การวิเคราะห์ความถดถอยเชิงซ้อน (Multiple Regression Analysis)

การวิเคราะห์ความถดถอยเชิงซ้อนเป็นการศึกษาความสัมพันธ์ระหว่างตัวแปรตาม (y) 1 ตัวแปร กับ ตัวแปรอิสระ (x) ตั้งแต่ 2 ตัวแปรขึ้นไป โดยที่รูปแบบและการศึกษาจะคล้ายกับการวิเคราะห์ความถดถอยอย่างง่าย

การประมาณค่าคงที่ และ ค่าสัมประสิทธิ์ของตัวแปรอิสระ

โดยเริ่มจาก ถ้ากำหนดให้สมการถดถอยเชิงเส้น คือ

$$\hat{y} = a + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

เมื่อ k = จำนวนตัวแปรอิสระ และ $k \geq 2$

เพื่อให้ง่ายแก่การวิเคราะห์ จึงกำหนดให้มีตัวแปรอิสระ 2 ตัว ($k = 2$)

$$\hat{y} = a + b_1x_1 + b_2x_2$$

โดยที่ $SSE = \sum (y - a - b_1x_1 - b_2x_2)^2$

$$= \sum (y - a - b_1(x'_1 + \bar{x}_1) - b_2(x'_2 + \bar{x}_2))^2$$

เมื่อ $x_1 = x'_1 + \bar{x}_1$ และ $x_2 = x'_2 + \bar{x}_2$

$$= \sum (y - (a + b_1\bar{x}_1 + b_2\bar{x}_2) - b_1x'_1 - b_2x'_2)^2$$

$$= \sum (y - a' - b_1x'_1 - b_2x'_2)^2$$

เมื่อ $a' = a + b_1\bar{x}_1 + b_2\bar{x}_2$

หาค่า a', b_1 และ b_2 ที่ทำให้ SSE มีค่าน้อยที่สุด

$$\frac{\partial SSE}{\partial a'} = -2\sum(y - a' - b_1x'_1 - b_2x'_2) = 0 \text{ \_\_\_\_\_\_\_\_\_\_ สมการ 1}$$

จะได้ว่า $= \sum(y - a' - b_1x'_1 - b_2x'_2) = 0$

$$= \sum y - na' - b_1\sum x'_1 - b_2\sum x'_2 = 0$$

เนื่องจาก $x'_1 = x_1 - \bar{x}_1$ และ $x_2 - \bar{x}_2$ จึงทำให้

$\sum x'_1 = \sum(x_1 - \bar{x}_1) = 0$ และ $\sum x'_2 = \sum(x_2 - \bar{x}_2) = 0$

จะได้ว่า $\sum y - na' = 0$ __________ สมการ 2

$$a' = \bar{y}$$

ดังนั้น $a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2$ __________ สมการ 3

$$\frac{\partial SSE}{\partial b_1} = -2\sum(y - a' - b_1x'_1 - b_2x'_2)\,x'_1 = 0$$

จะได้ว่า $\sum x'_1\, y - a'\sum x'_1 - b_1\sum x^{2\prime}_1 - b_2\sum x'_1x'_2 = 0$ __________ สมการ 4

เนื่องจาก $\sum x_1' = 0$

จึงทำให้ $\sum x_1'y = b_1\sum x'^{2}_1 + b_2\sum x'_1x'_2$ __________ สมการ 5

$$\frac{\partial SSE}{\partial b_2} = -2\sum(y - a' - b_1x'_1 - b_2x'_2)\,x'_2 = 0$$

จะได้ว่า $\sum x_2'y - a'\sum x_2' - b_1\sum x_1'x_2' - b_2\sum x'^{2}_2 = 0$ __________ สมการ 6

