การเปรียบเทียบพหุคูณ (Multiple Comparison Procedures)

ฉัตรศิริ ปิยะพิมลสิทธิ์

การเปรียบเทียบพหุคูณจะเกี่ยวข้องกับการเปรียบเทียบค่าเฉลี่ยระหว่างกลุ่ม ซึ่งในเนื้อหาที่ผ่านมาจะเป็นการวิเคราะห์ความแปรปรวนแบบหนึ่งองค์ประกอบ ใช้เมื่อต้องการเปรียบเทียบค่าเฉลี่ยของกลุ่มตั้งแต่ 2 กลุ่มขึ้น ถ้าผลการทดสอบ F-test มีนัยสำคัญหรือปฏิเสธ H_0 แล้ว หากเป็นกรณีเปรียบเทียบเฉพาะกลุ่มตัวอย่าง 2 กลุ่ม และเป็นการปฏิเสธ H_0 แสดงว่าค่าเฉลี่ย 2 กลุ่มมีความแตกต่างกัน กรณีที่เป็นกลุ่มตัวอย่างมากกว่า 2 กลุ่ม และปฏิเสธ H_0 กรณีนี้ควรจะต้องใช้กระบวนการเปรียบเทียบพหุคูณ (Multiple Comparison Procedures : MCP) ดังนั้น ในกรณีนี้จะต้องมีค่าเฉลี่ยอย่างน้อย 3 กลุ่มและ H_0 ถูกปฏิเสธในการทดสอบ ANOVA ดังนั้นควรจะดำเนินการ MCP ในขั้นถัดไป เพื่อเปรียบเทียบค่าเฉลี่ยของแต่ละกลุ่มว่าแตกต่างกันหรือไม่ หากเป็นกรณีที่ผู้วิจัยไม่ได้สนใจการวิเคราะห์ความแปรปรวนแบบทางเดียว แต่สนใจเฉพาะการเปรียบเทียบค่าเฉลี่ยเฉพาะกลุ่ม นั่นคือกรณีที่ MCP จะใช้ในการเปรียบเทียบค่าเฉลี่ยเฉพาะบางกลุ่มที่สนใจ

ในเนื้อหานี้จะนำเสนอแนวคิดที่สำคัญ เช่น การนิยามถึงการเปรียบเทียบ (Contrast) การเปรียบเทียบแบบวางแผนหรือเปรียบเทียบก่อนกับการเปรียบเทียบหลัง ระดับความคลาดเคลื่อนแบบที่ 1 และการเปรียบเทียบแบบอิสระ (orthogonal contrasts) จะกล่าวถึงแนวคิดของการเปรียบเทียบแบบต่าง ๆ และรวมไปถึงการประยุกต์ใช้ MCP จะใช้สำหรับการเปรียบเทียบในแต่ละระดับของตัวแปรอิสระ

แนวคิดของกระบวนการเปรียบเทียบพหุคูณ

หัวข้อนี้จะอธิบายลักษณะสำคัญของการเปรียบเทียบพหุคูณ เราจะเริ่มต้นด้วยการนิยามการเปรียบเทียบ และการเปรียบเทียบวางแผนกับการเปรียบเทียบหลัง ระดับความคลาดเคลื่อนแบบที่ 1 และการเปรียบเทียบอิสระ (orthogonal contrasts)

การเปรียบเทียบ (Contrasts)

การเปรียบเทียบ คือ การถ่วงน้ำหนักค่าเฉลี่ยแล้วนำมารวมกัน ตัวอย่าง การเปรียบเทียบเป็นรายคู่อาจจะเป็นดังนี้ 1) กลุ่ม 1 กับกลุ่ม 2 2) การรวมกันของกลุ่ม 1 และ 2 กับกลุ่ม 3 ในเชิงสถิติ การเปรียบเทียบจะนิยามว่า

$$\psi_i = c_1\mu_1 + c_2\mu_2 + \ldots + c_j\mu_j$$

เมื่อ c_j เป็นสัมประสิทธ์การเปรียบเทียบ (หรือน้ำหนัก) ซึ่งอาจจะเป็นบวกหรือลบและนิยามการเปรียบเทียบว่า ψ_i ในอีกกรณีหนึ่ง การเปรียบเทียบอย่างง่ายจะมีหลักการรวมกันของ

ค่าเฉลี่ยของกลุ่ม ขึ้นอยู่กับค่าเฉลี่ยที่ผู้วิจัยสนใจในการเปรียบเทียบ การเปรียบเทียบสามารถเขียนได้ในรูปซับซ้อนดังนี้

$$\psi_i = \Sigma(c_j\mu_j)$$

จะเห็นว่า ค่าเฉลี่ยแต่ละกลุ่ม μ_j ถูกถ่วงน้ำหนัก (หรือถูกคูณ) โดยสัมประสิทธิ์การเปรียบเทียบ c_j และนำมารวมกันในแต่ละกลุ่มตั้งแต่ $j = 1,\ldots,J$ ดังนั้น รูปแบบการเปรียบเทียบนี้ $\Sigma c_j = 0$ สำหรับกรณีที่ n แต่ละกลุ่มมีจำนวนเท่ากัน และ $\Sigma(c_j\mu_j) = 0$ สำหรับกรณี n แต่ละกลุ่มไม่เท่ากัน

ตัวอย่าง สมมติว่าต้องการเปรียบเทียบค่าเฉลี่ยของกลุ่ม 1 และ 3 สำหรับ J = 4 และเรียกว่า การเปรียบเทียบชุดที่ 1 (Contrast 1) ซึ่งการเปรียบเทียบจะเขียนได้ว่า

$$\begin{aligned}\psi_1 &= \Sigma(c_j\mu_j)\\ &= c_1\mu_1 + c_2\mu_2 + c_3\mu_3 + c_4\mu_4\\ &= (+1)\mu_1 + (0)\mu_2 + (-1)\mu_3 + (0)\mu_4\\ &= \mu_1 - \mu_3\end{aligned}$$

อะไรคือสมมติฐานทดสอบเมื่อต้องการเปรียบเทียบ สมมติฐานศูนย์และสมมติฐานอื่นของการเปรียบเทียบนี้สามารถเขียนในรูปอย่างง่ายได้ว่า

$H_0 : \psi_i = 0$

และ $H_1 : \psi_i \neq 0$

ดังนั้น หากทำการทดสอบการรวมกันของค่าเฉลี่ยที่คูณด้วยสัมประสิทธิ์การเปรียบเทียบในรูปของความแตกต่าง ซึ่งจะไปเกี่ยวพันย้อนไปถึงการเปรียบเทียบ F-test สมมติฐานศูนย์และสมมติฐานอื่นสำหรับการทดสอบ F-test สามารถเขียนในเทอมของการเปรียบเทียบได้ว่า

H_0 : ψ_i ทั้งหมด = 0

และ H_1 : มี ψ_i อย่างน้อย 1 ค่าที่ $\neq 0$

การทดสอบ F-test จะมีผลสำหรับการเปรียบเทียบทุก ๆ ชุด ว่าควรจะมีนัยสำคัญในชุดของค่าเฉลี่ยที่มี J ค่า

ก่อนที่เราได้จะทราบชนิดของการเปรียบเทียบ จะอธิบายถึงความเกี่ยวข้องกับแนวคิดโมเดลเชิงเส้น ซึ่งสมการสำหรับโมเดลเชิงเส้นของ ANOVA คือ

$$Y_{ij} = \mu + \alpha_j + \varepsilon_{ij}$$

เมื่อ $\alpha_j = (\mu_j - \mu)$ การเปรียบเทียบสามารถเขียนได้ว่า

$$\psi_i = \Sigma(c_j\mu_j)$$

แต่เพราะว่า $\mu_j = (\mu + \alpha_j)$ แล้ว การเปรียบเทียบสามารถเขียนได้ว่า

$$\begin{aligned}\psi_i &= \Sigma[c_j(\mu + \alpha_j)]\\ &= \mu\Sigma(c_j) + \Sigma(c_j\alpha_j)\\ &= \mu(0) + \Sigma(c_j\alpha_j)\end{aligned}$$

$$= \Sigma(c_j\alpha_j)$$

เพราะว่า $\Sigma(c_j) = 0$ จากที่นิยามไว้ก่อนหน้านี้ จากรูปแบบของการเปรียบเทียบนี้ มี 2 ประการที่ต้องหลีกเลี่ยงคือ ประการแรก การเปรียบเทียบไม่ได้เกี่ยวข้องโดยตรงกับการทดสอบค่าเฉลี่ยทั้งหมด แต่จะเกี่ยวข้องโดยตรงเฉพาะ c_j และ α_j ประการที่สอง การเปรียบเทียบระหว่างค่าเฉลี่ยเป็นการเปรียบเทียบอิทธิพลของกลุ่ม α_j ซึ่งจะเป็นการเปรียบเทียบเฉพาะอิทธิพลของกลุ่มที่เราสนใจศึกษา

การเปรียบเทียบสามารถทำได้โดยการเปรียบเทียบอย่างง่ายหรือเปรียบเทียบรายคู่และการเปรียบเทียบที่ซับซ้อนหรือไม่ใช่รายคู่ การเปรียบเทียบอย่างง่ายหรือรายคู่จะเป็นการเปรียบเทียบที่เกี่ยวข้องกับค่าเฉลี่ย 2 ค่าเท่านั้น ตัวอย่างสถานการณ์การเปรียบเทียบกับกลุ่ม J = 3 กลุ่ม จะมีความเป็นไปได้ในการเปรียบเทียบเป็นรายคู่คือ 1) $\mu_1 - \mu_2 = 0$ 2) $\mu_1 - \mu_3 = 0$ และ 3) $\mu_2 - \mu_3 = 0$ ในการเขียนเพื่อการเปรียบเทียบรายคู่นั้นระหว่างกลุ่ม 1 และ 2 นั้นจะสามารถเขียนสลับได้คือ $\mu_1 - \mu_2 = 0$ หรือ $\mu_2 - \mu_1 = 0$ ในเทอมของสัมประสิทธิ์การเปรียบเทียบควรจะเขียนในรูปของตารางได้ดังนี้

	c_1	c_2	c_3
$\psi_1 : \mu_1 - \mu_2 = 0$	+1	−1	0
$\psi_2 : \mu_1 - \mu_3 = 0$	+1	0	−1
$\psi_3 : \mu_2 - \mu_3 = 0$	0	+1	−1

