การเปรียบเทียบความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน

A Comparison of Type I Error and Power of Statistics for Homogeneity of Variance Tests

ดวงพร หัชชะวณิช

Doungporn Hatchavanich

สาขาสถิติประยุกต์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยหอการค้าไทย

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบ Levene Modified-Levene Bartlett Box-Andersen Jackknife Z-variance O'Brien และ Overall-Woodward Modified Z-variance ซึ่งใช้ทดสอบความเป็นเอกพันธ์ของความแปรปรวน ในการศึกษาได้กำหนดรูปแบบการแจกแจงของประชากรและขนาดตัวอย่างในระดับต่างๆ โดยกำหนดให้มีอัตราส่วนของความแปรปรวนเท่ากับ 1:2:3:4 และ 1:2:3:4:5 จากการศึกษาโดยใช้วิธีมอนติคาร์โลพบว่า Levene ไม่ใช่สถิติทดสอบที่ดีที่สุด แต่ยังมีสถิติทดสอบอื่นที่ให้ผลที่ดีกว่าทั้งนี้ขึ้นกับการแจกแจงของข้อมูล ในกรณีที่ประชากรมีการแจกแจงปกติ Bartlett Box-Andersen และ Z-variance เป็นสถิติทดสอบที่ดีเพราะไม่ว่าจะใช้ขนาดตัวอย่างน้อยหรือมากก็ให้อำนาจการทดสอบมาก ในกรณีที่ประชากรมีการแจกแจงยูนิฟอร์ม ไคสแควร์ และการแจกแจงที่มีความเบ้ต่ำพบว่า Levene O’Brien และ Bartlett ให้อำนาจการทดสอบมากที่สุดตามลำดับ

คำสำคัญ : สถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวน ความคลาดเคลื่อนชนิดที่ 1 อำนาจการทดสอบ

Abstract

The main objective of this research was to compare the probability of type I error and power of the statistics for homogeneity of variance test. The tests considered Levene, Modified Levene, Bartlet, Box-Andersen, Jackknife, Z-variance, O'Brien and Overall-Woodward Modified Z-variance.

E-mail address: doungporn_hat@utcc.ac.th

The researcher compared the performance of these tests under the different types of distributions and sample sizes. The group variances followed the ratio of 1:2:3:4 and 1:2:3:4:5. Monte-Carlo methods were used in this study to calculate the empirical powers and the type I errors under the different settings. In conclusion, the Levene test was not the best choice. There were better tests that could be used, and some were more preferable depending on the distributional shape of data. For a normal distribution, it was found that Bartlett, Box-Andersen and Z-variance test were good choice for testing homogeneity of variances since they were not affected by sample sizes. For a uniform distribution, it was found that the Levene test was the best; for a low skew distribution, Bartlett test; for a chi-square distribution, O'Brien test.

Keywords : statistics for homogeneity of variance tests, type I error, power of the tests

1. บทนำ

สถิติทดสอบหลายๆวิธีมักจะมีข้อกำหนดที่จะต้องทดสอบสมมติฐานว่าตัวอย่างสุ่มมาจากประชากรที่มีค่าความแปรปรวนแตกต่างกันหรือไม่ เช่น ในการทดสอบสมมติฐานเพื่อทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร หลายกลุ่มด้วยวิธีวิเคราะห์ความแปรปรวน (ANOVA) นั้นมีข้อสมมติเบื้องต้นที่จำเป็นต้องตรวจสอบ 3 ประการได้แก่ ข้อมูลต้องสุ่มมาจากประชากรที่มีการแจกแจงปกติ มีค่าความแปรปรวนของประชากรไม่แตกต่างกัน และค่าความคลาดเคลื่อนเป็นอิสระกัน หากข้อสมมติเหล่านี้ไม่เป็นจริงจะทำให้ได้ข้อสรุปที่ผิดพลาดได้ ซึ่งสถิติทดสอบหลายวิธีที่สามารถใช้ในการทดสอบสมมติฐานว่าตัวอย่างสุ่มจากประชากรที่มีค่าความแปรปรวนแตกต่างกันหรือไม่