เนื่องจาก $\sum x_2' = 0$

จึงทำให้ $\sum x_2'y = b_1\sum x_1'x_2' + b_2\sum x'^{2}_2$ __________ สมการ 7

นำสมการที่ 5 และ 7 เขียนในรูป Matrix

$$\begin{bmatrix} \sum x_1'y \\ \sum x_2'y \end{bmatrix} = \begin{bmatrix} \sum x_1'^2 & \sum x_1'x_2' \\ \sum x_1'\,x_2' & \sum x_2'^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

โดยที่ $b_1 = \dfrac{|A_1|}{|A|}$ และ $b_2 = \dfrac{|A_2|}{|A|}$

เมื่อ $|A| = \begin{vmatrix} \sum x'^{2}_1 & \sum x_1'x_2' \\ \sum x_1'x_2' & \sum x'^{2}_2 \end{vmatrix}$

$|A_1| = \begin{vmatrix} \sum x'_1y & \sum x_1'x_2' \\ \sum x_2'y & \sum x'^{2}_2 \end{vmatrix}$

$|A_2| = \begin{vmatrix} \sum x'^{2}_1 & \sum x_1'y \\ \sum x_2'x_2' & \sum x'_2y \end{vmatrix}$

ตัวอย่างที่ 1

ห้างสรรพสินค้าแห่งหนึ่งต้องการศึกษาความสัมพันธ์ระหว่างยอดขายรายเดือน กับ จำนวนพนักงานขาย และ ค่าโฆษณา โดยมีข้อมูลดังนี้

เดือนที่	ยอดขาย (แสนบาท)	จำนวนพนักงานขาย (คน)	ค่าโฆษณา (พันบาท)
1	20	70	125
2	24	72	130
3	16	68	120
4	28	74	135
5	32	75	140
6	19	68	120
7	24	71	130
8	29	72	135
9	35	75	145

วิธีทำ กำหนดให้ y = ยอดขาย

x_1 = จำนวนพนักงานขาย

x_2 = ค่าโฆษณา

Y	x_1	x_2	x_1'	x_2'	$X_1'^2$	$x_2'^2$	$x_1'x_2'$	$x_1'y$	$x_2'y$
20	70	125	-1.67	-5.56	2.779	30.869	9.262	-33.33	-111.12
24	72	130	0.33	-0.56	0.111	0.309	-0.185	8	-13.334
16	68	120	-3.67	-10.56	13.447	111.429	38.709	-58.67	-168.896
28	74	135	2.33	4.44	5.443	19.749	10.368	65.33	124.432
32	75	140	3.33	9.44	11.109	89.189	31.477	106.67	302.208
19	68	120	-3.67	-10.56	13.447	111.429	38.709	-69.67	-200.564
24	71	130	-0.67	-0.56	0.445	0.309	0.371	-16	-13.344
29	72	135	0.33	4.44	0.111	19.749	1.480	9.67	128.876
35	75	140	3.33	9.44	11.109	89.189	31.477	111.67	330.54
227	645	1,175	0	0	58	472.222	161.667	128.66	378.878

$$\bar{x}_1 = 645/9 = 71.667$$

$$\bar{x}_2 = 1{,}175/9 = 130.556$$

$$\bar{y} = 227/9 = 25.222$$

$$|A| = \begin{vmatrix} \Sigma x'^2_1 & \Sigma x_1'x_2' \\ \Sigma x_1'x_2' & \Sigma x'^2_2 \end{vmatrix} = \begin{vmatrix} 58 & 161.667 \\ 161.667 & 472.222 \end{vmatrix} = 1{,}252.778$$

$$|A_1| = \begin{vmatrix} \Sigma x'_1 y & \Sigma x_1'x_2' \\ \Sigma x_2'y & \Sigma x'^2_2 \end{vmatrix} = \begin{vmatrix} 128.66 & 161.667 \\ 378.878 & 472.222 \end{vmatrix} = -496.387$$

$$|A_2| = \begin{vmatrix} \Sigma x'^2_1 & \Sigma x_1'y \\ \Sigma x_2'x_2' & \Sigma x'_2 y \end{vmatrix} = \begin{vmatrix} 58 & 128.66 \\ 161.667 & 378.878 \end{vmatrix} = 1{,}175.08$$

