บทที่ 2

วรรณคดีที่เกี่ยวข้อง

วิธีการเปรียบเทียบภายแรกพหุคูณของดันน์

ผู้ริเริ่มขบวนการนี้เป็นคนแรก ไม่ทราบแน่ชัด แต่ในปี 1961 โอลีฟจีนดันน์ (Olive Jean Dunn) ได้ทำการทดสอบคุณสมบัติของกระบวนการนี้ และได้จัดเตรียมตารางขึ้นเฉพาะเพื่อใช้สำหรับการคำนวณ ซึ่งต่อมากระบวนนี้เป็นที่รู้จักในฐานะเป็นวิธีการเปรียบเทียบพหุคูณของดันน์

วิธีการของดันน์ เน้นที่ปัญหาของความเชื่อมั่นที่เกิดจากการที่ผู้ทดลองต้องการที่จะประมาณช่วงความเชื่อมั่นค่าเฉลี่ยของการเปรียบเทียบ โดยการใช้การแจกแจง ที่ซึ่งให้ช่วงที่สั้นกว่าดันน์ได้เปรียบเทียบวิธีการนี้กับวิธีเอสของเชฟเฟย์ และพบว่าค่า t' ของวิธีเปรียบเทียบของดันน์จะใช้ได้เหมาะสมในกรณีที่

1. ค่าระดับการทดลอง หรือจำนวนกลุ่มตัวอย่างเพิ่มขึ้น (k)
2. ค่าดีกรีของความเป็นอิสระ (ν) เพิ่มขึ้น
3. $1 - \alpha$ มีค่าเพิ่มขึ้น

วิธีของดันน์จะใช้ได้ดีเมื่อเป็นการวางแผนทดสอบสมมติฐานล่วงหน้าและทราบจำนวนสมมติฐานที่จะทดสอบแน่นอน ในขณะที่วิธีเอสของเชฟเฟย์เหมาะสมที่จะใช้ภายหลังการทดสอบข้อมูลด้วยการทดสอบเอฟแล้ว ช่วงของความเชื่อมั่นด้วยวิธีของดันน์คือ

$$\hat{\Psi} \pm t'\alpha\sqrt{MS_e \Sigma \frac{a_j^2}{n_j}}$$

เมื่อ $$\hat{\Psi} = \sum_{j=1}^{J} a_j \bar{x}_j$$

และค่า t' มาจาก $$\int_{-\alpha}^{t'} f(t;\nu)\, dt = 1 - \frac{\alpha}{2n}$$

f(t;ν) คือ density function ของการแจกแจงแบบสติวเดนท์ ทีด้วยดีกรีของความเป็นอิสระ ν ตารางของ t' สำหรับ α = .05 และ .01 ที่ดันน์คิดขึ้น พิมพ์ใน "Multiple Comparisons Among Means" Jounal of American Statistical Association 56 (1961) p. 55

ในปี 1966 Miller ได้ใช้สถิติบอนเฟอโรนิที (Bonferroni t Statistics) ภายใต้กระบวนการเดียวกัน ทั้งกระบวนการเปรียบเทียบพหุคูณของดันน์และบอนเฟอโรนิทีนั้นให้เป็นประโยชน์สำหรับการเปรียบเทียบค่าเฉลี่ยภายแรกหรือการเปรียบเทียบที่ได้วางแผนไว้ก่อนซึ่งไม่จำเป็นจะต้องเป็นอิสระต่อกัน และกระบวนการเหล่านี้ขึ้นอยู่กับการแจกแจงสติวเดนท์ ที ที่สามารถแบ่งแยกระดับนัยสำคัญไปตามเซ็ทของการเปรียบเทียบที่ได้วางแผนไว้ และไม่จำเป็นต้องทดสอบ ความมีนัยสำคัญของค่าเฉลี่ยด้วยการทดสอบเอฟก่อน

