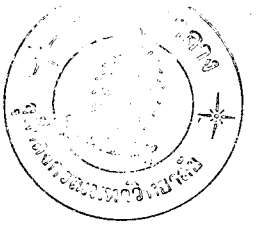
บทที่ 2

ทฤษฎีที่ใช้ในการวิจัย



ตัวแปรและสัญญลักษณ์

Y_i คือ ค่าใช้จ่ายต่อเดือนในหมวดที่ i ของครัวเรือน i = 1,2,3,4 เป็นตัวแปรตาม (dependent variable) ในการศึกษาครั้งนี้เลือกใช้จำนวนที่ครัวเรือนใช้จ่ายในการอุปโภคและบริโภคในหมวดต่าง ๆ แทนที่จะใช้ปริมาณสินค้า เนื่องจากค่าใช้จ่ายเป็นตัวแปรที่มีความสัมพันธ์กับรายได้ค่อนข้างดีให้เส้นโค้งที่แสดงความสัมพันธ์ค่อนข้างเรียบ นอกจากนั้นยังให้ความสะดวกในการรวมสินค้าหมวดเดียวกันแต่มีหน่วยต่างกันเข้าด้วยกัน เช่น สินค้าในหมวดอาหาร สามารถที่จะรวมไข่ 5 ฟอง กับนม 1 กระป๋อง เข้าด้วยกันได้

X คือ รายได้ต่อเดือนของครัวเรือน เป็นตัวแปรอิสระ (independent variable) ในที่นี้ใช้ข้อมูลยอดรวมค่าใช้จ่ายทุกหมวดต่อเดือนแทนข้อมูลรายได้ต่อเดือนของครัวเรือน เนื่องจากรายได้ของครัวเรือนมีการเปลี่ยนแปลงตลอดเวลา ข้อมูลรายได้ที่สำรวจในช่วงเวลาหนึ่ง ไม่ได้เป็นตัวแทนที่ดีที่จะบ่งชี้ถึงระดับการครองชีพ แต่ข้อมูลยอดรวมค่าใช้จ่ายทุกหมวดแสดงฐานะทางเศรษฐกิจที่ถาวรและใกล้เคียงกับความเป็นจริงมากกว่าข้อมูลรายได้ ดังนั้น ต่อไปนี้จะใช้ x เป็นสัญญลักษณ์ ของยอดรวมค่าใช้จ่ายทุกหมวด การใช้ข้อมูลยอดรวมค่าใช้จ่ายทุกหมวดเป็นตัวแปรอิสระจะต้อง

มีข้อสมมติว่า ขณะที่ยอดรวมค่าใช้จ่ายทุกหมวดขึ้นกับรายได้ที่คาดว่าจะได้รับและรสนิยม การกระจายของการใช้จ่ายเพื่อการอุปโภคบริโภคไปสู่หมวดต่าง ๆ จะขึ้นอยู่กับระดับของยอดรวมค่าใช้จ่ายทุกหมวด [1, 2]

N คือ ขนาดของครัวเรือน (household size) ซึ่งวัดจากจำนวนสมาชิกในครัวเรือน เป็นตัวแปรอิสระ วิธีที่พิจารณาใช้ยอดรวมค่าใช้จ่ายทุกหมวดเพียงตัวเดียวเป็นตัวแปรที่มีอิทธิพลต่อการใช้จ่ายในหมวดต่าง ๆ จะใช้ได้ดีถ้าสามารถแบ่งประชากร (population) ออกเป็นกลุ่ม ๆ ที่มีส่วนประกอบอื่น ๆ เหมือนกัน (homogeneous) แต่ต่างกันที่รายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด แล้วหารูปแบบทางคณิตศาสตร์ที่สามารถกระชับ (fit) ได้ดีกับแต่ละกลุ่มในประชากร แต่โดยเหตุที่ไม่สามารถแบ่งประชากรออกเป็นกลุ่ม ๆ ได้ก็จำเป็นต้องเพิ่มตัวแปรเข้าไป ตัวแปรที่สำคัญตัวหนึ่งคือ ขนาดของครัวเรือน

ข้อสมมติเบื้องต้น [1]

ในการวิเคราะห์ข้อมูลการใช้จ่ายซึ่งเป็นข้อมูลของครัวเรือนต่าง ๆ ที่สังเกตุได้จากช่วงเวลาเดียวกัน (cross-sectional data) เพื่อหารูปแบบทางคณิตศาสตร์นำมาใช้อธิบายพฤติกรรมของผู้บริโภคนั้น เนื่องจากข้อมูลเหล่านี้

---

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed. p. 80, p. 8.

2 Lawrence R. Klein, An Introduction to Econometrics, (New Delhi: Prentice-Hall of India Private Ltd., 1969), p. 58.

แสดงความแปรปรวนหรือความแตกต่างของการใช้จ่ายในระหว่างครัวเรือนต่าง ๆ ณ ระยะเวลาหนึ่ง จึงจำเป็นต้องมีข้อสมมติ คือ

จากการสังเกตุครัวเรือนในสภาวะต่าง ๆ กัน ณ เวลาเดียวกัน จะได้รายละเอียดข้อมูลที่เกี่ยวข้องหรือสามารถใช้ในการคาดคะเนหรือทำนายพฤติกรรมของครัวเรือนใดครัวเรือนหนึ่งเมื่อสภาวะแวดล้อมเดิมได้เปลี่ยนแปลงไป เช่น สังเกตุการใช้จ่ายของครัวเรือน 2 ครัวเรือนที่มีรายได้ต่างกัน และในปีต่อมารายได้ของครัวเรือนแรกเปลี่ยนแปลงไปเท่ากับรายได้ของครัวเรือนที่สอง ดังที่ได้สำรวจไว้ในปีฐานแล้ว ครัวเรือนแรกจะมีลักษณะการใช้จ่ายเช่นเดียวกับที่ครัวเรือนที่สองได้ใช้จ่ายไปในปีฐาน

คำว่า "สภาวะต่าง ๆ กัน" ได้แก่ รายได้ที่ต่างกัน ขนาดของครัวเรือนที่แตกต่างกัน ส่วนประกอบของครัวเรือนในรูปของอายุ เพศของสมาชิกที่แตกต่างกัน เป็นต้น

004522

<u>การวิเคราะห์เชิงสถิติของสมการความถดถอย (Regression Analysis)</u>

การวิเคราะห์ความถดถอย หมายถึง การศึกษาความสัมพันธ์ระหว่างตัวแปรตั้งแต่สองตัวขึ้นไป ในการศึกษาความสัมพันธ์นี้สามารถประมาณค่าของตัวแปรหนึ่งจากการรู้ค่าของตัวแปรอิสระอีกตัวหนึ่งหรือมากกว่าหนึ่งตัวขึ้นไปได้ และสามารถวัดค่าความคลาดเคลื่อน (errors) ในการประมาณค่าได้ ตัวแปรที่ต้องการประมาณค่าเป็นตัวแปรตาม ส่วนตัวแปรที่ใช้ในการประมาณค่าเป็นตัวแปรที่กำหนดค่าไว้ล่วงหน้า (predetermined variable) เรียกว่า ตัวแปรอิสระ วัตถุประสงค์ของการ

วิเคราะห์ความถดถอย [1] คือ เพื่อประมาณค่าของตัวแปรตามจากการรู้ค่าของตัวแปรอิสระ เครื่องมือที่ใช้ในการประมาณ คือ เส้นความถดถอย (regression line) ซึ่งเป็นเส้นที่กระชับ (fit) พอดีกับข้อมูลโดยวิธีกำลังสองน้อยที่สุด (method of least squares) เส้นความถดถอยจะอธิบายความสัมพันธ์โดยเฉลี่ยที่มีอยู่ระหว่างตัวแปรตามกับตัวแปรอิสระ และเป็นเส้นซึ่งแสดงค่าเฉลี่ยของตัวแปรตามเมื่อกำหนดค่าตัวแปรอิสระ สมการของเส้นความถดถอย เรียกว่า สมการความถดถอย (regression equation) ซึ่งจะให้ค่าประมาณของตัวแปรตามเมื่อแทนค่าของตัวแปรอิสระลงในสมการ วัตถุประสงค์อีกประการหนึ่งของการวิเคราะห์ความถดถอย ก็คือ เพื่อวัดค่าความคลาดเคลื่อน เนื่องจากการใช้เส้นความถดถอยเป็นฐานในการประมาณค่า ด้วยเหตุนี้ จึงหาค่าความคลาดเคลื่อนมาตรฐาน (standard error) ของค่าประมาณ และค่าความคลาดเคลื่อนมาตรฐานของเส้นความถดถอย ซึ่งค่าความคลาดเคลื่อนมาตรฐานของเส้นความถดถอยนี้ถ้ายกกำลังสองจะเป็นค่ามัชฌิมกำลังสองของความแตกต่างของค่าสังเกตุจากตัวอย่างกับค่าจากเส้นความถดถอย (error variance around the regression line) เป็นการวัดการกระจายของค่าสังเกตุของตัวแปรตามรอบ ๆ ค่าที่ประมาณจากเส้นความถดถอย เส้นที่กระชับเข้ากับข้อมูลดีมากจะสามารถประมาณค่าตัวแปรตามได้ดี หรือกล่าวอีกนัยหนึ่งว่า ถ้าค่าสังเกตุของตัวแปรตามกระจายรอบเส้นความถดถอยน้อยมากจะสามารถประมาณค่าตัวแปรตามได้ดี แต่ถ้าค่าสังเกตุของตัวแปรตามกระจายรอบเส้นความถดถอยมาก

