Research Article

Received: May 5, 2019; Accepted: November 8, 2019

การเปรียบเทียบประสิทธิภาพสำหรับวิธีการเปรียบเทียบพหุคูณแบบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ของแผนแบบสุ่มสมบูรณ์

Comparison of Efficiency for Parametric and Nonparametric Tests in Multiple Comparisons of a Completely Randomized Designs

ธารทิพย์ โนภาศ และสายชล สินสมบูรณ์ทอง*

ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ถนนฉลองกรุง เขตลาดกระบัง กรุงเทพมหานคร 10520

Tantip Nopas and Saichon Sinsomboonthong*

Department of Statistics, Faculty of Science, King Mongkut's Institute of Technology Ladkrabang,

Chalongkrung Road, Ladkrabang, Bangkok 10520

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบของวิธีการเปรียบเทียบพหุคูณแบบใช้พารามิเตอร์และไม่ใช้พารามิเตอร์ของการวางแผนแบบสุ่มสมบูรณ์ ซึ่งมีตัวสถิติทดสอบที่ศึกษา 6 การทดสอบ คือ ตัวสถิติทดสอบสติวเดนต์-นิวแมน-คูล ตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคน ตัวสถิติทดสอบวอลเลอร์ดันแคน ตัวสถิติทดสอบโคโนเวอร์ ตัวสถิติทดสอบแวนเดอแวร์เดน และตัวสถิติทดสอบด้วยค่าลำดับที่นีเมนยี โดยศึกษาจากข้อมูลที่สุ่มมาจากประชากรที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติในการคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และการคำนวณกำลังการทดสอบ กำหนดขนาดตัวอย่างที่มีขนาดเท่ากันและไม่เท่ากัน ซึ่งแบ่งเป็นกลุ่มตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ คือ (10,10,10), (30,30,30), (50,50,50), (5,10,15), (25,30,35) และ (45,50,55) กำหนดระดับนัยสำคัญในการวิจัย 2 ระดับ คือ 0.01 และ 0.05 ผลการวิจัยเมื่อพิจารณาตามวัตถุประสงค์ที่ 1 พบว่าตัวสถิติทดสอบสติวเดนต์-นิวแมน-คูล ตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคน ตัวสถิติทดสอบวอลเลอร์ดันแคน ตัวสถิติทดสอบโคโนเวอร์ และตัวสถิติทดสอบแวนเดอแวร์เดน มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ในทุกสถานการณ์ส่วนการพิจารณากำลัง การทดสอบพบว่าตัวสถิติทดสอบวอลเลอร์ดันแคน มีกำลังการทดสอบสูงสุดในทุกสถานการณ์ และเมื่อพิจารณาตามวัตถุประสงค์ที่ 2 พบว่าตัวสถิติทดสอบวอลเลอร์ดันแคนมีความเหมาะสมมากที่สุดสำหรับใช้ทดสอบข้อมูลที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติ

***ผู้รับผิดชอบบทความ :** tantipnopas2537@hotmail.com **doi:** 10.14456/tstj.2020.152

คำสำคัญ : การเปรียบเทียบพหุคูณ; ความน่าจะเป็นของความผิดพลาดแบบที่ 1; กำลังการทดสอบ

Abstract

This research purposed to study and compare the efficiencies of multiple comparisons tests with and without parameter of completely randomized design consisting of 6 tests, i.e., Student-Newman-Keul's test (SNK test), Duncan's new multiple range test, Waller-Duncan test, Conover test, van der Waerden test and Nemenyi test. The study was conducted in data randomly obtained from normal distribution, gamma distribution and log-normal distribution. For the probability calculation of type I error and power of a test using identical and different sample sizes, divided into small size, medium size and large size, i.e., (10,10,10), (30,30,30), (50,50,50), (5,10,15), (25,30,35) and (45,50,55), and there were 2 levels of significance: 0.01 and 0.05. The results of the study, according to objective 1, indicated that for every situation, the best probability control capability of type I error was found in Student-Newman-Keul's test, Duncan's new multiple range test, Waller-Duncan test, Conover test and van der Waerden test. As Weller-Duncan test provided the best power of a test for every situation and the results of the study, according to objective 2, Waller-Duncan test is most suitable for testing data with normal distribution, gamma distribution and log-normal distribution.

