ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตรมหาบัณฑิต (สถิติ)

ปริญญา

สถิติ | สถิติ

สาขา | **ภาควิชา**

เรื่อง สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติ

Test Statistic for Multiple Comparison of Nonnormal Distribution

นามผู้วิจัย นางสาวสาลินี เขื่อนคำ

ได้พิจารณาเห็นชอบโดย

อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

(ผู้ช่วยศาสตราจารย์วินัย โพธิ์สุวรรณ, Ph.D.)

หัวหน้าภาควิชา

(รองศาสตราจารย์ประสิทธิ์ พยัคฆพงษ์, M.S.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญจนา ธีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ **เดือน** **พ.ศ.**

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิทยานิพนธ์

เรื่อง

สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติ

Test Statistic for Multiple Comparison of Nonnormal Distribution

โดย

นางสาวสาลินี เขื่อนคำ

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ)
พ.ศ. 2555

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สาลินี เขื่อนคำ 2555: สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติ ปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก: ผู้ช่วยศาสตราจารย์วินัย โพธิ์สุวรรณ์, Ph.D. 107 หน้า

งานวิจัยนี้เป็นการศึกษาสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติด้วยวิธีการทดสอบ 3 วิธี คือ วิธีการทดสอบด้วยสถิติทดสอบ Tukey-Karmer วิธีการทดสอบด้วยสถิติทดสอบ Baumgartner-Weiß-Schindler และวิธีการทดสอบด้วยสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner โดยข้อมูลที่ใช้ในการศึกษาครั้งนี้ได้จากการจำลองข้อมูลด้วยภาษาโปรแกรม R 2.13.1 ให้มีการแจกแจงแบบปกติ การแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล และทำการเปรียบเทียบพหุคูณด้วยโปรแกรม SAS 9.1 ทำซ้ำแต่ละสถานการณ์ 500 รอบ

ผลการศึกษาพบว่า กรณีที่ข้อมูลมีขนาดเท่ากันและในกรณีขนาดตัวอย่างไม่เท่ากันสถิติทดสอบ Baumgartner-Weiß-Schindler มีประสิทธิภาพในการควบคุมอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-comparison error rate: PCER) ดีกว่าสถิติทดสอบอื่น ส่วนสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner มีประสิทธิภาพในการควบคุมอัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) ดีกว่าสถิติทดสอบอื่น ยกเว้นกรณีที่ข้อมูลมีขนาดตัวอย่างไม่เท่ากันและมีการแจกแจงแบบปกติพบว่า สถิติทดสอบ Tukey-Karmer มีประสิทธิภาพในการควบคุม FWER ดีกว่าสถิติทดสอบอื่น และเมื่อจำนวนคู่การเปรียบเทียบเพิ่มขึ้นค่า FWER ของสถิติทดสอบทั้ง 3 วิธี มีแนวโน้มเพิ่มขึ้น สำหรับกรณีข้อมูลมีคู่ที่แตกต่างกันพบว่า ค่า PCER และค่า FWER ของสถิติทดสอบ Tukey-Karmer มีแนวโน้มเพิ่มขึ้นเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ส่วนค่า PCER และค่า FWER ของสถิติทดสอบ Baumgartner-Weiß-Schindler และสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner มีแนวโน้มลดลงเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ยกเว้นกรณีที่ข้อมูลมีการแจกแจงแบบเอกซ์โพเนนเชียลที่ค่า PCER และค่า FWER ของสถิติทดสอบทั้ง 2 วิธี มีแนวโน้มเพิ่มขึ้นเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ส่วนอำนาจการทดสอบพบว่า สถิติทดสอบทั้ง 3 วิธี มีอำนาจการทดสอบที่ค่อนข้างสูง

_____ _____ ___ / ___ / ____

ลายมือชื่อนิสิต ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Salinee Khuankham 2012: Test Statistic for Multiple Comparison of Nonnormal Distribution. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Assistant Professor Winai Bodhisuwan, Ph.D. 107 pages.

The purpose of this study was to investigate of the multiple comparison among three different methods: Tukey-Karmer, Baumgartner-Weiß-Schindler and Dwass-Steel and Critchlow-Fligner when data are taken from a nonnormal distribution. Data were simulated by using programming language R 2.13.1 and then data analysis has done by SAS 9.1 program. Each situation under specified conditions was repeated 500 times.

The result show that when sample size are equal and not equal, Baumgartner-Weiß-Schindler has per-comparison error rate (PCER) near 0.05 than another methods and Dwass-Steel and Critchlow-Fligner has familywise-comparison error rate (FWER) near 0.05 than another methods except when data are normal distribution and sample size are not equal Tukey-Karmer has FWER rate near 0.05 than another methods and this study was found that the FWER are direct variation with number of comparisons. When data has different means this study was found that PCER and FWER of Tukey-Karmer are reverse with number of pairs with different means, but Baumgartner-Weiß-Schindler and Dwass-Steel and Critchlow-Fligner are direct variation with number of pairs with different means except when data are exponential distribution PCER and FWER of Baumgartner-Weiß-Schindler and Dwass-Steel and Critchlow-Fligner are reverse. Power of test of all methods are roughly high.

_____ _____ ___ / ___ / ___

Student's signature Thesis Advisor's signature

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

กิตติกรรมประกาศ

วิทยานิพนธ์นี้สำเร็จได้ด้วยความช่วยเหลือของผู้ช่วยศาสตราจารย์ ดร.วินัย โพธิ์สุวรรณ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ที่ได้กรุณาให้คำปรึกษา คำแนะนำ ตรวจทานแก้ไขข้อบกพร่องต่าง ๆ เป็นอย่างดีมาโดยตลอด ขอขอบคุณรองศาสตราจารย์นิดา ชาญบรรยง ประธานการสอบ และผู้ช่วยศาสตราจารย์ ดร.เสาวณิต สุขภารังษี ผู้ทรงคุณวุฒิภายนอก ที่ได้พิจารณาตรวจทานแก้ไขข้อบกพร่องและให้คำแนะนำเพื่อให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น และขอขอบคุณคณาจารย์ภาควิชาสถิติทุกท่านที่ได้ประสิทธิ์ประสาทวิชาความรู้แก่ผู้วิจัยเป็นอย่างดี ผู้วิจัยขอขอบคุณด้วยความรู้สึกซาบซึ้งและสำนึกในพระคุณเป็นอย่างสูง

ขอน้อมรำลึกถึงครอบครัวที่เป็นสิ่งสำคัญยิ่งสำหรับผู้วิจัย ขอขอบคุณพี่ที่ให้กำลังใจ ให้คำปรึกษาและการสนับสนุนด้วยดีเสมอมา ขอกราบขอบพระคุณบิดา มารดาที่เป็นทั้งที่ปรึกษา และผู้สนับสนุนทั้งทุนทรัพย์และกำลังใจ ตลอดจนการอบรม เลี้ยงดูเป็นอย่างดีเสมอมา ประโยชน์อันใดที่ได้จากวิทยานิพนธ์ฉบับนี้ขอมอบแด่ครอบครัวที่มีส่วนสำคัญยิ่งที่ทำให้ผู้วิจัยมีทุกวันนี้

ขอขอบคุณเจ้าหน้าที่ภาควิชาสถิติที่ให้ความช่วยเหลือ แนะนำ และสนับสนุนการทำวิจัยครั้งนี้เป็นอย่างดี และขอขอบคุณเพื่อนนิสิตปริญญาโท ทุกคนที่ให้กำลังใจ คำแนะนำ และความช่วยเหลือด้วยดีเสมอมา

สาลินี เขื่อนคำ

เมษายน 2555

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(3)
คำอธิบายสัญลักษณ์และคำย่อ	(5)
คำนำ	1
วัตถุประสงค์	3
การตรวจเอกสาร	5
อุปกรณ์และวิธีการ	40
อุปกรณ์	40
วิธีการ	40
ผลและวิจารณ์	42
ผล	42
วิจารณ์	88
สรุปและข้อเสนอแนะ	89
สรุป	89
ข้อเสนอแนะ	91
เอกสารและสิ่งอ้างอิง	92
ภาคผนวก	95
ประวัติการศึกษาและการทำงาน	107

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญตาราง

ตารางที่		หน้า
1	แสดงความคลาดเคลื่อนที่เกี่ยวข้องกับการทดสอบสมมติฐาน	10
2	แสดงอัตราความคลาดเคลื่อนของจำนวนการทดสอบ เมื่อ $\alpha = .05$	11
3	แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	43
4	แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	46
5	แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	49
6	แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	52
7	แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	55
8	แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	58
9	แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	61
10	แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	64
11	แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	67
12	แสดงค่า PCER กรณีขนาดตัวอย่างไม่เท่ากัน	70
13	แสดงค่า FWER กรณีขนาดตัวอย่างไม่เท่ากัน	73
14	แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างไม่เท่ากัน	76
15	แสดงค่า PCER กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่	79
16	แสดงค่า FWER กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่	82
17	แสดงอำนาจการทดสอบ กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่	85

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ

ภาพที่		หน้า
1	แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	44
2	แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	47
3	แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	50
4	แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	53
5	แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	56
6	แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	59
7	แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม	62
8	แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม	65
9	แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม	68
10	แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน	71
11	แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน	74
12	แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน	77
13	แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่	80

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สารบัญภาพ (ต่อ)

ภาพที่ **หน้า**

14 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ 83

15 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ 86

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำอธิบายสัญลักษณ์และคำย่อ

TK แทน สถิติทดสอบ Tukey-Karmer

BWS แทน สถิติทดสอบ Baumgartner-Weiß-Schindler

DSCF แทน สถิติทดสอบ Dwass-Steel และ Critchlow-Fligner

PCER แทน อัตราความคลาดเคลื่อนต่อการเปรียบเทียบ

FWER แทน อัตราความคลาดเคลื่อนต่อการทดสอบ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติ

Test Statistic for Multiple Comparison of Nonnormal Distribution

คำนำ

การเปรียบเทียบเป็นหนึ่งในวิธีการที่พบได้ในการศึกษาวิจัยทั้งทางวิทยาศาสตร์และทางสังคมศาสตร์ ตัวอย่างเช่น การเปรียบเทียบผลการรักษาผู้ป่วยที่ได้รับการรักษาด้วยยาชนิดเก่ากับผู้ป่วยที่ได้รับการรักษาด้วยยาชนิดใหม่ หรือการเปรียบเทียบผลผลิตทางการเกษตรที่ได้จากพืชต่างสายพันธุ์กัน ซึ่งการศึกษาวิจัยที่ดีจะต้องมีการวางแผนการทดลอง การเก็บข้อมูล และการวิเคราะห์ที่ดี เพื่อให้ผลสรุปที่ได้มีความถูกต้อง และเป็นประโยชน์ต่อผู้วิจัยและผู้ที่สนใจศึกษาอย่างแท้จริง จึงจำเป็นต้องอาศัยเครื่องมือที่มีประสิทธิภาพ ซึ่งสถิติเป็นเครื่องมือหนึ่งที่มีบทบาทสำคัญในการวิจัย เนื่องจากเป็นกระบวนการที่มีขั้นตอนและระเบียบการที่น่าเชื่อถือ โดยการเลือกใช้วิธีการทางสถิตินั้นจำเป็นจะต้องเลือกใช้ให้เหมาะสมกับวัตถุประสงค์และข้อตกลงเบื้องต้นของสถิติทดสอบแต่ละวิธี จึงจะทำให้ผลสรุปที่ได้มีความถูกต้อง แม่นยำ

เนื่องจากการศึกษาวิจัยบางครั้งไม่สามารถศึกษากับประชากรทุกหน่วยได้ จึงต้องอาศัยการศึกษาจากหน่วยตัวอย่าง และทำการวิเคราะห์ด้วยวิธีการทางสถิติเพื่ออ้างอิงไปสู่ค่าพารามิเตอร์ของประชากร โดยหลายครั้งผู้วิจัยสนใจเปรียบเทียบค่าเฉลี่ยอิทธิพลของปัจจัยที่มีต่อหน่วยทดลองเพื่อหาข้อสรุปเกี่ยวกับอิทธิพลของปัจจัยที่มีต่อประชากรว่ามีความแตกต่างกันอย่างไร การเปรียบเทียบค่าเฉลี่ยระหว่างปัจจัยตั้งแต่สามกลุ่ม ($k \geq 3$) โดยทั่วไปจะใช้การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) และหากผลการทดสอบปฏิเสธสมมติฐานหลัก (H_0) แสดงว่า มีอย่างน้อยหนึ่งกลุ่มมีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น แต่จะไม่สามารถสรุปได้ว่าประชากรกลุ่มใด หรือปัจจัยใดที่มีอิทธิพลต่างจากกลุ่มอื่น หากต้องการรายละเอียดว่ามีกลุ่มใดบ้างที่มีความแตกต่างจากกลุ่มอื่นจะต้องทำการเปรียบเทียบรายคู่ (Pairwise Comparison) ในขั้นตอนต่อไปด้วยการเปรียบเทียบพหุคูณ (Multiple Comparison) โดยอาจเลือกเปรียบเทียบเฉพาะคู่ที่สนใจศึกษาหรือเปรียบเทียบทุกคู่ ซึ่งมีจำนวนการเปรียบเทียบได้มากที่สุด $\binom{k}{2}$ คู่ วิธีการหนึ่งที่สามารถทำได้คือการเปรียบเทียบรายคู่ด้วยการทดสอบด้วยสถิติที (t-test) แบบสองทางที่ระดับนัยสำคัญ α แต่ระดับนัยสำคัญจะเพิ่มขึ้นตามจำนวนครั้งของการทดสอบ (Hochberg and Tamhane, 1987)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบพหุคูณ เป็นการเปรียบเทียบค่าเฉลี่ยของประชากร k กลุ่ม โดยเป็นการเปรียบเทียบในรูปของผลรวมเชิงเส้นของค่าเฉลี่ยซึ่งเป็นความแตกต่างระหว่างค่าเฉลี่ย 2 ค่าของค่าเฉลี่ย k กลุ่ม หรือความแตกต่างระหว่างค่าเฉลี่ยกลุ่มหนึ่งกับค่าเฉลี่ยของกลุ่มอื่นที่เหลือ การเปรียบเทียบพหุคูณสามารถแบ่งได้เป็น 2 ประเภท (ฉัตรศิริ, 2554)ได้แก่

1. การเปรียบเทียบแบบวางแผน (Planed Comparison) เป็นการเปรียบเทียบที่อยู่บนพื้นฐานของทฤษฎีการวิจัยเบื้องต้นและการตั้งสมมติฐาน การเปรียบเทียบแบบวางแผนมักมีข้อจำกัดเกี่ยวกับจำนวนของการเปรียบเทียบ และการเปรียบเทียบไม่เป็นอิสระต่อกัน หากมีจำนวนการเปรียบเทียบมากจะทำให้ช่วงความเชื่อมั่นกว้างและอำนาจการทดสอบต่ำ

2. การเปรียบเทียบภายหลัง (Post hoc Comparison) เป็นการเปรียบเทียบที่ขึ้นอยู่กับการวิเคราะห์ความแปรปรวน คือ ใช้การวิเคราะห์ความแปรปรวนเพื่อทดสอบความมีนัยสำคัญของค่าเฉลี่ยประชากร หากผลการทดสอบมีนัยสำคัญแสดงว่า มีอย่างน้อยหนึ่งกลุ่มมีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น จึงจะทำการเปรียบเทียบภายหลัง โดยการเปรียบเทียบภายหลังได้มีนักสถิติเสนอและพัฒนาขึ้นหลายวิธี ซึ่งส่วนใหญ่มีข้อตกลงเบื้องต้นเช่นเดียวกับการวิเคราะห์ความแปรปรวน ดังนี้

1. ตัวอย่างถูกสุ่มมากจากประชากรที่มีการแจกแจงแบบปกติหรือกลุ่มตัวอย่างมีขนาดใหญ่
2. ความแปรปรวนของประชากรเท่ากัน
3. ตัวอย่างถูกสุ่มอย่างอิสระ

ตัวอย่างสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ เช่น LSD, Tukey, Bonferroni, Dunnett, Duncan และ Scheffe และกรณีที่ความแปรปรวนของประชากรไม่เท่ากันมีสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ เช่น Tamhane's T2, Games-Howell, Brown-Forsythe, Dunnett's T3 และ Dunnett's C (ธีรศักดิ์, 2551) ส่วนกรณีที่ประชากรไม่ได้มีการแจกแจงแบบปกติทางเลือกหนึ่งในการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยอาจทำได้โดยการใช้สถิติทดสอบแบบไม่ใช้พารามิเตอร์ ซึ่งเป็นสถิติที่ไม่มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงแบบปกติของประชากร

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วัตถุประสงค์

1. เพื่อศึกษาผลการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติด้วยวิธีการทดสอบ 3 วิธี คือ วิธีการทดสอบด้วยสถิติทดสอบ Tukey-Karmer (TK) วิธีการทดสอบด้วยสถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และวิธีการทดสอบด้วยสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF)

2. เพื่อเปรียบเทียบอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ อัตราความคลาดเคลื่อนต่อการทดสอบ และอำนาจการทดสอบของการเปรียบเทียบพหุคูณด้วยวิธีการทดสอบ 3 วิธี คือ วิธีการทดสอบด้วยสถิติทดสอบ TK วิธีการทดสอบด้วยสถิติทดสอบ BWS และวิธีการทดสอบด้วยสถิติทดสอบ DSCF

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อให้ทราบความแตกต่างของผลการทดสอบเปรียบเทียบพหุคูณด้วยวิธีการทดสอบแบบใช้พารามิเตอร์ ซึ่งได้แก่ วิธีการทดสอบด้วยสถิติทดสอบ TK และวิธีการทดสอบแบบไม่ใช้พารามิเตอร์ ซึ่งได้แก่ วิธีการทดสอบด้วยสถิติทดสอบ BWS และวิธีการทดสอบด้วยสถิติทดสอบ DSCF

2. เพื่อเป็นแนวทางในการเลือกวิธีการเปรียบเทียบพหุคูณให้เหมาะสมกับลักษณะของข้อมูลเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติ

สมมติฐาน

1. ผลการเปรียบเทียบพหุคูณด้วยวิธีการทดสอบที่แตกต่างกันเมื่อตัวอย่างมีการแจกแจงที่แตกต่างกันจะให้จำนวนผลการปฏิเสธสมมติฐานหลักแตกต่างกัน

2. ผลการเปรียบเทียบพหุคูณด้วยวิธีการทดสอบที่แตกต่างกันเมื่อขนาดตัวอย่างแต่ละกลุ่มแตกต่างกันจะให้จำนวนผลการปฏิเสธสมมติฐานหลักแตกต่างกัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ขอบเขตการวิจัย

1. สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณที่ใช้ในการศึกษาครั้งนี้มี 3 วิธี คือ

1.1 สถิติทดสอบ Tukey-Karmer (TK) คือ สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบใช้พารามิเตอร์โดยมีข้อตกลงเบื้องต้น คือ ตัวอย่างถูกสุ่มอย่างอิสระ มีความแปรปรวนเท่ากัน และมีการแจกแจงแบบปกติ

1.2 สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) คือ สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบไม่ใช้พารามิเตอร์ ซึ่งไม่มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงแบบปกติ

1.3 สถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF) คือ สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบไม่ใช้พารามิเตอร์ ซึ่งไม่มีข้อตกลงเบื้องต้นเกี่ยวกับการแจกแจงแบบปกติ

2. ข้อมูลที่ใช้ในการศึกษาครั้งนี้ได้จากการจำลองสถานการณ์ด้วยภาษาโปรแกรม R เวอร์ชัน 2.13.1 (R Development Core Team, 2011) และทดสอบสมมติฐานด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณด้วยโปรแกรม SAS (Statistical Analysis System) เวอร์ชัน 9.1 โดยทำซ้ำในแต่ละสถานการณ์ 500 รอบ

3. จำนวนกลุ่มของตัวอย่างที่ใช้ในการศึกษาครั้งนี้แบ่งเป็น กรณีขนาดตัวอย่างเท่ากันโดยใช้จำนวนกลุ่มของตัวอย่างเท่ากับ 3 4 และ 5 กลุ่ม สำหรับกรณีขนาดตัวอย่างเท่ากันและกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ ศึกษาเฉพาะจำนวนกลุ่มของตัวอย่างเท่ากับ 3 กลุ่ม

4. ขนาดตัวอย่างที่ใช้ในการศึกษาครั้งนี้แบ่งเป็น กรณีขนาดตัวอย่างเท่ากันโดยใช้ตัวอย่างขนาด 5 7 และ 10 ส่วนกรณีขนาดตัวอย่างไม่เท่ากันใช้ตัวอย่างขนาด 5 ถึง 10 และกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ ศึกษาเฉพาะขนาดตัวอย่างเท่ากับ 10

5. การแจกแจงของตัวอย่างที่ใช้ในการศึกษาครั้งนี้ ได้แก่

5.1 การแจกแจงแบบปกติ

5.2 การแจกแจงแบบเบตา

5.3 การแจกแจงแบบไวบุลล์

5.4 การแจกแจงแบบเอกซ์โพเนนเชียล

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การตรวจเอกสาร

การตรวจเอกสารประกอบด้วย 5 ส่วน คือ ทฤษฎีที่เกี่ยวข้องกับการเปรียบเทียบพหุคูณ สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ สถิติที่เกี่ยวข้อง การแจกแจงที่ใช้ในการศึกษา และงานวิจัยที่เกี่ยวข้อง ซึ่งมีรายละเอียดดังต่อไปนี้

ทฤษฎีที่เกี่ยวข้องกับการเปรียบเทียบพหุคูณ

การเปรียบเทียบระหว่างค่าเฉลี่ย (Contrast Among Means)

การเปรียบเทียบ (contrast or comparison) ระหว่างค่าเฉลี่ยประชากร เป็นการเปรียบเทียบความแตกต่างของค่าเฉลี่ยประชากร เมื่อใช้สัญลักษณ์ทางพีชคณิตคือ ψ_i แทนความแตกต่างของค่าเฉลี่ยประชากรกลุ่มที่ i และ $\hat{\psi}_i$ แทนความแตกต่างของค่าประมาณค่าเฉลี่ยของตัวอย่างกลุ่มที่ i ตัวอย่างเช่น $\psi_i = \mu_j - \mu_{j'}$ เป็นความแตกต่างระหว่างค่าเฉลี่ยประชากรกลุ่มที่ j และ j' ส่วน $\hat{\psi}_i = \bar{Y}_{.j} - \bar{Y}_{.j'}$ แทนความแตกต่างของค่าประมาณค่าเฉลี่ยตัวอย่างกลุ่มที่ j และ j' ถ้าหากในการทดลองมีปัจจัยที่ใช้ในการทดลอง 3 ระดับ (k = 3) การเปรียบเทียบความแตกต่างของค่าเฉลี่ยจะเกี่ยวข้องกับค่าเฉลี่ย 2 หรือ 3 ตัว (Kirk, 1995) ตัวอย่างเช่น

$$\hat{\psi}_1 = \bar{Y}_{.1} - \bar{Y}_{.2} \qquad \hat{\psi}_4 = \frac{\bar{Y}_{.1} + \bar{Y}_{.2}}{2} - \bar{Y}_{.3}$$

$$\hat{\psi}_2 = \bar{Y}_{.1} - \bar{Y}_{.3} \qquad \hat{\psi}_5 = \frac{\bar{Y}_{.1} + \bar{Y}_{.3}}{2} - \bar{Y}_{.2}$$

$$\hat{\psi}_3 = \bar{Y}_{.2} - \bar{Y}_{.3} \qquad \hat{\psi}_6 = \frac{\bar{Y}_{.2} + \bar{Y}_{.3}}{2} - \bar{Y}_{.1} \qquad (1)$$

โดย $\hat{\psi}_1 - \hat{\psi}_3$ เป็นการเปรียบเทียบระหว่างค่าเฉลี่ยสองกลุ่ม ส่วน $\hat{\psi}_4 - \hat{\psi}_6$ เป็นการเปรียบเทียบที่ใช้ค่าเฉลี่ยของสองค่าเฉลี่ยกับอีกหนึ่งค่าเฉลี่ย ซึ่งสามารถใช้ในการเปรียบเทียบค่าเฉลี่ยของสองกลุ่มการทดลองกับกลุ่มควบคุม ส่วนใหญ่การเปรียบเทียบระหว่างค่าเฉลี่ยจะอยู่ในรูปของผลรวมเชิงเส้นที่มีค่าน้ำหนัก หรือค่าสัมประสิทธิ์ (coefficients) ใช้สัญลักษณ์คือ c_j แทนเมื่ออย่างน้อยมีหนึ่งค่าสัมประสิทธิ์ที่ไม่เท่ากับศูนย์ ($c_j \neq 0$) และผลรวมของค่าสัมประสิทธิ์เท่ากับศูนย์ ($\sum_{j=1}^{k} c_j = 0$) ตัวอย่างเช่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\psi_i = c_1\mu_1 + c_2\mu_2 + ... + c_k\mu_k$$

และ

$$\hat{\psi}_i = c_1\bar{Y}_{.1} + c_2\bar{Y}_{.2} + ... + c_k\bar{Y}_{.k}$$

ถ้า $c_j \neq 0$ ในบางกลุ่ม j และ $\sum_{j=1}^{k} c_j = 0$ การเปรียบเทียบ $\hat{\psi}_1 - \hat{\psi}_6$ จะสามารถเขียนในรูปของผลรวมเชิงเส้นของ $\bar{Y}_{.j}$ โดยการเลือกค่าสัมประสิทธิ์ เช่น

$$\hat{\psi}_i = c_1\bar{Y}_{.1} + c_2\bar{Y}_{.2} + c_3\bar{Y}_{.3}$$

$$\hat{\psi}_1 = (1)\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + (0)\bar{Y}_{.3} = \bar{Y}_{.1} - \bar{Y}_{.2}$$

$$\hat{\psi}_2 = (1)\bar{Y}_{.1} + (0)\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \bar{Y}_{.1} - \bar{Y}_{.3}$$

$$\hat{\psi}_3 = (0)\bar{Y}_{.1} + (1)\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \bar{Y}_{.2} - \bar{Y}_{.3}$$

$$\hat{\psi}_4 = \left(\frac{1}{2}\right)\bar{Y}_{.1} + \left(\frac{1}{2}\right)\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \frac{\bar{Y}_{.1} + \bar{Y}_{.2}}{2} - \bar{Y}_{.3}$$

$$\hat{\psi}_5 = \left(\frac{1}{2}\right)\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + \left(\frac{1}{2}\right)\bar{Y}_{.3} = \frac{\bar{Y}_{.1} + \bar{Y}_{.3}}{2} - \bar{Y}_{.2}$$

$$\hat{\psi}_6 = (-1)\bar{Y}_{.1} + \left(\frac{1}{2}\right)\bar{Y}_{.2} + \left(\frac{1}{2}\right)\bar{Y}_{.3} = \frac{\bar{Y}_{.2} + \bar{Y}_{.3}}{2} - \bar{Y}_{.1} \qquad (2)$$

เพื่อความสะดวกในการเปรียบเทียบค่าสัมประสิทธิ์ของแต่ละการเปรียบเทียบอาจเขียนในรูปของผลรวมของค่าสัมบูรณ์เท่ากับสอง นั่นคือ

$$\sum_{j=1}^{k} |c_j| = 2$$

เมื่อ $|c_j|$ เป็นค่าสัมประสิทธิ์สัมบูรณ์ และ $\sum_{j=1}^{k} |c_j| = 2$ ทุกการเปรียบเทียบ ตัวอย่างสำหรับ $\hat{\psi}_1$ และ $\hat{\psi}_6$ จะมีค่าสัมประสิทธิ์ ดังนี้

$$|1| + |-1| + |0| = 1 + 1 + 0 = 2$$

$$\left|\frac{1}{2}\right| + \left|\frac{1}{2}\right| + |-1| = \frac{1}{2} + \frac{1}{2} + 1 = 2$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบรายคู่และการเปรียบเทียบที่ไม่ใช่รายคู่

การเปรียบเทียบรายคู่ (Pairwise Comparison)

การเปรียบเทียบรายคู่ คือ การเปรียบเทียบที่ทุกค่าสัมประสิทธิ์ยกเว้นสัมประสิทธิ์ของสองกลุ่มที่ทำการเปรียบเทียบมีค่าเท่ากับศูนย์ จำนวนการเปรียบเทียบค่าเฉลี่ยที่เป็นไปได้ทั้งหมดสำหรับ k กลุ่ม เท่ากับ k(k-1)/2 ตัวอย่างเช่น เมื่อมีจำนวนกลุ่มที่ต้องการเปรียบเทียบค่าเฉลี่ย 3 กลุ่ม จะมีจำนวนคู่เปรียบเทียบทั้งหมดเท่ากับ 3(3-1)/2 = 3 คู่เปรียบเทียบ และสามารถเขียนในรูปผลบวกเชิงเส้นได้ ดังนี้

$$\hat{\psi}_1 = (1)\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + (0)\bar{Y}_{.3} = \bar{Y}_{.1} - \bar{Y}_{.2}$$

$$\hat{\psi}_2 = (1)\bar{Y}_{.1} + (0)\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \bar{Y}_{.1} - \bar{Y}_{.3}$$

$$\hat{\psi}_3 = (0)\bar{Y}_{.1} + (1)\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \bar{Y}_{.2} - \bar{Y}_{.3}$$

การเปรียบเทียบที่ไม่ใช่รายคู่ (Nonpairwise Comparison)

การเปรียบเทียบที่ไม่ใช่รายคู่ คือ การเปรียบเทียบค่าเฉลี่ยที่แต่ละครั้งของการเปรียบเทียบมีได้มากกว่า 2 กลุ่ม และจำนวนการเปรียบเทียบที่เป็นไปได้ทั้งหมดมีจำนวนมาก ตัวอย่างเช่น เมื่อมีจำนวนกลุ่มที่ต้องการเปรียบเทียบค่าเฉลี่ย 3 กลุ่ม อาจพิจารณาผลรวมเชิงเส้นได้ในรูป

$$\hat{\psi}_4 = \frac{1}{2}\bar{Y}_{.1} + \frac{1}{2}\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \frac{\bar{Y}_{.1} + \bar{Y}_{.2}}{2} - \bar{Y}_{.3}$$

$$\hat{\psi}_5 = \frac{1}{2}\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + \frac{1}{2}\bar{Y}_{.3} = \frac{\bar{Y}_{.1} + \bar{Y}_{.3}}{2} - \bar{Y}_{.2}$$

$$\hat{\psi}_6 = (-1)\bar{Y}_{.1} + \frac{1}{2}\bar{Y}_{.2} + \frac{1}{2}\bar{Y}_{.3} = \frac{\bar{Y}_{.2} + \bar{Y}_{.3}}{2} - \bar{Y}_{.1}$$

$$\hat{\psi}_7 = \frac{1}{3}\bar{Y}_{.1} + \frac{2}{3}\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \frac{(1)\bar{Y}_{.1} + (2)\bar{Y}_{.2}}{3} - \bar{Y}_{.3}$$

$$\hat{\psi}_8 = \frac{1}{4}\bar{Y}_{.1} + \frac{3}{4}\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \frac{(1)\bar{Y}_{.1} + (3)\bar{Y}_{.2}}{4} - \bar{Y}_{.3}$$

$$\hat{\psi}_9 = \frac{1}{5}\bar{Y}_{.1} + \frac{4}{5}\bar{Y}_{.2} + (-1)\bar{Y}_{.3} = \frac{(1)\bar{Y}_{.1} + (4)\bar{Y}_{.2}}{5} - \bar{Y}_{.3} \qquad (3)$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบอิสระ (Orthogonal Contrasts)

การเปรียบเทียบอิสระ คือ การเปรียบเทียบที่มีชุดของการเปรียบเทียบที่ไม่ซ้ำซ้อนและเป็นอิสระจากกันในแหล่งของความแปรปรวน หรืออาจกล่าวได้ว่า สารสนเทศที่ได้จากการเปรียบเทียบที่ไม่ซ้ำซ้อนกันและเป็นอิสระกัน เช่น พิจารณา $\hat{\psi}_2 = \bar{Y}_{.1} - \bar{Y}_{.3}$ และ $\hat{\psi}_3 = \bar{Y}_{.2} - \bar{Y}_{.3}$ จากสมการ (1) และ (2) ในรูป

$$\hat{\psi}_2 = \frac{1}{2}\hat{\psi}_1 + \hat{\psi}_4 = \left[\frac{1}{2}\left(\bar{Y}_{.1} - \bar{Y}_{.2}\right)\right] + \left[\frac{1}{2}\bar{Y}_{.1} + \frac{1}{2}\bar{Y}_{.2} - \bar{Y}_{.3}\right] = \bar{Y}_{.1} - \bar{Y}_{.3}$$

$$\hat{\psi}_3 = -\frac{1}{2}\hat{\psi}_1 + \hat{\psi}_4 = \left[\left(-\frac{1}{2}\right)\left(\bar{Y}_{.1} - \bar{Y}_{.2}\right)\right] + \left[\frac{1}{2}\bar{Y}_{.1} + \frac{1}{2}\bar{Y}_{.2} - \bar{Y}_{.3}\right] = \bar{Y}_{.2} - \bar{Y}_{.3}$$

พบว่า $\hat{\psi}_1$ และ $\hat{\psi}_4$ ไม่ได้ให้สารสนเทศเพิ่มเติมแก่ $\hat{\psi}_2$ และ $\hat{\psi}_3$ หรืออาจกล่าวได้ว่า $\hat{\psi}_1$ และ $\hat{\psi}_4$ มีค่าซ้ำซ้อนกับ $\hat{\psi}_2$ และ $\hat{\psi}_3$ เนื่องจากผลรวมเชิงเส้นของ $\hat{\psi}_2$ และ $\hat{\psi}_3$ สามารถเขียนในรูปผลรวมเชิงเส้นของ $\hat{\psi}_1$ และ $\hat{\psi}_4$ ได้ จึงไม่เป็นการเปรียบเทียบอิสระ

