วิทยานิพนธ์

ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร

The Factors Affect to Retail Oil Price by Multivariate Regression Analysis

นางสาวภัทริณี คงชู

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

พ.ศ. 2551

ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตรมหาบัณฑิต (สถิติ)

ปริญญา

สถิติ — **สาขา**

สถิติ — **ภาควิชา**

เรื่อง ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร

The Factors Affect to Retail Oil Price by Multivariate Regression Analysis

นามผู้วิจัย นางสาวภัทริณี คงชู

ได้พิจารณาเห็นชอบโดย

ประธานกรรมการ (ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมที, Ph.D.)

กรรมการ (อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

กรรมการ (อาจารย์วสันต์ ทองไทย, กศ.ด.)

หัวหน้าภาควิชา (อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

(รองศาสตราจารย์กัญจนา ธีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ 4 **เดือน** มิถุนายน **พ.ศ.** 2551

ใบรับรองวิทยานิพนธ์

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์

วิทยาศาสตรมหาบัณฑิต (สถิติ)

ปริญญา

สถิติ | สถิติ

สาขา | **ภาควิชา**

เรื่อง ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร

The Factors Affect to Retail Oil Price by Multivariate Regression Analysis

นามผู้วิจัย นางสาวภัทริณี คงชู

ได้พิจารณาเห็นชอบโดย

ประธานกรรมการ .....

(ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมที, Ph.D.)

กรรมการ .....

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

กรรมการ .....

(อาจารย์วสันต์ ทองไทย, กศ.ด.)

หัวหน้าภาควิชา .....

(อาจารย์อำไพ ทองธีรภาพ, Ph.D.)

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์รับรองแล้ว

.....

(รองศาสตราจารย์กัญจนา ธีระกุล, D.Agr.)

คณบดีบัณฑิตวิทยาลัย

วันที่ ..... **เดือน** ..... **พ.ศ.** .....

วิทยานิพนธ์

เรื่อง

ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก
โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร

The Factors Affect to Retail Oil Price by
Multivariate Regression Analysis

โดย

นางสาวภัทริณี คงชู

เสนอ

บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์
เพื่อความสมบูรณ์แห่งปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ)
พ.ศ. 2551

ภัทริณี คงชู 2551: ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร ปริญญาวิทยาศาสตรมหาบัณฑิต (สถิติ) สาขาสถิติ ภาควิชาสถิติ ประธานกรรมการที่ปรึกษา: ผู้ช่วยศาสตราจารย์บุญอ้อม โฉมที, Ph.D. 134 หน้า

งานวิจัยนี้ มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก 3 ประเภท ในกรุงเทพมหานคร คือ น้ำมันดีเซล (Y_1) น้ำมันเบนซิน 91 (Y_2) และน้ำมันเบนซิน 95 (Y_3) และศึกษารูปแบบความสัมพันธ์ระหว่างราคาน้ำมันขายปลีกกับปัจจัยต่าง ๆ ที่คาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลงของราคาน้ำมันขายปลีก 3 ประเภทนี้ ซึ่งศึกษาข้อมูลทุติยภูมิ เป็นรายเดือน ตั้งแต่เดือนมกราคม 2545 – เดือนธันวาคม 2549 จากหน่วยงาน ได้แก่ บริษัท ปตท. จำกัดมหาชน กรมธุรกิจพลังงาน สำนักงานสถิติ เป็นต้น และวิธีการวิเคราะห์ทางสถิติ คือ การวิเคราะห์สหสัมพันธ์ การวิเคราะห์แคนนอลนิคอล การวิเคราะห์ปัจจัย และการวิเคราะห์การถดถอยแบบหลายตัวแปร การศึกษานี้ประกอบด้วยตัวแปรตาม 3 ตัวแปร และตัวแปรอิสระ 74 ตัวแปร

ผลการวิเคราะห์สหสัมพันธ์ พบว่าจากตัวแปรอิสระทั้งหมด 74 ตัวนั้นมีจำนวน 33 ตัวแปรที่มีความสัมพันธ์กับตัวแปรตามทั้ง 3 ตัว อย่างมีนัยสำคัญที่ระดับนัยสำคัญ 0.05 และเนื่องจากตัวแปรอิสระทั้ง 33 ตัวแปรดังกล่าวมีความไม่เป็นอิสระกัน (Multicollinearity) ผู้วิจัยจึงแก้ปัญหาโดยเทคนิคการวิเคราะห์ปัจจัย ใช้การหมุนแบบ Orthogonal ด้วยวิธี Varimax ซึ่งสามารถจำแนกตัวแปรอิสระ 33 ตัวเป็นปัจจัย 3 ปัจจัย คือ ปัจจัยด้านราคา (Factor 1) ปัจจัยด้านการผลิตและส่งออกน้ำมัน (Factor 2) และปัจจัยด้านเศรษฐกิจ (Factor 3) ซึ่งปัจจัยทั้งหมดนี้สามารถอธิบายความผันแปรของตัวแปรอิสระทั้ง 33 ตัวได้ 91.041% โดยที่

Factor 1 (ปัจจัยด้านราคา) = $-0.003X_3-0.001X_4+0.005X_5-0.003X_6-0.001X_7+0.006X_8+0.029X_9+0.031X_{10}+0.066X_{11}+0.266X_{19}+0.137X_{20}+0.000X_{22}+0.144X_{23}+0.024X_{25}+0.017X_{26}+0.003X_{27}-0.041X_{28}+0.023X_{29}+0.037X_{30}-0.115X_{31}+0.153X_{34}+0.088X_{36}+0.003X_{40}+0.041X_{65}+0.032X_{69}+0.050X_{72}$

Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) = $0.196X_{37}+0.054X_{40}+0.411X_{47}+0.290X_{48}$

Factor 3 (ปัจจัยด้านเศรษฐกิจ) = $0.171X_{52}+0.435X_{68}+0.370X_{71}$

หลังจากแก้ปัญหา Multicollinearity แล้วจึงทำการวิเคราะห์ถดถอย เพื่อหารูปแบบความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภทกับปัจจัยทั้ง 3 ปัจจัย และพบว่าสมการถดถอย คือ

$\hat{Y}_1 = 17.547+4.239$Factor 1$+2.057$Factor 2$+1.733$Factor 3 ; $R_1^2 = 96.2$, $S.D_1 = 1.027$,

$\hat{Y}_2 = 19.619+3.66$Factor 1$+2.536$Factor 2$+2.192$Factor 3 ; $R_2^2 = 98.3$, $S.D_2 = 0.673$,

$\hat{Y}_3 = 20.489+3.627$Factor 1$+2.504$Factor 2$+2.126$Factor 3 ; $R_3^2 = 98.3$, $S.D_3 = 0.667$ ตามลำดับ

จะเห็นว่าสมการถดถอยทั้ง 3 สมการ สามารถอธิบายการเปลี่ยนแปลงของราคาน้ำมันขายปลีก 3 ประเภท โดยเฉลี่ยประมาณ 97.6% ซึ่งมีความถูกต้องแม่นยำในการประมาณค่าสูง ส่วนค่าความคลาดเคลื่อนในการประมาณค่าโดยเฉลี่ย 0.789 ซึ่งมีค่าต่ำมาก ดังนั้นผลการศึกษาที่ได้จึงมีความน่าเชื่อถือมากพอสมควรแก่การนำไปใช้ประโยชน์ในการพยากรณ์ราคาน้ำมันขายปลีกทั้ง 3 ประเภท

_____ / / _____

ลายมือชื่อนิสิต ลายมือชื่อประธานกรรมการ

Patrinee Khongchoo 2008: The Factors Affect to Retail Oil Price by Multivariate Regression Analysis. Master of Science (Statistics), Major Field: Statistics, Department of Statistics. Thesis Advisor: Assistant Professor Boonorm Chomtee, Ph.D. 134 pages.

The purpose of this research is to study the factors that affect to the three categories of the retail oil prices in Bangkok, i.e., Diesel, Unleaded Regular Gasolin 91, and Unleaded Gasolin 95. The study is based on the secondary data from the governmental and private organizations collected by month from the period January 2002 to December 2006. The Correlation Analysis, the Canonical Correlation Analysis, the Factor Analysis, and the Multivariate Regression Analysis are the statistical methodologies in this research. The dependent and independent variables in the study are comprised of 3 and 74 variables, respectively.

The study finds that among the total 74 independent variables, only 33 independent variables have significant relationship to the three dependent variables at the significant level of 0.05. Such 33 independent variables, however, encounter the Multicollinearity problem. Therefore, the study turns to the Factor Analysis by Orthogonal Method via Varimax. From the factor analysis, the three related factors are determined, these are,the price (Factor 1), the production and export of oil (Factor 2), and the economic factors (Factor 3). Such triple factors elucidate the variance of the 33 independent variables by 91.041% as follows:

Factor 1 (the Price) = $-0.003X_3-0.001X_4+0.005X_5-0.003X_6-0.001X_7+0.006X_8+0.029X_9+0.031X_{10}+0.066X_{11}+0.266X_{19}+0.137X_{20}+0.000X_{22}+0.144X_{23}+0.024X_{25}+0.017X_{26}+0.003X_{27}-0.041X_{28}+0.023X_{29}+0.037X_{30}-0.115X_{31}+0.153X_{34}+0.088X_{36}+0.003X_{40}+0.041X_{65}+0.032X_{69}+0.050X_{72}$

Factor 2 (the Production and Export of Oil) = $0.196X_{37}+0.054X_{40}+0.411X_{47}+0.290X_{48}$

Factor 3 (the Economic Factors) = $0.171X_{52}+0.435X_{68}+0.370X_{71}$.

Then, the Multivariate Regression Analysis is used to find out for the relationship model between the 3 categories of the retail oil price and the 3 factors. The regression equations are,

$\hat{Y}_1$ =17.547+4.239Factor 1+2.057Factor 2+1.733Factor 3 ; R_1^2 = 96.2 , $S.D_1$ = 1.027,

$\hat{Y}_2$ =19.619+3.66Factor 1+2.536Factor 2+2.192Factor 3 ; R_2^2 = 98.3 , $S.D_2$ = 0.673,

and $\hat{Y}_3$ = 20.489+3.627Factor 1+2.504Factor 2+2.126Factor 3 ; R_3^2 = 98.3 , $S.D_3$ = 0.667, respectively.

The above regression equations can explain approximately the change in the 3 kinds of the retail oil prices by 97.6%, with the average error 0.789. Hence, the results of this paper provide the useful regression equations for forecasting these retail oil prices.

_____ _____ ___ / ___ / ___

Student's signature Thesis Advisor's signature

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.บุญอ้อม โฉมที ประธานกรรมการที่ปรึกษาวิทยานิพนธ์ ที่ได้ช่วยเหลือในการวางแผนงานจัดทำวิทยานิพนธ์ฉบับนี้ ตลอดจนการให้คำปรึกษา แนะนำ และตรวจแก้ไขข้อบกพร่องต่าง ๆ ขอกราบขอบพระคุณ ดร.อำไพ ทองธีรภาพ กรรมการที่ปรึกษาวิชาเอก ดร.วสันต์ ทองไทย กรรมการที่ปรึกษาวิชารอง และ ดร.วรัทยา ธรรมกิตติภพ ผู้แทนบัณฑิตวิทยาลัย ที่กรุณาให้คำปรึกษา แนะนำ และช่วยเหลือตรวจแก้ไขวิทยานิพนธ์ให้สำเร็จลุล่วงตามวัตถุประสงค์และสมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณ เจ้าหน้าที่ บริษัท ปตท. จำกัดมหาชน, กรมธุรกิจพลังงาน, กรมสรรพสามิต สำนักงานสถิติแห่งชาติ, ธนาคารแห่งประเทศไทย, และกระทรวงพลังงาน ที่กรุณาให้ข้อมูล และคำแนะนำ พร้อมทั้งอำนวยความสะดวกเกี่ยวกับการได้มาซึ่งข้อมูล

ขอขอบพระคุณ คุณพ่อ คุณแม่ และพี่ ๆ น้อง ๆ ทุกคน ที่คอยให้ความช่วยเหลือ ให้กำลังใจ ชี้แนะสนับสนุน การทำวิทยานิพนธ์จนสำเร็จลุล่วงไปได้ด้วยดี

ภัทริณี คงชู

เมษายน 2551

สารบัญ

หน้า

สารบัญ	(1)
สารบัญตาราง	(2)
สารบัญภาพ	(4)
คำนำ	1
วัตถุประสงค์	4
การตรวจเอกสาร	9
อุปกรณ์และวิธีการ	59
อุปกรณ์	59
วิธีการ	59
ผลและวิจารณ์	62
ผล	62
วิจารณ์	91
สรุปและข้อเสนอแนะ	92
สรุป	92
ข้อเสนอแนะ	95
เอกสารและสิ่งอ้างอิง	97
ภาคผนวก	100
ภาคผนวก ก ข้อมูลตัวอย่างใหม่ที่ใช้ในการพยากรณ์	101
ภาคผนวก ข ผลการวิเคราะห์ข้อมูล โดยใช้โปรแกรม SAS	105
ภาคผนวก ค ตัวอย่างโปรแกรม SAS ที่ใช้ในการวิเคราะห์ข้อมูล	127
ประวัติการศึกษาและการทำงาน	134

สารบัญตาราง

ตารางที่		หน้า
1	เมตริกซ์ข้อมูลเบื้องต้นสำหรับการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล	29
2	ข้อมูลเบื้องต้นของตัวแปรตาม (Y_1, Y_2, Y_3)	62
3	ความสัมพันธ์ระหว่างตัวแปรอิสระ 74 ตัวแปรกับตัวแปรตาม 3 ตัวแปร	63
4	การทดสอบค่าสหสัมพันธ์แคนนอลนิคอลสำหรับ Y_1, Y_2, Y_3 และตัวแปรอิสระทั้ง 33 ตัวแปร โดยวิธี Wilks' Lambda	67
5	การทดสอบค่าสหสัมพันธ์แคนนอลนิคอลทีละคู่ โดยวิธี Wilks' Lambda	67
6	ค่าสหสัมพันธ์ระหว่างตัวแปรแคนนอลนิคอลกับตัวแปรอิสระและตัวแปรตาม	68
7	ผลการทดสอบการแจกแจงปกติหลายตัวแปรของความคลาดเคลื่อนด้วยวิธีการของMardia	71
8	ค่าความผันแปรจากการสกัดปัจจัย	72
9	ค่าน้ำหนักปัจจัย โดยการหมุนด้วยวิธี Varimax	73
10	คะแนนปัจจัยโดยวิธีการวิเคราะห์การถดถอย	77
11	สัมประสิทธิ์คะแนนปัจจัย	79
12	การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัว กับตัวแปรอิสระ 3 ตัว หรือปัจจัย 3 ปัจจัย	82
13	การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัว กับตัวแปรอิสระ (Factor 1)	83
14	การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัว กับตัวแปรอิสระ (Factor 2)	83
15	การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัว กับตัวแปรอิสระ (Factor 3)	84
16	ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันดีเซล	84
17	ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันเบนซิน 91	85
18	ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันเบนซิน 95	86
19	ค่าสถิติเบื้องต้นสำหรับตัวแปรตาม 3 ตัวแปร ที่ทำการศึกษาจากหน่วยใหม่ (ข้อมูลเดือนมกราคม 2550 – เดือนตุลาคม 2550)	88
20	ผลการพยากรณ์ของหน่วยตัวอย่างใหม่ เทียบกับข้อมูลจริง	90

สารบัญตาราง (ต่อ)

ตารางผนวกที่		หน้า
ก1	ข้อมูลตัวอย่างใหม่ที่ใช้ในการพยากรณ์	102
ข3	ผลการวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรอิสระ 33 ตัว	111

สารบัญภาพ

ภาพที่		หน้า
1	ความสัมพันธ์เชิงเส้นและความสัมพันธ์ไม่ใช่เชิงเส้น	11
2	แผนผังของการศึกษาวิจัย	61
3	การพล็อตค่าคลาดเคลื่อนกับค่าพยากรณ์	71

ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก
โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร

The Factors Affect to Retail Oil Price by
Multivariate Regression Analysis

คำนำ

น้ำมันปิโตรเลียมเป็นทรัพยากรที่มีบทบาทสำคัญในการผลิตสินค้า การคมนาคมขนส่ง และเป็นปัจจัยที่สะท้อนถึงการพัฒนาเศรษฐกิจที่สำคัญมาก เนื่องจากราคาน้ำมันจะส่งผลโดยตรงต่อต้นทุนการผลิตสินค้ารวมถึงภาวะเงินเฟ้อ หากประเทศมีการพัฒนาทางเศรษฐกิจที่เพิ่มขึ้นความต้องการน้ำมันปิโตรเลียมจะเพิ่มขึ้นเช่นเดียวกัน แต่เนื่องจากประเทศไทยไม่สามารถผลิตน้ำมันให้เพียงพอต่อความต้องการภายในประเทศ จึงเป็นผลให้ประเทศไทยต้องมีการนำเข้าน้ำมันจากต่างประเทศ ทั้งในรูปน้ำมันดิบและน้ำมันสำเร็จรูป โดยน้ำมันดิบจะนำเข้าจากแหล่งใหญ่ 2 แหล่ง คือ กลุ่มประเทศตะวันออกกลาง ได้แก่ ประเทศอิหร่าน อิรัก อิสราเอล จอร์แดน คูเวต เลบานอน โอมาน ซาอุดีอาระเบีย ซีเรีย เยเมน และดินแดนปาเลสไตน์และประเทศตะวันออกไกล ได้แก่ ประเทศมาเลเซีย สิงคโปร์ ในขณะที่น้ำมันสำเร็จรูปจะนำเข้าจากประเทศสิงคโปร์ ประเทศไทยนำเข้าน้ำมันดิบและน้ำมันสำเร็จรูปบางส่วนรวมกันมี สัดส่วนประมาณร้อยละ 90 ของความต้องการใช้น้ำมันเชื้อเพลิงภายในประเทศ ซึ่งตลาดสิงคโปร์เป็นศูนย์กลางการซื้อขายน้ำมันที่สำคัญที่สุดในเอเชีย (สำนักงานคณะกรรมการนโยบายพลังงานแห่งชาติ, 2547)

ประเทศไทยยังคงต้องพึ่งพาการนำเข้าน้ำมันดิบและน้ำมันสำเร็จรูปจากต่างประเทศ เนื่องจากประเทศไทยยังไม่มีความสามารถในการผลิตและจัดหาน้ำมันให้มีปริมาณเพียงพอต่อความต้องการในประเทศได้ จึงเป็นผลทำให้ราคาน้ำมันรวมถึงต้นทุนในการกลั่นน้ำมัน มีการเปลี่ยนแปลงตามราคาต่างประเทศ หรือตามราคาตลาดโลก ซึ่งในปัจจุบันประเทศไทยได้ยึดราคาประกาศ ณ ประเทศสิงคโปร์ เป็นหลักเกณฑ์ในการกำหนดราคาหน้าโรงกลั่น (Ex-refinery Price) ประเทศสิงคโปร์มีโรงกลั่นน้ำมันถึง 6 แห่ง ถือได้ว่าเป็นแหล่งผลิตที่สำคัญในภูมิภาคนี้ จึงเป็นผลให้ราคาน้ำมันขายปลีกในประเทศมีความเคลื่อนไหวตามราคาน้ำมันของตลาดสิงคโปร์ (นุชชฎา, 2547)

สาเหตุอีกประการที่ทำให้ประเทศไทยต้องมีการนำเข้าน้ำมันสำเร็จรูปจากต่างประเทศ เนื่องจากในอดีตที่ผ่านมาประเทศไทยมีโรงกลั่นน้ำมันในประเทศเพียง 3 โรงกลั่นได้แก่ โรงกลั่นไทยออยล์ โรงกลั่นบางจาก และโรงกลั่นเอสโซ่ มีกำลังการผลิตรวม 52.1 ล้านลิตร ต่อวัน (เมื่อ พ.ศ. 2536) ซึ่งไม่เพียงพอต่อความต้องการภายในประเทศ เป็นผลให้ต้องมีการนำเข้าน้ำมันสำเร็จรูปจากสิงคโปร์ ปัจจุบันมีโรงกลั่นเพิ่มขึ้นอีก 3 โรงกลั่น ได้แก่ โรงกลั่นของ TPI โรงกลั่นสตาร์ปิโตรเลียม โรงกลั่นระยอง ซึ่งมีกำลังการผลิตรวมกันเป็น 815,000 บาร์เรลต่อวัน หรือประมาณ 129,585,000 ลิตรต่อวัน (นุชชฎา, 2547)

สถานการณ์น้ำมันของโลกมีการเปลี่ยนแปลงอยู่ตลอดเวลา เนื่องจากความไม่สงบทางการเมือง สงครามและความขัดแย้งทางด้านการเมืองและเศรษฐกิจของประเทศมหาอำนาจผู้ส่งออกน้ำมันของโลกทำให้ประเทศผู้ส่งออกน้ำมันทั้งหลายทำการลดปริมาณการผลิตและการส่งออกน้ำมัน เป็นผลให้เกิดการขาดแคลนน้ำมันและราคาน้ำมันปรับตัวสูงขึ้นทั่วโลก และเป็นที่แน่นอนว่าประเทศผู้นำเข้าน้ำมันอย่างประเทศไทยต้องได้รับผลกระทบต่อเศรษฐกิจของประเทศอย่างหลีกเลี่ยงไม่ได้ ทั้งผลต่อดุลการค้าระหว่างประเทศ ผลต่อภาวะเงินเฟ้อที่สูงขึ้นจากราคาน้ำมันที่สูงขึ้น (เกียรติชัย, 2536) และส่งผลถึงภาคเอกชนที่ยังต้องรับภาระต้นทุนที่สูงขึ้นและผันผวนจากราคาน้ำมันที่เปลี่ยนแปลง ดังนั้นรัฐบาลจึงมีนโยบายที่จะแก้ปัญหาและลดผลกระทบจากความผันผวนของระดับราคาน้ำมันในตลาดโลกที่มีต่อเศรษฐกิจของประเทศ โดยจัดตั้งกองทุนน้ำมันขึ้นในปี พ.ศ.2520 เพื่อช่วยรักษาระดับราคาขายปลีกน้ำมันเชื้อเพลิงภายในประเทศ ในระยะสั้นที่ได้รับผลกระทบจากการเปลี่ยนแปลงของราคาน้ำมันในตลาดโลกและมีการใช้มาตรการทางภาษีเข้าแทรกแซงและรักษาระดับราคาน้ำมันเชื้อเพลิงภายในประเทศ แต่จากการควบคุมและแทรกแซงราคาน้ำมันเชื้อเพลิงของรัฐบาลนี้เอง ทำให้ราคาน้ำมันเชื้อเพลิงถูกใช้เป็นเครื่องมือทางการเมืองของนักการเมืองและรัฐบาล ในการหาเสียงและรักษาสถานภาพทางการเมือง โดยใช้เงินจำนวนมากในการพยุงราคาน้ำมันให้อยู่ในระดับที่ต่ำและไม่สะท้อนต่อราคาน้ำมันในตลาดโลก เป็นผลให้เกิดปัญหาเงินทุนสำรองของรัฐบาล และสถานะทางการเงินของประเทศ (โสภณ, 2533) ดังนั้นเพื่อแก้ปัญหาที่เกิดขึ้นรัฐบาลจึงยกเลิกการควบคุมราคาน้ำมันเชื้อเพลิง โดยนำระบบราคาน้ำมันลอยตัวมาใช้ตั้งแต่วันที่ 19 สิงหาคม พ.ศ.2534 เพื่อให้ราคาขายปลีกน้ำมันเชื้อเพลิงภายในประเทศมีการปรับตัวขึ้นลงตามสภาวะของตลาดและต้นทุนที่แท้จริงมากขึ้น และเป็นอิสระต่อการตัดสินใจทางการเมืองของรัฐบาล นอกจากการนำระบบราคาน้ำมันลอยตัวมาใช้แล้ว ยังมีการส่งเสริมให้มีการเพิ่มกำลังการผลิตน้ำมันสำเร็จรูปภายในประเทศ เพื่อลดการพึ่งพิงการนำเข้าน้ำมัน

สำเร็จรูปและปัญหาการขาดดุลการค้าอีกด้วย โดยส่งเสริมให้มีการขยายกำลังการผลิตของโรงกลั่นภายในประเทศ รวมถึงการเปิดเสรีให้มีการแข่งขันในธุรกิจน้ำมันมากขึ้น

ตั้งแต่ประเทศไทยได้นำระบบราคาน้ำมันลอยตัวเข้ามาใช้เมื่อ พ.ศ.2534 ในตลาดค้าปลีกน้ำมันเชื้อเพลิง มีการแข่งขันของผู้ค้าน้ำมันในระดับสูง การเพิ่มขึ้นของผู้ค้าน้ำมันจากเดิมมีเพียงผู้ค้ารายใหญ่ 4 ราย ได้เพิ่มขึ้นเป็น 29 ราย ปัจจุบันทำให้จำนวนสถานีบริการน้ำมันได้เพิ่มขึ้นอย่างรวดเร็ว โดยเพิ่มขึ้นมากที่สุดถึง 37.2% และ 39.0% ในปี 2537 และ 2538 ตามลำดับ เมื่อเปรียบเทียบกับปี 2532 มีเพียง 15.0% ทำให้การแข่งขันในระดับค้าปลีกของไทยอยู่ในระดับที่สูงมาก ติดอันดับ 1 ใน 5 ของโลก (สำนักงานคณะกรรมการนโยบายพลังงานแห่งชาติ, 2547) ส่งผลให้ผู้ค้าน้ำมันรายย่อยไม่สามารถกำหนดราคาขายปลีกในระดับที่แตกต่างกันได้มาก จากการศึกษาของสำนักงานคณะกรรมการนโยบายพลังงานแห่งชาติพบว่า การกำหนดราคาน้ำมันขายปลีกระดับที่แตกต่างกัน 50 สตางค์ต่อลิตร จะส่งผลให้ปริมาณการจำหน่ายของผู้ค้าน้ำมันลดลงครึ่งหนึ่ง ค่าการตลาดของผู้ค้าน้ำมันนับตั้งแต่ปี 2539 ได้ปรับลดลงในระดับทรงตัว ซึ่งเป็นผลจากการแข่งขันที่อยู่ในระดับสูง

เนื่องจากความเคลื่อนไหวของราคาน้ำมันนั้นส่งผลกระทบต่อดุลการค้าและดุลการชำระเงินของไทย และน้ำมันเป็นปัจจัยหนึ่งที่มีความจำเป็นต่อระบบเศรษฐกิจของประเทศไทย ทั้งด้านที่เป็นความต้องการพื้นฐานการพัฒนาอุตสาหกรรม และกิจกรรมทางเศรษฐกิจของประเทศไทย ดังนั้นผู้วิจัยจึงเล็งเห็นความสำคัญและสนใจรูปแบบความสัมพันธ์ของราคาน้ำมันและวิเคราะห์ปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก

ในการศึกษาครั้งนี้ ผู้วิจัยจะทำการศึกษาเฉพาะราคาน้ำมันขายปลีกในเขตกรุงเทพมหานคร โดยศึกษาราคาน้ำมัน 3 ประเภท คือ น้ำมันดีเซล (Diesel) น้ำมันเบนซิน 91 (Unleaded Regular Gasoline 91) และ น้ำมันเบนซิน 95 (Unleaded Gasoline 95) เนื่องจากน้ำมันทั้ง 3 ประเภทนี้มีปริมาณการใช้ค่อนข้างมากและใช้กันอย่างกว้างขวาง ส่วนการวิเคราะห์จะใช้วิธีการวิเคราะห์การถดถอยแบบหลายตัวแปร (Multivariate Regression Analysis) เพราะเป็นวิธีการทางสถิติที่นิยมใช้กันมากวิธีหนึ่ง โดยกำหนดตัวแปรที่สนใจเป็นตัวแปรตาม (Dependent Variables) คือราคาน้ำมันเชื้อเพลิงขายปลีก 3 ประเภท ได้แก่ ราคาน้ำมันดีเซล ราคาน้ำมันเบนซิน 91และราคาน้ำมันเบนซิน 95 ส่วนตัวแปรอิสระ (Independent Variables) ได้แก่ อัตราแลกเปลี่ยนดอลลาสหรัฐ ราคาน้ำมันหน้าโรงกลั่น ราคาน้ำมันดิบในตลาดโลก ราคาน้ำมันในตลาดสิงคโปร์ เป็นต้น

วัตถุประสงค์

1. เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อราคาขายปลีกของน้ำมันเชื้อเพลิง 3 ประเภท คือ น้ำมันดีเซล (Diesel) น้ำมันเบนซิน 91 (Unleaded Regular Gasoline 91) และน้ำมันเบนซิน 95 (Unleaded Gasoline 95)

2. เพื่อศึกษารูปแบบความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภท ดังกล่าวกับปัจจัยต่าง ๆ ที่คาดว่าจะมีผลต่อราคาน้ำมันขายปลีก

ประโยชน์ที่คาดว่าจะได้รับ

ผลการศึกษาครั้งนี้ ทำให้ทราบเกี่ยวกับปัจจัยที่มีอิทธิพลต่อราคาขายปลีกน้ำมันเชื้อเพลิง 3 ประเภท คือ น้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 รวมทั้งรูปแบบความสัมพันธ์ระหว่างราคาน้ำมันขายปลีกทั้ง 3 ประเภทกับปัจจัยต่างๆที่คาดว่าจะมีผลต่อราคาน้ำมันขายปลีกทั้ง 3 ประเภท เพื่อใช้เป็นข้อมูลในการพยากรณ์การเปลี่ยนแปลงของราคาน้ำมันเชื้อเพลิงขายปลีกทั้ง 3 ประเภท อีกทั้งยังเป็นประโยชน์แก่หน่วยงานทั้งภาครัฐและเอกชน สำหรับการวางแผนนโยบายด้านพลังงาน เพื่อป้องกันและแก้ไขปัญหาวิกฤตทางเศรษฐกิจอันมีสาเหตุมาจากการเปลี่ยนแปลงของราคาน้ำมันรวมทั้งทราบแนวทางการลดผลกระทบของปัญหาราคาน้ำมันผันผวนและเป็นประโยชน์ต่อประชาชนและหน่วยงานที่เกี่ยวข้องด้วย

ขอบเขตการวิจัย

ในการศึกษาเกี่ยวกับปัจจัยที่มีอิทธิพลต่อราคาน้ำมันเชื้อเพลิงครั้งนี้ ผู้วิจัยกำหนดขอบเขตของการวิจัยดังนี้

1. ศึกษาเฉพาะราคาน้ำมันขายปลีก 3 ประเภท คือ น้ำมันดีเซล (Diesel) น้ำมันเบนซิน 91 (Unleaded Regular Gasoline 91) และ น้ำมันเบนซิน 95 (Unleaded Gasoline 95)ในเขตกรุงเทพมหานคร

2. การศึกษานี้ใช้ข้อมูลทุติยภูมิ (Secondary Data) ซึ่งเก็บรวบรวมข้อมูลเป็นรายเดือน ตั้งแต่เดือนมกราคม 2545 – เดือนธันวาคม 2549 รวมระยะเวลา 5 ปี หรือขนาดของข้อมูลเป็น 60 เดือน โดยมีการรวบรวมจากแหล่งข้อมูลต่าง ๆ เช่น บริษัท ปตท. จำกัดมหาชน กรมธุรกิจพลังงาน ธนาคารแห่งประเทศไทย กรมสรรพสามิต สำนักงานสถิติแห่งชาติ และกระทรวงพลังงาน

3. ตัวแปรที่ใช้ในการศึกษา ประกอบด้วย ตัวแปรตาม และตัวแปรอิสระดังนี้

3.1 ตัวแปรตาม (Dependent Variables) คือ ตัวแปรที่เปลี่ยนแปลงหรือขึ้นอยู่กับตัวแปรอิสระ ซึ่งในการศึกษานี้ คือ ราคาน้ำมันขายปลีก 3 ประเภท ในเขตกรุงเทพมหานคร ได้แก่ น้ำมันดีเซล (Y_1) น้ำมันเบนซิน 91 (Y_2) และน้ำมันเบนซิน 95 (Y_3)

3.2 ตัวแปรอิสระ (Independent Variables) คือ ตัวแปรที่สามารถควบคุมได้ในการทดลองหรือวัดค่าได้ และคาดว่าจะมีอิทธิพลต่อการเปลี่ยนแปลงของตัวแปรตาม ซึ่งในที่นี้มีตัวแปรอิสระที่นำมาศึกษาทั้งสิ้น 74 ตัวแปร ดังนี้

X_1 = อัตราแลกเปลี่ยนดอลลาสหรัฐ(ซื้อ)
X_2 = อัตราแลกเปลี่ยนดอลลาสหรัฐ(ขาย)
X_3 = ราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_4 = ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_5 = ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_6 = ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_7 = ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_8 = ราคาน้ำมันดีเซล นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_9 = ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{10} = ราคาน้ำมันเบนซิน 91 ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{11} = ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{12} = ปริมาณการผลิตน้ำมันสำเร็จรูป (ล้านลิตร)
X_{13} = ปริมาณการส่งออกน้ำมันสำเร็จรูป (ล้านลิตร)
X_{14} = ปริมาณการนำเข้าน้ำมันสำเร็จรูป (ล้านลิตร)
X_{15} = ปริมาณการจำหน่ายน้ำมันสำเร็จรูป (ล้านลิตร)

X_{16} = ปริมาณการนำเข้าน้ำมันดิบ (ล้านลิตร)

X_{17} = ปริมาณการนำเข้าน้ำมันดิบในตะวันออกกลาง (ล้านลิตร)

X_{18} = ปริมาณการนำเข้าน้ำมันดิบในตะวันออกไกล (ล้านลิตร)

X_{19} = อัตราภาษีส่งออกน้ำมันเบนซิน 95 (บาท/ลิตร)

X_{20} = อัตราภาษีส่งออกน้ำมันเบนซิน 91 (บาท/ลิตร)

X_{21} = อัตราภาษีส่งออกน้ำมันดีเซล (บาท/ลิตร)

X_{22} = อัตราภาษีนำเข้าน้ำมันเบนซิน 95 (บาท/ลิตร)

X_{23} = อัตราภาษีนำเข้าน้ำมันเบนซิน 91 (บาท/ลิตร)

X_{24} = อัตราภาษีนำเข้าน้ำมันดีเซล (บาท/ลิตร)

X_{25} = ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (เหรียญสหรัฐ/บาร์เรล)

X_{26} = ราคาน้ำมันดิบในตลาดโลก (โอมาน) (เหรียญสหรัฐ/บาร์เรล)

X_{27} = ราคาน้ำมันเบนซิน 95 ในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)

X_{28} = ราคาน้ำมันเบนซิน 92 ในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)

X_{29} = ราคาน้ำมันดีเซลในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)

X_{30} = ปริมาณการใช้ไฟฟ้า (GWh)

X_{31} = ประชากรรวมในกรุงเทพฯ (พันคน)

X_{32} = จำนวนผู้มีงานทำในกรุงเทพฯ (พันคน)

X_{33} = จำนวนผู้ว่างงานในกรุงเทพฯ (พันคน)

X_{34} = ผลิตภัณฑ์มวลรวมในประเทศ (ล้านบาท)

X_{35} = อัตราเงินเฟ้อ

X_{36} = อัตราดอกเบี้ยเงินกู้ (12 เดือน)

X_{37} = กองทุนน้ำมัน (ล้านบาท)

X_{38} = แหล่งกำเนิดไฟฟ้าจากดีเซล (GWh)

X_{39} = ปริมาณการผลิตน้ำมันดีเซล (ล้านลิตร)

X_{40} = ปริมาณการผลิตน้ำมัน LPG (ล้านลิตร)

X_{41} = ปริมาณการผลิตน้ำมันเบนซิน 95 (ล้านลิตร)

X_{42} = ปริมาณการผลิตน้ำมันเบนซิน 91 (ล้านลิตร)

X_{43} = ปริมาณการใช้น้ำมันเบนซิน 95 (ล้านลิตร)

X_{44} = ปริมาณการใช้น้ำมันเบนซิน 91 (ล้านลิตร)

X_{45} = ปริมาณการใช้น้ำมันดีเซล (ล้านลิตร)

X_{46} = ปริมาณการนำเข้าน้ำมันดีเซล (ล้านลิตร)

X_{47} = ปริมาณการส่งออกน้ำมันเบนซิน 95 (ล้านลิตร)

X_{48} = ปริมาณการส่งออกน้ำมันเบนซิน 91 (ล้านลิตร)

X_{49} = ปริมาณการส่งออกน้ำมันดีเซล (ล้านลิตร)

X_{50} = ปริมาณการจำหน่ายน้ำมันเบนซิน (ล้านลิตร)

X_{51} = ปริมาณการจำหน่ายน้ำมันดีเซล (ล้านลิตร)

X_{52} = จำนวนรถจักรยานยนต์ที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{53} = จำนวนรถยนต์นั่งส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{54} = จำนวนรถโดยสารประจำทางที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{55} = จำนวนรถโดยสารไม่ประจำทางที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{56} = จำนวนรถโดยสารส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{57} = จำนวนรถบรรทุกไม่ประจำทางที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{58} = จำนวนรถบรรทุกส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{59} = มูลค่าการนำเข้ายานพาหนะและอุปกรณ์ขนส่ง (ล้านบาท)

X_{60} = มูลค่าการนำเข้าน้ำมันดิบ (ล้านบาท)

X_{61} = มูลค่าการนำเข้าน้ำเบนซิน (ล้านบาท)

X_{62} = มูลค่าการนำเข้าน้ำมันดีเซล (ล้านบาท)

X_{63} = มูลค่าการนำเข้าก๊าซธรรมชาติ (ล้านบาท)

X_{64} = มูลค่าการนำเข้าก๊าซปิโตรเลียม (ล้านบาท)

X_{65} = มูลค่าการส่งออกรถยนต์ (ล้านบาท)

X_{66} = รายได้ภาษีการนำเข้าน้ำมัน (ล้านบาท)

X_{67} = ปริมาณการจำหน่ายรถยนต์ในประเทศ (คัน)

X_{68} = ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศ (คัน)

X_{69} = มูลค่าการส่งออกผลิตภัณฑ์ปิโตรเลียม (ล้านบาท)

X_{70} = ปริมาณการผลิตรถยนต์ (คัน)

X_{71} = ปริมาณการผลิตรถจักรยานยนต์ (พันคัน)

X_{72} = ปริมาณการผลิตน้ำมันดิบ (1,000 บาร์เรล)

X_{73} = รายได้ภาษีน้ำมันเบนซินในประเทศ (ล้านบาท)

X_{74} = รายได้ภาษีน้ำมันดีเซลในประเทศ (ล้านบาท)

นิยามศัพท์

โรงกลั่น หมายถึง โรงกลั่นน้ำมันเชื้อเพลิง สถานที่ผลิตและจำหน่ายน้ำมันเชื้อเพลิงภายในประเทศ และให้หมายรวมถึงโรงแยกก๊าซในประเทศที่ผลิตและจำหน่ายก๊าซเพื่อใช้ในประเทศและโรงอุตสาหกรรมเคมีปิโตรเลียมและสารละลายด้วย