Keywords: multiple comparisons; probability of type I error; power of a test

1. บทนำ

การวิเคราะห์ข้อมูลในการวางแผนการทดลองซึ่งเป็นข้อมูลที่ได้จากการอนุมานเชิงสถิติที่เกี่ยวกับประชากรนิยมใช้การทดสอบสมมติฐาน ซึ่งการทดสอบสมมติฐานจะนำไปสู่ข้อสรุปของงานวิจัย ในที่นี้ผู้วิจัยได้ทดสอบสมมติฐานเกี่ยวกับการเปรียบเทียบความแตกต่างของค่าเฉลี่ยของประชากรตั้งแต่ 3 กลุ่มขึ้นไป โดยทั่วไปแล้วจะนิยมใช้การวิเคราะห์ความแปรปรวน (analysis of variance) โดยใช้ตัวสถิติทดสอบเอฟ ภายหลังจากการทดสอบสมมติฐานพบว่าหากการทดสอบสมมติฐานดังกล่าวเกิดการปฏิเสธสมมติฐานว่าง H_0 คือ มีความแตกต่างของค่าเฉลี่ยอย่างมีนัยสำคัญทางสถิติ จะต้องเปรียบเทียบพหุคูณต่อไปเพื่อหาว่ามีค่าเฉลี่ยของประชากรคู่ใดบ้างที่แตกต่างกัน ซึ่งตัวสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณมีทั้งแบบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ ซึ่งการเลือกใช้ตัวสถิติทดสอบนั้นจะขึ้นอยู่กับข้อมูลที่นำมาวิเคราะห์ว่ามีลักษณะเป็นไปตามข้อกำหนดเบื้องต้นของตัวสถิติทดสอบเอฟ คือ ข้อมูลแต่ละประชากรสุ่มมาจากประชากรที่มีการแจกแจงปรกติข้อมูลแต่ละประชากรจะต้องสุ่มมาจากประชากรที่มีความแปรปรวนเท่ากัน และข้อมูลแต่ละประชากรจะต้องสุ่มมาจากประชากรที่อิสระกัน การเปรียบเทียบพหุคูณจะใช้ตัวสถิติทดสอบแบบอิงพารามิเตอร์ หากไม่เป็นไปตามข้อกำหนดเบื้องต้นของตัวสถิติทดสอบเอฟก็จะใช้ตัวสถิติทดสอบแบบไม่อิงพารามิเตอร์ โดยการเลือกใช้ตัวสถิติทดสอบ

ควรที่จะต้องมีแนวทางในการเลือกใช้ให้เหมาะสมกับข้อมูลของผู้วิจัย [1]

การเลือกแผนแบบการทดลองนั้น ผู้วิจัยได้ศึกษาจากงานวิจัยต่าง ๆ ตำรา และวรรณกรรมที่เกี่ยวข้อง พบว่าแผนแบบสุ่มสมบูรณ์เป็นแผนที่ได้รับความนิยมนักวิจัยใช้กันอย่างแพร่หลาย เป็นแผนแบบการทดลองที่ยืดหยุ่นได้ วิธีการคำนวณง่าย และสะดวก การวิจัยครั้งนี้ผู้วิจัยได้จำลองข้อมูลจากโปรแกรมอาร์ (R) เวอร์ชั่น 3.5.2 โดยจำลองข้อมูลให้มีแต่ละขนาดจากประชากรที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติให้มีพารามิเตอร์ตามที่ต้องการ รวมทั้งได้เลือกตัวสถิติทดสอบที่มีความเหมาะสมที่จะนำมาวิจัย โดยผู้วิจัยได้ค้นคว้าและศึกษาองค์ความรู้และทฤษฎีต่าง ๆ เกี่ยวกับตัวสถิติทดสอบทั้งตัวสถิติทดสอบอิงพารามิเตอร์และตัวสถิติทดสอบไม่อิงพารามิเตอร์จากงานวิจัย ตำรา และวรรณกรรมที่เกี่ยวข้อง เพื่อเป็นแนวทางในการเลือกตัวสถิติทดสอบมาใช้ในการวิจัย ผลการศึกษางานวิจัยพบว่าตัวสถิติทดสอบสติวเดนต์-นิวแมน-คูลมีกำลังการทดสอบสูงที่สุดเมื่อประชากรมีลักษณะใกล้เคียงกับการแจกแจงปรกติที่มีความโด่งไม่มาก [2,3] และตัวสถิติทดสอบสติวเดนต์-นิวแมน-คูลสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในทุกสถานการณ์ [4,5]

ตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคนควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้มากที่สุดเมื่อเทียบกับตัวสถิติทดสอบอื่น ๆ และตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในทุกกรณี [4-8] และเมื่อตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคนมีค่าเฉลี่ยที่เพิ่มขึ้นจะทำให้เกิดความผิดพลาดแบบที่ 1 เพิ่มขึ้นด้วย [9] โดยที่ตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคนมีกำลังการทดสอบสูงที่สุด [3,10]

ตัวสถิติทดสอบวอลเลอร์ดันแคนมีกำลังการทดสอบสูงที่สุดในทุกกรณีเมื่อจำนวนกลุ่มตัวอย่าง 3-4 กลุ่ม [6] และตัวสถิติทดสอบวอลเลอร์ดันแคนสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในทุกสถานการณ์ [4]

สำหรับตัวสถิติทดสอบโคโนเวอร์ ตัวสถิติทดสอบแวนเดอแวร์เดน และตัวสถิติทดสอบด้วยค่าลำดับที่นีเมนยีนั้น ผู้วิจัยได้ศึกษาค้นคว้าจากงานวิจัย เอกสาร และวรรณกรรมต่าง ๆ ที่เกี่ยวข้องแล้ว พบว่าไม่มีงานวิจัยใดที่เป็นที่ยอมรับที่จะนำตัวสถิติทดสอบเหล่านี้มาทำงานวิจัยหรือศึกษาโดยทั่วไป ผู้วิจัยจึงมีความสนใจที่จะนำตัวสถิติทดสอบนี้มาทดสอบรวมกับตัวสถิติทดสอบอื่น ๆ

ดังนั้นผู้วิจัยจึงสนใจศึกษาการเปรียบเทียบความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของวิธีการเปรียบเทียบพหุคูณแบบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ของแผนแบบสุ่มสมบูรณ์โดยเปรียบเทียบตัวสถิติทดสอบ 6 การทดสอบ คือ ตัวสถิติทดสอบสติวเดนต์-นิวแมน-คูล (Student-Newman-Keul's test, SNK) ตัวสถิติทดสอบแบบพิสัยพหุคูณของคันแคน (Duncan's new multiple range test, DC) ตัวสถิติทดสอบวอลเลอร์ดันแคน (Waller-Duncan test, WD) ตัวสถิติทดสอบโคโนเวอร์ (Conover test, CNV) ตัวสถิติทดสอบแวนเดอแวร์เดน (van der Waerden test, VD) และตัวสถิติทดสอบด้วยค่าลำดับที่นีเมนยี (Nemenyi test, NM) โดยใช้โปรแกรมอาร์ (R) เวอร์ชั่น 3.5.2 ในการจำลองและวิเคราะห์ข้อมูล

2. วัตถุประสงค์ของงานวิจัย

2.1 เพื่อเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบของวิธีการ

เปรียบเทียบพหุคูณแบบอิงพารามิเตอร์และไม่อิงพารามิเตอร์ทั้ง 6 การทดสอบของแผนแบบสุ่มสมบูรณ์

2.2 เพื่อหาตัวสถิติทดสอบที่เหมาะสมสำหรับใช้ทดสอบกับข้อมูลที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติ

3. การดำเนินการวิจัย

3.1 การวางแผนการวิจัย

การวิจัยครั้งนี้กำหนดสถานการณ์การศึกษาเปรียบเทียบ ดังนี้

3.1.1 กำหนดจำนวนประชากรเท่ากับ 3 ประชากร

3.1.2 กำหนดขนาดตัวอย่างสุ่มจากแต่ละประชากรขนาดเท่ากันและไม่เท่ากันดังตารางที่ 1

Table 1 Sample sizes for study of 3 populations

Sample sizes (n_1,n_2,n_3)	
Equal	Unequal
(10,10,10)	(5,10,15)
(30,30,30)	(25,30,35)
(50,50,50)	(45,50,55)

