# การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิดสำหรับค่าเฉลี่ยประชากร ภายใต้โครงสร้างของสหสัมพันธ์แบบไม่เท่ากัน

**หทัยรัตน์ วงศ์ชัยสุวัฒน์ [1] และ กมล บุษบา [2]**
มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต คลองหลวง ปทุมธานี 12121

*รับเมื่อ 4 เมษายน 2550 ตอบรับเมื่อ 29 มิถุนายน2550*

## บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบวิธีการทดสอบตัวแปรพหุคูณแบบปิดโดยใช้วิธีแบบขั้นบันได สำหรับการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างประชากรสองกลุ่มจำนวน 5 วิธี คือ วิธีโฮเทลลิงทีสแควร์ วิธีบอนเฟอร์โรนี-โฮล์ม วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก วิธีเวสต์ฟอล-ยัง บูทสแตรป และวิธีการเรียงสับเปลี่ยนที่แท้จริง โดยพิจารณาจากความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 และกำลังการทดสอบ เมื่อประชากรทั้งสองกลุ่มมีการแจกแบบปกติหลายตัวแปร ซึ่งมีเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมเท่ากันและเท่ากับเมทริกซ์สหสัมพันธ์ โดยมีจำนวนตัวแปรตามเท่ากับ 3, 5 และ 7 ตัว ขนาดตัวอย่างเท่ากับ 10, 30, และ 50 โครงสร้างของสหสัมพันธ์เป็นแบบไม่เท่ากันที่มีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.0, 0.3, 0.5, 0.7 และ 0.9 ณ ระดับนัยสำคัญ 0.05 ศึกษาด้วยการจำลองด้วยเทคนิคมอนติคาร์โล ทำซ้ำ 1,000 ครั้งในแต่ละสถานการณ์ ผลที่ได้จากการวิจัยพบว่า ส่วนใหญ่วิธีโฮเทลลิงทีสแควร์มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนด ส่วนวิธีอื่นๆ มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 ขนาดตัวอย่างไม่มีผลต่อความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 แต่จำนวนตัวแปรตามมีผลต่อความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 นอกจากนี้วิธีเวสต์ฟอล-ยัง บูทสแตรปและวิธีการเรียงสับเปลี่ยนที่แท้จริงมีอัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 มากที่สุด ภายใต้สถานการณ์ส่วนใหญ่วิธีเวสต์ฟอล-ยัง บูทสแตรปมีกำลังการทดสอบเชิงประจักษ์สูงสุด แต่วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลักจะมีกำลังการทดสอบเชิงประจักษ์สูงสุดเมื่อสัมประสิทธิ์สหสัมพันธ์มีค่าต่ำ และวิธีโฮเทลลิงทีสแควร์มีกำลังการทดสอบเชิงประจักษ์ต่ำสุดในทุกกรณี วิธีเวสต์ฟอล-ยัง บูทสแตรป และวิธีการเรียงสับเปลี่ยนที่แท้จริงมีกำลังการทดสอบเชิงประจักษ์ใกล้เคียงกันในทุกสถานการณ์ นอกจากนี้กำลังการทดสอบเชิงประจักษ์แปรผันตามขนาดตัวอย่างและจำนวนตัวแปรตาม แต่แปรผกผันกับค่าสัมประสิทธิ์สหสัมพันธ์

---

[1] *นักศึกษาบัณฑิตศึกษา ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี*

[2] *รองศาสตราจารย์ ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี*

# An Efficiency Comparison of Closed Multiple Test Methods for Population Means under Unequal Correlation Matrix

**Hathairat Wongchaisuwat [1] and Kamon Budsaba [2]**

Thammasat University, Rangsit Center, Khlong Luang, Pathum Thani 12121

*Received 4 April 2007 ; accepted 29 June 2007*

## Abstract

The objective of this research is to compare five closed multiple test methods with stepwise procedure for testing the difference between two population means : Hotelling's $T^2$ method, Bonferroni-Holm method, Hommel's method based on Simes' test, Westfall-Young bootstrap method and Exact Permutational method, by considering their capacity of controlling type I error rate and their power of the test under multivariate normal distributions with the same covariance matrix which equals to the correlation matrix for 3, 5 and 7 dependent variables; 10, 30, 50 and 70 equal sample sizes; unequal correlation design matrix with correlation coefficient  equals to 0.0, 0.3, 0.5, 0.7 and 0.9 for the case of unequal correlation design matrix at 0.05 significant level ($\alpha$). Monte Carlo simulations was performed and repeated 1,000 times for each scenario. The results showed that in most situations, Hotelling's $T^2$ method has empirical type I error rate less than lower bound of the tolerance type I error rate controllable criterion, while others have empirical type I error rate lies in the interval of the tolerance criterion. Sample size do not affect the capacity of controlling type I error rate but the number of dependent variables affect the capacity of controlling type I error rate.   In addition, Westfall-Young bootstrap method and Exact Permutational method have empirical type I error rate lies in the interval of the tolerance criterion. For almost every situations Westfall-Young bootstrap method has the highest empirical power but Hommel's method based on Simes' test has the the highest empirical power when the correlation coefficient is low. Hotelling's $T^2$ method has the lowest empirical power in all situations. Westfall-Young bootstrap method and Exact Permutational method has similar empirical power in all situations. In addition, empirical power varies according to the sample size and the number of dependent variables but varies inversely with the correlation coefficient.

---

*[1] Graduate Student, Department of Mathematics and Statistics, Faculty of Science and Technology.*

*[2] Associate Professor, Department of Mathematics and Statistics, Faculty of Science and Technology.*

## 1. บทนำ

ในการศึกษาวิจัยทางการแพทย์หรือทางสังคมศาสตร์ส่วนใหญ่มักเกี่ยวข้องกับการเปรียบเทียบระหว่างวิธีการปฏิบัติ 2 วิธี เช่น การเปรียบเทียบระหว่างวิธีการรักษาแบบใหม่กับวิธีการรักษาแบบมาตรฐาน ซึ่งโดยปกติแล้วการเปรียบเทียบความแตกต่างของวิธีการรักษา 2 วิธีนี้ไม่ควรเปรียบเทียบโดยพิจารณาจากตัวแปรเพียงตัวเดียว แต่ควรวัดจากตัวแปรหลายๆ ตัว ซึ่งน่าจะใช้วัดประสิทธิภาพของวิธีการรักษาได้ดีกว่า เช่น การศึกษาประสิทธิภาพของยาในการรักษาโรคหืดโดยเปรียบเทียบระหว่างกลุ่มที่รับยารักษากับกลุ่มที่ได้ยาลวง (placebo) โดยวัดผลจากตัวแปร 3 ตัว ได้แก่ ปริมาณการหายใจออกในหนึ่งวินาที อัตราการหายใจออกสูงสุดต่อหนึ่งนาที และคะแนนแสดงระดับของอาการ ซึ่งนักวิจัยต้องการทราบว่าการใช้ยารักษาและยาลวงให้ผลแตกต่างกันหรือไม่ [1] ดังนั้นวิธีการทางสถิติที่จะนำมาใช้กับสถานการณ์นี้ควรเป็นวิธีที่คำนึงถึงความหลากหลายของตัวแปร กล่าวคือ ตัวแปรทุกตัวที่เราพิจารณาควรจะมีอิทธิพลต่อการตัดสินใจในการทดสอบสมมติฐานระหว่างวิธีการปฏิบัติ 2 วิธี และควรคำนึงถึงความสัมพันธ์ระหว่างตัวแปรตามเหล่านี้ด้วย

