บทที่ 2

ทฤษฎีและสถิติทดสอบที่เกี่ยวข้องในการวิจัย

การศึกษาเกี่ยวกับความสัมพันธ์ระหว่างกลุ่ม เป็นการศึกษาความสัมพันธ์ระหว่างกลุ่มตัวแปรที่มีการวัดซ้ำ และกลุ่มตัวแปรที่ไม่มีการวัดซ้ำ ในทางปฏิบัติมักต้องการตรวจสอบดูว่ามีความสัมพันธ์ระหว่างกลุ่มหรือไม่ การวิจัยนี้ศึกษาเปรียบเทียบการทดสอบความสัมพันธ์ระหว่างกลุ่ม 3 วิธีคือ

1. The Adjusted Pairwise Test
2. The Sib-Mean Test
3. The Ensemble Test

เนื่องจากประมาณการแจกแจงของตัวสถิติทดสอบได้ด้วยการแจกแจงปกติมาตรฐาน (Standard Normal distribution) และการแจกแจงที (Student's t distribution) และจำนวนข้อมูลที่เก็บซ้ำของแต่ละตัวอย่างมีการแจกแจงนิเสธทวินาม (Negative Binomial distribution) ส่วนข้อมูลมีการแจกแจงปกติพหุ(Multivariate Normal distribution) ในบทนี้จะกล่าวถึงการแจกแจงปกติมาตรฐาน การแจกแจงที การแจกแจงนิเสธทวินาม และการแจกแจงปกติพหุ ทฤษฎีทางสถิติที่ใช้ในการวิจัย และรายละเอียดของการทดสอบแต่ละวิธี พร้อมทั้งตัวอย่างของการคำนวณ มีรายละเอียดดังนี้

2.1 <u>การแจกแจงที่เกี่ยวข้องในการศึกษา</u>

2.1.1 <u>การแจกแจงปกติพหุ</u> (Multivariate Normal Distribution)

การแจกแจงความน่าจะเป็นของตัวแปรสุ่มที่สำคัญ และมีการใช้มากที่สุดในวิชาสถิติคือ การแจกแจงปกติ ซึ่งเป็นการแจกแจงแบบต่อเนื่องอย่างหนึ่ง ลักษณะที่เห็นเด่นชัดคือ มีลักษณะเป็นรูปโค้งสมมาตรคล้ายระฆัง (Bell-shaped) ซึ่งเรียกว่า โค้งปกติ (Normal curve) ลักษณะการแจกแจงแบบนี้ใช้อธิบายข้อมูลที่เกิดขึ้นทั่วไปในธรรมชาติ เช่น ความสูง น้ำหนัก และคะแนนการทดสอบทางปัญญา เป็นต้น

ถ้า χ เป็นตัวแปรสุ่มที่มีการแจกแจงปกติพหุ

$$\underset{p\times 1}{\chi} \sim N_p(\underset{\sim}{\mu}, \Sigma)$$

ฟังก์ชั่นความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงปกติพหุ มีรูปแบบดังนี้

$$f(\chi) = \{(2\pi)^p|\Sigma|\}^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\chi-\underset{\sim}{\mu})' \Sigma^{-1} (\chi-\underset{\sim}{\mu})\} \quad ; \quad \chi \varepsilon E_p$$

เมื่อ χ คือ เวกเตอร์ตัวแปรสุ่มขนาด pxl

$\underset{\sim}{\mu}$ คือ เวกเตอร์ค่าเฉลี่ยขนาด pxl

Σ คือ เมตริกซ์ความแปรปรวนขนาด pxp และ rank $(\Sigma) = p$

E_p คือ เซตของเวกเตอร์ตัวแปรสุ่มที่มีขนาด pxl

$$\chi = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_p \end{bmatrix}_{p\times 1} , \quad \underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_p \end{bmatrix}_{p\times 1}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2p} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \sigma_{p1} & \sigma_{p2} & \sigma_{p3} & \cdots & \sigma_p^2 \end{bmatrix}_{p\times p}$$

โดยที่ μ_1 , μ_2, ... μ_p เป็นค่าเฉลี่ยของตัวแปร $X_1, X_2, \ldots, X_p$ ตามลำดับ

σ_1^2 , σ_2^2, ... σ_p^2 เป็นค่าความแปรปรวนตัวแปรของ $X_1, X_2, \ldots X_p$ ตามลำดับ

