

ความสำคัญของการใช้ Z-score ในการปรับเทียบคะแนน

ผู้ช่วยศาสตราจารย์ ดร.วินิตา แก้วเกื้อ

บทคัดย่อ

ในระบบการประเมินผลหรือการวิจัยที่เกี่ยวข้องกับข้อมูลเชิงปริมาณ การเปรียบเทียบคะแนนจากกลุ่มตัวอย่างที่มีลักษณะแตกต่างกันในด้านค่ากลางและความแปรปรวนมักเป็นประเด็นสำคัญที่ส่งผลต่อการตีความข้อมูลอย่างถูกต้อง การใช้ ค่า Z-score เป็นหนึ่งในวิธีการทางสถิติที่มีประสิทธิภาพในการปรับเทียบคะแนนให้อยู่ในมาตรฐานเดียวกัน บทความนี้มีวัตถุประสงค์เพื่ออธิบายแนวคิดเกี่ยวกับการปรับเทียบคะแนน ความหมายของการปรับเทียบคะแนน เหตุผลในการปรับเทียบคะแนน การใช้ Z-score เพื่อเปรียบเทียบคะแนน ประโยชน์ของการใช้ Z-score และตัวอย่างการประยุกต์ใช้ Z-score ในบริบทการศึกษา เพื่อสร้างความเข้าใจที่ชัดเจนยิ่งขึ้น

แนวคิดเกี่ยวกับการปรับเทียบคะแนน

ความสำเร็จของการเปรียบเทียบคะแนนขึ้นอยู่กับแผนการรวบรวมข้อมูล ซึ่งต้องเชื่อมโยงกับวัตถุประสงค์การใช้งานของแบบทดสอบที่มีความเชื่อมโยงกัน โดยการปรับเทียบคะแนนอาจใช้วิธีการที่หลากหลาย (Angoff, 1971; Kolen และ Brennan, 2014) การเปรียบเทียบคะแนนที่ได้จากการประเมินที่แตกต่างกันหรือการจัดระดับคะแนนบนมาตรวัดเดียวกันโดยเป็นแบบวัดที่มีมาตรฐานในด้านจิตวิทยาและการศึกษา (Kolen, 2004) คำศัพท์ต่าง ๆ ที่ใช้สำหรับการปรับเทียบคะแนน เช่น การสอบเทียบ การเชื่อมโยง การวัดระดับ และการเปรียบเทียบคะแนน มักถูกใช้ในความหมายแคบหรือกว้าง ขึ้นอยู่กับมุมมองทางทฤษฎีที่นักวิจัยใช้ ทั้งนี้การเชื่อมโยงคะแนนหมายถึงการจัดระดับคะแนนของข้อสอบบนมาตรวัดเดียวกัน (Hambleton และ Swaminathan, 1985) ดังนั้น ฟังก์ชันข้อมูลของแบบทดสอบและเส้นโค้งลักษณะเฉพาะของแบบทดสอบ (test information functions and test characteristic curves) ทั้งสองแบบจึงควรมีความคล้ายคลึงกัน (McDonald, 1999)

ความหมายของการปรับเทียบคะแนน

การปรับเทียบคะแนน หรือที่เรียกว่า "การเทียบเคียงคะแนน" หรือ "การปรับคะแนน" เป็นกระบวนการที่ใช้ในหลายบริบท เช่น การสอบ การประเมินผล หรือการวิจัย เพื่อให้คะแนนมีความยุติธรรมและสามารถเปรียบเทียบกันได้อย่างถูกต้อง (Field, 2013)

เหตุผลในการปรับเทียบคะแนน

1. ความแตกต่างของข้อสอบ เนื่องจากข้อสอบแต่ละชุดที่นำมาใช้ในการสอบอาจมีความยากง่ายต่างกัน การจะนำคะแนนที่ได้จากแบบทดสอบ 2 ฉบับที่มีความยากง่ายแตกต่างกันแล้วนำคะแนนมาเปรียบเทียบกัน จึงไม่สามารถเปรียบเทียบกันได้โดยตรง

2. ความแตกต่างของผู้ประเมิน ผู้ประเมินแต่ละคนอาจมีมาตรฐานการให้คะแนนไม่เหมือนกัน แม้ว่าจะมีเกณฑ์การให้คะแนน แต่ความเป็นปัจเจกของแต่ละบุคคลย่อมส่งผลต่อการให้คะแนนแตกต่างกันออกไป โดยเฉพาะการให้คะแนนสำหรับแบบทดสอบอัตนัย

3. เพื่อความยุติธรรม การปรับคะแนนทำให้ผู้เข้าสอบจากกลุ่มต่าง ๆ ได้รับการประเมินอย่างเท่าเทียมกัน และเป็นการลดอคติหรือความไม่เที่ยงตรงจากระบบการให้คะแนนจากข้อ 1 และ 2 ข้างต้น

4. การเปรียบเทียบระหว่างกลุ่ม ผู้เกี่ยวข้องสามารถนำคะแนนที่ได้จากการปรับเทียบไปใช้ในการเปรียบเทียบผลการเรียนระหว่างโรงเรียน หรือระหว่างปีการศึกษา และใช้ในการวิจัยเพื่อวิเคราะห์ข้อมูลจากหลายแหล่ง