สำหรับการเปรียบเทียบที่ไม่ซ้ำซ้อน หรือการเปรียบเทียบอิสระ พิจารณาให้ $\hat{\psi}_i$ และ $\hat{\psi}_{i'}$ แทนการเปรียบเทียบชุดที่ i และ i' โดยมีค่าสัมประสิทธิ์คือ c_{ij} และ $c_{i'j}$ ตามลำดับ เมื่อ $j = 1,\ldots,k$ การเปรียบเทียบทั้งสองจะอิสระเมื่อ

$$\sum_{j=1}^{k} c_{ij}c_{i'j} = 0 \qquad \text{กรณีขนาดตัวอย่างเท่ากัน}$$

และ

$$\sum_{j=1}^{k} \frac{c_{ij}c_{i'j}}{n_j} = 0 \qquad \text{กรณีขนาดตัวอย่างไม่เท่ากัน}$$

นั่นคือ ความเป็นอิสระจะพิจารณาจากผลรวมของผลคูณระหว่างค่าสัมประสิทธิ์เท่ากับศูนย์ และกรณีที่จำนวนกลุ่มการเปรียบเทียบทั้งหมด k กลุ่ม ในการเปรียบเทียบหนึ่งจะมีการเปรียบเทียบอิสระได้มากที่สุด k-1 กลุ่ม ตัวอย่างเช่น กรณีขนาดตัวอย่างเท่ากัน พิจารณา $\hat{\psi}_1$ และ $\hat{\psi}_4$ จากสมการ (1) และ (2)

$$\hat{\psi}_1 = 1\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + 0\bar{Y}_{.3}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$\hat{\psi}_4 = (1/2)\bar{Y}_{.1} + (1/2)\bar{Y}_{.2} + (-1)\bar{Y}_{.3}$$

และ $$\sum_{j=1}^{k} c_{1j}c_{4j} = 1(1/2) + (-1)(1/2) + (0)(-1) = 0$$

เนื่องจาก $\hat{\psi}_1$ และ $\hat{\psi}_4$ มีผลรวมของผลคูณระหว่างค่าสัมประสิทธิ์เท่ากับศูนย์ ดังนั้นทั้งสองการเปรียบเทียบจึงเป็นอิสระกัน สำหรับกรณีของ $\hat{\psi}_1$ และ $\hat{\psi}_2$ จากสมการ (1) และ (2)

$$\hat{\psi}_1 = 1\bar{Y}_{.1} + (-1)\bar{Y}_{.2} + 0\bar{Y}_{.3}$$

$$\hat{\psi}_2 = 1\bar{Y}_{.1} + 0\bar{Y}_{.2} + (-1)\bar{Y}_{.3}$$

และ $$\sum_{j=1}^{k} c_{1j}c_{2j} = (1)(1) + (-1)(0) + (0)(-1) = 1$$

เนื่องจาก $\hat{\psi}_1$ และ $\hat{\psi}_2$ มีผลรวมของผลคูณระหว่างค่าสัมประสิทธิ์ไม่เท่ากับศูนย์ ดังนั้นทั้งสองการเปรียบเทียบจึงไม่เป็นอิสระกัน

ความคลาดเคลื่อนและอำนาจการทดสอบ

ในการทดสอบสมมติฐานอาจเกิดความคลาดเคลื่อนขึ้นได้ โดยจำแนกเป็น 2 ชนิด คือ ความคลาดเคลื่อนชนิดที่ 1 (Type I error) และความคลาดเคลื่อนชนิดที่ 2 (Type II error)

ความคลาดเคลื่อนชนิดที่ 1 (Type I error) คือ ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานหลัก (H_0) เมื่อสมมติฐานหลักเป็นจริง โดยกำหนดให้ α เป็นความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1

ความคลาดเคลื่อนชนิดที่ 2 (Type II error) คือ ความคลาดเคลื่อนที่เกิดจากการไม่ปฏิเสธสมมติฐานหลัก (H_0) เมื่อสมมติฐานหลักไม่เป็นจริง โดยกำหนดให้ β เป็นความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 2

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 1 แสดงความคลาดเคลื่อนที่เกี่ยวข้องกับการทดสอบสมมติฐาน

สรุปผลการทดสอบสมมติฐาน	สถานการณ์	
	สมมติฐานหลักถูกต้อง	สมมติฐานหลักไม่ถูกต้อง
ปฏิเสธสมมติฐานหลัก	ความคลาดเคลื่อนชนิดที่ 1	สรุปถูกต้อง
ไม่ปฏิเสธสมมติฐานหลัก	สรุปถูกต้อง	ความคลาดเคลื่อนชนิดที่ 2

อำนาจการทดสอบ คือ ความน่าจะเป็นหรือโอกาสที่จะปฏิเสธสมมติฐานหลัก (H_0) เมื่อสมมติฐานหลักไม่เป็นจริง ซึ่งมีค่าเท่ากับ $1-\beta$ โดยสถิติทดสอบที่มีอำนาจการทดสอบเข้าใกล้ 1 จะเป็นสถิติทดสอบที่มีอำนาจการทดสอบสูง

เมื่อขนาดตัวอย่างคงที่ค่า α และ β จะมีความสัมพันธ์ในทิศตรงข้าม คือ เมื่อลดขนาด α จะเป็นการเพิ่มขนาดของ β ในทางตรงข้าม หากเพิ่มขนาด α จะเป็นการลดขนาดของ β ซึ่งในทางปฏิบัติจะมีการควบคุมขนาด α และการเพิ่มขนาดตัวอย่างซึ่งจะช่วยลดโอกาสของการเกิดความคลาดเคลื่อนชนิดที่ 1 และความคลาดเคลื่อนชนิดที่ 2 ได้

อัตราความคลาดเคลื่อน (Error Rate)

ในการเปรียบเทียบปัจจัยที่มากกว่า 2 ปัจจัย ผู้วิจัยอาจสนใจทำการทดสอบสมมติฐานหลายครั้ง เช่น การเปรียบเทียบระดับของการใช้ยารักษาผู้ป่วย 4 ระดับ กับกลุ่มควบคุม ซึ่งจะทำการเปรียบเทียบได้ 4 รูปแบบ หรือการเปรียบเทียบผลผลิตทางการเกษตรที่ได้จากการใช้ปุ๋ยที่ระดับต่างกัน ซึ่งจำนวนการเปรียบเทียบจะขึ้นอยู่กับจำนวนปัจจัยและความสนใจในการเปรียบเทียบโดยอาจเลือกเปรียบเทียบเฉพาะปัจจัยที่สนใจศึกษา หรือเปรียบเทียบทุกปัจจัยซึ่งจะมีคู่เปรียบเทียบได้มากที่สุด $\binom{k}{2}$ คู่ โดยจำนวนการเปรียบเทียบส่งผลต่ออัตราความคลาดเคลื่อน ซึ่งอัตราความคลาดเคลื่อนที่เกี่ยวข้องกับการเปรียบเทียบพหุคูณ (อำไพ, 2554) แบ่งได้เป็น

1. อัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) คือ จำนวนความผิดพลาดที่เกิดจากการเปรียบเทียบต่อจำนวนการเปรียบเทียบทั้งหมด

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2. อัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) คือ อัตราความคลาดเคลื่อนที่เกิดจากการเปรียบเทียบต่อการทดสอบ หรือจำนวนความผิดพลาดจากการทดสอบสมมติฐานอย่างน้อยหนึ่งครั้งต่อจำนวนครั้งในการทดสอบ ซึ่ง FWER ≥ PCER (Hochberg and Tamhane, 1987)

ตารางที่ 2 แสดงอัตราความคลาดเคลื่อนของจำนวนการทดสอบ เมื่อ $\alpha = .05$

จำนวนการทดสอบ	PCER	FWER
1	0.05	0.05
5	0.05	0.23
10	0.05	0.40
20	0.05	0.64
50	0.05	0.92

ที่มา: Hochberg and Tamhane (1987)

จากตารางที่ 2 แสดงให้เห็นว่า เมื่อจำนวนการทดสอบเพิ่มขึ้นอัตราความคลาดเคลื่อนต่อการเปรียบเทียบจะไม่มีการเปลี่ยนแปลง แต่อัตราความคลาดเคลื่อนต่อการทดสอบจะเพิ่มขึ้นเมื่อจำนวนการทดสอบเพิ่มมากขึ้น

สถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ

สถิติทดสอบ Tukey–Karmer (TK)

เป็นสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบใช้พารามิเตอร์ที่มีพื้นฐานบนการแจกแจง studentized range statistic โดยมีข้อตกลงเบื้องต้น คือ ตัวอย่างถูกสุ่มมาอย่างอิสระ มีความแปรปรวนเท่ากัน และมีการแจกแจงแบบปกติ แบ่งเป็น 2 กรณี คือ

- กรณีขนาดตัวอย่างเท่ากันมีสูตรดังนี้

$$T = q_{\alpha;k,N-k}\sqrt{\frac{\text{MSE}}{n}}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เมื่อ $q_{\alpha;k,N-k}$ แทน ค่าวิกฤติที่ได้จากตาราง studentized range

MSE แทน ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

n แทน ขนาดตัวอย่าง

k แทน จำนวนกลุ่มตัวอย่าง

เกณฑ์การตัดสินจะปฏิเสธสมมติฐานหลักเมื่อ $|\bar{Y}_i - \bar{Y}_j| > T$

- กรณีขนาดตัวอย่างไม่เท่ากันมีสูตรดังนี้

$$T = q_{\alpha;k,N-k}\sqrt{\frac{\text{MSE}}{2}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

เมื่อ $q_{\alpha;k,N-k}$ แทน ค่าวิกฤติที่ได้จากตาราง studentized range

MSE แทน ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย

n_i, n_j แทน ขนาดตัวอย่างกลุ่มที่ i และ j ตามลำดับ

N แทน ขนาดตัวอย่างทั้งหมด

k แทน จำนวนกลุ่มตัวอย่าง

เกณฑ์การตัดสินจะปฏิเสธสมมติฐานหลักเมื่อ $|\bar{Y}_i - \bar{Y}_j| > T$

ตัวอย่างที่ 1 จากข้อมูลระยะห่างการวางไข่ของนกกาเหว่าในรัง 3 แบบ ดังต่อไปนี้

Tree pipit	Hedge-sparrow	Meadow pipit
21.1	20.9	19.6
21.8	21.7	20.1
22.1	22.0	20.6
22.4	22.8	21.6
22.7	23.0	21.9
23.2	23.1	22.2
23.3	23.5	22.3
23.4	23.8	22.5
23.6	23.9	22.6

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Tree pipit	Hedge-sparrow	Meadow pipit
24.0	25.0	22.9

ที่มา: Neuhauser and Bretz (2001)

สามารถเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Tukey-Karmer ได้ดังนี้

H_o : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit แบบ Hedge-sparrow และแบบ Meadow pipit ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit แบบ Hedge-sparrow และแบบ Meadow pipit อย่างน้อยหนึ่งค่าแตกต่างจากค่าอื่น

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

คำนวณค่าต่าง ๆ ได้ดังนี้

Tree pipit ผลรวม = 227.6 ค่าเฉลี่ย = 22.76

Hedge-sparrow ผลรวม = 229.7 ค่าเฉลี่ย = 22.97

Meadow pipit ผลรวม = 216.3 ค่าเฉลี่ย = 21.63

$$SSE = \sum Y_i^2 - \sum \frac{Y_{i.}^2}{n_i}$$

$$= (21.1^2 + 21.8^2 + \cdots + 22.6^2) - \left[\frac{227.6^2 + 229.7^2 + 216.3^2}{10}\right]$$

$$= 31.9060$$

$$MSE = \left(\frac{SSE}{N-k}\right)$$

$$= \frac{31.9060}{30-3}$$

$$= 1.1817037$$

ค่าวิกฤติจากตาราง $q_{0.05;3,27} = 3.50643$

$$T = 3.50643\sqrt{\frac{1.181704}{10}}$$

$$= 1.2054$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ผลการเปรียบเทียบรายคู่

Tree pipit และ Hedge-sparrow	| 22.76 – 22.97 |	=	0.21	< T
Tree pipit และ Meadow pipit	| 22.76 – 21.63 |	=	1.13	< T
Hedge-sparrow และ Meadow pipit	| 22.97 – 21.63 |	=	1.34	> T *

เนื่องจากค่าสัมประสิทธิ์ผลต่างค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Hedge-sparrow และรังแบบ Meadow pipit มากกว่าค่าวิกฤติ จึงปฏิเสธสมมติฐานหลัก ดังนั้นสรุปได้ว่า ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Hedge-sparrow และแบบ Meadow pipit แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS)

เป็นสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบไม่ใช้พารามิเตอร์ที่ดัดแปลงจากสถิติทดสอบ Wilcoxson rank sum test โดยการใช้สถิติ Baumgartner ในการคำนวณแทนสถิติ Wilcoxon มีสูตรดังนี้

$$B_{ij} = \frac{1}{2}\left(B_X^{ij} + B_Y^{ij}\right)$$

โดยที่

$$B_X^{ij} = \frac{1}{n_i}\sum_{r=1}^{n_i} \frac{\left(R_r - \frac{n_i + n_j}{n_i}.r\right)^2}{\frac{r}{n_i+1}\left(1-\frac{r}{n_i+1}\right).\frac{n_j(n_i+n_j)}{n_i}}$$

และ

$$B_Y^{ij} = \frac{1}{n_j}\sum_{s=1}^{n_j} \frac{\left(H_s - \frac{n_i + n_j}{n_j}.s\right)^2}{\frac{s}{n_j+1}\left(1-\frac{s}{n_j+1}\right).\frac{n_i(n_i+n_j)}{n_j}}$$

เมื่อ R_r แทน ลำดับร่วมของค่าสังเกตในกลุ่มที่ i โดย $R_1 < \ldots < R_{n_i}$

H_s แทน ลำดับร่วมของค่าสังเกตในกลุ่มที่ j โดย $H_1 < \ldots < H_{n_j}$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

n_i แทน ขนาดตัวอย่างในกลุ่มที่ i

n_j แทน ขนาดตัวอย่างในกลุ่มที่ j

เกณฑ์การตัดสินจะปฏิเสธสมมติฐานหลัก นั่นคือ $\tau_i \neq \tau_j$ เมื่อ $|B_{ij}| \geq b_\alpha$ โดย b_α หาได้จาก $\Pr_{H_0}\left(\max_{i<j}|B_{ij}| > b_\alpha\right) < \alpha$ โดยอาจใช้วิธี permutation test ซึ่งเป็นวิธีการหนึ่งของสถิติแบบไม่ใช้พารามิเตอร์

ตัวอย่างที่ 2 จากข้อมูลระยะห่างการวางไข่ของนกกาเหว่าในรัง 3 แบบ จากตัวอย่างที่ 1 สามารถเปรียบเทียบพหุคูณด้วยสถิติทดสอบ BWS ได้ดังนี้

การเปรียบเทียบ Tree pipit กับ Hedge-sparrow

H_0 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Hedge-sparrow ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Hedge-sparrow แตกต่างกัน

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

Tree pipit	ลำดับร่วม	B_X	Hedge-sparrow	ลำดับร่วม	B_Y
21.1	2	0.000000	20.9	1	0.060500
21.8	4	0.000000	21.7	3	0.033611
22.1	6	0.000000	22.0	5	0.025208
22.4	7	0.021607	22.8	9	0.021607
22.7	8	0.080667	23.0	10	0.000000
23.2	12	0.000000	23.1	11	0.020167
23.3	13	0.021607	23.5	15	0.021607
23.4	14	0.100833	23.8	17	0.025208
23.6	16	0.134444	23.9	18	0.000000
24.0	19	0.060500	25.0	20	0.000000
รวม	101	0.419659	**รวม**	109	0.207909

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำนวณค่า B_X^{ij} และ B_Y^{ij} ได้ดังนี้

$$B_X^{ij} = \frac{1}{n_i}\sum_{r=1}^{n_i} \frac{\left(R_r - \frac{n_i + n_j}{n_i}.r\right)^2}{\frac{r}{n_i+1}\left(1-\frac{r}{n_i+1}\right).\frac{n_j(n_i+n_j)}{n_i}}$$

$$B_X^{ij} = \frac{1}{10}\left[\frac{\left(2-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(19-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

$$= 0.0000000 + 0.000000 + \cdots + 0.060500$$

$$= 0.419659$$

$$B_Y^{ij} = \frac{1}{10}\left[\frac{\left(1-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(20-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

$$= 0.060500 + 0.033611 + \cdots + 0.000000$$

$$= 0.207909$$

สถิติทดสอบ BWS

$$B_{12} = \frac{1}{2}\left(B_X^{ij} + B_Y^{ij}\right)$$

$$= \frac{1}{2}(0.419659 + 0.207909)$$

$$= 0.313784$$

ค่าวิกฤติ คือ 3.497

เนื่องจากค่าสถิติทดสอบ BWS = 0.313784 ซึ่งน้อยกว่าค่าวิกฤติ จึงไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Tree pipit และรังแบบ Hedge-sparrow ไม่แตกต่างกันที่ระดับนัยสำคัญ .05

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบ Tree pipit กับ Meadow pipit

H_o : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Meadow pipit ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Meadow pipit แตกต่างกัน

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

Tree pipit	ลำดับร่วม	B_X	Meadow pipit	ลำดับร่วม	B_Y
21.1	4	0.242000	19.6	1	0.060500
21.8	6	0.134444	20.1	2	0.134444
22.1	8	0.100833	20.6	3	0.226875
22.4	11	0.194464	21.6	5	0.194464
22.7	14	0.322667	21.9	7	0.181500
23.2	16	0.322667	22.2	9	0.181500
23.3	17	0.194464	22.3	10	0.345714
23.4	18	0.100833	22.5	12	0.403333
23.6	19	0.033611	22.6	13	0.840278
24.0	20	0.000000	22.9	15	1.512500
รวม	133	1.645984	**รวม**	77	4.081109

คำนวณค่า B_X^{ij} และ B_Y^{ij} ได้ดังนี้

$$B_X^{ij} = \frac{1}{10}\left[\frac{\left(4-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(20-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

$$= 0.242000 + 0.134444 + \cdots + 0.00000$$

$$= 1.645984$$

$$B_Y^{ij} = \frac{1}{10}\left[\frac{\left(1-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(15-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

$$= 0.060500 + 0.134444 + \cdots + 1.512500$$

$$= 4.081109$$

สถิติทดสอบ BWS

$$B_{13} = \frac{1}{2}\left(B_X^{ij} + B_Y^{ij}\right)$$

$$= \frac{1}{2}(1.645984 + 4.081109)$$

$$= 2.863547$$

ค่าวิกฤติ คือ 3.497

เนื่องจาก ค่าสถิติทดสอบ BWS = 2.863547 ซึ่งน้อยกว่าค่าวิกฤติ จึงไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Tree pipit และรังแบบ Meadow pipit ไม่แตกต่างกันที่ระดับนัยสำคัญ .05

การเปรียบเทียบ Hedge-sparrow กับ Meadow pipit

H_o : ค่าเฉลี่ยระยะห่างการวางไข่ในรังแบบ Hedge-sparrow และแบบ Meadow pipit ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างการวางไข่ในรังแบบ Hedge-sparrow และแบบ Meadow pipit แตกต่างกัน

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

Hedge-sparrow	ลำดับร่วม	B_X	Meadow pipit	ลำดับร่วม	B_Y
20.9	4	0.242000	19.6	1	0.060500
21.7	6	0.134444	20.1	2	0.134444
22.0	8	0.100833	20.6	3	0.226875
22.8	13	0.540179	21.6	5	0.194464
23.0	15	0.504167	21.9	7	0.181500
23.1	16	0.322667	22.2	9	0.181500
23.5	17	0.194464	22.3	10	0.345714
23.8	18	0.100833	22.5	11	0.630208

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Hedge-sparrow	ลำดับร่วม	B_X	Meadow pipit	ลำดับร่วม	B_Y
23.9	19	0.033611	22.6	12	1.210000
25.0	20	0.000000	22.9	14	2.178000
รวม	136	2.173198	**รวม**	74	5.343206

คำนวณค่า B_X^{ij} และ B_Y^{ij} ได้ดังนี้

$$B_X^{ij} = \frac{1}{10}\left[\frac{\left(4-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(20-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

$$= 0.242000 + 0.134444 + \cdots + 0.00000$$

$$= 2.173198$$

$$B_Y^{ij} = \frac{1}{10}\left[\frac{\left(1-\frac{10+10}{10}.1\right)^2}{\frac{1}{10+1}\left(1-\frac{1}{10+1}\right).\frac{10(10+10)}{10}} + \cdots + \frac{\left(14-\frac{10+10}{10}.10\right)^2}{\frac{10}{10+1}\left(1-\frac{10}{10+1}\right).\frac{10(10+10)}{10}}\right]$$

$$= 0.060500 + 0.134444 + \cdots + 2.178000$$

$$= 5.343206$$

สถิติทดสอบ BWS

$$B_{12} = \frac{1}{2}\left(B_X^{ij} + B_Y^{ij}\right)$$

$$= \frac{1}{2}(2.173198 + 5.343206)$$

$$= 3.758202$$

ค่าวิกฤติ คือ 3.497

เนื่องจากค่าสถิติทดสอบ BWS = 3.758202 ซึ่งมากกว่าค่าวิกฤติ จึงปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Hedge-sparrow และรังแบบ Meadow-pipit แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบ Dwass - Steel และ Critchlow - Fligner (DSCF)

เป็นสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบไม่ใช้พารามิเตอร์ที่ใช้ค่า Wilcoxon มาประยุกต์ในการหาตัวสถิติทดสอบ โดยสถิติทดสอบ DSCF มีสูตรดังนี้

$$W_{ij} = \frac{S_{ij} - n_i(n_i + n_j + 1)/2}{\sqrt{n_i n_j (n_i + n_j + 1)/24}}$$

หรือ อาจกล่าวได้ว่าสถิติทดสอบ DSCF คือ ค่าสถิติ Wilcoxon คูณกับ $\sqrt{2}$

เมื่อ S_{ij} แทน ผลรวมลำดับร่วมของค่าสังเกตในกลุ่ม j

n_i แทน ขนาดตัวอย่างกลุ่ม i

n_j แทน ขนาดตัวอย่างกลุ่ม j

เกณฑ์การตัดสินจะปฏิเสธสมมติฐานหลัก นั่นคือ $\tau_i \neq \tau_j$ เมื่อ $|W_{ij}| \geq w_\alpha$ โดย w_α หาได้จาก $\Pr_{H_0}\left(\max_{i<j} |W_{ij}| > w_\alpha\right) < \alpha$ กรณีตัวอย่างขนาดเล็กและมีขนาดไม่เท่ากันสามารถหาค่าวิกฤติจากตารางของ Critchlow และ Fligner ปี 1991 ส่วนกรณีตัวอย่างขนาดใหญ่ เวคเตอร์ ($W^*_{12}, W^*_{13}, \ldots, W^*_{k-1,k}$) ข้อมูลจะมีการแจกแจงเข้าใกล้การแจกแจงแบบปกติหลายตัวแปร (Multivariate normal distribution) โดยมีเวคเตอร์ของตัวแปรเฉลี่ยคือ $\underset{\sim}{\theta} = (\theta, \theta, \ldots, \theta)$ เมื่อ $\min(n_1, n_2, \ldots, n_k) \to \infty$ แล้วอาจประมาณ w_α ด้วย q_α โดยที่ค่า q_α คือเปอร์เซ็นต์บน (upper percentage) สำหรับการแจกแจงของพิสัยที่มีการแจกแจงแบบปกติมาตรฐาน N(0,1) ที่เป็นอิสระ k ตัว

ตัวอย่างที่ 3 จากข้อมูลระยะห่างการวางไข่ของนกกาเหว่าในรัง 3 แบบ จากตัวอย่างที่ 1 สามารถเปรียบเทียบพหุคูณด้วยสถิติทดสอบ DSCF ได้ดังนี้

การเปรียบเทียบ Tree pipit กับ Hedge-sparrow

H_o : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Hedge-sparrow ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Hedge-sparrow แตกต่างกัน

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Tree pipit	ลำดับร่วม	Hedge-sparrow	ลำดับร่วม
21.1	2	20.9	1
21.8	4	21.7	3
22.1	6	22.0	5
22.4	7	22.8	9
22.7	8	23.0	10
23.2	12	23.1	11
23.3	13	23.5	15
23.4	14	23.8	17
23.6	16	23.9	18
24.0	19	25.0	20
รวม	101	**รวม**	109

สถิติทดสอบ DSCF

$$W_{12} = \frac{S_{12} - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/24}}$$

$$= \frac{101 - 10(10 + 10 + 1)/2}{\sqrt{(10)(10)(10 + 10 + 1)/24}}$$

$$= -0.4276$$

ค่าวิกฤติ คือ 3.314

เนื่องจากค่าสถิติทดสอบ DSCF = - 0.4276 ซึ่งน้อยกว่าค่าวิกฤติ จึงไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Tree pipit และรังแบบ Hedge-sparrow ไม่แตกต่างกันที่ระดับนัยสำคัญ .05

การเปรียบเทียบ Tree pipit กับ Meadow pipit

H_o : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Meadow pipit ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างของการวางไข่ในรังแบบ Tree pipit และแบบ Meadow pipit แตกต่างกัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

Tree pipit	ลำดับร่วม	Meadow pipit	ลำดับร่วม
21.1	4	19.6	1
21.8	6	20.1	2
22.1	8	20.6	3
22.4	11	21.6	5
22.7	14	21.9	7
23.2	16	22.2	9
23.3	17	22.3	10
23.4	18	22.5	12
23.6	19	22.6	13
24.0	20	22.9	15
รวม	133	**รวม**	77

สถิติทดสอบ DSCF

$$W_{13} = \frac{S_{12} - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/24}}$$

$$= \frac{133 - 10(10 + 10 + 1)/2}{\sqrt{(10)(10)(10 + 10 + 1)/24}}$$

$$= 2.993326$$

ค่าวิกฤติ คือ 3.314

เนื่องจากค่าสถิติทดสอบ DSCF = 2.993326 ซึ่งน้อยกว่าค่าวิกฤติ จึงไม่ปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Tree pipit และรังแบบ Meadow pipit ไม่แตกต่างกันที่ระดับนัยสำคัญ .05

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การเปรียบเทียบ Hedge-sparrow กับ Meadow pipit

H_o : ค่าเฉลี่ยระยะห่างการวางไข่ในรังแบบ Hedge-sparrow และแบบ Meadow pipit ไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยระยะห่างการวางไข่ในรังแบบ Hedge-sparrow และแบบ Meadow pipit แตกต่างกัน

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

Hedge-sparrow	ลำดับร่วม	Meadow pipit	ลำดับร่วม
20.9	4	19.6	1
21.7	6	20.1	2
22.0	8	20.6	3
22.8	13	21.6	5
23.0	15	21.9	7
23.1	16	22.2	9
23.5	17	22.3	10
23.8	18	22.5	11
23.9	19	22.6	12
25.0	20	22.9	14
รวม	136	**รวม**	74

สถิติทดสอบDSCF

$$W_{23} = \frac{S_{12} - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/24}}$$

$$= \frac{136 - 10(10 + 10 + 1)/2}{\sqrt{(10)(10)(10 + 10 + 1)/24}}$$

$$= 3.314039$$

ค่าวิกฤติ คือ 3.314

เนื่องจากค่าสถิติทดสอบ DSCF = 3.314039 ซึ่งมากกว่าค่าวิกฤติ จึงปฏิเสธสมมติฐานหลัก ดังนั้น ค่าเฉลี่ยระยะห่างการวางไข่ของนกกาเหว่าในรังแบบ Hedge-sparrow และรังแบบ Meadow pipit แตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ .05

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติที่เกี่ยวข้อง

สถิติทดสอบที (t-test)

เป็นสถิติทดสอบความแตกต่างของค่าเฉลี่ยระหว่างตัวอย่าง 2 กลุ่ม เมื่อขนาดตัวอย่างน้อยกว่า 30 หน่วย โดยมีสมมติฐานคือ

H_0 : ค่าเฉลี่ยของทั้ง 2 กลุ่มไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยของทั้ง 2 กลุ่มแตกต่างกัน

สามารถแบ่งเป็น 2 กรณี คือ

- กรณีความแปรปรวนเท่ากัน $(\sigma_i^2 = \sigma_j^2)$ จะมีองศาอิสระ$(\mathrm{df}) = n_i + n_j - 2$

$$t = \frac{\left(\overline{x}_i - \overline{x}_j\right) - \left(\mu_i - \mu_j\right)}{\sqrt{s_p^2\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}$$

โดยที่ $$s_p^2 = \frac{s_i^2(n_j - 1) + s_j^2(n_i - 1)}{n_i + n_j - 2}$$

เมื่อ t แทน สถิติทดสอบ t ที่ได้จากการคำนวณ

s_i^2, s_j^2 แทน ค่าประมาณความแปรปรวนของตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

$\overline{x}_i, \overline{x}_j$ แทน ค่าเฉลี่ยจากตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

μ_i, μ_j แทน ค่าคาดหวังของค่าเฉลี่ยตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

n_i, n_j แทน ขนาดตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

- กรณีความแปรปรวนไม่เท่ากัน ($\sigma_i^2 \neq \sigma_j^2$)

$$t = \frac{\left(\overline{x}_i - \overline{x}_j\right) - \left(\mu_i - \mu_j\right)}{\sqrt{\left(\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}\right)}}$$

โดยที่ $$\mathrm{df} = \frac{\left(\frac{s_i^2}{n_i} + \frac{s_j^2}{n_j}\right)^2}{\frac{(s_i^2/n_i)}{n_i - 1} + \frac{(s_j^2/n_j)}{n_j - 1}}$$

เมื่อ t แทน สถิติทดสอบ t ที่ได้จากการคำนวณ

s_i^2, s_j^2 แทน ค่าประมาณความแปรปรวนของตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

$\overline{x}_i, \overline{x}_j$ แทน ค่าเฉลี่ยจากตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

μ_i, μ_j แทน ค่าคาดหวังของค่าเฉลี่ยตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

n_i, n_j แทน ขนาดตัวอย่างกลุ่มที่ i และกลุ่มที่ j ตามลำดับ

เกณฑ์การทดสอบจะเปรียบเทียบค่าสถิติทดสอบ t ที่ได้จากการคำนวณกับค่าสถิติ $t_{\alpha,df}$ ที่ได้จากตารางสถิติ t ที่ระดับนัยสำคัญ α และองศาอิสระเท่ากับ df โดยจะปฏิเสธสมมติฐานหลัก เมื่อ $|t| \geq t_{\alpha,df}$ ซึ่งจะสรุปได้ว่า ค่าเฉลี่ยของทั้ง 2 กลุ่มแตกต่างกัน

สถิติทดสอบเอฟ (F-test)

ในการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของปัจจัยตั้งแต่ 3 ปัจจัยขึ้นไป จะใช้การวิเคราะห์ความแปรปรวน (Analysis of variance: ANOVA) ซึ่งเป็นการแบ่งความแปรปรวน หรือความผันแปรทั้งหมดของข้อมูลตามสาเหตุที่ทำให้ข้อมูลแตกต่างกัน เช่น เมื่อข้อมูลมีการจำแนกทางเดียวจะแบ่งความผันแปรออกเป็น 2 ส่วน คือ ความผันแปรภายในกลุ่ม (Within Groups Variance) และความผันแปรระหว่างกลุ่ม (Among Groups Variance) เมื่อสัดส่วนระหว่างค่าความผันแปร

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ระหว่างกลุ่มกับค่าความผันแปรภายในกลุ่มมาก จะแสดงว่า มีค่าเฉลี่ยของประชากรอย่างน้อยหนึ่งกลุ่มที่ต่างจากประชากรกลุ่มอื่น โดยในการวิเคราะห์ความแปรปรวนจะมีข้อตกลงเบื้องต้น คือ

1. ประชากรแต่ละกลุ่มมีการแจกแจงแบบปกติ
2. ความแปรปรวนของประชากรเท่ากัน
3. ตัวอย่างสุ่มจากประชากรที่เป็นอิสระกัน

จากข้อตกลงเบื้องต้นของการวิเคราะห์ความแปรปรวนนี้เป็นการกำหนดว่าประชากร k กลุ่มที่ศึกษามีการแจกแจงแบบปกติที่เป็นอิสระต่อกัน โดยที่ประชากรกลุ่มที่ i มีการแจกแจงแบบ $N(\mu_i, \sigma^2)$ และมีสมมติฐานการทดสอบ คือ

H_0 : ค่าเฉลี่ยของทุกกลุ่มไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยอย่างน้อย 1 กลุ่มแตกต่างจากกลุ่มอื่น

มีสูตรดังนี้

$$F = \frac{\sum_{j=1}^{j} n_j \left(\overline{X}_{.j} - \overline{X}_{..}\right)^2 / \left(k-1\right)}{\sum_{j=1}^{k}\sum_{i=1}^{n_i}\left(X_{ij} - \overline{X}_{.j}\right) / (N-k)}$$

เมื่อ F แทน สถิติทดสอบ F ที่ได้จากการคำนวณ

X_{ij} แทน ค่าสังเกตลำดับที่ i ที่ได้รับปัจจัย j

$\overline{X}_{.j}$ แทน ค่าเฉลี่ยของตัวอย่างที่ได้รับปัจจัย j

$\overline{X}_{..}$ แทน ค่าเฉลี่ยของตัวอย่างทั้งหมด

n_j แทน ขนาดตัวอย่างที่ได้รับปัจจัย j

N แทน ขนาดตัวอย่างทั้งหมด

k แทน จำนวนปัจจัย

เกณฑ์การทดสอบจะเปรียบเทียบค่า F ที่ได้จากการคำนวณกับค่าสถิติ $F_{(\alpha,k-1,k(n-1))}$ ที่ได้จากตารางการแจกแจง F ที่ระดับนัยสำคัญ α และมีองศาอิสระเท่ากับ $k-1, k(n-1)$ โดยจะปฏิเสธ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สมมติฐานเมื่อ F ที่ได้จากการคำนวณมากกว่าค่าสถิติ $F_{(\alpha,k-1,k(n-1))}$ หรือสรุปผลได้ว่าค่าเฉลี่ยอย่างน้อยหนึ่งกลุ่มแตกต่างจากกลุ่มอื่น

สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง (Anderson-Darling Test: AD)

สถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง เป็นสถิติสำหรับทดสอบการแจกแจงของข้อมูลที่พัฒนามาจากสถิติทดสอบ Kolmogorov-Smirnov (KS) ซึ่งมีค่าวิกฤติที่ไม่ขึ้นอยู่กับการแจกแจงของข้อมูลที่ต้องการทดสอบ แต่สำหรับสถิติทดสอบ AD ค่าวิกฤติจะขึ้นอยู่กับการแจกแจงที่ต้องการทดสอบและมีความอ่อนไหว (sensitive) ต่อการแจกแจงแบบหางยาว (Anderson and Darling, 1954) มีสมมติฐานการทดสอบคือ

H_0 : ข้อมูลมีการแจกแจงแบบที่ต้องการทดสอบ

H_1 : ข้อมูลไม่ได้มีการแจกแจงแบบที่ต้องการทดสอบ

โดยค่าสถิติทดสอบ AD หาได้จากสูตร

$$AD = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)[\log u_{(j)} + \log(1-u_{(n-j+1)})]$$

เมื่อ $u_{(j)}$ แทน ฟังก์ชันการแจกแจงที่คาดหมาย โดยที่ $u_{(j)} = F^0(x_{(j)})$ และ $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ เป็นค่าสังเกตในตัวอย่างลำดับ

เกณฑ์การตัดสินจะปฏิเสธสมมติฐานหลักเมื่อค่าสถิติที่ได้จากการคำนวณมากกว่าค่าวิกฤติซึ่งเป็นค่าวิกฤติเฉพาะของการแจกแจง

ตัวอย่างที่ 4 จำนวนการสั่งซื้อสินค้าของร้านค้าแห่งหนึ่งจำนวน 10 วัน คือ 94 100 112 117 121 125 126 128 149 และ 165 สามารถทดสอบการแจกแจงของจำนวนการสั่งซื้อสินค้าของร้านค้าแห่งนี้ว่ามีการแจกแจงแบบปกติหรือไม่ด้วยสถิติทดสอบ AD ได้ ดังนี้

ให้ X แทน จำนวนการสั่งซื้อสินค้า

H_0 : จำนวนการสั่งซื้อสินค้าของร้านค้าแห่งนี้มีการแจกแจงแบบปกติ

H_1 : จำนวนการสั่งซื้อสินค้าของร้านค้าแห่งนี้ไม่ได้มีการแจกแจงแบบปกติ

โดยกำหนดระดับนัยสำคัญของการทดสอบ $\alpha = .05$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

คำนวณค่าต่างๆ ได้ดังนี้

$$\bar{x} = \frac{94 + 100 + \cdots + 165}{10} = 123.7$$

$$s^2 = \frac{\sum_{j=1}^{10} (x_j - \bar{x})}{n-1}$$

$$= \frac{(94-123.7)^2 + (100-123.7)^2 + \cdots + (165-123.7)^2}{10-1} = 444.8999$$

ดังนั้น $s = 21.0927$

แปลงค่า x เป็นค่ามาตรฐานด้วย $Z = \frac{X_j - 123.4}{21.0927}$

ตัวอย่างการคำนวณค่า u_i

$$u_1 = P(X \le 94)$$

$$= P\left(Z \le \frac{94-123.7}{21.0927} \right)$$

$$= P(Z \le -1.41)$$

$$= 0.5 - 0.4207$$

$$= 0.0793$$

คำนวณค่าต่างๆ ได้ดังนี้

(j)	X_j	Z_j	$u_j=F_0(x_j)$	$\log u_j$
1	94	-1.41	0.0793	-2.5345
2	100	-1.12	0.1314	-2.0295
3	112	-0.55	0.2912	-1.2337
4	117	-0.32	0.3745	-0.9822
5	121	-0.13	0.4483	-0.8023
6	125	0.06	0.5239	-0.6465
7	126	0.11	0.5438	-0.6092

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

(j)	X_j	Z_j	$u_j=F_0(x_j)$	logu_j
8	128	0.20	0.5793	-0.5459
9	149	1.20	0.8849	-0.1223
10	165	1.96	0.9750	-0.0253

การคำนวณค่าต่างๆ เพื่อหาสถิติทดสอบ AD

(j)	u_{n-j+1}	1- u_{n-j+1}	log(1- u_{n-j+1})	(2j-1)(logu_j+ log(1- u_{n-j+1})
1	u_{10}	0.0250	-3.6889	-6.2234
2	u_9	0.1151	-2.1620	-12.5744
3	u_8	0.4207	-0.8658	-10.4979
4	u_7	0.4562	-0.7848	-12.3689
5	u_6	0.4761	-0.7421	-13.8998
6	u_5	0.5517	-0.5948	-13.6533
7	u_4	0.6255	-0.4692	-14.0189
8	u_3	0.7088	-0.3442	-13.3518
9	u_2	0.8686	-0.1409	-4.4736
10	u_1	0.9207	-0.0826	-2.0508
			รวม	-103.1127

$$AD = -10 - \frac{1}{10}\sum_{j=1}^{10}(2j-1)[\log u_{(j)} + \log(1-u_{(n-j+1)})]$$

$$= -10 - \frac{1}{10}(-103.1127) = 0.3113$$

ค่าวิกฤติจากตาราง AD คือ 2.492

เนื่องจากค่าสถิติ AD = 0.3113 ซึ่งน้อยกว่าค่าวิกฤติจึงไม่ปฏิเสธสมมติฐานหลัก ดังนั้น จำนวนการสั่งซื้อสินค้าของร้านค้าแห่งนี้มีการแจกแจงแบบปกติที่ระดับนัยสำคัญ .05

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิธี permutation test

วิธีการ permutation test เป็นวิธีการหนึ่งของสถิติไม่ใช้พารามิเตอร์ ผู้นำเสนอคนแรก คือ Fisher ในปี ค.ศ. 1935 โดยสมมติให้จำนวนการ permutation test ที่เป็นไปได้ทั้งหมดสามารถนับได้ วิธีการ permutation test มักใช้เมื่อไม่ทราบการแจกแจงของตัวอย่างชัดเจน (Smyth and Phipson, 2010) โดยใช้การแจกแจงแบบ permutation ภายใต้สมมติฐานหลักเป็นจริง เพื่อหา p-value ซึ่งจำนวน permutation ที่เป็นไปได้ทั้งหมดสำหรับตัวอย่าง 2 กลุ่มคือ $\binom{n_1+n_2}{n_1}$ ชุด

ตัวอย่างที่ 5 การใช้ permutation test ในการหา p-value ของสถิติทดสอบ DSCF (Hollander and Wolfe, 1999) จากข้อมูลต่อไปนี้

x_1=1.3 x_2=1.7 x_3=2.3

y_1=1.2 y_2=1.7

ขั้นตอนที่ 1 หาสถิติทดสอบจากข้อมูลข้างต้น

- หาลำดับร่วมของสมาชิกทั้ง 2 กลุ่ม

y_1	x_1	x_2	y_2	x_3
1.2	1.3	1.7	1.7	2.3
(1)	(2)	(3.5)	(3.5)	(5)

- สูตรของสถิติทดสอบ DSCF คือ $W_{test}=\dfrac{S_{ij}-n_i(n_i+n_j+1)/2}{\sqrt{n_i n_j(n_i+n_j+1)/24}}$

เมื่อ S_{ij} คือ ผลรวมลำดับร่วมของสมาชิกกลุ่ม i ดังนั้น $S_{ij}=2+3.5+5=10.5$ แทนค่าในสูตรจะได้ว่า

$$W_{test}=\frac{10.5-3(3+2+1)/2}{\sqrt{(3)(2)(3+2+1)/24}}$$
$$=1.2247$$

ขั้นตอนที่ 2 หาสถิติทดสอบจากวิธี permutation test ดังนี้

- จำนวน permutation test ที่เป็นไปได้ทั้งหมดคือ $\binom{5}{3}=10$ ชุด

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

- สมาชิกใหม่ของกลุ่มที่ i และ j ได้จากการนำข้อมูลทั้งหมดมาเรียงสับเปลี่ยน โดยสมาชิกกลุ่มที่ i มีจำนวน 3 ตัว และสมาชิกกลุ่มที่ j มี 2 ตัว จะได้ทั้งหมด 10 ชุด

- คำนวณค่าสถิติทดสอบของสมาชิกใหม่ได้ดังนี้

กลุ่ม i	ลำดับร่วม	กลุ่ม j	ลำดับร่วม	S_{ij}	$\lvert W_{ij}\rvert$	ความน่าจะเป็น
y_1,x_1,x_2	1,2,3.5	y_2,x_3	3.5,5	6.5	2.0412	1/10
y_1,x_1,y_2	1,2,3.5	x_2,x_3	3.5,5	6.5	2.0412	1/10
y_1,x_1,x_3	1,2,5	x_2,y_2	3.5,3.5	8	0.8165	1/10
y_1,x_2,y_2	1,3.5,3.5	x_1,x_3	2,5	8	0.8165	1/10
y_1,x_2,x_3	1,3.5,5	x_1,y_2	2,3.5	9.5	0.4082	1/10
y_1,y_2,x_3	1,3.5,5	x_1,x_2	2,3.5	9.5	0.4082	1/10
x_1,x_2,y_2	2,3.5,3.5	y_1,x_3	1,5	9	0.0000	1/10
x_1,x_2,x_3	2,3.5,5	y_1,y_2	1,3.5	10.5	1.2247	1/10
x_1,y_2,x_3	2,3.5,5	y_1,x_2	1,3.5	10.5	1.2247	1/10
x_2,y_2,x_3	3.5,3.5,5	y_1,x_1	1,2	12	2.4495	1/10

จะได้ $P(|W_{ij}| \geq 2.4495) = 1/10$

$P(|W_{ij}| \geq 2.0412) = 3/10$

$P(|W_{ij}| \geq 1.2247) = 5/10$

$P(|W_{ij}| \geq 0.8165) = 7/10$

$P(|W_{ij}| \geq 0.4082) = 9/10$

$P(|W_{ij}| \geq 0) = 1$

ขั้นตอนที่ 3 ค่า p-value หาได้จาก $P(|W_{ij}| \geq |W_{test}|)$ นั่นคือ

$$P(|W_{ij}| \geq |W_{test}|) = P(|W_{ij}| \geq 1.2247)$$
$$=5/10$$
$$=0.5$$

ดังนั้น p-value ของสถิติทดสอบ DSCF ที่ได้จากวิธี permutation test ของข้อมูลข้างต้นคือ 0.5

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การแจกแจงที่ใช้ในการศึกษา

ในการศึกษาครั้งนี้เป็นการศึกษาผลการเปรียบเทียบพหุคูณเมื่อกลุ่มตัวอย่างมีขนาดเล็กและมีการแจกแจงไม่ใช่แบบปกติ สำหรับการแจกแจงที่ใช้ในการศึกษาครั้งนี้ประกอบด้วยการแจกแจงแบบปกติ การแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล ซึ่งมีรายละเอียดดังนี้

การแจกแจงแบบปกติ (Normal Distribution)

การแจกแจงแบบปกติเป็นการแจกแจงความน่าจะเป็นของตัวแปรสุ่มชนิดต่อเนื่องที่มีความสำคัญมาก และยังเป็นข้อตกลงเบื้องต้นของสถิติทดสอบส่วนใหญ่ (จิรัชย์, 2548)

บทนิยาม 1 ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ โดยมีพารามิเตอร์ μ และ σ^2 มีฟังก์ชันความหนาแน่น ดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad ; -\infty < x < \infty; -\infty < \mu < \infty; \sigma > 0$$

$$= 0 \quad ; \text{เมื่อ x มีค่าอื่น ๆ}$$

โดยมี μ เป็นค่าเฉลี่ย และ σ^2 เป็นความแปรปรวนของ X

สำหรับการจำลองข้อมูลจากการแจกแจงแบบปกติโดยใช้ภาษาโปรแกรม R (Rizzo, 2008) สามารถใช้คำสั่ง rnorm โดยมีรูปแบบคำสั่งดังนี้

x=rnorm(N, mean=mean, sd=sd)

เช่น การจำลองข้อมูลจากประชากรที่มีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยเท่ากับ 50 และความแปรปรวนเท่ากับ 4 จำนวน 10 ค่า สามารถเขียนคำสั่งได้ดังนี้

x=rnorm(10, mean=50, sd=2)

ได้ค่า x ดังนี้

47.5859 50.5549 52.1689 45.3086 50.8583 51.0121 48.8505 48.9067 48.8711 48.2199

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การแจกแจงแบบเบตา (Beta Distribution)

การแจกแจงแบบเบตาเป็นการแจกแจงที่ค่าของตัวแปรสุ่มอยู่ในช่วง (0,1) ในการหาฟังก์ชันความหนาแน่นของตัวแปรสุ่มที่มีการแจกแจงแบบนี้ต้องอาศัยฟังก์ชันเบตา (Beta function) เขียนแทนด้วยสัญลักษณ์ $B(\alpha,\beta)$ ซึ่ง

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx \qquad ; \ \alpha > 0 \ , \ \beta > 0$$

$$= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

ตัวแปรสุ่มที่มีการแจกแจงแบบเบตา ได้แก่ สัดส่วนหรือเปอร์เซ็นต์ของลักษณะที่สนใจ ซึ่งได้จากการนับหรือการวัด

บทนิยาม 2 ตัวแปรสุ่ม X มีการแจกแจงแบบเบตา โดยมีพารามิเตอร์ α และ β มีฟังก์ชันความหนาแน่นของ X ดังนี้

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad ; \ 0 < x < 1, \alpha > 0 \ , \ \beta > 0$$

$$= 0 \qquad ; \text{เมื่อ x มีค่าอื่น ๆ}$$

โดยมี $\frac{\alpha}{\alpha+\beta}$ เป็นค่าเฉลี่ย และ $\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ เป็นความแปรปรวนของ X

สำหรับการจำลองข้อมูลจากการแจกแจงแบบเบตา โดยใช้ภาษาโปรแกรม R (Rizzo, 2008) สามารถใช้คำสั่ง rbeta โดยมีรูปแบบคำสั่ง ดังนี้

x=rbeta(N, shape1 = α , shape1 = β)

เช่น การจำลองข้อมูลจากประชากรที่มีการแจกแจงแบบเบตา โดยมีค่าพารามิเตอร์ shape 1 = 0.77 และ shape 2 = 0.1 จำนวน 10 ค่า สามารถเขียนคำสั่งได้ ดังนี้

x=rbeta(10, shape1=0.77, shape2=0.33)

ได้ค่า x ดังนี้

0.8996691 0.9497791 0.9652076 0.1276624 0.1847987 0.0147857 0.7611615 0.9966021 0.5777879 0.1606840

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การแจกแจงแบบไวบุลล์ (Weibull Distribution)

การแจกแจงแบบไวบุลล์เป็นการแจกแจงที่ถูกใช้มากทางด้านวิศวกรรม โดยเฉพาะทางด้านการทดลองเกี่ยวกับอายุการใช้งานของเครื่องจักร หรือการประยุกต์ใช้กับทฤษฎีของความน่าเชื่อถือ (Reliability Theory) และทางด้านคณิตศาสตร์ประกันภัยเกี่ยวกับจำนวนเงินของการประกัน โดยเฉพาะการประกันวินาศภัย (ประสิทธิ์, 2545)

บทนิยาม 3 ตัวแปรสุ่ม X มีการแจกแจงแบบไวบุลล์ โดยมีพารามิเตอร์ α และ β มีฟังก์ชันความหนาแน่นของ X ดังนี้

$$f(x) = \alpha\beta x^{\beta-1} e^{-\left(\alpha x^{\beta}\right)} \quad ; x > 0, \alpha > 0, \beta > 0$$
$$= 0 \quad ; \text{เมื่อ x มีค่าอื่น ๆ}$$

โดยมี $\alpha^{-1/\beta}\Gamma\left(1+\frac{1}{\beta}\right)$ เป็นค่าเฉลี่ย และ $\alpha^{-2/\beta}\left[\Gamma\left(1+\frac{2}{\beta}\right)-\left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right]$ เป็นความแปรปรวนของ X

สำหรับการจำลองข้อมูลจากการแจกแจงแบบไวบุลล์โดยใช้ภาษาโปรแกรม R (Scarpa and Alberini, 2005) สามารถใช้คำสั่ง rweibull โดยมีรูปแบบคำสั่ง ดังนี้

x=rweibull(N, shape = α , scale = β)

เช่น การจำลองข้อมูลจากประชากรที่มีการแจกแจงแบบไวบุลล์ โดยมีพารามิเตอร์ shape = 0.5 และพารามิเตอร์ scale = 10 สามารถเขียนคำสั่งได้ ดังนี้

x=rweibull(10, shape=0.5, scale=10)

ได้ค่า x ดังนี้

58.00271753 4.29661124 9.14736287 70.67677314 12.93729918 1.62188044 0.05844397 5.63952307 37.93162044 3.70041399

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

การแจกแจงแบบเอกซ์โพเนนเชียล (Exponential Distribution)

การแจกแจงแบบเอกซ์โพเนนเชียลเป็นกรณีเฉพาะของการแจกแจงแบบแกมมาที่มี $\alpha = 1$ และ $\beta = \frac{1}{\lambda}$ โดยตัวแปรสุ่มที่มักมีการแจกแจงแบบเอกซ์โพเนนชียล ได้แก่ อายุการใช้งานของเครื่องจักร เวลาในการรอคอยเวลาในการให้บริการ

บทนิยาม 4 ตัวแปรสุ่ม X มีการแจกแจงแบบเอกซ์โพเนนเชียล โดยมีพารามิเตอร์ λ มีฟังก์ชันความหนาแน่นของ X ดังนี้

$$f(x) = \lambda e^{-\lambda x} \quad ; x > 0 \ , \ \lambda > 0$$
$$= 0 \quad ; \text{เมื่อ x มีค่าอื่น ๆ}$$

โดยมี $\frac{1}{\lambda}$ เป็นค่าเฉลี่ย และ $\frac{1}{\lambda^2}$ เป็นความแปรปรวนของ X

สำหรับการจำลองข้อมูลจากการแจกแจงแบบเอกซ์โพเนนเชียลโดยใช้ภาษาโปรแกรม R (Rizzo, 2008) สามารถใช้คำสั่ง rexp โดยมีรูปแบบคำสั่ง ดังนี้

x=rexp(N, 1/mean)

เช่น การจำลองข้อมูลจากประชากรที่มีการแจกแจงแบบเอกซ์โพเนนเชียล โดยมีค่าเฉลี่ยเท่ากับ 50 จำนวน 10 ค่า สามารถเขียนคำสั่งได้ ดังนี้

x=rexp(10,1/50)

ได้ค่า x ดังนี้

125.0879302 12.3379442 0.3290978 87.1373045 19.3591292 4.4974836 41.2040757 10.1308950 41.9020160 38.0215150

งานวิจัยที่เกี่ยวข้อง

วิชชุดา (2539) ได้ศึกษาเปรียบเทียบวิธีการเปรียบเทียบพหุคูณในแผนการทดลองแบบสุ่มสมบูรณ์ด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ 4 วิธี ได้แก่ วิธี Unrestricted LSD, วิธี MurphysGap LSD, วิธี Tukey (H) และวิธี Murphys Gap Unrestricted LSD โดยข้อมูลที่ใช้มีการแจกแจงแบบปกติ มีค่าสัมประสิทธิ์ความผันแปร 5%, 10%, 15%, 20%, และ 30% จำนวนปัจจัยที่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ใช้ทดลองเท่ากับ 3, 4, 5, 6 และ 10 ปัจจัย ขนาดตัวอย่าง 4 ระดับ คือ 5, 10, 15 และ 20 และมีจำนวนซ้ำในแต่ละปัจจัยไม่เท่ากัน 3 ระดับ ซึ่งข้อมูลที่ใช้ในการศึกษาได้จากการจำลองด้วยเทคนิคมอนติคาร์โล ทำซ้ำ 1,000 รอบในแต่ละสถานการณ์ โดยพิจารณาจากค่าความคลาดเคลื่อนชนิดที่ 1 และหาอำนาจการทดสอบเมื่อการทดสอบผ่านเกณฑ์ในการควบคุมความคลาดเคลื่อนชนิดที่ 1 พบว่า วิธี Unrestricted LSD สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีและมีอำนาจการทดสอบสูงสุดในทุกสถานการณ์

สิรินุช (2544) ได้ศึกษาเปรียบเทียบผลการเปรียบเทียบพหุคูณเมื่อประชากรมีค่าความแปรปรวนไม่เท่ากันด้วยสถิติทดสอบ 3 วิธี ได้แก่ วิธี Dunnett's T3 วิธี Games-Howell และวิธี Brown-Forsythe จากตัวอย่างขนาด 20, 40, 80 และ 120 พบว่า เมื่อกลุ่มตัวอย่างมีขนาดใหญ่ จำนวนการปฏิเสธสมมติฐานหลักแตกต่างกันอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01 และผลการเปรียบเทียบพหุคูณระหว่างขนาดตัวอย่างพบว่า ทั้ง 3 วิธีมีจำนวนการปฏิเสธสมมติฐานหลักแตกต่างกันอย่างมีนัยสำคัญที่ระดับนัยทางสถิติสำคัญ 0.01

ปุณยนุช (2548) ได้ศึกษาเปรียบเทียบวิธีการเปรียบเทียบพหุคูณสำหรับแผนการทดลองแบบสุ่มสมบูรณ์ด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ 14 วิธี ได้แก่ วิธี LSD, วิธี Tukey's HSD, วิธี Bonferroni, วิธี Tukey's b, วิธี Sidak, วิธี Duncan, วิธี Scheffe's, วิธี Hochberg's GT2, วิธี R-E-G-WF, วิธี Gabriel, วิธี R-E-G-WQ, วิธี Waller-Duncan, วิธี S-N-K และ วิธี Dunnett โดยข้อมูลที่ใช้ในการศึกษาได้จากการจำลองข้อมูลด้วยเทคนิคมอนติคาร์โล ทำซ้ำ 10,000 รอบ ภายใต้เงื่อนไขข้อมูลมีการแจกแจงแบบปกติ และความแปรปรวนเท่ากัน โดยมีจำนวนปัจจัย 3 ถึง 8 กลุ่ม พบว่า กรณีขนาดตัวอย่างเท่ากัน วิธี LSD และวิธี Duncan สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ทุกกรณี ส่วน วิธี Sidak, วิธี Dunnett, วิธี Tukey's b, วิธี Waller-Duncan, วิธี S-N-K, วิธี Gabriel และวิธี R-E-G-WF สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้บางกรณี ส่วนกรณีขนาดตัวอย่างไม่เท่ากัน พบว่า วิธี LSD, วิธี Waller-Duncan และวิธี Duncan สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ทุกกรณี ส่วนวิธี Dunnett, วิธี Tukey's b, วิธี S-N-K, วิธีGabriel และวิธี R-E-G-W-F สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้บางกรณี และเมื่อพิจารณาอำนาจการทดสอบพบว่า วิธี LSD และวิธี Waller-Duncan มีอำนาจการทดสอบสูงสุดเมื่อมีจำนวนปัจจัยเท่ากับ 3 และ 4 ปัจจัย และวิธี Gabriel มีอำนาจการทดสอบสูงสุดเมื่อมีจำนวนปัจจัยเท่ากับ 5 ถึง 8 ปัจจัย

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ธีรศักดิ์ (2551) ได้ศึกษาการเปรียบเทียบรายคู่ด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณ 3 วิธี คือ วิธี Dunnett's T3, วิธี Games–Howell และวิธี Brown-Forsythe ซึ่งเป็นวิธีการเปรียบเทียบรายคู่สำหรับประชากรที่สองกลุ่มที่ไม่มีข้อสมมติเกี่ยวกับความเท่ากันของความแปรปรวน โดยเปรียบเทียบวิธีทดสอบทั้ง 3 วิธี เมื่อใช้กลุ่มตัวอย่างที่มีการแจกแจงรูปทรงที่แตกต่างกัน และการทดสอบเดียวกันกับกลุ่มตัวอย่างที่มีขนาดต่างกัน พบว่า มีความแตกต่างกันระหว่างวิธีการเปรียบเทียบในข้อมูลที่มีลักษณะเป็นโค้งปกติ และโค้งสูงแหลม ส่วนขนาดตัวอย่างพบว่ามีผลต่อการปฏิเสธสมมติฐานหลักในแต่ละวิธีแตกต่างกัน สำหรับอำนาจการทดสอบพบว่าอำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดเพิ่มขึ้น นอกจากนี้ยังพบว่า การแจกแจงของข้อมูลที่มีลักษณะเป็นโค้งปกติจะมีอำนาจการทดสอบมากกว่าอำนาจการทดสอบจากการแจกแจงลักษณะอื่น และวิธี Brown-Forsythe มีอำนาจการทดสอบมากที่สุด

นิภาพร (2552) ได้ศึกษาเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบใช้พารามิเตอร์ 5 วิธี ได้แก่ วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan, วิธี Scheffe' และวิธี Fisher's LSD โดยใช้ขนาดตัวอย่างเท่ากันสำหรับวิธีการต่างกัน และขนาดตัวอย่างต่างกันสำหรับวิธีการเดียวกัน ภายใต้ความแปรปรวนที่เท่ากัน พบว่า กรณีขนาดตัวอย่างเท่ากัน อำนาจการทดสอบของทุกวิธีการไม่แตกต่างกัน สำหรับการทดสอบในกรณีขนาดตัวอย่างไม่เท่ากันพบว่า วิธี Tukey-Kramer, วิธี Newman Keuls, วิธี Duncan และวิธี Scheffe' อำนาจการทดสอบขึ้นอยู่กับขนาดของตัวอย่าง ส่วนวิธี Fisher's LSD อำนาจการทดสอบไม่ต่างกันในทุกขนาดตัวอย่าง และมีอำนาจการทดสอบสูงที่สุด และนอกจากนี้ยังพบว่า อำนาจการทดสอบมีแนวโน้มสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น

Neuhäuser and Bretz (2001) ได้ศึกษาเปรียบเทียบสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณแบบไม่ใช้พารามิเตอร์ 2 วิธี คือ วิธี Steel-Dwass และ Critchlow-Fligner (DSCF) ซึ่งเป็นวิธีที่ประยุกต์ใช้สถิติ Wilcoxon และวิธี Baumgartner-Weiß-Schindler (BWS) เมื่อพิจารณาจากค่าความคลาดเคลื่อนชนิดที่ 1 พบว่า กรณีขนาดตัวอย่างแต่ละกลุ่มเท่ากับ 5 หน่วย ทั้งสองวิธีไม่มีความแตกต่างกัน ส่วนกรณีขนาดตัวอย่างแต่ละกลุ่มเท่ากับ 10 หน่วย เมื่อจำนวนกลุ่มเท่ากับ 3 ถึง 8 กลุ่ม วิธี BWS มีค่าความคลาดเคลื่อนชนิดที่ 1 เท่ากับระดับนัยสำคัญที่แท้จริง และเมื่อพิจารณาความสามารถในการตรวจพบความแตกต่างของแต่ละกลุ่มตัวอย่าง พบว่า เมื่อพิจารณาจากจำนวนความผิดพลาดต่อการทดสอบ หรือจำนวนครั้งของตรวจพบความแตกต่างอย่างน้อยหนึ่งคู่ของการเปรียบเทียบต่อจำนวนครั้งของการทดสอบทั้งหมด พบว่า กรณีของการแจกแจงแบบสมมาตร

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

วิธี DSCF มีความสามารถในการตรวจพบความแตกต่างมากกว่าวิธี BWS เล็กน้อย สำหรับกรณีการแจกแจงที่มีหางหนา และการแจกแจงที่มีความโด่งวิธี BWS มีความสามารถในการตรวจพบความแตกต่างได้ดีกว่า และเมื่อพิจารณาจากความผิดพลาดต่อการเปรียบเทียบ หรือจำนวนครั้งในการตรวจพบความแตกต่างทุกคู่ของการเปรียบเทียบต่อจำนวนครั้งของการทดสอบทั้งหมดพบว่า กรณีของการแจกแจงแบบสมมาตรทั้งสองวิธีไม่มีความแตกต่างกัน แต่ในกรณีการแจกแจงที่มีหางหนาและการแจกแจงที่มีความโด่ง วิธี BWS มีความสามารถในการตรวจพบความแตกต่างได้ดีกว่า

Neuhäuser (2003) ได้ศึกษาสถิติทดสอบแบบไม่ใช้พารามิเตอร์สำหรับทดสอบความแตกต่างของกลุ่มตัวอย่างสองกลุ่ม ได้แก่ วิธี Wilcoxon, วิธี Kolmogorov-Smirnov, วิธี χ^2 และวิธี exact โดยใช้สถิติทดสอบ Wilcoxon, สถิติทดสอบ Kolmogorov-Smirnov, สถิติทดสอบ χ^2 และสถิติทดสอบ Baumgartner เมื่อข้อมูลมีค่าซ้ำ (ties) ซึ่งต้องใช้ลำดับเฉลี่ยในการคำนวณ พบว่า วิธี Wilcoxon สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ใกล้เคียงกับระดับนัยสำคัญที่แท้จริงมากที่สุด ส่วนวิธี exact พบว่า เมื่อใช้สถิติทดสอบ Wilcoxon และสถิติทดสอบ Kolmogorov-Smirnov จะสามารถตรวจพบความแตกต่างได้ดีกว่าเมื่อใช้สถิติทดสอบ Baumgartner และเมื่อใช้สถิติทดสอบ χ^2 มีบางกรณีที่ให้ผลดีกว่าเมื่อใช้สถิติทดสอบ Baumgartner หรือกล่าวได้ว่าสถิติทดสอบ Baumgartner ไม่สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดีนัก และเมื่อพิจารณาอำนาจการทดสอบพบว่า วิธี exact เมื่อใช้สถิติทดสอบ Baumgartner มีอำนาจการทดสอบมากกว่าทั้งสถิติทดสอบ Wilcoxon และสถิติทดสอบ Kolmogorov-Smirnov แต่ไม่สามารถสรุปได้ว่าอำนาจการทดสอบมากกว่าหรือน้อยกว่าสถิติทดสอบ χ^2

Neuhäuser and Senske (2004) ได้ศึกษาเปรียบเทียบวิธีการทดสอบความแตกต่างของค่าเฉลี่ยสำหรับข้อมูล microarray ซึ่งปกติจะไม่ได้มีการแจกแจงแบบปกติ ด้วยสถิติทดสอบแบบไม่ใช้พารามิเตอร์ คือ วิธี Baumgartner-Weiß-Schindler, วิธี Wilcoxon rank sum test และวิธี Fisher-Pitman permutation test และสถิติทดสอบแบบใช้พารามิเตอร์ คือ สถิติทดสอบที (t-test) พบว่า วิธี Wilcoxon rank sum test มีอำนาจการทดสอบมากกว่าสถิติทดสอบ t และวิธี Fisher-Pitman permutation test เมื่อข้อมูลมีการแจกแจงแบบไม่สมมาตร และการแจกแจงที่มีหางหนาหรือข้อมูลที่มีค่า outlier ส่วนวิธี Baumgartner-Weiß-Schindler มีอำนาจการทดสอบต่ำกว่าสถิติทดสอบ t และวิธี Fisher-Pitman permutation test เมื่อข้อมูลมีการแจกแจงแบบสมมาตร แต่เมื่อข้อมูลมีการแจกแจงแบบไม่สมมาตรพบว่า วิธี Baumgartner-Weiß-Schindler จะมีอำนาจการทดสอบสูงกว่า และไม่มีความอ่อนไหวต่อค่า outlier

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Neuhäuser and Lam (2004) ได้ศึกษาเปรียบเทียบวิธีการทางสถิติแบบไม่ใช้พารามิเตอร์เพื่อเปรียบเทียบความแตกต่างระหว่างประชากรสองกลุ่มสำหรับข้อมูล microarray ซึ่งโดยทั่วไปจะไม่ได้มีการแจกแจงแบบปกติ และมีความแปรปรวนของแต่ละกลุ่มไม่เท่ากัน โดยเปรียบเทียบวิธี Wilcoxon rank sum test, วิธี Fisher-Pitman permutation และวิธี Baumgartner-Weiß-Schindler พบว่า ทั้ง 3 วิธีมีความอ่อนไหว (sensitive) เมื่อความแปรปรวนของประชากรทั้งสองกลุ่มไม่เท่ากัน โดยเมื่อความแปรปรวนมีค่าต่างกันมากขึ้นสัดส่วนของการปฏิเสธสมมติฐานหลักเมื่อทดสอบด้วยวิธี Baumgartner-Weiß-Schindler จะมากขึ้น หรือกล่าวได้ว่าการตรวจพบความแตกต่างเป็นผลมาจากความแตกต่างของความแปรปรวนของประชากร ส่วนวิธี Fisher-Pitman permutation และวิธี Wilcoxon rank sum test มีความน่าจะเป็นที่จะปฏิเสธสมมติฐานหลักเมื่อขนาดตัวอย่างไม่เท่ากัน และจะแม่นยำขึ้นเมื่อตัวอย่างมีขนาดเล็กและมีความแปรปรวนขนาดใหญ่ อย่างไรก็ตามหากพบว่าความแปรปรวนของประชากรไม่เท่ากันแล้ว ไม่จำเป็นต้องทดสอบความแตกต่างของตำแหน่ง (location) แต่เนื่องจากข้อมูล microarray มักจะมีความแปรปรวนไม่เท่ากัน จึงควรทำการทดสอบแบบ 2 ขั้นตอน หากผลการทดสอบด้วยวิธี Baumgartner-Weiß-Schindler ในขั้นตอนแรกพบว่ามีความแตกต่างระหว่างประชากรสองกลุ่ม ในขั้นตอนต่อไปอาจทำการทดสอบความแตกต่างระหว่าง location เพียงอย่างเดียว ด้วยวิธีบูสแตปโดยใช้สถิติทดสอบ Welch (bootstrap test based on the Welch)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

อุปกรณ์และวิธีการ

อุปกรณ์

1. เครื่องไมโครคอมพิวเตอร์
2. โปรแกรม SAS เวอร์ชัน 9.1 และ ภาษาโปรแกรม R เวอร์ชัน 2.13.1

