การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิด เมื่อประยุกต์ใช้กับข้อมูลที่จัดลำดับ

A Comparative Study on the Efficiency of Closed Multiple Testing Applied to Ordinal Data

กมล บุษบา

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

kamon@mathstat.sci.tu.ac.th

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิด โดยใช้วิธีแบบขั้นบันไดในการทดสอบความแตกต่างระหว่างกลุ่ม 2 กลุ่มของข้อมูลที่จัดลำดับจำนวน 5 วิธี คือ วิธีโฮเทลลิงทีสแควร์ วิธีบอนเฟอร์โรนี - โฮล์ม วิธีเวสต์ฟอล - ยัง บูทสแตรป วิธีการเรียงสับเปลี่ยนที่แท้จริง และวิธีของโฮม-เมลที่ใช้การทดสอบของซิมส์เป็นหลัก โดยพิจารณาจากความสามารถในการควบคุมความน่าจะเป็นของคลาดเคลื่อนประเภทที่หนึ่ง และกำลังการทดสอบเชิงประจักษ์ เมื่อประชากรของ ตัวแปรพหุคูณมีการแจกแจงแบบพหุนามโดยมีค่าที่เป็นไปได้คือ 1 ถึง 5 หรือ 1 ถึง 9 และมีการแจกแจงแบบสมมาตรหรือแบบเบ้ซ้าย ภายใต้จำนวนตัวแปรตามเท่ากับ 3, 5 และ 7 ขนาดตัวอย่างเท่ากับ 10 และ 30 ที่ระดับนัยสำคัญของการทดสอบ 0.05 ข้อมูลที่ใช้ในการวิจัยได้จากการจำลองด้วยเทคนิคมอนติคาร์โล โดยกระทำซ้ำทั้งหมด 500 ครั้งในแต่ละสถานการณ์

ผลการวิจัยพบว่า ค่าที่เป็นไปได้ของลำดับและลักษณะการแจกแจงของประชากรมีผลต่อความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง และกำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบ ภายใต้สถานการณ์ส่วนใหญ่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลักมีแนวโน้มว่าจะเป็นวิธีที่มีประสิทธิภาพมากที่สุด นอกจากนี้วิธีการทดสอบเกือบทุกวิธียกเว้นวิธีโฮเทลลิงทีสแควร์ มีแนวโน้มที่จะมีประสิทธิภาพไม่แตกต่างกันเมื่อประชากรมีลักษณะการแจกแจงแบบเบ้ซ้าย ขนาดตัวอย่างมากขึ้นและจำนวนตัวแปรตามที่มากขึ้น

คำสำคัญ: การทดสอบพหุคูณแบบปิด วิธีโฮเทลลิงทีสแควร์ วิธีบอนเฟอร์โรนี-โฮล์ม วิธีเวสต์ฟอล-ยัง บูทสแตรป วิธีการเรียงสับเปลี่ยนที่แท้จริง วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก

Abstract

In this study we aim to compare the efficiency of five closed multiple test method i.e Hotelling's T^2, Bonferroni – Holm, Westfall – Young bootstrap, Exact Permutational and Hommel's Method Based on Simes' Test method, by using step - wise procedure for testing the difference between two groups of ranked data. We consider the capacities of controlling type I error rates and their powers when population have multinomial 1 to 5 or 1 to 9 and have symmetric or left skewed distribution under 3, 5 and 7 dependent variables, equal sample size 10 and 30, and level of significance is 0.05. Monte Carlo simulation was performed and repeated 500 times for each scenario.

The results of this study are as follows; The possible values of multinomial variable and the shape of distribution affected the capacity of controlling the type I error rates and their powers. In most situations, Hommel's Method Based on Simes' Test method tends to be the best efficiency method. Almost every methods except Hotelling's T^2 tends to be indifferent efficiency under left skewed distribution, large sample size, and large number of dependent variables.

Keyword: Closed Multiple Test, Hotelling's T^2, Bonferroni – Holm, Westfall – Young bootstrap, Exact Permutational and Hommel's

1. บทนำ

ในการศึกษาวิจัยโดยทั่วไป โดยเฉพาะอย่างยิ่งในงานวิจัยด้านสุขศาสตร์ วิทยาศาสตร์และเทคโนโลยี ส่วนใหญ่มักเกี่ยวข้องกับการเปรียบเทียบระหว่างวิธีการปฏิบัติ 2 วิธี เช่น การเปรียบเทียบระหว่างวิธีการรักษาแบบใหม่กับวิธีการรักษาแบบมาตรฐาน ซึ่งโดยปกติแล้วการเปรียบเทียบความแตกต่างของวิธีการรักษา 2 วิธีนี้ไม่ควรเปรียบเทียบโดยพิจารณาจากความแตกต่างของตัวแปรเพียงตัวเดียว แต่ควรวัดความแตกต่างด้วยตัวแปรหลายๆ ตัว ซึ่งน่าจะใช้วัดประสิทธิภาพของวิธีการรักษาได้ดีกว่า เช่น การศึกษาประสิทธิภาพของยาในการรักษาโรคหืดโดยเปรียบเทียบระหว่างกลุ่มที่รับยารักษากับกลุ่มที่ได้ยาลวง (placebo) โดยวัดผลจากการประเมินการหายใจออกในหนึ่งวินาที การวัดค่าหรือประเมินอัตราการหายใจออกสูงสุดต่อหนึ่งนาที และคะแนนที่แสดงระดับของอาการ หรือในการพัฒนาผลิตภัณฑ์ เช่นน้ำผักรวม พลาสเจอไรส์ ในขั้นตอนสุดท้าย สมมุติว่ามีสูตรน้ำผักรวมสองสูตรและต้อง การทำการทดสอบทางประสาทสัมผัส อาจให้ผู้ทดสอบทำการประเมินความชอบโดยพิจารณาจาก ความใส สี และกลิ่น (ผัก) ของผลิตภัณฑ์รวมทั้งความชอบโดยรวม ซึ่งในกรณีดังกล่าวนี้ นักวิจัยต้องการเปรียบเทียบผลของการใช้ยารักษาและยาลวงว่าให้ผลแตกต่างกันหรือไม่ มีผู้ประเมินด้านประสาทสัมผัสของน้ำผักรวมสองสูตรแตกต่างกันหรือไม่ วิธีการทางสถิติที่จะนำมาใช้กับสถานการณ์ในลักษณะนี้ควรเป็นวิธีที่คำนึงถึงตัวแปรที่เกี่ยวข้องมากกว่าหนึ่งตัวโดยตัวแปรทุกตัวที่พิจารณาจะร่วม กันมีผลต่อกระบวนการตัดสินใจ หรือใช้ตัวแปรหลายๆ ตัว

