

ปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย

FACTOR AFFECTING THE EMPLOYEE'S INCOME
IN THAILAND



ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรบัณฑิต (สถิติประยุกต์)

ภาควิชาสถิติ คณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2566

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

FACTOR AFFECTING THE EMPLOYEE'S INCOME
IN THAILAND



A SPECIAL PROBLEM SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENT FOR
THE DEGREE OF BACHELOR OF SCIENCE (APPLIED STATISTICS)
DEPARTMENT OF STATISTICS, SCHOOL OF SCIENCE
KING MONGUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ **ACADEMIC YEAR 2023** ภาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อปัญหาพิเศษ	ปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย		
	Factor Affecting the Employee's Income in Thailand		
ชื่อนักศึกษา	นางสาวสิยาพัทธ์ โพธิ์ทอง	รหัสนักศึกษา	63050674
	นายชัยญนันต์ เร่งกำเหนิด	รหัสนักศึกษา	63050614
	นายธนพล อยู่ทอง	รหัสนักศึกษา	63050626
ปริญญา	วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)		
ภาควิชา	สถิติ		
ปีการศึกษา	2566		
อาจารย์ที่ปรึกษา	รศ.สายชล สินสมบูรณ์ทอง		

คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.) อนุมัติให้
 ปัญหาพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรบัณฑิต (สถิติประยุกต์)
 ประจำปีการศึกษา 2566

คณะกรรมการสอบ	ลายมือชื่อ
ผศ.ดร.กนกวรรณ ลิโรจนาประภา ประธานกรรมการ	
ดร.ธัญรดา ชัยขจรวัฒน์ กรรมการ	
รศ.สายชล สินสมบูรณ์ทอง กรรมการและอาจารย์ที่ปรึกษา	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามเผยแพร่ไปยังสื่อออนไลน์ และต้องยกย่องเกียรติของเอกสารทุกครั้งที่มีการนำไปใช้

ลิขสิทธิ์ของคณะวิทยาศาสตร์
 สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

หัวข้อปัญหาพิเศษ	ปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย		
	Factor Affecting the Employee's Income in Thailand		
ชื่อนักศึกษา	นางสาวสิยาพัฐ โพธิ์ทอง	รหัสนักศึกษา	63050674
	นายชัยญนันต์ เร่งกำเหนิด	รหัสนักศึกษา	63050614
	นายธนพล อยู่ทอง	รหัสนักศึกษา	63050626
ปริญญา	วิทยาศาสตร์บัณฑิต (สถิติประยุกต์)		
ภาควิชา	สถิติ		
คณะ	วิทยาศาสตร์		
มหาวิทยาลัย	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (สจล.)		
ปีการศึกษา	2566		
อาจารย์ที่ปรึกษา	รศ.สายชล สิ้นสมบุญรณทอง		

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย เนื่องด้วยการทราบถึงที่มาของรายได้จะทำให้เราทราบว่า มีปัจจัยใดบ้างที่ส่งผลต่อรายได้ของประชาชนในปัจจุบัน โดยการวิจัยครั้งนี้ได้นำข้อมูลที่รวบรวมมาจากสำนักงานสถิติแห่งชาติ ประเทศไทย นำมาวิเคราะห์การถดถอยเชิงเส้นพหุคูณ โดยวิธีการถดถอยทีละขั้น ใช้เป็นวิธีในการคัดเลือกตัวแปรเข้าสู่สมการเพื่อใช้ในการหาสมการถดถอย

ผลการวิจัยพบว่าจากตัวแปรอิสระทั้งหมด 15 ตัวแปร คือ ภูมิภาค เพศ เขตการปกครอง สถานภาพสมรส สังกัดของลูกจ้าง ประเภทค่าจ้าง อายุ จำนวนชั่วโมงในการทำงาน จำนวนสมาชิกในครัวเรือน ค่าล่วงเวลา โบนัส เงินอื่นๆ เงินค่าอาหารที่ได้รับ เงินค่าเสื้อผ้าที่ได้รับ และเงินค่าที่อยู่อาศัยที่ได้รับ หลังจากใช้วิธีการถดถอยแบบเป็นขั้นตอนแล้ว พบว่ามีตัวแปรอิสระ 10 ตัวแปร ที่มีความสัมพันธ์กับตัวแปรตามแล้วนำไปวิเคราะห์เพื่อหาตัวแปรสำหรับสร้างสมการถดถอยเชิงเส้นพหุตัวแปรอิสระทั้ง 10 ตัว คือ ภูมิภาค เพศ เขตการปกครอง สถานภาพสมรส จำนวนคนในที่ทำงาน ประเภทค่าจ้าง อายุ จำนวนชั่วโมงในการทำงาน จำนวนสมาชิกในครัวเรือน และค่าล่วงเวลา มีความสัมพันธ์กับรายได้ของลูกจ้างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยมีค่า $R^2 = 0.362$ และสามารถสร้างสมการทำนายรายได้ของลูกจ้างได้ร้อยละ 36.2 ที่เหลืออีกร้อยละ 63.8 เป็นอิทธิพลจากตัวแปรอื่นๆ

คำสำคัญ : การถดถอยเชิงเส้นพหุคูณ การถดถอยทีละขั้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Thesis Title	Factor Affecting the Employee's Income in Thailand		
Student Name	Miss.Siyaphat Phothong	Student ID	63050674
	Mr.Chaiyanan Rangkumnued	Student ID	63050614
	Mr.Thanaphon Youthong	Student ID	63050626
Degree	Bachelor of Science (Applied Statistics)		
Department	Statistics		
School	Science		
University	King Mongkut's Institute of Technology Ladkrabang (KMITL)		
Academic Year	2023		
Advisor	Assoc. Prof. Saichon Sinsomboonthong		

Abstract

This research aims to study the factors that influence the employee's income in Thailand. Due to knowing the origin of income, it will let us know what factors affect the income of the current people. This research has taken the data collected from the National Statistical Office of Thailand to analyze the linear regression. Stepwise Regression method is used as a way to select variables into the equation to find regression equations.

The results showed that from a total of 15 independent variables, namely regions, gender, marital status, working status, type of wages, age, number of working hours, number of household members, overtime, bonuses, other money, food, clothing, and housing. After using the stepwise regression method it was found that there are 10 independent variables that are related to the following variable and then analyzed to find a model for creating multiple linear regression analysis. All 10 independent variables are region, gender, administrative district, marital status, number of people at work, type of wages, age, number of working hours worked, number of household members and overtime. It is statistically significant relationship with the income at the level of 0.05 with a value of $R^2 = 0.362$ and can create a prediction of income of the employed person's income of 36.2 percent. The remaining 63.8 percent is influenced by other variables.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าการ **Keywords** : Multiple Linear Regression, Stepwise Regression เอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

ปัญหาพิเศษเล่มนี้สำเร็จลุล่วงไปได้ด้วยดีและมีความถูกต้องในเนื้อหา เนื่องด้วยได้รับความอนุเคราะห์จาก รศ.สายชล สินสมบูรณ์ทอง ซึ่งเป็นอาจารย์ที่ปรึกษา เป็นผู้ซึ่งให้คำแนะนำ คำปรึกษา เอื้อเพื่อเอกสารต่าง ๆ และหนังสืออ้างอิง ที่ใช้ในการวิเคราะห์ข้อมูลและตรวจทานแก้ไขความถูกต้องตลอดจนติดตามผลงานทุกขั้นตอนของการดำเนินงานในการทำปัญหาพิเศษเล่มนี้จนกระทั่งเสร็จสมบูรณ์ จึงขอกราบขอบพระคุณด้วยความเคารพเป็นอย่างสูงไว้ ณ ที่นี้ด้วย

ขอขอบพระคุณ ผศ.ดร.กนกวรรณ ลีโรจนประภา และ ดร.ธัญรดา ชัยขจรวัฒน์ ผู้ซึ่งเป็นประธาน และกรรมการ ที่กรุณาให้คำปรึกษาแนะนำมาตลอดจนแก้ไขข้อผิดพลาดเพิ่มเติม ทำให้ปัญหาพิเศษเล่มนี้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณคณาจารย์และบุคลากรภาควิชาสถิติ สาขาวิชาสถิติประยุกต์ทุกท่านที่ได้มอบความรู้และคำแนะนำที่มีประโยชน์ รวมถึงการให้ความช่วยเหลือในเรื่องต่าง ๆ มาโดยตลอด

สุดท้ายนี้ขอขอบพระคุณบิดามารดาของผู้จัดทำปัญหาพิเศษที่ให้การสนับสนุนและเป็นกำลังใจให้เสมอมา และขอขอบคุณเพื่อน ๆ ทุกคนที่ให้คำปรึกษา ช่วยเหลือในการทำงานมาโดยตลอด จนปัญหาพิเศษเล่มนี้สำเร็จไปได้ด้วยดี

สียาพัฐ โพธิ์ทอง
ชัยญันต์ เร่งกำหนด
ธนพล อยู่ทอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
บทคัดย่อภาษาอังกฤษ.....	ข
กิตติกรรมประกาศ.....	ค
สารบัญ.....	ง
สารบัญตาราง.....	ช
สารบัญรูป.....	ซ
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์ของงานวิจัย.....	2
1.3 ขอบเขตของงานวิจัย.....	2
1.4 นิยามศัพท์เฉพาะ.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	5
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	6
2.1 สมมติฐานรายได้สมบูรณ์.....	6
2.2 สมมติฐานรายได้เปรียบเทียบ.....	7
2.3 ทฤษฎีข้อมูลสูญหายและการแทนค่าข้อมูลสูญหาย.....	8
2.3.1 ข้อมูลสูญหาย.....	8
2.3.2 การแทนค่าข้อมูลสูญหาย.....	9
2.3.2.1 ค่าเฉลี่ยของค่าใกล้เคียง.....	9
2.3.2.2 ค่ามัธยฐานของค่าใกล้เคียง.....	9
2.3.3 ทฤษฎี Mahalanobis.....	10
2.4 สถิติที่ใช้ในการวิเคราะห์ข้อมูล.....	10
2.4.1 การสุ่มตัวอย่าง.....	10
2.4.2 การวิเคราะห์การถดถอยพหุคูณ.....	11
2.4.2.1 ตัวแบบการวิเคราะห์การถดถอยพหุคูณ.....	11
2.4.2.2 การตรวจสอบข้อสมมติ (Assumption)	
ก่อนสร้างสมการถดถอย.....	12
2.4.2.3 การสร้างสมการถดถอย.....	16
2.4.2.4 การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย.....	18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่ควรเผยแพร่ไปโดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
2.5 งานวิจัยที่เกี่ยวข้อง.....	22
บทที่ 3 วิธีการดำเนินงานวิจัย.....	25
3.1 การจัดเตรียมข้อมูล.....	26
3.1.1 การเก็บรวบรวมข้อมูล.....	26
3.1.2 การศึกษาข้อมูล.....	26
3.1.3 การเลือกตัวแปรที่จะนำมาวิเคราะห์.....	28
3.1.4 การทำความสะอาดและตรวจสอบข้อมูล.....	29
3.1.5 การกำจัดค่า.....	30
3.1.6 การจัดเก็บข้อมูล.....	30
3.2 การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย.....	30
3.2.1 การตรวจสอบการแจกแจงปกติของตัวแปรตาม.....	30
3.2.2 การตรวจสอบความสัมพันธ์เชิงเส้น.....	30
3.2.3 การตรวจสอบความสัมพันธ์เชิงเส้นพหุ.....	30
3.3 การสร้างสมการถดถอย.....	30
3.4 การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย.....	31
3.4.1 การตรวจสอบการแจกแจงปกติของส่วนเหลือ.....	31
3.4.2 การตรวจสอบความเป็นอิสระของส่วนเหลือ.....	31
3.4.3 การตรวจสอบค่าเฉลี่ยของส่วนเหลือ.....	31
3.4.4 การตรวจสอบความเท่ากันของความแปรปรวนของส่วนเหลือ.....	32
3.5 ขั้นตอนการดำเนินงานวิจัย.....	32
บทที่ 4 ผลการวิเคราะห์ข้อมูล.....	35
4.1 ผลการวิเคราะห์ข้อมูลทางสถิติของตัวแปรอิสระ.....	35
4.2 การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย.....	38
4.2.1 การตรวจสอบการแจกแจงปกติของตัวแปรตาม.....	38
4.2.2 การตรวจสอบความสัมพันธ์เชิงเส้น.....	39
4.2.3 การตรวจสอบความสัมพันธ์เชิงเส้นพหุ.....	43
4.3 การสร้างสมการถดถอย.....	44
4.3.1 การหาสมการถดถอย.....	44

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

	หน้า
4.4 การตรวจสอบสมมติหลังการวิเคราะห์การถดถอย.....	47
4.4.1 การตรวจสอบส่วนเหลือมีการแจกแจงปรกติ.....	47
4.4.2 การตรวจสอบส่วนเหลือเป็นอิสระกัน.....	47
4.4.3 การตรวจสอบส่วนเหลือมีความแปรปรวนเท่ากันหรือคงที่.....	48
4.4.4 การตรวจสอบค่าเฉลี่ยส่วนเหลือ.....	48
4.5 การอภิปรายผล.....	49
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ.....	51
5.1 สรุปผลงานวิจัย.....	51
5.2 ข้อเสนอแนะ.....	52
เอกสารอ้างอิง.....	53
ภาคผนวก.....	57
ภาคผนวก ก.....	58
ภาคผนวก ข.....	65

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
3.1 ตัวแปร ชื่อตัวแปร และประเภทตัวแปร.....	28
3.2 ตัวแปรที่มีค่าสูญหาย สถิติเชิงพรรณนาที่แทนค่าสูญหายและค่าสูญหายที่ถูกแทนที่.....	29
4.1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานและตัวแปรตามของตัวแปรอิสระเชิงปริมาณ.....	35
4.2 ความถี่และร้อยละของตัวแปรอิสระเชิงคุณภาพ.....	36
4.3 ผลการทดสอบการแจกแจงปกติของเงินเดือนด้วยวิธีลิลีโฟร์ส (Lilliefors).....	38
4.4 ผลการทดสอบการแจกแจงปกติของ $\frac{1}{\sqrt{Y}}$ ด้วยวิธีลิลีโฟร์ส.....	39
4.5 ค่าสหสัมพันธ์ระหว่างตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณและตัวแปรตาม.....	40
4.6 ค่าสหสัมพันธ์ของสเปียร์แมนระหว่างตัวแปรอิสระเชิงคุณภาพและตัวแปรตามที่เป็นเชิงปริมาณ.....	42
4.7 ค่า VIF และ Tolerance ของตัวแปรอิสระ.....	43
4.8 ผลการวิเคราะห์ความแปรปรวนของตัวแปรตามจากตัวแปรอิสระ 15 ตัวแปร.....	44
4.9 ผลการเลือกตัวแปรอิสระเข้าสมการถดถอยโดยวิธีการถดถอยทีละขั้น.....	45
4.10 ค่าสหสัมพันธ์ (R) สัมประสิทธิ์การกำหนด (R Square) สัมประสิทธิ์การกำหนดแบบปรับ (Adjusted R Square) และค่าสถิติทดสอบเดอ์บิน-วัตสัน ของตัวแปรตามจากตัวแปรอิสระ 15 ตัวแปร.....	46
4.11 ผลการทดสอบการแจกแจงปกติของส่วนเหลือ.....	47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่		หน้า
2.1	สมมติฐานรายได้สมบูรณ์และฟังก์ชันการบริโภค.....	8
3.1	แผนผังแสดงขั้นตอนการดำเนินงานวิจัย.....	34
4.1	อีส์โตรแกรมของรายได้ของลูกจ้าง.....	37
4.2	ผลการแปลงค่าเงินเดือนของลูกจ้างด้วยวิธีการแปลงบ็อก-ค็อก.....	38
4.3	พล็อตกราฟระหว่างส่วนเหลือมาตรฐานและค่าทำนายมาตรฐาน.....	48



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปี 2566 (6 เดือนแรก) คร่าวเรือนทั่วประเทศมีรายได้ทั้งสิ้นเฉลี่ยเดือนละ 29,502 บาท ส่วนใหญ่เป็นรายได้จากการทำงาน 20,799 บาท ซึ่งได้แก่ ค่าจ้างและเงินเดือน 12,930 บาท กำไรสุทธิจากการทำธุรกิจ 5,473 บาท และกำไรสุทธิจากการทำการเกษตร 2,396 บาท ส่วนรายได้ที่ไม่เป็นตัวเงินซึ่งอยู่ในรูปสวัสดิการ/สินค้า บริการต่าง ๆ ที่ได้รับมาโดยไม่ต้องซื้อ (รวมประเมินค่าเช่าบ้าน/ บ้านของตนเอง) 4,232 บาท รายได้ที่ไม่ได้เกิดจากการทำงาน เช่น รายได้จากเงินที่ได้รับความช่วยเหลือจากรัฐหรือบุคคลอื่นนอกครัวเรือน 3,835 บาท และรายได้ไม่ประจำและจากทรัพย์สิน 636 บาท (สำนักงานสถิติแห่งชาติ, 2566) การทราบถึงที่มาของรายได้จะทำให้เราทราบว่าปัจจัยอะไรบ้างที่จะส่งผลต่อค่าจ้าง เพราะจากเศรษฐกิจทุกวันนี้ที่มีราคาแพงขึ้นนำมาซึ่งการดำรงชีวิตที่ต้องมีค่าใช้จ่ายมากมายตามมา การหารายได้ที่เพิ่มขึ้นจึงเป็นสิ่งที่มนุษย์ทุกคนสนใจ

จากการศึกษาของศศิธรและจรงค์ (2557) ได้ศึกษาปัจจัยที่มีผลต่อรายได้ของแรงงานในสถานประกอบการขนาดกลางและขนาดย่อมในอำเภอเมือง จังหวัดขอนแก่น รวบรวมข้อมูลโดยแบบสอบถามและวิเคราะห์ด้วยวิธีการถดถอยแบบกำลังสองน้อยที่สุด และการวิเคราะห์สมการถดถอยแบบสองชั้น ผลการวิจัยพบว่า ปัจจัยด้านทักษะทางสติปัญญาที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกของภาคการบริการและภาคการค้า ได้แก่ จำนวนปีการศึกษา และประสบการณ์ในการทำงาน ปัจจัยทักษะเชิงพฤติกรรมที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกของภาคบริการคือ บุคลิกภาพแบบเปิดกว้าง และปัจจัยทักษะเชิงพฤติกรรมที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกภาคการค้า ได้แก่ บุคลิกภาพแบบเปิดกว้าง และบุคลิกภาพแบบเน้นประสิทธิภาพ สำหรับบุคลิกภาพแบบเจ้าอารมณ์มีผลต่อรายได้ของแรงงานในเชิงลบ ที่ระดับนัยสำคัญทางสถิติ 0.01 ส่วนขององอาจ (2558) ได้ศึกษาปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ หนองคาย มุกดาหาร แม่สอด มีวัตถุประสงค์เพื่อศึกษาผลกระทบต่อรายได้ของคนในพื้นที่จากการพัฒนาเขตเศรษฐกิจพิเศษ และประเมินศักยภาพ โอกาส ข้อจำกัด อุปสรรคของการพัฒนาเขตเศรษฐกิจพิเศษ รวมทั้งวิเคราะห์การเปลี่ยนแปลงความเป็นอยู่ของผู้ที่อยู่อาศัยในบริเวณชายแดน จากการใช้ประโยชน์ในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ โดยใช้ข้อมูลรายได้ครัวเรือนจากสำมะโนประชากร และข้อมูลการสำรวจภาวะแรงงาน ของสำนักงานสถิติแห่งชาติ รวมทั้งข้อมูลภูมิสารสนเทศดินและการใช้ประโยชน์ที่ดิน ของกรมพัฒนาที่ดินโดยวิเคราะห์สภาพการใช้ที่ดิน โดยใช้ข้อมูลทั้ง 3 หน่วยงาน ตั้งแต่ปี พ.ศ. 2548–2557 รวมเป็นระยะเวลา 10 ปี จึงนำมา

วิเคราะห์ด้วยวิธีการวิเคราะห์การถดถอยพหุคูณ และนำมาสร้างตัวแบบการศึกษาเพื่อทำการ
เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
วิเคราะห์ข้อมูลและทดสอบว่าปัจจัยแต่ละตัวมีความสัมพันธ์กับการเปลี่ยนแปลงรายได้ของประชาชน
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีเหตุแต่สงสัยและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษหนองคาย มุกดาหาร แม่สอด จากการศึกษาข้อมูลในภาคครัวเรือนพบว่าปัจจัยที่มีผลต่อการเปลี่ยนแปลงรายได้ของครัวเรือนในพื้นที่เขตเศรษฐกิจพิเศษของหนองคาย มุกดาหาร และแม่สอด คือจำนวนสมาชิกของครัวเรือน อายุ การศึกษา และปีประกาศเขตเศรษฐกิจพิเศษ

จากการศึกษางานวิจัยที่กล่าวมาข้างต้น คณะผู้จัดทำปัญหาพิเศษสนใจที่จะศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย โดยในการทำปัญหาพิเศษนี้จะนำข้อมูลที่รวบรวมมาจากสำนักงานสถิติแห่งชาติประเทศไทย มาศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างโดยใช้การวิเคราะห์การถดถอยพหุคูณ เพื่อหาปัจจัยที่มีอิทธิพลต่อรายได้ของผู้มีงานทำ

1.2 วัตถุประสงค์ของงานวิจัย

เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย

1.3 ขอบเขตของงานวิจัย

ขอบเขตด้านข้อมูล เป็นข้อมูลที่ทางสำนักงานสถิติแห่งชาติได้เก็บรวบรวมข้อมูลไว้แล้วเกี่ยวกับตัวแปรต่างๆ ที่นำมาใช้ในการวิเคราะห์ข้อมูล เป็นข้อมูลทุติยภูมิ (Secondary Data) โดยได้รวบรวมไว้เป็นรายไตรมาส ซึ่งข้อมูลที่น่ามาใช้คือ ปีพ.ศ. 2566 ไตรมาสที่ 2

1.4 นิยามศัพท์เฉพาะ

1.4.1 การถดถอยพหุคูณ (Multiple Regression) คือ การศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวกับตัวแปรตาม 1 ตัว เพื่อศึกษาว่ามีตัวแปรอิสระตัวใดบ้างที่ร่วมกันทำนายหรืออธิบายการผันแปรของตัวแปรตามได้ (นงลักษณ์, 2542)

นิยามศัพท์อ้างอิงจากสำนักงานสถิติแห่งชาติ ดังนี้

1.4.2 กำลังแรงงาน หมายถึง จำนวนผู้ที่มีอายุ 15 ปีขึ้นไป ที่มีความพร้อมและต้องการทำงาน (ธิดารัตน์, 2566)

1.4.3 ลูกจ้าง หมายถึง บุคคลที่มีอายุ 15 ปีขึ้นไป ที่มีลักษณะอย่างใดอย่างหนึ่งดังนี้

1) ได้ทำงานตั้งแต่ 1 ชั่วโมงขึ้นไป โดยได้รับค่าจ้าง เงินเดือน ผลกำไร เงินเป็นผลค่าตอบแทนที่เป็นเงินสดหรือสิ่งของ

2) ทำงานอย่างน้อย 1 ชั่วโมง โดยไม่ได้รับค่าจ้างในวิสาหกิจหรือไร่นาของหัวหน้าครัวเรือนหรือสมาชิกในครัวเรือน

3) ไม่ได้ทำงานหรือทำงานน้อยกว่า 1 ชั่วโมง (เป็นผู้ปกติมีงานประจำ)

3.1 ยังคงได้รับค่าตอบแทน ค่าจ้างผลประโยชน์อื่น ๆ ผลกำไรจากงานหรือธุรกิจ
 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ธุรกิจในระหว่างที่ไม่ได้ทำงาน
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 ไม่ได้รับค่าตอบแทน ค่าจ้างผลประโยชน์อื่น ๆ ผลกำไรจากงานหรือธุรกิจในช่วงที่ไม่ได้ทำงาน แต่ยังมีงานหรือธุรกิจที่จะกลับไปทำ

โดยให้หมายเหตุไว้ว่า คำนิยามในกรณีนี้ นิยามตามมาตรฐานการเก็บรวบรวมข้อมูลสถิติแรงงานขององค์การสหประชาชาติ และ องค์การแรงงานระหว่างประเทศ เพื่อให้ประเทศสมาชิกมีมาตรฐานเดียวกันในการเก็บข้อมูลเพื่อมาเปรียบเทียบสถิติระหว่างประเทศ

1.4.4 เขตแดนนับ (Enumerating Area: EA) หมายถึง เขตพื้นที่หรืออาณาบริเวณที่กำหนดขึ้นโดยแบ่งเป็นส่วนสำหรับใช้ในการปฏิบัติงานโครงการสำมะโน/สำรวจตัวอย่าง เพื่อแสดงให้พนักงานเก็บรวบรวมข้อมูลได้ทราบถึงอาณาเขตหรือบริเวณที่รับผิดชอบในการเก็บรวบรวมข้อมูล โดยพื้นที่ในเขตเทศบาล 1 EA จะประกอบด้วยบ้านประมาณ 100–200 หลัง สำหรับพื้นที่นอกเขตเทศบาลจะแบ่ง EA ตามเขตการปกครองของกรมการปกครอง กระทรวงมหาดไทย โดยกำหนดให้ 1 หมู่บ้านคือ 1 EA ยกเว้นหมู่บ้านขนาดใหญ่ (มีบ้านมากกว่า 450 หลัง) จะแบ่งหมู่บ้านออกเป็นหลาย EA

1.4.5 จำนวนลูกจ้างหารายได้ หมายถึง จำนวนสมาชิกของครัวเรือนซึ่งมีรายได้ที่เป็นตัวเงินจากแหล่งต่าง ๆ เช่น กำไรจากการประกอบธุรกิจ ค่าแรงและเงินเดือน ค่าเช่า หรือเงินช่วยเหลือจากบุคคลภายนอกครัวเรือน ผู้รับเงินรายได้อาจจะเป็นผู้ปฏิบัติงานในเชิงเศรษฐกิจหรือไม่ได้ปฏิบัติงานในเชิงเศรษฐกิจก็ได้

1.4.6 จำนวนสมาชิกในครัวเรือน หมายถึง จำนวนสมาชิกทั้งหมดของครัวเรือน (มีชื่ออยู่ในทะเบียนบ้านหรือไม่ก็ได้) เป็นผู้ที่อยู่เป็นประจำในครัวเรือนและอยู่ในวันสำรวจ มีลักษณะดังนี้คือ ผู้ที่อยู่ประจำในครัวเรือน แต่ได้จากไปที่อื่นชั่วคราว โดยไม่มีวัตถุประสงค์จะไปอยู่ประจำที่อื่น และมีลักษณะอย่างใดอย่างหนึ่ง คือ

1) การไปที่อื่นชั่วคราวไม่เกิน 3 เดือน เช่น ไปเยี่ยมญาติ ผู้ป่วยที่เข้ารับการรักษาในโรงพยาบาล ผู้ที่ไปบวชหรือผู้ไปทำงาน เรียนหนังสือ เป็นต้น (ไม่รวมผู้ที่จากไปเพื่อเรียนหนังสือหรือประกอบอาชีพ ซึ่งมีที่อยู่ประจำที่อื่น ผู้ต้องโทษ ผู้ถูกเกณฑ์ทหาร และคนไข้ ซึ่งต้องอยู่เกิน 3 เดือน)

2) การไป 3 เดือนขึ้นไป แต่ไม่มีที่อยู่อาศัยประจำที่อื่น เช่น ไปทำงานในเรือ 4 เดือน เซลส์แมน ไปฝึกภาคซ้อมรบ เป็นต้น

3) การไปเพื่อศึกษา อบรม ดูงานต่างประเทศไม่ถึง 6 เดือน

1.4.7 รายได้ของครัวเรือน หมายถึง เงินหรือสิ่งของที่ครัวเรือนได้รับมาจากการทำงานหรือผลิตเองหรือจากทรัพย์สิน หรือได้รับความช่วยเหลือจากผู้อื่น ซึ่งทั้งที่เป็นรายได้ประจำและไม่ใช่อายรายได้ประจำ

1.4.8 สถานภาพการทำงาน หมายถึง สถานะของบุคคลในการทำงานในเชิงเศรษฐกิจทุกประเภท

1.4.9 สถานภาพสมรส หมายถึง ความผูกพันระหว่างชายกับหญิงในการเป็นสามีภรรยา (เฉพาะผู้มีอายุ 15 ปีขึ้นไป)

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ของสำนักงานคณะกรรมการการศึกษาศึกษาแห่งชาติ เมื่อผู้ยืมหรือเห็นชอบใช้เอกสารนี้โดยไม่ได้รับอนุญาต
ไม่ว่าการแก้ไขเปลี่ยนแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.4.10 ลูกจ้าง หมายถึง ผู้ที่ทำงานโดยได้รับค่าจ้างเป็น รายเดือน รายสัปดาห์ รายวัน ราย ชั่วโมง หรือ เหมายจ่าย ค่าตอบแทนที่ได้รับจากการทำงาน อาจจะเป็นเงิน หรือสิ่งของ ลูกจ้าง

แบ่งออกเป็น 3 ประเภท

1) ลูกจ้างรัฐบาล หมายถึง ข้าราชการ พนักงานเทศบาล พนักงานองค์การบริหาร ส่วนจังหวัด เจ้าหน้าที่องค์การระหว่างประเทศ ตลอดจนลูกจ้างประจำ และชั่วคราวของรัฐบาล

2) ลูกจ้างรัฐวิสาหกิจ หมายถึง ผู้ที่ทำงานให้กับหน่วยงานรัฐวิสาหกิจ

3) ลูกจ้างเอกชน หมายถึง ผู้ที่ทำงานให้กับเอกชน หรือธุรกิจของเอกชน รวมทั้งผู้ที่รับจ้าง ทำงานบ้าน

1.4.11 ค่าจ้าง หมายถึง เงินที่นายจ้างหรือสถานประกอบการจ่ายให้ลูกจ้าง (ก่อนหักภาษีเงิน ได้บุคคลธรรมดา เงินสมทบกองทุนประกันสังคมในกรณีที่ลูกจ้างต้องจ่ายค่าประกันชีวิตและรายจ่าย อื่น ๆ ของลูกจ้าง) ตามข้อตกลงการจ้างแรงงาน โดยอาจจ่ายตามเงื่อนไขของระยะเวลา หรือจ่ายตาม ปริมาณงาน

1.4.12 ค่าล่วงเวลา โบนัส เงินรางวัลพิเศษ เงินเพิ่มค่าครองชีพ ค่านายหน้า หมายถึง เงิน นอกเหนือจากค่าจ้าง เงินเดือนที่สถานประกอบการจ่ายให้แก่ลูกจ้างเป็นค่าตอบแทนในการทำงาน

1.4.13 สวัสดิการและผลประโยชน์ตอบแทนอื่น ๆ หมายถึง ผลประโยชน์ตอบแทนแรงงานที่ นายจ้างหรือสถานประกอบการจ่าย หรือบริการให้แก่ลูกจ้าง ได้แก่ ค่าตอบแทนแรงงานที่จ่ายเป็น สิ่งของหรือผลผลิต และสวัสดิการที่นายจ้างจัดหาหรือบริการลูกจ้าง เช่น อาหาร เครื่องดื่ม บ้านพัก คนงาน ค่าเช่าบ้าน ค่าซ่อมแซมที่อยู่อาศัย ค่ารักษาพยาบาล บริการดูแลบุตร พาหนะรับส่งมาทำงาน บ้านเทียงหรือนันทนาการต่างๆ ที่จัดให้ลูกจ้าง เป็นต้น ทั้งนี้ไม่รวมเสื้อผ้า เครื่องแบบที่ใช้เฉพาะในการ ปฏิบัติงาน (เงินค่าเสื้อผ้าที่ได้รับ เครื่องแบบ ที่ใช้เฉพาะในการปฏิบัติงาน ถือเป็นค่าใช้จ่ายอื่น ๆ ของ สถานประกอบการ)

1.4.14 เงินที่สถานประกอบการจ่ายสมทบเข้ากองทุนเพื่อการประกันสังคมฯ หมายถึง เงินที่สถานประกอบการจ่ายสมทบเข้ากองทุนเพื่อการประกันสังคมของทั้งรัฐบาลและเอกชน เพื่อมิให้ ลูกจ้างได้รับความเดือดร้อนเมื่อต้องขาดรายได้ไปบางส่วนหรือทั้งหมดอันเนื่องมาจากการเจ็บป่วย หรือประสบอันตรายทั้งในและนอกเวลาท างาน การคลอดบุตร ทูพพลภาพ การว่างงาน ชราภาพ และเสียชีวิต เงินดังกล่าว เช่น เงินสมทบกองทุนประกันสังคม กองทุนเงินทดแทน กองทุนสำรองเลี้ยงชีพ

1.4.15 ครัวเรือนแบ่งได้เป็น 3 ประเภท คือครัวเรือนส่วนบุคคล ครัวเรือนพิเศษ และ ครัวเรือนสถาบัน

1. ครัวเรือนส่วนบุคคล ได้แก่ ครัวเรือนที่มีบุคคลคนเดียว หรือหลายคนอาศัยอยู่ใน บ้านเดียวกัน บ้านที่อาศัยอยู่นั้นอาจเป็นบ้านชนิดใดก็ได้ เช่น เรือนไม้ ตึกแถว ห้องแถว เรือ แพ หรือ บ้านในบริเวณโรงงาน ห้องหรือที่อาศัย ซึ่งแยกต่างหากเฉพาะบุคคลหรือเฉพาะครอบครัวในโรงงาน อุตสาหกรรม โรงพยาบาล เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อผู้ยู่ใช้เห็นเป็นชอบหรือเห็นว่ามีประโยชน์
ไม่ว่ากรณีใดๆ ทั้งสิ้น ยกเว้นหากไม่พบเห็นแต่เปลี่ยนแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ครั้วเรือนพิเศษ หมายถึง ครั้วเรือนซึ่งบุคคลหลายคนอยู่รวมกันในสถานที่อยู่แห่งหนึ่ง เพื่อผลประโยชน์ของตนเอง แบ่งออกเป็น 3 ประเภท คือ

2.1 ครั้วเรือนซึ่งบุคคลกินอยู่รวมกันมีจำนวนตั้งแต่ 6 คนขึ้นไป โดยที่บุคคลเหล่านั้นไม่มีความสัมพันธ์ทางญาติหรือมีความสัมพันธ์ทางญาติไม่เกิน 3 คน และบุคคลเหล่านั้นต่างก็ออกค่าใช้จ่ายตามที่ตกลงกัน

