Math 326 – Homework 09 (11.10 – 11.12 and 13.2-3.5)
Due (via upload to Canvas) Friday, April 29, 2022 at 11:59 PM

1. The results that follow were obtained from an analysis of data obtained in a study to assess the relationship between percent increase in yield (Y) and base saturation (x_1, pounds/acre), phosphate saturation (x_2, BEC%), and soil pH (x_3). Fifteen responses were analyzed in the study The least-squares equation and other useful information follow.

$$\hat{y} = 38.83 - 0.0092x_1 - 0.92x_2 + 11.56x_3, \quad S_{yy} = 10965.46, \quad \text{SSE} = 1107.01,$$

$$10000(\mathbf{X}^t\mathbf{X})^{-1} = \begin{bmatrix} 151401.8 & 2.6 & 100.5 & -28082.9 \\ 2.6 & 1.0 & 0.0 & 0.4 \\ 100.5 & 0.0 & 8.1 & 5.2 \\ -28082.9 & 0.4 & 5.2 & 6038.2 \end{bmatrix}$$

   (a) Assuming all the x_i's are independent, is there sufficient evidence that $\beta_2 < 0$? What can you conclude about the usefulness of the x_2 term in the model?

   (b) Give a 95% confidence interval for the mean percent increase in yield if $x_1 = 914$, $x_2 = 65$, and $x_3 = 6$.

2. Data is collected from three populations (A, B, C) which have normal distributions with a common variance. The data is as follows:

A	1	3	2
B	6	4	
C	6	8	4

   (a) This data set is small enough to do by hand. Determine $\hat{\mu}_i$'s, $\hat{\mu}_0$, SSA and SSW.

   (b) Fill out the ANOVA table for this experiment.

   (c) Is there sufficient evidence at level $\alpha = 0.10$ that at least one of the means is different from the others? Briefly explain.

   (d) Give bounds for the p-value in this experiment.

3. Students at a large high-school preparing for the ACT exam have the option of attending an ACT tutorial session. Some students chose to do this and some do not. School administrators want to determine if the tutorial helps. The ACT scores of students at this school are summarized below:

attended tutorial:	$\bar{y} = 26$	$S = 3.5$	$n = 50$
did not attend:	$\bar{y} = 25$	$S = 3$	$n = 80$

   (a) Determine $\hat{\mu}_0$ and then compute SSA (or SST, using the book's notation) and SSW (or SSE in the book).

(b) Fill out an ANOVA table and test the administrators' conjecture at the $\alpha = 0.05$ significance level. Fully justify your conclusion.

(c) Determine a 95% confidence interval for the difference in means.

(d) Explain how your answer in (b) supports your conclusion in (a).

4. An experiment was conducted to compare the starch content of tomato plants grown in sandy soil supplemented by one of four different nutrient packages, A, B, C, and D.

A	45	47	67	61	56	72
B	39	39	43	46	59	27
C	60	42	55	44	30	18
D	56	61	60	47	69	42

(a) Using the technology of your choice, construct an ANOVA table.

(b) Report the p-value and interpret the p-value in the context of the problem

Some theoretical ANOVA odds-n-ends and a little review.

5. In class we stated the MLE for $\hat{\mu}_i$ and S_a^2 under the alternative hypothesis that all the means are not equal. Using the likelihood function $L(\text{all } Y_{ij}|\hat{\mu}_1, \ldots, \hat{\mu}_k, S_a^2)$ that we derived in class, show that the MLE is as stated in class. (Hint: Recall that when using this method to determine $\hat{\theta}$, we differentiate with respect to the whole of the estimator, $d/d\hat{\theta}$.)

6. In class we stated that if $X \sim \chi^2(n) = \Gamma(\frac{n}{2}, 2)$, that the expected value of $E\left[\frac{1}{X}\right] = \frac{1}{n-2}$. We seek to prove this.

(a) Using the definition of the gamma function (the *function*, not distribution), write the following integral as a multiple of a gamma function:

$$\int_0^\infty x^{\frac{n}{2}-2} e^{-x/2}\, dx.$$

(b) Via the definition, compute the expected value of $E[1/X]$.

(c) Let $X_1 \sim \chi^2(m)$ and $X_2 \sim \chi^2(n)$ be independent random variables. Define the random variable $F = \dfrac{X_1/m}{X_2/n}$. Use the previous results to compute the mean of F. (Note $F \sim F(m, n)$.)