<u>การคำนวณเพื่อใช้สำหรับเปรียบเทียบรายคู่</u>

คำนวณหาความแตกต่างระหว่างค่าเฉลี่ยเป็นรายคู่ (pairwise) แล้วเรียงตามลำดับจากค่าน้อยไปหามาก จำนวนของการเปรียบเทียบค่าเฉลี่ยเป็นรายคู่ทั้งหมดจะมีค่าเท่ากับ C ซึ่งเท่ากับ k(k-1)/2 คู่ เมื่อ k หมายถึง จำนวนค่าเฉลี่ยของทรีทเมนท์แล้วคำนวณค่าสถิติวิกฤตจากสูตร

$$d = t'_{D\alpha/z;C,\nu} \sqrt{MS_e \; \Sigma \frac{a_j^2}{n_j}}$$

เมื่อ $t'_{D\alpha/\alpha;C,\nu}$ ค่าวิกฤตจากตารางที่ดันน์ได้คำนวณไว้

สัมประสิทธิ์ของความเป็นอิสระสำหรับค่าเฉลี่ยที่ i

MS_e ค่าประมาณไม่ลำเอียงของความแปรปรวนความคลาดเคลื่อนของประชากร

n_j จำนวนกลุ่มตัวอย่างในแต่ละทรีทเมนท์

จากนั้นก็เปรียบเทียบค่าความแตกต่างของค่าเฉลี่ย แต่ละคู่ทั้งหมดกับค่าสถิติวิกฤตที่คำนวณด้วยวิธีของดันน์ (d) ค่าความแตกต่างของค่าเฉลี่ยจะมีนัยสำคัญก็ต่อเมื่อค่าความแตกต่างของค่าเฉลี่ยนั้นจะต้องมากกว่าค่าสถิติวิกฤต d ที่คำนวณได้

<u>วิธีทดสอบของทูกี้</u>

เป็นขบวนการเปรียบเทียบพหุคูณสำหรับการเปรียบเทียบหรือการวิเคราะห์ภายหลังที่คล้ายกับการทดสอบ LSD ของฟิชเชอร์ ซึ่งคิดขึ้นโดย J.W. Tukey ในปี 1953 โดยใช้การแจกแจงของสติวเดนท์ไทซ์เรนจ์ และเรียกว่า HSD (Honesty Significant Difference Test หรือ w procedure โดยตั้งอัตราความคลาดเคลื่อนที่ระดับ α จุดประสงค์ของวิธีทดสอบของทูกี้ คือ การเปรียบเทียบความแตกต่างค่าเฉลี่ยรายคู่ทั้งหมด ภายใต้ข้อตกลงเบื้องต้นของความเป็นปกติและความเท่ากันของความแปรปรวนเช่นเดียวกับข้อตกลงของการทดสอบด้วยอัตราส่วนเอฟ นอกจากนั้นขนาดของกลุ่มตัวอย่างต้องเท่ากันหรือควรจะเท่ากันโดยประมาณ

วิธีทดสอบของทูกี้สามารถให้ช่วงของความเชื่อมั่นของความแตกต่างค่าเฉลี่ยรายคู่ทั้งหมด $k(k-1)/2$ คู่ ถ้า $\{\Theta_i\}$ หมายถึง เซ็ตของค่าเฉลี่ย $\{\hat{\Theta}\}$ หมายถึง ค่าประมาณไม่ลำเอียงของ $\{\Theta_i\}$ $\{J_i\}$ หมายถึงขนาดของกลุ่มตัวอย่างที่ต้องเท่ากันทุกกลุ่ม และ S_e หมายถึงค่าประมาณความคลาดเคลื่อนของความแปรปรวนประชากร ข้อตกลงเบื้องต้นของค่าสถิติ $\{\hat{\Theta}_1\}$ และ S_e คือ

$$
\Omega \; : \; \left\{ \begin{array}{ll} & \{\hat{\Theta}_i\} \text{ มีความเป็นอิสระเชิงสถิติ และ } \hat{\Theta}_i \sim N(\Theta_i, a^2\sigma^2), \\ & i = 1,\ldots,k \\ a & \text{เป็นค่าคงที่บวก } S_e \text{ เป็นค่าประมาณกำลังสองอิสระของ } \sigma^2 \\ & \text{ด้วยดีกรีความเป็นอิสระ } \nu \end{array} \right.
$$