เมื่อในแต่ละการเปรียบเทียบสามารถสลับค่าเฉลี่ยกันได้ในการกำหนดสัมประสิทธิ์การเปรียบเทียบ ตัวอย่างเช่น ในการเปรียบเทียบชุดแรก ψ_1 จะไม่เกี่ยวข้องกับกลุ่ม 3 เพราะว่าสัมประสิทธิ์การเปรียบเทียบเป็นศูนย์ แต่จะเกี่ยวข้องกับกลุ่ม 1 และ 2 เพราะว่าสัมประสิทธิ์การเปรียบเทียบไม่เป็นศูนย์ สัมประสิทธิ์การเปรียบเทียบจะเป็น +1 สำหรับกลุ่ม 1 และ −1 สำหรับกลุ่ม 2 การเปรียบเทียบทั้งสามเขียนได้ดังนี้

$$\begin{aligned}
\psi_1 &= (+1)\mu_1 + (-1)\mu_2 + (0)\mu_3 \\
&= \mu_1 - \mu_2 \\
\psi_2 &= (+1)\mu_1 + (0)\mu_2 + (-1)\mu_3 \\
&= \mu_1 - \mu_3 \\
\psi_3 &= (0)\mu_1 + (+1)\mu_2 + (-1)\mu_3 \\
&= \mu_2 - \mu_3
\end{aligned}$$

จำนวนที่เป็นไปได้ในการเปรียบเทียบเป็นรายคู่สามารถคำนวณได้ด้วยสูตร $[J(J-1)]/2$ ดังนั้นถ้า J = 3 จะได้จำนวนคู่ในการเปรียบเทียบคือ 3 คู่ เมื่อ J = 5 จะได้จำนวนคู่ในการเปรียบเทียบคือ 10 คู่

การเปรียบเทียบแบบซับซ้อนเป็นการเปรียบเทียบที่เกี่ยวข้องกับค่าเฉลี่ยตั้งแต่ 3 กลุ่มขึ้นไป ต่อจากตัวอย่างข้างต้นที่ J = 3 กลุ่ม เราอาจจะสนใจในการทดสอบเปรียบเทียบ μ_1 - [(μ_2

$+ \mu_3)/2]$ การเปรียบเทียบนี้เป็นการเปรียบเทียบของค่าเฉลี่ยของกลุ่ม 1 กับการเฉลี่ยของค่าเฉลี่ยกลุ่ม 2 และ 3 ในเทอมของสัมประสิทธิ์การเปรียบเทียบจะเขียนได้ว่า

	c_1	c_2	c_3
$\psi_4 : \mu_1 - \mu_2/2 - \mu_3/2 = 0$	+1	−1/2	−1/2

สามารถเขียนในรูปของสัมประสิทธิ์การเปรียบเทียบได้ว่า

$$\psi_4 = (+1)\mu_1 + (-1/2)\mu_2 + (-1/2)\mu_3$$

$$= \mu_1 - \mu_2/2 - \mu_3/2$$

จำนวนของการเปรียบเทียบที่ซับซ้อนจะมีจำนวนมากกว่าหรือเท่ากับ [J(J-1)]/2 เมื่อ J ควรจะมีอย่างน้อย 3 กลุ่ม ในอีกกรณีหนึ่ง จำนวนของการเปรียบเทียบมีจำนวนมากเมื่อมีกลุ่มตั้งแต่ 3 กลุ่มขึ้นไป จำนวนโดยรวมของการเปรียบเทียบรายคู่และซับซ้อนจะเท่ากับ $[1 + ((3^J - 1)/2) - 2^J]$ ดังนั้นสำหรับ J = 4 ควรจะมีการเปรียบเทียบรายคู่และซับซ้อนจำนวน 25 ชุด

กระบวนการเปรียบเทียบพหุคูณอยู่บนพื้นฐานเดียวกันในการทดสอบทางสถิติ แม้ว่ากระบวนการเหล่านี้จะไม่อยู่บนพื้นฐานค่าวิกฤติเดียวกัน ซึ่งสถิติแรกที่จะแนะนำคือ "standard t" อัตราส่วนมาตรฐาน t สำหรับการเปรียบเทียบมีว่า

$$t = \frac{\psi'}{s_{\psi'}}$$

เมื่อ $$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2 / n_j)]\}}$$

สัญลักษณ์ ' บ่งชี้ว่า การเปรียบเทียบอยู่บนพื้นฐานของค่าเฉลี่ยของกลุ่มตัวอย่าง $s_{\psi'}$ แสดงถึงค่าความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบและ n_j แสดงถึงจำนวนของค่าสังเกตในกลุ่ม j ถ้าจำนวนของค่าสังเกตต่อกลุ่มมีจำนวนเท่ากันซึ่งเท่ากับ n แล้ว รูปแบบโดยทั่วไปของความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบสามารถเขียนในรูปอย่างง่ายได้ว่า

$$s_{\psi'} = \sqrt{\{[MS_{error} / n][\Sigma(c_j^2)]\}}$$

สำหรับการเปรียบเทียบรายคู่ค่าความคลาดเคลื่อนมาตรฐานสามารถเขียนในรูปอย่างง่ายได้ว่า

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_1) + (1/n_2)]\}}$$

โดยทั่วไปกรณีที่มีจำนวนเท่ากันทั้งสองกลุ่มที่นำมาเปรียบเทียบจะเขียนได้ว่า

$$s_{\psi'} = \sqrt{\{2MS_{error} / n\}}$$

ถ้าไม่ต้องการเปรียบเทียบเฉพาะคู่ที่สนใจแล้วจะใช้รูปแบบทั่ว ๆ ไปในการเปรียบเทียบ

การเปรียบเทียบแบบวางแผนกับการเปรียบเทียบภายหลัง (Planned Versus Posthoc Comparisons)

ในหัวข้อนี้เป็นการอธิบายชนิดของการเปรียบเทียบ แบ่งออกเป็นการเปรียบเทียบแบบวางแผนที่ดำเนินการก่อนทำการวิจัยและการเปรียบเทียบภายหลังจากการทดสอบ F-test เรียบร้อยแล้ว การเปรียบเทียบแบบวางแผน จะเกี่ยวข้องกับการเปรียบเทียบเพื่อตรวจสอบก่อนที่จะเก็บรวบรวมข้อมูล การเปรียบเทียบแบบวางแผนนี้จะอยู่บนพื้นฐานของทฤษฎี การวิจัยนำร่อง และการตั้งสมมติฐาน นักวิจัยที่สนใจจะเปรียบเทียบแบบวางแผน โดยปกติจะมีจำนวนชุดของการเปรียบเทียบน้อย การเปรียบเทียบแบบวางแผนจะเกี่ยวข้องไปถึงการทดสอบ F-test ในบางกรณีที่นักวิจัยสนใจเปรียบเทียบแบบวางแผน แต่ไม่ใช้ F-test ในการตรวจสอบการเปรียบเทียบที่เป็นไปได้ทั้งหมด ในกรณีนี้นักวิจัยควรจะเอาใจใส่เกี่ยวกับการเปรียบเทียบที่น่าจะเป็นไปได้ และแน่นอนว่ามีการเปรียบเทียบเพียงไม่กี่คู่เท่านั้นที่สนใจ นอกจากนี้ นักวิจัยอาจจะไม่สนใจการเกิดระดับความคลาดเคลื่อนแบบ the family-wise ในการเปรียบเทียบแบบวางแผน เพราะว่าจะมีไม่กี่คู่เท่านั้นที่เปรียบเทียบ การเปรียบเทียบแบบวางแผนมีจำนวนน้อยกว่าการเปรียบเทียบภายหลัง การเปรียบเทียบแบบวางแผนโดยทั่วไปจะมีช่วงความเชื่อมั่นที่แคบ มีอำนาจการทดสอบที่มากกว่า และมีความคลาดเคลื่อนแบบที่ 1 มากกว่าการเปรียบเทียบภายหลัง

การเปรียบเทียบภายหลังจะมีสูตรที่ผู้วิจัยสามารถดำเนินการในการทดสอบได้ ซึ่งการเปรียบเทียบนี้จะมีอยู่หลายชนิดและการทดสอบนัยสำคัญนั้นจะใช้ F-test เป็นหลัก คำว่า Posthoc เป็นภาษาละติน แปลว่า "หลังจากข้อเท็จจริง" (after the fact) ดังนั้นเป็นการอ้างอิงถึงการเปรียบเทียบที่ทดสอบหลังจาก F-test ใน ANOVA มีนัยสำคัญ

โรเซนทอล และโรสนาว (Resenthal and Rosnow) ได้อธิบายถึงวิธีการเปรียบเทียบสำหรับนักวิจัยที่สนใจในความแตกต่างระหว่างค่าเฉลี่ยของกลุ่ม การทดสอบ F-test นิยามโดย โรเซนทอล และโรสนาว ว่าเป็นวิธีสำหรับผู้วิจัยที่ไม่ได้สนใจหรือไม่เน้นที่จะตอบคำถามเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยของกลุ่ม นักสถิติเช่น เชฟเฟ่ (Scheffe), โรเซนทอลและโรสนาว (Rosenthal and Rosnow) ฮอคเบิร์ก และแทมแฮน (Hochberg and Tamhane) และวิลคอก (Wilcox) เชื่อว่า การเปรียบเทียบพหุคูณ (MCP) ไม่ได้มีเงื่อนไขบนความมีนัยสำคัญของ F-test นั่นคือผลของการทดสอบ F-test ไม่เกี่ยวข้องกับการเปรียบเทียบรายคู่ วิลคอก (Wilcox) ได้แสดงถึงการได้มาของวิธีการเปรียบเทียบรายคู่โดยมากจะไม่ได้อยู่บนพื้นฐานของข้อตกลงเบื้องต้นที่ว่า F-test ต้องมีนัยสำคัญ ในข้อเท็จจริง เบิร์นฮาร์ดสัน (Bernhardson) ได้แสดงให้เห็นว่า หากใช้เฉพาะการเปรียบเทียบพหุคูณหลังจาก F มีนัยสำคัญแล้ว ระดับความคลาดเคลื่อนแบบที่ 1 จะมีค่าน้อยกว่า และจะมีอำนาจการทดสอบน้อยกว่าถ้าใช้การเปรียบเทียบพหุคูณโดยปราศจากการทดสอบ F-test ในเบื้องต้นได้แนะนำว่า วิธีการเปรียบเทียบรายคู่ด้วยวิธีเชฟเฟ่ จะไม่อยู่บนพื้นฐานของการมีหรือไม่มีนัยสำคัญของ F-test

ระดับความคลาดเคลื่อนแบบที่ 1 (The Type I Error Rate)