วิธีการ

การวิจัยครั้งนี้ศึกษาผลการทดสอบสมมติฐานด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติด้วยวิธีการทดสอบ 3 วิธี คือ วิธีการทดสอบด้วยสถิติทดสอบ Tukey - Karmer (TK) วิธีการทดสอบด้วยสถิติทดสอบ Baumgartner - Weiß - Schindler (BWS) และวิธีการทดสอบด้วยสถิติทดสอบ Dwass - Steel และ Critchlow - Fligner (DSCF) โดยพิจารณาจากค่าอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ อัตราความคลาดเคลื่อนต่อการทดสอบและอำนาจการทดสอบ สำหรับการแจกแจงที่ใช้ในการศึกษาครั้งนี้ ได้แก่ การแจกแจงแบบปกติ การแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล ซึ่งมีรายละเอียดขั้นตอนการดำเนินการดังนี้

1. ศึกษาวิธีการเปรียบเทียบพหุคูณด้วยสถิติทดสอบดังนี้
   1.1. สถิติทดสอบ Tukey-Karmer (TK)
   1.2. สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS)
   1.3. สถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF)

2. จำลองข้อมูลสำหรับใช้ในการศึกษาครั้งนี้ด้วยภาษาโปรแกรม R เวอร์ชัน 2.13.1 ดังนี้

   2.1. กำหนดจำนวนกลุ่มตัวอย่างที่ใช้ในการศึกษาดังนี้ กรณีขนาดตัวอย่างเท่ากัน คือ 3 4 และ 5 กลุ่ม สำหรับกรณีขนาดตัวอย่างไม่เท่ากันและกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ กำหนดจำนวนกลุ่มตัวอย่างที่ใช้ศึกษาคือ 3 กลุ่ม

   2.2. กำหนดขนาดตัวอย่างที่ใช้ในการศึกษาดังนี้ กรณีขนาดตัวอย่างเท่ากันกำหนดขนาดตัวอย่างคือ 5 7 และ 10 สำหรับกรณีขนาดตัวอย่างไม่เท่ากันกำหนดขนาดตัวอย่างคือ 5 ถึง 10 และกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ กำหนดขนาดตัวอย่างคือ 10

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

2.3. กำหนดการแจกแจงที่ใช้ในการศึกษาครั้งนี้คือ การแจกแจงแบบปกติ การแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล

2.4. ทดสอบการแจกแจงด้วยสถิติทดสอบแอนเดอร์สัน-ดาร์ลิ่ง

2.5. ทดสอบความเท่ากันของค่าเฉลี่ยด้วยสถิติทดสอบเอฟ และสถิติทดสอบทีกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

3. ทดสอบสมมติฐานด้วยสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF โดยใช้โปรแกรม SAS เวอร์ชัน 9.1 โดยสถิติทดสอบ BWS และสถิติทดสอบ DSCF จะใช้เทคนิค all possible permutation test ในการทดสอบสมมติฐาน ทำซ้ำในแต่ละสถานการณ์ 500 รอบ

4. คำนวณอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) อัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) และอำนาจการทดสอบ โดยมีรายละเอียดดังนี้

4.1. อัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (PCER) คำนวณจากผลการทดสอบสมมติฐานที่ได้จากสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF ดังนี้

$$\text{PCER} = \frac{\text{จำนวนการเปรียบเทียบที่ผิดพลาด}}{\text{จำนวนการเปรียบเทียบทั้งหมด}}$$

4.2. อัตราความคลาดเคลื่อนต่อการทดสอบ (FWER) คำนวณจากผลการทดสอบสมมติฐานที่ได้จากสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF ดังนี้

$$\text{FWER} = \frac{\text{จำนวนการทดสอบที่ผิดพลาดอย่างน้อย 1 การเปรียบเทียบ}}{\text{จำนวนการทดสอบทั้งหมด}}$$

4.3. อำนาจการทดสอบคำนวณจากผลการทดสอบสมมติฐานที่ได้จากสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF เมื่อสมมติฐานหลักไม่จริง ดังนี้

$$\text{อำนาจการทดสอบ} = \frac{\text{จำนวนการปฏิเสธสมมติฐานหลักเมื่อสมมติฐานหลักไม่จริง}}{\text{จำนวนการทดสอบทั้งหมด}}$$

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ผลและวิจารณ์

ผล

การวิจัยครั้งนี้เป็นการศึกษาผลการทดสอบสมมติฐานด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติด้วยวิธีการทดสอบ 3 วิธี ได้แก่ วิธีการทดสอบด้วยสถิติทดสอบ Tukey-Karmer (TK) วิธีการทดสอบด้วยสถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และวิธีการทดสอบด้วยสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF) โดยจำลองสถานการณ์ด้วยภาษาโปรแกรม R เวอร์ชัน 2.13.1 ให้ข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล กำหนดจำนวนกลุ่มตัวอย่างที่ใช้ในการศึกษากรณีขนาดตัวอย่างเท่ากัน คือ 3 4 และ 5 กลุ่ม และกำหนดขนาดตัวอย่าง คือ 5 7 และ 10 สำหรับกรณีขนาดตัวอย่างไม่เท่ากันศึกษาเฉพาะจำนวนกลุ่มตัวอย่างเท่ากับ 3 กลุ่ม ขนาดตัวอย่างเท่ากับ 5 ถึง 10 และกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ ศึกษาเฉพาะจำนวนกลุ่มตัวอย่างเท่ากับ 3 กลุ่ม ขนาดตัวอย่างเท่ากับ 10 โดยสร้างข้อมูลให้มีค่าเฉลี่ยเท่ากันสำหรับศึกษาอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) และอัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) และสร้างข้อมูลให้มีสมมติฐานหลักไม่จริงสำหรับศึกษาอำนาจการทดสอบ (Power of test) จากนั้นจึงทดสอบสมมติฐานด้วยสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณทั้ง 3 วิธี โดยใช้โปรแกรม SAS เวอร์ชัน 9.1 โดยกรณีของสถิติทดสอบแบบไม่ใช้พารามิเตอร์จะใช้วิธี permutation test ซึ่งผลการศึกษาแสดงรายละเอียดดังนี้

1. ผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Tukey-Karmer (TK) สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และสถิติทดสอบ Dwass - Steel และ Critchlow - Fligner (DSCF) กรณีขนาดตัวอย่างเท่ากัน

การศึกษาผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF พิจารณาจากอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ อัตราความคลาดเคลื่อนต่อการทดสอบ และอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน แสดงรายละเอียดดังต่อไปนี้

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

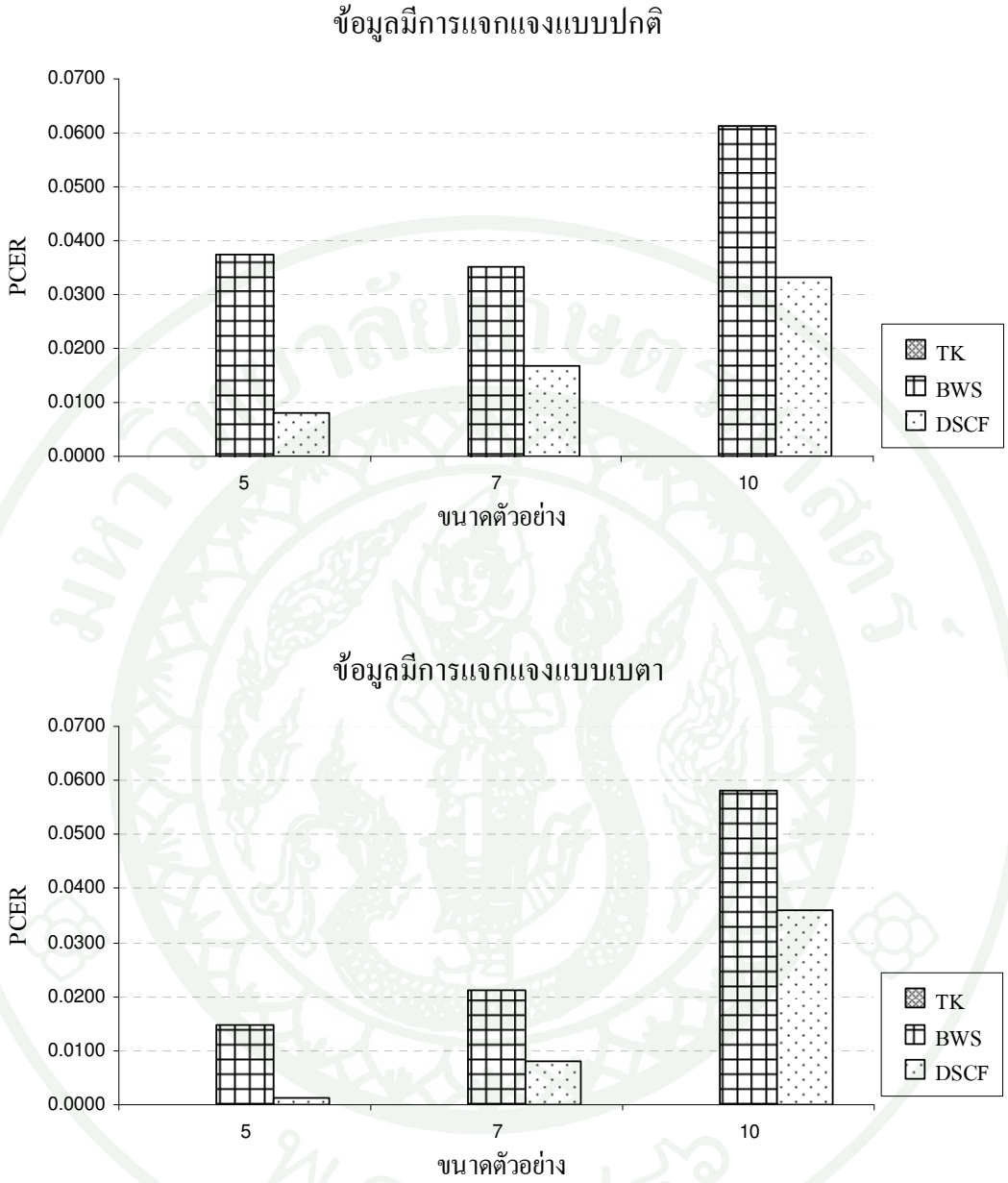
1.1 การศึกษาอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (PCER) ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF

ตารางที่ 3 แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า PCER		
		TK	BWS	DSCF
ปกติ	5	0.0000	0.0373	0.0080
	7	0.0000	0.0353	0.0167
	10	0.0000	0.0613	0.0333
เบตา	5	0.0000	0.0147	0.0013
	7	0.0000	0.0213	0.0080
	10	0.0000	0.0580	0.0360
ไวบุลล์	5	0.0000	0.0140	0.0013
	7	0.0000	0.0153	0.0047
	10	0.0000	0.0920	0.0520
เอกซ์โพเนนเชียล	5	0.0000	0.0100	0.0047
	7	0.0000	0.0220	0.0147
	10	0.0073	0.0647	0.0593

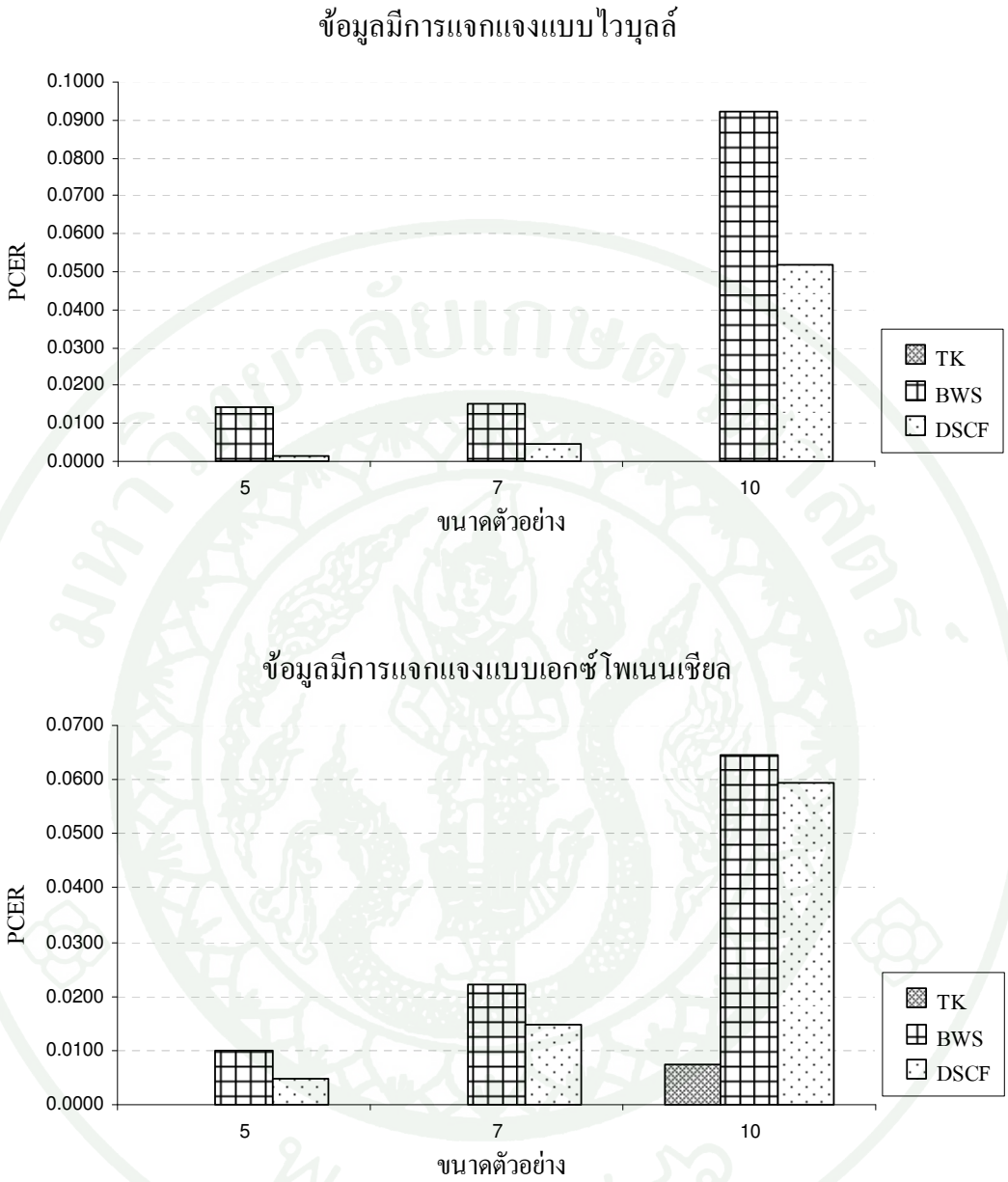
จากตารางที่ 3 พบว่า ค่า PCER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.0073 สำหรับค่า PCER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.01 ถึง 0.092 และค่า PCER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0013 ถึง 0.0593 และโดยภาพรวมพบว่า สถิติทดสอบ BWS มีค่า PCER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 1 แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 1 (ต่อ)

จากภาพที่ 1 แสดงให้เห็นว่า เมื่อข้อมูลมีขนาดใหญ่ขึ้นค่า PCER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้น และพบว่า สถิติทดสอบ BWS มีค่า PCER มากกว่าสถิติทดสอบอื่น

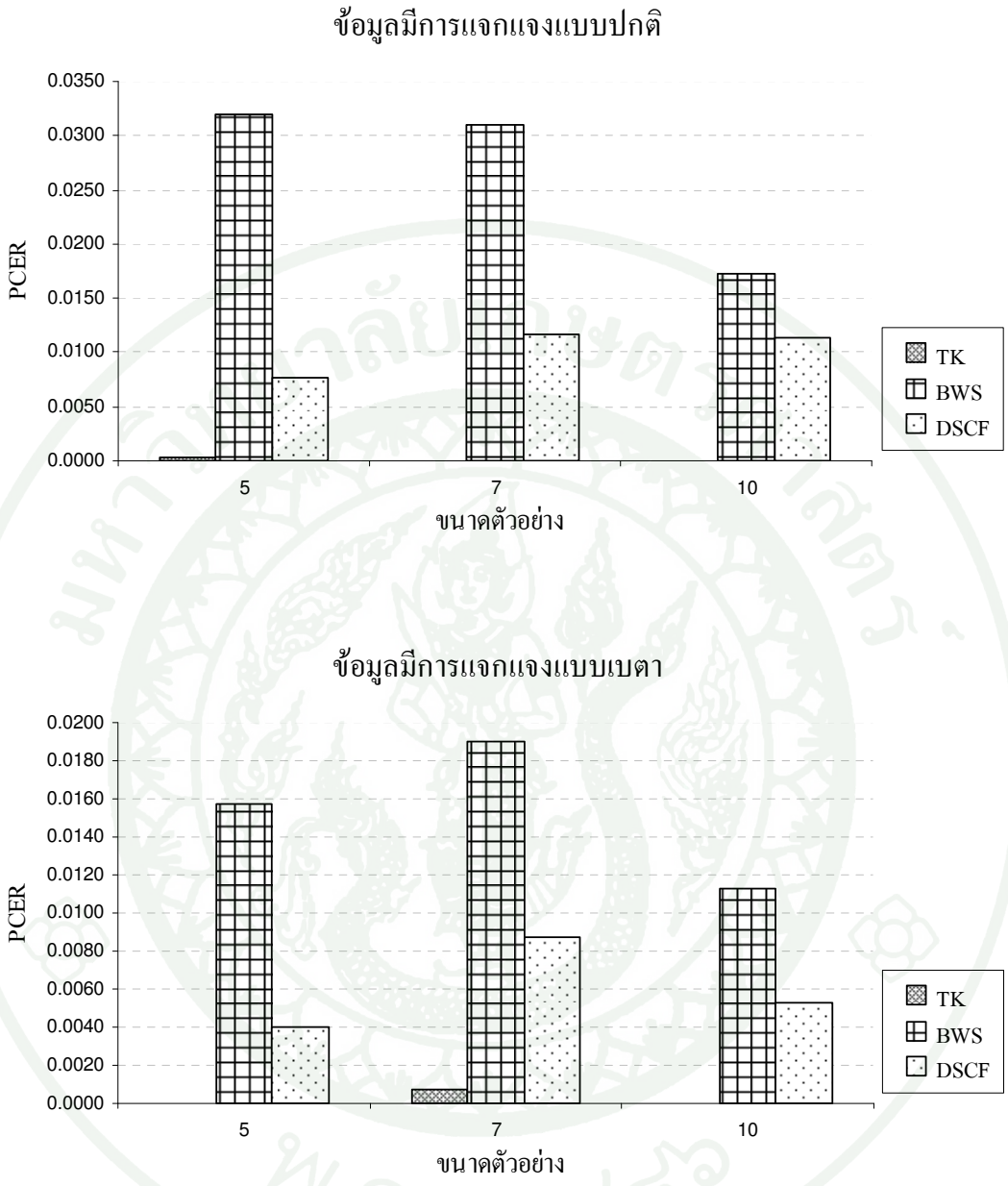
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 4 แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากันสำหรับการเปรียบเทียบ 4 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า PCER		
		TK	BWS	DSCF
ปกติ	5	0.0003	0.0320	0.0077
	7	0.0000	0.0310	0.0117
	10	0.0000	0.0173	0.0113
เบตา	5	0.0000	0.0157	0.0040
	7	0.0007	0.0190	0.0087
	10	0.0000	0.0113	0.0053
ไวบุลล์	5	0.0000	0.0137	0.0033
	7	0.0000	0.0213	0.0107
	10	0.0000	0.0113	0.0083
เอกซ์โพเนนเชียล	5	0.0000	0.0143	0.0027
	7	0.0007	0.0260	0.0150
	10	0.0000	0.0183	0.0143

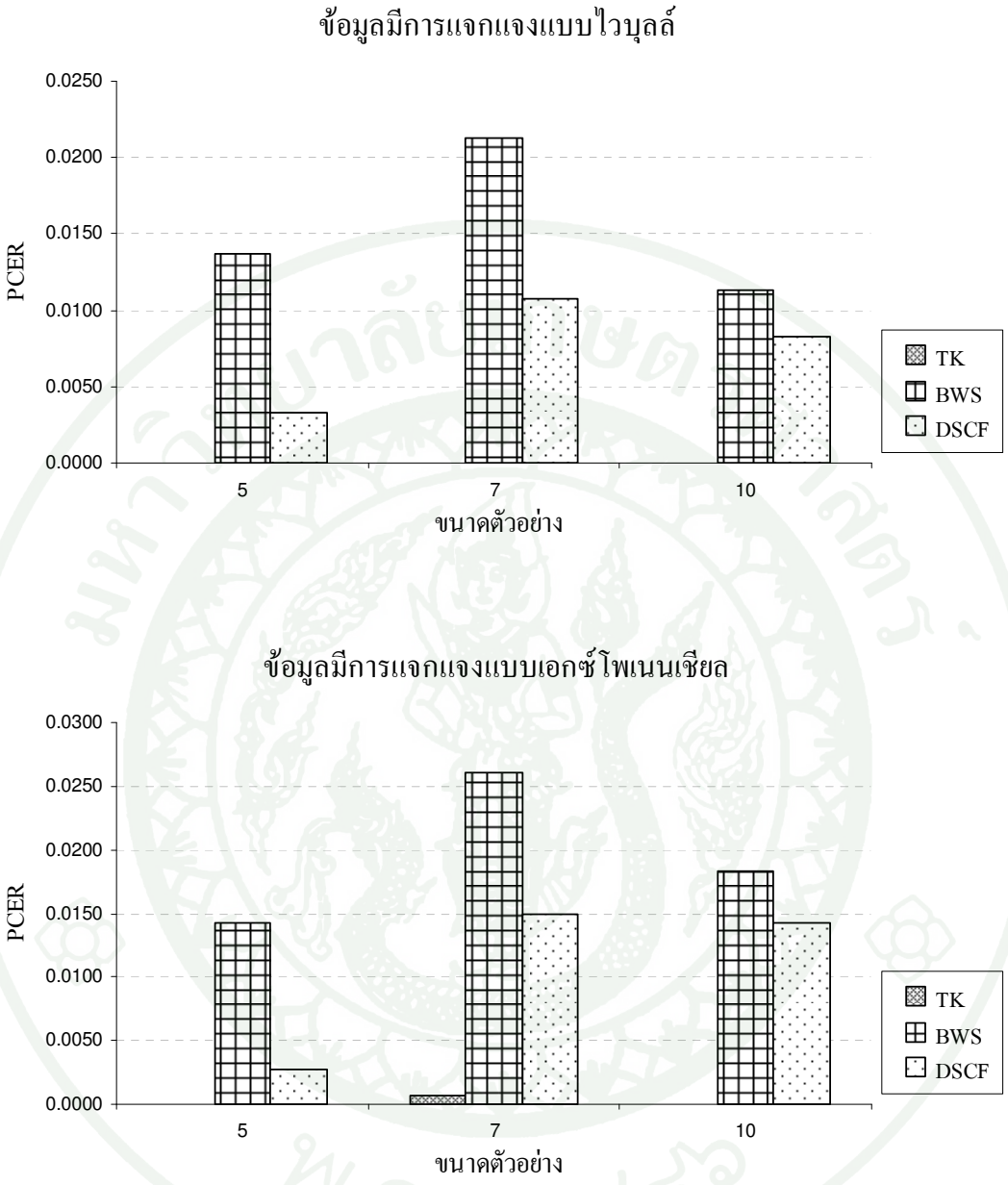
จากตารางที่ 4 พบว่า ค่า PCER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.0007 สำหรับค่า PCER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.0113 ถึง 0.0320 และค่า PCER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0027 ถึง 0.0150 และโดยภาพรวมพบว่า สถิติทดสอบ BWS มีค่า PCER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 2 แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 2 (ต่อ)

จากภาพที่ 2 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติค่า PCER ของสถิติทดสอบ BWS มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนค่า PCER ของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น เมื่อข้อมูลมีการแจกแจงแบบปกติและการแจกแจงแบบเอกซ์โพเนนเชียล

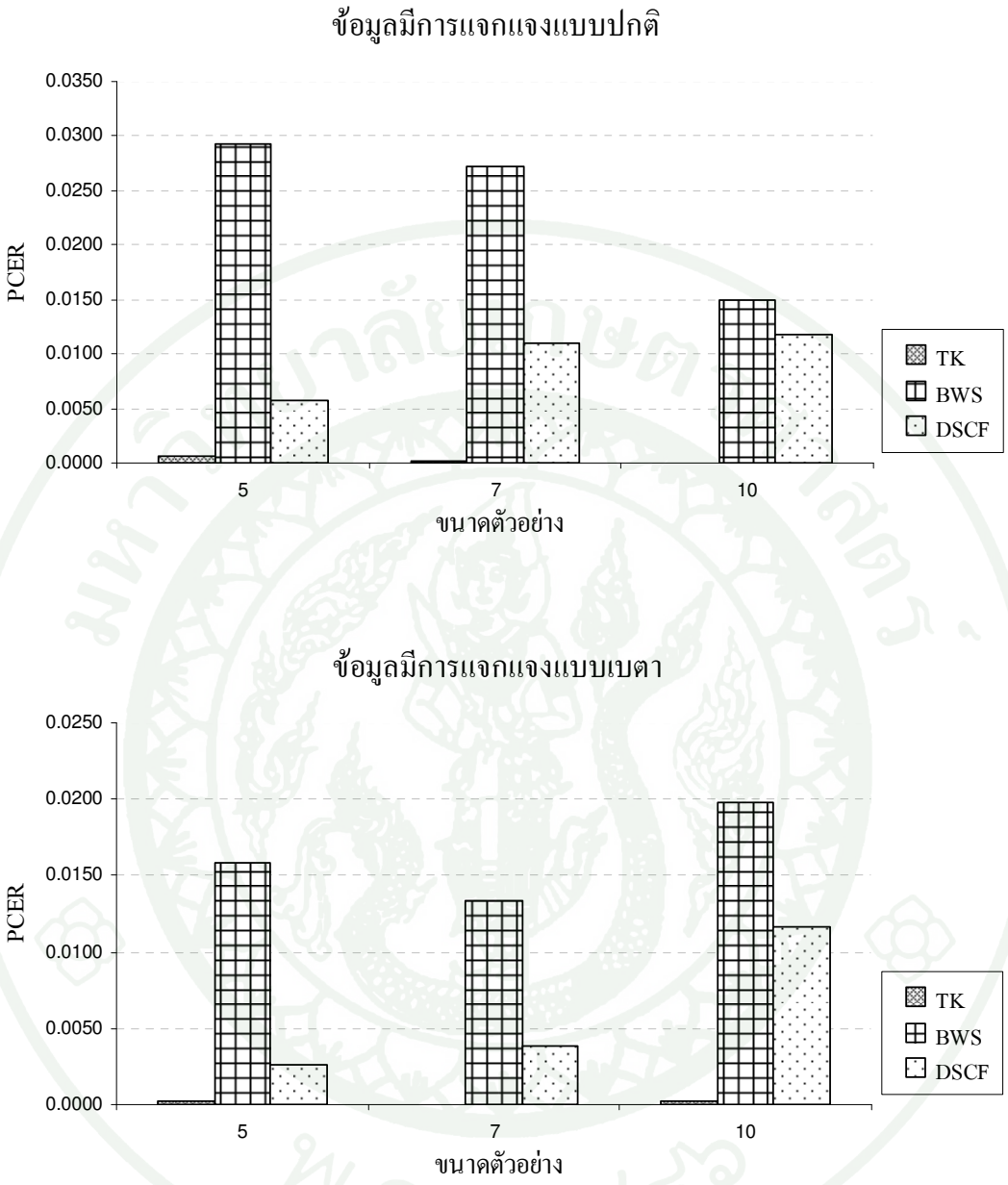
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 5 แสดงค่า PCER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า PCER		
		TK	BWS	DSCF
ปกติ	5	0.0006	0.0292	0.0058
	7	0.0002	0.0272	0.0110
	10	0.0000	0.0150	0.0118
เบตา	5	0.0002	0.0158	0.0026
	7	0.0000	0.0134	0.0038
	10	0.0002	0.0198	0.0116
ไวบุลล์	5	0.0006	0.0168	0.0042
	7	0.0000	0.0120	0.0046
	10	0.0000	0.0160	0.0092
เอกซ์โพเนนเชียล	5	0.0000	0.0202	0.0054
	7	0.0000	0.0136	0.0088
	10	0.0000	0.0186	0.0180

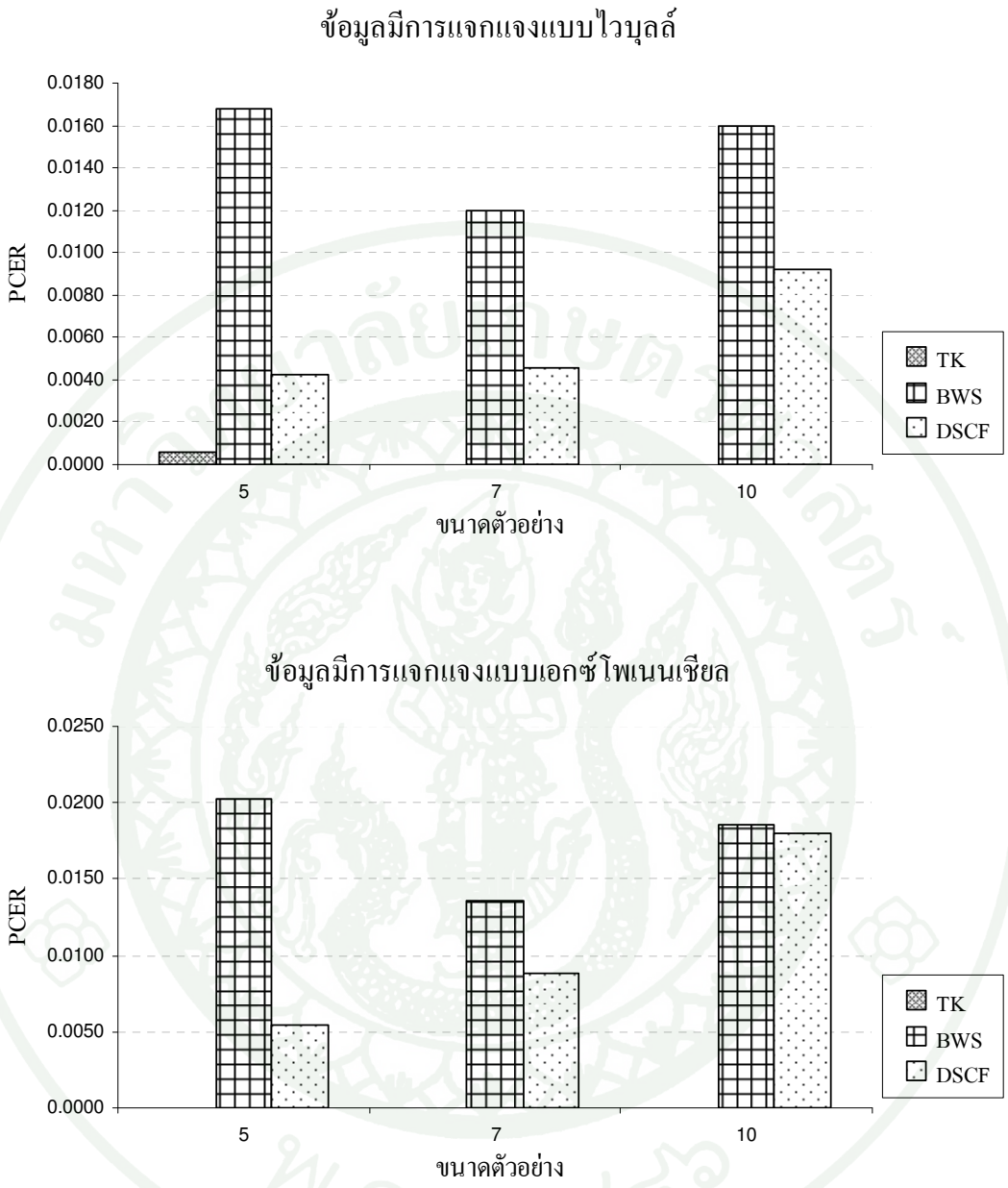
จากตารางที่ 5 พบว่า ค่า PCER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.0006 สำหรับค่า PCER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.0120 ถึง 0.029 และค่า PCER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0026 ถึง 0.0180 และโดยภาพรวมพบว่า สถิติทดสอบ BWS มีค่า PCER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 3 แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 3 (ต่อ)

จากภาพที่ 3 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติค่า PCER ของสถิติทดสอบ BWS มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนค่า PCER ของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อข้อมูลมีขนาดเพิ่มขึ้นสำหรับทุกการแจกแจง และพบว่า สถิติทดสอบ BWS มีค่า PCER มากกว่าสถิติทดสอบอื่นสำหรับทุกการแจกแจง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

