Psych 315, Winter 2021, Homework 7

Due Friday, February 19th by midnight.

Name _____ ID _____

Section [AA] (Natalie), [AB] (Natalie), [AC] (Ryan), [AD] (Ryan), [AE] (Kelly), [AE] (Kelly)

Problem 1

According to the CDC, the average height of US women increased from 62 to 63 inches from 1964 to 2015. Let's assume that the population of heights for both years are normally distributed with a standard deviation of 3 inches, so we'll be working with z-distributions.

This first problem is about the power associated with detecting this 1 inch increase in mean height. We'll be filling in the probabilities in the following table:

Probabilities ($\alpha = 0.05$)		
	H_0 True	H_0 False
Fail to reject H_0		
Reject H_0		

In order to determine if women are significantly taller in 2015 than they were in 1964 we will let the null hypothesis be that heights are coming from the 1964 distribution (mean of 62 inches and a standard deviation of 3 inches).

a) Suppose the null hypothesis is true and we're drawing means from the 1964 population. Let $\alpha = 0.05$. What is the probability of rejecting H_0? We know the answer to this because the probability of rejecting H_0 when it is true is defined to be α. Since we will either reject or fail to reject H_0, the probability of failing to reject H_0 when H_0 is true is $1 - \alpha$. Fill in those numbers in the table above.

b) Suppose you draw a mean from a sample size of 16 from the 1964 distribution of heights. What is the standard error of the mean? Remember, each height comes from a normal distribution that has a standard deviation of $\sigma_x = 3$ inches, so the standard error of the mean is $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$, and n is the sample size.

c) Any mean drawn the null hypothesis distribution (1964 population) can be converted to a z-score by subtracting the null hypothesis mean (62 inches) from this mean and dividing by the standard error of the mean, calculated in part **b**.

What is the z-score for a mean of 63 inches, which is the mean of heights in 2015? That is, divide the difference the mean heights between 1964 and 2015 by the standard error of the mean from part **b**.

d) Next you'll be calculating power by drawing normal distributions and filling in areas, much like the tutorial on power. The bell-curve in the figure below is the standard normal, or z-distribution. This is the distribution of z-scores if the null hypothesis is true. Mark the z-score from problem **c** (The z-score for the mean of the 2015 population) on the z-axis below.
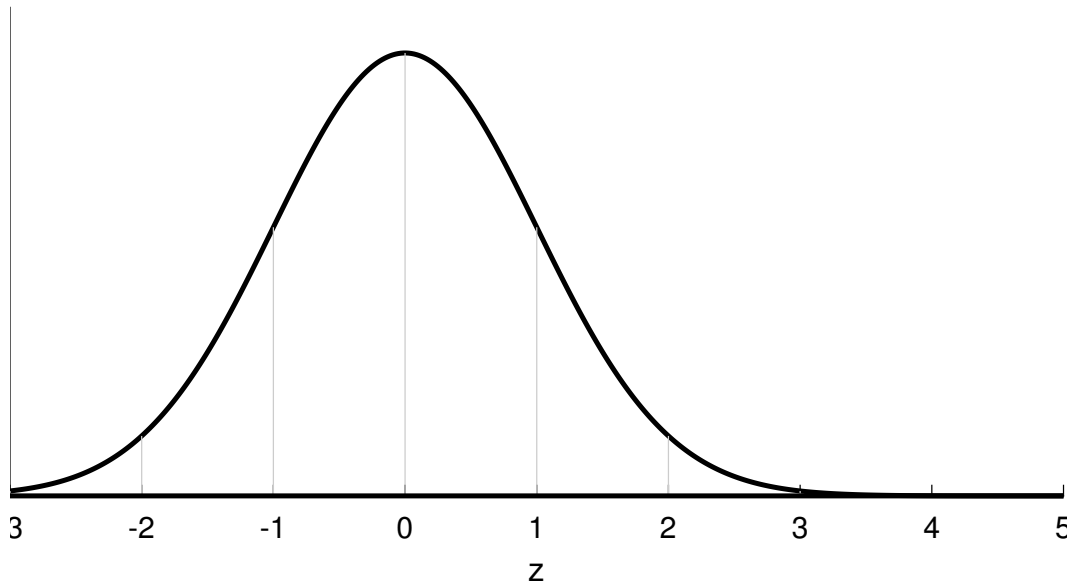
e) Find the z-score for the upper tail of this null hypothesis distribution which has an area of $\alpha = 0.05$. Shade the area under the bell curve above this z-score.

