

บทที่ 11

การวิเคราะห์การถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ

1. ตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ

ในบทก่อนเราอธิบายการวิเคราะห์การถดถอยที่มีตัวแปรตามเป็นตัวแปรเชิงปริมาณและตัวแปรอิสระเป็นตัวแปรเชิงปริมาณเท่านั้น สำหรับในบทนี้จะอธิบายการวิเคราะห์การถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ โดยการใช้ตัวแปรดัมมี่ในสมการถดถอย

ตัวแปรเชิงคุณภาพซึ่งไม่สามารถวัดค่าออกมาได้ว่ามากหรือน้อย ตัวอย่างเช่น เพศ เชื้อชาติ ศาสนา อาชีพ เป็นต้น เราจึงกำหนดค่าที่แตกต่างกันให้แก่แต่ละกลุ่มที่แตกต่างกันของตัวแปรเชิงคุณภาพที่เป็นตัวแปรแบบแบ่งประเภท (nominal variable) โดยปกติจะกำหนดค่าให้เป็น 0 และ 1 ซึ่งไม่ได้มีความหมายเป็นการวัดค่าแต่หมายถึงชั้นต่าง ๆ ที่สนใจของตัวแปรนั้น และเรียกว่าตัวแปรดัมมี่

ตัวอย่างของตัวแปรดัมมี่เช่น ตัวแปรเพศ เชื้อชาติ ศาสนา และอาชีพ ทำให้เป็นตัวแปรดัมมี่ได้คือ

$$\begin{aligned} X_1 &= \begin{cases} 1 & \text{ถ้าตัวอย่างเป็นผู้หญิง} \\ 0 & \text{ถ้าตัวอย่างเป็นผู้ชาย} \end{cases} \\ X_2 &= \begin{cases} 1 & \text{ถ้าตัวอย่างเป็นเชื้อชาติไทย} \\ 0 & \text{ถ้าตัวอย่างเป็นเชื้อชาติอื่น ๆ} \end{cases} \\ X_3 &= \begin{cases} 1 & \text{ถ้าตัวอย่างนับถือศาสนาพุทธ} \\ 0 & \text{ถ้าตัวอย่างนับถือศาสนาอื่น ๆ} \end{cases} \\ X_4 &= \begin{cases} 1 & \text{ถ้าตัวอย่างเป็นข้าราชการ} \\ 0 & \text{ถ้าตัวอย่างมีอาชีพอื่น ๆ} \end{cases} \end{aligned}$$

สำหรับตัวแปรแบบแบ่งประเภทที่สนใจศึกษาที่มีจำนวนชั้นที่สนใจมากกว่า 2 ชั้น ตัวอย่างเช่น ตัวแปรศาสนา (พุทธ, คริสต์, อิสลาม) , ตัวแปรอาชีพ (ข้าราชการ, รัฐวิสาหกิจ,

รับจ้าง, ส่วนตัว) ทำให้เป็นตัวแปรคัมมีได้โดยมีกฎคือ ถ้าตัวแปรแบบแบ่งประเภทที่สนใจศึกษา มี k ชั้น สามารถกำหนดตัวแปรคัมมีได้ $k-1$ ตัว

ตัวอย่างของตัวแปรศาสนา (พุทธ, คริสต์, อิสลาม) ทำให้เป็นตัวแปรคัมมี X_1 และ X_2 คือ

$$X_1 = \begin{cases} 1 & \text{ถ้าตัวอย่างนับถือศาสนาพุทธ} \\ 0 & \text{ถ้าตัวอย่างนับถือศาสนาอื่น ๆ} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{ถ้าตัวอย่างนับถือศาสนาคริสต์} \\ 0 & \text{ถ้าตัวอย่างนับถือศาสนาอื่น ๆ} \end{cases}$$

ตัวอย่างของตัวแปรอาชีพ (ข้าราชการ, รัฐวิสาหกิจ, รับจ้าง, ส่วนตัว) ทำให้เป็นตัวแปรคัมมี มี $X_3, X_4,$ และ X_5 คือ

$$X_3 = \begin{cases} 1 & \text{ถ้าตัวอย่างมีอาชีพข้าราชการ} \\ 0 & \text{ถ้าตัวอย่างมีอาชีพอื่น ๆ} \end{cases}$$

$$X_4 = \begin{cases} 1 & \text{ถ้าตัวอย่างมีอาชีพรัฐวิสาหกิจ} \\ 0 & \text{ถ้าตัวอย่างมีอาชีพอื่น ๆ} \end{cases}$$

$$X_5 = \begin{cases} 1 & \text{ถ้าตัวอย่างมีอาชีพรับจ้าง} \\ 0 & \text{ถ้าตัวอย่างมีอาชีพอื่น ๆ} \end{cases}$$

