STA674: Regression Analysis and Design of Experiments

Assignment #4
Experimental Design

<u>Submission:</u>

You must format your assignments as a pdf. Handwritten assignments will not be accepted.

When are ready to submit your assignment, copy your R (or RStudio work) or SAS code and paste it at the end of your document. *Don't forget to address the "software initiative" portion of this assignment, via comments to help the grader follow your work, re-doing variable names, etc.* Collaboration during the process of solving the problems is not only allowed but encouraged; that said, the submissions are each expected to be an individual effort reflecting the individual's work. Identical submissions or even submissions found to be **"too close to be coincidental" will be flagged and given no credit**.

Problems are worth the indicated points. Turn your completed assignment in, as is customary, via electronic submission on Canvas by the due date on the syllabus.

Homework Questions

1. (10 points) State whether each of the following scenarios represents a comparative experiment or an observational study and for each identify:

    *   the factors,

    *   the treatments,

    *   the experimental units, and

    *   the observational units.

    a. A researcher in pet therapy is interested to know if visits by dogs improve the recovery of patients after surgery. First, a sample of 30 hospitals is selected and a sample of 10 patients undergoing surgery is selected from each hospital. The selected patients within 15 of the hospitals (selected at random) receive visits from therapy dogs after their surgery and the patients at the remaining 15 hospitals do not. Patients are then compared as to the number of days until they are released from hospital.

    b. To study the effects of school funding on the performance of elementary school students, a researcher in education selects 15 classes of grade 5 students in schools across Lexington. She gives the students in each class a standardized test and then compares the average scores for each of the classes.

    c. An agricultural scientist wants to study the effects of different fertilizers on the production of different grains and what factors might affect the potency of the fertilizers. In carefully controlled greenhouses he grows wheat, rice, and spelt under a variety of conditions including 3 types of fertilizer, high or low simulated precipitation

levels, and high or low temperatures. Many plants are grown in each greenhouse and the conditions are the same for all plants in one greenhouse. At the end of the study the total amounts of grain produced by the plants of each type in each greenhouse are compared.

2. (4 points) An experiment is planned to compare three treatments applied to shirts in a test of durable press fabric treatments to produce wrinkle-free fabrics. In the past formaldehyde had been used to produce wrinkle-free fabric, but it was considered an undesirable chemical treatment. This study is to consider three alternative chemicals: (a) PCA (1-2-3 propane tricarbolic acid), (b) BTCA (butane tetracarboxilic acid), and (c) CA (citric acid). Four shirts will be used for each of the treatments. First, the treatments are applied to the shirts, which are then subjected to simulated wear and washing in a simulation machine. The chemical treatments will not contaminate one another if they are all placed in the same washing machine during the test. The machine can hold one to four shirts in a single simulation run. At the end of the simulation run each of the shirts is measured for tear and breaking strength of the fabric and how wrinkle-free they are after being subjected to the simulated wear and washing. The comparisons among the treatments can be affected by (a) the natural variation from shirt to shirt; (b) measurement errors; (c) variation in the application of the durable press treatment; and (d) variation in the run of the simulation of wear and washing by the simulation machine. Following is a brief description of three proposed methods of conducting this simple experiment.

- Method I. The shirts are divided randomly into three groups of four shirts. Each group receives a durable press treatment as one batch and then each batch is processed in one run of the simulation machine. Each run of the simulation machine has four shirts that have received the same treatment. There are three runs of the simulation machine.

- Method II. The shirts are divided randomly into three treatment groups of four shirts each, and the durable press treatments are applied independently to single shirts. The shirts are grouped into four sets of three, one shirt from each durable press treatment in each of the four sets, and each set of three so constructed is used in one run of the simulation machine. There are four runs of the simulation machine.

- Method Ill. The shirts are divided randomly into three groups of four shirts. The durable press treatments are applied independently to single shirts. The simulation of wear and washing is done as in Method I.

    a. Which method do you favor, and why do you favor the method you have chosen?

    b. Briefly, what are the disadvantages of the other two methods?

3. (2 points) An experiment is planned to compare three methods of instruction. Each is tested with a single classroom of 25 students. A different instructor is to be used for each classroom and consequently each instruction method.

    a. Write a short critique of the proposed experiment

    b. How could the experiment be improved?

4. (4 points) An experiment was conducted to test the effects of nitrogen fertilizer on lettuce production. Five rates of ammonium nitrate were applied to four replicate plots in a completely randomized design. The data (Source: Dr. B. Gardner) in the table are the number of heads of lettuce harvested from the plot (these are also in the file on Canvas.)

Treatment (lb N/acre)	Heads of lettuce/plot
0	104, 114, 90, 140
50	134, 130, 144, 174
100	146, 142, 152, 156
150	147, 160, 160, 163
200	131, 148, 154, 163

    a. Write the statistical model for this experiment and explain the components.

    b. Compute the analysis of variance for the data and test the hypothesis of no difference between the treatments using the .05 level of significance. Clearly state your hypotheses, the value of the test statistic, and your conclusion.

    c. Compute the estimated mean response and 95% confidence intervals for each treatment. Provide an interpretation of these values for the treatment of 200 lb/acre.

    d. State the assumptions necessary for the analysis of variance to be valid. Are these assumptions satisfied?