

เฉลยแบบฝึกหัดบทที่ 6

6.1 จากชายฉิวขาว 32 คน ซึ่งมีอายุเกิน 40 ปี ได้บันทึกข้อมูลต่อไปนี้ไว้คือ

(Y) SBP = ความดันโลหิต

(X_1) AGE = อายุ

(X_2) SMK = ประวัติการสูบบุหรี่ (=1 ถ้ายังสูบบุหรี่อยู่หรือเคยสูบบุหรี่)

(X_3) QUET = ขนาดของร่างกาย (QUET = "Quetelet index" คืออัตราส่วนวัดขนาดของร่างกาย โดยที่ $QUET = 100(\text{weight}/\text{height}^2)$)

คนที่	Y	X_1	X_2	X_3	คนที่	Y	X_1	X_2	X_3
1	135	45	0	2.876	17	145	49	1	3.360
2	122	41	0	3.251	18	142	46	1	3.024
3	130	49	0	3.100	19	135	57	0	3.171
4	148	52	0	3.768	20	142	56	0	3.401
5	146	54	1	2.979	21	150	56	1	3.628
6	129	47	1	2.790	22	144	58	0	3.751
7	162	60	1	3.668	23	137	53	0	3.296
8	160	48	1	3.612	24	132	50	0	3.210
9	144	44	1	2.368	25	149	54	1	3.301
10	180	64	1	4.637	26	132	48	1	3.017
11	166	59	1	3.877	27	120	43	0	2.789
12	138	51	1	4.032	28	126	43	1	2.956
13	152	64	0	4.116	29	161	63	0	3.800
14	138	56	0	3.673	30	170	63	1	4.132
15	140	54	1	3.562	31	152	62	0	3.962
16	134	50	1	2.998	32	164	65	0	4.010

ในการศึกษาถึงความสัมพันธ์ในรูปการถดถอยแบบพหุคูณของ SBP(Y) กับ AGE(X_1), SMK(X_2) และ QUET(X_3) ได้พิจารณาแบบจำลองการถดถอยแบบพหุคูณ 3 แบบจำลองคือ

- 1) y บน x_1
- 2) y บน x_1, x_2
- 3) y บน x_1, x_2, x_3

ได้ค่าประมาณแบบกำลังสองน้อยที่สุด และตาราง ANOVA ต่อไปนี้

แบบจำลอง	ตัวแปรอิสระ ในแบบจำลอง	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$S_{F_1}^2$	$S_{F_2}^2$	$S_{F_3}^2$
1	AGE (X_1)	59.092	1.605	-	-	.2387	-	-
2	AGE (X_1), SMK (X_2)	48.050	1.709	10.294	-	.2018	2.7681	-
3	AGE (X_1), SMK (X_2), QUET(X_3)	45.103	1.213	9.946	8.592	.3238	2.6561	4.4987

ตาราง ANOVA

แบบจำลอง 1		
SV	df	SS
Regression (X_1)	1	3861.630
Error	30	2564.338

แบบจำลอง 2		
SV	df	SS
Regression (X_1, X_2)	2	4689.684
Error	29	1736.285

แบบจำลอง 3		
SV	df	SS
Regression (X_1, X_2, X_3)	3	4889.826
Error	28	1536.143

- 1) จงใช้แบบจำลอง 3
 - 1.1 ทำนาย SBP ของชายอายุ 50 ปี สวมบุหรี่ และ ขนาดของร่างกาย = 3.5
 - 1.2 ทำนาย SBP ของชายอายุ 50 ปี ไม่สวมบุหรี่ และ ขนาดของร่างกาย = 3.5
 - 1.3 สำหรับชายอายุ 50 ปี สวมบุหรี่ จงประมาณความแตกต่างของ SBP เมื่อขนาดของร่างกายเพิ่มจาก 3.0 เป็น 3.5
- 2) จากตาราง ANOVA จงคำนวณหา R^2 ของแบบจำลองทั้ง 3
- 3) จงทำ overall F-test ของแต่ละแบบจำลอง ระบุ H_0 ที่ทดสอบให้ชัดเจนด้วย
- 4) จงทำ partial F-test สำหรับการ
 - 4.1 เพิ่ม SMK เข้าไปในแบบจำลองที่มี AGE อยู่แล้ว
 - 4.2 เพิ่ม QUET เข้าไปในแบบจำลองที่มี AGE และ SMK อยู่แล้ว
- 5) จงทำ t-test ซึ่งสมนัยกับ partial F-test ใน 4) และในแต่ละกรณีแสดงด้วยว่าค่าที่คำนวณได้สมนัยกันจริง ($t_c^2 = f_c$)

วิธีทำ

1) แบบจำลอง 3: ได้ $\hat{y} = 45.103 + 1.213x_1 + 9.946x_2 + 8.592x_3$

1.1) $x_1 = 50, x_2 = 1, x_3 = 3.5$

$$\begin{aligned} \text{ดังนั้น } \hat{y} &= 45.103 + 1.213(50) + 9.946(1) + 8.592(3.5) \\ &= 145.771 \end{aligned}$$

1.2) $x_1 = 50, x_2 = 0, x_3 = 3.5$

$$\begin{aligned} \text{ดังนั้น } \hat{y} &= 45.103 + 1.213(50) + 9.946(0) + 8.592(3.5) \\ &= 135.825 \end{aligned}$$

1.3) $x_1 = 50, x_2 = 1, x_3 = 3.5, \hat{y} = 145.771$ } ต่างกัน 4.296
 $x_1 = 50, x_2 = 1, x_3 = 3.0, \hat{y} = 141.475$ }

2) แบบจำลอง 1: $R^2 = 3861.630/6425.968 = 0.6010$

แบบจำลอง 2: $R^2 = 4689.684/6425.968 = 0.7298$

แบบจำลอง 3: $R^2 = 4889.826/6425.968 = 0.7609$

3) Overall F-Test

แบบจำลอง 1: $y = \beta_0 + \beta_1x_1 + e$

$$H_0: \beta_1 = 0$$

$$f_c = (3861.630/1)/(2564.338/30) = 45.177^{**}$$

$$f_{(1,30),.01} = 7.56$$

แบบจำลอง 2: $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + e$

$$H_0: \beta_1 = \beta_2 = 0$$

$$f_c = (4689.684/2)/(1736.285/29) = 39.164^{**}$$

$$f_{(2,29),.01} = 5.42$$

แบบจำลอง 3: $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + e$

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

$$f_c = (4889.826/3)/(1536.143/28) = 29.710^{**}$$

$$f_{(3, 28), .01} = 4.57$$

4) Partial F-Test

4.1) สำหรับการเพิ่ม SMK(X_2) เข้าไปในแบบจำลองที่มี AGE(X_1) อยู่แล้ว

$$\begin{aligned} \text{Partial } f_c &= SS(X_2 | X_1) / \text{MSE}(X_1, X_2) \\ &= [SS(X_1, X_2) - SS(X_1)] / \text{MSE}(X_1, X_2) \\ &= [4689.684 - 3861.631] / (1736.285/29) \\ &= 13.83^{**} \end{aligned}$$

$$f_{(1, 29), .01} = 7.6$$

4.2) สำหรับการเพิ่ม QUET(X_3) เข้าไปในแบบจำลองที่มี AGE(X_1)

และ SMK(X_2) อยู่แล้ว

$$\begin{aligned} \text{Partial } f_c &= SS(X_3 | X_1, X_2) / \text{MSE}(X_1, X_2, X_3) \\ &= [SS(X_1, X_2, X_3) - SS(X_1, X_2)] / \text{MSE}(X_1, X_2, X_3) \\ &= [4889.826 - 4689.684] / (1536.143/28) \\ &= 3.648(\text{n.s.}) \end{aligned}$$

