

การเปรียบเทียบประสิทธิภาพของสถิติสำหรับทดสอบความเท่ากันของความแปรปรวน*

An efficiency comparison of testing statistics for homogeneity of variances

อาณัติ แสงสว่าง**

บทคัดย่อ

การศึกษาวิจัยครั้งนี้ ผู้วิจัยมีความสนใจศึกษา และเปรียบเทียบประสิทธิภาพของตัวสถิติสำหรับทดสอบความเท่ากันของความแปรปรวนของประชากร 3 และ 4 กลุ่ม จำนวน 3 วิธี คือ สถิติบราวน์ – ฟอร์สตี (BF) สถิติ ANOMV (AV) และสถิติจี (GT) โดยพิจารณาประสิทธิภาพของตัวสถิติทั้ง 3 วิธี จากการเปรียบเทียบกำลังการทดสอบของตัวสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อกำหนดการแจกแจงของประชากรเป็นแบบปกติมีค่าเฉลี่ยเท่ากับ 0 ขนาดตัวอย่างเล็ก ($n \leq 15$) กลาง ($15 < n \leq 120$) และใหญ่ ($n > 120$) ยกเว้นสถิติ AV ที่จะทำการศึกษาเฉพาะตัวอย่างขนาดกลาง โดยกำหนดให้สำหรับแต่ละกลุ่มประชากรที่เปรียบเทียบมีขนาดตัวอย่างเท่ากัน เมื่อกำหนดระดับความแตกต่างของความแปรปรวน (ϕ) 3 ระดับคือ น้อย ($0 < \phi < 1.5$) ปานกลาง ($1.5 \leq \phi < 3.0$) และมาก ($\phi \geq 3.0$) ที่ระดับนัยสำคัญ (α) 0.01 และ 0.05 จำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ทำซ้ำ 1000 ครั้งในแต่ละสถานการณ์ ผลการศึกษาพบว่า เมื่อตัวอย่างขนาดเล็ก สถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีประชากร 3 และ 4 กลุ่ม ทุกระดับนัยสำคัญ เมื่อตัวอย่างขนาดกลาง สถิติ BF และสถิติ AV สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีประชากร 3 และ 4 กลุ่ม ทุกระดับนัยสำคัญ ส่วนสถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้กรณีประชากร 4 กลุ่ม ที่ระดับนัยสำคัญ 0.01 เมื่อตัวอย่างขนาดใหญ่ สถิติ BF สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีประชากร 3 และ 4 กลุ่ม ที่ทุกระดับนัยสำคัญ ส่วนสถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้กรณีประชากร 3 กลุ่ม ที่ระดับนัยสำคัญ 0.01 ส่วนเมื่อกรณีประชากร 4 กลุ่ม สถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ ทุกระดับนัยสำคัญ และเมื่อเปรียบเทียบกำลังการทดสอบของตัวสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ พบว่า เมื่อตัวอย่างขนาดเล็ก สถิติ GT ให้ค่ากำลังการทดสอบต่ำ เมื่อตัวอย่างขนาดกลาง สถิติ AV ให้ค่ากำลังการทดสอบสูงกว่าสถิติ BF ที่ทุกระดับของ ϕ ยกเว้นกรณีประชากร 4 กลุ่ม และระดับนัยสำคัญ 0.01 สถิติ GT จะให้ค่ากำลังการทดสอบสูงสุดในทุกระดับของ ϕ เมื่อตัวอย่างขนาดใหญ่ สถิติ BF และสถิติ GT ให้ค่ากำลังการทดสอบสูงใกล้เคียงกัน และเมื่อค่าระดับนัยสำคัญ จำนวนกลุ่มประชากร และระดับความแตกต่างของความแปรปรวนเพิ่มขึ้น สถิติทั้ง 3 วิธี มีแนวโน้มจะให้ค่ากำลังการทดสอบเพิ่มขึ้น

คำสำคัญ: ความเท่ากันของความแปรปรวน สถิติบราวน์ – ฟอร์สตี สถิติจี

* บทความนี้เป็นส่วนหนึ่งของวิทยานิพนธ์ระดับปริญญาโทบัณฑิต เรื่อง การเปรียบเทียบประสิทธิภาพของสถิติสำหรับทดสอบความเท่ากันของความแปรปรวน

** นักศึกษาหลักสูตรวิทยาศาสตรมหาบัณฑิต ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ E-mail : dumtamin@hotmail.com โดยมี ผศ.ดร.บุญอ้อม โหมที เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก และมี รศ.ดร.อภิญา หิรัญวงษ์ เป็นอาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม

ABSTRACT

The objectives of this research are to study and compare power of the test for the three test statistics of homogeneity of variance of 3 and 4 populations which are Brown Forsythe's test (BF), Analysis of means test for variances (AV) and Gini's test (GT). The probability of type I error and power of the test statistics were considered. The normality distributed population with mean equal to zero, small ($n \leq 15$), medium ($15 < n \leq 120$) and large ($n > 120$) sample sizes are studied, except for AV, which are studied, only for medium sample size. Each compared populations had equal sample sizes. The three difference levels of variance (ϕ) : low ($0 < \phi < 1.5$), medium ($1.5 \leq \phi < 3.0$) and high ($\phi \geq 3.0$) for significant levels (α) 0.01 and 0.05 are studied. The data were generated by the Monte Carlo technique repeated 1,000 times for each situation. The results showed that, for small sample size, GT can control the probability of type I error in every cases. For medium sample size, both BF and AV can control the probability of type I error in every cases, but GT can control the probability of type I error for 4 populations groups at 0.01 significant level. For large sample size, BF can control the probability of type I error in every cases, GT can control the probability of type I error for 3 populations groups at 0.01 significant level, however for 4 populations groups GT can control the probability of type I error every significant levels. Consequently, only the test statistics that can control the probability of type I error were compared power of the test. In case of small sample size, GT had low power of the test. When medium sample size, AV had power of the test more than BF for every difference levels of variance, except for 4 populations groups at 0.01 significant level, GT had maximum power of the test for every difference levels of variance. When large sample size, both BF and GT had high power of the test very close. Moreover, when the difference of population variances increases, the power of the three test statistics increase. Also, when the significant levels, the number of population groups and the difference levels of variance increases, power of the test for the three tests tends to increase.

Keywords: homogeneity of variances, Brown Forsythe's test, Gini's test, analysis of means test

บทนำ

การอนุมานเชิงสถิติเกี่ยวกับประชากร โดยเฉพาะการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของประชากร ซึ่งได้แก่ การเปรียบเทียบความแตกต่างของค่าเฉลี่ยประชากร 2 กลุ่ม อาจจะใช้สถิติทดสอบแบบใช้พารามิเตอร์ (Parametric Testing) ได้แก่ สถิติทดสอบ Z สถิติทดสอบ t หรือ สถิติทดสอบแบบไม่ใช้พารามิเตอร์ (Nonparametric Testing) เช่น การทดสอบแมน-วิทนี (Mann-Witney Test) ส่วนการเปรียบเทียบความแตกต่างของค่าเฉลี่ยประชากรมากกว่า 2 กลุ่ม สถิติทดสอบแบบใช้พารามิเตอร์ที่นิยม คือ สถิติ

ทดสอบ F ภายใต้การวิเคราะห์ความแปรปรวน ซึ่งจำเป็นต้องตรวจสอบข้อกำหนดเบื้องต้นเกี่ยวกับความเท่ากันของความแปรปรวนของประชากรแต่ละกลุ่ม เนื่องจากหากมีการฝ่าฝืนข้อกำหนดเบื้องต้นเกี่ยวกับความเท่ากันของความแปรปรวนจะทำให้เกิดความคลาดเคลื่อนในการสรุปผล และมีผลต่อความเชื่อถือได้ในการอ้างอิงไปยังประชากร

นับถึงปัจจุบัน ได้มีนักสถิติหลายท่านได้คิดค้นและพัฒนาตัวสถิติทดสอบความเท่ากันของความแปรปรวนของประชากรมากกว่า 2 กลุ่ม ทั้งที่เป็นสถิติทดสอบแบบที่ใช้พารามิเตอร์ และแบบที่ไม่ใช้พารามิเตอร์ สถิติทดสอบที่ใช้พารามิเตอร์ เช่น สถิติทดสอบบาร์ตเลต (Bartlett's test) สถิติทดสอบคอคเครน (Cochran's test) สถิติทดสอบบราวน์-ฟอร์ลิต (Brown-Forsythe's test) เป็นต้น และสถิติทดสอบที่ไม่ใช้พารามิเตอร์ เช่น สถิติทดสอบ ANOMV (Analysis of means test for variances) สถิติทดสอบจีนิ (Gini's test) สถิติทดสอบคล็อทซ์ (Klotz's test) และสถิติทดสอบมูด (Mood's test) เป็นต้น

ในการวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาและเปรียบเทียบประสิทธิภาพของตัวสถิติสำหรับทดสอบความเท่ากันของความแปรปรวนของประชากร 3 และ 4 กลุ่ม จำนวน 3 วิธี คือ สถิติบราวน์-ฟอร์ลิต ซึ่งเป็นสถิติแบบใช้พารามิเตอร์ สถิติ ANOMV และสถิติจีนิ ซึ่งเป็นสถิติแบบไม่ใช้พารามิเตอร์ โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และเปรียบเทียบกำลังการทดสอบของตัวสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

วิธีการวิจัย

1. สถิติทดสอบ 3 วิธีที่ศึกษา คือ

1.1 สถิติทดสอบบราวน์ – ฟอร์ลิต (Brown Forsythe's test : BF) ซึ่งผู้นำเสนอคือ Brown, M.B. และForsythe, A.B. (1974) คำนวณดังนี้

$$BF = \frac{\sum_{i=1}^k n_i (\bar{z}_i - \bar{z}_{..})^2 / (k-1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} (z_{ij} - \bar{z}_i)^2 / \sum_{i=1}^k (n_i - 1)}$$

เมื่อ $z_{ij} = |x_{ij} - x_i^{med}|$ โดย x_i^{med} เป็นค่ามัธยฐานของกลุ่มตัวอย่างที่ i

$$\bar{z}_i = \sum_{j=1}^{n_i} z_{ij} / n_i, \quad \bar{z}_{..} = \sum_{i=1}^k \sum_{j=1}^{n_i} z_{ij} / \sum_{i=1}^k n_i$$

k คือ จำนวนกลุ่มตัวอย่าง และ n_i คือ ขนาดตัวอย่างกลุ่มที่ i

โดยสถิติทดสอบบราวน์-ฟอร์ลิต มีการแจกแจงแบบ F และมีองศาอิสระ df_1 และ df_2 เท่ากับ