2.2 ครั้วเรือนพิเศษประเภทคนงาน ได้แก่ ครั้วเรือนของคนทำงานตั้งแต่ 6 คนขึ้นไป (โดยไม่นับครอบครัวของเจ้าของ) ที่อยู่อาศัยเป็นประจำในบริเวณที่ทำงานหรือที่โรงงานหรือสถานประกอบการจัดให้อยู่และไม่ได้แยกเฉพาะครั้วเรือน

2.3 ครั้วเรือนพิเศษประเภทหอพัก คือ หอพักซึ่งมีบุคคลทั่วไปเช่าอยู่ ซึ่งไม่ใช่หอพักนักเรียนในโรงเรียน วิทยาลัย หรือมหาวิทยาลัย และไม่ได้แยกอยู่เฉพาะครั้วเรือน

3. ครั้วเรือนสถาบัน ได้แก่ ครั้วเรือนซึ่งบุคคลหลายคนอยู่รวมกัน โดยที่บุคคลเหล่านั้นจะมีความสัมพันธ์ฉันญาติหรือไม่ก็ตาม แต่จำเป็นต้องอยู่รวมกันโดยมีกฎข้อบังคับ ได้แก่

3.1 นักโทษในเรือนจำ

3.2 ผู้อยู่ในสถานสงเคราะห์ต่างๆ

3.3 ทหาร หรือตำรวจที่อยู่ในค่ายกรมกองทหาร ตำรวจ

3.4 นักเรียนประจำในโรงเรียนกึ่งนอน หรือในหอพักนักเรียน หรือนักเรียนพยาบาล

3.5 สงฆ์ สามเณร ชี ลูกศิษย์วัด ที่อยู่ในวัด สำนักสงฆ์

3.6 ผู้จำคุกในโรงเจ

3.7 คนไข้ที่รักษาตัวอยู่ในโรงพยาบาล

3.8 ผู้ที่พักอยู่ประจำในโรงแรมหรือเกสต์เฮาส์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

จากการศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย จะทำให้เห็นถึงความสัมพันธ์ของปัจจัยต่าง ๆ ที่ส่งผลและไม่ส่งผลต่อรายได้ของลูกจ้าง นำมาซึ่งการวางแผนหาแนวทางและนโยบายในการแก้ไขปัญหาทางเศรษฐกิจ คณะผู้จัดทำหวังว่าจะมีประโยชน์ต่อผู้ที่ศึกษาหรือองค์กรที่เกี่ยวข้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในการศึกษาเรื่องปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย คณะผู้ทำปัญหาพิเศษ ได้ศึกษาเอกสารและงานวิจัยที่เกี่ยวข้อง ดังนี้

2.1 สมมติฐานรายได้สมบูรณ์ (Absolute Income Hypothesis)

John Maynard Keynes (1936) (สำนักงานเศรษฐกิจการคลัง, 2546) ได้เขียนทฤษฎีการบริโภคแบบรายได้สมบูรณ์ อธิบายโดยสรุปพบว่าจากการวิเคราะห์ตามกฎจิตวิทยาขั้นพื้นฐาน พบกับการศึกษาข้อเท็จจริงเกี่ยวกับธรรมชาติของมนุษย์ โดยทั่วไปบุคคลจะบริโภคเพิ่มขึ้น เมื่อรายได้เพิ่มขึ้น แต่การบริโภคที่เพิ่มขึ้นนั้นจะน้อยกว่ารายได้ที่เพิ่มขึ้น และเมื่อระดับรายได้ของบุคคลสูงขึ้น บุคคลจะบริโภคในสัดส่วนที่มีต่อรายได้ลดลง นั่นคือค่าเฉลี่ยในการบริโภค (Average Propensity to Consume : APC) ลดลง นั่นคือทำให้สัดส่วนการออมที่มีต่อรายได้เพิ่มขึ้นหรือค่าเฉลี่ยในการออม (Average Propensity to Save : APS) เพิ่มขึ้น แปลว่าเมื่อรายได้สูงขึ้น บุคคลมีแนวโน้มออมมากขึ้น แต่ในทางตรงข้าม ถ้าระดับรายได้ต่ำลง บุคคลจะบริโภคในสัดส่วนที่มีต่อรายได้สูงหรือค่า APC สูง นั่นคือมีแนวโน้มออมน้อยลงเมื่อมีรายได้ต่ำลง เหตุผลที่เป็นเช่นนี้เพราะระดับรายได้ต่ำนั้นการใช้จ่ายพื้นฐานที่จำเป็นต่อการดำรงชีวิตของตัวเองและครอบครัวย่อมมีความสำคัญและจำเป็นเหนือกว่าการออม แสดงว่าบุคคลจะออมมากขึ้นก็ต่อเมื่อตัวเองและครอบครัวมีปัจจัยพื้นฐานได้รับการบำบัดอย่างเพียงพอแล้วเท่านั้น (สำนักงานเศรษฐกิจการคลัง, 2546)

ดังนั้นจึงอาจเขียนในรูปของสมการแสดงความสัมพันธ์ของการบริโภคได้ดังนี้

$$C = a + bY; \quad a > 0, 0 < b < 1 \quad (2.1)$$

โดยที่ C คือ ค่าใช้จ่ายเพื่อการบริโภค

a คือ ค่าใช้จ่ายเพื่อการบริโภคในขณะที่ยังไม่มีรายได้ แม้ว่าบุคคลจะยังไม่มีรายได้ ($Y = 0$)

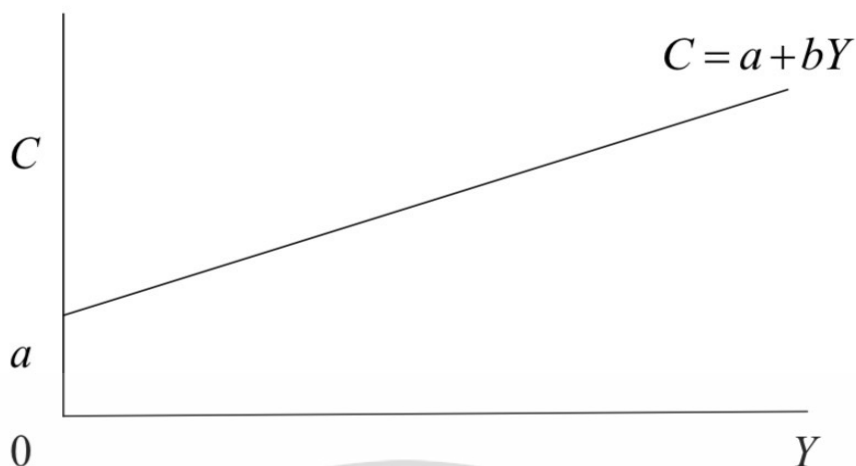
แต่เพื่อความอยู่รอดก็ยังคงต้องกินต้องใช้

b คือ ค่าความโน้มเอียงหน่วยสุดท้ายในการบริโภค (MPC)

Y คือ ระดับรายได้สามารถนำไปใช้จ่ายใช้สอยได้จริง

จากสมการที่ 2.1 ข้างต้นสามารถแสดงได้ดังรูปที่ 2.1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 สมมติฐานรายได้สมบูรณ์และฟังก์ชันการบริโภค

จากสมมติฐานข้างต้นค่าใช้จ่ายเพื่อบริโภค (C) กับรายได้ (Y) มีความสัมพันธ์ในทิศทางเดียวกันและพฤติกรรมการบริโภคจะขึ้นอยู่กับส่วนที่สัมพันธ์กับรายได้และส่วนที่เป็นรายได้ เส้นลาดเอียงจากซ้ายไปขวามีค่าความชันเท่ากับ b เมื่อรายได้เพิ่มขึ้น การบริโภคจะเพิ่มขึ้นและค่าโน้มเอียงเฉลี่ยการบริโภคลดลง โดยผู้บริโภครายได้ต่ำจะมีค่าใช้จ่ายในการบริโภคเป็นสัดส่วนที่มากกว่ารายได้ ทำให้มีค่าความโน้มเอียงเฉลี่ยในการบริโภคสูง ในขณะที่ผู้มีรายได้สูงจะมีค่าใช้จ่ายในการบริโภคเป็นสัดส่วนที่น้อยกว่ารายได้ ทำให้มีค่าความโน้มเอียงเฉลี่ยในการบริโภคต่ำ

2.2 สมมติฐานรายได้เชิงเปรียบเทียบ (Relative Income Hypothesis)

James S. Duesenberry (1949) (นันทกา, 2552) เชื่อว่าค่าใช้จ่ายเพื่อการบริโภคไม่ได้ขึ้นอยู่กับระดับรายได้สมบูรณ์ตามแนวคิดของ John Maynard Keynes เท่านั้น แต่จากการที่มนุษย์มีพฤติกรรมการเลียนแบบในการบริโภคของสังคม (Demonstration Effect) จึงทำให้ค่าใช้จ่ายเพื่อการบริโภคควรจะขึ้นอยู่กับระดับรายได้ เปรียบเทียบกับระดับรายได้โดยเฉลี่ยของคนในสังคมด้วย กล่าวคือครัวเรือนที่มีระดับรายได้ซึ่งเปรียบเทียบแล้วต่ำกว่าระดับรายได้โดยเฉลี่ยของคนในสังคมมีแนวโน้มที่ครัวเรือนเหล่านี้จะมีสัดส่วนของค่าใช้จ่ายเพื่อการบริโภคต่อรายได้ค่อนข้างสูง นั่นคือค่าความโน้มเอียงเฉลี่ยในการบริโภคสูง (Average Propensity to Consume : APC) ทั้งนี้เพราะพฤติกรรมของครัวเรือนจะพยายามรักษาระดับของการบริโภคให้ใกล้เคียงกับระดับการบริโภคโดยเฉลี่ยของคนในสังคม ส่วนครัวเรือนที่มีระดับรายได้ซึ่งเปรียบเทียบแล้วสูงกว่าระดับรายได้เฉลี่ยของคนในสังคมมีแนวโน้มที่ครัวเรือนเหล่านี้จะมีค่า APC ต่ำ ทั้งนี้เพราะครัวเรือนในกลุ่มนี้จะมีระดับของการบริโภคสูงอยู่แล้วดังนั้นจึงสามารถที่จะทำการใช้จ่ายเพื่อการบริโภคโดยจัดสรรเงินจากรายได้เพียง

ส่วนน้อยก็เพียงพอสำหรับการบริโภคสินค้าและบริการให้ใกล้เคียงกับระดับการบริโภคโดยเฉลี่ยของคนในสังคมได้แล้ว นอกจากนี้ Duesenberry จะเชื่อว่าการใช้จ่ายเพื่อการบริโภคขึ้นอยู่กับระดับรายได้ไม่เท่ากันแต่ทั้งนี้ทั้งนั้น ยังเห็นด้วยกับแนวคิดแบบสังคมนิยมและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มาใช้

เปรียบเทียบกับระดับรายได้โดยเฉลี่ยของสังคมแล้ว เขายังเชื่ออีกว่าค่าใช้จ่ายเพื่อการบริโภคในงวดปัจจุบันยังขึ้นอยู่กับระดับค่าใช้จ่ายเพื่อการบริโภคในงวดเวลาก่อนๆ อีกด้วย ทั้งนี้เพราะครัวเรือนที่มีระดับค่าใช้จ่ายเพื่อการบริโภคและระดับรายได้ได้อยู่ในระดับหนึ่งแล้วจะเป็นการยากที่จะทำให้ขาดระดับค่าใช้จ่ายเพื่อการบริโภคให้ต่ำกว่าเดิมเมื่อระดับรายได้เปรียบเทียบกับระดับรายได้ที่เคยได้รับสูงสุด (Previous Peak Income) แยก ค่า APC จึงสูงขึ้น (นันทกา, 2552)

2.3 ทฤษฎีที่ใช้ในการจัดเตรียมข้อมูล

2.3.1 ข้อมูลสูญหาย (Missing Data)

ข้อมูลสูญหาย คือ ข้อมูลที่ไม่ครบถ้วนหรือไม่สมบูรณ์ สามารถเกิดขึ้นได้กับงานวิจัยทุกประเภท โดยเฉพาะอย่างยิ่งงานวิจัยเชิงปริมาณที่ใช้แบบสอบถามเป็นเครื่องมือในการเก็บรวบรวมข้อมูล การดำเนินการกับข้อมูลสูญหายมักเป็นดุลยพินิจของนักวิจัย ซึ่งอาจทำให้ผลการวิจัยมีความคลาดเคลื่อนหรือมีความลำเอียง ดังนั้น นักวิจัยจึงจำเป็นต้องมีแนวทางในการป้องกันเพื่อลดการเกิดข้อมูลสูญหาย และต้องมีความรู้เกี่ยวกับการดำเนินการเกี่ยวกับข้อมูลสูญหาย เพื่อเลือกใช้วิธีดำเนินการกับข้อมูลสูญหายอย่างเหมาะสม (วุฒิ, 2558)

ประเภทของข้อมูลสูญหาย (Type of Missing Data)

การพิจารณาประเภทของข้อมูลสูญหายเป็นขั้นตอนที่สำคัญ เนื่องจากหากสามารถทราบถึงลักษณะของข้อมูลสูญหายจะช่วยในการพิจารณาแนวทางสำหรับจัดการกับปัญหาความไม่สมบูรณ์หรือการสูญหายของข้อมูลได้อย่างเหมาะสม ซึ่งโดยทั่วไป สามารถจำแนกข้อมูลสูญหายออกเป็น 3 ประเภท ดังนี้ (ปิยะภรณ์และสุคนธ์, 2552)

1) การสูญหายแบบสุ่มสมบูรณ์ (Missing Completely at Random : MCAR) เป็นลักษณะของข้อมูลสูญหายที่เกิดขึ้นอย่างสุ่มจากค่าสังเกตทั้งหมด นั่นคือที่สูญหายเป็นอิสระจากตัวแปรต่างๆ สามารถทำการตรวจสอบลักษณะของข้อมูลสูญหายกลุ่มนี้โดยการแบ่งกลุ่มของค่าสังเกตเป็นกลุ่มข้อมูลปกติและข้อมูลสูญหายในกรณีนี้เมื่อทำการทดสอบจะไม่พบความแตกต่างอย่างมีนัยสำคัญระหว่างทั้งสองกลุ่มสำหรับตัวแปรต่างๆ ในฐานข้อมูล

2) การสูญหายแบบสุ่ม (Missing at Random : MAR) เป็นลักษณะของข้อมูลสูญหายซึ่งไม่ได้เกิดขึ้นอย่างสุ่มจากค่าสังเกตทั้งหมด แต่เกิดขึ้นอย่างสุ่มภายในบางส่วนหรือบางกลุ่มของค่าสังเกต นั่นคือค่าของข้อมูลสูญหายขึ้นอยู่กับตัวแปรอื่นๆ ในฐานข้อมูลซึ่งไม่ได้เป็นตัวแปรที่เกิดข้อมูลสูญหาย

3) การสูญหายแบบไม่สุ่ม (Not Missing at Random : NMAR) เป็นลักษณะของข้อมูลซึ่งไม่ได้เกิดขึ้นอย่างสุ่ม โดยค่าของข้อมูลสูญหายขึ้นอยู่กับค่าของข้อมูลสมบูรณ์ในตัวแปรเดียวกันรวมถึงตัวแปรตัวอื่นด้วย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.3.2 การแทนค่าข้อมูลสูญหาย (Missing Value Replacement)

2.3.2.1 ค่าเฉลี่ยของค่าใกล้เคียง (Mean of Nearby Point) คือการแทนที่ค่าข้อมูลสูญหายด้วยค่าเฉลี่ยของค่าที่อยู่ใกล้เคียง ช่วงของจุดใกล้เคียงคือค่าที่อยู่ด้านบนและด้านล่างของค่าที่สูญหายไปจะใช้การคำนวณค่าเฉลี่ย (ฐนัฐ, 2559)

กำหนดให้ $k = a$ โดยที่ k คือ ค่ารอบจุดที่อยู่ต่ำกว่าจุด p จำนวน a จุด และอยู่สูงกว่าจุด p จำนวน a จุด

p คือ ตำแหน่งที่ข้อมูลสูญหาย

a คือ ข้อมูลสูญหายที่มีตำแหน่ง p

$$L = \frac{\sum_{i=1}^b X_{p-i}}{a}; b = k \quad (2.1)$$

โดยที่ L คือ ค่าเฉลี่ยของจุดจำนวน a จุด ที่มีตำแหน่งต่ำกว่าจุด p

b คือ จำนวนจุด a จุด ที่มีตำแหน่งต่ำกว่า p

ถ้าหากตำแหน่ง $p-i$ ตรงกับข้อมูลสูญหายและเลื่อนตำแหน่งไปใช้ข้อมูลตัวใหม่ที่ต่ำกว่า จะได้ b ตัวใหม่ เป็น $b = b + 1$ ทุกๆ ครั้งที่พบค่าสูญหาย และไม่นำค่าข้อมูลสูญหายที่พบมาคำนวณ

$$U = \frac{\sum_{i=1}^b X_{p+i}}{a}; c = k \quad (2.2)$$

โดยที่ U คือ ค่าเฉลี่ยของจุดจำนวน a จุด ที่มีตำแหน่งสูงกว่าจุด p

c คือ จำนวนจุด a จุด ที่มีตำแหน่งต่ำกว่า p

ถ้าหากตำแหน่ง $p+i$ ตรงกับข้อมูลสูญหายและเลื่อนตำแหน่งไปใช้ข้อมูลตัวใหม่ที่สูงกว่า จะได้ c ตัวใหม่ เป็น $c = c + 1$ ทุกๆ ครั้งที่พบค่าสูญหาย และไม่นำค่าข้อมูลสูญหายที่พบมาคำนวณ

$$\hat{X}_p = \frac{L+U}{2} \quad (2.3)$$

โดยที่ \hat{X}_p คือ ค่าประมาณของข้อมูลสูญหายที่ตำแหน่ง p ด้วยค่าเฉลี่ยรอบจุด

2.3.2.2 ค่ามัธยฐานของค่าใกล้เคียง (Median of Nearby Point) คือการแทนค่าที่สูญหายด้วยค่ามัธยฐานของค่าที่อยู่ใกล้เคียง ช่วงของจุดใกล้เคียงคือค่าที่อยู่ด้านบนและด้านล่างของค่าสูญหายไปจะใช้ในการคำนวณค่ามัธยฐาน (ฐนัฐ, 2559)

กำหนดให้ $k = a$ โดยที่ k คือ ค่ารอบจุดที่อยู่ต่ำกว่าจุด p จำนวน a จุด และอยู่สูงกว่าจุด p จำนวน a จุด

p คือ ตำแหน่งที่ข้อมูลสูญหาย

X_p คือ ข้อมูลสูญหายที่มีตำแหน่ง p

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จะทำการเลือกข้อมูลเป็นดังเซต S

$$S = \{X_{p-b}, X_{p-b+1}, X_{p-b+2}, \dots, X_{p-1}, X_{p+1}, X_{p+2}, X_{p+3}, X, \dots, X_{p-c}\} \quad b = a, c = a$$

โดยที่ b คือ จำนวนจุด a จุด ที่มีตำแหน่งต่ำกว่า p

c คือ จำนวนจุด a จุด ที่มีตำแหน่งสูงกว่า p

- ถ้าพบข้อมูลสูญหายในตำแหน่ง $[p-b, p-1]$ จะเลื่อนไปเลือกข้อมูลที่อยู่ตำแหน่งต่ำกว่าจะได้

b ใหม่คือ $b = b+1$ ทุกๆ ครั้งที่พบข้อมูลสูญหาย และไม่นำค่าข้อมูลที่พบมาคำนวณ

- ถ้าพบข้อมูลสูญหายในตำแหน่ง $[p-b, p-1]$ จะเลื่อนไปเลือกข้อมูลที่อยู่ตำแหน่งสูงกว่าจะได้

c ใหม่คือ $c = c+1$ ทุกๆ ครั้งที่พบข้อมูลสูญหาย และไม่นำค่าข้อมูลที่พบมาคำนวณ

กำหนดให้ T คือ เซตลำดับของ S

$$T = \{T_1, T_2, T_3, \dots, T_{2a}\}; T_1 < T_2 < T_3 < \dots < T_{2a}$$

จะได้
$$\hat{X}_p = \frac{T_a + T_{a+1}}{2}$$

โดยที่ \hat{X}_p คือ ค่าประมาณของข้อมูลสูญหายที่ตำแหน่ง p ด้วยค่ามัธยฐาน

2.3.3. ทฤษฎี Mahalanobis คือ Mahalanobis จะแสดงค่า Mahalanobis distance

โดยที่ Mahalanobis distance = $D_i = \left(\frac{X_i - \bar{X}}{S_x} \right)^2$ ถ้าค่า D_i ของกรณีใดมีค่ามาก แสดงว่า ค่า นั้นห่างจากค่าเฉลี่ยมาก จะถือเป็นค่าผิดปกติ (กัลยาและรัฐิตา, 2564)

2.4 สถิติที่ใช้ในการวิเคราะห์ข้อมูล

2.4.1 การสุ่มตัวอย่าง (Sampling)

แผนการเลือกตัวอย่างที่ใช้เป็นแบบแบ่งชั้นภูมิ 2 ชั้นตอน (Stratified Two Stage Sampling) โดยมีจังหวัดเป็นสตราตัม (Stratum) มีในเขตเทศบาลและนอกเขตเทศบาลเป็นสตราตัมย่อย (Subs Stratum) ในการเลือกตัวอย่างขั้นที่หนึ่ง (Primary Sampling Stage) หน่วยตัวอย่างขั้นที่หนึ่งคือเขตแฉงนั้บ โดยกรอบตัวอย่างสำหรับการเลือกหน่วยขั้นที่หนึ่งได้มาจากโครงการสำมะโนประชากรและเคหะ พ.ศ. 2553 ที่ได้มีการปรับปรุงให้มีความทันสมัย จำนวนเขตแฉงนั้บตัวอย่างทั้งสิ้น 5,430 เขตแฉงนั้บ สำหรับการเลือกตัวอย่างขั้นที่สอง (Secondary Sampling Stage) หน่วยตัวอย่างขั้นที่สอง คือ ครั้วเรือนส่วนบุคคลและสมาชิกในครั้วเรือนกลุ่มบุคคลประเภทคนงาน โดยกำหนด 16 ครั้วเรือนตัวอย่างต่อเขตแฉงนั้บตัวอย่างทั้งในเขตเทศบาลและนอกเขตเทศบาลตามลำดับ ในกรณีของครั้วเรือนกลุ่มบุคคลประเภทคนงานจะเป็นการเลือกสมาชิกตัวอย่างจากครั้วเรือนกลุ่มบุคคลประเภทคนงานทุกครั้วเรือน โดยก่อนดำเนินการเลือกครั้วเรือนตัวอย่างได้มีการ

สร้างกรอบครั้วเรือนตัวอย่างด้วยการนับและจดครั้วเรือนและอาคาร/สิ่งปลูกสร้างทุกแห่งในแต่ละเขตแฉงนั้บตัวอย่าง หลังจากนั้นทำการจัดเรียงรายชื่อครั้วเรือนส่วนบุคคลใหม่ตามขนาดครั้วเรือน ซึ่งวัดไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรอบไปใช้

ด้วยจำนวนสมาชิกในครัวเรือน แล้วจึงดำเนินการเลือกครัวเรือนตัวอย่างด้วยวิธีการสุ่มแบบมีระบบ (Systematic Sampling) (สำนักงานสถิติแห่งชาติ, 2566)

2.4.2 การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis)

2.4.2.1 ตัวแบบการวิเคราะห์การถดถอยพหุคูณ

ความหมายของสัมประสิทธิ์การถดถอยเชิงเส้น

ถ้ามีตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม (Y) จำนวน k ตัว คือ X_1, X_2, \dots, X_k สมการการถดถอยพหุคูณ คือ $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + e$

ค่าประมาณของ Y คือ $\hat{Y} = a + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$

จากสมการ ค่าประมาณของ $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ คือ $b_0, b_1, b_2, \dots, b_k$ ตามลำดับ

โดยที่ b_0 คือ ส่วนหรือระยะตัดแกน Y หรือค่าเฉลี่ยของ Y ซึ่งเมื่อกำหนดให้

$$X_1 = X_2 = \dots = X_k = 0$$

b_1, b_2, \dots, b_k คือ ค่าประมาณของสัมประสิทธิ์การถดถอยเชิงเส้นซึ่งมีหน่วยเหมือนกับ Y และมีความหมายดังนี้

b_1 คือ ค่าซึ่งแสดงถึงความสัมพันธ์ระหว่าง Y และ X_1 หมายถึง ถ้า X_1 เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไปโดยเฉลี่ย b_1 หน่วย (ขึ้นอยู่กับเครื่องหมายของ b_1) โดยที่ควบคุมให้ตัวแปรอิสระอื่นๆ คือ X_2, \dots, X_k มีค่าคงที่

b_2 คือ ค่าซึ่งแสดงถึงความสัมพันธ์ระหว่าง Y และ X_2 หมายถึง ถ้า X_2 เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไปโดยเฉลี่ย b_2 หน่วย โดยที่ควบคุมให้ตัวแปรอิสระ X_1, X_3, \dots, X_k มีค่าคงที่

ในทำนองเดียวกัน b_k จะแสดงถึงความสัมพันธ์ระหว่าง Y กับ X_k โดยที่ถ้า X_k เพิ่มขึ้น 1 หน่วย จะทำให้ Y เปลี่ยนไปโดยเฉลี่ย b_k หน่วย โดยที่ควบคุมให้ตัวแปรอิสระ X_1, X_2, \dots, X_{k-1} มีค่าคงที่

ฟังก์ชันความน่าจะเป็น (likelihood function) คือ

$$\ln L(\beta_0, \beta_1) = \prod_{i=1}^n p^{y_i} (1-p)^{1-y_i}$$

ดังนั้น $\log_e L(\beta_0, \beta_1) = \ln L(\beta_0, \beta_1)$

$$= \sum_{i=1}^n \{Y_i \ln[P(Y_i)] + (1-Y_i) Y_i \ln[1-P(Y_i)]\}$$

สมการนี้เรียกว่า ฟังก์ชันล็อก-ความน่าจะเป็น (Log-likelihood function)

การประมาณค่า β_0 และ β_1 จะใช้วิธีความน่าจะเป็นสูงสุด (Maximum likelihood Method) โดยการประมาณค่า β_0 และ β_1 ที่ทำให้ $\ln L(\beta_0, \beta_1)$ มีค่ามากที่สุด ใช้การหาอนุพันธ์บางส่วนเทียบกับ β_0 และ β_1 ให้เท่ากับศูนย์ โดยที่ค่าประมาณของ β_0 คือ b_0 และค่าประมาณ

ของ β_1 คือ b_1 ดังนั้น b_0 และ b_1 ใช้วิธีการประมาณค่าความน่าจะเป็นสูงสุด (Maximum Likelihood estimation : MLE)

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4.2.2 การตรวจสอบข้อสมมติ (Assumption) ก่อนสร้างสมการถดถอย

ขั้นตอนที่ 1 การตรวจสอบการแจกแจงปกติของตัวแปรตาม (Normality) โดยใช้สถิติทดสอบลิลลี่โฟร์ส (Lilliefors Test) เป็นการทดสอบการแจกแจงของประชากรที่มีการแจกแจงปกติหรือไม่ โดยต่างจากสถิติทดสอบคอลโมโกรอฟ-สมิร์นอฟ (Kolmogorov-Smirnov Test) คือสถิติทดสอบคอลโมโกรอฟ-สมิร์นอฟจะต้องกำหนดค่าเฉลี่ย $\mu = \mu_0$ และส่วนเบี่ยงเบนมาตรฐาน $\sigma = \sigma_0$ ไว้ในสมมติฐาน H_0 ด้วย แต่สถิติทดสอบลิลลี่โฟร์สจะไม่กำหนดค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของประชากร จึงต้องประมาณ μ ด้วย \bar{X} และประมาณ σ ด้วย S สถิติทดสอบลิลลี่โฟร์สจะเหมือนกับสถิติทดสอบคอลโมโกรอฟ-สมิร์นอฟเกือบทุกประการ ยกเว้นการใช้คะแนนมาตรฐาน (Normalized value) แทนคะแนนดิบ (สายชล, 2563)

สมมติฐาน

H_0 : ตัวอย่างมาจากประชากรที่มีการแจกแจงปกติ

H_1 : ตัวอย่างมาจากประชากรที่ไม่ได้มีการแจกแจงปกติ

สถิติทดสอบ

$$D = \max |S(x) - F(x)| \quad (2.4)$$

เมื่อ $S(x)$ คือ ความน่าจะเป็นสะสมของตัวอย่าง

$F(x)$ คือ ความน่าจะเป็นสะสมภายใต้สมมติฐานว่าง

เขตวิกฤต

จะปฏิเสธ H_0 ถ้าค่า D มีค่ามากกว่าค่าวิกฤตจากในตารางคอลโมโกรอฟ-

สมิร์นอฟ

การแก้ไขปัญหากรณีตัวแปรตามไม่มีการแจกแจงปกติ

กรณีที่พบว่าตัวแปรตามไม่มีการแจกแจงปกติหรือส่วนเหลือไม่มีการแจกแจงปกติสามารถแก้ไขได้โดยทำการแปลงข้อมูลของตัวแปรตามด้วยวิธีการแปลงบ็อก-ค็อก (Box-Cox Transformation) (Neter et al., 1996) ซึ่งได้เสนอวิธีช่วยในการแก้ปัญหาดังกล่าว โดยการแปลงสมการดังนี้

$$Y' = Y^\lambda \quad (2.5)$$

เมื่อ λ คือ ค่าพารามิเตอร์ที่กำหนดจากข้อมูล ซึ่งมีเงื่อนไขในการกำหนด λ ที่ทำให้ค่าผลบวกกำลังสองของค่าคลาดเคลื่อน (Sum Square of Error : SSE) มีค่าน้อยที่สุดในการนำ Y' ไปวิเคราะห์การถดถอยเชิงเส้น โดยที่ Y' จะแตกต่างกันขึ้นอยู่กับค่า λ ดังนี้

$$\lambda = -1, Y' = \frac{1}{Y}$$

$$\lambda = -0.5, Y' = \frac{1}{\sqrt{Y}}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและตีพิมพ์อ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำ (2.6)

$$\lambda = 0.5, Y' = \sqrt{Y}$$

$$\lambda = 2, Y' = Y^2$$

ถ้า นำข้อมูลที่แปลงแล้วไปทดสอบการแจกแจงปกติพบว่าข้อมูลไม่ได้มีการแจกแจงปกติ จะใช้ทฤษฎีขีดจำกัดส่วนกลางมาอธิบายว่า ถ้าสุ่มตัวอย่างขนาดใหญ่จากประชากรที่มีการแจกแจงใดๆ จะพบว่าค่าเฉลี่ยตัวอย่างจะมีการแจกแจงเข้าใกล้การแจกแจงปกติ ดังนี้

ทฤษฎีบทขีดจำกัดส่วนกลาง (Central Limit Theorem : CLT)

ให้ X_1, X_2, \dots, X_n เป็นตัวอย่างสุ่มขนาด n จากการแจกแจงที่มีค่าเฉลี่ย μ และค่าแปรปรวน σ^2 โดย $-\infty < \mu < \infty$ และ $0 < \sigma^2 < \infty$ แล้วค่าเฉลี่ยตัวอย่าง $\frac{1}{n} \sum_{i=1}^n X_i$ จะมีการ

แจกแจงปกติที่มีค่าเฉลี่ย μ และค่าแปรปรวน $\frac{\sigma^2}{n}$ ดังนั้นตัวแปรสุ่ม $\frac{\frac{1}{n} \sum_{i=1}^n X_i - \mu}{\frac{\sigma}{\sqrt{n}}}$ จะเข้าสู่เชิงการแจกแจงสู่ตัวแปรสุ่ม Z ซึ่งมีการแจกแจงปกติมาตรฐาน เมื่อ $n \rightarrow \infty$ (สุชาดา และคณะ, 2558)

ขั้นตอนที่ 2 การตรวจสอบความสัมพันธ์เชิงเส้น (Linearity) โดยใช้สถิติสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Product-Moment Correlation Coefficient) เป็นการศึกษาความสัมพันธ์เชิงเส้น (Linear Relationship) ระหว่างตัวแปร 2 ตัว หรือข้อมูล 2 ชุด โดยมีข้อตกลงเบื้องต้นคือตัวแปรทั้งสองตัวต้องเป็นตัวแปรต่อเนื่อง อยู่ในมาตราอันตรภาคชั้น (Interval Scale) ขึ้นไป และมีการแจกแจงปกติสองตัวแปร (Bivariate Normal Distribution) โดยค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สันจะมีค่าอยู่ระหว่าง -1 ถึง +1 ถ้าเข้าใกล้ +1 หมายความว่าตัวแปรทั้งสองตัวมีความสัมพันธ์อยู่ในระดับมาก แต่ถ้าสัมประสิทธิ์สหสัมพันธ์มีค่าเข้าใกล้ 0 หมายความว่าตัวแปรทั้งสองตัวมีความสัมพันธ์กันอยู่ระดับน้อยหรือไม่มีความสัมพันธ์กันเลย (ชนินันท์, 2563)

สมมติฐาน

$H_0: \rho = 0$ (ตัวแปรทั้งสองตัวไม่มีความสัมพันธ์กัน)

$H_1: \rho \neq 0$ (ตัวแปรทั้งสองตัวมีความสัมพันธ์กัน)

สถิติทดสอบ

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}, \quad df = n-2 \quad (2.7)$$

เขตวิกฤต

ค่าสถิติทดสอบ t ที่คำนวณได้มีค่ามากกว่า $t_{\frac{\alpha}{2}; n-2}$ หรือน้อยกว่า $-t_{\frac{\alpha}{2}; n-2}$ ที่เปิดจากตาราง

หรือค่า p (p-value) มีค่าน้อยกว่า α จะปฏิเสธ H_0 แสดงว่าตัวแปรทั้งสองตัวมีความสัมพันธ์กัน
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการสงวนลิขสิทธิ์อื่นใด หากมีผู้ใดละเมิดลิขสิทธิ์ในการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การทดสอบความเป็นอิสระกันด้วยค่าลำดับที่สเปียร์แมน

ใช้ทดสอบความสัมพันธ์ระหว่างตัวแปร 2 ตัว ว่ามีความสัมพันธ์กันหรือไม่ หรือมีความสัมพันธ์กันในทิศทางเดียวกันหรือมีความสัมพันธ์กันในทิศทางตรงกันข้าม สัมประสิทธิ์สหสัมพันธ์ชนิดนี้เป็นที่นิยมใช้กันมากที่สุด โดยที่ตัวแปรทั้ง 2 ตัว ต้องมีมาตรวัดอย่างน้อยเป็นแบบเรียงลำดับ (Ordinal scale)

ข้อสมมติ (Assumption)