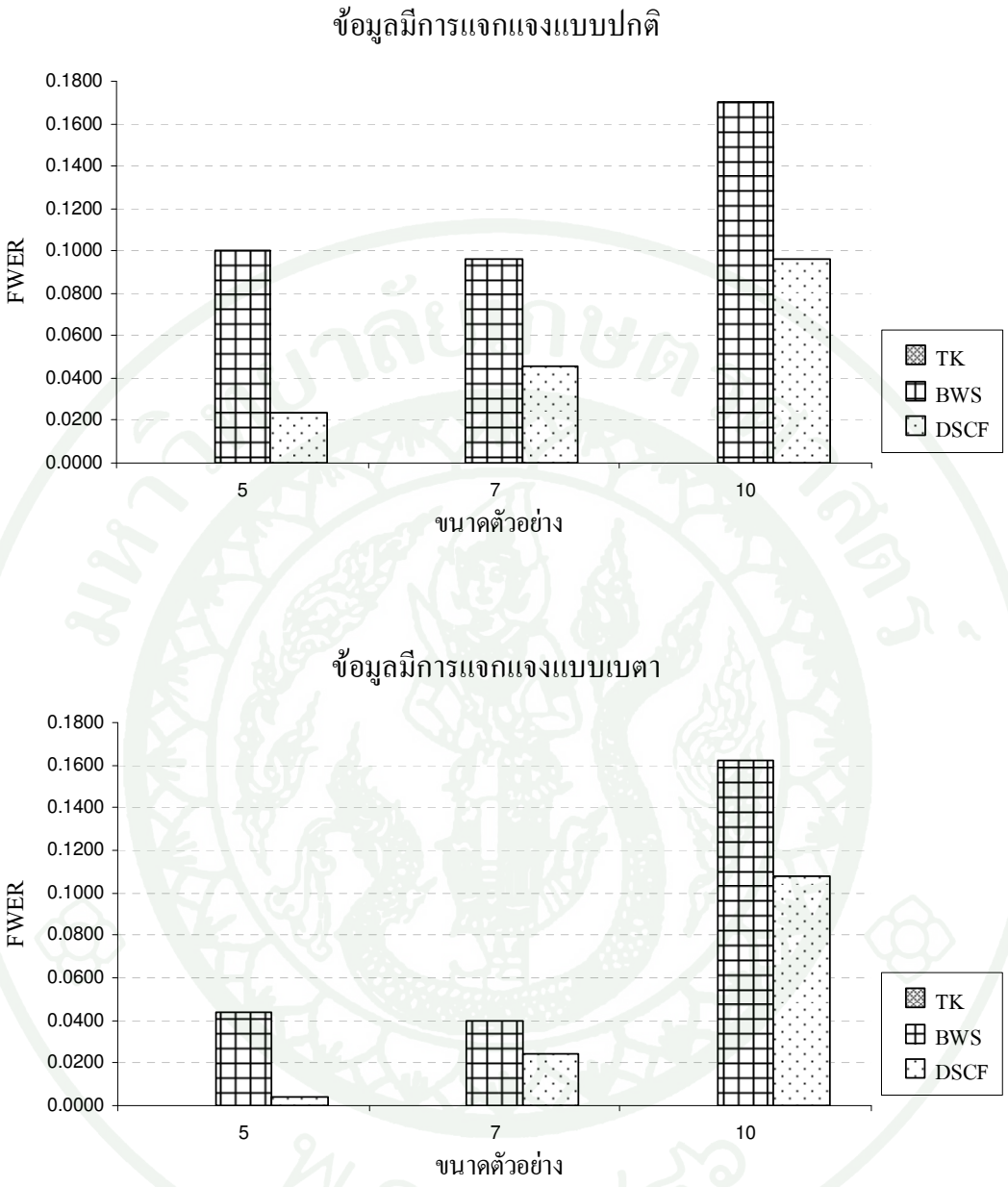
1.2 การศึกษาอัตราความคลาดเคลื่อนต่อการทดสอบ (FWER) ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF

ตารางที่ 6 แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า FWER		
		TK	BWS	DSCF
ปกติ	5	0.0000	0.1000	0.0240
	7	0.0000	0.0960	0.0460
	10	0.0000	0.1700	0.0960
เบตา	5	0.0000	0.0440	0.0040
	7	0.0000	0.0400	0.0240
	10	0.0000	0.1620	0.1080
ไวบุลล์	5	0.0000	0.0420	0.0040
	7	0.0000	0.0440	0.0140
	10	0.0000	0.2440	0.1480
เอกซ์โพเนนเชียล	5	0.0000	0.0280	0.0140
	7	0.0000	0.0580	0.0380
	10	0.0220	0.1740	0.1600

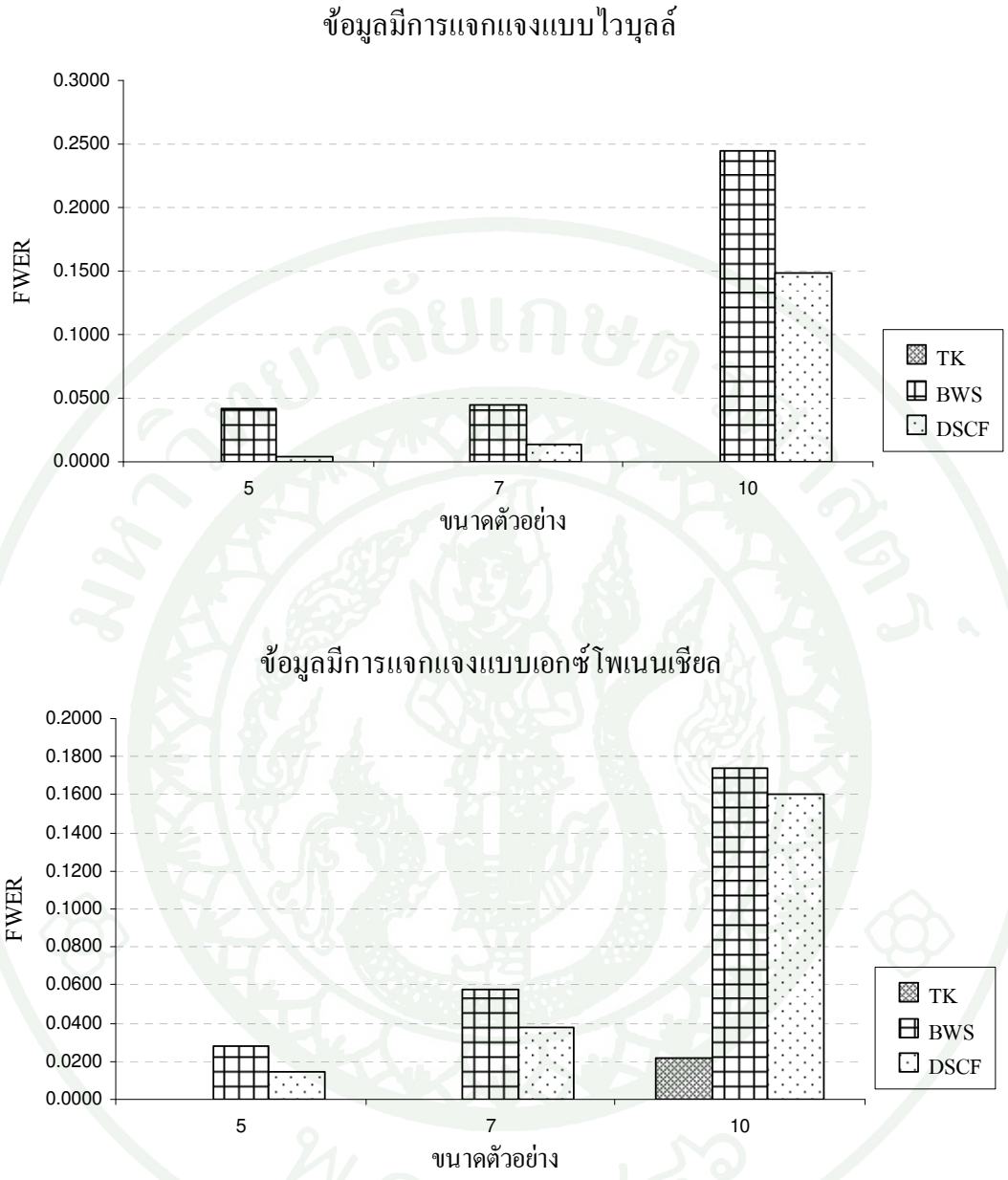
จากตารางที่ 6 พบว่า ค่า FWER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.0220 สำหรับค่า FWER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.0280 ถึง 0.2440 และค่า FWER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0040 ถึง 0.1600 เมื่อข้อมูลมีการแจกแจงแบบปกติพบว่า สถิติทดสอบ DSCF มีค่า FWER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น สำหรับกรณีที่ข้อมูลมีการแจกแจงแบบเบตา และการแจกแจงแบบไวบุลล์ พบว่า สถิติทดสอบ BWS จะมีค่า FWER ที่ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น และโดยภาพรวมพบว่า สถิติทดสอบ DSCF มีค่า FWER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 4 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 4 (ต่อ)

จากภาพที่ 4 แสดงให้เห็นว่า ค่า FWER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นสำหรับทุกการแจกแจง และพบว่า สถิติทดสอบ BWS มีค่า FWER มากกว่าสถิติทดสอบอื่น

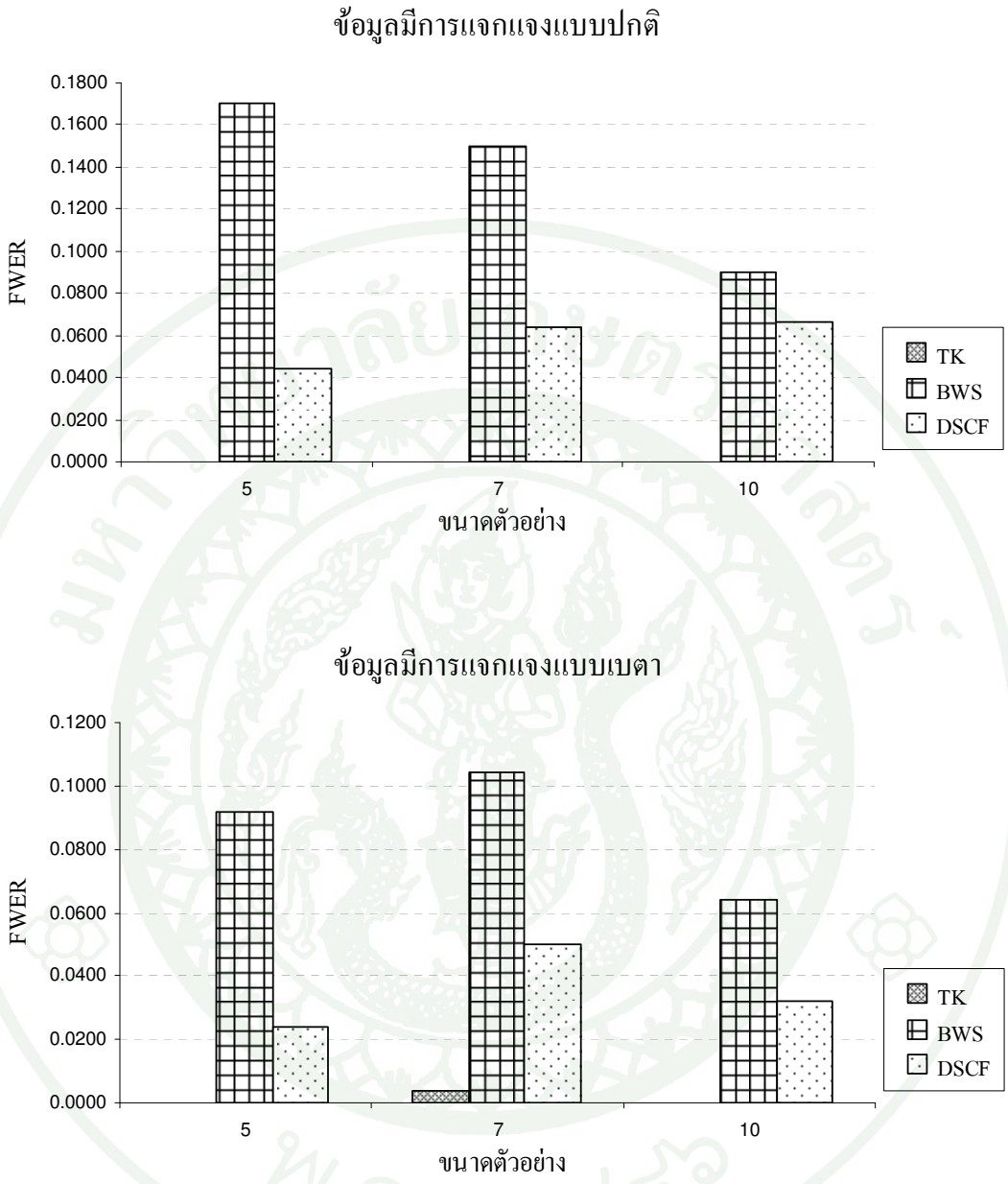
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 7 แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า FWER		
		TK	BWS	DSCF
ปกติ	5	0.0000	0.1700	0.0440
	7	0.0000	0.1500	0.0640
	10	0.0000	0.0900	0.0660
เบตา	5	0.0000	0.0920	0.0240
	7	0.0040	0.1040	0.0500
	10	0.0000	0.0640	0.0320
ไวบุลล์	5	0.0000	0.0780	0.0180
	7	0.0000	0.1120	0.0580
	10	0.0000	0.0620	0.0480
เอกซ์โพเนนเชียล	5	0.0000	0.0840	0.0160
	7	0.0040	0.1300	0.0760
	10	0.0000	0.0960	0.0720

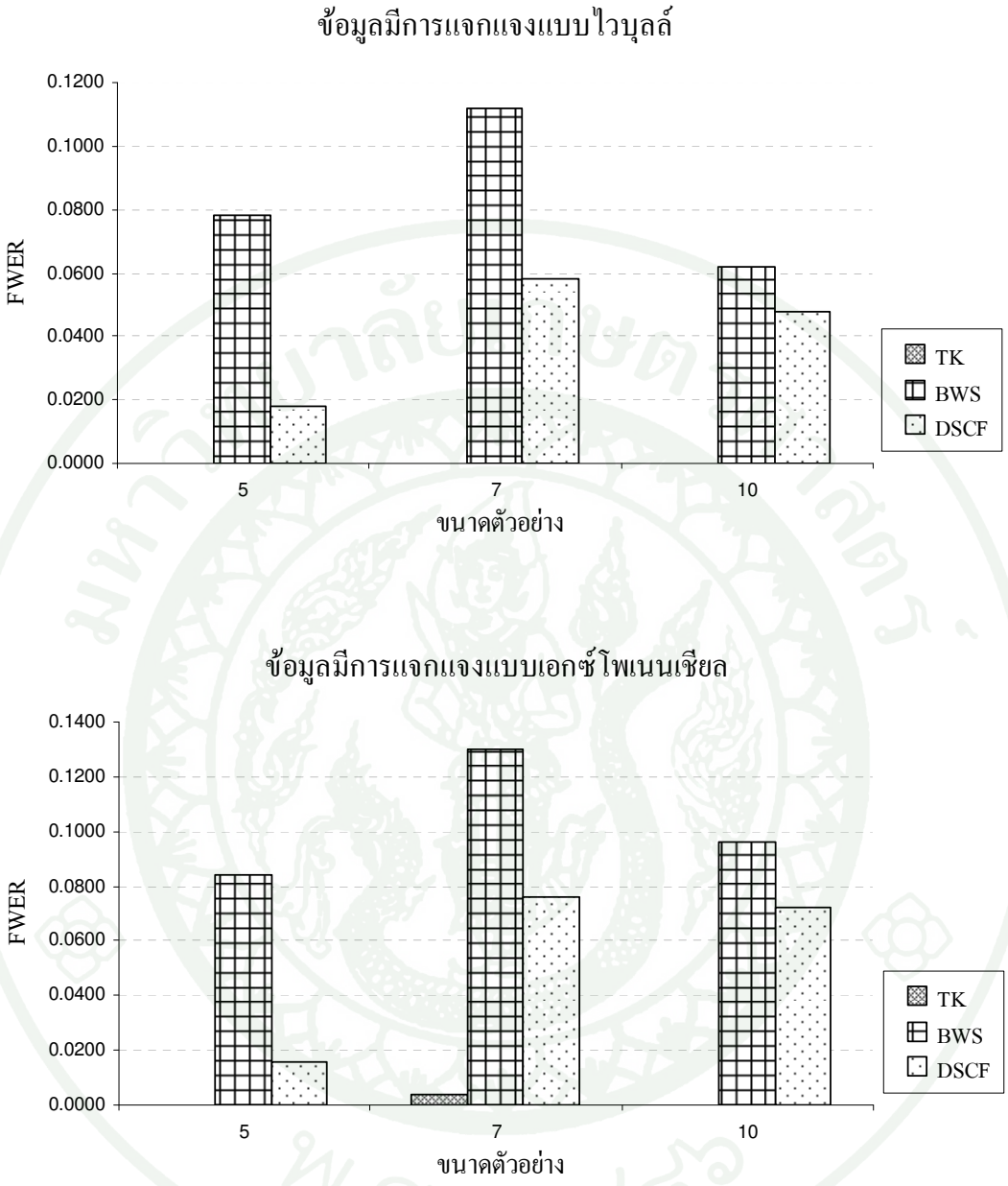
จากตารางที่ 7 พบว่า ค่า FWER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.004 สำหรับค่า FWER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.0620 ถึง 0.1700 และค่า FWER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0160 ถึง 0.0760 และโดยภาพรวมพบว่า สถิติทดสอบ DSCF มีค่า FWER ที่ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่นทุกการแจกแจง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 5 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 5 (ต่อ)

จากภาพที่ 5 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติค่า FWER ของสถิติทดสอบ BWS มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น สำหรับค่า FWER ของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และพบว่า สถิติทดสอบ BWS มีค่า FWER มากกว่าสถิติทดสอบอื่นทุกการแจกแจง

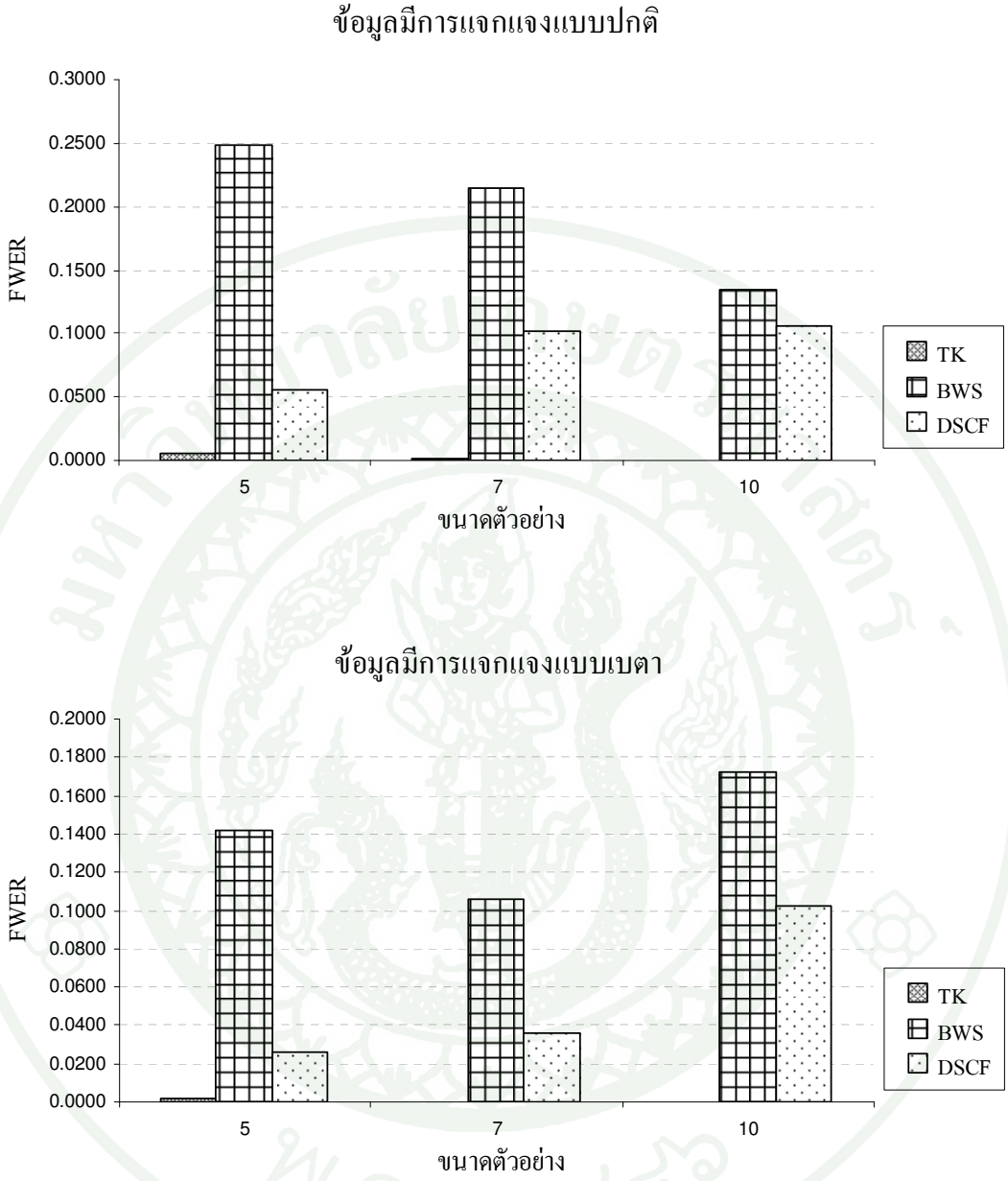
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 8 แสดงค่า FWER กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	ค่า FWER		
		TK	BWS	DSCF
ปกติ	5	0.0060	0.2480	0.0560
	7	0.0020	0.2140	0.1020
	10	0.0000	0.1340	0.1060
เบตา	5	0.0020	0.1420	0.0260
	7	0.0000	0.1060	0.0360
	10	0.0000	0.1720	0.1020
ไวบุลล์	5	0.0060	0.1640	0.0400
	7	0.0000	0.1020	0.0380
	10	0.0000	0.1260	0.0760
เอกซ์โพเนนเชียล	5	0.0000	0.1640	0.0520
	7	0.0000	0.1200	0.0800
	10	0.0000	0.1360	0.1400

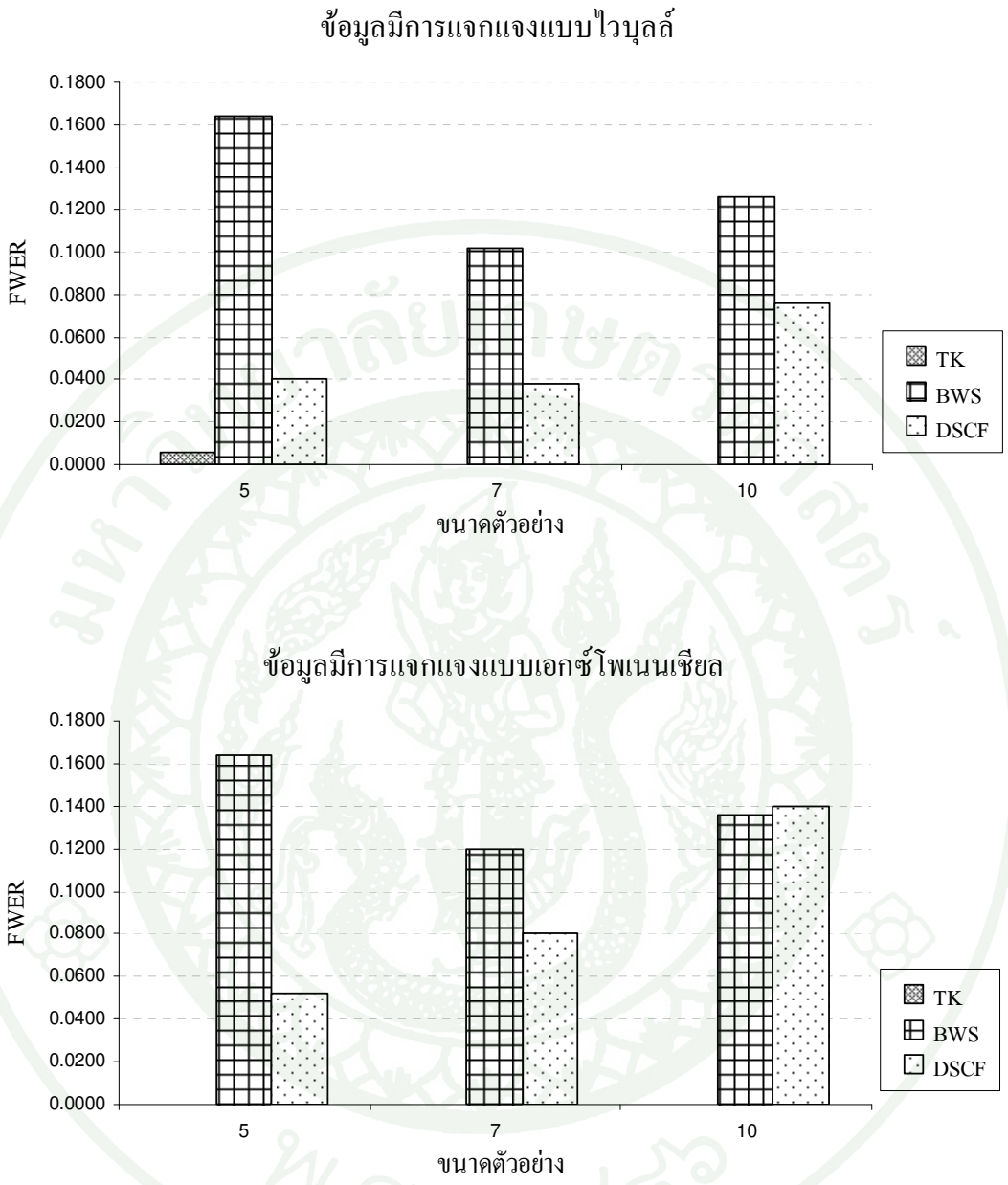
จากตารางที่ 8 พบว่า ค่า FWER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.006 สำหรับค่า FWER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.1020 ถึง 0.2480 และค่า FWER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0260 ถึง 0.1400 และโดยภาพรวมพบว่า สถิติทดสอบ DSCF มีค่า FWER ที่ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 6 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 6 (ต่อ)

จากภาพที่ 6 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติค่า FWER ของสถิติทดสอบ BWS มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนค่า FWER ของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นสำหรับทุกการแจกแจง และพบว่า ค่า FWER ของสถิติทดสอบ BWS มีค่ามากกว่าสถิติทดสอบอื่นยกเว้นกรณีการแจกแจงแบบเอกซ์โพเนนเชียลเมื่อขนาด

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตัวอย่างเท่ากับ 10 พบว่า ค่า FWER ของสถิติทดสอบ DSCF สูงกว่าค่า FWER ของสถิติทดสอบ BWS เล็กน้อย

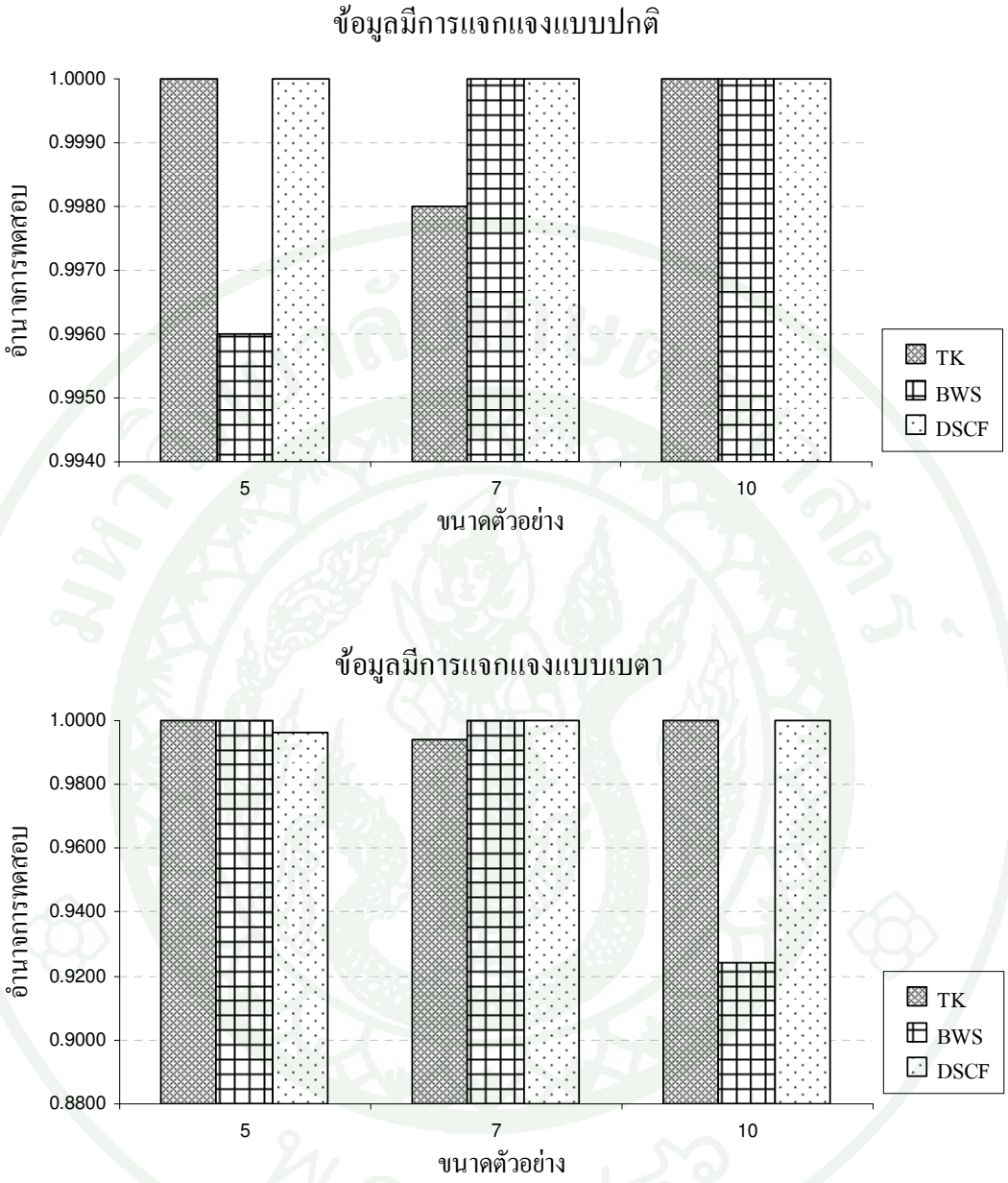
1.3 การศึกษาอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF

ตารางที่ 9 แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	อำนาจการทดสอบ		
		TK	BWS	DSCF
ปกติ	5	1.0000	0.9960	1.0000
	7	0.9980	1.0000	1.0000
	10	1.0000	1.0000	1.0000
เบตา	5	1.0000	1.0000	0.9960
	7	0.9940	1.0000	1.0000
	10	1.0000	0.9240	1.0000
ไวบุลล์	5	1.0000	1.0000	1.0000
	7	0.9800	1.0000	0.9980
	10	1.0000	1.0000	1.0000
เอกซ์โพเนนเชียล	5	1.0000	1.0000	0.9920
	7	0.9860	1.0000	0.9920
	10	1.0000	1.0000	1.0000

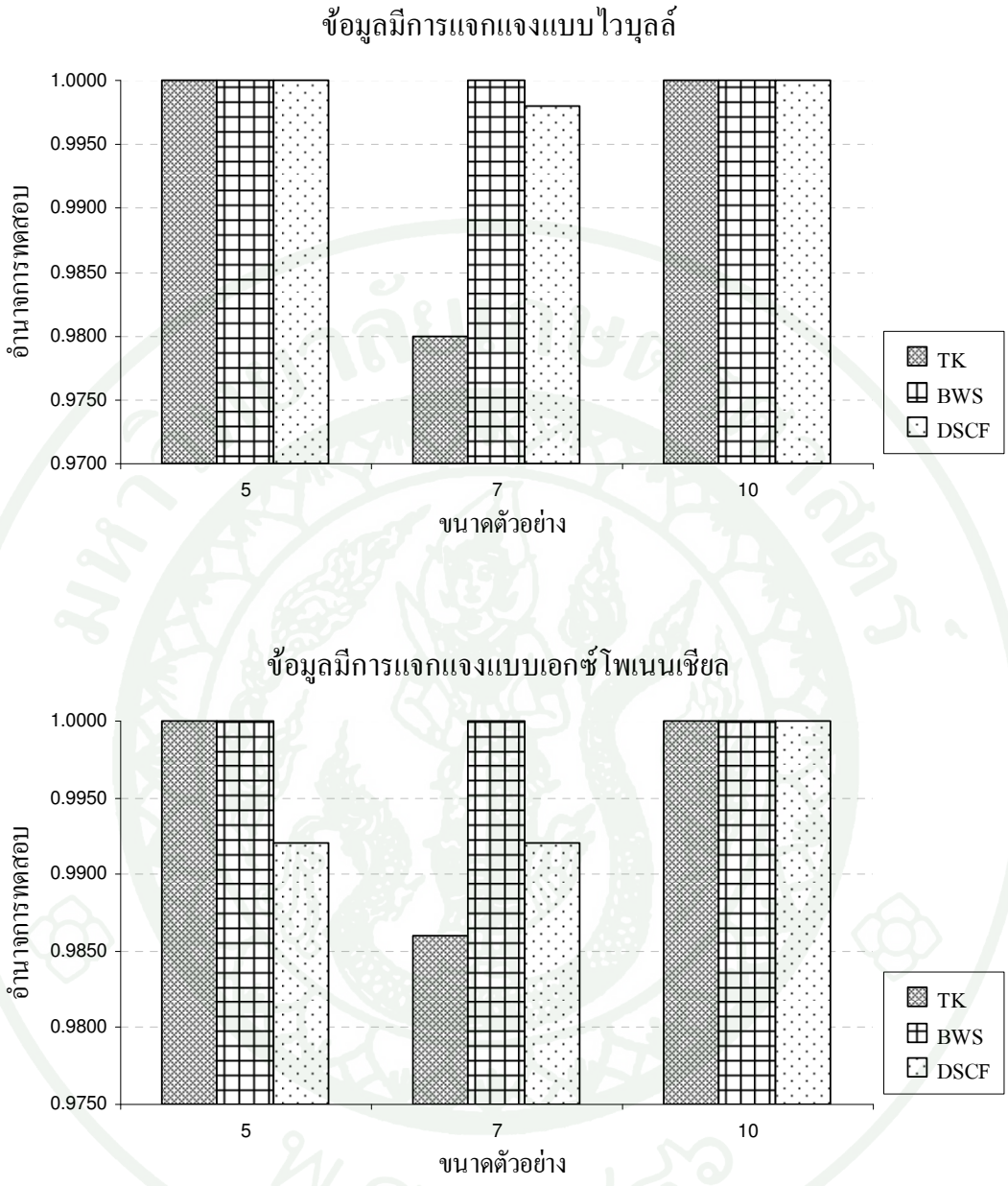
จากตารางที่ 9 พบว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าระหว่าง 0.9800 ถึง 1 สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีค่าระหว่าง 0.9240 ถึง 1 และอำนาจการทดสอบของสถิติทดสอบ DSCF มีค่าระหว่าง 0.992 ถึง 1 และโดยภาพรวมพบว่า สถิติทดสอบทั้ง 3 วิธีมีอำนาจการทดสอบที่ค่อนข้างสูง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 7 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 3 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 7 (ต่อ)

