บทที่ 4

การคำนวณโดยใช้ทฤษฎี Multiple Linear Regression

4.1 การคำนวณของเครื่องคอมพิวเตอร์

ในการหาความสัมพันธ์ของตัวแปรต่าง ๆ โดยใช้ทฤษฎี MULTIPLE LINEAR REGRESSION มีวิธีคำนวณที่นิยมใช้กัน 2 วิธีคือ

วิธีที่ 1 เมื่อเครื่องคอมพิวเตอร์รับข้อมูลต่าง ๆ เรียบร้อยแล้ว ขั้นแรกจะคำนวณหาความสัมพันธ์ของตัวแปรอิสระ (X) หลาย ๆ ตัว กับตัวแปรตาม (Y) โดยคำนวณหาความสัมพันธ์ระหว่าง Y กับ X ตัวที่มีความสำคัญมากที่สุด แล้วจึงหาความสัมพันธ์ของ Y กับ X ตัวอื่น ที่มีความสำคัญรองลงมา เพิ่มเข้าไปในสมการทีละตัวจนครบ k ตัว

ดังนั้นสมการที่แสดงความสัมพันธ์ของ Y กับ X มีจำนวน k สมการตามลำดับความสำคัญของตัวแปร X ดังนี้คือ

ขั้นแรก เมื่อได้ x_1 เป็นตัวแปรที่มีความสำคัญต่อ y มากที่สุดจะได้สมการ

$$y = a + b_1x_1$$

เมื่อ b_1 เป็นค่า Regression Coefficient

x_1 เป็น Independent variable ตัวที่ 1

y เป็น Dependent variable

a เป็น intercept ของสมการ

ขั้นต่อไป เพิ่ม $x_2,\ldots x_k$ ที่มีความสำคัญรองลงไปเข้าในสมการจะได้

$y = a' + b_1'x + b_2x_2$ เมื่อเพิ่ม x_2 เข้าในสมการ

$y = a'' + b_1''x_1 + b_2'x_2 + b_3x_3$ เมื่อเพิ่ม x_3 เข้าในสมกา

$y = a+b_1x_1+b_2x_2+\ldots\ldots+b_kx_k$ เมื่อเพิ่ม x_k เข้าในสมการ

สมการสุดท้ายแสดงความสัมพันธ์ของ X ทั้ง k ตัว กับ Y ตามลำดับความสำคัญของ X

ค่า a, b_i ในแต่ละสมการจะไม่เท่ากัน เนื่องจากจำนวนตัวแปรที่นำมาพิจารณาในแต่ละสมการมีไม่เท่ากัน

<u>วิธีที่ 2</u> การดำเนินการขั้นแรกเหมือนกับวิธีที่ 1 แต่ในการหา Regression Coefficient แต่ละตัวนั้นจะทดสอบค่า b โดยใช้การทดสอบแบบ t ในแต่ละขั้นที่เพิ่มตัวแปร X เข้าในสมการ ค่า b ตัวใดที่ไม่มีนัยสำคัญ (non significant) จะตัดออกจากสมการ ดังนั้นเมื่อเครื่องคำนวณถึงค่า b ตัวใดที่ไม่มีนัยสำคัญ เครื่องจะหยุดคำนวณ เพราะว่า X ตัวถัดไปย่อมไม่มีนัยสำคัญด้วย เนื่องจากได้มีการจัดเรียงลำดับความสำคัญของ X ไว้ก่อนแล้ว ดังนั้นตัวแปร X ทั้ง k ตัว จะไม่ปรากฏในสมการสุดท้ายทุกตัว จะปรากฏเฉพาะตัวที่ทดสอบค่า b แล้วมีนัยสำคัญเท่านั้น สมการแสดงความสัมพันธ์ของตัวแปร X และ Y จึงมีขนาดเล็กกว่าสมการที่ได้จากวิธีแรก

