การออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนา และการวิเคราะห์ความถดถอยเชิงเส้น

นาย ชาญวิทย์ สินธพพันธุ์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมศาสตร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2543

ISBN 974-347-021-2

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A DESIGN AND DEVELOPMENT OF OBJECT ORIENTED SOFTWARE COMPONENTS FOR DESCRIPTIVE STATISTICS AND LINEAR REGRESSION ANALYSIS

Mr. Chanvit Sintoppun

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2000

ISBN 974-347-021-2

หัวข้อวิทยานิพนธ์	การออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น
โดย	นาย ชาญวิทย์ สินธพพันธุ์
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษา	อาจารย์ ดร. พรศิริ หมื่นไชยศรี

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญามหาบัณฑิต

--คณบดีคณะวิศวกรรมศาสตร์

(ศาสตราจารย์ ดร. สมศักดิ์ ปัญญาแก้ว)

คณะกรรมการสอบวิทยานิพนธ์

--ประธานกรรมการ

(รองศาสตราจารย์ ดร. วันชัย ริ้วไพบูลย์)

--อาจารย์ที่ปรึกษา

(อาจารย์ ดร. พรศิริ หมื่น ไชยศรี)

--กรรมการ

(อาจารย์ ดร. ธาราทิพย์ สุวรรณศาสตร์)

--กรรมการ

(ผู้ช่วยศาสตราจารย์ วิวัฒน์ วัฒนาวุฒิ)

--กรรมการ

(อาจารย์ ดร. ทวิตีย์ เสนีวงศ์ ณ อยุธยา)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ชาญวิทย์ สินธพพันธุ์ : การออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนา (A DESIGN AND DEVELOPMENT OF OBJECT ORIENTED SOFTWARE COMPONENTS FOR DESCRIPTIVE STATISTICS AND LINEAR REGRESSION ANALYSIS) อาจารย์ที่ปรึกษา: อาจารย์ ดร.พรศิริ หมื่นไชยศรี ; 181 หน้า.
ISBN 974-347-021-2

การวิเคราะห์ทางสถิติมีความสำคัญสำหรับงานวิจัยจำนวนมาก ถึงแม้ว่าปัจจุบันนี้จะมีเครื่องมือสำหรับการวิเคราะห์ทางสถิติอยู่มากมาย แต่ก็ยังขาดในส่วนของส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติ งานวิจัยนี้จึงได้ทำการออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น นักพัฒนาโปรแกรมต่างๆสามารถนำส่วนประกอบซอฟต์แวร์ทางสถิตินี้ไปใช้ร่วมเป็นส่วนหนึ่งในการพัฒนาโปรแกรมเชิงวัตถุของตนเมื่อต้องการคำนวณค่าสถิติเชิงพรรณนาหรือการวิเคราะห์ความถดถอยเชิงเส้นได้ ช่วยให้ลดเวลาที่นักพัฒนาโปรแกรมเชิงวัตถุจะต้องเขียนชุดคำสั่งด้วยตัวเอง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชา วิศวกรรมศาสตร์ ลายมือชื่อนิสิต ....
สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่ออาจารย์ที่ปรึกษา ....
ปีการศึกษา 2543

4171421421 : MAJOR COMPUTER SCIENCE

KEYWORD: OBJECT ORIENTED / SOFTWARE COMPONENT / DESCRIPTIVE STATISTICS / LINEAR REGRESSION ANALYSIS

CHANVIT SINTOPPUN: A DESIGN AND DEVELOPMENT OF OBJECT ORIENTED SOFTWARE COMPONENTS FOR DESCRIPTIVE STATISTICS AND LINEAR REGRESSION ANALYSIS. THESIS ADVISOR: DR. PORNSIRI MUENCHAISRI. 181 pp. ISBN 974-347-021-2

Statistics analysis is very important for many researches. Although now there are various tools for statistics analysis, there is still lack of object oriented software components for statistics analysis. This research designs and develops object oriented software components for descriptive statistics and linear regression analysis. Programmers can reuse these software components as part of their object oriented software when they need computation of descriptive statistics or linear regression analysis with fewer manuals coding time and effort.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Department Engineering ....   Student's signature ....

Field of study Computer Science ....   Advisor's signature ....

Academic year 2543 ....

กิตติกรรมประกาศ

ข้าพเจ้าใคร่ขอกราบขอบพระคุณอาจารย์ ดร. พรศิริ หมื่นไชยศรี อาจารย์ที่ปรึกษาวิทยานิพนธ์ของข้าพเจ้า ที่ท่านเป็นผู้ให้ความรู้ คำปรึกษา ความช่วยเหลือต่างๆ ตลอดจนคอยดูแลการทำวิจัยของข้าพเจ้าอย่างดียิ่งจนสำเร็จลุล่วงลงด้วยดี

และขอกราบขอบพระคุณบิดา มารดา และพี่ๆ ที่ได้ให้โอกาสและสนับสนุนในด้านการเงินและกำลังใจแก่ข้าพเจ้าเสมอมา

ชาญวิทย์ สินธพพันธุ์

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ

หน้า

บทคัดย่อภาษาไทย .... ง

บทคัดย่อภาษาอังกฤษ .... จ

กิตติกรรมประกาศ .... ฉ

สารบัญตาราง .... ฐ

สารบัญภาพ .... ฑ

บทที่ 1 บทนำ 1

1.1 ความเป็นมาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ของงานวิจัย 2

1.3 ขอบเขตของการวิจัย 2

1.4 ขั้นตอนและวิธีดำเนินงาน 2

1.5 ประโยชน์ที่คาดว่าจะได้รับ 3

บทที่ 2 แนวคิดและทฤษฎีที่เกี่ยวข้อง 4

2.1 ยูเอ็มแอล 4

2.2 ทฤษฎีทางสถิติ

2.2.1 สถิติเชิงพรรณนา(Descriptive Statistics) 6

2.2.1.1 การนำเสนอข้อมูล 8

2.2.1.2 การแจกแจงความถี่ 9

2.2.1.3 การวัดค่ากลางของข้อมูล 12

2.1.1.3.1 ค่าเฉลี่ย (Mean) 12

2.1.1.3.2 มัธยฐาน (Median) 13

2.1.1.3.3 ฐานนิยม (Mode) 13

2.1.1.3.4 เปอร์เซ็นต์ไทล์ (Percentile) 14

2.1.1.3.5 ควอไทล์ (Quartile) 14

2.2.1.4 การวัดการกระจายของข้อมูล 15

2.1.1.4.1 พิสัย (Range) 15

2.1.1.4.2 พิสัยควอไทล์ (Interquartile Range) 15

สารบัญ(ต่อ)

หน้า

2.2.1.4.3 ค่าแปรปรวน (Variance) 15

2.2.1.4.4 ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) 16

2.2.1.4.5 สัมประสิทธิ์ความแปรผัน (Coefficient of Variation : CV) 16

2.2.1.4.6 ค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (Standard Error of Mean) 17

2.2.1.4.7 การวัดความเบ้ (Skewness) 17

2.2.1.4.8 การวัดความโด่ง (Kurtosis) 17

2.2.2 การวิเคราะห์ความถดถอยเชิงเส้น(Linear Regression Analysis) 18

2.2.2.1 การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย(Simple Linear Regression) 18

2.2.2.1.1 สมมติฐานของการวิเคราะห์ความถดถอย 20

2.2.2.1.2 การประมาณค่าพารามิเตอร์ของสมการความถดถอย 20

2.2.2.1.3 การประมาณค่าแปรปรวนของความคลาดเคลื่อน 22

2.2.2.1.4 การประมาณค่าแบบช่วงและการทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์ความถดถอย 22

2.2.2.1.5 การทดสอบสัมประสิทธิ์ความถดถอยโดยใช้การวิเคราะห์ความแปรปรวน 24

2.2.2.1.6 สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination) 25

2.2.2.1.7 สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) 25

2.2.2.1.8 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์ 26

2.2.2.2 การวิเคราะห์ความถดถอยเชิงเส้นเชิงซ้อน(Multiple Linear Regression) 26

2.2.2.2.1 สมมติฐานของการวิเคราะห์ความถดถอยเชิงซ้อน 27

2.2.2.2.2 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงซ้อน 27

2.2.2.2.3 การทดสอบสมการความถดถอยเชิงเส้นเชิงซ้อนโดยใช้การวิเคราะห์ความ แปรปรวน 28

2.2.2.2.4 การประมาณค่าคลาดเคลื่อนของความถดถอย 29

2.2.2.2.5 การประมาณค่าสัมประสิทธิ์ความถดถอยแบบช่วง 29

2.2.2.2.6 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์ความถดถอย 30

2.2.2.2.7 สัมประสิทธิ์การตัดสินใจเชิงซ้อน 30

สารบัญ(ต่อ)

หน้า

2.2.2.2.8 การตรวจสอบสมมติฐานของการวิเคราะห์ความถดถอย 31

2.2.2.2.9 การเกิดปัญหามัลติคอลลิเนียริตี (Multicollinearity) 33

2.2.2.2.10 การเลือกตัวแปรอิสระเข้าสมการความถดถอย 34

บทที่ 3 การวิเคราะห์และออกแบบระบบ 38

3.1 ความสามารถในการคำนวณค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 38

3.2 การวิเคราะห์และออกแบบคลาส 39

3.2.1 คลาสสำหรับสถิติเชิงพรรณนา 39

3.2.2 คลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 57

3.2.3 คลาสสำหรับตารางสถิติ 87

3.3แผนผังแสดงลำดับการคำนวณของส่วนประกอบซอฟต์แวร์สำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 91

3.3.1 แผนผังลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม 91

3.3.2 แผนผังลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบไม่จัดกลุ่ม 91

3.3.3 แผนผังลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม 92

3.3.4 แผนผังลำดับของการคำนวณค่าความถี่ของข้อมูลแบบไม่จัดกลุ่ม 92

3.3.5 แผนผังลำดับของการคำนวณค่าการกระจายของข้อมูลแบบจัดกลุ่ม 93

3.3.6 แผนผังลำดับของการคำนวณค่าการกระจายของข้อมูลแบบไม่จัดกลุ่ม 93

3.3.7 แผนผังลำดับของการคำนวณค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น 94

บทที่ 4 การพัฒนาระบบ 96

4.1 การพัฒนาโปรแกรมสำหรับคลาส DescriptiveStatistics 96

4.2 การพัฒนาโปรแกรมสำหรับคลาส FreqDistribution 97

4.3 การพัฒนาโปรแกรมสำหรับคลาส GroupFreqDistribution 97

4.4 การพัฒนาโปรแกรมสำหรับคลาส UngroupFreqDistribution 98

4.5 การพัฒนาโปรแกรมสำหรับคลาส CentralTendency 100

4.6 การพัฒนาโปรแกรมสำหรับคลาส GroupCentralTendency 101

4.7 การพัฒนาโปรแกรมสำหรับคลาส UngroupCentralTendency 103

สารบัญ(ต่อ)

หน้า

4.8 การพัฒนาโปรแกรมสำหรับคลาส Dispersion 105
4.9 การพัฒนาโปรแกรมสำหรับคลาส GroupDispersion 106
4.10 การพัฒนาโปรแกรมสำหรับคลาส UngroupDispersion 107
4.11 การพัฒนาโปรแกรมสำหรับคลาส LinearRegression 104
4.12 การพัฒนาโปรแกรมสำหรับคลาส Enter 110
4.13 การพัฒนาโปรแกรมสำหรับคลาส Forward 111
4.14 การพัฒนาโปรแกรมสำหรับคลาส Backward 112
4.15 การพัฒนาโปรแกรมสำหรับคลาส Stepwise 114
4.16 การพัฒนาโปรแกรมสำหรับคลาส ModelSummary 116
4.17 การพัฒนาโปรแกรมสำหรับคลาส Correlations 119
4.18 การพัฒนาโปรแกรมสำหรับคลาส Anova 120
4.19 การพัฒนาโปรแกรมสำหรับคลาส Coefficients 124
4.20 การพัฒนาโปรแกรมสำหรับคลาส ResidualsStatis 125
4.21 การพัฒนาโปรแกรมสำหรับคลาส LinearRegressionOption 130
4.22 การพัฒนาโปรแกรมสำหรับคลาส StatTable 132

บทที่ 5 การใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 137
5.1 การจัดเก็บไฟล์ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 137
5.2 ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 137
5.2.1 ส่วนประกอบสำหรับการนับค่าความถี่ของข้อมูล 138
5.2.2 ส่วนประกอบสำหรับการคำนวณสถิติแสดงค่ากลางของข้อมูล 138
5.2.3 ส่วนประกอบสำหรับสำหรับคำนวณสถิติที่วัดการกระจายของข้อมูล 138
5.2.4 ส่วนประกอบสำหรับสำหรับวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีเอนเตอร์ 138

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญ(ต่อ)

หน้า

5.2.5 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีฟอร์เวิร์ด 139
5.2.6 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีแบ็คเวิร์ด 139
5.2.7 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีสเตปไวส์ 139
5.3 การใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 140
5.4 ตัวอย่างการใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 140
5.4.1 การกำหนดชื่อคลาสที่จะใช้งาน 140
5.4.2 การสร้างอินสแตนซ์ของคลาส 141
5.4.3 การรับค่าคุณสมบัติของอินสแตนซ์ 141

บทที่ 6 การใช้งาน 142
6.1 การเข้าสู่โปรแกรมสำหรับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น 142
6.1.1 เมนู File 143
6.1.2 เมนู Statistics 143
6.1.3 เมนู Help 144
6.2 การวิเคราะห์สถิติเชิงพรรณนา 145
6.2.1 การกำหนดค่าสถิติสำหรับการวิเคราะห์สถิติเชิงพรรณนา 147
6.2.2 การกำหนดค่าแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา 149
6.2.3 การแสดงแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา 150
6.3 การวิเคราะห์ความถดถอยเชิงเส้น 151
6.3.1 การกำหนดค่าสถิติสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 153
6.3.2 การกำหนดค่าสถิติที่ต้องการบันทึกค่า 154
6.3.3 การกำหนดเงื่อนไขการเลือกตัวแปรอิสระเข้าสมการความถดถอย 155
6.3.4 การแสดงแผนภูมิสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 156

สารบัญ(ต่อ)

หน้า

บทที่ 7 การทดสอบ 158
7.1 การทดสอบส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนา 158
7.2 การทดสอบส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 161
7.2.1 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นวิธีเอนเตอร์ 161
7.2.1.1 ทดสอบการพิจารณานำตัวแปรอิสระเข้าหรือออกจากสมการความถดถอยเชิงเส้น 161
7.2.1.2 การทดสอบสถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น 162
7.2.1.3 การทดสอบการวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น 163
7.2.1.4 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น 164
7.2.1.5 การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว 165
7.2.1.6 การทดสอบการคำนวณค่าคลาดเคลื่อน 166
7.2.2 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นวิธีฟอร์เวิร์ด 167
7.2.2.1 ทดสอบการพิจารณานำตัวแปรอิสระเข้าหรือออกจากสมการความถดถอยเชิงเส้น 167
7.2.2.2 การทดสอบค่าสถิติสรุปแต่ละรูปแบบของสมการความถดถอย 169
7.2.2.3 การทดสอบการวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น 170
7.2.2.4 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น 171
7.2.2.5 การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว 173
7.2.3 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นวิธีแบ็คเวิร์ด 174
7.2.4 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นวิธีสเตปไวส์ 175

บทที่ 8 บทสรุปและข้อเสนอแนะ 178
8.1 บทสรุป 178
8.2 ข้อเสนอแนะ 179

รายการอ้างอิง 180
ประวัติผู้เขียนวิทยานิพนธ์ 181

สารบัญตาราง

หน้า

ตารางที่ 2.1 การแจกแจงความถี่ของลักษณะที่สนใจที่เป็นไปได้ทั้งหมด 9
ตารางที่ 2.2 การแจกแจงความถี่สำหรับช่วงค่าของลักษณะที่สนใจ 10
ตารางที่ 2.3 ขอบเขตจำกัดชั้นของการแจกแจงความถี่สำหรับช่วงค่าของลักษณะที่สนใจ 10
ตารางที่ 3.1 รายละเอียด ของคลาส DescriptiveStatistics 41
ตารางที่ 3.2 รายละเอียด ของคลาส FreqDistribution 42
ตารางที่ 3.3 รายละเอียด ของคลาส GroupFreqDistribution 44
ตารางที่ 3.4 รายละเอียด ของคลาส UngroupFreqDistribution 46
ตารางที่ 3.5 รายละเอียด ของคลาส CentralTendency 47
ตารางที่ 3.6 รายละเอียด ของคลาส GroupCentralTendency 48
ตารางที่ 3.7 รายละเอียด ของคลาส UngroupCentralTendency 49
ตารางที่ 3.8 รายละเอียด ของคลาส Dispersion 51
ตารางที่ 3.9 รายละเอียด ของคลาส GroupDispersion 53
ตารางที่ 3.10 รายละเอียด ของคลาส UngroupDispersion 54
ตารางที่ 3.11 รายละเอียด ของคลาส LinearRegression 59
ตารางที่ 3.12 รายละเอียด ของคลาส Enter 62
ตารางที่ 3.13 รายละเอียด ของคลาส Backward 63
ตารางที่ 3.14 รายละเอียด ของคลาส Forward 64
ตารางที่ 3.15 รายละเอียด ของคลาส Stepwise 66
ตารางที่ 3.16 รายละเอียด ของคลาส ModelSummary 68
ตารางที่ 3.17 รายละเอียด ของคลาส Correlations 71
ตารางที่ 3.18 รายละเอียด ของคลาส Anova 74
ตารางที่ 3.19 รายละเอียด ของคลาส Coefficients 78
ตารางที่ 3.20 รายละเอียด ของคลาส ResidualStatistics 81
ตารางที่ 3.21 รายละเอียด ของคลาส LinearRegressionOption 83
ตารางที่ 3.22 รายละเอียด ของคลาส StatTable 88

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญรูป

หน้า

รูปที่ 2. 1 ตัวแบบต่างๆที่ได้นำข้อดีของแต่ละตัวแบบมาประกอบไว้ในยูเอ็มแอล 4
รูปที่ 2.2 ตัวอย่างผังยูสเคส 6
รูปที่ 2.3 ตัวอย่างผังโครงสร้างคลาส 6
รูปที่ 2.4 การนำเสนอในรูปตาราง 8
รูปที่ 2.5 การนำเสนอข้อมูลในรูปแผนภาพการกระจาย (Scatter Plot) 8
รูปที่ 2.6 ตัวอย่างการนำเสนอข้อมูลในรูปฮิสโตแกรม 9
รูปที่ 2.7 ตัวอย่างแผนภาพการกระจายซึ่ง ความสัมพันธ์ระหว่าง X และ Y 19
รูปที่ 2.8 กราฟความสัมพันธ์ระหว่าง e กับ X e กับ Y หรือ e กับ $\hat{Y}$ 19
รูปที่ 2.9 กราฟความสัมพันธ์ระหว่า ง e กับ t 31
รูปที่ 3.1 ยูสเคสแสดงความสามารถของส่วนประกอบทางซอฟต์แวร์เชิงวัตถุ
สำหรับสถิติเชิงพรรณนาและการวิเคราะห์ ความถดถอยเชิงเส้น 38
รูปที่ 3.2 ความสัมพันธ์ระหว่างคลาสสำหรับสถิติเชิงพรรณนา 40
รูปที่ 3.3 ส่วนประกอบของคลาส DescriptiveStatistics 40
รูปที่ 3.4 ส่วนประกอบของคลาส FreqDistribution 42
รูปที่ 3.5 ส่วนประกอบของคลาส GroupFreqDistribution 44
รูปที่ 3.6 ส่วนประกอบของคลาส UngroupFreqDistribution 45
รูปที่ 3.7 ส่วนประกอบของคลาส CentralTendency 46
รูปที่ 3.8 ส่วนประกอบของคลาส GroupCentralTendency 48
รูปที่ 3.9 ส่วนประกอบของคลาส UngroupCentralTendency 49
รูปที่ 3.10 ส่วนประกอบของคลาส Dispersion 51
รูปที่ 3.11 ส่วนประกอบของคลาส GroupDispersion 53
รูปที่ 3.12 ส่วนประกอบของคลาส UngroupDispersion 54
รูปที่ 3.13 ความสัมพันธ์ของคลาสต่างๆสำหรับสถิติเชิงพรรณนา 56
รูปที่ 3.14 ความสัมพันธ์ระหว่างคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 58
รูปที่ 3.15 ส่วนประกอบของคลาส LinearRegression 59
รูปที่ 3.16 ส่วนประกอบของคลาส Enter 62

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

สารบัญรูป (ต่อ)

หน้า

รูปที่ 3.17 ส่วนประกอบของคลาส Backward 63
รูปที่ 3.18 ส่วนประกอบของคลาส Forward 64
รูปที่ 3.19 ส่วนประกอบของคลาส Stepwise 66
รูปที่ 3.20 ส่วนประกอบของคลาส ModelSummary 68
รูปที่ 3.21 ส่วนประกอบของคลาส Correlations 71
รูปที่ 3.22 ส่วนประกอบของคลาส Anova 73
รูปที่ 3.23 ส่วนประกอบของคลาส Coefficients 77
รูปที่ 3.24 ส่วนประกอบของคลาส ResidualStatistics 81
รูปที่ 3.25 ส่วนประกอบของคลาส LinearRegressionOption 83
รูปที่ 3.26 รายละเอียดและความสัมพันธ์ของคลาสต่างๆสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 86
รูปที่ 3.27 ส่วนประกอบของคลาส StatTable 87
รูปที่ 3.28 รายละเอียดและความสัมพันธ์ของคลาสทั้งหมดในขั้นตอนการวิเคราะห์และออกแบบ 90
รูปที่ 3.29 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม 91
รูปที่ 3.30 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบไม่จัดกลุ่ม 92
รูปที่ 3.31 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม 92
รูปที่ 3.32 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม 93
รูปที่ 3.33 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม 93
รูปที่ 3.34 แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบไม่จัดกลุ่ม 94
รูปที่ 3.35 แผนผังแสดงลำดับของการคำนวณค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น 95
รูปที่ 4.1 แพ็กเกจและประเภทของคลาส DescriptiveStatistics 96
รูปที่ 4.2 คุณสมบัติของคลาส DescriptiveStatistics 96
รูปที่ 4.3 แพ็กเกจและประเภทของคลาส FreqDistribution 97
รูปที่ 4.4 คุณสมบัติของคลาส FreqDistribution 97
รูปที่ 4.5 เมทธอดที่สำคัญๆของคลาส FreqDistribution 97
รูปที่ 4.6 แพ็กเกจและประเภทของคลาส GroupFreqDistribution 98
รูปที่ 4.7 คุณสมบัติของคลาส GroupFreqDistribution 98

สารบัญรูป (ต่อ)

หน้า

รูปที่ 4.8 เมทธอดที่สำคัญๆของคลาส GroupFreqDistribution 98
รูปที่ 4.9 แพ็กเกจและประเภทของคลาส UngroupFreqDistribution 94
รูปที่ 4.10 คุณสมบัติของคลาส UngroupFreqDistribution 99
รูปที่ 4.11 เมทธอดที่สำคัญๆของคลาส UngroupFreqDistribution 99
รูปที่ 4.12 แพ็กเกจและประเภทของคลาส CentralTendency 100
รูปที่ 4.13 คุณสมบัติของคลาส CentralTendency 100
รูปที่ 4.14 เมทธอดที่สำคัญๆของคลาส CentralTendency 100
รูปที่ 4.15 แพ็กเกจและประเภทของคลาส GroupCentralTendency 101
รูปที่ 4.16 เมทธอดที่สำคัญๆของคลาส GroupCentralTendency 103
รูปที่ 4.17 แพ็กเกจและประเภทของคลาส UngroupCentralTendency 103
รูปที่ 4.18 เมทธอดที่สำคัญๆของคลาส UngroupCentralTendency 105
รูปที่ 4.19 แพ็กเกจและประเภทของคลาส Dispersion 105
รูปที่ 4.20 คุณสมบัติของคลาส Dispersion 105
รูปที่ 4.21 เมทธอดที่สำคัญๆของคลาส Dispersion 106
รูปที่ 4.22 แพ็กเกจและประเภทของคลาส GroupDispersion 106
รูปที่ 4.23 เมทธอดที่สำคัญๆของคลาส GroupDispersion 107
รูปที่ 4.24 แพ็กเกจและประเภทของคลาส UngroupDispersion 108
รูปที่ 4.25 คุณสมบัติของคลาส UngroupDispersion 108
รูปที่ 4.26 เมทธอดที่สำคัญๆของคลาส UngroupDispersion 109
รูปที่ 4.27 แพ็กเกจและประเภทของคลาส LinearRegression 109
รูปที่ 4.28 คุณสมบัติของคลาส LinearRegression 110
รูปที่ 4.29 เมทธอดที่สำคัญๆของคลาส LinearRegression 110
รูปที่ 4.30 แพ็กเกจและประเภทของคลาส Enter 110
รูปที่ 4.31 เมทธอดที่สำคัญๆของคลาส Enter 111
รูปที่ 4.32 แพ็กเกจและประเภทของคลาส Forward 111
รูปที่ 4.33 คุณสมบัติของคลาส Forward 111
รูปที่ 4.34 เมทธอดที่สำคัญๆของคลาส Forward 112
รูปที่ 4.35 แพ็กเกจและประเภทของคลาส Backward 112

สารบัญรูป (ต่อ)

หน้า

รูปที่ 4.36 คุณสมบัติของคลาส Backward 113
รูปที่ 4.37 เมทธอดที่สำคัญๆของคลาส Backward 114
รูปที่ 4.38 แพ็กเกจและประเภทของคลาส Stepwise 114
รูปที่ 4.39 คุณสมบัติของคลาส Stepwise 115
รูปที่ 4.40 เมทธอดที่สำคัญๆของคลาส Stepwise 116
รูปที่ 4.41 แพ็กเกจและประเภทของคลาส ModelSummary 116
รูปที่ 4.42 คุณสมบัติของคลาส ModelSummary 117
รูปที่ 4.43 เมทธอดที่สำคัญๆของคลาส ModelSummary 118
รูปที่ 4.44 แพ็กเกจและประเภทของคลาส Correlations 119
รูปที่ 4.45 คุณสมบัติของคลาส Correlations 119
รูปที่ 4.46 เมทธอดที่สำคัญๆของคลาส Correlations 120
รูปที่ 4.47 แพ็กเกจและประเภทของคลาส Anova 120
รูปที่ 4.48 คุณสมบัติของคลาส Anova 121
รูปที่ 4.49 เมทธอดที่สำคัญๆของคลาส Anova 123
รูปที่ 4.50 แพ็กเกจและประเภทของคลาส Coefficients 124
รูปที่ 4.51 คุณสมบัติของคลาส Coefficients 124
รูปที่ 4.52 เมทธอดที่สำคัญๆของคลาส Coefficients 126
รูปที่ 4.53 แพ็กเกจและประเภทของคลาส ResidualsStatistics 127
รูปที่ 4.54 คุณสมบัติของคลาส ResidualsStatistics 127
รูปที่ 4.55 เมทธอดที่สำคัญๆของคลาส ResidualStatistics 130
รูปที่ 4.56 แพ็กเกจและประเภทของคลาส LinearRegressionOption 130
รูปที่ 4.57 คุณสมบัติของคลาส LinearRegressionOption 131
รูปที่ 4.58 เมทธอดที่สำคัญๆของคลาส LinearRegressionOption 131
รูปที่ 4.59 แพ็กเกจและประเภทของคลาส StatTable 132
รูปที่ 4.60 คุณสมบัติของคลาส StatTable 132
รูปที่ 4.61 เมทธอดที่สำคัญๆของคลาส StatTable 136
รูปที่ 6.1 หน้าจอเมนูหลักโปรแกรม Stat 2000 142
รูปที่ 6.2 รายการต่างๆของเมนู File 143

สารบัญรูป (ต่อ)

หน้า

รูปที่ 6.3 รายการต่างๆของเมนู Statistics 144
รูปที่ 6.4 รายการของเมนู Help 144
รูปที่ 6.5 หน้าจอ About 145
รูปที่ 6.6 การเลือกรายการวิเคราะห์สถิติเชิงพรรณนา (Descriptive) 146
รูปที่ 6.7 หน้าจอกำหนดตัวแปรในการวิเคราะห์สถิติเชิงพรรณนา 147
รูปที่ 6.8 หน้าจอสำหรับกำหนดค่าสถิติสำหรับการวิเคราะห์สถิติเชิงพรรณนา 148
รูปที่ 6.9 หน้าจอกำหนดค่าแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา 149
รูปที่ 6.10 แสดงการเลือกตัวแปรที่ต้องการแสดงผลลัพธ์ในรูปแผนภูมิ 150
รูปที่ 6.11 แสดงผลลัพธ์แผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา 150
รูปที่ 6.12 การเลือกรายการการวิเคราะห์ความถดถอยเชิงเส้น 152
รูปที่ 6.13 หน้าจอกำหนดตัวแปรในการวิเคราะห์ความถดถอยเชิงเส้น 152
รูปที่ 6.14 หน้าจอสำหรับกำหนดเลือกค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น 154
รูปที่ 6.15 หน้าจอสำหรับกำหนดค่าสถิติที่ต้องการบันทึกค่า 155
รูปที่ 6.16 หน้าจอกำหนดเงื่อนไขการเลือกตัวแปรอิสระเข้าสมการความถดถอย 156
รูปที่ 6.17 แสดงการเลือกตัวแปรที่ต้องการแสดงผลลัพธ์ในรูปแผนภูมิ 156
รูปที่ 6.18 แสดงผลลัพธ์แผนภูมิสำหรับการวิเคราะห์ความถดถอยเชิงเส้น 157
รูปที่ 7.1 ผลลัพธ์จากการวิเคราะห์สถิติเชิงพรรณนาโดยโปรแกรม SPSSS 158
รูปที่ 7.2 ผลลัพธ์จากการวิเคราะห์สถิติเชิงพรรณนาโดยโปรแกรม Stat2000 159
รูปที่ 7.3 ผลลัพธ์จากการวิเคราะห์ค่าความถี่ของตัวแปร Salary โดยโปรแกรม SPSS 160
รูปที่ 7.4 ผลลัพธ์จากการวิเคราะห์ค่าความถี่ของตัวแปร Salary โดยโปรแกรม Stat2000 160
รูปที่ 7.5 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีเอนเตอร์
โดยโปรแกรม SPSS 161
รูปที่ 7.6 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีเอนเตอร์
โดยโปรแกรม Stat2000 162
รูปที่ 7.7 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณ
โดยโปรแกรม SPSS 162

สารบัญรูป (ต่อ)

หน้า

รูปที่ 7.8 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม Stat2000 163
รูปที่ 7.9 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 163
รูปที่ 7.10 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม Stat2000 163
รูปที่ 7.11 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 164
รูปที่ 7.12 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม Stat2000 164
รูปที่ 7.13 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดย
โปรแกรม SPSS 165
รูปที่ 7.14 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดย
โปรแกรม Stat2000 166
รูปที่ 7.15 ค่าคลาดเคลื่อนคำนวณโดย โปรแกรม SPSS 166
รูปที่ 7.16 ค่าคลาดเคลื่อนคำนวณโดยโปรแกรม Stat2000 167
รูปที่ 7.17 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วย
วิธีฟอร์เวิร์ดโดยโปรแกรม SPSS 168
รูปที่ 7.18 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วย
วิธีฟอร์เวิร์ดโดยโปรแกรม Stat2000 168
รูปที่ 7.19 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 169
รูปที่ 7.20 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 169
รูปที่ 7.21 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 170
รูปที่ 7.22 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม Stat2000 171

สารบัญรูป (ต่อ)

หน้า

รูปที่ 7.23 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม SPSS 172
รูปที่ 7.24 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดย
โปรแกรม Stat2000 172
รูปที่ 7.25 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดย
โปรแกรม SPSS 173
รูปที่ 7.26 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดย
โปรแกรม Stat2000 174
รูปที่ 7.27 ตัวแปรอิสระที่ตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีแบ็คเวิร์ดโดย
โปรแกรม SPSS 175
รูปที่ 7.28 ตัวแปรอิสระที่ตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีแบ็คเวิร์ดโดย
โปรแกรม Stat2000 175
รูปที่ 7.29 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วย
วิธีสเตปไวส์โดยโปรแกรม SPSS 176
รูปที่ 7.30 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วย
วิธีสเตปไวส์โดยโปรแกรม Stat2000 177

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ส่วนประกอบซอฟต์แวร์ (Software Component) [6] คือส่วนของชุดคำสั่งที่กระบวนการในการพัฒนาและทดสอบเป็นอิสระจากระบบที่จะนำชุดคำสั่งเหล่านี้ไปใช้ โปรแกรมเชิงวัตถุได้รับการออกแบบเพื่อรองรับการเขียนโปรแกรมในรูปแบบของส่วนประกอบซอฟต์แวร์ เพื่อให้เกิดการนำส่วนประกอบซอฟต์แวร์ที่ได้พัฒนาและตรวจสอบแล้วกลับมาใช้ในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆ

ปัจจุบันการออกแบบและพัฒนาโปรแกรมเชิงวัตถุได้รับการยอมรับมากขึ้น เนื่องจากนักพัฒนาโปรแกรมสามารถนำส่วนประกอบซอฟต์แวร์เชิงวัตถุที่ได้พัฒนาและทดสอบแล้วกลับมาใช้ร่วมในการพัฒนาโปรแกรมใหม่ๆ ทำให้การพัฒนาโปรแกรมเป็นไปอย่างรวดเร็วยิ่งขึ้น อีกทั้งผู้วิจัยพบว่าไม่มีนักวิจัยอื่นๆที่พัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติ ดังนั้นผู้ทำวิจัยจึงได้ออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติโดยประกอบด้วยส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ความถดถอยเชิงเส้น ทั้งนี้โดยในการพัฒนายึดหลักการพัฒนาเพื่อให้ผู้อื่นสามารถที่จะนำส่วนประกอบซอฟต์แวร์เชิงวัตถุเหล่านี้ไปใช้ร่วมในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆได้

1.2 วัตถุประสงค์

งานวิจัยนี้มีวัตถุประสงค์เพื่อจัดทำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

1.3 ขอบเขตของงานวิจัย

วิทยานิพนธ์มีขอบเขตของงานวิจัยดังนี้

1. ออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติประกอบด้วย
   - สถิติเชิงพรรณนา

     การวิเคราะห์สถิติเชิงพรรณนาเป็นการบรรยายลักษณะของข้อมูลที่ได้เก็บรวบรวม โดยสรุปลักษณะที่สำคัญของข้อมูลของกลุ่มที่ศึกษา ซึ่งประกอบด้วย การวัดค่ากลางของข้อมูลการวัดค่า

การกระจายของข้อมูล การนับค่าความถี่ของข้อมูล รวมทั้งการนำเสนอข้อมูลค่าความถี่ของข้อมูลในรูปของกราฟความถี่

- การวิเคราะห์ความถดถอยเชิงเส้น

  การวิเคราะห์ความถดถอยเชิงเส้นเป็นการศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไปที่มีความสัมพันธ์กันเชิงเส้นโดยมีวัตถุประสงค์ที่จะพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์ การวิเคราะห์ความถดถอยแบ่งเป็น 2 ประเภทคือ การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่ายและการวิเคราะห์ความถดถอยเชิงเส้นเชิงซ้อนหรือเชิงพหุ โดยในการวิเคราะห์มีวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอย 4 วิธี คือ วิธีเอนเตอร์ (Enter) วิธีฟอร์เวิร์ด (Forward) วิธีแบ็คเวิร์ด (Backward) และวิธีสเตปไวส์ (Stepwise)

2. ใช้โปรแกรมภาษาจาวาในการพัฒนาชุดคำสั่ง
3. ใช้ข้อมูลตัวอย่างจากหนังสือทางสถิติจำนวน 5 ชุดขึ้นไปเพื่อทดสอบโปรแกรม
4. โปรแกรมที่ใช้ในการเปรียบเทียบในผลลัพธ์ได้แก่ โปรแกรมเอสพีเอสเอส (SPSS) สำหรับวินโดวส์ เวอร์ชัน8.0
5. ซอฟต์แวร์ที่ใช้ในการพัฒนาระบบ ประกอบด้วย
   - ระบบปฏิบัติการวินโดวส์ 98
   - โปรแกรมจาวาบิวเดอร์ (Java Builder) เวอร์ชัน 3.0 หรือสูงกว่า
   - จาวาดีวีลอปเมนท์คิท (Java Development Kit) เวอร์ชัน 1.2 หรือสูงกว่า

1.4 ขั้นตอนและวิธีดำเนินงาน

ขั้นตอนและวิธีดำเนินงานวิจัยมีดังนี้

1. การเก็บรวบรวมความต้องการ
2. ศึกษาทฤษฎีการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นรวมทั้งทฤษฎีที่เกี่ยวข้อง
3. ศึกษาการใช้งานโปรแกรม SPSS สำหรับวินโดวส์ เพื่อเป็นแบบอย่างในการเปรียบเทียบ
4. ศึกษายูเอ็มแอล
5. วิเคราะห์และออกแบบ
6. ศึกษาแนวคิดและทฤษฎีการวิเคราะห์และออกแบบเชิงวัตถุ
7. วิเคราะห์และออกแบบโปรแกรมตามทฤษฎีการวิเคราะห์และออกแบบโปรแกรมเชิงวัตถุ
8. ศึกษาโปรแกรมภาษาจาวา
9. พัฒนาโปรแกรมตามโครงร่างที่ได้ออกแบบไว้

10. ตรวจสอบความถูกต้องของโปรแกรมโดยเปรียบเทียบค่าทางสถิติกับโปรแกรมเอสพีเอสเอส
11. ตรวจสอบความครบถ้วนของโปรแกรมตามขอบเขตที่กำหนดในยูสเคส
12. สรุปผล จัดทำเอกสารวิทยานิพนธ์

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นไปใช้ร่วมในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆได้
2. ผู้ใช้สามารถนำโปรแกรมไปใช้ในการคำนวนค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นได้
3. ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นที่ได้พัฒนาขึ้นจะเป็นแบบอย่างสำหรับการออกแบบและพัฒนาส่วนประกอบทางซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติอื่นๆ เช่น การประมาณค่าสถิติ การทดสอบสมมติฐาน และการทดสอบสมมติฐานที่ไม่ใช้พารามิเตอร์ เป็นต้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

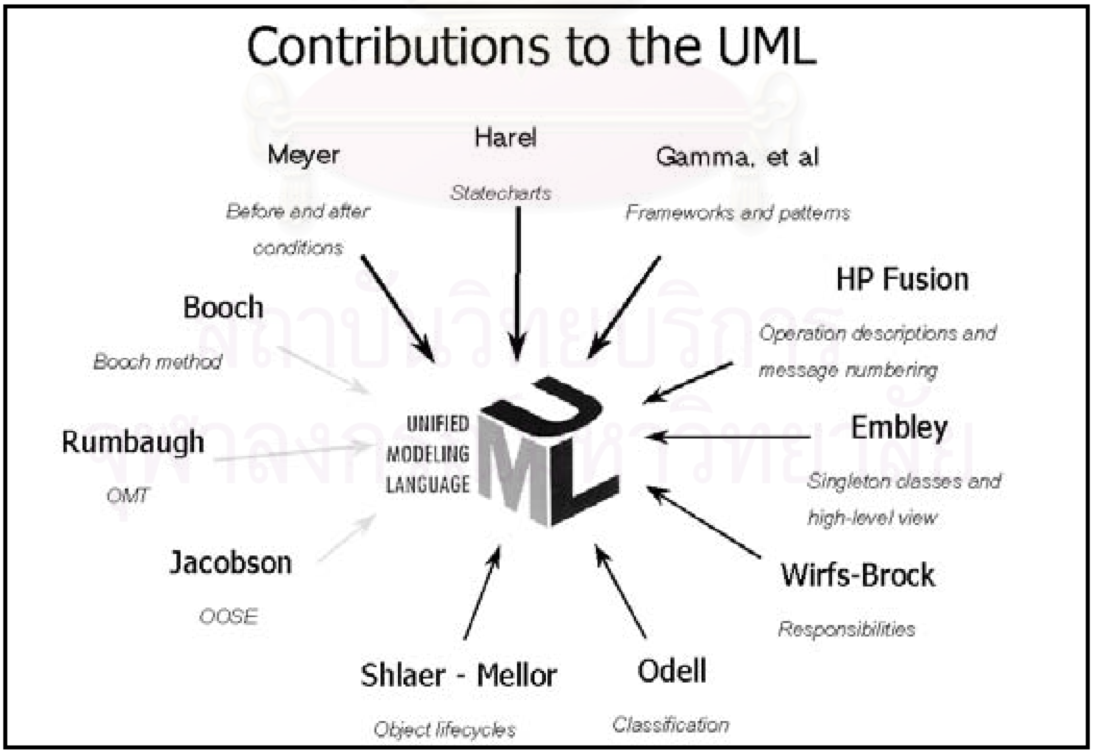
บทที่ 2

แนวคิดและทฤษฎีที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงแนวการวิเคราะห์และออกแบบโปรแกรมตามหลักการวิเคราะห์และออกแบบโปรแกรมเชิงวัตถุ โดยในการวิเคราะห์และออกแบบจะใช้ภาษาเพื่อการวิเคราะห์และออกแบบโปรแกรมเชิงวัตถุที่เรียกว่ายูเอ็มแอล (Unified Modeling Language หรือ UML) ซึ่งเป็นภาษาที่กำหนดขึ้นเป็นมาตรฐานสำหรับการวิเคราะห์และออกแบบโปรแกรมเชิงวัตถุ[4] จากนั้นจะกล่าวถึงรายละเอียดของทฤษฎีทางสถิติเกี่ยวกับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นรวมทั้งสูตรที่ใช้ในการคำนวนค่าสถิติต่างๆ

2.1 ยูเอ็มแอล

ยูเอ็มแอลถูกกำหนดขึ้นเพื่อเป็นมาตรฐานในการวิเคราะห์และออกแบบโปรแกรมเชิงวัตถุ โดย Grady Booch, James Rambaugh และ Ivar Jacobson [3]ได้ร่วมกันกำหนดตัวแบบ (Model)โดยได้เลือกข้อดีจากตัวแบบต่างๆ เช่น Booch, OMT, OOSE, Fusion, Coad และ Yourdon แต่ทั้งนี้ได้ยึดตัวแบบของBooch, OMT-2 และ OOSE เป็นหลัก ดังแสดงในรูปที่ 2.1



รูปที่ 2. 1 ตัวแบบต่างๆที่ยูเอ็มแอลได้นำเอาข้อดีของแต่ละตัวแบบมาประกอบไว้ [5]

ยูเอ็มแอลถูกออกแบบโดยมีวัตถุประสงค์ดังนี้

1. เพื่อการออกแบบระบบโดยใช้ความคิดเชิงวัตถุ
2. เพื่อสร้างความเสมือนระหว่างแนวความคิดกับวัตถุที่ทำงานได้
3. ใช้ได้กับระบบที่มีความซับซ้อน และมีขนาดใหญ่
4. สร้างต้นแบบที่มนุษย์หรือเครื่องคอมพิวเตอร์สามารถเข้าใจได้

กระบวนการในการออกแบบมีความสำคัญมาก โดยเฉพาะอย่างยิ่งสำหรับระบบที่มีขนาดใหญ่ การออกแบบที่ดีจะทำให้สามารถที่จะปรับปรุงหรือเพิ่มขยายระบบได้โดยง่าย ซึ่งตัวแบบที่ดีจะต้องมีคุณสมบัติต่างๆ คือ

1. สามารถอธิบายระบบได้อย่างถูกต้อง
2. ในมุมมอง (View) ต่างๆ จะต้องไม่มีสิ่งที่ขัดแย้งกัน (Consistent)
3. สามารถสื่อสารเชื่อมโยงกับตัวแบบอื่นๆได้
4. สามารถเข้าใจได้ง่าย

ในการพัฒนาโปรแกรมผู้วิจัยจะพัฒนาตามขั้นตอนคือ การเก็บรวบรวมความต้องการ การวิเคราะห์เชิงวัตถุ การออกแบบเชิงวัตถุ การเขียนโปรแกรมและการทดสอบโปรแกรม โดยในแต่ละขั้นตอนจะใช้ยูเอ็มแอล ประกอบดังรายละเอียดต่อไปนี้

1. การเก็บรวบรวมความต้องการ (Requirements Gathering)

ในขั้นตอนนี้เป็นขั้นตอนในการรวบรวมความต้องการของผู้ใช้ โดยการศึกษาจากทฤษฎีทางสถิติเกี่ยวกับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น ในขั้นตอนนี้ผู้วิจัยใช้ยูสเคส (Use Case) เป็นตัวแบบความต้องการของผู้ใช้เท่านั้น และจะยังไม่พิจารณาถึงแนวทางในการสร้างระบบ (System Implementation) ยูสเคสเป็นตัวแบบหนึ่งของยูเอ็มแอลซึ่งแสดงถึงความสัมพันธ์ระหว่างผู้กระทำนอกระบบ (External Actor) ซึ่งอาจจะเป็นผู้ใช้งาน ระบบอื่น หรืออุปกรณ์ต่างๆ กับฟังก์ชันการทำงานของระบบดังแสดงในรูปที่ 2.2

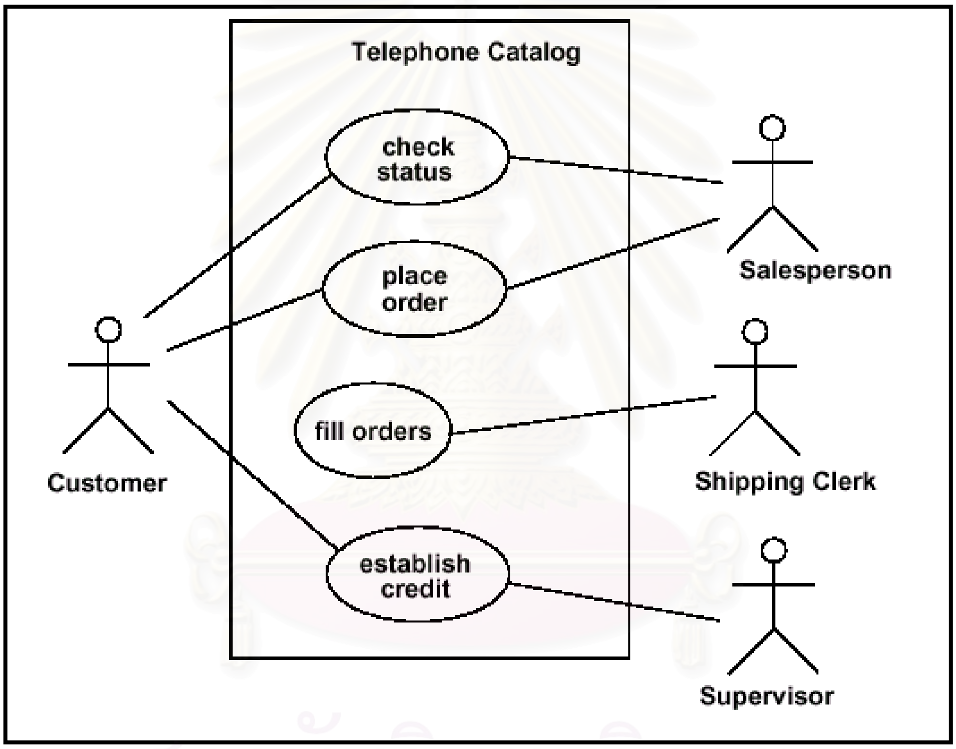
2. การวิเคราะห์เชิงวัตถุ (Object-Oriented Analysis)

ขั้นตอนการวิเคราะห์เป็นขั้นตอนที่พิจารณาเฉพาะขอบเขตของปัญหา (Problem Domain) สำหรับงานวิจัยนี้มีขอบเขตการวิจัยเกี่ยวกับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น โดยการวิเคราะห์จะใช้ผังโครงสร้างคลาส (Class Diagram) มาประกอบการวิเคราะห์ ผังโครงสร้างคลาสเป็นผังที่แสดงรายละเอียดของคลาสซึ่งประกอบด้วย 3 ส่วนคือ ชื่อคลาส คุณสมบัติ

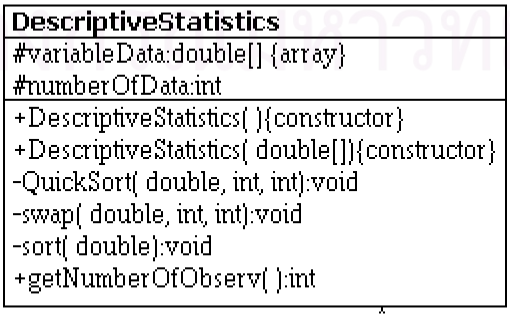
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

และ เมธทอด ดังแสดงในรูปที่ 2.3 สำหรับคุณสมบัติและเมธทอดของคลาส สามารถกำหนดขอบเขตการมองเห็น (Scope of Visibility) ได้เป็น 3 ระดับดังนี้

1. ขอบเขตการมองเห็นระดับพับลิค (Public) สมาชิกคลาสจะถูกอ้างอิงได้จากคลาสใดๆ ที่มองเห็นคลาสนั้นได้ ขอบเขตการมองเห็นระดับพับลิคแทนได้ด้วยเครื่องหมาย '+'
2. ขอบเขตการมองเห็นระดับโพรเทค (Protected) สมาชิกคลาสจะถูกอ้างอิงได้จากคลาสใดๆใน เพคเกจ (Package) เดียวกัน หรือคลาสที่ขยายออกมาจากคลาสนั้น ขอบเขตการมองเห็นระดับโพรเทคแทนได้ด้วยเครื่องหมาย '#'
3. ขอบเขตการมองเห็นระดับไพรเวท (Private) สมาชิกคลาสจะถูกอ้างถึงได้จากภายในคลาสเท่านั้น ขอบเขตการมองเห็นระดับไพรเวทแทนได้ด้วยเครื่องหมาย '-'



รูปที่ 2.2 ตัวอย่างผังยูสเคส [5]



รูปที่ 2.3 ตัวอย่างผังโครงสร้างคลาส

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3. การออกแบบเชิงวัตถุ (Object-Oriented Design)

ขั้นตอนนี้เป็นขั้นตอนที่เพิ่มขยายจากขั้นตอนการวิเคราะห์ โดยมีการเพิ่มเติมคลาสต่างๆทางด้านเทคนิคเช่น คลาสเกี่ยวกับการติดต่อกับผู้ใช้งาน (User Interface) คลาสเกี่ยวกับการรองรับการจัดการข้อมูล (Database Handling) และคลาสเกี่ยวกับการติดต่อเชื่อมโยงกับอุปกรณ์ต่างๆเป็นต้น

4. การเขียนโปรแกรม (Programming)

ขั้นตอนนี้เป็นขั้นตอนในการสร้างระบบ โดยการแปลงต้นแบบที่ได้มาจากการวิเคราะห์และออกแบบมาเป็นโปรแกรมเชิงวัตถุ แต่ทั้งนี้ควรหลีกเลี่ยงการแปลงโดยอ้างอิงจากต้นแบบทั้งหมดที่ได้จากการวิเคราะห์และออกแบบมาเป็นโปรแกรม เนื่องจากในขั้นตอนของการวิเคราะห์และออกแบบ ต้นแบบถูกจัดทำขึ้นเพื่อทำความเข้าใจระบบมากกว่าออกแบบมาเพื่อการแปลงต้นแบบเป็นโปรแกรม

5. การทดสอบ (Testing)

กระบวนการทดสอบเป็นสิ่งจำเป็นในการพัฒนาโปรแกรม การทดสอบสามารถทดสอบโดยอ้างถึงตัวแบบที่ได้วิเคราะห์และออกแบบไว้ดังนี้

- การทดสอบโปรแกรมในระดับคลาสหรือกลุ่มของคลาสโดยเปรียบเทียบกับคลาสที่ได้วิเคราะห์และออกแบบไว้ ว่าโปรแกรมมีคุณสมบัติและเมธทอดครบถ้วนหรือไม่
- การทดสอบการทำงานร่วมกันของคลาสต่างๆว่าเป็นไปตามข้อกำหนด (Specification) ที่ระบุหรือไม่ โดยพิจารณาจากผังโครงสร้างคลาส
- การทดสอบระบบว่ามีฟังก์ชันการทำงานครบถ้วนตามความต้องการของผู้ใช้งานหรือไม่ โดยพิจารณาว่าโปรแกรมมีฟังก์ชันครบถ้วนดังที่กำหนดไว้ในยูสเคสหรือไม่

2.2 ทฤษฎีทางสถิติ

หัวข้อนี้จะกล่าวถึงรายละเอียดของสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น [1, 2] ซึ่งเป็นส่วนที่ได้นำมาจัดทำเป็นส่วนประกอบซอฟต์แวร์เชิงวัตถุดังนี้

2.2.1 สถิติเชิงพรรณนา (Descriptive Statistics)

สถิติเชิงพรรณนา (Descriptive Statistics) ประกอบไปด้วยสถิติเบื้องต้นซึ่งอธิบายลักษณะของข้อมูลตัวอย่างที่ได้รวบรวม และ การนำเสนอข้อมูลในรูปตารางสถิติและในรูปกราฟต่างๆ

2.2.1.1 การนำเสนอข้อมูล

การนำเสนอข้อมูลประกอบด้วย

- การนำเสนอในรูปตาราง
- การนำเสนอในรูปกราฟ เช่น การนำเสนอในรูปกราฟเส้น กราฟแท่ง กราฟวงกลม เป็นต้น

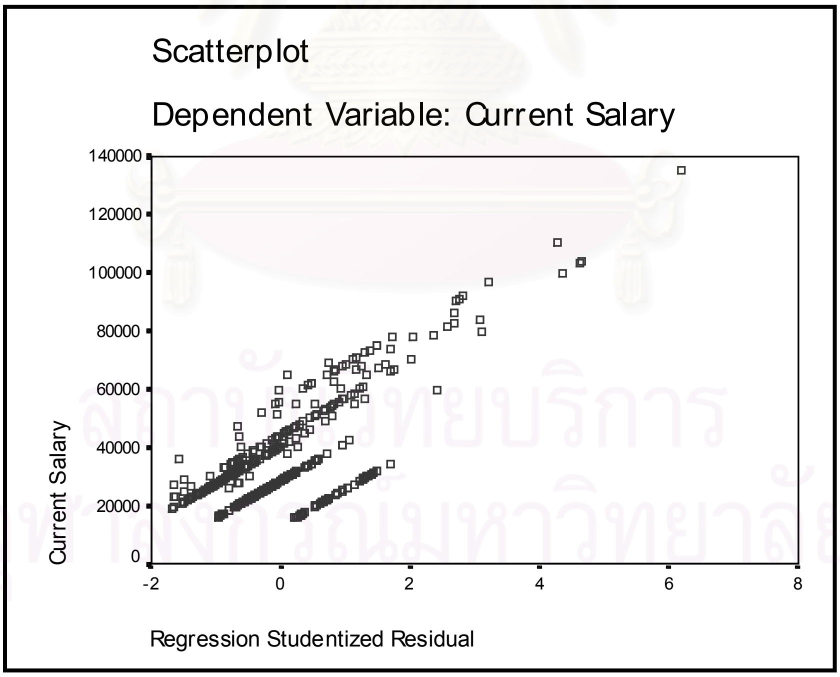
หมายเหตุ วิทยานิพนธ์นี้จะครอบคลุม การนำเสนอข้อมูลใน 3 รูปแบบคือ การนำเสนอข้อมูลในรูปตาราง แผนภาพการกระจาย และ ฮิสโตแกรม โดยมีตัวอย่างดังแสดงในรูปที่ 2.4 รูปที่ 2.5 และรูปที่ 2.6 ตามลำดับ

