6-4 Other Aspects of Regression

Q 6-56

6-19 Data from a patient satisfaction survey in a hospital are shown in the following table:

Observation	Age	Severity	Surg-Med	Anxiety	Satisfaction
1	39	42	0	3.5	83
2	70	41	1	7.0	59

The regressor variables are the patient's age, an illness severity index (larger values indicate greater severity), an indicator variable denoting whether the patient is a medical patient (0) or a surgical patient (1), and an anxiety index (larger values indicate greater anxiety)

6-56. Consider the patient satisfaction survey data of Exercise 6-19. Using only first-order terms, build regression models using the following techniques:

(a) All possible regressions. Find the C_p and S values.
(b) Forward selection.
(c) Backward elimination.
(d) Comment on the models obtained. Which model would you prefer?

Age	Sev	Surg	Anx	Sat
39	42	0	3.5	83
70	41	1	7	59
62	62	0	7.2	46

```
ods graphics off;
DATA q656;
Infile 'C:\Users\korea\Dropbox\수업 자료\학부\imen214-응용통계
\sas\ex656.txt';
INPUT X1 X2 X3 X4 Y;
if x3=0 then x3d=1; else x3d=0;
label Y='Satisfaction' X1='Age' X2='Severity'
  X3='Surg-Med' X4='Anxiety';

proc reg data=q656;
 model y=x1 x2 x3d x4/selection=rsquare adjrsq cp rmse best=2;
 title 'question 6-56 (a)';

proc reg data=q656;
 model y=x1 x2 x3d x4/selection=forward slentry=0.25;;
 title 'question 6-56 (b)';

proc reg data=q656;
 model y=x1 x2 x3d x4/selection=backward;
 title 'question 6-56 (c)';

proc reg data=q656 outest=est;
 m1: model y=x1 x2/rsquare adjrsq cp press;
m2: model y=x1 x2 x4/rsquare adjrsq cp press ;
m3: model y=x1 x2 x3d/rsquare adjrsq cp press ;
PROC PRINT data=est ;
 title 'question 6-56 (d)';

RUN;
```

6-4 Other Aspects of Regression

question 6-56 (a)

R-Square Selection Method

Number of Observations Read	25
Number of Observations Used	25

Number in Model	R-Square	Adjusted R-Square	C(p)	Root MSE	Variables in Model
1	0.8146	0.8066	15.3764	9.29483	X1
1	0.4689	0.4458	83.2353	15.73399	X2
2	0.8896	0.8795	2.6748	7.33605	X1 X2
2	0.8167	0.8001	16.9647	9.44980	X1 X4
3	0.8980	0.8835	3.0123	7.21496	X1 X2 X4
3	0.8896	0.8738	4.6696	7.50777	X1 X2 x3d
4	0.8981	0.8777	5.0000	7.39087	X1 X2 x3d X4

6-4 Other Aspects of Regression

question 6-56 (b)

No other variable met the 0.2500 significance level for entry into the model.

Summary of Forward Selection								
Step	Variable Entered	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	X1	Age	1	0.8146	0.8146	15.3764	101.08	<.0001
2	X2	Severity	2	0.0749	0.8896	2.6748	14.92	0.0008
3	X4	Anxiety	3	0.0085	0.8980	3.0123	1.74	0.2008

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	143.85960	6.12234	28742	552.13	<.0001
X1	−1.18183	0.13953	3734.51659	71.74	<.0001
X2	−0.52649	0.12868	871.40181	16.74	0.0005
X4	1.47264	1.11493	90.81673	1.74	0.2008

6-4 Other Aspects of Regression

question 6-56 (c)

Summary of Backward Elimination								
Step	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x3d		3	0.0001	0.8980	3.0123	0.01	0.9129
2	X4	Anxiety	2	0.0085	0.8896	2.6748	1.74	0.2008

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	143.03657	6.19276	28711	533.49	<.0001
X1	-1.07933	0.11791	4509.85941	83.80	<.0001
X2	-0.49856	0.12906	803.07461	14.92	0.0008

6-4 Other Aspects of Regression

question 6-56 (d)

OBS	_MODEL_	_TYPE_	_DEPVAR_	_RMSE_	_PRESS_	Intercept	X1	X2	Y	X4	x3d	_IN_	_P_	_EDF_	_RSQ_	_ADJRSQ_	_CP_
1	m1	PARMS	Y	7.33605	1581.11	143.037	−1.07933	−0.49856	−1	.	.	2	3	22	0.88955	0.87951	3
2	m2	PARMS	Y	7.21496	1602.92	143.860	−1.18183	−0.52649	−1	1.47264	.	3	4	21	0.89803	0.88346	4
3	m3	PARMS	Y	7.50777	1708.81	142.813	−1.07665	−0.49891	−1	.	0.22845	3	4	21	0.88958	0.87381	4