5. การแปลงคะแนนเป็นมาตรฐานเดียวกัน เช่น การแปลงคะแนนดิบเป็นคะแนน T-score, Z-score หรือ Percentile เพื่อใช้ในการคัดเลือกหรือจัดอันดับคะแนนของผู้เข้าสอบ

6. การควบคุมคุณภาพของการสอบ ใช้ในการตรวจสอบว่าข้อสอบมีความน่าเชื่อถือและวัดผลได้ตรงตามวัตถุประสงค์ (Gravetter & Wallnau, 2016)

การใช้ Z-score เพื่อเปรียบเทียบคะแนน

การใช้ **Z-score** เป็นวิธีหนึ่งในการปรับเทียบคะแนน เพื่อให้สามารถเปรียบเทียบคะแนนของผู้เรียนในบริบทที่ต่างกันได้อย่างเป็นธรรมและแม่นยำ โดยเฉพาะเมื่อข้อสอบมีความยากง่ายต่างกัน หรือเมื่อเปรียบเทียบคะแนนจากหลายกลุ่ม โดย **Z-score** คือ ค่าที่บอกว่า คะแนนหนึ่ง ๆ ห่างจากค่าเฉลี่ยของกลุ่มมากน้อยแค่ไหน โดยวัดเป็นจำนวน ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) (McDonald, 1999)

สูตรการคำนวณ Z-score คือ:

$$Z = \frac{X - \mu}{\sigma}$$

โดยที่: Z = ค่า Z-score, X = คะแนนดิบของผู้เรียน, μ (mu) = ค่าเฉลี่ยของคะแนนทั้งหมดในกลุ่ม
 σ (sigma) = ส่วนเบี่ยงเบนมาตรฐานของคะแนนในกลุ่ม

ความหมายของค่า Z-score

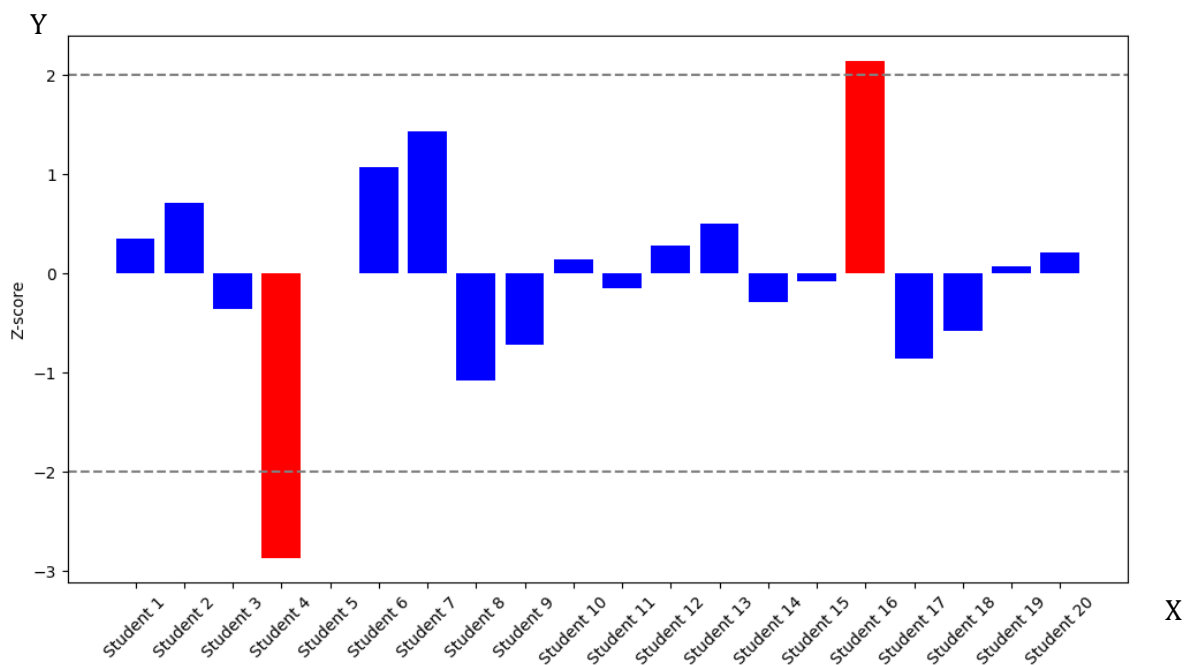
- ถ้า $Z > 0$ หมายถึง คะแนนสูงกว่าค่าเฉลี่ย
- ถ้า $Z < 0$ หมายถึง คะแนนต่ำกว่าค่าเฉลี่ย
- ถ้า $Z = 0$ หมายถึง คะแนนเท่ากับค่าเฉลี่ย
- ถ้า $Z > 2$ หรือ $Z < -2$ หมายถึง คะแนนผิดปกติ (สูงหรือต่ำกว่าค่าเฉลี่ยมาก)

