

การเปรียบเทียบอำนาจของวิธีการทดสอบการแจกแจงแบบปกติจากการจำลองข้อมูลด้วยวิธีเพาเวอร์



บทคัดย่อและแฟ้มข้อมูลฉบับเต็มของวิทยานิพนธ์ตั้งแต่ปีการศึกษา 2554 ที่ให้บริการในคลังปัญญาจุฬาฯ (CUIR)
เป็นแฟ้มข้อมูลของนิสิตเจ้าของวิทยานิพนธ์ ที่ส่งผ่านทางบัณฑิตวิทยาลัย

The abstract and full text of theses from the academic year 2011 in Chulalongkorn University Intellectual Repository (CUIR)
are the thesis authors' files submitted through the University Graduate School.

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาสถิติ ภาควิชาสถิติ

คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2558

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A POWER COMPARISON OF NORMALITY TESTS VIA POWER METHOD SIMULATION

Miss Siraporn Khemthit



A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science Program in Statistics

Department of Statistics

Faculty of Commerce and Accountancy

Chulalongkorn University

Academic Year 2015

Copyright of Chulalongkorn University

ศิราภรณ์ เข้มทิต : การเปรียบเทียบอำนาจของวิธีการทดสอบการแจกแจงแบบปกติจาก การจำลองข้อมูลด้วยวิธีเพาเวอร์ (A POWER COMPARISON OF NORMALITY TESTS VIA POWER METHOD SIMULATION) อ.ที่ปรึกษาวิทยานิพนธ์หลัก: ผศ. ดร. วิรุธา พึ่งพาพงศ์, 73 หน้า.

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อเปรียบเทียบอำนาจการทดสอบของวิธีการทดสอบการแจกแจงแบบปกติ 3 วิธี ได้แก่ วิธีการทดสอบชาปิโร-วิลค์ วิธีไรอัน-จอยเนอร์ และวิธีซุรุกุซี โดยการหาค่าพี-แวลูสำหรับวิธีการทดสอบไรอัน-จอยเนอร์และวิธีการทดสอบซุรุกุซีนั้นจะใช้วิธีบูตสเตรป สำหรับการเปรียบเทียบอำนาจการทดสอบนั้นผู้วิจัยใช้ข้อมูลจำลองด้วยวิธีเพาเวอร์ซึ่งเป็นวิธีจำลองข้อมูลที่ง่ายและสามารถจำลองข้อมูลที่มีระดับความเบ้และความโด่งที่ต้องการ ในการศึกษานี้จะแบ่งกรณีศึกษาออกเป็น 4 กลุ่ม คือกลุ่มการแจกแจงลักษณะสมมาตรและโด่งต่ำ กลุ่มการแจกแจงลักษณะสมมาตรและโด่งสูง กลุ่มการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ และกลุ่มการแจกแจงลักษณะไม่สมมาตรและโด่งสูง และมีขนาดตัวอย่างแตกต่างกันตั้งแต่ $n=10, 20, 30, 50$ และ 100

จากการศึกษาพบว่า วิธีการทดสอบ ชาปิโร-วิลค์ ให้อำนาจการทดสอบสูงสำหรับทุกกลุ่มเมื่อเปรียบเทียบกับวิธีการทดสอบไรอัน-จอยเนอร์และซุรุกุซีในทุกกรณีศึกษา นอกจากนี้ยังพบว่าอำนาจการทดสอบมีค่ามากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น ดังนั้นจึงสรุปได้ว่าวิธีชาปิโร-วิลค์เป็นวิธีที่มีอำนาจการทดสอบสูงสุดเมื่อเทียบกับวิธีไรอัน-จอยเนอร์และวิธีซุรุกุซี และการหาค่าพี-แวลูยังสามารถทำได้ง่ายกว่าเนื่องจากตัวสถิติชาปิโร-วิลค์นั้นมีการแจกแจงอย่างชัดเจน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ภาควิชา สถิติ

ลายมือชื่อนิสิต

สาขาวิชา สถิติ

ลายมือชื่อ อ.ที่ปรึกษาหลัก

ปีการศึกษา 2558

5781586826 : MAJOR STATISTICS

KEYWORDS: NORMALITY TESTING / SHAPIRO-WILK TEST / RYAN-JOINER TEST / SURUCU "C" TEST / POWER METHOD / BOOTSTRAP

SIRAPORN KHEMTHIT: A POWER COMPARISON OF NORMALITY TESTS VIA POWER METHOD SIMULATION. ADVISOR: ASST. PROF. VITARA PUNGPAPONG, Ph.D., 73 pp.

This research aims to compare power of the test for three normality tests including Shapiro-Wilk test, Ryan-Joiner test, and Sürücü "C" test. Bootstrap method was employed to find p-value for the Ryan-Joiner and Sürücü "C" tests. In comparing power of the test of these three tests, simulation studies were done using Fleishman's Power Method to generate the data. The Fleishman's Power method is simple and we can simulate the data with any desired levels of skewness and kurtosis. In this study, we generated the data in four cases: (1) symmetric-platykurtic distribution, (2) symmetric-leptokurtic distribution, (3) asymmetric-platykurtic distribution, and (4) asymmetric-leptokurtic distribution. In addition, for each case, the varying sample sizes were considered: $n=10, 20, 30, 50,$ and 100 .

The result of this study indicates that the Shapiro-Wilk statistical test has the largest power of the test comparing to the Ryan-Joiner and the Sürücü "C" tests for all cases. Furthermore, we found that larger sample size, the larger power of the test. Hence, the Shapiro-Wilk is a preferred method which gives the highest power of the test comparing to the Ryan-Joiner and the Sürücü "C" tests. The calculation of p-value for the Shapiro-Wilk statistical test is also easier than the other two methods since the sampling distribution of the Shapiro-Wilk test statistic is known

Department: Statistics

Student's Signature

Field of Study: Statistics

Advisor's Signature

Academic Year: 2015

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้เสร็จสมบูรณ์ได้ ด้วยความช่วยเหลือและความเอาใจใส่อย่างดียิ่งจาก ผู้ช่วยศาสตราจารย์ ดร. วิฐรา พึ่งพาพงศ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ท่านอาจารย์ได้กรุณาใช้เวลาให้คำปรึกษา คำแนะนำ แนวคิดที่เป็นประโยชน์ และข้อคิดเห็นต่างๆ เพื่อปรับปรุงแก้ไข วิทยานิพนธ์ ตลอดจนให้ความช่วยเหลือ อบรมสั่งสอน และคอยเป็นกำลังใจในการทำงาน จนกระทั่งวิทยานิพนธ์นี้เสร็จสมบูรณ์ ผู้วิจัยขอขอบพระคุณท่านอาจารย์เป็นอย่างสูงด้วยความเคารพอย่างยิ่ง

ผู้วิจัยขอขอบพระคุณท่าน อาจารย์ ดร. อัครินทร์ ไพบูลย์พานิช ประธานกรรมการสอบ วิทยานิพนธ์ อาจารย์ ดร. นัท กุลวานิช และอาจารย์ ดร. อรุณี กำลัง กรรมการสอบวิทยานิพนธ์ ท่านอาจารย์ทั้งสามท่านได้กรุณาใช้เวลาเป็นกรรมการสอบครั้งนี้ ตลอดจนช่วยตรวจสอบและให้ คำแนะนำเพื่อแก้ไขวิทยานิพนธ์ฉบับนี้ให้สมบูรณ์ยิ่งขึ้น อีกทั้งขอขอบพระคุณคณาจารย์ประจำ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัยทุกท่านที่ให้โอกาส ทางการศึกษา ถ่ายทอดวิชาความรู้ และอบรมสั่งสอนทั้งในการเรียนและการดำรงชีวิตให้แก่ผู้วิจัย เสมอมาจนสำเร็จการศึกษาในครั้งนี้ และขอขอบคุณเจ้าหน้าที่ภาควิชาสถิติทุกท่านที่ให้ความ ช่วยเหลือในการติดต่อประสานงานในเรื่องต่าง ๆ ทำให้การดำเนินงานเป็นไปอย่างราบรื่น

สุดท้ายนี้ผู้วิจัยขอกราบขอบพระคุณและมอบความสำเร็จนี้ให้แก่ครอบครัว ผู้ซึ่งคอย ห่วงใย อยู่เคียงข้าง ให้กำลังใจ ส่งเสริม สนับสนุนในทุกด้าน และขอขอบคุณเพื่อน ๆ ทุกคนที่ ช่วยเหลือ ให้คำแนะนำ และเป็นกำลังใจตลอดมา ซึ่งกำลังใจและความรักอย่างต่อเนื่อง ได้เป็น แรงบันดาลใจทำให้ผู้วิจัยมาถึง ณ จุดนี้ เพื่อได้จัดทำงานวิจัยที่เป็นประโยชน์ขึ้นมา

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญ.....	ช
สารบัญตาราง.....	ญ
สารบัญภาพ.....	ฎ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์	4
1.3 ข้อตกลงเบื้องต้น.....	4
1.4 ขอบเขตของการวิจัย.....	4
1.6 เกณฑ์ที่ใช้ในการตัดสินใจ.....	6
1.7 วิธีการศึกษา.....	7
1.8 ประโยชน์ที่คาดว่าจะได้รับ.....	7
บทที่ 2 ทฤษฎีและตัวสถิติที่เกี่ยวข้อง.....	8
2.1 ทฤษฎีที่เกี่ยวข้องกับการแจกแจง.....	8
2.1.1 ความเบ้ (Skewness).....	9
2.1.2 ความโด่ง (Kurtosis).....	10
2.2 การทดสอบสมมติฐานการแจกแจงแบบปกติ (Normality Hypothesis Testing).....	11
2.2.1 วิธีการทดสอบของชาปิโร-วิลค์ (Shapiro - Wilk / W Test).....	11
2.2.2 วิธีทดสอบของไรอัน-จอยเนอร์ (Ryan - Joiner / RJ Test).....	12
2.2.3 ทดสอบของซุรุคู“ซี” (Sürücü “C” / C Test).....	13

2.3 วิธีเพาเวอร์ (Power Method).....	14
บทที่ 3 วิธีการดำเนินการศึกษา.....	18
3.1 ขอบเขตของการศึกษา.....	19
3.2 ขั้นตอนในการดำเนินการศึกษา.....	20
3.3 ขั้นตอนการทำงานของโปรแกรม.....	22
บทที่ 4 ผลการวิจัย.....	25
4.1 ผลการจำลองจากวิธีเพาเวอร์ของ Fleishman (1978).....	26
4.2 ผลการวิเคราะห์ข้อมูลที่จำลองด้วยกราฟ.....	27
4.2.1 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งต่ำ.....	28
4.2.2 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งสูง.....	32
4.2.3 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ.....	36
4.2.4 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งสูง.....	40
4.3 ผลการเปรียบเทียบอำนาจการทดสอบ (Power of Test) ของสถิติทดสอบการแจกแจง ปกติ ระหว่างสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C”	44
4.3.1 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ.....	45
4.3.2 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะสมมาตรและโด่งสูง.....	45
4.3.3 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ.....	46
4.3.4 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง.....	46
4.4 การวิเคราะห์อำนาจการทดสอบ	53
4.4.1 วิเคราะห์ผลจากวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C”	53
4.4.2 วิเคราะห์ผลจากค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C”.....	57
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	60
5.1 สรุปผลการวิจัย.....	60

5.1.1 การเปรียบเทียบอำนาจทดสอบ	61
5.1.2 ปัจจัยที่ส่งผลต่อประสิทธิภาพของอำนาจทดสอบ	63
5.2 อภิปรายผลการวิจัย	63
5.3 ข้อเสนอแนะ	64
5.3.1 ข้อเสนอแนะในการนำไปประยุกต์ใช้งาน	64
5.3.2 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป	64
รายการอ้างอิง	65
ภาคผนวก.....	67
ประวัติผู้เขียนวิทยานิพนธ์	73



สารบัญตาราง

ตารางที่	หน้า
ตารางที่ 2. 1 Lower Bounds of γ_2	15
ตารางที่ 2. 2 Two real solutions for extreme values of γ_2	16
ตารางที่ 2. 3 Four real solutions for $\gamma_1=1$ and $\gamma_2=3.75$	17
ตารางที่ 4. 1 แสดงค่า a, b, c และ d ของคู่อันดับ (0,-1) (0,1) (0,2) (0,3)	26
ตารางที่ 4. 2 แสดงค่า a, b, c และ d ของคู่อันดับ (0,5) (0,10) (0,20) (0,50)	26
ตารางที่ 4. 3 แสดงค่า a, b, c และ d ของคู่อันดับ (-1,2) (-1,3) (1,2) (1,3)	26
ตารางที่ 4. 4 แสดงค่า a, b, c และ d ของคู่อันดับ (-2,10) (-2,50) (1,10) (1,50)	27
ตารางที่ 4. 5 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ (0,-1) (0,1) (0,2) (0,3)	45
ตารางที่ 4. 6 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ (0,5) (0,10) (0,20) (0,50)	45
ตารางที่ 4. 7 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ (-1,2) (-1,3) (1,2) (1,3)	46
ตารางที่ 4. 8 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงความเบ้และความโด่งเป็นคู่อันดับ (-2,10) (-2,50) (1,10) (1,50)	46
ตารางที่ 5. 1 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]	61

ตารางที่	หน้า
ตารางที่ 5. 2 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะสมมาตรและโด่งสูง โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]	61
ตารางที่ 5. 3 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]	62
ตารางที่ 5. 4 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]	62



สารบัญภาพ

ภาพที่	หน้า
ภาพที่ 2. 1 แสดงลักษณะการแจกแจงเบ้ซ้าย (a) และเบ้ขวา (b)	9
ภาพที่ 2. 2 แสดงลักษณะความโด่งแบบ Platykurtic Mesokurtic และ Leptokurtic	10
ภาพที่ 2. 3 Possible Values of γ_1 and γ_2	16
ภาพที่ 4. 1 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=-1$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	28
ภาพที่ 4. 2 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=1$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	29
ภาพที่ 4. 3 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	30
ภาพที่ 4. 4 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	31
ภาพที่ 4. 5 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=5$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	32
ภาพที่ 4. 6 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	33
ภาพที่ 4. 7 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=20$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	34
ภาพที่ 4. 8 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	35
ภาพที่ 4. 9 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=-1, \gamma_2=2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	36
ภาพที่ 4. 10 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=-1, \gamma_2=3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	37
ภาพที่ 4. 11 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	38

ภาพที่	หน้า
ภาพที่ 4. 12 แสดง Histogram และNormal Q-Q Plotของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	39
ภาพที่ 4. 13 แสดง Histogram และNormal Q-Q Plotของการแจกแจงลักษณะ $\gamma_1=-2, \gamma_2=10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	40
ภาพที่ 4. 14 แสดง Histogram และNormal Q-Q Plotของการแจกแจงลักษณะ $\gamma_1=-2, \gamma_2=50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	41
ภาพที่ 4. 15 แสดง Histogram และNormal Q-Q Plotของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	42
ภาพที่ 4. 16 แสดง Histogram และNormal Q-Q Plotของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100.....	43
ภาพที่ 4. 17 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อ กำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงสมมาตรและโด่งต่ำ.....	47
ภาพที่ 4. 18 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อ กำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงสมมาตรและโด่งสูง.....	48
ภาพที่ 4. 19 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อ กำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงไม่สมมาตรและโด่งต่ำ..	49
ภาพที่ 4. 20 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อ กำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงไม่สมมาตรและโด่งสูง..	50
ภาพที่ 4. 21 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,-1) และ n=30.....	53
ภาพที่ 4. 22 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,5) และ n=30.....	54
ภาพที่ 4. 23 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (-1,2) และ n=30.....	55
ภาพที่ 4. 24 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (-2,10) และ n=30	56
ภาพที่ 4. 25 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,-1) และ n=30.....	57

ภาพที่	หน้า
ภาพที่ 4. 26 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,5) และ n=30.....	57
ภาพที่ 4. 27 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (-1,2) และ n=30.....	58
ภาพที่ 4. 28 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (-2,10) และ n=30	58



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

การอนุมานทางสถิติที่เกี่ยวข้องกับข้อมูลทางสถิติส่วนใหญ่มักจะมีข้อสมมติเบื้องต้นว่าข้อมูลต้องมาจากประชากรที่มีการแจกแจงแบบปกติ โดยในการตรวจสอบข้อมูลที่กำลังศึกษาว่ามีคุณสมบัติสอดคล้องกับข้อสมมติดังกล่าวหรือไม่นั้น จำเป็นต้องอาศัยเทคนิคทางสถิติต่าง ๆ เข้ามาช่วย

วิธีการทางสถิติในการตรวจสอบข้อมูลสำหรับประชากรที่มีการแจกแจงแบบปกติ จะทำการทดสอบสมมติฐานด้วยสถิติไม่อิงพารามิเตอร์ (Non-Parametric Statistics) ซึ่งมีหลายวิธีที่รู้จักกันดีและใช้กันอย่างแพร่หลาย ซึ่งวิธีที่ใช้กันอย่างกว้างขวางในการตรวจสอบการแจกแจงแบบปกติของข้อมูล ได้แก่ Shapiro-Wilk

Ryan และ Joiner (Ryan & Joiner, 1976) เสนอวิธีการทดสอบการแจกแจงแบบปกติของข้อมูล โดยตัวสถิติทดสอบของวิธี Ryan-Joiner ก็คือความสัมพันธ์ระหว่างข้อมูลและควอนไทล์ของข้อมูล แม้ว่าตัวสถิติทดสอบจะง่ายในการคำนวณ การแจกแจงของตัวสถิติทดสอบนี้มีความซับซ้อนแต่สามารถประมาณการแจกแจงได้ด้วยวิธี Simulation จากงานวิจัยนี้ทำการเปรียบเทียบประสิทธิภาพของสถิติทดสอบสำหรับการแจกแจงแบบปกติ 2 วิธี ได้แก่ สถิติทดสอบ Shapiro-Wilk และ Ryan-Joiner โดยศึกษาข้อมูลจากการจำลองการแจกแจงแบบต่าง ๆ อาทิเช่น การแจกแจง Uniform, Right Triangle, Exponential, Weibull, Lognormal, Cauchy และ Contaminated Normal ด้วยขนาดตัวอย่าง 10 และ 20 ผลการวิจัยในการเปรียบเทียบอำนาจการทดสอบ พบว่าการแจกแจงแบบ Uniform และ Right Triangle อำนาจการทดสอบของ Shapiro-Wilk สูงกว่า Ryan-Joiner ส่วนการแจกแจงแบบ Cauchy และ Contaminated Normal อำนาจการทดสอบของ Ryan-Joiner สูงกว่า Shapiro-Wilk จึงสรุปได้ว่าการแจกแจงที่เหมาะสมสำหรับ Shapiro-Wilk คือการแจกแจงแบบ short-tailed distributions และการแจกแจงที่เหมาะสมสำหรับ Ryan-Joiner คือการแจกแจงแบบ long-tailed distributions นั้นเอง อีกทั้งค่าอำนาจการทดสอบของทั้ง 2 วิธีในแต่ละการแจกแจงยังมีค่าใกล้เคียงกัน ซึ่งสถิติทดสอบวิธี Ryan-Joiner มีอำนาจการทดสอบที่สูงเช่นเดียวกัน แต่อาจจะยังไม่เป็นที่แพร่หลายมากเมื่อเทียบกับวิธี Shapiro-Wilk

สถิติทดสอบนี้เป็นวิธีใหม่และน่าสนใจ ศึกษาโดย Baris Sürücü (Baris Sürücü, 2008) คือสถิติทดสอบ C เป็นส่วนผสมระหว่างสถิติทดสอบ Shapiro - Wilk และค่าความสัมพันธ์หรือสถิติทดสอบ R โดยการแจกแจงของตัวสถิติทดสอบ C มีความซับซ้อน แต่สามารถประมาณการแจกแจงได้

ด้วยวิธี Simulation จากงานวิจัยนี้ทำการเปรียบเทียบอำนาจของสถิติทดสอบ 6 วิธี คือ สถิติทดสอบ Sürücü “C” (C), Shapiro – Wilk (W), Ryan-Joiner (R), modified EDF tests (\hat{D} and \hat{A}) และ Reduces to Z for exponential (Z^*) แบ่งการแจกแจง 2 ลักษณะ ได้แก่ การแจกแจงที่มีความเบ้ (Skew distributions) : χ^2 , Lognormal, Weibull(2), Beta(2,1), Tukey(10,3.1) และ Johnson SB(0.533,0.5) เป็นต้น และการแจกแจงที่สมมาตร (Symmetric distributions) : Normal, Student t(2), Student t(4), Logistic, Tukey(1,1.5), Uniform, และ JarqueBeraJB(0,1) เป็นต้น ด้วยขนาดตัวอย่าง 10, 20, 50 พบว่าลักษณะ Skew distributions สถิติทดสอบ C ให้อำนาจสูงที่ Tukey(10,3.1) และ JohnsonSB(0.533,0.5) ซึ่งการแจกแจงอื่น ๆ ในกลุ่มนี้ สถิติทดสอบ C มีค่าสูงใกล้เคียงกับ W และ Z^* ส่วนลักษณะ Symmetric distributions สถิติทดสอบ C ให้ค่าสูงสุดที่ Tukey(1,1.5) และ Uniform และยังพบว่าเมื่อขนาดตัวอย่างใหญ่ขึ้นอำนาจการทดสอบสำหรับการทดสอบก็สูงขึ้นอีกด้วย ดังนั้นในงานวิจัยนี้สถิติทดสอบวิธี Sürücü “C” เป็นสถิติทดสอบที่มีอำนาจทดสอบที่ดี อีกทั้งยังไม่เป็นที่แพร่หลายมากนักเมื่อเทียบกับวิธี Shapiro-Wilk เช่นเดียวกันกับวิธี Ryan-Joiner

สถิติทดสอบ Sürücü “C” สำหรับการทดสอบการแจกแจงแบบปกติ เป็นส่วนหนึ่งในงานวิจัยของ Fernando และ Jorge (Fernando & Jorge, 2013) ที่ได้นำตัวสถิติทดสอบนี้มาเปรียบเทียบอำนาจการทดสอบกับสถิติทดสอบอีก 6 วิธี คือสถิติทดสอบ Shapiro-Wilk, Kolmogorov-Smirnov, Anderson-Darling, Doornik-Hansen, Chen-Shapiro และ Robust Jarque-Beta โดยศึกษาข้อมูลจากการจำลองบน Ex-Gaussian Distributions ที่ขนาดตัวอย่างแตกต่างกัน ในการประมาณจากความเบ้และความโด่งสำหรับการแบ่งแต่ละกลุ่มของการแจกแจง เป็น Positively asymmetric long-tailed, medium-tailed และ short-tailed distributions จากนั้นทำการจำลองข้อมูลด้วย Weibull (2,1) และ Log-Normal (0,1) ผลของการทดสอบการแจกแจงแบบปกติของรายงานการวิจัยนี้พบว่าวิธี Shapiro-Wilk, Chen-Shapiro และ Sürücü “C” มีอำนาจการทดสอบที่สูง ซึ่งเป็นสถิติทางเลือกที่ดีในการทดสอบการแจกแจงแบบปกติ

จากการศึกษาของ Hadi และ Naser (Hadi & Naser, 2011) ได้เปรียบเทียบประสิทธิภาพการทดสอบสถิติทดสอบ 7 วิธี คือสถิติทดสอบ Cramer von Mises (CvM), Kolmogorov-Smirnov (KS), Kuiper (KL), Shapiro-Wilk (W), Anderson-Darling (AD), Vasicek (V) และ Jarque-Bera (JB) เป็นงานวิจัยที่ศึกษาถึงความเหมาะสมของสถิติทดสอบต่อการแจกแจงแบบปกติ โดยทำการเปรียบเทียบอำนาจการทดสอบ ซึ่งการจำลองข้อมูลจะแบ่งออกเป็น 4 กลุ่ม แต่ละกลุ่มจะใช้ขนาดตัวอย่าง 10, 20, 30 และ 50 สำหรับทุกการแจกแจงของกลุ่มที่กำหนดไว้ เช่นกลุ่มที่ 1 มีการแจกแจง t (1), t (3), Logistic, Double exponential และ Average เป็นต้น งานวิจัยนี้ได้สรุปผล

ของแต่ละกลุ่มที่พบอำนาจการทดสอบสูงสุดเป็นดังนี้ ผลของกลุ่มที่ 1 คือ JB กลุ่มที่ 2 คือ W กลุ่มที่ 3 คือ W, KL และกลุ่มที่ 4 คือ KL จะเห็นได้ว่า สถิติทดสอบ Shapiro-Wilk หรือ W ให้ผลอำนาจการทดสอบที่ดีที่สุดถึง 2 กลุ่ม ถือได้ว่า Shapiro-Wilk เป็นสถิติทดสอบที่มีอำนาจการทดสอบที่ดี

จากการศึกษาของ Mehmet และ Akin (Mehmet & Akin, 2013) ได้เปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ 3 วิธี คือสถิติทดสอบ Kolmolkorov-Smirnov, Shapiro-Wilk และ Lilliefors โดยข้อมูลที่ใช้ในการจำลองมาจากการแจกแจงแบบต่าง ๆ อาทิเช่น การแจกแจง $t(30)$, Gamma (2,3), Weibul (1.5,1), Exp (0.50) และ Beta (2,5) เป็นต้น จากผลการรายงานพบว่า การแจกแจงแบบ Exponential มีประสิทธิภาพสำหรับทั้ง 3 การทดสอบ แต่ถ้าไม่คำนึงถึงการแจกแจงและขนาดตัวอย่าง พบว่าสถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงกว่าอีก 2 การทดสอบข้างต้น สำหรับการทดสอบการแจกแจงแบบปกตินี้ ถือได้ว่าสถิติทดสอบ Shapiro-Wilk เป็นสถิติทดสอบที่มีอำนาจการทดสอบที่ดี ซึ่งให้ผลเช่นเดียวกับงานวิจัยของ Hadi และ Naser (2011)

จะเห็นได้ว่าผู้ที่ศึกษาเกี่ยวกับการเปรียบเทียบประสิทธิภาพเกี่ยวกับการแจกแจงแบบปกติของข้อมูลโดยส่วนใหญ่จะจำลองข้อมูลจากการแจกแจงแบบต่าง ๆ มีผู้วิจัยได้ทำการศึกษามากขึ้นอย่างต่อเนื่อง ซึ่งวิธีการตรวจสอบดังกล่าว ถือว่ามีความจำเป็นอย่างมากในทางสถิติ แต่สำหรับในการเปรียบเทียบประสิทธิภาพการทดสอบของสถิติทดสอบด้วยวิธีการที่ใช้ความเบ้และความโด่งในการจำลองข้อมูลยังมีรายงานการศึกษาที่ไม่กว้างขวางนัก ซึ่งเป็นสิ่งที่ผู้วิจัยต้องการศึกษา