ราคา ณ โรงกลั่น หมายถึง ราคาน้ำมันที่โรงกลั่นเป็นผู้กำหนดเอง ซึ่งจะเป็นไปตามต้นทุนที่เกิดขึ้นจริง เนื่องจากโรงกลั่นจะต้องมีการแข่งขันกับต้นทุนของการนำเข้าการกำหนดราคาของโรงกลั่น ซึ่งกำหนดให้สอดคล้องและใกล้เคียงกับต้นทุนการนำเข้าให้มากที่สุด เพื่อดึงดูดให้ผู้ค้าน้ำมันซื้อน้ำมันจากโรงกลั่นแทนการนำเข้า

ผลิตภัณฑ์มวลรวมในประเทศ (Gross Domestic Product : GDP) หมายถึง มูลค่าของสินค้าและบริการขั้นสุดท้ายที่ผลิตขึ้นภายในประเทศ ในระยะเวลาหนึ่ง โดยไม่คำนึงถึงว่าทรัพยากรที่ใช้ในการผลิตสินค้า และบริการจะเป็นทรัพยากรของพลเมืองในประเทศหรือเป็นของชาวต่างประเทศ ในทางตรงกันข้ามทรัพยากรของพลเมืองในประเทศแต่ไปทำการผลิตในต่างประเทศก็ไม่นับรวมไว้ในผลิตภัณฑ์ในประเทศ

กองทุนน้ำมัน หมายถึง การเรียกเก็บเงินเข้ากองทุนน้ำมันจากน้ำมันบางประเภท หรือชดเชย (จ่ายออกจากกองทุน) ให้กับน้ำมันบางประเภท เพื่อรักษาระดับราคาค้าปลีกไม่ให้มีการเปลี่ยนแปลงมากนัก โดยอัตราการจัดเก็บเงินเข้ากองทุนของแต่ละผลิตภัณฑ์จะอยู่ในอัตรา 10 สตางค์ต่อลิตร และเป็นอัตราเดียวกับเงินที่หักให้กองทุนสภาพแวดล้อม

ภาษีนำเข้า หมายถึง เงินภาษีที่บริษัทผู้ค้าน้ำมันต้องจ่ายให้แก่รัฐ ในกรณีที่นำเข้าน้ำมันสำเร็จรูปจากต่างประเทศ

น้ำมันดิบ หมายถึง น้ำมันดิบทุกประเภทที่มีการใช้ภายในประเทศและคอนเดนเสทที่ผลิตได้จากอ่าวไทย

น้ำมันสำเร็จรูป หมายถึง ผลิตภัณฑ์ปิโตรเลียมที่ได้จากการกลั่นน้ำมันดิบ ได้แก่ น้ำมันเบนซิน น้ำมันดีเซล น้ำมันก๊าด น้ำมันอากาศยานและน้ำมันเตา

การตรวจเอกสาร

การตรวจเอกสารฉบับนี้แบ่งออกเป็น 2 ส่วน ส่วนแรกกล่าวถึงวิธีการทางสถิติ และส่วนที่สองกล่าวถึงงานวิจัยที่เกี่ยวข้อง โดยมีรายละเอียดดังนี้

วิธีการทางสถิติ

1. การวิเคราะห์สหสัมพันธ์ (Correlation Analysis)

การวิเคราะห์สหสัมพันธ์เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรเชิงปริมาณ 2 ตัวว่ามีความสัมพันธ์กันหรือไม่ โดยไม่มีการกำหนดตัวแปรตาม หรือตัวแปรอิสระ ขนาดความสัมพันธ์ระหว่างตัวแปรแสดงด้วยค่าสัมประสิทธิ์สหสัมพันธ์เพียร์สัน (Pearson Correlation Coefficient) คำนวณได้จากสูตรต่อไปนี้ (สุรินทร์, 2548)

$$r_{xy} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{I=1}^{n}(Y_i - \overline{Y})^2}}$$

เมื่อ r_{xy} แทนค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลที่ได้จากกลุ่มตัวอย่างระหว่างตัวแปร X และ ตัวแปรY ซึ่ง $-1 \leq r_{xy} \leq 1$ ค่าของ r_{xy} เป็นบวกแสดงว่าข้อมูลสองชุดนั้นมีความสัมพันธ์ในทิศทางเดียวกัน ถ้าค่า r_{xy} เป็นลบแสดงว่าข้อมูลสองชุดนั้นมีความสัมพันธ์ในทิศทางตรงกันข้าม

1.1 การทดสอบสมมติฐาน

การทดสอบสมมติฐานว่าตัวแปรทั้งสองมีความสัมพันธ์กันหรือไม่มีสมมติฐานการทดสอบดังนี้

H_0 : ตัวแปรทั้งสองไม่มีความสัมพันธ์กัน

H_1 : ตัวแปรทั้งสองมีความสัมพันธ์กัน

สถิติทดสอบ คือ $t = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}}$; $df = n-2$

ซึ่งจะปฏิเสธสมมติฐาน H_0 ถ้า $t > t_{0.05,n-2}$ และถ้าการทดสอบนี้ปฏิเสธสมมติฐาน H_0 จะสรุปผลการทดสอบว่าตัวแปรทั้งสองตัวแปรมีความสัมพันธ์กัน ที่ระดับนัยสำคัญ 0.05

2. การวิเคราะห์การถดถอย (Regression Analysis)

2.1 การวิเคราะห์การถดถอยตัวแปรเดียว (Univariate Regression Analysis)

ธีระศักดิ์ (2546) ได้กล่าวว่า การวิเคราะห์การถดถอยเป็นการศึกษา เพื่อหาฟังก์ชันหรือรูปแบบความสัมพันธ์ เพื่อใช้ในการพยากรณ์ค่าของตัวแปรตามที่ต้องการศึกษา โดยอาศัยความรู้หรือการเปลี่ยนแปลงเกี่ยวกับค่าของตัวแปรอิสระตัวเดียวหรือหลายตัว ซึ่งสมการที่ต้องการนั้น ควรจะเป็นสมการที่ไม่ซับซ้อน และในขณะเดียวกันก็สามารถประมาณความสัมพันธ์ระหว่างตัวแปรเหล่านั้นให้ได้ใกล้เคียงมากที่สุดด้วย พร้อมทั้งวัดค่าของความคลาดเคลื่อนที่มีความเกี่ยวข้องในกระบวนการของการประมาณด้วย ในกรณีที่มีตัวแปรอิสระ (X) เพียงตัวเดียว ตัวแปร Y กับ X สามารถเขียนตัวแบบการถดถอยในรูป

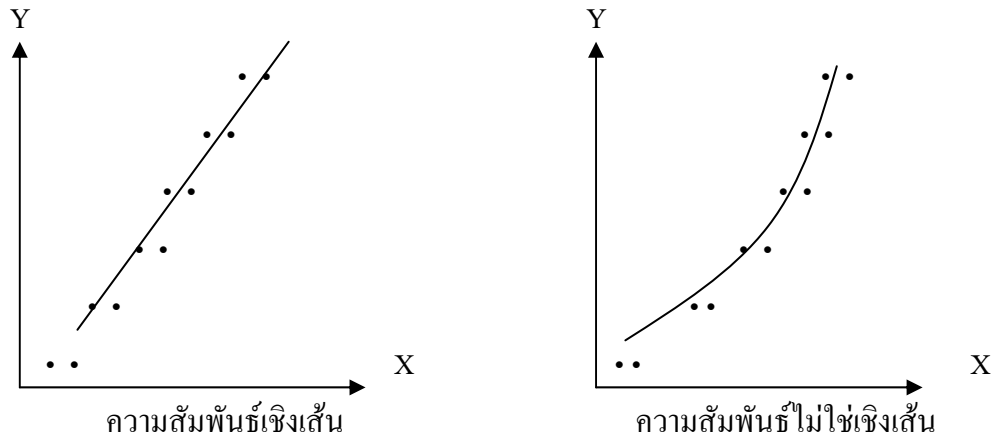
$$Y_j = \beta_0 + \beta_1 x + \varepsilon \quad ; \; j = 1, 2, \ldots, n$$

โดยที่ Y = ค่าของตัวแปรตาม
x = ค่าของตัวแปรอิสระ
β_0 = ค่า y - Intercept
β = ค่าสัมประสิทธิ์การถดถอย
ε = ค่าความคลาดเคลื่อน

เรียกสมการดังกล่าวว่า ตัวแบบการถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression Model)

ในการหาสมการที่ใช้กำหนดความสัมพันธ์ระหว่างตัวแปร X กับ Y นั้น ขั้นแรกให้รวบรวมข้อมูลที่ต้องการจะพิจารณา เช่น $x_1, x_2, ..., x_k$ แทนตัวแปรอิสระและ $y_1, y_2, ..., y_n$ แทน

ตัวแปรตาม ขั้นต่อไปนำคู่อันดับ (x_1,y_1), (x_2,y_2),...,(x_n,y_n) ไปเขียนกราฟ จะได้จุดซึ่งกระจายทั่วๆ ไปบนกราฟ เรียกว่า แผนภาพกระจาย (Scatter Diagram) จากแผนภาพการกระจายจะมองเห็นแนวโน้ม (Trend) ซึ่งอาจจะเป็นความสัมพันธ์ในลักษณะเส้นตรง เรียกว่าความสัมพันธ์เชิงเส้น (Linear Relationship) หรือไม่เป็นเส้นตรง เรียกว่าความสัมพันธ์ไม่ใช่เชิงเส้น (Non-Linear Relationship) ดังแสดงในภาพที่ 1



ภาพที่ 1 ความสัมพันธ์เชิงเส้นและความสัมพันธ์ไม่ใช่เชิงเส้น

ในการหาสมการเพื่อประมาณเส้นโค้งที่เหมาะสมกับข้อมูลที่กำหนดให้ โดยทั่วไปเรียกว่า การปรับเส้นโค้ง (Curve Fitting)

จากที่กล่าวข้างต้นนั้น เป็นการสร้างสมการที่มีตัวแปรอิสระเพียงหนึ่งตัว และตัวแปรตามหนึ่งตัว บางครั้งในสถานการณ์ความเป็นจริง ตัวแปรตามตัวหนึ่งอาจขึ้นอยู่กับตัวแปรอิสระมากกว่าหนึ่งตัว ซึ่งเรียกความสัมพันธ์นี้ว่า การถดถอยเชิงเส้นพหุคูณ และในทางปฏิบัติสิ่งที่สนใจศึกษาไม่ว่าจะเป็นการทดลองหรือในงานวิจัย จะพบว่าตัวแปรที่เข้ามามีบทบาทมักจะมีมากกว่าหนึ่งตัวเสมอ ในที่นี้จะกล่าวถึงความสัมพันธ์ของตัวแปรตามเพียงหนึ่งตัว และตัวแปรอิสระมากกว่าหนึ่งตัว ในลักษณะเชิงเส้นตรง ซึ่งเป็นการสร้างสมการที่เรียกว่า สมการถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression) สามารถเขียนตัวแบบความสัมพันธ์ดังนี้

$$Y_j = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad ; \; j = 1, 2, \ldots, n$$

2.1.1 ข้อสมมติเบื้องต้นสำหรับการวิเคราะห์การถดถอย มีดังนี้

ก. ค่าความคลาดเคลื่อน ε มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวนคงที่ เท่ากับ σ^2

ข. ความคลาดเคลื่อน ε_i และ ε_j เป็นอิสระกัน โดยมีความแปรปรวนร่วมระหว่าง ε_i และ ε_j เท่ากับ 0

ค. ตัวแปรอิสระมีความเป็นอิสระซึ่งกันและกัน

2.1.2 การประมาณค่าสัมประสิทธิ์การถดถอย

ในการประมาณค่าสัมประสิทธิ์การถดถอย โดยวิธีกำลังสองน้อยที่สุด เมื่อกำหนดให้ b_0 และ b_1 เป็นตัวประมาณของ β_0 และ β_1 ตามลำดับ ที่ให้ผลรวมกำลังสองของความคลาดเคลื่อน หรือ SSE มีค่าน้อยที่สุด ซึ่งตัวแบบความสัมพันธ์และสมการถดถอยเชิงเส้นสามารถเขียนได้ดังนี้

$$Y_j = \beta_0 + \beta_1 x + \varepsilon \quad , \quad \hat{Y}_j = b_0 + b_1 x \quad ; \quad j = 1, 2, \ldots, n \qquad \text{ตามลำดับ}$$

ดังนั้น $Y_j - \hat{Y}_j = e_j$ ซึ่งเป็นความคลาดเคลื่อนและผลบวกของค่าคลาดเคลื่อนกำลังสอง (SSE) แสดงได้ดังนี้

$$\begin{aligned}\sum_{j=1}^{n} e_j^2 &= \sum_{j=1}^{n} (Y_j - \hat{Y}_j)^2 \\ &= \sum_{j=1}^{n} (Y_j - b_0 - b_1 x)^2\end{aligned}$$

หาค่า b_0 และ b_1 จากการคำนวณอนุพันธ์ย่อยของ SSE เทียบกับ b_0 และ b_1 แล้วกำหนดให้เท่ากับศูนย์ ดังสมการที่ (1) และ(2) ตามลำดับ

$$\frac{\partial SSE}{\partial b_0} = \frac{\partial \sum_{j=1}^{n} (Y - b_0 - b_1 x)^2}{\partial b_0}$$

$$= -2\sum_{j=1}^{n} (Y - b_0 - b_1 x) = 0 \quad (1)$$

$$\frac{\partial SSE}{\partial b_1} = \frac{\partial \sum_{j=1}^{n} (Y - b_0 - b_1 x)^2}{\partial b_1}$$

$$= -2\sum_{j=1}^{n} x(Y - b_0 - b_1 x) = 0 \quad (2)$$

จากสมการที่ (1) และ (2) จะได้สมการปกติดังนี้

$$\sum_{j=1}^{n} Y_j = nb_0 + b \sum_{j=1}^{n} x_1 \quad (3)$$

$$\sum_{j=1}^{n} xY_j = b_0 \sum_{j=1}^{n} x + b_1 \sum_{j=1}^{n} x_1^2 \quad (4)$$

จากการแก้สมการปกติที่ (3) และ (4) จะได้

$$b_1 = \frac{\sum_{j=1}^{n} (x_i - \overline{x_i})(Y_j - \overline{Y_j})}{\sum_{i=1}^{k} (x_i - \overline{x_i})^2} = \frac{\sum_{j=1}^{n} x_1 Y_j}{\sum_{i=1}^{k} x^2}$$

$$b_0 = \overline{Y} - b_1 \overline{x}$$

และสามารถใช้เมตริกซ์เข้ามาช่วยในการคำนวณค่า b_0 และ b_1 ได้ดังนี้

ตัวแบบสมการถดถอยเชิงเส้นอย่างง่ายสามารถเขียนในรูปเมตริกซ์ดังนี้

$$Y = x\beta + \varepsilon$$

โดยที่

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{nx1}, \quad x = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}_{nx2}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}_{2x1}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}_{nx1}$$

ถ้าให้เมตริกซ์ $\hat{\beta}$ เป็นเมตริกซ์ของค่าประมาณของ β จะได้

$$\hat{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

ซึ่งหาได้จากสมการปกติ ซึ่งเขียนเป็นรูปเมตริกซ์ได้คือ

$$(x'x)\hat{\beta} = x'Y$$

โดยที่

$$x'x = \begin{bmatrix} n & \sum_{j=1}^{n} x_i \\ \sum_{j=1}^{n} x_i & \sum_{j=1}^{n} x^2 \end{bmatrix}$$

และ

$$x'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} Y \\ \sum_{j=1}^{n} xY \end{bmatrix}$$

จะได้ $\hat{\beta} = (x'x)^{-1}x'Y$ ซึ่งเป็นค่าสัมประสิทธิ์การถดถอย เพื่อใช้ในการพยากรณ์ค่าของตัวแปรตาม Y

2.1.3 การทดสอบค่าสัมประสิทธิ์การถดถอย

ในกรณีที่มีตัวแปรอิสระหลายตัวในสมการถดถอย สามารถทดสอบความมีนัยสำคัญของความสัมพันธ์ระหว่างตัวแปรตาม Y และตัวแปรอิสระ $X_1, X_2, X_3, ..., X_k$ โดยใช้การวิเคราะห์ความแปรปรวน และใช้ค่าสถิติ F ทดสอบ ซึ่งเรียกว่า Partial F-Test โดยมีสมมติฐานการทดสอบดังนี้

$H_0: \beta_1 = \beta_2 = \ldots = \beta_k$

$H_1:$ มี β_i อย่างน้อย 1 ค่าที่มีค่าไม่เท่ากับ 0 ; $i = 1, 2, \ldots, k$

สถิติทดสอบ คือ $F = \frac{MSR}{MSE}$

ซึ่งจะปฏิเสธสมมติฐาน H_0 ที่ระดับนัยสำคัญ α ถ้า $F > F_{\alpha,\ k,\ n-k-1}$ และจะสรุปผลการทดสอบได้ว่า ตัวแปรตาม (Y) มีความสัมพันธ์กับตัวแปรอิสระ (X) ทั้ง k ตัว ในรูปของเส้นตรง หมายความว่า สัมประสิทธิ์การถดถอยบางส่วนอย่างน้อย 1 ตัวมีค่าไม่เท่ากับ 0 ดังนั้นจึงควรทดสอบสัมประสิทธิ์การถดถอยบางส่วนแต่ละตัวต่อไป

การวิเคราะห์ความแปรปรวนในการวิเคราะห์ถดถอย

Sov	SS	df	MS	F
การถดถอย (R)	SSR	k	$MSR = \frac{SSR}{k}$	$F = \frac{MSR}{MSE}$
ความคลาดเคลื่อน (E)	SSE	n – k – 1	$MSE = \frac{SSE}{n-k-1}$	
ทั้งหมด	SST	n – 1		

โดยที่ SST (Sum Square of Total) คือ ค่าแปรปรวนทั้งหมดของ Y

$$SST = \sum_{j=1}^{n}(Y_j - \overline{Y}) = Y'Y - n\overline{Y}^2$$

SSR (Sum Square of Regression) คือ ค่าแปรปรวนของ Y เนื่องจากอิทธิพลของ $X_1, X_2, X_3, ..., X_k$ คำนวณจาก

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

SSE (Sum Square of Error) คือ ค่าแปรปรวนของ Y เนื่องจากอิทธิพลอื่น ๆ คำนวณจาก

$$SSE = SST - SSR = Y'Y - b'X'Y$$

2.1.4 การตรวจสอบข้อสมมติของการวิเคราะห์การถดถอย

การตรวจสอบข้อสมมติของการวิเคราะห์การถดถอยโดยส่วนใหญ่จะพิจารณาจากการวิเคราะห์ค่าความคลาดเคลื่อน (Residual Analysis) โดยค่าความคลาดเคลื่อนเป็นค่าความแตกต่างระหว่างค่าจริงกับค่าประมาณ นั้นคือ $Y_j - \hat{Y}_j = e_j$; j = 1, 2, …, n ซึ่งมีรายละเอียดดังต่อไปนี้ (ทรงศิริ, 2548)

ก. การตรวจสอบการแจกแจงปกติของความคลาดเคลื่อนสุ่ม โดยทำ Normal Probability Plot (NPP) ซึ่งเป็นการพล็อตค่าระหว่าง e_j กับค่าคาดหวังของ e_j หรือ $E(e_j)$ เมื่อ e_j เป็นค่าลำดับที่ j ของ e_j นั่นคือ $e_1 \le e_2 \le \ldots \le e_n$ และถ้าจุดที่ได้จากการพล็อตอยู่ในแนวเส้นตรงแสดงว่าความคลาดเคลื่อนมีการแจกแจงปกติ

ข. การตรวจสอบความเป็นอิสระของความคลาดเคลื่อน สามารถตรวจสอบโดยการพล็อต e_j กับค่าประมาณ $\hat{Y}_j$ หรือกับค่าของตัวแปรอิสระ X_j หากค่าความคลาดเคลื่อนเป็นอิสระต่อกัน ค่าความคลาดเคลื่อนจะมีการกระจายรอบค่า 0 อย่างไรก็ตามการพิจารณาจากกราฟจะ

ให้ข้อสรุปที่ไม่ชัดเจน จึงอาจจะทำการทดสอบ โดยใช้การทดสอบของ Durbin และ Watson ซึ่งพิจารณาค่าความคลาดเคลื่อนที่ได้จากการใช้สมการถดถอยที่สร้างด้วยวิธีกำลังสองน้อยที่สุด

ค. การตรวจสอบค่าความคลาดเคลื่อนมีความแปรปรวนคงที่หรือไม่ ทำโดยการพล็อต $\hat{Y}_j$ กับ e_j ถ้าภาพการพล็อตของคู่ลำดับ $(\hat{Y}_j, e_j)$ กระจายในลักษณะเป็นรูปกรวยปลายเปิด แสดงว่าความคลาดเคลื่อนสุ่มมีความแปรปรวนไม่คงที่ เนื่องจากค่า e_j มีการแปรผันตามค่า $\hat{Y}_j$ นอกจากการพิจารณาจากการพล็อตแล้ว ผู้วิเคราะห์อาจจะทดสอบโดยวิธีทางสถิติ ได้แก่ วิธีการทดสอบของบาร์ทเลท (Bartlett's Test) เป็นต้น

ง. การตรวจสอบความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) ซึ่งเป็นปัญหาที่เกิดขึ้นจากการที่ตัวแปรอิสระมีความสัมพันธ์กันในอัตราค่อนข้างสูง ปัญหาดังกล่าวจะทำให้ค่าประมาณของสัมประสิทธิ์การถดถอย และค่าความคลาดเคลื่อนมาตรฐานของตัวประมาณ (S_b) ไม่ถูกต้อง การตรวจสอบปัญหา Multicollinearity อาจทำได้หลายวิธี ดังนี้

1) โดยการทดสอบสมมติฐานว่าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระ x_i กับ x_j ว่ามีค่าเป็น 0 หรือไม่ ถ้าพบว่ามีนัยสำคัญแสดงว่าเกิดปัญหา Multicollinearity

2) โดยพิจารณาขนาดและเครื่องหมายของค่าสัมประสิทธิ์การถดถอย กล่าวคือ ถ้าเพิ่มตัวแปรอิสระใหม่เข้าสู่ตัวแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระแล้วพบว่าค่าสัมประสิทธิ์การถดถอยเดิมเปลี่ยนแปลงทั้งขนาดและเครื่องหมายแสดงว่าอาจเกิดปัญหา Multicollinearity

3) โดยสร้างรูปแบบการถดถอยลดรูปด้วยการสับเปลี่ยนตัวแปรอิสระออกจากรูปแบบการถดถอยเต็มรูปคราวละ 1 ตัว แล้วเปรียบเทียบค่า R^2 หากค่า R^2 จากสมการถดถอยเต็มรูปและ R^2 จากสมการถดถอยลดรูปมีค่าใกล้เคียงกันแสดงว่า อาจเกิดปัญหา Multicollinearity

4) โดยพิจารณาจากการทดสอบสัมประสิทธิ์การถดถอยทั้งหมดในรูปแบบ Overall F-test และการทดสอบสัมประสิทธิ์การถดถอยครั้งละหนึ่งตัว หากพบว่าการทดสอบแรกมีนัยสำคัญและการทดสอบหลังไม่มีนัยสำคัญแสดงว่าอาจเกิดจากปัญหา Multicollinearity

5) ตรวจสอบโดยการพิจารณาค่า Variance Inflation Factor (VIF) ซึ่งคำนวณโดย

$$VIF_j = \frac{1}{1 - R_j^2} \quad ; j = 1, 2, \ldots, n$$

โดย R_j^2 คือ ค่าสัมประสิทธิ์การตัดสินใจของ x_j กับตัวแปรอิสระอื่น ๆ ที่เหลือ และถ้าค่า VIF มีค่ามากกว่า 10 แสดงว่าเกิดปัญหา Multicollinearity

จ. การทดสอบความไม่เหมาะสม (lack of fit test) เป็นการทดสอบทางสถิติวิธีหนึ่ง สำหรับทดสอบรูปแบบการถดถอยว่ารูปแบบที่ใช้มีความเหมาะสมกับข้อมูลหรือไม่ ซึ่งมีรายละเอียดดังนี้

H_0 : รูปแบบการถดถอยเหมาะสมกับข้อมูล

H_1 : รูปแบบการถดถอยไม่เหมาะสมกับข้อมูล

การทดสอบความไม่เหมาะสมของตัวแบบ

Sov	SS	df	F
การถดถอย (R)	SSR	1	
ความคลาดเคลื่อน (E)	SSE	n-2	
ความไม่เหมาะสม	SSLF	m-2	$F = \frac{MSLF}{MSPE}$
ความคลาดเคลื่อนจริง	SSPE	n-m	
ทั้งหมด	SST	n – 1	

สถิติสำหรับทดสอบ คือ $F = \frac{SSLF/(m-2)}{SSPE/(n-m)}$

โดยที่ SSPE คือ $\sum_{j=1}^{m}\sum_{i=1}^{n}(Y_{ij}-\bar{Y}_j)^2$ เป็นผลบวกกำลังสองเนื่องจาก pure error ซึ่งเป็นความคลาดเคลื่อนที่เกิดขึ้นจริงในการทดลอง

SSLF คือ $\sum_{j=1}^{m} n_j(\bar{Y}_j-\hat{Y}_j)^2$ เป็นผลบวกกำลังสอง เนื่องจาก lack of fit ซึ่งเป็นความคลาดเคลื่อนที่เกิดจากการสร้างสมการถดถอยที่มีรูปแบบไม่เหมาะสม

ถ้า $F > F_{\alpha,m-2,n-m}$ จะปฏิเสธ H_0 แสดงว่า รูปแบบการถดถอยเชิงเส้นไม่เหมาะสม จึงต้องหารูปแบบใหม่

การแก้ไขรูปแบบการถดถอยให้เหมาะสม ถ้าตรวจพบว่ารูปแบบการถดถอยไม่มีความเหมาะสมกับข้อมูลที่มีอยู่จะแก้ไขได้ 2 วิธี คือ การยกเลิกรูปแบบการถดถอยนั้นแล้วหารูปแบบการถดถอยใหม่ แต่อาจจะได้รูปแบบการถดถอยที่ยุ่งยาก และมีความลำบากในการประมาณค่าพารามิเตอร์ หรืออีกทางเลือกหนึ่งคือ การแปลงข้อมูล (Transformation) ให้มีความเหมาะสมกับข้อมูลที่มีอยู่โดยมีหลักการพิจารณาดังนี้

1) กรณีที่รูปแบบการถดถอยไม่เป็นเส้นตรง อาจจะใช้รูปแบบการถดถอยโพลิโนเมียลลำดับต่าง ๆ

2) กรณีความคลาดเคลื่อนมีค่าความแปรปรวนไม่คงที่อาจทำการแปลงตัวแปร Y เป็นตัวแปร Y' เพื่อให้ค่าความคลาดเคลื่อนจากรูปแบบการถดถอยที่ได้จากการแปลงมีค่าความแปรปรวนคงที่

3) กรณีความคลาดเคลื่อนไม่มีการแจกแจงแบบปกติ อาจแก้ไขโดยการแปลงจาก Y เป็น $Y' = \log Y$ หรือ $Y' = \sqrt{Y}$ เป็นต้น การแปลงค่าของตัวแปรตามนี้นอกจากจะเป็นการแก้ปัญหากรณีค่าความคลาดเคลื่อนมีการแจกแจงแบบไม่เป็นปกติแล้วยังเป็นการแก้ปัญหากรณีความแปรปรวนของค่าความคลาดเคลื่อนไม่คงที่ด้วย

2.1.5 การเลือกตัวแปรอิสระ

ในการวิเคราะห์การถดถอยนั้นจะเห็นได้ว่า ถ้าเป็นการวิเคราะห์การถดถอยเชิงเส้นอย่างง่ายจะประกอบด้วยตัวแปรอิสระเพียงตัวเดียว แต่ถ้าเป็นการวิเคราะห์การถดถอยเชิงเส้นพหุคูณจะประกอบด้วยตัวแปรอิสระหลายตัว โดยทั่วไปในการวิเคราะห์การถดถอยนั้นจะพบว่าถ้าเพิ่มตัวแปรอิสระเข้าไปในสมการถดถอยมากขึ้นอิทธิพลของตัวแปรอิสระร่วมกันที่มีต่อตัวแปรตามก็จะมากขึ้น ทั้งนี้อาจพิจารณาได้จากค่าสัมประสิทธิ์การตัดสินใจจะเพิ่มมากขึ้นหรือค่าเฉลี่ยของกำลังสองของความคลาดเคลื่อน (MSE) อาจจะลดลง แต่การเพิ่มตัวแปรอิสระเข้าไปในรูปแบบการถดถอยมาก ๆ จะทำให้การคำนวณค่าต่าง ๆ มีความลำบากมากขึ้น ดังนั้นจึงควรมีการพิจารณาเลือกเฉพาะตัวแปรอิสระที่มีความสำคัญและมีความจำเป็นเท่านั้นเข้าไปในรูปแบบการถดถอยที่จะทำการวิเคราะห์ (สุรินทร์, 2548)

การเลือกตัวแปรอิสระที่ดีที่สุดนั้นควรจะต้องมีจำนวนที่พอดีที่ทำให้การวิเคราะห์สามารถทำได้อย่างสะดวกและผลการวิเคราะห์จะสามารถนำไปใช้ประโยชน์ได้อย่างดี การเลือกตัวแปรอิสระที่ดีจะทำให้

1. ค่าสัมประสิทธิ์การตัดสินใจ (R^2) มีค่าสูง
2. ค่าเฉลี่ยของกำลังสองของความคลาดเคลื่อน (MSE) มีค่าต่ำ
3. ค่าสัมประสิทธิ์การถดถอยต้องมีค่าแตกต่างจากค่าศูนย์อย่างมีนัยสำคัญ

วิธีการเลือกตัวแปรอิสระ อาจทำหลายวิธี ได้แก่ Forward Selection, Backward Elimination หรือ Stepwise Regression ซึ่งแต่ละวิธีมีรายละเอียดดังนี้

วิธีการเลือกตัวแปรอิสระแบบ Forward Selection เป็นวิธีการพิจารณาตัวแปรอิสระเข้าในรูปแบบทีละตัวด้วยหลักการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย โดยใช้ตัวแปรอิสระแต่ละตัวที่มีอยู่ทั้งหมด k ตัว และทดสอบสัมประสิทธิ์การถดถอยของตัวแปรอิสระ $\beta_i = 0$ หรือไม่ โดยสถิติที่ใช้ทดสอบคือ $t = \frac{b_i}{s_b}$ หรือ F – test แล้วเลือก t_i หรือ F ที่มีค่าสูงที่สุด ซึ่งหมายถึงว่าตัวแปรอิสระนั้นจะมีความสำคัญมากที่สุดเมื่อเทียบกับตัวแปรอิสระอื่น ๆ ถ้าหากการทดสอบ

สัมประสิทธิ์การถดถอยพบว่ามีนัยสำคัญ ตัวแปรอิสระนั้นจะถูกใส่เข้าไปในรูปแบบการถดถอยเป็นตัวแรก

ขั้นต่อไปจะพิจารณาตัวแปรอิสระที่ยังเหลืออยู่อีก k-1 ตัว ว่าควรจะใส่ตัวแปรอิสระตัวใดเป็นตัวต่อไปเข้าไปในรูปแบบการถดถอย คือจะพิจารณารูปแบบการถดถอยที่มีตัวแปรอิสระที่ยังเหลืออยู่เพิ่มเข้าไปทีละตัวในรูปแบบการถดถอยที่มีตัวแปรอิสระตัวแรกอยู่แล้ว และคำนวณค่าสถิติที่ใช้ทดสอบ คือ $t = \frac{b_i}{s_b}$ หรือ partial F – test แล้วเลือก t_i ที่มีค่าสูงที่สุด (หรือเลือก partial F – test ที่มีค่าสูงที่สุด) หมายถึงว่าตัวแปรอิสระนั้นจะมีความสำคัญมากที่สุดตัวต่อไปเมื่อเทียบกับตัวแปรอิสระอื่น ๆ ที่ยังเหลืออยู่ ถ้าหากการทดสอบสัมประสิทธิ์การถดถอยพบว่ามีนัยสำคัญ ตัวแปรอิสระนั้นจะถูกใส่เข้าไปในรูปแบบการถดถอยเป็นตัวที่สอง

ขั้นต่อไปจะพิจารณาตัวแปรอิสระที่ยังเหลืออยู่ทำนองเดียวกันนี้ไปเรื่อย ๆ จนกว่าจะพบว่า t_i ที่มีค่าสูงสุด นั้นไม่มีนัยสำคัญก็จะหยุดไม่ใส่ตัวแปรอิสระเข้าไปในรูปแบบการถดถอยอีก และใช้รูปแบบการถดถอยนั้นเป็นรูปแบบการถดถอยที่เลือกได้

วิธีการเลือกตัวแปรอิสระแบบ Backward Elimination นี้ จะเริ่มจากกาพิจารณารูปแบบการถดถอยที่มีตัวแปรอิสระทุกตัว (k ตัว) ก่อน แล้วพิจารณาว่าตัวแปรอิสระใดมีความสำคัญน้อยที่สุดจะถูกตัดออกไป โดยพิจารณาจากการทดสอบสมมติฐานที่เกี่ยวกับสัมประสิทธิ์การถดถอยของตัวแปรอิสระแต่ละตัวดังนี้คือ $H_0 : \beta_i = 0$ และ $H_1 : \beta_i \neq 0$ เมื่อ $i = 1, 2, \ldots, k$ ตัวสถิติที่ใช้ทดสอบคือ $t = \frac{b_i}{s_b}$ (หรืออาจใช้ partial F - test)

การทดสอบสมมติฐานข้างต้น จะเลือก t_i ที่มีค่าน้อยที่สุด (หรือเลือก partial F – test ที่มีค่าน้อยที่สุด) ซึ่งหมายถึงว่าตัวแปรอิสระนั้นจะมีความสำคัญน้อยที่สุดเมื่อเทียบกับตัวแปรอิสระอื่น ๆ ที่อยู่ในรูปแบบ ถ้าหากผลการทดสอบพบว่าสัมประสิทธิ์การถดถอยไม่มีนัยสำคัญตัวแปรอิสระนั้นจะถูกตัดออกไปจากรูปแบบการถดถอย ขั้นต่อไปก็จะพิจารณารูปแบบการถดถอยที่มีตัวแปรอิสระเพียง k-1 ตัว และจะทดสอบสมมติฐานสำหรับสัมประสิทธิ์การถดถอยของตัวแปรอิสระที่มีความสำคัญน้อยที่สุดต่อไป ถ้าหากการทดสอบพบว่าสัมประสิทธิ์การถดถอยไม่มีนัยสำคัญ ตัวแปรอิสระนั้นจะถูกตัดออกไปจากรูปแบบการถดถอยอีก ขั้นต่อไปก็จะพิจารณารูปแบบการถดถอยที่มีตัวแปรอิสระเพียง k-2 ตัว เป็นเช่นนี้ไปเรื่อย ๆ จนกว่าจะพบว่าการทดสอบ

สัมประสิทธิ์การถดถอยมีนัยสำคัญจึงจะหยุด ซึ่งแสดงว่าตัวแปรอิสระที่มีค่า t_i หรือค่า partial F – test มีค่าน้อยที่สุดนั้นมีความสำคัญที่ควรจะยังคงไว้ในรูปแบบการถดถอยและใช้รูปแบบการถดถอยนั้นเป็นรูปแบบการถดถอยที่เลือกได้

วิธีการเลือกตัวแปรอิสระแบบ Stepwise Regression เป็นวิธีการเลือกตัวแปรอิสระผสมผสานระหว่างวิธีการเลือกตัวแปรอิสระแบบ Forward Selection และวิธี Backward Elimination เข้าด้วยกัน โดยวิธี Stepwise Regression นั้นจะคัดเลือกตัวแปรอิสระเข้าไปในรูปแบบการถดถอยทีละตัวด้วยหลักการเดียวกับวิธี Forward Selection ด้วยการเริ่มจากการพิจารณาการถดถอยเชิงเส้นอย่างง่ายโดยใช้ตัวแปรอิสระทุกตัวที่มีอยู่ทั้งหมด k ตัว และทดสอบสัมประสิทธิ์การถดถอยว่า $\beta_i = 0$ หรือไม่ โดยการคำนวณค่าสถิติที่ใช้ทดสอบคือ $t = \frac{b_i}{s_b}$ หรือ F แล้วเลือก t_i หรือ F ที่มีค่าสูงที่สุด ถ้าหากพบว่าการทดสอบสัมประสิทธิ์การถดถอยมีนัยสำคัญ ตัวแปรอิสระนั้นจะถูกใส่เข้าไปในรูปแบบการถดถอยเป็นตัวแรก

ขั้นต่อไปก็จะพิจารณาตัวแปรอิสระซึ่งยังเหลืออยู่อีก k-1 ตัว ว่าควรใส่ตัวแปรอิสระตัวใดเป็นตัวต่อไปเข้าไปในรูปแบบการถดถอย ในขั้นนี้จะพิจารณารูปแบบการถดถอยที่เสมือนมีตัวแปรอิสระที่ยังเหลืออยู่เพิ่มเข้าไปทีละตัวในรูปแบบการถดถอยที่มีตัวแปรอิสระตัวแรกอยู่แล้ว และทดสอบสัมประสิทธิ์การถดถอยคือ $t = \frac{b_i}{s_b}$ หรือ partial F – test แล้วเลือก t_i หรือ partial F – test ที่มีค่าสูงที่สุด ถ้าหากการทดสอบสัมประสิทธิ์การถดถอยพบว่ามีนัยสำคัญ ตัวแปรอิสระนั้นจะถูกใส่เข้าไปในรูปแบบการถดถอยเป็นตัวที่สอง แต่ในขั้นนี้วิธี Stepwise Regression จะตรวจสอบตัวแปรอิสระที่มีอยู่ในรูปแบบเหมือนวิธี Backward Elimination กล่าวคือจะทำการตรวจสอบตัวแปรอิสระที่เข้าไปก่อนหน้านี้ทุกตัวหลังจากที่ได้เพิ่มตัวแปรอิสระเข้าไปใหม่ ทั้งนี้เพราะว่าการที่เพิ่มตัวแปรอิสระเข้าไปใหม่จะทำให้ค่าของ t_i หรือ partial F – test เปลี่ยนแปลงไป การตรวจสอบนี้จะทำการคำนวณค่าสถิติในการทดสอบสัมประสิทธิ์การถดถอยคือ t_i หรือ partial F ของตัวแปรอิสระทุกตัวที่เข้าไปก่อนหน้านี้ หากการทดสอบสัมประสิทธิ์การถดถอยพบว่าไม่มีนัยสำคัญตัวแปรอิสระนั้นจะถูกตัดออกไปจากรูปแบบ แต่ถ้าหากการทดสอบพบว่ามีนัยสำคัญตัวแปรอิสระนั้นก็จะยังคงอยู่ในรูปแบบต่อไป

