รายงานการวิจัย

เรื่อง

การเปรียบเทียบประสิทธิภาพของการทดสอบการแจกแจงปกติโดยการใช้โปรแกรมสำเร็จรูปทางสถิติ

(A Comparison of the Efficiency for Normality Test by Statistical Packages)

โดย

รศ.อุมาพร จันทศร

ได้รับทุนสนับสนุนงานวิจัยจากเงินงบรายได้ ประจำปีงบประมาณ 2551

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ส่วนที่ 1 รายละเอียดเกี่ยวกับโครงการ

ชื่อโครงการ (ภาษาไทย) การเปรียบเทียบประสิทธิภาพของการทดสอบการแจกแจงปกติโดยใช้โปรแกรมสำเร็จรูปทางสถิติ

(ภาษาอังกฤษ) A Comparison of the Efficiency for Normality Test by Statistical Packages

ได้รับทุนอุดหนุนการวิจัยจาก งบรายได้คณะวิทยาศาสตร์ สจล.

ประจำปี 2551 จำนวนเงิน 20,000 บาท

ระยะเวลาทำการวิจัย 1 ปี ตั้งแต่ 1 ต.ค. 2550 ถึง ก.ย. 2551

หน่วยงานและผู้ดำเนินการวิจัยพร้อมหน่วยงานที่สังกัดและเลขหมายโทรศัพท์

สาขาวิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

รศ.อุมาพร จันทศร โทร. 02-3264111 ต่อ 6278

ส่วนที่ 2 บทคัดย่อ

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อทำการเปรียบเทียบประสิทธิภาพในการทดสอบการแจกแจงปกติจากสถิติทดสอบ 3 แบบ คือ Kolmogorov, Lilliefors และ Shapiro – Wilk ที่มีอยู่ใน SPSS v.13 และจากสถิติทดสอบ 3 แบบ คือ Kolmogorov, Anderson-Darling และ Ryan-Joiner ที่มีอยู่ใน MINITAB 14 ข้อมูลที่นำมาศึกษาเป็นข้อมูลจากการจำลองแบบจากประชากรที่มีการแจกแจงปกติ และที่เบี่ยงเบนไปเล็กน้อยจากการแจกแจงปกติ โดยใช้เมนู Calc-Random Data จากโปรแกรมสำเร็จรูปทางสถิติ MINITAB ในแต่ละกรณีมีการทำซ้ำ 500 ครั้ง ของการจำลองแบบข้อมูล มีขนาดตัวอย่างเท่ากับ 10, 20, 30, 50 และ 100 ทำการเปรียบเทียบอำนาจการทดสอบ และความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ I

ผลการวิจัย พบว่าสถิติทดสอบ Ryan-Joiner (คล้าย Shapiro-Wilk) ใน MINITAB มีอำนาจการทดสอบสูงที่สุดในทุกกรณี และทุกขนาดตัวอย่าง ที่ระดับนัยสำคัญ 0.10 โดยเฉพาะเมื่อขนาดตัวอย่างใหญ่เท่ากับ 100 จะมีอำนาจการทดสอบใกล้ 1 และมีความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ I ได้เกือบทั้งหมด ตามเกณฑ์ของ Cochran และยังพบว่าสถิติ Kolmogorov ใน MINITAB แท้ที่จริงคือสถิติ Lilliefors

The purpose of this study is to compare the efficiency of three tests statistics for testing normality ; Kolmogorov, Lilliefors and Shapiro – Wilk, which are available in SPSS v13.0. and three tests statistics; Kolmogorov, Anderson-Darling and Ryan – Joiner, which are available in MINITAB 14. The data for this study was yielded from simulation, by the method of Monte Carlo, under conditions of Normal distribution and slightly depart from the Normal distribution, by using Menu Calc-Random Data

from MINITAB 14. In each situation was done 500 iterations with different sample size ; 10, 20, 30, 50 and 100. Comparisons of Type I error and power of the test among the six test statistics were done.

It was found that Ryan-Joiner Test (Similar to Shapiro-Wilk) in MINITAB Test had the highest power of the test in all cases and all sample sizes, at 0.10 significance level, especially for the sample size of 100, it had the power of the test almost to be 1. And it also had ability to control probability of Type I error in almost situations under criteria of Cochran. Even though K-S test in MINITAB appears under the appellation of the K-S test, it is really the Lilliefors test.

คำสำคัญ (Keywords) : Normality Test, Kolmogorov Test, Lilliefors Test, Shapiro-Wilk Test, Anderson-Darling, Ryan-Joiner, SPSS, MINITAB

คำนำ

ปัจจุบันโปรแกรมสำเร็จรูปทางสถิติจะถูกใช้เป็นเครื่องมือที่สำคัญในการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติ โปรแกรม SPSS และ MINITAB เป็นโปรแกรมที่แพร่หลายในประเทศไทย เพราะจัดหาได้ง่าย และมีคู่มือการใช้ที่แปลด้วยภาษาไทยมากมาย สำหรับการวิเคราะห์เพื่อตรวจสอบค่าตัวแปรสุ่มที่สนใจว่ามีการแจกแจงปกติหรือไม่ โปรแกรมทั้งสองมีสถิติทดสอบให้เลือกใช้ 3 แบบ จึงน่าจะหาข้อสรุปว่า สถิติใดจึงจะมีประสิทธิภาพสูงสุด สามารถนำไปใช้ได้ในสถานการณ์ต่าง ๆ

ผลของงานวิจัยนี้จะเป็นประโยชน์สำหรับนักวิจัยทั่ว ๆ ไป และจะทำให้เลือกใช้สถิติวิเคราะห์ได้ถูกต้อง เหมาะสมกับลักษณะข้อมูล และข้อกำหนดเบื้องต้นอย่างแท้จริง

สุดท้ายนี้ผู้วิจัยขอขอบคุณภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ที่อุดหนุนงบประมาณจำนวน 20,000 บาท สำหรับงานวิจัยนี้ รวมทั้งนักศึกษาชั้นปีที่ 4 ปีการศึกษา 2551 คือ นางสาววรพรรณ เจริญขำ และนายวทัญญู ทองสุข ที่สละเวลาวิเคราะห์ข้อมูลด้วยโปรแกรม SPSS และ MINITAB จนสำเร็จลุล่วง

รศ.อุมาพร จันทศร
11 พฤษภาคม 2552

สารบัญ

สารบัญตาราง

เรื่อง		หน้า

สารบัญรูป

บทที่ 1
บทนำ

1.1 ความสำคัญของปัญหา

วิธีการทางสถิติมีบทบาทสำคัญมากในงานวิจัยในสาขาต่าง ๆ กล่าวคือสามารถสรุปผลถึงกลุ่มใหญ่ด้วยการใช้ตัวแทนของกลุ่มใหญ่นั้น โดยจะสรุปถึงกลุ่มใหญ่ในแง่ความน่าจะเป็น จึงมีการนำวิธีการทางสถิติไปใช้อย่างกว้างขวาง นอกจากผู้ใช้จะต้องมีความรู้ความเข้าใจในเรื่องที่ศึกษาเป็นอย่างดีแล้ว จำเป็นต้องมีความรู้ทางด้านสถิติพอสมควร เช่น ทราบถึงข้อจำกัดและความสอดคล้องระหว่างสถิติที่ใช้กับวัตถุประสงค์ของงานวิจัยนั้น ๆ เป็นต้น ในทางทฤษฎีสถิติอาจแยกการวิเคราะห์ได้ 2 ประเภทใหญ่ ๆ คือ สถิติแบบพาราเมตริก (Parametric Statistics) และสถิติแบบไม่ใช้พารามิเตอร์ (Nonparametric Statistics) โดยสถิติแบบแรกต้องมีข้อจำกัดที่สำคัญคือ ตัวแปรสุ่มต้องมีการแจกแจงปกติ (1) ดังนั้นการแจกแจงปกติจึงเป็นการแจกแจงที่สำคัญสำหรับการวิเคราะห์แบบพาราเมตริก เมื่อผู้วิจัยต้องการใช้วิธีการจากสถิติแบบนี้ จึงจำเป็นต้องตรวจสอบในลำดับแรกก่อนว่าตัวแปรสุ่มนั้นมีการแจกแจงปกติหรือไม่ ถ้าเป็นจริงก็เลือกใช้สถิติแบบพาราเมตริก แต่ถ้าไม่จริง วิธีที่นิยมใช้คือการแปลงข้อมูล (Transformation) เช่น แปลงเป็นค่าล็อค (Log) แล้วทดสอบอีกครั้งว่ามีการแจกแจงปกติหรือไม่ ขั้นสุดท้ายต้องเลือกใช้สถิติแบบไม่ใช้พารามิเตอร์ เมื่อข้อกำหนดเบื้องต้นไม่เป็นจริง (2)

การตรวจสอบการแจกแจงปกติ มีหลายวิธีจากง่ายที่สุดคือดูจากกราฟ ไปจนถึงการใช้สถิติทดสอบ (3) เช่น R.A. Fisher (1923-1930) เป็นผู้คิดค้นสถิติทดสอบตัวแรกที่ใช้ในการทดสอบการแจกแจงปกติ เมื่อไม่ทราบค่าเฉลี่ย และความแปรปรวนของประชากร โดยเรียกสถิติทดสอบนั้นว่า Standard Third Moment $\left(\sqrt{b_1}\right)$ และ Standard Fourth Moment $\left(b_2\right)$ และได้มีการพัฒนาสถิติทดสอบขึ้นมาอีกมากมาย เช่น Anderson & Darling (AD) (4) หรือ Kolmogorov – Smirnov (K – S) (5) หรือ Shapiro – Wilk (S – W) (6) ที่เป็นที่รู้จักกันโดยทั่วไปในหมู่นักวิจัย นอกจากนี้มีการคิดค้นสถิติใหม่ ๆ อีกมาก เช่น Chen & Shapiro (1995) หรือ Zhang (1999) (7) ซึ่งใช้หลักการของการทดสอบการถดถอยและสหสัมพันธ์เหมือนการทดสอบของ S – W ในขณะที่สองการทดสอบแรกใช้หลักการของ Empirical Distribution Function Test (EDF) และยังมีการพัฒนาตัวสถิติทดสอบที่ใช้ความเบ้และความโด่งของโค้งการแจกแจง (Test based on Skewness and Kurtosis) มาทดสอบการแจกแจงปกติ คือ D’ Agostine (1990), Park (1999)

สถิติทดสอบที่โปรแกรมสำเร็จรูปทางสถิติมีให้ผู้ใช้เลือกส่วนใหญ่ คือ สถิติ Kolmogorov-smirnov (K-S), Shapiro-Wilk (S-W), Anderson-Darling (AD), Lilliefors, Cramer-von Mises, Shapiro-Francia เป็นต้น โปรแกรมสำเร็จรูปทางสถิติ SPSS และ MINITAB เป็นโปรแกรมที่แพร่หลายในประเทศไทย และมีสถิติทดสอบที่ใช้ทดสอบการแจกแจงปกติหลายแบบ คือ SPSS มีสถิติทดสอบ 1-sample K-S, Lilliefors และ S-W ส่วนโปรแกรม MINITAB จะมีสถิติทดสอบ AD, RJ (Ryan-Joiner ที่กล่าวว่าคล้ายสถิติ Shapiro-Wilk) และ K-S

ตามทฤษฎีสถิติทดสอบ K-S ควรจะใช้กับตัวอย่างขนาดเล็ก เพราะจะได้ค่าวิกฤตจากการแจกแจงที่แท้จริง (Conover หน้า 77 (8), Gibbons หน้า 176 (9)) และการตั้งสมมติฐานเบื้องต้น เมื่อจะใช้สถิติทดสอบ K-S ตามทฤษฎีจำเป็นต้องระบุพารามิเตอร์ให้ครบถ้วน (เช่น H_0 : X มีการแจกแจงปกติด้วยค่าเฉลี่ย 100 ความแปรปรวน = 25 เป็นต้น) ในกรณีอื่นนอกจากนี้ไม่แนะนำให้ใช้สถิติทดสอบนี้อย่างเด็ดขาด (Sprent หน้า 68) แต่ในโปรแกรม SPSS ด้วยเมนู Nonparametric Test ผู้ใช้ไม่ต้องระบุค่าพารามิเตอร์ แต่ผลการวิเคราะห์จะหมายเหตุว่าใช้ค่าเฉลี่ยและความแปรปรวนของตัวอย่างเป็นค่าประมาณแทน และในเมนู Help ก็กล่าวว่าถ้าใช้วิธีการเช่นนี้ การแจกแจงของสถิติทดสอบ K – S จะเปลี่ยนไป แนะนำให้ผู้ใช้เปลี่ยนไปใช้เมนู Explore แทน ซึ่งวิธีการเช่นนี้จะทำให้ผู้ใช้ที่ไม่มีความรู้ทางสถิติเพียงพอ อาจเลือกใช้สถิติทดสอบที่ไม่เหมาะสม ในกรณีเช่นนี้เมนู Nonparametric Test ควรที่จะมีการปรับปรุง Algorithm ให้ผู้ใช้ระบุพารามิเตอร์ หรือไม่เช่นนั้นก็ไม่ควรจะมีการตรวจสอบการแจกแจงปกติในเมนูนี้ เช่นเดียวกันในโปรแกรม MINITAB ก็ไม่ได้ให้ความสำคัญในเรื่องนี้ และไม่มีคำอธิบายเพิ่มเติม

ตามทฤษฎีสถิติทดสอบ K – S เหมาะสมที่จะใช้ทดสอบการแจกแจงแบบต่อเนื่องเท่านั้น (Sprent หน้า 77) แต่ใน SPSS สามารถเลือกทดสอบการแจกแจง Poisson ซึ่งเป็นการแจกแจงแบบไม่ต่อเนื่อง ผลสรุปที่ได้จะเชื่อถือได้หรือไม่

สถิติทดสอบแบบ Lilliefors ได้ถูกพัฒนาเพื่อใช้แทนที่สถิติทดสอบ K-S แต่ไม่จำเป็นต้องระบุค่าพารามิเตอร์ (คือค่าเฉลี่ยและความแปรปรวน) ที่ครบถ้วนในสมมติฐานเบื้องต้น แต่นักวิจัยทั่ว ๆ ไปไม่คุ้นเคยกับสถิติทดสอบนี้ การแจกแจงของสถิติทดสอบจะมีหลายกรณี เช่น ทราบค่าเฉลี่ย แต่ไม่ทราบค่าความแปรปรวน และกรณีสุดท้ายไม่ทราบค่าทั้งคู่ แต่ในโปรแกรม SPSS ผู้ใช้จะใช้ได้ในกรณีเดียวคือไม่ทราบค่าพารามิเตอร์ทั้งคู่ SPSS ได้คำนวณค่า p (p – value) จากการแจกแจงที่ถูกต้องหรือไม่

สถิติทดสอบ K-S และ Lilliefors ควรจะมีการปรับค่า (Correction for the Test Statistic) เพื่อพิจารณาความแตกต่างของโค้งการแจกแจงความน่าจะเป็นสะสม 2 เส้น ทางขวามือด้วยสูตร

$$D = \max_{1 \leq i \leq r} \left\{ \max\left[\left|S(x_i) - F_0(x_i)\right|, \left|S(x_{i-1}) - F_0(x_i)\right| \right] \right\}$$

ซึ่งจะทำให้เห็นความแตกต่างอย่างแท้จริง

สถิติทดสอบทั้งสอง ถูกนำไปใช้ในโปรแกรมสำเร็จรูปทางสถิติหลาย ๆ โปรแกรม แต่อาจจะใช้ชื่อที่ทำให้ผู้ใช้สับสน เช่น โปรแกรม S-Plus แม้จะใช้ชื่อว่าสถิติ K-S แต่แท้ที่จริงคือสถิติ Lilliefors แต่ในโปรแกรม R และ Matlab แยกสถิติทดสอบทั้งสองออกจากกันให้ผู้ใช้เลือกใช้ (10) และในโปรแกรม SPSS จะเรียกสถิติทดสอบแบบ K-S ว่า 1-Sample K-S (เลือกจากเมนู Analyze-Nonparametric Test) และเรียกสถิติ Lilliefors ว่า สถิติ K-S โดยมีหมายเหตุว่า คือ Lilliefors (เลือกจากเมนู Analyze-Descriptive Statistics-Explore-Plots-Normality Plots with tests) (3) ในโปรแกรม MINITAB จะใช้ชื่อว่าสถิติ K-S แต่แท้ที่จริงคือสถิติใด (ไม่มีการอธิบายเพิ่มเติมใน Help) โดยใช้เมนู Basic-Statistics-Normality Test

มีผลงานวิจัยเกี่ยวกับการตรวจสอบการแจกแจงปกติด้วยสถิติ K-S ซึ่งได้ผลสรุปสอดคล้องกันมากมายว่าสถิติทดสอบ K-S ไม่ควรถูกนำมาใช้เพราะมีประสิทธิภาพต่ำ (10), (11) แม้จะใช้ตัวอย่างขนาดใหญ่และกำหนดระดับนัยสำคัญ (α) ให้ใหญ่ขึ้นก็ตาม คือสถิติ K-S มักจะยอมรับว่าตัวแปรสุ่มนั้นมีการแจกแจงปกติ แม้ว่าที่แท้จริงจะมีการแจกแจงแบบอื่น ๆ และมีงานวิจัยที่พบว่าจากโปรแกรม SPSS ที่มีสถิติทดสอบการแจกแจงปกติให้เลือกใช้ 3 แบบนั้น สถิติ K-S มีประสิทธิภาพต่ำที่สุด เมื่อเทียบกับสถิติทดสอบ S-W และ Lilliefors (12)

สำหรับสถิติ S-W งานวิจัยส่วนใหญ่พบว่ามีประสิทธิภาพสูงสามารถใช้ได้ในสถานการณ์ทั่ว ๆ ไป (12) (13) (14) โปรแกรม SPSS แนะนำให้ใช้สถิติ S-W นี้ เมื่อขนาดตัวอย่างน้อยกว่า 50 ซึ่งขัดแย้งกับผลงานวิจัยที่พบว่าถ้าต้องการให้มีอำนาจการทดสอบเข้าใกล้ 1 ควรใช้ขนาดตัวอย่างเข้าใกล้ 100 (12)

ส่วนสถิติทดสอบ AD นั้น การแจกแจงของสถิติทดสอบจะต่างกันไปตามกรณีต่าง ๆ เช่น ทราบค่าเฉลี่ยและความแปรปรวนของประชากร หรือทราบเพียงค่าใดค่าหนึ่ง หรือไม่ทราบทั้งคู่ ซึ่งในกรณีนี้ควรใช้สถิติแบบปรับค่า (Correction) (11 หน้า 90) แต่ในโปรแกรม MINITAB จะมีเพียงกรณีสุดท้าย คือ ประมาณค่าพารามิเตอร์จากตัวอย่าง และให้การทดสอบนี้เป็น Default ของโปรแกรม ซึ่งหมายความว่าโปรแกรมนี้ให้ความสำคัญกับสถิติทดสอบนี้มากกว่าสถิติอื่น ๆ ถ้าผู้ใช้ไม่มีความรู้มากพอ นำผลวิเคราะห์ของสถิติ AD นี้ไปใช้ อาจจะได้ค่าอำนาจการทดสอบต่ำกว่าสถิติทดสอบอื่น ๆ ก็ได้

จากคุณลักษณะของสถิติทดสอบต่าง ๆ ที่กล่าวมาข้างต้น จึงน่าสงสัยว่าโปรแกรมสำเร็จรูป SPSS และ MINITAB ได้คำนึงถึงหรือไม่ มีการสร้าง Algorithm ถูกต้องตามทฤษฎีสถิติหรือไม่ โปรแกรมทั้งสองมีสถิติ K-S รวมอยู่ด้วย แม้จะมีผลงานวิจัยมากมายสนับสนุนว่า K-S มีประสิทธิภาพต่ำ อาจจะทำให้ผู้ใช้ที่ไม่ทราบรายละเอียดนำผลสรุปที่คลาดเคลื่อนไปใช้ รวมทั้งใน MINITAB สถิติ K-S ที่แท้จริงคือสถิติใดระหว่าง Kolmogorov-Smirnov หรือ Lilliefors นอกจากนี้ในโปรแกรมทั้งสองยังมีสถิติทดสอบที่เหมือนกันอีก 1 แบบ คือ สถิติ SW (ใน SPSS) และสถิติ RJ (ใน MINITAB) ดังนั้น จึงน่าสงสัยว่าการสร้าง Algorithm ของโปรแกรมทั้งสองได้คำนึงถึงการแจกแจงที่แท้จริงของสถิติทดสอบหรือไม่ สถิติตัวใดจะมีอำนาจการทดสอบสูงที่สุดในสถานการณ์ต่าง ๆ และโปรแกรมใดจะให้ผลสรุปที่น่าเชื่อถือมากกว่ากัน เพื่อผู้ใช้จะได้เลือกใช้และมั่นใจในผลสรุปมากที่สุด

ผลการวิเคราะห์ทางสถิติของโปรแกรมสำเร็จรูปต่าง ๆ พบว่ายังมีข้อผิดพลาด หรือทำให้เข้าใจผิด (Misleading) ในหลายประเด็น เช่น การศึกษาของกมล (15,16) หรือของ Bergmann, Ludbrook and Spooren (17) เกี่ยวกับการปรับค่าซ้ำ (Correction for ties) ของสถิติทดสอบ Wilcoxon – Mann – Whitney หรือการศึกษาของ Leo Knüsel (18) ที่พบว่าค่าความน่าจะเป็น (Probability) จากการแจกแจงต่าง ๆ ที่คำนวณจาก Microsoft Excel 97 และ 2003 ยังไม่ถูกต้องในหลายกรณี และการศึกษาของ McCullough และ Wilson (19) พบว่าการใช้เทคนิคสถิติ 3 เทคนิค คือ การวิเคราะห์การถดถอย (ทั้งกรณีเชิงเส้นและไม่ใช่เชิงเส้น) การผลิตเลขสุ่ม และการแจกแจงแบบต่าง ๆ ยังไม่ถูกต้อง จากการใช้ Microsoft Excel 97 และแนะนำว่าไม่ควรวิเคราะห์สถิติเหล่านี้โดยใช้โปรแกรมสำเร็จรูป Excel และกรณีสุดท้ายเป็นการศึกษาของกุศยา (20) ที่

เปรียบเทียบประสิทธิภาพของการจำลองแบบข้อมูลด้วยโปรแกรม SAS และ MINITAB ที่พบว่าสำหรับตัวอย่างขนาดเล็กทั้ง 2 โปรแกรมให้ผลไม่ต่างกัน แต่สำหรับตัวอย่างขนาดใหญ่ แนะนำให้ใช้โปรแกรม SAS มากกว่า

1.2 วัตถุประสงค์

1.2.1 จากสถิติทดสอบ 3 แบบในโปรแกรม MINITAB และโปรแกรม SPSS หาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และพิจารณาว่ามีความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ตามเกณฑ์ของ Cochran และ Bradley ได้หรือไม่

1.2.2 จากสถิติทดสอบ 3 แบบในโปรแกรม MINITAB และโปรแกรม SPSS หาค่าอำนาจการทดสอบของสถิติแต่ละแบบ (พิจารณาเฉพาะสถิติทดสอบที่มีความสามารถในการควบคุมคลาดเคลื่อนชนิดที่ 1) ในแต่ละสถานการณ์ของขนาดตัวอย่างและระดับนัยสำคัญของการทดสอบ

1.2.3 เปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบ 3 แบบในโปรแกรม MINITAB และ SPSS เพื่อหาข้อสรุปว่า สถิติตัวใดมีอำนาจการทดสอบสูงที่สุดในแต่ละสถานการณ์

1.2.4 เปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบในโปรแกรม MINITAB และ SPSS ที่คล้ายกัน คือ สถิติ K-S (ใน MINITAB) กับสถิติ 1-sample K-S หรือ Lilliefors (ใน SPSS) และสถิติ RJ (ใน MINITAB) กับสถิติ S-W (ใน SPSS)

1.2.5 หาผลสรุปว่า สถิติ K-S ใน MINITAB ที่แท้จริงคือสถิติใดระหว่าง K-S กับ Lilliefors (โดยใช้ผลจากค่าความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 และค่าอำนาจการทดสอบจากสถิติ 1-sample K-S และ Lilliefors ใน SPSS เปรียบเทียบ)

1.3 ข้อตกลงเบื้องต้น

1.3.1 การวิจัยครั้งนี้ถือว่า ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบทั้ง 3 แบบ ใน SPSS และ MINITAB เป็นดัชนีสำคัญที่ผู้วิจัยใช้เป็นเกณฑ์ในการสรุปผลว่าสถิติทดสอบใดควรจะถูกนำไปใช้

1.3.2 ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จะใช้เกณฑ์ของ Cochran และ Bradley ดังนี้

เกณฑ์ของ Cochran จะมีเกณฑ์การพิจารณาว่าสามารถควบคุมความน่าจะเป็นได้ เมื่อ

ที่ระดับนัยสำคัญ .10 ถ้าค่าที่ได้มีค่าในช่วง (.08 - .12)

ที่ระดับนัยสำคัญ .05 ถ้าค่าที่ได้มีค่าในช่วง (.04 - .06)

เกณฑ์ของ Bradley จะมีเกณฑ์การพิจารณาว่าสามารถควบคุมความน่าจะเป็นได้ เมื่อ

ที่ระดับนัยสำคัญ .10 ถ้าค่าที่ได้มีค่าในช่วง (.05 - .15)

ที่ระดับนัยสำคัญ .05 ถ้าค่าที่ได้มีค่าในช่วง (.025 - .075)

1.3.3 การสร้างข้อมูลให้มาจากประชากรแบบต่าง ๆ ได้จากการจำลองข้อมูลจากโปรแกรม MINITAB ด้วยเมนู Calc-Random Data โดยเชื่อมั่นว่าข้อมูลที่ได้มาจากประชากรที่มีการแจกแจงตามที่กำหนด ด้วยผลงานวิจัยของกุศยา (15) ที่ได้ผลสรุปงานวิจัยว่า MINITAB มีประสิทธิภาพใกล้เคียงกับ SAS ในการจำลองข้อมูลให้มาจากการแจกแจงแบบต่าง ๆ เมื่อข้อมูลมีขนาดเล็ก

1.4 ขอบเขตของการวิจัย

1.4.1 ประชากรที่สุ่มตัวอย่างมาเพื่อทดสอบการแจกแจงปกติ จะมีการแจกแจงสมมาตรที่เบี่ยงเบนไปจากการแจกแจงปกติเล็กน้อย ดังต่อไปนี้

