บทที่ ๒

วิธีดำเนินการวิจัย

เนื่องจากข้อมูลซึ่งเก็บรวบรวมมาเพื่อการศึกษานี้เป็นข้อมูลประเภทอนุกรมเวลา ซึ่งโดยทั่วไปแล้ว อาจเขียนในรูปของโมเดล (Model) ได้ดังนี้คือ

$$Y = TSCI$$

ซึ่งเรียกว่าเป็นโมเดลเชิงคูณ (Multiplicative Model)

เมื่อ Y คือ ตัวแปรตาม (Dependent Variable) ที่เราต้องการศึกษา เป็นตัวแปรที่เราคาดว่าจะขึ้นอยู่กับแนวโน้ม (T) การเปลี่ยนแปลงเนื่องจากฤดูกาล (S) การเปลี่ยนแปลงแบบวัฏจักร (C) และการเปลี่ยนแปลงเนื่องจากเกิดเหตุการณ์ผิดปกติ เช่น ราคาสินค้าออก ปริมาณสินค้าออก ปริมาณนักท่องเที่ยว ฯลฯ

T คือ แนวโน้ม (Trend หรือ Secular Trend) ซึ่งหมายถึงแบบแผนซึ่งแสดงถึงการเคลื่อนไหวของข้อมูลภายในช่วงระยะเวลายาว (ประมาณ ๑๐ ปีขึ้นไป)

S คือการเปลี่ยนแปลงเนื่องจากฤดูกาลหรือเทศกาล (Seasonal Variation) ข้อมูลที่ศึกษาเป็นช่วงเวลาที่เล็กกว่าปี เช่น เป็นรายวัน รายสัปดาห์ รายเดือนหรือราย ๓ เดือน เป็นต้น ในช่วงเวลาที่เหมือนกัน ข้อมูลมีการเปลี่ยนแปลงคล้ายคลึงกัน

C คือ การเปลี่ยนแปลงแบบวัฏจักร (Cyclical Movement) อันหมายถึงเหตุการณ์ที่เกิดขึ้นเหมือนกันโดยอาศัยรอบของวัฏจักรนานหลายปีคล้ายแนวโน้ม แต่รูปร่างของการเปลี่ยนแปลงนั้นแตกต่างกันไปตั้งแต่ระยะรุ่งเรืองที่สุดถึงจุดต่ำสุด

I คือ การเปลี่ยนแปลงเนื่องจากเกิดเหตุการณ์ผิดปกติ (Irregular Movement) อันเป็นไปตามธรรมชาติ เช่น แผ่นดินไหว น้ำท่วม ไฟไหม้ หรือเป็นวิกฤตการณ์ทางเศรษฐกิจ

ในโมเดลเชิงฤดูกาลดังกล่าวข้างต้น มักจะพบค่าแนวโน้มและการเปลี่ยนแปลงเนื่องจากฤดูกาลปรากฏอยู่เสมอ ดังนั้น ผู้วิจัยจึงจะทำการวิเคราะห์หาค่าแนวโน้ม ซึ่งเป็นการศึกษาหาความสัมพันธ์ระหว่างเวลากับข้อมูลที่สนใจศึกษา โดยมีเวลา (X) เป็นตัวแปรอิสระ (Dependent Variable) และข้อมูลที่สนใจศึกษา (Y) เป็นตัวแปรตาม (Independent Variable) ในกรณีที่ข้อมูลที่สนใจศึกษามีการเคลื่อนไหวแบบฤดูกาล (Seasonal Movement) คือ มีแนวโน้มที่จะสูงขึ้นหรือต่ำลงในบางเดือน ก็จะทำการหาดัชนีฤดูกาล (Seasonal Index) ด้วย

๒.๑ โมเดลของแนวโน้ม

ในการศึกษาเพื่อทราบว่าโมเดลแนวโน้มของข้อมูลที่สนใจเป็นแบบใดนั้น ควรจะมีการทดสอบก่อนว่าข้อมูลดังกล่าวมีแนวโน้มหรือไม่ วิธีการที่ใช้ทดสอบคือการพิจารณาจำนวนชุดของข้อมูลที่อยู่สูงกว่าและต่ำกว่ามัธยฐาน (Runs above and below the Median)[1] ซึ่งดำเนินการได้ดังนี้

๑. นำข้อมูลอนุกรมเวลาที่ต้องการศึกษามาหาค่ามัธยฐาน (Median)

๒. พิจารณาข้อมูลและให้สัญลักษณ์ a หรือ b แก่ข้อมูลอนุกรมเวลาชุดนี้

เมื่อ a คือ ค่าข้อมูลที่สูงกว่าค่ามัธยฐาน

b คือ ค่าข้อมูลที่ต่ำกว่าค่ามัธยฐาน

---

[1] John E.Freund and Frank J. Williams, Elementary Business Statistics: The Modern Approach(Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1965),p. 409.

