การเปรียบเทียบประสิทธิภาพระหว่างตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis

โดย
นางสาวพวงทิพย์ พูลสวัสดิ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติประยุกต์

ภาควิชาสถิติ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2547

ISBN 974-464-654-3

THE EFFICIENCY COMPARISONS AMONG BI-ASPECT, F AND KRUSKAL – WALLIS TEST STATISTICS

By

Paungtip Poonsawat

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

Department of Statistics

Graduate School

SILPAKORN UNIVERSITY

2004

ISBN 974-464-654-3

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง "การเปรียบเทียบประสิทธิภาพระหว่างตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis" เสนอโดย นางสาวพวงทิพย์ พูลสวัสดิ์ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาสถิติประยุกต์

……….……………………………………..
(รองศาสตราจารย์ ดร. จิราวรรณ คงคล้าย)
คณบดีบัณฑิตวิทยาลัย
วันที่ .... เดือน .... พ.ศ. ....

ผู้ควบคุมวิทยานิพนธ์

รองศาสตราจารย์ ไพบูลย์ รัตนประเสริฐ

คณะกรรมการตรวจสอบวิทยานิพนธ์

.... ประธานกรรมการ
(อาจารย์ ดร. กมลชนก พานิชการ)
............/.........................../............

.... กรรมการ
(รองศาสตราจารย์ ไพบูลย์ รัตนประเสริฐ)
............/.........................../............

.... กรรมการ
(อาจารย์ ดร. กุศยา ปลั่งพงษ์พันธ์)
............/.........................../............

.... กรรมการ
(อาจารย์ ดร. บุญอ้อม โฉมที)
............/.........................../............

K45304202 : สาขาสถิติประยุกต์
คำสำคัญ : การเรียงสับเปลี่ยน / ค่ากลาง / กำลังการทดสอบ

พวงทิพย์ พูลสวัสดิ์ : การเปรียบเทียบประสิทธิภาพระหว่างตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis (THE EFFICIENCY COMPARISONS AMONG BI-ASPECT, F AND KRUSKAL –WALLIS TEST STATISTICS) อาจารย์ผู้ควบคุมวิทยานิพนธ์ : รศ.ไพบูลย์ รัตนประเสริฐ.103 หน้า. ISBN 974-464-654-3

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อ เปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม ได้แก่ ตัวสถิติทดสอบ Bi- aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal-Wallis โดยประสิทธิภาพวัดจากความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อสมมติฐาน H_0 จริง และกำลังการทดสอบเมื่อสมมติฐาน H_1 จริง การศึกษาได้ทำการจำลองแบบข้อมูลโดยวิธีมอนติคาร์โล จำนวน 4,000 ซ้ำ จากประชากรที่มีการแจกแจงใน 3 ลักษณะคือการแจกแจงที่มีลักษณะสมมาตร ได้แก่การแจกแจงแบบปกติ และยูนิฟอร์ม การแจกแจงที่มีลักษณะเบ้ ได้แก่การแจกแจงแบบไคสแควร์ และเอกซ์โปเนนเชียล และการแจกแจงแบบปลอมปน ซึ่งกำหนดให้เปอร์เซ็นต์ของการปลอมปนเท่ากับ 10% และ 30% โดยศึกษาใน 2 กรณีคือ กรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน และกรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกันที่ระดับ 5% และ 20% กำหนดจำนวนกลุ่มของประชากรที่ศึกษาเท่ากับ 3 และ 5 กลุ่ม โดยทำการทดสอบที่ระดับนัยสำคัญ 0.05 และ 0.10

ผลการศึกษาพบว่า สำหรับประชากร 3 กลุ่ม ในกรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน การแจกแจงแบบปกติ และยูนิฟอร์ม ตัวสถิติทดสอบทั้ง 3 ชนิดมีค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เข้าใกล้ระดับนัยสำคัญที่ตั้งไว้ การแจกแจงแบบไคสแควร์ และเอกซ์โปเนนเชียล ตัวสถิติทดสอบ Kruskal-Wallis จะควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบตัวอื่นเล็กน้อย ส่วนการแจกแจงแบบปลอมปน 10% และ 30% ตัวสถิติทดสอบ Kruskal-Wallis จะควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ Bi-aspect ในขณะที่ตัวสถิติทดสอบ F ไม่สามารถควบคุมได้ กรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกัน ศึกษาใน 2 กรณีย่อย คือ เมื่อมีตัวอย่างเพียง 1 กลุ่มมีค่าเฉลี่ยแตกต่างจากกลุ่มอื่น 5% และ 20% พบว่า ในการแจกแจงแบบปกติและยูนิฟอร์ม ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบต่ำสุด สำหรับการแจกแจงอื่น ๆ ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และตัวสถิติทดสอบ F มีกำลังการทดสอบต่ำสุด ส่วนกรณีตัวอย่างทุกกลุ่มมีค่าเฉลี่ยแตกต่างกัน 5% และ 20% ได้ผลเช่นเดียวกันแต่มีกำลังการทดสอบมากกว่ากรณีที่มีค่าเฉลี่ยของตัวอย่าง 1 กลุ่มแตกต่างจากกลุ่มอื่น สำหรับกรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มจำนวน 5 กลุ่มได้ผลการศึกษาทำนองเดียวกับกรณีที่มีประชากร 3 กลุ่ม และพบว่ากำลังการทดสอบของตัวสถิติทดสอบทั้ง 3 ชนิดมีค่าเพิ่มเมื่อขนาดตัวอย่างเพิ่มและค่าเฉลี่ยของตัวอย่างมีความแตกต่างมากขึ้น

ภาควิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร ปีการศึกษา 2547
ลายมือชื่อนักศึกษา ………………….…………
ลายมือชื่ออาจารย์ผู้ควบคุมวิทยานิพนธ์ ………………………………..

K45304202 : MAJOR : APPLIED STATISTICS

KEYWORD : PERMUTATION / LOCATION / POWER OF THE TEST

PAUNGTIP POONSAWAT : THE EFFICIENCY COMPARISONS AMONG BI-ASPECT, F AND KRUSKAL – WALLIS TEST STATISTICS. THESIS ADVISOR : ASSO.PROF.PAIBOOL RATANAPRASERT. 103 pp. ISBN 974-464-654-3.

The objective of this research is to compare the efficiencies among Bi-aspect, F and Kruskal-Wallis test statistics, which are used for testing hypothesis about the difference of the population location parameters. The efficiency is measured by the probability of type I error when H_0 is true and power of the test when H_1 is true. The study is conducted by simulating data using Monte Carlo Technique with 4,000 replications under each conditions. Three types of distribution are studied; symmetry distribution are represented by normal and uniform distribution, skew distribution are represented by chi-squared and exponential distribution and contaminated distribution. The percent of the contamination is set at 10 and 30 . The studies are divided into two cases : all of the population means are equal and the case when there is different means at 5% and 20%. The groups of population studied are three and five and there will be tested at level of significance 0.05 and 0.10.

The results of the study of 3 groups of population are in the followings. The case when all of the population means are equal, normal and uniform distribution, all the three test statistics have the probability of type I error which approach the designed level of significance. Chi-squared and exponential distribution, Kruskal-Wallis test performed better control of the probability of type I error than others. Under 10% and 30% contaminated distribution, Kruskal-Wallis test provided better results than Bi-aspect whereas F test is too conservative. The case when there is the different mean, the two following sub-cases will be studied; the first case when the mean of one group differs 5% and 20% from other groups, the finding are as follow : in normal and uniform distribution, the F test statistics have the hightest power of the test, the Kruskal-Wallis and Bi-aspect decrease respectively. As for other distributions, Kruskal-Wallis test statistics has the hightest power of test , the Bi-aspect and F decrease respectively. The second case when the mean of every group differs 5% and 20%, the results are the same as the first case but the power of the test is higher. As for 5 groups, the results are the same as the study of 3 groups. However, the power of the test, are improved with increasing sample sizes and difference in mean.

Department of Statistics Graduate School ,Silpakorn University Academic Year 2004

Student’s signature ………………………………

Thesis Advisor’s signature ………………………………….

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้เป็นงานวิจัยชิ้นแรกในชีวิตของผู้วิจัย และเป็นส่วนสำคัญในการศึกษาระดับบัณฑิตศึกษาสาขาสถิติประยุกต์ ซึ่งผู้วิจัยต้องอาศัยความรู้ ประสบการณ์ ตลอดความอดทนเพื่อให้สำเร็จเป็นรูปเล่ม อย่างไรก็ตาม ความสำเร็จครั้งนี้จะเกิดขึ้นไม่ได้ หากปราศจากคณะบุคคลที่คอยให้ความช่วยเหลือ ในโอกาสนี้ผู้วิจัยขอกราบขอบพระคุณ รองศาสตราจารย์ไพบูลย์ รัตนประเสริฐ อาจารย์ผู้ที่เสียสละเวลาให้คำแนะนำ ชี้แนะ ตรวจสอบแก้ไข ด้วยดีมาโดยตลอด

กราบขอบพระคุณ อาจารย์ ดร. กมลชนก พานิชการ ประธานกรรมการสอบวิทยานิพนธ์ อาจารย์ ดร. กุศยา ปลั่งพงษ์พันธ์ และ อาจารย์ ดร. บุญอ้อม โฉมที กรรมการผู้ตรวจสอบวิทยานิพนธ์ที่คอยให้คำชี้แนะที่เป็นประโยชน์ยิ่งต่อผู้วิจัย

กราบขอบพระคุณอาจารย์ทุกท่านในภาควิชาสถิติที่คอยประสิทธิ์ประสาทวิชาความรู้ เพื่อให้สามารถนำไปใช้ในการดำเนินชีวิตต่อไปในอนาคต

ขอขอบคุณ คุณประสพชัย พสุนนท์ และคุณอาฟีฟี ลาเต๊ะ ที่คอยให้คำแนะนำและช่วยเหลือด้านการเขียนโปรแกรม และขอขอบคุณ คุณสุมณฑา เขียวเสน คุณภาณุพงศ์ พนมวันคุณพนิดา คงแจ่ม ที่คอยให้กำลังใจและช่วยเหลือตลอดมา รวมทั้งขอบคุณเพื่อน พี่ และน้องในสาขาสถิติประยุกต์ทุกท่าน

สุดท้ายขอกราบขอบพระคุณ คุณพ่อ และ คุณแม่ ในความรัก ความอบอุ่นที่มีให้ตลอดเวลา กำลังใจในยามท้อถอยและการสนับสนุนอย่างดีมาโดยตลอด

สารบัญ

สารบัญตาราง

ตารางที่ หน้า

ตารางที่	หน้า

สารบัญรูป

รูปที่ หน้า

รูปที่		หน้า

รูปที่		หน้า

รูปที่ **หน้า**

รูปที่ **หน้า**

บทที่1

บทนำ

ความเป็นมาและความสำคัญของปัญหา

ในการทดสอบสมมติฐานเชิงสถิติเกี่ยวกับค่าเฉลี่ยของประชากร 1 กลุ่ม หรือ 2 กลุ่ม มักดำเนินการโดยใช้ตัวสถิติทดสอบ t (t-test) โดยมีข้อตกลงเบื้องต้นว่าประชากรที่ตัวอย่างถูกสุ่มมาต้องมีการแจกแจงแบบปกติ แต่ถ้าเราสนใจที่จะเปรียบเทียบค่าเฉลี่ยของประชากรมากกว่า 2 กลุ่มพร้อมกัน การใช้ตัวสถิติทดสอบ t เพื่อทดสอบความแตกต่างของค่าเฉลี่ยของประชากรทีละคู่ อาจไม่เหมาะสมเนื่องจากจะต้องทำการทดสอบหลายครั้ง ทำให้เกิดความยุ่งยาก นอกจากนี้ยังทำให้ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เพิ่มมากขึ้น ดังนั้นในกรณีที่มีตัวอย่างสุ่มจากประชากรมากกว่า 2 กลุ่มจึงนิยมใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way Analysis of Variance : Anova) และทดสอบสมมติฐานด้วยตัวสถิติทดสอบ F (F-test) ในการใช้ตัวสถิติทดสอบ F มีข้อตกลงเบื้องต้นคือ ตัวอย่างถูกสุ่มอย่างเป็นอิสระจากประชากร C กลุ่ม ประชากรแต่ละกลุ่มมีการแจกแจงแบบปกติ และมีความแปรปรวนเท่ากัน ดังนั้นจึงเป็นปัญหาว่าเมื่อตัวอย่างที่สุ่มมาจากประชากรไม่ได้เป็นไปตามข้อตกลงเบื้องต้น การใช้ตัวสถิติทดสอบ t และตัวสถิติทดสอบ F ในการทดสอบสมมติฐาน อาจนำไปสู่ผลสรุปที่ผิดพลาดได้ ถึงแม้บางครั้งมีการแก้ไขปรับข้อมูล โดยวิธีการแปลงข้อมูลเพื่อให้มีคุณสมบัติเป็นไปตามข้อตกลงเบื้องต้น แต่ในการแปลงข้อมูลมีขั้นตอนการทดสอบที่ยุ่งยาก หรือในบางกรณีไม่สามารถหาวิธีการแปลงข้อมูลที่เหมาะสมได้ จึงมีผู้คิดหาวิธีทดสอบที่ไม่ต้องมีข้อตกลงเบื้องต้นมากนักและนำไปใช้ได้สะดวก เนื่องจากใช้กระบวนการทางคณิตศาสตร์ง่าย ๆ เช่นการจัดลำดับ หรือ การนับความถี่ ตัวอย่าง เช่น Mood (1950) และ Westenberg (1948) เสนอตัวสถิติทดสอบมัธยฐาน (The median test) เพื่อใช้ทดสอบกลุ่มตัวอย่าง 2 กลุ่มที่อิสระกันว่าสุ่มมาจากประชากรที่มีค่ามัธยฐานเท่ากันหรือไม่ โดยนำข้อมูลจากตัวอย่างทั้งสองกลุ่มมาหาค่ามัธยฐานร่วม แล้วแยกข้อมูลในแต่ละตัวอย่างออกเป็น 2 พวกคือพวกหนึ่งมีค่าสูงกว่าค่ามัธยฐานร่วมหรืออยู่เหนือมัธยฐานร่วมและอีกพวกหนึ่งมีค่าต่ำกว่าค่ามัธยฐานร่วมหรืออยู่ใต้มัธยฐานร่วม ถ้าค่าสังเกตมีค่าเท่ากับค่ามัธยฐานร่วมพอดีสามารถแก้ปัญหานี้ได้ 2 วิธีคือ (1) ถ้าขนาดตัวอย่างมีมากพอและมีค่าสังเกตน้อยค่าที่เท่ากับมัธยฐานร่วม

ให้ตัดค่าเหล่านั้นทิ้งไป (2) ให้แยกประเภทของข้อมูลเสียใหม่เป็นอยู่เหนือมัธยฐานร่วม (สูงกว่ามัธยฐานร่วม) และไม่อยู่เหนือมัธยฐานร่วม (ต่ำกว่าหรือเท่ากับมัธยฐานร่วม) ซึ่งถ้าประชากรทั้งสองกลุ่มมีค่ามัธยฐานเท่ากันจริง ครึ่งหนึ่งของค่าสังเกตในแต่ละตัวอย่างจะอยู่เหนือมัธยฐานร่วมและอีกครึ่งหนึ่งอยู่ใต้มัธยฐานร่วม ดังนั้นการทดสอบสมมติฐาน H_0 จึงพิจารณาจากขนาดของความแตกต่างระหว่างสัดส่วนที่ค่าสังเกตจะอยู่เหนือและใต้มัธยฐานร่วม ถ้ามีขนาดความแตกต่างของสัดส่วนมากก็จะปฏิเสธสมมติฐาน H_0 และขยายการทดสอบไปสู่การทดสอบความแตกต่างของค่ามัธยฐานสำหรับประชากรหลายกลุ่มด้วย Terpstra (1952) เสนอตัวสถิติ Jonckheere-Terpstra test เป็นตัวสถิติทดสอบความแตกต่างของค่ามัธยฐานสำหรับประชากรหลายกลุ่มว่ามีความแตกต่างกันหรือไม่ โดยสมมติฐาน H_0 คือประชากรทุกกลุ่มมีค่ามัธยฐานเท่ากัน และสมมติฐาน H_1 เป็นแบบอันดับคือกลุ่มที่ 1 มีค่ามัธยฐานน้อยกว่ากลุ่มที่ 2 และกลุ่มที่ 2 มีค่ามัธยฐานน้อยกว่ากลุ่มที่ 3 ไปจนกระทั่งกลุ่มที่ C ตัวสถิติทดสอบจะเป็นผลรวมของจำนวนที่ค่าจากตัวอย่างกลุ่มที่ i มีค่ามากกว่ากลุ่มที่ j โดยที่ $i < j$ สำหรับทุกคู่ที่เป็นไปได้จำนวน $\binom{C}{2} = \frac{C(C-1)}{2}$ คู่ ปฏิเสธสมมติฐาน H_0 เมื่อผลรวมมีค่ามากพอ Van der Waerden (1953) เสนอตัวสถิติทดสอบ Normal Scores โดยใช้ค่าของ inverse-normal score $\left(Z^{(i)}\right)$ แทนค่าตัวอย่างโดยทำการแปลงค่าของตัวอย่างให้อยู่ในรูปเปอร์เซ็นต์ไทล์ของอันดับ ตัวสถิติทดสอบจะเป็นผลรวมของคะแนนปกติผกผันของทุกกลุ่ม เมื่อขนาดตัวอย่างมีขนาดมากพอจะมีลิมิตการแจกแจงเป็นการแจกแจงแบบไคสแควร์ ด้วยองศาความเป็นอิสระ $C-1$ เมื่อสมมติฐาน H_0 จริง และตัวสถิติทดสอบที่นิยมใช้ทดสอบความแตกต่างของค่ากลางของประชากร C กลุ่มซึ่งใช้ได้ดีกับตัวอย่างขนาดเล็ก คือตัวสถิติทดสอบ Kruskal Wallis ซึ่งเสนอโดย Kruskal and Wallis (1952) ตัวสถิติดังกล่าวนี้สามารถวิเคราะห์ข้อมูลที่ได้จากการวางแผนทดลองชนิดสุ่มสมบูรณ์ จึงเรียกว่าการวิเคราะห์ความแปรปรวนทางเดียวชนิดอันดับ แต่เนื่องจากตัวสถิติทดสอบ Kruskal - Wallis มีข้อเสียอยู่บ้างตรงที่หากข้อมูลที่นำมาทดสอบตรงตามข้อตกลงของตัวสถิติทดสอบ F กล่าวคือเมื่อตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ และมีความแปรปรวนเท่ากันประสิทธิภาพของตัวสถิติทดสอบ Kruskal - Wallis จะด้อยกว่าตัวสถิติทดสอบ F

Macro Marozzi (2003) ได้เสนอตัวสถิติทดสอบที่สามารถทดสอบความแตกต่างของค่ากลางในประชากร 2 กลุ่มได้ดีในทุก ๆ กรณี เพียงแต่ตัวอย่างที่สุ่มมาต้องสุ่มจากประชากรที่มีลักษณะต่อเนื่อง เรียกตัวสถิติทดสอบนี้ว่าตัวสถิติทดสอบ Bi-aspect (Bi-aspect test) ประกอบด้วยตัวสถิติทดสอบ 2 ตัวเพื่อทดสอบข้อมูล 2 ด้านไปพร้อม ๆ กัน คือทดสอบเชิงตัวเลข (numerical aspect) และทดสอบเชิงกลุ่ม (categorical aspect) โดยใช้หลักการของการ

ทดสอบแบบ permutation test และในปี 2004 เขาได้พัฒนาตัวสถิติทดสอบ Bi-aspect ให้สามารถทดสอบได้กับประชากรหลายกลุ่มได้ด้วย

เพื่อเป็นการแสดงให้เห็นว่า ตัวสถิติทดสอบความแตกต่างค่ากลางของประชากรหลายกลุ่มที่ Macro Marozzi เสนอ มีประสิทธิภาพเป็นอย่างไรเมื่อเทียบกับตัวสถิติทดสอบที่นิยมใช้กันในปัจจุบันคือตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis ในงานวิจัยนี้ ผู้วิจัยจึงสนใจทำการเปรียบเทียบตัวสถิติทดสอบ Bi-aspect กับ ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis โดยจะทำการเปรียบเทียบในกรณีที่ตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงในลักษณะต่างๆ การพิจารณาประสิทธิภาพของตัวสถิติจะเปรียบเทียบจากความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อสมมติฐาน H_0 เป็นจริง และกำลังการทดสอบของตัวสถิติ (Power of Test) ของตัวสถิติทดสอบแต่ละตัวเมื่อสมมติฐาน H_1 เป็นจริง

วัตถุประสงค์ของการวิจัย

เพื่อเปรียบเทียบประสิทธิภาพ ของตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม คือ ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และ ตัวสถิติทดสอบ Kruskal- Wallis

ขอบเขตของการศึกษา

ในการเปรียบเทียบตัวสถิติทดสอบ 3 ตัว คือ ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และ ตัวสถิติทดสอบ Kruskal- Wallis กำหนดการแจกแจงของประชากรที่ศึกษา 3 ลักษณะคือ

1. การแจกแจงที่มีลักษณะสมมาตร ประกอบด้วย

การแจกแจงแบบปกติและการแจกแจงแบบยูนิฟอร์ม ซึ่งในที่นี้กำหนดให้การแจกแจงแบบปกติมี $\mu = 10$ และ $\sigma^2 = 1$ ส่วนการแจกแจงแบบยูนิฟอร์ม กำหนดให้มีการแจกแจงอยู่ในช่วง $(0, 2\sqrt{3})$ เขียนสัญลักษณ์แทนด้วย $N(10,1)$ และ $U(0, 2\sqrt{3})$ ตามลำดับ

2. การแจกแจงที่มีลักษณะเบ้ ประกอบด้วย

การแจกแจงแบบไคสแควร์ และการแจกแจงแบบเอกซ์โปเนนเชียล ซึ่งในที่นี้กำหนดให้การแจกแจงแบบเอกซ์โปเนนเชียล มีค่าพารามิเตอร์ $\beta = 1$ ส่วนการแจกแจงแบบไคสแควร์ กำหนดให้มีองศาอิสระเท่ากับ 4

3. การแจกแจงแบบปลอมปน เป็นการแจกแจงผสมระหว่างตัวอย่างที่มีการแจกแจงแบบปกติที่มี $\mu_1 = 10, \sigma_1^2 = 1$ กับตัวอย่างที่มีการแจกแจงแบบปกติที่มี $\mu_2 = 1, \sigma_2^2 = 100$ ด้วยสัดส่วน p ซึ่งในที่นี้กำหนดให้ $p = 0.1$ และ $p = 0.3$ เขียนแทนด้วย $CN(\mu_1, \sigma_1^2, p, \mu_2, \sigma_2^2)$

สำหรับขนาดตัวอย่างกำหนดให้ตัวอย่างในแต่ละกลุ่มเท่ากัน แบ่งออกเป็น 2 กรณี

1. กรณีที่มีตัวอย่าง 3 กลุ่มถูกสุ่มอย่างเป็นอิสระจากประชากรแต่ละกลุ่ม กำหนดให้ขนาดตัวอย่างในแต่ละกลุ่มคือ 10, 20 และ 40 ตามลำดับ สามารถเขียนแทนด้วย $(10,10,10)$, $(20,20,20)$ และ $(40,40,40)$

2. กรณีที่มีขนาดตัวอย่าง 5 กลุ่ม ถูกสุ่มอย่างเป็นอิสระจากประชากรแต่ละกลุ่ม กำหนดให้ขนาดตัวอย่างที่ใช้คือ 10, 20 และ 40 ตามลำดับ สามารถเขียนแทนด้วย $(10,10,10,10,10)$, $(20,20,20,20,20)$ และ $(40,40,40,40,40)$

โดยในแต่ละกรณีพิจารณาใน 2 กรณี คือ

กรณีที่ 1 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มที่มีค่าเฉลี่ยเท่ากัน ซึ่งกรณีดังกล่าวนี้ประสิทธิภาพของตัวสถิติทดสอบจะพิจารณาจากความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ว่าใกล้เคียงระดับนัยสำคัญที่กำหนดหรือไม่

กรณีที่ 2 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มที่มีค่าเฉลี่ยแตกต่างกัน แบ่งออกเป็น 2 กรณีย่อยคือ

กรณี 2.1 มีค่าเฉลี่ยของประชากรเพียง 1 กลุ่มที่แตกต่างจากกลุ่มอื่น

กรณี 2.2 มีค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน

ประสิทธิภาพของตัวสถิติทดสอบจะพิจารณาจากกำลังการทดสอบของตัวสถิติ

การศึกษาใช้การจำลองแบบโดยวิธีการของมอนติคาร์โล (Monte Carlo Method) ด้วยโปรแกรม Fortran power station 4.0 ทำซ้ำ 4,000 รอบ และหาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบแต่ละตัวเปรียบเทียบกันในข้อมูลที่มีการแจกแจงแบบต่าง ๆ ตามที่กำหนด

ประโยชน์ที่คาดว่าจะได้รับ

เพื่อเป็นแนวทางสำหรับการเลือกตัวสถิติทดสอบความแตกต่างค่ากลางของประชากรหลายกลุ่มให้เหมาะสมกับข้อมูลที่มีลักษณะต่าง ๆ

นิยามศัพท์

ความคลาดเคลื่อนชนิดที่ 1 (Type I Error) หมายถึง ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานว่าง H_0 เมื่อสมมติฐานว่าง H_0 จริง และมักจะเรียกความน่าจะเป็นของความคลาดเคลื่อนชนิดนี้ว่า ระดับนัยสำคัญ (level of significance) แทนด้วย สัญลักษณ์ α

ความคลาดเคลื่อนชนิดที่ 2 (Type II Error) หมายถึง ความคลาดเคลื่อนที่เกิดจากการยอมรับสมมติฐานว่าง H_0 เมื่อสมมติฐานแย้ง H_1 จริง และมักแทนความน่าจะเป็นของความคลาดเคลื่อนชนิดนี้ด้วยสัญลักษณ์ β

กำลังการทดสอบ (Power of the test) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง H_0 เมื่อสมมติฐานแย้ง H_1 จริง แทนด้วย สัญลักษณ์ $1-\beta$

ประสิทธิภาพการทดสอบ (Efficiency) พิจารณาโดยใช้เกณฑ์ดังนี้

(1) กรณีที่ H_0 จริง พิจารณาจากความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 โดยความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ควรมีค่าเข้าใกล้และไม่ควรเกินระดับนัยสำคัญที่ตั้งไว้

(2) กรณี H_1 จริง พิจารณาจากกำลังการทดสอบ $(1-\beta)$ ถ้าการทดสอบใดมีค่า $1-\beta$ สูงกว่าถือว่ามีประสิทธิภาพดีกว่า

ค่ากลาง (Location) หมายถึง ค่าที่เป็นตัวแทนของข้อมูลทุก ๆ หน่วยในกลุ่มของประชากร ซึ่งค่ากลางที่ใช้ในงานวิจัยนี้มี 2 ชนิดประกอบด้วย

1. ค่าเฉลี่ย (Mean) คำนวณได้จากค่าคะแนนของข้อมูล แบ่งออกเป็น 2 ชนิดคือ
   - ค่าเฉลี่ยของประชากร เขียนสัญลักษณ์แทนด้วย μ
   - ค่าเฉลี่ยของตัวอย่าง เขียนสัญลักษณ์แทนด้วย $\overline{X}$

2. ค่ามัธยฐาน (Mode) เป็นค่าที่อยู่ตำแหน่งตรงกึ่งกลางของข้อมูล เมื่อเรียงลำดับข้อมูลแล้ว ซึ่งแบ่งออกเป็น 2 ชนิดคือ
   - ค่ามัธยฐานของประชากร เขียนสัญลักษณ์แทนด้วย τ
   - ค่ามัธยฐานของตัวอย่าง เขียนสัญลักษณ์แทนด้วย $\widetilde{M}$

บทที่ 2

ทฤษฎีและวรรณกรรมที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีและวรรณกรรมที่เกี่ยวข้องกับการวิจัย ประกอบด้วย 3 ส่วน ดังนี้คือ
ส่วนที่ 1 การแจกแจงที่เกี่ยวข้อง
ส่วนที่ 2 ทฤษฎีที่เกี่ยวข้อง
ส่วนที่ 3 งานวิจัยที่เกี่ยวข้อง

ส่วนที่ 1 การแจกแจงที่เกี่ยวข้องกับการวิจัย

ประกอบด้วยการแจกแจง 5 การแจกแจง ดังนี้

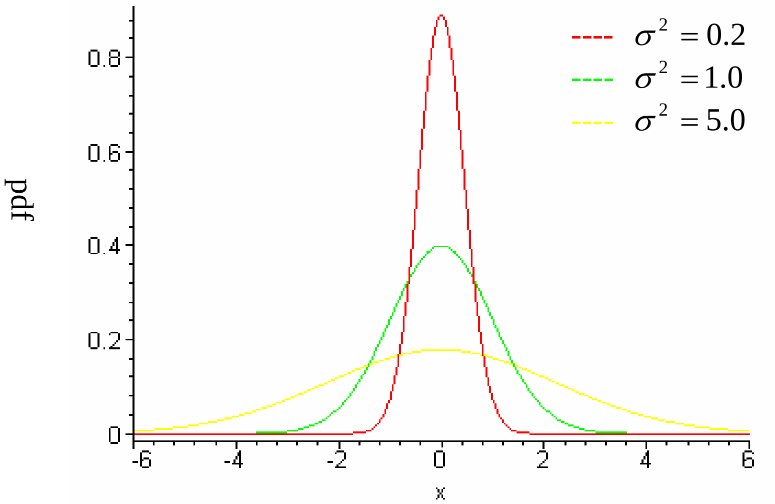
1. การแจกแจงแบบปกติ (Normal Distribution)

การแจกแจงแบบปกติ เป็นการแจกแจงที่สำคัญมากในทางสถิติ เนื่องจากการแจกแจงของค่าที่ได้จากการสังเกตต่างๆมักจะสามารถอธิบายได้ด้วยการแจกแจงแบบปกติ นอกจากนี้ยังสามารถใช้การแจกแจงแบบปกติเพื่อ ประมาณการแจกแจงแบบอื่น ๆ โดยใช้ทฤษฎีลิมิตเข้าสู่ส่วนกลาง (central limit theorem)

ตัวแปรสุ่ม X ซึ่งมีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ย μ และความแปรปรวน σ^2 มีฟังก์ชันความน่าจะเป็น คือ $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$ $\quad -\infty < x < \infty$
กราฟของฟังก์ชันความน่าจะเป็นมีลักษณะเป็นรูประฆังคว่ำ โดยการแจกแจงจะขึ้นอยู่กับค่า μ และ σ^2 ดังรูปที่ 1

โดยรูปที่ 1 เป็นกราฟของฟังก์ชันความน่าจะเป็นเมื่อมีค่า $\mu = 0$ และ $\sigma^2 = 0.2, 1.0$ และ 5.0 ตามลำดับ

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ กราฟของ $f(x)$ จะมีค่าเข้าใกล้ 0 เมื่อ x เข้าใกล้ $-\infty$ หรือ $+\infty$ และจะสมมาตรที่ $x = \mu$ โดยมีค่าเฉลี่ย $E(X) = \mu$ และความแปรปรวน $Var(X) = \sigma^2$

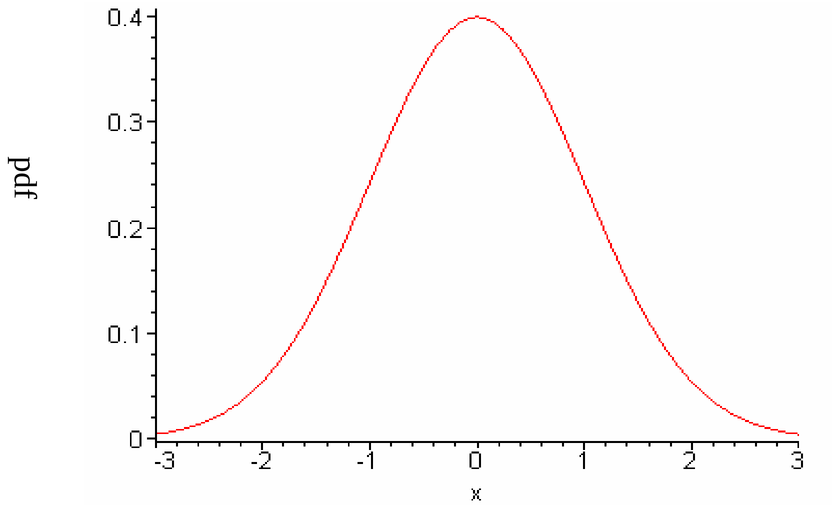


รูปที่ 1 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบปกติ

การแจกแจงแบบปกติที่มี $\mu = 0$ และ $\sigma^2 = 1$ จะเรียกว่า การแจกแจงแบบปกติมาตรฐาน ถ้า Z เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติมาตรฐาน จะได้ว่า ตัวแปรสุ่ม Z จะมีฟังก์ชันความน่าจะเป็นคือ

$$f(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-z^2/2} \qquad -\infty < z < \infty$$

เราสามารถแปลงตัวแปรสุ่ม X ซึ่งมีการแจกแจงแบบปกติด้วยค่าเฉลี่ย μ และความแปรปรวน σ^2 ให้เป็นตัวแปรสุ่มที่มีการแจงแจงแบบปกติมาตรฐาน โดยใช้การแปลงในรูป $Z = \dfrac{X-\mu}{\sigma}$ กราฟของฟังก์ชันความน่าจะเป็นที่มีการแจกแจงแบบปกติมาตรฐานมีลักษณะเป็นรูประฆังคว่ำ ดังรูปที่ 2



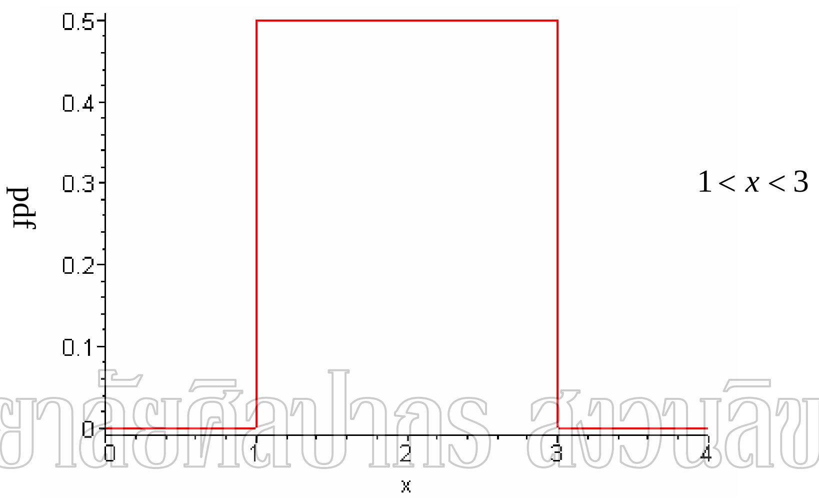
รูปที่ 2 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบปกติมาตรฐาน

2. การแจกแจงแบบยูนิฟอร์ม (Uniform Distribution)

ตัวแปรสุ่ม X มีการแจกแจงความน่าจะเป็นแบบยูนิฟอร์ม ด้วยพารามิเตอร์ a และ b จะมีฟังก์ชันความน่าจะเป็น คือ

$$f(x) = \frac{1}{b-a} \qquad \text{เมื่อ} \quad a < x < b$$
$$= 0 \qquad \text{ที่อื่น ๆ}$$