พร้อมกันเพื่อทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างวิธีการสองวิธี หรือ สูตรสองสูตร โดยคำนึงถึงความสัมพันธ์กันระหว่างตัวแปรตามที่บันทึกผล หรือคำนึงถึงสหสัมพันธ์ระหว่างตัวแปรตามเหล่านี้ด้วย นอกจากนั้นค่าที่ได้จากการสังเกตไม่จำเป็นต้องเป็นค่าที่แท้จริง แต่สามารถอยู่ในลักษณะของการจัดอันดับหรือมีมาตราวัดแบบอันดับมาตรา เช่น มีค่าเป็น 1, 2, 3, 4, 5 สำหรับมาตราวัดแบบลิเคอร์ท (Likert scale) ซึ่งเป็นที่นิยมใช้โดยทั่วไป หรือ 1, 2, 3,..., 9 สำหรับมาตราวัดแบบฮีดอนิก (Hedonic scale) ที่นิยมใช้ในการประเมินทางประสาทสัมผัส ในสาขาวิทยาศาสตร์การอาหาร เป็นต้น

ปัญหาในลักษณะดังกล่าวนั้น เรียกว่า "การทดสอบพหุคูณ (Multiple Testing)" ซึ่งเป็นการเปรียบเทียบความแตกต่างระหว่างกลุ่มสองกลุ่ม หรือมากกว่าสองกลุ่ม โดยพิจารณาจากตัวแปรตามหลายๆ ตัวพร้อมกัน ซึ่งต่างจาก "การเปรียบเทียบพหุคูณ (Multiple Comparisons)" ซึ่งโดยทั่วไปหมายถึงการเปรียบเทียบค่าเฉลี่ยของตัวแปรตามตัวแปรเดียวแต่มีหลายๆ กลุ่มที่ต้องการเปรียบเทียบซึ่งใช้บ่อยในการวิเคราะห์ความแปรปรวน นอกจากนี้ ปัญหาข้างต้นยังแตกต่างจากปัญหาในการทดสอบพหุคูณทั่วไป กล่าวคือ ตัวแปรตามที่วัดได้นั้นทุกๆ ตัวมีมาตราวัดแบบอันดับมาตรา

จากตัวอย่างของลักษณะปัญหาข้างต้นจะเห็นว่า จะต้องทำการทดสอบสมมุติฐานหลายสมมุติฐานด้วยกันและต้องทำการทดสอบสมมุติฐานเหล่านี้แบบพร้อมกัน(simultaneous tests) ซึ่งเป็นที่ทราบดีว่า ในการทดสอบสมมุติฐานทีละสมมุติฐานนั้นจะมีความผิดพลาดแบบที่หนึ่ง (Type I error) เกิดขึ้นด้วยความน่าจะเป็นค่าหนึ่งเกี่ยวข้องด้วยเสมอ ดังนั้นจุดมุ่งหมายของการใช้การทดสอบพหุคูณจึงเพื่อควบคุม "ค่ามากที่สุดของอัตราการเกิดความผิดพลาดแบบที่หนึ่งโดยรวม (maximum overall Type I error rate)" ซึ่งหมายถึงค่ามากที่สุดของความน่าจะเป็นที่สมมุติฐานว่างอย่างน้อยหนึ่งสมมุติฐานจะถูกปฏิเสธเมื่อจริงๆ แล้วสมมุติฐานว่างนั้นเป็นจริง อัตราความผิดพลาดดังกล่าวมีชื่อเรียกว่า "อัตราความผิดพลาดสูงสุดชนิดต่อการทดลอง (Maximum Experimentwise Error Rate: MEER)" หรือ "อัตราความผิดพลาดสูงสุดต่อวงศ์ (Maximum Familywise Error Rate: FWR)" [1]

ในกรณีที่มีตัวแปรตามที่ต้องการทดสอบ m ตัว จะมีเซตของสมมุติฐานว่างที่ต้องการทดสอบ คือ $H = \{H_{01},\ H_{02}, ..., H_{0m}\}$ การทดสอบพหุคูณจะควบคุม FWR ได้อย่างเข้มงวด ถ้าความน่าจะเป็นของความผิดพลาดแบบที่หนึ่งในการทดสอบอย่างน้อย 1 การทดสอบ มีค่าไม่เกินค่าระดับนัยสำคัญ α โดยไม่คำนึงถึงว่าจริงๆ แล้วมีสมมติฐานว่างใดบ้างในเซต H ที่เป็นจริง ทั้งการทดสอบพหุคูณและการเปรียบเทียบพหุคูณนั้นในปัจจุบันจะใช้ "การทดสอบแบบปิด (closed testing)" เข้ามาประยุกต์ใช้ ซึ่งมีทั้งการทดสอบแบบขั้นบันได (step-wise methods) และแบบขั้นตอนเดียว (single-step methods) ตัวอย่างของการทดสอบแบบขั้นบันไดที่ทราบกันอย่างแพร่หลาย ได้แก่ วิธีความแตก ต่างอย่างน้อยสุดที่มีนัยสำคัญ (Least Significant Difference : LSD) ของฟิชเชอร์ (Fisher) ซึ่งจะทำการทดสอบสมมุติฐานตัวแปรแต่ละตัวโดยใช้ตัวสถิติ t แบบพหุ (Multiple t-test) ที่ระดับนัยสำคัญ α ก็ต่อเมื่อมีการปฏิเสธสมมติฐานว่างในการวิเคราะห์ความแปรปรวนด้วยตัวสถิติ F ซึ่งจะเห็นได้ว่าวิธีนี้ควบคุม FWR ได้ไม่เข้มงวด เนื่องจากจะควบคุมได้เฉพาะกรณีที่ทุกค่าใน

$H = \{H_{01}, H_{02}, ..., H_{0m}\}$ เป็นจริงเท่านั้น ส่วนการทดสอบแบบขั้นตอนเดียวที่ทราบกันอย่างแพร่หลายวิธีการหนึ่ง และควบคุม FWR ได้อย่างเข้มงวด คือ วิธีบอนเฟอร์โรนี (Bonferroni) ซึ่งจะทำการทดสอบแต่ละสมมติฐานในเซต H ด้วยสถิติ t โดยใช้ระดับนัยสำคัญ α / m ได้เลยโดยไม่ต้องพิจารณาผลการทดสอบด้วยสถิติ F ในการวิเคราะห์ความแปรปรวน (Hochberg and Tamhane, 1987 อ้างใน [1]) อนึ่ง วิธีการแบบทีละขั้นโดยทั่วไปจะมีกำลังการทดสอบสูงกว่าวิธีแบบขั้นตอนเดียว [1]

