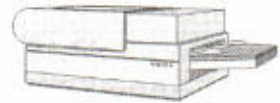




การนำโปรแกรม MATHCAD มาใช้ในการวิเคราะห์การถดถอย



บันทึก น. สมอ*

ในการวิเคราะห์การถดถอย จะใช้โปรแกรมคอมพิวเตอร์ SPSS มาช่วยในการวิเคราะห์ แต่เราสามารถนำโปรแกรม MATHCAD มาช่วยในการวิเคราะห์ได้ ซึ่งจะช่วยในการหาค่าสถิติในแต่ละขั้นตอนได้อย่างละเอียด เช่นหาค่า $X^T X$, $(X^T X)^{-1}$

การวิเคราะห์การถดถอยเป็นเครื่องมือทางสถิติที่นำมาใช้เพื่อศึกษาและสร้างรูปแบบความสัมพันธ์ระหว่างตัวแปร 2 ตัว หรือมากกว่า 2 ตัว เมื่อทราบลักษณะความสัมพันธ์ระหว่างตัวแปรดังกล่าวแล้ว ก็จะสามารถพยากรณ์ค่าของตัวแปรหนึ่งจากตัวแปรอีกกลุ่มหนึ่งได้ I กรณีศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระ 1 ตัว (X) กับตัวแปรตามหรือตัวแปรที่เป็นผลตอบสนอง (Y) ที่มีแนวโน้มเป็นเส้นตรง เรียกว่า simple linear regression

ดังนั้น รูปแบบการถดถอยคือ

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

ซึ่ง β_0 , β_1 เป็นค่าพารามิเตอร์ และ ε เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยสุ่ม
สูตรต่าง ๆ ในการวิเคราะห์การถดถอยอย่างง่ายเชิงเส้นตรง

- หาค่าสถิติ

ค่าเฉลี่ย ($\bar{X}$) :

$$\bar{X} = \frac{\sum X_i}{n}$$

ส่วนเบี่ยงเบนมาตรฐานของ X :

$$s_x = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

สัมประสิทธิ์สหสัมพันธ์ (r):

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}$$

สัมประสิทธิ์การถดถอย :

$$a = \bar{Y} - b\bar{X}, \quad b = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

- การทดสอบค่าพารามิเตอร์ β_1 หาค่า F

$$\text{SS Total} \quad \text{SST} = \sum (Y_i - \bar{Y})^2$$

$$\text{SS Regression} \quad \text{SSR} = b^2 \sum (X_i - \bar{X})^2$$

$$\text{SS Error} \quad \text{SSE} = \text{SST} - \text{SSR}$$

$$\text{MS Regression} \quad \text{MSR} = \text{SSR} / 1$$

$$\text{MS Error} \quad \text{MSE} = \text{SSE} / (n-2)$$

$$\text{F-test} \quad F = \text{MSR} / \text{MSE}$$

$$\text{ถ้า } F > F_{1, n-2, \infty} \quad \text{แสดงว่า } \beta_1 \neq 0$$

- การทดสอบค่าพารามิเตอร์ β_1 หาค่า t

ค่าประมาณความแปรปรวนของ b :

$$s_b^2 = \text{MSE} / \sum (X_i - \bar{X})^2$$

$$t = b / s_b$$

$$\text{ถ้า } |t| > t_{n-2} \quad \text{แสดงว่า } \beta_1 \neq 0$$

ตัวอย่าง ต้องการหาความสัมพันธ์ของ X กับ Y โดยวิเคราะห์การถดถอย

(ตัวแปรอิสระ)	X	6	4	3	5	4	2
(ตัวแปรตาม) Y	1	2	3	4	5	6	

ขั้นตอนในการวิเคราะห์การถดถอย

1. กำหนดเมตริกซ์ Y

- พิมพ์ Y คลิกไปที่เครื่องหมาย : = จะได้ y : =

- สร้างเมตริกซ์ เลือกจากเมนู MATH คลิกที่ MATRICES จะแสดงไดอะล็อก BOX

-	ท	E	MATRICES.....
ROWS :		COLUMNS :	
CREATE	INSERT	DELETE	

พิมพ์ ROWS : 6

COLUMNS : 1

แล้วใช้เมาส์คลิกไปที่ CREATE จะสร้าง MATRICES ขนาด 6 x 1

- พิมพ์ค่าของ Y ลงใน Placeholder แล้วกด Tab เพื่อเลื่อนไปยังตำแหน่งถัดไปจนครบทุกตำแหน่งในเมตริกซ์