๓. หาค่า u คือ จำนวนชุดของสัญลักษณ์ a หรือ b ที่อยู่ต่อเนื่องกัน

๔. ตั้งสมมติฐานแรก (Null hypothesis-H_0) ว่าการจัดข้อมูลเป็นไปแบบสุ่ม (random arrangement) โดยมีสมมติฐานรอง (Alternative hypothesis -H_1) ว่าข้อมูลที่จัดมามีแนวโน้ม

๕. กำหนดระดับนัยสำคัญ

๖. นำค่า u ที่หาได้มาเปรียบเทียบกับค่า u_α จากตารางที่ ๒.๑

๗. ถ้า u ที่หาได้ตามข้อ ๓ น้อยกว่าหรือเท่ากับค่า u_α ที่ได้จากตารางให้ปฏิเสธสมมติฐานแรก (H_0)นั่นคือยอมรับสมมติฐานรอง (H_1) สรุปว่าข้อมูลที่ศึกษามีแนวโน้มชัดแจ้ง ณ ระดับนัยสำคัญ α

ในกรณีที่ข้อมูลมีแนวโน้ม ค่า uหรือจำนวนชุดควรจะน้อย เนื่องจากข้อมูลมีการเรียงลำดับกันจากน้อยไปหามาก เป็นแนวโน้มแบบเพิ่มขึ้น (upward trend) หรือข้อมูลมีการเรียงลำดับกันจากมากไปหาน้อย เป็นแนวโน้มแบบลดลง (downward trend)แต่ถ้าค่า u หรือจำนวนชุดมาก แสดงว่าข้อมูลไม่มีการเรียงลำดับอย่างแน่นอนมีค่าขึ้นลงแบบสุ่ม(random) ที่ไม่อาจหาแนวโน้มได้

การทดสอบที่กล่าวมาข้างต้นนี้ เป็นการทดสอบอย่างคร่าว ๆ เท่านั้นเพื่อทำให้ทราบว่าข้อมูลมีแนวโน้มจริงหรือไม่ ดังนั้นแม้ว่าการทดสอบดังกล่าวจะทำให้เราได้ข้อสรุปที่ว่าข้อมูลนั้น ไม่มีแนวโน้มก็ตาม ผู้วิจัยก็จะดำเนินการหาโมเดลแนวโน้มของข้อมูลในแบบต่าง ๆ และทำการทดสอบค่าพารามิเตอร์โดยการวิเคราะห์ค่าแปรปรวน (Analysis of Variance) และหาค่าสัมประสิทธิ์แห่งการตัดสินใจ (Coefficient of Determination -R^2) เพื่อยืนยันผลการทดสอบอีกครั้งหนึ่งว่าข้อมูลดังกล่าวนั้นไม่มีแนวโน้มจริง ๆ

ในบางกรณี พบว่าเราอาจหาโมเดลของแนวโน้มที่เหมาะสมกับข้อมูลได้ทั้ง ๆ ที่ผลของการทดสอบเบื้องต้นดังที่กล่าวมาแล้ว สรุปว่าข้อมูลชุดนั้นไม่มีแนวโน้มทั้งนี้ก็เนื่องมาจากโมเดลดังกล่าวให้ค่าสัมประสิทธิ์แห่งการตัดสินใจค่อนข้างต่ำนั่นเอง

ตารางที่ ๒.๑ ค่าวิกฤต(critical value)ของข้อมูลที่อยู่สูงกว่าและต่ำกว่าค่ามัธยฐาน (u)[1]

จำนวน a = จำนวน b	$u_{.05}$	$u_{.01}$
5	3	2
6	3	2
7	4	3
8	5	4
9	6	4
10	6	5
11	7	6
12	8	6
13	9	7
14	10	8
15	11	9
16	11	10
17	12	10
18	13	11
19	14	12
20	15	13

---

[1] Freund and Williams, Elementary Business Statistics ....., Table VIII, p. 481.

โมเดลของแนวโน้มที่จะใช้ในการศึกษารายได้ที่ได้รับจากอุตสาหกรรมท่องเที่ยวและรายได้จากการส่งสินค้าออกที่สำคัญทั้ง ๔ ประเภทมีดังนี้ คือ

๒.๑.๑ <u>โมเดลเส้นตรง</u> (Linear Model)

<u>โมเดล</u> $Y = \alpha_0 + \alpha_1 X + \varepsilon_1$

เมื่อ X คือตัวแปรอิสระ (Independent Variable) ในที่นี้ได้แก่ระยะเวลา (ปี)

Y คือตัวแปรตาม (Dependent Variable) ขึ้นกับ X ในที่นี้ได้แก่ข้อมูลที่ต้องการศึกษา เช่นจำนวนนักท่องเที่ยว

α_0 คือค่าเฉลี่ย (General mean) ที่ไม่นับค่าเปลี่ยนแปลงอันเนื่องจากการเปลี่ยนแปลงค่า X แล้ว

α_1 คืออัตราการเปลี่ยนแปลงของค่า Y ต่อหนึ่งหน่วยของค่า X

α_0, α_1 คือตัวพารามิเตอร์ (Parameter) ของโมเดล

ε_1 คือค่าความคลาดเคลื่อนแบบสุ่ม (Random error) ซึ่งเนื่องมาจากความผิดพลาดในการวัด (Measurement error) และเนื่องมาจากตัวประกอบอื่น ๆ นอกจาก X