$b_1 = |A_1| / |A| = -496.387 / 1,252.778 = -0.396$

$b_2 = |A_2| / |A| = 1,175.08 / 1,252.778 = 0.938$

$\therefore\ a = y - b_1\overline{x}_1 - b_2\overline{x}_2$

$= 25.22 - (-0.396)(71.667) - (0.938)(130.556)$

$= -68.861$

ดังนั้น สมการถดถอยที่ได้ คือ $\hat{y} = 68.861 - 0.396\,x_1 + 0.938\,x_2$

การหาค่า R^2 , $\overline{R}^2$ และ SEE

สำหรับการหาค่า R^2 , $\overline{R}^2$ และ SSE ในการวิเคราะห์สมการถดถอยเชิงซ้อนนั้น จะใช้หลักการเหมือนกับการวิเคราะห์ความถดถอยอย่างง่ายในบทที่ 5

ตัวอย่างที่ 6.2

จากตัวอย่างที่ 6.1 จงหาค่า R^2 , $\overline{R}^2$ และ SSE พร้อมทั้งอธิบายค่าที่ได้

จากตัวอย่างที่ 6.1 $\hat{y} = 68.861 - 0.396\,x_1 + 0.938\,x_2$

x_1	x_2	Y	$\hat{y} = -68.861 - 0.396x_1 + 0.938x_2$	$y-\overline{y}$	$(y-\overline{y})^2$	$\hat{y}-\overline{y}$	$(\hat{y}-\overline{y})^2$
70	125	20	20.669	-5.222	27.271	-4.553	20.730
72	130	24	24.567	-1.222	1.494	-0.655	0.429
68	120	16	16.771	-9.222	85.049	-8.451	71.419
74	135	28	28.465	2.778	7.716	3.243	10.517
75	140	32	32.759	6.778	45.939	7.531	56.806
68	120	19	16.771	-6.222	38.716	-8.451	71.419
71	130	24	24.963	-1.222	1.494	-0.259	0.067
72	135	29	29.257	3.778	14.272	4.035	16.281
75	140	35	32.759	9.778	95.605	7.537	56.806
645	1,175	227	226.981	0.00	317.556	-0.017	304.476

โดยที่ $\overline{Y} = 227/9 = 25.222$

$$R^2 = \frac{\Sigma(\hat{y}-\overline{y})^2}{\Sigma(y-\overline{y})^2} = \frac{SSR}{SST}$$

$$= 304.476 / 317.556 = 0.9588$$

$R^2 = 0.9588$ หมายความว่า ความผันแปรทั้งหมดของยอดขาย มีสาเหตุมาจากจำนวนพนักงานขายและค่าโฆษณา 95.88%

$$\overline{R}^2 = 1 - \frac{SSE}{SST} \cdot \frac{(n-1)}{(n-k-1)}$$

$$= 1 - \left[\frac{(1-R^2) \times (n-1)}{(n-k-1)}\right]$$

$$= 1 - \left[\frac{(1-0.9588) / (8)}{6}\right]$$

$$= 0.9451$$

R^2 = 0.9451 หมายความว่า เมื่อได้มีการปรับค่าสัมประสิทธิ์การตัดสินใจแล้ว ความผันแปรทั้งหมดของยอดขาย มีสาเหตุมาจากจำนวนพนักงานขายและค่าโฆษณา 94.51%

$$SEE = \sqrt{\frac{SSE}{(n-k-1)}}$$

โดยที่ $SSE = SST - SSR \; \Sigma(y - \overline{y})^2 - \Sigma(\hat{y} - \overline{y})^2$

$$= 317.556 - 304.476 = 13.08$$

$$SEE = \sqrt{13.08 / 6} = 1.476$$

SEE = 1.476 หมายความว่า ยอดขายที่เกิดขึ้นจริง กับ ยอดขายที่ได้จากการประมาณมีความแตกต่างกันประมาณ 1.476 แสนบาท