มีนักวิจัยที่คิดเกี่ยวกับระดับของความคลาดเคลื่อนแบบที่ 1 แบบ the family-wise ซึ่งขึ้นอยู่กับการเลือกวิธีการเปรียบเทียบพหุคูณ อาจจะมีเพียง α เดียวสำหรับในแต่ละการเปรียบเทียบ หรือชุดของ α สำหรับการเปรียบเทียบทั้งหมด วิธีการเปรียบเทียบพหุคูณในกลุ่มนี้จะอยู่บนพื้นฐานของการเปรียบเทียบ (contrast - based) เราจะออกแบบระดับของ α สำหรับวิธีการที่อยู่บนพื้นฐานของการเปรียบเทียบ (contrast - based procedure) จะเรียกว่า α_{pc} ซึ่งจะแสดงถึงระดับความคลาดเคลื่อนแบบที่ 1 ต่อหนึ่งชุดของการเปรียบเทียบ ดังนั้น α_{pc} คือความน่าจะเป็นที่จะเกิดระดับความคลาดเคลื่อนแบบที่ 1 สำหรับชุดของการเปรียบเทียบ ถัดมาเรียกว่า α คือชุดสำหรับการเปรียบเทียบทั้งหมด ซึ่งวิธีการเปรียบเทียบพหุคูณในกลุ่มนี้จะเรียก the family-wise ซึ่งระดับ α สำหรับ family-wise ใช้สัญลักษณ์ว่า α_{fw} ซึ่งจะแสดงถึง the family-wise Type I error rate ดังนั้น α_{fw} คือความน่าจะเป็นที่เกิดความคลาดเคลื่อนแบบที่ 1 อย่างน้อย 1 ชุดในชุดของการเปรียบเทียบทั้งหมด สำหรับการเปรียบเทียบแบบอิสระ (orthogonal contrasts) จะเขียนได้ว่า

$$\alpha_{fw} = 1 - (1 - \alpha_{pc})^c$$

เมื่อ c คือจำนวนของการเปรียบเทียบอิสระ และ c = J - 1 การเปรียบเทียบอิสระจะอธิบายในหัวข้อถัดไป สำหรับการเปรียบเทียบอิสระนั้น

$$\alpha_{fw} \leq c\alpha_{pc}$$

ซึ่งระดับของความคลาดเคลื่อนแบบที่ 1 นี้จะเป็นไปในทำนองเดียวกับการทดสอบ t-test ที่เป็นอิสระจากกันหลายครั้ง

การเปรียบเทียบอิสระ (Orthogonal Contrasts)

การนิยามการเปรียบเทียบอิสระ คือชุดของการเปรียบเทียบที่เป็นอิสระ ถ้าชุดการเปรียบเทียบเหล่านั้นไม่ซ้ำซ้อนกันและเป็นอิสระจากกันในแหล่งของความแปรปรวน (ถ้าโดยปกติแล้วต้องเป็นไปตามข้อตกลงเบื้องต้นของ ANOVA) อีกกรณีหนึ่ง สารสนเทศหรือผลลัพธ์ที่ได้จากการเปรียบเทียบไม่ซ้ำซ้อนกันและอิสระจากกัน สำหรับกลุ่ม J กลุ่ม จะมีการเปรียบเทียบอิสระเพียง J - 1 ชุดเท่านั้น

นำมาทำให้ง่ายขึ้น โดยจะพิจารณากรณีแรกก่อนคือ n ของแต่ละกลุ่มเท่ากัน จะมีการเปรียบเทียบ 2 ชุดที่เป็นอิสระจากกัน ถ้าผลของสัมประสิทธิ์การเปรียบเทียบมีค่าศูนย์ โดยมากการเปรียบเทียบ 2 ชุดจะเป็นอิสระจากกันถ้า

$$\Sigma(c_j c_{j'}) = c_1 c_{1'} + c_2 c_{2'} + \ldots + c_j c_{j'} = 0$$

เมื่อ j และ j' ตัวห้อยจะแสดงถึงการเปรียบเทียบ 2 ชุดที่แตกต่างกัน ดังนั้นจะเห็นว่าความเป็นอิสระขึ้นอยู่กับสัมประสิทธ์การเปรียบเทียบ เช่น ถ้า J = 3 แล้วสามารถเขียนในชุดของการเปรียบเทียบที่อิสระจากกันได้ว่า

	c_1	c_2	c_3	
$\psi_1 : \mu_1 - \mu_2 = 0$	+1	−1	0	
$\psi_2 : \mu_1/2 - \mu_2/2 - \mu_3 = 0$	+1/2	+1/2	−1	
	+1/2 +	−1/2 +	0	= 0

ชุดของการเปรียบเทียบ 2 ชุดจะไม่เป็นอิสระจากกันถ้า

	c_1	c_2	c_3	
$\psi_3 : \mu_1 - \mu_2 = 0$	+1	−1	0	
$\psi_4 : \mu_1 - \mu_3 = 0$	+1	0	−1	
	+1 +	0 +	0	= +1

ถ้าผลรวมของผลคูณในสัมประสิทธิ์การเปรียบเทียบทั้ง 2 ชุดเท่ากับศูนย์แล้ว นิยามว่าเป็นชุดของการเปรียบเทียบที่เป็นอิสระจากกัน (mutually orthogonal set of contrasts)

พิจารณาสำหรับบางกรณีเมื่อมีกลุ่ม 3 กลุ่ม และเราได้ตัดสินใจจะเปรียบเทียบ 3 ชุด และไม่สามารถจะทำให้เป็นอิสระจากกันได้ การเปรียบเทียบอาจจะได้รูปแบบว่า

	c_1	c_2	c_3
$\psi_1 : \mu_1 - \mu_2 = 0$	+1	−1	0
$\psi_2 : \mu_2 - \mu_3 = 0$	0	+1	−1
$\psi_3 : \mu_1 - \mu_2 = 0$	+1	0	−1

สมมติว่าค่าเฉลี่ยของกลุ่ม $\mu_1 = 30$, $\mu_2 = 24$ และ $\mu_3 = 20$ จะพบว่า $\psi_1 = 6$ สำหรับการเปรียบเทียบคู่แรก และ $\psi_2 = 4$ สำหรับการเปรียบเทียบคู่สอง เพราะว่าการเปรียบเทียบไม่เป็นอิสระจากกันและโดยรวมแล้วข้อมูลที่ได้จะซ้ำซ้อนกัน $\psi_3 = 10$ สำหรับการเปรียบเทียบคู่สาม ดังนั้นในการเปรียบเทียบคู่ที่ 3 ไม่ได้ช่วยให้ข้อมูลเพิ่มเติมใด ๆ มากกว่าการเปรียบเทียบ 2 คู่แรก

สุดท้ายกรณีที่ n แต่ละกลุ่มไม่เท่ากันแล้ว การเปรียบเทียบ 2 ชุดจะเป็นอิสระถ้า

$$\Sigma[(c_j c_{j'})/n_j] = 0$$

ตัวส่วน n_j ทำให้มีความยากมากกว่า

โดยสรุปแล้วหัวข้อนี้จะแนะนำถึงแนวคิดที่สำคัญของวิธีการเปรียบเทียบพหุคูณ ในแนวคิดใหม่นี้จะนิยามการเปรียบเทียบอย่างง่ายและซับซ้อน การเปรียบเทียบวางแผนและภายหลังและการเปรียบเทียบแบบอิสระ ซึ่งจะอธิบายในแต่ละวิธีในรายละเอียดต่อไป

การเลือกวิธีการเปรียบเทียบพหุคูณ

ในหัวข้อนี้จะอธิบายถึงการเลือกวิธีการเปรียบเทียบพหุคูณ (MCP) ซึ่งจะนำเสนอวิธีการที่ดีที่สุดที่มีประโยชน์ (Utility) เป็นที่นิยม (Popularity) และมีความสามารถในการการควบคุมระดับความคลาดเคลื่อนแบบที่ 1 และ 2 (Control of Type I and II error rate) สำหรับวิธีอื่น ๆ

จะอธิบายถึงเพียงคร่าว ๆ เท่านั้น วิธีที่จะนำเสนอนี้เป็นการเปรียบเทียบแบบวางแผน การเปรียบเทียบแบบภายหลัง โดย จะอธิบายลักษณะเฉพาะของแต่ละวิธี การทดสอบทางสถิติกับตัวอย่างที่ใช้

ในแต่ละวิธีของ MCP มีข้อตกลงเบื้องต้นร่วมกัน คือ การแจกแจงต้องเป็นโค้งปกติ (normality) มีความเป็นเอกพันธ์ของความแปรปรวน (homogeneity of variance) และมีความเป็นอิสระของกลุ่มตัวอย่าง (independence of observations) ในบางวิธีการจะมีข้อตกลงเบื้องต้นเพิ่มเติม (เช่น จำนวน n แต่ละกลุ่มต้องเท่ากัน) MCP ส่วนใหญ่จะเป็นสมมติฐานแบบสองทาง มีบาง MCP ที่สามารถจะใช้สมมติฐานแบบทางเดียว โดยทั่วไป MCP จะมีความแกร่งต่อการไม่เป็นโค้งปกติ (nonnormality) (แต่ต้องไม่มีค่าผิดสังเกต) แต่ความไม่แกร่งของสถิติจะขึ้นอยู่กับความเป็นเอกพันธ์ของความแปรปรวนหรือความเป็นอิสระของกลุ่มตัวอย่าง

การวิเคราะห์แนวโน้ม (Trend Analysis)

การวิเคราะห์แนวโน้มเป็น MCP แบบวางแผนที่มีประโยชน์เมื่อกลุ่มแต่ละกลุ่มแสดงถึงความแตกต่างกันเชิงปริมาณที่แบ่งออกเป็นระดับต่าง ๆ (ตัวแปรอิสระอยู่ในระดับช่วงหรืออัตราส่วน) ตัวอย่างเช่น ตัวแปรอิสระอาจจะเป็นอายุ ปริมาณการใช้ยา และระยะเวลาที่ใช้ เป็นต้น ผู้วิจัยมีความสนใจในค่าเฉลี่ยของกลุ่มที่แปรเปลี่ยนในเชิงปริมาณของตัวแปรอิสระ ในหัวข้อนี้จะนิยาม การวิเคราะห์แนวโน้มในรูปของ orthogonal polynomials และสมมติว่าระดับของตัวแปรอิสระมีช่วงห่างเท่ากัน และจำนวนของค่าสังเกตต่อกลุ่มเท่ากัน