Bartlett เป็นสถิติทดสอบที่ใช้ในการทดสอบสมมติฐานว่าตัวอย่างสุ่มมาจากประชากรที่มีค่าความแปรปรวนแตกต่างกันหรือไม่ โดยมีข้อจำกัดว่าตัวอย่างต้องสุ่มมาจากประชากรที่มีการแจกแจงปกติ วิธีนี้มีความไวหากประชากรไม่ได้แจกแจงปกติ [1-2] ถึงกระนั้นก็มีผู้ที่นำสถิติทดสอบ Bartlett ไปใช้อย่างกว้างขวางแม้ว่าตัวอย่างจะไม่ได้ถูกสุ่มจากประชากรที่มีการแจกแจงปกติ

Levene [3] ได้เสนอสถิติทดสอบซึ่งใช้ทดสอบสมมติฐานว่าตัวอย่างสุ่มมาจากประชากรที่มีค่าความแปรปรวนที่เท่ากันหรือไม่ โดยใช้การแปลงข้อมูลให้อยู่ในรูปค่าความแตกต่างระหว่างค่าของข้อมูลกับค่าเฉลี่ยของตัวอย่างที่ถูกสุ่มมาจากประชากรแต่ละชุด Neuhauser [4] ได้เสนอว่าค่าสถิติทดสอบนี้มีความแกร่งเมื่อข้อมูลไม่ได้ถูกสุ่มมาจากประชากรที่มีการแจกแจงปกติ และให้อำนาจการทดสอบสูง Levene จึงเป็นที่นิยมใช้มากขึ้นในทางสถิติและในการประยุกต์ใช้ในด้านอื่นๆ

Lim และ Loh [5] ได้ศึกษาเปรียบเทียบการทดสอบความเป็นเอกพันธ์ของความแปรปรวนโดยใช้

ข้อมูลจากการจำลองให้มีตัวอย่างขนาดเล็กถึงขนาดปานกลาง จากการเปรียบเทียบความแกร่งและอำนาจการทดสอบของสถิติทดสอบ Levene กับ Bartlett ในกรณีที่มีการปรับค่าความโด่งและไม่ได้ปรับค่าความโด่ง Box-Andersen และJackknife ซึ่งประมาณค่าพารามิเตอร์โดยใช้ Bootstrap และไม่ใช้ Bootstrap พบว่า Levene ซึ่งประมาณค่าพารามิเตอร์ด้วยวิธี Bootstrap ให้อำนาจการทดสอบมากที่สุด

Katz Restori และLee [6] ได้ศึกษาเปรียบเทียบความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบ Levene Modified- Levene Z-variance Overall-Woodward modified Z-variance O'Brien Samiuddin sube-root และ Fmax ซึ่งพบว่า Levene Modified- Levene Z-variance Overall-Woodward modified Z-variance และ O'Brien ให้ผลดีกว่า Samiuddin sube-root และ Fmax และพบว่าไม่มีสถิติทดสอบตัวใดที่ให้ผลดีที่สุดในทุกกรณี

จึงเป็นที่น่าสนใจว่าสถิติทดสอบ Levene Modified-Levene Z-variance Overall-Woodward modified Z-variance และO'Brien จะให้ความคลาดเคลื่อนชนิดที่ 1 ซึ่งอยู่ในขอบเขตที่ยอมรับได้และมีอำนาจการทดสอบมากกว่า Bartlett Box-Andersen และ Jackknife หรือไม่ ในกรณีข้อมูลมีการแจกแจงที่เป็นไปตามข้อสมมติและไม่เป็นไปตามข้อสมมติ ซึ่งจะทำให้ได้ข้อเสนอแนะที่สามารถใช้เป็นแนวทางในการเลือกใช้สถิติทดสอบที่เหมาะสมในการทดสอบความเป็นเอกพันธ์ของความแปรปรวนของประชากร