กราฟของฟังก์ชันความน่าจะเป็น ดังรูปที่ 3



รูปที่ 3 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบยูนิฟอร์ม

ตัวแปรสุ่ม X ซึ่งมีการแจกแจงแบบยูนิฟอร์มในช่วง (a,b) มีค่าเฉลี่ยและความแปรปรวนคือ

$$E(X) = \frac{b+a}{2} \qquad \text{และ} \quad Var(X) = \frac{(b-a)^2}{12}$$

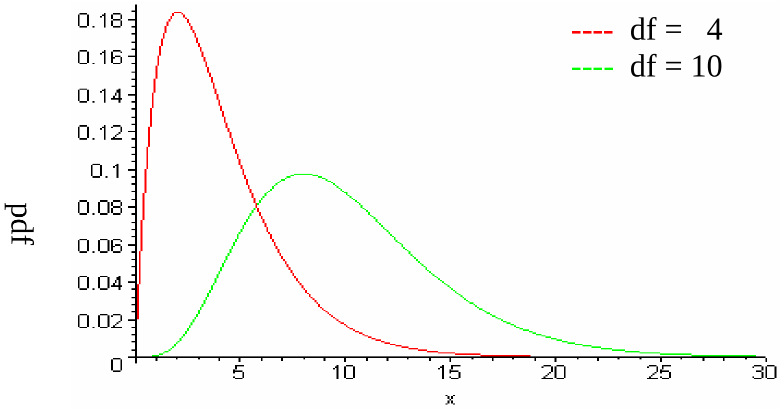
3. การแจกแจงแบบไคสแควร์ (Chi-square Distribution)

ให้ $Y_1, Y_2, ..., Y_\upsilon$ เป็นตัวแปรสุ่มอิสระ มีการแจกแจงแบบปกติมาตรฐาน $N(0,1)$ เมื่อ χ^2 แทนผลบวกของกำลังสองของ Y_i ; $i = 1,...,\upsilon$ จะได้

$\chi^2 = Y_1^2 + Y_2^2 + ... + Y_\upsilon^2$ เป็นตัวแปรสุ่ม จะมีฟังก์ชันความน่าจะเป็นคือ

$$f(\chi^2) = \frac{1}{2^{\upsilon/2}\Gamma(\upsilon/2)} e^{-\chi^2/2} (\chi^2)^{\upsilon/2-1} \qquad ; \qquad (\chi^2 \geq 0)$$

เราเรียกการแจกแจงของตัวแปรสุ่มที่มีฟังก์ชันความน่าจะเป็นดังกล่าวว่าการแจกแจงแบบไคสแควร์ มีค่า $E(\chi^2) = \upsilon$ และ $Var(\chi^2) = 2\upsilon$ โดยที่ υ แทน องศาอิสระ (degrees of freedom) กราฟของฟังก์ชันความน่าจะเป็น ดังรูปที่ 4



รูปที่ 4 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบไคสแควร์

4. การแจกแจงแบบเอกซ์โปเนนเชียล (Exponential Distribution)

เป็นการแจกแจงรูปแบบพิเศษของการแจกแจงแบบแกมมา (Gamma Distribution) ซึ่งมีฟังก์ชันความน่าจะเป็นคือ

$$f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad ; \quad x > 0$$

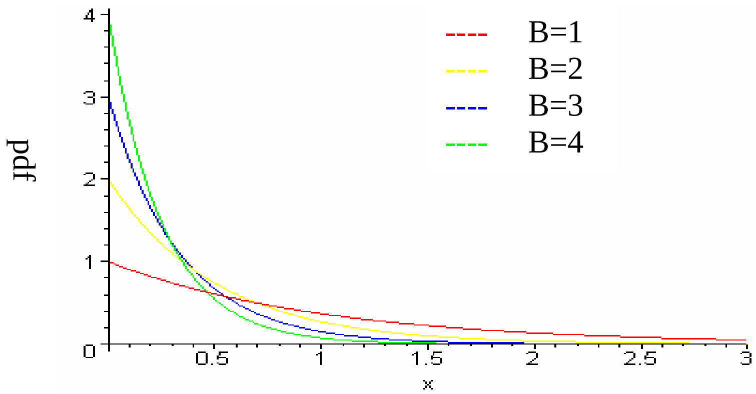
มีค่า $E(X) = \dfrac{\alpha}{\beta}$ และ $Var(X) = \alpha/\beta^2$

กรณีที่การแจกแจงแบบแกมมามีพารามิเตอร์ $\alpha = 1$ จะคือการแจกแจงแบบเอกซ์โปเนนเชียลที่มีพารามิเตอร์ $\beta > 0$ มีฟังก์ชันความน่าจะเป็น คือ

$$f(x) = \beta e^{-\beta x} \qquad x > 0$$
$$= 0 \qquad \text{ที่อื่น ๆ}$$

มีค่า $E(X) = \dfrac{1}{\beta}$ และ $Var(X) = 1/\beta^2$

โดยกราฟของฟังก์ชันความน่าจะเป็นจะขึ้นอยู่กับค่า β ดังตัวอย่างในรูปที่ 5 เมื่อกำหนด $\beta = 1,2,3$ และ 4 ตามลำดับ



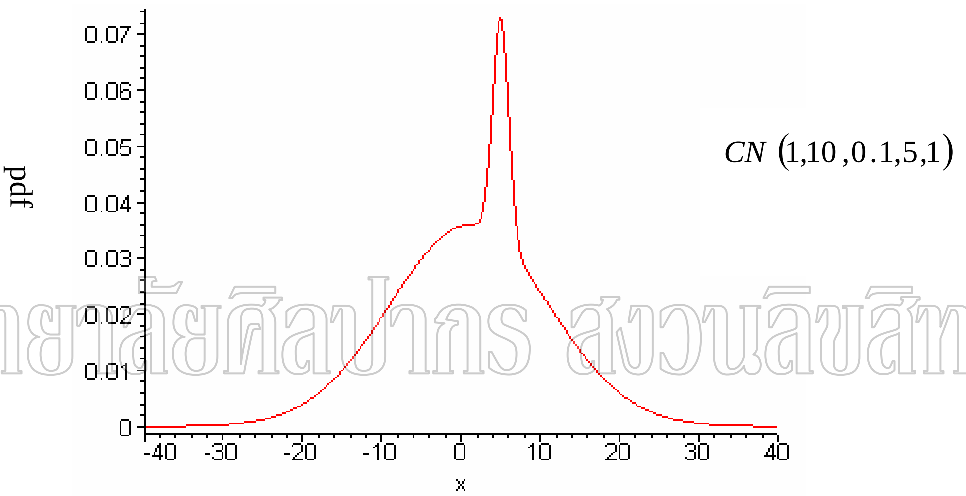
รูปที่ 5 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบเอกซ์โปเนนเชียล

5. การแจกแจงแบบปลอมปน (Contraminate Distribution)

เป็นการแจกแจงที่สร้างขึ้นจากข้อมูล 2 ส่วน ข้อมูลส่วนแรกสุ่มมาจากการแจกแจงแบบปกติด้วยค่าเฉลี่ย μ_1 ความแปรปรวน σ_1^2 และข้อมูลส่วนที่ 2 สุ่มมาจากการแจกแจงแบบปกติด้วยค่าเฉลี่ย μ_2 ความแปรปรวน σ_2^2 ด้วยสัดส่วน p เขียนแทนด้วย $CN\left(\mu_1,\sigma_1^2,p,\mu_2,\sigma_2^2\right)$ ฟังก์ชันความน่าจะเป็นคือ

$$f(x)=(1-p)N\left(\mu_1,\sigma_1^2\right)+pN\left(\mu_2,\sigma_2^2\right)$$

กราฟของฟังก์ชันความน่าจะเป็น ดังรูปที่ 6



รูปที่ 6 กราฟฟังก์ชันความน่าจะเป็นของการแจกแจงแบบปลอมปน

ส่วนที่ 2 ทฤษฎีที่เกี่ยวข้อง

1. ตัวสถิติทดสอบที่ใช้ในการวิจัย

1.1. ตัวสถิติทดสอบเอฟ (F – Test)

การทดสอบสมมติฐานเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยของประชากร 2 กลุ่ม ตัวสถิติทดสอบที่นิยมใช้คือตัวสถิติทดสอบ t (t-test) แต่ในทางปฏิบัติเราอาจสนใจทำการศึกษาหรือทดลองกับกลุ่มตัวอย่างที่สุ่มจากประชากรมากกว่า 2 กลุ่มขึ้นไป และถ้าเราต้องการทดสอบว่าค่าเฉลี่ยของประชากรเหล่านี้แตกต่างกันหรือไม่ ถ้าหากเราทำการทดสอบด้วยตัวสถิติทดสอบ t โดยการเปรียบเทียบค่าเฉลี่ยทีละคู่ เช่นต้องการทดสอบค่าเฉลี่ยของตัวอย่าง 4 กลุ่ม จะต้องทำการทดสอบค่าเฉลี่ยถึง 6 คู่ด้วยกัน หรือถ้ากลุ่มตัวอย่าง 15 กลุ่ม จะต้องทำการทดสอบถึง 105 คู่

ซึ่งจะเห็นว่าถ้ากลุ่มตัวอย่างมีจำนวนมาก การทดสอบความแตกต่างของค่าเฉลี่ยเกิดความยุ่งยากและใช้เวลามาก นอกจากนี้ยังพบว่าในการดำเนินการทดสอบค่าเฉลี่ยทีละคู่ทำให้ความคลาดเคลื่อนชนิดที่ 1 (Type I Error) เพิ่มมากขึ้นกว่าที่กำหนด เช่นถ้าในการทดสอบสมมติฐานกำหนดระดับความมีนัยสำคัญที่ระดับ α ความน่าจะเป็นที่จะไม่เกิดความคลาดเคลื่อนชนิดที่ 1 ในการเปรียบเทียบความแตกต่างค่ากลางของตัวอย่างแต่ละกลุ่มที่อิสระเท่ากับ $1-\alpha$ ถ้ามีการเปรียบเทียบความแตกต่างที่เป็นอิสระจำนวน C กลุ่ม ความน่าจะเป็นที่จะไม่เกิดความคลาดเคลื่อนชนิดที่ 1 ของการเปรียบเทียบความแตกต่างค่ากลางของประชากรที่อิสระเท่ากับ $(1-\alpha)^{C-1}$ ซึ่งคือความน่าจะเป็นที่จะยอมรับสมมติฐานว่าง H_0 เมื่อสมมติฐานว่างทั้ง C กลุ่มเป็นจริง ดังนั้นความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 ในการเปรียบเทียบอย่างน้อย 1 ครั้งในการเปรียบเทียบความแตกต่างจำนวน C ครั้งเท่ากับ $1-(1-\alpha)^{C-1}$ เช่น ถ้าทดสอบความแตกต่างระหว่างค่าเฉลี่ยจำนวน 4 ครั้ง กำหนดระดับนัยสำคัญในแต่ละครั้งของการทดสอบเท่ากับ 0.01 ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 อย่างน้อย 1 ครั้งในการเปรียบเทียบความแตกต่างมีค่าเท่ากับ $1-(1-0.01)^{4-1}=0.029$ ซึ่งมากกว่าระดับนัยสำคัญที่ตั้งไว้ 0.01

Fisher (1923) ได้คิดวิธีการที่เหมาะสม ซึ่งสามารถสรุปผลการทดสอบเกี่ยวกับความแตกต่างระหว่างค่าเฉลี่ยของประชากรหลายกลุ่มได้อย่างถูกต้อง และรวดเร็ว โดยทำการทดสอบสมมติฐานเพียงครั้งเดียวด้วยวิธีการที่เรียกว่า การวิเคราะห์ความแปรปรวน ซึ่งหลักการวิเคราะห์ความแปรปรวนสำหรับข้อมูลจำแนกทางเดียวมีดังนี้

กำหนดให้ $\underline{X}_1=(X_{11},...,X_{1n_1}), \underline{X}_2=(X_{21},...,X_{2n_2}),..., \underline{X}_C=(X_{1C},...,X_{Cn_C})$ เป็นตัวอย่างสุ่มที่สุ่มอย่างอิสระขนาด n_j ; $j=1,2,3,...,C$ จากประชากร C กลุ่มที่มีการแจกแจงแบบปกติ และมีความแปรปรวนเท่ากัน โดยที่ $C\geq 2$ ต้องการทดสอบว่าประชากร C กลุ่มมีค่าเฉลี่ยเท่ากันหรือไม่ สมมติฐานที่ต้องการทดสอบ

$H_0:\ \mu_1=\mu_2=...=\mu_C$ แย้งกับ $H_1:\ \mu_h\neq\mu_j$ สำหรับบางคู่ $(h\neq j)$

โดยที่ $\mu_1,\mu_2,...,\mu_C$ เป็นค่ากลางซึ่งในกรณีนี้คือค่าเฉลี่ยประชากรกลุ่มที่ j

ถ้ากำหนดให้ x_{ji} แทน ค่าสังเกตที่ i ของตัวอย่างสุ่มขนาด n_j จากประชากรกลุ่มที่ j

$x_{j.}=\sum_{i=1}^{n_j}x_{ji}$ แทน ผลรวมของกลุ่มที่ j

$\bar{x}_{j.}=\dfrac{\sum_{i=1}^{n_j}x_{ji}}{n_j}$ แทน ค่าเฉลี่ยของกลุ่มที่ j

$x_{..}=\sum_{j=1}^{C}\sum_{i=1}^{n_j}x_{ji}$ แทน ผลรวมของค่าสังเกตทุก ๆ ค่าของทุก ๆ กลุ่ม

$\bar{x}_{..} = \frac{1}{n}\sum_{j=1}^{C}\sum_{i=1}^{n_j} x_{ji}$ แทน ค่าเฉลี่ยรวม

$n = \sum_{j=1}^{C} n_j$ แทน จำนวนหน่วยตัวอย่างทั้งหมด

เราสามารถใช้วิธีการวิเคราะห์ความแปรปรวน แยกความผันแปรของข้อมูลออกตามสาเหตุที่ทำให้ข้อมูลแตกต่างกัน คือความผันแปรระหว่างกลุ่มตัวอย่างกับความผันแปรภายในกลุ่มตัวอย่าง ได้ดังนี้

$$x_{ji} - \bar{x}_{..} = \left(\bar{x}_{j.} - \bar{x}_{..}\right) + \left(x_{ji} - \bar{x}_{j.}\right)$$

ยกกำลังสองทั้งสองข้าง ผลคูณร่วมมีค่าเป็น 0 ได้เป็น

$$\sum_{j=1}^{C}\sum_{i=1}^{n_j}\left(x_{ji} - \bar{x}_{..}\right)^2 = \sum_{i=1}^{n_j} n_j\left(\bar{x}_{j.} - \bar{x}_{..}\right)^2 + \sum_{j=1}^{C}\sum_{i=1}^{n_j}\left(x_{ji} - \bar{x}_{j.}\right)^2$$

เขียนสัญลักษณ์แทนได้ดังนี้

ผลรวมกำลังสองของทั้งหมด = ผลรวมกำลังสองระหว่างกลุ่ม + ผลรวมกำลังสองภายในกลุ่ม

$$SST = SSB + SSW$$

เราสามารถพิสูจน์ได้ว่า SSB และ SSW มีการแจกแจงแบบไคสแควร์ด้วยจำนวนองศาแห่งความมีอิสระ $C-1$ และ $n-C$ ตามลำดับ

เมื่อนำผลบวกกำลังสองของความผันแปรจากแหล่งต่าง ๆ หารด้วยองศาอิสระจะได้ค่าเฉลี่ยกำลังสอง นั่นคือ

ค่าเฉลี่ยกำลังสองระหว่างกลุ่ม = ผลรวมกำลังสองระหว่างกลุ่ม / องศาอิสระ

ใช้สัญลักษณ์แทนด้วย MSB นั่นคือ $MSB = \frac{SSB}{df_b}$; $df_b = C-1$ ทำนองเดียวกัน

ค่าเฉลี่ยกำลังสองภายในกลุ่ม = ผลรวมกำลังสองภายในกลุ่ม / องศาอิสระ

หรือ $MSW = \frac{SSW}{df_w}$; $df_w = n-C$

ในการทดสอบสมมติฐานจะพิจารณาอัตราส่วนของความแปรปรวนระหว่างกลุ่มตัวอย่างและความแปรปรวนภายในกลุ่มตัวอย่าง ดังนี้

$$\frac{(n-C)\sum_{j=1}^{C} n_j\left(\bar{x}_{j.} - \bar{x}..\right)^2}{(C-1)\sum_{j=1}^{C}\sum_{i=1}^{n_i}\left(x_{ji} - \bar{x}_{j.}\right)^2}$$

ซึ่งอัตราส่วนดังกล่าวจะมีการแจกแจงแบบ F ด้วยองศาแห่งความมีอิสระ $df = (C-1),(n-C)$ ดังนั้นการทดสอบสมมติฐานจึงขึ้นอยู่กับตัวสถิติทดสอบ F

ถ้าสมมติฐาน $H_0: \quad \mu_1 = \mu_2 = ... = \mu_C$ เป็นจริงค่าเฉลี่ยกำลังสองระหว่างกลุ่มจะมีค่าเข้าใกล้ 0 ทำให้ตัวสถิติทดสอบ F มีค่าน้อย แต่ถ้าสมมติฐานไม่ถูกต้องความแตกต่างระหว่าง $\bar{x}_{j.}$ กับ $\bar{x}_{..}$ มีค่ามาก ทำให้ค่าตัวสถิติทดสอบ F มีค่ามากตามไปด้วย จนกระทั่งมีค่ามากกว่าค่าที่เปิดจากตารางจึงปฏิเสธสมมติฐาน H_0 สามารถแสดงเป็นตารางการวิเคราะห์ความแปรปรวนสำหรับข้อมูลจำแนกทางเดียวได้ดังตารางที่ 1

ตารางที่ 1 การวิเคราะห์ความแปรปรวนจำแนกทางเดียว

Sourt of variation (SV)	degree of freedom (df)	sum of squares (SS)	Mean square (MS)	F value (CR_F)
Between sample	$C-1$	SSB	MSB	$\frac{MSB}{MSW} = F$
With in sample	$n-C$	SSW	MSW	
total	$n-1$	$SSB+SSW$		

ปฏิเสธสมมติฐาน H_0 เมื่อ $F_{cal} \geq F_{(1-\alpha),(C-1),(n-C)}$

1.2. ตัวสถิติทดสอบครัสคัล-วอลลิส (Kruskal - Wallis Test)

ในการทดสอบการเท่ากันของค่าเฉลี่ยสำหรับประชากร 2 กลุ่มเมื่อตัวอย่างถูกสุ่มมาจากประชากรที่ไม่ได้มีการแจกแจงแบบปกติ โดยใช้ตัวสถิติทดสอบ t อาจไม่ถูกต้องมากนัก Wilcoxon (1945) เสนอตัวสถิติทดสอบค่ากลางซึ่งในที่นี้จะพิจารณาจากค่ามัธยฐาน กรณีตัวอย่างมีขนาดเท่ากัน โดยใช้การเรียงลำดับ (ranks) ของข้อมูล และพิจารณาว่าผลรวมของลำดับของข้อมูลทั้งสองกลุ่มต่างมากเกินไปหรือไม่ ถ้าประชากรทั้ง 2 กลุ่มมีค่ามัธยฐานไม่แตกต่างกัน ผลรวมของอันดับของข้อมูลทั้งสองกลุ่มจะไม่แตกต่างกันมากนัก สามารถดำเนินการทดสอบโดยเริ่มจากรวมข้อมูลทั้งสองกลุ่มเข้าด้วยกันแล้วจัดลำดับ (ranks) จากค่าน้อยสุดเป็นลำดับที่ 1 ค่าที่ถัดมาเป็นลำดับที่ 2 จนถึงค่าสูงสุดเป็นลำดับสุดท้าย แล้วหาผลรวมลำดับของข้อมูลแต่ละกลุ่ม คำนวณค่าตัวสถิติ S เปรียบเทียบกับค่าจากตาราง ถ้าค่าตัวสถิติ S มากกว่าค่าตามตารางจะปฏิเสธสมมติฐาน ต่อมา Mann and Whitney (1947) เสนอตัวสถิติ Mann-Whitney U Test ซึ่งเป็นตัวสถิติทดสอบที่พัฒนามาจากตัวสถิติ Wilcoxon(1945) ใช้ในกรณีที่ขนาดตัวอย่างไม่เท่ากัน โดยดำเนินการทดสอบเช่นเดียวกัน เพียงแต่พิจารณาตัวสถิติ U จากผลรวมของจำนวนที่ตัวอย่างกลุ่มแรกมากกว่าตัวอย่างกลุ่มที่สอง ถ้าค่าตัวสถิติ U มากกว่าค่าตามตารางจะปฏิเสธสมมติฐาน

Kruskal and Wallis (1952) ได้พัฒนาการทดสอบสำหรับประชากร C กลุ่มมาจาก Mann-Whitney U Test เพื่อใช้ทดสอบว่ากลุ่มตัวอย่างอิสระ C กลุ่ม มีการแจกแจงเดียวกันหรือไม่ หรือกลุ่มตัวอย่างอิสระ C กลุ่ม ถูกสุ่มมาจากประชากรที่มีค่ากลางเท่ากันหรือไม่ ซึ่งเป็นวิธีการที่เทียบได้กับการวิเคราะห์ความแปรปรวนของข้อมูลจำแนกทางเดียว แต่ไม่จำเป็นจะต้องมีข้อตกลงเบื้องต้น ว่าประชากรที่ตัวอย่างสุ่มมาจะต้องมีการแจกแจงปกติและมีความแปรปรวนเท่ากัน เพียงแต่ตัวอย่างต้องมาจากประชากรที่มีการแจกแจงเหมือนซึ่งอาจมีค่ากลางแตกต่างกันได้ ค่าสถิติที่ใช้ในการทดสอบของครัสคัล-วอลลิส เรียกว่า ค่า H ดังนั้นจึงเรียกการทดสอบชนิดนี้ว่า Kruskal - Wallis H Test หรือ H-Test โดยมีวิธีการดังนี้

กำหนดให้ $\underline{X}_1 = \left(X_{11},\ldots,X_{1n_1}\right), \underline{X}_2 = \left(X_{21},\ldots,X_{2n_2}\right),\ldots,\underline{X}_C = \left(X_{1C},\ldots,X_{Cn_C}\right)$ เป็นตัวอย่างสุ่ม ขนาด n_j ; $j = 1,2,3,\ldots,C$ ที่สุ่มมาอย่างอิสระจากประชากร C กลุ่มเมื่อ $C \geq 2$ ต้องการทดสอบว่าประชากร C กลุ่มมีค่ามัธยฐานเท่ากันหรือไม่ โดยข้อมูลมีมาตราการวัดอย่างน้อยต้องเป็นมาตราอันดับ

สมมติฐานทดสอบ $H_0 :\ \tau_1 = \tau_2 = \ldots = \tau_C$ แย้งกับ

$H_1 :\ \tau_h \neq \tau_j$ สำหรับบางคู่ $(h \neq j)$

เมื่อ $\tau_1 = \tau_2 = \ldots = \tau_C$ เป็นค่ามัธยฐานของตัวอย่างกลุ่มที่ $1,2,3,\ldots,C$ นำข้อมูลทั้ง C กลุ่มมารวมกันแล้วเรียงลำดับ (Rank) จากค่าต่ำสุดเป็นอันดับ 1 ค่าถัดมาเป็นอันดับ 2 จนถึงค่าสูงสุดเป็นอันดับสุดท้าย แล้วหาผลบวกของลำดับของข้อมูลแต่ละกลุ่ม

ให้ R_{ij} แทน Rank ของ X_{ij} ในตัวอย่างที่รวมกัน C กลุ่ม

$R_{.j} = \sum_{i=1}^{n_j} R_{ij}$ แทน ผลรวมของ rank ที่ได้จากกลุ่มที่ j โดย $j = 1,2,\ldots,C$

เมื่อ $\overline{R}_{.j} = \dfrac{R_{.j}}{n_j}$ และ $n = \sum_{j=1}^{C} n_j$

โดยที่ค่าของ Rank จะต้องไม่ซ้ำกัน ถ้าค่าซ้ำกันต้องทำการเฉลี่ยค่าของ rank

ถ้าค่ามัธยฐานของแต่ละกลุ่มไม่แตกต่างกัน ค่าเฉลี่ยของอันดับในแต่ละกลุ่มจะไม่ต่างจากค่าเฉลี่ยของอันดับของทุกกลุ่มมากนัก หรือ $\overline{R}_{.j} - (n+1)/2$ ไม่ควรแตกต่างมากนัก เนื่องจาก $\dfrac{\overline{R}_{.j} - (n+1)/2}{\sqrt{Var\left(\overline{R}_{.j}\right)}}$ มีลิมิตการแจกแจงลู่เข้าสู่ $N(0,1)$

โดย $VAR\left(\overline{R}_{.j}\right) = \left(n - n_j\right)(n+1)/12n_j$ และเมื่อ $\overline{R}_{.j}$, $j = 1,2,\ldots,C$ มีความอิสระกัน กำลังสองของตัวแปรสุ่มแบบปกติมาตรฐานจะประมาณได้ด้วยการแจกแจงแบบไคสแควร์ ดังนั้น ตัวสถิติทดสอบ H จะอยู่ในรูป

$$H = \sum D_{jn}^2 \left\{ \frac{\overline{R}_{.j} - (n+1)/2}{\sqrt{Var(\overline{R}_{.j})}} \right\}^2$$

เนื่องจาก $\overline{R}_{.1}, \overline{R}_{.2}, ..., \overline{R}_{.C}$ ไม่ได้เป็นอิสระกันทั้งหมด จะมีความเป็นอิสระ $C-1$ กลุ่ม ดังนั้น เพื่อให้การแจกแจงของตัวสถิติ H ลู่เข้าสู่การแจกแจงแบบไคสแควร์ เมื่อตัวอย่างมีขนาดใหญ่ จึงจำเป็นต้องเลือก D_{jn}^2 ที่เหมาะสมเพื่อเป็นตัวปรับให้การแจกแจงของ H มีการแจกแจงเข้าสู่การแจกแจงแบบไคสแควร์ ด้วยจำนวนองศาอิสระ $C-1$ ซึ่งค่า D_{jn}^2 ที่เหมาะสมคือ $D_{jn}^2 = 1 - \frac{n_j}{n}$ จะได้ว่า

$$H = \sum_{j=1}^{C} \left(1 - \frac{n_j}{n}\right) \left\{ \frac{\overline{R}_{.j} - (n+1)/2}{\sqrt{\frac{(n-n_j)(n+1)}{12 n_j}}} \right\}^2$$

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{C} \frac{R_{.j}^2}{n_j} - 3(n+1)$$

ปฏิเสธสมมติฐาน H_0 เมื่อ $H > \chi^2_{1-\alpha, C-1}$

1.3. ตัวสถิติทดสอบ Bi-aspect

โดยทั่วไปตัวสถิติทดสอบที่ใช้ทดสอบสมมติฐาน มักทดสอบด้านเชิงตัวเลข (numerical aspect) หรือ ด้านเชิงกลุ่ม (categorical aspect) เพียงด้านเดียว เช่นตัวสถิติทดสอบ t ใช้ทดสอบความแตกต่างของค่าเฉลี่ยสำหรับประชากร 2 กลุ่มที่มีการแจกแจงแบบปกติ ตัวสถิติทดสอบ F ก็เช่นเดียวกันเป็นตัวสถิติทดสอบความแตกต่างของค่าเฉลี่ยในกรณีประชากรหลายกลุ่มที่มีการแจกแจงแบบปกติ เป็นตัวสถิติทดสอบทางด้านเชิงตัวเลข ส่วนตัวสถิติที่สนใจทดสอบเฉพาะด้านเชิงกลุ่ม (categorical aspect) เป็นตัวสถิติทดสอบที่แบ่งข้อมูลออกเป็น 2 กลุ่ม คือกลุ่มที่สูงกว่ามัธยฐานร่วมและกลุ่มที่ต่ำกว่าหรือเท่ากับมัธยฐานร่วม โดยไม่สนใจการแจกแจงของข้อมูล เช่น ตัวสถิติ Median test ใช้สำหรับทดสอบความแตกต่างของค่ามัธยฐานของประชากร 2 กลุ่มและขยายไปสู่การทดสอบประชากรมากกว่า 2 กลุ่มได้ โดยเริ่มจากการหามัธยฐานร่วมของข้อมูลแล้วทำการแบ่งข้อมูลออกเป็น 2 กลุ่มคือกลุ่มที่มีค่าสูงกว่ามัธยฐานกับกลุ่มที่มีค่าต่ำกว่ามัธยฐานตามตารางการณ์จร ดังตารางที่ 2

ตารางที่ 2 การแบ่งกลุ่มของข้อมูล

จำนวนข้อมูลที่มีค่า	กลุ่มตัวอย่างที่						รวม
	1	2	...	j	…	C	
สูงกว่ามัธยฐาน	a_1	a_2	...	a_j		a_C	a
ต่ำกว่ามัธยฐาน	b_1	b_2	...	b_j	...	b_C	b
รวม	n_1	n_2	…	n_j	…	n_C	n

แล้วจึงเปรียบเทียบสัดส่วนของตัวอย่างที่อยู่ในกลุ่มที่ j เทียบกับสัดส่วนคาดหวัง ตัวสถิติทดสอบอยู่ในรูป

$$\chi^2_{C-1} = \frac{(n-1)}{ab}\sum_{j=1}^{C}\frac{(na_j - n_j a)^2}{nn_j}$$

เมื่อ a_j, b_j แทน จำนวนข้อมูลที่อยู่สูงกว่าและต่ำกว่ามัธยฐานในกลุ่มที่ j ตามลำดับ

a, b แทน ผลรวมของจำนวนข้อมูลที่อยู่สูงกว่าและต่ำกว่ามัธยฐาน ตามลำดับ

n_i แทน จำนวนข้อมูลกลุ่มที่ j , $j = 1,2,...,C$

n แทน จำนวนข้อมูลทั้งหมด

$n = n_1 + n_2 + ... + n_C$ หรือ $n = a + b$

$a = a_1 + a_2 + ... + a_C$

$b = b_1 + b_2 + ... + b_C$

ซึ่งตัวสถิติทดสอบนี้จะมีการแจกแจงแบบไคสแควร์ ด้วยองศาอิสระ $C-1$ การทดสอบสมมติฐานจะเปรียบเทียบค่าตัวสถิติกับค่าจากตารางไคสแควร์ซึ่งจะปฏิเสธสมมติฐาน H_0 เมื่อ $\chi^2_{cal} > \chi^2_{table}$

ในปี ค.ศ. 2003 Macro Marozzi ได้เสนอตัวสถิติทดสอบที่เป็นการรวมตัวสถิติทดสอบเชิงตัวเลขและตัวสถิติทดสอบเชิงกลุ่มเข้าไว้ด้วยกัน และพิจารณาตัวสถิติทั้งสองพร้อมกัน โดยไม่จำเป็นต้องกำหนดลักษณะการแจกแจงของประชากรที่ตัวอย่างถูกสุ่มมา ใช้ทดสอบความแตกต่างของค่ากลาง สำหรับประชากร 2 กลุ่ม โดยทำการทดสอบภายใต้วิธีการ permutation test (ดูรายละเอียดหัวข้อ 2) และเรียกตัวสถิตินี้ว่า Bi-aspect test ซึ่งสามารถทดสอบทั้งเชิงตัวเลขและเชิงกลุ่มพร้อมกัน ซึ่งวิธีการของตัวสถิติทดสอบ Bi-aspect test สำหรับประชากร 2 กลุ่ม มีวิธีการดังต่อไปนี้

กำหนดให้ $\underline{X}_1 = \left(X_{11},\ldots,X_{1n_1}\right)$ และ $\underline{X}_2 = \left(X_{21},\ldots,X_{2n_2}\right)$ เป็นตัวอย่างที่สุ่มมาจากประชากรที่มีการแจกแจงแบบต่อเนื่อง มีการแจกแจงเดียวกัน และมีความแปรปรวนเท่ากัน นั่นคืออาจจะมีความแตกต่างระหว่างค่ากลางเพียงอย่างเดียว $n_1 + n_2 = n$ ต้องการทดสอบว่าประชากร 2 กลุ่มมีค่ากลางเท่ากันหรือไม่

มีสมมติฐานการทดสอบ คือ

$H_0 : \ \vartheta = 0$ แย้งกับ $H_1 : \ \vartheta > 0$

เมื่อ $\vartheta = \delta_1 - \delta_2$

และ δ_j แทน ค่ากลางของประชากรกลุ่มที่ j , $j = 1,2$

ซึ่งในการดำเนินการทดสอบตัวสถิติมีขั้นตอน 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 ทำการทดสอบตัวสถิติแต่ละด้าน คือ ด้านเชิงตัวเลข กับ ด้านเชิงกลุ่ม

1.1 ด้านเชิงตัวเลข (numerical aspect) จะดูว่าผลรวมของค่าสังเกตจากตัวอย่างที่ถูกสุ่มมาว่าทั้ง 2 กลุ่มเท่ากันหรือไม่ นิยามดังนี้

$PT = T_a = \sum_{i=1}^{n_1} X_{1i}$; X_{1i} แทน ค่าสังเกตของหน่วยตัวอย่างที่ i ของกลุ่ม 1

n_1 แทน จำนวนตัวอย่างของกลุ่ม 1

สมมติฐานการทดสอบ

$H_{0a} : \ \delta_1 = \delta_2$ แย้งกับ $H_{1a} : \ \delta_1 > \delta_2$ หรือ

$H_{0a} : \ \vartheta = 0$ แย้งกับ $H_{1a} : \ \vartheta > 0$

1.2 ตัวสถิติทดสอบตัวที่ 2 เป็นตัวสถิติทดสอบเชิงกลุ่มเพื่อดูว่าข้อมูลทั้ง 2 กลุ่มมีค่ามัธยฐานเท่ากันหรือไม่ โดยแบ่งกลุ่มค่าสังเกตออกเป็น 2 กลุ่มดังนี้

1. กลุ่มใหญ่ คือ กลุ่มที่ค่าสังเกตมีค่ามากกว่าค่ามัธยฐานร่วม
2. กลุ่มเล็ก คือ กลุ่มที่ค่าสังเกตมีค่าน้อยกว่าหรือเท่ากับค่ามัธยฐานร่วม

ซึ่งค่ามัธยฐานร่วมของค่าสังเกตหาได้จาก

$\widetilde{M} = X_{(n+1)/2}$ เมื่อจำนวนค่าสังเกตเป็นจำนวนคู่

$\widetilde{M} = \frac{1}{2}\left(X_{n/2} + X_{(n/2)+1}\right)$ เมื่อจำนวนค่าสังเกตเป็นจำนวนคี่

แสดงได้ดังตารางที่ 3

ตารางที่ 3 การแบ่งข้อมูลแต่ละกลุ่มออกเป็น 2 พวก คือพวกที่มีค่าสูงกว่ามัธยฐานและพวกที่มีค่าต่ำกว่าหรือเท่ากับมัธยฐาน

จำนวนข้อมูลที่มีค่า	กลุ่มตัวอย่างที่		รวม
	กลุ่มที่ 1	กลุ่มที่ 2	
ต่ำกว่าหรือเท่ากับมัธยฐาน	n_{1s}	n_{2s}	n_s
สูงกว่ามัธยฐาน	n_{1g}	n_{2g}	n_g
รวม	n_1	n_2	n