วิธีการที่ผู้วิจัยใช้ศึกษาเพื่อจำลองข้อมูลในงานวิจัยนี้ คือวิธีเพาเวอร์ของ Fleishman (1978) นำเสนอโดย Hao Luo (Hao Luo, 2011) สำหรับการจำลองตัวแปรที่ไม่ปกติ (Non-Normal) ซึ่งวิธีการจำลองของข้อมูลที่ไม่ปกติของ Fleishman's เริ่มจากการสุ่มตัวแปร x ในการแจกแจงแบบปกติมาตรฐาน จะได้ว่า Y มีการแจกแจงที่มีค่าของความเบ้ (Skewness) และค่าของความโด่ง (Kurtosis) เป็น γ_1 และ γ_2 ตามลำดับ โดยผู้วิจัยจะนำไปค่าดังกล่าวนี้ไปจำลองข้อมูลเพื่อทำการเปรียบเทียบตัวสถิติทดสอบของแต่ละวิธีต่อไป

จากงานวิจัยของนักวิจัยหลายท่านที่กล่าวมานั้น พบว่าสถิติทดสอบของ (S. S. Shapiro & M. B. Wilk, 1965) เป็นที่รู้จักกันมานานและรู้จักกันอย่างแพร่หลาย โดยมีอำนาจการทดสอบที่ดี ส่วนสถิติทดสอบของ (Ryan & Joiner, 1976) ยังไม่เป็นที่แพร่หลายมากนัก แต่ก็ให้ค่าอำนาจการทดสอบที่สูง ซึ่งสถิติทดสอบ Ryan-Joiner มีการแจกแจงของตัวสถิติที่มีความซับซ้อน นั่นคือไม่ทราบลักษณะของการแจกแจง ต่างกับสถิติทดสอบของ Shapiro-Wilk ที่ทราบการแจกแจงชัดเจน จึงจำเป็นต้องทำการจำลองวิธี Simulation โดยใช้วิธีการ Bootstrap คำนวณค่า p-value เช่นเดียวกับสถิติทดสอบของ (Baris Sürücü, 2008) จะใช้วิธีดังกล่าวเช่นกัน ดังนั้นสถิติทดสอบทั้ง 3 วิธีนี้เป็นตัวสถิติที่ผู้วิจัยต้องการศึกษาเพื่อนำมาทดสอบสมมติฐานเกี่ยวกับการแจกแจงแบบปกติ

วิธี Bootstrap เสนอโดย Efron (B. Efron, 1979) เป็นวิธีการที่มีหลักเกณฑ์ว่าตัวอย่างที่ถูกเก็บมาเปรียบเสมือนประชากรจริง จึงทำการสุ่มตัวอย่างในจำนวนที่มากพอ แต่ละหน่วยตัวอย่างมีโอกาสถูกเลือกเท่า ๆ กัน ซึ่งวิธี Bootstrap แบ่งออกเป็น 2 แบบ คือ แบบใช้พารามิเตอร์โดยจะต้องทราบการแจกแจงประชากรที่สุ่มตัวอย่างมา ส่วนแบบไม่ใช้พารามิเตอร์จะจำลองค่าสังเกตโดยเป็นการสุ่มตัวอย่างแบบคืนที่ (Resampling with replacement) นั่นคือมีหน่วยตัวอย่างซ้ำได้นั่นเอง ในงานวิจัยนี้ทำการคำนวณค่า p-value จาก Bootstrap แบบไม่ใช้พารามิเตอร์

การศึกษาครั้งนี้ ผู้วิจัยมีความสนใจในการทดสอบสมมติฐานของข้อมูลว่ามีการแจกแจงแบบปกติหรือไม่ โดยจะทำการหาค่า p-value ของสถิติทดสอบทั้ง 3 วิธี ได้แก่ สถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” จากนั้นนำค่า p-value ที่ได้ มาวิเคราะห์ค่าอำนาจการทดสอบ (Power of Test) เพื่อเปรียบเทียบว่าสถิติทดสอบวิธีใดมีประสิทธิภาพมากที่สุด รวมถึงศึกษาว่าการทดสอบแต่ละวิธีมีความอ่อนไหวกับการเบี่ยงเบนของข้อมูลที่เกิดจากการแจกแจงแบบปกติหรือไม่ อย่างไร นอกจากนี้ งานวิจัยนี้ยังสามารถเป็นแนวทางในการเลือกสถิติทดสอบที่เหมาะสมเพื่อใช้ทดสอบการแจกแจงปกติของข้อมูลอีกด้วย

1.2 วัตถุประสงค์

เพื่อศึกษาว่ากรณีความเบ้ (Skewness) และความโด่ง (Kurtosis) รูปแบบใดที่ส่งผลสำหรับการทดสอบการแจกแจงปกติของข้อมูลในแต่ละวิธี แล้วทำการเปรียบเทียบประสิทธิภาพสถิติทดสอบดังกล่าว พร้อมทั้งศึกษาถึงขนาดกลุ่มตัวอย่างว่าขนาดกลุ่มตัวอย่างใดเหมาะสมกับวิธีทดสอบใดบ้าง

1.3 ข้อตกลงเบื้องต้น

1. ข้อมูลที่สุ่มมา เป็นอิสระต่อกัน
2. ข้อมูลอยู่ในระดับอันดับ (Ordinal Scale) ขึ้นไป

1.4 ขอบเขตของการวิจัย

ในการวิจัยครั้งนี้จะทำการศึกษาภายใต้ขอบเขต ดังต่อไปนี้

1. ศึกษาภายใต้วิธีเพาเวอร์ของ Fleishman (1978) เพื่อใช้ในการจำลองข้อมูล ดังนี้
 - กำหนดค่าความเบ้ (Skewness) และความโด่ง (Kurtosis) เป็นคู่อันดับ คือ (γ_1, γ_2) โดยให้ γ_1 แทน ค่าของความเบ้ และ γ_2 แทน ค่าของความโด่ง

- การให้ค่าของคู่อันดับ จะกำหนดค่าของ γ_1 และ γ_2 ต้องมาจากเงื่อนไข $\gamma_2 \geq \gamma_1^2 - 2!$ จึงจะสามารถจำลองหาค่า a, b, c และ d ได้ โดยที่ $\gamma_1, \gamma_2 \in \mathbb{R}$

จะได้ว่า การให้ค่าของคู่อันดับ (γ_1, γ_2) สามารถแบ่งออกเป็น 4 กรณี ดังนี้

กรณีที่ 1 : การแจกแจงแบบสมมาตรและโด่งต่ำ

กำหนด : คู่อันดับ (0,-1) (0,1) (0,2) (0,3)

กรณีที่ 2 : การแจกแจงแบบสมมาตรและโด่งสูง

กำหนด : คู่อันดับ (0,5) (0,10) (0,20) (0,50)

กรณีที่ 3 : การแจกแจงแบบไม่สมมาตรและโด่งต่ำ

กำหนด : คู่อันดับ (-1,2) (-1,3) (1,2) (1,3)

กรณีที่ 4 : การแจกแจงแบบไม่สมมาตรและโด่งสูง

กำหนด : คู่อันดับ (-2,10) (-2,50) (1,10) (1,50)

หมายเหตุ : การแจกแจงสมมาตร เมื่อ $\gamma_1 = 0$

การแจกแจงโด่งต่ำ เมื่อ $\gamma_2 \leq 3$ และโด่งสูง เมื่อ $\gamma_2 > 3$

2. ศึกษาภายใต้กลุ่มขนาดตัวอย่าง (n) ดังนี้

- กลุ่มตัวอย่างมีขนาดเล็ก คือ มีขนาดตัวอย่างไม่เกิน 30 ($n \leq 30$)

กำหนด : $n = 10, 20, 30$

- กลุ่มตัวอย่างมีขนาดใหญ่ คือ มีขนาดตัวอย่างมากกว่า 30 ($n > 30$)

กำหนด : $n = 50, 100$

3. สถิติทดสอบที่ศึกษา คือ สถิติทดสอบ Shapiro-Wilk สถิติทดสอบ Ryan-Joiner และ สถิติทดสอบ Sürücü "C" โดยกำหนดระดับนัยสำคัญที่ 0.05

4. จากขอบเขตการวิจัยข้างต้น กำหนดการจำลองทำซ้ำข้อมูลในแต่ละกรณีจำนวน 500 รอบ

1.5 คำจำกัดความที่ใช้ในงานวิจัย

ประสิทธิภาพทดสอบ (Efficiency of Testing)

คือ เกณฑ์ในการตัดสินว่าตัวทดสอบใดดีที่สุดสุดในบรรดาสถิติทดสอบที่สนใจศึกษา โดยวัดประสิทธิภาพจากค่าอำนาจการทดสอบ (Power of the test)

ค่าพี-แวลู (P-value)

คือ ความน่าจะเป็นที่ค่าสถิติทดสอบมีค่าเกินกว่าสถิติทดสอบที่คำนวณได้จากข้อมูลตัวอย่าง ภายใต้สมมติฐานว่างที่เป็นจริง

ความเบ้ (Skewness)

คือ การบอกความสมมาตร หรือไม่สมมาตรของการแจกแจง โดยการแจกแจงที่สมมาตรนั้น จะมีค่าความเบ้เป็น 0

ความโด่ง (Kurtosis)

คือ การบอกลักษณะการแจกแจงว่าโด่งมากหรือโด่งน้อย

วิธีบูตสเตรป (Bootstrap Method)

คือ ตัวอย่างที่ถูกเก็บมาเปรียบเสมือนประชากรจริง ทำการสุ่มตัวอย่างด้วยจำนวนที่มากพอ แต่หน่วยตัวอย่างมีโอกาสถูกเลือกเท่า ๆ กัน โดยใช้การสุ่มตัวอย่างแบบคืนที่ (Resampling with Replacement) ขนาด n จากชุดข้อมูลเดิมที่มีเพื่อสร้างชุดขนาดตัวอย่าง n ที่เป็นไปได้ กล่าวคือ เปรียบเสมือนการสร้างการแจกแจงของสถิติตัวอย่าง (Sampling Distribution) ขึ้นมา

1.6 เกณฑ์ที่ใช้ในการตัดสินใจ

เกณฑ์ที่ใช้ในการตัดสินใจว่าสถิติทดสอบวิธีใดมีประสิทธิภาพและเหมาะสมสำหรับการทดสอบสมมติฐานการแจกแจงแบบปกติ คืออำนาจของการทดสอบ (Power of Test) และในทางสถิติการทดสอบที่ดีควรมีค่าอำนาจของการทดสอบที่สูง

อำนาจของการทดสอบ (Power of Test) : $(1 - \beta)$

โดยที่ สัญลักษณ์ β คือ ความน่าจะเป็นที่จะเกิดความผิดพลาดชนิดที่ 2 (Type II Error) เป็นความผิดพลาดในการตัดสินใจยอมรับ H_0 ที่ไม่จริง

$$\begin{aligned} \text{จะได้ว่า} \quad \text{อำนาจของการทดสอบ} &= 1 - \beta \\ &= 1 - P(\text{Type II Error}) \\ &= P(\text{Reject } H_0 \mid H_0 \text{ is false}) \end{aligned}$$

$$\text{ดังนั้น} \quad \text{อำนาจของการทดสอบ} = \frac{\text{จำนวนครั้งของการปฏิเสธ } H_0}{500}$$

หมายเหตุ :

สำหรับการวิจัยนี้ ทางผู้วิจัยทำการจำลองข้อมูล 500 ครั้ง ซึ่งเป็นจำนวน H_0 ที่เป็นเท็จทั้งหมด

1.7 วิธีการศึกษา

ในการวิจัยครั้งนี้มีวิธีการศึกษา ดังต่อไปนี้

1. ค้นคว้า ศึกษาข้อมูล เอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้อง
2. กำหนดเงื่อนไขและขอบเขตในการวิจัย
 - 2.1 คู่อันดับ (γ_1, γ_2) คือ ค่าความเบ้ (Skewness : γ_1) และความโด่ง (Kurtosis : γ_2)
 - 2.2 กำหนดค่า (γ_1, γ_2) เมื่อ $\gamma_1, \gamma_2 \in \mathbb{R}$
 - 2.3 กำหนดกลุ่มขนาดตัวอย่าง (n) คือ ขนาด 10, 20, 30, 50 และ 100
3. จำลองข้อมูล (Simulation) โดยใช้วิธีเพาเวอร์ของ Fleishman (1978) ดังนี้
 - 3.1 จำลอง x จาก Standard Normal
 - 3.2 นำค่า γ_1 และ γ_2 ที่กำหนด มาจำลองในโปรแกรม R เพื่อหาค่า a, b, c และ d
 - 3.3 ทำการจำลองตัวแปรสุ่ม Y ด้วยสมการ $Y = a + bx + cx^2 + dx^3$
4. ขั้นตอนการทดสอบสมมติฐานทางสถิติ
 - 4.1 ตั้งสมมติฐาน
 - 4.2 คำนวณค่าสถิติทั้ง 3 วิธี ได้แก่
 - สถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü "C"
 - 4.3 หาค่า p-value ของการทดสอบ
 - 4.4 ตัดสินใจว่าปฏิเสธสมมติฐานว่าง (H_0) หรือไม่
5. เปรียบเทียบประสิทธิภาพของสถิติทดสอบจากอำนาจการทดสอบ
6. วิเคราะห์และสรุปผลเกี่ยวกับความเหมาะสมของสถิติทดสอบสำหรับการแจกแจงแบบปกติ

1.8 ประโยชน์ที่คาดว่าจะได้รับ

จากงานวิจัยนี้ จะเป็นแนวทางในการเลือกสถิติทดสอบที่เหมาะสม เพื่อใช้ในการทดสอบการแจกแจงแบบปกติของข้อมูล

บทที่ 2

ทฤษฎีและตัวสถิติที่เกี่ยวข้อง

การอนุมานทางสถิติที่เกี่ยวข้องกับข้อมูลทางสถิติส่วนใหญ่มักจะมีข้อกำหนดเบื้องต้นว่า ข้อมูลหรือกลุ่มตัวอย่างที่สุ่มมาจากประชากรมีลักษณะการแจกแจงแบบใด โดยในการตรวจสอบข้อมูลที่กำลังศึกษาว่ามีการแจกแจงแบบใดนั้น หรือการตรวจสอบข้อมูลสำหรับประชากรที่ศึกษาว่ามีการแจกแจงเป็นไปตามที่คาดหวังไว้หรือไม่ จะทำการทดสอบสมมติฐานด้วยสถิติไม่อิงพารามิเตอร์ (Non-Parametric Statistics) โดยสถิติที่ใช้ในการทดสอบการแจกแจงมีหลายวิธีให้เลือกใช้ แต่เมื่อสถานการณ์ของการแจกแจงเปลี่ยนแปลงไปจะส่งผลให้ต่อประสิทธิภาพของสถิติทดสอบนั้น ๆ ด้วย จึงยังไม่สามารถบอกได้ว่าสถิติทดสอบวิธีใดดีที่สุดหรือมีประสิทธิภาพที่สุด

ในหัวข้อนี้จะนำเสนอเกี่ยวกับรายละเอียดของทฤษฎีที่เกี่ยวข้องกับการแจกแจง รวมไปถึงวิธีการทดสอบสมมติฐานการแจกแจงแบบปกติ 3 การทดสอบ และขั้นตอนของวิธีเพาเวอร์ซึ่งเป็นวิธีการที่ใช้ในการจำลองข้อมูลเพื่อใช้ในการเปรียบเทียบประสิทธิภาพหรืออำนาจของสถิติทดสอบ

2.1 ทฤษฎีที่เกี่ยวข้องกับการแจกแจง

ลักษณะการแจกแจงของตัวแปรสุ่ม (วราพรณ เจริญชา, 2556) แบ่งออกเป็น 2 ลักษณะ คือ

1. ลักษณะแบบสมมาตร (Symmetric Distribution)

การแจกแจงของตัวแปรสุ่มที่มีลักษณะสมมาตร จะมีการกระจายของข้อมูลสม่ำเสมอโดยจะเบี่ยงเบนจากค่ากลางไปในทางบวกและลบพอ ๆ กัน ซึ่งข้อมูลที่มีลักษณะสมมาตร พบว่าค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) และฐานนิยม (Mode) จะเท่ากันหรืออยู่ที่จุดเดียวกัน คือจุดที่มีความถี่สูงสุด เช่น การแจกแจงปกติ (Normal distribution) เส้นโค้งจะมีลักษณะเป็นรูปประฆังคว่ำ

2. ลักษณะแบบไม่สมมาตร (Asymmetric Distribution)

การแจกแจงของตัวแปรสุ่มที่มีลักษณะไม่สมมาตร จะมีการกระจายของข้อมูลไม่สม่ำเสมอ ซึ่งข้อมูลที่มีลักษณะไม่สมมาตร พบว่าค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) และฐานนิยม (Mode) ของตัวแปรสุ่มไม่อยู่ในตำแหน่งเดียวกันทั้งหมด อาจมีลักษณะการแจกแจงเป็นแบบเบ้ขวาหรือเบ้ซ้าย

ลักษณะการแจกแจงที่เกี่ยวข้องข้องในการศึกษา แบ่งออกเป็นความเบ้และความโด่ง ดังนี้

2.1.1 ความเบ้ (Skewness)

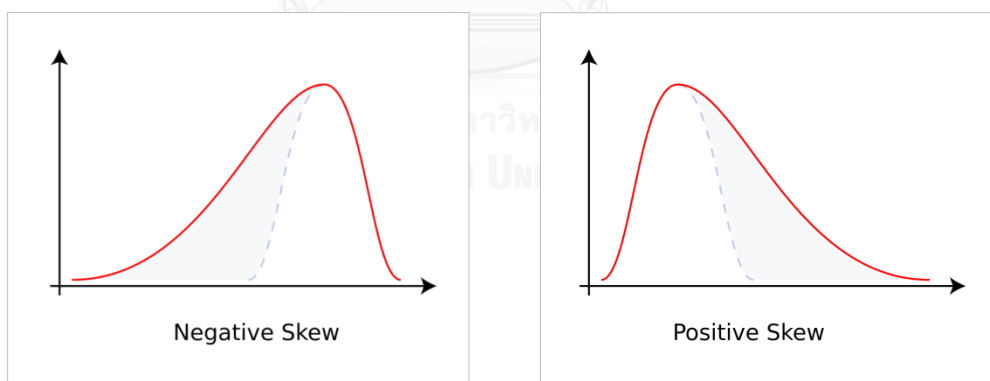
ความเบ้ (Skewness, 2013) คือ ความไม่สมมาตรของการแจกแจง โดยทั่วไปนิยามวัดด้วยอัตราส่วนของโมเมนต์ที่เป็นการวัดขนาดของความเบ้ในรูปค่าสัมประสิทธิ์ความเบ้ (Coefficient of skewness) ถ้าค่าสัมประสิทธิ์ความเบ้เท่ากับศูนย์หมายถึงโค้งมีลักษณะสมมาตร ซึ่งความเบ้ที่ลักษณะการแจกแจงของตัวแปรสุ่มที่โค้งด้านซ้ายและด้านขวาไม่สมมาตร แบ่งออกเป็น 2 แบบ ได้แก่

1. เบ้ซ้าย (Left skewness หรือ Negative skewness)

ลักษณะการแจกแจงของตัวแปรสุ่มที่มีเส้นโค้งด้านซ้ายหางยาวกว่าปกติ โดยหางของเส้นโค้งทอดยาวไปทางซ้ายมากกว่าทางขวา ซึ่งในการวัดค่าสัมประสิทธิ์ความเบ้ค่าทดสอบจะเป็นลบ และกล่าวได้ว่าการแจกแจงที่ข้อมูลเบ้ซ้ายจะมีค่าความเบ้น้อยกว่า 0

2. เบ้ขวา (Right skewness หรือ Positive skewness)

ลักษณะการแจกแจงของตัวแปรสุ่มที่มีเส้นโค้งด้านขวาหางยาวกว่าปกติ โดยหางของเส้นโค้งทอดยาวไปทางขวามากกว่าทางซ้าย ซึ่งในการวัดค่าสัมประสิทธิ์ความเบ้ค่าทดสอบจะเป็นบวก และกล่าวได้ว่าการแจกแจงที่ข้อมูลเบ้ขวามีค่าความเบ้มากกว่า 0 (ศุภกิจ วงศ์วิวัฒน์นุกิจ, 2550)



(a) การแจกแจงเบ้ซ้าย

(b) การแจกแจงเบ้ขวา

ภาพที่ 2. 1 แสดงลักษณะการแจกแจงเบ้ซ้าย (a) และเบ้ขวา (b)

ที่มา : <https://en.wikipedia.org/wiki/Skewness> (Skewness, 2013)

2.1.2 ความโด่ง (Kurtosis)

ความโด่ง คือ ภาวะยอดมนของเส้นโค้งความถี่ กรณีการแจกแจงปกติความโด่งมีค่าเท่ากับ 3 โดยลักษณะความโด่งของการแจกแจงแบ่งเป็น 3 แบบ (สำนักงานราชบัณฑิตยสภา, 2558) ดังนี้

1. ความโด่งแบบพลาติเคอร์ติก (Platykurtic)

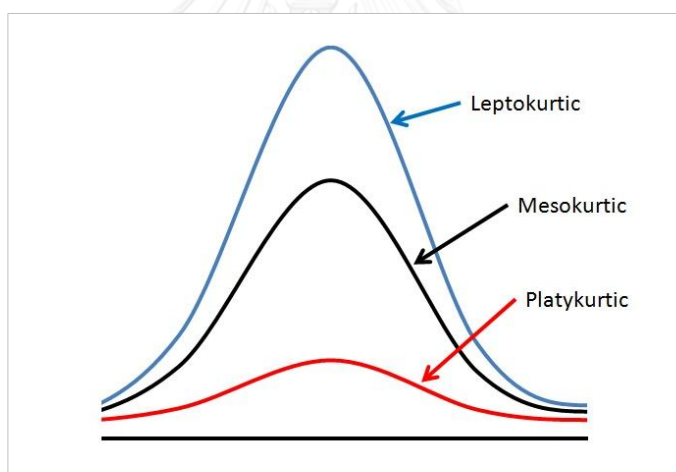
เป็นการแจกแจงที่มีความโด่งน้อยกว่าความโด่งของการแจกแจงปกติ นั่นคือเส้นโค้งความถี่มีค่าความโด่งน้อยกว่า 3 หรือเรียกว่า ความโด่งต่ำ

2. ความโด่งแบบมีโซเคอร์ติก (Mesokurtic)

เป็นการแจกแจงที่มีความโด่งเท่ากับความโด่งของการแจกแจงปกติ (Normal) นั่นคือเส้นโค้งความถี่มีค่าความโด่งเท่ากับ 3 หรือเรียกว่า ความโด่งปกติ

3. ความโด่งแบบเลปโทเคอร์ติก (Leptokurtic)

เป็นการแจกแจงที่มีความโด่งมากกว่าความโด่งของการแจกแจงปกติ นั่นคือเส้นโค้งความถี่มีค่าความโด่งมากกว่า 3 หรือเรียกว่า ความโด่งสูง



ภาพที่ 2. 2 แสดงลักษณะความโด่งแบบ Platykurtic Mesokurtic และ Leptokurtic

ที่มา : <http://schaal15.blog.sbc.edu> (Spotlight, 2014)

2.2 การทดสอบสมมติฐานการแจกแจงแบบปกติ (Normality Hypothesis Testing)

2.2.1 วิธีการทดสอบของชาปิโร-วิลค์ (Shapiro - Wilk / W Test)

การทดสอบนี้เป็นวิธีการทดสอบที่เหมาะสมกับขนาดตัวอย่างที่ค่อนข้างต่ำ นั่นคือสามารถใช้ได้กับกรณีขนาดตัวอย่างไม่เกิน 30 ได้ และลักษณะการแจกแจงข้อมูลเป็นแบบต่อเนื่อง

- สมมติฐาน

H_0 : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ

H_a : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบไม่ปกติ

- การคำนวณ

คำนวณหาตัวสถิติทดสอบโดยเรียงลำดับค่าสังเกต n ค่า จากน้อยไปหามาก

$$x_{(1)} \leq x_{(2)} < \dots \leq x_{(n)}$$

ตัวสถิติทดสอบ คือ

$$W_{cal} = \frac{b^2}{\sum_{i=1}^k (x_i - \bar{x})^2}$$

$$\text{เมื่อ } b = \sum_{i=1}^k a_{n-i+1} (x_{n-i+1} - x_i)$$

โดย a_{n-i+1} คือ สัมประสิทธิ์ของ W-test ในการทดสอบการแจกแจงแบบปกติ

โดยเปิดจากตาราง Coefficients $\{a_{n-i+1}\}$ for the W test for Normality

สำหรับ $n \leq 50$

$$\text{และ } k = \begin{cases} \frac{n}{2} & \text{เมื่อ } n \text{ เป็นเลขคู่} \\ \frac{n-1}{2} & \text{เมื่อ } n \text{ เป็นเลขคี่} \end{cases}$$

- เกณฑ์การตัดสินใจ

จะปฏิเสธ H_0 เมื่อ W_{cal} มีค่าน้อย ๆ โดยค่า W_{cal} มีค่ามากกว่าศูนย์เสมอ

ดังนั้น $p\text{-value} = P(W < W_{cal})$

หรือสามารถหาค่าวิกฤตได้จากตาราง Percentage point of the W test for Normality

สำหรับ n มีค่าตั้งแต่ 3 ถึง 50

2.2.2 วิธีทดสอบของไรอัน-จอยเนอร์ (Ryan - Joiner / RJ Test)

การทดสอบที่หาความสัมพันธ์ระหว่างข้อมูลและควอนไทล์ของข้อมูล ซึ่งตัวสถิติทดสอบจะง่ายต่อการคำนวณ แต่การแจกแจงของตัวสถิติทดสอบนี้มีความซับซ้อน โดยค่าสถิติหรือค่าสัมประสิทธิ์สหสัมพันธ์ (R) อยู่ใกล้ 1 ประชากรที่มีจะมีแนวโน้มเป็นปกติ วิธีการทดสอบนี้คล้ายกับวิธีทดสอบของ Shapiro - Wilk เป็นอย่างมาก

- สมมติฐาน

H_0 : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ

H_a : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบไม่ปกติ

- การคำนวณ

ตัวสถิติทดสอบ คือ

$$R_{cal} = \frac{\sum_{i=1}^n x_i b_i}{\sqrt{s^2(n-1) \sum_{i=1}^n b_i}}$$

เนื่องจาก R คือ ค่าความสัมพันธ์ระหว่างแต่ละคู่อันดับ (b_i, x_i)

จะได้ว่า

x_i คือ ข้อมูลที่ศึกษา โดยเรียงลำดับค่าสังเกต n ค่า จากน้อยไปหามาก

$$x_{(1)} \leq x_{(2)} < \dots \leq x_{(n)}$$

b_i คือ ค่าของข้อมูลที่ศึกษาแปลงเป็นคะแนนปกติ

$$\text{โดยที่ } b_i = \phi^{-1}\left(\frac{i - 0.375}{n + 0.25}\right), \quad i = 1, 2, \dots, n$$

และ $\phi^{-1}(\cdot)$ คือ inverse CDF ของตัวแปรสุ่มแบบปกติมาตรฐาน

s^2 คือ ค่าความแปรปรวนของตัวอย่าง

- เกณฑ์การตัดสินใจ

จะปฏิเสธ H_0 เมื่อ R_{cal} หรือค่าสัมประสิทธิ์สหสัมพันธ์มีค่าน้อยมากๆ

ดังนั้น $p\text{-value} = P(R < R_{cal})$

เนื่องจากการแจกแจงของตัวสถิติ R มีความซับซ้อน จะหาค่า $p\text{-value}$ โดยใช้วิธี Bootstrap

