนั้น ได้ถูกต้องด้วยความน่าจะเป็นเท่ากับ 0.5 โดยถือว่าคำตอบที่ถูกไม่ใช่เกิดจากการเดา โดยทฤษฎีค่าความยากของข้อสอบมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -2 ถึง +2 ถ้า $b_i = -2$ แสดงว่าข้อสอบง่ายมาก และ $b_i = +2$ แสดงว่าข้อสอบยากมาก

ค่า c_i เป็นค่าการเดา ซึ่งหมายถึง โอกาสของการตอบข้อสอบถูกของคนที่ไม่มีความสามารถเลย หรือคือ โอกาสของการเดาถูกนั่นเอง (Guessing Parameter or Pseudo-chance Score)

ค่า θ เป็นค่าแสดงระดับความสามารถของผู้สอบ โดยทฤษฎีค่าความสามารถมีค่าตั้งแต่ $-\infty$ ถึง $+\infty$ แต่ในทางปฏิบัติจะมีค่าอยู่ระหว่าง -3 ถึง +3 ถ้า $\theta = -3$ แสดงว่ามีความสามารถต่ำ และ $\theta = +3$ แสดงว่าความสามารถสูง

2. <u>ราชค์โมเดล</u> (Rasch Model)

2.1 <u>ประวัติความเป็นมาของราชค์โมเดล</u>

ราชค์โมเดล เป็นรูปแบบของการทดสอบอย่างหนึ่งของลอจิสติกโมเดล ที่มักจะเรียกกันว่า ลอจิสติกโมเดลที่ใช้พารามิเตอร์ 1 ตัว (One-parameter Logistic Model) แนวคิดนี้ ราชค์ (Rasch) นักคณิตศาสตร์ชาวเดนมาร์ค (Denmark) ได้คิดขึ้นและได้เสนอแนวคิดนี้ในปีค.ศ.1960 และได้ตีพิมพ์บทความในทำนองเดียวกันอีกในปีค.ศ.1961,1966 ต่อมาเบนจามินไรท์ (Bemjamin Wright) ได้เผยแพร่แนวคิดนี้ให้เป็นที่แพร่หลายทั่วไปตั้งแต่ปีค.ศ.1967 เป็นต้นมา

แนวคิดที่สำคัญของราชค์โมเดลก็คือ เป็นแนวคิดที่เชื่อว่าความเป็นปรนัยของการวัด (Objectivity of Measurement) ที่ไม่สามารถจะหาได้จากคลาสสิคคอลโมเดล (Classical Models) ทั่วๆไปก็คือ

2.1.1 ความเป็นอิสระของกลุ่มตัวอย่าง (Sample-free Test Calibration) กล่าวคือ การคำนวณหาค่าพารามิเตอร์ของข้อสอบ เป็นอิสระจากกลุ่มตัวอย่าง ค่าต่างๆของข้อสอบ เช่น ระดับความยากของข้อสอบจะไม่แปรเปลี่ยนไปตามลักษณะของกลุ่มตัวอย่าง

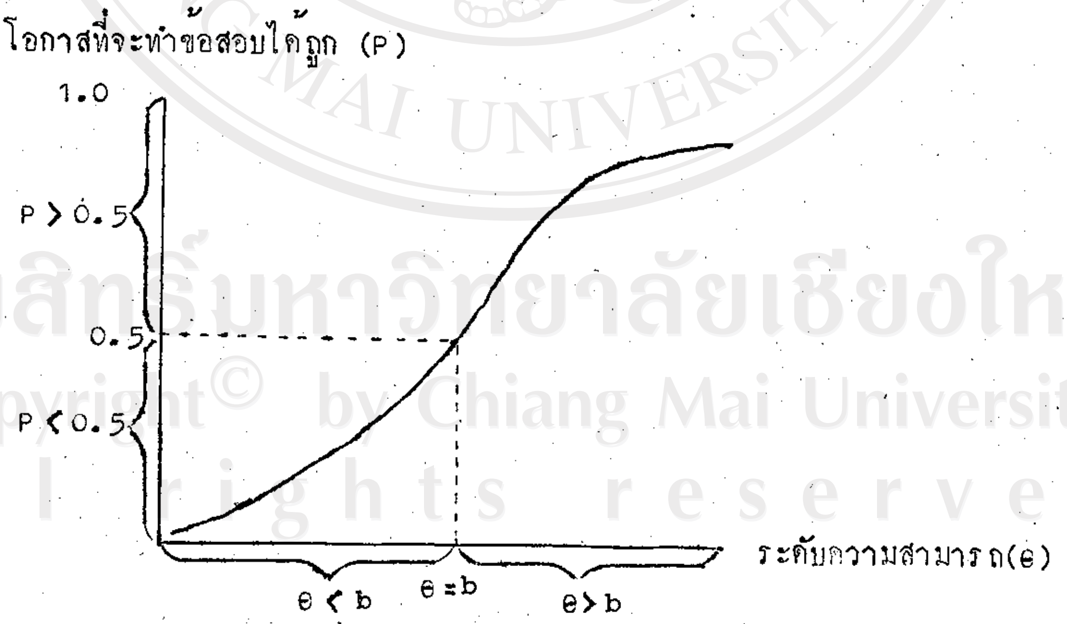
2.1.2 ความเป็นอิสระของข้อสอบ (Item-free Person Measurement) กล่าวคือ การคำนวณหาค่าพารามิเตอร์ของบุคคลเป็นอิสระจากข้อสอบ เช่น ความสามารถของบุคคลจะไม่แปรเปลี่ยนไปตามลักษณะของข้อสอบ ความสามารถของบุคคลจะคงที่ไม่ว่าจะวัดด้วยข้อสอบใด ข้อสอบหนึ่งๆที่คนเคยทำมาแล้วจะมีค่าคงที่เมื่อวัดบุคคลเดิม ไม่ว่าข้อสอบนั้นจะไปปรากฏอยู่ที่ส่วนใดของแบบทดสอบ