f) Suppose the you drew a mean from this null hypothesis distribution that fell in this shaded region. What would your decision be? Would this be a correct decision? If not, what type of error would it be?

g) Let's assume that H_0 is false and that the 'true' distribution of heights has a mean of 63 inches (from the 2015 distribution). Draw a normal distribution on the graph above that represents z-scores from this 'true' distribution of **mean** heights. This should be a normal distribution with mean z-score from part **c** and with a standard deviation of 1.

h) Take the rejection region that you found from problem **e** and lightly shade the area under this 'true' distribution.

i) If you were to draw a mean from this 'true' distribution which had a z-score that landed in the this shaded region, what would you decision be? Would this be a correct decision? If not, what type of error would we be making?

j) What is the area of this newly shaded region?

k) The area you just calculated in part **j** is the power of our test. It's the probability of rejecting H_0 when H_0 is false. That is, it is the probability of obtaining a mean from the 2015 distribution that is significantly greater than the null hypothesis mean of 62 inches.

Now fill in the other two numbers in the table from part **a** above.

l) Now, redo the math and calculate the power for the same means and sample size but using an α value of 0.01. Did power go up or down? Explain why. Here's a new table and graph for you to fill out. To get you started, I've drawn both the null hypothesis and the 'true' distributions. Note that z-score for the mean of the 'true' distribution is the same as before ($z = 1.3333$), since all we've changed is α.

Probabilities (alpha = 0.01)		
	H_0 True	H_0 False
Fail to reject H_0		
Reject H_0		

Z
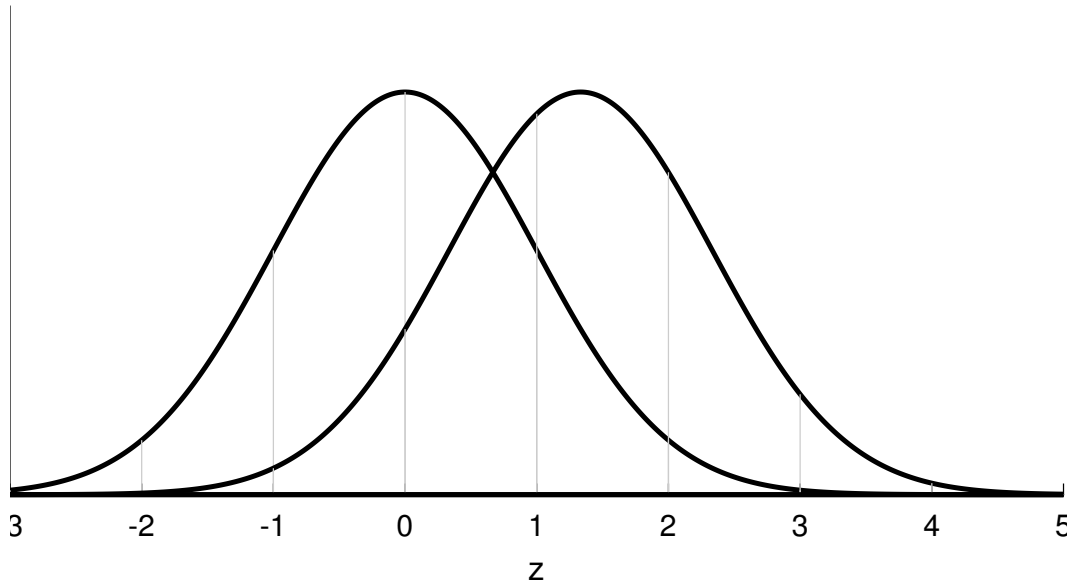
Problem 2 Suppose you think there might be a difference in self confidence between genders. In the survey, I asked what you thought your score would be on Exam 1. Splitting the class by gender, the estimated Exam 1 score for the 122 female students had a mean of 84.8 and a standard deviation of 8.2271. Exam 1 score predictions for the 29 male students had a mean of 86.9 and a standard deviation of 7.4082. Let's see if these two means are significantly different from each other. We'll use an α value of 0.05.

a) Calculate the pooled standard deviation:

$$s_p = \sqrt{\frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{(n_x-1)+(n_y-1)}}$$

b) Calculate the pooled standard error of the mean:

$$s_{\bar{x}-\bar{y}} = s_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}$$

c) Calculate the t-statistic:

$$t = \frac{\bar{x}-\bar{y}}{s_{\bar{x}-\bar{y}}}$$

**d)** Find the critical value of t

e) State your decision as a sentence in APA format.

