

การประเมินผลก่อนการเขียน

ตัวอย่างข้อสอบ

ข้อสอบชุดนี้ 5 ข้อ 100 คะแนน

ข้อ 1. จงเติมคำหรือข้อความให้ถูกต้องและได้ความ

1.1 เมื่อแบบจำลอง $Y_{n \times 1} = X_{n \times p} \beta_{p \times 1} + e_{n \times 1}$ มี $\text{rank}(X) = p$ เราเรียกแบบจำลองนี้ว่า General linear hypothesis model of full rank (แบบจำลอง a)

แต่ถ้า $\text{rank}(X) < p$ เราจะเรียกแบบจำลอง (แบบจำลอง b) ว่า

1.2 จากข้อ 1.1 แบบจำลอง (a) หรือ (b) ที่คือ Experimental Design Model

1.3 Regression Model เป็นแบบจำลอง (a) หรือแบบจำลอง (b) ใน 1.1 _____

1.4 จากข้อ 1.1 สำหรับแบบจำลอง (a) จงเขียน normal equations

ในรูป matrix _____

ดังนั้น least square estimator ของ β คือ $\hat{\beta} =$ _____

1.5 จงเขียน

(a) Cubic polynomial model: _____

และ (b) Quartic polynomial model: _____

1.6 จาก Regression model $y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + e$

R^2 (Coefficient of multiple determination) = $SS(\beta_1, \beta_2, \beta_3 | \beta_0) / SST$

ให้อธิบาย _____

1.7 จาก 1.6 $R^2 = R$ คือสหสัมพันธ์อันดับหนึ่งซึ่งมีชื่อว่า _____

1.8 จาก 1.6 ถ้า partial $f_c(X_3 | X_1, X_2)$ ไม่มีนัยสำคัญ (n.s.) เราสรุปผลว่า _____

1.9 ใน (a x b) factorial design นั้น ถ้าระดับของปัจจัย A ถูกสุ่มมาจากระดับทั้งหมดของ A และระดับของปัจจัย B ถูกกำหนดขึ้นตามความสนใจ เราเรียกแบบจำลองของการวิเคราะห์ว่า _____

1.10 ใน (5 x 5) Greco latin square design (1 หน่วยทดลอง/cell) เราสามารถทดสอบอิทธิพลหลักได้เป็นจำนวน _____ อิทธิพล แต่เราใช้หน่วยทดลองสำหรับ design นี้เพียง _____ หน่วยทดลองเท่านั้น

ข้อ 2. 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

จาก Computer Printer จงเติมตาราง และ ตอบปัญหาต่อไปนี้

(ให้ใช้ทศนิยม 1 ตำแหน่ง)

Least Squares estimates

แบบจำลอง	ตัวแปรอิสระในแบบจำลอง	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$S_{\hat{\beta}_1}^2$	$S_{\hat{\beta}_2}^2$	$S_{\hat{\beta}_3}^2$
1	AGE (X_1)							
2	AGE (X_1), SMK (X_2)							
3	AGE (X_1), SMK (X_2), QUET (X_3)							

ตาราง ANOVA

แบบจำลอง 1		
sv	df	ss
Regression (X_1)		
Error		

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

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

1) จงใช้แบบจำลอง 3

1.1 ทำนาย SBP ของชายอายุ 50 ปี สูบบุหรี่ และ ขนาดของร่างกาย = 3.5

1.2 ทำนาย SBP ของชายอายุ 50 ปี ไม่สูบบุหรี่ และ ขนาดของร่างกาย = 3.5

1.3 สำหรับชายอายุ 50 ปี สูบบุหรี่ จงประมาณความแตกต่างของ SBP เมื่อขนาดของร่างกายเพิ่มจาก 3.0 เป็น 3.5

2) R^2 ของแบบจำลองทั้ง 3 คือ

3) จงทำ overall F-test ของแต่ละแบบจำลอง ระบุ H_0 ที่ทดสอบให้ชัดเจนด้วย

4) จงทำ partial F-test โดยให้ $\alpha = .05$ สำหรับการ

4.1 เพิ่ม SMK เข้าไปในแบบจำลองที่มี AGE อยู่แล้ว

4.2 เพิ่ม QUET เข้าไปในแบบจำลองที่มี AGE และ SMK อยู่แล้ว

ข้อนี้ นักศึกษาอาจทำ t-test ซึ่งสมมูลกับ partial F-test ก็ได้

```

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.