การทดสอบพหุคูณโดยวิธีการทดสอบแบบปิดที่ใช้วิธีแบบขั้นบันไดนั้นมีหลายวิธี วิธีที่มีให้เลือกใช้ในโปรแกรมสำเร็จรูปทางสถิติ SAS เช่น วิธีบอนเฟอร์โรนี-โฮล์ม (Bonferoni-Holm minP) วิธีเวสต์ฟอล-ยัง บูทสแตรป (Westfall-Young Bootstrap minP) และวิธีการเรียงสับเปลี่ยนที่แท้จริง (Exact Permutational minP) เป็นต้น เนื่องจากวิธีเหล่านี้มีกำลังการทดสอบมากกว่าการใช้วิธีการทดสอบแบบขั้นตอนเดียว โดยสามารถควบคุม FWR ได้ โดย หทัยรัตน์และกมล ([2], [3]) ได้ศึกษาการควบคุมอัตราความผิดพลาดของวิธีการทดสอบพหุคูณแบบปิดโดยใช้วิธีแบบขั้นบันไดเมื่อประชากรทั้งสองกลุ่มมีการแจกแจงแบบปกติหลายตัวแปร พบว่าจำนวนตัวแปรตามมีผลต่อความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 และวิธีเวสต์ฟอล-ยัง บูทสแตรปและวิธีการเรียงสับเปลี่ยนที่แท้จริงมีอัตราความผิดพลาดแบบที่ 1 ของการทดลองอยู่ในช่วงของการทดสอบความสามารถในการควบคุมอัตราความผิดพลาดแบบที่ 1 มากที่สุด ส่วนการเปรียบเทียบกำลังการทดสอบนั้น เมื่อโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรตามเป็นแบบเท่ากัน ภายใต้สถานการณ์ส่วนใหญ่วิธีการเรียงสับเปลี่ยนที่แท้จริงมีกำลังการทดสอบสูงสุด เมื่อโครงสร้างของสหสัมพันธ์ระหว่างตัวแปรตามเป็นแบบไม่เท่ากันภายใต้สถานการณ์ส่วนใหญ่ วิธีเวสต์ฟอล-ยัง บูทสแตรปมีกำลังการทดสอบสูงสุด อย่างไรก็ตามยังไม่มีผู้เคยทำการศึกษาเปรียบเทียบประสิทธิภาพของการทดสอบพหุคูณแบบปิด ที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดอันดับ หรือมีมาตราวัดแบบอันดับมาตราซึ่งเป็นมาตราวัดที่นำ มาใช้บ่อยในทางปฏิบัติ ดังนั้นผู้วิจัยจึงสนใจเปรียบเทียบกำลังการทดสอบของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับ จำนวน 5 วิธี ได้แก่

1. วิธีโฮเทลลิงทีสแควร์ (Hotelling's T^2)
2. วิธีบอนเฟอร์โรนี-โฮล์ม (Bonferroni-Holm)
3. วิธีเวสต์ฟอล-ยัง บูทสแตรป (Westfall-Young Bootstrap)
4. วิธีเรียงสับเปลี่ยนที่แท้จริง (Exact Permutational)
5. วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (Hommel's Method Based on Simes'Test) โดยการสรุปผลว่า จะปฏิเสธสมมติฐานว่างหนึ่ง ๆ (H_{0k}) หรือไม่นั้นจะพิจารณาจากค่า p-value ที่ปรับแล้ว (adjusted p-value) ซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาสมมติฐานที่ทำการทดสอบตัวแปรที่ k

2. วัตถุประสงค์ของการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิดโดยใช้วิธีแบบขั้นบันไดในการทดสอบความแตกต่างของข้อมูลที่จัดลำดับระหว่างกลุ่มสองกลุ่ม จำนวน 5 วิธี

3. ขอบเขตของการวิจัย

3.1 ทำการเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิดวิธีต่าง ๆ โดยพิจารณาจากความ สามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง และกำลังการทดสอบเชิงประจักษ์แบบเฉลี่ย

3.2 ให้สมมติฐานของตัวแปรที่ k โดยที่ $k=1,2,..,m$ คือ

$$H_0^k : m_{1k} - m_{2k} = 0$$

$$H_1^k : m_{1k} - m_{2k} \neq 0$$

ให้สมมติฐานอินเตอร์เซกชันระหว่าง H^k โดยที่ $k=1,2,..,m$ เป็นดังตัวอย่างต่อไปนี้

$$H_0^{12} : [m_{11}, m_{12}]' - [m_{11}, m_{12}]' = \underline{0}$$

$$H_1^{12} : [m_{11}, m_{12}]' - [m_{11}, m_{12}]' \neq \underline{0}$$

$$H_0^{123} : [m_{11}, m_{12}, m_{13}]' - [m_{11}, m_{12}, m_{13}]' = \underline{0}$$

$$H_1^{123} : [m_{11}, m_{12}, m_{13}]' - [m_{11}, m_{12}, m_{13}]' \neq \underline{0}$$

3.3 กำหนดระดับนัยสำคัญเท่ากับ 0.05

3.4 กำหนดให้ขนาดตัวอย่าง (n) เท่ากันทั้ง 2 กลุ่มประชากร เท่ากับ 10 หรือ 30 ในแต่ละกรณี

3.5 กำหนดจำนวนตัวแปรตาม (m) เท่ากับ 3, 5 หรือ 7 ตัว ในแต่ละกรณี

3.6 ตัวแปรตามมีค่าที่เป็นไปได้คือ ลำดับที่ 1 ถึง 5 หรือ ลำดับที่ 1 ถึง 9

3.7 กำหนดลักษณะของการแจกแจงเป็นแบบสมมาตรหรือ เบ้ซ้าย (กำหนดโดยระบุค่าความน่าจะเป็นของการเกิดค่าที่เป็นไปได้ที่ทำให้การแจกแจงมีลักษณะที่ต้องการ)

(1) กรณีที่ตัวแปรตามมีค่าที่เป็นไปได้ คือ ลำดับที่ 1 ถึง 5

การแจกแจงแบบสมมาตร กำหนดให้ $P_{(i)} = 0.2,\ i = 1,2,...,5$

การแจกแจงแบบเบ้ซ้าย กำหนดให้ $P_{(1)} = 0.05, P_{(2)} = 0.1, P_{(3)} = 0.15, P_{(4)} = 0.5,$ และ $P_{(5)} = 0.2$