$k-1$ และ $\sum_{i=1}^k (n_i - 1)$ ตามลำดับ

เกณฑ์ในการตัดสินใจ จะปฏิเสธสมมติฐานหลัก เมื่อสถิติทดสอบบราวน์-ฟอร์ลิตี มีค่ามากกว่าค่าวิกฤติจากตารางการแจกแจง $F_{\alpha, k-1, \sum_{i=1}^k (n_i-1)}$

1.2 สถิติทดสอบ *ANOMV* (Analysis of means test for variances : AV) ซึ่งผู้นำเสนอคือ Wludyka, Peter S และ Nelson, Peter R. (1997) คำนวณดังนี้

คำนวณค่าความแปรปรวนของตัวอย่างกลุ่มที่ i ; $i = 1, 2, \dots, k$ และค่าเฉลี่ยความ

แปรปรวน k กลุ่ม ($\bar{S}^2$) โดย $S_i^2 = \frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}{n_i - 1}$ เมื่อ $\bar{x}_i$ คือ ค่าเฉลี่ยของกลุ่มตัวอย่างที่ i

$$\text{และ } \bar{S}^2 = \frac{\sum_{i=1}^k S_i^2}{k}$$

เกณฑ์ในการตัดสินใจ จะปฏิเสธสมมติฐานหลัก เมื่อค่าของสถิติทดสอบ S_i^2 อย่างน้อย 1 ค่า อยู่นอกขอบเขตบน (Upper decision line : UDL) หรือขอบเขตล่าง (Lower decision line : LDL) โดยค่า $U_{\alpha, k, \nu}$ และ $L_{\alpha, k, \nu}$ เป็นค่าวิกฤติจากตารางของ Wludyka และ Nelson (1997) ที่องศาอิสระ k กับ $\nu = n_i - 1$ และคำนวณ UDL และ LDL ดังนี้ $UDL = U_{\alpha, k, \nu} k \bar{S}^2$, $LDL = L_{\alpha, k, \nu} k \bar{S}^2$

1.3 สถิติทดสอบจีนิ (Gini's test : GT) ผู้นำเสนอคือ B.R. Bhat, M.N. Badade และ K. Aruna Rao (2002) คำนวณดังนี้

คำนวณค่าความต่างเฉลี่ยของจีนิ (Gini's mean difference : T_i) โดยเรียงลำดับข้อมูลจากมากไปน้อยสำหรับแต่ละกลุ่มที่ i ; $i = 1, 2, \dots, k$

$$T_i = \frac{\sqrt{\pi}}{(n_i(n_i-1))} \sum_{j < l}^{n_i} (x_{i(j)} - x_{i(l)}) , \quad i = 1, 2, 3, \dots, k , \quad j = 1, 2, 3, \dots, n_i , \quad l = 1, 2, 3, \dots, n_i$$

เมื่อ $x_{i(j)}$ คือ ค่าสังเกตกลุ่มที่ i ลำดับที่ j
 $x_{i(l)}$ คือ ค่าสังเกตกลุ่มที่ i ลำดับที่ l
 n_i คือ ขนาดตัวอย่างกลุ่มที่ i

และ สถิติ GT(T_G) คำนวณได้จาก

$$T_G = \frac{E(T_w)T_g}{T_w E(T_g)} \quad \text{หรือ} \quad T_G = \frac{T_g}{T_w} \frac{2E(T_w)}{\sum_{i=1}^k D_i \left(1 - \left(\frac{n_i}{N}\right)\right)}$$

โดยที่

$$T_w = \frac{1}{N} \sum_{i=1}^k n_i T_i, \quad N = \sum_{i=1}^k n_i, \quad T_g = \sum_{i=1}^k n_i \ln T_w - \sum_{i=1}^k n_i \ln T_i$$

$$D_i = \frac{1}{n_i - 1} \left[\frac{\pi(n_i + 1)}{3} + 2\sqrt{3}(n_i - 2) - 2(2n_i - 3) \right]$$

และ $E(T_w) = \frac{\sum_{i=1}^k n_i \sigma_i}{N}$, $Var(T_w) = \frac{\sum_{i=1}^k n_i D_i \sigma_i^2}{N^2}$, $E(T_g) = \frac{1}{2} \sum_{i=1}^k D_i \left(1 - \frac{n_i}{N} \right)$

$$Var(T_g) = \frac{1}{2N^2} \left[N^2 \sum_{i=1}^k D_i^2 + \left(\sum_{i=1}^k n_i D_i \right)^2 - 2N \sum_{i=1}^k n_i D_i^2 \right]$$

เมื่อ σ_i^2 คือ ค่าความแปรปรวนของประชากรกลุ่มที่ i

σ_i คือ ค่าเบี่ยงเบนมาตรฐานของประชากรกลุ่มที่ i

n_i คือ ขนาดตัวอย่างกลุ่มที่ i

n คือ ขนาดตัวอย่างทั้งหมด

สถิติทดสอบจีนี้ มีการแจกแจงแบบ F_{ρ_1, ρ_2} โดยมี ρ_1 และ ρ_2 เป็นองศาอิสระของเศษและส่วน ซึ่งคำนวณดังนี้

$$\rho_1 = \frac{2[E(T_g)]^2}{Var(T_g)}, \quad \rho_2 = \frac{2[E(T_w)]^2}{Var(T_w)}$$