สมการแนวโน้มคือ $\hat{Y} = a_0 + a_1 X$

เมื่อ $\hat{Y}$ คือค่าประมาณของ Y

a_0, a_1 คือค่าประมาณของตัวพารามิเตอร์ α_0, α_1 ตามลำดับ

ค่า a_0, a_1 นี้ประมาณขึ้นได้โดยอาศัยวิธีการที่เรียกว่า Least Squares Method ซึ่งเป็นวิธีการที่ทำให้ผลรวมของผลต่างระหว่างข้อมูลที่ศึกษากับค่าของสมการแนวโน้มยกกำลังสองแล้ว ($\Sigma (Y - \hat{Y})^2$) ได้ค่าน้อยที่สุด

ด้วยวิธีการของ Least Squares สามารถเขียน Normal Equations เพื่อใช้คำนวณค่า a_0, a_1 ได้ดังนี้

$$n a_0 + a_1 \Sigma X = \Sigma Y$$

000295

$$a_0 \Sigma X + a_1 \Sigma X^2 = \Sigma XY$$

ในที่นี้ค่าของ X คือค่าที่ จึงสามารถกำหนดค่า X ใหม่ซึ่งทำให้ผลรวมของ X มีค่าเท่ากับ 0 เพื่อสะดวกในการคำนวณ ดังนั้น Normal Equations จะลดรูปลงเหลือเพียง

$$na_0 = \Sigma Y \qquad (2.1)$$

$$a_1 \Sigma X^2 = \Sigma XY \qquad (2.2)$$

นั่นคือ

$$a_0 = \frac{\Sigma Y}{n}$$

$$a_1 = \frac{\Sigma XY}{\Sigma X^2}$$

๒.๑.๒ <u>โมเดลเส้นโค้ง</u> (Nonlinear Model) มีหลายชนิดด้วยกัน ในที่นี้ขอยกมากล่าวเฉพาะโมเดลที่เป็นที่นิยมและเหมาะสมกับข้อมูลโดยทั่ว ๆ ไป เพียง ๒ ชนิด คือ

๒.๑.๒.๑ <u>สมการโพลิโนเมียล (Polynomial) กำลังตั้งแต่ ๒ ขึ้นไป</u> เป็นสมการที่รวมการเปลี่ยนค่าที่เนื่องจาก X^2, X^3,...

ก. <u>สมการกำลังสอง</u>

<u>โมเดล</u> $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_2$

<u>สมการแนวโน้ม</u> $\hat{Y} = b_0 + b_1 X + b_2 X^2$

ด้วยวิธีการของ Least Squares Method เช่นกันทำให้ได้ Normal Equations เป็น

$$nb_0 + b_1 \Sigma X + b_2 \Sigma X^2 = \Sigma Y$$

$$b_0 \Sigma X + b_1 \Sigma X^2 + b_2 \Sigma X^3 = \Sigma XY$$

$$b_0 \Sigma X^2 + b_1 \Sigma X^3 + b_2 \Sigma X^4 = \Sigma X^2 Y$$

เนื่องจากได้กำหนดค่า X ใหม่ทำให้ $\Sigma X = 0$, $\Sigma X^3 = 0$ ดังนั้น Normal Equations จะลดรูปลงเหลือเพียง

$$nb_0 + b_2 \Sigma X^2 = \Sigma Y \qquad (2.3)$$

$$b_1 \Sigma X^2 = \Sigma XY \qquad (2.4)$$

$$b_0 \Sigma X^2 + b_2 \Sigma X^4 = \Sigma X^2 Y \qquad (2.5)$$

จากสมการ Normal Equations ทั้ง ๓ สามารถคำนวณหาค่า b_0, b_1, b_2 ซึ่งเป็นค่าประมาณของ $\beta_0, \beta_1, \beta_2$ ตามลำดับได้ด้วยวิธีการแก้สมการทางพีชคณิต

ข. <u>สมการกำลังสาม</u>

<u>โมเดล</u> $Y = \gamma_0 + \gamma_1 X + \gamma_2 X^2 + \gamma_3 X^3 + \varepsilon_3$

<u>สมการแนวโน้ม</u> $\hat{Y} = c_0 + c_1 X + c_2 X^2 + c_3 X^3$

จะได้ Normal Equations ดังนี้

$$nc_0 + c_1 \Sigma X + c_2 \Sigma X^2 + c_3 \Sigma X^3 = \Sigma Y$$

$$c_0 \Sigma X + c_1 \Sigma X^2 + c_2 \Sigma X^3 + c_3 \Sigma X^4 = \Sigma XY$$

$$c_0 \Sigma X^2 + c_1 \Sigma X^3 + c_2 \Sigma X^4 + c_3 \Sigma X^5 = \Sigma X^2Y$$

$$c_0 \Sigma X^3 + c_1 \Sigma X^4 + c_2 \Sigma X^5 + c_3 \Sigma X^6 = \Sigma X^3Y$$