เราไม่ปฏิเสธ H_0 ดังนั้นไม่ควรเพิ่ม QUET เข้าไปในแบบจำลอง

$$[f_{(1, 28), .05} = 4.2, f_{(1, 28), .01} = 7.64]$$

5) แบบจำลอง 2: $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e$

$$H_0: \beta_2 = 0, H_1: \beta_2 \neq 0, CR: |T| > t_{29, .005} = 2.756$$

$$\alpha = .01, t_c = \hat{\beta}_2 / s_{\hat{\beta}_2} = 10.294 / 2.7681 = 3.719^{**}$$

$$t_c^2 = f_c = 13.83 \text{ (เทียบกับ 4.1)}$$

แบบจำลอง 3: $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

$$H_0: \beta_3 = 0, H_1: \beta_3 \neq 0, CR: |T| > t_{28, .005} = 2.763$$

$$\alpha = .01, t_c = \hat{\beta}_3 / s_{\hat{\beta}_3} = 8.592 / 4.4987 = 1.91(\text{n.s.})$$

$$t_c^2 = f_c = 3.648 \text{ (เทียบกับ 4.2)}$$

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 6.1 และ 6.6

```
SET DISK=ON.
DATA LIST FREE/ SBP,AGE,SMK,QUET.
BEGIN DATA.
135 45 0 2.876
122 41 0 3.251
130 49 0 3.1
148 52 0 3.768
146 54 1 2.979
129 47 1 2.79
162 60 1 3.668
160 48 1 3.612
144 44 1 2.368
180 64 1 4.637
166 59 1 3.877
138 51 1 4.032
152 64 0 4.116
138 56 0 3.673
140 54 1 3.562
134 50 1 2.998
145 49 1 3.36
142 46 1 3.024
135 57 0 3.171
142 56 0 3.401
150 56 1 3.628
144 58 0 3.751
137 53 0 3.296
132 50 0 3.21
149 54 1 3.301
132 48 1 3.017
120 43 0 2.789
126 43 1 2.956
161 63 0 3.8
170 63 1 4.132
152 62 0 3.962
164 65 0 4.01
END DATA.
REGRESSION DESCRIPTIVES=DEFAULTS SIG COV XPROD
/VARIABLES=SBP,AGE,SMK,QUET
/CRITERIA= POUT(.055)
/DEPENDENT=SBP
/METHOD=BACKWARD.
REGRESSION VARIABLES=SBP,AGE,SMK,QUET
/CRITERIA=PIN(.05)
/DEPENDENT=SBP
/METHOD=FORWARD.
REGRESSION VARIABLES=SBP,AGE,SMK,QUET
/CRITERIA=PIN(.05) POUT(.055)
/DEPENDENT=SBP
/METHOD=STEPWISE.
```

Output สำหรับแบบฝึกหัดข้อ 6.1 และ 6.6

Page 2

SPSS/PC+

2/1/91

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

	Mean	Std Dev	Label
SBP	144.531	14.398	
AGE	53.250	6.956	
SMK	.531	.507	
QUET	3.441	.497	

N of Cases = 32

Page 3

SPSS/PC+

2/1/91

*** MULTIPLE REGRESSION ***

Correlation, Covariance, 1-tailed Sig, Cross-Product:

	SBP	AGE	SMK	QUET
SBP	1.000	.775	.247	.742
	207.289	77.637	1.805	5.310
	.999	.000	.086	.000
	6425.969	2406.750	55.969	164.619
AGE	.775	1.000	-.139	.803
	77.637	48.387	-.492	2.776
	.000	.999	.223	.000
	2406.750	1500.000	-15.250	86.046
SMK	.247	-.139	1.000	-.071
	1.805	-.492	.257	-.018
	.086	.223	.999	.349
	55.969	-15.250	7.969	-.558

Page 4

SPSS/PC+

2/1/91

*** MULTIPLE REGRESSION ***

	SBP	AGE	SMK	QUET
QUET	.742	.803	-.071	1.000
	5.310	2.776	-.018	.247
	.000	.000	.349	.999
	164.619	86.046	-.558	7.660

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

- 1.. QUET
- 2.. SMK
- 3.. AGE

Multiple R .87232
 R Square .76095
 Adjusted R Square .73533
 Standard Error 7.40691

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	3	4889.82570	1629.94190
Residual	28	1536.14305	54.86225

F = 29.70972 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
QUET	8.59245	4.49868	.29666	1.910	.0664
SMK	9.94557	2.65606	.35023	3.744	.0008
AGE	1.21271	.32382	.58592	3.745	.0008
(Constant)	45.10319	10.76488		4.190	.0003

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 2. Method: Backward

Variable(s) Removed on Step Number
4.. QUET

Multiple R .85428
R Square .72980
Adjusted R Square .71117
Standard Error 7.73769

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	4689.68423	2344.84211
Residual	29	1736.28452	59.87188

F = 39.16433 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
SMK	10.29439	2.76811	.36252	3.719	.0009
AGE	1.70916	.20176	.82577	8.471	.0000
(Constant)	48.04960	11.12956		4.317	.0002

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
QUET	.29666	.33951	.34880	1.910	.0664

End Block Number 2 POUT = .055 Limits reached.

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 1. Method: Forward

Page 11 SPSS/PC+ 2/1/91

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

1.. AGE

Multiple R .77520
 R Square .60094
 Adjusted R Square .58764
 Standard Error 9.24543

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	3861.63038	3861.63038
Residual	30	2564.33837	85.47795

F = 45.17692 Signif F = .0000

Page 12 SPSS/PC+ 2/1/91

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
AGE	1.60450	.23872	.77520	6.721	.0000
(Constant)	59.09162	12.81626		4.611	.0001

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
SMK	.36252	.56825	.98054	3.719	.0009
QUET	.33665	.31778	.35559	1.805	.0815

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

2.. SMK

Multiple R .85428
R Square .72980
Adjusted R Square .71117
Standard Error 7.73769

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	4689.68423	2344.84211
Residual	29	1736.28452	59.87188

F = 39.16433 Signif F = .0000

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
AGE	1.70916	.20176	.82577	8.471	.0000
SMK	10.29439	2.76811	.36252	3.719	.0009
(Constant)	48.04960	11.12956		4.317	.0002

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
QUET	.29666	.33951	.34880	1.910	.0664

End Block Number 1 PIN = .050 Limits reached.

This procedure was completed at 12:00:52

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 1. Method: Stepwise

 Page 17 SPSS/PC+ 2/1/91

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

1.. AGE

Multiple R .77520
 R Square .60094
 Adjusted R Square .58764
 Standard Error 9.24543

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	3861.63038	3861.63038
Residual	30	2564.33837	85.47795

F = 45.17692 Signif F = .0000

 Page 18 SPSS/PC+ 2/1/91

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
AGE	1.60450	.23872	.77520	6.721	.0000
(Constant)	59.09162	12.81626		4.611	.0001

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
SMK	.36252	.56825	.98054	3.719	.0009
QUET	.33665	.31778	.35559	1.805	.0815

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

2.. SMK

Multiple R .85428
 R Square .72980
 Adjusted R Square .71117
 Standard Error 7.73769

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	4689.68423	2344.84211
Residual	29	1736.28452	59.87188

F = 39.16433 Signif F = .0000

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
AGE	1.70916	.20176	.82577	8.471	.0000
SMK	10.29439	2.76811	.36252	3.719	.0009
(Constant)	48.04960	11.12956		4.317	.0002

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
QUET	.29666	.33951	.34880	1.910	.0664

End Block Number 1 PIN = .050 Limits reached.

This procedure was completed at 12:01:05

6.2 จากคนไข้ 25 คน (with hyperlipoproteinemia) ได้บันทึกข้อมูลเกี่ยวกับ X_1 , X_2 และ Y = ระดับ plasma lipid ของโคเลสเตอรอลรวม ไว้ดังนี้

คนไข้คนที่	ระดับ plasma lipid (Y) (mg/100 ml)	น้ำหนัก (X_1) (kg)	อายุ (X_2) (years)
1	354	84	46
2	190	73	20
3	405	65	52
4	263	70	30
5	451	76	57
6	302	69	25
7	288	63	28
8	385	72	36
9	402	79	57
10	365	75	44
11	209	27	24
12	290	89	31
13	346	65	52
14	254	57	23
15	395	59	60
16	434	69	48
17	220	60	34
18	374	79	51
19	308	75	50
20	220	82	34
21	311	59	46
22	181	67	23
23	274	85	37
24	303	55	40
25	244	63	30

1) จากตาราง ANOVA ที่กำหนดให้สำหรับการถดถอยเชิงเส้นตรงของ y บน x_1 และของ y บน x_2 ตัวแปรอิสระ 2 ตัวคือ X_1 และ X_2 นั้น ตัวใดจะเป็นตัวทำนายค่า Y ได้ดีกว่า (ตัวใดสำคัญกว่า)