- การแจกแจงปกติโดยสนใจศึกษาเมื่อพารามิเตอร์ $\mu = 100$ และ $\sigma = 10, 50, 100$
- การแจกแจงสมมาตรที่มีหางยาวกว่าปกติ (Long tailed distribution) คือ การแจกแจงแบบที d.f = 1 และ 5
- การแจกแจงสมมาตรแต่โด่งกว่าปกติมาก (Distributions with high kurtosis) เช่น การแจกแจงCauchy ด้วยพารามิเตอร์ 0, 2 คือ C(0, 2) และ C(0, 0.5) และการแจกแจง Laplace คือ L(0, 0.5)
- การแจกแจงสมมาตรแต่โด่งกว่าปกติเล็กน้อย (Distribution with kurtosis slightly higher than the Normal) เช่น Logistic คือ L(0, 0.5) และ L(0, 1)
- การแจกแจงสมมาตรที่มีหางสั้น (Short tailed distribution) เช่น Uniform แบบ U(0 – 1) และ U (10 – 50)

1.4.2 จะสุ่มตัวอย่างข้อมูลจากประชากรที่มีการแจกแจงต่าง ๆ ด้วยขนาดตัวอย่างเป็น 10, 20, 30, 50 และ 100

1.4.3 ทดลองใช้เมนูของ SPSS คือ Explore (ซึ่งมีสถิติทดสอบ K-S แบบ Lilliefors และ Shapiro-Wilk, S-W) และเมนู 1-Sample K-S (มีสถิติทดสอบ K-S) ทดสอบการแจกแจงปกติจากข้อมูลตัวอย่างที่สุ่มมาจากประชากรแบบต่าง ๆ เป็นจำนวนซ้ำ 500 ครั้ง หรือใช้ข้อมูลตัวอย่างจากประชากรหนึ่ง ๆ 500 ชุด

1.4.4 ทดลองใช้เมนูของ MINITAB คือ Basic Statistics-Normality Test (ซึ่งมีสถิติทดสอบ K-S, AD และ RJ) ทดสอบการแจกแจงปกติด้วยข้อมูลตัวอย่างในตอนต้น (ซึ่งใช้ในหัวข้อ 1.4.3 กับ SPSS) เป็นจำนวนซ้ำ 500 ครั้ง

1.5 คำนิยามเชิงปฏิบัติการ

ความคลาดเคลื่อนชนิดที่ 1 (Type I error) หมายถึง ความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมติฐานเบื้องต้น เมื่อสมมติฐานเบื้องต้นเป็นจริง ความน่าจะเป็นหรือโอกาสที่จะเกิดความคลาดเคลื่อนชนิดที่ 1 จะกำหนดเบื้องต้นไว้ที่ $\alpha = .05$ หรือ .10 ในงานวิจัยนี้จะเกิดความคลาดเคลื่อน

ชนิดที่ 1 นี้ เมื่อมีผลการทดสอบพบว่า ด้วยเมนูต่าง ๆ ของ SPSS และ MINITAB แล้ว มีนัยสำคัญทางสถิติของการทดสอบจากข้อมูลตัวอย่างของประชากรที่มีการแจกแจงปกติ

ความคลาดเคลื่อนชนิดที่ 2 (Type II error) หมายถึงความคลาดเคลื่อนที่เกิดจากการยอมรับสมมติฐานเบื้องต้น เมื่อสมมติฐานเบื้องต้นเป็นเท็จ ความน่าจะเป็นหรือโอกาสที่จะเกิดความคลาดเคลื่อนชนิดที่ 2 กำหนดด้วย β ในงานวิจัยนี้จะเกิดความคลาดเคลื่อนชนิดที่ 2 นี้ เมื่อมีผลการทดสอบด้วยเมนูต่าง ๆ ของ SPSS และ MINITAB แล้ว พบว่าไม่มีนัยสำคัญทางสถิติของการทดสอบจากข้อมูลตัวอย่างของประชากรที่มีการแจกแจงแบบอื่น ๆ ที่ไม่ใช่การแจกแจงปกติ

อำนาจการทดสอบ (Power of the Test) หมายถึง ความน่าจะเป็นที่จะปฏิเสธสมมติฐานเบื้องต้น เมื่อสมมติฐานเบื้องต้นเป็นเท็จ ซึ่งมีค่าเท่ากับ $1-\beta$ ในงานวิจัยนี้จะคือความน่าจะเป็นของการมีนัยสำคัญทางสถิติจากการทดสอบด้วยเมนูต่าง ๆ ของ SPSS และ MINITAB เมื่อใช้ข้อมูลตัวอย่างจากประชากรที่มีการแจกแจงอื่น ๆ ที่ไม่ใช่การแจกแจงปกติ

ความคลาดเคลื่อนชนิดที่ 1 จากผลการวิจัยนี้ หมายถึง สัดส่วนของจำนวนครั้งของการปฏิเสธสมมติฐานเบื้องต้น (คือได้ค่า p (p-value) น้อยกว่า .05 หรือ .10) จากข้อมูลตัวอย่างจากประชากรที่มีการแจกแจงปกติ ต่อจำนวนครั้งที่ทดลองทั้งหมด คือ 500 ครั้ง

อำนาจการทดสอบ จากผลการวิจัยนี้ หมายถึง สัดส่วนของการปฏิเสธสมมติฐานเบื้องต้น (คือได้ค่า p (p-value) น้อยกว่า .05 หรือ .10) จากข้อมูลตัวอย่างจากประชากรที่มีการแจกแจงอื่น ๆ ที่ไม่ใช่การแจกแจงปกติ ต่อจำนวนครั้งที่ทดลองทั้งหมด คือ 500 ครั้ง

1.6 ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย

ภายใต้ขั้นตอนวิธี (Algorithm) ของการสร้างระบบของโปรแกรม SPSS และ MINITAB

1.6.1 ทำให้สามารถเลือกใช้สถิติทดสอบเพื่อทดสอบการแจกแจงปกติจากโปรแกรม SPSS ในแต่ละกรณี (ขนาดตัวอย่าง ระดับนัยสำคัญ) ให้ได้ผลที่น่าไว้วางใจอย่างแท้จริง (valid test)

1.6.2 ทำให้สามารถเลือกใช้สถิติทดสอบเพื่อทดสอบการแจกแจงปกติจากโปรแกรม MINITAB ในแต่ละกรณี (ขนาดตัวอย่าง ระดับนัยสำคัญ) ให้ได้ผลที่น่าไว้วางใจอย่างแท้จริง (Valid test)

1.6.3 เพื่อเป็นข้อสรุปให้นักวิจัยที่ต้องการทดสอบการแจกแจงปกติ ว่าควรจะใช้โปรแกรมใดระหว่าง SPSS และ MINITAB จึงจะได้อำนาจการทดสอบสูงที่สุด ภายใต้สถานการณ์ต่าง ๆ (ขนาดตัวอย่าง และระดับนัยสำคัญต่าง ๆ)

บทที่ 2
เอกสาร และงานวิจัยที่เกี่ยวข้อง

ในบทนี้ผู้วิจัยจะนำเสนอแยกเป็น 4 ตอน ดังนี้

ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับการแจกแจงปกติ

ตอนที่ 2 สถิติทดสอบสำหรับการทดสอบการแจกแจงปกติ จากเมนูของ SPSS และ MINITAB คือสถิติทดสอบ 1 sample Kolomogorov-Smirnov (1-Sample K-S), Lilliefors, Shapiro-Wilk (S-W) และ Anderson Darling (AD)

ตอนที่ 3 การแจกแจงที่เกี่ยวข้องกับงานวิจัยนี้

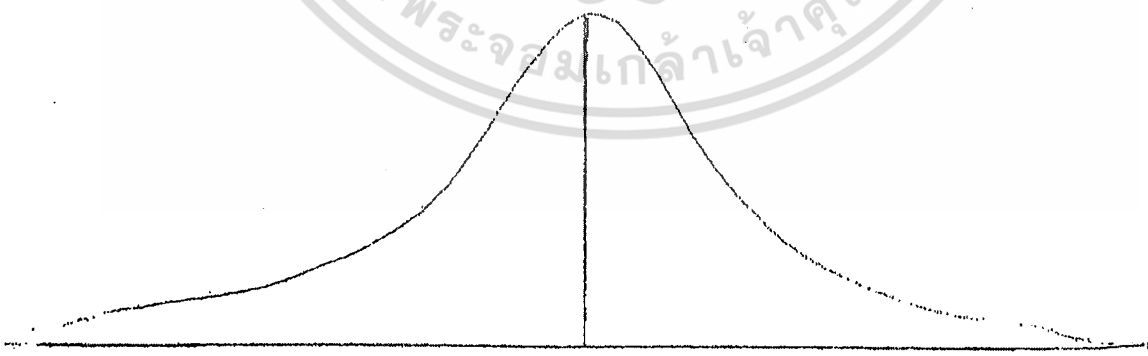
ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

ตอนที่ 1 ความรู้ทั่วไปเกี่ยวกับการแจกแจงปกติ

การแจกแจงแบบปกติ (Normal Distribution)

เป็นการแจกแจงชนิดพื้นฐานของตัวแปรสุ่มชนิดต่อเนื่อง (continuous random variable) ซึ่งมีความสำคัญมากในทฤษฎีสถิติสมัยใหม่ และนิยมใช้กันอย่างแพร่หลายในงานวิจัยทั่วไป

กราฟของการแจกแจงแบบปกติเรียกว่า เส้นโค้งปกติ (Normal curve) มีลักษณะคล้ายระฆังคว่ำ (Bell shape) ชนิดสมมาตร ดังรูป 2.1 ซึ่งมีค่าเฉลี่ย = ฐานนิยม = มัธยฐาน อยู่ตรงกลางของโค้งปกติ มักจะพบว่าข้อมูลที่เกิดตามธรรมชาติทั่ว ๆ ไป เช่น อายุของผู้ป่วย ความสูงของคนในช่วงอายุหนึ่ง ความดันโลหิต ปริมาณน้ำฝนในช่วงฤดูฝนของท้องที่หนึ่ง มีการแจกแจงแบบปกติ



รูปที่ 2.1 แสดงโค้งปกติ

ถ้า X เป็นตัวแปรสุ่มชนิดต่อเนื่อง ซึ่งมีค่า $-\alpha < x < \alpha$ และมีฟังก์ชันความหนาแน่นของความน่าจะเป็น (pdf)

คือ $$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left\{(x-\mu)^2/2\sigma^2\right\}} \quad \text{เมื่อ} \quad -\alpha < x < \alpha$$

โดยที่ X เป็นตัวแปรสุ่มแบบต่อเนื่อง

μ เป็นค่าเฉลี่ยของตัวแปรสุ่ม X

σ เป็นส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X

e และ π เป็นค่าคงที่ โดย e = 2.71828 และ π = 3.1416

แล้วจะเรียกการแจกแจงของตัวแปรสุ่ม X นี้ว่า มีการแจกแจงปกติ (normal distribution) ด้วยพารามิเตอร์ ค่าเฉลี่ย = μ และความแปรปรวน = σ^2 โดย $-\alpha < \mu < \alpha$ และ $\sigma > 0$ นิยมใช้สัญลักษณ์ $X \sim N(\mu, \sigma^2)$ แทนความหมายที่ว่า ตัวแปรสุ่ม X มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย = μ และความแปรปรวน = σ^2

คุณสมบัติทั่วไปของการแจกแจงแบบปกติ

1. $f(x) \geq 0$
2. $\int_{-\alpha}^{\alpha} f(x)dx = 1$ นั่นคือพื้นที่ทั้งหมดภายใต้เส้นโค้งปกติเท่ากับ 1
3. ความน่าจะเป็นของค่า X ในช่วงหนึ่ง ๆ เช่น $P(a \leq X \leq b) = \int_a^b f(x)dx$ หาได้จากพื้นที่ภายใต้โค้งระหว่างค่า X = a ถึง X = b
4. ความน่าจะเป็นของค่า X ที่ X_0 ใด ๆ จะมีค่าเป็น 0 เนื่องจาก $\int_{x_0}^{x_0} f(x)dx = 0$

คุณสมบัติของโค้งปกติ

1. ลักษณะของโค้งเป็นรูประฆังคว่ำที่มีลักษณะสมมาตร (Symmetry) คือ ถ้าแบ่งครึ่งตามแนวตั้งแล้วพื้นที่ครึ่งซ้าย = พื้นที่ครึ่งขวา = 0.5

2. ค่าสูงสุดของเส้นโค้งปกติ มีเพียงจุดเดียว คือ จุดที่อยู่ตรงกลางพอดี ซึ่งมีค่าเท่ากับค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม

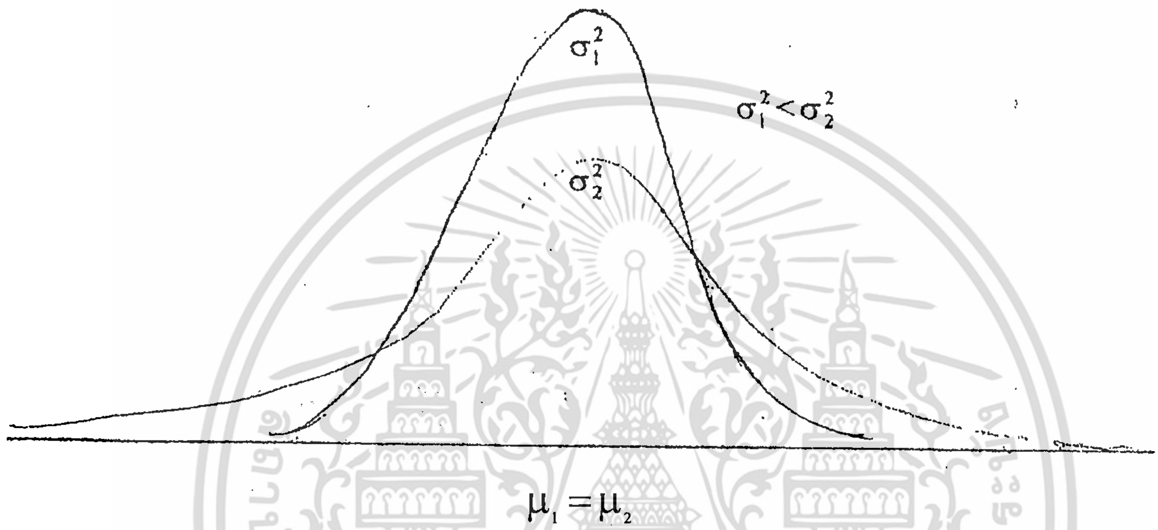
3. ค่าเฉลี่ยของตัวแปรสุ่ม X = μ ความแปรปรวน = σ^2

4. โดยทางทฤษฎีแล้ว เส้นโค้งจะค่อย ๆ แผ่ออกไปทั้งสองข้าง และเข้าใกล้แกนระนาบมากขึ้น แต่จะไม่ตัดแกน คือวิ่งเข้าสู่ค่าอนันต์ทั้งสองข้าง

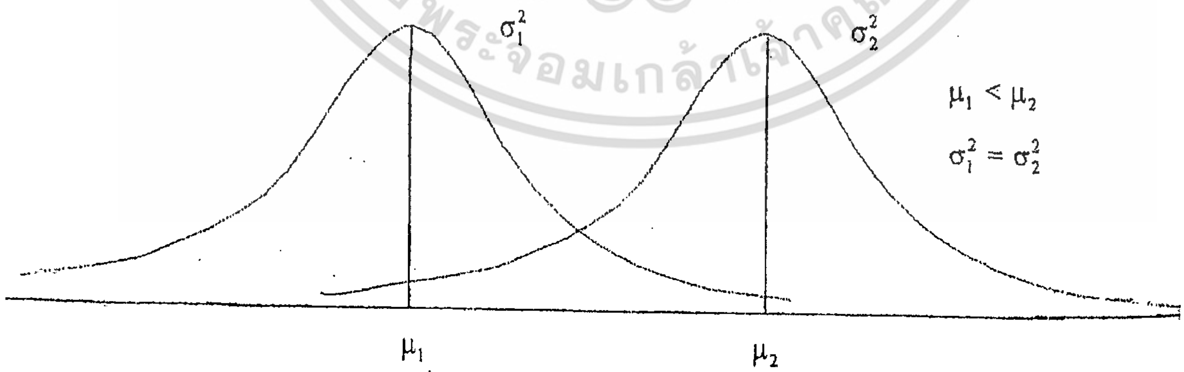
5. สำหรับข้อมูลที่ต่างก็มีการแจกแจงแบบปกติ แต่มีค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานต่างกัน จะมีรูปโค้งปกติที่ต่างกัน เช่น ข้อมูลชุดที่มีความแปรปรวนน้อยกว่าจะมีความโด่งของข้อมูลมากกว่า ข้อมูลที่มีความแปรปรวนสูงกว่า

อาจแยกได้หลายกรณี ดังนี้

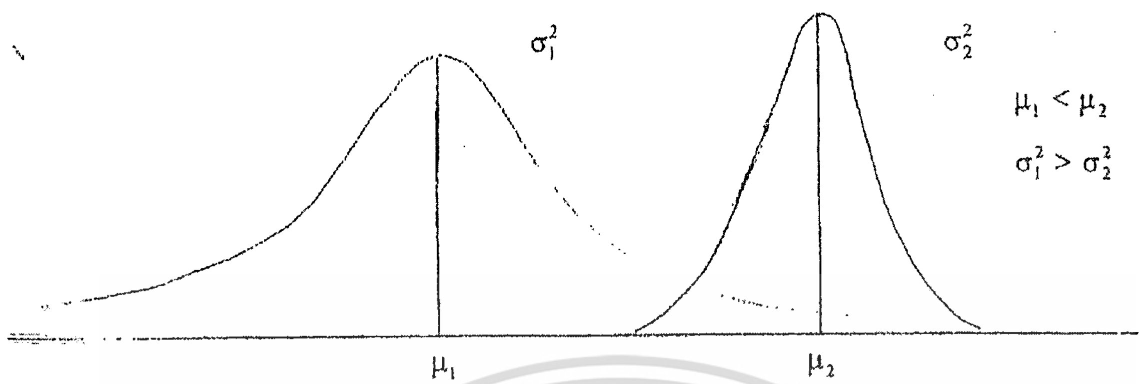
ก. ข้อมูล 2 ชุด ที่มีค่าเฉลี่ยเท่ากัน แต่ความแปรปรวนต่างกัน รูปโค้งปกติที่ได้คือ



ข. ข้อมูล 2 ชุด ที่มีค่าเฉลี่ยต่างกัน แต่ส่วนเบี่ยงเบนมาตรฐานเท่ากัน รูปโค้งปกติที่ได้คือ



ค. ข้อมูล 2 ชุด ที่มีค่าเฉลี่ยต่างกัน และส่วนเบี่ยงเบนมาตรฐานต่างกัน รูปโค้งปกติที่ได้ คือ



ในความเป็นจริงแล้ว โดยทั่วไปไม่ค่อยพบโค้งปกติตามคุณสมบัติที่กำหนดให้ทางคณิตศาสตร์ที่แท้จริงเลย (ที่เห็นเด่นชัดคือคุณสมบัติข้อ 4) ฉะนั้นจึงเป็นเพียงการประมาณให้เป็นโค้งปกติ เพราะโค้งปกติสามารถอธิบายเกี่ยวกับรูปร่างของการแจกแจงความถี่ของข้อมูลจำนวนมาก และสามารถนำไปประยุกต์ใช้กับข้อมูลใด ๆ ก็ได้ นอกจากนี้ขบวนการทางสถิติ เช่น การประมาณค่า การทดสอบสมมติฐานและอื่น ๆ **ส่วนใหญ่มีข้อตกลงกันว่าการแจกแจงของข้อมูลต้องเป็นการแจกแจงแบบปกติ**

การแจกแจงแบบปกติมาตรฐาน (Standard Normal Distribution)

จากตัวแปรสุ่ม X ซึ่งมีการแจกแจงแบบปกติด้วยค่าเฉลี่ย μ ความแปรปรวน = σ^2 ที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็น (pdf) คือ

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\{(x-\mu)^2/2\sigma^2\}}$$

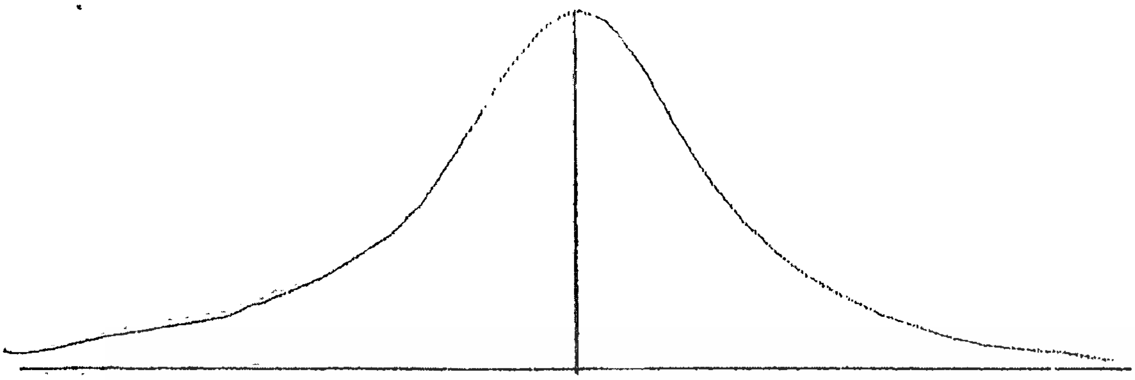
จะเห็นว่า การอินทิเกรทเพื่อหาค่าความน่าจะเป็นของค่า x ในช่วงใด ๆ จะทำได้ยุ่งยาก และขึ้นกับค่า μ และ σ^2 ที่ต่างกัน เพื่อที่จะทำให้ง่ายขึ้น สามารถแปลงตัวแปรสุ่ม X (transform) ให้เป็นตัวแปรใหม่ที่เป็นมาตรฐาน หรือ Z ซึ่งเรียกว่าคะแนนมาตรฐาน (Standard score)

โดย $Z = \dfrac{x-\mu}{\sigma}$

แล้วจะได้ว่าตัวแปร Z มีการแจกแจงแบบปกติด้วย โดยมีค่าเฉลี่ย = 0 ความแปรปรวน = 1 และมีฟังก์ชันความหนาแน่นของความน่าจะเป็น คือ

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{1/2} z^2$$

และเรียกการแจกแจงของ Z นี้ว่าเป็นการแจกแจงแบบปกติมาตรฐาน (Standard Normal Distribution) กราฟของรูปโค้งของ f(z) ซึ่งเรียกว่า รูปโค้งปกติมาตรฐานแสดงได้ดังรูป 2.2



รูปที่ 2.2 แสดงโค้งปกติมาตรฐาน ซึ่งมีค่าเฉลี่ย = 0 ความแปรปรวน = 1

ดังนั้น อาจเขียนเป็นสัญลักษณ์ได้ว่า Z~N(0, 1)

พิสูจน์

1. E(Z) = 0
2. Var(Z) = 1

$$E(Z) = \frac{E[x-\mu]}{\sigma}$$

$$= \frac{1}{\sigma}E(x-\mu)$$

$$= \frac{1}{\sigma}[E(x)-\mu]$$

$$= \frac{1}{\sigma}(\mu-\mu)$$

$$= 0$$

$$Var(Z) = E(Z^2)-[E(Z)]^2$$

$$= E(Z^2)-0$$

$$= E\left[\frac{x-\mu}{\sigma}\right]^2$$

$$= \frac{1}{\sigma^2}E(x-\mu)^2$$

$$= \frac{\sigma^2}{\sigma^2}$$

$$= 1$$

คะแนนมาตรฐานมีประโยชน์คือสามารถใช้เปรียบเทียบข้อมูลต่างชุด ซึ่งมีค่าเฉลี่ยความแปรปรวนต่างกันได้ ดังตัวอย่างต่อไปนี้

ตัวอย่าง ในการสอบครั้งหนึ่ง วิชาสถิติมีคะแนนเฉลี่ย 48 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 10 คะแนน ส่วนวิชาอังกฤษมีคะแนนเฉลี่ย 59 คะแนน ส่วนเบี่ยงเบนมาตรฐาน 8 คะแนน ถ้านาย ก. สอบวิชาอังกฤษได้คะแนน = 72 และสอบวิชาสถิติได้คะแนน 69 จะกล่าวได้หรือไม่ว่านาย ก. สอบวิชาอังกฤษได้ดีกว่าวิชาสถิติ

วิธีทำ เนื่องจากคะแนนเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของ 2 วิชาต่างกัน ไม่สามารถเปรียบเทียบกันด้วยคะแนนดิบได้ จำเป็นต้องเปลี่ยนเป็นคะแนนมาตรฐานก่อนจะได้คะแนนมาตรฐานของแต่ละวิชาของนาย ก. คือ

คะแนนมาตรฐานของวิชาสถิติ $= \dfrac{69-48}{10} = 2.1$

คะแนนมาตรฐานของวิชาอังกฤษ $= \dfrac{72-59}{8} = 1.62$

คะแนนมาตรฐานของวิชาสถิติสูงกว่าคะแนนมาตรฐานของวิชาอังกฤษ ดังนั้น นาย ก. สอบวิชาสถิติได้ดีกว่าวิชาอังกฤษ

พื้นที่ภายใต้เส้นโค้งปกติมาตรฐาน

ดังกล่าวแล้วว่า $\int_{-\alpha}^{\alpha} f(x)\,dx$ = พื้นที่ภายใต้โค้งปกติ = 1 และเมื่อได้ Z ที่เป็นตัวแปรที่มีการแจกแจงแบบปกติด้วย $\mu = 0$, $\sigma^2 = 1$ ดังนั้น พื้นที่ภายใต้โค้งปกติมาตรฐาน = 1 ด้วย ถ้าต้องการหาความน่าจะเป็นของตัวแปรสุ่ม x ในช่วงใดช่วงหนึ่งก็สามารถคำนวณได้โดยการอินทิเกรท (Integrate) หาพื้นที่ภายใต้เส้นโค้งในช่วงนั้น ๆ โดยอาจอินทิเกรทจาก f(x) หรือจาก f(z) ก็ได้