การเปรียบเทียบแบบ orthogonal polynomial ใช้การทดสอบ t-test ซึ่งเปรียบเทียบกับค่าวิกฤติ $\pm_{1-\alpha/2}t_{df(error)}$ โดยเปิดจากตาราง t ในภาคผนวก การเปรียบเทียบ orthogonal polynomial ได้รวมสองแนวคิดเข้าด้วยกัน คือ orthogonal contrasts และ polynomial regression สำหรับกลุ่ม J กลุ่ม มีเฉพาะ J - 1 กลุ่มที่เปรียบเทียบแบบอิสระได้ ส่วน polynomial regression นั้น จะอยู่ในรูปของโมเดลแนวโน้มเชิงเส้น (a linear trend) แนวโน้มควอดราติก (a quadratic trend) และแนวโน้มคิวบิค (a cubic trend) เขียนได้ว่า

$$\overline{Y}_{.j} = b_0 + b_1X + b_2X^2 + \ldots + b_{(J-1)}X^{(J-1)}$$

เมื่อ b_0 แสดงถึงจุดตัดบนแกน Y และ b แสดงถึงสัมประสิทธิ์แนวโน้ม

ดังนั้นรูปแบบของการเปรียบเทียบคือ การเปรียบเทียบชุดแรกจะใช้ linear trend ชุดที่สองใช้ quadratic trend และชุดที่สามใช้ cubic trend ดังนั้นสำหรับกลุ่ม J กลุ่ม จะมีชุดของ polynomial สูงสุดเป็น J - 1 ชุด และสำหรับกลุ่ม 4 กลุ่ม ชุดของ orthogonal contrasts จะใช้ linear, quadratic และ cubic trend โดยการวิเคราะห์จะเชื่อมโยงระหว่างการวิเคราะห์ความแปรปรวนและการวิเคราะห์การถดถอย

สัมประสิทธิ์การเปรียบเทียบสำหรับ linear, quadratic และ cubic trends คือ

	c_1	c_2	c_3	c_4
linear :	-3	-1	$+1$	$+3$
quadratic :	$+1$	-1	-1	$+1$
cubic :	-1	$+3$	-3	$+1$

เมื่อการเปรียบเทียบสามารถเขียนได้เป็น

$$\psi_{\text{linear}} = (-3)\mu_1 + (-1)\mu_2 + (+1)\mu_3 + (+3)\mu_4$$

$$\psi_{\text{quadratic}} = (+1)\mu_1 + (-1)\mu_2 + (-1)\mu_3 + (+1)\mu_4$$

$$\psi_{\text{cubic}} = (-1)\mu_1 + (+3)\mu_2 + (-3)\mu_3 + (+1)\mu_4$$

สัมประสิทธิ์การเปรียบเทียบสามารถเปิดได้ในภาคผนวก สำหรับจำนวนของค่าที่แตกต่างกันของ J หากมองหาในตารางสัมประสิทธิ์การเปรียบเทียบสำหรับค่า J ที่มากกว่า 5 ก็จะเห็นสัมประสิทธิ์สำหรับ polynomial ในลำดับอื่น ๆ ที่สูงกว่า ตัวอย่างเช่น J = 6 สัมประสิทธิ์จะมีไปจนถึง quintic trend สัมประสิทธิ์เหล่านี้ได้มาโดยเปิดจากตาราง และการทดสอบในลำดับที่สูงกว่าของ polynomial โดยปกติจะไม่เป็นที่สนใจสำหรับนักวิจัย เป็นการยากที่จะใช้ polynomial ในลำดับอื่น ๆ ที่สูงขึ้นไปกว่า cubic เพราะเป็นการยากที่นักวิจัยจะทำความเข้าใจและแปลความหมาย การเปรียบเทียบจะถูกทดสอบเป็นลำดับเริ่มต้นจาก linear trend ไปจนถึง trend ที่สูงกว่า

เพื่อเพิ่มความเข้าใจยิ่งขึ้น ระดับของ polynomial เป็นจำนวนครั้งของการเปลี่ยนสัญลักษณ์สำหรับการเปรียบเทียบ ในการเปรียบเทียบเชิงเส้นจะแสดงในรูปอย่างง่าย นั่นคือมีระดับแรกของ polynomial (เช่น ใช้ X^1 หรือ X) และสัญลักษณ์ที่เปลี่ยนแปลงไป 1 ครั้ง ในการเปรียบเทียบ quadratic จะมีระดับ polynomial ที่สอง (เช่น ใช้ X^2) สัญลักษณ์จะเปลี่ยนไป 2 ครั้ง ในการเปรียบเทียบ cubic เป็นระดับที่สาม (เช่น ใช้ X^3) เมื่อสัญลักษณ์เปลี่ยนแปลงไป 3 ครั้ง

ใช้ตัวอย่างในการวิเคราะห์ความแปรปรวนแบบทางเดียว โดยการทดสอบ linear, quadratic และ cubic trends การวิเคราะห์แนวโน้มจะใช้ได้กับข้อมูลตัวอย่างในเนื้อหาที่ผ่านมา เพราะว่าตัวแปรอิสระในตัวอย่างเป็นความแตกต่างเชิงปริมาณ ซึ่ง J = 4 สามารถใช้สัมประสิทธิ์การเปรียบเทียบที่นำเสนอไว้ข้างต้นได้ สามารถคำนวณได้ดังนี้

ค่าวิกฤติ

$$\pm_{1-\alpha/2}t_{df(error)} = \pm_{.975}t_{28} = \pm 2.05$$

ความคลาดเคลื่อนมาตรฐานของ linear trend

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2/n_j)]\}}$$

$$= \sqrt{36.11[(9/8)+(1/8)+(1/8)+(9/8)]}$$

$$= 9.50$$

ความคลาดเคลื่อนมาตรฐานสำหรับ quadratic trend

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2 / n_j)]\}}$$
$$= \sqrt{36.11[(1/8)+(1/8)+(1/8)+(1/8)]}$$
$$= 4.25$$

ความคลาดเคลื่อนมาตรฐานสำหรับ cubic trend

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2 / n_j)]\}}$$
$$= \sqrt{36.11[(1/8)+(9/8)+(9/8)+(1/8)]}$$
$$= 9.50$$

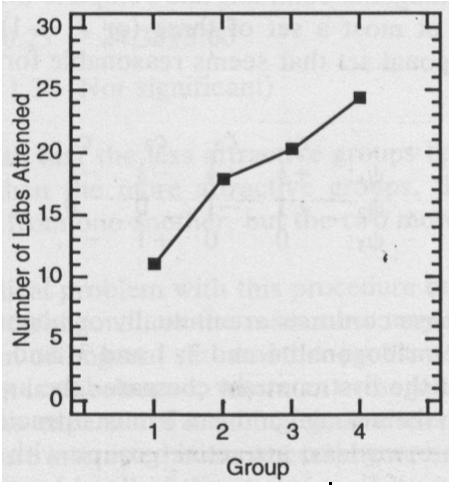
การทดสอบทางสถิติ

$$t_1 = (-3\overline{Y}_1 - 1\overline{Y}_2 + 1\overline{Y}_3 + 3\overline{Y}_4)/s_{\psi'}$$
$$= (-3(11.13) - 1(17.88) + 1(20.25) + 3(24.38))/9.50$$
$= 4.43$ (มีนัยสำคัญ)

$$t_2 = (1\overline{Y}_1 - 1\overline{Y}_2 - 1\overline{Y}_3 + 1\overline{Y}_4)/s_{\psi'}$$
$$= (1(11.13) - 1(17.88) - 1(20.25) + 1(24.38))/4.25$$
$= -0.62$ (ไม่มีนัยสำคัญ)

$$t_3 = (-1\overline{Y}_1 + 3\overline{Y}_2 - 3\overline{Y}_3 + 1\overline{Y}_4)/s_{\psi'}$$
$$= (-1(11.13) + 3(17.88) - 3(20.25) + 1(24.38))/9.50$$
$= 0.64$ (ไม่มีนัยสำคัญ)

ผลการเปรียบเทียบ พบการมีนัยสำคัญของ linear trend แต่ไม่สูงไปกว่า trend ในระดับอื่น ๆ เป็นสิ่งที่ไม่น่าแปลกใจ หากเมื่อลองนำค่าเฉลี่ยมาพล็อตกราฟ ดังภาพประกอบ 1 จะเห็นว่าแผนภาพมีลักษณะเกือบจะเป็น linear trend หรือในอีกกรณีหนึ่ง การเพิ่มขึ้นของค่าเฉลี่ยคงที่ในแต่ละระดับของตัวแปรอิสระ การพล็อตค่าเฉลี่ยจะสามารถช่วยในการแปลความหมายการเปรียบเทียบได้



ภาพประกอบ 1 แสดงค่าเฉลี่ยของกลุ่ม

การเปรียบเทียบอิสระแบบวางแผน (Planned Orthogonal Contrasts)

การเปรียบอิสระแบบวางแผนคือ MCP ที่เปรียบเทียบก่อนดำเนินการวิจัยและชุดของการเปรียบเทียบแต่ละชุดเป็นอิสระจากกัน วิธี POC (Planned Orthogonal Contrasts) เป็นการเปรียบเทียบที่ผู้วิจัยไม่ชัดเจนเกี่ยวกับระดับความคลาดเคลื่อนแบบ family-wise ชุดของการเปรียบเทียบแต่ละชุดเป็นอิสระจากกัน ดังนั้นจำนวนของการเปรียบเทียบจะมีน้อย

การคำนวณนั้น จะใช้ t-test ในการเปรียบเทียบกับค่าวิกฤต $\pm_{1-\alpha/2}t_{df(error)}$ ที่หาได้จากตาราง t ในภาคผนวก ใช้ตัวอย่างข้อมูลข้างต้นเราจะได้ชุดของการเปรียบเทียบอิสระและมีความสมบูรณ์ในการคำนวณ โดยมี J = 4 พบว่ามีชุดเปรียบเทียบ 3 ชุด (J - 1) ที่เป็นอิสระกัน ดังนี้

	c_1	c_2	c_3	c_4
ψ_1 :	+1/2	+1/2	−1/2	−1/2
ψ_2 :	+1	−1	0	0
ψ_3 :	0	0	+1	−1

เราจะเห็นว่าการเปรียบเทียบชุดแรกค่าเฉลี่ยจะมีอย่างน้อย 2 กลุ่มที่มีเครื่องหมายลบและอีก 2 กลุ่มมีเครื่องหมายบวกในการเปรียบเทียบ การเปรียบเทียบชุดสองจะมีสองกลุ่มแรกที่เปรียบเทียบกัน ส่วนการเปรียบเทียบชุดที่สามจะมี 2 กลุ่มหลังที่เปรียบเทียบกัน ในตัวอย่างเป็นออกแบบที่สมดุล (มีจำนวน n เท่ากันในทุกกลุ่ม) ในการคำนวณค่าวิกฤตจะได้

$$\pm_{1-\alpha/2}t_{df(error)} = \pm_{.975}t_{28} = \pm 2.05$$

ความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบชุดที่ 1

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2/n_j)]\}}$$

$$= \sqrt{\{36.11[(1/4)/8+(1/4)/8+(1/4)/8+(1/4)/8]\}}$$

$$= 2.12$$

ความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบชุดที่ 2 และ 3

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_j)+(1/n_{j'})]\}}$$
$$= \sqrt{\{36.11[1/8+1/8]\}}$$
$$= 3.00$$