การเลือกตัวแปรอิสระเข้าหรือตัดตัวแปรอิสระออกจากรูปแบบจะดำเนินไปเช่นนี้เรื่อย ๆ จนกว่าจะไม่สามารถเลือกตัวแปรอิสระใดเข้าหรือตัดตัวแปรอิสระใดออกจากรูปแบบการถดถอยได้ก็จะหยุดและใช้รูปแบบการถดถอยสุดท้ายนั้นเป็นรูปแบบการถดถอยที่เลือกได้ จะเห็นได้ว่าวิธีการเลือกตัวแปรอิสระแบบ Stepwise Regression นั้นตัวแปรอิสระที่เข้าไปในรูปแบบการถดถอยในขั้นก่อนอาจจะถูกตัดออกในขั้นต่อ ๆ มาได้ และในทำนองเดียวกันตัวแปรอิสระที่ถูกตัดออกไปในขั้นก่อนก็อาจจะถูกเลือกกลับเข้ามาใหม่อีกในขั้นต่อมาได้

2.2 การวิเคราะห์ถดถอยหลายตัวแปร (Multivariate Regression Analysis)

Johnson, R.A. and Wichem, D.W.(2002) ในสถานการณ์ความเป็นจริงการวิเคราะห์การถดถอยอาจศึกษาข้อมูลที่มีตัวแปรตามมากกว่าหนึ่งตัวกับตัวแปรอิสระหลาย ๆ ตัวและผู้ศึกษาต้องการศึกษารูปแบบความสัมพันธ์ระหว่างตัวแปรตามเหล่านี้กับตัวแปรอิสระพร้อม ๆ กันจะเรียกว่า การวิเคราะห์การถดถอยหลายตัวแปรเขียนรูปแบบการถดถอย ดังนี้

$$\underset{(nxp)}{\mathbf{Y}} = \underset{(nx(k+1))}{\mathbf{X}}\ \underset{((k+1)xp)}{\boldsymbol{\beta}} + \underset{(nxp)}{\boldsymbol{\varepsilon}}$$

โดยที่ Y คือ เมตริกซ์ของตัวแปรตามที่มีขนาด nxp

X คือ เมตริกซ์ของตัวแปรอิสระที่มีขนาด (nx(k+1))

β คือ ค่าสัมประสิทธิ์การถดถอยที่มีขนาด ((k+1)xp)

ε คือ เมตริกซ์ความคลาดเคลื่อนที่มีขนาด nxp

p คือ จำนวนตัวแปรตามที่ศึกษา

k คือ จำนวนตัวแปรอิสระที่ศึกษา

n คือ จำนวนข้อมูลตัวอย่าง

มีข้อสมมติเบื้องต้น คือ $E(\varepsilon_i) = 0$ และ $Cov(\varepsilon_i, \varepsilon_j) = \sigma_{ij}^2 I$; i, j = 1, 2, …, p

เมื่อ σ_{ij}^2 คือ ความแปรปรวน และ I คือ เมตริกซ์เอกลักษณ์

และสามารถเขียนรูปแบบการถดถอยในรูปของเมตริกซ์ได้ดังนี้

$$\underset{(nxp)}{\mathbf{Y}} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{bmatrix}, \quad \underset{(nx(k+1))}{\mathbf{X}} = \begin{bmatrix} x_{10} & x_{11} & \cdots & x_{1k} \\ x_{20} & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n0} & x_{n1} & \cdots & x_{nk} \end{bmatrix}$$

$$\underset{((k+1)xp)}{\beta} = \begin{bmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{kp} \end{bmatrix}, \quad \underset{(nxp)}{\varepsilon} = \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{bmatrix}$$

2.2.1 การประมาณค่าสัมประสิทธิ์การถดถอย

การสร้างสมการถดถอยเริ่มจากการหาตัวประมาณ $\hat{\beta}$ ของพารามิเตอร์ β ด้วยวิธีกำลัง-สองน้อยที่สุด (Least Square Method) ซึ่งเป็นการทำให้ผลรวมกำลังสองของความคลาดเคลื่อน SSE น้อยที่สุด และจะได้

$$\hat{\boldsymbol{\beta}} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{Y}$$

โดย

$$\hat{\boldsymbol{\beta}}_{\mathbf{i}} = (\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'\mathbf{Y}_{\mathbf{i}} \qquad \text{โดยที่} \quad i = 1, 2, \ldots, k$$

และอาจแสดงตัวอย่างการหาค่าสัมประสิทธิ์การถดถอยดังนี้

ตัวอย่าง การคำนวณค่าสัมประสิทธิ์การถดถอยสำหรับข้อมูลต่อไปนี้

Y_1	6	6	0	2	4
Y_2	9	3	9	5	1
X_1	7	3	6	6	7
X_2	9	8	6	0	2
X_3	8	7	7	1	3

ในที่นี้ n = 5 k = 3 และ p = 2

จากข้อมูลสามารถเขียนในรูปเมตริกซ์ ได้ดังนี้

$$Y=\begin{bmatrix}6 & 9\\ 6 & 3\\ 0 & 9\\ 2 & 5\\ 4 & 1\end{bmatrix}, \qquad X=\begin{bmatrix}1 & 7 & 9 & 8\\ 1 & 3 & 8 & 7\\ 1 & 6 & 6 & 7\\ 1 & 6 & 0 & 1\\ 1 & 7 & 2 & 3\end{bmatrix}$$

$$X'=\begin{bmatrix}1 & 1 & 1 & 1 & 1\\ 7 & 3 & 6 & 6 & 7\\ 9 & 8 & 6 & 0 & 2\\ 8 & 7 & 7 & 1 & 3\end{bmatrix}, \qquad X'X=\begin{bmatrix}5 & 29 & 25 & 32\\ 29 & 179 & 137 & 146\\ 25 & 137 & 185 & 176\\ 32 & 146 & 176 & 172\end{bmatrix}$$

$$(X'X)^{-1}=\begin{bmatrix}-0.0379 & -0.0283 & -0.1312 & 0.1652\\ -0.0283 & 0.0210 & 0.0084 & -0.0212\\ -0.1312 & 0.0084 & 0.0181 & -0.0013\\ 0.1653 & -0.0212 & -0.0013 & -0.0055\end{bmatrix}$$

$$X'Y=\begin{bmatrix}18 & 27\\ 100 & 163\\ 110 & 161\\ 104 & 164\end{bmatrix},$$

ดังนั้น $$\hat{\beta}=\begin{bmatrix}-0.763 & 0.333\\ 0.316 & 0.544\\ 0.341 & 0.537\\ 0.122 & -0.133\end{bmatrix}_{4x2}$$

โดยที่ $\hat{\beta}_1 = \begin{bmatrix} -0.763 \\ 0.316 \\ 0.341 \\ 0.122 \end{bmatrix}_{4x1}$, และ $\hat{\beta}_2 = \begin{bmatrix} 0.333 \\ 0.544 \\ 0.537 \\ -0.133 \end{bmatrix}_{4x1}$

ดังนั้นสมการถดถอย คือ $\hat{Y}_1 = -0.763+0.316X_1+0.341X_2+0.122X_3$

$$\hat{Y}_2 = 0.333+0.544X_1+0.537X_2-0.133X_3$$

2.2.2 การทดสอบค่าสัมประสิทธิ์การถดถอย

การวิเคราะห์การถดถอยนั้นจะทำการทดสอบสมมติฐานเกี่ยวกับ β โดยมีสมมติฐานการทดสอบดังนี้

$H_0: \beta_1 = \beta_2 = ... = \beta_k$

H_1 : มี β_i อย่างน้อย 1 ค่าที่มีค่าไม่เท่ากับ 0 ; i = 1, 2, …, k

สถิติทดสอบ คือ Wilks' lambda (λ) $= \dfrac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|}$

โดยที่ $\mathbf{E} = \mathbf{Y'Y} - \mathbf{Y'X}\hat{\boldsymbol{\beta}}$ = ค่าผันแปรภายในกลุ่ม

$\mathbf{H} = \mathbf{Y'X}\hat{\boldsymbol{\beta}} - \dfrac{1}{\mathbf{n}}\mathbf{Y'11'Y}$

$\mathbf{Y}$ = เมตริกซ์ของตัวแปรตามที่มีขนาด nxp

$\mathbf{X}$ = เมตริกซ์ของตัวแปรอิสระที่มีขนาด (nx(k+1))

$\hat{\boldsymbol{\beta}}$ = เมตริกซ์ค่าสัมประสิทธิ์การถดถอยที่มีขนาด ((k+1)xp)

สามารถแปลงสถิติทดสอบ Wilks' lambda (λ) เป็นสถิติไคสแควส์ได้ดังนี้

$$\chi^2 = -(n-1-\frac{p+k}{2})\ln\lambda$$
$$= -(n-1-\frac{p+k}{2})\ln(\frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|})$$

χ^2 จะมีการแจกแจงโดยประมาณแบบไคสแควส์ที่องศาอิสระ p(k-1)

ซึ่งจะปฏิเสธสมมติฐาน H_0 ถ้า $\chi^2 > \chi^2_{\alpha,p(k-1)}$ และ ถ้าในการทดสอบนี้ ปฏิเสธสมมติฐาน H_0 จะสรุปผลการทดสอบได้ว่า ตัวแปรตามทุกตัว (Y) มีความสัมพันธ์กับตัวแปรอิสระ (X) ทั้ง k ตัว ในรูปแบบเชิงเส้นตรง หมายความว่า สัมประสิทธิ์การถดถอยบางส่วนอย่างน้อย 1 ตัวมีค่าไม่เท่ากับศูนย์ ดังนั้นจึงควรทดสอบสัมประสิทธิ์การถดถอยบางส่วนสำหรับตัวแปรอิสระแต่ละตัวต่อไปดังนี้

สมมติฐานในการทดสอบ $H_0: \beta_i = 0$

$H_1: \beta_i \neq 0 \quad ; i = 1, 2, \ldots, k$

สถิติทดสอบ คือ Wilks' lambda (λ) $= \frac{|\mathbf{E}|}{|\mathbf{E}+\mathbf{H}|}$

2.2.3 การตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยหลายตัวแปร

สำหรับข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยหลายตัวแปรนั้นเป็นข้อสมมติเกี่ยวกับค่าคลาดเคลื่อน คือ $\varepsilon_i \sim MID(0,\Sigma)$ ซึ่งในการตรวจสอบข้อสมมติดังกล่าวทำได้ดังนี้

ก. การตรวจสอบว่าค่าคลาดเคลื่อนจะต้องมีการแจกแจงแบบปกติหลายตัวแปรที่มีค่าเฉลี่ยเป็นศูนย์ ในที่นี้จะใช้วิธีการของ Mardia ซึ่งพิจารณาจากค่าความเบ้และค่าความโด่ง ดังนี้

สมมติฐานในการทดสอบ H_0 : ค่าคลาดเคลื่อนมีการแจกแจงปกติหลายตัวแปร

H_1 : ค่าคลาดเคลื่อนไม่มีการแจกแจงปกติหลายตัวแปร

สถิติทดสอบ ค่าความเบ้ $b_{1,p} = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}^3$

ค่าความโด่ง $b_{2,p} = \frac{1}{n}\sum_{i=1}^{n} g_{ii}^2$

โดยที่ $g_{ij} = (y_i - \bar{y})' \sum (y_j - \bar{y})$

$g_{ij} = (y_i - \bar{y})' \sum (y_{ii} - \bar{y})$

$\sum$ = เมตริกซ์ค่าความแปรปรวนร่วมขนาด nxp

p = จำนวนตัวแปรตามที่ศึกษา

n = จำนวนข้อมูลที่ศึกษา

y_i = ข้อมูลของตัวอย่างที่ i

$\bar{y}$ = ค่าเฉลี่ยของตัวอย่าง

$b_{1,p}$ = ค่าสัมประสิทธิ์ความเบ้

$b_{2,p}$ = ค่าสัมประสิทธิ์ความโด่ง

ข. การตรวจสอบว่าความแปรปรวนของค่าคลาดเคลื่อนมีค่าเท่ากันและเท่ากับ $\sum$ การตรวจสอบจะพิจารณาจากกราฟ โดยดูความสัมพันธ์ระหว่าง ค่าความคลาดเคลื่อนกับค่าพยากรณ์ ทำโดยการพล็อต $\hat{Y}_j$ กับ e_j ถ้าภาพการพล็อตของคู่ลำดับ $(\hat{Y}_j, e_j)$ กระจายในลักษณะเป็นรูปกรวยปลายเปิด แสดงว่าความคลาดเคลื่อนสุ่มมีความแปรปรวนไม่คงที่ เนื่องจากค่า e_j มีการแปรผันตามค่า $\hat{Y}_j$ และถ้าภาพการพล็อตของคู่ลำดับ $(\hat{Y}_j, e_j)$ กระจายในลักษณะรอบๆ แกนอ้างอิง แสดงว่าความคลาดเคลื่อนสุ่มมีความแปรปรวนคงที่

3. การวิเคราะห์สหสัมพันธ์แคนนอลนิคอล (Canonical Correlation Analysis)

การวิเคราะห์สหสัมพันธ์แคนนอลนิคอลเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปร 2 เซต โดยตัวแปรในแต่ละเซตมีมากกว่า 1 ตัว จำนวนตัวแปรในแต่ละเซตจะเท่ากันหรือไม่เท่ากันก็ได้ สถิติวัดความสัมพันธ์ของตัวแปร 2 เซต คือ ค่าสหสัมพันธ์แคนนอลนิคอล ซึ่งเป็นค่าสหสัมพันธ์ซึ่งมีคุณสมบัติใช้วัดได้ทั้งขนาดและทิศทางความสัมพันธ์ และมีค่าระหว่าง –1 ถึง +1

3.1 การคำนวณค่าสหสัมพันธ์แคนนอลนิคอล

หลังจากที่รวบรวมข้อมูลที่เป็นค่าของตัวแปร 2 เซต ซึ่งในที่นี้คือ ตัวแปรอิสระ (X) ซึ่งมี k ตัว และตัวแปรตาม (Y) ซึ่งมี p ตัว นำมาจัดอยู่ในรูปเมตริกซ์ดังนี้

ตารางที่ 1 เมตริกซ์ข้อมูลเบื้องต้นสำหรับการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล

สมาชิกของกลุ่มตัวอย่าง	ตัวแปรอิสระ	ตัวแปรตาม
1	$X_{11}\ X_{12}\ldots X_{1k}$	$Y_{11}\ Y_{12}\ldots Y_{1p}$
2	$X_{21}\ X_{22}\ldots X_{2k}$	$Y_{21}\ Y_{22}\ldots Y_{2p}$
.	.	.
.	.	.
.	.	.
n	$X_{N1}\ X_{N2}\ldots X_{Nk}$	$Y_{N1}\ Y_{N2}\ldots Y_{Np}$

จากเมตริกซ์ข้อมูลเบื้องต้นในตารางที่ 1 คำนวณหาค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายโดยจับคู่ระหว่างตัวแปรทุกตัวเป็นคู่ๆ แล้วนำค่าสัมประสิทธิ์สหสัมพันธ์ทั้งหมดมาจัดในรูปของเมตริกซ์ และแบ่งเมตริกซ์นั้นเป็น 4 ส่วน โดย

กำหนดส่วนย่อยของเมตริกซ์ X กับ Y ให้อยู่ในรูปของเมตริกซ์ ดังนี้

$$\mathbf{R} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}$$

เมื่อ $\mathbf{R}$ คือ เมตริกซ์สหสัมพันธ์ทั้งหมด

$\mathbf{R}_{xx}$ คือ เมตริกซ์สหสัมพันธ์ระหว่างตัวแปรอิสระ

$\mathbf{R}_{yy}$ คือ เมตริกซ์สหสัมพันธ์ระหว่างตัวแปรตาม

$\mathbf{R}_{xy}$ และ $\mathbf{R}_{yx}$ คือ เมตริกซ์สหสัมพันธ์ระหว่างตัวแปรอิสระ X กับตัวแปรตาม Y

จากนั้นหาค่าเมตริกซ์ของ $\mathbf{R}_{yy}^{-1}, \mathbf{R}_{yx}, \mathbf{R}_{xx}^{-1}, \mathbf{R}_{xy}$ แล้วนำไปสร้างสมการดีเทอร์มิแนนท์ ดังนี้

$$| \mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy} - \lambda \mathbf{I} | = 0$$

เมื่อ $\mathbf{R}_{yy}^{-1}$ แทน อินเวอร์เมตริกซ์ของ $\mathbf{R}_{yy}$

$\mathbf{R}_{xx}^{-1}$ แทน อินเวอร์เมตริกซ์ของ $\mathbf{R}_{xx}$

$\mathbf{I}$ แทน เมตริกซ์เอกลักษณ์

λ แทน ค่าไอเกน (eigenvalue)

จากสมการดีเทอร์มิแนนท์จะได้สมการกำลังสอง คือ

$$a^2 - b + c = 0$$

และคำนวณค่า $a^2 - b + c$ จาก $\lambda = \frac{\pm b\sqrt{b^2 - 4ac}}{2a}$

จากนั้นคำนวณค่าสหสัมพันธ์แคนนอลนิคอล โดยการถอดรากที่สอง λ ดังนี้

$$\mathbf{R} = \sqrt{\lambda}$$

3.2 การทดสอบสมมติฐานของค่าสหสัมพันธ์แคนนอลนิคอล

ในการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล ควรทดสอบสมมติฐาน 2 ส่วนคือ ส่วนที่ 1 เป็นการทดสอบความสัมพันธ์ทั้งหมด ว่ามีความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรสองชุดเหล่านั้นหรือไม่ ถ้าพบว่ามีนัยสำคัญก็จะไปทดสอบส่วนที่ 2 คือ การทดสอบว่ามีน้ำหนักความสำคัญแคนนอลนิคอลค่าใดบ้างที่มีนัยสำคัญ คือมีกลุ่มตัวแปรอิสระและกลุ่มตัวแปรตามกี่คู่ที่มีนัยสำคัญ ซึ่งการทดสอบทั้งสองส่วนนี้นิยมใช้วิธีการทดสอบโดยใช้การแจกแจงแบบไคสแควส์ตามวิธีของ Bartlett โดยมีสมมติฐานในการทดสอบดังนี้ (กัลยา, 2548)

สมมติฐานในการทดสอบ $H_0 : R_1 = R_2 = \ldots = R_i = \ldots = R_k = 0 \qquad i = 1, 2, \ldots, k$

H_1 : มี R_i อย่างน้อย 1 ค่าที่ไม่เท่ากับ 0

สถิติทดสอบ $V = -[N-1-0.5(p+k+1)]\ln\lambda$; $df = pk$

เมื่อ **N** แทน จำนวนข้อมูลในกลุ่มตัวอย่างทั้งหมด

k แทน จำนวนตัวแปรอิสระ

p แทน จำนวนตัวแปรตาม

λ แทน ผลคูณของผลต่างระหว่าง 1 กับ eigenvalue แต่ละตัว จนครบ m ตัว หรือค่า Wilks' Lambda ดังนั้น $\lambda = (1-\lambda_1)(1-\lambda_2)\ldots(1-\lambda_p)$

เมื่อคำนวณค่า V แล้วนำไปเปรียบเทียบกับค่าวิกฤติของการแจกแจงไคสแควส์ ที่องศาอิสระเท่ากับ pk และถ้าค่า V มากกว่าหรือเท่ากับค่าวิกฤต จะสรุปว่ามีนัยสำคัญทางสถิติ คือมีความสัมพันธ์เชิงเส้นตรงระหว่างตัวแปรทั้งสองกลุ่มนั้น แต่ถ้าค่าสถิติทดสอบไม่มีนัยสำคัญนั้น คือ ค่า V น้อยกว่าค่าวิกฤต แสดงว่า ไม่มีความสัมพันธ์เชิงเส้นตรงอย่างมีนัยสำคัญทางสถิติระหว่างตัวแปรทั้งสองกลุ่ม สามารถแสดงตัวอย่างการคำนวณค่าสหสัมพันธ์แคนนอลนิคอล และการทดสอบสมมติฐานได้ดังตัวอย่างต่อไปนี้

ตัวอย่าง ถ้ามีตัวแปร 2 เซต คือเซตของตัวแปรอิสระ X กับเซตของตัวแปรตาม Y ในที่นี้ $p = k = 2$

คำนวณค่าสัมประสิทธิ์สหสัมพันธ์ของตัวแปรทั้ง 4 ตัว ได้ดังนี้

	Y_1	Y_2	X_1	X_2
Y_1	1.0			
Y_2	-0.307	1.0		
X_1	0.221	0.316	1.0	
X_2	0.445	0.168	-0.176	1.0

ในที่นี้เมตริกซ์สหสัมพันธ์ (**R**) แบ่งเป็น 4 เมตริกซ์ย่อยคือ $\mathbf{R}_{yy}$, $\mathbf{R}_{xx}$,$\mathbf{R}_{yx}$ และ$\mathbf{R}_{xy}$ ดังนี้

$$\mathbf{R} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}$$

เมื่อ $\mathbf{R}_{yy} = \begin{bmatrix} 1.0 & -0.307 \\ -0.307 & 1.0 \end{bmatrix}$, $\mathbf{R}_{xx} = \begin{bmatrix} 1.0 & -0.176 \\ -0.176 & 1.0 \end{bmatrix}$

$\mathbf{R}_{yx} = \begin{bmatrix} 0.221 & 0.445 \\ 0.316 & 0.168 \end{bmatrix}$, $\mathbf{R}_{xy} = \begin{bmatrix} 0.221 & 0.316 \\ 0.445 & 0.168 \end{bmatrix}$

คำนวณค่าไอเก็นของเมตริกซ์ $\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy}$ และได้ว่า

$$\mathbf{R}_{yy}^{-1}\mathbf{R}_{yx}\mathbf{R}_{xx}^{-1}\mathbf{R}_{xy} = \begin{bmatrix} 0.374 & 0.417 \\ 0.555 & 0.313 \end{bmatrix}$$

ค่าไอเก็นของเมตริกซ์ข้างต้นได้ 0.59 และ 0.02 ในตัวอย่างนี้สามารถสร้างตัวแปรแคนนอลนิคอลได้ 2 คู่ เนื่องจาก $\mathbf{R} = \sqrt{\lambda}$

ดังนั้น ค่าสหสัมพันธ์แคนนอลนิคอลคู่ที่ 1 $= \sqrt{0.59} = 0.77$

ตัวแปรแคนนอลนิคอลคู่ที่ 1 มีค่าไอเก็น = 0.59 หมายถึงค่าแปรปรวนหรือความผันแปรเดิมสามารถอธิบายได้ด้วยตัวแปรแคนนอลนิคอลคู่ที่1 ร้อยละ 59

ค่าสหสัมพันธ์แคนนอลนิคอลคู่ที่ 2 $= \sqrt{0.02} = 0.14$

ตัวแปรแคนนอลนิคอลคู่ที่ 2 มีค่าไอเก็น = 0.02 หมายถึงค่าแปรปรวนหรือความผันแปรเดิมสามารถอธิบายได้ด้วยตัวแปรแคนนอลนิคอลคู่ที่ 2 ร้อยละ 2

จากนั้นทำการทดสอบสมมติฐานค่าสหสัมพันธ์แคนนอลนิคอล

จากตัวอย่างสร้างตัวแปรแคนนอลนิคอลได้ 2 คู่ และคำนวณค่าสหสัมพันธ์แคนนอลนิคอลได้ 2 ค่า คือ 0.77 และ 0.14 ตามลำดับ สมมติให้ $n = 100$ ดังนั้น

สมมติฐานในการทดสอบ

$H_0 : R_1 = R_2 = 0$

$H_1 : R_i$ อย่างน้อย 1 ค่าที่ไม่เท่ากับ 0

สถิติทดสอบ $V = -[n-1-0.5(p+k+1)]\ln\lambda$; $df = pk$

$= -[(100-1)-0.5(2+2+1)]\ln(0.399)$

$= 88.684$

$$\lambda = \prod_{i=1}^{2}(1-\lambda_i^2) = (1-(0.77)^2)(1-(0.14)^2) = 0.399$$

จากตารางการแจกแจงไคสแควส์ได้ค่า $\chi^2_{0.05,4} = 9.49$ จึงปฏิเสธสมมติฐาน H_0 นั่นคือ ค่าสหสัมพันธ์แคนนอลนิคอลของตัวแปรแคนนอลนิคอลคู่ที่ 1 ไม่เท่ากับ 0 แสดงว่า ตัวแปรแคนนอลนิคอลคู่ที่ 1 มีความสัมพันธ์กัน

4. การเกิดปัญหาตัวแปรอิสระไม่เป็นอิสระกัน (Multicollinearity)

กัลยา (2546) กล่าวว่า การวิเคราะห์การถดถอยหลายตัวแปรเป็นการศึกษาถึงความสัมพันธ์ระหว่างตัวแปรตามหลายตัวกับตัวแปรอิสระหลายตัวนั้น มีข้อกำหนดว่าตัวแปรอิสระต้องไม่มีความสัมพันธ์กัน แต่ในทางปฏิบัติจะพบว่าตัวแปรอิสระมักจะมีความสัมพันธ์กันเอง การที่ตัวแปรมีความสัมพันธ์กันจะทำให้เกิดปัญหาที่เรียกว่า Multicollinearity การเกิดปัญหา Multicollinearity จะมากหรือน้อยจะขึ้นอยู่กับความสัมพันธ์ระหว่างตัวแปรอิสระ ถ้าตัวแปรอิสระมีความสัมพันธ์กันมากปัญหา Multicollinearity จะมากด้วย

การที่ตัวแปรอิสระมีความสัมพันธ์กันจะทำให้ผลของการวิเคราะห์ผิดปกติไป คือ จะทำให้ค่าเบี่ยงเบนมาตรฐานของสัมประสิทธิ์ความถดถอยมีค่าสูงมาก จะทำให้เครื่องหมายของสัมประสิทธิ์ความถดถอยตรงข้ามกับที่ควรจะเป็น และจะทำให้ค่าสัมประสิทธิ์ไม่คงที่ เมื่อมีตัวแปรอิสระมากขึ้น

การแก้ไขเมื่อเกิดปัญหา Multicollinearity อาจทำได้ดังนี้ คือ คำนวณหาสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรอิสระต่าง ๆ แล้วทำการทดสอบสมมติฐาน ถ้าพบว่ายอมรับ H_0 แสดงว่าตัวแปรอิสระต่าง ๆ ไม่มีความสัมพันธ์กัน ในทางปฏิบัติการที่จะหาตัวแปรที่เป็นอิสระต่อกันทุกคู่นั้นเป็นไปได้ยาก กรณีที่ตัวแปรอิสระบางคู่ที่มีความสัมพันธ์กัน อาจแก้ปัญหาโดยการตัดตัวแปรอิสระตัวใดตัวหนึ่งที่สัมพันธ์กันออกจากสมการถดถอยหรืออาจแก้ปัญหาโดยการใช้เทคนิคการวิเคราะห์ปัจจัย

5. การวิเคราะห์ปัจจัย (Factor Analysis)

การวิเคราะห์ปัจจัยเป็นเทคนิคการวิเคราะห์ตัวแปรหลายตัว (Multivariate statistical technique) ที่มีการนำไปใช้อย่างกว้างขวางแทบทุกวงการวิชาการ ไม่ว่าจะเป็นในด้านสังคมศาสตร์ เช่น สังคมวิทยา รัฐศาสตร์ ประชากรศาสตร์ จิตวิทยาสังคม และการวิจัยตลาด หรือในด้านวิทยาศาสตร์ สาธารณสุข พยาบาลศาสตร์ เกษตรศาสตร์ เป็นต้น

การวิเคราะห์ปัจจัย นักวิชาการบางกลุ่มอาจเรียกว่า การวิเคราะห์ตัวประกอบ ซึ่งแต่เดิมการเรียกเช่นนี้อาจจะถูกต้อง เพราะเดิมนั้นเทคนิคหลักของการวิเคราะห์ปัจจัยคือ การหาองค์ประกอบหลัก (Principal Components) ของตัวแปรต่าง ๆ แต่ในปัจจุบันนี้การวิเคราะห์ปัจจัยได้ก้าวหน้าไปมาก มีเทคนิคต่าง ๆ มากมายในการแยกปัจจัยต่าง ๆ จากข้อมูลที่นอกเหนือไปจากวิธีการตัวประกอบหรือองค์ประกอบหลัก (สุชาติ, 2540)

5.1 วัตถุประสงค์หลักของการวิเคราะห์ปัจจัย มีรายละเอียดดังนี้

5.1.1. การลดจำนวนตัวแปรให้น้อยลง หรือการศึกษาความสัมพันธ์ระหว่างตัวแปร

การลดจำนวนตัวแปรให้น้อยลงทำได้โดยอาศัยความสัมพันธ์ระหว่างตัวแปรหรือความร่วมกัน (Communality) ระหว่างตัวแปรเป็นฐานในการเปลี่ยนสภาพตัวแปรหลายตัวให้มารวมกันเป็นปัจจัยเพียงไม่กี่ปัจจัย แต่ละปัจจัยจะประกอบด้วยตัวแปรหลายตัวที่มีความร่วมกันสูง ในการเปลี่ยนสภาพตัวแปรให้เป็นปัจจัยจะต้องพยายามให้ปัจจัยที่ได้สามารถอธิบายการผันแปรของตัวแปรให้ได้มากที่สุด สมมติว่ามีตัวแปร 20 ตัว เทคนิคการวิเคราะห์ปัจจัยอาจแยกปัจจัยได้ 4 ปัจจัย ปัจจัยทั้ง 4 นี้ ควรที่จะสามารถอธิบายการผันแปรของตัวแปรทั้ง 20 ตัวที่ได้ให้ได้มาก

ที่สุด เช่น ร้อยละ 60 ขึ้นไป หากผู้วิจัยต้องการนำปัจจัยทั้ง 4 ไปใช้แทนตัวแปรทั้งหมด แต่ถ้าหากต้องการศึกษาการรวมกลุ่มของตัวแปรเป็นปัจจัยว่าจะรวมกันได้หรือไม่และเป็นกี่กลุ่ม ก็ไม่มีความจำเป็นว่าจะต้องอธิบายได้กี่เปอร์เซ็นต์ของการผันแปรทั้งหมด

สมมติว่าตัวแปร 20 ตัวนี้ เป็นตัวแปรที่ถูกแปลงค่าเป็นค่ามาตรฐาน คือมีค่าเฉลี่ยเท่ากับ 0 และค่าการผันแปรเท่ากับ 1 การผันแปรรวมทั้งหมดของตัวแปร 20 ตัวนี้ คือ ค่าไอเก็น(Eigenvalue) ซึ่งเท่ากับค่าการผันแปรของแต่ละตัวแปรรวมกันคือเท่ากับ 20 ถ้าปัจจัยทั้ง 4 ที่สกัดได้สามารถอธิบายตัวแปรทั้งหมดได้ครบถ้วน ก็หมายความว่าปัจจัยทั้ง 4 นี้ อธิบายค่าการผันแปรของตัวแปรทั้งหมดได้ 20 ค่าไอเก็น

ในทางปฏิบัติปัจจัยที่สกัดได้จากการวิเคราะห์อาจจะมีหลายปัจจัยหรือเท่ากับจำนวนตัวแปรจึงจะสามารถอธิบายได้ครบ ซึ่งในหลักการและแนวปฏิบัติ ผู้วิจัยจะไม่ใช้ทุกปัจจัยที่รวมกันแล้วอธิบายการผันแปรได้ทั้งหมดทุกตัว แต่จะเลือกใช้ปัจจัยที่สามารถอธิบายการผันแปรของตัวแปรได้อย่างน้อยที่สุดไม่น้อยกว่า 1 ตัวแปรหรือไม่น้อยกว่า 1 ไอเก็น

การที่เทคนิคการวิเคราะห์ปัจจัยสามารถหาปัจจัยเพียงไม่กี่ปัจจัยมาแทนตัวแปรจำนวนมากได้ ทำให้ผู้วิจัยสามารถที่จะนำปัจจัยเพียงไม่กี่ปัจจัยไปใช้ได้สะดวกมากกว่าการใช้ตัวแปรชุดเดิมซึ่งมีจำนวนมาก นอกจากนั้นแล้วเทคนิคการวิเคราะห์ทางสถิติบางวิธี เช่น การวิเคราะห์ถดถอยเมื่อนำตัวแปรจำนวนมากมาใช้ในการวิเคราะห์มักจะก่อให้เกิดปัญหาความสัมพันธ์ระหว่างตัวแปรอิสระ (Multicollinearity) แต่หากนำตัวแปรเหล่านี้มารวมกันเหลือเพียงไม่กี่ปัจจัยที่ไม่มีความสัมพันธ์กัน อาจจะใช้ปัจจัยเหล่านี้แทนตัวแปรทั้งหมดในการวิเคราะห์ถดถอยได้

การที่ตัวแปรที่สัมพันธ์กันหรือมีส่วนที่มีความร่วมกันสูงมารวมกันเป็นปัจจัยเดียวกันทำให้ปัจจัยแต่ละปัจจัยมีความหมายขึ้น ความหมายของปัจจัยเหล่านี้ขึ้นอยู่กับคุณลักษณะหรือคุณสมบัติของตัวแปรเหล่านี้ที่มารวมเป็นปัจจัยเดียวกัน เช่น หากตัวแปรที่มารวมกันเป็นปัจจัยเดียวกัน เป็นตัวแปรเกี่ยวกับรายได้ อาชีพ ทรัพย์สินในครอบครอง ภาวะการเป็นหนี้สินอาจเรียกปัจจัยเหล่านี้ว่าเป็นปัจจัยเศรษฐกิจ หรือหากตัวแปรเหล่านี้เกี่ยวข้องกับสีของผลิตภัณฑ์ การบรรจุห่อ กลิ่นและรูปของสินค้า อาจเรียกปัจจัยเหล่านี้ว่า ปัจจัยที่ว่าด้วยรสนิยมของผู้บริโภค เป็นต้น

5.1.2. การทดสอบยืนยันความถูกต้องของมาตรวัด

ในการวิจัยหนึ่ง ๆ ถ้าผู้วิจัยได้สร้างมาตรวัดขึ้นมาใหม่ที่ประกอบด้วยตัวแปรหลายตัวและมีการให้น้ำหนักมากน้อยของแต่ละตัวแปรที่นำมารวมกัน ปัญหาที่ตามมาคือมาตรวัดที่สร้างขึ้นมานั้นควรรวมตัวแปรต่าง ๆ นี้หรือไม่ และการให้น้ำหนักของตัวแปรเหล่านี้มีความถูกต้องหรือไม่ เทคนิคการวิเคราะห์ปัจจัยจะเป็นสิ่งหนึ่งที่จะทดสอบความถูกต้องของการสร้างมาตรวัดดังกล่าว จึงเรียกว่าเป็นการวิเคราะห์ปัจจัยเพื่อการยืนยัน

สมมติว่าในการสร้างมาตรวัดที่เรียกว่า ปัจจัยความเป็นเมือง (Urban factor, UF) ซึ่งประกอบด้วยความหนาแน่นของประชากร (Population density, PD) สัดส่วนของแรงงานที่ประกอบอาชีพบริการ (Percentage of labour employed in service sector, PS) และสัดส่วนของประชากรที่มีการศึกษามัธยมศึกษาขึ้นไป (Percentage of population with high school education or more, PH) ผู้วิจัยอาจกำหนดน้ำหนัก (weight, W) ของ PD เป็น w_1 ของ PS เป็น w_2 และ PH เป็น w_3 ดังนั้น $UF = w_1PD + w_2PS + w_3PH$ และปัญหาก็คือน้ำหนัก w_1, w_2 และ w_3 ที่ให้นั้นถูกต้องหรือไม่ และ UF ควรประกอบ ด้วย PD, PS และ PH จริงหรือไม่ โดยเฉพาะอย่างยิ่งหากมีตัวแปรตัวอื่น ๆ มากกว่านั้น ซึ่งการวิเคราะห์ปัจจัยสามารถให้คำตอบ เพื่อพิสูจน์ยืนยันหรือลบล้างว่าปัจจัย UF ประกอบด้วยตัวแปร ทั้งสามตัวนั้นหรือไม่ โดยเฉพาะอย่างยิ่งหากมีตัวแปรอื่น ๆ ที่นอกเหนือกว่านั้นมาร่วมทดสอบด้วย จะพบว่าการวิเคราะห์ปัจจัยจะให้ผลที่จะบอกให้ผู้วิจัยทราบว่าปัจจัยที่อาจเรียกว่า UF ควรประกอบด้วยตัวแปรใดและตัวแปรต่าง ๆ ที่ประกอบเป็นปัจจัยนั้นแต่ละตัวควรมีน้ำหนักเชิงสัมพัทธ์อย่างไร

5.2 ประโยชน์ของการวิเคราะห์ปัจจัย

5.2.1. การรวมกลุ่มตัวแปร เทคนิคการวิเคราะห์ปัจจัยเป็นเทคนิคหนึ่งที่ใช้ในการจัดกลุ่มหรือจัดหมวดหมู่ของตัวแปร เมื่อผู้วิจัยมีตัวแปรจำนวนมากและไม่ทราบว่าจะจัดหมวดหมู่หรือกลุ่มของตัวแปรเหล่านี้อย่างไร หรือไม่แน่ใจว่าตัวแปรใดควรจะอยู่ในกลุ่มใด

ในการจัดหมวดหมู่หรือกลุ่มของตัวแปร นอกจากเทคนิคการวิเคราะห์ปัจจัยแล้ว ยังมีเทคนิคการวิเคราะห์กลุ่ม (Cluster Analysis) ที่สามารถนำมาวิเคราะห์หมวดหมู่ของตัวแปรได้ ความแตกต่างระหว่างการวิเคราะห์ปัจจัยและการวิเคราะห์กลุ่มอยู่ที่ว่า เทคนิคการ

วิเคราะห์จัดกลุ่มใช้ตัวแปรที่มีความสัมพันธ์กันในเชิงบวกเท่านั้นมารวมเป็นกลุ่มเดียวกัน ส่วนเทคนิคการวิเคราะห์ปัจจัยจะประกอบด้วยตัวแปรที่มีความสัมพันธ์กันมากไม่ว่าจะเป็นในเชิงลบหรือ เชิงบวก การตัดสินใจว่าจะใช้เทคนิคใดระหว่างสองเทคนิคนี้จึงอยู่ที่แนวคิดของผู้วิจัยเอง หากผู้วิจัยไม่ต้องการตัดสินใจและต้องการใช้ตัวแปรทั้งหมดก็ควรใช้เทคนิคการวิเคราะห์ปัจจัย แต่หากผู้วิจัยมีแนวคิดว่าตัวแปรที่รวมกันเป็นกลุ่มควรเป็นตัวแปรที่มีความสัมพันธ์ในเชิงบวก ก็ควรใช้เทคนิคการวิเคราะห์กลุ่ม