2. วิธีดำเนินงานวิจัย

2.1 จำลองข้อมูลโดยใช้โปรแกรม R โดยประชากรมีการแจกแจงปกติ ที ยูนิฟอร์ม ไคสแควร์ เลขชี้กำลัง และการแจกแจงที่มีความเบ้ต่ำ (ค่าความเบ้ไม่เกิน 1) โดยกำหนดให้ประชากรมีสัดส่วนของค่าความแปรปรวน $\sigma_1^2:\sigma_2^2:\sigma_3^2:\sigma_4^2 = 1:2:3:4$ และ $\sigma_1^2:\sigma_2^2:\sigma_3^2:\sigma_4^2:\sigma_5^2 = 1:2:3:4:5$ สุ่มตัวอย่างจากประชากรโดยกำหนดให้มีขนาดตัวอย่างมีจำนวนเท่ากันและแตกต่างกัน เพื่อทดสอบสมมติฐาน $H_0:\sigma_1^2=\sigma_2^2=\ldots=\sigma_k^2$ $H_1:\sigma_i^2\neq\sigma_j^2$ อย่างน้อย 1 คู่ เมื่อ $i \neq j$ (i,j=1,2,…,k) โดยกำหนดระดับนัยสำคัญ 0.05 ทำซ้ำ 10,000 รอบ นับจำนวนครั้งที่ปฏิเสธ H_0 เมื่อ H_0 เป็นจริง เพื่อประมาณค่าความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 และนับจำนวนครั้งที่ปฏิเสธ H_0 เมื่อ H_0 เป็นไม่จริง เพื่อประมาณค่าอำนาจการทดสอบของสถิติทดสอบดังนี้

2.1.1 Levene

$$L_{SQ} = \frac{(n-k)\sum_{i=1}^{k} n_i\left(\bar{z}_{i.}-\bar{z}\right)^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_j}\left(z_{ij}-\bar{z}_{i.}\right)^2}$$

โดย n = จำนวนข้อมูลทั้งหมด n_i = ขนาดตัวอย่างของกลุ่ม i $z_{ij} = (x_{ij} - \bar{x}_i)^2$

$\bar{x}_i$ = ค่าเฉลี่ยของกลุ่ม i $\bar{z}_{i\bullet} = \dfrac{\sum_{j=1}^{n_i} z_{ij}}{n_i}$, $\bar{z} = \dfrac{\sum_{i=1}^{k} n_i \bar{z}_{i\bullet}}{n}$, k = จำนวนประชากร

ซึ่งสถิติทดสอบ L_{SQ} มีการแจกแจงใกล้เคียงกับการแจกแจงเอฟ [5] ซึ่งมีระดับขั้นความเสรีเท่ากับ k-1 และ n-k

2.1.2 Modified-Levene

$$L_{ABS} = \frac{(n-k)\sum_{i=1}^{k} n_i (\bar{z}_{i\bullet} - \bar{z})^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_j} (z_{ij} - \bar{z}_{i\bullet})^2}$$

$z_{ij} = |x_{ij} - \bar{x}_{i\bullet}|$, x_{ij} = ค่าข้อมูลดิบ n = จำนวนข้อมูลทั้งหมด n_i = ขนาดตัวอย่างของกลุ่ม i

$\bar{x}_i$ = ค่าเฉลี่ยกลุ่มที่ i $\bar{z}_{i\bullet} = \dfrac{\sum_{j=1}^{n_i} z_{ij}}{n_i}$; $\bar{z} = \dfrac{\sum_{i=1}^{k} n_i \bar{z}_{i\bullet}}{n}$; k = จำนวนประชากร

ซึ่งสถิติทดสอบ L_{ABS} มีการแจกแจงใกล้เคียงกับการแจกแจงเอฟ [6] ซึ่งมีระดับขั้นความเสรีเท่ากับ k-1 และ n-k

2.1.3 Bartlett

$$\text{Bartlett1} = B_1 = \frac{M}{C} \; ; \; M = (N-I)\log\left\{\sum_{i=1}^{k}(n_i - 1)s_i^2 / (N-I)^{-1}\right\} - \sum_{i=1}^{k}(n_i - 1)\log s_i^2$$

$$C = 1 + \left\{\left[\sum_{i=1}^{I}(n_i - 1)^{-1}\right] - (N-I)^{-1}\right\} / \{3(I-1)\}$$

ซึ่ง B_1 มีการแจกแจงไคสแควร์ (Lim และ Loh [5]) โดยจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า B_1 มีค่ามากกว่า $\chi^2_{1-\alpha}(I-1)$