เกณฑ์ในการตัดสินใจ จะปฏิเสธสมมติฐานหลัก ที่ระดับนัยสำคัญ α เมื่อค่า $T_G > F_{\alpha, \rho_1, \rho_2}$ โดยที่ $F_{\alpha, \rho_1, \rho_2}$ ได้จากการเปิดตารางการแจกแจง F

2. การกำหนดระดับความแตกต่างของความแปรปรวน

งานวิจัยนี้ใช้หลักการกำหนดระดับความแปรปรวน (ϕ) ซึ่งเสนอโดย Games, P.A. Winkler, H.B. และ Probert, D.A (1972) ดังนี้

ค่า ϕ อยู่ในช่วง $0 < \phi < 1.5$ เรียกว่า ความแปรปรวนของประชากรแตกต่างกัน้อย

ค่า ϕ อยู่ในช่วง $1.5 \leq \phi < 3.0$ เรียกว่า ความแปรปรวนของประชากรแตกต่างกันปานกลาง

ค่า $\phi \geq 3.0$ เรียกว่า ความแปรปรวนของประชากรแตกต่างมาก

3. เกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1

เสาวลักษณ์ (2540) กล่าวว่า การทดสอบสมมติฐานโดยสถิติ Z นั้น เกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และจำนวนการทดลองเท่ากับ 1,000 ครั้ง แบบทดสอบจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ถ้าค่า Z ตกอยู่ในช่วง $[-1.96 \leq Z \leq 1.96]$ ซึ่งสามารถสรุปผลได้ 2 กรณีคือ

กรณีที่ 1 กำหนดระดับนัยสำคัญของการทดลองเท่ากับ 0.01 ดังนั้น ตัวสถิติที่ศึกษาจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง หรือสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก เมื่อสมมติฐานหลักเป็นจริง กับจำนวนครั้งที่ทำการทดลอง ตกอยู่ในช่วง $[0.00383, 0.01617]$

กรณีที่ 2 กำหนดระดับนัยสำคัญของการทดลองเท่ากับ 0.05 ดังนั้น ตัวสถิติที่ศึกษาจะสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ที่ได้จากการทดลอง หรือสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก เมื่อสมมติฐานหลักเป็นจริง กับจำนวนครั้งที่ทำการทดลอง ตกอยู่ในช่วง $[0.0365, 0.0635]$

4. การเปรียบเทียบกำลังการทดสอบของตัวสถิติ

ในที่นี้จะเปรียบเทียบกำลังการทดสอบสำหรับตัวสถิติที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ โดยตัวสถิติใดให้กำลังการทดสอบสูงที่สุด ตัวสถิตินั้นจะเป็นสถิติที่มีประสิทธิภาพสูงที่สุด และสามารถคำนวณความน่าจะเป็นของกำลังการทดสอบ เท่ากับสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานหลัก เมื่อสมมติฐานหลักไม่เป็นจริง กับจำนวนครั้งที่ทำการทดลอง

ผลการศึกษา

ตารางที่ 1 ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของสถิติ BF AV และ GT ที่ระดับนัยสำคัญ 0.01

จำนวน กลุ่มประชากร	ตัวอย่างขนาดเล็ก			ตัวอย่างขนาดกลาง			ตัวอย่างขนาดใหญ่		
	BF	AV	GT	BF	AV	GT	BF	AV	GT
3 กลุ่ม	0	•	0.011*	0.004*	0.012*	0.027	0.004*	•	0.013*
4 กลุ่ม	0	•	0.011*	0.006*	0.007*	0.010*	0.009*	•	0.011*

หมายเหตุ * หมายถึง สถิติทดสอบสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

• หมายถึง ไม่ได้ทำการศึกษาเนื่องจากข้อจำกัดของค่าวิกฤตที่ใช้ในการเปรียบเทียบ

จากตารางที่ 1 ที่ระดับนัยสำคัญ 0.01 เมื่อตัวอย่างขนาดเล็ก สถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ส่วนสถิติ BF ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อตัวอย่างขนาดกลาง สถิติ BF และสถิติ AV สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ส่วนสถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้กรณีจำนวนประชากร 4 กลุ่ม เมื่อตัวอย่างขนาดใหญ่ สถิติ BF และสถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม

ตารางที่ 2 ค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ของสถิติ BF AV และ GT ที่ระดับนัยสำคัญ 0.05

จำนวน กลุ่มประชากร	ตัวอย่างขนาดเล็ก			ตัวอย่างขนาดกลาง			ตัวอย่างขนาดใหญ่		
	BF	AV	GT	BF	AV	GT	BF	AV	GT
3 กลุ่ม	0	•	0.060*	0.037*	0.057*	0.075	0.037*	•	0.068
4 กลุ่ม	0	•	0.052*	0.037*	0.046*	0.071	0.054*	•	0.058*

หมายเหตุ * หมายถึง สถิติทดสอบสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้