---

1 Morris Hamburg, Statistical Analysis for Decision Making, (New York: Harcourt, Brace & World, Inc., 1970), p. 464-466.

เส้นความถดถอยจะไม่ให้ค่าประมาณที่ถูกต้องแน่นอนนัก

ในการกำหนดรูปแบบทางคณิตศาสตร์ของ ระนาบ Engel จำเป็นต้องคำนึงถึงความสะดวกในการคิดคำนวณตัวเลขเพื่อประมาณค่าสัมประสิทธิ์ความถดถอย จึงมักกำหนดรูปแบบซึ่งเป็น linear หรือ non-linear ที่สามารถแปลงรูปเป็น linear ได้โดยใช้วิธีแปลงรูปตัวแปร (transform variable) ที่เหมาะสม ซึ่งจะช่วยให้ได้ตัวคลาดเคลื่อน (error term) ที่มีค่าแปรปรวน (variance) เป็นค่าคงที่ (homoscedastic) [1]

รูปแบบทางคณิตศาสตร์ที่ใช้ในการตรวจสอบเพื่อหาความสัมพันธ์ระหว่างค่าใช้จ่ายของครัวเรือนในหมวดต่าง ๆ กับยอดรวมค่าใช้จ่ายทุกหมวดและขนาดของครัวเรือน จะใช้รูปแบบที่ได้มีผู้เคยศึกษามาแล้วในต่างประเทศ [2] คือ

Linear form $$Y_i = A_i + B_iX + C_iN + E_i$$

ซึ่งสมการนี้ Allen และ Bowley ได้เคยใช้ศึกษามาก่อน โดยใช้สมการในการประมาณ คือ

$$\hat{Y}_i = a_i + b_iX + c_iN$$

---

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed. p. 86.

2 Houthakker, "An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel's Law," in Econometrica, vol.25 (1957) p. 532-551.

Double-logarithmic form $\log Y_i = A_i + B_i \log X + C_i \log N + E_i$

สมการนี้มีรูปเดิมเป็น non-linear ในรูปของผลคูณ (multiplicative form)

$$Y_i = A_i' X^{B_i} N^{C_i} E_i'$$

แล้วแปลงรูปโดยใส่ log ฐาน e โดยตลอด

H.S. Houthakker ได้ศึกษา double-logarithmic form โดยใช้สมการในการกะประมาณ คือ

$$\log \hat{Y}_i = a_i + b_i \log X + c_i \log N$$

หรือ $$\hat{Y}_i = a_i' X^{b_i} N^{c_i}$$

Semi-logarithmic form $Y_i = A_i + B_i \log X + C_i \log N + E_i$

สมการนี้รูปเดิมเป็น non-linear ในรูปของ exponential form แล้วแปลงรูปโดยใส่ log ฐาน e โดยตลอด

Prais และ H.S. Houthakker ได้ศึกษา semi-logarithmic form โดยใช้สมการในการประมาณ คือ

$$\hat{Y}_i = a_i + b_i \log X + c_i \log N$$

หรือ $$e^{\hat{Y}_i - a_i} = X^{b_i} N^{c_i}$$

2.1 การประมาณค่าสัมประสิทธิ์ความถดถอยด้วยวิธีกำลังสองน้อยที่สุด

ในการวิจัยครั้งนี้จะศึกษาหาความสัมพันธ์ระหว่างตัวแปรสามตัว คือ ค่าใช้จ่ายต่อเดือนในหมวดที่ i ของครัวเรือน(Y_i) ซึ่งเป็นตัวแปรตาม ยอดรวมค่าใช้จ่ายทุกหมวดต่อเดือนของครัวเรือน(X) และขนาดของครัวเรือน(N) ซึ่งตัวแปรสองตัวหลังนี้เป็นตัวแปรอิสระ ดังนั้นความสัมพันธ์ระหว่างตัวแปรทั้งสามนี้จะแสดงออกมาในลักษณะ

ของระนาบ (plane) ที่มีรูปร่างต่างกัน เช่น

Linear form $Y_{ij} = A_i + B_iX_j + C_iN_j + E_{ij}$

Estimated linear form $\hat{Y}_{ij} = a_i + b_iX_j + c_iN_j$

โดย $\hat{Y}_{ij}$ คือ ค่าประมาณของค่าใช้จ่ายต่อเดือนในหมวดที่ i ของครัวเรือน j (Y_{ij}) เมื่อกำหนดยอดรวมค่าใช้จ่ายทุกหมวดต่อเดือนของครัวเรือนที่ j (X_j) และขนาดของครัวเรือนที่ j (N_j)

a_i, b_i และ c_i คือ ค่าประมาณของ A_i, B_i และ C_i ซึ่งเป็นค่าสัมประสิทธิ์ความถดถอย

E_{ij} คือ ตัวคลาดเคลื่อน (error term) ในหมวดที่ i ของครัวเรือนหรือค่าสังเกตุที่ j, j = 1,2,....m เมื่อm คือ จำนวนค่าสังเกตุ

การประมาณระนาบความถดถอย มีข้อสมมติเบื้องต้นเกี่ยวกับตัวแปรอิสระและ ตัวคลาดเคลื่อน [1] คือ

1. Normality : E_{ij} มีการกระจายปกติ ทุก ๆ j
2. Zero mean : $E(E_{ij}) = 0$ ค่าเฉลี่ยของ E_{ij} เป็น 0 ทุก ๆ j
3. Homoscedastic : $E(E_{ij}^2) = \sigma^2$ ค่าแปรปรวนของ E_{ij} มีค่าคงที่ทุก ๆ j
4. Nonautocorrelation : $E(E_{ij}E_{ik}) = 0$, $j \neq k$ นั่นคือ E_{ij} และ E_{ik} ไม่มีความสัมพันธ์กัน (uncorrelate)

---

1 Jan Kmenta, Elements of Econometrics, (New York: The Macmillan Company, 1971), p. 348.

5. Nonstochastic independent variable : X_j และ N_j ไม่ใช่ตัวแปรสุ่ม แต่ถูกกำหนดค่าไว้ล่วงหน้าโดยที่ค่าเฉลี่ยของผลต่างกำลังสองระหว่างตัวแปรอิสระกับค่าเฉลี่ยของตัวแปรนั้นมีค่าต่างจาก 0

6. จำนวนค่าสังเกตุในตัวอย่างจะต้องมากกว่าจำนวนสัมประสิทธิ์ของความถดถอยที่ต้องการประมาณค่า

7. ไม่มีความสัมพันธ์กันโดยตรงอย่างแน่ชัดในเชิงเส้นตรง (exact linear relation) ระหว่างตัวแปรอิสระทั้งหลาย

ในการประมาณระนาบความถดถอย $\hat{Y}_j = a + bX_j + cN_j$ ของค่าใช้จ่ายหมวดที่สนใจ ด้วยวิธีกำลังสองน้อยที่สุดจะได้ normal equation คือ

$$\sum_j Y_j = ma + b\sum_j X_j + c\sum_j N_j$$

$$\sum_j X_j Y_j = a\sum_j X_j + b\sum_j X_j^2 + c\sum_j X_j N_j$$

$$\sum_j N_j Y_j = a\sum_j N_j + b\sum_j N_j X_j + c\sum_j N_j^2$$

เพื่อความสะดวกในการหาค่า a, b และ c จะแปลง normal equation ใหม่ โดยให้ $y_j = Y_j - \bar{Y}$

$$x_j = X_j - \bar{X}$$

$$n_j = N_j - \bar{N}$$

จะได้ normal equation ชุดใหม่ ดังนี้

$$\sum_j x_j y_j = b\sum_j x_j^2 + c\sum_j x_j n_j$$

$$\sum_j n_j y_j = b\sum_j x_j n_j + c\sum_j n_j^2$$

ซึ่งจะได้ค่า a, b และ c ดังนี้

$$b = \frac{\sum_j x_j y_j \sum_j n_j^2 - \sum_j n_j y_j \sum_j x_j n_j}{\sum_j x_j^2 \sum_j n_j^2 - (\sum_j x_j n_j)^2}$$

$$c = \frac{\sum_j n_j y_j \sum_j x_j^2 - \sum_j x_j y_j \sum_j x_j n_j}{\sum_j x_j^2 \sum_j n_j^2 - (\sum_j x_j n_j)^2}$$

$$a = \bar{Y} - b\bar{X} - c\bar{N}$$

Double-logarithmic form $\log Y_{ij} = A_i + B_i \log X_j + C_i \log N_j + E_{ij}$

Estimated double-logarithmic form $\log \hat{Y}_{ij} = a_i + b_i \log X_j + c_i \log N_j$

ให้ $Z_{ij} = \log Y_{ij}$, $U_j = \log X_j$, $V_j = \log N_j$

ดังนั้น $\hat{Z}_{ij} = a_i + b_i U_j + c_i V_j$

Semi-logarithmic form $Y_{ij} = A_i + B_i \log X_j + C_i \log N_j + E_{ij}$

Estimated semi-logarithmic form $\hat{Y}_{ij} = a_i + b_i \log X_j + c_i \log N_j$