3.1.3 กำหนดระดับนัยสำคัญ คือ 0.01 และ 0.05

3.1.4 กำหนดข้อมูลที่สุ่มจากประชากรที่มีการแจกแจงปรกติด้วยพารามิเตอร์ (μ,σ^2) การแจกแจงแกมมาด้วยพารามิเตอร์ (α,β) และการแจกแจงล็อกปรกติด้วยพารามิเตอร์ (μ,σ^2)

3.1.5 กำหนดจำนวนทรีทเมนต์ที่ศึกษา คือ 3 ทรีทเมนต์

3.1.6 สร้างแผนแบบสุ่มสมบูรณ์ ดังนี้ (1) แต่ละทรีทเมนต์มีการทำซ้ำจำนวน r ซ้ำ คือ (1.1) กรณีที่มีจำนวนซ้ำเท่ากัน คือ 10, 30 และ 50 และ (1.2) กรณีที่มีจำนวนซ้ำไม่เท่ากัน คือ 5, 10, 15, 25, 30, 35, 45, 50 และ 55 (2) แต่ละทรีทเมนต์จะประกอบด้วยจำนวนหน่วยทดลองที่เท่ากันและไม่เท่ากัน คือ (2.1) กรณีที่มีจำนวนหน่วยทดลองที่เท่ากัน คือ 10, 30 และ 50 และ (2.2) กรณีที่มีจำนวนหน่วยทดลองที่ไม่เท่ากัน คือ 5, 10, 15, 25, 30, 35, 45, 50 และ 55 (3) หน่วยทดลองทุกหน่วยต่างก็มีโอกาสที่จะได้รับทรีทเมนต์ใดทรีทเมนต์หนึ่งโดยเท่าเทียมกัน และ (4) หน่วยทดลองที่ใช้ในการทดลองมีลักษณะเหมือน ๆ กัน

Table 2 Completely randomized trial plan for equal and unequal sample sizes

Treatments	Number of observations						Replications					
	Equal sample size			Unequal sample size			Equal replication			Unequal replication		
1	10	30	50	5	25	45	10	30	50	5	25	45
2	10	30	50	10	30	50	10	30	50	10	30	50
3	10	30	50	15	35	55	10	30	50	15	35	55

3.1.7 การคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบจะกำหนดค่าเฉลี่ยและความแปรปรวนของแต่ละประชากรเพื่อใช้ในการหาค่าพารามิเตอร์ สำหรับแต่ละ

การแจกแจงโดยแบ่งออกเป็น 4 สถานการณ์ ซึ่งในสถานการณ์ที่มีความแปรปรวนต่างกันจะกำหนดค่าระดับความแตกต่างของความแปรปรวนโดยใช้ค่าพารามิเตอร์ไม่ศูนย์กลาง ϕ (non-centrality parameter) เป็นเกณฑ์วัดความแตกต่างของความแปรปรวน ซึ่งเกณฑ์ที่ใช้ในการวัดความแตกต่างของความแปรปรวนของประชากร คือ (1) ถ้าค่า ϕ อยู่ในช่วง $0 < \phi < 1.5$ แสดงว่าความแปรปรวนของประชากรมีความแตกต่างกันน้อย (2) ถ้าค่า ϕ อยู่ในช่วง $0 \leq \phi < 3$ แสดงว่าความแปรปรวนของประชากรมีความแตกต่างกันปานกลาง และ (3) ถ้าค่า $\phi \geq 3$ แสดงว่าความแปรปรวนของประชากรมีความแตกต่างกันมาก [11]

(1) การเปรียบเทียบความน่าจะเป็นของความผิดพลาดแบบที่ 1 สำหรับการแจกแจงปรกติจะกำหนดค่าพารามิเตอร์ (μ,σ^2) ได้ดังนี้

(1.1) สถานการณ์ที่ 1 กรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 3

(1.2) สถานการณ์ที่ 2 กรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 4

Table 3 Parameters for calculation the probability of type I error of normal distribution in situation 1

Model	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(4,8)	(4,8)	(4,8)	4,4,4	8,8,8

Table 4 Parameters for calculation the probability of type I error of normal distribution in situation 2

Models	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameters (ϕ)
	Population 1	Population 2	Population 3			
1	(4,1)	(4,2)	(4,3)	4,4,4	1,2,3	0.816
2	(8,1)	(8,3)	(8,5)	8,8,8	1,3,5	1.633
3	(12,1)	(12,6)	(12,14)	12,12,12	1,6,14	5.354

