One/Two Way ANOVA

Due: Monday 11/15 (11.00PM)
Submission: On Gradescope

Part I: Practice Questions

1. A recent study claims that using music in a class enhances the concentration and consequently helps students absorb more information. To test this claim, three different groups of ten randomly selected students (all of the same age) from three different classrooms were selected. Each classroom was provided with a different environment for students to study. *Classroom A* had constant music being played in the background, *Classroom B* had variable music being played and *Classroom C* was a regular class with no music playing. After one month, the students were given a test and their scores were collected (out of 10). The test scores were:

		Test Scores j										Mean
Classroom i	A: constant sound	7	9	5	8	6	8	6	10	7	4	7
	B: variable sound	4	3	6	2	7	5	5	4	1	3	4
	C: no sound	6	1	3	5	3	4	6	5	7	3	4.3

Using the information above, answer the following questions:

(a) Estimate the mean scores of the students in each type of classroom.

(b) Draw an illustration of the ANOVA model.

(c) Write down the ANOVA model for this problem, including the assumptions. Explain your notation.

(d) Fill in the degrees of freedom, MS and F-value in the ANOVA Table:

Source of Variation	*df*	*SS*	*MS*	*F*	*p*-value
Between Groups		54.6			0.0017
Within Groups		90.1			
Total		144.7.56			

(e) Test whether the classroom environment affects students concentration. Use $\alpha = 5\%$. State the null/alternative hypotheses, decision rule and conclusion in the context of the problem.

(f) Construct 90% family confidence intervals for all pairwise comparisons of classroom environments. State your conclusions.

(g) Estimate the following contrast with a 95% confidence interval:

$$L = 2\mu_1 - \mu_2 - \mu_3$$

2. Consider the ANOVA model and the difference estimator $D = \mu_i - \mu_j$. Show that $E(\hat{D}) = \mu_i - \mu_j$ and that its estimated variance is

$$s^2_{\hat{D}} = MSE\left(\frac{1}{n_i} + \frac{1}{n_{i'}}\right).$$

3. An experiment was conducted to determine the effects of four different pesticides on the yield of fruit from three different varieties of a citrus tree. Eight trees of each variety were randomly selected from an orchard. The four pesticides were randomly assigned to two trees of each variety and applications were made according to recommended levels. Yields of fruit (in bushels) were obtained after the test period.

   (a) Write down the factor effects model that corresponds to this experiment.

   (b) Prepare an estimated interaction plot. What is your conclusion about the presence of interactions.

   (c) Fit an ANOVA model with *yield* as the response, and test whether the interaction term is statistically significant. Use $\alpha = 0.05$. State the alternatives, decision rule and conclusion.

   (d) Test whether the main effects are statistically significant. Use $\alpha = 0.05$. State the alternatives, decision rule and conclusion.

   (e) Estimate the mean difference in yields of fruit between Variety 1 and Variety 2 with a 95% confidence interval.

   (f) Estimate the following contrast with a 95% confidence interval

$$L = \frac{\mu_{A,\cdot} + \mu_{B,\cdot}}{2} - \frac{\mu_{C,\cdot} + \mu_{D,\cdot}}{2}$$

Part II: Homework Questions – to be submitted

1. A manufacturer of television sets is interested in the effect of tube conductivity of **four** different types of coating color picture tubes. An experiment is conducted and the following conductivity data are obtained:

Coating Type	Conductivity			
1	143	141	150	146
2	152	149	137	143
3	134	136	132	127
4	129	127	132	129

   (a) Is there a difference in conductivity due to coating type? Use $\alpha = 0.05$.

   (b) Compute a 95% confidence interval for the mean of coating type 4.

   (c) Compute a 99% confidence interval for the for the mean difference between coating types 1 and 4.

   (d) Test all pairwise differences in means with a family confidence coefficient 90% (you can choose Bonferroni, Scheffe or Tukey). Based on the results, which coating type produces the highest conductivity?

   (e) Assuming that coating type 4 is currently in use, what are your recommendations to the manufacturer if they wish to minimize conductivity?

2. Consider the *butterfat* data set in the *Faraway* library. This data set contains information about the percent of butter fat (more is better) in the milk taken from 100 cows. In the study, there are 5 different breeds of cows and 2 different ages. We are interested in assessing if *Age* and *Breed* affect the butterfat content.

(a) Write down the factor effects model that corresponds to this problem.

(b) Prepare an estimated interaction plot. What is your conclusion about the presence of interactions.

(c) Fit an ANOVA model with *log(butterfat)* are the response, and test whether the interaction term is statistically significant. Use $\alpha = 0.05$. State the alternatives, decision rule and conclusion.

(d) Test whether the main effects are statistically significant. Use $\alpha = 0.05$. State the alternatives, decision rule and conclusion.

(e) Estimate the mean difference in butterfat content between Mature and 2year cows with a 95% confidence interval.

(f) Estimate the following contrast with a 95% confidence interval

$$L = \frac{\mu_{Ayrshire,\cdot} + \mu_{Canadian,\cdot}}{2} - \frac{\mu_{Guernsey,\cdot} + \mu_{Jersey,\cdot}}{2}$$