ให้ $U_j = \log X_j$, $V_j = \log N_j$

ดังนั้น $\hat{Y}_{ij} = a_i + b_i U_j + c_i V_j$

เนื่องจาก double-logarithmic form และ semi-logarithmic form สามารถจัดรูปให้มีลักษณะเดียวกับ linear form ได้ ดังนั้นการประมาณค่าของสัมประสิทธิ์ของความถดถอย (regression coefficient) และการทดสอบสมมติฐานเกี่ยวกับค่าของสัมประสิทธิ์ความถดถอยในประชากร ซึ่งจะกล่าวต่อไปในหัวข้อ 2.2 คงทำตามลำดับขั้นตอนเช่นเดียวกับ linear form สำหรับข้อสมมติเบื้องต้นเกี่ยวกับตัวแปรอิสระและตัวคลาดเคลื่อนยังคงเดิม

2.2 การทดสอบสมมติฐานเกี่ยวกับค่าสัมประสิทธิ์ความถดถอยในประชากร

การทดสอบสมมติฐานเกี่ยวกับค่าสัมประสิทธิ์ความถดถอยในประชากร สามารถทดสอบได้ในสองลักษณะ คือ

1. ทดสอบอิทธิพลร่วม (joint influence) ของตัวแปรอิสระสองตัว คือ ยอดรวมค่าใช้จ่ายทุกหมวด และขนาดของครัวเรือน ที่มีต่อการใช้จ่ายในหมวดต่าง ๆ โดยตั้งสมมติฐานไว้ว่า ค่าสัมประสิทธิ์ความถดถอยในประชากรของตัวแปรอิสระทั้งสองมีค่าเป็น 0 (H_0 : B = C = 0) และตั้งสมมติฐานแย้งไว้ว่า ค่าสัมประสิทธิ์ความถดถอยในประชากรของตัวแปรอิสระทั้งสองไม่จำเป็นต้องมีค่าเป็น 0 หมดทุกตัว (H_1 : B ≠ 0, C ≠ 0) วิธีการทดสอบสมมติฐานจะทดสอบด้วย F-test โดยใช้ค่าสถิติ F จากตารางวิเคราะห์ค่าแปรปรวน (analysis of variance table หรือ ANOVA table)

ตาราง 2.1

ANOVA Table

Source of Vaiation	df	SS	MS	F-Value
Regression (X,N)	2	$Q_1 = b\sum_j x_j y_j + c\sum_j n_j y_j$	$Q_1/2 = Q_3$	Q_3/S^2
Residual	m-3	$Q_2 - Q_1$	$Q_2-Q_1/m-3 = S^2$	
Total (corrected for mean)	m-1	$Q_2 = \sum_j y_j^2$		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

$H_o : B = C = 0$, $H_1 : B \neq 0, C \neq 0$

เปรียบเทียบค่า F กับ F_α ในตารางราง F-distribution ที่ระดับความมีนัยสำคัญ α องศาอิสระ (degree of freedom) (2, m-3)

ถ้า F มีค่ามากกว่า F_α อย่างมีนัยสำคัญจะปฏิเสธสมมติฐาน ซึ่งหมายความว่า จะต้องมีสัมประสิทธิ์ความถดถอยของตัวแปรอิสระอย่างน้อยหนึ่งตัวที่มีค่าแตกต่างจาก 0 อย่างมีนัยสำคัญ หรือจะกล่าวได้อีกอย่างว่า จะต้องมีตัวแปรอิสระอย่างน้อยหนึ่งตัวที่มีอิทธิพลต่อการใช้จ่ายในหมวดต่าง ๆ ของครัวเรือน คือ อาจจะเป็นตัวแปรยอดรวมค่าใช้จ่ายทุกหมวด หรือ ตัวแปรขนาดของครัวเรือน หรือทั้งสองตัวก็ได้

ถ้า F มีค่าไม่มากกว่า F_α อย่างมีนัยสำคัญจะยอมรับสมมติฐาน ซึ่งหมายความว่า สัมประสิทธิ์ความถดถอยของตัวแปรอิสระทั้งสองไม่มีค่าแตกต่างจาก 0 อย่างมีนัยสำคัญ หรือจะกล่าวได้ว่า ไม่มีตัวแปรอิสระตัวใดที่มีอิทธิพลต่อการใช้จ่ายในหมวดต่าง ๆ นั่นคือ การใช้จ่ายในหมวดต่าง ๆ ของแต่ละครัวเรือนที่มีค่าแปรเปลี่ยนไปนั้น มิได้มีสาเหตุเนื่องจากอิทธิพลของยอดรวมค่าใช้จ่ายทุกหมวด หรือรายได้ของครัวเรือน หรือขนาดครัวเรือน แต่เป็นเพราะการใช้จ่ายในหมวดต่าง ๆ นั้นมีค่าสุ่ม (random)

2. ทดสอบอิทธิพลของตัวแปรอิสระแต่ละตัว (seperate influence) โดยตั้งสมมติฐานไว้ว่า ค่าสัมประสิทธิ์ความถดถอยแต่ละตัวในประชากรมีค่าเป็น 0 ($H_o : A = 0$, $H_o : B = 0$, $H_o : C = 0$) และตั้งสมมติฐานแย้งไว้ว่า ค่าสัมประสิทธิ์ความถดถอยแต่ละตัวในประชากรมีค่าไม่เท่ากับ 0 ($H_1 : A \neq 0$, $H_1 : B \neq 0$, $H_1 : C \neq 0$) วิธีทดสอบสมมติฐาน จะทดสอบด้วย t-test

$H_o : A = 0, H_1 : A \neq 0$

$$t = a/S.E.(a)$$

$$S.E.(a) = \left[\frac{1}{m} + \frac{\bar{X}^2 \sum_j n_j^2 + \bar{N}^2 \sum_j x_j^2}{\sum_j x_j^2 \sum_j n_j^2 - (\sum_j x_j n_j)^2} \right]^{\frac{1}{2}} S$$

เปรียบเทียบค่าสัมบูรณ์ของ t กับ $t_{(m-3,\ 1-\frac{1}{2}\alpha)}$ ในตาราง t-distribution ที่ระดับความมีนัยสำคัญ α องศาอิสระ m-3

ถ้า t ที่คำนวณได้มีค่ามากกว่าในตารางอย่างมีนัยสำคัญ จะปฏิเสธสมมติฐาน ซึ่งหมายความว่าระนาบความถดถอยของประชากรตัดแกนตั้งฉากห่างจากจุดเริ่มต้น (origin) เท่ากับ A

ถ้า t ที่คำนวณได้มีค่าไม่มากกว่าในตารางอย่างมีนัยสำคัญ จะยอมรับสมมติฐาน ซึ่งหมายความว่า ถ้าให้ตัวแปรยอดรวมค่าใช้จ่ายทุกหมวด และตัวแปรขนาดครัวเรือนอยู่ในแกนนอน และแกนทั้งสองตั้งฉากกัน โดยมีตัวแปรค่าใช้จ่ายในแต่ละหมวดอยู่ในแกนตั้งและตั้งฉากกับสองแกนแรก ระนาบความถดถอยของประชากรในสามมิติจะไม่ตัดแกนตั้งฉากแต่จะผ่านจุดเริ่มต้น หรือ intecept เท่ากับ 0

$H_o : B = 0, H_1 : B \neq 0$

$$t = b/S.E.(b)$$

$$S.E.(b) = \left[\frac{\sum_j n_j^2}{\sum_j x_j^2 \sum_j n_j^2 - (\sum_j x_j n_j)^2} \right]^{\frac{1}{2}} S$$

$$H_o : C = 0, \; H_1 : C \neq 0$$

$$t = c/S.E.(c)$$

$$S.E.(c) = \left[\frac{\sum_j x_j^2}{\sum_j x_j^2 \sum_j n_j^2 - \left(\sum_j x_j n_j\right)^2} \right]^{\frac{1}{2}} S$$

เปรียบเทียบค่าสัมบูรณ์ t เช่นเดียวกับการทดสอบสมมติฐานเกี่ยวกับ A ถ้า t ที่คำนวณได้มีค่ามากกว่าในตารางอย่างมีนัยสำคัญ จะปฏิเสธสมมติฐาน ซึ่งหมายความว่า ตัวแปรยอดรวมค่าใช้จ่ายทุกหมวด หรือรายได้ มีอิทธิพลต่อการใช้จ่ายของครัวเรือนในแต่ละหมวด และทำนองเดียวกัน ขนาดครัวเรือนหรือจำนวนสมาชิกในครัวเรือนก็มีอิทธิพลต่อการใช้จ่ายในแต่ละหมวด

ถ้า t ที่คำนวณได้มีค่าไม่มากกว่าในตารางอย่างมีนัยสำคัญ จะยอมรับสมมติฐาน ซึ่งหมายความว่า ตัวแปรยอดรวมค่าใช้จ่ายทุกหมวด หรือรายได้ ไม่มีอิทธิพลต่อการใช้จ่ายในแต่ละหมวด ทำนองเดียวกัน ขนาดของครัวเรือนก็ไม่มีอิทธิพลต่อการใช้จ่ายในแต่ละหมวดของครัวเรือน