5.2.2. การกำหนดน้ำหนักเชิงสัมพันธ์ของตัวแปรที่นำมาสร้างมาตรวัดประกอบ

บ่อยครั้งที่ผู้วิจัยมีแนวคิดล่วงหน้าว่าตัวแปรหลายตัวที่นำมาสร้างมาตรวัดประกอบนั้น มีตัวแปรอะไรบ้าง ปัญหามีแต่เพียงว่าจะให้น้ำหนักแก่ตัวแปรเหล่านี้อย่างไร การนำเอาตัวแปรเหล่านี้มารวมกันเฉย ๆ ก็เท่ากับเป็นการให้น้ำหนักเท่า ๆ กัน ซึ่งหากตัวแปรแต่ละตัวมีขนาดไม่เท่ากัน ก็เท่ากับให้ความสำคัญอย่างมากกับตัวแปรที่มีขนาดใหญ่กว่า ทั้งนี้เพราะตัวแปรที่มีขนาดใหญ่กว่าจะกลืนความสำคัญของตัวแปรเล็กโดยสิ้นเชิง ซึ่งตัวแปรจะรวมกันได้ โดยไม่มีปัญหาดังกล่าว ก็ต่อเมื่อตัวแปรเหล่านั้นมีขนาดเท่ากัน หรือถูกทำให้มีขนาดเท่ากัน เช่น การทำให้เป็น Z-score

หากผู้วิจัยไม่ต้องการทำการตัดสินล่วงหน้าว่าจะให้น้ำหนักแก่ตัวแปรแต่ละตัวอย่างไร ผู้วิจัยอาจใช้เทคนิคการวิเคราะห์ปัจจัยช่วยกำหนด ซึ่งจะช่วยให้ทราบน้ำหนักของตัวแปรแต่ละตัวที่มีต่อปัจจัย (Factor loading)

5.2.3. การแก้ปัญหาความสัมพันธ์ระหว่างตัวแปรอิสระด้วยกัน เมื่อใช้เทคนิคการวิเคราะห์ ทางสถิติที่มี ข้อจำกัดห้ามมิให้ตัวแปรอิสระมีความสัมพันธ์กัน

เทคนิคการวิเคราะห์ทางสถิติบางเทคนิคมีข้อจำกัดที่ห้ามมิให้ตัวแปรอิสระมีความสัมพันธ์กัน โดยเฉพาะอย่างยิ่งการวิเคราะห์ถดถอยพหุ (Multiple Regression) และการวิเคราะห์เส้นทาง (Path Analysis) ดังนั้นเมื่อจะใช้เทคนิคดังกล่าวผู้วิจัยจะต้องทดสอบว่าตัวแปรต่าง ๆ ที่นำมาใช้เป็นตัวแปรอิสระมีความสัมพันธ์กันหรือไม่ หากไม่สัมพันธ์กันมากจึงจะใช้เทคนิคดังกล่าวได้ แต่ถ้าหากสัมพันธ์กันมาก ผู้วิจัยสามารถแก้ปัญหาดังกล่าวโดย

ก. ตัดตัวแปรใดตัวแปรหนึ่งที่มีความสัมพันธ์กันมากออกไป โดยมีแนวความคิดว่าตัวแปรที่เก็บไว้สามารถใช้เป็นตัวแทนตัวแปรอีกตัวหนึ่งได้

ข. การสร้างตัวแปรใหม่ เพื่อหลีกเลี่ยงปัญหาการตัดตัวแปรใดตัวแปรหนึ่งทิ้งไป ผู้วิจัยอาจสร้างตัวแปรใหม่ที่ประกอบด้วยตัวแปรที่มีความสัมพันธ์กันมาก หากมีเพียงสองตัวก็อาจทำได้โดยการรวมตัวแปรสองตัวเข้าด้วยกันโดยทำให้เป็นหน่วยที่เท่ากันก่อน ซึ่งวิธีนี้สามารถทำได้โดยใช้เทคนิคการวิเคราะห์ปัจจัยนั่นเอง

ค. การขจัดอิทธิพลของตัวแปรหนึ่งจากอีกตัวแปรหนึ่งด้วยการวิเคราะห์ถดถอย

5.2.4. การยืนยันความถูกต้องของการสร้างมาตรวัด

ผู้วิจัยทำการสร้างมาตรวัดโดยการนำคะแนนต่าง ๆ มารวมกันโดยให้น้ำหนักแก่ตัวแปร แต่ละตัวตามแนวความคิดที่ผู้วิจัยคิดว่าเหมาะสมและมีเหตุผลมากที่สุด แต่ปรากฏว่าเมื่อนำมาตรวัดนั้นมาใช้ในการวิจัยแล้วพบว่าไม่เป็นไปตามทฤษฎีที่คาดหวัง ปัญหาก็คือว่ามาตรวัดที่สร้างขึ้นมาหรือน้ำหนักที่ให้แก่ตัวแปรแต่ละตัวนั้นถูกต้องหรือไม่

ในการทดสอบความถูกต้องของมาตรวัดที่สร้าง ผู้วิจัยอาจจะใช้เทคนิคการวิเคราะห์ปัจจัยกำหนด น้ำหนักตัวแปรต่าง ๆ เหล่านี้ แล้วทำการเปรียบเทียบว่า สอดคล้องกับแบบแผนของการให้น้ำหนักตามที่ผู้วิจัยกำหนดไว้หรือไม่ ถ้าสอดคล้องกันก็แสดงว่าการให้น้ำหนักนั้นมีความถูกต้องอย่างน้อยที่สุดก็ในด้านแบบแผนการให้น้ำหนักเชิงสัมพัทธ์ระหว่างตัวแปรต่าง ๆ ที่นำมาสร้างเป็นมาตรวัดประกอบ

ในการวิเคราะห์ปัจจัย ประกอบด้วยคำต่าง ๆ ซึ่งอาจให้ความหมายได้ดังต่อไปนี้

1. ปัจจัยร่วม (Common Factor) หมายถึงปัจจัยที่ประกอบด้วยตัวแปรตั้งแต่ 2 ตัวขึ้นไปมาอยู่ร่วมกันในปัจจัยเดียวกัน โดยใช้ความสัมพันธ์ที่มีต่อกัน ปัจจัยที่ประกอบด้วยตัวแปรที่มีความสัมพันธ์กันมาก จะเป็นปัจจัยที่มีความหมาย

2. ปัจจัยเฉพาะ (Specific Factor) หมายถึงปัจจัยที่ประกอบด้วยตัวแปรตัวเดียว

3. ความร่วมกัน (Communality) หมายถึง ค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรหนึ่งกับตัวแปรอื่น ๆ ทั้งหมดที่เหลือ มีค่าอยู่ระหว่าง 0 กับ 1

4. น้ำหนักปัจจัย (Factor loading) หมายถึง ค่าความสัมพันธ์ระหว่างตัวแปรกับปัจจัย

5. คะแนนปัจจัย (Factor Score) หมายถึง คะแนนที่ได้จากน้ำหนักปัจจัยและค่าของตัวแปรในปัจจัยนั้น เพื่อใช้เป็นค่าของตัวแปรใหม่ที่เรียกว่าปัจจัย

6. ค่าไอเก็น (Value Eigen) หมายถึง ค่าความผันแปรรวมของตัวแปรทั้งหมดในแต่ละปัจจัย ในการวิเคราะห์ปัจจัย ปัจจัยร่วมที่ได้ปัจจัยแรกจะเป็นปัจจัยที่แยกความผันแปรออกมาจากปัจจัยอื่นได้มากที่สุด จึงมีค่าไอเก็นสูงที่สุด

5.3 หลักเกณฑ์ของเทคนิคการวิเคราะห์ปัจจัย

การวิเคราะห์ปัจจัยจะมีข้อสมมติว่าเวกเตอร์ตัวแปร x สามารถเขียนเป็นฟังก์ชันของตัวแปรร่วม (Common Factor) และ ค่าเฉพาะของตัวแปรแต่ละตัว โดยที่ปัจจัยร่วมจะแสดงถึงความร่วมกันของตัวแปรต่าง ๆ หลาย ๆ ตัว หรือปัจจัยร่วมจะเป็นปัจจัยที่ประกอบด้วยรายละเอียดหรือความ ผันแปรของตัวแปรหลายตัว (กัลยา , 2548)

กรณีที่มีตัวแปร p ตัว ($X_1, X_2, \ldots, X_p$) และถ้ามีปัจจัยร่วม m ปัจจัย ($F_1, F_2, \ldots, F_m$) รูปแบบของสมการของการวิเคราะห์ปัจจัย คือ

$$
\begin{aligned}
X_1 &= L_{11}F_1 + L_{12}F_2 + \cdots + L_{1m}F_m + e_1 \\
X_2 &= L_{21}F_1 + L_{22}F_2 + \cdots + L_{2m}F_m + e_2 \\
&\vdots \\
X_p &= L_{p1}F_1 + L_{p2}F_2 + \cdots + L_{pm}F_m + e_p
\end{aligned}
$$

หรือเขียนในรูปเมทริกซ์และเวกเตอร์ได้ดังนี้

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = \begin{bmatrix} L_{11} & L_{12} & \cdots & L_{1m} \\ L_{21} & L_{22} & \cdots & L_{2m} \\ \vdots & & & \\ L_{p1} & L_{p2} & \cdots & L_{pm} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

หรือ $\underset{(px1)}{\mathbf{X}} = \underset{(pxm)}{\mathbf{L}} \underset{(mx1)}{\mathbf{F}} + \underset{(px1)}{\mathbf{e}}$

โดยที่ L_{ij} แทนค่าสัมประสิทธิ์หรือน้ำหนักของตัวแปร (factor loading) ซึ่งเป็นค่าที่แสดงความมีส่วนร่วมของตัวแปร X_i ที่มีอยู่

m แทนจำนวนปัจจัยร่วม และ p แทนจำนวนตัวแปร

ตัวอย่างเช่น การศึกษาปัจจัยร่วมที่ 2 จะพิจารณาแถวตั้งที่ 2 ของเมตริกซ์ L ได้ค่า factor loading คือ $L_{12}, L_{22}, \ldots, L_{p2}$ ถ้าค่า L_{i2} ค่าใดมาก แสดงว่า X_i จะแสดงความหมายของปัจจัยร่วมที่ 2 ได้ดี

การวิเคราะห์ปัจจัยสรุปได้ดังนี้

1. ค่าแปรปรวนของตัวแปรหนึ่ง ๆ จะแบ่งเป็น 2 ส่วน คือค่าความร่วมกัน (x_i) และค่าแปรปรวนของค่าเฉพาะ ($Var(X_i)$) ดังนี้

$$x_i = L_{i1}F_1 + L_{i2}F_2 + \ldots + L_{im}F_m + e_i$$

$$Var(x_i) = L_{i1}^2 + L_{i2}^2 + \ldots + L_{im}^2 + Var(e_i) = h_i^2 + \theta_i$$

โดยที่ค่าความร่วมกัน (Communality) ของตัวแปร X_i คือ h_i^2 เมื่อ $h_i^2 = L_{i1}^2 + L_{i2}^2 + \ldots + L_{im}^2$

2. ค่า factor loading (L_{il}) เป็นค่าที่แสดงความสัมพันธ์ระหว่างตัวแปรเดิม (X_i) กับปัจจัยร่วม (F_j) กรณีที่ปรับข้อมูลให้อยู่ในรูปมาตรฐาน ค่า factor loading (L_{il}) จะเป็นค่าสหสัมพันธ์ของ X_i และ F_j

3. ถ้าปรับข้อมูลให้อยู่ในรูปมาตรฐานแล้ว ค่าสหสัมพันธ์ของตัวแปรเดิม X_i และ X_j จะเป็นผลคูณของค่า factor loading

5.4 ขั้นตอนการวิเคราะห์ปัจจัย

ขั้นตอนในการวิเคราะห์ปัจจัยแบ่งได้ 5 ส่วน คือ

5.4.1 การตรวจสอบความเหมาะสมในการใช้เทคนิคการวิเคราะห์ปัจจัย

การที่จะสร้างปัจจัยร่วม ซึ่งเป็นตัวแทนของตัวแปรหลาย ๆ ตัวได้นั้นแสดงว่าตัวแปรเหล่านั้นจะต้องมีความสัมพันธ์กัน ถ้าตัวแปรไม่มีความสัมพันธ์กันจะทำให้ไม่สามารถสร้างปัจจัยร่วมได้ จึงต้องมีการตรวจสอบความสัมพันธ์ของตัวแปรก่อนที่จะทำการสร้างปัจจัยร่วม การตรวจสอบความสัมพันธ์ของตัวแปรทำได้หลายวิธี ดังนี้ (กัลยา, 2548)

วิธีที่ 1 ตรวจสอบโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์

ตรวจสอบความสัมพันธ์ระหว่างตัวแปรโดยใช้เมตริกซ์ค่าสหสัมพันธ์ ซึ่งเป็นการตรวจสอบความสัมพันธ์ของตัวแปรทุกคู่

- ถ้าตัวแปรคู่ใดมีความสัมพันธ์มาก คือเข้าสู่ +1 หรือ –1 แสดงว่าตัวแปรคู่นั้นมีความสัมพันธ์กันมาก หรือมีส่วนร่วมกันมากในปัจจัยเดียวกัน ดังนั้นตัวแปรหลาย ๆ ตัวที่มีความสัมพันธ์กันมาก ควรจะมีความผันแปรร่วมกันมาก ทำให้ปัจจัยร่วมสามารถอธิบายหรือเป็นตัวแทนของตัวแปรกลุ่มที่มีความสัมพันธ์กันมากได้ดี
- ถ้าตัวแปรคู่ใดมีค่าสัมประสิทธิ์สหสัมพันธ์ต่ำหรือใกล้ศูนย์ แสดงว่าไม่มีความสัมพันธ์กัน หรือกล่าวได้ว่าไม่มีส่วนที่ร่วมกัน
- ถ้าตัวแปรใดที่ไม่มีความสัมพันธ์กับตัวแปรอื่น ๆ เลย ควรตัดตัวแปรนั้นออก ก่อนที่จะทำการวิเคราะห์ปัจจัย

ดังนั้นจึงสรุปได้ว่าเทคนิคการวิเคราะห์ปัจจัยเป็นเทคนิคที่ทำการแยกหรือแบ่งกลุ่มตัวแปร โดยให้ตัวแปรที่อยู่กลุ่มเดียวกันมีส่วนที่ร่วมกันมากหรือมีความสัมพันธ์กันมาก และเรียกแต่ละกลุ่มว่าปัจจัยร่วม

การตรวจสอบโดยการใช้เมตริกซ์สัมประสิทธิ์สหสัมพันธ์จะยุ่งยากกรณีที่มีจำนวนตัวแปรมาก และเป็นวิธีที่ผู้ศึกษาจะต้องตัดสินใจเองว่าข้อมูลที่อยู่ควรจะใช้เทคนิคการวิเคราะห์ปัจจัยหรือไม่

วิธีที่ 2 ตรวจสอบโดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วน

วิธีนี้จะตรวจสอบระดับความสัมพันธ์ระหว่างตัวแปร โดยใช้ค่าสัมประสิทธิ์สหสัมพันธ์บางส่วน ทำการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนของตัวแปรแต่ละคู่ โดยการควบคุมหรือกำจัดผลกระทบในรูปเชิงเส้นของตัวแปรอื่น ๆ ออกไป ถ้าตัวแปรมีความสัมพันธ์กัน หรือมีส่วนร่วมกันมาก ค่าสัมประสิทธิ์สหสัมพันธ์จะมีค่าต่ำ นอกจากค่าสัมประสิทธิ์สหสัมพันธ์บางส่วนจะทำให้สามารถประมาณค่าสหสัมพันธ์ระหว่างค่าเฉพาะ (e_i และ e_j) ได้ จากข้อสมมติของการวิเคราะห์ปัจจัยที่ว่าค่าเฉพาะของตัวแปรไม่มีความสัมพันธ์กัน ดังนั้นถ้าค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างค่าเฉพาะมีค่าใกล้ศูนย์ แสดงว่าข้อมูลชุดนั้นสามารถใช้เทคนิคการวิเคราะห์ปัจจัยได้

วิธีที่ 3 ตรวจสอบโดยใช้ Bartlett’s Sphericity Test

เพื่อตรวจสอบว่าเมตริกซ์สัมประสิทธิ์สหสัมพันธ์ของตัวแปร p ตัว สามารถแบ่งเป็นปัจจัยร่วมหรือกลุ่มตัวแปรที่มีส่วนร่วมกันหรือสัมพันธ์กันได้หรือไม่ โดยที่ Bartlett’s Sphericity Test มีการแจกแจงโดยประมาณแบบไคสแควส์ ซึ่งเป็นฟังก์ชันของค่าดีเทอร์มิแนนต์ของเมตริกซ์สัมประสิทธิ์สหสัมพันธ์ (R)

$$\chi^2 = -\left[(n-1)-\frac{(2p+5)}{6}\right]\ln|R|$$

โดยที่ χ^2 มีองศาอิสระเท่ากับ $\frac{1}{2}p(p-1)$

$\ln|R|$ เป็นค่าล็อกของดีเทอร์มิแนนของเมตริกซ์สหสัมพันธ์ R

p เป็นจำนวนตัวแปร

n เป็นจำนวนข้อมูล

สมมติฐาน H_0 : เมตริกซ์สหสัมพันธ์เป็นเมตริกซ์เอกลักษณ์

H_1 : เมตริกซ์สหสัมพันธ์ไม่เป็นเมตริกซ์เอกลักษณ์

สถิติทดสอบ $$\chi^2 = -\left[(n-1)-\frac{(2p+5)}{6}\right]\ln|R|$$

ถ้าตัวแปรไม่มีความสัมพันธ์กันจะได้ว่า $R = I$ ซึ่งทำให้ค่าไอเก็นของตัวแปรมีค่าใกล้ 1 หรือ $|R|$ มีค่าใกล้ 1 ซึ่งมีผลทำให้ $\ln|R|$ ลู่เข้าสู่ศูนย์ และค่า χ^2 จะต่ำจึงไม่สามารถปฏิเสธสมมติฐานหลักได้ นั่นคือ ไม่สมควรใช้การวิเคราะห์ปัจจัย

ถ้าค่าสัมประสิทธิ์สหสัมพันธ์มีค่ามาก หรือตัวแปรมีความสัมพันธ์กัน จะทำให้ค่าไอเก็นของบางตัวแปรมีค่ามากกว่าหนึ่ง ขณะที่ค่าไอเก็นของบางตัวแปรมีค่าใกล้ศูนย์ จะทำให้ $\ln|R|$ มีค่าติดลบมาก ซึ่งมีผลให้ค่า χ^2 สูง จึงปฏิเสธสมมติฐานหลัก แสดงว่าสมควรใช้การวิเคราะห์ปัจจัย

การสรุปผลการทดสอบ จะปฏิเสธ H_0 ถ้า $\chi^2 > \chi^2$ ที่องศาอิสระ $\frac{1}{2}p(p-1)$

ซึ่งแสดงการคำนวณได้ดังตัวอย่างต่อไปนี้

ตัวอย่าง ในกรณีที่ตัวแปรอิสระ 6 ตัว ($X_1, X_2, X_3, X_4, X_5, X_6$) ขนาดตัวอย่าง เท่ากับ 86 และต้องการตรวจสอบว่าควรจะใช้การวิเคราะห์ปัจจัยหรือไม่ โดยใช้สถิติ Bartlett's Sphericity Test สามารถแสดงได้ดังนี้

เมตริกซ์สัมประสิทธิ์สหสัมพันธ์

	X_1	X_2	X_3	X_4	X_5	X_6
X_1	1.0					
X_2	-0.09	1.0				
X_3	0.02	0.31	1.0			
X_4	0.07	0.01	0.10	1.0		
X_5	0.04	0.09	-0.24	-0.06	1.0	
X_6	0.21	0.13	0.12	0.23	0.08	1.0

จากเมตริกซ์สัมประสิทธิ์สหสัมพันธ์ข้างต้น คำนวณค่าไอเก็น ของตัวแปรได้ดังนี้

$\lambda_1 = 1.491$, $\lambda_2 = 1.266$, $\lambda_3 = 1.122$, $\lambda_4 = 0.905$, $\lambda_5 = 0.671$ และ $\lambda_6 = 0.547$

ค่าดีเทอร์มิแนนต์ของ $\mathbf{R} = |\mathbf{R}| = \prod_{i=1}^{6} \lambda_i$

$= (1.491)(1.266)(1.122)(0.905)(0.671)(0.547) = 0.701$

$\therefore \ln|\mathbf{R}| = \ln(0.701) = -0.355$

$$\chi^2 = -\left[(86-1) - \frac{(2(6)+5)}{6} \right](-0.355) = 29.2$$

จากตารางไคสแควส์ ที่ระดับนัยสำคัญ 0.05 ค่าองศาอิสระ เท่ากับ $\frac{1}{2}p(p-1) = 15$ ได้ค่า $\chi^2_{0.5,15} = 25$ จึงปฏิเสธ H_0 แสดงว่า ควรใช้การวิเคราะห์ปัจจัย

5.4.2 การสกัดปัจจัย (Factor Extraction)

การสกัดปัจจัยเป็นการสร้างหรือหาปัจจัยร่วมจำนวนหนึ่งซึ่งมีจำนวนน้อยกว่าจำนวนตัวแปร โดยให้ปัจจัยร่วมสามารถแทนตัวแปรเดิม หรือสามารถสกัดความผันแปรต่างๆ ของตัวแปรเดิมไว้ในปัจจัยร่วม ทำให้ปัจจัยร่วมแต่ละปัจจัยเป็นตัวแทนที่ดีของตัวแปรเดิมหลาย ๆ ตัวที่

มีส่วนร่วมในปัจจัยหลักแต่ละปัจจัยมาก วิธีการสกัดปัจจัยมีหลายวิธี เช่น วิธีองค์ประกอบหลัก (Principal Component Analysis) วิธีแกนหลัก (Principal Axis Factoring) วิธีกำลังสองน้อยที่สุด (Generalized Least Square) วิธีเงา (Image) เป็นต้น

การเลือกว่าวิธีใดเหมาะสม จะขึ้นกับหลักเกณฑ์ของแต่ละวิธี โดยทุกวิธีมีวัตถุประสงค์เพื่อหา Factor loading ซึ่งในที่นี้จะกล่าวถึงเฉพาะวิธีองค์ประกอบหลัก (Principal Component Analysis) ดังนี้

ก. หลักการวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis)

การวิเคราะห์องค์ประกอบหลักเป็นเทคนิคการลดจำนวนตัวแปรที่มีจำนวนมาก โดยการสร้างตัวแปรใหม่ซึ่งเป็นฟังก์ชันของตัวแปรเดิม p ตัว และจะต้องสกัดหรือดึงรายละเอียดหรือค่าแปรปรวนจากตัวแปรเดิมมาไว้ในตัวแปรใหม่ให้มากที่สุด

ในที่นี้จะใช้ PC_i แทนองค์ประกอบหลักที่ i ; i = 1, 2, …, p และการสร้าง PC_i มีขั้นตอนดังนี้

ขั้นที่ 1 การสร้าง PC_1 หรือองค์ประกอบหลักที่ 1 ซึ่งเป็นตัวแปรใหม่ตัวแรก โดยให้ PC_1 เป็นฟังก์ชันเชิงเส้นของตัวแปรเดิมทั้ง p ตัว และจะต้องสกัดหรือดึงค่าความแปรปรวนจากตัวแปรทั้ง p ตัวมาไว้ใน PC_1 ให้มากที่สุด ซึ่งจะทำให้ PC_1 มีค่าแปรปรวนมากที่สุด โดยที่

$$PC_1 = w_{11}X_1 + w_{12}X_2 + \ldots + w_{1p}X_p = w_1'x$$

เมื่อ $Var(w_1'x)$ มีค่ามากที่สุด และ $w_1'w_1 = 1$

ขั้นที่ 2 การสร้าง PC_2 จะเป็นฟังก์ชันเชิงเส้นของตัวแปรเดิม p ตัว และสกัดรายละเอียดหรือความแปรปรวนที่เหลือจาก PC_1 มาไว้ใน PC_2 ให้มากที่สุดเท่าที่จะทำได้ และ PC_2 จะต้องไม่มีความสัมพันธ์กับ PC_1 หรือจะต้องตั้งฉากกับ PC_1 และ

$$PC_2 = w_{21}X_1 + w_{22}X_2 + \ldots + w_{2p}X_p = w_2'x$$

เมื่อ $Var(w_2'x)$ มีค่ามากที่สุด โดยมีเงื่อนไข $w_2'w_2 = 1$, $w_1'w_2 = 0$ และ $Cov(w_1'x, w_1'x) = 0$

ขั้นที่ p การสร้าง PC_p ซึ่งเป็นฟังก์ชันเชิงเส้นของตัวแปรเดิม p ตัว โดย PC_p จะมีความผันแปรของตัวแปรเดิมที่เหลือจาก $PC_1, PC_2, \dots, PC_{p-1}$ และไม่มีความสัมพันธ์กับ $PC_1, PC_2, \dots, PC_{p-1}$ โดยที่

$$PC_p = w_{p1}X_1 + w_{p2}X_2 + \dots + w_{pp}X_p = w_p'x$$

เมื่อ $Var(w_p'x)$ มีค่ามากที่สุด โดยมีเงื่อนไข $w_p'w_p = 1$ และ $Cov(w_p'x, w_j'x) = 0$ สำหรับ $j < p$

จากวิธีการข้างต้นทำให้ ค่าแปรปรวนของ PC_1 จะมากที่สุด และค่าแปรปรวนของ PC อื่น ๆ จะลดลงเรื่อย ๆ จนทำให้ค่าแปรปรวนของ PC_p ต่ำสุด

5.4.3 การกำหนดปัจจัยร่วมที่เหมาะสม

ในการวิเคราะห์ปัจจัย เพื่อหาจำนวนปัจจัยร่วมซึ่งเป็นตัวแทนของตัวแปรต่าง ๆ โดยปัจจัยร่วมดังกล่าวจะใช้อธิบายความสัมพันธ์ระหว่างตัวแปรหลาย ๆ ตัว การพิจารณาว่าควรมีปัจจัยร่วมกี่ปัจจัยนั้น จะพิจารณาจาก

ก. ค่าไอเก็น โดยที่ปัจจัยร่วมแต่ละปัจจัยควรมีค่าไอเก็นมากกว่า 1

ข. ใช้กราฟ scree plot

Velicer (1986) ได้เสนอวิธีพิจารณาจำนวนปัจจัยร่วมที่เหมาะสม โดยพิจารณาจากค่าเฉลี่ยของสัมประสิทธิ์สหสัมพันธ์บางส่วน ถ้ามีจำนวนปัจจัยร่วม m ปัจจัย แล้วทำให้ค่าเฉลี่ยของค่าสัมประสิทธิ์บางส่วนต่ำสุด แสดงว่าควรมีปัจจัยร่วม m ปัจจัย โดยวิธีนี้จะใช้ได้ดีกรณีที่ในแต่ละปัจจัยร่วมมีจำนวนตัวแปรมาก ๆ

5.4.4 การหมุนแกนปัจจัย (Factor Rotation)

ดังที่กล่าวมาแล้วว่ากรณีที่ค่า loading ของตัวแปรต่าง ๆ ในปัจจัยร่วมไม่ชัดเจน กล่าวคือไม่มีค่า loading ของตัวแปรใดที่มีค่ามาก มีค่าใกล้ +1 หรือ –1 ในปัจจัยร่วมหนึ่งหรือมีค่าน้อย มีค่าใกล้ 0 ในปัจจัยร่วมอื่น ๆ จะทำให้ไม่สามารถให้ความหมายของปัจจัยร่วมได้ชัดเจน จึงต้องทำการหมุนแกนของปัจจัยร่วม โดยวัตถุประสงค์การหมุนแกนปัจจัยร่วม เพื่อทำให้ค่า loading ของตัวแปรแต่ละตัวมีค่าเพิ่มขึ้นในปัจจัยร่วมใดปัจจัยหนึ่ง และมีค่าลดลงในปัจจัยอื่น ๆ จึงทำให้ทราบว่าตัวแปรใดมีส่วนร่วมกับตัวแปรอื่น ๆ ตัวใดบ้างในปัจจัยร่วมแต่ละปัจจัย หรือการหมุนปัจจัยเพื่อที่จะทำให้เห็นโครงสร้างของปัจจัยร่วมได้ชัดเจน จนสามารถให้ความหมายหรืออธิบายความหมายของปัจจัยร่วมแต่ละปัจจัยได้

ก. วิธีการหมุนแกนปัจจัยร่วม

วิธีการหมุนแกนปัจจัยร่วมมีหลายวิธี ทำให้ผลลัพธ์ที่ได้จากการหมุนแกนอาจจะแตกต่างกันเมื่อใช้วิธีการหมุนที่แตกต่างกัน ถึงแม้จะทำให้ปัจจัยร่วมแตกต่างกัน หรือค่า loading ในเมตริกซ์ปัจจัยร่วมแตกต่างกันหรือเปลี่ยนไปเมื่อเทียบกับก่อนทำการหมุนแกนก็ตาม แต่ค่าความร่วมกันและสัดส่วนของความผันแปรทั้งหมดจะไม่เปลี่ยนแปลง และสัดส่วนของความผันแปรของตัวแปรแต่ละตัวที่มีส่วนร่วมในปัจจัยร่วมแต่ละปัจจัยไม่เปลี่ยนแปลง

วิธีการหมุนแกนปัจจัยร่วมมี 2 วิธีใหญ่ ๆ คือ วิธีการหมุนโดยให้แกนปัจจัยร่วมยังคงตั้งฉากกันและแบบไม่ตั้งฉากกันหรือที่เรียกว่าแบบเฉียง ดังมีรายละเอียดดังนี้

1) การหมุนแกนปัจจัยร่วมให้ตั้งฉากกัน (Orthogonal Rotation)

เป็นวิธีการหมุนแกนปัจจัยร่วมที่เมื่อหมุนแกนปัจจัยร่วมแล้วทำให้แกนปัจจัยร่วมตั้งฉากกันหรือเป็นอิสระกันเพื่อทำให้ค่า loading เพิ่มขึ้นในปัจจัยร่วมบางปัจจัยและลดลงในปัจจัยร่วมอื่น วิธีการหมุนแกนปัจจัยร่วมให้ตั้งฉากหรือให้ปัจจัยร่วมเป็นอิสระกัน เช่น วิธี Varimax , Quartimax และ Equamax โดยมีรายละเอียดดังนี้

1.1. วิธี Varimax การหมุนแกนปัจจัยร่วมโดยวิธี Varimax มีวัตถุประสงค์หลักเพื่อให้ตัวแปรแต่ละตัวมีค่าloading สูงในปัจจัยร่วมเพียงปัจจัยเดียวและมีค่า loading ต่ำมากหรือใกล้ ๆ ศูนย์ในปัจจัยร่วมอื่น ๆ หรือกล่าวได้ว่าเป็นวิธีที่ทำให้มีจำนวนตัวแปรน้อยที่สุดที่มีค่า loading สูงในปัจจัยร่วมที่ 1 ปัจจัย

1.2. วิธี Quartimax จะทำการหมุนแกนปัจจัยร่วมเพื่อให้ได้ factor loading ที่มีคุณสมบัติดังนี้ ตัวแปรแต่ละตัวควรจะมีค่า loading สูง 1 ค่าใน 1 ปัจจัยร่วมและมีค่าใกล้ศูนย์ในปัจจัยร่วมที่เหลือ หรือตัวแปรแต่ละตัวควรมีค่า loading ใกล้ +1 หรือ –1 ในปัจจัยร่วมปัจจัยใดปัจจัยหนึ่ง จากคุณสมบัติของ Quartimax จะทำให้มีปัจจัยร่วมหนึ่งที่เป็นปัจจัยร่วมของทุกตัวแปร ส่วนปัจจัยร่วมที่เหลือจะเป็นค่าเฉพาะ ดังนั้นวิธี Quartimax จึงเหมาะสมในกรณีที่ผู้ศึกษาคาดว่า มีปัจจัยหนึ่งซึ่งเป็นปัจจัยร่วมของทุกตัวแปร ซึ่งเรียกว่าเป็นปัจจัยทั่วไป

1.3. วิธี Equamax เป็นวิธีการหมุนแกนปัจจัยร่วมที่พิจารณาทั้งทางด้านคือแถวนอนและแถวตั้งของเมตริกซ์ค่า loading

2) การหมุนแกนแบบเฉียง (Oblique Rotation)

เป็นการหมุนแกนปัจจัยร่วมแล้วไม่ทำให้แกนปัจจัยร่วมตั้งฉากกัน ได้แก่วิธี Oblimax, Covarimin และ Oblimin

2.1 วิธี Oblimax เป็นการหมุนแกนปัจจัยร่วมที่ทำให้จำนวนค่า loading ที่มีค่ามากและมีค่าน้อยเพิ่มมากขึ้น ซึ่งจะเป็นการลดค่า loading ที่มีค่ากลาง ๆ ซึ่งทำให้จัดตัวแปรหรือให้ความหมายกับปัจจัยร่วมทำได้ง่ายขึ้น

2.2 วิธี Covarimin เป็นวิธีการแบบวิธี Varimax แต่แกนปัจจัยร่วมไม่ตั้งฉากกัน

5.4.5 การสร้างคะแนนปัจจัย (Factor Scores)

จากการหมุนแกนปัจจัยแล้ว ทำให้สามารถจัดตัวแปรหรือให้ความหมายแก่ปัจจัยร่วมแต่ละปัจจัยได้ ทำให้ปัจจัยร่วมที่ได้เป็นตัวแปรที่สร้างขึ้นใหม่ จึงคำนวณค่าตัวแปรใหม่เพื่อนำไปใช้ในการวิเคราะห์เทคนิคอื่น ๆ ต่อไป เช่นการวิเคราะห์การถดถอย เป็นต้น

ค่าของปัจจัยร่วมเรียกว่า คะแนนปัจจัย (Factor Score) โดยแทนค่าตัวแปรลงในสมการ

$$\underset{(nxm)}{\hat{\mathbf{F}}} = \underset{(nxp)}{\mathbf{X}} \quad \underset{(pxm)}{\mathbf{B}}$$

หรือ $F_{ij} = b_1X_{i1} + b_2X_{i2} + \ldots + b_3X_{i3}$; $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$

โดยที่ F = ค่าประมาณของคะแนนปัจจัย

X = ค่าของตัวแปร

b = สัมประสิทธิ์การถดถอยเชิงเส้น

ในกรณีที่ข้อมูลที่นำมาศึกษามีหน่วยที่ต่างกัน จะต้องทำการแปลงข้อมูลให้อยู่ในรูปค่ามาตรฐานก่อนทำการวิเคราะห์ ซึ่งคำนวณได้จากสูตรต่อไปนี้

$$Z = \frac{X - \mu}{\sigma}$$

โดยที่ Z = ค่ามาตรฐานของตัวแปร

X = ค่าของตัวแปร

μ = ค่าเฉลี่ยของตัวแปร

σ = ค่าเบี่ยงเบนมาตรฐานของตัวแปร

เอกสารงานวิจัยที่เกี่ยวข้อง

วาสนา (2523) ทำการศึกษาถึงระบบการตลาดและการขนส่งน้ำมันเชื้อเพลิงในประเทศไทย โดยทำการศึกษาระบบตลาดและการขนส่งน้ำมันเชื้อเพลิงจากผู้ขายส่งไปสู่ผู้บริโภคกลุ่มสำคัญทางเศรษฐกิจ 3 ประเภทคือ กิจการเกษตร กิจการประมง และกิจการขนส่งโดยรถบรรทุก เพื่อศึกษาปัญหาด้านราคา ความขาดแคลนน้ำมันเชื้อเพลิง ปัญหาและอุปสรรคต่าง ๆ ตลอดจนศึกษาถึงมาตรการการประหยัดน้ำมันเชื้อเพลิง และการแสวงหาพลังงานทดแทนที่อาจนำมาใช้ได้ในอนาคต

ในส่วนของการศึกษาทางด้านราคาน้ำมันเชื้อเพลิง ได้ทำการศึกษาและวิเคราะห์โครงสร้างราคาน้ำมันเชื้อเพลิงในประเทศไทย ปัจจัยที่มีผลกระทบต่อการเปลี่ยนแปลงน้ำมันเชื้อเพลิงในประเทศไทย และสร้างตัวแบบราคาน้ำมันเชื้อเพลิงขึ้น โดยทำการสัมภาษณ์เจ้าหน้าที่จากฝ่ายต่างๆ ทั้งทางโรงกลั่น ผู้ค้าน้ำมัน และเจ้าหน้าที่ของรัฐ ประกอบกับการเก็บข้อมูลทุติยภูมิเพื่อใช้ในการวิเคราะห์ จากการศึกษาพบว่าโครงสร้างราคาน้ำมันนับตั้งแต่การกลั่นน้ำมันดิบ จนถึงราคาขายปลีกในขณะนั้นประกอบด้วย

1. ราคาออกจากโรงกลั่นประกอบด้วย ราคาน้ำมันดิบ ค่าใช้จ่ายในการกลั่น ส่วนแบ่งของรัฐบาล ภาษีการค้า ภาษีรายได้ ฯลฯ

2. ราคาขายปลีก ประกอบด้วย ราคาออกจากโรงกลั่น ภาษีสรรพสามิต ส่วนของกองทุนน้ำมัน และค่าการตลาด

จากการศึกษาพบว่า ปัจจัยที่มีผลกระทบต่อการเปลี่ยนแปลงราคาน้ำมันภายในประเทศคือการเปลี่ยนแปลงของราคาน้ำมันดิบ ราคาที่ขายจริงจากโรงกลั่น รวมเงินเก็บเข้ากองทุนน้ำมัน และเงินชดเชยจากกองทุนน้ำมัน ภาษีการค้าและสรรพสามิต และส่วนแบ่งของรัฐบาล และยังมีการสร้างตัวแบบราคาน้ำมันเชื้อเพลิงในประเทศ ซึ่งประกอบด้วยตัวแปร 2 ประเภท คือ ประเภทที่ 1 ตัวแปรผันที่สามารถควบคุมได้ (Controllable Variable) คือ ค่าขนส่ง ระบบภาษี ข้อกำหนด การผสมน้ำมัน ส่วนแบ่งของรัฐ ค่าใช้จ่ายในการจัดจำหน่าย และค่าการตลาด และประเภทที่ 2 ตัวแปรผันที่ไม่สามารถควบคุมได้ (Uncontrollable Variable) คือราคาน้ำมันดิบ ค่าใช้จ่ายในการกลั่น และค่าขนส่งน้ำมันดิบจากต่างประเทศ