Y บน X_1		
SV	df	SS
Regression (X_1)	1	10231.7
Error	23	135144.3

Y บน X_2		
SV	df	SS
Regression (X_2)	1	101932.7
Error	23	43444.3

2) กำหนดสมการการทำนายของ y บน x_1 และ x_2 , y บน x_1 และ y บน x_2 ให้ดังนี้

$$a) \hat{y} = 77.983 + 0.417 x_1 + 5.217 x_2$$

$$b) \hat{y} = 199.2975 + 1.622 x_1$$

$$c) \hat{y} = 102.5751 + 5.321 x_2$$

จากแต่ละแบบจำลอง จงทำนายระดับ plasma lipid (Y) สำหรับคนไข้คนที่ 4 (ซึ่งมี $y = 263$, $x_1 = 70$, $x_2 = 30$) แล้วเปรียบเทียบกับค่าสังเกต (หา \hat{y}_4 เทียบ $y_4 = 263$)

3) กำหนด ANOVA ของแบบจำลองแรกใน 2) ให้ จงทำ overall F-test และทำ partial F-test ในการเพิ่ม X_1 เข้าไปในแบบจำลองที่มี X_2 อยู่แล้ว

SV	df	SS
Regression (X_1, X_2)	2	102570.8
Error	22	42806.2

4) คำนวณค่า R^2 สำหรับแต่ละแบบจำลองใน 2)

5) จากการทำในข้อ 1)-4) ท่านพอจะสรุปได้ไหมว่าแบบจำลองใดใน 3 แบบจำลองนั้น จะช่วยในการทำนาย Y ได้ดีที่สุด

วิธีทำ

1) แบบจำลอง $y = \beta_0 + \beta_1 X_1 + e$

จากตาราง ANOVA Y บน X_1

ทดสอบ $H_0: \beta_1 = 0, H: \beta_1 \neq 0$

$$\alpha = .01, CR: F > f_{(1, 23), .01} = 7.88$$

$$\alpha = .05, CR: F > f_{(1, 23), .05} = 4.28$$

$$f_c = 10231.7 / (135144.3 / 23) = 1.74 \text{ (n.s.)}$$

นั่นคือไม่ปฏิเสธ H_0 ที่ $\alpha = .05$

แบบจำลอง $y = \beta_0 + \beta_2 X_2 + e$

จากตาราง ANOVA Y บน X_2

ทดสอบ $H_0: \beta_2 = 0, H: \beta_2 \neq 0$

$$\alpha = .01, CR: F > f_{(1, 23), .01} = 7.88$$

$$f_c = 101932.7 / (43444.3 / 23) = 53.96^{**}$$

นั่นคือปฏิเสธ H_0 ที่ $\alpha = .01$

ดังนั้น X_2 จึงเป็นตัวทำนาย Y ได้ดีกว่า X_1

2) $x_1 = 70, x_2 = 30, y = 263$

จาก a) $\hat{y} = 77.983 + 0.417 x_1 + 5.217 x_2$

$x_1 = 70, x_2 = 30, \hat{y} = 263.473$ ----> ประมาณได้ใกล้เคียง

จาก b) $\hat{y} = 199.2975 + 1.622 x_1$

$x_1 = 70, x_2 = 30, \hat{y} = 312.838$ ----> ประมาณได้ไม่ใกล้เคียง

จาก c) $\hat{y} = 102.5751 + 5.321 x_2$

$x_1 = 70, x_2 = 30, \hat{y} = 262.205$ ----> ประมาณได้ใกล้เคียง

3) Overall F-Test

แบบจำลอง $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e$

$H_0: \beta_1 = \beta_2 = 0, f_c = (102570.8/2)/(42806.2/22) = 26.36^{**}$

$f_{(2, 22), .01} = 5.72$

นั่นคือปฏิเสธ H_0 ที่ $\alpha = .01$

Partial F-Test สำหรับการเพิ่ม X_1 เข้าไปในแบบจำลองที่มี X_2 อยู่แล้ว

$$\begin{aligned} \text{Partial } f_c &= SS(X_1 | X_2) / MS(X_1, X_2) \\ &= [SS(X_1, X_2) - SS(X_2)] / MS(X_1, X_2) \\ &= (102570.8 - 101932.7) / (42806.2/22) \\ &= 0.328 (\text{n.s.}) \end{aligned}$$

$f_{(1, 22), .01} = 7.95$

4) แบบจำลอง a): $R^2 = 102570.8/145377.0 = 0.706$

แบบจำลอง b): $R^2 = 10231.7/145377.0 = 0.070$

แบบจำลอง c): $R^2 = 101932.7/145377.0 = 0.701$

ส่วนของความแปรปรวนของ Y เนื่องจาก X_2 มีมาก การเพิ่ม X_1 เข้าไปในแบบจำลองจะช่วยทำนาย Y อีกเล็กน้อย

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 6.2

```
SET DISK=ON.
DATA LIST FREE/Y XI X2.
BEGIN DATA,
354 84 46 190 73 20 405 65 52 263 70 30 451 76 51
302 69 25 288 63 28 385 72 36 402 79 57 365 75 44
209 27 24 290 89 31 346 65 52 254 57 23 395 59 60
434 69 48 220 60 34 374 79 51 308 75 50 220 82 34
311 59 46 181 67 23 274 85 37 303 55 40 244 63 30
END DATA.
REGRESSION DESCRIPTIVES=DEFAULTS SIG COV XPROD
/VARIABLES=Y XI X2
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION /VARIABLES=Y XI
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION /VARIABLES=Y X2
/DEPENDENT=Y
/METHOD=ENTER.
```

Output สำหรับแบบฝึกหัดที่ 6.2

Page 2

SPSS/PC+

2/6/91

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

	Mean	Std Dev	Label
Y	310.720	77.829	
X1	68.680	12.727	
X2	39.120	12.249	

N of Cases = 25

Page 3

SPSS/PC+

2/6/91

*** MULTIPLE REGRESSION ***

Correlation, Covariance, 1-tailed Sig, Cross-Product:

	Y	X1	X2
Y	1.000	.265	.837
	6057.377	262.782	798.243
	.999	.100	.000
	145377.040	6306.760	19157.840
X1	.265	1.000	.240
	262.782	161.977	37.415
	.100	.999	.124
	6306.760	3887.440	897.960
X2	.837	.240	1.000
	798.243	37.415	150.027
	.000	.124	.999
	19157.840	897.960	3600.640

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number

1.. X2
2.. X1

Multiple R .83997
R Square .70555
Adjusted R Square .67878
Standard Error 44.11051

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	102570.81465	51285.40733
Residual	22	42806.22535	1945.73752

F = 26.35782 Signif F = .0000

Page 5

SPSS/PC+

2/6/91

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	5.21659	.75724	.82097	6.889	.0000
X1	.41736	.72878	.06825	.573	.5727
(Constant)	77.98254	52.42964		1.487	.1511

End Block Number 1 All requested variables entered.

**** MULTIPLE REGRESSION ****

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X1

Multiple R .26529
 R Square .07038
 Adjusted R Square .02996
 Standard Error 76.65431

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	10231.72620	10231.72620
Residual	23	135145.31380	5875.88321

F = 1.74131 Signif F = .2000

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X1	1.62234	1.22943	.26529	1.320	.2000
(Constant)	199.29750	85.81792		2.322	.0294

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X2

Multiple R .83735
 R Square .70116
 Adjusted R Square .68817
 Standard Error 43.46131

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	101932.66571	101932.66571
Residual	23	43444.37429	1888.88584

F = 53.96444 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	5.32068	.72429	.83735	7.346	.0000
(Constant)	102.57514	29.63757		3.461	.0021

End Block Number 1 All requested variables entered.

5) จากผลที่ได้ในข้อ 1)-4) สรุปว่าแบบจำลองที่สมควรเป็นแบบจำลองที่มี X_2 ตัวเดียว (แบบจำลอง c)

6.3 นักสังคมวิทยาได้ทำการสำรวจเพื่อดูการเปลี่ยนแปลงของอัตราการฆ่าตัวตายในสหรัฐอเมริกา โดยพิจารณาจากอัตราการฆ่าตัวตาย (ต่อ 1000,000 คน) Y ว่าสัมพันธ์กับขนาดของประชากร X_1 , เปอร์เซ็นต์ของครอบครัวที่มีรายได้ต่อปีต่ำกว่า 5000 เหรียญ X_2 และอัตราการว่างงาน X_3 ของประชากรในเมือง ข้อมูลต่อไปนี้เป็นตัวอย่างจากเมือง 20 เมือง

