



รายงานการวิจัยฉบับสมบูรณ์

การเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้น

A Comparison of Statistical Packages in Data Analysis of Linear Regression

นางสาว ชูใจ กุهارัตนไชย

ได้รับทุนสนับสนุนงานวิจัยจากทุนอุดหนุนการวิจัย (เงินรายได้) ประจำปีงบประมาณ พ.ศ.2560

คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อโครงการวิจัย(ภาษาไทย) การเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ข้อมูลการ
ถดถอยเชิงเส้น

แหล่งเงินงบประมาณ เงินรายได้ประจำปีงบประมาณพ.ศ.2560

จำนวนเงินที่ได้รับการสนับสนุน 40,000 บาท

ระยะเวลาทำการวิจัย 1 ปี ตั้งแต่วันที่ 1 ตุลาคม 2559 ถึงวันที่ 30 กันยายน 2560

ชื่อ – สกุล นางสาวชอุใจ กุฬารัตนไชย หัวหน้าโครงการ ภาควิชาสถิติ คณะวิทยาศาสตร์

บทคัดย่อ

โปรแกรมสำเร็จรูปทางสถิติ SPSS Minitab และ Microsoft Excel ถือเป็นโปรแกรมที่นิยมใช้ในการวิเคราะห์ข้อมูล ในการทำวิจัยครั้งนี้จึงมีวัตถุประสงค์เพื่อเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ SPSS Minitab และ Microsoft Excel ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่ายและการถดถอยเชิงเส้นพหุคูณ โดยการจำลองข้อมูลเมื่อกำหนดค่าความแปรปรวน (σ^2) มีค่าเท่ากับ 10 , 5 , 2 , 1.5 , 1 , 0.5 และขนาดตัวอย่าง(n) มีค่าเป็น 100 , 60 , 40 , 20 ค่าMSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ SPSS MINITAB และ Microsoft Excel จะให้ค่าใกล้เคียงกันทุกสถานการณ์ที่จำลองข้อมูล และเมื่อนำค่า MSE ดังกล่าวมาทำการวิเคราะห์ความแปรปรวน (ANOVA) จะให้ค่า P-value เข้าใกล้ 1 ทุกสถานการณ์เช่นกัน

Statistical packages including SPSS, Minitab and Microsoft Excel are highly preferable packages for data analysis. The objective of this research is to compare SPSS, Minitab and Microsoft Excel when used in data analysis by linear regression and multiple linear regression methods using data simulation with the variance (σ^2) values of 10, 5, 2, 1.5, 1, and 0.5 while the sample sizes are 100, 60, 40, and 20. The MSE values, calculated by the SPSS, Minitab and Microsoft Excel statistical packages, were similar in every situation of the simulation. The Analysis of Variance (ANOVA) of such MSE values also results in P-value close to 1 in every situation.

Keywords : Regression analysis , Statistical packages

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำนำ

รายงานการวิจัยเรื่อง การเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้น เป็นงานวิจัยที่ศึกษาจากการจำลองข้อมูลให้มีรูปแบบการถดถอยเชิงเส้นอย่างง่ายและเชิงเส้นพหุคูณ เพื่อนำเอาข้อมูลในแต่ละสถานการณ์จำลองมาใช้ในการเปรียบเทียบค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) ที่คำนวณโดยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 โดยมีวัตถุประสงค์ในการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่ายและเชิงเส้นพหุคูณ จากผลการวิจัยที่ได้จะทำให้ทราบถึงความแตกต่างของผลลัพธ์ที่ได้จากการใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลทางด้านการถดถอยเชิงเส้นอย่างง่ายและเชิงเส้นพหุคูณ เพื่อใช้เป็นแนวทางในการเลือกโปรแกรมที่เหมาะสมกับข้อมูลที่ใช้ในการวิเคราะห์

ผู้วิจัยหวังเป็นอย่างยิ่งว่า รายงานวิจัยนี้จะเป็นประโยชน์ต่อคณาจารย์ นักศึกษา และผู้ที่เกี่ยวข้องทั่วไป

ผศ. ชูใจ กุฬารัตนไชย

23 มกราคม 2562

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

เรื่อง	หน้า
บทคัดย่อภาษาไทย	ก
บทคัดย่อภาษาอังกฤษ	ก
คำนำ	ข
สารบัญตาราง	จ
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญ	1
1.2 วัตถุประสงค์ของการวิจัย	2
1.3 ประโยชน์ที่คาดว่าจะได้รับ	2
1.4 ขอบเขตของงานวิจัย	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	
2.1 ความรู้ทั่วไปเกี่ยวกับโปรแกรมสำเร็จรูปทางสถิติ	4
2.2 การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย	14
2.3 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ	18
2.4 ค่าวัดประสิทธิภาพของรูปแบบ	21
2.5 การทดสอบความแปรปรวน	23
2.6 การทดสอบการแจกแจงแบบปกติของประชากร	24
2.7 การทดสอบของเลวีเน (Levene's Test)	25
2.8 การทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส	26
2.8 งานวิจัยที่เกี่ยวข้อง	27
บทที่ 3 การดำเนินงานวิจัย	
3.1 ข้อมูลที่ใช้ในงานวิจัย	31
3.2 สถิติที่ใช้ในการวิเคราะห์	33
บทที่ 4 ผลการวิเคราะห์ข้อมูล	
4.1 ผลการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่าย	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

เรื่อง	หน้า
4.2 ผลการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นพหุคูณง่าย	40
บทที่ 5 สรุปผลการวิเคราะห์และข้อเสนอแนะ	
5.1 ข้อมูลการถดถอยเชิงเส้นอย่างง่าย	47
5.2 ข้อมูลการถดถอยเชิงเส้นพหุคูณ	48
5.3 ข้อเสนอแนะ	48
บรรณานุกรม	50



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

เรื่อง	หน้า
ตารางที่ 4.1 แสดงผลการทดสอบการแจกแจงปกติของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย	35
ตารางที่ 4.2 แสดงผลการทดสอบความแปรปรวนของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย	37
ตารางที่ 4.3 แสดงผลการทดสอบค่าเฉลี่ยของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย	39
ตารางที่ 4.4 แสดงผลการทดสอบการแจกแจงปกติของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ	41
ตารางที่ 4.5 แสดงผลการทดสอบความแปรปรวนของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ	43
ตารางที่ 4.6 แสดงผลการทดสอบค่าเฉลี่ยของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ	45

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญ

โปรแกรมประยุกต์ทางสถิติถือเป็นโปรแกรมสำเร็จรูปที่ใช้กันอย่างกว้างขวางในวงการศึกษา ซึ่งบุคลากรที่ใช้งานได้แก่ นักศึกษา นักวิจัย นักสถิติ และนักวิชาการ ฯลฯ ที่ต้องการคำนวณหรือวิเคราะห์หาค่าทางสถิติต่างๆ ทั้งสถิติพรรณนา และสถิติเชิงอ้างอิง เช่น การหาค่าเฉลี่ย มัชยฐาน ฐานนิยม การทดสอบค่าเฉลี่ยโดยใช้ T-Test หรือ ANOVA การวิเคราะห์การถดถอย การแสดงกราฟ และแผนภูมิต่างๆ เป็นต้น ซึ่งโปรแกรมสำเร็จรูปทางสถิติที่นิยมใช้กัน คือ โปรแกรม SPSS , Minitab , SAS , R และ Microsoft Excel และโปรแกรมที่หาใช้ได้ง่ายและมีรูปแบบในการใช้คำสั่งที่ไม่ยุ่งยากคือ โปรแกรม SPSS , Minitab และ Microsoft Excel

การวิเคราะห์การถดถอย(Regression analysis) เป็นวิธีหนึ่งที่น่ามาใช้มากในการวิเคราะห์ทางด้านสถิติอยู่ตลอดเวลา ซึ่งวิธีนี้เป็นการศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะประมาณหรือพยากรณ์ค่าของตัวแปรตัวหนึ่ง ในที่นี้จะให้เป็นตัวแปรตาม Y จากตัวแปรอื่นๆ หรือตัวแปรอิสระ X_1, X_2, \dots, X_k ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์ ถ้าตัวแปรตามกับตัวแปรอิสระมีความสัมพันธ์ในรูปแบบเส้นตรง จะเรียกว่า การถดถอยเชิงเส้น(Linear regression) ซึ่งสามารถแบ่งเป็น 2 แบบ คือ การถดถอยเชิงเส้นอย่างง่าย(Simple linear regression) เป็นการวิเคราะห์การถดถอยของตัวแปรอิสระ 1 ตัวและตัวแปรตาม 1 ตัว โดยตัวแปรทั้งสองจะมีความสัมพันธ์เชิงเส้นตรง และการถดถอยเชิงเส้นพหุคูณ(Multiple linear regression) เป็นการวิเคราะห์การถดถอยของตัวแปรอิสระมากกว่า 1 ตัวและตัวแปรตาม 1 ตัว โดยตัวแปรอิสระกับตัวแปรตามมีความสัมพันธ์เชิงเส้นตรง

จากงานวิจัยในเรื่องการเปรียบเทียบความน่าเชื่อถือของโปรแกรมสำเร็จรูปสำหรับงานวิเคราะห์ทางสถิติ ได้ทำการประมาณค่าทางสถิติ 3 ชนิดคือ การวิเคราะห์สถิติตัวแปรเดียว(Univariate Summary Statistic), การวิเคราะห์ความแปรปรวนจำแนกทางเดียว(One-Way Analysis of Variance) และการวิเคราะห์การถดถอยเชิงเส้น(Linear Regression Analysis) ของโปรแกรมสำเร็จรูปทั้ง 5 โปรแกรม ประกอบด้วย โปรแกรม SAS 9 , R 2.6.2 , Microsoft Excel 2007 , SPSS 15 และ Minitab 15 โดยใช้ข้อมูลจาก The National Institute of Standards and Technology (NIST) ซึ่งแบ่งข้อมูลตามระดับความยาก

3 ระดับคือ ระดับต่ำ, ระดับปานกลาง และระดับสูง ซึ่งผลการวิจัยสรุปในส่วนของการวิเคราะห์การเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทดลองพบว่า ค่าความแม่นยำของสถิติทดสอบเอฟของโปรแกรม SAS 9, R 2.6.2 และ Microsoft Excel 2007 มีค่าความแม่นยำในระดับสูง ส่วนโปรแกรม SPSS 15 และ Minitab 15 เมื่อข้อมูลมีความยาก ระดับสูง จะให้ค่าความแม่นยำในระดับต่ำ ส่วนค่าความแม่นยำของค่าสัมประสิทธิ์การถดถอย เมื่อข้อมูลมีความยากระดับต่ำและปานกลาง ทั้ง 5 โปรแกรม จะให้ค่าความแม่นยำในระดับสูง แต่เมื่อข้อมูลมีความยากระดับสูง พบว่า โปรแกรม SAS 9, R 2.6.2, Microsoft Excel 2007 และ Minitab 15 ให้ค่าความแม่นยำไม่แตกต่างกัน แต่โปรแกรม SPSS 15 ให้ค่าความแม่นยำต่ำกว่าโปรแกรมอื่น [1] และจากการศึกษา งานวิจัย การเปรียบเทียบการใช้โปรแกรม Minitab, InStat และ R Commander กรณีการทดสอบด้วยไคกำลังสอง โดยใช้โจทย์ปัญหา 3 ตัวอย่าง มาทำการทดสอบด้วยไคกำลังสอง พบว่า การคำนวณค่าทางสถิติของการทดสอบด้วยไคกำลังสองจากทั้ง 3 โปรแกรม ให้ผลลัพธ์ในข้อสรุปที่เหมือนกัน เพียงแต่มีความแตกต่างกันในเรื่องการนำเข้าสู่ข้อมูลและการแจ้งข้อความเตือนเมื่อความถี่คาดหวังมีค่าน้อยกว่า 5 เท่านั้น[2]

จากที่กล่าวมาจะเห็นได้ว่า ในการวิเคราะห์ข้อมูลโดยใช้โปรแกรมสำเร็จรูปทางสถิติที่แตกต่างกัน จะให้ผลลัพธ์ทางสถิติในเรื่องเดียวกันที่เหมือนหรือแตกต่างกันก็ได้ และจากการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายและเชิงเส้นพหุคูณ สามารถใช้โปรแกรมสำเร็จรูปทางสถิติ SPSS Minitab และ Microsoft Excel มาช่วยในการวิเคราะห์ข้อมูล ผู้วิจัยจึงเห็นถึงความสำคัญในการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายและพหุคูณ เพื่อให้ผู้ใช้โปรแกรมสำเร็จรูปทางสถิติสามารถเลือกใช้โปรแกรมได้อย่างเหมาะสมกับลักษณะของข้อมูลที่ใช้ในการวิเคราะห์

1.2 วัตถุประสงค์ของการวิจัย

- 1.2.1 เพื่อเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่าย
- 1.2.2 เพื่อเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นพหุคูณ

1.3 ประโยชน์ที่คาดว่าจะได้รับ

ทำให้ทราบถึงความแตกต่างของผลลัพธ์ที่ได้จากการใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลทางด้านการถดถอยเชิงเส้นอย่างง่ายและเชิงเส้นพหุคูณ

เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 ขอบเขตของงานวิจัย

งานวิจัยนี้จะแบ่งงานออกเป็น 2 ส่วน คือ

ส่วนที่ 1 จะทำการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่าย โดยมีขั้นตอนดังนี้

1. จากรูปแบบการถดถอยเชิงเส้นอย่างง่าย $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ จะทำการจำลองค่า X_i และค่า ϵ_i ในการสร้างค่า Y_i
2. นำค่า X_i และค่า Y_i ที่ได้ มาหาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) โดยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007
3. ศึกษาขนาดตัวอย่าง 20 , 40 , 60 และ 100

ส่วนที่ 2 จะทำการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นพหุคูณ โดยมีขั้นตอนดังนี้

1. จากรูปแบบการถดถอยเชิงเส้นพหุคูณ $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$ จะทำการจำลองค่า X_{1i} , X_{2i} และค่า ϵ_i ในการสร้างค่า Y_i
2. นำค่า X_{1i} , X_{2i} และค่า Y_i ที่ได้ มาหาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) โดยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007
3. ศึกษาขนาดตัวอย่าง 20 , 40 , 60 และ 100

ในการวิจัยครั้งนี้ จะทำการจำลองข้อมูลจากโปรแกรม Minitab 16 ซึ่งจะกระทำซ้ำกัน 100 ครั้งในแต่ละสถานการณ์

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ความรู้ทั่วไปเกี่ยวกับโปรแกรมสำเร็จรูปทางสถิติ

โปรแกรมสำเร็จรูปทางสถิติที่นิยมใช้ ได้แก่

1. กลุ่มโปรแกรมสำเร็จรูปที่เรียกสั้นๆ ว่า SPSS ได้แก่ SPSS SPSS^x SPSS/PC และ SPSS for Windows โดยที่ SPSS ย่อมาจากคำว่า Statistical Package for the Social Sciences เป็นชุดโปรแกรมสถิติที่นิยมใช้ทางด้านสังคมศาสตร์ ออกเผยแพร่ครั้งแรกในปีพ.ศ. 2511 หลังจากถูกพัฒนาโดย Norman H. Nie และ C. Hadlai Hull โปรแกรมๆในกลุ่มนี้ สร้างขึ้นในประเทศสหรัฐอเมริกา โดยเริ่มพัฒนาให้ใช้งานบนเครื่องไมโครคอมพิวเตอร์ขนาดใหญ่ ต่อมาก็สามารถใช้งานบนเครื่องไมโครคอมพิวเตอร์ และล่าสุดได้รับการพัฒนาให้ใช้ร่วมกับโปรแกรมจัดระบบงานที่ใช้รูปภาพและเมนูเป็นหลัก ที่เรียกว่าวินโดว (Windows) และเมื่อเดือนมิถุนายน พ.ศ.2552 บริษัทจัดจำหน่าย SPSS ถูกซื้อโดยบริษัท IBM

2. โปรแกรมสำเร็จรูปมินิแท็บ(Minitab) เป็นโปรแกรมสำเร็จรูปที่พัฒนาโดยบริษัท Minitab Inc. ให้สามารถใช้ได้กับคอมพิวเตอร์หลายขนาดทั้งไมโครคอมพิวเตอร์ มินิคอมพิวเตอร์ จนถึงระดับซูเปอร์คอมพิวเตอร์ การพัฒนาโปรแกรมนี้เริ่มขึ้นโดยมีวัตถุประสงค์เพียงเพื่อประกอบการสอนวิชาสถิติ โดยใช้ประมวลผลค่าสถิติต่างๆ และได้รับการปรับปรุงให้มีประสิทธิภาพเพิ่มขึ้น จนใช้งานได้กว้างขวางทั้งทางธุรกิจ การเกษตร การวิจัยและการเรียนการสอน