ดังนั้นวิธีทดสอบของทูกี้ จะให้ช่วงความเชื่อมั่นดังต่อไปนี้

<u>ทฤษฎี</u> ภายใต้ข้อตกลง Ω โอกาส $1-\alpha$ ของความแตกต่างค่าเฉลี่ยรายคู่ $(\Theta_i - \Theta_{i'})$ ทั้งหมด $k(k-1)/2$ คู่ จะอยู่ในช่วง

$$\hat{\Theta}_i - \hat{\Theta}_{i'} - TS_e \le \Theta_i - \Theta_{i'} \le \hat{\Theta}_i - \hat{\Theta}_{i'} + TS_e$$

เมื่อ $T = aq_{\alpha;k,\nu}$

$q_{\alpha;k,\nu}$ มาจากการแจกแจงของสติวเดนท์ไทซ์เรนจ์

การคำนวณเพื่อใช้สำหรับเปรียบเทียบรายคู่

คำนวณหาค่าความแตกต่างของค่าเฉลี่ยเป็นรายคู่ทั้งหมดและคำนวณค่าสถิติวิกฤตจากสูตร

$$T = q_{\alpha;k,\nu} \sqrt{\frac{S_e}{n}}$$

เมื่อ $q_{\alpha;k,\nu}$ เป็นค่าวิกฤตที่ได้จากตารางการแจกแจงสติวเดนซ์ไทซ์เรนจ์

k หมายถึงระดับของการทดลอง หรือจำนวนกลุ่มตัวอย่าง

α หมายถึงระดับนัยสำคัญสำหรับการทดสอบ

n ขนาดของกลุ่มตัวอย่าง

ν ดีกรีความเป็นอิสระของ

S_e ค่าประมาณความคลาดเคลื่อนประชากร

เปรียบเทียบค่าสถิติวิกฤต T ที่คำนวณได้กับค่าความแตกต่างค่าเฉลี่ยรายคู่ทั้งหมดค่าความแตกต่างค่าเฉลี่ยรายคู่ใดจะเกิดนัยสำคัญก็ต่อเมื่อค่าความแตกต่างค่าเฉลี่ยรายคู่นั้นจะต้องมีค่ามากกว่าค่าสถิติวิกฤต T ที่คำนวณได้

การเปรียบเทียบพหุคูณด้วยวิธีเอสของเชฟเฟ่

Scheff'e (1953) ได้เสนอวิธีที่เรียกว่าวิธีเอส เพื่อเปรียบเทียบความแตกต่างของค่าเฉลี่ยทั้งหมด วิธีนี้มีความไวน้อยกว่าการทดสอบ HSD ของทูกี้ แต่จะสามารถนำไปใช้ในการเปรียบเทียบที่ซับซ้อนยิ่งขึ้น ดังนั้นวิธีของเชฟเฟ่จึงเหมาะสำหรับใช้เปรียบเทียบจำนวนมากๆ จำนวนเท่าใดก็ได้ไม่มีขีดจำกัดทั้งที่เป็นรายคู่และไม่ใช่รายคู่ โดยตั้งอัตราความคลาดเคลื่อนของการทดลองเท่ากับ α วิธีของเชฟเฟ่มักจะทำภายหลังที่การทดสอบเอฟมีนัยสำคัญ และพื้นฐานทางทฤษฎีวิธีเอฟของเชฟเฟ่มาจากการใช้การแจกแจงเอฟซึ่งให้ช่วงความเชื่อมั่นสำหรับการเปรียบเทียบทั้งหมด $1 - \alpha$ ดังนี้

$$\hat{\Psi} - S\hat{\sigma}^2_{\hat{\Psi}} \leq \Psi \leq \hat{\Psi} + S\hat{\sigma}^2_{\hat{\Psi}}$$

เมื่อ S มาจาก $S^2 = (k-1)F_{\alpha,k-1,N-k}$

$\hat{\sigma}^2_{\hat{\Psi}} = S^2_p \sum_i (c_i^2 / n_i)$ $\hat{\alpha}^2_{\hat{\Psi}}$ คือ MS_e

Ψ และ $\hat{\Psi}$ คือการเปรียบเทียบของประชากรและตัวประมาณค่าของการเปรียบเทียบ

การคำนวณเพื่อใช้สำหรับเปรียบเทียบรายคู่

คำนวณค่าสถิติวิกฤตจาก

$$S = \sqrt{(k-1)F_{\alpha,\nu_1,\nu_2}} \sqrt{MS_e \left[\Sigma \left(\frac{c_i^2}{n_i}\right)\right]}$$

