Math 420 HW 2

Fall 2023

Washington University in St. Louis

Due date: Saturday, 10/7/2023

Instruction:

Please type your answers clearly and show your work neatly. You are encouraged to use the Rmarkdown version of this assignment as a template to submit your work. Unless stated otherwise, all programming references in the assignment will be in R. For this assignment, problems roughly covers content from Factorial Designs, Random Block Designs, Variance Components and Fractional Factorial Designs

Problem 1

Kenett and Steinberg (1987) described a two-level factorial experiment conducted by students to study the time required to boil 1 qt of water. Factors were A=flame level (low or high), B=pan size (small or large), C=pan cover (none or glass cover), and D=salt added to water (no or yes).

(a) If the standard deviation in boiling time (tested at the same conditions) was found to be $\sigma^2 = 0.236$ minutes, use the shortcut formula to determine how many experiments you will need to perform in order to have power of 0.95 for detecting effects of size $\Delta = 0.50$ minutes. Would this answer change if you decided to only perform an experiment with 3 of the 4 factors?

(b) Create a list of experiments in random order for performing these experiments.

Problem 2

Lew (2007) presents the data from an experiment to determine whether cultured cells respond to two drugs. The experiment was conducted using a stable cell line plated onto Petri dishes, with each experimental run involving assays of responses in three Petri dishes: one treated with drug 1, one treated with drug 2, and one untreated serving as a control. The data are shown in the table below:

	Control	Drug1	Drug2
Exp1	1147	1169	1009
Exp2	1283	1323	1260
Exp3	1216	1276	1143
Exp4	1046	1240	1099
Exp5	1108	1432	1385
Exp6	1265	1562	1164

(a) Analyze the data as if it came from a completely randomized design using the model $y_{ij} = \mu + \tau_i + \epsilon_{ij}$. Is there a significant difference between the treatment groups?

(b) Analyze the data as an RCB design, where experiment number represents a blocking factor.
(c) Is there any difference in the results you obtain in (a) and (b)? If so explain what may be the cause of the difference in the results and which method would you recommend?

Problem 3

Consider the data in Table 5.20 (p.216 in DAE with R book) from Smith and Beverly (1981) taken from a staggered nested design to investigate the sources of variability in impurities in raw materials received at a plant in trailer loads. Two samples of material were taken from each of nine trailer loads of pellets. Two measurements of impurities were made on the first sample from each trailer but only one measurement for the second sample from each trailer.

(a) Write the model for the data.

(b) Analyze the data and estimate the three variance components using the method of moments.

(c) Analyze the data using REML and check to see if your estimates remain the same.

(d) Make half-normal plots of the square root of the variances pooled to get the mean squares for sample(trailer) and measurement(sample). Does the assumption of homogeneous variances appear reasonable?

(e) Calculate the EBLUPs for the random trailer effect and make a normal plot to check the normality assumption. What is your conclusion?

Problem 4

Reanalyze the data from the golf experiment, presented in the Appendix of Chapter 4 (or dataset `rcb` in the `daewr` R package) using the `lmer` function. Check to see if you get the same P-values and conclusions shown in Section 4.7.

Problem 5

Consider the experiment to study factors that affect the percent removal of arsenic from drinking water using a iron coated sand filter. (a) Modify the R code in Section 6.5.1 to produce the design shown in Table 6.8. (b) Determine the defining relation for the experiments in the first block and show the confounding of main effects with strings of two-factor interactions. (c) Calculate the effects and make a half-normal plot of them using only the data from block 1 in Table 6.8. (d) Determine the defining relation for the combined experiments. (e) Calculate the effects, including the block effect, and make a half-normal plot of them using all the data in Table 6.8. What interactions are confounded with the block effect?