



SAS,SPSS를 활용 (7강~10강 중심)

정보통계학과 김성수교수

✓ 학습목차

1

SAS를 이용한 변수선택, 모형개발, 진단

2

SPSS를 이용한 변수선택, 모형개발, 진단

1

SAS 활용

1. 다중공선성 구하기 : 94페이지 자료

<해군병원의 인력 자료>

X_1	X_2	X_3	X_4	X_5	Y
15.57	2463	472.92	18.0	4.45	566.52
44.02	2048	1339.75	9.5	6.92	696.82
20.42	3940	620.25	12.8	4.28	1033.15
18.74	6505	568.33	36.7	3.90	1603.62
49.20	5723	1497.60	35.7	5.50	1611.37
44.92	11520	1365.83	24.0	4.60	1613.27
55.48	5779	1687.00	43.3	5.62	1854.17
59.28	5969	1639.92	46.7	5.15	2160.55
94.39	8461	2872.33	78.7	6.18	2305.58
128.02	20106	3655.08	180.5	6.15	3503.93
96.00	13313	2912.00	60.9	5.88	3571.89
131.42	10771	3921.00	103.7	4.88	3741.40
127.21	15543	3865.67	126.8	5.50	4026.52
252.90	36194	7684.10	157.7	7.00	10343.81
409.20	34703	12446.33	169.4	10.78	11732.17
463.70	39204	14098.40	331.4	7.05	15414.94
510.22	86533	15524.00	371.6	6.35	18854.45

Y : 월간 의사 연 근무시간

X_1 : 일 평균 환자수

X_2 : 월간 X-ray 촬영 횟수

X_3 : 월간 이용 병석수

X_4 : 해당지역의 병원 이용가능 인구 / 1000

X_5 : 평균입원일

hospital_sas - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

15.57	2463	472.92	18.0	4.45	566.52
44.02	2048	1339.75	9.5	6.92	696.82
20.42	3940	620.25	12.8	4.28	1033.15
18.74	6505	568.33	36.7	3.90	1603.62
49.20	5723	1497.60	35.7	5.50	1611.37
44.92	11520	1365.83	24.0	4.60	1613.27
55.48	5779	1687.00	43.3	5.62	1854.17
59.28	5969	1639.92	46.7	5.15	2160.55
94.39	8461	2872.33	78.7	6.18	2305.58
128.02	20106	3655.08	180.5	6.15	3503.93
96.00	13313	2912.00	60.9	5.88	3571.89
131.42	10771	3921.00	103.7	4.88	3741.40
127.21	15543	3865.67	126.8	5.50	4026.52
252.90	36194	7684.10	157.7	7.00	10343.81
409.20	34703	12446.33	169.4	10.78	11732.17

1) 다중공선성-①

```
DATA navyhospital ;  
  INFILE "c:\data\reg\hospital_sas.txt " ;  
  INPUT X1-X5 Y ;  
  RUN;  
PROC REG ;  
  MODEL Y=X1-X5/ vif ;  
RUN;
```

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	490177488	98035498	237.79	<.0001
Error	11	4535052	412277		
Corrected Total	16	494712540			

Root MSE	642.08838	R-Square	0.9908
Dependent Mean	4978.48000	Adj R-Sq	0.9867
Coeff Var	12.89728		

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
Intercept	1	1962.94816	1071.36170	1.83	0.0941	0
X1	1	-15.85167	97.65299	-0.16	0.8740	9597.57076
X2	1	0.05593	0.02126	2.63	0.0234	7.94059
X3	1	1.58962	3.09208	0.51	0.6174	8933.08650
X4	1	-4.21867	7.17656	-0.59	0.5685	23.29386
X5	1	-394.31412	209.63954	-1.88	0.0867	4.27984

2) 다중공선성-② : X1 제거시

```
DATA navyhospital ;
  INFILE "c:\data\reg\hospital_sas.txt " ;
  INPUT X1-X5 Y ;
  RUN;
PROC REG ;
  MODEL Y=X1-X5/ vif ;
  RUN;
PROC REG ;
  MODEL Y=X2-X5/ vif ;
  RUN;
```

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	490166625	122541656	323.48	<.0001
Error	12	4545916	378826		
Corrected Total	16	494712540			

Root MSE	615.48868	R-Square	0.9908
Dependent Mean	4978.48000	Adj R-Sq	0.9877
Coeff Var	12.36298		