จากภาพที่ 7 แสดงให้เห็นว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าต่ำสุดเมื่อขนาดตัวอย่างเท่ากับ 7 สำหรับทุกการแจกแจง สำหรับกรณีเมื่อข้อมูลมีการแจกแจงแบบปกติอำนาจการทดสอบของสถิติทดสอบ BWS มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น และมีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้นและข้อมูลมีการแจกแจงแบบเบตา และอำนาจการทดสอบของสถิติ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

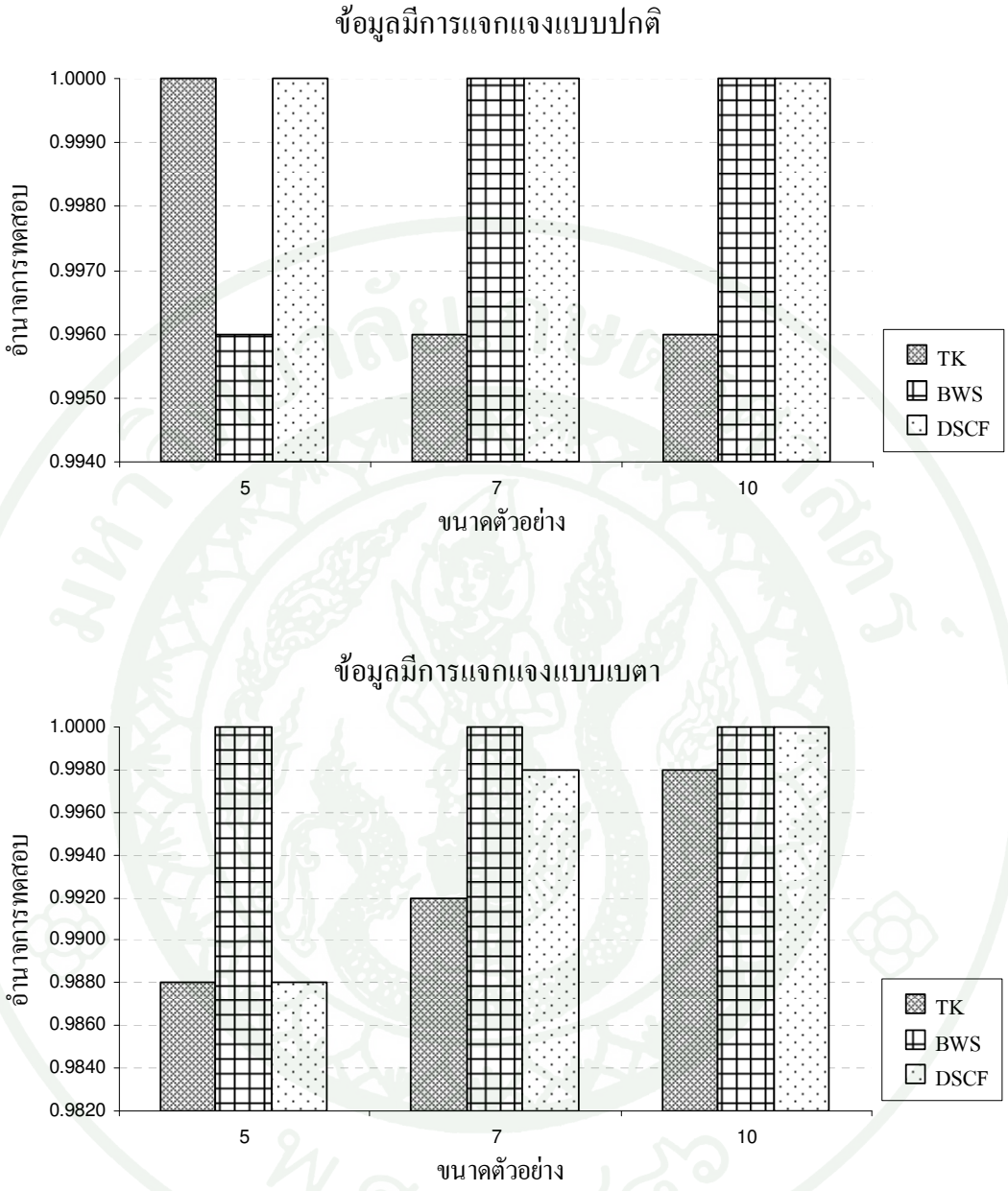
ทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเมื่อข้อมูลแจกแจงแบบเบตา และการแจกแจงแบบเอกซ์โพเนนเชียล

ตารางที่ 10 แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	อำนาจการทดสอบ		
		TK	BWS	DSCF
ปกติ	5	1.0000	0.9960	1.0000
	7	0.9960	1.0000	1.0000
	10	0.9960	1.0000	1.0000
เบตา	5	0.9880	1.0000	0.9880
	7	0.9920	1.0000	0.9980
	10	0.9980	1.0000	1.0000
ไวบุลล์	5	0.9980	1.0000	0.9980
	7	0.9660	1.0000	1.0000
	10	0.9620	0.9980	0.9880
เอกซ์โพเนนเชียล	5	0.9720	1.0000	0.9900
	7	0.9240	1.0000	0.9900
	10	0.7900	0.9520	0.9240

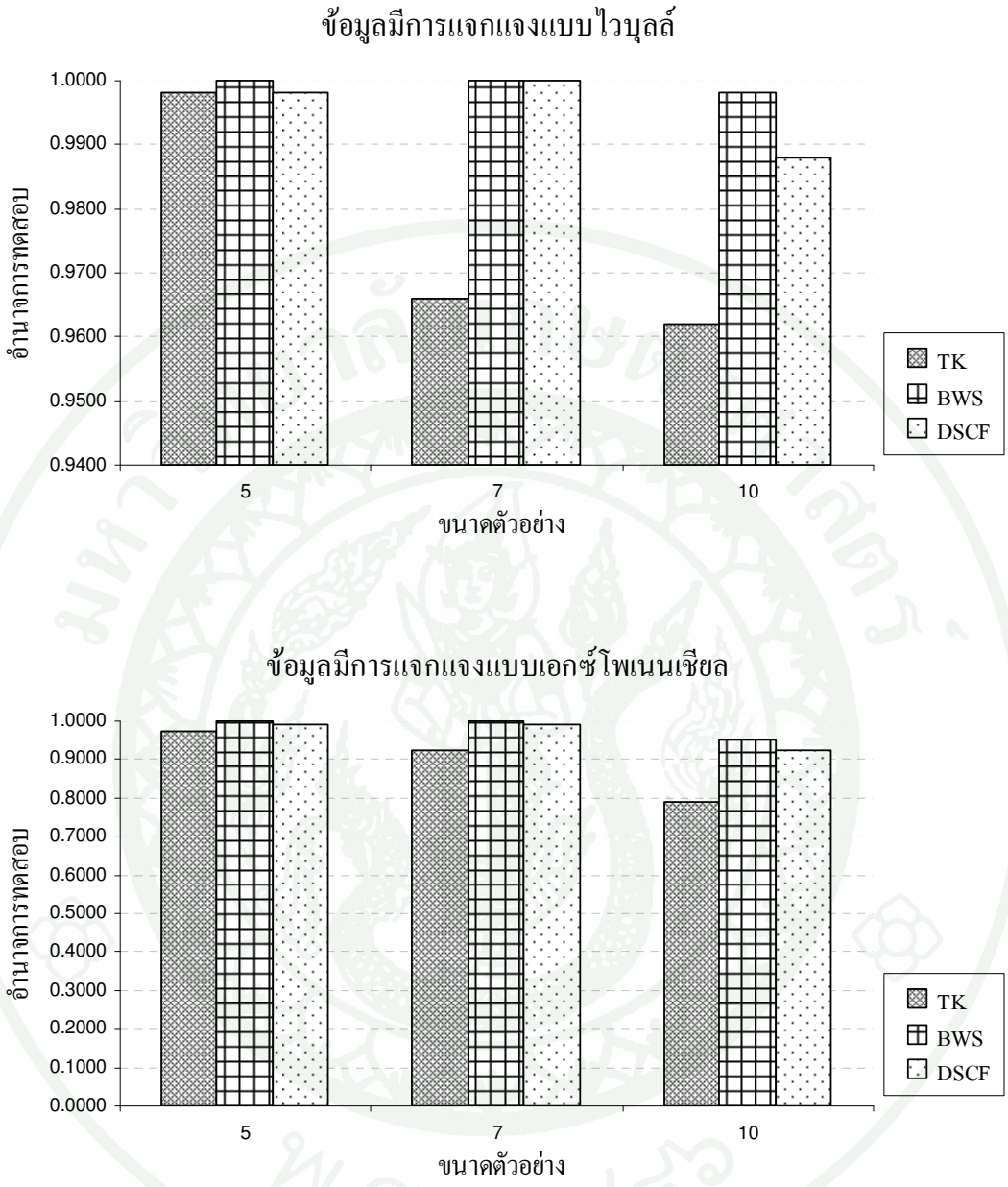
จากตารางที่ 10 พบว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าระหว่าง 0.7900 ถึง 1 สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีค่าระหว่าง 0.952 ถึง 1 และอำนาจการทดสอบของสถิติทดสอบ DSCF มีค่าระหว่าง 0.924 ถึง 1 และโดยภาพรวมพบว่า สถิติทดสอบทั้ง 3 วิธี มีอำนาจการทดสอบที่ค่อนข้างสูงและอำนาจการทดสอบของสถิติทดสอบ TK ต่ำกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 8 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 4 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 8 (ต่อ)

จากภาพที่ 8 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติอำนาจการทดสอบของสถิติทดสอบ TK มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนกรณีที่ข้อมูลมีการแจกแจงแบบเบตาอำนาจการทดสอบของสถิติทดสอบ TK และสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่าง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

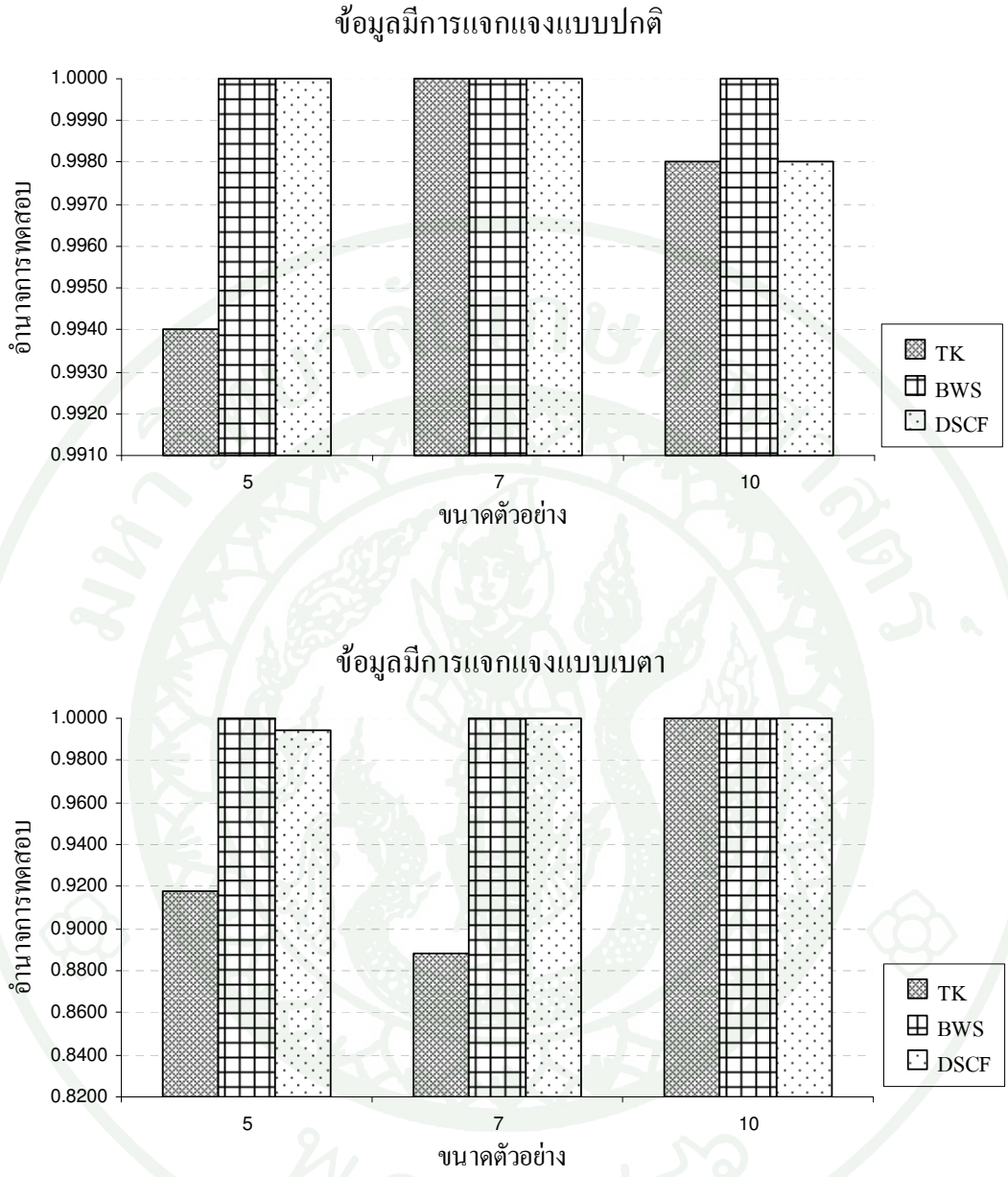
เพิ่มขึ้น และกรณีที่ข้อมูลมีการแจกแจงแบบไวบุลล์และแบบเอกซ์โพเนนเชียลพบว่า อำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธี มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

ตารางที่ 11 แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

การแจกแจง	ขนาดตัวอย่าง	อำนาจการทดสอบ		
		TK	BWS	DSCF
ปกติ	5	0.9940	1.0000	1.0000
	7	1.0000	1.0000	1.0000
	10	0.9980	1.0000	0.9980
เบตา	5	0.9180	1.0000	0.9940
	7	0.8880	1.0000	1.0000
	10	1.0000	1.0000	1.0000
ไวบุลล์	5	0.9960	1.0000	1.0000
	7	0.9840	1.0000	1.0000
	10	0.9980	1.0000	1.0000
เอกซ์โพเนนเชียล	5	0.5560	0.9980	0.9920
	7	0.6160	1.0000	0.9980
	10	1.0000	0.8640	0.8600

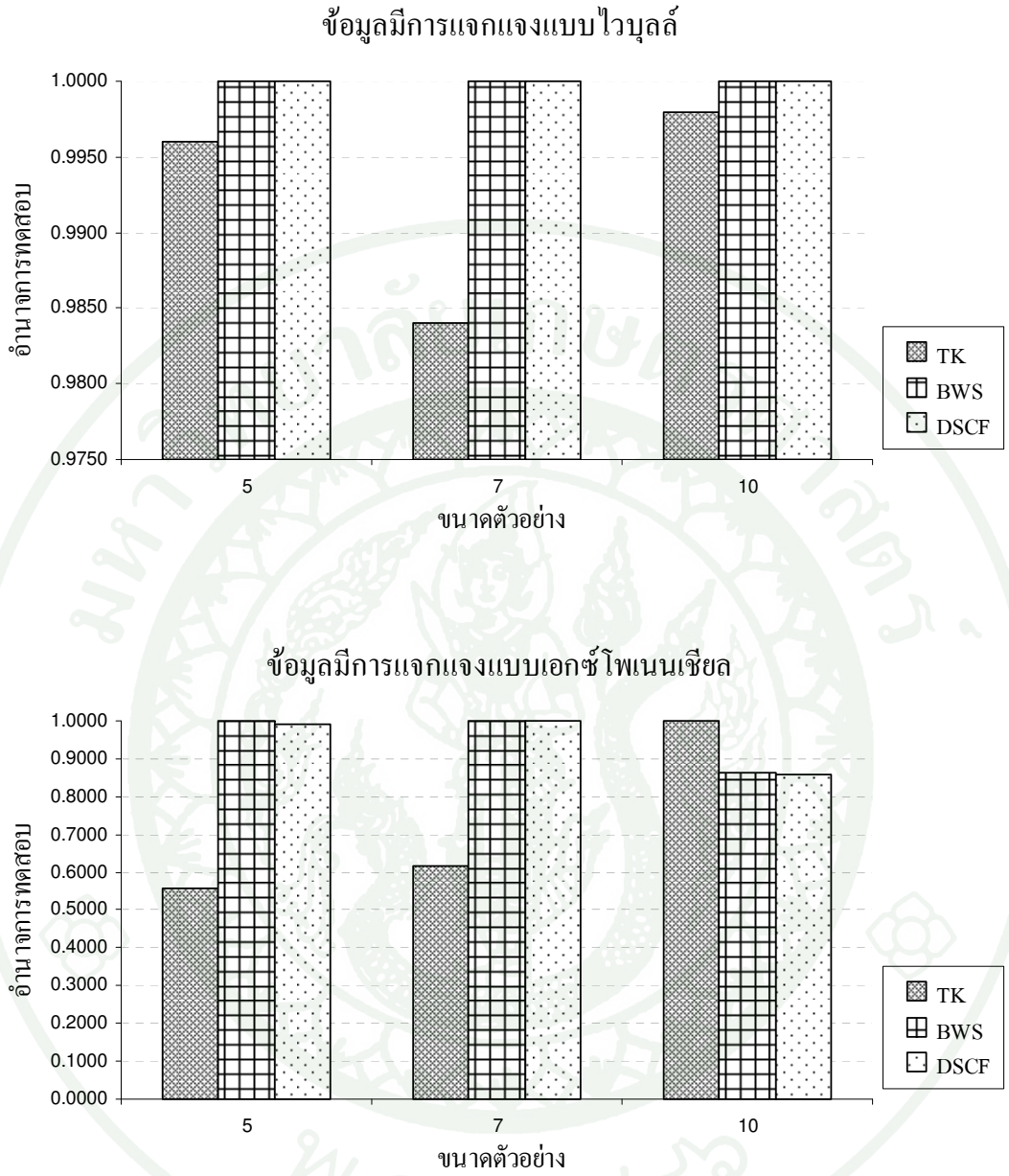
จากตารางที่ 11 พบว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าระหว่าง 0.5560 ถึง 1 สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีค่าระหว่าง 0.864 ถึง 1 และอำนาจการทดสอบของสถิติทดสอบ DSCF มีค่าระหว่าง 0.860 ถึง 1 และโดยภาพรวมพบว่า อำนาจการทดสอบของสถิติทดสอบ TK ต่ำกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 9 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างเท่ากัน สำหรับการเปรียบเทียบ 5 กลุ่ม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 9 (ต่อ)

จากภาพที่ 9 แสดงให้เห็นว่า เมื่อข้อมูลมีการแจกแจงแบบปกติอำนาจการทดสอบของสถิติทดสอบ DSCF มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนกรณีข้อมูลมีการแจกแจงแบบเบตาอำนาจการทดสอบของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น สำหรับกรณีที่ข้อมูลมีการแจกแจงแบบเอกซ์โพเนนเชียลพบว่า อำนาจการทดสอบของสถิติทดสอบ TK มี

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

แนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น ส่วนอำนาจการทดสอบของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มลดลงเมื่อขนาดตัวอย่างเพิ่มขึ้น

2. ผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Tukey-Karmer (TK) สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และสถิติทดสอบ Dwass - Steel และ Critchlow - Fligner (DSCF) กรณีขนาดตัวอย่างไม่เท่ากัน

การศึกษาผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF พิจารณาจากอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ อัตราความคลาดเคลื่อนต่อการทดสอบ และอำนาจการทดสอบ กรณีขนาดตัวอย่างไม่เท่ากัน แสดงรายละเอียดดังต่อไปนี้

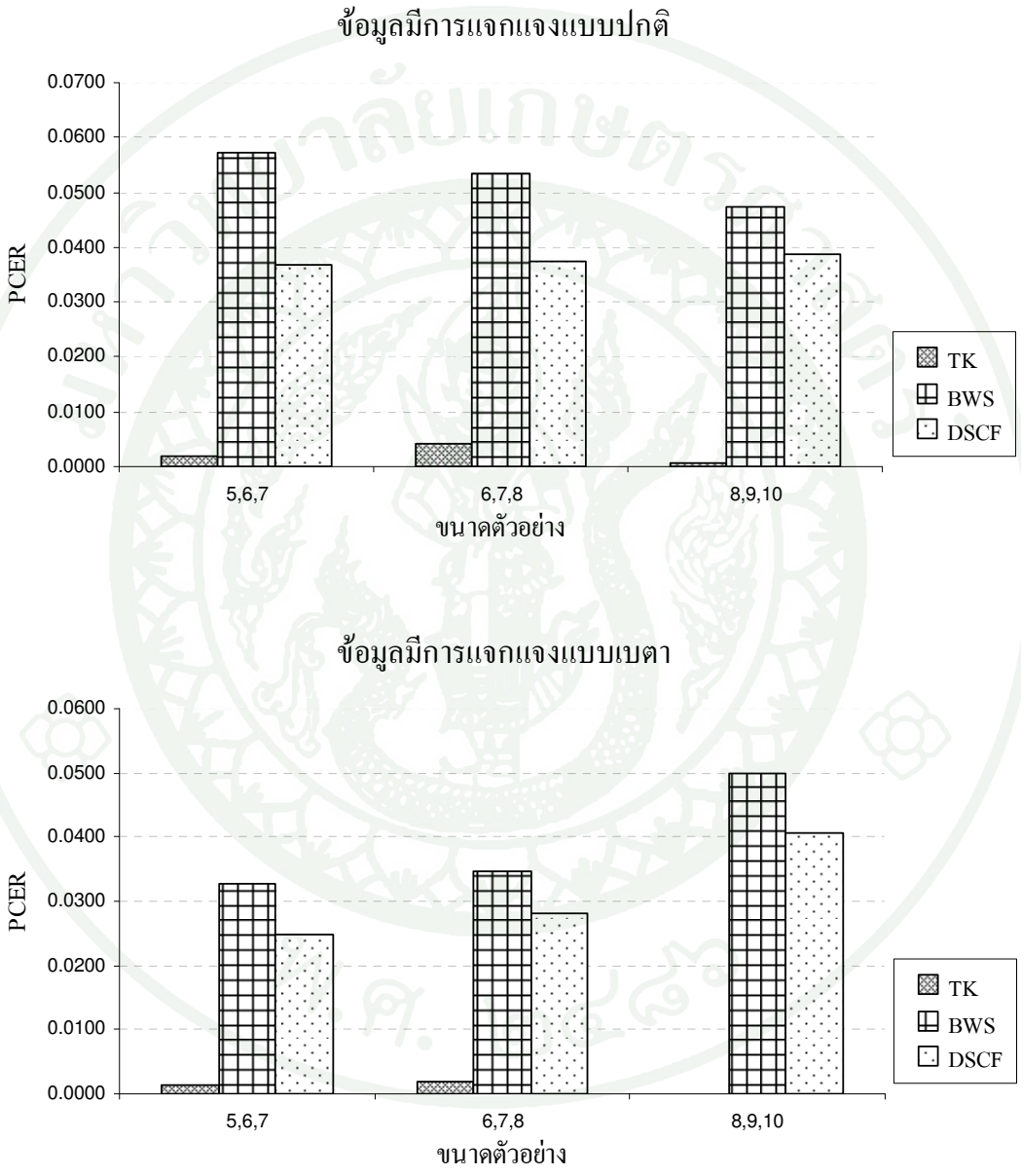
ตารางที่ 12 แสดงค่า PCER กรณีขนาดตัวอย่างไม่เท่ากัน

การแจกแจง	ขนาดตัวอย่าง	ค่า PCER		
		TK	BWS	DSCF
ปกติ	5,6,7	0.0020	0.0573	0.0367
	6,7,8	0.0040	0.0533	0.0373
	8,9,10	0.0007	0.0473	0.0387
เบตา	5,6,7	0.0013	0.0327	0.0247
	6,7,8	0.0020	0.0347	0.0280
	8,9,10	0.0000	0.0500	0.0407
ไวบุลล์	5,6,7	0.0000	0.0327	0.0180
	6,7,8	0.0000	0.0340	0.0253
	8,9,10	0.0000	0.0313	0.0260
เอกซ์โพเนนเชียล	5,6,7	0.0000	0.0347	0.0253
	6,7,8	0.0020	0.0260	0.0160
	8,9,10	0.0007	0.0373	0.0340

จากตารางที่ 12 พบว่า ค่า PCER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.0040 สำหรับค่า PCER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.026 ถึง 0.0573 และค่า PCER ของสถิติทดสอบ

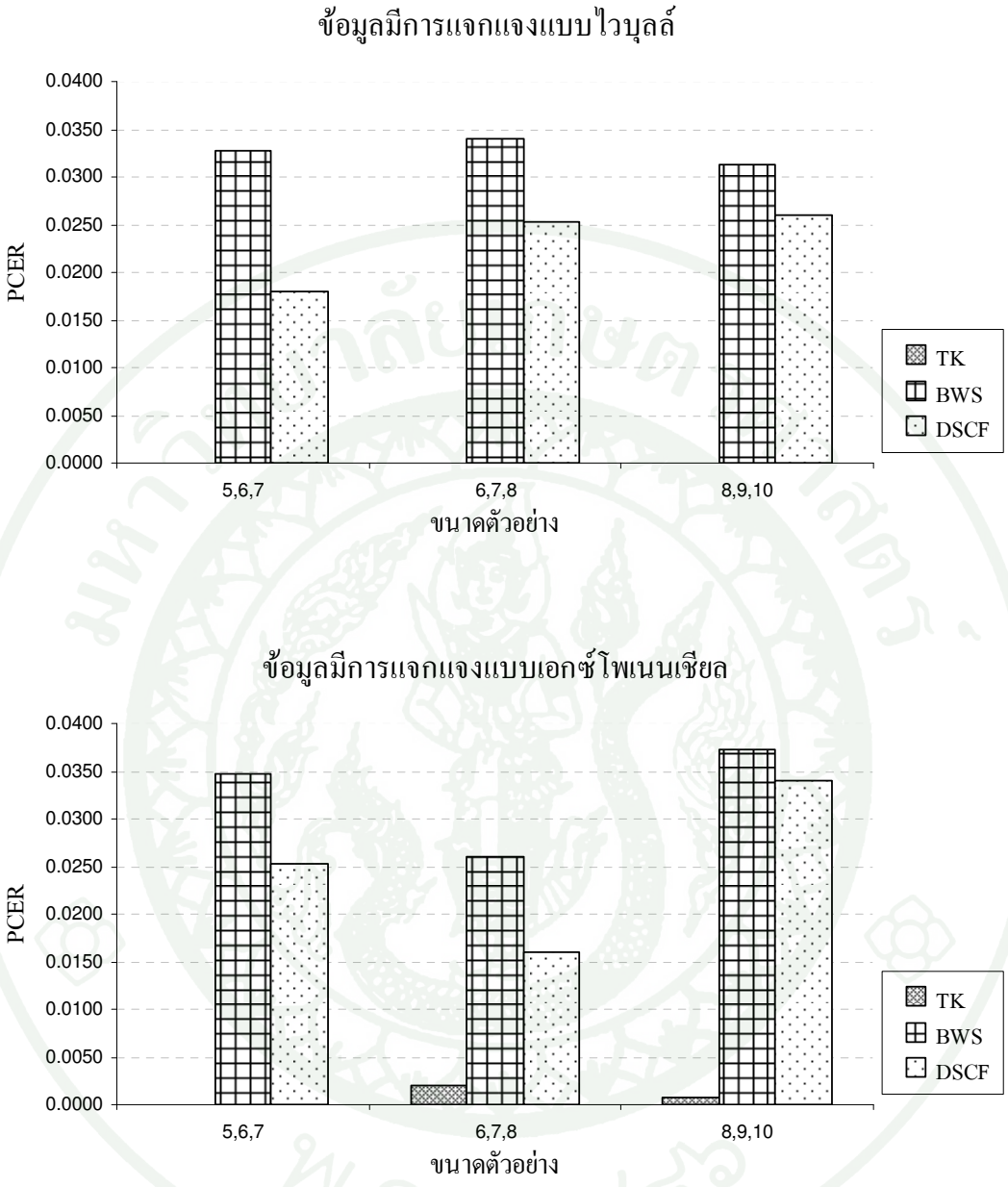
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

DSCF มีค่าระหว่าง 0.016 ถึง 0.0407 โดยภาพรวมพบว่า ค่า PCER ของสถิติทดสอบ BWS มีค่าใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่นทุกการแจกแจง และค่า PCER ของสถิติทดสอบ DSCF ส่วนใหญ่มีค่าใกล้เคียง 0.05 เช่นกัน



ภาพที่ 10 แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 10 (ต่อ)

จากภาพที่ 10 แสดงให้เห็นว่า ค่า PCER ของสถิติทดสอบ BWS มีค่ามากกว่าสถิติทดสอบอื่นทุกกรณี

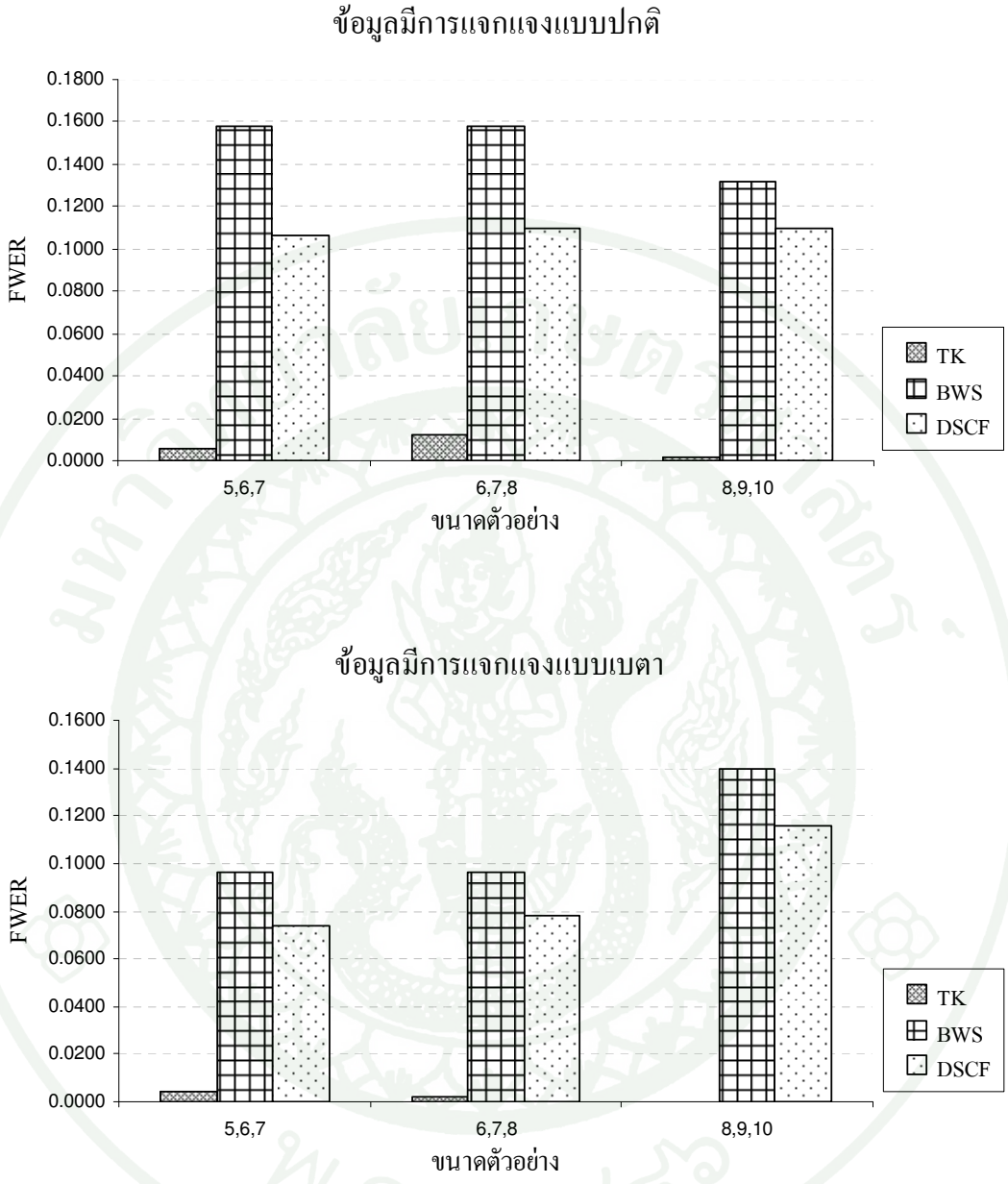
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 13 แสดงค่า FWER กรณีขนาดตัวอย่างไม่เท่ากัน

การแจกแจง	ขนาดตัวอย่าง	ค่า FWER		
		TK	BWS	DSCF
ปกติ	5,6,7	0.0060	0.1580	0.1060
	6,7,8	0.0120	0.1580	0.1100
	8,9,10	0.0020	0.1320	0.1100
เบตา	5,6,7	0.0040	0.0960	0.0740
	6,7,8	0.0020	0.0960	0.0780
	8,9,10	0.0000	0.1400	0.1160
ไวบุลล์	5,6,7	0.0000	0.0960	0.0540
	6,7,8	0.0000	0.0940	0.0700
	8,9,10	0.0000	0.0920	0.0760
เอกซ์โพเนนเชียล	5,6,7	0.0000	0.1000	0.0740
	6,7,8	0.0040	0.0760	0.0480
	8,9,10	0.0020	0.0980	0.0900