Descriptives

Descriptive Statistics

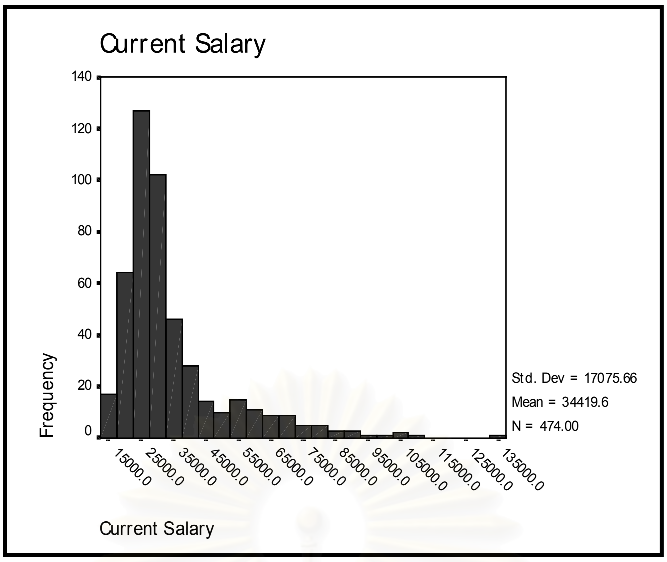
	N	Minimum	Maximum	Mean	Std. Deviation
Current Salary	474	$15,750	$135,000	*********	$17,075.66
Valid N (listwise)	474				

รูปที่ 2.4 การนำเสนอในรูปตาราง



รูปที่ 2.5 การนำเสนอข้อมูลในรูปแผนภาพการกระจาย (Scatter Plot)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 2.6 ตัวอย่างการนำเสนอข้อมูลในรูปฮิสโตแกรม

2.2.1.2 การแจกแจงความถี่

การแจกแจงความถี่มี 2 แบบคือ

ก การแจกแจงความถี่ของลักษณะที่สนใจที่เป็นไปได้ทั้งหมด

การแจกแจงความถี่แบบนี้จะใช้กับข้อมูลที่มีจำนวนลักษณะที่เป็นไปได้ทั้งหมดไม่มากนัก เช่น จำแนกตามเพศ คือ ชาย หญิง หรือจำแนกตามความคิดเห็น เห็นด้วย ไม่เห็นด้วยไม่แสดงความคิดเห็น เป็นต้น ตัวอย่างของตารางแจกแจงความถี่ของลักษณะที่เป็นไปได้ทั้งหมดดังตารางที่ 2.1

ตารางที่ 2.1 การแจกแจงความถี่ของลักษณะที่สนใจที่เป็นไปได้ทั้งหมด

เพศ	จำนวนคนไข้ (คน)
ชาย	214
หญิง	125
รวม	339

ข การแจกแจงความถี่สำหรับค่าในแต่ละช่วงของลักษณะที่สนใจ

การแจกแจงความถี่แบบนี้ใช้ในกรณีที่ค่าของข้อมูลที่เป็นไปได้ทั้งหมดของลักษณะที่สนใจมีเป็นจำนวนมาก เช่น หาจำนวนคนไทยในแต่ละอายุ 0,1,2,.... ปี ซึ่งจะเป็นการเสียเวลาและละเอียดเกินไปสำหรับการวิเคราะห์ในขั้นต่อๆไป เพื่อให้ง่ายจึงแบ่งค่าข้อมูลออกเป็นช่วงๆ ที่ต่อเนื่องกัน โดยในแต่ละ

ช่วงประกอบด้วยข้อมูลหลายๆค่า ทำให้จำนวนค่าที่เป็นไปได้ทั้งหมดลดลง ตัวอย่างของตารางแจกแจงความถี่สำหรับค่าในแต่ละช่วงของลักษณะที่สนใจแสดงดังตารางที่ 2.2

ตารางที่ 2.2 การแจกแจงความถี่สำหรับช่วงค่าของลักษณะที่สนใจ

ความสูงของนิสิตชาย	จำนวนนิสิต
135 – 144	4
145 - 154	19
155 – 164	40
165 – 174	29
175 - 184	8
รวม	100

จากตารางที่ 2.2 ช่วงของแต่ละชั้นเช่น 135-144 เรียกว่าขีดจำกัดชั้น (Class Limit) ค่าต่ำสุดของแต่ละชั้น เช่น 135 เรียกว่า ขีดจำกัดล่าง (Lower Class Limit) และค่าสูงสุดของแต่ละชั้น เช่น 144 เรียกว่า ขีดจำกัดบน (Upper Class Limit) แต่โดยทั่วไปเรามักจะกำหนดค่าต่ำสุดและค่าสูงสุดของแต่ละชั้นให้มีทศนิยมมากกว่าค่าของข้อมูล 1 ตำแหน่งและจะเรียกช่วงในแต่ละชั้นว่าขอบเขตจำกัดชั้น (Class Boundary) และค่าต่ำสุดของแต่ละชั้นคือเช่น 134.5 เรียกว่าขอบเขตจำกัดล่าง (Lower Class Boundary) และค่าสูงสุดของแต่ละชั้นเช่น 144.5 เรียกว่าขอบเขตจำกัดบน (Upper Class Boundary) ดังแสดงในตารางที่ 2.3

ตารางที่ 2.3 ขอบเขตจำกัดชั้นของการแจกแจงความถี่สำหรับช่วงค่าของลักษณะที่สนใจ

ความสูงของนิสิตชาย	จำนวนนิสิต
134.5 – 144.5	4
144.5 – 154.5	19
154.5 – 164.5	40
164.5 – 174.5	29
174.5 – 184.5	8
รวม	100

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<u>ขั้นตอนการสร้างตารางแจกแจงความถี่</u>

การสร้างตารางแจกแจงความถี่มีขั้นตอนดังนี้

1. หาค่าพิสัยของข้อมูล

พิสัย (Range) = ค่าสูงสุด – ค่าต่ำสุด

ถ้ามีข้อมูล N ค่า คือ $X_1, X_2, X_3, \ldots\ldots, X_N$

และให้ X_{max} = ค่าสูงสุดของข้อมูล = Max $(X_1, X_2, X_3, \ldots\ldots, X_N)$

X_{min} = ค่าต่ำสุดของข้อมูล = Min $(X_1, X_2, X_3, \ldots\ldots, X_N)$

จะได้ พิสัย = $R = X_{max} - X_{min}$

2. กำหนดจำนวนชั้น (k) โดยทั่วไปมักจะกำหนดจำนวนชั้นเป็น 5-20 ชั้น หรืออาจใช้สูตรเพื่อหาจำนวนชั้นที่เหมาะสมดังนี้

$k = 1+3.3\log(N)$

3. คำนวณความกว้างของชั้น (Class Interval : I)

I = พิสัย / จำนวนชั้น $= R/k$

ถ้าค่า I เป็นเลขไม่ลงตัวจะปัดให้เป็นเลขจำนวนแต็มไม่ว่าค่าเศษจะเป็นเท่าไร

โดยที่

ค่ากลางชั้น (Mid Point) = (ขอบเขตบน+ขอบเขตล่าง)/2

= (ขีดจำกัดบน+ขีดจำกัดล่าง)/2

4. คำนวณหาขีดจำกัดชั้น (Class Limit)

โดยจะกำหนดให้ขีดจำกัดล่างของชั้นแรก (ชั้นที่มีค่าต่ำสุด) ครอบคลุมข้อมูลที่มีค่าต่ำสุดและให้ขีดจำกัดบนของชั้นสุดท้าย (ชั้นที่มีค่าสูงสุด) ครอบคลุมข้อมูลที่มีค่าสูงสุดหรืออาจใช้สูตรต่อไปนี้

ขีดจำกัดล่างของชั้นแรก = ค่าต่ำสุด – (I x k – R)/2

และปัดเศษให้มีลักษณะเหมือนข้อมูลจริง เช่นข้อมูลมีจำนวนหลักหลังจุดทศนิยมสองหลักจะปัดเศษให้ขีดจำกัดชั้นเป็นเลขทศนิยมสองหลักเช่นกัน

5. คำนวณหาขอบเขตจำกัดชั้น (Class Boundary) การหาขอบเขตชั้นนั้นจะกำหนดให้ขอบเขตชั้นมีจำนวนทศนิยมมากกว่าค่าของข้อมูลจริงอยู่ 1 หลักเสมอ เช่น ถ้าข้อมูลจริงเป็นเลขจำนวนเต็ม ของเขตจำกัดชั้นจะมีจุดทศนิยม 1 หลัก
   ขอบเขตจำกัดชั้น = (ขีดจำกัดบนของชั้น+ขีดจำกัดล่างของชั้นถัดไป)/2
6. นับจำนวนค่าของข้อมูล (ความถี่)ในแต่ละชั้น
   หลังจากสร้างขอบเขตจำกัดชั้นแล้ว จึงตรวจสอบว่าข้อมูลค่าใดอยู่ในชั้นใดบ้างแล้วนับจำนวนข้อมูลในแต่ละชั้น

2.2.1.3 การวัดค่ากลางของข้อมูล

การวัดค่ากลางของข้อมูลคือสถิติที่แสดงค่ากลางของข้อมูลมี 5 แบบคือ ค่าเฉลี่ย มัธยฐาน ฐานนิยม เปอร์เซ็นต์ไทล์ และควอไทล์

2.2.1.3.1 ค่าเฉลี่ย (Mean)

ค่าเฉลี่ยคือสถิติที่แสดงค่ากลางของข้อมูลมี 2 แบบคือ

ก. การหาค่าเฉลี่ยสำหรับข้อมูลที่ไม่ได้จัดกลุ่ม

ค่าเฉลี่ยสำหรับข้อมูลที่ไม่ได้จัดกลุ่มคือค่าสถิติที่แสดงค่ากลางของข้อมูลที่ไม่ได้จัดกลุ่ม มีสูตรในการคำนวณดังนี้

ค่าเฉลี่ยตัวอย่าง $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ โดยที่ n = จำนวนข้อมูลตัวอย่าง

ข. การหาค่าเฉลี่ยสำหรับข้อมูลที่จัดกลุ่ม

ค่าเฉลี่ยสำหรับข้อมูลที่จัดกลุ่มคือค่าสถิติที่แสดงค่ากลางของข้อมูลที่จัดกลุ่มแล้วมีสูตรในการคำนวณดังนี้

- สูตรในการคำนวณค่าเฉลี่ย $\bar{x}$ ของข้อมูลตัวอย่างคือ

$$\bar{x} = \frac{\sum_{i-1}^{k} x_i f_i}{n}$$

โดยที่ k = จำนวนชั้นในตารางแจกแจงความถี่

X_i = ค่ากึ่งกลางชั้นที่ i

f_i = ความถี่ชั้นที่ i

n = จำนวนข้อมูลตัวอย่าง

- สูตรในการคำนวณค่าเฉลี่ยของประชากร μ คือ

$$\mu = \frac{\sum_{i=1}^{k} x_i f_i}{N} \quad \text{โดยที่ } N = \text{จำนวนประชากร}$$

2.2.1.3.2 มัธยฐาน (Median)

ค่ามัธยฐานคือค่าของข้อมูลที่อยู่ตำแหน่งกลางของข้อมูลที่เรียงลำดับจากน้อยไปมากมี 2 แบบคือ

ก มัธยฐานของข้อมูลที่ไม่ได้มีการจัดกลุ่ม

ค่ามัธยฐานของข้อมูลที่ไม่ได้จัดกลุ่มคือคือค่าของข้อมูลที่อยู่ตำแหน่งกลางของข้อมูลที่เรียงลำดับจากน้อยไปมากโดยใช้หลักการคำนวณดังนี้

ถ้า n เป็นเลขคี่ มัธยฐานคือค่าของข้อมูลที่อยู่ตำแหน่งกลางหรือตำแหน่ง (n+1)/2

ถ้า n เป็นเลขคู่ ค่ามัธยฐานคือค่าเฉลี่ยของข้อมูลที่อยู่ตำแหน่งกลาง 2 ตำแหน่ง

ข มัธยฐานของข้อมูลที่จัดกลุ่มแล้ว

ค่ามัธยฐานของข้อมูลที่จัดกลุ่มแล้วคือคือค่าของข้อมูลที่อยู่ตำแหน่งกลางของข้อมูลที่เรียงลำดับจากน้อยไปมากโดยใช้มีสูตรในการคำนวณดังนี้

<u>**สูตร**</u>

$$Median = Med = L + \frac{(\frac{n}{2} - CF)}{f_m} \times I$$

โดยที่	L	=	ค่าขอบเขตจำกัดล่างของชั้นที่มีมัธยฐานอยู่
	CF	=	ความถี่สะสมของชั้นก่อนชั้นที่มีมัธยฐานอยู่
	f_m	=	ความถี่ของชั้นที่มีมัธยฐานอยู่
	I	=	ช่วงกว้างของชั้น

2.2.1.3.3 ฐานนิยม (Mode)

ฐานนิยมคือข้อมูลที่มีความถี่สูงมี 2 แบบคือ

ก. ฐานนิยมของข้อมูลที่ไม่ได้มีการจัดกลุ่ม

ฐานนิยมของข้อมูลที่ไม่ได้จัดกลุ่มคือค่าของข้อมูลที่เกิดขึ้นบ่อยสุดหรือมีความถี่สูงสุด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข. ฐานนิยมของข้อมูลที่จัดกลุ่มแล้ว

ฐานนิยมของข้อมูลที่จัดกลุ่มแล้วคือค่าของข้อมูลที่เกิดขึ้นบ่อยสุดหรือมีความถี่สูงสุดโดยใช้สูตรการคำนวณดังนี้

$$Mode = M_o = L_o + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times I$$

โดยที่ L_0 = ขอบเขตจำกัดล่างของชั้นที่มีความถี่สูงสุด

f_m = ความถี่ของชั้นที่มีฐานนิยมอยู่ (ชั้นที่มีความถี่สูงสุด)

f_1 = ความถี่ของชั้นก่อนชั้นที่มีฐานนิยมอยู่

f_2 = ความถี่ของชั้นหลังชั้นที่มีฐานนิยมอยู่

I = ช่วงกว้างของชั้น

2.2.1.3.4 เปอร์เซ็นต์ไทล์ (Percentile)

เปอร์เซ็นต์ไทล์คือค่าของข้อมูลที่มีข้อมูลอย่างน้อย p% การหาเปอร์เซ็นต์ไทล์มีขั้นตอนดังนี้

1. จัดเรียงข้อมูลจากน้อยไปมาก
2. คำนวณค่า i = (p/100)n

   p = เปอร์เซ็นต์ไทล์ที่ต้องการหา และ n = จำนวนข้อมูล
3. ถ้าค่า i ไม่ใช่จำนวนเต็มให้ปัดขึ้นเป็นเลขจำนวนเต็ม และค่าของเปอร์เซ็นต์ไทล์ที่ p คือ ค่าของข้อมูลที่อยู่ตำแหน่ง i ที่ปัดค่าแล้ว ถ้าค่า i เป็นเลขจำนวนเต็ม เปอร์เซ็นต์ไทล์ที่ p คือ ค่าเฉลี่ยของข้อมูลตำแหน่งที่ i และตำแหน่งที่ i+1

2.2.1.3.5 ควอไทล์ (Quartile)

ควอไทล์เป็นการแบ่งข้อมูลออกเป็น 4 ส่วนส่วนละ 25 เปอร์เซ็นต์

Q_1: ควอไทล์ที่ 1 = เปอร์เซ็นต์ไทล์ที่ 25

Q_2: ควอไทล์ที่ 2 = เปอร์เซ็นต์ไทล์ที่ 50

Q_3: ควอไทล์ที่ 3 = เปอร์เซ็นต์ไทล์ที่ 75

Q_4: ควอไทล์ที่ 4 = เปอร์เซ็นต์ไทล์ที่ 100

2.2.1.4 การวัดการกระจายของข้อมูล

การวัดการกระจายของข้อมูลคือการพิจารณาสรุปลักษณะของข้อมูลโดยพิจารณาถึงการกระจายของข้อมูล ได้แก่ พิสัย พิสัยควอไทล์ ค่าแปรปรวน ค่าเบี่ยงเบนมาตรฐาน สัมประสิทธิ์ความแปรผัน ค่าคลาดเคลื่อนของค่าเฉลี่ย ค่าความเบ้ และ ค่าความโด่ง

2.2.1.4.1 พิสัย (Range)

พิสัยเป็นค่าที่ใช้การวัดการกระจายของข้อมูลโดยใช้ข้อมูลเพียง 2 ค่าคือ ค่าสูงสุดและค่าต่ำสุดของข้อมูล มีสูตรในการคำนวณดังนี้

<u>**สูตร**</u>

พิสัย	=	ค่าสูงสุดของข้อมูล – ค่าต่ำสุดของข้อมูล

2.2.1.4.2 พิสัยควอไทล์ (Interquartile Range)

พิสัยควอไทล์คือค่าแตกต่างระหว่างควอไทล์บน (Q_3) และควอไทล์ล่าง (Q_1)มีสูตรในการคำนวณดังนี้

<u>**สูตร**</u>

$$IQ_R = Q_3 - Q_1$$

2.2.1.4.3 ค่าแปรปรวน (Variance)

ค่าแปรปรวนเป็นค่าที่ใช้ในการวัดการกระจายของข้อมูลโดยพิจารณาจากความแตกต่างระหว่างข้อมูลแต่ละค่ากับค่าเฉลี่ย มีสูตรในการคำนวณดังนี้

<u>**สูตร**</u>

- ค่าแปรปรวนของตัวอย่าง (S^2)

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

- ค่าแปรปรวนของประชากร (σ^2)

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ค่าแปรปรวนจะมีหน่วยเป็นหน่วยของข้อมูลยกกำลังสอบ เช่น ข้อมูลเป็นรายได้หน่วยบาท ค่าแปรปรวนของรายได้จะมีหน่วย (บาท)^2

2.2.1.4.4 ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation)

ค่าเบี่ยงเบนมาตรฐานคือค่ารากที่สองของค่าแปรปรวนมีสูตรในการคำนวณดังนี้

<u>**สูตร**</u>

- ค่าเบี่ยงเบนมาตรฐานของตัวอย่าง (S)

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

- ค่าเบี่ยงเบนมาตรฐานของประชากร (σ)

$$s = \sqrt{s^2}$$

2.2.1.4.5 สัมประสิทธิ์ความแปรผัน (Coefficient of Variation : CV)

สัมประสิทธิ์ความแปรผันเป็นค่าที่ใช้วัดการกระจายของข้อมูลที่ไม่มีหน่วย ซึ่งต่างกับค่าสถิติที่ใช้วัดค่าการกระจายของข้อมูลตัวอื่นๆ ค่าสัมประสิทธิ์ความแปรผัน คือ ค่าเบี่ยงเบนมาตรฐานหารด้วยค่าเฉลี่ย ดังนี้

<u>**สูตร**</u>

C.V. ของตัวอย่าง = $\frac{s}{\bar{x}}$

C.V. ของประชากร = $\frac{\sigma}{\mu}$

ในการเปรียบเทียบการกระจายของข้อมูลที่มีหน่วยต่างกัน จะพิจารณาจาก C.V. ข้อมูลที่มีค่า C.V. มากจะมีการกระจายมากกว่าข้อมูลชุดที่มีค่า C.V. ต่ำ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2.1.4.6 ค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (Standard Error of Mean)

<u>**สูตร**</u>

ค่าคลาดเคลื่อนมาตรฐานของ $\overline{X} = \sigma_{\bar{x}}^2 = \frac{\sum_{i=1}^{k}(\bar{x}_i - \mu)^2}{k}$

หรือ ค่าคลาดเคลื่อนมาตรฐานของ $\overline{X} = S.E.mean = \sigma_{\bar{x}}$

โดยที่ $\sigma_{\bar{x}}$ ใช้วัดการกระจายของ $\overline{X}$

2.2.1.4.7 การวัดความเบ้ (Skewness)

ความเบ้เป็นค่าที่ใช้ในการวัดลักษณะของเส้นโค้งหรือลักษณะของข้อมูลว่าเบ้หรือไม่ สำหรับเส้นโค้งปกติจะมีความเบ้เป็นศูนย์ ความเบ้จะวัดผลต่างระหว่างค่าเฉลี่ยกับค่ามัธยฐาน

<u>**สูตร**</u>

ความเบ้ $= \frac{\bar{x} - Mode}{\sigma}$

หรือ $= \frac{3(\bar{x} - Median)}{\sigma}$

หรือ $= \frac{\sum f_i (x_i - \bar{x})^2 / N}{\sigma^2}$

- ถ้าความเบ้ที่คำนวณได้มีค่าบวก แสดงว่าเบ้ขวา
- ถ้าความเบ้ที่คำนวณได้มีค่าเป็นศูนย์ แสดงว่ากราฟสมมาตรหรือข้อมูลมีการแจกแจงแบบปกติ
- ถ้าความเบ้ที่คำนวณได้มีค่าเป็นลบ แสดงว่าเบ้ซ้าย

2.2.1.4.8 การวัดความโด่ง (Kurtosis)

ความโด่งเป็นค่าที่ใช้วัดความโด่งของเส้นโค้ง

<u>**สูตร**</u>

ความโด่ง $= \frac{\sum f_i (x_i - \bar{x})^4 / N}{\sigma^4}$

- ถ้าค่าความโด่งเป็นศูนย์ หมายถึงข้อมูลมีการแจกแจงแบบปกติ
- ถ้าค่าความโด่งเป็นบวก หมายถึงข้อมูลมีการแจกแจงที่ค่อนข้างป้านหรือโค้งน้อย
- ถ้าค่าความโด่งเป็นลบ หมายถึงข้อมูลมีการแจกแจงที่มียอดแหลมหรือโค้งมาก

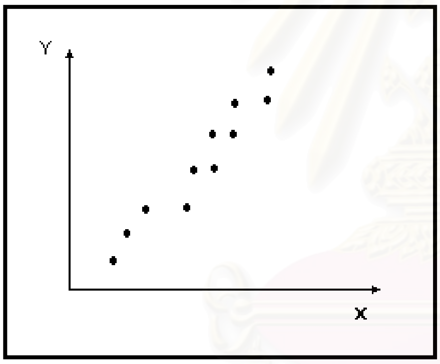
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2.2 การวิเคราะห์ความถดถอยเชิงเส้น (Linear Regression Analysis)

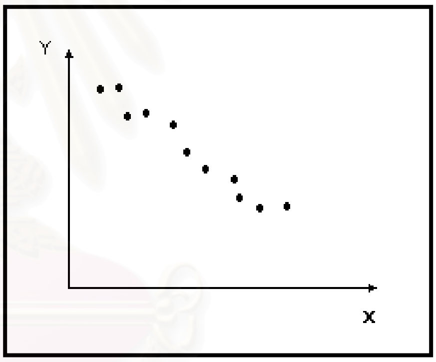
การศึกษาการวิเคราะห์ความถดถอยเป็นการศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์ การวิเคราะห์ความถดถอยแบ่งเป็น 2 ประเภทคือ การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย และการวิเคราะห์ความถดถอยเชิงเส้นเชิงซ้อน หรือเชิงพหุ

2.2.2.1 การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

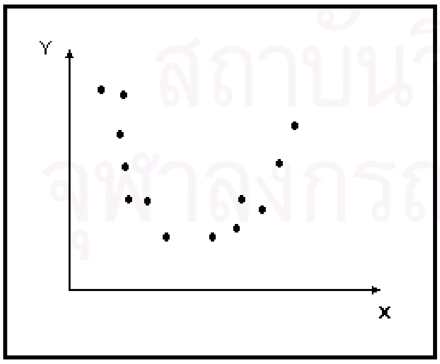
การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่าย เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ (Y) และตัวแปรตาม (X)โดยในการพิจารณาความสัมพันธ์มีวัตถุประสงค์เพื่อพิจารณาว่าตัวแปรทั้งสองมีความสัมพันธ์กันมากเพียงใดและใช้ความสัมพันธ์ที่ได้มาประมาณค่าหรือพยากรณ์ค่าตัวแปรอิสระ สำหรับการหารูปแบบความสัมพันธ์ระหว่างตัวแปร Y และ X สามารถพิจารณาจากแผนภาพการกระจาย (Scatter Plot) ตัวอย่างแผนภาพการกระจายของตัวแปร X และ Y ได้แสดงไว้ดังรูปที่ 2.7



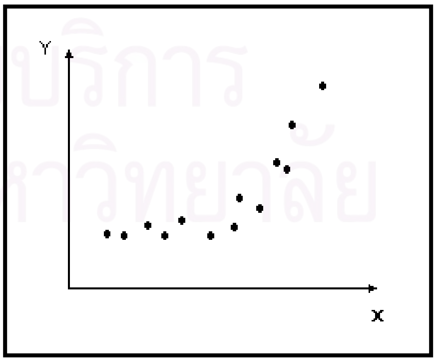
ก. ความสัมพันธ์อยู่ในรูปเส้นตรงและเป็นบวก



ข. ความสัมพันธ์อยู่ในรูปเส้นตรงและเป็นลบ



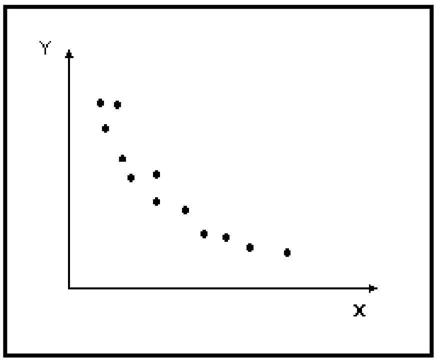
ค. ความสัมพันธ์อยู่ในรูปพาราโบลา



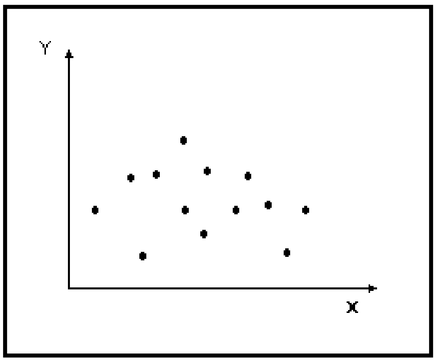
ง. ความสัมพันธ์อยู่ในรูปเอกซ์โพเนนเชียลและเป็นบวก (หรือเป็นเส้นโค้งและเป็นบวก)

รูปที่ 2.7 ตัวอย่างแผนภาพการกระจายซึ่งแสดงความสัมพันธ์ระหว่าง X และ Y

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



จ. ความสัมพันธ์อยู่ในรูปเอกซ์โพเนนเชียล และเป็นลบ



ฉ. ไม่มีความสัมพันธ์กัน

รูปที่ 2.7 ตัวอย่างแผนภาพการกระจายซึ่งแสดงความสัมพันธ์ระหว่าง X และ Y (ต่อ)

การวิเคราะห์ความถดถอยเชิงเส้นเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว ซึ่งสามารถแสดงความสัมพันธ์ในรูปสมการเชิงเส้นดังนี้

$$Y_i \quad = \quad \beta_0+\beta_1 X_i+e_i \quad ; \quad i = 1, 2, \ldots, N$$

Y_i: ตัวแปรตาม (Dependent Variable) เนื่องจากค่าของ Y ขึ้นอยู่กับค่าของ X

X: ตัวแปรอิสระ (Independent Variable)

β_0: ส่วนตัดแกน Y หรือ คือค่าของ Y เมื่อ X มีค่าเป็นศูนย์

e : ความคลาดเคลื่อนอย่างสุ่ม (Random Error)

β_1: ความชัน (Slope) ของเส้นตรง ซึ่งเป็นค่าที่แสดงถึงอัตราการเปลี่ยนแปลงของ Y เมื่อ X เปลี่ยนไป 1 หน่วย และจะเรียก β_1 ว่าสัมประสิทธิ์ความถดถอย (Regression Coefficient) โดยที่ค่าของ β_1 มีความหมายดังนี้

ก. $\beta_1 > 0$ แสดงว่า X และ Y มีความสัมพันธ์ในทางเดียวกัน คือถ้า X มีค่าเพิ่ม Y จะมีค่าเพิ่มขึ้นด้วย

ข. $\beta_1 < 0$ แสดงว่า X และ Y มีความสัมพันธ์ตรงข้ามกัน คือถ้า X มีค่าเพิ่ม Y จะมีค่าลดลง

ค. $\beta_1 = 0$ หรือเข้าใกล้ 0 แสดงว่า X และ Y ไม่มีความสัมพันธ์กันเลยหรือมีความสัมพันธ์กันน้อย

โดยที่หน่วยของ β_0 และ β_1 จะมีหน่วยเหมือนกับหน่วยของ Y

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2.2.1.1 สมมติฐานของการวิเคราะห์ความถดถอย

1. ค่า X จะต้องเป็นค่าที่กำหนดไว้ล่วงหน้าหรือทราบค่า
2. ความคลาดเคลื่อน e_i เป็นตัวแปรที่มี ค่าเฉลี่ย = 0 หรือ $E(e_i) = 0$

   ค่าแปรปรวนของ e_i มีค่าเท่ากันทุกค่าของ i และมีค่าเท่ากับค่าแปรปรวนของ Y

   $V(e_i) \quad = \quad V(Y) \quad = \quad \sigma^2$
3. e_i และ e_j เป็นอิสระกัน นั่นคือ $Cov(e_i, e_j) = E(e_i, e_j) = 0$; i ไม่เท่ากับ j

   e_i มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์และค่าแปรปรวน = σ^2

   และจะได้ว่า $E(Y_i) \quad = \quad \beta_0 + \beta_1 X_I$
4. e_j มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์และค่าแปรปรวน = σ^2 นั่นคือ

   $e_j \quad \sim \quad \text{Normal}(0, \sigma^2)$

   จากข้อสมมติฐานข้างต้น จะได้ว่า

   $Y_i \quad = \quad \text{Normal}(E(Yi), \sigma^2)$

   โดยที่

$$
\begin{aligned}
E(Y_i) &= E(\beta_0 + \beta_1 X_i + e_j) \\
&= \beta_0 + \beta_1 X_i + E(e_j) \\
&= \beta_0 + \beta_1 X_i
\end{aligned}
$$

2.2.2.1.2 การประมาณค่าพารามิเตอร์ของสมการความถดถอย

เมื่อพิจารณาความสัมพันธ์ระหว่างตัวแปร X และ Y จากแผนภาพการกระจายและพบว่าความสัมพันธ์อยู่ในรูปเส้นตรง จะต้องคำนวณหาค่าของ β_0 และ β_1 ซึ่งจะทำให้ทราบความสัมพันธ์ของตัวแปร X และ Y ว่ามีความสัมพันธ์กันมากน้อยเพียงใดและมีความสัมพันธ์ตามกันหรือตรงข้ามกัน

การหาค่า β_0 และ β_1 ได้จะอาศัยการประมาณค่าจากค่า X และ Y ทุกค่าของข้อมูลตัวอย่าง ดังนี้

$Y_i \quad = \quad \beta_0 + \beta_1 X_i + e_j$

หรือ

$$\boxed{\hat{Y}_i = a + bX_i}$$

โดยที่ a และ b เป็นค่าที่ได้จากวิธีกำลังสองน้อยสุด (Least Square Method) ซึ่งเป็นวิธีที่ต้องการหาค่า a และ b ที่ทำให้ผลบวกของค่าคลาดเคลื่อนยกกำลังสองมีค่าน้อยสุดนั่นคือ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลบวกค่าคลาดเคลื่อนกำลังสอง = $\sum_{i=1}^{n} e_i^{\ 2} = \sum (Y_i - \hat{Y}_i)^2$

ในการหาค่า a และ b ที่ทำให้ ให้ผลบวกของค่าคลาดเคลื่อนยกกำลังสองมีค่าน้อยสุดทำได้โดยการใช้ อนุพันธ์เชิงส่วน (Partial Derivative) ดังนี้

$$\frac{\partial}{\partial a}\left[\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right] = \frac{\partial}{\partial a}\left[\sum_{i=1}^{n}(Y_i - a - bX_i)^2\right] = 0$$

หรือ $-2\sum (Y_i - a - bX_i) = 0$

$-2\sum Y_i + 2na + 2b\sum X_i = 0$

$an + b\sum X_i = \sum Y_i$ ……………………..(1.1)

และ

$$\frac{\partial}{\partial a}\left[\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2\right] = -2\sum_{i=1}^{n}(Y_i - a - bX_i)X_i = 0$$

หรือ

$$a\sum_{i=1}^{n}x_i + b\sum_{i=1}^{n}x_i^2 = \sum_{i=1}^{n}x_i y_i \text{ .... } (1.2)$$

จากสมการที่ 1.1 และ 1.2 สามารถหาค่า a และ b ได้ดังนี้

$$a = \bar{y} - b\bar{x}$$

$$b = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

หรือ

$$b = \frac{SS_{xy}}{SS_{xx}}$$

$$a = \bar{y} - b\bar{x}$$

โดยที่

$$SS_{xx} = \sum (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}$$

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n}$$

$$SS_{yy} = \sum (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{(\sum_{i=1}^{n} y_i)^2}{n}$$

การที่ประมาณค่าโดยใช้วิธีกำลังสองน้อยสุดจะทำให้

1. ผลรวมของค่าคลาดเคลื่อนในการประมาณค่า y_i ด้วย $\hat{y}_i$ มีค่าเป็นศูนย์
2. จุด $(\bar{x}, \bar{y})$ เป็นจุดที่อยู่บนเส้นความถดถอย

2.2.2.1.3 การประมาณค่าแปรปรวนของความคลาดเคลื่อน

ในการประมาณค่าแปรปรวนของความคลาดเคลื่อนจะใช้สูตรดังนี้

$$S^2 = SSE / n\text{-}2$$

โดยที่ $$SSE = SS_{yy} - (SS_{xy})^2 / SS_{xx}$$

2.2.2.1.4 การประมาณค่าแบบช่วงและการทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์ความถดถอย

<u>การประมาณค่า β_1 แบบช่วง</u>

สำหรับการประมาณค่าช่วงความเชื่อมั่นของ β_1 ที่ระดับความเชื่อมั่น (1-α) 100% (α คือระดับนัยสำคัญ)โดยใช้สถิติทดสอบ t คือ

$$b - t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{SS_{xx}}} < \beta_1 < b + t_{1-\frac{\alpha}{2}} \frac{S}{\sqrt{SS_{xx}}}$$

โดยที่ t มีองศาอิสระ (Degree of Freedom: df) n-2

<u>การทดสอบสมมติฐานเกี่ยวกับ β_1</u>

การทดสอบสมมติฐานเกี่ยวกับค่า β_1 เป็นการทดสอบว่าตัวแปร X และ Y มีความสัมพันธ์ ในลักษณะเชิงเส้นหรือไม่ โดยเป็นการทดสอบสมมติฐานแบบ 2 ข้าง

จากสมการ

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

ถ้า $\beta_1 = 0$ แสดงว่า X และ Y ไม่มีความสัมพันธ์กันในลักษณะเชิงเส้น

ดังนั้นจึงทดสอบ

H_0 : $\beta_1 = 0$: Y และ X ไม่มีความสัมพันธ์กันในรูปเชิงเส้น

H_1 : $\beta_1 \neq 0$: Y และ X มีความสัมพันธ์กันในรูปเชิงเส้น

<u>สถิติทดสอบ</u>

$$t = \frac{b}{\frac{S}{\sqrt{SS_{xx}}}}$$

<u>เขตปฏิเสธสมมติฐาน H_0</u> คือ

$$t < -t_{1-\alpha/2;n-2} \quad \text{หรือ} \quad t > t_{1-\alpha/2;n-2}$$

ถ้าปฏิเสธ H_0 แสดงว่า X และ Y สัมพันธ์กันในลักษณะเชิงเส้นหรือ การเปลี่ยนแปลงของ X จะมีอิทธิพลต่อการเปลี่ยนแปลงของ Y ในรูปเชิงเส้นนั่นเอง

<u>การประมาณค่า β_0 แบบช่วง</u>

สำหรับการประมาณค่าช่วงความเชื่อมั่นของ β_0 ที่ระดับความเชื่อมั่น $(1-\alpha)100\%$ โดยใช้สถิติทดสอบ t คือ

$$a - t_{1-\alpha/2;n-2} S\sqrt{1/n + \bar{x}^2/SS_{xx}} < a < a + t_{1-\alpha/2;n-2} S\sqrt{1/n + \bar{x}^2/SS_{xx}}$$

<u>การทดสอบสมมติฐานเกี่ยวกับ β_0</u>

<u>สมมติฐาน</u>

H_0 : $\beta_1 = 0$: กราฟเส้นตรงผ่านจุดกำเนิด (0, 0)

H_1 : $\beta_1 \neq 0$: กราฟเส้นตรงไม่ผ่านจุดกำเนิด (0, 0)

<u>สถิติทดสอบ</u>

$$t = \frac{a}{S_a}$$

โดยที่

$$S_a = S\sqrt{1/n + \bar{x}^2/SS_{xx}}$$

<u>เขตปฏิเสธสมมติฐาน</u> จะปฏิเสธ H_0 เมื่อ $|t| > t_{1-\alpha;n-2}$

2.2.2.1.5 การทดสอบสัมประสิทธิ์ความถดถอยโดยใช้การวิเคราะห์ความแปรปรวน

ในการตรวจสอบความสัมพันธ์ของตัวแปร Xและ Y นอกจากจะใช้สถิติ t ในการทดสอบแล้วยังสามารถใช้หลักการของการวิเคราะห์ความแปรปรวนมาทดสอบ โดยพิจารณาว่าค่าของ Y มีการเปลี่ยนแปลงจาก 2 ปัจจัยคือ

1 ความเปลี่ยนแปลงอันเกิดจากค่าของ X ที่เปลี่ยนไป

2 ความเปลี่ยนแปลงเนื่องจากอิทธิพลของปัจจัยอื่นๆที่มีอิทธิพลต่อตัวแปร Y

นั่นคือ ค่าแปรปรวนของตัวแปร Y = SST = SSR + SSE

SST: ค่าแปรปรวนของตัวแปร Y (Sum Square of Total)

SSR: ค่าแปรปรวนของตัวแปร Y อันเนื่องมาจากอิทธิพลของ X

SSE: ค่าแปรปรวนของตัวแปร Y อันเนื่องมาจากอิทธิพลอื่นๆ

โดยที่

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

และ

$$SSE = SST - SSR$$

$$SSR = \frac{(SS_{xy})^2}{SS_{xx}}$$

Mean Square of Regression:	MSR	=	SSR / df	โดยที่ df = 1
หรือ	MSR	=	SSR	
Mean Square of Error:	MSE	=	SSR / df	โดยที่ df = n-2
หรือ	MSE	=	SSE / n-2	

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

<u>สมมติฐาน</u>

H_0 : $\beta_1 = 0$: YและX ไม่มีความสัมพันธ์ในรูปเส้นตรง

H_1 : $\beta_1 \neq 0$: YและX สัมพันธ์ในรูปเส้นตรง

<u>สถิติทดสอบ</u>

F = MSR / MSE

<u>เขตปฏิเสธสมมติฐาน</u>

จะปฏิเสธ H_0 ถ้า $F > F_{1-\alpha; n-2}$

2.2.2.1.6 สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination : r^2 หรือ R^2)

สัมประสิทธิ์การตัดสินใจเป็นสัดส่วนที่ใช้ในการอธิบายการเปลี่ยนแปลงของ Y เมื่อ X เปลี่ยนแปลงค่า นั่นคือถ้า r^2 มีค่ามากแสดงว่า Y และ X มีความสัมพันธ์กันมากโดยที่

$$r^2 = 1 - SSR / SST$$

ดังนั้น $0 \leq r^2 \leq 1$ เนื่องจาก SST > SSR

<u>คุณสมบัติของ r^2</u>

- r^2 ไม่มีหน่วย
- ถ้า r^2 มีค่าเข้าใกล้ 1 แสดงว่า X และ Y มีความสัมพันธ์กันมาก
- ถ้า r^2 มีค่าเข้าใกล้ 0 แสดงว่า X และ Y มีความสัมพันธ์กันน้อย

2.2.2.1.7 สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient)

$$\text{สัมประสิทธิ์สหสัมพันธ์} = \sqrt{r^2}$$

หรือ

$$r = \sqrt{r^2}$$

<u>ความหมายของค่า r</u>

1. ค่า r เป็นลบ แสดงว่า X และ Y มีความสัมพันธ์ในทิศทางตรงข้าม คือ ถ้า Xมีค่าเพิ่ม Y จะมีค่าลดลง
2. ค่า r เป็นบวก แสดงว่า X และ Y มีความสัมพันธ์ในทิศทางเดียวกัน คือ ถ้า Xมีค่าเพิ่ม Y จะมีค่าเพิ่มขึ้นด้วย
3. ค่า r มีค่าเข้าใกล้ 1 หมายถึง X และ Y มีความสัมพันธ์กันมากในทิศทางเดียวกัน

4. ค่า r มีค่าเข้าใกล้ -1 หมายถึง X และ Y มีความสัมพันธ์กันมากในทิศทางตรงข้ามกัน
5. ค่า r มีค่าเข้าใกล้ 0 แสดงว่า X และ Y มีความสัมพันธ์กันน้อย

2.2.2.1.8 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์

การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์ เป็นการทดสอบว่า X และ Y มีความสัมพันธ์กันหรือไม่

สมมติฐาน

H_0 : $\rho = 0$: YและX ไม่มีความสัมพันธ์ในรูปเส้นตรง

H_1 : $\rho \neq 0$: YและX สัมพันธ์ในรูปเส้นตรง

สถิติทดสอบ

$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}}$$

เขตปฏิเสธสมมติฐาน

จะปฏิเสธ H_0 เมื่อ

$$t < -t_{1-\alpha/2;n-2} \quad \text{หรือ} \quad t > t_{1-\alpha/2;n-2}$$

2.2.2.2 การวิเคราะห์ความถดถอยเชิงเส้นเชิงซ้อน (Multiple Linear Regression)

ในหัวข้อที่ 1 ได้กล่าวถึงความสัมพันธ์ของตัวแปร Y กับตัวแปร X เพียงตัวเดียว แต่โดยทั่วๆไป ปัจจัยที่มีอิทธิพลต่อ Y จะมีหลายปัจจัย คือมีตัวแปรอิสระ k ตัว ($X_1, X_2, X_3, \ldots, X_k$) ที่ความสัมพันธ์กับ Y ในรูปเชิงเส้นดังนี้

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + e$$

โดยที่ β_0 = ส่วนตัดแกน Y เมื่อกำหนดให้ $X_1 = X_2 = \ldots = X_k = 0$

และ $\beta_0, \beta_1, \beta_2, \ldots \beta_k$ เป็นสัมประสิทธิ์ความถดถอยเชิงส่วน (Partial Regression Coefficient)

โดยที่ β_i เป็นค่าคงที่แสดงถึงการเปลี่ยนแปลงของตัวแปรตาม Y เมื่อตัวแปรอิสระ X_i เปลี่ยนไป 1 หน่วย โดยที่ตัวแปรอิสระ X ตัวอื่นๆมีค่าคงที่

2.2.2.2.1 สมมติฐานของการวิเคราะห์ความถดถอยเชิงซ้อน

สมมติฐานของการวิเคราะห์ความถดถอยเชิงซ้อนจะเหมือนกับสมมติฐานของการวิเคราะห์ความถดถอยเชิงเส้นอย่างง่ายดังนี้

1. ความคลาดเคลื่อน e เป็นตัวแปรที่มีการแจกแจงแบบปกติ
2. ค่าเฉลี่ยของความคลาดเคลื่อนเป็นศูนย์ นั่นคือ E (e) = 0
3. ค่าแปรปรวนของความคลาดเคลื่อนเป็นค่าคงที่ไม่ทราบค่า V (e) = β^2_e
4. e_i และ e_j เป็นอิสระต่อกัน ;$i \neq j$ นั่นคือ covariance(e_i, e_j) = 0

2.2.2.2.2 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงซ้อน

ในกรณีที่ k มากกว่า 2 การคำนวณหาค่าพารามิเตอร์ของสมการความถดถอยสามารถทำได้โดยการคำนวณโดยใช้เมตริกซ์

โดยที่เมตริกซ์

$$X = \begin{bmatrix} 1 & X_{11} & \dots & X_{k1} \\ 1 & X_{12} & \dots & X_{k2} \\ \dots & \dots & \dots & \dots \\ 1 & X_{k2} & \dots & X_{kn} \end{bmatrix} \qquad b = \begin{bmatrix} a \\ b_1 \\ \dots \\ b_k \end{bmatrix}$$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_k \end{bmatrix}$$

และสามารถคำนวณหาค่าพารามิเตอร์ของสมการความถดถอยจากสูตร

$$\boxed{b = (X'X)^{-1}X'Y}$$

2.2.2.2.3 การทดสอบสมการความถดถอยเชิงเส้นเชิงซ้อนโดยใช้การวิเคราะห์ความแปรปรวน

จากสมการความถดถอยเชิงซ้อน

Y = $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + e$

ค่าแปรปรวนของ Y = ค่าแปรปรวนที่เกิดจากอิทธิพลของ $X_1, X_2, X_3, \ldots X_k$ + ค่าแปรปรวนอย่างสุ่ม

หรือ SST = SSR + SSE

โดยที่ SST (Sum Square of Total) คือค่าแปรปรวนทั้งหมดของ Y

SSR (Sum Square of Regression) คือค่าแปรปรวนของ Y เนื่องจากอิทธิพลของ $X_1, X_2, \ldots X_k$

SSE (Sum Square of Error) คือค่าแปรปรวนของ Y เนื่องจากอิทธิพลอื่นๆ หรือเรียกว่าค่าแปรปรวนแบบสุ่ม

โดยที่

$$SST = \sum_{i=1}^{n} (y_i - \bar{y})^2 = y'y - n\bar{y}^2$$

$$SSR = b'X'Y - n\bar{Y}^2$$

$$SSE = \sum_{i=1}^{n} \left[Y_i - (a + b_1 X_1 + b_2 X_2 + \ldots\ldots + b_k X_k)\right]^2$$

หรือ SSE = SST – SSR

$$= Y'Y - b'X'Y$$

<u>สมมติฐาน</u>

H_0 : $\beta_0 = \beta_1 = \beta_2 = \ldots\ldots = \beta_k = 0$ นั่นคือ Y และX ไม่มีความสัมพันธ์ในรูปเส้นตรง

H_1 : มี β_i อย่างน้อย 1 ค่าที่ $\neq 0$ นั่นคือ YและX สัมพันธ์ในรูปเส้นตรง

<u>สถิติทดสอบ</u>

$$F = \frac{MSR}{MSE} = \frac{(b'X'Y - n\bar{Y}^2)/k}{(Y'Y - b'X'Y)/(n-k-1)}$$

<u>เขตปฏิเสธสมมติฐาน</u> จะปฏิเสธสมมติฐาน H_0 เมื่อ $F > F_{k,\, n-k-1;\, 1-\alpha}$

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผลการทดสอบสมมติฐานอาจจะเป็น

ก. ยอมรับสมมติฐาน H_0 จะสรุปได้ว่า Y ไม่มีความสัมพันธ์กับ X ทั้ง k ตัวในรูปเชิงเส้น

ข. ปฏิเสธสมมุติฐาน H_0 หรือยอมรับสมมติฐาน H_1 ซึ่งสรุปได้ว่ามี X อย่างน้อย 1 ตัวมีความสัมพันธ์กับ Y ในรูปเชิงเส้น และจะต้องทำการทดสอบต่อไปว่า X ตัวใดบ้างที่มีความสัมพันธ์กับ Y ในรูปเช้งเส้นซึ่งจะกล่าวถึงในหัวข้อต่อไป

2.2.2.2.4 การประมาณค่าคลาดเคลื่อนของความถดถอย

ความคลาดเคลื่อนหรือค่าเบี่ยงเบนมาตรฐานของค่าประมาณ คือ

โดยที่

$$S = \sqrt{S^2} = \sqrt{SSE/(n-k-1)} = \sqrt{MSE}$$

$$S^2 = \frac{SSE}{n-k-1} = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-k-1}$$

การประมาณค่าเบี่ยงเบนมาตรฐานของ b_i เราจะต้องทราบค่าเบี่ยงเบนมาตรฐานของ b และจะได้ว่า

$$S^2(b) = (X'X)^{-1}S^2$$

โดยที่

$$S_b^2 = \begin{bmatrix} S_a^2 & & & \\ & S_{b_1}^2 & & \\ & & \cdots & \\ & & & S_{b_k}^2 \end{bmatrix}$$

และ

$$S_{b_i} = \sqrt{S_{b_i}^2}$$

2.2.2.2.5 การประมาณค่าสัมประสิทธิ์ความถดถอยแบบช่วง

ค่าประมาณแบบช่วงของ β_i ที่ระดับความเชื่อมั่น (1-α)100% คือ

$$b_i \pm t_{1-\alpha/2;n-k-1}S_{bi}$$

2.2.2.2.6 การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์ความถดถอย

ถ้ามีตัวแปรอิสระ k ตัว (X_1, X_2, X_3, … X_k) ที่มีความสัมพันธ์กับตัวแปรตาม Y และเมื่อได้สถิติทดสอบ F และยอมรับสมมติฐาน H_0 จะต้องทำการทดสอบต่อว่ามี β_i ตัวใดบ้างที่มีความสัมพันธ์กับ Y โดยการทดสอบสมมติฐานดังต่อไปนี้

<u>สมมติฐาน</u> H_0 : β_i = 0: i = 1, 2, 3,…, k

H_1 : $\beta_i \neq 0$

<u>สถิติทดสอบ</u>

$$t = \frac{b_i}{S_{b_i}}$$

<u>เขตปฏิเสธสมมติฐาน H_0</u>

$$t < -t_{1-\alpha/2;n-k-1} \quad \text{หรือ} \quad t > t_{1-\alpha/2;n-k-1}$$

2.2.2.2.7 สัมประสิทธิ์การตัดสินใจเชิงซ้อน (Multiple Coefficient of Determination : R^2 หรือ r^2)

สัมประสิทธิ์การตัดสินใจเชิงซ้อนเป็นสัดส่วนหรือเปอร์เซ็นต์ที่ตัวแปรอิสระ (X_1, X_2, X_3, … X_k) สามารถอธิบายการเปลี่ยนแปลงของ Y ได้โดยมีรูปแบบสมการดังนี้

r^2 = ความผันแปรของ Y เนื่องจากอิทธิพลของ X_1, X_2, X_3, … X_k ความผันแปรทั้งหมด

= SSR/SST

หรือ

r^2 = (SST – SSE)/SST

โดยที่ $0 \leq r^2 \leq 1$

ถ้าค่า r^2 ใกล้ 1 จะหมายถึง X_1, X_2, X_3, … X_k มีความสัมพันธ์กับ Y มาก

แต่ถ้าค่า r^2 ใกล้ 0 จะหมายถึง X_1, X_2, X_3, … X_k มีความสัมพันธ์กับ Y น้อย

เนื่องจาก SSR จะมีค่าเพิ่มขึ้นถ้าเพิ่มตัวแปรอิสระ ดังนั้นเมื่อเพิ่มตัวแปรอิสระเข้าสมการความถดถอยจะทำให้ค่า r^2 มากขึ้นทั้งๆที่ตัวแปรอิสระ X ที่เพิ่มขึ้นอาจไม่มีความสัมพันธ์กับ Y เลย ทั้งนี้จึงมีการปรับค่า r^2 ให้ถูกต้องขึ้น เรียกว่า Adjust r^2 โดยที่

$$AdjustR^2 = R_a^2 = 1 - \frac{SSE/(n-k-1)}{SST/(n-1)}$$

2.2.2.2.8 การตรวจสอบสมมติฐานของการวิเคราะห์ความถดถอย

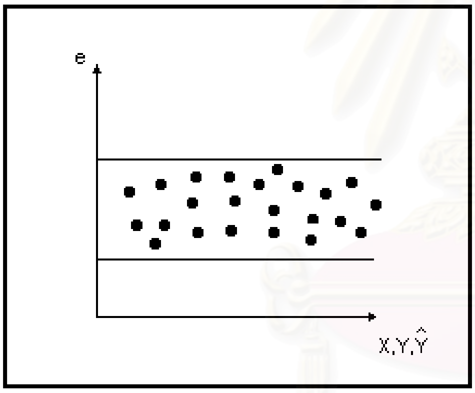
การที่พยากรณ์ค่า Y โดยใช้สมการความถดถอย โดยกำหนดค่า X ได้นั้น ค่าคลาดเคลื่อนจะต้องมีคุณสมบัติตามสมมติฐาน ถ้าไม่เป็นไปตามสมมติฐานที่กล่าวไว้ก็ไม่สามารถนำสมการความถดถอยไปพยากรณ์ Y ได้ ดังนั้นก่อนที่จะประมาณหรือพยากรณ์ค่า Y จะต้องตรวจสอบคุณสมบัติทั้ง 4 ข้อดังนี้

<u>1. การตรวจสอบว่า E (e) = 0</u>

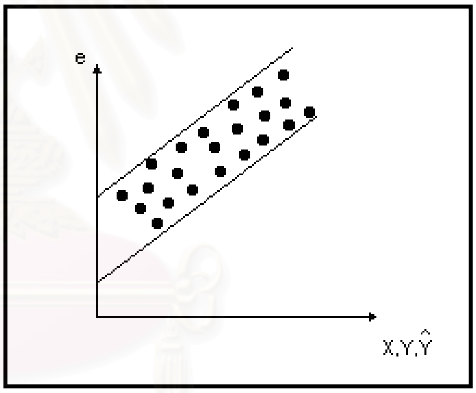
เนื่องจากเราใช้วิธีกำลังสองน้อยสุดในการประมาณ β_0 ด้วย a และ β_i ด้วย b_i ซึ่งทำให้ $\sum e_i = 0$ ดังนั้น E(e) = 0

