

ความหมาย ของคะแนนจาก แบบทดสอบแบบเลือกตอบที่มีการเดา

ศส.ดร.อำพล รสรมเจริญ*

บทความนี้เขียนขึ้นโดยมีวัตถุประสงค์เพื่อเผยแพร่ผลการวิจัย และเพื่อเป็นการนำผลวิจัยไปใช้ประโยชน์ในทางปฏิบัติ โดยความรู้ที่ได้จากการวิจัยนี้จะนำไปใช้เพื่อการตัดสินใจที่ถูกต้อง การวิจัยที่กล่าวถึงนี้เป็นการวิเคราะห์และคำนวณหาคะแนนที่ได้จากความรู้จริงจากการสอบที่ใช้ข้อสอบแบบเลือกตอบ ซึ่งมีคะแนนที่ได้จากการเดาเข้าไปเกี่ยวข้องด้วย

เป็นที่ทราบกันดีว่า ในการสอบโดยใช้ข้อสอบแบบเลือกตอบ ผู้สอบสามารถเดาตอบได้ ซึ่งมีโอกาสถูกและได้คะแนนจากการเดาตอบนั้นซึ่งเราจะเรียกว่าคะแนนจากการเดา ดังนั้น คะแนนรวมที่ได้จากการสอบด้วยข้อสอบแบบเลือกตอบจึงประกอบด้วยสองส่วน คือคะแนนจากความรู้จริงส่วนหนึ่งและคะแนนจากการเดาอีกส่วนหนึ่ง โดยปกติแล้ว เราจะไม่ทราบว่ามีผู้สอบมีความรู้จริงกี่ข้อและเดาตอบกี่ข้อ เราทราบแต่เพียงคะแนนรวมเท่านั้น เมื่อเราใช้คะแนนรวมเป็นตัวตัดสินใจจึงเกิดความคลาดเคลื่อนขึ้นดังต่อไปนี้

1. การสอบแข่งขันและสอบคัดเลือกต่าง ๆ เช่น สอบคัดเลือกเข้ามหาวิทยาลัย สอบแข่งขันเข้ารับราชการ สอบแข่งขันเพื่อรับทุน
2. การสอบไล่ที่มีการตัดสินใจได้-ตก
3. การใช้คะแนนไปวิเคราะห์ข้อสอบ

ในการสอบแข่งขันและสอบคัดเลือกโดยใช้ข้อสอบแบบเลือกตอบและมีการเดา ทำให้เกิดความไม่ยุติธรรม คือ ผู้ที่มีความรู้น้อยกว่าอาจเดาตอบได้

คะแนนมากกว่า และทำให้ได้คะแนนรวมมากกว่าผู้ที่มีความรู้มากกว่า ทำให้ผู้ที่มีความรู้น้อยกว่าเป็นผู้สอบได้และผู้ที่มีความรู้มากกว่าเป็นผู้สอบตก (ไม่ได้รับคัดเลือก) อีกประการหนึ่ง การเดาทำให้เกิดความไม่เหมาะสม คือ ผู้ที่มีความรู้จริงไม่ถึงเกณฑ์ที่จะสอบได้แต่เดาถูกได้คะแนนรวมสูง และกลายเป็นผู้สอบได้ ความไม่ยุติธรรมและความไม่เหมาะสมทั้งสองประการนี้ ผู้เขียนได้ทำการวิจัยไว้ในเรื่อง "ความยุติธรรมและความเหมาะสมของข้อสอบปรนัยแบบเลือกตอบที่ใช้ในการสอบคัดเลือกหรือสอบแข่งขัน" (อำพล รสรมเจริญ (1)) โดยเป็นการคำนวณหาเปอร์เซ็นต์ของการเกิดความไม่ยุติธรรมและความไม่เหมาะสมของข้อสอบแบบเลือกตอบแบบต่าง ๆ เช่น ชนิดที่มีจำนวนตัวเลือก 2 3 4 หรือ 5 ตัวเลือก และความยากง่ายต่าง ๆ กัน ปรากฏผลว่า โดยเฉลี่ยแล้วจะเกิดความไม่ยุติธรรมประมาณ 15% และความไม่เหมาะสมประมาณ 25% หมายความว่า ในจำนวนผู้ที่สอบคัดเลือกเข้ามาได้นั้น มีประมาณ 15% ที่คะแนนความรู้จริงน้อยกว่าผู้ที่สอบเข้าไม่ได้ และมีประมาณ 25% ที่คะแนนความรู้จริงน้อยกว่าเกณฑ์ที่ควรจะต้องสอบเข้าได้ จะเห็นว่าการใช้คะแนนรวม (จากข้อสอบดังกล่าว) เป็นเกณฑ์ตัดสินใจต่าง ๆ ทำให้เกิดข้อผิดพลาด ส่วนในรายละเอียดอื่น ๆ นั้นผลวิจัยบ่งว่าข้อสอบชนิดที่มีจำนวนตัวเลือกมากกว่า มีเปอร์เซ็นต์ของความไม่ยุติธรรมและความไม่เหมาะสมต่ำกว่า และข้อสอบที่ยากทำให้เปอร์เซ็นต์ของความไม่ยุติธรรมและความไม่เหมาะสมสูงกว่าข้อสอบที่ง่าย ผู้

* ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยบูรพา

อ่านอาหารรายละเอียดเพิ่มเติมจากรายงานการวิจัย (1) และในบทความเรื่อง "แนวทางการเลือกแบบของข้อสอบ และการตรวจให้คะแนนในการสอบคัดเลือกหรือสอบแข่งขัน" (อำพล ธรรมเจริญ (2))

ในการสอบไล่ที่มีการตัดสินใจได้-ตกโดยใช้คะแนนรวมเป็นเกณฑ์การตัดสิน อาจเกิดความผิดพลาดในแง่ของความหมายของคะแนน กล่าวคือ สมมุติว่าเราตั้งเกณฑ์ว่าผู้ที่สอบได้ไม่น้อยกว่า 50% ของคะแนนเต็มเป็นผู้สอบได้ เราอาจหมายความว่าหรือเข้าใจว่า ผู้ที่สอบผ่านจะต้องมีความรู้ไม่น้อยกว่าครึ่งหนึ่งของคำถามในข้อสอบ ความเข้าใจเช่นนี้ผิดพลาดอย่างสิ้นเชิง ทั้งนี้เพราะเราคิดเปอร์เซ็นต์ของคะแนนรวมซึ่งมีคะแนนจากการเดาส่วนหนึ่งคะแนนจากความรูจริงจึงไม่ถึง 50% ดังที่เข้าใจ นั่นคือ โดยแท้จริงแล้วผู้ที่สอบผ่านบางคนมีคะแนนความรูจริงไม่ถึงเกณฑ์ที่จะสอบผ่าน เกิดความไม่เหมาะสมดังที่กล่าวมาแล้ว

การหาคะแนนความรูจริงจากคะแนนรวมนั้น ผู้เขียนได้ทำการวิจัย (เชิงคณิตศาสตร์) เพื่อหาสูตรต่าง ๆ ที่ใช้ในการแปลงคะแนนรวมให้เป็นคะแนนจากความรูจริง เป็นการวิจัย 2 เรื่องคือ "การแจกแจงของคะแนนจากความรูจริงของการสอบที่ใช้ข้อสอบแบบเลือกตอบ และมีการเดา" (อำพล ธรรมเจริญ (3)) และ "การใช้กำหนดการเชิงกำลังสองในการปรับค่าการแจกแจงของคะแนนจากการสอบที่ใช้ข้อสอบแบบเลือกตอบและมีการเดา" (อำพล ธรรมเจริญ (4)) ซึ่งจะกล่าวถึงรายละเอียดอย่างย่อในที่นี้

ให้ n เป็นจำนวนข้อของข้อสอบ

S เป็นจำนวนตัวเลือกของข้อสอบแต่ละข้อ

T เป็นคะแนนรวมของผู้สอบ i คน (T คือ จำนวนข้อที่ตอบถูก)

R เป็นคะแนนจากความรูจริง

X เป็นคะแนนจากการเดา

ปกติเราทราบ T แต่ไม่ทราบ R และ X แต่เราทราบความสัมพันธ์ว่า

$$T = R + X$$

ในการสอบเมื่อทราบคะแนนรวมของผู้สอบแต่ละคนแล้ว

เรานำคะแนนมาแจกแจงความถี่ และหาความถี่สัมพัทธ์ (ความถี่ของแต่ละคะแนนหารด้วยจำนวนผู้เข้าสอบ) ผลที่ได้คือตารางแสดงการแจกแจงความน่าจะเป็นของ T เราใช้สัญลักษณ์ $f(t)$ เป็นการแจกแจงของ T นอกจากตารางแล้ว การแจกแจงของ T ยังอาจหาได้จากสูตร

$$(1) f(t) = \sum_{r=0}^t g(r) \binom{n-r}{t-r} p^{t-r} (1-p)^{n-t} \quad t=r, r+1, \dots, n.$$

