Homework 9

Insert Name

Math 141, Week 9

Due: 11:59pm, Friday April 1

Instructions

Work through the problems below and submit this document as a knitted .pdf to the Math 141 S22 Wells Lecture gradescope page.

For each problem, put your solution between the bars of red stars.

Acknowledgements

If you work with a classmate, please write a note acknowledging this.

Exercise 1

The consumption of caffeine to benefit alertness is a common activity practiced by 90% of adults in North America. Often, caffeine is used in order to replace the need for sleep. One recent study compares student's ability to recall memorized information after either the consumption of caffeine or a brief nap. A random sample of 35 adults were randomly divided into 3 groups and verbally given a list of 24 words to memorize. During a break, one of the groups was instructed to take a nap for an hour and a half, another group is kept awake and given a caffeine pill an hour prior to testing, and the third group is kept awake and given a placebo. The response variable of interest is the number of words participants are able to recall following the break.

Summary statistics for the three groups are presented in the table below. We are interested in testing whether there is evidence of a difference in average recall ability between any two of the treatments. Thus, we have 3 possible tests: Sleep vs. Caffeine, Sleep vs. Placebo, Caffeine vs Placebo.

Group	Sample Size	Mean	St. Dev
Sleep	12	15.25	3.3
Caffeine	12	12.25	3.5
Placebo	11	13.70	3.0

a. In the test comparing the sleep to caffeine group, the p-value is 0.003. What is the conclusion of the test? In the sample, which group had the better recall ability? According to the hypothesis test, can you conclude that sleep is actually better than caffeine for recall ability (carefully consider the implied alternate hypothesis in the problem statement)?

b. In the test comparing the sleep group to the placebo group, the p-value is 0.06. What is the conclusion of the test using a 5% significance level? Using a 10% significance level? How strong is the evidence of a difference in mean recall ability between these two groups?

c. In the test comparing the caffeine group to the placebo group, the p-value is 0.22. What is the conclusion of the test? In the sample, which group had better recall ability? According to the hypothesis test, are we justified in concluding that caffeine consumption has no effect on recall ability?

d. Just based on the results of these samples alone, what seems to be the best strategy to implement before a memory-intensive exam? Explain why this doesn't mean the result you chose has the highest likelihood of increasing your exam score.

Exercise 2

For each of the following situations, describe the effects of a Type I and Type II error in context. Then indicate whether it makes more sense to use a larger significance level (such as alpha = 0.10) or a smaller significance level (such as alpha = 0.01), and briefly explain your reasoning.

a. A new drug with potentially dangerous side effects is tested to see if it is better than the drug currently in use. If it is found to be effective, it will be prescribed to millions of people.

b. Two slightly different versions of a statistics exercise are given to a 9am and 10am section Math 141 and the proportions of students in each class who get the question correct are recorded. We wish to determine whether one question was more difficult than the other.

c. A consumer advocacy group wishes to test if a popular company is engaging in false advertising. If there is evidence the company is lying, the Federal Trade Commission will file a lawsuit against the company.

d. You test whether taking a vitamin supplement reduces the duration of symptoms of the common cold. There are no (known) harmful side effects of the supplement.

e. A government intelligence agency notices increased activity in online chatrooms often frequented by members of terrorist group. If they forward the report to the executive branch, a pre-emptive military strike will be ordered against a suspected terrorist compound in a foreign country.

Exercise 3

The Stanford University Heart Transplant Study was conducted to determine whether an experimental heart transplant program increased lifespan. Each patient entering the program was designated an official heart transplant candidate, meaning that they were gravely ill and would most likely benefit from a new heart. Some patients got a transplant and some did not. The variable `transplant` indicates which group the patients were in; patients in the treatment group got a transplant and those in the control group did not.The variable `survtime` indicates the number of days patients were alive after the date they were determined to be a candidate for a heart transplant.

The data from this study can be loaded with the following code chunk:

```
library(openintro)
data("heart_transplant")
```

a. Based on the information in the statement of the problem, what is the research question that the Stanford University Heart Transplant Study sought to answer?

b. State the implied null and alternative hypotheses for this study.

e. What are the consequences of a type I and type II error in this case?

d. Based on your response to the previous part, which significance level seems most appropriate? (Choose from among $\alpha = 0.1, 0.05, 0.01$, where $\alpha = 0.05$ represents the approach that balances Type I and Type II errors).

e. Use `dplyr` verbs to create a single data frame with 2 rows corresponding to the treatment and control groups in the experiment, and with 4 columns: the name of the group, the number of patients in the group, the mean survival time, and the standard deviation of survival time.

f. Use `infer` to conduct a hypothesis test at the level you selected in part (d). What is the p-value for your test?

g. Based on your p-value, make a conclusion about your hypotheses.

Exercise 4

In March 2020, biomedical researchers in France conducted a study to test whether a popular anti-malaria drug was effective in reducing symptoms of the novel coronavirus. The report can be found here:

https://www.sciencedirect.com/science/article/pii/S0924857920300996

The results of the study were widely disseminated immediately after the report was released, and soon several world leaders suggested that a promising cure for coronavirus had been identified. However, many in the biomedical statistics community expressed skepticism about the results.

In this problem, we will first perform a hypothesis testing imitating the procedure used by the study's authors, and then will investigate some methodological flaws in the original study.

The original study separated patients testing positive for coronavirus into two groups: a treatment group consisting of 20 patients receiving a daily doses of hydroxychloroquine sulfate 200mg, and a control group consisting of 16 patients receiving "standard" coronavirus care. Patients in each group were tracked for 6 days, and the number of patients in each group no longer test positive for the virus on the 6th day were recorded.

Run the following code to load the results in the `coronavirus` data frame.

```
coronavirus <- tibble(
  test_corona = c(replicate(6, "positive"), replicate(14, "negative") , replicate(14, "positive"), repli
  group = c(replicate(20, "treatment"), replicate(16, "control"))
                        )
```

a. State the researchers null and alternative hypotheses, both in words and mathematical symbols.

b. Use `infer` to conduct a hypothesis test at the alpha = 0.01 significance level.

c. Based on the P-value obtained in the test, make a conclusion about the researcher's hypotheses.

d. In the experiment design, the researchers elected to use as a control the patients who refused treatment or who had an exclusion criterion (i.e could not safely be given the treatment medication due to pre-existing conditions like high blood pressure). Explain why should cause skepticism about your conclusion in part (c). Does the size of the P-value in computed in the previous part overcome this skepticism?

e. Read the following March 30, 2020 article published in the *Annals of Internal Medicine*:

https://www.acpjournals.org/doi/10.7326/M20-1223

What are two other methodological flaws the authors of this article identified in the original study?