2.1.2 <u>การแจกแจงปกติมาตรฐาน</u> (Standard Normal Distribution)

การแจกแจงปกติมาตรฐาน (Standard Normal distribution) คือการแจกแจงปกติที่มีค่าเฉลี่ย (μ) เท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐาน (σ) เท่ากับ 1 ใช้สัญลักษณ์ N(0,1) การเปลี่ยนการแจกแจงปกติทั่วไปให้เป็นการแจกแจงปกติมาตรฐาน สามารถทำได้โดยใช้สูตรแปลงค่าดังนี้

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

ฟังก์ชั่นความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงปกติมาตรฐาน มีรูปแบบดังนี้

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} \quad ; \quad -\infty < Z < \infty$$

และเนื่องจากการแจกแจงปกติมาตรฐาน เป็นกรณีเฉพาะของการแจกแจงปกติ จึงมีคุณสมบัติเหมือนการแจกแจงปกติ

2.1.3 <u>การแจกแจงที</u> (Student's t Distribution)

การแจกแจงทีจะมีลักษณะใกล้เคียงกับการแจกแจงปกติ แต่มีลักษณะหางยาวอย่างเห็นได้ชัด โดยข้อมูลที่มีค่ามากและน้อยซึ่งแตกต่างจากข้อมูลส่วนใหญ่จะมีจำนวนเท่า ๆ กัน

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงที

ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงที มีรูปแบบดังนี้

$$f(X) = \frac{\left(1 + \frac{X^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}}{\sqrt{n}\ B\left(\frac{1}{2},\frac{n}{2}\right)} \quad ; -\infty < X < \infty \quad ; \quad n = 1,2,3,...$$

โดยที่ $B(a,b)$ เป็นเบต้าฟังก์ชั่นและ $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$

การแจกแจงทีมีความสัมพันธ์กับการแจกแจงปกติมาตรฐานและการแจกแจงไคสแควร์ (Chi-Square distribution) ดังนี้

$$t(n) = \frac{Z}{\sqrt{\frac{\chi^2(n)}{n}}}$$

; n = องศาความเป็นอิสระ (degree of freedom)

$\chi^2(n)$ เป็นตัวแปรสุ่มที่มีการแจกแจงไคสแควร์ มีองศาความเป็นอิสระ $n-1$

ลักษณะการแจกแจงทีจะแปรเปลี่ยนตามขนาดตัวอย่าง และเมื่อขนาดตัวอย่างมากขึ้นการแจกแจงทีจะมีลักษณะเข้าใกล้การแจกแจงปกติ

2.1.4 <u>การแจกแจงนิเสธทวินาม</u> (Negative Binomial Distribution)

ถ้า X มีการแจกแจงนิเสธทวินาม

$$X \sim NB\ (k,p)$$

ฟังก์ชั่นความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงนิเสธทวินาม มีรูปแบบดังนี้

$$p(X) = \binom{k+X-1}{X} p^k q^X \quad ; \quad X = 0,1,2,...$$

เมื่อ k คือ จำนวนครั้งความสำเร็จในการทดลอง k+x ครั้ง

x คือ จำนวนครั้งความล้มเหลวก่อนประสบความสำเร็จครั้งที่ k

p คือ ความน่าจะเป็นที่ประสบความสำเร็จ

$p = 1-q$

ตัวแปรสุ่ม X มีค่าเฉลี่ยเท่ากับ $\frac{kq}{p}$ และมีความแปรปรวนเท่ากับ $\frac{kq}{p^2}$

2.2 ทฤษฎีทางสถิติที่ใช้ในการวิจัย

2.2.1 การแจกแจงแบบมีเงื่อนไข (Conditional Distribution)

การแจกแจงแบบมีเงื่อนไขเป็นการแจกแจงที่มีข้อกำหนดเพิ่มเติม มีทฤษฎีที่สำคัญดังนี้

เมื่อ $\underset{\sim}{X}$ เป็นตัวแปรสุ่มที่มีการแจกแจงปกติพหุ

$$\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma) \text{ และ } \operatorname{rank}(\Sigma) = p$$

$$\underset{\sim}{X} = \begin{bmatrix} \underset{\sim}{X}_1 \\ \underset{\sim}{X}_2 \end{bmatrix}_{p\times 1}, \quad \underset{\sim}{\mu} = \begin{bmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{bmatrix}_{p\times 1}$$