จาก Normal Equations ข้างบน เมื่อค่า $\Sigma X = 0, \Sigma X^3 = 0, \Sigma X^5 = 0$ จะทำให้สมการลดลงเหลือเพียง

$$nc_0 + c_2 \Sigma X^2 = \Sigma Y \quad (2.6)$$

$$c_1 \Sigma X^2 + c_3 \Sigma X^4 = \Sigma XY \quad (2.7)$$

$$c_0 \Sigma X^2 + c_2 \Sigma X^4 = \Sigma X^2Y \quad (2.8)$$

$$c_1 \Sigma X^4 + c_3 \Sigma X^6 = \Sigma X^3Y \quad (2.9)$$

จากสมการ Normal Equations ทั้ง ๔ สามารถคำนวณหาค่า c_0, c_1, c_2, c_3 ซึ่งเป็นค่าประมาณของ $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ ตามลำดับได้ด้วยวิธีการแก้สมการทางพีชคณิตเช่นกัน

สำหรับสมการโพลิโนเมียลกำลังสูงกว่านี้ขึ้นไป อาจสร้าง Normal Equations และคำนวณค่าประมาณในสมการแนวโน้มได้ในทำนองเดียวกัน

๒.๑.๒.๒ <u>สมการเอ็กซ์โปเนนเชียล</u> (Exponential)

<u>โมเดล</u> $Y = AB^X + \varepsilon$

<u>สมการแนวโน้ม</u> $\hat{Y} = ab^X$

หอสมุดกลาง

a, b เป็นค่าประมาณของตัวพารามิเตอร์ A, B ตามลำดับ

เนื่องจากสมการเป็นรูปของเอ็กซโพเนนเชียล ดังนั้นในการประมาณค่า A, B อาจทำให้ง่ายขึ้นโดยใช้การแปรรูป (Transformation) เข้าช่วยเพื่อให้สมการเดิมเปลี่ยนรูปเป็นสมการเส้นตรง

<u>โมเดล</u> $\log Y = \log A + X \log B + \log \varepsilon$

<u>สมการแนวโน้ม</u> $\widehat{\log Y} = \log a + X \log B$

จะได้ Normal Equations ดังนี้

$$n \log a + \log b \sum X = \sum \log Y \qquad (2.10)$$

$$\log a \sum X + \log b \sum X^2 = \sum X \log Y \qquad (2.11)$$

ด้วยวิธีการกำหนดค่า X ใหม่จึงทำให้ $\Sigma X = 0$ ฉะนั้นสมการ Normal Equations จะลดรูปลงเหลือเพียง

$$n \log a = \sum \log Y$$

$$\log b \quad X^2 = \sum X \log Y$$

และจะสามารถคำนวณค่า log a, log b ได้จากสูตร

$$\log a = \frac{\sum \log Y}{n}$$

$$\log b = \frac{\sum X \log Y}{\sum X^2}$$

ดังนั้น $a = \text{antilog}\left(\frac{\sum \log Y}{n}\right)$

$$b = \text{antilog}\left(\frac{\sum X \log Y}{\sum X^2}\right)$$

สมการแนวโน้มในรูปของเอ็กซโพเนนเชียล คือ

$$\hat{Y} = ab^X$$

๒.๒ การทดสอบโมเดล

การทดสอบโมเดลจะทำขึ้นภายหลังจากที่ได้นำข้อมูลมาทดลองสร้างกราฟคู่ทั้งในกระดาษที่มีสเกลแบบธรรมดาโดยมีเวลา (X) เป็นแกนนอนและข้อมูลที่ศึกษา (Y) เป็นแกนตั้งและในกระดาษที่มีสเกลแบบ Semi - log โดยมีเวลา(X) เป็นแกนนอน และ log ของข้อมูลที่ศึกษา (logY) เป็นแกนตั้ง

ในกรณีที่ข้อมูลให้แนวโน้มเป็นเส้นตรงในตาราง Semi - log ก็จะทำการทดสอบโมเดลของแนวโน้มรูปเอ็กซโพเนนเชียล (Exponential) โดยการวิเคราะห์ค่าแปรปรวน (Analysis of Variance) และใช้การทดสอบที่เรียกว่า F-test ในการพิจารณาว่าข้อมูลนั้นเหมาะสมกับโมเดลแนวโน้มรูปเอ็กซโพเนนเชียลหรือไม่

ในกรณีที่โมเดลของแนวโน้มไม่อยู่ในรูปเอ็กซโพเนนเชียลดังกล่าว ก็จะทำการพิจารณาทดสอบโมเดลที่เหมาะสมกับข้อมูลว่าควรจะเป็นโมเดลโพลิโนเมียลกำลังใด โดยการวิเคราะห์ค่าแปรปรวนและใช้ F-test ในการทดสอบว่าการเพิ่มกำลังของสมการโพลิโนเมียลมีความสำคัญพอหรือไม่ ทั้งนี้โดยอาศัยการพิจารณาค่าสัมประสิทธิ์แห่งการตัดสินใจ (Coefficient of Determination -R^2) และค่าความคลาดเคลื่อนมาตรฐาน(Standard error - s.e.) มาประกอบการพิจารณาโมเดลของสมการโพลิโนเมียลที่เหมาะสมกับข้อมูลด้วย