เมื่อ $F_{\alpha;\nu_1,\nu_2}$ เป็นค่าที่ได้จากตารางการแจกแจงเอฟ

ν_1 หมายถึง ดีกรีของความเป็นอิสระมีค่าเท่ากับ $k - 1$

ν_2 หมายถึง ดีกรีของความเป็นอิสระมีค่าเท่ากับ $N - k$

k หมายถึง จำนวนกลุ่มตัวอย่าง

c_i หมายถึง สัมประสิทธิ์ของการเปรียบเทียบ

n_i หมายถึง ขนาดของกลุ่มตัวอย่าง

การเปรียบเทียบความแตกต่างของค่าเฉลี่ยจะแสดงนัยสำคัญก็ต่อเมื่อความแตกต่างค่าเฉลี่ยนั้นมีค่ามากกว่าสถิติวิกฤต S ที่คำนวณได้

การเปรียบเทียบพหุคูณด้วยวิธีทดสอบของดันเนตต์

วิธีทดสอบของดันเนตต์มีจุดมุ่งหมายเพื่อเปรียบเทียบค่าเฉลี่ยของทรีทเมนต์หลายๆทรีทเมนต์ในระดับการทดลองกับค่าเฉลี่ยทรีทเมนต์ควบคุม ในปี 1955 ดันเนตต์ได้พัฒนาการทดสอบช่วงพหุคูณสำหรับจำนวนการเปรียบเทียบ $k - 1$ การเปรียบเทียบของค่าเฉลี่ย k ค่า อัตราความคลาดเคลื่อน

การทดสอบการเปรียบเทียบ k - 1 ครั้ง คือ α ซึ่งดันเนตต์กล่าวว่าการใช้กระบวนวิธีทดสอบของทูกี้หรือเชฟเฟย์นั้นให้ลิมิตของความเชื่อมั่นเกินความจำเป็น รากฐานทฤษฎีของดันเนตต์มีดังนี้

<u>ทฤษฎี</u>

สมมติการทดลอง N ทรีทเมนต์ ซึ่ง N_o เป็นทรีทเมนต์ที่ควบคุม ตามข้อตกลงปกติของการวิเคราะห์ความแปรปรวน คือ x_{ij} เป็นการแจกแจงแบบปกติและอิสระด้วยความแปรปรวน σ^2 และค่าเฉลี่ย $m_{i.}$ ถ้าการทดลองมีเพียงหนึ่งทรีทเมนต์ สำหรับเปรียบเทียบกับทรีทเมนต์ควบคุม การทดสอบสมมติฐานของความแตกต่างจะขึ้นอยู่กับการแจกแจงที ที่ $t = \frac{z}{s}$ เมื่อ

$$z = \frac{\bar{x}_1 - \bar{x}_o - (m_1 - m_o)}{\sqrt{\frac{1}{N_1} + \frac{1}{N_o}}}$$

ที่มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย 0 และความแปรปรวน σ^2 แล้วค่าของ $t = \frac{z}{s}$ จะมีการแจกแจงทีด้วยดีกรีของความเป็นอิสระ n ช่วงความเชื่อมั่นของ $m_1 - m_o$ คือ

$$\bar{x}_1 - \bar{x}_o \pm d'' S\sqrt{\frac{1}{N_1} + \frac{1}{N_o}}$$

เมื่อ d'' มาจาก $\text{Prob}\ (|t| < d'') = P$ ....(1)

ถ้าการทดลองมี p ทรีทเมนต์และทรีทเมนต์ควบคุม

$$z = \frac{\bar{x}_i - \bar{x}_o - (m_i - m_o)}{\sqrt{\frac{1}{n_i} - \frac{1}{n_o}}}$$

และ $t_i = \frac{z_i}{s}$, $i = 1, 2, \ldots, p$ ดังนั้นค่าของลิมิตความเชื่อมั่น 2 ทาง คือ

$$\bar{x}_i - \bar{x}_o \pm d''_i\ S\sqrt{\frac{1}{N_i} + \frac{1}{N_o}} \quad (i > 1, 2, \ldots, p)$$