- หมายถึง ไม่ได้ทำการศึกษาเนื่องจากข้อจำกัดของค่าวิกฤตที่ใช้ในการเปรียบเทียบ

จากตารางที่ 2 ที่ระดับนัยสำคัญ 0.05 เมื่อตัวอย่างขนาดเล็ก สถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ส่วนสถิติ BF ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อตัวอย่างขนาดกลาง สถิติ BF และสถิติ AV สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ส่วนสถิติ GT ไม่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ เมื่อตัวอย่างขนาดใหญ่ สถิติ BF สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีจำนวนประชากร 3 และ 4 กลุ่ม ส่วนสถิติ GT สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้กรณีจำนวนประชากร 4 กลุ่ม

ตารางที่ 3 ค่ากำลังการทดสอบของสถิติ BF AV และ GT ที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ กรณีประชากร 3 กลุ่ม ที่ระดับนัยสำคัญ 0.01

ระดับความแตกต่าง ของความแปรปรวน(ϕ)	ตัวอย่างขนาดเล็ก	ตัวอย่างขนาดกลาง		ตัวอย่างขนาดใหญ่	
	GT	BF	AV	BF	GT
แตกต่างกัน้อย ($0 < \phi < 1.5$)	0.009	0.118	0.233	0.810	0.947
แตกต่างปานกลาง ($1.5 \leq \phi < 3.0$)	0.018	0.502	0.798	1	1
แตกต่างมาก ($\phi \geq 3.0$)	0.028	0.847	0.983	1	1

จากตารางที่ 3 สำหรับกรณีประชากร 3 กลุ่ม ที่ระดับนัยสำคัญ 0.01 เมื่อตัวอย่างขนาดเล็ก สถิติ GT ให้ค่ากำลังการทดสอบต่ำ เมื่อตัวอย่างขนาดกลาง สถิติ AV ให้ค่ากำลังการทดสอบสูงกว่าสถิติ BF ที่ทุกระดับความแตกต่างของความแปรปรวน เมื่อตัวอย่างขนาดใหญ่ สถิติ BF และสถิติ GT ให้ค่ากำลังการทดสอบสูง และเมื่อระดับความแตกต่างของความแปรปรวนเพิ่มขึ้น สถิติ GT สถิติ AV และสถิติ BF มีแนวโน้มให้ค่ากำลังการทดสอบเพิ่มขึ้นด้วย

ตารางที่ 4 ค่ากำลังการทดสอบของสถิติ BF AV และ GT ที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ กรณีประชากร 4 กลุ่ม ที่ระดับนัยสำคัญ 0.01

ระดับความแตกต่าง ของความแปรปรวน(ϕ)	ตัวอย่างขนาดเล็ก	ตัวอย่างขนาดกลาง			ตัวอย่างขนาดใหญ่	
	GT	BF	AV	GT	BF	GT
แตกต่างกัน้อย($0 < \phi < 1.5$)	0.020	0.262	0.409	0.581	0.966	0.992
แตกต่างปานกลาง($1.5 \leq \phi < 3.0$)	0.020	0.694	0.919	0.969	1	1
แตกต่างมาก($\phi \geq 3.0$)	0.018	0.875	0.989	0.998	1	1

จากตารางที่ 4 สำหรับกรณีประชากร 4 กลุ่ม ที่ระดับนัยสำคัญ 0.01 เมื่อตัวอย่างขนาดเล็ก สถิติ GT ให้ค่ากำลังการทดสอบต่ำ เมื่อตัวอย่างขนาดกลาง สถิติ GT ให้ค่ากำลังการทดสอบสูงกว่าสถิติ BF และสถิติ AV ที่ทุกระดับความแตกต่างของความแปรปรวน เมื่อตัวอย่างขนาดใหญ่ สถิติ BF และสถิติ GT ให้ค่ากำลังการทดสอบสูง และเมื่อระดับความแตกต่างของความแปรปรวนเพิ่มขึ้น สถิติ GT สถิติ AV และสถิติ BF มีแนวโน้มให้ค่ากำลังการทดสอบเพิ่มขึ้นด้วย

ตารางที่ 5 ค่ากำลังการทดสอบของสถิติ BF AV และ GT ที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ กรณีประชากร 3 กลุ่ม ที่ระดับนัยสำคัญ 0.05

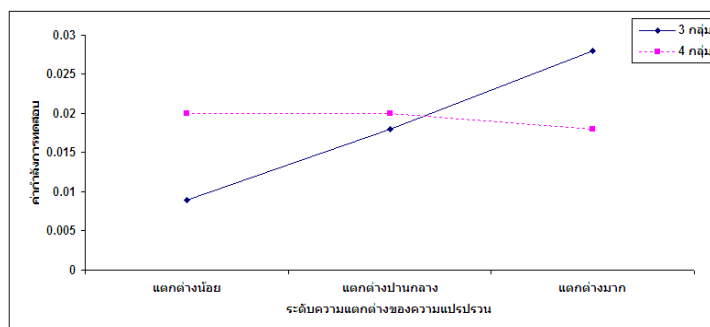
ระดับความแตกต่าง ของความแปรปรวน(ϕ)	ตัวอย่างขนาดเล็ก	ตัวอย่างขนาดกลาง		ตัวอย่างขนาดใหญ่
	GT	BF	AV	BF
แตกต่างกัน้อย($0 < \phi < 1.5$)	0.059	0.387	0.526	0.934
แตกต่างปานกลาง($1.5 \leq \phi < 3.0$)	0.106	0.812	0.940	1
แตกต่างมาก($\phi \geq 3.0$)	0.136	0.982	0.999	1

