GLM Example: One-Way Analysis of Covariance

Understanding Design and Analysis of Research Experiments

An animal scientist is interested in determining the effects of four different feed plans on hogs. Twenty four hogs of a breed were chosen and randomly assigned to one of the four feeding plans for certain period. Initial weight (X) of the hogs and gains in weight (Y) in pounds at the end of the experiment are given below:

Feed Plan							
1		2		3		4	
X	Y	X	Y	X	Y	X	Y
30	165	24	180	34	156	41	201
27	170	31	169	32	189	32	173
20	130	20	171	35	138	30	200
21	156	26	161	35	190	35	193
33	167	20	180	30	160	28	142
29	151	25	170	29	172	36	189

Since differences in initial weight may contribute to the differences in gain in weight, it is decided to use initial weight as a covariate in an analysis of covariance.

Linear model: $Y_{ij} = \mu + \alpha_i + \beta(X_{ij} - \overline{X}_{..}) + e_{ij}$

Assumptions:

- Initial weight (X_{ij}) are fixed, measured without error and independent of treatments (feed plans);
- Regression of gain in weight, after removal of treatment effect (i.e., Y_{ij} - $Y_{i.}$), on initial weight, X_{ij}, is linear and independent of treatments;
- The residuals (e_{ij}) are normally and independently distributed with zero mean and a common variance, i.e., $e_{ij} \sim N(0, \ ^2)$.

ANOVA (Gain in weight unadjusted):

Source	df	SS	MS	F
Treatment	3	2163.1	721.0	2.43
Error	20	5937.8	296.9	

CV = 10.2% SEM = 17.2

ANOVA (Gain in weight adjusted for initial weight):

		Sums of products			Y adjusted for X			
	df	(X,X)	(X,Y)	(Y,Y)	df	SS	MS	F
Trt.	3	365.5	451.2	2163.1				
Error	20	361.5	496.8	5937.8	19	5255.1	276.6	
T + E	23	727.0	948.0	8100.9	22	6864.7		
Trt adj.					3	1609.6	536.5	1.94

Test for regression coefficient = 0 (i.e., initial weight has nothing to do with gain in weight):

$$F = \frac{(5937.8 - 5255.1)/1}{(5255.1)/19} = \frac{682.7/1}{5255.1/19} = 2.47$$

Test for homogeneity of within-treatment regressions (cf. Steel & Torrie 1980, Table 17.6):

$$A = \text{pooled error} = \sum_{i=1}^{4} SS_i \text{ (residuals)} = 4196.0$$

$$B = \sum_{i=1}^{4} E_{yy}(i) - \frac{\left[\sum_{i=1}^{4} E_{xy}(i)\right]^2}{\sum_{i=1}^{4} E_{xx}(i)} = 5937.8 - \frac{(496.8)^2}{361.5} = 5937.8 - 682.7 = 5254.1$$

$$F = \frac{(B - A)/(4 - 1)}{A/(24 - 8)} = \frac{(5254.1 - 4196.0)/(4 - 1)}{(4196.0)/(24 - 8)} = \frac{(1058.1)/3}{(4196)/16} = \frac{352.7}{262.3} = 1.35$$

Adjusted means:

$$\hat{\beta} = \frac{496.8}{361.5} = 1.37$$

$$\bar{Y}'_{1.} = 156.5 - (1.37)(26.7 - 29.3) = 160.1$$

$$SE(\bar{Y}'_{1.}) = \sqrt{276.6\left[\frac{1}{6} + \frac{(26.7 - 29.3)^2}{361.5}\right]} = 7.16$$

SAS programs to carry out the analysis of covariance for the above data

```
options nodate nonumber;
title 'GLM Example: One-Way Analysis of Covariance';
data raw;
  input trt x y;
cards;
1      30      165
1      27      170
1      20      130
1      21      156
1      33      167
1      29      151
2      24      180
2      31      169
2      20      171
2      26      161
2      20      180
2      25      170
3      34      156
3      32      189
3      35      138
3      35      190
3      30      160
3      29      172
4      41      201
4      32      173
4      30      200
4      35      193
4      28      142
4      36      189
;

title2 'ANOVA (Gain in weight unadjusted)';
proc anova manova outstat=newstat data=raw;
  class trt;
  model x y=trt;
run;

title2 'Output cross-product matrics';
proc print data=newstat;
run;

title2 'ANOVA (Gain in weight adjusted): SAS Type III SS should be used';
proc glm data=raw;
```

```
  class trt;
  model y=trt x/solution ss1 ss2 ss3 ss4;
  lsmeans trt / stderr pdiff cov out=adjmeans;
run;

title2 'Output adjusted means';
proc print data=adjmeans;

run;

/*****************************************************************
 Compute necessary statistics to test for homogeneity of regressions
 *****************************************************************/
proc sort data=raw out=new;
  by trt;

proc corr data=new outp=sscp csscp noprint;
  var x y;
  by trt;
run;

data sscp; set sscp;
  if _type_='CSSCP';

data s1; set sscp(keep=trt _type_ _name_ x);
  if _name_='X';

data s2(drop=_name_ rename=(x=xy)); set sscp(keep= _name_ x);
  if _name_='Y';

data s3(drop=_name_); set sscp(keep= _name_ y);
  if _name_='Y';

data snew; merge s1 s2 s3;
  drop _name_ _type_;
  rss=y-(xy*xy/x);

title2 'Output statistics needed to test for homogeneity of regressions';
proc means data=snew n sum;
  var x xy y rss;

run;
```

