

**WINTER 2019/20 TERM 1 EPSE 592: ASSIGNMENT 2****Due: Week 7 in class: Feb. 27**

- Please make sure you write your answers to these questions in your own words. Even if you work with a group to formulate your responses, do not just copy someone else's sentences/words.
- There is no need to record more than 3 decimal places for any of these problems.
- All problem data are available online in .csv format.

**Question 1: (Interpretation of p-values: the traditional threshold approach versus the orders of magnitude rule)** In this problem, you will explore some of the dangers of using an arbitrary significance threshold to declare results “significant” or not.

Suppose you are given the following data on 8 patients that volunteered to try a new kind of smoking cessation therapy. The Time 0 data record how many cigarettes per week each patient would smoke on average before administration of the treatment. Each patient then started the new cessation therapy at the same time and recorded how many cigarettes they smoked in the first week (Time 1), in the second week (Time 2), and in the third week (Time 3) of therapy.

	Time 0	Time 1	Time 2	Time 3
Patient 1	70	73	76	66
Patient 2	120	113	105	80
Patient 3	85	70	55	44
Patient 4	50	55	59	49
Patient 5	40	33	26	18
Patient 6	100	95	90	71
Patient 7	95	88	81	53
Patient 8	115	107	99	57

Table 1: Data for smoking cessation study.

- Perform a paired samples t-test between baseline and Time 1. Report the p-value, mean difference between the two time points (i.e. treatment effect), and the 95% confidence interval for the mean difference. Interpret this output using the common, traditional rule that “p-values below 0.05 indicate the presence of a significant effect.”
- Now perform a paired samples t-test between Time 1 and Time 2. Again, report the p-value, mean difference, and 95% confidence interval for the mean difference. Interpret this output using the same traditional threshold rule as before.
- What are the actual differences between the treatment effects from baseline to Time 1 and from Time 1 to Time 2? For each patient, compute their change in weekly cigarette intake from baseline to Time 1, and from Time 1 to Time 2. Do you find evidence that the smoking cessation therapy has a different effect during the first week as it does during the second week?
- Perform a paired samples t-test between Time 2 and Time 3. Report the p-value, mean difference, and 95% confidence interval for the mean difference. Interpret using the same traditional threshold rule as before.

- (e) Review your quantitative output in parts (a), (b), and (d). Using the *orders of magnitude rule* (i.e. is their weak, moderate, strong evidence for a non-zero effect, given the standard 0.05 Type I error rate? See Class 4 notes.), interpret the effects of the smoking cessation therapy over the full three weeks of treatment.

**Question 2: (Practice with the basics of ANOVA)** Suppose we are testing the effects of four different diet regiments on hemoglobin levels in adult patients with iron deficiency anemia. 40 patients are selected for our study, half male and half female. All patients have hemoglobin levels that fall about 1 gram per deciliter of blood below the healthy range for their respective sex (for women, a healthy range is about 12.0 to 15.5 g/dL; for men, a healthy range is about 13.5 to 17.5 g/dL). We randomly assign 5 men and 5 women to one of four diets, and after 6 months, we record their new hemoglobin levels. The first few data points are recorded below (full dataset is on the website).

	(g) Hemoglobin / (dL) blood	Sex	Diet
Patient 1	12.0	F	1
Patient 2	12.6	F	2
Patient 3	12.2	M	1
Patient 4	12.6	M	2

Table 2: Partial data for anemia study.

- (a) Graph the data by sex and by diet simultaneously (a collection of boxplots might be a good choice). Looking at the graphic, does sex appear to have any (mean) effect on the response variable? What about diet? Do you see any signs of a possible interaction between the sex and diet variables?
- (b) Perform a formal (full) two-way ANOVA to test the average effect of sex, diet, and their interaction on hemoglobin levels. Which of these three explains the most variation in the response? Which explains the least?
- (c) Now perform some model diagnostics. As in class, use the following code in the Rj editor of Jamovi to create a plot of the residuals (observed errors) vs. the fitted values (estimated treatment effects) and a qq-plot (see full instructions and code on webpage):

```
mod1 <- aov(formula=data$Hemoglobin~data$Sex*data$Diet)
plot(mod1)
```

[Note: this code spits out 4 different diagnostic plots, but you only need to consider the first two.] Assess the homoskedasticity and normality assumptions; i.e. how reasonable is the assumption that the errors in each treatment group are distributed as  $N(0, \sigma^2)$ , for some fixed  $\sigma^2$ ?

- (d) If you are confident that your modelling assumptions are reasonably satisfied, now it's time to dig deeper into the analysis.
- i. Look at the pairwise post hoc comparisons of the marginal diet categories. Which pairwise comparisons seem to indicate a substantial difference in hemoglobin levels between diet regimens?
  - ii. Look at the pairwise post hoc comparisons for the interaction effect. Realize that many of these comparisons are not very meaningful (e.g. comparing males in diet 1 to females in diet 4); we are interested in comparisons of adjacent categories, as these are easiest to interpret (e.g. comparing males in diet 1 to males in diet 2, or comparing males in diet 3 to females in diet 3). Which of these adjacent pairwise comparisons appear significant?

- iii. Make a pair of interaction plots to better visualize how the sex and diet variables interact to explain variation in the hemoglobin response. In Jamovi, open the “Estimated Marginal Means” tab in your ANOVA analysis window. Then drag the two explanatory variables over to the right (inside the “Term 1” box). To make the other interaction plot, simply change the order of these two variables by dragging one above or below the other inside this same box. On these plots, indicate where the significant pairwise comparisons are.