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Variance Inflation
Intercept	1	2032.18806	942.07483	2.16	0.0520	0
X2	1	0.05608	0.02036	2.75	0.0175	7.92580
X3	1	1.08837	0.15340	7.10	<.0001	23.92684
X4	1	-5.00407	5.08071	-0.98	0.3441	12.70597
X5	1	-410.08296	178.07810	-2.30	0.0400	3.36087

2. 모든 가능한 회귀 : 105페이지 자료

```
DATA hald ;
  infile "c:\data\wreg\hald_sas.txt" ;
  input x1-x4 y @@;
run;
proc reg ;
  model y = x1-x4 / selection=rsquare
                                mse adjrsq cp aic ;
run;
```

Number in Model	R-Square	Adjusted R-Square	C(p)	AIC	MSE	Variables in Model
1	0.6745	0.6450	138,7308	58,8516	80,35154	x4
1	0.6663	0.6359	142,4864	59,1780	82,39421	x2
1	0.5339	0.4916	202,5488	63,5195	115,06243	x1
1	0.2859	0.2210	315,1543	69,0674	176,30913	x3
2	0.9787	0.9744	2,6782	25,4200	5,79045	x1 x2
2	0.9725	0.9670	5,4959	28,7417	7,47621	x1 x4
2	0.9353	0.9223	22,3731	39,8526	17,57380	x3 x4
2	0.8470	0.8164	62,4377	51,0371	41,54427	x2 x3
2	0.6801	0.6161	138,2259	60,6293	86,88801	x2 x4
2	0.5482	0.4578	198,0947	65,1167	122,70721	x1 x3
3	0.9823	0.9764	3,0182	24,9739	5,33030	x1 x2 x4
3	0.9823	0.9764	3,0413	25,0112	5,34562	x1 x2 x3
3	0.9813	0.9750	3,4968	25,7276	5,64846	x1 x3 x4
3	0.9728	0.9638	7,3375	30,5759	8,20162	x2 x3 x4
4	0.9824	0.9736	5,0000	26,9443	5,98295	x1 x2 x3 x4

파일(F)	편집(E)	서식(O)	보기(V)	도움말(H)
7	26	6	60	78.5
1	29	15	52	74.3
11	56	8	20	104.3
11	31	8	47	87.6
7	52	6	33	95.9
11	55	9	22	109.2
3	71	17	6	102.7
1	31	22	44	72.5
2	54	18	22	93.1
21	47	4	26	115.9
1	40	23	34	83.8
11	66	9	12	113.3
10	68	8	12	109.4

3. 앞으로부터 선택 : 105페이지 자료

```
DATA hald ;  
  infile "c:\data\wreg\hald_sas.txt" ;  
  input x1-x4 y @@;  
  run;  
proc reg ;  
  model y = x1-x4 / selection=forward ;  
run;
```

Summary of Forward Selection

Step	Variable Entered	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x4	1	0,6745	0,6745	138,731	22,80	0,0006
2	x1	2	0,2979	0,9725	5,4959	108,22	<,0001
3	x2	3	0,0099	0,9823	3,0182	5,03	0,0517

4. 뒤로부터의 제거 : 105페이지 자료

```
DATA hald ;  
  infile "c:\data\wreg\hald_sas.txt" ;  
  input x1-x4 y @@;  
run;  
proc reg ;  
  model y = x1-x4 / selection=backward ;  
run;
```

Summary of Backward Elimination							
Step	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x3	3	0,0000	0,9823	3,0182	0,02	0,8959
2	x4	2	0,0037	0,9787	2,6782	1,86	0,2054

5. 단계별 선택 : 105페이지 자료

```
DATA hald ;  
  infile "c:\data\wreg\hald_sas.txt" ;  
  input x1-x4 y @@;  
  run;  
proc reg ;  
  model y = x1-x4 / selection=stepwise ;  
run;
```

Summary of Stepwise Selection								
Step	Variable Entered	Variable Removed	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	x4		1	0,6745	0,6745	138,731	22,80	0,0006
2	x1		2	0,2979	0,9725	5,4959	108,22	<,0001
3	x2		3	0,0099	0,9823	3,0182	5,03	0,0517
4		x4	2	0,0037	0,9787	2,6782	1,86	0,2054

