

MATH 338

FINAL EXAM - LAB PORTION

DUE THURSDAY, DECEMBER 20, 2018

Your name: _____

Your scores (to be filled in by Dr. Wynne):

Problem 1: ____/4

Problem 2: ____/6

Problem 3: ____/4

Problem 4: ____/4

Total: ____/18

You have one week to complete this exam, save your answers (as a .docx or .pdf file) and upload the file to Titanium.

You may refer to your notes, your textbook, and any pre-existing online reference (eBook, R/Rguroo help, anything on Titanium). You may search for help online, but you must cite any source found through the search. You may ask Dr. Wynne to clarify what a question is asking for. You may not ask other people for help or use any other resources.

For full credit, show all work except for final numerical calculations (which can be done using a scientific/graphing calculator or R).

1. If you're as bad at chem lab as I was, you will have a lot of measurement error affecting your results. Suppose that I am terrible at pipetting and, when I attempt to pipette 10 mL of water, the actual amount I pipette is normally distributed with mean 10 mL and standard deviation 0.3 mL.

For parts (a) and (b) below, round your answers to 3 decimal places. Also, include some way of indicating to me how you got software to compute your answer (Rguroo screenshots and/or R code).

a) [1.5 pts] What is the probability that I pipette less than 9.9 mL of water?

Write your answer and include your software code/output below.

b) [2 pts] What is the probability that, over 10 independent attempts, I average less than 9.9 mL of water in the pipette?

Write your answer and include your software code/output below.

c) [0.5 pts] Go to the following site: http://fullerton.qualtrics.com/jfe/form/SV_8wUu57lprlFPH1P. Write your name and the date in the boxes at the top of the survey and take a screenshot. Paste the screenshot below. Then read the rest of the page and either accept or decline participation in this research study. If you do participate in the study, fill out the rest of the survey. (I cannot give credit for participating because I'm not supposed to know who is in the study)

2. Having prolonged high blood pressure can have many negative effects on general health and can potentially lead to stroke and/or death. There are many drugs available that are known to help lower blood pressure, however, these drugs often come with negative side effects. A clinical trial was set up to explore a more natural treatment of taking a potassium, magnesium, and calcium (PMC) complex tablet. These minerals are known to counteract the blood pressure raising effects of high sodium intake.

A random sample of 200 adult males (29-49 years of age) currently diagnosed with hypertension (high blood pressure) were obtained from a pool of subjects who volunteered to be part of the study. The subjects were then randomly assigned to one of two groups: one group received the PMC complex tablet and the other received a sugar tablet.

The *PMC.csv* and *PMC1.csv* files contains the change in blood pressure, to the nearest mmHg, over the course of the study in both the PMC and Sugar groups. (These files contain the same data, presented in two different forms. You may use either form you are comfortable with) Note that positive values mean that the blood pressure went up over the course of the study, and negative values mean that the blood pressure went down.

a) [1.5 pts] Perform exploratory analysis on this dataset – describe the distribution of the change in blood pressure in each group in the observed sample of 200 adult males and compare the two distributions.

Write a short paragraph describing and comparing the distributions below. Include any relevant graphs and other software code/output you produced.

b) Perform an appropriate statistical hypothesis test to draw conclusions about the effectiveness of the PMC complex. Write a short paragraph describing your conclusions. Full credit will be given in this part for a writeup that:

[0.5 pts] Uses background analysis and the exploratory analysis from part (a) to inform your decision about the correct hypothesis test to use

[1 pt] Specifies an appropriate null hypothesis, alternative hypothesis, and significance level

[1 pt] Contains relevant, necessary and correct (R code and) software output

[2 pts] Makes a statistically justified conclusion and interprets that conclusion in the context of the problem.

Write your answer and include your software code/output below.

2. A 2018 study investigated whether macaque monkeys can perform statistical reasoning. In the study, 11 macaques watched a researcher select food from two buckets. On each trial, the researcher drew a piece of food from a bucket containing 80% grapes with one hand and drew from a bucket containing only 20% grapes with the other hand. After looking at the drawing, the monkeys indicated which hand they wanted food from. Each monkey selected a hand on 12 trials.

The *monkeysee-monkeydoinference.csv* file contains the name of the monkey (Individual) and the number of “correct” choices out of 12 that the monkey made (Score).

Estimate with 95% confidence the proportion of the time that Sophie would make the “correct” choice if presented with very many trials. Do you believe that Sophie is picking a hand entirely at random, or do you think that she is thinking about which hand is more likely to have the grapes?

Hint: You don’t need to actually import this file into the software! You can just figure out what the appropriate numbers are!

Full credit will be given in this problem for a writeup that:

[1 pt] Includes appropriate background and/or exploratory analysis that informs your decision about the correct confidence interval procedure to use

[1 pt] Contains relevant, necessary and correct (R code and) software output

[0.5 pts] Correctly identifies the confidence interval from the output

[1.5 pts] Correctly interprets the confidence interval in the context of the question asked.

Write your answer and include your software code/output below.

4. The *lumefantrine.csv* dataset contains the concentrations (in ng/mL) of the anti-malarial drug lumefantrine two hours after ingestion in 31 pregnant women. The variable *Vc_2h* contains the concentrations in venous plasma (veins) and the variable *Cc_2h* contains the concentrations in capillary plasma.

Researchers would like to predict the concentration of lumefantrine in capillary plasma (response variable) from the concentration of lumefantrine in venous plasma (predictor variable). When researchers make a scatterplot of the data for the 31 women, they notice that the points have an approximately linear relationship and that the slope of the regression line is approximately 1.

Suppose that a new pregnant woman is given a dose of lumefantrine and, two hours later, the concentration in her venous plasma is measured to be 1700 ng/mL. Predict with 95% confidence the concentration of lumefantrine in her capillary plasma. Do you have faith in this method of prediction, or are assumptions necessary for this prediction violated?

Full credit will be given in this problem for a writeup that:

[1 pt] Contains relevant, necessary and correct (R code and) software output

[1 pt] Provides a correct interval estimate from the output

[2 pts] Includes appropriate background and/or exploratory analysis and arrives at a reasonable conclusion about the appropriateness of the inference

Write your answer and include your software code/output below.