เมื่อใดที่ทดสอบสมมติฐานแล้ว ปรากฏว่าสัมประสิทธิ์ความถดถอยในประชากรของตัวแปรอิสระตัวใดตัวหนึ่งเป็น 0 ระนาบความถดถอยจะเปลี่ยนเป็นเส้นความถดถอยในสองมิติ แต่ถ้าทดสอบแล้วพบว่า ยอมรับสมมติฐาน ทั้ง B = 0 และ C = 0 จะหมายความว่า การเปลี่ยนแปลงของการใช้จ่ายในแต่ละหมวดมีค่าสุ่มไม่ได้ขึ้นกับอิทธิพลของตัวแปรอิสระทั้งสองดังกล่าว

ค่าความคลาดเคลื่อนมาตรฐาน หรือ S.E. ของค่าประมาณของสัมประสิทธิ์ความถดถอย สามารถนำมาหาอันตรภาคแห่งความเชื่อมั่น (confidence interval) ด้วยความเชื่อมั่น 100 $(1-\alpha)$ % ดังนี้คือ

$$a \pm t_{(m-3,\ 1-\frac{1}{2}\alpha)}\ S.E.(a)$$

$$b \pm t_{(m-3,\ 1-\frac{1}{2}\alpha)}\ S.E.(b)$$

$$c \pm t_{(m-3,\ 1-\frac{1}{2}\alpha)}\ S.E.(c)$$

2.3 การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เชิงซ้อน (Multiple Correlation Coefficient Analysis)

การวิเคราะห์ค่าสัมประสิทธิ์สหสัมพันธ์เชิงซ้อน หมายถึง วิธีการวัดความสัมพันธ์ระหว่างตัวแปรมากกว่าสองตัวขึ้นไปว่ามีความสัมพันธ์กันมากน้อยเพียงใด โดยถือว่าตัวแปรนั้นเป็นตัวแปรสุ่ม (random variable) เครื่องมือที่ใช้วัดความสัมพันธ์คือ สัมประสิทธิ์สหสัมพันธ์เชิงซ้อน (coefficient of multiple correlation) ซึ่งเมื่อยกกำลังสองจะเป็นค่าของ coefficient of determination

ในการวิเคราะห์ค่าสหสัมพันธ์ระหว่างค่าใช้จ่ายในแต่ละหมวดต่อเดือนของครัวเรือน(Y) กับยอดรวมค่าใช้จ่ายทุกหมวดต่อเดือนของครัวเรือน(X) และขนาดของครัวเรือน(N) วิเคราะห์ได้จาก

$$R = \left[\frac{(\Sigma x_j y_j)^2}{\Sigma x_j^2 \Sigma y_j^2} + \frac{(\Sigma n_j y_j)^2}{\Sigma n_j^2 \Sigma y_j^2} \right]^{\frac{1}{2}}$$

หรือ

$$= \left[\frac{b\Sigma x_j y_j + c\Sigma n_j y_j}{\Sigma y_j^2} \right]^{\frac{1}{2}}$$

R มีค่าอยู่ระหว่าง −1 กับ 1 เป็นตัวที่ใช้วัดความสัมพันธ์ระหว่าง X และ N กับ Y ว่ามีมากน้อยเพียงใด

ถ้า R เท่ากับ 1 หมายความว่า X และ N มีความสัมพันธ์กับ Y ในทางบวกอย่างสมบูรณ์ คือ Y มีค่าแปรตาม X และ N และทุก ๆ จุดของแผนภาพการกระจาย (scatter diagram) คือ (X,N;Y) จะอยู่บนระนาบตรงที่มีความชัน (slope) เป็นบวก ถ้า R เท่ากับ −1 ก็มีความหมายเช่นเดียวกัน แต่ความชัน (slope) เป็นลบ

ถ้า R เท่ากับ 0 หมายความว่า X และ N ไม่มีความสัมพันธ์กับ Y ในเชิงเส้นตรง (linear unassociated) แต่ก็ไม่ได้หมายความว่า X, N เป็นอิสระจาก Y

กำลังสองของค่าสัมประสิทธิ์เชิงซ้อน R^2 จะเป็นตัวแสดงให้ทราบว่าหลังจากที่สามารถหารูปแบบของสมการความถดถอยแสดงความสัมพันธ์ระหว่าง X, N กับ Y ได้แล้ว สมการความถดถอยนั้นสามารถอธิบายค่าแปรปรวนทั้งหมดของ Y (total variation of Y) ได้มากน้อยเพียงใด R^2 สามารถเขียนได้ในรูปของ

$$R^2 = \frac{\text{ค่าแปรปรวนของ Y ในส่วนที่อธิบายได้ด้วยสมการความถดถอย}}{\text{ค่าแปรปรวนทั้งหมดของ Y}}$$

หรือ $$= \sum_j (Y_j - \hat{Y}_j)^2 \;/\; \sum_j (Y_j - \bar{Y})^2$$

R^2 มีค่าสูงสุดเป็น 1 และมีค่าต่ำสุดเป็น 0

ถ้า R^2 เท่ากับ 0 หมายความว่า สมการความถดถอยอธิบายค่าแปรปรวนของ Y ไม่ได้เลย

ถ้า R^2 เท่ากับ 1 หมายความว่า สมการความถดถอยสามารถอธิบายค่าแปรปรวนของ Y ได้อย่างสมบูรณ์

การทดสอบสมมติฐานว่าสมการความถดถอยจะสามารถอธิบายค่าแปรปรวนของ Y ได้หรือไม่ หรือมีความสัมพันธ์ระหว่าง X, N กับ Y หรือไม่ ตั้งสมมติฐานไว้ว่า มีความสัมพันธ์ระหว่างตัวแปร และตั้งสมมติฐานแย้งว่าไม่มีความสัมพันธ์ระหว่างตัวแปร โดยให้ ρ เป็นตัวที่แสดงความสัมพันธ์ระหว่างตัวแปรในประชากร ทดสอบโดยใช้ Fisher's Z distribution โดย

$$z = \tfrac{1}{2} \log \left[(1+R)/(1-R)\right]$$

$$\sigma_z = 1/(m-k-2)^{\frac{1}{2}}$$ โดย k คือ จำนวนตัวแปรอิสระ

$$Z = z/\sigma_z$$

$H_o : \rho = 0$, $H_1 : \rho \neq 0$

เปรียบเทียบค่า Z ที่คำนวณได้กับ $Z_{(1-\frac{1}{2}\alpha)}$ ในตาราง Z-distribution ที่ระดับความมีนัยสำคัญ α

ถ้า Z มีค่ามากกว่าในตารางอย่างมีนัยสำคัญ จะปฏิเสธสมมติฐาน

ถ้า Z มีค่าไม่มากกว่าในตารางอย่างมีนัยสำคัญ จะยอมรับสมมติฐาน

<u>การวิเคราะห์เชิงเศรษฐศาสตร์</u> [1]

รูปแบบทางคณิตศาสตร์ของเส้นโค้ง Engel สำหรับค่าใช้จ่ายแต่ละหมวด มีคุณสมบัติดังนี้

1. จะมีค่าใช้จ่ายบางหมวดที่ครัวเรือนจะยังไม่ใช้จ่ายเงินเพื่อการอุปโภคบริโภค เมื่อเริ่มมีรายได้ในระยะเริ่มแรก (initial income)

---

1 Prais and Houthakker, <u>The Analysis of Family Budgets</u>, 2d ed. p. 82-84.

2. มีระดับอิ่มตัว (satiety level) คือปริมาณสูงสุดที่ครัวเรือนจะใช้จ่ายในหมวดนั้น ซึ่งจะไม่เกินไปกว่านี้อีกแม้ว่าจะมีรายได้เพิ่มขึ้น

โดยทั่วไปแล้ว ไม่จำเป็นที่ค่าใช้จ่ายทุก ๆ หมวดจะต้องมีคุณสมบัติทั้งสองข้อนี้ แต่สำหรับค่าใช้จ่ายหมวดที่สำคัญ ๆ มักจะมีคุณสมบัติดังกล่าว

รูปแบบทางคณิตศาสตร์ที่มีคุณสมบัติทั้งสองข้ออย่างชัดเจนคือ รูป hyperbola

$$Y_i = A_i - B_i/X + E_i$$

โดย Y_i คือ ค่าใช้จ่ายของครัวเรือนในหมวดที่ i

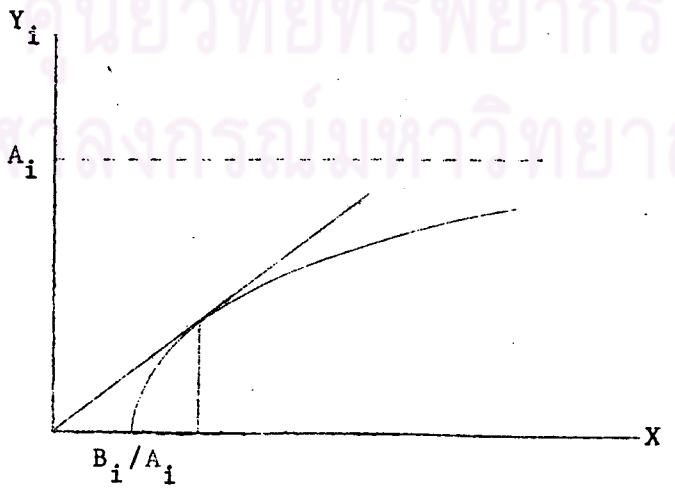
X คือ ยอดรวมค่าใช้จ่ายทุกหมวดหรือรายได้ของครัวเรือน