ถ้า p เป็นค่าคงที่ d_1 ก็จะสอดคล้องกับสมการ

$$\text{Prob}\ (\ |t_1| < d''_1\ ,\ |t_2| < d''_2\ ,\ \ldots\ldots\ldots,\ |t_p| < d''_i)\ \ldots\ldots\ldots \quad (2)$$

เพื่อที่จะหาค่าของ d จากค่ารากของสมการ (2) ดังนั้นสมการ (2) สามารถเขียนได้ในรูป

$$P = \text{Prob}\ (|z_i| < d''_1\ S,\ |z_2| < d''_2\ S\ ,\ldots\ldots\ldots,\ |z_p| < d''_p\ S)$$

$$P = \int_{-\infty}^{+\infty} G\ (d''_i S\ ,\ d''_2 S\ ,\ldots\ldots\ldots,\ d''_p S)\ p(s)\ ds$$

เมื่อ $G(z_1,z_2,\ldots,z_{pp})$ เป็น c.d.f. ของ $|z_i|$

การคำนวณเพื่อใช้สำหรับเปรียบเทียบรายคู่

$$d'' \quad = \quad tD_{\alpha/2;k,v}\ \sqrt{\frac{2S}{n}}$$

ค่า $tD_{\alpha/2;k,v}$ มาจากตารางของดันเนตต์ วิธีทดสอบของดันเนตต์จะแสดงนัยสำคัญก็ต่อเมื่อ ความแตกต่างค่าเฉลี่ยมีค่ามากสถิติวิกฤต d''

การเปรียบเทียบพหุคูณด้วยวิธีทดสอบของนิวแมนคูลส์

การทดสอบของนิวแมนคูลส์ เป็นวิธีที่มีกระบวนการแตกต่างไปจากการเปรียบเทียบพหุคูณด้วยวิธีอื่นๆ อันเกิดจากข้อเสนอแนะของสติวเดนท์ (1927) ซึ่งนิวแมน (1939) เป็นผู้พัฒนาการทดสอบนี้ขึ้นและคูลส์ (1952) ได้ขยายวิธีการทดสอบของนิวแมน ดังนั้นการทดสอบนี้จึงเรียกว่า วิธีทดสอบของนิวแมนคูลส์ การทดสอบเริ่มจากวิธีการเรียงลำดับของค่าเฉลี่ยจากน้อยไปมาก และค่าความแตกต่างของค่าเฉลี่ยตามลำดับขั้นค่าเฉลี่ยที่ห่างกัน เพื่อเป็นการป้องกันขีดจำกัดล่างของ 1-α สำหรับการเรียงลำดับค่าเฉลี่ยทั้งหมด ค่าสถิติวิกฤตของความแตกต่างค่าเฉลี่ยรายคู่ของการทดสอบจะเปลี่ยนแปลงไปตามลำดับขั้นของค่าเฉลี่ยที่ห่างกันในแต่ละชุดการเปรียบเทียบ การทดสอบของนิวแมนคูลส์นี้ไม่อาจที่จะสร้างช่วงของความเชื่อมั่น เพราะกระบวนการทดสอบไม่เกี่ยวข้องกับกระบวนการหาความเชื่อมั่นอย่างใด

การคำนวณเพื่อใช้สำหรับเปรียบเทียบรายคู่

คำนวณหาความแตกต่างระหว่างค่าเฉลี่ยเป็นรายคู่ ซึ่งเรียกตามลำดับค่าเฉลี่ยจากน้อยไปหามากและคำนวณค่าสถิติวิกฤตจากสูตร

$$w_r \quad = \quad q_{r,\alpha;\nu}\ \sqrt{\frac{MS_e}{n}}$$

$q_{r,\alpha;\nu}$ หมายถึงค่าวิกฤตจากตารางของสติวเดนท์ไทซ์เรนจ์

r หมายถึงจำนวนลำดับขั้นของค่าเฉลี่ยที่ห่างกัน

MS_e หมายถึงค่าประมาณความคลาดเคลื่อนประชากร

n หมายถึงขนาดกลุ่มตัวอย่าง

การเรียงลำดับของค่าเฉลี่ยจากน้อยไปหามาก ค่าเฉลี่ย 2 ค่า ที่เปรียบเทียบกันสามารถอธิบายลำดับขั้นของความห่างของค่าเฉลี่ยด้วยแผนภาพข้างล่างนี้

$$\bar{x}_n \quad \bar{x}_j \quad \bar{x}_k \quad \bar{x}_l \quad \bar{x}_m \quad \bar{x}_n \quad \bar{x}_o$$

