การเปรียบเทียบประสิทธิภาพของสถิติทดสอบการเปรียบเทียบพหุคูณ ในกรณีความแปรปรวนเท่ากันสำหรับ 3 ประชากร โดยใช้โปรแกรมอาร์

An Efficiency Comparison of Multiple Comparison Test in Case of Homogeneity of Variance for Three Populations using R Program

อัชฌา อระวีพร ปารณัท สุขเจริญ กฤตพล ธีรธนิตนันท์ กำชัย สุภัทรกุล และ โศภน พงษ์ชาลี
Autcha Araveeporn Paranut Sukcharoen Krittapon Theratanitnan Kumchai Suphatarakul and Sopon Pongchalee
ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

บทคัดย่อ

การวิจัยนี้เป็นการวิจัยเชิงจำลองมีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของสถิติทดสอบของฟิชเชอร์โดยใช้ความแตกต่างที่มีนัยสำคัญน้อยที่สุด สถิติทดสอบของบอนเฟอร์โรนี สถิติทดสอบของทูกีย์ สถิติทดสอบแบบพิสัยพหุคูณของดันแคน สถิติทดสอบของสติวเดนต์–นิวแมน–คูล และสถิติทดสอบของเชฟเฟ สำหรับทดสอบการเปรียบเทียบพหุคูณในกรณีความแปรปรวนเท่ากัน สำหรับ 3 ประชากร โดยศึกษาจากข้อมูลที่สุ่มจากประชากรที่มีการแจกแจงปรกติ กำหนดขนาดตัวอย่างเท่ากับ (5,5,5) (10,10,10) (30,30,30) (5,6,7) (10,12,14) และ (30,34,38) ในการคำนวณค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 กำหนดค่าเฉลี่ยของประชากรเท่ากับ (4,4,4) และการคำนวณค่าประมาณกำลังการทดสอบ กำหนดค่าเฉลี่ยของประชากรเท่ากับ (4,8,12) โดยที่ความแปรปรวนของแต่ละประชากรกำหนดเป็น 2 4 8 และ 16 ตามลำดับ กำหนดระดับนัยสำคัญ 3 ระดับ คือ 0.01 0.05 และ 0.1 ใช้โปรแกรมอาร์ในการจำลองและวิเคราะห์ข้อมูล ทำการจำลองข้อมูลซ้ำ 5,000 รอบในแต่ละสถานการณ์ ผลการวิจัยพบว่าสถิติทดสอบของบอนเฟอร์โรนี สถิติทดสอบของทูกีย์ สถิติทดสอบของสติวเดนต์–นิวแมน–คูล และสถิติทดสอบของเชฟเฟ สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ทุกสถานการณ์ที่ศึกษา เมื่อพิจารณาค่าประมาณกำลังการทดสอบ พบว่าสถิติทดสอบของสติวเดนต์–นิวแมน–คูล มีค่าประมาณกำลังการทดสอบสูงที่สุดในทุกสถานการณ์ที่ศึกษา และพบว่าค่าประมาณกำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นหรือความแปรปรวนลดลง

คำสำคัญ : การเปรียบเทียบพหุคูณ ค่าประมาณกำลังการทดสอบ ค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 สถิติทดสอบของเชฟเฟ สถิติทดสอบของทูกีย์ สถิติทดสอบของบอนเฟอร์โรนี สถิติทดสอบของฟิชเชอร์โดยใช้ความแตกต่างที่มีนัยสำคัญน้อยที่สุด สถิติทดสอบของสติวเดนต์–นิวแมน–คูล สถิติทดสอบแบบพิสัยพหุคูณของดันแคน

*ที่อยู่ติดต่อ. E-mail address : kaautcha@hotmail.com

Abstract

This research is a simulating research that aimed to study and to compare the efficiency of Fisher's least significant difference test, Bonferroni's test, Tukey's test, Duncan's new multiple range test, Student–Newman–Keul's test, and Scheffe's test for multiple comparison testing in case of homogeneity of variance for three populations. In this case, we randomize data from three populations that have a normal distribution. The sample sizes are set equal to (5,5,5), (10,10,10), (30,30,30), (5,6,7), (10,12,14), and (30,34,38). The population mean are set equal to (4,4,4) for calculating the estimated probability of type I error, and set equal to (4,8,12) for calculating the estimated power of a test. The population variances of each population are set equal to 2, 4, 8, and 16. The significant levels are considered on three levels at 0.01, 0.05, and 0.1. R program is used for simulation and data analysis with 5,000 times for each situation. The results revealed that Bonferroni's test, Tukey's test, Student–Newman–Keul's test and Scheffe's test can control probability of type I error in all situations. Considering the power of a test, Student–Newman–Keul's test shows the highest power of a test in all situations. Power of a test increases as sample size increased or variance decreased.

Keywords : Multiple Comparison, Power of a Test, Probability of Type I Error, Scheffe's Test, Tukey's Test, Bonferroni's Test, Fisher's Least Significant Difference Test, Student–Newman–Keul's Test, Duncan's New Multiple Range Test

1. บทนำ

ในการทดสอบสมมุติฐานเพื่อทดสอบความเท่ากันของค่าเฉลี่ยในกรณีตั้งแต่ 3 ประชากรขึ้นไป นิยมใช้การวิเคราะห์ความแปรปรวน (Analysis of Variance : ANOVA) โดยใช้สถิติทดสอบเอฟ ซึ่งมีข้อกำหนดเบื้องต้น (Assumption) คือ ข้อมูลแต่ละประชากรสุ่มจากประชากรที่มีการแจกแจงปรกติ ข้อมูลแต่ละประชากรสุ่มจากประชากรที่มีความแปรปรวนเท่ากันหรือคงที่ และข้อมูลแต่ละประชากรสุ่มจากประชากรที่เป็นอิสระกัน ภายหลังการทดสอบแล้วถ้าค่าสถิติทดสอบเอฟ ไม่มีนัยสำคัญทางสถิติ แสดงว่าค่าเฉลี่ยของแต่ละประชากรไม่มีความแตกต่างกัน แต่ถ้าค่าสถิติทดสอบเอฟ มีนัยสำคัญทางสถิติ ก็แสดงว่ามีค่าเฉลี่ยของประชากรอย่างน้อย 1 คู่ ที่มีความแตกต่างกัน การสรุปผลที่ได้จากการวิเคราะห์ความแปรปรวนยังไม่สามารถบอกได้ว่ามีประชากรคู่ไหนบ้างที่มีค่าเฉลี่ยแตกต่างกัน ผู้วิจัยต้องทำการเปรียบเทียบพหุคูณ (Multiple Comparison) เพื่อตรวจสอบว่าค่าเฉลี่ยของประชากรคู่ใดบ้างที่แตกต่างกันหรือไม่แตกต่างกัน [1]