A_i, B_i คือ ค่าสัมประสิทธิ์ความถดถอย

E_i คือ ตัวคลาดเคลื่อนในหมวดที่ i

รูปที่ 2.1

A Hyperbolic Engel Curve



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Hyperbolic Engel curve นี้ มีรายได้ในระยะเริ่มแรกที่ยังไม่นำไปใช้จ่ายในการอุปโภคบริโภคหมวดที่ i เท่ากับ B_i/A_i มี asymptote หรือระดับอิ่มตัวเท่ากับ A_i ค่าความยืดหยุ่นของการใช้จ่ายเมื่อเปรียบเทียบกับรายได้ (income elasticity of demand) ลดลงอย่างต่อเนื่องขณะที่รายได้เพิ่มขึ้น และจะมีรายได้ระดับหนึ่งที่มีค่าความยืดหยุ่นเท่ากับ 1 จุดนี้หาได้โดยลากเส้นตรงจากจุดเริ่มต้นสัมผัสกับเส้นโค้ง ณ จุดสัมผัส $dY_i/dX = Y_i/X$ นั่นคือ $XdY_i/Y_i dX = 1$ จุดนี้เรียกว่า จุดฟุ่มเฟือย (luxury point) เนื่องจาก ณ ระดับต่าง ๆ ของรายได้ที่อยู่ทางซ้ายมือของจุดนี้ ความยืดหยุ่นของการใช้จ่ายเมื่อเปรียบเทียบกับรายได้มีค่ามากกว่า 1 จึงเรียกว่าเป็นค่าใช้จ่ายฟุ่มเฟือย (luxury) และ ณ ระดับต่าง ๆ ของรายได้ที่อยู่ทางขวามือของ luxury point ความยืดหยุ่นของการใช้จ่ายเมื่อเปรียบเทียบกับรายได้มีค่าน้อยกว่า 1 จึงเรียกว่าเป็นค่าใช้จ่ายที่จำเป็น (necessity) ค่าใช้จ่ายในหมวดเดียวกันเป็นได้ทั้งค่าใช้จ่ายที่ฟุ่มเฟือยและค่าใช้จ่ายที่จำเป็น ขึ้นอยู่กับรายได้

ระดับอิ่มตัวจะนำมาใช้กับค่าใช้จ่ายหมวดที่ผู้บริโภคจะไม่มีการเปลี่ยนแปลงไปใช้สินค้าอื่นที่มีคุณภาพดีกว่า แม้ว่าจะมีรายได้เพิ่มขึ้น แต่โดยทั่วไปแล้วมักจะมีการเปลี่ยนแปลงไปใช้สินค้าที่มีคุณภาพดีกว่าอยู่เสมอ

3. รูปแบบทางคณิตศาสตร์ของเส้นโค้ง Engel มีข้อบังคับที่กำหนดขึ้นเกี่ยวกับค่าสัมประสิทธิ์ความถดถอย เพื่อให้ผลบวกของค่าใช้จ่ายทุกหมวดรวมกันเท่ากับรายได้ ณ ทุก ๆ ระดับของรายได้ หรือ ระดับของยอดรวมค่าใช้จ่ายทุกหมวด เมื่อใช้ข้อมูลยอดรวมค่าใช้จ่ายแทนรายได้ เช่น ใน linear form

$$Y_i = A_i + B_i X + E_i$$

และถ้าให้ $W_i = Y_i/X$ คือ สัดส่วนของค่าใช้จ่ายหมวดที่ i ในยอดรวมค่าใช้จ่ายทุกหมวด X

แล้วจะได้ว่า $\sum_i Y_i = X$ $\sum_i W_i = 1$

ซึ่งคุณสมบัติทั้งสองประการนี้ เรียกว่า additive property ซึ่ง linear form ที่มีคุณสมบัติ additive property ต้องมีข้อจำกัดเกี่ยวกับค่าประมาณของสัมประสิทธิ์ความถดถอย หรือที่เรียกว่า additive criterion คือ

$$\left.\begin{aligned}\sum_i A_i &= 0\\ \sum_i B_i &= 1\end{aligned}\right\} \text{จึงจะได้} \quad \begin{aligned}\sum_i Y_i &= X\\ \sum_i W_i &= 1\end{aligned}$$

สิ่งที่ตามมาจาก additive property คือ เส้นโค้ง Engel ไม่สามารถมี asymptote หรือระดับอิ่มตัวทุกหมวด ถ้าบางหมวดมีระดับอิ่มตัว จะมีบางหมวดที่ไม่มีระดับอิ่มตัว เช่น สินค้าที่มีคุณภาพต่ำจะไม่มีระดับอิ่มตัว เมื่อผู้บริโภคมีรายได้เพิ่มขึ้นจะหันไปใช้สินค้าอื่นแทน หากทุกหมวดมีระดับอิ่มตัวแล้ว จะมีรายได้บางระดับที่ไม่มีการใช้จ่ายเพื่อการอุปโภคบริโภคเลย ซึ่งย่อมเป็นไปไม่ได้

คุณสมบัติของ additive property จะช่วยป้องกันมิให้เกิดพฤติกรรมที่หลอกลวงขึ้น แต่ในการศึกษาเพื่อหารูปแบบทางคณิตศาสตร์ที่เหมาะสมของเส้นโค้ง Engel สำหรับค่าใช้จ่ายในหมวดต่าง ๆ ไม่ควรที่จะผูกมัดโดยใช้แต่เพียงรูปแบบเดียวเพื่อให้ $\sum_i A_i = 0$ และ $\sum_i B_i = 1$ รูปแบบทางคณิตศาสตร์ที่แสดงพฤติกรรมของผู้บริโภคได้มากกว่านั้นควรได้จากการใช้หลาย ๆ รูปแบบสำหรับแต่ละหมวดค่าใช้จ่ายที่แตกต่างกันออกไป [1]

---

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed. p. 84-86.

ในความเป็นจริง คุณสมบัติ additive property ไม่ได้ใช้ประโยชน์ในการตรวจสอบรูปแบบทางคณิตศาสตร์เท่าใดนัก มีทฤษฎีที่เกี่ยวกับการประมาณค่าสัมประสิทธิ์ความถดถอยของเส้นโค้ง Engel โดยวิธีกำลังสองน้อยที่สุด แล้วผลที่ได้มีคุณสมบัติตรงตาม additive property [1] ซึ่งสาระสำคัญของทฤษฎีนี้ได้พิสูจน์โดย Nicholson [2]

ทฤษฎีนี้กล่าวว่า " ถ้าใช้รูปแบบทางคณิตศาสตร์แบบเดียวกระชับ (fit) เข้ากับเส้นโค้ง Engel ของค่าใช้จ่ายทุก ๆ หมวด โดยให้รูปแบบนั้นมีคุณสมบัติ additive property แล้ว ค่าประมาณของเส้นโค้ง ที่ได้โดยวิธีกำลังสองน้อยที่สุด จะมีคุณสมบัติ additive property ด้วย นั่นคือถ้าค่าสังเกตุเบื้องต้นมี $\sum_i Y_i = X$ แล้วจะได้ $\sum_i \hat{Y}_i = X$ ด้วยเมื่อ $\hat{Y}_i$ คือค่าประมาณของ Y_i "

หรือจะกล่าวได้ว่า เมื่อต้องการสมการชุดหนึ่งซึ่งอยู่ในลักษณะที่ค่าประมาณจากสมการมีคุณสมบัติ additive property เช่นเดียวกับข้อมูลที่สังเกตุได้เบื้องต้น สมการชุดนี้จะหาได้โดยใช้วิธีกำลังสองน้อยที่สุดกับกลุ่มตัวแปรอิสระชุดเดิม แม้ว่าค่าประมาณของเส้นโค้งที่หาได้จะไม่มีคุณสมบัติ additive property อย่างแน่ชัดโดยตรงทีเดียว ก็ยังสามารถเรียกได้ว่ามีคุณสมบัติ additive property โดยประมาณ ถ้าข้อมูลเบื้องต้น (basic data) ที่ใช้หาสมการชุดนี้มีคุณสมบัติ additive property และเส้นโค้งนั้นสามารถกระชับกับข้อมูลได้ดี เราก็สามารถที่จะให้ความสำคัญของคุณสมบัตินี้ลดน้อยลงได้

---

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed. p. 84-86.