ภายใต้ความยาวของเส้นที่เท่ากัน ค่าสถิติวิกฤตของการทดสอบความแตกต่างค่าเฉลี่ยก็จะมีค่าเหมือนกัน วิธีทดสอบของนิวแมนคูลส์นี้จะแสดงนัยสำคัญก็ต่อเมื่อค่าความแตกต่างค่าเฉลี่ยรายคู่มีค่ามากกว่าสถิติวิกฤตของขั้นที่ r (W_r)

<u>การวิจัยที่เกี่ยวข้อง</u>

Olive J. Dunn[1] ได้เปรียบเทียบขบวนการวิเคราะห์ที่เกิดขึ้นกับขบวนการเปรียบเทียบภายหลังด้วยวิธีทดสอบของทูกี้ และเชฟเฟ่ และได้แสดงให้เห็นว่าการเปรียบเทียบพหุคูณของดันน์ ขึ้นอยู่กับจำนวนการเปรียบเทียบ ถ้าจำนวนการเปรียบเทียบมีมากกว่าจำนวนค่าเฉลี่ยที่เกิดจากทรีทเมนต์หลายๆ วิธีในการทดลองแล้ว ขบวนการของวิเคราะห์ด้วยวิธีของดันน์จะให้ช่วงความเชื่อมั่นมากกว่าขบวนการเปรียบเทียบภายหลังด้วยวิธีเอสทดสอบของทูกี้และเชฟเฟ่ ที่เป็นเช่นนี้ก็เพราะการเปรียบเทียบด้วยวิธีทดสอบของทูกี้และเชฟเฟ่ที่ขึ้นอยู่กับจำนวนค่าเฉลี่ย ดังนั้น ถ้าผู้วิจัยได้ทราบถึงจำนวนเปรียบเทียบแบบอิสระที่น้อยกว่าจำนวนค่าเฉลี่ยแล้ว การเปรียบเทียบพหุคูณของดันน์ย่อมมีอำนาจมากกว่ากระบวนการเปรียบเทียบภายหลังทั้ง 2 วิธี

Kirk[2] ได้แสดงการเปรียบเทียบอำนาจของกระบวนการเปรียบเทียบพหุคูณวิธีต่างๆ โดยพิจารณาจากสถิติวิกฤตของการทดสอบการเปรียบเทียบพหุคูณ แต่ละวิธีจากการเปรียบเทียบพบว่า การเปรียบเทียบภายแรกแบบอิสระนั้นมีอำนาจมากกว่าการเปรียบเทียบภายแรกแบบไม่อิสระ

---

[1]Kirk, Expermental Design Procedure for the Behavioral Sciences, 81.

[2]Ibnd., p. 95

และการเปรียบเทียบภายหลัง นั่นคือวิธีการเปรียบเทียบภายแรกแบบอิสระนั้น สามารถจับความแตกต่างค่าเฉลี่ยที่แท้จริงได้ วิธีทดสอบของทูกี้และเชฟเฟย์มีอำนาจน้อย เมื่อเทียบกับวิธีอื่นๆ แต่มีคุณสมบัติในการช่วยทดสอบสมมติฐานได้มากกว่า

Boardman และ Moffitt[1] ได้เปรียบเทียบวิธีวิเคราะห์พหุคูณ 5 วิธี คือ LSD HSD วิธีเอส การทดสอบช่วงพหุคูณของดันแคน และวิธีทดสอบของนิวแมนคูลส์ ด้วยกลุ่มตัวอย่างที่มีการแจกแจงแบบปกติขนาด 5,10 และ 15 ในระดับการทดลองตั้งแต่ 2-11 และทำการทดลองซ้ำ 1,000 ครั้ง เปรียบเทียบอัตราความคลาดเคลื่อน 2 แบบ คือ อัตราความคลาดเคลื่อนต่อการเปรียบเทียบและอัตราความคลาดเคลื่อนต่อการทดลอง จากการวิจัยได้แสดงให้เห็นเด่นชัดว่าอัตราความคลาดเคลื่อนของวิธี LSD และการทดสอบช่วงพหุคูณของดันแคนเพิ่มขึ้นตามจำนวนค่าเฉลี่ย ส่วนวิธีเอสของเชฟเฟย์เป็นวิธีเปรียบเทียบพหุคูณที่มีอัตราความคลาดเคลื่อนที่ นุรักษ์นิยมมากที่สุด