(2) กรณีที่ตัวแปรตามมีค่าที่เป็นไปได้ คือ ลำดับที่ 1 ถึง 9

การแจกแจงแบบสมมาตร กำหนดให้ $P_{(i)} = \frac{1}{9}$, $i = 1,2,...,9$

การแจกแจงแบบเบ้ซ้าย กำหนดให้

$$P_{(1)} = \frac{0.1}{9}, P_{(2)} = \frac{0.2}{9}, P_{(3)} = \frac{0.3}{9}, P_{(4)} = \frac{0.4}{9}, P_{(5)} = \frac{1}{9}, P_{(6)} = \frac{2}{9}, P_{(7)} = \frac{3}{9}, P_{(8)} = \frac{1.5}{9}, P_{(9)} = \frac{0.5}{9}$$

3.8 กำหนดให้ $\Delta_k = m_{1k} - m_{2k}$

(1) กรณีพิจารณาความน่าจะเป็นของความคลาด เคลื่อนประเภทที่ 1

$\Delta_k = 0$ สำหรับทุกค่า k

(2) กรณีพิจารณากำลังการทดสอบ กำหนดการแจกแจงของประชากรกลุ่มที่สองให้มีลักษณะเบ้ขวา จะได้ว่า

- กรณีที่ตัวแปรตามมีค่าที่เป็นไปได้ คือ ลำดับที่ 1 ถึง 5 กำหนดให้

$$P_{(1)} = 0.2, P_{(2)} = 0.5, P_{(3)} = 0.15, P_{(4)} = 0.1, P_{(5)} = 0.05$$

ทำให้ได้ว่า $\Delta_k = 0.7$ สำหรับกรณีสมมาตร

และ $\Delta_k = 1.4$ สำหรับกรณีเบ้ซ้าย

- กรณีที่ตัวแปรตามมีค่าที่เป็นไปได้ คือ ลำดับที่ 1 ถึง 9 กำหนดให้

$$P_{(1)} = \frac{0.5}{9}, P_{(2)} = \frac{1.5}{9}, P_{(3)} = \frac{3}{9}, P_{(4)} = \frac{2}{9}, P_{(5)} = \frac{1}{9}, P_{(6)} = \frac{0.4}{9}, P_{(7)} = \frac{0.3}{9}, P_{(8)} = \frac{0.2}{9}, P_{(9)} = \frac{0.1}{9}$$

ทำให้ได้ว่า $\Delta_k = 1.38889$ สำหรับกรณีสมมาตร

และ $\Delta_k = 2.77778$ สำหรับกรณีเบ้ซ้าย

3.9 การวิจัยนี้จะจำลองสถานการณ์ต่างๆ โดยใช้เทคนิคการจำลองแบบมอนติคาร์โล โดยใช้โปรแกรม SAS รุ่น 8.0 และทำการจำลองซ้ำ 500 รอบ ในแต่ละสถานการณ์

4. สถิติโฮเทลลิงทีสแควร์ (Hotelling's T^2)

ตัวสถิติหลักที่ใช้ในการทดสอบสมมติฐาน คือ

$$T^2 = [(\bar{X}_1 - \bar{X}_2) - 0]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) S_{pooled} \right]^{-1} [(\bar{X}_1 - \bar{X}_2) - 0]$$

โดยที่

$$\bar{X}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1} X_{1j} \quad , \quad S_1 = \frac{1}{n_1 - 1}\sum_{j=1}^{n_1}(X_{1j} - \bar{X}_1)(X_{1j} - \bar{X}_1)'$$

และ $$S_{pooled} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \qquad [4]$$

5. วิธีการดำเนินการวิจัย

5.1 วิธีการทดสอบแบบปิดสำหรับการทดสอบพหุคูณ

ถ้าหากต้องการทดสอบสมมติฐาน H_{01}, H_{02} และ H_{03} โดยที่ H_{0k} คือ สมมติฐานของการเปรียบเทียบวิธีการปฏิบัติของตัวแปรที่ k ซึ่งอาจจะเป็นการเปรียบเทียบวิธีการปฏิบัติ 3 กลุ่มกับกลุ่มควบคุม 1 กลุ่ม หรือเปรียบเทียบวิธีการปฏิบัติ 1 กลุ่ม กับกลุ่มควบคุม 1 กลุ่ม วิธีการทดสอบแบบปิดมีขั้นตอนดังต่อไปนี้

1. ในการทดสอบแต่ละสมมติฐาน H_{01}, H_{02} และ H_{03} จะใช้ระดับนัยสำคัญของการทดสอบ (α) ที่เหมาะสม

2. สร้างเซตของอินเตอร์เซกชันที่เป็นไปได้ทั้งหมดระหว่าง H_{01}, H_{02} และ H_{03} ในกรณีนี้จะได้สมมติฐานคือ $H_{012}, H_{013}, H_{023}$ และ H_{0123}

3. การทดสอบแต่ละอินเตอร์เซกชันต้องใช้ระดับนัยสำคัญของการทดสอบที่เหมาะสม ซึ่งในการทดสอบอาจใช้ F-tests, Hotelling's T^2 หรือวิธีอื่นๆ ที่ใช้สำหรับการทดสอบสมมติฐานอินเตอร์เซกชัน ซึ่งวิธีการทดสอบแบบปิดแต่ละวิธีก็ให้ผลลัพธ์ที่แตกต่างกันไป

4. ในการทดสอบจะปฏิเสธสมมติฐาน H_{0k} โดยการควบคุมอัตราความผิดพลาดสูงสุดต่อการทดลอง (Maximum Experiment wise Error Rate: MEER) ที่เป็นไปตามเงื่อนไขทั้ง 2 ข้อ คือ

(1) การทดสอบ H_{0k} มีนัยสำคัญ

(2) การทดสอบทุก ๆ สมมติฐานอินเตอร์เซกชันต้องประกอบด้วย H_{0k} ที่มีนัยสำคัญ

และในการหาขอบเขตวิกฤตจะพิจารณาค่า p-value ที่ปรับแล้ว (adjusted p-value) ซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาค่า p-value ที่ได้จากการทดสอบสมมติฐานที่ประกอบด้วยตัวแปรที่ k ถ้าค่า p-value ที่ปรับแล้วน้อยกว่าระดับนัยสำคัญ α แล้วจะปฏิเสธสมมติฐานว่าง

