Homework 9 - MATH 141

Due Date: Wednesday 12/08/2021, 11:59 PM

Instructions:

- Please provide complete answers/solutions for each question/problem.

- **If it involves mathematical computations, please provide your reasoning and/or detailed solutions.**

- There are two ways you can write your answers, a: by handwriting (either physically or digitally), or b: by typing on a template document with file type options, Word or RMarkdown, which can be downloaded from the course website.

- If you had handwritten your answers/solutions on a physical paper, make sure to label it properly and please scan your document using a scanner app for convenience. Suggestions: (1) "Tiny Scanner" for Android or (2) "Scanner App" for iOS.

- If a problem asks you to show your R code, R outputs, or R plots, please provide them as additional pages into your current homework pdf while labeling them properly. This means that, **if you have handwritten your homework solutions and saved it as pdf, you would need to merge the separate pdf which contains your R code, R outputs, or R plots. Note that all of the problems that require R does not require you to show your R code - unless the problem specifically says so.**

- If you have questions or concerns, please feel free to ask the instructor.

- **Please save your work as one pdf file, don't put your name in any part of the document, and submit it to the Gradescope page for this course. Your document upload will correspond to your name automatically in Gradescope.**

I. Inference for Comparing Two Means

The exercise problems shown below was taken and slightly modified from your textbook OpenIntro: Introduction to Modern Statistics Sections 20.6 and 21.5

1. Gaming, distracted eating, and intake.

Consider the following statement and answer the following questions. Note that when performing inference, you must include the null and hypothesis statements (words and mathematical symbols), and interpretations of the values - p-value and confidence interval - in the context of the problem.

> A group of researchers who are interested in the possible effects of distracting stimuli during eating, such as an increase or decrease in the amount of food consumption, monitored food intake for a group of 44 patients who were randomized into two equal groups. The treatment group ate lunch while playing solitaire, and the control group ate lunch without any added distractions. Oldham-Cooper et al. 2011

Patients in the treatment group ate an average of 52.1 grams of biscuits, with a standard deviation of 45.1 grams, and patients in the control group ate an average of 27.1 grams of biscuits, with a standard deviation of 26.4 grams.

    a. Do these data provide convincing evidence that the average food intake (measured in amount of biscuits consumed) is different for the patients in the treatment group compared to the control group? Assuming that conditions for conducting inference using mathematical models are satisfied, conduct a hypothesis test to answer the given question. Use significance value of 0.05.

    b. How much of a difference are there for the patients in the treatment group compared to the control group? Assuming that conditions for conducting inference using mathematical models are satisfied, calculate the 95% confidence interval to answer the given question.

2. Paired or Not Paired?

In each of the following scenarios, determine if the data are paired or not paired. Explain your reasoning.

    a. Compare pre- (beginning of semester) and post-test (end of semester) scores of students.

    b. Assess gender-related salary gap by comparing salaries of randomly sampled men and women.

    c. We would like to know if Intel's stock and Southwest Airlines' stock have similar rates of return. To find out, we take a random sample of 50 days, and record Intel's and Southwest's stock on those same days.

    d. We randomly sample 50 items from Target stores and note the price for each. Then we visit Walmart and collect the price for each of those same 50 items.

II. Inference for Linear Regression

The exercise problems shown below was taken and slightly modified from your textbook OpenIntro: Introduction to Modern Statistics Section 24.8.

1. I Heart Cats I.

Consider the following statement and answer the following questions.

> Researchers collected data on heart and body weights of 144 domestic adult cats. The table below shows the output of a linear model predicting heart weight (measured in grams) from body weight (measured in kilograms) of these cats.

The `cats` data used in this exercise can be found in the **MASS** R package.
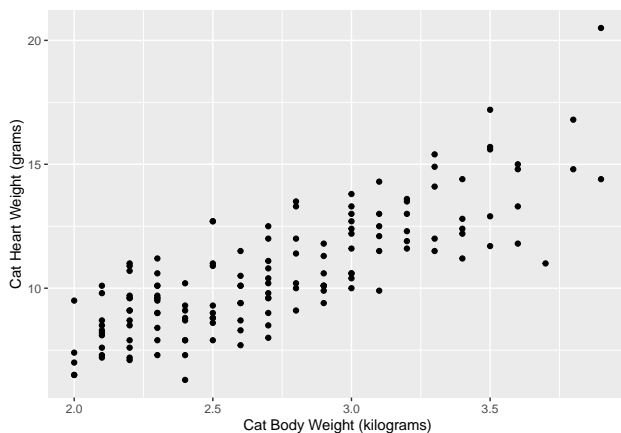
term	estimate	std.error
(Intercept)	−0.3567	0.6923
Bwt	4.0341	0.2503

    a. Write the equation of the regression line. Interpret the slope and intercept in context.

    b. State the null and alternative hypotheses.

    c. Assuming that the conditions for inference for linear regression is satisfied, compute the T statistic and the p-value (using the theoretical method), and state your conclusion. Use significance value of 0.05.

2. I Heart Cats II.

Use the information provided in problem 1.

    a. Assuming that the conditions for inference for linear regression is satisfied, use the theoretical method to calculate a 95% confidence interval for the slope of body weight, and interpret it in context of the data.

    b. Do your results from the hypothesis test in problem 1 and the confidence interval you computed in part (a) agree? Explain.

    c. Describe the relationship between cat body weight and heart weight. The correlation coefficient for the body and heart weight is 0.8. Calculate R^2 and interpret it.