Carmer และ Swanson[2] ใช้เทคนิคของซีวมูเลชั่นเพื่อศึกษาอัตราความคลาดเคลื่อนประเภทที่ 1 และอัตราความคลาดเคลื่อนประเภทที่ 2 ของวิธีเปรียบเทียบพหุคูณแบบต่างๆด้วยแผนการทดลองแบบ B-k โดยเปรียบเทียบค่าเฉลี่ย 10 คู่ ซึ่งมีพารามิเตอร์ $\mu = 100$, $\sigma^2 = 100$ ขนาดกลุ่มตัวอย่างเท่ากับ 5,10 และ 20 ตามลำดับทำการจำลองการทดลองซ้ำ 4,000 ครั้ง และตั้งระดับนัยสำคัญเฉพาะที่ α เท่ากับ .05 เท่านั้น ผลการวิจัยแสดงอัตราความคลาดเคลื่อน 2 แบบ คือ อัตราความคลาดเคลื่อนต่อการเปรียบเทียบและอัตราความคลาดเคลื่อนต่อการทดลอง อัตราความคลาดเคลื่อนที่เกิดจากวิธีการเปรียบเทียบพหุคูณด้วยวิธีเอสของเชฟเฟย์ มีค่าน้อยกว่าวิธีอื่นๆ วิธีทดสอบของทูกี้ และวิธีทดสอบของนิวแมนคูลส์ให้อัตราความคลาดเคลื่อนสูงกว่าอัตราความคลาดเคลื่อนที่ระบุในระดับ .05 แต่อย่างไรก็ตามวิธีทดสอบของทูกี้ ยังให้อัตราความคลาดเคลื่อนที่น้อยกว่าวิธีทดสอบของนิวแมนคูลส์ และวิธีทดสอบของทูกี้ใช้ได้ดีแม้ขนาดกลุ่มตัวอย่างจะมีขนาดใหญ่ขึ้น

---

[1] Thomas J. Boardman and Donald R. Moffitt, "Graphical Monte Carb Type I Error Rates for Multiple Comparison Procedures," Biometric, (September 1971), pp. 738-744.

[2] S.G. Carmer and M.R. Swanson, "An Evaluation of Ten Pairwise Multiple Comparison Procedures Monte Caulo Methods," Journal of the Americass Statistical Association, (March 1973), pp. 66-74

Bernhardson[1] (1975) ได้เปรียบเทียบอัตราความคลาดเคลื่อนประเภทที่ 1 ของการเปรียบเทียบพหุคูณ 5 วิธี คือ LSD HSD วิธีเอสของเชฟเฟย์ การทดสอบของนิวแมนคูลส์ และการทดสอบช่วงพหุคูณของดันแคน การวิจัยครั้งนี้ทำการวิจัยเฉพาะขนาดกลุ่มตัวอย่าง 15 และกระทำภายหลังการวิเคราะห์ความแปรปรวนด้วยการทดสอบเอส กำหนดพารามิเตอร์ $\mu = 50$, $\sigma = 15$ ภายใต้การแจกแจงของประชากรแบบปกติโดยใช้สับรูทีน โปรแกรม Gauss ผลการวิจัยสอดคล้องกับการวิจัยของ Boardman และ Maffitt แต่ Bernhavdson สรุปว่าถ้าทำการเปรียบเทียบพหุคูณภายหลังการวิเคราะห์ความแปรปรวนด้วยการทดสอบเอฟแล้วผลของอัตราความคลาดเคลื่อนประเภทที่ 1 จะลดน้อยลงอันเนื่องมาจากการทดสอบเอฟนั้นได้ป้องกันอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ

---

[1]Clemens S. Bernhardson, "Type I Error Rates When Multiple Comparison Procedures Follow a Significant F test of Anova," Biometrics 31 (March 1975) pp. 229-232.