ตัวสถิติทดสอบคือ

$$T_b = n_{1g} = \sum_{i=1}^{n_1} I\left(X_{1i} > \tilde{M}\right)$$ เมื่อ $\tilde{M}$ แทน ค่ามัธยฐานของข้อมูล

$$I\left(X_{1i} > \tilde{M}\right) = 1 \quad \text{เมื่อ} \quad X_{1i} > \tilde{M}$$
$$= 0 \quad \text{ที่อื่น ๆ}$$

สมมติฐานการทดสอบ

$H_{0b} : P\{X_1 > \tilde{M}\} = P\{X_2 > \tilde{M}\}$ แย้งกับ $H_{1b} : P\{X_1 > \tilde{M}\} > P\{X_2 > \tilde{M}\}$

ภายใต้ H_{0b} จริงค่า n_{1g}/n_1 และ n_{2g}/n_2 จะต้องมีค่าไม่แตกต่างกันในขณะที่ภายใต้ H_{1b} จริงค่า n_{1g}/n_1 และ n_{2g}/n_2 จะมีค่าแตกต่างกัน (การ permutation ไม่ทำให้ค่าของ n_g เปลี่ยนแปลง)

วิธีการคำนวณค่า p-value ของตัวสถิติทดสอบ

จากตัวอย่างที่สุ่มมา $\underline{X} = \left(\underline{X}_1, \underline{X}_2\right)$ หาค่าตัวสถิติทดสอบเชิงตัวเลข ${}_0PT$ จาก ${}_0PT = {}_0T_a = \sum_{i=1}^{n_1} X_{1i}$ และตัวสถิติทดสอบเชิงกลุ่ม ${}_0T_b$ จาก ${}_0T_b = n_{1g} = \sum_{i=1}^{n_1} I\left(X_{1i} > \tilde{M}\right)$ ทำการ permutation โดยนำตัวอย่างทั้งสองกลุ่มมารวมกันทำการจัดเป็น 2 กลุ่มใหม่คือ $\underline{X}^* = \left(\underline{X}_1^*, \underline{X}_2^*\right)$ เรียกว่า permutation ครั้งที่ 1 หาค่าตัวสถิติทดสอบ ${}_1PT = {}_1T_a = \sum_{i=1}^{n_1} X_{1i}$ และ ${}_1T_b = n_{1g} = \sum_{i=1}^{n_1} I\left(X_{1i} > \tilde{M}\right)$ นำข้อมูลทั้งสองกลุ่มมารวมกันอีกและทำการ permutation และแบ่งกลุ่มใหม่จนกระทั่งครบจำนวน K ครั้ง แล้วคำนวณหาค่าสัดส่วนของจำนวน permutation ที่มีค่าตัวสถิติมากกว่าหรือเท่ากับค่าของตัวสถิติที่สังเกตได้จากตัวอย่างคือ ${}_0PT$ และ ${}_0T_b$ คำนวณค่า $\hat{L}_{PT}$ และ $\hat{L}_{T_b}$ ตามวิธีการประมาณของ pesarin(2001) ดังนี้

$$\hat{L}_{PT}\left({}_0PT\right)=\frac{1}{K}\sum_{k=1}^{K}I\left({}_K PT \geq {}_0PT\right)$$

โดย ${}_kPT$ แทน ค่าPT ของ permutation ครั้งที่ k $\left(k=1,\ldots,K\right)$ ของ $\underline{X}$

$I(.)$ แทน ฟังก์ชัน $I\left({}_kPT \geq {}_0PT\right)$ = 1 ; ${}_kPT \geq {}_0PT$

= 0 ; ที่อื่น ๆ

$$\hat{L}_{T_b}\left({}_0T_b\right)=\frac{1}{K}\sum_{k=1}^{K}I\left({}_kT_b \geq {}_0T_b\right)$$

โดย ${}_kT_b$ แทน ค่า T_b ของ permutation ครั้งที่ k $\left(k=1,\ldots,K\right)$ ของ $\underline{X}$

เมื่อ $I\left({}_kT_b > {}_0T_b\right)$ = 1 เมื่อ ${}_kT_b \geq {}_0T_b$

= 0 ที่อื่น ๆ

แต่เนื่องจากความไม่ต่อเนื่องและค่า p-value ไม่ควรจะเป็น 0 หรือ 1 จึงต้องทำการปรับค่าโดยการเพิ่มเศษอีก 0.5 และเพิ่มตัวส่วนอีก 1 เพื่อให้ค่าอยู่ในช่วง $]0,1[$ (Pesarin, 2001, p.144) ดังนี้

$$\hat{L}_{T_a}\left({}_0T_a\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_a \geq {}_0T_a\right)\right]\Big/(K+1)$$

เมื่อ $I\left({}_kT_a > {}_0T_a\right)$ = 1 เมื่อ ${}_kT_a \geq {}_0T_a$

= 0 ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0<\alpha<1$ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อ $\hat{L}_{PT} \leq \alpha$

$$\hat{L}_{T_b}\left({}_0T_b\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_b \geq {}_0T_b\right)\right]\Big/(K+1)$$

เมื่อ $I\left({}_kT_b > {}_0T_b\right)$ = 1 เมื่อ ${}_kT_b \geq {}_0T_b$

= 0 ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0<\alpha<1$ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อ $\hat{L}_{T_b} \leq \alpha$

ขั้นตอนที่ 2 เป็นการสร้างตัวสถิติทดสอบที่ใช้ตัวสถิติ 2 ตัวร่วมกันเรียกว่า Bi-aspect test มีสมมติฐานการทดสอบ คือ

$$H_0: \ \vartheta = 0 \quad \text{แย้งกับ} \quad H_1: \ \vartheta > 0$$

คำนวณตัวสถิติได้จากฟังก์ชันที่รวม L_{T_a} และ L_{T_b} เข้าไว้ด้วยกัน เลือกตัวสถิติทดสอบที่มีค่า p-value น้อยกว่า ดังนี้

$$T_{ab} = max\left(1-L_{T_a}, 1-L_{T_b}\right)$$

ซึ่งค่าประมาณตัวสถิติ T_{ab} ที่ได้จากค่าสังเกตคือ

$${}_0\hat{T}_{ab} = max\left(1-\hat{L}_{T_a}\left({}_0T_a\right), 1-\hat{L}_{T_b}\left({}_0T_b\right)\right)$$

และจากการทำการ permutation ครั้งที่ 1 ตัวสถิติ T_{ab} ใน permutation แรก จะได้

$${}_1\hat{T}_{ab} = max\left(1-\hat{L}_{T_a}\left({}_1T_a\right), 1-\hat{L}_{T_b}\left({}_1T_b\right)\right)$$

แล้วทำการคำนวณ permutation อีกจนครบจำนวน K ครั้ง จะได้เวคเตอร์ของตัวสถิติทดสอบ ดังนี้ $T_{ab}\left({}_1T_{ab}, {}_2T_{ab}, \ldots, {}_KT_{ab}\right)$

หาค่า p-value ของตัวสถิติได้จาก

$$\hat{L}_{T_{ab}} = \frac{1}{K}\sum_{k=1}^{K} I\left({}_kT_{ab} \geq {}_0T_{ab}\right)$$

เมื่อ $I\left({}_kT_{ab} > {}_0\hat{T}_{ab}\right) = 1$ เมื่อ ${}_kT_{ab} \geq {}_0\hat{T}_{ab}$

$= 0$ ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0 < \alpha < 1$ จะปฏิเสธสมมติฐาน H_0 เมื่อ $\hat{L}_{T_{ab}} \leq \alpha$

สรุปวิธีการหาค่า p-value ของตัวสถิติ T_{ab} มีขั้นตอนดังต่อไปนี้

1. นำตัวอย่างกลุ่มที่ 1 และตัวอย่างกลุ่มที่ 2 มารวมกันหาค่ามัธยฐานร่วม ($\tilde{M}$)
2. คำนวณหาค่าตัวสถิติทดสอบจากค่าสังเกตแต่ละกลุ่ม ดังนี้

ตัวสถิติทดสอบเชิงตัวเลข ${}_0T_a = \sum_{i=1}^{n_1} X_{1i}$

ตัวสถิติทดสอบเชิงกลุ่ม ${}_0T_b = \sum_{i=1}^{n_1} I\left(X_{1i} > \tilde{M}\right)$

3. ทำการ permutation จะได้ $\underline{X}^*$ เป็นครั้งแรกจาก $\underline{X}$
4. ทำการคำนวณค่าสถิติ T_a และ T_b จาก permutation แรก โดย

ตัวสถิติทดสอบเชิงตัวเลข ${}_1T_a = \sum_{i=1}^{n_1} X_{1i}$

ตัวสถิติทดสอบเชิงกลุ่ม ${}_1T_b = \sum_{i=1}^{n_1} I\left(X_{1i} > \tilde{M}\right)$

5. ทำการคำนวณตัวสถิติทดสอบทั้ง 2 ด้าน คือทำการ permutation $K-1$ ครั้ง โดยดำเนินการเช่นเดียวกับขั้นตอนที่ 3 และขั้นตอนที่ 4
6. คำนวณหาจำนวนตัวสถิติ T_a ที่มีค่ามากกว่าหรือเท่ากับ ${}_0T_a$ หาค่าประมาณค่า p-value $\hat{L}_{T_a}\left({}_0T_a\right)$ ของตัวสถิติ T_a คือ

$$\hat{L}_{T_a}\left({}_0T_a\right) = \left[\frac{1}{2} + \sum_{k=1}^{K} I\left({}_kT_a \geq {}_0T_a\right)\right] \Big/ (K+1)$$

และในทำนองเดียวกันสำหรับตัวสถิติ T_b ค่า p-value ของตัวสถิติ $\hat{L}_{T_b}\left({}_0T_b\right)$ หาได้ดังนี้

$$\hat{L}_{T_b}\left({}_0T_b\right) = \left[\frac{1}{2} + \sum_{k=1}^{K} I\left({}_kT_b \geq {}_0T_b\right)\right] \Big/ (K+1)$$

7. หาค่าประมาณของตัวสถิติทดสอบ T_{ab} (Bi-aspect) โดยเลือกตัวสถิติ ${}_1T_a$ กับ ${}_1T_b$ จากฟังก์ชัน ${}_0\hat{T}_{ab} = max\left(1 - \hat{L}_{T_a}\left({}_0T_a\right), 1 - \hat{L}_{T_b}\left({}_0T_b\right)\right)$

8. คำนวณหาจำนวน permutation ของตัวสถิติ T_a ที่มีค่ามากกว่าหรือเท่ากับ ${}_1T_a$ โดยมีค่า p-value ของตัวสถิติ $\hat{L}_{T_a}\left({}_1T_a\right)$ ดังนี้

$$\hat{L}_{T_a}\left({}_1T_a\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_a\geq{}_1T_a\right)\right]\Big/(K+1)$$

และในทำนองเดียวกันสำหรับตัวสถิติ T_b ค่า p-value ของตัวสถิติ $\hat{L}_{T_b}\left({}_1T_b\right)$ หาได้ดังนี้

$$\hat{L}_{T_b}\left({}_1T_b\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_b\geq{}_1T_b\right)\right]\Big/(K+1)$$

9. หาค่าประมาณของตัวสถิติทดสอบ Bi-aspect ของ permutation แรกจากฟังก์ชัน

$${}_1\hat{T}_{ab}=max\left(1-\hat{L}_{T_a}\left({}_1T_a\right),1-\hat{L}_{T_b}\left({}_1T_b\right)\right)$$

10. ทำเช่นเดียวกับขั้นตอนที่ 8-9 อีก K-1 permutation โดยใช้ ${}_kT_a$ และ ${}_kT_b$ ซึ่ง $k=2,3,\ldots,K$ แทนที่ ${}_1T_a$ และ ${}_1T_b$ ได้เวกเตอร์ของตัวสถิติ T_{ab} คือ $\left({}_1T_{ab},{}_2T_{ab},\ldots,{}_KT_{ab}\right)$

11. คำนวณสัดส่วนของจำนวน permutation ที่ตัวสถิติ T_{ab} มีค่ามากกว่าหรือเท่ากับ ตัวสถิติ ${}_0T_{ab}$ และหาค่า p-value ของตัวสถิติ Bi-aspect ดังนี้

$$\hat{L}_{T_{ab}}=\frac{1}{K}\sum_{k=1}^{K}I\left({}_kT_{ab}\geq{}_0\hat{T}_{ab}\right)$$

12. ปฏิเสธสมมติฐาน H_0 ที่ระดับ $0<\alpha<1$ ถ้า $\hat{L}_{T_{ab}}\leq\alpha$

Macro Marozzi (2004) ได้พัฒนาตัวสถิติทดสอบ Bi-aspect ที่ใช้ทดสอบความแตกต่างของค่ากลางสำหรับประชากร 2 กลุ่มให้สามารถใช้ทดสอบได้กับประชากรหลายกลุ่ม โดยมีข้อสมมติและหลักการเดียวกัน วิธีการทดสอบตัวสถิติ Bi-aspect test สำหรับประชากรหลายกลุ่มมีวิธีการดังต่อไปนี้

กำหนดให้ $\underline{X}_1=\left(X_{11},\ldots,X_{1n_1}\right),\underline{X}_2=\left(X_{21},\ldots,X_{2n_2}\right),\ldots,\underline{X}_C=\left(X_{1C},\ldots,X_{Cn_c}\right)$ เป็นตัวอย่างสุ่มที่สุ่มมาอย่างอิสระจากประชากรที่มีการแจกแจงแบบต่อเนื่อง มีการแจกแจงเดียวกันและมีความแปรปรวนเท่ากัน นั่นคืออาจจะมีความแตกต่างระหว่างค่ากลางเพียงอย่างเดียว จำนวน C กลุ่มโดยที่ $C\geq 2$ ต้องการทดสอบว่าประชากร C กลุ่มมีค่ากลางเท่ากันหรือไม่

สมมติฐานทดสอบ

$H_0:\ \delta_1=\delta_2=\ldots=\delta_C$ แย้งกับ $H_1:\ \delta_h\neq\delta_j$ สำหรับบางคู่ $(h\neq j)$

โดยที่ $\delta_1,\delta_2,\ldots,\delta_C$ เป็นค่ากลางของตัวอย่างกลุ่มที่ j $(j=1,2,\ldots,C)$

ซึ่งในการดำเนินการทดสอบตัวสถิติมีขั้นตอน 2 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 ทำการทดสอบตัวสถิติแต่ละด้าน 2 ด้านคือ ด้านเชิงตัวเลข กับ ด้านเชิงกลุ่ม

1.1 ด้านเชิงตัวเลข (numerical aspect) จะเป็นผลรวมกำลังสองของค่าเฉลี่ยคูณด้วยจำนวนค่าสังเกตทุกกลุ่ม นิยามดังนี้

$$PT = T_A = \sum_{j=1}^{C} n_j \bar{X}_j^2$$

โดย $\bar{X}_j = \sum_{i=1}^{n_j} X_{ji} / n_j$ และ n_j แทนขนาดตัวอย่างกลุ่มที่ j

เทียบได้กับ $SSB(X) = \sum_{j=1}^{C} n_j \left(\bar{X}_j - \bar{X}\right)^2$ เมื่อ $\bar{X} = \sum_{j=1}^{C} n_j \bar{X}_j / n$ แต่เนื่องจาก $\bar{X}$ เป็นค่าคงที่สำหรับทุกกลุ่ม จึงพิจารณาเฉพาะส่วนที่ทำให้แต่ละกลุ่มมีความแตกต่างคือ $\sum_{j=1}^{C} n_j \bar{X}_j^2$ เป็นตัวสถิติ T_A

สมมติฐานการทดสอบ

$$H_{0A} : \delta_1 = \delta_2 = \ldots = \delta_C \quad \text{แย้งกับ} \quad H_1 : \ \delta_h \neq \delta_j \ \text{ สำหรับบางคู่ } (h \neq j)$$

1.2 ตัวสถิติทดสอบตัวที่ 2 เป็นตัวสถิติเชิงกลุ่มเพื่อดูว่าข้อมูลทั้ง C กลุ่มมีค่ามัธยฐานเท่ากันหรือไม่ โดยแบ่งกลุ่มค่าสังเกตออกเป็น 2 พวกดังนี้

1. กลุ่มใหญ่ คือ กลุ่มที่ค่าสังเกตมีค่ามากกว่าค่ามัธยฐานร่วม
2. กลุ่มเล็ก คือ กลุ่มที่ค่าสังเกตมีค่าน้อยกว่าหรือเท่ากับค่ามัธยฐานร่วม

ดังแสดงในตารางที่ 4

ตารางที่ 4 การแบ่งข้อมูลแต่ละกลุ่มออกเป็น 2 พวก คือพวกที่มีค่าสูงกว่ามัธยฐานและพวกที่มีค่าต่ำกว่าหรือเท่ากับมัธยฐาน

จำนวนข้อมูลที่มีค่า	กลุ่มตัวอย่างที่				รวม
	กลุ่มที่ 1	กลุ่มที่ 2	...	กลุ่มที่ C	
ต่ำกว่าหรือเท่ากับมัธยฐาน	n_{1s}	n_{2s}	...	n_{Cs}	n_s
สูงกว่ามัธยฐาน	n_{1g}	n_{2g}	...	n_{Cg}	n_g
รวม	n_1	n_2	...	n_C	n

ตัวสถิติทดสอบเชิงกลุ่มนิยามได้ดังนี้

$$T_B = \sum_{j=1}^{C} \left[\left(\sum_{i=1}^{n_i} I\left(X_{ji} > \tilde{M}\right) \right)^2 \bigg/ n_j \right]$$

$$\text{เมื่อ} \quad I\left(X_{ji} > \tilde{M}\right) = 1 \quad \text{เมื่อ } X_{ji} > \tilde{M}$$

$$= 0 \quad \text{ที่อื่น ๆ}$$

หรืออาจกล่าวได้ว่า

$$T_B = \sum_{j=1}^{C} n_{jg}^2 / n_j \qquad \text{โดยที่} \qquad n_{jg} = \sum_{i=1}^{n_j} I\left(X_{ji} > \widetilde{M}\right)$$

ภายใต้สมมติฐาน H_0 จริง

$$E(Z_1) = E(Z_2) = \ldots = E(Z_C)$$

โดยที่
$$Z_{ji} = \begin{cases} 0 \;, & X_{ji} \le \widetilde{M} \\ 1 \;, & X_{ji} > \widetilde{M} \end{cases}$$

และ $SSB(Z) = \sum_{j=1}^{C} n_j \left(\overline{Z}_j - \overline{Z}_{..}\right)^2$ เมื่อ $\overline{Z}_j = n_{jg} / n_j$ และ $\overline{Z}_{..} = n_g / n$

จะได้ว่า $SSB(Z) = \sum_{j=1}^{C} n_{jg}^2 / n_j - n_g^2 / n$ เนื่องจาก n_g^2 / n เป็นค่าคงที่สำหรับทุกกลุ่มจึงพิจารณาเฉพาะส่วนที่ทำให้แต่ละกลุ่มมีความแตกต่างคือ $\sum_{j=1}^{C} n_{jg}^2 / n_j$ เป็นตัวสถิติ T_B

สมมติฐานการทดสอบ

H_{0B} : $Pr\left(X_1 > \widetilde{M}\right) = Pr\left(X_2 > \widetilde{M}\right) = \ldots = Pr\left(X_C > \widetilde{M}\right)$ แย้งกับ

H_{1B} : $Pr\left(X_h > \widetilde{M}\right) \ne Pr\left(X_j > \widetilde{M}\right)$ สำหรับบางคู่ $(h \ne j)$

วิธีการคำนวณค่า p-value ของตัวสถิติทดสอบ คือ

จากตัวอย่างที่สุ่มมา $\underline{X} = \left(X_1, X_2, \ldots, X_C\right)$ หาค่าตัวสถิติทดสอบเชิงตัวเลข ${}_0PT$ จาก ${}_0PT = {}_0T_A = \sum_{j=1}^{C} n_j \overline{X}_j^2$ และตัวสถิติทดสอบเชิงกลุ่ม ${}_0T_B$ จาก ${}_0T_B = \sum_{j=1}^{C} n_{jg}^2 / n_j$ ทำการ permutation กลุ่มตัวอย่างโดยรวมตัวอย่างทั้ง C กลุ่มไว้ด้วยกันจัดกลุ่มใหม่คือ $\underline{X}^* = \left(X_1^*, X_2^*, \ldots, X_C^*\right)$ เป็น permutation ครั้งที่ 1 หาค่าตัวสถิติทดสอบเชิงตัวเลขได้เป็น ${}_1PT = {}_1T_A = \sum_{j=1}^{C} n_j \overline{X}_j^2$ และตัวสถิติทดสอบเชิงกลุ่มได้เป็น ${}_1T_B = \sum_{j=1}^{C} n_{jg}^2 / n_j$ ทำการ permutation จนกระทั่งครบจำนวน K ครั้ง แล้วคำนวณหาค่าสัดส่วนของจำนวน permutation ที่มีค่าตัวสถิติมากกว่าหรือเท่ากับตัวสถิติของค่าสังเกต ${}_0PT$ และ ${}_0T_B$ โดยการคำนวณ $\hat{L}_{PT}$ และ $\hat{L}_{T_B}$ ตามวิธีการประมาณของ pesarin (2001) ดังนี้

$$\hat{L}_{PT}\left({}_0PT\right) = \frac{1}{K} \sum_{k=1}^{K} I\left({}_K PT \ge {}_0PT\right)$$

โดย ${}_kPT$ แทน ค่า PT ของ permutation ครั้งที่ k $\left(k = 1, \ldots, K\right)$ ของ $\underline{X}$

$I(.)$ แทน ฟังก์ชัน $I\left({}_k PT \ge {}_0PT\right) = 1 \quad ; \quad {}_kPT \ge {}_0PT$

$= 0 \quad ;$ ที่อื่น ๆ

$$\hat{L}_{T_B}\left({}_0T_B\right)=\frac{1}{K}\sum_{k=1}^{K}I\left({}_kT_B \geq {}_0T_B\right)$$

โดย ${}_kT_B$ แทน ค่า T_B ของ permutation ครั้งที่ k $(k=1,\ldots,K)$ ของ $\underline{X}$

$I(.)$ แทน ฟังก์ชัน $I\left({}_kT_B \geq {}_0T_B\right) = 1$; ${}_kT_B \geq {}_0T_B$

$= 0$; ที่อื่น ๆ

และทำการปรับค่าดังนี้

$$\hat{L}_{T_A}\left({}_0T_A\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_A \geq {}_0T_A\right)\right]\Big/(K+1)$$

เมื่อ $I\left({}_kT_A > {}_0T_A\right) = 1$ เมื่อ ${}_kT_A \geq {}_0T_A$

$= 0$ ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0<\alpha<1$ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อ $\hat{L}_{PT} \leq \alpha$

$$\hat{L}_{T_B}\left({}_0T_B\right)=\left[\frac{1}{2}+\sum_{k=1}^{K}I\left({}_kT_B \geq {}_0T_B\right)\right]\Big/(K+1)$$

เมื่อ $I\left({}_kT_B > {}_0T_B\right) = 1$ เมื่อ ${}_kT_B \geq {}_0T_B$

$= 0$ ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0<\alpha<1$ จะปฏิเสธสมมติฐานว่าง H_0 เมื่อ $\hat{L}_{T_B} \leq \alpha$

ขั้นตอนที่ 2 เป็นการสร้างตัวสถิติทดสอบที่ใช้ตัวสถิติ 2 ตัวร่วมกัน เรียกว่า Bi-aspect test (T_{AB}) มีสมมติฐานการทดสอบ คือ

$H_0:\ \delta_1=\delta_2=\ldots=\delta_C$ แย้งกับ $H_1:\ \delta_h \neq \delta_j$ สำหรับบางคู่ $(h \neq j)$

คำนวณตัวสถิติได้จากฟังก์ชันที่รวม L_{T_A} และ L_{T_B} เข้าไว้ด้วยกัน ตัวสถิติทดสอบ Bi-aspect คือตัวสถิติที่มีค่า p-value น้อยกว่า ดังนี้

$$T_{AB}=max\left(1-L_{T_A},1-L_{T_B}\right)$$

ซึ่งค่าประมาณของตัวสถิติ T_{AB} ที่ได้จากค่าสังเกตคือ

$${}_0\hat{T}_{AB}=max\left(1-\hat{L}_{T_A}\left({}_0T_A\right),1-\hat{L}_{T_B}\left({}_0T_B\right)\right)$$

และจากการทำการ permutation ครั้งที่ 1 ตัวสถิติ T_{AB} ใน permutation แรก จะได้

$${}_1T_{AB}=\max\left(1-L_{T_A}\left({}_1T_A\right),1-L_{T_B}\left({}_1T_B\right)\right)$$

แล้วทำการคำนวณ permutation อีกจนครบจำนวน K ครั้ง จะได้เวคเตอร์ของตัวสถิติทดสอบ ดังนี้ $T_{AB}\left({}_1T_{AB},{}_2T_{AB},\ldots,{}_KT_{AB}\right)$

หาค่า p-value ของตัวสถิติได้จาก

$$\hat{L}_{T_{AB}} = \frac{1}{K}\sum_{k=1}^{K} I\left({}_kT_{AB} \geq {}_0\hat{T}_{AB}\right)$$

โดย $\quad I\left({}_kT_{AB} > {}_0\hat{T}_{AB}\right) \quad = \quad 1 \quad$ เมื่อ $\quad {}_kT_{AB} \geq {}_0\hat{T}_{AB}$

$\qquad\qquad\qquad\qquad\qquad = \quad 0 \quad$ ที่อื่น ๆ

ที่ระดับนัยสำคัญ $0 < \alpha < 1$ จะปฏิเสธสมมติฐาน H_0 เมื่อ $\hat{L}_{T_{AB}} \leq \alpha$

สรุปขั้นตอนการคำนวณตัวสถิติทดสอบ Bi-aspect ในกรณีที่มีตัวอย่างสุ่มจากประชากรมากกว่าหรือเท่ากับ 2 กลุ่มมีวิธีดำเนินการดังนี้

1. นำตัวอย่างทุกกลุ่มมารวมกันหาค่ามัธยฐานร่วม ($\tilde{M}$)
2. จากค่าสังเกตของตัวอย่างที่สุ่มมาอย่างอิสระคำนวณหาค่าตัวสถิติทดสอบ ดังนี้

ตัวสถิติทดสอบเชิงตัวเลข $\quad PT = {}_0T_A = \sum_{j=1}^{C} n_j \bar{X}_j^2$

ตัวสถิติทดสอบเชิงกลุ่ม $\quad {}_0T_B = \sum_{j=1}^{C} n_{jg}^2 / n_j$

3. ทำการ permutation ตัวอย่างสุ่มที่เป็นไปได้ทั้งหมด K ครั้ง ในการทำ permutation ครั้งที่ 1 แทนด้วย $\underline{X}^*$ ซึ่ง $\underline{X}^* = \left(X_1^*, X_2^*, \ldots, X_C^*\right)$

4. คำนวณหาค่าตัวสถิติทดสอบของ permutation ครั้งที่ 1 ดังนี้

ตัวสถิติทดสอบเชิงตัวเลข $\quad {}_1T_A = \sum_{j=1}^{C} n_j \left(\bar{X}_j^*\right)^2$

ตัวสถิติทดสอบเชิงกลุ่ม $\quad {}_1T_B = \sum_{j=1}^{C} \left(n_{jg}^*\right)^2 / n_j$

5. ทำการคำนวณตัวสถิติทดสอบทั้ง 2 ตัวสำหรับอีก $K-1$ permutation โดยทำเช่นเดียวกับขั้นตอนที่ 3 และ 4
6. คำนวณหาจำนวน permutation ของตัวสถิติ T_A ที่มีค่ามากกว่าหรือเท่ากับ ${}_0T_A$ และหาค่า p-value ของตัวสถิติ $\hat{L}_{T_A}\left({}_0T_A\right)$ ดังนี้

$$\hat{L}_{T_A}\left({}_0T_A\right) = \left[\frac{1}{2} + \sum_{k=1}^{K} I\left({}_kT_A \geq {}_0T_A\right)\right] \Big/ (K+1)$$

ในทำนองเดียวกันสำหรับตัวสถิติ T_B ค่า p-value ของตัวสถิติ $\hat{L}_{T_B}\left({}_0T_B\right)$ หาได้ดังนี้

$$\hat{L}_{T_B}\left({}_0T_B\right) = \left[\frac{1}{2} + \sum_{k=1}^{K} I\left({}_kT_B \geq {}_0T_B\right)\right] \Big/ (K+1)$$

7. หาค่าประมาณของตัวสถิติทดสอบ Bi-aspect เลือกตัวสถิติ ${}_1T_A$ กับ ${}_1T_B$ จากฟังก์ชัน

$${}_0\hat{T}_{AB} = max\left(1 - \hat{L}_{T_A}\left({}_0T_A\right), 1 - \hat{L}_{T_B}\left({}_0T_B\right)\right)$$

8. คำนวณหาจำนวน permutation ของตัวสถิติ T_A ที่มีค่ามากกว่าหรือเท่ากับ ${}_1T_A$ และหาค่า p-value ของตัวสถิติ $\hat{L}_{T_A}\left({}_1T_A\right)$ ดังนี้

$$\hat{L}_{T_A}\left({}_1T_A\right)=\left[\frac{1}{2}+\sum_{k=1}^{K} I\left({}_kT_A \geq {}_1T_A\right)\right]\Big/(K+1)$$

ในทำนองเดียวกันสำหรับตัวสถิติ T_B ค่า p-value ของตัวสถิติ $\hat{L}_{T_B}\left({}_1T_B\right)$ หาได้ดังนี้

$$\hat{L}_{T_B}\left({}_1T_B\right)=\left[\frac{1}{2}+\sum_{k=1}^{K} I\left({}_kT_B \geq {}_1T_B\right)\right]\Big/(K+1)$$

9. หาค่าประมาณของตัวสถิติทดสอบ Bi-aspect ของ permutation แรกจากฟังก์ชัน

$${}_1T_{AB}=\max\left(1-L_{T_A}\left({}_1T_A\right),1-L_{T_B}\left({}_1T_B\right)\right)$$

10. อีก K-1 permutation ทำเช่นเดียวกับขั้นตอนที่ 8-9

11. คำนวณสัดส่วนของจำนวน permutation ที่ตัวสถิติ T_{AB} มีค่ามากกว่าหรือเท่ากับตัวสถิติ ${}_0T_{AB}$ และหาค่า p-value ของตัวสถิติ Bi-aspect ดังนี้

$$\hat{L}_{T_{AB}}=\frac{1}{K}\sum_{k=1}^{K} I\left({}_kT_{AB} \geq {}_0\hat{T}_{AB}\right)$$

12. ปฏิเสธสมมติฐาน H_0 ที่ระดับ $0<\alpha<1$ ถ้า $\hat{L}_{T_{AB}} \leq \alpha$

2. ทฤษฎี Permutation Test

Permutation test เริ่มมีมานานแล้วแต่ได้รับการยอมรับจากนักสถิติเมื่อราว ค.ศ. 1930 ซึ่งเสนอโดย Fisher และในปี ค.ศ. 1937 Pitman ได้นำวิธีการ Permutation มาใช้ในการทดสอบความแตกต่างของค่า mean ของประชากร 2 กลุ่ม ว่ามีค่าเท่ากันหรือไม่หรือเพื่อทดสอบว่าประชากรมาจากประชากรที่มีการแจกแจงเดียวกันหรือไม่นั่นเอง โดยที่ไม่จำเป็นต้องมีข้อตกลงเบื้องต้นเกี่ยวกับประชากร ซึ่งในบางครั้งเรียกวิธีการนี้ว่า randomization test

สำหรับหลักการของ permutation test นั้น Berry (1982) เสนอดังนี้ สมมติตัวอย่างที่สุ่มมาจากประชากร K กลุ่ม แต่ละกลุ่มมีขนาด R_i เมื่อ $(i=1,2,\ldots,K)$ โดยที่ $R_1+R_2+\ldots+R_K=N$ ไม่มีการจัดลำดับกลุ่มและไม่มีกลุ่มใดที่เป็นเซตว่างเริ่มดำเนินการ permutation โดยการนำตัวอย่างทั้งหมดมารวมกันทำการสุ่มแบบไม่ใส่คืนจากตัวอย่างขนาด N สุ่มครั้งแรก R_1 ตัว เหลือตัวอย่างขนาด $N-R_1$ สุ่มครั้งที่ 2 R_2 ตัว เหลือตัวอย่างขนาด $N-R_1-R_2$ สุ่มครั้งที่ 3 R_3 ตัว สุ่มจนกระทั่งครั้งที่ $k-1$ มีขนาดตัวอย่าง $N-R_1-R_2-\ldots-R_{k-2}$ สุ่มมา R_{k-1} ตัว จะได้จำนวน permutation ทั้งหมดเท่ากับ $\binom{N}{R_1 R_2 \ldots R_K}$ วิธี นั่นคือ

$$\binom{N}{R_1 R_2 \ldots R_K} = \binom{N}{R_1}\binom{N-R_1}{R_2}\cdots\binom{N-R_1-R_2-\ldots-R_{K-2}}{R_{K-1}}$$

จากวิธีการ permutation นำตัวอย่างกลุ่มใหม่ที่ได้มาหาค่าตัวสถิติที่ต้องการ แต่ในบางครั้งการนำทุกกลุ่มที่ได้จากการ permutation มาศึกษาอาจไม่สะดวกมากนักเนื่องจากมีจำนวนมากจึงอาจจะต้องเลือกมาในจำนวนที่จำกัด

3. งานวิจัยที่เกี่ยวข้อง

Fligner and Policella (1981) ศึกษาเปรียบเทียบตัวสถิติที่ใช้ทดสอบค่ากลางของประชากร 2 กลุ่ม โดยพิจารณาประสิทธิภาพของตัวสถิติจากกำลังการทดสอบของตัวสถิติทดสอบ t ตัวสถิติทดสอบ Welch t ตัวสถิติทดสอบ Mann-Whitney-Wilcoxon (U-test) และ ตัวสถิติทดสอบ modified U และพิจารณาจากการแจกแจง 5 การแจกแจงคือ การแจกแจงแบบยูนิฟอร์ม การแจกแจงแบบปกติ การแจกแจงแบบปกติที่มีการปลอมปน การแจกแจงแบบดับเบิลเอกซ์โปเนนเชียล และ การแจกแจงแบบโคชี โดยทำการศึกษาใน 2 กรณี คือกรณีที่มีค่าสเกลแตกต่างกัน และ กรณีที่มีค่ากลางแตกต่างกัน ซึ่งจากการศึกษาพบว่าในกรณีที่ค่าสเกลแตกต่างกันตัวสถิติทดสอบ U จะให้กำลังการทดสอบสูงสุดในทุกการแจกแจง ส่วนในกรณีที่มีค่ากลางแตกต่างกัน การแจกแจงแบบปกติและการแจกแจงแบบยูนิฟอร์ม ตัวสถิติทดสอบ t จะมีกำลังการทดสอบมากกว่าตัวสถิติทดสอบอื่นรองลงมาเป็นตัวสถิติทดสอบ Welch t แต่เมื่อข้อมูลมีการแจกแจงเป็นแบบปกติที่มีการปลอมปน การแจกแจงแบบดับเบิลเอกซ์โปเนนเชียล และ การแจกแจงแบบโคชี พบว่าตัวสถิติทดสอบ U มีกำลังการทดสอบมากกว่าตัวสถิติทดสอบอื่นรองลงมาเป็นตัวสถิติทดสอบ modified U