Table 5 Parameters for calculation the probability of type I error of gamma distribution in situation 1

Model	Parameters (α,β)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(2,2)	(2,2)	(2,2)	4,4,4	8,8,8

(2) การเปรียบเทียบความน่าจะเป็นของความผิดพลาดแบบที่ 1 สำหรับการแจกแจงแกมมาจะกำหนดค่าพารามิเตอร์ (α,β) ได้ดังนี้

(2.1) สถานการณ์ที่ 1 กรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 5

(2.2) สถานการณ์ที่ 2 กรณีค่าเฉลี่ย

เท่ากัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 6

(3) การเปรียบเทียบความน่าจะเป็นของความผิดพลาดแบบที่ 1 สำหรับการแจกแจงล็อกปรกติจะกำหนดค่าพารามิเตอร์ (α,β) ได้ดังนี้

(3.1) สถานการณ์ที่ 1 กรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 7

(3.2) สถานการณ์ที่ 2 กรณีค่าเฉลี่ยเท่ากัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 8

(4) การคำนวณกำลังการทดสอบสำหรับการแจกแจงปรกติจะกำหนดค่าพารามิเตอร์ได้ดังนี้

(4.1) สถานการณ์ที่ 3 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 9

(4.2) สถานการณ์ที่ 4 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 10

(5) การคำนวณกำลังการทดสอบสำหรับการแจกแจงแกมมาจะกำหนดค่าพารามิเตอร์ (α,β) ได้ดังนี้

(5.1) สถานการณ์ที่ 3 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 11

(5.2) สถานการณ์ที่ 4 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 12

Table 6 Parameters for calculation the probability of type I error of gamma distribution in situation 2

Models	Parameters (α,β)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameter (ϕ)
	Population 1	Population 2	Population 3			
1	(16,1/4)	(8,1/2)	(4,1)	4,4,4	1,2,4	1.247
2	(16,1/2)	(8,1)	(4,2)	8,8,8	4,8,16	2.494
3	(16,1)	(8,2)	(4,4)	16,16,16	16,32,64	4.989

Table 7 Parameters for calculation the probability of type I error of lognormal distribution in situation 1

Model	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(1.1065,0.5596)	(1.1065,0.5596)	(1.1065,0.5596)	4,4,4	12,12,12

Table 8 Parameters for calculation the probability of type I error of lognormal distribution in situation 2

Models	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameter (ϕ)
	Population 1	Population 2	Population 3			
1	(1.3560, 0.3560)	(1.3274 ,0.1178)	(1.2747, 0.2231)	4,4,4	1,2,4	1.247
2	(2.7648, 0.0155)	(2.7572, 0.0308)	(2.7423, 0.0606)	16,16,16	4,8,16	2.494
3	(2.7423, 0.0606)	(2.7137, 0.1178)	(2.6610, 0.2231)	16,16,16	16,32,64	4.989

Table 9 Parameters for calculation the power of a test of normal distribution in situation 3

Model	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(4,4)	(5,4)	(6,4)	4,5,6	4,4,4

Table 10 Parameters for calculation the power of a test of normal distribution in situation 4

Models	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameter (ϕ)
	Population 1	Population 2	Population 3			
1	(4,1)	(5,2)	(6,3)	4,5,6	1,2,3	0.816
2	(8,1)	(9,3)	(10,5)	8,9,10	1,3,5	1.633
3	(12,1)	(13,6)	(14,14)	12,13,14	1,6,14	5.354

Table 11 Parameters for calculation the power of a test of gamma distribution in situation 3

Model	Parameters (α,β)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(2,2)	(8,1)	(32,1/2)	4,8,16	8,8,8

Table 12 Parameters for calculation the power of a test of gamma distribution in situation 4

Models	Parameters (α,β)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameter (ϕ)
	Population 1	Population 2	Population 3			
1	(2,1)	(3,1)	(4,1)	2,3,4	2,3,4	0.577
2	(4,1)	(16,1)	(18,1)	4,16,18	8,16,18	1.528
3	(2,1)	(4,1)	(4,2)	2,4,8	2,4,16	4.372

