

# Practice Session 9\_answers

## Question 1: Type I, Type II, or Neither?

- a) Type I Error
- b) No error
- c)

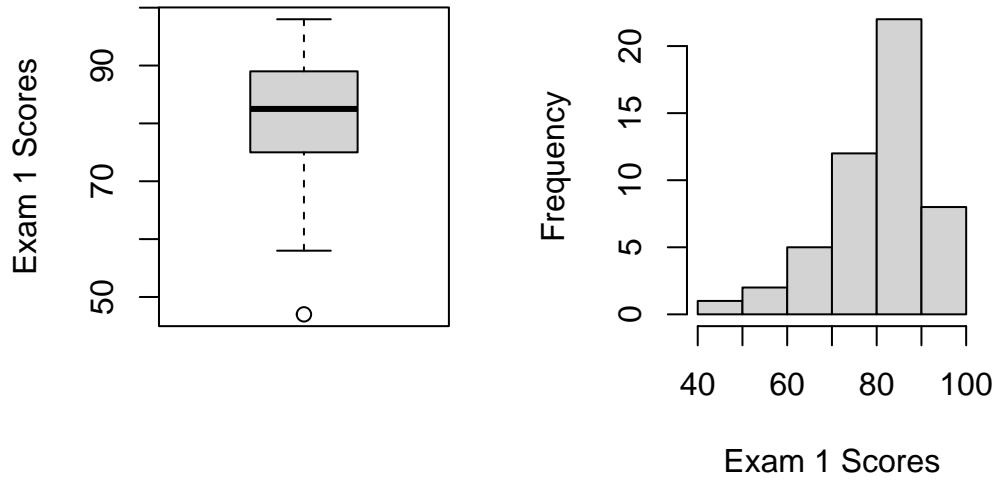
- No error
- Type I Error
- Type II Error
- No error

If the disease is dangerous, then a Type II error (saying a patient doesn't have the disease when they really do) can be a big issue. We would want to make the Type II error lower. However, this may make the Type I error larger as a result.

## Question 2: Single Sample T Test

- a)

```
library(Lock5Data)
par(mfrow = c(1, 2))
boxplot(StatGrades$Exam1, main = "", ylab = "Exam 1 Scores")
hist(StatGrades$Exam1, main = "", xlab = "Exam 1 Scores")
```



There appears to be one outlier. We might want to investigate this score further.

b)

- $H_0 : \mu = 80$
- $H_A : \mu \neq 80$

c)

```
n1 = length(StatGrades$Exam1) # get sample size (n = 50)

mean_exam1 = mean(StatGrades$Exam1) # compute sample mean
sd_exam1 = sd(StatGrades$Exam1)
se_exam1 = sd_exam1 / sqrt(n1) # compute standard error

T_stat_exam1 = (mean_exam1 - 80) / se_exam1
T_stat_exam1
```

[1] 0.7145975

d)

```
# df = 50 - 1 = 49
pt(T_stat_exam1, df = n1 - 1, lower.tail = F)
```

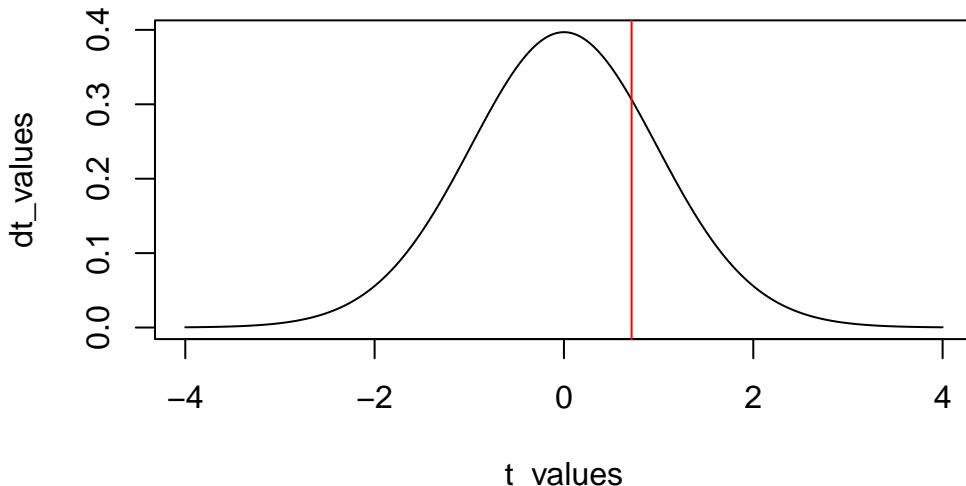
[1] 0.2391243

Since our p-value is greater than 0.05, we will fail to reject the null hypothesis. We therefore do not have evidence that this year's exam scores are significantly higher than 80.

e)

```
# Create sequence of t-values
t_values = seq(-4, 4, length.out = 1000)
dt_values = dt(t_values, df = 49) # convert to density values

# plot and make line plot with type = "l"
plot(t_values, dt_values, type = "l")
abline(v = T_stat_exam1, col = "red") # add vertical red line
```