ต่อมา Boos และ Brownie (1989) [5] ได้เสนอสถิติทดสอบ Bartlett 2 ซึ่งปรับปรุงจาก Bartlett 1 โดย

Bartlett 2 = $B_2 = dB_1$; $d = 2/\left(\hat{\beta}_2 - 1\right)$ โดย $\hat{\beta}_2 = \dfrac{N \sum_{i=1}^{I} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_{i.}\right)^4}{\left\{\sum_{i=1}^{I} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_{i.}\right)^2\right\}^2}$

เนื่องจาก $B_1 \rightarrow \frac{1}{2}\left(\beta_2 - 1\right)\chi^2_{I-1}$ โดย $\beta_2 = E\left(X - \mu\right)^4 / \sigma^4$ เป็นค่าความโด่ง ดังนั้นค่าวิกฤตของ B_2 เหมือนกับ B_1 [5]

2.1.4 Box-Andersen

เป็นสถิติทดสอบที่ปรับปรุงจาก Bartlett โดย

$$B_3 = dM \;;\; M = (N - I)\log\left\{\sum_{i=1}^{I} (n_i - 1)s_i^2 / (N - I)^{-1}\right\} - \sum_{i=1}^{I} (n_i - 1)\log s_i^2$$

B_3 มีการแจกแจงไคสแควร์ [5] โดยจะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า B_3 มีค่ามากกว่า $\chi^2_{1-\alpha}(I-1)$

2.1.5 Jackknife

$$J_1 = \frac{\sum_{i=1}^{I} n_i \left(\bar{\mu}_{i.} - \bar{\mu}_{..}\right)^2 / (I - 1)}{\sum_{i=1}^{I} \sum_{j=1}^{n_i} \left(\mu_{ij} - \bar{\mu}_{i.}\right)^2 / (N - I)}$$

โดย $\mu_{ij} = n_i \log s_i^2 - \left(n_i - 1\right)s^2_{i(j)}$ $\qquad s^2_{i(j)} = \dfrac{\sum_{i \neq j} \left(x_{ij} - \bar{x}_{i(j)}\right)^2}{(n_i - 2)}$

$\bar{x}_{i(j)} = \dfrac{\sum_{i \neq j} x_{ij}}{(n_i - 1)}$, $\bar{\mu}_{i.} = \dfrac{\sum_{j=1}^{n_i} u_{ij}}{n_i}$, $\bar{\mu}_{..} = \dfrac{\sum_{i-1}^{I} \sum_{j=1}^{n_i} \mu_{ij}}{N}$ สำหรับ $j = 1,2,3,\ldots,n_i$ $\quad i = 1,2,3,\ldots,I$

ซึ่งจะปฏิเสธ H_0 ถ้า J_1 มีค่ามากกว่า $F_{(1-\alpha)}(I - 1, N - I)$ [5]

2.1.6 Z-variance

$F = \dfrac{\sum_{i=1}^{k} z_i^2}{k - 1}$ โดย $z_i = \sqrt{\dfrac{c(n_i - 1)s_i^2}{MSE}} - \sqrt{c(n_i - 1) - \dfrac{c}{2}}$ $\quad c = 2 + \dfrac{1}{n_i}$

s_i^2 = ค่าประมาณที่ไม่เอนเอียงของความแปรปรวนในกลุ่มที่ i $\quad n_i$ = ขนาดตัวอย่างกลุ่มที่ i

MSE = ค่าความแปรปรวนของความคลาดเคลื่อนภาย ในกลุ่ม ซึ่ง z_i สามารถประมาณได้ด้วยการแจกแจงปกติที่มีค่าเฉลี่ยเท่ากับศูนย์ โดยจะปฏิเสธ H_0 ถ้า $F > F(\alpha, k-1, \infty)$ [7]

2.1.7 Overall-Woodward Modified Z-variance

$$c = 2.0\left(\frac{2.9 + 0.2/n_i}{k}\right)^{1.6(n_i - 1.8k + 14.7)/n_i}$$ โดย n_i = ขนาดตัวอย่างของกลุ่มที่ i

k = ค่าเฉลี่ยของความโด่งที่คำนวณจากตัวอย่างทุกกลุ่ม $k = \frac{\sum z_{ij}^4}{n_i - 2}$ โดย $z_{ij} = \frac{x_{ij} - \bar{x}_i}{\sqrt{\left(\frac{n_i - 1}{n_i}\right)s_i^2}}$ [7]