(6) ในการคำนวณกำลังการทดสอบสำหรับการแจกแจงล็อกปรกติจะกำหนดค่าพารามิเตอร์ (μ,σ^2) ได้ดังนี้

(6.1) สถานการณ์ที่ 3 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนเท่ากัน แสดงดังตารางที่ 13

(6.2) สถานการณ์ที่ 4 กรณีค่าเฉลี่ยต่างกัน ความแปรปรวนต่างกัน แสดงดังตารางที่ 14

3.1.8 หาจำนวนครั้งที่ใช้ในการวิจัย การหาจำนวนครั้งในการคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของสถาณการณ์ดังกล่าวข้างต้น พบว่าความน่าจะเป็นของความความผิดพลาดแบบที่ 1 ของตัวสถิติทดสอบทั้ง 6 การทดสอบ จะมีค่าคงที่เมื่อจำนวนครั้งทั้งหมดเท่ากับ 16,013 ครั้ง และจำนวนครั้งที่ผ่านการทดสอบการวิเคราะห์ความแปรปรวน

(ANOVA) เท่ากับ 800 ครั้ง ดังนั้นผู้วิจัยจึงกำหนดให้การจำลองข้อมูลเพื่อนำมาทดสอบต้องผ่านการวิเคราะห์ความแปรปรวน 800 ครั้ง

3.1.9 คำนวณความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของตัวสถิติทดสอบตามเกณฑ์ของแบรดลีย์ (Bradley) [12] ในแต่ละสถานการณ์ โดยมีรายละเอียดของเกณฑ์ของแบรดลีย์ คือ ที่ระดับนัยสำคัญ 0.01 ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองอยู่ในช่วง (0.005,0.015) ที่ระดับนัยสำคัญ 0.05 ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองอยู่ในช่วง (0.025,0.075) จะถือว่าตัวสถิติทดสอบนั้นสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

3.1.10 คำนวณกำลังการทดสอบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 เท่านั้น

3.1.11 เปรียบเทียบกำลังการทดสอบของตัวสถิติทดสอบ ถ้าพบว่าตัวสถิติทดสอบใดมีกำลังการทดสอบสูงที่สุดจะเป็นตัวสถิติทดสอบที่ดีที่สุด

Table 13 Parameters for calculation the power of a test of lognormal distribution in situation 3

Model	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$
	Population 1	Population 2	Population 3		
1	(1.2747,0.2231)	(1.5352,0.1484)	(1.7391,0.1054)	4,5,6	4,4,4

Table 14 Parameters for calculation the power of a test of lognormal distribution in situation 4

Models	Parameters (μ,σ^2)			Averages	Variances $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$	Non-centrality parameter (ϕ)
	Population 1	Population 2	Population 3			
1	(1.3560, 0.0606)	(1.5710, 0.0770)	(1.7391, 0.1054)	4,5,6	1,2,4	1.247
2	(2.0491, 0.0606)	(2.1501, 0.0942)	(2.2284, 0.1484)	8,9,10	4,8,16	2.494
3	(2.4814, 0.0069)	(2.5475, 0.0349)	(2.6046, 0.0690)	12,13,14	1,6,14	5.354

4. ผลการวิจัย

4.1 ความน่าจะเป็นของความผิดพลาดแบบที่ 1

4.1.1 ที่ระดับนัยสำคัญ 0.01

ตารางที่ 15 เมื่อพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ระดับนัยสำคัญ 0.01 ในสถานการณ์ที่มีขนาดตัวอย่างเท่ากัน พบว่าการแจกแจงปรกติและการแจกแจงแกมมา ตัวสถิติทดสอบทั้ง 6 การทดสอบ มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และการแจกแจงล็อกปรกติ ตัวสถิติทดสอบ SNK, DC, WD, CNV และ VD มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ส่วนในสถานการณ์ที่มีขนาดตัวอย่างต่างกันพบว่าการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจง ล็อกปรกติ ตัวสถิติทดสอบ SNK, DC, WD, CNV และ VD มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

Table 15 The ability to control the probability of type I error for randomized data with a normal, gamma and lognormal distributions at the level of significance 0.01

Distributions	Sample sizes	Variances			
		Equal variances	Unequal variances		
			Low	Moderate	High
Normal	Equal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD
Gamma	Equal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD
Lognormal	Equal	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD

Table 16 The ability to control the probability of type I error for randomized data with a normal, gamma and lognormal distributions at the level of significance 0.05

Distributions	Sample sizes	Variances			
		Equal variances	Unequal variances		
			Low	Moderate	High
Normal	Equal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
Gamma	Equal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
Lognormal	Equal	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD	SNK, DC, WD, CNV, NM, VD
	Unequal	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD	SNK, DC, WD, CNV, VD

4.1.2 ที่ระดับนัยสำคัญ 0.05

ตารางที่ 16 เมื่อพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ระดับนัยสำคัญ 0.05 ในสถานการณ์ที่มีขนาดตัวอย่างเท่ากัน พบว่าการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติ ตัวสถิติทดสอบทั้ง 6 การทดสอบ มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ส่วนสถานการณ์ที่มีขนาดตัวอย่างต่างกัน การแจกแจงปรกติและการแจกแจงแกมมาตัวสถิติทดสอบทั้ง 6 การทดสอบ มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ส่วนการแจกแจงล็อกปรกติพบว่าตัวสถิติทดสอบ SNK, DC, WD, CNV และ VD มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

4.2 กำลังการทดสอบ

4.2.1 ที่ระดับนัยสำคัญ 0.01

ตารางที่ 17 เมื่อพิจารณากำลังการทดสอบที่ระดับนัยสำคัญ 0.01 ในสถานการณ์ที่มีขนาดตัวอย่างเท่ากันและต่างกัน พบว่าการแจกแจงปรกติตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมา คือ ตัวสถิติทดสอบ CNV การแจกแจงแกมมา ตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมา คือ ตัวสถิติทดสอบ CNV, NM และ DC ส่วนการแจกแจงล็อกปรกติ ตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมา คือ ตัวสถิติทดสอบ CNV

4.2.2 ที่ระดับนัยสำคัญ 0.05

ตารางที่ 18 เมื่อพิจารณากำลังการทดสอบที่ระดับนัยสำคัญ 0.05 ในสถานการณ์ที่มีขนาดตัวอย่างเท่ากันและต่างกัน พบว่าในการแจกแจงปรกติตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมาคือ ตัวสถิติทดสอบ CNV, NM และ SNK ในการแจกแจงแกมมาตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมาคือ ตัวสถิติทดสอบ CNV, NM และ VD ส่วนในการแจกแจงล็อกปรกติตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด รองลงมาคือ ตัวสถิติทดสอบ CNV, SNK และ NM

Table 17 The power of a test of test statistics for randomized data with a normal, gamma and lognormal distributions at the level of significance 0.01

Distributions	Sample sizes	Variances			
		Equal variances	Unequal variances		
			Low	Moderate	High
Normal	Equal	WD	WD	WD	WD
	Unequal	WD	WD	WD	CNV
Gamma	Equal	WD	NM	DC, WD, NM	NM
	Unequal	WD	WD, CNV	WD, CNV	WD, CNV
Lognormal	Equal	WD	WD	WD	WD
	Unequal	WD	CNV	WD, CNV	CNV

Table 18 The power of a test of test statistics for randomized data with a normal, gamma and lognormal distributions at the level of significance 0.05

Distributions	Sample sizes	Variances			
		Equal variances	Unequal variances		
			Low	Moderate	High
Normal	Equal	WD	WD	SNK, WD, NM	NM
	Unequal	WD	WD, CNV	CNV	CNV
Gamma	Equal	WD	WD	WD, CNV	NM
	Unequal	WD	WD	WD	WD, CNV, VD
Lognormal	Equal	WD	SNK, WD, CNV, NM, VD	WD	SNK, WD, CNV, NM
	Unequal	WD	WD	CNV	CNV

5. สรุปและวิจารณ์

5.1 สรุป

ผลจากการวิจัยจากการพิจารณาถึงความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบทั้ง 6 การทดสอบ สำหรับแผนแบบสุ่มสมบูรณ์โดยศึกษาข้อมูลที่สุ่มมาจากประชากรที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติที่มีขนาดกลุ่มตัวอย่างเท่ากันและไม่เท่ากัน และลักษณะความแปรปรวนแบบต่าง ๆ ที่ระดับนัยสำคัญ 0.01 และ 0.05 พบว่าตัวสถิติทดสอบ SNK, DC, WD, CNV และ VD มีความความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในทุกสถานการณ์ และเมื่อพิจารณากำลังการทดสอบพบว่าตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงสุดในทุกสถานการณ์ และมีความเหมาะสมมากที่สุดสำหรับใช้ทดสอบข้อมูลที่มีการแจกแจงปรกติ การแจกแจงแกมมา และการแจกแจงล็อกปรกติ