แต่การอินทิเกรท f(x) จะยุ่งยากมากเพราะขึ้นกับค่า μ และ σ จึงได้มีนักคณิตศาสตร์คำนวณหาพื้นที่ภายใต้เส้นโค้ง f(z) ในช่วง z ต่าง ๆ และได้สร้างตารางสำเร็จรูปที่แสดงค่าความน่าจะเป็นของ z ในช่วงต่าง ๆ จึงเป็นการสะดวกในการหาความน่าจะเป็นในช่วงใดช่วงหนึ่งของตัวแปรสุ่ม x ที่มีการแจกแจงแบบปกติ โดยการเปลี่ยนค่าของตัวแปรสุ่ม x ให้เป็นคะแนนมาตรฐาน Z และหาพื้นที่ภายใต้เส้นโค้ง f(z)

ตอนที่ 2 สถิติทดสอบสำหรับการทดสอบการแจกแจงปกติ

1. การทดสอบของโคลโมโกรอฟ (The Kolmogorov test หรือเรียกว่า one-sample Kolmogorov-Smirnov Test)

แบบทดสอบที่จะใช้เพื่อทดสอบสารูปสนิทดีจากข้อมูลที่เป็นแบบต่อเนื่อง หรือกับข้อมูลที่วัดมาอย่างน้อยจากมาตราแบบอันดับนั้น ถูกคิดค้นขึ้นในปี ค.ศ. 1933 โดยนักคณิตศาสตร์ชาวรัสเซียชื่อ โคลโมโกรอฟ (A.N. Kolmogorov) ได้แนะนำขบวนการสำหรับใช้ทดสอบสารูปสนิทดีกับข้อมูลของ 1 ตัวอย่าง ซึ่งเป็นขบวนการที่จะทดสอบว่าตัวอย่างที่สุ่มมาจากประชากรมีการแจกแจงตามที่คาดหวังหรือไม่ และต่อมาในปี ค.ศ. 1939 นักคณิตศาสตร์ชาวรัสเซียอีกคนหนึ่งชื่อ สมิร์นอฟ (N.V Smirnov) ได้แนะนำขบวนการสำหรับใช้ทดสอบสารูปสนิทดีสำหรับข้อมูลของ 2 ตัวอย่าง ซึ่งเป็นขบวนการที่จะทดสอบว่า 2 ตัวอย่างมาจากประชากรเดียวกันหรือมาจาก 2 ประชากรที่เหมือนกัน โดยขบวนการทดสอบของนักคณิตศาสตร์ทั้งสองท่านจะพิจารณาจากฟังก์ชันการแจกแจงแบบเอ็มไพริคัล (Empirical distribution) หรือฟังก์ชันการแจกแจงแบบสะสม (Cumulative distribution) ซึ่งหมายความว่าถ้าสุ่มตัวอย่างจากประชากรที่ไม่ทราบฟังก์ชันการแจกแจง F(x) สามารถจะสรุปได้หรือไม่ว่า $F(x) = F^*(x)$ สำหรับทุก ๆ ค่าของ X เมื่อ $F^*(x)$ เป็นฟังก์ชันการแจกแจงที่ทราบการแจกแจง และถ้า $F(x) = F^*(x)$ แล้วสามารถคาดหวังว่า $F^*(x)$ และ S(x) ซึ่งเป็นฟังก์ชันการแจกแจงเอ็มไพริคัลของตัวอย่าง จะมีความสอดคล้องใกล้เคียงกันมากนอกเหนือจากความผันแปรที่เกิดจากการสุ่ม

ข้อมูล ข้อมูลประกอบด้วยค่าสังเกตของตัวอย่างสุ่ม $X_1, X_2, ..., X_n$ ขนาด n ที่มีฟังก์ชันการแจกแจง F(x) ซึ่งไม่ทราบรูปแบบการแจกแจง โดยตัวแปรสุ่ม X_i ต้องมีมาตราวัดข้อมูลอย่างน้อยแบบเรียงลำดับ (ordinal Scale)

ข้อสมมุติ ตัวอย่างต้องเป็นตัวอย่างสุ่ม

สถิติทดสอบ

ให้ S(x) เป็นฟังก์ชันการแจกแจงแบบเอ็มไพริคัลจากตัวอย่างสุ่ม $X_1, X_2, ..., X_n$ ซึ่งนิยามได้โดย

$$S(x) = \frac{\#(X \le x)}{n}$$

เมื่อ $\#(X \le x)$ แทนจำนวนของ X ที่น้อยกว่าหรือเท่ากับ x ซึ่งการแจกแจงของ S(x) เป็นดังนี้

$$S(x) = \begin{cases} 0, & x < x_{(1)} \\ \frac{k}{n}, & x_{(k)} \le x \le x_{(k+1)}; k = 2,3,...,n-1 \\ 1, & x \ge x_{(n)} \end{cases}$$

เมื่อ $X_{(k)}$ เป็นสถิติลำดับที่ k ของ X สถิติทดสอบนิยามได้แตกต่างกันเป็น 3 กลุ่มตามลักษณะของสมมุติฐานที่จะกล่าวถึงต่อไป และให้ $F^*(x)$ เป็นฟังก์ชันการแจกแจงตามสมมุติฐานและทราบรูปแบบการแจกแจง

ก. การทดสอบแบบสองหาง (Two – tailed test)

$H_0 \quad : \quad F(x) = F^*(x)$ สำหรับทุกค่าของ $x, -\infty < x < \infty$

$H_1 \quad : \quad F(x) \neq F^*(x)$ สำหรับบางค่าของ x

ให้สถิติทดสอบ T เป็นระยะห่างที่กว้างที่สุดตามแนวตั้งระหว่าง S(x) และ $F^*(x)$ นั่นคือ

$$T = \sup_x \left| F^*(x) - S(x) \right|$$

ข. การทดสอบแบบหางเดียวด้านล่าง (One – sided test)

$H_0 \quad : \quad F(x) \leq F^*(x), -\infty < x < \infty$

$H_1 \quad : \quad F(x) > F^*(x)$ อย่างน้อย 1 ค่าของ x

สถิติทดสอบนิยามได้โดย

$$T^- = \sup_x \left[S(x) - F^*(x) \right]$$

(พิจารณาเฉพาะระยะห่างตามแนวตั้งที่มากที่สุด เมื่อ S(x) อยู่ข้างบน $F^*(x)$)

ค. การทดสอบแบบหางเดียวด้านบน (One – sided test)

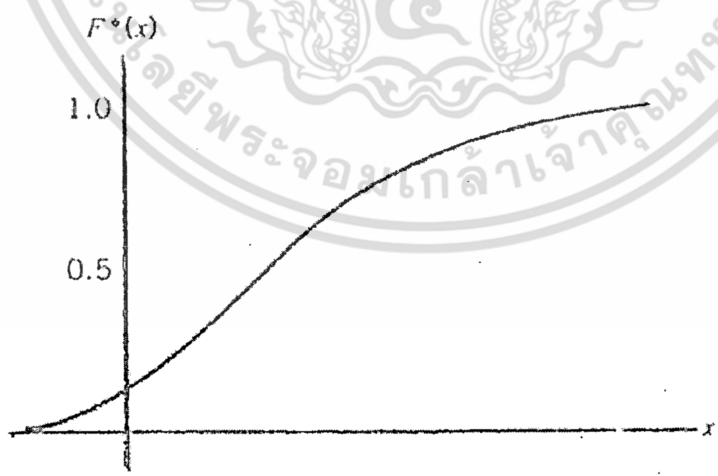
$H_0 \quad : \quad F(x) \geq F^*(x), -\infty < x < \infty$

$H_1 \quad : \quad F(x) < F^*(x)$ อย่างน้อย 1 ค่าของ x

สถิติทดสอบกำหนดได้โดย

$$T^+ = \sup_x \left| F^*(x) - S(x) \right|$$

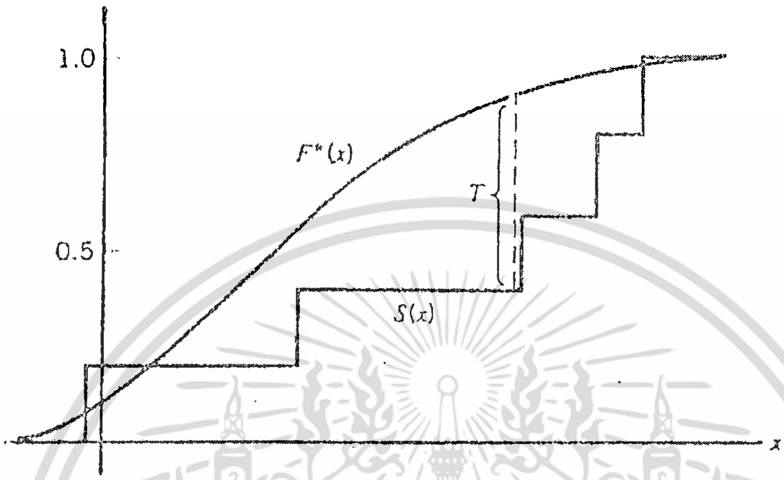
(พิจารณาเฉพาะระยะห่างตามแนวตั้งที่มากที่สุดเมื่อฟังก์ชัน $F^*(x)$ อยู่ข้างบน S(x))



รูปที่ 2.3 ฟังก์ชันการแจกแจงตามสมมุติฐาน

ถ้าข้อมูลตัวอย่าง สุ่มมาจากการแจกแจงที่ระบุไว้ในสมมุติฐาน ความแตกต่างระหว่าง S(x) และ $F^*(x)$ ของค่าสังเกต X ควรที่จะไม่มาก นั่นคือมีความสอดคล้องระหว่าง S(x) และ $F^*(x)$ สำหรับทุกค่าสังกต X

เมื่อ H_0 เป็นจริง ในทางตรงกันข้าง ถ้า H_0 ไม่จริงหรือตัวอย่างไม่ได้มาจากการแจกแจงที่ระบุไว้ในสมมุติฐาน ความแตกต่างระหว่าง S(x) และ $F^*(x)$ ควรที่จะมีค่ามากจนสามารถสังเกตเห็นได้ รูปการแจกแจงที่ใช้แสดงฟังก์ชันการแจกแจงของสมมุติฐาน $F^*(x)$ ฟังก์ชันการแจกแจงเอ็มไพริคัล S(x) และสถิติของโคลโมโกรอฟ แสดงได้ดังนี้



รูปที่ 2.4 ฟังก์ชันการแจกแจงตามสมมุติฐาน $F^*(x)$, ฟังก์ชันการแจกแจงเอ็มไพริคัล S(x), และสถิติโคลโมโกรอฟ

การพิจารณาการประมาณการแจกแจงของสถิติทดสอบ อาจพิจารณาได้จากการที่ F(x) มีการแจกแจงแบบต่อเนื่องและสมมุติฐานว่างเป็นจริง ซึ่งฟังก์ชันการแจกแจงที่แท้จริงของ T^+ และ T^- กำหนดโดย

$$G(x) = 1 - x \sum_{j=0}^{[n(1-x)]} \binom{n}{j} \left(1 - x - \frac{j}{n}\right)^{n-j} \left(x + \frac{j}{n}\right)^{j-1}$$

ซึ่ง $[n(1-x)]$ คือจำนวนเต็มที่ใหญ่ที่สุดที่มีค่าน้อยกว่าหรือเท่ากับ $n(1-x)$ และทั้ง T^+ และ T^- จะเป็นการแจกแจงเหมือนกัน ฟังก์ชันการแจกแจงแบบแอสซิมโทติก $(n \to \infty)$ ของ $\sqrt{n}T^+$ และ $\sqrt{n}T^-$ หาได้จาก

$$H(x) = \lim_{n \to \infty} G\left(\frac{x}{\sqrt{n}}\right) = 1 - C^{-2x^2}$$

และฟังก์ชันประมาณการแจกแจงของ T คือ

$$\Pr(T \le x) \square [G(x)]^2$$

ซึ่ง T จะน้อยกว่า x เมื่อทั้ง T^+ และ T^- มีค่าน้อยกว่า x เท่านั้น

ควอนไทล์ที่แท้จริงของ T ในการทดสอบแบบสองหางและค่าควอนไทล์โดยประมาณของ T^+ และ T^- ในการทดสอบแบบหางเดียวจะแสดงในตารางที่ 1 ในภาคผนวกสำหรับ $n \leq 40$ และการประมาณแบบแอสซิมโทติกจะใช้เมื่อ $n > 40$

หมายเหตุ ตารางการแจกแจงของค่าควอนไทล์ของ T จะให้ผลถูกต้องเมื่อ F(x) เป็นการแจกแจงแบบต่อเนื่องเท่านั้น แต่ถ้า F(x) มีการแจกแจงเป็นแบบไม่ต่อเนื่อง จะใช้วิธีอื่นหาการแจกแจงของ T

การทดสอบสมมุติฐาน

ก. การทดสอบแบบสองหาง (Two – tailed test)

$H_0 \quad : \quad F(x) = F^*(x)$ สำหรับทุกค่าของ $x, -\infty < x < \infty$

$H_1 \quad : \quad F(x) \neq F^*(x)$ สำหรับบางค่าของ x

จะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า T มากกว่าควอนไทล์ที่ $1-\alpha$ ที่หาได้จากตารางที่ 1 และค่าพี (p-value) หาได้จาก

ค่าพี = 2 x (ค่าระดับนัยสำคัญสังเกตหาได้จากการทดสอบแบบหางเดียว)

ค่าพีที่แท้จริงจากการทดสอบแบบหางเดียว คือ

$$t \sum_{j=0}^{[n(1+t)]} \binom{n}{i} \left(1-t-\frac{j}{n}\right)^{n-j} \left(t+\frac{j}{n}\right)^{j-1}$$

เมื่อ t คือค่าสังเกตของสถิติทดสอบและ $[n(1+t)]$ คือจำนวนเต็มที่ใหญ่ที่สุดที่มีค่าน้อยกว่าหรือเท่ากับ n(l+t)

หมายเหตุ ค่าประมาณค่าพีอาจหาได้โดยใช้ตารางที่ 1

ข. การทดสอบแบบหางเดียวด้านล่าง (One – tailed test)

$H_0 \quad : \quad F(x) \leq F^*(x), -\infty < x < \infty$

$H_1 \quad : \quad F(x) > F^*(x)$ อย่างน้อย 1 ค่าของ x

จะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า T^- มากกว่าควอนไทล์ที่ $1-\alpha$ ของการทดสอบแบบหางเดียวจากตารางที่ 1 ค่าพีที่แท้จริงหาได้จากสูตรหาระดับนัยสำคัญสังเกตข้างต้น หรืออาจประมาณได้จากตารางที่ 1

ค. การทดสอบแบบหางเดียวด้านบน (One – tailed test)

$H_0 \quad : \quad F(x) \geq F^*(x), -\infty < x < \infty$

$H_1 \quad : \quad F(x) < F^*(x)$ อย่างน้อย 1 ค่าของ x

จะปฏิเสธ H_0 ที่ระดับนัยสำคัญ α ถ้า T^+ มากกว่าควอนไทล์ที่ $1-\alpha$ ที่หาได้จากตารางที่ 1 และค่าระดับนัยสำคัญสังเกตหาได้ในทำนองเดียวกับข้อ ข.

หมายเหตุ การทดสอบสารูปสนิทดีของโคลโมโกรอฟจะใช้ได้ดีเมื่อตัวอย่างสุ่มมีการแจกแจงเป็นไปตามฟังก์ชันการแจกแจงบางอย่างที่ระบุไว้ชัดเจน นั่นคือเมื่อฟังก์ชันการแจกแจงในสมมุติฐานไม่มีตัวพารามิเตอร์ที่ไม่ทราบค่าที่จะต้องประมาณจากตัวอย่าง แต่ถ้าฟังก์ชันการแจกแจงในสมมุติฐานมีตัวพารามิเตอร์ที่ไม่ทราบค่าที่จะต้องประมาณจากตัวอย่าง แล้ว การทดสอบแบบนี้จะให้ผลไม่ถูกต้อง

ตัวอย่าง ตัวอย่างสุ่มขนาด 10 มีค่าสังเกตดังนี้ $X_1 = 0.621$, $X_2 = 0.503$, $X_3 = 0.203$, $X_4 = 0.477$, $X_5 = 0.710$, $X_6 = 0.581$, $X_7 = 0.329$, $X_8 = 0.480$, $X_9 = 0.554$ และ $X_{10} = 0.382$ จงทดสอบว่าข้อมูลนี้มาจากประชากรที่มีการแจกแจงแบบเอกรูป

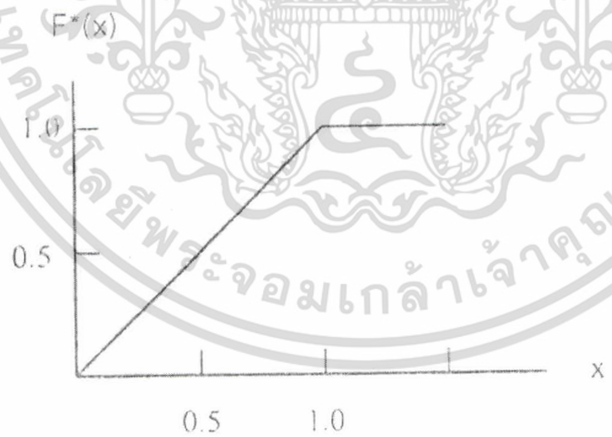
จากสิ่งกำหนดให้ข้างต้น สมมุติฐานของการทดสอบเป็น

H_0 : ประชากรมีการแจกแจงแบบเอกรูป

ดังนั้นฟังก์ชันการแจกแจงภายใต้สมมุติฐานว่าง H_0 คือ

$$F^*(x) = \begin{cases} 0 & \text{ถ้า} \quad x < 0 \\ x & \text{ถ้า} \quad 0 \le x < 1 \\ 1 & \text{ถ้า} \quad x \ge 1 \end{cases}$$

ซึ่งมีกราฟเป็น



รูปที่ 2.5 ฟังก์ชันการแจกแจงของสมมติฐาน

ดังนั้นสมมุติฐานของการทดสอบเขียนใหม่ได้เป็น

H_0 : $F(x) = F^*(x)$ สำหรับทุก ๆ ค่าของ x

H_1 : $F(x) \neq F^*(x)$ อย่างน้อย 1 ค่าของ x

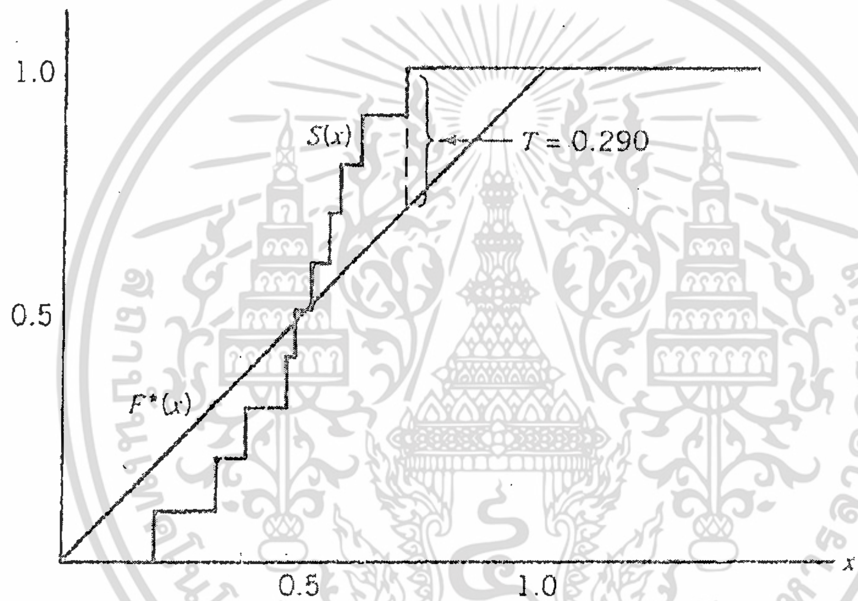
เมื่อ F(x) เป็นฟังก์ชันการแจกแจงที่ไม่ทราบรูปแบบสำหรับทุก ๆ x_i และ $F^*(x)$ เป็นฟังก์ชันการแจกแจงภายใต้สมมุติฐานว่าง จากสมมุติฐานใช้ การทดสอบสารูปสนิทดีของโคลโมโกรอฟแบบสองหางและจาก

ตารางที่ 1 ควอนไทล์ของสถิติทดสอบโคลโมโกรอฟ เมื่อ n = 10 บริเวณวิกฤตที่ระดับนัยสำคัญ α = 0.05 คือ T มากกว่าควอนไทล์ที่ 0.95 ซึ่งจากตารางที่ 1 พบว่ามีค่าเท่ากับ 0.409 และค่าสถิติทดสอบสังเกตคำนวณได้จาก

$$T = \sup_x \left| F^*(x) - S(x) \right|$$

เมื่อ $$S(x) = \frac{\text{จำนวนของค่าสังเกต } X \le x}{n}$$

เขียนกราฟของ $F^*(x)$ และ S(x) ได้เป็น



รูปที่ 2.6 กราฟของ $F^*(x)$ และ S(x) และค่าของ T

จากรูประยะห่างที่มากที่สุดตามแนวตั้งที่แยกระหว่าง 2 กราฟ คือ 0.290 ซึ่งเกิดขึ้นที่ x = 0.710 เนื่องจากว่าค่าของฟังก์ชันการแจกแจงเอ็มไพริคัลที่ x = 0.710 มีค่าเท่ากับ 1.00 นั่นคือ S(0.710) = 1.00 และค่าของฟังก์ชันการแจกแจงสมมุติฐานว่างเป็น $F^*(0.710) = 0.710$ ดังนั้น

$$\begin{aligned} T &= \sup_x \left| F^*_{(x)} - S(x) \right| \\ &= \left| F^*_{(0.710)} - S(0.710) \right| \\ &= 0.290 \end{aligned}$$

และ T = 0.290 < 0.409 ดังนั้น จึงไม่ปฏิเสธ H_0 และสรุปว่าข้อมูลนี้มาจากประชากรที่มีการแจกแจงเอกรูป ซึ่งค่าพีหาได้โดยใช้ตารางที่ 1 เมื่อ n = 10 และมีค่าพี = Pr(T > 0.290) > Pr(T > 0.323) = 0.20

ถ้าเปลี่ยนสมมุติฐานของการทดสอบเป็น

H_0 : $F(x) \geq F^*(x)$ สำหรับทุก ๆ x

H_1 : $F(x) < F^*(x)$ สำหรับบางค่าของ x

บริเวณวิกฤตที่ $\alpha = 0.05$ จะเป็นที่ T^+ มากกว่า 0.95 ควอนไทล์ สำหรับการทดสอบแบบหางเดียว ซึ่งจากตารางที่ 1 และ n = 10 พบว่าบริเวณวิกฤตคือ T > 0.369 ค่าของสถิติทดสอบหาได้จาก

$$\begin{aligned} T^+ &= \sup_x \left[F^*_{(x)} - S_{(x)}\right] \\ &= F^*_{(0.3289)} - S_{(0.3289)} \\ &= 0.3289 - 0.200 \\ &= 0.1289 \end{aligned}$$

เพราะว่า $T^+ = 0.1289 < 0.369$ จึงทำให้ไม่สามารถปฏิเสธ H_0 จึงให้ผลสรุปเหมือนกันกับการทดสอบสองหางข้างต้นด้วย ค่าพี > 0.10

ถ้าสมมุติฐานเพื่อการทดสอบเปลี่ยนเป็น

H_0 : $F(x) \leq F^*(x)$ สำหรับทุก ๆ ค่า x

H_1 : $F(x) > F^*(x)$ สำหรับบางค่าของ x

สถิติทดสอบจึงเป็น

$$\begin{aligned} T^- &= \sup_x \left[S(x) - F^*(x)\right] \\ &= S(0.710) - F^*(0.710) \\ &= 1.00 - 0.710 \\ &= 0.290 \end{aligned}$$

ด้วยค่า ค่าพี > 0.10 จึงยอมรับ H_0 และสรุปว่าข้อมูลดังกล่าวมาจากประชากรที่มีการแจกแจงเอกรูป

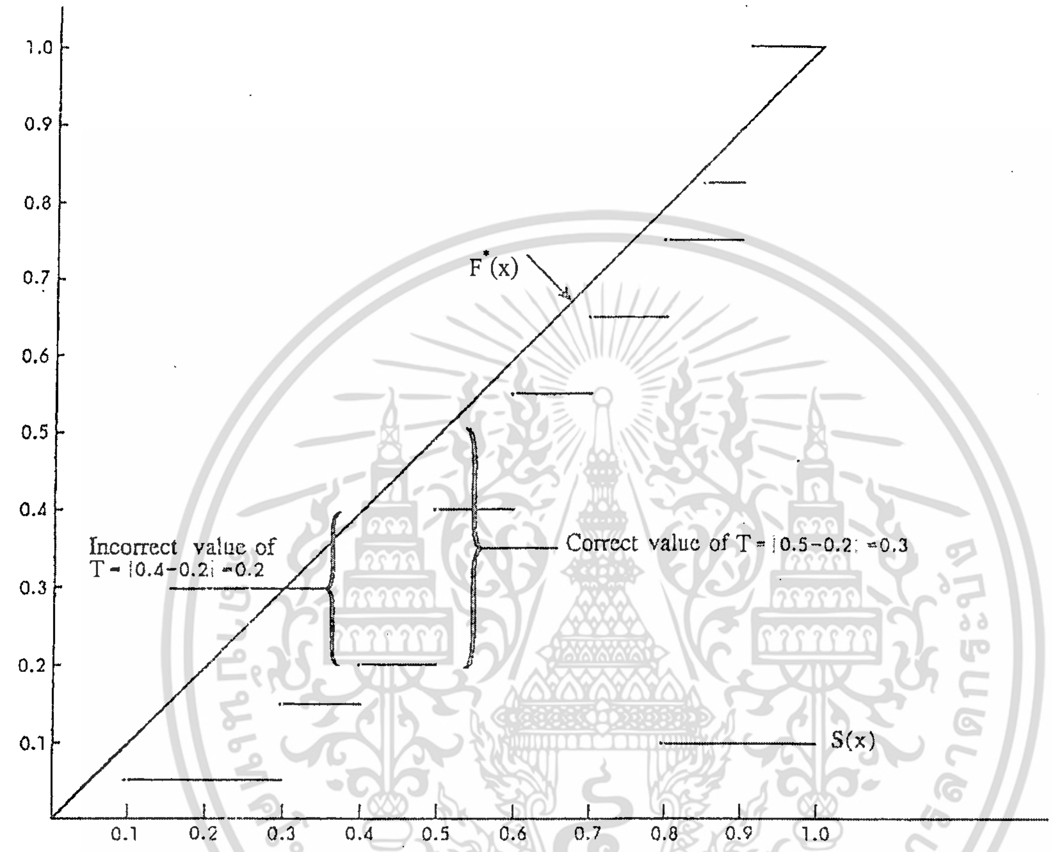
หมายเหตุ ค่าพีที่แท้จริงของการทดสอบแบบสองหางหาได้จาก

$$\text{ค่าพี} = 2(0.29)\sum_{j=0}^{7}\binom{10}{j}\left(0.71 - \frac{j}{10}\right)^{10-j}\left(0.29 + \frac{j}{10}\right)^{j-1}$$

$$= 0.307$$

สำหรับ $F^*(x)$ ที่ไม่ต่อเนื่อง ค่าพีที่ได้จากตารางที่ 1 จะไม่ถูกต้อง

การเปรียบเทียบค่า $F^*(x)$ และ $S(x)$ ด้วยวิธีการดังกล่าวข้างต้น บางครั้งจะพบว่าไม่เพียงพอที่จะหาค่า $T = \sup_x \left|F^*_{(x)} - S(x)\right|$ เนื่องจากระยะห่างมากที่สุด (ในแนวตั้ง) ระหว่าง $F^*(x)$ และ $S(x)$ ไม่ได้เกิดขึ้นที่ค่า x แต่เกิดขึ้นที่ค่าอื่น ให้พิจารณาจากกราฟของ $F^*(x)$ และ $S(x)$ ของตัวอย่างต่อไปนี้