ปรเมสฐ์ (2533) ได้ทำการวิเคราะห์ผลกระทบของนโยบายการค้าภายในประเทศ และภายนอกประเทศที่มีผลต่อระดับราคาผลิตภัณฑ์ปิโตรเลียมของประเทศไทย โดยศึกษาผลกระทบจากนโยบายการค้าน้ำมัน และปัจจัยอื่น ๆ ที่มีอิทธิพลต่อระดับราคาผลิตภัณฑ์ปิโตรเลียมภายในประเทศ ทำการเก็บข้อมูลทุติยภูมิตั้งแต่ปี พ.ศ. 2503 – 2530 โดยพิจารณาถึงผลิตภัณฑ์ปิโตรเลียม 5 ชนิด คือ น้ำมันเบนซินซุปเปอร์ น้ำมันเบนซินธรรมดา น้ำมันดีเซลหมุนเร็ว น้ำมันดีเซลหมุนช้า และน้ำมันก๊าด และทำการวิเคราะห์โดยใช้วิธีการทางสถิติถึงความสัมพันธ์ระหว่างระดับราคาผลิตภัณฑ์ปิโตรเลียมภายในประเทศกับปัจจัยต่าง ๆ ที่มีอิทธิพล ซึ่งได้แก่ นโยบายการค้าน้ำมันภายในและภายนอกประเทศ โดยใช้สมการถดถอยตัวแปรเดียว (Simple Regression) และสร้างเป็นแบบจำลองผลจากนโยบายของกลุ่มโอเปคต่อราคาผลิตภัณฑ์ปิโตรเลียมในประเทศไทย โดยใช้อัตราแลกเปลี่ยนของเงินสกุลต่าง ๆ ของประเทศในกลุ่มโอเปคเป็นตัวแทนของนโยบายของประเทศ และใช้อัตราภาษีและกองทุนน้ำมันเป็นตัวแทนของนโยบายของรัฐบาลภายในประเทศ

จากการศึกษาผลกระทบจากปัจจัยภายนอกพบว่า นโยบายการค้าภายนอกประเทศเป็นปัจจัยสำคัญ โดยเกิดจากการกระทำของกลุ่มโอเปค ได้แก่ การร่วมกันประกาศขึ้นราคา ภาษี ค่าธรรมเนียม ฯลฯ ส่วนปัจจัยภายในประเทศที่สำคัญได้แก่ นโยบายการค้าน้ำมันภายในประเทศ เช่น การใช้กองทุนน้ำมัน หรือการใช้นโยบายควบคุมราคาโดยรัฐบาล

ธานินทร์ (2537) ได้ทำการศึกษาเรื่องการนำเข้าน้ำมันดิบ และผลิตภัณฑ์น้ำมันของประเทศไทย โดยศึกษาถึงปัจจัยที่มีผลกระทบต่อการนำเข้าน้ำมันดิบ และผลิตภัณฑ์น้ำมันของประเทศไทย โดยเก็บข้อมูลทุติยภูมิทางด้านปริมาณ และนำมาวิเคราะห์ในรูปสมการถดถอยเชิงซ้อน (Multiple Regression) โดยแยกการพิจารณาออกเป็นน้ำมันดิบ และผลิตภัณฑ์น้ำมันต่างๆ

จากผลการศึกษา พบว่า GDP เป็นปัจจัยสำคัญที่มีอิทธิพลต่อปริมาณความต้องการนำเข้าวัตถุดิบมากที่สุด รองลงมาคือ ปริมาณน้ำมันดิบที่ได้จากแหล่งภายในประเทศ และกำลังการกลั่นภายในประเทศ และในส่วนของปัจจัยที่มีผลกระทบต่อปริมาณการนำเข้าน้ำมันดีเซลของประเทศไทย คือ จำนวนรถยนต์ที่ใช้เครื่องดีเซล ปริมาณน้ำมันดีเซลที่ผลิตได้ภายในประเทศ และแหล่งนำเข้าน้ำมันดิบที่สำคัญคือ กลุ่มประเทศตะวันออกกลาง ส่วนแหล่งนำเข้าผลิตภัณฑ์น้ำมันคือ กลุ่มประเทศอาเซียน โดยเฉพาะประเทศสิงค์โปร์

ปิยสวัสดิ์ (2538) ศึกษาเรื่องนโยบายพลังงานและการพัฒนาเศรษฐกิจไทยไว้ว่า ในเรื่องของนโยบายพลังงานมี 2 เรื่องที่สำคัญ คือ การจัดหาพลังงานให้เพียงพอต่อความต้องการใช้ภายในประเทศ และ เรื่องการใช้พลังงานอย่างมีประสิทธิภาพ และในเรื่องของการจัดหายังต้องมองในเรื่องของราคาที่จัดหาว่าเหมาะสมมากน้อยเพียงใด ตลอดจนเรื่องของการใช้พลังงานจะต้องมีมาตรการหลัก คือ มาตรการด้านราคา ต้องเป็นราคาที่ถูก มิฉะนั้น ผู้ใช้จะไม่สามารถใช้ พลังงานได้อย่างมีประสิทธิภาพ ในด้านโครงสร้างราคาน้ำมัน พบว่ายังมีการบิดเบือนสูงจากการที่ใช้แบบจำลอง cross-subsidize พบว่า ราคาน้ำมันชนิดต่าง ๆ มีปัญหามาก จึงต้องมีการนำหลักเศรษฐศาสตร์เข้ามาแก้ปัญหา โดยในขั้นแรกต้องมองว่าการบิดเบือนราคาเกิดจากอะไร ได้มีการประมาณฟังก์ชั่น อุปสงค์น้ำมัน เพื่อคำนวณหาค่าความยืดหยุ่น เพื่อการทดแทนกันระหว่างน้ำมันชนิดต่างๆ พบว่ามีค่าสูงมากและกลไกตลาดและการแข่งขันเป็นเรื่องสำคัญ และจากประสบการณ์ของปิยสวัสดิ์เอง พบว่า ระบบการแข่งขันแบบเสรีและกลไกตลาดจะเป็นตัวที่นำไปสู่จุดที่ทำให้เกิดประสิทธิภาพสูงสุด การที่รัฐเข้าไปควบคุมในหลาย ๆ เรื่องจะทำให้เกิดผลเสียเหมือนในอดีต

ในแง่ของการใช้พลังงาน พบว่า ราคามีบทบาทค่อนข้างชัดเจน กล่าวคือ ถ้าราคาสูงปริมาณการใช้จะลดลงอย่างชัดเจน และเป็นผลให้มีการเปลี่ยนแปลงเครื่องยนต์จากเบนซินไปเป็นเครื่องดีเซล ทั้ง ๆ ที่ราคาน้ำมันดีเซลซึ่งรวมภาษี ในอดีตสูงกว่าน้ำมันเบนซิน แต่ราคาขายปลีกกลับต่ำกว่าราคาน้ำมันเบนซิน นอกจากนี้ยังกล่าวอีกว่า ราคาในช่วงปี 2536 เป็นไปตามต้นทุนค่อนข้างมาก และการยกเลิกการควบคุมราคาน้ำมันเป็นมาตรการที่ทำให้ระดับราคาน้ำมันชนิดต่าง ๆ สะท้อนถึงต้นทุนทางเศรษฐศาสตร์ ซึ่งการยกเลิกการควบคุมราคาน้ำมันนั้นเป็นผลดีต่อรัฐบาล คือทำให้สามารถแยกการตัดสินใจทางการเมืองออกจากการกำหนดราคาน้ำมันได้

นอกจากนั้นยังพบว่าราคาน้ำมันภายในประเทศปรับตัวขึ้นลงตามราคาตลาดโลก แต่ไม่สามารถปรับราคาได้ทุก ๆ วัน เนื่องจากประเทศเรามีการนำเข้าน้ำมันจากประเทศสิงคโปร์เป็นสำคัญ แต่ราคาขายที่ประเทศสิงคโปร์มีการเปลี่ยนแปลงทุกชั่วโมง จึงทำให้ผู้ค้าภายในประเทศไม่สามารถเปลี่ยนแปลงตามราคาขายที่สิงคโปร์ได้ จึงเป็นผลให้ราคาขายปลีกนานครั้งถึงจะมีการเปลี่ยนแปลง

ในตอนสุดท้ายเป็นการกล่าวถึงเรื่องของพลังงานว่า เป็นสาขาที่น่าสนใจสำหรับนักเศรษฐศาสตร์มาก เนื่องจากมีบางสิ่งบางอย่างที่ผิดไปจากทฤษฎี เช่น การควบคุมราคา การควบคุมผู้ค้าน้ำมัน นอกจากนั้นแล้วในแง่ของการแข่งขัน การที่จะมีการส่งเสริมการแข่งขัน

เนื่องจากตราบใดที่ยังคงมีการผูกขาดในการผลิตและการจัดหาพลังงาน การแข่งขันก็จะยังไม่เกิดขึ้น และเป็นผลให้การดำเนินงานเป็นไปอย่างไม่มีประสิทธิภาพ

สุภิญญา (2538) ศึกษาเรื่องการประเมินผลนโยบายลอยตัวเสรีราคาน้ำมันว่านโยบายลอยตัวน้ำมันนี้ประสบผลสำเร็จตามวัตถุประสงค์ของรัฐบาล ผลจากการศึกษาพบว่านโยบายดังกล่าว บรรลุตามวัตถุประสงค์ของรัฐบาล กล่าวคือ การเปลี่ยนแปลงของราคาน้ำมันขายปลีกไม่ขึ้นอยู่กับการตัดสินใจของรัฐ บริษัทน้ำมันสามารถกำหนดราคาขายได้ด้วยตนเอง มีความฉับไวในการเปลี่ยนแปลงราคาให้ทันกับสถานการณ์โลก ราคาน้ำมันจะปรับขึ้นลงตามราคาตลาดโลก แต่ยังคงมีความล่าช้าอยู่ประมาณ 3 อาทิตย์

สภาพการแข่งขันในตลาดน้ำมันเชื้อเพลิงของประเทศพบว่า ภายหลังที่ปล่อยให้น้ำมัน มีราคาที่ลอยตัว สภาพการแข่งขันเปลี่ยนแปลงไปอย่างมาก เมื่อเทียบกับช่วงที่รัฐบาลควบคุมราคา กล่าวคือ มีผู้ค้าน้ำมันรายใหม่ ๆ เข้าสู่ตลาดเพิ่มมากขึ้น ซึ่งเข้ามาแย่งส่วนแบ่งการตลาดจากผู้ค้า รายเก่ามากขึ้น ได้ทำการศึกษาโดยใช้ Size Ratio พบว่า ความแตกต่างในขนาดของบริษัทผู้ค้า รายใหญ่และผู้ค้ารายย่อยเริ่มลดลง แสดงว่าโอกาสที่ผู้ค้ารายเล็กจะเข้ามาแข่งขันมีมากขึ้น

ในตอนสุดท้ายพบว่ารัฐบาลได้มีมาตรการในการแก้ไขปัญหาต่าง ๆ เพื่อส่งเสริมให้มีการแข่งขันมากขึ้น เช่น การเปิดสถานีบริการขนาดเล็ก ตลอดจนยกเลิกการกำหนดโควต้าการนำเข้าน้ำมันสำเร็จรูป เป็นผลให้ผู้ค้ารายใหม่ ๆ สามารถเข้าสู่ตลาดได้มากขึ้น นอกจากนั้นแล้ว การแข่งขันที่เป็นการแข่งขันด้านอื่น ๆ ที่ไม่ใช่ราคา เช่น การโฆษณาความแตกต่างของผลิตภัณฑ์ กล่าวโดยรวมได้ว่า สภาพตลาดน้ำมันภายหลังใช้นโยบายราคาน้ำมันลอยตัวพบว่ามีการแข่งขันเพิ่มมากขึ้น

พีรณัฐ (2540) ได้ทำการศึกษาเรื่องการวิเคราะห์การกำหนดราคาน้ำมันในระดับค้าปลีก มีจุดประสงค์เพื่อศึกษาถึงโครงสร้างราคาขายปลีกน้ำมัน ว่าประกอบด้วยตัวแปรอะไรบ้าง และตัวแปรใดมีความสำคัญมากที่สุด เพื่อศึกษาการเปลี่ยนแปลงราคาขายปลีกมีความสอดคล้องกับต้นทุนมากน้อยเพียงใด ซึ่งได้ศึกษาถึงน้ำมัน 3 ชนิดที่มีการใช้อย่างแพร่หลาย ได้แก่ น้ำมันเบนซินพิเศษไร้สารตะกั่ว น้ำมันเบนซินพิเศษสูตรทดแทนสารตะกั่ว และน้ำมันดีเซลหมุนเร็ว

จากการศึกษาพบว่า ในโครงสร้างราคาน้ำมันในปัจจุบัน ตัวแปรที่สามารถกำหนดราคาขายปลีก คือ ราคา ณ โรงกลั่น หรือราคานำเข้าอย่างเดียว แต่ตัวแปรด้านอื่น ๆ เช่น ภาษีต่าง ๆ ที่รัฐเป็นผู้กำหนด ไม่มีส่วนสำคัญทางเชิงสถิติในการกำหนดโครงสร้างแต่อย่างใด แสดงถึงโอกาสที่ผู้ค้าน้ำมันไม่สามารถผลักภาระไปให้ผู้บริโภค

ศุภกร (2541) ศึกษาเรื่องโครงสร้างทางการค้า และปัจจัยที่มีผลกระทบต่อราคาน้ำมันดีเซลในประเทศไทยโดยศึกษาถึงปัจจัยสำคัญที่มีผลกระทบต่อระดับราคาขายปลีกน้ำมันดีเซล ข้อมูลที่ใช้ทำการศึกษาเป็นข้อมูลทุติยภูมิ (Secondary Data) ซึ่งเป็นอนุกรมเวลา ข้อมูลเกี่ยวกับโครงสร้างทางการค้า ทำการจัดเก็บเป็นรายปี ตั้งแต่ปี พ.ศ. 2535 - 2539 รวม 5 ปี และข้อมูลเกี่ยวกับราคาน้ำมันดีเซลทำการจัดเก็บเป็นรายเดือน ตั้งแต่เดือนมกราคม พ.ศ. 2535 - เดือนธันวาคม พ.ศ. 2539 รวมทั้งสิ้น 60 เดือน แบ่งการวิเคราะห์ออกเป็น 2 ส่วน คือ การวิเคราะห์เชิงพรรณา (Descriptive Method) และการวิเคราะห์เชิงปริมาณ โดยใช้สมการถดถอยเชิงซ้อน

การศึกษาโครงสร้างทางการค้าน้ำมันดีเซล จะทำการศึกษาครอบคลุมในด้านการจัดหาการจัดจำหน่ายน้ำมันดีเซล รวมถึงการจัดหาน้ำมันดิบที่ใช้เป็นวัตถุดิบในการผลิต ส่วนการศึกษาราคาน้ำมันดีเซลได้ทำการศึกษาครอบคลุมถึงระบบราคาน้ำมันลอยตัว การกำหนดราคา บทบาทของรัฐในการกำหนดราคา และปัจจัยสำคัญที่มีผลกระทบต่อราคาน้ำมันดีเซลในประเทศไทย ศึกษาเฉพาะปัจจัยทางด้านอุปทาน

จากการศึกษาพบว่า การค้าน้ำมันดีเซลในประเทศไทยมีการขยายตัวอย่างมากตามปริมาณความต้องการใช้ที่เพิ่มขึ้น โครงสร้างทางการค้าของน้ำมันดีเซลได้มีการเปลี่ยนแปลงจากการนำเข้ามาเป็นการผลิตภายในประเทศมากขึ้น ทำให้สัดส่วนการผลิตต่อการนำเข้าของน้ำมันดีเซลมีแนวโน้มสูงขึ้น ในส่วนของการจำหน่ายภายในประเทศก็มีปริมาณเพิ่มขึ้นและมีการส่งออกไปยังประเทศอีกด้วย รัฐบาลไม่ได้มีการกำหนดราคาน้ำมันเชื้อเพลิงภายในประเทศอีกต่อไป ทำให้ราคาขายปลีกน้ำมันดีเซลภายในประเทศมีการเปลี่ยนแปลงอยู่ตลอดเวลา โดยปัจจัยสำคัญที่มีผลกระทบต่อราคาขายปลีกน้ำมันดีเซลในประเทศ ได้แก่ ราคาน้ำมันดีเซลในตลาดโลก และอัตราแลกเปลี่ยนใน 1 เดือนที่ผ่านมา

สาธิต (2541) ศึกษาปัจจัยที่มีอิทธิพลต่อการนำเข้าน้ำมันดีเซลและน้ำมันเบนซินของประเทศไทย ตั้งแต่ปี พ.ศ. 2515 – 2539 รวม 26 ปี แบ่งการศึกษาออกเป็น 2 ส่วน ในส่วนที่ 1 ศึกษาปัจจัยที่มีอิทธิพลต่อการนำเข้าน้ำมันดีเซลและน้ำมันเบนซินจากแบบจำลองทางเศรษฐมิติ โดยปัจจัยที่มีอิทธิพลต่อการนำเข้าประกอบด้วย ราคานำเข้าของน้ำมัน ผลิตภัณฑ์มวลรวมภายในประเทศ อัตราภาษีเฉลี่ยของน้ำมัน ปริมาณรถยนต์ทุกประเภทภายในประเทศ จำนวนประชากรภายในประเทศ และในการศึกษาน้ำมันดีเซล พิจารณาปริมาณเรือภายในประเทศเพิ่มเติมอีกปัจจัยหนึ่ง ในการหาความสัมพันธ์ จะใช้วิธีกำลังสองน้อยที่สุดแบบธรรมดา (OLS) โดยสมการมีความสัมพันธ์อยู่ในรูปของ Static Log Linear model ส่วนที่ 2 การพยากรณ์อุปสงค์การนำเข้าน้ำมันดีเซล

ผลจากการศึกษาส่วนที่ 1 ในการศึกษาน้ำมันดีเซล พบว่าปัจจัยที่มีอิทธิพลต่อการนำเข้าน้ำมันดีเซลอย่างมีนัยสำคัญทางสถิติ ได้แก่ ราคานำเข้าน้ำมันดีเซล ผลิตภัณฑ์มวลรวมภายในประเทศ ปริมาณรถยนต์ภายในประเทศ และอัตราภาษีเฉลี่ยของน้ำมันดีเซล ส่วนปริมาณเรือที่ใช้ น้ำมันดีเซลและจำนวนประชากรภายในประเทศ พบว่าไม่มีนัยสำคัญทางสถิติ ส่วนในการศึกษาปัจจัยที่มีอิทธิพลต่อการนำเข้าน้ำมันเบนซิน พบว่าราคานำเข้าน้ำมันเบนซิน ผลิตภัณฑ์มวลรวมภายในประเทศ ปริมาณรถยนต์ภายในประเทศ และอัตราภาษีเฉลี่ยน้ำมันเบนซิน และพบว่าราคานำเข้ามีผลต่ออุปสงค์การนำเข้ามากที่สุด

ผลการศึกษาส่วนที่ 2 พบว่า การพยากรณ์แนวโน้มความต้องการนำเข้าน้ำมันดีเซลและน้ำมันเบนซิน ในช่วงปี พ.ศ. 2540 ถึงปี พ.ศ. 2545 โดยใช้ปี พ.ศ.2515 ถึงปี พ.ศ. 2539 เป็นปีฐาน ด้วยวิธีหาค่าเฉลี่ย มีอัตราการเพิ่มความต้องการนำเข้าน้ำมันดีเซลและนำมันเบนซิน ร้อยละ 10 และ 3 ต่อปี ตามลำดับ และการพยากรณ์ด้วยสมการถดถอย มีอัตราการเพิ่มความต้องการนำเข้า น้ำมันดีเซลและน้ำมันเบนซินร้อยละ 10 และ 2 ต่อปี ตามลำดับ

ศราวุธ (2546) ได้ศึกษาเรื่องการวิเคราะห์ปัจจัยที่กำหนดปริมาณการผลิตก๊าซแอลพีจีในประเทศไทย ได้อาศัยหลักการทางสถิติเป็นเครื่องมือ และทำการสร้างแบบจำลองอุปสงค์ด้วยสมการพหุคูณเชิงเส้นโดยการคัดเลือกตัวแปรที่สำคัญ คือมูลค่าผลิตภัณฑ์มวลรวมภายในประเทศเบื้องต้น ราคาขายถัวเฉลี่ยของน้ำมันเตา ราคาขายถัวเฉลี่ยของก๊าซแอลพีจี ปริมาณการใช้ก๊าซแอลพีจีไทยในประเทศ ราคาขายถัวเฉลี่ยของน้ำมันเบนซินและ Dummy ช่วงระยะเวลาการใช้นโยบายต่าง ๆ ของรัฐบาล เพื่อให้ได้สมการที่เหมาะสม ซึ่งสามารถนำมาอธิบายความสัมพันธ์ระหว่าง

ปริมาณการผลิตก๊าซแอลพีจีกับปัจจัยต่าง ๆ ที่กำหนดปริมาณการผลิตก๊าซแอลพีจีได้ พร้อมทั้งทำการพยากรณ์แนวโน้มปริมาณการผลิตก๊าซแอลพีจีช่วงปีพ.ศ. 2546-2550

จากสมการเมื่อพิจารณาค่าความยืดหยุ่นของตัวแปรแต่ละตัว โดยพิจารณาจากค่าของสัมประสิทธิ์ที่อยู่หน้าตัวแปรแต่ละตัว พบว่าค่าความยืดหยุ่นของมูลค่าผลิตภัณฑ์มวลรวมภายในประเทศเบื้องต้น ราคาขายถัวเฉลี่ยของน้ำมันเตา ราคาขายถัวเฉลี่ยของก๊าซแอลพีจี ปริมาณการใช้ก๊าซแอลพีจีไทยในประเทศ ราคาขายถัวเฉลี่ยของน้ำมันเบนซินและ Dummy ช่วงระยะเวลาการใช้นโยบายต่าง ๆ ของรัฐบาลมีค่าเท่ากับ 0.223, 2.359, 0.063, 0.793, 2.991,0.430 ตามลำดับ

สำหรับการพยากรณ์แนวโน้มปริมาณการผลิตก๊าซแอลพีจีช่วงปี พ.ศ. 2546-2550 โดยใช้สูตรอัตราดอกเบี้ยทบต้น จากการคำนวณปรากฏว่า มูลค่าผลิตภัณฑ์มวลรวมภายในประเทศเบื้องต้น ราคาขายถัวเฉลี่ยของน้ำมันเตา ราคาขายถัวเฉลี่ยของก๊าซแอลพีจี ปริมาณการใช้ก๊าซแอลพีจีไทยในประเทศ ราคาขายถัวเฉลี่ยของน้ำมันเบนซิน และปริมาณการผลิตก๊าซแอลพีจีมีแนวโน้มเพิ่มขึ้นในอัตรา 3.3%, 6.8%, 2.7%, 7.3%, 6.6%, 5.2% และ 7.7% ตามลำดับ

นุชชฎา (2547) ได้ทำการศึกษาเรื่องโครงสร้างและการกำหนดราคาค้าปลีกน้ำมันเชื้อเพลิง โดยมีวัตถุประสงค์เพื่อศึกษาโครงสร้างราคาค้าปลีกน้ำมัน โดยวิเคราะห์ปัจจัยต่าง ๆ ที่มีผลกระทบต่อราคาค้าปลีกน้ำมัน ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิรายไตรมาส ตั้งแต่ปี 2536-2546 ซึ่งได้ศึกษาถึงน้ำมัน 3 ชนิดที่มีการใช้อย่างแพร่หลาย ได้แก่ น้ำมันเบนซินพิเศษไร้สารตะกั่ว น้ำมัน-เบนซินพิเศษสูตรทดแทนสารตะกั่ว และน้ำมันดีเซลหมุนเร็ว ทำการวิเคราะห์โดยใช้วิธีกำลังสองน้อยที่สุด สมการโครงสร้างราคาเป็นแบบ Price Linkage Model

จากผลการศึกษาพบว่าในโครงสร้างราคาค้าปลีกน้ำมันเชื้อเพลิงในปัจจุบันตัวแปรทุกตัวในสมการโครงสร้างราคาสามารถกำหนดราคาค้าปลีกได้อย่างมีนัยสำคัญ ซึ่งประกอบด้วย ราคา ณ โรงกลั่น ราคานำเข้า ภาษีสรรพสามิต กองทุนน้ำมันเชื้อเพลิง กองทุนเพื่อการส่งเสริมการอนุรักษ์พลังงาน โดยที่ราคา ณ โรงกลั่นและราคานำเข้า จะมีปัจจัยอื่น ๆ ที่มากำหนดด้วย นั่นคือ ราคาน้ำมันดิบ ราคาน้ำมันสำเร็จรูปที่ตลาดจรสิงคโปร์ อัตราแลกเปลี่ยนและภาษีนำเข้า ซึ่งทุกปัจจัยดังกล่าวมีความสัมพันธ์แบบมีนัยสำคัญ

Ashakul (1985) ศึกษาอุปสงค์ของผลิตภัณฑ์น้ำมันในส่วนของน้ำมันเชื้อเพลิงที่ใช้ในการคมนาคมขนส่ง 4 ชนิด ประกอบด้วยน้ำมันเบนซินพิเศษ น้ำมันเบนซินธรรมดา น้ำมันดีเซล และก๊าซปิโตรเลียมเหลว (LPG) ในการศึกษาได้พิจารณาความยืดหยุ่นต่อราคาและความ ยืดหยุ่นต่อสินค้าทดแทน ข้อมูลที่ใช้ศึกษาเป็นรายไตรมาสในช่วงปี ค.ศ. 1978-1984 จากการศึกษาพบว่าราคาปัจจัยที่มีอิทธิพลต่ออุปสงค์ผลิตภัณฑ์น้ำมัน โดย LPG มีค่าความยืดหยุ่นต่อราคามากที่สุด คือ –2.78 ในขณะที่น้ำมันดีเซลมีค่าความยืดหยุ่นต่อราคาต่ำสุด คือ –0.24 ส่วนค่าความยืดหยุ่นต่อราคาของน้ำมันเบนซินพิเศษและน้ำมันเบนซินธรรมดา มีค่ามากกว่าค่าความ ยืดหยุ่นของน้ำมันดีเซล โดยมีค่าเท่ากับ –1.25 และ –1.27 ตามลำดับ สำหรับคาความยืดหยุ่นของสินค้าทดแทนระหว่างน้ำมันเบนซินพิเศษและน้ำมันเบนซินธรรมดาจะมีค่าความยืดหยุ่นมากที่สุด คือ –3.90 ในขณะที่ความยืดหยุ่นของสินค้าทดแทนระหว่างน้ำมันเบนซินธรรมดาและน้ำมันดีเซลมีค่าต่ำสุดเท่ากับ 0.73

จากการศึกษางานวิจัยที่เกี่ยวข้องข้างต้น ผู้วิจัยจึงใช้ตัวแปรที่จะทำการศึกษาดังนี้ อัตราแลกเปลี่ยนดอลลาสหรัฐ(ซื้อ) อัตราแลกเปลี่ยนดอลลาสหรัฐ(ขาย) ราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น ราคาน้ำมันดีเซล นำเข้าหน้าโรงกลั่น ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น ราคาน้ำมันเบนซิน 91 ขายส่งหน้าโรงกลั่น ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น ปริมาณการผลิตน้ำมันสำเร็จรูป ปริมาณการส่งออกน้ำมันสำเร็จรูป ปริมาณการนำเข้าน้ำมันสำเร็จรูป ปริมาณการจำหน่ายน้ำมันสำเร็จรูป ปริมาณการนำเข้าน้ำมันดิบ ปริมาณการนำเข้าน้ำมันดิบในตะวันออกกลาง ปริมาณการนำเข้าน้ำมันดิบในตะวันออกไกล อัตราภาษีส่งออกน้ำมันเบนซิน 95 อัตราภาษีส่งออกน้ำมันเบนซิน 91 อัตราภาษีส่งออกน้ำมันดีเซล อัตราภาษีนำเข้าน้ำมันเบนซิน 95 อัตราภาษีนำเข้าน้ำมันเบนซิน 91 อัตราภาษีนำเข้าน้ำมันดีเซล ราคาน้ำมันดิบในตลาดโลก (โอมาน) ราคาน้ำมันเบนซิน 95 ในตลาดสิงคโปร์ ราคาน้ำมันเบนซิน 92 ในตลาดสิงคโปร์ ราคาน้ำมันดีเซลในตลาดสิงคโปร์ ปริมาณการใช้ไฟฟ้า ประชากรรวมในกรุงเทพฯ จำนวนผู้มีงานทำในกรุงเทพฯ จำนวนผู้ว่างงานในกรุงเทพฯ ผลิตภัณฑ์มวลรวมในประเทศ อัตราเงินเฟ้อ อัตราดอกเบี้ยเงินกู้ กองทุนน้ำมัน แหล่งกำเนิดไฟฟ้าจากดีเซล ปริมาณการผลิตน้ำมันดีเซล ปริมาณการผลิตน้ำมัน LPG ปริมาณการผลิตน้ำมันเบนซิน 95 ปริมาณการผลิตน้ำมันเบนซิน 91 ปริมาณการใช้น้ำมันเบนซิน 95 ปริมาณการใช้น้ำมันเบนซิน 91 ปริมาณการใช้น้ำมันดีเซล ปริมาณการนำเข้าน้ำมันดีเซล ปริมาณการส่งออกน้ำมันเบนซิน 95 ปริมาณการส่งออกน้ำมันเบนซิน 91 ปริมาณการส่งออกน้ำมันดีเซล

ปริมาณการจำหน่ายน้ำมันเบนซิน ปริมาณการจำหน่ายน้ำมันดีเซล จำนวนรถจักรยานยนต์ที่จดทะเบียนในกรุงเทพฯ จำนวนรถยนต์นั่งส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ จำนวนรถโดยสารประจำทางที่ จดทะเบียนในกรุงเทพฯ จำนวนรถโดยสารไม่ประจำทางที่จดทะเบียนในกรุงเทพฯ จำนวนรถโดยสารส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ จำนวนรถบรรทุกไม่ประจำทางที่จดทะเบียนในกรุงเทพฯ จำนวนรถบรรทุกส่วนบุคคลที่จดทะเบียนในกรุงเทพฯ มูลค่าการนำเข้ายานพาหนะและอุปกรณ์ขนส่ง มูลค่าการนำเข้าน้ำมันดิบ มูลค่าการนำเข้าน้ำเบนซิน มูลค่าการนำเข้าน้ำมันดีเซล มูลค่าการนำเข้าก๊าซธรรมชาติ มูลค่าการนำเข้าก๊าซปิโตรเลียม มูลค่าการส่งออกรถยนต์ รายได้ภาษีการนำเข้าน้ำมัน ปริมาณการจำหน่ายรถยนต์ในประเทศ ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศ มูลค่าการส่งออกผลิตภัณฑ์ปิโตรเลียม ปริมาณการผลิตรถยนต์ ปริมาณการผลิตรถจักรยานยนต์ ปริมาณการผลิตน้ำมันดิบ รายได้ภาษีน้ำมันเบนซินในประเทศ รายได้ภาษีน้ำมันดีเซลในประเทศ

อุปกรณ์และวิธีการ

อุปกรณ์

1. เครื่องไมโครคอมพิวเตอร์ ที่ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

2. วิเคราะห์ข้อมูลโดยใช้โปรแกรม SAS Version 9.1

วิธีการ

การเก็บรวบรวมข้อมูล

ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิ (Secondary Data) เป็นข้อมูลเกี่ยวกับราคาน้ำมันขายปลีกในเขตกรุงเทพมหานคร เช่น อัตราดอกเบี้ย อัตราภาษี ปริมาณการนำเข้าน้ำมัน ราคา ณ โรงกลั่นน ปริมาณการผลิตน้ำมันในประเทศ ฯลฯ ซึ่งทำการจัดเก็บรวบรวมข้อมูลรายเดือน ตั้งแต่เดือนมกราคม 2545 - เดือนธันวาคม 2549 รวมระยะเวลา 5 ปี ซึ่งทำการรวบรวมข้อมูลจากหน่วยงานต่าง ๆ ได้แก่ บริษัท ปตท. จำกัดมหาชน กรมธุรกิจพลังงาน ธนาคารแห่งประเทศไทย กรมสรรพสามิต กระทรวงพลังงาน และสำนักงานสถิติแห่งชาติ

การวิเคราะห์ข้อมูล

1. เก็บรวบรวมข้อมูลราคาน้ำมันขายปลีกในเขตกรุงเทพมหานคร และปัจจัยต่าง ๆ ที่คาดว่าจะมีผลต่อการเปลี่ยนแปลงของราคาน้ำมันขายปลีก ซึ่งในการศึกษานี้ทำการเก็บขอมูลของตัวแปรตาม 3 ตัวแปร คือ ราคาน้ำมันดีเซล ราคาน้ำมันเบนซิน 91 และราคาน้ำมันเบนซิน 95 ส่วนตัวแปรอิสระ 74 ตัวแปร เช่น ราคาน้ำมันหน้าโรงกลั่น ราคาน้ำมันในตลาดโลก อัตราภาษี เป็นต้น

2. ทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระทั้ง 74 ตัวแปร กับตัวแปรตามทั้ง 3 ตัวแปร โดยการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล

3. ตรวจสอบข้อสมมติเบื้องต้นของข้อมูล ว่าสอดคล้องกับข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอยแบบหลายตัวแปรหรือไม่ โดยการตรวจสอบการแจกแจงปกติหลายตัวแปรของความคลาดเคลื่อน โดยวิธีการของ Mardia และตรวจสอบความคงที่ของความแปรปรวนของค่าคลาดเคลื่อน โดยพิจารณาจากกราฟ

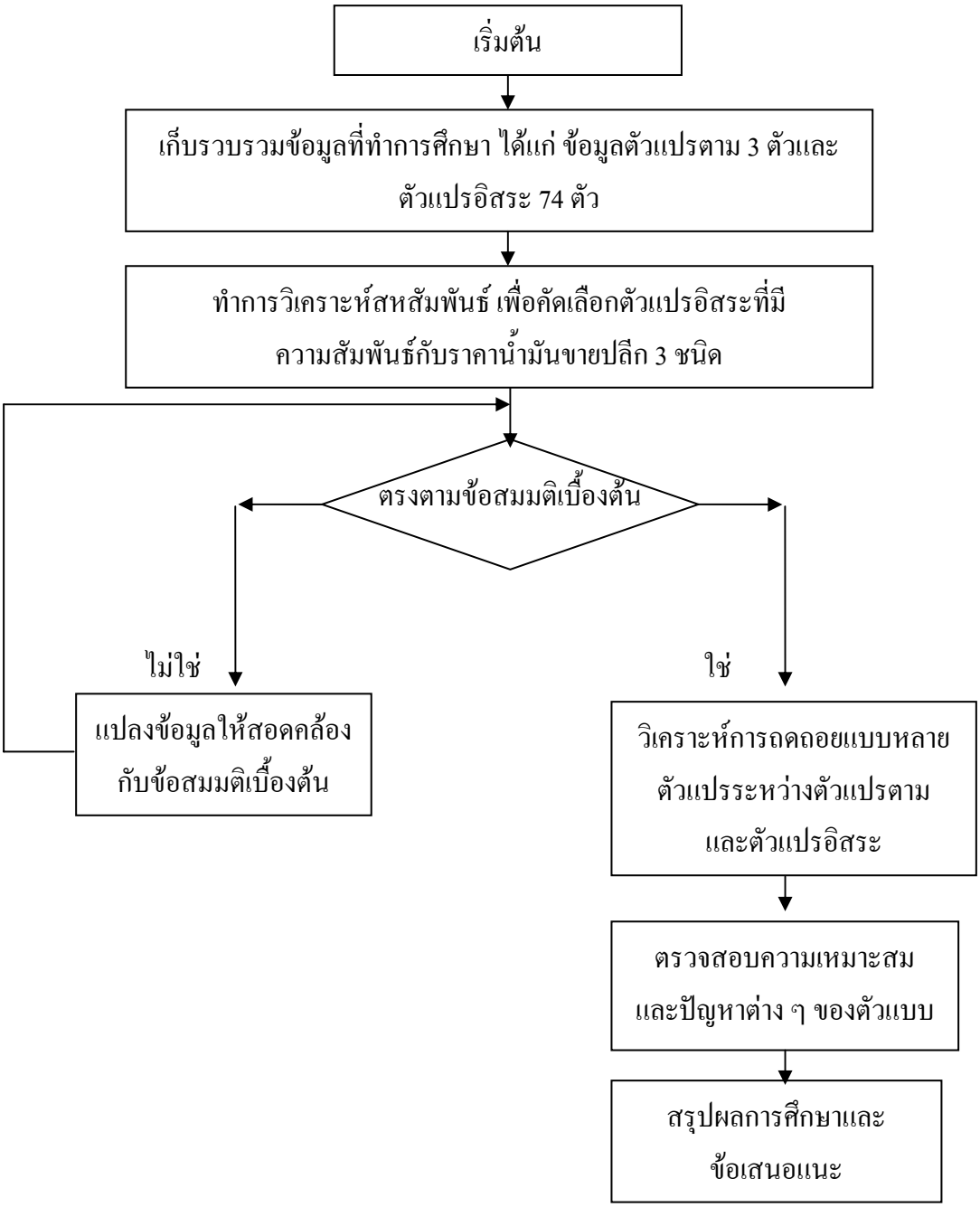
4. ถ้าตรวจสอบแล้วไม่ตรงตามข้อสมมติเบื้องต้น จะทำการแปลงข้อมูลให้สอดคล้องกับข้อสมมติเบื้องต้น

5. นำข้อมูลที่สอดคล้องกับข้อสมมติเบื้องต้นนั้นไปวิเคราะห์การถดถอยแบบหลายตัวแปร

6. เมื่อทำการวิเคราะห์การถดถอยหลายตัวแปรแล้ว จึงทำการตรวจสอบความเหมาะสมของรูปแบบความสัมพันธ์ที่ได้

7. สรุปผลที่ได้จากการวิเคราะห์และข้อเสนอแนะที่ได้จากการศึกษาวิจัยในครั้งนี้

จากวิธีการดังกล่าวข้างต้นสามารถสรุปแผนผังของการศึกษาวิจัยได้ดังนี้



ภาพที่ 2 แผนผังของการศึกษาวิจัย

สถานที่ทำการวิจัย

ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตบางเขน

ผลและวิจารณ์

ผล

งานวิจัยนี้ได้ทำการศึกษารูปแบบความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภท กับปัจจัยต่าง ๆ ที่คาดว่าจะมีผลต่อราคาน้ำมันขายปลีก โดยพิจารณาจากตัวแปรอิสระทั้ง 74 ตัว ที่นำมาทำการศึกษาในงานวิจัยนี้ โดยอาศัยเทคนิคการวิเคราะห์ทางสถิติต่าง ๆ ได้แก่ การวิเคราะห์สหสัมพันธ์ การวิเคราะห์สหสัมพันธ์แคนนอลนิคอล การวิเคราะห์ปัจจัย และการวิเคราะห์การถดถอยแบบหลายตัวแปร ผลการศึกษาแบ่งเป็น 7 ส่วนใหญ่ ๆ คือ

1. ผลการวิเคราะห์ข้อมูลเบื้องต้น
2. ผลการวิเคราะห์สหสัมพันธ์
3. ผลการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล
4. ผลการตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอย
5. ผลการวิเคราะห์ปัจจัย
6. ผลการวิเคราะห์การถดถอยแบบหลายตัวแปร
7. การนำสมการถดถอยที่ได้จากการวิจัยไปใช้ในการพยากรณ์

1. ผลการวิเคราะห์ข้อมูลเบื้องต้น

ตารางที่ 2 ข้อมูลเบื้องต้นของตัวแปรตาม (Y_1, Y_2, Y_3)

ตัวแปร	N	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ค่าต่ำสุด	ค่าสูงสุด
Y_1	60	17.547	5.12	11.48	27.79
Y_2	60	19.619	5.01	12.56	29.27
Y_3	60	20.488	4.94	13.56	30.07