การทดสอบสมมติฐาน

สำหรับการทดสอบสมมติฐานในการวิเคราะห์ความถดถอยเชิงซ้อน จะแบ่งเป็น 2 ส่วน คือ การทดสอบโดยค่าสถิติ F กับ การทดสอบโดยค่าสถิติ t

การทดสอบโดยค่าสถิติ F เป็นการทดสอบสมมติฐานว่า มีตัวแปรอิสระอย่างน้อย 1 ตัว มีความสัมพันธ์กับ ตัวแปรตามหรือไม่ ซึ่งมีขั้นตอนดังนี้

(1) การตั้งสมมติฐาน

H_0 : $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ (ตัวแปรอิสระทุกตัวไม่มีความสัมพันธ์กับตัวแปรตาม)

H_1 : มี β_i อย่างน้อย 1 ตัว $\neq 0$ เมื่อ i = 1 , 2 , … , k (มีตัวแปรอิสระอย่างน้อย 1 ตัว ที่มีความสัมพันธ์กับตัวแปรตาม)

(2) กำหนดระดับนัยสำคัญ (α) พร้อมทั้งเปิดตารางหาค่า $F_{(\alpha, v1, v2)}$

เมื่อ $v_1 = k$ และ $v_2 = n-k-1$

k = จำนวนตัวแปรอิสระ

n = จำนวนตัวอย่าง

(3) คำนวณหาค่า F โดยอาศัยตาราง One-Way Anova

สาเหตุของความแปรปรวน	ผลรวมกำลังสอง	**จำนวนองศาอิสระ**	**ค่าเฉลี่ยกำลังสอง**	F
- Regression	SSR	k	MSR = SSR/K	F = MSR/MSE
- Residual	SSE	n-k-1	MSE = SSE/(n-k-1)	
ผลรวม	SST	n-1		

(4) เปรียบเทียบค่า F ที่ได้จากการคำนวณ กับ $F_{(\alpha,v1,v2)}$

- ถ้าค่า $F < F_{(\alpha,v1,v2)}$ จะยอมรับ H_0
- ถ้าค่า $F > F_{(\alpha,v1,v2)}$ จะปฏิเสธ H_0

(5) สรุปคำตอบ

การทดสอบโดยค่าสถิติ t เป็นการทดสอบว่า ตัวแปรอิสระแต่ละตัวในสมการถดถอย มีความสัมพันธ์กับตัวแปรตามหรือไม่ ซึ่งมีขั้นตอนดังนี้

(1) ตั้งสมมติฐาน (กรณีที่ทดสอบทางเดียวและมีตัวแปรอิสระ 2 ตัว)

กรณีที่ค่า $b_i > 0$

เมื่อ i = 1 , 2

H_0 : $\beta_i \leq 0$ (x_i กับ y ไม่มีความสัมพันธ์ในทิศทางเดียวกัน)

H_1 : $\beta_i > 0$ (x_i กับ y มีความสัมพันธ์ในทิศทางเดียวกัน)

กรณีที่ค่า $b_i < 0$

H_0 : $\beta_i \geq 0$ (x_i กับ y ไม่มีความสัมพันธ์ในทิศทางตรงกันข้าม)

H_1 : $\beta_i < 0$ (x_i กับ y มีความสัมพันธ์ในทิศทางตรงกันข้าม)