2.2 <u>แนวคิดของราชค์โมเดล</u>

ตามแนวคิดของราชค์โมเดลนั้น โอกาสที่คนจะทำข้อสอบได้หรือไม่ขึ้นอยู่กับระดับความสามารถของคนเอง (Ability Parameter : θ) และระดับความยากของข้อสอบ (Difficulty Parameter : β) เช่น ถ้า β = .5 และ θ = .5 โอกาสที่คนผู้นั้นจะสามารถทำข้อสอบข้อนั้นได้ถูกต้องประมาณ 50% ถ้าความสามารถของบุคคล (θ) น้อยกว่าความสามารถของข้อสอบ (ความยาก-ง่าย หรือ β) แล้ว โอกาสที่จะทำข้อสอบข้อนั้นก็ย่อมจะน้อยกว่า 50% และในทำนองเดียวกัน ถ้าหากว่าค่า θ มากกว่าค่า β แล้ว โอกาสที่คนผู้นั้นจะทำข้อสอบได้ถูกต้องก็มีมากกว่า 50% (ดูภาพที่ 3 ประกอบ)

ดังนั้น ตามแนวความคิดของราชค์โมเดลนั้น ค่าพารามิเตอร์ที่เกี่ยวข้องมีเพียง 2 ค่าเท่านั้น ไม่มีอำนาจจำแนกหรือโอกาสการเดาเข้ามาเกี่ยวข้องด้วย โดยเป็นข้อตกลงเบื้องต้นว่า ข้อสอบที่นำมาใช้กับราชค์โมเดลนั้นจะต้องมีอำนาจจำแนกเท่ากัน หรือมีลักษณะที่ทำให้เกิดการเดาได้น้อยที่สุด แต่อย่างไรก็ตามในทางปฏิบัติลักษณะทั้งสองนี้ความแกร่ง (Robustness)ของโมเดลสามารถแก้ได้

ภาพ 3 แสดงโค้งลักษณะของข้อสอบ (ICC) ของราชค์โมเดล



2.3 <u>พัฒนาการของสูตรของราชค์โมเดล</u>

ดังได้กล่าวมาแล้ว แนวคิดของ ราชค์โมเดล เกี่ยวข้องกับความสัมพันธ์ของความสามารถของบุคคล (θ) และความยากของข้อสอบ (β) เท่านั้น ความสัมพันธ์นี้เป็นความสัมพันธ์กันเชิงคณิตศาสตร์ กล่าวคือ โอกาสที่บุคคล V ที่มีระดับความสามารถ θ (หรือ θ_v) จะทำข้อสอบ C ที่มีระดับความยาก β (หรือ β_c) จะทำข้อสอบได้ถูกมากน้อยเพียงใดขึ้นอยู่กับระดับความแตกต่างของ ($\theta_v-\beta_c$) นั่นคือ

โอกาสของความสำเร็จ (Odd of success) = ($\theta_v-\beta_c$)

แต่เนื่องจาก ($\theta_v-\beta_c$) นั้นมีค่าระหว่าง $+\infty$ แต่ว่าโอกาสของความสำเร็จมีได้ระหว่าง 1 กับ 0 เท่านั้น เพื่อให้ค่าของ ($\theta_v-\beta_c$)เป็นค่าที่มีหน่วยเล็กลง และคงที่เหมาะแก่การนำไปใช้ จึงใช้ค่าเอกซ์โปเนนท์ (Exponent) ของ ($\theta_v-\beta_c$) แทน ซึ่งจะมีค่าอยู่ระหว่าง 0 กับ $+\infty$ ดังนี้

$$\text{โอกาสของความสำเร็จ} = e^{(\theta_v-\beta_e)} = \exp(\theta_v-\beta_e)$$

เมื่อ $e = 2.71828...$

เพื่อทำให้ค่า $\exp(\theta_v-\beta_e)$ มีค่าอยู่ระหว่าง 0 และ 1 อาจเขียนได้เป็น

$$\exp(\theta_v-\beta_e) / 1 + \exp(\theta_v-\beta_e)$$

ดังนั้นสูตรที่แสดงว่า โอกาสที่บุคคล v ที่มีความสามารถ θ จะทำข้อสอบ c ที่มีความยาก β ได้ถูกต้อง (ได้คะแนน = 1 หรือ x_{v1}) คือ

$$P\{x_{v1}=1/\theta_v, \beta_e\} = \exp(\theta_v-\beta_e)/[1+\exp(\theta_v-\beta_e) \text{ -- } (11)$$