2.2.3 ทดสอบของซุรูกู“ซี” (Sürücü “C” / C Test)

การทดสอบนี้เป็นสถิติทดสอบของวิธีการทดสอบภาวะสารูปสนิทธิ (Goodness of fit test) อีกวิธีหนึ่ง โดยเป็นส่วนผสมระหว่างสถิติทดสอบ Shapiro - Wilk และค่าความสัมพันธ์หรือสถิติทดสอบ R ซึ่งสถิติทดสอบนี้จะไม่ทราบการแจกแจงของตัวสถิติทดสอบ แต่โดยรวมจะมีประสิทธิภาพมากกว่าการทดสอบอื่นๆใน Goodness of fit test

- สมมติฐาน

H_0 : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ

H_a : $\{x_1, x_2, \dots, x_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบไม่ปกติ

- การคำนวณ

ตัวสถิติทดสอบ C แบ่งวิธีการคำนวณออกเป็น 2 ส่วน ดังนี้

1. จาก Shapiro - Wilk Statistic (W Test)

ให้ $x_{(1)} \leq x_{(2)} < \dots \leq x_{(n)}$ เป็นสถิติอันดับของตัวอย่างสุ่ม x_1, x_2, \dots, x_n

โดยที่ ตัวสถิติทดสอบ W คือ

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{X})^2}$$

เมื่อ $0 < W < \infty$ และ $\sum_{i=1}^n a_i^2 = 1$

โดย a_i คือ ค่าของสัมประสิทธิ์ W Test ($1 < i < n$)

2. จาก Correlation Statistic (R)

โดยที่ ตัวสถิติทดสอบ R คือ

$$R = 1 - \hat{\rho}^2 \quad ; (0 < R < 1)$$

เนื่องจาก $\hat{\rho}$ คือ สัมประสิทธิ์สหสัมพันธ์ระหว่าง $x_{(i)}$ และ t_i ; ($1 < i < n$)

จะได้ว่า t_i คือ quantiles ของ ฟังก์ชันความหนาแน่น f_0

$$\text{เมื่อ } t_{(i)} = \phi^{-1}\left(\frac{i}{n+1}\right) \quad ; 1 \leq i \leq n$$

และ $\phi^{-1}(\cdot)$ คือ inverse CDF ของตัวแปรสุ่มแบบปกติมาตรฐาน

Combined statistic: คือ การรวมกันระหว่างสถิติทดสอบ W และ R เพื่อให้ตัวสถิติมีประสิทธิผลการทดสอบมากขึ้น

ดังนั้น ตัวสถิติทดสอบ C คือ

$$C_{cal} = 1 - \{[1 + a_1(a_2 - 1)]W + a_1(1 - a_2)(1 - R)\}$$

เมื่อ $0 < a_1, a_2 < 1$ โดยที่ สัมประสิทธิ์ a_1 และ a_2 คำนวณจากสมการ ต่อไปนี้

$$a_1 = \exp(-(b_1 / 0.6)^5)$$

$$a_2 = \exp(-(b_2 / 3.5)^5)$$

เนื่องจาก $\sqrt{b_1}$ และ b_2 คือ ค่าความเบ้และความโด่งของตัวอย่าง ตามลำดับ สำหรับค่าของ $b_1 < 0.6$ และ $b_2 \leq 3.5$

- เกณฑ์การตัดสินใจ

จะปฏิเสธ H_0 เมื่อ C_{cal} มีค่ามากๆ

ดังนั้น $p\text{-value} = P(C > C_{cal})$

เนื่องจากการแจกแจงของตัวสถิติ C มีความซับซ้อน จะหาค่า p-value โดยใช้วิธี Bootstrap

2.3 วิธีเพาเวอร์ (Power Method)

Fleishman's (Fleishman, 1978) เสนอวิธีการจำลองตัวแปรที่ไม่ปกติ (Non-Normal) โดยการกำหนดของความเบ้ (Skewness) และค่าของความโด่ง (Kurtosis)

ให้ x เป็นตัวแปรสุ่มแบบปกติ ($x \sim N(0,1)$)

จาก Polynomial Transformation ;

$$Y = a + bx + cx^2 + dx^3 \quad (1)$$

เมื่อ a, b, c และ d เป็นค่าคงที่

สมมติให้ Y มีโมเมนต์อันดับ 1 ถึง 4 ดังนี้

$$E(Y) = 0, E(Y^2) = 1, E(Y^3) = \gamma_1 \text{ และ } E(Y^4) = \gamma_2 + 3$$

เมื่อ γ_1 คือ ค่าของความเบ้ (Skewness)

γ_2 คือ ค่าของความโด่ง (Kurtosis)

พบว่า $a = -c$

ส่วน b, c และ d สร้างจากการจำลองข้อมูล Y ที่หาจากสมการ Fleishman's ต่อไปนี้

$$F_1(b, c, d) = b^2 + 6bd + 2c^2 + 15d^2 - 1 = 0, \quad (2)$$

$$F_2(b, c, d) = 2c(b^2 + 24bd + 105d^2 + 2) - \gamma_1 = 0, \quad (3)$$

$$F_3(b, c, d) = 24(bd + c^2 [1 + b^2 + 28bd]) + d^2 [12 + 48bd + 141c^2 + 25d^2] - \gamma_2 = 0. \quad (4)$$

จากสมการ (3) ; $c = \frac{\gamma_1}{2(b^2 + 24bd + 105d^2 + 2)}$ นำไปแทนค่าในสมการ (2) และ (4)

จะทำให้เหลือ 2 สมการ เพื่อหาค่า b และ d

Fleishman ใช้ Newton Method แก้ปัญหาดังกล่าวข้างต้น ซึ่งปรากฏค่าของ b, c และ d สำหรับความเบ้ (γ_1) และความโด่ง (γ_2) ในตาราง 1 ถ้าต้องการช่วงค่า γ_1 และ γ_2 จำนวนมากขึ้น นอกเหนือที่ระบุในตารางที่ 1 จะมีวิธีหาค่า b, c และ d ดังนี้

$$F(b, c, d) = F_1^2(b, c, d) + F_2^2(b, c, d) + F_3^2(b, c, d), \quad (5)$$

จากสมการ (2), (3) และ (4) จะเห็นได้ว่า

$$F_1^2(b, c, d) = F_2^2(b, c, d) = F_3^2(b, c, d) = 0, \quad (6)$$

ดังนั้น $F(b, c, d) = 0$ ณ จุดที่ต่ำสุด

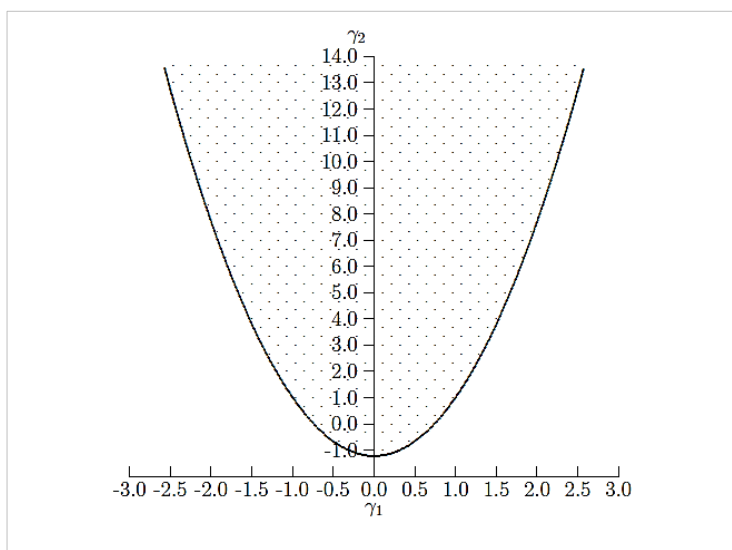
เนื่องจาก $F(b, c, d)$ ในสมการ (5) เป็นพหุนามอันดับสูง ในการที่จะหาค่า b, c และ d ต้องทำการหาอนุพันธ์ครั้งที่ 1 และ 2 โดยค่าที่สามารถหาได้นั้นมีเกณฑ์ที่ใช้ คือ $F(b, c, d) < 0.5 \times 10^{-30}$ จะให้ผลลัพธ์ที่ถูกต้อง ซึ่งแสดงในตารางของ Fleishman's

γ_1	Bound γ_2
0.0	-1.2264489
0.5	-0.8161896
1.0	0.4145884
1.5	2.4658850
2.0	5.3377003
2.5	9.0300342
3.0	13.5428868

ตารางที่ 2. 1 Lower Bounds of γ_2

Fleishman (1978) ชี้ให้เห็นว่า คำตอบจริงสำหรับ b, c และ d ไม่สามารถเกิดขึ้นได้ทุกค่าของ γ_1 และ γ_2 การแก้ปัญหาดังกล่าวข้างต้นนี้จะเป็นไปได้เฉพาะค่า γ_1 และ γ_2 ในรูปที่ 2.1 ซึ่งได้แสดงอาณาเขตของขอบเขตด้านล่างในรูปพาราโบลา นี้ นิยามโดย

$$\gamma_2 = -1.2264489 + 1.6410373\gamma_1^2 \quad (7)$$



ภาพที่ 2. 3 Possible Values of γ_1 and γ_2

(ที่มา : A Study of Fleishman's Power Method, Hao Luo, 2011)

จากตารางที่ 2.1 Lower Bounds of γ_2 เกิดจากการแทนค่าในสมการ (7) ซึ่งมีขอบเขตของค่าความโค้ง (γ_2) สำหรับบางค่าความเบ้ (γ_1) ที่อยู่ระหว่าง 0 ถึง 3 จะพบว่า γ_2 ที่มีค่าน้อย ๆ เท่านั้น ที่จะสามารถนำมาหาค่า γ_1 ในขอบเขตที่แคบเช่นนี้ได้

แต่ Fleishman's ได้ใช้วิธีการหนึ่งด้วยการหาค่า b, c และ d ที่สามารถแก้ปัญหาค่าของความโค้ง (γ_2) ที่มีค่ามากขึ้น ดังตัวอย่างในตารางที่ 2 โดยแสดงค่าสำหรับ γ_2 ขนาดใหญ่

γ_1	γ_2	b	c	d	$F(b, c, d)$
1	10	0.5642607	0.08771825	0.12620922	0.770372×10^{-32}
1	50	-0.0814358	0.0536441	0.27339828	0.986076×10^{-32}

ตารางที่ 2. 2 Two real solutions for extreme values of γ_2

จากตารางที่ 2.2 พบว่า คำตอบจริงสำหรับ b, c และ d ที่เกิดขึ้นมีได้หลายค่า เนื่องจากสมการ (4) เป็นสมการ Polynomial อันดับ 3 จะมีคำตอบจริงถึง 4 คำตอบ

จากสมการ (2)-(4) ถ้า b, c และ d เป็นคำตอบจริงสำหรับ (γ_1, γ_2) แล้ว $-b, c$ และ $-d$ เป็นอีกคำตอบของ (γ_1, γ_2) ค่าเดียวกัน ตัวอย่างเช่น ที่ $\gamma_1 = 1$ และ $\gamma_2 = 3.75$ โดยมีค่า c จำนวน 2 ค่าที่แตกต่างกัน แสดงในตารางที่ 2.3 ดังนี้

b	c	d	$F(b, c, d)$
0.7894207441645	0.1194238366286	0.0615396192450	$0.34281553 \times 10^{-32}$
-0.7894207441645	0.1194238366286	-0.0615396192450	$0.34281553 \times 10^{-32}$
1.3498589798986	0.3187035006281	-0.2026375792547	$0.61629758 \times 10^{-32}$
-1.3498589798986	0.3187035006281	0.2026375792547	$0.61629758 \times 10^{-32}$

ตารางที่ 2. 3 Four real solutions for $\gamma_1=1$ and $\gamma_2=3.75$

สรุปได้ว่า วิธีเพาเวอร์นี้ที่ซึ่งทำการกำหนดค่าของความเบ้ (γ_1) และค่าของความโด่ง (γ_2) ในสมการ Fleishman's จะสามารถหา b, c และ d โดยที่ $a = -c$ ได้จำนวนมากที่สุด 4 คำตอบที่แตกต่างกัน

บทที่ 3

วิธีการดำเนินการศึกษา

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบการแจกแจงปกติ วิธีใดมีประสิทธิภาพมากที่สุด โดยศึกษาจากกรณีความเบ้ (Skewness) และความโด่ง (Kurtosis) ในรูปแบบต่าง ๆ เพื่อนำค่าดังกล่าวไปจำลองข้อมูลด้วยวิธีเพาเวอร์ของ Fleishman (1978) จากนั้นนำข้อมูลที่จำลองไปหาค่า p-value เพื่อทดสอบสมมติฐานว่าข้อมูลมีการแจกแจงแบบปกติหรือไม่ โดยวัดประสิทธิภาพของสถิติทดสอบจากค่าอำนาจการทดสอบ (Power of Test) ใน 3 การทดสอบ ได้แก่ สถิติทดสอบ Shapiro-Wilk Ryan-Joiner และ Sürücü “C”

โดยมีสมมติฐานในการทดสอบ คือ

H_0 : $\{Y_1, Y_2, \dots, Y_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ

H_a : $\{Y_1, Y_2, \dots, Y_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบไม่ปกติ

พร้อมทั้งศึกษาถึงขนาดกลุ่มตัวอย่างว่าขนาดกลุ่มตัวอย่างใดเหมาะสมกับวิธีทดสอบใดบ้าง โดยตัวอย่างขนาดเล็กแทนขนาดตัวอย่าง 10, 20 และ 30 และตัวอย่างขนาดใหญ่แทนขนาดตัวอย่าง 50 และ 100 ภายใต้กรณีความเบ้และความโด่งที่แตกต่างกัน เพื่อการเลือกสถิติทดสอบที่เหมาะสม จะทำการใช้โปรแกรม R เวอร์ชัน 3.2.2 สำหรับการจำลองและวิเคราะห์ข้อมูลทั้งหมด โดยบทนี้กล่าวถึงวิธีการดำเนินการศึกษา ซึ่งมีรายละเอียดดังหัวข้อต่อไปนี้

3.1 ขอบเขตของการศึกษา

3.2 ขั้นตอนในการดำเนินการศึกษา

และในส่วนท้ายแสดงขั้นตอนสำหรับการเขียนโปรแกรมไว้ โดยอธิบายเป็นแผนภาพตั้งแต่เริ่มต้นการทำงานจนสิ้นสุดการทำงาน นั่นคือ

3.3 ขั้นตอนการทำงานของโปรแกรม

3.1 ขอบเขตของการศึกษา

ในการวิจัยครั้งนี้ เพื่อกำหนดสถานการณ์ต่าง ๆ จะทำการศึกษาภายใต้ขอบเขต ดังต่อไปนี้

- ศึกษาภายใต้วิธีเพาเวอร์ของ Fleishman (1978) เพื่อใช้ในการจำลองข้อมูล ดังนี้
 - กำหนดค่าความเบ้ (Skewness) และความโด่ง (Kurtosis) เป็นคู่อันดับ คือ (γ_1, γ_2) โดยให้ γ_1 แทน ค่าของความเบ้
 γ_2 แทน ค่าของความโด่ง
 - การให้ค่าของคู่อันดับ จะกำหนดค่าของ γ_1 และ γ_2 ต้องมาจากเงื่อนไข $\gamma_2 \geq \gamma_1^2 - 2!$ จึงจะสามารถจำลองหาค่า a, b, c และ d ได้ โดยที่ $\gamma_1, \gamma_2 \in \mathbb{R}$

จะได้ว่า การให้ค่าของคู่อันดับ (γ_1, γ_2) สามารถแบ่งออกเป็น 4 กรณี ดังนี้

กรณีที่ 1 : การแจกแจงแบบสมมาตรและโด่งต่ำ

กำหนด : คู่อันดับ (0,-1) (0,1) (0,2) (0,3)

กรณีที่ 2 : การแจกแจงแบบสมมาตรและโด่งสูง

กำหนด : คู่อันดับ (0,5) (0,10) (0,20) (0,50)

กรณีที่ 3 : การแจกแจงแบบไม่สมมาตรและโด่งต่ำ

กำหนด : คู่อันดับ (-1,2) (-1,3) (1,2) (1,3)

กรณีที่ 4 : การแจกแจงแบบไม่สมมาตรและโด่งสูง

กำหนด : คู่อันดับ (-2,10) (-2,50) (1,10) (1,50)

หมายเหตุ : การแจกแจงสมมาตร เมื่อ $\gamma_1 = 0$

การแจกแจงโด่งต่ำ เมื่อ $\gamma_2 \leq 3$ และโด่งสูง เมื่อ $\gamma_2 > 3$

- ศึกษาภายใต้กลุ่มขนาดตัวอย่าง (n) ดังนี้
 - กลุ่มตัวอย่างมีขนาดเล็ก คือ มีขนาดตัวอย่างไม่เกิน 30 ($n \leq 30$)
กำหนด : $n = 10, 20, 30$
 - กลุ่มตัวอย่างมีขนาดใหญ่ คือ มีขนาดตัวอย่างมากกว่า 30 ($n > 30$)
กำหนด : $n = 50, 100$
- สถิติทดสอบที่ศึกษา คือ สถิติทดสอบ Shapiro-Wilk สถิติทดสอบ Ryan-Joiner และ สถิติทดสอบ Sürücü "C" โดยกำหนดระดับนัยสำคัญที่ 0.05
- จากขอบเขตการวิจัยข้างต้น กำหนดการจำลองทำซ้ำข้อมูลในแต่ละกรณีจำนวน 500 รอบ

3.2 ขั้นตอนในการดำเนินการศึกษา

ในการวิจัยครั้งนี้มีขั้นตอนในการดำเนินการศึกษา ดังต่อไปนี้

1. ค้นคว้า ศึกษาข้อมูล เอกสาร ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

2. กำหนดการจำลองข้อมูลตามขอบเขตที่ศึกษา

ในการจำลองข้อมูลครั้งนี้ ศึกษาภายใต้วิธีเพาเวอร์ของ Fleishman (1978)

2.1 ให้ค่าของคู่อันดับ (γ_1, γ_2) ตามขอบเขตการวิจัยออกเป็น 4 กรณีข้างต้น

2.2 จำลองตัวแปรสุ่ม x จาก Standard Normal ซึ่งขนาดของ x ที่จำลองขึ้นมาในแต่ละครั้งเป็นไปตามค่า n ที่กำหนด

2.4 นำค่า γ_1 และ γ_2 ทั้ง 4 กรณีมาจำลองในโปรแกรม R เพื่อหาค่า a, b, c และ d

2.4 ทำการจำลอง Y ด้วยสมการ $Y = a + bx + cx^2 + dx^3$

3. การทดสอบสมมติฐาน

ในการทดสอบสมมติฐาน มีขั้นตอนดังต่อไปนี้

3.1 ตั้งสมมติฐาน

H_0 : $\{Y_1, Y_2, \dots, Y_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ

H_a : $\{Y_1, Y_2, \dots, Y_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบไม่ปกติ

3.2 คำนวณค่าสถิติทดสอบ ได้แก่

1. Shapiro-Wilk สถิติทดสอบ W

2. Ryan-Joiner สถิติทดสอบ R

3. Sürücü "C" สถิติทดสอบ C

3.3 หาค่า p-value ของการทดสอบ

1. Shapiro-Wilk หาได้ง่าย เนื่องจากทราบการแจกแจงตัวสถิติชัดเจน

2. Ryan-Joiner ไม่ทราบการแจกแจงตัวสถิติ จึงใช้วิธี Bootstrap Simulation

3. Sürücü "C" ไม่ทราบการแจกแจงตัวสถิติ จึงใช้วิธี Bootstrap Simulation

3.4 ตัดสินใจว่าปฏิเสธสมมติฐานว่าง (H_0) หรือไม่

พิจารณาจากค่า p-value ของการทดสอบ โดยกำหนดระดับนัยสำคัญที่ 0.05

3.5 กระทำซ้ำ 500 ครั้ง ในแต่ละกรณี

3.6 บันทึกจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง (H_0)

4. การเปรียบเทียบประสิทธิภาพอำนาจการทดสอบ

ในการเปรียบเทียบประสิทธิภาพของสถิติทดสอบ จะทำการวัดค่าอำนาจการทดสอบ โดยในทางสถิติถ้าอำนาจการทดสอบของตัวสถิติยิ่งสูง แสดงว่าการทดสอบยิ่งมีประสิทธิภาพดี

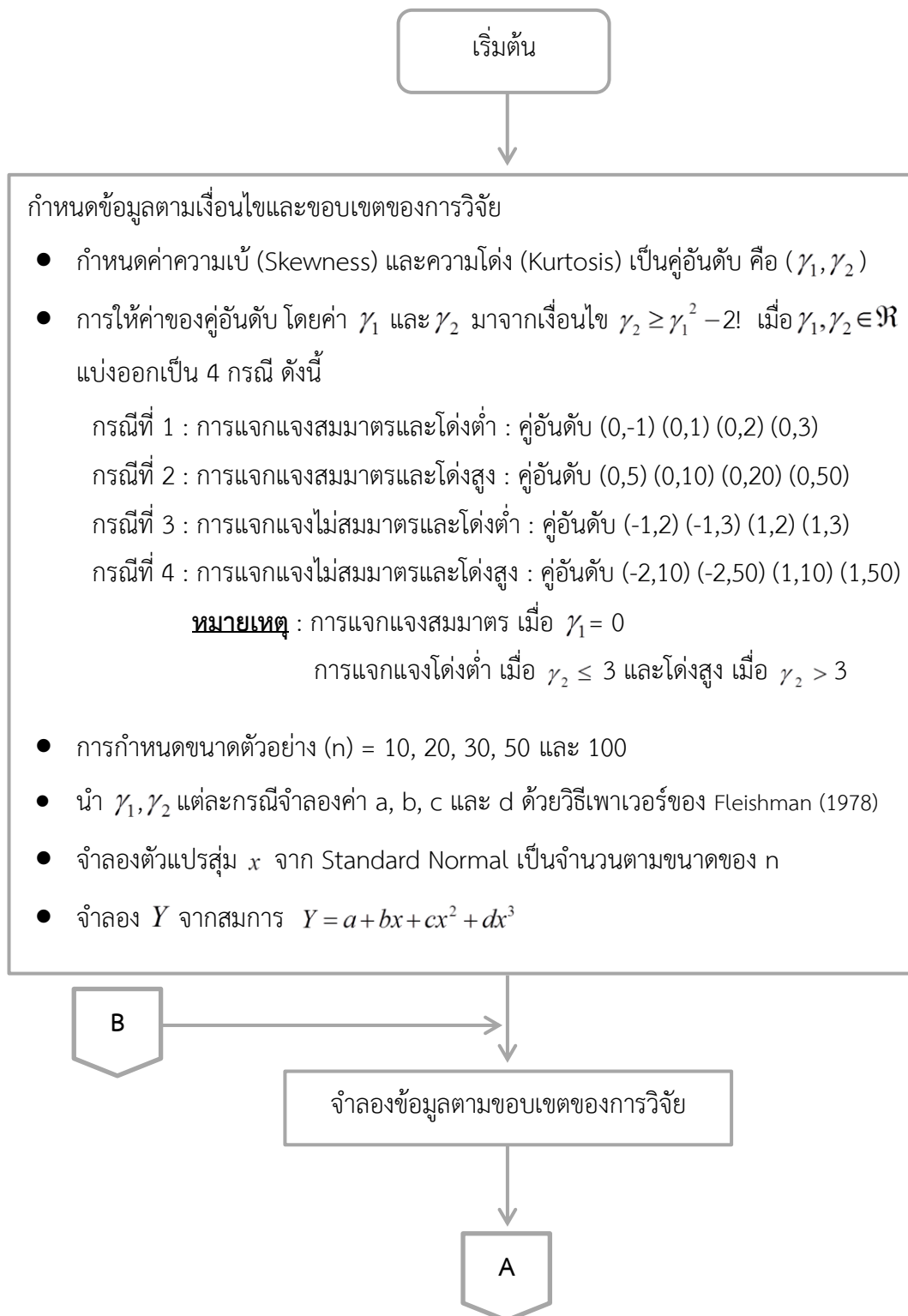
ทำการหาค่าอำนาจการทดสอบ (Power of Test) จะใช้ผลในการบันทึกจำนวนครั้งที่ปฏิเสธสมมติฐานว่าง (H_0) (ข้อ 3.6) ทารด้วย 500 เนื่องจากการวิจัยนี้ผู้วิจัยทำการจำลองข้อมูล 500 ครั้ง ซึ่งเป็นจำนวน H_0 ที่เป็นเท็จทั้งหมด

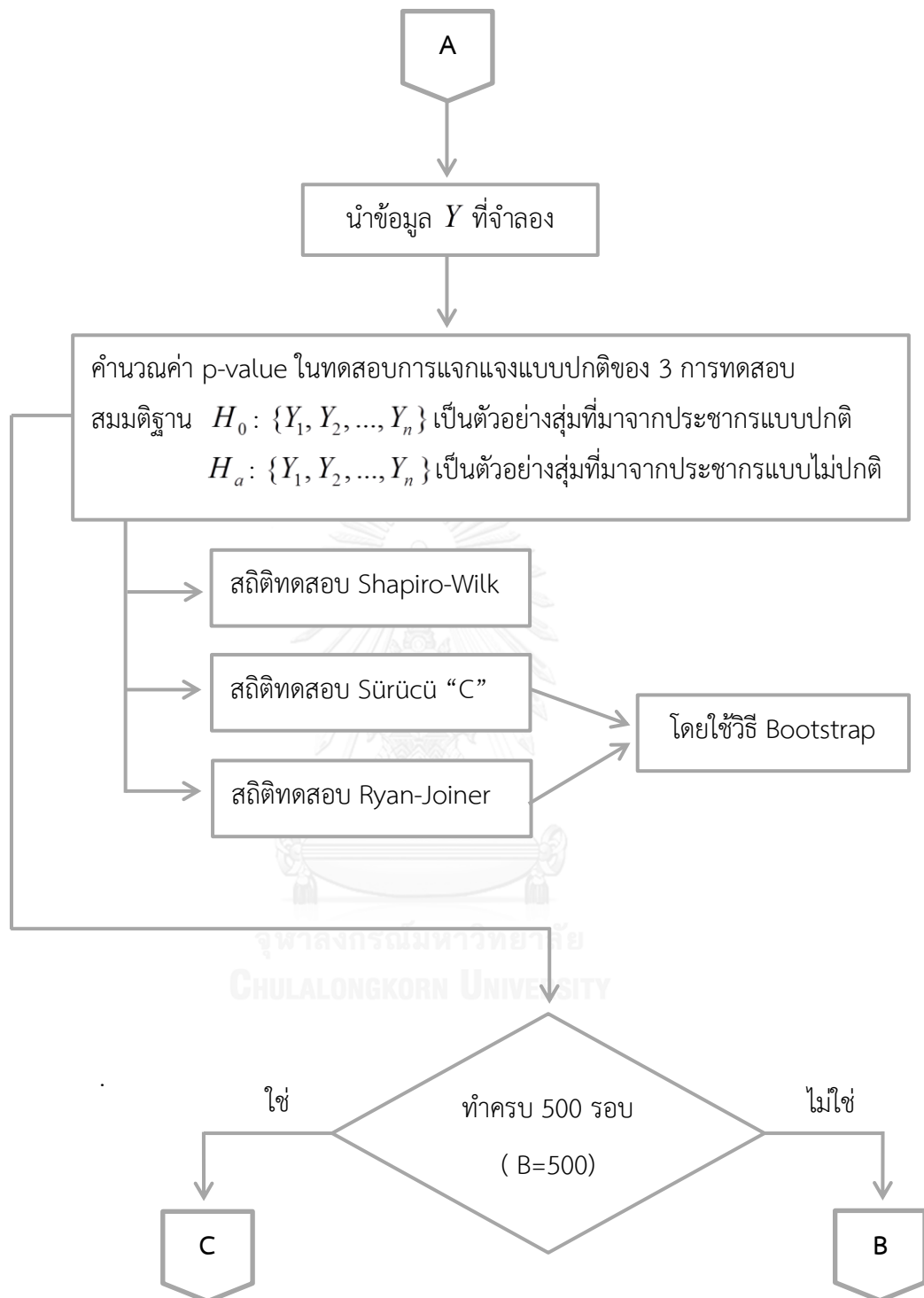
5. การวิเคราะห์และสรุปผลเกี่ยวกับความเหมาะสมของสถิติทดสอบ