2.1.8 O'Brien

$$v_{ij} = \frac{(n_i - 1.5)n_i(y_{ij} - \bar{y}_i)^2 - 0.5s_i^2(n_i - 1)}{(n_i - 1)(n_i - 2)}$$ โดย y_{ij} เป็นค่าของข้อมูลดิบ $\bar{y}_i = \frac{\sum_{j=1}^{n_i} y_{ij}}{n_i}$

เป็นค่าเฉลี่ยของข้อมูลในกลุ่มที่ i $s_i^2 = \frac{\sum_{j=1}^{n_i}(y_{ij} - \bar{y})^2}{n_i - 1}$ ค่าเฉลี่ยของ v ในแต่ละกลุ่มคำนวณได้จาก

$\bar{v}_{i.} = \frac{\sum v_{ij}}{n_i} = s_i^2$ จากนั้นนำค่า v_{ij} ไปคำนวณค่า F ด้วยวิธีเดียวกับการวิเคราะห์ความแปรปรวน [8]

2.2 เปรียบเทียบประสิทธิภาพของสถิติทดสอบ

วัดประสิทธิภาพของสถิติทดสอบโดยพิจารณาจากค่าประมาณความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 และค่าประมาณของอำนาจการทดสอบ

ค่าประมาณความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 (Estimated probability of type I error) เป็นค่าประมาณความน่าจะเป็นของการเกิดความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักเป็นจริง คำนวณได้จาก $\frac{\text{จำนวนครั้งที่ได้ผลการทดสอบปฏิเสธ}H_0\text{ เมื่อ }H_0\text{เป็นจริง}}{\text{จำนวนครั้งที่ทำซ้ำ(10,000 ครั้ง)}}$

อำนาจการทดสอบ (Power of the test) เป็นความน่าจะเป็นจากการปฏิเสธสมมติฐานหลักเมื่อสมมติฐานทางเลือกเป็นจริง คำนวณได้จาก $\frac{\text{จำนวนครั้งที่จะปฏิเสธ }H_0\text{ เมื่อ }H_0\text{ ไม่จริง}}{\text{จำนวนครั้งที่ทำซ้ำ (10,000 ครั้ง)}}$

ในการสรุปผลการทดสอบว่าสถิติทดสอบที่นำมาศึกษาวิธีใดจะสามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ จะใช้เกณฑ์ของ Cochran [9] โดยกำหนดระดับนัยสำคัญ 0.05 ซึ่งหากค่าประมาณความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 จากการทดลองอยู่ในช่วง (0.04,0.06) จะสรุปว่าค่าสถิติทดสอบนั้นสามารถควบคุมความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ จากนั้นจึงเปรียบเทียบค่าประมาณอำนาจการทดสอบของสถิติทดสอบที่สามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้โดยอำนาจการทดสอบของสถิติทดสอบตัวใดที่มากที่สุดเป็นวิธีที่มีประสิทธิภาพมากที่สุด

3. ผลการวิจัย

ผลการศึกษาพบว่าสถิติทดสอบแต่ละวิธีสามารถควบคุมความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 ได้แตกต่างกันโดยขึ้นกับการแจกแจงของประชากรและขนาดตัวอย่าง ซึ่งสามารถสรุปได้ในตารางที่ 1 ดังนั้นจึงจำเป็นต้องเปรียบเทียบอำนาจการทดสอบเพื่อให้ได้ข้อสรุปที่สามารถใช้เป็นข้อเสนอแนะในการเลือกใช้สถิติทดสอบที่สามารถควบคุมความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 ได้และให้ค่าประมาณอำนาจการทดสอบมาก ดังแสดงในตารางที่ 2-ตารางที่ 8

ตารางที่ 1. สถิติทดสอบที่สามารถควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้

