STA130H1S – Fall 2022

Problem Set 10

() and STA130 Professors

Instructions

Complete the exercises in this .Rmd file and submit your .Rmd and .pdf output through Quercus on Thursday, December 1st by 5:00 p.m. ET.

Part 1: Lumosity

Lumosity is a brain training app thought to help cognitive skills – for example, memory, reasoning and focus. A large randomized trial was conducted to evaluate the impact of Lumosity training on cognitive skills. The study and results are presented in ["Enhancing Cognitive Abilities with Comprehensive Training: A Large, Online, Randomized, Active-Controlled Trial" (2015, Hardy et al.)[https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134467].

Thousands of participants were recruited from Lumosity's free users (i.e., people who set up free Lumosity accounts but did not pay for full access) and randomly assigned to either:

- Lumosity training (Treatment): complete Lumosity training online for approximately 15 minutes at a time, at least 5 times a week for 10 weeks, or
- Crossword puzzles (Control): complete crossword puzzles online for approximately 15 minutes at a time, at least 5 times a week for 10 weeks.

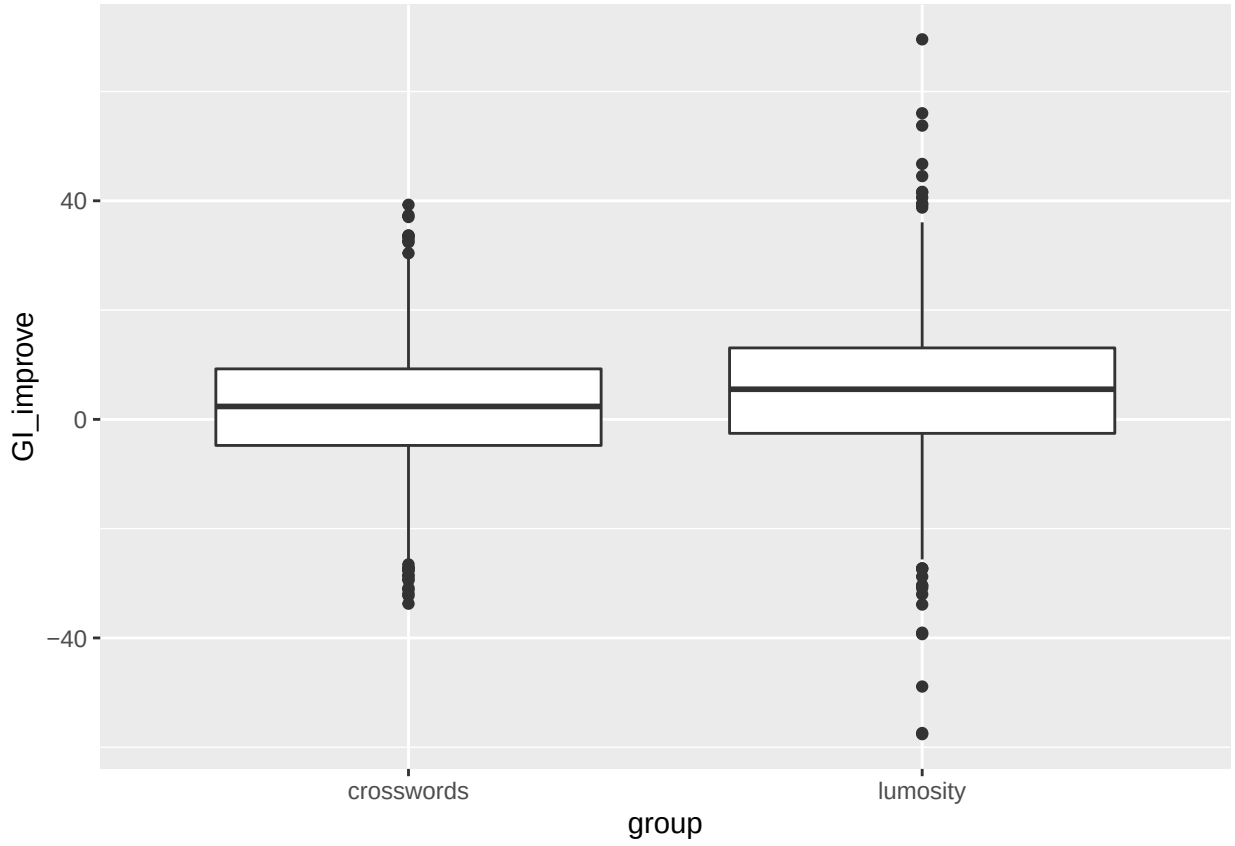
The data on the improvement (i.e., after-before) in GI Scores is `GI_improve`, and this as well as several other potentially useful variables from the study are stored in the data frame called `study_dat` for the 5045 Lumosity users who participated in the study.

```
library(tidyverse)
study_dat<-read_csv("lumosity_study_data.csv")
glimpse(study_dat)
```

```
## Rows: 5,045
## Columns: 6
## $ participant_id     <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ~
## $ group              <chr> "crosswords", "lumosity", "crosswords", "crosswords~
## $ age_round          <dbl> 51, 24, 19, 21, 31, 25, 45, 27, 40, 50, 46, 52, 21,~
## $ GI_improve         <dbl> -5.992334, 9.030539, -15.030917, -9.265834, 3.18986~
## $ concentration_post <dbl> 1, 3, 5, 2, 4, 1, 4, 5, 5, 4, 4, 4, 5, 4, 4, 2, 4, ~
## $ active_days        <dbl> 56, 54, 6, 21, 43, 15, 0, 40, 39, 69, 64, 32, 21, 4~
```