ปัญหาในลักษณะดังกล่าวนั้น เรียกว่า "การทดสอบพหุคูณ (Multiple Testing)" ซึ่งเป็นการเปรียบเทียบความแตกต่างระหว่างกลุ่มสองกลุ่ม (หรือมากกว่า) โดยพิจารณาจากตัวแปรตามหลายตัวพร้อมกัน โดยมีจุดมุ่งหมายเพื่อควบคุม "ค่ามากที่สุดของอัตราการเกิดความผิดพลาดแบบที่หนึ่งโดยรวม (maximum overall Type I error rate)" ซึ่งหมายถึงค่ามากที่สุดของความน่าจะเป็นที่สมมติฐานว่างอย่างน้อยหนึ่งสมมติฐานจะถูกปฏิเสธเมื่อจริงๆ แล้วสมมติฐานว่างนั้นเป็นจริง อัตราความผิดพลาดดังกล่าวมีชื่อเรียกว่า "อัตราความผิดพลาดสูงสุดต่อการทดลอง (Maximum Experimentwise Error Rate: MEER)" หรือ "อัตราความผิดพลาดสูงสุดต่อวงศ์ (Maximum Familywise Error Rate: FWE)" [2]

ในกรณีที่มีตัวแปรตามที่ต้องการทดสอบ $m$ ตัว จะมีเซตของสมมติฐานว่างที่ต้องการทดสอบ คือ $H = \{H_{01}, H_{02},..., H_{0m}\}$ การทดสอบพหุคูณจะควบคุม FWE อย่างเข้มงวด ถ้าความน่าจะเป็นของความผิดพลาดแบบที่ 1 (Type I Error) ในการทดสอบอย่างน้อย 1 การทดสอบมีค่าไม่เกินระดับนัยสำคัญ $\alpha$ โดยไม่คำนึงถึงว่าจริงๆ แล้วมีสมมติฐานใดบ้างใน $H$ ที่เป็นจริง การทดสอบพหุคูณในปัจจุบันจะใช้ "การทดสอบแบบปิด (closed testing)" เข้ามาประยุกต์ใช้ ซึ่งมีทั้งการทดสอบแบบขั้นบันได (step-wise method) และแบบขั้นตอนเดียว (single-step method)

การทดสอบพหุคูณโดยวิธีการทดสอบแบบปิดที่ใช้วิธีแบบขั้นบันไดมีหลายวิธี วิธีแบบนี้มีกำลังการทดสอบมากกว่าการใช้วิธีการทดสอบแบบขั้นตอนเดียว โดยสามารถควบคุม FWE ได้ [2] ผู้วิจัยจึงสนใจเปรียบเทียบกำลังการทดสอบของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดที่มีให้เลือกใช้ในโปรแกรมสำเร็จรูป SAS จำนวน 5 วิธี ได้แก่ (1) วิธีโฮเทลลิงทีสแควร์ (HOTE) (2) วิธีบอนเฟอร์โรนี-โฮล์ม (BON) (3) วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (HOM) (4) วิธีเวสต์ฟอล-ยังบูทสแตรป (BOOT) และ (5) วิธีการเรียงสับเปลี่ยนที่แท้จริง (PERM)

## 2. วัตถุประสงค์ของการวิจัย

เพื่อศึกษาเปรียบเทียบวิธีการทดสอบตัวแปรพหุคูณแบบปิดโดยใช้วิธีแบบขั้นบันไดที่ใช้ในการทดสอบความแตกต่างของค่าเฉลี่ยระหว่างกลุ่มประชากร 2 กลุ่ม ที่มีตัวแปรหลายตัวจำนวน 5 วิธี ภายใต้สถานการณ์ต่างๆ

## 3. วิธีการดำเนินการวิจัย

การวิจัยครั้งนี้ใช้ข้อมูลที่ได้จากการจำลองด้วยเทคนิคมอนติคาร์โลทำซ้ำ 1,000 รอบ ในแต่ละสถานการณ์ โดยใช้โปรแกรม SAS รุ่น เวอร์ชัน 8.02 ประชากรที่ทำการศึกษามี 2 กลุ่ม แต่ละกลุ่มมีการแจกแจงแบบปกติของตัวแปรตามพหุคูณ และกำหนดให้

(1) ระดับนัยสำคัญเท่ากับ 0.05

(2) ขนาดตัวอย่างของประชากรทั้ง 2 กลุ่มเท่ากัน และเท่ากับ 10, 30, 50 และ 70

(3) จำนวนตัวแปรตามเท่ากับ 3, 5 และ 7 ตัวแปร

(4) ค่าสัมประสิทธิ์สหสัมพันธ์ $\rho_{kk'} = \rho^{|k-k'|}$ โดยที่ $\rho$ = 0.0, 0.3, 0.5, 0.7 และ 0.9

(5) กรณีพิจารณากำลังการทดสอบ กำหนดให้ $\Delta_k = \mu_{1k} - \mu_{2k}$ โดยที่ $\Delta_k = 0.1 \times (m-k+1)$

### 3.1 วิธีการทดสอบแบบปิดสำหรับการทดสอบพหุคูณ

ถ้าต้องการทดสอบสมมติฐาน $H_{01}, H_{02}$ และ $H_{03}$ โดยที่ $H_{01}$ คือ สมมติฐานของการเปรียบเทียบวิธีการปฏิบัติของตัวแปรที่ 1 วิธีการทดสอบแบบปิดมีขั้นตอนดังนี้

(1) ทดสอบแต่ละสมมติฐาน $H_{01}, H_{02}$ และ $H_{03}$ โดยใช้ระดับนัยสำคัญของการทดสอบ $(\alpha)$ ที่เหมาะสม

(2) สร้างเซตของอินเตอร์เซกชันที่เป็นไปได้ทั้งหมดระหว่าง $H_{01}, H_{02}$ และ $H_{03}$ ในกรณีนี้จะได้สมมติฐาน $H_{012}, H_{013}, H_{023}$ และ $H_{0123}$

(3) การทดสอบแต่ละอินเตอร์เซกชันต้องใช้ระดับนัยสำคัญของการทดสอบที่เหมาะสม อาจใช้ F-tests, MANOVA tests หรือวิธีอื่นๆ ที่ใช้สำหรับการทดสอบสมมติฐานอินเตอร์เซกชัน