6. 다항회귀 : 122페이지 자료

```
data tcrime ;
    infile "c:\data\reg\tcrime_sas.txt" ;
    input yesr tcratio motor ;
run;
data crime ;
    set tcrime ;
    smotor = motor*motor ;
run;
proc gplot data=crime;
    plot tcratio*motor ;
run;
proc reg data=crime ;
    model tcratio = motor smotor ;
run;
```

tcrime_sas - 메모장

국립(가)	범죄(가)	서식(가)	분기(가)	도움말(가)
1966	0.9	59.5		
1967	2.0	77.1		
1968	4.1	107.4		
1969	5.6	159.5		
1970	12.0	188.2		
1971	23.6	205.5		
1972	22.5	209.7		
1973	24.2	229.7		
1974	21.3	220.4		
1975	40.8	236.7		
1976	74.3	257.8		
1977	91.2	345.0		
1978	118.3	500.1		
1979	128.8	643.2		
1980	84.6	653.4		
1981	51.1	691.1		
1982	267.7	777.6		
1983	398.1	954.6		
1984	436.6	1151.2		
1985	497.2	1364.2		
1986	519.1	1612.8		

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	2830430	1415215	454.45	<.0001
Error	22	68511	3114.15786		
Corrected Total	24	2898941			

Root MSE	55.80464	R-Square	0.9764
Dependent Mean	263.38400	Adj R-Sq	0.9742
Coeff Var	21.18756		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-74.50324	18.56260	-4.01	0.0006
motor	1	0.45395	0.03041	14.93	<.0001
smotor	1	-0.00004149	0.00000687	-6.04	<.0001

7. 가변수회귀모형 : 128페이지 자료

```
data soup ;
  infile "c:\data\reg\soup_sas.txt" ;
  input Y X D ;
run;
data nsoup ;
  set soup ;
  XD = X*D ;
run;
proc reg;
  model Y=X D ;
  model Y = X D XD ;
run;
```

soup_sas - 메모장

파일(F)	편집(E)	서식(O)	보기(V)	도움
218	100	1		
248	125	1		
360	220	1		
351	205	1		
470	300	1		
394	255	1		
332	225	1		
321	175	1		
410	270	1		
260	170	1		
241	155	1		
331	190	1		
275	140	1		
425	290	1		
367	265	1		
140	105	0		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	27,28179	15,40701	1,77	0,0893
X	1	1,23074	0,06555	18,77	<,0001
D	1	53,12920	8,21003	6,47	<,0001

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	7,57446	20,86970	0,36	0,7200
X	1	1,32205	0,09262	14,27	<,0001
D	1	90,39086	28,34573	3,19	0,0041
XD	1	-0,17666	0,12884	-1,37	0,1835

8. 자료진단 : 157페이지 자료

```
data soil ;
  infile "c:\Wdata\reg\soil_sas.txt" ;
  input id SL SG LOBS PGC ;
run;
proc reg ;
  model SL = SG LOBS PGC
    / r influence ;
run;
```

soil_sas - 메모창

파일(F)	편집(E)	서식(O)	보기(V)	도움말(H)
1	27.1	0.43	1.95	0.34
2	35.6	0.47	5.13	0.32
3	31.4	0.44	3.98	0.29
4	37.8	0.48	6.25	0.30
5	40.2	0.48	7.12	0.25
6	39.8	0.49	6.50	0.26
7	55.5	0.53	10.67	0.10
8	43.6	0.50	7.08	0.16
9	52.1	0.55	9.88	0.19
10	43.8	0.51	8.72	0.18
11	35.7	0.48	4.96	0.28