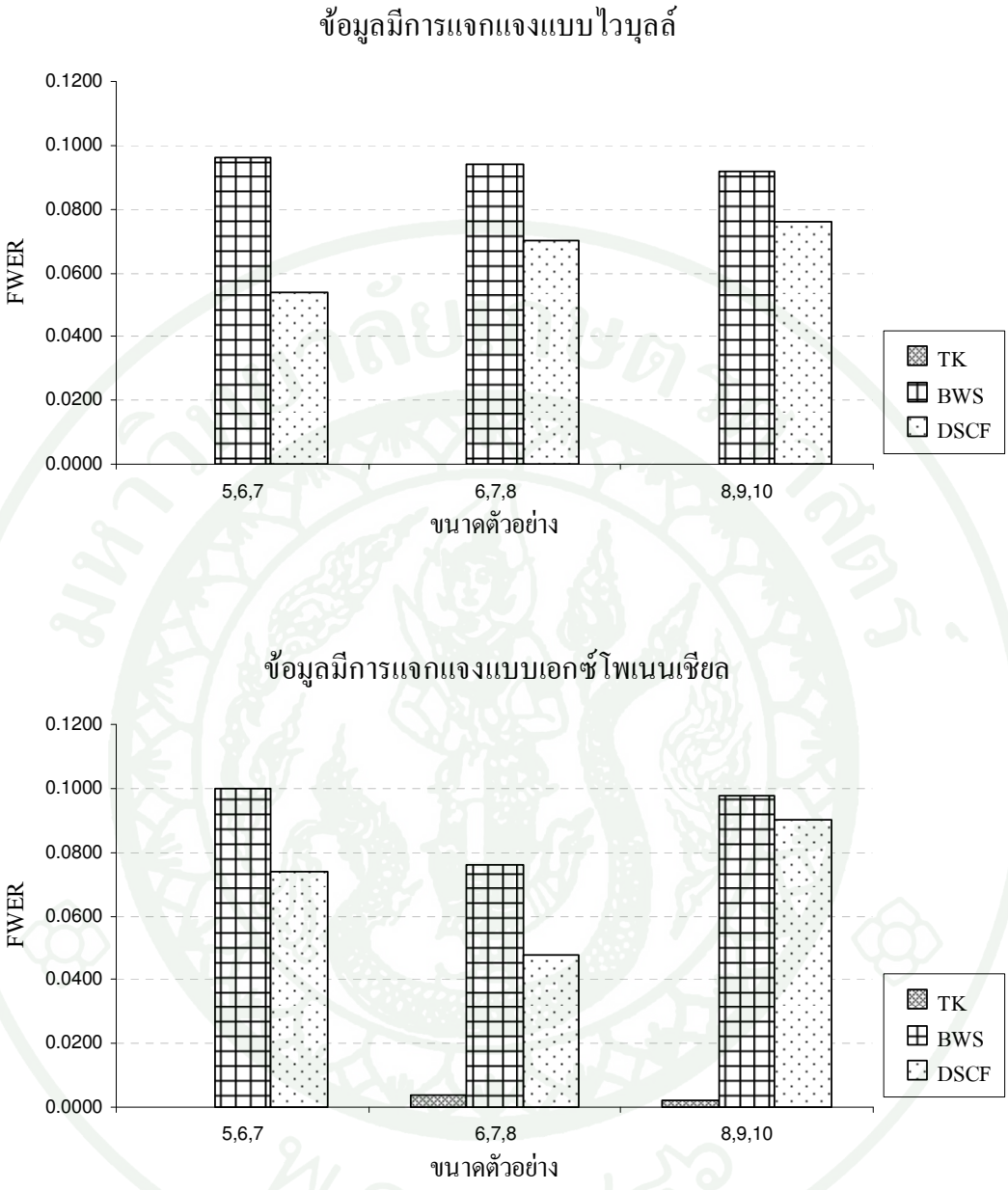
จากตารางที่ 13 พบว่า ค่า FWER ของสถิติทดสอบ TK มีค่าระหว่าง 0 ถึง 0.012 สำหรับค่า FWER ของสถิติทดสอบ BWS มีค่าระหว่าง 0.0760 ถึง 0.1580 และค่า FWER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.048 ถึง 0.1160 และโดยภาพรวมพบว่า เมื่อข้อมูลมีการแจกแจงแบบปกติสถิติทดสอบ TK มีค่า FWER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น ส่วนกรณีข้อมูลมีการแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล สถิติทดสอบ DSCF มีค่า FWER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 11 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 11 (ต่อ)

จากภาพที่ 11 แสดงให้เห็นว่า ค่า FWER ของสถิติทดสอบ BWS มีค่ามากกว่าสถิติทดสอบอื่นทุกกรณี

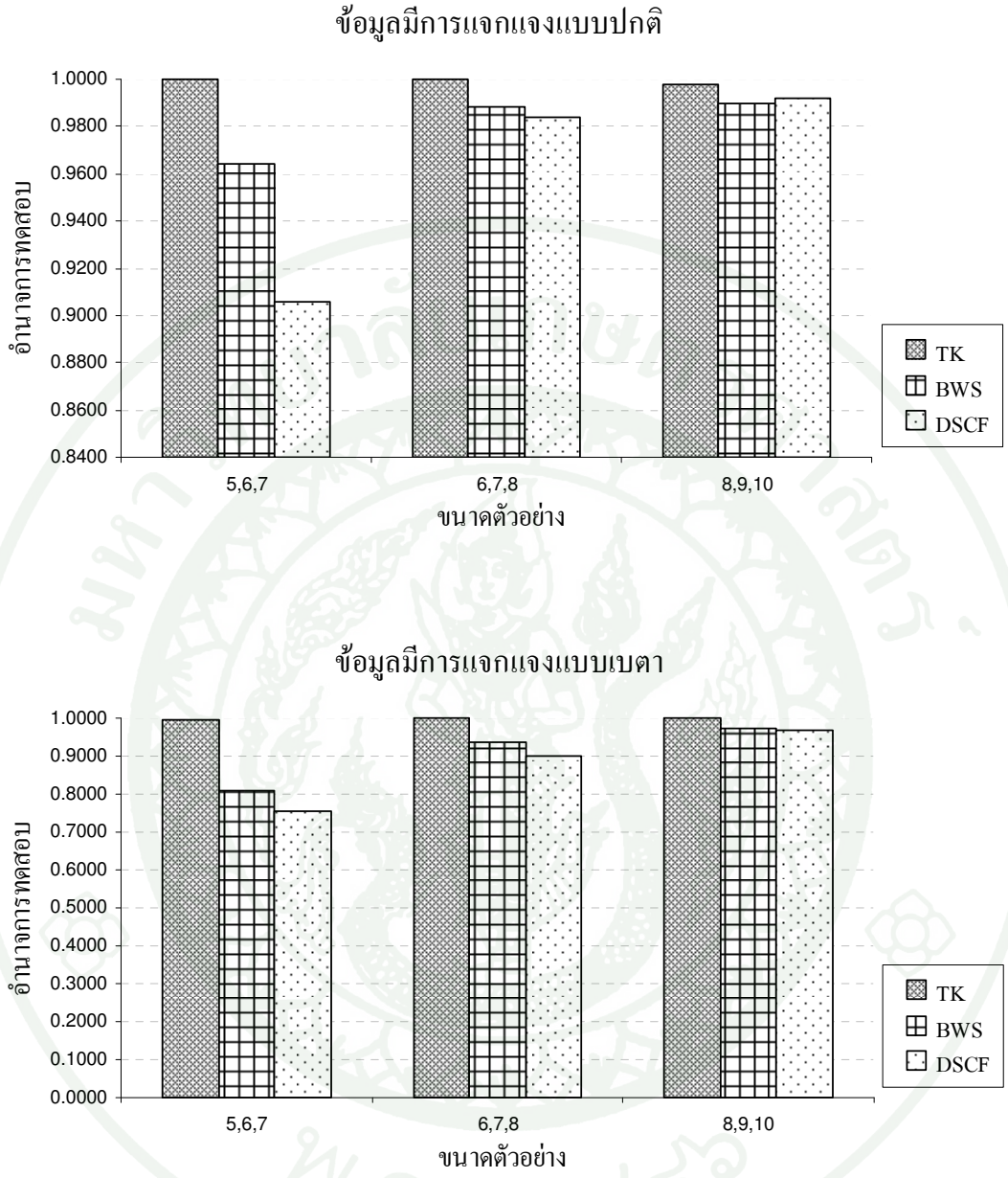
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ตารางที่ 14 แสดงอำนาจการทดสอบ กรณีขนาดตัวอย่างไม่เท่ากัน

การแจกแจง	ขนาดตัวอย่าง	อำนาจการทดสอบ		
		TK	BWS	DSCF
ปกติ	5,6,7	1.0000	0.9640	0.9060
	6,7,8	1.0000	0.9880	0.9840
	8,9,10	0.9980	0.9900	0.9920
เบตา	5,6,7	0.9940	0.8100	0.7540
	6,7,8	0.9980	0.9380	0.9000
	8,9,10	1.0000	0.9720	0.9700
ไวบุลล์	5,6,7	0.9880	0.9220	0.8940
	6,7,8	1.0000	0.9840	0.9720
	8,9,10	0.9900	0.9600	0.9580
เอกซ์โพเนนเชียล	5,6,7	0.9720	0.8320	0.7860
	6,7,8	0.9860	0.8780	0.8060
	8,9,10	0.9640	0.9300	0.9060

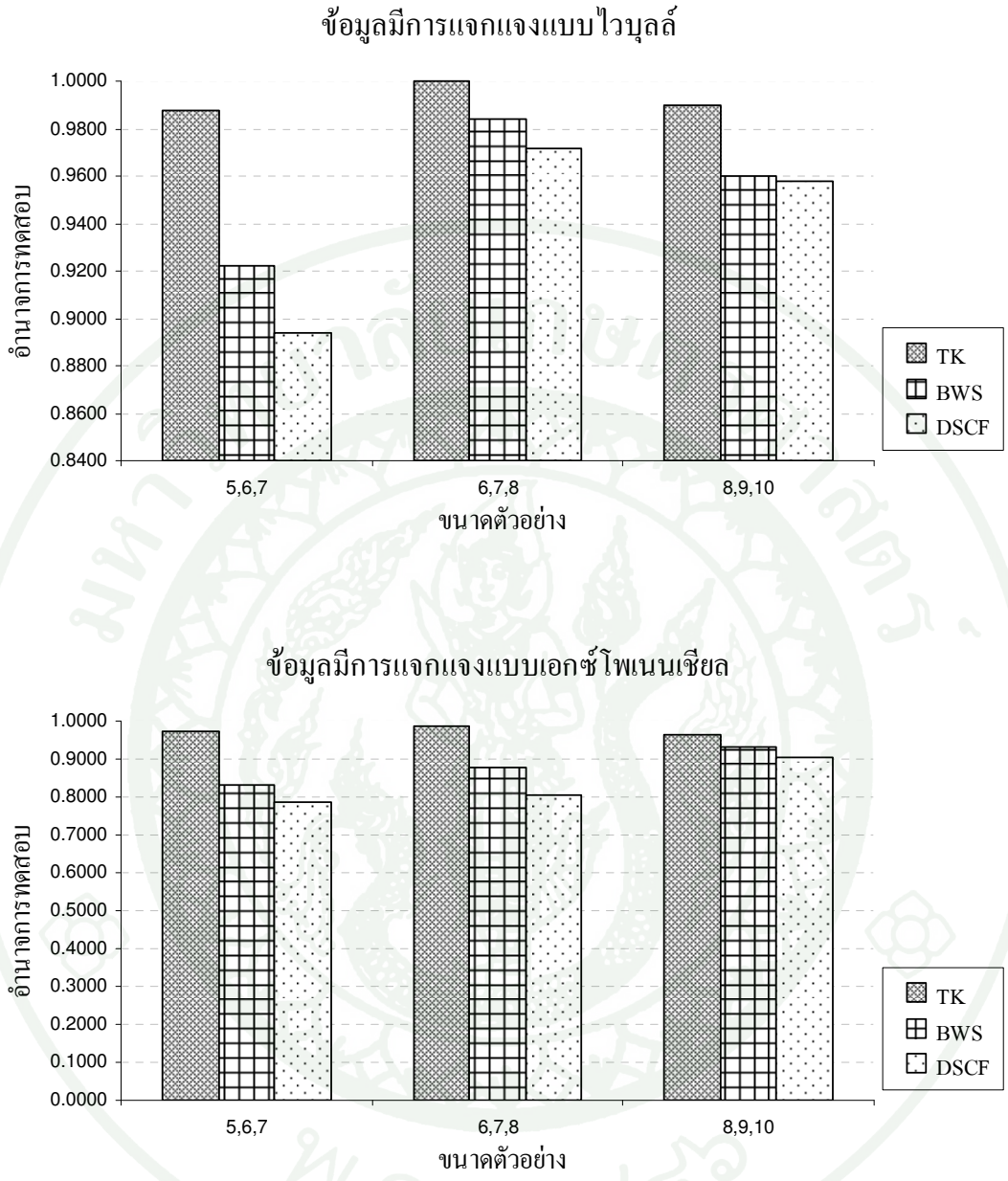
จากตารางที่ 14 พบว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าระหว่าง 0.964 ถึง 1 สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีค่าระหว่าง 0.81 ถึง 0.99 และอำนาจการทดสอบของสถิติทดสอบ DSCF มีค่าระหว่าง 0.754 ถึง 0.992 และโดยภาพรวมพบว่า สถิติทดสอบ TK มีอำนาจการทดสอบของสูงกว่าสถิติทดสอบอื่น ส่วนสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีอำนาจการทดสอบใกล้เคียงกัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 12 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีขนาดตัวอย่างไม่เท่ากัน

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 12 (ต่อ)

จากภาพที่ 12 แสดงให้เห็นว่า อำนาจการทดสอบของสถิติทดสอบ TK สูงกว่าสถิติทดสอบอื่น รองลงมาคือสถิติทดสอบ BWS สำหรับทุกการแจกแจง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

3. ผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Tukey-Karmer (TK) สถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF) กรณีข้อมูลมีค่าเฉลี่ยที่ต่างกันอย่างน้อย 1 คู่

การศึกษาผลการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ TK สถิติทดสอบ BWS และ สถิติทดสอบ DSCF พิจารณาจากอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ อัตราความคลาดเคลื่อนต่อการทดสอบ และอำนาจการทดสอบ กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ แสดงรายละเอียดดังต่อไปนี้

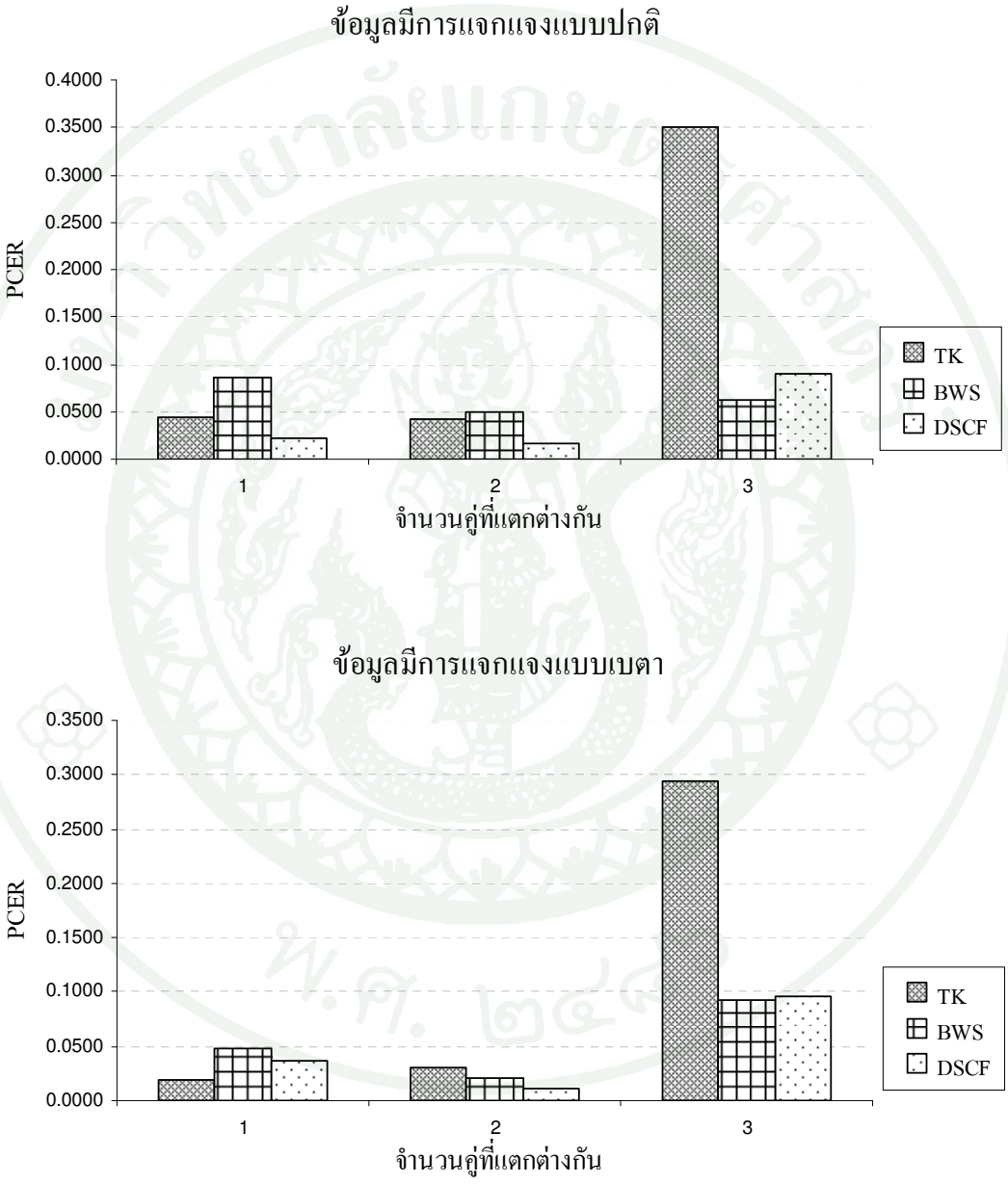
ตารางที่ 15 แสดงค่า PCER กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

การแจกแจง	จำนวนคู่ที่แตกต่างกัน	ค่า PCER		
		TK	BWS	DSCF
ปกติ	1	0.0447	0.0867	0.0227
	2	0.0420	0.0500	0.0170
	3	0.3500	0.0627	0.0907
เบตา	1	0.0190	0.0487	0.0360
	2	0.0300	0.0207	0.0106
	3	0.2947	0.0920	0.0960
ไวบุลล์	1	0.0633	0.0600	0.0090
	2	0.0520	0.0150	0.0087
	3	0.3093	0.0000	0.0007
เอกซ์โพเนนเชียล	1	0.0427	0.0453	0.0447
	2	0.0513	0.0093	0.0060
	3	0.3367	0.1720	0.2173

จากตารางที่ 15 พบว่า ค่า PCER ของสถิติทดสอบ TK มีค่าระหว่าง 0.0190 ถึง 0.3500 สำหรับค่า PCER ของสถิติทดสอบ BWS มีค่าระหว่าง 0 ถึง 0.1720 และค่า PCER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.0007 ถึง 0.2173 และโดยภาพรวมพบว่า สถิติทดสอบ TK จะมีค่า PCER สูงสุดเมื่อมีจำนวนคู่ที่แตกต่างกันเท่ากับ 3 คู่ สำหรับทุกการแจกแจง ส่วนสถิติทดสอบ BWS และ

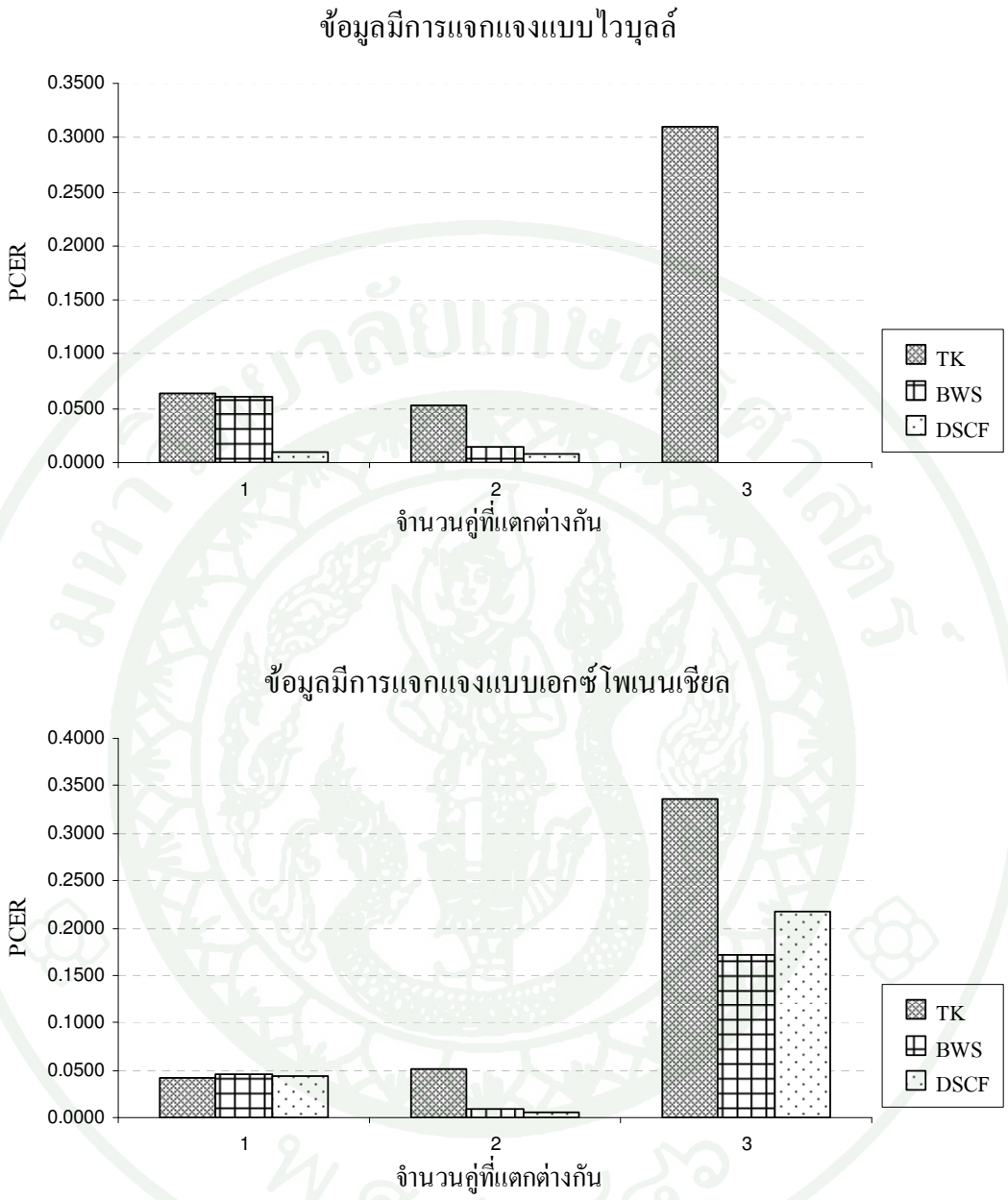
ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สถิติทดสอบ DSCF จะมีค่า PCER ลดลงเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้นเมื่อข้อมูลแตกต่างกัน 1 ถึง 2 คู่ แต่จะมีค่าสูงขึ้นเมื่อจำนวนคู่ที่แตกต่างกันเท่ากับ 3 คู่ ยกเว้นกรณีที่ข้อมูลมีการแจกแจงแบบไวบุลล์



ภาพที่ 13 แสดงการเปรียบเทียบค่า PCER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 13 (ต่อ)

จากภาพที่ 13 แสดงให้เห็นว่า ค่า PCER ของสถิติทดสอบ TK มีแนวโน้มสูงขึ้นเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น สำหรับค่า PCER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มลดลงเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ยกเว้นกรณีที่ข้อมูลแตกต่างกันทั้ง 3 คู่ เมื่อข้อมูลมีการแจกแจงแบบเบตาและการแจกแจงแบบเอกซ์โพเนนเชียลค่า PCER ของสถิติ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

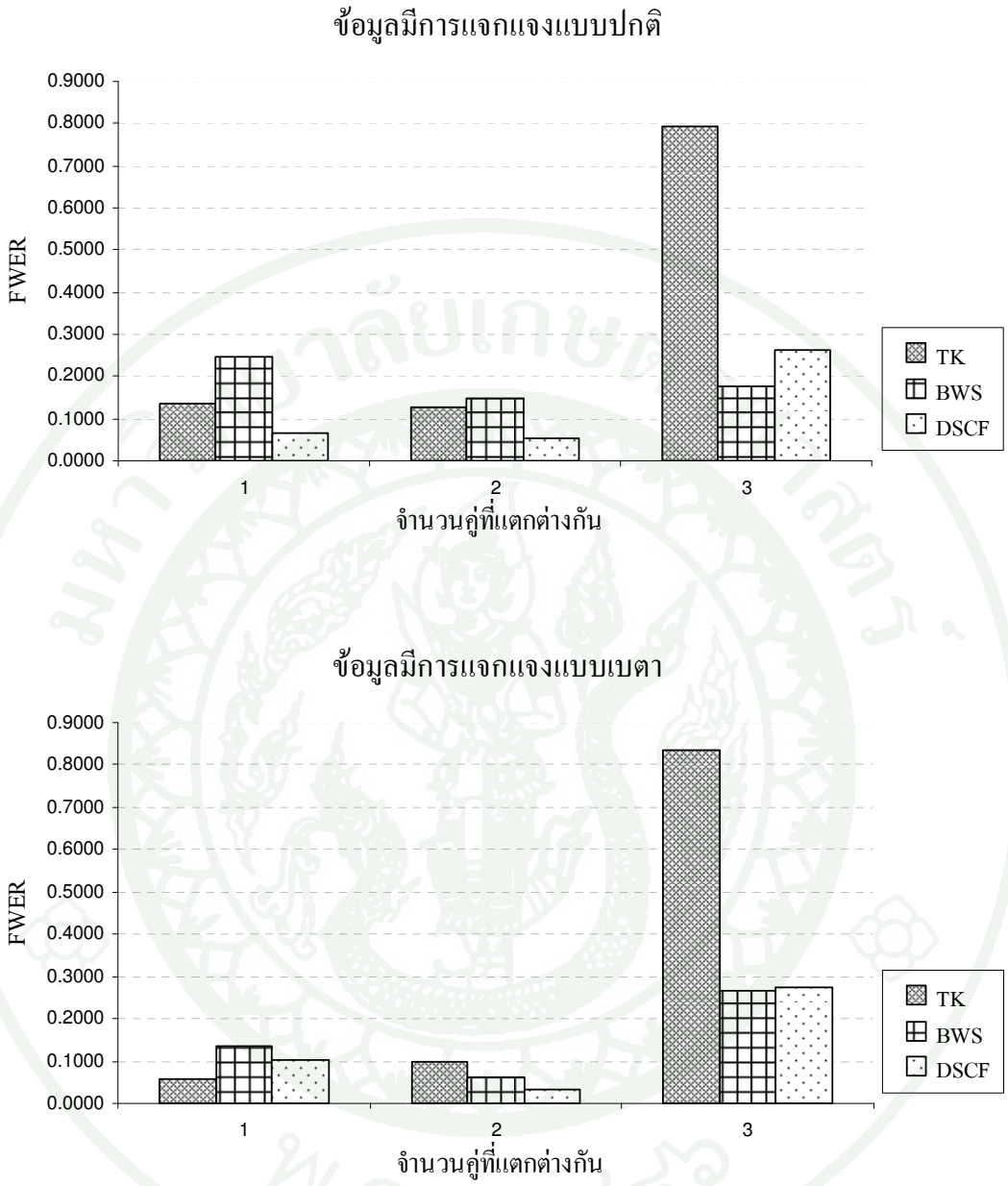
ทดสอบ BWS และสถิติทดสอบ DSCF จะมีค่าสูงที่สุดเมื่อเทียบกับกรณีที่ข้อมูลมีจำนวนคู่แตกต่างน้อยกว่าและมีการแจกแจงเดียวกัน

ตารางที่ 16 แสดงค่า FWER กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

การแจกแจง	จำนวนคู่ที่แตกต่างกัน	ค่า FWER		
		TK	BWS	DSCF
ปกติ	1	0.1340	0.2480	0.0660
	2	0.1260	0.1500	0.0520
	3	0.7920	0.1780	0.2620
เบตา	1	0.0580	0.1360	0.1020
	2	0.1000	0.0620	0.0320
	3	0.8360	0.2640	0.2760
ไวบุลล์	1	0.1900	0.1800	0.0280
	2	0.1560	0.0460	0.0260
	3	0.9280	0.0000	0.0020
เอกซ์โพเนนเชียล	1	0.1280	0.1340	0.1320
	2	0.1420	0.0280	0.0180
	3	0.9840	0.4420	0.5480

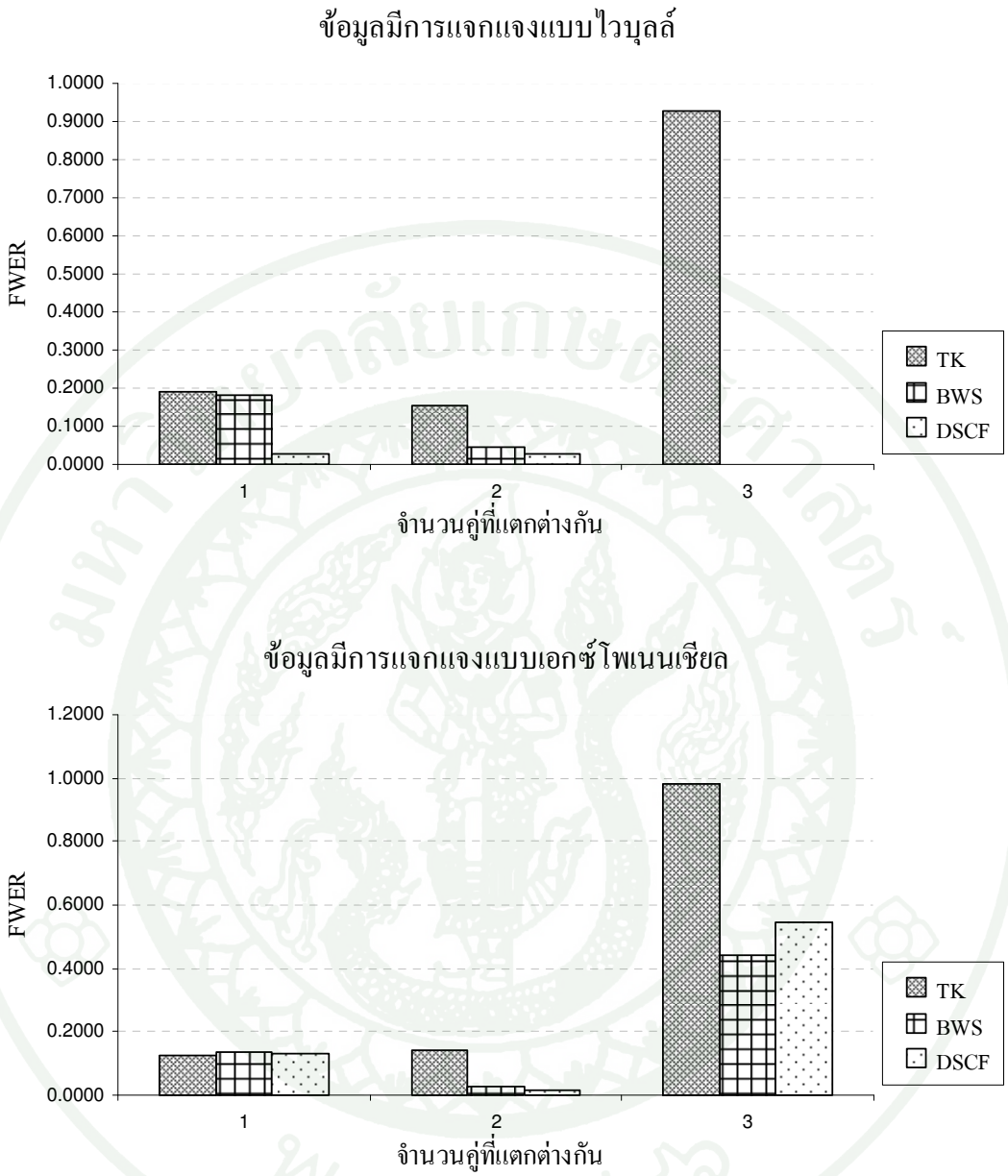
จากตารางที่ 16 พบว่า ค่า FWER ของสถิติทดสอบ TK มีค่าระหว่าง 0.058 ถึง 0.984 สำหรับค่า FWER ของสถิติทดสอบ BWS มีค่าระหว่าง 0 ถึง 0.442 และค่า FWER ของสถิติทดสอบ DSCF มีค่าระหว่าง 0.002 ถึง 0.548 ซึ่งโดยภาพรวมพบว่า สถิติทดสอบ TK จะมีค่า FWER สูงสุดเมื่อมีจำนวนคู่ที่แตกต่างกัน 3 คู่ สำหรับทุกการแจกแจง ส่วนสถิติทดสอบ BWS จะมีค่า FWER สูงสุดเมื่อมีจำนวนคู่ที่แตกต่างกัน 3 คู่ เฉพาะกรณีที่ข้อมูลมีการแจกแจงแบบเบตา และการแจกแจงแบบเอกซ์โพเนนเชียล และสถิติทดสอบ DSCF มีค่า FWER สูงสุดเมื่อมีจำนวนคู่ที่แตกต่างกัน 3 คู่ เมื่อข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบเบตาและการแจกแจงแบบเอกซ์โพเนนเชียล

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 14 แสดงการเปรียบเทียบค่า FWER ของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 14 (ต่อ)

จากภาพที่ 14 แสดงให้เห็นว่า ค่า FWER ของสถิติทดสอบ TK มีแนวโน้มเพิ่มขึ้นเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ส่วนค่า FWER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มลดลงเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ยกเว้นกรณีที่ข้อมูลมีค่าเฉลี่ยแตกต่างกันทั้ง 3 คู่ เมื่อข้อมูลมีการแจกแจงแบบเบตาและการแจกแจงแบบเอกซ์โพเนนเชียล