(4) ขอบเขตวิกฤตของการทดสอบ คือ จะปฏิเสธสมมติฐานด้วยการควบคุมอัตราความผิดพลาดของการทดลอง MEER เมื่อเป็นไปตามเงื่อนไขทั้ง 2 ข้อ ดังต่อไปนี้

(i) ผลการทดสอบ $H_{ok}$ มีนัยสำคัญ

(ii) การทดสอบทุกๆ สมมติฐานอินเตอร์เซกชันต้องประกอบด้วย $H_{ok}$ ที่มีนัยสำคัญ

ในการหาขอบเขตวิกฤตจะต้องพิจารณาค่า p-value ที่ปรับแล้ว ซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาค่า p-value ที่ได้จากการทดสอบสมมติฐานที่ประกอบด้วยตัวแปรที่ k ถ้าค่า p-value ที่ปรับแล้วน้อยกว่าระดับนัยสำคัญ $\alpha$ แล้วจะปฏิเสธสมมติฐานว่างนั้น

### 3.2 วิธีโฮเทลลิงทีสแควร์

วิธีโฮเทลลิงทีสแควร์เป็นวิธีการทดสอบสมมติฐานเชิงพหุคูณที่ต้องหาค่า p-value ของทุกสมมติฐานเชิงเดี่ยวและสมมติฐานอินเตอร์เซกชัน เช่น ต้องการทดสอบค่าเฉลี่ยของตัวแปร 3 ตัว จะต้องหาค่า p-value ของสมมติฐาน $H_{01}, H_{02}, H_{03}, H_{012}, H_{013}, H_{023}$ และ $H_{0123}$ แล้วจึงไปหาขอบเขตวิกฤตของการทดสอบ โดยพิจารณาค่า p-value ที่ปรับแล้ว ถ้าค่า p-value ที่ปรับแล้วน้อยกว่าระดับนัยสำคัญ $\alpha$ แล้วจะปฏิเสธสมมติฐานว่าง [2]

### 3.3 วิธีบอนเฟอร์โรนี-โฮล์ม

วิธีบอนเฟอร์โรนี-โฮล์มเป็นวิธีการทดสอบที่พิจารณาค่า p-value ที่มีค่าน้อยที่สุดของการทดสอบแต่ละตัวแปรกับ $\alpha / k^*$ เมื่อ $k^*$ คือ จำนวนตัวแปรในสมมติฐานอินเตอร์เซกชันและ $\alpha$ คือ ระดับนัยสำคัญที่กำหนดขึ้น และจะปฏิเสธสมมติฐานเมื่อ $\min p \le \alpha / k^*$ โดยที่ $\min\ p$ คือ ค่า p-value ที่มีค่าน้อยที่สุดของการทดสอบแต่ละตัวแปร หรือสามารถปฏิเสธสมมติฐานอินเตอร์เซกชัน เมื่อ $k^* \times \min p \le \alpha$ ดังนั้น $k^* \times \min p$ คือ ค่า p-value ของสมมติฐานอินเตอร์เซกชันโดยที่การทดสอบสมมติฐานเชิงเดี่ยวจะมีค่า p-value ที่ได้มาจากการทดสอบสมมติฐาน ด้วยตัวสถิติโฮเทลลิงทีสแควร์ [2]

### 3.4 วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก

ซิมส์ [3] ได้เสนอวิธีการทดสอบสมมติฐานเชิงพหุคูณโดยพิจารณาค่า p-value ที่เรียงลำดับจากน้อยไปหามาก ถ้ามีสมมติฐาน m สมมติฐาน ให้ $P_{(1)} \le P_{(2)} \le \ldots \le P_{(\mathrm{m})}$ เป็นลำดับของค่า p-value ที่สมมติฐาน $H_{(0i)}$ มีค่า p-value เป็น $P_{(i)}$ โดยที่ $i = 1, 2, \ldots, m$ จะปฏิเสธสมมติฐานอินเตอร์เซกชันได้ถ้า $p_{(\mathrm{i})} \le i\alpha / k$ สำหรับ $i$ อย่างน้อยหนึ่งค่า โดยที่ $k$ คือ จำนวนตัวแปร หรือ จะปฏิเสธสมมติฐานอินเตอร์เซกชันถ้า $\min(k^*\ p_{(i)} / i) \le \alpha$ ดังนั้น $\min(k^*\ p_{(i)} / i)$ คือ ค่า p-value สำหรับการทดสอบสมมติฐานอินเตอร์เซกชัน [2] เมื่อทำการทดสอบโดยวิธีของซิมส์แบบปิดจะสามารถสร้างผลสรุปเกี่ยวกับสมมติฐานเชิงเดี่ยวได้ ซึ่งผลของวิธีนี้ถูกเรียกว่า วิธีของโฮมเมล [4]

### 3.5 วิธีเวสต์ฟอล-ยัง บูทสแตรป

ในปี ค.ศ. 1989 เวสต์ฟอลและยัง [5] ได้เสนอวิธีการทดสอบโดยใช้เทคนิคบูทสแตรปในการปรับค่า P สำหรับข้อมูลที่มีหลายตัวแปร และได้เสนอวิธีที่ขยายจากวิธีการดั้งเดิมในปี ค.ศ. 1993 ซึ่งมีขั้นตอนดังนี้

(1) สุ่มตัวอย่างขนาด $n_1$ และ $n_2$ (แบบแทนที่) สำหรับกลุ่มวิธีการปฏิบัติที่ 1 และ 2 ตามลำดับ จาก

ข้อมูลดั้งเดิมที่รวมตัวอย่างของทั้ง 2 กลุ่ม เข้าไว้ด้วยกัน โดยถือว่า $H_0$ เป็นจริง นั่นคือ ไม่มีความแตกต่างระหว่างกลุ่มวิธีการปฏิบัติ

(2) คำนวณค่าบูทสแตรป $P$ ได้แก่ $P^*_{[1]}, P^*_{[2]},\ldots, P^*_{[m]}$ ด้วยวิธีการเดียวกับการคำนวณค่า $P$ ที่ได้จากข้อมูลดั้งเดิม แต่ค่าบูทสแตรป $P$ ที่คำนวณได้นี้ไม่จำเป็นจะต้องเรียงลำดับเหมือนกับค่า $P$ ดั้งเดิม ทำเช่นเดียวกันนี้ซ้ำๆ กัน N ครั้ง ซึ่งจะได้เวกเตอร์ขนาด m ของค่าบูทสแตรป $P$ ทั้งหมด N เวกเตอร์