```


*** 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

*** 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

*** 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 12 SPSS/PC+ 2/13/91

*** 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 13 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
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

Page 15

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
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.

Page 16

SPSS/PC+

2/13/91

This procedure was completed at 8:33:35

Listwise Deletion of Missing Data

Equation Number 1 Dependent Variable.. SBP

Beginning Block Number 1. Method: Stepwise

 Page 18 SPSS/PC+ 2/13/91

***** 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 19 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
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.

ข้อ 3. จากโจทก์และ Computer printout ในข้อ 2. จงอธิบายวิธีเลือกแบบจำลองการถดถอยที่ดีที่สุด (Best Regression model) สำหรับการทำนาย Y (SBP) เลือกอธิบายอย่างละเอียด 1 วิธี จาก 3 วิธีที่กำหนดให้ คือ

- 1) วิธี Forward selection
- 2) วิธี Backward elimination
- 3) วิธี Stepwise

หมายเหตุ ให้ใช้ค่าจาก Computer printout ประกอบคำอธิบายทุกขั้นตอน

ข้อ 4.1 ตารางต่อไปนี้จะแสดงระยะเวลารอคอยเฉลี่ย (นาที) ที่คนไข้ 16 คนต้องรอนหมอ คนไข้ถูกจำแนกโดย ชนิดของการฝึกหัดของหมอ (Type of practice: TYPRAC) และชนิดของหมอ (Type of physician: PHYSTY)

ชนิดของหมอ	ชนิดของการฝึกหัด	
	แบบกลุ่ม (GROUP)	แบบเดี่ยว (SOLO)
รักษาโรคทั่วไป (GP)	15, 20, 25, 20	20, 25, 30, 25
เฉพาะทาง (SPEC)	30, 25, 30, 35	25, 20, 30, 30

กำหนดบางส่วนของตาราง ANOVA ให้ จงเติมตาราง ANOVA ให้สมบูรณ์ และ จงทดสอบที่ $\alpha = .05$ สำหรับอิทธิพลหลัก และอิทธิพลร่วม โดยใช้ Random effect model

ANOVA

SV	df	SS	MS	f
PHYSTY		126.56	(1) _____	$f_1 =$ _____
TYPRAC		1.5625	(2) _____	$f_2 =$ _____
PHYSTY x TYPRAC		76.563	(3) _____	$f_3 =$ _____
Error		218.7525	(4) _____	
Total		_____		

หมายเหตุ ที่คอลัมน์ MS ให้ใช้เลขตัวเต็ม ในที่นี้เพื่อให้หาค่า f_c ได้ง่าย (เฉพาะการสอบ)

ให้ระบุให้ชัดเจนว่าการหา f_c แต่ละตัวหามาได้อย่างไร (อะไรเป็นเศษ อะไรเป็นส่วน)

ข้อ 4.2 ที่ปรึกษาเกี่ยวกับ MIS (Management Information System) ได้ทำ small-scale study เกี่ยวกับรายงานสรุปรายวัน 5 แบบ (A = มีรายละเอียดมากที่สุด, B, C, D, E = มีรายละเอียดน้อยที่สุด) เขาใช้ผู้จัดการฝ่ายขาย (Sales executive) 5 คน แต่ละคนจะได้รับรายงานสรุปรายวันแบบหนึ่งโดยตลอดในเดือนหนึ่ง ๆ แล้วจะต้องให้คะแนนว่ารายงานนั้นช่วยการตัดสินใจมากน้อยเพียงใดตาม 25-point scale (0 = ไม่ช่วยเลย, ..., 25 = ช่วยมากที่สุด) ในช่วงเวลา 5 เดือน ผู้จัดการฝ่ายขายแต่ละคนจะได้รับรายงานแบบหนึ่งในเดือนหนึ่ง ๆ ผลปรากฏในตารางต่อไปนี้