3. โปรแกรมไมโครซอฟท์ เอกซ์เซล (Microsoft Excel) เป็นโปรแกรมประเภท สเปรดชีต (spreadsheet) หรือตารางคำนวณอิเล็กทรอนิกส์ที่ใช้เก็บบันทึกข้อมูลในลักษณะต่างๆ ซึ่งส่วนใหญ่มักเก็บข้อมูลประเภทการคำนวณ โดยจะเก็บข้อมูลลงในตารางสี่เหลี่ยมที่เรียกว่า เซล (Cell)ที่สามารถนำเอาเซลมาอ้างอิงใส่ในสูตร เพื่อให้โปรแกรมคำนวณหาผลลัพธ์จากข้อมูลที่บันทึกไว้ได้ และยังนิยมใช้ในด้าน การคำนวณทางคณิตศาสตร์ สถิติ และทางการเงิน

2.1.1 โปรแกรมสำเร็จรูป SPSS[3]

SPSS เป็นโปรแกรมสำเร็จรูปทางสถิติ เพื่อการวิเคราะห์ข้อมูลทางสถิติ โดย SPSSย่อมาจาก Statistical Package for the Social Sciences ซึ่งเป็นลิขสิทธิ์ของบริษัท SPSS จำกัด แห่งประเทศสหรัฐอเมริกา โดยแนวคิดของ SPSS ในรุ่น (version) ปัจจุบัน ได้ปรับเปลี่ยนเป็น Statistical Product and Service Solutions

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

SPSS เป็นโปรแกรมคอมพิวเตอร์ที่สามารถใช้กับเครื่องไมโครคอมพิวเตอร์ (Personal Computer) หรือเครื่องขนาดใหญ่ก็ได้ โดย SPSS เป็นโปรแกรมที่มีการพัฒนาและเปลี่ยนแปลงปรับปรุงมาโดยตลอด ดังนี้

1. SPSS^x เป็นโปรแกรมที่ใช้กับเครื่องขนาดใหญ่ที่เป็นประเภท Mini-computer หรือ Main-frame computer
2. SPSS/PC เป็นโปรแกรมที่ใช้กับเครื่อง Microcomputer หรือ Personal Computer ที่ทำงานบนระบบ DOS
3. SPSS for Windows เป็นโปรแกรมประเภทเดียวกับ SPSS/PC เพียงแต่ทำงานบนระบบ windows

โปรแกรม SPSS for Windows ได้ถูกพัฒนาให้สามารถทำงานบนโปรแกรมควบคุมระบบ Microsoft windows โดยมีการปรับปรุงรูปแบบการใช้งานให้ดูง่ายสำหรับผู้ใช้งาน (Graphic user interface) และมีประสิทธิภาพสูง สะดวกในการทำงาน วิเคราะห์ข้อมูลทางสถิติได้อย่างละเอียด พร้อมทั้งสามารถนำเสนอข้อมูลในรูปของกราฟและตารางในรูปแบบที่หลากหลาย และยังสามารถรับข้อมูลที่สร้างจากโปรแกรมประเภทอื่นได้ด้วย เช่น Microsoft Excel, LOTUS, dBASE, SPSS/PC หรือข้อมูลในรูปแบบของแอสกี (Text file) นอกจากนี้ SPSS for windows ยังสามารถบันทึกไฟล์ข้อมูลที่ใช้งานให้อยู่ในรูปแบบของไฟล์ที่นำไปใช้กับโปรแกรมประเภทอื่นๆ ได้อีกด้วย ส่วนวิธีการใช้งานของ SPSS for windows เป็นรูปแบบการใช้งานที่ง่าย เนื่องจากมีระบบเมนูให้โต้ตอบในรูปแบบที่เป็น Dialog box และเป็นระบบหน้าต่างให้เลือกชนิด Check box หรือ Option button นอกจากนี้ผู้ใช้งานยังสามารถบันทึกขั้นตอนการทำงานที่ผู้ใช้เลือกมาเป็นชุดคำสั่ง (Syntax) คล้ายกับการทำงานเป็นโปรแกรมตามแบบของ SPSS/PC ส่วนการใช้ SPSS for Windows นั้น ผู้ใช้ควรมีพื้นฐานการใช้โปรแกรมประยุกต์บน Microsoft windows มาบ้างแล้ว ซึ่งจะทำให้การใช้งาน SPSS for windows คล่องตัวมากขึ้น

ปัจจุบัน โปรแกรม SPSS มีการพัฒนาโปรแกรมรุ่นใหม่ออกมาอยู่เป็นประจำ และมีขอบเขตงานวิเคราะห์ทางสถิติ ที่ครอบคลุมกว้างขวางขึ้น ไม่ว่าจะเป็นงานการวิเคราะห์ข้อมูลทางด้านธุรกิจ การตลาด การควบคุมการผลิต ตลอดจนการวิเคราะห์ข้อมูลทางด้านวิทยาศาสตร์ ทางอุตสาหกรรม ทำให้ความหมายของ SPSS เปลี่ยนมาเป็น Statistical Product and Service Solutions

โปรแกรม SPSS for windows เป็นโปรแกรมคอมพิวเตอร์ที่สามารถติดตั้งบนเครื่องไมโครคอมพิวเตอร์และทำงานบน Microsoft windows ซึ่งโปรแกรม SPSS for Windows ประกอบด้วยชุดการทำงานต่างๆ (Module) ซึ่งเรียกว่า SPSS Family of Product ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น เมื่อผู้ผู้ใดให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1. ชุด BASE System
2. ชุด Professional Statistics
3. ชุด Advanced Statistics
4. ชุด Tables
5. ชุด Trends
6. ชุด Categories
7. ชุด CHAID
8. ชุด LISREL
9. ชุด Exact Tests
10. ชุด Developers Kit
11. ชุด Mapinfo
12. ชุด Teleform

ความสามารถของโปรแกรม SPSS for Windows [4]

1. ความสามารถในการวิเคราะห์ข้อมูล

ความสามารถในการวิเคราะห์ข้อมูล เป็นความสามารถในการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติดังต่อไปนี้

1.1 การคำนวณค่าสถิติเบื้องต้น (Descriptive Statistics) สามารถคำนวณค่าสถิติพื้นฐาน เช่น ค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) ค่าฐานนิยม (Mode) ค่าพิสัย (Range) ค่าความแปรปรวน (Variance) และค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เป็นต้น

1.2 การแจกแจงความถี่ (Frequency Distributions) สามารถทำการแจกแจงค่าของตัวแปรตามจำนวนที่นับได้ทั้งแบบทางเดียวและแบบหลายทาง (Crosstabs) พร้อมทั้งแสดงค่าสถิติที่เกี่ยวข้อง เช่นค่าเฉลี่ย (Mean) ค่ามัธยฐาน (Median) ค่าฐานนิยม (Mode) ค่าพิสัย (Range) ค่าความแปรปรวน (Variance) ค่าส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เปอร์เซ็นไทล์ (Percentiles) กราฟแท่งหรือค่าสถิติที่เกี่ยวข้องกับการทดสอบทางสถิติ เช่น Chi-Squares, Phi เป็นต้น

1.3 การเปรียบเทียบค่าเฉลี่ย (Mean Groups Comparison) สามารถเปรียบเทียบและทำการทดสอบค่าเฉลี่ยระหว่างกลุ่ม 2 กลุ่มตัวอย่างโดยค่าสถิติ t (Student's t) และสำหรับหลายกลุ่มตัวอย่างโดยค่าสถิติ F ด้วยการวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA) ทั้งแบบทางเดียวและแบบ

หลายทาง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4 การหาความสัมพันธ์ระหว่างตัวแปร (Correlation) สามารถคำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรแบบต่างๆ เช่น Pearson, Kendall, Spearman เป็นต้น

1.5 การวิเคราะห์การถดถอย (Regression Analysis) สามารถทำการหาความสัมพันธ์เพื่อการพยากรณ์แบบการถดถอยเชิงเส้น (Linear Regression Analysis) ทั้ง 1 ตัวแปรอิสระและตัวแปรอิสระมากกว่า 1 ตัว และสามารถดูรูปแบบความสัมพันธ์ ในลักษณะอื่นที่ไม่ใช่เส้นตรง เช่น Linear, Quadratic, Logarithmic เป็นต้น

1.6 การทดสอบแบบนอนพาราเมตริก (Nonparametric Test) สามารถวิเคราะห์ข้อมูลโดยวิธีของนอนพาราเมตริกสำหรับการทดสอบแบบต่างๆ เช่น Sign Test, Wilcoxon signed-rank test, Friedman test เป็นต้น

1.7 การวิเคราะห์ข้อมูลแบบหลายคำตอบ (Multiple Response Analysis) สามารถทำการวิเคราะห์ข้อมูลจากแบบสอบถามที่มีตัวเลือกมาให้และผู้ตอบสามารถตอบได้มากกว่า 1 คำตอบ

2. ความสามารถในการนำเสนอด้วยกราฟ

โปรแกรม SPSS for Windows สามารถนำเสนอข้อมูลในรูปแบบของกราฟหรือตารางแบบต่างๆ เช่น กราฟแท่ง (Bar, Histogram) กราฟเส้น (Line) กราฟวงกลม (Pie) และกราฟชนิดอื่นๆ

3. ความสามารถในการทำงานด้านอื่นๆ

ในการใช้งานโปรแกรม SPSS for Windows นอกจากจะทำการวิเคราะห์ข้อมูลด้วยวิธีการทางสถิติแล้ว ผู้ใช้อาจจะมีการดำเนินการกับข้อมูลในลักษณะต่างๆ เช่น การสร้างตัวแปรเพิ่ม เรียงลำดับข้อมูล คัดเลือกข้อมูลมาทำการวิเคราะห์ ฯลฯ ซึ่งสามารถแบ่งเป็นประเภทต่างๆ ได้ดังนี้

3.1 การเปลี่ยนรูปแบบข้อมูล (Data Transformation) โดยการเปลี่ยนค่าใหม่ จัดค่าใหม่ หรือสร้างตัวแปรใหม่ด้วยฟังก์ชันพิเศษต่างๆ ทางคณิตศาสตร์ที่มีใน โปรแกรม SPSS

3.2 การจัดกลุ่มตัวแปร (Define Set of Variable) โดยการเลือกตัวแปรหรือจัดกลุ่มตัวแปรไว้เป็นชุดต่างๆ เพื่อนำมาวิเคราะห์เป็นชุดๆ ในภายหลัง

3.3 การเลือกข้อมูล (Select Case) โดยการกำหนดเงื่อนไขต่างๆ หรือการเลือกข้อมูลแบบสุ่มตัวอย่าง

3.4 การสร้างข้อมูลแบบอนุกรมเวลา (Create Time Series) โดยการสร้างข้อมูลที่เกิดขึ้นตามเวลา เช่น วัน เดือน ไตรมาส ฯลฯ สำหรับการวิเคราะห์แบบอนุกรมเวลา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 การดำเนินการกับข้อมูลในลักษณะอื่นๆ โดยการเรียงลำดับข้อมูล การให้น้ำหนักหรือ ความสำคัญแก่ชุดข้อมูล การสลับที่ข้อมูลระหว่างแถวและคอลัมน์

3.6 การจัดการกับข้อมูล โดยการรวมเพิ่มข้อมูลตั้งแต่ 2 แฟ้ม เช่น รวมตัวแปร รวมชุด ข้อมูล

4. ความสามารถในการเชื่อมโยงข้อมูลกับโปรแกรมอื่นๆ

การทำงานของโปรแกรม SPSS for Windows version 12 เป็นการทำงานภายใต้ระบบการ ปฏิบัติงาน Windows ดังนั้นสามารถใช้ความสามารถขั้นพื้นฐาน เช่น การเลือกบริเวณเพื่อคัดลอก(copy), ตัดทิ้ง(cut) หรือวาง(paste) ฯลฯ แล้วนำข้อมูลนั้นไปใช้กับโปรแกรมอื่นๆ เช่น Excel, Microsoft Word, MathCAD หรือนำเอาข้อมูลจาก Excel, Microsoft Word, MathCAD มาใช้กับ SPSS for Windows ตัวอย่างเช่น ข้อมูลในรูปแบบของคอลัมน์ (Column) สามารถนำมาเป็นข้อมูลในรูปแบบของตัวแปร SPSS for Windows ได้ หรือข้อมูลที่วิเคราะห์ได้จาก SPSS for Windows สามารถ copy รูปแบบตารางไปเป็น ตารางของ Microsoft Word ได้ทันที

ข้อดีของโปรแกรม SPSS

1. โปรแกรมเป็นที่รู้จักแพร่หลาย นิยมใช้งานทางด้านสังคมศาสตร์เป็นส่วนใหญ่ และยังนิยม ใช้ทางด้านการแพทย์และด้านสาธารณสุข ในกรณีที่ไม่ต้องใช้สถิติขั้นสูงสามารถหาหนังสืออ่านได้ง่าย กว่า STATA
2. โปรแกรมใช้งานง่าย สามารถเลือกคำสั่งได้จากเมนูได้ โดยเฉพาะผู้ที่ไม่ได้เชี่ยวชาญทางด้าน สถิติมากนัก และสามารถบันทึกเป็น syntax สำหรับเก็บคำสั่งไว้ใช้ภายหลังได้
3. dataset ใน SPSS ไม่จำกัดว่าสร้างมาจาก version ไหน เพราะสามารถที่จะเปิดได้ทุกเวอร์ชัน เช่น สร้าง dataset มาจากเวอร์ชัน 12 สามารถที่จะนำ dataset มาเปิดที่เวอร์ชัน 10 ได้

ข้อเสียของโปรแกรม SPSS

1. ในกรณีที่เป็นสถิติขั้นสูง SPSS จะมีปัญหาไม่สามารถวิเคราะห์ได้ เช่น Survival
2. อาจเกิดปัญหา garbage in garbage out หมายความว่า ถ้าไม่สามารถแยกแยะได้ว่าตัวไหน เป็นตัวแปรต้น ตัวแปรตาม หรือตัวแปรชนิดใดได้แน่ชัด แล้วใส่ตัวแปรนั้นเข้าไป โปรแกรมก็จะคำนวณ ออกมาตามข้อมูลที่เราใส่เข้าไป ทำให้ผลการวิเคราะห์ ข้อมูลผิดพลาด ตามมาด้วยการอธิบายผลผิดอีก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.2 โปรแกรมสำเร็จรูป Minitab[3]

โปรแกรม Minitab เป็นโปรแกรมสำเร็จรูปในการนำมาใช้ประมวลผลข้อมูลทางด้านสถิติ โดยพัฒนาจากกลุ่มนักวิชาการทางด้านสถิติมากกว่า 30 ปีแล้ว ซึ่งสามารถใช้กับระบบปฏิบัติการวินโดวส์ โดย Minitab เข้ามามีบทบาทสำหรับผู้สถิติในส่วนของ การประมวลผล และการแสดงผลข้อมูลใน ลักษณะของตัวเลขและกราฟ ประกอบกับเทคโนโลยีทางด้านคอมพิวเตอร์ได้พัฒนาและมีบทบาทต่อ ชีวิตประจำวัน ดังนั้น Minitab จึงถูกเลือกใช้ด้วยเหตุผลหลัก 3 ประการ คือ

1. ความซับซ้อนในการประมวลผล (Complexity)
2. ความเที่ยงตรงและแม่นยำในการประมวลผลข้อมูล (Accuracy)
3. ความรวดเร็วและความสามารถในการทำซ้ำ (Repeatability)

โปรแกรม Minitab เป็นโปรแกรมที่มีความโดดเด่นในด้านการใช้งานที่ง่าย และยังมี การพัฒนาปรับปรุงฟังก์ชันต่างๆ ให้สอดคล้องกับความรู้และทฤษฎีใหม่ๆ รวมถึงการประยุกต์ทางด้านสถิติ โดยเฉพาะในงานด้านคุณภาพอย่างต่อเนื่อง ดังนั้น โปรแกรม Minitab จึงเป็นที่รู้จักกันเป็นอย่างดีสำหรับ กลุ่มผู้ที่พัฒนาปรับปรุงคุณภาพด้วยหลักการ “ซิกซ์ ซิกมา” เนื่องจากโปรแกรม Minitab เป็นเครื่องมือที่ สำคัญสำหรับกลุ่มผู้ใช้งานนี้ แต่ไม่ได้หมายความว่า จะมีเฉพาะกลุ่มผู้ใช้งานนี้เท่านั้นที่เข้าถึง โปรแกรม Minitab ยังเป็นโปรแกรมที่นิยมใช้กันอย่างแพร่หลายในกลุ่มนักวิชาการ นักวิจัย นักวิทยาศาสตร์ และ ผู้ใช้สถิติทั่วไป ถึงแม้ว่าโปรแกรม Minitab จะเป็นโปรแกรมที่ช่วยให้การทำงานสะดวกขึ้นมาก แต่ ความรู้ความเข้าใจในการทำงานของโปรแกรมก็ยังเป็นสิ่งสำคัญสำหรับผู้ใช้งาน

ขั้นตอนการใช้งานโปรแกรม MINITAB[5]

