Lab 7: Color Interference

Chapter 7: Comparing a Numerical Variable across Two Groups

Team Captain: NAME HERE Facilitator: NAME HERE
Recorder: NAME HERE Resource Manager: NAME HERE

Invalid Date

The abstract of the article "Studies of interference in serial verbal reactions" in the *Journal of Experimental Psychology* reads:

> In this study pairs of conflicting stimuli, both being inherent aspects of the same symbols, were presented simultaneously (a name of one color printed in the ink of another color—a word stimulus and a color stimulus). The difference in time for reading the words printed in colors and the same words printed in black is the measure of interference of color stimuli upon reading words. The interference of conflicting color stimuli upon the time for reading 100 words (each word naming a color unlike the ink-color of its print) caused an increase of 2.3 seconds or 5.6% over the normal time for reading the same words printed in black.

The article reports on the results of a study in which seventy college undergraduates were given forms with 100 names of colors written in black ink, and the same 100 names of colors written in another color (i.e., the word purple written in green ink). The total time (in seconds) for reading the 100 words printed in black, and the total time (in seconds) for reading the 100 words printed in different colors were recorded for each subject. The order in which the forms (black or color) were given was randomized to the subjects.

The data below was simulated to reflect the results found in the paper.

```
color <- read.csv("data/interference.csv")
head(color)
```

```
  Subj DiffCol Black
1    1   39.64 37.53
2    2   44.26 42.02
3    3   33.84 31.99
4    4   53.88 39.48
```

```
5     5   38.94 43.09
6     6   48.77 36.00
```

Research Question Does printing the name of colors in a different color increase the time it takes to read the words?

1. What is the observational unit?
2. Should these observations be considered paired or independent? Explain your answer.
3. Based on your answer to question 2, is the appropriate observed summary statistic to be used to analyze these data the difference in mean times or the mean difference in times? Explain your answer.
4. Write your parameter of interest.
5. Write out the null and alternative hypotheses (in words), in the context of this study.
6. Create two histograms of the following:

- the distribution of reading times for words printed in black `Black`
- the distribution of reading times for words printed in a different color `DiffCol`

```
## Code for histogram 1: Reading times printed in black
ggplot(data = color,
       mapping = aes(x = ________)) +
  geom_histogram(binwidth = _____) +
  labs(title = "_______________",
       x = "_______________",
       y = "_______________")

## Code for histogram 2: Reading times printed in a different color
ggplot(data = color,
       mapping = aes(x = ________)) +
  geom_histogram(binwidth = _____) +
  labs(title = "_______________",
       x = "_______________",
       y = "_______________")
```

```
Error: <text>:4:27: unexpected input
3: ggplot(data = color,
4:        mapping = aes(x = __
                            ^
```

If you look at the data, you will notice there are three columns: `Subj`, `DiffCol`, and `Black`. There is not, however a column with the differences between the different color read time (`DiffCol`) and the black read time (`Black`). Let's make a new column with the differences!

```
color <- color |>
  mutate(Difference = DiffCol - Black)
head(color)
```

```
  Subj DiffCol Black Difference
1    1   39.64 37.53       2.11
2    2   44.26 42.02       2.24
3    3   33.84 31.99       1.85
4    4   53.88 39.48      14.40
5    5   38.94 43.09      -4.15
6    6   48.77 36.00      12.77
```

7. Create a histogram to visualize the difference in read time between the different colored text and the black text.

```
ggplot(data = color,
       mapping = aes(x = ________)) +
  geom_histogram(binwidth = _____) +
  labs(title = "_______________",
       x = "_______________",
       y = "_______________")
```

```
Error: <text>:2:27: unexpected input
1: ggplot(data = color,
2:        mapping = aes(x = __
                            ^
```

8. Use the `favstats()` function to obtain the summary of the observed data. *Be careful about what you are summarizing!*

```
# favstats() code goes here
```

9. What is the observed mean difference in reading times? Assign a symbol to this value.

10 Check the normality assumption for using the t-test to investigate the research question.

11. Calculate the t test-statistic. *Hint: (Feel free to use the code block below as a calculator)*

12. Fill in the code with these necessary pieces to conduct the paired t-test.

```r
t_test(color,
       response = __________,
       mu = ____,
       alternative = "________",
       conf.int = TRUE,
       conf.level = 0.99
      )
```

```
Error: <text>:2:20: unexpected input
1: t_test(color,
2:        response = __
                     ^
```

13. What is the p-value for this hypothesis test? With an $alpha = 0.01$ what decision would you make?
14. What does the p-value you obtained mean, in the context of the study? *Hint: it is the probability of what…assuming what?*
15. Report the 99% confidence interval for the mean difference.
16. Interpret the confidence interval in context of the study.
17. The abstract states, that the conflicting color stimuli "*caused* an increase of 2.3 seconds or 5.6% over the normal time for reading the same words printed in black." Is this statement valid? Explain.