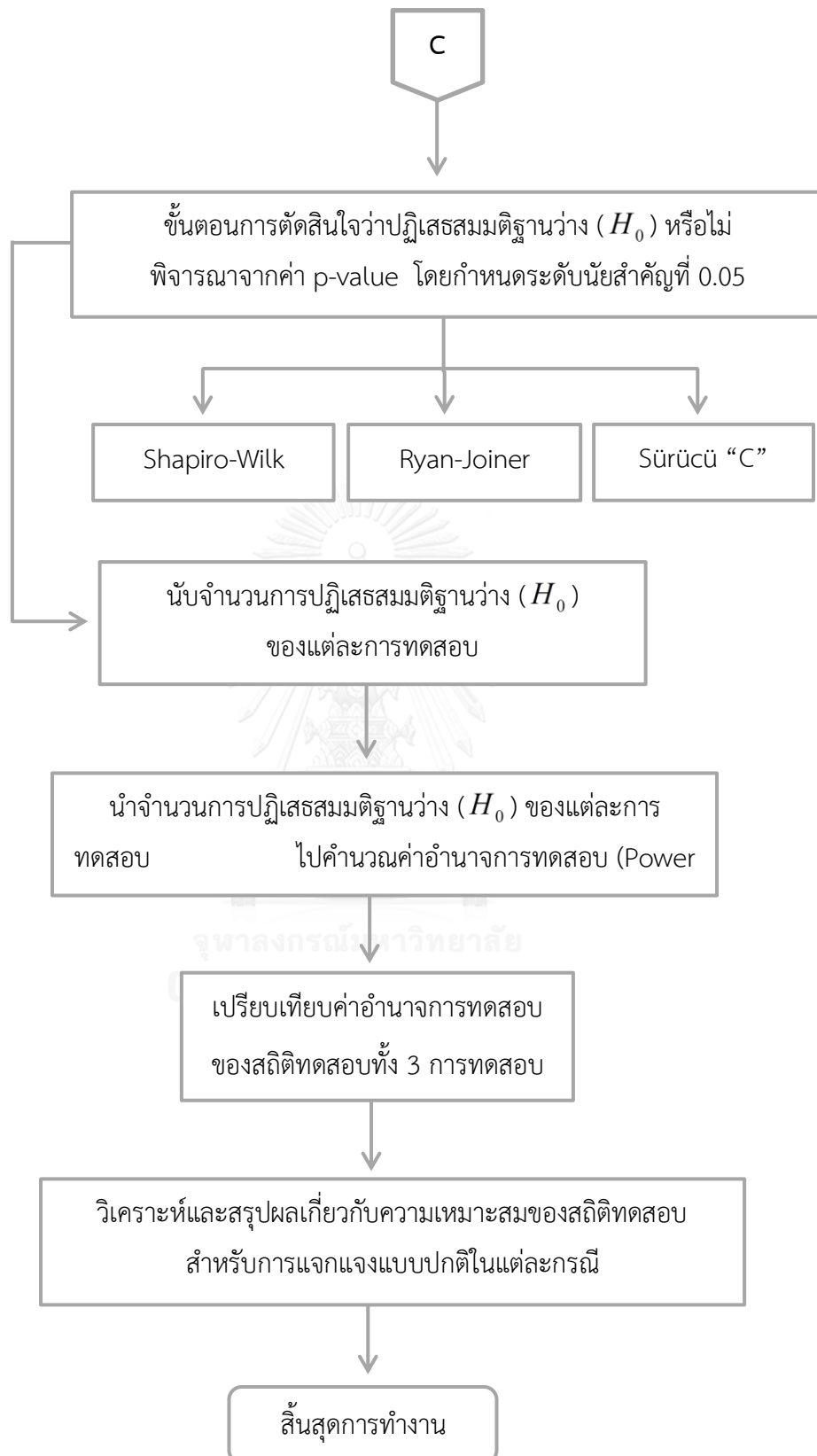
นำค่าอำนาจการทดสอบมาวิเคราะห์ประสิทธิภาพของสถิติทดสอบ เพื่อสรุปหาสถิติทดสอบที่เหมาะสมสำหรับการแจกแจงแบบปกติของข้อมูลในแต่ละกรณี



3.3 ขั้นตอนการทำงานของโปรแกรม







บทที่ 4

ผลการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบการแจกแจงแบบปกติวิธีใดมีประสิทธิภาพมากที่สุด โดยสถิติทดสอบที่ทำการศึกษาคือ

1. สถิติทดสอบ Shapiro-Wilk
2. สถิติทดสอบ Ryan-Joiner
3. สถิติทดสอบ Sürücü “C”

ในการสร้างข้อมูลการจำลอง จะศึกษาการสร้างข้อมูลจากกรณีความเบ้ (Skewness) และความโด่ง (Kurtosis) รูปแบบต่าง ๆ ด้วยวิธีเพาเวอร์ของ Fleishman (1978) ซึ่งสามารถสร้างข้อมูลภายใต้การแจกแจงแบบต่างๆ ทั้งหมด 16 ลักษณะ ประกอบด้วยการแจกแจงสมมาตรและโด่งต่ำ 4 ลักษณะ การแจกแจงสมมาตรและโด่งสูง 4 ลักษณะ การแจกแจงไม่สมมาตรและโด่งต่ำ 4 ลักษณะ และการแจกแจงไม่สมมาตรและโด่งสูงอีก 4 ลักษณะ ซึ่งขนาดตัวอย่างที่ทำการศึกษามี 2 กลุ่มได้แก่กลุ่มตัวอย่างขนาดเล็กคือ 10, 20 และ 30 และกลุ่มตัวอย่างขนาดใหญ่คือ 50 และ 100

การพิจารณาผลการวิเคราะห์ จะกำหนดอักษรย่อและสัญลักษณ์ต่างๆที่ปรากฏอยู่ในผลการวิจัยทั้งในตารางและข้อความต่าง ๆ แทนความหมายดังนี้

n	แทน ขนาดของตัวอย่าง
W	แทน สถิติทดสอบ Shapiro-Wilk
RJ	แทน สถิติทดสอบ Ryan-Joiner
C	แทน สถิติทดสอบ Sürücü “C”
γ_1	แทน ค่าของความเบ้ (Skewness)
γ_2	แทน ค่าของความโด่ง (Kurtosis)
Fleishman's	แทน สมการที่ใช้คำนวณหาค่า a, b, c และ d สำหรับวิธีเพาเวอร์
Power	แทน อำนาจการทดสอบ (Power of Test)

สำหรับงานวิจัยนี้จะนำเสนอโดยแบ่งผลการวิจัยออกเป็น 4 ส่วน ดังต่อไปนี้

4.1 ผลการจำลองจากวิธีเพาเวอร์ของ Fleishman (1978)

ในการจำลองข้อมูลด้วยวิธีเพาเวอร์สำหรับการทดสอบการแจกแจงแบบปกตินั้น เป็นการนำค่าความเบ้และค่าความโด่งเพื่อจำลองหาค่า a, b, c และ d ในสมการ Fleishman's ก่อน แล้วนำค่า a, b, c และ d ดังกล่าวนี้ไปจำลองข้อมูลสำหรับการทดสอบการแจกแจงแบบปกติ แบ่งได้ 4 กรณีดังนี้

- การแจกแจงลักษณะสมมาตรและโด่งต่ำ

ตารางที่ 4. 1 แสดงค่า a, b, c และ d ของคู่อันดับ (0,-1) (0,1) (0,2) (0,3)

γ_1	γ_2	a	b	c	d
0	-1	0.000000000	1.2210090	0.000000000	-0.08015816
0	1	0.000000000	0.9029766	0.000000000	0.03135643
0	2	0.000000001	0.8356646	-0.000000001	0.05205739
0	3	0.000000000	0.7823562	0.000000000	0.06790456

- การแจกแจงลักษณะสมมาตรและโด่งสูง

ตารางที่ 4. 2 แสดงค่า a, b, c และ d ของคู่อันดับ (0,5) (0,10) (0,20) (0,50)

γ_1	γ_2	a	b	c	d
0	5	0.000000000	0.69800478	0.000000000	0.09207639
0	10	0.000000000	0.54668238	0.000000000	0.13293821
0	20	0.000000000	0.33871787	0.000000000	0.18446114
0	50	0.000000001	-0.08544226	-0.000000001	0.27491009

- การแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ

ตารางที่ 4. 3 แสดงค่า a, b, c และ d ของคู่อันดับ (-1,2) (-1,3) (1,2) (1,3)

γ_1	γ_2	a	b	c	d
-1	2	0.1472108	0.9047583	-0.1472108	0.02386092
-1	3	0.1283967	0.8322164	-0.1283967	0.04803206
1	2	-0.1472108	0.9047584	0.1472108	0.02386091
1	3	-0.1283968	0.8322165	0.1283968	0.04803200

● การแจกแจงลักษณะไม่สมมาตรและโด่งสูง

ตารางที่ 4. 4 แสดงค่า a, b, c และ d ของคู่อันดับ (-2,10) (-2,50) (1,10) (1,50)

γ_1	γ_2	a	b	c	d
-2	10	0.20344876	0.63044360	-0.20344876	0.09874214
-2	50	0.10943102	-0.06917559	-0.10943102	0.26867306
1	10	-0.08771825	0.56426070	0.08771825	0.12620922
1	50	-0.05364415	-0.08143568	0.05364415	0.27339827

จากตาราง 4.1 – 4.4 แสดงผลลัพธ์ค่า a, b, c และ d จากการจำลองค่าความเบ้และความโด่งใน โดยค่าดังกล่าวจะนำไปสร้างข้อมูลในสมการ Fleishman เพื่อเป็นข้อมูลสำหรับการวิเคราะห์คุณลักษณะของการแจกแจงด้วยกราฟ และนำข้อมูลดังกล่าวนี้ไปใช้ในการทดสอบการแจกแจงแบบปกติ เพื่อเปรียบเทียบประสิทธิภาพด้วยอำนาจการทดสอบของสถิติทดสอบแต่ละวิธีต่อไป

4.2 ผลการวิเคราะห์ข้อมูลที่จำลองด้วยกราฟ

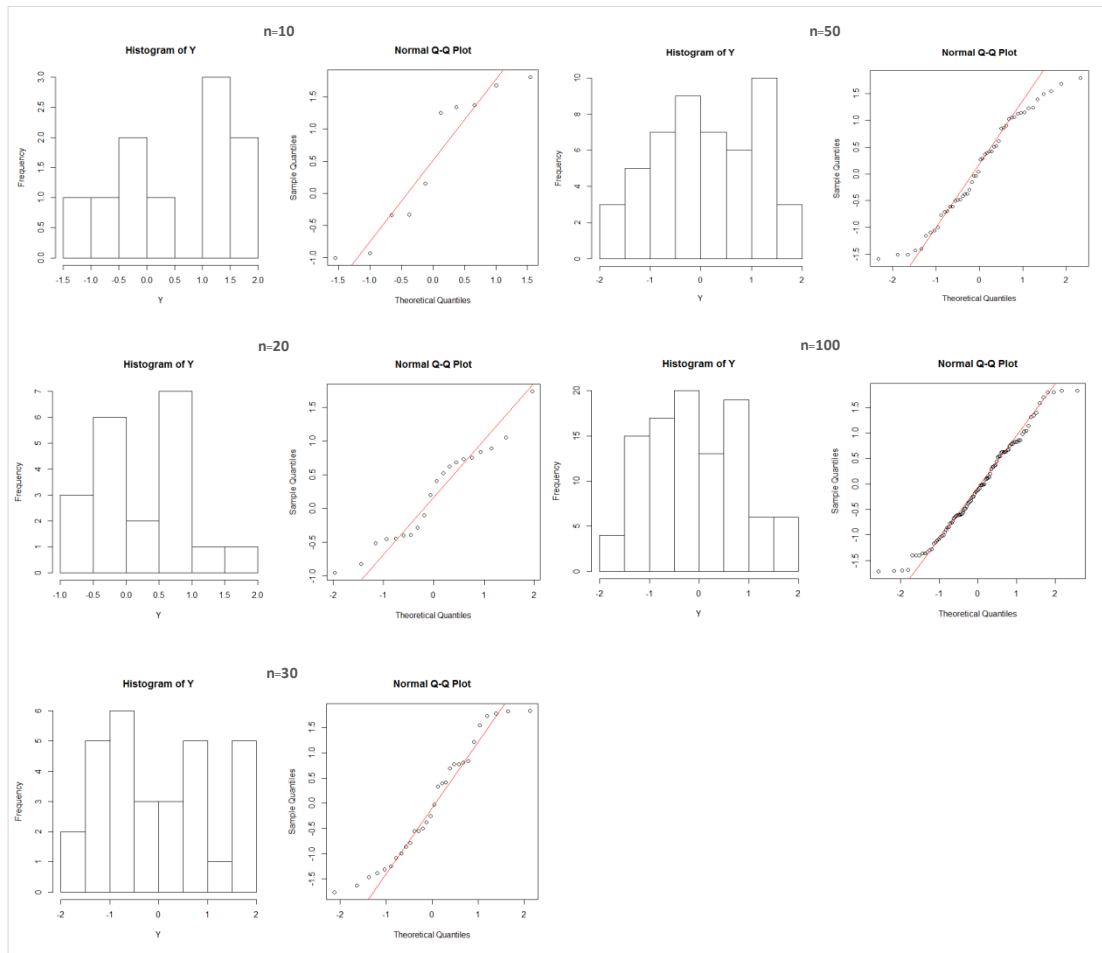
ในทางสถิติสำหรับการวิเคราะห์ข้อมูลเบื้องต้นเพื่อดูคุณลักษณะการแจกแจงของข้อมูลนั้น ขั้นแรกของการวิเคราะห์ (Asghar & Saleh, 2012) คือการวัดจากสายตา โดยดูกราฟของข้อมูลว่ามีลักษณะใดอย่างคร่าว ๆ เพื่อตรวจสอบว่าลักษณะรูปร่างนั้นเป็นการแจกแจงเป็นแบบปกติหรือไม่ จากนั้นจึงนำข้อมูลดังกล่าวไปทดสอบด้วยสถิติทดสอบต่อไป ในงานวิจัยนี้ได้นำข้อมูลจากการจำลองของค่าความเบ้และความโด่งลักษณะต่าง ๆ ด้วยวิธีเพาเวอร์ ได้ผลลัพธ์คือ ค่า a, b, c และ d โดยนำค่าดังกล่าวนี้ไปจำลองในสมการ Fleishman's เพื่อสร้าง Histogram และ Normal Q-Q Plot ดังนี้

กราฟที่จำลองในแต่ละลักษณะของการแจกแจงแต่ละแบบจะทำการจำลองด้วยขนาดตัวอย่างที่แตกต่างกัน ซึ่งแบ่งกราฟของการจำลองข้อมูลจากการแจกแจงเป็น 4 ลักษณะ แต่ละลักษณะแบ่งออกเป็นความเบ้และความโด่ง 4 แบบ แต่ละชุดของความเบ้และความโด่งจะนำไปสร้างกราฟด้วยจำนวนตัวอย่างที่แตกต่างกัน โดยที่ $n=10, 20, 30, 50$ และ 100 มีผลการวิเคราะห์ ดังนี้

- 4.2.1 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งต่ำ
- 4.2.2 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งสูง
- 4.2.3 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ
- 4.2.4 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งสูง

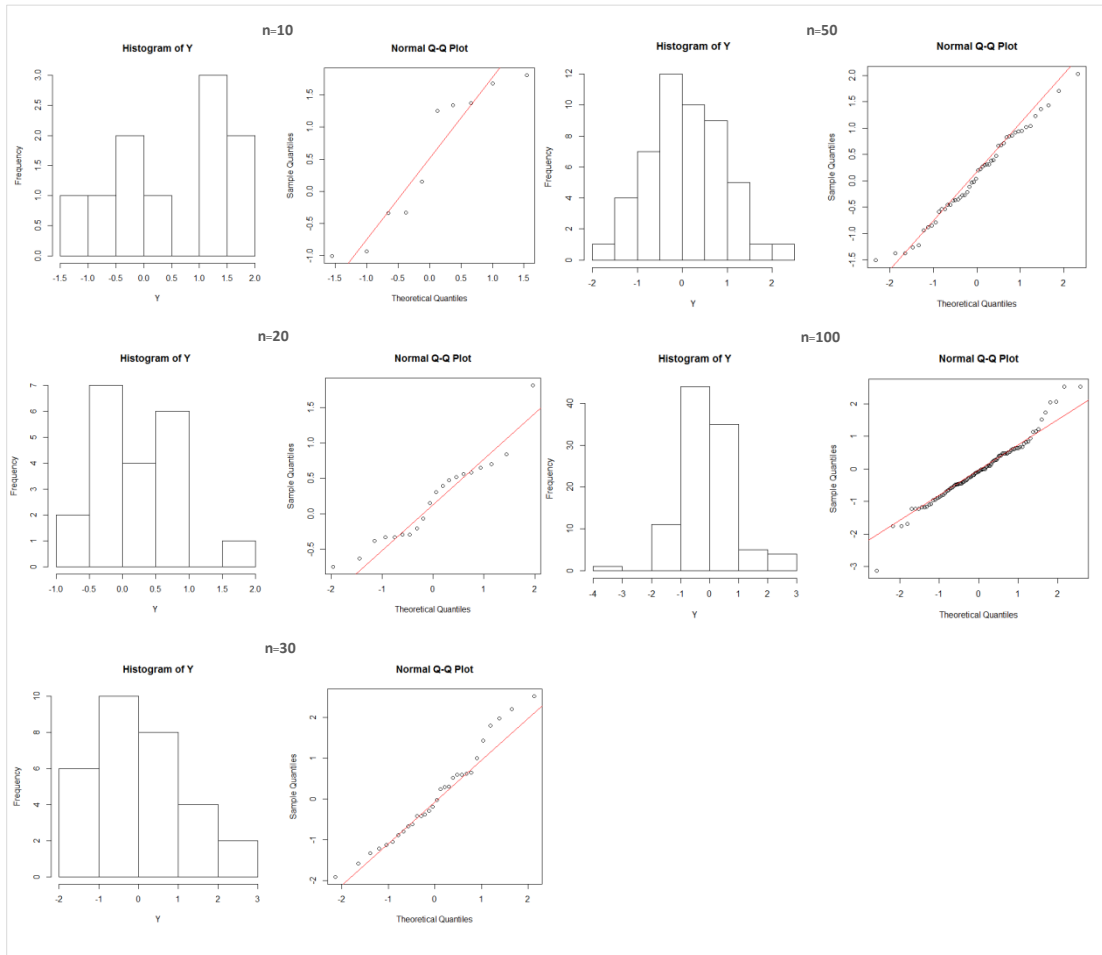
4.2.1 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งต่ำ

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=-1$



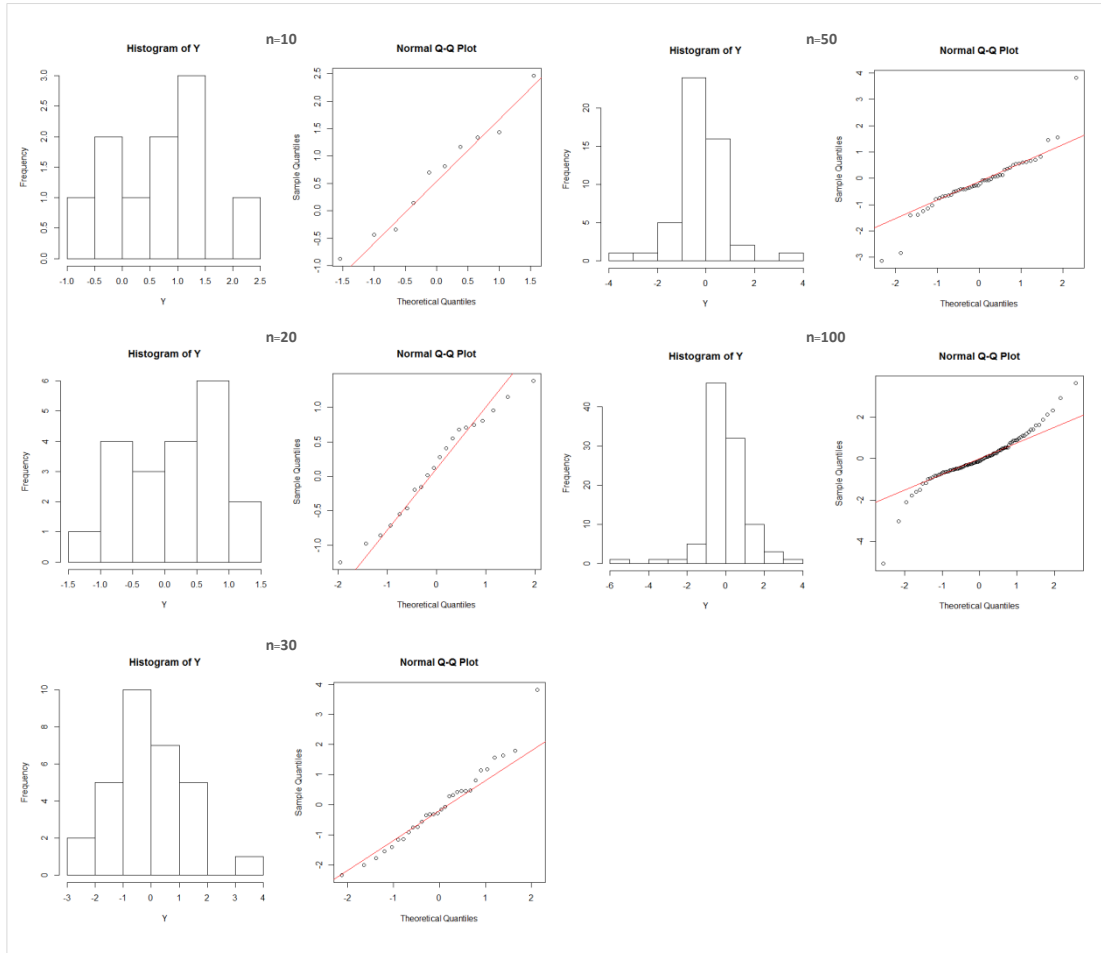
ภาพที่ 4. 1 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=-1$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=1$



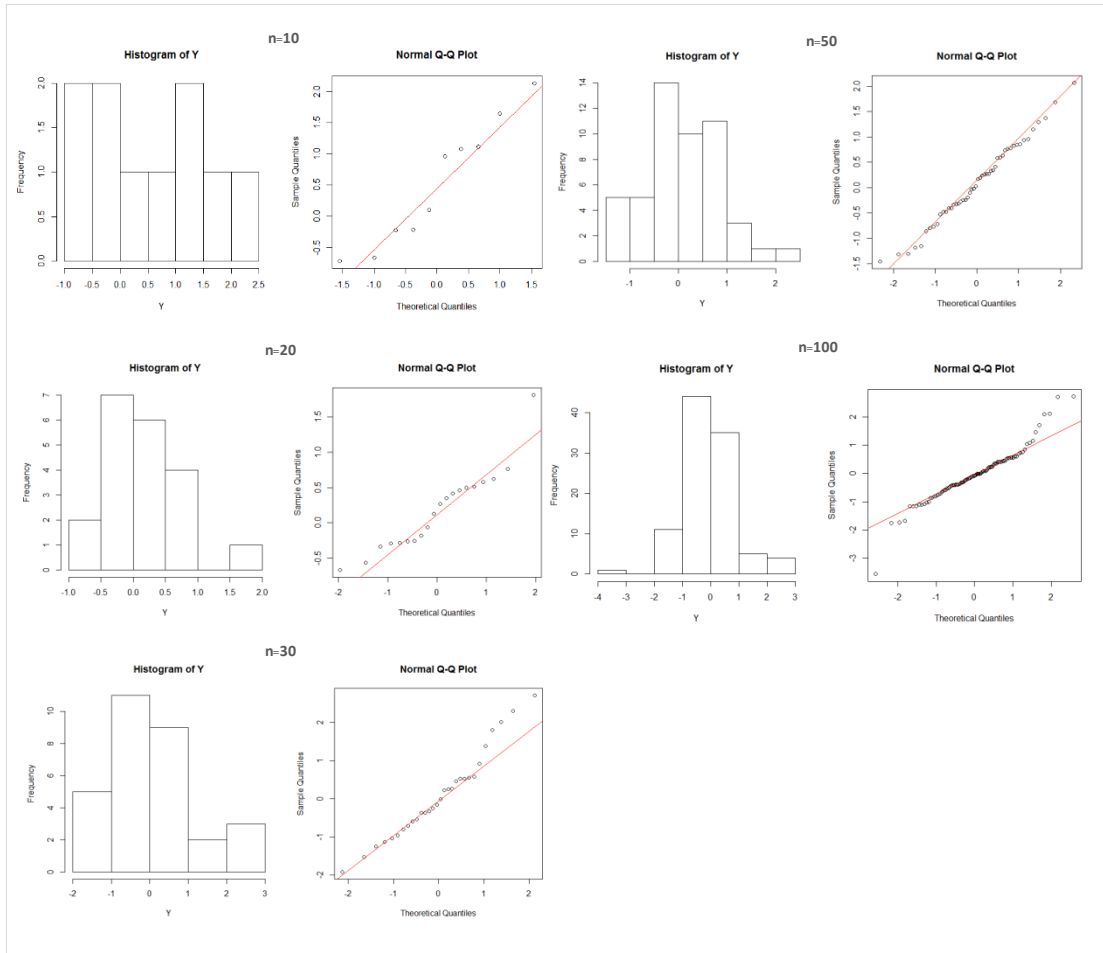
ภาพที่ 4. 2 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=1$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=2$