จากตารางที่ 5 สำหรับกรณีประชากร 3 กลุ่ม ที่ระดับนัยสำคัญ 0.05 เมื่อตัวอย่างขนาดเล็ก สถิติ GT ให้ค่ากำลังการทดสอบต่ำ เมื่อตัวอย่างขนาดกลาง สถิติ AV ให้ค่ากำลังการทดสอบสูงกว่าสถิติ BF ที่ทุกระดับความแตกต่างของความแปรปรวน เมื่อตัวอย่างขนาดใหญ่ สถิติ BF ให้ค่ากำลังการทดสอบสูง และเมื่อระดับความแตกต่างของความแปรปรวนเพิ่มขึ้น สถิติ GT สถิติ AV และสถิติ BF มีแนวโน้มให้ค่ากำลังการทดสอบเพิ่มขึ้นด้วย

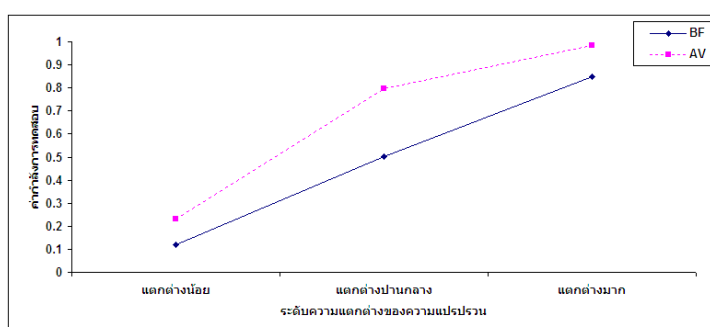
ตารางที่ 6 ค่ากำลังการทดสอบของสถิติ BF AV และ GT ที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 ได้ กรณีประชากร 4 กลุ่ม ที่ระดับนัยสำคัญ 0.05

ระดับความแตกต่าง ของความแปรปรวน(ϕ)	ตัวอย่างขนาดเล็ก	ตัวอย่างขนาดกลาง		ตัวอย่างขนาดใหญ่	
	GT	BF	AV	BF	GT
แตกต่างกัน้อย($0 < \phi < 1.5$)	0.093	0.552	0.668	0.998	1
แตกต่างปานกลาง($1.5 \leq \phi < 3.0$)	0.124	0.926	0.989	1	1
แตกต่างมาก($\phi \geq 3.0$)	0.155	0.982	0.999	1	1

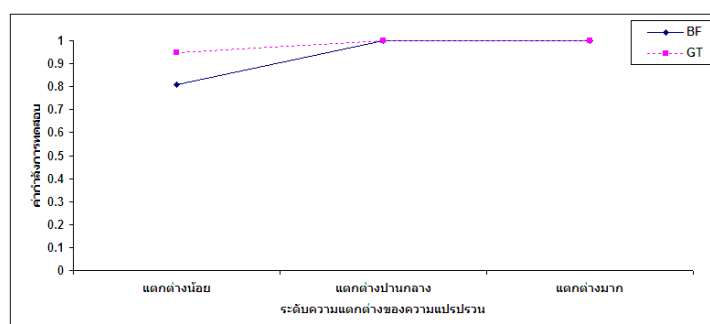
จากตารางที่ 6 สำหรับกรณีประชากร 4 กลุ่ม ที่ระดับนัยสำคัญ 0.05 เมื่อตัวอย่างขนาดเล็ก สถิติ GT ให้ค่ากำลังการทดสอบต่ำ เมื่อตัวอย่างขนาดกลาง สถิติ AV ให้ค่ากำลังการทดสอบสูงกว่าสถิติ BF ที่ทุกระดับความแตกต่างของความแปรปรวน เมื่อตัวอย่างขนาดใหญ่ สถิติ BF และสถิติ GT ให้ค่ากำลังการทดสอบสูง และเมื่อระดับความแตกต่างของความแปรปรวนเพิ่มขึ้น สถิติ GT สถิติ AV และสถิติ BF มีแนวโน้มให้ค่ากำลังการทดสอบเพิ่มขึ้นด้วย



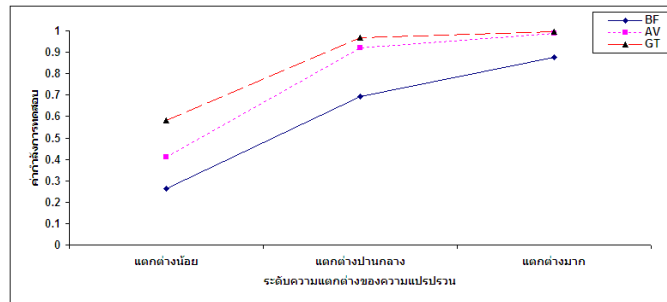
ภาพที่ 1 กำลังการทดสอบของสถิติ GT กรณีประชากร 3 และ 4 กลุ่ม
เมื่อตัวอย่างขนาดเล็ก $n_1 = 3, n_2 = 3, n_3 = 3, n_4 = 3$ ที่ระดับนัยสำคัญ 0.01