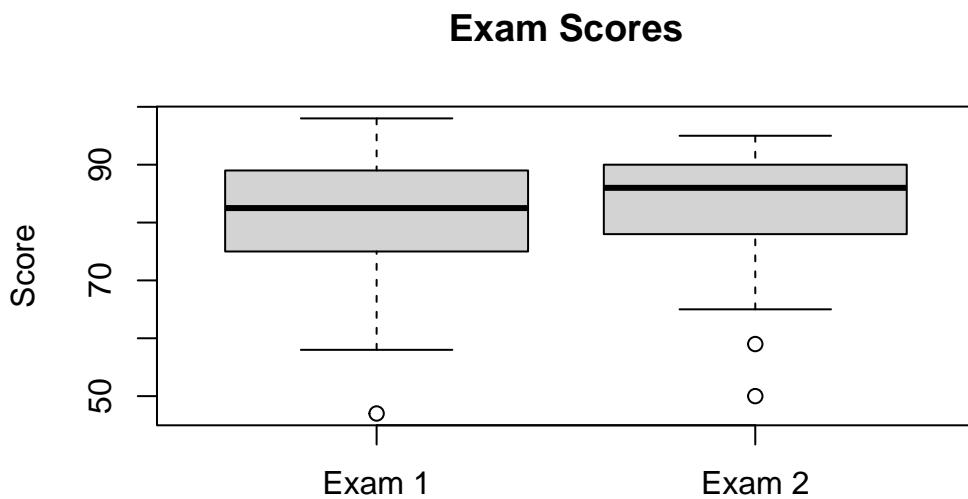


It seems that the sample mean wasn't very far from the null value of 80, which supports the lack of significance in our test.

### Question 3: Two Sample T Test

a)

```
boxplot(StatGrades$Exam1, StatGrades$Exam2,
        names = c("Exam 1", "Exam 2"),
        main = "Exam Scores",
        ylab = "Score")
```



The median scores appear to be similar for both exams. Exam 1 scores appear to have more variability than exam 2 scores. There are two outliers in the exam 2 scores.

b)

- $H_0 : \mu_1 = \mu_2$
- $H_A : \mu_1 \neq \mu_2$

c)

```

n2 = length(StatGrades$Exam2) # get sample size (n = 50)

mean_exam2 = mean(StatGrades$Exam2) # compute sample mean
sd_exam2 = sd(StatGrades$Exam2)

# Compute standard error for two means
se_two_means = sqrt(sd_exam1^2/n1 + sd_exam2^2/n2)
T_stat_two_means = (mean_exam2 - mean_exam1) / se_two_means
T_stat_two_means

```

[1] 1.026024

d)

Since we are doing a two-sided test, we need to multiply the p-value by 2.

```
2 * pt(T_stat_two_means, df = 50 - 1, lower.tail = F)
```

[1] 0.3099189

Since our p-value is greater than 0.05, we will fail to reject the null hypothesis. We therefore do not have evidence that this year's students performed differently on the two exams.

e)

- $H_0 : \mu_1 = \mu_{final}$
- $H_A : \mu_1 < \mu_{final}$

Note that we are doing a one-sided test, with the alternative being improvement on the final compared to the first exam.

```

n_final = length(StatGrades$Final) # get sample size (n = 50)

mean_final = mean(StatGrades$Final) # compute sample mean
sd_final = sd(StatGrades$Final)

# Compute standard error for two means
se_one_final = sqrt(sd_exam1^2/n1 + sd_final^2/n_final)
T_stat_one_final = (mean_final - mean_exam1) / se_one_final
T_stat_one_final

```

```
[1] 2.264897
```

```
pt(T_stat_one_final, df = 50 - 1, lower.tail = F)
```

```
[1] 0.01398826
```

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have evidence that student scores improved on the final compared to the first exam.

## Question 4: Paired T Test

a)

Let  $D$  be the true difference between the exam 1 and exam 2 scores. We will perform the following two-sided test:

- $H_0 : D = 0$
- $H_A : D \neq 0$

b)

```
diff_exams = StatGrades$Exam2 - StatGrades$Exam1
```

c)

```
n_diff = length(diff_exams)  
  
mean_diff = mean(diff_exams)  
se_diff = sd(diff_exams) / sqrt(n_diff)
```

d)

```
T_stat_diff = mean_diff / se_diff
```

e)

```
2 * pt(T_stat_diff, df = n_diff - 1, lower.tail = F)
```

```
[1] 0.01411695
```

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have evidence that this year's students performed differently on the two exams.

## **Part 2 Permutations tests for one mean and difference between two means:**

### **Test for the difference between two means with permutations:**

The same local statistics instructor believes that students perform the same on the first exam (`Exam1`) and the second exam (`Exam2`). Again, the data is available in the `StatsGrades` data set from the `Lock5Data` library.