และโอกาสที่บุคคลดังกล่าว จะทำข้อสอบดังกล่าวผิด ก็คือ

$$P\{X_{vc} = 0 / \theta_v, \beta_c\} = 1+\exp(\theta_v-\beta_c)/[1+\exp(\theta_v-\beta_c)]$$
$$= 1/[1+\exp(\theta_v-\beta_c)] \text{-------} \quad (12)$$

สมการที่ (11) และ (12) อาจเขียนรวมกันได้ดังนี้

$$P\{X_{vc}/\theta_v, \beta_c\} = \exp[X_{vc}(\theta_v-\beta_c)]/[1+\exp(\theta_v-\beta_c)] \text{---} \quad (13)$$

ในกรณีที่มีคนทำข้อสอบหลายๆคน และข้อสอบมีหลายๆข้อ ถ้าสมมุติให้ $((X_{vc}))$ แทนเดต้าแมทริกซ์ (Data Matrix) ของโอกาสที่คนจำนวน N คน ทำข้อสอบที่มีความยาว L ข้อ อาจเขียนได้ดังนี้

$$P\{((X_{vc}))/(\theta_v),(\beta_c)\} = \prod_{v}^{N}\prod_{c}^{L} \frac{\exp[X_{vc}(\theta_v-\beta_c)]}{1+\exp(\theta_v-\beta_c)} \text{-} \quad (14)$$

ถ้ากำหนดให้คะแนนรวมของบุคคลที่ v คือ $r_v = \sum_{c}^{L} X_{vc}$ (person score)

และจำนวนคนที่ทำข้อสอบข้อที่ c คือ $s_c = \sum_{v}^{N} X_{vc}$ (item score)

และเนื่องจาก

$$\prod_{v}^{N}\prod_{c}^{L} \exp[X_{vc}(\theta_v-\beta_c)] = \exp[\sum_{v}^{N}\sum_{c}^{L} X_{vc}(\theta_v-\beta_c)]$$

ดังนั้นสมการที่ (14) อาจเขียนใหม่ได้ดังนี้

$$P\{((X_{vc}))/(\theta_v),(\beta_c)\} = \frac{\exp[\sum_{v}^{N} r_v\theta_v - \sum_{c}^{L} S_c\beta_c]}{\prod_{v}^{N}\prod_{c}^{L}[1 + (\exp(\theta_v - \beta_c)]} \quad - (15)$$

สมการที่ (15) มีความสำคัญมาก เพราะแสดงว่า

(1) ในการที่จะคำนวณหาความยากของแต่ละข้อ (β_c) นั้น เราอาศัยแต่เพียงคะแนนรวมของแต่ละบุคคล (person score : r_v) เท่านั้นก็เพียงพอในการกำจัดค่า θ_v ออกจากสมการ ด้วยวิธีการคำนวณหา จึงเป็นอิสระจากกลุ่มตัวอย่าง (Sample-free Item Calibration)

(2) ในการหาความสามารถของแต่ละบุคคล (θ_v) เราอาศัยแต่เพียงคะแนนรวมของแต่ละข้อ (item score : S_c) เท่านั้นก็เพียงพอในการกำจัด β_c ออกจากสมการ ดังนั้นการคำนวณหาค่า θ_v จึงเป็นอิสระจากกลุ่มข้อสอบ (Item-free Person Measurement)

(3) การคำนวณหาค่า θ_v และ β_c เป็นอิสระจากกันและกัน

วิธีการคำนวณหาค่า θ_v และ β_c โดยวิธีการกำจัดค่าใดค่าหนึ่งออกไปจากสมการเรียกว่าวิธี Condition Maximum Likelihood Procedure เป็นที่ยุ่งยากสับสนในการคำนวณ อีกทั้ง ไรท์ (Wright) และ ดักลาส (Douglass) (Wright,1978:17 อ้างถึงใน ภาวิณี ศรีสุขวัฒนานนท์ และคณะ 2523 หน้า 4/5) ได้ทำการศึกษาแล้วพบว่าจะใช้วิธีนี้กับการวิเคราะห์ข้อสอบที่ยาวเกินกว่า 25 ข้อทำไม่ได้ ดังนั้น ไรท์ (Wright,1978:17 อ้างถึงใน ภาวิณี ศรีสุขวัฒนานนท์ และคณะ 2523 หน้า 4/5) จึงได้พัฒนาวิธีคำนวณหาค่า θ_v และ β_c โดยวิธี Unconditional Maximum Likelihood Procedure (UCON) ขึ้น ซึ่งสามารถใช้ได้กับการวิเคราะห์ข้อสอบที่มีขนาดใหญ่และเล็กได้ด้วย แม้ว่า