5.2 วิธีโฮเทลลิงทีสแควร์

วิธีโฮเทลลิงทีสแควร์เป็นวิธีการทดสอบสมมติฐานเชิงพหุคูณที่ต้องหาค่า p-value ของทุกสมมติฐานเชิงเดี่ยวและสมมติฐานอินเตอร์เซกชัน ตัวอย่างเช่น ต้องการทดสอบค่าเฉลี่ยของตัวแปร 3 ตัว จะต้องหาค่า p-value ของสมมติฐาน $H_{01}, H_{02}, H_{03}, H_{012}, H_{013}, H_{023}, H_{0123}$ แล้วจึงไปหาขอบเขตวิกฤตของการทดสอบ โดยพิจารณาค่า p-value ที่ปรับแล้ว ซึ่งเป็นค่า p-value ที่มากที่สุดในบรรดาค่า p-value ที่ได้จากการทดสอบสมมติฐานที่

ประกอบด้วยตัวแปรที่ k ถ้าค่า p-value ที่ปรับแล้วน้อยกว่าระดับนัยสำคัญ α แล้วจะปฏิเสธสมมติฐานว่างเชิงเดี่ยว [1]

5.3 วิธีบอนเฟอร์โรนี-โฮล์ม

วิธีบอนเฟอร์โรนี-โฮล์มเป็นวิธีการทดสอบที่พิจารณาค่า p-value ที่มีค่าน้อยที่สุดของการทดสอบแต่ละตัวแปรกับ α / k^* เมื่อ k^* คือ จำนวนตัวแปรในสมมติฐานอินเตอร์เซกชันและ α คือ ระดับนัยสำคัญที่กำหนดขึ้น และจะปฏิเสธสมมติฐานเมื่อ $\min p \le \alpha / k^*$ โดยที่ $\min p$ คือ ค่า p-value ที่มีค่าน้อยที่สุดของการทดสอบแต่ละตัวแปร หรือเท่ากับว่าสามารถปฏิเสธสมมติฐานอินเตอร์เซกชัน เมื่อ $k^* \times \min p \le \alpha$ ดังนั้น $k^* \times \min p$ คือ ค่า p-value ของสมมติฐานอินเตอร์เซกชันโดยที่การทดสอบสมมติฐานเชิงเดี่ยวจะมีค่า p-value ที่ได้มาจากการทดสอบสมมติฐานด้วยตัวสถิติโฮเทลลิงที-สแควร์ [6] ตัวอย่างเช่น ถ้าค่า p-value ของสมมติฐาน H_{01} เท่ากับ 0.0982 และค่า p-value ของสมมติฐาน H_{02} เท่ากับ 0.0262 แล้วค่า p-value ของสมมติฐาน H_{012} จะเท่ากับ $2 \times 0.0262 = 0.0524$

5.4 วิธีเวสต์ฟอล-ยัง บูทสแตรป

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

(1) สุ่มตัวอย่างขนาด n_1 และ n_1 (แบบแทนที่) สำหรับ กลุ่มวิธีการปฏิบัติที่ 1 และ 2 ตามลำดับ จากข้อมูลดั้งเดิมที่รวมตัวอย่างของทั้ง 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 ที่คำนวณได้นี้ไม่จำเป็นจะต้องเรียง ลำดับเหมือนกับค่า P ดั้งเดิม ทำเช่นเดียวกันนี้ซ้ำๆ กัน N ครั้ง ซึ่งจะได้เวกเตอร์ขนาด m ของค่าบูทสแตรป P ทั้งหมด N เวกเตอร์

(3) ประมาณค่า P ซึ่งปรับค่าแล้วสำหรับตัวแปรที่สอดคล้องกับสมมติฐาน $H_{0[1]}$ (ซึ่งคือตัวแปรที่ให้ค่า P น้อยที่สุดจากข้อมูลดั้งเดิม) ได้จากสัดส่วนของเหตุการณ์ $\min_{1 \le \tilde{k} \le m} P^*_{[\tilde{k}]} \le P_{[1]}$ แทนด้วย $\tilde{P}^{adj}_{[1]}$ สำหรับตัวแปรที่สอดคล้องกับ $H_{0[2]}$ และสามารถประมาณค่าบูทสแตรปที่ปรับแล้ว $\tilde{P}^{adj}_{[2]}$ ได้จากสัดส่วนของเหตุการณ์ $\min_{2 \le \tilde{k} \le m} P^*_{[\tilde{k}]} \le P_{[2]}$ สำหรับสำหรับตัวแปรที่สอดคล้องกับ $H_{0[k]}$ สามารถประมาณค่าบูทสแตรปที่ปรับแล้ว $\tilde{P}^{adj}_{[k]}$ ได้จากสัดส่วนของเหตุการณ์ $\min_{k \le \tilde{k} \le m} P^*_{[\tilde{k}]} \le P_{[k]}$

(4) หลังจากคำนวณค่า P ที่ปรับค่าโดยใช้เทคนิคบูทสแตรป $\tilde{P}^{adj}_{[1]}, \tilde{P}^{adj}_{[2]}, \ldots, \tilde{P}^{adj}_{[m]}$ (ซึ่งไม่จำเป็นต้องเรียงตาม ลำดับ) 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 $\min_{1 \le \tilde{k} \le m} P^*_{[\tilde{k}]} \le P_{[k]}$ โดยที่ $k = 1, 2, \ldots, m$ [5]

5.5 วิธีเรียงสับเปลี่ยนที่แท้จริง

เป็นวิธีที่เหมือนกับวิธีเวสต์ฟอล - ยัง บูทสแตรป แต่ต่างกันตรงการสุ่มตัวอย่าง โดยวิธีการเรียงสับเปลี่ยนที่แท้จริงจะเป็นการสุ่มตัวอย่างแบบไม่แทนที่ [5]

5.6 วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก

ซิมส์ (Simes, 1986) [6] ได้เสนอวิธีการทดสอบสมมติฐานเชิงพหุคูณโดยพิจารณาค่า p-value ที่เรียงลำดับจากน้อยไปหามาก ซึ่งมีหลักการดังนี้ สมมติว่ามีสมมติฐาน m สมมติฐาน ให้ $P_{(1)} \le P_{(2)} \le \ldots \le P_{(m)}$ เป็นลำดับของค่า p-value ที่สมมติฐาน $H_{(0i)}$ มีค่า p-value เป็น $P_{(i)}$ โดยที่ $i = 1, 2, \ldots, m$ จะปฏิเสธสมมติฐานอินเตอร์เซกชันได้ถ้า $p_{(i)} \le i\alpha / k$ 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 [1] ตัวอย่างเช่น ถ้าค่า p-value ของสมมติฐาน $H_{(01)}$ เท่ากับ 0.0262 และค่า p-value ของสมมติฐาน $H_{(02)}$ เท่ากับ 0.0982 แล้วจะได้ว่าค่า p-value ของสมมติฐาน H_{012} มีค่าเท่ากับ $\min(2 \times 0.0262/1,\ 2 \times 0.0982/2 =$ 0.0524