John Ludbrook and Hugh Dudley (1998) ทำการศึกษาประสิทธิภาพของตัวสถิติทดสอบ t และตัวสถิติทดสอบ F เปรียบเทียบกับตัวสถิติทดสอบภายใต้ permutation test เนื่องจากเป็นที่ถกเถียงกันระหว่างนักสถิติกับนักวิทยาศาสตร์ทางการแพทย์ คือนักสถิติเชื่อว่างานวิจัยทางด้านการแพทย์จำเป็นต้องทำการสุ่มตัวอย่างและดำเนินการวิเคราะห์ด้วยวิธีการทางสถิติ ซึ่งในความเป็นจริงตัวอย่างมักไม่ได้จากการสุ่มตามวิธีการทางสถิติ แต่จากการศึกษางานวิจัยพบว่าการใช้ตัวอย่างโดยไม่สุ่มได้ผลดีกว่าตัวอย่างที่สุ่มด้วยวิธีการทางสถิติและดำเนินการทดสอบภายใต้วิธีการ permutation test ซึ่งเขาได้ศึกษาการทดลองวัดระดับคลอเรสเตอรอลในเลือดของผู้ชาย 12 คนแบ่งออกเป็นผู้ชายกินปลา (ไม่กินเนื้อ) 7 คน และผู้ชายกินเนื้อ (ไม่กินปลา) 5 คน ว่ามีความแตกต่างกันหรือไม่ ทำการทดสอบด้วยวิธีการ permutation test พบว่ามีความแตกต่างระหว่างผู้ชาย 2 กลุ่มและเมื่อทำการเปรียบเทียบกับตัวสถิติอื่น ๆ คือตัวสถิติทดสอบ t ตัวสถิติ

ทดสอบ Welch ตัวสถิติทดสอบ Exact WMW และตัวสถิติทดสอบ Exact permutation พบว่าที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบ Exact permutation ปฏิเสธสมมติฐานที่ระดับ p-value น้อยสุด รองลงมาเป็นตัวสถิติทดสอบ Exact WMW และ ตัวสถิติทดสอบ t ตามลำดับ ส่วนตัวสถิติทดสอบ Welch ไม่ปฏิเสธสมมติฐาน และเมื่อทำการ transform ข้อมูลโดยการ take log ได้ผลเช่นเดียวกันเพียงแต่มีค่า p-value เปลี่ยนไปเล็กน้อยยกเว้นตัวสถิติทดสอบ Exact WMW ที่มีค่า p-value ไม่เปลี่ยนแปลง

Macro Marozzi (2003) เสนอตัวสถิติทดสอบ Bi-aspect สำหรับทดสอบความแตกต่างค่ากลางของประชากร 2 กลุ่มภายใต้ permutation test ด้วยตัวสถิติทดสอบดังกล่าวนี้ใช้ทดสอบ 2 ด้านคือด้านเชิงตัวเลขและด้านเชิงกลุ่ม โดยทำการศึกษาตัวอย่างที่สุ่มอย่างอิสระจากประชากรที่มีการแจกแจง 10 การแจกแจงคือ การแจกแจงแบบปกติมาตรฐาน การแจกแจงแบบยูนิฟอร์ม การแจกแจงแบบดับเบิ้ลเอกซ์โปเนนเชียล การแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบไคสแควร์ การแจกแจงแบบเอกซ์โปเนนเชียล การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว การแจกแจงแบบไบโมดอล (bimodol) ที่มีหางบาง การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 10% การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 30% ด้วยขนาดตัวอย่างเท่ากัน และไม่เท่ากัน ดังนี้

ขนาดตัวอย่างเท่ากัน (10,10) , (20,20) และ (40,40)

ขนาดตัวอย่างไม่เท่ากัน (10,20) , (10,40) , (20,10) และ (40,10)

โดยทำการจำลองข้อมูล 4,000 รอบ ทำการ permutation จำนวน 1,000 ครั้ง ได้ผลลัพธ์ดังนี้

1. ขนาดความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของตัวอย่างเปรียบเทียบระหว่างตัวสถิติทดสอบ $T_a\ (PT)$ กับ ตัวสถิติทดสอบ Bi-aspect (T_{ab}) พบว่าสำหรับทุกการแจกแจงตัวสถิติทดสอบ T_a และ T_{ab} มีความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เข้าใกล้ระดับนัยสำคัญ 0.05 แต่สำหรับการแจกแจงแบบโคชีมาตรฐาน และการแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว มีความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ห่างจากระดับนัยสำคัญที่กำหนด

2. กำลังการทดสอบของตัวอย่างโดยเปรียบเทียบจากตัวสถิติทดสอบ 4 ตัว คือ t , PT , T_b และ T_{ab} พบว่า ภายใต้การแจกแจงแบบปกติและการแจกแจงแบบไคสแควร์ตัวสถิติทดสอบ T_{ab} และ PT จะให้ผลดีเท่า ๆ กันและให้ผลใกล้เคียงกับตัวสถิติทดสอบ t ภายใต้การแจกแจงแบบยูนิฟอร์มและไบโมดอลตัวสถิติทดสอบ PT จะให้ผลดีกว่าตัวสถิติทดสอบ T_{ab} เล็กน้อย ภายใต้การแจกแจงแบบเอกซ์โปเนนเชียลและดับเบิ้ลเอกซ์โปเนนเชียลตัวสถิติทดสอบ T_{ab} จะให้ผลดีกว่าตัวสถิติทดสอบ PT ทั้งหมด และสิ่งที่น่าสนใจมาก

สำหรับการแจกแจงอีก 4 การแจกแจงที่มีลักษณะหางหนา ได้แก่การแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 10% และการแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 30% นั่นคือตัวสถิติทดสอบ T_{ab} จะให้ผลดีกว่าตัวสถิติทดสอบ PT อย่างมาก

3. ตัวสถิติทดสอบ T_a จะทดสอบได้ไม่ดีเมื่อค่ากลางของตัวอย่าง 2 กลุ่มมีความแตกต่างกันมาก ในขณะที่ตัวสถิติทดสอบ T_b จะทดสอบได้ดีกว่า แต่ตัวสถิติทดสอบ T_b จะมีประสิทธิภาพน้อยมากเมื่อข้อมูลมีการแจกแจงแบบปกติ การแจกแจงแบบยูนิฟอร์ม การแจกแจงแบบไคสแควร์ การแจกแจงแบบเอกซ์โปเนนเชียล และการแจกแจงแบบไบโมดอล ดังนั้นจึงนิยมใช้ตัวสถิติทดสอบ T_{ab} มากกว่าเนื่องจากตัวสถิติทดสอบ T_{ab} จะเป็นผลรวมของตัวสถิติ PT และ T_b ความแตกต่างของค่ากลางและการแจกแจงที่มีลักษณะเบ้จะไม่มีผลกับตัวสถิติทดสอบ T_{ab} มากนัก

4. กำลังการทดสอบสัมพัทธ์ของตัวสถิติทดสอบ T_{ab} เทียบกับตัวสถิติทดสอบ PT พบว่าจะได้ผลดีมากเมื่อมีการแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 10% และการแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 30% ส่วนการแจกแจงแบบที่เหลือจะให้ผลใกล้เคียงกัน

Macro Marozzi (2004) ได้ทำการเสนอตัวสถิติทดสอบ Bi-aspect สำหรับทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม ในทำนองเดียวกับการศึกษาในกรณีของตัวอย่าง 2 กลุ่มโดยศึกษาในกรณีตัวอย่าง 3 กลุ่ม 4 กลุ่ม และ 5 กลุ่ม ด้วยขนาดต่าง ๆ กัน และทำการศึกษาในกรณีค่ากลางแตกต่าง 2 ประเภทคือ

1. กรณีที่มี 1 กลุ่มเท่านั้นที่มีค่ากลางแตกต่างจากกลุ่มอื่น
2. กรณีที่ทุกกลุ่มมีค่ากลางแตกต่างกันหมด

ซึ่งค่ากลางที่แตกต่างกำหนดจากกำลังการทดสอบที่ระดับ 50% ของตัวสถิติ PT เมื่อ $\alpha = .05$ ผลจากการศึกษาพบว่า

1. ขนาดของความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 มีค่าใกล้เคียงกับระดับนัยสำคัญ 0.05 ในทุกการแจกแจงยกเว้นการแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว

2. กำลังการทดสอบในกรณีของตัวอย่าง 3 กลุ่มให้ผลเช่นเดียวกับการศึกษาในกรณีตัวอย่าง 2 กลุ่มคือภายใต้การแจกแจงแบบปกติ การแจกแจงแบบยูนิฟอร์ม และการแจกแจงแบบไบโมดอล ตัวสถิติทดสอบ T_{AB} จะให้ผลดีกว่าตัวสถิติทดสอบ PT การแจกแจงแบบ

ไคสแควร์ ตัวสถิติทดสอบ T_{AB} จะให้ผลดีกว่าตัวสถิติทดสอบ PT เล็กน้อย การแจกแจงแบบเอกซ์โปเนนเชียลและดับเบิ้ลเอกซ์โปเนนเชียล ตัวสถิติทดสอบ T_{AB} จะให้ผลดีกว่าตัวสถิติทดสอบ PT มาก แต่สิ่งที่น่าสนใจมากสำหรับการแจกแจงอีก 4 การแจกแจงที่มีลักษณะหางหนา ได้แก่การแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 10% และการแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 30% นั่นคือตัวสถิติทดสอบ T_{AB} จะให้ผลดีกว่าตัวสถิติทดสอบ PT อย่างมาก

3. กำลังการทดสอบในกรณีของตัวอย่าง 4 กลุ่มและ 5 กลุ่มพิจารณาจากกำลังการทดสอบสัมพัทธ์ของตัวสถิติทดสอบ T_{AB} เทียบกับตัวสถิติทดสอบ PT ได้ผลเช่นเดียวกับกรณีการศึกษาในตัวอย่าง 2 กลุ่มคือจะได้ผลดีมากเมื่อมีการแจกแจงแบบโคชีมาตรฐาน การแจกแจงแบบโคชีมาตรฐานแบบครึ่งเดียว การแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 10% และการแจกแจงแบบปกติที่มีค่าผิดปกติปนอยู่ 30% ส่วนการแจกแจงแบบที่เหลือจะให้ผลใกล้เคียงกัน

บทที่ 3

วิธีการดำเนินการวิจัย

ในการศึกษาเพื่อเปรียบเทียบประสิทธิภาพ ของตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม ได้แก่ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis Test จะใช้การจำลองแบบข้อมูลโดยวิธีมอนติคาร์โล (Monte Carlo Method) ด้วยโปรแกรม Fortran power station 4.0 โดยกำหนดให้มีตัวอย่างสุ่มอย่างอิสระจำนวน 3 กลุ่ม และ 5 กลุ่มจากประชากรซึ่งมีการแจกแจงในลักษณะต่าง ๆ จำนวน 5 แบบ คือการแจกแจงแบบปกติ การแจกแจงแบบยูนิฟอร์ม การแจกแจงแบบไคสแควร์ การแจกแจงแบบเอกซ์โปเนนเชียล และการแจกแจงแบบปลอมปน โดยในการศึกษาเปรียบเทียบจะพิจารณาด้านขนาดของกลุ่มตัวอย่าง ประกอบด้วยตัวอย่างขนาด 3 กลุ่มและ 5 กลุ่ม ซึ่งสามารถแบ่งการศึกษาออกเป็น 2 กรณี ดังต่อไปนี้

กรณีที่ 1 กำหนดให้ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน

ในกรณีดังกล่าวนี้ต้องการเปรียบเทียบประสิทธิภาพ ในด้านของความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของตัวสถิติทดสอบต่างๆ ซึ่งมีขั้นตอนดำเนินการดังต่อไปนี้

เมื่อตัวอย่างถูกสุ่มจากการแจกแจงแบบปกติ

1. ทำการสุ่มตัวอย่างจำนวน 3 กลุ่มจากการแจกแจงแบบปกติซึ่งมีค่าเฉลี่ยเท่ากับ 10 และความแปรปรวน 1 ใช้สัญลักษณ์แทนด้วย $N(10,1)$

2. กำหนดค่าระดับความมีนัยสำคัญของการทดสอบ $\alpha = .05$

3. คำนวณค่า p-value ของตัวสถิติทดสอบทั้ง 3 ตัว คือ ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis โดยตัวสถิติทดสอบ Bi-aspect กำหนดให้จำนวน permutation ที่ใช้เท่ากับ 1,000 ครั้ง

4. ทำการเปรียบเทียบค่า p-value ของตัวสถิติทดสอบที่คำนวณได้กับระดับนัยสำคัญที่ตั้งไว้ ถ้าค่า p-value ที่คำนวณได้ มีค่าน้อยกว่าระดับนัยสำคัญที่กำหนดจะปฏิเสธสมมติฐานว่าง

5. ดำเนินการ ข้อ 1-4 จำนวน 4,000 รอบ หาค่าประมาณความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 โดยนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นจริงหารด้วยจำนวนครั้งทั้งหมดของการจำลองแบบคือ 4000 ครั้ง

6. ดำเนินการตามข้อ 1-5 โดยเปลี่ยนระดับความมีนัยสำคัญของการทดสอบเป็น $\alpha = .10$ และเปลี่ยนขนาดตัวอย่างและจำนวนกลุ่มดังที่กำหนดในตารางที่ 5

7. ดำเนินการตามข้อ 1-6 สำหรับการแจกแจงของประชากรในลักษณะอื่น ๆ ได้แก่ การแจกแจงแบบยูนิฟอร์ม การแจกแจงแบบไคสแควร์ การแจกแจงแบบเอกซ์โปเนนเชียล และการแจกแจงแบบปลอมปน โดยกำหนดค่าต่าง ๆ ของแต่ละการแจกแจงดังนี้

- การแจกแจงแบบยูนิฟอร์ม เขียนแทนด้วย $U(a,b)$ กำหนดพารามิเตอร์ a และ b คือ $a = 0$ และ $b = 2\sqrt{3}$ โดยการแจกแจงที่กำหนดดังกล่าวนี้มีค่าเฉลี่ยและความแปรปรวนเป็น $\sqrt{3}$ และ 1 ตามลำดับ

- การแจกแจงแบบไคสแควร์ เขียนแทนด้วย $\chi^2(\upsilon)$ เมื่อ υ เป็นองศาอิสระ ในที่นี้กำหนดให้ $\upsilon = 4$ ซึ่งการแจกแจงดังกล่าวนี้มีค่าเฉลี่ยและความแปรปรวนเป็น 4 และ 8 ตามลำดับ

- การแจกแจงแบบเอกซ์โปเนนเชียล เขียนแทนด้วย $Exp(\beta)$ โดยกำหนดให้ $\beta = 1$ ดังนั้นการแจกแจงมีค่าเฉลี่ยและความแปรปรวนเป็น 1

- การแจกแจงแบบปลอมปน กำหนดให้ผสมการแจกแจงผสมระหว่างการแจกแจงแบบปกติที่มี $\mu_1 = 10, \sigma_1^2 = 1$ กับการแจกแจงแบบปกติที่มี $\mu_2 = 1, \sigma_2^2 = 100$ ด้วยสัดส่วน p ซึ่งในที่นี้กำหนดให้ p = 0.1 และ 0.3 เขียนแทนด้วย $CN(\mu_1, \sigma_1^2, p, \mu_2, \sigma_2^2)$

กรณีที่ 2 กำหนดให้ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกัน

ในกรณีดังกล่าวนี้ต้องการศึกษาว่าตัวสถิติทดสอบที่ศึกษาทั้ง 3 ตัว มีกำลังการทดสอบ (power of the test) แตกต่างกันอย่างไร การศึกษาในกรณีดังกล่าวนี้แบ่งการศึกษาออกเป็น 2 กรณีย่อย ตามการศึกษาของ Macro Marozzi (2004) คือ

กรณีที่ 2.1 มีค่าเฉลี่ยของประชากรเพียง 1 กลุ่มที่แตกต่างจากประชากรกลุ่มอื่น ในที่นี้จะกำหนดให้ค่าเฉลี่ยของประชากรกลุ่มที่แตกต่างจากประชากรกลุ่มอื่นมีค่าเฉลี่ยเพิ่มขึ้น 5% และ 20% โดยมีขั้นตอนของการดำเนินการดังนี้

เมื่อตัวอย่างจำนวน 3 กลุ่มถูกสุ่มจากการแจกแจงแบบปกติ

ก. ทำการสุ่มตัวอย่าง 2 กลุ่ม จากการแจกแจงแบบปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ 10 และความแปรปรวน 1 ส่วนตัวอย่างอีก 1 กลุ่ม สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติที่มีความแปรปรวนเท่ากับ 1 เช่นเดียวกัน แต่เพิ่มค่าเฉลี่ยของตัวอย่างอีก 5% และ 20%

ข. กำหนดค่าระดับความมีนัยสำคัญของการทดสอบ $\alpha = .05$

ค. คำนวณค่า p-value ของตัวสถิติทดสอบทั้ง 3 ตัว คือ ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis โดยตัวสถิติทดสอบ Bi-aspect กำหนดให้จำนวน permutation ที่ใช้เท่ากับ 1,000 ครั้ง

ง. ทำการเปรียบเทียบค่า p-value ของตัวสถิติทดสอบที่คำนวณได้กับระดับนัยสำคัญที่ตั้งไว้ ถ้าค่า p-value ที่คำนวณได้มีค่าน้อยกว่าระดับนัยสำคัญที่กำหนดจะปฏิเสธสมมติฐานว่าง

จ. ดำเนินการ ข้อ ก. - ง. จำนวน 4,000 รอบ หาค่าประมาณของกำลังการทดสอบโดยนับจำนวนครั้งของการปฏิเสธสมมติฐานว่าง เมื่อสมมติฐานว่างเป็นเท็จหารด้วยจำนวนครั้งทั้งหมดของการจำลองแบบคือ 4,000 ครั้ง

ฉ. ดำเนินการตามข้อ ก. - จ. โดยเปลี่ยนระดับความมีนัยสำคัญของการทดสอบเป็น $\alpha = .10$ และเปลี่ยนขนาดตัวอย่างและเพิ่มจำนวนกลุ่มดังที่กำหนดในตารางที่ 5

ช. ดำเนินการตามข้อ ก. - ฉ. สำหรับการแจกแจงของประชากรในลักษณะอื่น ๆ โดยกำหนดการแจกแจงของประชากรเริ่มต้น เช่นเดียวกับกรณีที่กำหนดให้ประชากรแต่ละกลุ่มมีค่ากลางเท่ากัน

กรณีที่ 2.2 ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน

ในกรณีนี้กำหนดให้ค่าเฉลี่ยของกลุ่มถัดไป เพิ่มขึ้นจากค่าเฉลี่ยของกลุ่มก่อนหน้า 5% และ 20% โดยในกรณีค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน 5% ดำเนินการโดยตัวอย่างกลุ่มแรกสุ่มมาจากประชากรที่มีค่าเฉลี่ยคือ μ ตัวอย่างกลุ่มที่ 2 จะเป็นตัวอย่างที่สุ่มมาจากประชากรที่มีค่าเฉลี่ยเพิ่มขึ้นจากประชากรกลุ่มแรกอีก 5% สำหรับตัวอย่างกลุ่มที่ 3 และกลุ่มต่อ ๆ ไปจะเป็นตัวอย่างที่ถูกสุ่มมาจากประชากรที่มีค่าเฉลี่ยเพิ่มขึ้น จากประชากรกลุ่มแรกครั้งละ 5% นั่นคือ ตัวอย่างกลุ่มที่ 3 จะเป็นตัวอย่างที่สุ่มจากประชากรที่มีค่าเฉลี่ยเพิ่มขึ้นจากตัวอย่างกลุ่มแรก 10% และตัวอย่างกลุ่มที่ 4 จะมาจากประชากรที่มีค่าเฉลี่ยเพิ่มขึ้น 15% จากค่าเฉลี่ยของประชากรกลุ่มแรก เช่นนี้ เรื่อย ๆ ไปสำหรับในกรณีค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน 20% มีลักษณะเช่นเดียวกับในกรณีค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน 5% โดยมีขั้นตอนในการดำเนินงานต่อไปลักษณะเดียวกับกรณี 2.1 ซึ่งสรุปเป็นตารางการดำเนินการสำหรับกรณีต่าง ๆ ได้ดังตารางที่ 5-7

ตารางที่ 5 ขนาดตัวอย่างและค่ากลางที่ใช้ในการทดลองสำหรับการแจกแจงลักษณะต่าง ๆ

การแจกแจง	กลุ่ม	ขนาดตัวอย่าง	ค่า location $(\mu_1, \mu_2, ..., \mu_k)$				
			เท่าทุกกลุ่ม	แตกต่างกัน 1 กลุ่ม		ทุกกลุ่มแตกต่าง	
				5%	20%	5%	20%
N(10,1) $\mu=10$ $\sigma^2=1$	3	(10,10,10)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		(20,20,20)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		(40,40,40)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
	5	(10,10,10,10,10)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		(20,20,20,20,20)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		(40,40,40,40,40)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
U(0,3.46) $\mu=1.73$ $\sigma^2=1$	3	(10,10,10)	(1.73,1.73,1.73)	(1.82,1.73,1.73)	(2.08,1.73,1.73)	(1.73,1.82,1.91)	(1.73,2.08,2.43)
		(20,20,20)	(1.73,1.73,1.73)	(1.82,1.73,1.73)	(2.08,1.73,1.73)	(1.73,1.82,1.91)	(1.73,2.08,2.43)
		(40,40,40)	(1.73,1.73,1.73)	(1.82,1.73,1.73)	(2.08,1.73,1.73)	(1.73,1.82,1.91)	(1.73,2.08,2.43)
	5	(10,10,10,10,10)	(1.73,1.73,1.73,1.73,1.73)	(1.82,1.73,1.73,1.73,1.73)	(2.08,1.73,1.73,1.73,1.73)	(1.73,1.82,1.91,2.0,2.09)	(1.73,2.08,2.43,2.78,3.13)
		(20,20,20,20,20)	(1.73,1.73,1.73,1.73,1.73)	(1.82,1.73,1.73,1.73,1.73)	(2.08,1.73,1.73,1.73,1.73)	(1.73,1.82,1.91,2.0,2.09)	(1.73,2.08,2.43,2.78,3.13)
		(40,40,40,40,40)	(1.73,1.73,1.73,1.73,1.73)	(1.82,1.73,1.73,1.73,1.73)	(2.08,1.73,1.73,1.73,1.73)	(1.73,1.82,1.91,2.0,2.09)	(1.73,2.08,2.43,2.78,3.13)
χ^2_4 $\mu=4$ $\sigma^2=8$	3	(10,10,10)	(4,4,4)	(4.2,4,4)	(4.8,4,4)	(4,4.2,4.4)	(4,4.8,5.6)
		(20,20,20)	(4,4,4)	(4.2,4,4)	(4.8,4,4)	(4,4.2,4.4)	(4,4.8,5.6)
		(40,40,40)	(4,4,4)	(4.2,4,4)	(4.8,4,4)	(4,4.2,4.4)	(4,4.8,5.6)
	5	(10,10,10,10,10)	(4,4,4,4,4)	(4.2,4,4,4,4)	(4.8,4,4,4,4)	(4,4.2,4.4,4.6,4.8)	(4,4.8,5.6,6.4,7.2)
		(20,20,20,20,20)	(4,4,4,4,4)	(4.2,4,4,4,4)	(4.8,4,4,4,4)	(4,4.2,4.4,4.6,4.8)	(4,4.8,5.6,6.4,7.2)
		(40,40,40,40,40)	(4,4,4,4,4)	(4.2,4,4,4,4)	(4.8,4,4,4,4)	(4,4.2,4.4,4.6,4.8)	(4,4.8,5.6,6.4,7.2)
Exp(1) $\mu=1$ $\sigma^2=1$	3	(10,10,10)	(1,1,1)	(1.05,1,1)	(1.2,1,1)	(1,1.05,1.1)	(1,1.20,1.4)
		(20,20,20)	(1,1,1)	(1.05,1,1)	(1.2,1,1)	(1,1.05,1.1)	(1,1.20,1.4)
		(40,40,40)	(1,1,1)	(1.05,1,1)	(1.2,1,1)	(1,1.05,1.1)	(1,1.20,1.4)
	5	(10,10,10,10,10)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.20,1.40,1.60,1.80)
		(20,20,20,20,20)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.20,1.40,1.60,1.80)
		(40,40,40,40,40)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.20,1.40,1.60,1.80)

ตารางที่ 6 จำนวนและค่าเฉลี่ยของตัวอย่างที่ถูกนำไปผสมในการแจกแจงแบบปลอมปนตามสัดส่วนต่าง ๆ พร้อมทั้งจำนวนและค่าเฉลี่ยของการแจกแจงหลัก

จำนวนกลุ่ม	ขนาดตัวอย่าง	จำนวนตัวอย่างการแจกแจง หลัก/ปลอมปน P=0.1		ค่า location $(\mu_1, \mu_2, ..., \mu_C)$				
				เท่ากันทุกกลุ่ม	แตกต่าง 1 กลุ่ม		ทุกกลุ่มแตกต่าง	
					5%	20%	5%	20%
3	10	หลัก	(9,9,9)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(1,1,1)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
	20	หลัก	(18,18,18)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(2,2,2)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
	40	หลัก	(36,36,36)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(4,4,4)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
5	10	หลัก	(9,9,9,9,9)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		ปลอมปน	(1,1,1,1,1)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)
	20	หลัก	(18,18,18,18,18)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		ปลอมปน	(2,2,2,2,2)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)
	40	หลัก	(36,36,36,36,36)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		ปลอมปน	(4,4,4,4,4)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)

ตารางที่ 7 จำนวนและค่าเฉลี่ยของตัวอย่างที่ถูกนำไปผสมในการแจกแจงแบบปลอมปนตามสัดส่วนต่าง ๆ พร้อมทั้งจำนวนและค่าเฉลี่ยของการแจกแจงหลัก

จำนวนกลุ่ม	ขนาดตัวอย่าง	จำนวนตัวอย่างการแจกแจง หลัก/ปลอมปน P=0.3		ค่า location $(\mu_1, \mu_2, ..., \mu_C)$				
				เท่ากันทุกกลุ่ม	แตกต่าง 1 กลุ่ม		ทุกกลุ่มแตกต่าง	
					5%	20%	5%	20%
3	10	หลัก	(7,7,7)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(3,3,3)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
	20	หลัก	(14,14,14)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(6,6,6)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
	40	หลัก	(28,28,28)	(10,10,10)	(10.5,10,10)	(12,10,10)	(10,10.5,11)	(10,12,14)
		ปลอมปน	(12,12,12)	(1,1,1)	(1.5,1,1)	(1.2,1,1)	(1,1.05,1.10)	(1,1.2,1.4)
5	10	หลัก	(7,7,7,7,7)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(7,7,7,7,7)
		ปลอมปน	(3,3,3,3,3)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)
	20	หลัก	(14,14,14,14,14)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		ปลอมปน	(6,6,6,6,6)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)
	40	หลัก	(28,28,28,28,28)	(10,10,10,10,10)	(10.5,10,10,10,10)	(12,10,10,10,10)	(10,10.5,11,11.5,12)	(10,12,14,16,18)
		ปลอมปน	(12,12,12,12,12)	(1,1,1,1,1)	(1.05,1,1,1,1)	(1.2,1,1,1,1)	(1,1.05,1.10,1.15,1.20)	(1,1.2,1.4,1.6,1.8)

บทที่ 4

ผลการวิจัย

ในการศึกษาประสิทธิภาพของตัวสถิติทดสอบ ความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม ได้แก่ตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal - Wallis ซึ่งกำหนดประชากรที่มีการแจกแจงในลักษณะต่าง ๆ 3 ลักษณะ คือ การแจกแจงแบบสมมาตรได้แก่ การแจกแจงแบบปกติ และ ยูนิฟอร์ม การแจกแจงแบบเบ้ ได้แก่การแจกแจงแบบไคสแควร์ และ เอกซ์โปเนนเชียล การแจกแจงแบบปลอมปน ได้แก่การแจกแจงแบบปกติปลอมปน 10% และ 30% ซึ่งมีรายละเอียดการดำเนินการในบทที่ 3 ได้ผลการศึกษาเป็นดังนี้

กรณีกลุ่มตัวอย่างถูกสุ่มอย่างอิสระจากประชากร 3 กลุ่ม

ซึ่งมีกรณีย่อย 2 กรณีคือ

กรณีที่ 1 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน จากการศึกษาโดยใช้ตัวอย่างสุ่มจากประชากรซึ่งมีการแจกแจงแบบต่าง ๆ ตามที่กำหนด และมีค่าเฉลี่ยเท่ากันเมื่อกำหนดระดับความมีนัยสำคัญสำหรับการทดสอบของแต่ละตัวสถิติเท่ากับ 0.05 และ 0.10 พบความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ดังตารางที่ 8

ตารางที่ 8 ความน่าจะเป็นของความคลาดเลื่อนชนิดที่ 1 ของตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis test เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

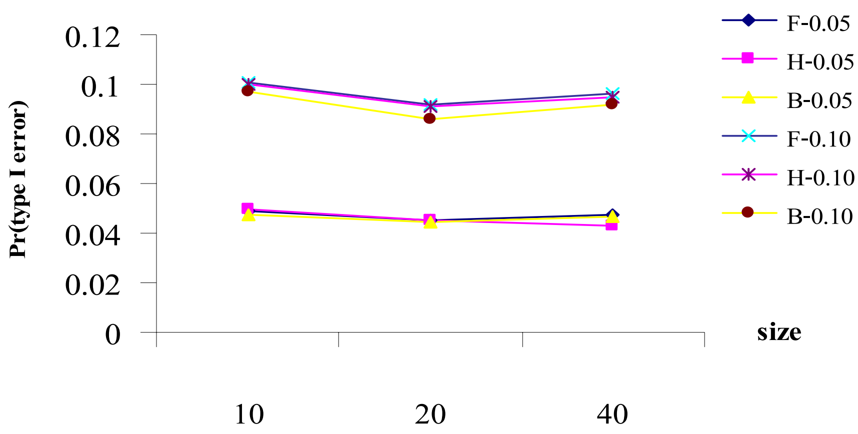
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.0487	0.0452	0.0477	0.0500	0.0450	0.0428	0.0472	0.0445	0.0465
	0.10	0.1007	0.0915	0.0965	0.1000	0.0908	0.0945	0.0967	0.0860	0.0918
Uniform	0.05	0.0547	0.0468	0.0457	0.0500	0.0448	0.0425	0.0472	0.0450	0.0475
	0.10	0.0988	0.0897	0.0945	0.1000	0.0913	0.0948	0.0983	0.0878	0.0935

ตารางที่ 8(ต่อ)

การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Chi-squared	0.05	0.0457	0.0428	0.0452	0.0492	0.0472	0.0505	0.0452	0.0428	0.0500
	0.10	0.1035	0.0908	0.0957	0.0960	0.1005	0.1023	0.1032	0.0922	0.0938
Exponential	0.05	0.0385	0.0437	0.0420	0.0500	0.0450	0.0428	0.0437	0.0425	0.0448
	0.10	0.0932	0.0970	0.0935	0.1000	0.0908	0.0940	0.0983	0.0920	0.0908
10% outlier	0.05	0.0033	0.0030	0.0040	0.0365	0.0305	0.0322	0.0195	0.0162	0.0205
	0.10	0.0103	0.0153	0.0168	0.0728	0.0660	0.0705	0.0417	0.0413	0.0483
30% outlier	0.05	0.0058	0.0045	0.0085	0.0217	0.0162	0.0155	0.0122	0.0095	0.0153
	0.10	0.0223	0.0192	0.0260	0.0477	0.0420	0.0408	0.0330	0.0322	0.0340

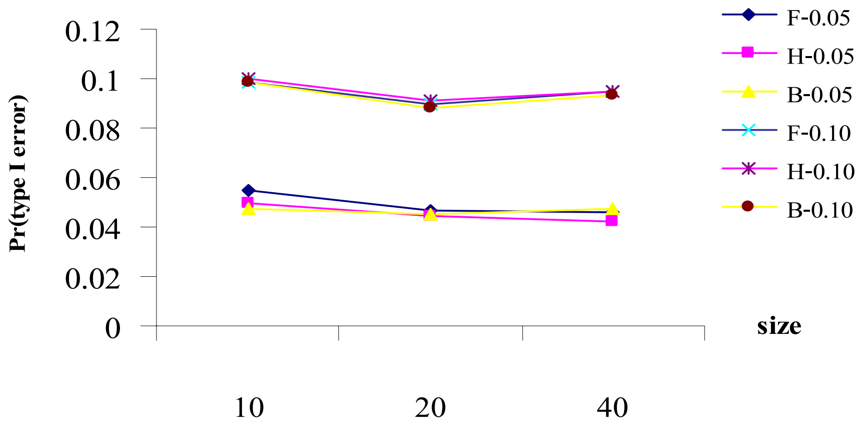
เพื่อความสะดวกในการพิจารณาตารางที่ 8 จึงแสดงความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ทั้งระดับนัยสำคัญที่ 0.05 และ 0.10 สำหรับแต่ละการแจกแจง ด้วยกราฟโดยกำหนดสัญลักษณ์แทน ดังนี้

F-0.05 แทน ตัวสถิติทดสอบ F ที่ระดับนัยสำคัญ 0.05
H-0.05 แทน ตัวสถิติทดสอบ Kruskal-Wallis ที่ระดับนัยสำคัญ 0.05
B-0.05 แทน ตัวสถิติทดสอบ B-aspect ที่ระดับนัยสำคัญที่ 0.05
F-0.10 แทน ตัวสถิติทดสอบ F ที่ระดับนัยสำคัญ 0.10
H-0.10 แทน ตัวสถิติทดสอบ Kruskal-Wallis ที่ระดับนัยสำคัญ 0.10
B-0.10 แทน ตัวสถิติทดสอบ B-aspect ที่ระดับนัยสำคัญที่ 0.10



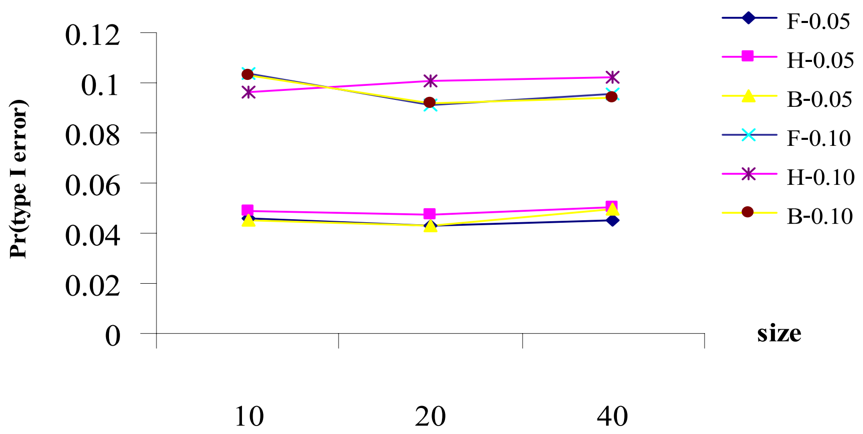
รูปที่ 7 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบปกติ กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 7 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีแจกแจงแบบปกติ พบว่า ที่ระดับนัยสำคัญ 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิด มีความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ใกล้ระดับนัยสำคัญที่กำหนด