**f)** What is the effect size, $d = \frac{|\bar{x} - \bar{y}|}{s_p}$

g) Suppose this were the true effect size. What would be the power of this test? (Use the power curves for $\alpha = 0.05$, two tailed, two means). Remember, each power curve corresponds to the average sample size for each mean (about 80 for this example).
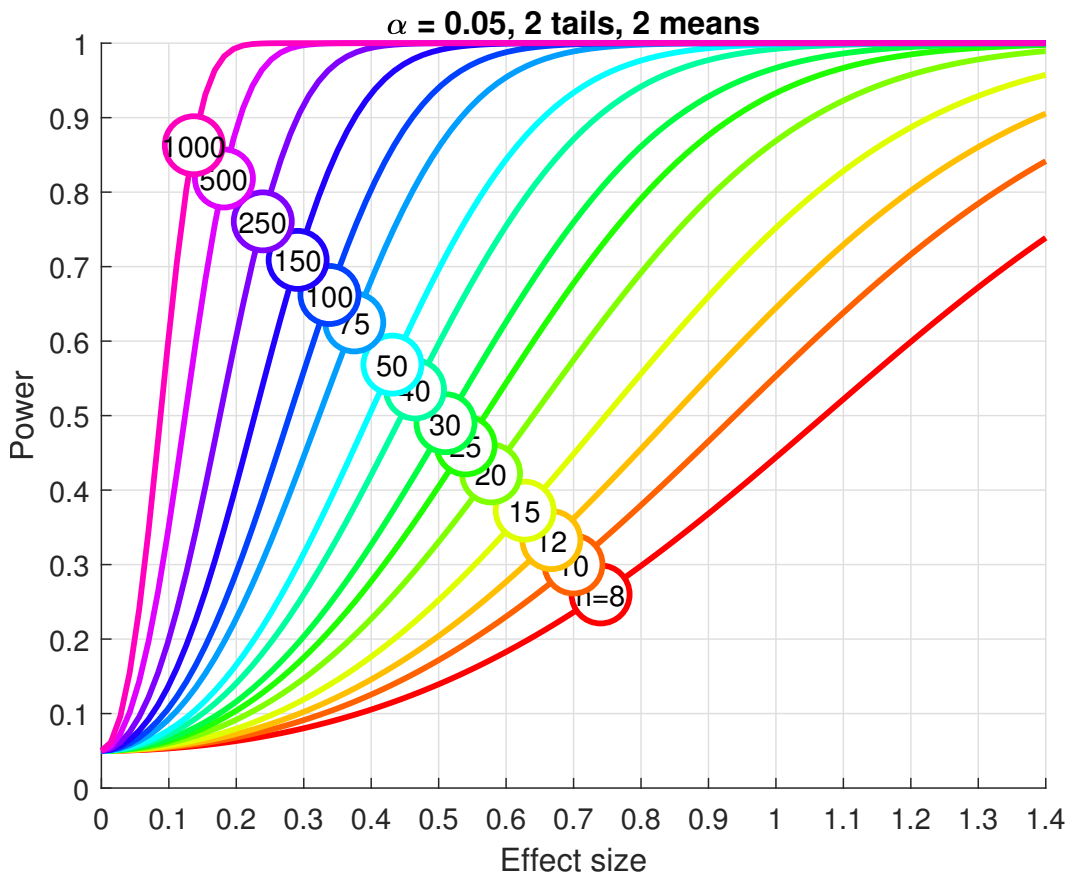
h) Use the power curve below to determine, for this effect size, how large the sample size for each group would need to be to get a power of 0.8