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots\ldots$$

การคำนวณหาความสัมพันธ์ของ X, Y ทั้ง 2 วิธีมีข้อสังเกตดังนี้คือ

<u>ในกรณีที่ตัวแปร X มีจำนวนมาก</u>

สมการจากวิธีที่ 1 ทำให้เกิดความยุ่งยากในการคำนวณเนื่องจากต้องคำนวณผลจาก X ทุกตัวที่มีต่อ Y ไม่ว่า X ตัวนั้นจะมีอิทธิพลต่อ Y น้อยมากจนอาจคิดว่าไม่มีผลกระทบกระเทือนค่า Y เลยก็ตาม การนำตัวแปรที่ไม่มีนัยสำคัญเข้ามาในสมการจึงไม่เกิดประโยชน์ในการประมาณค่า Y เท่าใดนัก เพราะตัวแปร X มีจำนวนมาก ตัวแปรแต่ละตัวจึงมีอิทธิพลต่อส่วนรวมน้อยมาก การตัดตัวแปรที่มีความสำคัญน้อยออก จึงไม่เกิดผลกระทบกระเทือนกับการประมาณค่า Y ถ้าใช้วิธีที่ 1 จะ เพิ่มความรบกวน (disturbance) ในการคำนวณ เมื่อตัวแปร X มีจำนวนมาก จึงควรเลือกวิธีที่ 2 ในการคำนวณ

ในกรณีที่ X มีจำนวนน้อย

การหาความสัมพันธ์ของตัวแปรโดยวิธีที่ 1 จะช่วยให้ประมาณค่า Y ได้ถูกต้องกว่าวิธีที่ 2 เนื่องจากตัวแปรมีจำนวนน้อย การนำตัวแปรทุก ๆ ตัวเข้ามาพิจารณาในสมการจะให้ผลในการประมาณที่ถูกต้องกว่า การพิจารณาจากสมการที่ตัดตัวแปรบางตัวออก ในขณะที่ตัวแปรมีจำนวนน้อยอยู่แล้ว อาจเกิดความผิดพลาดในการประมาณได้ง่ายกว่าถ้าจะหาความสัมพันธ์ของตัวแปรโดยวิธีที่ 2 ดังนั้นเมื่อตัวแปร (X) มีจำนวนน้อย จึงเลือกวิธีที่ 1 ในการหาความสัมพันธ์ของตัวแปรต่าง ๆ

4.2 ผลการคำนวณและการวิเคราะห์

ข้อมูลที่ใช้ในการคำนวณเป็นข้อมูลของผู้รับทุนจำนวน 159 คน ข้อมูลของแต่ละคนประกอบด้วยตัวแปร 4 ตัวคือ

1. ตัวแปรอิสระ X_1 ได้แก่อายุของผู้รับทุน
2. ตัวแปรอิสระ X_2 ได้แก่คะแนนภาษาอังกฤษของผู้รับทุน
3. ตัวแปรอิสระ X_3 ได้แก่ระยะเวลาตั้งแต่ผู้รับทุนจบมหาวิทยาลัยในประเทศจนถึงวันที่ผู้รับทุนเดินทางไปศึกษาต่อในต่างประเทศ
4. ตัวแปรตาม Y ได้แก่คะแนนเฉลี่ยของผู้รับทุนในมหาวิทยาลัยต่างประเทศ

การคำนวณหาความสัมพันธ์ของ Y และ X ดังกล่าว ไม่สามารถคำนวณด้วยมือตามวิธีธรรมดา เนื่องจากข้อมูลมีจำนวนมาก จึงต้องใช้เครื่องคอมพิวเตอร์ช่วยในการคำนวณ

1. <u>ขั้นแรก</u> เครื่องคอมพิวเตอร์เลือกได้ X_2 เป็นตัวแปรที่มีความสำคัญต่อ Y มากที่สุด

จากสมการ $Y = a + b_2X_2$

ได้สมการ $Y = 1.38378 + 0.000595X_2$

เมื่อกำหนดให้ a เป็นค่า intercept = 1.38378

b_2 เป็นค่า **Partial Regression coefficient** ของ X_2 ต่อ Y = + 0.00595

หมายความว่าเมื่อ X_2 เพิ่มขึ้น 1 หน่วยจะมีผลทำให้ Y เพิ่มขึ้น 0.000595 หน่วย นั่นคือ เมื่อผู้รับทุนได้คะแนนภาษาอังกฤษเพิ่มขึ้น 1 หน่วย มีผลทำให้คะแนนเฉลี่ยในต่างประเทศเพิ่มขึ้น = 0.000595 หน่วยโดยประมาณ