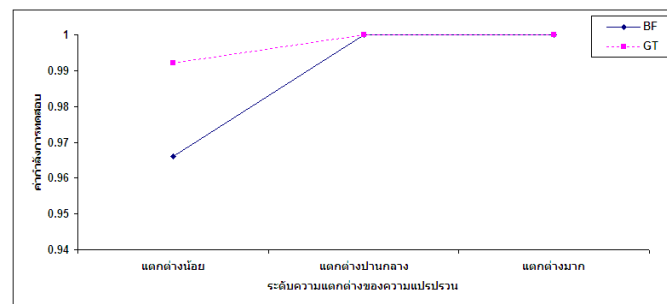
ภาพที่ 2 กำลังการทดสอบของสถิติ BF และสถิติ AV กรณีประชากร 3 กลุ่ม
เมื่อตัวอย่างขนาดกลาง $n_1 = 20, n_2 = 20, n_3 = 20$ ที่ระดับนัยสำคัญ 0.01



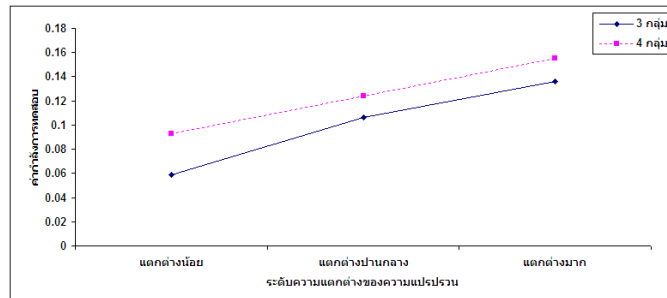
ภาพที่ 3 กำลังการทดสอบของสถิติ BF และสถิติ GT กรณีประชากร 3 กลุ่ม
เมื่อตัวอย่างขนาดใหญ่ $n_1 = 60, n_2 = 60, n_3 = 60$ ที่ระดับนัยสำคัญ 0.01



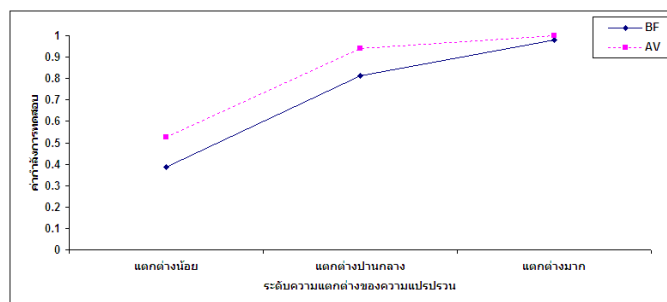
ภาพที่ 4 กำลังการทดสอบของสถิติ BF สถิติ AV และสถิติ GT กรณีประชากร 4 กลุ่ม เมื่อตัวอย่างขนาดกลาง $n_1 = 20, n_2 = 20, n_3 = 20, n_4 = 20$ ที่ระดับนัยสำคัญ 0.01



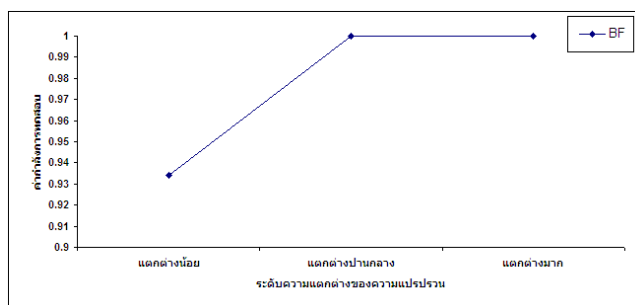
ภาพที่ 5 กำลังการทดสอบของสถิติ BF และสถิติ GT กรณีประชากร 4 กลุ่ม เมื่อตัวอย่างขนาดใหญ่ $n_1 = 60, n_2 = 60, n_3 = 60, n_4 = 60$ ที่ระดับนัยสำคัญ 0.01



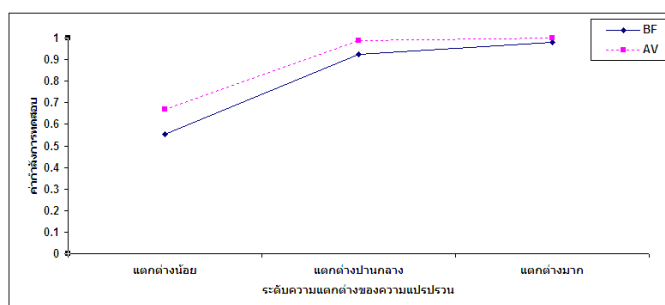
ภาพที่ 6 กำลังการทดสอบของสถิติ GT กรณีประชากร 3 และ 4 กลุ่ม เมื่อตัวอย่างขนาดเล็ก $n_1 = 3, n_2 = 3, n_3 = 3, n_4 = 3$ ที่ระดับนัยสำคัญ 0.05



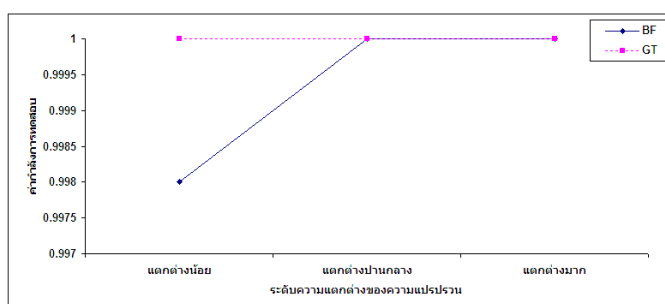
ภาพที่ 7 กำลังการทดสอบของสถิติ BF และสถิติ AV กรณีประชากร 3 กลุ่ม เมื่อตัวอย่างขนาดกลาง $n_1 = 20, n_2 = 20, n_3 = 20$ ที่ระดับนัยสำคัญ 0.05



