Statistics 135 - Spring 2022

Take Home Final Exam

Due: June 9, 2022 at 5:30pm

Name:

Note, the exam is open book. You are not allowed to search for solutions online or communicate with another person about the content of this exam. Any attempt to do so will be considered academic dishonesty and will be reported. The exam has 5 problems, each worth 60 points. When you are asked to extract a random subset please, submit your selected data set with your final exam.

P1 The data in table 6.13 on page 349 contain measurements of skull sizes for 3 different time periods. Carry out a Manova as follows. First randomly select 15 observations from each of the 3 time periods. Then carry out a manova.

(a) Create the Manova table and calculate Wilk's Lambda Λ^*. Test at $\alpha = .05$ if there are any differences over time.

(b) Identify which, if any of the 4 variables exhibits changes over time.

P2 Subjects were asked to judge whether two two numbers were both even or odd or whether this was not the case. Numbers were presented as either Arabic numerals or words. Reaction times were recorded for all subjects. The type of number is indicated in the variable name as Word or Arabic, the type as same (both numbers even or odd) or diff (one even, one odd number). Select a random sample of 20 subjects for your analysis.

(a) Compare, using Hotellings' T^2 statistic the reaction times for same vs different types of numbers

(b) Compare using Hotellings' T^2 statistic the reaction times for Arabic vs word numbers.

P3 For the census tract data (table 8.5) carry out a principal components analysis on a randomly selected subset of 40 census tracts.

(a) Obtain the eigenvalues of the sample covariance matrix and determine the number of principal components needed from a screeplot. Obtain the associated eigenvectors and calculate the principal components for your sample data.

(b) Plot the first two principal components in a scatterplot.

(c) Find a 95% confidence region for the first two principal components.

P4 For the data in problem 3, carry out a factor analysis on the random sample on the random sample you selected.

(a) Construct the matrix $\hat{L}$, $\hat{L}\hat{L}^{prime}$ and $\hat{\Psi}$. Calculate the communalities for all factors and the specific variances.

(b) Calculate the factors scores by one of the methods presented in the book on pages 514-517. Clearly state the method you chose to estimate factor loadings.

P5 For the data in table 11.4 on bankruptcy, randomly select 11 observations from population π_1 and 15 observations from population π_2.

(a) Obtain the linear discriminant function for your data.

(b) For the data you did select and for the data you did not select,separately, calculate the confusion matrix. Compare the error rates for both subsets of data.