๒.๒.๑ การวิเคราะห์ค่าแปรปรวนเพื่อทดสอบโมเดลเอ็กซโพเนนเชียล

<u>โมเดล</u> $Y = AB^X + \varepsilon$

<u>สมการแนวโน้ม</u> $\hat{Y} = ab^X$

เนื่องจากสมการเป็นรูปของเอ็กซโพเนนเชียล จึงใช้การแปรรูป (transform) เข้าช่วยเพื่อให้สมการเปลี่ยนรูปเป็นสมการเส้นตรง

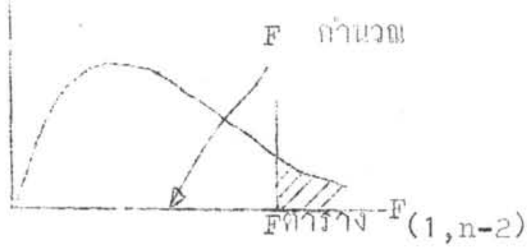
$$\widehat{\log Y} = \log a + X \log b$$

เมื่อสามารถคำนวณหาค่า log a, log b ในสมการแนวโน้มได้แล้ว ก็จะสร้างตารางวิเคราะห์ค่าแปรปรวน เพื่อทดสอบสมมติฐานเกี่ยวกับค่า log B ว่ามีค่าเท่ากับ ๐ หรือไม่

ตารางที่ ๒.๒ สูตรในการวิเคราะห์ค่าแปรปรวนเพื่อทดสอบสมมติฐาน H_0: log B = 0 ของโมเดลเอ็กซโพเนนเชียล

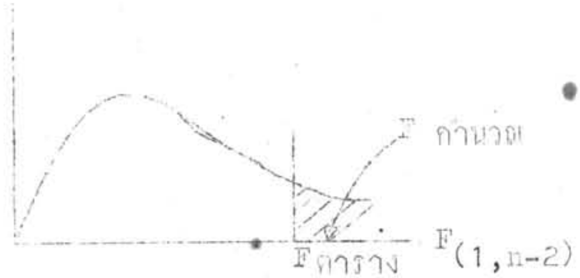
S.O.V.	d.f.	S.S.	M.S.	F
log A, log B	2	log a(Σlog Y)+log b(ΣXlogY) =(1)		
log A	1	log a(Σ log Y) =(2)		
log B/ log A	1	(1) - (2) =(3)	(3)	$F = \frac{(3)}{\frac{(5)}{(n-2)}}$
Residuals	n-2	(4) - (1) =(5)	$\frac{(5)}{(n-2)}$	
Total (uncorrected)	n	Σ (log Y)^2 = (4)		

ที่ระดับนัยสำคัญ α ถ้าค่า F ที่ได้จากการคำนวณในตารางวิเคราะห์ค่าแปรปรวนน้อยกว่าค่า F ที่ได้จากตารางสถิติเมื่อ degrees of freedom(d.f.) เป็น 1 และ n-2 หรือเขียนเป็น $F_{(1,n-2,1-\alpha)}$ แสดงว่าค่า F ที่ได้จากการคำนวณอยู่ในเกณฑ์ยอมรับสมมติฐาน



สรุปได้ว่า ณ ระดับนัยสำคัญ α มีเหตุผลเพียงพอที่จะยอมรับสมมติฐานว่า log B = 0 กล่าวคือ ข้อมูลดังกล่าวไม่ได้มีความสัมพันธ์แบบเอ็กซโพเนนเชียล

ถ้าค่า F ที่ได้จากการคำนวณในตารางวิเคราะห์ค่าแปรปรวนมากกว่าค่า F ที่ได้จากตารางสถิติเมื่อ degrees of freedom เป็น 1 และ n-2 แสดงว่าค่า F ที่ได้จากการคำนวณอยู่ในเกณฑ์ปฏิเสธสมมติฐาน แสดงว่า log B $\neq 0$



สรุปได้ว่า ณ ระดับนัยสำคัญ α เรามีเหตุผลเพียงพอที่จะปฏิเสธสมมติฐาน นั่นคือ ยอมรับว่าข้อมูลดังกล่าวมีความสัมพันธ์แบบเอ็กซโพเนนเชียล

๒.๒.๒ การวิเคราะห์ค่าแปรปรวนเพื่อทดสอบโมเดลโพลิโนเมียล[1]

การทดสอบโมเดลโพลิโนเมียลโดยอาศัยการวิเคราะห์ค่าแปรปรวนจะทำเป็นขั้น ๆ คือ จะทำการวิเคราะห์ค่าแปรปรวนในสมการโพลิโนเมียลกำลังหนึ่ง ซึ่งมีรูปโมเดลเป็น

$$Y = \alpha_0 + \alpha_1 X + \varepsilon_1$$

แล้วทดสอบสมมติฐานว่า $\alpha_1 = 0$ หรือไม่ ถ้ายอมรับสมมติฐานว่า $\alpha_1 = 0$ แสดงว่าข้อมูลชุดนี้ไม่มีแนวโน้ม หรือมีแนวโน้มเป็น $Y = \alpha_0 + \varepsilon$ แต่ถ้าปฏิเสธสมมติฐานคือได้ว่า $\alpha_1 \neq 0$ แสดงว่า α_1 มีความสำคัญพอที่จะอยู่ในโมเดลได้ เมื่อได้ว่า α_1 พอที่จะอยู่ในโมเดลได้แล้วก็จะต้องพิจารณาสมการโพลิโนเมียลที่มีกำลังสูงขึ้นไป คือสมการโพลิโนเมียลกำลังสองที่มีรูปโมเดลเป็น