ขั้นตอนที่ 1 การเปิดโปรแกรมหรือเริ่มต้นเข้าสู่การทำงานบนแฟ้มงาน

ขั้นตอนที่ 2 ป้อนข้อมูลเข้าสู่เวิร์คชีท เนื่องจากโปรแกรม Minitab เป็นโปรแกรมสำหรับ การวิเคราะห์และประมวลผลข้อมูล ดังนั้น เมื่อเราเปิดแฟ้มงานแล้ว สิ่งที่เป็นมากสำหรับการใช้งาน โปรแกรมคือ ข้อมูล ซึ่งจะเป็นการป้อนจากแป้นพิมพ์ การนำข้อมูลจากฐานข้อมูลชนิดอื่น เช่น Excel, Access หรือ Text เป็นต้น หรือการเปิดจากไฟล์เวิร์คชีทของโปรแกรม Minitab ซึ่งขึ้นอยู่กับผู้ใช้งาน

ขั้นตอนที่ 3 เป็นส่วนการจัดการข้อมูล ซึ่งเป็นขั้นตอนที่สามารถข้ามไปยังขั้นตอนต่อไป ได้ ถ้าหากว่าข้อมูลที่อยู่ในเวิร์คชีทมีความพร้อมและอยู่ในรูปแบบที่ถูกต้อง สำหรับการวิเคราะห์โดยใช้ โปรแกรม Minitab แต่ถ้าหากผู้ใช้งานไม่จำเป็นต้องจัดการข้อมูลดังกล่าว อาจใช้ฟังก์ชันที่มีในโปรแกรม Minitab หรือจัดการจากโปรแกรมอื่นก่อนนำเข้าสู่เวิร์คชีทในโปรแกรม Minitab ก็ได้ ซึ่งจะขึ้นอยู่กับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัญหาและประสบการณ์ในการจัดการข้อมูลของผู้ใช้งาน โปรแกรม ตัวอย่างขั้นตอนนี้ คือ การรวมข้อมูล (Stack) การเปลี่ยนทิศทางการเรียงข้อมูล (Transpose Data) การคำนวณข้อมูล (Calculate) การสร้างชุดข้อมูลย่อย (Subset) เป็นต้น

ขั้นตอนที่ 4 เป็นขั้นตอนการวิเคราะห์ทางสถิติ โดยผู้ใช้งานจะเลือกฟังก์ชันหรือตัวสถิติที่เราจะทำการวิเคราะห์ เช่น การวิเคราะห์ความแปรปรวน (ANOVA) การออกแบบการทดลอง (DOE) เป็นต้น

ขั้นตอนที่ 5 เป็นส่วนในรายละเอียดของการวิเคราะห์ใดๆ ที่ผู้ใช้งานเลือก โดยปกติแล้ว ถ้าผู้ใช้งานใช้การตั้งค่าเริ่มต้นของโปรแกรม(Default) ซึ่งเป็นค่าที่มีการใช้ทั่วไปแล้ว เมื่อผู้ใช้งานใส่ข้อมูลครบถ้วนตามที่โปรแกรมต้องการแล้ว ผู้ใช้งานสามารถข้ามขั้นตอนนี้ไปได้ แต่ในบางครั้งผู้ใช้งานต้องการปรับเปลี่ยนค่าดังกล่าว เช่น ค่าความเชื่อมั่น 95% เป็น 99% การสั่งสร้างกราฟ หรือจะเป็นการสั่งให้โปรแกรมแสดงค่าหรือเก็บค่าใดๆ ที่ผู้ใช้งานต้องการ เป็นต้น

ขั้นตอนที่ 6 เป็นการจัดการกับผลการประมวลผลที่โปรแกรม Minitab ดำเนินการให้ ซึ่งจะออกมาใน 2 รูปแบบคือ ผลทางตัวเลขหรือตัวอักษร และกราฟ ตัวอย่างของการทำงานในขั้นตอนนี้ เช่น การสร้างรายงาน ตัดข้อมูลผิดปกติ การเปลี่ยนสีกราฟ เป็นต้น

จากขั้นตอนโดยรวมทั้งหมดที่กล่าวมานี้ โดยปกติผู้ใช้งานสามารถที่จะทำซ้ำหรือการวนหรือย้อนกลับขั้นตอนต่างๆ ได้ตลอดเวลา トラバิดที่เรากำลังทำงานบนแท็บงาน

ความสามารถด้านฟังก์ชันทางสถิติและกราฟ

1. Basic Statistics เป็นฟังก์ชันพื้นฐานทางด้านสถิติ ซึ่งประกอบด้วยสถิติเชิงพรรณนา (descriptive) และสถิติเชิงอนุมาน (inferential) เช่น การหาค่าพารามิเตอร์ทางสถิติ เช่น ค่าเฉลี่ย (mean), ค่าความแปรปรวน (variance), ค่าพิสัย (range) เป็นต้น นอกจากนี้ยังมีชุดคำสั่งในการหาช่วงความเชื่อมั่น (confidence interval) และการทดสอบสมมติฐาน (hypothesis testing) โดยผลการคำนวณจะให้ทั้งผลลัพธ์บน Session และกราฟ

2. Regression Analysis เป็นฟังก์ชันในการวิเคราะห์การถดถอยเชิงเส้นได้ทั้งรูปแบบสมการเส้นตรง (linear) และสมการกำลัง (polynomial) หรือรูปแบบอื่นๆ ที่ต้องการ รวมถึงการเก็บค่าเศษเหลือ (residual) และกราฟประกอบการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. **ANOVA** เป็นฟังก์ชันการวิเคราะห์ความแปรปรวน โดยสามารถวิเคราะห์ปัจจัยตั้งแต่ 1 ปัจจัย (one-way) 2 ปัจจัย(two-way) หรือมากกว่า 2 ปัจจัย รวมถึงการแสดงผลกราฟปัจจัยอิทธิพลหลัก (main effect plot) และปัจจัยอิทธิพลร่วม (interaction effect plot) และอื่นๆ ที่เกี่ยวข้อง

4. **Statistical Quality Tools** เป็นฟังก์ชันที่สนับสนุนงานทางการวิเคราะห์ด้านคุณภาพ ประกอบด้วยหัวเรื่องหลัก 4 เรื่อง คือ

- **Quality Tools** สามารถทำการสร้างกราฟพาร์โต (pareto), รันชาร์ต(run chart) เป็นต้น
- **Control Charts** ชุดคำสั่งสร้างแผนภูมิควบคุมมากมาย หลากหลาย ครอบคลุมข้อมูลทุกประเภท รวมถึงคำสั่งเพิ่มเติมช่วยในการวิเคราะห์ข้อมูลผิดปกติ
- **Capability Analysis** เป็นคำสั่งวิเคราะห์ความสามารถกระบวนการ โดยการนำเอาความผันแปรข้อมูลเทียบกับข้อกำหนด (specification) ซึ่งสามารถจัดการได้ทั้งกรณีข้อมูลเป็นปกติ (normal)หรือไม่ปกติ (non-normal)
- **Measurement System** เป็นคำสั่งวิเคราะห์และประเมินความสามารถกระบวนการวัด เช่น Stability Bias Linearity และ Gage R&R เป็นต้น

5. **Design of Experiment** เป็นชุดคำสั่งที่ช่วยในการออกแบบการทดลองประกอบด้วย factorial, response surface, mixture และ taguchi สำหรับคำสั่งการออกแบบการทดลองในโปรแกรม Minitab จะช่วยเหลือตั้งแต่การออกแบบการทดลอง การจัดเก็บข้อมูล และการวิเคราะห์ผลที่ได้และกราฟประกอบการแปลผล รวมถึงการตั้งค่าเพื่อผลลัพธ์ที่ต้องการ (response optimizer)

6. **Reliability** เป็นฟังก์ชันสำหรับวิเคราะห์ค่าความน่าเชื่อถือของผลิตภัณฑ์ โดยอาศัยหลักการทางสถิติ โดยมีฟังก์ชันช่วยในการหาฟังก์ชันความน่าจะเป็น (fit distribution) แบบต่างๆ และทำการวิเคราะห์ผ่านฟังก์ชัน โดยอาศัยหลักการความน่าจะเป็น

7. **Power and Sample Size** เป็นฟังก์ชันเพื่อช่วยในการหาขนาดทดสอบ (sample size) หรือความเชื่อมั่นในการทดสอบ (power) สำหรับการทดสอบสมมติฐานแต่ละแบบ

8. **Multivariate Analysis** เป็นฟังก์ชันสำหรับการวิเคราะห์ตัวแปรหลายตัวแปร เช่น คำสั่ง Principal Component, factor analysis, cluster analysis เป็นต้น

9. **Time Series and Forecasting** เป็นฟังก์ชันการวิเคราะห์ข้อมูลแปรผันตามเวลา เพื่อดูแนวโน้ม ลักษณะกราฟหรือทิศทาง เพื่อใช้ในการทำนาย

10. **Nonparametric** เป็นฟังก์ชันการวิเคราะห์ที่ไม่ใช้พารามิเตอร์

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเท่านั้น เมื่อผู้ใช้ได้พิมพ์เอกสารนี้โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

11. Tables เป็นฟังก์ชันที่ใช้ในการจัดการข้อมูลที่เป็นอยู่ในรูปแบบสำหรับตารางของข้อมูล นับจำนวนความถี่ พร้อมคำสั่งการวิเคราะห์ เช่น การทดสอบ Chi-square

12. Simulation and Distribution เป็นคำสั่งช่วยในการสุ่มชุดข้อมูลผ่านฟังก์ชันความน่าจะเป็น (distribution) เพื่อประโยชน์ในการทำการทดสอบโมเดลหรือการวิเคราะห์ที่ต้องการ

13. EDA (explore data analysis) เป็นหมวดฟังก์ชันในการประเมินข้อมูล เพื่อดูลักษณะการกระจายของข้อมูล เช่น คำสั่ง Stem-and leaf หรือ boxplot เป็นต้น

2.1.3 โปรแกรมสำเร็จรูป Microsoft Excel[6]

โปรแกรม Microsoft Excel เป็นโปรแกรมที่อยู่ในรูปกระดาษทำการ (Spread Sheet) ใช้กับระบบปฏิบัติการ Windows สามารถบันทึกข้อความลงในตารางแล้วคำนวณค่าต่างๆตลอดจนค่าสถิติ ทำกราฟ โปรแกรมสำเร็จรูป Microsoft Excel เป็นโปรแกรมหนึ่งในชุดของ Microsoft Office ที่ผลิตโดยบริษัท ไมโครซอฟท์ จำกัด ผู้ใช้จะเลือกใช้งานจากแถบของเมนู โปรแกรม Microsoft Excel สามารถใช้งานได้ดัง

1. คำนวณค่าสถิติเบื้องต้น เช่น ค่ากลาง หาค่าสูงสุด หาค่าต่ำสุด
2. การกระจายของข้อมูล
3. การวิเคราะห์ความถดถอย
4. การหาค่าสหสัมพันธ์
5. การนำเสนอข้อมูลในรูปแบบของตารางและกราฟแบบต่างๆ

คุณสมบัติที่สำคัญในโปรแกรม Excel

1. ความสามารถด้านการคำนวณ โปรแกรม Excel สามารถป้อนสูตรด้านการคำนวณทางคณิตศาสตร์ เช่น บวก ลบ คูณ หาร เป็นต้น รวมทั้งสูตรคำนวณด้านอื่นๆ และจุดเด่นของการคำนวณคือผลลัพธ์ของการคำนวณจะเปลี่ยนแปลงตามเมื่อข้อมูลอินพุตที่นำมาเปลี่ยนค่า ทำให้เราไม่ต้องเสียเวลาเปลี่ยนแปลงค่าผลการคำนวณใหม่

2. ความสามารถด้านการใช้ฟังก์ชัน นอกจากในการป้อนสูตรทางคณิตศาสตร์แล้ว โปรแกรม Excel ยังสามารถป้อนฟังก์ชันอื่นๆ ได้อีก เช่น ฟังก์ชันเกี่ยวกับตัวอักษร ตัวเลข วันที่ ฟังก์ชันเกี่ยวกับการเงินหรือการตัดสินใจ

3. ความสามารถในการสร้างกราฟ โปรแกรม Excel สามารถนำข้อมูลที่ป้อนลงในตาราง มาสร้างเป็นกราฟได้ทันที มีรูปภาพให้เลือกใช้งานหลายรูปแบบตามความเหมาะสม เช่น กราฟแท่ง แสดงยอดขายแต่ละเดือน กราฟวงกลม แสดงส่วนแบ่งการตลาด เป็นต้น

4. ความสามารถในการตกแต่งตารางข้อมูล โปรแกรม Excel สามารถตกแต่งตารางข้อมูล หรือกราฟข้อมูลด้วยภาพสีและรูปแบบตัวอักษรต่างๆ เพื่อให้เกิดความสวยงามและแยกแยะข้อมูลได้ง่ายขึ้น

5. ความสามารถในการจัดเรียงลำดับ โปรแกรม Excel ยังสามารถคัดเลือกเฉพาะข้อมูลที่ต้องการมาวิเคราะห์ได้

6. ความสามารถในการพิมพ์งานออกทางเครื่องพิมพ์ โปรแกรม Excel สามารถพิมพ์งาน ทั้งข้อมูลและรูปภาพหรือกราฟออกทางเครื่องพิมพ์ได้ทันที ซึ่งทำให้ง่ายต่อการสร้างรายงาน

7. ความสามารถในการแปลงข้อมูลในตารางให้เป็นเว็บเพจ เพื่อนำมาแสดงในโฮมเพจ

ประโยชน์ของโปรแกรม Microsoft Excel[7]

1. สร้างตารางทำงาน จัดตารางสวยงาม ในรูปแบบต่างๆ
2. สร้างเอกสารที่ต้องมีการคำนวณ และยังสามารถเชื่อมโยงสูตร หมายความว่าสามารถเชื่อมโยงในไฟล์เดียวกันข้ามไฟล์ หรือข้ามเครื่องก็ได้
3. งานจัดเก็บข้อมูลเบื้องต้นที่มีจำนวนข้อมูลไม่เกิน 1 ล้านแถว ซึ่งในทางปฏิบัติ แนะนำว่าไม่ควรเกินหลักแสน เพราะจะทำงานได้คล่องตัวกว่า
4. สร้างรายงานสรุปผลในมุมมองต่างๆ เช่น ตารางสรุปยอดขาย ตารางสรุปข้อมูลสินค้า สรุปงบดุล สรุปแผนการผลิต สรุปข้อมูล ขาดลามาสาย ของพนักงาน เป็นต้น
5. สร้างกราฟ นำเสนอข้อมูล ในรูปแบบต่างๆ ไม่ว่าจะเป็นกราฟแท่ง กราฟเส้น วงกลม จุด ทั้ง 2 มิติ และ 3 มิติ มีรูปแบบต่างๆ มากมาย

ลักษณะงานที่เหมาะสมกับโปรแกรม Microsoft Excel

- งานด้านบัญชี (Accounting)
- งานด้านการเงิน (Financial)
- งานด้านการวางแผน (Planning)
- งานด้านงบประมาณ (Budgeting)
- งานด้านสถิติ (Statistic)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- งานด้านวิศวกรรมศาสตร์ (Engineering)

2.2 การวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Analysis) [8]

ในกรณีนี้ จะทำการศึกษาตัวแปรตามและตัวแปรอิสระเพียงตัวเดียวในรูปเส้นตรง ถ้ากำหนดให้ Y เป็นตัวแปรตาม และ X เป็นตัวแปรอิสระ จะศึกษาตัวแปร X มีอิทธิพลต่อตัวแปร Y มากน้อยเพียงใดในเชิงเส้นตรง และเรียกลักษณะนี้ว่า การถดถอยเชิงเส้นอย่างง่าย

ถ้ากำหนดให้ ε (Epsilon) เป็นผลต่างระหว่างค่า Y ที่แท้จริงกับค่า $\mu_{Y,X}$ จะได้ว่า

$$\varepsilon_i = Y_i - \mu_{Y,X}$$

$$\begin{aligned} \text{หรือ } Y_i &= \mu_{Y,X} + \varepsilon_i \\ &= \beta_0 + \beta_1 X_i + \varepsilon_i \end{aligned}$$

ดังนั้น ค่า ε_i จะเป็นค่าความคลาดเคลื่อน (Error Term) และมีข้อสมมติเบื้องต้นดังนี้

1) ผลรวมของ ε_i มีค่าเท่ากับ 0

$$\sum \varepsilon_i = 0 \quad \text{หรือ} \quad E(\varepsilon_i) = 0$$

2) ค่าความแปรปรวนของ ε_i ในแต่ละประชากรย่อยจะต้องเท่ากัน

$$V(\varepsilon_i) = \sigma^2$$

3) ค่าของ ε_i จะต้องมีการแจกแจงแบบปกติ (Normal Distribution)

$$\varepsilon_i \sim N(0, \sigma^2)$$

4) ค่าของ ε_i กับ ε_j จะต้องเป็นอิสระกัน เมื่อ $i \neq j$

แต่ในทางปฏิบัติจะไม่สามารถศึกษาข้อมูลประชากรได้ จึงทำการศึกษาข้อมูลที่ได้จากกลุ่มตัวอย่างเท่านั้น ดังนั้น ตัวแบบการถดถอยเชิงเส้นแบบง่ายของข้อมูลตัวอย่าง (Sample Simple Linear Regression) คือ

$$Y_i = b_0 + b_1 X_i + e_i \quad ; \quad i = 1, 2, \dots, n$$

และสมการการถดถอยเชิงเส้นแบบง่ายของข้อมูลตัวอย่าง (Sample Simple Linear Regression Equation) คือ

$$\hat{Y}_i = b_0 + b_1 X_i$$

เมื่อ b_0 คือ ค่าที่เส้นการถดถอยตัดแกน Y (Y -intercept)

b_1 คือ ค่าความชัน (Slope) ของเส้นการถดถอย และเรียก b_1 ว่า สัมประสิทธิ์ของการถดถอย (Regression Coefficient)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.1 การประมาณค่าพารามิเตอร์โดยวิธีกำลังสองน้อยที่สุด (Least Square Method : LSM)

วิธี LSM เป็นวิธีการที่ใช้หาค่า b_0 , b_1 ของสมการ $\hat{Y}_i = b_0 + b_1X_i$ ที่ให้คุณสมบัติความไม่เอนเอียงและความแปรปรวนต่ำสุด ตามที่กล่าวมาข้างต้น โดยอาศัยหลักการที่ว่า

1. $\sum e_i = 0$
2. $\sum e_i^2$ มีค่าน้อยที่สุด

ดังนั้น ในการหาค่า b_0 , b_1 จะมีขั้นตอนดังนี้

ขั้นที่ 1 สร้างสมการปกติ (Normal Equation)

จากตัวแบบ

$$Y_i = b_0 + b_1X_i + e_i$$

$$\therefore e_i = Y_i - b_0 - b_1X_i$$

อาศัยคุณสมบัติของวิธีกำลังสองน้อยที่สุด โดยพิจารณาจากผลบวกกำลังสองของค่าความคลาดเคลื่อนจะต้องมีค่าน้อยที่สุด จะได้

$$e_i = Y_i - b_0 - b_1X_i$$

$$e_i^2 = (Y_i - b_0 - b_1X_i)^2$$

$$\sum e_i^2 = \sum (Y_i - b_0 - b_1X_i)^2 \quad \text{จะต้องมีค่าน้อยที่สุด}$$

โดยอาศัยความรู้ทางด้านคณิตศาสตร์ ด้วยการดิฟเฟอเรนทิเอท (Differentiate) ค่า $\sum e_i^2$ เทียบกับ b_0 , b_1 แล้วกำหนดให้เท่ากับ 0 ก็จะได้คุณสมบัติตามที่ต้องการ ดังนี้

$$\frac{\partial \sum e_i^2}{\partial b_0} = \frac{\partial \sum (Y_i - b_0 - b_1X_i)^2}{\partial b_0} = 0$$

$$\text{จะได้} \quad -2\sum (Y_i - b_0 - b_1X_i) = 0$$

$$\sum Y_i - \sum b_0 - \sum b_1X_i = 0$$

$$\sum Y_i - nb_0 - b_1\sum X_i = 0$$

$$\therefore \quad \sum Y_i = nb_0 + b_1\sum X_i \quad \dots\dots\dots(1)$$

$$\frac{\partial \sum e_i^2}{\partial b_1} = \frac{\partial \sum (Y_i - b_0 - b_1X_i)^2}{\partial b_1} = 0$$

$$\text{จะได้} \quad -2\sum X_i(Y_i - b_0 - b_1X_i) = 0$$

$$\sum X_i Y_i - \sum b_0 X_i - \sum b_1 X_i^2 = 0$$

$$\sum X_i Y_i - b_0\sum X_i - b_1\sum X_i^2 = 0$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$\therefore \sum x_i Y_i = b_0 \sum x_i + b_1 \sum x_i^2 \dots\dots\dots(2)$$

ดังนั้น สมการปกติ คือ

$$\sum Y_i = nb_0 + b_1 \sum X_i$$

$$\sum x_i Y_i = b_0 \sum x_i + b_1 \sum x_i^2$$

ขั้นที่ 2 จากสมการปกติ นำมาคำนวณหาค่า b_0 , b_1 โดยวิธีพีชคณิต จะได้

$$b_1 = \frac{\sum(x_i - \bar{x})(y - \bar{y})}{\sum(x_i - \bar{x})^2}$$

หรือ
$$b_1 = \frac{n \sum x_i Y_i - \sum x_i \sum Y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

และ
$$b_0 = \bar{Y} - b_1 \bar{X}$$

2.2.2 คุณสมบัติของ b_0 และ b_1

- 1) b_0 และ b_1 เป็นตัวประมาณค่าที่ไม่เอนเอียงของ β_0 และ β_1 ตามลำดับ
- 2) ความแปรปรวนของ b_0 และ b_1 คือ $V(b_0)$ และ $V(b_1)$ มีค่าดังนี้

$$V(b_1) = \frac{\sigma^2}{\sum(x_i - \bar{x})^2}$$

$$V(b_0) = \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2} \right)$$

เนื่องจาก σ^2 จะเป็นค่าที่ไม่ทราบค่า จึงประมาณค่า σ^2 ด้วย $S_{Y.X}^2$ จะได้ค่าประมาณของ $V(b_0)$ และ $V(b_1)$ ดังนี้

$$\hat{V}(b_0) = S_{b_0}^2 = S_{Y.X}^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{\sum(x_i - \bar{x})^2} \right)$$

และ
$$\hat{V}(b_1) = S_{b_1}^2 = \frac{S_{Y.X}^2}{\sum(x_i - \bar{x})^2}$$

เมื่อ
$$S_{Y.X}^2 = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}$$

ในการหาค่า $S_{Y.X}^2$ ยังสามารถหาได้จากความสัมพันธ์ ดังนี้

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า $\sum(Y_i - \bar{Y})^2$ เรียกว่า ผลบวกกำลังสองของการกระจายทั้งหมด (Total Sum Square)
เขียนแทนด้วย SST

ค่า $\sum(\hat{Y}_i - \bar{Y})^2$ เรียกว่า ผลบวกกำลังสองของการถดถอย (Regression Sum Square)
เขียนแทนด้วย SSR

ค่า $\sum(Y_i - \hat{Y}_i)^2$ เรียกว่า ผลบวกกำลังสองของความคลาดเคลื่อน (Error Sum Square)
เขียนแทนด้วย SSE

$$\text{จาก } \sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

$$\text{SST} = \text{SSR} + \text{SSE}$$

$$\therefore \text{SSE} = \text{SST} - \text{SSR}$$

$$\text{และ } S_{Y.X}^2 = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}$$

$$= \frac{\text{SSE}}{n-2}$$

$$= \frac{\text{SST} - \text{SSR}}{n-2}$$

2.2.3 การวิเคราะห์การถดถอยด้วยการวิเคราะห์ความแปรปรวน

จากการกระจายของข้อมูลต่างๆ ออกเป็นหลายส่วน ดังนี้

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2$$

$$\text{SST} = \text{SSR} + \text{SSE}$$

สามารถนำมาเขียนในรูปตารางการวิเคราะห์ความแปรปรวนได้ดังนี้

ตารางการวิเคราะห์ความแปรปรวน (ANOVA)

Source of Variation	df	SS	MS	F
ความถดถอย (Regression)	1	SSR	$MSR = \frac{SSR}{1}$	$F = \frac{MSR}{MSE}$
ความคลาดเคลื่อน (Error)	$n - 2$	SSE	$MSE = \frac{SSE}{n - 2}$	
Total	$n - 1$	SST		

สมมติฐานในการทดสอบ

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

ที่ระดับนัยสำคัญ α ค่าวิกฤต คือ $F_{\alpha(1, n-2)}$

สรุปผล : จะปฏิเสธสมมติฐาน H_0 เมื่อ $F > F_{\alpha(1, n-2)}$

2.3 การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

ในการหาความสัมพันธ์ของตัวแปรอิสระกับตัวแปรตามนั้น หากพบว่าจำนวนตัวแปรอิสระที่เราสนใจในการศึกษามีมากกว่าหนึ่งตัวแปร จึงจำเป็นต้องใช้การวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple linear regression analysis) ซึ่งเป็นเทคนิคในการวิเคราะห์การถดถอยที่เกี่ยวข้องกับตัวแปรอิสระที่มากกว่าหนึ่งตัวแปร ในการเพิ่มตัวแปรอิสระที่เกี่ยวข้องเข้าไปในการวิเคราะห์จะส่งผลให้ความถูกต้องของการวิเคราะห์เพิ่มมากขึ้น และค่าคลาดเคลื่อนมาตรฐานของตัวประมาณค่า (Standard error of estimates) ลดลง

2.3.1 รูปแบบของสมการถดถอยเชิงเส้นพหุคูณ

จากหัวข้อที่ผ่านมานั้น สมการถดถอยเชิงเส้นอย่างง่ายจะมีตัวแปรอิสระเพียงตัวเดียว และมีความสัมพันธ์ในเชิงเส้นตรง สำหรับกรณีที่มีตัวแปรอิสระมากกว่า 1 ตัว ที่มีความสัมพันธ์ในเชิงเส้นสามารถเขียนรูปแบบของสมการได้ดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

เมื่อ β_0 คือส่วนตัดแกน Y เมื่อกำหนดให้ $X_1 = X_2 = \dots = X_k = 0$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$\beta_1, \beta_2, \dots, \beta_k$ เป็นสัมประสิทธิ์การถดถอย

ซึ่ง β_i เป็นค่าที่แสดงถึงการเปลี่ยนแปลงของตัวแปรตาม Y เมื่อตัวแปรอิสระ X_i เปลี่ยนไป 1 หน่วย โดยที่ตัวแปรอิสระ X ตัวอื่นๆ มีค่าคงที่

สำหรับค่า ε คือค่าความคลาดเคลื่อน และมีข้อสมมติเบื้องต้นเช่นเดียวกับการถดถอยเชิงเส้นอย่างง่าย

2.3.2 การประมาณค่าพารามิเตอร์ของสมการถดถอยเชิงเส้นพหุ

ในที่นี้จะขอกล่าวเฉพาะตัวแปรอิสระเพียง 2 ตัวเท่านั้น เพื่อเป็นแนวทางในการวิเคราะห์สมการถดถอยเชิงเส้นพหุ และสามารถเขียนรูปแบบการถดถอยเชิงเส้นพหุของข้อมูลตัวอย่างได้ดังนี้

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + e_i$$

และสามารถประมาณค่า Y หรือ \hat{Y} ได้จาก

$$\hat{Y} = b_0 + b_1X_{1i} + b_2X_{2i}$$

วิธีกำลังสองน้อยที่สุดเป็นวิธีที่ถูกเลือกนำมาใช้เช่นเดียวกับการถดถอยเชิงเส้นแบบง่าย โดยอาศัยหลักการเดียวกันคือ

1. $\sum e_i = 0$

2. $\sum e_i^2$ มีค่าน้อยที่สุด

ดังนั้น ในการหาค่า b_0, b_1 และ b_2 จะมีขั้นตอนดังนี้

จากตัวแบบ

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + e_i$$

$$e_i = Y_i - \hat{Y}$$

$$= Y_i - b_0 - b_1X_{1i} - b_2X_{2i}$$

อาศัยคุณสมบัติของวิธีกำลังสองน้อยที่สุด โดยพิจารณาจากผลบวกกำลังสองของค่าความคลาดเคลื่อนจะต้องมีค่าน้อยที่สุด จะได้

$$e_i = Y_i - b_0 - b_1X_{1i} - b_2X_{2i}$$

$$e_i^2 = (Y_i - b_0 - b_1X_{1i} - b_2X_{2i})^2$$

$$\sum e_i^2 = \sum (Y_i - b_0 - b_1X_{1i} - b_2X_{2i})^2 \quad \text{จะต้องมีค่าน้อยที่สุด}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยอาศัยความรู้ทางด้านคณิตศาสตร์ ด้วยการดิฟเฟอเรนทิเอท (Differentiate) ค่า $\sum e_i^2$ เทียบกับ b_0 , b_1 , b_2 แล้วกำหนดให้เท่ากับ 0 ก็จะได้คุณสมบัติตามที่ต้องการ ดังนี้

$$\frac{\partial \sum e_i^2}{\partial b_0} = 0 \text{ จะได้ } \sum Y_i = nb_0 + b_1 \sum X_{1i} + b_2 \sum X_{2i}$$

$$\frac{\partial \sum e_i^2}{\partial b_1} = 0 \text{ จะได้ } \sum X_{1i}Y_i = b_0 \sum X_{1i} + b_1 \sum X_{1i}^2 + b_2 \sum X_{1i}X_{2i}$$

$$\frac{\partial \sum e_i^2}{\partial b_2} = 0 \text{ จะได้ } \sum X_{2i}Y_i = b_0 \sum X_{2i} + b_1 \sum X_{1i}X_{2i} + b_2 \sum X_{2i}^2$$

นั่นคือ ถ้ามีตัวแปรอิสระ 2 ตัวในสมการ จะได้สมการปกติจำนวน 3 สมการ ดังนี้

$$\sum Y_i = nb_0 + b_1 \sum X_{1i} + b_2 \sum X_{2i}$$

$$\sum X_{1i}Y_i = b_0 \sum X_{1i} + b_1 \sum X_{1i}^2 + b_2 \sum X_{1i}X_{2i}$$

$$\sum X_{2i}Y_i = b_0 \sum X_{2i} + b_1 \sum X_{1i}X_{2i} + b_2 \sum X_{2i}^2$$

นอกจากการหาค่า b_0 , b_1 และ b_2 ในรูปพีชคณิตแล้ว ยังสามารถหาอยู่ในรูปของเมทริกซ์ได้

2.3.3 การวิเคราะห์ความแปรปรวน

การวิเคราะห์ความแปรปรวนเป็นการทดสอบสัมประสิทธิ์การถดถอยทั้งหมดมีค่าเท่ากับ 0 หรือไม่ ในกรณีที่ตัวแปรอิสระมีเพียง 2 ตัว สมมติฐานที่ตั้งคือ

$$H_0 : \beta_1 = \beta_2 = 0$$

$$H_1 : \text{มี } \beta \text{ อย่างน้อย 1 ค่าที่ไม่เท่ากับ 0}$$

และใช้การทดสอบแบบ F เช่นเดียวกับการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย หากปฏิเสธสมมติฐานหลัก แสดงว่า มีตัวแปรอิสระอย่างน้อย 1 ตัวที่มีความสัมพันธ์เชิงเส้นตรงกับตัวแปรตาม สามารถวิเคราะห์ในรูปแบบของความแปรปรวน ได้ดังนี้

$$\text{จาก } SST = SSR + SSE$$

$$\text{เมื่อ } SST = \sum (Y_i - \bar{Y})^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

$$SSR = \sum (\hat{Y}_i - \bar{Y})^2 = b_1 \sum (X_{1i} - \bar{X}_1)(Y_i - \bar{Y}) + b_2 \sum (X_{2i} - \bar{X}_2)(Y_i - \bar{Y})$$

$$SSE = SST - SSR$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางวิเคราะห์ความแปรปรวนในกรณีตัวแปรอิสระ 2 ตัว สามารถเขียนได้ดังนี้

Source of variation	df	SS	MS	F
Regression	$k - 2$	SSR	MSR	$\frac{MSR}{MSE}$
Error	$n - k - 1$	SSE	MSE	
Total	$n - 1$	SST		

จะปฏิเสธสมมติฐาน H_0 เมื่อค่า $F > F_{\alpha(2,n-k-1)}$

2.4 ค่าวัดประสิทธิภาพของรูปแบบ[9]

เมื่อกำหนดรูปแบบการถดถอยให้กับข้อมูลตัวอย่างแล้ว จะทำการสร้างสมการถดถอยตามรูปแบบการถดถอยนั้น อย่างไรก็ตามรูปแบบการถดถอยที่กำหนดไว้ อาจจะเหมาะสมกับข้อมูลตัวอย่างหรือไม่ก็ได้ ดังนั้น ในการพิจารณาความเหมาะสมของรูปแบบการถดถอย จะได้จากค่าสถิติที่ใช้วัดประสิทธิภาพของรูปแบบและจากการทดสอบสมมติฐาน ค่าวัดประสิทธิภาพของรูปแบบมีอยู่หลายค่า เช่น ผลรวมกำลังสองของความคลาดเคลื่อน(SSE) ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน(MSE) ค่าสัมประสิทธิ์ตัวกำหนด(R^2) และค่าสัมประสิทธิ์ตัวกำหนดปรับแล้ว (R_a^2) เป็นต้น ค่าสถิติเหล่านี้จะใช้วัดประสิทธิภาพของรูปแบบมีรายละเอียดดังนี้

1. ค่าผลรวมกำลังสองของความคลาดเคลื่อน(Sum Square Error : SSE) เป็นค่าวัดที่ยังไม่มีเกณฑ์แน่นอนว่า รูปแบบที่เหมาะสมจะต้องมีค่า SSE เป็นเท่าใด แต่รูปแบบการถดถอยที่เหมาะสมควรจะเป็นรูปแบบที่มีค่า SSE น้อยที่สุด