เมื่อทำการทดสอบโดยวิธีของซิมส์แบบปิด จะสามารถสร้างผลสรุปเกี่ยวกับสมมติฐานเชิงเดี่ยวได้ ซึ่งผลที่ได้นี้เรียกว่าวิธีของโฮมเมล (Hommel's method, 1988) [7]

5.7 การจำลองข้อมูลแบบลำดับที่ของตัวแปรพหุ

การจำลองข้อมูลนั้นเริ่มต้นโดยใช้การแจกแจงแบบสม่ำเสมอเป็นพื้นฐาน โดยใช้โปรแกรม SAS และใช้คำสั่ง RANUNI(seed) เพื่อสร้างตัวแปรสุ่มแบบสม่ำเสมอในแต่ละตัวแปร จำนวน m ตัว โดยที่ตัวแปรแต่ละตัวเป็นอิสระกัน

สำหรับการแปลงข้อมูลให้อยู่ในรูปของลำดับที่ 1 ถึง 5 หรือ 1 ถึง 9 จะใช้คำสั่งในโปรแกรม SAS เช่น คำสั่ง IF ในการกำหนดลำดับที่ตามความน่าจะเป็นของการเกิดขึ้นของค่าที่เป็นไปได้แต่ละค่า

5.8 การหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง

จำลองข้อมูลแบบลำดับที่ศึกษาให้มีประชากร 2 กลุ่ม แล้วทดสอบสมมติฐานเพื่อหาค่า p-value จากนั้นหาค่า p-value ที่ปรับแล้ว แล้วหาความน่าจะเป็นของความคลาด เคลื่อนประเภทที่หนึ่ง จากสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่างเชิงพหุคูณ เมื่อสมมติฐานว่างเป็นจริง จากการทำซ้ำจำนวน 500 ครั้ง

5.9 การทดสอบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง

เปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งของตัวสถิติแต่ละตัวตามเกณฑ์ โดยใช้การทดสอบแบบปกติมาตรฐาน (z-test) ในการทดสอบสมมติฐานแบบสองด้าน ซึ่งตัวสถิติทดสอบสามารถควบคุมความน่าจะเป็นของความผิดพลาดแบบที่ 1 ที่ระดับนัยสำคัญ 0.05 ได้ถ้าอัตราความผิดพลาดแบบที่ 1 โดยเฉลี่ยมีค่าอยู่ในช่วง [0.031, 0.069]

5.10 การหากำลังการทดสอบเชิงประจักษ์

จำลองข้อมูลเช่นเดียวกับการหาค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง แต่ให้สมมติฐานว่างไม่เป็นจริง แล้วทดสอบสมมติฐานเพื่อหาค่า p-value จากนั้นหาค่า p-value ที่ปรับแล้ว แล้วหากำลังการทดสอบจากสัดส่วนของจำนวนครั้งที่ปฏิเสธสมมติฐานว่างเชิงพหุคูณ เมื่อสมมติฐานว่างไม่เป็นจริง จากการทำซ้ำจำนวน 500 ครั้ง

5.11 การเปรียบเทียบกำลังการทดสอบเชิงประจักษ์

พิจารณาเลือกวิธีการทดสอบพหุคูณแบบปิดจากวิธีการทดสอบที่ให้กำลังการทดสอบที่สูงที่สุดในแต่ละสถานการณ์

6. ผลการวิจัย

ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง ของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับ สรุปได้ในตารางที่ 1 ถึง 4 โดย แต่ละวิธีการจะใช้อักษรย่อดังนี้

T หมายถึง วิธีโฮเทลลิงทีสแควร์

B หมายถึง วิธีบอนเฟอร์โรนี-โฮล์ม

W หมายถึง วิธีเวสต์ฟอล-ยัง บูทสแตรป

P หมายถึง วิธีเรียงสับเปลี่ยนที่แท้จริง

H หมายถึง วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก

ตาราง 1 อัตราความคลาดคลื่อนประเภทที่หนึ่งโดยเฉลี่ยเมื่อตัวแปรตามมีค่า 1 ถึง 5 และประชากรมีการแจกแจงแบบสมมาตร ที่ระดับนัยสำคัญเท่ากับ .05

m	n	T	B	W	**P**	H
3	10	.056	.068	.066	**.060**	.068
	30	.028*	.042	.040	**.038**	.042
5	10	.016*	.054	.052	**.044**	.054
	30	.018*	.050	.048	**.044**	.052
7	10	.010*	.058	.054	**.048**	.029
	30	.012*	.072*	.072*	**.068**	.072*

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

ตาราง 2 อัตราความคลาดคลื่อนประเภทที่หนึ่งโดยเฉลี่ยเมื่อตัวแปรตามมีค่า 1 ถึง 5 และประชากรมีการแจกแจงแบบเบ้ซ้าย ที่ระดับนัยสำคัญเท่ากับ .05

m	n	T	**B**	**W**	P	**H**
3	10	.026*	**.048**	**.048**	.042	**.048**
	30	.022*	**.034**	**.034**	.024*	**.034**
5	10	.006*	**.042**	**.046**	.036	**.042**
	30	.026*	**.058**	**.060**	.058	**.058**
7	10	.006*	**.036**	**.044**	.038	**.036**
	30	.016*	**.050**	**.052**	.046	**.050**

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

ตาราง 3 อัตราความคลาดคลื่อนประเภทที่หนึ่งโดยเฉลี่ยเมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบสมมาตร ที่ระดับนัยสำคัญเท่ากับ .05

m	n	T	**B**	**W**	**P**	H
3	10	.038	**.054**	**.056**	**.044**	.056
	30	.038	**.064**	**.064**	**.064**	.064
5	10	.024*	**.066**	**.064**	**.056**	.070*
	30	.028*	**.048**	**.048**	**.046**	.048
7	10	.008*	**.056**	**.054**	**.044**	.056
	30	.010*	**.050**	**.048**	**.046**	.050