เมื่อ $X_1 = 1, X_2 = 0$ หมายถึง นับถือศาสนาพุทธ

$X_1 = 0, X_2 = 1$ หมายถึง นับถือศาสนาคริสต์

$X_1 = 0, X_2 = 0$ หมายถึง นับถือศาสนาอิสลาม

และ $X_3 = 1, X_4 = 0, X_5 = 0$ หมายถึง อาชีพข้าราชการ

$X_3 = 0, X_4 = 1, X_5 = 0$ หมายถึง อาชีพรัฐวิสาหกิจ

$X_3 = 0, X_4 = 0, X_5 = 1$ หมายถึง อาชีพรับจ้าง

$X_3 = 0, X_4 = 0, X_5 = 0$ หมายถึง อาชีพส่วนตัว

2. ตัวแบบการถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ

ตัวอย่างการศึกษาเกี่ยวกับปัจจัยที่มีความสัมพันธ์กับน้ำหนักแรกเกิด วัตถุประสงค์คือ อยากรู้อยากทราบว่า การสูบบุหรี่มีอิทธิพลต่อน้ำหนักแรกเกิดหรือไม่ ข้อมูลได้มาจากการสุ่มตัวอย่างน้ำหนักแรกเกิดที่ลูกบันทึกไว้แล้วจำนวน 32 คน แต่ละตัวอย่างมีตัวแปรอิสระ 2 ตัวแปรคือ

จำนวนสัปดาห์ที่เด็กอยู่ในครรภ์มารดา (Gestation) ซึ่งเป็นตัวแปรเชิงปริมาณและสภาพการ
สูบบุหรี่ของมารดาซึ่งเป็นตัวแปรเชิงคุณภาพ (จาก Daniel, 1995) ข้อมูลอยู่ในตาราง
ต่อไปนี้คือ

ตารางที่ 11.1 ข้อมูลน้ำหนักเด็กแรกเกิดจำนวน 32 คน

ตัวอย่าง	Y น้ำหนักแรกเกิด (กรัม)	X ₁ ระยะเวลาที่เด็กอยู่ในครรภ์มารดา (สัปดาห์)	X ₂ สภาพการสูบบุหรี่ของมารดา (S = สูบ, N =ไม่สูบ)
1	2940	38	S
2	3130	38	N
3	2420	36	S
4	2450	34	N
5	2760	39	S
6	2440	35	S
7	3226	40	N
8	3301	42	S
9	2729	37	N
10	3410	40	N
11	2715	36	s
12	3095	39	N
13	3130	39	S
14	3244	39	N
15	2520	35	N
16	2928	39	S
17	3523	41	N
18	3446	42	S
19	2920	38	N
20	2957	39	S
21	3530	42	N
22	2580	38	S
23	3040	37	N
24	3500	42	S
25	3200	41	S
26	3322	39	N
27	3459	40	N
28	3346	42	S
29	2619	35	N
30	3175	41	S
31	2740	38	S
32	2841	36	N

ในการวิเคราะห์การถดถอยเราจะกำหนดตัวแปรตามและตัวแปรอิสระสำหรับตัวอย่างนี้ ตัวแปรตามคือ น้ำหนักแรกเกิด แทนด้วย Y ตัวแปรอิสระ มี 2 ตัว คือ ตัวแปรระยะเวลาที่เด็กอยู่

ในกรณีมารดา แทนด้วย X_1 และตัวแปรสภาพการสูบบุหรี่ของมารดาให้เป็นตัวแปรดัมมี่ได้ คือ

$$X_2 = \begin{cases} 0 & \text{ถ้ามารดาไม่สูบบุหรี่} \\ 1 & \text{ถ้ามารดาสูบบุหรี่} \end{cases}$$

ดังนั้นตัวแบบการถดถอยของการศึกษานี้คือ

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

สำหรับ X_2 ที่เป็นตัวแปรเชิงคุณภาพ

β_0 คือ ค่าเฉลี่ยของน้ำหนักแรกเกิดสำหรับกลุ่มมารดาที่ไม่สูบบุหรี่ (μ_N) โดยที่ X_2 มีค่าที่เป็นไปได้ 2 ระดับคือ (0 = ไม่สูบบุหรี่ และ 1 = สูบบุหรี่) โดยที่ β_0 จะแทนค่าเฉลี่ยของตัวแปรตามที่มีตัวแปรคุณภาพที่ถูกกำหนดค่าให้เป็น 0 เสมอ เรียกว่า ระดับฐาน (base level)