ผู้จัดการ	เดือน				
	มี.ค.	เม.ย.	พ.ค.	มิ.ย.	ก.ค.
นาย ก	D 21	A 8	C 17	B 9	E 16
นาย ข	A 5	E 10	B 3	C 12	D 15
นาย ค	C 20	B 10	E 15	D 21	A 12
นาย ง	B 4	D 15	A 3	E 9	C 10
นาย จ	E 17	C 16	D 20	A 8	B 11

กำหนดบางส่วนของตาราง ANOVA ให้ จงเติมตาราง ANOVA ให้สมบูรณ์ และ จงทดสอบที่ $\alpha = .025$ ว่า แบบของรายงานมีอิทธิพลต่างกันหรือไม่

SV	df	SS	MS	f
ผู้จัดการ		233.04		$f_1 = 62.87$
เดือน		12.24		$f_2 = 3.30$
รายงาน				$f_3 = 129.13$
Error		11.12		
Total		735.04		

ข้อ 5.1 จากแบบจำลอง $y_{ij} = \mu_i + e_{ij}$, $i = 1, \dots, 3$; $j = 1, \dots, n_i$

ค่าสังเกต:

วิธีการ				
	1	2	3	
	y_{11}	y_{21}	y_{31}	
	y_{12}	y_{22}	y_{32}	
	y_{13}		y_{33}	
	y_{14}			
n_i	4	2	3	$n = 9$
T_i	T_1	T_2	T_3	G

ถ้าแบบจำลองในรูปเมตริกซ์ คือ $Y = X\beta + e$

จงเขียนสิ่งต่อไปนี้ (โดยแสดงสมาชิกให้ชัดเจน และบอกขนาดของเมตริกซ์ หรือ เวกเตอร์ด้วย) Y , X , β , e , $X'X$, $X'Y$, $Y'Y$ และ $e'e$

ข้อ 5.2 จากข้อมูลเกี่ยวกับปริมาณน้ำฝนในปีต่าง ๆ สมมติว่าเราต้องการหา polynomial ที่ fit กับข้อมูลดังต่อไปนี้คือ

ปีที่ i	1	2	3	4	5	6	7	8
y: ปริมาณน้ำฝน	30.2	32.2	35.1	34.2	39.1	41.3	36.1	30.1

ปีที่ i	9	10	11	12
y: ปริมาณน้ำฝน	30.5	26.1	24.8	28.2

$$\sum_{i=1}^{12} y_i = 387.9, \quad \sum_{i=1}^{12} y_i^2 = 12815.99$$

จากตาราง ANOVA ที่กำหนดให้

1. จงทดสอบสมมติฐานที่ $\alpha = .02$ ว่า linear term, quadratic term, cubic term และ quartic term มีนัยสำคัญหรือไม่ (ให้แสดงการทดสอบให้ชัดเจน ประกอบการสรุปผล)

2. จงอภิบาลสรุปว่า polynomial degree เท่าใด fit กับข้อมูลข้างต้นดีที่สุด

ANOVA

S.V.	df.	SS	MS	f_c	Pr
Total (uncorrected)	12	12815.99			
Reduction for mean ($n\bar{y}^2$)	1	12538.87			
Linear	1	71.548	71.548	3.48	>2%
Error	10	205.576	20.558		
Quadratic	1	103.963	103.963	9.21	<2%
Error	9	101.613	11.290		
Cubic	1	38.484	38.484	4.88	>2%
Error	8	63.130	7.891		
Quartic	1	34.432	34.432	8.40	>2%
Error	7	28.691	4.100		