จะมีความลำเอียง (Bias) ในการคำนวณเพียงเล็กน้อยแต่ก็แก้ไขได้ วิธีดังกล่าวนี้ ในโปรแกรมไบคาล (BICAL) แนะนำให้ใช้เมื่อข้อสอบมีขนาดสั้น (ประมาณ 25 ข้อ) และการกระจายของคะแนนมีลักษณะเบ้ แม้ว่าจะมีขนาดของกลุ่มตัวอย่างมากหรือน้อยก็ตาม

ในกรณีที่ข้อสอบมีขนาดยาว (มากกว่า 25 ข้อขึ้นไป) และมีกลุ่มตัวอย่างขนาดใหญ่ (มากกว่า 400 คน) และเชื่อว่าการกระจายของคะแนนมีแนวโน้มว่า เป็นโค้งปกติแล้ว การคำนวณหาค่า θ_v และ β_i ควรจะใช้วิธี Cohen's Approximation (PROX) จะได้ผลดี

นอกจากนี้ในราวปีค.ศ.1950-1960 นั้น ราชค์ โมเดล ใช้วิธี Least Square Method ในการคำนวณหาค่าพารามิเตอร์ (Parameter) (Wright,1977:219 อ้างถึงใน ภาวิณี ศรีสุขวัฒนานนท์ และคณะ 2523 หน้า 5/5) แต่ว่าต่อมาไม่เป็นที่นิยม เพราะว่าต้องอาศัยกลุ่มตัวอย่างขนาดใหญ่มาก และให้ค่าไม่ถูกต้องมากนัก

2.4 ข้อตกลงเบื้องต้นของราชค์โมเดล

ราชค์โมเดลนอกจากจะมีข้อตกลงเบื้องต้นเหมือนกับทฤษฎีการตอบข้อสอบ (IRT) แล้ว ยังมีข้อตกลงเบื้องต้นเพิ่มเติมอีกดังนี้ (อวยพร วิบูลย์กาญจน์ 2526 หน้า 19-21)

(1) คะแนนเป็นแบบถูก 1 ผิด 0

(2) ข้อสอบแต่ละข้อมีอำนาจจำแนกเท่ากัน ในการคำนวณจึงให้เป็น 1 และไม่มีการเดาเกิดขึ้น

2.5 วิธีการประมาณค่าพารามิเตอร์ในราชค์โมเดล

ไรท์ (Wright;Afterward in Rasch , 1980 : 188) เสนอไว้หลายวิธีดังนี้ (อวยพร วิบูลย์กาญจน์ 2526 หน้า 19-21)

(1) The LOG Method เป็นวิธีการประมาณค่าพารามิเตอร์ในรูปของ log ของความสำเร็จ ซึ่งเป็นวิธีการที่เหมาะสมกับแบบทดสอบที่มีค่าอำนาจจำแนกเท่ากันคือ 1 และใช้จำนวนในการประมาณค่าพารามิเตอร์เป็นจำนวนมาก

(2) The PAIR Method เป็นวิธีการประมาณค่าพารามิเตอร์ โดยใช้การจับคู่ข้อสอบ เช่น 1-2 , 1-3 , 1-4 , 2-4 , ... เป็นต้น เอาจำนวนคนที่พยายามตอบ 2 ข้อ แต่ทำถูกเพียงข้อเดียวมาวิเคราะห์เพื่อหา Sample Free ดังนั้น จึงเหมาะที่จะใช้ประมาณค่าแบบทดสอบที่มีจำนวนข้อสอบไม่มาก

(3) The FCON Method เป็นวิธีประมาณค่าพารามิเตอร์ โดยใช้ข้อสอบทุกข้อและคนทุกคนมาวิเคราะห์ เหมาะสำหรับแบบทดสอบที่มีจำนวนข้อสอบน้อยกว่า 30 ข้อ ถ้ามากกว่าจะทำให้การประมาณค่ามีความคลาดเคลื่อน แต่ภายหลังสามารถใช้กับข้อสอบจำนวน 60-70 ข้อได้

(4) The UCON Method เป็นวิธีการประมาณค่าพารามิเตอร์ ด้วยวิธี Unidimensional Maximum Likelihood Procedure ใช้กับแบบทดสอบที่มีการกระจายของความสามารถของคนและความยากของข้อสอบเป็นการกระจายแบบโค้งปกติ และใช้กับแบบทดสอบที่มีจำนวนข้อตั้งแต่ 25 ข้อขึ้นไป โดยมากจะใช้กับแบบทดสอบที่มี 1,000 ข้อขึ้นไป ถ้าจำนวนน้อยใช้ FCON

(5) The PROX Method เป็นวิธีการประมาณค่าพารามิเตอร์ ด้วยวิธีของ Cohen´s Approximation ใช้กับแบบทดสอบที่มีการกระจายของความสามารถของคนและความยากของข้อสอบเป็นลักษณะเบ้ คือมีแนวโน้มเป็นปกติ โดยมาก PROX และ UCON จะให้ผลเหมือนกัน จะต่างกันที่ค่าความคลาดเคลื่อนมาตรฐานเพียงเล็กน้อย