(3) สามารถประมาณค่า $P$ ซึ่งปรับค่าแล้วสำหรับตัวแปรที่สอดคล้องกับสมมติฐาน $H_{0[1]}$ (ซึ่งคือตัวแปรที่ให้ค่า $P$ น้อยที่สุดจากข้อมูลดั้งเดิม) ได้จากสัดส่วนของเหตุการณ์ $\min\limits_{1\le\tilde{k}\le m} P^*_{[\tilde{k}]} \le P_{[1]}$ แทนด้วย $\tilde{P}^{adj}_{[1]}$ สำหรับ ตัวแปรที่สอดคล้องกับ $H_{0[2]}$ เราสามารถประมาณค่าบูทสแตรปที่ปรับแล้ว $\tilde{P}^{adj}_{[2]}$ ได้จากสัดส่วนของเหตุการณ์ $\min\limits_{2\le\tilde{k}\le m} P^*_{[\tilde{k}]} \le P_{[2]}$ และสำหรับตัวแปรที่สอดคล้องกับ $H_{0[k]}$ และสามารถประมาณค่าบูทสแตรปที่ปรับแล้ว $\tilde{P}^{adj}_{[k]}$ ได้จากสัดส่วนของเหตุการณ์ $\min\limits_{k\le\tilde{k}\le m} P^*_{[\tilde{k}]} \le P_{[k]}$

(4) หลังจากคำนวณค่า $P$ ที่ปรับค่าโดยใช้เทคนิคบูทสแตรป $\tilde{P}^{adj}_{[1]}, \tilde{P}^{adj}_{[2]},\ldots, \tilde{P}^{adj}_{[m]}$, (ซึ่งไม่จำเป็นต้องเรียงตามลำดับ) แล้ว จะปฏิเสธ $H_{0k}$ ก็ต่อเมื่อ $\tilde{P}^{adj}_{[\tilde{k}]}, \le \alpha$ สำหรับทุก $\tilde{k} \le k$

วิธีการดั้งเดิมของเวสต์ฟอลและยังซึ่งเสนอในปี ค.ศ. 1989 นั้น ค่า $P$ ที่ปรับค่าโดยใช้เทคนิคบูทสแตรปประมาณได้จากสัดส่วนของเหตุการณ์ $\min\limits_{k\le\tilde{k}\le m} P^*_{[\tilde{k}]} \le P_{[k]}$ โดยที่ $k = 1,2,\ldots,m$ [5]

### 3.6 วิธีการเรียงสับเปลี่ยนที่แท้จริง

วิธีการเรียงสับเปลี่ยนที่แท้จริงจะเหมือนกับวิธีเวสต์ฟอล-ยัง บูทสแตรป ยกเว้นการสุ่มตัวอย่างเป็นแบบไม่แทนที่ โดยที่วิธีเวสต์ฟอล-ยัง บูทสแตรปมีการสุ่มตัวอย่างเป็นแบบแทนที่ [2] ซึ่งมีขั้นตอนเหมือนกับวิธีเวสต์ฟอล-ยัง บูทสแตรป

### 3.7 วิธีการสร้างข้อมูลให้มีการแจกแจงปกติ

การสร้างข้อมูลให้มีการแจกแจงปกติของตัวแปรพหุนั้น จำเป็นต้องสร้างตัวเลขสุ่มที่การแจกแจงปกติมาตรฐานของตัวแปรพหุก่อน โดยใช้โปรแกรม SAS ด้วยการใช้คำสั่ง rannor (seed) จากนั้นจึงทำการสร้างตัวเลขสุ่มที่มีการแจกแจงแบบปกติของตัวแปรพหุจากสมการ $\underset{\sim}{y} = C\underset{\sim}{z}$

เมื่อ $\underset{\sim}{y}' = [y_1, y_2,\ldots,y_m]$ มีการแจกแจงแบบปกติของตัวแปรตาม m ตัวแปร และมีเมทริกซ์ความแปรปรวนร่วมเป็น $\Sigma$ และ $\underset{\sim}{z}' = [z_1,z_2,\ldots, z_m]$ มีการแจกแจงแบบปกติมาตรฐานของ m ตัวแปร โดยมีเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{0}$ และเมทริกซ์ความแปรปรวนร่วม คือ $I_m$ และ $I_m$ เป็นเมทริกซ์เอกลักษณ์ขนาด $m\times m$ ให้ $C$ เป็นเมทริกซ์สามเหลี่ยมล่างเพียงเมทริกซ์เดียว (unique lower triangular matrix) ที่ทำให้

$$CC'=\Sigma \quad \text{และ} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{1m} \\ \sigma_{21} & \sigma_{22}\ldots & \sigma_{2m} \\ \vdots & \vdots \ddots & \vdots \\ \sigma_{m1} & \sigma_{m2}\ldots & \sigma_{mm} \end{bmatrix}$$

การเขียน $C$ ในรูปผลคูณของเมทริกซ์สามเหลี่ยมล่าง เรียกว่า Cholesky decomposition การหาเมทริกซ์ $C$ มีสูตรในการหาดังนี้ [5]

$$c_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{jk}}{\left(\sigma_{jj} - \sum_{k=1}^{j-1} c^2_{jk}\right)^{1/2}}$$

เมื่อ $\sum_{k=1}^{0} c_{ik}c_{jk} = 0 \quad , 1 \le j \le i \le m$

นั่นคือ $c_{i1} = \dfrac{\sigma_{i1}}{\sqrt{\sigma_{11}}} \quad , 1 \le i \le m$

$$c_{ii} = \sqrt{\sigma_{i1} - \sum_{k=1}^{i-1} c^2_{ik}} \quad , 1 < i \le m$$

$$c_{ij} = \frac{\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{jk}}{c_{jj}} \quad , 1 < j < i \le m$$

$$c_{ij} = 0 \quad , 1 \le i < j \le m$$

### 3.8 การหาอัตราความผิดพลาดแบบที่ 1 จากการทดลอง

ขั้นตอนในการคำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 คือ

(1) กำหนดจำนวนตัวแปรตาม และขนาดตัวอย่างตามสถานการณ์ต่างๆ

(2) กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.05

(3) กำหนดโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรต่างๆ และหาเมทริกซ์ $C$

(4) จำลองข้อมูลให้มีการแจกแจงแบบปกติมาตรฐานที่มีจำนวนตัวแปรตามและขนาดตัวอย่างตามแต่ละสถานการณ์ให้อยู่ในรูปของเมทริกซ์จำนวน 2 กลุ่ม

(5) สร้างข้อมูลที่มีการแจกแจงแบบปกติที่ตัวแปรตามมีสหสัมพันธ์กันจากสมการ $\underset{\sim}{y} = C\underset{\sim}{z}$ ในรูปของเมทริกซ์จำนวน 2 กลุ่มเป็น $Y_1$ และ $Y_2$

(6) ทดสอบสมมติฐานเพื่อหาค่า p-value ของสมมติฐานว่างเชิงเดี่ยว $H_{0k}$ และสมมติฐานอินเตอร์เซกชันทุกสมมติฐาน

(7) หาค่า p-value ที่ปรับแล้วของแต่ละตัวแปร ซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาสมมติฐานที่ประกอบด้วยการทดสอบตัวแปรที่ k แล้วพิจารณาว่ายอมรับหรือปฏิเสธสมมติฐานว่าง $H_{0k}$ โดยที่ $k = 1,2,...,m$