β_2 คือ ความแตกต่างระหว่างค่าเฉลี่ยของน้ำหนักแรกเกิดของกลุ่มมารดาที่สูบบุหรี่ และกลุ่มมารดาที่ไม่สูบบุหรี่ ($\mu_S - \mu_N$)

$$\text{โดยที่ } \mu_S - \mu_N = (\beta_0 + \beta_2) - \beta_0 = \beta_2$$

สำหรับ X_1 ที่เป็นตัวแปรเชิงปริมาณ

β_1 คือ สัมประสิทธิ์การถดถอยบางส่วน แทนค่าการเปลี่ยนแปลงของ y สำหรับ X_1 ที่เปลี่ยนแปลงไป 1 หน่วย เมื่อตัวแปร X_2 คงที่ สมการทำนายของตัวแบบคือ

$$\hat{Y} = b_0 + b_1 X_1 + b_2 X_2$$

การพิจารณาความหมายของพารามิเตอร์ของตัวแบบนี้ซึ่งได้แก่ สัมประสิทธิ์การถดถอย (β_i) กรณีที่มารดาไม่สูบบุหรี่ ($X_2 = 0$) จะได้สมการทำนาย ซึ่งเป็นสมการเส้นตรงคือ

$$\begin{aligned} \hat{Y} &= b_0 + b_1 X_1 + b_2(0) \\ &= b_0 + b_1 X_1 \end{aligned}$$

และกรณีที่มารดาสูบบุหรี่ ($X_2 = 1$) จะได้สมการทำนายซึ่งเป็นสมการเส้นตรงด้วยเหมือนกันคือ

^

$$Y = b_0 + b_1X_1 + b_2(1) = (b_0 + b_2) + b_1X_1$$

3. การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอย

จากการศึกษาเรื่องนี้ให้นักแรกเกิด สมมติว่าข้อมูลเป็นไปตามข้อตกลงเบื้องต้นของการวิเคราะห์การถดถอยพหุ การทดสอบสมมติฐานเกี่ยวกับสัมประสิทธิ์การถดถอยที่อยู่ในตัวแบบการถดถอยพหุ มี 2 แบบ คือ

(1) การทดสอบเกี่ยวกับสัมประสิทธิ์การถดถอยทั้งหมดในตัวแบบ คือ $H_0 : \beta_1 = \beta_2 = 0$ คู่กับ $H_1 : \text{มี } \beta_k \text{ อย่างน้อย 1 ค่า ไม่เท่ากับ } 0 ; k = 1, 2$ สถิติทดสอบคือ

$$F^* = \frac{SSR(X_1, X_2)/2}{SSE(X_1, X_2)/(32 - 3)} = \frac{MSR}{MSE}$$

ซึ่งมีการแจกแจงแบบ F ที่มีจำนวนชั้นอิสระ 2 , และ 29 ที่ $\alpha = .05$ จุดวิกฤติคือ $F_{.95,2,29} = 3.33$ จะสรุปว่าปฏิเสธ H_0 ถ้าค่าสถิติ F^* มากกว่าจุดวิกฤติ

(2) การทดสอบสัมประสิทธิ์การถดถอยทีละตัวคือ $H_0 : \beta_2 = 0$ คู่กับ $H_1 : \beta_2 \neq 0$ เพื่อทดสอบว่าการสูบบุหรี่มีอิทธิพลต่อน้ำหนักแรกเกิดหรือไม่ สถิติทดสอบคือ

$$t^* = \frac{b_2 - 0}{s_{b_2}}$$

ซึ่งมีการแจกแจงแบบ t ที่มีจำนวนชั้นอิสระเท่ากับ $n - k - 1 = 32 - 2 - 1 = 29$ ที่ $\alpha = .05$ จุดวิกฤติคือ $t_{.025,29} = 2.0452$ จะสรุปว่าปฏิเสธ H_0 ถ้าค่าสถิติ t^* มากกว่าหรือเท่ากับ 2.0452 หรือน้อยกว่าหรือเท่ากับ -2.0452