จากตารางที่ 2 เป็นผลการวิเคราะห์ข้อมูลเบื้องต้นของตัวแปรตามทั้ง 3 ตัวแปร ซึ่งพบว่า ค่าเฉลี่ยของราคาน้ำมันดีเซล (Y_1) คือ 17.547 (บาท/ลิตร) ค่าเฉลี่ยของราคาน้ำมันเบนซิน 91 (Y_2) คือ 19.619 (บาท/ลิตร) และ ค่าเฉลี่ยของราคาน้ำมันเบนซิน 95 (Y_3) คือ 20.488 (บาท/ลิตร)

2. ผลการวิเคราะห์สหสัมพันธ์

ศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตามว่ามีความสัมพันธ์กันหรือไม่ โดยใช้การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์ สามารถคำนวณได้ดังตารางต่อไปนี้

ตารางที่ 3 ความสัมพันธ์ระหว่างตัวแปรอิสระ 74 ตัว กับตัวแปรตาม 3 ตัว

	Y_1	Y_2	Y_3
X_1	-0.224 (0.125)	-0.385 **(0.007)**	-0.372 **(0.009)**
X_2	-0.224 (0.125)	-0.386 **(0.007)**	-0.373 **(0.009)**
X_3	0.883 **(0.000)**	0.974 **(0.000)**	0.973 **(0.000)**
X_4	0.883 **(0.000)**	0.973 **(0.000)**	0.971 **(0.000)**
X_5	0.891 **(0.000)**	0.986 **(0.000)**	0.945 **(0.000)**
X_6	0.865 **(0.000)**	0.970 **(0.000)**	0.967 **(0.000)**
X_7	0.872 **(0.000)**	0.972 **(0.000)**	0.970 **(0.000)**
X_8	0.893 **(0.000)**	0.986 **(0.000)**	0.986 **(0.000)**
X_9	0.915 **(0.000)**	0.993 **(0.000)**	0.993 **(0.000)**
X_{10}	0.914 **(0.000)**	0.984 **(0.000)**	0.985 **(0.000)**
X_{11}	0.990 **(0.000)**	0.926 **(0.000)**	0.927 **(0.000)**
X_{12}	0.426 **(0.003)**	0.575 **(0.000)**	0.570 **(0.000)**
X_{13}	0.296 (0.410)	0.339 **(0.018)**	0.342 **(0.017)**
X_{14}	0.238 (0.103)	0.167 (0.257)	0.160 (0.278)
X_{15}	0.449 **(0.001)**	0.614 **(0.000)**	0.603 **(0.000)**
X_{16}	0.218 (0.136)	0.381 **(0.008)**	0.376 **(0.008)**
X_{17}	0.166 (0.260)	0.333 **(0.021)**	0.329 **(0.023)**
X_{18}	-0.393 **(0.006)**	-0.448 **(0.001)**	-0.445 **(0.002)**

ตารางที่ 3 (ต่อ)

	Y_1	Y_2	Y_3
X_{19}	0.899 **(0.000)**	0.774 **(0.000)**	0.778 **(0.000)**
X_{20}	0.890 **(0.000)**	0.769 **(0.000)**	0.773 **(0.000)**
X_{21}	0.485 **(0.000)**	0.403 **(0.005)**	0.405 **(0.004)**
X_{22}	0.899 **(0.000)**	0.774 **(0.000)**	0.778 **(0.000)**
X_{23}	0.814 **(0.000)**	0.697 **(0.000)**	0.701 **(0.000)**
X_{24}	0.485 **(0.000)**	0.403 **(0.005)**	0.405 **(0.004)**
X_{25}	0.926 **(0.000)**	0.973 **(0.000)**	0.972 **(0.000)**
X_{26}	0.921 **(0.000)**	0.977 **(0.000)**	0.976 **(0.000)**
X_{27}	0.854 **(0.000)**	0.961 **(0.000)**	0.958 **(0.000)**
X_{28}	0.851 **(0.000)**	0.960 **(0.000)**	0.957 **(0.000)**
X_{29}	0.872 **(0.000)**	0.980 **(0.000)**	0.978 **(0.000)**
X_{30}	0.626 **(0.000)**	0.749 **(0.000)**	0.744 **(0.000)**
X_{31}	-0.700 **(0.000)**	-0.573 **(0.000)**	0.580 **(0.000)**
X_{32}	-0.504 **(0.000)**	-0.382 **(0.007)**	-0.390 **(0.006)**
X_{33}	-0.496 **(0.000)**	-0.560 **(0.000)**	-0.559 **(0.000)**
X_{34}	0.675 **(0.000)**	0.803 **(0.000)**	0.796 **(0.000)**
X_{35}	0.953 **(0.000)**	0.988 **(0.000)**	0.988 **(0.000)**
X_{36}	0.801 **(0.000)**	0.774 **(0.000)**	0.779 **(0.000)**
X_{37}	0.863 **(0.000)**	0.926 **(0.000)**	0.927 **(0.000)**
X_{38}	-0.239 (0.102)	-0.332 **(0.021)**	-0.328 **(0.023)**
X_{39}	0.344 **(0.017)**	0.498 **(0.000)**	0.491 **(0.000)**
X_{40}	0.686 **(0.000)**	0.800 **(0.000)**	0.799 **(0.000)**
X_{41}	0.472 **(0.001)**	0.524 **(0.000)**	0.523 **(0.000)**
X_{42}	0.377 **(0.008)**	0.495 **(0.000)**	0.493 **(0.000)**
X_{43}	-0.243 (0.096)	-0.305 **(0.035)**	-0.313 **(0.030)**
X_{44}	-0.276 (0.058)	-0.231 (0.115)	-0.240 (0.100)
X_{45}	0.298 **(0.040)**	0.513 **(0.000)**	0.502 **(0.000)**

ตารางที่ 3 (ต่อ)

	Y_1	Y_2	Y_3
X_{46}	-0.155 (0.292)	-0.184 (0.211)	-0.188 (0.200)
X_{47}	0.715 **(0.000)**	0.740 **(0.000)**	0.743 **(0.000)**
X_{48}	0.758 **(0.000)**	0.824 **(0.000)**	0.825 **(0.000)**
X_{49}	-0.006 (0.966)	-0.066 (0.656)	-0.066 (0.657)
X_{50}	-0.244 (0.094)	-0.305 **(0.035)**	-0.314 **(0.030)**
X_{51}	0.257 (0.078)	0.479 **(0.001)**	0.468 **(0.001)**
X_{52}	0.771 **(0.000)**	0.843 **(0.000)**	0.840 **(0.000)**
X_{53}	0.165 (0.251)	0.207 (0.158)	0.202 (0.168)
X_{54}	-0.113 (0.444)	-0.150 (0.307)	-0.146 (0.323)
X_{55}	0.101 (0.494)	0.135 (0.361)	0.126 (0.392)
X_{56}	-0.036 **(0.008)**	-0.027 (0.853)	-0.031 (0.835)
X_{57}	0.118 (0.424)	0.370 **(0.010)**	0.360 **(0.012)**
X_{58}	-0.007 (0.961)	0.121 (0.411)	0.116 (0.433)
X_{59}	0.492 **(0.000)**	0.428 **(0.002)**	0.426 **(0.003)**
X_{60}	0.826 **(0.000)**	0.857 **(0.000)**	0.855 **(0.000)**
X_{61}	0.370 **(0.010)**	0.407 **(0.004)**	0.403 **(0.005)**
X_{62}	0.132 (0.372)	0.155 (0.293)	0.154 (0.297)
X_{63}	0.378 **(0.008)**	0.459 **(0.001)**	0.458 **(0.001)**
X_{64}	0.139 (0.346)	0.237 (0.105)	0.232 (0.112)
X_{65}	0.932 **(0.000)**	0.923 **(0.000)**	0.924 **(0.000)**
X_{66}	0.135 (0.360)	0.305 **(0.035)**	0.296 **(0.041)**
X_{67}	0.288 **(0.047)**	0.383 **(0.007)**	0.375 **(0.009)**
X_{68}	0.563 **(0.000)**	0.621 **(0.000)**	0.614 **(0.000)**
X_{69}	0.754 **(0.000)**	0.846 **(0.000)**	0.845 **(0.000)**
X_{70}	0.476 **(0.001)**	0.608 **(0.000)**	0.599 **(0.000)**
X_{71}	0.656 **(0.000)**	0.735 **(0.000)**	0.728 **(0.000)**
X_{72}	0.836 **(0.000)**	0.774 **(0.000)**	0.778 **(0.000)**

ตารางที่ 3 (ต่อ)

	Y_1	Y_2	Y_3
X_{73}	-0.282 (0.130)	-0.221 (0.131)	-0.226 (0.223)
X_{74}	-0.530 **(0.000)**	-0.355 **(0.013)**	-0.363 **(0.011)**

หมายเหตุ ค่าในวงเล็บคือค่า p-value ที่ระดับนัยสำคัญ 0.05

ผลการวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวแปรกับตัวแปรอิสระแต่ละตัวทั้ง 74 ตัวแปร พบว่าที่ระดับนัยสำคัญ 0.05 มีตัวแปรทั้งหมด 33 ตัวแปร ที่มีความสัมพันธ์กับตัวแปรตามทั้ง 3 ตัวแปร ได้แก่ ตัวแปรราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (X_3), ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (X_4), ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น(X_5), ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (X_6), ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (X_7), ราคาน้ำมันดีเซลนำเข้าหน้าโรงกลั่น (X_8), ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (X_9), ราคาน้ำมันเบนซิน 91 ขายส่งหน้าโรงกลั่น (X_{10}), ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น (X_{11}), อัตราภาษีส่งออกน้ำมันเบนซิน 95 (X_{19}), อัตราภาษีส่งออกน้ำมันเบนซิน 91 (X_{20}), อัตราภาษีนำเข้าน้ำมันเบนซิน 95 (X_{22}), อัตราภาษีนำเข้าน้ำมันเบนซิน 91 (X_{23}), ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (X_{25}), ราคาน้ำมันดิบในตลาดโลก (โอมาน) (X_{26}), ราคาน้ำมันเบนซิน 95 ในตลาดจรสิงคโปร์ (X_{27}), ราคาน้ำมันเบนซิน 92 ในตลาดจรสิงคโปร์ (X_{28}), ราคาน้ำมันดีเซลในตลาดจรสิงคโปร์ (X_{29}), ปริมาณการใช้ไฟฟ้า (X_{30}), ประชากรรวมในกรุงเทพฯ (X_{31}) , ผลิตภัณฑ์มวลรวมในประเทศ (X_{34}) , อัตราดอกเบี้ยเงินกู้ (X_{36}), กองทุนน้ำมัน (X_{37}), ปริมาณการผลิตน้ำมัน LPG (X_{40}), ปริมาณ การส่งออกน้ำมันเบนซิน 95 (X_{47}), ปริมาณการส่งออกน้ำมันเบนซิน 91 (X_{48}), จำนวนรถจักรยานยนต์ที่จดทะเบียนในกรุงเทพฯ (X_{52}), มูลค่าการนำเข้าน้ำมันดิบ (X_{60}), มูลค่า การส่งออกรถยนต์ (X_{65}), ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศ (X_{68}), มูลค่าการส่งออกผลิตภัณฑ์ปิโตรเลียม (X_{69}), ปริมาณการผลิตรถจักรยานยนต์ (X_{71}), ปริมาณการผลิตน้ำมันดิบ (X_{72})

หลังจากพบว่าตัวแปรอิสระทั้ง 33 ตัวแปรมีความสัมพันธ์กับตัวแปรตามทั้ง 3 ตัวแปร ผู้วิจัยจึงนำตัวแปรอิสระทั้ง 33 ตัวมาตรวจสอบดูความสัมพันธ์ในระดับหลายตัวแปรโดยการวิเคราะห์สหสัมพันธ์แคนนอลนิคอลนั่นคือ ดูค่าความสัมพันธ์ระหว่าง Y_i ; $i = 1,2,3$ และ X_k; $k = 1,2,\ldots,33$

3. ผลการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล

การวิเคราะห์สหสัมพันธ์แคนนอลนิคอล เพื่อศึกษาความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภทกับปัจจัยต่าง ๆ ทั้ง 33 ตัวแปร ว่ามีความสัมพันธ์กันหรือไม่ โดยใช้วิธีการทดสอบ Wilks' Lambda (Likelihood Ratio Test) และแสดงผลการวิเคราะห์ดังต่อไปนี้

สมมติฐานในการทดสอบ H_0 : ไม่มีความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภทกับตัวแปรปัจจัยต่าง ๆ ทั้ง 33 ตัวแปร

H_1 : มีความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภทกับตัวแปรปัจจัยต่าง ๆ ทั้ง 33 ตัวแปร

ตารางที่ 4 การทดสอบค่าสหสัมพันธ์แคนนอลนิคอลสำหรับ Y_1, Y_2, Y_3 และ ตัวแปรอิสระทั้ง 33 ตัวแปร โดยวิธี Wilks' Lambda

	F	df	p-value
Wilks' Lambda	33.40	96	0.0001

ผลการวิเคราะห์พบว่า ค่า Wilks' Lambda มีค่าเท่ากับ 33.40 และค่า P-Value เท่ากับ 0.0001 ที่ระดับนัยสำคัญ 0.05 จึงสรุปได้ว่ามีความสัมพันธ์ระหว่างราคาน้ำมันขายปลีก 3 ประเภทกับตัวแปรปัจจัยต่าง ๆ ทั้ง 33 ตัวแปร

ตารางที่ 5 การทดสอบค่าสหสัมพันธ์แคนนอลนิคอลทีละคู่ โดยวิธี Wilks' Lambda

ลำดับที่	ค่าไอเก็น	สหสัมพันธ์แคนนอลนิคอล	F	df	p-value
1	853.901	0.999	33.40	96	0.0001
2	20.288	0.976	7.25	62	0.0001
3	3.365	0.878	3.03	30	0.0024

จากตารางที่ 5 พบว่าค่าสหสัมพันธ์แคนนอลนิคอลระหว่างกลุ่มตัวแปรอิสระกับกลุ่มตัวแปรตามที่สูงที่สุด 3 คู่ มีค่าเท่ากับ 0.999, 0.976 และ0.878 ตามลำดับ และจากค่า p-value ที่ได้ดังตารางที่ 5 พบว่าค่าสหสัมพันธ์แคนนอลนิคอลทั้ง 3 ค่ามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ α เท่ากับ 0.05

และจากผลการวิเคราะห์แสดงในภาคผนวก ข หน้า 109 พบว่าตัวแปรแคนนอลนิคอลทั้ง 3 คู่นี้ ตัวแปรแคนนอลนิคอลคู่ที่ 1 อธิบายความผันแปรร่วมกันได้สูงสุด ร้อยละ 97.30 รองลงมาคือตัวแปรแคนนอลนิคอลคู่ที่ 2 และ 3 ซึ่งสามารถอธิบายความผันแปรแคนนอลนิคอลได้เท่ากับร้อยละ 2.31 และร้อยละ 0.38 ตามลำดับ ดังนั้นจึงเลือกใช้ตัวแปรแคนนอลนิคอลคู่ที่ 1 ที่สามารถอธิบายความผันแปรร่วมกันได้สูงสุด

ตารางที่ 6 ค่าสหสัมพันธ์ระหว่างตัวแปรแคนนอลนิคอลกับตัวแปรอิสระและตัวแปรตาม

ตัวแปร	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 1	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 2	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 3
Y_1	0.988	0.151	0.040
Y_2	0.989	-0.144	0.018
Y_3	0.990	-0.137	0.007
X_3	0.950	-0.217	0.060
X_4	0.951	-0.213	0.065
X_5	0.967	-0.201	0.051
X_6	0.946	-0.240	0.104
X_7	0.950	-0.228	0.096
X_8	0.968	-0.198	0.045
X_9	0.985	-0.143	-0.003
X_{10}	0.983	-0.129	-0.012
X_{11}	0.991	0.112	0.019
X_{19}	0.893	0.240	0.029
X_{20}	0.887	0.238	0.026
X_{22}	0.893	0.240	0.029

ตารางที่ 6 (ต่อ)

ตัวแปร	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 1	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 2	ค่าสหสัมพันธ์ แคนนอลนิคอลคู่ที่ 3
X_{23}	0.865	0.243	0.030
X_{25}	0.971	-0.115	0.113
X_{26}	0.967	-0.141	0.110
X_{27}	0.937	-0.246	0.170
X_{28}	0.866	-0.350	0.194
X_{29}	0.972	-0.182	0.089
X_{30}	0.835	-0.150	0.178
X_{31}	-0.854	-0.268	0.125
X_{34}	0.815	0.178	0.060
X_{36}	0.924	0.106	-0.098
X_{37}	0.833	-0.254	0.003
X_{40}	0.752	-0.280	0.094
X_{47}	0.453	-0.239	-0.108
X_{48}	0.673	-0.300	-0.088
X_{52}	0.566	-0.350	0.191
X_{60}	0.898	-0.098	0.155
X_{65}	0.945	0.006	0.110
X_{68}	0.512	-0.270	0.434
X_{69}	0.884	-0.130	0.029
X_{71}	0.626	-0.308	0.379
X_{72}	0.906	0.106	-0.041

จากผลการวิเคราะห์แสดงในตารางที่ 6 เมื่อพิจารณาค่าสหสัมพันธ์ระหว่างตัวแปรแคนนอนนิคอลคู่ที่ 1 กับตัวแปรอิสระที่มีค่าสูงกว่า 0.9 พบว่าจากตัวแปรอิสระทั้งหมด 33 ตัว (X_i ; i = 1,…,33) มีตัวแปร 16 ตัวแปรที่มีน้ำหนักหรือมีส่วนในการสร้างตัวแปรสูงกว่าตัวแปรอิสระอื่นๆ ซึ่งประกอบด้วย ตัวแปรดังต่อไปนี้

X_3 = ราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_4 = ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_5 = ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_6 = ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_7 = ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_8 = ราคาน้ำมันดีเซลนำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_9 = ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{10} = ราคาน้ำมันเบนซิน 91ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{11} = ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{25} = ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (เหรียญสหรัฐ/บาร์เรล)
X_{26} = ราคาน้ำมันดิบในตลาดโลก (โอมาน) (เหรียญสหรัฐ/บาร์เรล)
X_{27} = ราคาน้ำมันเบนซิน 95ในตลาดสิงคโปร์ (เหรียญสหรัฐ/ บาร์เรล)
X_{29} = ราคาน้ำมันดีเซลในตลาดสิงคโปร์ (เหรียญสหรัฐ/ บาร์เรล)
X_{36} = อัตราดอกเบี้ยเงินกู้
X_{65} = มูลค่าการส่งออกรถยนต์ (ล้านบาท) และ
X_{72} = ปริมาณการผลิตน้ำมันดิบ (1,000บาร์เรล)

จากนั้นผู้วิจัยจึงทำการศึกษาหารูปแบบความสัมพันธ์หรือสมการถดถอย เพื่อใช้ในการพยากรณ์ตัวแปรตาม Y_i ; i=1,2,3 โดยพิจารณาจากตัวแปรอิสระจำนวน 33 ตัวที่ได้จากผลการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล ดังแสดงต่อไปนี้

4. ผลการตรวจสอบข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอย

4.1 ตรวจสอบความคลาดเคลื่อนว่ามีการแจกแจงแบบปกติหลายตัวแปรหรือไม่

การตรวจสอบเกี่ยวกับการแจกแจงแบบปกติหลายตัวแปร โดยวิธีการของ Mardia ซึ่งพิจารณาจากค่าความเบ้และค่าความโด่ง ดังนี้

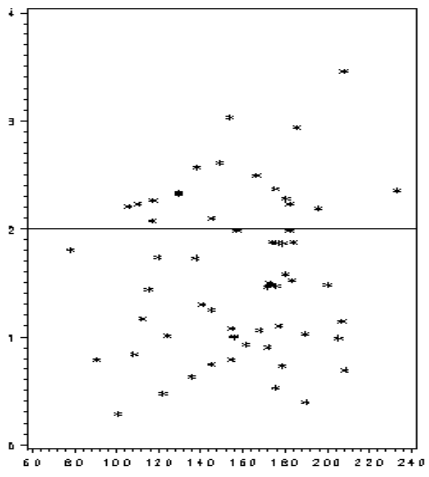
สมมติฐานในการทดสอบ H_0 : ความคลาดเคลื่อนมีการแจกแจงปกติหลายตัวแปร
H_1 : ความคลาดเคลื่อนมีการแจกแจงไม่ปกติหลายตัวแปร

ตารางที่ 7 ผลการทดสอบการแจกแจงปกติหลายตัวแปรของความคลาดเคลื่อนด้วยวิธีการของ Mardia

		B	p-value
ค่าความเบ้	5.258		0.089
ค่าความโด่ง		12.429	0.069

ผลการวิเคราะห์โดยวิธีการของ Mardia พบว่าสถิติสำหรับการทดสอบความเบ้เท่ากับ 5.258 และ ค่า P-value เท่ากับ 0.089 และสำหรับการทดสอบความโด่งเท่ากับ 12.429 และค่า P-value เท่ากับ 0.069 ซึ่งทั้งค่า P-value ของการทดสอบความเบ้และความโด่งมีค่ามากกว่าระดับนัยสำคัญ 0.05 ดังนั้น สรุปได้ว่า ความคลาดเคลื่อนมีการแจกแจงปกติหลายตัวแปร

4.2 ตรวจสอบว่าความแปรปรวนของความคลาดเคลื่อนมีค่าคงที่หรือไม่



ภาพที่ 3 การพล็อตค่าความคลาดเคลื่อนกับค่าพยากรณ์

จากภาพที่ 3 พิจารณารูปแบบการกระจายของค่าความคลาดเคลื่อนกับค่าพยากรณ์ พบว่ามีการกระจายที่ไม่เป็นแบบแผน หรือมีการกระจายรอบๆแกนอ้างอิงแสดงว่า ความคลาดเคลื่อนมีความแปรปรวนคงที่

จากการตรวจสอบข้อสมมติเบื้องต้นทั้ง 2 ข้อซึ่งพบว่าสอดคล้องกับข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอย ดังนั้นข้อมูลชุดนี้จึงเหมาะสมที่จะทำการวิเคราะห์ถดถอยต่อไปได้

จากการวิเคราะห์สหสัมพันธ์แคนนอลนิคอลพบว่าตัวแปรอิสระทั้ง 33 ตัวมีความสัมพันธ์กับตัวแปรตามทั้ง 3 ตัว และตัวแปรอิสระจำนวน 33 ตัวนี้มีความสัมพันธ์กันเองด้วย แสดงค่าสหสัมพันธ์ระหว่าง X_i ; i = 1, 2, …,33 ในภาคผนวก ข นั่นคือตัวแปรอิสระเหล่านี้ไม่เป็นอิสระต่อกัน ซึ่งจะเกิดปัญหา Multicollinearity ดังนั้นผู้วิจัยจึงใช้การวิเคราะห์ปัจจัยมาช่วยในการแก้ปัญหา Multicollinearity

5. ผลการวิเคราะห์ปัจจัย (Factor Analysis)

ผลจากการวิเคราะห์ปัจจัย โดยการหมุนแบบ Orthogonal ด้วยวิธี Varimax ซึ่งพบว่าตัวแปรอิสระทั้งหมด 33 ตัวมีปัจจัยร่วมกัน 3 ปัจจัย และปัจจัยทั้งหมดนี้อธิบายความผันแปรของตัวแปรอิสระทั้งหมดได้ร้อยละ 91.041 ดังแสดงในตารางที่ 8

ตารางที่ 8 ค่าความผันแปรจากการสกัดปัจจัย

ปัจจัย	ค่าความผันแปร (Eigenvalue)	สัดส่วนของความผันแปรรวม (% of variance)	สัดส่วนของความผันแปรรวมสะสม (Cumulative %)
1	14.753	44.706	44.706
2	8.146	24.684	69.390
3	7.145	21.651	91.041

โดยค่าสัดส่วนของความผันแปรรวม สามารถคำนวณได้ดังนี้

สัดส่วนของความผันแปรรวมของ ปัจจัยที่ 1 $= \frac{14.753}{33} \times 100 = 44.706\%$

หมายถึง ปัจจัยที่ 1 สามารถอธิบายความผันแปรทั้งหมดได้ 44.706%

สัดส่วนของความผันแปรรวมของ ปัจจัยที่ 2 $= \frac{8.146}{33} \times 100 = 24.684\%$

หมายถึง ปัจจัยที่ 2 สามารถอธิบายความผันแปรทั้งหมดได้ 24.684%

สัดส่วนของความผันแปรรวมของ ปัจจัยที่ 3 $= \frac{7.145}{33} \times 100 = 21.651\%$

หมายถึง ปัจจัยที่ 3 สามารถอธิบายความผันแปรทั้งหมดได้ 21.651%

ส่วนการจัดตัวแปรอิสระทั้ง 33 ตัวแปรให้เป็นปัจจัยต่าง ๆ จำนวน 3 ปัจจัย พิจารณาจากค่าน้ำหนักปัจจัยซึ่งแสดงในตารางที่ 9 ดังนี้

ตารางที่ 9 ค่าน้ำหนักปัจจัย โดยการหมุนด้วยวิธี Varimax

ตัวแปร	ปัจจัย		
	ปัจจัยที่ 1	ปัจจัยที่ 2	ปัจจัยที่ 3
X_3	0.605	0.579	0.508
X_4	0.610	0.574	0.506
X_5	0.637	0.570	0.497
X_6	0.602	0.565	0.525
X_7	0.609	0.534	0.518
X_8	0.640	0.566	0.497
X_9	0.714	0.537	0.423
X_{10}	0.721	0.534	0.408
X_{11}	0.813	0.427	0.344
X_{19}	0.949	0.182	0.177

ตารางที่ 9 (ต่อ)

ตัวแปร	ปัจจัย		
	ปัจจัยที่ 1	ปัจจัยที่ 2	ปัจจัยที่ 3
X_{20}	0.956	0.167	0.169
X_{22}	0.949	0.182	0.177
X_{23}	0.957	0.104	0.177
X_{25}	0.699	0.526	0.465
X_{26}	0.980	0.542	0.475
X_{27}	0.920	0.531	0.540
X_{28}	0.458	0.621	0.585
X_{29}	0.690	0.504	0.489
X_{30}	0.640	0.270	0.549
X_{31}	-0.874	-0.318	-0.038
X_{34}	0.943	-0.009	0.205
X_{36}	0.846	0.431	0.139
X_{37}	0.435	0.740	0.395
X_{40}	0.459	0.513	0.457
X_{47}	0.021	0.871	0.105
X_{48}	0.215	0.822	0.320
X_{52}	0.069	0.539	0.673
X_{60}	0.591	0.487	0.534
X_{65}	0.726	0.462	0.411
X_{68}	0.108	0.154	0.930
X_{69}	0.668	0.498	0.337
X_{71}	0.207	0.276	0.900
X_{72}	0.720	0.388	0.389

จากตารางที่ 9 สามารถอธิบายได้ว่าตัวแปรอิสระทั้งหมดจำนวน 33 ตัวแปรมีความสัมพันธ์กันจัดได้เป็น 3 ปัจจัยและโดยพิจารณาจากค่าน้ำหนักปัจจัย สามารถสรุปองค์ประกอบของแต่ละปัจจัยได้ดังนี้

ปัจจัยที่ 1 ประกอบด้วยตัวแปรอิสระจำนวน 26 ตัวแปร ดังนี้

X_3 = ราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_4 = ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_5 = ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_6 = ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_7 = ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_8 = ราคาน้ำมันดีเซล นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_9 = ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{10} = ราคาน้ำมันเบนซิน 91ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{11} = ราคาน้ำมันดีเซล ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{19} = อัตราภาษีส่งออกน้ำมันเบนซิน 95 (บาท/ลิตร)
X_{20} = อัตราภาษีส่งออกน้ำมันเบนซิน 91 (บาท/ลิตร)
X_{22} = อัตราภาษีนำเข้าน้ำมันเบนซิน 95 (บาท/ลิตร)
X_{23} = อัตราภาษีนำเข้าน้ำมันเบนซิน 91 (บาท/ลิตร)
X_{25} = ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (เหรียญสหรัฐ/บาร์เรล)
X_{26} = ราคาน้ำมันดิบในตลาดโลก (โอมาน) (เหรียญสหรัฐ/บาร์เรล)
X_{27} = ราคาน้ำมันเบนซิน 95ในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)
X_{28} = ราคาน้ำมันเบนซิน 92ในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)
X_{29} = ราคาน้ำมันดีเซลในตลาดสิงคโปร์ (เหรียญสหรัฐ/บาร์เรล)
X_{30} = ปริมาณการใช้ไฟฟ้า (GWh)
X_{31} = ประชากรรวมในกรุงเทพฯ (พันคน)
X_{34} = ผลิตภัณฑ์มวลรวมในประเทศ (ล้านบาท)
X_{36} = อัตราดอกเบี้ยเงินกู้
X_{60} = มูลค่าการนำเข้าน้ำมันดิบ (ล้านบาท)
X_{65} = มูลค่าการส่งออกรถยนต์ (ล้านบาท)

X_{69} = มูลค่าการส่งออกผลิตภัณฑ์ปิโตรเลียม (ล้านบาท)

X_{72} = ปริมาณการผลิตน้ำมันดิบ (1,000บาร์เรล)

ผู้วิจัยจะเรียก ปัจจัยที่ 1 นี้ว่า **ปัจจัยด้านราคา (Factor 1)**

ปัจจัยที่ 2 ประกอบด้วยตัวแปรอิสระจำนวน 4 ตัวแปร ดังนี้

X_{37} = กองทุนน้ำมัน (ล้านบาท)

X_{40} = ปริมาณการผลิตน้ำมัน LPG (ล้านลิตร)

X_{47} = ปริมาณการส่งออกน้ำมันเบนซิน 95 (ล้านลิตร)

X_{48} = ปริมาณการส่งออกน้ำมันเบนซิน 91 (ล้านลิตร)

ผู้วิจัยจะเรียกปัจจัยที่ 2 นี้ว่า **ปัจจัยด้านการผลิตและส่งออกน้ำมัน (Factor 2)**

ปัจจัยที่ 3 ประกอบด้วยตัวแปรอิสระจำนวน 3 ตัวแปร ดังนี้

X_{52} = จำนวนรถจักรยานยนต์ที่จดทะเบียนในกรุงเทพฯ (คัน)

X_{68} = ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศ (คัน)

X_{71} = ปริมาณการผลิตรถจักรยานยนต์ (พันคัน)

ผู้วิจัยจะเรียกปัจจัยที่ 3 นี้ว่า **ปัจจัยด้านเศรษฐกิจ (Factor 3)**

การสร้างคะแนนปัจจัย (Factor scores)

หลังจากทำการสกัดปัจจัยแล้ว ผู้วิจัยจึงทำการสร้างคะแนนปัจจัยด้วยวิธีวิเคราะห์การถดถอย ซึ่งสามารถเขียนสมการถดถอยสำหรับการประมาณปัจจัยแต่ละตัวได้ดังนี้

$$F = \beta_{p1}Z_1+\beta_{p2}Z_2+\ldots+\beta_{pj}Z_j+\ldots+\beta_{mn}Z_n \;;\; p = 1, 2, \ldots, m \;;\; j = 1, 2, \ldots, n$$

โดยที่ F = ค่าประมาณของคะแนนปัจจัย

Z_j = ค่าคะแนนมาตรฐานของตัวแปรที่ j

β_{pj} = สัมประสิทธิ์การถดถอยเชิงเส้นของปัจจัยที่ p ของตัวแปรที่ j

m = จำนวนตัวแปรอิสระ

n = จำนวนข้อมูล

แสดงคะแนนปัจจัยดังตารางที่ 10

ตารางที่ 10 คะแนนปัจจัยโดยวิธีการวิเคราะห์การถดถอย

Factor 1	Factor 2	Factor 3
-0.445	0.569	-2.735
-0.255	-0.167	-2.324
-0.232	-0.602	-1.513
-0.194	-0.213	-1.862
-0.324	0.167	-1.957
-0.288	0.209	-2.143
-0.227	-0.722	-1.213
-0.454	-0.283	-1.204
-0.386	-0.465	-1.018
-0.483	-0.639	-0.499
-0.616	-0.674	-0.321
-0.531	-0.331	-0.820
-0.500	-0.746	-0.028
-0.472	-0.848	0.018
-0.468	-1.356	0.740
-0.510	-0.538	-0.532
-0.710	-0.542	-0.112
-0.668	-0.431	-0.358
-0.482	-1.055	0.062
-0.593	0.186	-0.846

ตารางที่ 10 (ต่อ)

Factor 1	Factor 2	Factor 3
-0.379	-0.627	-0.624
-0.277	-0.805	-0.435
-0.423	-1.197	0.144
-0.475	-0.963	0.159
-0.550	-1.106	0.577
-0.585	-1.438	0.949
-0.816	-1.004	1.344
-0.553	-0.577	0.192
-0.755	-1.054	1.730
-0.948	-0.180	0.945
-0.704	-0.479	0.988
-0.417	-0.097	0.484
-0.686	0.465	0.275
-0.741	0.695	0.622
-0.980	0.210	1.473
-1.135	1.097	0.404
-0.795	0.302	0.927
-0.734	0.417	0.457
-0.687	0.759	0.945
-0.779	2.270	-0.494
-0.556	0.405	1.545
-0.479	0.810	1.604
-0.212	1.905	0.515
-0.109	3.289	-0.455
0.094	2.839	0.006
0.149	1.679	0.684
0.209	0.998	0.885

ตารางที่ 10 (ต่อ)

Factor 1	Factor 2	Factor 3
0.774	1.172	-0.009
1.314	0.182	0.588
1.624	0.080	-0.084
1.572	-0.935	1.615
1.772	0.719	-0.462
1.598	0.463	0.711
1.676	0.014	0.877
1.801	0.602	0.147
1.776	0.579	0.250
1.828	0.008	-0.075
2.292	-0.897	-0.052
2.470	-0.860	-0.394
2.660	-1.261	-0.293

โดยการนำค่าประมาณน้ำหนักปัจจัย หลังจากหมุนแกนปัจจัย โดยวิธี Varimax มาประมาณค่าคะแนนปัจจัย โดยวิธีการวิเคราะห์การถดถอย ได้สัมประสิทธิ์คะแนนปัจจัยดังแสดงในตารางที่ 11

ตารางที่ 11 สัมประสิทธิ์คะแนนปัจจัย

ตัวแปร	Factor 1 (ปัจจัยด้านราคา)	Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน)	Factor 3 (ปัจจัยด้านเศรษฐกิจ)
X_3	-0.003	0.054	0.025
X_4	-0.001	0.052	0.026
X_5	0.005	0.049	0.020
X_6	-0.003	0.041	0.040
X_7	-0.001	0.042	0.035

ตารางที่ 11 (ต่อ)

ตัวแปร	Factor 1 (ปัจจัยด้านราคา)	Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน)	Factor 3 (ปัจจัยด้านเศรษฐกิจ)
X_8	0.006	0.047	0.021
X_9	0.029	0.048	-0.014
X_{10}	0.031	0.051	-0.022
X_{11}	0.066	0.006	-0.028
X_{19}	0.266	-0.160	-0.090
X_{20}	0.137	-0.086	-0.044
X_{22}	0.000	0.000	0.000
X_{23}	0.144	-0.121	-0.019
X_{25}	0.024	0.030	0.013
X_{26}	0.017	0.037	0.014
X_{27}	0.003	0.017	0.057
X_{28}	-0.041	0.067	0.064
X_{29}	0.023	0.012	0.034
X_{30}	0.037	-0.124	0.149
X_{31}	-0.115	-0.046	0.159
X_{34}	0.153	-0.187	0.035
X_{36}	0.088	0.074	-0.141
X_{37}	-0.044	0.196	-0.076
X_{40}	-0.017	0.054	0.033
X_{47}	-0.110	0.411	-0.241
X_{48}	-0.087	0.290	-0.125
X_{52}	-0.107	0.042	0.171
X_{60}	0.003	-0.004	0.072
X_{65}	0.041	0.012	0.004
X_{68}	-0.075	-0.249	0.435
X_{69}	0.032	0.062	-0.043

ตารางที่ 11 (ต่อ)

ตัวแปร	Factor 1 (ปัจจัยด้านราคา)	Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน)	Factor 3 (ปัจจัยด้านเศรษฐกิจ)
X_{71}	-0.070	-0.188	0.370
X_{72}	0.050	-0.018	0.017

จากค่าสัมประสิทธิ์คะแนนปัจจัยโดยวิธีวิเคราะห์การถดถอย สามารถนำมาเขียนเป็นสมการถดถอยของแต่ละปัจจัยได้ดังนี้

Factor 1 (ปัจจัยด้านราคา) = $-0.003X_3-0.001X_4+0.005X_5-0.003X_6-0.001X_7+0.006X_8+0.029X_9+0.031X_{10}+0.066X_{11}+0.266X_{19}+0.137X_{20}+0.000X_{22}+0.144X_{23}+0.024X_{25}+0.017X_{26}+0.003X_{27}-0.041X_{28}+0.023X_{29}+0.037X_{30}-0.115X_{31}+0.153X_{34}+0.088X_{36}+0.003X_{40}+0.041X_{65}+0.032X_{69}+0.050X_{72}$

Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) = $0.196X_{37}+0.054X_{40}+0.411X_{47}+0.290X_{48}$

Factor 3 (ปัจจัยด้านเศรษฐกิจ) = $0.171X_{52}+0.435X_{68}+0.370X_{71}$

หลังจากทำการแก้ปัญหา Multicollinearity โดยการลดจำนวนตัวแปรอิสระ 33 ตัวเป็นปัจจัย 3 ปัจจัย แล้วจึงทำการวิเคราะห์การถดถอยดังนี้

6. ผลการวิเคราะห์ถดถอยแบบหลายตัวแปร (Multivariate Regression Analysis)

จากผลการวิเคราะห์ปัจจัยด้วยวิธีหมุนแบบ Varimax สามารถสกัดปัจจัยได้ 3 ปัจจัย และนำค่าคะแนนปัจจัยที่ได้มาทดสอบ สามารถทดสอบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระโดยสถิติ Wilks' lambda (λ) ทดสอบและมีสมมติฐานในการทดสอบดังนี้

6.1 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวกับตัวแปรอิสระ 3 ตัว

$$H_0 : \begin{bmatrix} \beta_{01} & \beta_{02} & \beta_{03} \\ \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} = 0 \qquad H_1 : \begin{bmatrix} \beta_{01} & \beta_{02} & \beta_{03} \\ \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \neq 0$$

ตารางที่ 12 การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวกับตัวแปรอิสระ 3 ตัวหรือปัจจัย 3 ปัจจัย