2. ค่าเฉลี่ยกำลังสองของความคลาดเคลื่อน(Mean Square Error : MSE) เป็นค่าวัดที่เป็นฟังก์ชันของ SSE นั่นคือ เป็นค่า SSE ที่ปรับด้วยชั้นแห่งความเป็นอิสระ ซึ่ง $MSE = \frac{SSE}{n - k - 1}$ สำหรับ

รูปแบบการถดถอยที่มีจำนวนตัวแปรอิสระต่างกัน แต่มี SSE เท่ากัน รูปแบบที่มีจำนวนตัวแปรอิสระน้อยกว่าจะให้ค่า MSE ที่ต่ำกว่า ค่า MSE จากรูปแบบที่มีจำนวนตัวแปรอิสระมากกว่า นอกจากการใช้ค่า SSE และ MSE เพื่อวัดประสิทธิภาพของรูปแบบแล้ว ยังมีผู้ใช้ค่ารากที่สองของ MSE หรือ RMSE ซึ่ง $RMSE = \sqrt{MSE}$ ในการพิจารณาความเหมาะสมของรูปแบบ กรณีที่ใช้ค่า MSE หรือ RMSE ในการพิจารณารูปแบบที่เหมาะสม รูปแบบที่ให้ค่า MSE หรือ RMSE ต่ำที่สุดจะเป็นรูปแบบที่เหมาะสมที่สุด ซึ่งเป็นทำนองเดียวกับการใช้ค่า SSE

3. ค่าสัมประสิทธิ์ตัวกำหนด(Coefficient of determination : R^2) เป็นค่าสถิติที่ใช้ในการวัดค่าว่า ตัวแปรอิสระที่อยู่ในรูปแบบการถดถอยมีส่วนในการอธิบายความผันแปรรวม $\sum(Y_i - \bar{Y})^2$ มากน้อยเท่าใด รูปแบบที่เหมาะสมที่สุดจะเป็นรูปแบบที่ให้ค่า R^2 สูงสุด ค่า R^2 จะเป็นสัดส่วนของ SSR กับ SST เนื่องจาก $SST = SSR + SSE$ ทำให้ค่า R^2 จึงมีค่าอยู่ระหว่าง 0 ถึง 1

$$\begin{aligned} R^2 &= \frac{SSR}{SST} \\ &= 1 - \frac{SSE}{SST} \end{aligned}$$

โดยทั่วไปจะอธิบายค่าของ R^2 เป็นเปอร์เซ็นต์แทนการอธิบายด้วยสัดส่วน เช่น สำหรับรูปแบบการถดถอยที่มีตัวแปรอิสระ X_1 และ X_2 ที่มี $R^2 = 0.8921$ จะอธิบายได้ว่าตัวแปรอิสระ X_1 และ X_2 มีส่วนในการอธิบายความผันแปรรวม $\sum(Y_i - \bar{Y})^2$ ได้ดี 89.21 เปอร์เซ็นต์ เป็นต้น เนื่องจาก R^2 แปรผันกับ SSE ดังนั้นเมื่อ SSE มีค่าน้อย R^2 จะมีค่ามากหรือเข้าใกล้ 1 และเมื่อ SSE มีค่ามาก R^2 จะมีค่าน้อยหรือเข้าใกล้ 0

4. ค่าสัมประสิทธิ์ตัวกำหนดที่ปรับแล้ว (Adjusted coefficient of determination : R_a^2) เป็นค่าสถิติที่นำมาใช้วัดว่าตัวแปรอิสระที่อยู่ในรูปแบบการถดถอย มีส่วนในการอธิบายความแปรปรวน (S_y^2) มากน้อยเพียงใด รูปแบบที่เหมาะสมที่สุดจะเป็นรูปแบบที่ให้ค่า R_a^2 สูงสุด ค่า R_a^2 จะแตกต่างจากค่า R^2 ที่ค่า R_a^2 คำนึงถึงชั้นแห่งความเป็นอิสระของ SSE และ SST นั่นคือ จะพิจารณา MSE แทน SSE และ S_y^2 แทน SST

$$\begin{aligned} R_a^2 &= 1 - \frac{SSE / (n - k - 1)}{SST / (n - 1)} \\ &= 1 - \frac{MSE}{S_y^2} \end{aligned}$$

เมื่อ n มีขนาดใหญ่ ค่า R_a^2 จะใกล้เคียงกับค่า R^2 ในการเปรียบเทียบรูปแบบการถดถอยสองรูปแบบที่มีจำนวนตัวแปรอิสระต่างกันแต่มีค่า SSE เท่ากัน รูปแบบที่มีจำนวนตัวแปรอิสระมากจะมีค่า R_a^2 น้อยกว่าค่า R_a^2 ของรูปแบบที่มีจำนวนตัวแปรอิสระน้อย เนื่องจาก R_a^2 แปรผกผันกับ MSE หรือ R_a^2 เป็นฟังก์ชันของ MSE

2.5 การทดสอบความแปรปรวน (Analysis of Variance) [10]

การวิเคราะห์ความแปรปรวนจะใช้ทดสอบความแตกต่างของค่าเฉลี่ยประชากรตั้งแต่ 3 ประชากรขึ้นไป ซึ่งหลักเกณฑ์สำคัญที่ใช้ในการทดสอบสมมติฐาน คือ การแยกความแปรปรวนทั้งหมดของข้อมูลออกตามสาเหตุที่ทำให้ข้อมูลแตกต่างกัน นั่นคือ แยกตามความแปรปรวนทั้งหมดของข้อมูลออกเป็น

1. ความแปรปรวนหรือความแตกต่างระหว่างประชากร
2. ความแปรปรวนหรือความแตกต่างภายในประชากรเดียวกัน

โดยมีข้อตกลงเบื้องต้นคือ

1. ประชากรทั้ง k ประชากรมีการแจกแจงแบบปกติ
2. ค่าความแปรปรวนของแต่ละประชากรต้องเท่ากัน คือ $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$
3. การสุ่มตัวอย่างแต่ละชุดจากแต่ละประชากรจะเป็นอิสระกัน

สมมติฐานในการทดสอบ

H_0 : ค่าเฉลี่ยระหว่างประชากรทั้ง k กลุ่มไม่แตกต่างกัน

H_1 : ค่าเฉลี่ยของประชากรอย่างน้อยสองประชากรแตกต่างกัน

หรือ $H_0 : \mu_1 = \mu_2 = \dots = \mu_k$

$H_1 : \mu_i \neq \mu_j$ อย่างน้อย 1 คู่ ; $i \neq j$

จากการแยกความแปรปรวนของข้อมูลทั้งหมด สามารถสรุปเป็นตารางได้ดังนี้

Source of variation	df	SS	MS	F
Between Groups	k - 1	SSB	$MSB = \frac{SSB}{k - 1}$	$F = \frac{MSB}{MSW}$
Within Groups	n - k	SSW	$MSW = \frac{SSW}{n - k}$	
Total	n - 1	SST		

$$\text{เมื่อ } SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$$

$$SSB = \sum_{i=1}^k n_i (\bar{X}_i - \bar{X})^2$$

$$SSW = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2$$

โดยที่ \bar{X}_i = ค่าเฉลี่ยของตัวแปรตามในกลุ่มย่อยที่ i ; $i = 1, 2, \dots, k$
 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานในเพียงการศึกษาเท่านั้น เมื่ออนุญาตให้เผยแพร่ไปยังประชาชนโดยไม่เสียค่าใช้จ่าย
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

\bar{X} = ค่าเฉลี่ยของข้อมูลทั้งหมด

ดังนั้น จะทำการปฏิเสธสมมติฐานหลัก เมื่อค่า F ที่คำนวณได้ มีค่ามากกว่าค่า F ที่เปิดจากตาราง F_{α} ที่อาศาศอิสระ $(k-1)$ และ $(n-k)$

2.6 การทดสอบการแจกแจงแบบปกติของประชากร[11]

2.6.1 การทดสอบของ Lilliefors

ปรับปรุงมาจากการทดสอบของ Kolmogorov – Smirnov เพื่อใช้ทดสอบการแจกแจงแบบปกติ กรณีที่ไม่ทราบค่าเฉลี่ยและความแปรปรวนของประชากร โดยจะใช้ค่าเฉลี่ยและความแปรปรวนของตัวอย่างแทน ใช้กับข้อมูลที่มีมาตรวัดอย่างน้อยแบบอันตรภาค (Interval scale) ขึ้นไป โดยมีขั้นตอนในการทดสอบสมมติฐานดังนี้

สมมติฐานในการทดสอบคือ

H_0 : ประชากรมีการแจกแจงแบบปกติ

H_1 : ประชากรไม่มีการแจกแจงแบบปกติ

สถิติที่ใช้ในการทดสอบ คือ

$$D = \max |F_0(x) - S_n(x)|$$

โดยที่ $F_0(x) = P(X \leq x) = P\left(Z < \frac{x - \bar{x}}{s}\right)$

เมื่อ x คือ ค่าสังเกตของตัวอย่าง

S คือ ส่วนเบี่ยงเบนมาตรฐานของตัวอย่าง

เกณฑ์การตัดสินใจคือ จะปฏิเสธสมมติฐานหลัก (H_0) เมื่อ D มากกว่าค่าวิกฤตที่ได้จากตาราง

Lilliefors test

2.6.2 การทดสอบของ Shapiro – Wilk

การทดสอบการแจกแจงแบบปกติสามารถทำได้หลายการทดสอบ แต่การทดสอบที่ได้รับความนิยมมาก คือ Shapiro-Wilk test หรือกล่าวสั้นๆ ว่า W test เมื่อตัวอย่างสุ่มมีขนาด $n \leq 50$

สมมติฐานของการทดสอบคือ

H_0 : ประชากรมีการแจกแจงแบบปกติ

H_1 : ประชากรไม่มีการแจกแจงแบบปกติ

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น เมื่อผู้เผยแพร่เห็นประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$X_1 < X_2 < \dots < X_n$$

ตัวทดสอบสถิติคือ

$$W = \frac{b^2}{\sum_{i=1}^n (e_i - \bar{e})^2}$$

เมื่อ $b = \sum_{i=1}^k a_{n-i+1} (e_{n-i+1} - e_i)$

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

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

ทั้งนี้ W มีค่ามากกว่าศูนย์เสมอ และ W ที่มีค่ามาก จะสนับสนุน H_0 ดังนั้น

P-value = $P(W < w)$ โดยได้จากตาราง Percentage point of the W test เมื่อ n มีค่าตั้งแต่ 3

ถึง 50

2.7 การทดสอบของเลวี (Levene's Test) [12]

การทดสอบนี้คิดขึ้นโดย Levene (1940) ใช้ทดสอบความแปรปรวนของประชากรตั้งแต่ 3 กลุ่มขึ้นไป โดยมีวิธีการทดสอบดังนี้

1. สมมติฐานในการทดสอบ

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

$$H_1 : \sigma_i^2 \neq \sigma_j^2 \text{ อย่างน้อย 1 คู่ ; } i \neq j$$

2. กำหนดค่าเฉลี่ยของแต่ละกลุ่มตัวอย่าง (\bar{X}_i)
3. กำหนดค่าสัมบูรณ์ของส่วนเบี่ยงเบนจากค่าเฉลี่ยของแต่ละกลุ่มตัวอย่าง นั่นคือ

$$e_{ij} = |X_{ij} - \bar{X}_i|$$

4. นำค่า e_{ij} มาวิเคราะห์ความแปรปรวนแบบทางเดียว (One-way ANOVA)
5. จะได้สถิติทดสอบคือ

$$F = \frac{MSB}{MSW}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. จะทำการปฏิเสธสมมติฐานหลัก เมื่อค่า F ที่คำนวณได้ มีค่ามากกว่าค่า F ที่เปิดจากตาราง F_{α} ที่อาสีสระ $(k-1)$ และ $(n-k)$

2.8 การทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Ranks) [13]

การทดสอบครัสคัลและวอลลิส มุ่งทดสอบความแตกต่างของประชากรที่มากกว่า 2 กลุ่มซึ่งไม่ทราบรูปแบบของการแจกแจงของประชากร เพื่อตรวจสอบว่า ประชากรเหล่านั้นมีคุณลักษณะเหมือนกันหรือไม่ เป็นการวิเคราะห์ความแปรปรวนแบบทางเดียวโดยใช้อันดับ การทดสอบนี้สามารถใช้แทนการวิเคราะห์ความแปรปรวนแบบทางเดียว ในการทดสอบค่าเฉลี่ยของประชากรตั้งแต่สามกลุ่มขึ้นไป เมื่อตัวอย่างจำนวน k กลุ่มถูกสุ่มมาจากประชากรกลุ่มเดียวกัน หรือหลายกลุ่มก็ตาม การใช้การทดสอบครัสคัล-วอลลิส เป็นการทดสอบว่าทั้ง k กลุ่มนี้ มีคุณลักษณะแตกต่างกันหรือไม่ หรือมาจากประชากรกลุ่มเดียวกันหรือไม่ ซึ่งการทดสอบนี้อาจสรุปในรูปของความแตกต่างระหว่างค่าเฉลี่ยก็ได้ แต่อย่างไรก็ตามหลักการสำคัญของวิธีการนี้ก็เช่นเดียวกับการทดสอบแมนน์-วิตนีย์ คือใช้การวิเคราะห์จากอันดับของข้อมูล โดยมีข้อตกลงเบื้องต้นคือ ตัวอย่างทั้ง k กลุ่มที่นำมาศึกษาต้องเป็นตัวอย่างที่ได้มาโดยการสุ่ม ตัวอย่างทั้ง k กลุ่มเป็นอิสระกัน ข้อมูลเป็นข้อมูลแบบต่อเนื่องที่อย่างต่ำต้องอยู่ในระดับเรียงอันดับ

สมมติฐาน

H_0 : ค่ามัธยฐานของประชากรทั้งหมด k ประชากรเท่ากัน

H_1 : มีประชากรอย่างน้อย 1 คู่ ที่มีค่ามัธยฐานไม่เท่ากัน

ค่าสถิติที่ใช้ในการทดสอบสมมติฐาน

$$H = \left[\frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(n+1) \quad \text{โดยที่} \quad n = \sum_{i=1}^k n_i$$

เมื่อ k คือ จำนวนกลุ่มประชากร

n_i คือ จำนวนข้อมูลหรือขนาดตัวอย่างชุดที่ i ; $i = 1, 2, \dots, k$

R_i คือ ผลรวมของอันดับของข้อมูลในตัวอย่างที่ i ; $i = 1, 2, \dots, k$

ในกรณีที่อันดับที่เท่ากันหรือซ้ำกัน จะกำหนดให้อันดับที่ให้เท่ากับค่าเฉลี่ย และให้

$$H = \frac{\left[\frac{12}{n(n+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(n+1)}{1 - \frac{\sum (t_i^2 - t_i)}{n^3 - n}}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อ t คือ จำนวนหน่วยตัวอย่างที่มีค่าเท่ากันในแต่ละลำดับที่
การตัดสินใจ

- ในกรณีที่ $k = 3$ หรือมีจำนวนตัวอย่าง 3 กลุ่ม โดยแต่ละกลุ่มมีจำนวนข้อมูลหรือขนาดไม่เกิน 5 ($n_i \leq 5$) ให้ใช้ค่า H ที่คำนวณได้เปรียบเทียบกับค่าวิกฤต H จากตาราง
- เมื่อตัวอย่างมีจำนวนมากขึ้นและขนาดตัวอย่างใหญ่ขึ้น ค่า H ที่คำนวณได้ จะมีการแจกแจงใกล้เคียงกับการแจกแจงไคสแควร์ (χ^2) ที่มี $df = k - 1$ จึงใช้ค่า H ที่คำนวณได้เปรียบเทียบกับค่าวิกฤต χ^2 จากตาราง