α = 0.05, 2 tails, 2 means
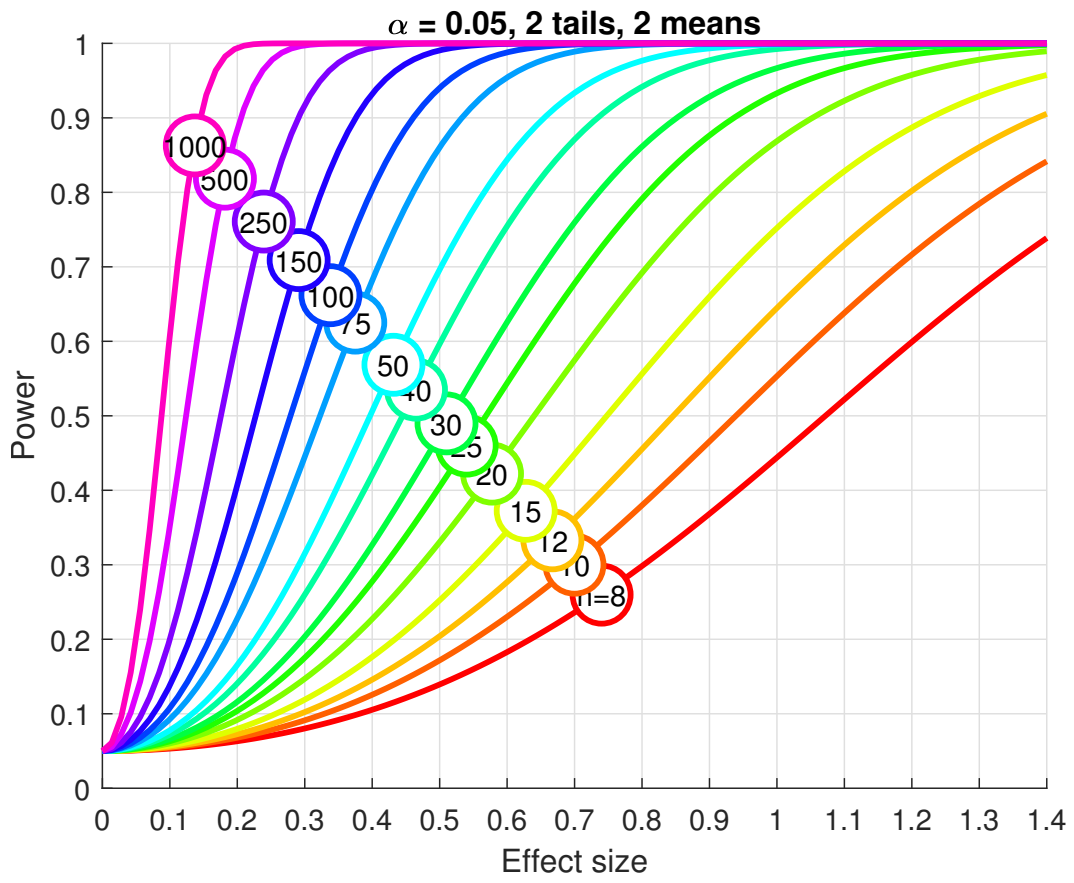
i) Draw a bar graph showing the two means with error bars representing ± one standard error of the mean. You'll have to calculate the standard errors of the mean by dividing each standard deviation by the square root of each sample size.
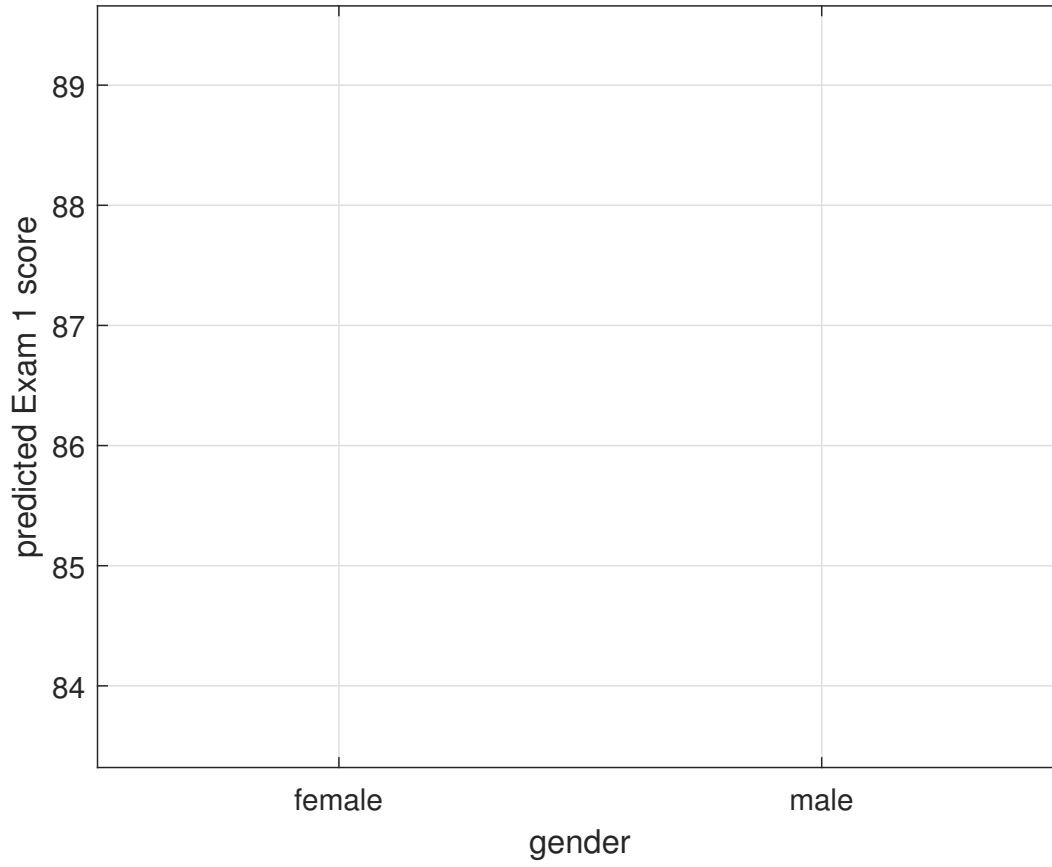
Problem 3 a) Use R to run the same test that you did in problem 2, comparing expected score in Exam 1 across gender. Use $\alpha = 0.05$.