เมือง	Y	X_1 (พัน)	X_2	X_3	เมือง	Y	X_1 (พัน)	X_2	X_3
1	11.2	587	16.5	6.2	11	14.5	7895	18.1	6.0
2	13.4	643	20.5	6.4	12	26.9	762	23.1	7.4
3	40.7	635	26.3	9.3	13	15.7	2793	19.1	5.8
4	5.3	692	16.5	5.3	14	36.2	741	24.7	8.6
5	24.8	1248	19.2	7.3	15	18.1	625	18.6	6.5
6	12.7	643	16.5	5.9	16	28.9	854	24.9	8.3
7	20.9	1964	20.2	6.4	17	14.9	716	17.9	6.7
8	35.7	1531	21.3	7.6	18	25.8	921	22.4	8.6
9	8.7	713	17.2	4.9	19	21.7	595	20.2	8.4
10	9.6	749	14.3	6.4	20	25.7	3353	16.9	6.7

1) จากตาราง ANOVA ที่กำหนดให้ (แบบจำลองที่มีตัวแปรอิสระ 2 ตัว) จงทำ overall F-test ของแต่ละแบบจำลอง และทำ partial F-test ของตัวแปรที่เข้าไปที่หลังในแต่ละแบบจำลอง

1) Y บน X_1 และ X_2		
SV	df	SS
X_2	1	1308.34
$X_1 X_2$	1	9.46
Error	17	537.40

2) Y บน X_1 และ X_3		
SV	df	SS
X_3	1	1387.60
$X_1 X_3$	1	35.63
Error	17	431.97

3) Y บน X_2 และ X_3		
SV	df	SS
X_3	1	1387.60
$X_2 X_3$	1	100.26
Error	17	367.34

- 2) จากผลของ 1) ทำแผนแนะนำให้ใช้แบบจำลองใดระหว่าง 3 แบบจำลองข้างต้น
- 3) คำนวณ R^2 ของทั้ง 3 แบบจำลองใน 1) แล้วโยงกับผลสรุปในข้อ 2)
- 4) กำหนดตาราง ANOVA ของแบบจำลองที่มีตัวแปรอิสระ 3 ตัว ให้ทดสอบว่าการเพิ่ม X_1 เข้าไปในแบบจำลองที่มี X_2 และ X_3 อยู่แล้ว ช่วยทำนาย Y อย่างมีนัยสำคัญหรือไม่

SV	df	SS
Regression (X_1, X_2, X_3)	3	1518.14
Error	16	337.06

5) หาค่า R^2 สำหรับแบบจำลองใน 4) เทียบกับ R^2 ของแบบจำลองที่มี X_2 และ X_3 เท่านั้น ว่าการเพิ่มขึ้นเป็นอย่างไร

6) ตาราง ANOVA ของแบบจำลองที่มีตัวแปรอิสระ X_2, X_3 , และ $X_4 = X_2X_3$ จงทดสอบว่าการเพิ่ม X_4 เข้าไปในแบบจำลองที่มี X_2 และ X_3 อยู่แล้วช่วยการทำนาย Y อย่างมีนัยสำคัญ

SV	df	SS	MS
X_3	1	1387.60	1387.60
$X_2 X_3$	1	100.26	100.26
$X_4 X_2X_3$	1	0.09	0.09
Error	16	367.25	22.95

7) จากผลของการกระทำข้างต้น ตัวแปรอิสระ X_1, X_2, X_3 และ X_4 ตัวใดที่ท่าน เห็นว่าจะเป็นตัวทำนายที่ดีที่สุดของ Y และจงลำดับความสำคัญของตัวแปรทั้ง 4 ตัว ถ้าพิจารณา R^2 ตัวแปรใดที่สำคัญ

วิธีทำ

1) จากตาราง ANOVA ทำ Overall F-test สำหรับแต่ละแบบจำลอง

จากแบบจำลอง 1) Y บน X_1 และ X_2 : $y = \beta_0 + \beta_1X_1 + \beta_2X_2 + e$

$H_0: \beta_1 = \beta_2 = 0$, $f_c = [(1308.34+9.46 +)/2]/(537.40/17)$

$= 20.844^{**}$, $f_{(2,17), .01} = 6.11$

จากแบบจำลอง 3) Y บน X_1 และ X_3 : $y = \beta_0 + \beta_1 X_1 + \beta_3 X_3 + e$

$$H_0: \beta_1 = \beta_3 = 0, f_c = [(1387.6+35.63)/2]/(431.97/17) \\ = 28.005^{**}$$

$$f_{(2,17),.01} = 6.11$$

จากแบบจำลอง 3) Y บน X_2 และ X_3 : $y = \beta_0 + \beta_2 X_2 + \beta_3 X_3 + e$

$$H_0: \beta_2 = \beta_3 = 0, f_c = [(1387.6+100.26)/2]/(367.34/17) \\ = 34.428^{**}$$

$$f_{(2,17),.01} = 6.11$$

ทำ Partial F- test ของตัวแปรที่เข้าไปที่หลังใน 3 แบบจำลอง

$$\text{จากแบบจำลอง 1) Partial } f_c = SS(X_1 | X_2) / MS(X_1, X_2) \\ = 9.46 / (537.40/17) = 0.299(\text{n.s.})$$

$$\text{จากแบบจำลอง 2) Partial } f_c = SS(X_1 | X_3) / MS(X_1, X_3) \\ = 35.63 / (431.97/17) = 1.402(\text{n.s.})$$

$$\text{จากแบบจำลอง 3) Partial } f_c = SS(X_2 | X_3) / MS(X_2, X_3) \\ = 100.26 / (367.34/17) = 4.64^*$$

$$f_{(1,17),.05} = 4.45, f_{(1,17),.01} = 8.40$$

2) จากผลของ 1) Overall F-test ของแบบจำลอง 3 Y บน X_2 และ X_3 สูงสุด และแม้ว่า Partial F-test ($X_2 | X_3$) จะไม่มีนัยสำคัญที่ $\alpha = .01$ ก็ตาม เราแนะนำให้ใช้แบบจำลอง 3 [ที่ขอบระหว่างแบบจำลอง 1-3 ใน 1)]

3) หา R^2 ของทั้ง 3 แบบจำลองใน 1)

$$R^2(X_1, X_2) = (1308.34 + 9.46)/1855.20 = 0.710$$

$$R^2(X_1, X_3) = (1387.6 + 35.63)/1855.20 = 0.767$$

$$R^2(X_2, X_3) = (1387.6 + 100.26)/1855.20 = 0.802 \text{ แสดงว่า}$$

แบบจำลอง 3 จะใช้ในการทำนายดีกว่าแบบจำลอง 1 และ แบบจำลอง 2

$$\begin{aligned} 4) \text{ Partial } f_c &= SS(X_1 | X_2, X_3) / MSE(X_1, X_2, X_3) \\ &= SS(X_1, X_2, X_3) - SS(X_2, X_3) \\ &= [1518.14 - (1387.6 + 100.26)] / (337.06/16) \\ &= 1.44(\text{n.s.}) \end{aligned}$$

$$f_{(1,16), .05} = 4.49$$

เราไม่ปฏิเสธ $H_0: \beta_1 = 0$

ในแบบจำลอง $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

ดังนั้นการเพิ่ม X_1 ไม่ช่วย X_2 และ X_3 ทำนาย Y

$$5) R^2(X_1, X_2, X_3) = 1518.14/1855.20 = 0.818 \quad \dots(1)$$

$$R^2(X_2, X_3) = 0.802 \quad \dots(2)$$

(1) และ (2) ต่างกัน 0.016 ซึ่งไม่มาก ดังนั้นการเพิ่ม X_1 จึงไม่มีความสำคัญ

$$\begin{aligned} 6) \text{ Partial } f_c &= SS(X_4 | X_2, X_3) / MSE(X_2, X_3, X_4) \\ &= 0.09/21.07 = 0.004(\text{n.s.}), (1,16) \text{ df.} \end{aligned}$$

ดังนั้นการเพิ่ม X_4 ไม่ช่วย X_2 และ X_3 ทำนาย Y

7) จากการกระทำข้างต้น เราพบว่า X_2 และ X_3 เป็นตัวแปรที่สำคัญ 2 ตัว X_3 นั้นสำคัญกว่า X_2

[พิจารณาจาก 1) $SS(X_3) = 1387.6 > SS(X_2) = 1308.34$]

เมื่อพิจารณาจากค่า R^2 เราพบว่าจากแบบจำลอง 3 (X_2, X_3)