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

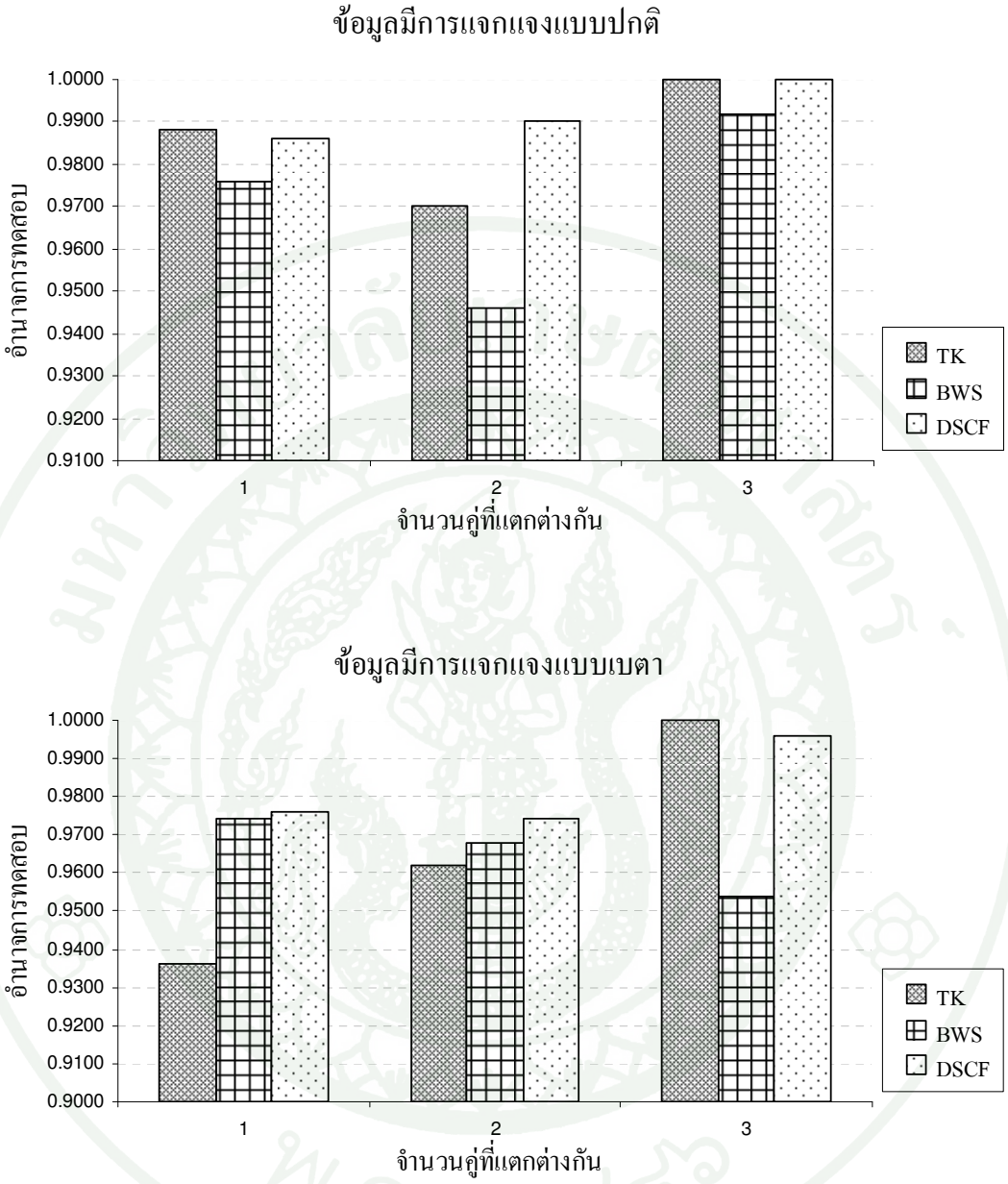
ค่า FWER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF จะมีค่าสูงที่สุดเมื่อเทียบกับกรณีที่ข้อมูลมีจำนวนคู่แตกต่างน้อยกว่าและมีการแจกแจงแบบเดียวกัน

ตารางที่ 17 แสดงอำนาจการทดสอบกรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

การแจกแจง	จำนวนคู่ที่แตกต่างกัน	อำนาจการทดสอบ		
		TK	BWS	DSCF
ปกติ	1	0.9880	0.9760	0.9860
	2	0.9700	0.9460	0.9900
	3	1.0000	0.9920	1.0000
เบตา	1	0.9360	0.9740	0.9760
	2	0.9620	0.9680	0.9740
	3	1.0000	0.9540	0.9960
ไวบุลล์	1	0.9980	0.9800	0.9820
	2	1.0000	0.9900	0.9900
	3	1.0000	0.9760	0.9860
เอกซ์โพเนนเชียล	1	0.9720	0.9700	0.9640
	2	1.0000	0.9840	0.9760
	3	1.0000	0.9980	0.9980

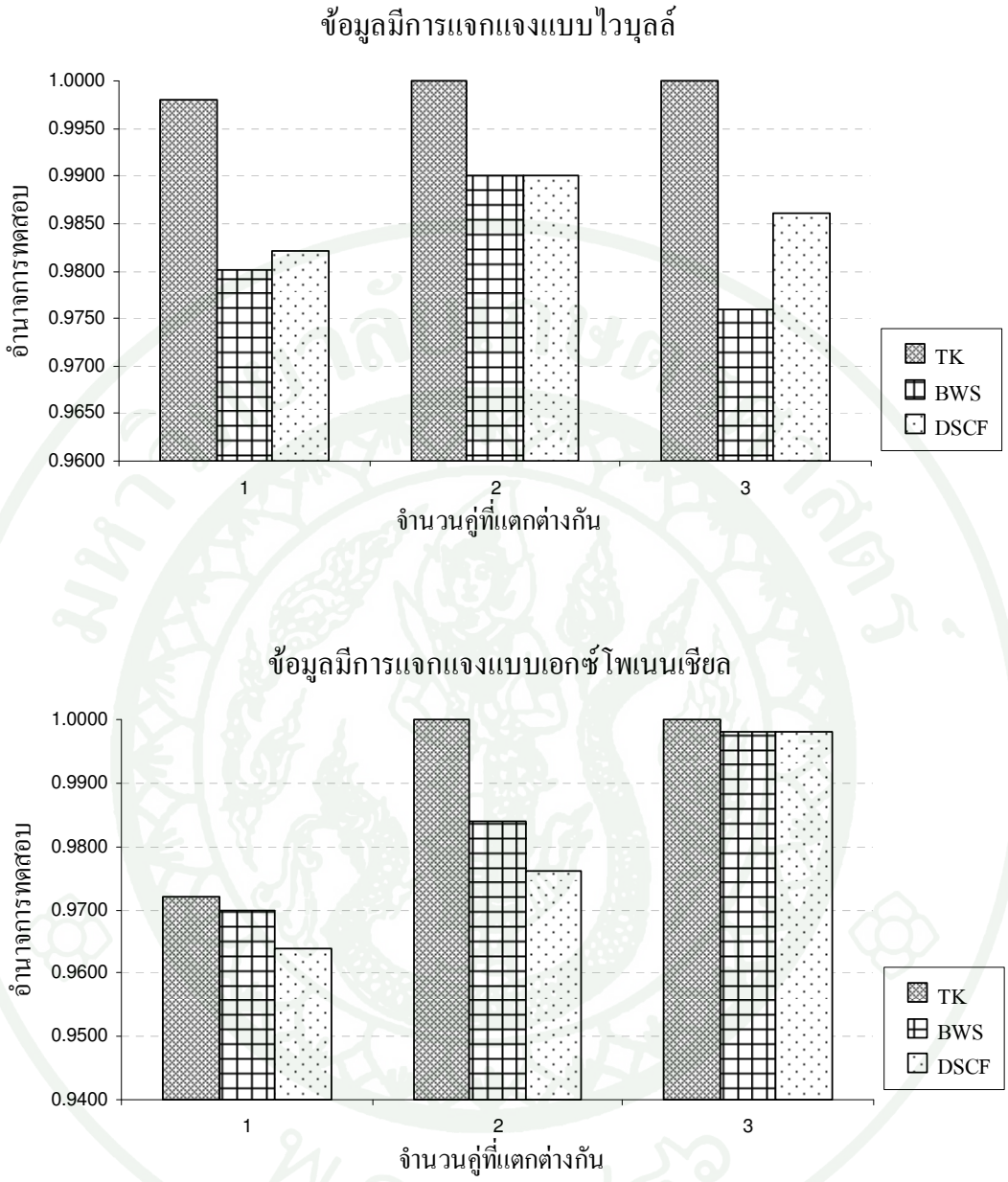
จากตารางที่ 17 พบว่า อำนาจการทดสอบของสถิติทดสอบ TK มีค่าระหว่าง 0.936 ถึง 1 สำหรับอำนาจการทดสอบของสถิติทดสอบ BWS มีค่าระหว่าง 0.946 ถึง 0.998 และอำนาจการทดสอบของสถิติทดสอบ DSCF มีค่าระหว่าง 0.964 ถึง 1 และโดยส่วนใหญ่พบว่า สถิติทดสอบ TK จะมีอำนาจการทดสอบสูงกว่าสถิติทดสอบอื่น ยกเว้นกรณีที่ข้อมูลมีการแจกแจงแบบเบตาพบว่าสถิติทดสอบ DSCF จะมีอำนาจการทดสอบของสูงกว่าสถิติทดสอบอื่น

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 15 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ TK สถิติทดสอบ BWS และสถิติทดสอบ DSCF กรณีข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์



ภาพที่ 15 (ต่อ)

จากภาพที่ 15 แสดงให้เห็นว่า อำนาจการทดสอบของสถิติทดสอบ TK มีแนวโน้มเพิ่มขึ้นเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ส่วนอำนาจการทดสอบของสถิติทดสอบ BWS มีแนวโน้มลดลงเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้นและมีการแจกแจงแบบเบตา สำหรับอำนาจการทดสอบของสถิติทดสอบ DSCF มีแนวโน้มเพิ่มขึ้นเมื่อข้อมูลมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้นและมีการแจกแจงแบบปกติ และสำหรับกรณีที่ข้อมูลมีการแจกแจงแบบเอกซ์โพเนนเชียล

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

พบว่า อำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธี มีแนวโน้มเพิ่มขึ้นเมื่อมีจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น

วิจารณ์

1. การเปรียบเทียบที่มีจำนวนการเปรียบเทียบมากจะส่งผลทำให้ค่า FWER มากขึ้น ซึ่งสอดคล้องกับการศึกษาของ Hochberg and Tamhane (1987) แต่จะไม่ส่งผลต่อค่า PCER จึงทำให้ค่าทั้งสองแตกต่างกันมากขึ้นเมื่อมีจำนวนการเปรียบเทียบที่มากขึ้น

2. กรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันทุกคู่ สถิติทดสอบทั้ง 3 วิธี จะมีค่า PCER และค่า FWER สูงกว่ากรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกัน 1 และ 2 คู่ ยกเว้นเมื่อข้อมูลมีการแจกแจงแบบไวบุลล์

3. ค่า PCER และค่า FWER ของสถิติทดสอบ TK มีค่าต่ำทุกกรณี เนื่องจากข้อมูลที่ใช้ในการวิจัยครั้งนี้มีขนาดเล็กและมีการแจกแจงที่ไม่ได้เป็นไปตามข้อตกลงเบื้องต้นของสถิติทดสอบ TK จึงทำให้ประสิทธิภาพของการทดสอบสมมติฐานด้วยสถิติทดสอบ TK ต่ำ

4. อำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธีส่วนใหญ่มีค่าสูง เนื่องจากเมื่อตรวจพบความแตกต่างอย่างน้อย 1 คู่ จะปฏิเสธสมมติฐานหลักแม้ในการทดสอบสมมติฐานบางครั้งที่สถิติทดสอบทั้ง 3 วิธี จะให้ผลการปฏิเสธสมมติฐานหลักต่างคู่การเปรียบเทียบกันก็ตาม

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

สรุปและข้อเสนอแนะ

สรุป

การศึกษาสถิติทดสอบสำหรับการเปรียบเทียบพหุคูณเมื่อข้อมูลมีการแจกแจงไม่ใช่แบบปกติด้วยวิธีการทดสอบ 3 วิธี คือ วิธีการทดสอบด้วยสถิติทดสอบ Tukey-Karmer (TK) วิธีการทดสอบด้วยสถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) และวิธีการทดสอบด้วยสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF) ซึ่งผลการวิจัยสรุปได้ดังนี้

1. การเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Tukey-Karmer (TK) มีอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) และอัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) ต่ำมาก นั่นคือไม่สามารถควบคุมอัตราความคลาดเคลื่อนทั้ง 2 ชนิดได้ ทั้งกรณีขนาดตัวอย่างเท่ากันและกรณีขนาดตัวอย่างไม่เท่ากัน ยกเว้นกรณีขนาดตัวอย่างไม่เท่ากันเมื่อข้อมูลมีการแจกแจงแบบปกติพบว่า สถิติทดสอบ TK มีค่า FWER ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น ส่วนกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ พบว่า ค่า PCER และค่า FWER มีแนวโน้มเพิ่มขึ้นตามจำนวนคู่ที่แตกต่างกันที่เพิ่มขึ้นทุกการแจกแจง นอกจากนี้ยังพบว่า กรณีที่ข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล เมื่อข้อมูลมีค่าเฉลี่ยที่แตกต่างกัน 1 และ 2 คู่ สถิติทดสอบ TK มีค่า PCER และค่า FWER ใกล้เคียง 0.05 และเมื่อพิจารณาจากอำนาจการทดสอบพบว่ามีแนวโน้มเพิ่มขึ้นตามจำนวนคู่ที่แตกต่างกันที่เพิ่มขึ้นด้วย หรืออาจกล่าวได้ว่า สถิติทดสอบ TK มีประสิทธิภาพในการตรวจพบความแตกต่างต่ำกว่าสถิติทดสอบอื่น

2. การเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Baumgartner-Weiß-Schindler (BWS) มีอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) ที่ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น ทั้งกรณีที่ขนาดตัวอย่างเท่ากันและกรณีที่ขนาดตัวอย่างไม่เท่ากัน ส่วนอัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) เมื่อข้อมูลมีการแจกแจงแบบเบตา การแจกแจงแบบไวบุลล์ และการแจกแจงแบบเอกซ์โพเนนเชียล กรณีขนาดตัวอย่างเท่ากัน เมื่อจำนวนกลุ่มตัวอย่างเท่ากับ 3 กลุ่ม มีค่าส่วนใหญ่ใกล้เคียง 0.05 และเมื่อจำนวนกลุ่มตัวอย่างมากขึ้นจะมีค่าเพิ่มขึ้นซึ่งสอดคล้องกับการศึกษาของ Hochberg and Tamhane (1987) สำหรับกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ พบว่า ค่า PCER และค่า FWER เมื่อข้อมูลมีการแจกแจงแบบ

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ปกติ และการแจกแจงแบบไวบุลล์ มีแนวโน้มลดลงเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น สำหรับอำนาจการทดสอบพบว่า กรณีที่ข้อมูลมีขนาดไม่เท่ากันและกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ มีค่าใกล้เคียง 0.95 มากกว่าสถิติทดสอบอื่น

3. การเปรียบเทียบพหุคูณด้วยสถิติทดสอบ Dwass-Steel และ Critchlow-Fligner (DSCF) อัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (Per-Comparison error rate: PCER) มีแนวโน้มเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้นเมื่อจำนวนกลุ่มตัวอย่างเท่ากับ 3 และ 5 กลุ่ม นอกจากนี้ยังพบว่า เมื่อจำนวนกลุ่มตัวอย่างเพิ่มขึ้นค่า PCER จะลดลง สำหรับกรณีที่ขนาดตัวอย่างไม่เท่ากันพบว่าโดยส่วนใหญ่ค่า PCER มีค่าใกล้เคียง 0.05 ทุกการแจกแจง สำหรับกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกัน 1 คู่ เมื่อข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบเบตา และการแจกแจงแบบเอกซ์โพเนนเชียล สถิติทดสอบ DSCF มีค่า PCER ใกล้เคียง 0.05 สำหรับอัตราความคลาดเคลื่อนต่อการทดสอบ (Familywise error rate: FWER) พบว่า กรณีจำนวนกลุ่มตัวอย่างเท่ากับ 4 และ 5 กลุ่ม เมื่อข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบเบตา และการแจกแจงแบบไวบุลล์ สถิติทดสอบ DSCF มีค่า FWER ส่วนใหญ่ใกล้เคียง 0.05 มากกว่าสถิติทดสอบอื่น นอกจากนี้ยังพบว่า เมื่อมีจำนวนคู่การเปรียบเทียบเพิ่มขึ้นค่า FWER จะมากขึ้นสอดคล้องกับการศึกษาของ Hochberg and Tamhane (1987) ส่วนกรณีที่ข้อมูลมีค่าเฉลี่ยแตกต่างกันอย่างน้อย 1 คู่ พบว่า ค่า FWER เมื่อข้อมูลมีการแจกแจงแบบไวบุลล์ มีแนวโน้มลดลงเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น สำหรับอำนาจการทดสอบพบว่า กรณีที่ข้อมูลมีขนาดไม่เท่ากันจะมีอำนาจการทดสอบมีค่าต่ำกว่าสถิติทดสอบอื่น

ดังนั้นจึงสามารถสรุปได้ว่า กรณีที่ข้อมูลมีขนาดเท่ากันและกรณีขนาดตัวอย่างไม่เท่ากัน สถิติทดสอบ BWS มีประสิทธิภาพในการควบคุม PCER ได้ดีกว่าสถิติทดสอบอื่น ส่วนสถิติทดสอบ DSCF มีประสิทธิภาพในการควบคุม FWER ได้ดีกว่าสถิติทดสอบอื่น และเมื่อจำนวนคู่การเปรียบเทียบเพิ่มขึ้นค่า FWER ของสถิติทดสอบทั้ง 3 วิธี มีแนวโน้มเพิ่มขึ้น สำหรับกรณีที่ข้อมูลมีค่าเฉลี่ยที่แตกต่างกันอย่างน้อย 1 คู่ พบว่า ค่า PCER และค่า FWER ของสถิติทดสอบ TK มีแนวโน้มเพิ่มขึ้นเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น ส่วนค่า PCER และค่า FWER ของสถิติทดสอบ BWS และสถิติทดสอบ DSCF มีแนวโน้มลดลงเมื่อจำนวนคู่ที่แตกต่างกันเพิ่มขึ้น แต่จะมีค่าเพิ่มขึ้นเมื่อข้อมูลแตกต่างกันทุกคู่และมีการแจกแจงแบบปกติ การแจกแจงแบบเบตา และการแจกแจงแบบเอกซ์โพเนนเชียล ส่วนอำนาจการทดสอบพบว่า สถิติทดสอบทั้ง 3 วิธี มีค่าค่อนข้างสูง

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลการวิจัยไปใช้

1.1. สำหรับการเปรียบเทียบพหุคูณที่ต้องการควบคุมอัตราความคลาดเคลื่อนต่อการเปรียบเทียบ (PCER) ควรเลือกใช้สถิติทดสอบ BWS เนื่องจากมีประสิทธิภาพในการควบคุมค่า PCER ได้ดีกว่าสถิติทดสอบอื่น ยกเว้นกรณีที่จำนวนคู่การเปรียบเทียบมากควรเลือกใช้สถิติทดสอบ DSCF ซึ่งจะมีประสิทธิภาพในการควบคุมค่า PCER ได้ดีกว่า

1.2. สำหรับการเปรียบเทียบพหุคูณที่ต้องการควบคุมอัตราความคลาดเคลื่อนต่อการทดสอบ (FWER) ควรเลือกใช้สถิติทดสอบ DSCF เนื่องจากจะมีประสิทธิภาพในการควบคุมค่า FWER ได้ดีว่าสถิติทดสอบอื่น

1.3. หากข้อมูลมีความแตกต่างกันไม่มากสถิติทดสอบ BWS และสถิติทดสอบ DSCF จะมีประสิทธิภาพในการตรวจพบความแตกต่างดีกว่าสถิติทดสอบ TK

2. ข้อเสนอแนะในการทำวิจัยต่อไป

2.1. ในการศึกษาครั้งต่อไปควรศึกษาสถิติทดสอบทั้ง 3 วิธีกับข้อมูลที่มีการแจกแจงแบบสมมาตร และแบบไม่สมมาตร และพิจารณาค่าความแปรปรวนของข้อมูล ซึ่งอาจมีผลต่อการทดสอบ

2.2. ในการเปรียบเทียบพหุคูณด้วยสถิติทดสอบ BWS และ สถิติทดสอบ DSCF อาจเลือกใช้เทคนิคอื่นแทนวิธี permutation test ที่ใช้ในงานวิจัยนี้ เช่น วิธีบูสแตป หรือ วิธี randomize permutation test

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

เอกสารและสิ่งอ้างอิง

จิรัชย์ สุขะเกตุ. 2548. **ความน่าจะเป็นและทฤษฎีสถิติเบื้องต้น.** สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

ฉัตรศิริ ปิยะพิมลสิทธิ์. 2554. **การเปรียบเทียบพหุคูณ (Multiple Comparison Procedures)**. แหล่งที่มา: http://www.watpon.com, 18 กุมภาพันธ์ 2554.

ธีรศักดิ์ จันทร์กระจ่าง. 2551. **อำนาจการทดสอบ (Power of Test) ของวิธีการเปรียบเทียบพหุคูณตามข้อตกลงความแปรปรวนวิวิธพันธ์ของข้อมูลที่แจกแจงต่างกัน และขนาดกลุ่มตัวอย่างต่างกัน**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒ.

นิภาพร ขำสอาด. 2552. **อำนาจการทดสอบของการใช้สถิติการทดสอบความแตกต่างของค่าเฉลี่ย.** วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒ.

ปุณยนุช พินชู. 2548. **การเปรียบเทียบค่าความคลาดเคลื่อนประเภทที่ 1 และอำนาจการทดสอบของวิธีการเปรียบเทียบค่าเฉลี่ยรายคู่ สำหรับแผนการทดลองแบบสุ่มสมบูรณ์**. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

ประสิทธิ์ พยัคฆพงษ์. 2545. **สถิติเชิงคณิศาสตร์: ทฤษฎีและการประยุกต์.** สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

วิชชุดา ศรีโสภา. 2539. **การศึกษาเปรียบเทียบวิธีการเปรียบเทียบพหุ**. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

สิรินุช เขียวเอี่ยม. 2544. **ผลการเปรียบเทียบพหุคูณรายคู่บางวิธีภายใต้ความแปรปรวนที่แตกต่างกันจากกลุ่มตัวอย่างขนาดต่างกัน**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒ.

อำไพ ทองธีรภาพ. 2554. **วิธีการทางสถิติ**. สำนักพิมพ์ มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Anderson, T.W. and D.A. Darling. 1954. A Test of Goodness of fit. **Journal of the American Statistical Association** 49: 765-769.

Hochberg, Y. and A. Tamhane. 1987. **Multiple Comparison Procedures.** John Wiley & Sons, Canada.

Hollander, M. and D.A. Wolfe. 1999. **Nonparametric statistical methods**. 2nd ed. John Wiley & Sons, Canada.

Kirk, R.E. 1995. **Experimental Design: Procedures for the Behavioral Science**. 3rd ed. Californian Print, United States of American.

Neuhäuser, M. 2003. A Note on the Exact Test Based on the Baumgartner-Weiß-Schindler statistic in the presence of ties. **Computational Statistics & Data Analysis**. 42: 561–568.

_____ and F. Bretz. 2001. Nonparametric All-Pairs Multiple Comparisons. **Biometrical Journal**. 43: 571–580.

_____ and F.C. Lam. 2004. Nonparametric Approaches to Detecting Differentially Expressed Genes in Replicated Microarray Experiments. **Australian Computer Society**. 29: 139–143.

_____ A. Schulz and D. Czech. 2009. A SAS/IML Algorithm for an Exact Permutation Test. **GSM Medizinische Informatik, Biometrie und Epidemiologie** 5 (2).

_____ and R. Senske. 2004. The Baumgartner-Weiß-Schindler Test for the Detection of Differentially Expressed Genes in Replicated Microarray Experiments. **Bioinformatics**. 20: 3553–3564.

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

R Development Core Team. 2011. **R: A Language and Environment for Statistical computing**. R Foundation for Statistical Computing. Source: http://www.R-project.org, July 17, 2011.

Rizzo M.L. 2008. **Statistical Computing with R.** Chapman and Hall, United States of American.

Scarpa, R. and A. Alberini. 2005. **Applications of Simulation Methods in Environmental and Resource Economics.** Spinger, Netherlands.

Smyth, G. K. and B. Phipson. 2010. Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn. **Statistical Applications in Genetics and Molecular Biology** 9 (1).

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ภาคผนวก

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

โปรแกรมที่ใช้ในการทดสอบสมมติฐานด้วยสถิติทดสอบ TK

```
data testtk;
infile 'D:\multipledata\data\datae3gp12.txt';
input rep 1-3 group 4-5 value;
title1 'Tukey test ' ;
        proc glm;
        class group ;
        model value=group;
        lsmeans group/ pdiff adjust= tukey;
proc print;
run;
```

ผลลัพธ์จากโปรแกรม SAS

Tukey test

The GLM Procedure

Class Level Information

Class	Levels	Values
Group	3	1 2 3
Number of Observations Read		30
Number of Observations Used		30

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Dependent Variable: value

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	18146.06512	9073.03256	10.47	0.0004
Error	27	23393.32583	866.41948		
Corrected Total	29	41539.39095			

R-Square	Coeff Var	Root MSE	value Mean
0.436840	51.46119	29.43500	57.19845

Source	DF	Type I SS	Mean Square	F Value	Pr > F
group	2	18146.06512	9073.03256	10.47	0.0004

Source	DF	Type III SS	Mean Square	F Value	Pr > F
group	2	18146.06512	9073.03256	10.47	0.0004

The GLM Procedure

Least Squares Means

Adjustment for Multiple Comparisons: Tukey

group	value LSMEAN	LSMEAN Number
1	23.5348236	1
2	81.6047883	2
3	66.4557401	3

Least Squares Means for effect group

Pr > |t| for H0: LSMean(i)=LSMean(j)

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

Dependent Variable: value

i/j	1	2	3
1		0.0004	0.0082
2	0.0004		0.4921
3	0.0082	0.4921	

โปรแกรมที่ใช้ในการทดสอบสมมติฐานด้วยสถิติทดสอบ BWS

ดัดแปลงมาจากโปรแกรม Permtest (Neuhäuser *et al*.,2009)

```
/* case Baumgartner */
%MACRO Permtest(indata);
proc iml;
/* Reading the data */
USE &indata;
READ ALL INTO currdata;

/* Computation of ranks */
ranks=RANKTIE(currdata[ ,3]);

/* Calculation of the sample sizes per group */
N_total=Nrow(currdata[ ,3]);
n2=currdata[+,2];
n1=N_total-n2;
print N_total n1 n2;

/* Creation of all possible permutations */
start perm(n,n_1);
matrix = shape(0,(gamma(n+1)/(gamma(n_1+1)*gamma(n-n_1+1))),n);
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
index = 1;
vektor=shape(-1,1,n);
pos = 1;
ok = 1;
do while(ok=1);
        if pos > n then do;
                if vektor[,+] = n_1 then do;
                        matrix[index,]= vektor;
                        index = index + 1;
                end;
                pos = pos-1;
        end;
        else do;
                if vektor[,pos] < 1 then do;
                        vektor[,pos] = vektor[,pos]+1;
                        pos = pos+1;
                end;
        else do;
                vektor[,pos]=-1;
                pos = pos-1;
        end;
end;
if pos < 1 then ok = 0;
end;
return (matrix);
finish;
permutations = perm(N_total,n1);
P=Nrow(permutations);

/* Calculation of test statistic */
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
start test_sta(R1, R2, N_total, n1, n2);
b=R1;
R1[,rank(R1)]=b;
b=R2;
R2[,rank(R2)]=b;
i=1:n1;
j=1:n2;
Bx=(sum( (R1-(N_total/n1)#i)##2/( (i/(n1+1))#(1-(i/(n1+1)))#((n2#N_total)/n1 ) ))#(1/n1);
By=(sum( (R2-(N_total/n2)#j)##2/( (j/(n2+1))#(1-(j/(n2+1)))#((n1#N_total)/n2) ) )#(1/n2);
B=(Bx+By)/2;
return (B);
finish;

/* Carrying out the test */
Tab=REPEAT(T(ranks),P,1);
R1=choose(permutations=0,.,Tab);
R2=choose(permutations=1,.,Tab);
R1g=R1[loc(R1^=.)];
R2g=R2[loc(R2^=.)];
R1z=shape(R1g,P, n1);
R2z=shape(R2g,P, n2);
test_st0=
test_sta(T(ranks[1:n1]),T(ranks[(n1+1):N_total]), N_total, n1, n2);
Pval=0;

do i=1 to P by 1;
        B = test_sta(R1z[ i , ], R2z[ i , ], N_total, n1, n2);
if B >= test_st0 then Pval=Pval+1;
end;
Pval=Pval/P;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
/* Definition of output */
x=( N_Total || N1 || N2 || Pval || test_st0 || P);
cols={N_Total n1 n2  P_value test_statistic total_Perms};
print x[colname=cols];

/* optional: Creation of an output dataset called results */
CREATE results FROM x[colname=cols];
APPEND FROM x;
CLOSE results;
/*********************************************************/
quit;
%MEND Permtest;
data testBWS;
infile 'D:\multipledata\data\datae3gp12.txt';
input rep 1-3 group 4-5  value;
title1 'BWS group1 VS group 2'
if rep=100 then
        output;
        %Permtest(testBWS);
proc print;
run;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ผลลัพธ์จากโปรแกรม SAS

BWS group1 VS group 2

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.0004763	7.2685228	184756

BWS group1 VS group 3

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.0007145	6.7979673	184756

BWS group2 VS group 3

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.5344238	0.6892679	184756

โปรแกรมที่ใช้ในการทดสอบสมมติฐานด้วยสถิติทดสอบ DSCF

ดัดแปลงมาจากโปรแกรม Permtest (Neuhäuser *et al*.,2009)

```
/*DSCF */
/* case : group 1 vs group2*/
%MACRO Permtest(indata);
proc iml;
/* Reading the data */
USE &indata;
READ ALL INTO currdata;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
/* Computation of ranks */
ranks=RANKTIE(currdata[ ,3]);

/* Calculation of the sample sizes per group */
N_total=Nrow(currdata[ ,3]);
n2=currdata[+,2];
n1=N_total-n2;
/*print N_total n1 n2;*/

/* Creation of all possible permutations */
start perm(n,n_1);
matrix = shape(0,(gamma(n+1)/(gamma(n_1+1)*gamma(n-n_1+1))),n);
index = 1;
vektor=shape(-1,1,n);
pos = 1;
ok = 1;
do while(ok=1);
        if pos > n then do;
                if vektor[,+] = n_1 then do;
                        matrix[index,]= vektor;
                        index = index + 1;
                end;
                pos = pos-1;
        end;
        else do;
                if vektor[,pos] < 1 then do;
                        vektor[,pos] = vektor[,pos]+1;
                        pos = pos+1;
                end;
        else do;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
                vektor[,pos]=-1;
                pos = pos-1;
        end;
end;
if pos < 1 then ok = 0;
end;
return (matrix);
finish;
permutations = perm(N_total,n1);
P=Nrow(permutations);

/* Calculation of test statistic */
start test_sta(R1, R2, N_total, n1, n2);
b=R1;
R1[,rank(R1)]=b;
b=R2;
R2[,rank(R2)]=b;
i=1:n1;
j=1:n2;
W=((sum(R1)-(n1#((n1+n2+1)/2))))/(((n1#n2)#((n1+n2+1)/24)))##(1/2));
B=abs(W);
return (B);
finish;

/* Carrying out the test */
Tab=REPEAT(T(ranks),P,1);
R1=choose(permutations=0,.,Tab);
R2=choose(permutations=1,.,Tab);
R1g=R1[loc(R1^=.)];
R2g=R2[loc(R2^=.)];
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
R1z=shape(R1g,P, n1);
R2z=shape(R2g,P, n2);
test_st0=
test_sta(T(ranks[1:n1]),T(ranks[(n1+1):N_total]), N_total, n1, n2);
Pval=0;

do i=1 to P by 1;
        B = test_sta(R1z[ i , ], R2z[ i , ], N_total, n1, n2);
if B >= test_st0 then Pval=Pval+1;
end;
Pval=Pval/P;

/* Definition of output */
x=(N_Total || N1 || N2 || Pval || test_st0 || P);
cols={N_Totla n1 n2 P_value test_statistic total_Perms};
print x[colname=cols];

/* optional: Creation of an output dataset called results */
CREATE results FROM x[colname=cols];
APPEND FROM x;
CLOSE results;
/*************************************************************/
quit;
%MEND Permtest;
data testDSCF;
infile 'D:\multipledata\data\datae3gp12.txt';
input rep 1-3 group 4-5  value;
title1 'DSCF group1 VS group 2'
if rep=100 then
        output;
```

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

```
%Permtest(testDSCF);
proc print;
run;
```

ผลลัพธ์ที่ได้จากโปรแกรม SAS

DSCF group 1 VS group 2

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.0007253	4.4899889	184756

DSCF group1 VS group 3

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.00105	4.3830844	184756

DSCF group2 VS group 3

N_TOTAL	N1	N2	P_VALUE	TEST_STATISTIC	TOTAL_PERMS
20	10	10	0.4358722	1.1759495	184756

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นางสาวสาลินี เขื่อนคำ
วัน เดือน ปี ที่เกิด	8 กันยายน พ.ศ. 2529
สถานที่เกิด	อำเภอแม่เมาะ จังหวัดลำปาง
ประวัติการศึกษา	วท.บ. (สถิติ) มหาวิทยาลัยเกษตรศาสตร์
ตำแหน่งหน้าที่การงานปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-

ลิขสิทธิ์ มหาวิทยาลัยเกษตรศาสตร์