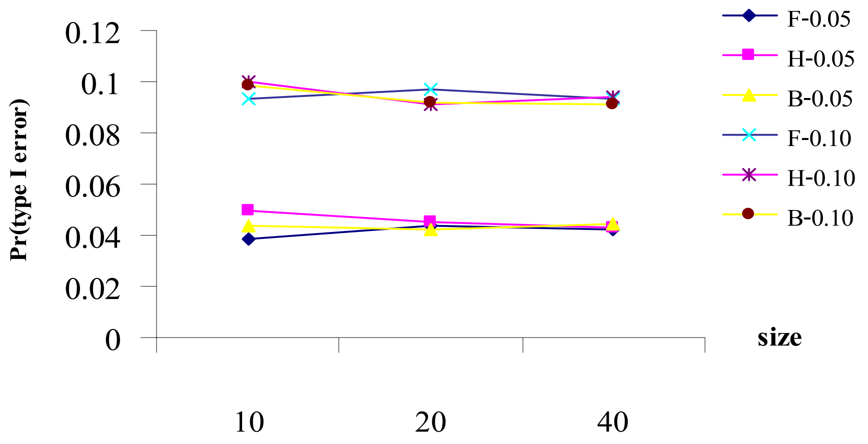
รูปที่ 8 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบยูนิฟอร์ม กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 8 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์ม พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิดยังควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดี ความแตกต่างน่าจะเกิดจาก sampling error ไม่มีรูปแบบที่แน่นอน



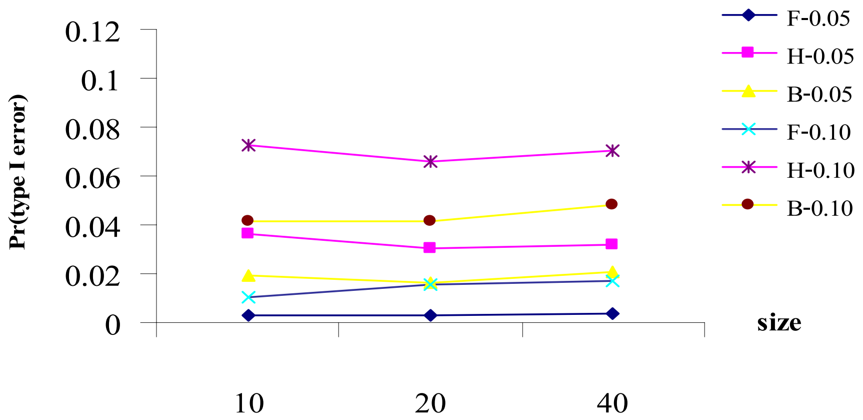
รูปที่ 9 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบไคสแควร์ กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 9 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis ควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบตัวอื่น เล็กน้อย



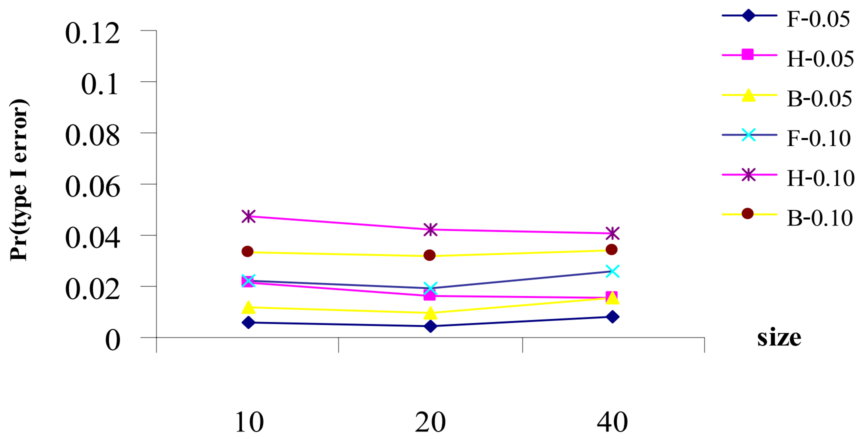
รูปที่ 10 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 10 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียล พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิด ควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ใกล้เคียงกัน



รูปที่ 11 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบปลอมปน 10% กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 11 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 10% พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis ควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ Bi-aspect ส่วนตัวสถิติทดสอบ F จะไม่สามารถควบคุมได้



รูปที่ 12 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบปลอมปน 30% กรณีกลุ่มตัวอย่าง 3 กลุ่ม

จากรูปที่ 12 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 30% พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis ควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ Bi-aspect ส่วนตัวสถิติทดสอบ F จะไม่สามารถควบคุมได้

กรณีที่ 2 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกัน

จากการศึกษาโดยใช้ตัวอย่างสุ่มจากประชากรซึ่งมีการแจกแจงแบบต่าง ๆ ตามที่กำหนด โดยมีค่าเฉลี่ยแตกต่างกันซึ่งแบ่งเป็นกรณีย่อย 2 กรณีคือ

กรณี 2.1 มีค่าเฉลี่ยของประชากรเพียง 1 กลุ่มที่แตกต่างจากกลุ่มอื่น

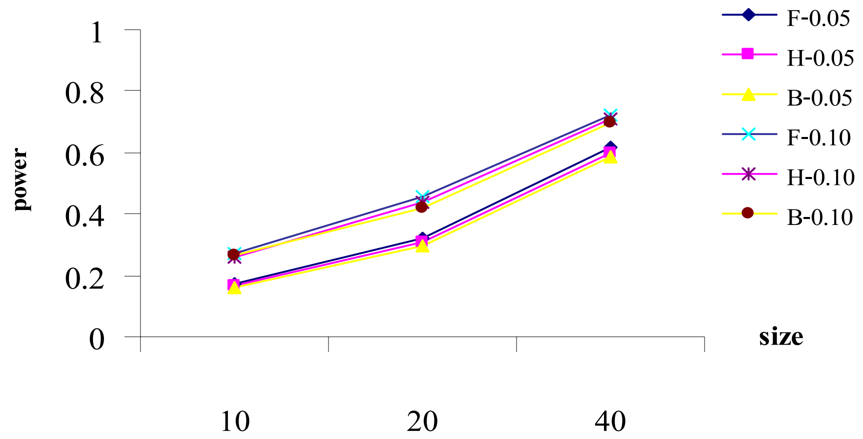
กรณี 2.1.1 มีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 9

ตารางที่ 9 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

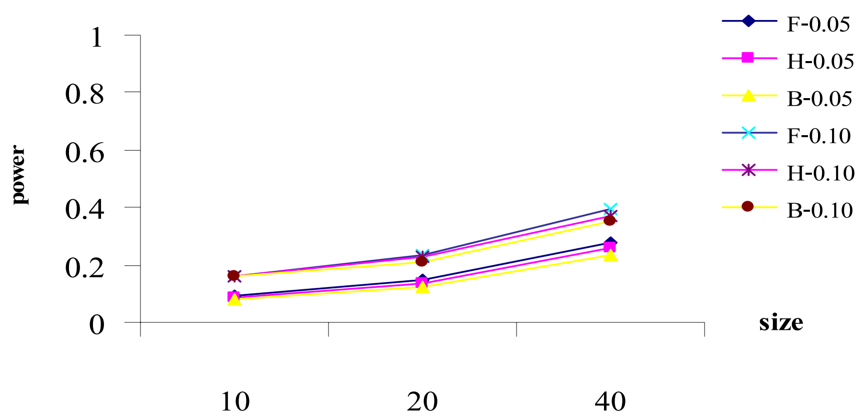
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.1702	0.3220	0.6155	0.1640	0.3065	0.5972	0.1575	0.2988	0.5893
	0.10	0.2697	0.4540	0.7232	0.2605	0.4397	0.7097	0.2650	0.4223	0.6990
Uniform	0.05	0.0925	0.1510	0.2750	0.0860	0.1377	0.2573	0.0780	0.1240	0.2345
	0.10	0.1610	0.2352	0.3938	0.1593	0.2280	0.3730	0.1580	0.2090	0.3510
Chi-squared	0.05	0.0510	0.0463	0.0545	0.0500	0.0582	0.0642	0.0477	0.0475	0.0599
	0.10	0.1040	0.1018	0.1112	0.1015	0.1117	0.1203	0.1070	0.1042	0.1065
Exponential	0.05	0.0428	0.0463	0.0498	0.0567	0.0520	0.0600	0.0460	0.0463	0.0495
	0.10	0.0983	0.0990	0.1002	0.1095	0.1055	0.1157	0.1020	0.0978	0.1058
10% outlier	0.05	0.0172	0.0393	0.1682	0.1125	0.3230	0.5095	0.0693	0.1208	0.3955
	0.10	0.0527	0.0932	0.2892	0.1953	0.2138	0.6478	0.1270	0.2120	0.5222
30% outlier	0.05	0.0235	0.0298	0.0450	0.0642	0.1002	0.1845	0.0457	0.0667	0.1445
	0.10	0.0600	0.0725	0.0920	0.1177	0.1700	0.2912	0.0908	0.1390	0.2335

ผลจากตารางแสดงด้วยกราฟสำหรับแต่ละการแจกแจงดังนี้



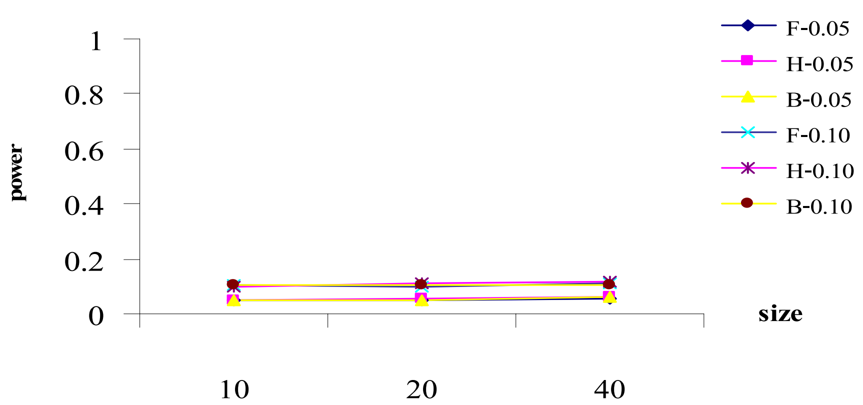
รูปที่ 13 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 13 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่เพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้นและมีกำลังการทดสอบใกล้เคียงกัน โดยตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect ตามลำดับ



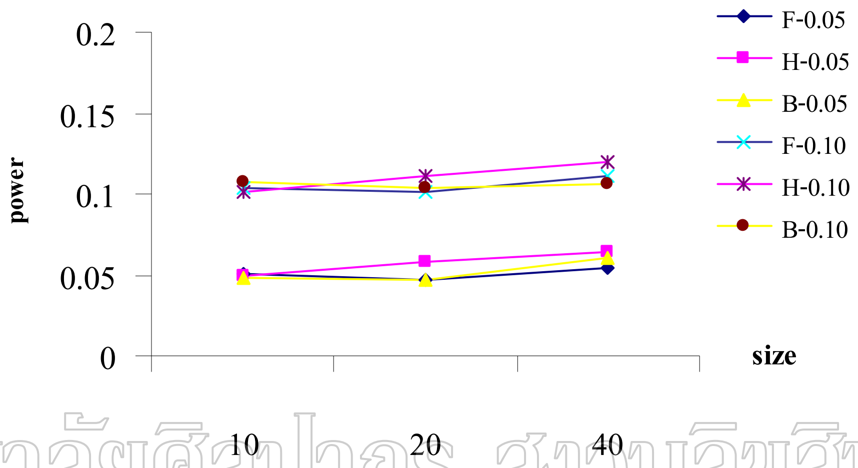
รูปที่ 14 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 14 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์มที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่เพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้น โดยตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect ตามลำดับ



รูปที่ 15 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 15 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันเพิ่มขึ้นเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มมากขึ้น และมีค่าน้อยเมื่อเทียบกับกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ในกรณีการแจกแจงแบบปกติและยูนิฟอร์ม ทั้งนี้อาจเนื่องมาจากเป็นการแจกแจงแบบไคสแควร์มีความแตกต่างของค่าเฉลี่ยน้อยมาก เมื่อเทียบกับความแปรปรวนทำให้มีกำลังการทดสอบต่ำ จึงทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นดังนี้



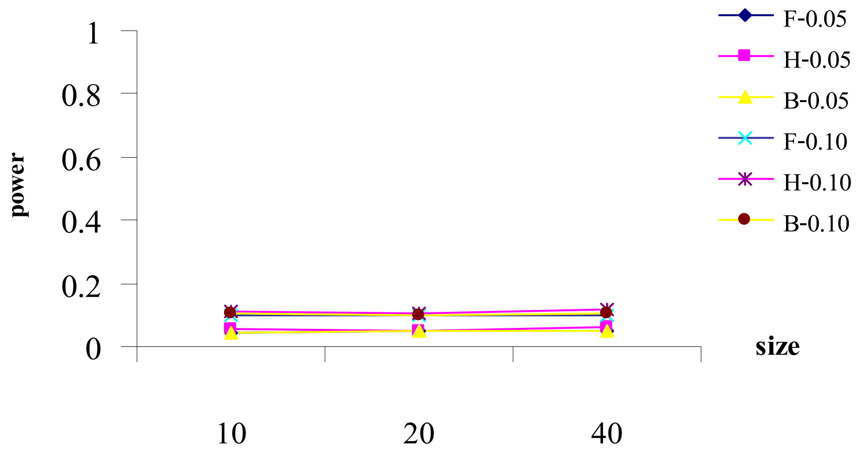
รูปที่ 16 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์ กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 16 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มมาจากประชากรที่มีการแจกแจงแบบไคสแควร์ ที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า

เมื่อขนาดตัวอย่างเป็น 10 ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ตามลำดับ ส่วนที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ F และ Kruskal-Wallis ตามลำดับ

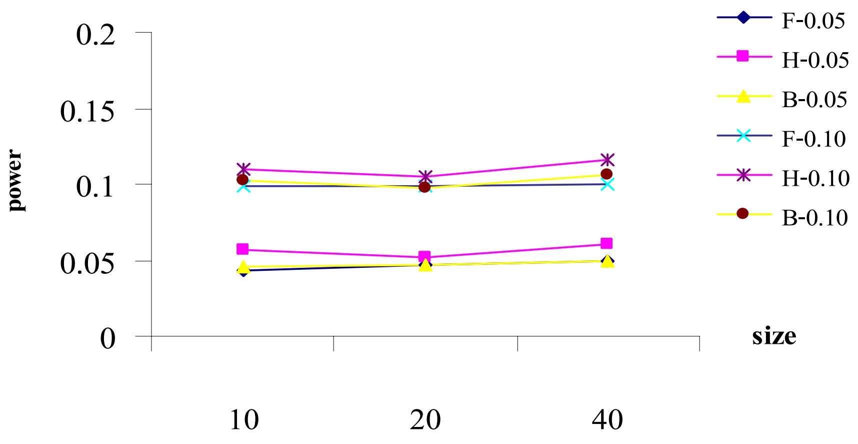
เมื่อขนาดตัวอย่างเป็น 20 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ ทั้งระดับนัยสำคัญ 0.05 และ 0.10

เมื่อขนาดตัวอย่างเพิ่มเป็น 40 ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ ส่วนที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ตามลำดับ



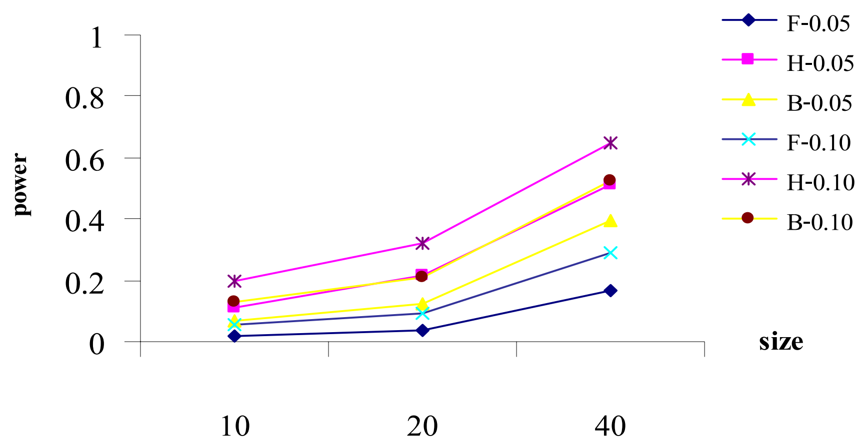
รูปที่ 17 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 17 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันเพิ่มขึ้นเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มมากขึ้น เมื่อทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นดังนี้



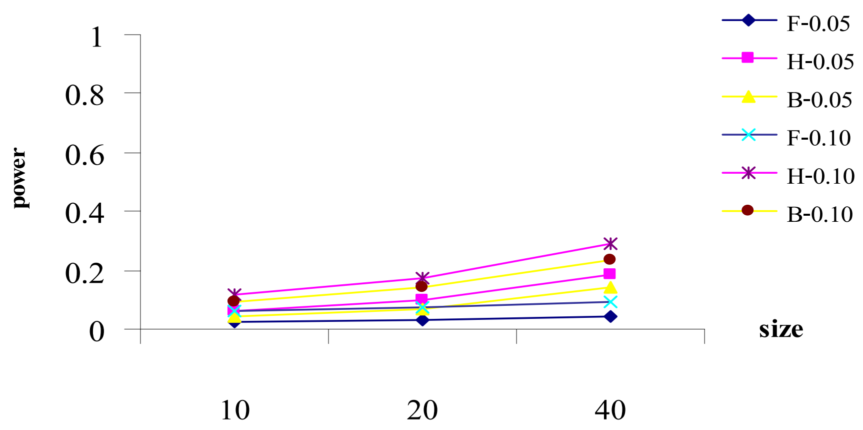
รูปที่ 18 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 18 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ



รูปที่ 19 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10%
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 19 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 10% ที่ระดับ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ



รูปที่ 20 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30%
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

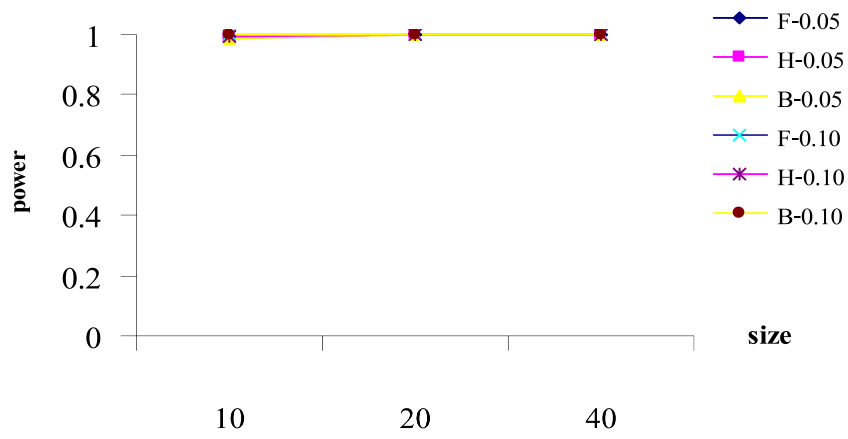
สำหรับกำลังการทดสอบภายใต้การแจกแจงแบบปลอมปน 30% ในรูปที่ 20 พบว่า ที่ระดับ 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ

กรณี 2.1.2 มีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20% พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 10

ตารางที่ 10 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

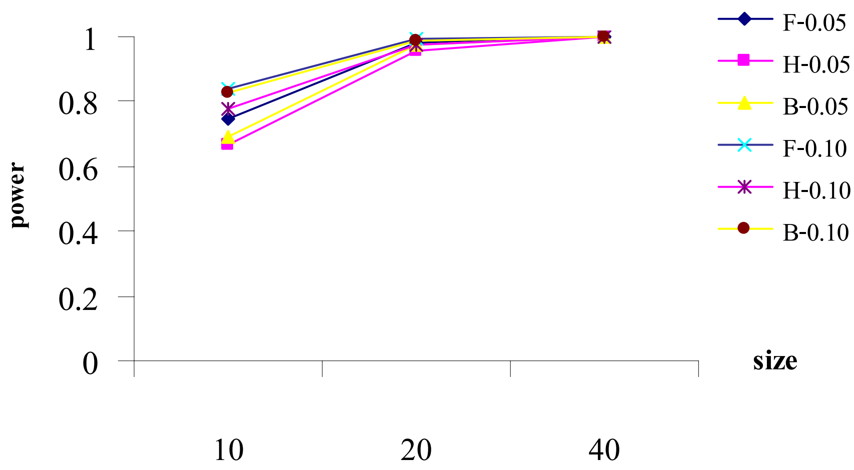
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.9940	1	1	0.9887	1	1	0.9893	1	1
	0.10	0.9973	1	1	0.9960	1	1	0.9973	1	1
Uniform	0.05	0.7465	0.9818	1	0.6645	0.9545	0.9995	0.6905	0.9735	1
	0.10	0.8420	0.9927	1	0.7795	0.9770	0.9998	0.8267	0.9887	1
Chi-squared	0.05	0.0927	0.1330	0.2420	0.1015	0.1795	0.3270	0.0900	0.1353	0.2450
	0.10	0.1632	0.2240	0.3473	0.1815	0.2775	0.4480	0.1758	0.2305	0.3490
Exponential	0.05	0.0718	0.0880	0.1447	0.1023	0.1458	0.2918	0.0737	0.0930	0.1468
	0.10	0.1398	0.1583	0.2358	0.1715	0.2418	0.4085	0.1460	0.1720	0.2415
10% outlier	0.05	0.1287	0.2068	0.4058	0.9068	0.9998	1	0.6835	0.9747	1
	0.10	0.2007	0.3153	0.5625	0.9607	1	1	0.7635	0.9950	1
30% outlier	0.05	0.0160	0.0265	0.0662	0.2675	0.7452	0.9945	0.0978	0.4325	0.9560
	0.10	0.0408	0.0700	0.1385	0.4178	0.8765	0.9977	0.1838	0.6683	0.9868

ผลจากตารางแสดงด้วยกราฟสำหรับแต่ละการแจกแจงดังนี้



รูปที่ 21 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

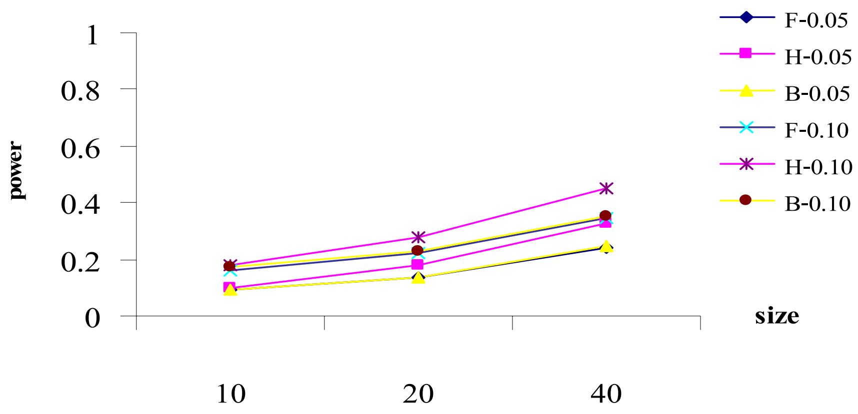
จากรูปที่ 21 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันและมีค่าสูงมากคือมีค่าเท่ากับ 1 ยกเว้นที่ขนาดตัวอย่างเป็น 10 ที่มีกำลังการทดสอบต่ำกว่า 1 เล็กน้อย ทั้งนี้เพราะค่าเฉลี่ยแตกต่างกันมากขึ้น



รูปที่ 22 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

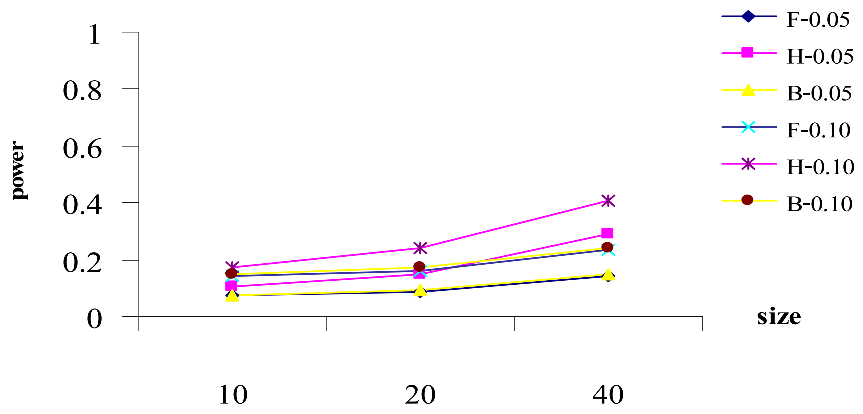
จากรูปที่ 22 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์ม ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการ

ทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น และมีกำลังการทดสอบเป็น 1 เมื่อตัวอย่างมีขนาด 40 ที่ขนาดตัวอย่างเป็น 10 และ 20 ตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุดรองลงมาเป็น ตัวสถิติทดสอบ Bi-aspect และ Kruskal-Wallis ตามลำดับ



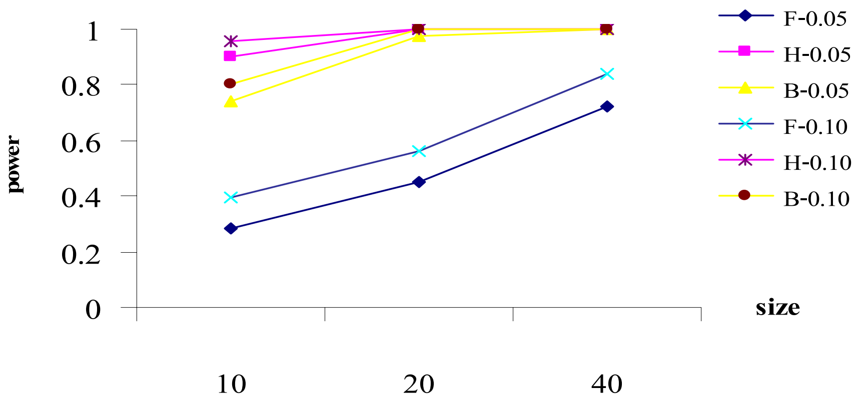
รูปที่ 23 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 23 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ซึ่งมีค่าใกล้เคียงกันมาก แต่ตัวสถิติทดสอบ Bi-aspect ดีกว่าเล็กน้อย



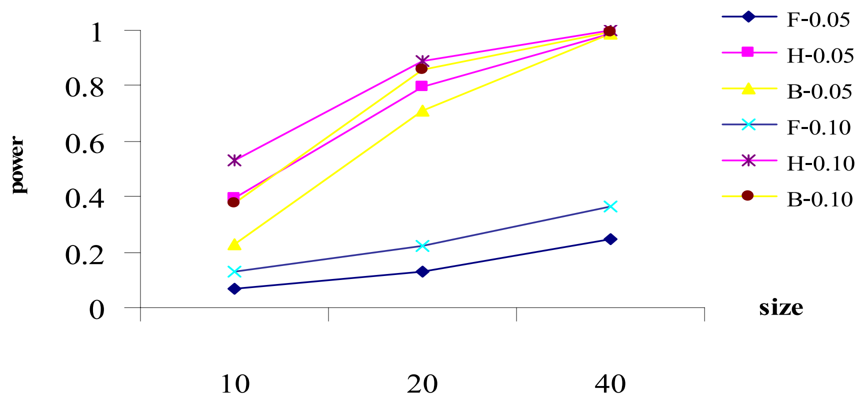
รูปที่ 24 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 24 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ซึ่งมีค่าใกล้เคียงกัน



รูปที่ 25 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10%
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 25 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 10% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น เมื่อขนาดตัวอย่างเป็น 10 และ 20 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ เมื่อขนาดตัวอย่างเป็น 40 ตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีกำลังการทดสอบเท่ากับ 1



รูปที่ 26 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30%
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 26 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ

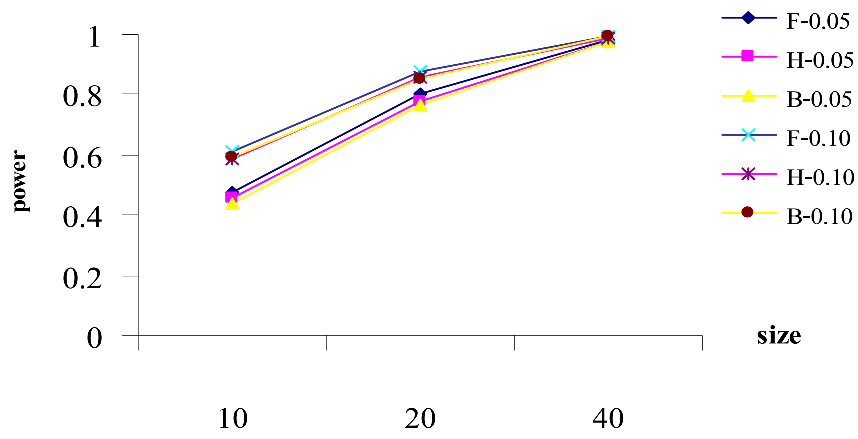
กรณี 2.2 ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน โดยทำการศึกษาที่ระดับนัยสำคัญ 0.05 และ 0.10 ใน 2 กรณีย่อยดังนี้

กรณี 2.2.1 ค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้นจากค่าเฉลี่ยของกลุ่มก่อนหน้า 5% พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 11

ตารางที่ 11 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

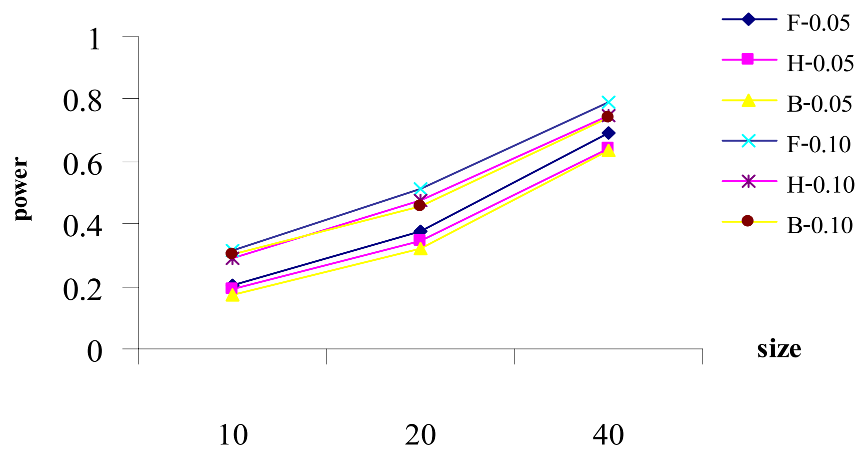
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.4730	0.7995	0.9818	0.4550	0.7753	0.9762	0.4360	0.7645	0.9755
	0.10	0.6085	0.8777	0.9935	0.5842	0.8608	0.9900	0.5953	0.8540	0.9918
Uniform	0.05	0.2023	0.3758	0.6940	0.1898	0.3470	0.6415	0.1710	0.3210	0.6373
	0.10	0.3135	0.5142	0.7928	0.2930	0.4780	0.7490	0.3005	0.4572	0.7427
Chi-squared	0.05	0.0558	0.0585	0.0795	0.0560	0.0718	0.0970	0.0535	0.0627	0.0880
	0.10	0.1098	0.1155	0.1447	0.1165	0.1290	0.1733	0.1130	0.1125	0.1472
Exponential	0.05	0.0422	0.0518	0.0553	0.0573	0.0630	0.0905	0.0455	0.0515	0.0622
	0.10	0.0985	0.1065	0.1233	0.1130	0.1233	0.1670	0.1040	0.1037	0.1152
10% outlier	0.05	0.0233	0.0265	0.0505	0.2710	0.5615	0.8928	0.1532	0.3250	0.7045
	0.10	0.0523	0.0670	0.1080	0.4013	0.6967	0.9465	0.2358	0.4680	0.8055
30% outlier	0.05	0.0080	0.0085	0.0140	0.0710	0.1608	0.4230	0.0500	0.1200	0.3575
	0.10	0.0225	0.0270	0.0405	0.1417	0.2727	0.5832	0.0890	0.2188	0.4913

ผลจากตารางแสดงด้วยกราฟของแต่ละการแจกแจง ดังนี้



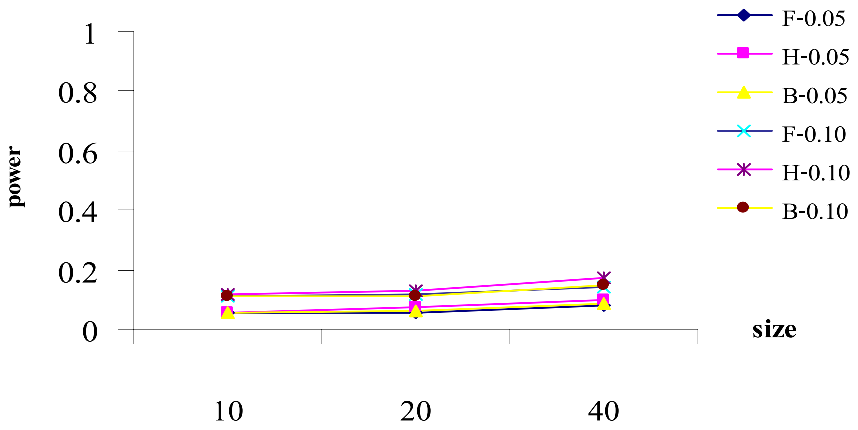
รูปที่ 27 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 27 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้นและมีค่ากำลังการทดสอบใกล้เคียงกัน โดยตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect



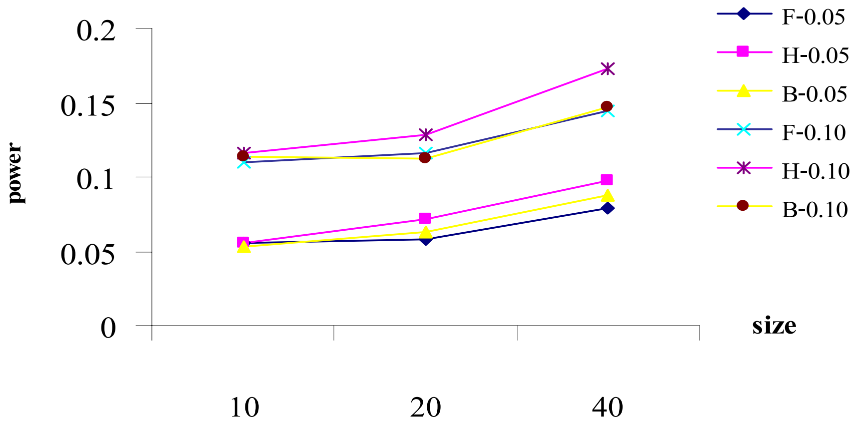
รูปที่ 28 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 28 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์มที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีค่าใกล้เคียงกัน



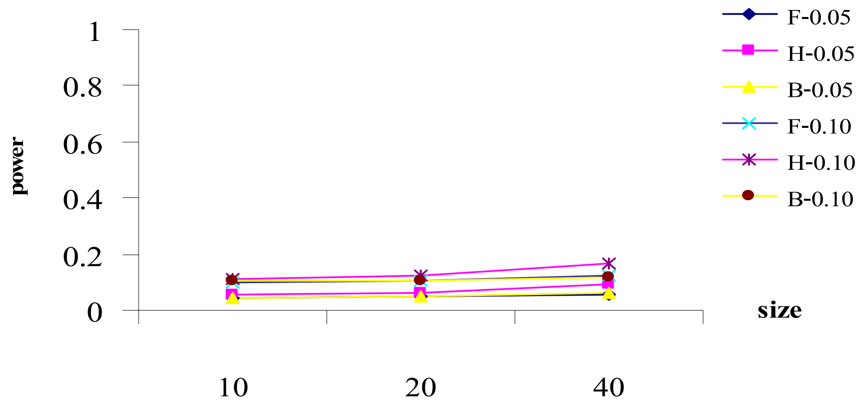
รูปที่ 29 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์ กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 29 แสดงกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันมาก และมีกำลังการทดสอบต่ำ ทั้งนี้เพราะความแตกต่างของค่าเฉลี่ยประชากรที่ตัวอย่างถูกสุ่มมีน้อยเมื่อเทียบกับค่าความแปรปรวน จึงทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นดังนี้