<u>2. การตรวจสอบ V (e) = σ^2</u>

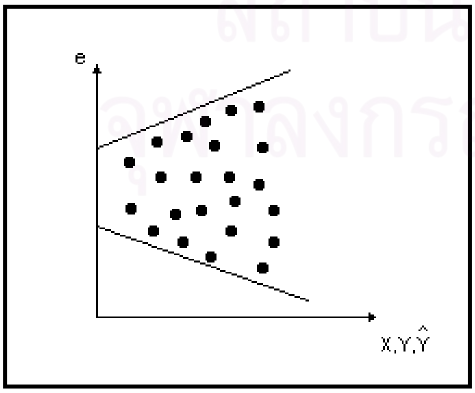
การตรวจสอบว่า V (e) = V (Y) = σ^2 = ค่าคงที่ สามารถทำได้โดยการเขียนกราฟแสดงความสัมพันธ์ระหว่าง e กับ X, e กับ Y หรือ e กับ $\hat{Y}$ ถ้า V (e) ไม่เท่ากับค่าคงที่จะเรียกว่าเกิดปัญหา Heteroscedastic ดังแสดงในรูปที่ 2.8.ค และ 2.8.ง แต่ถ้า V (e) = ค่าคงที่ตามสมมติฐานที่ตั้งไว้ จะเรียกว่า Homoscedastic ดังแสดงในรูปที่ 2.8.ก และ 2.8.ข



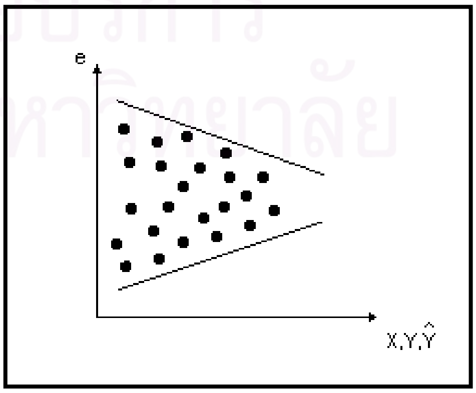
รูปที่ 2.8.ก



รูปที่ 2.8.ข



รูปที่ 2.8.ค



รูปที่ 2.8.ง

รูปที่ 2.8 กราฟความสัมพันธ์ระหว่าง e กับ X, e กับ Y หรือ e กับ $\hat{Y}$

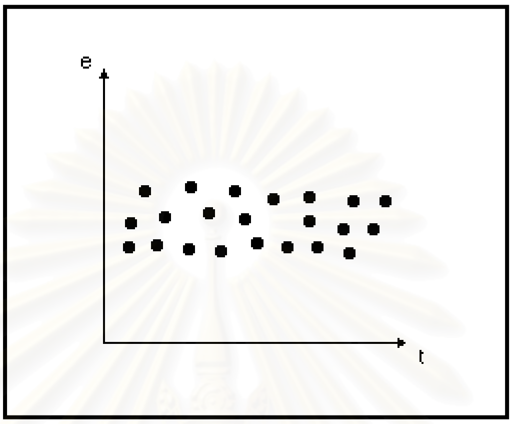
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากรูปที่ 2.8.ง. ค่า σ^2 จะมีค่าน้อยเมื่อ $\hat{Y}$ มีค่ามากขึ้น ในขณะที่รูป 2.8.ก. และ 2.8.ข ค่า σ^2 จะมีค่าคงที่เมื่อ $\hat{Y}$ มีค่ามากขึ้น ส่วนรูป 2.8.ค. ค่า σ^2 จะมีค่ามากเมื่อ $\hat{Y}$ มีค่ามากขึ้น

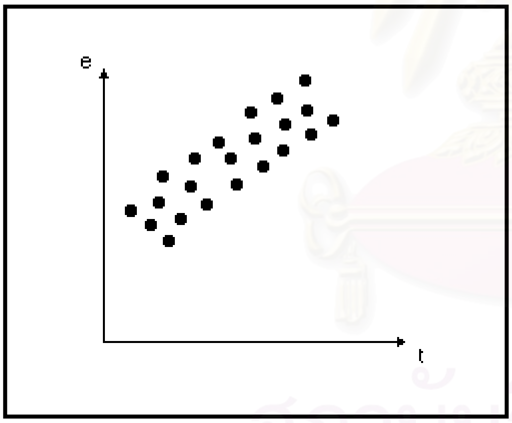
<u>3. การตรวจสอบ e_i และ e_j ว่าเป็นอิสระต่อกันหรือไม่</u>

การทดสอบความเป็นอิสระกันของค่าคลาดเคลื่อน e_i และ e_j ทำได้ 2 วิธีคือ

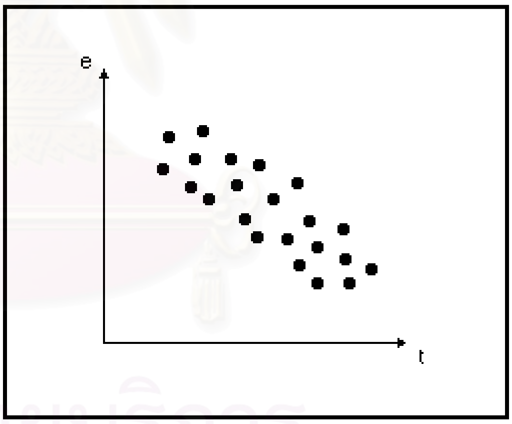
ก. โดยเขียนกราฟความสัมพันธ์ระหว่าง e_t กับ t



รูปที่ 2.9.ก e_i และ e_j เป็นอิสระกันหรือเรียกว่า No Autocorrelation



รูปที่ 2.9.ข Positive Autocorrelation



รูปที่ 2.9.ค Negative Autocorrelation

รูปที่ 2.7 กราฟความสัมพันธ์ระหว่า ง e กับ t

รูปที่ 2.9.ข แสดงว่า e_i และ e_j มีความสัมพันธ์กันในทิศทางเดียวกันส่วนรูปที่ 2.9.ค e_i และ e_j สัมพันธ์ในทิศตรงข้ามกัน

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ข.ใช้สถิติทดสอบ เดอร์บิน-วัทสัน (Durbin-Watson)

สถิติทดสอบเดอร์บิน-วัทสัน = d = $\dfrac{\sum_{2}^{n}(e_t - e_{t-1})^2}{\sum_{1}^{n} e_t^2}$

โดยที่ $0 \leq d \leq 4$

การทดสอบ e_i และ e_j เป็นอิสระต่อกันหรือไม่ โดยใช้สถิติทดสอบ Durbin-Watson จะพิจารณาจากค่าสถิติทดสอบ เดอร์บิน-วัทสัน ดังนี้

- ถ้าค่าสถิติทดสอบเดอร์บิน-วัทสัน มีค่าเข้าใกล้ 2 จะสรุปว่า e_i และ e_j อิสระต่อกัน
- ถ้าค่าสถิติทดสอบเดอร์บิน-วัทสัน < 2 แสดงว่าความสัมพันธ์ของ e_i และ e_j อยู่ในทิศทางบวก และถ้าสถิติทดสอบเดอร์บิน-วัทสัน มีค่าเข้าใกล้ศูนย์แสดงว่า e_i และ e_j มีความสัมพันธ์กันมาก
- ถ้าค่าสถิติทดสอบ เดอร์บิน-วัทสัน > 2 แสดงว่าความสัมพันธ์ของ e_i และ e_j อยู่ในทิศทางลบ และถ้า Durbin-Watson มีค่าเข้าใกล้ 4 แสดงว่า e_i และ e_j มีความสัมพันธ์กันมาก

หรือจะพิจารณาจากค่านัยสำคัญของสถิติทดสอบสถิติทดสอบ เดอร์บิน-วัทสัน ถ้าค่าน้อยกว่าระดับนัยสำคัญที่กำหนด จะสรุปได้ว่า e_i และ e_j มีความสัมพันธ์กัน

4.การทดสอบว่า e_t มีการแจกแจงแบบปกติหรือไม่ ทำได้ 2 วิธีดังนี้

1. ใช้ การทดสอบไคสแควร์ (Chi-Square Test)
2. เขียนฮิสโตแกรมแสดงความสัมพันธ์ระหว่าง e_t กับความถี่ที่เกิด

2.2.2.2.9 การเกิดปัญหามัลติคอลลิเนียริตี (Multicollinearity)

การวิเคราะห์ความถดถอยเชิงซ้อนมีข้อกำหนดว่าตัวแปรอิสระ (X_1, X_2, X_3, … X_k; k>1) จะต้องไม่มีความสัมพันธ์กัน การที่ตัวแปรอิสระมีความสัมพันธ์กันเองมากจะทำให้เกิดปัญหามัลติคอลลิเนียริตี ซึ่งจะทำให้ผลลัพธ์ของการวิเคราะห์ความถดถอยผิดปกติไปดังนี้

1. ทำให้ค่าเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอยมีค่าสูงมาก ซึ่งทำให้สถิติทดสอบ t และ F ได้ข้อสรุปที่ขัดแย้งกัน
2. ทำให้เครื่องหมายของสัมประสิทธิ์ความถดถอยตรงข้ามกับที่ควรจะเป็น
3. ทำให้ค่าสัมประสิทธิ์ความถดถอยเปลี่ยนแปลงไป (ไม่คงที่) เมื่อมีตัวแปรอิสระเพิ่มขึ้น

เมื่อเกิดปัญหามัลติคอลลิเนียริตี จะทำให้สมการความถดถอยไม่ถูกต้อง ดังนั้นผู้วิเคราะห์จึงต้องหาทางป้องกันการเกิดปัญหามัลติคอลลิเนียริตี ดังนี้

1. คำนวณหาสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปร X ต่างๆ และทำการทดสอบสมมติฐานว่าสัมประสิทธิ์สหสัมพันธ์ของ X แต่ละคู่เป็นศูนย์หรือไม่ กรณีที่มี X บางคู่ที่มีความสัมพันธ์กันจะต้องตัดตัวใดตัวหนึ่งออกจากสมการความถดถอย
2. ใช้วิธีสเตปไวส์ซึ่งเป็นวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยจะนำตัวแปรอิสระเข้าสมการทีละ 1 ตัวถ้าตัวแปรที่นำเข้ามีความสัมพันธ์กับตัวแปรอิสระที่มีอยู่แล้วในสมการถดถอย ก็จะตัดตัวแปรตัวใดตัวหนึ่งออกจากสมการความถดถอย โดยจะได้กล่าววิธีการโดยละเอียดต่อไป

2.2.2.2.10 การเลือกตัวแปรอิสระเข้าสมการความถดถอย

การเลือกตัวแปรอิสระ X เข้าสมการความถดถอย ผู้วิเคราะห์จะต้องพิจารณาว่าตัวแปรอิสระใดบ้างที่น่าจะมีความสัมพันธ์กับ Y สำหรับเทคนิคที่ใช้ในการพิจารณาว่าตัวแปรอิสระใดบ้างที่มีความสัมพันธ์กับ Y มี 4 วิธีดังนี้

1. วิธีเอนเตอร์ (Enter หรือ All Possible Regression)
2. วิธีแบ็คเวิร์ด (Backward Elimination)
3. วิธีฟอร์เวิร์ด (Forward Selection)
4. วิธีสเตปไวส์ (Stepwise Regression)

ไม่ว่าผู้วิเคราะห์จะใช้วิธีการใด ผู้วิเคราะห์จะต้องใช้การทดสอบสมมติฐาน โดยใช้สถิติทดสอบ F และ t

1) วิธีเอนเตอร์

วิธีเอนเตอร์เป็นวิธีที่เลือกตัวแปรอิสระเข้าสมการความถดถอยโดยพิจารณาจากสมการความถดถอยที่เป็นไปได้ทั้งหมด เช่น ถ้ามีตัวแปรอิสระ 3 ตัว คือ X_1, X_2 และ X_3 สมการความถดถอยที่จะเป็นไปได้จะมี 8 สมการดังนี้

1. ไม่มีตัวแปรอิสระใดเลยที่มีความสัมพันธ์กับ
2. มีตัวแปร X1 เพียงตัวเดียวที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_1 X_1 + e$
3. มีตัวแปร X2 เพียงตัวเดียวที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_1 X_2 + e$
4. มีตัวแปร X3 เพียงตัวเดียวที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_1 X_3 + e$
5. มีตัวแปร X1, X2 ที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e$
6. มีตัวแปร X1,X3 ที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + e$
7. มีตัวแปร X2, X3 ที่สัมพันธ์กับ Y: $Y = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + e$
8. ตัวแปรทั้ง 3ที่สัมพันธ์กับ Y : $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

การพิจารณาว่าควรเลือกสมการความถดถอยใดจาก 8 สมการข้างต้น จะพิจารณาจาก

1. สถิติทดสอบ F
2. สถิติทดสอบ t
3. สัมประสิทธิ์การตัดสินใจ r^2
4. สัมประสิทธิ์สหสัมพันธ์ r
5. ความคลาดเคลื่อนมาตรฐาน $S_{y.x}$

นั่นคือจากสมการความถดถอยที่ผลการทดสอบทั้งจากสถิติทดสอบ t และ F เป็นปฏิเสธสมมุติฐาน H_0จะเลือกสมการความถดถอยที่มีค่า r_2 และ $S_{y.x}$ มีค่าน้อย

2) วิธีแบ็คเวิร์ด

วิธีแบ็คเวิร์ดเป็นวิธีเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยให้สมการความถดถอยประกอบด้วยตัวแปรอิสระทั้งหมด (k ตัว) ที่คาดว่าจะมีความสัมพันธ์กับตัวแปรตาม Y แล้วใช้สถิติทดสอบ t ทดสอบความสัมพันธ์ระหว่าง Y กับ X ทีละตัว

$H_0 \quad : \quad \beta_i \quad = \quad 0:$ YและX ไม่มีความสัมพันธ์ในรูปเส้นตรง

$H_1 \quad : \quad \beta_i \quad \neq \quad 0:$ YและX สัมพันธ์ในรูปเส้นตรง

และจะปฏิเสธสมมุติฐานเมื่อ $|t| > t_{1-\alpha/2;\, n-k-1}$

โดยจะมีการพิจารณา X ที่มีค่าสถิติ | t | มากที่สุดออกไปจากสมการความถดถอยทีละตัว จนกระทั่งไม่มีตัวแปรอิสระตัวใดที่ไม่มีความสัมพันธ์กับ Y ก็จะได้สมการความถดถอยที่เหมาะสม เนื่องจากตัวแปรอิสระทุกตัวมีความสัมพันธ์กับตัวแปรตาม Y อย่างมีนัยสำคัญ

3) วิธีฟอร์เวิร์ด

วิธีฟอร์เวิร์ดเป็นวิธีที่มีหลักเกณฑ์ตรงข้ามกับวิธีแบ็คเวิร์ด นั่นคือจะเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัแปรตามเข้าสมการทีละตัว โดยจะเลือกตัวแปรที่มีความสัมพันธ์กับตัวแปรตามมากที่สุดเข้าสมการเป็นลำดับแรก การเลือกตัวแปรอิสระตัวแรกเข้าสมการจะพิจารณาสมการความถดถอยอย่างง่าย นั่นคือให้สร้างสมการความถดถอยระหว่าง X และตัวแปรอิสระทีละตัว

$$Y = \beta_0 + \beta_i X_i + e; \quad i = 1, 2, 3, \ldots, k$$

แล้วเลือกตัวแปรอิสระ X ที่มีความสัมพันธ์กับ Y มากที่สุด โดยพิจารณาจากสถิติทดสอบ t หรือ F ในการทดสอบ(t และ F จะให้ผลการทดสอบเหมือนกันสำหรับสมการความถดถอยอย่างง่าย)และทำการเลือกตัวแปรอื่นๆเข้าสมการโดยใช้สมมติฐาน

H_0 : β_i = 0: YและX ไม่มีความสัมพันธ์ในรูปเส้นตรง

H_1 : β_i $\neq$ 0: YและX สัมพันธ์ในรูปเส้นตรง

<u>ผลการทดสอบ</u>

1. ยอมรับ H_0 ทุกค่าของ i แสดงว่าไม่มีตัวแปรอิสระตัวใดมีความสัมพันธ์กับ Y จึงไม่ต้องทำการทดสอบต่อไป
2. ปฏิเสธ H_0 สำหรับบางค่าของ i นั่นคือมี X บางตัวที่มีความสัมพันธ์กับ Y และจะเลือก X ตัวที่มีความสัมพันธ์กับ Yมากที่สุดเข้าสมการ แล้วจึงทำการเลือกตัวแปรอิสระตัวต่อไปจนกระทั่งไม่มีตัวแปรอิสระตัวใดที่มีความสัมพันธ์อย่างมีนัยสำคัญกับตัวแปรตามก็จะได้สมการความถดถอยที่เหมาะสม

4) วิธีสเตปไวส์

วิธีสเตปไวส์ เป็นวิธีเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยใช้หลักเกณฑ์ของทั้งแบ็คเวิร์ดและวิธีฟอร์เวิร์ดรวมกันนั่นคือ

<u>ขั้นที่ 1</u> จะใช้เกณฑ์ของวิธีฟอร์เวิร์ดโดยการเลือกตัวแปรอิสระ 1 ตัวที่มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญมากที่สุดเข้าสมการ

<u>ขั้นที่ 2</u> ใช้เกณฑ์ของวิธีแบ็คเวิร์ดและวิธีฟอร์เวิร์ด นั่นคือจากตัวแปรอิสระที่เหลือ k-1 ตัว เลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญมากที่สุดเข้าสมการ ในขณะเดียวกัน จะต้องทดสอบความสัมพันธ์ของตัวแปรตามกับตัวแปรอิสระที่เลือกไว้ทุกตัวและทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระตัวใหม่ที่เลือกซึ่งอาจจะทำให้ต้องตัดตัวแปรบางตัวออก

ทำซ้ำขั้นตอนที่ 2 จนกระทั่งไม่มีตัวแปรอิสระใดที่ควรนำเข้าสมการแล้วจะได้สมการความถดถอยที่เหมาะสม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 3

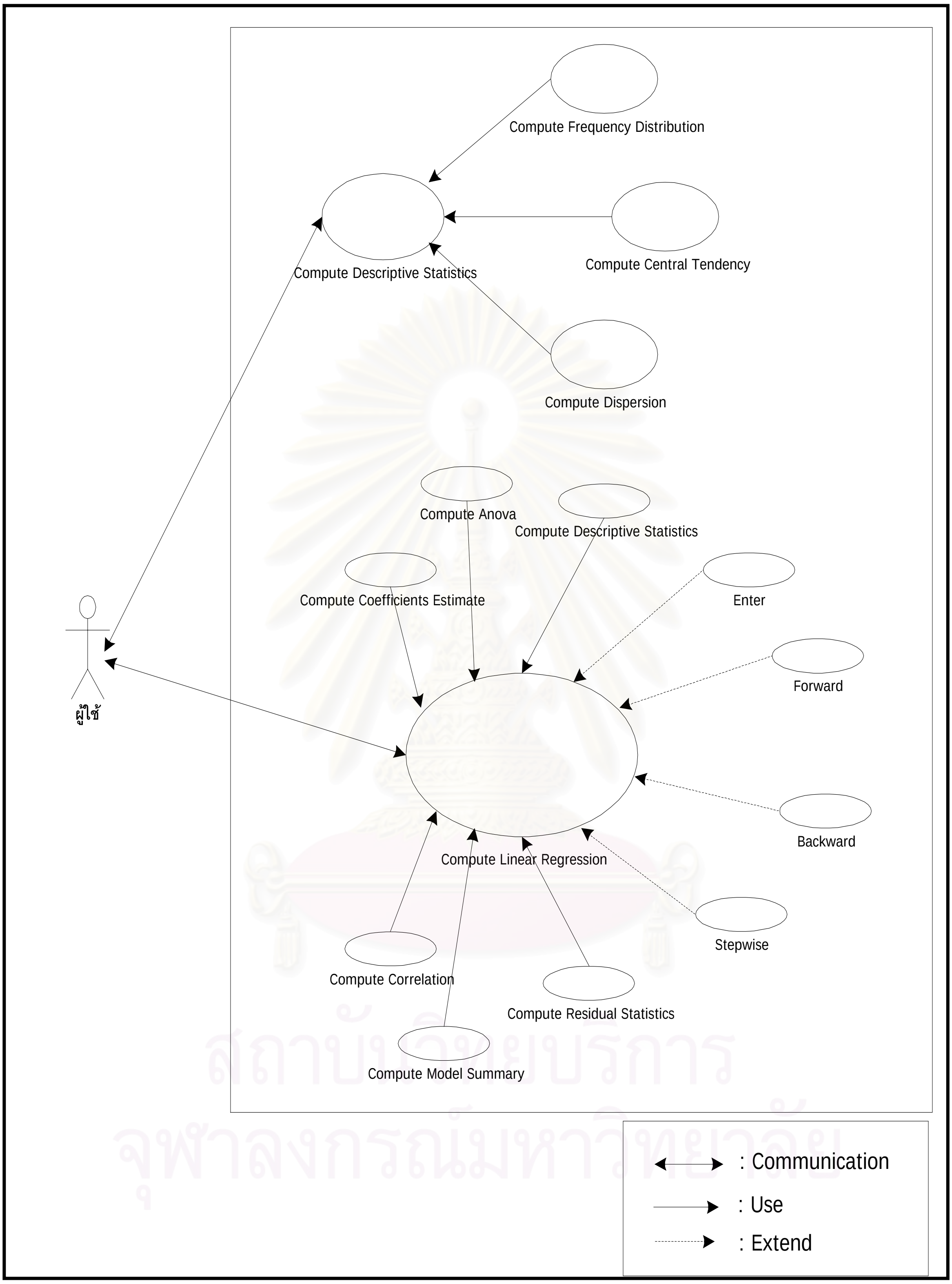
การวิเคราะห์และออกแบบระบบ

ในบทนี้จะนำเสนอ แผนผังยูสเคส แผนผังคลาสและแผนผังความสัมพันธ์ของคลาส โดยเริ่มจากการอธิบายถึงรายละเอียดของคลาสแต่ละคลาสที่เกี่ยวข้องกับสถิติเชิงพรรณนา คลาสที่เกี่ยวข้องกับการวิเคราะห์ความถดถอยเชิงเส้นและคลาสที่เกี่ยวข้องกับตารางสถิติ ในการอธิบายจะแสดงรายละเอียดของคลาสแต่ละคลาสและความสัมพันธ์ระหว่างคลาส โดยความสัมพันธ์ของคลาสจะพิจารณาความสัมพันธ์ของคลาสที่เกี่ยวข้องกับสถิติเชิงพรรณนา ความสัมพันธ์ของคลาสที่เกี่ยวข้องกับการวิเคราะห์ความถดถอยเชิงเส้น รวมถึงพิจารณาความสัมพันธ์ของคลาสทั้งหมดในขั้นตอนการวิเคราะห์และออกแบบระบบ

3.1 ความสามารถในการคำนวณค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ผู้วิจัยได้พัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุให้มีความสามารถในการคำนวณค่าสถิติเชิงพรรณและการวิเคราะห์ความถดถอยเชิงเส้น โดยความสามารถของส่วนประกอบซอฟแวร์เชิงวัตถุประกอบด้วยความสามารถในการคำนวณค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นโดยมีรายละเอียดแทนได้ด้วยผังยูเอ็มแอลดังแสดงในรูปที่ 3.1

จากรูปที่ 3.1 แสดงถึงความสามารถในการวิเคราะห์สถิติเชิงพรรณนาซึ่งประกอบด้วย ความสามารถในการนับค่าความถี่ของข้อมูล ความสามารถในการคำนวณสถิติแสดงค่ากลางของข้อมูล และความสามารถในการคำนวณค่าการกระจายของข้อมูล ส่วนความสามารถในการคำนวณค่าสถิติการวิเคราะห์ความถดถอยเชิงเส้นประกอบด้วยการคำนวณค่าสถิติต่างๆได้แก่ การคำนวณค่าสถิติสรุปของสมการถดถอย การคำนวณค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว การคำนวณค่าความแปร ปรวนของการวิเคราะห์ความถดถอยเชิงเส้น การคำนวณค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้นและการคำนวณค่าสถิติของค่าคลาดเคลื่อน โดยในการวิเคราะห์ความถดถอยเชิงเส้นสามารถกำหนดมีวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นได้ 4 วิธีคือ วิธีเอนเตอร์ วิธีฟอร์เวิร์ด วิธีแบ็คเวิร์ดและ วิธีสเตปไวส์



รูปที่ 3.1 ยูสเคสแสดงความสามารถของส่วนประกอบทางซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ ความถดถอยเชิงเส้น

3.2 การวิเคราะห์และออกแบบคลาส

ขั้นตอนของการวิเคราะห์และออกแบบเป็นขั้นตอนที่พิจารณาคลาสที่เกี่ยวข้องในขอบเขตของปัญหา (Problem Domain) รวมทั้งคลาสอื่นๆที่เกี่ยวข้องทั้งหมด โดยในแต่ละคลาสจะพิจารณาสมาชิกคลาสซึ่งได้แก่ คุณสมบัติและหน้าที่ของคลาสรวมทั้งความสัมพันธ์ระหว่างคลาส ในขั้นตอนของการกำหนดรายละเอียดของคลาสจะต้องกำหนดขอบเขตการมองเห็น (Scope of Visibility) สำหรับสมาชิกภายในคลาสซึ่งประกอบด้วยคุณสมบัติและเมทธอด ในการกำหนดขอบเขตการมองเห็นสามารถกำหนดได้เป็น 3 ระดับคือ ระดับพับบลิค ระดับโพรเทค และระดับ ไพรเวท

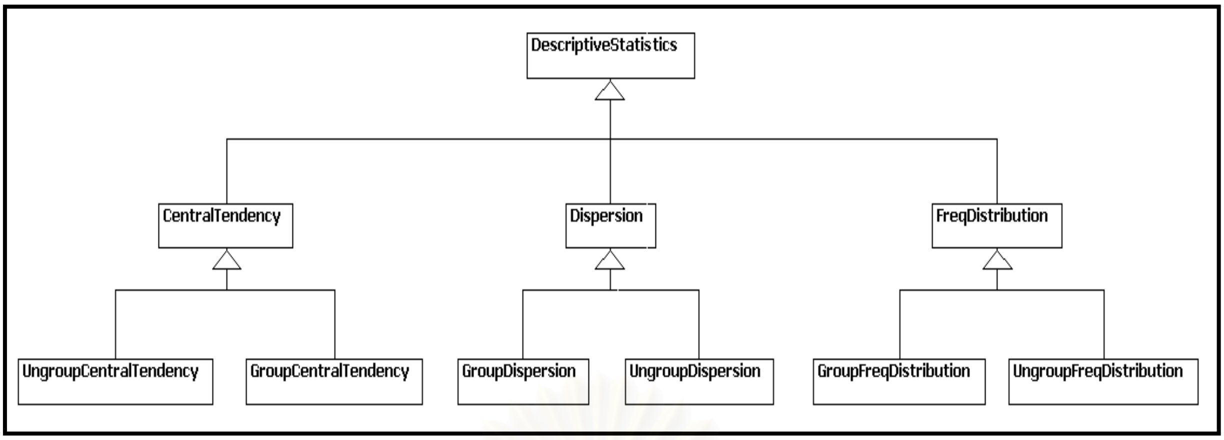
จากการวิเคราะห์โดยการศึกษาเนื้อหาทางสถิติจะได้คลาสต่างๆสำหรับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น รวมทั้งคลาสที่ทำหน้าที่คำนวณค่าสถิติทดสอบซึ่งมีความเกี่ยวข้องกับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น ดังนี้

3.2.1 คลาสสำหรับสถิติเชิงพรรณนา

สถิติเชิงพรรณนาเป็นหลักการในการบรรยายลักษณะของข้อมูลที่ได้เก็บรวบรวม โดยสรุปลักษณะเฉพาะที่สำคัญของข้อมูลกลุ่มที่ศึกษา สถิติเชิงพรรณนาประกอบด้วย การวัดค่ากลางของข้อมูล การวัดค่าการกระจายของข้อมูล การนับค่าความถี่ของข้อมูล และการนำเสนอข้อมูลค่าความถี่ของข้อมูลในรูปของกราฟความถี่คลาสสำหรับสถิติเชิงพรรณนาประกอบด้วยคลาสต่างๆได้แก่

1. คลาส DescriptiveStatistics
2. คลาส FreqDistribution
3. คลาส GroupFreqDistribution
4. คลาส UngroupFreqDistribution
5. คลาส CentralTendency
6. คลาส GroupCentralTendency
7. คลาส UngroupCentralTendency
8. คลาส Dispersion
9. คลาส GroupDispersion
10. คลาส UngroupDispersion

ความสัมพันธ์ของคลาสต่างๆสำหรับการวิเคราะห์สถิติเชิงพรรณนามีรายละเอียดดังแสดงในรูปที่ 3.2



รูปที่ 3.2 ความสัมพันธ์ระหว่างคลาสสำหรับสถิติเชิงพรรณนา

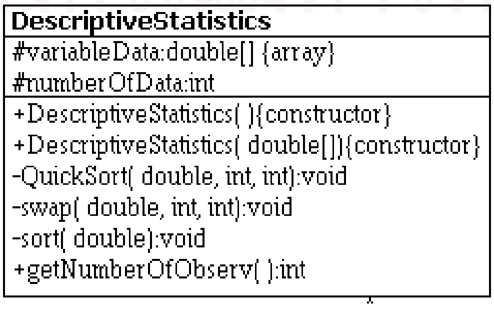
จากรูปที่ 3.2 คลาส DescriptiveStatistics เป็นคลาสหลัก (Superclass) ซึ่งประกอบด้วยคลาสย่อย (Subclass) 3 คลาสคือ

1. คลาส CentralTendency ประกอบด้วยคลาสย่อย 2 คลาสคือ คลาส GroupCentralTendency และ คลาส UngroupCentralTendency
2. คลาส Dispersion ประกอบด้วยคลาสย่อย 2 คลาสคือ คลาส GroupDispersion และ คลาส UngroupDispersion
3. คลาส FreqDistribution ประกอบด้วยคลาสย่อย 2 คลาสคือ คลาส GroupFreqDistribution และ คลาส UngroupFreqDistribution

รายละเอียดของคลาสต่างๆสำหรับการวิเคราะห์สถิติเชิงพรรณนามีดังนี้

3.2.1.1 คลาส DescriptiveStatistics

คลาส DescriptiveStatistics เป็นคลาสสำหรับสถิติเชิงพรรณนาซึ่งประกอบด้วย การแจกแจงความถี่ การวัดค่ากลางของข้อมูลและการวัดการกระจายของข้อมูล คลาส DescriptiveStatistics ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.3



รูปที่ 3.3 ส่วนประกอบของคลาส DescriptiveStatistics

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

จากรูปที่ 3.3 คลาส DescriptiveStatistics ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงใน ตารางที่ 3.1

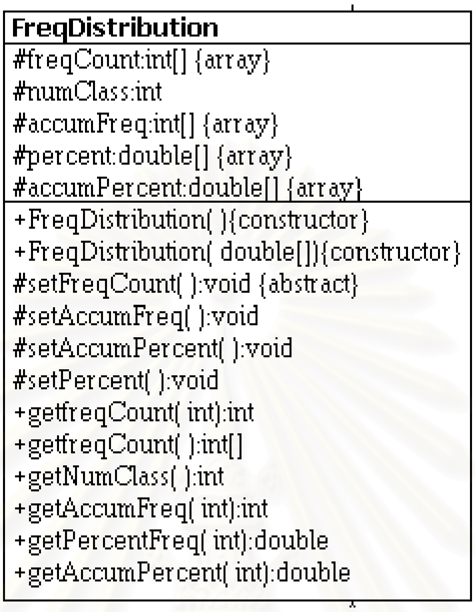
ตารางที่ 3.1 รายละเอียด ของคลาส DescriptiveStatistics

ชื่อคลาส	**DescriptiveStatistics**
ประเภท	Abstract Base Class
คำอธิบาย	สถิติเชิงพรรณนา(Descriptive Statistics) ประกอบด้วย 1. การแจกแจงความถี่ 2. การวัดค่ากลางของข้อมูล 3. การวัดการกระจายของข้อมูล
ตัวแปร	▪ variableData ค่าข้อมูลของตัวแปร ▪ numberOfData จำนวนข้อมูลตัวอย่าง
เมทธอด	▪ DescriptiveStatistics() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์ ▪ DescriptiveStatistics(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวิเคราะห์สถิติเชิงพรรณนา ▪ sort() จัดเรียงข้อมูลจากค่าน้อยไปมาก *พารามิเตอร์ที่ใช้* a หมายถึง ค่าตัวแปรที่ต้องการเรียงลำดับ ▪ QuickSort(double a[], int lo0, int hi0) เมทธอดนี้ใช้สำหรับเรียงข้อมูลจากค่าน้อยไปมาก โดยวิธี Quick Sort *พารามิเตอร์ที่ใช้* a หมายถึง ค่าตัวแปรที่ต้องการเรียงลำดับ lo0 หมายถึง ค่าด้านซ้ายของอาร์เรย์ที่แบ่งแล้ว hi0 หมายถึง ค่าด้านซ้ายของอาร์เรย์ที่แบ่งแล้ว ▪ swap(double a[], int i, int j) เมทธอดนี้ใช้สำหรับสลับค่าของตัวแปรตามพารามิเตอร์ที่ระบุ *พารามิเตอร์ที่ใช้* a หมายถึง ค่าตัวแปรที่ต้องการเรียงลำดับ i หมายถึง ตำแหน่งค่าที่ i ของตัวแปรอาร์เรย์ a j หมายถึง ตำแหน่งค่าที่ j ของตัวแปรอาร์เรย์ a ▪ getNumberOfObserv() รับค่าจำนวนข้อมูลตัวอย่าง *ค่าส่งกลับ* จำนวนข้อมูลตัวอย่าง

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.2.1.2 คลาส FreqDistribution

คลาส FreqDistribution เป็นคลาสสำหรับการแจกแจงความถี่ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.4



รูปที่ 3.4 ส่วนประกอบของคลาส FreqDistribution

จากรูปที่ 3.4 คลาส FreqDistribution ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 รายละเอียด ของคลาส FreqDistribution

ชื่อคลาส	**FreqDistribution**
ประเภท	คลาสย่อย ของ คลาส DescribtiveStatistics
คำอธิบาย	การนับค่าความถี่ของข้อมูล(Frequency Distribution)
ตัวแปร	▪ freqCount ค่าความถี่ของข้อมูล ▪ numClass จำนวนชั้นของค่าความถี่ ▪ accumFreq ค่าความถี่สะสม ▪ percent ค่าเปอร์เซ็นต์ความถี่ ▪ accumPercent ค่าเปอร์เซ็นต์ความถี่สะสม
เมทธอด	▪ FreqDistribution() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์ (Default Constructor)

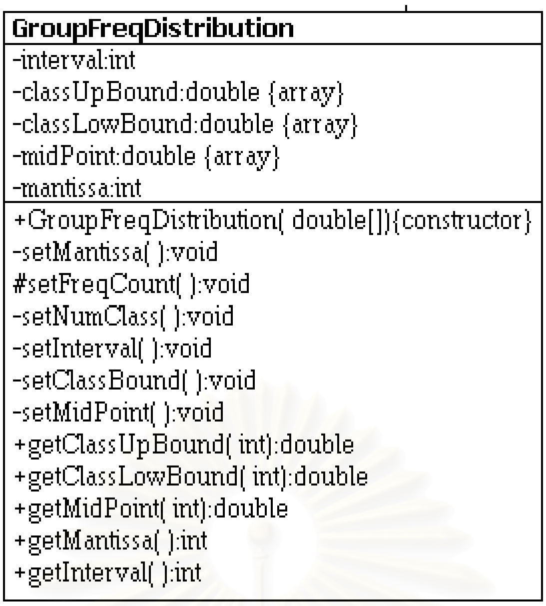
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.2 รายละเอียด ของคลาส FreqDistribution (ต่อ)

เมทธอด	■ FreqDistribution(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการนับค่าความถี่ ■ setFreqCount() นับค่าความถี่ ■ setAccumFreq() คำนวณค่าความถี่สะสม ■ setAccumPercent() คำนวณเปอร์เซ็นต์ความถี่สะสม ■ setPercent() คำนวณค่าเปอร์เซ็นต์ความถี่ ■ getFreqCount() รับค่าความถี่ของทุกชั้นความถี่ *ค่าส่งกลับ* ค่าความถี่ในแต่ละชั้นความถี่ ■ getfreqCount(int index) เมทธอดนี้ใช้สำหรับรับค่าความถี่ *ค่าส่งกลับ* ค่าความถี่ของชั้นที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ■ getNumClass() รับค่าจำนวนชั้นความถี่ *ค่าส่งกลับ* จำนวนชั้นความถี่ ■ getAccumFreq() รับค่าความถี่สะสม *ค่าส่งกลับ* ค่าความถี่สะสมของชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ■ getPercentFreq() รับค่าเปอร์เซ็นต์ความถี่ *ค่าส่งกลับ* ค่าเปอร์เซ็นต์ของค่าความถี่ของชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ■ getAccumPercent() รับค่าเปอร์เซ็นต์ความถี่สะสม *ค่าส่งกลับ* ค่าเปอร์เซ็นของค่าความถี่สะสมถึงชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่

3.2.1.3 คลาส GroupFreqDistribution

คลาส GroupFreqDistribution เป็นคลาสสำหรับการแจกแจงความถี่ของข้อมูลที่จัดกลุ่มแล้ว ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.5



รูปที่ 3.5 ส่วนประกอบของคลาส GroupFreqDistribution

จากรูปที่ 3.4 คลาส GroupFreqDistribution ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.3

ตารางที่ 3.3 รายละเอียด ของคลาส GroupFreqDistribution

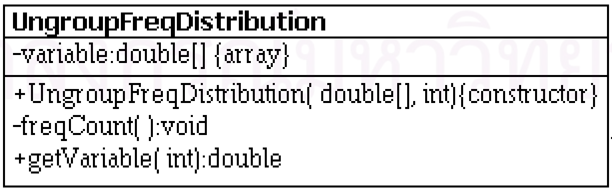
ชื่อคลาส	**GroupFreqDistribution**
ประเภท	สืบทอด จาก คลาส FreqDistribution
คำอธิบาย	การนับค่าความถี่ของข้อมูลแบบจัดกลุ่ม(Group Frequency Distribution)
ตัวแปร	▪ interval ความกว้างของชั้น(Class interval) ▪ classUpBound ขีดจำกัดบน ▪ classLowBound ขีดจำกัดล่าง ▪ midPoint จุดกึ่งกลางชั้น ▪ mantissa จำนวนค่าหลังจุดทศนิยม
เมทธอด	▪ GroupFreqDistribution(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการนับค่าความถี่แบบจัดกลุ่ม ▪ setMantissa() เมทธอดนี้ใช้สำหรับนับจำนวนค่าหลังจุดทศนิยม ▪ setFreqCount() เมทธอดนี้ใช้สำหรับนับค่าความถี่ ▪ setNumClass() เมทธอดนี้ใช้สำหรับคำนวณจำนวนชั้นความถี่ที่เหมาะสม

ตารางที่ 3.3 รายละเอียด ของคลาส GroupFreqDistribution (ต่อ)

เมทธอด	▪ setInterval() เมทธอดนี้ใช้สำหรับคำนวณค่าความกว้างของชั้นความถี่ ▪ setClassBound() เมทธอดนี้ใช้สำหรับคำนวณค่าขีดจำกัดชั้นความถี่ ▪ setMidPoint() เมทธอดนี้ใช้สำหรับคำนวณค่าจุดกึ่งกลางชั้นความถี่ ▪ getClassUpBound(int index) เมทธอดนี้ใช้สำหรับรับค่าขอบเขตบนของชั้นที่กำหนด *ค่าส่งกลับ* ค่าขอบเขตบนของชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ▪ getClassLowBound(int index) เมทธอดนี้ใช้สำหรับรับค่าขอบเขตล่างของชั้นความถี่ที่กำหนด *ค่าส่งกลับ* ค่าขอบเขตล่างของชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ▪ getMidPoint(int index) เมทธอดนี้ใช้สำหรับรับค่าจุดกึ่งกลางชั้นของชั้นความถี่ที่กำหนด *ค่าส่งกลับ* ค่ากึ่งกลางชั้นของชั้นความถี่ที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่ ▪ getMantissa() เมทธอดนี้ใช้สำหรับรับค่าจำนวนค่าหลังจุดทศนิยม *ค่าส่งกลับ* จำนวนค่าหลังจุดทศนิยม ▪ public int getInterval() เมทธอดนี้ใช้สำหรับรับค่าความกว้างของชั้นความถี่ *ค่าส่งกลับ* ความกว้างของชั้นความถี่

3.2.1.4 คลาส UngroupFreqDistribution

คลาส UngroupFreqDistribution เป็นคลาสสำหรับการแจกแจงความถี่ของข้อมูลที่ไม่ได้จัดกลุ่ม ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.6



รูปที่ 3.6 ส่วนประกอบของคลาส UngroupFreqDistribution

จากรูปที่ 3.6 คลาส UngroupFreqDistribution ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.4

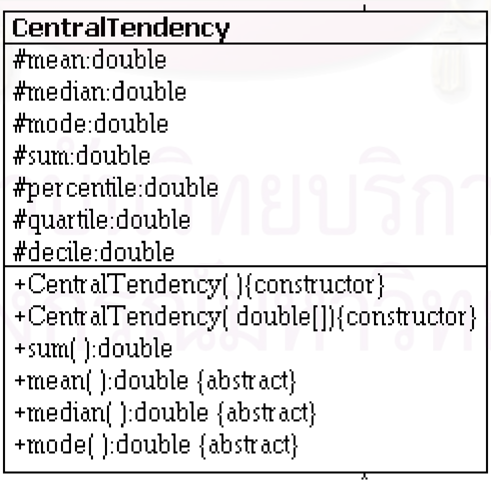
สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.4 รายละเอียด ของคลาส UngroupFreqDistribution

ชื่อคลาส	**UngroupFreqDistribution**
ประเภท	สืบทอดจากคลาส FreqDistribution
คำอธิบาย	การนับค่าความถี่ของข้อมูลที่ไม่ได้จัดกลุ่ม(Ungroup Frequency Distribution)
ตัวแปร	▪ variable ค่าของตัวแปรที่นับค่าความถี่
เมทธอด	▪ **UngroupFreqDistribution(double[] varValue)** เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการนับค่าความถี่แบบไม่จัดกลุ่ม ▪ **setFreqCount()** เมทธอดนี้ใช้สำหรับคำนวณค่าความถี่ ▪ **getVariable(int index)** เมทธอดนี้ใช้สำหรับรับค่าตัวแปรของชั้นที่ระบุ *ค่าส่งกลับ* ค่าตัวแปรของชั้นที่ระบุ *พารามิเตอร์ที่ใช้* index หมายถึง ชั้นความถี่

3.2.1.5 คลาส CentralTendency

คลาส CentralTendency เป็นคลาสสำหรับการวัดค่ากลางของข้อมูลประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.7



รูปที่ 3.7 ส่วนประกอบของคลาส CentralTendency

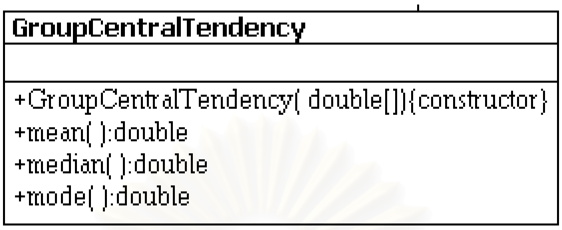
จากรูปที่ 3.7 คลาส CentralTendency ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงใน ตารางที่ 3.5

ตารางที่ 3.5 รายละเอียด ของคลาส CentralTendency

ชื่อคลาส	**CentralTendency**
ประเภท	คลาสย่อย ของ คลาส DescribtiveStatistics
คำอธิบาย	เป็นสถิติที่แสดงค่ากลางของข้อมูล
ตัวแปร	▪ mean ค่าเฉลี่ยเลขคณิต ▪ median ค่ามัธยฐาน ▪ mode ค่าฐานนิยม ▪ sum ผลบวกของข้อมูล ▪ percentile ค่าเปอร์เซ็นต์ไทล์ ▪ quartile ค่าควอไทล์ ▪ decile ค่าเดไซล์
เมทธอด	▪ CentralTendency() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์(Default Constructor) ▪ CentralTendency(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวัดค่ากลางทางสถิติ ▪ sum() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกของข้อมูล *ค่าส่งกลับ* ค่าผลบวกของข้อมูล ▪ mean() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยเลขคณิต *ค่าส่งกลับ* ค่าเฉลี่ยเลขคณิต ▪ double median() เมทธอดนี้ใช้สำหรับคำนวณค่ามัธยฐาน *ค่าส่งกลับ* ค่ามัธยฐาน ▪ double mode() เมทธอดนี้ใช้สำหรับคำนวณค่าฐานนิยม *ค่าส่งกลับ* ค่าฐานนิยม

3.2.1.6 คลาส GroupCentralTendency

คลาส GroupCentralTendency เป็นคลาสสำหรับการวัดค่ากลางของข้อมูลที่จัดกลุ่มแล้วประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.8



รูปที่ 3.8 ส่วนประกอบของคลาส GroupCentralTendency

จากรูปที่ 3.8 คลาส GroupCentralTendency ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.6

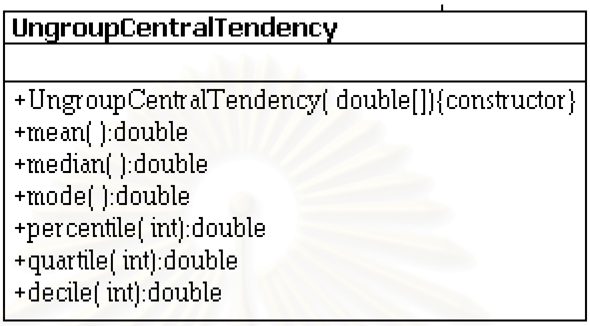
ตารางที่ 3.6 รายละเอียด ของคลาส GroupCentralTendency

ชื่อคลาส	**GroupCentralTendency**
ประเภท	สืบทอดจากคลาส CentralTendency
คำอธิบาย	เป็นสถิติที่แสดงค่ากลางของข้อมูลที่จัดกลุ่ม
ตัวแปร	ตัวแปรที่สืบทอดจากคลาส Central Tendency
เมทธอด	■ GroupCentralTendency(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวัดค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม ■ public double mean() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยเลขคณิต *ค่าส่งกลับ* ค่าเฉลี่ยเลขคณิต ■ public double median() เมทธอดนี้ใช้สำหรับคำนวณค่ามัธยฐาน *ค่าส่งกลับ* ค่ามัธยฐาน ■ public double mode() เมทธอดนี้ใช้สำหรับคำนวณค่าฐานนิยม *ค่าส่งกลับ* ค่าฐานนิยม

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3.2.1.7 คลาส UngroupCentralTendency

คลาส UngroupCentralTendency เป็นคลาสสำหรับการวัดค่ากลางของข้อมูลที่ไม่ได้จัดกลุ่มประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.9 จากรูปที่ 3.9 คลาส UngroupCentralTendency ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.7



รูปที่ 3.9 ส่วนประกอบของคลาส UngroupCentralTendency

ตารางที่ 3.7 รายละเอียด ของคลาส UngroupCentralTendency

ชื่อคลาส	**UngroupCentralTendency**
ประเภท	สืบทอดจากคลาส CentralTendency
คำอธิบาย	เป็นสถิติที่แสดงค่ากลางของข้อมูลที่ไม่ได้จัดกลุ่ม
ตัวแปร	ตัวแปรที่สืบทอดจากคลาส Central Tendency
เมทธอด	■ public UngroupCentralTendency(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวัดค่ากลางทางสถิติของข้อมูลที่ไม่ได้จัดกลุ่ม ■ public double mean() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยเลขคณิต *ค่าส่งกลับ* ค่าเฉลี่ยเลขคณิต ■ public double median() เมทธอดนี้ใช้สำหรับคำนวณค่ามัธยฐาน *ค่าส่งกลับ* ค่ามัธยฐาน ■ public double mode() เมทธอดนี้ใช้สำหรับคำนวณค่าฐานนิยม *ค่าส่งกลับ* ค่าฐานนิยม ■ public double percentile(int percent) เมทธอดนี้ใช้สำหรับคำนวณค่าเปอร์เซ็นต์ไทล์

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

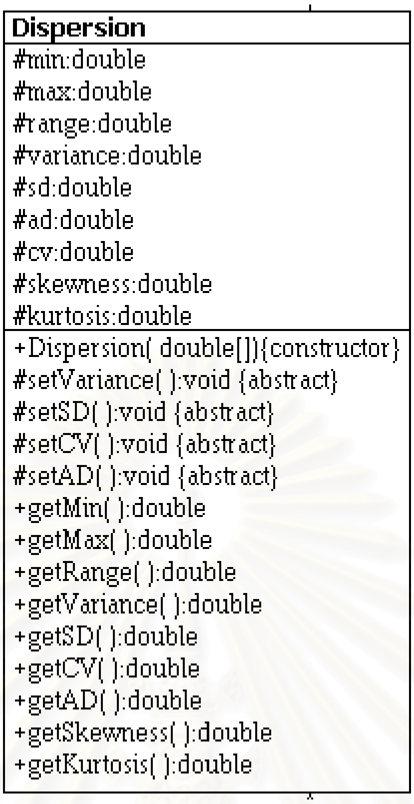
ตารางที่ 3.7 รายละเอียด ของคลาส UngroupCentralTendency (ต่อ)

เมทธอด	*ค่าส่งกลับ* ค่าเปอร์เซ็นต์ไทล์ *พารามิเตอร์ที่ใช้* percent หมายถึง ลำดับค่าเปอร์เซ็นต์ไทล์ ■ public double quartile(int quarter) เมทธอดนี้ใช้สำหรับคำนวณค่าควอไทล์ *ค่าส่งกลับ* ค่าควอไทล์ *พารามิเตอร์ที่ใช้* quarter หมายถึง ลำดับค่าควอไทล์ มีความหมายดังนี้ ลำดับที่ 1 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 25 ลำดับที่ 2 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 50 ลำดับที่ 3 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 75 ลำดับที่ 4 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 100 ■ public double decile(int dec) เมทธอดนี้ใช้สำหรับคำนวณค่าเดไซล์ *ค่าส่งกลับ* ค่าเดไซล์ *พารามิเตอร์ที่ใช้* dec หมายถึง ลำดับค่าเดไซล์ มีความหมายดังนี้ ลำดับที่ 1 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 10 ลำดับที่ 2 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 20 ลำดับที่ 3 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 30 ลำดับที่ 4 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 40 ลำดับที่ 5 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 50 ลำดับที่ 6 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 60 ลำดับที่ 7 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 70 ลำดับที่ 8 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 80 ลำดับที่ 9 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 90 ลำดับที่ 10 หมายถึง ค่าเปอร์เซ็นต์ไทล์ที่ 100

3.2.1.8 คลาส Dispersion

คลาส Dispersion เป็นคลาสสำหรับการวัดการกระจายของข้อมูลประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.10

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.10 ส่วนประกอบของคลาส Dispersion

จากรูปที่ 3.10 คลาส Dispersion ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.8

ตารางที่ 3.8 รายละเอียด ของคลาส Dispersion

ชื่อคลาส	**Dispersion**
ประเภท	คลาสย่อยของคลาส DescribtiveStatistics
คำอธิบาย	ค่าสถิติที่วัดการกระจายของข้อมูล
ตัวแปร	■ Dispersion() Default constructor ■ min ค่าต่ำสุดของข้อมูล(Minimum) ■ max ค่าสูงสุดของข้อมูล(Maximum) ■ range พิสัยของข้อมูล ■ variance ค่าแปรปรวน(Variance) ■ sd ค่าเบี่ยงเบนมาตรฐานของข้อมูล(Standard Deviation) ad ค่าส่วนเบี่ยงเบนเฉลี่ย

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

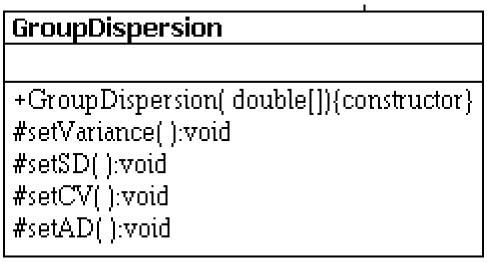
ตารางที่ 3.8 รายละเอียด ของคลาส Dispersion (ต่อ)

ตัวแปร	▪ cv ค่าสัมประสิทธิ์ความแปรผัน
เมทธอด	▪ Dispersion(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวิเคราะห์การกระจาย ▪ setVariance() เมทธอดนี้ใช้สำหรับคำนวณค่าแปรปรวน ▪ setSD() เมทธอดนี้ใช้สำหรับคำนวณค่าเบี่ยงเบนมาตรฐาน ▪ setCV() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์ความแปรผัน ▪ setAD() เมทธอดนี้ใช้สำหรับคำนวณค่าส่วนเบี่ยงเบนเฉลี่ย ▪ getMin() เมทธอดนี้ใช้สำหรับคำนวณค่าต่ำสุดของข้อมูล *ค่าส่งกลับ* ค่าต่ำสุดของข้อมูล ▪ getMax() เมทธอดนี้ใช้สำหรับคำนวณค่าสูงสุดของข้อมูล *ค่าส่งกลับ* ค่าสูงสุดของข้อมูล ▪ getRange() เมทธอดนี้ใช้สำหรับรับค่าพิสัย *ค่าส่งกลับ* ค่าพิสัย ▪ getVariance() เมทธอดนี้ใช้สำหรับรับค่าแปรปรวน *ค่าส่งกลับ* ค่าแปรปรวน ▪ getSD() เมทธอดนี้ใช้สำหรับรับค่าเบี่ยงเบนมาตรฐาน *ค่าส่งกลับ* ค่าเบี่ยงเบนมาตรฐาน ▪ getCV() เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธิ์ความแปรผัน *ค่าส่งกลับ* ค่าสัมประสิทธิ์ความแปรผัน ▪ getAD() เมทธอดนี้ใช้สำหรับรับค่าส่วนเบี่ยงเบนเฉลี่ย *ค่าส่งกลับ* ค่าส่วนเบี่ยงเบนเฉลี่ย

3.2.1.9 คลาส GroupDispersion

คลาส GroupDispersion เป็นคลาสสำหรับการวัดการกระจายของข้อมูลที่จัดกลุ่มแล้วประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.11

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.11 ส่วนประกอบของคลาส GroupDispersion

