

## Analysis of a Repeated-Measures Experiment

An experiment was designed to compare the effect of three drugs (A, B, and C) on the heart rate of women. Five women were randomly assigned to each drug. The heart rate (in beats per minute) of each woman was measured at 0, 5, 10, and 15 minutes after the drug was administered. The data are provided in the table below.

Drug	Woman	Time in minutes			
		0	5	10	15
A	1	72	86	81	77
	2	78	83	88	81
	3	71	82	81	75
	4	72	83	83	69
	5	66	79	77	66
B	6	85	86	83	80
	7	82	86	80	84
	8	71	78	70	75
	9	83	88	79	81
	10	86	85	76	76
C	11	69	73	72	74
	12	66	62	67	73
	13	84	90	88	87
	14	80	81	77	72
	15	72	72	69	70

1. Write down a model or SAS *proc glm* or *proc mixed* code that might be used to analyze the data if we view this as a split-plot experiment.
2. Provide an ANOVA table with SOURCE and DF columns for the model in question 1.
3. Give a formula (in terms of mean squares) for testing each hypothesis concerning fixed factors or interactions between fixed factors.
4. Give a formula (in terms of mean squares) for estimating each variance component in your model. (Expected mean squares are provided on the next page.)

The GLM Procedure  
 Dependent Variable: y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	23	2449.500000	106.500000	12.73	<.0001
Error	36	301.100000	8.363889		
Corrected Total	59	2750.600000			

drug	2	337.600000	168.800000	20.18	<.0001
woman(drug)	12	1498.500000	124.875000	14.93	<.0001
time	3	256.333333	85.444444	10.22	<.0001
drug*time	6	357.066667	59.511111	7.12	<.0001

Source	Type III Expected Mean Square
drug	Var(Error) + 4 Var(woman(drug)) + Q(drug, drug*time)
woman(drug)	Var(Error) + 4 Var(woman(drug))
time	Var(Error) + Q(time, drug*time)
drug*time	Var(Error) + Q(drug*time)

In the classic split-plot model,

- the variance of an observation from any split plot is equal to some constant value that does not depend on the split plot,
- observations from split plots on different whole plots are independent,
- observations from split plots within the same whole plot are correlated, and
- correlation between any two observations from split plots within the same whole plot is equal to some constant positive value that does not depend on the two split plots or their whole plot.

It is often reasonable to analyze a repeated-measures experiment like a split-plot experiment. Sometimes, however, different assumptions about the correlation and/or variance of split-plot observations are more appropriate. When response variables are measured over time (as in the example on the front of this page), correlation between split plots corresponding to adjacent time points may be greater than correlation between split plots that are well separated in time. In the classic split plot experiment, the levels of the split-plot factor are randomly ordered within each whole plot. That is not the case when time is playing the role of the split-plot factor because it is not possible to randomize the levels of time to the split-plots. Time 1 and time 2 measurements, for example, are always right next to each other in time. Thus, while the bulleted assumptions above may be appropriate for a split-plot experiment, they might not be appropriate for a repeated-measures experiment.

The *repeated statement* in *proc mixed* is needed for more general assumptions about the correlation and/or variance of split-plot observations. The code below is an alternative way to fit the classic split-plot model.

```
proc mixed method=type3;
  class drug woman time;
  model y=drug time drug*time;
  repeated / type=cs subject=woman(drug) r rcorr;
```

If *type=cs* is replaced by *type=ar(1)*, the model allows for correlation between to split-plot observations from a single whole-plot experimental unit to increase as the observations become closer together in time. Specifically the correlation between observations from a single whole-plot experimental unit that are *m* time points apart is modeled as  $\rho^m$  for some  $\rho$  between 0 and 1. If the time points are not equally spaced, it may be better to use *type=sp(pow)(time)*, which models the correlation between observations at times  $t_1$  and  $t_2$  from a single whole-plot experimental unit as  $\rho^{|t_1-t_2|}$ . For either *type=cs*, *type=ar(1)*, or *type=sp(pow)(time)*, the variance of split-plot observations is assumed to be constant over time. If *type=un* is used, the model allows for a unique variance for each time point and a unique correlation between any two time points. Many other choices for the correlation structure are possible. We would like to choose the simplest correlation structure that adequately fits the data. An informal way to select an appropriate correlation structure is to use the model with the lowest measure of the *Bayesian Information Criterion (BIC)*. BIC is a statistic that measures how well a model fits the data with a penalty for making a model too complex. See the SAS code and output accompanying this handout.