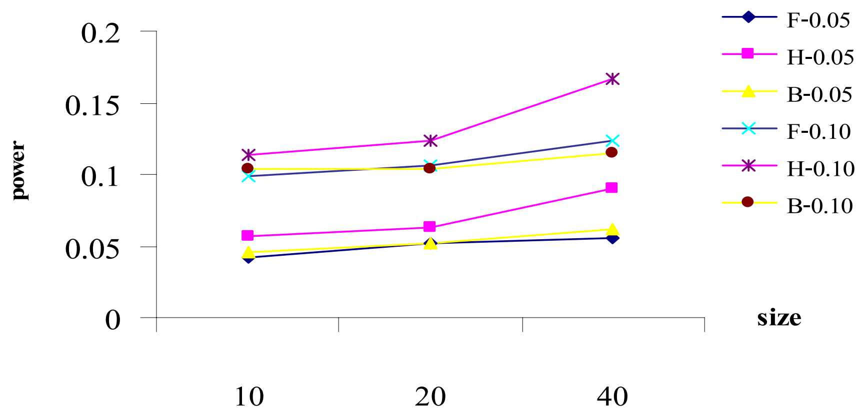
รูปที่ 30 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์ กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 30 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ที่ระดับนัยสำคัญ 0.05 แต่ที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบ Bi-aspect และ F มีกำลังการทดสอบที่ใกล้เคียงกันมาก ซึ่งตัวสถิติทดสอบ Bi-aspect ดีกว่าตัวสถิติทดสอบ F เล็กน้อย



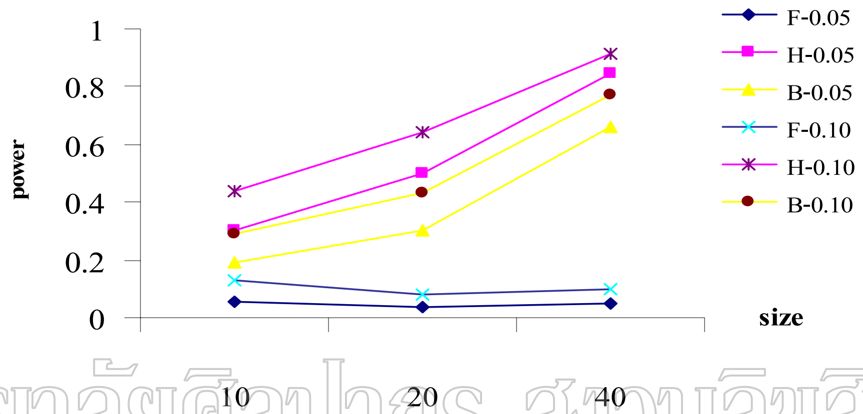
รูปที่ 31 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 31 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียลที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันมาก เนื่องจากกำลังการทดสอบน้อยจึงทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นดังนี้

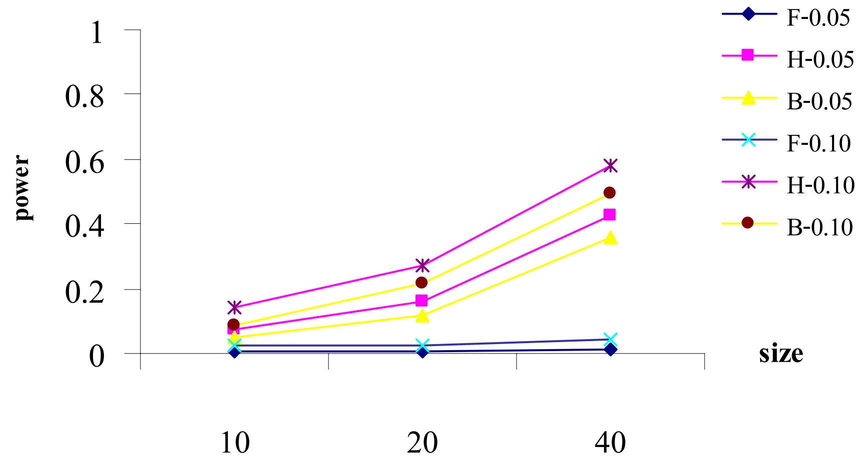


รูปที่ 32 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 32 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียลที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด ซึ่งมากกว่าตัวสถิติทดสอบ F และ Bi-aspect อย่างชัดเจน รองลงมาเมื่อขนาดตัวอย่างเป็น 10 และ 20 ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบ F เมื่อขนาดตัวอย่าง 40 ที่ระดับนัยสำคัญ 0.05 ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบ F แต่ที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบ Bi-aspect



รูปที่ 33 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10% กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%



รูปที่ 34 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30% กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

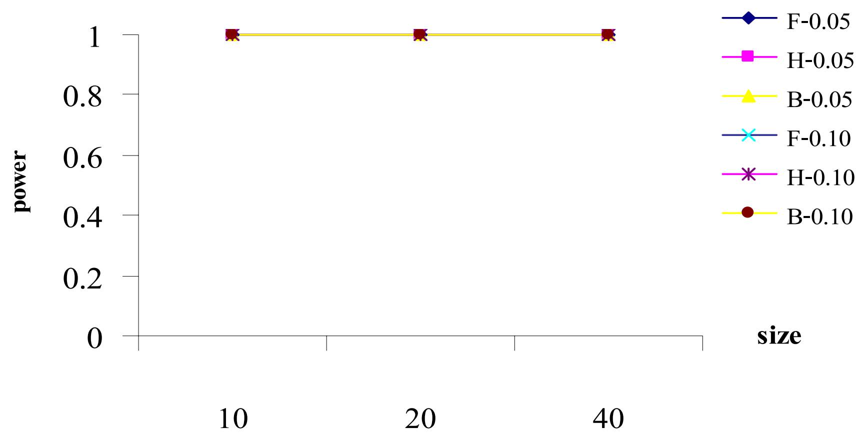
จากรูปที่ 33 และ 34 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 10% และ 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ F มีกำลังการทดสอบน้อยมากส่วนตัวสถิติทดสอบ Kruskal-Wallis และ ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงกว่าตัวสถิติทดสอบ Bi-aspect

กรณี 2.2.2 ค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้นจากค่าเฉลี่ยของกลุ่มก่อนหน้า 20%

ตารางที่ 12 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

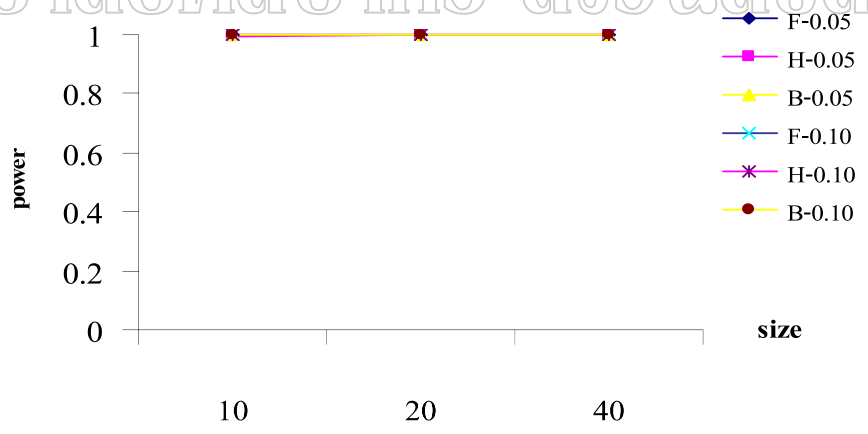
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	1	1	1	1	1	1	1	1	1
	0.10	1	1	1	1	1	1	1	1	1
Uniform	0.05	0.9987	1	1	0.9962	1	1	0.9998	1	1
	0.10	1	1	1	0.9998	1	1	0.9998	1	1
Chi-squared	0.05	0.1893	0.3485	0.6058	0.2352	0.4582	0.7602	0.1893	0.3467	0.6018
	0.10	0.2880	0.4780	0.7215	0.3487	0.5813	0.8505	0.2990	0.4780	0.7207
Exponential	0.05	0.1198	0.1928	0.3550	0.1865	0.3498	0.6655	0.1240	0.1965	0.3828
	0.10	0.1990	0.2950	0.4837	0.2833	0.4767	0.7725	0.2130	0.3033	0.5200
10% outlier	0.05	0.3225	0.5552	0.9030	0.9995	1	1	0.9910	1	1
	0.10	0.4310	0.6957	0.9535	0.9998	1	1	0.9983	1	1
30% outlier	0.05	0.0377	0.0830	0.2323	0.5820	0.9785	1	0.5880	0.9872	1
	0.10	0.0852	0.1663	0.3745	0.7655	0.9927	1	0.6730	0.9980	1

จากตารางที่ 12 แสดงกำลังการทดสอบสำหรับแต่ละการแจกแจงด้วยกราฟ ดังนี้



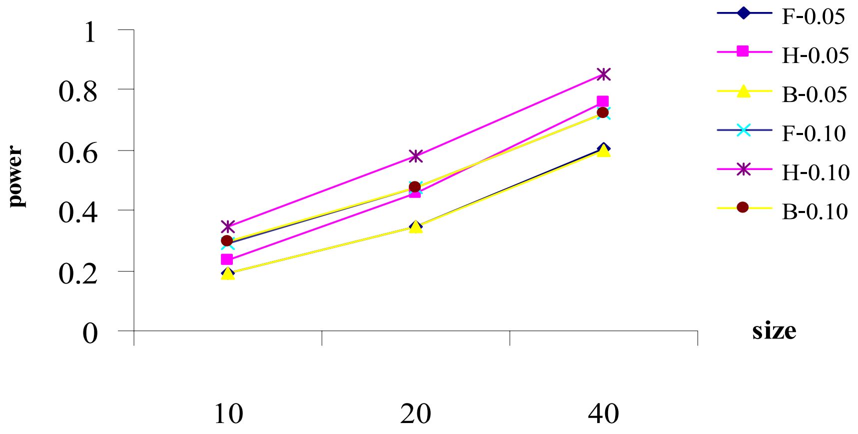
รูปที่ 35 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 35 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่ดีมากคือมีค่าเท่ากับ 1



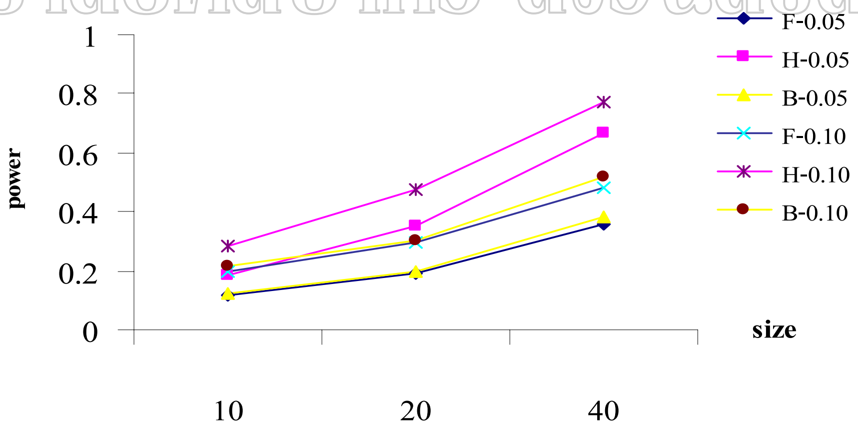
รูปที่ 36 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 36 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์ม ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่ดีมากคือมีค่าเท่ากับ 1 ยกเว้นที่ขนาดตัวอย่างเท่ากับ 10 ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเข้าใกล้ 1



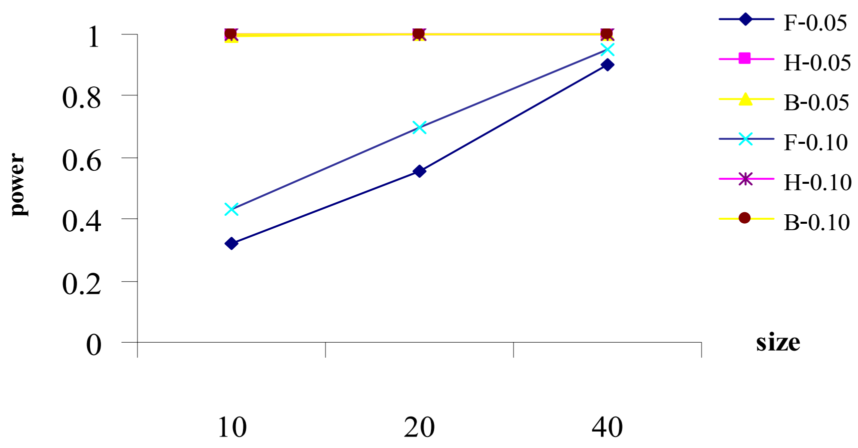
รูปที่ 37 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 37 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ที่มีกำลังการทดสอบใกล้เคียงกันมาก



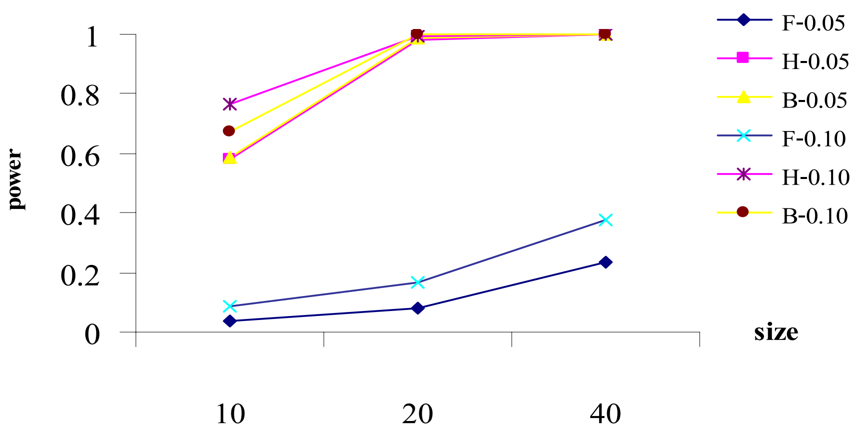
รูปที่ 38 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 38 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น โดยตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ซึ่งมีกำลังการทดสอบใกล้เคียงกันมาก โดยตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบสูงกว่าเล็กน้อย



รูปที่ 39 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10%
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 39 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 10% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีกำลังการทดสอบใกล้เคียงกันมากคือเป็น 1 เมื่อตัวอย่างมีขนาด 20 และ 40 ส่วนตัวสถิติทดสอบ F มีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น



รูปที่ 40 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30%
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 40 แสดงกราฟกำลังการทดสอบเมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปลอมปน 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ และเมื่อขนาดตัวอย่างเพิ่มเป็น 20 และ 40 ตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีกำลังการทดสอบใกล้เคียงกันมากคือเป็น 1 ส่วนตัวสถิติทดสอบ F มีกำลังการทดสอบที่น้อยกว่าอย่างชัดเจนแต่จะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

กรณีมีกลุ่มตัวอย่าง 5 กลุ่ม ได้แบ่งการศึกษาออกเป็น 2 กรณีคือ

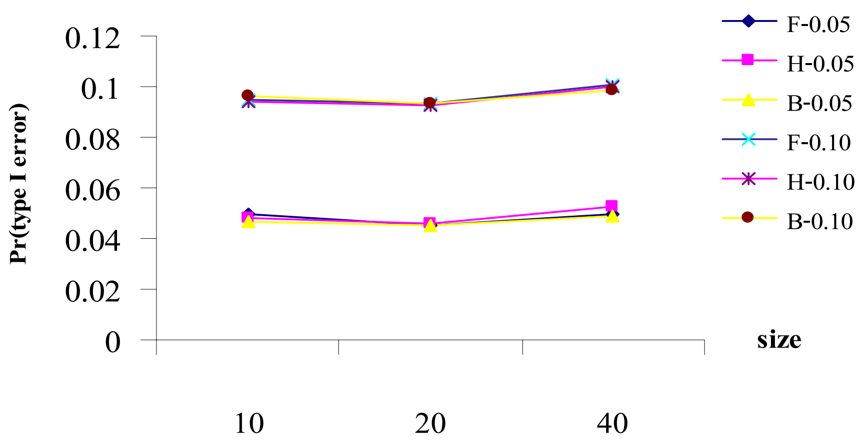
กรณีที่ 1 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน

จากการศึกษาโดยใช้ตัวอย่างสุ่มจากประชากรที่มีการแจกแจงแบบต่างๆ ตามที่กำหนดและมีค่าเฉลี่ยเท่ากันเมื่อกำหนดระดับความมีนัยสำคัญสำหรับการทดสอบของแต่ละตัวสถิติเท่ากับ 0.05 และ 0.10 ของแต่ละการแจกแจงพบความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ดังตารางที่ 13

ตารางที่ 13 ความน่าจะเป็นของความคลาดเลื่อนชนิดที่ 1 ของตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis test เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

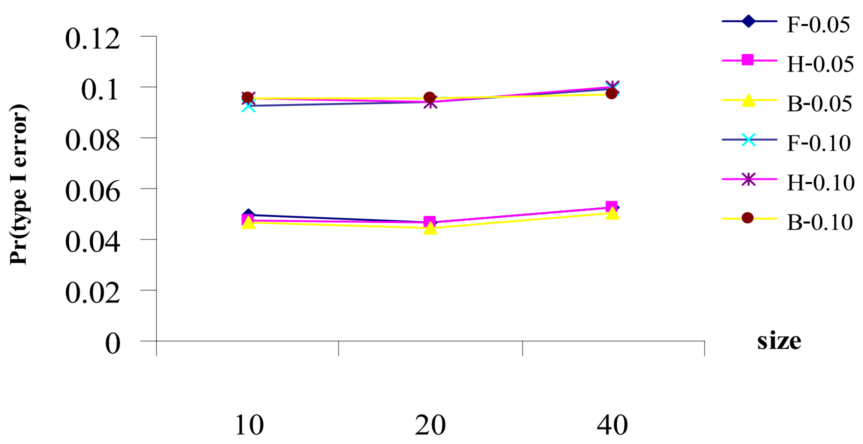
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.0498	0.0452	0.0498	0.0485	0.0457	0.0527	0.0465	0.0450	0.0492
	0.10	0.0950	0.0932	0.1005	0.0943	0.0927	0.1002	0.0962	0.0930	0.0983
Uniform	0.05	0.0498	0.0465	0.0523	0.0477	0.0468	0.0523	0.0463	0.0443	0.0503
	0.10	0.0927	0.0940	0.0990	0.0953	0.0938	0.1002	0.0953	0.0955	0.0970
Chi-squared	0.05	0.0450	0.0483	0.0550	0.0500	0.0498	0.0555	0.0505	0.0455	0.0535
	0.10	0.0922	0.0962	0.1035	0.0990	0.0960	0.1053	0.0990	0.0960	0.1070
Exponential	0.05	0.0428	0.0435	0.0487	0.0483	0.0465	0.0523	0.0498	0.0472	0.0487
	0.10	0.0900	0.0840	0.0935	0.0950	0.0932	0.1002	0.0983	0.0913	0.0985
10% outlier	0.05	0.0003	0.0018	0.0012	0.0282	0.0385	0.0347	0.0215	0.0240	0.0223
	0.10	0.0015	0.0072	0.0075	0.0615	0.0790	0.0787	0.0452	0.0503	0.0425
30% outlier	0.05	0.0010	0.0030	0.0040	0.0160	0.0110	0.0135	0.0077	0.0137	0.0103
	0.10	0.0093	0.0113	0.0150	0.0335	0.0292	0.0330	0.0245	0.0258	0.0310

เพื่อความสะดวกจากตารางที่ 13 แสดงความน่าจะเป็นของการเกิดความคลาดเคลื่อนชนิดที่ 1 ทั้งระดับนัยสำคัญที่ 0.05 และ 0.10 ของแต่ละการแจกแจงแสดงด้วยกราฟดังนี้



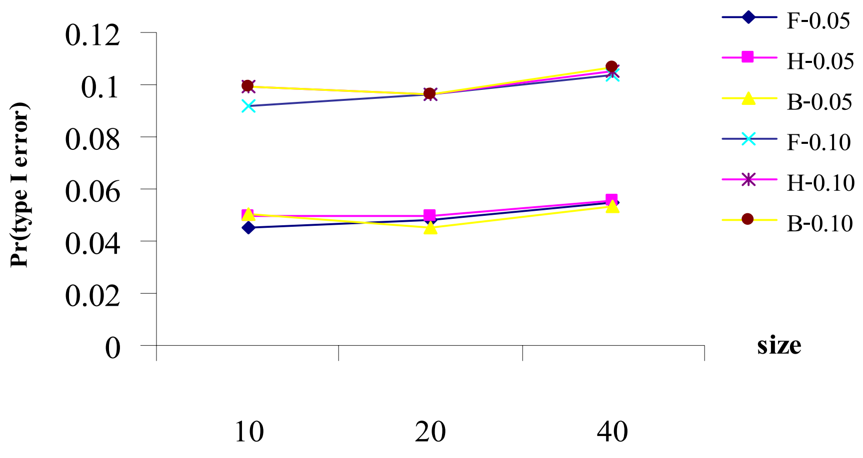
รูปที่ 41 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ภายใต้การแจกแจงแบบปกติ กรณีกลุ่มตัวอย่าง 5 กลุ่ม

จากรูปที่ 41 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบปกติ พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิด มีความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ใกล้ระดับนัยสำคัญที่กำหนด



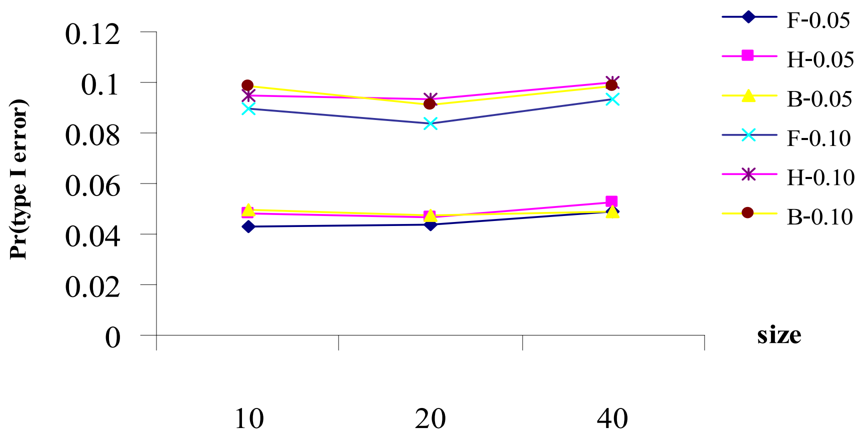
รูปที่ 42 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ภายใต้การแจกแจงแบบยูนิฟอร์ม กรณีกลุ่มตัวอย่าง 5 กลุ่ม

จากรูปที่ 42 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบยูนิฟอร์ม พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบทั้ง 3 ชนิดยังควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดี



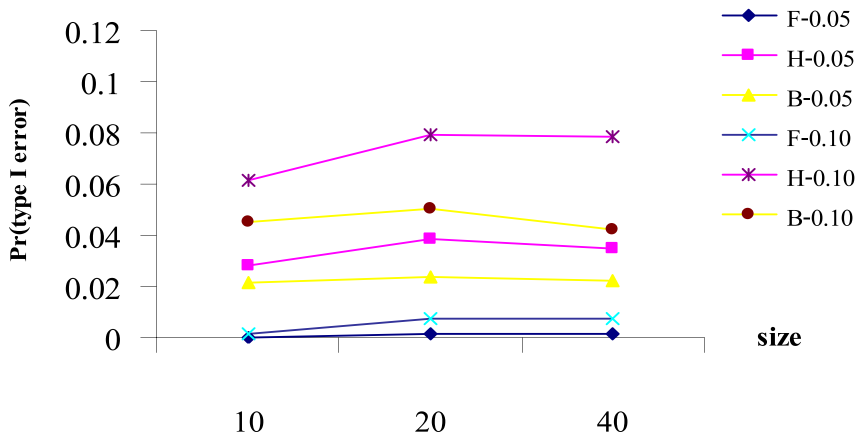
รูปที่ 43 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบไคสแควร์ กรณีกลุ่มตัวอย่าง 5 กลุ่ม

จากรูปที่ 43 แสดงกราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบไคสแควร์ พบว่า ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis และ ตัวสถิติทดสอบ Bi-aspect สามารถควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ F

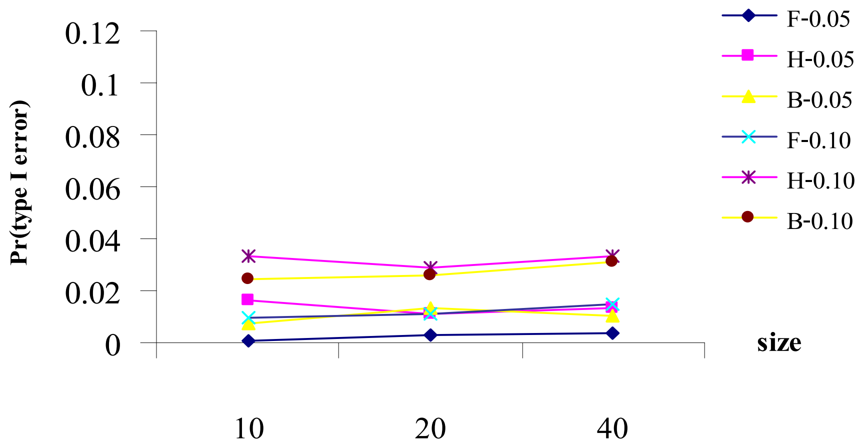


รูปที่ 44 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1
ภายใต้การแจกแจงแบบเอกซ์โปเนนเชียล กรณีกลุ่มตัวอย่าง 5 กลุ่ม

สำหรับค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อประชากรที่ตัวอย่างสุ่มมามีการแจกแจงแบบเอกซ์โปเนนเชียล ดังรูปที่ 44 ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 พบว่าตัวสถิติทดสอบ Kruskal-Wallis และ ตัวสถิติทดสอบ Bi-aspect ยังคงสามารถควบคุมการเกิดความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ F เช่นกัน



รูปที่ 45 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ภายใต้การแจกแจงแบบปลอมปน 10% กรณีกลุ่มตัวอย่าง 5 กลุ่ม



รูปที่ 46 กราฟความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ภายใต้การแจกแจงแบบปลอมปน 30% กรณีกลุ่มตัวอย่าง 5 กลุ่ม

ในกรณีของการแจกแจงแบบปลอมปน ทั้งการแจกแจงแบบปลอมปน 10% และ 30% ที่ระดับนัยสำคัญเป็น 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ Bi-aspect ส่วนตัวสถิติทดสอบ F ไม่สามารถควบคุมได้ ดังรูปที่ 45 และ 46

กรณีที่ 2 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกันศึกษาใน 2 กรณีคือ

กรณี 2.1 ค่าเฉลี่ยของประชากรเพียง 1 กลุ่มที่แตกต่างจากกลุ่มอื่น

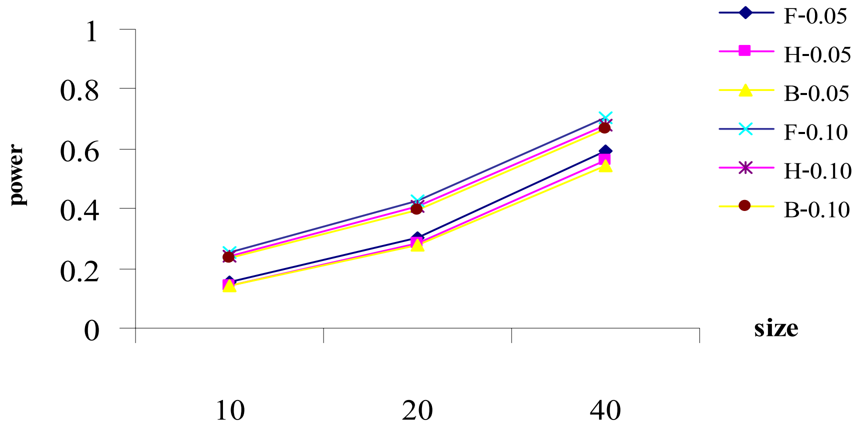
กรณี 2.1.1 มีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 14

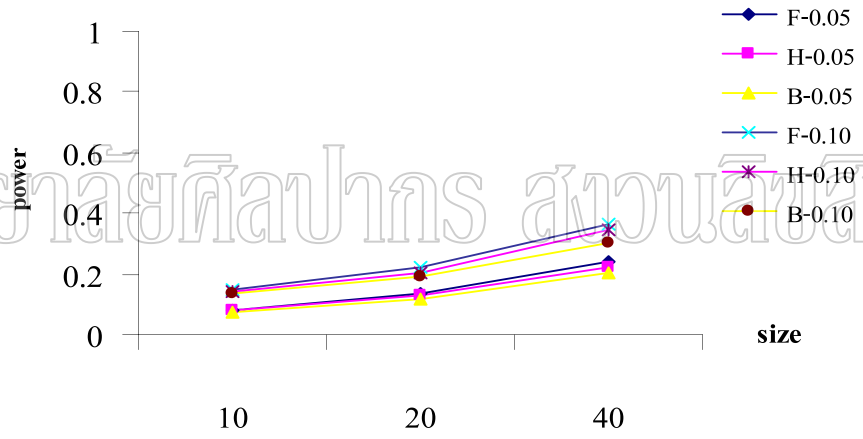
ตารางที่ 14 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.1560	0.3018	0.5928	0.1415	0.2850	0.5642	0.1405	0.2750	0.5460
	0.10	0.2533	0.4265	0.7067	0.2410	0.4085	0.6768	0.2368	0.3950	0.6693
Uniform	0.05	0.0833	0.1330	0.2400	0.0790	0.1270	0.2235	0.0747	0.1163	0.2010
	0.10	0.1485	0.2195	0.3618	0.1425	0.2060	0.3447	0.1355	0.1912	0.3043
Chi-squared	0.05	0.0477	0.0485	0.0630	0.0538	0.0560	0.0655	0.0512	0.0512	0.0600
	0.10	0.0960	0.1018	0.1142	0.1018	0.1060	0.1238	0.1040	0.1042	0.1130
Exponential	0.05	0.0460	0.0460	0.0515	0.0535	0.0523	0.0605	0.0532	0.0503	0.0515
	0.10	0.0900	0.0910	0.1000	0.0993	0.1053	0.1157	0.1018	0.0988	0.0988
10% outlier	0.05	0.0080	0.0037	0.0072	0.0755	0.1580	0.3480	0.0382	0.0887	0.1970
	0.10	0.0047	0.0155	0.0210	0.1430	0.2582	0.4855	0.0840	0.1475	0.2932
30% outlier	0.05	0.0010	0.0033	0.0043	0.0225	0.0413	0.0880	0.0113	0.0290	0.0632
	0.10	0.0093	0.0135	0.0172	0.0515	0.0850	0.1650	0.0385	0.0565	0.1225

จากตารางที่ 14 เป็นการเปรียบเทียบเทียบประสิทธิภาพตัวสถิติทดสอบ ในกรณีที่มีค่าเฉลี่ย 1 กลุ่มที่แตกต่างจากกลุ่มอื่นโดยมากกว่ากลุ่มอื่น 5% เพื่อความสะดวกจึงจะแสดงค่ากำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ในแต่ละการแจกแจงด้วยกราฟ ดังนี้

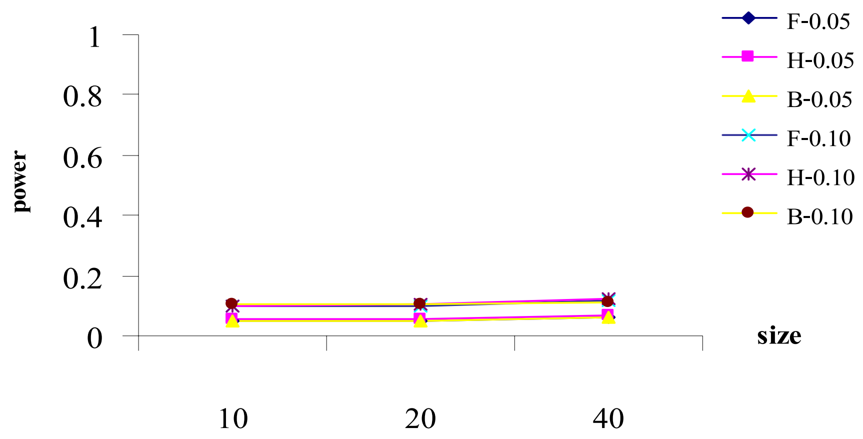


รูปที่ 47 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%



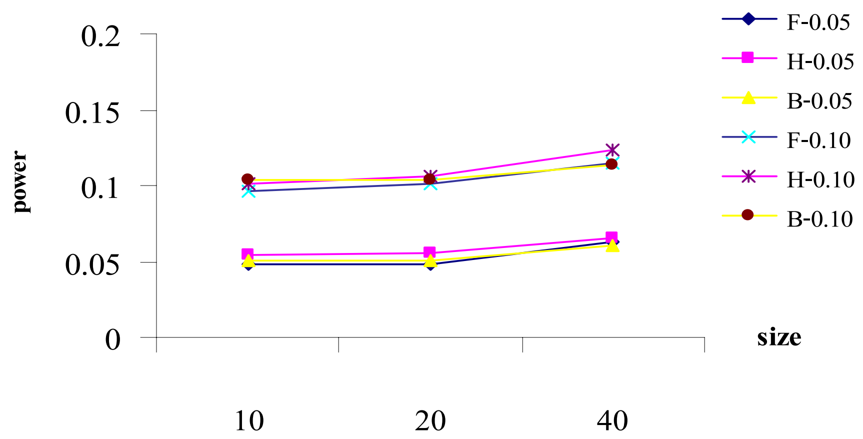
รูปที่ 48 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

ในการแจกแจงแบบปกติและการแจกแจงแบบยูนิฟอร์ม ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่เพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มมากขึ้น และมีค่ากำลังการทดสอบใกล้เคียงกัน โดยตัวสถิติทดสอบ F มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect ตามลำดับ ซึ่งแสดงให้เห็นได้ดังรูปที่ 47 และ 48



รูปที่ 49 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

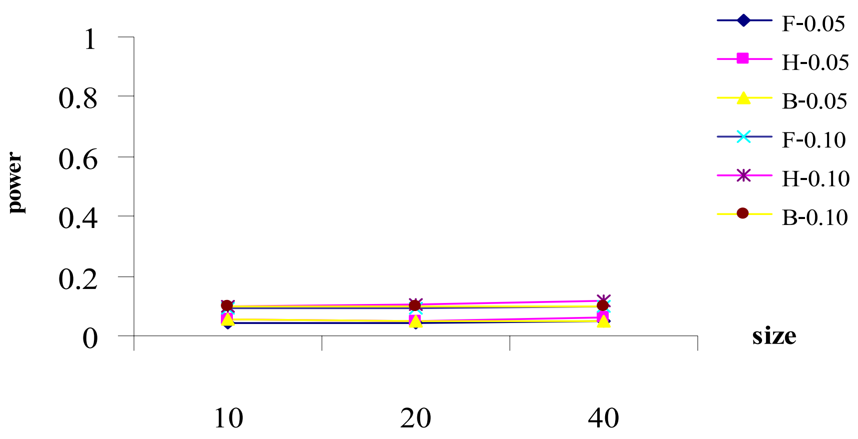
จากรูปที่ 49 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันเพิ่มขึ้นเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มมากขึ้น แต่เนื่องจากกำลังการทดสอบน้อยจึงทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นดังนี้