เมื่อเพิ่ม X_1 หรือ X_4 เข้าไป X_1 หรือ X_4 ไม่ช่วยการทำนาย Y อย่างมีนัยสำคัญ

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 6.3

```
DATA LIST FREE/ Y X1 X2 X3.
BEGIN DATA.
11.2 587 16.5 6.2 13.4 643 20.5 6.4 40.7 635 26.3 9.3
5.3 692 16.5 5.3 24.8 1248 19.2 7.3 12.7 643 16.5 5.9
20.9 1964 20.2 6.4 35.7 1531 21.3 7.6 8.7 713 17.2 4.9
9.6 749 14.3 6.4 14.5 7895 18.1 6.0 26.9 762 23.1 7.4
15.7 2793 19.1 5.8 36.2 741 24.7 8.6 18.1 625 18.6 6.5
28.9 854 24.9 8.3 14.9 716 17.9 6.7 25.8 921 22.4 8.6
21.7 595 20.2 8.4 25.7 3353 16.9 6.7
END DATA.
REGRESSION DESCRIPTIVES=DEFAULTS SIG COV XPROD
/VARIABLES=Y X1 X2 X3
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X1 X2
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X1 X3
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X2 X3
/DEPENDENT=Y
/METHOD=ENTER.
```

Output สำหรับแบบฝึกหัดข้อ 6.3

Page 2

SPSS/PC+

2/6/91

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

	Mean	Std Devi	Label
Y	20.570	9.881	
X1	1433.000	1703.726	
X2	19.720	3.242	
X3	6.935	1.207	

N of Cases = 20

* * * * MULTIPLE REGRESSION * * * *

Correlation, Covariance, 1-tailed Sig, Cross-Product:

	Y	X1	X2	X3
Y	1.000	-.067	.840	.865
	97.642	-1129.616	26.905	10.316
	.999	.389	.000	.000
	1855.202	-21462.700	511.192	196.001
X1	-.067	1.000	-.164	-.233
	-1129.616	2902683.053	-904.705	-480.021
	.389	.999	.245	.161
	-21462.700	55150978.000	-17189.400	-9120.400
X2	.840	-.164	1.000	.815
	26.905	-904.705	10.512	3.191
	.000	.245	.999	.000
	511.192	-17189.400	199.732	60.636

* * * * MULTIPLE REGRESSION * * * *

	Y	X1	X2	X3
X3	.865	-.233	.815	1.000
	10.316	-480.021	3.191	1.457
	.000	.161	.000	.999
	196.001	-9120.400	60.636	27.686

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3

2.. X1

3.. X2

Multiple R .90461

R Square .81832

Adjusted R Square .78425

Standard Error 4.58978

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	3	1518.14494	506.04831
Residual	16	337.05706	21.06607

F = 24.02197 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	4.71982	1.53048	.57658	3.084	.0071
X1	7.629369E-04	6.36301E-04	.13154	1.199	.2480
X2	1.19217	.56165	.39117	2.123	.0497
(Constant)	-36.76493	7.01093		-5.244	.0001

End Block Number 1 All requested variables entered.

***** MULTIPLE REGRESSION *****

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

Page 10

SPSS/PC+

2/6/91

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X2
2.. X1

Multiple R .84281
R Square .71033
Adjusted R Square .67625
Standard Error 5.62245

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	1317.79841	658.89921
Residual	17	537.40359	31.61198

F = 20.84334 Signif F = .0000

Page 11

SPSS/PC+

2/6/91

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	2.59552	.40328	.85163	6.436	.0000
X1	4.198061E-04	7.67456E-04	.07238	.547	.5915
(Constant)	-31.21522	8.30060		-3.761	.0016

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3
2.. X1

Multiple R .87588
R Square .76716
Adjusted R Square .73976
Standard Error 5.04083

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	1423.23203	711.61601
Residual	17	431.96997	25.41000

F = 28.00535 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	7.35187	.98524	.89811	7.462	.0000
X1	8.266270E-04	6.98055E-04	.14252	1.184	.2526
(Constant)	-31.59977	7.22124		-4.376	.0004

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3
2.. X2

Multiple R .89554
R Square .80199
Adjusted R Square .77870
Standard Error 4.64848

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	1487.85941	743.92970
Residual	17	367.34259	21.60839

F = 34.42782 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	4.39894	1.52617	.53738	2.882	.0103
X2	1.22393	.56820	.40159	2.154	.0459
(Constant)	-34.07253	6.72655		-5.065	.0001

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X1

Multiple R .06710
 R Square .00450
 Adjusted R Square -.05080
 Standard Error 10.12930

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	8.35248	8.35248
Residual	18	1846.84952	102.60275

F = .08141 Signif F = .7787

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X1	-3.89163E-04	1.36396E-03	-.06710	-.285	.7787
(Constant)	21.12767	2.99173		7.062	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X2

Multiple R .83978
 R Square .70523
 Adjusted R Square .68885
 Standard Error 5.51192

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	1308.33948	1308.33948
Residual	18	546.86252	30.38125

F = 43.06404 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	2.55939	.39001	.83978	6.562	.0000
(Constant)	-29.90116	7.78918		-3.839	.0012

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3

Multiple R .86484
 R Square .74795
 Adjusted R Square .73395
 Standard Error 5.09685

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	1387.59972	1387.59972
Residual	18	467.60228	25.97790

F = 53.41461 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	7.07955	.96867	.86484	7.309	.0000
(Constant)	-28.52671	6.81372		-4.187	.0006

End Block Number 1 All requested variables entered.

6.4 จากตัวอย่าง blood serum ของคน 42 คน ได้ศึกษาถึงสิ่งที่พบใน high density lipoprotein (HDL) ตัวแปรที่คิดว่าจะทำนายค่า HDL ที่วัดได้ (Y) คือ

X_1 = โคลเลสเตอรอลรวม

X_2 = total triglyceride concentrations

และ X_3 = การปรากฏหรือไม่ปรากฏของส่วน "sticky" ใน serum ซึ่ง
เรียกว่า "sinking pre-beta หรือ SPB"

1) จงทดสอบว่า X_1 , X_2 หรือ X_3 ตัวใดตัวหนึ่งช่วยทำนาย Y อย่างมีนัยสำคัญหรือไม่

2) จงทดสอบว่า (X_1 และ X_2 และ X_3) ช่วยทำนาย Y อย่างมีนัยสำคัญหรือไม่

3) จงทดสอบว่า $\beta_{13} = \beta_{23} = 0$ หรือไม่ ในแบบจำลองที่มี X_1, X_2, X_3, X_1X_3 และ X_2X_3 (β_{13}, β_{23} เป็นสัมประสิทธิ์ของ X_1X_3 และ X_2X_3 ตามลำดับ) เขียน H_0 ในเทอมของสัมประสิทธิ์สหสัมพันธ์ชนิด multiple และถ้าเราไม่ปฏิเสธ H_0 ท่านจะสรุปเกี่ยวกับความสัมพันธ์ของ Y กับ X_1 และ X_2 เมื่อ $X_3 = 1$ เทียบกับเมื่อ $X_3 = 0$ ว่าอย่างไร

4) จงทดสอบ ($\alpha = .05$) ว่า X_3 มีความสัมพันธ์กับ Y หลังจากควบคุมอิทธิพลของ X_1 และ X_2 แล้ว จงเขียน H_0 ในเทอมของสัมประสิทธิ์สหสัมพันธ์ชนิด partial

5) จงสรุปผลเกี่ยวกับความสัมพันธ์ของ Y กับตัวแปรอิสระทั้ง 3 ตัว (X_1, X_2, X_3)

กำหนดข้อมูลและตาราง ANOVA ให้ดังต่อไปนี้

Y	X ₁	X ₂	X ₃	Y	X ₁	X ₂	X ₃
47	287	111	0	57	192	115	1
38	236	135	0	42	349	408	1
47	255	98	0	54	263	103	1
39	135	63	0	60	223	102	1
44	121	46	0	33	316	274	0
64	171	103	0	55	288	130	0
58	260	227	0	36	256	149	0
49	237	157	0	36	318	180	0
55	261	266	0	42	270	134	0
52	397	167	0	41	262	154	0
49	295	164	0	42	264	86	0
47	261	119	1	39	325	148	0
40	258	145	1	27	388	191	0
42	280	247	1	31	260	123	0
63	339	168	1	39	284	135	0
40	161	68	1	56	326	236	1
59	324	92	1	40	248	92	1
56	171	56	1	58	285	153	1
76	265	240	1	43	361	126	1
67	280	306	1	40	248	226	1
57	248	93	1	46	280	176	1