ช่วงความเชื่อมั่นของ β_2 คือ $b_2 \pm ts_{b_2}$

ที่ $\alpha = .05$ จะสรุปว่าการสูบบุหรี่มีอิทธิพลต่อน้ำหนักแรกเกิด ถ้าช่วงความเชื่อมั่นของ β_2 ไม่คลุมค่าศูนย์

4. การวิเคราะห์การถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ

ในบทที่ 9 ได้มีการอธิบายขั้นตอนการวิเคราะห์การถดถอยไว้แล้ว ดังนั้นในบทนี้จึงจะแสดงการวิเคราะห์การถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ ข้อมูลอยู่ในแฟ้มข้อมูลชื่อ `dummyvar.sav` มีขั้นตอนการใช้โปรแกรม SPSS ช่วยในการวิเคราะห์ข้อมูลดังนี้

(1) ไปที่เมนูบาร์ คลิกที่ **Graphs** , **Scatter...** จะได้หน้าต่าง **Scatterplot**

คลิกที่คำสั่ง **Simple** แล้วคลิกที่ปุ่ม **Define** จะได้หน้าต่าง **Simple Scatterplot**

(2) ในหน้าต่าง **Simple Scatterplot**

คลิกที่ตัวแปรตาม `bweight` ให้ย้ายไปอยู่ในช่อง **Y axis :**

คลิกที่ตัวแปรอิสระ `gestat` ให้ย้ายไปอยู่ในช่อง **X axis :**

คลิกที่ตัวแปรอิสระ `smoke` ให้ย้ายไปอยู่ในช่อง **Label Cases by :** เพื่อต้องการให้แบ่งการกระจายของข้อมูลออกเป็นกลุ่ม ๆ คือ กลุ่มที่มารดาสูบบุหรี่ และกลุ่มที่มารดาไม่สูบบุหรี่

แล้วคลิกปุ่ม **Options...** จะได้หน้าต่าง **Options....**

คลิกที่คำสั่ง **Display chart with case labels** เพื่อให้ระบุแต่ละจุดในกราฟด้วยว่าเป็นข้อมูลของกลุ่มใด สำหรับตัวอย่างนี้ กลุ่มที่มารดาสูบบุหรี่คือ 1 และกลุ่มที่มารดาไม่สูบบุหรี่คือ 0

แล้วคลิกปุ่ม **Continue** หน้าต่างนี้จะถูกปิดไป

แล้วคลิกปุ่ม **OK** จะได้ผลลัพธ์เป็นกราฟการกระจายของข้อมูลของกลุ่มที่มารดาสูบบุหรี่และกลุ่มที่มารดาไม่สูบบุหรี่ ดังรูปที่ 11.1

(3) ไปที่เมนูบาร์ คลิกที่ **Analyze, Regression, Linear ...** จะได้หน้าต่าง **Linear Regression**

คลิกที่ตัวแปรตาม `bweight` ให้ย้ายไปอยู่ในช่อง **Dependent :**

คลิกที่ตัวแปรอิสระ `gestat` และ `smoke` ให้ย้ายเข้าไปอยู่ในช่อง **Independent(s) :**

ที่คำสั่งย่อย **Method :** เลือก **Enter** ซึ่งโดยปกติโปรแกรมจะเลือกให้อยู่แล้ว

คลิกที่ปุ่ม **Statistics ...** จะได้หน้าต่าง **Linear Regression : Statistics**

(4) ในหน้าต่าง **Linear Regression : Statistics**

ในกรอบคำสั่ง **Regression Coefficients** เลือกคำสั่ง **Estimates** และคำสั่ง **Confidence intervals**

เลือกคำสั่ง **Model fit**

แล้วคลิกปุ่ม **Continue** หน้าต่างนี้จะถูกปิดไป

(5) ในหน้าต่าง **Linear Regression**

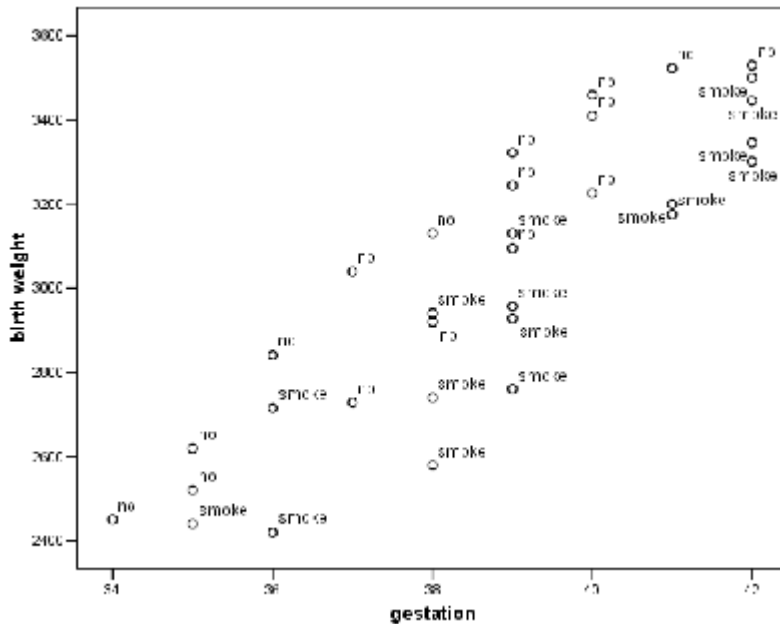
คลิกปุ่ม **OK** จะได้ผลลัพธ์เป็นผลการวิเคราะห์ความแปรปรวน (**ANOVA**) และค่าประมาณของสัมประสิทธิ์การถดถอยของตัวแบบการถดถอย ดังภาพที่ 11.2

5. ผลการวิเคราะห์การถดถอยที่มีตัวแปรอิสระเป็นตัวแปรเชิงคุณภาพ

จากการศึกษาเรื่องน้ำหนักแรกเกิด ได้ผลการวิเคราะห์คือ

(1) กราฟ แสดงการกระจายของข้อมูลระหว่างน้ำหนักแรกเกิดบนแกน **Y** และระยะเวลาที่เด็กอยู่ในครรภ์มารดาบนแกน **X** สำหรับกรณีที่มีมารดาไม่สูบบุหรี่ และกรณีที่มีมารดาสูบบุหรี่ ดังภาพที่ 11.1

Graph



ภาพที่ 11.1 กราฟการกระจายของข้อมูลระหว่างน้ำหนักแรกเกิดและระยะเวลาที่เด็กอยู่ในครรภ์มารดา

(2) ผลการวิเคราะห์การถดถอย ดังภาพที่ 11.2 ประกอบด้วย

- ตาราง Coefficients ที่แสดงสัมประสิทธิ์การถดถอย
- ตารางการวิเคราะห์ความแปรปรวนแบบทางเดียว ANOVA
- ตารางสรุปค่าสัมประสิทธิ์ต่าง ๆ ของสมการถดถอยอยู่ในตาราง Model Summary

Regression

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	smoking, ^a gestation ^a	.	Enter

a. All requested variables entered.

b. Dependent Variable: birth weight

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.947 ^a	.896	.889	115.530

a. Predictors: (Constant), smoking, gestation

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3348719.7	2	1674359.837	125.45	.000 ^a
	Residual	387069.827	29	13347.235		
	Total	3735789.5	31			

a. Predictors: (Constant), smoking, gestation

b. Dependent Variable: birth weight

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	-2389.573	349.206		-6.843	.000	-3103.779	-1675.366
	gestation	143.100	9.128	.963	15.677	.000	124.431	161.769
	smoking	-244.544	41.982	-.358	-5.825	.000	-330.406	-158.682

a. Dependent Variable: birth weight

ภาพที่ 11.2 ผลการวิเคราะห์การถดถอยของตัวอย่างเรื่องน้ำหนักแรกเกิด

6. การแปลความหมายการวิเคราะห์การถดถอย

(1) จากกราฟการกระจายของข้อมูลระหว่างน้ำหนักแรกเกิดและระยะเวลาที่เด็กอยู่ในครรภ์มารดา แสดงว่ามีลักษณะความสัมพันธ์เป็นเชิงเส้นตรงทั้ง 2 กรณี คือ มารดาไม่สูบบุหรี่ และมารดาสูบบุหรี่

(2) จากผลการวิเคราะห์การถดถอยได้สมการถดถอยคือ

$$\hat{bweight} = -2389.573 + 143.100 \text{ gestat} - 244.544 \text{ smoke}$$

ทดสอบความสัมพันธ์ของตัวแปรตาม bweight กับตัวแปรอิสระ gestat และ smoke สมมติฐานที่ต้องการทดสอบคือ $H_0 : \beta_{\text{gestat}} = \beta_{\text{smoke}} = 0$