มัชฌิมเลขคณิต 1/ ของคะแนนภาษาอังกฤษ ($\overline{X}_2$) = 322.01886 คะแนน

ส่วนเบี่ยงเบนมาตรฐาน 2/ ของคะแนนภาษาอังกฤษ = 16.14160

1/ มัชฌิมเลขคณิต ($\overline{X}_2$) หาได้จากผลรวมของคะแนนภาษาอังกฤษหารด้วยจำนวนผู้รับทุน

จากสูตร $\overline{X} = \frac{\sum x_i}{n}$ เมื่อ x_i คือคะแนนภาษาอังกฤษของผู้รับทุนคนที่ i

n คือจำนวนผู้รับทุนทั้งหมด

2/ ส่วนเบี่ยงเบนมาตรฐานหาได้จากสูตร $SE = \sqrt{\frac{\sum(x_i - \overline{x})^2}{n}}$ หารด้วย n เพราะถือว่าผู้รับทุนกลุ่มนี้เป็นประชากร (population)

COEFFICIENT OF VARIATION ของ X_2 (C.V.)$^{3/}$ = 5%

แสดงว่า ความกระจายของคะแนนภาษาอังกฤษเปรียบเทียบกับมัชฌิมเลขคณิตของคะแนนภาษาอังกฤษ 5% ซึ่งนับว่า คะแนนภาษาอังกฤษมีการกระจายน้อยมาก การที่คะแนนภาษาอังกฤษมีความกระจายน้อยมากเพราะผู้รับทุนแต่ละคนจะต้องสอบภาษาอังกฤษให้สูงกว่า 290 จึงจะได้รับพิจารณาให้ไปศึกษาต่อ ณ ต่างประเทศได้

$$S_{b_2} = 0.00260$$

$$\therefore \quad t = \frac{0.000595}{0.00260} = 2.288^{*}$$

$$|t_{.05}| = 1.96, \quad |t_{.01}| = 2.57$$

$\therefore$ Reject $H_0 : B_2 = 0$ ที่ $\alpha = .05$

นั่นคือ X_2 มีอิทธิพลต่อ Y อย่างมีนัยสำคัญ ที่ $\alpha = .05$

2. ขั้นสอง เมื่อเพิ่มตัวแปร X_3 ที่มีความสำคัญรองลงมาเข้าในสมการ X_3 คือระยะเวลาตั้งแต่จบมหาวิทยาลัยในประเทศ จนถึงวันที่ไปศึกษาต่อต่างประเทศ

จากสมการ $Y = a + b_2X_2 + b_3X_3$

ได้สมการ $Y = 1.48308 + 0.00575X_2 - 0.00043X_3$

โดยกำหนดให้ a เป็น intercept ของสมการ = 1.48308

$^{3/}$C.V. = $\frac{S.E}{\bar{X}} \times 100$ มีประโยชน์ในการเปรียบเทียบส่วนเบี่ยงเบนมาตรฐานเป็นเปอร์เซ็นต์ของมัชฌิมเลขคณิต C.V. ของข้อมูลแต่ละชุดมักมีค่าคงที่ในขณะที่ S.E , $\bar{X}$ ของข้อมูลมักเปลี่ยนแปลงไปตามความกระจายของข้อมูล นอกจากนี้ C.V. ยังมีประโยชน์ในการเปรียบเทียบความกระจายของข้อมูลที่มีหน่วยต่างกัน เนื่องจาก C.V. ไม่มีหน่วย

b_2 เป็น Partial Regression Coefficient ของ X_2 ต่อ Y เมื่อ X_3 คงที่

มีค่า $= 0.00575$ 4/

b_3 เป็น Partial Regression Coefficient ของ X_3 ต่อ Y เมื่อ X_2 คงที่

มีค่า $= -0.00043$ 5/

มัชฌิมเลขคณิตของ X_3 ($\bar{X}_3$) = 78.333 เดือน

ส่วนเบี่ยงเบนมาตราฐาน (S.E.) = 42.79544

$\therefore$ Coefficient of Variation (C.V.) = 54.86%

แสดงว่าการกระจายของระยะเวลาตั้งแต่จบมหาวิทยาลัยจนถึงวันไปศึกษาต่อต่างประเทศของผู้รับทุนเปรียบเทียบกับมัชฌิมเลขคณิต = 54.86%