ANOVA

SV	df	SS
X_1	1	46.2356
Error	40	4567.3835

SV	df	SS
X_2	1	21.3397
Error	40	4592.2793

SV	df	SS
X_3	1	735.2054
Error	40	3878.4136

SV	df	SS
X_1, X_2	2	135.3820
Error	39	4478.2369

SV	df	SS
X_1, X_3	2	783.1691
Error	39	3830.4500

SV	df	SS
X_2, X_3	2	737.8069
Error	39	3875.8122

SV	df	SS	SV	df	SS
X_1, X_2, X_3	3	819.7473	X_1, X_3	2	783.1691
$X_1 X_3, X_2 X_3 \mid X_1, X_2, X_3$	2	74.7443	$X_1 X_3 \mid X_1, X_3$	1	62.4247
Error	36	3719.0517	Error	38	3768.0252

SV	df	SS
Regression (X_1, X_2)	2	737.8069
$X_1 X_2 \mid X_1 X_2$	1	1.5539
Error	38	3874.2583

วิธีทำ

1) แบบจำลอง Y บน X_1 : $y = \beta_0 + \beta_1 X_1 + e$

$$H_0: \beta_1 = 0$$

$$f_e(X_1) = 46.2356 / (4567.3835/40) = 0.405(\text{n.s.}), (1,40) \text{ df.}$$

แบบจำลอง Y บน X_2 : $y = \beta_0 + \beta_2 X_2 + e$

$$H_0: \beta_2 = 0$$

$$f_e(X_2) = 21.3397 / (4592.2793/40) = 0.186(\text{n.s.}), (1,40) \text{ df.}$$

แบบจำลอง Y บน X_3 : $y = \beta_0 + \beta_3 X_3 + e$

$$H_0: \beta_3 = 0$$

$$f_e(X_3) = 735.2054 / (3878.4136/40) = 7.583^{**}, (1,40) \text{ df.}$$

$$f_{(1,40), .01} = 7.31$$

2) Overall F-test

แบบจำลอง Y บน X_1, X_2 และ X_3 : $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

$$f_e = [SS(X_1, X_2, X_3) / 3] / MSE(X_1, X_2, X_3)$$

$$= (819.7473/3) / (3793.8718/38) = 2.737(\text{n.s.})$$

$$SST = 4613.6191, SSE(X_1, X_2, X_3) = 3793.8718,$$

$$f_{(3,38), .05} = 2.856$$

3) แบบจำลองมีตัวแปร $X_1, X_2, X_3, X_1X_3, X_2X_3$ (X_3 คือ dummy variable)

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + e$$

$$H_0: \beta_{13} = \beta_{23} = 0 \quad \text{หรือ} \quad H_0: f_{Y.(X_1 X_3 \cdot X_2 X_3)} \Big|_{X_1 \cdot X_2 \cdot X_3} = 0$$

$$\begin{aligned} \text{Multiple-partial } f_c &= (74.7443/2)/(3719.0517/36) \\ &= 0.362(\text{n.s.}), (2, 36) \text{ df.} \end{aligned}$$

นั่นคือเราไม่ปฏิเสธ H_0 ที่ $\alpha = .05$

สรุปได้ว่า สมการ Y บน X_1 และ X_2 เมื่อ $X_3 = 1$ ขนาน (parallel)

กับสมการ Y บน X_1 และ X_2 เมื่อ $X_3 = 0$ นั่นคือ ส.ป.ส. การทดลอง
ของตัวแปรเดียวกันเท่ากันในแบบจำลองทั้งสอง

4) แบบจำลอง Y บน X_1, X_2 และ X_3 : $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

$$H_0: f_{Y X_3} \Big|_{X_1 \cdot X_2} = 0$$

$$\begin{aligned} \text{Partial } f_c &= SS(X_3 \Big| X_1, X_2) / \text{MSE}(X_1, X_2, X_3) \\ &= [SS(X_1, X_2, X_3) - SS(X_1, X_2)] / \text{MSE}(X_1, X_2, X_3) \\ &= (819.7473 - 135.3820) / (3793.8718/38) \\ &= 6.855^* \end{aligned}$$

$$f_{(1, 38), .01} = 7.36, \quad f_{(1, 38), .05} = 4.10$$

สรุปได้ว่า สมการ Y บน X_1 และ X_2 เมื่อ $X_3 = 1$ ไม่ coincident กับ
สมการ Y บน X_1 และ X_2 เมื่อ $X_3 = 0$ ($\alpha = .05$)

5) X_3 (SBP) มีความสัมพันธ์กับ Y (HDL measurement) แต่อีก 2 ตัวไม่มี
ดังนั้นไม่ควรเอา X_1 และ X_2 ไว้ในแบบจำลอง

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 6.4

```
DATA LIST FREE/ Y X1 X2 X3.  
BEGIN DATA.  
47 287 111 0  
38 236 135 0  
47 255 98 0  
39 135 63 0  
44 121 46 0  
64 171 103 0  
58 260 227 0  
49 237 157 0  
55 261 266 0  
52 397 167 0  
49 295 164 0  
47 261 119 1  
40 258 145 1  
42 280 247 1  
63 339 168 1  
40 161 68 1  
59 324 92 1  
56 171 56 1  
76 265 240 1  
67 280 306 1  
57 248 93 1  
57 192 115 1  
42 349 408 1  
54 263 103 1  
60 223 102 1  
33 316 274 0  
55 288 130 0  
36 256 149 0  
36 318 180 0  
42 270 134 0  
41 262 154 0  
42 264 86 0  
39 325 148 0  
27 388 191 0  
31 260 123 0  
39 284 135 0  
56 326 236 1  
40 248 92 1  
58 285 153 1  
43 361 126 1  
40 248 226 1  
46 280 176 1  
END DATA.
```

```

REGRESSION DESCRIPTIVES=DEFAULTS SIG COV XPROD
/VARIABLES=Y X1 X2 X3
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X1 X2
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X1 X3
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X2 X3
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X1
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X2
/DEPENDENT=Y
/METHOD=ENTER.
REGRESSION VARIABLES=Y X3
/DEPENDENT=Y
/METHOD=ENTER.

```

Output สำหรับแบบฝึกหัดข้อ 6.4

42 cases are written to the uncompressed active file.

This procedure was completed at 14:31:33

Page 2

SPSS/PC+

2/6/91

***** MULTIPLE REGRESSION *****

Listwise Deletion of Missing Data

	Mean	Std Dev	Label
Y	47.762	10.608	
X1	267.810	60.403	
X2	155.048	73.750	
X3	.476	.505	

N of Cases = 42

* * * * MULTIPLE REGRESSION * * * *

Correlation, Covariance, 1-tailed Sig, Cross-Product:

	Y	X1	X2	X3
Y	1.000	-.100	.068	.399
	112.527	-64.144	53.207	2.141
	.999	.264	.334	.004
	4613.619	-2629.905	2181.476	87.762
X1	-.100	1.000	.513	.005
	-64.144	3648.548	2284.204	.142
	.264	.999	.000	.488
	-2629.905	149590.476	93652.381	5.810
X2	.068	.513	1.000	.111
	53.207	2284.204	5439.120	4.148
	.334	.000	.999	.242
	2181.476	93652.381	223003.905	170.048

* * * * MULTIPLE REGRESSION * * * *

	Y	X1	X2	X3
X3	.399	.005	.111	1.000
	2.141	.142	4.148	.256
	.004	.488	.242	.999
	87.762	5.810	170.048	10.476

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3
2.. X1
3.. X2

Multiple R .42152
R Square .17768
Adjusted R Square .11276
Standard Error 9.99193

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	3	819.74732	273.24911
Residual	38	3793.87172	99.83873

F = 2.73690 Signif F = .0568

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	8.14831	3.11224	.38828	2.618	.0126
X1	-.02731	.03015	-.15552	-.906	.3707
X2	.01504	.02485	.10456	.605	.5486
(Constant)	48.86453	7.24985		6.740	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X2
2.. X1

Multiple R .17130
R Square .02934
Adjusted R Square -.02043
Standard Error 10.71572

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	135.38213	67.69106
Residual	39	4478.23692	114.82659

F = .58951 Signif F = .5595

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	.02329	.02643	.16191	.881	.3837
X1	-.03216	.03227	-.18313	-.997	.3251
(Constant)	52.76401	7.60916		6.934	.0000

End Block Number 1 All requested variables entered.