โดยทั่วไปแล้วการเปรียบเทียบพหุคูณจะกระทำภายหลังเมื่อทราบว่าค่าสถิติทดสอบเอฟมีนัยสำคัญทางสถิติ และสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ มีอยู่หลายสถิติทดสอบ เช่น สถิติ

ทดสอบของฟิชเชอร์โดยใช้ความแตกต่างที่มีนัยสำคัญน้อยที่สุด (Fisher's Least Significant Difference Test : LSD) สถิติทดสอบของบอนเฟอร์โรนี (Bonferroni's Test) สถิติทดสอบของทูกีย์ (Tukey's Test) สถิติทดสอบแบบพิสัยพหุคูณของดันแคน (Duncan's New Multiple Range Test) สถิติทดสอบของสติวเดนต์–นิวแมน–คูล (Student–Newman–Keul's Test : SNK Test) และสถิติทดสอบของเชฟเฟ (Scheffe's Test) เป็นต้น และโดยทั่วไปผู้วิเคราะห์จะใช้สถิติทดสอบใดสถิติทดสอบหนึ่งเท่านั้น ซึ่งเป็นที่น่าสนใจว่าถ้าใช้สถิติทดสอบอื่น แล้วพบว่าผลการทดสอบไม่สอดคล้องกัน เราควรเลือกใช้ผลการทดสอบจากสถิติทดสอบใด

จากการศึกษางานวิจัยที่เกี่ยวข้องพบว่า Boardman และ Moffitt [2] ได้ทำการศึกษาเปรียบเทียบวิธีวิเคราะห์พหุคูณ คือ วิธี เช่น สถิติทดสอบของฟิชเชอร์โดยใช้ความแตกต่างที่มีนัยสำคัญน้อยที่สุด สถิติทดสอบของทูกีย์ สถิติทดสอบของเชฟเฟ สถิติทดสอบแบบพิสัยพหุคูณของดันแคน และ สถิติทดสอบของ สติวเดนต์–นิวแมน–คูล จากการแจกแจงปรกติ นอกจากนี้ Bernhardson [3] ได้ศึกษาเพิ่มเติมโดยใช้ตัวสถิติเดิมแต่ทำภายหลังจากการวิเคราะห์ความแปรปรวนด้วยสถิติทดสอบเอฟ และ สุญานี [4] ได้ศึกษาตัวสถิติเดิม 5 ตัว และเพิ่มรูปแบบการแจกแจงเป็นแบบยูนิฟอร์ม และการแจกแจงแบบสูงโด่ง

ดังนั้นในการทำงานวิจัยครั้งนี้ จึงสนใจทำการศึกษาและเปรียบเทียบประสิทธิภาพของสถิติทดสอบของฟิชเชอร์ สถิติทดสอบของบอนเฟอร์โรนี สถิติทดสอบของทูกีย์ สถิติทดสอบของดันแคน สถิติทดสอบของสติวเดนต์–นิวแมน–คูล และสถิติทดสอบของเชฟเฟ สำหรับทดสอบการเปรียบเทียบพหุคูณ ในกรณีความแปรปรวนเท่ากันสำหรับ 3 ประชากร โดยใช้โปรแกรมอาร์ ในการจำลองและวิเคราะห์ข้อมูล

2. วิธีการวิจัย

การวิจัยนี้เป็นการวิจัยเชิงจำลองเพื่อศึกษาและเปรียบเทียบประสิทธิภาพของสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ ในกรณีความแปรปรวนเท่ากันสำหรับ 3 ประชากร ทั้ง 6 สถิติทดสอบ โดยมีวิธีการดำเนินการวิจัย ดังต่อไปนี้

2.1 ขอบเขตของการวิจัย

2.1.1 กำหนดจำนวนประชากรเท่ากับ 3 ประชากร

2.1.2 กำหนดระดับนัยสำคัญของการทดสอบสมมุติฐานคือ 0.01 0.05 และ 0.1 ตามลำดับ

2.1.3 กำหนดขนาดตัวอย่างสุ่มจากแต่ละประชากรขนาดเท่ากันและไม่เท่ากัน ดังตารางที่ 1

ตารางที่ 1. ขนาดตัวอย่างที่ใช้ในการวิจัย

ขนาดตัวอย่าง (n_1, n_2, n_3)	
เท่ากัน	ไม่เท่ากัน
(5,5,5) , (10,10,10) , (30,30,30)	(5,6,7) , (10,12,14) , (30,34,38)

2.1.4 ในทุกขนาดตัวอย่างจากจากตารางที่ 1 จะศึกษาจากข้อมูลที่สุ่มจากประชากรที่มีการแจกแจงปรกติ ด้วยพารามิเตอร์ (μ, σ^2)

2.1.5 ในการคำนวณค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 จะจำลองข้อมูลและสุ่มตัวอย่าง โดยกำหนดให้ข้อมูลสุ่มจากประชากรที่มีค่าเฉลี่ยเท่ากันทั้ง 3 ประชากร กำหนดค่าเฉลี่ยของแต่ละประชากรเท่ากับ 4 และความแปรปรวนเท่ากับ 2 4 8 และ 16 ตามลำดับ โดยกำหนดค่าพารามิเตอร์ที่สอดคล้องกับค่าเฉลี่ยและความแปรปรวนของแต่ละประชากร ดังตารางที่ 2