การแจกแจงของประชากร	**ประชากร 4 ชุด**	**ประชากร 5 ชุด**
ปกติ	Levene Bartlett1 Bartlett2 Z-variance Box-Andersen O'Brien	Levene Bartlett1 Bartlett2 Z-variance Box-Andersen O'Brien
ยูนิฟอร์ม	Levene Jacknife	Levene
มีความเบ้ต่ำ	Bartlett2 O'Brien	Bartlett2
ไคสแควร์	O'Brien	-
ที และ เลขชี้กำลัง	-	-

ตารางที่ 2. ค่าประมาณอำนาจการทดสอบของสถิติทดสอบ ในกรณีที่ประชากร 4 ชุดมีการแจกแจงปกติ

ขนาดตัวอย่าง	**ขนาดตัวอย่างเฉลี่ย**	**Levene**	**Bartlett1**	**Bartlett2**	**Box-Andersen**	**Z-variance**	**O'Brien**
ขนาดตัวอย่างไม่เท่ากัน							
10,15,20,25	17.5	0.7636	0.9948	0.9625	0.9669	0.9844	0.7142
35,40,45,52	43	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
35,50,65,80	57.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
30,65,90,150	83.75	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ขนาดตัวอย่างไม่เท่ากัน							
16,16,16,16	16	0.8982	0.9988	0.9774	0.9805	0.9980	0.8633
20,20,20,20	20	0.9698	0.9999	0.9964	0.9968	0.9999	0.9586
30,30,30,30	30	0.9996	1.0000	0.9999	0.9999	1.0000	0.9991
50,50,50,50	50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
60,60,60,60	60	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

ตารางที่ 3. ค่าประมาณอำนาจการทดสอบของสถิติทดสอบ ในกรณีที่ประชากร 5 ชุดมีการแจกแจงปกติ

ขนาดตัวอย่าง	**ขนาดตัวอย่างเฉลี่ย**	**Levene**	**Bartlett1**	**Bartlett2**	**Box-Andersen**	**Z-variance**	**O'Brien**
ขนาดตัวอย่างไม่เท่ากัน							
4,10,18,22,26	16	0.5719	0.9547	0.8623	0.8871	0.8725	0.5094
10,13,16,19,22	16	0.8446	0.9996	0.9861	0.9893	0.9983	0.7975
12,14,16,18,20	16	0.8929	0.9998	0.9908	0.9932	0.9997	0.8584
14,16,20,24,26	20	0.9612	1.0000	0.9984	0.9989	1.0000	0.9472
16,18,20,22,24	20	0.9753	1.0000	0.9983	0.9984	1.0000	0.9657
10,20,30,40,50	30	0.9860	1.0000	0.9999	0.9999	1.0000	0.9796
26,28,30,32,34	30	0.9993	1.0000	1.0000	1.0000	1.0000	0.9999
24,26,30,34,36	30	0.9996	1.0000	1.0000	1.0000	1.0000	0.9994
35,40,45,52,60	46.4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
30,65,90,100,150	87	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
ขนาดตัวอย่างเท่ากัน							
16,16,16,16,16	16	0.9511	1.0000	0.9941	0.9953	0.9999	0.9296
20,20,20,20,20	20	0.9906	1.0000	0.9993	0.9995	1.0000	0.9852
30,30,30,30,30	30	0.9997	1.0000	1.0000	1.0000	1.0000	0.9997
60,60,60,60,60	60	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

ตารางที่ 4. ค่าประมาณอำนาจการทดสอบของสถิติทดสอบกรณีประชากร 4 ชุดมีการแจกแจงยูนิฟอร์ม

ขนาดตัวอย่าง	ขนาดตัวอย่างเฉลี่ย	Levene	Jacknife
ขนาดตัวอย่างไม่เท่ากัน			
10,15,20,25	17.5	0.7212	0.3626
35,40,45,52	43	1.0000	0.9976
35,50,65,80	57.5	1.0000	0.9999
30,65,90,150	83.75	1.0000	1.0000
ขนาดตัวอย่างเท่ากัน			
16,16,16,16	16	0.8011	0.5258
20,20,20,20	20	0.9199	0.7229
30,30,30,30	30	0.9946	0.9622
50,50,50,50	50	1.0000	0.9998
60,60,60,60	60	1.0000	1.0000