**** MULTIPLE REGRESSION ****

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3
2.. X1

Multiple R .41201
R Square .16975
Adjusted R Square .12717
Standard Error 9.91043

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	783.16907	391.58453
Residual	39	3830.44998	98.21667

F = 3.98695 Signif F = .0266

**** MULTIPLE REGRESSION ****

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	8.38720	3.06193	.39967	2.739	.0092
X1	-.01791	.02562	-.10196	-.699	.4888
(Constant)	48.56351	7.17378		6.770	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

Page 18

SPSS/PC+

2/6/91

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3

2.. X2

Multiple R .39990
R Square .15992
Adjusted R Square .11684
Standard Error 9.96894

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	737.80689	368.90345
Residual	39	3875.81216	99.37980

F = 3.71206 Signif F = .0334

Page 19

SPSS/PC+

2/6/91

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	8.32149	3.09922	.39654	2.685	.0106
X2	3.436833E-03	.02124	.02389	.162	.8723
(Constant)	43.26642	3.78287		11.437	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X1

Multiple R .10011
 R Square .01002
 Adjusted R Square -.01473
 Standard Error 10.68572

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	46.23556	46.23556
Residual	40	4567.38349	114.18459

F = .40492 Signif F = .5282

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X1	-.01758	.02763	-.10011	-.636	.5282
(Constant)	52.47018	7.58057		6.922	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X2

Multiple R .06801
 R Square .00463
 Adjusted R Square -.02026
 Standard Error 10.71480

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	21.33971	21.33971
Residual	40	4592.27934	114.80698

F = .18587 Signif F = .6687

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X2	9.782233E-03	.02269	.06801	.431	.6687
(Constant)	46.24519	3.88712		11.897	.0000

End Block Number 1 All requested variables entered.

*** MULTIPLE REGRESSION ***

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. Y

Beginning Block Number 1. Method: Enter

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

Variable(s) Entered on Step Number

1.. X3

Multiple R .39919
 R Square .15936
 Adjusted R Square .13834
 Standard Error 9.84684

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	735.20541	735.20541
Residual	40	3878.41364	96.96034

F = 7.58254 Signif F = .0088

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. Y

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
X3	8.37727	3.04225	.39919	2.754	.0088
(Constant)	43.77273	2.09935		20.851	.0000

End Block Number 1 All requested variables entered.

6.5 จากข้อ 6.1 การถดถอยเชิงเส้นตรงของ SBP บน QUET ถูกเปรียบเทียบกันระหว่างพวกที่สูบบุหรี่ (SMK = 1) และพวกที่ไม่สูบบุหรี่ (SMK = 0)

1) จงเขียนแบบจำลองการถดถอยแบบพหุคูณ 1 แบบจำลอง สำหรับพวกที่สูบบุหรี่และพวกที่ไม่สูบบุหรี่ และเขียนแบบจำลองเชิงเส้นตรงสำหรับ 2 กลุ่ม กำหนด intercept และ slope ของแต่ละเส้นตรงในเทอมของสัมประสิทธิ์การถดถอยของแบบจำลองการถดถอยแบบพหุคูณ

2) ใช้ output จากคอมพิวเตอร์ที่กำหนดให้เพื่อหาและพลอตเส้นตรงของการถดถอยที่คำนวณได้ (fitted straight lines) ทั้งสอง

3) จงทดสอบ 3.1) H_0 : เส้นตรงทั้งสองขนานกัน (parallel)

3.2) H_0 : เส้นตรงทั้งสองเป็นเส้นตรงเดียวกัน (coincident)

ให้เขียนสมมติฐานในรูปของสัมประสิทธิ์การถดถอย

ตัวแปรที่เข้าไปในขั้นที่ 1: QUET				ANOVA (ขั้นที่ 1)		
ตัวแปร	$\hat{\beta}$	$s_{\hat{\beta}}$	Partial F	SV	df	SS
QUET	21.49167	3.54515	36.751	Regression	1	3537.94538
ค่าคงที่	70.57643			Error	30	2888.02337

ตัวแปรที่เข้าไปในขั้นที่ 2: SMK				ANOVA (ขั้นที่ 2)		
ตัวแปร	$\hat{\beta}$	$s_{\hat{\beta}}$	Partial F	SV	df	SS
QUET	22.11560	3.22996	46.882	Regression	2	4120.36603
SMK	8.57101	3.16670	7.326	Error	29	2305.60272
ค่าคงที่	63.87606					

ตัวแปรที่เข้าไปในขั้นที่ 3: QUET.SMK				ANOVA (ขั้นที่ 3)		
ตัวแปร	$\hat{\beta}$	$s_{\hat{\beta}}$	Partial F	SV	df	SS
QUET	26.30282	5.70349	21.268	Regression	3	4184.10718
SMK	29.94357	24.16355	1.536	Error	28	2241.86157
QUET.SMK	-6.18479	6.93171	0.796			
ค่าคงที่	49.31178					

วิธีทำ

$$1) \text{SBP} = \beta_0 + \beta_1(\text{QUET}) + \beta_2(\text{SMK}) + \beta_3(\text{QUET.SMK}) + e$$

โดยที่ SMK = 1 ถ้าสูบบุหรี่ (smoker)

0 ถ้าสูบบุหรี่ (nonsmoker)

แทน SMK = 1 และ SMK = 0 ตามลำดับ เราจะได้

สำหรับ smoker : $\text{SBP} = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)\text{QUET} + e$

สำหรับ nonsmoker : $\text{SBP} = \beta_0 + \beta_1(\text{QUET}) + e$

$$2) \widehat{\text{SBP}} = 49.31178 + 26.30282(\text{QUET}) + 29.94357(\text{SMK}) - 6.18479(\text{QUET.SMK})$$

$$\text{smoker (SMK = 1): } \widehat{\text{SBP}} = 79.253 + 20.118(\text{QUET})$$

$$\text{nonsmoker (SMK = 0): } \widehat{\text{SBP}} = 49.312 + 26.303(\text{QUET})$$

แล้วนำสมการทั้งสองไปพลอตบนแกน xy

3) 3.1) H_0 : เส้นตรงทั้งสองขนานกัน (parallel) หรือ $H_0: \beta_2 = 0$

$$\begin{aligned} \text{Partial } f_c &= \text{SS}(\text{QUET.SMK} \mid \text{QUET, SMK}) / \text{MSE}(\text{QUET, SMK, QUET.SMK}) \\ &= [\text{SS}(\text{QUET, SMK, QUET.SMK}) - \text{SS}(\text{QUET, SMK})] / \\ &\quad \text{MSE}(\text{QUET, SMK, QUET.SMK}) \\ &= [4184.10718 - 4120.36603] / (2241.86157/28) \\ &= 0.769(\text{n.s.}), (1, 28) \text{ df.} \end{aligned}$$

สรุป ไม่ปฏิเสธ H_0 ที่ $\alpha = .05$ นั่นคือเส้นตรงทั้งสองขนานกัน

3.2) H_0 : เส้นตรงทั้งสองเป็นเส้นตรงเดียวกัน (coincident)

หรือ $H_0: \beta_2 = \beta_3 = 0$

$$\begin{aligned} \text{Multiple-partial } f_c &= \text{SS}(\text{SMK, QUET.SMK} \mid \text{QUET}) / \\ &\quad \text{MSE}(\text{QUET, SMK, QUET.SMK}) \\ &= [\text{SS}(\text{QUET, SMK, QUET.SMK}) - \text{SS}(\text{QUET})] / \\ &\quad \text{MSE}(\text{QUET, SMK, QUET.SMK}) \\ &= [4184.10718 - 3537.94538] / 21 / \\ &\quad (2241.86157/28) \\ &= 4.03^* \end{aligned}$$

$$f_{(2, 28), .05} = 3.34, f_{(2, 28), .01} = 5.45$$

สรุป ไม่ยอมรับ H_0 ที่ $\alpha = .05$ แต่ยอมรับ H_0 ที่ $\alpha = .01$

นั่นคือถ้าทดสอบที่ $\alpha = .01$ เส้นตรงทั้งสองเป็นเส้นตรงเดียวกัน

6.6 จากข้อมูลในข้อ 6.1 และตัวแปรทั้ง 4 ตัว จงหาแบบจำลองการถดถอยที่ดีที่สุด (best regression model) ให้ใช้ $\alpha = .05$ โดย

- 1) วิธี forward selection
 - 2) วิธี Backward elimination
- และ 3) วิธี All-possible regressions

SPSS/PC commands สำหรับแบบฝึกหัดข้อ 6.5

คำสั่งส่วนต้นจากคำสั่งในข้อ 6.1 (ถึง END DATA.)

```
COMPUTE QUETSMK=QUET*SMK.  
REGRESSION VARIABLES=SBP,QUET,SMK,QUETSMK  
    /DEPENDENT=SBP  
    /METHOD=ENTER QUET  
    /METHOD=ENTER SMK  
    /METHOD=ENTER QUETSMK.
```

Output สำหรับแบบฝึกหัดข้อ 6.5

Page 2 SPSS/PC+ 2/13/91

***** MULTIPLE REGRESSION *****

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 1. Method: Enter QUET

Page 3 SPSS/PC+ 2/13/91

***** MULTIPLE REGRESSION *****

Equation Number 1 Dependent Variable.. SBP

Variable(s) Entered on Step Number

1.. QUET

Multiple R .74200
R Square .55057
Adjusted R Square .53559
Standard Error 9.81160

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	1	3537.94574	3537.94574
Residual	30	2688.02301	96.26743

F = 36.75122 Signif F = .0000

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
QUET	21.49167	3.54515	.74200	6.062	.0000
(Constant)	70.57640	12.32187		5.728	.0000

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
SMK	.30183	.44907	.99491	2.707	.0113
QUETSMK	.28739	.42564	.98581	2.533	.0170

End Block Number 1 All requested variables entered.