รูปที่ 2.7 แสดงความแตกต่างของ $F^*(x)$ และ $S(x)$ ในด้านซ้ายและขวา

จากรูปกราฟข้างต้น ถ้าหาค่า $\left|F^*(x) - S(x)\right|$ เฉพาะที่จุดด้านซ้ายของเส้นตามแกนนอน (---) เราจะได้ว่า $T = |0.2 - 0.4| = 0.2$ ซึ่งไม่ใช่ค่าที่ถูกต้องของ T แต่ถ้าพิจารณากราฟให้ละเอียด จะพบว่าระยะห่างที่มากที่สุด (ในแนวตั้ง) จะเกิดขึ้นทางด้านขวาของเส้นตามแกนนอน (---) ที่จุด x = 0.4 ดังนั้นค่า T ที่ถูกต้อง คือ $|0.5 - 0.2| = 0.3$

ดังนั้นวิธีการหาค่าที่ถูกต้องของค่า T จากตัวเลขสามารถทำได้เช่นกันโดยหาค่า $\left|S(x_{i-1}) - F^*(x_i)\right|$ สำหรับทุกค่าของ i = 1, 2,…, r+1 เมื่อ r = จำนวนของค่า x ที่แตกต่างกัน

และกำหนดให้ $S(X_0) = 0$ ค่าสถิติทดสอบที่ถูกต้อง : T คือ

$$T = \sup_{1 \le i \le r} \left\{ \sup \left[\left| S(X_i) - F^*(X_i) \right|, \left| S(X_{i-1}) - F^*(X_i) \right| \right] \right\}$$

ให้ลองพิจารณาตัวอย่างต่อไปนี้

ตัวอย่าง จงทดสอบว่าข้อมูลต่อไปนี้ที่ถูกสุ่มมาด้วยขนาด 36 มาจากประชากรที่มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย 85 และส่วนเบี่ยงเบนมาตรฐาน = 15 หรือไม่

58	78	84	90	97	70	90	86	82
59	90	70	74	83	90	76	88	84
68	93	70	94	70	110	67	68	75
80	68	82	104	92	112	84	98	80

วิธีทำ H_0 : ข้อมูลนี้มาจากการแจกแจงปกติ ด้วยค่าเฉลี่ย = 85 และ $\sigma = 15$

H_1 : ข้อมูลนี้ไม่ได้มาจากการแจกแจงปกติ ด้วยค่าเฉลี่ย = 85 และ $\sigma = 15$

จะใช้การทดสอบของ Kolmogorov-Smirnov และคำนวณหาค่า $S(x)$, $F^*(x)$, T ดังต่อไปนี้

X_i	$S(x_i)$	$F^*(x_i)$	$\left\|S(x_i) - F^*(x_i)\right\|$	$\left\|S(x_{i-1}) - F^*(x_i)\right\|$
58	.0278	.0359	.0081	.0359
59	.0556	.0418	.0138	.0140
67	.0833	.1151	.0318	.0595
68	.1667	.1292	.0375	.0459
70	.2778	.1587	.1191	.0080
74	.3056	.2327	.0729	.0451
75	.3333	.2514	.0819	.0542
76	.3611	.2743	.0868	.0590
78	.3889	.3192	.0697	.0419
80	.4444	.3707	:0737	.0182
82	.5000	.4207	.0793	.0237

X_i	$S(x_i)$	$F^*(x_i)$	$\lvert S(x_i)-F^*(x_i)\rvert$	$\lvert S(x_{i-1})-F^*(x_1)\rvert$
83	.5278	.4483	.0795	.0517
84	.6111	.4721	.1390	.0557
86	.6339	.5279	.1110	.0832
88	.6667	.5793	.0874	.0596
90	.7778	.6293	.1485	.0374
92	.8056	.6808	.1248	.0970
93	.8333	.7019	.1314	.1037
94	.8611	.7257	.1354	.1076
97	.8889	.7881	.1008	.0730
98	.9167	.8078	.1089	.0811
104	.9444	.8980	.0464	.0187
110	.9722	.9525	.0197	.0081
112	1.0000	.9641	.0359	.0081

จะพบว่าค่าใหญ่ที่สุดของ $\lvert S(x_i)-F^*(x_i)\rvert = .1485$ แต่ไม่มีค่าใดของ $\lvert S(x_{i-1})-F^*(x_1)\rvert$ ที่มีค่ามากกว่า .1485 ดังนั้น T = .1485 จากตารางค่าวิกฤตที่ 1 ที่ N = 36 , α = .05 ได้ค่าวิกฤต = .221 ดังนั้นจึงยอมรับ H_0 นั่นคือ ตัวอย่างสุ่มชุดนี้ถูกสุ่มมาจากประชากรปกติด้วยค่าเฉลี่ย = 85 และส่วนเบี่ยงเบนมาตรฐาน = 15

หมายเหตุ จากตัวอย่างข้างต้นจะพบว่า ไม่มีประโยชน์ที่หาค่า $\lvert S(x_{i-1})-F^*(x_1)\rvert$ แต่กรณีเช่นนี้ไม่เกิดขึ้นเสมอไป ดังนั้นในการทดสอบการแจกแจงชนิดต่อเนื่อง แนะนำให้คำนวณค่าทั้งสอง คือ $\lvert S(x_i)-F^*(x_i)\rvert$ และ $\lvert S(x_{i-1})-F^*(x_1)\rvert$ แต่ถ้านำไปทดสอบการแจกแจงชนิดไม่ต่อเนื่อง ก็คำนวณหาค่า $\lvert S(x_i)-F^*(x_i)\rvert$ เพียงอย่างเดียว ซึ่งการทดสอบชนิดนี้จะเป็นการทดสอบแบบ Conservative (Noether and Slakter)

2. การทดสอบของ Lilliefors (The Lilliefors test)

Lilliefors (1967) ได้ปรับปรุงการทดสอบของ Kolmogorov-Smirnov ในกรณีที่ต้องการทดสอบเกี่ยวกับการแจกแจงปกติที่ไม่ได้ระบุค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน อาจเรียกได้ว่าเป็น “การทดสอบสำหรับการแจกแจงปกติ” (A Test For Normality) การทดสอบของ Lilliefors จะมีลักษณะคล้ายการทดสอบของ Kolmogorov-Smirnov เพียงแต่ตารางค่าวิกฤตจะต่างกัน

การทดสอบของ Lilliefors จะเหมือนกับการทดสอบของ Kolmogorov-Smirnov เกือบทุกประการ ยกเว้นการใช้คะแนนมาตรฐาน (Normolized value) แทนคะแนนดิบ กล่าวคือ จากข้อมูลตัวอย่างคำนวณหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐาน $(\overline{X},S)$ ด้วย

สูตร $\overline{X} = \dfrac{\sum_{i=1}^{N} X_i}{N}$

และ $S = \sqrt{\dfrac{1}{(N-1)}\sum_{i=1}^{N}(X_i-\overline{X})^2}$

และแปลงค่า X_i เป็น Z_i ด้วยสูตร $Z_i = \dfrac{X_i-\overline{X}}{S}$, $i = 1, \ldots, N$ การคำนวณหาสถิติทดสอบจะคำนวณจากค่า Z_i แทน X_i ซึ่งเป็นข้อมูลดิบ นั่นคือ หาค่า $S(Z_i) = \dfrac{k}{N}$ เมื่อ k = จำนวนข้อมูลค่า Z ที่มีค่าน้อยกว่าหรือเท่ากับ Z_i และคำนวณหาค่า $F_0(Z_i)$ จากความน่าจะเป็นสะสมของการแจกแจงปกติมาตรฐาน

การหาสถิติทดสอบคงเหมือนกับการทดสอบของ Kolmogorov-Smirnov แต่การหาค่าวิกฤตจะใช้ตารางค่าวิกฤตของ Lilliefors แทนดังตารางที่ 2 ซึ่งแยกได้ 3 ตาราง คือ

ตาราง 2_I ใช้ในกรณีที่ไม่ทราบค่าเฉลี่ยแต่ทราบค่าความแปรปรวน

ตาราง 2_{II} ใช้ในกรณีที่ทราบค่าเฉลี่ยแต่ไม่ทราบค่าความแปรปรวน

ตาราง 2_{III} ใช้ในกรณีที่ไม่ทราบค่าเฉลี่ยและความแปรปรวน

ตัวอย่าง ข้อมูลอายุของผู้เสียชีวิตเพศชายในสุสานแห่งหนึ่งใน Water Ross, Scotland จำนวน 117 คน ได้ถูกบันทึกไว้ และสุ่มมา 30 คน ได้ค่าอายุ (เรียงลำดับ) ดังนี้

11	13	14	22	29	30	41	41	52	55	56	59	65	65	66
74	74	75	77	81	82	82	82	82	83	85	85	87	87	88

มีเหตุผลเพียงพอหรือไม่ ที่จะกล่าวว่าอายุของผู้เสียชีวิตมีการแจกแจงแบบปกติ

วิธีทำ

H_0 : อายุผู้เสียชีวิตมีการแจกแจงแบบปกติ

H_1 : อายุผู้เสียชีวิตไม่มีการแจกแจงแบบปกติ

จากข้อมูลตัวอย่างขนาด 30 คำนวณหา $\overline{X}$ และ S ได้ค่า $\overline{X} = 61.43$ และ $S = 25.04$ แล้วเปลี่ยนค่าข้อมูลอายุ (X_i) ให้เป็นคะแนนมาตรฐาน Z_i ดังนี้

ที่ $X_1 = 11$, $Z_1 = \dfrac{11-61.43}{25.04} = -2.014$ และทำนองเดียวกันที่ i อื่น ๆ

คำนวณหาค่า $F_0(Z_i)$ จากตารางการแจกแจงปกติมาตรฐาน ได้พื้นที่ด้านซ้ายของ $Z_1 = -2.014$ $= .022 = P(Z < -2.014)$ และทำนองเดียวกันที่ i อื่น ๆ

ส่วนค่า $S(Z_i) = \frac{k}{N}$ เช่น $S(Z_1) = \frac{1}{30} = .033$ สามารถสรุปได้ดังตารางต่อไปนี้

X_i	Z_i	$F_0(Z_i)$	$S(Z_i)$	$\lvert F_0(Z_i) - S(Z_i) \rvert$	$\lvert F_0(Z_i) - S(Z_{i-1}) \rvert$
11	-2.014	.022	.033	0.011	.022
13	-1.934	.026	.067	0.044	.007
14	-1.894	.029	.100	0.071	.038
22	-1.575	.058	.133	0.075	.042
29	-1.295	.098	.167	0.069	.035
30	-1.255	.105	.200	0.095	.062
41*	-0.816	.207	.267	0.060	.007
52	-0.377	.353	.300	0.053	.086
55	-0.257	.399	.333	0.066	.099
56	-0.217	.414	.367	0.047	.081
59	-0.097	.461	.400	0.061	.094
65*	0.142	.556	.467	0.089	.156
66	0.183	.572	.500	0.072	.105
74*	0.502	.692	.567	0.125	.192*
75	0.542	.706	.600	0.106	.139
77	0.622	.733	.633	0.100	.133
81	0.781	.782	.667	0.115	.149
82*	0.821	.794	.800	0.006	.127
83	0.861	.805	.833	0.028	.005
85*	0.942	.827	.900	0.073	.006
87*	1.021	.846	.967	0.121	.054
88	1.061	.856	1.000	0.144	.111

* หมายถึง มีข้อมูลซ้ำ

จากคอลัมน์ที่ 5 และ 6 พบว่า ได้ค่า T = .192

และหาค่าวิกฤตจากตารางที่ 2_{III} (เนื่องจากไม่ทราบค่า μ และ σ) เมื่อกำหนด $\alpha = 0.01$ ที่ N = 30

ได้ค่าวิกฤต = .183

ดังนั้นจึงตกในอาณาเขตวิกฤต ปฏิเสธ H_0

นั่นคือ อายุคนเสียชีวิตเพศชายของสุสานนี้ ไม่มีการแจกแจงแบบปกติ

หมายเหตุ ยังคงสามารถใช้การเปรียบเทียบกราฟของ $F_0(Z_i)$ และ $S(Z_i)$ ได้เช่นกัน คือหา D = ระยะห่างที่มากที่สุด (ในแนวตั้ง) ระหว่าง $F_0(Z_i)$ และ $S(Z_i)$

3. การทดสอบของ Shapiro-Wilk

กำหนดให้
$$W = \frac{\left\{\sum_{i=1}^{k} a_{n+1-i}\left(x_{n+1-i} - x_{(i)}\right)\right\}^2}{\sum_{i=1}^{n}\left(x_{(i)} - \bar{x}\right)^2}$$

เมื่อ n แทน ขนาดตัวอย่าง

k แทน จำนวนเต็มที่เล็กที่สุดที่มากกว่าหรือเท่ากับ n/2

a_i แทน ค่าสัมประสิทธิ์ที่ได้จากการเปิดตารางที่ 3 ในภาคผนวก เมื่อ $n \leq 50$

$x_{(i)}$ แทน order sample

จะปฏิเสธสมมติฐานว่างที่ว่า "ประชากรมีการแจกแจงแบบปกติ" เมื่อค่าสถิติที่คำนวณได้มีค่าน้อยกว่าค่า W ที่ได้จากตารางในภาคผนวก ที่ขนาดตัวอย่าง n และระดับนัยสำคัญที่ต้องการ

ตัวสถิติ W เป็นตัวสถิติที่ใช้สัดส่วนของ $\hat{\sigma}^2/s^2$ เมื่อ $\hat{\sigma}^2$ เป็นค่าประมาณของ σ^2 ขั้นตอนในการหาตัวสถิติ W มีดังนี้

กำหนดให้ $m' = (m_1, m_2, ..., m_n)$ เป็นเวคเตอร์ค่าคาดหวังของ $x_1, x_2, ..., x_n$

$v = (v_{ij})$ เป็น Covariance Matrix ขนาด n x n

$y_1 \leq y_2 \leq ... \leq y_n$ เป็น order sample ขนาด n จาก N(0, 1)

ดังนั้น $E(x_i) = m_i$

$Cov(y_i, y_j) = v_{ij} \quad (i, j = 1, 2, \ldots, n)$

ให้ $x' = \left(x_{(1)}, x_{(2)}, ..., x_{(n)}\right)$ เป็นเวคเตอร์ของ order sample ของตัวอย่างขนาด n ต้องการทดสอบว่าตัวอย่างนี้สุ่มมาจากการแจกแจงแบบปกติ $N(\mu, \sigma^2)$

ถ้า $\{x_i\}$ มีการแจกแจงแบบปกติ

จะได้
$$\frac{x_i-\mu}{\sigma} = y_i$$
$$x_i = \mu+\sigma y_i \qquad (i = 1, 2, \ldots, n)$$
$$E(x_i) = \mu+\sigma E(y_i)$$
$$E(x) = 1+\sigma m \qquad (1 = \underset{\sim n\times 1}{1})$$

ดังนั้น
$$E(x) = p\theta \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

เมื่อ p เป็นเมตริกซ์ $(1,m_i)$ ซึ่งมีขนาด n x 2

θ' เป็นเวกเตอร์ (μ,σ)

และ
$$\mathrm{Var}(x_i) = \sigma^2\mathrm{Var}(y_i)$$
$$\mathrm{Var}(x) = \sigma^2 v$$

จาก (1) โดยทฤษฎี Generalized Least-Sqrares
$$\hat{\theta} = \left(p'v^{-1}p\right)^{-1}p'v^{-1}x \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$p'v^{-1}p$ สามารถกระจายในรูปของ
$$p'v^{-1}p = \begin{bmatrix} 1'v^{-1}1 & 1'v^{-1}m \\ 1'v^{-1}m & m'v^{-1}m \end{bmatrix}$$
$$\left(p'v^{-1}p\right)^{-1} = \frac{1}{\Delta}\begin{bmatrix} m'v^{-1}m & -1'v^{-1}m \\ -1'v^{-1}m & 1'v^{-1}1 \end{bmatrix}$$

เมื่อ
$$\Delta = 1'v^{-1}1m'v^{-1}m-1'v^{-1}m1'v^{-1}m$$
$$= 1'v^{-1}1m'v^{-1}m-\left(1'v^{-1}m\right)^2$$

ในทำนองเดียวกันสามารถกระจาย $p'v^{-1}x$ ในรูปของ
$$p'v^{-1}x = \begin{bmatrix} 1'v^{-1}x \\ m'v^{-1}x \end{bmatrix}$$

จาก (2) ดังนั้น
$$\hat{\theta} = \frac{1}{\Delta}\begin{bmatrix} m'v^{-1}m & -1'v^{-1}m \\ -1'v^{-1}m & 1'v^{-1}1 \end{bmatrix}\begin{bmatrix} 1'v^{-1}x \\ m'v^{-1}x \end{bmatrix}$$
$$\begin{bmatrix} \hat{\mu} \\ \hat{\sigma} \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} m'v^{-1}m1'v^{-1}x-1'v^{-1}mm'v^{-1}x \\ -1'v^{-1}m1'v^{-1}x+1'v^{-1}1m'v^{-1}x \end{bmatrix}$$
$$\hat{\mu} = \frac{m'v^{-1}\left(m1'+1m'\right)v^{-1}x}{1'v^{-1}1m'v^{-1}m-\left(1'v^{-1}m\right)^2}$$

$$\hat{\sigma} = \frac{1'v^{-1}(1m'-m1')v^{-1}x}{1'v^{-1}1m'v^{-1}m-(1'v^{-1}m)^2} \qquad (\because 1'v^{-1}m=0)$$

$$= \frac{m'v^{-1}x}{m'v^{-1}m}$$

เนื่องจาก $s^2 = \sum_{i=1}^{n}(y_i-\bar{y})^2$ เป็นตัวประมาณที่ไม่เอนเอียงของ $(n-1)\sigma^2$

ดังนั้นตัวสถิติ W ที่ใช้ทดสอบการแจกแจงแบบปกติ คือ

$$W = \frac{R^4\hat{\sigma}^2}{C^2S^2} = \frac{b^2}{S^2} = \frac{(a'x)^2}{S^2} = \frac{(\sum a_i x_i)^2}{\sum(x_i-x)^2}$$

เมื่อ $R^2 = m'v^{-1}m$

$C^2 = m'v^{-1}v^{-1}m$

$$a' = (a_1, a_2, \ldots, a_n) = \frac{m'v^{-1}}{(m'v^{-1}v^{-1}m)^{1/2}} = \frac{m'v^{-1}}{C}$$

$$b = \frac{R^2\hat{\sigma}}{C}$$

แต่เนื่องจาก $-a_i = a_{n+1-i}$ จาก Shapiro-Wilk (1965 : 593)

ดังนั้น

$$W = \frac{\left\{\sum_{i=1}^{k} a_{n+1-i}\left(x_{(n+1-i)} - x_{(i)}\right)\right\}^2}{\sum_{i=1}^{n}\left(x_{(i)} - \bar{x}\right)^2}$$

Sarhan และ Greenberg ได้คำนวณค่าของ a' เมื่อขนาดตัวอย่างเล็กกว่า 20 แต่สำหรับขนาดตัวอย่างที่ใหญ่กว่า 20 Shapiro และ Wilk ได้ประมาณค่า a' โดย

$$a' = \frac{m'v^{-1}}{(m'v^{-1}v^{-1}m)^{1/2}}$$

$$\therefore \quad a'a = 1$$

ให้ $a^* = m'v^{-1}$

ดังนั้น $c^2 = a^*a^{*\prime}$

และประมาณค่า a^* โดย $a_i^* = 2m_i$ เมื่อ $i = 2, 3, \ldots, n-1$

$$\text{และ} \quad \left(a_i^*\right)^2 = \left(a_n^*\right)^2 = \begin{cases} \dfrac{\tau\left\{\left(1/2\right)n\right\}}{\sqrt{2\tau\left\{\left(1/2\right)\left(n+1\right)\right\}}} & (n \leq 20) \\[2ex] \dfrac{\tau\left\{\left(1/2\right)\left(n+1\right)\right\}}{\sqrt{2\tau\left\{\left(1/2\right)n+1\right\}}} & (n > 20) \end{cases}$$

ตัวอย่าง การใช้สถิติของ Shapiro และ Wilk ในการทดสอบการแจกแจงแบบปกติของข้อมูลต่อไปนี้

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
-3	-5	4	6.5	0	2	1	3.5	7	10

H_0 : ประชากรมีการแจกแจงแบบปกติ

H_1 : ประชากรไม่มีการแจกแจงแบบปกติ

เรียงลำดับข้อมูลใหม่

i	$X_{(i)}$	$\left(X_{(i)}-\overline{X}\right)^2$
1	-5	57.76
2	-3	31.36
3	0	6.76
4	1	2.56
5	2	0.36
6	3.5	0.81
7	4	1.96
8	6.5	15.21
9	7	19.36
10	10	54.76
		190.90

$$s^2 = \sum_{i=1}^{10}\left(X_{(i)}-\overline{X}\right)^2$$

$$= 190.90$$

จากตารางที่ 3 ในภาคผนวก เมื่อ n = 10

$a_{10} = 0.5739, \quad a_9 = 0.3291, \quad a_8 = 0.2141, \quad a_7 = 0.1224$ และ $a_6 = 0.0399$

$$b = \sum_{i=1}^{k} a_{n+1-i}\left(x_{(n+1-i)} - x_{(i)}\right)$$

$$= (0.5739)(10-(-5)) + (0.3291)(7-(-3)) + (0.2141)(6.5-0) + (0.1224)(4-1) + (0.0399)(3.5-2)$$

$$= 13.7182$$

$$W = b^2 / s^2 = (13.7182)^2 / 190.90 = 0.9858$$

จากตารางที่ 3 พบว่า ถ้ากำหนดระดับนัยสำคัญ 0.05 ค่า W ที่ขนาดตัวอย่าง 10 มีค่าเท่ากับ 0.842 ดังนั้นยอมรับ H_0 เนื่องจากค่า W ที่คำนวณได้มีค่ามากกว่า 0.842

4. การทดสอบของ Anderson และ Darling

Anderson และ Darling (1954) ได้ปรับปรุงสถิติทดสอบของ Kolmogorov-Smirnov โดยให้ความสำคัญที่ส่วนหางของโค้งการแจกแจง สามารถนำไปทดสอบได้หลายการแจกแจง คือ การแจกแจงปกติ การแจกแจงลอกนอร์มอล และการแจกแจงเอ็กโปเนนเชียล โดยที่ค่าวิกฤตของแต่ละการทดสอบ ต้องคำนวณหาจากแต่ละการแจกแจงนั้น ๆ

กำหนดให้ สถิติทดสอบคือ

$$A^2 = -n - \left[\sum_{i=1}^{n}(2i-1)\left\{\ln z_i + \ln\left(1 - z_{n+1-i}\right)\right\} / n\right]$$

เมื่อ n แทน ขนาดตัวอย่าง

z_i แทน ความน่าจะเป็นสะสมของการแจกแจงแบบปกติมาตรฐาน

ซึ่ง $z_i = \phi\left(\left(x_{(i)} - \bar{x}\right)/s\right)$

X_i แทน order sample

$$\bar{x} = \sum_{i=1}^{n} x_i / n$$

$$s^2 = \sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 / n-1$$

ค่าวิกฤตของตัวสถิติ A^2 หาจากตารางของ M.A.Stephens

สถิติทดสอบ A^2 ดังกล่าวข้างต้นจะใช้ได้ดีกับสมมติฐานเบื้องต้นเกี่ยวกับการแจกแจงปกติ ที่อาจจะเป็นกรณีต่าง ๆ ของค่าพารามิเตอร์ของการแจกแจง ดังนี้

ก. ทราบค่าพารามิเตอร์ทั้งสอง คือ ทราบค่า μ และ σ^2

ข. ทราบเพียงค่าพารามิเตอร์ μ แต่พารามิเตอร์ σ^2 ถูกประมาณด้วยค่า s^2

ค. ทราบเพียงค่าพารามิเตอร์ σ^2 แต่พารามิเตอร์ μ ถูกประมาณด้วยค่า $\overline{X}$

สำหรับกรณีที่สมมติฐานเบื้องต้นเกี่ยวกับการแจกแจงปกติที่ไม่ทราบค่าพารามิเตอร์ทั้งสองซึ่งต้องประมาณจากค่าสถิติ $\overline{X}$ และ s^2 ของตัวอย่าง จำเป็นต้องปรับค่า (Modify) ค่าสถิติทดสอบเป็น $T^*\left(A^2\right) = A^2\left(1+\frac{4}{n}-\frac{25}{n^2}\right)$ ซึ่งสถิติทดสอบทั้งสองมีค่าวิกฤตด้านบน ดังตารางต่อไปนี้ (18)

ตาราง ค่าวิกฤตของสถิติทดสอบ A^2 หรือ $T^*\left(A^2\right)$ จากสมมติฐานเบื้องต้นจากการแจกแจงปกติ เมื่อกำหนดระดับนัยสำคัญ (ด้านบน)

