STAT 8200 — Design of Experiments for Research Workers
Lab 10 – Due: Tuesday, November 12

Example:

Seven different hardwood concentrations are being studied to determine their effect on the paper produced. However, the pilot plant can only produce three runs each day. As days may differ, the analyst uses the balanced incomplete block design that follows.

Hardwood	Days						
Concentration	1	2	3	4	5	6	7
2	114	–	–	–	120	–	117
4	126	120	–	–	–	119	–
6	–	137	117	–	–	–	134
8	141	–	129	149	–	–	–
10	–	145	–	150	143	–	–
12	–	–	120	–	118	123	–
14	–	–	–	136	–	130	127

Obtain the file paper.sas from the course web page, run it and examine the program and its associated output.

This design is a balanced incomplete block design with $a = 7$ treatments (the hardwood concentrations) and $b = 7$ blocks (the days). Each block is of size $k = 3$ because only three runs could be completed per day. Each treatment occurs $r = 3$ times in the design and each pair of treatment occurs in the design $\lambda = 1$ time.

Examine the output from paper.sas. In the first section of paper.sas the classical intrablock analysis is performed. The Type I sums of squares are given for DAY and CONC. These SSs are really $SS_{Day(unadj)}$ and $SS_{Conc(adj)}$ respectively. The Type III SSs are really $SS_{Day(adj)}$ and $SS_{Conc(adj)}$. Only the Type I SSs sum to SS_T, but the Type III SSs are the appropriate ones for inference on both DAYs and CONCentrations. We see that we should reject the null hypothesis that the means are the same for all levels of CONC ($F = 10.42$, $p = .0021$), and, informally, there seems to be mild evidence of differences across days ($F = 3.12$).

Because the levels of CONC are evenly spaced and quantitative, it is of interest here to ask whether strength changes with concentration in a linear fashion. The contrast statements in paper.sas (orthogonal polynomial contrast coefficients obtained from Table D.6 of our text) indicate that the relationship between strength and concentration is at least linear ($F = 10.99$) but that there is significant lack of fit from linearity ($F = 10.31$) meaning that the relationship is not adequately described by a line (the relationship conforms to a higher polynomial). We also see in the contrast results that the relationship is at least quadratic ($F = 32.97$), but there is significant lack-of-fit from a quadratic relationship, too ($F = 4.64$). It is not worth proceeding any further in this manner because our results indicate that there is not a simple (linear or quadratic) relationship

between strength and concentration. From the profile plot in the output we see why this is the case. The relationship seems to be linear from CONC=2 to CONC=10, but then further increasing the concentration of hardwood reduces the strength of paper dramatically.

The second part of paper.sas features a call to PROC MIXED in which days are modeled with random effects. This seems more appropriate here where we would like to generalize to the population of all days on which the treatments could be observed. PROC MIXED gives the combined inter- and intra-block analysis. Notice that the results differ somewhat, but not greatly, from the intrablock analysis given by PROC GLM.

Exercise:

Consider a study in which a tasting panel is convened to score the quality of steaks produced by progeny of 10 different bulls. Since a judge's ability to discriminate tastes diminishes as the number of items to be judged increases, only 5 steaks are presented to each judge at one time. Complete balance for this design requires an integer multiple of 9 ratings for each bull. In the study actually performed, 9 ratings per bull were were obtained. The sample plan, which shows assignment of bulls to judges, appears below.

Judge	Bulls					Judge	Bulls				
(J1)	B1	B2	B3	B4	B5	(J10)	B2	B3	B4	B8	B10
(J2)	B1	B2	B3	B6	B7	(J11)	B2	B3	B5	B9	B10
(J3)	B1	B2	B4	B6	B9	(J12)	B2	B4	B7	B8	B9
(J4)	B1	B2	B5	B7	B8	(J13)	B2	B5	B6	B8	B10
(J5)	B1	B3	B6	B8	B9	(J14)	B2	B6	B7	B9	B10
(J6)	B1	B3	B7	B8	B10	(J15)	B3	B4	B6	B7	B10
(J7)	B1	B4	B5	B6	B10	(J16)	B3	B4	B5	B7	B9
(J8)	B1	B4	B8	B9	B10	(J17)	B3	B5	B6	B8	B9
(J9)	B1	B5	B7	B9	B10	(J18)	B4	B5	B6	B7	B8

The bulls were assigned the labels B1–B10 at random and the judges were assigned the labels J1–J18 at random. Judges rated each of the five steaks for flavor and tenderness on a scale from 0 to 8, with 8 being the best possible score.

- This experiment is an example of a common application of incomplete block designs. In situations where judges are used to assess several treatments, it is typical to block by judge, and only observe a relatively small subset of the treatments for each judge. The reason for this is that it is often difficult for judges to differentiate between more than a few treatments.

STAT 8200 — Lab 10

Name:_____

The data collected from the above design appear below and in file steak.dat on the course web page.

Judge	Scores					Judge	Scores				
(J1)	7	8	8	6	7	(J10)	7	7	5	4	8
(J2)	5	7	6	6	7	(J11)	8	8	7	8	8
(J3)	6	8	5	6	7	(J12)	7	4	7	4	6
(J4)	6	7	5	7	4	(J13)	6	4	4	2	5
(J5)	4	5	4	2	4	(J14)	6	5	7	6	6
(J6)	4	6	6	3	6	(J15)	8	6	7	8	8
(J7)	5	5	6	6	7	(J16)	6	4	5	7	6
(J8)	6	5	4	8	8	(J17)	7	5	5	4	6
(J9)	4	4	6	6	6	(J18)	5	6	7	8	5

1. Identify the treatments and the blocks.

2. What are the values of k, r, a, b and λ?

3. Why does the design require a multiple of 9 ratings per bull? (Hint: consider the computation of λ.)

Write a SAS program to answer the following questions. If judge effects are most appropriate considered to be random, then use the combined inter- and intra-block analysis. Otherwise, do the intrablock analysis using PROC GLM.

4. Are there significant differences between the mean ratings for the 10 bulls? Report the appropriate F statistic and $p-$value.

5. Which bull's progeny yields the best (most tender and flavorful) steak? Which bull's progeny yields the worst steak? What are the estimated population mean ratings for these bulls?

6. Suppose that bulls B1, B5, B6 and B8 are of one variety and bulls B2, B3, B4, B7, B9 and B10 are of a different variety. Form and test a contrast to compare the population means for the two varieties. State your conclusion.

Please hand in pp.3–4, including your answers. Remember to write your name at the top. You may keep pages 1–2 for your notes.