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 2. Method: Enter SMK

Variable(s) Entered on Step Number
2.. SMK

Multiple R .80075
 R Square .64121
 Adjusted R Square .61646
 Standard Error 8.91647

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	2	4120.36649	2060.18325
Residual	29	2305.60226	79.50353

F = 25.91311 Signif F = .0000

Page 6

SPSS/PC+

2/13/91

* * * * MULTIPLE REGRESSION * * * *

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
QUET	22.11560	3.22996	.76355	6.847	.0000
SMK	8.57101	3.16670	.30183	2.707	.0113
(Constant)	63.87603	11.46811		5.570	.0000

----- Variables not in the Equation -----

Variable	Beta In	Partial	Min Toler	T	Sig T
QUETSMK	-.76272	-.16627	.01705	-.892	.3799

End Block Number 2 All requested variables entered.

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable., SBP

Beginning Block Number 3. Method: Enter QUETSMK

Variable(s) Entered on Step Number
3. QUETSMK

Multiple R .80692
R Square .65112
Adjusted R Square .61375
Standard Error 8.94799

Analysis of Variance

	DF	Sum of Squares	Mean Square
Regression	3	4184.10759	1394.70253
Residual	28	2241.86116	80.06647

F = 17.41931 Signif F = .0000

Page 8

SPSS/PC+

2/13/91

*** MULTIPLE REGRESSION ***

Equation Number 1 Dependent Variable.. SBP

----- Variables in the Equation -----

Variable	B	SE B	Beta	T	Sig T
QUET	26.30283	5.70349	.90811	4.612	.0001
SMK	29.94357	24.16355	1.05446	1.239	.2256
QUETSMK	-6.18478	6.93171	-.76272	-.892	.3799
(Constant)	49.31176	19.97235		2.469	.0199

End Block Number 3 All requested variables entered.

Page 9

SPSS/PC+

2/13/91

This procedure was completed at 8:53:37

1) วิธี forward selection

ตัวแปรตาม (Dependent variable) : Y(SBP)

ตัวแปรอิสระ (Independent variables): AGE, QUET, SMK

ขั้นที่ 1 หา ส.ป.ส. สหสัมพันธ์ของ Y กับตัวแปรอิสระแต่ละตัว

$$r_{SBP, AGE} = 0.775$$

$$r_{SBP, QUET} = 0.742$$

$$r_{SBP, SMK} = 0.247$$

$r_{SBP, AGE}$ มีค่าสูงสุด ดังนั้นตัวแปรตัวแรกที่จะเข้าไปในแบบจำลองคือ AGE

สมการการทำนายคือ $\widehat{SBP} = 59.092 + 1.605(AGE)$

ANOVA

SV	df	SS	MS	f_c
AGE	1	3861.6304	3861.6304	45.1769**
Error	30	2564.3384	85.4779	
Total	31	6425.9688		

$$f_{(1,30), .05} = 4.17, f_{(1,30), .01} = 7.56$$

ขั้นที่ 2 คำนวณ partial f_c ของตัวแปรที่เหลืออยู่ (ที่ยังไม่ได้เข้าไปในแบบจำลอง) จากแบบจำลองซึ่งมีตัวแปรตัวนั้น ๆ และตัวแปรที่เข้าไปแล้วในขั้นที่ 1

$$\text{partial } f_c (\text{QUET} | \text{AGE}) = 3.2576, (1,29) \text{ df.}$$

$$\text{partial } f_c (\text{SMK} | \text{AGE}) = 13.8304^{**}, (1,29) \text{ df.}$$

partial f_c (SMK | AGE) สูงสุด เราเพิ่ม SMK เข้าไปในแบบจำลอง

ขั้นที่ 3 partial $f_c(\text{QUET} | \text{AGE, SMK}) = 3.6481(\text{n.s.})$, (1,28) df.
 เราไม่เพิ่ม QUET เข้าไปในแบบจำลอง

ดังนั้นแบบจำลองที่เลือกได้จากวิธี Forward selection คือ
 แบบจำลองที่มี AGE และ SMK

$$\widehat{\text{SBP}} = 48.0496 + 1.7092(\text{AGE}) + 10.2944(\text{SMK})$$

2) วิธี Backward elimination

ขั้นที่ 1 จากการใช้แบบจำลองที่มีตัวแปรอิสระทุกตัวที่มีอยู่

$$\widehat{\text{SBP}} = 45.1032 + 1.2127(\text{AGE}) + 8.5925(\text{QUET}) + 9.9456(\text{SMK})$$

ANOVA

SV	df	SS	MS
Regression	3	4889.8258	1629.9419
Error	28	1536.1430	54.8623
Total	31	6425.9688	

ขั้นที่ 2 คำนวณ partial f_c ของตัวแปรทุกตัวในแบบจำลอง โดยถือว่าตัวแปรตัวนั้น
 เข้ามาในแบบจำลองเป็นตัวสุดท้าย

ตัวแปร	partial f_c (1,28) df.	
AGE	$f_c(\text{AGE} \text{SMK, QUET}) = 14.0253$	
QUET	$f_c(\text{QUET} \text{AGE, SMK}) = 3.6481$	<---- ต่ำที่สุด
SMK	$f_c(\text{SMK} \text{AGE, QUET}) = 14.0212$	

ขั้นที่ 3 $f_c(\text{QUET} | \text{AGE, SMK})$ มีค่าต่ำสุด และการทดสอบไม่มีนัยสำคัญที่ $\alpha = .05$
 ดังนั้นเราเอา QUET ออกจากแบบจำลอง

ขั้นที่ 4 คำนวณ partial f_c ของตัวแปรที่เหลืออยู่ในแบบจำลอง

ตัวแปร	partial f_c (1,29) df.
AGE	$f_c(\text{AGE} \text{SMK}) = 71.763$
SMK	$f_c(\text{SMK} \text{AGE}) = 13.830$ ← ต่ำที่สุด

ขั้นที่ 5 $f_c(\text{SMK} | \text{AGE})$ มีค่าต่ำที่สุด และการทดสอบมีนัยสำคัญที่ $\alpha = .05$
 ดังนั้นเราเก็บทั้ง AGE และ SMK ไว้ในแบบจำลอง

ดังนั้นแบบจำลองที่เลือกได้จากวิธี Backward elimination
 คือแบบจำลองที่มี AGE และ SMK

3) วิธี All-possible regressions

โดยการหาค่าสัมประสิทธิ์การถดถอยทุกสมการที่จะเป็นไปได้จากการใช้ตัวแปรอิสระที่เรา
 พิจารณาอยู่ แล้วคำนวณ multiple R^2 ของแต่ละสมการการถดถอย
 ได้ผลดังตารางต่อไปนี้

แบบจำลอง	ตัวแปรในแบบจำลอง	R ²	หมายเหตุ
1	AGE	0.601	สูงสุดในกลุ่ม
2	QUET	0.551	
3	SMK	0.061	
4	AGE, QUET	0.641	สูงสุดในกลุ่ม
5	AGE, SMK	0.730	
6	QUET, SMK	0.641	
7	AGE, SMK, QUET	0.761	สูงสุด

แม้ว่าแบบจำลองที่มีตัวแปรอิสระ 3 ตัว (แบบจำลอง 7) จะมีค่า R² สูงสุด (R² = 0.761) แต่อย่างไรก็ดีแบบจำลองที่มีตัวแปร AGE และ SMK (แบบจำลอง 5) มีค่า R² ซึ่งมีค่าสูงเกือบเท่า ๆ กัน (R² = 0.73) แสดงว่าแบบจำลอง 5 นั้นเหมาะสมดีแล้ว

สรุปผลการเลือกสมการการถดถอยที่ดีที่สุดนั้นเหมือนกับ 2 วิธีข้างต้น