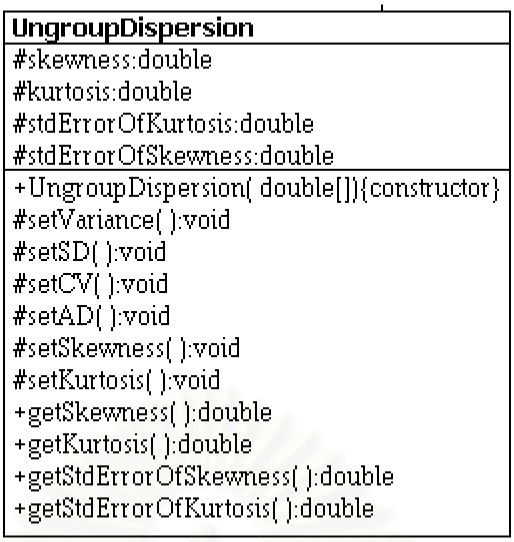
จากรูปที่ 3.11 คลาส GroupDispersion ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.9

ตารางที่ 3.9 แสดงรายละเอียด ของคลาส GroupDispersion

ชื่อคลาส	**GroupDispersion**
ประเภท	สืบทอด จาก คลาส Dispersion
คำอธิบาย	ค่าสถิติที่วัดการกระจายของข้อมูลแบบจัดกลุ่ม
ตัวแปร	ตัวแปรที่สืบทอดจากคลาส Dispersion
เมทธอด	▪ GroupDispersion(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวิเคราะห์การกระจายของข้อมูลแบบจัดกลุ่ม ▪ setVariance()เมทธอดนี้ใช้สำหรับคำนวณค่าแปรปรวนของข้อมูล ▪ setSD() เมทธอดนี้ใช้สำหรับคำนวณค่าเบี่ยงเบนมาตรฐานของข้อมูล ▪ setCV() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์ความแปรผัน ▪ setAD() เมทธอดนี้ใช้สำหรับคำนวณค่าส่วนเบี่ยงเบนเฉลี่ย

3.2.1.10 UngroupDispersion

คลาส UngroupDispersion เป็นคลาสสำหรับการวัดการกระจายของข้อมูลที่ไม่ได้จัดกลุ่มประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.12



รูปที่ 3.12 ส่วนประกอบของคลาส UngroupDispersion

จากรูปที่ 3.12 คลาส UngroupDispersion ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.10

ตารางที่ 3.10 รายละเอียด ของคลาส UngroupDispersion

ชื่อคลาส	**UngroupDispersion**
ประเภท	สืบทอด จาก คลาส Dispersion
คำอธิบาย	ค่าสถิติที่วัดการกระจายของข้อมูลแบบไม่จัดกลุ่ม
ตัวแปร	▪ skewness ค่าความเบ้ของการแจกแจงที่ไม่สมมาตร ▪ kurtosis ค่าความโด่งของการแจกแจงข้อมูล
เมทธอด	▪ UngroupDispersion(double[] varValue) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* varValue หมายถึง ค่าของตัวแปรที่ต้องการวิเคราะห์การกระจายของข้อมูลแบบจัดกลุ่ม ▪ setVariance() เมทธอดนี้ใช้สำหรับคำนวณค่าแปรปรวน ▪ setSD() เมทธอดนี้ใช้สำหรับคำนวณค่าเบี่ยงเบนมาตรฐาน ▪ setCV() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์ความแปรผัน ▪ setAD() เมทธอดนี้ใช้สำหรับคำนวณค่าส่วนเบี่ยงเบนเฉลี่ย ▪ setSkewness() เมทธอดนี้ใช้สำหรับคำนวณค่าความเบ้ของการแจกแจงที่ไม่สมมาตร

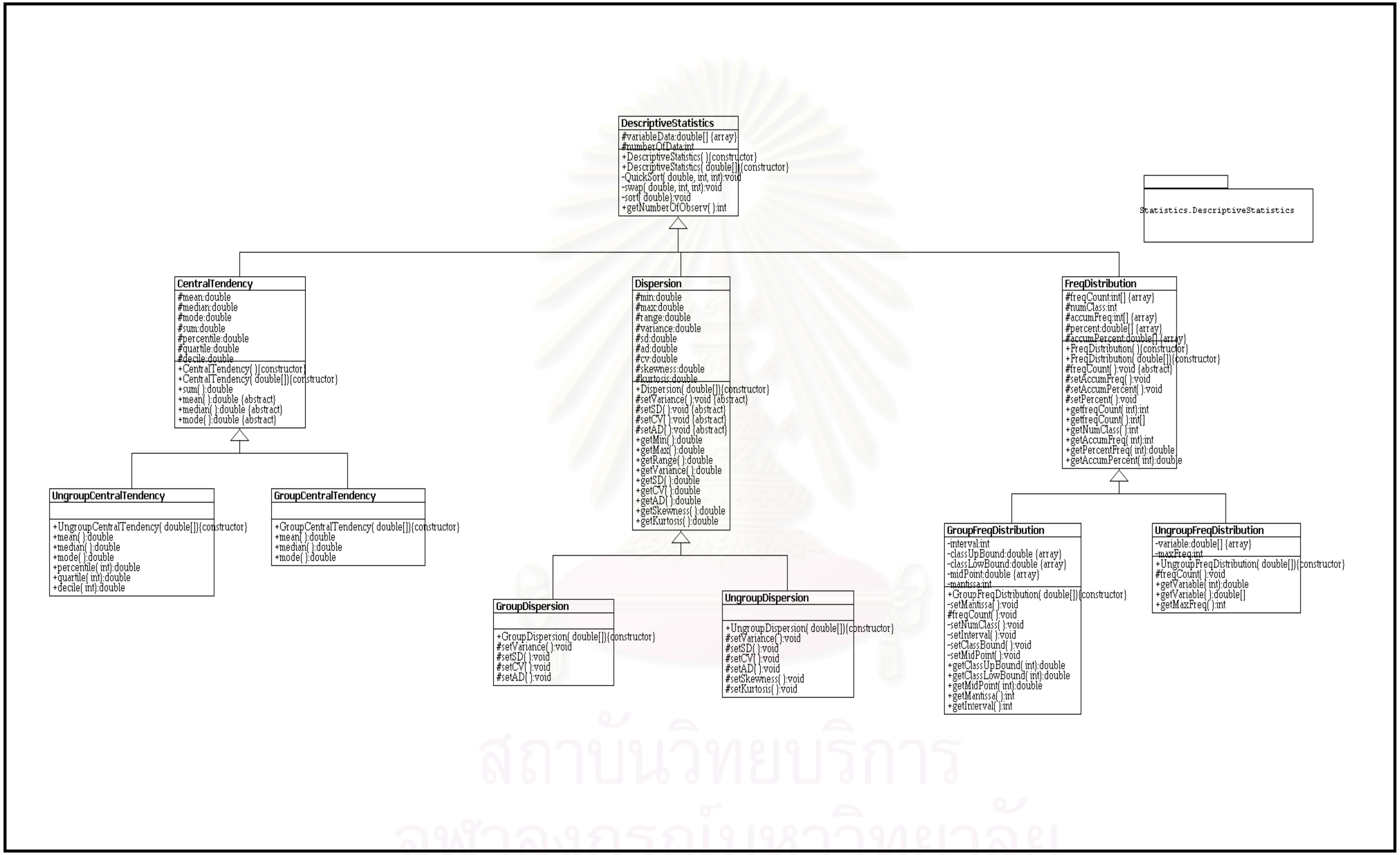
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.10 รายละเอียด ของคลาส UngroupDispersion (ต่อ)

เมทธอด	▪ setKurtosis() เมทธอดนี้ใช้สำหรับคำนวณค่าความโด่งของการแจกแจงข้อมูล ▪ getSkewness() เมทธอดนี้ใช้สำหรับรับค่าความเบ้ของการแจกแจงที่ไม่สมมาตร *ค่าส่งกลับ* ค่าความเบ้ ▪ getKurtosis() เมทธอดนี้ใช้สำหรับรับค่าความโด่งของการแจกแจงข้อมูล *ค่าส่งกลับ* ค่าความโด่ง

จากผังแสดงความสัมพันธ์ระหว่างคลาสสำหรับสถิติเชิงพรรณนาในรูปที่ 3.2 และรายละเอียดของคลาสแต่ละคลาสทำให้ได้ผังความสัมพันธ์ระหว่างคลาสและรายละเอียดของคลาสแต่ละคลาสสำหรับสถิติเชิงพรรณนาดังแสดงในรูปที่ 3.13

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.13 ความสัมพันธ์ของคลาสต่างๆสำหรับสถิติเชิงพรรณา

3.2.2 คลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

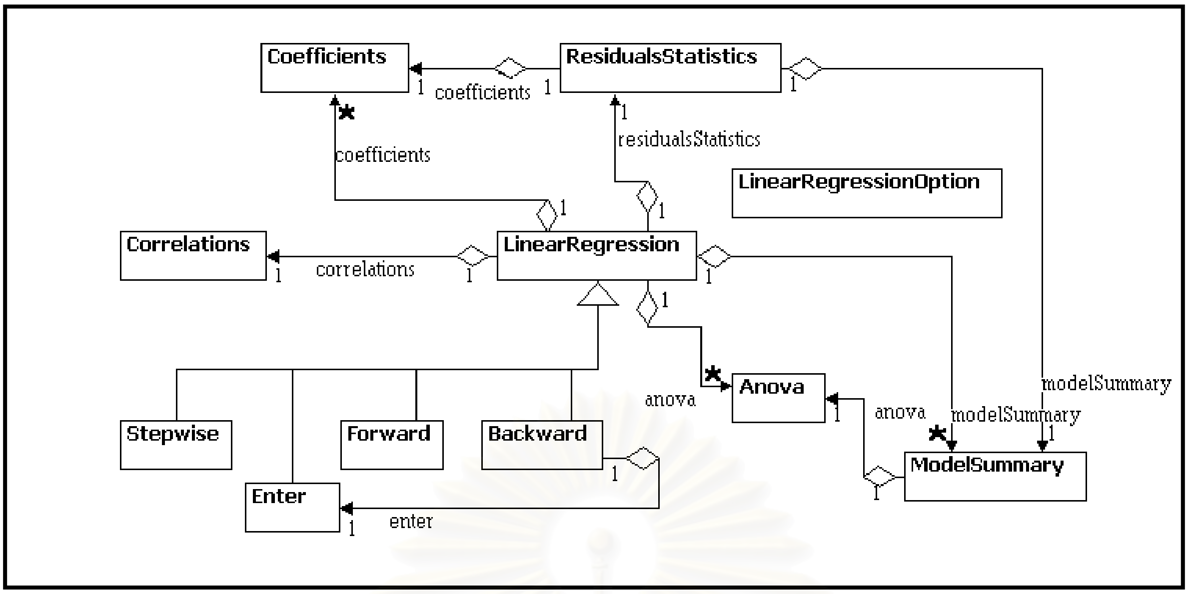
การวิเคราะห์ความถดถอยเป็นการศึกษาความสัมพันธ์เชิงเส้นของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์โดยการเลือกตัวแปรอิสระเข้าสมการความถดถอย ผู้วิเคราะห์จะต้องพิจารณาว่าตัวแปรอิสระใดบ้างที่น่าจะมีความสัมพันธ์กับตัวแปรตาม สำหรับเทคนิคที่ใช้ในการพิจารณาว่าตัวแปรอิสระใดบ้างที่มีความสัมพันธ์กับตัวแปรตามมี 4 วิธีดังนี้

1. วิธีเอนเตอร์ เป็นวิธีการนำตัวแปรอิสระทุกตัวเข้าสมการความถดถอย
2. วิธีแบ็คเวิร์ด เป็นวิธีการตัดตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามออกจากสมการความถดถอยเชิงเส้นทีละตัว
3. วิธีฟอร์เวิร์ด เป็นวิธีการนำตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเข้าสมการความถดถอยเชิงเส้นทีละตัว
4. วิธีสเตปไวส์ เป็นวิธีการนำตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด และในขณะเดียวกัน ก็พิจารณาตัวตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามออกจากสมการ โดยวิธีแบ็คเวิร์ด

คลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นประกอบด้วยคลาสต่างๆได้แก่

1. คลาส LinearRegression
2. คลาส Enter
3. คลาส Forward
4. คลาส Backward
5. คลาส Stepwise
6. คลาส ModelSummary
7. คลาส Correlations
8. คลาส Anova
9. คลาส Coefficients
10. คลาส ResidualsStatistics
11. คลาส LinearRegressionOption

ความสัมพันธ์ของคลาสต่างๆสำหรับการวิเคราะห์ความถดถอยเชิงเส้นมีรายละเอียดดังแสดงในรูปที่ 3.14

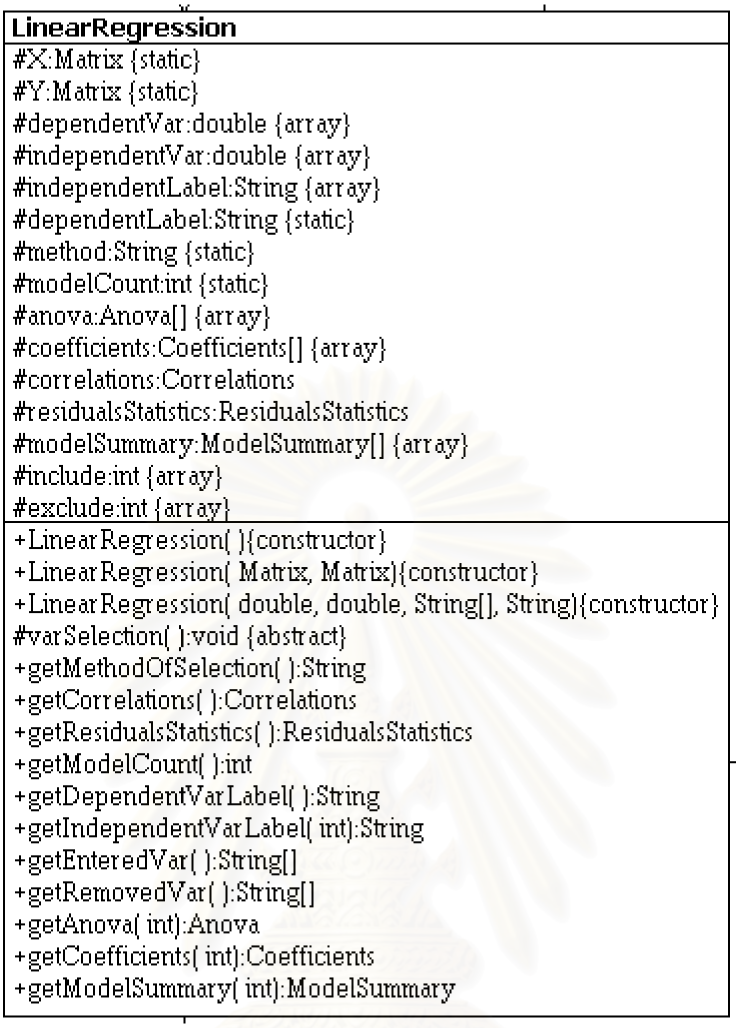


รูปที่ 3.14 ความสัมพันธ์ระหว่างคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

จากรูปที่ 3.14 คลาส LinearRegression ประกอบด้วยคลาสสมาชิกคือ คลาส ModelSummary คลาส Correlations คลาส Anova คลาส Coefficients และ คลาส ResidualsStatistics นอกจากนั้นคลาส LinearRegression ยังเป็นคลาสหลักซึ่งประกอบด้วยคลาสย่อย 4 คลาสคือ คลาส Enter คลาส Forward คลาส Backward และ คลาส Stepwise ซึ่งคลาสทั้ง 4 คลาสจะได้รับการถ่ายทอดคุณสมบัติจากคลาส LinearRegression นั่นคือ ในแต่ละคลาสจะประกอบด้วยคลาสสมาชิกต่างๆ คือ คลาส ModelSummary คลาส Correlations คลาส Anova คลาส Coefficients และ คลาส ResidualsStatistics และสำหรับคลาส LinearRegressionOption เป็นคลาสที่ไม่มีความสัมพันธ์กับคลาสอื่นๆ รายละเอียดของคลาสต่างๆสำหรับการวิเคราะห์ความถดถอยเชิงเส้นมีดังนี้

3.2.2.1 คลาส LinearRegression

คลาส LinearRegression เป็นคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.15 คลาส LinearRegression ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงใน ตารางที่ 3.11



รูปที่ 3.15 ส่วนประกอบของคลาส LinearRegression

ตารางที่ 3.11 รายละเอียด ของคลาส LinearRegression

ชื่อคลาส	**LinearRegression**
ประเภท	Base Case
คำอธิบาย	การวิเคราะห์ความถดถอยเชิงเส้น เป็นการศึกษาความสัมพันธ์เชิงเส้นของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์โดยการเลือกตัวแปรอิสระเข้าสมการความถดถอย โดยเทคนิกที่ใช้ในการพิจารณาว่าตัวแปรอิสระใดบ้างที่มีความสัมพันธ์กับตัวแปรตามมี 4 วิธีคือ วิธีเอนเตอร์ วิธีฟอร์เวิร์ด วิธีแบ็คเวิร์ด และ วิธีสเตปไวส์

ตารางที่ 3.11 รายละเอียด ของคลาส LinearRegression (ต่อ)

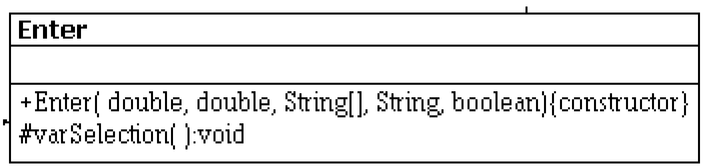
ตัวแปร	▪ X เมตริกซ์ค่าตัวแปรอิสระ ▪ Y เมตริกซ์ค่าตัวแปรตาม ▪ dependentVar ค่าตัวแปรตาม ▪ independentVar ค่าตัวแปรอิสระ ▪ independentLabel ชื่อตัวแปรอิสะ ▪ modelCount จำนวนแบบของสมการความถดถอยเชิงเส้น ▪ anova ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น ▪ coefficients ค่าประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น ▪ correlations ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว ▪ residualsStatistics สถิติวิเคราะห์ของค่าคลาดเคลื่อน ▪ modelSummary สถิติสรุปของสมการความถดถอยเชิงเส้น ▪ include ตัวแปรที่นำเข้าสมการความถดถอยเชิงเส้น ▪ exclude ตัวแปรที่ตัดออกจากสมการความถดถอยเชิงเส้น
เมทธอด	▪ LinearRegression() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์ (Default Constructor) ▪ public LinearRegression(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ▪ LinearRegression(double dependentVar[], double independentVar[][], String[] independentItems) *พารามิเตอร์ที่ใช้* dependentVar หมายถึง ค่าตัวแปรตาม independentVar หมายถึง ค่าตัวแปรอิสระ independentItems หมายถึง ชื่อตัวแปรอิสระ ▪ getAnova(int index) เมทธอดนี้ใช้สำหรับรับค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของรูปแบบสมการความถดถอยเชิงเส้น

ตารางที่ 3.11 รายละเอียด ของคลาส LinearRegression (ต่อ)

เมทธอด	▪ getCoefficients(int index) เมทธอดนี้ใช้สำหรับรับค่าประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของรูปแบบสมการความถดถอยเชิงเส้น ▪ getModelSummary(int index) เมทธอดนี้ใช้สำหรับรับค่าสถิติสรุปของสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสถิติสรุปของสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของรูปแบบสมการความถดถอยเชิงเส้น ▪ getCorrelations() เมทธอดนี้ใช้สำหรับรับค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว *ค่าส่งกลับ* ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว ▪ getResidualsStatistics() เมทธอดนี้ใช้สำหรับรับค่าสถิติวิเคราะห์ของค่าคลาดเคลื่อน *ค่าส่งกลับ* ค่าสถิติวิเคราะห์ของค่าคลาดเคลื่อน ▪ getModelCount() เมทธอดนี้ใช้สำหรับรับค่าจำนวนแบบของสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าจำนวนแบบของสมการความถดถอยเชิงเส้น ▪ getEnteredVar() เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่นำเข้าสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าตัวแปรที่นำเข้าสมการความถดถอยเชิงเส้น ▪ getRemovedVar() เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ตัดออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าตัวแปรที่ตัดออกจากสมการความถดถอยเชิงเส้น

3.2.2.2 คลาส Enter

คลาส Enter เป็นคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีเอนเตอร์ ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.16



รูปที่ 3.16 ส่วนประกอบของคลาส Enter

จากรูปที่ 3.16 คลาส Enter ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.12

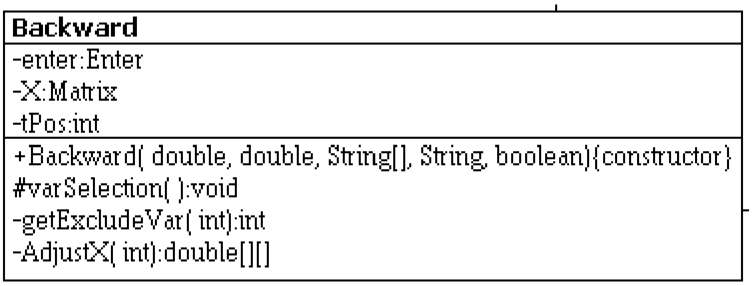
ตารางที่ 3.12 รายละเอียด ของคลาส Enter

ชื่อคลาส	**Enter**
ประเภท	สืบทอดจากคลาส LinearRegression
คำอธิบาย	การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยวิธี เอนเตอร์
ตัวแปร	ตัวแปรที่สืบทอดจากคลาส LinearRegression
เมทธอด	▪ Enter(double dependentVar[], double independentVar[][], String[] independentLabel, boolean resInd) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* dependentVar หมายถึง ค่าตัวแปรตาม independentVar หมายถึง ค่าตัวแปรอิสระ independentLabel หมายถึง ชื่อตัวแปรอิสระ resInd หมายถึง ตัวบ่งชี้ว่าต้องการคำนวณค่าความคลาดเคลื่อนหรือไม่ มีค่าจริงหากต้องการคำนวณค่าและมีค่าเท็จหากไม่ต้องการคำนวณค่า ▪ varSelection() เมทธอดนี้ใช้สำหรับเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยใช้วิธี เอนเตอร์

3.2.2.3 คลาส Backward

คลาส Backward เป็นคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีแบ็คเวิร์ด ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.17 จากรูปที่ 3.17 คลาส Backward ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.13

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 3.17 ส่วนประกอบของคลาส Backward

ตารางที่ 3.13 รายละเอียด ของคลาส Backward

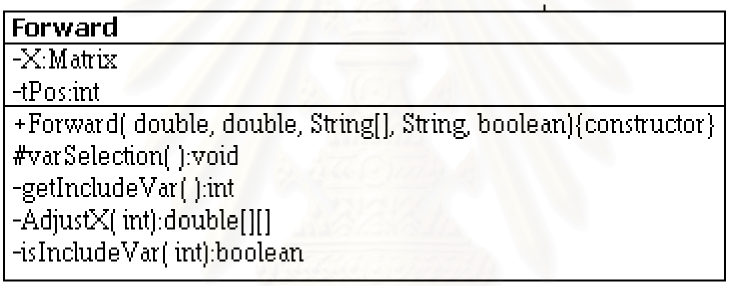
ชื่อคลาส	**Backward**
ประเภท	สืบทอด จาก คลาส VarSelection
คำอธิบาย	การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยวิธีแบ็คเวิร์ด
ตัวแปร	■ ตัวแปรที่สืบทอดจากคลาส LinearRegression ■ X เมตริกซ์ค่าตัวแปรอิสระเมื่อมีการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้นโดยวิธีแบ็คเวิร์ด ■ enter การคำนวณค่าสถิติความถดถอยเชิงเส้น โดยเทคนิคการเลือกตัวแปรอิสระเข้าสมการความถดถอยแบบเอนเตอร์
เมทธอด	■ Backward(double dependentVar[], double independentVar[][], independentItems, boolean resInd) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* dependentVar หมายถึง ค่าตัวแปรตาม independentVar หมายถึง ค่าตัวแปรอิสระ independentItems หมายถึงชื่อตัวแปรอิสระ resInd หมายถึง ตัวบ่งชี้ว่าต้องการคำนวณค่าความคลาดเคลื่อนหรือไม่ มีค่าจริงหากต้องการคำนวณค่าและมีค่าเท็จหากไม่ต้องการคำนวณค่า ■ varSelection() เมทธอดนี้ใช้สำหรับเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยใช้วิธีแบ็คเวิร์ด ■ getExcludeVar(int modelNum) เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ต้องการตัดออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าตัวแปรที่ต้องการตัดออกจากสมการความถดถอยเชิงเส้น

ตารางที่ 3.13 รายละเอียด ของคลาส Backward (ต่อ)

เมทธอด	▪ *พารามิเตอร์ที่ใช้* modelNum หมายถึง ลำดับของรูปแบบสมการความถดถอยเชิงเส้น ▪ AdjustX(int excludeVar) เมทธอดนี้ใช้สำหรับปรับเมตริกซ์ค่าตัวแปรอิสระ *ค่าส่งกลับ* เมตริกซ์ของตัวแปรอิสระที่ปรับค่าหลังจากการนำตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ *พารามิเตอร์ที่ใช้* excludeVar หมายถึง ตัวแปรที่ตัดออกจากสมการความถดถอยเชิงเส้น

3.2.2.4 คลาส Forward

คลาส Forward เป็นคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.18



รูปที่ 3.18 ส่วนประกอบของคลาส Forward

จากรูปที่ 3.18 คลาส Forward ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.14

ตารางที่ 3.14 รายละเอียด ของคลาส Forward

ชื่อคลาส	**Forward**
ประเภท	สืบทอดจากคลาส LinearRegression
คำอธิบาย	การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด
ตัวแปร	▪ ตัวแปรที่สืบทอดจากคลาส LinearRegression ▪ X เมตริกซ์ค่าตัวแปรอิสระเมื่อมีการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยวิธีฟอร์เวิร์ด

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.14 รายละเอียด ของคลาส Forward (ต่อ)

เมทธอด	▪ Forward(double dependentVar[],double independentVar[][], String [] independentItems, boolean resind) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* dependentVar หมายถึง ค่าตัวแปรตาม independentVar หมายถึง ค่าตัวแปรอิสระ independentItems หมายถึง ชื่อตัวแปรอิสระ resInd หมายถึง ตัวบ่งชี้ว่าต้องการคำนวณค่าความคลาดเคลื่อนหรือไม่ มีค่าจริงหากต้องการคำนวณค่าและมีค่าเท็จหากไม่ต้องการคำนวณค่า ▪ VarSelection()เมทธอดนี้ใช้สำหรับเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยใช้วิธี ฟอร์เวิร์ด ▪ GetIncludeVar() เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ▪ *ค่าส่งกลับ* ค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ▪ GetIncludeVar() เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ▪ *ค่าส่งกลับ* ค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ▪ AdjustX(int includeVar) เมทธอดนี้ใช้สำหรับปรับเมตริกซ์ค่าตัวแปรอิสระ *ค่าส่งกลับ* เมตริกซ์ของตัวแปรอิสระที่ปรับค่าหลังจากการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* includeVar หมายถึง ค่าตัวแปรที่นำเข้าสมการความถดถอยเชิงเส้น ▪ isIncludeVar(int index) เมทธอดนี้ใช้ตรวจสอบว่าตัวแปรอิสระถูกนำเข้าสมการความถดถอยหรือไม่ ▪ *ค่าส่งกลับ* ค่าจริงถ้าตัวแปรนั้นอยู่ในสมการความถดถอย หรือ ค่าเท็จถ้าตัวแปรนั้นไม่อยู่ในสมการความถดถอย ▪ *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของตัวแปรอิสระ

3.2.2.5 คลาส Stepwise

คลาส Stepwise เป็นคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีสเตปไวส์ ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.19

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Stepwise
-X:Matrix -tPos:int
+Stepwise(double, double, String[], String, boolean){constructor} #varSelection( ):void -getExcludeVar(int):int -getIncludeVar( ):int -backwardAdjustX(int):double[][] -forwardAdjustX(int):double[][] -isIncludeVar(int):boolean

รูปที่ 3.19 ส่วนประกอบของคลาส Stepwise

จากรูปที่ 3.19 คลาส Stepwise ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงใน ตารางที่ 3.15

ตารางที่ 3.15 รายละเอียด ของคลาส Stepwise

ชื่อคลาส	**Stepwise**
ประเภท	สืบทอดจากคลาส LinearRegression
คำอธิบาย	การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธี สเตปไวส์
ตัวแปร	▪ ตัวแปรที่สืบทอดจากคลาส LinearRegression ▪ X เมตริกซ์ค่าตัวแปรอิสระเมื่อมีการเลือกและตัดตัวแปรอิสระเข้าและออกจากสมการความถดถอยเชิงเส้นโดยวิธี สเตปไวส์
เมทธอด	▪ Stepwise(double dependentVar[],double independentVar[][], String [] independentItems, boolean resInd) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ ▪ *พารามิเตอร์ที่ใช้* dependentVar หมายถึง ค่าตัวแปรตาม independentVar หมายถึง ค่าตัวแปรอิสระ independentItems หมายถึง ชื่อตัวแปรอิสระ resInd หมายถึง ตัวบ่งชี้ว่าต้องการคำนวณค่าความคลาดเคลื่อนหรือไม่ มีค่าจริงหากต้องการคำนวณค่าและมีค่าเท็จหากไม่ต้องการคำนวณค่า ▪ varSelection()เมทธอดนี้ใช้สำหรับเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นโดยใช้วิธี สเตปไวส์ getExcludeVar(int modelNum) เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ต้องการตัดออกจากสมการความถดถอยเชิงเส้น

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.15 รายละเอียด ของคลาส Stepwise (ต่อ)

เมทธอด	*ค่าส่งกลับ* ค่าตัวแปรที่ต้องการตัดออกจากสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* modelNum หมายถึง ลำดับของรูปแบบสมการความถดถอยเชิงเส้น ■ getIncludeVar() เมทธอดนี้ใช้สำหรับรับค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอย *ค่าส่งกลับ* ค่าตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ■ backwardAdjustX(int excludeVar) เมทธอดนี้ใช้สำหรับปรับเมตริกซ์ค่าตัวแปรอิสระหลังการตัดตัวแปรออกจากสมการความถดถอยเชิงเส้นโดยวิธีแบ็คเวิร์ด *ค่าส่งกลับ* เมตริกซ์ค่าตัวแปรอิสระหลังการตัดตัวแปรออกจากสมการความถดถอยเชิงเส้นโดยวิธีแบ็คเวิร์ด *พารามิเตอร์ที่ใช้* excludeVar หมายถึง ค่าตัวแปรที่ต้องการตัดออกจากสมการความถดถอยเชิงเส้น ■ forwardAdjustX(int includeVar) เมทธอดนี้ใช้สำหรับปรับเมตริกซ์ค่าตัวแปรอิสระอิสระหลังการนำตัวแปรเข้าสมการความถดถอยเชิงเส้นโดยวิธีฟอร์เวิร์ด *ค่าส่งกลับ* เมตริกซ์ค่าตัวแปรอิสระอิสระหลังการนำตัวแปรเข้าสมการความถดถอยเชิงเส้นโดยวิธีฟอร์เวิร์ด *พารามิเตอร์ที่ใช้* includeVar หมายถึง ลำดับที่ของตัวแปรที่ต้องการนำเข้าสมการความถดถอยเชิงเส้น ■ isIncludeVar(int index) เมทธอดนี้ใช้ตรวจสอบว่าตัวแปรอิสระถูกนำเข้าสมการความถดถอยหรือไม่ ■ *ค่าส่งกลับ* ค่าจริงถ้าตัวแปรนั้นอยู่ในสมการความถดถอย หรือ ค่าเท็จถ้าตัวแปรนั้นไม่อยู่ในสมการความถดถอย ■ *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของตัวแปรอิสระ

3.2.2.6 คลาส Model Summary

คลาส ModelSummary เป็นคลาสสำหรับการคำนวณสถิติสรุปของสมการความถดถอยเชิงเส้นประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.20 จากรูปที่ 3.20 คลาส ModelSummary ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงใน ตารางที่ 3.16

ModelSummary
-X:Matrix -Y:Matrix -r:double -rSqr:double -adjustRSqr:double -stdError:double -rSqrChange:double -fChange:double -df1:int -df2:int -fChangeSign:double -durbinWatson:double -anova:Anova -statTable:StatTable
+ModelSummary( ){constructor} +ModelSummary(Matrix, Matrix, Anova, double, int){constructor} +ModelSummary(Matrix, Matrix, Anova, double, int, int){constructor} -setRSquare( ):void -setR( ):void -setAdjustRSquare( ):void -setStdError( ):void -setDf2( ):void -setRSqrChange(double):void -setFChange( ):void -setFChangeSign( ):void -setDurbinWatson( ):void +getR( ):double +getRSqr( ):double +getAdjustRSqr( ):double +getStdError( ):double +getRSquareChange( ):double +getFChange( ):double +getFChangeSign( ):double +getDurbinWatson( ):double +getDf1( ):int +getDf2( ):int

รูปที่ 3.20 ส่วนประกอบของคลาส ModelSummary

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.16 รายละเอียด ของคลาส ModelSummary

ชื่อคลาส	**ModelSummary**
ประเภท	คลาสย่อยของ คลาส LinearRegression
คำอธิบาย	สถิติสรุปของสมการความถดถอยเชิงเส้น(Model Summary)
ตัวแปร	▪ X เมตริกซ์ค่าตัวแปรอิสระ ▪ Y เมตริกซ์ค่าตัวแปรตาม ▪ r ค่าสัมประสิทธิ์สหสัมพันธ์

ตารางที่ 3.16 รายละเอียด ของคลาส ModelSummary (ต่อ)

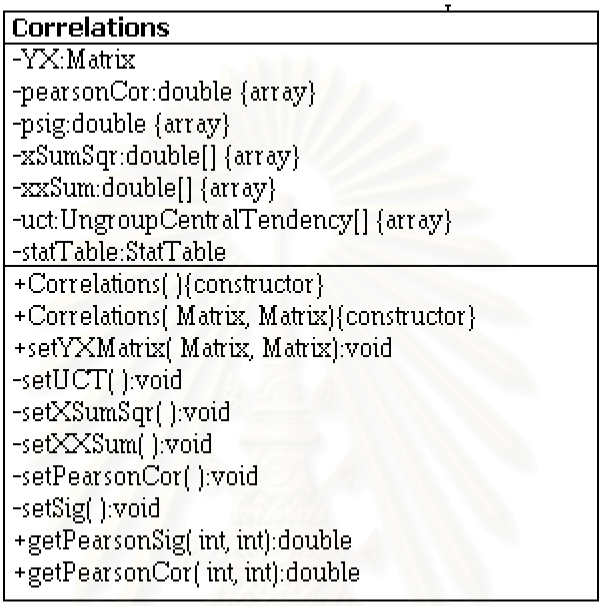
ตัวแปร	▪ rSqr สัมประสิทธิ์การตัดสินใจ (R Square) ▪ adjRSqr ค่าสัมประสิทธิ์การตัดสินใจที่ปรับแล้ว(Adjusted R Square) ▪ stdError ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า(Standard Error) ▪ rSqrChange ค่าสัมประสิทธิ์การตัดสินใจที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการ ▪ fChange ค่าสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการ ▪ df1 องศาอิสระ 1(จำนวนตัวแปรอิสระ) ▪ df2 องศาอิสระ 2(จำนวนของปัจจัยอื่น) ▪ fChangeSign ค่านัยสำคัญของสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการ ▪ anova ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น ▪ durbinWatson ค่าสถิติ Durbin-Watson ▪ statTable ค่าตารางทางสถิติ
เมทธอด	▪ ModelSummary(Matrix Y, Matrix X, Anova anova) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ Anova หมายถึง ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น ▪ ModelSummary(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ▪ setRSquare() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์การตัดสินใจ ▪ setR() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ ▪ setAdjustRSquare() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธ์การตัดสินใจที่ปรับแล้ว ▪ setStdError() เมทธอดนี้ใช้สำหรับคำนวณค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า ▪ setDf() เมทธอดนี้ใช้สำหรับ กำหนดค่าองศาอิสระ ▪ setRSqrChange() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธิ์การตัดสินใจที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ setFChange() เมทธอดนี้ใช้สำหรับคำนวณค่าสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น

ตารางที่ 3.16 (ต่อ)

เมทธอด	▪ setFChangeSign() เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ setDurbinWatson() เมทธอดนี้ใช้สำหรับรับค่าสถิติ Durbin-Watson ▪ getR() เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธิ์สหสัมพันธ์ *ค่าส่งกลับ* ค่าสัมประสิทธิ์สหสัมพันธ์ ▪ getRSqr() เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธิ์การตัดสินใจ *ค่าส่งกลับ* ค่าสัมประสิทธิ์การตัดสินใจ ▪ getAdjustRSqr() เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธ์การตัดสินใจที่ปรับแล้ว *ค่าส่งกลับ* ค่าสัมประสิทธิ์การตัดสินใจที่ปรับแล้ว ▪ getStdError() เมทธอดนี้ใช้สำหรับรับค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า *ค่าส่งกลับ* ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า ▪ getRSquareChange() เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธิ์การตัดสินใจที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสัมประสิทธิ์การตัดสินใจที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ getFChange() เมทธอดนี้ใช้สำหรับรับค่าสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ getFChangeSign() เมทธอดนี้ใช้สำหรับรับค่านัยสำคัญของสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่านัยสำคัญของสถิติ F ที่เปลี่ยนไปเมื่อนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ getDurbinWatson() เมทธอดนี้ใช้สำหรับรับค่าสถิติ Durbin-Watson *ค่าส่งกลับ* ค่าสถิติ Durbin-Watson ▪ getDf1() เมทธอดนี้ใช้สำหรับรับค่าองศาอิสระ 1 *ค่าส่งกลับ* ค่าองศาอิสระ 1 ▪ getDf2() เมทธอดนี้ใช้สำหรับรับค่าองศาอิสระ 2 *ค่าส่งกลับ* ค่าองศาอิสระ 1

3.2.2.7 คลาส Correlations

คลาส Correlations เป็นคลาสสำหรับการทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.21 จากรูปที่ 3.21 คลาส Correlations ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.17



รูปที่ 3.21 ส่วนประกอบของคลาส Correlations

ตารางที่ 3.17 รายละเอียด ของคลาส Correlations

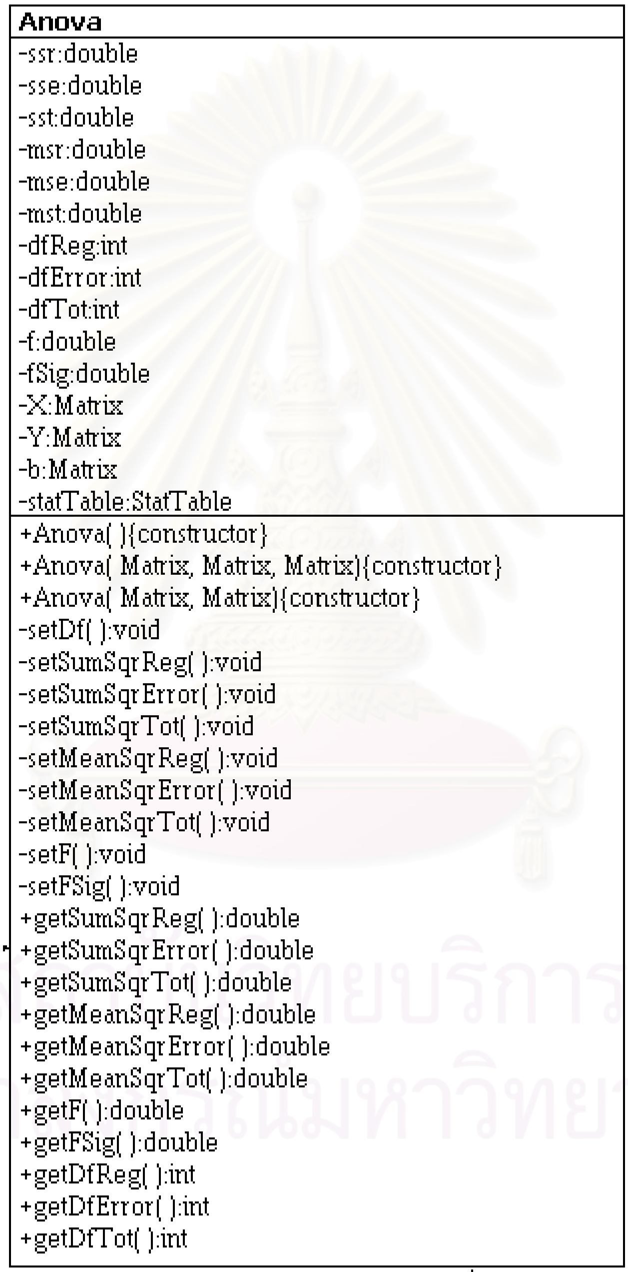
ชื่อคลาส	**Correlations**
ประเภท	คลาสย่อยของคลาส LinearRegression
คำอธิบาย	การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว
ตัวแปร	▪ YX เมตริกซ์ค่าตัวแปรตามและตัวแปรอิสระ ▪ pearsonCor สัมประสิทธ์สหสัมพันธ์ระหว่างตัวแปรอิสระ(Pearson Correlation) ▪ pSig ค่านัยสำคัญของสัมประสิทธ์สหสัมพันธ์ระหว่างตัวแปรอิสระ ▪ xSumSqr ค่าผลบวกกำลังสองของตัวแปรอิสระ ▪ xxSum ค่าผลบวกของเมตริกซ์ผลคูณ YX กับ YX ▪ uct ค่าสถิติวัดค่ากลางของข้อมูลที่ไม่ได้จัดกลุ่ม ▪ statTable ค่าตารางทางสถิติ

ตารางที่ 3.17 รายละเอียด ของคลาส ModelSummary (ต่อ)

เมทธอด	▪ Correlations() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์ (Default Constructor) ▪ Correlations(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ▪ setYXMatrix(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับ กำหนดค่าของเมตริกซ์ค่าตัวแปรตามและตัวแปรอิสระ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ▪ setUCT() เมทธอดนี้ใช้สำหรับ กำหนดค่าสถิติวัดค่ากลางของข้อมูลที่ไม่ได้จัดกลุ่ม ▪ setXSumSqr() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของตัวแปรอิสระ ▪ setXXSum() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกของเมตริกซ์ผลคูณ YX กับ YX ▪ setPearsonCor() เมทธอดนี้ใช้สำหรับคำนวณค่าสัมประสิทธ์สหสัมพันธ์ระหว่างตัวแปรอิสระ ▪ setSig() เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ ▪ setSig() เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ ▪ getPearsonCor(int index1, int index2) เมทธอดนี้ใช้สำหรับรับค่าสัมประสิทธ์สหสัมพันธ์ระหว่างตัวแปรอิสระ *ค่าส่งกลับ* ค่าสัมประสิทธิ์สัมพันธ์ระหว่างตัวแปรอิสระสองตัวที่ระบุ *พารามิเตอร์ที่ใช้* index1 หมายถึง ลำดับที่ของตัวแปรตัวที่ 1 index2 หมายถึง ลำดับที่ของตัวแปรตัวที่ 2 ▪ getPearsonSig(int index1, int index2) เมทธอดนี้ใช้สำหรับรับค่านัยสำคัญของสัมประสิทธ์สหสัมพันธ์ระหว่างตัวแปรอิสระ *ค่าส่งกลับ* ค่านัยสำคัญของสัมประสิทธิ์สัมพันธ์ระหว่างตัวแปรอิสระสองตัวที่ระบุ *พารามิเตอร์ที่ใช้* index1 หมายถึง ลำดับที่ของตัวแปรตัวที่ 1 index2 หมายถึง ลำดับที่ของตัวแปรตัวที่ 2

3.2.2.8 คลาส Anova

คลาส Anova เป็นคลาสสำหรับการวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.22 จากรูปที่ 3.22 คลาส Anova ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.18



รูปที่ 3.22 ส่วนประกอบของคลาส Anova

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.18 รายละเอียด ของคลาส Anova

ชื่อคลาส	**ANOVA**
ประเภท	คลาสย่อย ของ คลาส LinearRegression
คำอธิบาย	การวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น
ตัวแปร	▪ ssr ค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตามเนื่องจากอิทธิพลของ ▪ ตัวแปรอิสระ ▪ sse ค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตามเนื่องจากอิทธิพลอื่นๆ ▪ sst ค่าผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของตัวแปรตาม ▪ msr ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ ▪ mse ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนตัวแปรตามเนื่องจากอิทธิพลอื่นๆ ▪ mst ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของตัวแปรตาม ▪ dfReg ค่าองศาอิสระของตัวแปรอิสระ ▪ dfError ค่าองศาอิสระของค่าแปรปรวนสุ่ม ▪ dfTot ค่าองศาอิสระรวม ▪ f ค่าสถิติทดสอบ F ▪ fSign ค่าระดับนัยสำคัญของสถิติทดสอบ F ▪ X เมตริกซ์ค่าตัวแปรอิสระ ▪ Y เมตริกซ์ค่าตัวแปรตาม ▪ b เมตริกซ์ค่าคงที่และสัมประสิทธิ์ความถดถอย ▪ statTable ค่าตารางทางสถิติ
เมทธอด	▪ Anova() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์(Default Constructor) ▪ Anova(Matrix Y, Matrix X, Matrix b) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ b หมายถึง เมตริกซ์ค่าสัมประสิทธิ์ความถดถอยเชิงเส้น ▪ public Anova(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์

ตารางที่ 3.18 รายละเอียด ของคลาส Anova (ต่อ)

เมทธอด	*พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ▪ setDf() เมทธอดนี้ใช้สำหรับ กำหนดค่าองศาอิสระ ▪ setSumSqrReg() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลของตัวแปรอิสระ ▪ setSumSqrError() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลอื่นๆ ▪ setSumSqrTot() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของตัวแปรตาม ▪ setMeanSqrReg() เมทธอดนี้ใช้สำหรับคำนวณเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ ▪ setMeanSqrError() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลอื่นๆ ▪ setMeanSqrTot() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของ ตัวแปรตาม ▪ setF() เมทธอดนี้ใช้สำหรับคำนวณค่าสถิติทดสอบ F ▪ setFSig() เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสถิติทดสอบ F ▪ getSumSqrReg() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลของตัวแปรอิสระ *ค่าส่งกลับ* ค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลของตัวแปรอิสระ ▪ getSumSqrError() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลอื่นๆ *ค่าส่งกลับ* ค่าผลบวกกำลังสองของค่าแปรปรวนตัวแปรตาม เนื่องจากอิทธิพลอื่นๆ ▪ getSumSqrTot() เมทธอดนี้ใช้สำหรับคำนวณค่าผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของตัวแปรตาม *ค่าส่งกลับ* ค่าผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของตัวแปรตาม ▪ getMeanSqrReg() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ

ตารางที่ 3.18 รายละเอียด ของคลาส Anova (ต่อ)

เมทธอด	*ค่าส่งกลับ* ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ ▪ getMeanSqrError() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลอื่นๆ *ค่าส่งกลับ* ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลอื่นๆ ▪ getMeanSqrTot() เมทธอดนี้ใช้สำหรับคำนวณค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของ ตัวแปรตาม *ค่าส่งกลับ* ค่าเฉลี่ยของผลบวกกำลังสองของค่าแปรปรวนทั้งหมดของ ตัวแปรตาม ▪ getF() เมทธอดนี้ใช้สำหรับรับค่าสถิติทดสอบ F *ค่าส่งกลับ* ค่าสถิติทดสอบ F ▪ getFSig() เมทธอดนี้ใช้สำหรับรับค่านัยสำคัญของสถิติทดสอบ F *ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ F ▪ getDfReg() เมทธอดนี้ใช้สำหรับรับค่าองศาอิสระของตัวแปรอิสระ *ค่าส่งกลับ* ค่าองศาอิสระของตัวแปรอิสระ ▪ getDfError() เมทธอดนี้ใช้สำหรับรับค่าองศาอิสระของอิทธิพลอื่นๆ *ค่าส่งกลับ* ค่าองศาอิสระของอิทธิพลอื่นๆ ▪ getDfTot() เมทธอดนี้ใช้สำหรับรับค่าองศาอิสระของตัวแปรตาม *ค่าส่งกลับ* ค่าองศาอิสระของตัวแปรตาม

3.2.2.9 คลาส Coefficients

คลาส Coefficient เป็นคลาสสำหรับการประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้นประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.23 จากรูปที่ 3.23 คลาส Coefficients ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.19

สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

Coefficients
-b:Matrix -X:Matrix -Y:Matrix -e:Matrix -bLabels:String[] {array} -Ye:Matrix -stdDevReg:double -stdDevRegSqr:double -stdError:double[] {array} -bUpper:double[] {array} -bLower:double[] {array} -t:double[] {array} -tSig:double[] {array} -statTable:StatTable
+Coefficients( ){constructor} +Coefficients(Matrix, Matrix){constructor} +Coefficients(Matrix, Matrix, String[]){constructor} -setCoefficients( ):void -setstdDevRegSqr( ):void -setstdDevReg( ):void -setStdError( ):void -setConfInterval( ):void -setT( ):void -setTSig( ):void +getCoefficientsMatrix( ):Matrix +getEstimateY( ):Matrix +getCoefficients(int):double +getEstimateResidual( ):Matrix +getStdError(int):double +getUpperBound(int):double +getLowerBound(int):double +getT(int):double +getTSig(int):double +getBLabels(int):String

รูปที่ 3.23 ส่วนประกอบของคลาส Coefficients

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.19 รายละเอียด ของคลาส Coefficients

ชื่อคลาส	**Coefficients**
ประเภท	คลาสย่อยของคลาส LinearRegression
คำอธิบาย	การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น
ตัวแปร	■ b เมตริกซ์ค่าคงที่และสัมประสิทธิ์ความถดถอย ■ X เมตริกซ์ค่าตัวแปรอิสระ ■ Y เมตริกซ์ค่าตัวแปรตาม ■ e เมตริกซ์ค่าประมาณความผิดพลาด ■ bLabels ชื่อตัวแปรอิสะ ■ Ye เมตริกซ์ค่าประมาณของตัวแปรตาม ■ stdDevReg ค่าแปรปรวนมาตรฐานของสมการความถดถอยเชิงเส้น ■ stdDevRegSqr ค่ากำลังสองของค่าแปรปรวนมาตรฐานของสมการความถดถอยเชิงเส้น ■ stdError[ ] ค่าคลาดเคลื่อนมาตรฐานของค่าคงที่และสัมประสิทธิ์ความถดถอย ■ beta[ ] สัมประสิทธิ์ความถดถอยมาตรฐาน ■ bUpper ค่าขอบเขตบนของค่าความเชื่อมั่น ■ bLower ค่าขอบเขตล่างของค่าความเชื่อมั่น ■ t สถิติทดสอบ t ■ tSig ระดับนัยสำคัญของสถิติทดสอบ t
เมทธอด	■ Coefficients() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์(Default Constructor) ■ Coefficients(Matrix Y, Matrix X) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ ■ Coefficients(Matrix Y, Matrix X, String [] bLabels) เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ bLabels หมายถึง ชื่อของค่าคงที่และสัมประสิทธิ์ความถดถอยเชิงเส้น
เมทธอด	■ setCoefficients() เมทธอดนี้ใช้สำหรับคำนวณค่าประมาณสัมประสิทธิ์ความถดถอย

ตารางที่ 3.19 รายละเอียด ของคลาส Coefficients (ต่อ)

เมทธอด	■ setstdDevRegSqr() เมทธอดนี้ใช้สำหรับคำนวณค่ากำลังสองของค่าเบี่ยงเบนมาตรฐานของสมการความถดถอย ■ setstdDevReg() เมทธอดนี้ใช้สำหรับคำนวณค่าค่าเบี่ยงเบนมาตรฐานของสมการความถดถอย ■ setStdError() เมทธอดนี้ใช้สำหรับคำนวณค่าคลาดเคลื่อนมาตรฐานของค่าคงที่และสัมประสิทธิ์ความถดถอย ■ setConfInterval() เมทธอดนี้ใช้สำหรับคำนวณค่าของช่วงความเชื่อมั่น ■ setT() เมทธอดนี้ใช้สำหรับคำนวณค่าสถิติทดสอบ T ■ setTSig() คำนวณค่านัยสำคัญของสถิติทดสอบ T ■ getCoefficientsMatrix() เมทธอดนี้ใช้สำหรับรับเมตริกซ์ค่าประมาณสัมประสิทธิ์ความถดถอย ■ *ค่าส่งกลับ* เมตริกซ์ค่าประมาณสัมประสิทธิ์ความถดถอย ■ getCoefficients(int index) *ค่าส่งกลับ* ค่าประมาณสัมประสิทธิ์ความถดถอย *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ public Matrix getEstimateY() เมทธอดนี้ใช้สำหรับรับเมตริกซ์ค่าประมาณของตัวแปรตาม *ค่าส่งกลับ* เมตริกซ์ค่าประมาณของตัวแปรตาม ■ public Matrix getEstimateResidual() เมทธอดนี้ใช้สำหรับรับเมตริกซ์ค่าประมาณความผิดพลาด *ค่าส่งกลับ* เมตริกซ์ค่าประมาณความผิดพลาด ■ public double getStdError(int index) เมทธอดนี้ใช้สำหรับรับค่าคลาดเคลื่อนมาตรฐานของค่าคงที่และสัมประสิทธิ์ความถดถอย *ค่าส่งกลับ* ค่าคลาดเคลื่อนมาตรฐานของค่าคงที่และสัมประสิทธิ์ความถดถอย ■ *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ค่าคลาดเคลื่อนมาตรฐานของสัมประสิทธิ์ความถดถอย ■ public double getUpperBound(int index) เมทธอดนี้ใช้สำหรับรับค่าขอบเขตบนของค่าความเชื่อมั่น

ตารางที่ 3.19 รายละเอียด ของคลาส Coefficients (ต่อ)

เมทธอด	*ค่าส่งกลับ* ค่าขอบเขตบนของค่าความเชื่อมั่น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ public double getUpperBound(int index) เมทธอดนี้ใช้สำหรับรับค่าขอบเขตบนของค่าความเชื่อมั่น *ค่าส่งกลับ* ค่าขอบเขตบนของค่าความเชื่อมั่น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ public double getLowerBound(int index) เมทธอดนี้ใช้สำหรับรับค่าขอบเขตล่างของค่าความเชื่อมั่น *ค่าส่งกลับ* ค่าขอบเขตล่างของค่าความเชื่อมั่น *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ getStdDevRegSqr() เมทธอดนี้ใช้สำหรับรับค่ากำลังสองของค่าแปรปรวนมาตรฐาน *ค่าส่งกลับ* ค่ากำลังสองของค่าแปรปรวนมาตรฐาน ■ getT(int index) เมทธอดนี้ใช้สำหรับรับค่าสถิติทดสอบ T *ค่าส่งกลับ* ค่าสถิติทดสอบ T *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ public double getTSig(int index) เมทธอดนี้ใช้สำหรับรับค่านัยสำคัญของสถิติทดสอบ T *ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ T *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย ■ getBLabels(int index) เมทธอดนี้ใช้สำหรับรับค่าชื่อตัวแปรอิสะ *ค่าส่งกลับ* ชื่อตัวแปรอิสะ *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของค่าประมาณสัมประสิทธิ์ความถดถอย