(8) ทำซ้ำจำนวน 1,000 ครั้ง แล้วนับจำนวนการปฏิเสธสมมติฐานว่างเชิงพหุคูณ (ให้สมมติฐานว่างเชิงพหุคูณเป็นดังนี้ $H_0 = H_{01} \cap H_{02} \cap ... \cap H_{0m}$ และจะปฏิเสธสมมติฐานว่างเชิงพหุคูณเมื่อมีการปฏิเสธสมมติฐานเชิงเดี่ยว $H_{0k}$ อย่างน้อยหนึ่งสมมติฐาน)

(9) คำนวณความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลองจากอัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่างเชิงพหุคูณเมื่อสมมติฐานว่างเชิงพหุคูณเป็นจริง เทียบกับจำนวนครั้งที่ทำการทดลอง ซึ่งในที่นี้คือ 1,000 ครั้ง

### 3.9 การทดสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1

ใช้สถิติทดสอบซี (z-test) ในการทดสอบสมมติฐานแบบสองด้านที่ระดับนัยสำคัญ 0.05 โดยมีสมมติฐาน คือ

$$H_0 : \alpha = \alpha_0$$
$$H_1 : \alpha \neq \alpha_0$$

สถิติทดสอบ $$Z = \frac{\alpha_* - \alpha_0}{\sqrt{\dfrac{\alpha_0 (1-\alpha_0)}{N}}}$$

กำหนดให้ $\alpha$ แทน ความน่าจะเป็นของความผิดพลาดแบบที่ 1

$\alpha_*$ แทน ความน่าจะเป็นของความผิดพลาดแบบที่ 1 จากการทดลอง

$\alpha_0$ แทน ระดับนัยสำคัญที่กำหนดในการศึกษา คือ 0.05

N แทน จำนวนครั้งที่ทำการทดลอง ในที่นี้เท่ากับ 1,000 ครั้ง

ตัวสถิติทดสอบจะสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ได้ ถ้า $-1.96 \leq Z \leq 1.96$ ดังนั้น ด้วยความเชื่อมั่น 95% ตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ $\alpha_0 = 0.05$ ได้ถ้า $\alpha_*$ มีค่าอยู่ในช่วง [0.036, 0.064]

### 3.10 การหากำลังการทดสอบเชิงประจักษ์

ขั้นตอนในการหากำลังการทดสอบเชิงประจักษ์มีดังนี้

(1) กำหนดจำนวนตัวแปรตาม และขนาดตัวอย่างตามสถานการณ์ต่างๆ

(2) กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.05

(3) กำหนดโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรต่างๆ และหาเมทริกซ์ $C$

(4) จำลองข้อมูลที่มีการแจกแจงแบบปกติมาตรฐานโดยมีจำนวนตัวแปรตามและขนาดตัวอย่างตามแต่ละสถานการณ์ให้อยู่ในรูปของเมทริกซ์จำนวน 2 กลุ่ม

(5) สร้างข้อมูลที่มีการแจกแจงแบบปกติที่ตัวแปรตามมีสหสัมพันธ์กันจากสมการ $\underset{\sim}{y} = C\underset{\sim}{z}$ ในรูปของเมทริกซ์จำนวนสองกลุ่มเป็น $Y_1$ และ $Y_2$

(6) ให้สมมติฐานว่างไม่เป็นจริง โดยกำหนดให้ $\Delta_k = \mu_{1k} - \mu_{2k}$ โดยที่ $\Delta_k = 0.1\times(m\text{-}k+1)$

(7) ให้เมทริกซ์ของข้อมูลของกลุ่มที่ 1 เป็น $Y_1+\Delta_k$ และเมทริกซ์ของข้อมูลกลุ่มที่ 2 เป็น $Y_2$

(8) ทดสอบสมมติฐานเพื่อหาค่า p-value ของสมมติฐานว่างเชิงเดี่ยว $H_{0k}$ และสมมติฐานอินเตอร์เซกชันทุกสมมติฐาน

(9) หาค่า p-value ที่ปรับแล้วของแต่ละตัวแปรซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาสมมติฐานที่ประกอบด้วยการทดสอบตัวแปรที่ k แล้วพิจารณาว่ายอมรับหรือปฏิเสธสมมติฐานว่าง $H_{0k}$ โดยที่ $k = 1,2,...,m$

(10) ทำซ้ำจำนวน 1,000 ครั้ง แล้วนับจำนวนการปฏิเสธสมมติฐานว่างเชิงพหุคูณ

(11) คำนวณหากำลังการทดสอบเชิงประจักษ์จากอัตราส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่างเชิงพหุคูณเมื่อสมมติฐานว่างเชิงพหุคูณไม่เป็นจริง เทียบกับจำนวนครั้งที่ทำการทดลอง ซึ่งในที่นี้คือ 1,000 ครั้ง

## 4. ผลการวิจัย

### 4.1 อัตราความผิดพลาดแบบที่ 1

อัตราความผิดพลาดแบบที่ 1 ของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ เมื่อกำหนดจำนวนตัวแปรตาม 3, 5 และ 7 ตัว ณ ระดับนัยสำคัญ $\alpha = 0.05$ แสดงในตารางที่ 1, 2 และ 3 ตามลำดับ

**ตารางที่ 1** อัตราความผิดพลาดแบบที่ 1 จากการทดลองของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 3

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | 0.043 | 0.051 | 0.053 | 0.052 | 0.050 |
| | 0.3 | 0.032* | 0.045 | 0.045 | 0.048 | 0.047 |
| | 0.5 | 0.033* | 0.050 | 0.053 | 0.055 | 0.055 |
| | 0.7 | 0.034* | 0.050 | 0.053 | 0.062 | 0.065* |
| | 0.9 | 0.022* | 0.024* | 0.029* | 0.034* | 0.034* |
| 30 | 0.0 | 0.042 | 0.054 | 0.055 | 0.054 | 0.052 |
| | 0.3 | 0.038 | 0.053 | 0.054 | 0.055 | 0.054 |
| | 0.5 | 0.027* | 0.038 | 0.039 | 0.042 | 0.041 |
| | 0.7 | 0.034* | 0.045 | 0.047 | 0.051 | 0.051 |
| | 0.9 | 0.017* | 0.023* | 0.030* | 0.045 | 0.048 |
| 50 | 0.0 | 0.048 | 0.053 | 0.055 | 0.054 | 0.054 |
| | 0.3 | 0.039 | 0.048 | 0.048 | 0.050 | 0.052 |
| | 0.5 | 0.026* | 0.040 | 0.044 | 0.044 | 0.042 |
| | 0.7 | 0.033* | 0.045 | 0.046 | 0.051 | 0.050 |
| | 0.9 | 0.027* | 0.035* | 0.039 | 0.053 | 0.052 |