จะได้

$$y := \begin{bmatrix} 6 \\ 4 \\ 3 \\ 5 \\ 4 \\ 2 \end{bmatrix}$$

- สร้างกรอบเลือก ครอบคลุมทั้งเมตริกซ์ Y

2. สร้างเมตริกซ์ X โดยวิธีการเหมือนข้อ 1

จะได้

$$X := \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$$

3. หาค่าสถิติ โดยใช้ฟังก์ชันของโปรแกรม MATHCAD

หาค่าเฉลี่ย X พิมพ์ค่า $\bar{X} := \text{MEAN}(X)$ จะได้ $\bar{X} = 3.5$

หาค่าเฉลี่ย Y พิมพ์ค่า $\bar{Y} := \text{MEAN}(Y)$ จะได้ $\bar{Y} = 4$

หาค่าส่วนเบี่ยงเบนมาตรฐานของ X พิมพ์ค่า $s_x := \text{STDEV}(X)$ จะได้ $s_x = 1.708$

Y $s_y := \text{STDEV}(Y)$ จะได้ $s_y = 1.291$

หาค่าสัมประสิทธิ์สหสัมพันธ์ (r) พิมพ์ค่า $r := \text{CORR}(X,Y)$ จะได้ $r = -0.68$

หาค่าสัมประสิทธิ์การถดถอย a พิมพ์ค่า $a := \text{INTERCEPT}(X,Y)$ จะได้ $a = 5.8$

b พิมพ์ค่า $b := \text{SLOPE}(X,Y)$ จะได้ $b = -.514$

หาค่าสัมประสิทธิ์กำหนด r^2 พิมพ์ค่า $rr := r [\text{shift}]^2$ จะได้ $rr = 0.463$

4. การทดสอบค่าพารามิเตอร์ β , ค่า F

จำนวนตัวอย่าง 6 ชุด

หาค่า SS Total	พิมพ์ค่า	SST : = 6* sy[shift]^2 จะได้ SST = 10
หาค่า SS Regresion	พิมพ์ค่า	SSR := 6* (b* sx) [shift]^2 จะได้ SSR = 4.629
หาค่า SS Error	พิมพ์ค่า	SSE : = SST-SSR จะได้ SSE = 5.371
หาค่า MS Error	พิมพ์ค่า	MSE : = SSE/4 จะได้ MSE = 1.343
หาค่า MS Regresion	พิมพ์ค่า	MSR : = SSR จะได้ MSR = 4.629
หาค่า F - test	พิมพ์ค่า	F : = MSR/MSE จะได้ F = 3.447
หาค่า standard error of regression พิมพ์ค่า STDERROR : = MSE คลิกลูกศร $\sqrt{\quad}$		
จะได้ STDERROR = 1.159		

5. การทดสอบค่าพารามิเตอร์ β , ค่า t

หาค่า s(b) พิมพ์ค่า sb : = MSE/ (6*sx[shift]^2) นำกรอบเลือกแล้วคลิกลูกศร $\sqrt{\quad}$ จะได้ sb = 0.277

หาค่า t พิมพ์ค่า t : = b/sb จะได้ t = -1.857

II กรณีศึกษา ตัวแปรอิสระหลายตัว ($X_1, X_2, \dots, X_k$) กับตัวแปรตาม (Y) ที่มีแนวโน้มเป็นเส้นตรง เรียกว่า Multiple linear regression

รูปแบบการถดถอย คือ

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$$

ซึ่ง $\beta_0, \beta_1, \dots, \beta_k$ เป็นค่าพารามิเตอร์ และ เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยสุ่ม
สูตรต่าง ๆ ในการวิเคราะห์การถดถอยพหุคูณเชิงเส้นตรง

- หาค่าสัมประสิทธิ์การถดถอย :

$$b = (X^T X)^{-1} (X^T Y)$$

- การทดสอบค่าพารามิเตอร์ β ค่า F

$$SS \text{ Total} \quad SST = \sum (Y_i - \bar{Y})^2$$

$$SS \text{ Regression} \quad SSR = b^T X^T Y - n \bar{Y}^2$$

SS Error

$$SSE = SST - SSR$$

MS Regression

$$MSR = SSR / k$$

MS Error

$$MSE = SSE / [n - (k + 1)]$$

$$F = MSR / MSE$$

ถ้า $F > F_{k, n-(k+1), \alpha}$ แสดงว่า β_i อย่างน้อย 1 ตัวไม่เท่ากับ 0 ($i = 1, 2, \dots, k$)

ตัวอย่าง ต้องการทราบความสัมพันธ์ระหว่าง X_1, X_2 (ตัวแปรอิสระ) กับ Y (ตัวแปรตาม) โดยวิเคราะห์การถดถอย

