STAA 553: HW5

YOUR NAME HERE

See Canvas Calendar for due date.
48 points total, 2 points per problem unless otherwise noted.
Add or delete code chunks as needed.
Content for Q1-Q15 is from section 07.
Content for Q16-Q20 is from section 09.

Biomass (Q1 - Q15)

A greenhouse study was done to examine the effect of three herbicides (A, B or C) and two water regimes (Low or High) for two plant types (Grass or Forb). The response variable is biomass. There are three reps per treatment combination for a total of 36 observations. Each observation was a potted plant. The 36 pots were randomly assigned without restriction to locations in the greenhouse. The data is available from Canvas as "Biomass.csv".

Important notes:

- Remember to run str() and then define things as.factor where needed.
- Change contrasts options to get meaningful Type 3 tests (using Anova): options(contrasts=c("contr.sum", "contr.poly"))
- Diagnostic plots are considered for several questions. You do NOT need to include these plots in your assignment. But you do need to discuss your findings.

Q1

Fit the three-way model with all interactions and show the Type 3 ANOVA table. You should find evidence of a 3 way interaction.

Q2 (4 pts)

Use residual diagnostic plots to discuss whether model assumptions are satisfied. You do NOT need to include the plots in your assignment. But for full credit it should be clear which plot is being used to check which assumption.

Response

Q3

Create a summary graph (of emmeans) using code similar to what is provided.

Q4

Regardless of any concerns you may have about assumptions, use emmeans to calculate pairwise comparisons of Water (High vs Low) *for each level of Herb and Type*. Use code similar to what is provided.

Biomass: Forb Only (Q4 - Q9)

Now fit a two-way model (including interaction) for **Forb only**.

Q5

Show the Type 3 ANOVA table.

Q6

Consider the diagnostics plots and (briefly) discuss whether model assumptions are (better) satisfied.

Response

Q7

Use emmeans to calculate pairwise comparisons of Water (High vs Low) *for each level of Herb*.

Q8

Use emmeans to calculate the comparison of Water (High vs Low) *averaging over the levels of Herb*.

Q9

Considering the SE for the comparisons from Q7 (interaction comparisons) and Q8 (main effect comparison), which has higher power? Briefly discuss.

Response

Biomass: Grass Only (Q10 - Q13)

Now fit a two-way model (including interaction) for **Grass only**.

Q10

Show the Type 3 ANOVA table.

Q11

Consider the diagnostics plots and (briefly) discuss whether model assumptions are (better) satisfied.

Q12

Use emmeans to calculate pairwise comparisons of Water (High vs Low) *for each level of Herb.*

Q13

Would it be appropriate to calculate the comparison of Water (High vs Low) *averaging over the levels of Herb*? Briefly discuss.

Response

Biomass: Compare Models (Q14 - Q15)

Now we compare the three-way model to the separate two-way models.

Q14

Give (at least) *one benefit* of splitting the analysis by Type (running separate 2way ANOVAs for Grass and Forb). Your answer should be *based on specific output.*

Response

Q15

Give (at least) *one weakness* of splitting the analysis by Type as compared to the full 3way ANOVA model.

Response

Breakfast (Q16 - Q20)

We return to the breakfast data from HW3. A study was done to examine whether breakfast choice was associated with cholesterol levels in children. A total of n=35 fourth and fifth graders were included in the study. Based on survey response, children were identified as one of (g = 4) four (BKFST) breakfast types: Cereal_F (cereal with fiber), Cereal_O (other cereal), Other_Br (other breakfast) or Skip (no breakfast). Note that the sample sizes are unequal. The height and weight of each child was used to determine their Body Mass Index (BMI). BMI is not of direct research interest, but will be considered as a covariate in some models. The response variable is plasma total cholesterol (TC). The data is available from Canvas as Breakfast.csv.

Q16

Construct a scatterplot of TC (Y) vs BMI (X) for all BKFST groups on the same plot. Overlay a separate regression line for each BKFST group.

Q17 (0 pts)

Calculate a table of summary statistics including sample size, mean, sd by BKFST group. (0 pts, because we already did this for HW3).

Q18

Fit a one-way model (using BKFST as the predictor).

Q18A (0 pts)

Show the ANOVA table. (0 pts, because we already did this for HW3).

Q18B

Calculate Tukey adjusted pairwise comparisons for BKFST.

Q19

Now fit a model including both BKFST and BMI (but no interaction).

Q19A

Show the Type 3 ANOVA table.

Q19B

Show the emmeans for BKFST.

Q19C

Calculate Tukey adjusted pairwise comparisons for BKFST.

Q19D (4 pts)

Briefly summarize your findings from the previous question (using alpha = 0.05).

Response

Q20

Compare the results from the one-way model (Q18) vs the ANCOVA model (Q19). Briefly explain why we were able to detect differences using the ANCOVA model, when we did not detect differences using the one-way model. Your answer should be based on *specific output*. Hint: You may want to calculate MSResid.

Response

Appendix

```
#Retain this code chunk!!!
library(knitr)
knitr::opts_chunk$set(echo = FALSE)
knitr::opts_chunk$set(message = FALSE)
knitr::opts_chunk$set(warning = FALSE)
#Q1

#Q2

#Q3
library(emmeans)
#emmip(BM_3way, Water ~ Herb | Type, CIs = TRUE)
#Q4
#emout1 <- emmeans(BM_3way, ~ Water|Herb*Type)
#pairs(emout1)
#Q5

#Q6
```

```
#Q7

#Q8

#Q10

#Q11

#Q12

#Q16

#Q17

#Q18A

#Q18B

#Q19A

#Q19B

#Q19C
```