**หมายเหตุ** * หมายถึง อัตราความผิดพลาดแบบที่ 1 อยู่นอกช่วง [0.036, 0.064]

จากตารางที่ 1 จะเห็นได้ว่า

(1) วิธี HOTE มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.3 เมื่อ n=10, $\rho$=0.5 เมื่อ n=10, 30 และ 50, $\rho$=0.7 เมื่อ n=10, 30 และ 50 และ $\rho$=0.9 เมื่อ n=10, 30 และ 50

(2) วิธี BON มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.9 เมื่อ n=10, 30 และ 50

(3) วิธี HOM มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.9 เมื่อ n=10, 30

(4) วิธี BOOT และ PERM มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.9 เมื่อ n=10

(5) วิธี PERM มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองมากกว่าขอบเขตบนของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.7 เมื่อ n=10

**ตารางที่ 2** อัตราความผิดพลาดแบบที่ 1 จากการทดลองของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 5

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | 0.020* | 0.057 | 0.057 | 0.061 | 0.057 |
| | 0.3 | 0.022* | 0.049 | 0.049 | 0.049 | 0.049 |
| | 0.5 | 0.011* | 0.053 | 0.053 | 0.053 | 0.054 |
| | 0.7 | 0.011* | 0.031* | 0.033* | 0.038 | 0.038 |
| | 0.9 | 0.011* | 0.024* | 0.027* | 0.046 | 0.046 |
| 30 | 0.0 | 0.011* | 0.036 | 0.037 | 0.036 | 0.036 |
| | 0.3 | 0.024* | 0.045 | 0.045 | 0.047 | 0.045 |
| | 0.5 | 0.022* | 0.052 | 0.053 | 0.055 | 0.053 |
| | 0.7 | 0.018* | 0.034* | 0.038 | 0.045 | 0.044 |
| | 0.9 | 0.018* | 0.032* | 0.040 | 0.060 | 0.058 |
| 50 | 0.0 | 0.030* | 0.053 | 0.053 | 0.055 | 0.053 |
| | 0.3 | 0.024* | 0.044 | 0.044 | 0.046 | 0.044 |
| | 0.5 | 0.018* | 0.048 | 0.050 | 0.055 | 0.052 |
| | 0.7 | 0.013* | 0.042 | 0.042 | 0.053 | 0.051 |
| | 0.9 | 0.020* | 0.038 | 0.046 | 0.054 | 0.056 |

**หมายเหตุ** * หมายถึง อัตราความผิดพลาดแบบที่ 1 อยู่นอกช่วง [0.036, 0.064]

จากตารางที่ 2 จะเห็นได้ว่า

(1) วิธี HOTE อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในทุกกรณี

(2) วิธี BON อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.7 เมื่อ n=10 และ 30 และ $\rho$=0.9 เมื่อ n=10 และ 30

(3) วิธี HOM อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.7 เมื่อ n=10 และ $\rho$=0.9 เมื่อ n=10

(4) วิธี BOOT และ PERM อัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ให้ใกล้เคียง 0.05 ในทุกกรณี

**ตารางที่ 3** อัตราความผิดพลาดแบบที่ 1 จากการทดลองของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 7

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | 0.004* | 0.031* | 0.033* | 0.035* | 0.031* |
| | 0.3 | 0.015* | 0.053 | 0.053 | 0.058 | 0.058 |
| | 0.5 | 0.002* | 0.047 | 0.047 | 0.056 | 0.053 |
| | 0.7 | 0.002* | 0.042 | 0.044 | 0.058 | 0.053 |
| | 0.9 | 0.004* | 0.036 | 0.038 | 0.053 | 0.060 |
| 30 | 0.0 | 0.011* | 0.056 | 0.058 | 0.060 | 0.060 |
| | 0.3 | 0.018* | 0.051 | 0.051 | 0.053 | 0.053 |
| | 0.5 | 0.013* | 0.058 | 0.058 | 0.060 | 0.060 |
| | 0.7 | 0.000* | 0.044 | 0.047 | 0.049 | 0.049 |
| | 0.9 | 0.004* | 0.036 | 0.038 | 0.064 | 0.062 |
| 50 | 0.0 | 0.011* | 0.051 | 0.051 | 0.053 | 0.053 |
| | 0.3 | 0.016* | 0.040 | 0.040 | 0.042 | 0.040 |
| | 0.5 | 0.007* | 0.047 | 0.049 | 0.056 | 0.051 |
| | 0.7 | 0.007* | 0.020* | 0.020* | 0.024* | 0.027* |
| | 0.9 | 0.004* | 0.024* | 0.027* | 0.053 | 0.056 |

**หมายเหตุ** * หมายถึง อัตราความผิดพลาดแบบที่ 1 อยู่นอกช่วง [0.036, 0.064]

จากตารางที่ 3 จะเห็นได้ว่า

(1) วิธี HOTE อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในทุกกรณี

(2) วิธี BON และ HOM อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.0 เมื่อ n=10, $\rho$=0.7 เมื่อ n=50 และ $\rho$=0.9 เมื่อ n=50

(3) วิธี BOOT และ PERM อัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนดในกรณีที่ $\rho$=0.0 เมื่อ n=10 และ $\rho$=0.7 เมื่อ n=50

### 4.2 กำลังการทดสอบเชิงประจักษ์

กำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ เมื่อกำหนดจำนวนตัวแปรตามเท่ากับ 3, 5 และ 7 ตัว ณ ระดับนัยสำคัญ $\alpha$ = 0.05 แสดงในตารางที่ 4, 5 และ 6 ตามลำดับ

**ตารางที่ 4** กำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 3

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | $0.062_{(5)}$ | $0.088_{(3)}$ | $0.089_{(1)}$ | $0.088_{(3)}$ | $0.088_{(3)}$ |
| | 0.3 | $0.056_{(5)}$ | $0.087_{(4)}$ | $0.090_{(2)}$ | $0.091_{(1)}$ | $0.090_{(2)}$ |
| | 0.5 | $0.041_{(5)}$ | $0.060_{(4)}$ | $0.063_{(3)}$ | $0.067_{(1)}$ | $0.065_{(2)}$ |
| | 0.7 | $0.043_{(5)}$ | $0.068_{(4)}$ | $0.072_{(3)}$ | $0.084_{(2)}$ | $0.085_{(1)}$ |
| | 0.9 | $0.036_{(5)}$ | $0.054_{(4)}$ | $0.062_{(3)}$ | $0.070_{(2)}$ | $0.078_{(1)}$ |
| 30 | 0.0 | $0.161_{(5)}$ | $0.187_{(3)}$ | $0.191_{(1)}$ | $0.187_{(3)}$ | $0.189_{(2)}$ |
| | 0.3 | $0.142_{(5)}$ | $0.162_{(4)}$ | $0.165_{(3)}$ | $0.169_{(1)}$ | $0.167_{(2)}$ |
| | 0.5 | $0.109_{(5)}$ | $0.154_{(4)}$ | $0.158_{(3)}$ | $0.166_{(1)}$ | $0.162_{(2)}$ |
| | 0.7 | $0.095_{(5)}$ | $0.148_{(4)}$ | $0.155_{(3)}$ | $0.175_{(1)}$ | $0.175_{(1)}$ |
| | 0.9 | $0.094_{(5)}$ | $0.110_{(4)}$ | $0.119_{(3)}$ | $0.162_{(1)}$ | $0.161_{(2)}$ |
| 50 | 0.0 | $0.271_{(5)}$ | $0.280_{(2)}$ | $0.286_{(1)}$ | $0.280_{(2)}$ | $0.279_{(4)}$ |
| | 0.3 | $0.202_{(5)}$ | $0.248_{(4)}$ | $0.253_{(3)}$ | $0.256_{(2)}$ | $0.257_{(1)}$ |
| | 0.5 | $0.174_{(5)}$ | $0.237_{(4)}$ | $0.246_{(2)}$ | $0.248_{(1)}$ | $0.244_{(3)}$ |
| | 0.7 | $0.178_{(5)}$ | $0.218_{(4)}$ | $0.227_{(3)}$ | $0.246_{(1)}$ | $0.243_{(2)}$ |
| | 0.9 | $0.176_{(4)}$ | $0.172_{(5)}$ | $0.183_{(3)}$ | $0.228_{(1)}$ | $0.226_{(2)}$ |