5.2 วิจารณ์

ผลการวิจัยพบว่าตัวสถิติทดสอบ SNK, DC และ WD มีความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ในทุกสถานการณ์ ซึ่งสอดคล้องกับงานวิจัยของ ปุณยนุช [6] ที่กล่าวว่าแผนแบบสุ่มสมบูรณ์ที่ระดับนัยสำคัญ 0.05 เมื่อกลุ่มตัวอย่างมีขนาดเท่ากัน ตัวสถิติทดสอบที่ควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ในทุกกรณี คือ ตัวสถิติทดสอบ DC เมื่อกลุ่มตัวอย่างมีขนาดไม่เท่ากัน ตัวสถิติทดสอบที่ควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ในทุกกรณี คือ ตัวสถิติทดสอบ DC และ WD สอดคล้องกับงานวิจัยของ Armando และคณะ [4] ที่กล่าวว่าตัวสถิติทดสอบ WD และ DC สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 และสอดคล้องกับ Ozkaya [7] ที่กล่าวว่ากรณีที่กลุ่มตัวอย่างที่มีขนาดเท่ากันและไม่เท่ากัน ตัวสถิติทดสอบ DC สามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

ส่วนกำลังการทดสอบผลการวิจัยพบว่าตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุดในทุกสถานการณ์ ซึ่งสอดคล้องกับงานวิจัยของ ปุณยนุช [6] ที่กล่าวว่าแผนแบบสุ่มสมบูรณ์ที่ระดับนัยสำคัญ 0.05

ตัวสถิติทดสอบ WD มีกำลังการทดสอบสูงที่สุด แต่ขัดแย้งกับงานวิจัยของ Francisco และ Carlos [3] ที่กล่าวว่าที่ระดับนัยสำคัญ 0.05 และ 0.01 ตัวสถิติทดสอบ DC และ SNK มีกำลังการทดสอบสูงที่สุด

6. References

[1] Sinsomboonthong, S., 2015, Experimental Design 1, Chamchuree Products, Co., Ltd., Bangkok, 547 p. (in Thai)

[2] Worayangkun, S., 1988, Comparisons of the Power of a Test Between Parametric and Nonparametric Tests in Multiple Comparisons of a Randomized Complete Block Design, Master thesis, Chulalongkorn University, Bangkok. (in Thai)

[3] Francisco, D.A.S.S. and Carlos, A.V.D.A., 2016, Comparison of means of agricultural experimentation data through different tests using the software assistant, Afr. J. Agric. Res. 11: 3527-3531.

[4] Armando, C., Décio, B. and Clarice, G.B.D., 2008, Modifications for the Tukey test procedure and evaluation of the power and efficiency of multiple comparison procedures, Sci. Agric. 65: 428-432.

[5] Douglas, C.E., 2019, Multiple Comparisons: Philosophies and Illustrations, Available Source:http://www.physiology.org/journal/ajpregu, May 1, 2019.

[6] Pinchoo, P., 2007, Comparisons of type I error rates and power of the test in post hoc comparison procedures for completely randomized design, J. Res. Meth. 20(3): 331-351. (in Thai)

[7] Ozkaya, G. and Ercan, I., 2012, Examining multiple comparison procedures according to error rate, power type and false discovery rate, J. Mod. Appl. Stat. Met. 11: 348-360.

[8] David, J.S., 2012, Multiple comparison procedures- cutting the gordian knot, Agron. J. 107: 730-735.

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

[10] Khumsaard, N., 2009, Power of Test on Multiple Comparison, Master Thesis, Srinakharinwirot University, Bangkok, 112 p. (in Thai)

[11] Carmer, S.G. and Walker, W.M., 1985, Pairwise multiple comparisons of treatment means in agronomic research, J. Agron. Educ. 14: 19-26.

[12] Bradley, J.V., 1978, Robustness, J. Math. Stat. Psychol. 31: 144-152.