สถิติทดสอบ A^2 หรือ $T^*\left(A^2\right)$	Upper Tail Probability			
	.10	.05	.025	.01
(a) Case 0 : Fully Specified $N\left(\mu,\sigma^2\right)$	1.933	2.492	3.070	3.857
(b) Case 1 : $N\left(\mu,\sigma^2\right)$, only σ^2 known	0.894	1.087	1.285	1.551
Case 2 : σ^2 estimated by s^2, μ known	1.743	2.308	2.898	3.702
Case 3 : μ and σ^2 estimated, $T^*\left(A^2\right)$	0.631	0.752	0.873	1.035

ตัวอย่าง การใช้ตัวสถิติของ Anderson-Darling ในการทดสอบการแจกแจงแบบปกติของข้อมูลต่อไปนี้

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8	y_9	y_{10}
-3	-5	4	6.5	0	2	1	3.5	7	10

H_0 : ประชากรมีการแจกแจงแบบปกติ

H_1 : ประชากรไม่มีการแจกแจงแบบปกติ

เรียงลำดับข้อมูลใหม่

i	X_i	z_i	$\ln z_i$	$\ln(1-z_{n+1-i})$	$\ln z_i + \ln(1-z_{n+1-i})$	2i-1
1	-5	0.0495	-3.0262	-2.9243	-5.9505	1
2	-3	0.1112	-2.1893	-1.7808	-3.9700	3
3	0	0.2877	-1.2458	-1.6210	-2.8668	5
4	1	0.3632	-1.0128	-0.9621	-1.9748	7
5	2	0.4483	-0.8023	-0.8658	-1.6681	9
6	3.5	0.5793	-0.5459	-0.5947	-1.1406	11
7	4	0.6179	-0.4814	-0.4513	-0.9327	13
8	6.5	0.8023	-0.2203	-0.3393	-0.5595	15
9	7	0.8315	-0.1845	-0.1178	-0.3024	17
10	10	0.9463	-0.0552	-0.0507	-0.1059	19

i	$(2i-1)(\ln z_i + \ln(1-z_{n+1-i}))$
1	-5.9505
2	-11.9100
3	-14.3340
4	-13.8236
5	-15.0129
6	-12.5466
7	-12.1251
8	-8.3925
9	-5.1428
10	-2.0121
	-101.2481

$$A^2 = -n - \left[\sum_{i=1}^{n} (2i-1) \left\{ \ln z_i + \ln\left(1 - z_{n+1-i}\right) \right\} \Big/ n \right]$$

$$= -10-(-101.2481)/10$$

$$= 0.1281$$

ทำการปรับให้อยู่ในรูป Modified form ได้ดังนี้

$$T^*\left(A^2\right) = A^2\left(1 + 4/n - 25/n^2\right)$$

$$= (0.1281)(1-4/10-25/100)$$

$$= 0.1473$$

ค่า $T^*\left(A^2\right)$ จากตารางเท่ากับ 0.752 ที่ระดับนัยสำคัญ 0.05 ดังนั้น ยอมรับ H_0 นั่นคือ ประชากรมีการแจกแจงปกติ

ตอนที่ 3 การแจกแจงที่เกี่ยวข้องกับงานวิจัยนี้

การแจกแจงที (Student's t – Distribution)

การแจกแจงที คิดค้นโดย William Sealy Gosset ตีพิมพ์ครั้งแรกในปี ค.ศ. 1908 โดยใช้นามปากกาว่า "Student" ดังนั้นการแจกแจงที จึงมีชื่อว่า Student t – Distribution หรือ Student's t – Distribution หรือบางทีเรียกว่า t – Distribution (Freund, 1992)

สมการของการแจกแจงที

ถ้า Y และ Z เป็นตัวแปรสุ่มที่อิสระต่อกัน Y เป็นตัวแปรสุ่มที่มีการแจกแจงแบบไคสแควร์ ที่ชั้นแห่งความเป็นอิสระเท่ากับ ν และ Z เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปกติ ตัวแปรสุ่ม T จะมีสมการดังนี้

$$T = \frac{Z}{\sqrt{Y/\nu}}$$

ซึ่งได้รับจากฟังก์ชันความหนาแน่นของความน่าจะเป็นของตัวแปร T ดังนี้

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)} \cdot \left(1 + \frac{t^2}{\nu}\right)^{\frac{-\nu+1}{2}} \qquad ; -\infty < t < \infty$$

เมื่อ $\Gamma(P)$ คือ ฟังก์ชันแกรมม่า มีค่า $= \int_0^\infty x^{P-1} e^{-x} dx \quad ; \quad P > 0$

เรียกตัวแปรสุ่ม T ว่ามีการแจกแจงที ที่ชั้นแห่งความเป็นอิสระเท่ากับ ν

คุณสมบัติของการแจกแจงที

1. มีลักษณะสมมาตร เมื่อเทียบกับแกน $t = 0$ จุดนี้จะเป็นค่าเฉลี่ย มัธยฐาน และฐานนิยม
2. ในกรณีที่ $\nu = 1$ การแจกแจงจะไม่มีค่าเฉลี่ย แต่ถ้า $\nu = 2, 3, \ldots$ ค่าเฉลี่ยจะมีค่าเท่ากับ 0
3. เมื่อ $\nu = 1, 2$ การแจกแจงจะไม่มีค่าความแปรปรวน แต่ถ้า $\nu = 3, 4, \ldots$ ค่าความแปรปรวนจะเท่ากับ $\nu/\nu - 2$
4. ถ้า ν มีค่ามาก ๆ การแจกแจงทีจะใกล้เคียงการแจกแจงแบบปกติ
5. กำลังสองของ t ที่ $\nu = k$ จะมีค่าเท่ากับ F ที่ $\nu = 1$ กับ k

การแจกแจงแบบยูนิฟอร์ม (The Uniform Distribution)

เราจะเรียกฟังก์ชันความหนาแน่นของความน่าจะเป็นที่กำหนดโดย

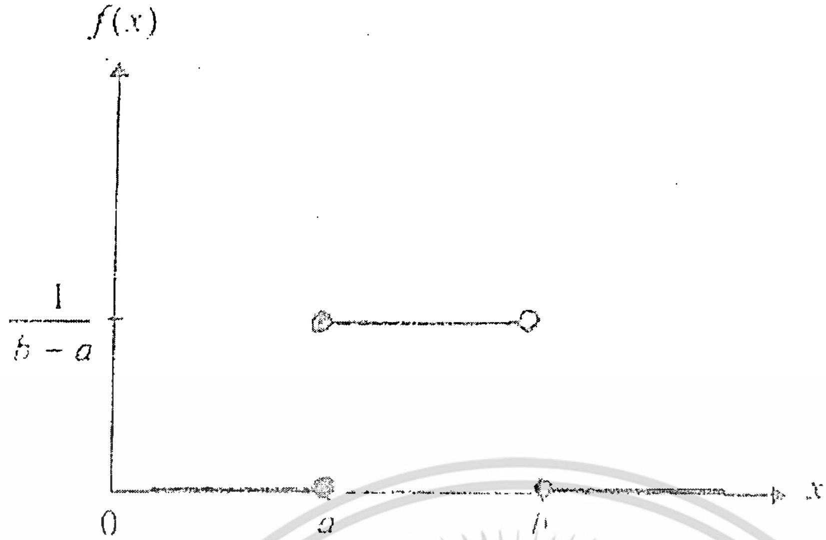
$$f(x) \quad = \begin{cases} \dfrac{1}{b-a} & \text{เมื่อ } a < x < b \\ 0 & \text{เมื่อ } x \text{ มีค่าอื่น ๆ} \end{cases}$$

โดยที่ a และ b เป็นค่าคงตัวที่ $a < b$ ว่า **การแจกแจงยูนิฟอร์ม** (uniform distribution) บนช่วง (a, b) และเรียกตัวแปรสุ่มที่มีการแจกแจงยูนิฟอร์มบนช่วง (a, b) ว่า **ตัวแปรสุ่มยูนิฟอร์ม** (uniform random variable) บนช่วง (a, b) โดยเราจะใช้สัญลักษณ์ "$X \sim U(a,b)$" แทนคำกล่าวที่ว่า "X เป็นตัวแปรสุ่มยูนิฟอร์มบนช่วง (a, b)"

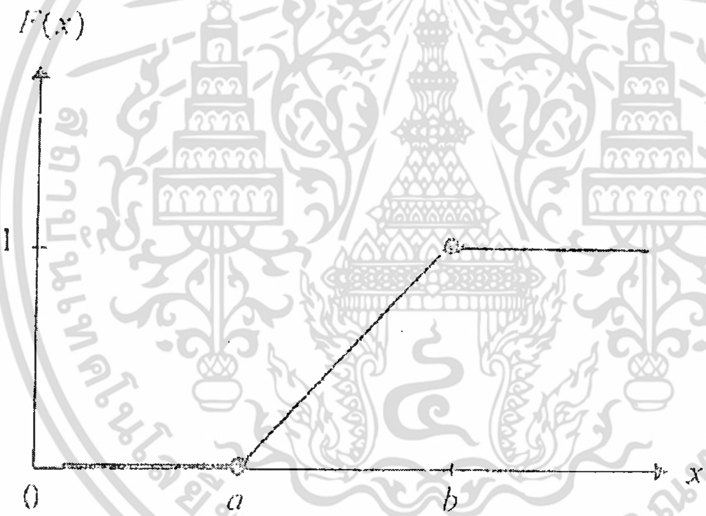
ข้อสังเกต 1. ถ้า $X \sim U(a,b)$ จะได้ว่า ฟังก์ชันการแจกแจง F ของ X จะกำหนดโดย

$$F(x) \quad = \begin{cases} 0 & \text{เมื่อ } x < a \\ \dfrac{x-a}{b-a} & \text{เมื่อ } a \le x \le b \\ 0 & \text{เมื่อ } x > b \end{cases}$$

2. ถ้า $X \sim U(a,b)$ จะได้ว่า ฟังก์ชันความน่าจะเป็น f และฟังก์ชันการแจกแจง F ของ X มีกราฟดังนี้



รูปที่ 2.8 ฟังก์ชันความหนาแน่นของความน่าจะเป็นของการแจกแจงแบบยูนิฟอร์มในช่วง (a, b)



รูปที่ 2.9 ฟังก์ชันการแจกแจงของการแจกแจงยูนิฟอร์มในช่วง (a, b)

ตัวอย่าง ให้ $X \sim U(0,10)$ จงหา

ก. $P(X<3)$

ข. $P(X>6)$

ค. $P(3<X<8)$

วิธีทำ เพราะว่า $X \sim U(0,10)$ ดังนั้นฟังก์ชันความน่าจะเป็น f ของ X กำหนดโดย

$$f(x) = \begin{cases} \frac{1}{10} & \text{เมื่อ } 0<x<10 \\ 0 & \text{เมื่อ } x \text{ มีค่าอื่น ๆ} \end{cases}$$

เพราะฉะนั้นจะได้ว่า

ก) $P(X<3) = \int_{-\infty}^{3} f(x)dx$

$= \int_{0}^{3} \frac{1}{10}dx$

$= \frac{3}{10}$

ข) $P(X>6) = 1-P(X\leq 6)$

$= 1-\int_{-\infty}^{6} f(x)dx$

$= 1-\int_{0}^{6} \frac{1}{10}dx$

$= 1-\frac{6}{10}$

$= \frac{2}{5}$

ค) $P(3<X<8) = \int_{3}^{8} f(x)dx$

$= \int_{3}^{8} \frac{1}{10}dx$

$= \frac{1}{2}$

ตัวอย่าง รถประจำทางจะมาจอดที่ป้ายในเวลา 7 :15 และ 7:30 ตามลำดับ ผู้โดยสารคนที่หนึ่งจะมาถึงป้ายรถประจำทางด้วยการแจกแจงยูนิฟอร์มบนช่วง (7, 7:30) จงหาความน่าจะเป็นที่

ก. ผู้โดยสารจะคอยรถประจำทางไม่เกิน 5 นาที

ข. ผู้โดยสารจะคอยรถประจำทางเกิน 10 นาที

วิธีทำ ให้ X เป็นตัวแปรสุ่มที่มีค่าเป็นเวลาที่ผู้โดยสาร จะมาถึงป้ายรถประจำทาง ดังนั้น $X \sim U(7,7:30)$ และ

ก) ความน่าจะเป็นที่ผู้โดยสารจะคอยรถประจำทางไม่เกิน 5 นาที

$= P(7:10<X<7:15)+P(7:25<X<7:30)$

$= \int_{7:10}^{7:15} \frac{1}{30}dx + \int_{7:25}^{7:30} \frac{1}{30}dx$

$$= \frac{5}{30} + \frac{5}{30}$$

$$= \frac{1}{3}$$

ข. ความน่าจะเป็นที่ผู้โดยสารจะคอยรถประจำทางเกิน 10 นาที

$$= P(7{:}00 < x < 7{:}05) + P(7{:}15 < x < 7{:}20)$$

$$= \int_{7:00}^{7:05} \frac{1}{30} dx + \int_{7:15}^{7:20} \frac{1}{30} dx$$

$$= \frac{5}{30} + \frac{5}{30}$$

$$= \frac{1}{3}$$

ทฤษฎี กำหนดให้ $X \sim U(a,b)$ จะได้ว่า

1. $E(X) = \dfrac{a+b}{2}$
2. $Var(X) = \dfrac{(a+b)^2}{12}$

พิสูจน์

1) $$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

$$= \int_{a}^{b} \frac{x}{b-a} dx$$

$$= \frac{a+b}{2}$$

2) เนื่องจาก $$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$= \int_{a}^{b} \frac{x^2}{b-a} dx$$

$$= \frac{a^2+ab+b^2}{3}$$

และ $E(X) = \dfrac{a+b}{2}$ จึงได้ว่า

$$Var(X) = E(X^2) - E^2(X)$$

$$= \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{(a-b)^2}{12}$$

ตัวอย่าง ให้ $X \sim U(0,10)$ จงหา $E(X)$

วิธีทำ จากทฤษฎีจะได้ว่า

$$E(X) = \frac{10+0}{2} = 5$$

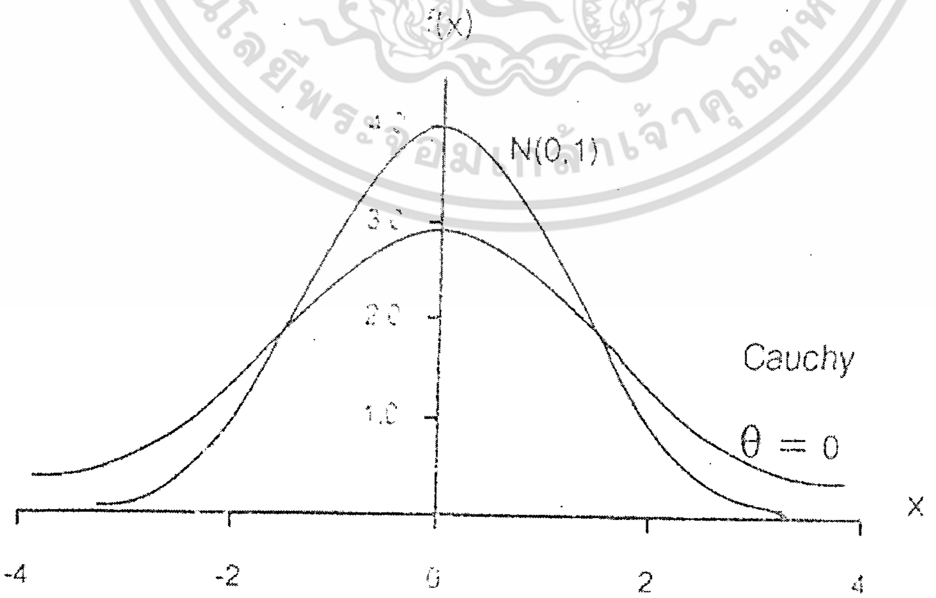
การแจกแจงแบบโคชี (The Cauchy Distribution)

นิยาม ถ้า X เป็นตัวแปรสุ่ม X จะมีการแจกแจงโคชี ถ้า X มีฟังก์ชันความหนาแน่นของความน่าจะเป็นดังนี้

$$f(x) = \frac{1}{\pi[1+(x-\theta)]^2} \qquad -\infty < x < \infty \;;\; -\infty < \theta < \infty$$

$$= 0 \qquad \text{นอกช่วงดังกล่าว}$$

การแจกแจงโคชีจะเหมือนกับการแจกแจงปกติ เส้นโค้งมีลักษณะเป็นรูประฆังคว่ำ และสมมาตรกับแกนตั้งที่ลากผ่าน θ แต่เส้นโค้งจะแบนกว่าโค้งปกติ ดังรูป



รูปที่ 2.10 ฟังก์ชันความหนาแน่นของความน่าจะเป็นของการแจกแจงแบบโคชี และแบบปกติมาตรฐาน

ถ้า X มีการแจกแจงโคชี แล้วฟังก์ชันก่อกำเนิดโมเมนต์ และทุกโมเมนต์ของ X รอบศูนย์ $\left(E\left(X^k\right)\right)$ จะไม่มีจริง (does not exist)

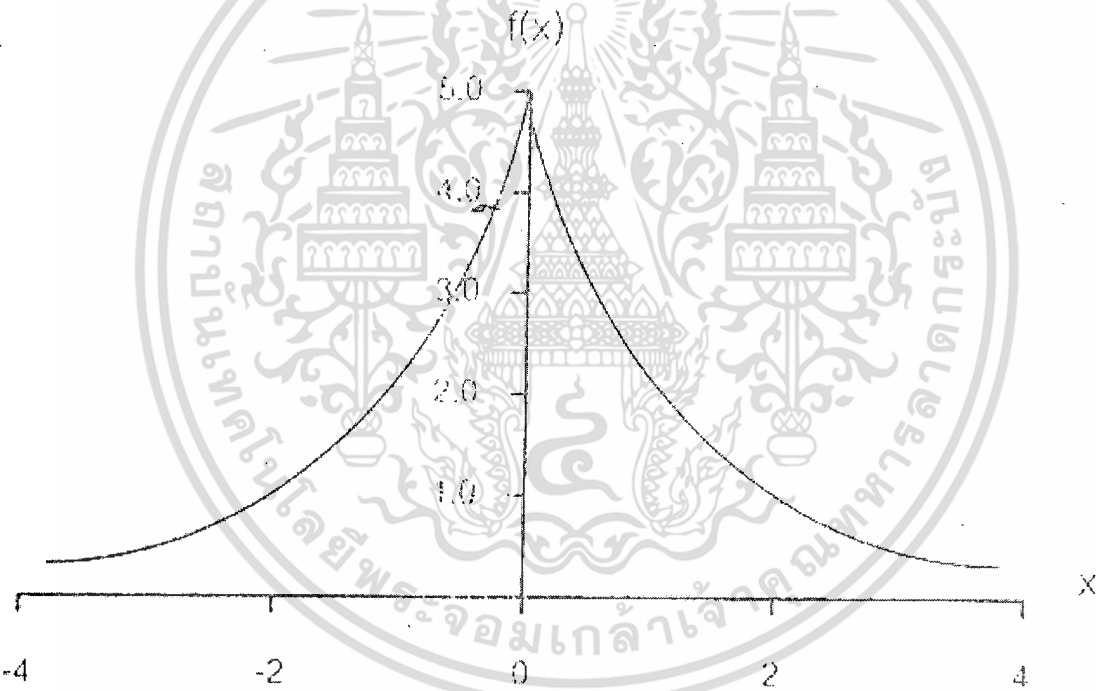
การแจกแจงแบบลาปลาซ (The Laplace Distribution)

ถ้า X เป็นตัวแปรสุ่ม X จะมีการแจกแจงลาปลาซ หรือการแจกแจงเอ็กซ์โพเนนเชียลสองด้าน (double exponential) ถ้า X มีฟังก์ชันความหนาแน่นของความน่าจะเป็นดังนี้

$$f(x) = \frac{1}{2\beta}e^{-\frac{|x-\theta|}{\beta}} \qquad -\infty < x < \infty \;;\; -\infty < \theta < \infty, \beta > 0$$

$$= 0 \qquad \text{นอกช่วงดังกล่าว}$$

กราฟของฟังก์ชัน สำหรับ $\theta = 0$ และ $\beta = 1$ ดังรูป



รูปที่ 2.11 ฟังก์ชันความหนาแน่นของความน่าจะเป็นของการแจกแจงแบบลาปลาซ ที่ $\theta = 0$ และ $\beta = 1$

ถ้า X มีการแจกแจงลาปลาซ ที่มีพารามิเตอร์ θ และ β แล้วจะได้

$$E(X) = \theta$$

$$V(X) = 2\beta^2$$

และ $$M_X(t) = \frac{e^{\theta t}}{1-\beta^2 t^2} \qquad ,|t| < \frac{1}{\beta}$$

ตอนที่ 4 งานวิจัยที่เกี่ยวข้อง

สำหรับงานวิจัยที่เกี่ยวข้องนั้น ได้มีนักสถิติจำนวนมากทำการศึกษาเกี่ยวกับอำนาจการทดสอบของสถิติทดสอบที่ใช้ในการทดสอบการแจกแจงแบบปกติ ซึ่งในส่วนนี้จะเสนอเฉพาะบางผลงานวิจัยที่สำคัญและเกี่ยวข้องกับงานวิจัยนี้เท่านั้น

Shapiro และคณะ (1968) เป็นผู้ริเริ่มในการศึกษาอำนาจการทดสอบของสถิติทดสอบต่าง ๆ ในการทดสอบการแจกแจงแบบปกติ โดยทำการศึกษาสถิติทดสอบ 9 แบบ คือ Shapiro-Wilk Statistic (W), $\sqrt{b_1}$, b_2, Kolmogorov-Smirnov Test (K), Cramer-von Mises $\left(W^2\right)$, Anderson-Darling $\left(A^2\right)$, Durbin (D), Chi-square Test $\left(X^2\right)$, และ Studentized Range Test (U) ภายใต้การแจกแจง 12 การแจกแจง ซึ่งมีค่าพารามิเตอร์ที่แตกต่างกัน รวมเป็น 45 การแจกแจง ได้ผลสรุปดังนี้

1. Shapiro-Wilk Statistic ใช้ได้ดีในการทดสอบทั่วไป
2. การทดสอบโดยใช้ Empirical Distribution Function ได้อำนาจการทดสอบต่ำ
3. Studentized Range Test (U) มีอำนาจการทดสอบสูงเมื่อประชากรมีการแจกแจงแบบ Symmetric Short-Tailed และมีอำนาจการทดสอบต่ำเมื่อประชากรมีการแจกแจงแบบ Asymmetric Short-Tailed และ Asymmetric Long-Tailed
4. $\sqrt{b_1}$ และ b_2 ใช้ในการทดสอบได้ดี แต่มีอำนาจการทดสอบต่ำกว่า W

M.A. Stephens (1974) ได้แสดงให้เห็นว่า ผลการศึกษาของ Shapiro และคณะ ใช้ค่าวิกฤตสำหรับสถิติทดสอบที่ใช้ Empirical Distribution Function ในการทดสอบการแจกแจงแบบปกติของประชากรที่ไม่ถูกต้องนัก ทั้งนี้เพราะค่าวิกฤตดังกล่าวคำนวณมาจากการสมมติว่าทราบค่าเฉลี่ยและความแปรปรวนของประชากร ดังนั้น Stephens จึงทำการคำนวณค่าวิกฤตของตัวสถิติเหล่านี้ขึ้นมาใหม่ โดยสมมติว่าไม่ทราบค่าเฉลี่ยและความแปรปรวนของประชากร

สมพิศ โชติวิทยะธารากร (2531) ได้ทำการเปรียบเทียบอำนาจการทดสอบของตัวสถิติที่ใช้ในการทดสอบการแจกแจงแบบปกติ 5 ตัว คือ Chi-Sqrare Test $\left(X^2\right)$ Studentized Range Test (U) Shapiro-Wilk Statistic (W) Probability Plot Correlation Coefficient Test (r) และ Hannu oja Statistic (T_1 และ T_2) ภายใต้การแจกแจง 2 ลักษณะที่สำคัญ คือ การแจกแจงแบบปกติ และการแจกแจงแบบไม่ปกติ ซึ่งพบว่าในด้านการนำไปใช้ประโยชน์แล้วควรเลือกใช้ตัวสถิติ W เนื่องจากให้อำนาจการทดสอบสูงเป็นส่วนใหญ่ และสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ดี

Edith Seirr (2004) (21) ได้ทำการศึกษาถึงสถิติทดสอบต่าง ๆ ที่ใช้ทดสอบการแจกแจงปกติ รวมทั้งสิ้น 10 สถิติทดสอบ โดยทำการจำลองข้อมูลให้มีขนาดตั้งแต่ 20 --> 100 ด้วยจำนวนซ้ำ 10,000 รอบ จากการแจกแจงแบบต่าง ๆ ที่มีลักษณะต่าง ๆ เช่น Bimodal, หางสั้น, โด่งกว่าปกติเล็กน้อย, โด่งมาก, มีความเบ้ ศึกษาถึงค่า empirical alpha และ power ที่ได้จากสถิติทดสอบแต่ละแบบ พบว่า สถิติทดสอบที่ใช้หลักการของ Regression test คือ D’Agostino (1972), Shapiro, Royston (ซึ่งคือ Shapiro corrected), Chen and Shapiro และ Zhang จะดีที่สุด คือมี power สูงที่สุด ในเกือบทุกแบบของการแจกแจง แต่ถ้ามีวัตถุประสงค์

เพื่อดูว่าการแจกแจงมีลักษณะสมมาตรที่มีความโด่งสูงหรือไม่ ควรใช้สถิติทดสอบแบบที่ใช้หลักของ Skewness และ Kurtosis Test คือ D'Agostino (1990) และ G-kurtosis ซึ่งประกอบด้วย The first statistic G_w^2 และ The second statistic G_w^{2*} ที่ให้ค่า power สูงที่สุด และยังแนะนำว่าโปรแกรมสำเร็จรูปควรจัดให้มีสถิติทดสอบเหล่านี้

กมล บุษบา (2547) ได้ทำการศึกษาถึงการคำนวณค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยในการทดลองแบบแฟกทอเรียล กรณีตัวแบบอิทธิพลผสมของโปรแกรม SPSS ซึ่งพบว่า ได้ค่าไม่ถูกต้อง ในขณะที่โปรแกรม SAS ให้ผลลัพธ์ที่ถูกต้อง และกมล ยังได้ศึกษาถึงการรายงานค่าพี (p-value) ด้วยสถิติทดสอบไค-สแควร์ จากตารางไขว้ของโปรแกรม SPSS ที่พบว่า เป็นค่าความน่าจะเป็นแบบด้านเดียว ไม่ใช่แบบสองด้านตามที่โปรแกรมรายงาน ดังนั้นผู้นำผลไปใช้ไม่จำเป็นต้องนำไปหารด้วยสอง เนื่องจากการทดสอบด้วยไค-สแควร์ เป็นการทดสอบแบบด้านเดียว

Reinhard Bergmann, John Ludbrook and Wilk P.J.M. Shooren (2000) ได้ทำการศึกษาถึงผลลัพธ์ของการใช้สถิติทดสอบ Wilcoxon-Mann Whitney จากโปรแกรมสำเร็จรูป 11 แบบ เช่น SPSS 8.0, StatXact 4.0, SigmaStat 2.03, S-Plus 2000, SAS 6.12 เป็นต้น พบว่า ผลลัพธ์ที่ได้จะแตกต่างกันในหลายด้าน เช่น การปรับ ties, การปรับค่าต่อเนื่อง, การปรับค่ากรณีตัวอย่างมีขนาดใหญ่ รวมทั้งการทำให้เข้าใจผิด หรือ Inadequate description of Algorithms

Leo Knüsel ได้ทำการศึกษาถึงความถูกต้องของค่าความน่าจะเป็นของการแจกแจงทวินาม พัวซอง ไฮเปอร์จีออเมตริก แกมมา และ Inverse Beta จากโปรแกรม Excel 2003 พบว่าได้ค่าความน่าจะเป็นของการแจกแจงทวินามและพัวซองจะถูกต้อง เมื่อค่าตัวแปรสุ่มมีค่าในช่วงกลาง ๆ ของการแจกแจงเท่านั้น แต่เมื่อตัวแปรสุ่มมีค่าในตอนต้นของการแจกแจง (extreme lower tail) Excel 2003 จะปัดค่าความน่าจะเป็นเหล่านั้นเป็น 0 ในขณะที่ Excel 97 ให้ค่าที่ถูกต้อง แสดงให้เห็นว่าการสร้าง Algorithm ใน Excel 2003 ก็ยังคงมีข้อบกพร่องอยู่ แม้ว่าจะมีการปรับปรุงจาก Excel 97 แล้วก็ตาม

บทที่ 3
วิธีดำเนินการวิจัย .