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_2$$

แล้วทดสอบสมมติฐานว่า $\beta_2 = 0$ หรือไม่ ถ้ายอมรับสมมติฐานว่า $\beta_2 = 0$ ให้หยุดการพิจารณาสมการโพลิโนเมียลกำลังสูงถัดไป และยอมรับว่าโมเดลที่เหมาะสมกับข้อมูลคือสมการโพลิโนเมียลกำลังหนึ่งหรือสมการเส้นตรง แต่ถ้าปฏิเสธสมมติฐานคือได้ว่า $\beta_2 \neq 0$ แสดงว่า β_2 มีความสำคัญพอที่จะอยู่ในโมเดลได้ ให้ทำการพิจารณาสมการโพลิโนเมียลกำลังสามที่มีรูปโมเดลเป็น

$$Y = \gamma_0 + \gamma_1 X + \gamma_2 X^2 + \gamma_3 X^3 + \varepsilon_3$$

แล้วทดสอบสมมติฐานว่า $\gamma_3 = 0$ ถ้าได้ว่า $\gamma_3 = 0$ ให้หยุดการพิจารณาสมการโพลิโนเมียลกำลังสูงถัดไป และยอมรับเอาสมการโพลิโนเมียลกำลังสองเป็นสมการที่เหมาะสมกับข้อมูล แต่ถ้าปฏิเสธสมมติฐานได้ว่า $\gamma_3 \neq 0$ ก็จำเป็นต้องพิจารณาสมการโพลิโนเมียลที่มีกำลังสูงถัดไป วิธีการพิจารณานั้นก็เป็นไปในทำนองเดียวกันกับที่กล่าวมาแล้วในสมการโพลิโนเมียลกำลังต่ำกว่า

---

[1] Franklin A. Graybill, An Introduction to Linear Statistical Models (New York: McGraw-Hill Book Co., Inc., 1961), pp.166-171.

นอกเหนือจากการวิเคราะห์ค่าแปรปรวนดังกล่าว ผู้วิจัยจะทำการพิจารณาค่าสัมประสิทธิ์แห่งการตัดสินใจ (R^2) และค่าความคลาดเคลื่อนมาตรฐาน(s.e.) ก่อนการตัดสินใจยอมรับโมเดลที่เหมาะสม

ขั้นตอนของการทดสอบโมเดลมีดังนี้ คือ

<u>ขั้นที่ ๑</u> การพิจารณาสมการโพลิโนเมียลกำลังหนึ่งและทดสอบสมมติฐาน $H_0: \alpha_1 = 0$

ในโมเดล $Y = \alpha_0 + \alpha_1 X + \varepsilon_1$

ตารางที่ ๒.๓ สูตรในการวิเคราะห์ค่าแปรปรวนเพื่อทดสอบสมมติฐาน $H_0: \alpha_1 = 0$
ในสมการโพลิโนเมียลกำลังหนึ่ง

S.O.V.	d.f.	S.S.	M.S.	F
α_0, α_1	2	$\hat{\alpha}_0 \Sigma Y + \hat{\alpha}_1 \Sigma XY$ =(1)		
α_0	1	$\hat{\alpha}_0 \Sigma Y$ =(2)		
α_1 / α_0	1	(1)-(2) =(3)	(3)	$F = \frac{(3)}{(5)/(n-2)}$
Residuals	n-2	(4)-(1) =(5)	(5)/(n-2)	
Total(uncorrected)	n	ΣY^2 =(4)		

<u>Null Hypothesis</u>

$$H_0:\ \alpha_1 = 0 \quad \text{หรือ} \quad Y = \alpha_0 + \varepsilon$$

<u>Alternative Hypothesis</u>

$$H_0:\ \alpha_1 \neq 0 \quad \text{หรือ} \quad Y = \alpha_0 + \alpha_1 X + \varepsilon_1$$

ที่ระดับสำคัญ α ถ้าค่า F ที่ได้จากการคำนวณในตารางวิเคราะห์ค่าแปรปรวน น้อยกว่าค่า F ที่ได้จากตารางสถิติเมื่อ degrees of freedom เป็น 1 และ n-2 แสดงว่าค่า F ที่ได้จากการคำนวณอยู่ในเกณฑ์ยอมรับสมมติฐาน

สรุปได้ว่า ณ ระดับนัยสำคัญ α เรามีเหตุผลเพียงพอที่จะยอมรับว่า $\alpha_1 = 0$ กล่าวคือ α_1 ไม่มีความสำคัญพอที่จะอยู่ในโมเดล ดังนั้น โมเดลที่เหมาะสมกับข้อมูลชุดนี้ คือ $Y = \alpha_0 + \varepsilon$