1. ข้อมูลประกอบด้วยตัวอย่างสุ่มของค่าสังเกตที่เป็นตัวเลขหรือไม่ใช่ตัวเลขก็ได้จำนวน n คู่
2. ค่าสังเกตในแต่ละคู่ถูกวัดใน 2 มาตรวัด

สมมติฐาน (Hypothesis)

ให้ $(X_1, Y_1), \dots, (X_n, Y_n)$ เป็นตัวอย่างสุ่มจากประชากร 2 ตัวแปร แบบต่อเนื่องที่มีฟังก์ชันการแจกแจงสะสมรวม $F(x, y)$ และฟังก์ชันการแจกแจงสะสมทางเดียว $F(x)$ และ $F(y)$

H_0 : X, Y เป็นอิสระกัน คือ $F(x, y) = F(x) F(y)$ ทุกค่า (x, y)

H_1 : X, Y ไม่เป็นอิสระกันหรือขึ้นอยู่กับกันหรือมีความสัมพันธ์กัน

วิธีการทดสอบ (Procedure)

1. เรียงลำดับ X_1, \dots, X_n จากค่าน้อยที่สุดไปหามากที่สุด และให้ R_i เป็นค่าลำดับที่ของ X_i ; $i=1, \dots, n$
2. เรียงลำดับ Y_1, \dots, Y_n จากค่าน้อยที่สุดไปหามากที่สุด และให้ S_i เป็นค่าลำดับที่ของ Y_i ; $i=1, \dots, n$
3. คำนวณสัมประสิทธิ์สหสัมพันธ์ค่าลำดับที่ของสเปียร์แมน (Spearman rank correlation coefficient)

$$r_s = \frac{12 \sum_{i=1}^n \left(R_i - \frac{n+1}{2} \right) \left(S_i - \frac{n+1}{2} \right)}{n(n^2 - 1)}$$

$$= 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)}$$

โดยที่ $D_i = S_i - R_i$; $i=1, \dots, n$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. การทดสอบแบบสองทาง (Two-Sided Test)

สมมติฐาน

$$H_0 : X, Y \text{ เป็นอิสระกัน}$$

$$H_1 : X, Y \text{ ไม่เป็นอิสระกันหรือขึ้นอยู่กับกัน}$$

หรือ $H_0 : \rho_s = 0$

$$H_1 : \rho_s \neq 0$$

เขตวิกฤต

จะปฏิเสธสมมติฐาน H_0 ถ้า $r_s \geq r_{\frac{\alpha}{2}}$ หรือ $r_s \leq -r_{\frac{\alpha}{2}}$

การแก้ปัญหากรณี X และ Y ไม่มีความสัมพันธ์กันในเชิงเส้น

กรณีที่กำหนดรูปแบบความสัมพันธ์ไม่เป็นเชิงเส้น เราสามารถใช้กระบวนการเชิงเส้น (Linearize) เพื่อแปลงข้อมูลไม่เป็นเชิงเส้น (Nonlinear) บางกรณีให้เป็นเชิงเส้น (Linear) ได้ เช่น take ln, take log, การ Invert ค่า y โดยส่วนใหญ่มักจะพบใน 3 รูปแบบ ได้แก่ รูปแบบ Power Model รูปแบบ Exponential Model และรูปแบบ Saturation Growth-Rate Model เป็นต้น (เรียงใจ, 2547)

ขั้นตอนที่ 3 การตรวจสอบความสัมพันธ์เชิงเส้นพหุ (Multicollinearity) โดยใช้สถิติทดสอบ VIF (Variance Inflation Factor) ค่า VIF เป็นค่าที่จะสะท้อนให้เห็นถึงอิทธิพลร่วมของตัวแปรอิสระในตัวแบบเส้นตรง ความสัมพันธ์เชิงเส้นพหุของตัวแปรในตัวแบบเส้นตรงคือการอธิบายว่าตัวแปรอิสระที่อยู่ในตัวแบบนั้นมีความสัมพันธ์กันอย่างไร โดยหากมีค่าไม่เกิน 10 แสดงว่าตัวแปรอิสระในสมการแต่ละตัวเป็นอิสระกัน (สมใจ, 2559)

1) ค่าสถิติทดสอบ VIF (Variance Inflation Factor) หาได้จาก

$$(VIF)_j = \frac{1}{1 - R_j^2} \quad (2.8)$$

โดยที่ R_j^2 คือค่าสัมประสิทธิ์การกำหนดของตัวแบบที่ไม่รวมตัวแปรอิสระตัวที่ j ถ้าค่า $(VIF)_j$ มีค่าเกิน 10 แสดงว่าตัวแปรอิสระมีความสัมพันธ์เชิงเส้นพหุ (สมใจ, 2559)

2) ค่าสถิติทดสอบความคลาดเคลื่อนยินยอม (Tolerance) หาได้จาก

$$(TOL)_j = \frac{1}{(VIF)_j} \quad (2.9)$$

ดังนั้นค่าความคลาดเคลื่อนยินยอมจึงเป็นส่วนกลับของ $(VIF)_j$ ทำให้มีการปิดค่า $(TOL)_j$

ที่ได้นำมาใช้เพื่อตรวจสอบความสัมพันธ์เชิงเส้นพหุ (Multicollinearity) ของตัวแปรอิสระในสมการแต่ละตัวเป็นอิสระกัน (สมใจ, 2559)

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) ค่าลักษณะเฉพาะ (Eigen Value) ถ้าค่าลักษณะเฉพาะมีค่าเข้าใกล้ 0 แสดงว่าตัวแปรอิสระมีความสัมพันธ์เชิงเส้นพหุ (สมใจ, 2559)

4) ค่าดัชนีเงื่อนไข (Condition Index) ถ้าค่าดัชนีเงื่อนไข > 30 แสดงว่าตัวแปรอิสระมีความสัมพันธ์เชิงเส้นพหุ (สมใจ, 2559)

การแก้ไขปัญหากรณีตัวแปรอิสระมีความสัมพันธ์เชิงเส้นพหุสูง

กรณีที่เกิดปัญหาตัวแปรอิสระมีความสัมพันธ์เชิงเส้นพหุสูง สามารถแก้ปัญหาได้โดยใช้วิธีการวิเคราะห์ปัจจัย (Factor Analysis) ซึ่งเป็นเทคนิคการรวมตัวแปรที่มีความสัมพันธ์กันไว้ในกลุ่มหรือปัจจัย (Factor) เดียวกัน ตัวแปรที่อยู่ในปัจจัยเดียวกันจะมีความสัมพันธ์กันมาก โดยความสัมพันธ์นี้อาจเป็นไปได้ทั้งทิศทางบวกหรือทิศทางลบก็ได้ ส่วนตัวแปรที่อยู่คนละปัจจัยจะไม่มีความสัมพันธ์กันหรือมีความสัมพันธ์กันน้อยมาก (กัลยา, 2548)

2.4.2.3 การสร้างสมการถดถอย

ขั้นตอนที่ 1 การเลือกตัวแปรโดยวิธีการถดถอยทีละขั้น (Stepwise Regression) เป็นวิธีการคัดเลือกที่ผสมผสานระหว่างวิธีการเลือกแบบไปข้างหน้า (Forward Selection) และวิธีการกำจัดถดถอยหลัง (Backward Selection) เข้าด้วยกัน วิธีการถดถอยทีละขั้นสามารถทำได้ดังนี้

1) ใช้วิธีการเลือกแบบไปข้างหน้า (Forward Selection) โดยคัดเลือกตัวแปรอิสระ 1 ตัวเข้าไว้ในตัวแบบการถดถอย โดยเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม Y สูงสุด หรือที่ทำให้ตัวแบบการถดถอยมีค่าผลบวกกำลังสองของค่าคลาดเคลื่อน (SSE) ที่ต่ำสุดไว้ในตัวแบบการถดถอยเป็นตัวแปรแรก จากนั้นคำนวณค่าสถิติเอฟบางส่วน (Partial F test) หรือค่าสถิติที (t-test) เพื่อทดสอบสมมติฐานว่าตัวแปรอิสระมีความสำคัญต่อการทำนายค่าของ Y หรือไม่ ถ้าตัวแปรอิสระนั้นไม่มีความสำคัญต่อการทำนายค่าของ Y จะสิ้นสุดการคัดเลือกตัวแปรอิสระเข้าไว้ในตัวแบบการถดถอย

2) ใช้วิธีการกำจัดถดถอยหลัง (Backward Elimination) โดยคัดตัวแปรอิสระนั้นออกจากรูปแบบการถดถอยหากเป็นไปได้ โดยจะสามารถคัดตัวแปรอิสระที่อยู่ในตัวแบบออกได้ ตัวแปรอิสระที่มีความสำคัญต่อการทำนายจะไม่สามารถคัดออกจากตัวแบบการถดถอยได้

3) พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรอิสระกับตัวแปรตาม Y โดยจะเลือกตัวแปรอิสระที่ให้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนกับตัวแปรตามสูงที่สุด โดยตัวแปรอิสระที่ถูกพิจารณาตัวต่อมามีความสำคัญต่อการทำนายค่าของ Y เมื่อมีตัวแปรอิสระก่อนหน้าอยู่ในตัวแบบการถดถอยก่อน แล้วจะคัดตัวแปรอิสระที่ถูกพิจารณาเข้ามาอยู่ในตัวแบบเป็นตัวต่อมา

4) ทำการทดสอบความมีนัยสำคัญของตัวแปรที่เข้ามาเป็นตัวสุดท้ายก่อนโดยพิจารณาจากค่าสถิติเอฟบางส่วน ถ้ามีค่าน้อยกว่าเกณฑ์ที่กำหนดไว้จะต้องนำตัวแปรนั้นออกจากตัวแบบ แต่ถ้าค่าสถิติเอฟบางส่วนมีค่ามากกว่าเกณฑ์ที่กำหนดไว้ ก็จะย้อนกลับไปทดสอบตัวแปรอิสระ

ตัวก่อนหน้าที่เข้ามาอยู่ในตัวแบบ โดยพิจารณาจากค่าสถิติเอฟบางส่วนเช่นเดียวกัน ทำซ้ำในขั้นตอนที่สามและสี่ต่อไปเรื่อยๆ จนกระทั่งไม่สามารถนำตัวแปรอิสระได้ออกจากตัวแบบได้ หรือไม่สามารถนำตัวแปรอิสระใดเข้าสู่ตัวแบบได้อีก ข้อดีของวิธีการนี้คือสามารถแก้จุดบกพร่องของวิธีการเลือกตัวแปรโดยวิธีการเลือกแบบไปข้างหน้าและวิธีการเลือกตัวแปรโดยวิธีการกำจัดถอยหลังได้ (Montgomery et al., 2012)

ขั้นตอนที่ 2 สัมประสิทธิ์สหสัมพันธ์พหุคูณ (Multiple Coefficient of Correlation) (กัลยา, 2558) ค่าของสัมประสิทธิ์สหสัมพันธ์พหุคูณได้จากการถอดรากที่สองของสัมประสิทธิ์การกำหนดพหุคูณ ดังนั้นสัมประสิทธิ์สหสัมพันธ์พหุคูณ $= R_{Y.12\dots k} = R = \sqrt{R^2_{Y.12\dots k}}$ โดยที่ $0 \leq R \leq 1$ ซึ่งสัมประสิทธิ์สหสัมพันธ์พหุคูณหมายถึงค่าสัมประสิทธิ์สหสัมพันธ์ระหว่าง Y กับ X_1, X_2, \dots, X_k ถ้าสัมประสิทธิ์สหสัมพันธ์พหุคูณมีค่ามากแสดงว่า Y กับ X_1, X_2, \dots, X_k มีความสัมพันธ์กันมากดังนี้

1. R มีค่าเข้าใกล้ศูนย์ แสดงว่า Y มีความสัมพันธ์กับ X_1, X_2, \dots, X_k น้อยมาก และถ้า $R=0$ แสดงว่า Y ไม่มีความสัมพันธ์กับ X_1, X_2, \dots, X_k เลย
2. R มีค่าเข้าใกล้ 1 แสดงว่า Y มีความสัมพันธ์กับ X_1, X_2, \dots, X_k มาก

ขั้นตอนที่ 3 สัมประสิทธิ์การกำหนดพหุคูณ (Multiple Coefficient of Determination: R^2 หรือ r^2) (กัลยา, 2558)

สัมประสิทธิ์การกำหนดพหุคูณ คือ สัดส่วนหรือร้อยละที่ตัวแปรอิสระ X_1, X_2, \dots, X_k สามารถอธิบายการเปลี่ยนแปลงของ Y ได้ หรือกล่าวได้ว่าสัมประสิทธิ์การกำหนดพหุคูณเป็นสัดส่วนหรือร้อยละของความผันแปรของ Y ที่มีสาเหตุเนื่องจากความผันแปรของ X_1, X_2, \dots, X_k โดยที่สัมประสิทธิ์การกำหนดพหุคูณโดยทั่วไปจะใช้สัญลักษณ์ R^2 โดยที่

$$R^2 = \frac{\text{ความผันแปรของ } Y \text{ เนื่องจากอิทธิพลของ } X_1, X_2, \dots, X_k}{\text{ความผันแปรทั้งหมด}} \quad (2.10)$$

$$r^2 = R^2 = \frac{SSR}{SST}$$

$$\text{โดยที่ } 0 \leq R^2, r^2 \leq 1$$

ถ้าค่า R^2 ที่ใกล้ 1 จะหมายถึง X_1, X_2, \dots, X_k มีความสัมพันธ์กับ Y มาก แต่ถ้า R^2 เข้าใกล้ 0 หมายถึงค่า X_1, X_2, \dots, X_k มีความสัมพันธ์กับ Y น้อย

เนื่องจาก SSR จะเพิ่มขึ้นถ้าเพิ่มตัวแปรอิสระ เช่น เดิมมี X_1 และ X_2 ที่มีความสัมพันธ์กับ Y แต่ถ้าเพิ่มตัวแปรอิสระ X_3 เข้าในสมการการถดถอย จะได้ว่า

$$SSR(X_1, X_2, X_3) > SSR(X_1, X_2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ $SSR(X_1, X_2, X_3)$ หมายถึง SSR ของสมการการถดถอยที่มีตัวแปรอิสระ X_1, X_2 และ X_3

และ $SSR(X_1, X_2)$ หมายถึง SSR ของสมการการถดถอยที่มีตัวแปรอิสระ X_1 และ X_2

ดังนั้นเมื่อเพิ่มตัวแปรอิสระเข้าสมการการถดถอย จะทำให้ค่า R^2 มากขึ้นทั้งที่ตัวแปรอิสระ X ที่เพิ่มอาจจะไม่มีความสัมพันธ์กับ Y เลยก็ได้ จึงมีการปรับค่า R^2 ให้ถูกต้องขึ้น เรียกว่า สัมประสิทธิ์การกำหนดพหุคูณแบบปรับ (Adjusted R^2) โดยที่

$$R_A^2 = AdjustedR^2 = 1 + \frac{(n-1)}{(n-k-1)}(R^2 - 1) \quad (2.11)$$

กรณีที่มีตัวแปรอิสระหลายตัว จะพิจารณาจากค่า R_A^2

ขั้นตอนที่ 4 สัมประสิทธิ์สหสัมพันธ์เชิงส่วน (Coefficients of Partial Correlation) สัมประสิทธิ์สหสัมพันธ์เชิงส่วนเป็นค่าที่แสดงความสัมพันธ์ระหว่าง Y กับ X ตัวใดตัวหนึ่งโดยให้ X ตัวอื่นๆ มีค่าคงที่ เช่น ถ้า Y ความสัมพันธ์กับตัวแปรอิสระ 3 ตัว (X_1, X_2, X_3) สัมประสิทธิ์สหสัมพันธ์เชิงส่วนระหว่าง Y กับ X_j โดยกำหนดให้ X_i และ X_k คงที่ ($i \neq j \neq k$) จะแสดงความสัมพันธ์ระหว่าง Y กับ X_j จริงๆ โดยกำจัดอิทธิพลของ X_i และ X_k ที่มีต่อ Y

สัญลักษณ์ของสัมประสิทธิ์สหสัมพันธ์เชิงส่วนที่ใช้คือ

$r_{Y1.23}$ คือ สัมประสิทธิ์สหสัมพันธ์เชิงส่วนระหว่าง Y กับ X_1 โดยกำหนดให้ X_2 และ X_3 มีค่าคงที่ โดยที่ $r_{Y1.23}$ เป็นค่าที่แสดงความสัมพันธ์ระหว่าง Y กับ X_1 โดยให้ X_2 และ X_3 มีค่าคงที่ จึงเป็นค่าที่แสดงความสัมพันธ์ระหว่าง Y กับ X_1 เท่านั้น มิใช่ความสัมพันธ์ของ X_2 และ X_3 กับ Y

$r_{Y2.13}$ คือ สัมประสิทธิ์สหสัมพันธ์เชิงส่วนระหว่าง Y กับ X_2 โดยกำหนดให้ X_1 และ X_3 มีค่าคงที่

$r_{Y3.12}$ คือ สัมประสิทธิ์สหสัมพันธ์เชิงส่วนระหว่าง Y กับ X_3 โดยกำหนดให้ X_1 และ X_2 มีค่าคงที่โดยที่ $-1 \leq r_{Yi.jk} \leq 1$

2.4.2.4 การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย

ขั้นตอนที่ 1 การตรวจสอบการแจกแจงปกติของส่วนเหลือ (Normality of Residual) โดยใช้สถิติทดสอบลิลล์ปีร์สเช่นเดียวกับการตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย (สายชล, 2563)

สมมติฐาน

H_0 : ส่วนเหลือมาจากประชากรที่มีการแจกแจงปกติ

H_1 : ส่วนเหลือมาจากประชากรที่ไม่ได้มีการแจกแจงปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถิติทดสอบ

$$D = \max |S(x) - F(x)| \quad (2.12)$$

เมื่อ $F(x)$ คือ ความน่าจะเป็นสะสมภายใต้สมมติฐานว่าง

$S(x)$ คือ ความน่าจะเป็นสะสมของตัวอย่าง

เขตวิกฤต

ที่ระดับนัยสำคัญ α จะกำหนดค่าวิกฤตด้านขวา นั่นคือ

$$D > w_{1-\alpha}$$

จะปฏิเสธ H_0 ถ้าค่า D มีค่ามากกว่าค่าวิกฤตจากในตารางลิลลีโพรส์

การแก้ปัญหากรณีส่วนเหลือไม่มีการแจกแจงปกติ

ปัญหาของส่วนเหลือไม่มีการแจกแจงปกติมักจะเกิดขึ้นพร้อมกับปัญหาส่วนเหลือมีความแปรปรวนไม่คงที่ ซึ่งวิธีการแปลง (Transformation) เป็นอีกวิธีหนึ่งที่จะช่วยให้ส่วนเหลือมีความแปรปรวนคงที่และมีการแจกแจงปกติเพิ่มมากขึ้น ซึ่งการแปลงที่นิยมใช้คือ การแปลงบ็อก-ค็อก (Box-Cox Transformations) และวิธีการถ่วงน้ำหนัก (เริงใจ, 2547)

ขั้นตอนที่ 2 การตรวจสอบความเป็นอิสระของส่วนเหลือ (Independence of Residual) โดยใช้สถิติทดสอบเดอร์บิน-วัตสัน (Durbin-Watson Test) ซึ่งเป็นผู้เสนอสถิติทดสอบสหสัมพันธ์ในตัวของส่วนเหลือที่มีรูปแบบเป็น AR(1) ในการวิเคราะห์สมการการถดถอย ซึ่งเป็นวิธีที่นิยมนำมาใช้กันอย่างกว้างขวาง ดังนั้นจะเห็นได้จากการที่โปรแกรมสำเร็จรูปทั่วไปจะมีค่าสถิติทดสอบเดอร์บิน-วัตสันแสดงไว้ (โชติรส, 2544)

สมมติฐาน

$H_0: \rho = 0$ (ส่วนเหลือไม่มีความสัมพันธ์กันหรือเป็นอิสระกัน)

$H_1: \rho \neq 0$ (ส่วนเหลือมีความสัมพันธ์กันหรือไม่เป็นอิสระกัน)

โดยที่ ρ คือ ค่าสัมประสิทธิ์สหสัมพันธ์ของส่วนเหลือ

สถิติทดสอบ

$$d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2} \quad (2.13)$$

เขตวิกฤต

ถ้าค่าสถิติทดสอบเดอร์บิน-วัตสันมีค่าใกล้ 2 โดยอยู่ระหว่าง 1.5 - 2.5 จะสรุปได้ว่าส่วนเหลือ e_t เป็นอิสระต่อกัน

การแก้ไขปัญหาส่วนเหลือมีสหสัมพันธ์ในตัว

หลักการแก้ปัญหาสหสัมพันธ์ในตัวระหว่างส่วนเหลือ คือการนำค่า ρ มาทำการปรับตัวแบบ (Model) การถดถอยเพื่อให้ความสัมพันธ์ระหว่าง e_t และ e_{t-1} หดไป โดย ρ หาได้จากวิธีไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Cochran–Orcutt 2 ขั้นตอน ซึ่งมีวิธีในการหาค่า $\hat{\rho}$ จำนวน 2 ขั้นตอนคือ ขั้นตอนที่ 1 ทำการคำนวณส่วนเหลือที่ประมาณค่าได้ ($e_i = Y_i - \hat{Y}_i$) ที่ได้จากการประมาณสมการการถดถอย และ ขั้นตอนที่ 2 นำค่า e_i ที่ได้ไปใช้ในการคำนวณค่า $\hat{\rho}$ ดังนี้ (บัณฑิต, 2558)

$$\hat{\rho} = \frac{\sum_{i=2}^n e_i e_{i-1}}{\sum_{i=2}^n e_i^2} \quad (2.14)$$

โดยที่ $-1 \leq \hat{\rho} \leq 1$

ถ้าทดสอบแล้วพบว่ามีความเป็นอิสระของส่วนเหลือ ทำการประมาณค่า $\hat{\rho}$ ใหม่ ถ้าทดสอบแล้วพบว่าไม่มีความเป็นอิสระของส่วนเหลือ ก็จะสิ้นสุดวิธี Cochran–Orcutt

ขั้นตอนที่ 3 การตรวจสอบค่าเฉลี่ยของส่วนเหลือ (Mean of Residual) โดยใช้สถิติทดสอบที (t-test)

วิธีการทางสถิติที่ใช้ในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของส่วนเหลือว่าเท่ากับ 0 หรือไม่ (กัลยา, 2561) มีขั้นตอนดังนี้

สมมติฐาน

$$H_0 : \mu = 0$$

$$H_1 : \mu \neq 0$$

หรือ

$$H_0 : \text{ค่าเฉลี่ยของส่วนเหลือเท่ากับศูนย์}$$

$$H_1 : \text{ค่าเฉลี่ยของส่วนเหลือไม่เท่ากับศูนย์}$$

สถิติทดสอบ

$$t = \frac{\bar{X} - \mu}{S / \sqrt{n}} \quad (2.15)$$

เขตวิกฤต

$$\text{จะปฏิเสธ } H_0 \text{ ถ้า } t < -t_{\frac{\alpha}{2}, n-1} \text{ หรือ } t > t_{\frac{\alpha}{2}, n-1}$$

ขั้นตอนที่ 4 การตรวจสอบความเท่ากันของความแปรปรวนของส่วนเหลือ (Homoscedasticity of Residual)

การที่ความแปรปรวนของส่วนเหลือไม่คงที่จะมีผลทำให้การหาช่วงความเชื่อมั่นและการทดสอบสมมติฐานทำได้ไม่ถูกต้อง การทดสอบความคงที่ของความแปรปรวนของส่วนเหลือทำได้โดยการพล็อตกราฟระหว่างค่ามาตรฐานของส่วนเหลือกับค่าประมาณของตัวแปรตาม ถ้าพบว่าจุดต่างๆ ในรูปกระจายเป็นแบบสุ่มขนานกับแกนนอน จะสรุปได้ว่าความแปรปรวนของส่วนเหลือคงที่ แต่ถ้าพบว่าจุดต่างๆ ในรูปกระจายเป็นรูปปากแตร จะสรุปได้ว่าความแปรปรวนของส่วนเหลือไม่คงที่ หรือพิจารณาการใช้สถิติทดสอบบรูซส์-พาแกน (Breusch - Pagan Test) (จุฑาทิพย์ และ สุภาพร เอกสารนี้, 2562) การที่ส่งวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สถิติทดสอบบรูซส์-พาแกนเป็นการทดสอบความคงที่ของความแปรปรวนของการแจกแจงของส่วนเหลือ มีข้อสมมติคือความไม่คงที่ของความแปรปรวนเกิดขึ้นจากตัวแปรต่างๆ หลายตัวแปร ดังนี้

$$\sigma_i^2 = f(Z_1, Z_2, \dots, Z_m) \quad (2.16)$$

โดยที่ Z_i ตัวแปรใดๆ ก็ได้ ไม่ว่าจะเป็นตัวแปรอิสระภายในตัวแบบ (Model) หรือตัวแปรอื่นๆ ภายนอกตัวแบบ $i=1, 2, \dots, m$ และ m คือจำนวนตัวแปรที่คาดว่าจะจะเป็นสาเหตุให้ความแปรปรวนของการแจกแจงส่วนเหลือไม่คงที่

การทดสอบกำหนดให้ความสัมพันธ์ระหว่าง σ_i^2 กับ Z_i เป็นความสัมพันธ์เชิงเส้นตรง เขียนได้ดังนี้

$$\sigma_i^2 = \alpha_0 + \alpha_1 Z_1 + \alpha_2 Z_2 + \dots + \alpha_m X_m + V_i \quad (2.17)$$

โดยที่ V_i คือ ส่วนเหลือ

ค่า σ_i^2 ในสมการข้างต้นใช้ ρ_i เป็นตัวประมาณค่า โดยที่ $\rho_i = \frac{e_i^2}{\sigma^2}$

โดยที่ e_i คือ ส่วนเหลือที่ประมาณค่าได้จากตัวแบบการถดถอยที่ต้องการทดสอบและ

$$\sigma^2 = \left(\sum_{i=1}^n e_i^2 \right) / n \quad \text{แทนค่า } \sigma_i^2 \text{ ด้วย } \rho_i \text{ ในสมการ } \rho_i = \alpha_0 + \alpha_1 X_{11} + \alpha_2 X_{12} + \dots + \alpha_m X_m + V_i$$

ประมาณสมการด้วยวิธีกำลังสองน้อยที่สุด แล้วคำนวณค่าผลบวกกำลังสองของค่าคลาดเคลื่อน (Sum Square of Error : SSE) เพื่อใช้ในการคำนวณค่าสถิติทดสอบ χ^2 ดังนี้

$$\chi_{m-1}^2 = \frac{SSE}{2} \quad (2.18)$$

โดยที่ $m-1$ คือ องศาเสรี (Degree of Freedom)

สมมติฐาน

H_0 : ส่วนเหลือมีความแปรปรวนเท่ากัน

H_1 : ส่วนเหลือมีความแปรปรวนไม่เท่ากัน

การแก้ไขปัญหากรณีความแปรปรวนของส่วนเหลือไม่คงที่

กรณีที่ความแปรปรวนของส่วนเหลือไม่คงที่สามารถแก้ไขได้โดยการใช้วิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนัก (Weighted Least Square Method : WLS) ในการประมาณค่าพารามิเตอร์ เนื่องจากวิธี WLS เป็นวิธีที่ให้ความสำคัญหรือให้น้ำหนักแก่ข้อมูลไม่เท่ากัน นั่นคือ จะให้น้ำหนักเป็น 1 หมายความว่าข้อมูลที่มีค่าความแปรปรวนหรือมีค่าการกระจายสูงจะมีความสำคัญหรือมีน้ำหนักน้อยกว่าข้อมูลที่มีค่าความแปรปรวนต่ำ โดยในการประมาณค่าน้ำหนักของข้อมูลนั้นสามารถทำได้โดยการประมาณค่าน้ำหนักจากตัวแปร (Estimating Weights from a Variable) โดยผู้วิเคราะห์จะต้องศึกษาความสัมพันธ์ระหว่างค่าความแปรปรวนของตัวแปรตาม Y กับตัวแปรอื่นๆ ที่เกี่ยวข้องว่ามีความสัมพันธ์กันอย่างไร เพื่อนำไปใช้ในการหาค่าน้ำหนักของข้อมูลที่จะนำไปใช้ในการวิเคราะห์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอญุ่ให้เินหาไปใช้ประโยชน์ด้านการค้าต่อไป (กัลยา, 2548)

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5 งานวิจัยที่เกี่ยวข้อง

พรพิมล และชูชีพ (2554) ศึกษาผลของระดับการศึกษาที่มีต่อรายได้ของครัวเรือนไทย การศึกษาเป็นส่วนหนึ่งของการเพิ่มทุนมนุษย์และการศึกษาจะทำให้เกิดประสิทธิภาพในการผลิตที่มีคุณภาพมากขึ้นซึ่งนำไปสู่รายได้ที่เพิ่มขึ้น งานวิจัยนี้จึงทำการคำนวณรายได้ตลอดช่วงชีวิตโดยการประมาณการเส้นอายุ-รายได้ (Age-earning Profile) ด้วยสมการการถดถอยแสดงความสัมพันธ์ระหว่างรายได้กับตัวแปรอิสระต่างๆ โดยใช้ข้อมูลการสำรวจภาวะเศรษฐกิจและสังคมครัวเรือน ปี พ.ศ.2554 และศึกษาความคุ้มค่าโครงการลงทุนทางการศึกษาจาก 3 ตัวชี้วัด ได้แก่ 1) มูลค่าปัจจุบันสุทธิ (NPV) 2) อัตราส่วนผลประโยชน์ต่อทุน (BCR) 3) อัตราผลตอบแทนภายในโครงการ (IRR) โดยใช้ข้อมูลต้นทุนการศึกษาจากกองทุนเงินให้กู้ยืมเพื่อการศึกษา และกำหนดอัตราคิดลด (Discount Rate) เท่ากับร้อยละ 2.14 มาจากอัตราดอกเบี้ยเงินฝากของธนาคารพาณิชย์ซึ่งจะศึกษาเฉพาะผู้ที่สำเร็จการศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูงและผู้สำเร็จการศึกษาระดับปริญญาตรี ผลการวิจัยพบว่าตัวแปรที่มีผลต่อรายได้คืออายุ ประสบการณ์ สถานภาพ และสถานที่ทำงาน ในส่วนรายได้ของผู้สำเร็จการศึกษาระดับ ป.ตรีมีแนวโน้มที่เพิ่มขึ้นมากกว่าผู้ที่สำเร็จการศึกษาระดับ ปวส. ตลอดช่วงอายุการทำงานและเพิ่มขึ้นอย่างต่อเนื่อง ความคุ้มค่าโครงการลงทุนในการศึกษาของผู้สำเร็จการศึกษาระดับปวส. และระดับ ป.ตรีทั้ง 3 ตัวชี้วัด ได้ผลดังนี้ 1) NPV เท่ากับ 7,822 และ 37,049 บาท ตามลำดับ 2) BCR เท่ากับ 1.18 และ 1.86 ตามลำดับ 3) IRR เท่ากับร้อยละ 46.19 และ 27.25 ตามลำดับ จากผลข้างต้นทำให้ทราบว่า การลงทุนศึกษาทั้งในระดับ ปวส. และระดับ ป.ตรีเป็นโครงการที่มีความคุ้มค่าต่อการลงทุนโดยที่การลงทุนการศึกษาในระดับ ป.ตรีคุ้มค่ามากกว่า

ศศิธร และจรงค์ (2557) ศึกษาปัจจัยด้านทักษะของแรงงานที่ส่งผลต่อรายได้ของแรงงานในสถานประกอบการขนาดกลางและขนาดย่อมในเขตอำเภอเมือง จังหวัดขอนแก่น การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อรายได้ของแรงงานในสถานประกอบการขนาดกลางและขนาดย่อมในอำเภอเมือง จังหวัดขอนแก่น รวบรวมข้อมูลโดยแบบสอบถามและวิเคราะห์ด้วยวิธีการถดถอยกำลังสองน้อยที่สุด (Ordinary Least Squares) และการวิเคราะห์สมการการถดถอยแบบสองขั้นตอน (Two-Stage Least Squares) ผลการวิจัยพบว่า ปัจจัยด้านทักษะทางสติปัญญา (Cognitive Skill) ที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกของภาคการบริการและภาคการค้า ได้แก่ จำนวนปีการศึกษา และจำนวนปีประสบการณ์ในการทำงาน ปัจจัยทักษะเชิงพฤติกรรม (Non-Cognitive Skill) ที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกของภาคบริการคือ บุคลิกภาพแบบเปิดกว้าง (Openness) และปัจจัยทักษะเชิงพฤติกรรม (Non-Cognitive Skill) ที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกภาคการค้า ได้แก่ บุคลิกภาพแบบเปิดกว้าง (Openness) และบุคลิกภาพแบบเน้นประสิทธิภาพ (Conscientiousness) สำหรับบุคลิกภาพแบบเจ้าอารมณ์ (Neuroticism) มีผลต่อรายได้ของแรงงานในเชิงลบ ที่ระดับนัยสำคัญทางสถิติ 0.01