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

ตาราง 4 อัตราความคลาดคลื่อนประเภทที่หนึ่งโดยเฉลี่ยเมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบเบ้ซ้าย ที่ระดับนัยสำคัญเท่ากับ .05

m	*n*	T	**B**	**W**	**P**	**H**
3	10	.032	**.040**	**.048**	**.044**	**.040**
	30	.028*	**.034**	**.038**	**.032**	**.036**
5	10	.008*	**.032**	**.036**	**.034**	**.032**
	30	.022*	**.046**	**.048**	**.046**	**.046**
7	10	.006*	**.032**	**.044**	**.040**	**.032**
	30	.012*	**.044**	**.050**	**.050**	**.046**

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

จากตารางที่ 3 เมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบสมมาตร วิธีบอนเฟอร์โรนี-โฮล์ม (B) วิธีเวสต์ฟอล-ยัง บูทสแตรป (W) และวิธีเรียงสับเปลี่ยนที่แท้จริง (P) สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้ในทุกสถานการณ์ แต่เมื่อประชากรมีการแจกแจงแบบเบ้ซ้าย (ตารางที่ 2) วิธีการเกือบทุกวิธี ยกเว้นวิธีโฮเทลลิงทีสแควร์ (T) สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้ในทุกสถานการณ์

การเปรียบเทียบกำลังการทดสอบเชิงประจักษ์ของวิธีการแต่ละวิธี สรุปได้ดังตารางที่ 5 ถึง 8

ตาราง 5 กำลังการทดสอบเชิงประจักษ์เมื่อตัวแปรตามมีค่า 1 ถึง 5 และประชากรมีการแจกแจงแบบสมมาตร ที่ระดับนัยสำคัญเท่ากับ .05

m	*n*	T	B	W	P	H
3	10	.276	.296	.296	.254	**.306**
	30	.808*	.758	.762	.738	**.784**
5	10	.262*	.340	.342	.298	**.346**
	30	.918*	.850	.848	.834	**.872**
7	10	.208*	.396	.392	.366	**.400**
	30	.912*	.882*	.884*	**.878**	.904*

หมายเหตุ * หมายถึง ไม่สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้

ตาราง 6 กำลังการทดสอบเชิงประจักษ์เมื่อตัวแปรตามมีค่า 1 ถึง 5 และประชากรมีการแจกแจงแบบเบ้ซ้าย ที่ระดับนัยสำคัญเท่ากับ .05

m	*n*	T	B	W	P	H
3	10	.964*	.948	.948	.918	**.960**
	30	1.00*	1.00	1.00	1.00*	1.00
5	10	.986*	.980	.984	.968	**.988**
	30	1.00*	1.00	1.00	1.00	1.00
7	10	.984*	.986	.988	.980	**.990**
	30	1.00*	1.00	1.00	1.00	1.00

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

จากตารางที่ 5 เมื่อตัวแปรตามมีค่า 1 ถึง 5 และประชากรมีการแจกแจงแบบสมมาตร วิธีการที่ควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้และมีกำลังการทดสอบเชิงประจักษ์สูงสุดเกือบทุกสถานการณ์ ได้แก่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (H) ยกเว้นเมื่อจำนวนตัวแปรตาม 7 ตัวและขนาดตัวอย่างเท่ากับ 30 วิธีที่มีกำลังการทดสอบเชิงประจักษ์สูงสุดได้แก่ วิธีเรียงสับเปลี่ยนที่แท้จริง (P) แต่เมื่อประชากรมีการแจกแจงแบบเบ้ซ้าย (ตารางที่ 6) เฉพาะกรณีที่ขนาดตัวอย่างเท่ากับ 10 เท่านั้นที่สามารถเปรียบเทียบกำลังการทดสอบเชิงประจักษ์ได้ และวิธีที่มีกำลังการ

ทดสอบเชิงประจักษ์สูงสุดได้แก่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (H)

ตาราง 7 กำลังการทดสอบเชิงประจักษ์เมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบสมมาตร ที่ระดับนัยสำคัญเท่ากับ .05

m	*n*	T	B	W	P	H
3	10	.368	.392	.392	.362	**.396**
	30	**.950**	.902	.902	.898	.924
5	10	.370*	.448	**.460**	.428	.460*
	30	.962*	.958	.960	.954	**.962**
7	10	.258*	.458	**.474**	.438	.470
	30	.992*	.972	.978	.972	**.984**

หมายเหตุ * หมายถึง ไม่สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้

ตาราง 8 กำลังการทดสอบเชิงประจักษ์เมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบเบ้ซ้าย ที่ระดับนัยสำคัญเท่ากับ .05

m	*n*	T	B	W	P	H
3	10	**.996**	**.996**	**.996**	.994	**.996**
	30	1.00*	1.00	1.00	1.00	1.00
5	10	.998*	.998	.998	.998	**1.00**
	30	1.00*	1.00	1.00	1.00	1.00
7	10	1.00*	1.00	1.00	1.00	1.00
	30	1.00*	1.00	1.00	1.00	1.00

หมายเหตุ * หมายถึง ค่าที่ได้อยู่นอกช่วง [0.031, 0.069]

จากตารางที่ 7 เมื่อตัวแปรตามมีค่า 1 ถึง 9 และประชากรมีการแจกแจงแบบสมมาตร วิธีการที่สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งได้และมีกำลังการทดสอบเชิงประจักษ์สูงสุดนั้นขึ้นกับจำนวนตัวแปรตามและขนาดตัวอย่าง เช่น เมื่อขนาดตัวอย่างเท่ากับ 10 วิธีการที่มีกำลังการทดสอบเชิงประจักษ์สูงสุด กรณีมีตัวแปรตาม 3 ตัวได้แก่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (H) แต่เมื่อจำนวนตัวแปรตามเป็น 5 หรือ 7 ตัว วิธีที่มีกำลังการทดสอบเชิงประจักษ์สูงสุดได้แก่ วิธีเวสต์ฟอล-ยัง บูทสแตรป (W) แต่ถ้าขนาดตัวอย่างเากับ 30 วิธีการที่มีกำลังการทดสอบเชิงประจักษ์สูงสุด กรณีมีตัวแปรตาม 3 ตัว ได้แก่ วิธีโฮเทลลิงทีสแควร์ (T) แต่เมื่อจำนวนตัวแปรตามเป็น 5 หรือ 7 ตัว วิธีที่มีกำลังการทดสอบเชิงประจักษ์สูงสุดได้แก่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (H)