ที่ระดับนัยสำคัญ α ถ้าค่า F ที่ได้จากการคำนวณในตารางวิเคราะห์ค่าแปรปรวนมากกว่าค่า $F_{(1,n-2)}$ ที่ได้จากตารางสถิติ แสดงว่าค่า F ที่ได้จากการคำนวณอยู่ในเกณฑ์ปฏิเสธสมมติฐาน

สรุปได้ว่า ณ ระดับนัยสำคัญ α เรามีเหตุผลเพียงพอที่จะปฏิเสธสมมติฐาน นั่นก็คือ $\alpha_1 \neq 0$ กล่าวคือ α_1 มีความสำคัญพอที่จะอยู่ในโมเดลได้ ให้ทำขั้นที่ ๒ ต่อไป

<u>ขั้นที่ ๒</u> การพิจารณาสมการโพลิโนเมียลกำลังสองและทดสอบสมมติฐาน $H_0: \beta_2 = 0$ ในโมเดล $Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_2$

<u>ตารางที่ ๒.๔</u> สูตรในการวิเคราะห์ค่าแปรปรวนเพื่อทดสอบสมมติฐาน $H_0: \beta_2 = 0$ ในสมการโพลิโนเมียลกำลังสอง

S.O.V.	d.f.	S.S.	M.S.	F
$\beta_0, \beta_1, \beta_2$	3	$\hat{\beta}_0 \Sigma Y + \hat{\beta}_1 \Sigma XY + \hat{\beta}_2 \Sigma X^2 Y = (1)$		
β_0, β_1	2	$\hat{\alpha}_0 \Sigma Y + \hat{\alpha}_1 \Sigma XY = (2)$		
$\beta_2 / \beta_0, \beta_1$	1	$(1)-(2) = (3)$	(3)	$F = \dfrac{(3)}{(5)/(n-3)}$
Residuals	n-3	$(4)-(1) = (5)$	(5)/(n-3)	
Total (uncorrected)	n	$\Sigma Y^2 = (4)$		

<u>Null Hypothesis</u>

$$H_0: \beta_2 = 0 \qquad Y = \alpha_0 + \alpha_1 X + \varepsilon_1$$

<u>Alternative Hypothesis</u>

$$H_0: \beta_2 \neq 0 \qquad Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_2$$

ที่ระดับนัยสำคัญ α ถ้าค่า F ที่คำนวณได้น้อยกว่าค่า $F_{(1,n-3)}$ จากตารางสถิติอยู่ในเกณฑ์ยอมรับสมมติฐานว่า $\beta_2 = 0$ เราก็จะหยุดและยอมรับว่า $Y = \alpha_0 + \alpha_1 X + \varepsilon_1$ เป็นโมเดลที่เหมาะสมกับข้อมูลที่สุด แต่ถ้าค่า F ที่คำนวณได้อยู่ในเกณฑ์ปฏิเสธสมมติฐานแสดงว่า β_2 สำคัญพอที่จะอยู่ในโมเดล ให้ทำขั้นที่ ๓ ต่อไป

ขั้นที่ ๓ การพิจารณาสมการโพลิโนเมียลกำลังสามและทดสอบสมมุติฐาน $H_0: \gamma_3 = 0$ ในโมเดล

ตารางที่ ๒.๕ สูตรในการวิเคราะห์ค่าแปรปรวนเพื่อทดสอบสมมุติฐาน $H_0: \gamma_3 = 0$ ในสมการโพลิโนเมียลกำลังสาม

S.O.V.	d.f.	S.S.	M.S.	F
$\gamma_0, \gamma_1, \gamma_2, \gamma_3$	4	$\hat{\gamma}_0 \Sigma Y + \hat{\gamma}_1 \Sigma XY + \hat{\gamma}_2 \Sigma X^2Y + \hat{\gamma}_3 \Sigma X^3Y = (1)$		
$\gamma_0, \gamma_1, \gamma_2$	3	$\hat{\beta}_0 \Sigma Y + \hat{\beta}_1 \Sigma XY + \hat{\beta}_2 \Sigma X^2Y \quad = (2)$		
$\gamma_3 / \gamma_0, \gamma_1, \gamma_2$	1	$(1) - (2) \quad = (3)$	(3)	$F = \frac{(3)}{(5)/(n-4)}$
Residuals	n-4	$(4) - (1) \quad = (5)$	$\frac{(5)}{(n-4)}$	
Total (uncorrected)	n	$\Sigma Y^2 \quad = (4)$		

Null Hypothesis

$$H_0: \gamma_3 = 0 \text{ หรือ } Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \varepsilon_2$$

Alternative Hypothesis

$$H_1: \gamma_3 \neq 0 \text{ หรือ } Y = \gamma_0 + \gamma_1 X + \gamma_2 X^2 + \gamma_3 X^3 + \varepsilon_3$$