ภาพที่ 8 กำลังการทดสอบของสถิติ BF กรณีประชากร 3 กลุ่ม
เมื่อตัวอย่างขนาดใหญ่ $n_1 = 60, n_2 = 60, n_3 = 60$ ที่ระดับนัยสำคัญ 0.05



ภาพที่ 9 กำลังการทดสอบของสถิติ BF และสถิติ AV กรณีประชากร 4 กลุ่ม
เมื่อตัวอย่างขนาดกลาง $n_1 = 20, n_2 = 20, n_3 = 20, n_4 = 20$ ที่ระดับนัยสำคัญ 0.05



ภาพที่ 10 กำลังการทดสอบของสถิติ BF และสถิติ GT กรณีประชากร 4 กลุ่ม
เมื่อตัวอย่างขนาดใหญ่ $n_1 = 60, n_2 = 60, n_3 = 60, n_4 = 60$ ที่ระดับนัยสำคัญ 0.05

สรุปผลการศึกษา

การศึกษานี้มีวัตถุประสงค์เพื่อศึกษา และเปรียบเทียบประสิทธิภาพของสถิติ 3 วิธี สำหรับทดสอบความเท่ากันของความแปรปรวนของประชากร 3 และ 4 กลุ่ม คือ สถิติบราวน์ - พอร์สตี สถิติ ANOMV และสถิติจีนี้ เมื่อกำหนดการแจกแจงของประชากรเป็นแบบปกติมีค่าเฉลี่ยเท่ากับ 0 สำหรับตัวอย่างขนาดเล็ก ขนาดกลาง และขนาดใหญ่ โดยที่ สถิติทดสอบ ANOMV จะทำการศึกษาเฉพาะกรณีตัวอย่างขนาดกลาง และกำหนดให้ประชากรทุกกลุ่มที่เปรียบเทียบมีขนาดตัวอย่างเท่ากัน กำหนดระดับความแตกต่างของแปรปรวน 3 ระดับ คือ แตกต่างน้อย ปานกลาง และมาก ที่ระดับนัยสำคัญ 0.01 และ 0.05 สามารถสรุปผลการศึกษาของตัวสถิติที่เหมาะสมแต่ละสถานการณ์ได้ ดังตารางที่ 7

ตารางที่ 7 ผลการศึกษาของตัวสถิติที่เหมาะสมแต่ละสถานการณ์

จำนวนกลุ่มประชากร	ขนาดตัวอย่าง(n)	ระดับนัยสำคัญ(α)	
		0.01	0.05
3 กลุ่ม	$n \leq 15$	GT	GT
	$15 < n \leq 120$	AV	AV
	$n > 120$	GT	BF
4 กลุ่ม	$n \leq 15$	GT	GT
	$15 < n \leq 120$	GT	AV
	$n > 120$	GT	GT

เอกสารอ้างอิง

ภาษาไทย

ปรีชา เครือโสม. “การจำลองแบบเพื่อเปรียบเทียบประสิทธิภาพของการวิเคราะห์ความแปรปรวนสำหรับข้อมูลจำแนกทางเดียวเทียบกับการวิเคราะห์ความแปรปรวนของข้อมูลจำแนกสองทางในแผนแบบการทดลองแบบบล็อกสุ่มสมบูรณ์” วารสารวิชาการ Vol. 4 September – December, 2554.

แหล่งที่มา: <http://www.ejournal.su.ac.th>

สุวิตา ศักดิ์ชัยนันท์. “การเปรียบเทียบวิธีการประมาณช่วงความเชื่อมั่นของผลต่างค่าเฉลี่ยสองประชากรในการแจกแจงแบบล็อกนอร์มอล” วารสารวิชาการ Vol. 4 September – December, 2554.

แหล่งที่มา: <http://www.ejournal.su.ac.th>

เสาวลักษณ์ สุนนางกูร. “การเปรียบเทียบกำลังการทดสอบของแบบทดสอบการแจกแจงปกติหลายตัวแปร.”

วิทยานิพนธ์ปริญญามหาบัณฑิต. สาขาสถิติประยุกต์. มหาวิทยาลัยธรรมศาสตร์, 2540.

ภาษาต่างประเทศ

Bhat, B.R., Badade, M.N. and K. Aruna Rao. “A New test for equality of variances for k normal population.” Communications in Statistics, 31(4): 567-587. 2002.

Brown, M.B. and A.B. Forsythe. “Robust tests for the equality of variances.” JASA, 69: 364-367. 1974.

Games, P.A., H.J. Winkle and D.A. Probert. “Robust tests for homogeneity of variance.” Educational and Psychological Measurment, 32: 887-909. 1972.

Wludyka, P.S. and P.R. Nelson. “An analysis of means type test for variances from normal populations.” Technometrics, 39: 274-285. 1997.

Wludyka, P.S. and P.R. Nelson. “Two non-parametric, analysis of means type tests for homogeneity of variances.” Journal of Applied Statistics, 26: 243-256. 1999.