**หมายเหตุ** (1), (2), (3), (4), (5) หมายถึง มีกำลังการทดสอบมากเป็นอันดับที่ 1, 2, 3, 4, 5 ตามลำดับ

จากตารางที่ 4 จะเห็นได้ว่า

ส่วนใหญ่วิธี BOOT มีกำลังการทดสอบเชิงประจักษ์สูงสุด รองลงมาคือ วิธี PERM ยกเว้นกรณีที่ $\rho$=0.0 เมื่อ n=10, 30 และ 50 วิธี HOM มีกำลังการทดสอบเชิงประจักษ์สูงสุด และเกือบทุกกรณีวิธี HOTE จะมีกำลังการทดสอบเชิงประจักษ์ต่ำสุด

**ตารางที่ 5** กำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 5

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | $0.072_{(5)}$ | $0.152_{(4)}$ | $0.158_{(3)}$ | $0.163_{(1)}$ | $0.160_{(2)}$ |
| | 0.3 | $0.050_{(5)}$ | $0.143_{(4)}$ | $0.148_{(2)}$ | $0.148_{(2)}$ | $0.150_{(1)}$ |
| | 0.5 | $0.047_{(5)}$ | $0.135_{(4)}$ | $0.137_{(3)}$ | $0.145_{(1)}$ | $0.142_{(2)}$ |
| | 0.7 | $0.048_{(5)}$ | $0.112_{(4)}$ | $0.113_{(3)}$ | $0.123_{(1)}$ | $0.123_{(1)}$ |
| | 0.9 | $0.040_{(5)}$ | $0.078_{(4)}$ | $0.092_{(3)}$ | $0.122_{(2)}$ | $0.123_{(1)}$ |
| 30 | 0.0 | $0.362_{(5)}$ | $0.455_{(3)}$ | $0.460_{(1)}$ | $0.457_{(2)}$ | $0.450_{(4)}$ |
| | 0.3 | $0.315_{(5)}$ | $0.423_{(4)}$ | $0.435_{(1)}$ | $0.430_{(3)}$ | $0.435_{(1)}$ |
| | 0.5 | $0.233_{(5)}$ | $0.330_{(4)}$ | $0.337_{(3)}$ | $0.348_{(1)}$ | $0.345_{(2)}$ |
| | 0.7 | $0.205_{(5)}$ | $0.313_{(4)}$ | $0.325_{(3)}$ | $0.358_{(1)}$ | $0.358_{(1)}$ |
| | 0.9 | $0.190_{(5)}$ | $0.273_{(4)}$ | $0.277_{(3)}$ | $0.352_{(1)}$ | $0.350_{(2)}$ |
| 50 | 0.0 | $0.655_{(5)}$ | $0.683_{(4)}$ | $0.690_{(2)}$ | $0.688_{(3)}$ | $0.692_{(1)}$ |
| | 0.3 | $0.535_{(5)}$ | $0.628_{(4)}$ | $0.632_{(1)}$ | $0.632_{(1)}$ | $0.632_{(1)}$ |
| | 0.5 | $0.428_{(5)}$ | $0.568_{(4)}$ | $0.587_{(2)}$ | $0.590_{(1)}$ | $0.585_{(3)}$ |
| | 0.7 | $0.371_{(5)}$ | $0.538_{(4)}$ | $0.547_{(3)}$ | $0.576_{(1)}$ | $0.571_{(2)}$ |
| | 0.9 | $0.437_{(5)}$ | $0.498_{(4)}$ | $0.507_{(3)}$ | $0.582_{(2)}$ | $0.583_{(1)}$ |

**หมายเหตุ** (1), (2), (3), (4), (5) หมายถึง มีกำลังการทดสอบมากเป็นอันดับที่ 1, 2, 3, 4, 5 ตามลำดับ

จากตารางที่ 5 จะเห็นได้ว่า ส่วนใหญ่วิธี BOOT มีกำลังการทดสอบเชิงประจักษ์สูงสุด รองลงมาคือ วิธี PERM ยกเว้นกรณีที่ $\rho$=0.0 เมื่อ n=30 และ $\rho$=0.3 เมื่อ n=30 และ 50 วิธี HOM มีกำลังการทดสอบเชิงประจักษ์สูงสุด

**ตารางที่ 6** กำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบทั้ง 5 วิธี จำแนกตามขนาดตัวอย่างและสัมประสิทธิ์สหสัมพันธ์ จำนวนตัวแปรตามเท่ากับ 7