ณ ระดับนัยสำคัญ α ถ้าค่า F ที่คำนวณได้น้อยกว่าค่า $F_{(1,n-4)}$ จากตารางสถิติ อยู่ในเกณฑ์ยอมรับสมมุติฐานว่า $\gamma_3 = 0$ เราก็จะหยุดและยอมรับว่าโมเดลโพลิโนเมียลกำลังสองเป็นโมเดลที่เหมาะสมกับข้อมูลที่สุด แต่ถ้าค่า F ที่คำนวณได้มากกว่าค่า $F_{(1,n-4)}$ จากตารางสถิติ อยู่ในเกณฑ์ปฏิเสธสมมุติฐาน แสดงว่า γ_3 มีความสำคัญพอที่จะอยู่ในโมเดลได้ ต้องทำการพิจารณาสมการโพลิโนเมียลกำลังสี่ในทำนองเดียวกัน แต่ในที่นี้ เราจะหยุดที่การพิจารณาสมการโพลิโนเมียลกำลังสาม และใช้การพิจารณาค่าสัมประสิทธิ์แห่งการตัดสินใจ คือ R^2 และค่าความคลาดเคลื่อนมาตรฐาน (s.e.) ช่วยในการทดสอบโมเดลด้วย

๒.๒.๓ การพิจารณาค่าสัมประสิทธิ์แห่งการตัดสินใจ (Coefficient of Determination – R^2) และค่าความคลาดเคลื่อนมาตรฐาน (Standard Error – s.e.)

ค่าสัมประสิทธิ์แห่งการตัดสินใจ (R^2) เป็นค่าที่ใช้วัดความสัมพันธ์ระหว่าง Y กับฟังก์ชันของ X ว่ามีความสัมพันธ์กันมากน้อยเพียงใด ในโมเดลโพลีโนเมียลกำลังต่าง ๆ สามารถคำนวณหาค่า R^2 ได้จากค่าในตารางวิเคราะห์ค่าแปรปรวน ดังนี้

สมการโพลีโนเมียลกำลังหนึ่ง

$$R^2 = \frac{\text{Sum of Squares due to } \alpha_1 / \alpha_0}{\text{Total Sum of Squares(corrected)}}$$

$$= \frac{(\alpha_0, \alpha_1)\ \text{S.S.} - (\alpha_0)\ \text{S.S.}}{\Sigma\, y^2}$$

สมการโพลีโนเมียลกำลังสอง

$$R^2 = \frac{\text{Sum of Squares due to } \beta_1, \beta_2 / \beta_0}{\text{Total Sum of Squares(corrected)}}$$

$$= \frac{(\beta_0, \beta_1, \beta_2)\ \text{S.S.} - (\alpha_0)\ \text{S.S.}}{\Sigma\, y^2}$$

สมการโพลีโนเมียลกำลังสาม

$$R^2 = \frac{\text{Sum of Squares due to } \gamma_1, \gamma_2, \gamma_3 / \gamma_0}{\text{Total Sum of Squares (corrected)}}$$

$$= \frac{(\gamma_0, \gamma_1, \gamma_2, \gamma_3)\ \text{S.S.} - (\alpha_0)\ \text{S.S.}}{\Sigma\, y^2}$$

โดยที่ $\qquad \Sigma y^2 = \Sigma (Y - \bar{Y})^2$

$$= \Sigma Y^2 - (\Sigma Y)^2 / n$$

การพิจารณาค่า R^2 จะช่วยยืนยันการทดสอบโมเดลโดยการวิเคราะห์ค่าแปรปรวนว่าการเพิ่มกำลังของสมการโพลิโนเมียลมีผลทำให้ความสัมพันธ์ระหว่าง Y กับ f(X) สูงขึ้นมากเพียงพอหรือไม่ ถ้าการเพิ่มกำลังของสมการโพลิโนเมียลมีผลทำให้ค่า R^2 เพิ่มขึ้นเพียงเล็กน้อย ก็ไม่มีความจำเป็นที่จะต้องเพิ่มกำลังของสมการโพลิโนเมียล แต่ถ้าค่า R^2 เพิ่มสูงกว่าเกินมากก็สมควรที่จะเพิ่มกำลังของสมการโพลิโนเมียลให้สูงขึ้น ในที่นี้จะใช้เกณฑ์การเพิ่มขึ้นของ R^2 เท่ากับ ๑๐% ในการพิจารณา

อย่างไรก็ตาม ต้องพิจารณาค่าความคลาดเคลื่อนมาตรฐาน (s.e.) ประกอบด้วยค่า s.e. จะเป็นค่าวัดความคลาดเคลื่อนของค่าประมาณจากข้อมูลจริง ซึ่งจะหาได้จากตารางวิเคราะห์ค่าแปรปรวนเช่นกัน คำนวณหา s.e. ได้จาก

$$s.e. = \sqrt{\text{Residual Mean Square}} \; .$$

เมื่อค่า R^2 เพิ่มขึ้น จะทำให้ Residual Sum of Square ต่ำลง แต่ค่า Residual Mean Square อาจไม่ต่ำลงก็ได้ ดังนั้นเราจึงจะเลือกใช้โมเดลที่ให้ค่า R^2 สูง และค่า s.e. ต่ำ