ตารางที่ 2. พารามิเตอร์สำหรับการคำนวณค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1

สถานการณ์	พารามิเตอร์ (μ, σ^2)			ค่าเฉลี่ย (μ_1, μ_2, μ_3)	ความแปรปรวน $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$
	ประชากรที่ 1	ประชากรที่ 2	ประชากรที่ 3		
1	(4,2)	(4,2)	(4,2)	(4,4,4)	(2,2,2)
2	(4,4)	(4,4)	(4,4)	(4,4,4)	(4,4,4)
3	(4,8)	(4,8)	(4,8)	(4,4,4)	(8,8,8)
4	(4,16)	(4,16)	(4,16)	(4,4,4)	(16,16,16)

2.1.6 ทำการทดสอบการเปรียบเทียบพหุคูณทั้ง 6 สถิติทดสอบ โดยนำค่าสถิติทดสอบที่คำนวณได้ไปคำนวณค่าพี (p–value) และเปรียบเทียบค่าพีกับระดับนัยสำคัญที่กำหนดเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมุติฐานว่าง บันทึกจำนวนครั้งที่ปฏิเสธสมมุติฐานว่าง และทำซ้ำจนครบ 5,000 ครั้ง

2.1.7 คำนวณค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยนำจำนวนครั้งที่ปฏิเสธสมมุติฐานว่างหารด้วย 5,000

2.1.8 ตรวจสอบความสามารถในการควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของ สถิติทดสอบตามเกณฑ์ของ Bradley [5] ในแต่ละสถานการณ์ โดยถ้าค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 อยู่ในช่วงที่กำหนดดังตารางที่ 3 จะสรุปว่าสถิติทดสอบนั้นสามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

ตารางที่ 3. เกณฑ์ของ Bradley จำแนกตามระดับนัยสำคัญ

ระดับนัยสำคัญ	ช่วงที่กำหนด
0.01	[0.005,0.015]
0.05	[0.025,0.075]
0.10	[0.050,0.150]

2.1.9 คำนวณค่าประมาณกำลังการทดสอบของสถิติทดสอบ เฉพาะสถิติทดสอบที่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley ในแต่ละสถานการณ์ โดยจำลองข้อมูลและสุ่มตัวอย่าง โดยกำหนดให้ข้อมูลสุ่มจากประชากรที่มีค่าเฉลี่ยแตกต่างกันทั้ง 3 ประชากร กำหนดระยะห่างของค่าเฉลี่ยของแต่ละประชากร เป็น 4 และความแปรปรวนเท่ากับ 2 4 8 และ 16 ตามลำดับ โดยกำหนดค่าพารามิเตอร์ที่สอดคล้องกับค่าเฉลี่ยและความแปรปรวนของแต่ละประชากร ดังตารางที่ 4

ตารางที่ 4. พารามิเตอร์สำหรับการคำนวณค่าประมาณกำลังการทดสอบ

สถานการณ์	พารามิเตอร์ (μ,σ^2)			ค่าเฉลี่ย (μ_1,μ_2,μ_3)	ความแปรปรวน $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	ประชากรที่ 1	ประชากรที่ 2	ประชากรที่ 3		
1	(4,2)	(8,2)	(12,2)	(4,8,12)	(2,2,2)
2	(4,4)	(8,4)	(12,4)	(4,8,12)	(4,4,4)
3	(4,8)	(8,8)	(12,8)	(4,8,12)	(8,8,8)
4	(4,16)	(8,16)	(12,16)	(4,8,12)	(16,16,16)

2.1.10 ทำการทดสอบการเปรียบเทียบพหุคูณทั้ง 6 สถิติทดสอบ โดยนำค่าสถิติทดสอบที่คำนวณได้ไปคำนวณค่าพี (p–value) และเปรียบเทียบค่าพีกับระดับนัยสำคัญที่กำหนดเพื่อสรุปว่าจะปฏิเสธหรือยอมรับสมมุติฐานว่าง บันทึกจำนวนครั้งที่ปฏิเสธสมมุติฐานว่าง เนื่องจากข้อมูลที่นำมาทดสอบทั้ง 3 ประชากร และทำซ้ำจนครบ 5,000 ครั้ง

2.1.11 คำนวณค่าประมาณกำลังการทดสอบโดยนำจำนวนครั้งที่ปฏิเสธสมมุติฐานว่างหารด้วย 5,000

2.1.12 เปรียบเทียบค่าประมาณกำลังการทดสอบของสถิติทดสอบ โดยสถิติทดสอบที่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley และมีค่าประมาณกำลังการทดสอบสูงที่สุดจะเป็นสถิติทดสอบที่มีประสิทธิภาพสูงที่สุด

2.2 สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ

ถ้ามีประชากรทั้งหมด 3 ประชากร สำหรับการเปรียบเทียบพหุคูณจะทำการเปรียบเทียบได้ทั้งหมด ${}^3C_2 = 3$ คู่การเปรียบเทียบ โดยจะดำเนินการเปรียบเทียบไปจนครบทั้ง 3 การเปรียบเทียบ และในแต่ละครั้งของการเปรียบเทียบจะมีสมมุติฐานทางสถิติดังนี้

$$H_0: \quad \mu_i = \mu_j$$

$$H_1: \quad \mu_i \neq \mu_j \quad ; \quad i, j = 1,2,3 \text{ และ } i \neq j$$

โดยข้อกำหนดเบื้องต้น (Assumption) ของสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณจะเหมือนกับข้อกำหนดเบื้องต้นของการวิเคราะห์ความแปรปรวน คือ