$\underset{\sim}{X}_1$ และ $\underset{\sim}{X}_2$ เป็นเวกเตอร์ขนาด $q\times 1$ และ $(p-q)\times 1$ ตามลำดับ

$\underset{\sim}{\mu}_1$ และ $\underset{\sim}{\mu}_2$ เป็นเวกเตอร์ขนาด $q\times 1$ และ $(p-q)\times 1$ ตามลำดับ

$$\Sigma_{\alpha} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}_{pxq}$$

Σ_{11}, Σ_{12} , Σ_{21} และ Σ_{22} เป็นเมตริกซ์ขนาด qxq, qx(p-q) , (p-q)xq และ (p-q)x(p-q) ตามลำดับ

การแจกแจงแบบมีเงื่อนไขของ χ_1 เมื่อกำหนด $\chi_2 = C_2$ จะเป็นการแจกแจงปกติที่มีค่าเฉลี่ย $\underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma^{-1}_{22}(C_2 - \underset{\sim}{\mu}_2)$ และมีความแปรปรวน $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12}\Sigma^{-1}_{22}\Sigma_{21}$ เมื่อ C_2 เป็นเวกเตอร์ค่าคงที่ขนาด (p-q)x1 ; $0<q<p$

2.2.2 <u>การทดสอบค่าสัมประสิทธิ์สหสัมพันธ์</u>

ในการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) เมื่อตั้งสมมติฐานว่าง $H_o : \rho = 0$ จะพบว่าการกระจายของค่า r จะมีลักษณะใกล้เคียงกับการแจกแจงที โดยมีค่าเฉลี่ยเท่ากับ ρ ความแปรปรวนเท่ากับ $\frac{1 - r^2}{n - 2}$ ตัวสถิติที่ใช้ในการทดสอบคือ

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

เมื่อ r คือ ค่าสัมประสิทธิ์สหสัมพันธ์จากตัวอย่าง

n คือ จำนวนคู่ของตัวอย่าง

t มีการแจกแจงที ที่มีองศาความเป็นอิสระ n-2

2.3 วิธีการทดสอบที่ใช้ในการศึกษา

การวิจัยนี้ศึกษาเปรียบเทียบตัวสถิติทดสอบที่ใช้ทดสอบความสัมพันธ์ระหว่างกลุ่มของการทดลอบ 3 วิธีดังนี้

1. The Adjusted Pairwise Test
2. The Sib-Mean Test
3. The Ensemble Test

2.3.1 The Adjusted Pairwise Test

การทดสอบนี้ใช้ตัวสถิติทดสอบ PAIR โดยสร้างจากตัวประมาณ Pairwise (Pairwise Estimator ; $\hat{\rho}_{ms,p}$)

$$\hat{\rho}_{ms,p} = \sum_{\alpha=1}^{N} (X_{o\alpha} - \tilde{X}_m) \sum_{i=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s) \Big/ \left[\sum_{\alpha=1}^{N} k_\alpha (X_{o\alpha} - \tilde{X}_m)^2 \sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)^2 \right]^{\frac{1}{2}}$$

$$\tilde{X}_m = \sum_{\alpha=1}^{N} k_\alpha X_{o\alpha} \Big/ \sum_{\alpha=1}^{N} k_\alpha \quad , \quad \tilde{X}_s = \sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} X_{i\alpha} \Big/ \sum_{\alpha=1}^{N} k_\alpha$$

$\hat{\rho}_{ms,p}$ มีรูปแบบเหมือนตัวประมาณสัมประสิทธิ์สหสัมพันธ์เชิงเส้นของ Pearson แต่ $\tilde{X}_s$ เป็นค่าเฉลี่ยที่คิดจากทุกค่าที่วัดซ้ำของทุกตัวอย่าง และ $\sum_{i=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)$ เป็นการหาผลรวมของผลต่างแต่ละค่าที่วัดซ้ำกับ $\tilde{X}_s$ ของแต่ละตัวอย่าง

$$\text{PAIR} = (D - 2)^{\frac{1}{2}} \hat{\rho}_{ms,p} \Big/ (1 - \hat{\rho}^2_{ms,p})^{\frac{1}{2}}$$