ขั้นที่ 2 หาค่า t ที่ได้จากการเปิดตาราง ที่ $t_{(\alpha\ ,\ n-k-1)}$

ขั้นที่ 3 หาค่า t ที่ได้จากการคำนวณ

เมื่อ $t = b_i / s_{bi}$

b_i คือ ค่าประมาณของ β_i

s_{bi} คือ ส่วนเบี่ยงเบนมาตรฐานของ b_i

โดยที่ $s_{bi} = \sqrt{Var(b_i)}$

เมื่อ $Var(b_i)$ คือ ความแปรปรวนของ b_i

โดยที่สามารถหาค่า $Var(b_i)$ ได้จาก Variance-Covarian Matrix ($V(\hat{\beta})$)

เมื่อ $V(\hat{\beta}) = \begin{bmatrix} Var(b_1) & Cov(b_1,b_2) \\ Cov(b_1,b_2) & Var(b_1) \end{bmatrix}$

โดยที่ $V(\hat{\beta}) = s^2 A^{-1}$

เมื่อ s คือ ความคลาดเคลื่อนมาตรฐานของการประมาณ (SEE)

เนื่องจาก $A = \begin{bmatrix} \Sigma x'^2_1 & \Sigma x'_1 x'_2 \\ \Sigma x'_1 x'_2 & \Sigma x'^2_2 \end{bmatrix}$

ดังนั้น $A^{-1} = \dfrac{1}{|A|} \begin{bmatrix} \Sigma x'^2_2 & -\Sigma x'_1 x'_2 \\ -\Sigma x'_1 x'_2 & \Sigma x'^2_1 \end{bmatrix}$

ดังนั้น $V(\hat{B}) = \dfrac{s^2}{|A|} \begin{bmatrix} \Sigma x'^2_2 & -\Sigma x'_1 x'_2 \\ -\Sigma x'_1 x'_2 & \Sigma x'^2_1 \end{bmatrix}$

จะได้ว่า $Var(b_1) = s^2 \Sigma x'^2_2 \ / \ |A|$

$Var(b_2) = s^2 \Sigma x'^2_1 \ / \ |A|$

หรือ $s_{b1} = s\sqrt{\Sigma x'^2_2 / |A|}$

$s_{b2} = s\sqrt{\Sigma x'^2_1 / |A|}$

ขั้นที่ 4 เปรียบเทียบค่า t ที่ได้จากตาราง กับ ค่า t ที่ได้จากการคำนวณ

กรณีที่ $b_i > 0$ ปฏิเสธ H_0 เมื่อ $t > t_{(\alpha\ ,\ n-k-1)}$

กรณีที่ $b_i < 0$ ปฏิเสธ H_0 เมื่อ $-t < -t_{(\alpha\ ,\ n-k-1)}$

ขั้นที่ 5 สรุปคำตอบ

ตัวอย่างที่ 3

จากตัวอย่างที่ 6.1 และ 6.2 จงทดสอบสมมติฐานสมการถดถอยเชิงซ้อน ที่ระดับนัยสำคัญ 0.05

1. ทดสอบโดยค่าสถิติ F

   1.1 ตั้งสมมติฐาน

   $H_0 : \beta_1 = \beta_2 = 0$

   H_1 : มี β_i อย่างน้อย 1 ตัว $\neq 0$ เมื่อ $i = 1 , 2$

   1.2 เปิดตารางหาค่า $F_{(0.05\ ,\ 2\ ,\ 6)} = 5.14$

   1.3 คำนวณหาค่า F โดยใช้ตาราง One-Way ANOVA

สาเหตุของความแปรปรวน	ผลรวมกำลังสอง	จำนวนองศาอิสระ	ค่าเฉลี่ยกำลังสอง	F
Regression	304.476	2	152.238	= 69.834
Residual	13.08	6	2.18	
รวม	317.556	8		

1.4 เนื่องจากค่า F ที่ได้จากการคำนวณ = 69.834 มีค่ามากกว่า F (0.05 , 2 , 6) = 5.14 จึงปฏิเสธ H_0

1.5 สรุปคำตอบ มีปัจจัยอย่างน้อย 1 ปัจจัย ที่มีความสัมพันธ์กับยอดขาย ที่ระดับนัยสำคัญ 0.05