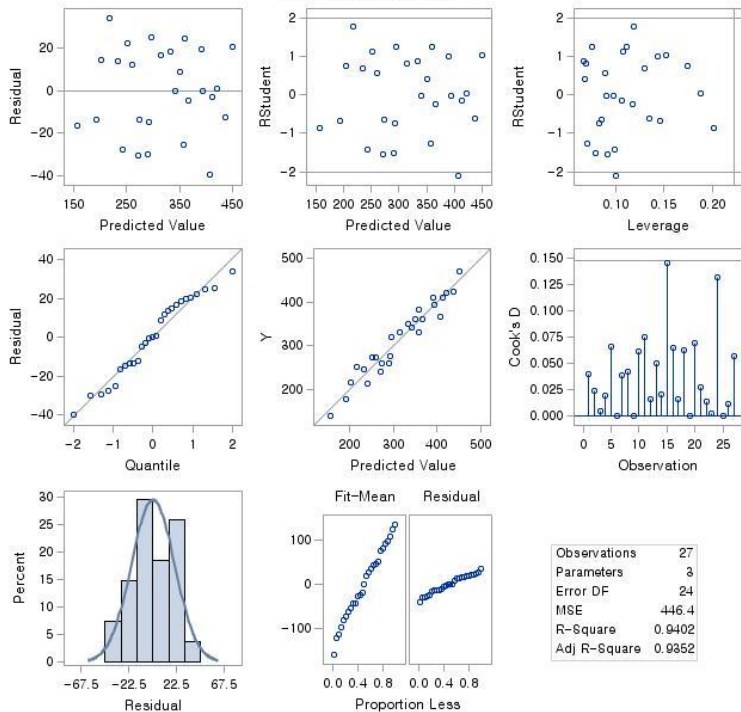
Output Statistics									
Obs	Dependent Variable	Predicted Value	Std Error Mean Predict	Residual	Std Error Residual	Student Residual	Cook's D	RStudent	Hat Diag H
1	27.1	26.2837	1.0332	0.8163	1.110	0.736	0.117	0.7089	0.4643
2	35.6	34.8128	0.7545	0.7872	1.315	0.599	0.029	0.5689	0.2476
3	31.4	31.4171	0.9129	-0.0171	1.211	-0.014	0.000	-0.0131	0.3625
4	37.8	37.8103	0.8292	-0.0103	1.269	-0.008	0.000	-0.007511	0.2991
5	40.2	40.3619	0.8731	-0.1619	1.240	-0.131	0.002	-0.1211	0.3316
6	39.8	39.9295	0.5204	-0.1295	1.424	-0.091	0.000	-0.0842	0.1178
7	55.5	53.3486	1.1066	2.1514	1.037	2.075	1.227	3.0976	0.5326
8	43.6	43.9974	1.1040	-0.3974	1.039	-0.382	0.041	-0.3577	0.5301
9	52.1	51.5121	1.2021	0.5879	0.924	0.636	0.171	0.6068	0.6285
10	43.8	46.8495	0.6582	-3.0495	1.366	-2.232	0.289	-3.8510	0.1884
11	35.7	36.2771	0.8271	-0.5771	1.271	-0.454	0.022	-0.4268	0.2976

9. 모형진단 : 디폴트 출력

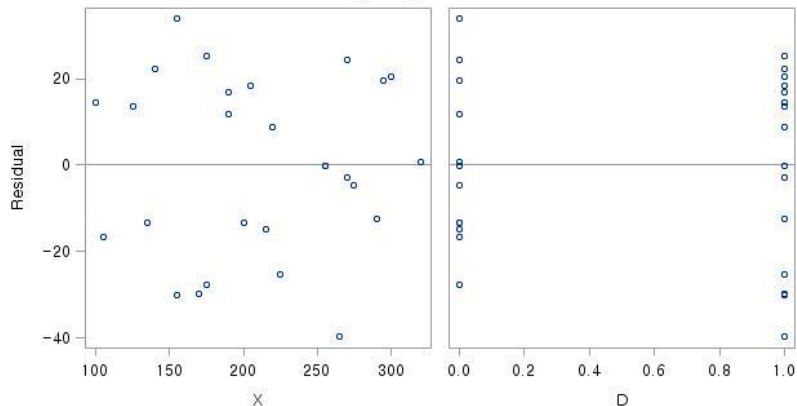
SAS 시스템

The REG Procedure
Model: MODEL1
Dependent Variable: Y

Fit Diagnostics for Y



Residual by Regressors for Y



2

SPSS 활용

1. 다중공선성 구하기 : 94페이지 자료

hospital_sas - 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

15.57	2463	472.92	18.0	4.45	566.52
44.02	2048	1339.75	9.5	6.92	696.82
20.42	3940	620.25	12.8	4.28	1033.15
18.74	6505	568.33	36.7	3.90	1603.12
49.20	5723	1497.60	35.7	5.50	1611.89
44.92	11520	1365.83	24.0	4.60	1613.39
55.48	5779	1687.00	43.3	5.62	1854.12
59.28	5969	1639.92	46.7	5.15	2160.55
94.39	8461	2872.33	78.7	6.18	2305.58
128.02	20106	3655.08	180.5	6.15	3503.93
96.00	13313	2912.00	60.9	5.88	3571.89
131.42	10771	3921.00	103.7	4.88	3741.40
127.21	15543	3865.67	126.8	5.50	4026.52
252.90	36194	7684.10	157.7	7.00	10343.81
409.20	34703	12446.33	169.4	10.78	11732.12

텍스트 가져오기 마법사 - 6단계 중 2단계

변수는 어떻게 배열되어 있습니까?