3.2.2.10 คลาส ResidualsStatistics

คลาส ResidualsStatistics เป็นคลาสสำหรับการวิเคราะห์ค่าสถิติของค่าคลาดเคลื่อนประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.24 และจากรูปที่ 3.24 คลาส ResidualStatistics ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.20

ResidualsStatistics
-X:Matrix -Y:Matrix -H:Matrix -coefficients:Coefficients -modelSummary:ModelSummary -dependent:double {array} -zpred:double {array} -zresid:double {array} -dresid:double {array} -sdresid:double {array} -sresid:double {array} -adjpred:double {array}
+ResidualsStatistics(Matrix, Matrix, Coefficients, ModelSummary){constructor} +ResidualsStatistics( ){constructor} #setResidual( ):void -delObservation(int):void -setDresid(Matrix, int):void -setAdjPred(Matrix, int):void -setSdresid(Matrix, int):void -setDependentVariable( ):void -setStandardPredictValues( ):void -setStandardResiduals( ):void -setStudentResidual( ):void +getResiduals(int):double[] +getResidualsType(int):String

รูปที่ 3.24 ส่วนประกอบของคลาส ResidualStatistics

ตารางที่ 3.20 รายละเอียด ของคลาส ResidualStatistics

ชื่อคลาส	**ResidualsStatistics**
ประเภท	คลาสย่อยของคลาส LinearRegression
คำอธิบาย	การวิเคราะห์ค่าสถิติของค่าคลาดเคลื่อน
ตัวแปร	▪ X เมตริกซ์ค่าตัวแปรอิสระ ▪ Y เมตริกซ์ค่าตัวแปรตาม ▪ H Leverage Matrix ▪ coefficients ค่าประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น ▪ modelSummary สถิติสรุปของสมการความถดถอยเชิงเส้น ▪ dependent ค่าตัวแปรตาม ▪ zpred ค่า Standardized Predicted ▪ zresid ค่า Standardized Residual

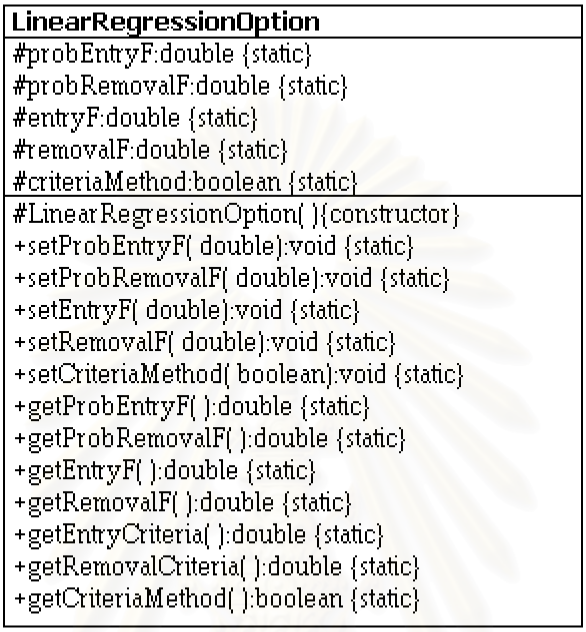
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ตารางที่ 3.20 รายละเอียด ของคลาส ResidualStatistics (ต่อ)

ตัวแปร	■ dresid ค่า Deleted Residual ■ sdresid ค่า Studentdized Residual ■ sresid ค่า Studentdized Delete Residual ■ adjpred ค่า Adjusted Residual
เมทธอด	■ ResidualsStatistics() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์(Default Constructor) ■ ResidualsStatistics(Matrix Y, Matrix X, Coefficients cf, ModelSummary md) ■ เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยระบุพารามิเตอร์ *พารามิเตอร์ที่ใช้* Y หมายถึง เมตริกซ์ค่าตัวแปรตาม X หมายถึง เมตริกซ์ค่าตัวแปรอิสระ Cf หมายถึง ค่าประมาณสัมประสิทธิ์ความถดถอย Md หมายถึง ค่าสถิติสรุปของสมการความถดถอยเชิงเส้น ■ setResidual() เมทธอดนี้ใช้สำหรับคำนวณค่าสถิติของค่าคลาดเคลื่อน ■ delObservation(int index) *พารามิเตอร์ที่ใช้* index หมายถึง ลำดับที่ของแถว ■ setDependentVariable() เมทธอดนี้ใช้สำหรับกำหนดค่าตัวแปรตาม ■ setStandardPredictValues() เมทธอดนี้ใช้สำหรับคำนวณค่า Standardized Predicted Value ■ setStandardResiduals() เมทธอดนี้ใช้สำหรับคำนวณค่า Standardized Residual ■ setStudentResidual() เมทธอดนี้ใช้สำหรับคำนวณค่า Studentdized Residual ■ getResiduals(int index) เมทธอดนี้ใช้สำหรับรับค่าสถิติของค่าคลาดเคลื่อน *ค่าส่งกลับ* ค่าสถิติของค่าคลาดเคลื่อน *พารามิเตอร์ที่ใช้* index หมายถึง ค่า 0-6 ที่ระบุเพื่อรับค่าความคลาดเคลื่อนโดยมีความหมายดังนี้ - ค่า 0 หมายถึง ค่า DEPENDENT - ค่า 1 หมายถึง ค่า ZPRED - ค่า 2 หมายถึง ค่า ZRESID - ค่า 3 หมายถึง ค่า DRESID - ค่า 4 หมายถึง ค่า ADJPRED - ค่า 5 หมายถึง ค่า SRESID - ค่า 6 หมายถึง ค่า SDRESID

3.2.2.11 คลาส LinearRegressionOption

คลาส LinearRegressionOption เป็นคลาสสำหรับกำหนดเงื่อนไขในการเลือกตัวแปรอิสระเข้าสมการความถดถอยประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.25 และจากรูปที่ 3.25 คลาส LinearRegressionOption ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.21



รูปที่ 3.25 ส่วนประกอบของคลาส LinearRegressionOption

ตารางที่ 3.21 รายละเอียด ของคลาส LinearRegressionOption

ชื่อคลาส	**LinearRegressionOption**
ประเภท	Base Class
คำอธิบาย	การกำหนดเงื่อนไขในการเลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น
ตัวแปร	▪ probEntryF ค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ probRemovalF ค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ entryF ค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ removalF ค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ criteriorMethod วิธีการที่เลือกใช้เป็นเงื่อนไขในการเลือกตัวแปรอิสระเข้า/ออกจากสมการความถดถอยเชิงเส้น

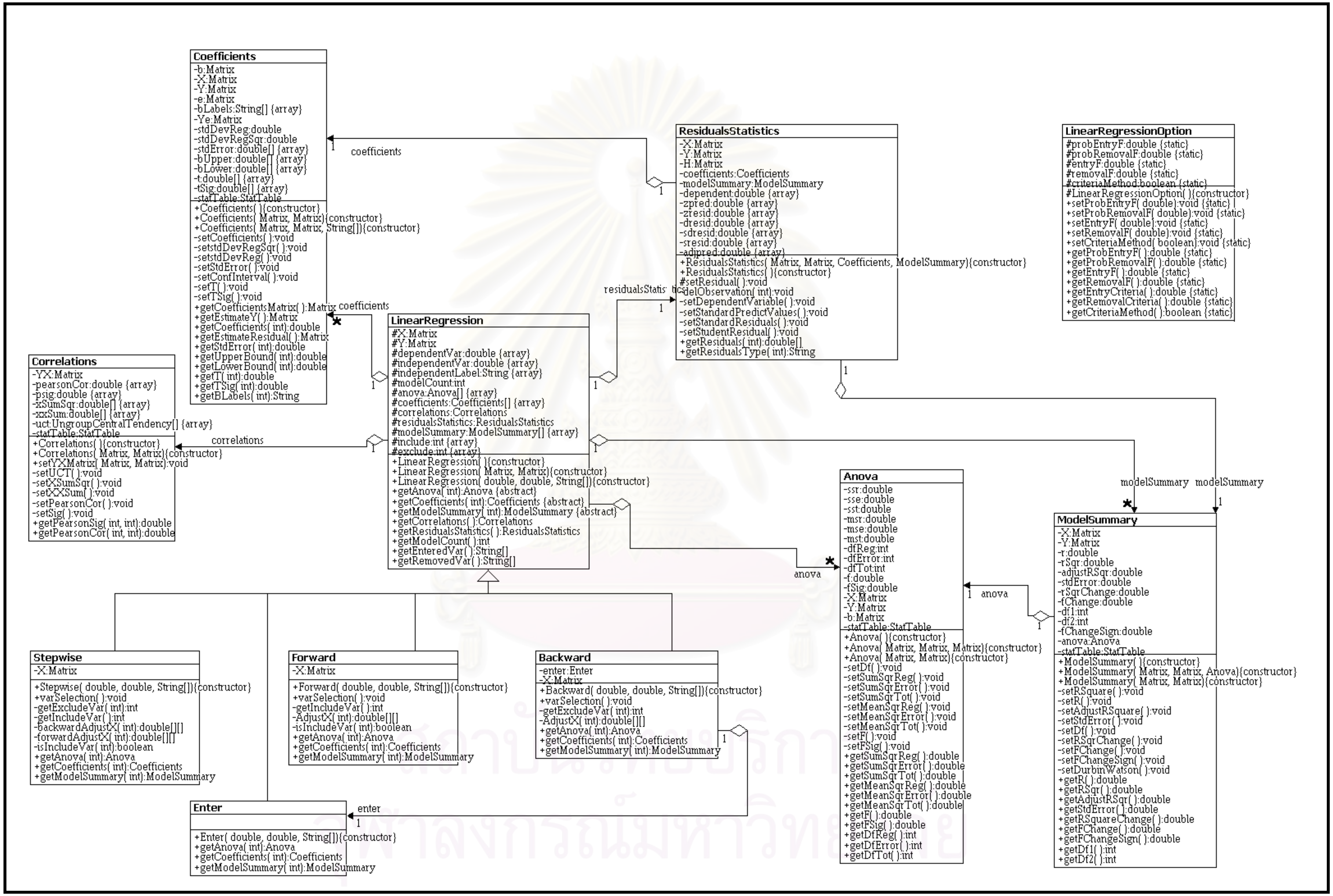
ตารางที่ 3.21 รายละเอียด ของคลาส LinearRegressionOption (ต่อ)

เมทธอด	▪ LinearRegressionOption() เมทธอดนี้ใช้สำหรับการสร้างคลาสโดยไม่ระบุพารามิเตอร์ (Default Constructor) ▪ setProbEntryF(double probEntry) เมทธอดนี้ใช้สำหรับ กำหนดค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* probEntry หมายถึง ค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ setProbRemovalF(double probRemoval) เมทธอดนี้ใช้สำหรับ กำหนดค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* ▪ probRemoval หมายถึง ค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ setEntryF(double entry) เมทธอดนี้ใช้สำหรับกำหนดค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* entry หมายถึง ค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ setRemovalF(double removal) เมทธอดนี้ใช้สำหรับกำหนดค่าสถิติ F ที่ใชเป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* removal หมายถึง ค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ setCriteriaMethod(boolean method) เมทธอดนี้ใช้สำหรับกำหนดวิธีการที่เลือกใช้เป็นเงื่อนไขในการเลือกตัวแปรอิสระเข้า/ออกจากสมการความถดถอยเชิงเส้น *พารามิเตอร์ที่ใช้* method หมายถึง วิธีการที่เลือกใช้เป็นเงื่อนไขในการเลือกตัวแปรอิสระเข้า/ออกจากสมการความถดถอยเชิงเส้น มีค่าเป็นจริงเมื่อกำหนดเงื่อนไขการตัวแปรเข้า/ออกจากสมการความถดถอยเชิงเส้นโดยพิจารณาจากค่าสถิติ F และมีค่าเป็นเท็จเมื่อกำหนดเงื่อนไขการตัวแปรเข้า/ออกจากสมการความถดถอยเชิงเส้นโดยพิจารณาจากค่าระดับนัยสำคัญของ F ▪ getProbEntryF() เมทธอดนี้ใช้สำหรับรับค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น

ตารางที่ 3.21 รายละเอียด ของคลาส LinearRegressionOption (ต่อ)

เมทธอด	*ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ F ที่ใช้เป็นเกณฑ์ในการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้น ▪ getProbRemovalF() เมทธอดนี้ใช้สำหรับรับค่าระดับนัยสำคัญของ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ getRemovalF() เมทธอดนี้ใช้สำหรับรับค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสถิติ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น ▪ getCriteriaMethod() เมทธอดนี้ใช้สำหรับรับวิธีการที่เลือกใช้เป็นเงื่อนไขในการเลือกตัวแปรอิสระเข้า/ออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* วิธีการที่เลือกใช้เป็นเงื่อนไขในการเลือกตัวแปรอิสระเข้า/ออกจากสมการความถดถอยเชิงเส้น ▪ getEntryCriteria() เมทธอดนี้ใช้สำหรับรับค่าสถิติทดสอบ F หรือ ค่านัยสำคัญของสถิติทดสอบ F ขึ้นอยู่กับเงื่อนไขในการเลือกตัวแปรอิสระเข้าจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสถิติทดสอบ F หรือ ค่านัยสำคัญของสถิติทดสอบ F ▪ getRemovalCriteria() เมทธอดนี้ใช้สำหรับรับค่าสถิติทดสอบ F หรือ ค่านัยสำคัญของสถิติทดสอบ F ขึ้นอยู่กับเงื่อนไขในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น *ค่าส่งกลับ* ค่าสถิติทดสอบ F หรือ ค่านัยสำคัญของสถิติทดสอบ F ที่ใช้เป็นเกณฑ์ในการตัดตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้น

จากผังแสดงความสัมพันธ์ระหว่างคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นในรูปที่ 3.14 และรายละเอียดของคลาสแต่ละคลาสทำให้ได้ผังความสัมพันธ์ระหว่างคลาสและรายละเอียดของคลาสแต่ละคลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นดังแสดงในรูปที่ 3.26



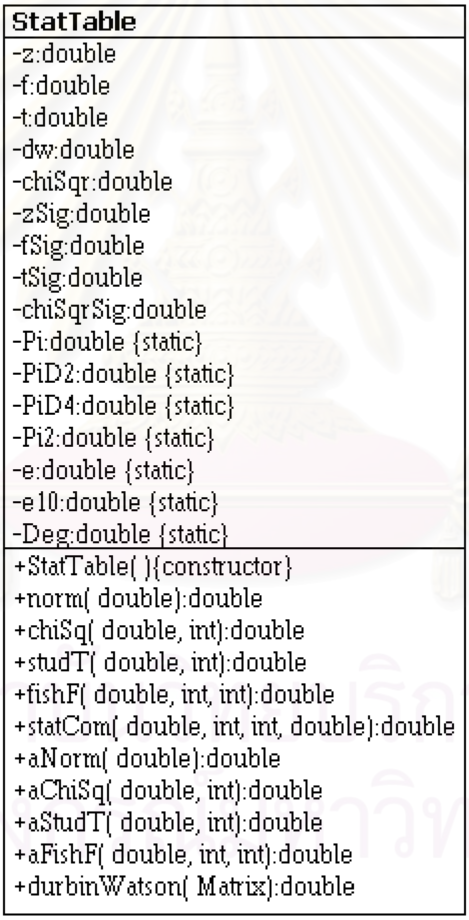
รูปที่ 3.26 ความสัมพันธ์ของคลาสต่างๆสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

3.2.3 คลาสสำหรับตารางสถิติ

ตารางสถิติ(Statistic table) แทนได้ด้วยคลาส StatTable ซึ่งประกอบด้วยตารางสถิติต่างๆคือ

1. ตารางความน่าจะเป็นแบบเอฟ (F Table)
2. ตารางความน่าจะเป็นแบบที (T Table)
3. ตารางความน่าจะเป็นแบบปรกติ (Z Table)
4. ตารางความน่าจะเป็นแบบไคสแควร์ (Chi – Squar Table)
5. ตารางค่าวิกฤตของ Durbin – Watson Statistic

คลาส StatTable ประกอบด้วยคุณสมบัติและเมทธอด ดังแสดงในรูปที่ 3.27 จากรูปที่ 3.27 คลาส StatTable ประกอบด้วยคุณสมบัติและเมทธอด โดยมีรายละเอียดดังแสดงในตารางที่ 3.22



รูปที่ 3.27 ส่วนประกอบของคลาส StatTable

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

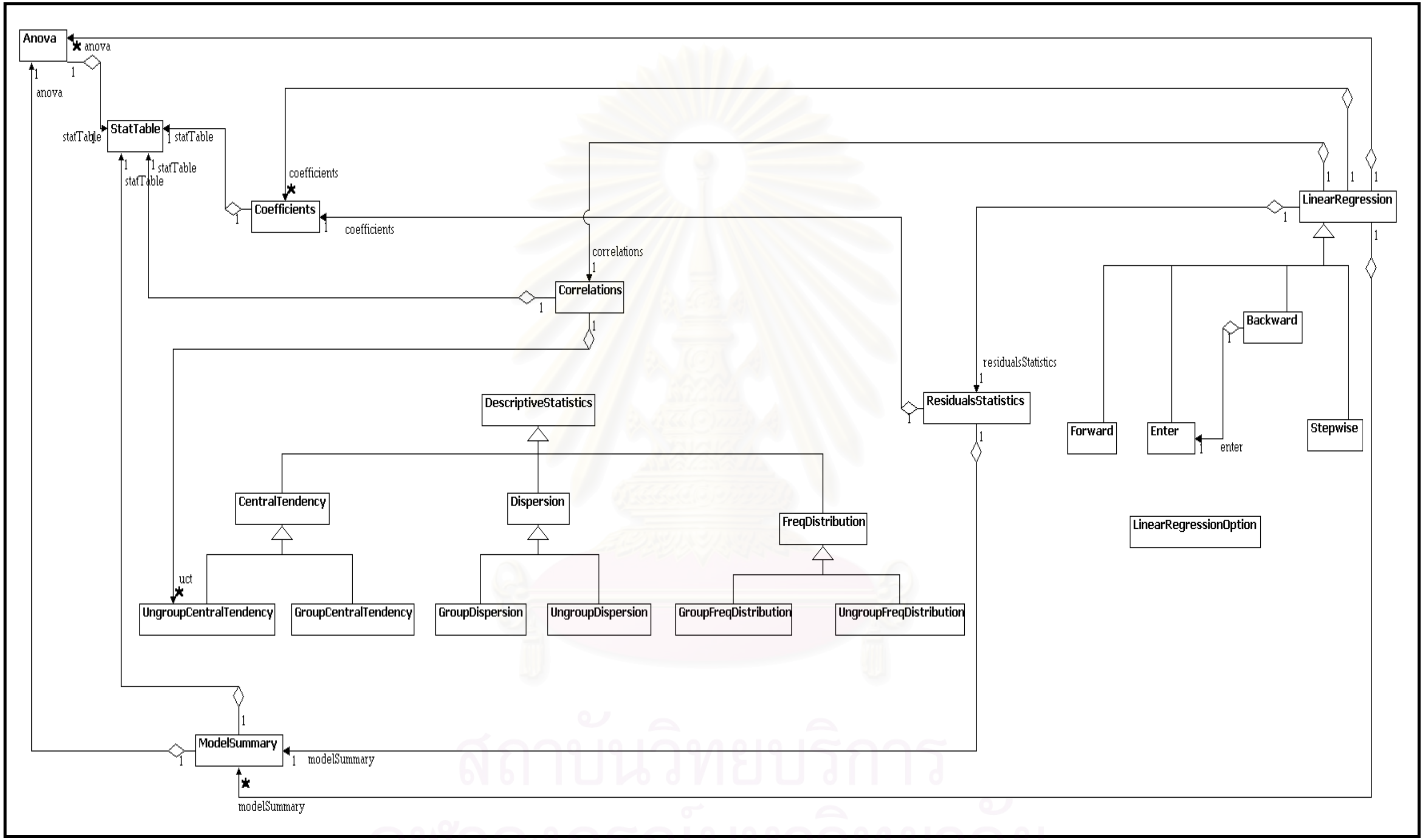
ตารางที่ 3.22 รายละเอียด ของคลาส StatTable

ชื่อคลาส	**StatTable**
ประเภท	Base Case
คำอธิบาย	ตารางสถิติ(Statistic table) ประกอบด้วย 1. ตารางความน่าจะเป็นแบบเอฟ (F Table) 2. ตารางความน่าจะเป็นแบบที (T Table) 3. ตารางความน่าจะเป็นแบบปรกติ (Z Table) 4. ตารางความน่าจะเป็นแบบไคสแควร์ (Chi - Squar Table) 5. ตารางค่าวิกฤตของ Durbin - Watson Statistic
ตัวแปร	▪ z ค่าสถิติทดสอบ z ▪ f ค่าสถิติทดสอบ f ▪ t ค่าสถิติทดสอบ t ▪ dw ค่าสถิติทดสอบ Durbin-Watson ▪ chiSqr ค่าไคสแควร์ ▪ zSig ค่านัยสำคัญของสถิติทดสอบ z ▪ fSig ค่านัยสำคัญของสถิติทดสอบ f ▪ tSig ค่านัยสำคัญของสถิติทดสอบ t ▪ chiSqrSig ค่านัยสำคัญของไคสแควร์
เมทธอด	▪ norm(double z) เมทธอดนี้ใช้สำหรับคำนวณขนาดของค่าสถิติทดสอบ z *ค่าส่งกลับ* ขนาดของค่าสถิติทดสอบ z *พารามิเตอร์ที่ใช้* z หมายถึง ค่าสถิติทดสอบ z ▪ chiSq(double x, int n) เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของค่าไคสแควร์ *ค่าส่งกลับ* ค่านัยสำคัญของค่าไคสแควร์ *พารามิเตอร์ที่ใช้* x หมายถึงค่าไคสแควร์ n หมายถึง ค่าองศาอิสระ ▪ studT(double t, int n) เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสถิติทดสอบ t *ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ t *พารามิเตอร์ที่ใช้* t หมายถึง ค่าสถิติทดสอบ t n หมายถึง ค่าองศาอิสระ

ตารางที่ 3.21 รายละเอียด ของคลาส LinearRegressionOption (ต่อ)

เมทธอด	■ fishF(double f, int n1, int n2) เมทธอดนี้ใช้สำหรับคำนวณค่านัยสำคัญของสถิติทดสอบ f *ค่าส่งกลับ* ค่านัยสำคัญของสถิติทดสอบ f *พารามิเตอร์ที่ใช้* f หมายถึง ค่าสถิติทดสอบ f n1 หมายถึง ค่าองศาอิสระที่ 1 n2 หมายถึง ค่าองศาอิสระที่ 2 ■ aNorm(double p) เมทธอดนี้ใช้สำหรับการคำนวณขนาดของค่าความน่าจะเป็น *ค่าส่งกลับ* ขนาดของค่าความน่าจะเป็น *พารามิเตอร์ที่ใช้* p หมายถึง ค่าความน่าจะเป็น ■ aChiSq(double p, int n) เมทธอดนี้ใช้สำหรับการคำนวณค่าไคสแควร์ *ค่าส่งกลับ* ค่าไคสแควร์ *พารามิเตอร์ที่ใช้* p หมายถึง ค่าความน่าจะเป็น n หมายถึง ค่าองศาอิสระ ■ aStudT(double p, int n) เมทธอดนี้ใช้สำหรับการคำนวณค่าสถิติทดสอบ t *ค่าส่งกลับ* ค่านัยสำคัญของค่าสถิติทดสอบ t *พารามิเตอร์ที่ใช้* p หมายถึง ค่าความน่าจะเป็น n หมายถึง ค่าองศาอิสระ ■ aFishF(double p, int n1, int n2) เมทธอดนี้ใช้สำหรับการคำนวณค่าสถิติทดสอบ f *ค่าส่งกลับ* ค่านัยสำคัญของค่าสถิติทดสอบ f *พารามิเตอร์ที่ใช้* p หมายถึง ค่าความน่าจะเป็น n1 หมายถึง ค่าองศาอิสระที่ 1 n2 หมายถึง ค่าองศาอิสระที่ 2 ■ durbinWatson(Matrix e) เมทธอดนี้ใช้สำหรับการคำนวณค่า Durbin-Watson *ค่าส่งกลับ* ค่า Durbin-Watson *พารามิเตอร์ที่ใช้* e หมายถึง เมตริกซ์ของค่าผิดพลาด

ผังความสัมพันธ์ระหว่างคลาสต่างๆทั้งหมด และรายละเอียดของคลาสแต่ละคลาสสำหรับสถิติเชิงพรรณนา การวิเคราะห์ความถดถอยเชิงเส้นและ ตารางสถิติแสดงดังรูปที่ 3.28



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

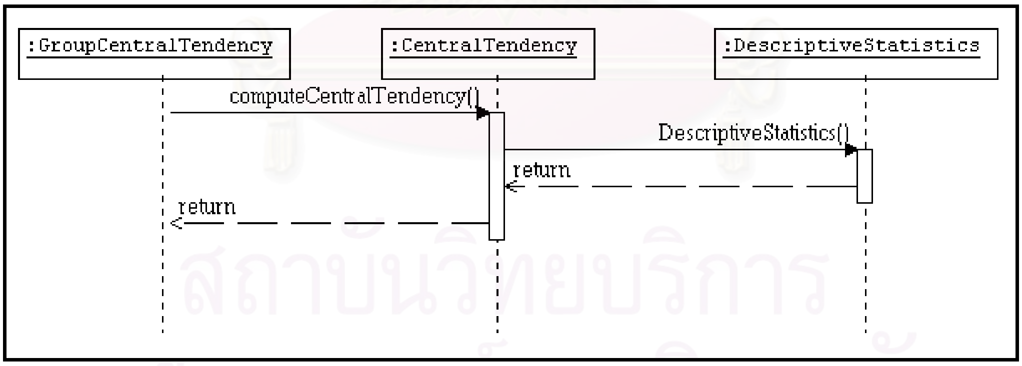
รูปที่ 3.28 ความสัมพันธ์ของคลาสทั้งหมดในขั้นตอนการวิเคราะห์และออกแบบ

3.3 แผนผังแสดงลำดับการคำนวณของส่วนประกอบซอฟต์แวร์สำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

แผนผังแสดงลำดับ (Sequence Diagram) แสดงถึงความสัมพันธ์ของวัตถุแบบไดนามิก โดยแผนผังแสดงถึงการส่งคำร้องขอระหว่างวัตถุกับวัตถุในขณะที่โปรแกรมทำงาน แผนผังแสดงลำดับการคำนวณค่าสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นประกอบด้วย แผนผังแสดงลำดับสำหรับการคำนวณค่ากลางทางสถิติ แผนผังแสดงลำดับสำหรับการคำนวณค่าความถี่ของข้อมูล แผนผังแสดงลำดับสำหรับการคำนวณค่าการกระจายของข้อมูล และ แผนผังแสดงลำดับสำหรับการวิเคราะห์ค่าความถดถอยเชิงเส้น ซึ่งประกอบด้วยการคำนวณค่าสถิติต่างๆได้แก่ การคำนวณค่าสถิติสรุปของสมการถดถอย การคำนวณค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว การคำนวณค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น การคำนวณค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น และการคำนวณค่าสถิติของค่าคลาดเคลื่อน

3.3.1 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม

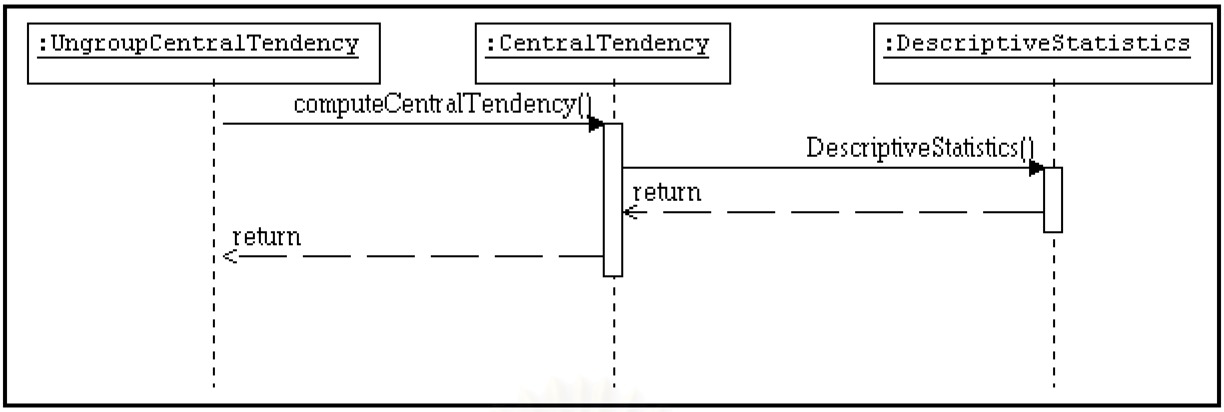
แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม แสดงการส่งคำร้องขอจากวัตถุ GroupCentralTendency ไปยังวัตถุ CentralTendency และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส GroupCentralTendency ดังแสดงในรูปที่ 3.29



รูปที่ 3.29 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบจัดกลุ่ม

3.3.2 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบไม่จัดกลุ่ม

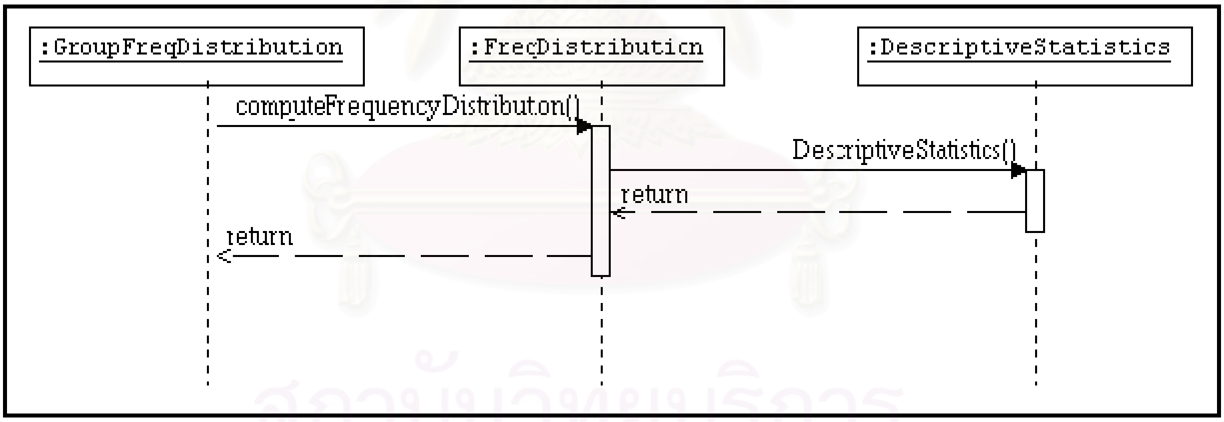
แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบไม่จัดกลุ่มแสดงการส่งคำร้องขอจากวัตถุ UngroupCentralTendency ไปยังวัตถุ CentralTendency และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส UngroupCentralTendency ดังแสดงในรูปที่ 3.30



รูปที่ 3.30 แผนผังแสดงลำดับของการคำนวณค่ากลางทางสถิติของข้อมูลแบบไม่จัดกลุ่ม

3.3.3 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม

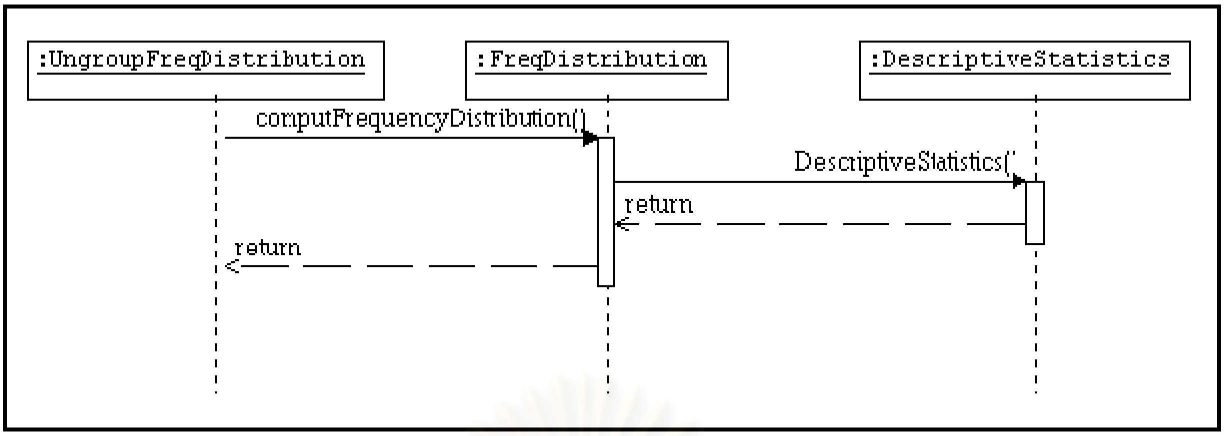
แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่มแสดงการส่งคำร้องขอจากวัตถุ GroupFreqDistribution ไปยังวัตถุ FreqDistribution และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส GroupFreqDistribution ดังแสดงในรูปที่ 3.31



รูปที่ 3.31 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม

3.3.4 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบไม่จัดกลุ่ม

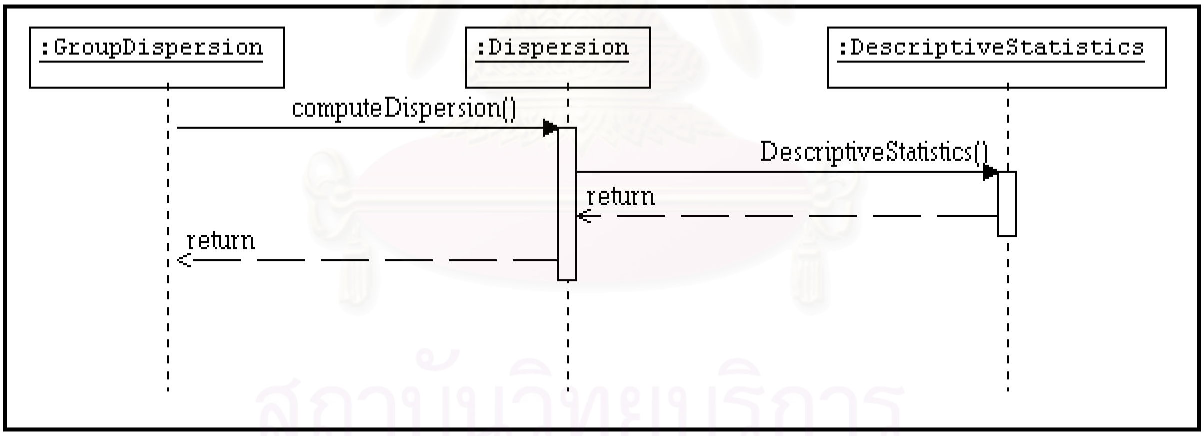
แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบไม่จัดกลุ่มแสดงการส่งคำร้องขอจากวัตถุ UngroupFreqDistribution ไปยังวัตถุ FreqDistribution และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส UngroupFreqDistribution ดังแสดงในรูปที่ 3.32



รูปที่ 3.32 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม

3.3.5 แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบจัดกลุ่ม

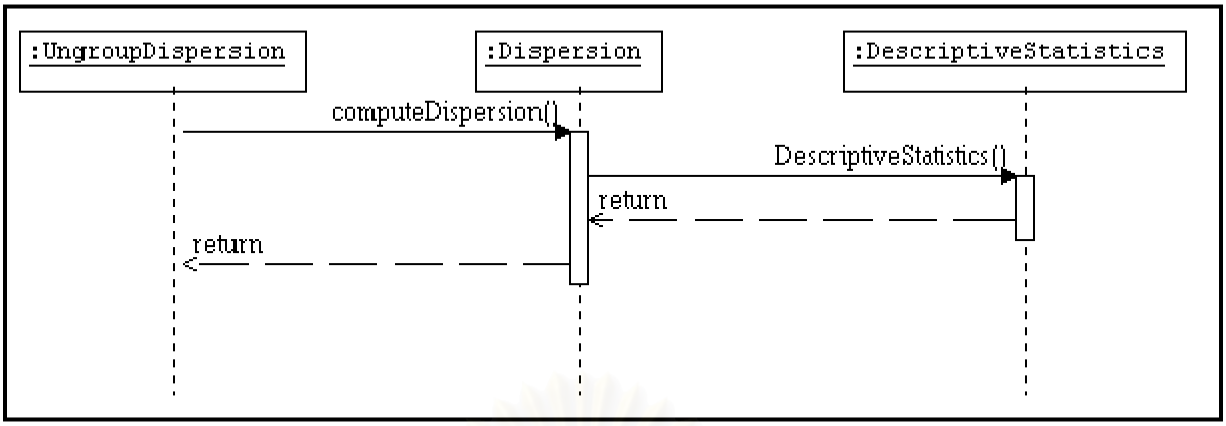
แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบจัดกลุ่มแสดงการส่งคำร้องขอจากวัตถุ GroupDispersion ไปยังวัตถุ Dispersion และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส GroupDispersion ดังแสดงในรูปที่ 3.33



รูปที่ 3.33 แผนผังแสดงลำดับของการคำนวณค่าความถี่ของข้อมูลแบบจัดกลุ่ม

3.3.6 แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบไม่จัดกลุ่ม

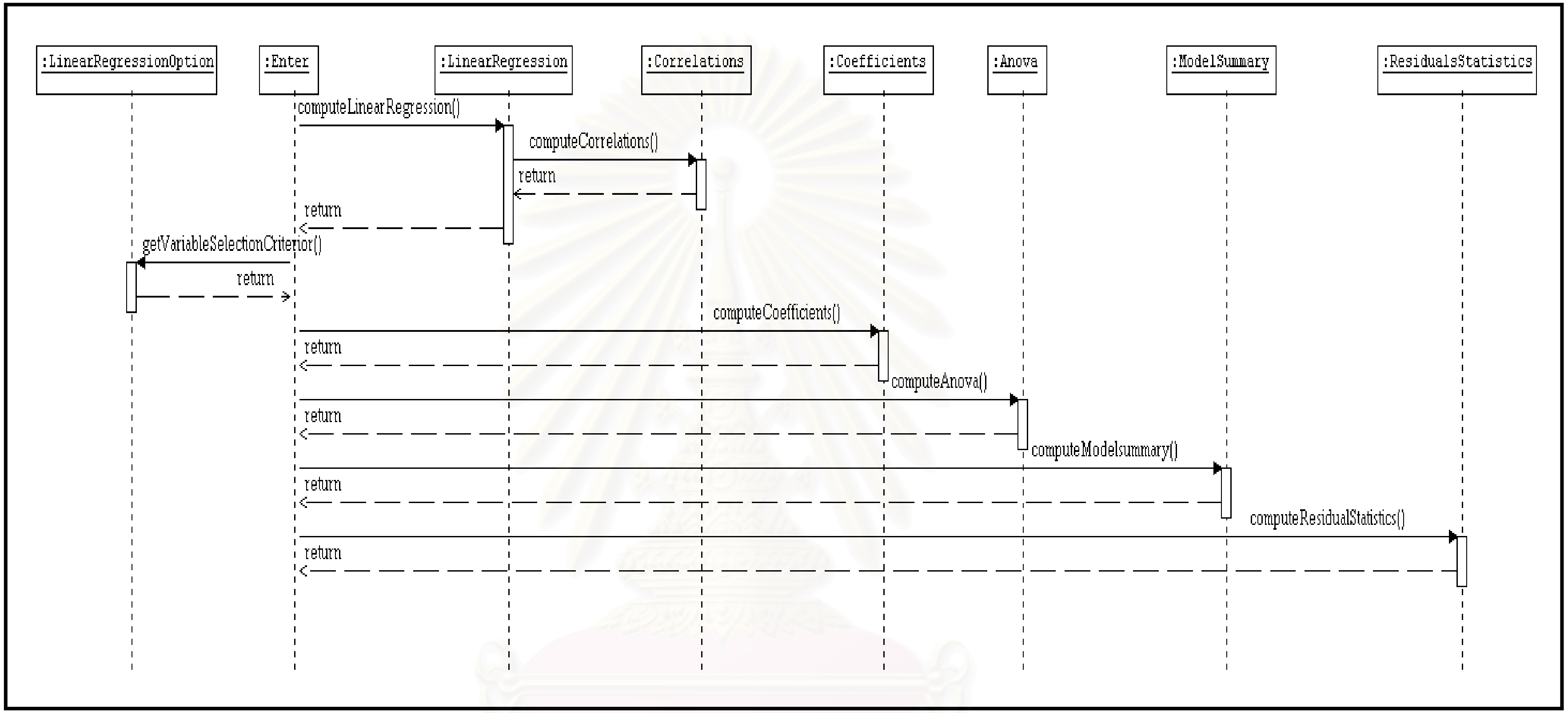
แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบไม่จัดกลุ่มแสดงถึงการส่งคำร้องขอจากวัตถุ UngroupDispersion ไปยังวัตถุ Dispersion และวัตถุนี้ส่งคำร้องขอต่อไปยังวัตถุ DescriptiveStatistics ซึ่งเป็นคลาสแม่ของคลาส UngroupDispersion ดังแสดงในรูปที่ 3.34



รูปที่ 3.34 แผนผังแสดงลำดับของการคำนวณค่าการกระจายของข้อมูลแบบไม่จัดกลุ่ม

3.3.7 แผนผังแสดงลำดับของการคำนวณค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น

แผนผังแสดงลำดับของการคำนวณค่าสถิติสำหรับการวิเคราะห์ความถดถอยเชิงเส้น แสดงถึงการส่งคำร้องขอจาก วัตถุ Enter Forward Backward หรือ Stepwise ไปยังวัตถุ LinearRegression ซึ่งเป็นคลาสแม่ และวัตถุสำหรับคำนวณค่าสถิติที่เกี่ยวข้องกับการวิเคราะห์ความถดถอยเชิงเส้นทั้งหมดซึ่งได้แก่วัตถุ ModelSummary วัตถุ Correlations วัตถุ Anova วัตถุ Coefficients และ วัตถุ ResidualsStatistics ส่วนวัตถุ LinearRegressionOption ถูกเรียกใช้ในการตรวจสอบเงื่อนไขที่จะนำตัวแปรอิสระเข้าหรือออกจากสมการความถดถอยเชิงเส้น แผนผังแสดงลำดับในรูปที่ 3.35 เป็นแผนผังแสดงลำดับของการคำนวณค่าสถิติการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยแบบเอนเตอร์ สำหรับการคำนวณค่าสถิติการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยอื่นๆ ก็มีการเรียกวัตถุต่างๆที่เกี่ยวข้องเช่นเดียวกันกับวิธีเอนเตอร์



รูปที่ 3.35 แผนผังลำดับของการคำนวณค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 4

การพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนา และการวิเคราะห์ความถดถอยเชิงเส้น

จากการวิเคราะห์และออกแบบเชิงวัตถุทำให้ได้รายละเอียดสมาชิกของคลาสและผังความสัมพันธ์ระหว่างคลาสต่างๆ ในบทนี้จะอธิบายขั้นตอนการพัฒนาโปรแกรมโดยอ้างอิงตามผังความสัมพันธ์และรายละเอียดที่ได้วิเคราะห์และออกแบบไว้ โดยในการพัฒนาแต่ละคลาสจะแสดงถึงการกำหนดประเภทคลาส คุณสมบัติของคลาส และ รายละเอียดโปรแกรมในส่วนของเมทธอดที่สำคัญๆของคลาส สำหรับโปรแกรมทั้งหมดสามารถดูได้จากภาคผนวก ก.

4.1 การพัฒนาโปรแกรมสำหรับคลาส DescriptiveStatistics

คลาส DescriptiveStatistics ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และกำหนดให้เป็นคลาสนามธรรม (Abstract Class) ดังแสดงในรูปที่ 4.1 คุณสมบัติภายในคลาสประกอบด้วยข้อมูลตัวอย่างและค่าจำนวนข้อมูลตัวอย่างซึ่งกำหนดขอบเขตการมองเห็นแบบโพรเทค ดังแสดงในรูปที่ 4.2 คลาส DescriptiveStatistics ประกอบด้วยเมทธอดที่สำคัญคือเมทธอดเรียงลำดับข้อมูล ซึ่งช่วยให้การคำนวณค่าสถิติ เช่น ค่าสูงสุด ค่าต่ำสุด ค่าความถี่ของข้อมูล และ ค่ามัธยฐานสามารถคำนวณค่าได้อย่างรวดเร็ว

```
package Statistics.DescriptiveStatistics
abstract public class DescriptiveStatistics
```

รูปที่ 4.1 แพ็กเกจและประเภทของคลาส DescriptiveStatistics

```
protected double[]  variableData;
protected int       numberOfData;
```

รูปที่ 4.2 คุณสมบัติของคลาส DescriptiveStatistics

4.2 การพัฒนาโปรแกรมสำหรับคลาส FreqDistribution

คลาส FreqDistribution ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคุณสมบัติจากคลาส DescriptiveStatistics ดังแสดงในรูปที่ 4.3 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งได้สืบทอดคุณสมบัติจากคลาส DescriptiveStatistics และคุณสมบัติเพิ่มเติมซึ่งกำหนดขอบเขตการมองเห็นแบบโพรเทค ดังแสดงในรูปที่ 4.4 คลาส FreqDistribution ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนามธรรมซึ่งทำหน้าที่นับค่าความถี่ดังแสดงในรูปที่ 4.5 โดยคลาสที่ต้องการสืบทอดคุณสมบัติจะต้องเขียนทับเมทธอดนี้

```
package Statistics.DescriptiveStatistics
abstract public class FreqDistribution extends DescriptiveStatistics
```

รูปที่ 4.3 แพ็กเกจและประเภทของคลาส FreqDistribution

```
protected  int[]     freqCount;          //Frequency Count
protected  int       numClass;           //Number of Class
protected  int[]     accumFreq;          //Accumulate Frequency
protected  double[]  percent;            //Percent
protected  double[]  accumPercent;       //Accumulate Percent
```

รูปที่ 4.4 คุณสมบัติของคลาส FreqDistribution

```
abstract protected void setFreqCount() throws Exception;
```

รูปที่ 4.5 เมทธอดที่สำคัญๆของคลาส FreqDistribution

4.3 การพัฒนาโปรแกรมสำหรับคลาส GroupFreqDistribution

คลาส GroupFreqDistribution ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส FreqDistribution ดังแสดงในรูปที่ 4.6 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งกำหนดขอบเขตการมองเห็นแบบไพรเวท ดังแสดงในรูปที่ 4.7 คลาส GroupFreqDistribution ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนับค่าความถี่ซึ่งเป็นเมทธอดที่ทำการเขียนทับเมทธอดที่ได้จากการสืบทอดคุณสมบัติจากคลาส FreqDistribution ดังแสดงในรูปที่ 4.8

```
package Statistics.DescriptiveStatistics
public class GroupFreqDistribution extends FreqDistribution
```

รูปที่ 4.6 แพ็กเกจและประเภทของคลาส GroupFreqDistribution

```
private int interval;
private double  classUpBound[];      //Class Upper Bound
private double  classLowBound[];   //Class Lower Bound
private double  midPoint[];             //Midpoint of Class
private int     mantissa;
```

รูปที่ 4.7 คุณสมบัติของคลาส GroupFreqDistribution

```
//Frequency count for grouping data
protected void setFreqCount()  {
  freqCount   =   new int[numClass];
  int i,j=0;
    for(i=0;i<numberOfData;i++)
        if(variableData[i] <= classUpBound[j])
            freqCount[j]++;
        else
            freqCount[++j]++;
}
```

รูปที่ 4.8 เมทธอดที่สำคัญๆของคลาส GroupFreqDistribution

4.4 การพัฒนาโปรแกรมสำหรับคลาส UngroupFreqDistribution

คลาส UngroupFreqDistribution ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส FreqDistribution ดังแสดงในรูปที่ 4.9 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งกำหนดขอบเขตการมองเห็นแบบไพรเวท ดังแสดงในรูปที่ 4.10 คลาส UngroupFreqDistribution ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนับค่าความถี่ซึ่งเป็นเมทธอดที่ทำการเขียนทับเมทธอดที่ได้จากการสืบทอดคุณสมบัติจากคลาส FreqDistribution ดังแสดงในรูปที่ 4.11

```
package Statistics.DescriptiveStatistics
public class UngroupFreqDistribution extends FreqDistribution
```

รูปที่ 4.9 แพ็กเกจและประเภทของคลาส UngroupFreqDistribution

```
 private double[]  variable;
 private int       maxFreq;
```

รูปที่ 4.10 คุณสมบัติของคลาส UngroupFreqDistribution

```
protected void setFreqCount() throws Exception {
  double[] tempVariable  = new double[numberOfData];
  int[] tempFreqCount  = new int[numberOfData];
  int k=0;
    tempVariable[k] = variableData[0];
    tempFreqCount[k] = 1;
    for(int i=1;i<numberOfData;i++)
        if(variableData[i] == variableData[i-1])
          tempFreqCount[k]++;
        else   {
            tempVariable[++k] = variableData[i];
            tempFreqCount[k] =  1;         }
    variable = new double[k+1];
    freqCount = new int[k+1];  numClass = k+1;
    for(int i =0;i<numClass;i++)      {
          variable[i] =  tempVariable[i];
          freqCount[i] = tempFreqCount[i];
          if(freqCount[i] > maxFreq)
              maxFreq = freqCount[i];  }
}
```

รูปที่ 4.11 เมทธอดที่สำคัญๆของคลาส UngroupFreqDistribution

4.5 การพัฒนาโปรแกรมสำหรับคลาส CentralTendency

คลาส CentralTendency ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคุณสมบัติจากคลาส DescriptiveStatistics ดังแสดงในรูปที่ 4.12 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดมาและคุณสมบัติเพิ่มเติมซึ่งกำหนดขอบเขตการมองเห็นแบบโพรเทค ดังแสดงในรูปที่ 4.13 คลาส CentralTendency ประกอบด้วยเมทธอดที่สำคัญได้แก่เมทธอดนามธรรมดังแสดงในรูปที่ 4.14 โดยคลาสที่ต้องการสืบทอดคุณสมบัติจะต้องเขียนทับเมทธอดเหล่านี้

```
package Statistics.DescriptiveStatistics
abstract public class CentralTendency extends DescriptiveStatistics
```

รูปที่ 4.12 แพ็กเกจและประเภทของคลาส CentralTendency

```
protected double mean;
protected double median;
protected double mode;
protected double sum;
protected double percentile;
protected double quartile;
protected double decile;
```

รูปที่ 4.13 คุณสมบัติของคลาส CentralTendency

```
//Calculate Mean
abstract public double mean();
//Calculate Median
abstract public double median();
//Calculate Mode
abstract public double mode() throws Exception;
```

รูปที่ 4.14 เมทธอดที่สำคัญๆของคลาส CentralTendency

4.6 การพัฒนาโปรแกรมสำหรับคลาส GroupCentralTendency

คลาส GroupCentralTendency ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส CentralTendency ดังแสดงในรูปที่ 4.15 คลาส GroupCentralTendency ประกอบด้วยเมทธอดที่สำคัญซึ่งเป็นเมทธอดที่ทำการเขียนทับเมทธอดที่ได้สืบทอดคุณสมบัติจากคลาส CentralTendency ดังแสดงในรูปที่ 4.16

```
package Statistics.DescriptiveStatistics
public class GroupCentralTendency extends CentralTendency
```

รูปที่ 4.15 แพ็กเกจและประเภทของคลาส GroupCentralTendency

```
//Calculate Mean
public double mean() {
 double sum = 0;
   try  {
     GroupFreqDistribution gfd = new GroupFreqDistribution(variableData);

   for(int i=0;i<gfd.numClass;i++)
     sum +=gfd.getMidPoint (i)*gfd.getfreqCount(i);
   mean = sum/ numberOfData;
   }
   catch(Exception e) {
   e.printStackTrace();}
   return(mean);
}

//Calculate Median
public double median()  {
   int medianClass=0;
```

รูปที่ 4.16 เมทธอดที่สำคัญๆของคลาส GroupCentralTendency

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
int previousAccumFreq;
  try  {
       GroupFreqDistribution gfd = new GroupFreqDistribution(variableData);
      for(int i=0;i<gfd.numClass;i++)
        if(gfd.getAccumFreq(i) > numberOfData/2)    {
            medianClass = i;
            i = gfd.numClass;
          }
      if(medianClass == 0)
        previousAccumFreq = 0;
      else
        previousAccumFreq = gfd.getAccumFreq(medianClass - 1);

      median =  gfd.getClassLowBound(medianClass) +
               (numberOfData/2 - previousAccumFreq * gfd.getInterval())
               /gfd.getfreqCount(medianClass);
    }
    catch(Exception e) {
    e.printStackTrace();}
    return(median);
}

//Calculate Mode
//Return Mode Value
public double mode() {
  int modeClass = 0;
  int previousFreq = 0;
  int nextFreq = 0;
  int maxFreq = 0;
```

รูปที่ 4.16 เมทธอดที่สำคัญๆของคลาส GroupCentralTendency (ต่อ)

```
  try  {
        GroupFreqDistribution gfd = new GroupFreqDistribution(variableData);
        for(int i=0;i<gfd.numClass;i++)
          if(gfd.getfreqCount(i) > maxFreq)        {
             maxFreq  =  gfd.getfreqCount(i);
             modeClass = i;
        }
    if(modeClass != 0)
      previousFreq = gfd.getfreqCount(modeClass-1);
    if(modeClass != gfd.getNumClass())
      nextFreq = gfd.getfreqCount(modeClass+1);
    mode  =  gfd.getClassLowBound(modeClass) +
         (gfd.getfreqCount(modeClass) -  previousFreq) / ( (gfd.getfreqCount(modeClass) -
         previousFreq) +  (gfd.getfreqCount(modeClass) -  nextFreq) ) *  gfd.getInterval();   }
  catch(Exception e) {
  e.printStackTrace();}
   return(mode);
}
```

รูปที่ 4.16 เมทธอดที่สำคัญๆของคลาส GroupCentralTendency (ต่อ)

4.7 การพัฒนาโปรแกรมสำหรับคลาส UngroupCentralTendency

คลาส UngroupCentralTendency ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส CentralTendency ดังแสดงในรูปที่ 4.17 คลาส UngroupCentralTendency ประกอบด้วยเมทธอดที่สำคัญซึ่งเป็นเมทธอดที่ทำการเขียนทับเมทธอดที่ได้จากการสืบทอดคุณสมบัติจากคลาส CentralTendency ดังแสดงในรูปที่ 4.18