เมื่อ $p=1/s$ และ $g(r)$ เป็นการแจกแจงความน่าจะเป็นของ R (ซึ่งยังไม่ทราบ) จากการวิจัยของผู้เขียน (1) การแจกแจงของ R หาได้จากสูตร

$$(2) g(r) = \frac{\sum_{k=0}^r g(k) \binom{n-k}{r-k} p^{r-k} (1-p)^{(n-r)}}{(1-p)^{n-r}} \quad r=0, 1, 2, \dots, n.$$

สูตร (1) เป็นสูตรพื้นฐาน ใช้ในการแปลงเพื่อหาสูตร (2) ซึ่งเป็นสูตรสำคัญที่จะใช้หาความคาดหวังของคะแนนจากความรูจริงเมื่อทราบคะแนนรวม

ถ้าให้ $E(T)$ และ $E(R)$ เป็นค่าความคาดหวัง (หรือค่าเฉลี่ย) ของคะแนนรวม T และของคะแนนจากความรูจริง R ตามลำดับ จะมีความสัมพันธ์

$$(3) E(T) = (1-p) E(R) + np$$

และถ้าให้ $V(T)$ และ $V(R)$ เป็นค่าความแปรปรวนของคะแนนรวม T และของคะแนนจากความรูจริง R ตามลำดับ จะมีความสัมพันธ์

$$(4) V(T) = (1-p)^2 V(R) + (n-E(R)) p(1-p)$$

ถ้าเราทราบคะแนนจากความรูจริง เช่น ทราบคำตอบที่ถูกต้องแล้ว r ข้อ เหลือที่จะต้องเดาตอบอีก $n-r$ ข้อ เราประมาณจำนวนข้อตอบถูกโดยเป็นค่าเฉลี่ย หรือความคาดหวัง เขียนสัญลักษณ์ว่า $E(T:R = r)$

จะได้สูตร

$$(5) \quad E(T|R=r) = (1-p)r + np$$

แต่คำถามที่เราต้องการทราบคำตอบจริงๆ ก็คือถ้าได้คะแนนรวม t คะแนน แล้วจะได้คะแนนจากความรู้จริงเท่าไร ในทางทฤษฎีนั้นเราไม่สามารถบอกได้ว่าได้คะแนนเท่าไรแน่นอน แต่เราสามารถบอกได้เป็นค่าความคาดหวัง (หรือค่าเฉลี่ย) ว่าควรเป็นเท่าใด ใช้สัญลักษณ์ $E(R:T=t)$ หาได้จาก

$$(6) E(R:T=t) = (1/f(t)) \sum_{r=0}^t r g(r)^{n-r} c_{t-r}^{t-r} p^{t-r} (1-p)^{n-t}$$

การคำนวณตามสูตรนี้ต้องทราบการแจกแจงของ R ก่อน

ในการคำนวณตามสูตรที่ได้มาดังกล่าวอาจต้องใช้คอมพิวเตอร์ช่วยคำนวณ เพราะสูตรยาวและมีหลายขั้นตอน ผู้วิจัยได้เขียนโปรแกรมช่วยคำนวณไว้แล้ว ซึ่งจะพัฒนาเป็นซอฟต์แวร์และเผยแพร่ต่อไป

ในกรณีที่ได้รับการแจกแจงของ T คือ $f(t)$ จาก

คะแนนสอบ ค่าที่ได้จะมีการเบี่ยงเบนจากการแจกแจงตามทฤษฎี ทั้งนี้เป็นไปตามธรรมชาติ แต่จะส่งผลให้การคำนวณหาการแจกแจงของ R เบี่ยงเบนจากที่ควรจะเป็น กล่าวคือ อาจทำให้ค่า $g(r)$ เป็นลบในบางค่าของ r ซึ่งขัดแย้งกับทฤษฎีที่ว่า การแจกแจงจะมีค่าไม่เป็นลบ ดังนั้นก่อนที่จะนำผลไปใช้งาน ควรปรับค่าการแจกแจงของ T เพื่อให้การแจกแจงของ R มีค่าไม่เป็นลบ วิธีปรับค่ามีหลายวิธี แต่วิธีหนึ่งที่ผู้เขียนคิดว่าดีที่สุดก็คือปรับค่าโดยให้มีการเปลี่ยนแปลงน้อยที่สุดผู้เขียนกล่าวถึงโดยละเอียดในการวิจัย (4) วิธีการจะไม่นำมากล่าวในที่นี้ เพราะกระบวนการดำเนินการนั้นยาว จำเป็นต้องใช้คอมพิวเตอร์ช่วยในการคำนวณ และผู้เขียนก็ได้เขียนโปรแกรมคำนวณไว้แล้วเช่นเดียวกัน ซึ่งจะพัฒนาเป็นซอฟต์แวร์ที่สมบูรณ์ต่อไป