	F	df	p-value
Wilks' Lambda	112.63	9	0.0001

ผลการวิเคราะห์ ได้ค่า Wilks' lambda เท่ากับ 112.63 และ ค่า P-value เท่ากับ 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 ดังนั้น สรุปได้ว่า ตัวแปรตามทั้ง 3 ตัวแปร (ราคาน้ำมันดีเซล, ราคาน้ำมันเบนซิน 91 และราคาน้ำมันเบนซิน 95) และตัวแปรอิสระทั้ง 3 ตัว (ปัจจัยด้านปัจจัยด้านราคา, ปัจจัยด้านการผลิตและส่งออกน้ำมัน และปัจจัยด้านเศรษฐกิจ)มีความสัมพันธ์กันในรูปแบบเชิงเส้นตรง ที่ระดับนัยสำคัญ 0.05 แสดงว่ามีตัวแปรอิสระอย่างน้อย 1 ตัวที่มีความสัมพันธ์กับตัวแปรตามทั้ง 3 ตัวจึงทำการทดสอบความสัมพันธ์ของตัวแปรอิสระกับตัวแปรตามแต่ละตัวต่อไปดังนี้

6.2 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ (Factor 1 ซึ่งเป็นปัจจัยด้านราคา)

สมมติฐานในการทดสอบ $H_0 : [\beta_{11} \ \beta_{12} \ \beta_{13}] = 0$

$H_1 : [\beta_{11} \ \beta_{12} \ \beta_{13}] \neq 0$

ตารางที่ 13 การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวกับตัวแปรอิสระ (Factor 1)

	F	df	p-value
Wilks' Lambda	960.96	3	0.0001

ผลการวิเคราะห์ได้ค่า Wilks' lambda เท่ากับ 960.96 และ ค่า P-value เท่ากับ 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 ดังนั้น สรุปได้ว่า ปัจจัยด้านราคามีความสัมพันธ์กับ ราคาขายปลีกน้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 ที่ระดับนัยสำคัญ 0.05

6.3 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ (Factor 2 ซึ่งเป็นปัจจัยด้านการผลิตและส่งออกน้ำมัน) ดังนี้

สมมติฐานในการทดสอบ $H_0 : [\beta_{21} \ \beta_{22} \ \beta_{23}] = 0$

$H_1 : [\beta_{21} \ \beta_{22} \ \beta_{23}] \neq 0$

ตารางที่ 14 การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวกับตัวแปรอิสระ (Factor 2)

	F	df	p-value
Wilks' Lambda	371.53	3	0.0001

ผลการวิเคราะห์ได้ค่า Wilks' lambda เท่ากับ 371.53 และ ค่า P-value เท่ากับ 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 ดังนั้น สรุปได้ว่า ปัจจัยด้านการผลิตและส่งออกน้ำมัน มีความสัมพันธ์กับราคาขายปลีก น้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 ที่ระดับนัยสำคัญ 0.05

6.4 ตรวจสอบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระ (Factor 3 ซึ่งเป็นปัจจัยด้านเศรษฐกิจ) ดังนี้

สมมติฐานในการทดสอบ $H_0 : [\beta_{31} \ \beta_{32} \ \beta_{33}] = 0$

$H_1 : [\beta_{31} \ \beta_{32} \ \beta_{33}] \neq 0$

ตารางที่ 15 การทดสอบความสัมพันธ์ระหว่างตัวแปรตาม 3 ตัวกับตัวแปรอิสระ (Factor 3)

	F	df	p-value
Wilks' Lambda	308.43	3	0.0001

ผลการวิเคราะห์ได้ค่า Wilks' lambda เท่ากับ 308.43 และ ค่า P-value เท่ากับ 0.0001 ซึ่งน้อยกว่าระดับนัยสำคัญ 0.05 ดังนั้น สรุปได้ว่า ปัจจัยด้านเศรษฐกิจ มีความสัมพันธ์กับ ราคาขายปลีกน้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 ที่ระดับนัยสำคัญ 0.05

จากการทดสอบความสัมพันธ์ของตัวแปรอิสระกับตัวแปรตามแต่ละตัว พบว่าตัวแปรอิสระทั้ง 3 ตัวมีความสัมพันธ์กับตัวแปรตาม แสดงว่า ปัจจัยด้านราคา (Factor 1) ปัจจัยด้านการผลิตและส่งออกน้ำมัน (Factor 2) และ ปัจจัยด้านเศรษฐกิจ (Factor 3) มีความสัมพันธ์กับราคาขายปลีกน้ำมันดีเซล(Y_1)น้ำมันเบนซิน 91 (Y_2) และน้ำมันเบนซิน 95 (Y_3) ที่ระดับนัยสำคัญ 0.05 และทำการทดสอบค่าสัมประสิทธิ์การถดถอยของแต่ละตัวแปรต่อไปดังนี้

6.5 การทดสอบค่าสัมประสิทธิ์การถดถอยของราคาขายปลีกน้ำมันดีเซล (Y_1) ดังนี้

ตารางที่ 16 ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันดีเซล

ตัวแปร	Estimate	Standard Error	F	Pr >F
Intercept	17.547	0.132	132.33	0.0001
Factor 1	4.239	0.134	31.70	0.0001
Factor 2	2.057	0.134	15.38	0.0001
Factor 3	1.733	0.134	12.96	0.0001

R-Square = 0.962 Root MSE = 1.027 $\bar{y}_1$ = 17.547

จากตัวแปร Intercept , Factor 1 (ปัจจัยด้านราคา), Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน)และ Factor 3 (ปัจจัยด้านเศรษฐกิจ) มีค่า P - value เท่ากับ 0.0001 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ 0.05 แสดงว่าตัวแปร Factor 1, Factor 2, Factor 3 มีความสัมพันธ์กับตัวแปรตาม Y_1 มีค่า R^2 เท่ากับ 0.962 และมีค่าความคลาดเคลื่อน เท่ากับ 1.027 ซึ่งหมายความว่า ตัวแปรอิสระนี้สามารถอธิบายการเปลี่ยนแปลงของตัวแปรตาม Y_1 ได้ 96.2 %

จากการวิเคราะห์ถดถอยนี้ สามารถเขียนสมการถดถอยได้ดังนี้

$$\hat{Y}_1 = 17.547+4.239\text{Factor 1}+2.057\text{Factor 2}+1.733\text{Factor 3}$$

6.6 การทดสอบค่าสัมประสิทธิ์การถดถอยของราคาขายปลีกน้ำมันเบนซิน 91 (Y_2) ดังนี้

ตารางที่ 17 ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันเบนซิน 91

ตัวแปร	Estimate	Standard Error	F	Pr > F
Intercept	19.620	0.087	225.79	0.0001
Factor 1	3.660	0.088	41.77	0.0001
Factor 2	2.536	0.088	28.94	0.0001
Factor 3	2.192	0.088	25.02	0.0001

R-Square = 0.983 Root MSE = 0.673 $\bar{y}_2$ = 19.619

จากตัวแปร Intercept , Factor 1 (ปัจจัยด้านราคา), Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) และ Factor 3 (ปัจจัยด้านเศรษฐกิจ) มีค่า P - value เท่ากับ 0.0001 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ 0.05 แสดงว่าตัวแปร Factor 1, Factor 2, Factor 3 มีความสัมพันธ์กับตัวแปรตาม Y_2 และมีค่า R^2 เท่ากับ 0.983 และมีค่าความคลาดเคลื่อน เท่ากับ 0.673 ซึ่งหมายความว่า ตัวแปรอิสระนี้สามารถอธิบายการเปลี่ยนแปลงของตัวแปรตาม Y_2 ได้ 98.3 % และมีสมการถดถอย คือ

$$\hat{Y}_2 = 19.619+3.66\text{Factor 1}+2.536\text{Factor 2}+2.192\text{Factor 3}$$

6.7 การทดสอบค่าสัมประสิทธิ์การถดถอยของราคาขายปลีกน้ำมันเบนซิน 95 (Y_3) ดังนี้

ตารางที่ 18 ผลการวิเคราะห์ปัจจัยที่มีผลต่อราคาขายปลีกน้ำมันเบนซิน 95

ตัวแปร	Estimate	Standard Error	F	Pr > F
Intercept	20.489	0.086	237.82	0.0001
Factor 1	3.627	0.087	41.75	0.0001
Factor 2	2.504	0.087	28.82	0.0001
Factor 3	2.126	0.087	24.48	0.0001

R-Square = 0.983 Root MSE = 0.667 $\bar{y}_3$ = 20.489

จากตัวแปร Intercept , Factor 1 (ปัจจัยด้านราคา), Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) และ Factor 3 (ปัจจัยด้านเศรษฐกิจ) มีค่า P - value เท่ากับ 0.0001 ซึ่งมีค่าน้อยกว่าค่าระดับนัยสำคัญ 0.05 แสดงว่าตัวแปร Factor 1, Factor 2, Factor 3 มีความสัมพันธ์กับตัวแปรตาม Y_3 และมีค่า R^2 เท่ากับ 0.983 และมีค่าความคลาดเคลื่อน เท่ากับ 0.667 ซึ่งหมายความว่า ตัวแปรอิสระนี้สามารถอธิบายความเปลี่ยนแปลงของ ตัวแปรตาม Y_3 ได้ 98.3 % และมีสมการถดถอย คือ

$$\hat{Y}_3 = 20.489+3.627\text{Factor 1}+2.504\text{Factor 2}+2.126\text{Factor 3}$$

จากข้อมูลที่นำมาศึกษาตั้งแต่เดือนมกราคม 2545- เดือนธันวาคม 2549 รวมระยะเวลา 60 เดือน เพื่อศึกษาปัจจัยที่มีอิทธิพลต่อราคาขายปลีกของน้ำมันเชื้อเพลิง 3 ประเภท คือ น้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 และศึกษารูปแบบความสัมพันธ์ระหว่างราคาขายปลีกน้ำมันทั้ง 3 ประเภท โดยใช้วิธีการทางสถิติได้แก่ การวิเคราะห์การถดถอย การวิเคราะห์สหสัมพันธ์แคนนอลนิคอล และการวิเคราะห์ปัจจัย เป็นต้น พบว่าลักษณะข้อมูลมีความสอดคล้องกับข้อสมมติเบื้องต้นของการวิเคราะห์การถดถอย และมีการนำเทคนิคการวิเคราะห์ปัจจัยมาช่วยในการแก้ปัญหา Multicollinearity ซึ่งได้สมการถดถอยที่สามารถนำไปใช้ในการพยากรณ์ปริมาณราคาขายปลีกของน้ำมันดีเซลน้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 ดังนี้

1. สมการถดถอยสำหรับราคาขายปลีกน้ำมันดีเซล คือ

$$\hat{Y}_1 = 17.547+4.239\text{Factor 1}+2.057\text{Factor 2}+1.733\text{Factor 3}$$

เมื่อ Factor 1, Factor 2 และFactor 3 คือ ปัจจัยด้านราคา, ปัจจัยด้านการผลิตและส่งออกน้ำมันและ ปัจจัยด้านเศรษฐกิจ ตามลำดับ โดยมีค่า R^2 เท่ากับ 0.962 และมีค่าความคลาดเคลื่อนเท่ากับ 1.027

2. สมการถดถอยสำหรับราคาขายปลีกน้ำมันเบนซิน 91 คือ

$$\hat{Y}_2 = 19.619+3.66\text{Factor 1}+2.536\text{Factor 2}+2.192\text{Factor 3}$$

เมื่อ Factor 1, Factor 2 และFactor 3 คือ ปัจจัยด้านราคา, ปัจจัยด้านการผลิตและส่งออกน้ำมันและ ปัจจัยด้านเศรษฐกิจ ตามลำดับ โดยมีค่า R^2 เท่ากับ 0.983 และมีค่าความคลาดเคลื่อนเท่ากับ 0.673

3. สมการถดถอยสำหรับราคาขายปลีกน้ำมันเบนซิน 95 คือ

$$\hat{Y}_3 = 20.489+3.627\text{Factor 1}+2.504\text{Factor 2}+2.126\text{Factor 3}$$

เมื่อ Factor 1, Factor 2 และFactor 3 คือ ปัจจัยด้านราคา, ปัจจัยด้านการผลิตและส่งออกน้ำมันและ ปัจจัยด้านเศรษฐกิจ ตามลำดับ โดยมีค่า R^2 เท่ากับ 0.983 และมีค่าความคลาดเคลื่อนเท่ากับ 0.667

โดยมีสมการการถดถอยของแต่ละปัจจัย สามารถคำนวณได้ดังนี้

Factor 1 (ปัจจัยด้านราคา) = $-0.003X_3-0.001X_4+0.005X_5-0.003X_6-0.001X_7+0.006X_8+0.029X_9+0.031X_{10}+0.066X_{11}+0.266X_{19}+0.137X_{20}+0.000X_{22}+0.144X_{23}+0.024X_{25}+0.017X_{26}+0.003X_{27}-0.041X_{28}+0.023X_{29}+0.037X_{30}-0.115X_{31}+0.153X_{34}+0.088X_{36}+0.003X_{40}+0.041X_{65}+0.032X_{69}+0.050X_{72}$

Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) = $0.196X_{37}+0.054X_{40}+0.411X_{47}+0.290X_{48}$

Factor 3 (ปัจจัยด้านเศรษฐกิจ) = $0.171X_{52}+0.435X_{68}+0.370X_{71}$

7. การนำสมการถดถอยที่ได้จากการวิจัยนี้ไปใช้ในการพยากรณ์

เพื่อแสดงการใช้ผลการศึกษาครั้งนี้ในการพยากรณ์ราคาน้ำมันขายปลีก 3 ประเภท ผู้วิจัยได้ทำการเก็บข้อมูลราคาน้ำมันและปัจจัยต่าง ๆ ที่เกี่ยวข้องซึ่งได้แก่ ตัวแปรอิสระ X_i จำนวน 33 ตัวแปร ในช่วงเดือนมกราคม 2550 – ตุลาคม 2550 และนำข้อมูลดังกล่าวมาใช้กับสมการถดถอยที่ได้จากการศึกษา เพื่อเป็นตัวอย่างในการพยากรณ์ แสดงรายละเอียดดังนี้

7.1 ผลการวิเคราะห์ข้อมูลเบื้องต้น

ตารางที่ 19 ค่าสถิติเบื้องต้นสำหรับตัวแปรตาม 3 ตัวแปร ที่ทำการศึกษาจากหน่วยใหม่ (ข้อมูลเดือนมกราคม 2550 – เดือนตุลาคม 2550)

ตัวแปร	ค่าเฉลี่ย	ค่าเบี่ยงเบนมาตรฐาน	ค่าต่ำสุด	ค่าสูงสุด
Y_1	25.009	1.489	22.76	27.47
Y_2	27.743	1.756	24.83	29.57
Y_3	28.526	1.748	25.63	30.31

จากตารางที่ 19 เป็นผลการวิเคราะห์ข้อมูลเบื้องต้นสำหรับตัวแปรตาม 3 ตัวแปร มีค่าเฉลี่ยของราคาน้ำมันดีเซล (Y_1) คือ 25.01 (บาท/ลิตร) ค่าเฉลี่ยของราคาน้ำมันเบนซิน 91 (Y_2) คือ 27.74 (บาท/ลิตร) และค่าเฉลี่ยของราคาน้ำมันเบนซิน 95 (Y_3) คือ 28.53 (บาท/ลิตร)

7.2 นำข้อมูลตัวอย่างใหม่ที่เก็บตัวแปรอิสระมาทั้ง 33 ตัวแปร แปลงเป็นค่ามาตรฐาน แล้วแทนค่าลงในสมการของ Factor 1, Factor 2, และ Factor 3 ได้ดังนี้

Factor 1 (ปัจจัยด้านราคา) = -0.003(-1.929)-0.001(-1.923)+0.005(-1.621)- 0.003(-1.678)-
0.001(-1.678)+0.006(-1.578)+0.029(-1.922)+ 0.031(-1.928)+
0.066(-1.616)+0.266(-0.615)+0.137(-0.617)+0.000(-0.615) +
0.144(-0.618) + 0.024(-1.636)+0.017(-1.631)+0.003(-1.943)–
0.041(-1.952)+0.023(-1.625)+ 0.037(-1.205)-0.115(0)+
0.153(0.939)+ 0.088(2.232)+ 0.003(-1.642)+0.041(-0.101) +
0.032(-0.817)+ 0.050(-1.452)
= -0.405

Factor 2 (ปัจจัยด้านการผลิต = 0.196(1.471)+0.054(1.215)+ 0.411(0.015)+ 0.290(0.129)
และส่งออกน้ำมัน) = 0.397

Factor 3 (ปัจจัยด้านเศรษฐกิจ) = 0.171(0.213)+0.435(0.724)+0.370(1.379)
= 0.862

7.3 นำค่า Factor 1, Factor 2, และ Factor 3 ที่คำนวณได้แทนลงในสมการถดถอยดังนี้

$\hat{Y}_1$ = 17.547+4.239Factor 1+2.057Factor 2+1.733Factor 3
= 17.547+4.239(-0.405)+2.057(0.397)+1.733(0.862)
= 18.140

$\hat{Y}_2$ = 19.619+3.66Factor 1+2.536Factor 2+2.192Factor 3
= 19.619+3.66(-0.405)+2.536(0.397)+2.192(0.862)
= 21.033

$$\hat{Y}_3 = 20.489+3.627\text{Factor 1}+2.504\text{Factor 2}+2.126\text{Factor 3}$$
$$= 20.489+3.627(-0.405)+2.504(0.397)+2.126(0.862)$$
$$= 21.847$$

7.4 ผลการวิเคราะห์ถดถอย

ตารางที่ 20 ผลการพยากรณ์ของหน่วยตัวอย่างใหม่ ($\hat{Y}_i$; i = 1,2,3) เทียบกับข้อมูลจริง (Y_1 ; i = 1,2,3)

เดือน	ค่าพยากรณ์ $\hat{Y}_1$	ค่าจริง Y_1	ค่าพยากรณ์ $\hat{Y}_2$	ค่าจริง Y_2	ค่าพยากรณ์ $\hat{Y}_3$	ค่าจริง Y_3
มกราคม	18.140	22.76	21.033	24.91	21.847	25.71
กุมภาพันธ์	15.612	23.1	17.833	24.83	18.731	25.63
มีนาคม	18.251	23.62	20.814	26.54	21.663	27.34
เมษายน	13.712	24.83	15.689	27.88	16.674	28.68
พฤษภาคม	19.251	25.34	22.597	29.19	23.371	29.99
มิถุนายน	16.850	25.34	18.922	29.2	19.759	30.01
กรกฎาคม	16.015	25.61	18.119	29.05	18.972	29.85
สิงหาคม	21.323	25.44	23.392	27.73	24.189	28.53
กันยายน	22.615	26.58	24.178	28.59	24.972	29.21
ตุลาคม	21.234	27.47	23.306	29.51	24.113	30.31

จากค่าที่คำนวณได้ เป็นผลการพยากรณ์ของหน่วยตัวอย่างใหม่ ที่สามารถนำไปใช้ในการพยากรณ์ราคาน้ำมันขายปลีกทั้ง 3 ประเภท

วิจารณ์

จากการศึกษาปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีกโดยการวิเคราะห์การถดถอยแบบหลายตัวแปรผลการศึกษาพบว่า

1. ปัจจัยที่มีอิทธิพลต่อราคาขายปลีกของน้ำมันเชื้อเพลิง 3 ประเภท คือ น้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 พบว่า ปัจจัยที่มีอิทธิพลหรือผลกระทบต่อราคาน้ำมันขายปลีก คือ ปัจจัยด้านราคา ซึ่งสอดคล้องกับผลการศึกษาของ พีรณัฐ แดงสกุล (2540) ที่ได้ทำการศึกษาเรื่อง การวิเคราะห์การกำหนดราคาน้ำมันในระดับค้าปลีก ผลการวิจัยพบว่าโครงสร้างราคาน้ำมันในปัจจุบันตัวแปรที่สามารถกำหนดราคาขายปลีกคือราคา ณ โรงกลั่นและนุชชฎา ถิ่นกมุท (2547) ได้ทำการศึกษา เรื่องโครงสร้างและการกำหนดราคาค้าปลีกน้ำมันเชื้อเพลิง ผลการวิจัยพบว่าโครงสร้างราคาค้าปลีกน้ำมันเชื้อเพลิงในปัจจุบันตัวแปรทุกตัวในสมการโครงสร้างราคาสามารถกำหนดราคาค้าปลีกซึ่งประกอบด้วยราคา ณ โรงกลั่น ราคานำเข้า ภาษีสรรพสามิต ราคาน้ำมันสำเร็จรูปที่ตลาดจรสิงคโปร์ซึ่งทุกปัจจัยดังกล่าวมีความสัมพันธ์แบบมีนัยสำคัญ

2. ในการศึกษาครั้งนี้มีจำนวนตัวแปรอิสระหลายตัวจึงทำให้เกิดปัญหา Multicollinearity ผู้วิจัยจึงได้นำเทคนิคการวิเคราะห์ปัจจัย มาช่วยในการแก้ปัญหาดังกล่าว เนื่องจากการวิเคราะห์ปัจจัยเป็นวิธีการอย่างหนึ่งในการแก้ปัญหา Multicollinearity โดยการรวมกลุ่มตัวแปรอิสระที่มีความสัมพันธ์กันไว้ด้วยกัน (กัลยา, 2548) และพบว่าสามารถจัดปัจจัยได้ 3 ปัจจัยคือปัจจัยด้านราคา, ปัจจัยด้านการผลิตและส่งออกน้ำมัน, และปัจจัยด้านเศรษฐกิจ ซึ่งทั้ง 3 ปัจจัยนี้สามารถอธิบายความผันแปรของตัวแปรอิสระทั้งหมดได้ถึงร้อยละ 90.041

3. จากการศึกษาพบว่าปัจจัยด้านราคามีตัวแปรอิสระจำนวน 26 ตัวแปร ในปัจจัยนี้มีตัวแปรที่ไม่ใช่ด้านราคา คือ จำนวนประชากรในกรุงเทพมหานคร และปริมาณการใช้ไฟฟ้าที่เป็นตัวแปรทั้งสองมีความเกี่ยวข้องกับราคา เพราะว่าถ้าประชากรในกรุงเทพมหานครมีจำนวนเพิ่มขึ้นทำให้มีความต้องการในการใช้น้ำมันเพิ่มขึ้น จึงทำให้ส่งผลต่อราคาน้ำมันเพิ่มขึ้นด้วย ปริมาณการใช้ไฟฟ้าก็เช่นเดียวกัน จึงทำให้ตัวแปรทั้งสองนี้มีความสัมพันธ์กับตัวแปรด้านราคา ซึ่งสอดคล้องกับหลักเศรษฐศาสตร์ในเรื่องของอุปสงค์และอุปทาน (ณรงค์, 2548) ที่กล่าวว่าการเพิ่มขึ้นของจำนวนประชากรย่อมส่งผลโดยตรงต่อการเพิ่มขึ้นของความต้องการสินค้าหรือบริการต่าง ๆ นั่นคือ จำนวนประชากรมักจะมีการเปลี่ยนแปลงในทิศทางเดียวกันกับปริมาณอุปสงค์ของสินค้าหรือบริการ

สรุปและข้อเสนอแนะ

สรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษา ปัจจัยที่มีอิทธิพลต่อราคาขายปลีกของน้ำมันเชื้อเพลิง 3 ประเภท คือ น้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 และศึกษารูปแบบความสัมพันธ์ระหว่างราคาขายปลีกน้ำมัน 3 ประเภท กับ ปัจจัยต่าง ๆ ที่คาดว่ามีอิทธิพลต่อราคาขายปลีกน้ำมันทั้ง 3 ประเภท กับตัวแปรอิสระทั้งหมด 74 ตัวแปร ผลการศึกษาสามารถสรุปได้ดังนี้

1. การวิเคราะห์สหสัมพันธ์

จากการศึกษาความสัมพันธ์ของตัวแปรตามทั้ง 3 ตัวแปร และตัวแปรอิสระทั้ง 74 ตัวแปร พบว่ามีตัวแปรอิสระเพียง 33 ตัวแปรที่มีความสัมพันธ์กับราคาขายปลีกน้ำมันทั้ง 3 ประเภท ได้แก่ ตัวแปรราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (X_3), ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (X_4), ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น(X_5), ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (X_6), ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (X_7), ราคาน้ำมันดีเซลนำเข้าหน้าโรงกลั่น (X_8), ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (X_9), ราคาน้ำมันเบนซิน 91 ขายส่งหน้าโรงกลั่น (X_{10}), ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น (X_{11}), อัตราภาษีส่งออกน้ำมันเบนซิน 95 (X_{19}), อัตราภาษีส่งออกน้ำมันเบนซิน 91 (X_{20}), อัตราภาษีนำเข้าน้ำมันเบนซิน 95 (X_{22}), อัตราภาษีนำเข้าน้ำมันเบนซิน 91 (X_{23}), ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (X_{25}), ราคาน้ำมันดิบในตลาดโลก (โอมาน) (X_{26}), ราคาน้ำมันเบนซิน 95 ในตลาดจรสิงคโปร์ (X_{27}), ราคาน้ำมันเบนซิน 92 ในตลาดจรสิงคโปร์ (X_{28}), ราคาน้ำมันดีเซลในตลาดจรสิงคโปร์ (X_{29}), ปริมาณการใช้ไฟฟ้า (X_{30}), ประชากรรวมในกรุงเทพฯ (X_{31}) , ผลิตภัณฑ์มวลรวมในประเทศ (X_{34}) , อัตราดอกเบี้ยเงินกู้ (X_{36}), กองทุนน้ำมัน (X_{37}), ปริมาณการผลิตน้ำมัน LPG (X_{40}), ปริมาณการส่งออกน้ำมันเบนซิน 95 (X_{47}), ปริมาณการส่งออกน้ำมันเบนซิน 91 (X_{48}), จำนวนรถจักรยานยนต์ที่จดทะเบียนในกรุงเทพฯ (X_{52}), มูลค่าการนำเข้าน้ำมันดิบ (X_{60}), มูลค่าการส่งออกรถยนต์ (X_{65}), ปริมาณการจำหน่ายรถจักรยานยนต์ในประเทศ (X_{68}), มูลค่าการส่งออกผลิตภัณฑ์ปิโตรเลียม (X_{69}), ปริมาณการผลิตรถจักรยานยนต์ (X_{71}), ปริมาณการผลิตน้ำมันดิบ (x_{72})

และจากการศึกษาสหสัมพันธ์แคนนอลนิคอลพบว่า ที่ระดับนัยสำคัญ 0.05 ตัวแปรอิสระทั้ง 33 ตัว มีความสัมพันธ์กับราคาน้ำมันขายปลีกทั้ง 3 ประเภท และสามารถหาค่าสหสัมพันธ์แคนนอลนิคอลได้ 3 คู่ ซึ่งมีค่าเท่ากับ 0.999, 0.976, และ 0.878 ตามลำดับ ซึ่งพบว่า ตัวแปรแคนนอลนิคอลทั้ง 3 คู่นี้ ตัวแปรแคนนอลนิคอลคู่ที่ 1 อธิบายความผันแปรร่วมกันได้สูงสุดร้อยละ 97.3 รองลงมาคือ ตัวแปรแคนนอลนิคอลคู่ที่ 2 และคู่ที่ 3 ซึ่งสามารถอธิบายความผันแปรแคนนอลนิคอลได้ร้อยละ 2.31 และร้อยละ0.38 ตามลำดับ ดังนั้นจึงเลือกใช้ตัวแปรแคนนอลนิคอลคู่ที่ 1 ที่สามารถอธิบายความผันแปรร่วมกันได้สูงสุด ดังนั้นจากตัวแปรแคนนอลนิคอลคู่ที่ 1 จึงได้ว่าปัจจัยที่เป็นสาเหตุหลักของการเปลี่ยนแปลงราคาน้ำมันขายปลีกทั้ง 3 ประเภทประกอบด้วยตัวแปรจำนวน 16 ตัวแปร ดังนี้

X_3 = ราคาน้ำมันเบนซิน 95 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_4 = ราคาน้ำมันเบนซิน 91 ส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_5 = ราคาน้ำมันดีเซลส่งออกหน้าโรงกลั่น (บาท/ลิตร)
X_6 = ราคาน้ำมันเบนซิน 95 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_7 = ราคาน้ำมันเบนซิน 91 นำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_8 = ราคาน้ำมันดีเซลนำเข้าหน้าโรงกลั่น (บาท/ลิตร)
X_9 = ราคาน้ำมันเบนซิน 95 ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{10} = ราคาน้ำมันเบนซิน 91ขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{11} = ราคาน้ำมันดีเซลขายส่งหน้าโรงกลั่น (บาท/ลิตร)
X_{25} = ราคาน้ำมันดิบในตลาดโลก (ดูไบ) (เหรียญสหรัฐ/บาร์เรล)
X_{26} = ราคาน้ำมันดิบในตลาดโลก (โอมาน) (เหรียญสหรัฐ/บาร์เรล)
X_{27} = ราคาน้ำมันเบนซิน 95ในตลาดสิงคโปร์ (เหรียญสหรัฐ/ บาร์เรล)
X_{29} = ราคาน้ำมันดีเซลในตลาดสิงคโปร์ (เหรียญสหรัฐ/ บาร์เรล)
X_{36} = อัตราดอกเบี้ยเงินกู้
X_{65} = มูลค่าการส่งออกรถยนต์ (ล้านบาท) และ
X_{72} = ปริมาณการผลิตน้ำมันดิบ (1,000บาร์เรล)

2. การวิเคราะห์การถดถอย

เมื่อได้ศึกษาความสัมพันธ์แล้วพบว่าตัวแปรอิสระทั้ง 33 ตัวมีความสัมพันธ์กับ ตัวแปรตามทั้ง 3 ตัว จึงใช้ตัวแปรอิสระทั้งหมด 33 ตัว ในการวิเคราะห์การถดถอย และผลการตรวจสอบข้อสมมติเบื้องต้นพบว่า ข้อมูลชุดนี้ความคลาดเคลื่อนมีการแจกแจงปกติหลายตัวแปร และความคลาดเคลื่อนมีความแปรปรวนคงที่ อย่างไรก็ตามพบว่าเกิดปัญหา Multicollinearity เนื่องจากข้อมูลประกอบด้วยตัวแปรอิสระจำนวนทั้งสิ้น 33 ตัวแปร ดังนั้นผู้วิจัยจึงได้แก้ปัญหา Multicollinearity โดยใช้เทคนิคการวิเคราะห์ปัจจัย และพบว่าจากการวิเคราะห์ปัจจัย โดยทำการหมุนแบบ Orthogonal ด้วยวิธี Varimax พบว่าจากตัวแปรอิสระทั้งหมด 33 ตัวแปร สามารถสกัดปัจจัยได้ 3 ปัจจัย และทั้ง 3 ปัจจัยนี้อธิบายความผันแปรของตัวแปรอิสระทั้งหมดได้ร้อยละ 91.041 ซึ่งปัจจัยที่ 1 สามารถอธิบายความผันแปรได้ ร้อยละ 44.706 ผู้วิจัยเรียกปัจจัยที่ 1 นี้ว่า **ปัจจัยด้านราคา** ปัจจัยที่ 2 สามารถอธิบายความผันแปรได้ ร้อยละ 24.684 ผู้วิจัยเรียกปัจจัยที่ 2 นี้ว่า **ปัจจัยด้านการผลิตและส่งออกน้ำมัน** และปัจจัยที่ 3 สามารถอธิบายความผันแปรได้ร้อยละ 21.651 ผู้วิจัยเรียกปัจจัยที่ 3 นี้ว่า **ปัจจัยด้านเศรษฐกิจ**

นำปัจจัยทั้ง 3 กับตัวแปรตามราคาน้ำมันขายปลีกทั้ง 3 ประเภทมาวิเคราะห์ พบว่าสมการถดถอยที่ใช้ในการพยากรณ์ ราคาขายปลีกน้ำมันดีเซล น้ำมันเบนซิน 91 และน้ำมันเบนซิน 95 มีดังนี้

สมการของราคาขายปลีกน้ำมันดีเซล มีค่า R^2 = 0.962 และมีค่าความคลาดเคลื่อน = 1.027

$$\hat{Y}_1 = 17.547+4.239\text{Factor 1}+2.057\text{Factor 2}+1.733\text{Factor 3}$$

สมการของราคาขายปลีกน้ำมันเบนซิน 91 มีค่า R^2 = 0.983 และมีค่าความคลาดเคลื่อน = 0.673

$$\hat{Y}_2 = 19.619+3.66\text{Factor 1}+2.536\text{Factor 2}+2.192\text{Factor 3}$$

สมการของราคาขายปลีกน้ำมันเบนซิน 95 มีค่า R^2 = 0.983 และมีค่าความคลาดเคลื่อน = 0.667

$$\hat{Y}_3 = 20.489+3.627\text{Factor 1}+2.504\text{Factor 2}+2.126\text{Factor 3}$$

เมื่อ Factor 1 คือ ปัจจัยด้านราคา, Factor 2 คือ ปัจจัยด้านการผลิตและส่งออกน้ำมัน และ Factor 3 คือ ปัจจัยด้านเศรษฐกิจ

โดยมีสมการการถดถอยของแต่ละปัจจัย สามารถคำนวณได้ดังนี้

Factor 1 (ปัจจัยด้านราคา) = $-0.003X_3-0.001X_4+0.005X_5-0.003X_6-0.001X_7+0.006X_8+0.029X_9+0.031X_{10}+0.066X_{11}+0.266X_{19}+0.137X_{20}+0.000X_{22}+0.144X_{23}+0.024X_{25}+0.017X_{26}+0.003X_{27}-0.041X_{28}+0.023X_{29}+0.037X_{30}-0.115X_{31}+0.153X_{34}+0.088X_{36}+0.003X_{40}+0.041X_{65}+0.032X_{69}+0.050X_{72}$

Factor 2 (ปัจจัยด้านการผลิตและส่งออกน้ำมัน) = $0.196X_{37}+0.054X_{40}+0.411X_{47}+0.290X_{48}$

Factor 3 (ปัจจัยด้านเศรษฐกิจ) = $0.171X_{52}+0.435X_{68}+0.370X_{71}$

ข้อเสนอแนะ

จากผลการศึกษาปัจจัยที่มีอิทธิพลต่อราคาน้ำมันขายปลีก โดยการวิเคราะห์การถดถอยแบบหลายตัวแปร มีข้อเสนอแนะดังต่อไปนี้

ข้อเสนอแนะเพื่อนำผลการวิจัยไปใช้

1. จากการวิจัย พบว่า ปัจจัยที่ส่งผลต่อราคาขายปลีกน้ำมันเชื้อเพลิงทั้ง 3 ประเภท มีปัจจัยด้านราคา ปัจจัยด้านการผลิตและส่งออกน้ำมัน และปัจจัยด้านเศรษฐกิจ ดังนั้นหน่วยงานทั้งภาครัฐและเอกชน ที่ทำหน้าที่ในการผลิต และจำหน่ายน้ำมันทั้ง 3 ประเภทดังกล่าว สามารถนำผลการศึกษาครั้งนี้ไปใช้ในการคาดการณ์ หรือพยากรณ์แนวโน้มราคาน้ำมันต่อไปในอนาคตได้

2. จากการศึกษาวิจัยในครั้งนี้พบว่า การมีจำนวนตัวแปรอิสระหลายตัว จะทำให้เกิดปัญหา Multicollinearity เทคนิคหนึ่งที่ผู้วิจัยสามารถนำมาแก้ปัญหาดังกล่าวได้ คือการวิเคราะห์ปัจจัย

ดังนั้นหากมีการศึกษางานวิจัยในลักษณะที่มีจำนวนตัวแปรอิสระหลายตัว อาจใช้เทคนิคดังกล่าวนี้เข้ามาช่วยขจัดปัญหา Multicollinearity ได้

ข้อเสนอแนะในการวิจัยครั้งต่อไป

1. ในการศึกษาครั้งนี้ใช้ข้อมูลเป็นรายเดือน ดังนั้นในการศึกษาครั้งต่อไปอาจจะใช้ข้อมูลเป็นรายวัน หรือรายสัปดาห์เนื่องจากราคาน้ำมันมีการเปลี่ยนแปลงไปในแต่ละวัน เพื่อให้มีความละเอียด ถูกต้อง แม่นยำในการพยากรณ์

2. ในการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวพบว่ามีความไม่เป็นอิสระกันหรือมีปัญหา Multicollinearity จึงได้ใช้เทคนิคการวิเคราะห์ปัจจัยมาช่วยแก้ปัญหาดังกล่าวซึ่งผลการสกัดปัจจัยที่ได้ยังมีความซับซ้อนอยู่ จึงต้องทำการหมุนแกนปัจจัยซึ่งในงานวิจัยครั้งนี้ผู้วิจัยได้ใช้การหมุนแกนแบบ Orthogonal ด้วยวิธี Varimax ทำให้สามารถแปลความหมายของปัจจัยร่วมได้ชัดเจนขึ้น แต่เนื่องจากยังมีวิธีการหมุนแกนปัจจัยแบบอื่นๆอีกหลายวิธี เช่น วิธีQuartimax หรือ วิธีEquamax ซึ่งอาจทำให้การแปลความหมายของปัจจัยร่วมแตกต่างออกไปจากงานวิจัยนี้ดังนั้นในการศึกษาครั้งต่อไปผู้ศึกษาอาจใช้วิธีการหมุนแกนแบบอื่นๆมาช่วยในการศึกษาและทำการเปรียบเทียบผลการวิจัยต่อไปได้

3. การวิเคราะห์แนวโน้มของราคาน้ำมันเป็นเรื่องที่ค่อนข้างซับซ้อน และเปลี่ยนแปลงอยู่ตลอดเวลา เนื่องจากน้ำมันเป็นสินค้าที่ต้องการของผู้บริโภคทั่วโลก เพื่อนำไปใช้ทั้งในงานอุตสาหกรรมและเกษตรกรรมของประเทศ ดังนั้นจะเห็นว่า ในการซื้อขายน้ำมันนั้นมีปัจจัยหลากหลายเข้ามาเกี่ยวข้องทั้งทางตรงและทางอ้อม นอกเหนือจากปัจจัยด้านราคา ด้านการผลิตและการส่งออก และด้านเศรษฐกิจที่ได้ทำการศึกษาในครั้งนี้ เช่น สถานการณ์ทางการเมือง ความไม่สงบภายในประเทศ เป็นต้น ดังนั้นในการศึกษาครั้งต่อไปอาจนำตัวแปรเกี่ยวกับสถานการณ์ทางการเมือง ความไม่สงบภายในประเทศเข้ามาร่วมเป็นตัวแปรอิสระในการศึกษาด้วย

4. ในการศึกษาครั้งนี้ใช้การวิเคราะห์การถดถอย ซึ่งทำให้ไม่สามารถมองเห็นแนวโน้มของการเปลี่ยนแปลงของราคาน้ำมันตามเวลาที่เปลี่ยนแปลงไป ดังนั้นในการศึกษาครั้งต่อไปอาจจะใช้การวิเคราะห์อนุกรมเวลาช่วยในการวิเคราะห์ทางสถิติ ซึ่งจะทำให้เห็นแนวโน้มของการเปลี่ยนแปลงดังกล่าวได้

เอกสารและสิ่งอ้างอิง

กัลยา วานิชย์บัญชา. 2546. **การวิเคราะห์สถิติสำหรับการบริหารและการวิจัย**. พิมพ์ครั้งที่ 6. จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

_____. 2548. **การวิเคราะห์ข้อมูลหลายตัวแปร**. บริษัทธรรมสาร, กรุงเทพฯ.

เกียรติชัย สงอินทร์. 2536. **การศึกษาระบบราคาน้ำมันลอยตัวในประเทศไทย**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์ .