```
package Statistics.DescriptiveStatistics
public class UngroupCentralTendency extends CentralTendency
```

รูปที่ 4.17 แพ็กเกจและประเภทของคลาส UngroupCentralTendency

```
//Calculate Mean
public double mean()
 {
  double sum = 0;
    for(int i=0;i<numberOfData;i++)
      sum +=variableData[i];
      mean = sum/ numberOfData;
    return(mean);
}

//Calculate Median
public double median()
{
//Check if numberOfData is odd
    if(numberOfData%2 == 1)
      median = variableData[(numberOfData+1)/2];
    else
      median = (variableData[(numberOfData/2)-1] +
               variableData[(numberOfData/2)])/2;
    return(median);
}

//Calculate Mode
public double mode() throws Exception
{
  int count;
  int maxFreq=0;
    UngroupFreqDistribution fd = new UngroupFreqDistribution(variableData);
    for(int i=0;i<fd.getNumClass();i++)
```

รูปที่ 4.18 เมทธอดที่สำคัญๆของคลาส UngroupCentralTendency

```
if(fd.getfreqCount(i) > maxFreq)      {
        mode = fd.getVariable (i);
        maxFreq = fd.getfreqCount(i);
}
   return(mode);
}
```

รูปที่ 4.18 เมทธอดที่สำคัญๆของคลาส UngroupCentralTendency (ต่อ)

4.8 การพัฒนาโปรแกรมสำหรับคลาส Dispersion

คลาส Dispersion ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส DescriptiveStatistics ดังแสดงในรูปที่ 4.19 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งกำหนดขอบเขตการมองเห็นแบบโพรเทค ดังแสดงในรูปที่ 4.20 คลาส Dispersion ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนามธรรมดังแสดงในรูปที่ 4.21 โดยคลาสที่ต้องการสืบทอดคุณสมบัติจะต้องเขียนทับเมทธอดเหล่านี้

```
package Statistics.DescriptiveStatistics
abstract public class Dispersion extends DescriptiveStatistics
```

รูปที่ 4.19 แพ็กเกจและประเภทของคลาส Dispersion

```
protected double min,max,range;
protected double variance;    // Variance
protected double sd;          // Standard deviation
protected double ad;          // Average deviation
protected double cv;          // Coefficient of variation
```

รูปที่ 4.20 คุณสมบัติของคลาส Dispersion

```
//Set Variance
abstract protected void setVariance() throws Exception;
//Set Standard Deviation
abstract protected void setSD() throws Exception;
//Set Coefficient of Variation
abstract protected void setCV() throws Exception;
//Set Average Deviation
abstract protected void setAD() throws Exception;
```

รูปที่ 4.21 เมทธอดที่สำคัญๆของคลาส Dispersion

4.9 การพัฒนาโปรแกรมสำหรับคลาส GroupDispersion

คลาส GroupDispersion ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส Dispersion ดังแสดงในรูปที่ 4.22 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดคุณสมบัติจากคลาส Dispersion คลาส GroupDispersion ประกอบด้วยเมทธอดที่สำคัญซึ่งเป็นเมทธอดที่ทำการเขียนทับเมทธอดที่ได้จากการสืบทอดคุณสมบัติจากคลาส Dispersion ดังแสดงในรูปที่ 4.23

```
package Statistics.DescriptiveStatistics
public class GroupDispersion extends Dispersion
```

รูปที่ 4.22 แพ็กเกจและประเภทของคลาส GroupDispersion

```
protected void setVariance() throws Exception {          //Calculate Variance
  double sum = 0;  GroupCentralTendency gct = new GroupCentralTendency(variableData);
  GroupFreqDistribution gfd = new GroupFreqDistribution(variableData);
    for(int i=0;i<gfd.getNumClass();i++)
      sum += Math.pow(gfd.getfreqCount(i)*(gfd.getMidPoint(i) - gct.mean()),2);
      variance = sum/ numberOfData;
```

รูปที่ 4.23 เมทธอดที่สำคัญๆของคลาส GroupDispersion

```
}

// Calculate Standard deviation
protected void setSD() throws Exception {
  setVariance();
  sd = Math.sqrt(variance);
}

//Calculate Average deviation
protected void setCV() throws Exception {
  GroupCentralTendency gct = new GroupCentralTendency(variableData);
    setSD();
    cv = sd / gct.mean();
}

//Calculate Average deviation
protected void setAD() throws Exception {
  double sum = 0;
  GroupCentralTendency gct = new GroupCentralTendency(variableData);
  GroupFreqDistribution gfd = new GroupFreqDistribution(variableData);
    for(int i=0;i<gfd.getNumClass();i++)
      sum += gfd.getfreqCount(i) * Math.abs(gfd.getMidPoint(i)-gct.mean());
    ad = sum/numberOfData;
}
```

รูปที่ 4.23 เมทธอดที่สำคัญๆของคลาส GroupDispersion (ต่อ)

4.10 การพัฒนาโปรแกรมสำหรับคลาส UngroupDispersion

คลาส UngroupDispersion ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.DescriptiveStatistics และเป็นคลาสนามธรรมที่สืบทอดคลาส Dispersion ดังแสดงในรูปที่ 4.24 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆที่สืบทอดมาจากคลาส Dispersion และคุณสมบัติเพิ่มเติมดังแสดงในรูปที่ 4.25 คลาส

UngroupDispersion ประกอบด้วยเมทธอดที่สำคัญดังแสดงในรูปที่ 4.26 โดยคลาสที่ต้องการสืบทอดคุณสมบัติจะต้องเขียนทับเมทธอดนี้

```
package Statistics.DescriptiveStatistics
public class UngroupDispersion extends Dispersion
```

รูปที่ 4.24 แพ็กเกจและประเภทของคลาส UngroupDispersion

```
protected double skewness;    // Skewness
  protected double kurtosis;    // Kurtosis
```

รูปที่ 4.25 คุณสมบัติของคลาส UngroupDispersion

```
//Set Variance
protected void setVariance() throws Exception {
  double sum = 0;
  UngroupCentralTendency uct = new UngroupCentralTendency(variableData);
    for(int i=0;i<numberOfData;i++)
       sum += Math.pow(variableData[i] - uct.mean(),2);
    variance = sum/ (numberOfData-1);
}

//Set Standard Deviation
protected void setSD() throws Exception {
  setVariance();
  sd = Math.sqrt(variance);
}

//Calculate Coefficient of Variation
protected void setCV() throws Exception {
```

รูปที่ 4.26 เมทธอดที่สำคัญๆของคลาส UngroupDispersion

```
  UngroupCentralTendency uct = new UngroupCentralTendency(variableData);
    setSD();
    cv = sd / uct.mean();
}

//Calculate Average Deviation
protected void setAD() throws Exception
{
  double sum = 0;
  UngroupCentralTendency uct = new UngroupCentralTendency(variableData);
  UngroupFreqDistribution ufd = new UngroupFreqDistribution(variableData);
    for(int i=0;i<ufd.getNumClass();i++)
        sum += Math.abs (variableData[i]-uct.mean());
    ad = sum/numberOfData;
}
```

รูปที่ 4.26 เมทธอดที่สำคัญๆของคลาส UngroupDispersion (ต่อ)

4.11 การพัฒนาโปรแกรมสำหรับคลาส LinearRegression

คลาส LinearRegression ถูกกำหนดให้อยู่ในแพ็กเกจ package Statistics.LinearRegression และเป็นคลาสนามธรรมดังแสดงในรูปที่ 4.27 คุณสมบัติภายในคลาสประกอบคุณสมบัติซึ่งกำหนดขอบเขตการมองเห็นแบบโพรเทค ดังแสดงในรูปที่ 4.28 คลาส LinearRegression ประกอบด้วยเมทธอดที่สำคัญซึ่งเป็นเมทธอดนามธรรมที่ทำหน้าที่เลือกตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นดังแสดงในรูปที่ 4.29 โดยคลาสที่ต้องการสืบทอดคุณสมบัติจะต้องเขียนทับเมทธอดนี้

```
package Statistics.LinearRegression
abstract public class LinearRegression {
```

รูปที่ 4.27 แพ็กเกจและประเภทของคลาส LinearRegression

```
static protected        Matrix X,Y;
static protected double  dependentVar[];
static protected double  independentVar[][];
static protected String  independentLabel[];
static protected String  dependentLabel;
static protected String  method;       // Method of variable selection
                                       // (Enter, Forward, Backward or Stepwise)
static protected int modelCount;
protected Anova[]                anova ;              //Anova statistics
protected Coefficients[]         coefficients;        //Coefficuents statistics
protected Correlations           correlations;        //Correlations
protected ResidualsStatistics    residualsStatistics; //Residuals Statistics
protected ModelSummary[]         modelSummary;        //Model Summary
protected int                    include[];           //Include variables
protected int                    exclude[];           //Exclude variables
```

รูปที่ 4.28 คุณสมบัติของคลาส LinearRegression

```
abstract protected void varSelection()throws Exception;
```

รูปที่ 4.29 เมทธอดที่สำคัญๆของคลาส LinearRegression

4.12 การพัฒนาโปรแกรมสำหรับคลาส Enter

คลาส Enter ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสนามธรรมที่สืบทอดคลาส LinearRegression ดังแสดงในรูปที่ 4.30 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดคุณสมบัติจากคลาส LinearRegression คลาส Enter ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนามธรรม ดังแสดงในรูปที่ 4.31

```
package Statistics.LinearRegression
public class Enter extends LinearRegression
```

รูปที่ 4.30 แพ็กเกจและประเภทของคลาส Enter

```
protected void varSelection() throws Exception
{
 //All variables are entered
   for(int i=0;i<include.length;i++)
     include[i] = i;
   coefficients[0]  = new Coefficients(Y,X,super.getEnteredVar());
   anova[0]         = new Anova(Y,X,coefficients[0].getCoefficientsMatrix());
   modelSummary[0]  = new ModelSummary(Y,X,anova[0],0,include.length-1);
}
```

รูปที่ 4.31 เมทธอดที่สำคัญๆของคลาส Enter

4.13 การพัฒนาโปรแกรมสำหรับคลาส Forward

คลาส Forward ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสนามธรรมที่สืบทอดคลาส LinearRegression ดังแสดงในรูปที่ 4.32 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดคุณสมบัติจากคลาส LinearRegression และคุณสมบัติเพิ่มเติมดังแสดงในรูปที่ 4.33 คลาส Enter ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนามธรรมดังแสดงในรูปที่ 4.34

```
package Statistics.LinearRegression
public class Forward extends LinearRegression
```

รูปที่ 4.32 แพ็กเกจและประเภทของคลาส Forward

```
private Matrix X;    //Matrix of independent for adjustment
private int   tPos;    //Position of t statistic
```

รูปที่ 4.33 คุณสมบัติของคลาส Forward

```
//Find appropriate regression equation : Forward selection
protected void varSelection()   throws Exception  {
    double    rSqrSave = 0;    //Rsquare of previous model
    //At start all variables are exclude
    for(int i=0;i<exclude.length;i++)
      exclude[i] = i;
    for(int i = 0;i<super.X.getColumnDimension();i++)     {
     int includeVar = getIncludeVar();
     if(includeVar != 0)        {
        coefficients[modelCount]  = new Coefficients(Y,X,super.getEnteredVar());
        anova[modelCount]         = new Anova(Y,X,
                        coefficients[modelCount].getCoefficientsMatrix());
        modelSummary[modelCount]  = new ModelSummary(Y,X,anova[modelCount],rSqrSave,1);
        rSqrSave = modelSummary[modelCount].getRSqr();
        modelCount++;     }
     else
        i =  super.X.getColumnDimension(); //Stop forward selection     }
}
```

รูปที่ 4.34 เมทธอดที่สำคัญๆของคลาส Forward

4.14 การพัฒนาโปรแกรมสำหรับคลาส Backward

คลาส Enter ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics. LinearRegression และเป็นคลาสนามธรรมที่สืบทอดคลาส LinearRegression ดังแสดงในรูปที่ 4.35 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดคุณสมบัติจากคลาส LinearRegression และคุณสมบัติเพิ่มเติมดังแสดงในรูปที่ 4.36 คลาส Backward ประกอบด้วยเมทธอดที่สำคัญคือ เมทธอดนามธรรมดังแสดงในรูปที่ 4.37

```
package Statistics.LinearRegression
public class Backward extends LinearRegression
```

รูปที่ 4.35 แพ็กเกจและประเภทของคลาส Backward

```
private  Enter enter;
private   Matrix X;         //Matrix of independent for adjustment
private   int tPos = 0;     //Position of t statistic
```

รูปที่ 4.36 คุณสมบัติของคลาส Backward

```
//Finding appropriate regression equation : Backward method
protected void varSelection()   throws Exception
{
   int     count = 0;
   //  rSqrSave : r square change of the previous model
   double  rSqrSave = modelSummary[0].getRSquareChange();
   double  df1Save  = modelSummary[0].getDf1();

   //At start all variables are include
   for(int i=0;i<include.length;i++)
     include[i] = i;

   enter = new Enter(dependentVar, independentVar,independentLabel
   ,dependentLabel,false);
   coefficients[0]  = enter.getCoefficients(0);
   anova[0]         = enter.getAnova(0);
   modelSummary[0]  = enter.getModelSummary(0);
   modelCount =1;

   for(int i = 1 ;i <super.X.getColumnDimension() ;i++)
    {
     int excludeVar = getExcludeVar(modelCount-1);
     if(excludeVar != 0)
```

รูปที่ 4.37 เมทธอดที่สำคัญๆของคลาส Backward

```
        {
         X  = new Matrix(AdjustX(excludeVar));
         coefficients[modelCount]  =
             new Coefficients(Y,X,super.getEnteredVar());
         anova[modelCount]         =
             new Anova(Y,X,coefficients[modelCount].getCoefficientsMatrix());
         modelSummary[modelCount]  = new ModelSummary(Y,X,anova[modelCount],
                   rSqrSave,1,anova[modelCount].getDfError()+1);
         rSqrSave  =   modelSummary[modelCount].getRSqr();
         df1Save   =   modelSummary[modelCount].getDf1();
         modelCount++;
        }
      else
         i =  super.X.getColumnDimension();  //No more variable to exclude
      }
  }
```

รูปที่ 4.37 เมทธอดที่สำคัญๆของคลาส Backward (ต่อ)

4.15 การพัฒนาโปรแกรมสำหรับคลาส Stepwise

คลาส Stepwise ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสนามธรรมที่สืบทอดคลาส LinearRegression ดังแสดงในรูปที่ 4.38 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆซึ่งสืบทอดคุณสมบัติจากคลาส LinearRegression และคุณสมบัติเพิ่มเติมดังแสดงในรูปที่ 4.39 คลาส Enter ประกอบด้วยเมทธอดที่สำคัญคือเมทธอดนามธรรมดังแสดงในรูปที่ 4.40

```
package Statistics.LinearRegression
public class Stepwise extends LinearRegression
```

รูปที่ 4.38 แพ็กเกจและประเภทของคลาส Stepwise

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
private   Matrix X;         //Matrix of independent for adjustment
private   int tPos = 0;     //Position of t statistic
```

รูปที่ 4.39 คุณสมบัติของคลาส Stepwise

```
//* Finding appropriate regression  equation : Stepwise
protected void varSelection()  throws Exception {
  int includeVar = 0;
  int excludeVar = 0;
  double    rSqrSave = 0;    //Rsquare of previous model

    //At start all variables are exclude
    for(int i=0;i<exclude.length;i++)
       exclude[i] = i;

    for(int i = 0;i<super.X.getColumnDimension();i++)      {
      //Forward selection
        includeVar = getIncludeVar();
        if(includeVar != 0)          {
           if(i != 0)
             modelCount++;
           coefficients[modelCount]  =
              new Coefficients(Y,X,super.getEnteredVar());
           anova[modelCount]         =
              new Anova(Y,X,coefficients[modelCount].getCoefficientsMatrix());
           modelSummary[modelCount] =  new ModelSummary(Y,X,anova[modelCount],rSqrSave,1);
           rSqrSave = modelSummary[modelCount].getRSqr();
           //Backward Elimination
     excludeVar = getExcludeVar(modelCount);
if(excludeVar != 0)       {
```

รูปที่ 4.40 เมทธอดที่สำคัญๆของคลาส Stepwise

```
                modelCount++;
            X  = new Matrix(backwardAdjustX(excludeVar));
            coefficients[modelCount]  = new Coefficients(Y,X,super.getEnteredVar());
            anova[modelCount]         = new Anova(Y,X,
                      coefficients[modelCount].getCoefficientsMatrix());
            modelSummary[modelCount]  =
                      new ModelSummary(Y,X,anova[modelCount],rSqrSave,1);
            rSqrSave = modelSummary[modelCount].getRSqr();
          }
        }
      else
        // No variable to enter
        i =  super.X.getColumnDimension(); //Stop Stepwise
      // Check if the same variable was entered then deleted
      if(includeVar == excludeVar)
        i =  super.X.getColumnDimension(); //Stop Stepwise
    }
}
```

รูปที่ 4.40 เมทธอดที่สำคัญๆของคลาส Stepwise (ต่อ)

4.16 การพัฒนาโปรแกรมสำหรับคลาส ModelSummary

คลาส ModelSummary ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสที่เป็นส่วนประกอบของคลาส LinearRegression ดังแสดงในรูปที่ 4.41 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆดังแสดงในรูปที่ 4.42 LinearRegression คลาส ModelSummary ประกอบด้วยเมทธอดที่ใช้ในการคำนวณค่าสถิติสรุปของสมการความถดถอยต่างๆดังแสดงในรูปที่ 4.43

```
package Statistics.LinearRegression
public class ModelSummary
```

รูปที่ 4.41 แพ็กเกจและประเภทของคลาส ModelSummary

```
private Matrix    X,Y;
private double    r;                  //Correlation Coefficient : R
private double    rSqr;               //Coefficient of Determination : R Square
private double    adjustRSqr;         //Adjust R Square
private double    stdError;           //Standard Error of Determine
private double    rSqrChange;         //R Square Change
private double    fChange;            //F change
private int       df1;                //Number of independent variables
private int       df2;                //Degree of Freedom : Residual(Error)
private double    fChangeSign;        //F change significance
private double    durbinWatson;       //Durbin-Watson value
private Anova     anova;
private StatTable statTable = new StatTable();
```

รูปที่ 4.42 คุณสมบัติของคลาส ModelSummary

```
//Set R Square :Coefficient of Determination
private void setRSquare()  throws Exception {
   rSqr   =  (anova.getSumSqrTot()-anova.getSumSqrError()) / anova.getSumSqrTot();
}

//Set R :Correlation Coefficient
private void setR() {
   r      =  Math.sqrt(rSqr);
}

//Set Adjust R Square
private void setAdjustRSquare(){
     adjustRSqr =  1 - (1 - rSqr)*anova.getDfTot()/anova.getDfError();
}
```

รูปที่ 4.43 เมทธอดที่สำคัญๆของคลาส ModelSummary

```
private void setStdError()  throws Exception{
   stdError = Math.sqrt(anova.getMeanSqrError());
}

//Set degree of freedom
private void setDf2()  {
    df2 = anova.getDfError();
}

//Set R Square Change
private void setRSqrChange(double rSqrSave) {
   rSqrChange = Math.abs(rSqr-rSqrSave);
}

//Set F Change
private void setFChange() {
   fChange =  (rSqrChange*df2) / (df1*(1-rSqr));
}

//Set F Change Significance
private void setFChangeSign() {
  fChangeSign = statTable.FishF(fChange,df1,df2);
}

//Set Durbin-Watson value
private void setDurbinWatson() {
  Coefficients coefficients;
  coefficients = new Coefficients(Y,X);
  durbinWatson = statTable.durbinWatson(coefficients.getEstimateResidual());
}
```

รูปที่ 4.43 เมทธอดที่สำคัญๆของคลาส ModelSummary (ต่อ)

4.17 การพัฒนาโปรแกรมสำหรับคลาส Correlations

คลาส Correlations ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสที่เป็นส่วนประกอบของคลาส LinearRegression ดังแสดงในรูปที่ 4.44 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆดังแสดงในรูปที่ 4.45 คลาส Correlations ประกอบด้วยเมทธอดที่ใช้ในการคำนวณค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวดังแสดงในรูปที่ 4.46

```
package Statistics.LinearRegression
public class Correlations
```

รูปที่ 4.44 แพ็กเกจและประเภทของคลาส Correlations

```
private Matrix      YX;
private double      pearsonCor[][];   //Pearson Correlations
private double      psig[][];            //Pearson Correlations Significance
private double[]    xSumSqr;
private double[]    xxSum;
private UngroupCentralTendency[] uct;
private StatTable statTable = new StatTable();
```

รูปที่ 4.45 คุณสมบัติของคลาส Correlations

```
//Set Pearson Correlations
private void setPearsonCor()  {
  int count = 0;
    count = 0;
    for(int i = 0;i < YX.getColumnDimension();i++)
      for(int j = i;j < YX.getColumnDimension();j++)
        {
          pearsonCor[j][i] = pearsonCor[i][j] =
```

รูปที่ 4.46 เมทธอดที่สำคัญๆของคลาส Correlations

```
xxSum[count]/Math.sqrt(xSumSqr[i]*xSumSqr[j]);
        count++;
}
}

//Set Pearson Correlations Significance
private void setSig() {
  double t;
  psig = new double[YX.getColumnDimension()][YX.getColumnDimension()];

   for(int i = 0;i<YX.getColumnDimension();i++)
    for(int j = i;j < YX.getColumnDimension();j++)
    {
     t =  pearsonCor[i][j] * Math.sqrt((YX.getRowDimension()-2)/(1-
         Math.pow(pearsonCor[i][j],2)));
     psig[i][j] = psig[j][i] = statTable.StudT(t,YX.getRowDimension())/2;
    }
}
```

รูปที่ 4.46 เมทธอดที่สำคัญๆของคลาส Correlations (ต่อ)

4.18 การพัฒนาโปรแกรมสำหรับคลาส Anova

คลาส Anova ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสที่เป็นส่วนประกอบของคลาส LinearRegression ดังแสดงในรูปที่ 4.47 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆ ดังแสดงในรูปที่ 4.48 คลาส Anova ประกอบด้วยเมทธอดที่ใช้ในการคำนวณค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นดังแสดงในรูปที่ 4.49

```
package Statistics.LinearRegression;
public class Anova
```

รูปที่ 4.47 แพ็กเกจและประเภทของคลาส Anova

```
private double  ssr;                // Sum Square of Regression
private double  sse;                // Sum Square of Error
private double  sst;                // Sum Square of Total
private double  msr;                // Mean Square of Regression
private double  mse;                // Mean Square of Error
private double  mst;                // Mean Square of Total
private int     dfReg;              // Degree of Freedom : Regression
private int     dfError;            // Degree of Freedom : Residual(Error)
private int     dfTot;              // Degree of Freedom : Total
private double  f;                  // F Test
private double  fSig;               // Significance of F
private Matrix  X,Y,b;
private StatTable statTable = new StatTable();
```

รูปที่ 4.48 คุณสมบัติของคลาส Anova

```
//Set Degree of Freedom
private void setDf() {
   dfTot   = Y.getRowDimension()-1;
   dfReg   = X.getColumnDimension()-1;
   dfError = dfTot - dfReg;
}

//Set Sum Square of Regression
private void setSumSqrReg()  throws Exception {
  double [] y = new double[Y.getRowDimension()];
  double sum =0;

   for(int i = 0; i< Y.getRowDimension();i++)
        y[i] = Y.get(i,0);
```

รูปที่ 4.49 เมทธอดที่สำคัญๆของคลาส Anova

```
UngroupCentralTendency uct = new UngroupCentralTendency(y);
    ssr  =  b.transpose().times(X.transpose()).times(Y).get(0,0)
             - Y.getRowDimension() * Math.pow(uct.mean(),2);
}

//Set Sum Square of Error
private void setSumSqrError() throws Exception {
    sse  = getSumSqrTot() - getSumSqrReg();
}

//Set Sum Square of Total
private void setSumSqrTot()  throws Exception {
  double [] y = new double[Y.getRowDimension()];
  sst  = 0;
    for(int i = 0; i< Y.getRowDimension();i++)
          y[i] = Y.get(i,0);
UngroupCentralTendency uct = new UngroupCentralTendency(y);
    for(int i = 0; i< Y.getRowDimension();i++)
          sst += Math.pow((Y.get(i,0) - uct.mean()),2);
}

//Set Mean Square of Regression
private void setMeanSqrReg() {
    msr     =  ssr/dfReg;
}

//Set Mean Square of Error
private void setMeanSqrError() {
    mse     =  sse/dfError;
}
```

รูปที่ 4.49 เมทธอดที่สำคัญๆของคลาส Anova (ต่อ)

```
UngroupCentralTendency uct = new UngroupCentralTendency(y);
   for(int i = 0; i< Y.getRowDimension();i++)
        sst += Math.pow((Y.get(i,0) - uct.mean()),2);
}

//Set Mean Square of Regression
private void setMeanSqrReg() {
   msr     = ssr/dfReg;
}

//Set Mean Square of Error
private void setMeanSqrError() {
   mse     = sse/dfError;
}

//Set Mean Square of Total
private void setMeanSqrTot() {
   mst     = sst/dfTot;
}

//Set F test value
private void setF() {
   f      =   msr/mse;
}

//Set F significance
private void setFSig() {
   fSig = statTable.FishF(f,dfReg,dfError);
}
```

รูปที่ 4.49 เมทธอดที่สำคัญๆของคลาส Anova (ต่อ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4.19 การพัฒนาโปรแกรมสำหรับคลาส Coefficients

คลาส Coefficients ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสที่เป็นส่วนประกอบของคลาส LinearRegression ดังแสดงในรูปที่ 4.50 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆดังแสดงในรูปที่ 4.51 คลาส Coefficients ประกอบด้วยเมทธอดที่ใช้ในการคำนวณของการประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้นดังแสดงในรูปที่ 4.52

```
package Statistics.LinearRegression
public class Coefficients
```

รูปที่ 4.50 แพ็กเกจและประเภทของคลาส Coefficients

```
private Matrix    b,X,Y,e;
private String[]  bLabels;
private Matrix    Ye;                //Estimate value of Y
private double    stdDevReg;         //Standard deviation of regression
private double    stdDevRegSqr;      //Standard deviation square
private double[]  stdError;          //Standard error
private double[]  bUpper;            //Upper bound of confidence interval of b
private double[]  bLower;            //Lower bound of confidence interval of b
private double[]  t;                 //t values
private double[]  tSig;              //t significance values
private StatTable statTable = new StatTable();
```

รูปที่ 4.51 คุณสมบัติของคลาส Coefficients

```
//Set Coefficients
private void setCoefficients() {
  b  = new  Matrix(X.getColumnDimension(),1);
  b  = X.transpose().times(X).inverse().times(X.transpose()).times(Y);
```

รูปที่ 4.52 เมทธอดที่สำคัญๆของคลาส Coefficients

```
}
//Set Standard deviation Square
private void setstdDevRegSqr() {
   Ye = X.times(b);
   e = Y.minus(Ye);    // Estimate Residuals Matrix
   stdDevRegSqr = e.transpose().times(e).get(0,0) /
               (X.getRowDimension()-X.getColumnDimension());
}

//Set Standard deviation
private void setstdDevReg() {
   stdDevReg = Math.sqrt(stdDevRegSqr);
}

//Set Standard Error of Bi : S(Bi)
private void setStdError() {
  Matrix  S; //Standard Error Square of B Matrix
  stdError = new double[X.getColumnDimension()];
   S = X.transpose().times(X).inverse().times(stdDevRegSqr);
   for(int i=0;i < X.getColumnDimension();i++)
     stdError[i] = Math.sqrt(S.get(i,i));
}

//Set confidence interval of B
private void setConfInterval() {
   bUpper  = new double[X.getColumnDimension()];
   bLower  = new double[X.getColumnDimension()];

   double tVal;  // t value
   tVal = statTable.AStudT(0.05,X.getRowDimension()-X.getColumnDimension()-2);
```

รูปที่ 4.52 เมทธอดที่สำคัญๆของคลาส Coefficients (ต่อ)

```
for(int i=0;i<X.getColumnDimension();i++) {
        bUpper[i] = b.get(i,0)+tVal*stdError[i];
        bLower[i] = b.get(i,0)-tVal*stdError[i];
      }
}

//Set t values
private void setT() {
    t = new double[X.getColumnDimension()];
    for(int i=0;i<X.getColumnDimension();i++)
      t[i] = b.get(i,0)/stdError[i];
}

//Set t significance values
private void setTSig() {
    tSig = new double[X.getColumnDimension()];
    for(int i=0;i<X.getColumnDimension();i++)
      tSig[i] = statTable.StudT(t[i],X.getRowDimension());
}
```

รูปที่ 4.52 เมทธอดที่สำคัญๆของคลาส Coefficients (ต่อ)

4.20 การพัฒนาโปรแกรมสำหรับคลาส ResidualsStatistics

คลาส ResidualsStatistics ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression และเป็นคลาสที่เป็นส่วนประกอบของคลาส LinearRegression ดังแสดงในรูปที่ 4.53 คุณสมบัติภายในคลาสประกอบด้วยคุณสมบัติต่างๆดังแสดงในรูปที่ 4.54 คลาส ResidualsStatistics ประกอบด้วยเมทธอดที่ใช้ในการคำนวณสถิติวิเคราะห์ของค่าคลาดเคลื่อนดังแสดงในรูปที่ 4.55

```
package Statistics.LinearRegression
public class ResidualsStatistics
```

รูปที่ 4.53 แพ็กเกจและประเภทของคลาส ResidualsStatistics

```
private Matrix          X,Y;
private Matrix          H;                  //High leverage Matrix
private Coefficients    coefficients;       //Coefficients statistics
private ModelSummary    modelSummary;       //Model Summary
private double          dependent[];        //Dependent Variable
private double          zpred[];            //Standardized Predicted Values
private double          zresid[];           //Standardized Residuals
private double          dresid[];           //Deleted Residuals
private double          sdresid[];          //Studentized deleted Residual
private double          sresid[];           //Studentized Residual
private double          adjpred[];          //Adjust Predicted Values
```

รูปที่ 4.54 คุณสมบัติของคลาส ResidualsStatistics

```
//Set deleted residuals : dresid
private void setDresid(Matrix Ye,int index)  {
   dresid[index] = Y.get(index,0) - Ye.get(index,0);
}

//Set Adjusted predicted values : adjpred
private void setAdjPred(Matrix Ye,int index) {
   adjpred[index] = Ye.get(index,0);
}

//Set Studentized delete residuals : sdresid
```

รูปที่ 4.55 เมทธอดที่สำคัญๆของคลาส ResidualStatistics

```
private void setSdresid(Matrix Ye,int index) {
 double  variance=0;
 double  sd;
   double  yeMean=0;

    for(int i=0;i<Ye.getRowDimension();i++)
       yeMean +=  Ye.get(i,0);

    yeMean = yeMean/Ye.getRowDimension();

   for(int i=0;i<Ye.getRowDimension();i++)
     variance  += Math.pow(Ye.get(i,0) - yeMean,2);
   variance = variance/(Ye.getRowDimension()-1);
   sd = Math.sqrt(variance);

   int cp =  X.getRowDimension()-X.getColumnDimension();

   double si = 1/Math.sqrt(cp-1)
            * Math.sqrt((cp* Math.pow(modelSummary.getStdError(),2) /
            (1-H.get(index,index)))- Math.pow(dresid[index],2));
   sdresid[index] =  dresid[index] / si;
}

 //Set Dependent variable
private void setDependentVariable()  {
   dependent = new double[Y.getRowDimension()];
   for(int i=0;i<Y.getRowDimension();i++)
      dependent[i] = Y.get(i,0);
}

//Set Standardize predicted values : zpred
```

รูปที่ 4.55 เมทธอดที่สำคัญๆของคลาส ResidualStatistics (ต่อ)

```
private void setStandardPredictValues()  {
double yeMean=0;    //Mean of estimated Y;
 double variance=0;
   double sd;
   Matrix Ye =  coefficients.getEstimateY();
   for(int i=0;i<Ye.getRowDimension();i++)
      yeMean  +=  Ye.get(i,0);
   yeMean  =   yeMean/Ye.getRowDimension();

   for(int i=0;i<Ye.getRowDimension();i++)
      variance  += Math.pow(Ye.get(i,0) - yeMean,2);
   variance = variance/(Ye.getRowDimension()-1);
   sd = Math.sqrt(variance);

   for(int i=0;i<Ye.getRowDimension();i++)
     zpred[i] = (Ye.get(i,0) - yeMean) / sd;
}

 //Set Standardize residual : zresid
private void setStandardResiduals()  {
   double   eMean  = 0; //Mean of EstimateResidual;
   Matrix Ye   =  coefficients.getEstimateY();
   Matrix e    =  coefficients.getEstimateResidual();

   for(int i=0;i<e.getRowDimension();i++)
      eMean +=  e.get(i,0);

   eMean = eMean/e.getRowDimension();

   for(int i=0;i<Ye.getRowDimension();i++)
       zresid[i] =  (e.get(i,0) - eMean) /
```

รูปที่ 4.55 เมทธอดที่สำคัญๆของคลาส ResidualStatistics (ต่อ)

```
modelSummary.getStdError();
}

//Set Studentized residual : sresid
private void setStudentResidual()  {
    Matrix Ye  =  coefficients.getEstimateY();
    Matrix e   =  coefficients.getEstimateResidual();

    sresid = new double[Y.getRowDimension()];
     for(int i=0;i<Ye.getRowDimension();i++)      {
        sresid[i] =  e.get(i,0) /(modelSummary.getStdError() *
                   Math.sqrt(1-H.get(i,i)));
     }
}
```

รูปที่ 4.55 เมทธอดที่สำคัญๆของคลาส ResidualStatistics (ต่อ)

4.21 การพัฒนาโปรแกรมสำหรับคลาส LinearRegressionOption

คลาส LinearRegressionOption ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.LinearRegression ดังรูปที่ 4.56 คลาสนี้ประกอบด้วยคุณสมบัติภายในคลาสเป็นแบบคงที่ (Static) ซึ่งสามารถเรียกใช้ได้โดยไม่ต้องสร้างคลาส คลาสนี้ถูกเรียกใช้ในการตรวจสอบเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด แบ็คเวิร์ด และสเตปไวส์ คุณสมบัติของคลาส LinearRegressionOption แสดงดังรูปที่ 4.57 คลาส LinearRegressionOption ประกอบด้วยเมทธอดที่ใช้ในการกำหนดค่าคุณสมบัติต่างๆของคลาสดังแสดงในรูปที่ 4.58

```
package Statistics.LinearRegression
public class LinearRegressionOption
```

รูปที่ 4.56 แพ็กเกจและประเภทของคลาส LinearRegressionOption

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
  static protected double  probEntryF;         // Probability of F to enter
  static protected double  probRemovalF;    // Probability of F to remove
  static protected double  entryF;               // Value of F to enter
  static protected double  removalF;           // Value of F to remove
  static protected boolean criteriaMethod;  // Type of criteria
```

รูปที่ 4.57 คุณสมบัติของคลาส LinearRegressionOption

```
//Set probability of F for entry criteria
public static void setProbEntryF(double probEntry) {
  probEntryF = probEntry;
}

//Set probability of F for removal criteria
public static void setProbRemovalF(double probRemoval) {
  probRemovalF = probRemoval;
}

//Set value of F for entry criteria
public static void setEntryF(double entry) {
  entryF = entry;
}

//Set value of F for removal criteria
public static void setRemovalF(double removal) {
  removalF = removal;
}
```

รูปที่ 4.58 เมทธอดที่สำคัญๆของคลาส LinearRegressionOption

4.22 การพัฒนาโปรแกรมสำหรับคลาส StatTable

คลาส StatTable ถูกกำหนดให้อยู่ในแพ็กเกจ Statistics.StatTable ดังรูปที่ 4.59 คลาสนี้ประกอบด้วยคุณสมบัติแสดงดังรูปที่ 4.60 และประกอบด้วยเมทธอดที่ใช้ในการกำหนดค่าคุณสมบัติต่างๆของคลาสดังแสดงในรูปที่ 4.61

```
package Statistics.StatTable
public StatTable()
```

รูปที่ 4.59 แพ็กเกจและประเภทของคลาส StatTable

```
private static double Pi=Math.PI;
private static double PiD2=Pi/2;
private static double PiD4=Pi/4;
private static double Pi2=2*Pi;
private static double e=2.718281828459045235;
private static double e10 = 1.105170918075647625;
private static double Deg=180/Pi;              //Degree
private double z;                              //z value
private double f;                              //f value
private double t;                              //t value
private double dw;                             //Durbin Watson value
private double chiSqr;                         //Chi-Square value
private double zSig;                           //z Significance
private double fSig;                           //F Significance
private double tSig;                           //t Significance
private double chiSqrSig;                      //Chi-Square Significance
```

รูปที่ 4.60 คุณสมบัติของคลาส StatTable

```
//Get norm of z
public double norm(double z) {
   z=Math.abs(z);
   double p  = 1+ z*(0.04986735+ z*(0.02114101+ z*(0.00327763+ z*(0.0000380036+
              z*(0.0000488906+ z*0.000005383)))));
   p=p*p; p=p*p; p=p*p;
   return 1/(p*p);
   }

//Get Chi-Square significance
public double chiSq(double x,int n) {
   double p=Math.exp(-0.5*x);
   if((n%2)==1) { p=p*Math.sqrt(2*x/Pi); }
   int k=n;
   while(k>=2) { p=p*x/k; k=k-2; }
   double t=p; int a=n;
   while(t>0.000001*p) { a=a+2; t=t*x/a; p=p+t; }
   chiSqrSig = 1-p;
   return(chiSqrSig);
   }

//Get t Significance
public double studT(double t,int n) {
   t=Math.abs(t);
   double w=t/Math.sqrt(n);
   double th=Math.atan(w);
   if(n==1) { return (tSig = 1-th/PiD2); }
   double sth=Math.sin(th);
   double cth=Math.cos(th);
```

รูปที่ 4.61 เมทธอดที่สำคัญๆของคลาส StatTable

```
if((n%2)==1)
       { tSig = 1-(th+sth*cth*StatCom(cth*cth,2,n-3,-1))/PiD2; }
       else
       { tSig = 1-sth*StatCom(cth*cth,1,n-3,-1); }
   return(tSig);
   }

//Get F Significance
public double fishF(double f,int n1, int n2) {
   double x=n2/(n1*f+n2);
   if((n1%2)==0){return fSig = StatCom(1-x,n2,n1+n2-4,n2-2)*Math.pow(x,n2/2); }
   double th=Math.atan(Math.sqrt(n1*f/n2));
   double a=th/PiD2;
   double sth=Math.sin(th);
   double cth=Math.cos(th);
   if(n2>1) { a=a+sth*cth*StatCom(cth*cth,2,n2-3,-1)/PiD2; }
   if(n1==1) { return fSig=1-a; }
   if((n2%2)==0)
   {return fSig = 1-StatCom(x,n1,n1+n2-4,n1-2)*Math.pow(1-x,n1/2);}
   double c=4*StatCom(sth*sth,n2+1,n1+n2-4,n2-2)*sth*Math.pow(cth,n2)/Pi;
   if(n2==1) { return fSig=1-a+c/2; }
   double k=2; while(k<=(n2-1)/2) {c=c*k/(k-.5); k=k+1; }
   return(fSig=1-a+c);
   }

public double statCom(double q,int i,int j,double b) {
   double zz=1; double z=zz; int k=i;
   while(k<=j) { zz=zz*q*k/(k-b); z=z+zz; k=k+2; }
   return(z);
   }
```

รูปที่ 4.61 เมทธอดที่สำคัญๆของคลาส StatTable (ต่อ)

```
 public double aNorm(double p) {
   double v=0.5; double dv=0.5;
   z=0;
   while(dv>1e-6)   {
     z=1/v-1; dv=dv/2;
     if(Norm(z)>p) { v=v-dv; } else { v=v+dv; }
   }
   return (z);
   }

//Get Chi-Square value
public double aChiSq(double p,int n) {
double v=0.5; double dv=0.5; double x=0;
   while(dv>1e-6)    {
     x=1/v-1; dv=dv/2;
     if(ChiSq(x,n)>p) { v=v-dv; } else { v=v+dv; }
   }
   return (chiSqr=x);
}

//Get t value
public double aStudT(double p,int n) {
double v=0.5; double dv=0.5;   t = 0;
   while(dv>1e-6)    {
     t=1/v-1; dv=dv/2;
     if(StudT(t,n)>p) { v=v-dv; } else { v=v+dv; }    }
   return(t);
}
```

รูปที่ 4.61 เมทธอดที่สำคัญๆของคลาส StatTable (ต่อ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
//Get F value
public double aFishF(double p,int n1,int n2) {
double v=0.5; double dv=0.5; f = 0;
while(dv>1e-6) { f=1/v-1; dv=dv/2;
   if(FishF(f,n1,n2)>p) { v=v-dv;} else { v=v+dv; } }
   return(f);
}

//Get Durbin-Watson value
public double durbinWatson(Matrix e) {
double etSqr=0;    //e(t) square
double eDif=0;     //e(t) - e(t-1) Square
   for(int t = 1;t<e.getRowDimension();t++)
      eDif += Math.pow(e.get(t,0)-e.get(t-1,0),2);
   for(int t = 0;t<e.getRowDimension();t++)
      etSqr += Math.pow(e.get(t,0),2);
   dw =  eDif/etSqr;
   return(dw);
}
```

รูปที่ 4.61 เมทธอดที่สำคัญๆของคลาส StatTable (ต่อ)

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

การใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ในบทนี้จะอธิบายถึงขั้นตอนในการนำส่วนประกอบซอฟต์แวร์เชิงวัตถุที่ได้พัฒนาไปใช้ร่วมในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆ โดยส่วนประกอบทางซอฟต์แวร์ที่ได้พัฒนาขึ้นนี้ ผู้วิจัยได้จัดทำให้มีความสะดวกในการนำไปใช้ดังจะได้กล่าวถึงรายละเอียดต่อไป

5.1 การจัดเก็บไฟล์ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ส่วนประกอบซอฟต์แวร์สำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นถูกรวบรวมจัดเก็บไว้ในไฟล์ statistics.jar ซึ่งเป็นไฟล์ที่เก็บคลาสสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นรวมถึงคลาสที่เกี่ยวข้องทั้งหมด ไฟล์ statistics.jar เป็นไฟล์ประเภทที่มีการบีบอัดทำให้ไฟล์มีขนาดเล็กรวมทั้งเป็นไฟล์ที่รวบรวมคลาสต่างๆที่เกี่ยวข้องทั้งหมดจึงสะดวกในการนำไปใช้งาน การนำส่วนประกอบซอฟต์แวร์สำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นไปใช้งานจะต้องคัดลอกไฟล์นี้ไปไว้ในเส้นทางของไลบรารี (Library Path) หรือ กำหนดตำแหน่งของไฟล์โดยการกำหนดเส้นทางของคลาส (Class Path) ไว้ในไฟล์ Autoexec.bat เพื่อให้จาวาอินเตอร์พรีเตอร์สามารถค้นหาคลาสได้จากไดเรกทอรีที่กำหนด

<u>ตัวอย่างเช่น</u>

ถ้าเก็บไฟล์ statistics.jar ไว้ในไดเรกทอรี C:\statistics\

จะต้องกำหนดเส้นทางของคลาสเพิ่มเติมโดยใช้ประโยคข้างล่างนี้ลงในไฟล์ autoexec.bat

```
Set class path = C:\statistics\statistics.jar
```

5.2 ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นประกอบด้วยส่วนประกอบย่อยๆดังนี้

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5.2.1 ส่วนประกอบสำหรับการนับค่าความถี่ของข้อมูล

ส่วนประกอบสำหรับการนับค่าความถี่ของข้อมูลจัดการได้ด้วยคลาส FreqDistribution จำแนกได้เป็น 2 ประเภทคือ

- การนับค่าความถี่ของข้อมูลในกรณีข้อมูลเป็นแบบจัดกลุ่มจัดการได้ด้วยคลาส GroupFreqDistribution
- การนับค่าความถี่ของข้อมูลในกรณีข้อมูลเป็นแบบไม่ได้จัดกลุ่มจัดการได้ด้วยคลาส UgroupFreqDistribution

5.2.2 ส่วนประกอบสำหรับการคำนวณสถิติแสดงค่ากลางของข้อมูล

ส่วนประกอบสำหรับการคำนวณสถิติแสดงค่ากลางของข้อมูลจัดการได้ด้วยคลาส CentralTendency จำแนกได้เป็น 2 ประเภทคือ

- สถิติแสดงค่ากลางของข้อมูลในกรณีข้อมูลเป็นแบบจัดกลุ่มจัดการได้ด้วยคลาส GroupCentralTendency
- สถิติแสดงค่ากลางของข้อมูลในกรณีข้อมูลเป็นแบบไม่ได้จัดกลุ่มจัดการได้ด้วยคลาส UngroupCentralTendency

5.2.3 ส่วนประกอบสำหรับสำหรับคำนวณสถิติที่วัดการกระจายของข้อมูล

ส่วนประกอบสำหรับสำหรับคำนวณสถิติที่วัดการกระจายของข้อมูล จัดการได้ด้วยคลาส Dispersion จำแนกได้เป็น 2 ประเภทคือ

- การกระจายของข้อมูลในกรณีข้อมูลเป็นแบบจัดกลุ่มจัดการได้ด้วยคลาส GroupDispersion
- การกระจายของข้อมูลในกรณีข้อมูลเป็นแบบไม่ได้จัดกลุ่มจัดการได้ด้วยคลาส UngroupDispersion

5.2.4 ส่วนประกอบสำหรับสำหรับวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีเอนเตอร์

ส่วนประกอบสำหรับสำหรับวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีเอนเตอร์ จัดการได้ด้วยคลาส Enter ซึ่งเป็นคลาสที่สืบทอดคุณสมบัติมาจากคลาส LinearRegression ประกอบด้วยคลาสที่คำนวณค่าสถิติเกี่ยวกับการวิเคราะห์ความถดถอยเชิงเส้นได้แก่

- คลาส ModelSummary คือคลาสสำหรับการคำนวณสถิติสรุปของสมการถดถอย
- คลาส Correlations คือคลาสสำหรับการคำนวณความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว
- คลาส Anova คือคลาสสำหรับการคำนวณความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น
- คลาส Coefficients คือคลาสสำหรับการคำนวณค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น
- คลาส ResidualsStatistics คือคลาสสำหรับการคำนวณค่าสถิติของค่าคลาดเคลื่อน

5.2.5 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีฟอร์เวิร์ด

การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีฟอร์เวิร์ด จัดการได้ด้วยคลาส Forward ซึ่ง ประกอบด้วยคลาสที่สืบทอดคุณสมบัติมาจากคลาส LinearRegression เช่นเดียวกับข้อ 5.2.4

5.2.6 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีแบ็คเวิร์ด

การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีแบ็คเวิร์ด จัดการได้ด้วยคลาส Backward ซึ่งประกอบด้วยคลาสที่สืบทอดคุณสมบัติมาจากคลาส LinearRegression เช่นเดียวกับข้อ 5.2.4

5.2.7 การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีสเตปไวส์

การวิเคราะห์ความถดถอยเชิงเส้นโดยการเลือกตัวแปรอิสระเข้าสมการความถดถอยด้วยวิธีสเตปไวส์ จัดการได้ด้วยคลาส Stepwise ซึ่งประกอบด้วยคลาสที่สืบทอดคุณสมบัติมาจากคลาส LinearRegression เช่นเดียวกับข้อ 5.2.4

5.3 การใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

การนำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นไปใช้ร่วมในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆ จะต้องประกาศให้โปรแกรมรู้จักคลาสของส่วนประกอบซอฟต์แวร์เชิงวัตถุดังนี้

- ในการพัฒนาโปรแกรมเชิงวัตถุหากต้องการเรียกใช้คลาสสำหรับการวิเคราะห์สถิติเชิงพรรณนาจะต้องระบุให้โปรแกรมรู้จักคลาสต่างๆสำหรับการวิเคราะห์สถิติเชิงพรรณนาโดยใช้คำสั่งดังนี้

```
import    Statistics.DescriptiveStatistics.*;
```

- ในการพัฒนาโปรแกรมเชิงวัตถุหากต้องการเรียกใช้คลาสสำหรับการวิเคราะห์ความถดถอยเชิงเส้นจะต้องระบุให้โปรแกรมรู้จักคลาสต่างๆ สำหรับการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้คำสั่งดังนี้

```
import    Statistics.LinearRegression.*;
```

หรือหากต้องการเรียกใช้เฉพาะส่วนประกอบย่อยๆ เช่นต้องการคำนวณค่ากลางของข้อมูลแบบไม่จัดกลุ่มก็สามารถใช้คำสั่งดังนี้

```
import    Statistics.DescriptiveStatistics.UngroupCentralTendency;
```

5.4 ตัวอย่างการใช้งานส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ผู้วิจัยได้จัดทำโปรแกรม Stat2000 ซึ่งเป็นโปรแกรมที่เรียกใช้ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น เพื่อทดสอบความถูกต้องของส่วนประกอบซอฟต์แวร์ที่ได้พัฒนาขึ้น ตัวอย่างการเรียกใช้ส่วนประกอบซอฟต์แวร์สำหรับโปรแกรม Stat200 มีรายละเอียดดังนี้

5.4.1 การกำหนดชื่อคลาสที่จะใช้งาน

การกำหนดชื่อคลาสที่จะใช้งานสามารถกำหนดได้โดยใช้คำสั่ง import ซึ่งช่วยให้คอมไพเลอร์มองหาคลาสนั้นพบและสามารถอ้างถึงคลาสในขณะทำงานได้อย่างถูกต้อง โดยมีตัวอย่างคำสั่งดังนี้

```
import Statistics.LinearRegression.*;
```

คำสั่งนี้จะทำให้สามารถอ้างถึงทุกคลาสในแพ็กเกจ Statistics.LinearRegression ได้

5.4.2 การสร้างอินสแตนซ์ของคลาส

การสร้างอินสแตนซ์ของคลาสสามารถสร้างโดยใช้คำสั่ง new ดังตัวอย่างข้างล่างนี้

```
Enter enter  = new Enter(ImportFile.getColumnData(dependentItem),
 ImportFile.getMultiColumn(independentItems),independentItems,
 dependentItem,resInd);
```

จากคำสั่งด้านบนจะได้ enter เป็นอินสแตนซ์ของคลาส Enter ซึ่งทำให้ enter สามารถเรียกใช้ฟังก์ชันเพื่อรับค่าที่ได้จากการคำนวณการวิเคราะห์ความถดถอยเชิงเส้นโดยวิธี Enter ได้

5.4.3 การรับค่าคุณสมบัติของอินสแตนซ์

การรับค่าคุณสมบัติของอินสแตนซ์สามารถทำได้โดยเรียกเมทธอดของคลาส Enter ดังตัวอย่างข้างล่างนี้

```
String [] enteredVar = enter.getEnteredVar();      //Entered Variables
String [] removedVar = enter.getRemovedVar();  //Removed Variables
```

เมทธอดด้านบนเป็นเมทธอดในการรับค่าของตัวแปรที่นำเข้าและตัวออกจากสมการความถดถอยนอก จากนี้แล้วสำหรับค่าสถิติอื่นๆ ได้แก่ สถิติสรุปของสมการถดถอย ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น และค่าสถิติของค่าคลาดเคลื่อน ก็สามารถรับค่าได้โดยใช้คำสั่งในทำนองเดียวกัน

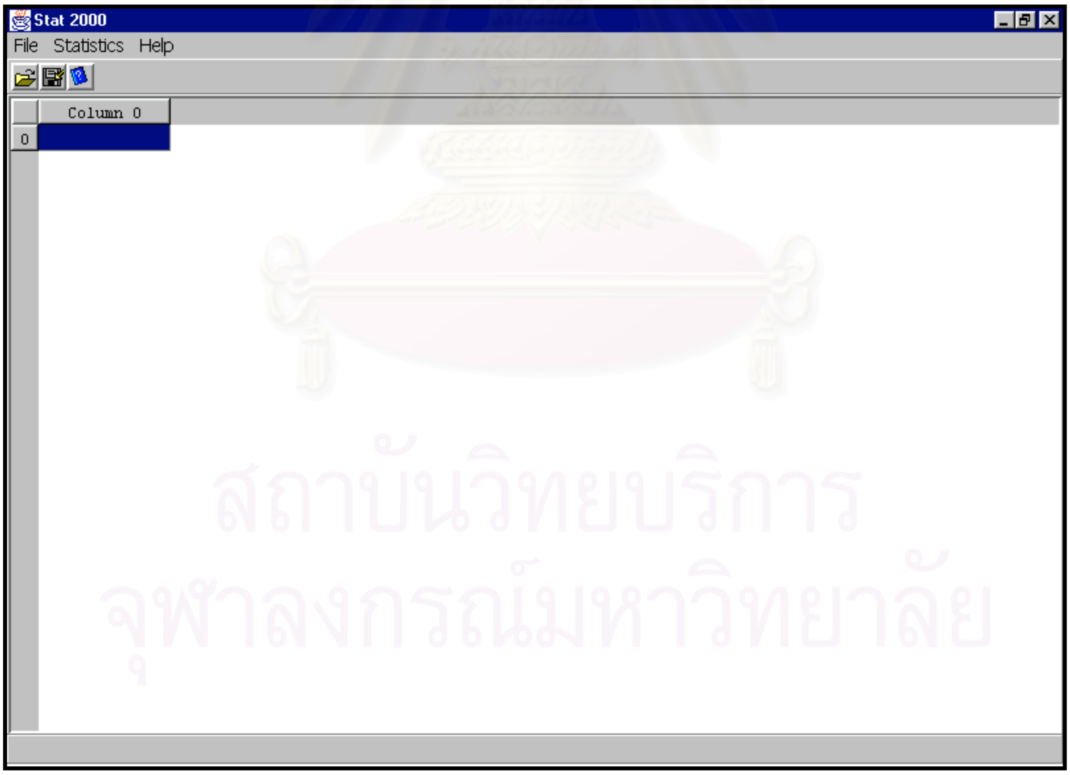
บทที่ 6

การใช้งานซอฟต์แวร์สำหรับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

ในบทนี้กล่าวถึงการใช้งานโปรแกรมในการทดสอบส่วนประกอบซอฟต์แวร์ที่เรียกว่า "โปรแกรม Stat2000" ซึ่งเป็นโปรแกรมสำหรับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น โปรแกรมที่จัดทำขึ้นเพื่อทดสอบการเรียกใช้ส่วนประกอบซอฟต์แวร์เชิงวัตถุ พร้อมทั้งอธิบายความหมายทางสถิติในแต่ละขั้นตอนอย่างละเอียด