ในการทดสอบสมมติฐานเราใช้การวิเคราะห์ความแปรปรวนซึ่งเป็นการเปรียบเทียบสัดส่วนของความแปรปรวนที่อธิบายได้ด้วยตัวแปรอิสระในตัวแบบการถดถอย กับความแปรปรวนที่อธิบายไม่ได้จากตัวแบบการถดถอย สัดส่วนของความแปรปรวนนี้คือ ค่าสถิติทดสอบ F ใช้เพื่อทดสอบว่าตัวแบบการถดถอยสามารถทำนายผลลัพธ์ได้ดีเพียงใด โดยเปรียบเทียบกับความคลาดเคลื่อนภายในตัวแบบ

ผลการวิเคราะห์ความแปรปรวน ได้ค่าสถิติทดสอบ F ที่ $df = 2, 29$ เท่ากับ 125.446 และค่า Sig. เท่ากับ .000 จึงสรุปว่าปฏิเสธ H_0 นั่นคือ มี β อย่างน้อย 1 ค่าที่ไม่เท่ากับ 0 หมายความว่า ตัวแปรอิสระ gestat และ smoke มีอย่างน้อย 1 ตัวที่มีอิทธิพลกับตัวแปรตาม bweight

(3) ทดสอบสัมประสิทธิ์การถดถอยที่ละตัว คือ $H_0 : \beta_{gestat} = 0$ และ $H_0 : \beta_{smoke} = 0$ เพื่อทดสอบว่าระยะเวลาที่เด็กอยู่ในครรภ์มารดา และสภาพการสูบบุหรี่ของมารดา มีอิทธิพลต่อน้ำหนักแรกเกิดหรือไม่

ผลการทดสอบได้ค่าสถิติทดสอบ $t = 15.677$ และ $t = -5.825$ ตามลำดับ และค่า Sig. เท่ากับ .000 และ .000 ตามลำดับ จึงสรุปว่า ปฏิเสธ H_0 ทั้ง 2 การทดสอบ นั่นคือ ระยะเวลาที่เด็กอยู่ในครรภ์มารดา และสภาพการสูบบุหรี่ของมารดามีอิทธิพลต่อน้ำหนักแรกเกิด

(4) เส้นถดถอยที่มีตัวแปรอิสระ smoke เป็นตัวแปรคัมมี จากสมการถดถอย

$$\hat{bweight} = -2389.573 + 143.100 \text{ gestat} - 244.544 \text{ smoke}$$

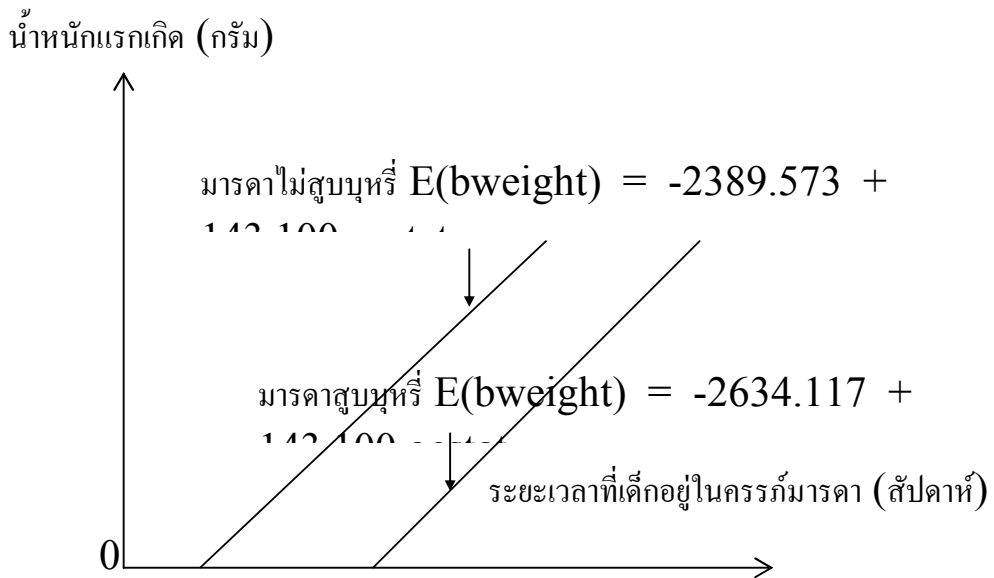
เมื่อตัวแปร smoke = 1 หมายถึง มารดาสูบบุหรี่จะได้สมการถดถอยคือ

$$\begin{aligned} \hat{bweight} &= -2389.573 + 143.100 \text{ gestat} - 244.544(1) \\ &= (-2389.573 - 244.544) + 143.100 \text{ gestat} \\ &= -2634.117 + 143.100 \text{ gestat} \end{aligned}$$