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

องอาจ (2558) ศึกษาปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ หนองคาย มุกดาหาร แม่สอด มีวัตถุประสงค์เพื่อศึกษาผลกระทบต่อรายได้ของคนในพื้นที่จากการพัฒนาเขตเศรษฐกิจพิเศษ และประเมินศักยภาพ โอกาส ข้อจำกัด อุปสรรคของการพัฒนาเขตเศรษฐกิจพิเศษ รวมทั้งวิเคราะห์การเปลี่ยนแปลงความเป็นอยู่ของผู้ที่อยู่อาศัยในบริเวณชายแดน จากการใช้ประโยชน์ในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ โดยใช้ข้อมูลรายได้ครัวเรือนจากสำมะโนประชากร (Household Socio-Economic Survey: SES) และข้อมูลการสำรวจภาวะแรงงาน (Labour Force Survey: LFS) ของสำนักงานสถิติแห่งชาติ รวมทั้งข้อมูลภูมิสารสนเทศและการใช้ประโยชน์ที่ดิน (Geographic Information System) ของกรมพัฒนาที่ดิน โดยวิเคราะห์สภาพการใช้ที่ดิน โดยใช้ข้อมูลทั้ง 3 หน่วยงาน ตั้งแต่ปี พ.ศ. 2548–2557 รวมเป็นระยะเวลา 10 ปี จึงนำมาวิเคราะห์ด้วยวิธีการวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) และนำมาสร้างตัวแบบการศึกษาเพื่อทำการวิเคราะห์ข้อมูลและทดสอบว่าปัจจัยแต่ละตัวมีความสัมพันธ์กับการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ หนองคาย มุกดาหาร แม่สอด จากการศึกษาข้อมูลในภาคครัวเรือนพบว่าปัจจัยที่มีผลต่อการเปลี่ยนแปลงรายได้ของครัวเรือนในพื้นที่เขตเศรษฐกิจพิเศษของหนองคาย มุกดาหาร และแม่สอด คือ จำนวนสมาชิกของครัวเรือน อายุ การศึกษา และปีประกาศเขตเศรษฐกิจพิเศษ ซึ่งเป็นไปตามสมมติฐานที่ว่า การพัฒนาเขตเศรษฐกิจพิเศษและรายได้ครัวเรือนมีความสัมพันธ์กัน เมื่อวิเคราะห์ข้อมูลรายได้ประชากรพบว่า ปัจจัยด้านแรงงานอื่นๆ ส่งผลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษทั้งสามพื้นที่และจังหวัดโดยรอบ โดยผลจากการวิเคราะห์ปัจจัยด้านแรงงานที่กำหนดในแบบจำลอง ได้แก่ ค่าแรงขั้นต่ำ จำนวนชั่วโมงทำงาน อายุ การศึกษา อาชีพ ประเภทอุตสาหกรรม ขนาดของอุตสาหกรรม และปีที่ประกาศเขตเศรษฐกิจพิเศษนั้นมีผลต่อรายได้ของประชากรในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ นอกจากนี้ ผลการเปรียบเทียบสัดส่วนการใช้ประโยชน์ที่ดินปี พ.ศ. 2549–2558 จากข้อมูล GIS ได้วิเคราะห์สัดส่วนการใช้ที่ดินของคนในพื้นที่ที่มีการเปลี่ยนแปลงจากนโยบายการพัฒนาเขตเศรษฐกิจพิเศษ โดยวิเคราะห์ข้อมูลภูมิสารสนเทศและการใช้ประโยชน์ที่ดิน Geographic Information System ของกรมพัฒนาที่ดิน เพื่อวิเคราะห์สภาพการใช้ที่ดิน เพื่อเปรียบเทียบการเปลี่ยนแปลงประเภทการใช้ประโยชน์ที่ดินและการเปลี่ยนแปลงถือครองที่ดินของประชาชนในพื้นที่ จากการวิเคราะห์ข้อมูลการใช้ที่ดินจังหวัดหนองคาย มุกดาหาร และแม่สอด การเปลี่ยนแปลงการใช้ประโยชน์ที่ระหว่างปี พ.ศ. 2549–2558 ซึ่งเป็นช่วงการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินที่คาบเกี่ยวในช่วงก่อนและหลังการประกาศเขตเศรษฐกิจพิเศษ จากผลการศึกษาพบว่า การใช้ประโยชน์ที่ดินในช่วงหลังการประกาศเขตเศรษฐกิจพิเศษของหนองคาย มุกดาหาร และแม่สอด ประเภทพื้นที่ชุมชนและสิ่งปลูกสร้างเพิ่มขึ้นก่อนการประกาศเขตเศรษฐกิจพิเศษ ซึ่งหมายความว่า การประกาศเขตเศรษฐกิจพิเศษอาจจะเป็นปัจจัยหนึ่งส่งผลให้เกิดการขยายตัวของภาคอุตสาหกรรม ทำให้มีการใช้พื้นที่ชุมชนและสิ่งปลูกสร้างมากขึ้นและการใช้พื้นที่เกษตรกรรมลดลง

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์หรือการเข้าถึงเพื่อการศึกษาค้นคว้าเท่านั้น มิใช่อยู่ใต้เงื่อนไขลิขสิทธิ์แบบใดก็ตาม
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อย่างไรก็ดี ควรมีการศึกษาเพิ่มเติมในปัจจัยอื่นๆ ที่เกี่ยวข้องและนำเครื่องมือทางเศรษฐมิติเชิงลึกมาใช้วิเคราะห์ในอนาคตต่อไป

สำนักงานเศรษฐกิจการเกษตร (2562) ศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ รายจ่าย การออม หนี้สิน ปัจจัยที่กำหนดการตัดสินใจก่อนนี้และความต้องการกู้ยืมของครัวเรือนเกษตร การเกษตร วิธีการคำนวณอาศัยเครื่องมือทางสถิติคือ การวิเคราะห์การถดถอยพหุคูณ (Multiple Regression Analysis) การวิเคราะห์โดยใช้แบบจำลองโลจิสติก (Logistic Model) การศึกษาวิจัยครั้งนี้เป็นการวิจัยในลักษณะที่มุ่งเน้นเพื่อให้ทราบปัจจัยที่มีอิทธิพลต่อรายได้ รายจ่าย การออม หนี้สิน ปัจจัยที่กำหนดการตัดสินใจก่อนนี้และความต้องการกู้ยืมของครัวเรือนเกษตร แนวทางการวิเคราะห์ปัจจัยที่ทางด้านเศรษฐกิจและสังคมของครัวเรือนเกษตรจำเป็นต้องคัดเลือกตัวแปร ทางด้านเศรษฐกิจและสังคมของครัวเรือนเกษตร เพื่อให้สมการสามารถคาดการณ์ตัวแปรที่สนใจได้ ผลการศึกษาในด้านรายจ่ายพบว่า รายได้เงินสดของครัวเรือน ความมั่งคั่งหรือทรัพย์สินสุทธิของครัวเรือน สวัสดิการมูลค่าหนี้สินทั้งสิ้นของครัวเรือน อายุของหัวหน้าครัวเรือน ระดับการศึกษาของหัวหน้าครัวเรือน สถานภาพสมรสของหัวหน้าครัวเรือน จำนวนสมาชิกในครัวเรือน ภูมิภาคของครัวเรือน หัวหน้าครัวเรือนปฏิบัติงานในเชิงเศรษฐกิจ และสมาชิกครัวเรือนที่พึ่งพิงมีอิทธิพลต่อค่าใช้จ่ายของครัวเรือนเกษตร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

วิธีการดำเนินงานวิจัย

ในการดำเนินงานวิจัยเรื่องปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทยนี้ ดำเนินตามขั้นตอนโดยแบ่งออกเป็น 6 ขั้นตอน คือ การจัดเตรียมข้อมูล การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย การสร้างสมการถดถอย การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย ตัวแปรที่ใช้ในการศึกษาและขั้นตอนการดำเนินงานวิจัย ตามรายละเอียดดังนี้

3.1 การจัดเตรียมข้อมูล (Data Preparation)

3.1.1 การเก็บรวบรวมข้อมูล การศึกษาในครั้งนี้ใช้ข้อมูลทุติยภูมิ (สำนักงานสถิติแห่งชาติ, 2566) จากสำนักงานสถิติแห่งชาติ เรื่องการสำรวจภาวะการทำงานของประชากร พ.ศ. 2566 ซึ่งทำการสำรวจอย่างต่อเนื่องเป็นประจำทุกปี ได้ดำเนินการจัดเก็บข้อมูลเป็นรายเดือน แล้วเสนอข้อมูลออกมาเป็นรายไตรมาส โดยที่ผู้อยู่ในวัยกำลังแรงงานมีอายุ 15 ปีขึ้นไป เพื่อให้สอดคล้องกับกฎหมายแรงงานเด็ก ปรับปรุงการจำแนกประเภทอาชีพและสถานภาพการทำงานให้สอดคล้องกับมาตรฐานสากล เพื่อให้สามารถเปรียบเทียบข้อมูลกันได้ และปรับเขตการปกครองจากเดิมสุขาภิบาลเป็นเขตเทศบาล ตามพระราชบัญญัติเปลี่ยนแปลงฐานะของสุขาภิบาลเป็นเทศบาล พ.ศ. 2542 ผู้เก็บรวบรวมข้อมูลระหว่างวันที่ 1-12 ของเดือนมกราคม-มีนาคม พ.ศ.2566 มีครัวเรือนที่เป็นตัวอย่างทั้งสิ้น 86,880 ครัวเรือน เป็นครัวเรือนตัวอย่างในเขตเทศบาล 49,248 ครัวเรือน (รวมครัวเรือนตัวอย่างในกรุงเทพมหานคร 4,800 ครัวเรือน) และนอกเขตเทศบาล 37,632 ครัวเรือน สำหรับวิธีการเก็บรวบรวมใช้การสัมภาษณ์หัวหน้าครัวเรือนหรือสมาชิกในครัวเรือนที่เป็นตัวอย่าง และใช้เครื่องคอมพิวเตอร์พกพาขนาดกลาง (Tablet) ในการบันทึกข้อมูลแทนแบบสอบถาม

3.1.2 การศึกษาข้อมูล จากข้อมูลที่ได้รวบรวมมีทั้งหมด 206,437 คน

โดยในแบบสอบถามจะแบ่งเป็น 5 ตอน และในแต่ละตอนจะมีคำถามย่อย ดังนี้

ตอนที่ 1 ลักษณะทั่วไปของสมาชิกในครัวเรือน

หัวข้อย่อยมีดังนี้ คือ ลำดับที่ ชื่อ-นามสกุล ความเกี่ยวข้องกับหัวหน้าครอบครัว เพศ อายุ และสถานภาพสมรส (ถามเฉพาะผู้ที่มีอายุตั้งแต่ 15 ปีขึ้นไป)

ตอนที่ 2 การศึกษา ถามเฉพาะผู้ที่มีอายุตั้งแต่ 6 ปี ขึ้นไป

หัวข้อย่อยมีดังนี้ คือ ขณะนี้เรียนชั้นใด จบการศึกษาสูงสุดระดับใด และจบการศึกษาสูงสุดสาขาวิชาใด

ตอนที่ 3 การทำงาน ถามเฉพาะผู้ที่มีอายุตั้งแต่ 15 ปีขึ้นไป

หัวข้อย่อยมีดังนี้ คือ ระหว่าง 7 วันก่อนการสัมภาษณ์ได้ทำงานที่ก่อให้เกิดรายได้/เงินเดือน/ค่าจ้าง/ผลกำไร/ผลตอบแทนหรือได้ทำงานในฐานะช่วยธุรกิจของครัวเรือนโดยไม่ได้รับค่าจ้างหรือไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าจ้างหรือไม่, ถึงแม้ว่าไม่ได้ทำงานระหว่าง 7 วันก่อนวันสัมภาษณ์ยังได้รับค่าจ้าง/เงินเดือนหรือผลประโยชน์อื่นๆ หรือผลกำไรจากงานหรือธุรกิจในช่วงที่ไม่ได้ทำงานหรือไม่ มีงาน/ธุรกิจกิจการที่จะกลับไปทำหรือไม่, หยุดทำงานมานานเท่าใด, ก่อนวันสัมภาษณ์ได้ทำงานทำ/สมัครงานบ้างหรือไม่, ได้ทำงานทำ/สมัครงาน อย่างไร, ระหว่าง 7 วัน ก่อนสัมภาษณ์พร้อมทำงานหรือไม่, เหตุผลที่ไม่พร้อมทำงาน, ทำไมจึงไม่ทำงานทำ, ทำงาน/พร้อมทำงานมานานเท่าใด, เคยทำงานมาก่อนหรือไม่, สาเหตุที่ออกจากงาน/หยุดทำงานครั้งสุดท้าย, ออกจากงาน/หยุดทำงานครั้งสุดท้ายมานานเท่าใด, ทำงานอะไร, กิจกรรมหลักที่ทำอยู่เป็นกิจกรรมหรือผลิอะไร, ทำงานในฐานะอะไร, กิจกรรมทำอยู่หรือทำครั้งสุดท้ายมีคนทำงานกี่คน, จำนวนชั่วโมงทำงานจริงของอาชีพที่บันทึก, จำนวนชั่วโมงทำงานจริงของอาชีพอื่นทุกอาชีพที่ทำในช่วง 7 วัน ก่อนสัมภาษณ์ และจำนวนชั่วโมงทำงานทั้งสิ้น

ตอนที่ 4 ความต้องการทำงานเพิ่ม

หัวข้อย่อยมีดังนี้ คือ ในช่วง 7 วัน ก่อนสัมภาษณ์ยังมีเวลาเหลือพอที่จะทำงานเพิ่มและต้องการจะทำงานเพิ่มหรือไม่, ต้องการทำงานเพิ่มขึ้นสัปดาห์ละกี่ชั่วโมง, ได้ทำงานทำ/ติดต่อสมัครงานเพิ่มบ้างหรือไม่ และทำไมไม่ได้ทำงานเพิ่ม

ตอนที่ 5 รายได้ของลูกจ้าง

หัวข้อย่อยมีดังนี้ คือ งานที่ทำอยู่ได้รับค่าจ้างในประเภทใด, ค่าจ้างที่ได้รับนั้นได้รับกี่บาทต่อชั่วโมงหรือต่อวันหรือต่อสัปดาห์, ค่าจ้างที่ได้รับประมาณเดือนละเท่าไร, ระหว่าง 30 วัน ก่อนสัมภาษณ์นอกจากค่าจ้างที่ได้รับแล้วยังได้รับโบนัสหรือไม่, ค่าล่วงเวลาเดือนละเท่าไร, เงินอื่นๆเดือนละเท่าไร, นอกจากค่าจ้างที่ได้รับยังได้รับรายได้หรือผลประโยชน์อื่นที่ไม่เป็นตัวเงินเป็นอาหาร, นอกจากค่าจ้างที่ได้รับยังได้รับรายได้หรือผลประโยชน์อื่นที่ไม่เป็นตัวเงินเป็นเสื้อผ้าหรือเครื่องแต่งกาย, นอกจากค่าจ้างที่ได้รับยังได้รับรายได้หรือผลประโยชน์อื่นที่ไม่เป็นตัวเงินเป็นที่อยู่อาศัย และนอกจากค่าจ้างที่ได้รับยังได้รับรายได้หรือผลประโยชน์อื่นที่ไม่เป็นตัวเงินเป็นอื่นๆ

ผู้วิจัยได้ทราบคำถามแยกประเภทคำถามเป็น 2 ประเภท คือ คนที่มีงานกับคนที่ไม่มีงานทำ จะได้รับชุดคำถามเริ่มต้นเหมือนกัน แล้วช่วงที่ 2 ของคำถามจะแยกเป็นคำถามของบุคคลแต่ละประเภท โดยคำถามทั้งหมดรวมมี 57 ข้อ แล้วข้อใดบ้างที่เราจะนำมาวิเคราะห์ โดยทางผู้ทำปัญหาพิเศษสนใจหัวข้อเกี่ยวกับรายได้ของลูกจ้างว่าปัจจัยใดบ้างที่มีอิทธิพล

3.1.3 การเลือกตัวแปรที่จะนำมาวิเคราะห์ จากข้อมูลทั้งหมดรวมมีตัวแปรทั้งสิ้น 60 ตัว

คณะผู้ทำปัญหาพิเศษได้คัดเลือกตัวแปรที่ใช้เบื้องต้นจากเงื่อนไขดังต่อไปนี้

- 1) เลือกผู้ตอบคำถามที่อยู่ในกลุ่มลูกจ้าง
- 2) ตัวแปรที่น่าจะมีความสัมพันธ์กับตัวแปรตาม โดยอ้างอิงจากงานวิจัยที่เกี่ยวข้อง

มีตัวแปร คือ อายุ จำนวนชั่วโมงในการทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเชิงวิชาการเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากเงื่อนไขที่กล่าวมานี้ทำให้ตัวแปรบางตัวที่อาจจะมีความสัมพันธ์กับตัวแปรตามถูกคัดออก เช่น ระดับการศึกษาจากงานวิจัยขององอาจ ดังนั้นจึงเหลือตัวแปรที่ใช้ได้ 15 ตัว แบ่งเป็นตัวแปรอิสระเชิงคุณภาพ 6 ตัว และตัวแปรอิสระเชิงปริมาณ 9 ตัว คือ ภูมิภาค เพศ เขตการปกครอง สถานภาพสมรส สังกัดของลูกจ้าง ประเภทค่าจ้าง อายุ จำนวนชั่วโมงในการทำงาน จำนวนสมาชิกในครัวเรือน ค่าล่วงเวลา โบนัส เงินอื่นๆ เงินค่าอาหารที่ได้รับ เงินค่าเสื้อผ้าที่ได้รับ และเงินค่าที่อยู่อาศัยที่ได้รับ ส่วนตัวแปรตามมีตัวแปรเชิงปริมาณ 1 ตัว ดังตาราง 3.1

ตารางที่ 3.1 ตัวแปร ชื่อตัวแปร และประเภทตัวแปร

ตัวแปร	ชื่อตัวแปร	หน่วย	ประเภทตัวแปร
X_1	ภูมิภาค	-	คุณภาพ
X_2	เพศ	-	คุณภาพ
X_3	เขตการปกครอง	-	คุณภาพ
X_4	สถานภาพสมรส	-	คุณภาพ
X_5	สังกัดของลูกจ้าง	-	คุณภาพ
X_6	ประเภทค่าจ้าง	-	คุณภาพ
X_7	อายุ	ปี	ปริมาณ
X_8	จำนวนชั่วโมงในการทำงาน	ชั่วโมงต่อสัปดาห์	ปริมาณ
X_9	จำนวนสมาชิกในครัวเรือน	คนต่อครัวเรือน	ปริมาณ
X_{10}	ค่าล่วงเวลา	บาทต่อเดือน	ปริมาณ
X_{11}	โบนัส	บาทต่อปี	ปริมาณ
X_{12}	เงินอื่นๆ	บาทต่อเดือน	ปริมาณ
X_{13}	เงินค่าอาหารที่ได้รับ	บาทต่อเดือน	ปริมาณ
X_{14}	เงินค่าเสื้อผ้าที่ได้รับ	บาทต่อเดือน	ปริมาณ
X_{15}	เงินค่าที่อยู่อาศัยที่ได้รับ	บาทต่อเดือน	ปริมาณ
Y	รายได้	บาทต่อเดือน	ปริมาณ

3.1.4 การทำความสะอาดและตรวจสอบข้อมูล จากตัวแปร 15 ตัว จำนวนคนตอบ 43,915 คน ทำการทำความสะอาดข้อมูลดังนี้

1) แทนค่าสูญหาย (Missing Value) โดยตัวแปรเชิงคุณภาพที่อยู่ในมาตราวัดเรียงลำดับ (Ordinal Scale) ใช้ค่ามัธยฐานของค่าใกล้เคียง (Median of Near by Points) ในการแทนที่ค่าสูญหาย ส่วนตัวแปรเชิงปริมาณที่อยู่ในมาตราวัดแบบช่วง (Interval Scale) หรือมาตราวัดแบบอัตราส่วน (Ratio Scale) ใช้ค่าเฉลี่ยของค่าใกล้เคียง (Mean of Near by Points) ในการแทนที่ค่าสูญหาย

โดยมีรายละเอียดดังตารางที่ 3.2

เอกสารนี้เป็นเอกสารสงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.2 ตัวแปรที่มีค่าสูญหาย สถิติเชิงพรรณนาที่แทนค่าสูญหายและค่าสูญหายที่ถูกแทนที่

ตัวแปรที่มี ค่าสูญหาย	จำนวน ทั้งหมด	จำนวนที่ สูญหาย	ประเภทของ ตัวแปร	สถิติเชิงพรรณนา ที่แทนค่าสูญหาย	ค่าสูญหายที่ถูก แทนที่
ประเภทค่าจ้าง	43,698	2	เรียงลำดับ	มัธยฐานของค่า ใกล้เคียง	4
เงินอื่นๆ	43,698	70	อัตราส่วน	ค่าเฉลี่ยของค่า ใกล้เคียง	684.6
เงินค่าอาหารที่ ได้รับ	43,698	42	อัตราส่วน	ค่าเฉลี่ยของค่า ใกล้เคียง	78
เงินค่าเสื้อผ้าที่ ได้รับ	43,698	33	อัตราส่วน	ค่าเฉลี่ยของค่า ใกล้เคียง	1.2
เงินค่าที่อยู่อาศัย ที่ได้รับ	43,698	8	อัตราส่วน	ค่าเฉลี่ยของค่า ใกล้เคียง	80.5

จากตารางที่ 3.2 จะพบว่า มีตัวแปรทั้งหมด 5 ตัวแปรที่มีค่าสูญหาย โดยทุกตัวมีจำนวนข้อมูลทั้งหมด 43,698 คน โดยที่ ประเภทค่าจ้าง (X_6) มีจำนวนค่าสูญหาย 2 คน ใช้มัธยฐานมาแทนค่า เนื่องจากค่าที่แทนคือ 4 เงินอื่นๆ (X_{12}) มีจำนวนค่าสูญหาย 70 คน ใช้ค่าเฉลี่ยมาแทนค่า โดยค่าที่แทนคือ 684.6 บาท เงินค่าอาหารที่ได้รับ (X_{13}) มีจำนวนค่าสูญหาย 42 คน ใช้ค่าเฉลี่ยมาแทนค่า โดยค่าที่แทนคือ 78 บาท เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) มีจำนวนค่าสูญหาย 33 คน ใช้ค่าเฉลี่ยมาแทนค่า โดยค่าที่แทนคือ 1.2 บาท เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) มีจำนวนค่าสูญหาย 8 คน ใช้ค่าเฉลี่ยมาแทนค่า โดยค่าที่แทนคือ 80.5 บาท

3.1.5 การกำจัดค่าผิดปกติ (Outlier) โดยพิจารณาจาก Mahalanobis แล้วตัดค่า 1-Mahalanobis ที่น้อยกว่า 0.01 ออกไป

3.1.6 การจัดเก็บข้อมูล เปลี่ยนชื่อตัวแปรอิสระแต่ละตัวให้เป็น X ตามด้วยเลขตามลำดับ กำหนด Y ให้เป็นตัวแปรตามที่เป็นเป้าหมาย แล้วบันทึกไฟล์เก็บไว้ใช้ข้อมูลต่อไป

3.2 การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย

3.2.1 การตรวจสอบการแจกแจงปกติของตัวแปรตาม (Normality) โดยใช้สถิติทดสอบ ลิลลีโพรส์ ซึ่งเป็นการทดสอบการแจกแจงของประชากรว่ามีการแจกแจงปกติหรือไม่ โดยต่างจากการทดสอบคอลโมโกรอฟ-สมิร์นอฟ คือ การทดสอบคอลโมโกรอฟ-สมิร์นอฟจะต้องกำหนดค่าเฉลี่ย เอกสารนี้ $\mu = \mu_0$ และส่วนเบี่ยงเบนมาตรฐาน $\sigma = \sigma_0$ ไว้ในสมมติฐาน H_0 ด้วย แต่การทดสอบลิลลีโพรส์จะไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไม่กำหนดค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของประชากร จึงต้องประมาณ μ ด้วย \bar{X} และประมาณ σ ด้วย S แทน

3.2.2 การตรวจสอบความสัมพันธ์เชิงเส้น (Linearity) โดยใช้สถิติทดสอบสัมประสิทธิ์สหสัมพันธ์ของเพียร์สัน ซึ่งเป็นการศึกษาความสัมพันธ์เชิงเส้นระหว่างตัวแปร 2 ตัวหรือข้อมูล 2 ชุด โดยมีข้อสมมติคือตัวแปรทั้งสองตัวต้องเป็นตัวแปรต่อเนื่อง อยู่ในมาตราอันตรภาคชั้นขึ้นไป และมีการแจกแจงปรกติสองตัวแปร โดยค่าสัมประสิทธิ์สหสัมพันธ์ของเพียร์สันจะมีค่าอยู่ระหว่าง -1 ถึง +1 ถ้ามีค่าเข้าใกล้ +1 หมายความว่าตัวแปรทั้งสองมีความสัมพันธ์อยู่ในระดับมาก แต่ถ้าสัมประสิทธิ์สหสัมพันธ์มีค่าเข้าใกล้ 0 หมายความว่าตัวแปรทั้งสองตัวมีความสัมพันธ์กันอยู่ระดับน้อยหรือไม่สัมพันธ์กันเลย

3.2.3 การตรวจสอบความสัมพันธ์เชิงเส้นพหุ (Multicollinearity) โดยใช้สถิติทดสอบ VIF ค่า VIF เป็นค่าที่จะสะท้อนให้เห็นถึงอิทธิพลรวมของตัวแปรอิสระในตัวแบบเส้นตรง ความสัมพันธ์เชิงเส้นพหุของตัวแปรในตัวแบบเส้นตรงคือการอธิบายว่าตัวแปรอิสระที่อยู่ในตัวแบบนั้นมีความสัมพันธ์กันอย่างไร โดยหากมีค่าไม่เกิน 10 แสดงว่าตัวแปรอิสระในสมการแต่ละตัวเป็นอิสระกัน

3.3 การสร้างสมการถดถอย

การเลือกตัวแปรโดยวิธีการถดถอยทีละขั้นเป็นวิธีการคัดเลือกที่ผสมผสานระหว่างวิธีการเลือกแบบไปข้างหน้าและวิธีการกำจัดถดถอยหลังเข้าด้วยกัน วิธีการถดถอยทีละขั้นสามารถทำได้ดังนี้

1) ใช้วิธีการเลือกแบบไปข้างหน้า คือ คัดเลือกตัวแปรอิสระ 1 ตัว เข้าไว้ในตัวแบบการถดถอย โดยเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตาม Y สูงสุด หรือที่ทำให้ตัวแบบการถดถอยมีค่าผลบวกกำลังสองของค่าคลาดเคลื่อน (SSE) ต่ำสุดไว้ในตัวแบบการถดถอยเป็นตัวแปรแรก จากนั้นคำนวณค่าสถิติเอฟบางส่วน (Partial F test) หรือค่าสถิติที (t -test) เพื่อทดสอบสมมติฐานว่าตัวแปรอิสระมีความสำคัญต่อการทำนายค่าของ Y หรือไม่ ถ้าตัวแปรอิสระนั้นไม่มีความสำคัญต่อการทำนายค่าของ Y จะสิ้นสุดการคัดเลือกตัวแปรอิสระเข้าไว้ในตัวแบบการถดถอย

2) ใช้วิธีการกำจัดแบบถดถอยหลัง คือคัดตัวแปรอิสระนั้นออกจากตัวแบบการถดถอยหากเป็นไปได้ โดยจะสามารถคัดตัวแปรอิสระที่อยู่ในตัวแบบออกได้ ตัวแปรอิสระนั้นมีความสำคัญต่อการทำนายจะไม่สามารถคัดออกจากตัวแบบการถดถอยได้

3) พิจารณาค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนระหว่างตัวแปรอิสระกับตัวแปรตาม Y โดยจะเลือกตัวแปรอิสระที่ให้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนกับตัวแปรตาม Y สูงที่สุด โดยตัวแปรอิสระที่ถูกพิจารณาตัวต่อมามีความสำคัญต่อการทำนายค่าของ Y เมื่อมีตัวแปรอิสระก่อนหน้าอยู่ในตัวแบบการถดถอยก่อน แล้วจะคัดตัวแปรอิสระที่ถูกพิจารณาเข้ามาอยู่ในตัวแบบเป็นตัวต่อมา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) ทำการทดสอบความมีนัยสำคัญของตัวแปรที่เข้ามาเป็นตัวสุดท้ายก่อน โดยพิจารณาจากค่าสถิติเอฟบางส่วน ถ้ามีค่าน้อยกว่าเกณฑ์ที่กำหนดไว้จะต้องนำตัวแปรนั้นออกจากตัวแบบ แต่ถ้าค่าสถิติเอฟบางส่วนมีค่ามากกว่าเกณฑ์ที่กำหนดไว้ ก็จะย้อนกลับไปทดสอบตัวแปรอิสระตัวก่อนหน้าที่ยังเข้ามาอยู่ในตัวแบบ โดยพิจารณาจากค่าสถิติเอฟบางส่วนเช่นเดียวกัน ทำซ้ำในขั้นตอนที่สามและสี่ต่อไปเรื่อยๆ จนกระทั่งไม่สามารถนำตัวแปรอิสระใดออกจากตัวแบบได้ หรือไม่สามารถนำตัวแปรอิสระใดเข้าสู่ตัวแบบได้อีก

3.4 การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย

3.4.1 การตรวจสอบการแจกแจงปกติของส่วนเหลือ (Normality of Residual) โดยใช้สถิติทดสอบลิลลิโพรส์เช่นเดียวกับการตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย

3.4.2 การตรวจสอบความเป็นอิสระของส่วนเหลือ (Independence of Residual) โดยใช้สถิติทดสอบเดอร์บิน-วัตสัน ซึ่งเป็นผู้เสนอสถิติทดสอบสหสัมพันธ์ในตัวของส่วนเหลือในการวิเคราะห์สมการถดถอย

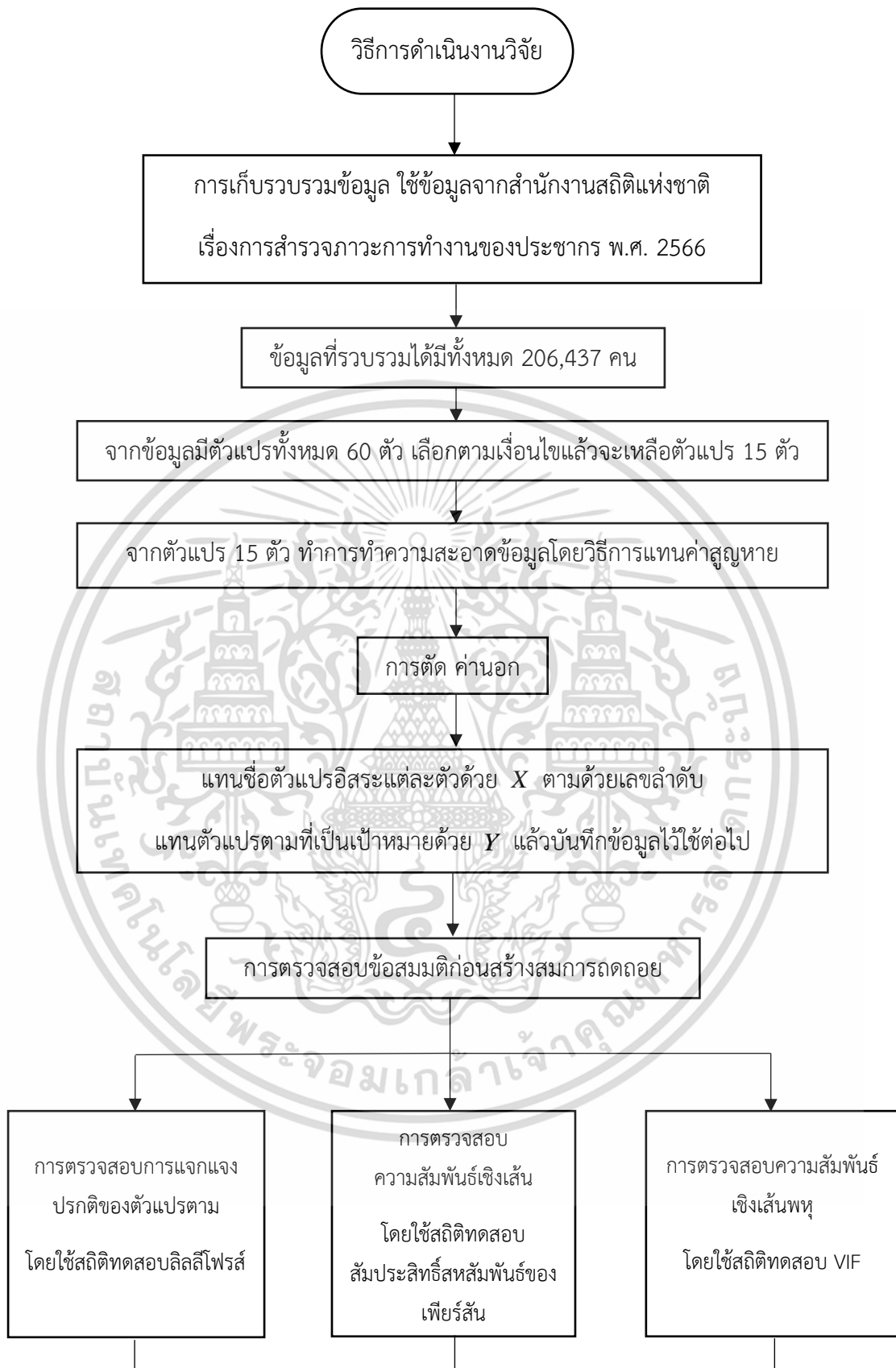
3.4.3 การตรวจสอบค่าเฉลี่ยของส่วนเหลือ (Mean of Residual) โดยใช้สถิติทดสอบที (t-test) วิธีการทางสถิติที่ใช้ในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของกลุ่มตัวอย่างหนึ่งกลุ่มกับประชากร หรือเปรียบเทียบระหว่างกลุ่มตัวอย่างสองกลุ่มที่อาจมีความสัมพันธ์กันหรือเป็นอิสระต่อกันก็ได้ โดยกลุ่มตัวอย่างต้องสุ่มมาจากประชากรที่มีการแจกแจงปกติและทราบค่าความแปรปรวนของประชากร

3.4.4 การตรวจสอบความเท่ากันของความแปรปรวนของส่วนเหลือ (Homoscedasticity of Residual) การที่ความแปรปรวนของส่วนเหลือไม่คงที่จะมีผลทำให้การหาช่วงความเชื่อมั่นและการทดสอบสมมติฐานทำได้ไม่ถูกต้อง การทดสอบความคงที่ของความแปรปรวนของส่วนเหลือทำได้โดยการพล็อตกราฟค่ามาตรฐานของส่วนเหลือกับค่าประมาณของตัวแปรตาม ถ้าพบว่าจุดต่างๆ ในรูปกระจายเป็นแบบสุ่มขนานกับแกนนอน จะสรุปได้ว่าความแปรปรวนของส่วนเหลือคงที่ แต่ถ้าพบว่าจุดต่างๆ ในรูปกระจายเป็นรูปปากแตร จะสรุปได้ว่าความแปรปรวนของส่วนเหลือไม่คงที่ หรือพิจารณาจากการใช้สถิติทดสอบบรูซส์-พาแกน