2 Nicholson, (1949), p. 388-389.

ดังนั้น ในการศึกษาเพื่อตรวจสอบรูปแบบทางคณิตศาสตร์ของสมการความถดถอย หรือระนาบ Engel สำหรับค่าใช้จ่ายหมวดต่าง ๆ ของครัวเรือนในภาคเหนือและกรุงเทพมหานคร ซึ่งมีพฤติกรรมปกติ และสามารถเลือกใช้รูปแบบต่าง ๆ ได้ ครั้งนี้ จะตรวจสอบรูปแบบอย่างเป็นอิสระจากข้อบังคับ additive criterion [1]

รูปแบบทางคณิตศาสตร์ของสมการความถดถอยทั้งสามรูปแบบ มีคุณสมบัติดังต่อไปนี้

---

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed. p. 84-86.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 2.2

คุณสมบัติของสมการความถดถอยจำแนกตามรูปแบบ

คุณสมบัติ	Linear form	Double-logarithmic form	Semi-logarithmic form
1. รายได้ในระยะเริ่มแรกที่ยังไม่มีการใช้จ่ายในหมวดที่ i (initial income)	$-(A_i+C_iN)/B_i$	0	$e^{-(A_i+C_i \log N)/B_i}$
2. ระดับอิ่มตัว (satiety level)	อนันต์ (infinity)	อนันต์	อนันต์
3. Additive property	มี additive property โดยมี additive criterion คือ $\sum_i A_i=0, \sum_i B_i=1, \sum_i C_i=0$	ไม่มี additive propety	ไม่มี additive property
4. ความแตกต่างของการใช้จ่ายเนื่องจากความแตกต่างของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด (marginal propensity to consume)	B_i	Y_iB_i/X	B_i/X

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

i 17231164

ตาราง 2.2 (ต่อ)

คุณสมบัติของสมการความถดถอยจำแนกตามรูปแบบ

คุณสมบัติ	Linear form	Double-logarithmic form	Semi-logarithmic form
5. ความยืดหยุ่นของการใช้จ่ายเนื่องจากรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด (income elasticity หรือ total expenditure elasticity)	XB_i/Y_i	B_i	B_i/Y_i
6. ความแตกต่างของการใช้จ่ายเนื่องจากความแตกต่างของขนาดครัวเรือน (economies of scale)	C_i	Y_iC_i/N	C_i/N

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ตาราง 2.2 (ต่อ)

คุณสมบัติของสมการความถดถอยจำแนกตามรูปแบบ [1, 2]

คุณสมบัติ	Linear form	Double-logarithmic form	Semi-logarithmic form
7. ความยืดหยุ่นของการใช้จ่ายเนื่องจากขนาดครัวเรือน (household size elasticity)	NC_i/Y_i	C_i	C_i/Y_i

1 Prais and Houthakker, The Analysis of Family Budgets, 2d ed., p. 82-88.

2 Leser, Econometric Techniques and Problems, (London: Charles Griffin & Co., Ltd., 1969), p. 77-81.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

ใน linear form มีรายได้ในระยะเริ่มแรกที่ยังไม่มีการใช้จ่ายเท่ากับ $-(A_i + C_iN)/B_i$ และ satiety level เท่ากับอนันต์ (infinity) หมายความว่า ครัวเรือนจะยังไม่ใช้จ่ายเงินเพื่อการอุปโภคบริโภคในหมวดที่ i ณ ระดับรายได้ที่น้อยกว่าหรือเท่ากับ $-(A_i + C_iN)/B_i$ หรือจะกล่าวได้อีกอย่างหนึ่งว่า ครัวเรือนจะเริ่มใช้จ่ายเงินเพื่อการอุปโภคบริโภคในหมวดที่ i เมื่อมีหนี้สินน้อยกว่า $-(A_i + C_iN)/B_i$ และปริมาณที่จะใช้จ่ายเงินในหมวดที่ i ไม่มีขีดจำกัดไม่ว่าจะมีรายได้มากสักเพียงใด linear form จะมีคุณสมบัติ additive property คือ ผลบวกของค่าใช้จ่ายทุกหมวดเท่ากับรายได้ หรือ ยอดรวมค่าใช้จ่ายทุกหมวด ณ ทุก ๆ ระดับของรายได้ก็จะต้องมีข้อจำกัดเกี่ยวกับค่าประมาณของสัมประสิทธิ์ความถดถอย หรือที่เรียกว่า additive criterion คือ ค่า A_i ของค่าใช้จ่ายทุกหมวดรวมกันเป็น 0 ค่าสัมประสิทธิ์ของตัวแปรยอดรวมค่าใช้จ่ายทุกหมวดในทุก ๆ หมวดค่าใช้จ่ายต้องมีผลบวกเท่ากับ 1 และค่าสัมประสิทธิ์ของตัวแปรขนาดครัวเรือนในค่าใช้จ่ายทุก ๆ หมวดรวมกันเป็น 0 ในด้านความแตกต่างของการใช้จ่ายอันเนื่องมาจากความแตกต่างของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด (marginal propensity to consume) มีค่าคงที่เท่ากับ B_i หมายความว่าถ้าครัวเรือนสองครัวเรือนมีรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวดต่างกัน 1 หน่วย แต่มีขนาดครัวเรือนเท่ากัน จะมีการใช้จ่ายหมวดที่ i แตกต่างกัน B_i หน่วย แต่ถ้าพิจารณาเปรียบเทียบอัตราความแตกต่างของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวดกับอัตราความแตกต่างของการใช้จ่ายในหมวดที่ i จะได้ว่าความยืดหยุ่นของการใช้จ่ายในหมวดที่ i อันเนื่องมาจากรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด (income elasticity หรือ total expenditure elasticity) เท่ากับ XB_i/Y_i นั่นคือ ณ ระดับรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด X ของครัวเรือนหนึ่ง มีการใช้จ่าย

ในหมวดที่ i เท่ากับ Y_i เปรียบเทียบการใช้จ่ายของครัวเรือนนี้กับอีกครัวเรือนหนึ่ง ที่มีขนาดของครัวเรือนเท่ากัน ครัวเรือนที่สองจะมีอัตราการใช้จ่ายที่แตกต่างจากครัวเรือนแรกเท่ากับ XB_i/Y_i เท่าของอัตรารายได้ที่แตกต่างกันของสองครัวเรือน จะเห็นได้ว่าค่าความยืดหยุ่นของการใช้จ่ายอันเนื่องจากรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด มีค่าแปรเปลี่ยนตามค่าของยอดรวมค่าใช้จ่ายทุกหมวดและส่วนกลับของค่าใช้จ่ายหมวดที่ i ทำนองเดียวกันความแตกต่างของการใช้จ่าย เนื่องจากความแตกต่างของขนาดครัวเรือน (economies of scale) มีค่าคงที่เท่ากับ C_i ซึ่งเป็นสัมประสิทธิ์ของขนาดครัวเรือน มีความหมายว่า ครัวเรือนสองครัวเรือนที่มีรายได้หรือยอดรวมค่าใช้จ่ายเท่ากัน แต่มีจำนวนสมาชิกในครัวเรือนแตกต่างกัน 1 คน จะมีการใช้จ่ายในหมวดที่ i แตกต่างกัน C_i หน่วย แต่ถ้าพิจารณาเปรียบเทียบอัตราความแตกต่างของจำนวนสมาชิกในครัวเรือนกับอัตราความแตกต่างของการใช้จ่าย จะได้ว่า ค่าความยืดหยุ่นของการใช้จ่ายในหมวดที่ i อันเนื่องมาจากขนาดครัวเรือน (household size elasticity) เท่ากับ NC_i/Y_i นั่นคือ ถ้าครัวเรือนหนึ่งมีจำนวนสมาชิก N คน มีการใช้จ่ายในหมวดที่ i เท่ากับ Y_i เปรียบเทียบการใช้จ่ายของครัวเรือนนี้กับอีกครัวเรือนหนึ่งที่มีรายได้หรือยอดรวมค่าใช้จ่ายเท่ากัน ครัวเรือนที่สองจะมีอัตราการใช้จ่ายที่แตกต่างจากครัวเรือนแรกเท่ากับ NC_i/Y_i เท่าของอัตราจำนวนสมาชิกที่แตกต่างกันของสองครัวเรือน ค่าความยืดหยุ่นของการใช้จ่ายอันเนื่องมาจากขนาดของครัวเรือนนี้ มีค่าแปรเปลี่ยนตามขนาดของครัวเรือนและส่วนกลับของค่าใช้จ่ายในหมวดที่ i จะเห็นได้ว่า linear form ให้ค่าความแตกต่างของการใช้จ่ายเนื่องจากความแตกต่างของรายได้และขนาดครัวเรือนที่มีค่าคงที่เสมอ

สำหรับ Double-logarithmic form ด้วยโครงสร้างของรูปแบบแสดงให้เห็นว่ารายได้ในระยะเริ่มแรกที่ยังไม่มีการใช้จ่ายเป็น 0 หมายความว่าครัวเรือนจะเริ่มใช้จ่ายเงินในหมวดที่ i เมื่อเริ่มมีรายได้ แต่จะไม่มีการใช้จ่ายเงินเพื่อการอุปโภคบริโภคในหมวดที่ i ถ้าครัวเรือนไม่มีรายได้ และปริมาณที่จะใช้จ่ายเงินในหมวดที่ i ไม่มีขีดจำกัดไม่ว่าจะมีรายได้มากสักเพียงใด Double-logarithmic form ไม่มีคุณสมบัติของ additive property เนื่องจากไม่สามารถหา additive criterion ที่ทำให้สมการมีคุณสมบัติของ additive property ได้ ในด้านความแตกต่างของการใช้จ่ายในหมวดที่ i เนื่องจากความแตกต่างของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด เมื่อกำหนดให้ขนาดของครัวเรือนคงที่ จะมีความแตกต่างเท่ากับ $Y_i B_i/X$ ซึ่งมีค่าแปรเปลี่ยนตามค่าใช้จ่ายในหมวดที่ i และส่วนกลับของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด แต่ความยืดหยุ่นของการใช้จ่ายในหมวดที่ i เนื่องจากรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวดมีค่าคงที่เท่ากับ B_i ซึ่งเป็นสัมประสิทธิ์ของ log ฐาน e ของยอดรวมค่าใช้จ่ายทุกหมวด ส่วนความแตกต่างของการใช้จ่ายอันเนื่องจากความแตกต่างของขนาดของครัวเรือน ในขณะที่รายได้หรือยอดรวมค่าใช้จ่ายทุกหมวดมีค่าคงที่นั้น มีความแตกต่างเท่ากับ $Y_i C_i/N$ ซึ่งมีค่าแปรเปลี่ยนตามค่าใช้จ่ายในหมวดที่ i และส่วนกลับของขนาดของครัวเรือน ความยืดหยุ่นของการใช้จ่ายในหมวดที่ i เนื่องจากขนาดครัวเรือนมีค่าคงที่เท่ากับ C_i ซึ่งเป็นสัมประสิทธิ์ของ log ฐาน e ของขนาดของครัวเรือน จะเห็นได้ว่า double-logarithmic form ให้ค่าความยืดหยุ่นที่มีค่าคงที่เสมอไม่ว่าจะเป็นค่าความยืดหยุ่นเนื่องจากรายได้หรือขนาดครัวเรือน