The main measure of cognitive skills was called the Grand Index (GI) Score, and higher GI values mean better cognitive skills. The cognitive skills of the participants who completed the study were scored before and after the 10-week study period, and the `GI_improve` measures by type of online training are as follows.

```
ggplot(study_dat, aes(x=group, y=GI_improve)) + geom_boxplot()
```




```
group_by(study_dat, group) %>%
  summarise(mean=mean(GI_improve), sd=sd(GI_improve), n=n())
```

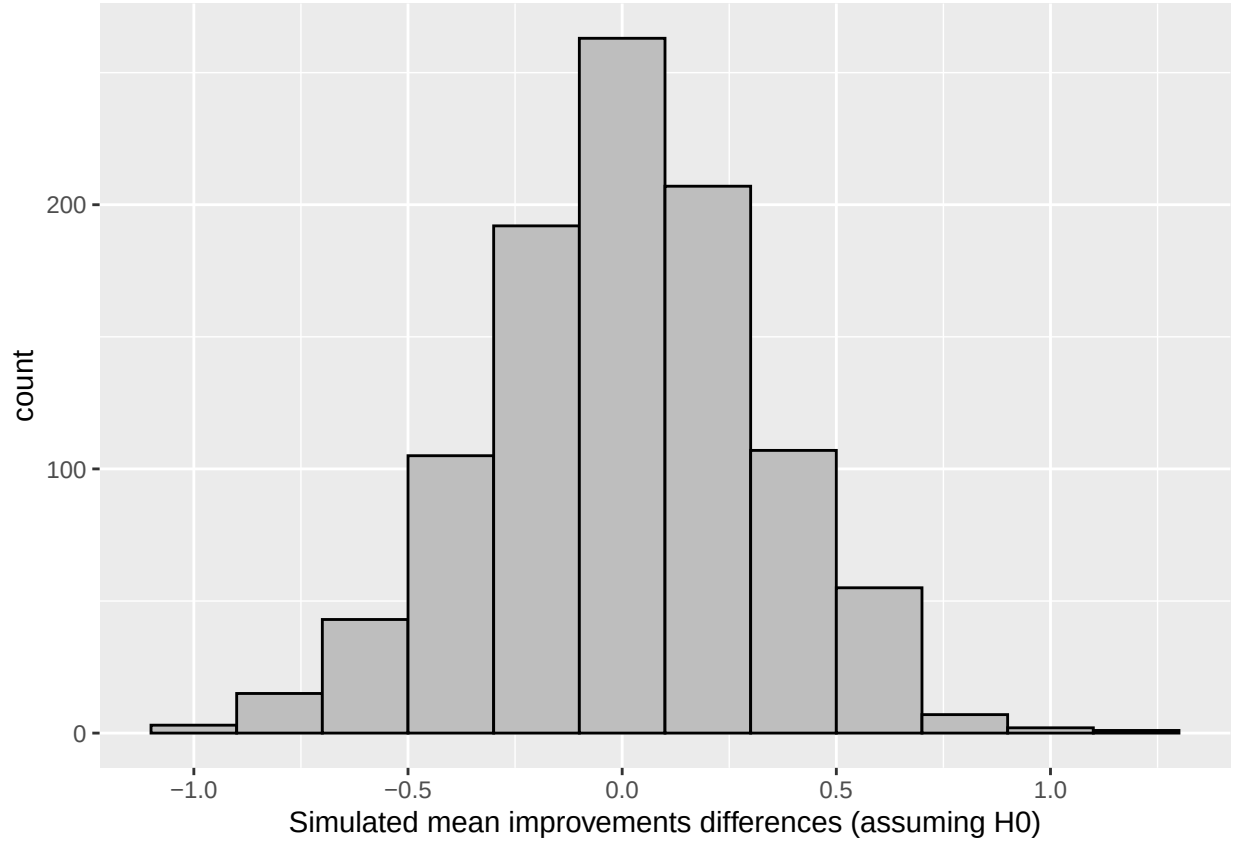
```
## # A tibble: 2 x 4
##   group       mean    sd     n
##   <chr>      <dbl> <dbl> <int>
## 1 crosswords  2.14  10.6  2378
## 2 lumosity    5.24  12.0  2667
```

Question 1: Cognative Performance Increases?