ภาพที่ 4.3 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

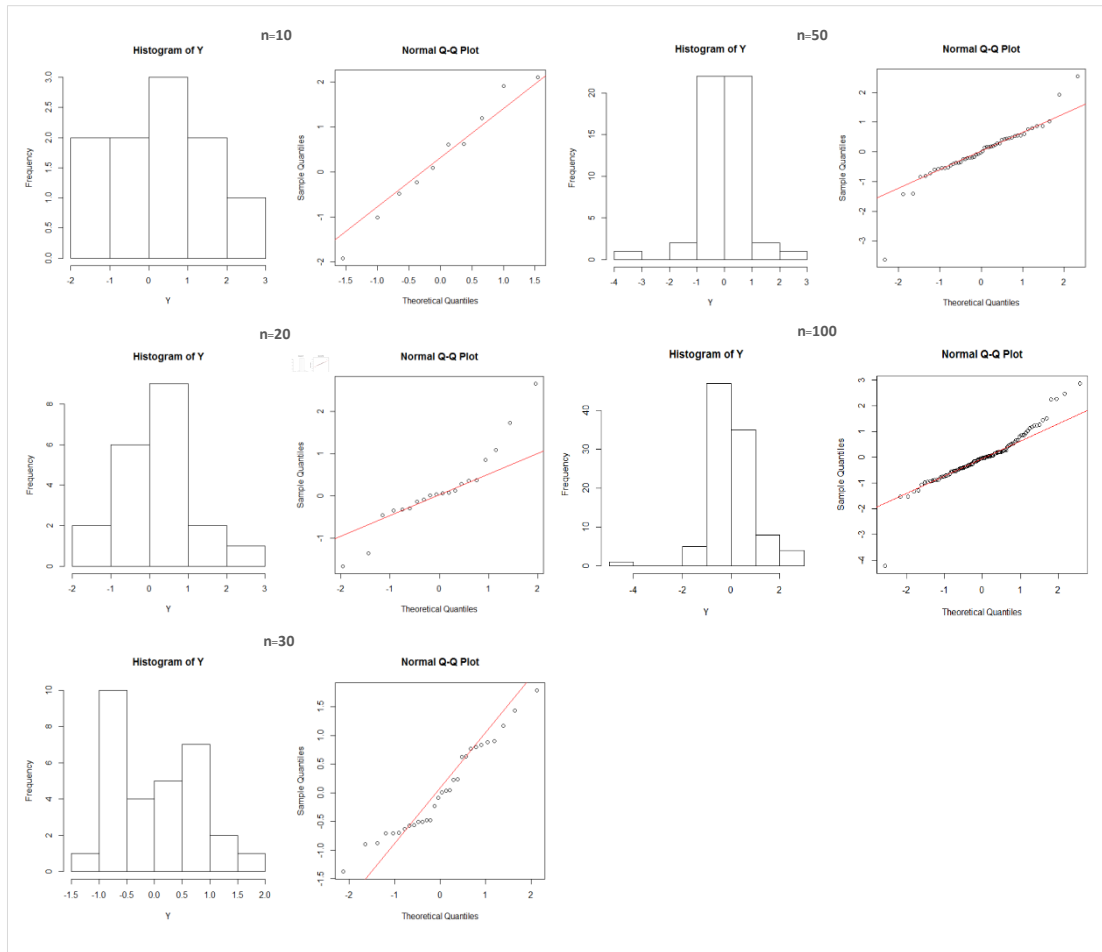
- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=3$



ภาพที่ 4. 4 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

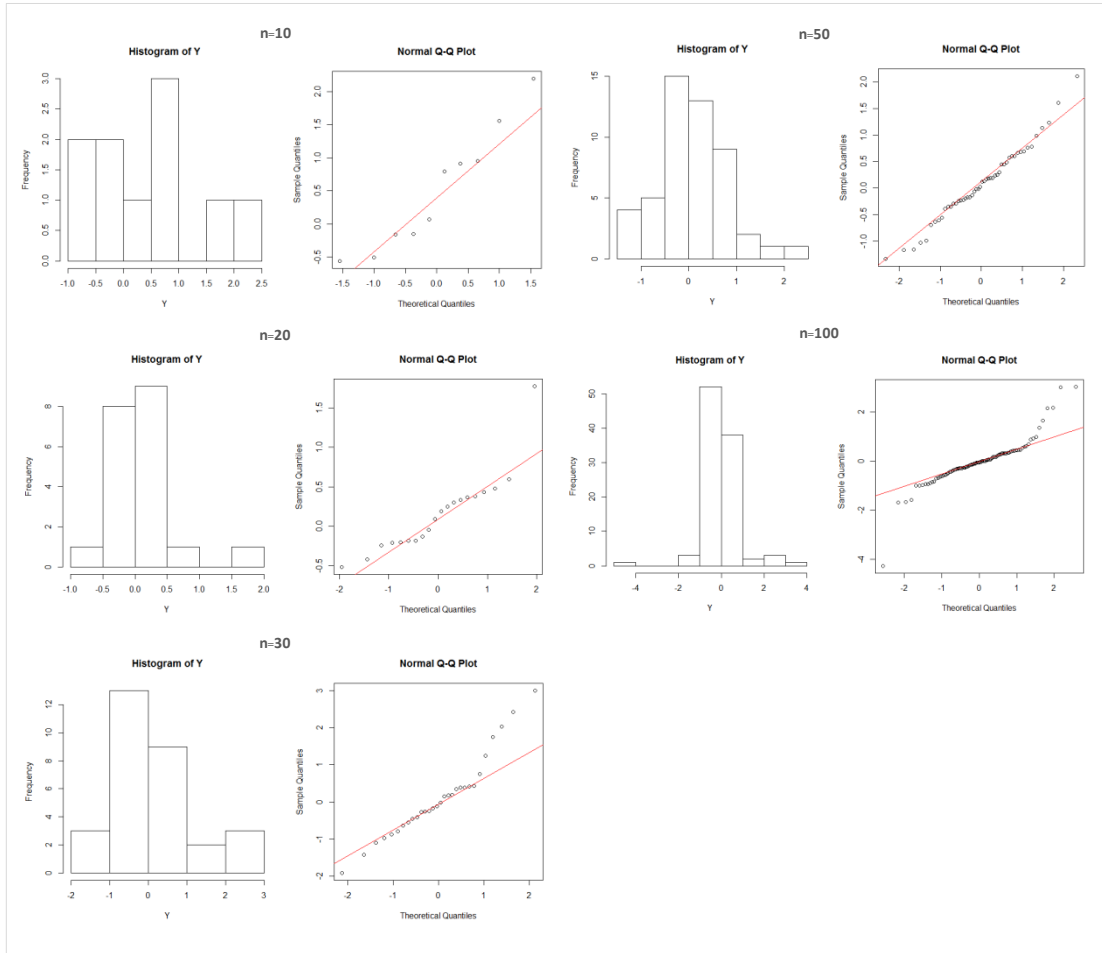
4.2.2 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะสมมาตรและโด่งสูง

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=5$



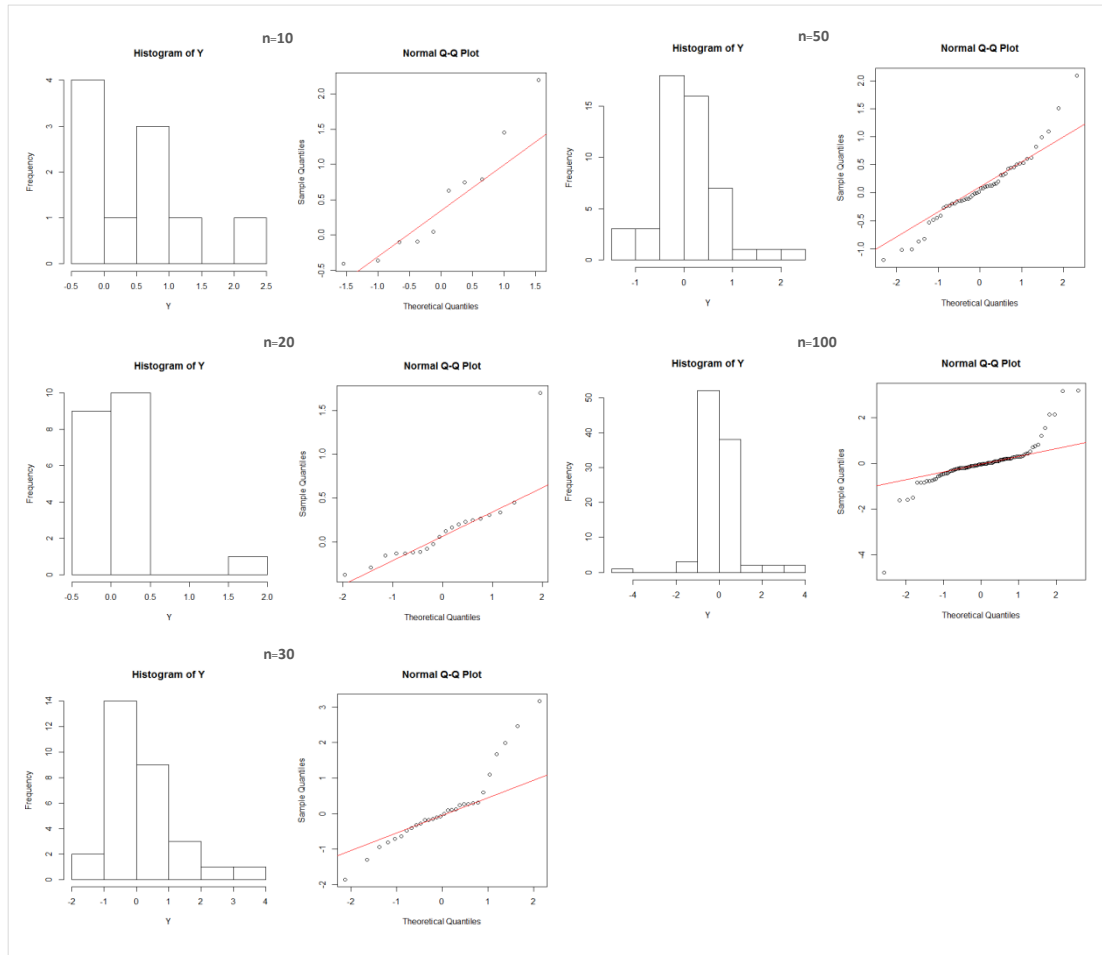
ภาพที่ 4.5 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=5$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=10$



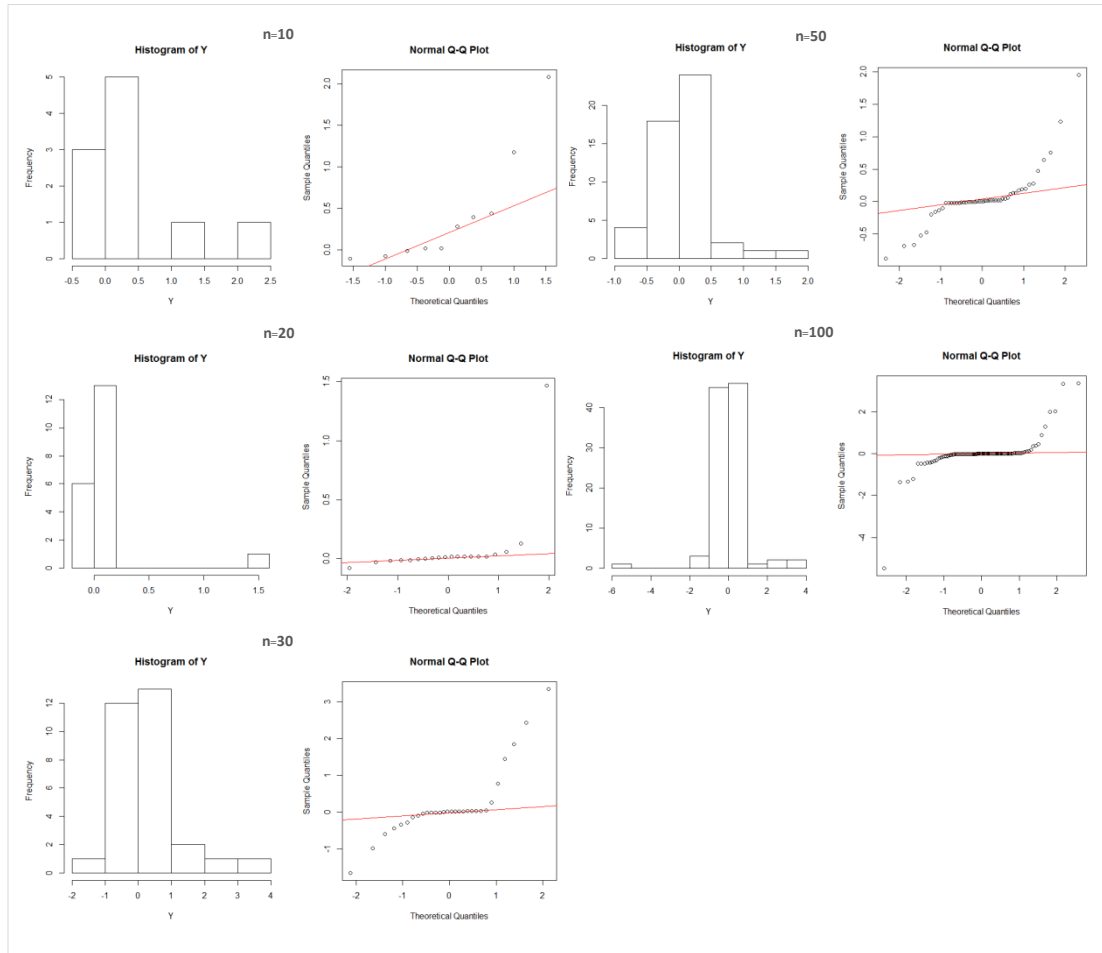
ภาพที่ 4. 6 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=20$



ภาพที่ 4. 7 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=20$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

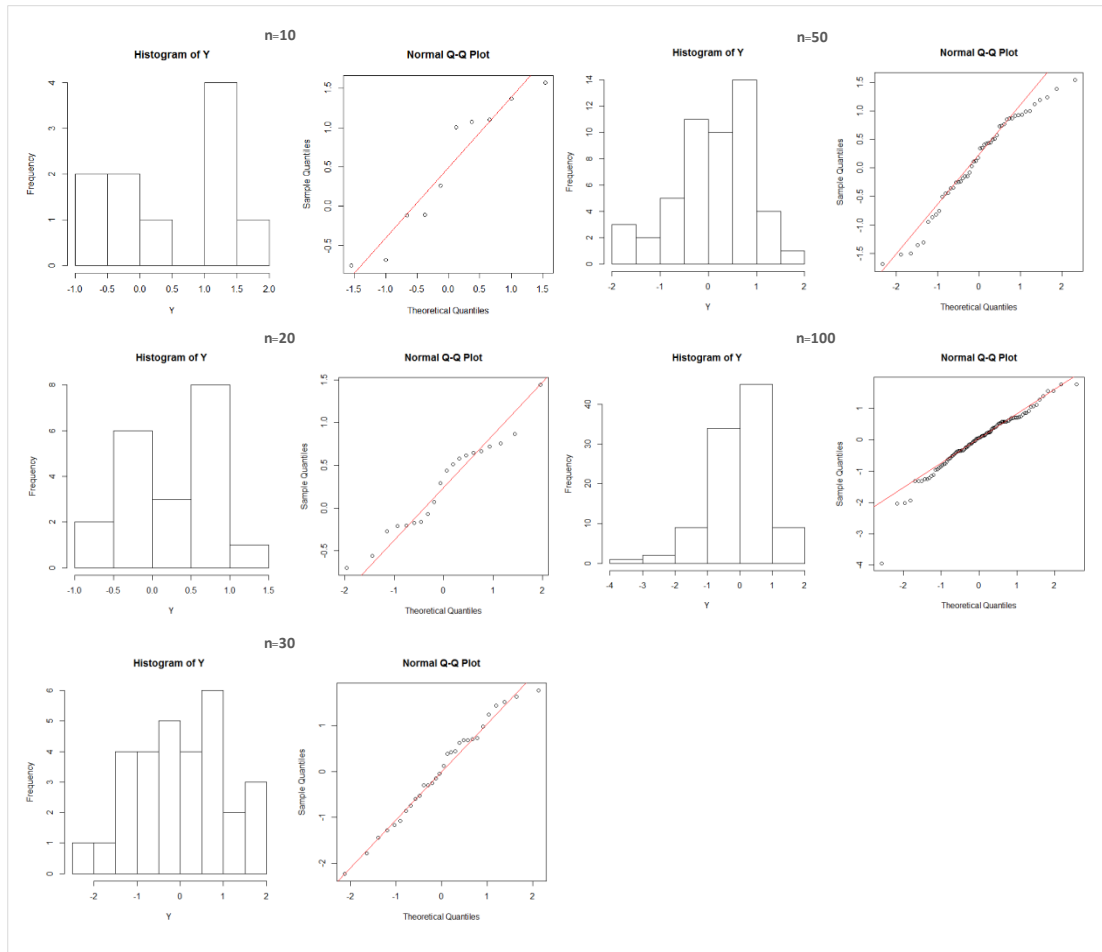
- การแจกแจงลักษณะ $\gamma_1=0, \gamma_2=50$



ภาพที่ 4. 8 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=0, \gamma_2=50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

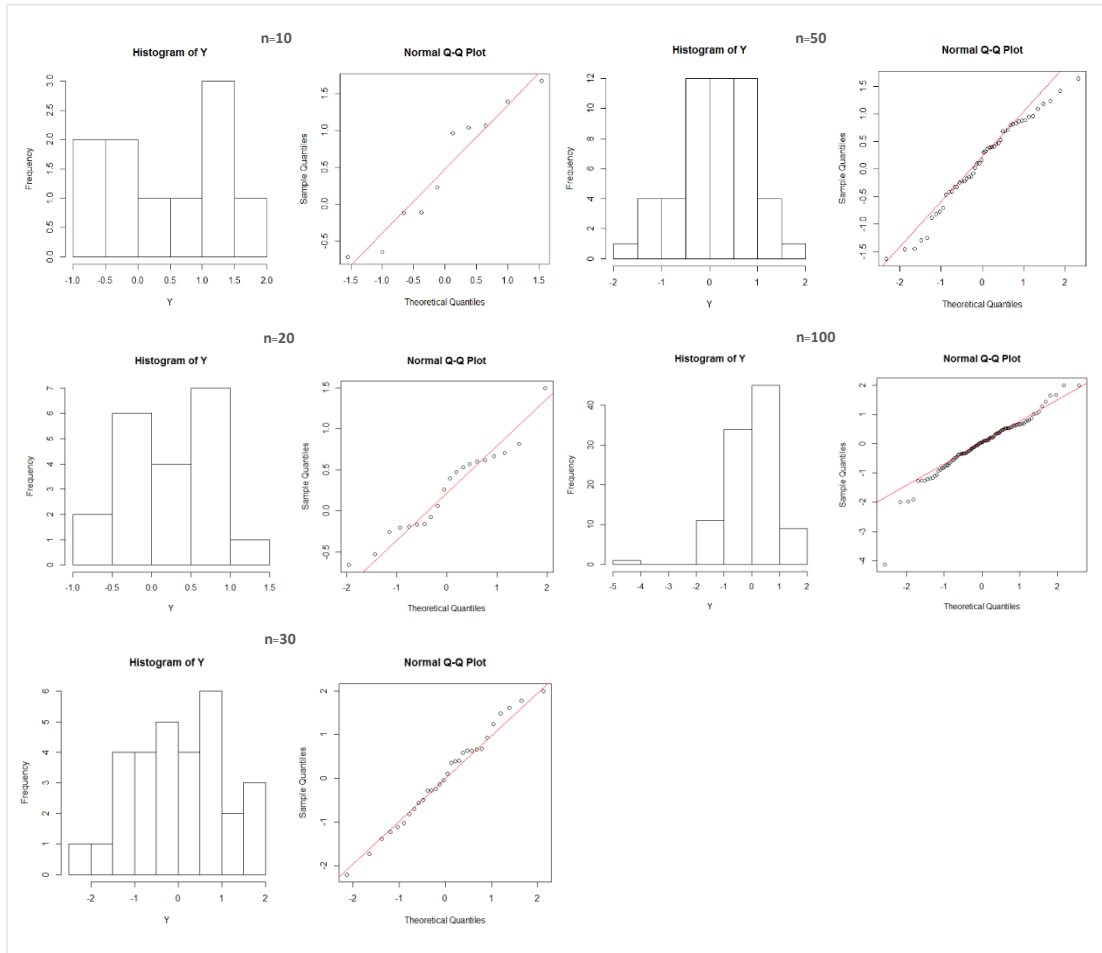
4.2.3 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ

- การแจกแจงลักษณะ $\gamma_1 = -1$, $\gamma_2 = 2$



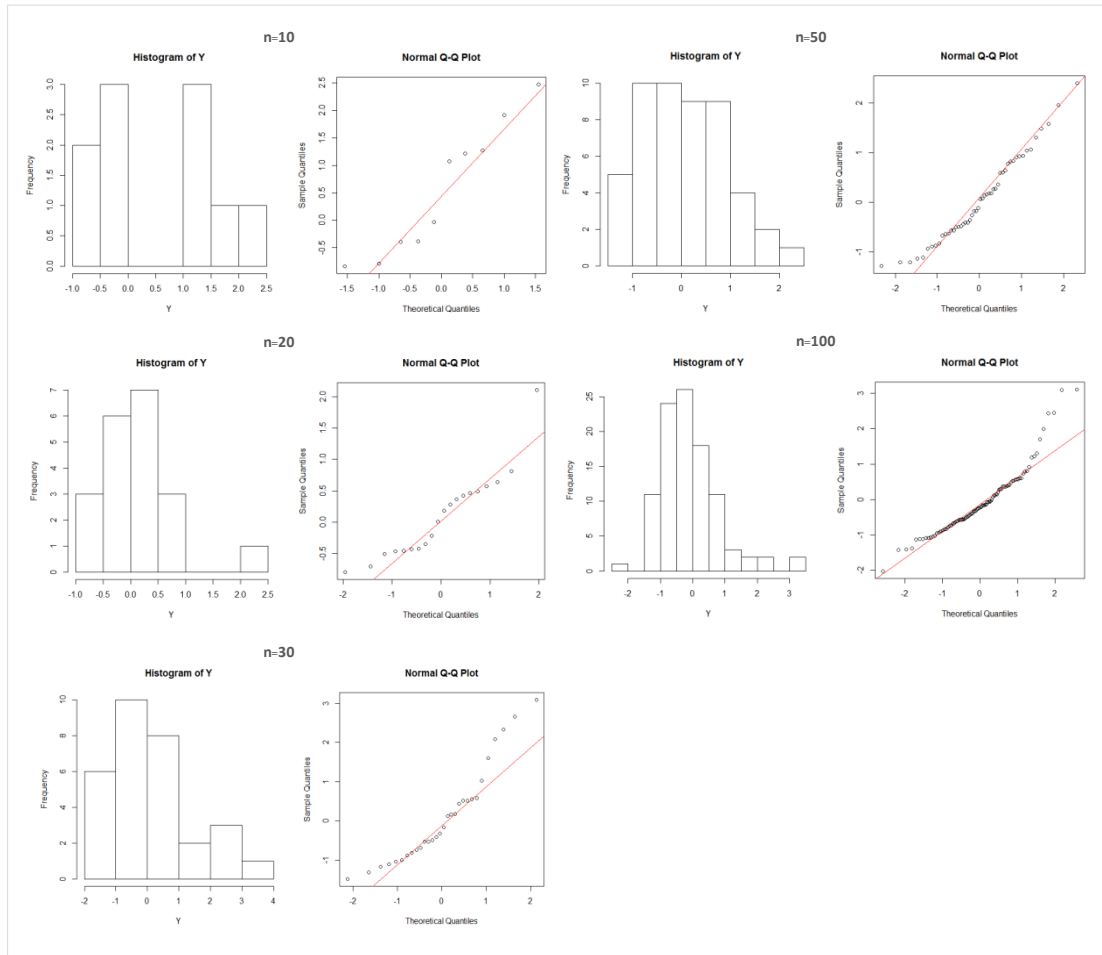
ภาพที่ 4. 9 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1 = -1$, $\gamma_2 = 2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1 = -1, \gamma_2 = 3$



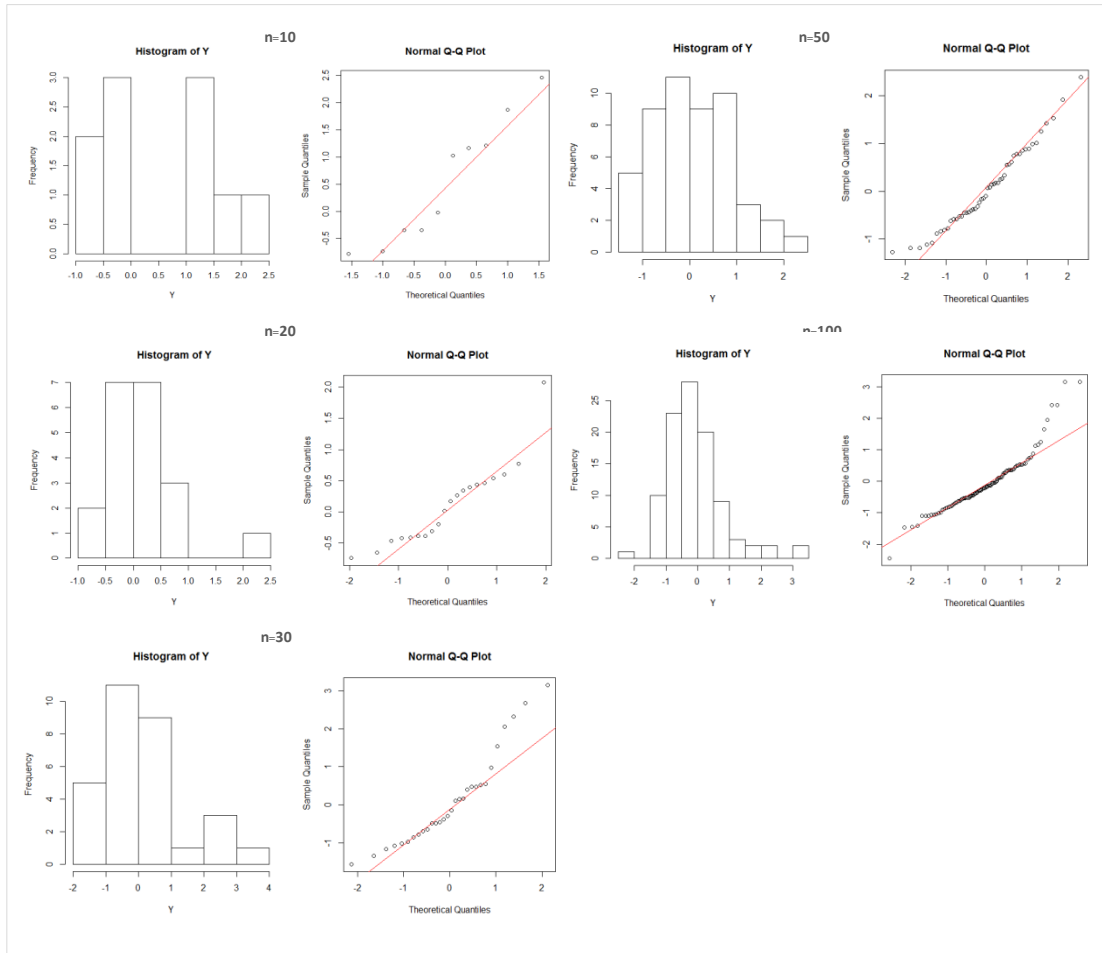
ภาพที่ 4. 10 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1 = -1, \gamma_2 = 3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=1, \gamma_2=2$