(6) The UFORM Method ใช้กับแบบทดสอบที่มีการกระจายความยากของข้อสอบเป็นแบบ Uniform

สำหรับวิธี PROX สามารถประมาณค่าด้วยมือ หรือประมาณค่าด้วยเครื่องคอมพิวเตอร์โดยใช้โปรแกรมสำเร็จรูปที่เรียกว่า BICAL (Wright and Mead , 1978) หรืออาจใช้โปรแกรม LOGIST (Wood and Lord, 1976;Wingersky and Lord , 1976) โดยกำหนดเงื่อนไขให้ค่าอำนาจจำแนกเท่ากับ 1 และค่าการเดาเป็น 0

สำหรับการประมาณค่าด้วยมือ ไรท์ (Wright) เสนอขั้นตอนไว้โดยสรุปดังนี้ (สุธรรม์ จันทน์หอม 2526 หน้า 45-60)

<u>ขั้นตอนที่ 1</u> หลังจากตรวจแบบทดสอบแล้ว ให้ตัดข้อสอบที่คนทำถูกและผิดหมดออก และตัดคนที่ได้คะแนนเต็มและคะแนนศูนย์ออก

<u>ขั้นตอนที่ 2</u> กำหนดให้ค่า S_i คือจำนวนคนที่ทำข้อสอบข้อที่ i ถูก เมื่อข้อสอบมีจำนวนข้อจาก i ถึง L และกำหนดให้ n_r คือความถี่หรือจำนวนคนที่ได้คะแนน r มีค่า $r-1$ ถึง $L-1$ (เพราะคนที่ได้คะแนนเต็มและศูนย์ถูกตัดออก)

<u>ขั้นตอนที่ 3</u> คำนวณค่า

$$X_i = \ln[(N-S_i)/S_i]$$
คือ ค่าอัตราส่วนล็อคของจำนวนคนทำผิดต่อจำนวนคนทำถูก ในแต่ละข้อ

$$\bar{X} = \sum X_i/L$$
คือ ค่าเฉลี่ยของ X_i จากข้อสอบจำนวน L ข้อ

$$U = \sum (X_i-\bar{X})^2/(L-1)$$
คือ ค่าความแปรปรวนของ X_i จากทุกข้อ

$$Y_r = \ln[r/(L-r)]$$
คือ ค่าอัตราส่วนล็อคของจำนวนทำถูกต่อคนทำผิดที่ L คะแนน

$$\bar{Y} = \sum_{r}^{L-1} n_r Y_r/n$$
คือ ค่าเฉลี่ยของ Y_r จากจำนวน N คน

$$V = \sum n_r(Y_r-\bar{Y})^2/(N-1)$$
คือ ค่าความแปรปรวนของ Y_r จำนวน N คน

$$X = \left[\frac{1+U/2.89}{1-UV/8.35}\right]^{1/2}$$
คือ ค่าปรับขยายของความยากของข้อสอบเมื่อคำนึงถึงความยาวของข้อสอบแล้ว

$Y = \left[\dfrac{1+V/2.89}{1-UV/8.35}\right]^{1/2}$	คือ	ค่าปรับขยายของความสามารถของคน เมื่อคำนึงถึงกลุ่มคนทั้งหมดแล้ว
$d_i = Y(X_i-\bar{X})$	คือ	ค่าความยากของข้อสอบแต่ละข้อที่ปรับแล้ว
$b_r = XY_r$	คือ	ค่าความสามารถของคนแต่ละระดับคะแนนที่ปรับแล้ว
$SE_{(d_i)} = Y[N/S_i(N-S_i)]^{1/2}$	คือ	ค่าความคลาดเคลื่อนมาตรฐานของความยากของข้อสอบ
$SE_{(b_r)} = X[L/r(L-r)]^{1/2}$	คือ	ค่าความคลาดเคลื่อนมาตรฐานของความสามารถของคน

หมายเหตุ ค่า X_i และ Y_r อาจใช้ตารางสำเร็จของไรท์ได้

2.6 การทดสอบความสอดคล้อง (Fit) ของข้อสอบและผู้สอบกับราชต์โมเดล

หลังจากหาค่าพารามิเตอร์ (d,b) แล้ว จำเป็นต้องมีการทดสอบว่าข้อสอบหรือคนนั้นๆมีความสอดคล้องกับโมเดลเพียงใด โดยการพิจารณาจากเหตุการณ์ที่แต่ละคนตอบข้อสอบข้อหนึ่งๆ เป็นไปตามความคาดหมายหรือไม่ กล่าวคือ ถ้าคนๆหนึ่งตอบข้อสอบข้อนั้นผิดทั้งๆที่เขามีความสามารถมากกว่าความยากของข้อสอบ หรือตอบข้อนั้นถูกทั้งๆที่ความสามารถของเขามีน้อยกว่าความยาก ซึ่งทั้งสองเหตุการณ์นี้เป็นลักษณะของการตอบที่ไม่เป็นไปตามคาด (Unexpected Responses) แสดงว่า จะต้องมีสิ่งใดสิ่งหนึ่งไม่เหมาะสมกับโมเดล คืออาจจะเป็นข้อสอบหรือคนซึ่งจะต้องทำการทดสอบต่อไป และไรท์ได้เสนอวิธีการซึ่งสรุปได้ดังนี้