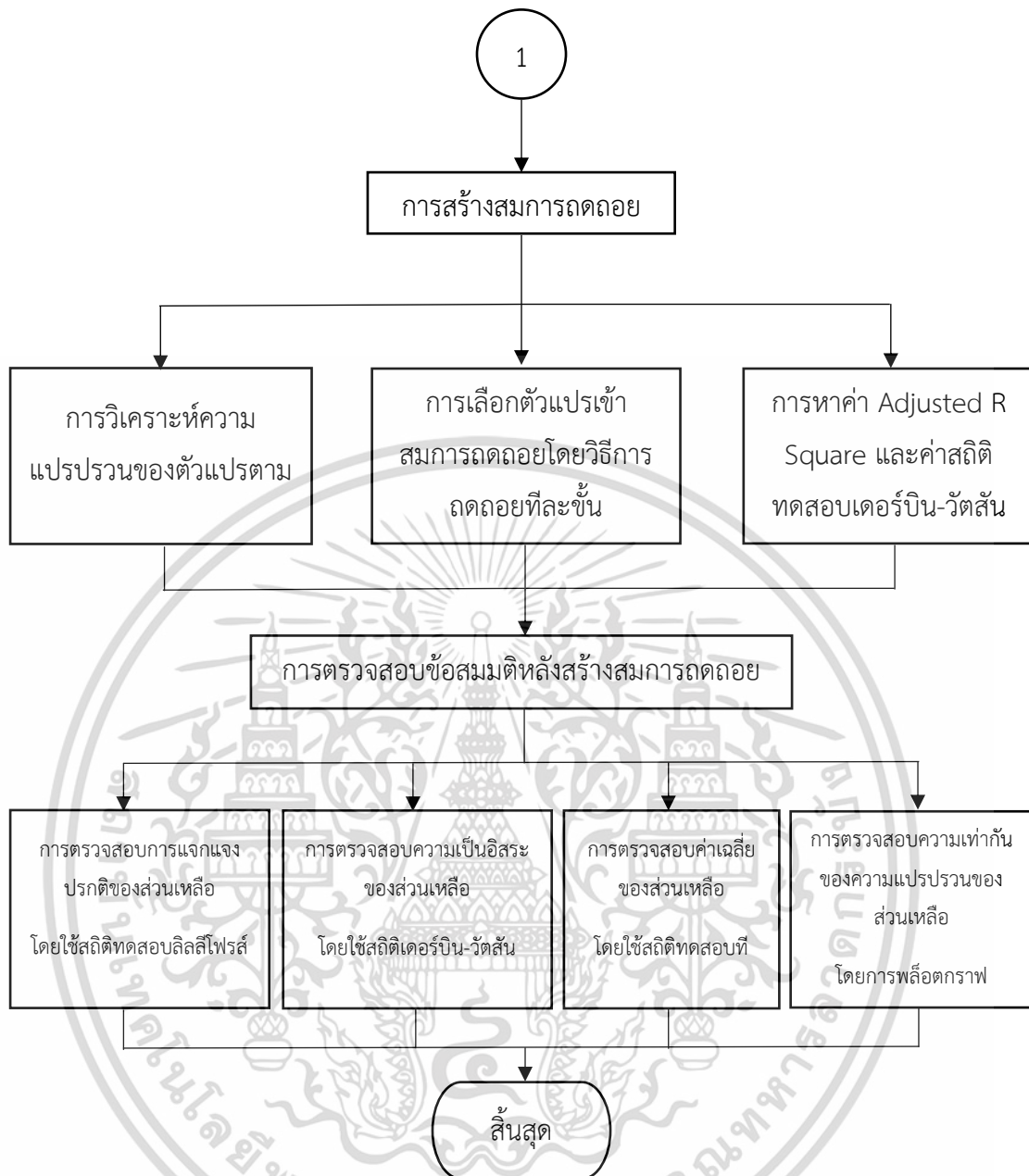
3.5 ขั้นตอนการดำเนินงานวิจัย

แผนผังรูปที่ 3.1 แผนผังแสดงขั้นตอนการดำเนินงานวิจัย ประกอบด้วย การจัดเตรียมข้อมูล การเก็บรวบรวมข้อมูล แผนการเลือกตัวอย่าง การประเมินข้อมูล การเลือกตัวแปรที่จะนำมาวิเคราะห์ การทำความสะอาดและตรวจสอบข้อมูล การแทนค่าสูญหาย การจัดเก็บข้อมูล การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย การสร้างสมการถดถอย การตรวจสอบข้อสมมติหลังสร้างสมการถดถอย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานี้เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.1 แผนผังแสดงขั้นตอนการดำเนินงานวิจัย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการวิจัยและการอภิปรายผล

4.1 ผลการวิเคราะห์ข้อมูลทางสถิติของตัวแปรอิสระ

จากปัจจัยที่มีอิทธิพลกับเงินเดือนของลูกจ้าง เป็นข้อมูลทุติยภูมิที่รวบรวมมาจากสำนักงานสถิติแห่งชาติ สถิติเชิงพรรณนาที่มีอิทธิพลต่อรายได้ของลูกจ้าง โดยจัดเป็นสถิติของตัวแปรอิสระและตัวแปรตาม ได้ผลดังตารางที่ 4.1 และ 4.2

ตารางที่ 4.1 ค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานของตัวแปรตามและตัวแปรอิสระเชิงปริมาณ

ตัวแปร	ค่าเฉลี่ย	ส่วนเบี่ยงเบนมาตรฐาน
อายุ (X_7)	40.47	11.870
จำนวนชั่วโมงในการทำงาน (X_8)	43.69	9.178
จำนวนสมาชิกในครัวเรือน (X_9)	3.11	1.568
ค่าล่วงเวลา (X_{10})	266.00	961.045
โบนัส (X_{11})	2.30	80.367
เงินอื่นๆ (X_{12})	520.12	1,799.436
เงินค่าอาหารที่ได้รับ (X_{13})	49.41	234.575
เงินค่าเสื้อผ้าที่ได้รับ (X_{14})	0.00	0.00
เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15})	39.93	254.485
รายได้ของลูกจ้าง (Y)	14,481.21	10,897.750

จากตารางที่ 4.1 พบว่า อายุ (X_7) มีค่าเฉลี่ย 40.47 ปี และมีส่วนเบี่ยงเบนมาตรฐาน 11.870 ปี จำนวนชั่วโมงในการทำงาน (X_8) มีค่าเฉลี่ย 43.69 ชั่วโมงต่อสัปดาห์ และมีส่วนเบี่ยงเบนมาตรฐาน 9.178 ชั่วโมงต่อสัปดาห์ จำนวนสมาชิกในครัวเรือน (X_9) มีค่าเฉลี่ย 3.11 คนต่อครัวเรือน และมีส่วนเบี่ยงเบนมาตรฐาน 1.568 คนต่อครัวเรือน ค่าล่วงเวลา (X_{10}) มีค่าเฉลี่ย 266.00 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 961.045 บาทต่อเดือน โบนัส (X_{11}) มีค่าเฉลี่ย 2.30 บาทต่อปี และมีส่วนเบี่ยงเบนมาตรฐาน 80.367 บาทต่อปี เงินอื่นๆ (X_{12}) มีค่าเฉลี่ย 520.12 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 1,799.436 บาทต่อเดือน เงินค่าอาหารที่ได้รับ (X_{13}) มีค่าเฉลี่ย 49.41 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 234.575 บาทต่อเดือน เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) มีค่าเฉลี่ย 0.00 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 0.00 บาทต่อเดือน เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) มีค่าเฉลี่ย 39.93 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 254.485 บาทต่อเดือน

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการศึกษาวิจัยเท่านั้น ไม่ควรนำข้อมูลไปใช้
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บาทต่อเดือน รายได้ของลูกค้า (Y) มีค่าเฉลี่ย 14,481.21 บาทต่อเดือน และมีส่วนเบี่ยงเบนมาตรฐาน 10,897.750 บาทต่อเดือน

ตารางที่ 4.2 ความถี่และร้อยละของตัวแปรอิสระเชิงคุณภาพ

ตัวแปร		ความถี่	ร้อยละ
ภูมิภาค (X_1)	กรุงเทพมหานคร	3,389	8.1
	ภาคกลาง	16,450	39.1
	ภาคเหนือ	6,636	15.8
	ภาคตะวันออกเฉียงเหนือ	7,585	18.0
	ภาคใต้	7,978	19.0
	ผลรวม	42,038	100%
เพศ (X_2)	ชาย	22,355	53.2
	หญิง	19,683	46.8
	ผลรวม	42,038	100%
เขตการปกครอง (X_3)	ในเขตเทศบาล	24,731	58.8
	นอกเขตเทศบาล	17,307	41.2
	ผลรวม	42,038	100%
สถานภาพสมรส (X_4)	โสด	26,167	62.2
	แต่งงานแล้ว	14,188	33.8
	ไม่ทราบสถานะ	1,683	4.0
	ผลรวม	42,038	100%
สังกัดของลูกค้า (X_5)	ลูกค้ารัฐ	10,134	24.1
	ลูกค้าวิสาหกิจ	741	1.8
	ลูกค้าเอกชน	31,163	74.1
	ผลรวม	42,038	100%
ประเภทค่าจ้าง (X_6)	รายชั่วโมง	50	1
	รายวัน	13,080	31.1
	รายเดือน	26,446	62.9
	อื่นๆ	2,462	5.9
	ผลรวม	42,038	100%

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตาราง 4.2 จะเห็นได้ว่าคนที่ตอบแบบสอบถามจะมีจำนวนคนและร้อยละแยกตามหัวข้อได้ดังนี้ ภูมิภาค (X_1) แบ่งเป็นกรุงเทพมหานครจำนวน 3,389 คน คิดเป็นร้อยละ 8.1 ภาคกลางจำนวน 16,450 คน คิดเป็นร้อยละ 39.1 ภาคเหนือจำนวน 6,636 คน คิดเป็นร้อยละ 15.8 ภาคตะวันออกเฉียงเหนือจำนวน 7,585 คน คิดเป็นร้อยละ 18.0 และภาคใต้จำนวน 7,978 คน คิดเป็นร้อยละ 19.0

เพศ (X_2) แบ่งเป็นเพศชายจำนวน 22,355 คน คิดเป็นร้อยละ 53.2 และเพศหญิงจำนวน 19,683 คน คิดเป็นร้อยละ 46.8

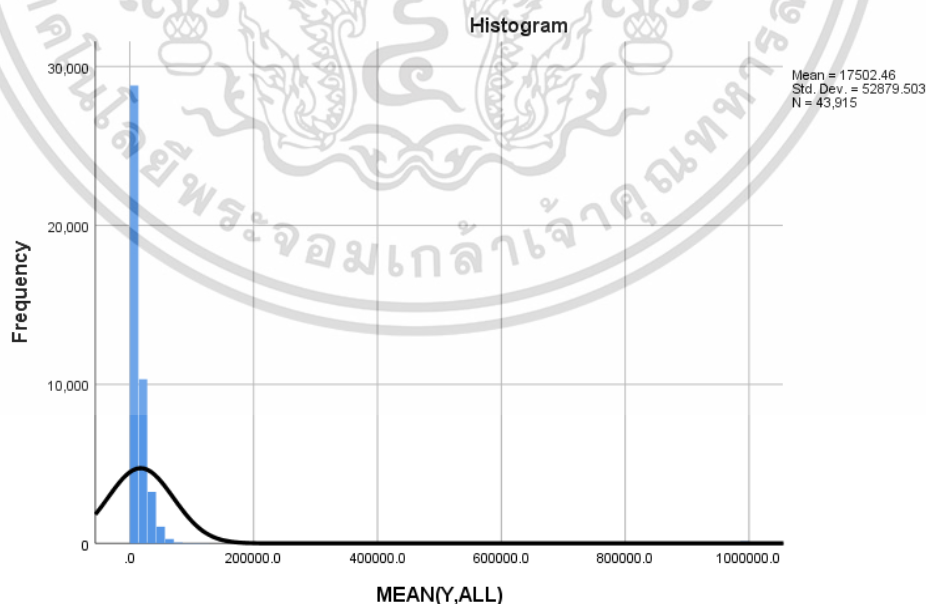
เขตการปกครอง (X_3) แบ่งเป็นในเขตเทศบาลจำนวน 24,731 คน คิดเป็นร้อยละ 58.8 และนอกเขตเทศบาลจำนวน 17,307 คน คิดเป็นร้อยละ 41.2

สถานภาพสมรส (X_4) แบ่งเป็นโสดจำนวน 26,167 คน คิดเป็นร้อยละ 62.2 แต่งงานแล้วจำนวน 14,188 คน คิดเป็นร้อยละ 33.8 และไม่ทราบสถานะจำนวน 1,683 คน คิดเป็นร้อยละ 4.0

สังกัดของลูกจ้าง (X_5) แบ่งเป็นลูกจ้างภาครัฐจำนวน 10,134 คน คิดเป็นร้อยละ 24.1 ลูกจ้างรัฐวิสาหกิจจำนวน 741 คน คิดเป็นร้อยละ 1.8 และลูกจ้างภาคเอกชนจำนวน 31,163 คน คิดเป็นร้อยละ 74.1

ประเภทค่าจ้าง (X_6) แบ่งเป็นค่าจ้างรายชั่วโมงจำนวน 50 คน คิดเป็นร้อยละ 1 ค่าจ้างรายวันจำนวน 13,080 คน คิดเป็นร้อยละ 31.1 ค่าจ้างรายเดือนจำนวน 26,446 คน คิดเป็นร้อยละ 62.9 และค่าจ้างอื่นๆ จำนวน 2,462 คน คิดเป็นร้อยละ 5.9

สถิติเชิงพรรณนาในรูปแบบฮิสโตแกรม (Histogram) ของรายได้ของลูกจ้าง ดังรูปที่ 4.1



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ 4.1 ฮิสโตแกรมของรายได้ของลูกจ้างนำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูป 4.1 จะเห็นได้ว่าจ้างรายได้ของลูกจ้างไม่มีการแจกแจงปกติ มีช่วงหนึ่งที่สูงกว่าช่วงอื่นๆ โดยคนส่วนใหญ่มีรายได้ประมาณ 17,502.46 บาท และคนที่มียาได้สูงกว่านั้นก็จะลดลงตามลำดับ โดยยังมีรายได้สูงขึ้น จำนวนคนก็จะยังมีจำนวนน้อยลง

4.2 การตรวจสอบข้อสมมติก่อนสร้างสมการถดถอย

4.2.1 การตรวจสอบการแจกแจงปกติของตัวแปรตาม (Normality)

การตรวจสอบว่ารายได้ของลูกจ้าง (Y) มีการแจกแจงปกติหรือไม่ โดยใช้สถิติทดสอบลิลีโพรส์ ได้ผลดังตารางที่ 4.1

ตารางที่ 4.3 ผลการทดสอบการแจกแจงปกติของเงินเดือนด้วยวิธีลิลีโพรส์

รายได้ของลูกจ้าง	Lilliefors	
	Statistic	p-value
Y	0.384	0.000

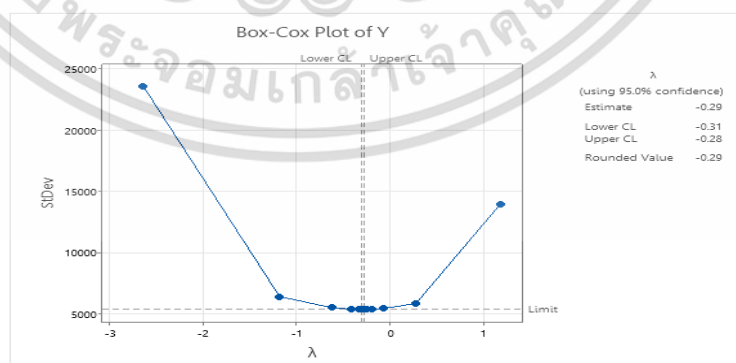
สมมติฐาน

H_0 : รายได้ของลูกจ้างมีการแจกแจงปกติ

H_1 : รายได้ของลูกจ้างไม่มีการแจกแจงปกติ

สถิติทดสอบ

จากตารางที่ 4.3 จะพบว่าStatistic = 0.384 และมี p-value = 0.000 มีค่าน้อยกว่า $\alpha = 0.05$ แสดงว่าปฏิเสธสมมติฐานว่าง H_0 นั่นคือรายได้ของลูกจ้างไม่มีการแจกแจงปกติ จึงแก้ไขปัญหาด้วยการแปลงข้อมูลรายได้ของลูกจ้างด้วยวิธีการแปลงบ็อก-ค็อก ดังรูปที่ 4.2



รูปที่ 4.2 ผลการแปลงค่ารายได้ของลูกจ้างด้วยวิธีการแปลงบ็อก-ค็อก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 4.2 ค่าประมาณของ $\lambda = -0.29$ ซึ่งใกล้เคียงกับ -0.50 จึงทำการแปลงรายได้ของลูกจ้างโดยใช้ส่วนกลับของรากที่สองของรายได้ของลูกจ้าง ได้เป็น $\frac{1}{\sqrt{Y}}$ และนำไปทำการทดสอบการแจกแจงปรกติอีกครั้งหนึ่งดังนี้

ตารางที่ 4.4 ผลการทดสอบการแจกแจงปรกติของ $\frac{1}{\sqrt{Y}}$ ด้วยวิธีลิลีโฟร์ส

รายได้ของลูกจ้าง	Lilliefors	
	Statistic	p-value
$\frac{1}{\sqrt{Y}}$	0.202	0.000

สมมติฐาน

$$H_0: \frac{1}{\sqrt{Y}} \text{ มีการแจกแจงปรกติ}$$

$$H_1: \frac{1}{\sqrt{Y}} \text{ ไม่มีการแจกแจงปรกติ}$$

สถิติทดสอบ

จากตารางที่ 4.4 จะพบว่า Statistic = 0.202 และ p-value = 0.000 มีค่าน้อยกว่า $\alpha = 0.05$ แสดงว่าปฏิเสธสมมติฐานว่าง H_0 นั่นคือส่วนกลับของรากที่สองของรายได้ของลูกจ้างไม่มีการแจกแจงปรกติ

จากทฤษฎีบทขีดจำกัดส่วนกลางที่กล่าวว่าถ้าสุ่มตัวอย่างขนาดใหญ่จากประชากรที่มีการแจกแจงใดๆ จะพบว่าค่าเฉลี่ยตัวอย่างจะมีการแจกแจงเข้าใกล้การแจกแจงปรกติ ในปัญหาพิเศษนี้สุ่มตัวอย่างขนาดใหญ่ จึงอนุมานว่ารายได้ของลูกจ้างมีการแจกแจงเข้าใกล้การแจกแจงปรกติ

4.2.2 การตรวจสอบความสัมพันธ์เชิงเส้น (Linearity)

สมมติฐาน

$$H_0: \text{รายได้ของลูกจ้างและภูมิภาคไม่มีความสัมพันธ์กันในเชิงเส้น}$$

$$H_1: \text{รายได้ของลูกจ้างและภูมิภาคมีความสัมพันธ์กันในเชิงเส้น}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.5 ค่าสหสัมพันธ์ของเพียร์สันระหว่างตัวแปรอิสระที่เป็นเชิงปริมาณและตัวแปรตามที่เป็นเชิงปริมาณ

ตัวแปร		X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	Y
X_7	Pearson Correlation	1	-.136	-.112	-.079	-.006	.085	-.051	.	-.007	.143
	p-value		.000	.000	.000	.218	.000	.000	.	.134	.000
X_8	Pearson Correlation	-.136	1	-.062	.363	.004	-.086	.110	.	-.015	-.127
	p-value	.000		.000	.000	.458	.000	.000	.	.002	.000
X_9	Pearson Correlation	-.112	-.062	1	-.081	-.002	-.042	-.020	.	-.124	-.079
	p-value	.000	.000		.000	.750	.000	.000	.	.000	.000
X_{10}	Pearson Correlation	-.079	.363	-.081	1	.002	.042	.058	.	-.005	.018
	p-value	.000	.000	.000		.614	.000	.000	.	.317	.000
X_{11}	Pearson Correlation	-.006	.004	-.002	.002	1	-.008	.006	.	-.004	.007
	p-value	.218	.458	.750	.614		.103	.217	.	.356	.176
X_{12}	Pearson Correlation	.085	-.086	-.042	.042	-.008	1	.020	.	.026	.395
	p-value	.000	.000	.000	.000	.103		.000	.	.000	.000
X_{13}	Pearson Correlation	-.051	.110	-.020	.058	.006	.020	1	.	.018	-.057
	p-value	.000	.000	.000	.000	.217	.000		.	.000	.000
X_{14}	Pearson Correlation	1	.	.
	p-value
X_{15}	Pearson Correlation	-.007	-.015	-.124	-.005	-.004	.026	.018	.	1	.034
	p-value	.134	.002	.000	.317	.356	.000	.000	.	.	.000
Y	Pearson Correlation	.143	-.127	-.079	.018	.007	.395	-.057	.	.034	1
	p-value	.000	.000	.000	.000	.176	.000	.000	.	.000	

จากตารางที่ 4.5 สามารถอธิบายได้ดังนี้

อายุ (X_7) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นโบนัส (X_{11}) มีค่า p-value = 0.218 > $\alpha = 0.05$ ค่าที่อยู่อาศัย (X_{15}) มีค่า p-value = 0.134 > $\alpha = 0.05$ แสดงว่าอายุ (X_7) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้นยกเว้น โบนัส (X_{11})

และค่าที่อยู่อาศัย (X_{15}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จำนวนชั่วโมงในการทำงาน (X_8) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นโบนัส (X_{11}) มีค่า $p\text{-value} = 0.458 > \alpha = 0.05$ แสดงว่าจำนวนชั่วโมงในการทำงาน (X_8) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้น โบนัส (X_{11}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

จำนวนสมาชิกในครัวเรือน (X_9) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นโบนัส (X_{11}) มีค่า $p\text{-value} = 0.750 > \alpha = 0.05$ แสดงว่าจำนวนสมาชิกในครัวเรือน (X_9) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้น โบนัส (X_{11}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

ค่าล่วงเวลา (X_{10}) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้น โบนัส (X_{11}) มีค่า $p\text{-value} = 0.614 > \alpha = 0.05$ เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) มีค่า $p\text{-value} = 0.317 > \alpha = 0.05$ แสดงว่าค่าล่วงเวลา (X_{10}) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้นโบนัส (X_{11}) และเงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

โบนัส (X_{11}) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นเพศ (X_2) มีค่า $p\text{-value} = 0.133 > \alpha = 0.05$ อายุ (X_7) มีค่า $p\text{-value} = 0.214 > \alpha = 0.05$ จำนวนสมาชิกในครัวเรือน (X_9) มีค่า $p\text{-value} = 0.162 > \alpha = 0.05$ ค่าล่วงเวลา (X_{10}) มีค่า $p\text{-value} = 0.841 > \alpha = 0.05$ เงินอื่นๆ (X_{12}) มีค่า $p\text{-value} = 0.988 > \alpha = 0.05$ เงินค่าอาหารที่ได้รับ (X_{13}) มีค่า $p\text{-value} = 0.798 > \alpha = 0.05$ เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) มีค่า $p\text{-value} = 0.967 > \alpha = 0.05$ และเงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) มีค่า $p\text{-value} = 0.400 > \alpha = 0.05$ แสดงว่าโบนัส (X_{11}) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้นเพศ (X_2) อายุ (X_7) จำนวนสมาชิกในครัวเรือน (X_9) ค่าล่วงเวลา (X_{10}) เงินอื่นๆ (X_{12}) เงินค่าอาหารที่ได้รับ (X_{13}) เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) และเงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

เงินอื่นๆ (X_{12}) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้น โบนัส (X_{11}) มีค่า $p\text{-value} = 0.103 > \alpha = 0.05$ แสดงว่าเงินอื่นๆ (X_{12}) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้น โบนัส (X_{11}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

เงินค่าอาหารที่ได้รับ (X_{13}) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้น โบนัส (X_{11}) มีค่า $p\text{-value} = 0.217 > \alpha = 0.05$ แสดงว่าเงินค่าอาหารที่ได้รับ (X_{13}) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น โบนัส (X_{11}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นอายุ (X_7) มีค่า $p\text{-value} = 0.134 > \alpha = 0.05$ ค่าล่วงเวลาต่อเดือน (X_{10}) มีค่า $p\text{-value} = 0.317 > \alpha = 0.05$ และโบนัส (X_{11})

มีค่า $p\text{-value} = 0.356 > \alpha = 0.05$ แสดงว่าเงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้น อายุ (X_7) ค่าล่วงเวลาต่อเดือน (X_{10}) และโบนัส (X_{11})

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์ของโรงเรียนเพื่อการศึกษาเท่านั้น เมื่อผู้ยืมได้เห็นใบแจ้งยืมเรียบร้อยแล้ว กรุณา
ไม่ว่ากรณใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ความสัมพันธ์กันในเชิงเส้น ยกเว้นอายุ (X_7) ค่าล่วงเวลาต่อเดือน (X_{10}) และโบนัส (X_{11}) ที่ไม่มี
ความสัมพันธ์กันในเชิงเส้น

รายได้ของลูกจ้าง (Y) จะพบว่า $p\text{-value} < \alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นโบนัส (X_{11}) มีค่า $p\text{-value} = 0.176 > \alpha = 0.05$ แสดงว่ารายได้ของลูกจ้าง (Y) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้นโบนัส (X_{11}) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

ตารางที่ 4.6 ค่าสหสัมพันธ์ของสเปียร์แมนระหว่างตัวแปรอิสระที่เป็นเชิงคุณภาพและตัวแปรตามที่เป็นเชิงปริมาณ

ตัวแปร		X_1	X_2	X_3	X_4	X_5	X_6	Y
X_1	Spearman's Correlation	1.000	-.022	.089	.089	-.148	.042	-.203
	p-value		.000	.000	.000	.000	.000	.000
X_2	Spearman's Correlation	-.022	1.000	-.036	.012	-.076	.087	.034
	p-value	.000		.000	.018	.000	.000	.000
X_3	Spearman's Correlation	.089	-.036	1.000	.025	.043	-.043	-.169
	p-value	.000	.000		.000	.000	.000	.000
X_4	Spearman's Correlation	.089	.012	.025	1.000	-.183	.080	.081
	p-value	.000	.018	.000		.000	.000	.000
X_5	Spearman's Correlation	-.148	-.076	.043	-.183	1.000	-.279	-.329
	p-value	.000	.000	.000	.000		.000	.000
X_6	Spearman's Correlation	.042	.087	-.043	.080	-.279	1.000	.360
	p-value	.000	.000	.000	.000	.000		.000
Y	Spearman's Correlation	-.203	.034	-.169	-.081	-.329	.360	1.000
	p-value	.000	.000	.000	.000	.000	.000	

เอกสารนี้เป็นเอกสารสงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปเผยแพร่ประโยชน์สาธารณะ
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากตารางที่ 4.6 สามารถอธิบายได้ดังนี้

ภูมิภาค (X_1) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว แสดงว่าภูมิภาค (X_1) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น

เพศ (X_2) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นสถานภาพสมรส (X_4) มีค่า p-value = 0.018 > $\alpha = 0.05$ แสดงว่าเพศ (X_2) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้นสถานภาพสมรส (X_4) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

เขตการปกครอง (X_3) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว แสดงว่าเขตการปกครอง (X_3) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น

สถานภาพสมรส (X_4) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว ยกเว้นเพศ (X_2) มีค่า p-value = 0.018 > $\alpha = 0.05$ แสดงว่าสถานภาพสมรส (X_4) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น ยกเว้นเพศ (X_2) ที่ไม่มีความสัมพันธ์กันในเชิงเส้น

สังกัดของลูกจ้าง (X_5) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว แสดงว่าสังกัดของลูกจ้าง (X_5) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น

ค่าจ้าง (X_6) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว แสดงว่าค่าจ้าง (X_6) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น

รายได้ของลูกจ้างในประเทศไทย (Y) จะพบว่า p-value < $\alpha = 0.05$ มีความสัมพันธ์กับตัวแปรทุกตัว แสดงว่ารายได้ของลูกจ้างในประเทศไทย (Y) กับตัวแปรทุกตัวมีความสัมพันธ์กันในเชิงเส้น

4.2.3 การตรวจสอบความสัมพันธ์เชิงเส้นพหุ (Multicollinearity)

ตารางที่ 4.7 ค่า VIF และ Tolerance ของตัวแปรอิสระ

ตัวแปรอิสระ	ค่า VIF	ค่า Tolerance
เงินอื่นๆต่อเดือน (X_{12})	1.078	0.928
สังกัดของลูกจ้าง (X_5)	1.396	0.716
ประเภทค่าจ้าง (X_6)	1.123	0.890
ภูมิภาค (X_1)	1.074	0.931
เขตการปกครอง (X_3)	1.029	0.972

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.7 ค่า VIF และ Tolerance ของตัวแปรอิสระ (ต่อ)

อายุ (X_7)	1.048	0.954
จำนวนสมาชิกในครัวเรือน (X_9)	1.049	0.953
เงินค่าอาหารที่ได้รับ (X_{13})	1.030	0.971
เพศ (X_2)	1.019	0.981
โบนัส (X_{11})	1.000	1.000
จำนวนชั่วโมงในการทำงาน (X_8)	1.290	0.775

จากตารางที่ 4.7 จะพบว่าค่า VIF ของตัวแปรอิสระทุกตัวมีค่าน้อยกว่า 10 และค่า Tolerance ของตัวแปรอิสระทุกตัวมีค่ามากกว่า 0.7 ซึ่งไม่ใกล้เคียง 0 แสดงว่าตัวแปรอิสระไม่มีความสัมพันธ์เชิงเส้นพหุ

4.3 การสร้างสมการถดถอย

4.3.1 การทดสอบการถดถอย

จากการวิเคราะห์ความแปรปรวน โดยทดสอบว่าตัวแปรอิสระมีส่วนในการอธิบายความผันแปรของตัวแปรตามหรือไม่ ดังตารางที่ 4.8

ตารางที่ 4.8 ผลการวิเคราะห์ความแปรปรวนของตัวแปรตามจากตัวแปรอิสระ 15 ตัวแปร

Sov	df	SS	MS	F	p-value
Regression	1.487E+12	11	1.352E+12	1621.237	.000
Residual	3.505E+12	42026	83401045.21		
Total	4.992E+12	42037			

สมมติฐาน

$$H_0 : \beta_i = \beta_2 = \dots = \beta_{15} = 0$$

$$H_1 : \beta_i \neq 0 \text{ อย่างน้อย 1 ตัว ; } i = 1, 2, \dots, 15$$

จากตารางที่ 4.8 จะพบว่าสถิติทดสอบคือ $F = \frac{MSR}{MSE} = 1,621.237$ และ p-value = 0.000 มีค่าน้อยกว่า $\alpha = 0.05$ แสดงว่าปฏิเสธสมมติฐานว่าง H_0 นั่นคือมีค่า b_i อย่างน้อย 1 ตัวที่แตกต่างจากศูนย์ หมายความว่า มีตัวแปรอิสระอย่างน้อย 1 ตัวแปร ที่มีส่วนในการอธิบายความผันแปรของตัวแปรตาม Y จึงต้องทำการทดสอบต่อไปว่าตัวแปรอิสระตัวใดมีส่วนในการอธิบายความผันแปรของตัวแปรตาม Y โดยใช้การทดสอบเอฟบางส่วน (partial F-test) หรือการทดสอบที (t-test)

เอกสารนี้เป็นเอกสารทสวงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอญาตเหนาไปไซประโยชน์ดานการคา โดยมีสมมติฐานดังนี้
ไม่ว่างรณเตฯ ฟงสน ยกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมติฐาน

$$H_0 : \beta_i = 0$$

$$H_1 : \beta_i \neq 0$$

จากตารางที่ 4.9 พบว่าตัวแปรอิสระทุกตัวมีค่า p-value อยู่ระหว่าง 0.000 - 0.028 ซึ่งน้อยกว่า $\alpha = 0.05$ แสดงว่าปฏิเสธสมมติฐานว่าง H_0 นั่นคือเมื่อเพิ่มตัวแปรอิสระแต่ละตัวในตัวแบบการถดถอยแล้ว พบว่าตัวแปรอิสระแต่ละตัวมีส่วนในการอธิบายความผันแปรของตัวแปรตาม เมื่อมีตัวแปรอิสระตัวอื่น ๆ อยู่ในตัวแบบการถดถอยแล้ว

เนื่องจากตัวแปรอิสระมีจำนวนมาก ดังนั้นจึงเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเข้าสมการโดยวิธีการถดถอยทีละขั้น (Stepwise Regression) ได้ผลดังตารางที่ 4.9

ตารางที่ 4.9 ผลการเลือกตัวแปรอิสระเข้าสมการถดถอยโดยวิธีการถดถอยทีละขั้น

Model	Unstandardized Coefficients		Standardized Coefficients	t	p-value
	B	Std. Error	Beta		
Constant	26471.530	534.171			0.000
X_{11}	1.118	0.554	0.008	2.016	0.044
X_8	10.992	5.513	0.009	1.994	0.046
X_{12}	1.902	0.026	0.314	74.018	0.000
X_1	- 1271.104	35.907	- 0.150	- 35.400	0.000
X_6	2007.756	47.393	0.184	42.364	0.000
X_3	- 1945.183	91.803	- 0.088	- 21.189	0.000
X_7	72.153	3.842	0.079	18.781	0.000
X_{13}	- 1.204	0.193	- 0.026	- 6.248	0.000
X_2	-231.750	90.111	-0.11	-2.572	0.010
X_5	- 2859.099	61.511	- 0.224	- 46.481	0.000
X_9	- 186.835	29.090	- 0.027	- 6.423	0.000

จากตารางที่ 4.9 จะพบว่าตัวแปรที่ถูกคัดเลือกเข้าสู่ตัวแบบ เงินอื่นๆ (X_{12}) สังกัดของลูกจ้าง (X_5) ประเภทค่าจ้าง (X_6) ภูมิภาค (X_1) เขตการปกครอง (X_3) อายุ (X_7) จำนวนสมาชิกในครัวเรือน (X_9) เงินค่าอาหารที่ได้รับ (X_{13}) เพศ (X_2) โบนัส (X_{11}) จำนวนชั่วโมงในการทำงาน (X_8) เมื่อพิจารณาค่าสัมประสิทธิ์การถดถอยของสมการทำนายพบว่าเงินอื่นๆ (X_{12}) สามารถทำนายรายได้ของลูกจ้างได้สูงสุด และมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยมีค่าสัมประสิทธิ์ไม่ต่ำกว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การถดถอยในรูปคะแนนมาตรฐานเป็น 0.314 รองลงมาเป็นสังกัดของลูกจ้าง (X_5) สามารถทำนายรายได้ของลูกจ้างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐานเป็น -0.224 ภูมิภาค (X_1) สามารถทำนายรายได้ของลูกจ้างอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยมีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐานเป็น -0.150 จำนวนชั่วโมงในการทำงาน (X_8) สามารถทำนายรายได้ของลูกจ้างได้น้อยที่สุด โดยมีนัยสำคัญทางสถิติที่ระดับ 0.05 มีค่าสัมประสิทธิ์การถดถอยในรูปคะแนนมาตรฐานเป็น 0.009 ดังนั้นสมการทำนายสำหรับรายได้ของลูกจ้าง คือ

$$\hat{Y} = 26,471.530 + 0.1.902X_{11} - 2,859.099X_8 + 2,007.756X_{12} - 1,271.104X_1 - 1,945.183X_6 + 72.153X_3 - 186.835X_7 - 1.204X_{13} - 231.750X_7 + 1.118X_5 + 10.992X_9$$

ค่า $b_1 = 0.314$ หมายถึงเงินอื่นๆเพิ่มขึ้น 1 บาทต่อเดือน จะทำให้รายได้ของลูกจ้างเพิ่มขึ้น 0.314 บาทต่อเดือน หรือถ้าเงินอื่นๆลดลง 1 บาทต่อเดือน จะทำให้รายได้ของลูกจ้างลดลง 0.314 บาทต่อเดือนเช่นกัน เป็นต้น

