MATH 338

MIDTERM 1 - LAB PORTION

TUESDAY, NOVEMBER 6, 2018

Your name: _____

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

Problem 1: ____/6

Problem 2: ____/3

Problem 3: ____/4

Total: ____/13

You have 50 minutes to complete this exam and an extra 10 minutes to save (as a .docx or .pdf file) and upload it 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. Post-menopausal women often develop sudden feelings of heat, skin redness, and sweating collectively known as hot flashes. Lambert and colleagues (2018) were interested in using Red Clover Extract to decrease the frequency and severity of these hot flashes in post-menopausal women. The rhfs.csv file on Titanium contains the self-reported change in hot flash intensity (HFI) over 12 weeks for their sample of 58 post-menopausal women.

Is there a significantly different effect of Red Clover Extract on hot flash intensity as compared to Placebo? Write 1-2 short paragraphs to answer this question. Support your answer by including and referring to software output.

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

a) [2 pts] Includes appropriate background and exploratory analysis that informs your decision about the correct hypothesis test to use

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

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

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

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

2. Suppose a gas station chain is interested in showing that a competitor's gas prices are higher. One morning, they decide to send people out to record the Regular Unleaded gasoline price at 50 randomly selected competitor's stations throughout Southern California. The company knows the following:

- On this day, the company's own stations in Southern California sell Regular Unleaded for, on average, $3.80 per gallon
- Previous studies have suggested that a sample standard deviation of $0.15 per gallon is a reasonable assumption

Is 50 a large enough number of gas stations to show that the competitor's prices are higher, if the competitor's true average price is $3.85 per gallon? Write a short report (1-2 sentences) answering this question. Support your answer by including and referring to software output.

Rguroo hint: you can type in values of the relevant statistics into the Mean Inference Basics and Details dialog; you don't need a dataset to do this problem.

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

a) [1 pt] Uses a correct statistical procedure

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

c) [1 pt] Correctly interprets the software output in the context of the problem to answer the question.

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

3. They'll measure anything these days. The disgust.csv contains scores of 159 Americans on various scales that measure how disgusting they find germs, pathogens, and similar things.

The variable BODS in the dataset represents people's disgust with body odor, normalized from 1 (not disgusting at all) to 5 (completely disgusted by it). Estimate with 95% confidence the mean score on this scale of body odor disgust if all Americans were to take the researchers' survey. Write a short sentence interpreting your estimate.

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

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

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

c) [2 pts] Correctly interprets the confidence interval in the context of the problem.

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