

การแปลงฟิชเชอร์ z' และการเปรียบเทียบความแตกต่างระหว่างค่าสหสัมพันธ์

ฉัตรศิริ ปิยะพิมลสิทธิ์*

ในการทดสอบนัยสำคัญของค่าสหสัมพันธ์ว่ามีค่าเป็นศูนย์หรือไม่นั้น เราจะใช้สูตร t-test ว่า

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} ; df = n - 2$$

แต่ในกรณีที่ต้องการเปรียบเทียบค่าสหสัมพันธ์ของกลุ่มตัวอย่างกับประชากรนั้น โดยปกติแล้วการแจกแจงของค่าสหสัมพันธ์นั้นมีลักษณะการแจกแจงแบบเบ้ ดังนั้น R. A. Fisher จึงได้พัฒนาการแปลงค่าสหสัมพันธ์ r เป็น z' ซึ่งจะมีการแจกแจงใกล้เคียงกับโค้งปกติและมีค่าความคลาดเคลื่อนมาตรฐานขึ้นอยู่กับกลุ่มตัวอย่าง

$$z' = \frac{1}{2} [\ln(1+r) - \ln(1-r)]$$

เมื่อ \ln คือลอการิทึมฐานธรรมชาติ ซึ่งในตาราง Fisher's Z จะแสดงการแปลงค่า r เป็น z' โดยไม่จำเป็นต้องคำนวณเอง ส่วนความคลาดเคลื่อนมาตรฐาน z' ก็คือ

$$sd_{z'} = \frac{1}{\sqrt{n-3}}$$

ดังนั้นในการทดสอบสมมติฐานระหว่างค่าสหสัมพันธ์ที่คำนวณจากกลุ่มตัวอย่างกับค่าสหสัมพันธ์ที่คำนวณจากประชากร มีสูตรว่า

$$z = \frac{z'_s - z'_h}{sd_{z'}}$$

เมื่อ z คือคะแนนมาตรฐานภายใต้โค้งปกติ และ z'_s คือค่าสหสัมพันธ์ของกลุ่มตัวอย่างที่แปลงเป็นฟิชเชอร์ z' และ z'_h คือค่าสหสัมพันธ์ของประชากรที่แปลงเป็นฟิชเชอร์ z' นำค่าที่คำนวณได้ไปเปิดตารางการแจกแจงปกติ จะได้ค่าระดับนัยสำคัญแบบทิศทางเดียว

ตัวอย่างในการคำนวณ

ในการศึกษาสหสัมพันธ์ระหว่างการศึกษากับรายได้ของกลุ่มตัวอย่างจำนวน 103 คน ได้ค่าสหสัมพันธ์ 0.47 และสมมติฐานที่เป็นกลาง ก็คือค่าสหสัมพันธ์ระหว่างสองตัวแปรนี้ของกลุ่มประชากรเท่ากับ 0.63 ไปเปิดตารางแปลงค่า r เป็นฟิชเชอร์ z' ได้ค่าเท่ากับ 0.510 และ 0.741 ตามลำดับ คำนวณความคลาดเคลื่อนมาตรฐานได้ค่า

$$sd_{z'} = \frac{1}{\sqrt{103-3}} = 0.10$$

* กศ.ม. (การวัดผลการศึกษา) <http://www.watpon.com>

และคำนวณค่า z ได้ดังนี้

$$z = \frac{0.510 - 0.741}{0.10} = -2.31$$

เปิดตารางการแจกแจงปกติในคอลัมน์ C เราจะพบ $z = 2.31$ จะให้ค่าระดับนัยสำคัญแบบทิศทางเดียวเท่ากับ 0.0104 ถ้าหากเป็นระดับนัยสำคัญแบบสองทิศทางจะได้ค่า 0.0208 ซึ่งมีนัยสำคัญทางสถิติที่ 0.05 (เพราะระดับนัยสำคัญของค่าที่คำนวณได้น้อยกว่า $\alpha = 0.05$) นั่นคือความสัมพันธ์ระหว่างการศึกษากับรายได้ของกลุ่มตัวอย่างมีค่าสูงกว่ากลุ่มประชากร

ทำนองเดียวกับการทดสอบความแตกต่างระหว่างค่าสหสัมพันธ์ 2 ค่าที่มาจากกลุ่มตัวอย่าง มีสูตรคำนวณคือ

$$z = \frac{z'_1 - z'_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

ตัวอย่างในการคำนวณ

จะใช้ค่าเดียวกับตัวอย่างที่ผ่านมาโดยสมมติว่ากลุ่มตัวอย่าง 2 กลุ่ม กลุ่มละ 103 คน คำนวณค่าสหสัมพันธ์ระหว่างการศึกษากับรายได้ มีค่าสหสัมพันธ์เท่ากับ 0.47 และ 0.63 เราจะเปรียบเทียบค่าสหสัมพันธ์สองค่านี้ โดยเราต้องแปลงค่า r ไปเป็นพิชเซอร์ z' แล้วแทนค่าในสูตรได้ดังนี้

$$\begin{aligned} z &= \frac{0.510 - 0.741}{\sqrt{\frac{1}{103 - 3} + \frac{1}{103 - 3}}} \\ &= \frac{0.510 - 0.741}{\sqrt{0.01 + 0.01}} \\ &= \frac{0.510 - 0.741}{0.1414} \\ &= -1.63 \end{aligned}$$