ภาพที่ 4.11 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=2$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

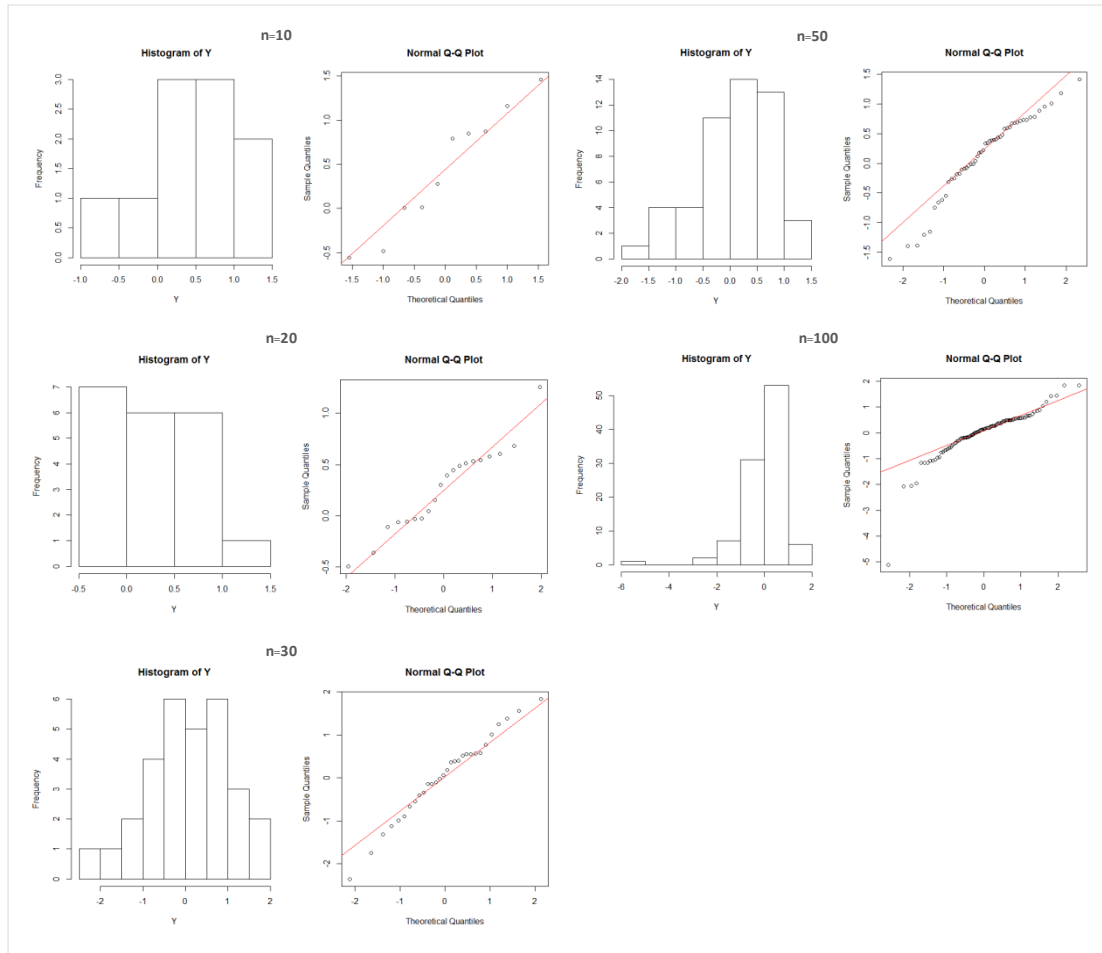
- การแจกแจงลักษณะ $\gamma_1=1, \gamma_2=3$



ภาพที่ 4. 12 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=3$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

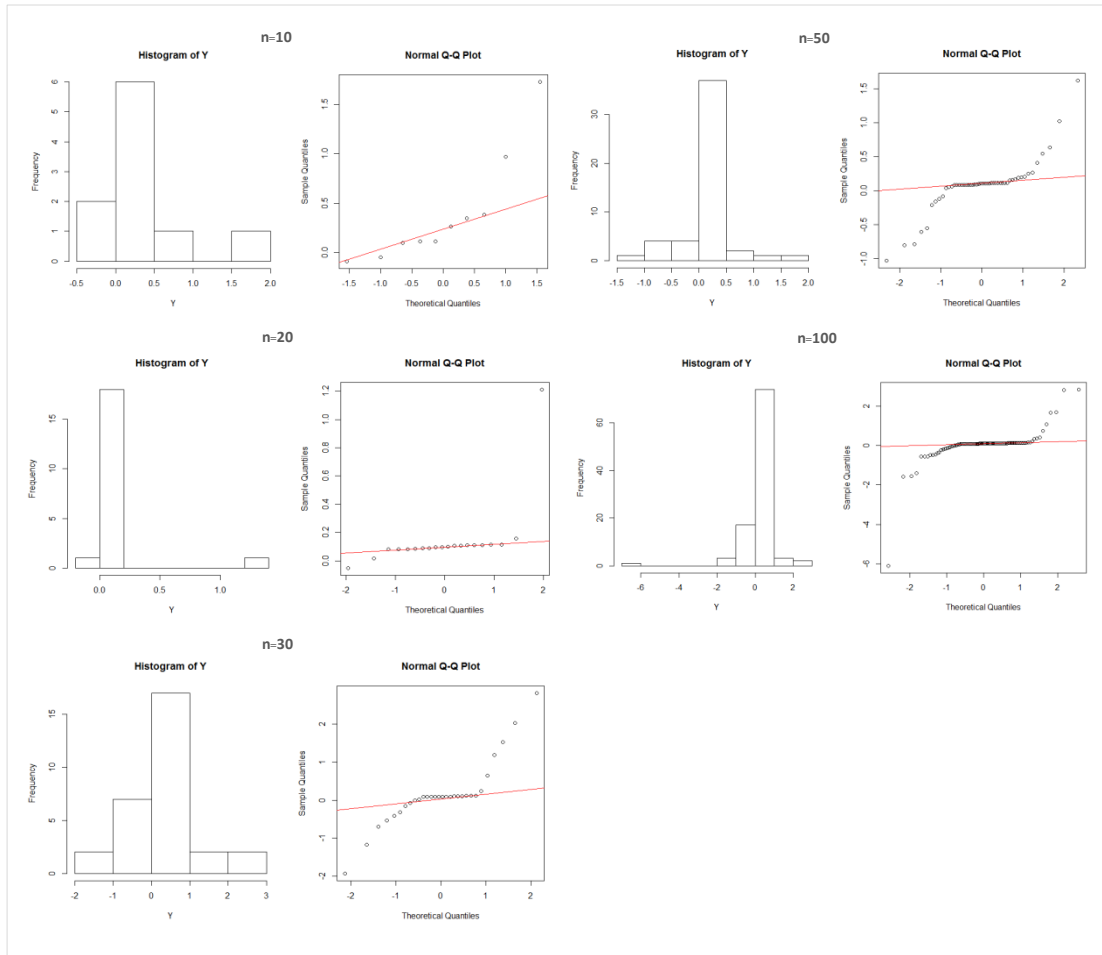
4.2.4 กราฟของการจำลองข้อมูลจากการแจกแจงลักษณะไม่สมมาตรและโด่งสูง

- การแจกแจงลักษณะ $\gamma_1 = -2$, $\gamma_2 = 10$



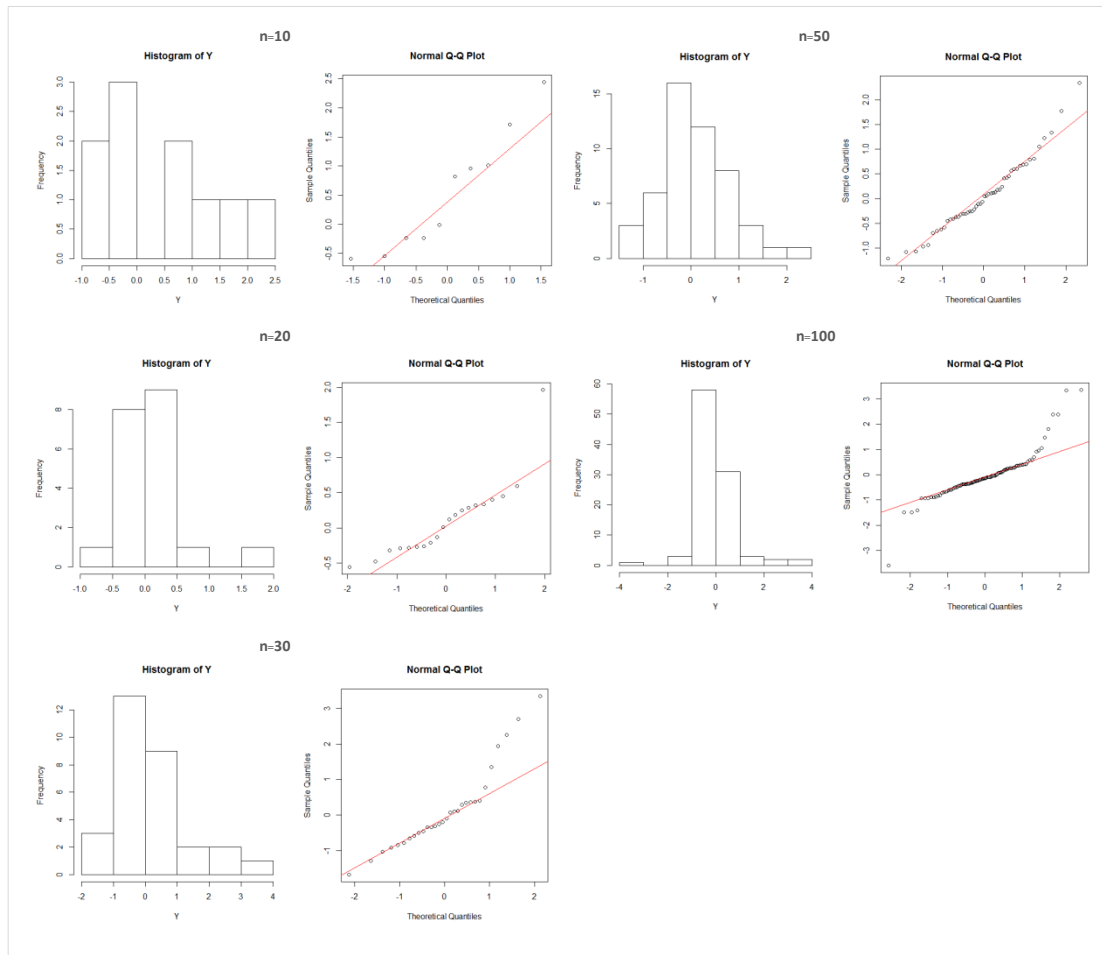
ภาพที่ 4.13 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1 = -2$, $\gamma_2 = 10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1 = -2, \gamma_2 = 50$



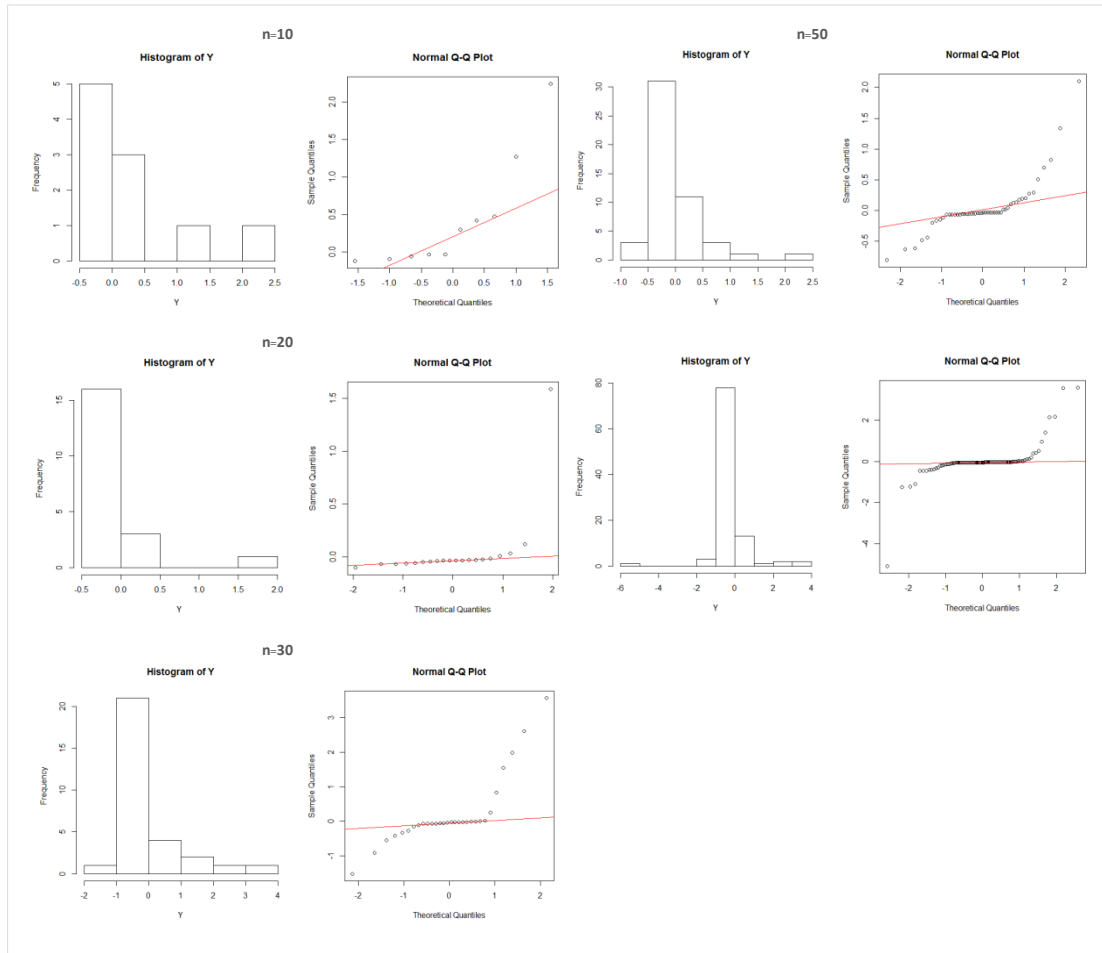
ภาพที่ 4.14 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1 = -2, \gamma_2 = 50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=1, \gamma_2=10$



ภาพที่ 4.15 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=10$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

- การแจกแจงลักษณะ $\gamma_1=1, \gamma_2=50$



ภาพที่ 4.16 แสดง Histogram และ Normal Q-Q Plot ของการแจกแจงลักษณะ $\gamma_1=1, \gamma_2=50$ ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

จากภาพที่ 4.1-4.16 Histogram และ Normal Q-Q Plot ของการแจกแจง 4 กรณี ประกอบด้วย ลักษณะการแจกแจงของความเบ้และความโด่งที่แตกต่างกันทั้งหมด 16 ลักษณะ ข้อมูลแต่ละลักษณะ จำลองกราฟด้วยวิธีไพเพาเวอร์ตามจำนวนขนาดตัวอย่างเป็น 10, 20, 30, 50 และ 100 จากกราฟของทุกกรณี พบว่าข้อมูลมีลักษณะที่ไม่เป็นปกติ โดยกราฟมีลักษณะเบ้ซ้าย เบ้ขวา ซึ่งไม่สมมาตรทุกกรณี แต่เมื่อดูภาพรวมของกราฟแต่ละกรณีสำหรับชุดขนาดตัวอย่างที่แตกต่างกัน พบว่าเมื่อจำลองข้อมูล ในจำนวนขนาดตัวอย่างที่เพิ่มขึ้น Histogram และ Normal Q-Q Plot มีลักษณะที่เข้าใกล้ความเป็นปกติมากยิ่งขึ้น

4.3 ผลการเปรียบเทียบอำนาจการทดสอบ (Power of Test) ของสถิติทดสอบการแจกแจงปกติ ระหว่างสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C”

ในส่วนนี้ผู้วิจัยต้องการเปรียบเทียบประสิทธิภาพของสถิติทดสอบการแจกแจงปกติ จะศึกษาจากอำนาจการทดสอบของแต่ละสถิติทดสอบเมื่อมีการเปลี่ยนแปลงค่าของความเบ้และความโด่งของข้อมูล การพิจารณาค่าอำนาจการทดสอบนั้นจะพิจารณาจากสถิติทดสอบ 3 วิธี คือ สถิติทดสอบ Shapiro-Wilk สถิติทดสอบ Ryan-Joiner และสถิติทดสอบ Sürücü “C” มีขนาดตัวอย่างตั้งแต่ 10, 20, 30, 50 และ 100 ที่ระดับนัยสำคัญ 0.05

โดยการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ 3 การทดสอบดังกล่าวข้างต้นนี้ ได้ทำการเปรียบเทียบในกลุ่มภายใต้เงื่อนไขของการแจกแจงที่มีลักษณะต่าง ๆ 4 กรณี ดังนี้

4.3.1 การแจกแจงลักษณะสมมาตรและโด่งต่ำ

คือ การแจกแจงที่มีความเบ้และความโด่ง เป็นคู่อันดับ (0,-1) (0,1) (0,2) (0,3) แสดงผลในตารางที่ 4.3.1

4.3.2 การแจกแจงลักษณะสมมาตรและโด่งสูง

คือ การแจกแจงที่มีความเบ้และความโด่ง เป็นคู่อันดับ (0,5) (0,10) (0,20) (0,50) แสดงผลในตารางที่ 4.3.2

4.3.3 การแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ

คือ การแจกแจงที่มีความเบ้และความโด่ง เป็นคู่อันดับ (-1,2) (-1,3) (1,2) (1,3) แสดงผลในตารางที่ 4.3.3

4.3.4 การแจกแจงลักษณะไม่สมมาตรและโด่งสูง

คือ การแจกแจงที่มีความเบ้และความโด่ง เป็นคู่อันดับ (-2,10) (-2,50) (1,10) (1,50) แสดงผลในตารางที่ 4.3.4

4.3.1 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ

ตารางที่ 4. 5 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ (0,-1) (0,1) (0,2) (0,3)

สถิติทดสอบ	ขนาด ตัวอย่าง (n)	อำนาจการทดสอบ ที่ระดับนัยสำคัญ 0.05			
		การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
		$\gamma_1=0, \gamma_2=-1$	$\gamma_1=0, \gamma_2=1$	$\gamma_1=0, \gamma_2=2$	$\gamma_1=0, \gamma_2=3$
Shapiro-Wilk (W)	10	0.060	0.072	0.080	0.098
	20	0.080	0.076	0.158	0.208
	30	0.152	0.114	0.200	0.262
	50	0.316	0.146	0.284	0.366
	100	0.868	0.264	0.490	0.630

*หมายเหตุ : ค่าอำนาจการทดสอบของสถิติ Ryan-Joiner และ Sürücü "C" เท่ากับ 0 ทุกกรณี

4.3.2 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะสมมาตรและโด่งสูง

ตารางที่ 4. 6 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ(0,5) (0,10) (0,20) (0,50)

สถิติทดสอบ	ขนาด ตัวอย่าง (n)	อำนาจการทดสอบ ที่ระดับนัยสำคัญ 0.05			
		การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
		$\gamma_1=0, \gamma_2=5$	$\gamma_1=0, \gamma_2=10$	$\gamma_1=0, \gamma_2=20$	$\gamma_1=0, \gamma_2=50$
Shapiro-Wilk (W)	10	0.150	0.258	0.412	0.842
	20	0.278	0.434	0.666	0.992
	30	0.376	0.564	0.818	1.000
	50	0.542	0.758	0.968	1.000
	100	0.812	0.980	1.000	1.000

*หมายเหตุ : ค่าอำนาจการทดสอบของสถิติ Ryan-Joiner และ Sürücü "C" เท่ากับ 0 ทุกกรณี

4.3.3 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ

ตารางที่ 4. 7 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงด้วยความเบ้และความโด่งเป็นคู่อันดับ (-1,2) (-1,3) (1,2) (1,3)

สถิติทดสอบ	ขนาด ตัวอย่าง (n)	อำนาจการทดสอบ ที่ระดับนัยสำคัญ 0.05 การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
		$\gamma_1=-1, \gamma_2=2$	$\gamma_1=-1, \gamma_2=3$	$\gamma_1=1, \gamma_2=2$	$\gamma_1=1, \gamma_2=3$
		Shapiro-Wilk (W)	10	0.158	0.160
	20	0.254	0.242	0.262	0.262
	30	0.352	0.322	0.392	0.370
	50	0.572	0.534	0.580	0.532
	100	0.862	0.804	0.864	0.816

*หมายเหตุ : ค่าอำนาจการทดสอบของสถิติ Ryan-Joiner และ Sürücü "C" เท่ากับ 0 ทุกกรณี

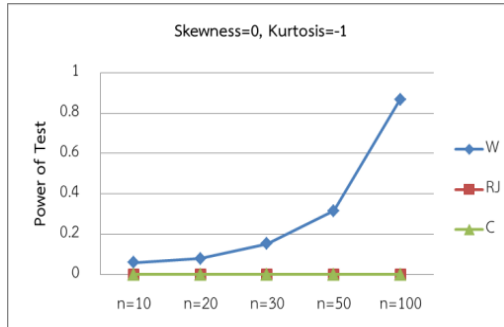
4.3.4 การเปรียบเทียบอำนาจการทดสอบกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง

ตารางที่ 4. 8 แสดงค่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk ที่ระดับนัยสำคัญ 0.05 โดยการสุ่มข้อมูลมาจากการแจกแจงความเบ้และความโด่งเป็นคู่อันดับ (-2,10) (-2,50) (1,10) (1,50)

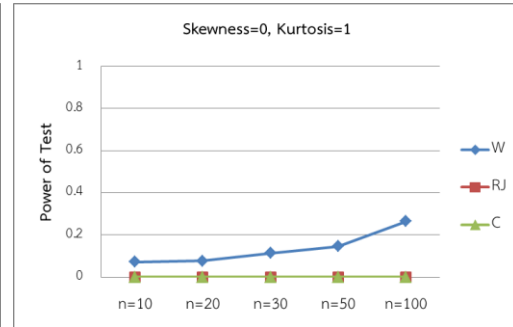
สถิติทดสอบ	ขนาด ตัวอย่าง (n)	อำนาจการทดสอบ ที่ระดับนัยสำคัญ 0.05 การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
		$\gamma_1=-2, \gamma_2=10$	$\gamma_1=-2, \gamma_2=50$	$\gamma_1=1, \gamma_2=10$	$\gamma_1=1, \gamma_2=50$
		Shapiro-Wilk (W)	10	0.316	0.818
	20	0.512	0.998	0.440	0.992
	30	0.692	1.000	0.584	1.000
	50	0.870	1.000	0.792	1.000
	100	0.982	1.000	0.974	1.000

*หมายเหตุ : ค่าอำนาจการทดสอบของสถิติ Ryan-Joiner และ Sürücü "C" เท่ากับ 0 ทุกกรณี

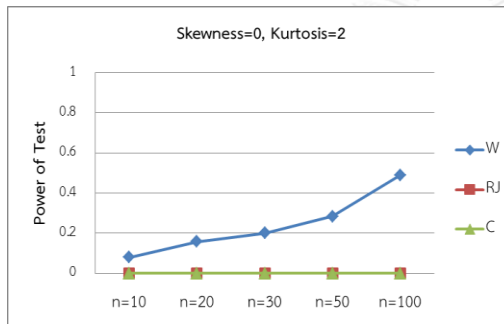
กราฟแสดงการเปรียบเทียบอำนาจการทดสอบของการแจกแจงลักษณะสมมาตรและโด่งต่ำ เพื่อดูแนวโน้มของขนาดตัวอย่าง และลักษณะการแจกแจงของความเบ้และความโด่ง ดังนี้



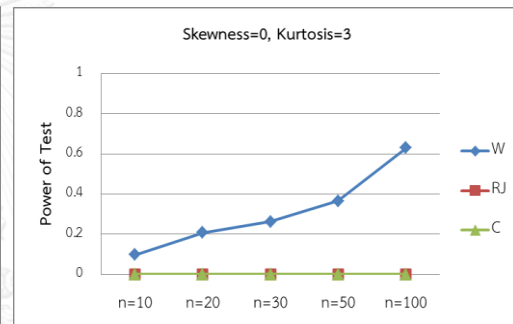
(a) อำนาจการทดสอบคู่อันดับ (0,-1)



(b) อำนาจการทดสอบคู่อันดับ (0,1)



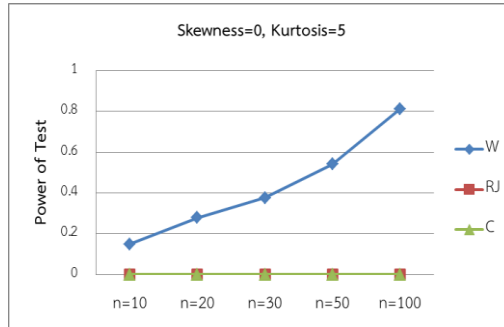
(c) อำนาจการทดสอบคู่อันดับ (0,2)



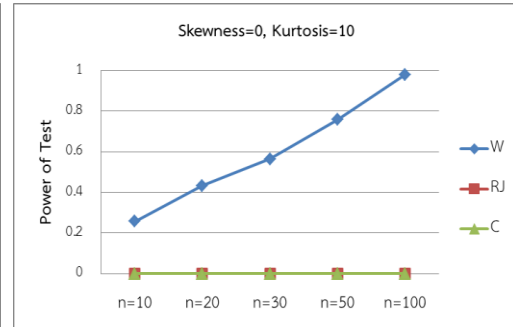
(d) อำนาจการทดสอบคู่อันดับ (0,3)

ภาพที่ 4.17 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อกำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงสมมาตรและโด่งต่ำ

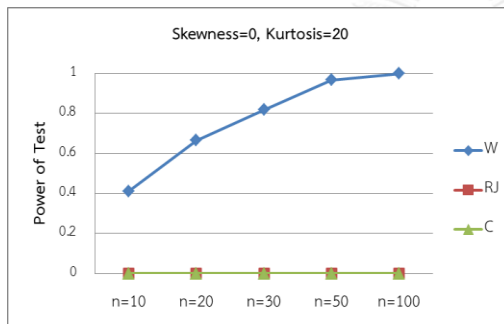
กราฟแสดงการเปรียบเทียบอำนาจการทดสอบของการแจกแจงลักษณะสมมาตรและโด่งสูง เพื่อดูแนวโน้มของขนาดตัวอย่าง และลักษณะการแจกแจงของความเบ้และความโด่ง ดังนี้



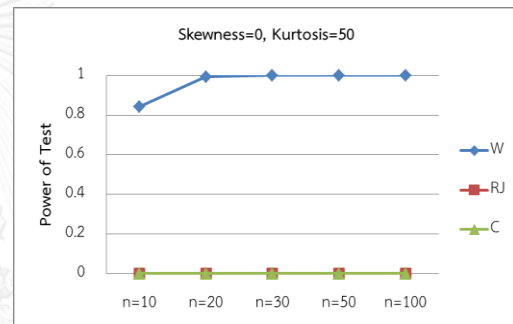
(a) อำนาจการทดสอบคู่อันดับ (0,5)