| n | $\rho$ | HOTE | BON | HOM | BOOT | PERM |
|---|---|---|---|---|---|---|
| 10 | 0.0 | $0.133_{(5)}$ | $0.247_{(3)}$ | $0.247_{(3)}$ | $0.251_{(2)}$ | $0.253_{(1)}$ |
| | 0.3 | $0.073_{(5)}$ | $0.236_{(2)}$ | $0.236_{(2)}$ | $0.\ 247_{(1)}$ | $0.236_{(2)}$ |
| | 0.5 | $0.064_{(5)}$ | $0.205_{(4)}$ | $0.207_{(3)}$ | $0.215_{(1)}$ | $0.213_{(2)}$ |
| | 0.7 | $0.033_{(5)}$ | $0.200_{(4)}$ | $0.204_{(3)}$ | $0.213_{(2)}$ | $0.216_{(1)}$ |
| | 0.9 | $0.042_{(5)}$ | $0.144_{(4)}$ | $0.158_{(3)}$ | $0.201_{(2)}$ | $0.207_{(1)}$ |
| 30 | 0.0 | $0.702_{(5)}$ | $0.791_{(3)}$ | $0.800_{(1)}$ | $0.793_{(2)}$ | $0.791_{(3)}$ |
| | 0.3 | $0.578_{(5)}$ | $0.751_{(4)}$ | $0.762_{(1)}$ | $0.760_{(3)}$ | $0.762_{(1)}$ |
| | 0.5 | $0.449_{(5)}$ | $0.691_{(4)}$ | $0.696_{(3)}$ | $0.709_{(1)}$ | $0.700_{(2)}$ |
| | 0.7 | $0.342_{(5)}$ | $0.584_{(4)}$ | $0.602_{(3)}$ | $0.627_{(1)}$ | $0.629_{(2)}$ |
| | 0.9 | $0.311_{(5)}$ | $0.518_{(4)}$ | $0.529_{(3)}$ | $0.638_{(1)}$ | $0.633_{(2)}$ |
| 50 | 0.0 | $0.958_{(5)}$ | $0.973_{(4)}$ | $0.980_{(1)}$ | $0.976_{(3)}$ | $0.978_{(2)}$ |
| | 0.3 | $0.858_{(5)}$ | $0.913_{(3)}$ | $0.922_{(1)}$ | $0.913_{(3)}$ | $0.\ 916_{(2)}$ |
| | 0.5 | $0.767_{(5)}$ | $0.907_{(4)}$ | $0.909_{(3)}$ | $0.911_{(1)}$ | $0.911_{(1)}$ |
| | 0.7 | $0.678_{(5)}$ | $0.873_{(3)}$ | $0.873_{(3)}$ | $0.887_{(2)}$ | $0.893_{(1)}$ |
| | 0.9 | $0.700_{(5)}$ | $0.780_{(4)}$ | $0.787_{(3)}$ | $0.847_{(2)}$ | $0.851_{(1)}$ |

**หมายเหตุ** (1), (2), (3), (4), (5) หมายถึง มีกำลังการทดสอบมากเป็นอันดับที่ 1, 2, 3, 4, 5 ตามลำดับ

จากตารางที่ 6 จะเห็นได้ว่า ส่วนใหญ่วิธี PERM มีกำลังการทดสอบเชิงประจักษ์สูงสุด รองลงมาคือ วิธี BOOT ยกเว้นกรณีที่ $\rho$=0.0 เมื่อ n=30 และ 50 และ $\rho$=0.3 เมื่อ n=30 และ 50 วิธี HOM มีกำลังการทดสอบเชิงประจักษ์สูงสุด

## 5. สรุปผลการวิจัยและข้อเสนอแนะ

ขนาดตัวอย่างที่ศึกษาไม่มีผลต่อความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ของวิธีการทดสอบทั้ง 5 วิธี ถ้ามีจำนวนตัวแปรตามน้อยวิธีการทดสอบทั้ง 5 วิธี มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 มากกว่าเมื่อมีจำนวนตัวแปรตามมาก

ถ้ามีจำนวนตัวแปรตามมากแล้ววิธีโฮเทลลิงทีสแควร์มีอัตราความผิดพลาดแบบที่ 1 ของการทดลองส่วนใหญ่น้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนด วิธีบอนเฟอร์โรนี-โฮล์ม, วิธีเวสต์ฟอล-ยัง บูทสแตรป, วิธีการเรียงสับเปลี่ยนที่แท้จริง และวิธีการของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลักมีอัตราความผิดพลาดแบบที่ 1 ของการทดลองส่วนใหญ่อยู่ในช่วงของการทดสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 อย่างไรก็ตามวิธีโฮเทลลิงทีสแควร์จะมีอัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ให้ใกล้เคียง 0.05 ก็ต่อเมื่อมีจำนวนตัวแปรตามน้อยและสัมประสิทธิ์สหสัมพันธ์มีค่าต่ำนอกจากนี้ถ้ามีจำนวนตัวแปรตามน้อยและสัมประสิทธิ์สหสัมพันธ์มีค่าสูงแล้ว วิธีบอนเฟอร์โรนี-โฮล์มจะมีอัตราความผิดพลาดแบบที่ 1 ของการทดลองน้อยกว่าขอบเขตล่างของเกณฑ์ที่กำหนด

ส่วนกำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบทั้ง 5 วิธีนั้น แปรผันตามขนาดตัวอย่าง และจำนวนตัวแปร แต่แปรผกผันกับสัมประสิทธิ์สหสัมพันธ์ ภายใต้

สถานการณ์ส่วนใหญ่ วิธีเวสต์ฟอล-ยัง บูทสแตรปมีกำลังการทดสอบสูงสุด รองลงมาคือ วิธีการเรียงสับเปลี่ยนที่แท้จริง ส่วนวิธีโฮเทลลิงทีสแควร์มีกำลังการทดสอบต่ำสุดในทุกกรณี วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลักมีกำลังการทดสอบสูงสุดเมื่อสัมประสิทธิ์สหสัมพันธ์มีค่าต่ำ ณ ทุกระดับของขนาดตัวอย่างและจำนวนตัวแปรตาม

สำหรับการวิจัยครั้งต่อไปควรศึกษาประสิทธิภาพของวิธีการทั้ง 5 วิธีในกรณีที่ข้อมูลไม่เป็นไปตามข้อตกลงเบื้องต้น เช่น ข้อมูลไม่ได้มีการแจกแจงแบบปกติ หรือประชากรสองกลุ่มมีเมทริกซ์ความแปรปรวนและความแปรปรวนร่วมไม่เท่ากัน เป็นต้น นอกจากนั้นอาจศึกษาในกรณีที่มีจำนวนประชากรมากกว่าสองกลุ่ม หรือ ศึกษาประสิทธิภาพของการทดสอบพหุคูณแบบปิดที่ใช้สำหรับทดสอบความแตกต่างของสัดส่วนระหว่างประชากร หรือเมื่อตัวแปรตามมีมาตราวัดแบบจัดอันดับ

## 6. เอกสารอ้างอิง

1. Ji Zhang, PhD, Hui Quan, PhD, Jennifer Ng, SD, and Michael E. Stepanavage, MS., Some Statistical Methods for Multiple Endpoints in Clinical Trials.,Merck Research Laboratories, Clinical Biostatistics and Research Data Systems, p. 18, 1997.

2. Peter H. Westfall and Pussell D. Wolfinger, Closed Multiple Testing Procedures and Proc Multtest, http://support.sas.com/documentation/ periodicals/obs/obswww23/.,2000.

3. Simes, R.J. An unproved Bonferroni procedure for multiple tests of significance, Biometrika, Vol. 73 ; pp. 751-754, 1986.

4. Hommel, G., A stagewise rejective multiple test procedure based on a modified Bonferroni test, Biometrika, Vol. 75 ; pp. 383-386, 1988.

5. ศศิประภา หิริโอตป์, การเปรียบเทียบวิธีทดสอบสำหรับความแตกต่างระหว่างค่าเฉลี่ยสองประชากร บนพื้นฐานของตัวแปรตามพหุคูณ, วิทยานิพนธ์มหาบัณฑิต ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 186 น., 2540.

6. SAS Institute Inc. SAS Macro Language Courses Notes, 1996