รูปที่ 50 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

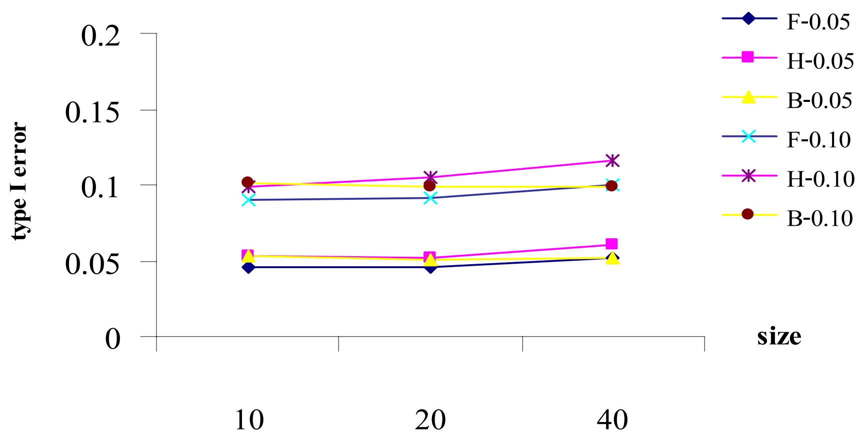
จากรูปที่ 50 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบไคสแควร์ ที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า เมื่อขนาดตัวอย่างเป็น 10 และ 20 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติ

ทดสอบ Bi-aspect และ F ตามลำดับ เมื่อขนาดตัวอย่างเพิ่มเป็น 40 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ตามลำดับ



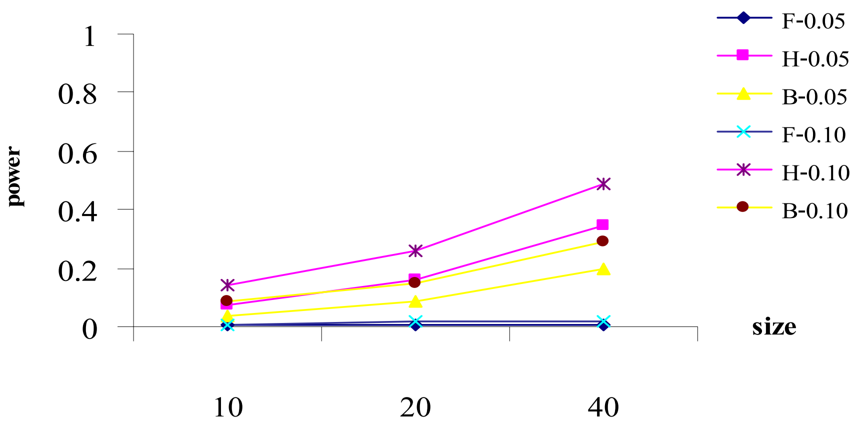
รูปที่ 51 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 51 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิด มีกำลังการทดสอบใกล้เคียงกันเพิ่มขึ้นเล็กน้อยเมื่อขนาดตัวอย่างเพิ่มมากขึ้น เมื่อทำการขยายภาพกราฟกำลังการทดสอบให้ชัดเจนขึ้นจะได้ดังนี้

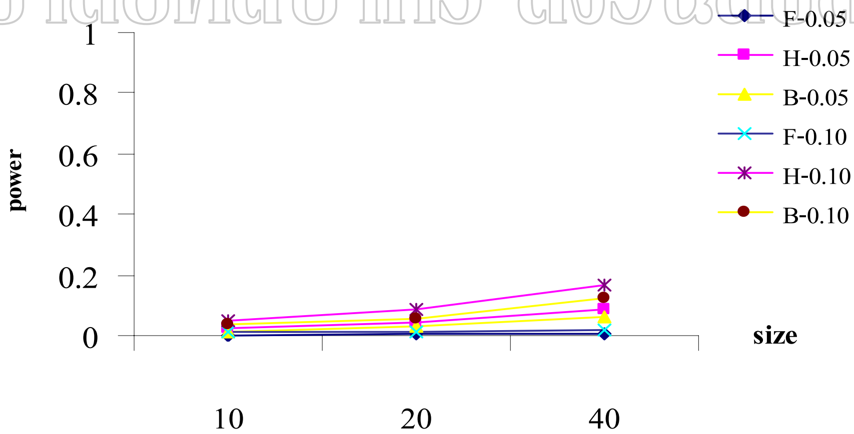


รูปที่ 52 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

จากรูปที่ 52 แสดงกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับ 0.05 และ 0.10 ซึ่งขยายให้ชัดเจนขึ้น พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ



รูปที่ 53 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10% กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%



รูปที่ 54 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30% กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 5%

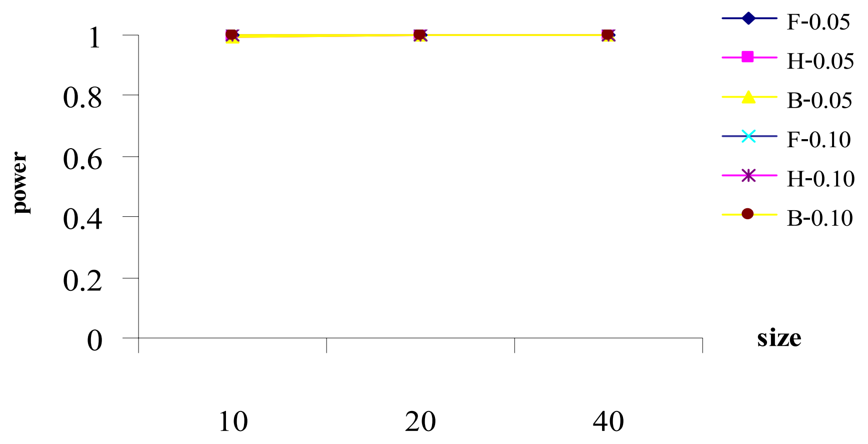
สำหรับกำลังการทดสอบ เมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปลอมปน 10% และ 30% ดังรูปที่ 53 และ 54 ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบมากที่สุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ ซึ่งกำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น ยกเว้นตัวสถิติทดสอบ F ที่มีกำลังการทดสอบน้อยมาก

กรณี 2.1.2 มีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20% พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 15

ตารางที่ 15 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

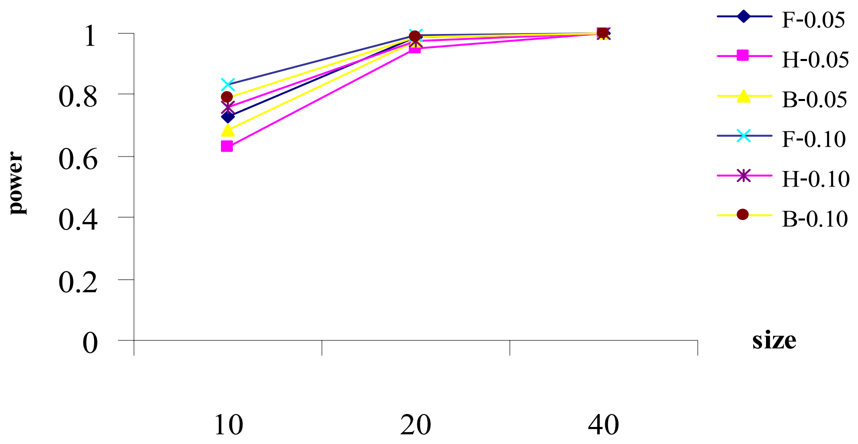
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.9977	1	1	0.9930	1	1	0.9955	1	1
	0.10	0.9995	1	1	0.9980	1	1	0.9987	1	1
Uniform	0.05	0.7305	0.9847	1	0.6270	0.9523	1	0.6865	0.9728	1
	0.10	0.8360	0.9940	1	0.7590	0.9780	1	0.7925	0.9887	1
Chi-squared	0.05	0.0790	0.1128	0.1970	0.0927	0.1500	0.2817	0.0828	0.1210	0.2055
	0.10	0.1435	0.1970	0.3170	0.1647	0.2460	0.4153	0.1535	0.2087	0.3135
Exponential	0.05	0.0620	0.0760	0.1200	0.0833	0.1325	0.2475	0.0695	0.0878	0.1325
	0.10	0.1135	0.1480	0.2020	0.1510	0.2278	0.3730	0.1318	0.1628	0.2272
10% outlier	0.05	0.0585	0.1260	0.3347	0.8940	0.9995	1	0.3927	0.9540	1
	0.10	0.1187	0.2167	0.4807	0.9535	0.9998	1	0.1312	0.9865	1
30% outlier	0.05	0.0060	0.0148	0.0428	0.1968	0.6378	0.9895	0.0550	0.2400	0.8188
	0.10	0.0207	0.0400	0.0798	0.3307	0.7905	0.9975	0.1312	0.4045	0.9243

ผลจากตารางที่ 15 สามารถแสดงด้วยกราฟสำหรับการแจกแจงต่าง ๆ ดังนี้



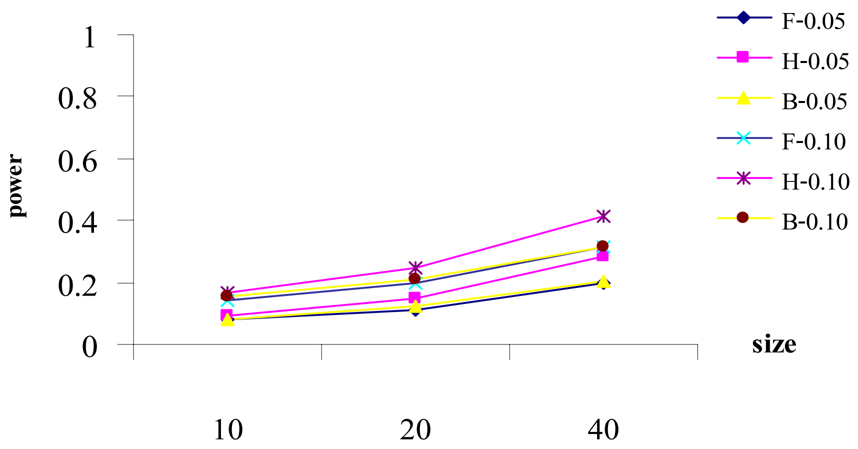
รูปที่ 55 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 55 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันโดยมีค่าเป็น 1 เมื่อขนาดตัวอย่างเป็น 20 และ 40

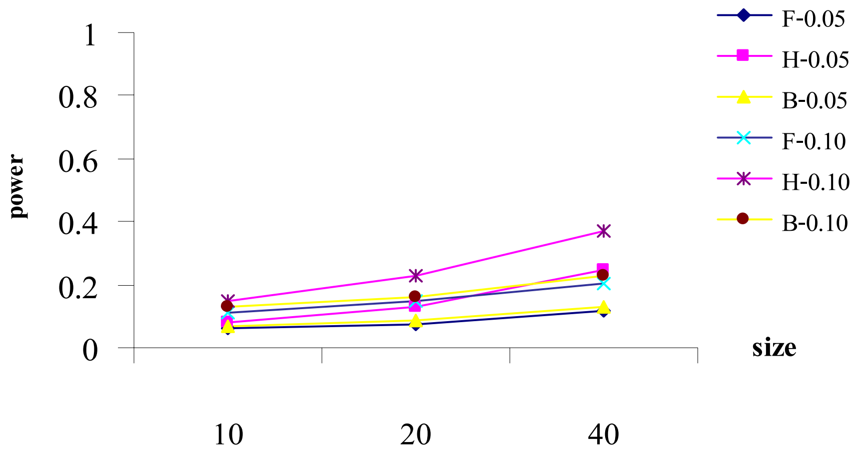


รูปที่ 56 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 56 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบยูนิฟอร์มที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ Kruskal-Wallis ตามลำดับ ซึ่งกำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น

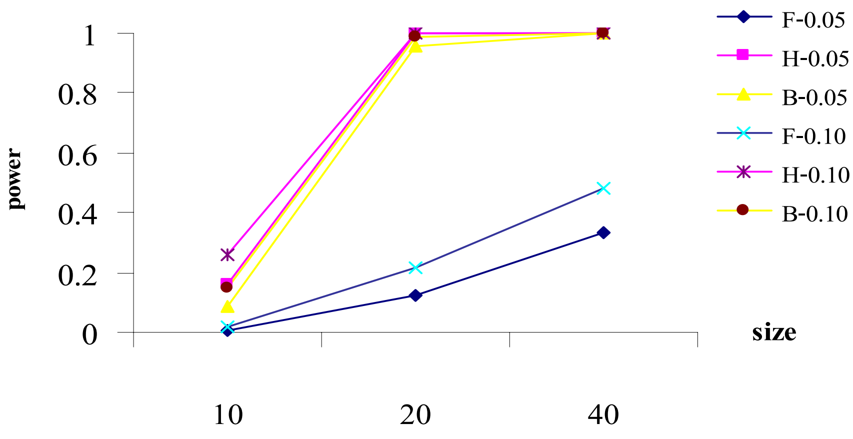


รูปที่ 57 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์ กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%



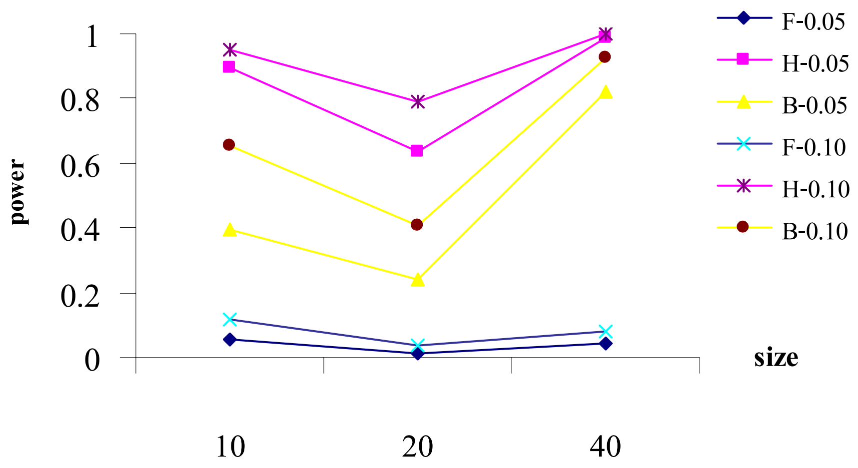
รูปที่ 58 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

จากรูปที่ 57 และ 58 แสดงกราฟกำลังการทดสอบภายใต้การแจกแจงแบบไคสแควร์ และการแจกแจงแบบเอกซ์โปเนนเชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ซึ่งมีกำลังการทดสอบใกล้เคียงกัน และเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น



รูปที่ 59 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10% กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

สำหรับกำลังการทดสอบ เมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปลอมปน 10% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุดรองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ ซึ่งกำลังการทดสอบจะเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น ดังรูปที่ 59



รูปที่ 60 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30%
กรณีมีค่าเฉลี่ย 1 กลุ่มมากกว่าค่าเฉลี่ยกลุ่มอื่น 20%

กำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปลอมปน 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุดรองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ ซึ่งกำลังการทดสอบจะลดลงเมื่อขนาดตัวอย่างเป็น 20 และเพิ่มขึ้นเมื่อขนาดตัวอย่างเป็น 40 ดังรูปที่ 60

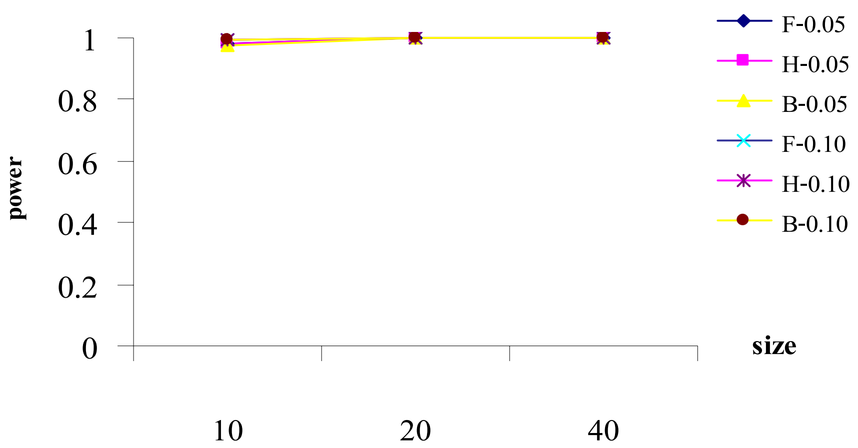
กรณี 2.2 ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน ศึกษาใน 2 กรณีย่อยดังนี้

กรณี 2.2.1 ค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้นจากค่าเฉลี่ยของกลุ่มก่อนหน้า 5% พบกำลังการทดสอบของตัวสถิติทดสอบต่าง ๆ ดังตารางที่ 16

ตารางที่ 16 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

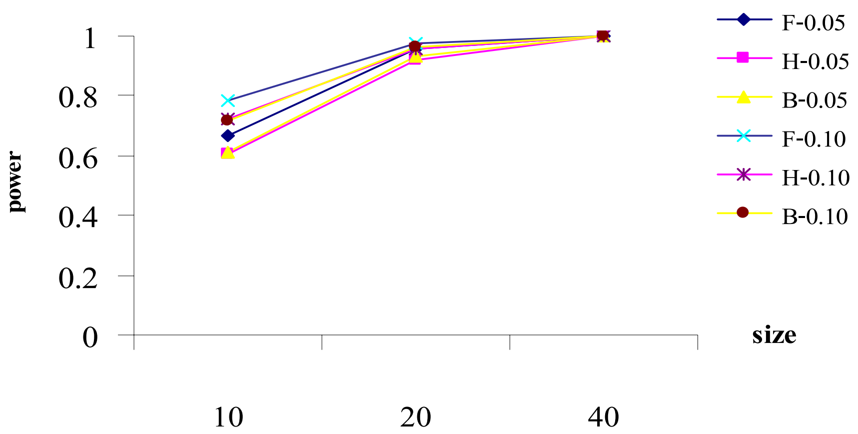
การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	0.9845	1	1	0.9803	1	1	0.9780	1	1
	0.10	0.9942	1	1	0.9927	1	1	0.9918	1	1
Uniform	0.05	0.6687	0.9538	1	0.6050	0.9172	0.9983	0.6130	0.9317	0.9995
	0.10	0.7845	0.9758	1	0.7230	0.9580	0.9998	0.7188	0.9638	1
Chi-squared	0.05	0.0688	0.0970	0.1883	0.0852	0.1270	0.2570	0.0712	0.1007	0.1733
	0.10	0.1335	0.1713	0.2825	0.1488	0.2163	0.3710	0.1370	0.1748	0.2738
Exponential	0.05	0.0535	0.0665	0.1025	0.0752	0.1208	0.2107	0.0645	0.0828	0.1112
	0.10	0.1045	0.1295	0.1755	0.1373	0.2065	0.3225	0.1115	0.1382	0.1898
10% outlier	0.05	0.0373	0.0782	0.1928	0.8388	0.9980	1	0.6198	0.9613	1
	0.10	0.0772	0.1440	0.3070	0.9165	0.9992	1	0.7575	0.9790	1
30% outlier	0.05	0.0027	0.0085	0.0160	0.2253	0.6310	0.9718	0.1865	0.6357	0.9760
	0.10	0.0153	0.0260	0.0498	0.3530	0.7663	0.9895	0.3155	0.7477	0.9877

ผลจากตารางแสดงด้วยกราฟของตัวสถิติทดสอบสำหรับการแจกแจงต่าง ๆ ดังนี้



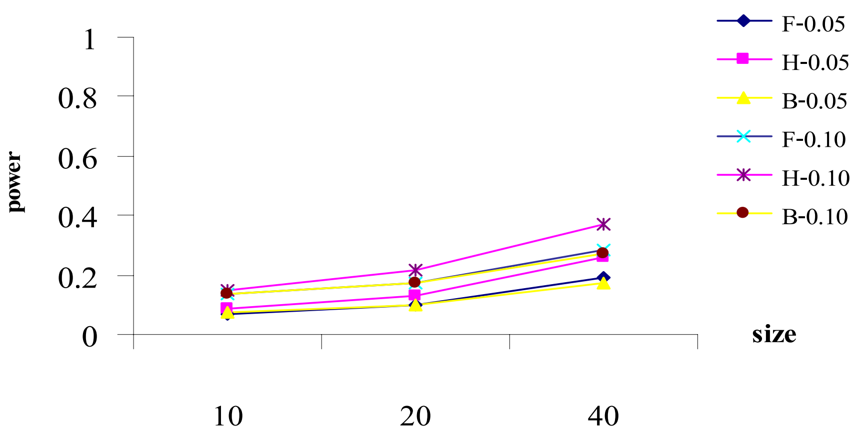
รูปที่ 61 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 61 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปกติที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันโดยมีค่าเป็น 1 เมื่อขนาดตัวอย่างเป็น 20 และ 40

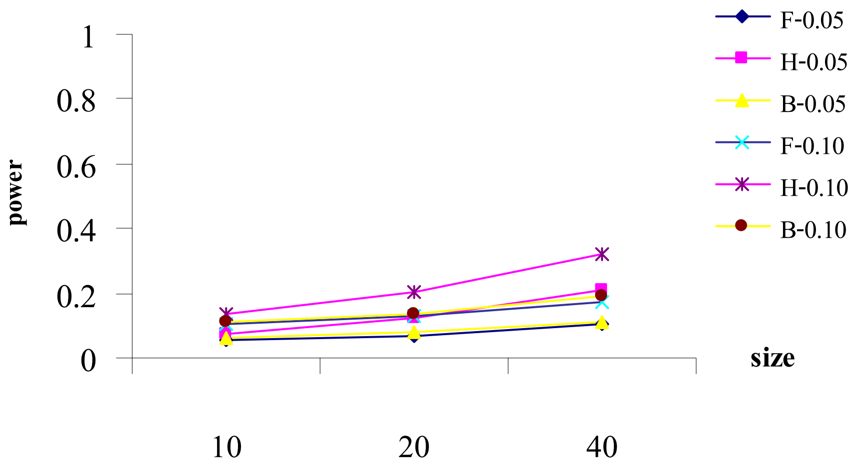


รูปที่ 62 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 62 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบยูนิฟอร์ม ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ Kruskal-Wallis ซึ่งมีกำลังการทดสอบใกล้เคียงกันแต่ตัวสถิติทดสอบ Bi-aspect จะดีกว่าเล็กน้อย จะมีค่าเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น และมีค่าเข้าใกล้ 1 เมื่อขนาดตัวอย่างเป็น 40

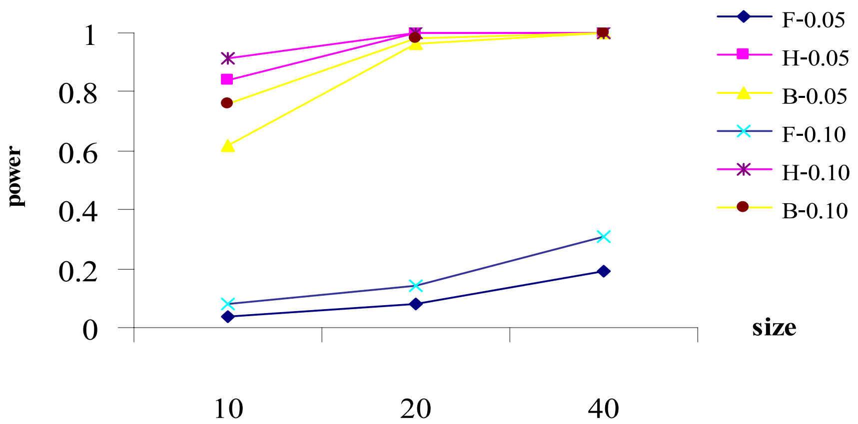


รูปที่ 63 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

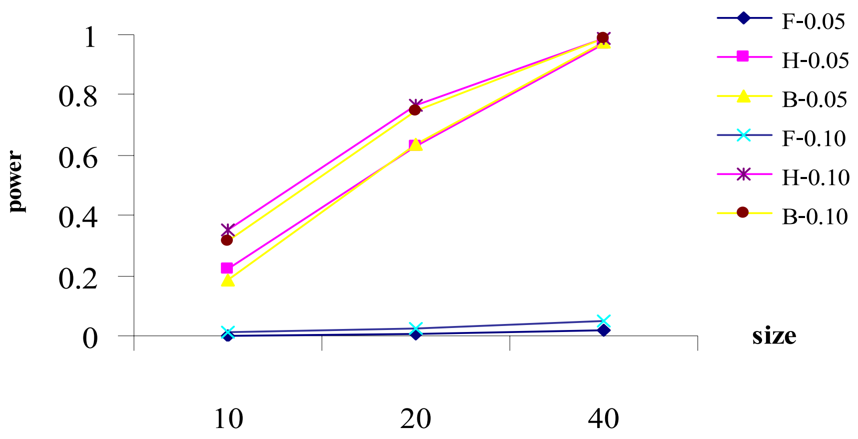


รูปที่ 64 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนเชียล กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

จากรูปที่ 63 และ รูปที่ 64 แสดงกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบไคสแควร์ และการแจกแจงแบบเอกซ์โปเนนเชียล ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ซึ่งมีค่าใกล้เคียงกันมากและจะเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น



รูปที่ 65 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10% กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%



รูปที่ 66 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30%
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 5%

สำหรับการแจกแจงแบบปลอมปน 10% และ 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ โดยตัวสถิติทดสอบ F มีกำลังการทดสอบที่น้อยกว่าอย่างชัดเจน แสดงได้ดังรูปที่ 65 และ 66

กรณี 2.2.2 ค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้นจากค่าเฉลี่ยของกลุ่มก่อนหน้า 20%

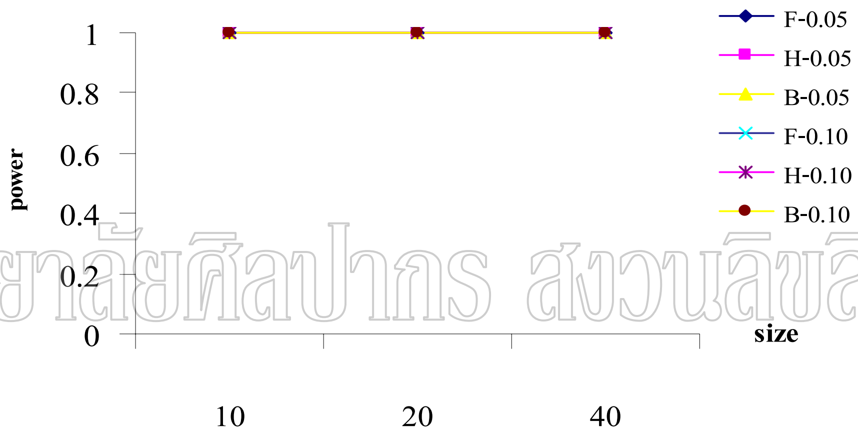
ตารางที่ 17 กำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect F และ Kruskal- Wallis เมื่อกำหนดระดับนัยสำคัญเท่ากับ 0.05 และ 0.10 สำหรับการแจกแจงของประชากรชนิดต่าง ๆ

การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
Normal	0.05	1	1	1	1	1	1	1	1	1
	0.10	1	1	1	1	1	1	1	1	1
Uniform	0.05	1	1	1	1	1	1	1	1	1
	0.10	1	1	1	1	1	1	1	1	1
Chi-squared	0.05	0.5623	0.8900	0.9973	0.6867	0.9553	0.9998	0.5732	0.8940	0.9973
	0.10	0.6925	0.9340	0.9995	0.7897	0.9783	1	0.7013	0.9427	0.9998
Exponential	0.05	0.0688	0.5935	0.8947	0.0852	0.8800	0.9973	0.0712	0.6485	0.7330
	0.10	0.1335	0.7067	0.9442	0.1488	0.9310	0.9990	0.1370	0.7632	0.9645

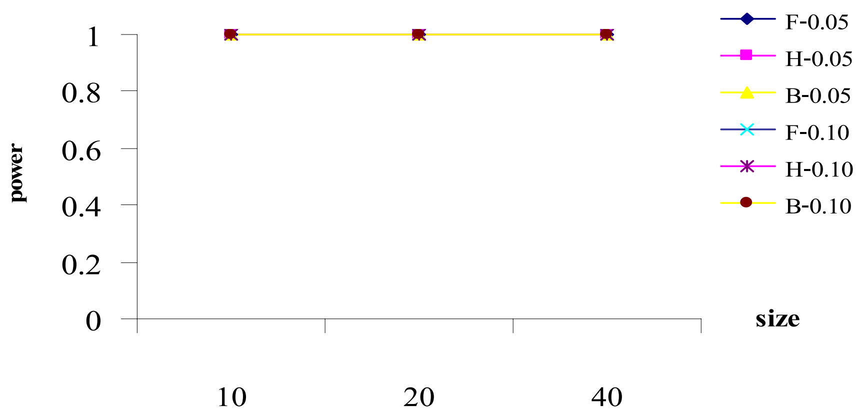
ตารางที่ 17(ต่อ)

การแจกแจง	ระดับนัยสำคัญ	ตัวสถิติทดสอบ								
		F			KW			Bi-aspect		
		10	20	40	10	20	40	10	20	40
10% outlier	0.05	0.7215	0.9862	1	1	1	1	1	1	1
	0.10	0.8195	0.9955	1	1	1	1	1	1	1
30% outlier	0.05	0.0927	0.3288	0.8158	0.9103	1	1	0.9613	1	1
	0.10	0.1785	0.4908	0.9072	0.9705	1	1	0.9908	1	1

ผลจากตารางแสดงด้วยกราฟของตัวสถิติทดสอบสำหรับการแจกแจงต่าง ๆ ดังนี้

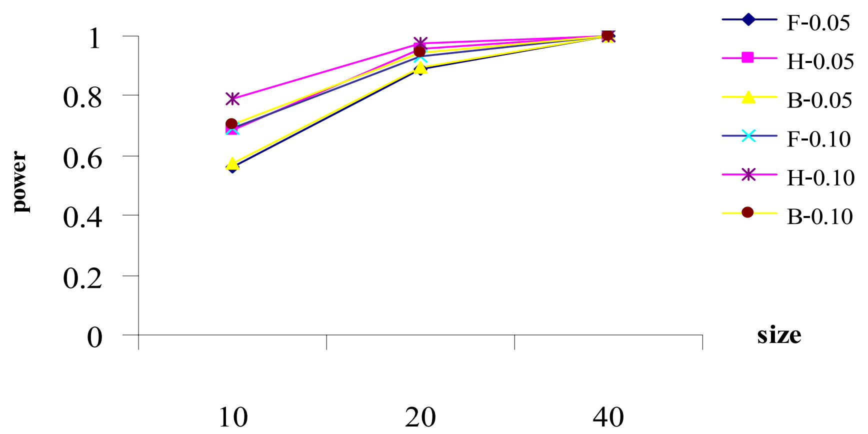


รูปที่ 67 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปกติ
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%



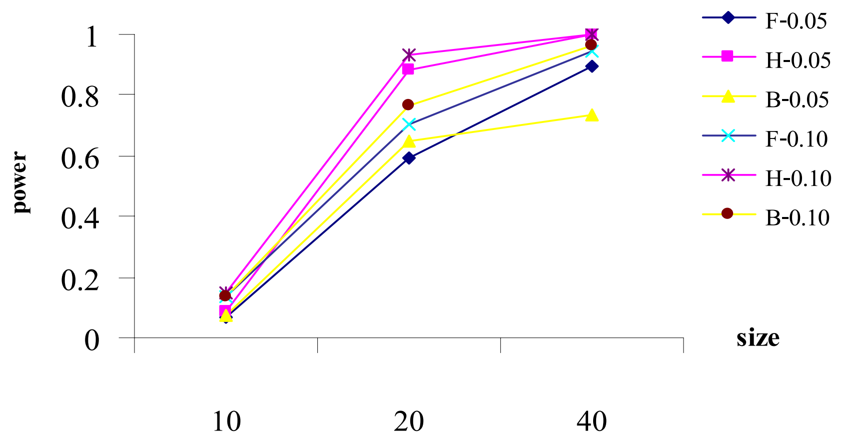
รูปที่ 68 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบยูนิฟอร์ม
กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 67 และ 68 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบปกติและการแจกแจงแบบยูนิฟอร์ม ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบที่ดีมากคือมีค่าเท่ากับ 1



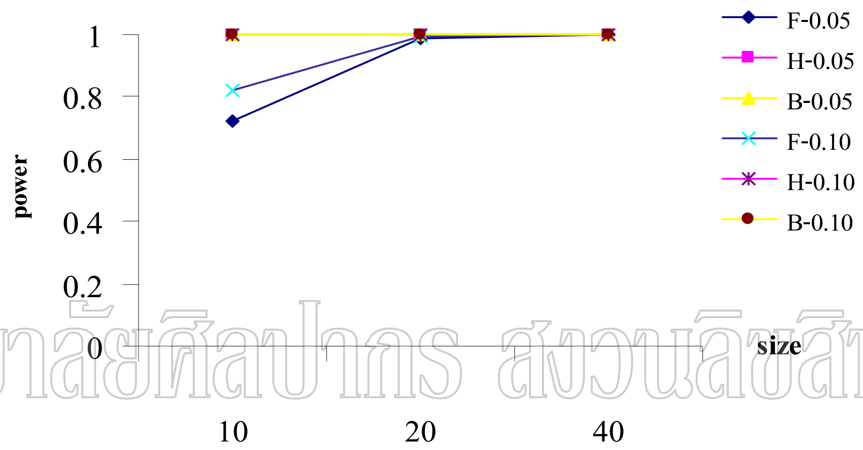
รูปที่ 69 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบไคสแควร์ กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 69 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบไคสแควร์ ที่ระดับนัยสำคัญ 0.05 และ 0.10 พบว่า ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ที่มีค่าใกล้เคียงกัน ซึ่งมีค่ามากขึ้นเมื่อขนาดตัวอย่างมากขึ้น



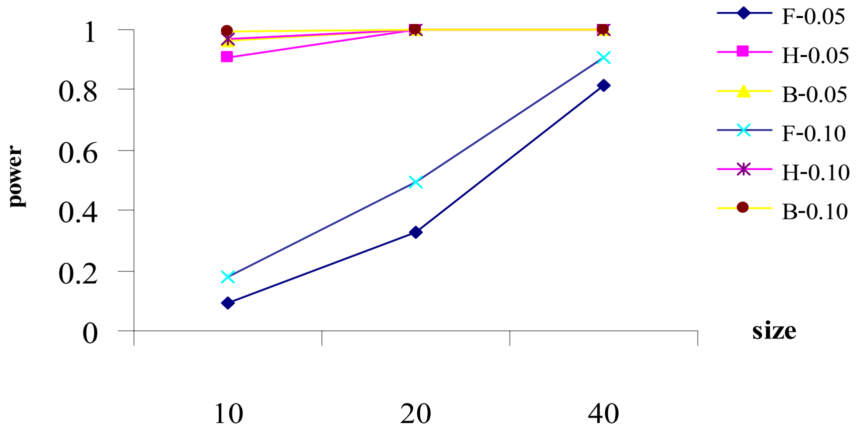
รูปที่ 70 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบเอกซ์โปเนนชียล กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 70 แสดงกราฟกำลังการทดสอบเมื่อตัวอย่างถูกสุ่มจากประชากรที่มีการแจกแจงแบบเอกซ์โปเนนชียล ที่ระดับนัยสำคัญ 0.05 เมื่อขนาดตัวอย่าง 10 และ 20 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ เมื่อขนาดตัวอย่าง 40 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ F และ Bi-aspect ตามลำดับ ที่ระดับนัยสำคัญ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ F ตามลำดับ