Hint: The expected score in Exam 1 is called 'Exam1'. Here's how to test for a significant difference in prefered temperature across gender:

```
survey <-read.csv("http://www.courses.washington.edu/psy315/datasets/Psych315W21survey.csv")
x <- survey$temperature[survey$gender == "Male"]
y <- survey$temperature[survey$gender == "Female"]
x <- x[!is.na(x)]
y <- y[!is.na(y)]
out <- t.test(x,y,
alternative = "two.sided",
var.equal = TRUE)
out$p.value
[1] 0.954474
```

b) Use R to calculate the effect size. Here's how (using the values from the temperature example), given that you already have your 'x' and 'y' variables.

```
nx <- length(x)
ny <- length(y)
sp <- sqrt( ((nx-1)*sd(x)^2 + (ny-1)*sd(y)^2)/(nx-1+ny-1))
sp
[1] 7.198934
d <- abs(mean(x)-mean(y))/sp
d
[1] 0.01181398
```

c) Calculate the observed power of this test given the effect size that you calculated in part b, the fact that this is a two-tailed test, and the value of $\alpha = 0.05$. Here's how (with values from the temperature example):

```
# Find observed power from d, alpha and n
nx <- length(x)
ny <- length(y)
out <- power.t.test(n = (nx+ny)/2,
                    d = d,
                    sig.level = 0.05,
                    power = NULL,
                    alternative = "two.sided",
                    type = "two.sample")
out$power
[1] 0.02952346
```

d) Run these lines in R to generate a bar graph with error bars representing plus or minus one standard error of the mean. This example is for preferred temperature across gender, so be sure to change the labels to 'Expected Exam 1 Score'. You don't have to turn in a printout of the plot - I'll trust you that you tried running the code yourself.

```
library(ggplot2)
# Generate a 'data frame' containing statistics for x and y:
summary <- data.frame(
mean <- c(mean(x),mean(y)),
n <- c(length(x),length(y)),
sd <- c(sd(x),sd(y)))
summary$sem <- summary$sd/sqrt(summary$n)
colnames(summary) = c("mean","n","sd","sem")
row.names(summary) = c("Male","Female")
# This was a bit of work, but it creates a nice table:
summary
# Once you have this summary table, the rest will give you a nice looking
# bar plot with error bars:
# Define y limits for the bar graph based on means and sem's
ylimit <-  c(min(summary$mean-1.5*summary$sem),
             max(summary$mean+1.5*summary$sem))
# Plot bar graph with error bar as one standard error (standard error of the mean/SEM)
ggplot(summary, aes(x = row.names(summary), y = mean)) +
  xlab("students") +
  geom_bar(position = position_dodge(), stat="identity", fill="blue") +
  geom_errorbar(aes(ymin=mean-sem, ymax=mean+sem),width = .5) +
  theme_bw() +
  theme(panel.grid.major = element_blank()) +
  scale_y_continuous(name = "Preferred temperature") +
  coord_cartesian(ylim=ylimit)
```