สถิติทดสอบ

$$t_1 = [(1/2)\overline{Y}_1 + (1/2)\overline{Y}_2 - (1/2)\overline{Y}_3 - (1/2)\overline{Y}_4]/s_{\psi'}$$
$$= [(1/2)11.13+(1/2)17.88-(1/2)(20.25)-(1/2)24.38]/2.12$$
$$= -3.68 \text{ (มีนัยสำคัญ)}$$
$$t_2 = (\overline{Y}_1 - \overline{Y}_2)/s_{\psi'}$$
$$= (11.13-17.88)/3.00$$
$$= -2.25 \text{ (มีนัยสำคัญ)}$$
$$t_3 = (\overline{Y}_3 - \overline{Y}_4)/s_{\psi'}$$
$$= (20.25-24.38)/3.00$$
$$= -1.37 \text{ (ไม่มีนัยสำคัญ)}$$

ผลที่ได้บ่งชี้ว่ามีค่าเฉลี่ยของ 2 กลุ่มแรกที่แตกต่างอย่างมีนัยสำคัญทางสถิติจากค่าเฉลี่ยของสองกลุ่มหลัง และมีความแตกต่างระหว่างค่าเฉลี่ยของกลุ่มที่หนึ่งและกลุ่มที่สอง แต่กลุ่มที่สามและสี่มีค่าเฉลี่ยไม่แตกต่างกัน

ปัญหาที่ใช้เป็นตัวอย่างจะเหมาะกับวิธีการนี้เพราะว่าการเปรียบเทียบที่สนใจอาจจะไม่อิสระจากกัน หรือนักวิจัยอาจจะไม่สนใจในการเปรียบเทียบทั้งหมดที่เป็นอิสระจากกัน ในอีกปัญหาหนึ่งจะกล่าวถึงคือกรณีที่การวิจัยออกแบบไม่สมดุล คือมี n แต่ละกลุ่มไม่เท่ากัน ซึ่งชุดของการเปรียบเทียบอิสระอาจจะถูกสร้างขึ้นมาให้เปรียบเทียบกันอย่างมีความหมาย มีคำแนะนำง่าย ๆ กรณีการเปรียบเทียบที่สนใจไม่เป็นอิสระแล้วจะใช้ MCP วิธีการอื่น ถ้าออกแบบการวิจัยให้ไม่สมดุลย์และการเปรียบเทียบไม่อิสระแล้วจะใช้ MCP วิธีการอื่นแล้วแต่กรณี ถ้าออกแบบ MCP แบบวางแผนแล้ว จะมีเฉพาะวิธีของ Dunnet หรือ Dun (Bonferroni) เท่านั้นที่ใช้ได้

โดยนิยามของ POC นั้นเป็นวิธีการเปรียบเทียบพื้นฐาน ที่ควรจะเปรียบเทียบแบบ family-wise อีกทางเหลือหนึ่งเมื่อระดับ α_{pc} ถูกหารด้วยจำนวนของชุดเปรียบเทียบ วิธีการนี้นิยามว่า $\alpha_{pc} = \alpha_{fw}/c$ เมื่อ c คือจำนวนของการเปรียบเทียบอิสระ (เช่น c = J - 1) ในแนวคิดจากวิธีการของ Dunn (Bonferroni) นั่นคือมีการโต้แย้งเกี่ยวกับการใช้ POC กับการนิยาม α อีกทางเลือกหนึ่ง ถ้าความแปรปรวนไม่เท่ากันในแต่ละกลุ่มแล้ว มีวิธีการประมาณค่าหลากหลายที่จะสามารถนำไปสู่การประมาณค่าความแปรปรวนของกลุ่มแต่ละกลุ่ม

วิธีการของดันเน็ตต์ (Dunnett's Method)

มีวิธีการ 3 วิธีสำหรับการเปรียบเทียบแบบวางแผน ซึ่งหนึ่งในนั้นคือ Dunnett ถูกออกแบบมาเพื่อเปรียบเทียบเป็นรายคู่เมื่อมีกลุ่มอ้างอิง (เช่น กลุ่มควบคุม) ไปเปรียบเทียบกับกลุ่มอื่น ๆ ที่เหลืออีก J - 1 กลุ่ม ดังนั้นชุดของการเปรียบเทียบรายคู่จะเฉพาะเจาะจง วิธีของดันเน็ตต์ ใช้สถิติทดสอบ t ที่มีความคลาดเคลื่อนมาตรฐานเท่ากับ

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_j) + (1/n_c)]\}}$$

เมื่อ c คือกลุ่มอ้างอิง และ j คือกลุ่มที่ถูกนำมาเปรียบเทียบ สถิติทดสอบจะเปรียบเทียบกับค่าวิกฤต $\pm_{1-\alpha/2}t_{df(error),J}$ ที่อ้างอิงจากตารางดันเน็ตต์ในภาคผนวก

ใช้ตัวอย่างที่อ้างอิงไว้ข้างต้น เปรียบเทียบกลุ่มที่ไม่มีผู้สอนแล็บ (กลุ่มอ้างอิง) กับกลุ่มอื่น ๆ ที่เหลืออีก 3 กลุ่ม แล้วคำนวณดังนี้

$$\pm_{1-\alpha/2}t_{df(error),J} = \pm_{.975}t_{28,4} = \pm 2.48$$

ความคลาดเคลื่อนมาตรฐาน

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_j) + (1/n_c)]\}}$$
$$= \sqrt{36.11[(1/8) + (1/8)]}$$
$$= 3.00$$

สถิติทดสอบ

$$t_1 = (\overline{Y}_1 - \overline{Y}_2)/s_{\psi'}$$
$$= (11.13-17.88)/3.00$$
$$= -2.25$$ (ไม่มีนัยสำคัญ)

$$t_2 = (\overline{Y}_1 - \overline{Y}_3)/s_{\psi'}$$
$$= (11.13-20.25)/3.00$$
$$= -3.04$$ (มีนัยสำคัญ)

$$t_3 = (\overline{Y}_1 - \overline{Y}_4)/s_{\psi'}$$
$$= (11.13-24.38)/3.00$$
$$= -4.41$$ (มีนัยสำคัญ)

ในที่นี้จะเห็นว่ากลุ่มที่สองไม่มีนัยสำคัญทางสถิติเมื่อเทียบกับกลุ่มควบคุม แต่กลุ่มที่สามและสี่มีนัยสำคัญทางสถิติแตกต่างจากกลุ่มอ้างอิง

ถ้าความแปรปรวนของกลุ่มอ้างอิงแตกต่างจากความแปรปรวนของกลุ่มอื่น ๆ J - 1 กลุ่ม ในทางประยุกต์ใช้วิธีการนี้จะอ่อนไหวน้อยเมื่อความแปรปรวนของกลุ่มแต่ละกลุ่มไม่เท่ากัน

วิธีของดันน์ (Dunn's Method : Bonferroni)

กระบวนการของดันน์ (Dunn) หรือเรียกอีกชื่อหนึ่งว่ากระบวนการของบอนเฟอร์โนนี (Bonferroni) เป็น MCP วางแผนแบบ family-wise เป็นการออกแบบการทดสอบเป็นรายคู่หรือ

การเปรียบเทียบที่ซับซ้อนสำหรับการออกแบบที่สมดุลหรือไม่สมดุล ซึ่งวิธีนี้จะมีความซับซ้อนมาก วิธีของดันน์ใช้สถิติทดสอบ t ระดับ α จะถูกแบ่งตามชุดของการเปรียบเทียบแบบวางแผน ระดับของ α ที่เปรียบเทียบจะมีค่า α/c เมื่อ c คือจำนวนของการเปรียบเทียบ ดังนั้น $\alpha_{pc} = \alpha_{fw}/c$ อธิบายสูตรนี้ ระดับความคลาดเคลื่อนแบบ family-wise จะเท่ากับ α ตัวอย่าง ถ้า α_{fw} = 0.05 ถูกออกแบบและมีการเปรียบเทียบ 5 ชุดที่ต้องทดสอบแล้วการเปรียบเทียบที่ถูกทดสอบจะมีระดับนัยสำคัญเท่ากับ 0.01 โรเซนทาล (Rosenthal) และโรสนาว (Rosnow) ได้อธิบายว่า α ไม่จำเป็นที่จะต้องกระจายให้เท่ากันในทุกชุดของการเปรียบเทียบ ผลรวมของ α_{pc} จะเท่ากับ α_{fw}

ในการคำนวณ วิธีของดันน์จะใช้การทดสอบ t ซึ่งจะนำไปเปรียบเทียบกับค่าวิกฤติ $\pm_{1-\alpha/c}t_{df(error)}$ สำหรับการทดสอบแบบสองทางจากตารางในภาคผนวก ใช้ข้อมูลจากตัวอย่าง เราจะได้ชุดการเปรียบเทียบดังนี้

	c_1	c_2	c_3	c_4
ψ_1 :	+1/2	+1/2	−1/2	−1/2
ψ_2 :	+1	−1	0	0
ψ_3 :	0	0	+1	−1

คำนวณได้ดังนี้

ค่าวิกฤติ

$$\pm_{1-\alpha/c}t_{df(error)} = \pm_{1-.05/3}t_{28} = \pm 2.54$$

ความคลาดเคลื่อนมาตรฐานสำหรับการเปรียบเทียบที่ 1

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2/n_j)]\}}$$
$$= \sqrt{\{36.11[(1/4)/8+(1/4)/8+(1/4)/8+(1/4)/8]\}}$$
$$= 2.12$$

ความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบชุดที่ 2 และ 3

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_j)+(1/n_{j'})]\}}$$
$$= \sqrt{\{36.11[1/8+1/8]\}}$$
$$= 3.00$$

สถิติทดสอบ

$$t_1 = [(1/2)\overline{Y}_1+(1/2)\overline{Y}_2-(1/2)\overline{Y}_3-(1/2)\overline{Y}_4]/s_{\psi'}$$
$$= [(1/2)11.13+(1/2)17.88-(1/2)(20.25)-(1/2)24.38]/2.12$$
$$= -3.68 \text{ (มีนัยสำคัญ)}$$

$$t_2 = (\overline{Y}_1-\overline{Y}_2)/s_{\psi'}$$
$$= (11.13-17.88)/3.00$$
$$= -2.25 \text{ (ไม่มีนัยสำคัญ)}$$

