Exam 1

Due Wednesday, July 8th, 2020 at 11:59pm

Instructions:

- Please turn in the final copy of your Exam 1 into your shared dropbox folder: "1 Turned-in Exams".
  - Name your file: "Lastname Firstname Exam1"
  - Please only turn in a single pdf
- Your answers must be written in your own words, extracting what you need from your software output to support your narrative.
- Please do not copy and paste R output.
- Any graphs/figures must be in the section where they help answer a question with an explanation.
- You may attach code and output to your assignment as an Appendix. The appendix may be looked at if something is not clear in your writing.

Problem 1. Typing Experiment

(a) (4 Points) For the typing experiment considered in Lecture #1, use a statistical model to quantify the gains from using randomization (as illustrated in the second sequence in the notes, slide #15) and from using balance in addition to randomization.

(b) (2 Points) Suppose that the following sequence is obtained from using balanced randomization:

AB, AB, AB, BA, BA, BA

Would you use it for the study? If not, what would you do? What aspect of the sequence makes you uneasy? Can you relate it to the possibility that the advantage of the learning effect may diminish over time and express it in more rigorous terms (Hint: The terms in the model should represent the effects you identified as potentially influencing the comparison.)

Problem 2. Pulp Experiment

(8 Points) For the pulp experiment (considered in Lecture #2) obtain 95% simultaneous confidence intervals for the six pairs of treatment differences using the Bonferroni method and the Tukey method. Which gives shorter intevals?
Hint:
Bonferroni 100% confidence interval for $\tau_i - \tau_j$: $\bar{y}_{i.} - \bar{y}_{j.} \pm t_{N-k,\alpha/2k'}\hat{\sigma}\sqrt{(1/n_i) + (1/n_j)}$
Tukey 100% confidence interval for $\tau_j - \tau_i$: $\bar{y}_{j.} - \bar{y}_{i.} \pm (1/\sqrt{2}) * q_{k,N-k,\alpha}\hat{\sigma}\sqrt{(1/n_j) + (1/n_i)}$

Problem 3. Bioactivity

The bioactivity of four different drugs, A, B, C, D for treating a particular illness was compared in a study and the following ANVOA table was given for the data:

Source	Sum of Squares	Degrees of Freedom	Mean Squares
Treatment	64.42	3	21.47
Residual	62.12	26	2.39
Total	126.54	29	

(a) (2 Points) Describe a proper design of the experiment to allow valid inferences to be made from the data.

(b) (2 Points) Use an F test to test at the 0.01 level to test the null hypothesis that the four treatments have the same bioactivity. Compute the p-value of the observed F statistic.

(c) (2 Points) The treatment averages are as follows $\bar{y_A} = 66.10$ (7 samples), $\bar{y_B} = 65.75$ (8 samples), $\bar{y_C} = 62.63$ (9 samples), $\bar{y_D} = 63.85$ (6 samples). Use the Tukey method to perform multiple comparisons of the four treatments at the 0.01 level.

(d) (2 Points) It turns out that A and B are brand-name drugs and C and D are generic drugs. To compare brand-name versus generic drugs, the contrast $\frac{1}{2}(\bar{y_A} + \bar{y_B}) - \frac{1}{2}(\bar{y_C} + \bar{y_D})$ is computed. Obtain the p-value of the computed contrast and test its significance at the 0.01 level. Comment on the difference between brand-name and generic drugs.

Problem 4. Cement Manufacturing

(6 Points) A cement manufacturing company has a packing plant containing a large number of packing machines. To test the consistency of packing performance across machines, three machines (denoted by M/C 1, M/C 2, and M/C 3) were selected from the plant, and 20 bags (each supposed to be of weight 50 lb) were filled in each of the three machines. The data (courtesy of T. Dasgupta) on titanium called *cement.csv* indicate the actual weight (in pounds) of each of the 60 bags. Suppose that interest lies in the comparison of the three chosen machines in the plant. Analyze the data using a one-way layout (with fixed effects). What can you conclude?