Stat 581 Exam 1 December 2, 2021

1. An experiment is conducted to study the effect of fitness level on ego strength. Random samples of college faculty members are selected from each fitness level, and an ego score is observed for each member in the sample. Higher values indicate greater ego. The data is provided as an attachment.

(a) State the hypotheses of interest. Provide an interpretation, stated in the context of the problem.
(b) Compute the t_o statistic and the p-value. Provide an interpretation, stated in the context of the problem.
(c) Create a Boxplot as a graphical display of the data. Is it true that all high fitness ego scores exceed all low fitness ego scores? In what sense can the experiment find that high fitness faculty members have greater egos than low fitness faculty members?

(d) Compute a 95% confidence interval for $\delta = \mu_1 - \mu_2$. Provide an interpretation, stated in the context of the problem.
(e) Explain how a confidence interval provides a complementary result to a hypothesis test.
(f) Explain how a confidence interval can be used in testing $H_o : \mu_1 = \mu_2$.

2. A completely randomized design is used to investigate the effect of drug dosage on the activity level of lab rats. Each dose level is applied to $n = 4$ rats, and an activity score is observed for each rat in the sample. Higher values indicate greater activity. The data is provided as an attachment.

(a) State the statistical hypotheses of interest. Briefly explain how the form of the alternative hypothesis requires a need for further investigation.

(b) Compute the F_o statistic and the p-value. Provide an interpretation, stated in the context of the problem. Create a Boxplot as a graphical display of the data.

(c) Compute and display 95% confidence intervals for all pairwise comparisons. Explain how computing multiple intervals impacts the probability of committing an error.

(d) Perform pairwise comparisons using the Fisher LSD method, and the Tukey method. Provide grouping information for each method. Comment on the seemingly contradictory nature of a pairwise comparisons analysis.

(e) Describe the defining characteristics for each of the above pairwise comparison methods.

(f) Compute the margin of error and comparison-wise error rate for the Tukey method in this problem.

(g) Compute the margin of error and family-wise error rate for the Fisher LSD method in this problem.

3. A factorial experiment is used to investigate the effect of pressure, temperature, and time on the yield from a chemical reaction. Two levels (low, high) of each factor are set and $n = 2$ runs of a 2^3 design are completed. The data is provided as an attachment.

(a) Perform tests for all main effects and for all interaction effects. State the F-statistic and p-value for each test of an effect deemed to be important. Fit a reduced model with the main effects and the statistically significant interaction effect.

(b) Provide a general definition of an interaction effect. Explain how an interaction plot is used in studying an interaction effect.

(c) Create a plot for the interaction effect deemed important. Provide an interpretation, stated in the context of the problem.

(d) Create a Boxplot showing the main effect for the remaining factor. Provide an interpretation, stated in the context of the problem.

(e) Create a plot of the fitted values for the reduced model. Which setting of the factors should be used if the goal is to maximize yield?

(f) Explain how the analysis is providing a simplification to the observed data.

4. An experiment to compare a new drug to a standard is in the planning stages. The response variable of interest is the clotting time (in minutes) of blood drawn from the subject. The experimenters want to perform a two sample t test at level $\alpha = .05$ with power $\pi = .90$ at $\delta_A = 0.5$, for standard deviation $\sigma = .7$

(a) Determine the sample size for each drug in order to achieve the stated test specifications.

(b) Graph the power curve for the chosen sample size. Explain how the power curve displays the desired properties of the test.

(c) Provide a general explanation of how δ_A can be determined.

(d) Briefly discuss some other issues that may provide additional insight to an experimental result beyond a finding of statistical significance.

(e) Briefly comment on the additional information provided by the p-value, beyond a determination of statistical significance alone.