ส่วน semi-logarithmic form มี initial income เท่ากับ $e^{-(A_i + C_i \log N)/B_i}$ ซึ่งจะมีค่าเป็นบวกเสมอไม่ว่าค่าของ A_i, B_i และ C_i จะเป็นบวกหรือลบ นั่นคือ ในครัวเรือนที่มีจำนวนสมาชิกเท่ากัน จะเริ่มใช้จ่ายเงินเพื่อการอุปโภคบริโภคในหมวดที่ i เมื่อมีรายได้มากกว่า $e^{-(A_i + C_i \log N)/B_i}$ และปริมาณที่จะใช้จ่ายเงินในหมวดที่ i ไม่มีขีดจำกัดไม่ว่าจะมีรายได้มากสักเพียงใด semi-logarithmic form ไม่มีคุณสมบัติของ additive property เช่นเดียวกับ double-logarithmic form ในด้านความแตกต่างของการใช้จ่ายอันเนื่องจากความแตกต่างของรายได้หรือยอดรวมค่าใช้จ่ายทุกหมวด เมื่อกำหนดให้ขนาดของครัวเรือนมีค่าคงที่ นั้น มีความแตกต่างเท่ากับ B_i/X ซึ่งมีค่าแปรเปลี่ยนตามส่วนกลับของยอดรวมค่าใช้จ่ายทุกหมวด ส่วนความยืดหยุ่นของการใช้จ่ายในหมวดที่ i เนื่องจากรายได้หรือยอดรวมค่าใช้จ่ายมีค่าเท่ากับ B_i/Y_i และความยืดหยุ่นของการใช้จ่ายในหมวดที่ i เนื่องจากขนาดครัวเรือนมีค่าเท่ากับ C_i/Y_i ซึ่งค่าความยืดหยุ่นทั้งสองนี้ มีค่าแปรเปลี่ยนตามส่วนกลับของค่าใช้จ่ายในหมวดที่ i ส่วนความแตกต่างของการใช้จ่ายในหมวดที่ i เนื่องจากความแตกต่างของขนาดครัวเรือนมีค่าเท่ากับ C_i/N ซึ่งมีค่าแปรเปลี่ยนตามส่วนกลับของขนาดครัวเรือน

<u>การวิเคราะห์ปัญหาทางเศรษฐมิติ</u>

การศึกษาวิจัยรูปแบบการใช้จ่ายของครัวเรือนในภาคเหนือและกรุงเทพ–มหานคร เป็นการนำวิธีการทางสถิติมาประยุกต์ใช้กับปัญหาทางเศรษฐกิจที่เกิดขึ้นในสังคม ตัวแปรที่นำมาวิเคราะห์หารูปแบบการใช้จ่ายของครัวเรือนนี้ ได้แนวทางจากความสัมพันธ์ทางด้านเศรษฐกิจที่ว่า รายได้หรือยอดรวมค่าใช้จ่ายทุกหมวดและจำนวนสมาชิกในครัวเรือน เป็นตัวบ่งชี้การใช้จ่ายของครัวเรือน หรือเป็นตัวแปรที่ทำให้เกิด

การเปลี่ยนแปลงในการใช้จ่าย ในระเบียบวิธีการทางสถิติ เมื่อใช้วิธีกำลังสองน้อยที่สุดหาสมการความถดถอย จะต้องมีข้อสมมติเบื้องต้นเกี่ยวกับตัวแปรอิสระและตัวคลาดเคลื่อน ดังที่ได้กล่าวไว้ในหัวข้อการวิเคราะห์เชิงสถิติของสมการความถดถอยแล้ว แต่ในการนำวิธีกำลังสองน้อยที่สุดมาประยุกต์ใช้กับปัญหาทางเศรษฐกิจมักประสบปัญหาที่ตัวแปรอิสระ และตัวคลาดเคลื่อน ไม่มีคุณสมบัติตามข้อสมมติ [1] ฉะนั้น จึงควรพิจารณาถึงปัญหาที่อาจเกิดขึ้น เช่น ปัญหาการเกิด autocorrelation ปัญหาการเกิด multicollinearity เป็นต้น

1. ปัญหาการเกิด Autocorrelation

จากข้อสมมติของตัวคลาดเคลื่อนที่ว่า ในประชากรตัวคลาดเคลื่อนแต่ละตัวไม่มีความสัมพันธ์กัน หรือค่าแปรปรวนร่วมเป็น 0 นั่นคือ $\text{Cov}(E_j, E_k) = E(E_j E_k) = 0$ เมื่อ $j \neq k$ ตัวคลาดเคลื่อนอยู่ในลักษณะที่เรียกว่า nonautocorrelation ซึ่งหมายความว่า ตัวคลาดเคลื่อนที่เกิดขึ้น ณ จุดหนึ่งของค่าสังเกตุ หรือของแต่ละหน่วย ตัวอย่างเช่น ครัวเรือน เป็นตัวคลาดเคลื่อนสุ่ม (random error) ที่มีการกระจายปกติ และเป็นอิสระจากกันเป็นคู่ ๆ (pairwise independent) สำหรับข้อมูล ณ ช่วงเวลาต่าง ๆ กัน (time series data) การเกิด autocorrelation หมายความว่า ตัวคลาดเคลื่อนในลำดับต่าง ๆ กัน เป็นอิสระจากกัน (serial independent) หรือ ตัวคลาดเคลื่อนที่เกิดขึ้นในช่วงเวลาหนึ่งจะไม่มีผลกระทบกระเทือนไปถึงตัวคลาดเคลื่อนที่เกิดขึ้นในอีกช่วงเวลาหนึ่ง

---

1 Leser, Econometric Techniques and Problems, p. 14, 27.

ส่วนข้อมูลของหน่วยต่าง ๆ ที่เกิดขึ้นในช่วงเวลาเดียวกัน (cross-sectional data) การเกิด autocorrelation จะหมายความว่า ในขณะใดขณะหนึ่ง ตัวคลาดเคลื่อนที่เกิดขึ้นกับค่าสังเกตุหนึ่งหรือครัวเรือนหนึ่ง จะไม่มีความสัมพันธ์ หรือถูกกระทบกระเทือน โดยตัวคลาดเคลื่อนที่เกิดขึ้นกับอีกค่าสังเกตุหรืออีกครัวเรือนหนึ่ง เช่น ในการศึกษาการใช้จ่ายของครัวเรือน การใช้จ่ายของครัวเรือนหนึ่งผิดไปจากปกติ สิ่งที่ผิดปกติไปนี้จะไม่มีผลกระทบกระเทือนต่อสิ่งผิดปกติของการใช้จ่ายของครัว-เรือนอื่น ๆ

เมื่อตัวคลาดเคลื่อนในประชากรไม่เป็นไปตามข้อสมมติจะเกิดปัญหา autocorrelation นั่นคือ ตัวคลาดเคลื่อนที่เกิดขึ้น ณ ค่าสังเกตุหนึ่งหรือครัวเรือนหนึ่ง มีความสัมพันธ์กับตัวคลาดเคลื่อนที่เกิดขึ้นกับอีกค่าสังเกตุหนึ่งหรือครัวเรือนอื่น เช่น มีความสัมพันธ์กันในรูป

<u>first order autocorrelation</u>[2] $E_j = \rho E_{j-1} + u_j$

$$E(E_j E_{j-1}) = \rho \sigma^2$$

<u>second order autocorrelation</u>[2] $E_j = \rho E_{j-2} + u_j$

$$E(E_j E_{j-2}) = \rho^2 \sigma^2$$

เมื่อเกิดปัญหา autocorrelation จะมีผลกระทบกระเทือนต่อค่าประมาณของสัมประสิทธิ์ความถดถอย ในสมการความถดถอยที่ประมาณได้ด้วยวิธีกำลังสองน้อยที่สุด [1, 2] คือ

---

1 J. Johnston, <u>Econometric Methods</u>, 2d ed. (New York: McGraw-Hill Book Company, 1960), p. 243-252.

2 Jan Kmenta, <u>Elements of Econometrics</u>, p. 269.

1. ยังคงได้ค่าประมาณของสัมประสิทธิ์ความถดถอยที่มีคุณสมบัติเป็นค่าประมาณที่ไม่เอนเอียง (unbiased estimates) [E(a) = A, E(b) = B และ E(c) = C] และค่าประมาณที่แนบนัย (consistency estimates) [ค่าแปรปรวนของค่าประมาณมีค่าเข้าใกล้ 0 เมื่อขนาดตัวอย่างเพิ่มขึ้นใกล้อนันต์ นั่นคือ $\lim_{m\to\infty} V(a) = 0$, $\lim_{m\to\infty} V(b) = 0$ และ $\lim_{m\to\infty} V(c) = 0$]