เมื่อตัวแปร smoke = 0 หมายถึง มารดาไม่สูบบุหรี่จะได้สมการถดถอยคือ

$$\begin{aligned} \hat{bweight} &= -2389.573 + 143.100 \text{ gestat} - 244.544(0) \\ &= -2389.573 + 143.100 \text{ gestat} \end{aligned}$$

กราฟของสมการถดถอยของน้ำหนักแรกเกิดสำหรับมารดาที่สูบบุหรี่ และมารดาที่ไม่สูบบุหรี่ ดังภาพที่ 11.3



ภาพที่ 11.3 กราฟเส้นถดถอยของน้ำหนักแรกเกิดสำหรับมารดาที่สูบบุหรี่และมารดาที่ไม่สูบบุหรี่

(5) 95% ช่วงความเชื่อมั่นของ β_{gestat} และ β_{smoke} ค่าประมาณของความแตกต่างเฉลี่ยของน้ำหนักแรกเกิดเมื่อระยะเวลาที่เด็กอยู่ในครรภ์มารดาคงที่ เท่ากับ 244.5 กรัม ระหว่างมารดาที่สูบบุหรี่แตกต่างจากมารดาที่ไม่สูบบุหรี่ เราสามารถหาช่วงของความแตกต่างของค่าเฉลี่ยระหว่างมารดาที่สูบบุหรี่และมารดาที่ไม่สูบบุหรี่ได้คือ

β	lower bound	upper bound
gestat	124.431	161.769
smoke	-330.406	-158.682

แสดงให้เห็นว่าช่วงความเชื่อมั่นของ β_{gestat} และ β_{smoke} ไม่คลุมค่าศูนย์ นั่นคือ ระยะเวลาที่เด็กอยู่ในครรภ์มารดามีอิทธิพลต่อน้ำหนักแรกเกิด และความแตกต่างระหว่างมารดาที่สูบบุหรี่ และไม่สูบบุหรี่ อยู่ระหว่าง 159 กรัม และ 330 กรัม

แบบฝึกหัดบทที่ 11

1. จากการศึกษาเกี่ยวกับผลิตภัณฑ์ที่ใช้ไล่แมลง (ที่รายงานโดย Consumer Reports, June. 2000 อ้างถึงใน Mendenhall, W. and Sincich, T. 2003) ผู้วิจัยอยากทราบว่าผลิตภัณฑ์ของบริษัทใดที่สามารถป้องกันการรบกวนจากยุงได้ดีที่สุด โดยทำการทดลองกับผลิตภัณฑ์จากบริษัทต่าง ๆ 14 บริษัทที่อ้างว่ามีประสิทธิภาพในการไล่ยุงได้ ทำการทดลองโดยแยกชนิดของผลิตภัณฑ์ทั้งหมดก่อน แบ่งออกได้ 2 แบบ คือ แบบครีมที่ใช้ทาและแบบฉีดพ่นเป็นละออง เก็บข้อมูลเกี่ยวกับค่าใช้จ่ายของผลิตภัณฑ์ต่อปริมาณ 1/3 ออนซ์ สำหรับการใช้ 1 ครั้ง แล้ววัดประสิทธิภาพของผลิตภัณฑ์เป็นจำนวนชั่วโมงที่สามารถป้องกันการรบกวนจากยุง กำหนดให้หน่วยของเวลานับเพิ่มขึ้นทีละครึ่งชั่วโมงทำการทดลองกับคนโดยวิธีการทาหรือฉีดพ่นผลิตภัณฑ์ที่เบน แล้วทดสอบกับยุง 200 ตัว แล้วบันทึกข้อมูลได้ดังตาราง

ตาราง ข้อมูลลักษณะของผลิตภัณฑ์ ค่าใช้จ่าย และจำนวนชั่วโมงที่สามารถป้องกันการรบกวนจากยุงของผลิตภัณฑ์จากบริษัทต่าง ๆ 14 บริษัท

ผลิตภัณฑ์	ลักษณะ	ค่าใช้จ่าย (ดอลลาร์)	จำนวนชั่วโมงที่สามารถ ป้องกันการรบกวน
1	ครีม	2.08	13.5
2	พ่น	0.67	0.5
3	ครีม	1.00	2.0
4	ครีม	0.75	7.0
5	ครีม	0.46	3.0
6	พ่น	0.11	6.0
7	พ่น	0.22	3.0
8	พ่น	0.19	5.5
9	พ่น	0.24	6.5
10	พ่น	0.27	1.0
11	ครีม	1.77	14.0
12	ครีม	0.67	3.0
13	ครีม	0.36	7.0
14	พ่น	2.75	24.0

แหล่งที่มา : “Buzz off.” Consumer Reports, June 2000. อ้างถึงใน Mendenhall, W. and Sincich, T. 2003.

