Homework II (2021)

Please solve the following problems and then submit the pdf copy of them.

1. The data in the table below show the measurements of hemoglobin (grams per 100 ml) in the blood of brown trout. The trout were placed at random in four different troughs (tanks). The fish food added to the troughs contained, respectively, 0, 5, 10, and 15 grams of sulfamerazine per 100 lbs. of fish. The measurements were made on ten randomly selected fish from each trough after 35 days.

Sulfamerazine	Hemoglobin									
0	6.7	7.8	5.5	8.4	7.0	7.8	8.6	7.4	5.8	7.0
5	9.9	8.4	10.4	9.3	10.7	11.9	7.1	6.4	8.6	10.6
10	10.4	8.1	10.6	8.7	10.7	9.1	8.8	8.1	7.8	8.0
15	9.3	9.3	7.2	7.8	9.3	10.2	8.7	8.6	9.3	7.2

Do parts (c)-(f) using hand calculations. Use the computer for part (g).

   (a) Plot the data (hemoglobin versus amount of sulfamerazine) and comment on the results. Does it look like the mean responses differs across the four treatment groups?

   (b) Write down a suitable model for the analysis of these data. Be sure to define all of the terms in your model and to state the basic assumptions of the model.

   (c) Estimate the treatment mean associated with 0 grams of sulfamerazine and form a 95% confidence interval for this quantity.

   (d) Conduct the ANOVA table and test the hypothesis of equal means across the four levels of sulfamerazine. Include columns for source of variability, degree of freedom, sums of squares, mean squares, $F-$test values, and $p-$values in your ANOVA table.

   (e) Define a contrast to compare no sulfamerazine with "some" sulfamerazine. Test this hypothesis at level $\alpha = .05$ and state the conclusion of your test.

   (f) Test for difference between all pairs of means using both the Bonferroni method and Tukey's HSD method. State your conclusions.

   (g) Write and run a SAS program to perform (c)-(f). Confirm your results from the hand calculations. Attach the relevant SAS output and a print out of the program.

2. A consumer testing agency obtains four cars from each of six makes: Ford, Chevrolet, Nissan, Lincoln, Cadillac, and Mercedes. Makes 3 and 6 are imported while the others are domestic; makes 4, 5, and 6 are expensive while 1, 2, and 3 are less expensive; 1 and 4 are Ford products, while 2 and 5 are GM products. We wish to compare the six makes on their oil use per 100,000 miles driven. The mean responses

by make of car were 4.6, 4.3, 4.4, 4.7, 4.8, and 6.2, and the sum of squares for error was 2.25.

(a) Compute the Analysis of Variance table for this experiment. What would you conclude?

(b) Design the three contrasts which are mentioned above (imported vs. domestic; expensive vs. less expensive; Ford vs. GM products). For each contrast, compute a 95% confidence interval and reach your conlcusion.

3. For more general one-way model for unbalanced experiment:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij},$$

where $\sum_{i=1}^{a} n_i \tau_i = 0$ for $i = 1, \ldots, a$; $j = 1, \ldots, n_i$. Then the estimates for parameters are given by:

$$\hat{\mu} = \bar{y}_{\cdot\cdot}, \quad \hat{\tau}_i = (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}), \quad \hat{\epsilon}_{ij} = y_{ij} - \bar{y}_{i\cdot}.$$

Prove the following results:

(a) $\sum_{i=1}^{a} n_i \hat{\tau}_i = 0$.

(b) $\sum_{j=1}^{n_i} \hat{\epsilon}_{ij} = 0$, for all i.

4. A researcher studied the effects of three experimental diets with varying fat contents on the total lipid (fat) level in plasma. Total lipid level is a widely used predictor of coronary heart disease. Fifteen male subjects who were within 20 percent of their ideal body weight were grouped into five groups according to age. within each group, the three experimental diets were randomly assigned to the three subjects. Data on reduction in lipid level (in grams per liter) after the subjects were on the diet for a fixed period of time follow.

		Fat Content of Diet		
Block		$j = 1$	$j = 2$	$j = 3$
i		Extremely Low	Fairly Low	Moderately Low
1	Ages 15-24	.73	.67	.15
2	Ages 25-34	.86	.75	.21
3	Ages 35-44	.94	.81	.26
4	Ages 45-54	1.40	1.32	.75
5	Ages 55-64	1.62	1.41	.78

(a) Identify the experimental design used here. Be very specific, and make sure that you identify the treatment and design structures, and which factors are treatment factors and which are blocking factors.

(b) Conduct the ANOVA table and test a proper hypothesis.

(c) Obtain the residuals for the model in (a) and plot them against the fitted values. Also prepare a normal probability plot of the residuals. What are your findings?

(d) Now we assume the randomized complete block design. Conduct the Tukey test for non-additivity of block and treatment effects.