1. เมื่อคนที่ v ทำข้อสอบข้อที่ i จะได้

$X_{vi} = 0$ เมื่อตอบผิด

และ $X_{vi} = 1$ เมื่อตอบถูก

2. ทดสอบความเหมาะสมของคนที่ v (Person Fit)

2.1 $v_v = \sum_{i}^{L} Z_{vi}^2/(L-1) \sim F_{L-1,}$ (L คือจำนวนข้อ)

2.2 $Z_{vi}^2 = \exp(b-d)$ เมื่อ $X_{vi} = 0$ หรือตอบผิด

และ $= \exp(d-b)$ เมื่อ $X_{vi} = 1$ หรือตอบถูก

เมื่อ Z_{vi}^2 คือ ค่าความคลาดเคลื่อนมาตรฐานกำลังสองของเหตุการณ์ที่คนที่ v ทำข้อสอบi

2.3 $t_v = [\ln(v_v)+v_v-1][(L-1)/8]^{1/2} \sim N(0,1)$

2.4 เปรียบเทียบค่า t_v กับค่า t จากตารางทางเดียว

3. ทดสอบความเหมาะสมของข้อสอบ i (Item Fit)

3.1 $v_i = \sum_{v}^{N} Z_{vi}^2/(N-1) \sim F_{N-1,}$ (N คือจำนวนคน)

3.2 $t_i = [\ln(v_i)+(v_i-1)][(N-1)/8]^{1/2} \sim N(0,1)$

3.3 เปรียบเทียบ t_i กับ t จากตารางทางเดียว

2.7 <u>การนำราชค์โมเคลไปใช้ประโยชน์</u>

ไรท์ (Wright , 1980) ได้รวบรวมเกี่ยวกับประโยชน์ของราชค์โมเคลไว้ดังนี้ (อวยพร วิบูลย์กาญจน์ 2526 หน้า 21)

1. วิเคราะห์ข้อสอบ (Item Analysis) จะช่วยแก้จุดอ่อนในการวิเคราะห์แบบเดิมได้ ในแง่ทำให้ค่าพารามิเตอร์มีค่าคงที่ไม่แปรเปลี่ยนไปตามกลุ่มตัวอย่าง

2. สร้างคลังข้อสอบ (Item Bank) เนื่องจากแบบทดสอบที่วิเคราะห์ด้วยราชค์โมเคล จะให้ค่าพารามิเตอร์คงที่ ดังนั้นจึงสามารถสร้างแบบทดสอบเป็นชุดๆ ซึ่งสามารถเลือกไปใช้สอบได้ตามต้องการ

3. สร้างแบบทดสอบที่ดีที่สุด (Best Test Design) ผลจากการวิเคราะห์ข้อสอบด้วยราชค์โมเคล สามารถนำไปใช้ในการออกแบบแบบทดสอบที่มีคุณลักษณะต่างๆตามต้องการได้

4. วัดระดับความสามารถของแต่ละบุคคล (Self-Tailoring) โดยสุ่มข้อสอบที่วิเคราะห์แล้ว และมีระดับความยากเรียงกันตามลำดับเพียงจำนวนเล็กน้อยมาทดสอบระดับความสามารถของแต่ละบุคคลได้

5. ตรวจสอบความลำเอียงของข้อสอบ (Item Bias) ในการวิเคราะห์หา ICC ของแต่ละกลุ่ม ถ้า ICC ของข้อสอบนั้นๆแตกต่างกันตามกลุ่มที่นำมาทดสอบ เช่น คนในเมือง คนนอกเมือง ชาย หญิง แล้วแสดงว่าข้อสอบนั้นมีความลำเอียงเกิดขึ้น

6. วินิจฉัยความสามารถของผู้สอบ (Individual Diagnosis) ถ้า ICC ของข้อสอบไม่สอดคล้อง (Fit) กับโมเคล แสดงว่าอาจจะมีบางสิ่งบางอย่างผิดปกติในตัวผู้สอบ

2.8 <u>ความแกร่งของราชค์โมเคล (The Robustness of Rasch)</u>

ในการใช้ราชค์เพื่อทดสอบนั้น ได้มีการใช้รูปแบบต่างๆกันเพื่อหาความแกร่งของโมเคลนี้ พบมีความแกร่งในเรื่องต่างๆดังนี้ (ภาวิณี ศรีสุขวัฒนานนท์ และคณะ 2525 หน้า 6/5)

