Math 338 Midterm 2 Lab Portion

November 20/21, 2019

NAME: JARED DYRESON

You have 50 minutes to complete this exam and an extra 10 minutes to Knit this file (click the button that says Knit) and upload the resulting pdf file to Titanium.

You may refer to your notes, your textbook, and any pre-existing online reference (eBook/Sapling, 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).

Problem 1

IQ scores are assumed to be normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$. Under these assumptions:

What is the probability that an individual person will have an IQ greater than 105? [1.5 pts]

```r
mu <- 100
value <- 105
sigma <- 15
pnorm(value, mean = mu, sd = sigma, lower.tail = TRUE)
```

```
## [1] 0.6305587
```

0.6305587 is the probability.

What is the probability that a randomly selected sample of 25 people will have a sample mean IQ between 95 and 105? [2.5 pts]

```r
bottom_iq <- 95
top_iq <- 105

mu <- 100
sigma <- 15

a <- pnorm(bottom_iq, mean = mu, sd = sigma)
b <- pnorm(top_iq, mean = mu, sd = sigma)
prob <- b - a
print(prob)
```

```
## [1] 0.2611173
```

0.2611173 is the probability.

Problem 2

Fazio, et al. (2016) investigated whether having students play an educational video game called Catch the Monster with Fractions could help them estimate the value of simple fractions. The monster.csv file on Titanium contains sample data from 26 fourth- and fifth-grade students. There are three variables:

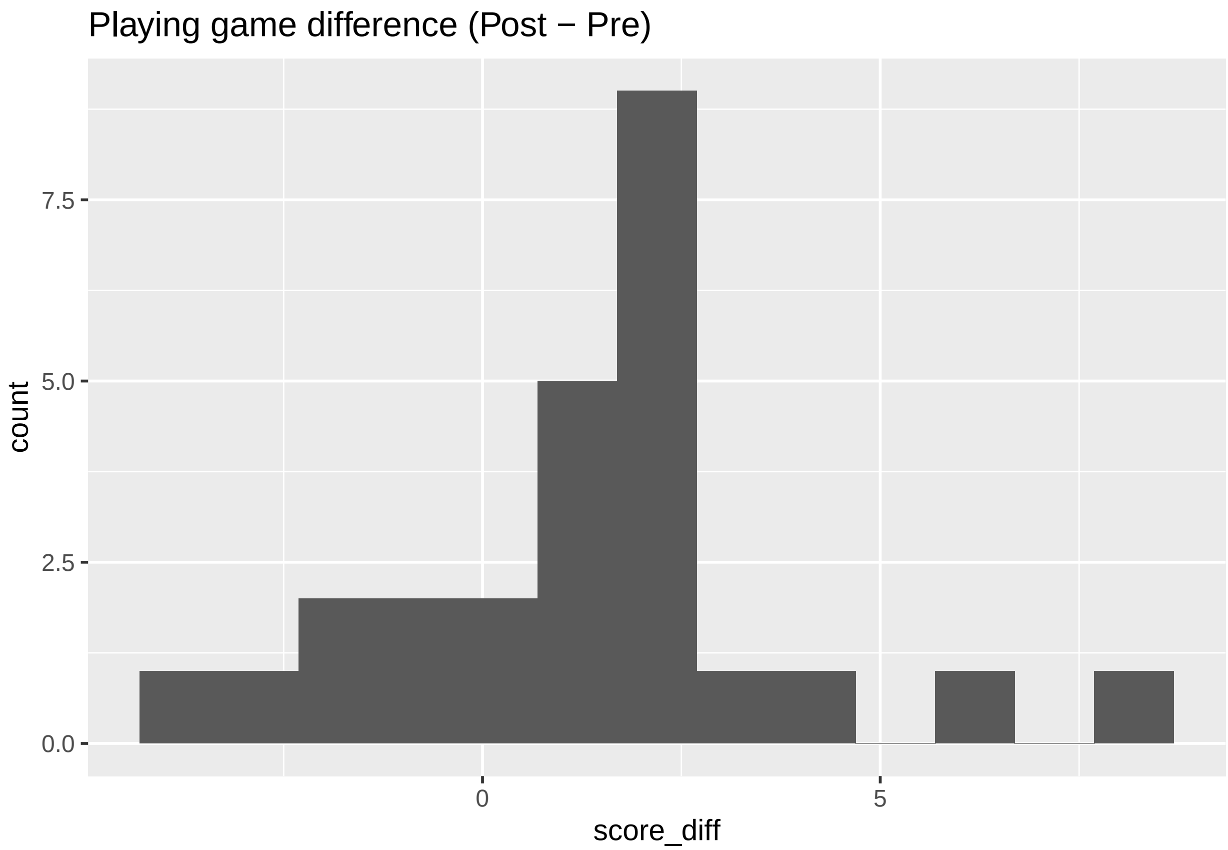
- Subject.ID: a unique three-digit number used to anonymize the student
- Pre: the score (number of problems correct) on a 15-question test about estimating fractions, taken before the student played the game
- Post: the score (number of problems correct) on a 15-question test about estimating fractions, taken after the student played the game