$t_3 = (\overline{Y}_3 - \overline{Y}_4)/s_{\psi'}$

$= (20.25-24.38)/3.00$

$= -1.37$ (ไม่มีนัยสำคัญ)

สำหรับชุดของการเปรียบเทียบนี้ จะเหมือนกับผลที่ได้จากวิธีของ POC ยกเว้นการเปรียบเทียบชุดที่ 2 ซึ่งไม่มีนัยสำคัญ เหตุผลสำหรับความแตกต่างนี้ค่าที่คำนวณน้อยกว่าค่าวิกฤติจากตาราง โดยธรรมชาติในวิธีการของดันนั้นค่าวิกฤติจะมีค่าสูงกว่าวิธี POC ดังนั้นจะทำให้ยากในการปฏิเสธ H_0

วิธีของเชฟเฟ่ (Scheffe's Method)

กระบวนการของเชฟเฟ่สามารถใช้ได้กับทุก ๆ การเปรียบเทียบที่เป็นไปได้ทั้งหมด ทั้งการเปรียบเทียบที่เป็นอิสระหรือไม่เป็นอิสระ การเปรียบเทียบเป็นรายคู่หรือซับซ้อน การเปรียบเทียบแบบวางแผนหรือเปรียบเทียบหลัง ซึ่งจะสามารถควบคุมระดับความคลาดเคลื่อน the family-wise โดยทั่วไปวิธีของเชฟเฟ่จะแนะนำให้ใช้เปรียบเทียบหลัง วิธีการของเชฟเฟ่โดยทั่วไปจะเป็นการทดสอบแบบธรรมเนียมนิยม โดยเฉพาะการเปรียบเทียบเป็นรายคู่

วิธีของเชฟเฟ่จะสอดคล้องกับผลของอัตราส่วน F ในการวิเคราะห์ความแปรปรวน ถ้า F มีนัยสำคัญแล้วจะมีอย่างน้อย 1 คู่จากการเปรียบเทียบที่ผลการทดสอบเชฟเฟ่จะมีนัยสำคัญ อย่างไรก็ตามถ้าหากมีชุดของการเปรียบเทียบจำนวนมากแต่อาจจะไม่ต้องเปรียบเทียบถ้า F ไม่มีนัยสำคัญ แม้ว่าหากเปรียบเทียบแล้วผลจากเชฟเฟ่จะมีนัยสำคัญก็ตาม

การทดสอบของเชฟเฟ่จะใช้การทดสอบ t นั่นคือเปรียบเทียบกับค่าวิกฤติ $\sqrt{[(J-1)_{1-\alpha}F_{J-1,df(error)}]}$ และค่า F ได้จากตาราง F ในภาคผนวก สังเกตว่า รากที่สองของค่า F จะถูกปรับแก้ด้วย J - 1 ซึ่งจะช่วยเพิ่มค่าวิกฤติของเชฟเฟ่ให้มากขึ้นทำให้มีนัยสำคัญยากขึ้น

พิจารณาตัวอย่างข้างต้นโดยการเปรียบเทียบด้วยวิธีของเชฟเฟ่ ได้ชุดของการเปรียบเทียบดังนี้

	c_1	c_2	c_3	c_4
ψ_1 :	+1/2	+1/2	−1/2	−1/2
ψ_2 :	+1	−1	0	0
ψ_3 :	0	0	+1	−1

คำนวณค่าต่าง ๆ ได้ดังนี้

ค่าวิกฤติ

$$\sqrt{[(J-1)_{1-\alpha}F_{J-1,df(error)}]} = \sqrt{[(3)_{.95}F_{3,28}]}$$

$$= \sqrt{[(3)2.95]}$$

$$= 2.97$$

ความคลาดเคลื่อนมาตรฐานสำหรับการเปรียบเทียบชุดที่ 1

$$s_{\psi'} = \sqrt{\{MS_{error}[\Sigma(c_j^2/n_j)]\}}$$
$$= \sqrt{\{36.11[(1/4)/8+(1/4)/8+(1/4)/8+(1/4)/8]\}}$$
$$= 2.12$$

ความคลาดเคลื่อนมาตรฐานของการเปรียบเทียบชุดที่ 2 และ 3

$$s_{\psi'} = \sqrt{\{MS_{error}[(1/n_j)+(1/n_{j'})]\}}$$
$$= \sqrt{\{36.11[1/8+1/8]\}}$$
$$= 3.00$$

สถิติทดสอบ

$$t_1 = [(1/2)\overline{Y}_1 + (1/2)\overline{Y}_2 - (1/2)\overline{Y}_3 - (1/2)\overline{Y}_4]/s_{\psi'}$$
$$= [(1/2)11.13+(1/2)17.88-(1/2)(20.25)-(1/2)24.38]/2.12$$
$$= -3.68 \text{ (มีนัยสำคัญ)}$$
$$t_2 = (\overline{Y}_1 - \overline{Y}_2)/s_{\psi'}$$
$$= (11.13-17.88)/3.00$$
$$= -2.25 \text{ (ไม่มีนัยสำคัญ)}$$
$$t_3 = (\overline{Y}_3 - \overline{Y}_4)/s_{\psi'}$$
$$= (20.25-24.38)/3.00$$
$$= -1.37 \text{ (ไม่มีนัยสำคัญ)}$$

การใช้วิธีของเชฟเฟ่ ผลที่ได้จะเหมือนกับวิธีการของดันน์ แต่มีความแตกต่างกันในค่าวิกฤติที่วิธีของเชฟเฟ่จะมีค่าวิกฤติสูงกว่าของดันน์และวิธี POC (วิธีของเชฟเฟ่ ได้ค่า 2.97 วิธีของดันน์ ได้ค่า 2.54 และวิธี POC ได้ค่า 2.05) วิธีของเชฟเฟ่จะเรียกได้ว่าเป็นแบบเป็นธรรมเนียมนิยมมากที่สุด นั่นคือการยากต่อการปฏิเสธ H_0

สำหรับกรณีที่ความแปรปรวนของกลุ่มไม่เท่ากัน วิธีของเชฟเฟ่จะอ่อนไหวน้อยต่อการไม่เท่ากันของความแปรปรวนในแต่ละกลุ่ม

การทดสอบแอลเอสดีของฟิชเชอร์ (Fisher's LSD Test)

ฟิชเชอร์ ได้นำเสนอการทดสอบ LSD (Least Significant Difference test) ใช้ t ในการทดสอบ เป็นวิธีที่ถูกพัฒนาให้เป็นการเปรียบเทียบภายหลังเป็นรายคู่ หลังจากอัตราส่วน F มีนัยสำคัญในการวิเคราะห์ความแปรปรวน การทดสอบเป็นรายคู่ทั้งหมด (หรือบางคู่) จะถูกทดสอบด้วยการทดสอบ t แล้วนำมาเปรียบเทียบกับค่าวิกฤติ $\pm_{1-\alpha/2}t_{df(error)}$ ปัญหาของวิธีการนี้พบว่าระดับความคลาดเคลื่อนแบบ the family-wise จะเพิ่มขึ้นตามจำนวนของการเปรียบเทียบที่เพิ่มขึ้น นั่นคือมีการควบคุมความคลาดเคลื่อนแบบที่ 1 น้อยกว่า MCP แบบอื่น ๆ ซึ่งเป็นเหตุผลว่าการทดสอบ LSD ของฟิชเชอร์จึงไม่แนะนำให้ใช้

การทดสอบเฮทเอสดีของตูกี้ (Tukey's HSD Test)

การทดสอบ HSD (Honestly Significant Difference) ของตู้กี้ (Tukey) เป็นวิธีหนึ่งที่นิยมกันมากสำหรับการเปรียบเทียบพหุคูณแบบภายหลัง ถูกพัฒนาโดย จอห์น ตูกี้ (John Tukey) ในปี 1953 และมีความเหมาะสมมากหากใช้ในการเปรียบเทียบเป็นรายคู่ทั้งหมดเมื่อจำนวน n ต่อกลุ่มเท่ากัน (รูปแบบสมดุล) หัวข้อถัดไปจะเป็นการประยุกต์กรณีที่ n ต่อกลุ่มไม่เท่ากัน ในกรณีที่มีการเปรียบเทียบอย่างซับซ้อน HSD ก็จะไม่เหมาะสม

การทดสอบ HSD บางครั้งจะอ้างอิงว่าเป็น the studentized range test เพราะว่าอยู่บนพื้นฐานของการแจกแจงอย่างสุ่มของสถิติ studentized range statistic ถูกพัฒนาโดย Willian Sealy Gosset ขั้นตอนแรกในการวิเคราะห์ต้องเรียงลำดับค่าเฉลี่ยจากมาก ($\overline{Y}_1$) ไปหาน้อย ($\overline{Y}_J$) การทดสอบทางสถิติจะใช้สูตรทดสอบดังนี้

$$q_i = (\overline{Y}_j - \overline{Y}_{j'})/s_{\psi'}$$

เมื่อ $$s_{\psi'} = \sqrt{(MS_{error}/n)}$$

i เป็นคู่ที่นำมาเปรียบเทียบ j และ j' เป็นค่าเฉลี่ย 2 กลุ่มที่นำมาเปรียบเทียบ n จะเป็นจำนวนของค่าสังเกตต่อกลุ่ม (ข้อตกลงเบื้องต้นคือ n ต่อกลุ่มมีจำนวนเท่ากัน) ผลการทดสอบทางสถิติจะเปรียบเทียบกับค่าวิกฤติ ${}_{1-\alpha}q_{df(error),J}$ เมื่อ df_{error} เท่ากับ J(n - 1) ตารางค่าวิกฤติเปิดได้จากภาคผนวก

การเปรียบเทียบแรกจะเกี่ยวกับการทดสอบคู่ของค่าเฉลี่ยที่แตกต่างกันมากที่สุด (q_1) หากผลการทดสอบไม่มีนัยสำคัญแล้วการวิเคราะห์ก็จะหยุด เพราะว่าความแตกต่างของคู่อื่น ๆ ที่ค่าเฉลี่ยต่างกันน้อยกว่าก็ควรจะไม่มีนัยสำคัญด้วย ถ้าผลการทดสอบค่าเฉลี่ยคู่ที่แตกต่างกันมากที่สุดมีนัยสำคัญแล้ว การทดสอบค่าเฉลี่ยก็จะดำเนินการต่อไปในคู่ที่สอง (q_2) ซึ่งเป็นคู่ที่มีค่าเฉลี่ยแตกต่างกันมากรองลงมา ถ้าผลการทดสอบไม่มีนัยสำคัญ ก็จะหยุดการเปรียบเทียบ เพราะว่าความแตกต่างของคู่อื่น ๆ ที่มีค่าเฉลี่ยต่างกันน้อยกว่าก็ควรจะไม่มีนัยสำคัญด้วย สังเกตว่าการทดสอบเป็นรายคู่อาจไม่จำเป็นต้องทดสอบทุกคู่เสมอไป