2.9 งานวิจัยที่เกี่ยวข้อง

กมล บุษบา สายทอง อมรวิเชษฐ์ และสารศรัย วัชรภรณ์ (2552) ได้ทำวิจัยเรื่อง การศึกษาเปรียบเทียบความน่าเชื่อถือของโปรแกรมสำเร็จรูปสำหรับงานวิเคราะห์ทางสถิติ ได้ทำการประมาณค่าทางสถิติ 3 ชนิดคือ การวิเคราะห์สถิติตัวแปรเดียว(Univariate Summary Statistic), การวิเคราะห์ความแปรปรวนจำแนกทางเดียว(One-Way Analysis of Variance) และการวิเคราะห์การถดถอยเชิงเส้น(Linear Regression Analysis) ของโปรแกรมสำเร็จรูป 5 โปรแกรม ประกอบด้วย โปรแกรม SAS 9 , R 2.6.2 , Microsoft Excel 2007 , SPSS 15 และ Minitab 15 โดยใช้ข้อมูลจาก The National Institute of Standards and Technology (NIST) ซึ่งแบ่งข้อมูลตามระดับความยาก 3 ระดับคือ ระดับต่ำ, ระดับปานกลาง และระดับสูง โดยกำหนดเลขนัยสำคัญของผลลัพธ์(Significant Digit) 15 ตำแหน่ง ซึ่งผลการวิจัย สามารถสรุปได้ดังนี้ การวิเคราะห์สถิติตัวแปรเดียวพบว่า ค่าเฉลี่ยให้ค่าความแม่นยำในระดับสูงทั้ง 5 โปรแกรม แต่ค่าความแม่นยำของส่วนเบี่ยงเบนมาตรฐานมีค่าอยู่ในระดับสูง เมื่อความยากของข้อมูลอยู่ในระดับต่ำ และค่าความแม่นยำจะมีค่าต่ำลงเมื่อข้อมูลมีระดับความยากสูงขึ้นเหมือนกันทั้ง 5 โปรแกรม ในการวิเคราะห์ความแปรปรวนจำแนกทางเดียว พบว่า โปรแกรม SAS 9 , R 2.6.2 , Microsoft Excel 2007 และ SPSS 15 ให้ค่าความแม่นยำไม่แตกต่างกัน ซึ่งค่าความแม่นยำของสถิติทดสอบเอฟ และค่าคลาดเคลื่อนกำลังสองเฉลี่ย ขึ้นอยู่กับจำนวนค่าสังเกตกับระดับความยากของข้อมูล และโปรแกรม SPSS 15 และ Microsoft Excel 2007 จะให้ค่าต่ำลงมาก เมื่อจำนวนหลักของค่าคงตัวหน้าทศนิยมมีมากขึ้น สำหรับโปรแกรม Minitab 15 ให้ค่าความแม่นยำสูงสุดเมื่อข้อมูลมีระดับความยากต่ำและระดับปานกลาง เนื่องจากขีดความสามารถของโปรแกรมซึ่งไม่สามารถแสดงผลได้ถึง 15 ตำแหน่ง และค่ามาตรฐานหลังทศนิยมมีค่าเป็นศูนย์ ในการวิเคราะห์การถดถอยพบว่า ค่าความแม่นยำของสถิติทดสอบเอฟของโปรแกรม SAS 9, R 2.6.2 และ Microsoft Excel 2007 มีค่าความแม่นยำในระดับสูง ส่วนโปรแกรม SPSS 15 และ Minitab 15 เมื่อข้อมูลมีความยากระดับสูง จะให้ค่าความแม่นยำในระดับต่ำ ส่วนค่าความแม่นยำของค่า

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัมประสิทธิ์การถดถอย เมื่อข้อมูลมีความยาวระดับต่ำและปานกลาง ทั้ง 5 โปรแกรม จะให้ค่าความแม่นยำในระดับสูง แต่เมื่อข้อมูลมีความยาวระดับสูง พบว่า โปรแกรม SAS 9 , R 2.6.2 , Microsoft Excel 2007 และ Minitab 15 ให้ค่าความแม่นยำไม่แตกต่างกัน แต่โปรแกรม SPSS 15 ให้ค่าความแม่นยำต่ำกว่าโปรแกรมอื่น [1]

นิตยา แม็คแนล (2552) ได้ศึกษาการพัฒนาโปรแกรมสำเร็จรูปทางสถิติสำหรับการเรียนการสอนและการวิจัย จะเห็นได้ว่า ปัจจุบันการเรียนการสอนวิชาสถิติจะเน้นทฤษฎีและการปฏิบัติจริง โดยวิเคราะห์ข้อมูลด้วยโปรแกรมสำเร็จรูป โปรแกรมสำเร็จรูปที่ใช้กันแพร่หลาย ส่วนใหญ่มีความซับซ้อนสำหรับผู้เริ่มใช้สถิติ ดังนั้น วัตถุประสงค์ของการวิจัยนี้ เพื่อพัฒนาเครื่องมือในการวิเคราะห์ข้อมูลเบื้องต้นที่ตรงตามวัตถุประสงค์ของการเรียนการสอนสถิติเบื้องต้น และทำการเผยแพร่โปรแกรมนี้สำหรับผู้สนใจทั่วไป โปรแกรมสำเร็จรูป StatLab นี้พัฒนาขึ้น โดยใช้ภาษา Visual Basic for Applications ที่เพิ่มเข้าไป(Add-in) ใน Microsoft Excel ประกอบด้วย 6 ฟังก์ชัน ได้แก่ ฟังก์ชันการนำเสนอข้อมูลด้วยกราฟและแผนภูมิ การวิเคราะห์ข้อมูลเบื้องต้น การทดสอบค่าเฉลี่ยและค่าสัดส่วนของข้อมูลกลุ่มเดียว การเปรียบเทียบค่าเฉลี่ยของข้อมูลสองกลุ่มหรือมากกว่าสองกลุ่มที่เป็นอิสระต่อกัน การทดสอบความสัมพันธ์ของข้อมูล และการทดสอบความเป็นอิสระของข้อมูล เมื่อนำโปรแกรมนี้มาทดลองใช้ในการเรียนการสอนและเผยแพร่ให้ผู้สนใจทั่วไปพบว่า โดยรวมผู้ใช้มีความพึงพอใจในระดับมาก[14]

อุมาพร จันทสร (2554) ได้ทำการวิจัยในเรื่อง “ การเปรียบเทียบประสิทธิภาพของการทดสอบการแจกแจงแบบปกติโดยใช้โปรแกรมสำเร็จรูปทางสถิติ ” ได้ทำการเปรียบเทียบประสิทธิภาพในการทดสอบการแจกแจงปกติของสถิติทดสอบ 3 แบบ คือ Kolmogorov , Lilliefors และ Shapiro – Wilk ที่มีอยู่ใน SPSS และจากสถิติทดสอบ 3 แบบ คือ Kolmogorov , Anderson-Darling และ Ryan-Joiner ที่มีอยู่ใน MINITAB โดยใช้ข้อมูลจากการจำลองแบบจากประชากรที่มีการแจกแจงแบบปกติและที่เบี่ยงเบนไปเล็กน้อยจากการแจกแจงแบบปกติ โดยใช้เมนู Calc-Random Data จากโปรแกรมสำเร็จรูปทางสถิติ MINITAB ในแต่ละกรณีที่มีการกระทำซ้ำ 500 ครั้ง ของการจำลองแบบข้อมูล ที่มีขนาดตัวอย่างเท่ากับ 10 , 20 , 30 , 50 และ 100 ทำการเปรียบเทียบอำนาจการทดสอบและความสามารถในการควบคุมความคลาดเคลื่อนชนิดที่ 1 ผลการวิจัยพบว่า สถิติทดสอบ Ryan-Joiner (คล้าย Shapiro-Wilk) ใน MINITAB มีอำนาจการทดสอบสูงที่สุดในทุกกรณี และทุกขนาดตัวอย่าง ที่ระดับนัยสำคัญ 0.10 โดยเฉพาะเมื่อขนาดตัวอย่างใหญ่เท่ากับ 100 จะมีอำนาจการทดสอบเข้าใกล้ 1 และมีความสามารถในการควบคุมความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้เกือบทั้งหมด ตามเกณฑ์ของ Cochran และยังพบว่า สถิติ Kolmogorov ใน MINITAB เป็นตัวสถิติตัวเดียวกันกับ Lilliefors ในโปรแกรม SPSS [15]

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์เพื่อการเผยแพร่งานวิจัยเท่านั้น เมื่อผู้เผยแพร่เห็นไปใช้ประโยชน์อื่นใด

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริศารักษ์ กาญจนสำราญวงศ์(2555) ได้ศึกษาการเปรียบเทียบการใช้โปรแกรม MINITAB , Instat และ R Commander กรณีการทดสอบด้วยไคกำลังสอง พบว่า การทดสอบด้วยไคกำลังสอง เป็นการทดสอบที่ใช้สำหรับการวิเคราะห์ข้อมูลที่อยู่ในรูปความถี่ ซึ่งมีการนำมาใช้กับงานวิจัยอย่างแพร่หลาย แต่เนื่องจากงานวิจัยส่วนใหญ่มักมีข้อมูลเป็นจำนวนมาก จึงจำเป็นต้องอาศัยโปรแกรมสำเร็จรูปทางสถิติ มาช่วยในการวิเคราะห์ข้อมูลเหล่านั้น โดยโปรแกรมสำเร็จรูปทางสถิติที่มีให้เลือกใช้หลากหลาย โปรแกรม ซึ่งแต่ละโปรแกรมมีข้อดีและข้อจำกัดที่แตกต่างกัน ในงานวิจัยนี้ได้กล่าวถึงการใช้โปรแกรม 3 โปรแกรม คือ MINITAB v. 15 , Instat v. 3.36 และ R Commander 2.9.0 โดยการใช้โจทย์ปัญหา 3 ตัวอย่าง มาทดสอบด้วยไคกำลังสอง พบว่า การคำนวณค่าทางสถิติของการทดสอบด้วยไคกำลังสองจาก ทั้ง 3 โปรแกรม ให้ผลลัพธ์ในข้อสรุปที่เหมือนกัน เพียงแต่มีความแตกต่างกันในเรื่องการนำเข้าข้อมูล และการแจ้งข้อความเตือนเมื่อความถี่คาดหวังมีค่าน้อยกว่า 5 เท่านั้น [2]

มัลลิกา ราชกิจ(2555) ได้ศึกษาการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ ข้อมูลอนุกรมเวลา โดยวิธีบ็อกซ์-เจนกินส์ โดยมีวัตถุประสงค์ของงานวิจัยครั้งนี้ เพื่อเปรียบเทียบ โปรแกรมสำเร็จรูปทางสถิติจำนวน 3 โปรแกรมคือ โปรแกรม MINITAB (รุ่นทดลองใช้งาน 30 วัน) โปรแกรม SPSS และโปรแกรม NCSS2007 (รุ่นทดลองใช้งาน 7 วัน) ในการวิเคราะห์ข้อมูลอนุกรมเวลา ทุกรูปแบบโดยใช้วิธีบ็อกซ์-เจนกินส์ และทำการเปรียบเทียบประสิทธิภาพของการพยากรณ์ด้วยค่าเฉลี่ย เปอร์เซนต์ความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percent error : MAPE) พบว่า โปรแกรม MINITAB เป็นโปรแกรมที่เหมาะสมกับข้อมูลอนุกรมเวลาทุกรูปแบบที่นำมาศึกษา เพราะเป็นวิธีที่ง่าย ในการวิเคราะห์และสะดวกกว่า[16]

อุมาพร จันทสร (2556) ได้ทำการศึกษาความน่าเชื่อถือของผลการวิเคราะห์ด้วยสถิติทดสอบ วิลคอกซัน-แมนวิทนีซ เมื่อคำนึงถึงข้อกำหนดเบื้องต้น จากโปรแกรมสำเร็จรูป SPSS และ MINITAB พบว่า สถิติทดสอบวิลคอกซัน-แมนวิทนีซ(WMW) สามารถใช้ทดสอบถึงความเท่ากันของค่ากลางจาก ประชากร 2 กลุ่มที่เป็นอิสระกัน มีข้อกำหนดเบื้องต้นที่สำคัญคือ ประชากรทั้งสองกลุ่มนั้นต้องมาจากการแจกแจงเดียวกัน แต่ผู้นำไปใช้มักละเลยถึงข้อกำหนดนี้ ดังนั้น วัตถุประสงค์ของงานวิจัยนี้ เพื่อหาผล สรุปว่าโปรแกรม SPSS และ MINITAB ได้คำนึงถึงข้อกำหนดดังกล่าวหรือไม่ ข้อมูลสำหรับการศึกษา ครั้งนี้ได้โดยวิธีมอนติคาร์โล เมื่อกำหนดให้ประชากรมีการแจกแจงเหมือนกันและต่างกัน และ กำหนดให้มีค่ากลางเหมือนกันและต่างกัน จากการทดลองสุ่มตัวอย่าง 1,000 ชุดจากแต่ละประชากรที่ กำหนดให้ เมื่อกำหนดขนาดตัวอย่างจากสองประชากร (n_1 , n_2) ด้วยขนาด (10 , 10) , (10 , 20) และ (10 ,

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

30) และใช้ระดับนัยสำคัญในการทดสอบคือ 0.05 และ 0.10 ทำการเปรียบเทียบค่าระดับนัยสำคัญที่แท้จริง(Empirical α) และอำนาจการทดสอบที่แท้จริง(Empirical power) ที่ได้จากสถิติ WMW จากโปรแกรม SPSS และโปรแกรม MINITAB การศึกษาพบว่า สถิติทดสอบ WMW จากโปรแกรมทั้ง 2 สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ตามเกณฑ์ของ Cochran และ Bradley ได้เฉพาะกรณีที่สองประชากรมีการแจกแจงเหมือนกันในเกือบทุกรูปแบบของการศึกษา แต่สำหรับในกรณีการแจกแจงต่างกันเกือบทั้งหมดของกรณีที่ศึกษาพบว่า สามารถควบคุมค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 ได้เพียงร้อยละ 9-20 เท่านั้น ส่วนอำนาจการทดสอบที่แท้จริง พบว่า ส่วนใหญ่มีค่าสูงเข้าใกล้ 1 ทั้งในกรณีการแจกแจงเหมือนกันหรือต่างกัน โดยโปรแกรม SPSS ให้ค่ามากกว่าโปรแกรม MINITAB เล็กน้อย [17]

[3] <https://sites.google.com/site/kanmanee1992/2-porkaerm-spss>

[4]<http://satitissicc.blogspot.com/p/7.html>

[5] <https://sites.google.com/site/kanmanee1992/3-porkaerm-minitab>

[6]http://www.thailandindustry.com/indust_newweb/news_preview.php?cid=13107

[7]<https://sites.google.com/site/kroopanumas/bth-reiyn-microsoft-excel/excel-khux-xari>

[8]<https://www.9experttraining.com/articles/>

[9] เอกสารประกอบคำสอนอ.ซูใจ

[10] อ.ทรงศิริ

[11](กัลยา วานิชย์บัญชา, 2557)

[12](กัลยา วานิชย์บัญชา, 2549)

[13] (รุ่งกานต์ ใจวงศ์ยะ, 2553)

[14] (พินันท์ คงลาเพชร, 2554)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การดำเนินงานวิจัย

ในการดำเนินการวิจัยครั้งนี้ เพื่อให้เป็นไปตามวัตถุประสงค์ที่ได้กำหนดไว้ ผู้วิจัยจึงได้ดำเนินการตามขั้นตอนในการวิจัยดังต่อไปนี้

3.1 ข้อมูลที่ใช้ในงานวิจัย

เป็นการจำลองข้อมูล โดยมีขั้นตอนดังต่อไปนี้

3.1.1 ศึกษารูปแบบการถดถอยเชิงเส้น

ในที่นี้ทำการศึกษา 2 รูปแบบ ดังนี้

รูปแบบที่ 1 รูปแบบการถดถอยเชิงเส้นอย่างง่าย เป็นรูปแบบกรณีที่ตัวแปรตาม Y มีความสัมพันธ์เชิงสถิติกับตัวแปรอิสระ X เพียงตัวแปรเดียว และลักษณะความสัมพันธ์เป็นแบบเส้นตรง

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

รูปแบบที่ 2 รูปแบบการถดถอยเชิงเส้นพหุคูณ เป็นรูปแบบกรณีที่ตัวแปรตาม Y มีความสัมพันธ์เชิงสถิติกับตัวแปรอิสระ X มากกว่าหนึ่งตัว และลักษณะความสัมพันธ์เป็นแบบเส้นตรง ในที่นี้จะศึกษาเฉพาะตัวแปรอิสระ X_1 และ X_2

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

เมื่อ β_0 คือ ส่วนตัดแกน Y

β_1, β_2 คือ สัมประสิทธิ์การถดถอย

โดยมีข้อสมมติของรูปแบบการถดถอยดังนี้ [3]

1. ค่าตัวแปร X หรือตัวแปร X_1, X_2 จะต้องเป็นค่าที่กำหนดไว้ล่วงหน้าหรือทราบค่า
2. ความคลาดเคลื่อน ε_i มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น 0 และค่าความแปรปรวนเท่ากับ σ^2 นั่นคือ $\varepsilon_i \sim N(0, \sigma^2)$
3. ε_i และ ε_j สำหรับ $i \neq j$ มีการแจกแจงที่เป็นอิสระกัน จะทำให้

$$\text{COV}(\varepsilon_i, \varepsilon_j) = 0$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 การจำลองข้อมูล

ในการจำลองครั้งนี้ จะแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 จะทำการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่าย จากรูปแบบ

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

กำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $X \sim N(3, 2.25)$ และ $\varepsilon \sim N(0, \sigma^2)$ เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1 สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20 โดยมีขั้นตอนดังนี้

1.1 จากรูปแบบการถดถอยเชิงเส้นอย่างง่าย $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ จะทำการจำลองข้อมูลของตัวแปรอิสระ $X \sim N(3, 2.25)$ และค่าคลาดเคลื่อน $\varepsilon \sim N(0, \sigma^2)$ จำนวน n ค่าในการสร้างค่า Y_i จำนวน n ค่า