3.1 ขั้นตอนการดำเนินการวิจัย

ขั้นตอนในการวิจัยแบ่งออกเป็น 4 ขั้นตอน ดังนี้

1. สร้างการแจกแจงของประชากรที่มีลักษณะสมมาตร ตามลักษณะที่กำหนดในสถานการณ์ ต่าง ๆ ดังนี้

1.1 กำหนดประชากรให้มีการแจกแจงที่สมมาตร ดังต่อไปนี้

- การแจกแจงปกติโดยสนใจศึกษาเมื่อพารามิเตอร์ $\mu = 100$ และ $\sigma = 10, 50, 100$
- การแจกแจงสมมาตรที่มีหางยาวกว่าปกติ (Long tailed distribution) คือ การแจกแจงแบบที d.f = 1 และ 5
- การแจกแจงสมมาตรแต่โด่งกว่าปกติมาก (Distributions with high kurtosis) เช่น การแจกแจงCauchy ด้วยพารามิเตอร์ 0, 2 คือ C(0, 2) และ C(0, 0.5) และการแจกแจง Laplace คือ L(0, 0.5)
- การแจกแจงสมมาตรแต่โด่งกว่าปกติเล็กน้อย (Distribution with kurtosis slightly higher than the Normal) เช่น Logistic คือ L(0, 0.5) และ L(0, 1)
- การแจกแจงสมมาตรที่มีหางสั้น (Short tailed distribution) เช่น Uniform แบบ U(0 – 1) และ U (10 – 50)
- ข้อมูลดังกล่าวข้างต้นได้จากการจำลองด้วยโปรแกรม MINITAB Version 14.0 ด้วยเมนู Calc-Random Data ซึ่งได้จากงานวิจัยของกุศยา (15) ที่พบว่าการจำลองข้อมูลให้มีการแจกแจงแบบต่าง ๆ โปรแกรม MINITAB มีประสิทธิภาพใกล้เคียงกับ SAS

1.2 กำหนดขนาดตัวอย่างเป็น 10, 20, 30, 50, 100

1.3 จำนวนทำซ้ำ = 500 รอบในแต่ละสถานการณ์

2. คำนวณค่าสถิติทดสอบและค่าพี (p – value)

โดยนำข้อมูลในข้อ 1.1 ไปวิเคราะห์ด้วยเมนูใน SPSS เพื่อทดสอบการแจกแจงปกติ คือเมนู Explore และเมนู Nonparametric Test และนำข้อมูลชุดเดิมนั้นวิเคราะห์ด้วยโปรแกรม MINITAB ด้วยเมนู Basic Statistics ในแต่ละสถานการณ์ในข้อ 1.2 และ 1.3

3. คำนวณค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1

เมื่อใช้ข้อมูลจากประชากรที่มีการแจกแจงปกติ ทดสอบด้วยสถิติทดสอบด้วยโปรแกรม SPSS และ MINITAB เป็นจำนวน 500 ชุดตัวอย่าง ในแต่ละสถานการณ์

หาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 โดยการนับจำนวนครั้งของการปฏิเสธสมมติฐานเบื้องต้น (H_0 : ตัวแปรสุ่มมาจากการแจกแจงปกติ) เมื่อกำหนด α = 0.5 หรือ .10 คือนับว่ามีจำนวนกี่ครั้งที่ได้ค่าพีน้อยกว่า .05 และน้อยกว่า .10 แล้วหารด้วย 500 และเทียบค่ากับความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ที่กำหนดไว้ = .05 และ .10 ว่าสามารถควบคุมค่าได้หรือไม่ โดยใช้เกณฑ์ของ Cochran และ Bradley ดังนี้

เกณฑ์ของ Cochran จะมีเกณฑ์การพิจารณาว่าสามารถควบคุมความน่าจะเป็นได้ เมื่อ

ที่ระดับนัยสำคัญ .10 ถ้าค่าที่ได้มีค่าในช่วง (.08 - .12)

ที่ระดับนัยสำคัญ .05 ถ้าค่าที่ได้มีค่าในช่วง (.04 - .06)

เกณฑ์ของ Bradley จะมีเกณฑ์การพิจารณาว่าสามารถควบคุมความน่าจะเป็นได้ เมื่อ

ที่ระดับนัยสำคัญ .10 ถ้าค่าที่ได้มีค่าในช่วง (.05 - .15)

ที่ระดับนัยสำคัญ .05 ถ้าค่าที่ได้มีค่าในช่วง (.025 - .075)

4. คำนวณค่าอำนาจการทดสอบ (Power of the test)

เมื่อใช้ข้อมูลจากประชากรที่มีการแจกแจงอื่น ๆ ที่ไม่ใช่การแจกแจงปกติ ทดสอบด้วยสถิติทดสอบจากโปรแกรม SPSS และ MINITAB เป็นจำนวน 500 ชุดตัวอย่าง ในแต่ละสถานการณ์ หาค่าอำนาจการทดสอบด้วยการนับจำนวนครั้งของการปฏิเสธสมมติฐานเบื้องต้น (H_0 : ตัวแปรสุ่มมาจากการแจกแจงปกติ) คือนับว่ามีจำนวนกี่ครั้งที่ได้ค่าพีน้อยกว่า .05 และ น้อยกว่า .10 แล้วหารด้วย 500 และจะเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบในแต่ละสถานการณ์

3.2 วิธีดำเนินการวิจัย

การวิจัยนี้ได้สร้างและจำลองการทดลอง โดยใช้เครื่องคอมพิวเตอร์ช่วยในการดำเนินการ โดยมีวิธีการดำเนินงานตามลำดับดังนี้

1. สร้างข้อมูลจากประชากรที่มีการแจกแจงต่าง ๆ โดยใช้โปรแกรมสำเร็จรูป MINITAB 14 ด้วยการใช้เมนูตามขั้นตอนต่อไปนี้

Calc

Random data

เลือกการแจกแจงตามที่ต้องการ เช่น การแจกแจงปกติ ซึ่งต้องกำหนด

ขนาดตัวอย่าง จาก .... rows of data

จำนวนชุดตัวอย่าง จาก .... Store in column

ค่าพารามิเตอร์ คือ ค่าเฉลี่ย และส่วนเบี่ยงเบนมาตรฐาน เช่น (0, 1)

(สำหรับการแจกแจงอื่น ๆ ก็ทำในทำนองเดียวกันนี้)

ตัวอย่างเช่น ต้องการสร้างข้อมูลจากประชากรปกติที่มีค่าเฉลี่ย = 0 และส่วนเบี่ยงเบนมาตรฐาน = 1 ด้วยขนาดตัวอย่าง = 10 โดยมี 3 ชุดตัวอย่าง จะได้ชุดตัวอย่างทั้ง 3 ดังนี้

	C1	C2	C3
1	-0.79781	0.09723	0.93731
2	0.68059	1.57570	0.54652
3	0.12719	0.93158	-0.91581
4	0.09112	0.93090	1.39447
5	-0.16258	1.60160	-0.66808
6	0.76615	0.35147	-0.36402
7	-0.02019	0.07582	-1.75371
8	-1.31629	0.04552	-0.57979
9	-1.18545	-0.99917	0.82800
10	-0.66199	0.86277	1.05859

เมื่อ C_1, C_2 และ C_3 คือชุดตัวอย่างที่ 1 ถึง 3 ในงานวิจัยนี้จะสร้างข้อมูล 500 ชุด นั่นคือ มี C_1 ถึง C_{500} แต่ละชุดตัวอย่างจะได้ค่าข้อมูลที่ต่างกัน

2. นำข้อมูลที่ได้ในขั้นที่ 1 ไปวิเคราะห์ด้วยโปรแกรม SPSS ด้วยสถิติทดสอบ 3 แบบ โดยใช้เมนูดังนี้

โดยใช้เมนูดังนี้	Analyze	และ	Analyze
	Descriptive Statistics		Nonparametric Test
	Explore		1 – Sample K-S (ซึ่งจะมีสถิติ K-S)
	Plot		
	Normality Plots with test (ซึ่งจะมีสถิติ Lilliefors และ S-W)		

ซึ่งจะได้ผลลัพธ์เป็นค่าสถิติทดสอบ และค่าพี (p-value)

3. นำข้อมูลที่ได้จากขั้นที่ 1 (คือข้อมูลชุดเดิมที่ใช้ในขั้นที่ 2) ไปวิเคราะห์ด้วยโปรแกรม MINITAB ด้วยสถิติทดสอบ 3 แบบ โดยใช้เมนู Basic Statistics1 ซึ่งจะมีสถิติ AD, RJ และ K-S โดยโปรแกรมได้กำหนดให้ Default อยู่ที่สถิติ AD

4. คำนวณหาค่า empirical alpha หรือความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ที่แท้จริง จากข้อมูลที่มาจากประชากรปกติ ที่ขนาดตัวอย่างหนึ่ง โดยจะนับจำนวนชุดตัวอย่างที่ทำการปฏิเสธสมมติฐานเบื้องต้น (คือได้ค่าพีน้อยกว่า .05 หรือน้อยกว่า .10) แล้วหารด้วย 500 แล้วนำ

ค่าที่ได้เทียบกับค่า Alpha ที่กำหนดเบื้องต้น (.05 หรือ .10) ว่ามีค่าที่อยู่ในขอบเขตที่ถือว่าควบคุมได้หรือไม่ ตามเกณฑ์ของ Cochran และ Bradley ดังกล่าวในตอนต้น

5. คำนวณหาค่าอำนาจการทดสอบ หรือความสามารถในการปฏิเสธสมมติฐานเบื้องต้นที่เป็นเท็จ จากข้อมูลที่มาจากการแจกแจงแบบอื่น ๆ ที่ไม่ใช่แบบปกติ ที่ขนาดตัวอย่างหนึ่ง โดยจะนับจำนวนชุดตัวอย่างของการปฏิเสธสมมติฐานเบื้องต้น (คือมีค่าพีน้อยกว่า .05 หรือ .10) แล้วหารด้วย 500

6. เปรียบเทียบค่าอำนาจการทดสอบจากสถิติทดสอบทั้งหมดจากโปรแกรม SPSS และ MINITAB ในกรณีที่สถิติทดสอบนั้น ๆ สามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้

บทที่ 4
ผลการวิจัย

จะเสนอผลการวิจัยโดยแยกผลจากโปรแกรม SPSS และผลจากโปรแกรม MINITAB และเปรียบเทียบผลจากโปรแกรม SPSS + MINITAB โดยแยกเป็น 2 ประเด็น คือ ความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 และค่าอำนาจการทดสอบ ดังนี้

4.1 **การเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1** เพื่อหาผลสรุปว่าสถิติทดสอบใดสามารถควบคุมความคลาดเคลื่อนชนิดที่ 1 ได้ ตามเกณฑ์ของ Cochran และ Bradley

4.1.1 ผลการเปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จากการทดลองใช้โปรแกรม **SPSS** เพื่อทดสอบการแจกแจงปกติ จะนำเสนอด้วยตาราง 4.1 ดังนี้

ตารางที่ 4.1 ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จากสถิติทดสอบต่าง ๆ ที่ใช้ทดสอบการแจกแจงปกติของโปรแกรม SPSS

การแจกแจง	สถิติทดสอบ	ขนาดตัวอย่าง	α	
			.05	.10
$N(100, 10^2)$	1 – Sample K-S	10	.0000	.0000
		20	.0000	.0000
		30	.0000	.0000
		50	.0000	.0020
		100	.0000	.0000
	Lilliefors	10	.0520	.1200
		20	.0540	.0920
		30	.0620^{*C}	.1160
		50	.0340^{*C}	.0740^{*C}
		100	.0400	.1060
	S-W	10	.0580	.1120
		20	.0540	.0900
		30	.0680^{*C}	.1120
		50	.0340^{*C}	.0940
		100	.0540	.1020
$N(100, 50^2)$	1 – Sample K-S	10	.0000	.0000
		20	.0000	.0020
		30	.0020	.0000
		50	.0000	.0000
		100	.0000	.0040
	Lilliefors	10	.0540	.1200
		20	.0400	.0980
		30	.0340^{*C}	.0800
		50	.0520	.1000
		100	.0540	.1140
	S-W	10	.0540	.1080
		20	.0360^{*C}	.0980
		30	.0440	.0920
		50	.0540	.1000
		100	.0500	.0960

ตารางที่ 4.1 (ต่อ)

การแจกแจง	สถิติทดสอบ	ขนาดตัวอย่าง	α	
			.05	.10
N(100, 100^2)	1 – Sample K-S	10	.0000	.0000
		20	.0000	.0000
		30	.0000	.0000
		50	.0000	.0020
		100	.0020	.0000
	Lilliefors	10	.0520	.0900
		20	.0500	.1020
		30	.0420	.0880
		50	.0480	.1060
		100	.0540	.1000
	S-W	10	.0480	.1000
		20	.0720*C	.1200
		30	.0560	.0980
		50	.0560	.1100
		100	.0420	.0860

*C มีค่านอกขอบเขตตามเกณฑ์ของ Cochran, *B มีค่านอกขอบเขตตามเกณฑ์ของ Bradley

จากตารางที่ 4.1 จะพบว่า ในทุกกรณีของการแจกแจงปกติ เมื่อใช้สถิติทดสอบ 1-Sample K-S จะยอมรับสมมติฐานเบื้องต้นเกือบทั้งหมด นั่นคือ จำนวนครั้งของการปฏิเสธสมมติฐานเบื้องต้นมีค่าเป็นศูนย์ หรือเข้าใกล้ 0 เมื่อกำหนดระดับนัยสำคัญที่ 0.05 และ 0.10 หรือสรุปได้ว่าสถิติ K-S มักจะยอมรับว่าตัวแปรสุ่มนั้นมีการแจกแจงปกติเกือบทั้งหมด

ส่วนการทดสอบด้วยสถิติทดสอบ Lilliefors จะพบว่ามีการปฏิเสธสมมติฐานเบื้องต้น เมื่อเป็นจริงด้วยค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ส่วนใหญ่อยู่ในขอบเขตที่ควบคุมได้ตามเกณฑ์ของ Cochran โดยมีค่านอกขอบเขตเพียงร้อยละ 13 และส่วนใหญ่เป็นกรณีที่มีค่าน้อยกว่าขอบเขตล่างของ Cochran คือต่ำกว่า 0.04 หรือต่ำกว่า 0.08

และสถิติทดสอบ S-W ก็ให้ผลลัพธ์ในทำนองเดียวกันกับสถิติทดสอบ Lilliefors

4.1.2 ผลการเปรียบเทียบความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จากการทดลองใช้โปรแกรม **MINITAB** เพื่อทดสอบการแจกแจงปกติ จะนำเสนอด้วยตาราง 4.2 ดังนี้

ตาราง 4.2 ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จากสถิติทดสอบต่าง ๆ ที่ใช้ทดสอบการแจกแจงปกติของโปรแกรม MINITAB

การแจกแจง	สถิติทดสอบ	ขนาดตัวอย่าง	α	
			.05	.10
N(100, 10^2)	AD	10	.040	.102
		20	.036[*C]	.092
		30	.052	.116
		50	.038[*C]	.096
		100	.046	.094
	RJ	10	.044	.096
		20	.042	.104
		30	.054	.118
		50	.038[*C]	.090
		100	.046	.088
	K-S	10	.050	.112
		20	.040	.086
		30	.052	.116
		50	.026[*C]	.072[*C]
		100	.032[*C]	.098
N(10, 50^2)	AD	10	.054	.114
		20	.042	.086
		30	.032[*C]	.086
		50	.046	.114
		100	.038[*C]	.100
	RJ	10	.052	.120
		20	.030[*C]	.076[*C]
		30	.050	.076[*C]
		50	.044	.096
		100	.046	.098
	K-S	10	.050	.118
		20	.034[*C]	.100
		30	.034[*C]	.100
		50	.040	.112
		100	.046	.114
N(100, 100^2)	AD	10	.048	.120
		20	.034[*C]	.114
		30	.028[*C]	.112
		50	.026[*C]	.114
		100	.030[*C]	.110
	RJ	10	.044	.126[*C]
		20	.028[*C]	.120
		30	.024[*C*B]	.108
		50	.024[*C*B]	.116
		100	.026[*C]	.114
	K-S	10	.048	.128[*C]
		20	.040	.126[*C]
		30	.034[*C]	.120
		50	.036[*C]	.118
		100	.032[*C]	.122[*C]

*C มีค่านอกขอบเขตตามเกณฑ์ของ Cochran, *B มีค่านอกขอบเขตตามเกณฑ์ของ Bradley

จากตารางที่ 4.2 จะพบว่าสถิติทดสอบ AD จะให้ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ส่วนใหญ่ (ร้อยละ 73) อยู่ในขอบเขตที่ควบคุมได้ตามเกณฑ์ของ Cochran ส่วนค่าที่อยู่นอกขอบเขตนั้น ทั้งหมด เป็นกรณีที่มีค่าน้อยกว่าขอบเขตล่างของ Cochran คือต่ำกว่า 0.04 และพบว่าจะเป็นค่าที่เกิดขึ้นเฉพาะกรณีที่ α = .05 เท่านั้น

สถิติทดสอบ RJ ได้ผลเช่นเดียวกับสถิติ AD แต่เกิดขึ้นทั้งที่ α = .05 และ .10 และมีบางค่าที่อยู่นอกขอบเขตตามเกณฑ์ของ Bradley (ร้อยละ 6) โดยมีค่าต่ำกว่าขอบเขตล่างของ Bradley คือต่ำกว่า 0.025

ส่วนสถิติทดสอบแบบสุดท้าย คือ K-S จะพบว่าได้ค่าใกล้เคียงกับสถิติ Lilliefors ใน SPSS เป็นส่วนใหญ่ ซึ่งเป็นที่ยืนยันได้ว่าสถิติ K-S ใน MINITAB ที่แท้จริงคือ สถิติแบบ Lilliefors และสถิตินี้ใน MINITAB สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนได้ตามเกณฑ์ของ Cochran

4.2 **การเปรียบเทียบค่าอำนาจการทดสอบ** เพื่อหาผลสรุปว่า สถิติทดสอบใดจะมีค่าสูงที่สุด

4.2.1 ผลการเปรียบเทียบค่าอำนาจการทดสอบ จากการทดลองใช้โปรแกรม **SPSS** เพื่อทดสอบการแจกแจงปกติ จะนำเสนอด้วยตาราง 4.3 ดังนี้

ตารางที่ 4.3 ค่าอำนาจการทดสอบของสถิติทดสอบต่าง ๆ ที่ใช้ทดสอบการแจกแจงปกติของโปรแกรม **SPSS**

การแจกแจงที่เป็นจริง	ขนาดตัวอย่าง	สถิติทดสอบ					
		1-Sample K-S		Lilliefors		S-W	
		α		α		α	
		.05	.10	.05	.10	.05	.10
เป็นการแจกแจงที่มีหางสั้น							
	10	.000	.000	.054	.104	.066	.152
	20	.000	.000	.084	.170	.214	.362
Uniform (0-1)	30	.000	.000	.124	.246	.376	.596
	50	.000	.006	.288	.456	.778	.912
	100	.016	.052	.610	.750	.992	.998
	10	.000	.000	.064	.140	.076	.182
	20	.000	.000	.120	.184	.210	.358
Uniform (10-50)	30	.000	.002	.166	.276	.396	.586
	50	.000	.004	.248	.404	.742	.880
	100	.008	.044	.586	.770	.998	1.000
เป็นการแจกแจงที่มีหางยาว							
	10	.126	.190	.546	.606	.572	.620
	20	.482	.500	.836	.880	.856	.892
t(1)	30	.706	.794	.934	.960	.962	.966
	50	.924	.952	.996	.996	.996	.996
	100	.998	.998	1.000	1.000	1.000	1.000
	10	.000	.000	.114	.164	.116	.164
	20	.002	.010	.148	.220	.208	.258
t(5)	30	.004	.012	.150	.228	.250	.326
	50	.006	.016	.200	.286	.352	.446
	100	.030	.054	.342	.468	.576	.632

ตารางที่ 4.3 (ต่อ)

การแจกแจงที่เป็นจริง	ขนาดตัวอย่าง	สถิติทดสอบ					
		1-Sample K-S		Lilliefors		S-W	
		α		α		α	
		.05	.10	.05	.10	.05	.10
เป็นการแจกแจงที่โด่งกว่าปกติเล็กน้อย							
	10	.000	.000	.070	.134	.082	.134
	20	.000	.004	.096	.174	.132	.200
Logistic (0,0.5)	30	.000	.010	.108	.196	.166	.234
	50	.000	.008	.118	.190	.214	.288
	100	.006	.010	.168	.252	.318	.392
	10	.000	.000	.080	.144	.066	.138
	20	.002	.006	.106	.152	.122	.214
Logistic (0,1)	30	.000	.000	.078	.148	.128	.198
	50	.000	.004	.092	.160	.178	.266
	100	.002	.008	.158	.262	.284	.408
เป็นการแจกแจงที่โด่งมาก							
	10	.000	.000	.122	.218	.132	.204
	20	.000	.000	.210	.318	.280	.370
Laplace (0,0.5)	30	.006	.016	.308	.418	.376	.464
	50	.018	.054	.424	.576	.520	.620
	100	.082	.192	.704	.796	.802	.864
	10	.124	.204	.568	.646	.588	.660
	20	.500	.606	.858	.898	.888	.910
Cauchy (0,0.5)	30	.682	.786	.950	.964	.972	.982
	50	.942	.956	.998	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	10	.104	.174	.554	.614	.564	.618
	20	.468	.572	.814	.868	.836	.872
Cauchy (0,2)	30	.728	.784	.948	.962	.970	.978
	50	.928	.964	.996	1.000	.998	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000

จากตารางที่ 4.3 จะพบว่า จากโปรแกรม SPSS สถิติทดสอบ S-W จะมีประสิทธิภาพสูงที่สุดในทุกรูปแบบการแจกแจง และมีค่าอำนาจจากการทดสอบสูงขึ้นตามลำดับ เมื่อขนาดตัวอย่างที่ใช้มากขึ้นด้วย และเมื่อกำหนดให้ระดับนัยสำคัญของการทดสอบมีค่าสูงขึ้น คือ ที่ $\alpha = .10$ จะได้ค่าอำนาจการทดสอบที่สูงกว่าที่ $\alpha = .05$ ในบางการแจกแจง เช่น การแจกแจงที่โด่งมาก มีอำนาจการทดสอบสูงสุดถึง 100 เปอร์เซ็นต์ นั่นหมายความว่า สถิติทดสอบนี้จะปฏิเสธสมมติฐานเบื้องต้นเกี่ยวกับการแจกแจงปกติ ในทุกชุดข้อมูลตัวอย่างที่มาจากการแจกแจงอื่น ๆ

ในขณะที่การทดสอบด้วยสถิติ Lilliefors มีอำนาจการทดสอบมากในลำดับถัดไป และมีค่าอำนาจการทดสอบสูงขึ้น เมื่อใช้ขนาดตัวอย่างมากขึ้นและระดับนัยสำคัญที่สูงขึ้นเช่นเดียวกัน

ส่วนสถิติทดสอบ 1-Sample K-S มีอำนาจการทดสอบน้อยที่สุด ในเกือบทุกแบบการแจกแจงยกเว้นการแจกแจงแบบ Cauchy ที่มีอำนาจการทดสอบสูงที่สุดเท่ากับสถิติทดสอบ Lilliefors

และ S-W ส่วนใหญ่ของการแจกแจงที่เข้าใกล้แบบปกติ สถิติทดสอบนี้ไม่สามารถแยกได้ว่าไม่เป็นแบบปกติ จึงได้อำนาจการทดสอบเข้าใกล้ 0

4.2.2 ผลการเปรียบเทียบค่าอำนาจการทดสอบ จากการทดลองใช้โปรแกรม **MINITAB** เพื่อทดสอบการแจกแจงปกติ จะนำเสนอด้วยตาราง 4.4 ดังนี้

ตารางที่ 4.4 ค่าอำนาจการทดสอบของสถิติทดสอบต่าง ๆ ที่ใช้ทดสอบการแจกแจงปกติของโปรแกรม **MINITAB**