เพื่อความเข้าใจ จะขอยกตัวอย่างในการคำนวณดังต่อไปนี้

มีข้อสอบแบบเลือกตอบชนิด 4 ตัวเลือก จำนวน 3 ข้อ มีผู้สอบ 15 คน ได้คะแนนต่างๆ กัน ซึ่งเมื่อแจกแจงความถี่แล้ว ปรากฏผลดังนี้

การแจกแจงความถี่					
คะแนน	0	1	2	3	รวม
จำนวนผู้ได้คะแนน	6	1	6	2	15
การแจกแจงของคะแนนรวม T					
คะแนน	0	1	2	3	รวม
ความถี่สัมพัทธ์	.4000 (6/15)	.0667 (1/15)	.4000 (6/15)	.1333 (2/15)	1

จากนี้ หากการแจกแจงของ R ดังในสูตร (2)

การแจกแจงของ R และ T

r, t	g(r)	f(t)
0	.9481	.4000
1	-.5926	.0667
2	.6591	.4000
3	-.0074	.1333
รวม	1.0000	1.000

จะเห็นว่า g (r) มีค่าเป็นลบสำหรับบางค่าของ r ต่อไปดำเนินการปรับค่าแจกแจงของ T เพื่อให้ g(r) มีค่าไม่เป็นลบโดยวิธีใน (4) ได้ผลลัพธ์

การแจกแจงของ R และ T เมื่อปรับค่าแล้ว

r, t	g (r)	f(t)
0	.5531	.2333
1	.0000	.2333
2	.4296	.4000
3	.0173	.1333
รวม	1.0000	1.000

หาค่าความคาดหวังของ T (หรือของ R ก่อนก็ได้) จากตารางได้

$$E(T) = 0x.2333 + 1x.2333 + 2x.4000 + 3x.1333 = 1.4333$$

หาค่าความคาดหวังของ R ตามสูตร (3) ได้ (เราทราบว่า $p = 1/4 = .25$)

$$E(R) = (E(T) - np) / (1-p) = (1.4333 - 3x.25) / (1-.25) = .9111$$

การสอบครั้งนี้ คะแนนรวมเฉลี่ยเท่ากับ 1.4333 คะแนน จากความรู้จริงเฉลี่ยเท่ากับ .9111

ถ้าทราบคะแนนความรู้จริงหาค่าความคาดหวังของ T จากสูตร (5)

$$\text{ถ้า } R = 0 \quad \text{ได้ } E(T|R=0) = (1-.25) \cdot 0 + 3(.25) = .75$$

$$R = 1 \quad E(T|R=1) = (1-.25) \cdot 1 + 3(.25) = 1.5$$

$$R = 2 \quad E(T|R=2) = (1-.25) \cdot 2 + 3(.25) = 2.25$$

$$R = 3 \quad E(T|R=3) = (1-.25) \cdot 3 + 3(.25) = 3$$

สมมติว่าเราสอบไล่และตั้งเกณฑ์ว่าผู้ที่มีคะแนนความรู้จริงอย่างน้อย 50% ของคะแนนรวมจึงจะเป็นผู้ที่สอบ

ได้ นั่นคือ $R = 1.5$ คะแนนนั้นตรงกับคะแนนรวม
 $(1-.25) \times 1.5 + 3(.25) = 1.875$ นั่นคือ ต้องสอบได้คะแนน
 รวมไม่น้อยกว่า 1.875 คะแนน จึงจะถือว่าเป็นผู้สอบได้

ถ้าทราบคะแนนรวม และต้องการหาค่าความ
 คาดหวังของคะแนนจากความรู้จริง เราคำนวณได้จาก
 สูตร (6) ซึ่งต้องทราบการแจกแจงของคะแนนจาก
 ความรู้จริงดังข้างบน จะได้