เมื่อ $D = \sum_{\alpha=1}^{N} \left[k_\alpha / (1 + (k_\alpha - 1) \hat{\rho}_{ss}) \right]$

$$\hat{\rho}_{ss} = \left[\sum_{\alpha=1}^{N} \sum_{i \neq j}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)(X_{j\alpha} - \tilde{X}_s) \Big/ \sum_{\alpha=1}^{N} k_\alpha (k_\alpha - 1) \right] \Big/ \left[\sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)^2 \Big/ \sum_{\alpha=1}^{N} k_\alpha \right]$$

เกณฑ์การตัดสินใจคือปฏิเสธสมมติฐานหลัก $H_o : \rho_{ms} = 0$ เทียบกับสมมติฐานแย้ง $H_1 : \rho_{ms} > 0$ เมื่อ PAIR > $t_{N-2,\alpha}$ รูปแบบของตัวสถิติทดสอบ PAIR เหมือนตัวสถิติ t ที่ใช้ในการทดสอบค่าสัมประสิทธิ์สหสัมพันธ์ที่กล่าวไว้ข้างต้น

<u>ตัวอย่างวิธีคำนวณค่าสถิติทดสอบ</u>

ข้อมูลในตารางที่ 2.1 เป็นข้อมูลที่สร้างขึ้นจากการทดลองโดยมีขนาดตัวอย่าง

ขนาดตัวอย่าง $(N) = 10$

ความสัมพันธ์ภายในกลุ่ม $(\rho_{ms}) = 0$

ความสัมพันธ์ระหว่างกลุ่ม $(\rho_{ss}) = 0.3$

ค่าเฉลี่ยของจำนวนการวัดซ้ำ $(kq/p) = 3$

<u>ตารางที่ 2.1</u> แสดงรายละเอียดของข้อมูลที่สร้างขึ้นจากการทดลองเมื่อกำหนด $N = 10$, $\rho_{ms} = 0$, $\rho_{ss} = .3$, $kq/p = 3$

ตัวอย่าง (α)	จำนวนครั้งการวัดซ้ำ (k_α)	$X_{o\alpha}$	$X_{1\alpha}$, $X_{2\alpha}$,..., $X_{k\alpha,\alpha}$
1	2	-1.7674	0.0942 , -0.2379
2	8	2.3711	0.9984 , -0.0880, 0.8767 , -0.7712,
			0.3195, -0.3002, 0.1322 , -0.3331
3	1	-1.3694	-0.9447
4	4	0.0126	1.1274, -0.2324, 0.5206 , 1,1879
5	4	0.2691	0.1487, 0.5566, 0.6591 , 0.3107
6	1	1.0255	0.1105
7	8	-0.9520	0.7146, 1.2333,-0.8136 , 1.4104
			0.9846, 0.9858, 0.1687 , 1.3440
8	2	1.0520	1.7809, -1.2124
9	7	1.0712	0.1439, 0.1360, 1.9462 , 1.4834
			0.4266, 0.7212, 1.1309
10	2	0.0078	-3.5161, -0.7768

จากข้อมูลในตารางที่ 2.1 สามารถนำมาคำนวณหาค่าสถิติทดสอบ PAIR ได้ดังนี้

$$\sum_{\alpha=1}^{N} k_\alpha X_{o\alpha} = 2(-1.7674) + 8(2.3711) + \ldots + 2(0.0078)$$

$$= 18.2183$$

$$\sum_{\alpha=1}^{N} k_\alpha = 2 + 8 + \ldots + 2 = 39$$

$$\tilde{X}_m = \sum_{\alpha=1}^{N} k_\alpha X_{o\alpha} / \sum_{\alpha=1}^{N} k_\alpha = \frac{18.2183}{39} = .4671$$

$$\sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} X_{i\alpha} = 0.0942 - 0.2379 + 0.9984 + \ldots -0.7768$$

$$= 12.4264$$

$$\tilde{X}_s = \sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} X_{i\alpha} / \sum_{\alpha=1}^{N} k_\alpha = \frac{12.4264}{39} = 0.3186$$

$$\sum_{\alpha=1}^{N} (X_{o\alpha} - \tilde{X}_m) \sum_{\alpha=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)$$