การแจกแจงที่เป็นจริง	ขนาดตัวอย่าง	สถิติทดสอบ					
		AD		RJ		K-S	
		α		α		α	
		.05	.10	.05	.10	.05	.10
เป็นการแจกแจงที่มีหางสั้น							
Uniform (0-1)	10	.068	.126	.042	.096	.058	.110
	20	.184	.287	.072	.202	.082	.174
	30	.306	.450	.160	.330	.122	.258
	50	.582	.757	.463	.690	.286	.457
	100	.932	.972	.962	.986	.620	.752
Uniform (10-50)	10	.076	.184	.046	.118	.062	.150
	20	.182	.294	.090	.192	.120	.188
	30	.344	.464	.188	.358	.170	.288
	50	.560	.726	.432	.648	.254	.412
	100	.944	.982	.962	.984	.584	.766
เป็นการแจกแจงที่มีหางยาว							
t(1)	10	.584	.646	.586	.668	.538	.610
	20	.870	.898	.800	.914	.836	.880
	30	.964	.972	.960	.974	.932	.960
	50	.996	1.000	.998	1.000	.996	.998
	100	1.00	1.000	1.000	1.000	1.000	1.000
t(5)	10	.122	.168	.138	.178	.112	.174
	20	.186	.272	.240	.312	.148	.224
	30	.224	.312	.298	.376	.148	.228
	50	.276	.375	.409	.483	.200	.288
	100	.443	.586	.618	.715	.350	.471
เป็นการแจกแจงที่โด่งกว่าปกติเล็กน้อย							
Logistic (0,0.5)	10	.077	.148	.081	.152	.071	.140
	20	.113	.189	.163	.225	.101	.179
	30	.138	.214	.192	.262	.110	.200
	50	.178	.250	.254	.336	.118	.190
	100	.242	.353	.373	.464	.167	.250
Logistic (0,1)	10	.080	.154	.096	.166	.080	.162
	20	.120	.187	.163	.245	.108	.157
	30	.106	.176	.160	.234	.082	.148
	50	.142	.218	.240	.322	.094	.166
	100	.264	.351	.371	.485	.158	.264

ตารางที่ 4.4 (ต่อ)

การแจกแจงที่เป็นจริง	ขนาดตัวอย่าง	สถิติทดสอบ					
		AD		RJ		K-S	
		α		α		α	
		.05	.10	.05	.10	.05	.10
เป็นการแจกแจงที่โด่งมาก							
	10	.146	.230	.180	.244	.128	.228
	20	.251	.345	.295	.397	.191	.289
Laplace (0,0.5)	30	.394	.520	.420	.576	.309	.464
	50	.550	.615	.583	.735	.427	.574
	100	.826	.898	.842	.922	.698	.820
	10	.589	.678	.606	.676	.575	.666
	20	.896	.914	.902	.918	.858	.892
Cauchy (0,0.5)	30	.992	.982	.996	.984	.969	.962
	50	1.000	1.000	1.000	1.000	.998	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	10	.590	.682	.602	.692	.560	.656
	20	.862	.892	.862	.892	.820	.892
Cauchy (0,2)	30	.972	.972	.978	.986	.946	.976
	50	.994	1.000	.996	1.000	.990	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000

จากตารางที่ 4.4 จะพบว่า จากโปรแกรม MINITAB สถิติทดสอบ RJ จะมีอำนาจการทดสอบสูงที่สุดเมื่อเทียบกับสถิติอีก 2 แบบ (โดยสถิติ RJ มีอำนาจการทดสอบต่ำกว่าสถิติ AD ในการแจกแจง Uniform เท่านั้น แต่เมื่อใช้ขนาดตัวอย่างเท่ากับ 100 จะมีอำนาจการทดสอบสูงกว่า) ในทุกขนาดตัวอย่างและทุกระดับนัยสำคัญ โดยมีอำนาจการทดสอบสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น ทั้งที่ระดับนัยสำคัญ .05 หรือ .10 และเมื่อขนาดตัวอย่างเท่ากับ 100 จะมีอำนาจการทดสอบสูงสุดถึงร้อยละ 100 ในบางการแจกแจง

สถิติทดสอบ AD มีอำนาจการทดสอบสูงรองจากสถิติ RJ และได้ค่าอำนาจการทดสอบในลักษณะเดียวกับสถิติ RJ ที่กล่าวข้างต้น คือขึ้นกับขนาดตัวอย่างและระดับนัยสำคัญ

สถิติทดสอบ K-S มีอำนาจการทดสอบต่ำที่สุด และได้ค่าอำนาจการทดสอบในทำนองเดียวกับสองการทดสอบแรก คือ มีค่าสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น และที่ระดับนัยสำคัญ .10 จะได้ค่าอำนาจการทดสอบสูงกว่าที่ .05

4.3 การเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของสถิติทดสอบต่าง ๆ จากโปรแกรม SPSS และโปรแกรม MINITAB

จากตารางที่ 4.1 และตารางที่ 4.2 สามารถสรุปได้ว่า

ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 จากสถิติ RJ (ใน MINITAB ที่กล่าวว่าคล้ายสถิติ Shapiro-Wilk) เมื่อเทียบกับค่าจาก S-W (ใน SPSS) ส่วนใหญ่จะมีค่าเล็กกว่า และสามารถควบคุมค่าได้ตามเกณฑ์ของ Cochran

ค่าจากสถิติ K-S (ใน MINITAB) เมื่อเทียบกับค่าจากสถิติ Lilliefors (ใน SPSS ด้วยเมนู Explore) จะพบว่าได้ค่าใกล้เคียงกันมากกว่า เมื่อเทียบกับสถิติ 1-sample K-S (ใน SPSS ด้วยเมนู Nonparametric Test) จึงอาจยืนยันได้ว่าสถิติ K-S ใน MINITAB แท้ที่จริงคือ สถิติ Lilliefors และพบว่าได้ผลเช่นเดียวกับสถิติ RJ ที่กล่าวข้างต้น

ส่วนค่าจากสถิติ AD ใน MINITAB พบว่าส่วนใหญ่มีค่าในขอบเขตที่ควบคุมได้ตามเกณฑ์ของ Cocharn ดังรายละเอียดจากท้ายตารางที่ 4.2

ส่วนค่าจากสถิติ 1-sample K-S ใน SPSS จะพบว่ามีค่าเป็นศูนย์หรือใกล้ศูนย์เกือบทั้งหมด ซึ่งแปลผลได้ว่าสถิติทดสอบนี้ เมื่อใช้ทดสอบกับข้อมูลที่มีการแจกแจงปกติ จะยอมรับว่ามีการแจกแจงปกติทั้งหมด ไม่มีความผิดพลาดเลย คือยอมรับร้อยละ 100

4.4 การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบต่าง ๆ จากโปรแกรม SPSS และโปรแกรม MINITAB

จากตารางที่ 4.3 และ 4.4 สามารถสรุปได้ว่า

สถิติทดสอบ RJ (ใน MINITAB) ได้ค่าอำนาจการทดสอบสูงกว่าค่าจากสถิติ S-W (ใน SPSS) เป็นส่วนใหญ่ (จาก 7 ใน 10 ของการแจกแจงแบบต่าง ๆ) ในทุกขนาดตัวอย่างทุกระดับนัยสำคัญ และสถิติทดสอบนี้มีค่าอำนาจการทดสอบสูงที่สุดเมื่อเทียบกับสถิติทดสอบอื่น ๆ ในทุกสถานการณ์

สถิติทดสอบ K-S (ใน MINITAB) ได้ค่าอำนาจการทดสอบใกล้เคียงกับค่าจากสถิติ Lilliefors (ใน SPSS ด้วยเมนู Explore) ซึ่งเป็นหลักฐานแสดงให้เห็นว่าสถิติ K-S ใน MINITAB แท้ที่จริงคือ Lilliefors อย่างแน่นอน แต่สถิตินี้ได้ค่าอำนาจการทดสอบต่ำกว่าสถิติ RJ (หรือ S-W ใน SPSS)

สถิติ AD ใน MINITAB ได้อำนาจการทดสอบต่ำกว่าสถิติ RJ ดังรายละเอียดท้ายตาราง 4.4 ที่กล่าวข้างต้น

สถิติ 1-sample K-S ใน SPSS มีอำนาจการทดสอบต่ำที่สุดเมื่อเทียบกับสถิติทดสอบทั้งหมด 6 แบบ (จาก SPSS จำนวน 3 แบบ และ MINITAB จำนวน 3 แบบ) คือได้ค่าอำนาจการทดสอบเป็นค่าศูนย์หรือเข้าใกล้ศูนย์เป็นส่วนมาก แม้จะเพิ่มขนาดตัวอย่างมากขึ้น และเพิ่มค่าระดับนัยสำคัญของการทดสอบเป็น 0.10 แทนที่ 0.05 ก็ตาม

บทที่ 5
สรุปผลการวิจัยและข้อเสนอแนะ

5.1 สรุปผลการวิจัย

การวิจัยครั้งนี้มีวัตถุประสงค์ เพื่อทำการเปรียบเทียบประสิทธิภาพในการทดสอบการแจกแจงปกติโดยใช้โปรแกรมสำเร็จรูปทางสถิติ SPSS V.13 ซึ่งมีสถิติที่ใช้ทดสอบ 3 แบบให้ผู้ใช้เลือกคือ Kolmogorov, Lilliefors และ Shapiro – Wilk และโดยใช้โปรแกรม MINITAB 14 ซึ่งมีสถิติที่ใช้ทดสอบ 3 แบบ ให้ผู้ใช้เลือกคือ Kolmogorov, Ryan-Joiner (RJ ที่กล่าวว่าคล้ายสถิติ Shapiro-Wilk) และ Anderson-Darling ข้อมูลที่นำมาศึกษาเป็นข้อมูลจากการจำลองแบบด้วยเมนู Calc-Random-Data จากโปรแกรม MINITAB โดยจำลองให้เป็นประชากรที่มีการแจกแจงปกติ และที่เบี่ยงเบนไปเล็กน้อยจากการแจกแจงปกติ ด้วยการแจกแจงที่มีลักษณะสมมาตร แต่เบี่ยงเบนจากปกติไปบ้าง เช่น หางสั้น คือ $u(0-1)$ และ $u(10-50)$ หางยาว เช่น $t(1)$ และ $t(5)$ โด่งกว่าปกติเล็กน้อย เช่น L(0,0.5) และ L(0,1) และโด่งกว่าปกติมาก เช่น C(0,2), C(0,0.5) และ L(0,0.5) ในแต่ละกรณีมีการทำซ้ำ 500 ครั้ง ของการจำลองแบบข้อมูล และให้มีขนาดตัวอย่างเท่ากับ 10, 20, 30, 50 และ 100 ทำการเปรียบเทียบค่าอำนาจการทดสอบ และความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ I ของสถิติทดสอบแต่ละแบบ เพื่อหาผลสรุปว่าสถิติทดสอบใดควรจะถูกเลือกใช้ในแต่ละสถานการณ์ เพื่อจะได้ผลสรุปที่ถูกต้องเชื่อถือได้

แผนการทดลอง

1. ทดลองหาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 (Empirical Alpha หรือ Actual Type I error) ของสถิติทดสอบทั้งหกแบบ เมื่อกำหนดระดับนัยสำคัญที่ .05 และ .10
2. ทดลองหาค่าอำนาจการทดสอบของสถิติทดสอบทั้งหกแบบ จากสถิติทดสอบที่สามารถควบคุมค่าความคลาดเคลื่อนชนิดที่ 1 ได้ ตามเกณฑ์ของ Cochran และ Bradley

แผนการทดลองนี้จะเสนอค่าความคลาดเคลื่อนชนิดที่ 1 จากผลการทดลองในกรณีต่าง ๆ ทั้งสิ้น 90 ค่า สำหรับ $\alpha = .05$ และอีก 90 ค่า สำหรับ $\alpha = .10$ และเสนอค่าอำนาจการทดสอบทั้งสิ้น 270 ค่า สำหรับ $\alpha = .05$ และ 270 ค่า สำหรับ $\alpha = .10$

วิธีดำเนินการทดลอง

การวิจัยนี้ใช้เทคนิคมอนติคาร์โลซิมมูเลชั่น โดยแต่ละกรณีกำหนดให้คอมพิวเตอร์จำลองการทดลอง 500 ครั้ง และนับจำนวนครั้งที่เกิดความคลาดเคลื่อนชนิดที่ 1 และอำนาจการทดสอบของสถิติทดสอบทั้ง 6 แบบ เพื่อคำนวณหาค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ที่แท้จริง และค่าอำนาจการทดสอบ

สรุปผลการวิจัย

การสรุปผลว่าสถิติทดสอบใดมีความเหมาะสมที่จะเลือกใช้จากโปรแกรม SPSS V.13 หรือจากโปรแกรม MINITAB 14 จะใช้วิธีเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบ 3 แบบในแต่ละโปรแกรม และเปรียบเทียบค่าอำนาจทดสอบของสถิติที่คล้ายกันระหว่างโปรแกรมว่าสถิติทดสอบใดจะมีค่าสูงสุด ซึ่งจะพิจารณาเฉพาะสถิติทดสอบที่มีความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 เท่านั้น ผลสรุปที่สำคัญของการวิจัยเป็นดังนี้

ผลสรุปจากโปรแกรม SPSS

1. การเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของสถิติทดสอบทั้ง 3 แบบ ที่ระดับนัยสำคัญ .05 และ .10 โดยใช้เกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ของ Cochran และ Bradley

1.1 สถิติทดสอบ 1-Sample K-S สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้เกือบทุกกรณีของการแจกแจงปกติ และ ณ ทุกระดับนัยสำคัญ .05 หรือ .10 โดยมีค่าน้อยกว่าความคลาดเคลื่อนที่กำหนดไว้ ในทุกกรณี และเกือบทั้งหมดมีค่าเป็น 0

1.2 สถิติทดสอบ Lilliefors สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ โดยค่าที่ได้ส่วนใหญ่อยู่ในขอบเขตที่ควบคุมได้ตามเกณฑ์ของ Cochran โดยมีค่าอยู่นอกขอบเขตเพียงร้อยละ 13 และค่านอกขอบเขตนี้ส่วนใหญ่มีค่าน้อยกว่าขอบเขตล่างของ Cochran คือต่ำกว่า .04 หรือต่ำกว่า .08

1.3 สถิติทดสอบ Shapiro-Wilk ให้ผลลัพธ์ในทำนองเดียวกับสถิติทดสอบ Lilliefors

ดังนั้นอาจสรุปได้ว่าสถิติทดสอบทั้ง 3 ใน SPSS เป็นสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ทั้งหมด ซึ่งจะนำผลลัพธ์จากอำนาจการทดสอบจากสถิติทดสอบทั้ง 3 มาเปรียบเทียบเพื่อหาผลสรุปว่าควรจะใช้สถิติทดสอบใดมากที่สุดในกรณีต่าง ๆ ที่ทดลองในครั้งนี้

2. เปรียบเทียบค่าอำนาจการทดสอบของสถิติทั้ง 3 แบบ ที่จะใช้ในการทดสอบการแจกแจงปกติ ที่ระดับนัยสำคัญ .05 และ .10 สำหรับการแจกแจงต่าง ๆ ที่สมมาตร แต่เบี่ยงเบนไปเล็กน้อยจากปกติ

2.1 สถิติทดสอบ 1-Sample K-S มีอำนาจการทดสอบค่อนข้างต่ำ และมีอำนาจการทดสอบต่ำที่สุด เมื่อเปรียบเทียบกับสถิติทดสอบอีก 2 แบบ โดยส่วนใหญ่มีค่าใกล้ 0 แม้จะใช้ตัวอย่างใหญ่ขึ้นถึง 100 ก็ตาม และมีค่าสูงกว่ากันเล็กน้อย ระหว่าง $\alpha = .05$ และ .10 ยกเว้นการแจกแจงแบบ Cauchy ที่มีอำนาจการทดสอบสูงที่สุด เท่ากับสถิติทดสอบ Lilliefors และ S-W

2.2 สถิติทดสอบ Lilliefors มีอำนาจการทดสอบที่สูงขึ้นกว่าสถิติทดสอบ 1-Sample K-S และมีค่าสูงขึ้นเมื่อใช้ตัวอย่างขนาดใหญ่ขึ้น มีค่าอำนาจการทดสอบสูงเมื่อใช้

$\alpha = .10$ มากกว่าที่ $\alpha = .05$ และมีอำนาจการทดสอบสูงที่สุด คือร้อยละ 100 เมื่อประชากรมาจากการแจกแจงแบบ t(1), Cauchy (0, 0.5) และ Cauchy (0, 2) จากตัวอย่างขนาด 100

2.3 สถิติทดสอบ S-W มีประสิทธิภาพสูงที่สุด เมื่อเทียบกับสถิติทดสอบอีก 2 แบบ กล่าวคือ มีค่าอำนาจการทดสอบสูงขึ้นเมื่อใช้ขนาดตัวอย่างใหญ่ขึ้น และที่ระดับนัยสำคัญ 0.10 จะมีค่าอำนาจการทดสอบสูงกว่าที่ระดับนัยสำคัญ .05 ส่วนใหญ่จากการแจกแจงแบบต่าง ๆ สถิติทดสอบนี้จะมีค่าอำนาจการทดสอบเข้าใกล้ร้อยละ 100 เมื่อใช้ขนาดตัวอย่างเท่ากับ 100

การสรุปผลเพื่อนำไปปฏิบัติใช้

จากผลการวิจัยครั้งนี้ ถ้าผู้ใช้จะใช้โปรแกรม SPSS เพื่อหาผลสรุปว่าตัวแปรสุ่มมีการแจกแจงปกติหรือไม่ ในสำคัญแรกให้ดูกราฟโค้งการแจกแจงว่าใกล้เคียงสมมาตรมากน้อยเพียงใด ถ้าสมมาตรแนะนำให้ใช้เมนู Explore ด้วยสถิติทดสอบ Shpairo-Wilk เมื่อกำหนดระดับนัยสำคัญ .10 และใช้ขนาดตัวอย่างเท่ากับ 100 จะทำให้ผู้ใช้เชื่อมั่นได้ถึงร้อยละ 100 ว่าสถิติทดสอบนี้จะปฏิเสธสิ่งที่ผิด ในกรณีที่ผู้ใช้มีข้อจำกัดของขนาดตัวอย่าง อาจลดขนาดตัวอย่างลงเป็น 50 (และใช้ $\alpha = .10$) เพราะอำนาจการทดสอบจะลดลงจากเมื่อใช้ขนาดตัวอย่างเท่ากับ 100 ไม่เกินร้อยละ 24 ในบางการแจกแจงทั้งสองขนาดตัวอย่างมีอำนาจการทดสอบเท่ากันเท่ากับ 100

ผลสรุปจากโปรแกรม MINITAB

1. การเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของสถิติทดสอบทั้ง 3 แบบ ที่ระดับนัยสำคัญ .05 และ .10 โดยใช้เกณฑ์การพิจารณาความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ของ Cochran และ Bradley

1.1 สถิติทดสอบ AD จะให้ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ส่วนใหญ่ (ร้อยละ 73) อยู่ในขอบเขตที่ควบคุมได้ตามเกณฑ์ของ Cochran ส่วนค่าที่อยู่นอกขอบเขตนั้นทั้งหมด เป็นกรณีที่มีค่าน้อยกว่าขอบเขตล่างของ Cochran คือต่ำกว่า 0.04 และพบว่าจะเป็นค่าที่เกิดขึ้นเฉพาะกรณีที่ $\alpha = .05$ เท่านั้น

1.2 สถิติทดสอบ RJ ได้ผลเช่นเดียวกับสถิติ AD แต่เกิดขึ้นทั้งที่ $\alpha = .05$ และ .10 และมีบางค่าที่อยู่นอกขอบเขตตามเกณฑ์ของ Bradley (ร้อยละ 6) โดยมีค่าต่ำกว่าขอบเขตล่างของ Bradley คือ ต่ำกว่า 0.025

1.3 สถิติทดสอบ K-S จะพบว่าได้ค่าใกล้เคียงกับสถิติ Lilliefors ใน SPSS เป็นส่วนใหญ่ จึงเป็นที่ยืนยันได้ว่าสถิติ K-S ใน MINITAB ที่แท้จริงคือ สถิติแบบ Lilliefors และสถิตินี้ใน MINITAB สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนได้ตามเกณฑ์ของ Cochran

ดังนั้นอาจสรุปได้ว่า สถิติทดสอบทั้ง 3 แบบใน MINITAB เป็นสถิติที่สามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้ทั้งหมด จึงจะหาอำนาจการทดสอบของสถิติแต่ละแบบต่อไป จากการแจกแจงต่าง ๆ ที่สมมาตรแต่เบี่ยงเบนเล็กน้อยจากปกติ

2. การเปรียบเทียบอำนาจการทดสอบของสถิติทั้ง 3 แบบ ที่ระดับนัยสำคัญ .05, .10 สำหรับการแจกแจงต่าง ๆ ที่สมมาตรแต่เบี่ยงเบนไปเล็กน้อยจากปกติ

2.1 สถิติทดสอบ RJ จะมีอำนาจการทดสอบสูงที่สุดเมื่อเทียบกับสถิติอีก 2 แบบ (โดยสถิติ RJ มีอำนาจการทดสอบต่ำกว่าสถิติ AD ในการแจกแจง Uniform เท่านั้น แต่เมื่อใช้ขนาดตัวอย่างเท่ากับ 100 จะมีอำนาจการทดสอบสูงกว่า) ในทุกขนาดตัวอย่างและทุกระดับนัยสำคัญ โดยมีอำนาจการทดสอบสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น ทั้งที่ระดับนัยสำคัญ .05 หรือ .10 และเมื่อขนาดตัวอย่างเท่ากับ 100 จะมีอำนาจการทดสอบสูงสุดถึงร้อยละ 100 ในบางการแจกแจง

2.2 สถิติทดสอบ AD มีอำนาจการทดสอบสูงรองจากสถิติ RJ และได้ค่าอำนาจการทดสอบในลักษณะเดียวกับสถิติ RJ ที่กล่าวข้างต้น คือขึ้นกับขนาดตัวอย่างและระดับนัยสำคัญ

2.3 สถิติทดสอบ K-S มีอำนาจการทดสอบต่ำที่สุด และได้ค่าอำนาจการทดสอบในทำนองเดียวกับสองการทดสอบแรก คือ มีค่าสูงขึ้นเมื่อขนาดตัวอย่างใหญ่ขึ้น และที่ระดับนัยสำคัญ .10 จะได้ค่าอำนาจการทดสอบสูงกว่าที่ .05 และยังพบว่าค่าอำนาจการทดสอบจากสถิติ K-S นี้จะได้ค่าใกล้เคียงกับค่าจากสถิติ Lilliefors ใน SPSS มาก จึงยืนยันได้ว่าที่แท้จริงคือสถิติ Lilliefors

ผลสรุปจากการเปรียบเทียบระหว่างโปรแกรม SPSS และ MINITAB

1. การเปรียบเทียบความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ของสถิติทดสอบทั้ง 6 แบบ จากโปรแกรม SPSS และ MINITAB

ได้ผลสรุปจากการวิจัยครั้งนี้ ว่า สถิติทดสอบทุกแบบให้ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 เกือบทั้งหมด เป็นไปตามเกณฑ์ของ Cochran เป็นส่วนใหญ่ คือที่ $\alpha = .05$ จะให้ค่าอยู่ในช่วง .04-.06 โดยค่าที่อยู่นอกขอบเขตจะเป็นกรณีที่มีค่าต่ำกว่าขอบเขตล่าง คือ ต่ำกว่า .04 ส่วนที่ $\alpha = .10$ จะให้ค่าอยู่ในช่วง .08-.12 เกือบทั้งหมดเช่นเดียวกัน

และพบว่าสถิติ 1-Sample K-S ใน SPSS จะให้ค่าเป็นศูนย์ (หรือเข้าใกล้ศูนย์) เกือบทุกแบบของการแจกแจงปกติ หมายความว่าสถิติ K-S นี้จะยอมรับว่าประชากรมีการแจกแจงปกติในทุกกลุ่มตัวอย่าง โดยไม่มีความคลาดเคลื่อนเลย (หรือมีน้อยมาก)

2. การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบทั้ง 6 แบบ จากโปรแกรม SPSS และ MINITAB

ได้ผลสรุปจากการวิจัยครั้งนี้ว่า สถิติ RJ (ใน MINITAB) จะเป็นสถิติทดสอบที่ให้ค่าอำนาจการทดสอบสูงที่สุดในทุกแบบการแจกแจงทุกขนาดตัวอย่าง ทุกระดับนัยสำคัญ (แม้ว่าจะให้ค่าต่ำกว่าค่าจากสถิติ S-W และ AD ในการแจกแจงแบบยูนิฟอร์ม ในทุกกรณีแต่พบว่าถ้าเพิ่มขนาด

ตัวอย่างเท่ากับ 100 จะได้ค่าต่างกันเล็กน้อย หรือแม้จะพบว่าได้ค่าเข้าใกล้ 1 เมื่อขนาดตัวอย่างเท่ากับ 100 เท่ากับค่าจากสถิติ AD และ S-W ก็ตาม แต่ที่ขนาดตัวอย่างต่ำกว่า 100 จะให้ค่าสูงกว่าค่าจากสถิติ AD และ S-W เกือบทั้งหมด) โดยค่าอำนาจการทดสอบนี้จะมีค่าสูงขึ้นเมื่อเพิ่มขนาดตัวอย่างตามลำดับ โดยมีค่าเข้าใกล้ 1 เมื่อขนาดตัวอย่าง = 100 และที่ระดับนัยสำคัญ = .10 จะให้ค่าอำนาจการทดสอบสูงกว่าที่ระดับนัยสำคัญ .05 ในทุกขนาดตัวอย่าง

3. การเปรียบเทียบค่าอำนาจการทดสอบของสถิติทดสอบที่คล้ายกันใน SPSS และ MINITAB

สถิติทดสอบที่มีเหมือนกันใน SPSS และ MINITAB มี 2 คู่ คือ

1. สถิติ S-W (ใน SPSS) กับสถิติ RJ (ใน MINITAB) พบว่าค่าอำนาจการทดสอบจากสถิติ RJ จะมีค่าสูงกว่าค่าจากสถิติ S-W เกือบทั้งหมดของการทดลองครั้งนี้

2. สถิติ K-S (ใน MINITAB) กับสถิติ 1-sample K-S หรือ Lilliefors (ใน SPSS) ค่าจากตารางที่ 4.1 – 4.4 สามารถยืนยันได้ว่าสถิติ K-S ใน MINITAB แท้ที่จริง คือ สถิติ Lilliefors โดยสถิติทั้งสองให้ค่าอำนาจการทดสอบที่ใกล้เคียงกันในเกือบทุกกรณีในการทดลองครั้งนี้