(b) อำนาจการทดสอบคู่อันดับ (0,10)



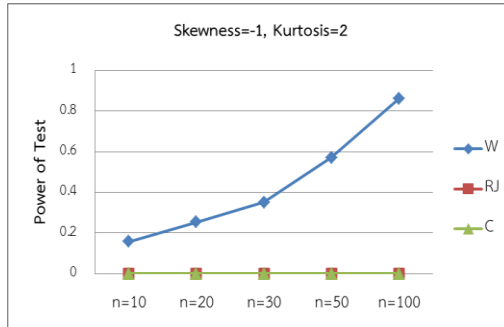
(c) อำนาจการทดสอบคู่อันดับ (0,20)



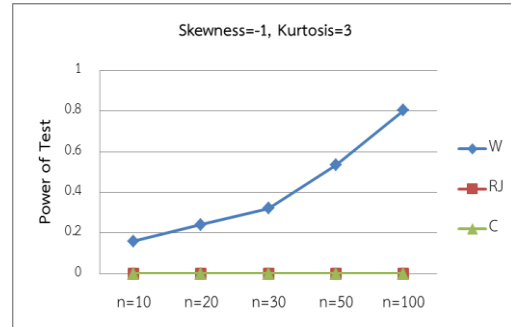
(d) อำนาจการทดสอบคู่อันดับ (0,50)

ภาพที่ 4. 18 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อกำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงสมมาตรและโด่งสูง

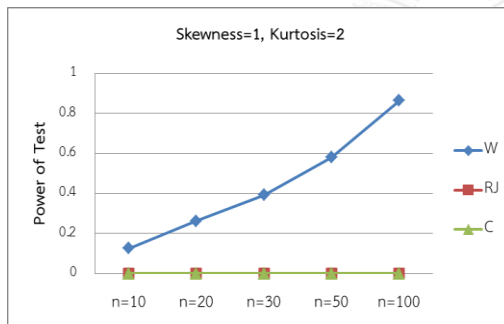
กราฟแสดงการเปรียบเทียบอำนาจการทดสอบของการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ เพื่อดูแนวโน้มของขนาดตัวอย่าง และลักษณะการแจกแจงของความเบ้และความโด่ง ดังนี้



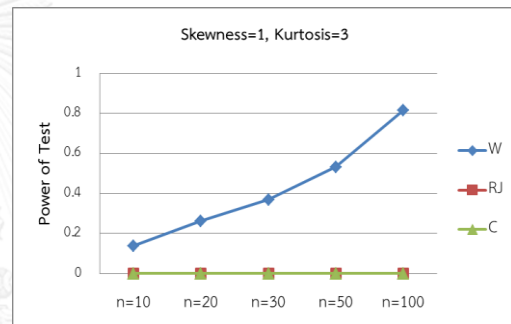
(a) อำนาจการทดสอบคู่อันดับ (-1,2)



(b) อำนาจการทดสอบคู่อันดับ (-1,3)



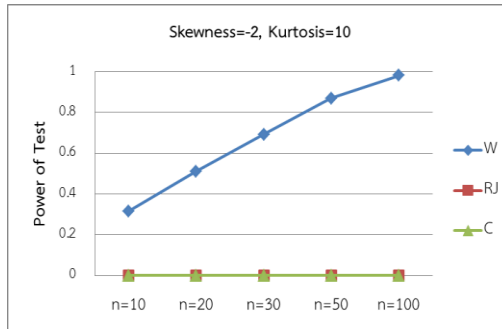
(c) อำนาจการทดสอบคู่อันดับ (1,2)



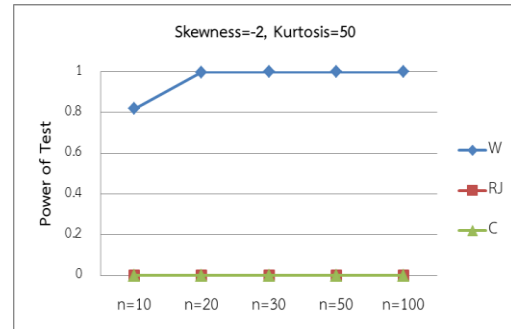
(d) อำนาจการทดสอบคู่อันดับ (1,3)

ภาพที่ 4.19 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อกำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงไม่สมมาตรและโด่งต่ำ

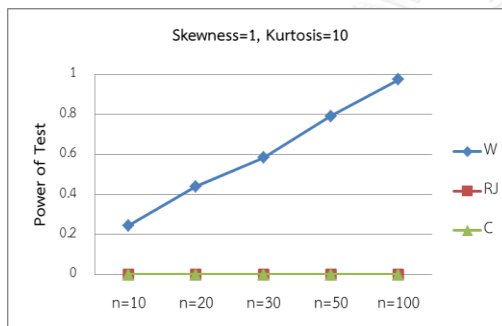
กราฟแสดงการเปรียบเทียบอำนาจการทดสอบของการแจกแจงลักษณะไม่สมมาตรและโด่งสูง เพื่อดูแนวโน้มของขนาดตัวอย่าง และลักษณะการแจกแจงของความเบ้และความโด่ง ดังนี้



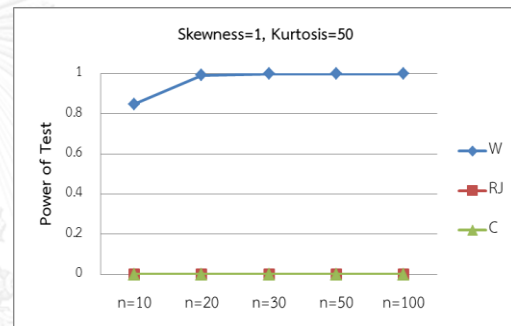
(a) อำนาจการทดสอบคู่อันดับ (-2,10)



(b) อำนาจการทดสอบคู่อันดับ (-2,50)



(c) อำนาจการทดสอบคู่อันดับ (1,10)



(d) อำนาจการทดสอบคู่อันดับ (1,50)

ภาพที่ 4. 20 แสดงการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบ W, RJ และ C เมื่อกำหนดระดับนัยสำคัญเป็น 0.05 ภายใต้ลักษณะการแจกแจงไม่สมมาตรและโด่งสูง

จากตารางที่ 4.5-4.8 สามารถวิเคราะห์ผลได้ดังนี้

1. พิจารณาตามขนาดตัวอย่าง

จากตาราง 4.5 กรณีการแจกแจงมีลักษณะสมมาตรและโด่งต่ำ พบว่าเมื่อขนาดตัวอย่างเท่ากับ 100 สำหรับทุกการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (0,-1) (0,1) (0,2) (0,3) โดยสถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงสุด โดยที่อำนาจการทดสอบของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เท่ากับ 0 ทุกกรณี และยังพบว่าอำนาจการทดสอบของสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น (แสดงดังภาพที่ 4.17)

จากตาราง 4.6 กรณีการแจกแจงมีลักษณะสมมาตรและโด่งสูง พบว่าเมื่อขนาดตัวอย่างเท่ากับ 100 สำหรับทุกการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (0,5) (0,10) (0,20) (0,50) สถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงสุด โดยที่อำนาจการทดสอบของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เท่ากับ 0 ทุกกรณี และยังพบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น (แสดงดังภาพที่ 4.18)

จากตาราง 4.7 กรณีการแจกแจงมีลักษณะไม่สมมาตรและโด่งต่ำ พบว่าเมื่อขนาดตัวอย่างเท่ากับ 100 สำหรับทุกการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (-1,2) (-1,3) (1,2) (1,3) สถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงสุด โดยที่อำนาจการทดสอบของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เท่ากับ 0 ทุกกรณี และยังพบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น (แสดงดังภาพที่ 4.19)

จากตาราง 4.8 กรณีการแจกแจงมีลักษณะไม่สมมาตรและโด่งสูง พบว่าเมื่อขนาดตัวอย่างเท่ากับ 100 สำหรับทุกการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (-2,10) (-2,50) (1,10) (1,50) สถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงสุด โดยที่อำนาจการทดสอบของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เท่ากับ 0 ทุกกรณี และยังพบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น (แสดงดังภาพที่ 4.20)

ดังนั้น ผลสรุปจาก 4 กรณีข้างต้น ได้ว่าสถิติทดสอบ Shapiro-Wilk ให้ค่าอำนาจทดสอบสูงเมื่อเปรียบเทียบกับวิธีการทดสอบ Ryan-Joiner และ Sürücü “C” ในทุกกรณีศึกษา หรือกล่าวได้ว่ามีประสิทธิภาพมากที่สุดในทุก ๆ สถานการณ์ที่จำลอง และเหมาะสมสำหรับตัวอย่างขนาดใหญ่

2. พิจารณาตามลักษณะการแจกแจงของความเบ้และความโด่ง

จากตาราง 4.5 กรณีการแจกแจงมีลักษณะสมมาตรและโด่งต่ำ พบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดความโด่งเพิ่มขึ้นสำหรับทุกขนาดตัวอย่างต่อการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (0,-1) (0,1) (0,2) (0,3) ทุกแบบ เช่น เมื่อขนาดตัวอย่างเท่ากับ 10 สำหรับความเบ้เท่ากับ 0 ทุกกรณีแล้วมีความโด่งที่เพิ่มขึ้นจาก -1, 1, 2 และ 3 สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเพิ่มขึ้น คือ 0.060, 0.072, 0.080 และ 0.098 ตามลำดับ กรณีขนาดตัวอย่างอื่น ๆ ของความโด่งที่เพิ่มขึ้นให้ผลเช่นเดียวกัน (แสดงดังภาพที่ 4.17)

จากตาราง 4.6 กรณีการแจกแจงมีลักษณะสมมาตรและโด่งสูง พบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดความโด่งเพิ่มขึ้นสำหรับทุกขนาดตัวอย่างต่อการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (0,5) (0,10) (0,20) (0,50) ทุกแบบ เช่น เมื่อขนาดตัวอย่างเท่ากับ 10 สำหรับความเบ้เท่ากับ 0 ทุกกรณีแล้วมีความโด่งที่เพิ่มขึ้นจาก 5, 10, 20 และ 50 ซึ่งเป็นความโด่งที่สูงมาก สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเพิ่มขึ้นมากเช่นกัน คือ 0.150, 0.258, 0.412 และ 0.842 ตามลำดับ ยังพบอีกว่าเมื่อขนาดตัวอย่างที่ความโด่งสูงมากเป็น 50 สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเท่ากับ 1 ซึ่งสูงสุด (แสดงดังภาพที่ 4.18)

จากตาราง 4.7 กรณีการแจกแจงมีลักษณะไม่สมมาตรและโด่งต่ำ พบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดความโด่งเพิ่มขึ้นสำหรับทุกขนาดตัวอย่างต่อการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (-1,2) (-1,3) (1,2) (1,3) ทุกแบบ เช่น เมื่อขนาดตัวอย่างเท่ากับ 10 สำหรับคู่อันดับ (-1,2) (-1,3) ที่ความเบ้เท่ากันคือ -1 แล้วความโด่งที่เพิ่มขึ้นจาก 2 เป็น 3 สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเพิ่มขึ้น คือ 0.158 และ 0.160 ตามลำดับ กรณีขนาดตัวอย่างอื่น ๆ ของความโด่งที่เพิ่มขึ้นให้ผลเช่นเดียวกัน (แสดงดังภาพที่ 4.19)

จากตาราง 4.8 กรณีการแจกแจงมีลักษณะไม่สมมาตรและโด่งสูง พบว่าอำนาจการทดสอบสถิติทดสอบ Shapiro-Wilk มีค่ามากขึ้นเมื่อขนาดความโด่งเพิ่มขึ้นสำหรับทุกขนาดตัวอย่างต่อการแจกแจงแต่ละความเบ้และความโด่งคู่อันดับ (-2,10) (-2,50) (1,10) (1,50) ทุกแบบ เช่น เมื่อขนาดตัวอย่างเท่ากับ 10 สำหรับคู่อันดับ (1,10) (1,50) ที่ความเบ้เท่ากันคือ 1 แล้วความโด่งที่เพิ่มขึ้นจาก 10 เป็น 50 สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเพิ่มขึ้น คือ 0.246 และ 0.848 ตามลำดับ ยังพบอีกว่าเมื่อขนาดตัวอย่างที่ความโด่งสูงมากเป็น 50 สถิติทดสอบ Shapiro-Wilk ให้อำนาจการทดสอบเท่ากับ 1 ซึ่งสูงสุด (แสดงดังภาพที่ 4.20)

ดังนั้น ผลสรุปจาก 4 กรณีข้างต้นเพิ่มเติม ได้ว่าในทุกกรณีศึกษาเหมาะสมสำหรับลักษณะการแจกแจงยิ่งโด่งสูงทั้งกรณีสมมาตรและไม่สมมาตร

4.4 การวิเคราะห์อำนาจการทดสอบ

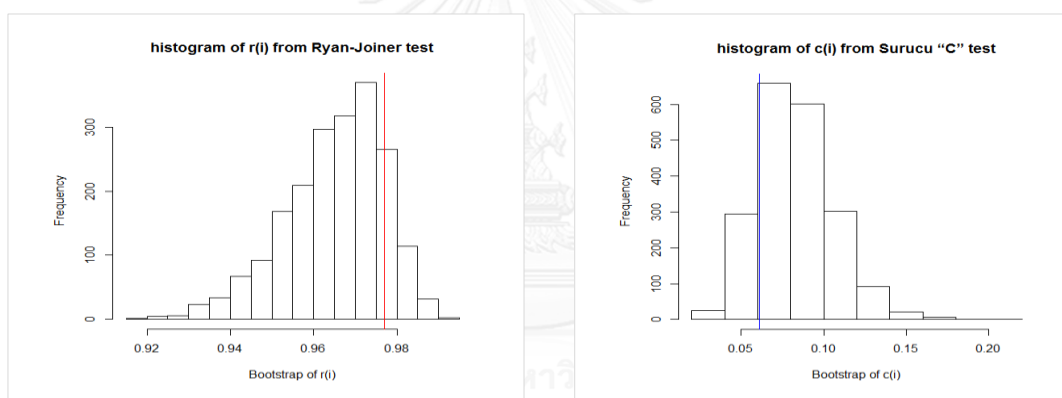
จากข้อมูลจำลองด้วยวิธีเพาเวอร์สำหรับเพื่อเปรียบเทียบอำนาจการทดสอบ ผลการทดสอบของสถิติ Ryan-Joiner และ Sürücü “C” มีค่าอำนาจการทดสอบเท่ากับ 0 ในทุกกรณี สามารถวิเคราะห์ผลออกเป็น 2 ส่วน ดังนี้

4.4.1 วิเคราะห์ผลจากวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C”

4.4.2 วิเคราะห์ผลจากค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C”

4.4.1 วิเคราะห์ผลจากวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C”

- กรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ



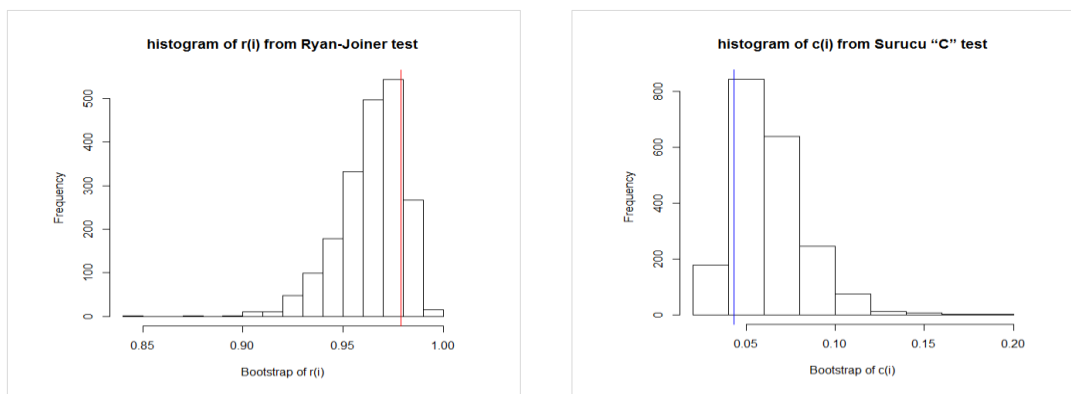
(a) Histogram of $r(i)$ from Ryan-Joiner

(b) Histogram of $c(i)$ from Sürücü “C”

ภาพที่ 4.21 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,-1) และ $n=30$

จากภาพที่ 4.21 (a) วิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner พบว่า **ค่าสถิติ $R_{cal}=0.9770$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ R_{cal} มีค่าน้อยๆ จะได้ว่า $p\text{-value} = P[R_{cal} < r(i)]$ โดยร้อยละพื้นที่ด้านซ้ายของ Histogram มาก ดังนั้นค่าประมาณ p-value จึงสูงมาก (แสดงดังภาพที่ 4.25 (b)) และจากภาพที่ 4.21 (b) วิธีบูตสเตรปของสถิติทดสอบ Sürücü “C” พบว่า **ค่าสถิติ $C_{cal}=0.0611$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ C_{cal} มีค่ามากๆ จะได้ว่า $p\text{-value} = P[C_{cal} > c(i)]$ โดยร้อยละพื้นที่ด้านขวา Histogram มาก ดังนั้นค่าประมาณ p-value จึงสูงมาก (แสดงดังภาพที่ 4.25 (c)) ซึ่งเป็นสาเหตุที่ค่าอำนาจการทดสอบเท่ากับ 0

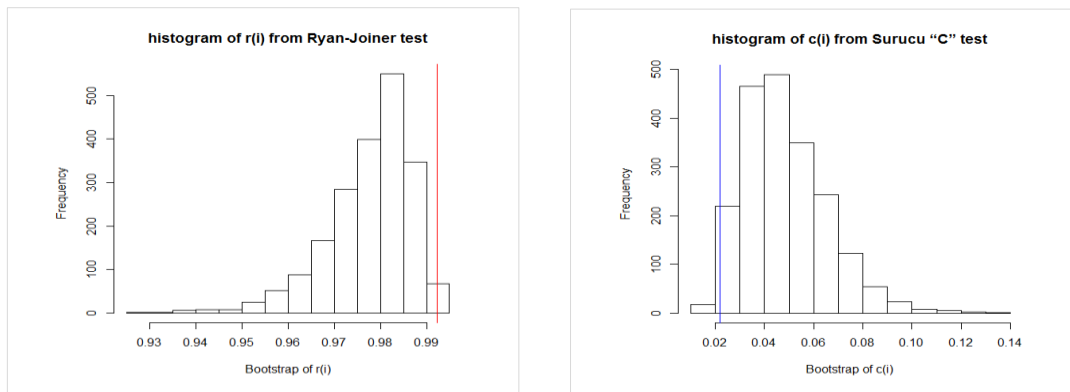
- กรณีการแจกแจงลักษณะสมมาตรและโด่งสูง

(a) Histogram of $r(i)$ from Ryan-Joiner(b) Histogram of $c(i)$ from Sürücü "C"

ภาพที่ 4.22 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü "C" ภายใต้ความเบ้และความโด่งคู่อันดับ (0,5) และ $n=30$

จากภาพที่ 4.22 (a) วิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner พบว่า **ค่าสถิติ $R_{cal}=0.9787$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ R_{cal} มีค่าน้อยๆ จะได้ว่า $p\text{-value} = P[R_{cal} < r(i)]$ โดยร้อยละพื้นที่ด้านซ้ายของ Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.26 (b)) และจากภาพที่ 4.22 (b) วิธีบูตสเตรปของสถิติทดสอบ Sürücü "C" พบว่า **ค่าสถิติ $C_{cal}=0.0431$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ C_{cal} มีค่ามากๆ จะได้ว่า $p\text{-value} = P[C_{cal} > c(i)]$ โดยร้อยละพื้นที่ด้านขวา Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.26 (c)) ซึ่งเป็นสาเหตุที่ค่าอำนาจการทดสอบเท่ากับ 0

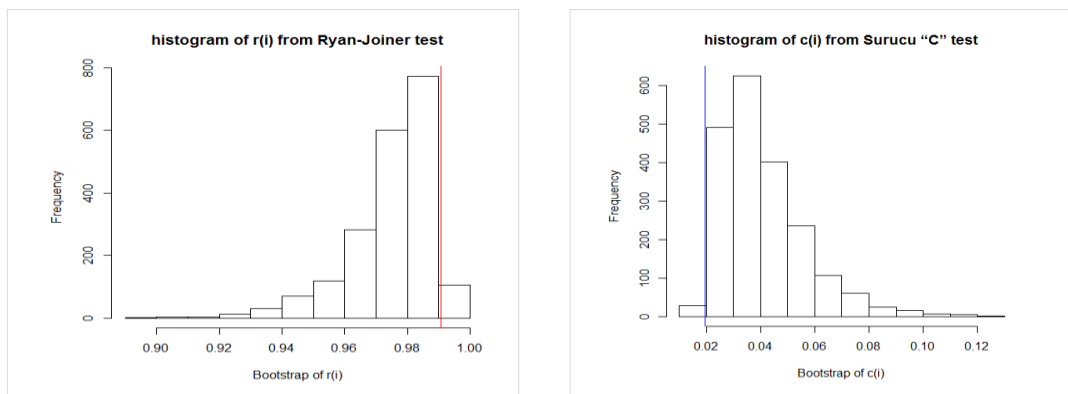
- กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ

(a) Histogram of $r(i)$ from Ryan-Joiner(b) Histogram of $c(i)$ from Sürücü "C"

ภาพที่ 4.23 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü "C" ภายใต้ความเบ้และความโด่งคู่อันดับ $(-1,2)$ และ $n=30$

จากภาพที่ 4.23 (a) วิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner พบว่า **ค่าสถิติ $R_{cal}=0.9923$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ R_{cal} มีค่าน้อยๆ จะได้ว่า $p\text{-value} = P[R_{cal} < r(i)]$ โดยร้อยละพื้นที่ด้านซ้ายของ Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.27 (b)) และจากภาพที่ 4.23 (b) วิธีบูตสเตรปของสถิติทดสอบ Sürücü "C" พบว่า **ค่าสถิติ $C_{cal}=0.0219$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ C_{cal} มีค่ามากๆ จะได้ว่า $p\text{-value} = P[C_{cal} > c(i)]$ โดยร้อยละพื้นที่ด้านขวา Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.27 (c)) ซึ่งเป็นสาเหตุที่ค่าอำนาจการทดสอบเท่ากับ 0

- กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง

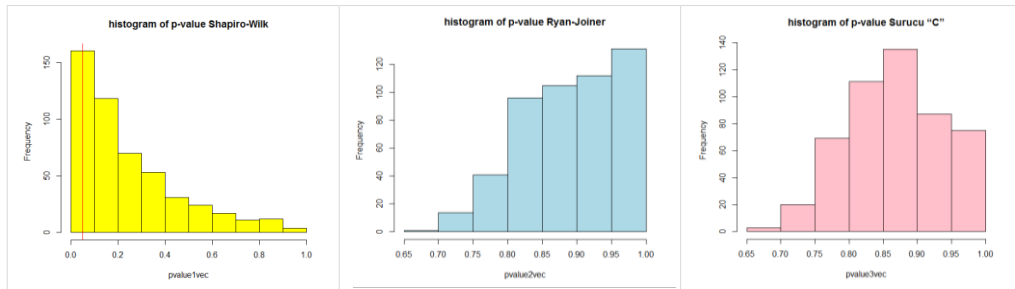
(a) Histogram of $r(i)$ from Ryan-Joiner(b) Histogram of $c(i)$ from Sürücü "C"

ภาพที่ 4. 24 แสดงฮิสโทแกรมของการจำลองด้วยวิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner และ Sürücü "C" ภายใต้ความเบ้และความโด่งคู่อันดับ $(-2,10)$ และ $n=30$

จากภาพที่ 4.24 (a) วิธีบูตสเตรปของสถิติทดสอบ Ryan-Joiner พบว่า **ค่าสถิติ $R_{cal}=0.9905$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ R_{cal} มีค่าน้อยๆ จะได้ว่า $p\text{-value} = P[R_{cal} < r(i)]$ โดยร้อยละพื้นที่ด้านซ้ายของ Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.28 (b)) และจากภาพที่ 4.24 (b) วิธีบูตสเตรปของสถิติทดสอบ Sürücü "C" พบว่า **ค่าสถิติ $C_{cal}=0.0194$** เกณฑ์ปฏิเสธสมมติฐานว่างเมื่อ C_{cal} มีค่ามากๆ จะได้ว่า $p\text{-value} = P[C_{cal} > c(i)]$ โดยร้อยละพื้นที่ด้านขวา Histogram มาก ดังนั้นค่าประมาณ $p\text{-value}$ จึงสูงมาก (แสดงดังภาพที่ 4.28 (c)) ซึ่งเป็นสาเหตุที่ทำให้ค่าอำนาจการทดสอบเท่ากับ 0

4.4.2 วิเคราะห์ผลจากค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C”

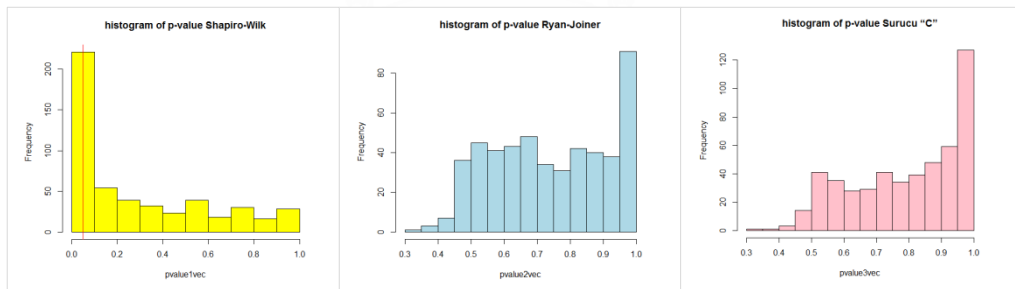
- กรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ



(a) Histogram of p-value Shapiro-Wilk (b) Histogram of p-value Ryan-Joiner (c) Histogram of p-value Sürücü “C”

ภาพที่ 4. 25 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,-1) และ $n=30$

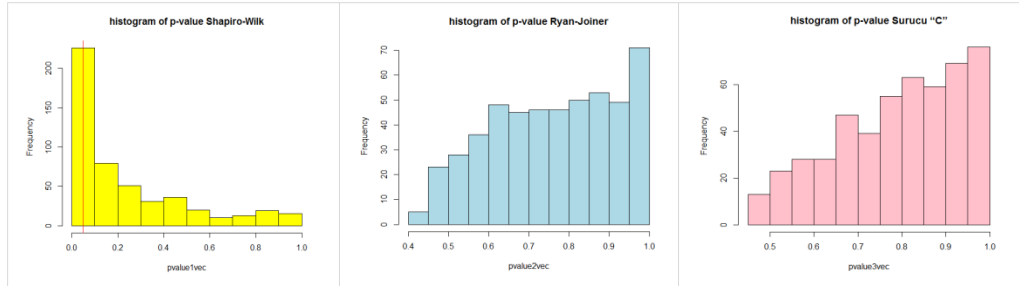
- กรณีการแจกแจงลักษณะสมมาตรและโด่งสูง



(a) Histogram of p-value Shapiro-Wilk (b) Histogram of p-value Ryan-Joiner (c) Histogram of p-value Sürücü “C”

