Statistics 511

**Final Take Home Portion**

**Fall 2022**

# Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Due Thursday, December 8th at 2:00 PM.**

**Ground Rules: -- Open book and notes. This should be your own work. Show your work!!! Do a complete analysis for each problem.**

# Please sign. This is all my own work \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. **(12 points)** On Canvas find the strength (take home problem 1) data relating a response, Strength (gm), to four predictors:

Force (gm), Power (mw), Temp (deg C), and Minutes (time).

This was a planned experiment with complete randomization.

**Find the best fitting regression equation for this data. Consider all main effects [Force (gm) , Power (mw) , Temp (deg C) , and Minutes (time)] and the second order terms (two-way interactions and squared terms). Do a complete job of data analysis. Use any of the tools we have learned that seem appropriate. Include relevant computer output.**

**Assume that high strength is desired. What approximate values of the predictors gives us the highest predicted strength? What is the predicted highest strength?**

1. **(14 Points)** The rise in algae in coastal waters is thought to be due to increases in nutrients such as nitrate (NO3) and other forms of nitrogen. It is theorized that the excessive amounts of nitrate are due to human influences. The data on Canvas contains the following variables: runoff, the average annual runoff from the watershed; depos, deposition of naturally occurring nitrate; nprec, nitrate containing precipitation; area, area of watershed; discharg, the annual average discharge of the river into an ocean; prec, annual precipitation; density, an estimate of human population density; NO3, nitrate concentration; and export, nitrate export as a function of runoff and nitrate. **Are either of the dependent variables, NO3 and/or Export, related to the independent variable density after adjusting for the other demographic (runoff – prec) variables? Do a complete analysis for both dependent variables.**

**(OVER)**

1. **(9 Points)** Consider the data for problem 3 on Canvas.
2. **Compute the least squares regression line of Y on X.**
3. **Compute the quadratic (keep X in the model also) regression equation. Is the quadratic term statistically significant?**
4. **Drop the quadratic term from the model and compute the multiple regression equation that allows for different intercepts for males and females but has the same slope. Is the difference in intercepts statistically significant?**
5. **Compute a single multiple regression equation that allows for both different slopes and different intercepts for males and females.**
6. **For the equation in part d) what are the slope and intercepts for both males and females?**
7. **(9 points)** Octane ratings for a particular brand of gasoline were sampled from service stations selected randomly from four regions of the northwestern United States, which we will designate simply as A, B, C, D. The data are on Canvas.

**Check the analysis assumptions. Show all calculations and the printouts from which the information for the calculations came. Using Extra-Sum-of-Squares F-tests, determine which of the following three models is preferable:**

1. **Constant octane for all four regions.**
2. **Four separate regions.**
3. **Two regions (A & C) versus (B & D)**
4. **(6 Points)** For your assigned chapter write three questions: 1 – true/false, 1 – fill in the blank, and 1 – multiple choice. Based on your last name, if it begins with A – C then write questions for chapter 9; D – J chapter 10; K – M chapter 11; and N – Z chapter 12. **Email these questions to me.**