1.2 นำค่า X_i และค่า Y_i ที่ได้มาวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย เพื่อหาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) โดยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007

1.3 ทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง

1.4 ในแต่ละสถานการณ์ จะทำการเปรียบเทียบค่า MSE ที่ได้จากโปรแกรมทั้ง 3 ด้วยการวิเคราะห์ความแปรปรวน (Analysis of Variance)

ส่วนที่ 2 จะทำการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นพหุคูณ จากรูปแบบ

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

กำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $\beta_2 = 3$, $X_1 \sim N(3, 2.25)$, $X_2 \sim N(5, 4)$ และ $\varepsilon \sim N(0, \sigma^2)$ เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1 สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20 โดยมีขั้นตอนดังนี้

2.1 จากรูปแบบการถดถอยเชิงเส้นพหุคูณ $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$ จะทำการจำลองข้อมูลของตัวแปรอิสระ $X_1 \sim N(3, 2.25)$, $X_2 \sim N(5, 4)$ และค่าคลาดเคลื่อน $\varepsilon \sim N(0, \sigma^2)$ จำนวน n ค่าในการสร้างค่า Y_i จำนวน n ค่า

- 2.2 นำค่า X_{1i} , X_{2i} และค่า Y_i ที่ได้มาวิเคราะห์การถดถอยเชิงเส้นพหุคูณ เพื่อหาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) โดยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007
- 2.3 ทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง
- 2.4 ในแต่ละสถานการณ์ จะทำการเปรียบเทียบค่า MSE ที่ได้จากโปรแกรมทั้ง 3 ด้วยการวิเคราะห์ความแปรปรวน (Analysis of Variance)

3.2 สถิติที่ใช้ในการวิเคราะห์

เมื่อคำนวณค่า MSE ที่ได้จากโปรแกรมทั้ง 3 ในแต่ละสถานการณ์จำลอง ด้วยการซ้ำทั้งหมด 100 ครั้ง หลังจากนั้นจะนำเอาค่า MSE ที่ได้มาทำการวิเคราะห์ทางสถิติ ดังนี้

1. การทดสอบการแจกแจงแบบปกติของประชากร โดยใช้การทดสอบของ Lilliefors
2. การทดสอบความแปรปรวนของประชากรตั้งแต่ 3 กลุ่มขึ้นไป โดยใช้การทดสอบของเลวี (Levene 's Test)
3. การทดสอบความแปรปรวน (Analysis of Variance)
4. การทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Ranks)

บทที่ 4

ผลการวิเคราะห์ข้อมูล

ในการวิจัยครั้งนี้ได้ทำการจำลองข้อมูล เพื่อใช้ในการเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่ายและการถดถอยเชิงเส้นพหุคูณ โดยการจำลองครั้งนี้ จะแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 เป็นการจำลองข้อมูลจากรูปแบบการถดถอยเชิงเส้นอย่างง่าย

เมื่อกำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $X \sim N(3, 2.25)$ และ $\varepsilon \sim N(0, \sigma^2)$

เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1

สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20

ส่วนที่ 2 เป็นการจำลองข้อมูลจากรูปแบบการถดถอยเชิงเส้นพหุคูณ

เมื่อกำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $\beta_2 = 3$, $X_1 \sim N(3, 2.25)$, $X_2 \sim N(5, 4)$

และ $\varepsilon \sim N(0, \sigma^2)$ เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1

สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20

โดยทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง ในแต่ละสถานการณ์ จะทำการเปรียบเทียบค่า MSE ที่ได้จากโปรแกรมทั้งสาม

4.1 ผลการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นอย่างง่าย

เมื่อทำการจำลองข้อมูลในรูปแบบการถดถอยเชิงเส้นอย่างง่ายแล้ว ในแต่ละสถานการณ์จำลอง จะนำเอาข้อมูลที่จำลองได้มาทำการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย โดยคำนวณค่าสถิติจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 แล้วนำเอาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) ที่ได้จากโปรแกรมทั้ง 3 มาเปรียบเทียบผลลัพธ์ที่ได้ ดังนั้น เมื่อทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง จึงได้ค่า MSE จำนวน 100 ค่า จากแต่ละโปรแกรม ได้ผลการวิเคราะห์ดังนี้

4.1.1 ผลการวิเคราะห์การแจกแจงปกติของประชากร

ตารางที่ 4.1 แสดงผลการทดสอบการแจกแจงปกติของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย

สถานการณ์จำลอง		Test of Normality					
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	ค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูป					
		SPSS 22		Minitab 16		Excel 2007	
		Statistic	p-value	Statistic	p-value	Statistic	p-value
10	100	0.097	0.021*	0.098	0.020*	0.096	0.024*
	60	0.087	0.057	0.080	0.115	0.087	0.058
	40	0.068	0.200	0.068	0.200	0.068	0.200
	20	0.064	0.200	0.064	0.200	0.064	0.200
5	100	0.196	0.00**	0.194	0.00**	0.196	0.00**
	60	0.145	0.00**	0.146	0.00**	0.145	0.00**
	40	0.172	0.00**	0.173	0.00**	0.173	0.00**
	20	0.146	0.00**	0.146	0.00**	0.146	0.00**
2	100	0.111	0.004**	0.111	0.004**	0.111	0.004**
	60	0.101	0.014*	0.102	0.013*	0.100	0.015*
	40	0.157	0.00**	0.157	0.00**	0.157	0.00**
	20	0.103	0.010**	0.103	0.010**	0.103	0.010**
1.5	100	0.130	0.00**	0.127	0.00**	0.130	0.00**
	60	0.151	0.00**	0.151	0.00**	0.152	0.00**
	40	0.182	0.00**	0.182	0.00**	0.181	0.00**
	20	0.137	0.00**	0.137	0.00**	0.137	0.00**
1	100	0.053	0.200	0.064	0.200	0.052	0.200
	60	0.078	0.135	0.077	0.154	0.078	0.135
	40	0.044	0.200	0.045	0.200	0.044	0.200
	20	0.078	0.136	0.078	0.136	0.078	0.136

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 (ต่อ)

สถานการณ์จำลอง		Test of Normality					
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	ค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูป					
		SPSS 22		Minitab 16		Excel 2007	
		Statistic	p-value	Statistic	p-value	Statistic	p-value
0.5	100	0.096	0.023*	0.105	0.008**	0.098	0.019*
	60	0.091	0.041*	0.104	0.009**	0.093	0.032*
	40	0.124	0.001**	0.124	0.0001**	0.124	0.001**
	20	0.272	0.00**	0.272	0.00**	0.273	0.00**
0.1	100	0.406	0.00**	0.408	0.00**	0.405	0.00**
	60	0.380	0.00**	0.385	0.00**	0.381	0.00**
	40	0.069	0.200	0.069	0.200	0.074	0.197
	20	0.073	0.200	0.071	0.200	0.074	0.200

จากตารางที่ 4.1 พบว่า ที่ระดับนัยสำคัญ 0.05 ข้อมูลที่มีการแจกแจงปกติ ได้แก่

- ข้อมูลที่มาจากโปรแกรม IBM SPSS 22 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ ค่า $\sigma^2 = 1$ กับ $n = 100, 60, 40, 20$ และค่า $\sigma^2 = 0.1$ กับ $n = 40, 20$
- ข้อมูลที่มาจากโปรแกรม Minitab 16 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ ค่า $\sigma^2 = 1$ กับ $n = 100, 60, 40, 20$ และค่า $\sigma^2 = 0.1$ กับ $n = 40, 20$
- ข้อมูลที่มาจากโปรแกรม Microsoft Excel 2007 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ ค่า $\sigma^2 = 1$ กับ $n = 100, 60, 40, 20$ และค่า $\sigma^2 = 0.1$ กับ $n = 40, 20$

4.1.2 ผลการวิเคราะห์ความแปรปรวนของประชากร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 แสดงผลการทดสอบความแปรปรวนของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย

สถานการณ์จำลอง		ค่าเบี่ยงเบนมาตรฐานของค่า MSE ที่คำนวณได้จากโปรแกรม			Test of Homogeneity of Variance	
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	SPSS 22	Minitab 16	Excel 2007	Statistic	p-value
10	100	1.4751	1.4748	1.4801	0.001	0.999
	60	1.6878	1.6832	1.6878	0.002	0.998
	40	2.1738	2.1752	2.1737	0.000	1.000
	20	3.2550	3.2549	3.2550	0.000	1.000
5	100	0.6416	0.6408	0.6416	0.000	1.000
	60	0.8274	0.8267	0.8274	0.000	1.000
	40	1.2239	1.2243	1.2239	0.000	1.000
	20	1.827	1.827	1.827	0.000	1.000
2	100	0.2597	0.2598	0.2597	0.000	1.000
	60	0.3799	0.3787	0.3800	0.000	1.000
	40	0.5171	0.5167	0.5171	0.000	1.000
	20	0.7911	0.7911	0.7911	0.000	1.000
1.5	100	0.2113	0.2103	0.2113	0.000	1.000
	60	0.2027	0.2032	0.2028	0.001	0.999
	40	0.1794	0.2796	0.2794	0.000	1.000
	20	0.4782	0.4782	0.4781	0.000	1.000
1	100	0.1307	0.1303	0.1307	0.000	1.000
	60	0.1814	0.1813	0.1814	0.000	1.000
	40	0.2116	0.2117	0.2116	0.000	1.000
	20	0.3356	0.3357	0.3356	0.000	1.000

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.2 (ต่อ)

สถานการณ์จำลอง		ค่าเบี่ยงเบนมาตรฐานของค่า MSE ที่ คำนวณได้จากโปรแกรม			Test of Homogeneity of Variance	
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	SPSS 22	Minitab 16	Excel 2007	Statistic	p-value
0.5	100	0.0589	0.0585	0.0590	0.028	0.973
	60	0.0655	0.0652	0.0656	0.005	0.995
	40	0.0941	0.0943	0.0939	0.001	0.999
	20	0.1467	0.1467	0.1467	0.000	1.000
0.1	100	0.0908	0.0908	0.0908	0.000	1.000
	60	0.1073	0.1070	0.1073	0.001	0.999
	40	0.0257	0.0257	0.0258	0.000	1.000
	20	0.0339	0.0339	0.0340	0.000	1.000

จากตารางที่ 4.2 พบว่า ที่ระดับนัยสำคัญ 0.05 ข้อมูลที่มาจากโปรแกรม IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 จะให้ค่าความแปรปรวนเท่ากันทุกสถานการณ์จำลอง

4.1.3 ผลการวิเคราะห์ค่าเฉลี่ยของประชากร

จากผลการวิเคราะห์ในตารางที่ 4.1 และ 4.2 ถ้าข้อมูลชุดใดเป็นไปตามข้อตกลงในการวิเคราะห์ความแปรปรวน (Analysis of Variance) จะใช้ตัวสถิติ F ในการทดสอบ แต่ถ้าข้อมูลชุดใดไม่เป็นไปตามข้อตกลงในการวิเคราะห์ความแปรปรวน จะใช้ตัวสถิติ H ในการทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Ranks)

ตารางที่ 4.3 แสดงผลการทดสอบค่าเฉลี่ยของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย

สถานการณ์จำลอง		ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จาก			ค่าสถิติ		p-value
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	โปรแกรม			ค่า F	ค่า H	
		SPSS 22	Minitab 16	Excel 2007			
10	100	10.11649	10.11662	10.11166	-	0.00	1.00
	60	10.08839	10.09860	10.08840	0.00122	-	0.99877
	40	10.06692	10.06538	10.06694	0.00001	-	0.99998
	20	10.29585	10.29590	10.29590	0.00	-	1.00
5	100	5.32990	5.33025	5.32983	-	0.065	0.968
	60	5.41450	5.41465	5.41454	-	0.004	0.998
	40	5.36350	5.36385	5.36352	-	0.003	0.999
	20	4.99750	4.99750	4.99752	-	0.020	0.990
2	100	2.07230	2.07150	2.07234	-	0.138	0.934
	60	2.10360	2.10330	2.10355	-	0.005	0.997
	40	2.15405	2.15390	2.15417	-	0.061	0.970
	20	2.14480	2.14480	2.14470	-	0.045	0.978
1.5	100	1.42540	1.42600	1.42542	-	0.047	0.977
	60	1.47395	1.47275	1.47391	-	0.180	0.914
	40	1.47575	1.47595	1.47579	-	0.012	0.994
	20	1.57555	1.57555	1.57550	-	0.020	0.990
1	100	1.00344	1.00370	1.0034	0.00013	-	0.99986
	60	1.01474	1.01466	1.01472	0.00	-	0.99999
	40	1.02193	1.02175	1.02195	0.00002	-	0.99997
	20	1.01873	1.01874	1.01879	0.00	-	1.00

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.3 (ต่อ)

สถานการณ์จำลอง		ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จาก			ค่าสถิติ		p-value
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	โปรแกรม			ค่า F	ค่า H	
		SPSS 22	Minitab 16	Excel 2007			
0.5	100	0.49085	0.49110	0.49083	-	0.062	0.969
	60	0.47532	0.47610	0.47534	-	0.227	0.893
	40	0.45483	0.45474	0.45479	-	0.006	0.997
	20	0.43806	0.43806	0.43798	-	0.009	0.996
0.1	100	0.11066	0.11120	0.11073	-	0.123	0.940
	60	0.11771	0.11726	0.11771	-	0.012	0.994
	40	0.10237	0.10237	0.10237	0.00	-	1.00
	20	0.10208	0.10208	0.10206	0.00	-	1.00

จากตารางที่ 4.3 จะพบว่า ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายด้วยการจำลองข้อมูล เมื่อกำหนดค่าความแปรปรวน (σ^2) มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5, 0.1 และขนาดตัวอย่าง(n) มีค่าเป็น 100, 60, 40, 20 ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 จะให้ค่าใกล้เคียงกันทุกสถานการณ์ที่จำลองข้อมูล และเมื่อนำค่า MSE ดังกล่าวมาทำการวิเคราะห์ความแปรปรวน (ANOVA) หรือทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส จะให้ค่า p-value เข้าใกล้ 1 ทุกสถานการณ์เช่นกัน

4.2 ผลการวิเคราะห์ข้อมูลการถดถอยเชิงเส้นพหุคูณ

เมื่อทำการจำลองข้อมูลในรูปแบบการถดถอยเชิงเส้นพหุคูณแล้ว ในแต่ละสถานการณ์จำลอง จะนำเอาข้อมูลที่จำลองได้มาทำการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย โดยคำนวณค่าสถิติจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 แล้วนำเอาค่าความคลาดเคลื่อนกำลังสอง (Mean Square Error ; MSE) ที่ได้จากโปรแกรมทั้ง 3 มาเปรียบเทียบผลลัพธ์ที่ได้ ดังนั้น เมื่อทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง จึงได้ค่า MSE จำนวน 100 ค่า จากแต่ละโปรแกรม ได้ผลการวิเคราะห์ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.1 ผลการวิเคราะห์การแจกแจงปกติของประชากร

ตารางที่ 4.4 แสดงผลการทดสอบการแจกแจงปกติของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

สถานการณ์จำลอง		Test of Normality					
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	ค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูป					
		SPSS 22		Minitab 16		Excel 2007	
		Statistic	p-value	Statistic	p-value	Statistic	p-value
10	100	0.100	0.016*	0.098	0.018*	0.100	0.016*
	60	0.055	0.200	0.059	0.200	0.055	0.200
	40	0.054	0.200	0.052	0.200	0.054	0.200
	20	0.058	0.200	0.058	0.200	0.058	0.200
5	100	0.216	0.00**	0.238	0.00**	0.216	0.00**
	60	0.128	0.00**	0.139	0.00**	0.128	0.00**
	40	0.153	0.00**	0.148	0.00*	0.153	0.00**
	20	0.125	0.001**	0.126	0.00**	0.125	0.001**
2	100	0.147	0.00**	0.128	0.00**	0.101	0.014*
	60	0.100	0.015*	0.116	0.002**	0.100	0.016*
	40	0.167	0.00**	0.169	0.00**	0.167	0.00**
	20	0.130	0.00**	0.128	0.00**	0.130	0.00**
1.5	100	0.128	0.00**	0.152	0.00**	0.128	0.00**
	60	0.089	0.047*	0.111	0.004**	0.090	0.046*
	40	0.151	0.00**	0.143	0.00**	0.151	0.00**
	20	0.122	0.001**	0.119	0.001**	0.122	0.001**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.4 (ต่อ)

สถานการณ์จำลอง		Test of Normality					
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	ค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูป					
		SPSS 22		Minitab 16		Excel 2007	
		Statistic	p-value	Statistic	p-value	Statistic	p-value
1	100	0.059	0.200	0.163	0.00**	0.059	0.200
	60	0.084	0.079	0.139	0.00**	0.083	0.083
	40	0.045	0.200	0.052	0.200	0.046	0.200
	20	0.055	0.200	0.057	0.200	0.055	0.200
0.5	100	0.144	0.00**	0.300	0.00**	0.145	0.00**
	60	0.136	0.00**	0.235	0.00**	0.137	0.00**
	40	0.114	0.003**	0.139	0.00**	0.114	0.003**
	20	0.207	0.00**	0.192	0.00**	0.206	0.00**
0.1	100	0.072	0.200	0.172	0.00**	0.066	0.200
	60	0.124	0.001**	0.160	0.00**	0.128	0.00**
	40	0.073	0.200	0.122	0.001**	0.074	0.200
	20	0.094	0.029*	0.118	0.002**	0.094	0.030*