6.1 การเข้าสู่โปรแกรมสำหรับการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น

การเข้าสู่โปรแกรมเริ่มต้นจากเลือกไอคอน Stat2000 จะปรากฏหน้าจอหลักของโปรแกรมดังแสดงในรูปที่ 6.1



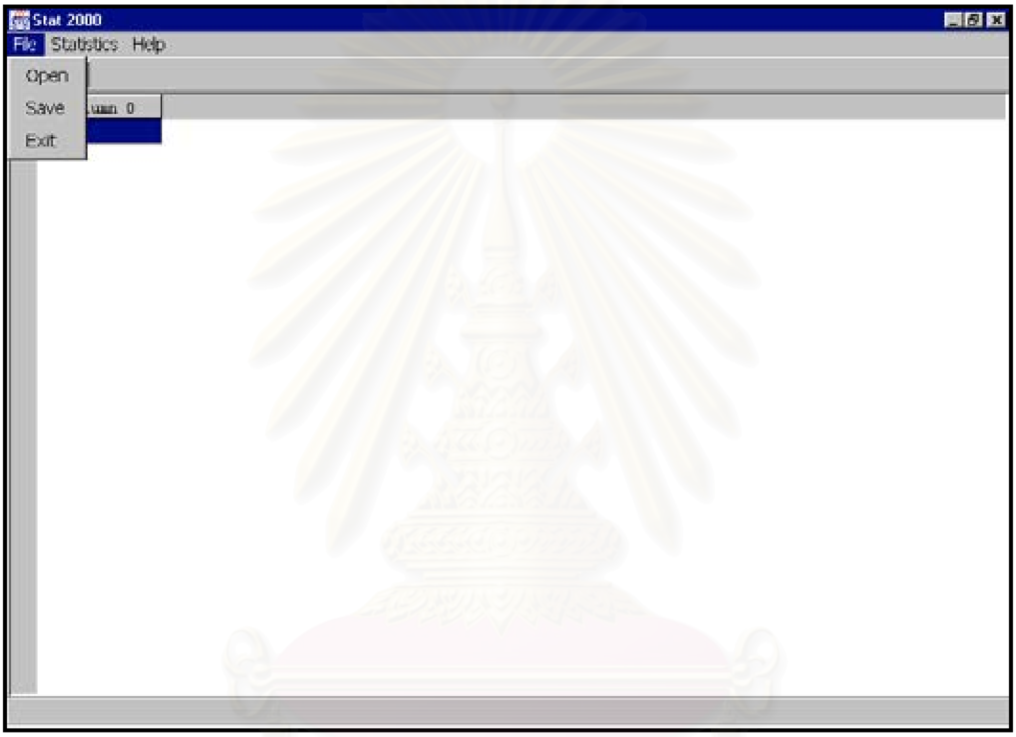
รูปที่ 6.1 หน้าจอเมนูหลักโปรแกรม Stat 2000

หน้าจอหลักประกอบด้วยเมนู 3 เมนูคือ เมนู File เมนู Statistics และเมนู Help โดยในแต่ละเมนูมีรายละเอียดต่างๆดังนี้

6.1.1 เมนู File

เมนู Fileประกอบด้วยคำสั่ง Open Save และ Exit ดังแสดงในรูปที่ 6.2 โดยในแต่ละคำสั่งมีความหมายดังนี้

1. คำสั่ง Open คำสั่งนี้เป็นคำสั่งที่ใช้สำหรับการเปิดแฟ้มข้อมูลเพื่อการวิเคราะห์ทางสถิติ
2. คำสั่ง Save คำสั่งนี้เป็นคำสั่งที่ใช้สำหรับการบันทึกแฟ้มข้อมูล
3. คำสั่ง Exit คำสั่งนี้ใช้สำหรับออกจากโปรแกรม Stat2000

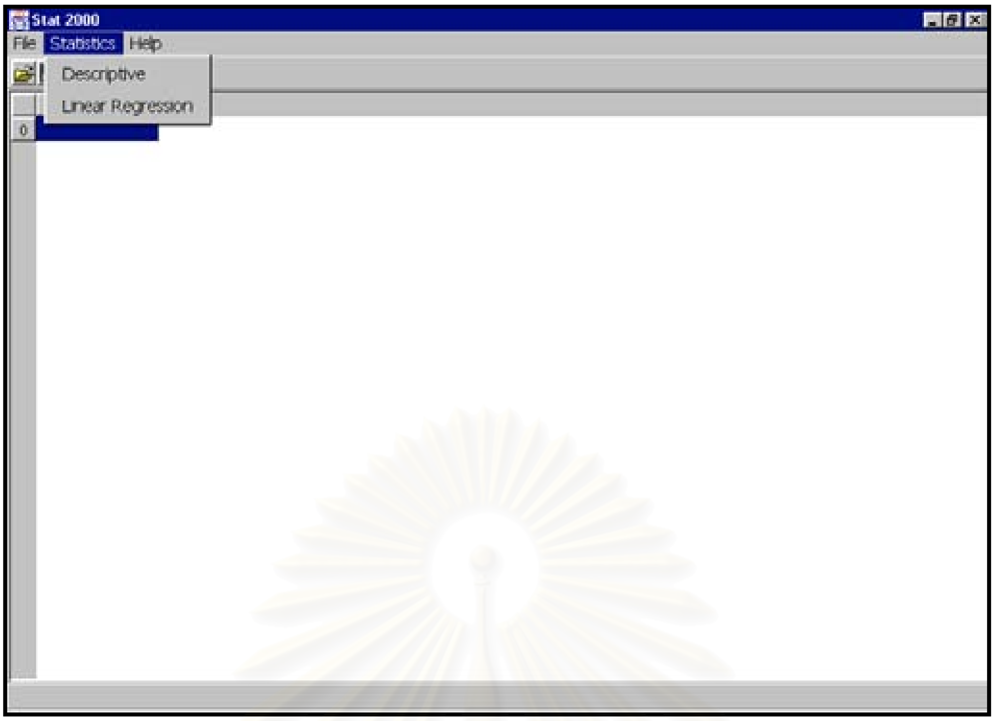


รูปที่ 6.2 รายการต่างๆ ของเมนู File

6.1.2 เมนู Statistics

เมนู Statistics ประกอบด้วยคำสั่ง Descriptive และ Linear Regression ดังแสดงในรูปที่ 6.3 โดยในแต่ละคำสั่งมีความหมายดังนี้

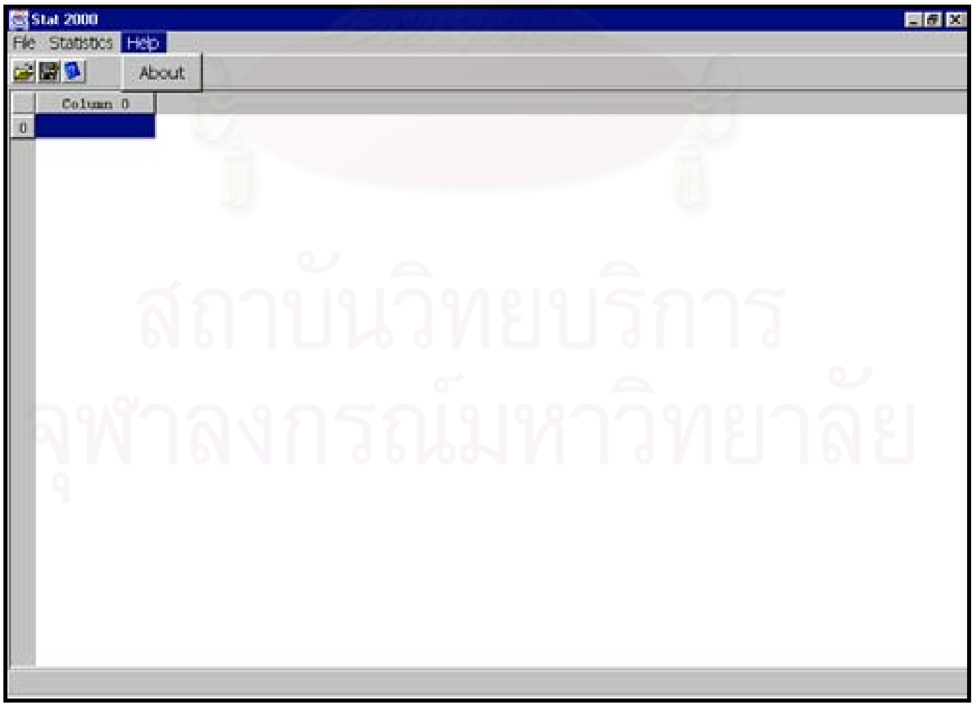
1. คำสั่ง Descriptive เป็นคำสั่งที่ใช้ในกรณีที่ต้องการวิเคราะห์ค่าสถิติเชิงพรรณนา
2. คำสั่ง Linear Regression เป็นคำสั่งที่ใช้ในกรณีที่ต้องการวิเคราะห์ความถดถอยเชิงเส้น



รูปที่ 6.3 รายการต่างๆของเมนู Statistics

6.1.3 เมนู Help

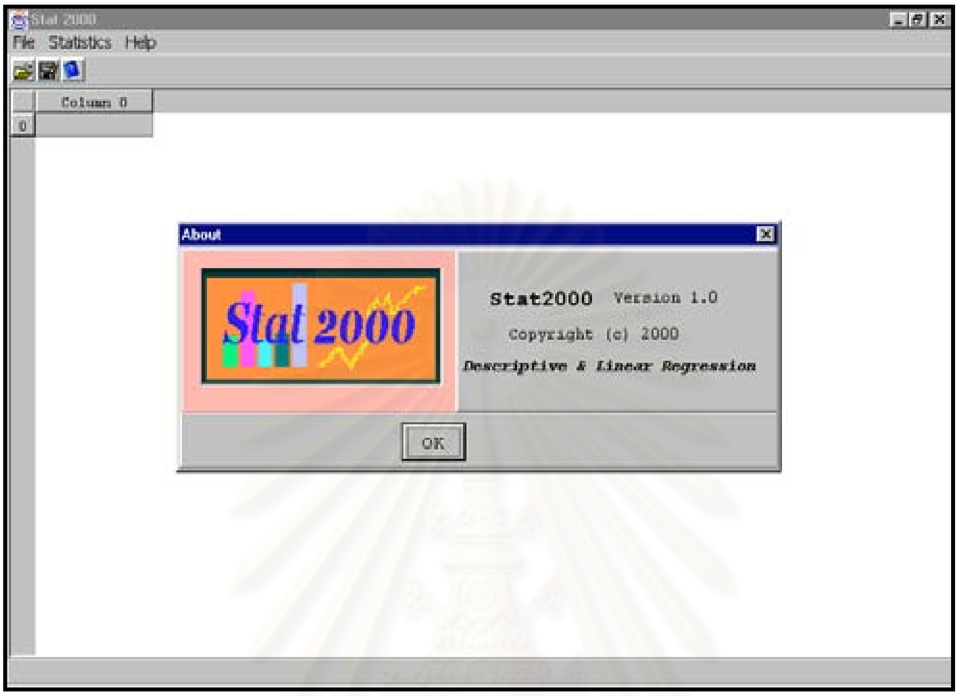
เมนู Help มีเมนูย่อย About ดังแสดงในรูปที่ 6.4 เมนูย่อย About เป็นคำสั่งสำหรับซึ่งใช้สำหรับดูรายละเอียดเกี่ยวกับโปรแกรม



รูปที่ 6.4 รายการของเมนู Help

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ผู้ใช้สามารถดูรายละเอียดของโปรแกรม Stat2000 ได้ โดยใช้คำสั่ง Help -> About จะปรากฏหน้าจอ About ดังแสดงในรูปที่ 6.5 หน้าจอนี้เป็นหน้าจอแสดงรายละเอียดเกี่ยวกับโปรแกรมประกอบด้วย ชื่อโปรแกรม เวอร์ชัน และรายละเอียดของโปรแกรม



รูปที่ 6.5 หน้าจอ About

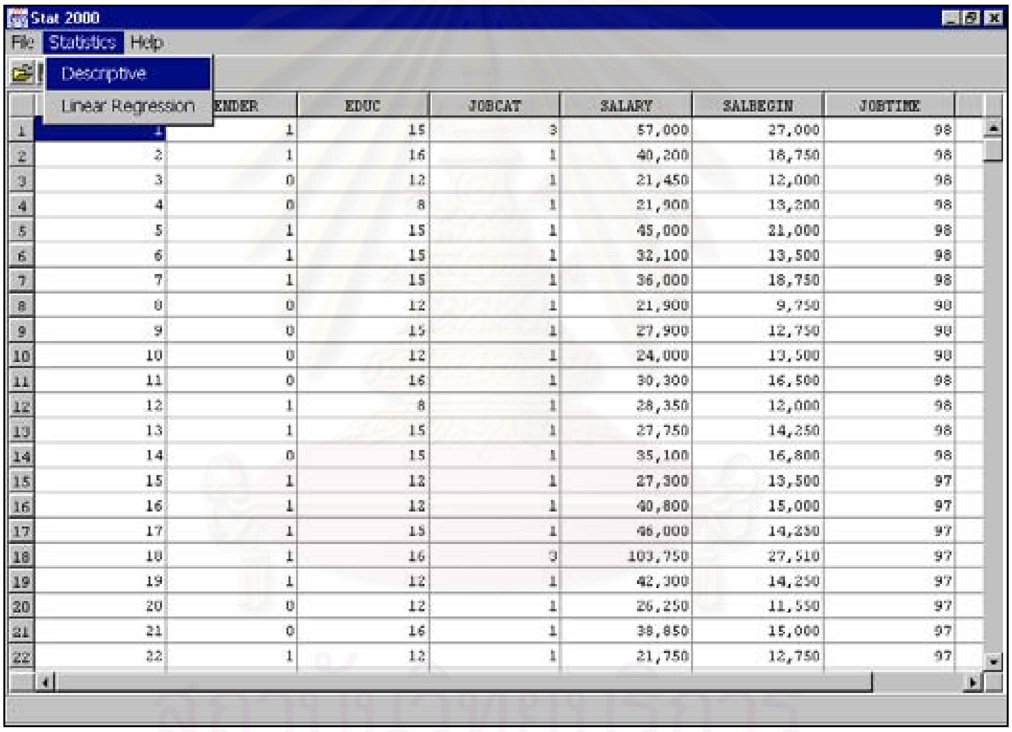
6.2 การวิเคราะห์สถิติเชิงพรรณนา

การวิเคราะห์สถิติเชิงพรรณนาเป็นหลักการในการบรรยายลักษณะของข้อมูลที่ได้เก็บรวบรวม ซึ่งไม่สามารถอ้างอิงถึงลักษณะของประชากรได้ จึงเป็นการสรุปลักษณะเฉพาะที่สำคัญของข้อมูลของกลุ่มที่ศึกษาเท่านั้น สถิติเชิงพรรณนาประกอบด้วย การวัดค่ากลางของข้อมูล การวัดค่าการกระจายของข้อมูล การนับค่าความถี่ของข้อมูล รวมทั้งการนำเสนอข้อมูลค่าความถี่ของข้อมูลในรูปของกราฟความถี่

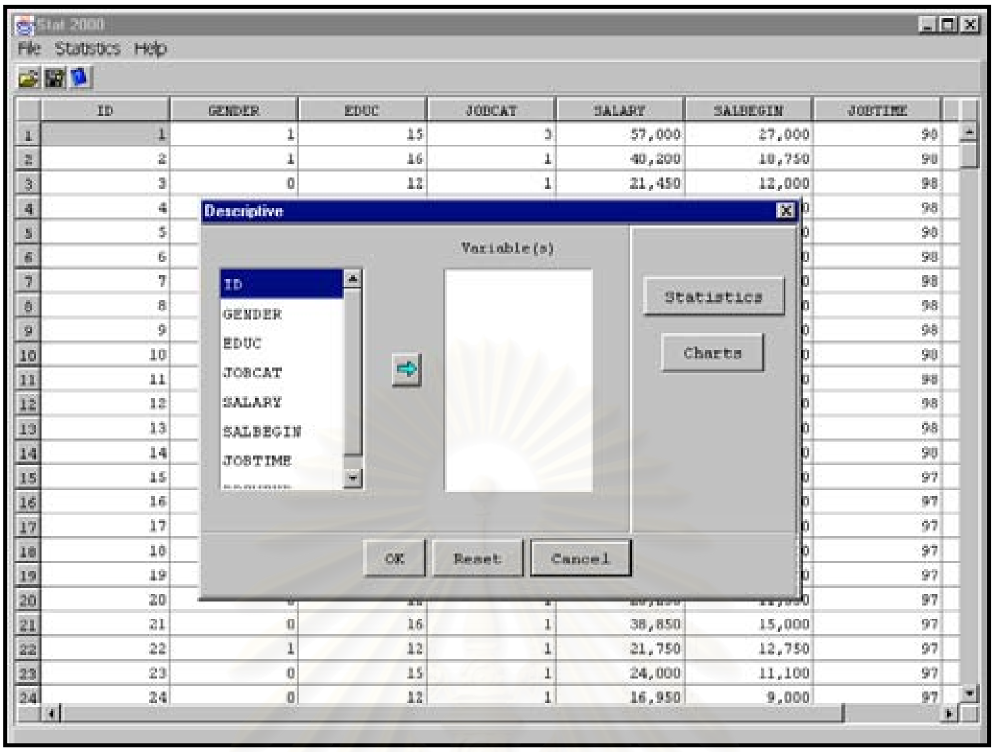
ในการคำนวณสถิติเชิงพรรณนา ผู้ใช้จะต้องเปิดแฟ้มข้อมูลก่อน แล้วจึงใช้คำสั่ง Statistics -> Descriptive
ดังแสดงในรูปที่ 6.6 และเมื่อใช้คำสั่ง Statistics -> Descriptive จะปรากฏหน้าจอดังรูปที่ 6.7 หน้าจอในรูปที่ 6.7 ประกอบด้วยส่วนต่างๆดังนี้

1. Variable(s) — ส่วนนี้แสดงตัวแปรเชิงปริมาณ(ชนิดตัวเลข)ที่ต้องการคำนวณค่าสถิติเชิงพรรณนาโดยสามารถเลือกตัวแปรได้มากกว่า 1 ตัวแปร
2. ปุ่ม Statistics — สำหรับให้ผู้ใช้เลือกเพื่อกำหนด ค่าสถิติที่ต้องการให้แสดงผลลัพธ์

3. ปุ่ม ลูกศรชี้ไปทางขวา — สำหรับเลือกตัวแปรที่ต้องการวิเคราะห์ค่าสถิติเชิงพรรณนา เมื่อเลือกตัวแปรในช่องด้านซ้ายซึ่งเป็นช่องแสดงรายชื่อตัวแปรทั้งหมดของแฟ้มที่เปิดและกดปุ่ม ตัวแปรที่เลือกจะย้ายมายังช่องแสดงรายการตัวแปรด้านขวาซึ่งเป็นช่องแสดงรายการของตัวแปรที่ต้องการวิเคราะห์ค่าสถิติเชิงพรรณนา
4. ปุ่ม Chart — สำหรับให้ผู้ใช้กำหนดประเภทของกราฟความถี่
5. ปุ่ม OK — สำหรับเลือกเพื่อการประมวลผลค่าสถิติเชิงพรรณนาและแสดงหน้าจอผลลัพธ์ค่าสถิติเชิงพรรณนา
6. ปุ่ม Reset — สำหรับเลือกเพื่อยกเลิกตัวแปรต่างๆทั้งหมดที่ได้เลือกไว้
7. ปุ่ม Cancel — สำหรับเลือกเพื่อยกเลิกการดำเนินการ และกลับไปเมนูหลัก



รูปที่ 6.6 การเลือกรายการวิเคราะห์สถิติเชิงพรรณนา (Descriptive)



รูปที่ 6.7 หน้าจอกำหนดตัวแปรในการวิเคราะห์สถิติเชิงพรรณนา

6.2.1 การกำหนดค่าสถิติสำหรับการวิเคราะห์สถิติเชิงพรรณนา

การกำหนดค่าสถิติสำหรับการวิเคราะห์สถิติเชิงพรรณนาเป็นการกำหนดค่าสถิติที่ต้องการให้แสดงผลลัพธ์ซึ่งเป็นค่าพยากรณ์ที่ได้จากการวิเคราะห์สถิติเชิงพรรณนาจากคำสั่ง Statistics -> Descriptive เมื่อกดปุ่ม Statistics จะได้หน้าจอดังรูปที่ 6.8 ซึ่งประกอบด้วยส่วนต่างๆดังนี้

1. Percentile Value : ผู้ใช้สามารถเลือกทางเลือกต่อไปนี้ ได้หลายๆทางเลือก ดังนี้
   - Quartiles เลือกเพื่อแสดงค่าเปอร์เซ็นต์ไทล์ที่ 25, 50, 75 และ 100 ตามลำดับ
   - Deciles เลือกเพื่อแสดงค่าเปอร์เซ็นต์ไทล์ที่ 10, 20, 30, 40, 50, 60, 70, 80, 90, และ 100 ตามลำดับ
   - Percentile(s) ผู้ใช้สามารถระบุค่าเปอร์เซนต์ไทล์ได้หลายค่าซึ่งเป็นเลขจำนวนเต็มที่มีค่าระหว่าง 1 ถึง 100โดยจะต้องระบุทีละค่าลงในช่องสี่เหลี่ยมด้านขวา
2. Central Tendency เป็นสถิติแสดงค่ากลางของข้อมูล ซึ่งประกอบด้วย
   - Mean เลือกเพื่อแสดงค่าเฉลี่ย
   - Median เลือกเพื่อแสดงค่ามัธยฐาน
   - Mode เลือกเพื่อแสดงฐานนิยม
   - Sum เลือกเพื่อแสดงผลบวกของข้อมูล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

3. Dispersion เป็นค่าสถิติที่วัดการกระจายของข้อมูล ซึ่งผู้วิเคราะห์เลือกได้หลายทางเลือกดังนี้

    Std. Deviation เลือกเพื่อแสดงค่าเบี่ยงเบนมาตรฐาน

    Variance เลือกเพื่อแสดงค่าแปรปรวน

    Range เลือกเพื่อแสดงค่าพิสัย

    Minimum เลือกเพื่อแสดงค่าต่ำสุดของข้อมูล

    Maximum เลือกเพื่อแสดงค่าสูงสุดของข้อมูล

    S.E. mean เลือกเพื่อแสดงค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยตัวอย่าง

4. Distribution เป็นค่าสถิติที่แสดงการแจกแจงของข้อมูล

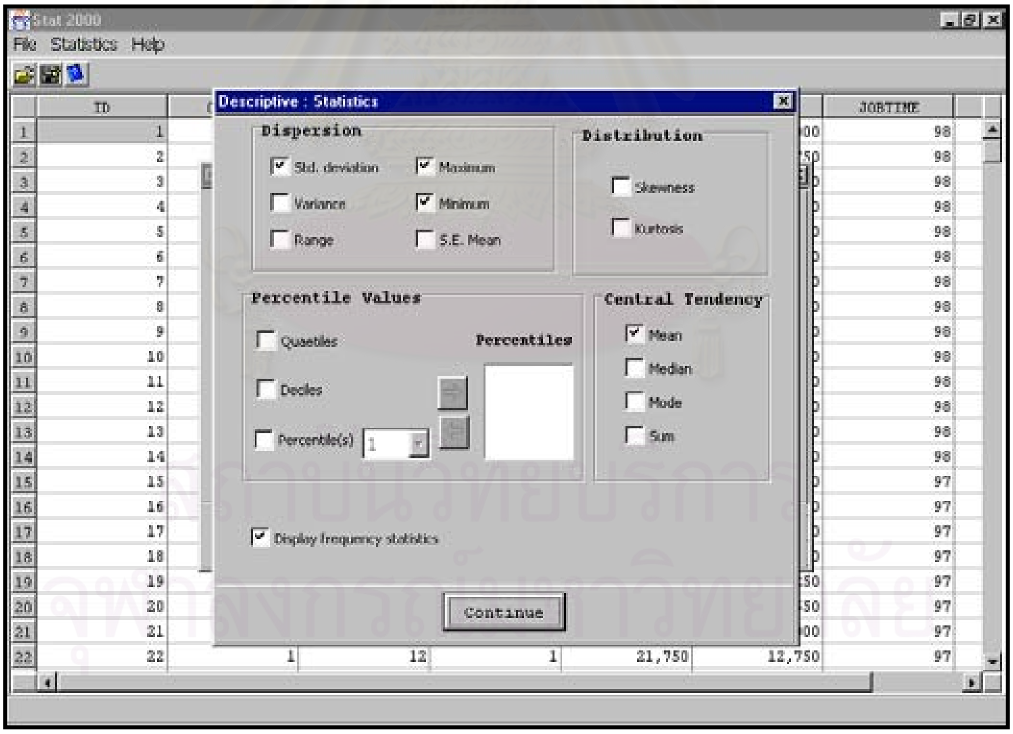
    Skewness เลือกเพื่อแสดงค่าความเบ้ของการแจกแจงที่ไม่สมมาตร

    Kurtosis เลือกเพื่อแสดงค่าความโด่งของการแจกแจงที่ไม่สมมาตร

5. Display Frequency Tables ส่วนนี้เลือกเพื่อแสดงค่าความถี่ของข้อมูล

6. ปุ่ม Continue ปุ่มสำหรับเลือกเพื่อดำเนินการต่อ โดยกำหนดการเลือกค่าสถิติเชิงพรรณนา

    ตามที่ผู้ใช้ได้กำหนดไว้และกลับไปยังหน้าจอในรูปที่ 6.6



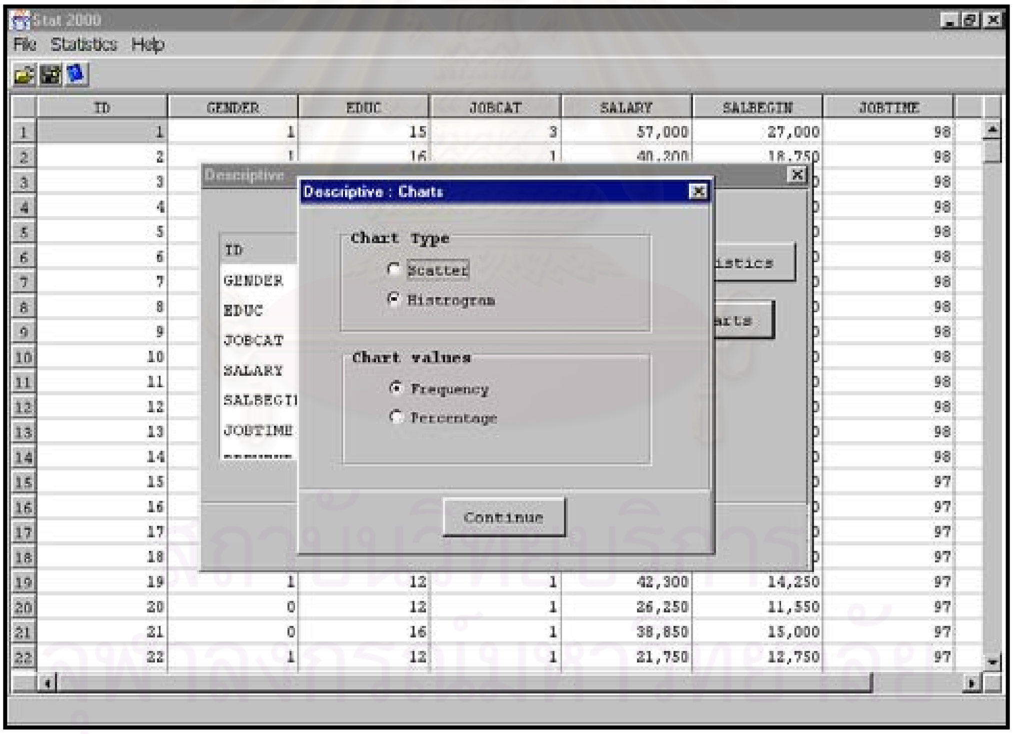
รูปที่ 6.8 หน้าจอสำหรับกำหนดค่าสถิติสำหรับการวิเคราะห์สถิติเชิงพรรณนา

6.2.2 การกำหนดค่าแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา

การกำหนดค่าแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนาเป็นการกำหนดประเภทและค่าที่ใช้สำหรับการแสดงผลในรูปแผนภูมิ จากคำสั่ง Statistics -> Descriptive เมื่อกดปุ่ม Charts จะปรากฏหน้าจอดังรูปที่ 6.9

หน้าจอในรูปที่ 6.9 ประกอบด้วยส่วนต่างๆดังนี้

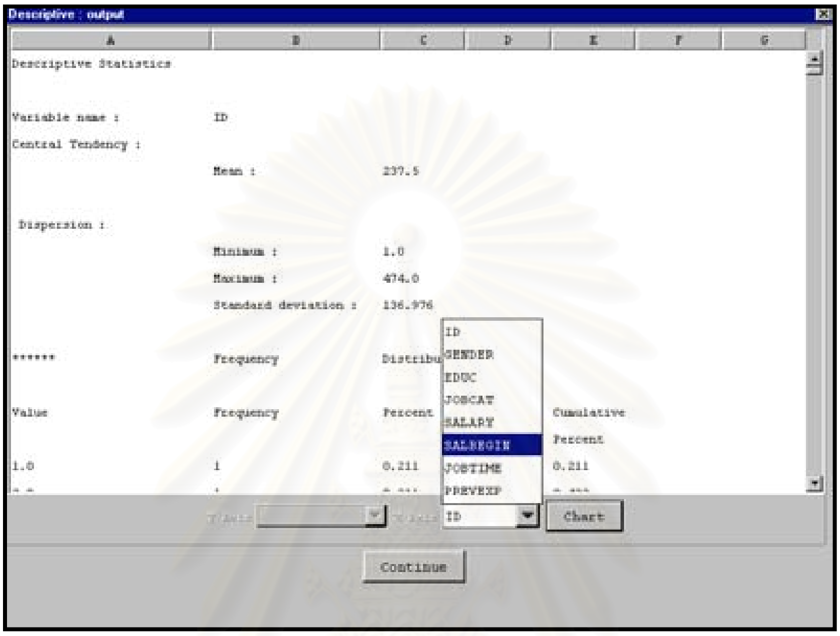
1. Chart Type ประเภทของแผนภูมิ ผู้ใช้สามารถเลือกได้ 2 อย่างคือ ฮิสโตแกรมและแผนภาพการกระจาย (Scatter Chart)
2. Chart Values ค่าที่แสดงบนแผนภูมิ ผู้ใช้สามารถเลือกได้ 2 อย่างคือ ค่าความถี่ของข้อมูล(Frequencies) และค่าร้อยละของข้อมูล (Percentages)
3. ปุ่ม Continue ปุ่มสำหรับเลือกเพื่อดำเนินการต่อ โดยกำหนดการเลือกตามที่ผู้ใช้ได้กำหนดไว้และกลับไปยังหน้าจอในรูปที่ 6.7



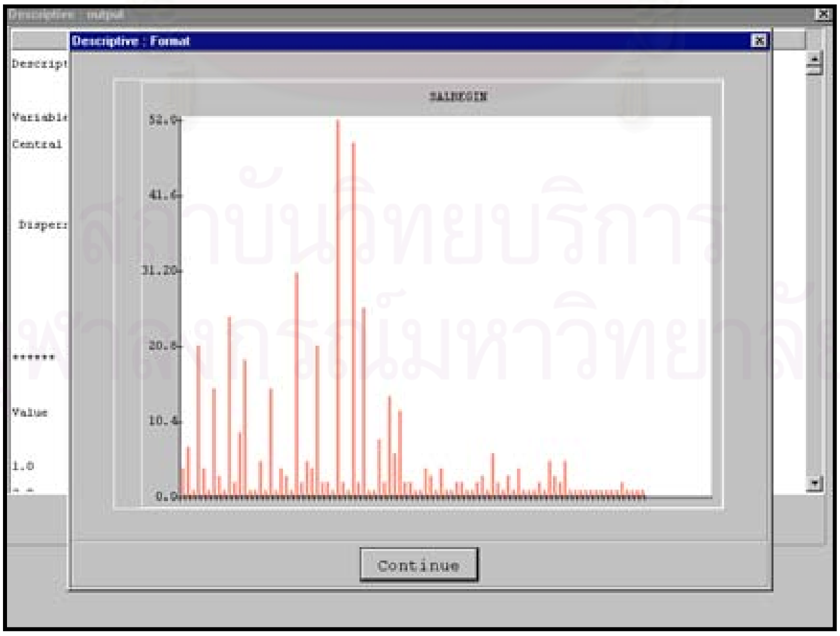
รูปที่ 6.9 หน้าจอกำหนดค่าแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา

6.2.3 การแสดงแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา

การแสดงแผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา สามารถเลือกการแสดงค่าจากรายการของตัวแปรซึ่งเป็นตัวแปรที่ทำการวิเคราะห์สถิติเชิงพรรณนาดังแสดงในรูปที่ 6.10 เมื่อเลือกตัวแปรแล้วกดปุ่ม Chart จะปรากฏหน้าจอแสดงแผนภูมิดังรูปที่ 6.11



รูปที่ 6.10 แสดงการเลือกตัวแปรที่ต้องการแสดงผลลัพธ์ในรูปแผนภูมิ



รูปที่ 6.11 แสดงผลลัพธ์แผนภูมิสำหรับการวิเคราะห์สถิติเชิงพรรณนา

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

6.3 การวิเคราะห์ความถดถอยเชิงเส้น

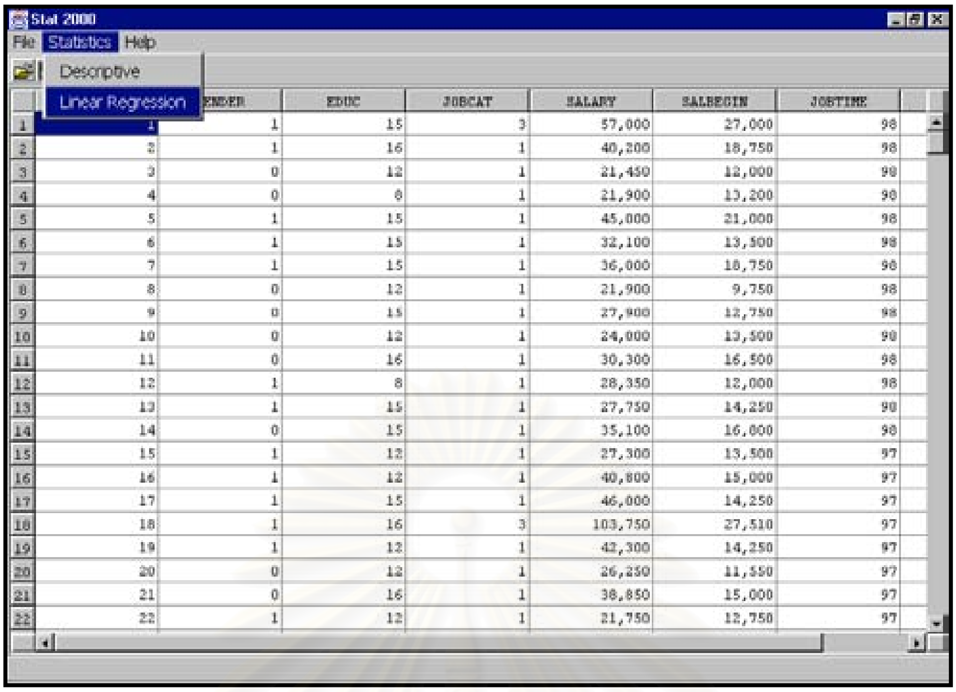
การวิเคราะห์ความถดถอยเป็นการศึกษาความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไป โดยมีวัตถุประสงค์ที่จะพยากรณ์ค่าของตัวแปรตัวหนึ่งจากตัวแปรอื่นๆ ที่มีความสัมพันธ์กับตัวแปรที่ต้องการพยากรณ์ โดยการเลือกตัวแปรอิสระเข้าสมการความถดถอย ผู้วิเคราะห์จะต้องพิจารณาว่าตัวแปรอิสระใดบ้างที่น่าจะมีความสัมพันธ์กับตัวแปรตาม สำหรับเทคนิคที่ใช้ในการพิจารณาว่าตัวแปรอิสระใดบ้างที่มีความสัมพันธ์กับตัวแปรตามมี 4 วิธีดังนี้

1. วิธีเอนเตอร์ เป็นวิธีการนำตัวแปรอิสระทุกตัวเข้าสมการความถดถอย
2. วิธีแบ็คเวิร์ด เป็นวิธีการตัดตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามออกจากสมการความถดถอยเชิงเส้นทีละตัว
3. วิธีฟอร์เวิร์ด เป็นวิธีการนำตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามเข้าสมการความถดถอยเชิงเส้นทีละตัว
4. วิธีสเตปไวส์ เป็นวิธีการนำตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด และในขณะเดียวกัน ก็พิจารณาตัวตัวแปรอิสระที่ไม่มีความสัมพันธ์กับตัวแปรตามออกจากสมการโดยวิธีแบ็คเวิร์ด

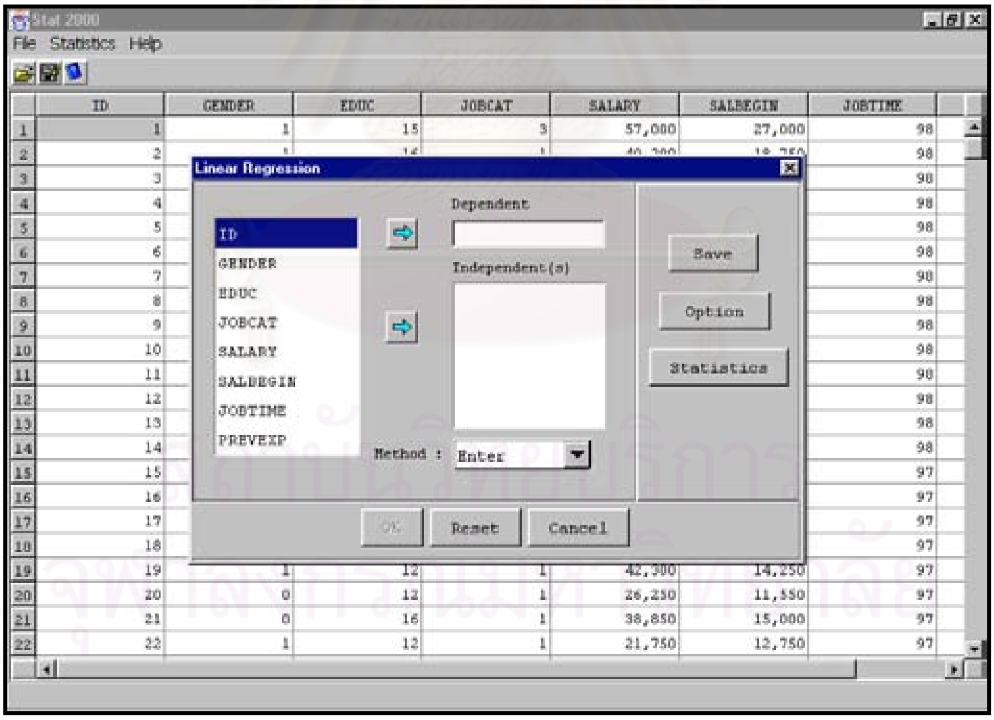
ในการวิเคราะห์ความถดถอยเชิงเส้น ผู้ใช้จะต้องเปิดแฟ้มข้อมูลก่อน แล้วจึงใช้คำสั่ง Statistics -> Linear Regression ดังแสดงในรูปที่ 6.12 และเมื่อใช้คำสั่ง Statistics -> Linear Regression จะได้หน้าจอดังรูปที่ 6.13 ซึ่งประกอบด้วยส่วนต่างๆดังนี้

1.	Dependent	สำหรับเลือกตัวแปรอิสระที่มีค่าเป็นตัวเลข 1 ตัวแปร
2.	Independent(s)	สำหรับเลือกตัวแปรอิสระอย่างน้อย 1 ตัวแปร
3.	Method	ระบุวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยซึ่งมี 4 วิธี คือ วิธีเอนเตอร์ วิธีฟอร์เวิร์ด วิธีแบ็คเวิร์ด และวิธีสเตปไวส์
4.	ปุ่ม Statistics	ปุ่มสำหรับให้ผู้ใช้เลือกเพื่อกำหนดค่าสถิติที่ต้องการให้แสดงผลลัพธ์
5.	ปุ่ม Save	ปุ่มสำหรับให้ผู้ใช้เลือกเพื่อกำหนดค่าสถิติที่ต้องการบันทึกค่าประกอบด้วยค่าสถิติต่างๆดังจะกล่าวรายละเอียดในหัวข้อที่ 6.3.2
6.	ปุ่ม OK	ปุ่มสำหรับเลือกเพื่อคำนวณค่าสถิติความถดถอยเชิงเส้นและแสดงหน้าจอผลลัพธ์ของการวิเคราะห์ความถดถอยเชิงเส้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 6.12 การเลือกรายการการวิเคราะห์ความถดถอยเชิงเส้น

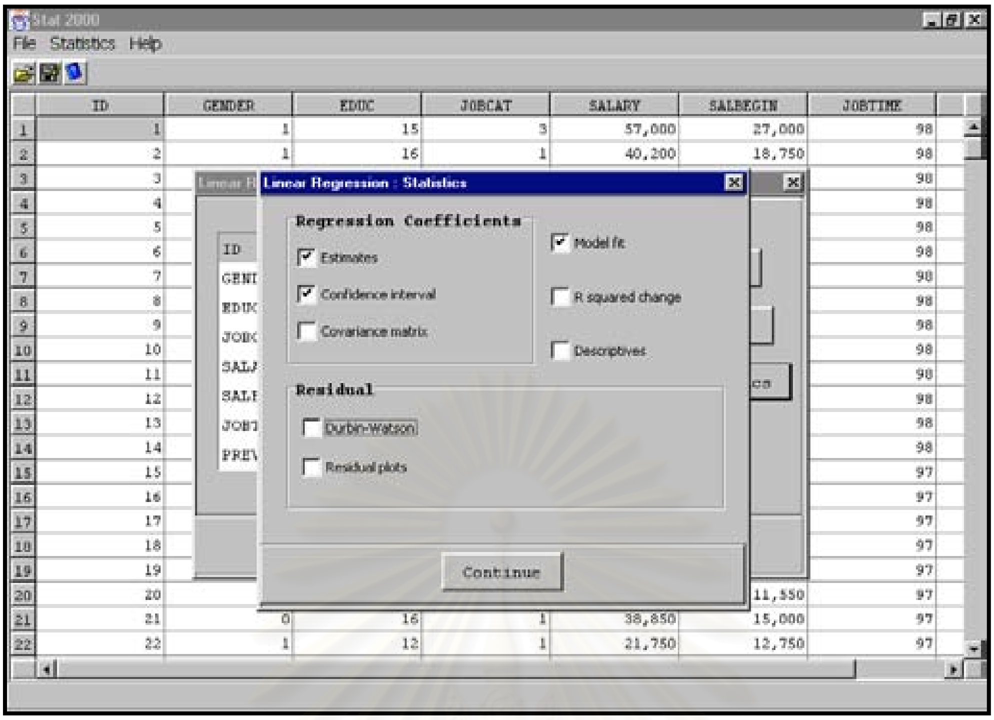


รูปที่ 6.13 หน้าจอกำหนดตัวแปรในการวิเคราะห์ความถดถอยเชิงเส้น

6.3.1 การกำหนดค่าสถิติสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

การกำหนดค่าสถิติสำหรับการวิเคราะห์ความถดถอยเชิงเส้นเป็นการกำหนดค่าสถิติที่ต้องการให้แสดงผลลัพธ์ซึ่งได้จากการพยากรณ์โดยใช้วิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยแบบเอนเตอร์ ฟอร์เวิร์ด แบ็คเวิร์ด หรือ สเตปไวส์ จากคำสั่ง Statistics -> Linear Regression เมื่อกดปุ่ม Statistics จะปรากฏหน้าจอดังรูปที่ 6.14 ซึ่งประกอบด้วย

1. Regression Coefficients — เลือกเพื่อแสดงค่าประมาณของสัมประสิทธิ์ความถดถอย ค่าคลาดเคลื่อนมาตรฐานของ b ค่าคลาดเคลื่อนมาตรฐานของค่าคงที่และสัมประสิทธิ์ความถดถอย ค่าสัมประสิทธิ์ความถดถอยมาตรฐาน ค่าสถิติทดสอบ t ค่านัยสำคัญของสถิติทดสอบ t สำหรับการทดสอบ 2 ข้าง การประมาณค่าแบบช่วงของ b_i และค่า เมตริกซ์ความสัมพันธ์ระหว่างตัวแปร
2. Model fit — เลือกเพื่อแสดงตัวแปรที่นำเข้าและเอาออกจากสมการความถดถอย และแสดงค่า สัมประสิทธิ์สหสัมพันธ์ ค่าสัมประสิทธิ์การตัดสินใจ ค่าสัมประสิทธิ์การตัดสินใจที่ปรับแล้ว ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่า และความแปรปรวนของวิเคราะห์ความถดถอยเชิงเส้น
3. R^2 Change — เลือกเพื่อแสดงค่าที่เปลี่ยนแปลงไป ของ สัมประสิทธิ์การตัดสินใจ ค่าสถิติทดสอบ F และ ค่านัยสำคัญของสถิติทดสอบ F
4. Descriptive — เลือกเพื่อแสดงค่าสถิติเชิงพรรณนาได้แก่ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐานของแต่ละตัวแปร ค่าเมตริกซ์ของสัมประสิทธิ์สหสัมพันธ์ที่คำนวณค่า ค่านัยสำคัญของการทดสอบแบบข้างเดียวเกี่ยวกับสัมประสิทธิ์สหสัมพันธ์
5. Residual — เลือกเพื่อแสดงค่าสถิติเดอร์บิน-วัทสัน สำหรับทดสอบความเป็นอิสระกันของค่าคลาดเคลื่อน และแสดงถึงตัวแปรที่มีค่าสูง หรือต่ำผิดปกติ
6. Residual plots — เลือกเพื่อแสดงค่าแผนภาพการกระจายของค่าคลาดเคลื่อน
7. ปุ่ม Continue — ปุ่มสำหรับเลือกเพื่อดำเนินการต่อ



รูปที่ 6.14 หน้าจอสำหรับกำหนดเลือกค่าสถิติของการวิเคราะห์ความถดถอยเชิงเส้น

6.3.2 การกำหนดค่าสถิติที่ต้องการบันทึกค่า

การกำหนดค่าสถิติที่ต้องการบันทึกค่าเป็นการกำหนดค่าสถิติที่ต้องการบันทึกค่าพยากรณ์ ในรูปแบบของตัวแปรใหม่ ในแฟ้มที่ใช้งานอยู่ จากคำสั่ง Statistics -> Linear Regression เมื่อกดปุ่ม Save จะปรากฏหน้าจอดังรูปที่ 6.15 ส่วนประกอบของรูปที่ 6.15 ประกอบด้วย

1. Predicted Values หมายถึงค่าพยากรณ์ โดยผู้ใช้สามารถเลือก

    1.1 Standardized หมายถึง ค่า Standardized Predicted Value

    1.2 Adjusted หมายถึง ค่า Adjusted Residual

2. Residual หมายถึงค่าคลาดเคลื่อน โดยที่ผู้ใช้สามารถเลือก

    2.1 Standardized หมายถึง ค่า Standardized Residual

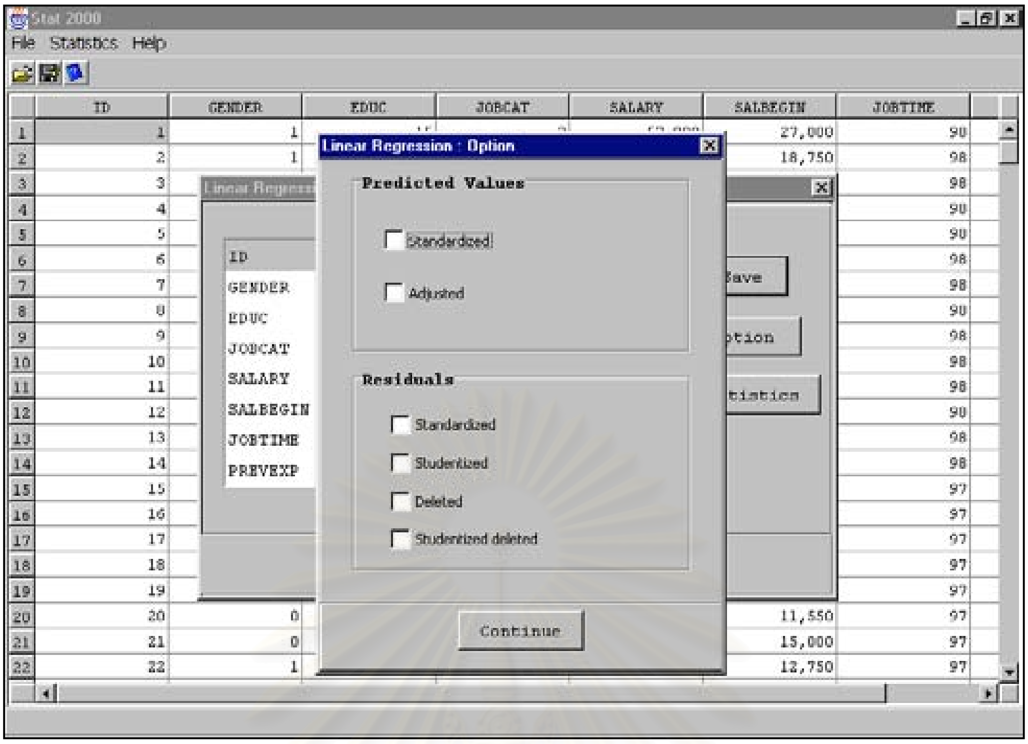
    2.2 Studentdized หมายถึง ค่า Studentdized Residual

    2.3 Deleted หมายถึง ค่า Deleted Residual

    2.4 Studentdized deleted หมายถึง ค่า Studentdized Delete Residual

3 ปุ่ม Continue เป็นปุ่มสำหรับเลือกเพื่อดำเนินการต่อ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



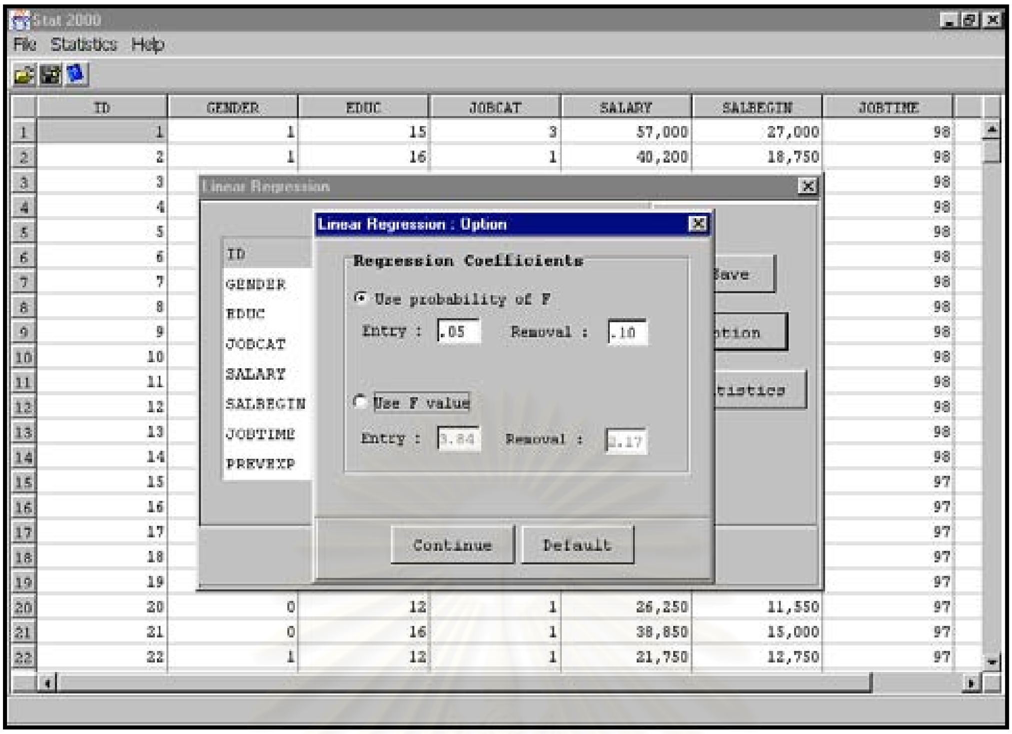
รูปที่ 6.15 หน้าจอสำหรับกำหนดค่าสถิติที่ต้องการบันทึกค่า

6.3.3 การกำหนดเงื่อนไขการเลือกตัวแปรอิสระเข้าสมการความถดถอย

การกำหนดเงื่อนไขในการเลือกตัวแปรอิสระเข้าสมการความถดถอยจะใช้เมื่อเลือกวิธีเลือกตัวแปรอิสระแบบ ฟอร์เวิร์ด แบ็คเวิร์ด หรือ สเตปไวส์ โดยเงื่อนไขที่กำหนดนี้จะเป็นตัวกำหนดค่าระดับนัยสำคัญของสถิติทดสอบ F สำหรับการเลือกตัวแปรเข้าหรือคัดตัวแปรออกจากสมการความถดถอย อาจจะกำหนดเป็นค่าสถิติทดสอบ F ก็ได้ จากคำสั่ง Statistics -> Linear Regression เมื่อกดปุ่ม Option จะปรากฏหน้าจอดังรูปที่ 6.16 ส่วนประกอบของรูปที่ 6.16 ประกอบด้วย

1. Use probability of F กำหนดใช้ระดับนัยสำคัญของ F เป็นเกณฑ์การเลือกตัวแปรอิสระ โดยทั่วไปจะให้ระดับนัยสำคัญของการเลือกเข้า (Entry) เป็น 0.5 และของการเลือกตัวแปรออก (Remove) เป็น 1.0
2. Use F value กำหนดใช้ค่าสถิติ F เป็นเกณฑ์ในการเลือกตัวแปรอิสระ โดยทั่วไป จะให้ค่า F ของการเลือก (Entry) เป็น 3.84 และของการเลือกตัวแปรออก (Remove) เป็น 2.71
3. ปุ่ม Continue ปุ่มสำหรับเลือกเพื่อดำเนินการต่อ