1. ข้อมูลแต่ละประชากรสุ่มจากประชากรที่มีการแจกแจงปรกติ

2. ข้อมูลแต่ละประชากรสุ่มจากประชากรที่มีความแปรปรวนเท่ากันหรือคงที่

3. ข้อมูลแต่ละประชากรสุ่มจากประชากรที่เป็นอิสระกัน

และอาจเลือกใช้สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ ได้ดังต่อไปนี้

2.2.1 สถิติทดสอบของฟิชเชอร์โดยใช้ความแตกต่างที่มีนัยสำคัญน้อยที่สุด (Fisher's Least Significant Difference Test : LSD) [1]

ตัวสถิติทดสอบ t = $\dfrac{\bar{y}_{i\bullet} - \bar{y}_{j\bullet}}{\sqrt{MSE\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|t| \geq t_{\frac{\alpha}{2};\nu}$

เมื่อ $t_{\frac{\alpha}{2};\nu}$ คือ ค่าวิกฤตที่เปิดจากตารางการแจกแจงทีที่ระดับนัยสำคัญ $\alpha/2$ และ ν เป็นจำนวนองศาเสรีของความคลาดเคลื่อน

MSE คือ ค่ากำลังสองเฉลี่ยของความคลาดเคลื่อนในตารางวิเคราะห์ความแปรปรวน

n_i, n_j คือ ขนาดตัวอย่างของประชากรที่ i และ j ที่ต้องการเปรียบเทียบ ตามลำดับ

$\bar{y}_{i\bullet}, \bar{y}_{j\bullet}$ คือ ค่าเฉลี่ยตัวอย่างของประชากรที่ i และ j ที่ต้องการเปรียบเทียบ ตามลำดับ

2.2.2 สถิติทดสอบของบอนเฟอร์โรนี (Bonferroni's Test : Bonferroni) [1]

ตัวสถิติทดสอบ tB = $\dfrac{\bar{y}_{i\bullet} - \bar{y}_{j\bullet}}{\sqrt{MSE\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|tB| \geq t_{\frac{\alpha}{2S};\nu}$

เมื่อ $t_{\frac{\alpha}{2S};\nu}$ คือ ค่าที่เปิดจากตารางค่าวิกฤตของบอนเฟอร์โรนี และ S เป็นจำนวนคู่การเปรียบเทียบ โดยใช้จำนวนองศาเสรีของความคลาดเคลื่อน ν

2.2.3 สถิติทดสอบของทูกีย์ (Tukey's Test : Tukey) (1956) [6]

ตัวสถิติทดสอบ qTK = $\dfrac{\bar{y}_{i\bullet} - \bar{y}_{j\bullet}}{\sqrt{\dfrac{MSE}{2}\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|qTK| \geq q_{\alpha}(p,\nu)$

เมื่อ $q_{\alpha}(p,\nu)$ คือ ค่าวิกฤตของพิสัยสติวเดนต์ไดซ์ ที่ระดับนัยสำคัญ α ที่มีจำนวนประชากร p ประชากร และองศาเสรีของความคลาดเคลื่อน เท่ากับ ν

2.2.4 สถิติทดสอบแบบพิสัยพหุคูณของดันแคน

(Duncan's New Multiple Range Test : Duncan) (1955) [7]

ตัวสถิติทดสอบ $qD = \dfrac{\bar{y}_{i\bullet} - \bar{y}_{j\bullet}}{\sqrt{\dfrac{MSE}{2}\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|qD| \geq r_{\alpha}(p,\nu)$

เมื่อ $r_{\alpha}(p,\nu)$ คือ ค่าวิกฤตของพิสัยพหุคูณของดันแคน ที่ระดับนัยสำคัญ α โดยที่ p คือ จำนวนค่าเฉลี่ยในช่วงการเปรียบเทียบ ซึ่งเท่ากับค่าสัมบูรณ์ของผลต่างของอันดับบวกด้วย 1 และมีจำนวนองศาเสรีของความคลาดเคลื่อนเท่ากับ ν

2.2.5 สถิติทดสอบของสติวเดนต์–นิวแมน–คูล

(Student–Newman–Keul's Test or SNK Test : SNK) (1952) [8]

ตัวสถิติทดสอบ $qSNK = \dfrac{\bar{y}_{i\bullet} - \bar{y}_{j\bullet}}{\sqrt{\dfrac{MSE}{2}\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|qSNK| \geq q_{\alpha}(p,\nu)$

เมื่อ $q_{\alpha}(p,\nu)$ คือ ค่าวิกฤตของพิสัยสติวเดนต์ไดซ์ ที่ระดับนัยสำคัญ α โดยที่ p คือจำนวนค่าเฉลี่ยในช่วงการเปรียบเทียบ ซึ่งเท่ากับค่าสัมบูรณ์ของผลต่างของอันดับบวกด้วย 1 และมีจำนวนองศาเสรีของความคลาดเคลื่อนเท่ากับ ν

2.2.6 สถิติทดสอบของเชฟเฟ (Scheffe's Test : Scheffe) [1]

ตัวสถิติทดสอบ $FS = \dfrac{\left(\bar{y}_{i\bullet} - \bar{y}_{j\bullet}\right)^2}{MSE\left(\dfrac{1}{n_i}+\dfrac{1}{n_j}\right)}$

เขตวิกฤต จะปฏิเสธสมมุติฐานว่างเมื่อ $|FS| \geq \nu_1 F_{\alpha;\nu_1,\nu_2}$

เมื่อ ν_1 คือ จำนวนองศาเสรีของระหว่างทรีตเมนต์ (Between Treatment)

ν_2 คือ จำนวนองศาเสรีของความคลาดเคลื่อน

$F_{\alpha;\nu_1,\nu_2}$ คือ ค่าที่เปิดจากตารางการแจกแจงเอฟ ที่ระดับนัยสำคัญ α โดยใช้จำนวนองศาเสรีของระหว่างทรีตเมนต์และความคลาดเคลื่อน ตามลำดับ