☐ 구분자에 의한 배열(D) - 변수는 특정 문자(예: 싹표, 탭)로 구분됩니다.

☒ 고정 너비로 배열(E) - 변수는 고정 너비 열로 정렬됩니다.

변수 이름이 파일의 처음에 있습니까?

☐ 예(Y)

☒ 아니오

텍스트 파일: C:\data\reg\hospital_sas.txt

	0	10	20	30	40	50	60
1	15.57	2463	472.92	18.0	4.45	566.52	
2	44.02	2048	1339.75	9.5	6.92	696.82	
3	20.42	3940	620.25	12.8	4.28	1033.15	

텍스트 가져오기 마법사 - 6단계 중 4단계(고정 너비로 배열)

각 변수가 시작되는 위치를 지정하십시오. 첫 번째 열은 열 0입니다.

변수 구분선을 삽입하려면 눈금자 또는 데이터 영역에서 원형 커를 사용하여 위치로 이동하거나 열 번호를 입력한 다음 구분

변수 구분선을 이동하려면 새 위치로 끄십시오.

변수 구분선을 삭제하려면 위치를 선택하거나 입력하십시오. 추를 누르십시오.

	0	10	20	30	40	
8	59.28	5969	1639.92	46.7	5.15	2160.55
9	94.39	8461	2872.33	78.7	6.18	2305.58
10	128.02	20106	3655.08	180.5	6.15	3503.93
11	96.00	13313	2912.00	60.9	5.88	3571.89
12	131.42	10771	3921.00	103.7	4.88	3741.40
13	127.21	15543	3865.67	126.8	5.50	4026.52
14	252.90	36194	7684.10	157.7	7.00	10343.81
15	409.20	34703	12446.33	169.4	10.78	11732.12

제목없음4 [데이터세트3] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A)

새 파일(N)

열기(O)

데이터베이스 열기(B)

☒ 텍스트 데이터 읽기(D)...

Cognos 데이터 읽기...

Triple-S 데이터 읽기

변수

1) 다중공선성

선형 회귀

종속변수(D): Y

블록(B) 1/1

미전(V) 다음(N)

독립변수(I): X1, X2, X3

방법(M): 입력

선택변수(E):

케이스 레이블(C):

WLS 가중값(H):

통계량(S)... 도표(T)... 저장(S)... 옵션(O)... 유형(L)...

확인 붙여넣기(P) 재설정(R) 취소 도움말

선형 회귀: 통계량

회귀계수

☒ 추정값(E) ☐ R 제곱 변화량(S)

☐ 신뢰구간(N) ☐ 기술통계(D)

수준(%): 95 ☐ 부분상관 및 편상관계수(P)

☐ 공분산 행렬(V) ☒ 공선성 진단(L)

잔차

☐ Durbin-Watson

☐ 케이스별 진단(C)

☒ 밖에 나타나는 이상값(O): 3 표준편차

☒ 모든 케이스(A)

계속(C) 취소 도움말

계수^a

모형		비표준화 계수		표준화 계수	t	유의확률	공선성 통계량	
		B	표준오차	베타			공차	VIF
1	(상수)	1962.948	1071.362		1.832	.094		
	X1	-15.852	97.653	-.459	-.162	.874	.000	9597.571
	X2	.056	.021	.214	2.631	.023	.126	7.941
	X3	1.590	3.092	1.403	.514	.617	.000	8933.087
	X4	-4.219	7.177	-.082	-.588	.569	.043	23.294
	X5	-394.314	209.640	-.112	-1.881	.087	.234	4.280

a. 종속변수: Y

2. 앞으로부터 선택 : 105페이지 자료

※ SPSS는 모든 가능한 회귀절차는 제공 않음

hald.sav [데이터세트] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 그래프(G) 유틸리티(U) 창(W) 도움말(H)

