MSDS UNIT 3 HW 3

Conceptual questions.

1. State the definition of factor, levels, treatment, and main effects.
2. What are the necessary assumptions of a TWO WAY ANOVA analysis in order to perform hypothesis testing?
3. Suppose we have a two factor ANOVA model where each factor has only two levels (Let be the factors and call the levels ZERO and ONE. Treating the factors as dummy variables, write out a multiple linear regression model for the additive and nonadditive models. What does the intercept represent in your model? What regression coefficient should be used to test for each of the single factors in the additive model? What regression coefficient would be used to test if an interaction exists in the nonadditive model?
4. TRUE or FALSE? Suppose we run a two way ANOVA model using two factors A (High Mediam Low) and B (assembly line 1, 2, and 3) and there is plenty of data for each factor level combination. All 3 type-III sum of squares F-tests are significant. It is valid to directly compare High versus Low via a t-test regardless of assembly line.
5. TRUE or FALSE? Since TWO WAY ANOVA is a special case of multiple regression, influential diagnostics such as leverage statstics should be examined in a 2Way ANOVA analysis.
6. TRUE or FALSE? Contrasts are specific hypothesis tests specified by the analyst to determine differences between specific factor levels or treatments in an ANOVA model.
7. The following boxplots summarize data for each factor combinations of Factor 1 (A,B) and Factor 2 (a,b,c) from a two way anova analysis. Using the median values of each boxplot, make a decision for each graph of whether or not an interaction potentially exists

  

Analysis Question

1. Back in Unit 1 we considered a study in which 4 different fertilizers were tested for their yield (in mm of growth) on a local grass: Red Fescue (Just for fun .. this is a real mountain grass! http://fescue.com/info/creepingred.html#.WIu6rbYrKu4). To conduct their study they had enough money to run three replicates of each fertilizer. They knew the red fescue was a mountain grass so they went out to the mountains and carefully fertilized plots of land as you see in the diagram below. The data is contained in the ***growth3*** data set.



Given the information you have available, run a simple ONE WAY ANOVA to test for the effects (if any) of the fertilizers. **From a previous study, we know that the yields from each fertilizer are normally distributed with equal variances. For now assume that independence is not a concern here. You may assume the assumptions are met for all questions in this homework.**

Obtain the following Deliverables: 1. ANOVA table. 2. Means Plot (Interaction Plot from SAS) 2. Conclusion for the appropriate ANOVA test. 3. Confidence intervals with multiple comparison corrections for **SIGNIFICANT** differences (between Fertilizers). 4. SAS Code: proc glm or Proc mixed code. Your answer should fit in the given box.

|  |
| --- |
| Place output here |

1. Now assume that you get to thinking about it and realize that we may really be looking at three different environments here: Sunny, Wetlands and Mostly Shady. This data has been recorded in the ***growth4*** data set. Conduct a similar analysis as in 1 but now account for the environment variable (ENV).



Deliverables: 1. ANOVA table. 2. Means Plot (Interaction Plot from SAS) 2. Conclusion for the ANOVA (only step 6 is required.) 3. The researchers were specifically interested in (a) Inference on the main effect of the fertilizer. (b) If Fertilizer 4 performed significantly better in one environment than another. (c) Which fertilizer will perform best in the wetlands. Answer each QOI (question of interest) with a confidence interval and a 1 sentence conclusion. Your answer should fit in the space below (on this page.) 4. SAS Code: Proc glm or proc mixed statement. Answer should fit in the given box.

|  |
| --- |
| Continued on next page… |

1. Consider the model you used in problem 2. Inspect the Means Plot (Interaction Plot). Does it look like there will be a significant interaction? Explain by interpreting what an interaction is and then comparing that to what you see in the plot.

|  |
| --- |
|  |

1. Fit the full model that includes the interaction term? What do you notice? Why? Discuss.

|  |
| --- |
|  |