3. ผลการวิจัย

3.1 ความสามารถในการควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1

- กำหนดระดับนัยสำคัญที่ระดับ 0.01

จากการพิจารณาความสามารถในการควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley เมื่อกำหนดระดับนัยสำคัญที่ระดับ 0.01 สรุปผลได้ดังตารางที่ 5

ตารางที่ 5. ค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.01

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	0.0254	0.0244	0.0264	0.0232	0.0230	0.0260
	Bonferroni	0.0090*	0.0086*	0.0084*	0.0072*	0.0074*	0.0078*
	Tukey	0.0108*	0.0096*	0.0088*	0.0090*	0.0086*	0.0086*
	Duncan	0.0210	0.0184	0.0182	0.0188	0.0176	0.0202
	SNK	0.0112*	0.0106*	0.0100*	0.0092*	0.0096*	0.0096*
	Scheffe	0.0072*	0.0064*	0.0068*	0.0066*	0.0056*	0.0056*
4	LSD	0.0218	0.0294	0.0278	0.0234	0.0290	0.0294
	Bonferroni	0.0062*	0.0086*	0.0094*	0.0076*	0.0088*	0.0092*
	Tukey	0.0080*	0.0096*	0.0114*	0.0084*	0.0094*	0.0104*
	Duncan	0.0160	0.0222	0.0208	0.0196	0.0226	0.0216
	SNK	0.0092*	0.0106*	0.0128*	0.0102*	0.0108*	0.0114*
	Scheffe	0.0052*	0.0076*	0.0078*	0.0064*	0.0070*	0.0074*
8	LSD	0.0278	0.0266	0.0228	0.0208	0.0244	0.0288
	Bonferroni	0.0106*	0.0070*	0.0086*	0.0086*	0.0104*	0.0090*
	Tukey	0.0120*	0.0086*	0.0090*	0.0094*	0.0108*	0.0102*
	Duncan	0.0238	0.0194	0.0184	0.0162	0.0196	0.0208
	SNK	0.0130*	0.0100*	0.0090*	0.0098*	0.0118*	0.0108*
	Scheffe	0.0086*	0.0052*	0.0074*	0.0074*	0.0080*	0.0064*
16	LSD	0.0232	0.0238	0.0236	0.0258	0.0254	0.0268
	Bonferroni	0.0094*	0.0094*	0.0074*	0.0094*	0.0090*	0.0082*
	Tukey	0.0110*	0.0100*	0.0078*	0.0100*	0.0096*	0.0086*
	Duncan	0.0192	0.0182	0.0182	0.0208	0.0200	0.0198
	SNK	0.0118*	0.0116*	0.0082*	0.0106*	0.0104*	0.0104*
	Scheffe	0.0088*	0.0086*	0.0054*	0.0082*	0.0072*	0.0066*

หมายเหตุ * สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley

จากตารางที่ 5 พบว่าตัวสถิติทดสอบของ Bonferroni Tukey SNK และ Scheffe สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ยกเว้นตัวสถิติ LSD และ Duncan

- กำหนดระดับนัยสำคัญที่ระดับ 0.05

จากการพิจารณาความสามารถในการควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley เมื่อกำหนดระดับนัยสำคัญที่ระดับ 0.05 สรุปผลได้ดังตารางที่ 6

ตารางที่ 6. ค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.05

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	0.1174	0.1090	0.1186	0.1234	0.1262	0.1182
	Bonferroni	0.0412*	0.0368*	0.0450*	0.0414*	0.0454*	0.0458*
	Tukey	0.0518*	0.0440*	0.0524*	0.0506*	0.0538*	0.0494*
	Duncan	0.0958	0.0868	0.1002	0.1050	0.1056	0.0944
	SNK	0.0590*	0.0504*	0.0608*	0.0566*	0.0590*	0.0572*
	Scheffe	0.0408*	0.0336*	0.0404*	0.0398*	0.0414*	0.0400*
4	LSD	0.1126	0.1228	0.1248	0.1156	0.1246	0.1224
	Bonferroni	0.0374*	0.0444*	0.0422*	0.0376*	0.0448*	0.0426*
	Tukey	0.0470*	0.0522*	0.0486*	0.0460*	0.0508*	0.0490*
	Duncan	0.0918	0.1028	0.0984	0.0946	0.1014	0.0990
	SNK	0.0574*	0.0612*	0.0558*	0.0544*	0.0612*	0.0586*
	Scheffe	0.0372*	0.0394*	0.0374*	0.0364*	0.0406*	0.0364*
8	LSD	0.1200	0.1194	0.1248	0.1162	0.1236	0.1222
	Bonferroni	0.0470*	0.0444*	0.0448*	0.0374*	0.0450*	0.0410*
	Tukey	0.0546*	0.0544*	0.0514*	0.0460*	0.0538*	0.0476*
	Duncan	0.1058	0.1034	0.1006	0.0952	0.1040	0.0976
	SNK	0.0636*	0.0622*	0.0594*	0.0548*	0.0632*	0.0562*
	Scheffe	0.0466*	0.0402*	0.0412*	0.0358*	0.0406*	0.0372*
16	LSD	0.1162	0.1132	0.1194	0.1178	0.1150	0.1276
	Bonferroni	0.0396*	0.0438*	0.0444*	0.0418*	0.0448*	0.0430*
	Tukey	0.0478*	0.0494*	0.0532*	0.0528*	0.0520*	0.0502*
	Duncan	0.0954	0.0942	0.0978	0.1020	0.0952	0.1070
	SNK	0.0554*	0.0554*	0.0590*	0.0596*	0.0584*	0.0580*
	Scheffe	0.0392*	0.0402*	0.0392*	0.0404*	0.0392*	0.0390*

หมายเหตุ * สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley

จากตารางที่ 6 พบว่าตัวสถิติทดสอบของ Bonferroni Tukey SNK และ Scheffe สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ยกเว้นตัวสถิติ LSD และ Duncan