표시: 5 / 5 변수

	X1	X2	X3	X4	Y
1	7	26	6	60	
2	1	29	15	52	
3	11	56	8	20	
4	11	31	8	47	
5	7	52	6	33	
6	11	55	9	22	
7	3	71	17	6	
8	1	31	22	44	
9	2	54	18	22	
10	21	47	4	26	
11	1	40	23	34	
12	11	66	9	12	
13	10	68	8	12	

선택 회귀

종속변수(D): Y

블록(B) 1/1

미전(V) 다음(N)

독립변수(I): X1, X2, X3

방법(M): 입력

선택변수(E):

케이스 레이블(O):

WLS 가중값(H):

통계량(S)... 도표(T)... 저장(S)... 옵션(O)... 유형(L)...

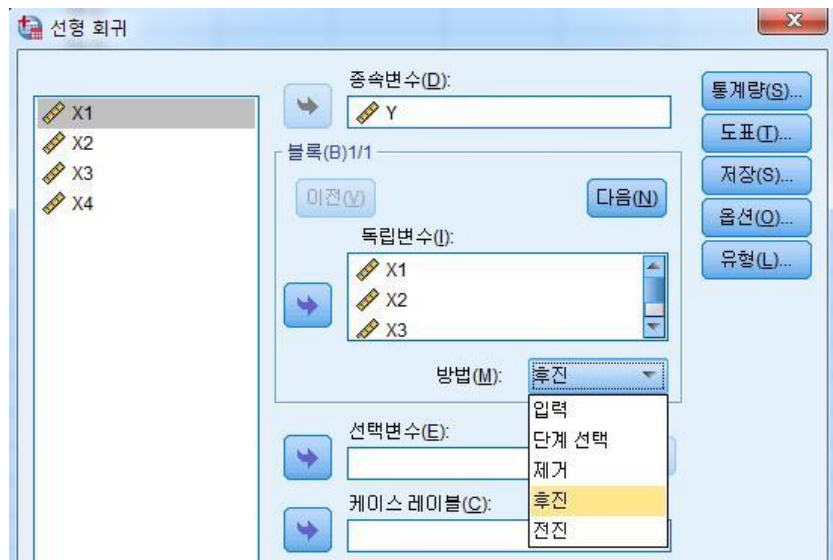
확인 붙여넣기(P) 재설정(R) 취소 도움말

계수^a

모형		비표준화 계수		표준화 계수	t	유의확률
		B	표준오차	베타		
1	(상수)	117.568	5.262		22.342	.000
	X4	-.738	.155	-.821	-4.775	.001
2	(상수)	103.097	2.124		48.540	.000
	X4	-.614	.049	-.683	-12.621	.000
	X1	1.440	.138	.563	10.403	.000

a. 종속변수: Y

3. 뒤로부터의 제거 : 105페이지 자료



계수^a

모형		비표준화 계수		표준화 계수	t	유의확률
		B	표준오차	베타		
1	(상수)	62.405	70.071		.891	.399
	X1	1.551	.745	.607	2.083	.071
	X2	.510	.724	.528	.705	.501
	X3	.102	.755	.043	.135	.896
	X4	-.144	.709	-.160	-.203	.844
2	(상수)	71.648	14.142		5.066	.001
	X1	1.452	.117	.568	12.410	.000
	X2	.416	.186	.430	2.242	.052
	X4	-.237	.173	-.263	-1.365	.205
3	(상수)	52.577	2.286		22.998	.000
	X1	1.468	.121	.574	12.105	.000
	X2	.662	.046	.685	14.442	.000

a. 종속변수: Y

4. 단계별 선택 : 105페이지 자료



ANOVA^a

모형		제곱합	자유도	평균제곱	F	유의확률
1	회귀	1831.896	1	1831.896	22.799	.001 ^b
	잔차	883.867	11	80.352		
	전체	2715.763	12			
2	회귀	2641.001	2	1320.500	176.627	.000 ^c
	잔차	74.762	10	7.476		
	전체	2715.763	12			

a. 종속변수: Y

b. 예측자: (상수), X4

c. 예측자: (상수), X4, X1

계수^a

모형		비표준화 계수		표준화 계수	t	유의확률
		B	표준오차	베타		
1	(상수)	117.568	5.262		22.342	.000
	X4	-.738	.155	-.821	-4.775	.001
2	(상수)	103.097	2.124		48.540	.000
	X4	-.614	.049	-.683	-12.621	.000
	X1	1.440	.138	.563	10.403	.000

a. 종속변수: Y

5. 다항회귀 : 122페이지 자료

IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 도움말(H)

