APPLIED MASTERS EXAM
SPRING 2024

Please prepare a written report to each of these problems (max 10 pages per problem) and provide an electronic copy to each member of the Master's Exam committee before **5:00 PM Tuesday 4/02**. The committee consists of Profs Craig, Chakraborty, Gu, and Xi. Your oral exam will take place on Thursday 4/04 or Friday 4/05 and should already be scheduled. **The oral exam lasts no more than 55 minutes** so we can stay on schedule.

At the oral exam, you will present a summary of your work (roughly a 30-minute presentation). You should prepare to discuss both problems. Because the same two problems are being addressed by everyone taking the exam, you will not need to go over the background of each problem in detail. Instead, devote most of your presentation to your design and/or analysis and conclusions. During these presentations you will likely be asked questions on what you have done, said, and written in your report. Other general questions regarding material covered in your courses may also be asked.

These reports and presentations should be **geared towards the client** but be prepared to expand on these answers should a committee member request it. You can assume the client has taken an introductory statistics course in the past but knows very little statistics beyond that. Your written and oral summaries can include excerpts from software output. However, make sure to incorporate these excerpts directly into the report. Also, do not submit large quantities of output that do not directly relate to your conclusions. Results should be illustrated graphically whenever possible.

You are not permitted to discuss these exam problems with anyone until all examinees have completed their oral exam. You can, however, email Prof Craig any "consulting client" questions regarding the problems **prior to Wednesday 3/27 at 11:59 PM.** He will answer them as if he were the client no later than Thurs 3/28 at 12:00 PM.

Be sure to number all of the pages of your report for easy reference.

Good Luck!!

1. Your client would like to examine the effects of three factors associated with autonomous driving on heart-rate variability. Each factor–transparency, recommended control mode, and system reliability–has a high and low level, so there are a total of eight factor combinations or treatments. The plan is for each subject to participate in two simulated drives, with each drive consisting of switches between the combinations of two of the three factors (the other factor being held constant). The client is planning there to be four such switches per drive so an example of a drive could be

$$\begin{array}{cccc} \text{Blk 1} & \text{Blk 2} & \text{Blk 3} & \text{Blk 4} \\ + - - & + - + & + + + & + + - \end{array}$$

where the first factor is held at the high level throughout and the combinations of the second and third factors are switched. The subjects stay in each combination for one to two minutes and heart rate variability is calculated using the middle 30 seconds of each block.

The client has questions regarding the design of an experiment that allows her to compare the eight treatment combinations (and thus the factors themselves) while also allowing for the possibility of first order residual or carryover effects. She'd love to only create two virtual drives and have each subject run through them but she's expecting that she will have to vary the order of factor combinations across subjects. Your job is to help construct a design (i.e., series of drives) that does the job yet is not overly complicated to implement. She expects to recruit around 50 subjects and does not want to have to create 100 unique drives.

Write up what you consider the best design for this study including the reasoning of why you chose the design, as well as a description of any randomization and the method of analysis you will use.

2. Professor Craig drives a Prius Prime, which has a hybrid system that runs on both electricity and gas. When charged the car can run for about 25-30 miles on electricity before changing over to gas. This means that he can drive around Lafayette/West Lafayette on electric most days and not use gas.

   Since he purchased the car, he's been keeping track of his mileage. Every time he fills up the tank, he records the computer generated total miles, overall mpg, number of miles since the last fill up and the mpg during that time. He also records the date and how much gas (in gallons) he needed to fill up the tank.

   You are being asked to analyze these data and see if there are any patterns. For example, one can also compute interval mpg using the miles driven during that interval and the gas needed to fill up the tank. This is often lower than the computer generated interval mpg. Is it consistently lower or does it vary based on certain factors like season and time since last fill up? The data can be found at

   `www.stat.purdue.edu/~bacraig/msexam/Prius Prime.xlsx`.