ตารางที่ 5. อำนาจการทดสอบของสถิติทดสอบ กรณีประชากร 5 ชุดมีการแจกแจงยูนิฟอร์ม

ขนาดตัวอย่าง	ขนาดตัวอย่างเฉลี่ย	Levene
ขนาดตัวอย่างไม่เท่ากัน		
4,10,18,22,26	16	0.5450
10,13,16,19,22	16	0.8040
12,14,16,18,20	16	0.8590
14,16,20,24,26	20	0.9480
16,18,20,22,24	20	0.9591
10,20,30,40,50	30	0.9796
26,28,30,32,34	30	0.9991
24,26,30,34,36	30	0.9995
35,40,45,52,60	46.4	1.0000
30,65,90,100,150	87	1.0000
ขนาดตัวอย่างเท่ากัน		
16,16,16,16,16	16	0.9057
20,20,20,20,20	20	0.9746
30,30,30,30,30	30	0.9998
60,60,60,60,60	60	1.0000

ตารางที่ 6. ค่าประมาณอำนาจการทดสอบ กรณีประชากร 4 ชุดมีการแจกแจงที่มีความเบ้ต่ำ

ขนาดตัวอย่าง	**ขนาดตัวอย่างเฉลี่ย**	**Bartlett2**	**O'Brien**
ขนาดตัวอย่างไม่เท่ากัน			
10,15,20,25	17.5	0.4133	0.2047
35,40,45,52	43	0.9084	0.8331
35,50,65,80	57.5	0.9515	0.8895
30,65,90,150	83.75	0.9774	0.9268
ขนาดตัวอย่างเท่ากัน			
16,16,16,16	16	0.4759	0.3524
20,20,20,20	20	0.5735	0.4533
30,30,30,30	30	0.7834	0.6898
50,50,50,50	50	0.9558	0.9309
60,60,60,60	60	0.9831	0.9700

ตารางที่ 7. ค่าประมาณอำนาจการทดสอบ ในกรณีประชากร 5 ชุดมีการแจกแจงที่มีความเบ้ต่ำ

ขนาดตัวอย่าง	**ขนาดตัวอย่างเฉลี่ย**	**Bartlett2**
ขนาดตัวอย่างไม่เท่ากัน		
4,10,18,22,26	16	0.2806
10,13,16,19,22	16	0.4970
12,14,16,18,20	16	0.5367
14,16,20,24,26	20	0.6516
16,18,20,22,24	20	0.6835
10,20,30,40,50	30	0.6855
26,28,30,32,34	30	0.8902
24,26,30,34,36	30	0.8784
35,40,45,52,60	46.4	0.9809
30,65,90,100,150	87	0.9973
ขนาดตัวอย่างเท่ากัน		
16,16,16,16,16	16	0.5940
20,20,20,20,20	20	0.7248
30,30,30,30,30	30	0.9059
60,60,60,60,60	60	0.9976

ตารางที่ 8. ค่าประมาณอำนาจการทดสอบกรณีประชากร 4 ชุดมีการแจกแจงไคสแควร์

ขนาดตัวอย่าง	**ขนาดตัวอย่างเฉลี่ย**	**O'Brien**
ขนาดตัวอย่างไม่เท่ากัน		
10,15,20,25	17.5	0.1327
35,40,45,52	43	0.4577
35,50,65,80	57.5	0.5038
30,65,90,150	83.75	0.5269
ขนาดตัวอย่างเท่ากัน		
16,16,16,16	16	0.2362
20,20,20,20	20	0.2796
30,30,30,30	30	0.3875
50,50,50,50	50	0.5821
60,60,60,60	60	0.6626

สำหรับทุกการแจกแจงของประชากรเมื่อพิจารณาที่ขนาดตัวอย่างเฉลี่ยเท่ากันพบว่าในกรณีที่ใช้ขนาดตัวอย่างเท่ากันทุกวิธีจะมีอำนาจการทดสอบมากกว่ากรณีที่ใช้ขนาดตัวอย่างแตกต่างกัน เมื่อมีขนาดตัวอย่างเฉลี่ยเพิ่มขึ้นทุกวิธีจะมีค่าประมาณอำนาจการทดสอบเข้าใกล้ 1 ในกรณีประชากรมีการแจกแจงปกติ พบว่า Bartlett Box-Andersen และZ-variance เป็นสถิติทดสอบที่ดีเพราะไม่ว่าจะใช้ขนาดตัวอย่างมากหรือน้อยจะมีอำนาจการทดสอบมาก ในกรณีประชากรมีการแจกแจงยูนิฟอร์ม พบว่า Levene สามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ เพียงวิธีเดียว