ถ้า $T = 0$ ได้	$E(R T=0) = 0$
$T = 1$	$E(R T=1) = 0$
$T = 2$	$E(R T=2) = 1.611$
$T = 3$	$E(R T=3) = 2$

ผู้ที่ได้คะแนนรวม 0 หรือ 1 คะแนน คะแนนจาก
 ความรู้จริงคือ 0 คะแนน ผู้ที่ได้คะแนนรวม 2 คะแนน
 คะแนนจากความรู้จริงคือ (เฉลี่ย) 1.611 คะแนน ผู้ที่ได้
 คะแนนรวมเต็ม 3 คะแนน จะมีคะแนนจากความรู้จริง
 2 คะแนนเท่านั้น

คะแนนจากความรู้จริง ทำให้เราทราบความ
 หมายของคะแนนต่าง ๆ ดีขึ้น และใช้ได้ถูกต้องทำให้การ

ตัดสินใจต่าง ๆ ที่ต้องใช้คะแนนเป็นเกณฑ์เป็นไปอย่าง
 ถูกต้อง ทั้งนี้รวมถึงการวิเคราะห์ข้อสอบ ซึ่งที่ปฏิบัติกัน
 นั้น วิเคราะห์โดยคะแนนรวม ซึ่งอาจทำให้มีความหมาย
 ต่างไปจากที่เข้าใจกัน ผู้เขียนเสนอว่าควรหาวิธีที่จะ
 วิเคราะห์โดยใช้คะแนนจากความรู้จริง (ผู้เขียนเชื่อว่า
 จะเป็นการวิจัยใหม่ที่ยังไม่มีผู้ใดทำมาก่อน อนึ่ง การวิจัย
 ที่ผู้เขียนกล่าวมาข้างต้นผู้เขียนเชื่อว่าเป็นการวิจัยที่ใหม่
 และไม่มีผู้ใดทำมาก่อนเช่นเดียวกัน)

อย่างไรก็ตามทั้งหมดยกกล่าวมานี้ก็เป็นผลมาจาก
 การที่ยอมให้ผู้สอบเดาตอบได้ ถ้าหากเราป้องกันการเดา
 ตอบของผู้สอบได้ การแปลความหมายของคะแนน การ
 ตัดสินใจต่าง ๆ ก็กระทำได้ตรงตามที่ต้องการ และไม่
 เกิดความไม่เหมาะสมหรือความไม่ยุติธรรมต่าง ๆ ผู้เขียน
 เสนอให้มีการหักคะแนนการเดาเพื่อเป็นมาตรการทาง
 จิตวิทยาที่จะไม่ให้ผู้สอบเดาตอบ เหตุผลรายละเอียด
 ต่าง ๆ มีอยู่ใน (2) หวังว่าผู้มีหน้าที่เกี่ยวข้องกับเรื่องต่าง ๆ
 นี้จะได้นำไปพิจารณาตามความเหมาะสมต่อไป



บรรณานุกรม

- อำพล ธรรมเจริญ 'ความยุติธรรมและความเหมาะสมของข้อสอบปรนัยแบบเลือกตอบที่ใช้
 ในการสอบคัดเลือกหรือสอบแข่งขัน' รายงานการวิจัย มหาวิทยาลัยบูรพา 2534.
- อำพล ธรรมเจริญ 'แนวทางการเลือกแบบของข้อสอบและการตรวจให้คะแนนในการสอบ
 คัดเลือกหรือสอบแข่งขัน' วารสารศรีนครินทร์วิโรจน์และพัฒนา เมษายน 2535.
- อำพล ธรรมเจริญ 'การแจกแจงของคะแนนจากความรู้จริงของการสอบที่ใช้ข้อสอบแบบ
 เลือกตอบและมีการเดา' หนังสือการประชุมวิชาการสถิติประยุกต์ ครั้งที่ 9
 จัดโดยคณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ ร่วมกับคณะ
 วิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ สิงหาคม 2535. เล่มที่ 2
- อำพล ธรรมเจริญ 'การใช้กำหนดการเชิงกำลังสองในการปรับค่าการแจกแจงของคะแนน
 จากการสอบที่ใช้ข้อสอบแบบเลือกตอบและมีการเดา' หนังสือการประชุมวิชาการ
 สถิติประยุกต์ครั้งที่ 9 จัดโดยคณะสถิติประยุกต์สถาบันบัณฑิตพัฒนบริหารศาสตร์
 ร่วมกับคณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่ สิงหาคม 2535. เล่มที่ 1