การสรุปผลเพื่อนำไปปฏิบัติใช้

ถ้าผู้ใช้สามารถเลือกใช้โปรแกรม MINITAB หรือ SPSS และข้อมูลมีรูปโค้งการแจกแจงใกล้เคียงโค้งที่สมมาตร ผลจากการวิจัยครั้งนี้ แนะนำให้ผู้ใช้ใช้โปรแกรม MINITAB ด้วยสถิติ RJ เมื่อกำหนดระดับนัยสำคัญ .10 โดยมีขนาดตัวอย่างเท่ากับ 100 จะทำให้ผู้ใช้มั่นใจว่าผลวิเคราะห์จากสถิติทดสอบนี้ จะถูกต้องเกือบร้อยละ 100 ในกรณีที่ไม่สามารถหาขนาดตัวอย่างได้ถึง 100 อาจลดลงเป็นขนาดเท่ากับ 50 ซึ่งผลการทดสอบจะมีประสิทธิภาพลดลงอยู่ในช่วงร้อยละ 16 ถึงร้อยละ 30

5.2 การอภิปรายผล

ผลการศึกษาครั้งนี้ได้ผลสอดคล้องกับงานวิจัยของ Gan และ Kochler (คศ. 1990) Edith Seier (คศ.2004) (18) และเกตุจันทร์ (พศ. 2534) ที่ได้ข้อสรุปว่าสถิติทดสอบ S-W มีอำนาจการทดสอบที่สูง แต่ขัดแย้งกับคู่มือการใช้โปรแกรม SPSS ที่แนะนำให้ใช้สถิติทดสอบ S-W เมื่อขนาดตัวอย่าง ≤ 50 จากการศึกษาครั้งนี้พบว่าสถิติทดสอบ S-W และ RJ จะมีค่าอำนาจการทดสอบเข้าใกล้ 1 เมื่อขนาดตัวอย่างเข้าใกล้ 100 ในเกือบทุกแบบของการแจกแจง และในทุกระดับของ α ที่ .05 หรือ .10 นั่นหมายความว่า ถ้านักวิจัยต้องการเชื่อมั่นในคำตอบของการทดสอบการแจกแจงปกติด้วยสถิติทดสอบ S-W หรือ RJ ในระดับสูง ควรจะใช้ขนาดตัวอย่างเข้าใกล้ 100

ส่วนสถิติทดสอบ Lilliefors (หรือ K-S ใน MINITAB) นั้นให้ผลการทดสอบใกล้เคียงกับสถิติทดสอบ S-W แต่ส่วนใหญ่มีอำนาจการทดสอบต่ำกว่าสถิติทดสอบ S-W ดังกล่าวในตอนต้นว่าการแจกแจงของสถิติทดสอบนี้แยกได้หลายกรณี เช่น กรณีที่ไม่ทราบค่าพารามิเตอร์ทั้งสองของ

การแจกแจงปกติ (คือ μ และ σ^2) หรือกรณีที่ทราบค่าพารามิเตอร์เพียงค่าเดียว (เช่น ทราบค่า μ หรือ σ^2) แต่ในโปรแกรม SPSS และ MINITAB ใช้เพียงกรณีเดียว คือ ไม่ทราบ μ และ σ^2 จึงน่าศึกษาต่อไปว่า Algorithm ของ SPSS และ MINITAB เลือกใช้การแจกแจงที่ถูกต้องหรือไม่ คือได้ค่าพีที่แท้จริงหรือไม่ ถ้า SPSS และ MINITAB ได้สร้างเมนูให้ผู้ใช้เลือกกรณี ต่าง ๆ ก็จะทำให้ได้การแจกแจงที่ถูกต้องอย่างแท้จริง

สถิติทดสอบ 1-Sample K-S มีประสิทธิภาพต่ำที่สุด ซึ่งสอดคล้องกับผลสรุปจากการศึกษาของ D' Agostino และ Stephens (1986) ว่าสถิติทดสอบนี้ไม่ควรจะนำไปใช้ทดสอบการแจกแจงปกติ เพราะมีอำนาจการทดสอบต่ำเมื่อเทียบกับการทดสอบอื่น ๆ (อ้างถึงใน Nonparametric Statistics with Applications to Science and Engineering By Kvam and Vidakovic หน้า 96) และสอดคล้องกับงานของ STeinskog et.al. ที่สรุปว่าสถิติทดสอบนี้มักจะยอมรับว่าตัวแปรสุ่มมีการแจกแจงปกติ (หรือยากที่จะปฏิเสธว่าไม่เป็นการแจกแจงปกติ) และให้ผู้ใช้ระวังการใช้สถิตินี้ในโปรแกรมสำเร็จรูปทางสถิติหลาย ๆ โปรแกรมที่ใช้ชื่อนี้ แต่อาจจะเป็นสถิติ Lilliefors

สถิติทดสอบ AD (ใน MINITAB) ที่โปรแกรมจัดให้เป็น Default ของโปรแกรมมีประสิทธิภาพสูงในลำดับที่สองรองจากสถิติ RJ (ใน MINITAB) แต่มีประสิทธิภาพใกล้เคียงกับ S-W (ใน SPSS) (คือมีค่าอำนาจการทดสอบต่ำกว่าสถิติ S-W ประมาณร้อยละ 50 ของค่าอำนาจการทดสอบที่วัดทั้งหมด 90 ค่า ที่เหลือมีค่าสูงกว่าค่าจาก S-W) ผลที่ได้นี้ขัดแย้งกันกับคู่มือการใช้ MINITAB → Home > Support > Answer ID : 1167 (22) ที่ตอบคำถาม "ควรจะเลือกใช้การทดสอบใดสำหรับทดสอบการแจกแจงปกติ" ว่าในทางปฏิบัตินักสถิติจะเลือกสถิติ AD เป็นอันดับแรก ถ้าให้ความสำคัญของการเบี่ยงเบนไปจากโค้งปกติที่หางของโค้งการแจกแจง และยังให้คำตอบว่าสถิติทดสอบทั้งสามมีประสิทธิภาพต่ำในการจำแนกการแจกแจงแบบ t หรือการแจกแจงที่มีความโด่งออกจากการแจกแจงปกติ ซึ่งขัดแย้งกับผลงานวิจัยครั้งนี้ที่ได้ทดลองกับการแจกแจงแบบ t ที่ t(1) และ t(5) ซึ่งพบว่าที่ t(1) สถิติทั้งสามมีค่าอำนาจการทดสอบสูงมาก คือ สูงกว่าร้อยละ 90 ที่ขนาดตัวอย่างเพียง 30 และมีค่าเป็นร้อยละ 100 เมื่อขนาดตัวอย่างเป็น 50 ส่วน t(5) สถิติ AD และ K-S มีค่าอำนาจการทดสอบค่อนข้างต่ำ (ต่ำกว่าร้อยละ 35-58) แม้จะใช้ขนาดตัวอย่างถึง 100 ก็ตาม แต่สถิติ RJ ให้ค่าสูงพอสมควรที่ขนาดตัวอย่างเท่ากับ 100 คือ ร้อยละ 72 (ที่ $\alpha = .10$) และร้อยละ 62 (ที่ $\alpha = .05$) สำหรับการแจกแจงที่มีความโด่งมาก คือ L(0,0.5), C(0,0.5) และ C(0,2) พบว่าสถิติทดสอบทั้งสามมีประสิทธิภาพสูงมากถึงร้อยละ 100 ที่ขนาดตัวอย่างเท่ากับ 100 เกือบทั้งหมด (มีค่าต่ำสุดเพียงร้อยละ 70 จาก K-S ที่ $\alpha = .05$ เมื่อ $n = 100$) โดยสถิติ RJ มีค่าอำนาจการทดสอบสูงที่สุด แต่งานวิจัยนี้ให้ผลสรุปสนับสนุนกับคำตอบจากคู่มือการใช้ MINITAB สำหรับการแจกแจงที่มีความโด่งเล็กน้อย (Logistic) คือพบว่ามีค่าอำนาจการทดสอบต่ำกว่าร้อยละ 50 แม้จะใช้ตัวอย่างเท่ากับ 100 และสถิติ RJ จะมีค่าสูงที่สุด (ร้อยละ 48.5 ที่ $n = 100$ และ $\alpha = .10$)

5.3 ข้อเสนอแนะ

เพื่อให้ผลสรุปครอบคลุมกว้างขวางขึ้น ควรจะมีการศึกษากับข้อมูลที่แตกต่างจากการแจกแจงปกติอย่างมาก เช่น มีความเบ้เล็กน้อย เบ้มาก หรือมีค่าฐานนิยม 2 ค่า (Bimodal) หรือ Scale Contaminated and Mixture of Normal Distribution และศึกษาในประเด็นของสถิติทดสอบ 1-Sample K-S ใน SPSS ว่าตามเมนูที่ผู้ใช้สามารถเลือกทดสอบการแจกแจง 4 แบบ นอกจาก Normal คือ Exponential, Uniform และ Poisson นั้น สถิติทดสอบนี้ให้ผลถูกต้องมากน้อยอย่างไร และประเด็นสุดท้ายเกี่ยวกับสถิติทดสอบ Lilliefors ว่าค่าพีที่ได้มาจากการแจกแจงที่ระบุพารามิเตอร์อย่างไร เพื่อทำให้ผู้ใช้มั่นใจได้ว่าผลการทดสอบได้จากการแจกแจงที่ถูกต้อง

บรรณานุกรม

1. James J. Higgins. **"Introduction to Modern Nonparametric Statistics"** Thomson Brook/cole, 2004

2. P.Sprent. **"Applied Nonparametric Statistical Methods"** 2nd edition, Chapman & Hall, London, 1993

3. กัลยา วาณิชย์บัญชา **"การใช้ SPSS for Windows ในการวิเคราะห์ข้อมูล เวอร์ชัน 7-10"** โรงพิมพ์ห้างหุ้นส่วนจำกัด ซี เค แอนด์ เอส โฟโต้สตูดิโอ, 2544.

4. Anderson, T. W. and Darling, D. A. "A Test of Goodness-of-fit." **Journal of the American Statistical Association.**, 1952, 49, 765-769

5. Kolmogorov, A.N. "Sulla determinazione empirica di una legge di distribuzione" **Giorna Ist. Attuari.**, 1933, 4, 83-91.

6. Shapiro, S. and Wilk, M.B. "An analysis of variance test for normality" **Biometrika**, 1965, 52, 591-611

7. Zhang, P. "Omnibus test of normality using the Q statistic." **Journal of Applied Statistics** , 1999, 26, 519-528

8. W.J. Conover. **"Practical Nonparametric Statistics"** John Wiley & Sons Inc. 1971

9. Jean Dickinson Gibbons. **"Nonparametric Methods for Quantitative Analysis"** Holt, Inehart and Winston, 1976

10. STeinskog, Tjostheim and Kvamsto "A Cautionary Note on the Use of Kolmogorov-Smirnov Test for Normality" **Journal of American Meteorological Society**, 2007, March, 1151-1156

11. Paul H. Kvam and Brani Vidakovic. **"Nonparametric Statistics with Applications to Science and Engineering"** John Wiley & Sons, Inc., Publications, 2007

12. อุมาพร จันทศร และ มนัส ไพฑูรย์เจริญลาภ "การเปรียบเทียบผลการทดสอบการแจกแจงปกติด้วยสถิติทดสอบจากโปรแกรมสำเร็จรูปทางสถิติ SPSS" **วารสารวิทยาศาสตร์ลาดกระบัง** ปีที่ 17 ฉบับที่ 2 กรกฎาคม – ธันวาคม 2551

13. Gan, F.F. and Kochler, K.J. "Goodness of Fit Test Based on P-P Probability Plots." **Technometrics** 32, 289-303, 1990

14. เกตุจันทร์ พัชรินทร์ศักดิ์. “การเปรียบเทียบวิธีการนอนพาราเมตริกสำหรับการทดสอบการแจกแจงปกติ” **วิทยานิพนธ์**ภาควิชาสถิติ บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย พ.ศ. 2534

15. กมล บุษบา. “การคำนวณค่าความคลาดเคลื่อนมาตรฐานที่ไม่ถูกต้องของโปรแกรม SPSS ในการทดลองแบบแฟกทอเรียลกรณีตัวแบบอิทธิพลผสม” **วารสารวิทยาศาสตร์และเทคโนโลยี ธรรมศาสตร์**, 2547 ปีที่ 12 ฉบับที่ 1 ม.ค. – เม.ย.

16. กมล บุษบา. “การรายงานค่า p-value ที่ทำให้เข้าใจผิดของโปรแกรม SPSS สำหรับการวิเคราะห์ข้อมูลในตารางไขว้ด้วย ไค – สแควร์” **วารสารวิทยาศาสตร์และเทคโนโลยี ธรรมศาสตร์**, 2546 ปีที่ 11 ฉบับที่ 2 ก.ค. – ธ.ค.

17. Reinhard Bergmann, John Ludbrook, and Will P.J.M. Spooren “Different Outcomes of the Wilcoxon-Mann-Whitney Test From Different Statistics Packages” **Journal of The American Statistician**, 2000, 54, No.1

18. Leo Knüsel. “On the accuracy of statistical distributions in Microsoft Excel 2003” **Computational Statistics & Data Analysis**, 48, 2005, 445 – 449

19. B.D. McCullough, Berry Wilson. “On the accuracy of Statistical procedures in Microsoft Excel 97” **Computational Statistics & Data Analysis**, 31, 1999, 27 – 37

20. กุศยา ปลั่งพงษ์พันธ์. “การเปรียบเทียบประสิทธิภาพของการจำลองแบบข้อมูลโดยการใช้โปรแกรมสำเร็จรูปทางสถิติ” **การประชุมวิชาการสถิติ และสถิติประยุกต์** ประจำปี 2550

21. Edith Seier. “Comparison of Test for Univariate Normality” Department of Mathematics, East Tennessee State University, Johnson City, TN 37614, 2004

22. http://www.minitab.com/support/answers/answer.aspx?ID=1167

ภาคผนวก

ตารางที่ 1 ตารางการแจกแจงของสถิติทดสอบ Kolmogorov-Smirnov

ตารางที่ 1 ควอนไทล์ของสถิติทดสอบโคลโมโกรอฟ (Quantiles of the Kolmogorov test statistic)

One-Sided Test											
	$p = 0.90$	0.95	0.975	0.99	0.995		$p = 0.90$	0.95	0.975	0.99	0.995
Two-Sided Test											
	$p = 0.80$	0.90	0.95	0.98	0.99		$p = 0.80$	0.90	0.95	0.98	0.99
$n = 1$	0.900	0.950	0.975	0.990	0.995	$n = 21$	0.226	0.259	0.287	0.321	0.344
2	0.684	0.776	0.842	0.900	0.929	22	0.221	0.253	0.281	0.314	0.337
3	0.565	0.636	0.708	0.785	0.829	23	0.216	0.247	0.275	0.307	0.330
4	0.493	0.565	0.624	0.689	0.734	24	0.212	0.242	0.269	0.301	0.323
5	0.447	0.509	0.563	0.627	0.669	25	0.208	0.238	0.264	0.295	0.317
6	0.410	0.468	0.519	0.577	0.617	26	0.204	0.233	0.259	0.290	0.311
7	0.381	0.436	0.483	0.538	0.576	27	0.200	0.229	0.254	0.284	0.305
8	0.358	0.410	0.454	0.507	0.542	28	0.197	0.225	0.250	0.279	0.300
9	0.339	0.387	0.430	0.480	0.513	29	0.193	0.221	0.246	0.275	0.295
10	0.323	0.369	0.409	0.457	0.489	30	0.190	0.218	0.242	0.270	0.290
11	0.308	0.352	0.391	0.437	0.468	31	0.187	0.214	0.238	0.266	0.285
12	0.296	0.338	0.375	0.419	0.449	32	0.184	0.211	0.234	0.262	0.281
13	0.285	0.325	0.361	0.404	0.432	33	0.182	0.208	0.231	0.258	0.277
14	0.275	0.314	0.349	0.390	0.418	34	0.179	0.205	0.227	0.254	0.273
15	0.266	0.304	0.338	0.377	0.404	35	0.177	0.202	0.224	0.251	0.269
16	0.258	0.295	0.327	0.366	0.392	36	0.174	0.199	0.221	0.247	0.265
17	0.250	0.286	0.318	0.355	0.381	37	0.172	0.196	0.218	0.244	0.262
18	0.244	0.279	0.309	0.346	0.371	38	0.170	0.194	0.215	0.241	0.258
19	0.237	0.271	0.301	0.337	0.361	39	0.168	0.191	0.213	0.238	0.255
20	0.232	0.265	0.294	0.329	0.352	40	0.165	0.189	0.210	0.235	0.252
				Approximation for $n > 40$			$\frac{1.07}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.52}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

SOURCE. Adapted from Table 1 of Miller (1956). Used with permission of the American Statistical Association.

[a] The entries in this table are selected quantiles w_p of the Kolmogorov test statistics T, T^+, and T^- as defined by Equation 6.1.1 for two-sided tests and by Equations 6.1.2 and 6.1.3 for one-sided tests. Reject H_0 at the level α if T exceeds the $1 - \alpha$ quantile given in this table. These quantiles are exact for $n \leq 40$ in the two-tailed test. The other quantiles are approximations that are equal to the exact quantiles in most cases. A better approximation for $n > 40$ results if $(n + \sqrt{n/10})^{1/2}$ is used instead of $\sqrt{n}$ in the denominator.

ตารางที่ 2 ตารางการแจกแจงของสถิติทดสอบ Lilliefors

TABLE 2 I

Critical values for Lilliefors test, normal case 1 (μ unknown, σ^2 known)

	α				
n	0.20	0.15	0.10	0.05	0.01
3	.392	.308	.428	.453	.495
4	.351	.366	.384	.410	.455
5	.318	.333	.350	.376	.423
6	.294	.307	.324	.348	.396
7	.276	.288	.305	.328	.374
8	.260	.272	.288	.311	.353
9	.246	.258	.272	.294	.334
10	.234	.245	.259	.280	.323
11	.225	.235	.249	.269	.309
12	.216	.226	.238	.259	.300
13	.209	.218	.230	.249	.285
14	.202	.211	.224	.242	.280
15	.195	.205	.217	.235	.270
16	.189	.197	.209	.227	.261
17	.184	.192	.203	.220	.256
18	.179	.187	.198	.215	.246
19	.174	.182	.194	.210	.242
20	.170	.178	.189	.205	.235
21	.166	.174	.184	.199	.230
22	.163	.171	.180	.195	.227
23	.160	.167	.177	.193	.221
24	.156	.164	.173	.188	.217
25	.154	.160	.170	.185	.214
26	.151	.158	.167	.181	.209
27	.147	.154	.163	.177	.205
28	.146	.153	.161	.174	.202
29	.143	.149	.158	.172	.198
30	.141	.147	.155	.169	.193

TABLE 2 II

Critical values for Lilliefors test, normal case 2 (μ known, σ^2 unknown)

	α				
n	0.20	0.15	0.10	0.05	0.01
2	.739	.770	.797	.820	.837
3	.551	.599	.657	.722	.798
4	.499	.529	.565	.621	.734
5	.440	.470	.507	.567	.660
6	.400	.429	.464	.514	.607
7	.375	.395	.429	.477	.566
8	.451	.374	.405	.450	.534
9	.332	.353	.382	.425	.505
10	.315	.335	.361	.401	.477
11	.300	.320	.346	.387	.466
12	.289	.307	.332	.371	.444
13	.277	.296	.320	.358	.428
14	.266	.284	.307	.341	.410
15	.259	.275	.297	.331	.397
16	.251	.257	.288	.322	.387
17	.244	.260	.282	.313	.377
18	.236	.251	.271	.302	.369
19	.231	.246	.266	.297	.357
20	.226	.241	.260	.290	.348
21	.219	.233	.252	.282	.337

TABLE 2 II

(*continued*)

n	α 0.20	0.15	0.10	0.05	0.01
22	.214	.228	.247	.278	.334
23	.210	.223	.242	.270	.319
24	.205	.218	.236	.263	.317
25	.202	.214	.231	.256	.308
26	.197	.210	.227	.255	.305
27	.194	.208	.224	.250	.302
28	.191	.203	.219	.244	.292
29	.188	.200	.217	.242	.290
30	.185	.198	.212	.236	.284
50	$1.02/\sqrt{n}$	$1.080/\sqrt{n}$	$1.170/\sqrt{n}$	$1.310/\sqrt{n}$	$1.595/\sqrt{n}$
100	$1.04/\sqrt{n}$	$1.100/\sqrt{n}$	$1.180/\sqrt{n}$	$1.320/\sqrt{n}$	$1.610/\sqrt{n}$
≥ 101	$1.06/\sqrt{n}$	$1.120/\sqrt{n}$	$1.190/\sqrt{n}$	$1.333/\sqrt{n}$	$1.625/\sqrt{n}$

TABLE 2 III

Critical values for Lilliefors test, normal case 3 (μ, σ^2 both unknown)

n	α 0.20	0.15	0.10	0.05	0.01
4	.303	.320	.344	.374	.414
5	.290	.302	.319	.344	.398
6	.268	.280	.295	.321	.371
7	.252	.264	.280	.304	.353
8	.239	.251	.266	.290	.333
9	.227	.239	.253	.275	.319
10	.217	.228	.241	.262	.303
11	.209	.219	.232	.252	.291
12	.201	.210	.223	.243	.281
13	.193	.203	.215	.233	.270
14	.187	.196	.209	.227	.264
15	.181	.190	.202	.219	.256
16	.176	.184	.195	.212	.248
17	.170	.179	.190	.207	.241
18	.166	.174	.185	.201	.234
19	.162	.171	.181	.197	.230
20	.159	.167	.177	.192	.223
21	.155	.163	.173	.188	.219
22	.152	.160	.170	.185	.214
23	.149	.156	.165	.181	.210
24	.145	.153	.162	.177	.205
25	.144	.151	.159	.173	.202
26	.141	.147	.156	.170	.198
27	.138	.145	.153	.166	.193
28	.136	.142	.151	.165	.191
29	.134	.140	.149	.162	.188
30	.132	.138	.146	.159	.183
31	$\frac{0.741}{d_n}$	$\frac{0.775}{d_n}$	$\frac{0.819}{d_n}$	$\frac{0.895}{d_n}$	$\frac{1.035}{d_n}$

$$d_n = (\sqrt{n} - 0.01 + 0.83/\sqrt{n})$$

Source: Andrew L. Mason and C. B. Bell, "New Lilliefors and Srinivasan Tables with Applications," *Communic. Statist.—Simul.*, Vol. 15, No. 2 (1986), pp. 457–459. Copyright (c) 1986 by Marcel Dekker, Inc.; reprinted by permission.

ตารางที่ 3 Percentage Points of the W Test* for n = 3(1)50

	Level								
n	0·01	0·02	0·05	0·10	0·50	0·90	0·95	0·98	0·99
3	0·753	0·756	0·767	0·789	0·959	0·998	0·999	1·000	1·000
4	·687	·707	·748	·792	·935	·987	·992	·996	·997
5	·686	·715	·762	·806	·927	·979	·986	·991	·993
6	0·713	0·743	0·788	0·826	0·927	0·974	0·981	0·986	0·989
7	·730	·760	·803	·838	·928	·972	·979	·985	·988
8	·749	·778	·818	·851	·932	·972	·978	·984	·987
9	·764	·791	·829	·859	·935	·972	·978	·984	·986
10	·781	·806	·842	·869	·938	·972	·978	·983	·986
11	0·792	0·817	0·850	0·876	0·940	0·973	0·979	0·984	0·986
12	·805	·828	·859	·883	·943	·973	·979	·984	·986
13	·814	·837	·866	·889	·945	·974	·979	·984	·986
14	·825	·846	·874	·895	·947	·975	·980	·984	·986
15	·835	·855	·881	·901	·950	·975	·980	·984	·987
16	0·844	0·863	0·887	0·906	0·952	0·976	0·981	0·985	0·987
17	·851	·869	·892	·910	·954	·977	·981	·985	·987
18	·858	·874	·897	·914	·956	·978	·982	·986	·988
19	·863	·879	·901	·917	·957	·978	·982	·986	·988
20	·868	·884	·905	·920	·959	·979	·983	·986	·988
21	0·873	0·888	0·908	0·923	0·960	0·980	0·983	0·987	0·989
22	·878	·892	·911	·926	·961	·980	·984	·987	·989
23	·881	·895	·914	·928	·962	·981	·984	·987	·989
24	·884	·898	·916	·930	·963	·981	·984	·987	·989
25	·888	·901	·918	·931	·964	·981	·985	·988	·989
26	0·891	0·904	0·920	0·933	0·965	0·982	0·985	0·988	0·989
27	·894	·906	·923	·935	·965	·982	·985	·988	·990
28	·896	·908	·924	·936	·966	·982	·985	·988	·990
29	·898	·910	·926	·937	·966	·982	·985	·988	·990
30	·900	·912	·927	·939	·967	·983	·985	·988	·900
31	0·902	0·914	0·929	0·940	0·967	0·983	0·986	0·988	0·990
32	·904	·915	·930	·941	·968	·983	·986	·988	·990
33	·906	·917	·931	·942	·968	·983	·986	·989	·990
34	·908	·919	·933	·943	·969	·983	·986	·989	·900
35	·910	·920	·934	·944	·969	·984	·986	·989	·990
36	0·912	0·922	0·935	0·945	0·970	0·984	0·986	0·989	0·990
37	·914	·924	·936	·946	·970	·984	·987	·989	·990
38	·916	·925	·938	·947	·971	·984	·987	·989	·990
39	·917	·927	·939	·948	·971	·984	·987	·989	·991
40	·919	·928	·940	·949	·972	·985	·987	·989	·991
41	0·920	0·929	0·941	0·950	0·972	0·985	0·987	0·989	0·991
42	·922	·930	·942	·951	·972	·985	·987	·989	·991
43	·923	·932	·943	·951	·973	·985	·987	·990	·991
44	·924	·933	·944	·952	·973	·985	·987	·990	·991
45	·926	·934	·945	·953	·973	·985	·988	·990	·991
46	0·927	0·935	0·945	0·953	0·974	0·985	0·988	0·990	0·991
47	·928	·936	·946	·954	·974	·985	·988	·990	·991
48	·929	·937	·947	·954	·974	·985	·988	·990	·991
49	·929	·937	·947	·955	·974	·985	·988	·990	·991
50	·930	·938	·947	·955	·974	·985	·988	·990	·991

* Based on fitted Johnson (1949) S_B approximation, see Shapiro & Wilk (1965a) for details.