ตารางที่ 4.10 ค่าสหสัมพันธ์ (R) สัมประสิทธิ์การกำหนด (R Square) สัมประสิทธิ์การกำหนดแบบปรับ (Adjusted R Square) และค่าสถิติทดสอบเดอ์บิน-วัตสัน ของตัวแปรตามจากตัวแปรอิสระ 15 ตัว

ตัวแปรตาม	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
Y	.546	.298	.298	9132.417	1.593

จากตารางที่ 4.10 ค่า Adjusted $R^2 = 0.298$ หมายความว่า ภูมิภาค (X_1) เขตการปกครอง (X_3) สังกัดของลูกจ้าง (X_5) ประเภทค่าจ้าง (X_6) อายุ (X_7) จำนวนชั่วโมงในการทำงาน (X_8) จำนวนสมาชิกในครัวเรือน (X_9) โบนัส (X_{11}) เงินอื่นๆ (X_{12}) และเงินค่าอาหารที่ได้รับ (X_{13}) มีความสัมพันธ์กับรายได้ของลูกจ้าง (Y) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยสามารถสร้างสมการทำนายรายได้ของลูกจ้างได้ร้อยละ 29.8 ส่วนที่เหลืออีกร้อยละ 70.2 เป็นอิทธิพลจากตัวแปรอื่นๆ

4.4 การตรวจสอบสมมติหลังการวิเคราะห์การถดถอย

ทำการตรวจสอบโดยใช้ค่าส่วนเหลือ $e_i = Y_i - \hat{Y}_i$ ซึ่งเป็นค่าประมาณของค่าความคลาดเคลื่อน (ε)

4.4.1 การตรวจสอบส่วนเหลือมีการแจกแจงปกติ (Normality of Residual)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตรวจสอบโดยใช้สถิติทดสอบด้วยวิธีลิลีฟอर्स (Lilliefors) เช่นเดียวกับการตรวจสอบการแจกแจงปกติของตัวแปรตาม

ตารางที่ 4.11 ผลการทดสอบการแจกแจงปกติของส่วนเหลือ

ส่วนเหลือ	Lilliefors	
	Statistic	p-value
e	0.145	0.000

สมมติฐาน

H_0 : ส่วนเหลือมีการแจกแจงปกติ

H_1 : ส่วนเหลือไม่มีการแจกแจงปกติ

สถิติทดสอบ

จากตารางที่ 4.11 จะพบว่า Statistic = 0.145 และ p-value = 0.000 < $\alpha = 0.05$ แสดงว่าปฏิเสธสมมติฐานว่าง H_0 ดังนั้นส่วนเหลือไม่มีการแจกแจงปกติ

4.4.2 การตรวจสอบส่วนเหลือเป็นอิสระกัน (Independence of Residual)

สมมติฐาน

H_0 : ส่วนเหลือไม่มีสหสัมพันธ์ในตัว

H_1 : ส่วนเหลือมีสหสัมพันธ์ในตัว

สถิติทดสอบ

จากตารางที่ 4.10 จะพบว่า DW = 1.593 มีค่าอยู่ในช่วง 1.5-2.5 แสดงว่าไม่สามารถปฏิเสธสมมติฐานว่าง H_0 ดังนั้นส่วนเหลือไม่มีสหสัมพันธ์ในตัวหรือส่วนเหลือเป็นอิสระกัน

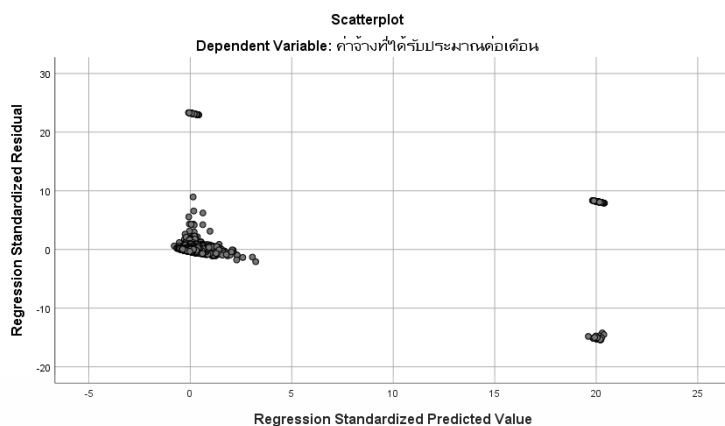
4.4.3 การตรวจสอบส่วนเหลือมีความแปรปรวนเท่ากันหรือคงที่ (Heteroscedasticity of Residual)

สมมติฐาน

H_0 : ส่วนเหลือมีความแปรปรวนเท่ากันหรือคงที่

H_1 : ส่วนเหลือมีความแปรปรวนไม่เท่ากันหรือไม่คงที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.3 พล็อตกราฟระหว่างส่วนเหลือมาตรฐานและค่าทำนายมาตรฐาน

จากรูปที่ 4.3 จะพบว่าส่วนเหลือมาตรฐานมีการกระจายอยู่รอบค่า 0 อย่างสม่ำเสมอ และขนานไปกับแกนนอน โดยที่ส่วนเหลือมาตรฐานมีค่าอยู่ระหว่าง -15 และ 25 ดังนั้นส่วนเหลือมีความแปรปรวนคงที่

4.4.4 การตรวจสอบค่าเฉลี่ยส่วนเหลือ (Mean of Residual)

สมมติฐาน

H_0 : ค่าเฉลี่ยของส่วนเหลือเท่ากับ 0

H_1 : ค่าเฉลี่ยของส่วนเหลือไม่เท่ากับ 0

ในกรณีนี้ไม่จำเป็นต้องตรวจสอบเนื่องจากการประมาณค่า β_i ด้วย b_i ด้วยวิธีกำลังสองน้อยที่สุด (Ordinary Least Square) จะได้ค่า $\sum e_i = 0$ จึงทำให้เงื่อนไขนี้เป็นจริงเสมอ แสดงว่าไม่สามารถปฏิเสธสมมติฐานว่าง H_0 ดังนั้นค่าเฉลี่ยของส่วนเหลือเท่ากับ 0

4.5 การอภิปรายผล

การศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทยจากการวิเคราะห์โดยใช้การถดถอยทีละขั้น (Stepwise Regression) จะพิจารณาพบว่า จากตัวแปรอิสระทั้งหมด 15 ตัว นำไปวิเคราะห์เพื่อหาตัวแบบสำหรับสร้างสมการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) ของปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย (Y) พบว่ามีตัวแปรอิสระ 11 ตัวที่มีความสัมพันธ์กับตัวแปรตามดังนี้คือ ภูมิภาค (X_1) เขตการปกครอง (X_3) สังกัดของลูกจ้าง (X_5) ประเภทค่าจ้าง (X_6) อายุ (X_7) จำนวนชั่วโมงในการทำงาน (X_8) จำนวนสมาชิกในครัวเรือน (X_9) โบนัส (X_{11}) เงินอื่นๆ (X_{12}) และเงินค่าอาหารที่ได้รับ (X_{13}) ซึ่งให้ผลสอดคล้องกับงานวิจัยของศศิธรและจรงค์ (2557) ได้ศึกษาปัจจัยที่มีผลต่อรายได้ของแรงงานในสถานประกอบการขนาดกลางและขนาดย่อมในอำเภอเมือง จังหวัดขอนแก่น รวบรวมข้อมูลโดยไม่ว่าจะเป็นแบบสอบถามและวิเคราะห์ด้วยวิธีการถดถอยแบบกำลังสองน้อยที่สุดและการวิเคราะห์สมการ

ถดถอยแบบสองชั้น ผลการวิจัยพบว่า ปัจจัยด้านทักษะทางสติปัญญาที่มีความสัมพันธ์ต่อรายได้ของแรงงานในเชิงบวกของภาคการบริการและภาคการค้า ได้แก่ จำนวนปีการศึกษา และประสบการณ์ในการทำงาน และให้ผลสอดคล้องกับองอาจ (2558) ที่ศึกษาปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ หนองคาย มุกดาหาร แม่สอด จากการศึกษาข้อมูลในภาคครัวเรือนพบว่าปัจจัยที่มีผลต่อการเปลี่ยนแปลงรายได้ของครัวเรือนในพื้นที่เขตเศรษฐกิจพิเศษของหนองคาย มุกดาหาร และแม่สอด คือ อายุ เมื่อวิเคราะห์ข้อมูลรายได้ประชากรพบว่า ปัจจัยด้านแรงงานอื่นๆ ส่งผลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษทั้งสามพื้นที่และจังหวัดโดยรอบ โดยผลจากการวิเคราะห์ปัจจัยด้านแรงงานที่กำหนดในแบบจำลอง ได้แก่ ค่าแรงขั้นต่ำ และจำนวนชั่วโมงทำงาน



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การวิจัยในครั้งนี้เพื่อศึกษาปัจจัยที่มีอิทธิพลรายได้ของลูกจ้างในประเทศไทย โดยใช้ข้อมูล ทศนิยมจากสำนักงานสถิติแห่งชาติ เรื่องการสำรวจภาวะการณมีงานทำ พ.ศ. 2567 โดยการวิเคราะห์ การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) ทำการเลือกตัวแปรอิสระด้วย วิธีเพิ่มตัวแปรอิสระแบบขั้นตอน (Stepwise Regression) เพื่อใช้ในการหาสมการถดถอย สามารถ สรุปผลการวิเคราะห์ข้อมูลได้ดังนี้

5.1 สรุปผลการวิจัย

จากการวิจัยพบว่า อายุ (X_7) มีค่าเฉลี่ย 40.47 ปี จำนวนชั่วโมงในการทำงาน (X_8) มี ค่าเฉลี่ย 43.69 ชั่วโมงต่อสัปดาห์ จำนวนสมาชิกในครัวเรือน (X_9) มีค่าเฉลี่ย 3.11 คนต่อครัวเรือน ค่าล่วงเวลา (X_{10}) มีค่าเฉลี่ย 266.00 บาทต่อเดือน โบนัส (X_{11}) มีค่าเฉลี่ย 2.30 บาทต่อปี เงิน อื่นๆ (X_{12}) มีค่าเฉลี่ย 520.12 บาทต่อเดือน เงินค่าอาหารที่ได้รับ (X_{13}) มีค่าเฉลี่ย 49.41 บาทต่อ เดือน เงินค่าเสื้อผ้าที่ได้รับ (X_{14}) มีค่าเฉลี่ย 0.00 บาทต่อเดือน เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15}) มี ค่าเฉลี่ย 39.93 บาทต่อเดือน รายได้ของลูกจ้าง (Y) มีค่าเฉลี่ย 14,481.21 บาทต่อเดือน

จะเห็นได้ว่าคนที่ตอบแบบสอบถามจะมีจำนวนคนและร้อยละที่มากที่สุดในแต่ละตัวแปรดังนี้ ภูมิภาค (X_1) โดยมีภาคกลางมากที่สุดจำนวน 16,450 คน คิดเป็นร้อยละ 39.1 เพศ (X_2) โดยมี เพศชายมากที่สุดจำนวน 22,355 คน คิดเป็นร้อยละ 53.2 เขตการปกครอง (X_3) โดยมีในเขต เทศบาลมากที่สุดจำนวน 24,731 คน คิดเป็นร้อยละ 58.8 สถานภาพสมรส (X_4) โดยสถานะโสด มากที่สุดจำนวน 26,167 คิดเป็นร้อยละ 62.2 สังกัดของลูกจ้าง (X_5) โดยมีลูกจ้างภาคเอกชนมาก ที่สุดจำนวน 31,163 คน คิดเป็นร้อยละ 74.1 ประเภทค่าจ้าง (X_6) โดยมีค่าจ้างรายเดือนมากที่สุด จำนวน 26,446 คน คิดเป็นร้อยละ 62.9

สรุปผลการวิเคราะห์ข้อมูลเพื่อศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย จากการศึกษาโดยใช้การถดถอยทีละขั้น (Stepwise Regression) จะพิจารณาพบว่า จากตัวแปร อิสระทั้งหมด 15 ตัว มีตัวแปรอิสระ 9 ตัวแปร ที่มีความสัมพันธ์กับตัวแปรตาม นำไปวิเคราะห์เพื่อหา ตัวแบบสำหรับสร้างสมการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) ของ ปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย (Y) โดยในสมการถดถอยเชิงเส้นพหุคูณมีตัว แปรเชิงคุณภาพ 6 ตัวแปร ดังนี้

$$\hat{Y} = 26,471.530 + 0.1.902X_{11} - 2,859.099X_8 + 2,007.756X_{12} - 1,271.104X_1 \\ - 1,945.183X_6 + 72.153X_3 - 186.835X_7 - 1.204X_{13} - 231.750X_7 + 1.118X_5 + 10.992X_9$$

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น มิใช่เพื่อเผยแพร่หรือใช้ประโยชน์ทางการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่า Adjusted $R^2 = 0.298$ หมายความว่า ภูมิภาค (X_1) เขตการปกครอง (X_3) สังกัดของลูกจ้าง (X_5) ประเภทค่าจ้าง (X_6) อายุ (X_7) จำนวนชั่วโมงในการทำงาน (X_8) จำนวนสมาชิกในครัวเรือน (X_9) โบนัส (X_{11}) เงินอื่นๆ (X_{12}) และเงินค่าอาหารที่ได้รับ (X_{13}) มีความสัมพันธ์กับรายได้ของลูกจ้าง (Y) อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 โดยสามารถสร้างสมการทำนายรายได้ของลูกจ้างได้ร้อยละ 29.8 ส่วนที่เหลืออีกร้อยละ 70.2 เป็นอิทธิพลจากตัวแปรอื่นๆ

5.2 ข้อเสนอแนะ

5.2.1 การศึกษาครั้งต่อไปอาจศึกษาตัวแปรอิสระอื่นที่น่าสนใจ หรือหัวข้ออื่นๆในแบบสอบถาม เพราะสามารถตัดแปลงได้หลายหัวข้อตามความสนใจและการวิเคราะห์ตัวแปร

5.2.2 อาจมีตัวแปรบางตัวที่ส่งผลต่อรายได้ของลูกจ้างที่ไม่ได้นำมาวิเคราะห์ เช่น การศึกษาสูงสุด สาขาที่เรียนจบ สายวิชาที่เรียน และขนาดของบริษัท เป็นต้น

5.2.3 ตัวแปรบางตัวมีข้อมูลสูญหายหลายค่า เช่น ประเภทค่าจ้าง ค่าล่วงเวลา รายได้ของลูกจ้าง เป็นต้น

5.2.4 ตัวแปรบางตัวมีค่าสูญหาย สำหรับตัวแปรเชิงปริมาณให้แทนที่ด้วยค่าเฉลี่ย สำหรับตัวแปรเชิงคุณภาพให้แทนที่ด้วยมัธยฐาน ทำให้ผลการวิเคราะห์อาจมีการคลาดเคลื่อน เนื่องจากไม่ใช่ข้อมูลที่แท้จริง แต่เกิดจากการแทนที่ค่าสูญหาย

เอกสารอ้างอิง

- กฤษณ์เลิศ สัมพันธ์รักษ์. (2560). การเงินครัวเรือนชนบท ผ่านสองทศวรรษงานวิจัย Townsend Thai Project. [Online]. Available: <https://thaipublica.org/2017/04/pier-19/>.
- กัลยา วานิชย์บัญชา. (2548). การวิเคราะห์สถิติขั้นสูงด้วย SPSS for Window. พิมพ์ครั้งที่ 4. กรุงเทพฯ: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย.
- กัลยา วานิชย์บัญชา. (2558). การใช้ SPSS for Window ในการวิเคราะห์ข้อมูล. พิมพ์ครั้งที่ 15. กรุงเทพฯ : โรงพิมพ์สามลดา.
- กัลยา วานิชย์บัญชา. (2561). สถิติสำหรับงานวิจัย. พิมพ์ครั้งที่ 12. กรุงเทพฯ : ห้างหุ้นส่วนจำกัด สามลดา.
- กัลยา วานิชย์บัญชาและฐิตา วานิชบัญชา. (2564). การใช้ SPSS for Window ในการวิเคราะห์ข้อมูล. พิมพ์ครั้งที่ 33. กรุงเทพฯ : โรงพิมพ์สามลดา.
- จุฑาทิพย์ แซ่ตัน และสุภาพร สอนสมนึก. (2562). ปัจจัยที่ส่งผลต่อค่าใช้จ่ายในการรักษาสุขภาพ ปัญหาพิเศษปริญญาบัณฑิต คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, หน้า 36.
- ชนินันท์ พงษ์ประมุข, นงนุช เจริญตะคุ และสิวะโชติ ศรีสุทียากร. (2563). ประสิทธิภาพของสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน สเปียร์แมน และเคนดอลล์ เมื่อข้อมูลแจกแจงไม่ปกติ [ข้อมูลอิเล็กทรอนิกส์]. วารสารอิเล็กทรอนิกส์ทางการศึกษา OJED, 2563(15), 2-16.
- โชติรส เทียนถาวร. (2544). การเปรียบเทียบประสิทธิภาพของวิธีประมาณพารามิเตอร์ของตัวแบบการถดถอยที่มีความคาดเคลื่อนแบบบอดิสถสัมพันธ์. วิทยานิพนธ์ปริญญาโทบริหารบัณฑิต คณะวิทยาศาสตร์ มหาวิทยาลัยศิลปากร. 2544. หน้า 24.
- ฐณัฐ วงศ์สายเชื้อ. (2559). Replace missing value - การแทนค่าสูญหายในโปรแกรม SPSS. [Online]. Available: https://www.youtube.com/watch?v=WzaeJ_HAqtk
- ธนาคารแห่งประเทศไทย. (2562). เครื่องชี้เศรษฐกิจมหภาคของไทย. [Online]. Available: https://www.bot.or.th/App/BTWS_STAT/statistics/ReportPage.aspx?reportID=409&language=th.
- ธิดารัตน์ เห็นพร้อม. (2566). อัตราการว่างงานคืออะไร สำคัญต่อการลงทุนอย่างไร. [Online]. Available: <https://thunhoon.com/article/273363>.
- นงลักษณ์ วัชชัย. (2542). โมเดลลิสมัล : สถิติวิเคราะห์สำหรับการวิจัย. พิมพ์ครั้งที่ 3. กรุงเทพฯ: จุฬาลงกรณ์มหาวิทยาลัย.
- นันทกา นันทวิสัย. (2552). การศึกษาเปรียบเทียบภาวะการออมของครัวเรือนในภาคเกษตรและนอกภาคเกษตร. [Online]. Available: http://thesis.swu.ac.th/swuthesis/Man_Econ/Nantaka_N.pdf. (ครั้งที่มีการนำไปใช้)

เอกสารอ้างอิง (ต่อ)

- บัณฑิต ชัยวิษณุชาติ. (2558). แบบจำลองเศรษฐมิติ: นอกเหนือจากแบบจำลองถดถอย. กรุงเทพฯ: สำนักพิมพ์มหาวิทยาลัยเกษตรศาสตร์.
- ปิยะภรณ์ ประสิทธิ์วัฒนเสรี และสุคนธ์ ประสิทธิ์วัฒนเสรี. (2552). ข้อมูลสูญหายและแนวทางการจัดการ. Data Management & Biostatistics Journal Vol.4 No.3, Page No.52-61.
- พรพิมล ทรงกิติพิศาล และชูชีพ พิพัฒน์ศิริ. (2554). ศึกษาผลของระดับการศึกษาที่มีต่อรายได้ของครัวเรือนไทย. [Online]. Available: https://esd.kps.ku.ac.th/kukconference/img/gallery/article_11/pdf_file/FULL%20PAPER_human/Oral/34_Pornpimon_783-790.pdf.
- เริงใจ ต้นสุชาติ. (2547). แบบจำลองถดถอย. [Online]. Available: <http://lms.mju.ac.th/courses/159/locker/Econometrics2/content10.htm>.
- วุฒิ สุขเจริญ. (2558). การดำเนินการกับข้อมูลขาดหาย. วารสารร่มพฤกษ์ มหาวิทยาลัยเกริก. 33(2) : 11-32.
- ศศิธร เห็นงาม และจงรักษ์ หงส์งาม. (2557). วารสารเศรษฐศาสตร์และกลยุทธ์การจัดการ, 2557 (2). Available: [file:///C:/Users/ADMIN/Downloads/chaiwat,+Journal+manager,+JEMS+Full+2014+V1\(2\)+P52-62%20\(1\).pdf](file:///C:/Users/ADMIN/Downloads/chaiwat,+Journal+manager,+JEMS+Full+2014+V1(2)+P52-62%20(1).pdf).
- สมใจ ชวนละคร. (2559). ปัจจัยที่มีอิทธิพลต่อการมีหนี้สินของครอบครัวข้าราชการพลเรือนสามัญในกรุงเทพมหานคร และปริมณฑล. วิทยานิพนธ์ปริญญาโทมหาบัณฑิต คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง, หน้า 33.
- สายชล สีนสมบูรณ์ทอง. (2563). สถิติเมืองพารามิเตอร์. พิมพ์ครั้งที่ 2 ฉบับปรับปรุง, กรุงเทพฯ: จามจรีโปรดักส์.
- สายชล สีนสมบูรณ์ทอง. (2563). สถิติเบื้องต้น. พิมพ์ครั้งที่ 12, กรุงเทพฯ: จามจรีโปรดักส์.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

สำนักงานเศรษฐกิจการเกษตร. (2562). ศึกษาปัจจัยที่มีอิทธิพลต่อรายได้ รายจ่าย การออม หนี้สิน ปัจจัยที่กำหนดการตัดสินใจก่อนหนี้และความต้องการกู้ยืมของครัวเรือนเกษตรกร [Online].

Available:

<https://www.oae.go.th/view/1/%E0%B8%A0%E0%B8%B2%E0%B8%A7%E0%B8%B0%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B9%80%E0%B8%A8%E0%B8%A3%E0%B8%A9%E0%B8%90%E0%B8%81%E0%B8%B4%E0%B8%88%E0%B8%AA%E0%B8%B1%E0%B8%87%E0%B8%84%E0%B8%A1%E0%B8%84%E0%B8%A3%E0%B8%B1%E0%B8%A7%E0%B9%80%E0%B8%A3%E0%B8%B7%E0%B8%AD%E0%B8%99%E0%B9%81%E0%B8%A5%E0%B8%B0%E0%B9%81%E0%B8%A3%E0%B8%87%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B9%80%E0%B8%81%E0%B8%A9%E0%B8%95%E0%B8%A3/31803/TH-TH>.

สำนักงานเศรษฐกิจการคลัง. (2546). ทฤษฎีการบริโภค. [Online]. Available :

<http://www.fpo.go.th/>.

สำนักงานสถิติแห่งชาติ. (2566). การสำรวจภาวะการทำงานของประชากรทั่วราชอาณาจักร ไตรมาสที่ 2 : เมษายน - มิถุนายน 2566 (รายงาน). [Online]. Available : [20230912113729_96988.pdf](https://www.nso.go.th/publication/20230912113729_96988.pdf) (nso.go.th).

สุชาดา กิระนันท์, ชินนะพงษ์ บำรุงทรัพย์, กมล บุขบา, มนต์ทิพย์ เทียนสุวรรณ, วีรานันท์ พงศาภักดี, พาชิตชนัด ศิริพานิช, จิราวัลย์ จิตรถเวช, วัลภา ประกอบผล, สุวาณี สุรเสียงสังข์ และสุดา ตระการเถลิงศักดิ์. (2558). **พจนานุกรมศัพท์สถิติศาสตร์ ฉบับราชบัณฑิตยสถาน. พิมพ์ครั้งที่ 2 ฉบับปรับปรุง กรุงเทพฯ: โรงพิมพ์แห่งจุฬาลงกรณ์มหาวิทยาลัย.**

องอาจ รุกขวัฒน์กุล. (2558). ศึกษาปัจจัยที่มีอิทธิพลต่อการเปลี่ยนแปลงรายได้ของประชาชนในพื้นที่การพัฒนาเขตเศรษฐกิจพิเศษ หนองคาย มุกดาหาร แม่สอด. [Online]. Available: https://ethesisarchive.library.tu.ac.th/thesis/2015/TU_2015_5704010700_4720_2858.pdf.

อมรทิพย์ แท้เที่ยงธรรม. (2547). **เศรษฐศาสตร์มหภาค. พิมพ์ครั้งที่ 3 กรุงเทพฯ: มหาวิทยาลัยเกษตรศาสตร์**

James S. Duesenberry. (1949). *Income, Saving and the Theory of Consumer Behavior.* Cambridge: Harvard University Press.

Montgomery, D.C. 2012. *Introduction to Linear Regression Analysis.* 5th ed. New

เอกสารนี้เป็นเอกสาร Jersey: Join Wiley. ารใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เอกสารอ้างอิง (ต่อ)

Neter, J. K., M.H., Nachtsheim, C.J. and Wasserman, W. (1996). Applied Linear Statistical Models. 4th ed. Chicago: Irwin.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

รายละเอียดและตัวอย่างข้อมูลที่นำมาวิเคราะห์

1. ข้อมูลตัวแปร

แผนการเลือกตัวอย่าง แผนการเลือกตัวอย่างที่ใช้เป็นแบบแบ่งชั้นภูมิ 2 ขั้นตอน (Stratified Two-Stage Sampling) โดยมีจังหวัดเป็นสตราตัม (Stratum) มีในเขตเทศบาลและนอกเขตเทศบาลเป็นสตราตัมย่อย (Substratum) ในการเลือกตัวอย่างขั้นที่หนึ่ง (Primary Sampling Stage) หน่วยตัวอย่างขั้นที่หนึ่งคือเขตแฉงนั้บ โดยกรอบตัวอย่างสำหรับการเลือกหน่วยขั้นที่หนึ่งได้มาจากโครงการสำมะโนประชากรและเคหะ พ.ศ.2553 ที่ได้มีการปรับปรุงให้มีความทันสมัย จำนวนเขตแฉงนั้บตัวอย่างทั้งสิ้น 5,430 เขตแฉงนั้บ สำหรับการเลือกตัวอย่างขั้นที่สอง (Secondary Sampling Stage) หน่วยตัวอย่างขั้นที่สองคือครัวเรือนส่วนบุคคลและสมาชิกในครัวเรือนกลุ่มบุคคลประเภทคนงาน โดยกำหนด 16 ครัวเรือนตัวอย่างต่อเขตแฉงนั้บตัวอย่างมีทั้งในเขตเทศบาลและนอกเขตเทศบาล ตามลำดับ ในกรณีของครัวเรือนกลุ่มบุคคลประเภทคนงานจะเป็นการเลือกสมาชิกตัวอย่างจากครัวเรือนกลุ่มบุคคลประเภทคนงานทุกครัวเรือน โดยก่อนดำเนินการเลือกครัวเรือนตัวอย่างได้มีการสร้างกรอบครัวเรือนตัวอย่างด้วยการนับและจดครัวเรือนและอาคาร/สิ่งปลูกสร้างทุกแห่งในแต่ละเขตแฉงนั้บตัวอย่าง หลังจากนั้นทำการจัดเรียงรายชื่อครัวเรือนส่วนบุคคลใหม่ตามขนาดครัวเรือน ซึ่งวัดด้วยจำนวนสมาชิกในครัวเรือน แล้วจึงดำเนินการเลือกครัวเรือนตัวอย่างด้วยวิธีการสุ่มแบบมีระบบ (Systematic Sampling) โดยผู้ตอบคำถาม คือ หัวหน้าครัวเรือนหรือสมาชิกในครัวเรือน(สามารถตอบแทนกันได้)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.1 ตัวแปรทั้งหมดก่อนเลือกมาวิเคราะห์

ลำดับ	ชื่อตัวแปรภาษาไทย	ชื่อตัวแปรภาษาอังกฤษ
1	ภาค	REG
2	จังหวัด	CWT
3	เขตปกครอง	AREA
4	ลำดับที่เขตแดนนับตัวอย่าง	PSU_NO
5	ชุด EA ตัวอย่าง	EA_SET
6	ชุดครัวเรือนตัวอย่าง	SAMSET
7	พ.ศ.	YR
8	ลำดับที่ครัวเรือนตัวอย่าง	HH_NO
9	ประเภทครัวเรือน	TYPE
10	จำนวนสมาชิกในครัวเรือนชั้นแรงงาน	MEMBERS
11	จำนวนสมาชิกในครัวเรือนชั้นนับจด	LISTING
12	ผลการแรงงาน	ENUM
13	ลำดับที่	NO
14	ความสัมพันธ์กับหัวหน้าครัวเรือน	RELATION
15	เพศ	SEX
16	อายุ	AGE
17	สถานภาพสมรส	MARITAL
18	ชั้นที่กำลังเรียน	GRADE_A
19	ระดับการศึกษา	GRADE_B
20	สาขาวิชา	SUBJECT
21	ระหว่าง 7 วันก่อนวันสัมภาษณ์ได้ทำงานที่ก่อให้เกิดรายได้/เงินเดือน/ค่าจ้าง/ผลกำไร/ผลตอบแทนหรือได้ทำงานในฐานะช่วยธุรกิจ	WK_7DAY
22	ไม่ได้ทำงานระหว่าง 7 วันก่อนวันสัมภาษณ์ ยังได้รับค่าจ้าง/เงินเดือนหรือผลประโยชน์อื่น ๆ หรือผลกำไรจากงานหรือธุรกิจ	RECEIVE
23	มีงาน/ธุรกิจ/กิจการที่จะกลับไปทำ	RETURN
24	ระยะเวลาที่หยุดทำงาน	ABSENT
25	ก่อนวันสัมภาษณ์ได้หางานทำ/สมัครงาน	SEEKING
26	วิธีหางานทำ/สมัครงาน	METHOD
27	ระหว่าง 7 วันก่อนวันสัมภาษณ์พร้อมทำงาน	AVAILABLE
28	เหตุผลที่ไม่พร้อมทำงาน	RE_UNAVAIL
29	เหตุผลที่ไม่หางานทำ	RE_NO_SEEK

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับครูและบุคลากรทางการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีกรณีไปใช้

ตารางที่ ก.1 ตัวแปรทั้งหมดก่อนเลือกมาวิเคราะห์ (ต่อ)

30	ระยะเวลาหางาน/พร้อมทำงาน	DR_SEEK
31	เคยทำงานมาก่อน	EVER_WK
32	สาเหตุที่ออกจากงาน/หยุดทำงาน	RE_UNEM
33	ระยะเวลาที่หยุดทำงาน	DR_UNEM
34	อาชีพ	OCCUP
35	เป็นกิจกรรมหรือผลิตภัณฑ์	INDUS
36	สังกัดของลูกจ้าง	STATUS
37	จำนวนคนทำงาน	SIZE
38	จำนวนชั่วโมงทำงานจริงของอาชีพที่บันทึกใน OCCUP	MAINHR
39	จำนวนชั่วโมงทำงานจริงของอาชีพอื่นทุก อาชีพที่ทำในระหว่าง 7 วันก่อนวันสัมภาษณ์	OTHER_HR
40	จำนวนชั่วโมงทำงานทั้งสิ้น	TOTAL_HR
41	มีเวลาเหลือและต้องการทำงานเพิ่ม	MORE_WK
42	จำนวนชั่วโมงที่ต้องการทำงานเพิ่มขึ้นต่อ สัปดาห์	MORE_HR
43	หางานทำ/ติดต่อสมัครงานเพิ่ม	FINDING
44	สาเหตุที่ไม่หางานเพิ่ม	RE_NOMORE
45	ประเภทค่าจ้าง	WAGE_TYPE
46	ค่าจ้างที่ได้รับเป็นเงินบาทต่อชั่วโมงหรือต่อวัน หรือต่อสัปดาห์	AMOUNT
47	ค่าจ้างที่ได้รับประมาณต่อเดือน	APPROX
48	โบนัสต่อปี (ในรอบ 12 เดือนก่อนวัน สัมภาษณ์)	BONUS
49	ค่าล่วงเวลาต่อเดือน	OT
50	เงินอื่นๆต่อเดือน	OTH_MONEY
51	อาหาร	FOOD
52	เสื้อผ้าหรือเครื่องแต่งกายอื่นๆ	CLOTH
53	ที่อยู่อาศัย	HOUSE
54	อื่นๆ	OTH_THING
55	เป็นสมาชิกคนสุดท้ายของครัวเรือน	LST_MEN
56	รหัสเข้าทำงานใหม่	RE_WK
57	ระดับการศึกษา	RE_ED
58	ค่าน้ำหนัก	WGT
59	ใครเป็นผู้ตอบแบบสัมภาษณ์	WHO
60	รหัสนำหมายเลข	RUN_NO

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับโครงการวิจัยที่ดำเนินการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ในที่สาธารณะโดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.2 คุณลักษณะและรายละเอียดตัวแปร

คุณลักษณะ	รายละเอียด
ภูมิภาค (X_1)	หน่วย : 0 แทนกรุงเทพมหานคร 1 แทนภาคกลาง 2 แทนภาคเหนือ 3 ภาคตะวันออกเฉียงเหนือ 4 ภาคใต้
เพศ (X_2)	หน่วย : 0 แทนเพศชาย 1 แทนเพศหญิง
เขตการปกครอง (X_3)	หน่วย : 0 แทนในเขตเทศบาล 1 แทนนอกเขตเทศบาล
สถานภาพสมรส (X_4)	หน่วย : 0 แทนสถานะโสด 1 แทนสถานะแต่งงานแล้ว 2 แทนไม่ทราบสถานะ
สังกัดของลูกจ้าง (X_5)	หน่วย : 0 แทนลูกจ้างรัฐ 1 แทนลูกจ้างวิสาหกิจ 2 แทนลูกจ้างเอกชน
ประเภทค่าจ้าง (X_6)	หน่วย : 0 แทนรายชั่วโมง 1 แทนรายวัน 2 แทนรายเดือน 3 แทนอื่นๆ
อายุ (X_7)	หน่วย : ปี
จำนวนชั่วโมงในการทำงาน (X_8)	หน่วย : ชั่วโมง
จำนวนสมาชิกในครัวเรือน (X_9)	หน่วย : คน
ค่าล่วงเวลา (X_{10})	หน่วย : บาท
โบนัส (X_{11})	หน่วย : บาท
เงินอื่นๆ (X_{12})	หน่วย : บาท
เงินค่าอาหารที่ได้รับ (X_{13})	หน่วย : บาท
เงินค่าเสื้อผ้าที่ได้รับ (X_{14})	หน่วย : บาท
เงินค่าที่อยู่อาศัยที่ได้รับ (X_{15})	หน่วย : บาท

