Psychology 105, Homework Assignment

Available 4/29/2021; due 5/6/2021

Unfortunately, for this assignment we cannot work with your Statlab data sets, because there is no suitable grouping variable that would produce a reasonable ANOVA design. We will work instead with completely made up data. You will find the data set at http://faculty.ucmerced.edu/jvevea/classes/105/data/Hwk5.csv. You can get the data set into your own R space either by downloading the file and using read.csv() with the path to where you have saved it, or by this easier command:

```
Hwk5data <- read.csv("http://faculty.ucmerced.edu/jvevea/classes/105/data/Hwk5.csv")
```

You will see that the dataset contains a variable called "Group" (Control, T1, T2, T3), and another called "Score." Imagine that the values of Score represent a Quality of Life index (where high scores are good), and that the data were collected from cancer patients under four different conditions.

Imagine, further, that the conditions are defined by several different treatments designed to improve mood in terminally ill cancer patients. The group designated "Control" represents a business-as-usual condition where patients receive no special mood-related intervention. T1 represents a traditional intervention. You, as a researcher, are considering two additional interventions designed to improve mood, T2 and T3.

In addition to the question of whether there are any differences among the four means, you want to know three things:

1. Does any treatment (i.e., the mean of the three "T" groups) represent a mood improvement over no treatment ("Control")?
2. Do your new treatments ("T2" and "T3") differ from the traditional treatment ("T1")?
3. Do your treatments ("T2" and "T3") differ from one another?

Part One

1. Create a dummy coding system (three 0,1 variables) that will allow you to do the ANOVA comparing the four groups. Run the regression, and examine the *F* statistic that compares the four groups. Is there evidence that there are differences among the four populations?
2. Let *R* create its own coding system [summary(lm(Score~Group))] and confirm that your result is the same.

Part Two

ANOVA requires two assumptions that you can check by examining data: equal variability in the populations, and normality within each population. Use methods demonstrated in class to assess those assumptions. Are they met? Do you have concerns?

Part Three

Use Tukey's HSD procedure to examine pairwise comparisons between group means. Which means differ significantly at the .05 level? For which comparisons is there no evidence of a difference?

Part Four

1. The three questions described above in bullet points 1, 2, and 3 imply a set of orthogonal contrasts. Using contrast-coefficient notation such as we have demonstrated in class, specify those contrasts. This will get you started: The first question can be summarized as
   3 -1 -1 -1.
2. Confirm that your set of contrasts is mutually orthogonal. (Show your work.)
3. Create a contrast coding system that represents those contrasts. Run the regression that uses the contrast codes, and
   a. Confirm that the overall ANOVA result is the same as before.
   b. Interpret each contrast. Do the three treatments differ from control? Are the other contrasts that you specified in 4.1 significant?
4. For extra credit, attach a contrast matrix to the "Group" variable, and duplicate your results from part 4.3 letting *R* do the ANOVA with a summary(lm(Score~Group)) command.