

STAT 512 Section 2 Assignment #7

See Canvas calendar for due date.

Reading: Ott & Longnecker section 14.3 (Factorial Trt Structure) and chapter 15 (Blocked Designs)
36 points Total, 2 points per question unless otherwise noted.

Important reminders for this assignment:

- 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"))`

1. In this group of questions we use the “PCB.csv” data available from Canvas. A researcher is interested in comparing PCB concentrations by sex (m, f) and species (1, 2). This corresponds to four groups (sp1f, sp1m, sp2f, sp2m). Note that depending on the analysis you will use group OR sex and species in the analysis but not all three!

- A. Create a table of summary statistics including sample size, mean and standard deviation for each sex, species combination. Then construct an interaction plot. For consistency, please put sex on the X axis. Include both the summary table and interaction plot in your assignment. **(4 pts)**

Questions B and C (PCB): Fit a **one-way** ANOVA model to the data using **group** as the predictor.

- B. Construct the Type3 ANOVA table.
- C. Use `emmeans` to calculate (Tukey adjusted) pairwise comparisons for all four groups.

Questions D through G (PCB): Fit a **two-way** ANOVA model to the data using **sex and species** as predictors. Be sure to include the **interaction**.

- D. Construct the Type3 ANOVA table.
- E. Use `emmeans(, pairwise ~ sex:species)` to calculate (Tukey adjusted) pairwise comparisons for all four groups.
- F. Use `emmeans(, pairwise ~ species)` to calculate the pairwise comparison corresponding to the main effect of species. Note that the p-value from this comparison should match the F-test corresponding to species from the ANOVA table from part (D).
- G. Consider the output from the two previous questions. From part (E), for the “m,1 - m,2” comparison you should have found an estimate = 2.67, p-value = 0.0857. From part (F), for the “1 - 2” comparison you should have found an estimate = 2.01, p-value = 0.0154. Briefly explain why we find a smaller p-value for the “1 - 2” comparison even though the estimated difference is smaller.

Note: The two models from above are equivalent. This can be seen by comparing the ANOVA tables (questions 2 and 4) and the pairwise comparisons (questions 3, 5). Either analysis approach is acceptable. However, one benefit of the two-way analysis is that since the interaction is not significant, we can easily discuss main effects.

2. For this problem use the data described in Ott & Longnecker problem 15.14 (p 907 in the 7th Edition). Use the **Ex15-14.csv** data file. (It is worth reading for some background on the research, but the wording actually isn't great in the text.)

Roadway	Sodium Chloride		Calcium Chloride		Sand	
	Low	High	Low	High	Low	High
1	37	49	43	47	27	33
2	39	50	42	48	27	31
3	48	52	47	50	36	37
4	44	57	45	54	34	37
5	54	68	56	63	45	44

- Describe the blocking and treatment structure.
- Create a table of summary statistics including sample size, mean and standard deviation for each Treatment*Concentration combination. Then construct an interaction plot. For consistency, please put Treatment on the X axis. Include both the summary table and interaction plot in your assignment. **(4 pts)**
- Considering your answer to part A, fit an appropriate model and include the Type 3 ANOVA table in your assignment. (You might considered checking diagnostics plot as a matter of habit even though they are not requested here). **(4 pts)**
- Discuss the value of the blocking for this data. Justify your response with appropriate test-statistic(s) and p-value(s).
- Compare mean response for **High vs Low Concentration separately for each Treatment**. This can be done using emmeans. Include the “emmeans contrasts” output in your assignment, but **also (briefly) summarize your findings**. **(4 pts)**
- Compare mean responses between the **3 Treatments for Concentration = Low**. (default Tukey adjustment is fine). Include these “emmeans contrasts” output in your assignment, but **also (briefly) summarize your findings**. **(4 pts)**