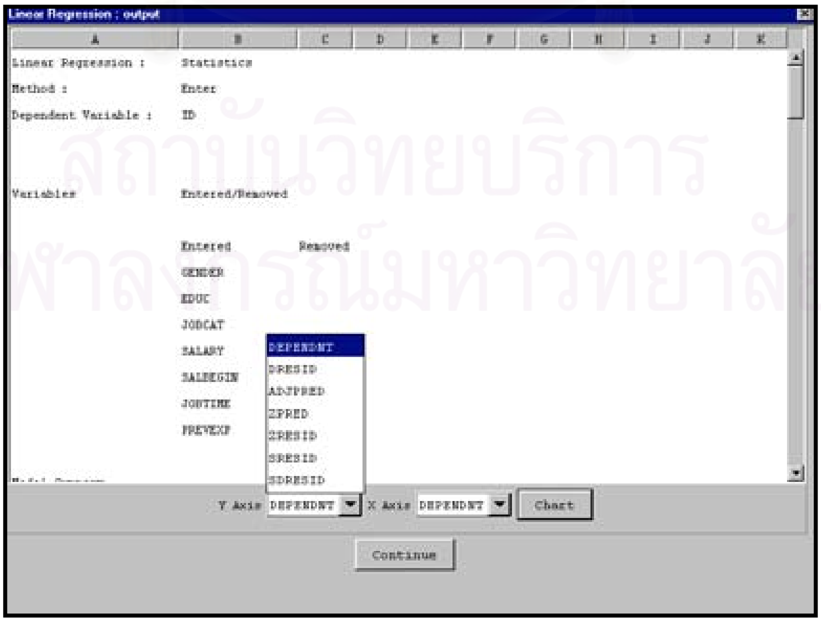
สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 6.16 หน้าจอกำหนดเงื่อนไขการเลือกตัวแปรอิสระเข้าสมการความถดถอย

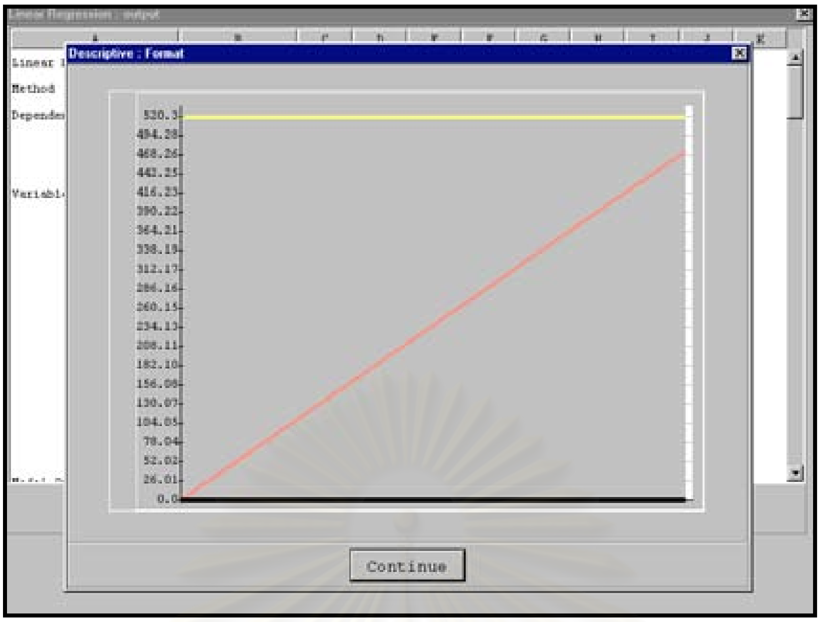
6.3.4 การแสดงแผนภูมิสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

การแสดงแผนภูมิสำหรับการวิเคราะห์ความถดถอยเชิงเส้น สามารถเลือกการแสดงค่าจากรายการค่าคลาดเคลื่อนดังแสดงในรูปที่ 6.17 โดยสามารถเลือกค่าตามแนวแกน Xและ Y เมื่อเลือกตัวแปรที่ต้องการแล้วกดปุ่ม Chart จะปรากฏหน้าจอแสดงแผนภูมิดังตามค่าแกน X และ Y ที่กำหนดดังแสดงในรูปที่ 6.18



รูปที่ 6.17 แสดงการเลือกตัวแปรที่ต้องการแสดงผลลัพธ์ในรูปแผนภูมิ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย



รูปที่ 6.18 แสดงผลลัพธ์แผนภูมิสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 7

การทดสอบส่วนประกอบซอฟต์แวร์

ผู้วิจัยได้พัฒนาโปรแกรม Stat2000 ขึ้น โปรแกรมนี้มีความสามารถครอบคลุมส่วนประกอบซอฟต์แวร์เชิงวัตถุที่ได้พัฒนาซึ่งมีความสามารถในการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น สำหรับบทนี้จะแสดงข้อเปรียบเทียบความถูกต้องของผลลัพธ์ที่ได้จากโปรแกรม Stat2000 กับผลลัพธ์ที่ได้จากโปรแกรม SPSS โดยจะแสดงผลลัพธ์จากการวิเคราะห์สถิติเชิงพรรณนาและผลลัพธ์จากการวิเคราะห์ความถดถอยเชิงเส้นโดยครอบคลุมวิธีการเลือกตัวแปรเข้าสมการความถดถอยเชิงเส้นทั้ง 4 แบบคือ วิธีเอนเตอร์ วิธีฟอร์เวิร์ด วิธีแบ็คเวิร์ด และ วิธีสเตปไวส์

สำหรับข้อมูลที่นำมาวิเคราะห์คือข้อมูลของพนักงาน โดยมีข้อมูลต่างๆประกอบด้วย รหัสพนักงาน เพศ การศึกษา ประเภทงาน เงินเดือน เงินเดือนเริ่มต้น ระยะเวลาที่ว่าจ้าง และประสบการณ์ก่อนว่าจ้าง

7.1 การทดสอบส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนา

สถิติเชิงพรรณนา เป็นสถิติสำหรับ การแจกแจงความถี่ การวัดค่ากลางของข้อมูลและการวัดการกระจายของข้อมูล การคำนวณค่ากลางของข้อมูลและค่าสถิติวัดการกระจายของข้อมูลโดยใช้โปรแกรมเอสพีเอสเอส ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.1 และ ค่ากลางของข้อมูลและค่าสถิติวัดการกระจายของข้อมูลที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.2

Statistics

SALARY

N	Valid	474
	Missing	0
Mean		34419.57
Std. Error of Mean		784.3111
Median		28875.00
Mode		30750.00
Std. Deviation		17075.66
Variance		2.9E+08
Skewness		2.125
Std. Error of Skewness		.112
Kurtosis		5.378
Std. Error of Kurtosis		.224
Range		119250.0
Minimum		15750.00
Maximum		135000.0
Sum		1.6E+07
Percentiles	25	24000.00
	50	28875.00
	75	37162.50

รูปที่ 7.1 ผลลัพธ์จากการวิเคราะห์สถิติเชิงพรรณนาโดยโปรแกรม SPSS

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
Descriptive Statistics


Variable name :           SALARY
Central Tendency :
                          Mean :                34419.568
                          Median :              28875.0
                          Mode :                30750.0
                          Sum :                 1.6314875E7

                          Quartile # 1 :        24000.0
                          Quartile # 2 :        28875.0
                          Quartile # 3 :        37162.5
                          Quartile # 4 :        135000.0
```

ก. ผลลัพธ์ค่าสถิติแสดงค่ากลางของข้อมูล

```
Dispersion :
                    Minimum :                15750.0
                    Maximum :                135000.0
                    Range :                  119250.0
                    Variance :               2.91578214453E8
                    Standard deviation :     17075.661
                    Skewness :               2.125
                    Kurtosis :               5.378
```

ข. ผลลัพธ์ค่าสถิติวัดการกระจายของข้อมูล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

รูปที่ 7.2 ผลลัพธ์จากการวิเคราะห์สถิติเชิงพรรณนาโดยโปรแกรม Stat2000

จากรูปที่ 7.1 และ 7.2 ผลลัพธ์ค่ากลางและสถิติวัดการกระจายที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

การคำนวณค่าความถี่โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.3 และการคำนวณค่าความถี่โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.4 จากรูปที่ 7.3 และ 7.4

จากผลลัพธ์วิเคราะห์ค่าความถี่ที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

Current Salary

		Frequency	Percent	Cumulative Percent
Valid	$15,750	1	.2	.2
	$15,900	1	.2	.4
	$16,200	3	.6	1.1
	$16,350	1	.2	1.3
	$16,500	1	.2	1.5
	$16,650	1	.2	1.7
	$16,800	1	.2	1.9
	$16,950	3	.6	2.5
	$17,100	2	.4	3.0
	$17,250	1	.2	3.2
	$17,400	2	.4	3.6
	$17,700	1	.2	3.8
	....	...	...	...

รูปที่ 7.3 ผลลัพธ์จากการวิเคราะห์ค่าความถี่ของตัวแปร Salaryโดยโปรแกรม SPSS

****** Frequency Distribution ******

Value	Frequency	Percent	Cumulative Frequency	Cumulative Percent
15750.0	1	0.211	1	0.211
15900.0	1	0.211	2	0.422
16200.0	3	0.633	5	1.055
16350.0	1	0.211	6	1.266
16500.0	1	0.211	7	1.477
16650.0	1	0.211	8	1.688
16800.0	1	0.211	9	1.899
16950.0	3	0.633	12	2.532
17100.0	2	0.422	14	2.954
17250.0	1	0.211	15	3.165
17400.0	2	0.422	17	3.586
17700.0	1	0.211	18	3.797
....	....	....	....	....

รูปที่ 7.4 ผลลัพธ์จากการวิเคราะห์ค่าความถี่ของตัวแปร Salaryโดยโปรแกรม Stat2000

7.2 การทดสอบส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ความถดถอยเชิงเส้น

การทดสอบส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ความถดถอยเชิงเส้นจะทดสอบเทคนิคการนำตัวแปรอิสระเข้าสมการความถดถอยเชิงเส้นทั้ง 4 วิธีคือ วิธีเอนเตอร์ วิธีฟอร์เวิร์ด วิธีแบ็คเวิร์ด และวิธีสเตปไวส์

7.2.1 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยโดยวิธีเอนเตอร์

การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีเอนเตอร์ได้ผลการทดสอบดังนี้

7.2.1.1 ทดสอบการพิจารณานำตัวแปรอิสระเข้าหรือออกจากสมการความถดถอยเชิงเส้น

การเลือกตัวแปรเข้าสมการความถดถอยโดยใช้วิธีการนำตัวแปรอิสระทุกตัวที่เลือกเข้าสมการความถดถอยเชิงเส้น โปรแกรม SPSS มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.5 และการเลือกตัวแปรเข้าสมการความถดถอยโดยใช้วิธีเอนเตอร์เป็นวิธีการที่นำตัวแปรอิสระทุกตัวที่เลือกเข้าสมการความถดถอยเชิงเส้น โปรแกรม Stat2000 มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.6

จากรูปที่ 7.5 และ 7.6 จะเห็นว่าการเลือกตัวแปรอิสระเข้าสมการความถดถอยวิธีเอนเตอร์ โดยใช้โปรแกรม SPSS และ Stat2000 ซึ่งทั้งสองโปรแกรมนำตัวแปรอิสระทุกตัวเข้าสมการความถดถอยเชิงเส้นเช่นเดียวกัน

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	PREVEXP, JOBTIME, SALBEGIN, EDUC, JOBCAT[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: SALARY

รูปที่ 7.5 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีเอนเตอร์โดยโปรแกรม SPSS

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
Linear Regression :   Statistics
Method :              Enter

Variables             Entered/Removed

                      Entered          Removed
                      EDUC
                      JOBCAT
                      SALBEGIN
                      JOBTIME
                      PREVEXP
```

รูปที่ 7.6 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีเอนเตอร์ โดยโปรแกรม Stat2000

7.2.1.2 การทดสอบสถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น

การคำนวณสถิติสรุปของสมการความถดถอยเชิงเส้นเป็นการคำนวนค่าสถิติเช่น ค่าสถิติสัมประสิทธิ์สหสัมพันธ์และค่าสัมประสิทธิ์การตัดสินใจของแต่ละรูปแบบของสมการความถดถอยเชิงเส้น ค่าสถิติสรุปของสมการความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.7 และค่าสถิติสรุปของสมการความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.8

จากรูปที่ 7.7 และ 7.8 ผลลัพธ์ค่าสถิติสรุปของสมการความถดถอยที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.917[a]	.840	.839	$6,856.79	.840	493.084	5	468	.000

a. Predictors: (Constant), Previous Experience (months), Months since Hire, Beginning Salary, Educational Level (years), Employment Category

รูปที่ 7.7 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Model Summary

Model	R	R Square	Adjusted R Square	Std. of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	0.917	0.84	0.839	6856.795	0.84	493.084	5	468	0.0

รูปที่ 7.8 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

7.2.1.3 การทดสอบการวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น

การวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นเป็นการคำนวณค่าสถิติต่างๆเช่น การคำนวณค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ การคำนวณค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลอื่นๆ และ การคำนวณค่าแปรปรวนรวม ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.9 และความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.10

จากรูปที่ 7.9 และ 7.10 ผลลัพธ์ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.16E+11	5	2.32E+10	493.084	.000[a]
	Residual	2.20E+10	468	47015636		
	Total	1.38E+11	473			

a. Predictors: (Constant), Previous Experience (months), Months since Hire, Beginning Salary, Educational Level (years), Employment Category

b. Dependent Variable: Current Salary

รูปที่ 7.9 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Anova

Model		Sum of Squares	df	Mean Square	F	t Sig.
1	Regression Sum Square :	1.15913177991258E11	5	2.3182635598252E10	493.084	-0.091
	Residual Sum Square :	2.2003317445082E10	468	4.7015635566E7		
	Total Sum Square :	1.3791649543634E11	473	2.91578214453E8		

รูปที่ 7.10 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

7.2.1.4 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น

การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้นเป็นการคำนวณค่าสถิติต่างๆ เช่น ค่าประมาณของพารามิเตอร์ของสมการความถดถอยเชิงเส้น ค่าช่วงความเชื่อมั่น และค่าสถิติทดสอบและค่านัยสำคัญของสถิติทดสอบ t ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.11 และ ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.12

จากรูปที่ 7.11 และ 7.12 ผลลัพธ์การประมาณค่าพารามิเตอร์ของสมการความถดถอยที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากันยกเว้นค่าขอบเขตบนและขอบเขตล่างของค่าความเชื่อมั่นที่ 95 เปอร์เซ็นต์ แต่ค่าที่ได้มีความใกล้เคียงกันมาก

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-15038.6	2992.525		-5.025	.000	-20919.023	-9158.125
	Educational Level (years)	539.642	152.735	.091	3.533	.000	239.511	839.773
	Employment Category	5859.585	624.945	.265	9.376	.000	4631.540	7087.631
	Beginning Salary	1.365	.069	.629	19.796	.000	1.230	1.501
	Months since Hire	154.698	31.464	.091	4.917	.000	92.870	216.526
	Previous Experience (months)	-19.553	3.250	-.120	-6.017	.000	-25.939	-13.167

a. Dependent Variable: Current Salary

รูปที่ 7.11 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Coefficients

Model		B	Std. Error	Beta	t	Sig.	95% Confidence Interval for B	
							Lower Bound	Upper Bound
1	Constant	-15038.574	2992.525		-5.025	0.0	-20919.017	-9158.13
	SALBEGIN	1.365	0.069	0.629	19.796	0.0	1.23	1.501
	JOBTIME	154.698	31.464	0.091	4.917	0.0	92.87	216.526
	PREVEXP	-19.553	3.25	-0.12	-6.017	0.0	-25.939	-13.167
	JOBCAT	5859.585	624.945	0.265	9.376	0.0	4631.541	7087.63
	EDUC	539.642	152.735	0.091	3.533	0.0	239.512	839.773

รูปที่ 7.12 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

7.2.1.5 การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว

การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวเป็นการคำนวณค่าความสัมพันธ์ของตัวแปรและตัวแปรอิสระและค่านัยสำคัญของความสัมพันธ์ของตัวแปรทีละคู่ การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.13 และการทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.14

จากรูปที่ 7.13 และ 7.14 ผลลัพธ์ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

Correlations

		Current Salary	Educational Level (years)	Employment Category	Beginning Salary	Months since Hire	Previous Experience (months)
Pearson Correlation	Current Salary	1.000	.661	.780	.880	.084	-.097
	Educational Level (years)	.661	1.000	.514	.633	.047	-.252
	Employment Category	.780	.514	1.000	.755	.005	.063
	Beginning Salary	.880	.633	.755	1.000	-.020	.045
	Months since Hire	.084	.047	.005	-.020	1.000	.003
	Previous Experience (months)	-.097	-.252	.063	.045	.003	1.000
Sig. (1-tailed)	Current Salary	.	.000	.000	.000	.034	.017
	Educational Level (years)	.000	.	.000	.000	.152	.000
	Employment Category	.000	.000	.	.000	.454	.087
	Beginning Salary	.000	.000	.000	.	.334	.163
	Months since Hire	.034	.152	.454	.334	.	.474
	Previous Experience (months)	.017	.000	.087	.163	.474	.
N	Current Salary	474	474	474	474	474	474
	Educational Level (years)	474	474	474	474	474	474
	Employment Category	474	474	474	474	474	474
	Beginning Salary	474	474	474	474	474	474
	Months since Hire	474	474	474	474	474	474
	Previous Experience (months)	474	474	474	474	474	474

รูปที่ 7.13 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดยโปรแกรม SPSS

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Correlations

		SALARY	EDUC	JOBCAT	SALBEGIN	JOBTIME	PREVEXP
Pearson Correlations :	SALARY	1.0	0.661	0.78	0.88	0.084	-0.097
	EDUC	0.661	1.0	0.514	0.633	0.047	-0.252
	JOBCAT	0.78	0.514	1.0	0.755	0.0050	0.063
	SALBEGIN	0.88	0.633	0.755	1.0	-0.02	0.045
	JOBTIME	0.084	0.047	0.0050	-0.02	1.0	0.0030
	PREVEXP	-0.097	-0.252	0.063	0.045	0.0030	1.0
Sig. (1-tailed) :	SALARY	0.0	0.0	-0.0	0.0	0.034	0.017
	EDUC	0.0	0.0	0.0	0.0	0.152	0.0
	JOBCAT	-0.0	0.0	0.0	0.0	0.454	0.087
	SALBEGIN	0.0	0.0	0.0	0.0	0.334	0.163
	JOBTIME	0.034	0.152	0.454	0.334	0.0	0.474
	PREVEXP	0.017	0.0	0.087	0.163	0.474	0.0

รูปที่ 7.14 ค่าคลาดเคลื่อนที่คำนวณโดยโปรแกรม Stat2000

7.2.1.6 การทดสอบการคำนวณค่าคลาดเคลื่อน

การคำนวณค่าคลาดเคลื่อนโดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.15 และการคำนวณค่าคลาดเคลื่อนโดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.16

จากรูปที่ 7.15 และ 7.16 ผลการคำนวณค่าคลาดเคลื่อนที่ได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

CUST1 - SPSS Data Editor

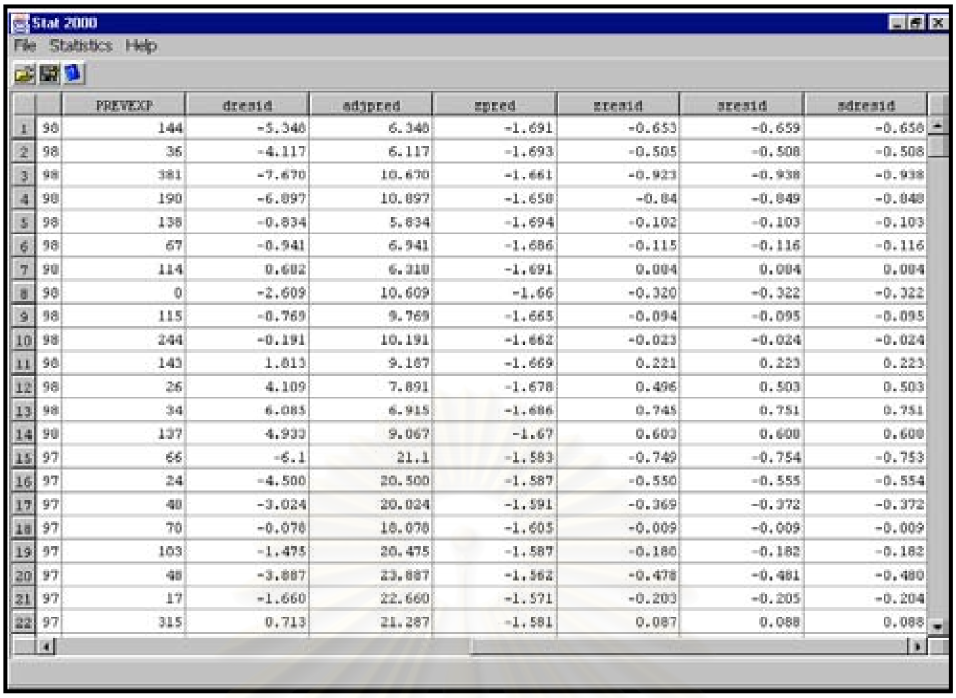
File Edit View Data Transform Statistics Graphs Utilities Window Help

14:sre_1 .608170933768663

	prevexp	dre_1	adj_1	zpr_1	zre_1	sre_1	sdr_1
1	144.00	-5.34840	6.34840	-1.69113	-.65274	-.65887	-.65847
2	36.00	-4.11746	6.11746	-1.69252	-.50477	-.50837	-.50796
3	381.00	-7.67014	10.67014	-1.66061	-.92300	-.93825	-.93812
4	190.00	-6.09603	10.09603	-1.65020	-.03961	-.04055	-.04029
5	138.00	-.83440	5.83440	-1.69426	-.10215	-.10295	-.10285
6	67.00	-.94065	6.94065	-1.68616	-.11545	-.11620	-.11608
7	114.00	.68164	6.31836	-1.69056	.08366	.08421	.08412
8	.00	-2.60889	10.60889	-1.65950	-.32004	-.32221	-.32190
9	115.00	-.76936	9.76936	-1.66547	-.09418	-.09492	-.09482
10	244.00	-.19124	10.19124	-1.66232	-.02337	-.02357	-.02355
11	143.00	1.81348	9.18652	-1.66938	.22101	.22324	.22301
12	26.00	4.10895	7.89105	-1.67824	.49592	.50337	.50296
13	34.00	6.08470	6.91530	-1.68560	.74545	.75100	.75065
14	137.00	4.93253	9.06747	-1.66991	.60305	.60817	.60776
15	66.00	-6.09851	21.09851	-1.58310	-.74869	-.75355	-.75320

SPSS Processor is ready

รูปที่ 7.15 ค่าคลาดเคลื่อนที่คำนวณโดยโปรแกรม SPSS



		PREVEXP	dresid	adjpred	zpred	zresid	sresid	sdresid
1	98	144	-5.348	6.348	-1.691	-0.653	-0.659	-0.658
2	98	36	-4.117	6.117	-1.693	-0.505	-0.508	-0.508
3	98	381	-7.670	10.670	-1.661	-0.923	-0.938	-0.938
4	98	190	-6.897	10.897	-1.658	-0.84	-0.849	-0.848
5	98	138	-0.834	5.834	-1.694	-0.102	-0.103	-0.103
6	98	67	-0.941	6.941	-1.686	-0.115	-0.116	-0.116
7	98	114	0.682	6.318	-1.691	0.084	0.084	0.084
8	98	0	-2.609	10.609	-1.66	-0.320	-0.322	-0.322
9	98	115	-0.769	9.769	-1.665	-0.094	-0.095	-0.095
10	98	244	-0.191	10.191	-1.662	-0.023	-0.024	-0.024
11	98	143	1.813	9.187	-1.669	0.221	0.223	0.223
12	98	26	4.109	7.891	-1.678	0.496	0.503	0.503
13	98	34	6.085	6.915	-1.686	0.745	0.751	0.751
14	98	137	4.933	9.067	-1.67	0.603	0.608	0.608
15	97	66	-6.1	21.1	-1.583	-0.749	-0.754	-0.753
16	97	24	-4.500	20.500	-1.587	-0.550	-0.555	-0.554
17	97	40	-3.024	20.024	-1.591	-0.369	-0.372	-0.372
18	97	70	-0.078	18.078	-1.605	-0.009	-0.009	-0.009
19	97	103	-1.475	20.475	-1.587	-0.180	-0.182	-0.182
20	97	48	-3.887	23.887	-1.562	-0.478	-0.481	-0.480
21	97	17	-1.660	22.660	-1.571	-0.203	-0.205	-0.204
22	97	315	0.713	21.287	-1.581	0.087	0.088	0.088

รูปที่ 7.16 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดยโปรแกรม Stat2000

7.2.2 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด

การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีฟอร์เวิร์ด ได้ผลการทดสอบดังนี้

7.2.2.1 ทดสอบการพิจารณานำตัวแปรอิสระเข้าหรือออกจากสมการความถดถอยเชิงเส้น

การเลือกตัวแปรเข้าสมการความถดถอยโดยใช้วิธีการนำตัวแปรอิสระที่มีนัยสำคัญเข้าสมการความถดถอยเชิงเส้น โปรแกรม SPSS มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.17 และการเลือกตัวแปรเข้าสมการความถดถอยโดยใช้วิธีการนำตัวแปรอิสระที่มีนัยสำคัญเข้าสมการความถดถอยเชิงเส้น โปรแกรม Stat2000 มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.18

รูปที่ 7.17 และ 7.18 แสดงการเลือกตัวแปรอิสระเข้าสมการความถดถอยวิธีฟอร์เวิร์ด โดยใช้โปรแกรม SPSS และ Stat2000 โปรแกรมพิจารณานำตัวแปรอิสระเข้าสมการทีละตัว และเนื่องจากตัวแปรอิสระแต่ละตัวล้วนแต่มีนัยสำคัญต่อสมการความถดถอยเชิงเส้น ดังนั้นตัวแปรทุกตัวจะถูกนำเข้าสมการความถดถอยเชิงเส้น จากผลลัพธ์โปรแกรม Stat2000 มีการเลือกตัวแปรอิสระตามวิธีฟอร์เวิร์ด ได้ผลเช่นเดียวกับโปรแกรม SPSS

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Beginning Salary	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
2	Employment Category	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
3	Previous Experience (months)	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
4	Months since Hire	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
5	Educational Level (years)	.	Forward (Criterion: Probability-of-F-to-enter <= .050)

a. Dependent Variable: Current Salary

รูปที่ 7.17 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีฟอร์เวิร์ด โดยโปรแกรม SPSS

```
Linear Regression :    Statistics
Method :               Forward

Variables              Entered/Removed

                       Entered          Removed
                       SALBEGIN
                       JOBCAT
                       PREVEXP
                       JOBTIME
                       EDUC
```

รูปที่ 7.18 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีฟอร์เวิร์ด โดยโปรแกรม Stat2000

7.2.2.2 การทดสอบค่าสถิติสรุปแต่ละรูปแบบของสมการความถดถอย

การคำนวณค่าสถิติสรุปของสมการความถดถอยเชิงเส้นเป็นการคำนวนค่าสถิติต่างๆเช่น ค่าสัมประสิทธิ์สหสัมพันธ์ และค่าสัมประสิทธิ์การตัดสินใจของแต่ละรูปแบบของสมการความถดถอยเชิงเส้น ค่าสถิติสรุปของสมการความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.19 และค่าสถิติสรุปของสมการความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.20

จากรูปที่ 7.19 และ 7.20 ผลลัพธ์ค่าสถิติสรุปของสมการความถดถอยที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.880[a]	.775	.774	$8,115.36	.775	1622.118	1	472	.000
2	.898[b]	.806	.805	$7,540.43	.031	75.719	1	471	.000
3	.909[c]	.827	.826	$7,127.04	.021	57.224	1	470	.000
4	.914[d]	.836	.835	$6,940.23	.009	26.642	1	469	.000
5	.917[e]	.840	.839	$6,856.79	.004	12.483	1	468	.000

a. Predictors: (Constant), Beginning Salary

b. Predictors: (Constant), Beginning Salary, Employment Category

c. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months)

d. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months), Months since Hire

e. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months), Months since Hire, Educational Level (years)

รูปที่ 7.19 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Model Summary

Model	R	R Square	Adjusted R Square	Std. of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	0.88	0.775	0.774	8115.356	0.775	1622.118	1	472	0.0
2	0.898	0.806	0.805	7540.433	0.031	75.719	1	471	-0.0
3	0.909	0.827	0.826	7127.04	0.021	57.224	1	470	0.0
4	0.914	0.836	0.835	6940.232	0.0090	26.642	1	469	0.0
5	0.917	0.84	0.839	6856.795	0.0040	12.483	1	468	0.0

รูปที่ 7.20 สถิติสรุปแต่ละรูปแบบของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

7.2.2.3 การทดสอบการวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น

การวิเคราะห์ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นเป็นการคำนวณค่าสถิติเช่น ค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลของตัวแปรอิสระ การคำนวณค่าแปรปรวนของตัวแปรตามเนื่องจากอิทธิพลอื่นๆ และ การคำนวณค่าแปรปรวนรวม ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.21 และความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.22

จากรูปที่ 7.21 และ 7.22 ผลลัพธ์ค่าความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้นที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

ANOVA[f]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.07E+11	1	1.07E+11	1622.118	.000[a]
	Residual	3.11E+10	472	65858997		
	Total	1.38E+11	473			
2	Regression	1.11E+11	2	5.56E+10	977.312	.000[b]
	Residual	2.68E+10	471	56858136		
	Total	1.38E+11	473			
3	Regression	1.14E+11	3	3.80E+10	748.392	.000[c]
	Residual	2.39E+10	470	50794697		
	Total	1.38E+11	473			
4	Regression	1.15E+11	4	2.88E+10	598.577	.000[d]
	Residual	2.26E+10	469	48166815		
	Total	1.38E+11	473			
5	Regression	1.16E+11	5	2.32E+10	493.084	.000[e]
	Residual	2.20E+10	468	47015636		
	Total	1.38E+11	473			

a. Predictors: (Constant), Beginning Salary

b. Predictors: (Constant), Beginning Salary, Employment Category

c. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months)

d. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months), Months since Hire

e. Predictors: (Constant), Beginning Salary, Employment Category, Previous Experience (months), Months since Hire, Educational Level (years)

f. Dependent Variable: Current Salary

รูปที่ 7.21 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Anova

Model		Sum of Squares	df	MeanSquare	F	t Sig.
1	Regression Sum Square :	1.06831048750124E11	1	1.06831048750124E11	1622.118	0.0
	Residual Sum Square :	3.1085446686215E10	472	6.5858997217E7		
	Total Sum Square :	1.37916495436 34E11	473	2.91578214453E8		
2	Regression Sum Square :	1.11136313278119E11	2	5.5568156639059E10	977.312	0.0
	Residual Sum Square :	2.6780182158221E10	471	5.6858136217E7		
	Total Sum Square :	1.3791649543634E11	473	2.91578214453E8		
3	Regression Sum Square :	1.14042988034361E11	3	3.8014329344787E10	748.392	-0.1
	Residual Sum Square :	2.3873507401979E10	470	5.07946966E7		
	Total Sum Square :	1.3791649543634E11	473	2.91578214453E8		
4	Regression Sum Square :	1.15326259146009E11	4	2.8831564786502E10	598.577	0.0
	Residual Sum Square :	2.2590236290331E10	469	4.8166815118E7		
	Total Sum Square :	1.3791649543634E11	473	2.91578214453E8		
5	Regression Sum Square :	1.15913177991258E11	5	2.3182635598252E10	493.084	-0.091
	Residual Sum Square :	2.2003317445082E10	468	4.7015635566E7		
	Total Sum Square :	1.3791649543634E11	473	2.91578214453E8		

รูปที่ 7.22 ความแปรปรวนของการวิเคราะห์ความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

7.2.2.4 การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้น

การประมาณค่าพารามิเตอร์ของสมการความถดถอยเชิงเส้นเป็นการคำนวณค่าสถิติเช่น ค่าประมาณของพารามิเตอร์ของสมการความถดถอยเชิงเส้น ค่าช่วงความเชื่อมั่น และค่าสถิติทดสอบและค่านัยสำคัญของสถิติทดสอบ t ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.23 และค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น ที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.24

จากรูปที่ 7.23 และ 7.24 ผลลัพธ์การประมาณค่าพารามิเตอร์ของสมการความถดถอยที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากันยกเว้นค่าขอบเขตบนและขอบเขตล่างของค่าความเชื่อมั่นที่ 95 เปอร์เซ็นต์ แต่ค่าที่ได้มีความใกล้เคียงกันมาก

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	1928.206	888.680		2.170	.031	181.947	3674.464
	Beginning Salary	1.909	.047	.880	40.276	.000	1.816	2.003
2	(Constant)	1036.931	832.051		1.246	.213	-598.060	2671.922
	Beginning Salary	1.469	.067	.677	21.873	.000	1.337	1.600
	Employment Category	5947.000	683.430	.269	8.702	.000	4604.050	7289.950
3	(Constant)	3039.205	829.783		3.663	.000	1408.661	4669.748
	Beginning Salary	1.467	.063	.676	23.117	.000	1.342	1.592
	Employment Category	6160.294	646.577	.279	9.528	.000	4889.753	7430.834
	Previous Experience (months)	-23.749	3.139	-.145	-7.565	.000	-29.918	-17.580
4	(Constant)	-10300.7	2707.813		-3.804	.000	-15621.623	-4979.726
	Beginning Salary	1.479	.062	.682	23.911	.000	1.357	1.600
	Employment Category	6060.446	629.927	.274	9.621	.000	4822.617	7298.274
	Previous Experience (months)	-23.789	3.057	-.146	-7.781	.000	-29.797	-17.782
	Months since Hire	163.826	31.739	.097	5.162	.000	101.457	226.195
5	(Constant)	-15038.6	2992.525		-5.025	.000	-20919.023	-9158.125
	Beginning Salary	1.365	.069	.629	19.796	.000	1.230	1.501
	Employment Category	5859.585	624.945	.265	9.376	.000	4631.540	7087.631
	Previous Experience (months)	-19.553	3.250	-.120	-6.017	.000	-25.939	-13.167
	Months since Hire	154.698	31.464	.091	4.917	.000	92.870	216.526
	Educational Level (years)	539.642	152.735	.091	3.533	.000	239.511	839.773

a. Dependent Variable: Current Salary

รูปที่ 7.23 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม SPSS

Coefficients

Model		B	Std. Error	Beta	t	Sig.	95% Confidence Interval for B	
							Lower Bound	Upper Bound
1	Constant	1928.206	888.68		2.17	0.031	181.94	3674.471
	SALBEGIN	1.909	0.047	0.88	40.276	0.0	1.816	2.003
2	Constant	1036.931	832.051		1.246	0.213	-598.058	2671.919
	SALBEGIN	1.469	0.067	0.677	21.873	0.0	1.337	1.6
	JOBCAT	5947.0	683.43	0.269	8.702	0.0	4604.052	7289.948
3	Constant	3039.205	829.783		3.663	0.0	1408.673	4669.737
	SALBEGIN	1.467	0.063	0.676	23.117	0.0	1.342	1.592
	JOBCAT	6160.294	646.577	0.279	9.528	-0.0	4889.762	7430.825
	PREVEXP	-23.749	3.139	-0.145	-7.565	0.0	-29.918	-17.58
4	Constant	-10300.675	2707.813		-3.804	0.0	-15621.648	-4979.702
	SALBEGIN	1.479	0.062	0.682	23.911	0.0	1.357	1.6
	JOBCAT	6060.446	629.927	0.274	9.621	0.0	4822.612	7298.28
	PREVEXP	-23.789	3.057	-0.146	-7.781	0.0	-29.797	-17.782
	JOBTIME	163.826	31.739	0.097	5.162	0.0	101.457	226.195
5	Constant	-15038.574	2992.525		-5.025	0.0	-20919.017	-9158.13
	SALBEGIN	1.365	0.069	0.629	19.796	-0.0	1.23	1.501
	JOBCAT	5859.585	624.945	0.265	9.376	0.0	4631.541	7087.63
	PREVEXP	-19.553	3.25	-0.12	-6.017	0.0	-25.939	-13.167
	JOBTIME	154.698	31.464	0.091	4.917	0.0	92.87	216.526
	EDUC	539.642	152.735	0.091	3.533	0.0	239.512	839.773

รูปที่ 7.24 ค่าประมาณพารามิเตอร์ของสมการความถดถอยเชิงเส้น คำนวณโดยโปรแกรม Stat2000

7.2.2.5 การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัว

การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวเป็นการคำนวณค่าความสัมพันธ์ของตัวแปรและตัวแปรอิสระและค่านัยสำคัญของความสัมพันธ์ของตัวแปรทีละคู่ การทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่ได้จากการวิเคราะห์โดยใช้โปรแกรม SPSS ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.25 และการทดสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่ได้จากการวิเคราะห์โดยใช้โปรแกรม Stat2000 ได้ผลลัพธ์ค่าสถิติดังแสดงในรูปที่ 7.26

จากรูปที่ 7.25 และ 7.26 ผลลัพธ์ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระที่คำนวณได้จากโปรแกรม SPSS และ Stat2000 ได้ค่าการคำนวณที่ถูกต้องและมีค่าเท่ากัน

Correlations

		Current Salary	Educational Level (years)	Employment Category	Beginning Salary	Months since Hire	Previous Experience (months)
Pearson Correlation	Current Salary	1.000	.661	.780	.880	.084	-.097
	Educational Level (years)	.661	1.000	.514	.633	.047	-.252
	Employment Category	.780	.514	1.000	.755	.005	.063
	Beginning Salary	.880	.633	.755	1.000	-.020	.045
	Months since Hire	.084	.047	.005	-.020	1.000	.003
	Previous Experience (months)	-.097	-.252	.063	.045	.003	1.000
Sig. (1-tailed)	Current Salary	.	.000	.000	.000	.034	.017
	Educational Level (years)	.000	.	.000	.000	.152	.000
	Employment Category	.000	.000	.	.000	.454	.087
	Beginning Salary	.000	.000	.000	.	.334	.163
	Months since Hire	.034	.152	.454	.334	.	.474
	Previous Experience (months)	.017	.000	.087	.163	.474	.
N	Current Salary	474	474	474	474	474	474
	Educational Level (years)	474	474	474	474	474	474
	Employment Category	474	474	474	474	474	474
	Beginning Salary	474	474	474	474	474	474
	Months since Hire	474	474	474	474	474	474
	Previous Experience (months)	474	474	474	474	474	474

รูปที่ 7.25 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดยโปรแกรม SPSS

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Correlations

		SALARY	EDUC	JOBCAT	SALBEGIN	JOBTIME	PREVEXP
Pearson Correlations :	SALARY	1.0	0.661	0.78	0.88	0.084	-0.097
	EDUC	0.661	1.0	0.514	0.633	0.047	-0.252
	JOBCAT	0.78	0.514	1.0	0.755	0.0050	0.063
	SALBEGIN	0.88	0.633	0.755	1.0	-0.02	0.045
	JOBTIME	0.084	0.047	0.0050	-0.02	1.0	0.0030
	PREVEXP	-0.097	-0.252	0.063	0.045	0.0030	1.0
Sig. (1-tailed) :	SALARY	0.0	0.0	-0.0	0.0	0.034	0.017
	EDUC	0.0	0.0	0.0	0.0	0.152	0.0
	JOBCAT	-0.0	0.0	0.0	0.0	0.454	0.087
	SALBEGIN	0.0	0.0	0.0	0.0	0.334	0.163
	JOBTIME	0.034	0.152	0.454	0.334	0.0	0.474
	PREVEXP	0.017	0.0	0.087	0.163	0.474	0.0

รูปที่ 7.26 ค่าความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระทีละตัวคำนวณโดยโปรแกรม Stat2000

<u>หมายเหตุ</u> การคำนวณค่าคลาดเคลื่อนได้ผลลัพธ์เช่นเดียวกันกับวิธีเอนเตอร์ เนื่องจากสมการความถดถอยเชิงเส้นที่เหมาะสมมีการนำค่าตัวแปรอิสระทุกตัวเข้าสมการความถดถอยเชิงเส้น

7.2.3 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยโดยวิธีแบ็ควอร์ด

การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธีแบ็คเวิร์ด เป็นวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยในขั้นแรกจะนำตัวแปรอิสระทุกตัวเข้าสมการความถดถอยแล้วพิจารณาหาตัวแปรอิสระที่ไม่มีนัยสำคัญออกจากสมการความถดถอยทีละตัว โปรแกรม SPSS มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.27 และการตัดออกตัวแปรอิสระออกจากสมการความถดถอยเชิงเส้นด้วยวิธีแบ็คเวิร์ด โดยโปรแกรม Stat2000 มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.28

จากรูปที่ 7.27 และ 7.28 แสดงการเลือกตัวแปรอิสระเข้าสมการความถดถอยวิธีแบ็คเวิร์ด โดยใช้โปรแกรม SPSS และ Stat2000 โปรแกรมพิจารณาตัดแปรอิสระออกจากสมการทีละตัว และเนื่องจากตัวแปรอิสระแต่ละตัวล้วนแต่มีนัยสำคัญต่อสมการความถดถอยเชิงเส้น ดังนั้นตัวแปรแต่ละตัวจึงไม่ถูกตัดออกจากสมการความถดถอยเชิงเส้น จากผลลัพธ์โปรแกรม Stat2000 มีการเลือกตัวแปรอิสระตามวิธีแบ็คเวิร์ดได้ผลเช่นเดียวกับโปรแกรม SPSS และเนื่องจากในการพิจารณาตัดตัวแปรออกจากสมการความถดถอยไม่พบตัวแปรที่ไม่มีนัยสำคัญ ดังนั้นตัวแปรอิสระทุกตัวถูกนำเข้าสมการความถดถอย ซึ่งทำให้ได้ค่าสถิติต่างๆเช่นเดียวกับวิธี เอนเตอร์ ดังแสดงในข้อที่ 1

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Previous Experience (months), Beginning Salary, Educational Level (years), Employment Category[a]	.	Enter

a. All requested variables entered.

b. Dependent Variable: Current Salary

รูปที่ 7.27 ตัวแปรอิสระที่ตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธี แบ็คเวร์ดโดยโปรแกรม SPSS

Linear Regression :	Statistics	
Method :	Backward	
Variables	Entered/Removed	
	Entered	Removed
	EDUC	
	JOBCAT	
	SALBEGIN	
	JOBTIME	
	PREVEXP	

รูปที่ 7.28 ตัวแปรอิสระที่ตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีแบ็ควอร์ดโดยโปรแกรม Stat2000

7.2.4 ทดสอบการวิเคราะห์ความถดถอยเชิงเส้นโดยใช้วิธีการเลือกตัวแปรเข้าสมการความถดถอยโดยวิธีสเตปไวส์

การวิเคราะห์ความถดถอยเชิงเส้นโดยวิธีการเลือกตัวแปรอิสระเข้าสมการความถดถอยโดยวิธี สเตปไวส์ เป็นการพิจารณานำตัวแปรอิสระเข้าสมการโดยวิธีฟอร์วอร์ดและพิจารณาตัดตัวแปรอิสระออกจากสมการโดยวิธีแบ็ควอร์ด โปรแกรม SPSS มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.29 และการนำตัวแปรอิสระเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีสเตปไวส์ โดยโปรแกรม Stat2000 มีการเลือกตัวแปรอิสระเข้าสมการดังแสดงในรูปที่ 7.30

จากรูปที่ 7.29 และ 7.30 แสดงการเลือกตัวแปรอิสระเข้าสมการความถดถอยวิธี สเต็ปไวส์ โดยใช้โปรแกรม SPSS และ Stat2000 โปรแกรมพิจารณานำตัวแปรอิสระเข้าสมการทีละตัวโดยวิธีฟอร์เวิร์ด และพิจารณาตัดตัวแปรอิสระออกจากสมการความถดถอยโดยวิธีแบ็คเวิร์ด เนื่องจากตัวแปรอิสระแต่ละตัวล้วนแต่มีนัยสำคัญต่อสมการความถดถอยเชิงเส้น ดังนั้นตัวแปรทุกตัวจะถูกนำเข้าสมการความถดถอยเชิงเส้น

จากผลลัพธ์โปรแกรม Stat2000 มีการเลือกตัวแปรอิสระตามวิธี สเต็ปไวส์ ได้ผลเช่นเดียวกับโปรแกรม SPSS

เนื่องจากในการพิจารณานำตัวแปรอิสระเข้าสมการความถดถอยพบว่าทุกตัวแปรมีนัยสำคัญจึงมีการนำตัวแปรอิสระเข้าสมการความถดถอยทีละตัว ซึ่งทำให้ได้ค่าสถิติต่างๆเช่นเดียวกับวิธีฟอร์เวิร์ด ดังแสดงในหัวข้อที่ 2

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Beginning Salary	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	Employment Category	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	Previous Experience (months)	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
4	Months since Hire	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
5	Educational Level (years)	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Current Salary

รูปที่ 7.29 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีสเต็ปไวส์โดยโปรแกรม SPSS

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```
Linear Regression :  Statistics
Method :             Stepwise

Variables            Entered/Removed

                     Entered           Removed
                     SALBEGIN
                     JOBCAT
                     PREVEXP
                     JOBTIME
                     EDUC
```

รูปที่ 7.30 ตัวแปรอิสระที่นำเข้าหรือตัดออกจากสมการความถดถอยเชิงเส้นด้วยวิธีสเต็ปไวส์ โดยโปรแกรม Stat2000

7.3 สรุปการทดสอบ

จากการตรวจสอบในแง่ผลลัพธ์ของการคำนวณค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น พบว่าค่าสถิติที่คำนวณโดยโปรแกรม Stat2000 ซึ่งเป็นโปรแกรมที่เรียกใช้ส่วนประกอบซอฟต์แวร์ที่ผู้วิจัยได้พัฒนาขึ้นเปรียบเทียบกับผลลัพธ์ที่คำนวณได้จากโปรแกรม SPSS เวอร์ชัน 8.0 พบว่าส่วนประกอบซอฟต์แวร์ที่ผู้วิจัยพัฒนาขึ้นนี้มีความถูกต้องสำหรับทุกค่าสถิติเมื่อเทียบกับผลการคำนวณค่าสถิติกับโปรแกรม SPSS ยกเว้นค่าขอบเขตบนและขอบเขตล่างของค่าความเชื่อมั่นที่ 95 เปอร์เซ็นต์ แต่อย่างไรก็ตาม ค่าที่ได้มีความใกล้เคียงกันมากจึงสามารถนำไปใช้อ้างอิงได้ เนื่องจากค่าสถิติดังกล่าวเป็นเพียงค่าประมาณทางสถิติเท่านั้น

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 8

บทสรุปและข้อเสนอแนะ

8.1 บทสรุป

งานวิจัยนี้ผู้วิจัยได้ออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้น เพื่อให้ผู้พัฒนาโปรแกรมเชิงวัตถุสามารถนำส่วนประกอบซอฟต์แวร์เชิงวัตถุเหล่านี้ไปใช้ร่วมกับการพัฒนาโปรแกรมเชิงวัตถุอื่นๆได้ นอกจากนี้ผู้ใช้ที่ต้องการวิเคราะห์สถิติเชิงพรรณนาหรือวิเคราะห์ความถดถอยเชิงเส้นยังสามารถนำโปรแกรมที่ใช้ทดสอบการเรียกใช้ส่วนประกอบซอฟต์แวร์ที่พัฒนาขึ้นไปคำนวณหาค่าสถิติแทนการใช้โปรแกรมสำเร็จรูปอื่นๆ

ส่วนประกอบซอฟต์แวร์ที่ได้พัฒนาขึ้นนี้ได้ผ่านตรวจสอบความถูกต้องโดยเปรียบเทียบผลลัพธ์การคำนวณค่าสถิติกับโปรแกรม SPSS เวอร์ชัน 8.0 และพบว่าค่าสถิติที่คำนวณโดยส่วนประกอบซอฟต์แวร์ที่พัฒนาขึ้นมีความถูกต้อง แต่อย่างไรก็ตามในการคำนวณค่าขอบเขตบนและขอบเขตล่างของค่าความเชื่อมั่นสำหรับการวิเคราะห์ความถดถอยเชิงเส้นพบว่ามีความคลาดเคลื่อนของผลลัพธ์ สาเหตุเนื่องจากการเรียกใช้ค่าสถิติทดสอบ t ที่คำนวณโดยคลาส StatTable ซึ่งเป็นคลาสที่ผู้วิจัยนำมาใช้ในการคำนวณค่าสถิติทดสอบ (คลาสนี้ผู้วิจัยไม่ได้พัฒนาขึ้นเอง) แต่อย่างไรก็ตามค่าขอบเขตบนและขอบเขตล่างเป็นเพียงค่าประมาณทางสถิติเท่านั้น ดังนั้นค่าสถิติที่คำนวณได้โดยส่วนประกอบซอฟต์แวร์ซึ่งมีความใกล้เคียงกับผลลัพธ์จากโปรแกรม SPSS มากจึงสามารถใช้เป็นค่าอ้างอิงได้ นอกจากนี้ผู้วิจัยได้ทดสอบความสามารถในการคำนวณค่าสถิติสำหรับกรณีที่ปริมาณของข้อมูลที่นำมาใช้เป็นข้อมูลจำนวนมาก โดยผู้วิจัยได้ทำการทดสอบส่วนประกอบซอฟต์แวร์โดยใช้ข้อมูลจำนวนถึง 25,000 เรคคอร์ดและเลือกการวิเคราะห์ความถดถอยเชิงเส้นแบบสเตปไวส์ซึ่งเป็นวิธีการที่ซับซ้อนที่สุดโดยเลือกตัวแปรอิสระในการวิเคราะห์ถึง 7 ตัว พบว่าโปรแกรม Stat 2000 ยังสามารถทำงานได้ โดยคุณสมบัติของเครื่องที่ใช้ทดสอบเป็นเครื่องเพนเทียมทู แรมขนาด 64 เมกะไบต์ ระบบปฏิบัติการวินโดวส์ 98

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

โดยสรุปแล้วงานวิจัยนี้มีประโยชน์ต่างๆคือ

1. สามารถนำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นไปใช้ร่วมในการพัฒนาโปรแกรมเชิงวัตถุอื่นๆได้

   ในการพัฒนาโปรแกรมเชิงวัตถุหากมีความจำเป็นในการคำนวณค่าสถิติเชิงพรรณนาหรือการคำนวณค่าสถิติความถดถอยเชิงเส้นก็สามารถนำส่วนประกอบซอฟต์แวร์ที่ได้ผ่านการทดสอบนี้ไปใช้ร่วมในการพัฒนาโปรแกรมได้ทันที ซึ่งช่วยลดเวลาในการพัฒนาโปรแกรมทำให้สามารถพัฒนาโปรแกรมได้อย่างรวดเร็ว

2. ผู้ใช้สามารถนำโปรแกรมไปใช้ในการคำนวณค่าสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นได้

   โปรแกรมที่ใช้ทดสอบการเรียกใช้ส่วนประกอบซอฟต์แวร์ที่ได้พัฒนาขึ้นสามารถนำไปใช้ในการวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นได้ เนื่องจากมีการตรวจสอบความถูกต้องของค่าสถิติที่คำนวณได้กับโปรแกรม SPSS เวอร์ชัน 8.0 แล้ว

3. ส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับสถิติเชิงพรรณนาและการวิเคราะห์ความถดถอยเชิงเส้นที่ได้พัฒนาขึ้นจะเป็นแบบอย่างสำหรับการออกแบบและพัฒนาส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติอื่นๆ

   เนื่องจากงานวิจัยนี้เป็นงานวิจัยแรกที่จัดทำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติ ผู้วิจัยหวังเป็นอย่างยิ่งที่จะให้งานวิจัยนี้เป็นแบบอย่างเพื่อเกิดการวิจัยเพิ่มเติมและได้ส่วนประกอบซอฟต์แวร์เชิงวัตถุที่ครอบคลุมการวิเคราะห์สถิติทั้งหมด

8.2 ข้อเสนอแนะ

งานวิจัยนี้เป็นแนวทางสำหรับผู้วิจัยอื่นๆ ที่สนใจในการจัดทำส่วนประกอบซอฟต์แวร์เชิงวัตถุสำหรับการวิเคราะห์ทางสถิติสามารถทำวิจัยเพื่อเพิ่มเติมความสามารถในการวิเคราะห์ทางสถิติอื่นๆลงในส่วนประกอบซอฟต์แวร์เชิงวัตถุของผู้วิจัยที่ได้จัดทำขึ้นทั้งนี้เพื่อให้ส่วนประกอบซอฟต์แวร์นี้มีความสมบูรณ์และสามารถนำไปใช้ในการวิเคราะห์ทางสถิติได้ครอบคลุมมากขึ้น

รายการอ้างอิง

[1] กัลยา วานิชย์บัญชา. การวิเคราะห์ข้อมูลด้วย SPSS for Windows กรุงเทพมหานคร: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2542.

[2] กัลยา วานิชย์บัญชา. สถิติเพื่อการตัดสินใจ กรุงเทพมหานคร: โรงพิมพ์จุฬาลงกรณ์มหาวิทยาลัย, 2542.

[3] Hans-Eric Erickson and Magnus Penker. UML Toolkit. (New York: John Wiley & Sons, Inc.): 1998

[4] G. Booch. Object-Oriented Analysis and Design with Applications. (Redwood City, CA : Benjamin Cummings): 1994.

[5] G. Booch, J. Rambaugh, I. Jacobson. Unified Modeling Language Semantics And Notation Guide 1.0 . (San Jose, CA: Rational Software Corporation): 1997.

[6] Clemens Szyperski. Component Software (Beyond Object-Oriented Programming). (Reading, MA: Addison-Wesley): 1998.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ประวัติผู้เขียนวิทยานิพนธ์

นาย ชาญวิทย์ สินธพพันธุ์ เกิดเมื่อวันที่ 26 มิถุนายน พ.ศ. 2515 สำเร็จการศึกษาหลักสูตรวิทยาศาสตร์บัณฑิต (วท.บ.) สาขา คณิตศาสตร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย เมื่อปีการศึกษา 2537 หลังจากนั้นได้ทำงานในบริษัท เทเลคอมเอเซียคอร์ปอเรชั่นจำกัด จนถึง พ.ศ. 2541 จึงเข้าศึกษาต่อหลักสูตรวิทยาศาสตร์มหาบัณฑิต (วท.ม.) ภาควิชาวิศวกรรมศาสตร์คอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย