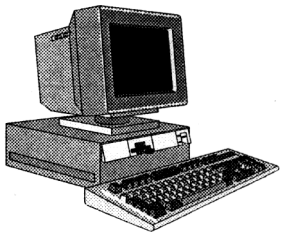
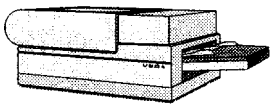


การนำโปรแกรม MATHCAD มาใช้ในการวิเคราะห์การถดถอย



นันทนา ณ ระนอง*

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

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 ($\bar{X}$) :

$$\bar{X} = X_i \mid n$$

ส่วนเบี่ยงเบนมาตรฐานของ X :

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

สัมประสิทธิ์สหสัมพันธ์ ( ):

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

สัมประสิทธิ์การถดถอย :

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

ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ ภาคใต้

การทดสอบค่าพารามิเตอร์ β_1 หาค่า F

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

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

SS Error $SSE = SST - SSR$

MS Regression $MSR = SSR/1$

MS Error $MSE = SSE / (n-2)$

F-test $F = MSR/MSE$

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

การทดสอบค่าพารามิเตอร์ β_1 หาค่า t

ค่าประมาณความแปรปรวนของ b

$$s^2_b = MSE/\Sigma(X_i - \bar{X})^2$$

$$t = b/s_b$$

ถ้า $|t| > t_{\frac{\alpha}{2}}$ แสดงว่า $\beta_1 \neq 0$

<u>ตัวอย่าง</u> ต้องการทราบความสัมพันธ์ของ X กับ Y โดยวิเคราะห์การถดถอย

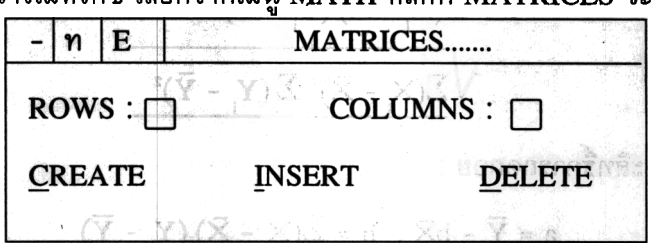
(ตัวแปรอิสระ) X	6	4	3	5	4	2
(ตัวแปรตาม) Y	2	3	4	5	6	

<u>ขั้นตอนในการวิเคราะห์การถดถอย</u>

1 *กำหนดเมตริกซ์ Y*

- พิมพ์ Y คลิกไปที่เครื่องหมาย = จะได้ y = □

- สร้างเมตริกซ์ เลือกจากเมนู MATH คลิกที่ MATRICES จะแสดงไดอะล็อก BOX



พิมพ์ ROWS : 6 COLUMNS : 1

แล้วใช้เมาส์คลิกไปที่ <u>C</u>REATE 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 XBAR : = MEAN (X) จะได้ XBAR = 3.5

หาค่าเฉลี่ย Y พิมพ์ค่า YBAR : = MEAN (Y) จะได้ YBAR = 4

หาค่าส่วนเบี่ยงเบนมาตรฐานของ X พิมพ์ค่า sx : = STDEV (X) จะได้ sx = 1.708

Y sy : = STDEV (Y) จะได้ sy = 1.291

หาค่าสัมประสิทธิ์สหสมพันธ์ (r) พิมพ์ค่า r : = CORR (X,Y) จะได้ r = -0.68

หาค่าสัมประสิทธิ์การถดถอย a พิมพ์ค่า a = INTERCEPT (X,Y) จะได้ a = 5.8

b พิมพ์ค่า b = SLOPE (X,Y) จะได้ b = -.514

หาค่าสัมประสิทธิ์กำหนด r^2 พิมพ์ค่า rr = r [shift]^ 2 จะได้ rr = 0.463

4. *การทดสอบค่าพารามิเตอร์ β_1 หาค่า 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 = .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 คลิกที่ $\sqrt{}$

จะได้ STDERROR = 1.159

5. *การทดสอบค่าพารามิเตอร์ β_1 หาค่า t*

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

หาค่า tพิมพ์ค่า t b/sb จะได้ t = -1.857

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

รูปแบบการถดถอย คือ

$$Y = \beta_0 + \beta_1 X + \beta_2 X_2 + \ldots\ldots + \beta_k X_k + \varepsilon$$

ซึ่ง $\beta_0, \beta_1, \ldots\ldots, \beta_k$ เป็นค่าพารามิเตอร์ และ เป็นความคลาดเคลื่อนที่เกิดขึ้นโดยสุ่ม

สูตรต่าง ๆ ในการวิเคราะห์การถดถอยพหุคูณเชิงเส้นตรง

- หาค่าสัมประสิทธิ์การถดถอย

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

- การทดสอบค่าพารามิเตอร์ β หาค่า F

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

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

SS Error $SSE = SST - SSR$

MS Regression $MSR = SSR / k$

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

$F = MSR/MSE$

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

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)

จะได้ $X :=$

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

2. หาค่า transpose ของ matrix X และ matrix Y

หา x^T - สร้างกรอบเลือก ครอบคลุม เมตริกซ์ X

- เลือกเมนู SYMBOLIC โหลด SYMBOLIC PROCESSOR
- เลือก TRANSPOSE MATRIX

จะได้

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

หา Y^T ทำนองเดียวกับ การหาค่า X^T

จะได้

yt = [6 8 9 10 13 15 7 18]

3. หาค่า X^TX , X^TY

พิมพ์ xtx : = xt * x จะได้ xtx = $\begin{bmatrix} 8 & 48 & 80 \\ 48 & 320 & 539 \\ 80 & 539 & 914 \end{bmatrix}$

พิมพ์ xty : = xt * y จะได้ xty = $\begin{bmatrix} 96 \\ 640 \\ 1.083 \times 10^3 \end{bmatrix}$

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

พิมพ์ ixtx : = xtx[shift]^-1

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: = yt * y - 8 * MEAN (y) ทำกรอบเลือกที่ MEAN (y) แล้วพิมพ์ [shift]^2 จะได้ SST = 136

หาค่า SS Regression พิมพ์ SSR: = bt * xty - 8 * mean (y) [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 สามารถหาได้หลายวิธี ขึ้นอยู่กับสูตรที่ใช้การวิเคราะห์ เนื้อหาข้างต้นก็เป็นวิธีหนึ่งในการวิเคราะห์การถดถอย นอกจากนี้ โปรแกรม MATHCAD สามารถหาค่าต่าง ๆ ในวิชาสถิติได้อีก

เอกสารอ้างอิง

เปรมใจ ตรีสรานุวัฒนา, การวิเคราะห์การถดถอย. ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์, 2531

2. วิทยา สงวนวรรณ, การใช้งาน MATHCAD เวอร์ชั่น 4.0 สำหรับวินโดวส์, บริษัทซีเอ็ดยูเคชั่น จำกัด (มหาชน), 2537