2. ทดสอบสมมติฐาน โดยค่าสถิติ t

2.1) ทดสอบสมมติฐาน β_1

2.1.1) $H_0 : \beta_1 \geq 0$

$H_1 : \beta_1 < 0$

2.1.2) เปิดตารางหาค่า $t_{(0.05 , 6)}$ = 1.943

2.1.3) คำนวณหาค่า $t = b_1/s_{b1}$

โดยที่ $b_1 = -0.396$

$S_{b1} = s\sqrt{\Sigma x'^2_2 / |A|}$

$= (1.476)\sqrt{\Sigma x'^2_2 / |A|}$

$= 0.9062$

$t = -0.396 / 0.9062 = -0.437$

2.1.4) เนื่องจากค่า t ที่ได้จากการคำนวณ = -0.437 มีค่ามากกว่า $-t_{(0.05 , 6)}$ = -1.943 ดังนั้น จึงยอมรับ H_0

2.1.5) สรุปคำตอบ จำนวนพนักงานขายไม่มีความสัมพันธ์ในทิศทางตรงกันข้ามกับยอดขาย ที่ระดับนัยสำคัญ 0.05

2.2) ทดสอบสมมติฐาน β_2

2.2.1) $H_0 : \beta_2 \leq 0$

$H_1 : \beta_2 > 0$

2.2.2) เปิดตารางหาค่า $t_{(0.05 , 6)}$ = 1.943

2.2.3) คำนวณหาค่า $t = b_2/s_{b2}$

โดยที่ $b_2 = 0.938$

$S_{b2} = s\sqrt{\Sigma x'^2_1 / |A|}$

$= (1.476) \sqrt{58 / 1,252.778}$

$= 0.3176$

$t = 0.938 / 0.3176 = 2.953$

2.2.4) เนื่องจากค่า t ที่ได้จากการคำนวณ = 2.953 มีค่ามากกว่า $t_{(0.05,6)}$ = 1.943 ดังนั้น จึงปฏิเสธ H_0

2.2.5) สรุปคำตอบ ค่าโฆษณามีความสัมพันธ์กับยอดขายในทิศทางเดียวกัน ณ ระดับนัยสำคัญ 0.05

ตัวอย่างที่ 4

ในการศึกษาถึงความสัมพันธ์ระหว่างราคาตึกแถว กับ จำนวนห้องนอน และ จำนวนห้องน้ำ ได้เลือกตึกแถวโดยการสุ่มมาจำนวน 8 ห้อง ผลปรากฏดังนี้

ราคา (พันบาท)	จำนวนห้องนอน	จำนวนห้องน้ำ
338	3	2
293	2	1
388	4	3
292	2	1
347	3	2
299	2	2
434	5	3
379	4	2

ก. จงหาสมการถดถอย

วิธีทำ กำหนดให้ Y = ราคาตึกแถว

x_1 = จำนวนห้องนอน

x_2 = จำนวนห้องน้ำ

Y	x_1	x_2	x_1'	x_2'	$x_1'^2$	$X_2'^2$	$x_1'x_2'$	$x_1'Y$	$x_1'Y$
338	3	2	-0.125	0	0.015625	0	0	-42.25	0
293	2	1	-1.125	-1	1.265625	1	1.125	-329.625	-293
388	4	3	0.875	1	0.765625	1	0.875	339.5	388
292	2	1	-1.125	-1	1.265625	1	1.125	-328.5	-292
347	3	2	-0.125	0	0.015625	0	0	-43.375	0
299	2	2	-1.125	0	1.265625	0	0	-336.375	0
434	5	3	1.875	1	3.515625	1	1.875	813.75	434
379	4	2	0.875	0	0.765625	0	0	331.625	0
2770	25	16	0	0	8.875	4	5	404.75	237