Should this dataset be analyzed using one-sample t procedures, matched pairs t procedures, or two-sample t procedures? Justify your answer. [1.5 pts]

This should be a matched pairs test because we are comparing the difference wheteher or not this game will help increase a student's test scores.

Based on your answer to part (A), produce an appropriate graph or set of graphs to visualize the sample data. Make sure to properly label your axes. [2.5 pts]

```
library(ggplot2, warn.conflicts = FALSE)
library(dplyr, warn.conflicts = FALSE)
data <- read.csv("./monsters.csv")
mu <- 1.192
diff <- data %>% mutate(score_diff = Post - Pre)
ggplot(diff, aes(x = score_diff)) + geom_histogram(binwidth = 1, center = mu) + labs(title = "Playing ga
```



Perform appropriate statistical inference to help the researchers determine whether the game helps students estimate fractions. [3 pts]

If you use a confidence interval, use a 95% confidence interval. If you use a hypothesis test, use NHST with a significance level of 0.05. Full credit will be given for:

- Defining (an) appropriate parameter(s) to estimate/test about and, if applicable, its value under the null hypothesis
- Performing statistical inference correctly in software and including the output
- Reporting either the confidence interval or the value of test statistic and p-value (as necessary)
- Making an appropriate conclusion in the real-world context and correctly justifying it based on the results of your inference
- $H_0 \implies \mu_{\text{d}} = 0$
- $H_1 \implies \mu_{\text{d}} \neq 0$
- $\alpha = 0.05$

```
library(ggplot2, warn.conflicts = FALSE)
library(dplyr, warn.conflicts = FALSE)
data <- read.csv("./monsters.csv")
mu <- 1.192
diff <- data %>% mutate(score_diff = Post - Pre)
n <- 13
alpha <- 0.05
```

```r
standard_deviation <- sd(diff$score_diff)

# t-stat value
t.test(diff$score_diff, mu = mu, alternative = "t")
```

```
##
##  One Sample t-test
##
## data:  diff$score_diff
## t = 0.00061249, df = 25, p-value = 0.9995
## alternative hypothesis: true mean is not equal to 1.192
## 95 percent confidence interval:
##  0.1576753 2.2269401
## sample estimates:
## mean of x
##  1.192308
```

```r
# p value
t.test(diff$score_diff, mu = mu, alternative = "t")$p.value
```

```
## [1] 0.9995162
```

Confidence interval of (0.1576753, 2.2269401) and a t stat value of 0.00061249

We arrive to the conclusion that we fail to reject the null hypothesis. That means there is no difference in using the game to help children to get better test scores. This is because the p-value far exceeds the significance level.

Problem 3

The resting potential of a neuron is -60 mV; however, due to randomness in ion channel activity, there are slight fluctuations around this average resting potential. A group of researchers believes that a drug increases the resting potential of the population of neurons in a particular area of the brain; suppose a practically relevant increase is +2 mV (that is, up to -58 mV).

The researchers believe they will only be able to measure the resting potential of a sample of 5 neurons after administering the drug. Based on previous experiments, they believe the sample standard deviation of their measurements will be around 1.65 mV.

Using a one-sample, right-tailed t-test with a Type I Error rate of 5%, is a sample of 5 neurons large enough to detect that 2 mV increase in resting potential? [3 pts]

```r
alpha <- 0.05
standard_deviation <- 1.65
n <- 5
init <- 60
final <- 58

power.t.test(n = n, delta = abs(init - final), sd = standard_deviation, sig.level = alpha, type = "one.
```

```
## [1] 0.5379911
```

Since we did not break the 80 percent minimum threshold, a sample size of five neurons is not enough to determine the increase.