- กำหนดระดับนัยสำคัญที่ระดับ 0.1

จากการพิจารณาความสามารถในการควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ตามเกณฑ์ของ Bradley เมื่อกำหนดระดับนัยสำคัญที่ระดับ 0.1 สรุปผลได้ดังตารางที่ 7

ตารางที่ 7. ค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.1

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	0.2186	0.2128	0.2250	0.2344	0.2294	0.2308
	Bonferroni	0.0776*	0.0814*	0.0892*	0.0826*	0.0828*	0.0878*
	Tukey	0.0970*	0.0974*	0.1058*	0.1072*	0.1006*	0.1040*
	Duncan	0.1930	0.1866	0.1960	0.2014	0.1954	0.1932
	SNK	0.1168*	0.1134*	0.1224*	0.1276*	0.1224*	0.1204*
	Scheffe	0.0828*	0.0816*	0.0884*	0.0862*	0.0828*	0.0858*
4	LSD	0.2172	0.2194	0.2200	0.2148	0.2172	0.2206
	Bonferroni	0.0748*	0.0732*	0.0786*	0.0746*	0.0780*	0.0828*
	Tukey	0.0932*	0.0930*	0.0948*	0.0960*	0.0952*	0.0986*
	Duncan	0.1910	0.1862	0.1898	0.1876	0.1816	0.1896
	SNK	0.1124*	0.1128*	0.1110*	0.1184*	0.1154*	0.1168*
	Scheffe	0.0790*	0.0734*	0.0768*	0.0778*	0.0778*	0.0808*
8	LSD	0.2108	0.2192	0.2192	0.2132	0.2282	0.2254
	Bonferroni	0.0854*	0.0798*	0.0792*	0.0754*	0.0832*	0.0800*
	Tukey	0.1068*	0.0998*	0.0932*	0.0982*	0.1002*	0.0954*
	Duncan	0.1872	0.1868	0.1852	0.1852	0.1956	0.1898
	SNK	0.1232*	0.1174*	0.1140*	0.1158*	0.1198*	0.1128*
	Scheffe	0.0906*	0.0802*	0.0760*	0.0800*	0.0830*	0.0770*
16	LSD	0.2106	0.2198	0.2408	0.2220	0.2216	0.2174
	Bonferroni	0.0798*	0.0798*	0.0846*	0.0814*	0.0852*	0.0798*
	Tukey	0.0986*	0.0974*	0.1006*	0.1022*	0.1010*	0.0956*
	Duncan	0.1844	0.1888	0.2036	0.1946	0.1866	0.1858
	SNK	0.1144*	0.1138*	0.1216*	0.1234*	0.1172*	0.1132*
	Scheffe	0.0844*	0.0808*	0.0828*	0.0838*	0.0848*	0.0766*

หมายเหตุ * สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ตามเกณฑ์ของ Bradley

จากตารางที่ 7 พบว่าตัวสถิติทดสอบของ Bonferroni Tukey SNK และ Scheffe สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ยกเว้นตัวสถิติ LSD และ Duncan

3.2 การเปรียบเทียบค่าประมาณกำลังการทดสอบ

การสรุปผลค่าประมาณกำลังการทดสอบของสถิติทดสอบจะแสดงในกรณีที่สถิติทดสอบสามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้เท่านั้น

- กำหนดระดับนัยสำคัญที่ระดับ 0.01

ตารางที่ 8. ค่าประมาณกำลังการทดสอบของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.01

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	–	–	–	–	–	–
	Bonferroni	0.5930	0.9964	1*	0.7882	0.9992	1*
	Tukey	0.6218	0.9976	1*	0.8066	0.9994	1*
	Duncan	–	–	–	–	–	–
	SNK	0.7934*	0.9990*	1*	0.9196*	1*	1*
	Scheffe	0.5676	0.9942	1*	0.7586	0.9990	1*
4	LSD	–	–	–	–	–	–
	Bonferroni	0.1360	0.7608	1*	0.2338	0.8756	1*
	Tukey	0.1560	0.7748	1*	0.2560	0.8848	1*
	Duncan	–	–	–	–	–	–
	SNK	0.2992*	0.8932*	1*	0.4332*	0.9562*	1*
	Scheffe	0.1246	0.7272	1*	0.2122	0.8530	1*
8	LSD	–	–	–	–	–	–
	Bonferroni	0.0188	0.2012	0.9874	0.0320	0.3190	0.9904
	Tukey	0.0210	0.2204	0.9882	0.0362	0.3378	0.9916
	Duncan	–	–	–	–	–	–
	SNK	0.0572*	0.3762*	0.9970*	0.0942*	0.5252*	0.9976*
	Scheffe	0.0160	0.1754	0.9832	0.0264	0.2796	0.9892
16	LSD	–	–	–	–	–	–
	Bonferroni	0.0024	0.0198	0.6062	0.0048	0.0382	0.7234
	Tukey	0.0024	0.0220	0.6198	0.0054	0.0416	0.7330
	Duncan	–	–	–	–	–	–
	SNK	0.0084*	0.0640*	0.7732*	0.0144*	0.1066*	0.8530*
	Scheffe	0.0018	0.0150	0.5588	0.0034	0.0292	0.6852

หมายเหตุ – ไม่พิจารณาค่าประมาณกำลังการทดสอบของสถิติทดสอบ เนื่องจากไม่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

* ค่าประมาณกำลังการทดสอบสูงที่สุดในสถานการณ์นั้น

จากตารางที่ 8 พบว่า ตัวสถิติทดสอบ SNK ส่วนใหญ่ให้ค่าประมาณกำลังการทดสอบสูงที่สุด รองลงมาคือ Bonferroni Tukey และ Scheffe ที่ระดับความแปรปรวน 2 และ 4

- กำหนดระดับนัยสำคัญที่ระดับ 0.05

ตารางที่ 9. ค่าประมาณกำลังการทดสอบของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.05