ใช้ตัวอย่างข้างต้นเปรียบเทียนพหุคูณด้วย HSD ได้ผลการคำนวณดังนี้

ค่าวิกฤติ

$${}_{1-\alpha}q_{df(error),J} = {}_{.95}q_{28,4} = 3.87$$

ความคลาดเคลื่อนมาตรฐาน

$$\begin{aligned} s_{\psi'} &= \sqrt{[MS_{error}/n]} \\ &= \sqrt{36.11/8} \\ &= 2.12 \end{aligned}$$

สถิติทดสอบ

$$\begin{aligned} q_1 &= (\overline{Y}_4 - \overline{Y}_1)/s_{\psi'} \\ &= (24.38-11.13)/2.12 \end{aligned}$$

$= 6.24$ (มีนัยสำคัญ)

$q_2 = (\overline{Y}_4 - \overline{Y}_2)/s_{\psi'}$

$= (24.38-17.88)/2.12$

$= 3.06$ (ไม่มีนัยสำคัญ)

$q_3 = (\overline{Y}_3 - \overline{Y}_1)/s_{\psi'}$

$= (20.25-11.13)/2.12$

$= 4.29$ (มีนัยสำคัญ)

$q_4 = (\overline{Y}_3 - \overline{Y}_2)/s_{\psi'}$

$= (20.28-17.88)/2.12$

$= 1.12$ (ไม่มีนัยสำคัญ)

$q_5 = (\overline{Y}_2 - \overline{Y}_1)/s_{\psi'}$

$= (17.88-11.13)/2.12$

$= 3.18$ (ไม่มีนัยสำคัญ)

ผลที่ได้ชี้ให้เห็นว่าค่าเฉลี่ยของกลุ่มที่แตกต่างกันคือกลุ่ม 1 กับ 4 และกลุ่ม 1 กับ 3 และคู่สุดท้ายสำหรับการตรวจสอบคือกลุ่ม 3 กับ 4 ซึ่งผลที่ได้พอจะคาดเดาได้ว่าไม่น่าจะมีนัยสำคัญทางสถิติ

$q_6 = (\overline{Y}_4 - \overline{Y}_3)/s_{\psi'}$

$= (24.38-20.25)/2.12$

$= 1.94$ (ไม่มีนัยสำคัญ)

อาจสรุปผลการเปรียบเทียบรายคู่ลงตาราง ดังตัวอย่างในตาราง 1 แสดงผลการทดสอบเปรียบเทียบรายคู่ด้วยวิธี HSD ของตูกี้ โดยในตารางจะเรียงลำดับค่าเฉลี่ยจากน้อยไปมาก ทั้งแนวแถวและสดมภ์

ตาราง 1 การทดสอบทางสถิติและผลของการเปรียบเทียบรายคู่ด้วยวิธี HSD

	กลุ่ม 1	กลุ่ม 2	กลุ่ม 3	กลุ่ม 4
กลุ่ม 1 (ค่าเฉลี่ย 11.13)	-	3.18	4.29*	6.24*
กลุ่ม 2 (ค่าเฉลี่ย 17.88)		-	1.12	3.06
กลุ่ม 3 (ค่าเฉลี่ย 20.25)			-	1.97
กลุ่ม 4 (ค่าเฉลี่ย 24.38)				-

การทดสอบ HSD ของตูกี้จะมีลักษณะเป็นแบบธรรมเนียมนิยมมากกว่าวิธีของ Newman-Keuls ซึ่งจะมีระดับความคลาดเคลื่อนแบบที่ 1 แบบ family-wise ต่ำ แต่มีระดับความคลาดเคลื่อนแบบที่ 2 สูงกว่า และมีอำนาจการทดสอบต่ำกว่า วิธี HSD จะมีอำนาจการทดสอบมากกว่าวิธีของดันน์ (Dunn) และเชฟเฟ่ (Scheffe) เมื่อมีการทดสอบทุกคู่ที่เป็นไปได้ทั้งหมด แต่ในการเปรียบเทียบบางคู่นั้น วิธีของดันน์ มีอำนาจการทดสอบสูงกว่า HSD และเชฟเฟ่ การทดสอบ HSD จะสามารถควบคุมระดับความคลาดเคลื่อนแบบที่ 1 ได้ดีกว่าวิธีของดันน์ หรือ ดันซิเด็ก (Dunn or the Dunn-Sidak procedure) เทคนิคการเปรียบเทียบด้วย HSD เป็นการเปรียบเทียบรายคู่ที่ดีเมื่อแต่ละกลุ่มมี n เท่ากัน มีความแกร่งต่อการไม่เป็นโค้งปกติแต่ไม่แกร่งเมื่อมีข้อมูลผิดปกติ extreme cases และไม่แกร่งเท่าวิธีของเชฟเฟ่

การทดสอบตูกี้-การ์เมอร์ (The Tukey-Kramer Test)

มีหลายทางเลือกสำหรับวิธีเปรียบเทียบพหุคูณกรณีที่ n แต่ละกลุ่มไม่เท่ากัน รวมไปถึงวิธีตูกี้-การ์เมอร์ ที่ใช้ค่าเฉลี่ยฮาร์โมนิคกรณีที่ n ไม่เท่ากัน วิธีนี้คำนวณคลายกับวิธีตูกี้แต่ต่างกันตรงที่

$$s_{\psi'} = \sqrt{\{MS_{error}[\frac{1}{2}(\frac{1}{n_1}+\frac{1}{n_2})]\}}$$

ค่าวิกฤติเหมือนกับวิธีตูกี้

เมื่อความแปรปรวนของกลุ่มแต่ละกลุ่มไม่เท่ากันหรือทั้งความแปรปรวนและจำนวนตัวอย่างในแต่ละกลุ่มไม่เท่ากัน มีหลายวิธีที่สามารถเลือกใช้ได้รวมไปถึง วิธี The Games-Howell วิธี Dunnett's T3 and C และวิธี The Tamhane ซึ่งรายละเอียดของวิธีเหล่านี้สามารถศึกษาได้จากหนังสือของเคิร์ก (Kirk. 1982)

วิธีนิวแมน-คูลส์ (Newman-Keuls Procedure)

การเปรียบเทียบพหุคูณด้วยวิธีนิวแมน-คูลส์ จะคล้ายกับการทดสอบ HSD ของตูกี้ แต่มีข้อแตกต่างบางประการคือ วิธีนิวแมน-คูลส์จะอยู่บนพื้นฐานของระดับความคลาดเคลื่อนแบบที่ 1 แบบ the family-wise สูงกว่าการทดสอบ HSD ดังนั้นวิธี วิธีนิวแมน-คูลส์จะมีนัยสำคัญได้ง่ายกว่าวิธี HSD ของตูกี้ แต่มีระดับการเกิดความคลาดเคลื่อนแบบที่ 2 ต่ำกว่าและมีอำนาจการทดสอบสูงกว่า วิธีนิวแมน-คูลส์ และ HSD จะเป็นการเปรียบเทียบภายหลังโดยเปรียบเทียบกันเป็นรายคู่กรณีที่มี n เท่ากันทุกกลุ่ม

วิธีนิวแมน-คูลส์จะคล้ายกับวิธี HSD ยกเว้นค่าวิกฤติ ซึ่งวิธีนิวแมน-คูลส์จะมีค่าวิกฤติเท่ากับ ${}_{1-\alpha}q_{df(error),r}$ เมื่อ r เท่ากับความแตกต่างของลำดับค่าเฉลี่ยของกลุ่มที่นำมาเปรียบเทียบ ดังนั้นค่าวิกฤติจะขึ้นอยู่กับการเปรียบเทียบ สำหรับการศึกษาค่าเฉลี่ย 10 ค่า การเปรียบเทียบจะเกี่ยวข้องกับค่าเฉลี่ยที่มากที่สุดและน้อยที่สุด ซึ่งความแตกต่างของลำดับที่คือ 10 ดังนั้น r ก็จะมี

ค่า 10 การเปรียบเทียบค่าเฉลี่ยที่มากที่สุดและค่าเฉลี่ยที่น้อยรองลงมา ค่า r จะมีค่า 9 ซึ่งสามารถเปิดได้จากตารางในภาคผนวก

จากตัวอย่างข้างต้นคำนวณวิธีนิวแมน-คูลส์ ผลจะได้ดังนี้

ค่าวิกฤติ

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,4} = 3.87$$

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,3} = 3.50$$

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,2} = 2.90$$

ความคลาดเคลื่อนมาตรฐาน

$$\begin{aligned} s_{\psi'} &= \sqrt{[MS_{error}/n]} \\ &= \sqrt{36.11/8} \\ &= 2.12 \end{aligned}$$

สถิติทดสอบ

$$\begin{aligned} q_1 &= (\overline{Y}_4 - \overline{Y}_1)/s_{\psi'} \\ &= (24.38-11.13)/2.12 \\ &= 6.24 \text{ (มีนัยสำคัญ)} \end{aligned}$$

$$\begin{aligned} q_2 &= (\overline{Y}_4 - \overline{Y}_2)/s_{\psi'} \\ &= (24.38-17.88)/2.12 \\ &= 3.06 \text{ (ไม่มีนัยสำคัญ)} \end{aligned}$$

$$\begin{aligned} q_3 &= (\overline{Y}_3 - \overline{Y}_1)/s_{\psi'} \\ &= (20.25-11.13)/2.12 \\ &= 4.29 \text{ (มีนัยสำคัญ)} \end{aligned}$$

$$\begin{aligned} q_4 &= (\overline{Y}_3 - \overline{Y}_2)/s_{\psi'} \\ &= (20.28-17.88)/2.12 \\ &= 1.12 \text{ (ไม่มีนัยสำคัญ)} \end{aligned}$$

$$\begin{aligned} q_5 &= (\overline{Y}_2 - \overline{Y}_1)/s_{\psi'} \\ &= (17.88-11.13)/2.12 \\ &= 3.18 \text{ (มีนัยสำคัญ)} \end{aligned}$$

ผลที่ได้นี้จะมีค่าเหมือนกับวิธี HSD ของตูกี้ ยกเว้นการเปรียบเทียบคู่ที่ 5 ที่มีนัยสำคัญซึ่งแตกต่างจากวิธี HSD ของตูกี้