$$\bar{x}_1 = 25/8 = 3.125$$
$$\bar{x}_2 = 16/8 = 2$$
$$\bar{y} = 2770/8 = 346.25$$

$$|A| = \begin{vmatrix} \sum x_1'^2 & \sum x_1'x_2' \\ \sum x_1'x_2' & \sum x_2'^2 \end{vmatrix} = \begin{vmatrix} 8.875 & 5 \\ 5 & 4 \end{vmatrix} = 10.5$$

$$|A_1| = \begin{vmatrix} \sum x_1'y & \sum x_1'x_2' \\ \sum x_2'y & \sum x_2'^2 \end{vmatrix} = \begin{vmatrix} 404.75 & 5 \\ 237 & 4 \end{vmatrix} = 434$$

$$|A_2| = \begin{vmatrix} \sum x_1'^2 & \sum x_1'y \\ \sum x_1'x_2' & \sum x_2'y \end{vmatrix} = \begin{vmatrix} 8.875 & 404.75 \\ 5 & 2.37 \end{vmatrix} = 79.625$$

$$b_1 = |A_1| / |A| = 434 / 10.5 = 41.3333$$
$$b_2 = |A_2| / |A| = 79.625 / 10.5 = 7.5833$$
$$a = \bar{y} - b_1\bar{x}_1 - b_2\bar{x}_2$$
$$= 346.25 - 41.333(3.125) - 7.5833(2)$$
$$= 201.9168$$

สมการถดถอยที่ได้ คือ $\hat{y} = 201.9168 + 41.3333\,x_1 + 7.5833\,x_2$

ข. จงหา R^2 , $\bar{R}^2$ และ SEE พร้อมทั้งอธิบายค่าที่ได้

จากข้อ ก.

$$\hat{y} = 201.9168 + 41.3333\, x_1 + 7.5833\, x_2$$
$$\bar{y} = 346.25$$

x_1	x_2	Y	$\hat{y}$ = ____	$y-\bar{y}$	$(y-\bar{y})^2$	$(\hat{y}-\bar{y})$	$(\hat{y}-\bar{y})^2$
3	2	338	341.0833	-8.25	341.0833	-5.1667	26.6948
2	1	293	292.1667	-53.25	292.1667	-54.0833	2925.0033
4	3	388	389.9999	41.75	1746.0625	43.7499	1914.0537
2	1	292	292.1667	-54.25	2943.0625	-54.0833	2925.0033
3	2	347	341.0833	0.75	0.5625	-5.1667	26.6948
2	2	299	299.75	-47.25	2232.5625	-46.5	2162.25
5	3	434	431.3332	87.75	7700.0625	85.0832	7239.1509
4	2	379	382.4166	32.75	1072.5625	36.1666	1308.0230
25	16	2770	2770	0	18595.5	-0.0003	18526.8739

SSR = 18526.8739 , SSE = 68.6261 , SST = 18595.5

$R^2 = \frac{SSR}{SST} = \frac{18526.8739}{18595.5} = 0.9963$ หมายความว่า ความผันแปรทั้งหมดของราคาตึกแถวมีสาเหตุมาจากจำนวนห้องนอนและจำนวนห้องน้ำ 99.48 %

$$\bar{R}^2 = 1 - \frac{(1-R^2)(n-1)}{(n-k-1)}$$
$$= 1 - \frac{(1-0.9963)(7)}{5} = 0.9948$$

$\bar{R}^2 = 0.9948$ หมายความว่าเมื่อมีการปรับค่าสัมประสิทธิ์การตัดสินใจแล้ว ความผันแปรทั้งหมดของราคาตึกแถวมีสาเหตุมาจากจำนวนห้องนอนและจำนวนห้องน้ำ 99.48%

$$SEE = \sqrt{\frac{SSE}{(n-k-1)}} = \sqrt{\frac{68.6261}{5}} = 3.7048$$

$SEE = 3.7048$ หมายความว่าราคาตึกแถวตามความเป็นจริงกับราคาตึกแถวที่ได้จากการประมาณมีความแตกต่างกันประมาณ 3.7048 (พันบาท)

