

Z-test และ t-test

ดร.มนตรี อนันตรักษ์*

ในการทำวิจัยหรือการสอนวิชาสถิติสำหรับการวิจัย มักจะพบคำถามว่า Z-test กับ t-test แตกต่างกันอย่างไร จะใช้ในเงื่อนไขหรือในกรณีใด บทความนี้จะชี้แจงเฉพาะการใช้ Z-test และ t-test กรณีทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ย

จากทฤษฎีบทขอบเขตกลาง (Central Limit Theorem) กล่าวว่าถ้าประชากรมีขนาดใหญ่มาก ไม่จำกัด ซึ่งมีการกระจายเป็นโค้งปกติ มีค่าเฉลี่ยของประชากร เท่ากับ μ มีความแปรปรวนของประชากรเท่ากับ σ^2 ถ้าสุ่มตัวอย่างมาขนาด n มาจากประชากร (สุ่มทุก ๆ กรณี) และมีค่าเฉลี่ยของกลุ่มตัวอย่างเท่ากับ $\bar{X}$ มีความแปรปรวนของกลุ่มตัวอย่างเท่ากับ S^2 แล้ว

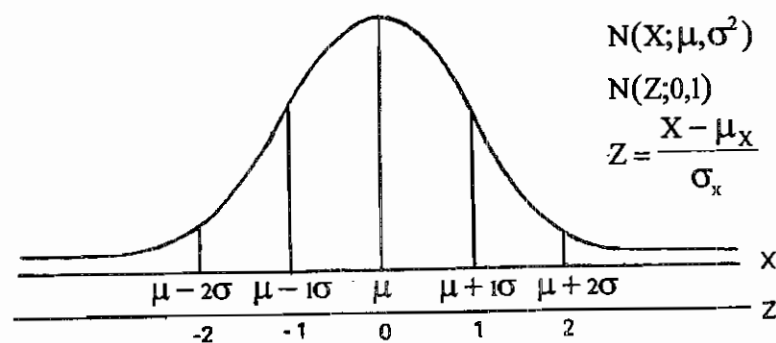
$$1. \mu_{\bar{X}} = \mu$$

(ค่าเฉลี่ยของค่าเฉลี่ยจากกลุ่มตัวอย่าง เท่ากับ ค่าเฉลี่ยของประชากร)

$$2. \sigma_{\bar{X}}^2 = \sigma^2 / n$$

(ความแปรปรวนของค่าเฉลี่ยจากกลุ่มตัวอย่าง เท่ากับ ความแปรปรวนของประชากรหารด้วยขนาดของกลุ่มตัวอย่าง)

แนวคิดของ Z-test มีดังนี้



*อาจารย์สังกัดภาควิชาวิจัยและพัฒนาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม

1. กรณีทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากรเท่ากับจำนวนจริง (Real Number)

$$H_0 : \mu = a \quad (a \text{ คือ จำนวนจริง})$$

$$H_1 : \mu \neq a$$

เช่น $H_0 : \mu = 120$

$$H_1 : \mu \neq 120$$

จาก
$$Z = \frac{X - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$$

ดังนั้น
$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$$

แต่
$$\begin{aligned} \mu_{\bar{X}} &= \mu \\ \sigma_{\bar{X}}^2 &= \sigma^2 / n \\ \sigma_{\bar{X}} &= \sigma / \sqrt{n} \end{aligned}$$

$\therefore$
$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

----- 1

หรือ
$$Z = \frac{\bar{X} - a}{\sigma / \sqrt{n}}$$

2. กรณีทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยของประชากร 2 ค่า

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

จาก
$$Z = \frac{X - \mu_x}{\sigma_x}$$

ดังนั้น
$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \mu_{(\bar{X}_1 - \bar{X}_2)}}{\sigma_{(\bar{X}_1 - \bar{X}_2)}}$$

แต่
$$\mu_{(\bar{X}_1 - \bar{X}_2)} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2 = 0 \quad (\text{ตาม } H_0)$$

และ $\sigma_{(X_1 - X_2)}^2 = \sigma_1^2 / n_1 + \sigma_2^2 / n_2$ (กรณีประชากรเป็นอิสระกัน)

$$\therefore \sigma_{(X_1 - X_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