변수 계산

목록변수(O): smotor = 숫자표현식(E): motor ** 2

유형 및 레이블(L)...

year tcratio motor

year tcratio motor

합수 집단(G): 모두 산술 CDF 및 비종심 CDF 변환 현재 날짜/시간 날짜 산술 날짜 작성

합수 및 특수변수(E):

조건(O)... (선택적 케이스 선택 조건)

확인 붙여넣기(B) 재설정(R) 취소 도움말

ANOVA^a

모형	제공함	자유도	평균제곱	F	유의확률
1 회귀	2830429.921	2	1415214.960	454.445	.000 ^b
잔차	68511.473	22	3114.158		
전체	2898941.394	24			

a. 종속변수: tcratio

b. 예측자: (상수), smotor, motor

계수^a

모형		비표준화 계수		표준화 계수		t	유의확률
		B	표준오차	베타			
1	(상수)	-74.503	18.563			-4.014	.001
	motor	.454	.030	1.572		14.930	.000
	smotor	-4.149E-5	.000	-.636		-6.036	.000

a. 종속변수: tcratio

6. 가변수회귀모형 : 128페이지 자료

soup.sav [데이터세트10] - IBM SPSS Statistics Data Editor

	Y	X	D
7	332	225	1
8	321	175	1
9	410	270	1
10	260	170	1
11	241	155	1
12	331	190	1
13	275	140	1
14	425	290	1
15	367	265	1
16	140	105	0
17	277	215	0
18	384	270	0

soup.sav [데이터세트10] - IBM SPSS Statistics Data Editor

	이름	유형	너비	소수점이...	레이블	값	정확값	범	맞춤	측도	역할
1	Y	숫자	3	0		없음	없음	8	오른쪽	척도	입력
2	X	숫자	8	0		없음	없음	8	오른쪽	척도	입력
3	D	숫자	19	0		없음	없음	6	오른쪽	명목형	입력

ANOVA^a

모형		제곱합	자유도	평균제곱	F	유의확률
1	회귀	168355.061	2	84177.531	188.568	.000 ^b
	잔차	10713.680	24	446.403		
	전체	179068.741	26			

a. 종속변수: Y

b. 예측자: (상수), D, X

계수^a

모형		비표준화 계수		표준화 계수	t	유의확률
		B	표준오차	베타		
1	(상수)	27.282	15.407		1.771	.089
	X	1.231	.066	.941	18.775	.000
	D	53.129	8.210	.324	6.471	.000

a. 종속변수: Y

7. 자료진단 : 157페이지 자료

soil.sav [데이터세트12] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 그래프(G) 유틸리티(U) 창(W) 도움말(H)

표시: 5 / 5 변수

	id	SL	SG	LOBS	PGC	변수	변수	변수	변수	변수	변수
1	1	27.1	.43	1.95	.34						

선택 회귀

종속변수(D): SL

블록(B) 1/1

이전(V) 다음(N)

독립변수(I): SG, LOBS, PGC

방법(M): 입력

선택변수(E):

케이스 레이블(C):

WLS 가중값(H):

통계량(S)...

도표(D)...

저장(S)...

옵션(O)...

유형(L)...

확인 불어넣기(B) 재설정(R) 취소 도움말

선택 회귀: 통계량

회귀계수

☒ 추정값(E) ☐ R 제곱 변화량(S)

☐ 신뢰구간(N) ☐ 기술통계(D)

수준(%): 95

☐ 공분산 행렬(V) ☐ 부분상관 및 편상관계수(P)

☐ 공선성 진단(L)

잔차

☐ Durbin-Watson

☒ 케이스별 진단(Q)

☐ 밖에 나타나는 이상값(O): 3 표준편차

☒ 모든 케이스(A)

계속(C) 취소 도움말

선택 회귀: 저장

예측값

☐ 비표준화(U)

☐ 표준화(R)

☐ 수정된(J)

☐ 평균예측 표준오차(P)

잔차

☐ 비표준화(N)

☒ 표준화(A)