สำหรับกรณีที่ n แต่ละกลุ่มไม่เท่ากัน การเปรียบเทียบจะใช้วิธีตูกี้-การ์เมอร์

การทดสอบ Duncan's New Multiple Range Test

การทดสอบของดันแคนนี้ถูกพัฒนาให้เป็นการเปรียบเทียบพหุคูณภายหลังเป็นรายคู่สำหรับกรณี n แต่ละกลุ่มเท่ากัน การทดสอบของดันแคนจะคล้ายกับของนิวแมน-คูลส์ คือเป็นวิธีการแบบ family-wise การทดสอบดันแคนจะมีนัยสำคัญได้ง่ายกว่าวิธีของนิวแมน-คูลส์ แต่มีอำนาจการทดสอบสูง ระดับนัยสำคัญของการทดสอบดันแคนจะคำนวณด้วยสูตร $1-(1-\alpha)^{r-1}$ เมื่อ r คือพิสัยของลำดับค่าเฉลี่ย ที่ $\alpha = 0.05$ ระดับนัยสำคัญจะเปลี่ยนแปลงไปตามค่าของ r เมื่อ r เพิ่มขึ้นระดับนัยสำคัญก็จะเพิ่มขึ้นด้วย เช่น

พิสัยของลำดับค่าเฉลี่ยคือ 2 : $1-(1-0.05)^1 = 0.05$

พิสัยของลำดัยค่าเฉลี่ยคือ 3 : $1-(1-0.05)^2 = 0.10$

พิสัยของลำดัยค่าเฉลี่ยคือ 4 : $1-(1-0.05)^3 = 0.14$

พิสัยของลำดัยค่าเฉลี่ยคือ 5 : $1-(1-0.05)^4 = 0.19$

กระบวนการคำนวณนั้นจะทำนองเดียวกับวิธีของนิวแมน-คูลส์ การทดสอบทางสถิติสำหรับการเปรียบเทียบแต่ละคู่จะถูกเปรียบเทียบกับค่าวิกฤติ ${}_{1-\alpha}q_{df(error),r}$ ที่เปิดจากตารางในภาคผนวก

จากตัวอย่างข้างต้นสามารถคำนวณได้ค่าต่าง ๆ ดังนี้

ค่าวิกฤติ

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,4} = 3.13$$

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,3} = 3.04$$

$${}_{1-\alpha}q_{df(error),r} = {}_{0.95}q_{28,2} = 2.90$$

ความคลาดเคลื่อนมาตรฐาน

$$s_{\psi'} = \sqrt{[MS_{error}/n]} = \sqrt{36.11/8} = 2.12$$

สถิติทดสอบ

$$q_1 = (\overline{Y}_4 - \overline{Y}_1)/s_{\psi'} = (24.38-11.13)/2.12 = 6.24$$ (มีนัยสำคัญ)

$$q_2 = (\overline{Y}_4 - \overline{Y}_2)/s_{\psi'} = (24.38-17.88)/2.12 = 3.06$$ (มีนัยสำคัญ)

$$q_{2'} = (\overline{Y}_4 - \overline{Y}_3)/s_{\psi'} = (24.38-20.25)/2.12 = 1.94$$ (ไม่มีนัยสำคัญ)

$q_3 = (\overline{Y}_3 - \overline{Y}_1)/s_{\psi'}$

$= (20.25\text{-}11.13)/2.12$

$=$ 4.29 (มีนัยสำคัญ)

$q_4 = (\overline{Y}_3 - \overline{Y}_2)/s_{\psi'}$

$= (20.28\text{-}17.88)/2.12$

$=$ 1.12 (ไม่มีนัยสำคัญ)

$q_5 = (\overline{Y}_2 - \overline{Y}_1)/s_{\psi'}$

$= (17.88\text{-}11.13)/2.12$

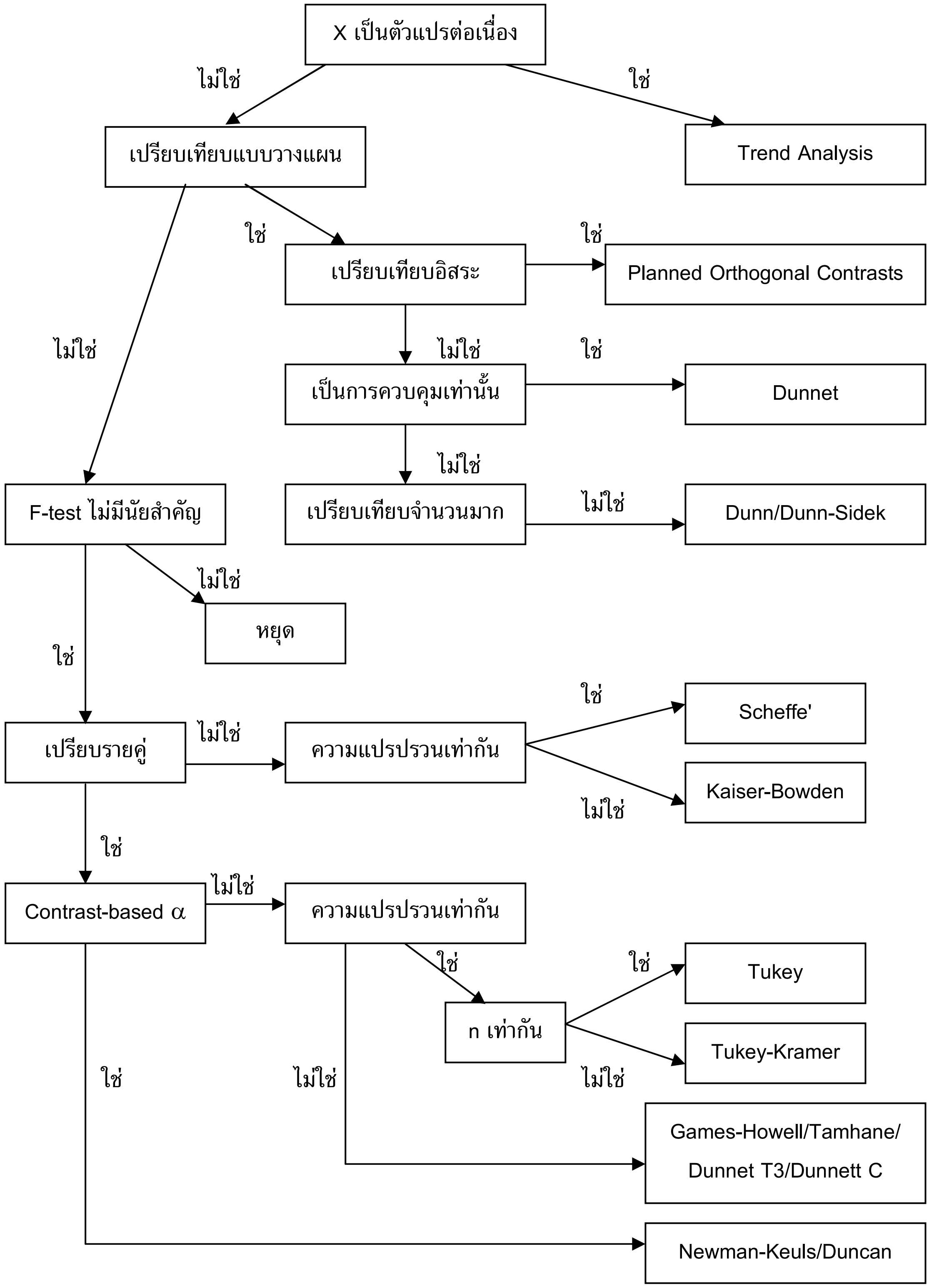
$=$ 3.18 (มีนัยสำคัญ)

จะเห็นว่าได้ผลทำนองเดียวกับวิธีของนิวแมน-คูลส์ แต่แตกต่างตรงการเปรียบเทียบคู่ที่ 2 ที่มีนัยสำคัญในวิธีของดันแคน ซึ่งผลนี้จะนำไปสู่การเปรียบเทียบคู่ที่ $q_{2'}$ ซึ่งไม่มีนัยสำคัญ การเปรียบเทียบกับค่าวิกฤติสำหรับค่าที่ r เหมือนกัน เราจะเห็นว่าการทดสอบดันแคนจะให้อำนาจการทดสอบสูงกว่าวิธีของนิวแมน-คูลส์ ยกเว้นการเปรียบเทียบพื้นฐานที่ให้อำนาจการทดสอบเท่ากัน

สรุป

ทั้งหมดที่กล่าวมาจะเกี่ยวข้องกับการเปรียบเทียบพหุคูณ ความแตกต่างระหว่างกลุ่มกรณีที่มีตัวแปรอิสระเพียงตัวแปรเดียว การเปรียบเทียบพหุคูณนี้จะกล่าวเกี่ยวกับ 1) การนิยามการเปรียบเทียบ 2) การเปรียบเทียบแบบวางแผนและการเปรียบเทียบแบบภายหลัง 3) อัตราการเกิดความคลาดเคลื่อนแบบที่ 1 แบบการเปรียบเทียบเป็นฐาน (contrast-bases) กับแบบ family-wise และ 4) การเปรียบเทียบอิสระ (Orthogonmal contrasts) ถัดมาเป็นการอธิบายวิธีการเปรียบเทียบพหุคูณในแต่ละวิธีรวมทั้งตัวอย่างการวิเคราะห์และการละเมิดข้อตกลงเบื้องต้นและเสนอแนะถึงการนำไปประยุกต์ใช้อย่างเหมาะสม

ภาพประกอบ 2 จะเป็นแผนภาพแสดงถึงการตัดสินใจใช้กระบวนการเปรียบเทียบพหุคูณในแต่ละวิธีการ ซึ่งไม่ใช่นักสถิติทุกคนที่จะเห็นด้วยกับแผนภาพนี้ อันเนื่องมาจากมีความขัดแย้งทางความคิดเห็นบางประการเกี่ยวกับกระบวนการเปรียบเทียบพหุคูณที่เหมาะสมในบางสถานการณ์ แต่แผนภาพนี้จะช่วยเป็นแนวทางในการตัดสินใจเลือกใช้วิธีการเปรียบเทียบพหุคูณกับงานวิจัย



ภาพประกอบ 2 แผนภาพประกอบการเปรียบเทียบพหุคูณ

แปลและเรียบเรียงจาก

Lomax, Richard G. (1992). **Statistical Concepts : A Second Course for Education and the Behavioral Sciences**. London : Lawrence Erlbaum Associates, Inc.