ในกรณีประชากร 4 ชุด มีการแจกแจงไคสแควร์พบว่า O'Brien สามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้เพียงวิธีเดียวเท่านั้นโดยมีค่าประมาณอำนาจการทดสอบไม่มากแม้ว่าจะใช้ขนาดตัวอย่างเฉลี่ยเพิ่มขึ้น ส่วนในกรณีที่ประชากร 5 ชุดมีการแจกแจงไคสแควร์ ประชากร 4 ชุดและ 5 ชุด มีการแจกแจงที และการแจกแจงเลขชี้กำลัง ไม่มีวิธีใดที่สามารถควบคุมความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้

4. สรุปผลการวิจัย

ผลการเปรียบเทียบค่าประมาณอำนาจการทดสอบของสถิติทดสอบที่สามารถควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ พบว่าเมื่อประชากรมีการแจกแจงปกติ ขนาดตัวอย่างเฉลี่ยน้อยกว่า 20 Bartlett Box-Andersen และ Z-variance มีค่าประมาณอำนาจการทดสอบมากกว่า Levene และ O'Brien เมื่อขนาดตัวอย่างเฉลี่ยเพิ่มขึ้นจะมีอำนาจการทดสอบเพิ่มขึ้น สำหรับกรณีประชากรมีการแจกแจงยูนิฟอร์ม Levene มีอำนาจการทดสอบมากที่สุด เมื่อประชากร 4 ชุดมีการแจกแจงไคสแควร์ O'Brien มี

อำนาจการทดสอบไม่มาก และในกรณีที่ประชากรมีการแจกแจงที่มีความเบ้ต่ำ Bartlett ให้อำนาจการทดสอบมากที่สุด

ดังนั้นผลการศึกษานี้จึงเป็นประโยชน์สำหรับนักวิจัยในการเลือกใช้สถิติทดสอบความเป็นเอกพันธ์ของความแปรปรวนของประชากร ซึ่งหากทราบว่าประชากรมีการแจกแจงอย่างไรจะทำให้สามารถเลือกสถิติทดสอบที่สามารถควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 และให้อำนาจการทดสอบมากที่สุด

กิตติกรรมประกาศ

งานวิจัยนี้ได้รับทุนสนับสนุนการวิจัยจากมหาวิทยาลัยหอการค้าไทย

เอกสารอ้างอิง (References)

[1] Box, G.E.P., 1953. Non-normality and tests on variances. *Biometrika,* 40, 318–335.

[2] Manoukian, E.B., Maurais, J., and Ouimet, R., 1986. Exact critical values of Bartlett's Test of homogeneity of variances for unequal sample sizes for two populations and power of the test. *Metrika,* 33, 275–289.

[3] Levene, H., 1960. Robust Testes for Equality of Variances in Contributions to Probability and Statistics. California: Stanford University Press.

[4] Neuhauser, M., 2007. A comparative study of nonparametric two- sample tests after Levene's transformation. *Journal of Statistical Computation and Simulation* ,77, 517–526.

[5] Lim, T.S. and Loh, W.Y., 1996. [on-line]. Available at <http://www.sciencedirect.com/science/article/pii/0167947395000542 > [Accessed 20 December 2013]

[6] Katz, G.S., Restori, A.F. and Lee, H.B. 2010. [on-line]. Available at <http://www.thescipub.com/abstract/10.3844/jmssp.2010.359.366>[Accessed 20 December 2013]

[7] Overall, J.E. and Woodward, J.A., 1974. A simple test for homogeneity of variance in complex factorial design. *Psychometrika,* 39, 311-218.

[8] O'Brien, R.G., 1981. A Simple test for variance effects in experimental designs. *Psychol. Bull.,* 89, 570-574.

[9] Cochran, W.G., 1954. The combination of estimates from different experiments. *Biometrics,* 10, 101-129.