ประโยชน์ของการใช้ Z-score

1. เปรียบเทียบคะแนนจากข้อสอบต่างชุดได้ แม้ข้อสอบจะมีความยากง่ายต่างกัน แต่สามารถเปรียบเทียบคะแนนได้อย่างเป็นธรรม
2. ใช้ในการคัดเลือกหรือจัดอันดับ เช่น การคัดเลือกนักเรียนเข้าศึกษาต่อ โดยใช้ Z-score แทนคะแนนดิบ
3. ใช้ในการสร้างเกณฑ์การตัดสิน เช่น การกำหนดการให้เกรด
4. ใช้ในการวิเคราะห์ข้อมูลทางสถิติ เช่น ตรวจสอบว่าคะแนนใดเป็น "คะแนนที่ผิดปกติ" (outlier) (Field, 2013)

ตัวอย่างการประยุกต์ใช้ Z-score ในบริบทการศึกษา

จากการคำนวณ Z-score จากคะแนนของนักเรียน 20 คน พบว่ามี 2 คน ที่มีคะแนนผิดปกติ (outliers) ตามเกณฑ์ Z-score ที่มากกว่า 2 หรือ น้อยกว่า -2 แสดงดังภาพที่ 1



ภาพที่ 1 กราฟแสดง Z-score ของคะแนนนักเรียน 20 คน

วิธีการอ่านกราฟ **Z-score** ช่วยให้เราเข้าใจได้ว่าแต่ละคะแนนของนักเรียนอยู่ในระดับใดเมื่อเทียบกับกลุ่มทั้งหมด โดยมีหลักการอ่านง่าย ๆ ดังนี้

ส่วนประกอบของกราฟ

1. แกน X (แนวนอน): แสดงลำดับของนักเรียน (เช่น นักเรียนคนที่ 1 ถึง 20)
2. แกน Y (แนวตั้ง): แสดงค่า **Z-score** ของคะแนนแต่ละคน
3. แถบสีฟ้า: คะแนนที่อยู่ในช่วงปกติ (Z-score ระหว่าง -2 ถึง 2)
4. แถบสีแดง: คะแนนที่ผิดปกติ (Z-score มากกว่า 2 หรือ น้อยกว่า -2)
5. เส้นประสีเทา: เส้นเกณฑ์ที่ $Z = 2$ และ $Z = -2$ ใช้เป็นจุดตัดสินว่าคะแนนใดผิดปกติ

วิธีอ่านและตีความ

ค่า Z-score	ความหมาย	ตัวอย่าง
ใกล้ 0	คะแนนใกล้เคียงค่าเฉลี่ยของกลุ่ม	นักเรียนทั่วไป
> 2	คะแนนสูงผิดปกติ	อาจเป็นนักเรียนที่มีความสามารถโดดเด่น
< -2	คะแนนต่ำผิดปกติ	อาจต้องการความช่วยเหลือเพิ่มเติม

การแปลงข้อมูลจากกราฟ

นักเรียนคนที่ 4 มีค่า Z-score = -2.67 (แถบสีแดง) แสดงว่ามีคะแนนต่ำกว่ากลุ่มมาก และนักเรียนคนที่ 16 มีค่า Z-score = 2.34 (แถบสีแดง) แสดงว่ามีคะแนนสูงกว่ากลุ่มมาก

ข้อเสนอแนะ

1. นักเรียนที่มีคะแนนต่ำผิดปกติควรได้รับการวิเคราะห์เพิ่มเติม เช่น การสัมภาษณ์เชิงลึก หรือ การตรวจสอบปัจจัยด้านการเรียนรู้
2. การใช้ Z-score ช่วยให้เราสามารถเปรียบเทียบคะแนนในบริบทต่าง ๆ ได้อย่างเป็นธรรม เช่น ระหว่างชั้นเรียน หรือระหว่างปีการศึกษา
3. ข้อมูลนี้สามารถนำไปใช้ในการปรับปรุงการออกแบบข้อสอบ การจัดกิจกรรมเสริม หรือการให้คำปรึกษาเฉพาะบุคคล

บรรณานุกรม

- Angoff, W. A. (1971). "Scales, norms, and equivalent scores" in *Educational Measurement*, ed. R. L. Thorndike (Washington, DC: American Council on Education), 508–600.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
- Gravetter, F. J., & Wallnau, L. B. (2016). *Statistics for the Behavioral Sciences*. Cengage Learning.
- Hambleton, R. K., and Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Leiden: Kluwer-Nijhoff. doi: 10.1007/978-94-017-1988-9
- Kolen, M. J. (2004). Linking assessments: concept and history. *Appl. Psychol. Meas.* 28, 219–226. doi: 10.1177/0146621604265030
- Kolen, M. J., and Brennan, R. L. (2014). *Test Equating, Scaling, and Linking: Methods and Practices*. Berlin: Springer Science and Business Media. doi: 10.1007/978-1-4939-0317-7
- McDonald, R. P. (1999). *Test Theory: a Unified Treatment*. New Jersey, NJ: Lawrence Erlbaum Associates.