ภาพที่ 4. 26 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” ภายใต้ความเบ้และความโด่งคู่อันดับ (0,5) และ $n=30$

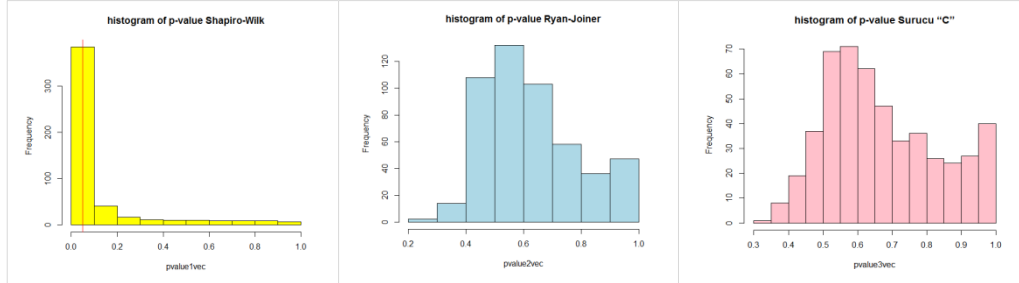
- กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ



(a) Histogram of p-value Shapiro-Wilk (b) Histogram of p-value Ryan-Joiner (c) Histogram of p-value Sürücü "C"

ภาพที่ 4. 27 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü "C" ภายใต้ความเบ้และความโด่งคู่อันดับ (-1,2) และ $n=30$

- กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง



(a) Histogram of p-value Shapiro-Wilk (b) Histogram of p-value Ryan-Joiner (c) Histogram of p-value Sürücü "C"

ภาพที่ 4. 28 แสดงฮิสโทแกรมค่า p-value ของสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü "C" ภายใต้ความเบ้และความโด่งคู่อันดับ (-2,10) และ $n=30$

จากภาพที่ 4.25 (a) Histogram สถิติทดสอบ Shapiro-Wilk พบว่าจำนวน p-value ทั้งหมด 500 ค่า มีค่า p-value < 0.05 ระดับนัยสำคัญที่กำหนด ซึ่งมีการปฏิเสธสมมติฐานว่างเกิดขึ้น ทำให้สามารถหาค่าอำนาจการทดสอบได้ แต่ในส่วนจากภาพที่ 4.25 (b) Histogram สถิติทดสอบ Ryan-Joiner และจากภาพที่ 4.25 (c) Histogram สถิติทดสอบ Sürücü “C” พบว่าค่า p-value ของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” สำหรับภายใต้ความเบ้และความโด่งคู่อันดับ (0,-1) และ $n=30$ ซึ่งอยู่ในกรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ จำนวน p-value ทั้ง 500 ค่าที่ประมาณด้วยวิธีбутแสดงมีค่าสูงมาก จึงไม่มีการปฏิเสธสมมติฐานว่างเกิดขึ้น รวมทั้งกรณีการแจกแจงลักษณะสมมาตรและโด่งสูง กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ และกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูงให้ผลในลักษณะเดียวกัน ผลดังกล่าวแสดงให้เห็นว่าทุกค่าของ p-value สำหรับสถิติทดสอบ Ryan-Joiner และ Sürücü “C” มากกว่าระดับนัยสำคัญ 0.05 จึงไม่มีโอกาสที่จะเกิดอำนาจการทดสอบเลย

ดังนั้นสรุปได้ว่า ในการทดสอบสมมติฐาน $H_0: \{Y_1, Y_2, \dots, Y_n\}$ เป็นตัวอย่างสุ่มที่มาจากประชากรแบบปกติ จากการประมาณค่า p-value ด้วยวิธีбутแสดง ผลแสดงให้เห็นในทุกกรณีว่าค่า p-value มากกว่าระดับนัยสำคัญ 0.05 ที่กำหนด ดังนั้นจึงทำให้อำนาจการทดสอบของสถิติ Ryan-Joiner และ Sürücü “C” เท่ากับ 0 ในทุกกรณีนั่นเอง

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในงานวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของสถิติทดสอบการแจกแจงปกติ วิธีใดมีประสิทธิภาพมากที่สุด การเปรียบเทียบประสิทธิภาพของสถิติทดสอบการแจกแจงปกติครั้งนี้ พิจารณาอำนาจการทดสอบของสถิติทดสอบ 3 วิธี ได้แก่ สถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” โดยลักษณะข้อมูลที่จำลองขึ้นศึกษาจากวิธีเพาเวอร์ของ Fleishman (1978) ซึ่งเป็นวิธีที่ง่ายสำหรับการกำหนดความเบ้และความโด่งในรูปแบบต่าง ๆ ได้แบ่งกรณีศึกษา 4 กรณี ประกอบด้วย การแจกแจงสมมาตรและโด่งต่ำ การแจกแจงสมมาตรและโด่งสูง การแจกแจงไม่สมมาตรและโด่งต่ำ และการแจกแจงไม่สมมาตรและโด่งสูง พร้อมทั้งศึกษาถึงขนาดกลุ่มตัวอย่างว่า ขนาดกลุ่มตัวอย่างใดเหมาะสมกับวิธีทดสอบใดบ้าง โดยกำหนดตัวอย่างขนาดเล็กคือจำนวนขนาดตัวอย่าง 10, 20 และ 30 และกำหนดตัวอย่างขนาดใหญ่คือจำนวนขนาดตัวอย่าง 50 และ 100 โดยแต่ละสถานการณ์จำลองข้อมูลมีการทำซ้ำค่า p-value 500 ครั้งของสถิติทดสอบแต่ละวิธี สำหรับสถิติทดสอบ Ryan-Joiner และ Sürücü “C” มีการแจกแจงซับซ้อน จึงประมาณค่า p-value ด้วยวิธี Bootstrap ผลการวิจัยนี้สามารถเป็นแนวทางในการเลือกสถิติทดสอบที่เหมาะสมสำหรับการทดสอบการแจกแจงแบบปกติ จะได้ว่าผลการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบทั้ง 3 วิธีข้างต้น มีข้อสรุปดังนี้

5.1 สรุปผลการวิจัย

จากการศึกษาเพื่อเปรียบเทียบอำนาจการทดสอบ (Power of Test) ของสถิติทดสอบการแจกแจงแบบปกติระหว่างสถิติทดสอบของ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” นั้น แบ่งเป็น 2 ส่วน คือ

5.1.1 การเปรียบเทียบอำนาจการทดสอบ

5.1.2 ปัจจัยที่ส่งผลต่อประสิทธิภาพของอำนาจการทดสอบ

5.1.1 การเปรียบเทียบอำนาจการทดสอบ

ในการเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบการแจกแจงปกติ 3 การทดสอบพบว่า สถิติทดสอบ Shapiro-Wilk มีอำนาจการทดสอบสูงเมื่อเทียบกับสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เนื่องจากสถิติทดสอบ Ryan-Joiner และ Sürücü “C” มีอำนาจการทดสอบเป็น 0 ทุกกรณี จึงทำการเปรียบเทียบเพียงอำนาจการทดสอบสถิติ Shapiro-Wilk ทั้ง 4 กรณีการแจกแจง โดยพิจารณาตามขนาดตัวอย่าง และลักษณะของความเบ้และความโด่ง สามารถสรุปผลได้ดังนี้

ตารางที่ 5. 1 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำ โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]

ขนาดตัวอย่าง (n)	ลำดับอำนาจการทดสอบสถิติ Shapiro-Wilk			
	การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
	$\gamma_1=0, \gamma_2=-1$	$\gamma_1=0, \gamma_2=1$	$\gamma_1=0, \gamma_2=2$	$\gamma_1=0, \gamma_2=3$
10	(5) [4]	(5) [3]	(5) [2]	(5) [1]
20	(4) [3]	(4) [4]	(4) [2]	(4) [1]
30	(3) [3]	(3) [4]	(3) [2]	(3) [1]
50	(2) [2]	(2) [4]	(2) [3]	(2) [1]
100	(1) [1]	(1) [4]	(1) [3]	(1) [2]

ตารางที่ 5. 2 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะสมมาตรและโด่งสูง โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]

ขนาดตัวอย่าง (n)	ลำดับอำนาจการทดสอบสถิติ Shapiro-Wilk			
	การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
	$\gamma_1=0, \gamma_2=5$	$\gamma_1=0, \gamma_2=10$	$\gamma_1=0, \gamma_2=20$	$\gamma_1=0, \gamma_2=50$
10	(5) [4]	(5) [3]	(5) [2]	(3) [1]
20	(4) [4]	(4) [3]	(4) [2]	(2) [1]
30	(3) [4]	(3) [3]	(3) [2]	(1) [1]
50	(2) [4]	(2) [3]	(2) [2]	(1) [1]
100	(1) [3]	(1) [2]	(1) [1]	(1) [1]

ตารางที่ 5. 3 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำ โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]

ขนาดตัวอย่าง (n)	ลำดับอำนาจการทดสอบสถิติ Shapiro-Wilk การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
	$\gamma_1=-1, \gamma_2=2$	$\gamma_1=-1, \gamma_2=3$	$\gamma_1=1, \gamma_2=2$	$\gamma_1=1, \gamma_2=3$
10	(5) [2]	(5) [1]	(5) [4]	(5) [3]
20	(4) [2]	(4) [3]	(4) [1]	(4) [1]
30	(3) [3]	(3) [4]	(3) [1]	(3) [2]
50	(2) [2]	(2) [3]	(2) [1]	(2) [4]
100	(1) [2]	(1) [4]	(1) [1]	(1) [3]

ตารางที่ 5. 4 แสดงลำดับอำนาจการทดสอบ กรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูง โดยพิจารณาตามขนาดตัวอย่าง (ลำดับแนวตั้ง) และความเบ้และความโด่ง [ลำดับแนวนอน]

ขนาดตัวอย่าง (n)	ลำดับอำนาจการทดสอบสถิติ Shapiro-Wilk การแจกแจงด้วยความเบ้และความโด่ง (γ_1, γ_2)			
	$\gamma_1=-2, \gamma_2=10$	$\gamma_1=-2, \gamma_2=50$	$\gamma_1=1, \gamma_2=20$	$\gamma_1=1, \gamma_2=50$
10	(5) [3]	(3) [2]	(5) [4]	(3) [1]
20	(4) [3]	(2) [1]	(4) [4]	(2) [2]
30	(3) [2]	(1) [1]	(3) [3]	(1) [1]
50	(2) [2]	(1) [1]	(2) [3]	(1) [1]
100	(1) [2]	(1) [1]	(1) [3]	(1) [1]

หมายเหตุ : เลข 1 หมายถึง ค่าอำนาจการทดสอบสูงสุด ตามด้วย 2, 3, 4 และ 5 ตามลำดับ

จากตารางที่ 5.1-5.4 สามารถสรุปผลได้ว่า จากลำดับค่าของอำนาจการทดสอบเมื่อพิจารณาภายใต้ลักษณะการแจกแจงเดียวกันด้วยตัวอย่างที่แตกต่างกัน พบว่าขนาดตัวอย่าง $n=100$ มีลำดับของอำนาจการทดสอบสูงสุดสำหรับทุกกรณีการแจกแจง นั่นคือสถิติ Shapiro-Wilk เป็นวิธีที่เหมาะสมกับตัวอย่างขนาดใหญ่ และเมื่อพิจารณาภายใต้ตัวอย่างเดียวกันด้วยลักษณะการแจกแจงแตกต่างกัน พบว่า ความโด่งสูง $\gamma_2=50$ มีลำดับของอำนาจการทดสอบสูงสุดสำหรับทุกขนาดตัวอย่าง นั่นคือสถิติ Shapiro-Wilk เป็นวิธีที่เหมาะสมกับข้อมูลความโด่งสูง

5.1.2 ปัจจัยที่ส่งผลต่อประสิทธิภาพของอำนาจการทดสอบ

จากผลการวิจัยอำนาจการทดสอบ (Power of Test) ของสถิติทดสอบ 3 วิธี คือสถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü “C” พบว่าขึ้นอยู่กับปัจจัยต่าง ๆ ดังนี้

1. การแจกแจงของตัวสถิติทดสอบ (Sampling Distribution)

พบว่า ตัวสถิติทดสอบ Ryan-Joiner และ Sürücü “C” มีการแจกแจงของตัวสถิติที่ซับซ้อน ต้องจำลองวิธี Simulation โดยใช้วิธีการ Bootstrap คำนวณค่า p-value โดยอำนาจการทดสอบทั้ง 2 วิธีให้ผลเป็นศูนย์สำหรับทุกกรณี จึงสรุปได้ว่าสถิติทดสอบ Shapiro-Wilk เป็นวิธีที่มีอำนาจการทดสอบสูงเมื่อเทียบกับวิธี Ryan-Joiner และ Sürücü “C” และการหาค่า p-value ของสถิติทดสอบ Shapiro-Wilk ยังสามารถทำได้ง่ายกว่า เนื่องจากตัวสถิติ Shapiro-Wilk นั้นมีการแจกแจงอย่างชัดเจน

2. ขนาดตัวอย่าง (Sample Size)

พบว่า เมื่อขนาดตัวอย่างเพิ่มขึ้นหรือขนาดตัวอย่างใหญ่ขึ้น ค่าอำนาจการทดสอบจะเพิ่มขึ้น ดังนั้น สถิติทดสอบ Shapiro-Wilk เหมาะสมกับตัวอย่างขนาดใหญ่

3. ลักษณะรูปร่างของการแจกแจง (Skewness and Kurtosis)

พบว่า เมื่อลักษณะรูปร่างการแจกแจงมีความใกล้เคียงกับการแจกแจงปกติ (ความเบ้เท่ากับ 0 และ ความโด่งเท่ากับ 3) แล้วจะมีอำนาจการทดสอบต่ำ แต่เมื่อลักษณะรูปร่างของการแจกแจงเบี่ยงเบนไปจากการแจกแจงปกติมากขึ้นแล้วจะมีอำนาจการทดสอบสูงตามด้วย จะได้ว่าจากกรณีการแจกแจงลักษณะสมมาตรและโด่งต่ำกับกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งต่ำนั้นจะให้อำนาจการทดสอบที่ต่ำ ในส่วนกรณีการแจกแจงลักษณะสมมาตรและโด่งสูงกับกรณีการแจกแจงลักษณะไม่สมมาตรและโด่งสูงนั้นจะให้อำนาจการทดสอบที่สูงมาก

5.2 อภิปรายผลการวิจัย

จากการศึกษาปัจจัยการเปลี่ยนแปลงของลักษณะการแจกแจงข้อมูลที่มีต่ออำนาจการทดสอบ พบว่าอำนาจการทดสอบของสถิติทดสอบมีแนวโน้มเพิ่มขึ้น เมื่อขนาดตัวอย่าง ความเบ้ หรือ ความโด่งที่เพิ่มขึ้น และเมื่อพิจารณาการเปรียบเทียบอำนาจการทดสอบเพื่อตรวจดูประสิทธิภาพของตัวสถิติทดสอบ พบว่าสถิติทดสอบ Shapiro-Wilk เป็นวิธีที่มีอำนาจการทดสอบสูงสุดเมื่อเทียบกับวิธี Ryan-Joiner และ Sürücü “C” เนื่องจากตัวสถิติ Shapiro-Wilk นั้นมีการแจกแจงอย่างชัดเจน ดังนั้นในงานวิจัยนี้สถิติทดสอบที่เหมาะสมสำหรับการแจกแจงแบบปกติของข้อมูล คือ สถิติทดสอบ Shapiro-Wilk ที่ขนาดตัวอย่างใหญ่ และลักษณะการแจกแจงมีความโด่งสูง

5.3 ข้อเสนอแนะ

การวิจัยครั้งนี้ มีข้อเสนอแนะแบ่งเป็น 2 ด้าน ได้แก่

5.3.1 ข้อเสนอแนะในการนำไปประยุกต์ใช้งาน

จากงานวิจัยนี้ในการเลือกใช้สถิติทดสอบสำหรับการแจกแจงแบบปกติ โดยสถิติทดสอบ Ryan-Joiner และ Sürücü “C” มีการแจกแจงของตัวสถิติที่ซับซ้อน ทำให้หาค่า p-value ยากในโปรแกรม R เนื่องจากไม่มีฟังก์ชันสำเร็จรูป ซึ่งผู้วิจัยได้ใช้วิธี Bootstrap ในการหาค่า p-value ของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” จึงน่าสนใจในการศึกษากรณีอื่น ๆ ที่สามารถจะนำวิธีбутแสดงรูปไปประยุกต์ใช้กับตัวสถิติที่ซับซ้อนมากยิ่งขึ้น หรืออาจเลือกใช้โปรแกรมสำเร็จรูปอื่น ๆ ที่มีในปัจจุบันเป็นทางเลือกในการวิเคราะห์ค่า p-value ของสถิติทดสอบ Ryan-Joiner และ Sürücü “C” เช่น โปรแกรม Minitab เป็นต้น

5.3.2 ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป

จากการศึกษาวิจัยครั้งนี้ ผลที่ได้จากงานวิจัยได้นำเสนอการแจกแจงของความเบ้และความโด่งในบางสถานการณ์ หากมีการศึกษาค้นคว้าเพิ่มเติมของตัวสถิติที่ใช้ในการทดสอบกลุ่มอื่น ๆ ที่มีการพัฒนาขึ้นมาใหม่จะทำให้เป็นประโยชน์มากยิ่งขึ้น อีกทั้งสามารถนำการจำลองด้วยวิธีเพาเวอร์ไปประยุกต์ใช้สร้างข้อมูลเพื่อทดสอบกับวิธีสถิติทดสอบใหม่ ๆ เพิ่มเติมอีกด้วย

รายการอ้างอิง

ภาษาไทย

- วรพรรณ เจริญขำ. (2556). การตรวจสอบค่านอกเกณฑ์ในตัวอย่างสุ่มจากประชากรปกติ. (วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต), สถาบันบัณฑิตพัฒนบริหารศาสตร์, คณะสถิติประยุกต์.
- ศุภกิจ วงศ์วิวัฒน์นุกิจ. (2550). พจนานุกรมศัพท์การวิจัยและสถิติ. กรุงเทพฯ: โครงการตำรา คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย.
- สำนักงานราชบัณฑิตยสภา. (2558). พจนานุกรมศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสภา. กรุงเทพฯ: สำนักพิมพ์คณะรัฐมนตรีและราชกิจจานุเบกษา.

ภาษาอังกฤษ

- Asghar, G., & Saleh, Z. (2012). Normality Tests for Statistical Analysis : A Guide for Non-Statisticians. *International Journal endocrinology Metabolism*, 10(2), 486–489.
- B. Efron. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annal of Statistics*, 7, 1-26.
- Baris Sürücü. (2008). A power comparison and simulation study of goodness-of-fit tests. *Computers and Mathematics with Applications*, 56(6), 1617-1625.
- Fernando, M.-R., & Jorge, G.-B. (2013). A Power Comparison of Various Tests of Univariate Normality on Ex-Gaussian Distributions. *Methodology*, 9, 137-149.
- Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, 43(4), 521–532.
- Hadi, A. N., & Naser, R. A. (2011). Monte Carlo comparison of seven normality tests. *Journal of Statistical Computation and Simulation*, 81(8), 965-972.
- Hao Luo. (2011). Generation of Non-normal Data - A Study of Fleishman's Power Method. *Working Paper Department of Statistics Uppsala University*, 1-29.
- Mehmet, M., & Akin, P. (2013). Type I Error Rate and Power of Three Normality Tests. *Pakistan Journal of Information and Technology*, 2(2), 135-139.

Ryan, T. A., & Joiner, B. L. (1976). Normal Probability Plots and Tests for Normality.

Statistics Department, The Pennsylvania State University, 1-13.

S. S. Shapiro, & M. B. Wilk. (1965). An Analysis of Variance Test for Normality

(Complete Samples). *Biometrika*, 52, 591-611.

Skewness. (2013). Retrieved June, 2016 from <http://en.wikipedia.org>

Spotlight. (2014). Kurtosis. Retrieved June, 2016 from <http://schaal15.blog.sbc.edu>





ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

การวิเคราะห์ข้อมูลด้วยโปรแกรม R

ส่วนที่ 1 : การจำลอง a, b, c และ d จากความเบ้และความโด่งที่แบ่งออกเป็น 4 กรณี ๆ ละ 4 แบบ
ด้วยวิธีเพาเวอร์ Fleishman (1978) ดังนี้

คำสั่ง

Skewness and kurtosis values of these variables are as follows:

```
library(BinOrdNonNor)
```

```
# (1) Symmetric-Platykurtic Distribution
```

```
skew.vec <- c(0,0,0,0)
```

```
kurto.vec <- c(-1,1,2,3)
```

```
coef.est1 <- Fleishman.coef.NN(skew.vec, kurto.vec)
```

```
coef.est1
```

```
# (2) Symmetric-Leptokurtic Distribution
```

```
skew.vec <- c(0,0,0,0)
```

```
kurto.vec <- c(5,10,20,50)
```

```
coef.est2 <- Fleishman.coef.NN(skew.vec, kurto.vec)
```

```
coef.est2
```

```
# (3) Asymmetric-Platykurtic Distribution
```

```
skew.vec <- c(-1,-1,1,1)
```

```
kurto.vec <- c(2,3,2,3)
```

```
coef.est3 <- Fleishman.coef.NN(skew.vec, kurto.vec)
```

```
coef.est3
```

```
# (4) Asymmetric-Leptokurtic Distribution
```

```
skew.vec <- c(-2,-2,1,1)
```

```
kurto.vec <- c(10,50,10,50)
```

```
coef.est4 <- Fleishman.coef.NN(skew.vec, kurto.vec)
```

```
coef.est4
```

ส่วนที่ 2 : จำลองตัวแปรสุ่ม X จาก Standard Normal ที่ขนาดตัวอย่าง 10, 20, 30, 50 และ 100

จำลองหา Y จากสมการ Fleishman's โดยที่ $Y = a + bx + cx^2 + dx^3$

นำข้อมูลจำลอง Y ไปคำนวณค่า p-value ของสถิติทดสอบ 3 การทดสอบ ได้แก่

สถิติทดสอบ Shapiro-Wilk, Ryan-Joiner และ Sürücü "C" แล้วหาอำนาจการทดสอบ

คำสั่ง

Find p-value of Shapiro-Wilk, Ryan-Joiner and Sürücü "C" for each test 500 times

```
simround<-500
```

```
n<-10
```

```
pvalue1vec<-c()
```

```
pvalue2vec<-c()
```

```
pvalue3vec<-c()
```

```
for(j in 1:simround) {
```

```
  ##generate x from standard normal##
```

```
  x<-rnorm(n)
```

```
  ##### Power Method #####
```

```
  # skewness=0 , kurtosis=-1
```

```
  # Find Value of a, b, c and d from Fleishman (1987) #
```

```
  a=0
```

```
  b=1.2210090
```

```
  c=0
```

```
  d=-0.08015816
```

```
  # generate y #
```

```
  sampleY<-matrix(0,n,1)
```

```
  for (i in 1:n){
```

```
    y<-a+b*(x)+c*(x^2)+d*(x^3)
```

```
    sampleY[i,]=y[i]
```

```
  }
```

```
  Y<-as.vector(sampleY)
```

```

# Normality testing #
##### Shapiro Wilk #####
#####
shapiro.test(Y)
result <- shapiro.test(Y)
pvalue1<-result$p.value
pvalue1vec<-c(pvalue1vec,pvalue1)

##### Ryan-Joiner #####
#####
p<-c()
for(i in 1:n){
  phi<-(i-0.375)/(n+0.25)
  p<-c(p,phi)
}
rcal<-cor(qnorm(p),sort(Y))
rcal
#using bootstrap simulation for Ryan-Joiner
B=2000
r<-c()
for(i in 1:B){
  resample<-sample(sort(Y),n,replace=TRUE)
  q<-cor(qnorm(p),sort(resample))
  r<-c(r,q)
}
count.r<-length(r[r<=rcal])
pvalue2<-count.r/B
pvalue2vec<-c(pvalue2vec,pvalue2)

```

```
##### Surucu "C" #####
#####

library(moments)
skewness(Y)
kurtosis(Y)

p2<-c()
for(i in 1:n){
  phi2<-i/(n+1)
  p2<-c(p2,phi2)
}
rho<-cor(sort(Y),qnorm(p2))
b1<-(skewness(Y))^2
b2<-kurtosis(Y)
resultCcal<- shapiro.test(sort(Y))
SW<-resultCcal$statistic
R<-1-(rho)^2
a1<-exp(-(b1/0.6)^5)
a2<-exp(-(b2/3.5)^5)
Ccal<-1-(((1+a1*(a2-1))*SW)+(a1*(1-a2)*(1-R)))
Ccal
#using bootstrap simulation for Surucu "C"
B=2000
C<-c()
for(i in 1:B){
  resample<-sample(sort(Y),n,replace=TRUE)
  resultC<-shapiro.test(sort(resample))
  SWc<-resultC$statistic
```

```

t<-qnorm((1:n)/(n+1))
Rc<-1-(cor(sort(Y),t))^2
C<-c(C,1-(((1+a1*(a2-1))*SWc)+((a1*(1-a2))*(1-Rc))))
}
count.C<-length(C[C>=Ccal])
pvalue3<-count.C/B
pvalue3vec<-c(pvalue3vec,pvalue3);
}
pvalue1vec
pvalue2vec
pvalue3vec

#####Power of test #####
#####
# alpha is 0.05 #
count.pvalue1<-length( pvalue1vec[ pvalue1vec<0.05])
powerW<-count.pvalue1/500
count.pvalue2<-length( pvalue2vec[ pvalue2vec<0.05])
powerRJ<-count.pvalue2/500
count.pvalue3<-length( pvalue3vec[ pvalue3vec<0.05])
powerC<-count.pvalue3/500
powerW
powerRJ
powerC

```

หมายเหตุ :

จากคำสั่งข้างต้น เป็นคำสั่งคำนวณอำนาจการทดสอบสำหรับกรณีการแจกแจงด้วยความเบ้และความโด่ง (0,-1) และ $n=10$ ที่ระดับนัยสำคัญ 0.05 ซึ่งในกรณีอื่น ๆ และขนาดตัวอย่างอื่น ๆ สามารถเปลี่ยนคำสั่งที่ค่า a, b, c และ d สำหรับการแจกแจงกรณีนั้น ๆ และทำการกำหนดขนาดตัวอย่างใหม่ตั้งแต่ $n= 10, 20, 30, 50$ และ 100

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวศิราภรณ์ เข้มทิศ เกิดวันพุธที่ 21 มีนาคม พ.ศ. 2533 สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาคณิตศาสตร์ ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ในปีการศึกษา 2556 และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาสถิติ ภาควิชาสถิติ คณะพาณิชยศาสตร์และการบัญชี จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2557