สำหรับกรณีที่ประชากรมีการแจกแจงแบบเบ้ซ้าย (ตารางที่ 8) เฉพาะกรณีที่ขนาดตัวอย่างเท่ากับ 10 และจำนวนตัวแปรตามเท่ากับ 3 หรือ 5 ตัวเท่านั้นที่สามารถเปรียบเทียบกำลังการทดสอบเชิงประจักษ์ได้ และวิธีที่มีกำลังการทดสอบเชิงประจักษ์สูงสุดได้แก่ วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก (H)

7. สรุปผลการวิจัย

จากการพิจารณาความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่งและกำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับทั้ง 5 วิธี ภายใต้สถานการณ์ที่ศึกษา สรุปได้ดังตารางที่ 9

ตาราง 9 วิธีการทดสอบที่มีประสิทธิภาพมากที่สุด จำแนกตามสถานการณ์ต่างๆ ที่ ระดับนัยสำคัญ .05

m	*n*	ตัวแปรตามมีค่าที่เป็นไปได้			
		1 ถึง 5		1 ถึง 9	
		สมมาตร	เบ้ซ้าย	สมมาตร	เบ้ซ้าย
3	10	H	H	H	ใกล้เคียง T,B,W,H
	30	H	ใกล้เคียง B,W,H	T	ใกล้เคียง B,W,P,H
5	10	H	H	W	H
	30	H	ใกล้เคียง B,W,P,H	H	ใกล้เคียง B,W,P,H
7	10	H	H	W	ใกล้เคียง B,W,P,H
	30	P	ใกล้เคียง B,W,P,H	H	ใกล้เคียง B,W,P,H

ข้อสังเกตที่ได้จากการศึกษา คือ

(1) ค่าที่เป็นไปได้ของตัวแปรตามที่อยู่ในรูปของลำดับที่ และลักษณะการแจกแจงของประชากรต่างก็มีผลต่อความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนประเภทที่หนึ่ง และกำลังการทดสอบเชิงประจักษ์ของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับ

(2) ภายใต้สถานการณ์ส่วนใหญ่ที่ศึกษา วิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลักมีแนวโน้มว่าจะเป็นวิธีที่มีประสิทธิภาพมากที่สุด

(3) วิธีบอนเฟอร์โรนี-โฮล์ม และวิธีเวสต์ฟอล-ยัง บูทสแตรป เป็นวิธีที่มีประสิทธิภาพไม่แตกต่างจากวิธีของโฮมเมลที่ใช้การทดสอบของซิมส์เป็นหลัก เมื่อประชากรมีลักษณะเบ้ซ้ายและขนาดตัวอย่างเท่ากับ 30 ทั้งค่าที่เป็นไปได้ของตัวแปรตามที่อยู่ระหว่าง 1 ถึง 5 และ 1 ถึง 9 ทุกค่าของจำนวนตัวแปรที่ศึกษา (3, 5 หรือ 7 ตัวแปร)

(4) วิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับเกือบทุกวิธียกเว้นวิธีโฮเทลลิงทีสแควร์ มีแนวโน้มที่จะมีประสิทธิภาพไม่แตกต่างกันเมื่อประชากรมีลักษณะการแจกแจงแบบเบ้ซ้าย ขนาดตัวอย่างมากขึ้นและจำนวนตัวแปรตามที่มากขึ้น

8. ข้อเสนอแนะ

จากการเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิดที่ใช้วิธีแบบขั้นบันไดเมื่อนำมาประยุกต์ใช้กับข้อมูลที่จัดลำดับ ภายใต้สถานการณ์ที่สนใจศึกษา ผู้วิจัยมีข้อเสนอแนะในการศึกษาเพื่อเปรียบเทียบประสิทธิภาพของแต่ละวิธีในการทดสอบความแตกต่างระหว่างประชากร บนพื้นฐานของตัวแปรพหุคูณ ดังนี้

1. ควรศึกษาในกรณีที่ขนาดตัวอย่างอยู่ระหว่าง 10 ถึง 30 และในกรณีที่ขนาดตัวอย่างมากกว่า 30 แต่กำหนดให้ความแตกต่างระหว่างประชากรสองกลุ่มมีค่าไม่มากนักเพื่อที่จะได้เปรียบเทียบกำลังการทดสอบเชิงประจักษ์ของแต่ละวิธีได้ชัดเจนยิ่งขึ้น

2. ควรศึกษาในกรณีที่ตัวแปรตามแต่ละตัวมีลักษณะการแจกแจงที่แตกต่างกันเล็กน้อย จนถึงการแจกแจงที่แตกต่างกันอย่างมาก

3. ควรศึกษาการทดสอบความแตกต่างระหว่างประชากรมากกว่าสองกลุ่ม

9. เอกสารอ้างอิง

[1] Westfall, P. H., and Wolfinger, R. D., Closed Multiple Testing Procedures and Proc MULTTEST,http://support.sas.com/kb/22/addl/fusion22950_1_multtest.pdf [September, 2208]

[2] หทัยรัตน์ วงศ์ชัยสุวัฒน์ และ กมล บุษบา, การควบคุมอัตราความผิดพลาดของวิธีการทดสอบพหุคูณแบบปิดสำหรับค่าเฉลี่ยประชากร, วารสารวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ 15(2), น. 22-32, 2550.

[3] หทัยรัตน์ วงศ์ชัยสุวัฒน์ และ กมล บุษบา, การเปรียบเทียบประสิทธิภาพของวิธีการทดสอบพหุคูณแบบปิดสำหรับค่าเฉลี่ยประชากร ภายใต้โครงสร้างของสหสัมพันธ์แบบไม่เท่ากัน, วารสารวิจัยและพัฒนา มจธ. 31(1), น. 201-213, 2551.

[4] Johnson R.A. and Wichern D.W. Applied Multivariate Statistical Analysis, Fourth Edition, Prentice-Hall, New Jersey, 1998.

[5] ศศิประภา หิริโอตปั์, การเปรียบเทียบวิธีทดสอบ สำหรับความแตกต่างระหว่างค่าเฉลี่ยสองประชากร บนพื้นฐานของตัวแปรตามพหุคูณ, วิทยานิพนธ์มหาบัณฑิต, ภาควิชาสถิติ, บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย, 186 น., 2540.

[6] Simes, R.J. An Unproved Bonferroni Procedure for Multiple Tests of Significance., Biometrika, Vol. 73, pp. 751-754, 1986.

[7] Hommel, G., A Stagewise Rejective Multiple test Procedure Based on a Modified Bonferroni Test., Biometrika, Vol. 75, pp. 383-386, 1988.