นั่นคือ X_3 มีการกระจายมากพอประมาณ

$S_{b_2} = 0.00265$, $t_2 = 2.169^*$

$\therefore$ Reject $H_0 : B_2 = 0$ ที่ $= .05$

$S_{b_3} = 0.00100$, $t_3 = -0.433$

$\therefore$ Accept $H_0 : \beta_3 = 0$ ที่ $\alpha = .05$

4/ ค่า b_2 ในสมการนี้ไม่เท่ากับ b_2 ในสมการแรกเนื่องจากเพิ่มตัวแปร X_3 เข้าในสมการ

5/ ค่า $b_3 = -0.00043$ แสดงว่าความสัมพันธ์ระหว่าง X_3 และ Y เป็นไปในทางกลับกัน กล่าวคือ เมื่อ X_3 เพิ่มขึ้น 1 หน่วยทำให้ค่า Y ลดลง 0.00043 หน่วยโดยประมาณ

นั่นคือใน step ที่ 2 เมื่อเพิ่ม X_3 เข้าในสมการแล้ว การเปลี่ยนแปลงของค่า X_3 มีอิทธิพลต่อการเปลี่ยนแปลงของ Y น้อยมากจนไม่มีนัยสำคัญทางสถิติ

3. <u>ขั้นที่สาม</u> เมื่อเพิ่ม X_1 เข้าในสมการ X_1 ได้แก่อายุของผู้รับทุน

จากสมการ $Y = a + b_2X_2 + b_3X_3 + b_1 X_1$

ได้สมการ $Y = 1.49946+0.00574X_2-0.00040X_3-0.00047X_1$

เมื่อกำหนดให้ a เป็น intercept = 1.49946

b_2 เป็น Partial Regression Coefficient ของ X_2 ต่อ Y

เมื่อ X_3, X_1 คงที่ มีค่า = 0.00574*

b_3 เป็น Partial Regression Coefficient ของ X_3 ต่อ Y

เมื่อ X_1, X_2 คงที่

มีค่า = -0.0040*

b_1 เป็น Partial Regression Coefficient ของ X_1 ต่อ Y

เมื่อ X_2, X_3 คงที่

มีค่า = -0.0047 6/

6/ ค่า $b_1 = -0.00047$ แสดงว่าความสัมพันธ์ระหว่าง X_1 และ Y_1 เป็นไปในทางกลับกัน กล่าวคือเมื่อ X_1 เพิ่มขึ้น 1 หน่วย ค่า Y จะลดลง .00047 หน่วยโดยประมาณ

* พึงสังเกตว่า ค่า b_2, b_3 เปลี่ยนไปจากเดิมแต่เครื่องหมายยังคงเดิม ทั้งนี้เนื่องจากการเพิ่มตัวแปรอิสระ X_1 เข้าในสมการ

มัชฌิมเลขคณิตของ X_1 ($\bar{X}_1$) = 32 ปี

ส่วนเบี่ยงเบนมาตรฐาน (S.E.) = 5.21

Coefficient of variation (C.V.)= 16.28%

แสดงว่าการกระจายของอายุผู้รับทุนเมื่อเปรียบเทียบกับมัชฌิมเลขคณิต

ของอายุ = 16.28%

นั่นคือผู้รับทุนแต่ละคนมีอายุต่างกันพอประมาณ

$S_{b_2} = .00267$, $t_2 = 2.150^*$

$\therefore$ Reject $H_0 : \beta_2 = 0$ ที่ $\alpha = .05$

$S_{b_3} = 0.00116$, $t_3 = -0.347$

$\therefore$ Accept $H_0 : \beta_3 = 0$ ที่ $\alpha = .05$

$S_{b_1} = 0.00953$, $t_1 = -0.049$

$\therefore$ Accept $H_0 : \beta_1 = 0$ ที่ $\alpha = .05$

สรุปผลจากการทดสอบค่า b_i ในสมการสุดท้ายได้ดังนี้คือ

1. การเปลี่ยนแปลงของ X_2 มีอิทธิพลต่อการเปลี่ยนแปลงของ Y อย่างมีนัยสำคัญที่ $\alpha = .05$

2. การเปลี่ยนแปลงของ X_3 ไม่มีอิทธิพลต่อการเปลี่ยนแปลงของ Y ที่ระดับนัยสำคัญ $\alpha = .05$

3. การเปลี่ยนแปลงของ X_1 ไม่มีอิทธิพลต่อการเปลี่ยนแปลงของ Y ที่ระดับนัยสำคัญ $\alpha = .05$

Multiple correlation coefficient $r_{y\hat{y}} = 0.183$

จากตาราง 7/ ค่า $R_{.05}$ = .225 , $R_{.01}$ = .270 เมื่อมีตัวแปรอิสระ 3 ตัว

∴ ค่า $r_{y\hat{y}}$ = 0.183 น้อยกว่าค่าในตาราง

ความสัมพันธ์ของ Y, $\hat{Y}$ ไม่เป็นที่ยอมรับที่ α = 0.05, .01

นั่นคือค่าประมาณ $\hat{Y}$ ไม่เป็นที่ยอมรับที่ α = 0.05, .01

<u>การหาค่าความสัมพันธ์ของตัวแปรต่าง ๆ จาก CORRELATION MATRIX</u>

	Y	X_1	X_2	X_3
Y	1.00000	-0.04981	0.17963	-0.06532
X_1		1.00000	-0.16381	0.52201
X_2			1.00000	-0.17691
X_3				1.00000

CORRELATION MATRIX มีลักษณะเป็น Symmetric Matrix

เมื่อกำหนดให้ a_{ij} เป็น element ของ matrix แถวที่ i สดมภ์ที่ j

a_{12} = a_{21} เป็นค่าแสดงความสัมพันธ์ของ Y และ X_1 = -0.04981

แสดงว่าความสัมพันธ์ของคะแนนเฉลี่ยในต่างประเทศและอายุเป็นไปในทางกลับกัน ตัวอย่างเช่น ผู้ที่รับทุนอายุมากจะมีคะแนนเฉลี่ยในต่างประเทศน้อย

" น้อย " " มาก

7/ ค่า R ได้จากตารางของหนังสือ PRINCIPLES AND PROCEDURES OF STATISTICS ของ STEEL & TORRIE หน้า 453

$a_{13} = a_{31}$ เป็นค่าแสดงความสัมพันธ์ของ Y และ X_2 = 0.17963

แสดงว่าความสัมพันธ์ของคะแนนเฉลี่ยในต่างประเทศและคะแนนภาษาอังกฤษเป็นไปในทางเดียวกัน คือผู้ที่มีคะแนนภาษาอังกฤษสูงย่อมมีคะแนนเฉลี่ยในต่างประเทศสูงด้วย

ตัวอย่างเช่น ผู้รับทุนที่มีคะแนนภาษาอังกฤษสูง ย่อมมีคะแนนเฉลี่ยในต่างประเทศสูง

" " ต่ำ " " ต่ำ

$a_{14} = a_{41}$ เป็นค่าแสดงความสัมพันธ์ของ Y และ X_3 = -0.6532

แสดงว่าความสัมพันธ์ของคะแนนเฉลี่ยในต่างประเทศและระยะเวลาตั้งแต่จบมหาวิทยาลัยในประเทศจนถึงวันที่เดินทางไปศึกษาต่อต่างประเทศมีความสัมพันธ์ในทางกลับกัน