รูปที่ 71 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 10% กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 71 แสดงกราฟกำลังการทดสอบภายใต้การแจกแจงแบบปลอมปน 10% ที่ระดับนัยสำคัญ 0.05 และ 0.10 ตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีกำลังการทดสอบที่ดีมาก คือมีค่าเป็น 1 ส่วนตัวสถิติทดสอบ F มีกำลังการทดสอบมากขึ้นเมื่อขนาดตัวอย่างเพิ่มขึ้น



รูปที่ 72 กราฟกำลังการทดสอบ ภายใต้การแจกแจงแบบปลอมปน 30% กรณีค่าเฉลี่ยของกลุ่มถัดไปเพิ่มขึ้น 20%

จากรูปที่ 72 แสดงกราฟกำลังการทดสอบภายใต้การแจกแจงแบบปลอมปน 30% ที่ระดับนัยสำคัญ 0.05 และ 0.10 ตัวสถิติทดสอบ Bi-aspect มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ F ตามลำดับ โดยตัวสถิติทดสอบ F มีกำลังการทดสอบน้อยกว่าอย่างชัดเจน แต่จะมีกำลังการทดสอบเพิ่มขึ้นเมื่อขนาดตัวอย่างมากขึ้น

บทที่ 5

สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

สรุปผลการวิจัย

ในการศึกษาประสิทธิภาพ ของตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม ได้แก่ ตัวสถิติทดสอบ Bi- aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal-Wallis โดยทำการศึกษาจากความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เมื่อสมมติฐาน H_0 จริงและกำลังการทดสอบเมื่อสมมติฐาน H_1 จริง โดยกำหนดประชากรที่มีการแจกแจง 3 ลักษณะคือ การแจกแจงแบบสมมาตรได้แก่ การแจกแจงแบบปกติ และ ยูนิฟอร์ม การแจกแจงที่มีลักษณะเบ้ ได้แก่ การแจกแจงแบบไคสแควร์ และ เอกซ์โปเนนเชียล การแจกแจงแบบปลอมปน ได้แก่การแจกแจงแบบปกติปลอมปน 10% และ 30% กำหนดระดับนัยสำคัญของการทดสอบเท่ากับ 0.05 และ 0.10 จำลองแบบข้อมูลด้วยวิธีมอนติคาร์โล ด้วยโปรแกรม Fortran power station 4.0 การศึกษาได้กำหนดให้มีตัวอย่างสุ่มอย่างอิสระจากประชากรจำนวน 3 และ 5 กลุ่ม โดยกำหนดให้ขนาดตัวอย่างในแต่ละกลุ่มเป็น 10, 20 และ 40 สำหรับจำนวนซ้ำในการสุ่มตัวอย่างสำหรับการจำลองแบบคือ 4,000 ทำการศึกษาใน 2 กรณีคือ กรณีตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน และในกรณีตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกัน ซึ่งในกรณี H_1 จริงแบ่งออกเป็นกรณีย่อย 2 กรณีคือ กรณีที่มีค่าเฉลี่ยของประชากรเพียง 1 กลุ่มแตกต่างจากประชากรกลุ่มอื่น 5% และ 20% และกรณีที่ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน โดยตัวอย่างกลุ่มถัดไปมีค่าเฉลี่ยแตกต่างจากกลุ่มก่อนหน้า 5% และ 20% จากผลการศึกษาพบว่า

กรณีกลุ่มตัวอย่างถูกสุ่มอย่างอิสระจากประชากร 3 กลุ่ม ศึกษาใน 2 กรณี ดังนี้

กรณีที่ 1 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยเท่ากัน พิจารณาความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 พบว่าในการแจกแจงแบบปกติและยูนิฟอร์ม ตัวสถิติทดสอบทั้ง 3 ชนิดมีค่าเข้าใกล้ระดับนัยสำคัญที่ตั้งไว้ ส่วนในการแจกแจงแบบไคสแควร์ และเอกซ์โปเนนเชียล ตัวสถิติทดสอบ Kruskal-Wallis จะควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบตัวอื่น ๆ เล็กน้อย สำหรับการแจกแจงแบบปกติปลอมปน 10% และ 30% ตัวสถิติทดสอบ Kruskal-Wallis ควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ดีกว่าตัวสถิติทดสอบ Bi-aspect ส่วนตัวสถิติทดสอบ F ไม่สามารถควบคุมได้

กรณีที่ 2 ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มซึ่งมีค่าเฉลี่ยแตกต่างกัน พิจารณาจากกำลังการทดสอบในกรณีย่อย 2 กรณี คือ

กรณีที่ 2.1 มีค่าเฉลี่ยของประชากรเพียง 1 กลุ่มแตกต่างจากประชากรกลุ่มอื่น

โดยศึกษาในกรณีค่าเฉลี่ยของประชากรกลุ่มหนึ่งเพิ่มขึ้น 5% และ 20% พบว่า ในกรณีค่าเฉลี่ยเพิ่มขึ้น 5 % การแจกแจงแบบปกติและยูนิฟอร์ม ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect ตามลำดับ การแจกแจงแบบไคสแควร์ และ เอกซ์โปเนนเชียล ตัวสถิติทดสอบทั้ง 3 ชนิดมีกำลังการทดสอบใกล้เคียงกันแต่ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และตัวสถิติทดสอบ F ตามลำดับ ส่วนการแจกแจงแบบปลอมปน 10% และ 30% ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และตัวสถิติทดสอบ F มีกำลังการทดสอบน้อยมาก สำหรับในกรณีค่าเฉลี่ยของประชากรกลุ่มหนึ่งแตกต่างจากกลุ่มอื่นเพิ่มขึ้น 20 % ได้ผลในทำนองเดียวกัน โดยในการแจกแจงแบบปกติและยูนิฟอร์ม มีกำลังการทดสอบที่ดีมากคือมีค่าเป็น 1 เมื่อขนาดตัวอย่างมากขึ้นเป็น 40 การแจกแจงแบบไคสแควร์ และ เอกซ์โปเนนเชียล ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบที่ดีกว่าตัวสถิติทดสอบ Bi-aspect และตัวสถิติทดสอบ F ซึ่งมีกำลังการทดสอบใกล้เคียงกันแต่ตัวสถิติทดสอบ Bi-aspect ดีกว่าเล็กน้อย ส่วนการแจกแจงแบบปลอมปน 10% และ 30% ตัวสถิติทดสอบ Kruskal-Wallis มีกำลังการทดสอบใกล้เคียงกับตัวสถิติทดสอบ Bi-aspect แต่ดีกว่าเล็กน้อย ในขณะที่ตัวสถิติทดสอบ F มีกำลังการทดสอบต่ำมาก

กรณีที่ 2.2 กรณีที่ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน โดยตัวอย่างกลุ่มถัดไปมีค่าเฉลี่ยแตกต่างจากกลุ่มก่อนหน้า 5% และ 20% พบว่า กำลังการทดสอบของตัวสถิติทดสอบทั้ง 3 ชนิดเพิ่มขึ้นในทุกการแจกแจง สำหรับกรณีค่าเฉลี่ยแต่ละกลุ่มแตกต่างกัน 5% การแจกแจงแบบปกติ และ ยูนิฟอร์ม ตัวสถิติทดสอบ F มีกำลังการทดสอบสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect ตามลำดับ การแจกแจงแบบไคสแควร์และเอกซ์โปเนนเชียล ตัวสถิติทดสอบ Kruskal-Wallis มีค่าสูงสุด รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และตัวสถิติทดสอบ F มีกำลังการทดสอบใกล้เคียงกันแต่ตัวสถิติทดสอบ Bi-aspect ดีกว่าเล็กน้อย ส่วนการแจกแจงแบบปลอมปน 10% และ 30% ตัวสถิติทดสอบ Kruskal-Wallis และ Bi-aspect มีค่าใกล้เคียงกันแต่ตัวสถิติทดสอบ Kruskal-Wallis ดีกว่าเล็กน้อย ในขณะที่ตัวสถิติทดสอบ F มีกำลังการทดสอบต่ำมาก สำหรับในกรณีค่าเฉลี่ยของประชากรเพิ่มขึ้น 20 % ได้ผลในทำนองเดียวกัน

กรณีกลุ่มตัวอย่างถูกสุ่มอย่างอิสระจากประชากร 5 กลุ่ม

พบว่า จากการศึกษาสามารถสรุปผลได้ในทำนองเดียวกันกับกรณีกลุ่มตัวอย่างถูกสุ่มอย่างอิสระจากประชากร 3 กลุ่ม โดยความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ในกรณีศึกษาเดียวกันมีค่าไม่แตกต่างกันมากนัก และกำลังการทดสอบในกรณีประชากร 5 กลุ่มโดยส่วนใหญ่มีค่าสูงกว่า สามารถแสดงผลสรุปเป็นตารางได้ดังตารางที่ 18 (หน้า 84)

อภิปรายผล

จากผลการวิจัยพบว่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และกำลังการทดสอบของตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal-Wallis ขึ้นอยู่กับปัจจัยหลายอย่างดังนี้

ลักษณะการแจกแจงของประชากร เมื่อประชากรมีลักษณะแตกต่างกันพบว่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และกำลังการทดสอบจะแตกต่างกันด้วย นั่นคือ เมื่อประชากรมีการแจกแจงแบบสมมาตรตัวสถิติทดสอบ F มีประสิทธิภาพสูงสุด เนื่องจากตัวสถิติทดสอบ F เป็นตัวสถิติสำหรับทดสอบค่าเฉลี่ย รองลงมาเป็นตัวสถิติทดสอบ Kruskal-Wallis และตัวสถิติทดสอบ Bi-aspect ตามลำดับ ส่วนการแจกแจงที่มีลักษณะเบ้ตัวสถิติทดสอบ Kruskal-Wallis มีประสิทธิภาพสูงสุด เนื่องจากตัวสถิติทดสอบ Kruskal-Wallis เป็นตัวสถิติทดสอบค่ามัธยฐานซึ่งเหมาะกับการแจกแจงที่มีลักษณะเบ้ รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect และ ตัวสถิติทดสอบ F ตามลำดับ สำหรับการแจกแจงแบบปกติปลอมปนตัวสถิติทดสอบ Kruskal-Wallis มีประสิทธิภาพสูงสุดเช่นเดียวกัน รองลงมาเป็นตัวสถิติทดสอบ Bi-aspect ส่วนตัวสถิติทดสอบ F มีประสิทธิภาพน้อยมาก

ขนาดตัวอย่างที่สุ่มมาอย่างอิสระ เมื่อขนาดตัวอย่างเพิ่มขึ้นพบว่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เข้าใกล้ระดับนัยสำคัญที่ตั้งไว้มากขึ้น และมีกำลังการทดสอบมากขึ้นสำหรับตัวสถิติทดสอบทั้ง 3 ชนิด

ข้อเสนอแนะ

สำหรับงานวิจัยนี้ ได้ทำการศึกษาเปรียบเทียบประสิทธิภาพของตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่มสำหรับประชากรที่มีการแจกแจง 5 การแจกแจงเท่านั้น จึงน่าสนใจที่จะศึกษาในกรณีที่ประชากรมีการแจกแจงลักษณะอื่น ๆ อีก หรืออาจศึกษาในกรณีที่ตัวอย่างถูกสุ่มจากประชากรแต่ละกลุ่มมีขนาดไม่เท่ากัน

ตารางที่ 18 ผลสรุปของการศึกษาตัวสถิติทดสอบ Bi-aspect F และ Kruskal-Wallis เมื่อประชากรมีลักษณะการแจกแจงแตกต่างกัน

กลุ่มตัวอย่าง	ลักษณะการแจกแจงของประชากร	ความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ประชากรทุกกลุ่มมีค่าเฉลี่ยเท่ากัน	กำลังการทดสอบ					
			มีค่าเฉลี่ยเพียง 1 กลุ่มแตกต่างจากกลุ่มอื่น			ค่าเฉลี่ยของประชากรทุกกลุ่มแตกต่างกัน		
			อันดับ 1	อันดับ 2	อันดับ3	อันดับ 1	อันดับ 2	อันดับ 3
3	สมมาตร	มีระดับนัยสำคัญใกล้เคียงกับระดับนัยสำคัญที่กำหนด	F	Kruskal-Wallis	Bi-aspect	F	Kruskal-Wallis	Bi-aspect
	เบ้	Kruskal-Wallis ควบคุมได้ดีกว่าตัวสถิติทดสอบอื่น ๆ	Kruskal-Wallis	Bi-aspect	F	Kruskal-Wallis	*Bi-aspect*	*F*
	ปลอมปน	Kruskal-Wallis ควบคุมดีสุด Bi-aspect รองลงมา F น้อยมาก	Kruskal-Wallis	Bi-aspect	F	*Kruskal-Wallis*	*Bi-aspect*	F
5	สมมาตร	มีระดับนัยสำคัญใกล้เคียงกับระดับนัยสำคัญที่กำหนด	F	Kruskal-Wallis	Bi-aspect	1	1	1
	เบ้	Kruskal-Wallis ควบคุมได้ดีกว่าตัวสถิติทดสอบอื่น ๆ	Kruskal-Wallis	*Bi-aspect*	*F*	Kruskal-Wallis	*Bi-aspect*	*F*
	ปลอมปน	Kruskal-Wallis ควบคุมดีสุด Bi-aspect รองลงมา F น้อยมาก	*Kruskal-Wallis*	*Bi-aspect*	F	*Kruskal-Wallis*	*Bi-aspect*	F

หมายเหตุ *อักษรเอียง* หมายถึง มีกำลังการทดสอบใกล้เคียงกันในลักษณะประชากรเดียวกัน

บรรณานุกรม

กมลทิพย์ ปรัชญชรินกร. “การเปรียบเทียบตัวสถิติทดสอบความแตกต่างของค่าเฉลี่ยของประชากรหลายกลุ่ม.” วิทยานิพนธ์ปริญญามหาบัณฑิต สาขาวิชาสถิติ บัณฑิตวิทยาลัย มหาวิทยาลัยจุฬาลงกรณ์, 2541.

ไพบูลย์ รัตนประเสริฐ. ระเบียบวิธีสถิติ. นครปฐม : โรงพิมพ์มหาวิทยาลัยศิลปากร ,2543.

สุชาดา กีระนันทน์. หลักสถิติ. กรุงเทพ : โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย , 2532.

Conover, W.J. Practical Nonparametric Statistics. Newyork : John Wiley, 1980.

Good, P. Permutation Tests, a Practical Guide to Resampling Methods for TestingHypotheses. New York : Springer-Verlag, 1994.

Jean Dickinson Gibbons, Ph.D. Nonparametric Statistical Inference. New York : McGraw-Hill Book Company, 1971.

Johnson, N.L., and S.Kotz. Continuous Univariate Distribution-1 . Boston : Houghton Mifflin, 1970.

Kenneth J. Berry. “Enumeration of All Permutations of Multi-Sets with Fixed Repetition Numbers.” Journal of the Royal Statistical Society (1982) : 169-173.

Lehmann E. L. Nonparametrics statistical Methods Based on Ranks. San Francisco : Holden - day, 1975.

___________. Testing Statistical Hypotheses. 2nd Edition. New York : Wiley , 1986.

Ludbrook, J., and H. Dudley. “Why permutation tests are superior to t and F tests in Biomedical research.” The American Statistician 52(1998) : 127-132.

Marozzi , M. “A bi-aspect nonparametric test for the two-sample location problem.” Computation statistics and data analysis 44(2003) : 639-648.

___________. “A bi-aspect nonparametric test for the multi-sample location problem.” Computation statistics and data analysis 46(2004) : 81-92.

Michael A. Fligner , and George E. Policella. “Robust Rank Procedures for the Behrens-fisher Problem.” Journal of the Statistical Association 76(1981) : 162-168.

Pitman , E.J.G. “Significance tests which may be applied to samples from any population” Journal of the Royal Statistical Society Ser. B 4(1937a) : 119-130.

Pitman , E.J.G. "Significance tests which may be applied to samples from any population.II. The correlation coefficient." Journal of the Royal Statistical Society Ser. B . 4(1937b) : 225-232.

__________ . "Significance tests which may be applied to samples from any population.III. The analysis of variance test." Biometika 29(1938) : 322-335.

Randles, Ronald H. Introduction to the theory of nonparametric statistics. New York : John Wiley, 1979.

Sheskin, David J. Handbook of parametric and nonparametric statistical procedures. Boca Raton : Fla, CRC Press, 2000.

Wayne W. Daniel. Applied Nonparametric Statistics. London : New Jersey, 1978.

ภาคผนวก

โปรแกรมคอมพิวเตอร์ที่ใช้ในงานวิจัย

โปรแกรมคอมพิวเตอร์ สำหรับคำนวณตัวสถิติทดสอบความแตกต่างของค่ากลางสำหรับประชากรหลายกลุ่ม คือตัวสถิติทดสอบ Bi-aspect ตัวสถิติทดสอบ F และตัวสถิติทดสอบ Kruskal-Wallis เขียนโดยใช้คำสั่งภาษา FORTRAN

โปรแกรมหลักประกอบด้วยตัวแปรดังนี้

1. ตัวแปรในโปรแกรมหลัก
2. ตัวแปรในโปรแกรมย่อย

ตัวแปรในโปรแกรมหลัก ประกอบด้วย

SZ	แทนขนาดตัวอย่างสุ่ม
B0	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.05
B1	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.10
X	แทนตัวอย่างสุ่ม
ZEED	แทนค่าใด ๆ ที่สุ่มมาจากการแจกแจงแบบเอกรูป มีค่าระหว่าง 0 ถึง 1
LDSTAT	แทนขนาดมิติของเมตริกซ์ผลลัพธ์
NGROUP	แทนจำนวนกลุ่มตัวอย่างตัวสถิติทดสอบเอฟ
NOBS	แทนขนาดตัวอย่างสุ่ม
IPRINT	แทนการแสดงผลลัพธ์
NI	แทนขนาดของเวกเตอร์ตัวอย่างสุ่มแต่ละกลุ่ม
NMISS	แทนจำนวนค่าสูญหาย
AOV	แทนเวกเตอร์ผลลัพธ์ของตัวสถิติทดสอบเอฟ
STAT	แทนเมตริกซ์ผลลัพธ์ของตัวสถิติทดสอบเอฟ
XOBS	แทนตัวอย่างสุ่มที่ใช้ในการคำนวณค่าสถิติ
FUZZ	แทนเกณฑ์ที่ใช้ตัดสินลำดับของข้อมูลเมื่อข้อมูลที่มีค่าเท่ากัน
NGROUP1	แทนจำนวนกลุ่มตัวอย่างสุ่มของตัวสถิติทดสอบครัสคัล-วอลลิส
STAT1	แทนเมตริกซ์ผลลัพธ์ของตัวสถิติทดสอบครัสคัล-วอลลิส
PTAI	แทนค่า p-value ของตัวสถิติทดสอบ Bi-aspect
ROBS1	แทนตัวอย่างสุ่มกลุ่มที่ 1
ROBS2	แทนตัวอย่างสุ่มกลุ่มที่ 2

ROBS3	แทนตัวอย่างสุ่มกลุ่มที่ 3
ROBS4	แทนตัวอย่างสุ่มกลุ่มที่ 4
ROBS5	แทนตัวอย่างสุ่มกลุ่มที่ 5
REF	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.05
REH	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.05
REB	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.05
PF	แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.05
PH	แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.05
PB	แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.05
H0	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.05
F0	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.05
REF1	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.10
REH1	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.10
REB1	แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.10
PF1	แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.10
PH1	แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.10

PB1 แทนความน่าจะเป็นที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.10

H1 แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบครัสคัล-วอลลิส เมื่อกำหนดระดับนัยสำคัญ 0.10

F1 แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบเอฟ เมื่อกำหนดระดับนัยสำคัญ 0.10

ตัวแปรในโปรแกรมย่อย ประกอบด้วย

N แทนขนาดตัวอย่างสุ่ม

POS แทนตำแหน่งกลางของข้อมูล

K แทนตำแหน่งของตัวอย่างสุ่ม

N1 แทนขนาดตัวอย่างสุ่มกลุ่มที่ 1

IPATH แทนลักษณะการเรียงสับเปลี่ยนข้อมูล

IPERMU แทนเวกเตอร์ของตัวอย่างสุ่ม

B0 แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.05

B1 แทนจำนวนครั้งที่ปฏิเสธสมมติฐานว่างตัวสถิติทดสอบ Bi-aspect เมื่อกำหนดระดับนัยสำคัญ 0.10

XPERMU แทนเวกเตอร์ของตัวอย่างสุ่มที่ทำการจัดกลุ่มใหม่

XOBS แทนตัวอย่างสุ่มที่ใช้ในการคำนวณค่าสถิติ

X แทนตัวอย่างสุ่ม

S แทนตัวแปรสำหรับการเปลี่ยนค่า

MD แทนมัธยฐานของตัวอย่างสุ่ม

N1G0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 1 ที่มีค่ามากกว่ามัธยฐานร่วม

N1S0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 1 ที่มีค่าน้อยกว่ามัธยฐานร่วม

N2G0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 2 ที่มีค่ามากกว่ามัธยฐานร่วม

N2S0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 2 ที่มีค่าน้อยกว่ามัธยฐานร่วม

N3G0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 3 ที่มีค่ามากกว่ามัธยฐานร่วม

N3S0 แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 3 ที่มีค่าน้อยกว่ามัธยฐานร่วม

N4G0	แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 4 ที่มีค่ามากกว่ามัธยฐานร่วม
N4S0	แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 4 ที่มีค่าน้อยกว่ามัธยฐานร่วม
N5G0	แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 5 ที่มีค่ามากกว่ามัธยฐานร่วม
N5S0	แทนจำนวนตัวอย่างสุ่มกลุ่มที่ 5 ที่มีค่าน้อยกว่ามัธยฐานร่วม
TA01	แทนค่าที่ได้จากนำค่าเฉลี่ยยกกำลังสองคูณด้วยจำนวนตัวอย่างสุ่มกลุ่ม 1
TA02	แทนค่าที่ได้จากนำค่าเฉลี่ยยกกำลังสองคูณด้วยจำนวนตัวอย่างสุ่มกลุ่ม 2
TA03	แทนค่าที่ได้จากนำค่าเฉลี่ยยกกำลังสองคูณด้วยจำนวนตัวอย่างสุ่มกลุ่ม 3
TA04	แทนค่าที่ได้จากนำค่าเฉลี่ยยกกำลังสองคูณด้วยจำนวนตัวอย่างสุ่มกลุ่ม 4
TA05	แทนค่าที่ได้จากนำค่าเฉลี่ยยกกำลังสองคูณด้วยจำนวนตัวอย่างสุ่มกลุ่ม 5
TB01	แทนค่าที่ได้จากนำจำนวนตัวอย่างสุ่มที่มากกว่ามัธยฐานร่วมยกกำลังสองหารด้วยจำนวนตัวอย่างกลุ่มที่ 1
TB02	แทนค่าที่ได้จากนำจำนวนตัวอย่างสุ่มที่มากกว่ามัธยฐานร่วมยกกำลังสองหารด้วยจำนวนตัวอย่างกลุ่มที่ 2
TB03	แทนค่าที่ได้จากนำจำนวนตัวอย่างสุ่มที่มากกว่ามัธยฐานร่วมยกกำลังสองหารด้วยจำนวนตัวอย่างกลุ่มที่ 3
TB04	แทนค่าที่ได้จากนำจำนวนตัวอย่างสุ่มที่มากกว่ามัธยฐานร่วมยกกำลังสองหารด้วยจำนวนตัวอย่างกลุ่มที่ 4
TB05	แทนค่าที่ได้จากนำจำนวนตัวอย่างสุ่มที่มากกว่ามัธยฐานร่วมยกกำลังสองหารด้วยจำนวนตัวอย่างกลุ่มที่ 5
TA	แทนตัวสถิติทดสอบ T_A
TB	แทนตัวสถิติทดสอบ T_B
BS	แทนตัวสถิติทดสอบ T_{AB}

โปรแกรมหลักประกอบด้วยโปรแกรมย่อย ดังนี้

RNSET	โปรแกรมย่อยสำหรับกำหนดจำนวนเริ่มต้น
RNNOR	โปรแกรมย่อยสำหรับสร้างข้อมูลตัวอย่างที่มาจากประชากรที่มาจากการแจกแจงแบบปกติมาตรฐาน
AONEW	โปรแกรมย่อยสำหรับคำนวณตัวสถิติทดสอบความแปรปรวนทางเดียว
KRSKL	โปรแกรมย่อยสำหรับคำนวณตัวสถิติทดสอบครัสคัล-วอลลิส
BIASPECT	โปรแกรมย่อยสำหรับคำนวณตัวสถิติทดสอบ Bi-aspect

โปรแกรมหลัก

```
            USE  MSIMSL
             INTEGER          I, SZ ,B0,B1
             REAL             X(50) ,ZEED
             INTEGER          LDSTAT, NGROUP, NOBS
             PARAMETER  (NGROUP=5, NOBS=50, LDSTAT=NGROUP)
            INTEGER           IPRINT, NI(NGROUP), NMISS
            REAL              AOV(15), STAT(LDSTAT,4), XOBS(NOBS)
            REAL              FUZZ
            PARAMETER  (FUZZ=0.001, NGROUP1=5)
            INTEGER           NGROUP1
            REAL               STAT1(4)
            REAL              PTAI(1)
            REAL              ROBS1(10),ROBS2(10),ROBS3(10),ROBS4(10),ROBS5(10)
            REAL              REF ,REH,REB,PF,PH,PB
            INTEGER           H0,F0
            REAL              REF1 ,REH1,REB1,PF1,PH1,PB1
            INTEGER           H1,F1
            OPEN (37,FILE='.\RETEST.TXT')
            OPEN (38,FILE='.\TOTAL.TXT')
            OPEN (35,FILE='.\XOBS.TXT')
            OPEN (39,FILE='.\TOL01.TXT')
*******************************************************************


DO T=1,4000
              AVE=10
              SZ=50
            DO 10  I=1, 10
              CALL RANDOM_NUMBER(ZEED)
              ISEED=ZEED*2147483646*T
```

```
          CALL RNSET (ISEED)
          CALL RNNOR (10,X)
          ROBS1(I) = X(I)+AVE
10      CONTINUE
        DO  11  I=1, 10
          CALL RANDOM_NUMBER(ZEED)
          ISEED=ZEED*2147483646*T
          CALL RNSET (ISEED)
          CALL RNNOR (10,X)
          ROBS2(I) = X(I)+AVE
11      CONTINUE
        DO  12  I=1, 10
          CALL RANDOM_NUMBER(ZEED)
          ISEED=ZEED*2147483646*T
          CALL RNSET (ISEED)
          CALL RNNOR (10,X)
          ROBS3(I) = X(I)+AVE
12      CONTINUE
        DO  13  I=1, 10
          CALL RANDOM_NUMBER(ZEED)
          ISEED=ZEED*2147483646*T
          CALL RNSET (ISEED)
          CALL RNNOR (10,X)
          ROBS4(I) = X(I)+AVE
13      CONTINUE
        DO  14  I=1, 10
          CALL RANDOM_NUMBER(ZEED)
          ISEED=ZEED*2147483646*T
          CALL RNSET (ISEED)
          CALL RNNOR (10,X)
```

```
            ROBS5(I) = X(I)+AVE
14        CONTINUE
            DO  J=1,50
                  IF(L.LE.J.AND.J.LE.10)  THEN
                  XOBS(J)=ROBS1(J)
                  ELSE  IF  (11.LE.J.AND.J.LE.20)  THEN
                  XOBS(J)=ROBS2(J-10)
                  ELSE  IF  (21.LE.J.AND.J.LE.30)  THEN
                  XOBS(J)=ROBS3(J-20)
                  ELSE  IF  (31.LE.J.AND.J.LE.40)  THEN
                  XOBS(J)=ROBS4(J-30)
                  ELSE
                  XOBS(J)=ROBS5(J-40)
                  ELSE  IF
            END DO
            WRITE (35,959)  XOBS
959     FORMAT(10F8.4)

****************************************************************

        DATA NI/10, 10, 10,10,10/
        IPRINT = 3
        CALL AONEW (NGROUP, NI, XOBS, IPRINT, AOV, STAT, LDSTAT, NMISS)
        CALL KRSKL (NGROUP1, NI, XOBS, FUZZ, STAT1)
        CALL BIASPECT (XOBS,PTAI,B0,B1)
            IF(AOV(10).LE.0.05) THEN
                  F0=F0+1
            END IF
            IF(STAT1(4).LE.0.05) THEN
                  H0=H0+1
```

```
          END IF
          IF(AOV(10).LE.0.1) THEN
                F1=F1+1
          END IF
          IF(STAT1(4).LE.0.1) THEN
                H1=H1+1
          END IF
ENDDO
          REF=F0
          REH=H0
          REB=B0
          PF=REF/4000
          PH=REH/4000
          PB=REB/4000
          REF1=F1
          REH1=H1
          REB1=B1
          PF1=REF1/4000
          PH1=REH1/4000
          PB1=REB1/4000
          WRITE (38,989)  PF,PH,PB
989     FORMAT(3F8.4)
          WRITE (39,989)  PF1,PH1,PB1
999     FORMAT(3F8.4)
        END
```

```
**************************************************************
            CALCULATE   BI-ASPECT    TEST
**************************************************************
  SUBROUTINE      BIASPECT(XOBS,PTAI,B0,B1)
           IMPLICIT   NONE
   INTEGER    I,N,J,POS,K ,N1,L
   INTEGER    IPATH,R
   PARAMETER  (IPATH=1)
   INTEGER    IPERMU(50)    ,B0,B1
   REAL          XPERMU(50),XOBS(50)
   REAL         X(50),S,MD
   REAL         N1G0,N1S0,N2G0,N2S0,N3G0,N3S0,N4G0,N4S0,N5G0,N5S0
   REAL         TA01,TA02,TA03,TA04,TA05
   REAL         TB01,TB02,TB03,TB04,TB05
   REAL         TA(1001),TB(1001)
   REAL          BS(1001)
   REAL          PTAI(1),PTA0,PTB0,PVTA0,PVTB0
   REAL          SUM1, SUM2, SUM3, SUM4, SUM5, XBAR1, XBAR2,  XBAR3
                XBAR4,  XBAR5
   REAL          TAI
            N=50
            N1=10
            POS=N/2

  DO  L=1,1001
      IF(L==1)  THEN
            X = XOBS
      ELSE
            X = XPERMU
      END  IF
```

```
    SUM1=0
    SUM2=0
    SUM3=0
    SUM4=0
    SUM5=0
    XBAR1=0
    XBAR2=0
    XBAR3=0
    XBAR4=0
    XBAR5=0
    N1G0=0
    N1S0=0
    N2G0=0
    N2S0=0
    N3G0=0
    N3S0=0
    N4G0=0
    N4S0=0
    N5G0=0
    N5S0=0

DO I=1,10
    SUM1=SUM1+X(I)
     XBAR1=SUM1/10
END DO
DO I=11,20
    SUM2=SUM2+X(I)
   XBAR2=SUM2/10
END DO
```

```
DO I=21,30
     SUM3=SUM3+X(I)
     XBAR3=SUM3/10
END DO
DO I=31,40
     SUM4=SUM4+X(I)
     XBAR4=SUM4/10
END DO
DO I=41,50
     SUM5=SUM5+X(I)
     XBAR5=SUM5/10
END DO

*******************************************************************
      DO 50 I=1,N-1
          DO 50 J=1,N-I
               IF(X(J).GT.X(J+1))THEN
                    S=X(J)
                    X(J)=X(J+1)
                    X(J+1)=S
               END IF
50    CONTINUE

*********************************************************************

      IF(MOD(N,2)==0)THEN
          MD=(X(POS)+X(POS+1))/2
      ELSE
          MD=X(POS+1)
      END IF
```

```
        IF(L==1)  THEN
                X = XOBS
        ELSE
                X = XPERMU
        END IF
        DO 45 K=1,10
            IF(X(K).GT.MD)THEN
                N1G0=N1G0+1
              ELSE
                N1S0=N1S0+1
            END IF
45      CONTINUE
         DO 46 K=11,20
            IF(X(K).GT.MD)THEN
                N2G0=N2G0+1
            ELSE
                N2S0=N2S0+1
         END IF
46       CONTINUE
         DO 47 K=21,30
           IF(X(K).GT.MD)THEN
                N3G0=N3G0+1
           ELSE
                N3S0=N3S0+1
           END IF
47       CONTINUE
          DO 48  K=31,40
           IF(X(K).GT.MD)THEN
                N4G0=N4G0+1
           ELSE
```

```
                N4S0=N4S0+1
         END  IF
48      CONTINUE
         DO 49 K=41,50
             IF(X(K).GT.MD)THEN
                   N5G0=N5G0+1
             ELSE
                   N5S0=N5S0+1
             END IF
49      CONTINUE

*******************************************************************

         DO  I=1,10
              TA01=N1*(XBAR1**2)
              TB01=(N1G0**2)/N1
         END DO
         DO  I=11,20
              TA02=N1*(XBAR2**2)
              TB02=(N2G0**2)/N1
         END DO
         DO  I=21,30
              TA03=N1*(XBAR3**2)
              TB03=(N3G0**2)/N1
         END DO
         DO  I=31,40
              TA04=N1*(XBAR4**2)
              TB04=(N4G0**2)/N1
         END DO
```

```
      DO  I=41,50
          TA05=N1*(XBAR5**2)
          TB05=(N5G0**2)/N1
      END DO
         TA(l)=TA01+TA02+TA03+TA04+TA05
         TB(l)=TB01+TB02+TB03+TB04+TB05
        CALL RNPER(N,IPERMU)
        CALL PERMU(N,XOBS,IPERMU,IPATH,XPERMU)
      END   DO

*****************************************************************

   DO  16  R=1,1001
              PTA0=0
              PTB0=0
         DO  15  L=2,1001
              IF  (TA(L).GE.TA(R))  THEN
                    PTA0=PTA0+1
              END  IF
              IF  (TB(L).GE.TB(R))  then
                    PTB0=PTB0+1
              END  IF
15        CONTINUE
              PVTA0=  (0.5+PTA0)/1001
              PVTB0=  (0.5+PTB0)/1001
              BS(R)=MAX(1-PVTA0,1-PVTB0)
 16       CONTINUE
           TAI=0
   DO   25   R=2,1001
```

```
                IF  (BS(R).GE.BS(L))  THEN
                      TAI=TAI+1
                END IF
25        CONTINUE
                PTAI=TAI/1000
          IF(PTAI(1).LE.0.05) THEN
                B0=B0+1
          END IF
          IF(PTAI(L).LE.0.1) THEN
                B1=B1+1
          END IF
                RETURN
END
```

ประวัติผู้วิจัย

ชื่อ – สกุล	นางสาวพวงทิพย์ พูลสวัสดิ์
ที่อยู่	122/41 หมู่ 18 ตำบลประสงค์ อำเภอท่าชนะ จังหวัดสุราษฎร์ธานี 84170

ประวัติการศึกษา

พ.ศ. 2541	สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิตวิชาเอกคณิตศาสตร์ มหาวิทยาลัยทักษิณ
พ.ศ. 2545	ศึกษาต่อระดับปริญญามหาบัณฑิต สาขาสถิติประยุกต์ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ประวัติการทำงาน

พ.ศ. 2542 – พ.ศ. 2544	อาจารย์ โรงเรียนเซนต์โยเซฟเกาะสมุย จังหวัดสุราษฎร์ธานี
พ.ศ. 2544 – พ.ศ. 2545	อาจารย์ โรงเรียนสมุยบริหารธุรกิจ จังหวัดสุราษฎร์ธานี