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.4 ตัวอย่างหลังแปลงข้อมูล

	X_1	X_2	X_3	X_4	X_5	X_{6_1}	X_7	X_8	X_9	X_{10_1}	X_{11}	X_{12_1}	X_{13_1}	X_{14_1}	X_{15_1}	Y
1	0	1	1	1	2	4.0	31	48	4	0	0	5000	0	0	0	10000
2	0	2	1	1	2	4.0	35	48	4	0	0	0	0	0	0	13500
3	0	1	1	2	2	4.0	44	40	7	0	0	0	0	0	0	10000
4	0	1	1	1	2	4.0	35	48	6	0	0	0	0	0	0	16000
5	0	1	1	2	2	4.0	56	40	2	0	0	50000	0	0	0	70000
6	0	1	1	1	2	4.0	32	30	2	0	0	0	0	0	14000	30000
7	0	1	1	2	1	4.0	31	35	1	0	0	0	0	0	0	18000
8	0	2	1	2	2	4.0	64	40	2	0	0	0	0	0	0	40000
9	0	1	1	2	2	4.0	40	48	1	0	0	0	4500	0	3000	11000
10	0	1	1	2	2	2.0	72	40	2	0	0	0	0	0	0	8000
11	0	1	1	1	2	2.0	65	24	2	0	0	0	0	0	0	3600
12	0	2	1	1	2	4.0	39	48	5	0	0	0	0	0	0	18000
13	0	1	1	2	2	4.0	32	48	5	0	0	0	0	0	0	18000
14	0	1	1	1	2	4.0	37	40	2	0	0	0	0	0	0	19000
15	0	1	1	2	2	4.0	67	56	2	2000.0	0	0	0	0	0	10000
16	0	1	1	2	2	4.0	46	50	1	6000.0	0	1400	2600	0	0	25000
17	0	1	1	1	2	4.0	58	56	4	335.1	0	684.6	0	0	0	-
18	0	2	1	2	2	4.0	29	56	4	335.1	0	684.6	0	0	0	-
19	0	2	1	1	2	4.0	33	60	4	335.1	0	684.6	0	0	0	-
20	0	2	1	1	2	4.0	58	48	4	0	0	0	0	0	0	9000

หมายเหตุ : X_{10_1} ถึง X_{15_1} คือตัวแปรที่ทำการแทนค่าสูญหายด้วยค่า mean ยกเว้น X_{6_1} แทนค่าด้วย median

ภาคผนวก ข

1. การแทนค่าข้อมูลสูญหาย

ขั้นตอนการวิเคราะห์ข้อมูลสูญหาย

Case	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
965	1	2	1	6	6	4	26	48	1	0	0	0	0
966	1	1	1	2	6	5	41	48	1	0	0	0	1500
967	1	1	1	2	6	4	51	49	2	0	0	0	0
968	1	1	1	1	6	2	23	48	1	0	0	0	1200
969	1	2	1	1	4	4	34	35	1	0	0	0	0
970	1	2	1	2	4	4	32	35	2	0	0	0	0
971	1	1	1	2	4	4	28	42	2	1200	0	0	0
972	1	2	1	1	4	4	40	35	1	0	0	0	0
973	1	2	1	2	4	4	56	35	1	0	0	0	0
974	1	1	1	2	4	4	57	48	2	0	0	0	0
975	1	2	1	1	4	4	48	35	1	0	0	0	0
976	1	1	1	2	4	4	30	35	2	0	0	0	0
977	1	2	1	2	4	4	29	42	2	0	0	0	0
978	1	2	1	2	4	4	54	35	2	0	0	5000	0
979	1	1	1	2	4	4	55	35	2	0	0	0	0
980	1	1	1	2	4	4	57	35	3	0	0	0	0
981	1	1	1	1	4	4	27	35	3	0	0	0	0
982	1	2	1	1	4	4	31	35	2	0	0	0	0
983	1	1	1	1	6	5	43	60	1	0	0	0	0
984	1	2	1	2	6	4	55	54	1	0	0	0	0
985	1	1	1	1	6	4	30	40	1	0	0	0	0
986	1	1	1	1	6	4	19	42	1	0	0	0	2400
987	1	2	1	2	6	4	55	54	4	0	0	0	2000
988	1	1	1	1	6	4	31	45	4	0	0	0	2000
989	1	1	1	1	6	4	26	48	4	0	0	5000	0
990	1	2	1	1	6	4	45	48	1	0	0	0	0
991	1	1	1	1	6	4	25	48	1	2000	0	0	2000
992	1	1	1	1	6	2	66	48	1	0	0	0	0
993	1	1	1	2	6	4	56	48	2	0	0	0	0
994	1	2	1	1	6	4	52	48	1	0	0	0	0
995	1	2	1	1	6	4	55	48	1	0	0	0	1500
996	1	1	1	1	6	4	51	48	1	0	0	0	0
997	1	1	1	1	6	2	33	40	2	0	0	0	0
998	1	2	1	2	6	2	22	40	2	0	0	0	0
999	1	1	1	2	6	2	54	40	1	0	0	0	0
1000	1	1	1	2	6	4	34	40	5	0	0	5000	0

รูปที่ ข.1 ตัวอย่างข้อมูลที่ต้องการแทนค่าสูญหาย

1.1 ขั้นตอนการแทนค่าสูญหายด้วยวิธีค่าเฉลี่ย

การแทนค่าข้อมูลสูญหายด้วยวิธีค่าเฉลี่ย โดยใช้โปรแกรม SPSS

- 1) เปิดไฟล์ชุดข้อมูลที่ต้องการแทนค่าสูญหาย ใช้คำสั่ง Transform > Replace Missing Values

Case	X4	X5	X9	X10	X11
1	1		2	0	0
2	1		3	0	0
3	1		3	0	0
4	1		3	0	0
5	1		3	0	0
6	1		1	0	0
7	1		3	0	0
8	1		3	0	0
9	1		4	0	0
10	1		4	0	0
11	2		6	0	0
12	1		1	0	0
13	1		4	0	0
14	1		1	0	0
15	2		2	0	0
16	2	6	4	39	40
17	2	6	4	32	42

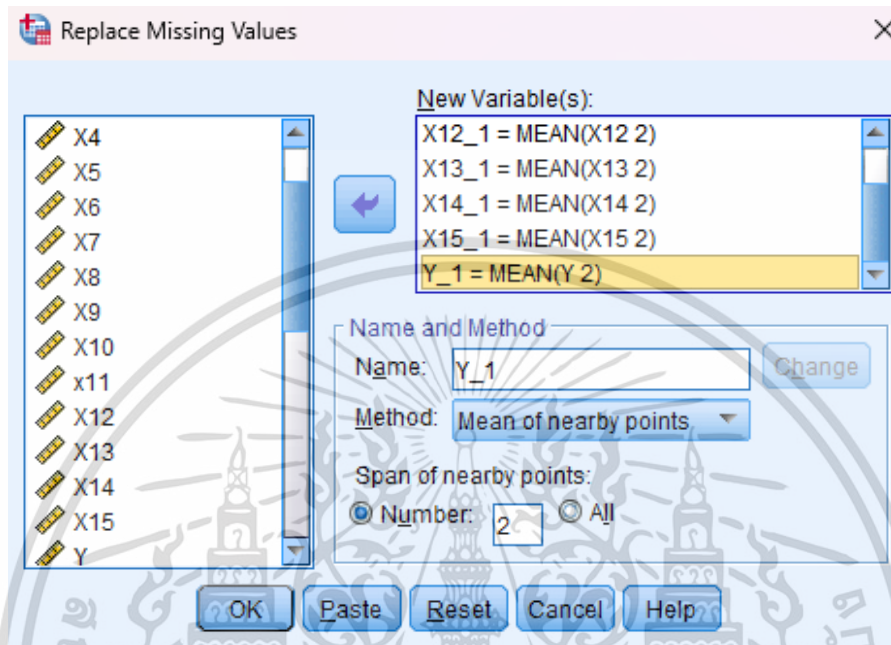
รูปที่ ข.2 การเลือกคำสั่งในการแทนค่าข้อมูลสูญหายด้วยวิธีการค่าเฉลี่ย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) เลือก Method: Mean of nearby point

เลือกข้อมูลที่ต้องการแทนค่าสูญหายไว้ในช่อง New Variables > ตั้งชื่อใหม่

คลิก OK



รูปที่ ข.3 การเลือกวิธีค่าเฉลี่ยกับตัวแปรเชิงปริมาณ และตัวแปรที่ต้องการแทนค่าสูญหาย

3) เลือกวิธีค่าเฉลี่ย และตัวแปรที่ต้องการแทนค่าสูญหาย

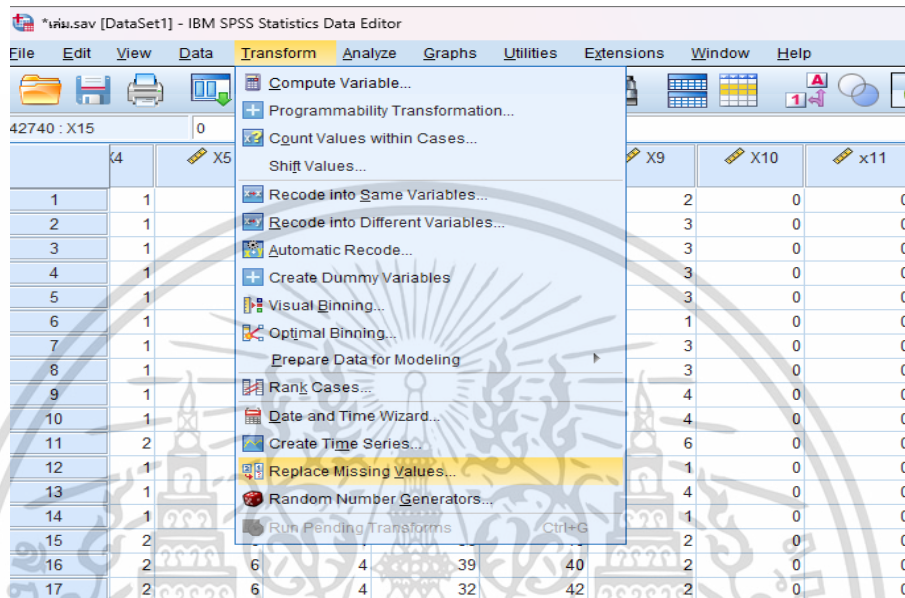
	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
11051	2	6	4	54	48	3	0	0	0	0	0
11052	2	4	4	40	40	1	0	0	0	0	0
11053	1	4	4	50	40	1	0	0	0	0	0
11054	1	4	4	40	40	1	0	0	0	0	0
11055	1	6	2	54	40	1	0	0	0	0	0
11056	1	6	2	40	60	3	3000	0	0	0	0
11057	1	6	4	22	60	3	3000	0	0	0	0
11058	2	6	2	63	48	4	0	0	0	0	0
11059	2	6	4	44	48	4	0	0	0	0	0
11060	1	6	4	31	48	1	0	0	0	0	0
11061	1	4	4	54	35	1	0	0	0	0	0
11062	2	6	4	23	56	7	0	0	0	0	0
11063	2	6	4	40	48	2	0	0	0	0	0
11064	2	6	4	41	48	2	0	0	0	0	0
11065	2	6	4	31	48	2	0	0	0	0	0
11066	2	6	4	37	48	2	0	0	0	0	0
11067	2	6	4	41	48	2	0	0	0	920	0
11068	2	6	4	37	48	2	0	0	0	920	0
11069	2	6	4	35	48	5	0	0	0	0	0
11070	1	4	4	30	40	3	0	0	0	0	0
11071	2	4	4	44	40	4	0	0	0	0	0
11072	2	4	4	44	40	4	0	0	0	0	0
11073	1	6	4	26	48	4	0	0	2000	0	0
11074	2	6	2	45	48	1	0	0	0	0	0
11075	2	6	4	35	60	3	4000	0	0	0	0
11076	2	6	4	33	60	3	4000	0	0	0	0
11077	2	6	4	38	40	3	0	0	0	0	0
11078	2	4	4	38	40	3	0	0	0	0	0
11079	2	6	4	43	48	3	0	0	0	0	0
11080	1	4	4	40	40	1	0	0	0	0	0
11081	2	4	4	41	40	1	0	0	0	0	0
11082	2	6	4	27	60	1	3000	0	0	0	0
11083	2	6	4	31	48	3	0	0	0	0	0
11084	2	6	4	42	48	2	0	0	2000	0	0
11085	2	6	4	44	48	2	0	0	2500	0	0
11086	2	6	4	44	48	2	0	0	1500	0	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น **รูปที่ ข.4 ผลลัพธ์การแทนค่าสูญหายด้วยวิธีค่าเฉลี่ย**
ไม่พบบทความนี้ในเอกสารฉบับนี้

1.2 ขั้นตอนการแทนค่าสูญหายด้วยวิธีมีฐานของค่าใกล้เคียง

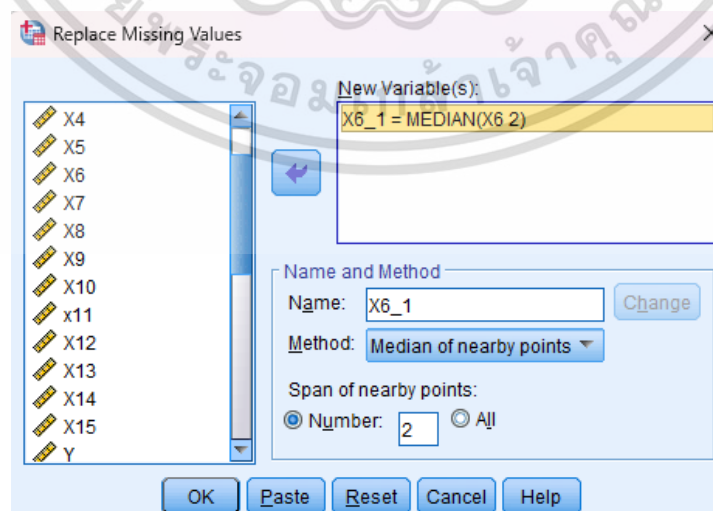
การหาค่าข้อมูลสูญหายด้วยวิธีมีฐานของค่าใกล้เคียง โดยใช้โปรแกรม SPSS

- 1) เปิดไฟล์ชุดข้อมูลที่ต้องการแทนค่าสูญหาย ใช้คำสั่ง Transform > Replace Missing Values



รูปที่ ข.5 เลือกคำสั่งในการแทนค่าข้อมูลสูญหาย ด้วยวิธีมีฐานของค่าใกล้เคียง

- 2) เลือก Method: Median of nearby point
เลือกข้อมูลที่ต้องการแทนค่าสูญหายใส่ในช่อง New Variables > ตั้งชื่อใหม่
คลิก OK



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
รูปที่ ข.6 การเลือกวิธีมีฐานของค่าใกล้เคียงกับตัวแปรเชิงคุณภาพ
ไม่ว่ากรณีใดๆ ทั้งสิ้น ยกเว้นห้ามมิให้ตีแบบสงวนสิทธิ์ และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และตัวแปรที่ต้องการแทนค่าสูญหาย

3) เลือกวิธีมีฐานของค่าใกล้เคียง และตัวแปรที่ต้องการแทนค่าสูญหาย

	X6_1	X10_1	x11_1	X12_1	X13_1	X14_1	X15_1	Y_1
2161	00	4.0	.0	.0	.0	.0	.0	12000.0
2162	00	4.0	.0	.0	.0	.0	.0	80000.0
2163	00	4.0	.0	.0	.0	.0	.0	70000.0
2164	00	4.0	.0	.0	.0	.0	.0	20000.0
2165	00	4.0	.0	.0	.0	.0	.0	40000.0
2166	00	4.0	.0	.0	.0	.0	.0	25000.0
2167	00	4.0	.0	.0	.0	.0	.0	80000.0
2168	00	4.0	.0	.0	.0	.0	.0	80000.0
2169	00	4.0	.0	.0	3000.0	.0	2000.0	5000.0
2170	00	4.0	.0	.0	3000.0	.0	2000.0	5000.0
2171	00	4.0	.0	.0	.0	.0	.0	9000.0
2172	00	4.0	.0	.0	.0	.0	.0	9000.0
2173	00	4.0	.0	.0	.0	.0	.0	20000.0
2174	00	4.0	.0	.0	.0	.0	.0	12000.0
2175	00	4.0	.0	.0	.0	.0	.0	10000.0
2176	99	4.0	.0	.0	.0	.0	.0	999999.0
2177	00	4.0	.0	.0	.0	.0	.0	32000.0
2178	00	4.0	.0	.0	.0	.0	.0	15500.0
2179	00	4.0	600.0	.0	.0	.0	.0	15000.0
2180	00	4.0	400.0	.0	.0	.0	.0	15000.0
2181	00	2.0	1500.0	.0	.0	.0	.0	9200.0
2182	00	4.0	.0	.0	.0	.0	.0	20800.0
2183	00	4.0	.0	.0	.0	.0	.0	15800.0
2184	00	4.0	3000.0	.0	.0	.0	.0	13000.0
2185	00	2.0	1500.0	.0	.0	.0	.0	9200.0
2186	00	4.0	.0	.0	.0	.0	.0	18000.0
2187	00	4.0	.0	.0	.0	.0	.0	15000.0
2188	00	4.0	.0	.0	.0	.0	.0	25000.0
2189	00	4.0	.0	.0	.0	.0	.0	30000.0
2190	00	4.0	.0	.0	.0	.0	.0	20000.0
2191	00	4.0	.0	.0	.0	.0	.0	15000.0
2192	00	4.0	.0	.0	.0	.0	.0	45000.0
2193	00	4.0	.0	.0	.0	.0	.0	11000.0
2194	00	4.0	.0	.0	.0	.0	.0	55000.0
2195	00	4.0	.0	.0	.0	.0	.0	50000.0
2196	00	4.0	.0	.0	.0	.0	.0	40000.0
2197	00	4.0	.0	.0	.0	.0	.0	45000.0

รูปที่ ข.7 ผลลัพธ์การแทนค่าสูญหายด้วยวิธีมีฐานของค่าใกล้เคียง

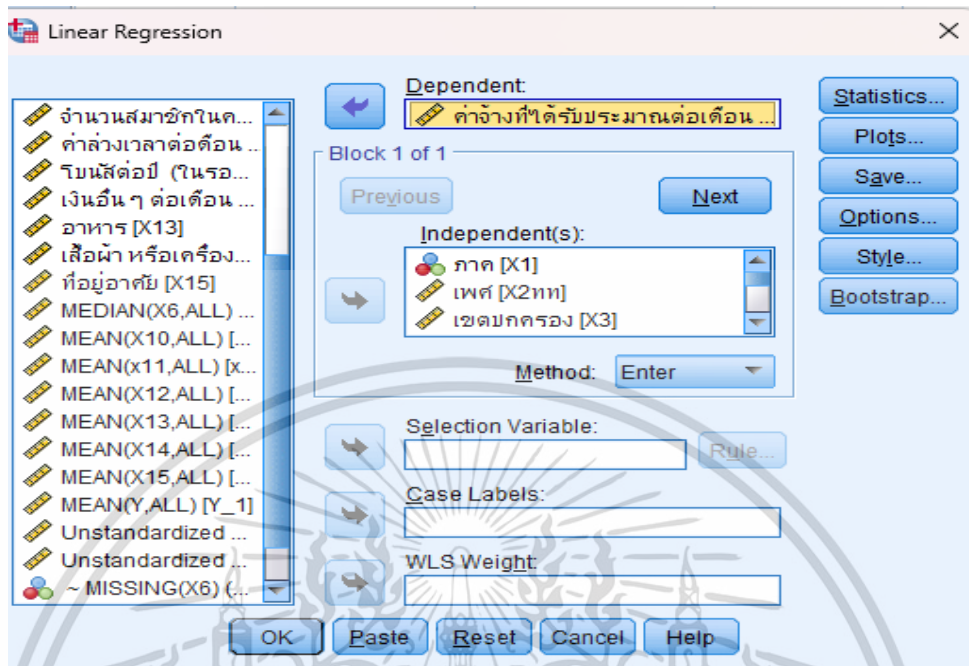
1.3 การตัดค่านอกเกณฑ์

1) คลิก Analyze เลือก Regression เลือก Linear

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ระบุว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

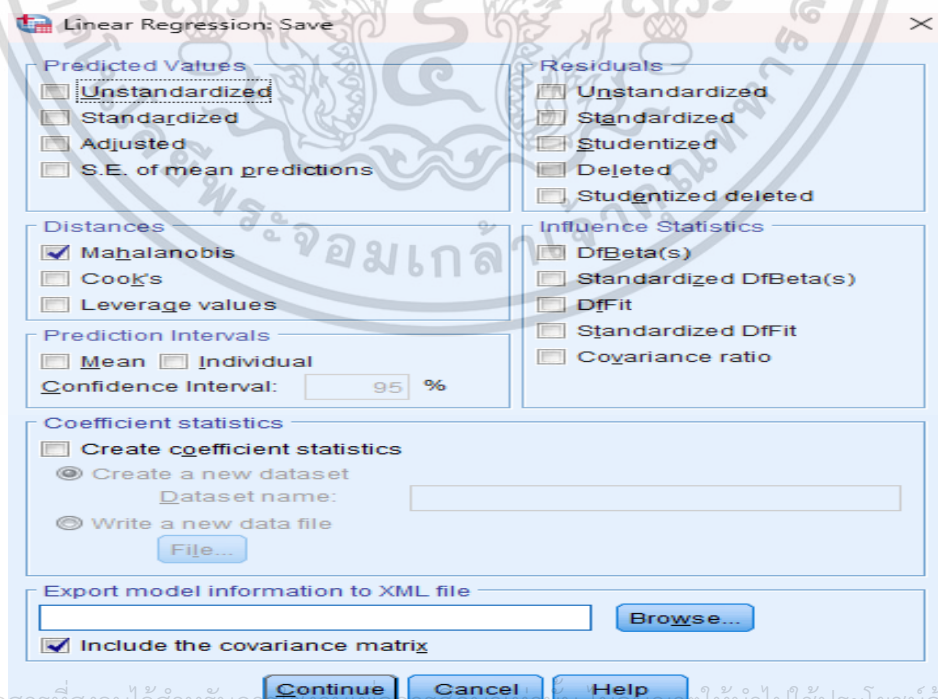
รูปที่ ข.8 การเลือกคำสั่งเพื่อหาค่านอกเกณฑ์

- 2) ช่อง Dependent ใส่ตัวแปรตาม ช่อง Independent ใส่ตัวแปรอิสระทุกตัว
คลิก save



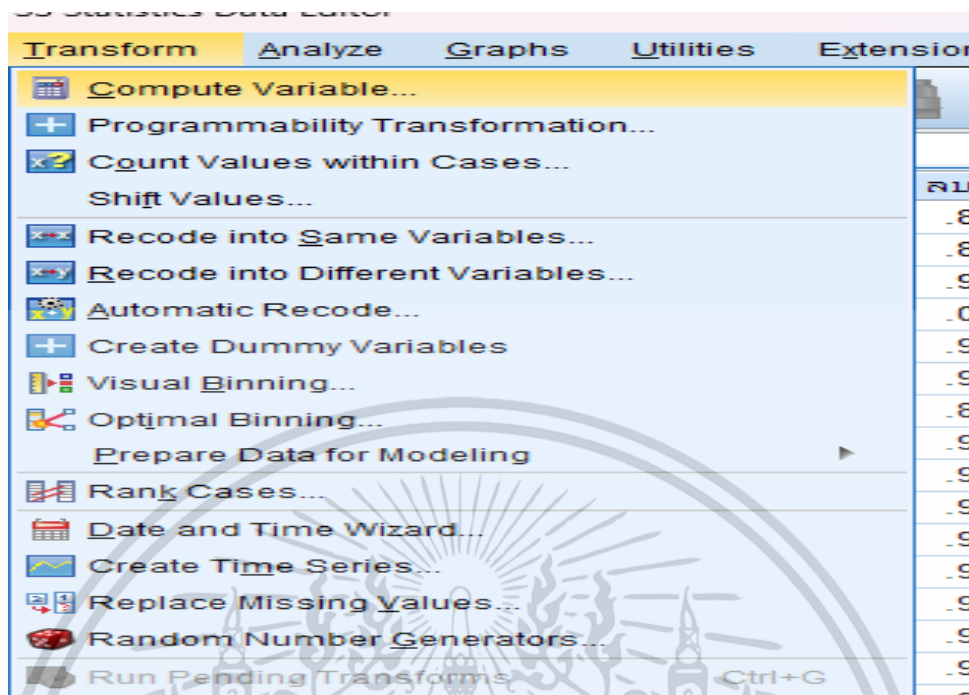
รูปที่ ข.9 การเลือกคำสั่งเพื่อหาค่านอกเกณฑ์

- 3) คลิก Mahatanobis
คลิก Continue
คลิก OK



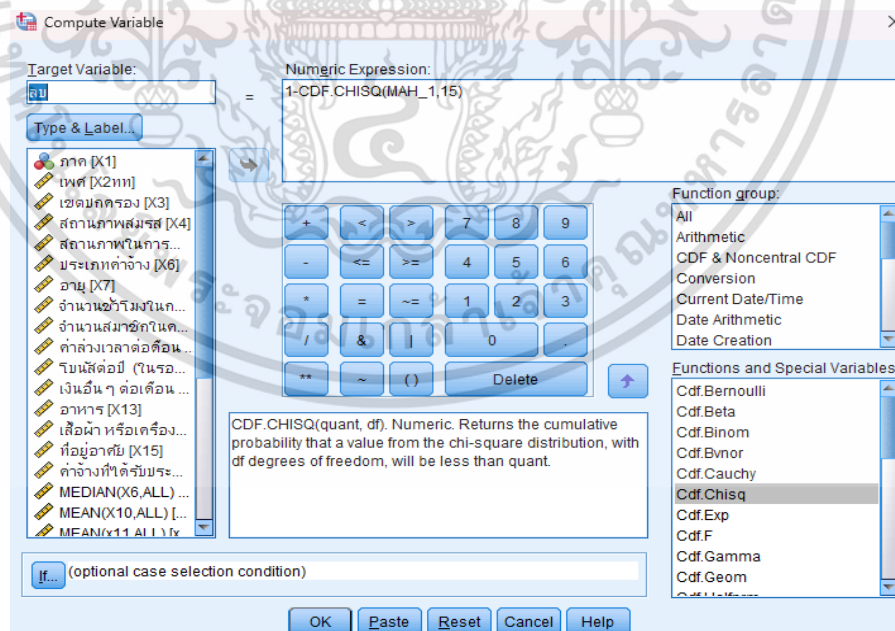
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมีรูปที่ ข.10 การเลือกคำสั่งเพื่อหาค่านอกเกณฑ์

4) คลิก Transform เลือก Compute Variable



รูปที่ ข.11 การเลือกคำสั่งเพื่อตัดค่านอกเกณฑ์

5) ช่อง Target variable ให้ตั้งชื่อตามที่เราต้องการ
 ช่อง Numeric Expression [X1]-CDF.CHISQ(MAH_1,15) หมายถึง 15 คือค่า df
 คลิก OK



รูปที่ ข.12 การเลือกคำสั่งเพื่อตัดค่านอกเกณฑ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6) กัดค่า Decimals ให้เป็น 5 แล้วตัดตัวที่ค่าต่ำกว่า 0.001 ออก

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	X1	Numeric	1	0	ภาค	{1, ๒}...	None	8	Right	Nominal
2	X2	Numeric	1	0	เพศ	{1, ๒}...	None	8	Right	Scale
3	X3	Numeric	1	0	เขตปกครอง	{1, ๒}...	None	8	Right	Scale
4	X4	Numeric	1	0	สถานภาพสมรส	{1, ๒}...	None	8	Right	Scale
5	X5	Numeric	1	0	สถานภาพในครอบครัว	{4, ๕}...	None	8	Right	Scale
6	X6	Numeric	1	0	ประเภทค่าจ้าง	{1, ๒}...	None	8	Right	Scale
7	X7	Numeric	3	0	อายุ	None	None	8	Right	Scale
8	X8	Numeric	2	0	จำนวนชั่วโมงทำงาน	None	None	8	Right	Scale
9	X9	Numeric	2	0	จำนวนสมาชิกใน...	None	None	8	Right	Scale
10	X10	Numeric	5	0	ค่าล่วงเวลาต่อ...	None	None	8	Right	Scale
11	X11	Numeric	6	0	วันหยุดต่อปี (หรือ...	None	None	8	Right	Scale
12	X12	Numeric	5	0	เงินอื่น ๆ ต่อเดือน	None	None	8	Right	Scale
13	X13	Numeric	5	0	อาหาร	None	None	8	Right	Scale
14	X14	Numeric	5	0	เสื้อผ้า หรือเครื่อง...	None	None	8	Right	Scale
15	X15	Numeric	5	0	ที่อยู่อาศัย	None	None	8	Right	Scale
16	Y	Numeric	6	0	ค่าจ้างที่เตรียมประ...	None	None	8	Right	Scale
17	มีงานทำ	Numeric	1	0	ระหว่าง 7 วันก่อน...	None	None	8	Right	Scale
18	filter_\$	Numeric	1	0	ลบ > 0.001 (FIL...	{0, Not Sele...	None	10	Right	Nominal
19	MAH_1	Numeric	11	5	Mahalanobis Di...	None	None	13	Right	Scale
20	ลบ	Numeric	8	5		None	None	10	Right	Scale
21	PRE_1	Numeric	11	5	Unstandardized...	None	None	13	Right	Scale
22	PRES_1	Numeric	11	5	Unstandardized...	None	None	13	Right	Scale

รูปที่ ข.13 การเลือกคำสั่งเพื่อตัดค่าออกเกิน

2. ขั้นตอนการตรวจสอบข้อสมมติก่อนการวิเคราะห์การถดถอยพหุคูณ

2.1 การตรวจสอบความสัมพันธ์เชิงเส้น

1) คลิก Analyze เลือก Correlate เลือก Bivariate

The screenshot shows the 'Analyze' menu in IBM SPSS Statistics. The 'Correlate' option is selected, and the 'Bivariate...' sub-menu is open. The 'Bivariate...' option is highlighted in yellow. The background shows a data editor window with a grid of data points.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในการเรียนการสอนเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสาร
รูปที่ ข.14 การเลือกคำสั่งในการวิเคราะห์ความสัมพันธ์

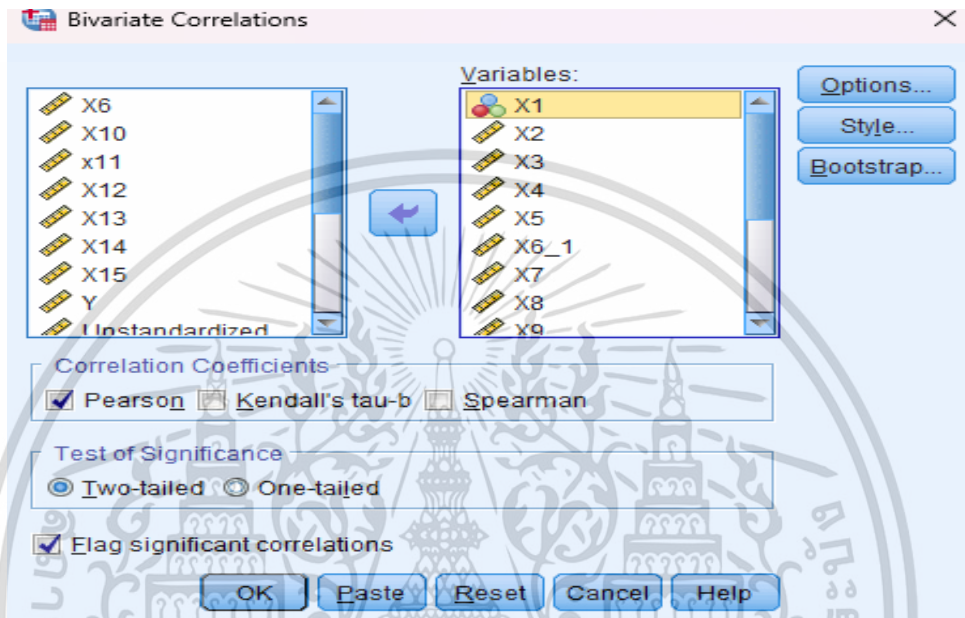
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) ใส่ตัวแปร $X_1, X_2, X_3, X_4, X_5, X_{6_1}, X_7, X_8, X_9, X_{10_1}, X_{11_1}, X_{12_1}, X_{13_1}, X_{14_1}, X_{15_1}$
 ในช่อง Variables

ในช่อง Correlation Coefficients เลือก Pearson

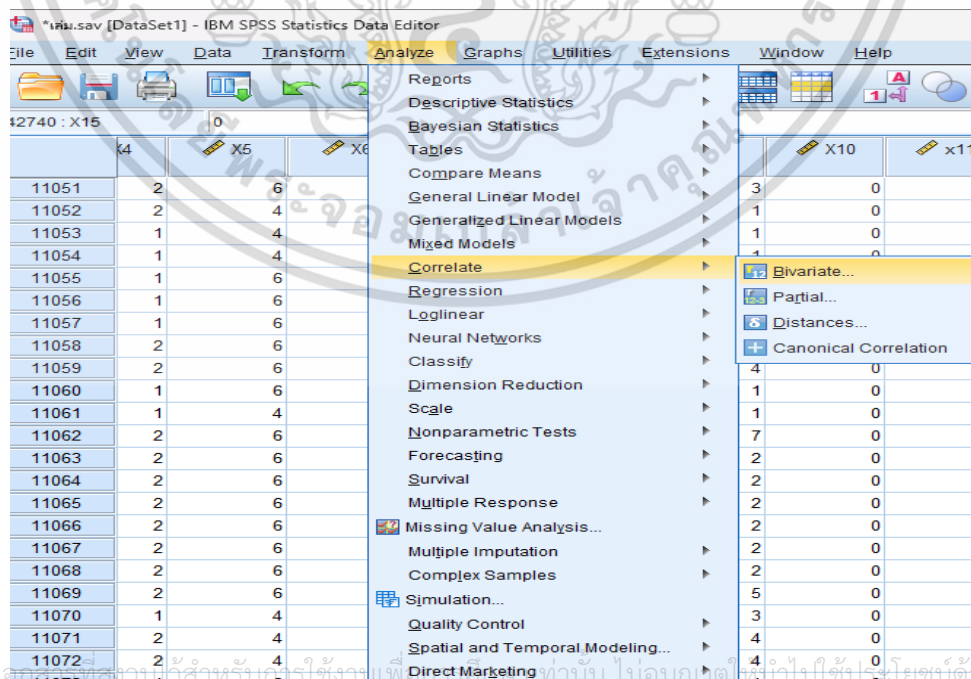
ในช่อง Test of Significance เลือก Two-tailed