SAS OUTPUT

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight unadjusted)

Analysis of Variance Procedure
Class Level Information

Class	Levels	Values
TRT	4	1 2 3 4

Number of observations in data set = 24

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight unadjusted)

Analysis of Variance Procedure

Dependent Variable: X

Source	DF	Sum of Squares	Mean Square	F Valu
Model	3	365.45833333	121.81944444	6.7
Error	20	361.50000000	18.07500000	
Corrected Total	23	726.95833333		

R-Square	C.V.	Root MSE
0.502723	14.51427	4.25147033

Source	DF	Anova SS	Mean Square	F Valu
TRT	3	365.45833333	121.81944444	6.7

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight unadjusted)

Analysis of Variance Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Valu
Model	3	2163.12500000	721.04166667	2.4
Error	20	5937.83333333	296.89166667	
Corrected Total	23	8100.95833333		

R-Square	C.V.	Root MSE
0.267021	10.15303	17.23054458

Source	DF	Anova SS	Mean Square	F Valu
TRT	3	2163.12500000	721.04166667	2.4

GLM Example: One-Way Analysis of Covariance
Output cross-product matrics

OBS	_NAME_	_SOURCE_	_TYPE_	X	Y	DF	SS	F
1	X	ERROR	ERROR	361.500	496.83	20	361.50	.
2	Y	ERROR	ERROR	496.833	5937.83	20	5937.83	.
3	X	TRT	ANOVA	365.458	451.21	3	365.46	6.73966
4	Y	TRT	ANOVA	451.208	2163.13	3	2163.13	2.42864

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight adjusted): SAS Type III SS should be used

General Linear Models Procedure
Class Level Information

Class	Levels	Values

TRT 4 1 2 3 4

Number of observations in data set = 24

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight adjusted): SAS Type III SS should be used

General Linear Models Procedure

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Valu
Model	4	2845.95587444	711.48896861	2.5
Error	19	5255.00245889	276.57907678	
Corrected Total	23	8100.95833333		

R-Square	C.V.	Root MSE
0.351311	9.799558	16.63066676

Source	DF	Type I SS	Mean Square	F Valu
TRT	3	2163.12500000	721.04166667	2.6
X	1	682.83087444	682.83087444	2.4

Source	DF	Type II SS	Mean Square	F Valu
TRT	3	1609.59477846	536.53159282	1.9
X	1	682.83087444	682.83087444	2.4

Source	DF	Type III SS	Mean Square	F Valu
TRT	3	1609.59477846	536.53159282	1.9
X	1	682.83087444	682.83087444	2.4

Source	DF	Type IV SS	Mean Square	F Valu
TRT	3	1609.59477846	536.53159282	1.9
X	1	682.83087444	682.83087444	2.4

Parameter		Estimate	T for H0: Parameter=0	Pr > |T|	Std Es
INTERCEPT		136.7296757 B	4.52	0.0002	30.
TRT	1	-16.8794375 B	-1.48	0.1547	11.
	2	1.6607500 B	0.13	0.8965	12.
	3	-13.8965729 B	-1.44	0.1664	9.
	4	0.0000000 B	.	.	.
X		1.3743661	1.57	0.1326	0.

NOTE: The X'X matrix has been found to be singular and a generalized inverse was use the normal equations. Estimates followed by the letter 'B' are biased, and a unique estimators of the parameters.

GLM Example: One-Way Analysis of Covariance
ANOVA (Gain in weight adjusted): SAS Type III SS should be used

General Linear Models Procedure
Least Squares Means

TRT	Y LSMEAN	Std Err LSMEAN	Pr > |T| H0:LSMEAN=0	Pr > |T| H0: LSMEAN(i)=LSMEAN i/j	1	2	3	
1	160.107711	7.167178	0.0001	1	.	0.0743	0.7868	0.
2	178.647898	8.056441	0.0001	2	0.0743	.	0.2092	0.
3	163.090576	7.346555	0.0001	3	0.7868	0.2092	.	0.
4	176.987148	7.793636	0.0001	4	0.1547	0.8965	0.1664	.

NOTE: To ensure overall protection level, only probabilities associated with pre-pla
comparisons should be used.

GLM Example: One-Way Analysis of Covariance
Output adjusted means

OBS	_NAME_	TRT	LSMEAN	STDERR	NUMBER	COV1	COV2	COV3
1	Y	1	160.108	7.16718	1	51.3684	9.9581	-6.4435
2	Y	2	178.648	8.05644	2	9.9581	64.9062	-12.1710
3	Y	3	163.091	7.34655	3	-6.4435	-12.1710	53.9719
4	Y	4	176.987	7.79364	4	-8.7866	-16.5968	10.7391

GLM Example: One-Way Analysis of Covariance
Output statistics needed to test for homogeneity of regressions

Variable	N	Sum
X	4	361.5000000
XY	4	496.8333333
Y	4	5937.83
RSS	4	4196.40

This information is maintained by Dr. Rong-Cai Yang
Last Revised/Reviewed July 14, 1998

[Top of Document]