☒ 스튜던트화(S)

☒ 제외 잔차(L)

☒ 삭제된 스튜던트화 잔차(E)

거리

☐ Mahalanobis의 거리

☒ Cook의 거리

☐ 레버리지 값(G)

영향력 통계량

☐ DFBETA

☐ 표준화 DFBETA(Z)

☐ DEFIT

☐ 표준화 DFFIT

☐ 공분산 비율(V)

예측구간

☐ 평균(M) ☐ 개별값(I)

신뢰구간(C): 95 %

계수 통계량

☐ 계수 통계량 만들기(Q)

☒ 새 데이터 세트 만들기

데이터 세트 이름(O):

☒ 새 데이터 파일 쓰기

파일(L)...

XML 파일로 모형정보 내보내기

찾아보기(W)...

☒ 공분산 행렬 포함(X)

계속(C) 취소 도움말

1) 자료진단 : 157페이지 자료

계수^a

모형	비표준화 계수		표준화 계수	t	유의확률
	B	표준오차	베타		
1 (상수)	-1.879	18.134		-.104	.920
SG	77.326	44.505	.326	1.737	.126
LOBS	1.559	.734	.479	2.123	.071
PGC	-23.904	13.429	-.216	-1.780	.118

a. 종속변수: SL

케이스별 진단^a

케이스 번호	표준화 잔차	SL	예측값	잔차
1	.538	27.1	26.284	.8163
2	.519	35.6	34.813	.7872
3	-.011	31.4	31.417	-.0171
4	-.007	37.8	37.810	-.0103
5	-.107	40.2	40.362	-.1619
6	-.085	39.8	39.929	-.1295
7	1.419	55.5	53.349	2.1514
8	-.262	43.6	43.997	-.3974
9	.388	52.1	51.512	.5879
10	-2.011	43.8	46.850	-3.0495
11	-.381	35.7	36.277	-.5771

a. 종속변수: SL

*soil.sav [데이터세트12] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 그래프(G) 유틸리티(U) 창(W) 도움말(H)

	LOBS	PGC	DRE_1	ZRE_1	SRE_1	SDR_1	COO_1
1	1.95	.34	1.52381	.53833	.73552	.70891	.11723
2	5.13	.32	1.04628	.51917	.59853	.56888	.02947
3	3.98	.29	-.02690	-.01131	-.01416	-.01311	.00003
4	6.25	.30	-.01469	-.00679	-.00811	-.00751	.00001
5	7.12	.25	-.24221	-.10678	-.13060	-.12106	.00212
6	6.50	.26	-.14677	-.08539	-.09092	-.08422	.00028
7	10.67	.10	4.60263	1.41887	2.07530	3.09761	1.22674
8	7.08	.16	-.84575	-.26208	-.38234	-.35773	.04123
9	9.88	.19	1.58267	.38775	.63618	.60679	.17118
10	8.72	.18	-3.75744	-2.01114	-2.23241	-3.85097	.28924
11	4.96	.28	-.82163	-.38062	-.45414	-.42679	.02184

8. 모형진단

*soil.sav [데이터세트12] - IBM SPSS Statistics Data Editor

파일(F) 편집(E) 보기(V) 데이터(D) 변환(T) 분석(A) 그래프(G) 유틸리티(U) 창(W) 도움말(H)

20 : SRE_1 표시: 10 / 10 변수

	LOBS	PGC	DRE_1	ZRE_1	SRE_1	SDR_1	COO_1	변수	변수
1	1.95	.34	1.52381	.53833	.73552	.70891	.11723		
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									

선형 회귀

종속변수(D): SL

블록(B) 1/1

이전(V) 다음(N)

독립변수(I): SG, LOBS, PGC

방법(M): 입력

선택변수(E):

케이스 레이블(C):

WLS 가중값(H):

확인 불어넣기(P) 재설정(R) 취소 도움말

선형 회귀: 도표

DEPENDNT

*ZPRED

*ZRESID

*DRESID

*ADJPRED

*SRESID

*SDRESID

산점도 1 / 1

이전(V) 다음(N)

Y: *DRESID

X: *ZPRED

표준화 잔차도표

☐ 히스토그램(H)

☒ 정규확률도표(R)

☒ 편회귀잔차도표 모두 출력(P)

계속(C) 취소 도움말



다음시간 안내

12강. 일반선형모형(1)