X_1	3	4	4	6	7	7	8	9
X_2	4	6	8	9	11	12	14	16
Y	6	8	9	10	13	15	17	18

ขั้นตอนการวิเคราะห์การถดถอย

1. สร้าง เมตริกซ์ X, Y (วิธีการทำเหมือนข้อ 1 ในการวิเคราะห์การถดถอยอย่างง่าย)

$$\text{จะได้ } X := \begin{bmatrix} 1 & 3 & 4 \\ 1 & 4 & 6 \\ 1 & 4 & 8 \\ 1 & 6 & 9 \\ 1 & 7 & 11 \\ 1 & 7 & 12 \\ 1 & 8 & 14 \\ 1 & 9 & 16 \end{bmatrix} \quad Y := \begin{bmatrix} 6 \\ 8 \\ 9 \\ 10 \\ 13 \\ 15 \\ 17 \\ 18 \end{bmatrix}$$

2. หาค่า transpose ของ matrix X และ matrix Y

หา x^T - สร้างกรอบเลือก กรอบคลุม เมตริกซ์ X

- เลือกเมนู SYMBOLIC โหลด SYMBOLIC PROCESSOR
- เลือก TRANSPOSE MATRIX

จะได้

$$x^T := \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 3 & 4 & 4 & 6 & 7 & 7 & 8 & 9 \\ 4 & 6 & 8 & 9 & 11 & 12 & 14 & 16 \end{bmatrix}$$

หา Y^T ทำนองเดียวกับ การหาค่า X^T

จะได้

$$y_t := [6 \ 8 \ 9 \ 10 \ 13 \ 15 \ 17 \ 18]$$

3. หาค่า $X^T X$, $X^T Y$

พิมพ์ $xtx := xt * x$

จะได้ $xtx =$

$$\begin{bmatrix} 8 & 48 & 80 \\ 48 & 320 & 539 \\ 80 & 539 & 914 \end{bmatrix}$$

พิมพ์ $xyt := xt * y$

จะได้ $xyt =$

$$\begin{bmatrix} 96 \\ 640 \\ 1.083 \times 10^3 \end{bmatrix}$$

4. หา INVERSE MATRIX คือ $(X^T X)^{-1}$

พิมพ์ $ixtx := xtx[\text{shift}]^{-1}$

$$\text{จะได้ } ixtx = \begin{bmatrix} 1.466 & -0.563 & 0.204 \\ -0.563 & 0.683 & -0.353 \\ 0.204 & -0.353 & 0.192 \end{bmatrix}$$

5. หาค่าสัมประสิทธิ์การถดถอย (b)

พิมพ์ $b := ixtx * xty$

จะได้ $b =$

$$\begin{bmatrix} 1.018 \\ 0.234 \\ 0.958 \end{bmatrix}$$

6. การทดสอบค่าพารามิเตอร์ β หาค่า F (เลือกตัวอย่าง 8 ชุด)

หาค่า b^T

จะได้ $bt := [1.018 \ 0.234 \ 0.958]$

หาค่า SS Total

พิมพ์ $SST := y_t * y - 8 * \text{MEAN}(y)$ ทำกรบนเลือกที่
MEAN (y) แล้วพิมพ์ $[\text{shift}]^2$ จะได้ $SST = 136$

หาค่า SS Regression

พิมพ์ $SSR := bt * xty - 8 * \text{mean}(y) [\text{shift}]^2$

จะได้ $SSR = 132.790$

หาค่า SS Error

พิมพ์ $SSE := SST - SSR$ จะได้ $SSE = 3.210$

หาค่า MS Error	พิมพ์ MSE = SSE/5	จะได้ MSE = .642
หาค่า MS Regression	พิมพ์ MSR = SSR/2	จะได้ MSR = 66.395
หาค่า F	พิมพ์ F : = MSR/MSE	จะได้ F = 103.433

หมายเหตุ กรอบเลือก มีลักษณะเหมือนกล่องที่มีสี่เหลี่ยม ขนาดขึ้นอยู่กับนิพจน์ที่กรอบเลือกคลุมอยู่

การวิเคราะห์การถดถอยโดยใช้โปรแกรม MATHCAD สามารถทำได้หลายวิธี ขึ้นอยู่กับสูตรที่ใช้การวิเคราะห์ เนื้อหาข้างต้นก็เป็นวิธีหนึ่งในการวิเคราะห์การถดถอย นอกจากนี้ โปรแกรม MATHCAD สามารถหาค่าต่าง ๆ ในวิชาสถิติได้อีก

เอกสารอ้างอิง

1. เปรมใจ ตีรสรานวัฒนา, การวิเคราะห์การถดถอย, ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2531
2. วิทยา สงวนวรรณ, การใช้งาน MATHCAD เวอร์ชัน 4.0 สำหรับวินโดวส์, บริษัทซีเอ็ดยูเคชั่น จำกัด (มหาชน), 2537