$$= (-1.7674 - 0.4671)(-0.7809) + \ldots + (0.0078 - 0.4671)(-4.9301)$$

$$= -0.4421$$

$$\sum_{\alpha=1}^{N} k_\alpha (X_{o\alpha} - \tilde{X}_m)^2$$

$$= 2(-1.7674 - 0.4671)^2 + \ldots + 2(0.0078 - 0.4671)^2$$

$$= 63.4270$$

$$\sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)^2$$

$$= (0.0942 - 0.3186)^2 + \ldots + (-0.7768 - 0.3186)^2$$

$$= 37.6437$$

$$\hat{\rho}_{ms,p} = -0.4421/\left[63.4270(37.6437)\right]^{1/2}$$

$$= -0.0090$$

$$\sum_{\alpha=1}^{N} \sum_{i \neq j}^{k_\alpha} (X_{i\alpha} - \tilde{X}_s)(X_{j\alpha} - \tilde{X}_s)$$

$$= 2\Big[\left[(0.0942 - 0.3186)(-0.2379 - 0.3186)\right]$$

$$+ \left[(0.9984 - 0.3186)(-0.0880 - 0.3186) + (0.9984 - 0.3186)\right.$$

$$\left.(0.8767 - 0.3186) + \ldots + (0.1322 - 0.3186)(-0.3331 - 0.3186)\right]$$

$$+ \left[(-3.5161 - 0.3186)(-0.7768 - 0.3186)\right]\Big]$$

$$= 20.0052$$

$$\sum_{\alpha=1}^{N} k_\alpha (k_\alpha - 1) = 2(1) + 8(7) + \ldots + 2(1) = 184$$

$$\hat{\rho}_{ss} = \left[\frac{20.0052}{184}\right]\left[\frac{37.6437}{39}\right]$$

$$= \frac{20.0052(39)}{184(37.6437)} = 0.1126$$

$$D = \sum_{\alpha=1}^{N} \left[k_\alpha / (1 + (k_\alpha - 1)\,\hat{\rho}_{ss}) \right]$$

$$= 1.7975 + 4.4730 + \ldots + 1.7975$$

$$= 26.4950$$

$$\text{PAIR} = \frac{(26.4950 - 2)^{1/2}(-0.0090)}{(1 - .000081)^{1/2}}$$

$$= \frac{-.0445}{.9999} = -0.0448$$

ซึ่ง PAIR $= -0.0448 < t_{8,0.05} = 1.860$

ดังนั้น ยอมรับสมมติฐานหลักที่ระดับนัยสำคัญ 0.05

2.3.2 <u>The Sib-Mean Test</u>

การทดสอบนี้ใช้ตัวสถิติทดสอบ MEAN โดยสร้างจากตัวประมาณ Sib-Mean (Sib-Mean Estimator ; $\hat{\rho}_{ms,s}$)

$$\hat{\rho}_{ms,s} = \sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)(\bar{X}_{s\alpha} - \bar{X}_s) / \left[\sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)^2 \sum_{\alpha=1}^{N} (\bar{X}_{s\alpha} - \bar{X}_s)^2\right]^{\frac{1}{2}}$$

$$\bar{X}_{s\alpha} = \sum_{i=1}^{k_\alpha} X_{i\alpha} / k_\alpha \quad , \quad \bar{X}_m = \sum_{\alpha=1}^{N} X_{o\alpha} / N$$

$$\bar{X}_s = \sum_{\alpha=1}^{N} \bar{X}_{s\alpha} / N$$

$\hat{\rho}_{ms,s}$ มีรูปแบบเหมือนตัวประมาณสัมประสิทธิ์สหสัมพันธ์เชิงเส้นของ Pearson แต่ $\bar{X}_s$ เป็นค่าเฉลี่ยของค่าเฉลี่ยที่คิดจากการวัดซ้ำของแต่ละตัวอย่าง และ $(\bar{X}_{s\alpha}-\bar{X}_s)$ เป็นการหาผลต่างของค่าที่คิดจากการวัดซ้ำของแต่ละตัวอย่างกับ $\bar{X}_s$

$$MEAN = (N - 1)^{\frac{1}{2}} \hat{\rho}_{ms,s}$$

เกณฑ์การตัดสินใจคือ ปฏิเสธสมมติฐานหลัก $H_o : \rho_{ms} = 0$ เทียบกับสมมติฐานแย้ง $H_1 : \rho_{ms} > 0$ เมื่อ MEAN $> Z_\alpha$