คลิก OK



รูปที่ ข.15 การเลือกวิธีตรวจสอบความสัมพันธ์เชิงเส้น สำหรับตัวแปรเชิงปริมาณ

3) คลิก Analyze เลือก Correlate เลือก Bivariate



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกรใช้งานเพื่อการศึกษาเท่านั้น ไมอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

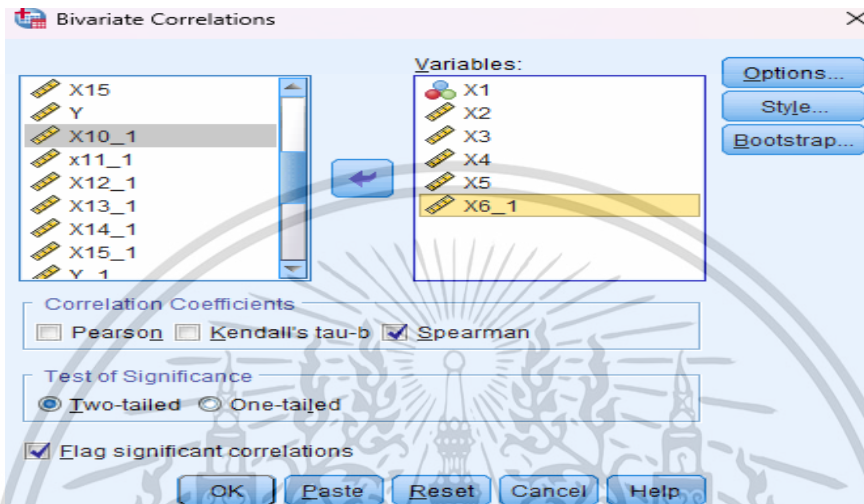
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกรูปที่ ข.16 การเลือกคำสั่งในการวิเคราะห์ความสัมพันธ์ สารทุกครั้งที่มีการนำไปใช้

4) ใส่ตัวแปร $X_1, X_2, X_3, X_4, X_5, X_{6_1}$ ในช่อง Variables

ในช่อง Correlation Coefficients เลือก Spearman

ในช่อง Test of Significance เลือก Two-tailed

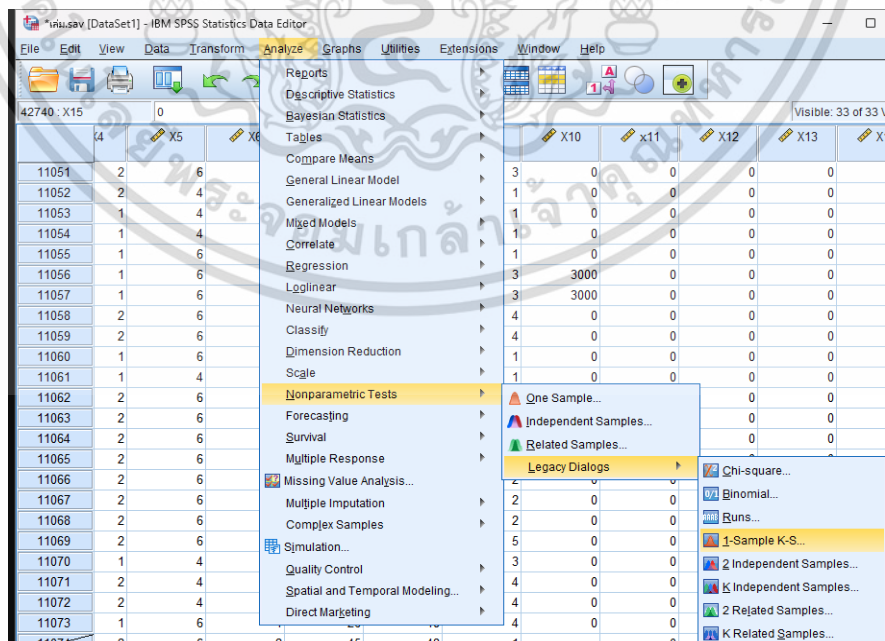
คลิก OK



รูปที่ ข.17 การเลือกวิธีตรวจสอบความสัมพันธ์เชิงเส้น สำหรับตัวแปรเชิงคุณภาพ

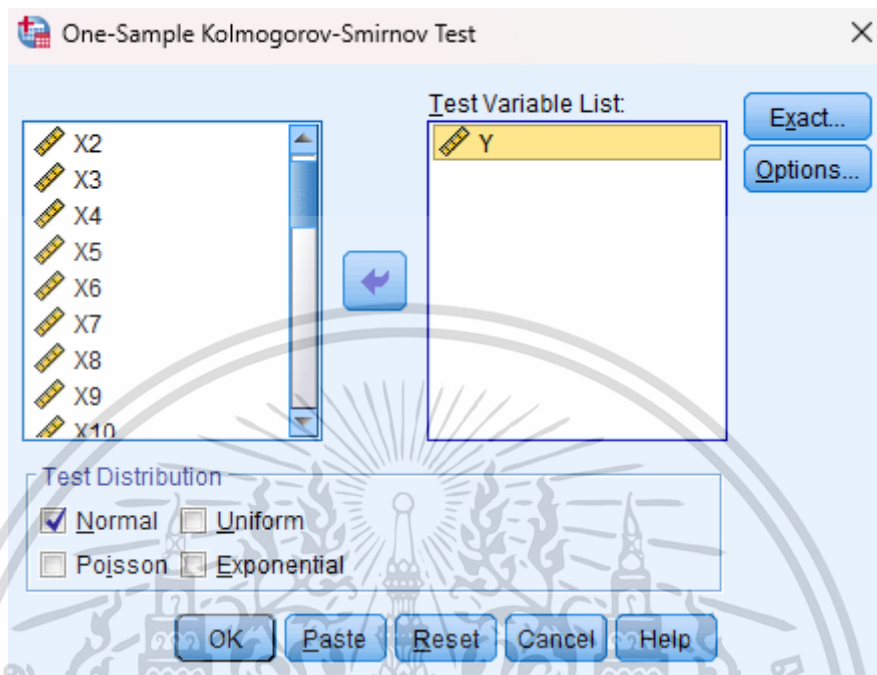
2.2 การตรวจสอบความแจกแจงปกติของตัวแปรตาม

1) กด Analyze เลือก Nonparametric Tests เลือก Legacy Dialogs เลือก 1 Sample K-S



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับรูปที่ ข.18 การเลือกวิธีตรวจสอบความแจกแจง
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

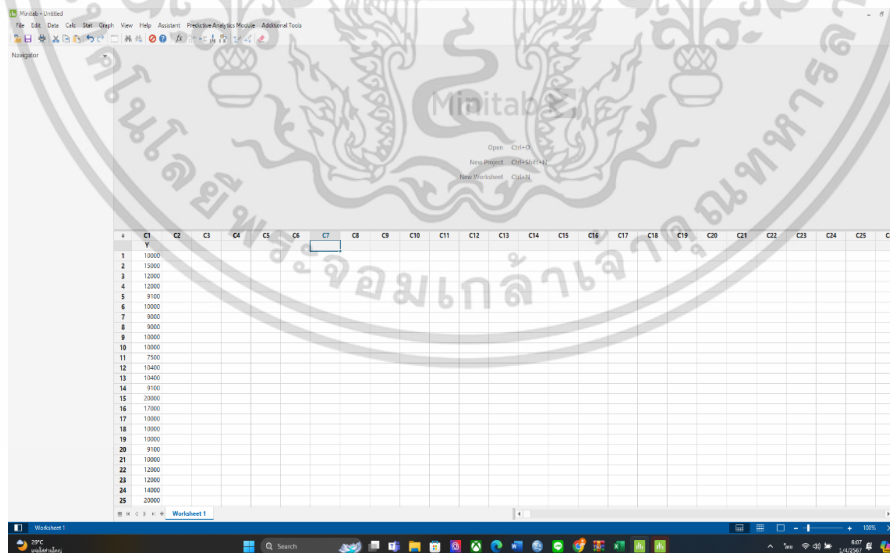
- 2) ในช่อง Test Variables List เลือก รายได้ของลูกจ้างในประเทศไทย (Y)
 ในช่อง Test Distribution เลือก Normal
 คลิก OK



รูปที่ ข.19 การเลือกวิธีตรวจสอบความแจกแจงปกติของตัวแปรตาม

การแปลงข้อมูลสำหรับการแจกแจงไม่เป็นปกติ

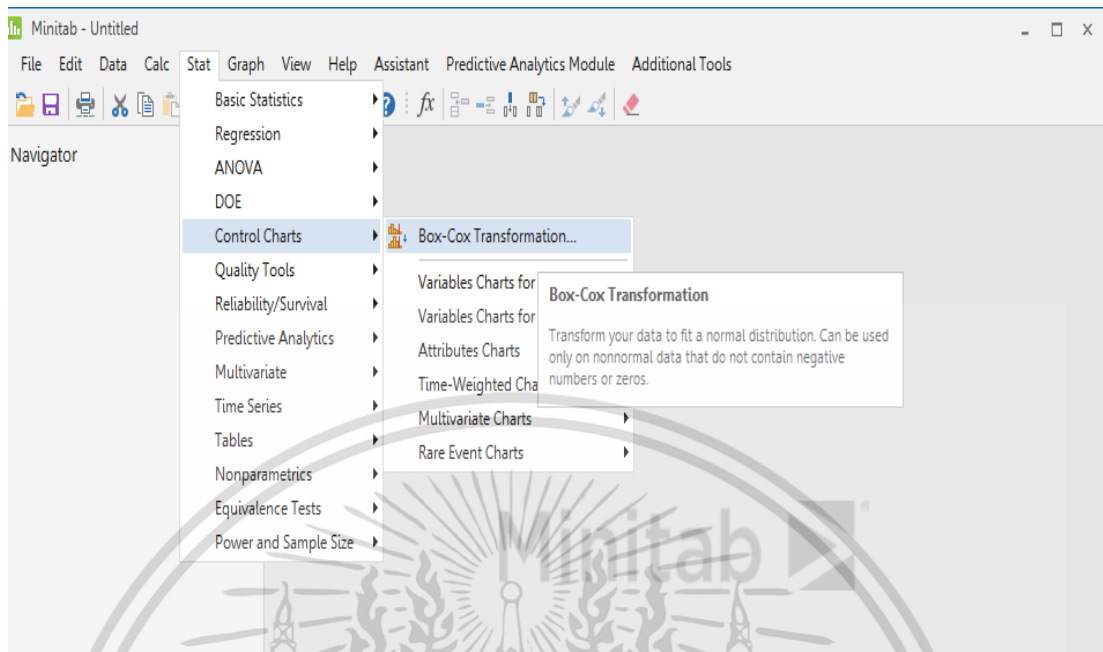
- 1) สำเนา (copy) ข้อมูลตัวแปร Y จากโปรแกรม SPSS มาใส่ไว้ในโปรแกรม MINITAB



รูปที่ ข.20 การสำเนาข้อมูลมาใส่ใน MINITAB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) คลิก Stat เลือก Control Charts เลือก Box-Cox Transformation

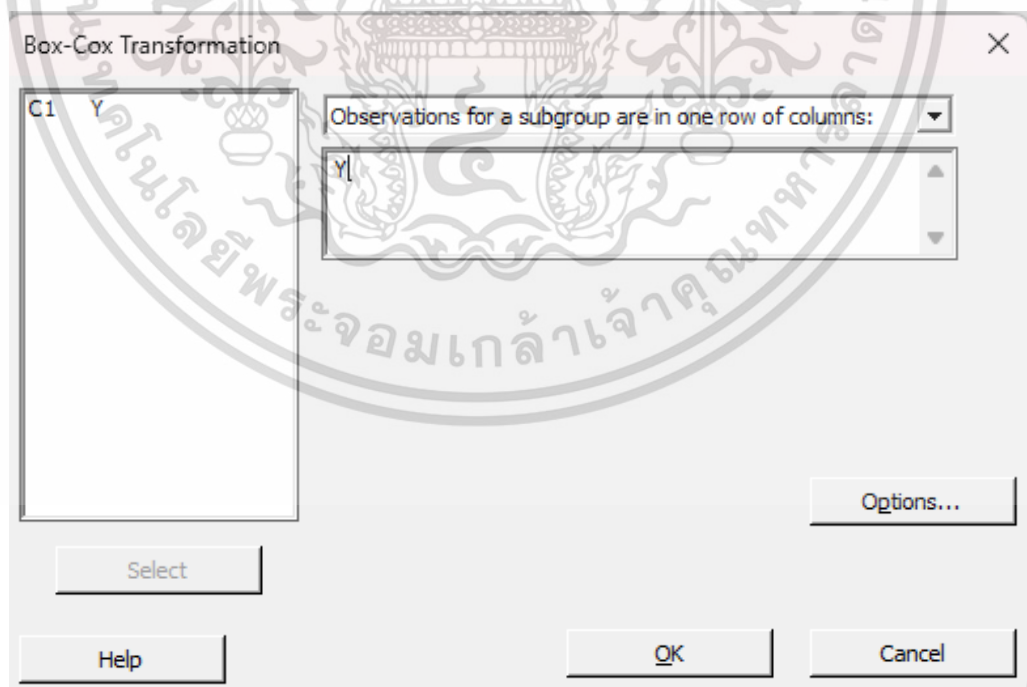


รูปที่ ข.21 วิธีแปลงข้อมูลใน MINITAB

3) เลือก Observations for a subgroup are in one row of columns

เลือก Y

คลิก OK



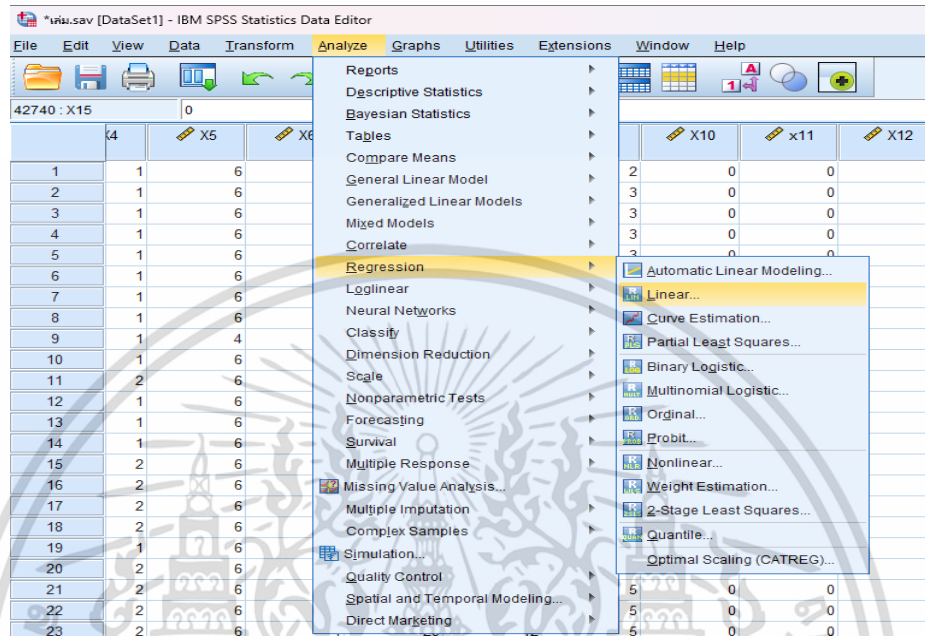
รูปที่ ข.22 การเลือกวิธีแปลงข้อมูลใน MINITAB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. ขั้นตอนการตรวจสอบความสัมพันธ์ของตัวแปรอิสระในสมการและการวิเคราะห์การถดถอยพหุคูณ

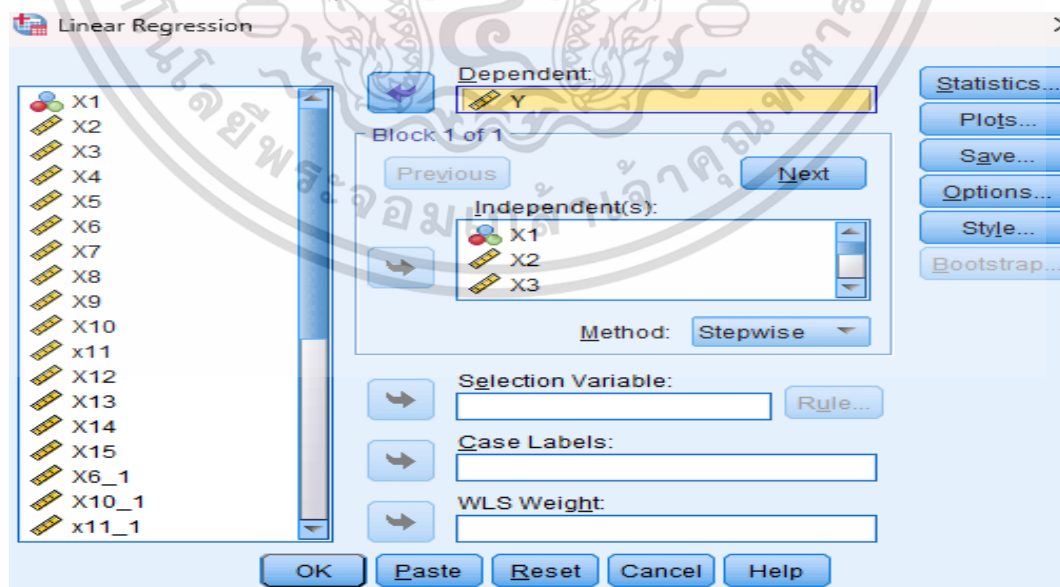
การวิเคราะห์การถดถอยพหุคูณ

1) คลิก Analyze เลือก Regression เลือก Linear



รูปที่ ข.23 การเลือกการทดสอบเชิงเส้น

2) ในช่อง Dependent เลือก Y
 ในช่อง Independent(s) เลือก X ทุกตัว
 ในส่วนของ Method เลือก Stepwise



รูปที่ ข.24 การเลือกการถดถอยทีละขั้น

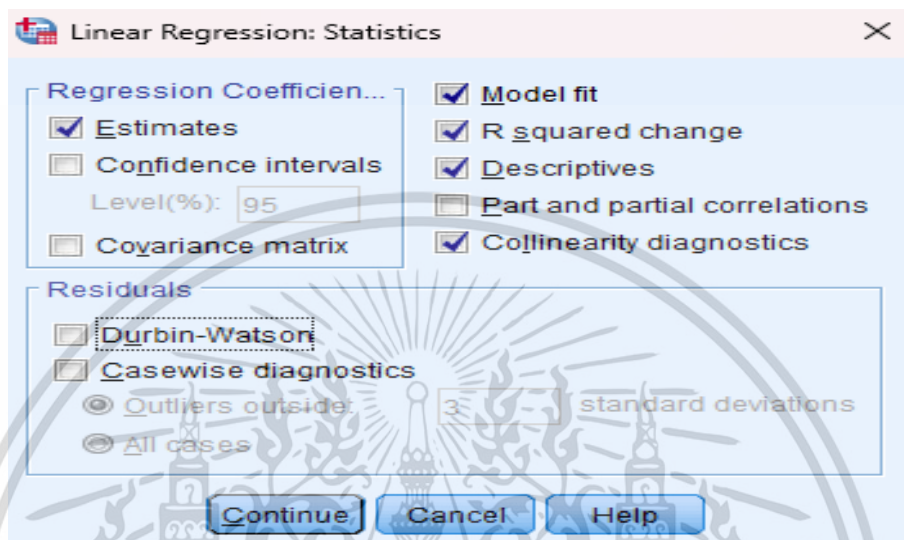
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) คลิก Statistics

ในส่วนของ Regression Coefficients เลือก Estimates, Model fit, R squared change, Descriptives และ Collinearity diagnostics

คลิก Continue

คลิก Save

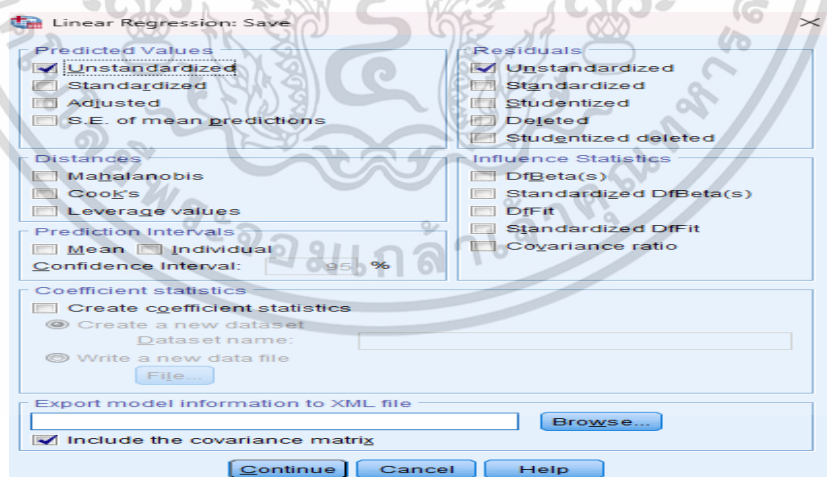


รูปที่ ข.25 เลือกวิธีวิเคราะห์การถดถอยพหุคูณโดยโปรแกรม SPSS

4) คลิก options ในช่อง Predicted Values เลือก Unstandardized

ในช่อง Residuals เลือก Unstandardized คลิก Continue

คลิก OK



รูปที่ ข.26 การเลือกค่าทำนายและส่วนเหลือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

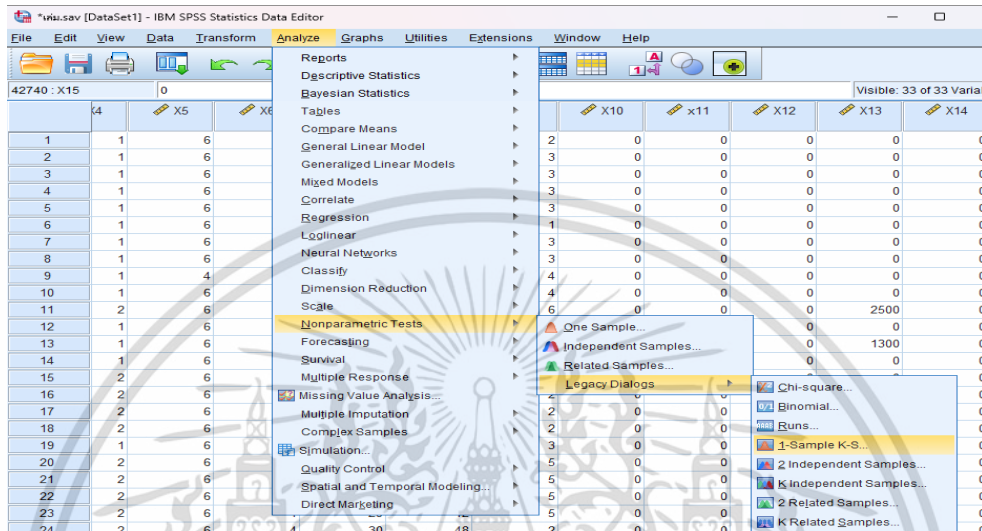
4. ขั้นตอนตรวจสอบข้อสมมติหลังการวิเคราะห์การถดถอยพหุคูณ

4.1 การตรวจสอบการแจกแจงปกติของส่วนเหลือ

1) คลิก Analyze เลือก Nonparametric Tests

เลือก Legacy Dialogs

เลือก 1 Sample K-S

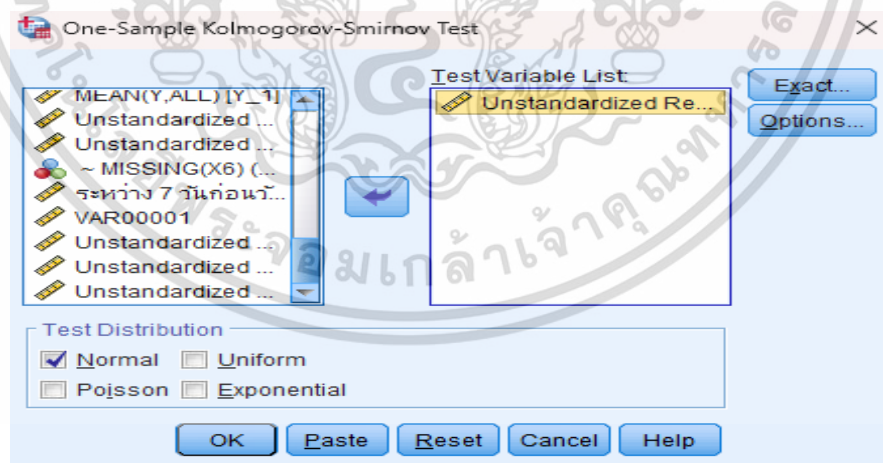


รูปที่ ข.27 การเลือกคำสั่งวิธีการตรวจสอบการแจกแจงปกติของส่วนเหลือ

2) ในช่อง Test Variables List เลือก Unstandardized Residual

ในช่อง Test Distribution เลือก Normal

คลิก OK

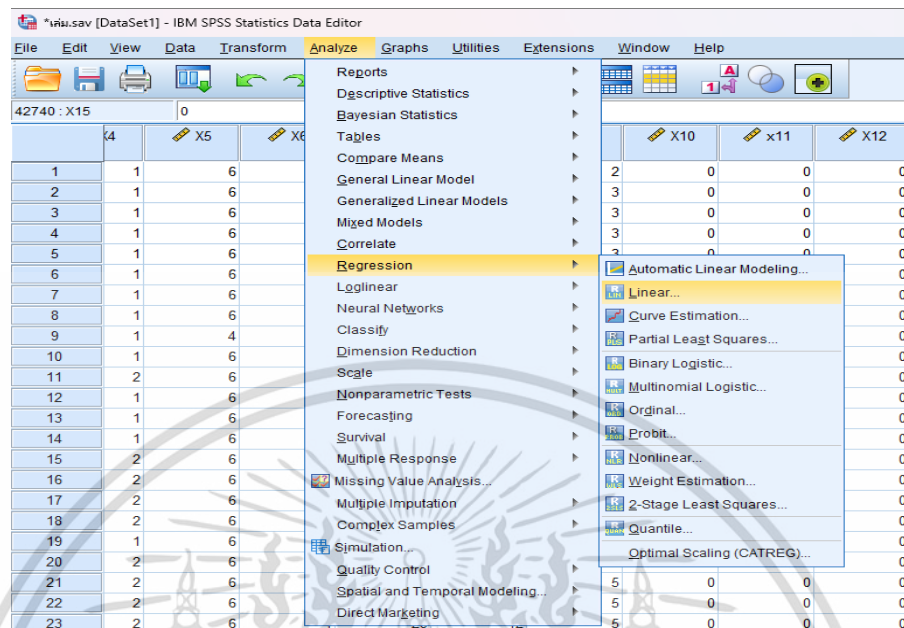


รูปที่ ข.28 การเลือกวิธีตรวจสอบการแจกแจงปกติของส่วนเหลือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 การตรวจสอบการเป็นอิสระกันของส่วนเหลือ

1) คลิก Analyze เลือก Regression เลือก Linear



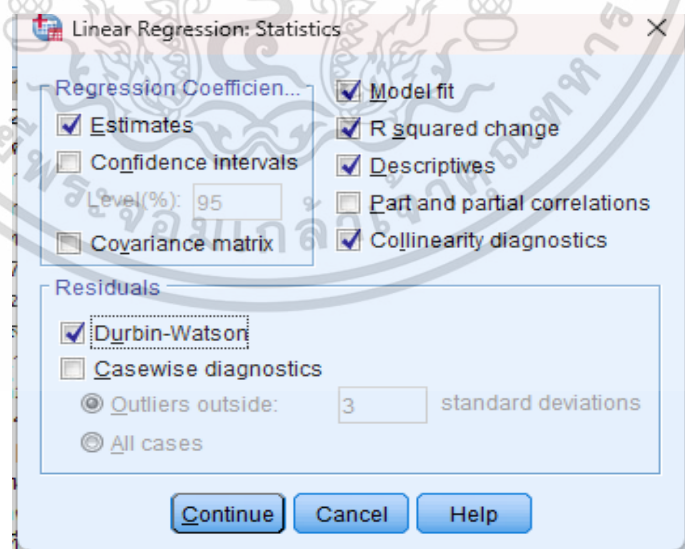
รูปที่ ข.29 การเลือกคำสั่งวิธีการตรวจสอบการเป็นอิสระของส่วนเหลือ

2) คลิก Statistics ในช่อง Regression Coefficient เลือก Estimates, Model fit, R squared change, Descriptives และ Collinearity diagnostics

ในช่อง Residuals เลือก Durbin-Watson

คลิก Continue

คลิก OK

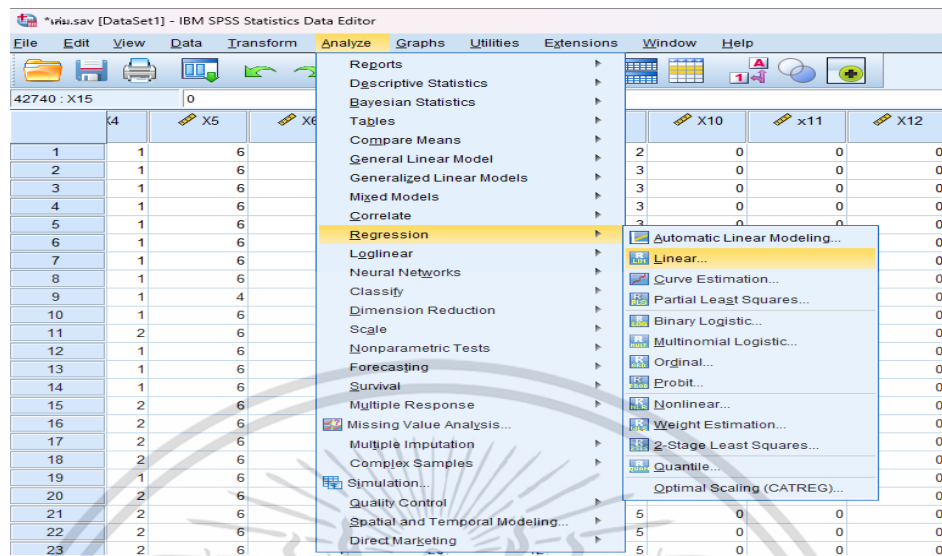


รูปที่ ข.30 การเลือกวิธีตรวจสอบการเป็นอิสระของส่วนเหลือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.3 การตรวจสอบความแปรปรวนเท่ากันของส่วนเหลือ

1) คลิก Analyze เลือก Regression เลือก Linear



รูปที่ ข.31 การเลือกคำสั่งวิธีการตรวจสอบความแปรปรวนของส่วนเหลือ

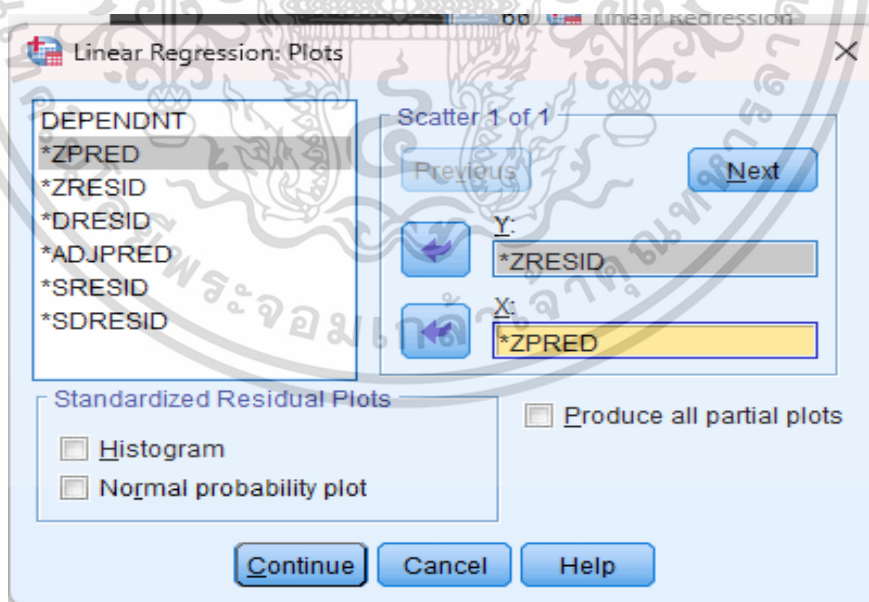
2) คลิก Plots

ในช่อง Y: เลือก ZRESID

ในช่อง X: เลือก ZPRED

คลิก Continue

คลิก OK



รูปที่ ข.32 การเลือกวิธีตรวจสอบการแปรปรวนของส่วนเหลือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



งานทะเบียนคณะวิทยาศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

คำรับรองเล่มปัญหาพิเศษ

วันที่ 28 เดือน พฤษภาคม พ.ศ.2567

ข้าพเจ้า นาย ชัยญนันต์ เร่งกำหนด รหัสประจำตัว 63050614

นาย ธนพล อยู่ทอง รหัสประจำตัว 63050626

นางสาว สียาพัฐ โพธิ์ทอง รหัสประจำตัว 63050674

นักศึกษาหลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชา สถิติประยุกต์ ภาควิชา วิทยาศาสตร์ ขอร้องว่าปัญหาพิเศษเรื่อง ชื่อภาษาไทย ปัจจัยที่มีอิทธิพลต่อรายได้ของลูกจ้างในประเทศไทย

ชื่อภาษาอังกฤษ Factor Affecting the Employee's Income in Thailand

ปีการศึกษา 2566

เป็นผลงานวิจัยที่ได้คัดลอกหรือละเมิดลิขสิทธิ์ของผู้อื่นและได้ผ่านการตรวจสอบความซ้ำซ้อนเรียบร้อยแล้ว และได้แนบเอกสารการตรวจสอบการลอกเลียนงานวรรณกรรมที่ตรวจสอบจากเล่มปัญหาพิเศษฉบับสมบูรณ์แล้ว

โปรแกรมอักขราวิสุทธิ์ 7.04 %

ลงชื่อ สียาพัฐ โพธิ์ทอง (สียาพัฐ โพธิ์ทอง)

นักศึกษา

ลงชื่อ ธนพล อยู่ทอง เร่งกำหนด (ธนพล อยู่ทอง เร่งกำหนด)

นักศึกษา

ลงชื่อ ชาญพล อยู่ทอง (ชาญพล อยู่ทอง)

นักศึกษา

ข้าพเจ้า รศ.สายชล สินสมบูรณ์ทอง อาจารย์ที่ปรึกษาปัญหาพิเศษ ได้ตรวจสอบปัญหาพิเศษของนักศึกษาข้างต้นแล้ว ขอร้องว่าเป็นผลงานวิจัยของนักศึกษาจริงและมีเนื้อหาสมบูรณ์ จึงลงชื่อไว้เป็นหลักฐาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดทอนข้อความใดๆ จากเอกสารฉบับนี้ และต้องแจ้งเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้ (รศ.สายชล สินสมบูรณ์ทอง)

อาจารย์ที่ปรึกษา