เปิดตารางการแจกแจงปกติในคอลัมน์ C พบว่ามีค่าระดับนัยสำคัญแบบทิศทางเดียวเท่ากับ 0.0516 ถ้าหากเป็นระดับนัยสำคัญแบบสองทิศทางมีค่าประมาณ 0.10 ซึ่งไม่มีนัยสำคัญทางสถิติ ซึ่งจะแตกต่างจากตัวอย่างข้างต้น ทั้งนี้เพราะตัวหามีจำนวนมากกว่า ทำให้ค่า z ที่คำนวณได้มีค่าลดลง

ในการแปลงค่า r เป็นพิชเซอร์ z' นั้น สามารถใช้ได้กับการทดสอบค่าสหสัมพันธ์ที่มากกว่า 2 ค่า โดยใช้สถิติไคสแควร์ที่มี $df = k - 1$ ซึ่ง k ก็คือจำนวนของสัมประสิทธิ์สหสัมพันธ์ที่เปรียบเทียบ มีสูตรว่า

$$\chi^2 = \sum (n_i - 3)z_i^2 - \frac{[\sum (n_i - 3)z_i]^2}{\sum (n_i - 3)}$$

ถ้าหากพบว่าค่า χ^2 มีนัยสำคัญทางสถิติ แสดงว่ามีค่าสหสัมพันธ์อย่างน้อย 1 คู่ที่แตกต่างกัน ถ้าต้องการทราบว่าคูไหนที่แตกต่างกันให้ทดสอบเป็นรายคู่ด้วย z ในสูตรข้างต้น

ในการทดสอบความแตกต่างระหว่างค่าสหสัมพันธ์ที่กล่าวข้างต้นนี้ ควรจะอยู่บนข้อตกลงเบื้องต้นที่ว่า ค่าสหสัมพันธ์แต่ละค่าที่นำมาเปรียบเทียบควรคำนวณจากกลุ่มตัวอย่างที่เป็นอิสระจากกัน และความแปรปรวนของกลุ่มประชากรควรมีความเป็นเอกพันธ์กัน

การทดสอบความแตกต่างระหว่างค่าสหสัมพันธ์ที่ไม่เป็นอิสระจากกัน

ในหัวข้อนี้จะกล่าวถึงกรณีที่มีตัวแปร 3 ตัว คือ V , X และ Y ที่วัดจากกลุ่มตัวอย่างกลุ่มเดียวกัน แล้วคำนวณหาสหสัมพันธ์ r_{VY} และ r_{XY} ซึ่งจะไม่เหมาะสมหากใช้วิธีแปลงค่าสหสัมพันธ์นี้เป็นพิชเซอร์ z' เพราะค่าสัมประสิทธิ์สหสัมพันธ์ไม่ได้คำนวณมาจากกลุ่มตัวอย่างที่เป็นอิสระจากกัน ฉะนั้นในการทดสอบกรณีที่กลุ่มตัวอย่างสัมพันธ์กันนั้นเราจะใช้สูตร t -test ที่มี $df = n - 3$ ดังนี้

$$t = \frac{(r_{XY} - r_{VY})\sqrt{(n-1)(1+r_{XV})}}{\sqrt{2\left(\frac{n-1}{n-3}\right)|\mathbf{R}| + \bar{r}^2(1-r_{XV})^3}}$$

$$\text{เมื่อ } \bar{r} = \frac{r_{XY} + r_{VY}}{2}$$

และ $|\mathbf{R}| = 1 - r_{XY}^2 - r_{VY}^2 - r_{XV}^2 + 2r_{XY}r_{VY}r_{XV}$ หรือก็คือค่าดีเทอร์มิแนนของเมตริกซ์สหสัมพันธ์ระหว่าง X , Y และ Z นั้นเอง

ตัวอย่างในการคำนวณ

ในการตรวจสอบสหสัมพันธ์ระหว่างเจตคติต่อการอยู่อาศัย (Y) ความกระตือรือร้น (V) และความสัมพันธ์กับเด็ก (X) กับกลุ่มตัวอย่าง 147 คน เราได้ค่าสหสัมพันธ์ดังนี้

$r_{XY} = 0.39$, $r_{VY} = 0.19$ และ $r_{XV} = 0.21$ และคำนวณ

$$\bar{r} = \frac{0.39 + 0.19}{2} = 0.29$$

$$|\mathbf{R}| = 1 - 0.39^2 - 0.19^2 - 0.21^2 + 2(0.39)(0.19)(0.21) = 0.7988$$

ดังนั้น

$$t = \frac{(0.39 - 0.19)\sqrt{(147-1)(1+0.21)}}{\sqrt{2\left(\frac{147-1}{147-3}\right)(0.7988) + 0.29^2(1-0.21)^3}} = \frac{2.6308}{1.2728} = 2.067$$

มี $df = n - 3 = 144$ มีนัยสำคัญทางสถิติที่ระดับ 0.05



บรรณานุกรม

Cohen, Jacob and Cohen, Patricia. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Second Edition. U.S.A. :
Lawrence Erlbaum Associates, Inc.