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	–	–	–	–	–	–
	Bonferroni	0.8734	0.9996	1*	0.9510	1*	1*
	Tukey	0.8960	0.9996	1*	0.9608	1*	1*
	Duncan	–	–	–	–	–	–
	SNK	0.9648*	1*	1*	0.9908*	1*	1*
	Scheffe	0.8712	0.9992	1*	0.9480	1*	1*
4	LSD	–	–	–	–	–	–
	Bonferroni	0.4138	0.9316	1*	0.5678	0.9732	1*
	Tukey	0.4626	0.9426	1*	0.6126	0.9790	1*
	Duncan	–	–	–	–	–	–
	SNK	0.6756*	0.9792*	1*	0.8004*	0.9926*	1*
	Scheffe	0.4108	0.9272	1*	0.5576	0.9698	1*
8	LSD	–	–	–	–	–	–
	Bonferroni	0.0958	0.4800	0.9990	0.1452	0.6284	0.9982
	Tukey	0.1168	0.5152	0.9994	0.1716	0.6556	0.9986
	Duncan	–	–	–	–	–	–
	SNK	0.2498*	0.7224*	1*	0.3418*	0.8250*	0.9996*
	Scheffe	0.0942	0.4648	0.9986	0.1410	0.6084	0.9980
16	LSD	–	–	–	–	–	–
	Bonferroni	0.0184	0.1042	0.8298	0.0262	0.1726	0.8990
	Tukey	0.0226	0.1240	0.8462	0.0328	0.1970	0.9118
	Duncan	–	–	–	–	–	–
	SNK	0.0646*	0.2816*	0.9344*	0.0904*	0.3720*	0.9678*
	Scheffe	0.0180	0.0950	0.8138	0.0252	0.1600	0.8856

หมายเหตุ – ไม่พิจารณาค่าประมาณกำลังการทดสอบของสถิติทดสอบ เนื่องจากไม่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

* ค่าประมาณกำลังการทดสอบสูงที่สุดในสถานการณ์นั้น

จากตารางที่ 9 พบว่า ตัวสถิติทดสอบ SNK ส่วนใหญ่ให้ค่าประมาณกำลังการทดสอบสูงที่สุด รองลงมาคือ Bonferroni Tukey และ Scheffe ที่ระดับความแปรปรวน 2 และ 4

- กำหนดระดับนัยสำคัญที่ระดับ 0.1

ตารางที่ 10. ค่าประมาณกำลังการทดสอบของสถิติทดสอบ ที่ระดับนัยสำคัญ 0.1

ความแปรปรวน	สถิติทดสอบ	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
2	LSD	–	–	–	–	–	–
	Bonferroni	0.9410	0.9998	1*	0.9838	1*	1*
	Tukey	0.9554	1*	1*	0.9888	1*	1*
	Duncan	–	–	–	–	–	–
	SNK	0.9888*	1*	1*	0.9978*	1*	1*
	Scheffe	0.9458	0.9998	1*	0.9848	1*	1*
4	LSD	–	–	–	–	–	–
	Bonferroni	0.5668	0.9670	1*	0.7158	0.9872	1*
	Tukey	0.6268	0.9744	1*	0.7646	0.9902	1*
	Duncan	–	–	–	–	–	–
	SNK	0.8172*	0.9932*	1*	0.8924*	0.9980*	1*
	Scheffe	0.5822	0.9674	1*	0.7250	0.9872	1*
8	LSD	–	–	–	–	–	–
	Bonferroni	0.1786	0.6284	0.9998	0.2606	0.7544	0.9996
	Tukey	0.2182	0.6802	0.9998	0.3128	0.7950	0.9996
	Duncan	–	–	–	–	–	–
	SNK	0.4252*	0.8432*	1*	0.5298*	0.9094*	1*
	Scheffe	0.1898	0.6288	0.9998	0.2690	0.7528	0.9994
16	LSD	–	–	–	–	–	–
	Bonferroni	0.0354	0.1990	0.9056	0.0586	0.2884	0.9472
	Tukey	0.0490	0.2420	0.9188	0.0766	0.3286	0.9562
	Duncan	–	–	–	–	–	–
	SNK	0.1386*	0.4568*	0.9716*	0.1892*	0.5498*	0.9862*
	Scheffe	0.0376	0.1992	0.9018	0.0612	0.2878	0.9468

หมายเหตุ – ไม่พิจารณาค่าประมาณกำลังการทดสอบของสถิติทดสอบ เนื่องจากไม่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้

* ค่าประมาณกำลังการทดสอบสูงที่สุดในสถานการณ์นั้น

จากตารางที่ 10 พบว่า ตัวสถิติทดสอบ SNK ส่วนใหญ่ให้ค่าประมาณกำลังการทดสอบสูงที่สุด รองลงมาคือ Bonferroni Tukey และ Scheffe ที่ระดับความแปรปรวน 2 และ 4

4. สรุปผลการวิจัย

4.1 การควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1

สถิติทดสอบของ Bonferroni สถิติทดสอบของ Tukey สถิติทดสอบของ SNK และสถิติทดสอบของ Scheffe สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ ทุกสถานการณ์ที่ศึกษา

แต่สถิติทดสอบ LSD และสถิติทดสอบของ Duncan ไม่สามารถควบคุมค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ ทุกสถานการณ์ที่ศึกษา

4.2 การเปรียบเทียบค่าประมาณกำลังการทดสอบ

สถิติทดสอบของ Bonferroni (B) สถิติทดสอบของ Tukey (T) สถิติทดสอบ SNK และสถิติทดสอบของ Scheffe (S) มีค่าประมาณกำลังการทดสอบสูงที่สุดร่วมกันที่ขนาดตัวอย่าง (30,30,30) และ (30,34,38) ความแปรปรวนเป็น 2 และ 4 และที่ขนาดตัวอย่าง (10,12,14) ความแปรปรวนเป็น 2 ที่ระดับนัยสำคัญ 0.05 และ 0.1 สถิติทดสอบของ Tukey และสถิติทดสอบของ SNK มีค่าประมาณกำลังการทดสอบสูงที่สุดร่วมกันที่ขนาดตัวอย่าง (10,10,10) ความแปรปรวนเป็น 2 ที่ระดับนัยสำคัญ 0.1 ส่วนในสถานการณ์อื่น ๆ สถิติทดสอบของ SNK มีค่าประมาณกำลังการทดสอบสูงที่สุด