A hypothesis test on two groups can be conducted to compare mean Grand Index score improvements after online training with Lumosity and crosswords.

```
# compute test statistic
test_stat<-as.numeric(study_dat %>%
  group_by(group) %>%
  summarise(means = mean(GI_improve), .groups='drop') %>%
    #.groups='drop' is included to avoid a warning message being
    # printed, but doesn't change behaviour
  summarise(value = diff(means)))

# conduct randomization test
simulated_values <- rep(NA, 1000)

for (i in 1:1000) {
  sim <- study_dat %>% mutate(group = sample(group))
```

```
  sim_value <- sim %>%
    group_by(group) %>%
    summarise(means = mean(GI_improve), .groups='drop') %>%
    summarise(value = diff(means))
  simulated_values[i] <- as.numeric(sim_value)
}

sim <- tibble(mean_diff = simulated_values)

ggplot(sim, aes(x=mean_diff)) +
  geom_histogram(col="black",fill="gray", binwidth=0.2) +
  labs(x = "Simulated mean improvements differences (assuming H0)")
```




```
sim <- tibble(mean_diff = simulated_values)
sim %>%
  filter(mean_diff >= abs(test_stat) |
           mean_diff <= -1*abs(test_stat)) %>%
  summarise(p_value = n() / 1000)
```

```
## # A tibble: 1 x 1
##   p_value
##     <dbl>
## 1       0
```

The estimated p-value based on this randomization test is 0 so there is very strong evidence against the hypothesis that the mean Grand Index Score improvement is the same for those training using the Lumosity app and those completing online crossword puzzles.

(a) Consider using a simple linear regression model instead to test for a difference in mean Grand Index score improvements for those who train using Lumosity and those who complete online crossword puzzles. *(i)* Write down the appropriate regression model. Explicitly define the terms you use.

REPLACE THIS TEXT WITH YOUR ANSWER

(ii) State the hypotheses to compare mean `GI_improve` when training using Lumosity and online crossword puzzles based on the model you specified in the previous part of this question.

REPLACE THIS TEXT WITH YOUR ANSWER

(iii) Use R to fit this model and interpret the estimated regression coefficients.

```
# Code your answer below
```

(iv) Interpret the p-value of this test to compare the mean improvement for Lumosity versus crossword puzzles. How does it compare to the p-value estimated using the randomization test earlier in this question? Is this surprising? Why or why not.

REPLACE THIS TEXT WITH YOUR ANSWER

(b) Consider the study design used by Hardy et al. (2015). *(i)* What type of study did Hardy et al. (2015) conduct?

REPLACE THIS TEXT WITH YOUR ANSWER

(ii) As reported in Hardy et al. (2015), 9919 participants consented to participate and were randomly assigned to a training type. However, only 5045 study participants actually completed the study. The dataset only included data on study participants who completed the study. Could this limit our conclusions? How so?

REPLACE THIS TEXT WITH YOUR ANSWER

(iii) Is the data we analyzed above (and got a *p-value* of 0 for) the full data of a controlled randomized trial? Or a subset of data that is based on individual behavior, and hence in some sense observational?

REPLACE THIS TEXT WITH YOUR ANSWER

(iv) With these pieces of information in mind, can we conclusively conclude that Lumosity training leads to more improvement in cognitive skills than completing crossword puzzles online based on these results? Explain your answer.

REPLACE THIS TEXT WITH YOUR ANSWER

(c) Perhaps age of the user is related to cognitive improvement as well. *(i)* Do you think user ages would be different between the Lumosity group and crossword groups? Why or why not?

REPLACE THIS TEXT WITH YOUR ANSWER

(ii) Produce an appropriate data summary to see if ages of the users differ for the Lumosity and crossword groups. Interpret your summary and comment on how this compares to your prediction about how ages would compare.

REPLACE THIS TEXT WITH YOUR ANSWER

```
# Code your answer below
```

(iii) Suppose that there was a big difference in the age distributions of the two treatment groups. For example, suppose that younger users were much more likely to drop out of the Lumosity group than to drop out of the crossword puzzle group, and so the mean age of individuals who completed the study was 50 for the Lumosity group and 38 for the crossword puzzle group. How might this limit the conclusions of the analysis, if at all?

REPLACE THIS TEXT WITH YOUR ANSWER

Part 2: OPTIONAL for Practice

You may complete these questions for practice if you wish. ***You are not required to complete these questions as they ARE NOT included as part of your mark.***

Question 2:

From Section 8.12 Problem 8 in "Modern Data Science with R Exercise, 2nd edition": A data analyst received permission to post a data set that was scraped from a social media site. The full data set included name, screen name, email address, geographic location, IP (internet protocol) address, demographic profiles, and preferences for relationships. Why might it be problematic to post a deidentified form of this data set where name and email address were removed?

REPLACE THIS TEXT WITH YOUR ANSWER

Question 3:

From Section 8.12 Problem 5 in "Modern Data Science with R Exercise, 2nd edition": A reporter carried out a clinical trial of chocolate where a small number of overweight subjects who had received medical clearance were randomized to either eat dark chocolate or not to eat dark chocolate. They were followed for a period and their change in weight was recorded from baseline until the end of the study. More than a dozen outcomes were recorded and one proved to be significantly different in the treatment group than the outcome. This study was publicized and received coverage from a number of magazines and television programs. Outline the ethical considerations that arise in this situation.

REPLACE THIS TEXT WITH YOUR ANSWER