ค. จงทดสอบสมมติฐานสมการถดถอย โดยค่าสถิติ F ที่ระดับนัยสำคัญ 0.05

(1)

$$H_0 : \beta_1 = \beta_2 = 0$$
H_1 : ÁÕβ_i ÍÂèÒ§¹éÍÂ1 µÑÇ≠ 0 àÁ×èÍ= 1,2

(2) ค่า F ที่ได้จากเปิดตาราง $= F(\alpha, V_1, V_2)$

$= (0.05, 2, 5)$

$= 5.79$

3. ค่า F ที่ได้จากการคำนวณ

สาเหตุของความแปรปรวน	Sum Square	d.f.	Mean Square	F
Regression	18526.8739	2	9263.4370	674.9218
Residual	68.6261	5	13.7252	
รวม	18595.5	7		

4. เปรียบเทียบ ค่า F ที่ได้จากการเปิดตาราง กับ F ที่ได้จากการคำนวณ

   Fคำนวณ = 674.9218 ซึ่งมากกว่า F(0.05,2,5) = ดังนั้นจึงปฏิเสธ H_0

5. สรุปคำตอบ : มีปัจจัยอย่างน้อย 1 ปัจจัย ที่มีความสัมพันธ์กับราคาตึกแถว ที่ระดับนัยสำคัญ 0.05

ง. จงทดสอบสมมติฐานค่าสัมประสิทธิ์ของตัวแปรอิสระ โดยค่าสถิติ t ที่ระดับนัยสำคัญ 0.05

1) ทดสอบสมมติฐาน β_1

1.1)

$$H_0 : \beta_1 \le 0$$
$$H_1 : \beta_1 > 0$$

1.2) t ที่ได้จากการเปิดตาราง = t(0.05,5) = 2.015

1.3) t ที่ได้จากการคำนวณ $= b_1 / s_{b1}$ $\quad s_{b1} = s\sqrt{\sum x_2'^2 / |A|}$

$$= \frac{41.3333}{2.2867} \qquad = 3.7048\sqrt{4/10.5}$$
$$= 18.0755 \qquad = 2.2867$$

1.4) เปรียบเทียบค่า t ที่ได้จากการเปิดตาราง กับ t ที่ได้จากการคำนวณ

t คำนวณ = 18.0755 ซึ่งมากกว่าค่า t(0.05,5) = 2.015 ดังนั้นจึงปฏิบัติ H_0

1.5) สรุปคำตอบ

จำนวนห้องนอนมีความสัมพันธ์กับราคาตึกแถวในทิศทางเดียวกัน ณ ระดับนัยสำคัญ 0.05

2) ทดสอบสมมติฐาน β_2

2.1)

$$H_0 : \beta_2 \le 0$$
$$H_1 : \beta_2 > 0$$

2.2) t ที่ได้จากการเปิดตาราง = t(0.05,5) = 2.015

2.3) t ที่ได้จากการคำนวณ $= b_2 / s_{b2}$; $s_{b2} = s\sqrt{\sum x_1'^2 / |A|}$

$$= \frac{7.5833}{3.4061} \qquad = 3.7048\sqrt{8.875 / 10.5}$$

$$= 2.2264 \qquad = 3.4061$$

2.4) เปรียบเทียบค่า t ที่ได้จากการเปิดตาราง กับ t ที่ได้จากการคำนวณ

t คำนวณ = 2.264 ซึ่งมากกว่า t(0.05,5) = 2.015 ดังนั้นปฏิเสธ H_0

2.5) สรุปคำตอบ

จำนวนห้องน้ำมีความสัมพันธ์ในทิศทางเดียวกับราคาตึกแถว ณ ระดับนัยสำคัญ 0.05