ณรงค์ ธนาวิภาส. 2548. **หลักเศรษฐศาสตร์.** บริษัทวิทยพัฒน์, กรุงเทพฯ.

ทรงศิริ แต้สมบัติ. 2549. **การพยากรณ์เชิงปริมาณ.** มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

ธานินทร์ ตั้งทวีพัฒนกุล. 2537. **การวิเคราะห์การนำเข้าน้ำมันดิบและผลิตภัณฑ์น้ำมันของประเทศไทย**.วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

ธีระศักดิ์ อุรัจนานนท์. 2546. **ความน่าจะเป็นและสถิติประยุกต์ เล่ม 2**. บริษัทสกายบุ๊กส์, กรุงเทพฯ.

นุชชฎา ถิ่นกมุท. 2547. **โครงสร้างและการกำหนดราคาค้าปลีกน้ำมันเชื้อเพลิง**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยรามคำแหง.

ปรเมสฐ์ บุญศรี. 2533. **ผลกระทบของนโยบายการค้าภายในและภายนอกประเทศที่มีผลต่อระดับราคาผลิตภัณฑ์ปิโตรเลียม : กรณีศึกษาของประเทศไทยในช่วงปี พ.ศ.2503-2530.** วิทยานิพนธ์ ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

ปิยสวัสดิ์ อัมระนันท์. 2538. **นโยบายพลังงานและการพัฒนาเศรษฐกิจไทยในศักยภาพและอนาคตของเศรฐษกิจไทย**. จุฬาลงกรณ์มหาวิทยาลัย, กรุงเทพฯ.

พีรณัฐ แดงสกุล. 2540. **การวิเคราะห์การกำหนดราคาน้ำมันในระดับค้าปลีก**.วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

รัตนาภรณ์ พงสะมั่ง. 2547. **การศึกษาความสัมพันธ์ระหว่างความฉลาดทางอารมณ์ สภาพแวดล้อมในมหาวิทยาลัย บรรยากาศในครอบครัว กับวิธีเผชิญปัญหาของนักเรียนระดับประกาศนียบัตรวิชาชีพวิทยาลัยเทคนิค สังกัดกรมอาชีวศึกษา**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยศรีนครินทรวิโรฒ.

วาสนา สิงหโกวินทร์. 2523**. การศึกษาระบบตลาดและขนส่งน้ำมันเชื้อเพลิงในประเทศไทย**. . วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

ศราวุธ ทองเนื้อห้า.2546. **การวิเคราะห์ปัจจัยที่กำหนดปริมาณการผลิตก๊าซปิโตรเลียมเหลวในประเทศไทย**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยรามคำแหง.

ศุภกร ขำล้ำเลิศ. 2541. **โครงสร้างทางการค้าและปัจจัยที่มีผลกระทบต่อราคาน้ำมันดีเซลในประเทศไทย**. วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยเกษตรศาสตร์.

สมฤทธิ์ ทรัพย์พุ่ม. 2539. **การวิเคราะห์เชิงสถิติทุนอุดหนุนการวิจัยของสำนักงานคณะกรรมการวิจัยแห่งชาติ**. วิทยานิพนธ์ปริญญาโท : มหาวิทยาลัยเกษตรศาสตร์.

สาธิต เถรแก้ว. 2541. **ปัจจัยที่มีอิทธิพลต่อการนำเข้าน้ำมันดีเซลและน้ำมันเบนซินของประเทศไทย**.วิทยานิพนธ์ปริญญาโท, มหาวิทยาลัยรามคำแหง.

สำนักงานคณะกรรมการนโยบายพลังงานแห่งชาติ. 2547. **สถานการณ์การค้าปลีกของสถานีบริการน้ำมันเชื้อเพลิง**. แหล่งที่มา : http://www.eppo.go.th/vrs/vrs50-07-oilretail.html, 6 กุมภาพันธ์ 2547.

สุชาติ ประสิทธิ์รัฐสินธุ์. 2540. **เทคนิคการวิเคราะห์ตัวแปรหลายตัวสำหรับการวิจัยทางสังคมศาสตร์และพฤติกรรมศาสตร์ หลักการ วิธีการ และการประยุกต์**. สถาบันบัณฑิตพัฒนบริหารศาสตร์, กรุงเทพฯ.

สุภิญญา พรหมภัทร. 2538. **การประเมินผลของนโยบายลอยตัวราคาน้ำมัน**. วิทยานิพนธ์ปริญญาโท, จุฬาลงกรณ์มหาวิทยาลัย.

สุรินทร์ นิยมางกูร. 2548. **สถิติวิจัย**. พิมพ์ครั้งที่ 2. มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ.

โสภณ สุภาพงษ์. 2533. “ระบบราคาน้ำมันลอยตัว ตัวแปรใหม่ด้านต้นทุนและการปรับตัวของภาคเอกชน”. **จุฬาลงกรณ์วารสาร.** 2(กรกฎาคม-ธันวาคม 2533) : 33-54.

Ashakul, Treera.1985. **The estimation of petroleum product demand function**. Paper Prepared for National Economic and Social Development Bord. Bangkok: National Economic and Social Development.

Johnson, R.A. and Wichem, D.W. 2002. **Applied multivariate statistical analysis**. 5 th Edition. N.J. : Prentice Hall Inc.

Khattree, Ravindra and Dayanand N. Naik, 1999.**Applied Multivariate Statistics with SAS Software, Second Edition**. Cary, NC: SAS Institute Inc.

ภาคผนวก

ภาคผนวก ก

ข้อมูลตัวอย่างใหม่ที่ใช้ในการพยากรณ์

ตารางผนวกที่ ก1 ข้อมูลตัวอย่างใหม่ที่ใช้ในการพยากรณ์

เดือน	Y_1	Y_2	Y_3	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
มกราคม(2550)	22.76	24.91	25.71	14.4	13.96	15.57	14.79	14.35	15.64	23.58	22.89	21.15
กุมภาพันธ์	23.1	24.83	25.63	15.4	14.94	16.41	14.9	14.45	16.25	24.61	23.93	22.06
มีนาคม	23.62	26.54	27.34	17.3	16.86	16.84	16.93	16.49	16.57	26.62	25.95	22.47
เมษายน	24.83	27.88	28.68	18.72	18.35	18.25	18.27	17.83	17.98	28.13	27.46	23.97
พฤษภาคม	25.34	29.19	29.99	19.67	19.25	18.42	19.58	19.15	18.36	29.2	28.53	24.17
มิถุนายน	25.34	29.2	30.01	18.84	18.41	18.41	18.99	18.56	18.44	28.3	27.64	24.19
กรกฎาคม	25.61	29.05	29.85	18.5	18.08	18.8	18.7	18.28	18.62	27.97	27.31	24.6
สิงหาคม	25.44	27.73	28.53	16.98	16.56	18.44	17.14	16.73	18.49	26.72	25.99	24.22
กันยายน	26.58	28.59	29.21	18.12	17.7	20.1	17.74	17.32	19.59	28.14	27.37	25.99
ตุลาคม	27.47	29.51	30.31	19.46	19.04	21.05	18.95	18.53	20.65	29.55	28.77	26.99

ตารางผนวกที่ ก1 (ต่อ)

เดือน	X_{19}	X_{20}	X_{22}	X_{23}	X_{25}	X_{26}	X_{27}	X_{28}	X_{29}	X_{30}	X_{31}	X_{34}
มกราคม(2550)	3.46	3.26	3.46	3.26	11.77	11.87	13.99	13.7	15	10203	-	2082.7
กุมภาพันธ์	3.46	3.26	3.46	3.26	12.57	12.62	15.06	14.82	15.92	10041	-	2082.7
มีนาคม	3.46	3.26	3.46	3.26	13.01	13.1	16.95	16.71	16.25	12028	-	2082.7
เมษายน	3.46	3.26	3.46	3.26	14.08	14.19	18.38	18.2	17.66	11348	-	2027.9
พฤษภาคม	3.46	3.26	3.46	3.26	14.12	14.21	19.4	19.22	17.86	11803	-	2027.9
มิถุนายน	3.46	3.26	3.46	3.26	14.36	14.42	18.51	18.3	17.86	11973	-	2027.9
กรกฎาคม	3.47	3.27	3.47	3.27	14.79	14.92	18.16	17.95	18.25	11817	-	2068.2
สิงหาคม	3.83	3.57	3.83	3.57	14.54	14.75	16.65	16.41	17.91	11835	-	2068.2
กันยายน	4	3.7	4	3.7	15.86	15.91	17.84	17.59	19.62	10124	-	2068.2
ตุลาคม	4	3.7	4	3.7	16.71	16.8	19.13	18.86	20.51	-	-	-

ตารางผนวกที่ ก1 (ต่อ)

เดือน	X_{36}	X_{37}	X_{40}	X_{47}	X_{48}	X_{52}	X_{60}	X_{65}	X_{68}	X_{69}	X_{71}	X_{72}
มกราคม(2550)	4.99	36696	690.2	92.9	33.8	31751	40670.78	7728.43	140	12429.38	180	3724
กุมภาพันธ์	4.65	33203	585.6	89.9	33.4	30129	56660.56	9092.61	124	11080.89	156	3751
มีนาคม	4.18	30058	574.5	126.6	35.5	35421	43950.75	7433.22	128	13458.01	152	4463
เมษายน	3.99	26336	615.7	129.8	36.8	26122	51372.19	6072.46	99	12615.47	129	4232
พฤษภาคม	3.9	21868	643.6	151.5	34.3	33809	61067.78	9004.39	154	15903.13	179	4114
มิถุนายน	3.86	18899	647.1	101.7	32.4	33864	54583.17	1460.01	143	14725.39	171	4090
กรกฎาคม	3.86	15616	643.9	50.4	30.1	30460	59837.65	8988.27	122	12911.4	160	4175
สิงหาคม	3.87	13100	684.5	85.5	33.8	31421	67342.6	9313.42	127	16528.84	155	3961
กันยายน	4.03	7968	659.5	84.2	31.8	28423	53749.67	9826.75	118	17766.11	147	3981
ตุลาคม	4.03	4291	676.6	10.4	-	30254	61540.09	11088.44	135	20539.94	160	3963

หมายเหตุ ตัวแปร X_{31} เป็นตัวแปรประชากรรวมในกรุงเทพมหานคร ซึ่งยังไม่ได้ทำการรวบรวม

ภาคผนวก ข

ผลการวิเคราะห์ข้อมูล โดยใช้โปรแกรม SAS

ภาคผนวก ข1 ผลการทดสอบการแจกแจงปกติหลายตัวแปร ด้วยวิธีของ Mardia

S

25.76796 24.10033 23.805136

24.10033 24.645471 24.299976

23.805136 24.299976 23.963476

S_INV

0.5673905 5.1925649 -5.829121

5.1925649 281.68685 -290.8006

-5.829121 -290.8006 300.71642

TESTS:

	BETA1HAT	KAPPA1	PVALSKEW
Based on skewness:	5.2582369	52.582369	**0.089007**

	BETA2HAT	KAPPA2	PVALKURT
Based on kurtosis:	12.42955	-1.817582	**0.069128**

ภาคผนวก ข2 ผลการวิเคราะห์สหสัมพันธ์แคนนอลนิคอล

Canonical Correlation Analysis

	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation
1	0.999415	0.999097	0.000152	0.998830
2	0.976231	0.963836	0.006115	0.953027
3	0.878018	0.816498	0.029824	0.770916

Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)

	Eigenvalue	Difference	Proportion	Cumulative
1	853.9012	833.6125	0.9730	0.9730
2	20.2887	16.9234	0.0231	0.9962
3	3.3652		0.0038	1.0000

Test of H0: The canonical correlations in the current row and all that follow are zero

	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.00001259	33.40	96	75.734	<.0001
2	0.01076083	7.25	62	52	<.0001
3	0.22908374	3.03	30	27	0.0024

Canonical Correlation Analysis

Multivariate Statistics and F Approximations

S=3 M=14 N=11.5

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.00001259	33.40	96	75.734
Pillai's Trace	2.72277317	8.29	96	81
Hotelling-Lawley Trace	877.55504807	217.85	96	57.339
Roy's Greatest Root	853.90116997	720.48	32	27

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

Canonical Structure

Correlations Between the VAR Variables and Their Canonical Variables

	V1	V2	V3
Y_1	0.9878	0.1505	0.0400
Y_2	0.9894	-0.1441	0.0181
Y_3	0.9905	-0.1374	0.0068

Correlations Between the WITH Variables and Their Canonical Variables

	W1	W2	W3
X_1	0.9498	-0.2165	0.0592
X_2	0.9505	-0.2131	0.0652
X_3	0.9667	-0.2012	0.0509
X_4	0.9457	-0.2398	0.1043
X_5	0.9492	-0.2279	0.0957
X_6	0.9676	-0.1979	0.0450
X_7	0.9849	-0.1429	-0.0034
X_8	0.9833	-0.1292	-0.0119
X_9	0.9905	0.1119	0.0187
X_{10}	0.8928	0.2395	0.0294
X_{11}	0.8873	0.2383	0.0260
X_{12}	0.8928	0.2395	0.0294
X_{13}	0.8651	0.2430	0.0304
X_{14}	0.9707	-0.1146	0.1130
X_{15}	0.9666	-0.1406	0.1097
X_{16}	0.9368	-0.2461	0.1703

X_{17}	0.8658	-0.3502	0.1943
X_{18}	0.9718	-0.1821	0.0885
X_{19}	0.8348	-0.1495	0.1785
X_{20}	-0.8541	-0.2681	0.1250
X_{21}	0.8154	0.1778	0.0600
X_{22}	0.9243	0.1056	-0.0980
X_{23}	0.8330	-0.2537	0.0033
X_{24}	0.7521	-0.2798	0.0937
X_{25}	0.4527	-0.2385	-0.1076
X_{26}	0.6728	-0.3004	-0.0883
X_{27}	0.5659	-0.3504	0.1908
X_{28}	0.8976	-0.0979	0.1554
X_{29}	0.9449	0.0061	0.1103
X_{30}	0.5123	-0.2700	0.4338
X_{31}	0.8842	-0.1299	0.0287
X_{32}	0.6261	-0.3080	0.3788
X_{33}	0.9055	0.1055	-0.0406

ตารางผนวก ข3 ผลการวิเคราะห์สหสัมพันธ์ระหว่างตัวแปรอิสระ 33 ตัว

	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{19}	X_{20}	X_{22}	X_{23}	X_{25}	X_{26}	X_{27}	X_{28}
X_3	1.000	1.000	0.985	0.992	0.923	0.980	0.980	0.973	0.922	0.744	0.737	0.744	0.706	0.962	0.964	0.974	0.949
		(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_4		1.000	0.984	0.993	0.993	0.979	0.980	0.973	0.923	0.747	0.740	0.747	0.709	0.964	0.966	0.975	0.949
			(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_5			1.000	0.980	0.980	0.997	0.983	0.977	0.935	0.782	0.773	0.782	0.747	0.978	0.979	0.970	0.938
				(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_6				1.000	0.999	0.979	0.974	0.966	0.914	0.740	0.733	0.740	0.706	0.959	0.961	0.977	0.953
					(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_7					1.000	0.979	0.977	0.969	0.920	0.744	0.738	0.744	0.711	0.959	0.961	0.977	0.951
						(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_8						1.000	0.980	0.974	0.935	0.784	0.775	0.784	0.750	0.976	0.979	0.970	0.937
							(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_9							1.000	0.997	0.963	0.835	0.829	0.835	0.801	0.975	0.975	0.966	0.917
								(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

ตารางผนวก ข3 (ต่อ)

	X_{29}	X_{30}	X_{31}	X_{34}	X_{36}	X_{37}	X_{40}	X_{47}	X_{48}	X_{52}	X_{60}	X_{65}	X_{68}	X_{69}	X_{71}	X_{72}
X_3	0.975	0.847	-0.722	0.671	0.818	0.862	0.751	0.532	0.758	0.668	0.897	0.902	0.602	0.868	0.709	0.847
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_4	0.975	0.849	-0.725	0.675	0.819	0.860	0.750	0.527	0.754	0.668	0.896	0.904	0.599	0.867	0.707	0.848
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_5	0.992	0.855	-0.727	0.705	0.852	0.891	0.796	0.538	0.752	0.669	0.921	0.914	0.597	0.879	0.174	0.847
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_6	0.976	0.854	-0.718	0.675	0.814	0.855	0.767	0.517	0.753	0.674	0.898	0.906	0.614	0.872	0.717	0.829
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_7	0.976	0.852	-0.725	0.680	0.818	0.853	0.763	0.513	0.753	0.668	0.899	0.907	0.608	0.873	0.711	0.823
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_8	0.991	0.854	-0.732	0.707	0.854	0.890	0.804	0.537	0.746	0.669	0.923	0.917	0.598	0.884	0.714	0.852
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_9	0.984	0.832	-0.804	0.759	0.887	0.861	0.766	0.495	0.729	0.612	0.894	0.929	0.542	0.886	0.657	0.887
	(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

ตารางผนวก ข3 (ต่อ)

	X_{10}	X_{11}	X_{19}	X_{20}	X_{22}	X_{23}	X_{25}	X_{26}	X_{27}	X_{28}	X_{29}	X_{30}	X_{31}	X_{34}	X_{36}	X_{37}	X_{40}
X_{10}	1.000	0.966	0.837	0.832	0.837	0.803	0.972	0.972	0.961	0.911	0.980	0.826	-0.808	0.759	0.890	0.861	0.766
		(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{11}		1.000	0.904	0.899	0.904	0.876	0.949	0.942	0.907	0.828	0.944	0.803	-0.877	0.823	0.920	0.794	0.715
			(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{19}			1.000	0.998	1.000	0.988	0.842	0.829	0.766	0.641	0.814	0.728	-0.891	0.915	0.909	0.636	0.640
				(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{20}				1.000	0.998	0.990	0.834	0.821	0.760	0.631	0.808	0.724	-0.904	0.924	0.911	0.623	0.627
					(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{22}					1.000	0.988	0.842	0.829	0.766	0.641	0.814	0.728	-0.891	0.915	0.909	0.636	0.640
						(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{23}						1.000	0.805	0.792	0.729	0.594	0.786	0.711	-0.872	0.929	0.898	0.590	0.597
							(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{25}							1.000	0.999	0.972	0.939	0.978	0.851	-0.781	0.741	0.880	0.897	0.802
								(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

ตารางผนวก ข3 (ต่อ)

	X_{26}	X_{27}	X_{28}	X_{29}	X_{30}	X_{31}	X_{34}	X_{36}	X_{37}	X_{40}	X_{47}	X_{48}	X_{52}	X_{60}	X_{65}	X_{68}	X_{69}	X_{71}	X_{72}
X_{26}	1.000	0.982	0.947	0.978	0.847	-0.770	0.725	0.874	0.907	0.808	0.522	0.734	0.646	0.919	0.939	0.584	0.863	0.716	0.873
		(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{27}		1.000	0.978	0.973	0.851	-0.718	0.693	0.810	0.871	0.793	0.468	0.722	0.666	0.889	0.915	0.621	0.864	0.739	0.834
			(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{28}			1.000	0.926	0.791	-0.614	0.530	0.715	0.887	0.786	0.573	0.767	0.734	0.858	0.862	0.659	0.813	0.770	0.771
				(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{29}				1.000	0.871	-0.744	0.763	0.865	0.861	0.794	0.478	0.710	0.608	0.918	0.913	0.584	0.883	0.696	0.843
					(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{30}					1.000	-0.585	0.729	0.693	0.677	0.633	0.312	0.524	0.471	0.843	0.804	0.548	0.754	0.653	0.740
						(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.015)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{31}						1.000	0.798	0.916	-0.613	-0.535	-0.310	-0.461	-0.338	-0.646	-0.847	-0.321	0.766	-0.339	-0.809
							(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.016)**	**(0.000)**	**(0.008)**	**(0.000)**	**(0.000)**	(0.075)	**(0.000)**	**(0.008)**	**(0.000)**
X_{34}							1.000	0.822	0.479	0.539	0.041	0.285	0.182	0.664	0.745	0.282	0.715	0.376	0.731
								(0.000)	**(0.000)**	**(0.000)**	(0.758)	**(0.028)**	(0.164)	**(0.000)**	**(0.000)**	**(0.029)**	**(0.000)**	**(0.003)**	**(0.000)**
X_{36}								1.000	0.793	0.635	0.407	0.607	0.399	0.789	0.871	0.328	0.785	0.458	0.797
									(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.002)**	**(0.000)**	**(0.000)**	**(0.011)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{37}									1.000	0.759	0.673	0.853	0.706	0.858	0.823	0.533	0.718	0.681	0.722
										(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000๗)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**

ตารางผนวก ข3 (ต่อ)

	X_{40}	X_{47}	X_{48}	X_{52}	X_{60}	X_{65}	X_{68}	X_{69}	X_{71}	X_{72}
X_{40}	1.000	0.545	0.632	0.615	0.808	0.719	0.548	0.783	0.643	0.709
		(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{47}		1.000	0.691	0.539	0.492	0.452	0.319	0.468	0.409	0.437
			(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.013)**	**(0.000)**	0.001	**(0.000)**
X_{48}			1.000	0.621	0.708	0.677	0.479	0.622	0.593	0.627
				(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{52}				1.000	0.631	0.608	0.699	0.576	0.774	0.560
					(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{60}					1.000	0.877	0.619	0.820	0.730	0.792
						(0.000)	**(0.000)**	**(0.000)**	**(0.000)**	**(0.000)**
X_{65}						1.000	0.538	0.854	0.659	0.886
							(0.000)	**(0.000)**	**(0.000)**	**(0.000)**
X_{68}							1.000	0.435	0.963	0.537
								(0.001)	**(0.000)**	**(0.000)**
X_{69}								1.000	0.546	0.789
									(0.000)	**(0.000)**
X_{71}									1.000	0.625
										(0.000)
X_{72}										1.000

หมายเหตุ ค่าในวงเล็บคือ ค่า p-value ที่ระดับนัยสำคัญ 0.05

ภาคผนวก ข4 ผลวิเคราะห์ปัจจัย โดยการหมุนแบบ Varimax

Rotation Method: Varimax

Orthogonal Transformation Matrix

	1	2	3
1	0.71675	0.51456	0.47063
2	-0.69624	0.49039	0.52419
3	0.03894	-0.70338	0.70975

Rotated Factor Pattern

	Factor1	Factor2	Factor3
X_1	0.60527	0.57866	0.50788
X_2	0.60999	0.57436	0.50628
X_3	0.63676	0.56977	0.49676
X_4	0.60230	0.56482	0.52529
X_5	0.60940	0.56381	0.51761
X_6	0.63986	0.56608	0.49688
X_7	0.71452	0.53656	0.42333
X_8	0.72147	0.53395	0.40765
X_9	0.81320	0.42729	0.34387
X_{10}	0.94850	0.18238	0.17663
X_{11}	0.95605	0.16650	0.16898
X_{12}	0.94850	0.18238	0.17663
X_{13}	0.95661	0.10442	0.17728
X_{14}	0.69877	0.52592	0.46476
X_{15}	0.67971	0.54215	0.47514
X_{16}	0.62005	0.53195	0.54028
X_{17}	0.45832	0.62145	0.58526
X_{18}	0.68987	0.50445	0.48888
X_{19}	0.64004	0.27018	0.54908

X_{20}	-0.87352	-0.31791	-0.03774
X_{21}	0.94258	-0.00855	0.20477
X_{22}	0.84622	0.43123	0.13922
X_{23}	0.43496	0.74048	0.39524
X_{24}	0.45853	0.51269	0.45794
X_{25}	0.02129	0.87071	0.10510
X_{26}	0.21509	0.82155	0.32016
X_{27}	0.06907	0.53901	0.67258
X_{28}	0.59088	0.48747	0.53417
X_{29}	0.72636	0.46209	0.41137
X_{30}	0.10790	0.15390	0.93034
X_{31}	0.66844	0.49773	0.33728
X_{32}	0.20689	0.27560	0.90016
X_{33}	0.71967	0.38844	0.38940

Variance Explained by Each Factor

Factor1	Factor2	Factor3
14.753048	8.145940	7.144825

Final Communality Estimates: Total = 30.043814

Standardized Scoring Coefficients

	Factor1	Factor2	Factor3
X_1	-0.00290	0.05382	0.02548
X_2	-0.00141	0.05157	0.02570
X_3	0.00471	0.04921	0.01999
X_4	-0.00318	0.04113	0.03970
X_5	-0.00117	0.04236	0.03537
X_6	0.00569	0.04689	0.02105
X_7	0.02858	0.04759	-0.01427

X_8	0.03142	0.05072	-0.02232
X_9	0.06584	0.00624	-0.02768
X_{10}	0.26625	-0.15961	-0.08990
X_{11}	0.13701	-0.08632	-0.04427
X_{12}	0.00000	0.00000	0.00000
X_{13}	0.14390	-0.12124	-0.01885
X_{14}	0.02367	0.02995	0.01272
X_{15}	0.01747	0.03711	0.01431
X_{16}	0.00272	0.01695	0.05739
X_{17}	-0.04061	0.06736	0.06425
X_{18}	0.02270	0.01177	0.03359
X_{19}	0.03706	-0.12357	0.14915
X_{20}	-0.11475	-0.04624	0.15885
X_{21}	0.15274	-0.18719	0.03524
X_{22}	0.08796	0.07364	-0.14088
X_{23}	-0.04360	0.19601	-0.07554
X_{24}	-0.01711	0.05433	0.03317
X_{25}	-0.10951	0.41131	-0.24070
X_{26}	-0.08652	0.28974	-0.12513
X_{27}	-0.10737	0.04204	0.17052
X_{28}	0.00336	-0.00043	0.07158
X_{29}	0.04063	0.01177	0.00362
X_{30}	-0.07513	-0.24877	0.43535
X_{31}	0.03204	0.06225	-0.04326
X_{32}	-0.06962	-0.18762	0.36993
X_{33}	0.05009	-0.01800	0.01740

ภาคผนวก ข5 ผลการวิเคราะห์การถดถอยแบบหลายตัวแปร

Number of Observations Read	60
Number of Observations Used	60

Dependent Variable: y1

Source	DF	Sum of Squares	Mean Square
Model	3	1486.995291	495.665097
Error	56	59.082327	1.055042
Corrected Total	59	1546.077618	

Source	F Value	Pr > F
Model	469.81	<.0001

R-Square	Coeff Var	Root MSE	y1 Mean
0.961786	5.853664	1.027152	17.54717

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	17.54716667	0.13260477	132.33	<.0001
Factor1	4.23903013	0.13372382	31.70	<.0001
Factor2	2.05694678	0.13372382	15.38	<.0001
Factor3	1.73288878	0.13372382	12.96	<.0001

Dependent Variable: y2

Source	DF	Sum of Squares	Mean Square
Model	3	1453.359701	484.453234
Error	56	25.368584	0.453010
Corrected Total	59	1478.728285	

Source	F Value	Pr > F
Model	1069.41	<.0001

R-Square	Coeff Var	Root MSE	y2 Mean
0.982844	3.430569	0.673060	19.61950

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	19.61950000	0.08689174	225.79	<.0001
Factor1	3.66024033	0.08762501	41.77	<.0001
Factor2	2.53597344	0.08762501	28.94	<.0001
Factor3	2.19196138	0.08762501	25.02	<.0001

Dependent Variable: y3

Source	DF	Sum of Squares	Mean Square
Model	3	1412.870576	470.956859
Error	56	24.937989	0.445321
Corrected Total	59	1437.808565	

Source	F Value	Pr > F
Model	1057.57	<.0001

R-Square	Coeff Var	Root MSE	y3 Mean
0.982656	3.257066	0.667324	20.48850

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Intercept	20.48850000	0.08615115	237.82	<.0001
Factor1	3.62702359	0.08687818	41.75	<.0001
Factor2	2.50402411	0.08687818	28.82	<.0001
Factor3	2.12638710	0.08687818	24.48	<.0001

Multivariate Test: test_all

Error Matrix (E)

59.082326954	-1.288932107	-0.111292094
-1.288932107	25.368584312	25.072927553
-0.111292094	25.072927553	24.937988975

Hypothesis Matrix (H)

1486.9952914	1447.3087471	1428.4194371
1447.3087471	1453.3597007	1432.9256274
1428.4194371	1432.9256274	1412.870576

Multivariate Statistics and F Approximations

S=3 M=-0.5 N=26

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.00516386	112.63	9	131.57
Pillai's Trace	1.59585324	21.22	9	168
Hotelling-Lawley Trace	91.16181701	539.92	9	81.733
Roy's Greatest Root	90.20366097	1683.80	3	56

Multivariate Statistics and F Approximations

S=3 M=-0.5 N=26

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

Model: MODEL1

Multivariate Test: test_int

Error Matrix (E)

59.082326954	-1.288932107	-0.111292094
-1.288932107	25.368584312	25.072927553
-0.111292094	25.072927553	24.937988975

Hypothesis Matrix (H)

18474.183482	20655.998185	21570.907455
20655.998185	23095.486815	24118.447545
21570.907455	24118.447545	25186.717935

Multivariate Statistics and Exact F Statistics

S=1 M=0.5 N=26

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.0006664	26990.8	3	54
Pillai's Trace	0.9993336	26990.8	3	54
Hotelling-Lawley Trace	1499.4912949	26990.8	3	54
Roy's Greatest Root	1499.4912949	26990.8	3	54

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

Multivariate Test: test_factor1

Error Matrix (E)

59.082326954	-1.288932107	-0.111292094
-1.288932107	25.368584312	25.072927553
-0.111292094	25.072927553	24.937988975

Hypothesis Matrix (H)

1060.1932106	915.43627397	907.12867418
915.43627397	790.44419762	783.27090306
907.12867418	783.27090306	776.16270626

Multivariate Statistics and Exact F Statistics

S=1 M=0.5 N=26

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.01838695	960.96	3	54
Pillai's Trace	0.98161305	960.96	3	54
Hotelling-Lawley Trace	53.38639551	960.96	3	54
Roy's Greatest Root	53.38639551	960.96	3	54

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

Multivariate Test: test_factor2

Error Matrix (E)

59.082326954	-1.288932107	-0.111292094
-1.288932107	25.368584312	25.072927553
-0.111292094	25.072927553	24.937988975

Hypothesis Matrix (H)

249.6307731	307.76538183	303.888015
307.76538183	379.43851665	374.6581794
303.888015	374.6581794	369.93806699

Multivariate Statistics and Exact F Statistics

S=1 M=0.5 N=26

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.04620958	371.53	3	54
Pillai's Trace	0.95379042	371.53	3	54
Hotelling-Lawley Trace	20.64053251	371.53	3	54
Roy's Greatest Root	20.64053251	371.53	3	54

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

Multivariate Test: test_factor3

Error Matrix (E)

59.082326954	-1.288932107	-0.111292094
-1.288932107	25.368584312	25.072927553
-0.111292094	25.072927553	24.937988975

Hypothesis Matrix (H)

177.17130765	224.1070913	217.40274791
224.1070913	283.47698641	274.99654499
217.40274791	274.99654499	266.76980278

Multivariate Statistics and Exact F Statistics

S=1 M=0.5 N=26

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.05514233	308.43	3	54
Pillai's Trace	0.94485767	308.43	3	54
Hotelling-Lawley Trace	17.13488899	308.43	3	54
Roy's Greatest Root	17.13488899	308.43	3	54

Statistic	Pr > F
Wilks' Lambda	<.0001
Pillai's Trace	<.0001
Hotelling-Lawley Trace	<.0001
Roy's Greatest Root	<.0001

ภาคผนวก ค

ตัวอย่างโปรแกรม SAS ที่ใช้ในการวิเคราะห์ข้อมูล

ภาคผนวก ค1 โปรแกรมการวิเคราะห์การแจกแจงปกติหลายตัวแปร ด้วยวิธีของ Mardia

/* This program is for testing the multivariate normality using Mardia's skewness and kurtosis measures. Application on C. R. Rao's cork data */ (Khattree, Ravindra and Dayanand N. Naik, 1999)

```
proc iml ;
y ={    11.48   12.56   13.56,
11.83   13.01   14.01,
12.33   13.87   14.87,
13.42   15.02   16.02,
13.47   15.06   16.06,
12.92   14.32   15.32,
12.79   14.05   15.05,
13.12   14.24   15.24,
13.78   14.78   15.78,
14.50   15.33   16.33,
13.90   14.49   15.49,
13.94   14.77   15.77,
14.65   15.80   16.79,
14.87   16.07   17.07,
14.79   15.99   16.99,
14.66   15.86   16.86,
14.01   15.25   16.25,
12.84   14.41   15.41,
13.04   15.35   16.35,
13.64   16.19   17.19,
13.63   15.44   16.36,
13.77   15.42   16.32,
14.08   15.78   16.58,
14.39   16.26   17.06,
```

14.55 16.31 17.11,

14.59 16.19 16.99,

14.59 16.19 16.99,

14.59 16.19 16.99,

14.59 16.67 17.47,

14.59 17.51 18.31,

14.59 18.05 18.85,

14.59 19.94 20.74,

14.59 20.99 21.79,

14.59 21.22 22.02,

14.59 20.82 21.62,

14.59 19.06 19.86,

14.59 18.57 19.37,

14.74 19.33 20.13,

16.06 21.11 21.91,

18.19 22.01 22.81,

19.48 21.60 22.40,

22.51 22.37 23.17,

22.51 24.77 25.57,

22.90 25.25 26.05,

23.87 26.59 27.39,

24.05 26.39 27.19,

22.96 24.64 25.44,

23.01 24.66 25.46,

24.17 25.92 26.72,

24.42 25.71 26.51,

25.36 26.1 26.9,

26.27 27.22 28.02,

26.45 28.31 29.11,

27.33 28.89 29.69,

```
27.79    29.27    30.07,
27.55    28.94    29.67,
26.05    26.08    26.88,
24.19    24.84    25.64,
23.87    24.75    25.55,
23.62    25.39    26.19} ;

    /* Matrix y can be created from a SAS data set as follows:
        data oil;
        infile 'oil.dat' ;
        input y1 y2 y3 ;
        run;
        proc iml;
        use cork;
        read all into y;
        See Appendix 1 for details.  */
    /* Here we determine the number of data points and the dimension of the vector. The
variable dfchi is the degrees of freedom for the chi square approximation of Multivariate
skewness. */
        n = nrow(y) ;
        p = ncol(y) ;
        dfchi = p*(p+1)*(p+2)/6 ;
    /* q is projection matrix, s is the maximum likelihood estimate of the variance covariance
matrix, g_matrix is n by n the matrix of g(i,j) elements, beta1hat and beta2hat are respectively the
Mardia's sample skewness and kurtosis measures, kappa1 and kappa2  are the test statistics based
on skewness and kurtosis to test for normality and pvalskew and pvalkurt are corresponding p
values. */

        q = i(n) - (1/n)*j(n,n,1);
        s = (1/(n))*y`*q*y ; s_inv = inv(s) ;
```

```
g_matrix = q*y*s_inv*y`*q;
beta1hat = (  sum(g_matrix#g_matrix#g_matrix)  )/(n*n);
beta2hat =trace(  g_matrix#g_matrix  )/n ;
kappa1 = n*beta1hat/6 ;
kappa2 = (beta2hat - p*(p+2) ) /sqrt(8*p*(p+2)/n) ;
pvalskew = 1 - probchi(kappa1,dfchi) ;
pvalkurt = 2*( 1 - probnorm(abs(kappa2)) );
print s ;
print s_inv ;
print 'TESTS:';
print 'Based on skewness: ' beta1hat kappa1 pvalskew ;
print 'Based on kurtosis: ' beta2hat kappa2 pvalkurt;
run;
```

ภาคผนวก ค2 โปรแกรมการวิเคราะห์ค่าสหสัมพันธ์แคนนอลนิคอล

```
dm'log'clear; dm'output'clear;
options ls=64 ps=45 nodate nonumber;
data oil;
infile 'oil.dat' ;
input  y1 y2 y3 x3 x4 x5 x6 x7 x8 x9 x10 x11 x19 x20 x22 x23 x25 x26 x27 x28 x29 x30 x31 x34
x36 x37 x40 x47 x48 x52 x60 x65 x68 x69 x71 x72 @@;
proc cancorr data=oil;
        var y1 y2 y3 ;
        with x3 x4 x5 x6 x7 x8 x9 x10 x11 x19 x20 x22 x23 x25 x26 x27 x28 x29 x30 x31 x34
x36 x37 x40 x47 x48 x52 x60 x65 x68 x69 x71 x72 ;
run;
```

ภาคผนวก ค3 โปรแกรมการวิเคราะห์ปัจจัย โดยการหมุนแบบ Varimax และโปรแกรมวิเคราะห์การถดถอยหลายตัวแปร (Multivariate Regression Analysis)

```
dm'log'clear; dm'output'clear;
options ls=64 ps=45 nodate nonumber;
data oil;
infile 'oil.dat' ;
input  y1 y2 y3 x3 x4 x5 x6 x7 x8 x9 x10 x11 x19 x20 x22 x23 x25 x26 x27 x28 x29 x30 x31 x34
x36 x37 x40 x47 x48 x52 x60 x65 x68 x69 x71 x72 @@;

proc standard data=oil mean=0 std=1 out=oil2 ;
 var x3 x4 x5 x6 x7 x8 x9 x10 x11 x19 x20 x22 x23 x25 x26 x27 x28 x29 x30 x31 x34 x36 x37
x40 x47 x48 x52 x60 x65 x68 x69 x71 x72;
run;
      /* โปรแกรมการวิเคราะห์ปัจจัย โดยการหมุนแบบ Varimax  */
proc factor data=oil2 method=p nfactor=3 rotate=varimax out=fscore;
var x3 x4 x5 x6 x7 x8 x9 x10 x11 x19 x20 x22 x23 x25 x26 x27 x28 x29 x30 x31 x34 x36 x37
x40 x47 x48 x52 x60 x65 x68 x69 x71 x72;
run;
proc print data=fscore;
run;
      /* โปรแกรมการวิเคราะห์การถดถอยหลายตัวแปร */
proc reg;
model y1 y2 y3 = factor1 factor2 factor3;
test_all:mtest factor1, factor2, factor3/print;
test_int:mtest intercept/print;
test_factor1:mtest factor1=0/print;
test_factor2:mtest factor2=0/print;
test_factor3:mtest factor3=0/print;
run;
```

ประวัติการศึกษา และการทำงาน

ชื่อ –นามสกุล	นางสาวภัทริณี คงชู
วัน เดือน ปี ที่เกิด	20 กรกฎาคม 2525
สถานที่เกิด	อ.เมือง จ.นครสวรรค์
ประวัติการศึกษา	วท.ม. (สถิติ) มหาวิทยาลัยเกษตรศาสตร์
ตำแหน่งหน้าที่การงานปัจจุบัน	-
สถานที่ทำงานปัจจุบัน	-
ผลงานดีเด่นและรางวัลทางวิชาการ	-
ทุนการศึกษาที่ได้รับ	-