<u>ตัวอย่างวิธีคำนวณค่าสถิติทดสอบ</u>

จากข้อมูลในตารางที่ 2.1 ของตัวอย่างที่ 1 สามารถนำมาคำนวณค่าสถิติทดสอบ MEAN ได้ดังนี้

$$\sum_{\alpha=1}^{N} X_{o\alpha} = -1.7674 + 2.3711 + \ldots + 0.0078$$

$$= 1.7204$$

$$\bar{X}_m = \sum_{\alpha=1}^{N} X_{o\alpha}/N = \frac{1.7204}{10} = 0.1720$$

$$\bar{X}_{s\alpha} = \sum_{i=1}^{k_\alpha} X_{i\alpha}/k_\alpha$$

$$\bar{X}_{s1} = \frac{0.0942 - 0.2379}{2} = \frac{-0.1437}{2} = -0.0718$$

$$\bar{X}_{s2} = 0.1043 , \ldots, \bar{X}_{s,10} = 2.1464$$

$$\sum_{\alpha=1}^{N} \bar{X}_{s\alpha} = -0.0718 + 0.1043 + \ldots + -2.1464$$

$$= 0.0146$$

$$\bar{X}_s = \sum_{\alpha=1}^{N} \bar{X}_{s\alpha}/N = \frac{0.0146}{10}$$

$$= 0.0015$$

$$\sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)(\bar{X}_{s\alpha} - \bar{X}_s)$$

$$= (-1.7674-0.1720)(-0.0718-0.0015)+\ldots+(0.0078-0.1720)$$

$$(-2.1464-0.0015)$$

$$= 2.3808$$

$$\sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)^2$$

$$= (-1.7674 - 0.1720)^2 + \ldots + (0.0078 - 0.1720)^2$$

$$= 14.6099$$

$$\sum_{\alpha=1}^{N} (\bar{X}_{s\alpha} - \bar{X}_s)^2$$

$$= (-0.0718 - 0.0015)^2 + \ldots + (-2.1464 - 0.0015)^2$$

$$= 7.5072$$

$$\hat{\rho}_{ms,s} = 2.3808 / \left[14.6099(7.5072)\right]^{\frac{1}{2}}$$

$$= 0.2273$$

$$\text{MEAN} = (10-1)^{\frac{1}{2}} (0.2273)$$

$$= 0.6820$$

ซึ่ง $\text{MEAN} = 0.6820 < Z_{.05} = 1.645$

ดังนั้น ยอมรับสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.3.3 The Ensemble Test

การทดสอบนี้ใช้ตัวสถิติ ENSE โดยสร้างจากตัวประมาณ Ensemble (Ensemble Estimator : $\hat{\rho}_{ms,e}$)

$$\hat{\rho}_{ms,e} = \frac{\sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)(\bar{X}_{s\alpha} - \bar{X}_s)}{\left[\sum_{\alpha=1}^{N} (X_{o\alpha} - \bar{X}_m)^2 \left[\frac{N-1}{N} \sum_{\alpha=1}^{N} \sum_{i=1}^{k_\alpha} (X_{i\alpha} - \bar{X}_{s\alpha})^2 / k_\alpha + \sum_{\alpha=1}^{N} (\bar{X}_{s\alpha} - \bar{X}_s)^2\right]\right]^{\frac{1}{2}}}$$

ตัวประมาณ $\hat{\rho}_{ms,e}$ เหมือนกับตัวประมาณ $\hat{\rho}_{ms,s}$ แต่ต่างกันตรงที่ตัวประมาณ $\hat{\rho}_{ms,e}$ มีตัวหารเพิ่มขึ้นคือ

$$\sum_{\alpha=1}^{N}(X_{o\alpha}-\bar{X}_m)^2\left[\frac{N-1}{N}\sum_{\alpha=1}^{N}\sum_{i=1}^{k_\alpha}(X_{i\alpha}-\bar{X}_{s\alpha})^2/k_\alpha\right]$$ ทำให้ $|\hat{\rho}_{ms,e}| < |\hat{\rho}_{ms,s}|$ เสมอ

$$\text{ENSE} \quad = \quad (N-1)^{\frac{1}{2}}\,\hat{\rho}_{ms,e}\Big/\left[N^{-1}\sum_{\alpha=1}^{N}(1/k_\alpha)(1-\hat{\rho}_{ss})+\hat{\rho}_{ss}\right]^{\frac{1}{2}}$$