- ก. จงสร้างตัวแปรดัมมี่สำหรับลักษณะของผลิตภัณฑ์ แล้วเขียนตัวแบบการถดถอยของค่าใช้จ่ายที่ถดถอยบนลักษณะของผลิตภัณฑ์
- ข. จงใช้ข้อมูลในตารางเพื่อหาสมการถดถอยตามตัวแบบในข้อ ก.
- ค. จงทดสอบสมมติฐานเกี่ยวกับสมการถดถอยที่ได้จากข้อ ข. และแปลความหมาย
- ง. จงทดสอบความเหมาะสมของเส้นถดถอยเชิงเส้นตรงที่ได้จากข้อ ค.
- จ. ทำซ้ำ ข้อ ก. ถึง ข้อ ง. อีกครั้งโดยให้ตัวแปรตามคือจำนวนชั่วโมงที่สามารถป้องกัน การรบกวนจากยูง
- ฉ. จงเขียนสมการถดถอยของจำนวนชั่วโมงที่สามารถป้องกันการรบกวนจากยูงสำหรับผลิตภัณฑ์ที่เป็นครีมและผลิตภัณฑ์ที่ใช้ฟัน และแปลความหมาย
2. ในการศึกษาประเมินการทำงานของเครื่องจักรดีเซลที่ใช้เชื้อเพลิงแตกต่างกัน ผู้วิจัยทำการทดลองเชื้อเพลิง 3 ชนิดคือ DF-2 , Blended , และ Advanced Timing ดำเนินการทดลองกับ เครื่องจักรที่มีกำลังแรงม้าแตกต่างกัน 5 ระดับ คือ 4, 6, 8, 10, และ 12 กิโลวัตต์ แล้วเก็บข้อมูล อัตราการเผาไหม้ของเชื้อเพลิง ได้ข้อมูลดังตาราง
- ถ้ากำหนดให้
- y คือ อัตราการเผาไหม้
- X_1 คือ กำลังแรงม้าเป็นกิโลวัตต์
- $$X_2 = \begin{cases} 1 & \text{ถ้าเชื้อเพลิงเป็น DF-2} \\ 0 & \text{ถ้าไม่ใช่} \end{cases}$$
- $$X_3 = \begin{cases} 1 & \text{ถ้าเชื้อเพลิงเป็น blended} \\ 0 & \text{ถ้าไม่ใช่} \end{cases}$$
- ก. จงเขียนตัวแบบการถดถอยสำหรับอัตราการเผาไหม้ที่ถดถอยบนกำลังแรงม้า และชนิดของเชื้อเพลิง
- ข. จงใช้ข้อมูลในตารางเพื่อหาสมการถดถอยตามตัวแบบในข้อ ก.
- ค. จงทดสอบสมมติฐานเกี่ยวกับสมการถดถอยที่ได้จากข้อ ข. และแปลความหมาย
- ง. จงทดสอบความเหมาะสมของเส้นถดถอยเชิงเส้นตรงที่ได้จากข้อ ค.
- จ. จงเขียนสมการถดถอยของอัตราการเผาไหม้สำหรับเชื้อเพลิง DF-2 , blended , และ advanced timing

ตาราง กำลังแรงม้า (กิโลวัตต์) ชนิดของเชื้อเพลิง และอัตราการเผาไหม้ของเชื้อเพลิง

จำนวนแรงม้า	ชนิดของเชื้อเพลิง	อัตราการเผาไหม้
4	DF-2	13.2
4	Blended	17.5
4	Advanced Timing	17.5
6	DF-2	26.1
6	Blended	32.7
6	Advanced Timing	43.5
8	DF-2	25.9
8	Blended	46.3
8	Advanced Timing	45.6
10	DF-2	30.7
10	Blended	50.8
10	Advanced Timing	68.9
12	DF-2	32.3
12	Blended	57.1

แหล่งที่มา : Litzinger, T.A., and Buzza, T.G. "Performance and emissions of a diesel engine using a coal-derived fuel." *Journal of Energy Resources Technology*, Vol. 112, Mar. 1990, p.32, Table 3. อ้างถึงใน Mendenhall, W. and Sincich, T. 2003.