2. ค่าแปรปรวนของค่าประมาณในสมการถดถอยมีค่าใหญ่มากโดยไม่มีเหตุผลเมื่อเปรียบเทียบกับค่าแปรปรวนของค่าประมาณได้มาโดยวิธีอื่น จึงทำให้การทำนายค่าในอนาคตมีความแม่นยำน้อย ประสิทธิภาพลดลง (inefficient prediction) การที่ ค่าแปรปรวนของค่าประมาณมีค่าเอนเอียง (biased) จะทำให้การทดสอบสมมติฐานเชื่อถือไม่ได้ และการหาอันตรภาคแห่งความเชื่อมั่นของค่าประมาณย่อมคลาดเคลื่อนไปด้วย

3. การหาค่าแปรปรวนของค่าประมาณในสมการความถดถอย ถ้ายังคงใช้สูตรเดิมที่เคยใช้กับสมการความถดถอยที่มีตัวคลาดเคลื่อนเป็นไปตามข้อสมมติคือเป็นตัวคลาดเคลื่อนสุ่มแล้ว จะได้ค่าแปรปรวนจากตัวอย่าง น้อยกว่าค่าที่แท้จริง

การทดสอบถึงการเกิด autocorrelation ใช้ Durbin-Watson test พิจารณาค่า Durbin-Watson "d" statistic โดยตั้งสมมติฐานว่า ไม่เกิด autocorrelation ($\rho = 0$) แล้วตั้งสมมติฐานแย้งว่าเกิด positive first-order autocorrelation ($\rho > 0$) คือตัวคลาดเคลื่อนในประชากรมีความสัมพันธ์กันในทางบวก

$H_o : \rho = 0$, $H_1 : \rho > 0$

$$d = \sum_{j=2}^{m} (e_j - e_{j-1})^2 / \sum_{j=1}^{m} e_j^2$$

เมื่อ e_j คือตัวคลาดเคลื่อนของครัวเรือนที่ j ในตัวอย่าง เป็นค่าผลต่างระหว่างค่าสังเกตุของครัวเรือนที่ j กับค่าที่ประมาณด้วยสมการความถดถอย นั่นคือค่าของ $(Y_j - \hat{Y}_j)$

d มีค่าตั้งแต่ 0 ถึง 4 มีค่ากึ่งกลางคือ 2

ในการทดสอบที่ตั้งสมมติฐานแย้งว่าเกิด positive autocorrelation นั้น d ต้องมีค่าน้อยกว่า หรือเท่ากับ 2 เปรียบเทียบค่า d ที่คำนวณได้กับค่า d ในตารางสถิติ ณ ระดับความมีนัยสำคัญ α

ถ้า d มีค่าน้อยกว่า d_L จะปฏิเสธสมมติฐาน

ถ้า d มีค่ามากกว่า d_U จะยอมรับสมมติฐาน

ถ้า d อยู่ระหว่าง d_L และ d_U การทดสอบสมมติฐานจะสรุปผลไม่ได้

ถ้า d ที่คำนวณได้จากตัวอย่างมีค่ามากกว่า 2 สมมติฐานแย้งจะเปลี่ยนเป็น H_1: negative first-order autocorrelation ($\rho < 0$) คือ ตัวคลาดเคลื่อนในประชากรมีความสัมพันธ์กันในทางลบ ให้หาค่า $4-d_L$, $4-d_U$ แล้วทดสอบเช่นเดียวกับกรณีแรก คือ

ถ้า d มีค่าน้อยกว่า $4-d_U$ จะยอมรับสมมติฐาน

ถ้า d มีค่ามากกว่า $4-d_L$ จะปฏิเสธสมมติฐาน

ถ้า d อยู่ระหว่าง $4-d_U$ และ $4-d_L$ การทดสอบสมมติฐานจะสรุปผลไม่ได้

2. ปัญหาการเกิด Multicollinearity

ข้อสมมติเกี่ยวกับตัวแปรอิสระในสมการความถดถอยกำหนดว่า จะต้องไม่มีความสัมพันธ์กันโดยตรงอย่างแน่ชัดในเชิงเส้นตรง (exact linear relation) ระหว่างตัวแปรอิสระทั้งหลาย เมื่อข้อสมมตินี้ถูกทำลายไปจึงเกิดปัญหา multicollinearity ซึ่งจะเกิดขึ้นในสมการความถดถอยที่มีตัวแปรอิสระมากกว่าหนึ่งตัว (multiple regression) โดยตัวแปรอิสระบางตัวหรือทั้งหมดมีความสัมพันธ์ซึ่งกันและกันอย่างสมบูรณ์ (exact linear relation) หรือมีความสัมพันธ์ค่อนข้างสูง (almost exact linear relation) และเนื่องจากตัวแปรอิสระเหล่านี้ได้สมมติไว้ว่าเป็นค่าที่กำหนดไว้ล่วงหน้า (nonstochastic) การศึกษาถึงความสัมพันธ์ของตัวแปรอิสระหรือปัญหา multicollinearity จึงเป็นเรื่องของตัวอย่างมิใช่ประชากร

ผลที่จะเกิดขึ้นอันเนื่องมาจากการเกิด multicollinearity คือ

1. ความถูกต้อง เที่ยงตรง และแม่นยำของการประมาณค่าในสมการความถดถอยลดน้อยลง เนื่องจากไม่สามารถลดความพัวพันที่มีกระทบกระเทือนถึงกันของตัวแปรอิสระหลาย ๆ ตัวได้ ความแม่นยำที่ลดน้อยลงหรือหมดไปนี้เป็นไปได้ 3 ลักษณะ คือ ค่าประมาณของสัมประสิทธิ์ความถดถอยแต่ละตัวผิดไปจากค่าที่แท้จริงอาจไม่แตกต่างจาก 0 อย่างมีนัยสำคัญ ความคลาดเคลื่อนของการประมาณแต่ละตัวมีความสัมพันธ์กันค่อนข้างสูง และค่าแปรปรวนของค่าประมาณของสัมประสิทธิ์ความถดถอยมีค่าสูง แต่ค่าแปรปรวนที่มีค่าสูงก็มิใช่ตัวบ่งชี้ถึงการเกิด multicollinearity เสมอไป บางครั้งผลการประมาณค่าอาจจะเชื่อถือไม่ได้ แต่ก็มิได้แสดงผลออกมาในรูปของค่าแปรปรวนที่มีค่าสูง

2. เมื่อเกิด multicollinearity บางครั้งผู้วิจัยอาจถูกชักจูงให้เข้าใจผิดไปตัดตัวแปรอิสระบางตัวออกไปจากการวิเคราะห์เพื่อให้ได้ค่าประมาณของสัมประสิทธิ์ความถดถอยที่เชื่อถือได้มากกว่าเดิม เนื่องจากค่าสัมประสิทธิ์ของตัวแปรอิสระนั้นมีค่าไม่แตกต่างจาก 0 อย่างมีนัยสำคัญ ทั้ง ๆ ที่ในความเป็นจริงตัวแปรอิสระนั้นมีความสัมพันธ์กับตัวแปรตามแต่ถูกพัวพันด้วยอิทธิพลของตัวแปรอิสระตัวอื่นและไม่สามารถแยกอิทธิพลที่พัวพันกันอยู่ออกจากกันได้

3. การประมาณค่าสัมประสิทธิ์ความถดถอย มีความไวต่อกลุ่มของข้อมูลจากตัวอย่าง นั่นคือ ค่าประมาณจะเปลี่ยนแปลงไปได้ง่าย เมื่อข้อมูลเปลี่ยนแปลงไป การเพิ่มจำนวนค่าสังเกตุเข้าไปเพียงเล็กน้อยบางครั้งทำให้เกิดการเปลี่ยนแปลงอย่างกระทันหันของค่าประมาณของสัมประสิทธิ์ความถดถอยบางตัว [1]

การวัดระดับของการเกิด multicollinearity มีวิธีการวัดได้หลายวิธี [2] ซึ่งจะศึกษาได้จากหนังสือเศรษฐมิติทั่ว ๆ ไป แต่ในสมการความถดถอยที่มีตัวแปรอิสระเพียงสองตัว การวัดการเกิด multicollinearity ทำได้ง่าย ๆ โดยทดสอบจากค่าอัมประสิทธิ์สหสัมพันธ์ และทดสอบสมมติฐาน โดยใช้ t-test ตั้งสมมติฐานว่า ไม่เกิด multicollinearity ($\rho = 0$) และตั้งสมมติฐานแย้งว่าเกิด multicollinearity ($\rho \neq 0$)

---

1 J. Johnston, Econometric Methods, 2d ed. (New York: McGraw-Hill Book Company, 1960), p. 160.

2 Jan Kmenta, Elements of Econometrics, p. 389.

$H_o : \rho = 0$, $H_1 : \rho \neq 0$

เปรียบเทียบค่าสัมบูรณ์ของ t ที่คำนวณได้กับ $t_{(n-2,\ \frac{1}{2}\alpha)}$ ในตาราง t- distribution ที่ระดับความมีนัยสำคัญ α

ถ้า t มีค่ามากกว่าในตารางอย่างมีนัยสำคัญ จะปฏิเสธสมมติฐาน

ถ้า t มีค่าไม่มากกว่าในตารางอย่างมีนัยสำคัญ จะยอมรับสมมติฐาน

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย