บทที่ 3

การวิเคราะห์ข้อมูล

ในบทนี้จะวิเคราะห์ข้อมูล โดยใช้วิธีต่าง ๆ ดังได้กล่าวไว้ในบทที่ 2 ข้อมูลที่ใช้เป็นข้อมูลทุติยภูมิ (Secondary data) รวบรวมจากสภาหอการค้าแห่งประเทศไทย เป็นข้อมูลที่เกี่ยวกับ

ราคาข้าวสารเจ้าชนิด 100 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

ราคาข้าวเปลือกเจ้าชนิด 100 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

ราคาข้าวสารเจ้าชนิด 5 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

และ ปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ เป็นจำนวนเมตริกตัน

ข้อมูลที่รวบรวมมาเพื่อวิเคราะห์นี้เป็นข้อมูลของเดือน มกราคม - ธันวาคม ของปี พ.ศ.2515, 2517 และ 2518

ในการวิเคราะห์การถดถอยเชิงเส้นอย่างง่าย โดยวิธีดังได้กล่าวในบทที่ 2 ได้นำข้อมูลที่รวบรวมมาได้มาทำให้ค่าสังเกตบางค่าขาดหายไปโดยวิธีสุ่มให้ค่าสังเกต Y และ X หายไปรวมกัน 20 %, 30 %, 35 % และ 40 % ของข้อมูลที่รวบรวมมาได้ทั้งหมด และใช้วิธีของบทที่ 2 ในการประมาณค่าที่ขาดหายไป แล้วจึงนำข้อมูลที่ได้ใหม่มาหาความสัมพันธ์เชิงพังก์ชั่นระหว่าง Y กับ X_1

ส่วนการวิเคราะห์การถดถอยเชิงเส้นพหุคูณนั้น ได้นำข้อมูลมาทำให้ค่าสังเกตขาดหายไปโดยวิธีสุ่มเช่นกัน โดยให้ค่าสังเกต Y, X_1 และ X_2 ขาดหายไปรวมกันประมาณ 35 % ของข้อมูลที่รวบรวมมาได้ทั้งหมด และใช้วิธีในบทที่ 2 ประมาณค่าที่ขาดหายไป แล้วจึงนำข้อมูลที่ได้ใหม่มาหาความสัมพันธ์เชิงพังก์ชั่นระหว่าง Y กับ X_1 และ X_2

การวิเคราะห์ข้อมูล

เนื่องจากในการวิเคราะห์ข้อมูลที่รวบรวมมานี้ ต้องทำหลายวิธีด้วยกัน ดังนั้น เพื่อป้องกันการผิดพลาดจากการคำนวณตัวเลข จึงได้ใช้เครื่องคอมพิวเตอร์ในการวิเคราะห์ข้อมูล โดยใช้เครื่องที่บริษัท บางกอกดาต้า เซนเตอร์ จำกัด

การวิเคราะห์ข้อมูลในบทนี้ได้ใช้ข้อมูลซึ่งสุ่มให้ค่า x และ y หายไปรวมกัน 40 % ตามตารางข้อมูล ในภาคผนวก โดยกำหนดให้

Y = ราคาข้าวสารเจ้าชนิด 100 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

X_1 = ราคาข้าวเปลือกเจ้าชนิด 100 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

X_2 = ปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ เป็นจำนวนเมตริกตัน

X_3 = ราคาข้าวสารเจ้าชนิด 5 % เป็นเงินบาทต่อหนึ่งเมตริกตัน

1. วิเคราะห์ข้อมูลโดยวิธีกำลังสองน้อยที่สุด

การวิเคราะห์การถดถอยโดยวิธีนี้จะใช้เฉพาะค่าที่มีอยู่ครบทั้งค่าตัวแปรอิสระ (X) และค่าตัวแปรตาม (Y) ทำขึ้นเพื่อจะดูผลการวิเคราะห์ว่าค่า R^2 ที่ได้จะต่างกับวิธีอื่น ๆ หรือไม่เพียงใด

1.1 การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}^{LS} = b_o^{LS} + b_1^{LS}X_{1i}$

เมื่อข้อมูลที่รวบรวมได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$$n_c = 22$$

$$\bar{X}_{1c} = \frac{\sum_{i=1}^{n_c} X_{1i}}{n_c} = 2160.11$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_1^{LS} = \frac{\sum_{i=1}^{n_c} X_{1i} Y_i - n_c \bar{X}_{1c} \bar{Y}_c}{\sum_{i=1}^{n_c} X_{1i}^2 - n_c \bar{X}_{1c}^2} = 1.53750$$

$$b_0^{LS} = \bar{Y}_c - b_1^{LS} \bar{X}_{1c} = 168.90651$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 3.90651 + 1.53750X_{1i}$

เราทดสอบสมมติฐาน $H_o : \beta_1 = 0$; $H_A : \beta_1 \neq 0$ โดยใช้ตารางทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 12

ANOVA

SOV	df	SS	MS	F
Reg	1	2.0381304×10^7	2.0381304×10^7	696.05875^{**}
Error	20	5.8562021×10^5	2.9281010×10^4	
Total	21	2.0966924×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.97207$$

หมายความว่าการเปลี่ยนแปลงราคาข้าวสารเจ้าชนิด 100 % (Y) ขึ้นอยู่กับราคาข้าวเปลือกเจ้าชนิด 100 % (X_1) = 97.207 % นอกนั้นเป็นอิทธิพลเนื่องมาจากสาเหตุอื่น

(2) เมื่อต้องการหา $\hat{Y}_i^{LS} = b_o^{LS} + b_2^{LS} X_{2i}$

เมื่อข้อมูลที่รวบรวมได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก $n_c = 22$

$$\bar{X}_{2c} = \frac{\sum_{i=1}^{n_c} X_{2i}}{n_c} = 109,107$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_2^{LS} = \frac{\sum_{i=1}^{n_c} X_{2i} Y_i - n_c \bar{X}_{2c} \bar{Y}_c}{\sum_{i=1}^{n_c} X_{2i}^2 - n_c \bar{X}_{2c}^2} = -0.01361$$

$$b_o^{LS} = \bar{Y}_c - b_2^{LS} \bar{X}_{2c} = 4,974.51320$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 4,974.51320 - 0.01361 X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta_2 = 0$; $H_A : \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 13

ANOVA

SOV	df	SS	MS	F
Reg	1	1.2579447×10^7	1.2579447×10^7	29.99579**
Error	20	8.3874764×10^6	4.1937382×10^5	
Total	21	2.0966924×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.59997$$

หมายความว่าการเปลี่ยนแปลงราคาข้าวสารเจ้าชนิด 100 % (Y) ขึ้นอยู่กับปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ (X_2) = 59.997 % นอกนั้นเป็นอิทธิพลเนื่องมาจากสาเหตุอื่น

(3) เมื่อต้องการหา $\hat{Y}_i^{LS} = b_o^{LS} + b_3^{LS} X_{3i}$

เมื่อข้อมูลที่รวบรวมได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22$

$$\bar{X}_{3c} = \frac{\sum_{i=1}^{n_c} X_{3i}}{n_c} = 3183.8636$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_3^{LS} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n_c\bar{X}_{3c}\bar{Y}_c}{\sum_{i=1}^{n_c} X_{3i}^2 - n_c\bar{X}_{3c}^2} = 1.07162$$

$$b_o^{LS} = \bar{Y}_c - b_3^{LS}\bar{X}_{3c} = 78.20098$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 78.20098 + 1.07162X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0$; $H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 14

ANOVA

SOV	df	SS	MS	F
Reg	1	2.0859295×10^7	2.0859295×10^7	3876.14243**
Error	20	1.0762914×10^5	5.3814572×10^3	
Total	21	2.0966924×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.99487$$

หมายความว่าการเปลี่ยนแปลงราคาข้าวสารเจ้าชนิด 100 % (Y) ขึ้นอยู่กับราคาข้าวสารเจ้าชนิด 5 % (X_3) = 99.487 % นอกนั้นเป็นอิทธิพลเนื่องมาจากสาเหตุอื่น

1.2 การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)

เมื่อต้องการหา $\hat{Y}_i^{LS} = b_o^{LS} + b_1^{LS}X_{1i} + b_2^{LS}X_{2i}$

เมื่อข้อมูลที่รวบรวมได้ตามตารางข้อมูลชุดที่ 5 ในภาคผนวก $n_c = 23$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$\hat{\beta}^{LS} = (X'X)^{-1}X'Y$$

$$= \begin{bmatrix} -97.80121 \\ 1.61629 \\ 0.0007 \end{bmatrix}$$

∴ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = -97.80121 + 1.61629X_{1i} + 0.0007X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta = 0$; $H_A : \beta \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 15

ANOVA

SOV	df	SS	MS	F
Reg	3	1.9940208×10^7	6.6467600×10^6	234.17626**
Error	20	5.6767154×10^5	2.8383577×10^4	
Total	23	2.0507879×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.97232$$

หมายความว่าการเปลี่ยนแปลงราคาข้าวสารเจ้าชนิด 100 % (Y) ขึ้นอยู่กับราคาข้าวเปลือกเจ้าชนิด 100 % (X_1) และปริมาณข้าวสารเจ้าส่งออกทางท่าเรือกรุงเทพฯ (X_2) = 97.232 % นอกนั้นเป็นอิทธิพลเนื่องมาจากสาเหตุอื่น

2. วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไปโดยวิธีกำลังสองน้อยที่สุด

การวิเคราะห์การถดถอยโดยวิธีนี้ ใช้เฉพาะค่าสังเกตที่มีครบทั้งตัวแปรตาม (Y) และตัวแปรอิสระ (X) หาสัมการถดถอยของ Y ในรูป X และใช้สมการถดถอยที่ได้นี้ประมาณค่า Y's ที่ขาดหายไป แล้วจึงนำข้อมูลที่ได้ใหม่ไปหาสมการถดถอย

2.1 การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}_i^{(1)} = b_1^{(1)} + b_1^{(1)}X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$$n_c = 22, \quad n_{x_1} = 29, \quad m_y = 7$$

$$\bar{X}_{1c} = \frac{\sum_{i=1}^{n_c} X_{1i}}{n_c} = 2,160.11$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_1^{LS} = \frac{\sum^{n_c} X_{1i}Y_i - n_c\bar{X}_{1c}\bar{Y}_c}{\sum_{i=1}^{n_c} X_{1i}^2 - n_c\bar{X}_{1c}^2} = 1.53750$$

$$b_o^{LS} = \bar{Y}_c - b_1^{LS}\bar{X}_{1c} = 168.90651$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 168.90651 + 1.53750X_{1i}$

ประมาณค่า Y_i's ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้ ซึ่งจะได้ค่า คือ

เมื่อ $X_1 = 2,607.50$; $\hat{Y} = 168.90651 + 1.53750(2607.50)$

$= 4,177.94$

$X_1 = 2,626.75$; $\hat{Y} = 168.90651 + 1.53750(2,626.75)$

$= 4,207.53$

$X_1 = 2,624.-$; $\hat{Y} = 168.90651 + 1.5375(2,624.-)$

$= 4,203.31$

$X_1 = 2,512.50$; $\hat{Y} = 168.90651 + 1.5375(2,512.50)$

$= 4,031.88$

$X_1 = 2,512.50$; $\hat{Y} = 168.90651 + 1.5375(2,512.50)$

$= 4,031.88$

$X_1 = 2,682.50$; $\hat{Y} = 168.90651 + 1.5375(2,682.50)$

$= 4,293.25$

$X_1 = 1,128.00$; $\hat{Y} = 168.90651 + 1.55375(1,128.00)$

$= 1,903.21$

นำค่า $\hat{Y}_i$'s ที่คำนวณได้นี้แทนค่า Y_i's ที่ขาดหายไป แล้วนำข้อมูลที่ได้ใหม่ซึ่งมีทั้ง X_1 และ Y หาค่า $b_1^{(1)}$ และ $b_o^{(1)}$ โดยวิธีกำลังสองน้อยที่สุด เช่นกัน n_c จะเพิ่มเท่ากับ $n_{x_1} = 29$ และจะได้

$$b_1^{(1)} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}Y_i - n_{x_1}\bar{X}_{n_{x_1}}\bar{Y}_{n_{x_1}}}{\sum_{i=1}^{n_{x_1}} X_{1i}^2 - n_{x_1}\bar{X}_{1n_{x_1}}^2} \quad ; \quad \bar{X}_{n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}}{n_{x_1}}$$

$$\bar{Y}_{n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} Y_i}{n_{x_1}}$$

$$= 1.53750$$

$$b_0^{(1)} = \bar{Y}_{n_x} - b_1^{(1)}\bar{X}_{1n_x}$$

$$= 168.90889$$

∴ สมการถดถอยที่ได้ใหม่คือ

$$\hat{Y}_i^{(1)} = 168.90889 + 1.53750X_{1i}$$

เราทดสอบสมมติฐาน $H_0 : \beta_1 = 0 ; H_A : \beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 16

ANOVA

SOV	df	SS	MS	F
Reg.	1	2.5426495×10^7	2.5426495×10^7	868.36126**
Error	20	5.8562021×10^5	2.9281010×10^4	
Total	21	2.6012115×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.97749$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(1)} = b_o^{(1)} + b_2^{(1)}X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22, \quad n_{x_2} = 29, \quad m_y = 7$

$$\bar{X}_{2c} = \frac{\sum_{i=1}^{n_c} X_{2i}}{n_c} = 109,107.-$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_2^{LS} = \frac{\sum_{i=1}^{n_c} X_{2i}Y_i - n_c\bar{X}_{2c}\bar{Y}_c}{\sum_{i=1}^{n_c} X_{2i}^2 - n_c\bar{X}_{2c}^2} = -0.01361$$

$$b_o^{LS} = \bar{Y}_c - b_2^{(1)}\bar{X}_{2c} = 4,974.51320$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 4,974.51320 - 0.01361X_{2i}$

ประมาณค่า Y_i's ที่ขาดหายไป โดยใช้สมการถดถอยที่ได้นี้

เมื่อ $X_2 = 83,111.-$; $\hat{Y} = 4,974.51320 - 0.01361(83,111.-)$

$= 3,843.37$

$X_2 = 47,457.-$; $\hat{Y} = 4,974.51320 - 0.01361(47,457.-)$

$= 4,328.62$

$X_2 = 93,625.-$; $\hat{Y} = 4,974.51320 - 0.01361(93,625.-)$

$= 3,700.28$

$X_2 = 87,906.-$; $\hat{Y} = 4,974.51320 - 0.01361(87,906.-)$

$= 3,778.11$

$X_2 = 79,629.-$; $\hat{Y} = 4,974.51320 - 0.01361(79,629.-)$

$= 3,890.76$

$X_2 = 35,423.-$; $\hat{Y} = 4,974.51320 - 0.01361(35,423.-)$

$= 4,492.41$

$X_2 = 170,303.-$; $\hat{Y} = 4,774.51320 - 0.01361(170,303.-)$

$= 2,656.69$

นำค่า $\hat{Y}_i$'s ที่คำนวณได้นี้แทนค่า Y_i's ที่ขาดหายไป แล้วนำข้อมูลที่ได้ใหม่ซึ่งมีทั้ง X_2 และ Y หาค่า $b_2^{(1)}$ และ $b_o^{(1)}$ โดยวิธีกำลังสองน้อยที่สุด เช่นกัน n_c จะเพิ่มเท่ากับ $n_{x_2} = 29$ และจะได้

$$b_2^{(1)} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}Y_i - n_{x_2}\bar{X}_{n_{x_2}}\bar{Y}_{n_{x_2}}}{\sum_{i=1}^{n_{x_2}} X_{2i}^2 - n_{x_2}\bar{X}_{n_{x_2}}^2} \quad ; \quad \bar{X}_{n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}}{n_{x_2}}$$

$$\bar{Y}_{n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} Y_i}{n_{x_2}}$$

$$= -0.01361$$

$$b_o^{(1)} = \bar{Y}_{n_{x_2}} - b_2^{(1)}\bar{X}_{2n_{x_2}}$$

$$= 4,974.41699$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(1)} = 4974.41699 - 0.01361X_{2i}$

เราทดสอบสมมติฐาน H_o : $\beta_2 = 0$; H_A : $\beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 17

ANOVA

SOV	df	SS	MS	F
Reg	1	1.5218192×10^7	1.5218192×10^7	36.287887**
Error	20	8.3874775×10^6	4.1937387×10^5	
Total	21	2.3605669×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.64468$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(1)} = b_o^{(1)} + b_3^{(1)}X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22,$ $n_{x_3} = 29,$ $m_y = 7$

$$\bar{X}_{3c} = \frac{\sum_{i=1}^{n_c} X_{3i}}{n_c} = 3,183.86$$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$b_3^{LS} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n_c\bar{X}_{3c}\bar{Y}_c}{\sum_{i=1}^{n_c} X_{3i}^2 - n_c\bar{X}_{3c}^2} = 1.07162$$

$$b_o^{LS} = \bar{Y}_c - b_3^{LS}\bar{X}_{3c} = 78.20098$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = 78.20098 + 1.07162X_{3i}$

ประมาณค่า Y_i's ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้

เมื่อ $X_3 = 3,692.-$; $\hat{Y} = 78.20098 + 1.07162(3,692.-)$

$= 4,034.62$

$X_3 = 3,694.50$; $\hat{Y} = 78.20098 + 1.07162(3,694.50)$

$= 4,037.30$

$X_3 = 3,960.-$; $\hat{Y} = 78.20098 + 1.07162(3,960.-)$

$= 4,321.82$

$X_3 = 4,062.50$; $\hat{Y} = 78.20098 + 1.07162(4062.50)$

$= 4,431.66$

$X_3 = 3,962.50$; $\hat{Y} = 78.20098 + 1.07162(3,962.50)$

$= 4,324.50$

$X_3 = 3,779.- \; ; \qquad \hat{Y} = 78.20098 + 1.07162(3,779.-)$

$= 4,127.85$

$X_3 = 1,602.- \; ; \qquad \hat{Y} = 78.20098 + 1.07162(1,602.-)$

$= 1,794.94$

นำค่า $\hat{Y}_i$'s ที่คำนวณได้นี้ แทนค่า Y_i's ที่ขาดหายไป แล้วนำข้อมูลที่ได้ใหม่ซึ่งมีทั้ง X_3 และ Y หาค่า $b_3^{(1)}$ และ $b_o^{(1)}$ โดยวิธีกำลังสองน้อยที่สุด เช่นกัน n_c จะเพิ่มเท่ากับ $n_{x_3} = 29$ และจะได้

$$b_3^{(1)} = \frac{\sum_{i=1}^{n_{x_3}} X_{3i}Y_i - n_{x_3}\bar{X}_{n_{x_3}}\bar{Y}_{n_{x_3}}}{\sum_{i=1}^{n_{x_3}} X_{3i}^2 - n_{x_3}\bar{X}_{3n_{x_3}}^2} \; ; \quad \bar{X}_{n_{x_3}} = \frac{\sum_{i=1}^{n_{x_3}} X_{3i}}{n_{x_3}}$$

$$\bar{Y}_{n_{x_3}} = \frac{\sum_{i=1}^{n_{x_3}} Y_i}{n_{x_3}}$$

$= 1.07113$

$b_o^{(1)} = \bar{Y}_{n_{x_3}} - b_3^{(1)}\bar{X}_{3n_{x_3}}$

$= 79.05205$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(1)} = 79.05205 + 1.07113X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0 \; ; H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 18

ANOVA

SOV	df	SS	MS	F
Reg	1	2.6768939×10^7	2.6768939×10^7	$4,953.85314^{**}$
Error	20	1.0806720×10^5	5.4036600×10^3	
Total	21	2.6877006×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.99598$$

2.2 การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)

เมื่อต้องการหา $\hat{Y}_i^{(1)} = b_o^{(1)} + b_1^{(1)} X_{1i} + b_2^{(1)} X_{2i}$

หรือเขียนในรูปเมตริกซ์ $\hat{Y}^{(1)} = X\beta^{(1)}$

เมื่อข้อมูลที่รวบรวมได้ตามตารางข้อมูลชุดที่ 5 ในภาคผนวก ซึ่งเขียนให้อยู่ในรูปเมตริกซ์ ได้ดังนี้

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} , \quad X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

โดยที่ Y_1 แทน เวคเตอร์ค่าสังเกต ที่มีอยู่ ซึ่งตรงกับ X_1

Y_2 แทน เวคเตอร์ค่าสังเกต ที่สูญหาย ซึ่งตรงกับ X_2

Y_3 แทน เวคเตอร์ค่าสังเกต ที่มีอยู่ ซึ่งตรงกับ X_3

Y_4 แทน เวคเตอร์ค่าสังเกต ที่สูญหาย ซึ่งตรงกับ X_4

X_1 แทน เมตริกซ์ที่ไม่มีค่าสังเกตตัวใดตัวหนึ่งสูญหาย

X_2 แทน เมตริกซ์ที่ไม่มีค่าสังเกตตัวใดตัวหนึ่งสูญหาย

X_3 แทน เมตริกซ์ที่มีค่าสังเกตสูญหายอย่างน้อยหนึ่งตัว

X_4 แทน เมตริกซ์ที่มีค่าสังเกตสูญหายอย่างน้อยหนึ่งตัว

ให้ $$X_0 = \begin{bmatrix} X_1 \\ X_2 \\ 0 \end{bmatrix}$$

จากข้อมูลที่รวบรวมมาได้จะหาสมการถดถอย

$$\hat{Y}^{LS} = X_1 \hat{\beta}^{LS}$$

ในการหา $\hat{\beta}^{LS}$ จะให้ Y_2, Y_4, X_3 และ X_4 มีค่าเท่ากับศูนย์ โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$\hat{\beta}^{LS} = (X_1' X_1)^{-1} X_1' Y$$

$$= \begin{bmatrix} -97.80121 \\ 1.61629 \\ 0.00070 \end{bmatrix}$$

สมการถดถอยที่ได้คือ $\hat{Y}_i^{LS} = -97.80121 + 1.61629X_{1i} + 0.0007X_{2i}$

ค่า Y_2 ที่ขาดหายไปจะประมาณด้วย $\hat{Y}_2$

เมื่อ $\hat{Y}_2 = C X_2 \hat{\beta}^{LS}$

โดยที่ $C = \left[I - X_2 (X_1' X_1)^{-1} X_2' \right]^{-1}$

$$X_2 = \begin{bmatrix} 2,667.50 & 113,916.- \\ 2,762.50 & 181,340.- \\ 1,445.- & 231,460.- \end{bmatrix}$$

จะได้ $C = \begin{bmatrix} 1.0729 & 0.0956 & 0.0939 \\ 0.0956 & 1.1669 & 0.2332 \\ 0.0939 & 0.2332 & 1.4157 \end{bmatrix}$

$\therefore \quad \hat{Y}_2 = C X_2 \hat{\beta}^{LS}$

$$= \begin{bmatrix} 5,384.82 \\ 6,358.22 \\ 5,018.97 \end{bmatrix}$$

นำค่า $\hat{Y}_2$ ที่ได้นี้ แทนค่า Y_2 ที่ขาดหายไป แล้วหาค่า $\hat{\beta}^{(1)}$ โดยแทนค่า

$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$ และ $X = X_o$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะได้

$$\underline{\beta}^{(1)} = (X_o' X_o)^{-1} X_o' Y$$

$$= \begin{bmatrix} -2,760.48676 \\ 2.36220 \\ 0.01193 \end{bmatrix}$$

∴ สมการถดถอยที่ได้คือ

$$\hat{Y}_i^{(1)} = -2760.48676 + 2.36220X_{1i} + 0.01193X_{2i}$$

เราทดสอบสมมติฐาน H_o : $\underline{\beta} = 0$; H_A : $\underline{\beta} \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 19

ANOVA

SOV	df	SS	MS	F
Reg	3	2.6262306×10^7	8.7541020×10^6	28.70263**
Error	20	6.0999860×10^6	3.0499930×10^5	
Total	23	3.2362292×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.81151$$

3. <u>วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไปโดยวิธีอันดับศูนย์</u>

การวิเคราะห์การถดถอยโดยวิธีนี้ จะประมาณค่า Y ที่ขาดหายไปด้วย $\bar{Y}_{n_y}$ และประมาณค่า X_i ที่ขาดหายไปด้วย $\bar{X}_{n_{x_i}}$ แล้วนำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{n_{x_i}}$ แทนค่า Y_i และ X_i ที่ขาดหายไป ตามลำดับ แล้วจึงนำข้อมูลที่ได้ใหม่ไปหาสมการถดถอย

3.1 <u>การถดถอยเชิงเส้นอย่างง่าย</u> (Simple Linear Regression)

(1) เมื่อต้องการ $\hat{Y}_i^{(2)} = b_o^{(2)} + b_1^{(2)} X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_1} = 29, \quad n_y = 29$

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{1n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}}{n_{x_1}} = 2,214.35$$

ค่า Y_1 และ X_{1i} ที่ขาดหายไป จะประมาณด้วย $\bar{Y}_{n_y}$ และ $\bar{X}_{1n_{x_1}}$ ที่หาได้นี้ ตามลำดับ แล้วนำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{1n_{x_1}}$ แทนค่า Y_i และ X_{1i} ที่ขาดหายไปตามลำดับ นำข้อมูลที่ได้หาค่า $b_1^{(2)}$ และ $b_o^{(2)}$ โดยวิธีอันดับศูนย์ในบทที่ 2 จะได้

$$b_1^{(2)} = \frac{\sum_{i=1}^{n} X_{1i} Y_i - n\bar{X}_{1n_{x_1}} \bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{1i} - n\bar{X}^2_{1n_{x_1}}}$$

$$= 1.22078$$

$$b_o^{(2)} = \bar{Y}_{n_y} - b_1^{(2)}\bar{\bar{X}}_{1 \cdot n_{x_1}}$$

$$= 681.89380$$

สมการถดถอยที่ได้คือ $\hat{Y}_i^{(2)} = 681.89380 + 1.22078X_{1i}$

เราทดสอบสมมติฐาน H_o : $\beta_1 = 0$; H_A : $\beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 20

ANOVA

SOV	df	SS	MS	F
Reg	1	1.6029940×10^7	1.6029940×10^7	22.34175^{**}
Error	20	1.4349757×10^7	7.1748785×10^5	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.52765$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(2)} = b_o^{(2)} + b_2^{(2)}X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36$, $n_c = 22$, $n_{x_2} = 29$, $n_y = 29$

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{2n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}}{n_{x_2}} = 103,372.69$$

ค่า Y_i และ X_{2i} ที่ขาดหายไป จะประมาณด้วย $\bar{Y}_{n_y}$และ $\bar{X}_{2n_{x_2}}$ ที่หาได้นี้ตามลำดับ แล้วนำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{2n_{x2}}$ แทนค่า Y_i และ X_{2i} ที่ขาดหายไปตามลำดับนำข้อมูลที่ได้หาค่า $b_2^{(2)}$และ $b_o^{(2)}$ โดยวิธีอันดับศูนย์ในบทที่ 2 จะได้

$$b_2^{(2)} = \frac{\sum_{i=1}^{n} X_{2i} Y_i - n\bar{X}_{2n_{x_2}} \bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{2i}^2 - n\bar{X}_{2n_{x_2}}^2}$$

$$= -0.01100$$

$$b_o^{(2)} = \bar{Y}_{n_y} - b_2^{(2)} \bar{X}_{2n_{x_2}}$$

$$= 4,531.03913$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(2)} = 4,531.03913 - 0.01109X_2$

เราทดสอบสมมติฐาน $H_o : \beta_2 = 0$; $H_A : \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 21

ANOVA

SOV	df	SS	MS	F
Reg	1	1.0102620×10^7	1.0102620×10^7	9.96457^{**}
Error	20	2.0277077×10^7	1.0138538×10^6	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.33255$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(2)} = b_o^{(2)} + b_3^{(2)} X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_3} = 29, \quad n_y = 29$

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{3n_{x_3}} = \frac{\sum_{i=1}^{n_{x_3}} X_{3i}}{n_{x_3}} = 3,268.88$$

ค่า Y_i และ X_{3i} ที่ขาดหายไป จะประมาณด้วย $\bar{Y}_{n_y}$ และ $\bar{X}_{3n_{x_3}}$ ที่หาได้นี้ตามลำดับ แล้วนำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{3n_{x_3}}$ แทนค่า Y_i และ X_{3i} ที่ขาดหายไปตามลำดับ นำข้อมูลที่ได้มาหาค่า $b_3^{(2)}$ และ $b_o^{(2)}$ โดยวิธีอันดับศูนย์ในบทที่ 2 จะได้

$$b_3^{(2)} = \frac{\sum_{i=1}^{n} X_{3i}Y_i - n\bar{X}_{3n_{x_3}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{3i}^2 - n\bar{X}^2_{2n_{x_3}}}$$

$$= 0.82661$$

$$b_o^{(2)} = \bar{Y}_{n_y} - b_3^{(2)}\bar{X}_{3n_{x_3}}$$

$$= 683.05230$$

.·. สมการถดถอยที่ได้คือ $\hat{Y}_i^{(2)} = 683.05230 + 0.82661X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0 \; ; H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 22

ANOVA

SOV	df	SS	MS	F
Reg	1	1.5927862×10^7	1.5927862×10^7	22.04268^{**}
Error	20	1.4451835×10^7	7.2259175×10^5	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.52429$$

3.2 <u>การถดถอยเชิงเส้นพหุคูณ</u> (Multiple Linear Regression)

เมื่อต้องการหา $\hat{Y}_i^{(2)} = b_o^{(2)} + b_1^{(2)} X_{1i} + b_2^{(2)} X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 5 ในภาคผนวก

$n = 36, \quad n_c = 23, \quad n_{x_1} = 31, \quad n_{x_2} = 30, \quad n_y = 29$

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,600.88$$

$$\bar{X}_{1n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}}{n_{x_1}} = 2,159.14$$

$$\bar{X}_{2n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}}{n_{x_2}} = 112,242.93$$

วน log

ค่า Y_i, X_{1i} และ X_{2i} ที่ขาดหายไป จะประมาณด้วย $\bar{Y}_{n_y}$, $\bar{X}_{1n_{x_1}}$ และ $\bar{X}_{2n_{x_2}}$ ที่หาได้นี้ตามลำดับ แล้วนำค่า $\bar{Y}_{n_y}$, $\bar{X}_{1n_{x_1}}$ และ $\bar{X}_{2n_{x_2}}$ แทนค่า Y_i, X_{1i} และ X_{2i} ตามลำดับ นำข้อมูลที่ได้ หาค่า $b_o^{(2)}$, $b_1^{(2)}$ และ $b_2^{(2)}$ โดยวิธีอันดับศูนย์ในบทที่ 2 จะได้

$$\hat{\beta}^{(2)} = \begin{bmatrix} b_o^{(2)} \\ b_1^{(2)} \\ b_2^{(2)} \end{bmatrix} = (X'X)^{-1}X'Y = \begin{bmatrix} 2,189.18186 \\ 0.87802 \\ -0.00431 \end{bmatrix}$$

∴ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(2)} = 2,189.18186 + 0.87802X_{1i} - 0.00431X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta = 0$; $H_A \; \beta \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 23

ANOVA

SOV	df	SS	MS	F
Reg	3	1.6586960×10^7	5.5289860×10^6	15.35707^{**}
Error	20	7.2005712×10^6	3.6002856×10^5	
Total	23	2.3787531×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.69730$$

4. <u>วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไป โดยวิธีอันดับศูนย์คัดแปลง</u>

การประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอยวิธีนี้ จะประมาณค่าสังเกต Y ที่ขาดหายไปด้วย u และประมาณค่าสังเกต X ที่ขาดหายไปด้วย v นำค่า u

และ v แทนค่าสังเกต Y และ X ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้ไปหาสมการถดถอย

การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}_i^{(3)} = b_o^{(3)} + b_1^{(3)} X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22,\ \ n_{x_1} = 29,\ \ n_y = 29,\ \ m_{x_1} = 7,\ \ m_y = 7$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3{,}490.09$$

$$\bar{X}_{1c} = \frac{\sum_{i=1}^{n_c} X_{1i}}{n_c} = 2{,}160.11$$

$$\bar{X}_{1m_y} = \frac{\sum_{i=1}^{m_y} X_{1i}}{m_y} = 2{,}384.82$$

$$\bar{Y}_{m_{x_1}} = \frac{\sum_{i=1}^{m_{x_1}} Y_i}{m_{x_1}} = 3{,}055.29$$

โดยวิธีอันดับศูนย์ดัดแปลงในบทที่ 2 จะได้

$$b_1' = \frac{\sum_{i=1}^{n_c}(X_{1i}-\bar{X}_{1c})(Y_i-\bar{Y}_c)}{\sum_{i=1}^{n_c}(X_{1i}-\bar{X}_{1c})^2 + \sum_{i=1}^{m_y}(X_{1i}-\bar{X}_{1m_y})^2}$$

$$= 1.26$$

$$b_o' = \bar{Y}_c - b_1'\bar{X}_{1c}$$

$$= 760.71063$$

$$u_1 = \bar{Y}_c + b_1'(\bar{X}_{1m_y} - \bar{X}_{1c})$$

$$= 3,773.23$$

$$v_1 = \bar{X}_{1c} + \frac{\bar{Y}_{m_{x_1}} - \bar{Y}_c}{b_1'}$$

$$= 1,815.03$$

ประมาณค่า Y_i และ X_{1i} ที่ขาดหายไปด้วย u_1 และ v_1 ที่หาได้นี้ตามลำดับ นำค่า u และ v แทนค่าสังเกต Y_i และ X_{1i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาหาค่า $b_1^{(3)}$ และ $b_o^{(3)}$ โดยวิธีอันดับศูนย์ดัดแปลงในบทที่ 2 จะได้

$$b_1^{(3)} = \frac{\sum_{i=1}^{n} X_{1i}Y_i - n\bar{X}_{1n}\bar{Y}_n}{\sum_{i=1}^{n} X_{1i}^2 - n\bar{X}_{1n}^2}$$

$$= 1.26358$$

$$b_o^{(3)} = \bar{Y}_n - b_1^{(3)}\bar{X}_{1n}$$

$$= 760.71063$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(3)} = 760.71063 + 1.26358X_{1i}$

เราทดสอบสมมติฐาน H_o : $\beta_1 = 0$; H_A : $\beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 24

ANOVA

SOV	df	SS	MS	F
Reg	1	1.8608376×10^7	1.8608376×10^7	29.48883**
Error	20	1.2620625×10^7	6.3103125×10^5	
Total	21	3.1229001×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.59587$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(3)} = b_o^{(3)} + b_2^{(3)} X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22$, $n_{x_2} = 29$, $n_y = 29$, $m_{x_2} = 7$, $m_y = 7$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{2c} = \frac{\sum_{i=1}^{n_c} X_{2i}}{n_c} = 109,107.-$$

$$\bar{X}_{2m_y} = \frac{\sum_{i=1}^{m_y} X_{2i}}{m_y} = 85,350.57$$

$$\bar{Y}_{m_{x_2}} = \frac{\sum_{i=1}^{m_{x_2}} Y_i}{m_{x_2}} = 3,055.29$$

โดยวิธีอันดับศูนย์ดัดแปลงในบทที่ 2 จะได้

$$b_2' = \frac{\sum_{i=1}^{n_c}(X_{2i}-\bar{X}_{2c})(Y_i-\bar{Y}_c)}{\sum_{i=1}^{n_c}(X_{2i}-\bar{X}_{2c})^2 + \sum_{i=1}^{m_y}(X_{2i}-\bar{X}_{2m_y})^2}$$

$$= -0.012$$

$$b_o' = \bar{Y}_c - b_2\bar{X}_{2c}$$

$$= 4,799.37$$

$$u_2 = \bar{Y}_c + b_2'(\bar{X}_{2m_y} - \bar{X}_{1c})$$

$$= 3,775.17$$

$$v_2 = \bar{X}_{2c} + \frac{\bar{Y}_{m_{x_2}} - \bar{Y}_c}{b_2'}$$

$$= 145,340.33$$

ประมาณค่า Y_i และ X_{2i} ที่ขาดหายไปด้วย u_2 และ v_2 ที่หาได้นี้ตามลำดับ นำค่า u_2 และ v_2 แทนค่าสังเกต Y_i และ X_{2i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาหาค่า $b_2^{(3)}$ และ $b_o^{(3)}$ โดยวิธีอันดับศูนย์ดัดแปลงในบทที่ 2 จะได้

$$b_2^{(3)} = \frac{\sum_{i=1}^{n} X_{2i}Y_i - n\bar{X}_{2n}\bar{Y}_n}{\sum_{i=1}^{n} X_{2i}^2 - n\bar{X}_{2n}^2}$$

$$= -0.01172$$

$$b_o^{(3)} = \bar{Y}_n - b_2^{(3)}\bar{X}_{2n}$$

$$= 4,767.89247$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(3)} = 4,767.89247 - 0.01172X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta_2 = 0$; $H_A : \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 25

ANOVA

SOV	df	SS	MS	F
Reg	1	1.2652244×10^7	1.2652244×10^7	13.61534**
Error	20	1.8585269×10^7	9.2926345×10^5	
Total	21	3.1237513×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.40503$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(3)} = b_o^{(3)} + b_3^{(3)} X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n_c = 22, \quad n_{x_3} = 29, \quad n_y = 29, \quad m_{x_3} = 7, \quad m_y = 7$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{3c} = \frac{\sum_{i=1}^{n_c} X_{3i}}{n_c} = 3,183.86$$

$$\bar{X}_{3m_y} = \frac{\sum_{i=1}^{m_y} X_{3i}}{m_y} = 3,536.07$$

$$\bar{Y}_{m_{x_3}} = \frac{\sum_{i=1}^{m_{x_3}} Y_i}{m_{x_3}} = 3,055.29$$

โดยวิธีอันดับศูนย์ดัดแปลงในบทที่ 2 จะได้

$$b_3' = \frac{\sum_{i=1}^{n_c}(X_{3i}-\bar{X}_{3c})(Y_i-\bar{Y}_c)}{\sum_{i=1}^{n_c}(X_{3i}-\bar{X}_{3c})^2 + \sum_{i=1}^{m_y}(X_{3i}-\bar{X}_{3m_y})^2}$$

$$= 0.859$$

$$b_o' = \bar{Y}_c - b_3'\bar{X}_{3c}$$

$$= 755.15$$

$$u_3 = \bar{Y}_c + b_3'(\bar{X}_{2m_y} - \bar{X}_{3c})$$

$$= 3,792.64$$

$$v_2 = \bar{X}_{3c} + \frac{\hat{\bar{Y}}_{m_{x_3}} - \bar{Y}_c}{b_3'}$$

$$= 2,677.69$$

ประมาณค่า Y_i และ X_{3i} ที่ขาดหายไปด้วย u_3 และ v_3 ที่หาได้นี้ตามลำดับ นำค่า u_3 และ v_3 แทนค่าสังเกต Y_i และ X_{3i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาหาค่า $b_3^{(3)}$ และ $b_o^{(3)}$ โดยวิธีอันดับศูนย์ตัดแปลงในบทที่ 2 จะได้

$$b_3^{(3)} = \frac{\sum_{i=1}^{n} X_{3i}Y_i - n\bar{X}_{3n}\bar{Y}_n}{\sum_{i=1}^{n} X_{3i}^2 - n\bar{X}_{3n}^2}$$

$$= 0.85928$$

$$b_o^{(3)} = \bar{Y}_n - b_3^{(3)}\bar{X}_{3n}$$

$$= 754.26361$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(3)} = 754.26361 + 0.85928X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0$; $H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 26

ANOVA

SOV	df	SS	MS	F
Reg	1	1.8667055×10^7	1.8667055×10^7	29.51540**
Error	20	1.2649024×10^7	6.3245120×10^5	
Total	21	3.1316079×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.59609$$

5. <u>วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไปโดยวิธีถดถอยอันดับหนึ่ง</u>

การประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอยวิธีนี้ มีวิธีการเป็นขั้น ๆ ดังนี้

ก. คำนวณการถดถอยของ Y ในเทอม X_i จากข้อมูลซึ่งมีทั้งค่าสังเกต Y และ X_i โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 และประมาณค่า Y_i's ที่ขาดหายไป โดยใช้สมการถดถอยที่ได้นี้

ข. คำนวณการถดถอยของ X_i ในเทอม Y จากข้อมูลซึ่งมีทั้งค่าสังเกต Y_i และ X_i โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 และประมาณค่า X_i's ที่ขาดหายไป โดยใช้สมการถดถอยที่ได้นี้

ค. นำค่า Y_i's ที่ประมาณได้จาก (ก) และค่า X_i's ที่ประมาณได้จาก (ข) แทนค่า Y_i's และ X_i's ที่ขาดหายไปตามลำดับ

ง. นำข้อมูลที่ได้มาคำนวณหาสมการถดถอย

การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}_i^{(4)} = b_o^{(4)} + b_1^{(4)} X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$X_{1c} = \frac{\sum_{i=1}^{n_c} X_{1i}}{n_c} = 2,160.11$$

โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2

$$b_1^{LS} = \frac{\sum_{i=1}^{n_c} X_{1i} Y_i - n_c \bar{X}_{1c} \bar{Y}_c}{\sum_{i=1}^{n_c} X_{1i}^2 - n_c \bar{X}_{1c}^2} = 1.53750$$

$$b_o^{LS} = \bar{Y}_c - b_1^{LS} \bar{X}_{1c} = 168.90651$$

$\therefore$ จะได้สมการถดถอยคือ $\hat{Y}_i^{LS} = 168.90651 + 1.53750X_{1i}$

จากข้อมูลเดียวกันนี้ คำนวณการถดถอยของ X_1 บน Y โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 เช่นกัน จะได้

$$\hat{d}_1^{LS} = \frac{\sum_{i=1}^{n_c} X_{1i}Y_i - n_c\bar{X}_{1c}\bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c\bar{Y}_c^2}$$

$$= 0.63224$$

$$\hat{d}_o^{LS} = \bar{X}_{1c} - \hat{d}_1^{LS}\bar{Y}_c$$

$$= -46.45576$$

$\therefore$ จะได้สมการถดถอยคือ $\hat{X}_{1i}^{LS} = -46.45576 + 0.63224Y_i$ (3.5.2)

ประมาณค่า Y_i ที่ขาดหายไปโดยใช้สมการถดถอยจาก (3.5.1) ซึ่งจะได้ค่าดังนี้

เมื่อ $X_1 = 2,607.50$; $\hat{Y} = 168.90651 + 1.53750(2,607.50)$

$= 4,177.94$

$X_1 = 2,626.75$; $\hat{Y} = 168.90651 + 1.53750(2,626.75)$

$= 4,207.53$

$X_1 = 2,624.-$; $\hat{Y} = 168.90651 + 1.53750(2,624.-)$

$= 4,203.31$

X_1 2,512.50 ; $\hat{Y} = 168.90651 + 1.53750(2,512.50)$

$= 4,031.88$

$X_1 = 2,512.50$; $\hat{Y} = 168.90651 + 1.53750(2,512.50)$

$= 4,031.88$

$X_1 = 2,682.50$; $\hat{Y} = 168.90651 + 1.53750(2,682.50)$

$= 4,293.25$

$X_1 = 1,128.-$; $\hat{Y} = 168.90651 + 1.53750(1,128.-)$

$= 1,903.21$

ประมาณค่า X_{1i} ที่ขาดหายไปโดยใช้สมการถดถอยจาก(3.5.2) ซึ่งจะได้ค่าดังนี้

เมื่อ $Y = 4,271.-$; $\hat{X}_1 = -46.45576 + 0.63224(4,271.-)$

$= 2,653.84$

$Y = 4,475.-$; $\hat{X}_1 = -46.45576 + 0.63224(4,475.-)$

$= 2,782.82$

$Y = 4,073.50$; $\hat{X}_1 = -46.45576 + 0.63224(4,073.50)$

$= 2,528.97$

$Y = 1,737.50$; $\hat{X}_1 = -46.45576 + 0.63224(1,737.50)$

$= 1,052.06$

$Y = 1,846.-$; $\hat{X}_1 = -46.45576 + 0.63224(1,846.-)$

$= 1,120.66$

$Y = 2,342.-$; $\hat{X}_1 = -46.45576 + 0.63224(2,342.-)$

$= 1,434.25$

$$Y = 2,642.- \quad ; \quad \hat{X}_1 = -46.45576 + 0.63224(2,642.-)$$

$$= 1,623.92$$

นำค่า $\hat{Y}_i$ และ $\hat{X}_{1i}$ ที่ได้นี้ แทนค่า Y_i และ X_{1i} ที่ขาดหายไปตามลำดับ แล้วหาค่า $b_1^{(4)}$ และ $b_0^{(4)}$ โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้ค่า

$$b_1^{(4)} = \frac{\sum_{i=1}^{n} X_{1i}Y_i - n\bar{X}_{1n}\bar{Y}_n}{\sum_{i=1}^{n} X_{1i}^2 - n\bar{X}_{1n}^2}$$

และ $b_0^{(4)} = \bar{Y}_n - b_1^{(4)}\bar{X}_{1n} = 141.53856$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(4)} = 141.53856 + 1.54915X_{1i}$

เราทดสอบสมมติฐาน $H_o : \beta_1 = 0 ; H_A : \beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 27

ANOVA

SOV	df	SS	MS	F
Reg	1	3.5344071×10^7	3.5344071×10^7	119.59394^{**}
Error	20	5.9106791×10^5	2.9553395×10^5	
Total	21	3.5935139×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.98355$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(4)} = b_o^{(4)} + b_2^{(4)} X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$X_{2c} = \frac{\sum_{i=1}^{n_c} X_{2i}}{n_c} = 109,107.-$$

โดยวิธีถดถอยอันดับหนึ่ง ในบทที่ 2 จะได้

$$b_2^{LS} = \frac{\sum_{i=1}^{n_c} X_{2i} Y_i - n_c \bar{X}_{2c} \bar{Y}_c}{\sum_{i=1}^{n_c} X_{2i}^2 - n_c \bar{X}_{2c}^2} = -0.01361$$

$$b_o^{LS} = \bar{Y}_c - b_2^{LS} \bar{X}_{2c} = 4,974.5132$$

$\therefore$ จะได้สมการถดถอยคือ $\hat{Y}_i^{LS} = 4,974.5132 - 0.01361 X_{2i}$ (3.5.3)

จากข้อมูลเดียวกันนี้ คำนวณการถดถอยของ X_2 ในเทอม Y โดยวิธีถดถอยอันดับหนึ่ง ในบทที่ 2 เช่นกัน จะได้

$$d_2^{LS} = \frac{\sum_{i=1}^{n_c} X_{2i} Y_i - n_c \bar{X}_{2c} \bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c \bar{Y}_c^2} = -44.09832$$

$$d_o^{LS} = \bar{X}_{2c} - d_2^{\prime LS}\bar{Y}_c = 263,014.12843$$

จะได้สมการถดถอย คือ $\hat{X}_{2i}^{LS} = 263,014.12843 - 44.09832Y_i$ (3.5.4)

ประมาณค่า Y_i ที่ขาดหายไปโดยใช้สมการถดถอยจาก (3.5.3.) ซึ่งจะได้ค่าดังนี้

เมื่อ $X_2 = 83,111.-$; $\hat{Y} = 4,974.51320 - 0.01361(83,111.-)$

$= 3,843.37$

$X_2 = 47,457.-$; $\hat{Y} = 4,974.51320 - 0.01361(47,457.-)$

$= 4,328.62$

$X_2 = 93,625.-$; $\hat{Y} = 4,974.51320 - 0.01361(93,625.-)$

$= 3,700.28$

$X_2 = 87,906.-$; $\hat{Y} = 4,974.51320 - 0.01361(87,906.-)$

$= 3,700.28$

$X_2 = 79,629.-$; $\hat{Y} = 4,974.51320 - 0.01361(79,629.-)$

$= 3,890.76$

$X_2 = 35,423.-$; $\hat{Y} = 4,974.51320 - 0.01361(35,423.-)$

$= 4,492.41$

$X_2 = 170,303.-$; $\hat{Y} = 4,974.51320 - 0.01361(170,303.-)$

$= 2,656.69$

ประมาณค่า X_{2i} ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้จาก (3.5.4) ซึ่งจะได้ค่าดังนี้

เมื่อ $Y = 4,271.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(4,271.-)$

$= 74,670.20$

$Y = 4,475.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(4,475.-)$

$= 65,674.15$

$Y = 4,073.50$; $\hat{X}_2 = 263,014.12843 - 44.09832(4,073.50)$

$= 83,379.62$

$Y = 1,737.50$; $\hat{X}_2 = 263,014.12843 - 44.09832(1,737.50)$

$= 186,393.30$

$Y = 1,846.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(1,846.-)$

$= 181,608.63$

$Y = 2,342.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(2,342.-)$

$= 159,735.86$

$Y = 2,642.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(2,642.-)$

$= 146,506.37$

นำค่า $\hat{Y}_i$'s และ $\hat{X}_{2i}$'s ที่ได้นี้ แทนค่า Y_i และ X_{2i} ที่ขาดหายไปตามลำดับ แล้วหาค่า $b_2^{(4)}$ และ $b_o^{(4)}$ โดยวิธีถดถอยอันดับหนึ่ง ในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้ค่า

$$b_2^{(4)} = \frac{\sum_{i=1}^{n} X_{2i}Y_i - n\bar{X}_{2n}\bar{Y}_n}{\sum_{i=1}^{n} X_{2i}^2 - n\bar{X}_{2n}^2} = -0.0153$$

$$b_o^{(4)} = \bar{Y}_n - b_2^{(4)}\bar{X}_{2n} = 5,123.77923$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(4)} = 5,123.77923 - 0.0153X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta_2 = 0 ; H_A : \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 28

ANOVA

SOV	df	SS	MS	F
Reg	1	2.3885758×10^7	2.3885758×10^7	49.70447**
Error	20	9.6111107×10^6	4.8055553×10^5	
Total	21	3.3496868×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.71307$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(4)} = b_o^{(4)} + b_3^{(4)}X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{3c} = \frac{\sum_{i=1}^{n_c} X_{3i}}{n_c} = 3,183.8636$$

โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 จะได้

$$b_3^{LS} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n_c\bar{X}_{3c}\bar{Y}_c}{\sum_{i=1}^{n_c} X_{3i}^2 - n_c\bar{X}_{3c}^2} = 1.07162$$

$$b_o^{LS} = \bar{Y}_c - b_3^{LS}\bar{X}_{3c} = 78.20098$$

จะได้สมการถดถอย คือ $\hat{Y}_i^{LS} = 78.20098 + 1.07162X_{3i}$ (3.5.5)

จากข้อมูลเดียวกันนี้ คำนวณการถดถอยของ X_3 on Y โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 เช่นกัน จะได้

$$d_3^{LS} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n_c\bar{X}_{3c}\bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c\bar{Y}_c^2} = 0.92838$$

$$d_o^{LS} = \bar{X}_{3c} - d_3^{LS}\bar{Y}_c = -56.25631$$

จะได้สมการถดถอย คือ $\hat{X}_{3i} = -56.25631 - 0.92838Y_i$ (3.5.6)

ประมาณค่า Y_i ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้จาก (3.5.5) ซึ่งจะได้ค่า ดังนี้

เมื่อ $X_3 = 3,692.-$; $\hat{Y} = 78.20098 + 1.07162(3,692.-)$

$= 4,034.62$

$X_3 = 3,694.50$; $\hat{Y} = 78.20098 + 1.07162(3,694.50)$

$= 4,037.30$

$X_3 = 3,960.-$; $\hat{Y} = 78.20098 + 1.07162(3,960.-)$

$= 4,321.82$

$X_3 = 4,062.50$; $\hat{Y} = 78.20098 + 1.07162(4,062.50)$

$= 4,431.66$

$X_3 = 3,962.50$; $\hat{Y} = 78.20098 + 1.07162(3,962.50)$

$= 4,324.50$

$X_3 = 3,779.00$; $\hat{Y} = 78.20098 + 1.07162(3,779.-)$

$= 4,127.85$

$X_3 = 1,602.00$; $\hat{Y} = 78.20098 + 1.07162(1,602.-)$

$= 1,794.94$

ประมาณค่า X_{3i} ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้จาก (3.5.6) ซึ่งจะได้ค่า ดังนี้

เมื่อ $Y_i = 4,271.-$; $\hat{X}_3 = -56.25631 + 0.92838(4,271.-)$

$= 3,908.85$

$Y = 4,475.-$; $\hat{X}_3 = -56.25631 + 0.92838(4,475.-)$

$= 4,098.24$

$Y = 4,073.50$; $\hat{X}_3 = -56.25631 + 0.92838(4,073.50)$

$= 3,725.50$

$Y = 1,737.50$; $\hat{X}_3 = -56.25631 + 0.92838(1,737.50)$

$= 1,556.80$

$Y = 1,846.-$; $\hat{X}_3 = -56.25631 + 0.92838(1,846.-)$

$= 1,657.53$

$Y = 2,342.-$; $\hat{X}_3 = -56.25631 + 0.92838(2,342.-)$

$= 2,118.01$

$Y = 2,642.-$ $\hat{X}_3 = -56.25631 + 0.92838(2,642.-)$

$= 2,396.52$

นำค่า $\hat{Y}_i$'s และ $\hat{X}_{3i}$'s ที่ได้นี้ แทนค่า Y_i และ X_{3i} ที่ขาดหายไปตามลำดับ แล้วหาค่า $b_3^{(4)}$ และ $b_o^{(4)}$ โดยวิธีถดถอยอันดับหนึ่งในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้ค่า

$$b_3^{(4)} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n\bar{X}_{3n}\bar{Y}_n}{\sum_{i=1}^{n} X_{3i}^2 - n\bar{X}_{3n}^2} = 1.07265$$

$$b_o^{(4)} = \bar{Y}_n - b_3^{(4)} \bar{X}_{3n} = 73.90506$$

∴ สมการถดถอยที่ได้ คือ $\hat{Y}_i^{(4)} = 73.90506 + 1.07265X_{3i}$

เราทดสอบสมมติฐาน $H_o: \beta_3 = 0$; $H_A: \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 29

ANOVA

SOV	df.	SS	MS	F
Reg	1	3.6737170×10^7	3.6737170×10^7	678.60310**
Error	20	1.0827292×10^5	5.4136460×10^3	
Total	21	3.6845443×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.99706$$

6. <u>วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไป โดยวิธีถดถอยสองชั้น</u>

การประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอยวิธีนี้ มีวิธีการดังนี้

ก. คำนวณการถดถอยของ X_i ในเทอม Y จากข้อมูลที่มีทั้งค่าสังเกต Y และ X_i โดยวิธีถดถอยสองชั้นในบทที่ 2 และประมาณค่า X_i's ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้

ข. นำค่า X_i's ที่ประมาณได้จาก (ก) แทนค่าสังเกต X_i's ที่ขาดหายไป

ค. นำข้อมูลที่ได้ซึ่งมีทั้งค่า X_i และ Y_i มาคำนวณหาสมการถดถอย

การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}_i^{(5)} + b_o^{(5)} + b_1^{(5)}X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_1} = 29$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{1c} = \frac{\sum_{i=1}^{n_c} X_{1i}}{n_c} = 2,160.11$$

โดยวิธีถดถอยสองชั้น ในบทที่ 2 จะได้

$$d_1^{LS} = \frac{\sum_{i=1}^{n_c} X_{1i}Y_i - n_c\bar{X}_{1c}\bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c\bar{Y}_c^2} = 0.63224$$

$$d_o^{LS} = \bar{X}_{1c} - d_1^{LS}\bar{Y}_c = -46.45576$$

สมการถดถอยที่ได้คือ $\hat{X}_{1i}^{LS} = -46.45576 + 0.63224Y_i$

ประมาณค่า X_{1i} ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้ ซึ่งจะได้ค่าดังนี้

เมื่อ $Y = 4,271.-$; $\hat{X}_1 = -46.45576 + 0.63224(4,271.-)$

$= 2,653.84$

$Y = 4,475.-$; $\hat{X}_1 = -46.45576 + 0.63224(4,475.-)$

$= 2,782.82$

$Y = 4,073.50$; $\hat{X}_1 = -46.45576 + 0.63224(4,073.50)$

$= 2,528.97$

$Y = 1,737.50$; $\hat{X}_1 = -46.45576 + 0.63224(1,737.50)$

$= 1,052.06$

$Y = 1,846.-$; $\hat{X}_1 = -46.45576 + 0.63224(1,846.-)$

$= 1,120.66$

$Y = 2,342.-$; $\hat{X}_1 = -46.45576 + 0.63224(2,342.-)$

$= 1,434.25$

$Y = 2,642.-$; $\hat{X}_1 = -46.45576 + 0.63224(2,642.-)$

$= 1,623.92$

นำค่า $\hat{X}_{1i}$'s ที่ได้นี้ แทนค่า X_{1i}'s ที่ขาดหายไป แล้วหาค่า $b_1^{(5)}$ และ $b_o^{(5)}$ โดยวิธีถดถอยสองชั้น ในบทที่ 2 n_c จะเพิ่มเท่ากับ $n_{x_1} = 29$ และจะได้ค่า

$$b_1^{(5)} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}Y_i - n_{x_1}\bar{X}_{1n_{x_1}}\bar{Y}_{n_{x_1}}}{\sum_{i=1}^{n_{x_1}} X_{1i}^2 - n_{x_1}\bar{X}_{1n_{x_1}}^2} = 1.55097$$

$$b_o^{(5)} = \bar{Y}_{n_{x_1}} - b_1^{(5)} \bar{X}_{1n_{x_1}} = 137,81948$$

∴ สมการถดถอยที่ได้ คือ $\hat{Y}_i^{(5)} = 137.81948 + 1.55097X_{1i}$

เราทดสอบสมมติฐาน $H_o: \beta_1 = 0$; $H_A: \beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 30

ANOVA

SOV	df	SS	MS	F
Reg	1	2.9788925×10^7	2.9788925×10^7	1008.47477**
Error	20	5.9077185×10^7	2.9538592×10^4	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.98055$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(5)} = b_o^{(5)} + b_2^{(5)} X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36$, $n_c = 22$, $n_{x_2} = 29$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{2c} = \frac{\sum_{i=1}^{n_c} X_{2i}}{n_c} = 109,107.-$$

โดยวิธีถดถอยสองชั้น ในบทที่ 2 จะได้

$$d_2^{LS} = \frac{\sum_{i=1}^{n_c} X_{2i}Y_i - n_c\bar{X}_{2c}\bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c\bar{Y}_c^2} = -44.09832$$

$$d_o^{LS} = \bar{X}_{2c} - d_2^{LS}\bar{Y}_c = 263,014.12843$$

สมการถดถอยที่ได้คือ $\hat{X}_{2i} = 263,014.12843 - 44.09832Y_i$

ประมาณค่า X_{2i}'s ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้ ซึ่งจะได้ค่าดังนี้

เมื่อ $Y = 4,271.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(4,271.-)$

$= 74,670.20$

$Y = 4,475.-$ $\hat{X}_2 = 263,014.12843 - 44.09832(4,475.-)$

$= 65,674.15$

$Y = 4,073.50$; $\hat{X}_2 = 263,014.12843 - 44.09832(4,073.50)$

$= 83,379.62$

$Y = 1,737.50$; $\hat{X}_2 = 263,014.12843 - 44.09832(1,737.50)$

$= 186,393.30$

$Y = 1,846.-$; $\hat{X}_2 = 263,014.12843 - 44.09832(1,846.-)$

$= 181,608.63$

$Y = 2,342.- \; ; \quad \hat{X}_2 = 263,014.12843 - 44.09832(2,342.-)$

$= 159,735.86$

$Y = 2,642.- \; ; \quad \hat{X}_2 = 263,014.12843 - 44.09832(2,642.-)$

$= 146,506.37$

นำค่า $\hat{X}_{2i}$'s ที่ได้นี้ แทนค่า X_{2i}'s ที่ขาดหายไป แล้วหาค่า $b_2^{(5)}$ และ $b_o^{(5)}$ โดยวิธีถดถอยสองชั้นในบทที่ 2 n_c จะเพิ่มเท่ากับ $n_{x_2} = 29$ และจะได้ค่า

$$b_2^{(5)} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}Y_i - n_{x_2}\bar{X}_{2n_{x_2}}\bar{Y}_{n_{x_2}}}{\sum_{i=1}^{n_{x_2}} X_{2i}^2 - n_{x_2}\bar{X}_{2n_{x_2}}^2} = -0.01553$$

$$b_o^{(5)} = \bar{Y}_{n_{x_2}} - b_2^{(5)}\bar{X}_{2n_{x_2}} = 5,151.45552$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i(5) = 5,151.45552 - 0.01553X_{2i}$

เราทดสอบสมมติฐาน $H_o: \beta_2 = 0; \; H_A: \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 31

ANOVA

SOV	df	SS	MS	F
Reg	1	2.0805548×10^7	2.0805548×10^7	43.46192**
Error	20	9.5741489×10^6	4.7870744×10^5	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.68485$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(5)} = b_o^{(5)} + b_3^{(5)}X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_3} = 29$

$$\bar{Y}_c = \frac{\sum_{i=1}^{n_c} Y_i}{n_c} = 3,490.09$$

$$\bar{X}_{3c} = \frac{\sum_{i=1}^{n_c} X_{3i}}{n_c} = 3,183.8636$$

โดยวิธีถดถอยสองชั้นในบทที่ 2 จะได้

$$d_3^{LS} = \frac{\sum_{i=1}^{n_c} X_{3i}Y_i - n_c\bar{X}_{3c}\bar{Y}_c}{\sum_{i=1}^{n_c} Y_i^2 - n_c\bar{Y}_c^2} = 0.92838$$

$$d_o^{LS} = \bar{X}_{3c} - d_3^{LS}\bar{Y}_c = 56.25631$$

สมการถดถอยที่ได้คือ $\hat{X}_{3i} = - 56.25631 + 0.92838Y_i$

ประมาณค่า X_{3i}'s ที่ขาดหายไปโดยใช้สมการถดถอยที่ได้นี้ ซึ่งจะได้ค่าดังนี้

เมื่อ $Y = 4,271.-$; $\hat{X}_3 = - 56.25631 + 0.92838(4,271.-)$

$= 3,908.85$

$Y = 4,475.-$; $\hat{X}_3 = - 56.25631 + 0.92838(4,475.-)$

$= 4,098.24$

$Y = 4,073.50$; $\hat{X}_3 = - 56.25631 + 0.92838(4,073.50)$

$= 3,725.50$

$Y = 1,737.50$; $\hat{X}_3 = - 56.25631 + 0.92838(1,737.50)$

$= 1,556.80$

$Y = 1,846.-$; $\hat{X}_3 = - 56.25631 + 0.92838(1,846.-)$

$= 1,657.53$

$Y = 2,342.-$; $\hat{X}_3 = - 56.25631 + 0.92838(2,342.-)$

$= 2,118.01$

$Y = 2,642.-$; $\hat{X}_3 = - 56.25631 + 0.92838(2,642.-)$

$= 2,396.52$

นำค่า $\hat{X}_{3i}$'s ที่ได้นี้ แทนค่า X_{3i}'s ที่ขาดหายไป แล้วหาค่า $b_3^{(5)}$ และ $b_o^{(5)}$ โดยวิธีถดถอยสองชั้น ในบทที่ 2 n_c จะเพิ่มเท่ากับ $n_{x_3} = 29$ และจะได้ค่า

$$b_3^{(5)} = \frac{\sum_{i=1}^{n_{x_3}} X_{3i}Y_i - n_{x_3}\bar{X}_{3n_{x_3}}\bar{Y}_{n_{x_3}}}{\sum_{i=1}^{n_{x_3}} X_{3i}^2 - n_{x_3}\bar{X}_{3n_{x_3}}^2} = 1.07333$$

$$b_o^{(5)} = \bar{Y}_{n_{x_3}} - b_3^{(5)}\bar{X}_{3n_{x_3}} = 72.39274$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(5)} = 72.39274 + 1.07333X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0$; $H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 32

ANOVA

SOV	df	SS	MS	F
Reg	1	3.0271896×10^7	3.0271896×10^7	5616.27962**
Error	20	1.0780053×10^5	5.3900265×10^3	
Total	21	3.0379697×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.99645$$

7. <u>วิเคราะห์ข้อมูลเมื่อมีการประมาณค่าของตัวแปรที่ขาดหายไปโดยวิธีผสม</u>

การประมาณค่าที่ขาดหายไปในการวิเคราะห์การถดถอยโดยวิธีนี้ มีวิธีการหลายขั้นตอนด้วยกัน คือ

7.1 <u>การถดถอยเชิงเส้นอย่างง่าย</u> (Simple Linear Regression)

(1) เมื่อต้องการหา $\hat{Y}_i^{(6)} = b_o^{(6)} + b_1^{(6)}X_{1i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_1} = 29, \quad n_y = 29$

นำข้อมูลที่รวบรวมได้มาหาค่าเฉลี่ยของค่าสังเกต Y_1 และ X_{1i} เท่าที่ข้อมูล Y และ X_1 มีอยู่ จะได้

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{1n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}}{n_{x_1}} = 2,214.35$$

นำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{1n_{x_1}}$ ซึ่งหาได้นี้ แทนค่า Y_i's และ X_{1i}'s ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาคำนวณหาสมการถดถอย

$$\hat{Y}_i^o = b_o^o + b_1^o X_{1i}$$

และ $$\hat{X}_{1i}^o = d_o^o + d_1^o Y_i$$

โดยวิธีผสมในบทที่ 2 จะได้

$$b_1^o = \frac{\sum_{i=1}^{n} X_{1i} Y_i - n\bar{X}_{1n_{x_1}} \bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{1i}^2 - n\bar{X}_{1n_{x_1}}^2}$$

$$= 1.22078$$

$$b_o^o = \bar{Y}_{n_y} - b_1^o \bar{X}_{n_{x_1}}$$

$$= 681.89380$$

สมการถดถอยที่ได้คือ $\hat{Y}_i^o = 681.89380 + 1.22078X_{1i}$ (3.7.1)

$$d_1^o = \frac{\sum_{i=1}^{n} X_{1i}Y_i - n\bar{X}_{1n_{x_1}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}_{n_y}^2}$$

$$= 0.43223$$

$$d_o^o = \bar{X}_{1n_{x_1}} - d_1^o \bar{Y}_{n_y}$$

$$= 751.21116$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{X}_{1i}^o = 751.21116 + 0.43223Y_i$ (3.7.2)

หาค่า $\hat{Y}_i$ และ $\hat{X}_{1i}$ ที่ขาดหายไปใหม่โดยใช้สมการถดถอย (3.7.1) และ (3.7.2) ตามลำดับ ซึ่งจะได้ค่าดังนี้

เมื่อ $X_1 = 2,607.50$; $\hat{Y}^o = 681.89380 + 1.22078(2,607.50)$

$= 3,865.08$

$X_1 = 2,626.75$; $\hat{Y}^o = 681.89380 + 1.22078(2,626.75)$

$= 3,888.58$

$X_1 = 2,624.- \; ; \quad \hat{Y}^o = 681.89380 + 1.22078(2,624.-)$

$= 3,885.22$

$X_1 = 2,512.50 \; ; \quad \hat{Y}^o = 681.89380 + 1.22078(2,512.50)$

$= 3,749.10$

$X_1 = 2,512.50 \; ; \quad \hat{Y}^o = 681.89380 + 1.22078(2,512.50)$

$= 3,749.10$

$X_1 = 2,682.50 \; ; \quad \hat{Y}^o = 681.89380 + 1.22078(2,682.50)$

$= 3,956.64$

$X_1 = 1,128.- \; ; \quad \hat{Y}^o = 681.89380 + 1.22078(1,128.-)$

$= 2,058.93$

$Y = 4,271.- \; ; \quad \hat{X}_1^o = 751.21116 + 0.43223(4,271.-)$

$= 2,597.27$

$Y = 4,475.- \; ; \quad \hat{X}_1^o = 751.21116 + 0.43223(4,475.-)$

$= 2,685.44$

$Y = 4,073.50 \; ; \quad \hat{X}_1^o = 751.21116 + 0.43223(4,073.50)$

$= 2,511.90$

$Y = 1,737.50 \; ; \quad \hat{X}_1^o = 751.21116 + 0.43223(1,737.50)$

$= 1,502.21$

$Y = 1,846.- \; ; \quad \hat{X}_1^o = 751.21116 + 0.43223(1,846.-)$

$= 1,549.11$

$Y = 2,342.-$; $\hat{X}_1^o = 751.21116 + 0.43223(2,342.-)$

$= 1,763.49$

$Y = 2,642.-$; $\hat{X}_1^o = 751.21116 + 0.43223(2,642.-)$

$= 1,893.16$

ค่า Y_i และ X_{1i} ที่ขาดหายไป จะประมาณโดยการนำค่า $\hat{Y}_i^o$ และ $\hat{X}_{1i}^o$ ที่คำนวณได้นี้ มาหาค่าเฉลี่ยกับ $\bar{Y}_{n_y}$ และ $\bar{X}_{1n_{x_1}}$ ตามลำดับ ซึ่งจะได้ค่าดังนี้

ให้ $$\hat{Y}_i' = \frac{Y_{n_y} + \hat{Y}_i^o}{2}$$

$$\hat{X}_{1i}' = \frac{X_{1n_{x_1}} + \hat{X}_{1i}^o}{2}$$

เมื่อ $\hat{Y}^o = 3,865.08$; $\hat{Y}' = \frac{3,385.14 + 3,865.08}{2} = 3,625.11$

$\hat{Y}^o = 3,888.58$; $\hat{Y}' = \frac{3,385.14 + 3,888.58}{2} = 3,636.86$

$\hat{Y}^o = 3,385.22$; $\hat{Y}' = \frac{3,385.14 + 3,385.22}{2} = 3,635.18$

$\hat{Y}^o = 3,749.10$; $\hat{Y}' = \frac{3,385.14 + 3,749.10}{2} = 3,567.12$

$\hat{Y}^o = 3,749.10$; $\hat{Y}' = \frac{3,385.14 + 3,749.10}{2} = 3,567.12$

$\hat{Y}^o = 3,956.64$; $\hat{Y}' = \frac{3,385.14 + 3,956.64}{2} = 3,670.89$

$\hat{Y}^o = 2,058.93$; $\hat{Y}' = \frac{3,385.14 + 2,058.93}{2} = 2,722.04$

$$\hat{X}_1^o = 2{,}597.27 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 2{,}597.27}{2} = 2{,}405.81$$

$$\hat{X}_1^o = 2{,}685.44 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 2{,}685.44}{2} = 2{,}449.90$$

$$\hat{X}_1^o = 2{,}511.90 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 2{,}511.90}{2} = 2{,}363.13$$

$$\hat{X}_1^o = 1{,}502.21 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 1{,}502.21}{2} = 1{,}858.28$$

$$\hat{X}_1^o = 1{,}549.11 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 1{,}549.11}{2} = 1{,}881.73$$

$$\hat{X}_1^o = 1{,}763.49 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 1{,}763.49}{2} = 1{,}988.92$$

$$\hat{X}_1^o = 1{,}893.16 \; ; \quad \hat{X}_1' = \frac{2{,}214.35 + 1{,}893.16}{2} = 2{,}053.76$$

นำค่า $\hat{Y}_i'$ และ $\hat{X}_{1i}'$ ซึ่งคำนวณได้ แทนค่า Y_i และ X_{1i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้ทั้งหมด หาสมการถดถอยโดยวิธีผสมในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้

$$b_1^{(6)} = \frac{\sum_{i=1}^{n} X_{1i} Y_i - n\bar{X}_{1n}\bar{Y}_n}{\sum_{1=1}^{n} X_{1i}^2 - n\bar{X}_{1n}} = 1.46599$$

$$b_o^{(6)} = \bar{Y}_n - b_1^{(6)} \bar{X}_{1n} = 179.47808$$

$\therefore$ สมการถดถอยที่ได้ คือ $\hat{Y}_i^{(6)} = 179.47808 + 1.46599X_{1i}$

เราทดสอบสมมติฐาน $H_o : \beta_1 = 0$; $H_A : \beta_1 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 33

ANOVA

SOV	df	SS	MS	F
Reg	1	2.4021702×10^{7}	2.4021702×10^{7}	67.53057^{**}
Error	20	7.1143186×10^{6}	3.5571593×10^{5}	
Total	21	3.1136021×10^{7}		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.77151$$

(2) เมื่อต้องการหา $\hat{Y}_i^{(6)} = b_o^{(6)} = b_2^{(6)} X_{2i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$$n = 36, \quad n_c = 22, \quad n_{x_2} = 29, \quad n_y = 29$$

นำข้อมูลที่รวบรวมได้ มาหาค่าเฉลี่ยของค่าสังเกต Y_i และ X_{2i} เท่าที่ข้อมูล Y และ X_{2i} มีอยู่ จะได้

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{2n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}}{n_{x_2}} = 103,372.69$$

นำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{2n_{x_2}}$ ซึ่งหาได้นี้ แทนค่า Y_i's และ X_{2i}'s ที่ขาดหายไป ตามลำดับ แล้วนำข้อมูลที่ได้มาคำนวณหาสมการถดถอย

$$\hat{Y}_i^o = b_o^o + b_2^o X_{2i}$$

$$\text{และ } \hat{X}_{2i}^o = d_o^o + d_2^o Y_i$$

โดยวิธีผสมในบทที่ 2 จะได้

$$b_2^o = \frac{\sum_{i=1}^{n} X_{2i}Y_i - n\bar{X}_{2n_{x_2}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{2i}^2 - n\bar{X}_{2n_{x_2}}^2} = -0.01109$$

$$b_o^o = \bar{Y}_{n_y} - b_o^o\bar{X}_{2n_{x_2}} = 4,531.03913$$

∴ สมการถดถอยที่ได้คือ $\hat{Y}_i^o = 4,531.03913 - 0.01109\, X_{2i}$ (3.7.3)

$$d_2^o = \frac{\sum_{i=1}^{n} X_{2i}Y_i - n\bar{X}_{2n_{x_2}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}_{n_y}^2} = -29.99918$$

$$d_o^o = \bar{X}_{2n_{x_2}} - d_2^o\bar{Y}_{n_y} = 204,924.05854$$

.'. สมการถดถอยที่ได้คือ $X_{2i}^o = 204,924.05854 - 29.99918Y_i$ (3.7.4)

หาค่า $\hat{Y}_i$ และ $\hat{X}_{2i}$ ที่ขาดหายไป โดยใช้สมการถดถอย (3.7.3) และ (3.7.4) ตามลำดับ ซึ่งจะได้ค่าดังนี้

เมื่อ $X_2 = 83,111.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(83,111.-)$

$= 3,609.34$

$X_2 = 47,457.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(47,457.-)$

$= 4,004.74$

$X_2 = 93,625.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(93,625.-)$

$= 3,492.74$

$X_2 = 87,906.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(87,906.-)$

$= 3,556.16$

$X_2 = 79,629.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(79,629.-)$

$= 3,647.95$

X_2 $35,423.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(35,423.-)$

$= 4,138.20$

X_2 $170,303.-$; $\hat{Y}^o = 4,531.03913 - 0.01109(170,303.-)$

$= 2,642.38$

$Y = 4,271.-$; $\hat{X}_2^o = 204,924.05854 - 29.99918(4,271.-)$

$= 76,797.56$

$Y = 4,475.-$; $\hat{X}_2^o = 204,924.05854 - 29.99918(4,475.-)$

$= 70,677.73$

$Y = 4,073.50$; $\hat{X}_2^o = 204,924.05854 - 29.99918(4,073.50)$

$= 82,722.40$

$Y = 1,737.50$; $\hat{X}_2^o = 204,924.05854 - 29.99918(1,737.50)$

$= 152,800.48$

$Y = 1,846.-$; $\hat{X}_2^o = 204,924.05854 - 29.99918(1,846.-)$

$= 149,545.57$

$Y = 2,342.-$; $\hat{X}_2^o = 204,924.05854 - 29.99918(2,342.-)$

$= 134,665.98$

$Y = 2,642.-$; $\hat{X}_2^o = 204,924.05854 - 29.99918(2,642.-)$

$= 125,666.22$

ค่า Y_i และ X_{2i} ที่ขาดหายไป จะประมาณค่าโดยการนำค่า $\hat{Y}_i^o$ และ $\hat{X}_{2i}^o$ ที่คำนวณได้นี้ มาหาค่าเฉลี่ยกับ $\bar{Y}_{n_y}$ และ $\bar{X}_{2n_{x_2}}$ ตามลำดับ ซึ่งจะได้ค่าดังนี้

ให้ $$\hat{Y}_i' = \frac{\bar{Y}_{n_y} + \hat{Y}_i^o}{2}$$

$$\hat{X}_{2i}' = \frac{\bar{X}_{2n_{x_2}} + \hat{X}_{2i}^o}{2}$$

เมื่อ $\hat{Y}^{o} = 3,609.34$; $\quad \hat{Y}' = \dfrac{3,385.14 + 3,609.34}{2} = 3,497.24$

$\hat{Y}^{o} = 4,004.74$; $\quad \hat{Y}' = \dfrac{3,385.14 + 4,004.74}{2} = 3,694.94$

$\hat{Y}^{o} = 3,492.74$; $\quad \hat{Y}' = \dfrac{3,385.14 + 3,492.74}{2} = 3,438.94$

$\hat{Y}^{o} = 3,556.16$; $\quad \hat{Y}' = \dfrac{3,385.14 + 3,556.16}{2} = 3,470.65$

$\hat{Y}^{o} = 3,647.95$; $\quad \hat{Y}' = \dfrac{3,385.14 + 3,647.95}{2} = 3,516.55$

$\hat{Y}^{o} = 4,138.20$; $\quad \hat{Y}' = \dfrac{3,385.14 + 4,138.20}{2} = 3,761.67$

$\hat{Y}^{o} = 2,642.38$; $\quad \hat{Y}' = \dfrac{3,385.14 + 2,642.38}{2} = 3,013.76$

$\hat{X}_2^{o} = 76,797.56$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 76,797.56}{2} = 90,085.13$

$\hat{X}_2^{o} = 70,677.73$; $\quad \hat{X}_2' = \dfrac{103,372.69 \quad 76,677.73}{2} = 87,025.21$

$\hat{X}_2^{o} = 82,722.40$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 82,722.40}{2} = 93,047.55$

$\hat{X}_2^{o} = 152,800.48$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 152,800.48}{2} = 128,086.59$

$\hat{X}_2^{o} = 149,545.57$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 149,545.57}{2} = 126,459.13$

$\hat{X}_2^{o} = 134,665.98$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 134,665.98}{2} = 119,019.33$

$\hat{X}_2^{o} = 125,666.22$; $\quad \hat{X}_2' = \dfrac{103,372.69 + 125,666.22}{2} = 114,519.46$

นำค่า $\hat{Y}_i'$ และ $\hat{X}_{2i}'$ ซึ่งคำนวณได้ แทนค่า Y_i และ X_{2i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้ทั้งหมดหาสมการถดถอยโดยวิธีผสมในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้

$$b_2^{(6)} = \frac{\sum_{i=1}^{n} X_{2i}Y_i - n\bar{X}_{2n}\bar{Y}n}{\sum_{i=1}^{n} X_{2i}^2 - n\bar{X}_{2n}^2} = -0.01335$$

$$b_o^{(6)} = \bar{Y}_n - b_2^{(6)}\bar{X}_{2n} = 4,797.29284$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}_i^{(6)} = 4,797.29284 - 0.01335X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta_2 = 0$; $H_A : \beta_2 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 34

ANOVA

SOV	df	SS	MS	F
Reg	1	1.5012273×10^7	1.5012273×10^7	19.03948**
Error	20	1.5769617×10^7	7.8848085×10^5	
Total	21	3.0781890×10^7		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.48770$$

(3) เมื่อต้องการหา $\hat{Y}_i^{(6)} = b_o^{(6)} + b_3^{(6)}X_{3i}$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 4 ในภาคผนวก

$n = 36, \quad n_c = 22, \quad n_{x_3} = 29, \quad n_y = 29$

นำข้อมูลที่รวบรวมได้มาหาค่าเฉลี่ยของค่าสังเกต Y_i และ X_{3i} เท่าที่ข้อมูล Y และ X_3 มีอยู่ จะได้

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,385.14$$

$$\bar{X}_{3n_{x_3}} = \frac{\sum_{i=1}^{n_{x_3}} X_{3i}}{n_{x_3}} = 3,268.88$$

นำค่า $\bar{Y}_{n_y}$ และ $\bar{X}_{3n_{x_3}}$ ซึ่งหาได้นี้ แทนค่า Y_i's และ X_{3i}'s ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาคำนวณหาสมการถดถอย

$$\hat{Y}_i^o = b_o^o + b_3^o X_{3i}$$

$$\text{และ } \hat{X}_{3i}^o = d_o^o + d_3^o Y_i$$

โดยวิธีผสมในบทที่ 2 จะได้

$$b_3^o = \frac{\sum_{i=1}^{n} X_{3i}Y_i - n\bar{X}_{3n_{x_3}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} X_{3i}^2 - n\bar{X}_{3n_{x_3}}^2} = 0.82661$$

$$b_o^o = \bar{Y}_{n_y} - b_3^o \bar{X}_{3n_{x_3}} = 683.05230$$

.'. สมการถดถอยที่ได้คือ $\hat{Y}_i^o = 683.05230 + 0.82661X_{3i}$ (3.7.5)

$$d_3 = \frac{\sum_{i=1}^{n} X_{3i}Y_i - n\bar{X}_{3n_{x_3}}\bar{Y}_{n_y}}{\sum_{i=1}^{n} Y_i^2 - n\bar{Y}_{n_y}^2}$$

$$= 0.63427$$

$$d_o^o = \bar{X}_{3n_{x_3}} - d_3^o \bar{Y}_{n_y}$$

$$= 1,121.78971$$

.'. สมการถดถอยที่ได้คือ $\hat{X}_{3i}^o = 1,121.78971 + 0.63427Y_i$ (3.7.6)

หาค่า $\hat{Y}_i$ และ $\hat{X}_{3i}$ ที่ขาดหายไปโดยใช้สมการถดถอย (3.7.5) และ (3.7.6) ตามลำดับ ซึ่งจะได้ค่าดังนี้

เมื่อ $X_3 = 3,692.-$; $\hat{Y}^o = 683.0523 + 0.82661(3,692.-)$

$= 3,734.90$

$X_3 = 3,694.50$; $\hat{Y}^o = 683.0523 + 0.82661(3,694.50)$

$= 3,736.96$

$X_3 = 3,960.-$; $\hat{Y}^o = 683.0523 + 0.82661(3,960.-)$

$= 3,956.43$

$X_3 = 4,062.50$; $\hat{Y}^o = 683.0523 + 0.82661(4,062.50)$

$= 4,041.16$

$X_3 = 3,962.50$; $\hat{Y}^o = 683.0523 + 0.82661(3,962.50)$

$= 3,958.49$

$X_3 = 3,779.-$; $\hat{Y}^o = 683.0523 + 0.82661(3,779.00)$

$= 3,806.81$

$X_3 = 1,602.-$; $\hat{Y}^o = 683.0523 + 0.82661(1,602.-)$

$= 2,007.28$

$Y = 4,271.-$; $\hat{X}_3^o = 1,121.78971 + 0.63427(4,271.-)$

$= 3,830.76$

$Y = 4,475.-$; $\hat{X}_3^o = 1,121.78971 \quad 0.63427(4,475.-)$

$= 3,960.15$

$Y = 4,073.50$; $\hat{X}_3^o = 1,121.78971 + 0.63427(4,073.50)$

$= 3,705.49$

$Y = 1,737.50$; $\hat{X}_3^o = 1,121.78971 + 0.63427(1,737.50)$

$= 2,859.92$

$Y = 1,846.-$; $\hat{X}_3^o = 1,121.78971 + 0.63427(1,846.-)$

$= 2,292.65$

$Y = 2,342.-$; $\hat{X}_3^o = 1,121.78971 + 0.63427(2,342.-)$

$= 2,607.25$

$$Y = 2,642.- \quad ; \quad \hat{X}_3^o = 1,121.78971 + 0.63427(2,642.-)$$
$$= 2,797.53$$

ค่า Y_i และ X_{3i} ที่ขาดหายไป จะประมาณค่าโดยการนำค่า $\hat{Y}_i^o$ และ $\hat{X}_{3i}^o$ ที่คำนวณได้นี้ มาหาค่าเฉลี่ยกับ $\overline{Y}_{n_y}$ และ $\overline{X}_{3n_{x_3}}$ ตามลำดับ ซึ่งจะได้ค่าดังนี้

$$\hat{Y}_i' = \frac{\overline{Y}_{n_y} + \hat{Y}_i^o}{2}$$

$$\hat{X}_{3i}' = \frac{\overline{X}_{3n_{x_3}} + \hat{X}_{3i}^o}{2}$$

เมื่อ $\hat{Y}^o = 3,734.90 \; ; \quad \hat{Y}' = \frac{3,385.14 + 3,734.90}{2} = 3,560.02$

$\hat{Y}^o = 3,736.96 \; ; \quad \hat{Y}' = \frac{3,385.14 + 3,736.96}{2} = 3,561.05$

$\hat{Y}^o = 3,956.43 \; ; \quad \hat{Y}' = \frac{3,385.14 + 3,956.43}{2} = 3,670.78$

$\hat{Y}^o = 4,041.16 \; ; \quad \hat{Y}' = \frac{3,385.14 + 4,041.16}{2} = 3,713.15$

$\hat{Y}^o = 3,958.49 \; ; \quad \hat{Y}' = \frac{3,385.14 + 3,958.49}{2} = 3,671.82$

$\hat{Y}^o = 3,806.81 \; ; \quad \hat{Y}' = \frac{3,385.14 + 3,806.81}{2} = 3,595.98$

$\hat{Y}^o = 2,007.28 \; ; \quad \hat{Y}' = \frac{3,385.14 + 2,007.28}{2} = 2,696.21$

$\hat{X}_3^o = 3,830.76 \; ; \quad \hat{X}_3' = \frac{3,268.88 + 3,830.76}{2} = 3,549.82$

$$\hat{X}_3^o = 3,960.15 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 3,960.15}{2} = 3,614.51$$

$$\hat{X}_3^o = 3,705.49 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 3,705.49}{2} = 3,487.18$$

$$\hat{X}_3^o = 2,859.92 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 2,859.92}{2} = 3,064.40$$

$$\hat{X}_3^o = 2,292.65 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 2,292.65}{2} = 2,780.77$$

$$\hat{X}_3^o = 2,607.25 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 2,607.25}{2} = 2,938.07$$

$$\hat{X}_3^o = 2,797.53 \ ; \quad \hat{X}_3' = \frac{3,268.88 + 2,797.53}{2} = 3,033.21$$

นำค่า $\hat{Y}_i'$ และ $\hat{X}_{3i}'$ ซึ่งคำนวณได้ แทนค่า Y_i และ X_{3i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้ทั้งหมด หาสมการถดถอยโดยวิธีผสมในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้

$$b_3^{(6)} = \frac{\sum_{i=1}^{n} X_{3i}Y_i - n\bar{X}_{3n}\ \bar{Y}_n}{\sum_{i=1}^{n} X_{3i}^2 - n\bar{X}_{3n}^2} = 0.98859$$

$$b_o^{(6)} = \bar{Y}_n - b_3^{(6)}\ \bar{X}_{3n} = 186.39124$$

สมการถดถอยที่ได้ คือ $\hat{Y}_i^{(6)} = 186.39124 + 0.98859X_{3i}$

เราทดสอบสมมติฐาน $H_o : \beta_3 = 0$; $H_A : \beta_3 \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 35

ANOVA

SOV	df	SS	MS	F
Reg	1	2.3452801×10^{7}	2.3452801×10^{7}	60.42767^{**}
Error	20	7.7622711×10^{6}	3.8811355×10^{5}	
Total	21	3.1215072×10^{7}		

$$R^2 = \frac{\text{ผลบวกกำลังสองเนื่องจากการถดถอย}}{\text{ผลบวกกำลังสองของทั้งหมด}} = 0.75133$$

7.2 การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression)

เมื่อต้องการหา $\hat{Y}_i^{(6)} = b_o^{(6)} + b_1^{(6)} X_{1i} + b_2^{(6)} X_{2i}$

ซึ่งจะเขียนในรูปเมตริกซ์ได้ว่า

$$Y^{(6)} = X\hat{\beta}^{(6)}$$

เมื่อข้อมูลที่รวบรวมมาได้ตามตารางข้อมูลชุดที่ 5 ในภาคผนวก

$n = 36, \quad n_c = 23, \quad n_y = 30, \quad n_{x_1} = 31, \quad n_{x_2} = 30$

ซึ่งสามารถเขียนข้อมูลดังกล่าวให้อยู่ในรูปเมตริกซ์ ดังนี้

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} , \quad X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{bmatrix}$$

เมื่อ Y_1 แทนเวคเตอร์ค่าสังเกต ที่มีอยู่ ซึ่งตรงกับ X_1

Y_2 แทนเวคเตอร์ค่าสังเกต ที่สูญหาย ซึ่งตรงกับ X_2

Y_3 แทนเวคเตอร์ค่าสังเกต ที่มีอยู่ ซึ่งตรงกับ X_3

Y_4 แทนเวคเตอร์ค่าสังเกต ที่สูญหาย ซึ่งตรงกับ X_4

X_1 แทนเมตริกซ์ที่ไม่มีค่าสังเกตตัวใดตัวหนึ่งสูญหาย

X_2 แทนเมตริกซ์ที่ไม่มีค่าสังเกตตัวใดตัวหนึ่งสูญหาย

X_3 แทนเมตริกซ์ที่มีค่าสังเกตที่สูญหายอย่างน้อย 1 ตัว

X_4 แทนเมตริกซ์ที่มีค่าสังเกตที่สูญหายอย่างน้อย 1 ตัว

นำข้อมูลที่รวบรวมได้มาหาค่าเฉลี่ยของค่าสังเกต Y_i, X_{1i} และ X_{2i} เท่าที่ข้อมูล Y_i, X_{1i} และ X_{2i} มีอยู่ จะได้

$$\bar{Y}_{n_y} = \frac{\sum_{i=1}^{n_y} Y_i}{n_y} = 3,600.88$$

$$\bar{X}_{1n_{x_1}} = \frac{\sum_{i=1}^{n_{x_1}} X_{1i}}{n_{x_1}} = 2,159.14$$

$$\bar{X}_{2n_{x_2}} = \frac{\sum_{i=1}^{n_{x_2}} X_{2i}}{n_{x_2}} = 112,242.93$$

นำค่า $\bar{Y}_{n_y}$, $\bar{X}_{1n_{x_1}}$ และ $\bar{X}_{2n_{x_2}}$ ซึ่งหาได้นี้ แทนค่า Y_i, X_{1i} และ X_{2i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้มาคำนวณสมการถดถอย

$$\hat{Y}_i^o = b_o^o + b_1^o X_{1i} + b_2^o X_{2i}$$

$$\hat{X}_{1i}^o = d_o^o + d_1^o Y_i + d_2^o X_{2i}$$

$$\hat{X}_{2i}^o = a_o^o + a_1^o Y + a_2^o X_{2i}$$

โดยวิธีกำลังสองน้อยที่สุดในบทที่ 2 จะคำนวณสมการถดถอยได้ดังนี้

$$\hat{Y}_i^o = 2,189.18186 + 0.87802X_{1i} - 0.00431X_{2i} \qquad (3.7.7)$$

$$\hat{X}_{1i}^o = 422.98131 + 0.52442Y_i - 0.00136X_{2i} \qquad (3.7.8)$$

$$\hat{X}_{2i}^o = 268,263.87471 - 32.93255Y_i - 17.33785X_{1i} \qquad (3.7.9)$$

หาค่า Y_i, X_{1i} และ X_{2i} ที่ขาดหายไปโดยใช้สมการถดถอย (3.7.7), (3.7.8) และ (3.7.9) ตามลำดับ จะได้ค่าดังนี้

เมื่อ

$X_1 = 2,667.50$; $X_2 = 113,916.-$; $\hat{Y}^o = 2,189.18186 + 0.087802$
$(2,667.50)-0.00431(113,916.-)$
$= 4,040.32$

$X_1 = 2,159.14$; $X_2 = 93,625.-$; $\hat{Y}^o = 2,189.18186 + 0.087802$
$(2,159.14)-0.00431(93,625.-)$
$= 3,681.43$

$X_1 = 2{,}762.50$; $X_2 = 181{,}340.-$; $\hat{Y}^o = 2{,}189.18186 + 0.087802$
$(2{,}762.50) - 0.0043(181{,}340.-)$

$= 3{,}833.14$

$X_1 = 2{,}159.14$; $X_2 = 171{,}002.-$; $\hat{Y}^o = 2{,}189.18186 + 0.087802$
$(2{,}159.14) - 0.0043(171{,}002.-)$

$= 3{,}347.93$

$X_1 = 1{,}010.-$; $X_2 = 112{,}242.93$; $\hat{Y}^o = 2{,}189.18186 + 0.087802$
$(1{,}010.-) - 0.0043(112{,}242.93)$

$= 2{,}592.22$

$X_1 = 1{,}128.-$; $X_2 = 112{,}242.93$; $\hat{Y}^o = 2{,}189.18186 + 0.087802$
$(1{,}128.-) - 0.0043(112{,}242.93)$

$= 2{,}695.82$

$X_1 = 1{,}445.-$; $X_2 = 231{,}460.-$; $\hat{Y}^o = 2{,}189.18186 + 0.087802$
$(1{,}445.-) - 0.0043(231{,}460.-)$

$= 2{,}460.33$

$Y = 4{,}271.-$; $X_2 = 81{,}589.-$; $\hat{X}_1^o = 422.98131 + 0.52442(4{,}271.-)$
$- 0.00136(81{,}589.-)$

$= 2{,}551.82$

$Y = 3,600.88$; $X_2 = 93,625.-$; $\hat{X}_1^o = 422.98131 + 0.52442(3,600.88) - 0.00136(93,625.-)$

$= 2,184.02$

$Y = 4,073.50$; $X_2 = 112,242.93$; $\hat{X}_1^o = 422.98131 + 0.52442(4,073.50) - 0.00136(112,242.93)$

$= 2,406.56$

$Y = 3,600.88$; $X_2 = 171,002.-$; $\hat{X}_1^o = 422.98131 + 0.52442(3,600.88) - 0.00136(171,002.-)$

$= 2,078.79$

$Y = 2,883.50$; $X_2 = 209,806.-$; $\hat{X}_1^o = 422.98131 + 0.52442(2,883.50) - 0.00136(209,806.-)$

$= 1,649.81$

$Y = 4,533.50$; $X_1 = 2,570.-$; $\hat{X}_2^o = 268,263.87471 - 32.93255(4,533.50) - 17.33785(2,570.-)$

$= 74,405.88$

$Y = 4,117.50$; $X_1 = 2,665.-$; $\hat{X}_2^o = 268,263.87471 - 32.93255(4,117.50) - 17.33785(2,665.-)$

$= 86,458.73$

$Y = 4,073.50;\ X_1 = 2,159.14;\ \hat{X}_2^o = 268,263.87471 - 32.93255$
$(4,073.50) - 17.33785(2,159.14)$

$= 96,678.29$

$Y = 3,600.88;\ X_1 = 1,010.- \ ;\ \hat{X}_2^o = 268,263.87471 - 32.93255$
$(3,600.88) - 17.33785(1,010.-)$

$= 132,166.49$

$Y = 3,600.88;\ X_1 = 1,128.- \ ;\ \hat{X}_2^o = 268,263.87471 - 32.93255$
$(3,600.88) - 17.33785(1,128.-)$

$= 130,120.62$

$Y = 2,642.- \ ;\ X_1 = 1,607.50;\ \hat{X}_2^o = 268,263.87471 - 32.93255$
$(2,642.-) - 17.33785(1,607.50)$

$= 153,385.48$

ประมาณค่า Y_i, X_{1i} และ X_{2i} ที่ขาดหายไปด้วย $\hat{Y}_i'$, $\hat{X}_{1i}'$ และ $\hat{X}_{2i}'$ ตามลำดับ

เมื่อค่า $\hat{Y}_i' = \dfrac{\bar{Y}_{n_y} + \hat{Y}_i^o}{2}$, $\hat{X}_{1i}' = \dfrac{\bar{X}_{1n_{x_1}} + \hat{X}_{1i}^o}{2}$, $\hat{X}_{2i}' = \dfrac{\bar{X}_{2n_{x_2}} + \hat{X}_{2i}^o}{2}$

เมื่อ $\hat{Y}^o = 4,040.32$; $\hat{Y}' = \dfrac{3,600.88 + 4,040.32}{2} = 3,820.60$

$\hat{Y}^o = 3,681.43$; $\hat{Y}' = \dfrac{3,600.88 + 3,681.43}{2} = 3,641.16$

$$\hat{Y}^{o} = 3,833.14 \; ; \quad \hat{Y}' = \frac{3,600.88 + 3,833.14}{2} = 3,717.01$$

$$\hat{Y}^{o} = 3,347.93 \; ; \quad \hat{Y}' = \frac{3,600.88 + 3,347.93}{2} = 3,474.41$$

$$\hat{Y}^{o} = 2,592.22 \; ; \quad \hat{Y}' = \frac{3,600.88 + 2,592.22}{2} = 3,096.55$$

$$\hat{Y}^{o} = 2,692.22 \; ; \quad \hat{Y}' = \frac{3,600.88 + 2,695.82}{2} = 3,148.35$$

$$\hat{Y}^{o} = 2,460.33 \; ; \quad \hat{Y}' = \frac{3,600.88 + 2,460.33}{2} = 3,030.61$$

$$\hat{X}_1^{o} = 2,551.82 \; ; \quad \hat{X}_1' = \frac{2,159.14 + 2,551.82}{2} = 2,355.48$$

$$\hat{X}_1^{o} = 2,184.02 \; ; \quad \hat{X}_1' = \frac{2,159.14 + 2,184.02}{2} = 2,171.58$$

$$\hat{X}_1^{o} = 2,406.56 \; ; \quad \hat{X}_1' = \frac{2,159.14 + 2,406.56}{2} = 2,282.85$$

$$\hat{X}_1^{o} = 2,078.79 \; ; \quad \hat{X}_1' = \frac{2,159.14 + 2,078.79}{2} = 2,118.97$$

$$\hat{X}_1^{o} = 1,649.81 \; ; \quad \hat{X}_1' = \frac{2,159.14 + 1,649.81}{2} = 1,904.48$$

$$\hat{X}_2^{o} = 77,405.88 \; ; \quad \hat{X}_2' = \frac{112,242.93 + 74,405.88}{2} = 93,324.41$$

$$\hat{X}_2^{o} = 86,458.73 \; ; \quad \hat{X}_2' = \frac{112,242.93 + 86,458.73}{2} = 99,350.83$$

$$\hat{X}_2^{o} = 96,678.29 \; ; \quad \hat{X}_2' = \frac{112,242.93 + 96,678.29}{2} = 104,460.61$$

$$\hat{X}_2^o = 132,166.49 \;;\; \hat{X}_2' = \frac{112,242.93 + 132,166.49}{2} = 122,204.71$$

$$\hat{X}_2^o = 130,120.62 \;;\; \hat{X}_2' = \frac{112,242.93 + 130,120.62}{2} = 121,181.78$$

$$\hat{X}_2^o = 153,385.48 \;;\; \hat{X}_2' = \frac{112,242.93 + 153,385.48}{2} = 132,814.21$$

นำค่า $\hat{Y}_i'$, $\hat{X}_{1i}'$ และ $\hat{X}_{2i}'$ ซึ่งคำนวณได้นี้ แทนค่า Y_i, X_{1i} และ X_{2i} ที่ขาดหายไปตามลำดับ แล้วนำข้อมูลที่ได้ใหม่ หาสมการถดถอยโดยวิธีผสมในบทที่ 2 n_c จะเพิ่มเท่ากับ $n = 36$ และจะได้

$$\hat{\beta}^{(6)} = (X'X)^{-1}X'Y = \begin{bmatrix} 1,693.81113 \\ 1.04928 \\ -0.00352 \end{bmatrix}$$

$\therefore$ สมการถดถอยที่ได้คือ $\hat{Y}^{(6)} = 1,693.81113 + 1.04928X_{1i} - 0.00352X_{2i}$

เราทดสอบสมมติฐาน $H_o : \beta = 0$; $H_A : \beta \neq 0$ โดยใช้การทดสอบ F (F-test) ตามตาราง ANOVA ดังนี้

ตารางที่ 36

ANOVA

SOV	df	SS	MS	F
Reg	3	2.0701969×10^7	6.9006563×10^6	35.35218**
Error	20	3.9039495×10^6	1.9519747×10^5	
Total	23	2.4605918×10^7		

$R^2 = 0.84134$