

8

Checking the models I: independence

New commands are introduced in just one section of this chapter.

8.2 Repeated measures

The main test discusses how to correctly analyse a repeated measures data set. The first option given is the single summary approach, as illustrated here.

SAS COMMANDS FOR BOX 8.3 **Analysing the pigs' final weight**

Commands `proc glm data=gandh.Chapter8;`
`class DIET;`
`model LOGWT60 = DIET /solution;`
`run;`

Menu route Statistics > Anova > Linear Models...

LOGWT60 → Dependent

DIET → Class

Statistics

Parameter Estimates

SAS OUTPUT FOR BOX 8.3 **Analysing the pigs' final weight**

The GLM Procedure

Dependent Variable: LOGWT60

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	0.06122750	0.06122750	4.32	0.0713
Error	8	0.11338770	0.01417346		
Corrected Total	9	0.17461519			

R-Square	Coeff Var	Root MSE	LOGWT60 Mean
0.350642	2.640918	0.119052	4.507992

Source	DF	Type I SS	Mean Square	F Value	Pr > F
DIET	1	0.06122750	0.06122750	4.32	0.0713

Source	DF	Type III SS	Mean Square	F Value	Pr > F
DIET	1	0.06122750	0.06122750	4.32	0.0713

Parameter	Estimate	Standard Error	t Value	Pr > t
Intercept	4.429744000 B	0.05324183	83.20	<.0001
DIET 1	0.156496000 B	0.07529532	2.08	0.0713
DIET 2	0.000000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular, and a generalized inverse was used to solve the normal equations. Terms whose estimates are followed by the letter 'B' are not uniquely estimable.

The second option is a multivariate analysis. SAS does provide a command to analyse these sorts of data sets, and a brief example is given here using the 'one-pig-one-datapoint' dataset (Table 8.2 in the main text). The dataset was first entered manually into SAS and then saved in the *gandh* library under the name *piglets*, but alternatively you can download it from the web site. Using the MANOVA statement, an analysis of variance with multiple Y variables will test the null hypothesis, that there is no difference between the two diets.

SAS COMMANDS FOR BOX 8.4 **Analysis of variance with multiple Y variables**

```

Commands  proc glm data=gandh.piglets;
           class DIET;
           model LOGWT3 LOGWT8 LOGWT20 LOGWT60 = DIET / nouni;
           manova H = DIET ;
           run;

```

Menu route Statistics > Anova > Linear Models...

LOGWT3 LOGWT8 LOGWT20 LOGWT60 → Dependent

DIET → Class

Tests

with Multivariate tab

Effects → H=effect E=error term

DIET → Add

[The phrase “H=DIET E=Residual error matrix” will now appear in the pane
“H=effect E=error term”]

In the command line route, it is possible to suppress the separate univariate analyses with the option ‘nouni’.

SAS OUTPUT FOR BOX 8.4 **Analysis of variance with multiple Y variables**

Multivariate Analysis of Variance

Characteristic Roots and Vectors of: E Inverse * H, where
H = Type III SSCP Matrix for DIET
E = Error SSCP Matrix

Characteristic Root	Percent	Characteristic Vector		V'EV=1	
		LOGWT3	LOGWT8	LOGWT20	LOGWT60
0.71796086	100.00	-2.90840526	1.27357198	1.02144255	3.23377180
0.00000000	0.00	1.75036153	-9.39005823	-2.96208605	7.25431040
0.00000000	0.00	0.31691474	-5.94641261	4.34463629	0.00000000
0.00000000	0.00	5.52247611	-4.36789749	0.00000000	0.00000000

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall DIET Effect
H = Type III SSCP Matrix for DIET
E = Error SSCP Matrix

Statistic	S=1 M=1 N=1.5			Den DF	Pr > F
	Value	F Value	Num DF		
Wilks' Lambda	0.58208544	0.90	4	5	0.5286
Pillai's Trace	0.41791456	0.90	4	5	0.5286
Hotelling-Lawley Trace	0.71796086	0.90	4	5	0.5286
Roy's Greatest Root	0.71796086	0.90	4	5	0.5286

8.6 Exercises

How non-independence can inflate sample size enormously

First of all the variable LUPRATE is calculated, and then analysed. If using the menu route, remember to get the data table into edit mode first (Edit > Mode > Edit).

SAS COMMANDS FOR BOX 8.7 Look-up rate in feeding sheep	
Commands	<pre> data; set gandh.Chapter8; LUPRATE=NLOOKUPS/DURATION; run; proc glm; class OBSPER SEX; model LUPRATE = OBSPER SEX / SS3; run; </pre>
Menu route	<p>Data > Transform > Compute</p> <p> Type LUPRATE to replace compl</p> <p> NLOOKUPS / DURATION → Main pane</p> <p>Statistics > Anova > Linear Models</p> <p> LUPRATE → Dependent</p> <p> OBSPER SEX → Class</p>

SAS OUTPUT FOR BOX 8.7 **Look-up rate in feeding sheep**

The GLM Procedure

Dependent Variable: luprate

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	20	0.32473357	0.01623668	3.87	<.0001
Error	99	0.41524366	0.00419438		
Corrected Total	119	0.73997723			

R-Square	Coeff Var	Root MSE	luprate Mean
0.438843	30.60150	0.064764	0.211637

Source	DF	Type III SS	Mean Square	F Value	Pr > F
OBSPER	19	0.19191758	0.01010093	2.41	0.0027
SEX	1	0.13281599	0.13281599	31.67	<.0001

Combining data from different experiments

See SAS output for this exercise in the answers for exercises.