จากตารางที่ 4.4 พบว่า ที่ระดับนัยสำคัญ 0.05 ข้อมูลที่มีการแจกแจงปกติ ได้แก่

- ข้อมูลที่มาจากโปรแกรม IBM SPSS 22 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ ค่า $\sigma^2 = 1$ กับ $n = 100, 60, 40, 20$ และค่า $\sigma^2 = 0.1$ กับ $n = 100, 40$
- ข้อมูลที่มาจากโปรแกรม Minitab 16 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ และค่า $\sigma^2 = 1$ กับ $n = 40, 20$
- ข้อมูลที่มาจากโปรแกรม Microsoft Excel 2007 เมื่อค่า $\sigma^2 = 10$ กับ $n = 60, 40, 20$ ค่า $\sigma^2 = 1$ กับ $n = 100, 60, 40, 20$ และค่า $\sigma^2 = 0.1$ กับ $n = 100, 40$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 ผลการวิเคราะห์ความแปรปรวนของประชากร

ตารางที่ 4.5 แสดงผลการทดสอบความแปรปรวนของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

สถานการณ์จำลอง		ค่าเบี่ยงเบนมาตรฐานของค่า MSE ที่คำนวณได้จากโปรแกรม			Test of Homogeneity of Variance	
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	SPSS 22	Minitab 16	Excel 2007	Statistic	Sig.
10	100	1.4930	1.4970	1.4930	0.001	0.999
	60	1.7575	1.7576	1.7575	0.000	1.000
	40	2.3224	2.3214	2.3224	0.000	1.000
	20	3.4384	3.4385	3.4384	0.000	1.000
5	100	0.6386	0.6247	0.6387	0.025	0.975
	60	0.8450	0.8414	0.8450	0.001	0.999
	40	1.2633	1.2638	1.2635	0.000	1.000
	20	1.8537	1.8543	1.8537	0.000	1.000
2	100	0.2008	0.2664	0.2606	1.443	0.238
	60	0.3747	0.3706	0.3747	0.002	0.998
	40	0.4509	0.4511	0.4509	0.001	0.999
	20	0.7360	0.7343	0.7360	0.001	0.999
1.5	100	0.2141	0.2170	0.2141	0.058	0.943
	60	0.2032	0.2068	0.2032	0.026	0.974
	40	0.2846	0.2903	0.2847	0.025	0.975
	20	0.5038	0.5040	0.5038	0.000	1.000
1	100	0.1303	0.1371	0.1303	0.060	0.942
	60	0.1849	0.1814	0.1850	0.041	0.960
	40	0.2165	0.2159	0.2166	0.002	0.998
	20	0.3420	0.3418	0.3419	0.000	1.000

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 (ต่อ)

สถานการณ์จำลอง		ค่าเบี่ยงเบนมาตรฐานของค่า MSE ที่ คำนวณได้จากโปรแกรม			Test of Homogeneity of Variance	
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	SPSS 22	Minitab 16	Excel 2007	Statistic	Sig.
0.5	100	0.0589	0.0646	0.0589	0.228	0.796
	60	0.0666	0.0702	0.0667	0.395	0.674
	40	0.1018	0.1036	0.1018	0.042	0.959
	20	0.1347	0.1343	0.1346	0.000	1.000
0.1	100	0.0139	0.0142	0.0139	0.068	0.934
	60	0.0211	0.0213	0.0211	0.025	0.975
	40	0.0263	0.0264	0.0264	0.003	0.997
	20	0.0353	0.0353	0.0354	0.001	0.999

จากตารางที่ 4.5 พบว่า ที่ระดับนัยสำคัญ 0.05 ข้อมูลที่มาจากโปรแกรม IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 จะให้ค่าความแปรปรวนเท่ากันทุกสถานการณ์จำลอง

4.2.3 ผลการวิเคราะห์ค่าเฉลี่ยของประชากร

จากผลการวิเคราะห์ในตารางที่ 4.4 และ 4.5 ถ้าข้อมูลชุดใดเป็นไปตามข้อตกลงในการวิเคราะห์ความแปรปรวน (Analysis of Variance) จะใช้สถิติ F ในการทดสอบ แต่ถ้าข้อมูลชุดใดไม่เป็นไปตามข้อตกลงในการวิเคราะห์ความแปรปรวน จะใช้สถิติ H ในการทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส (The Kruskal-Wallis One-Way Analysis of Variance by Ranks)

ตารางที่ 4.6 แสดงผลการทดสอบค่าเฉลี่ยของค่า MSE ที่คำนวณจากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ

สถานการณ์จำลอง		ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จาก			ค่าสถิติ		p-value
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	โปรแกรม			ค่า F	ค่า H	
		SPSS 22	Minitab 16	Excel 2007			
10	100	10.10041	10.09411	10.10045	-	0.010	0.995
	60	10.04161	10.04080	10.04159	0.00001	-	0.99999
	40	10.03807	10.03780	10.03812	0.00	-	1.00
	20	10.21130	10.21198	10.21136	0.00	-	1.00
5	100	5.33260	5.34000	5.33263	-	0.180	0.914
	60	5.432750	5.42450	5.43272	-	0.321	0.852
	40	5.35180	5.34800	5.35175	-	0.005	0.998
	20	4.98830	4.98900	4.98830	-	0.022	0.989
2	100	2.02910	2.08500	2.07911	-	1.828	0.401
	60	2.11665	2.12150	2.11676	-	0.155	0.925
	40	2.14475	2.14750	2.14484	-	0.060	0.971
	20	2.06930	2.06980	2.06936	-	0.043	0.979
1.5	100	1.42470	1.42000	1.42473	-	0.097	0.953
	60	1.46750	1.46800	1.46748	-	0.055	0.973
	40	1.47090	1.46500	1.47084	-	0.240	0.887
	20	1.57025	1.56950	1.57029	-	0.003	0.999
1	100	1.00381	1.00200	1.00384	-	0.002	0.999
	60	1.01559	1.01780	1.01534	-	0.003	0.999
	40	1.02090	1.02040	1.02094	0.00019	-	0.99981
	20	1.01567	1.01612	1.01565	0.00006	-	0.99994

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.6 (ต่อ)

สถานการณ์จำลอง		ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จากโปรแกรม			ค่าสถิติ		p-value
ค่าความแปรปรวน (σ^2)	ขนาดตัวอย่าง (n)	โปรแกรม			ค่า F	ค่า H	
		SPSS 22	Minitab 16	Excel 2007			
0.5	100	0.48810	0.48700	0.48819	-	0.866	0.648
	60	0.47210	0.47000	0.47207	-	0.620	0.733
	40	0.44962	0.45460	0.44970	-	0.237	0.888
	20	0.41510	0.41410	0.41500	-	0.052	0.974
0.1	100	0.10162	0.10185	0.10162	-	0.325	0.850
	60	0.10244	0.10279	0.10244	-	0.165	0.921
	40	0.10230	0.10222	0.10229	-	0.164	0.921
	20	0.10206	0.10199	0.10206	-	0.015	0.992

จากตารางที่ 4.6 จะพบว่า ในการวิเคราะห์การถดถอยเชิงเส้นพหุคูณด้วยการจำลองข้อมูล เมื่อกำหนดค่าความแปรปรวน (σ^2) มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5, 0.1 และขนาดตัวอย่าง(n) มีค่าเป็น 100, 60, 40, 20 ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 จะให้ค่าใกล้เคียงกันทุกสถานการณ์ที่จำลองข้อมูล และเมื่อนำค่า MSE ดังกล่าวมาทำการวิเคราะห์ความแปรปรวน (ANOVA) หรือทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส จะให้ค่า p-value เข้าใกล้ 1 ทุกสถานการณ์เช่นกัน

บทที่ 5

สรุปผลการวิเคราะห์และข้อเสนอแนะ

ในการวิจัยครั้งนี้ได้ทำการจำลองข้อมูลการถดถอยเชิงเส้นอย่างง่ายและการถดถอยเชิงเส้นพหุคูณ โดยการจำลองครั้งนี้ จะแบ่งออกเป็น 2 ส่วน ดังนี้

ส่วนที่ 1 เป็นการจำลองข้อมูลจากรูปแบบการถดถอยเชิงเส้นอย่างง่าย

เมื่อกำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $X \sim N(3, 2.25)$ และ $\varepsilon \sim N(0, \sigma^2)$

เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1

สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20

ส่วนที่ 2 เป็นการจำลองข้อมูลจากรูปแบบการถดถอยเชิงเส้นพหุคูณ

เมื่อกำหนดให้ $\beta_0 = 5$, $\beta_1 = 1$, $\beta_2 = 3$, $X_1 \sim N(3, 2.25)$, $X_2 \sim N(5, 4)$

และ $\varepsilon \sim N(0, \sigma^2)$ เมื่อ σ^2 มีค่าเท่ากับ 10, 5, 2, 1.5, 1, 0.5 และ 0.1

สำหรับขนาดตัวอย่าง(n) คือ 100, 60, 40 และ 20

โดยทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง ในแต่ละสถานการณ์ จะทำการเปรียบเทียบค่า MSE ที่ได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 สามารถสรุปผลการวิเคราะห์ได้ดังต่อไปนี้

5.1 ข้อมูลการถดถอยเชิงเส้นอย่างง่าย

เมื่อทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง ค่า MSE จำนวน 100 ค่าที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 สามารถสรุปได้ดังนี้

5.1.1 ค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 ส่วนใหญ่ไม่มีการแจกแจงปกติ ที่ระดับนัยสำคัญ 0.05

5.1.2 ค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 มีความแปรปรวนเท่ากันทุกสถานการณ์จำลอง ที่ระดับนัยสำคัญ 0.05

5.1.3 ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab

16 และ Microsoft Excel 2007 จะให้ค่าใกล้เคียงกันทุกสถานการณ์ที่จำลองข้อมูล และเมื่อ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นำค่า MSE ดังกล่าวมาทำการวิเคราะห์ความแปรปรวน (ANOVA) หรือทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส จะให้ค่า p-value เข้าใกล้ 1 ทุกสถานการณ์เช่นกัน

5.2 ข้อมูลการถดถอยเชิงเส้นพหุคูณ

เมื่อทำการจำลองแต่ละสถานการณ์จำนวน 100 ครั้ง ค่า MSE จำนวน 100 ค่าที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 สามารถสรุปได้ดังนี้

5.2.1 ค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 ส่วนใหญ่ไม่มีการแจกแจงปกติ ที่ระดับนัยสำคัญ 0.05

5.2.2 ค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 มีความแปรปรวนเท่ากันทุกสถานการณ์จำลอง ที่ระดับนัยสำคัญ 0.05

5.2.3 ค่าเฉลี่ยของค่า MSE ที่คำนวณได้จากโปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 จะให้ค่าใกล้เคียงกันทุกสถานการณ์ที่จำลองข้อมูล และเมื่อนำค่า MSE ดังกล่าวมาทำการวิเคราะห์ความแปรปรวน (ANOVA) หรือทดสอบความแปรปรวนด้วยอันดับครัสคัลและวอลลิส จะให้ค่า p-value เข้าใกล้ 1 ทุกสถานการณ์เช่นกัน

ดังนั้น ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายและการถดถอยเชิงเส้นพหุคูณ สามารถใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22, Minitab 16 และ Microsoft Excel 2007 มาช่วยในการวิเคราะห์ข้อมูลได้ โดยผลลัพธ์ที่ได้จากโปรแกรมทั้ง 3 จะให้ผลที่ไม่แตกต่างกัน

5.3 ข้อเสนอแนะ

ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายและการถดถอยเชิงเส้นพหุคูณ เมื่อผู้วิจัยใช้โปรแกรมสำเร็จรูปทางสถิติ IBM SPSS 22 , Minitab 16 และ Microsoft Excel 2007 มาทำการประมวลผล จะได้ค่า MSE มีทศนิยม 3 ตำแหน่ง , 1-2 ตำแหน่ง และ 6 ตำแหน่ง ตามลำดับ

บทที่ 6

สรุปผลผลิตงานวิจัย

6.1 สรุปรายชื่อและรายละเอียดผลผลิตงานวิจัยที่ผลิตได้

6.1.1 ได้ทำการตีพิมพ์ในวารสารวิทยาศาสตร์ ลาดกระบัง ปีที่ 28 ฉบับที่ 2 เดือนกรกฎาคม-ธันวาคม 2562 ในหัวข้อ การเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ข้อมูลการถดถอยเชิงเส้น (A COMPARISON OF STATISTICAL PACKAGES IN DATA ANALYSIS OF LINEAR REGRESSION)



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- [1] กมล บุษบา สายทอง อมรวิเชษฐ์ และ สารตริย์ วัชรภรณ์. 2552. การศึกษาเปรียบเทียบความน่าเชื่อถือของโปรแกรมสำเร็จรูปสำหรับงานวิเคราะห์ทางสถิติ. กรุงเทพฯ: มหาวิทยาลัยธรรมศาสตร์.
- [2] ปรีดาภรณ์ กาญจนสำราญวงศ์. 2555. การเปรียบเทียบการใช้โปรแกรม Minitab, Instat และ R Commander กรณีการทดสอบด้วยไคกำลังสอง. สงขลา: มหาวิทยาลัยทักษิณ.
- [3] กัญจุมณี ศรีเที่ยง. 2535. การใช้โปรแกรมสำเร็จรูปทางสถิติ เข้าถึงได้จาก : <https://sites.google.com/site/kanmanee1992/> [20 กุมภาพันธ์ 2561]
- [4] มีนมี สนอดอน. 2558. โปรแกรมสำเร็จรูปทางสถิติ เข้าถึงได้จาก : <http://satitissicc.blogspot.com/p/7.html> [20 กุมภาพันธ์ 2561]
- [5] ----- . 2554. ซอฟต์แวร์วิเคราะห์และประมวลผลทางสถิติ Minitab 16 เข้าถึงได้จาก : http://www.thailandindustry.com/indust_newweb/news_preview.php?cid=13107 [20 กุมภาพันธ์ 2561]
- [6] ภาณุมาศ นัตรทองแดง. 2560. บทเรียน Microsoft Excel เข้าถึงได้จาก : <https://sites.google.com/site/kroopanumas/bth-reiyn-microsoft-excel/excel-khux-xari> [20 กุมภาพันธ์ 2561]
- [7] ----- . 2559. ประโยชน์ของ Microsoft Excel เข้าถึงได้จาก : <https://www.9experttraining.com/articles/ประโยชน์-microsoft-excel> [20 กุมภาพันธ์ 2561]
- [8] ชูใจ คูหารัตนไชย. 2559. เอกสารประกอบคำสอนวิชาสถิติวิเคราะห์. กรุงเทพฯ: ภาควิชาสถิติ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- [9] ทรงศิริ แต่สมบัติ. 2542. การวิเคราะห์การถดถอย. กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.
- [10] กัลยา วานิชย์บัญชา. 2557. หลักสถิติ. พิมพ์ครั้งที่ 14. กรุงเทพฯ: โรงพิมพ์สามลดา.
- [11] กัลยา วานิชย์บัญชา. 2557. สถิติสำหรับงานวิจัย. พิมพ์ครั้งที่ 2. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [12] รุ่งกานต์ ใจวงศ์ยะ. 2553. สถิติวิเคราะห์ 2. เชียงใหม่: สาขาวิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้.
- [13] พินันท์ คงคาเพชร. 2554. สถิตินันพารามетริก. กรุงเทพฯ: แคนีกซ์อินเตอร์คอร์ปอเรชั่น.
- [14] นิตยา แม็คแนล. 2552. การพัฒนาโปรแกรมสำเร็จรูปทางสถิติสำหรับการเรียนการสอนและการวิจัย. สงขลา: มหาวิทยาลัยสงขลานครินทร์.
- [15] อุมพร จันทศร. 2554 การเปรียบเทียบประสิทธิภาพของการทดสอบการแจกแจงแบบปกติโดยใช้โปรแกรมสำเร็จรูปทางสถิติ. กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.
- [16] มลลิกา ราชกิจ. 2555. การเปรียบเทียบโปรแกรมสำเร็จรูปทางสถิติในการวิเคราะห์ข้อมูลอนุกรมเวลา โดยวิธีบ็อกซ์-เจนกินส์. เชียงใหม่: มหาวิทยาลัยเชียงใหม่.
- [17] อุมพร จันทศร. 2556. ความน่าเชื่อถือของผลการวิเคราะห์ด้วยสถิติทดสอบวิลคอกชัน-แมนวิทนีย์ เมื่อคำนึงถึงข้อกำหนดเบื้องต้น จากโปรแกรมสำเร็จรูป SPSS และ MINITAB กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.