1. ขนาดของกลุ่มตัวอย่างที่จะใช้กับราชค์โมเคลนั้นอาจจะมีขนาดเล็กประมาณ 100 คนเท่านั้นก็ได้ (Wright,1977)

2. การกระจายของความสามารถของกลุ่มตัวอย่างหรือความ

ยากของข้อสอบไม่จำเป็นต้องเป็นโค้งปกติก็ได้ เพียงแค่มีแนวโน้มว่าจะเป็นโค้งปกติเท่านั้นก็ได้ และตัวอย่างไม่จำเป็นที่จะได้มาเพราะการสุ่มก็ได้ (Wright and Stone , 1969)

3. ข้อสอบไม่จำเป็นต้องมีอำนาจจำแนกเท่ากันก็ได้ และไม่จำเป็นต้องคำนึงถึงค่าการเดาก็ได้ (Wright and Panchapakesan,1969:25)

4. ข้อสอบเป็แบบปรนัยเลือกตอบ (Multiple Choice) ก็ใช้กับราชค์โมเดลได้ (Willmott , 1980)

2.9 <u>งานวิจัยที่เกี่ยวกับราชค์โมเดล</u>

ไวท์ลี และเดวิส (Whitely and Dewis , 1974 cited by Yen and Allen ,1979 : 261) วิเคราะห์แบบทดสอบอุปมาอุปมัยทางภาษา จำนวน 60 ข้อ โดยวิธีราชค์โมเดล มีข้อสอบที่ไม่สอดคล้องกับโมเดล 30-40 เปอร์เซนต์

เรนซ์ และ บาชอร์ (Rentz and Bashaw 1975 cited by Rentz 1979 : 7) ได้วิจัยพบว่า ขนาดของกลุ่มตัวอย่างระหว่าง 500-1,000 คน จะทำให้การประมาณค่าพารามิเตอร์มีความคงที่เพิ่มขึ้นอย่างรวดเร็ว และจะค่อยๆเพิ่มขึ้นเมื่อกลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์มีจำนวน 2,000 และ 4,000 คน

ฟอร์สเตอร์ (Forster 1976 cited by Rentz , 1979 :7) พบว่า กลุ่มตัวอย่างที่ใช้ประมาณค่าพารามิเตอร์ 200 คน ก็เพียงพอที่จะทำให้ค่าพารามิเตอร์ที่ได้มีค่าคงที่ กลุ่มตัวอย่างที่มากกว่า 1,000 คน จะไม่เพิ่มความคงที่ในการประมาณค่าพารามิเตอร์มากนัก

เสวก โยธินประเสริฐ(Yotinprasert,Sawek,1987:4369-A) ได้ศึกษาขนาดของกลุ่มตัวอย่างที่ส่งผลกระทบต่อความคลาดเคลื่อน โดยการใช้วิธีเปรียบเทียบคะแนนของทัคเกอร์ (Tucher) และราชค์โมเดล (Rasch Model) ภายใต้การออกแบบโดยใช้ข้อสอบร่วม (Commom Items) และกลุ่มที่ไม่สุ่ม (Nonrandom Groups) กลุ่มตัวอย่างที่เขาใช้มี 5 ขนาดคือ 25,50,75,100 และ 500 คน ตามลำดับ พบว่า เมื่อใช้ขนาดของกลุ่มตัวอย่างน้อยวิธีการของราชค์โมเดลค่อนข้างจะมีความคลาดเคลื่อนและ

ลำเอียงมากกว่าวิธีการของทัคเกอร์

สตีเฟน เอช. โกลแมน และ แนมซูรี เอส. ราซู (Steven H. Goldman and Namsury S. Raju , 1986 : 11) ได้ใช้แบบทดสอบสำรวจทัศนคติ SRA (SRA Attitude Survey) ไปทดสอบกับกลุ่มตัวอย่าง 3,000 คน และวิเคราะห์โดยใช้โมเดล 1 และ 2 พารามิเตอร์ ได้ค้นพบสิ่งที่น่าสนใจ 2 ประการคือ 1) ค่าประมาณความสามารถของผู้สอบ และค่าความยากของข้อสอบที่วิเคราะห์จากโมเดล 1 และ 2 พารามิเตอร์ จะมีความแม่นยำเมื่อใช้กับกลุ่มตัวอย่างจำนวน 250 คน 2) ค่าอำนาจจำแนกของข้อสอบ (Item Discrimination Indeces) ของโมเดล 2 พารามิเตอร์ จะแม่นยำเมื่อใช้กลุ่มตัวอย่าง 1,000 คน

ออร์ (Orr , 1983 : 3885-A) ได้ใช้ผลสอบวัดทักษะพื้นฐาน (Basic Skills Examination)ซึ่งสอบกับนักเรียนระดับ 8 (Eighth Graders) แล้วนำมาวิเคราะห์โดยราชต์โมเดล พบว่า การกำหนดความเร็วในการทำข้อสอบต่างกัน และการสลับตำแหน่งของข้อสอบ ทำให้ค่าความยากของข้อสอบเปลี่ยนไป แต่ไม่พบว่าส่งผลต่อความสามารถของผู้สอบ