โดยสถิติทดสอบที่มีค่าประมาณกำลังการทดสอบสูงที่สุดในแต่ละสถานการณ์ สามารถสรุปได้ดังตารางที่ 11

ตารางที่ 11. สถิติทดสอบที่มีค่าประมาณกำลังการทดสอบสูงที่สุดในแต่ละสถานการณ์

ระดับนัยสำคัญ	ความแปรปรวน	ขนาดตัวอย่าง (n_1, n_2, n_3)					
		เท่ากัน			ไม่เท่ากัน		
		(5,5,5)	(10,10,10)	(30,30,30)	(5,6,7)	(10,12,14)	(30,34,38)
0.01	2	SNK	SNK	B,T,SNK,S	SNK	SNK	B,T,SNK,S
	4	SNK	SNK	B,T,SNK,S	SNK	SNK	B,T,SNK,S
	8	SNK	SNK	SNK	SNK	SNK	SNK
	16	SNK	SNK	SNK	SNK	SNK	SNK
0.05	2	SNK	SNK	B,T,SNK,S	SNK	B,T,SNK,S	B,T,SNK,S
	4	SNK	SNK	B,T,SNK,S	SNK	SNK	B,T,SNK,S
	8	SNK	SNK	SNK	SNK	SNK	SNK
	16	SNK	SNK	SNK	SNK	SNK	SNK
0.1	2	SNK	T,SNK	B,T,SNK,S	SNK	B,T,SNK,S	B,T,SNK,S
	4	SNK	SNK	B,T,SNK,S	SNK	SNK	B,T,SNK,S
	8	SNK	SNK	SNK	SNK	SNK	SNK
	16	SNK	SNK	SNK	SNK	SNK	SNK

4.3 ข้อเสนอแนะ

ผู้วิจัยได้ลองคำนวณค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 โดยกำหนดค่าเฉลี่ยของแต่ละประชากรเป็น 10 50 และ 100 พบว่าค่าประมาณของความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในแต่ละสถิติทดสอบไม่แตกต่างจากที่กำหนดค่าเฉลี่ยของแต่ละประชากรเป็น 4 และผู้วิจัยได้ลองคำนวณค่าประมาณกำลังการทดสอบ โดยกำหนดค่าเฉลี่ยของประชากรเป็น (4,12,8) (8,4,12) (8,12,4) (12,4,8) (12,8,4) (50,54,58) และ (100,104,108) พบว่าค่าประมาณกำลังการทดสอบในแต่ละสถิติทดสอบไม่แตกต่างจากที่กำหนดค่าเฉลี่ยของประชากรเป็น (4,8,12) แต่เมื่อผู้วิจัยกำหนดระยะห่างของค่าเฉลี่ยของแต่ละประชากรเป็น 1 2 และ 3 นั่นคือค่าเฉลี่ยของประชากรเป็น (4,5,6) (4,6,8) และ (4,7,10) ตามลำดับ พบว่าค่าประมาณกำลังการทดสอบในแต่ละสถิติทดสอบลดลงจากที่กำหนดระยะห่างของค่าเฉลี่ยของแต่ละประชากรเป็น 4 โดยเมื่อระยะห่างของค่าเฉลี่ยลดลงค่าประมาณกำลังการทดสอบจะลดลงตาม และเมื่อพิจารณาค่าประมาณกำลังการทดสอบในแต่ละสถานการณ์และทุกระยะห่างของค่าเฉลี่ยที่ผู้วิจัยลองศึกษา พบว่าสถิติทดสอบของสติวเดนต์–นิวแมน–คูล (SNK) มีค่าประมาณกำลังการทดสอบสูงที่สุด

เอกสารอ้างอิง (References)

[1] สายชล สินสมบูรณ์ทอง. 2558. การวางแผนแบบการทดลอง เล่ม 1. จามจุรีโปรดักท์, กรุงเทพฯ. [Saichon Sinsombunthong. 2015. Experimental Designs 1. Jamjuree Product, Bangkok. (in Thai)]

[2] Boardman, T.J. and Moffitt, D.R. 1971. Graphical Monte Carlo type I error rates for multiple comparison procedures. *Biometrics*, 27(3), 738-744.

[3] Bernrdson, C.S. 1975. Type I error rates when multiple comparison procedures follow a significant F test of ANOVA. *Biometrics*, 31(1), 229-232.

[4] สุญานี จิตตะยโศธร. 2525. การศึกษาโดยวิธีมอนติคาร์โล : การเปรียบเทียบความคลาดเคลื่อนประเภทที่ 1 จากข้อมูลที่ฝ่าฝืนข้อตกลงเบื้องต้นของการเปรียบเทียบพหุคูณ. วิทยานิพนธ์, สาขาวิจัยการศึกษา, จุฬาลงกรณ์มหาวิทยาลัย. [Suyanee Jittayasotorn. 1982. A Monte Carlo study : a comparisons of type I error rates from data violating basic assumptions of multiple comparisons, M. Ed. Thesis, Educational Research, Chulalongkorn University. (in Thai)]

[5] Bradley, J.V. 1978. Robustness?. *The British Journal of Mathematical and Statistical Psychology,* 31(2), 144–152.

[6] Kramer, C.Y. 1956. Extension of Multiple Range Tests to Group Means with Unequal Numbers of Replications. *Biometrics*, 12(3), 307–310.

[7] Duncan, D.B. 1955. Multiple Range and Multiple F tests. *Biometric*, 11(1), 1–42.

[8] Keuls, M. 1952. The Use of the Studentized Range in Connection with an Analysis of Variance. *Euphytica*, 1(2), 112–122.