ดังนั้น
$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{-----} \quad (2)$$

ดังนั้นการใช้ Z-test ในการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ย จะต้องทราบความแปรปรวนของประชากร (σ^2)

แต่ในทางปฏิบัติหรือในการทำวิจัย มักจะไม่ทราบค่าความแปรปรวนของประชากร (σ^2) แต่สามารถหาค่าความแปรปรวนของกลุ่มตัวอย่าง (S^2) ได้ จึงนำความแปรปรวนของกลุ่มตัวอย่างมาใช้แทนความแปรปรวนของประชากร ซึ่งเป็นที่มาขอ t-test

3. กรณีทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของประชากร เท่ากับจำนวนจริง (Real Number)

$$H_0 : \mu = a$$

$$H_1 : \mu \neq a$$

กรณีไม่ทราบค่าความแปรปรวนของประชากร (σ_X^2)

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}} \quad \text{-----} \quad (3)$$

หรือ

$$t = \frac{\bar{X} - a}{S / \sqrt{n}}$$

S^2 = ความแปรปรวนของกลุ่มตัวอย่าง

S = ส่วนเบี่ยงเบนมาตรฐานของกลุ่มตัวอย่าง

$$= \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}}$$

4. กรณีทดสอบสมมติฐานเกี่ยวกับความแตกต่างของค่าเฉลี่ยของประชากร 2 ค่า

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

4.1 กรณีกลุ่มตัวอย่างเป็นอิสระกัน (Independent Samples)

จาก
$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{(\bar{X}_1 - \bar{X}_2)}}$$

ดังนั้น
$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{(\bar{X}_1 - \bar{X}_2)}} \quad \text{-----} \quad (4)$$

แต่
$$S_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

ซึ่งนิยมเรียกว่า Pooled Variance เมื่อ $\sigma_1^2 = \sigma_2^2$

และ
$$S_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

ซึ่งนิยมเรียกว่า Separate Variance เมื่อ $\sigma_1^2 \neq \sigma_2^2$

ดังนั้น ก่อนใช้ t-test ทดสอบ (นิยมเรียก t-independent) สมมติฐาน ($H_0 : \mu_1 = \mu_2$) ต้องทดสอบว่าความแปรปรวนของประชากรทั้ง 2 กลุ่มเท่ากันหรือไม่ นั่นคือทดสอบ

$$H_0 : \sigma_1^2 = \sigma_2^2$$

โดยใช้
$$F = \frac{S_{Max}^2}{S_{Min}^2}$$

S_{Max}^2 = ความแปรปรวนของกลุ่มตัวอย่างที่มีค่ามาก

S_{Min}^2 = ความแปรปรวนของกลุ่มตัวอย่างที่มีค่าน้อย

ถ้า Reject. H_0 แสดงว่า $\sigma_1^2 \neq \sigma_2^2$ ให้ใช้ t-test กรณี Separate Variance กล่าวคือ

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \text{-----} (5)$$

แต่ถ้า Accept. H_0 แสดงว่า $\sigma_1^2 = \sigma_2^2$ ให้ใช้ t-test กรณี Pooled Variance กล่าวคือ

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \text{-----} (6)$$

4.2 กรณีกลุ่มตัวอย่างไม่อิสระ (Dependent Sample) เช่น ประชากรกลุ่มเดียวมีค่าเฉลี่ย 2 ค่า คือ μ_1 กับ μ_2

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{(\bar{X}_1 - \bar{X}_2)}}$$

$$t = \frac{\bar{d}}{S_d / \sqrt{n}} \text{-----} (7)$$

$$d = X_1 - X_2$$

นิยมเรียกว่า t-dependent

ดังนั้น สรุปได้ว่า การใช้ Z-test ทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ย ต้องทราบค่าความแปรปรวนของประชากร (σ^2) ถ้านำค่าความแปรปรวนของกลุ่มตัวอย่าง (S^2) มาใช้แทน ความแปรปรวนของประชากร จะใช้ t-test

เอกสารอ้างอิง

Glass, Gene V and Stanley, Jullan C. **Statistical Methods in Education and Psychology**. 1970.