ดักลาส (Douglass , 1981 :4000-A) ได้ใช้ข้อมูลผลสอบปลายภาค 100 ข้อ ที่สอบกับนักศึกษาระดับวิทยาลัย มาวิเคราะห์พบว่าการเปรียบเทียบคะแนนจริงระหว่างกลุ่มตัวอย่างที่มีขนาดและระดับความสามารถต่างกันโดยใช้ราชต์โมเดลให้ผลคงที่ และดีกว่า โมเดล 2 และ 3 พารามิเตอร์

โฮวาร์ด อี.เอ ทินสเลย์ และ เรอเน่ วี. เดวิส (Howard E.A. Tinsley and Rene V. Dawis , 1975 : 325) ได้นำแบบทดสอบที่ใช้คำ (Word) รูปภาพ (Picture) สัญญลักษณ์ (Symbol) และอุปมาอุปมัย (Analogies) ไปทดสอบกับนักเรียนระดับมัธยมศึกษา วิทยาลัย และข้าราชการพลเรือน ได้ข้อค้นพบที่น่าสนใจ 3 ประการคือ 1) ค่าความยากของข้อสอบตามราชต์โมเดล ไม่ทำให้ความสามารถของผู้สอบที่มีระดับความสามารถต่างกันแปรเปลี่ยน ถ้าหากใช้ขนาดของกลุ่มตัวอย่างพอเหมาะ 2) ค่าความยากของข้อสอบตามราชต์โมเดลที่ไม่แปรเปลี่ยนนั้นจะสัมพันธ์กับข้อสอบที่สอดคล้อง (Fit) กับโมเดล ซึ่งถ้าตัดข้อสอบที่มี

โอกาส (Probabilities) ต่ำออกไป ก็จะทำให้ค่าความยากของข้อสอบไม่แปรเปลี่ยนมากยิ่งขึ้น 3) การประมาณค่าความสามารถของผู้สอบ ซึ่งชี้ให้เห็นโดยคะแนนดิบของแบบทดสอบ จะไม่แปรเปลี่ยนในแต่ละระดับความสามารถของกลุ่มผู้สอบ เมื่อใช้ข้อสอบจำนวน 25 ข้อหรือมากกว่า แม้ว่าในการศึกษาจะใช้กลุ่มตัวอย่างน้อยกว่า 100 คนก็ตาม

อวยพร วิบูลย์กาญจน์ (2526 หน้า 32-45) ได้เปรียบเทียบผลการวิเคราะห์แบบทดสอบอุปมาอุปมัย ด้วยคลาสิคคอลโมเดลกับราชด์โมเดล โดยใช้กลุ่มตัวอย่าง 1,884 คน วิเคราะห์ข้อสอบด้วยเครื่องคอมพิวเตอร์โปรแกรม CIA (Classical Item Analysis) และโปรแกรม BICAL พบว่าข้อสอบที่คัดเลือกไว้ด้วยคลาสิคคอลโมเดล มีจำนวนมากกว่าข้อสอบที่คัดเลือกไว้ด้วยราชด์โมเดล อย่างมีนัยสำคัญที่ระดับ .05 หาความเชื่อมั่นโดยใช้สูตร K.R.20 ได้ค่าความเชื่อมั่นแบบเดิม (.85) และแบบราชด์ (.79) แตกต่างกันอย่างมีนัยสำคัญที่ระดับ .01 หาความเที่ยงตรงเชิงสภาพแบบเดิมได้ .49 และแบบราชด์ได้ .46

บนัดดา วัยวัฒนะ (2528 หน้า 60-70) ได้ทำการศึกษาผลการวิเคราะห์ข้อสอบโดยวิธีลอจิสติกโมเดลกับวิธีเดิม ในวิชาคณิตศาสตร์ ชั้นมัธยมศึกษาปีที่ 1 ใช้กลุ่มตัวอย่าง 1,687 คน วิเคราะห์ข้อสอบด้วยเครื่องคอมพิวเตอร์โปรแกรม CIA ,BICAL และ LOGIST พบว่า จำนวนข้อสอบที่คัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบเดิมกับพารามิเตอร์ตัวเดียว แบบเดิมกับพารามิเตอร์สองตัว และแบบเดิมกับพารามิเตอร์สามตัว มีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 หาความเชื่อมั่นโดยใช้สูตรแอลฟา (Coefficient Alpha) พบว่า ค่าความเชื่อมั่นของข้อสอบที่ได้รับการคัดเลือกไว้จากการวิเคราะห์ข้อสอบแบบเดิมกับพารามิเตอร์ตัวเดียว และแบบเดิมกับพารามิเตอร์สองตัว มีความแตกต่างกันอย่างมีนัยสำคัญที่ระดับ .05 ส่วนพารามิเตอร์ตัวเดียวกับสองตัว แบบพารามิเตอร์ตัวเดียวกับสามตัว แบบพารามิเตอร์สองตัวกับสามตัว และแบบเดิมกับพารามิเตอร์สามตัว มีความแตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