ตัวอย่างเช่น ผู้รับทุนที่จบมหาวิทยาลัยมานานย่อมมีคะแนนเฉลี่ยในต่างประเทศต่ำ

" ไม่นาน " " สูง

$a_{23} = a_{32}$ เป็นค่าแสดงความสัมพันธ์ของ X_1 และ X_2 = -0.16381

แสดงว่าความสัมพันธ์ของอายุและคะแนนภาษาอังกฤษเป็นไปในทางกลับกัน

ตัวอย่างเช่น ผู้รับทุนที่มีอายุสูงได้รับคะแนนภาษาอังกฤษต่ำ

" น้อย " สูง

$a_{24} = a_{42}$ เป็นค่าแสดงความสัมพันธ์ของ X_1 และ X_3 = 0.52201

แสดงว่าความสัมพันธ์ของอายุและระยะเวลาที่ผู้รับทุนจบจากมหาวิทยาลัยในประเทศจนถึงวันที่เดินทางไปศึกษาต่อ เป็นไปในทางเดียวกัน นั่นคือผู้ที่มีอายุมากย่อมจบจากมหาวิทยาลัยมานานกว่าผู้ที่มีอายุน้อย

ตัวอย่างเช่น ผู้รับทุนที่มีอายุสูงจบจากมหาวิทยาลัยมาเป็นเวลานาน

" น้อย " " ไม่นาน

$a_{34} = a_{43}$ เป็นค่าแสดงความสัมพันธ์ระหว่าง X_2 และ X_3 = -0.17691

แสดงว่าความสัมพันธ์ของคะแนนภาษาอังกฤษ และระยะเวลาตั้งแต่จบมหาวิทยาลัยจนถึงวันไปศึกษาต่อต่างประเทศ เป็นไปในทางกลับกัน

ตัวอย่างเช่น ผู้รับทุนที่จบมหาวิทยาลัยมานานได้รับคะแนนภาษาอังกฤษต่ำ

" ไม่นาน " สูง

แต่อย่างไรก็ตาม ค่า correlation coefficient ที่มีค่าน้อยกว่า

$$r_{.05} = 0.159 \;,\; r_{.01} = .208$$

จะถือว่าความสัมพันธ์ของตัวแปรคู่ที่กำลังพิจารณาไม่มีนัยสำคัญทางสถิติ

<u>ตารางสรุปผลค่าต่าง ๆ ของตัวแปร</u>

Parameter / ตัวแปร	$\overline{X}$	S.E.	C.V.(%)
Y	3.3	0.53	6.06
X_1	32	5.21	16.28
X_2	322	16.14	5.0
X_3	78	42.79	54.86

ค่าต่าง ๆ ในตารางมีความหมายดังต่อไปนี้

ก. <u>ค่ามัชฌิมเลขคณิตของตัวแปรต่าง ๆ</u>

ค่ามัชฌิมเลขคณิตของคะแนนเฉลี่ยในมหาวิทยาลัยต่างประเทศ		= 3.3	แต้มเฉลี่ย
"	อายุ ของผู้รับทุน	= 32	ปี
"	คะแนนภาษาอังกฤษ	= 322	คะแนน
"	ระยะเวลาตั้งแต่จบมหาวิทยาลัยในประเทศจนถึงวันไปศึกษาต่อในต่างประเทศ	= 78	เดือน

ข. <u>ค่าสัมประสิทธิ์การกระจาย (C.V.) ของตัวแปรต่าง ๆ</u>

1. ค่า C.V. ของคะแนนภาษาอังกฤษของผู้รับทุนมีค่าน้อยที่สุด แสดงว่าผู้รับทุนแต่ละคนมีคะแนนภาษาอังกฤษใกล้เคียงกัน ทั้งนี้เพราะผู้รับทุนทุกคนต้องสอบภาษาอังกฤษ (ALI/GU TEST) ให้ได้อย่างน้อย 290 คะแนน จึงจะได้รับพิจารณาให้เข้าศึกษาต่อในมหาวิทยาลัยต่างประเทศได้ ส่วนคะแนนขั้นสูงก็มีขีดจำกัดตามความสามารถของผู้รับทุนแต่ละคน ดังนั้นคะแนนภาษาอังกฤษของผู้รับทุนจึงกระจายอยู่ในช่วงหนึ่ง ซึ่งมีความแตกต่างกันไม่มาก

2. ค่า C.V. ของระยะเวลาตั้งแต่จบมหาวิทยาลัยในประเทศจนถึงวันไปศึกษาต่อต่างประเทศมีการกระจายมากที่สุดในบรรดาตัวแปรทั้ง 4 ตัว แสดงว่าผู้รับทุนแต่ละคนจบจากมหาวิทยาลัยมาเป็นระยะเวลานานแตกต่างกันมาก ทั้งนี้เนื่องจาก ไม่มีข้อกำหนดแน่นอนว่า ผู้รับทุนจะได้ไปศึกษาต่อต่างประเทศเมื่อทำงานมานานกี่ปี แต่ผู้รับทุนจะถูกคัดเลือกตามความเหมาะสมของหน้าที่การงานที่ปฏิบัติและความต้องการของหน่วยราชการนั้น ๆ ช่วงระยะเวลาตั้งแต่จบมหาวิทยาลัยจนถึงวันไปศึกษาต่อต่างประเทศของผู้รับทุนแต่ละคน จึงแตกต่างกันมาก

3. ค่า C.V. ของอายุของผู้รับทุนน้อยกว่าค่า C.V. ของระยะเวลาที่จบจากมหาวิทยาลัย ทั้ง ๆ ที่ค่าทั้งสองไม่ควรแตกต่างกันมาก แต่เหตุที่ค่า C.V. ของอายุมีค่าน้อยกว่าประมาณ 3 เท่า เนื่องจากถึงแม้ว่าผู้รับทุนจะมีอายุเท่ากัน ช่วงระยะเวลาตั้งแต่จบมหาวิทยาลัยจนถึงวันไปศึกษาต่อต่างประเทศของผู้รับทุนอาจแตกต่างกันมาก เพราะคนหนึ่งจบจากมหาวิทยาลัยเมื่ออายุน้อย แต่อีกคนหนึ่งจบจากมหาวิทยาลัยเมื่ออายุมาก จึงทำให้ C.V. ของอายุมีค่าน้อยกว่า C.V. ของช่วงระยะเวลาตั้งแต่จบมหาวิทยาลัย จนถึงวันไปศึกษาต่อต่างประเทศ

4. คะแนนเฉลี่ยในมหาวิทยาลัยต่างประเทศ มีค่า C.V. ไม่มากนัก เนื่องจากผู้รับทุนทุกคนต้องสอบให้ผ่านคะแนนเฉลี่ยขั้นต่ำที่กำหนด คือ 3 แต้มเฉลี่ย ส่วนคะแนนขั้นสูง ก็มีขีดจำกัดตามความสามารถของผู้รับทุนแต่ละคน C.V. ของ Y จึงมีความกระจายไม่มาก

<u>การประมาณค่า $\hat{Y}$ จากสมการ</u>

โดยปกติการพยากรณ์ค่า Y มักทำโดยการแทนค่าตัวแปรอิสระต่าง ๆ เพื่อหาค่าตัวแปรตามทางซ้ายมือ

จากสมการ $Y = 1.49946+0.00574X_2-0.00040X_3-0.00047X_1$

โดยการแทนค่า X_1, X_2, X_3 ก็จะได้ค่า Y ตามต้องการ

ในกรณีนี้ค่า $R = r_{y\hat{y}}$ ที่คำนวณได้มีค่าน้อยกว่า $r_{.05}, r_{.01}$ ซึ่งแสดงว่าความสัมพันธ์ของ Y ค่าจริง และ $\hat{Y}$ ค่าประมาณไม่อยู่ในระดับที่มีนัยสำคัญ (ที่ α =.05, .01) ดังนั้นถ้าเราจะใช้สมการดังกล่าวข้างต้นพยากรณ์ ความสำเร็จของผู้รับทุนก็จะทำให้เราไม่สามารถจำกัด ความผิดพลาด (error) ในการประมาณค่า Y ตามวิธีทางสถิติได้

อย่างไรก็ตาม ถ้าเรายอมรับความผิดพลาดที่เกิดขึ้นมากกว่า 5% ในการพยากรณ์ เราก็สามารถประมาณค่า Y ของผู้รับทุนอย่างคร่าว ๆ จากตัวแปร 3 ตัว ดังกล่าว (ดูรายละเอียดของตัวเลขเกี่ยวกับค่าตัวแปรต่าง ๆ และค่า $\hat{Y}$ ที่คำนวณได้ในภาคผนวก)

* * * * * * * *