STAT 475-575 **ASSIGNMENT 3** **NAME _____**
Fall 2020

Assignment: Due Friday, October 15, 2021 by 11:59pm.

1. The data for this problem come from a comparative bioassay of potency of insulin made by two different manufacturers, coded 1 and 2. The two insulins were supposed to have the same potency. A total of 54 rabbits were used in this experiment with 27 rabbits randomly assigned to each combination of dosage levels and manufacturer. Blood sugar concentrations, Y_1, Y_2, Y_3, Y_4, Y_5, were measured on each rabbit at 1, 2, 3, 4, and 5 hours after the insulin was administered. An initial measurement, Y_0, of blood sugar concentration was taken just before the insulin was administered. Percent change from the baseline concentration for each rabbit were computed as

$$X_i = 100(Y_i - Y_0)/Y_0 , \qquad i = 1, 2, 3, 4, 5 .$$

The data are posted on the course web page as **rabbits2.csv**. There is one line of data for each rabbit and seven numbers on each line in the following order: rabbit identification number, manufacturer, X_1, X_2, X_3, X_4, X_5.

A. Examine differences in mean response for the two manufacturers by constructing a profile plot of the percent changes from the baseline concentration versus time, with one profile for each manufacturer. Display the plot in you solutions to this assignment. Do the profiles appear to be the same for the two manufacturers?

B. Test the null hypothesis that profiles (mean vectors) are the same for the insulins made by the two manufacturers. Report the value of the two-sample Hoteling T^2 statistic, the corresponding F-value, degrees of freedom, and the p-value. State your conclusion.

C. Use the Bonferroni method to construct simultaneous 95% confidence intervals for the difference in the mean responses (percent change in blood sugar concentration from baseline) to the insulins made by the two manufacturers at each of the 5 time points. State your conclusions.

D. The Hoteling T^2 test was performed using the model assumption that the distribution of responses among rabbits at the five inspection times is a five-dimensional normal distribution for each insulin. Do the data conform to that condition? Perform an appropriate test and report the p-value and your conclusion.

2. Gerrild and Lantz (1969) chemically analyzed crude oil samples from three zones of sandstone :
   (zone=1) Wilhelm
   (zone=2) Sub-Mulinia
   (zone=3) Upper

   They measured the amounts of the five trace elements:

   X_1: vanadium (in percent ash)

   X_2: iron (in percent ash)

   X_3: beryllium (in percent ash)

   X_4: saturated hydrocarbons (in percent area)

   X_5: aromatic hydrocarbons (in percent area)

   The last two measurements are determined from areas under a gas-liquid chromatography curve. The data for 56 crude oil samples are posted on Blackboard as **crudeoil.csv**.

   (a) Perform a one-way MANOVA to test the null hypothesis of no differences among the three sandstone zones with respect to the mean vectors for the amounts of the five trace elements. Report values for Wilks Lambda and the corresponding F-value, degrees of freedom, and the p-value. State your conclusion.

   (b) Use the Bonferroni approach to determine which trace elements have significantly different means across the different sandstone zones, using an experiment wise type I error level of 0.05. Note that you must construct three confidence intervals (or do three t-tests) to compare the three pairs of sandstone zones for each trace element, a total of fifteen confidence intervals. For each trace element, report your conclusions about which sandstone zones have significantly higher or lower mean levels. (For example, crude oil extracted from the Wilhelm zone has a significantly lower mean level of vanadium than crude oil extracted either the Sub-Milina or the Upper zones. There is no difference in mean vanadium levels for crude oil extracted from the Sub-Milina and Upper zones.)