เกณฑ์การตัดสินใจคือปฏิเสธสมมติฐานหลัก $H_o : \rho_{ms} = 0$ เทียบกับสมมติฐานแย้ง $H_1 : \rho_{ms} > 0$ เมื่อ ENSE $> Z_\alpha$

ตัวอย่างวิธีคำนวณค่าสถิติทดสอบ

จากข้อมูลในตารางที่ 2.1 ของตัวอย่างที่ 1 เราสามารถนำมาคำนวณค่าสถิติทดสอบ ENSE ได้ดังนี้

$$\sum_{\alpha=1}^{N}\sum_{i=1}^{k_\alpha}(X_{i\alpha}-\bar{X}_{s\alpha})^2/k_\alpha$$

$$= \frac{(0.0942+0.0718)^2+(-0.2379+0.0718)^2}{2} + \frac{(0.9984-0.1043)^2+\ldots+(-0.3331-0.1043)^2}{8}$$

$$+\ldots+\frac{(-3.5161+2.1464)^2+(-0.7768+2.1464)^2}{2}$$

$$= \quad 5.7389$$

$$\hat{\rho}_{ms,e} \quad = \quad \frac{2.3808}{\left[14.6099\left(\frac{9}{10}\,(5.7384)+7.5072\right)\right]^{\frac{1}{2}}}$$

$$= \quad 0.1750$$

$$\sum_{\alpha=1}^{N}(1/k_\alpha)(1-\hat{\rho}_{ss})$$

$$= \quad (1-0.1126)\left(\frac{1}{2}+\frac{1}{8}+\ldots+\frac{1}{2}\right)$$

$$\doteq \quad 0.8874(4.3929) \quad = \quad 3.8983$$

$$ENSE = \frac{(10 - 1)^{\frac{1}{2}} (.1750)}{\left[\frac{1}{10} (3.8983) + .1126\right]^{\frac{1}{2}}}$$

$$= 0.7406$$

ซึ่ง $ENSE = 0.7406 < Z_{.05} = 1.645$

ดังนั้น ยอมรับสมมติฐานหลัก ที่ระดับนัยสำคัญ 0.05

2.4 <u>ผลงานวิจัยที่เกี่ยวข้อง</u>

<u>B.Rosner, A. Donner และ C.H. Hennekens (1977)</u>

ทำการทดลอบประสิทธิภาพของตัวประมาณความสัมพันธ์ระหว่างกลุ่มคือ

1. Pairwise Estimator ($\hat{\rho}_{ms,p}$)
2. Sib-Mean Estimator ($\hat{\rho}_{ms,s}$)
3. Ensemble Estimator ($\hat{\rho}_{ms,e}$)

โดยใช้ค่าเฉลี่ยของจำนวนการวัดซ้ำเท่ากับ 3.12 (W. Brass:1958) และขนาดตัวอย่าง (N) = 50 ความสัมพันธ์ภายในกลุ่ม (ρ_{ss}) = 0.1, 0.3, 0.5, 0.8 ความสัมพันธ์ระหว่างกลุ่ม (ρ_{ms}) = 0, 0.1, 0.3, 0.5, 0.8 และใช้ MSE (Mean Square Error) เป็นเกณฑ์ในการเปรียบเทียบ ได้ผลการทดลอบดังนี้

1. เมื่อความสัมพันธ์ภายในกลุ่มมีค่าน้อยกว่าหรือเท่ากับ 0.1 ตัวประมาณ Pairwise จะมีประสิทธิภาพมากกว่าตัวประมาณอื่น ๆ

2. เมื่อความสัมพันธ์ภายในกลุ่มมีค่ามากกว่าหรือเท่ากับ 0.5 ตัวประมาณ Ensemble จะมีประสิทธิภาพมากกว่าตัวประมาณอื่น ๆ

3. เมื่อความสัมพันธ์ภายในกลุ่มมีค่าเท่ากับ 0.3 ตัวประมาณ Pairwise และ Ensemble มีประสิทธิภาพใกล้เคียงกัน

B. Rosner และ A. Donner (1979)

เปรียบเทียบการทดสอบความสัมพันธ์ระหว่างกลุ่ม 4 วิธี ดังนี้

1. The Classical Pairwise Test

$$S_Q = (Q-1)^{\frac{1}{2}} \hat{\rho}_{ms,p} / (1 - \hat{\rho}^2_{ms,p})$$

เมื่อ $Q = \sum_{\alpha=1}^{N} k_\alpha$

เกณฑ์ในการตัดสินใจคือปฏิเสธสมมติฐานหลักเมื่อค่าสถิติ $S_Q >$ ค่าสถิติ $t_{Q-2,\alpha}$

2. The Conservative Pairwise Test

$$S_c = (N-2)^{\frac{1}{2}} \hat{\rho}_{ms,p} / (1 - \hat{\rho}^2_{ms,p})^{\frac{1}{2}}$$

เกณฑ์ในการตัดสินใจคือปฏิเสธสมมติฐานหลักเมื่อค่าสถิติ $S_c >$ ค่าสถิติ $t_{N-2,\alpha}$

3. The Adjusted Pairwise Test

$$PAIR = (D-2)^{\frac{1}{2}} \hat{\rho}_{ms,p} / (1 - \hat{\rho}^2_{ms,p})^{\frac{1}{2}}$$

$$D = \sum_{\alpha=1}^{N} \left[k_\alpha / \left[1 + (k-1) \hat{\rho}_{ss} \right] \right]$$

เกณฑ์ในการตัดสินใจคือปฏิเสธสมมติฐานหลักเมื่อค่าสถิติ PAIR > ค่าสถิติ $t_{N-2,\alpha}$

4. The Studentized Sib-Mean Test

$$S_M = (N-2)^{\frac{1}{2}} \hat{\rho}_{ms,p} / (1 - \hat{\rho}^2_{ms,p})^{\frac{1}{2}}$$

เกณฑ์ในการตัดสินใจคือปฏิเสธสมมติฐานหลักเมื่อค่าสถิติ $S_M >$ ค่าสถิติ $t_{N-2,\alpha}$

โดยใช้ค่าเฉลี่ยของจำนวนครั้งการวัดซ้ำ 3.12 (W. Brass:1958) และขนาดตัวอย่าง (N) = 20,50 ความสัมพันธ์ภายในกลุ่ม (ρ_{ss}) = 0.1, 0.2, 0.3, 0.4, 0.5, 0.8 ความสัมพันธ์ระหว่างกลุ่ม (ρ_{ms}) = 0, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4 ระดับนัยสำคัญ (α) = 0.05 และใช้ความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบ เป็นเกณฑ์ในการเปรียบเทียบ ได้ผลดังนี้

1. S_Q และ S_c ควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ไม่ได้ โดยมีค่ามากกว่าขอบเขตบนของเกณฑ์ที่ใช้ในการพิจารณา ส่วนตัวสถิติ S_M และตัวสถิติ PAIR สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

2. การทดสอบความสัมพันธ์ระหว่างกลุ่มโดยใช้ตัวสถิติทดสอบ PAIR มีอำนาจการทดสอบสูงกว่าการทดสอบที่ใช้ตัวสถิติทดสอบ S_M

SADANORI KONISHI (1985)

เปรียบเทียบการทดสอบความสัมพันธ์ระหว่างกลุ่ม 2 วิธีคือ

1. The Sib-Mean Test
2. The Ensemble Test

โดยใช้ค่าเฉลี่ยของจำนวนครั้งการวัดซ้ำ = 3,7 ขนาดตัวอย่าง (N) = 25 ความสัมพันธ์ภายในกลุ่ม (ρ_{ss}) = 0.1, 0.3, 0.5, 0.7 ความสัมพันธ์ระหว่างกลุ่ม (ρ_{ms}) = 0, 0.1, 0.2, 0.3 ระดับนัยสำคัญ (α) = 0.05 และใช้ความผิดพลาดประเภทที่ 1 และอำนาจการทดสอบ เป็นเกณฑ์ในการเปรียบเทียบ ได้ผลดังนี้

1. เมื่อค่าเฉลี่ยของจำนวนครั้งการวัดซ้ำมีค่าเพิ่มขึ้น อำนาจการทดสอบของตัวสถิติทดสอบจะมีค่าเพิ่มขึ้นด้วย

2. การทดสอบความสัมพันธ์ระหว่างกลุ่มโดยใช้ตัวสถิติทดสอบ ENSE มีอำนาจการทดสอบสูงกว่าการทดสอบที่ใช้ตัวสถิติทดสอบ MEAN