- 1.) Create a null distribution and create its histogram.
- 2.) Calculate the p-value

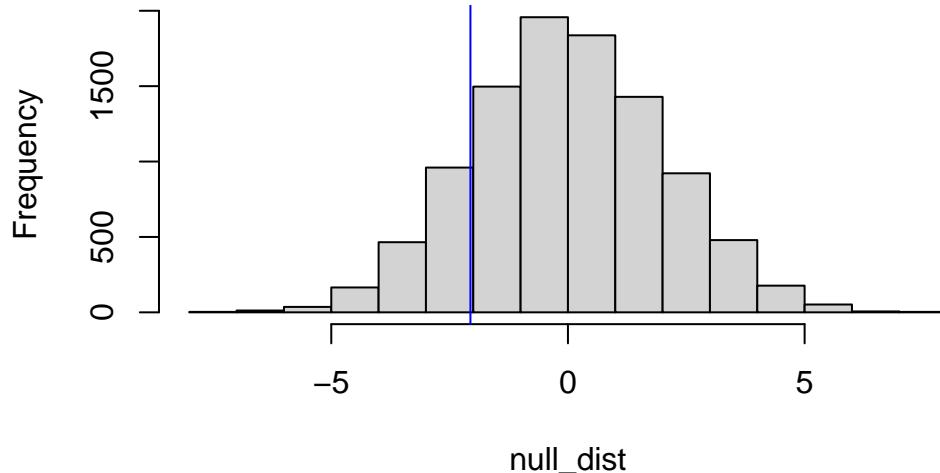
```
n=50
library(SDS1000)
combined_data <- c(StatGrades$Exam1, StatGrades$Exam2)
null_dist <- do_it (10000) * {
  shuff_data <- shuffle(combined_data)
  shuff_grp1 <- shuff_data[1:50]
  shuff_grp2 <- shuff_data[51:100]
  mean(shuff_grp2) - mean(shuff_grp1)
}
```

```
hist(null_dist)

obs_stat <- mean_exam1 -mean_exam2

abline(v = obs_stat, col = "blue")
```

### Histogram of null\_dist



4.) Calculate p-value

```
p_value <- pnull(obs_stat, null_dist, lower.tail = T)  
p_value
```

```
[1] 0.16
```

```
2*p_value
```

```
[1] 0.32
```

```
# 0.3072
```

#### Test for one mean:

(Let go back to he the question of testing one mean).

According to a local statistics instructor, the average score on his first exam of the semester (`Exam1`) is historically around 80%. The data is available in the `StatsGrades` data set from the `Lock5Data` library. let Create a bootstrap distribution with a mean of the hypothesised value

1.) Calculate the mean of the data.

```
mean_exam1 <- mean(StatGrades$Exam1) # compute sample mean
```

2.) Subtract this mean from the data.

```
subtract_mean <- StatGrades$Exam1 - mean_exam1
```

3.) Add the hypothesized value to the data.

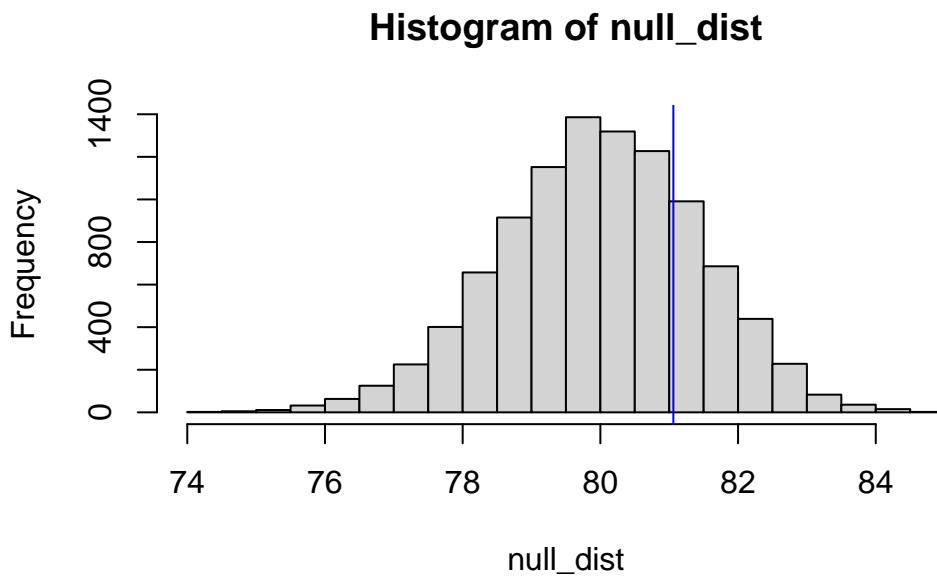
```
modified_data <- subtract_mean + 80
```

```
mean(modified_data )
```

[1] 80

4.) Now create a bootstrap distribution with the last modified data you have created in the previous step.

```
n=50  
my_sample <- modified_data  
  
boot_sample <- sample(my_sample, n, replace = TRUE) #sample with replacement  
null_dist <- do_it(10000) * {  
  curr_boot <- sample(my_sample, n, replace = TRUE)  
  mean(curr_boot)  
}  
  
hist( null_dist )  
  
obs_stat <- mean_exam1  
  
abline(v = obs_stat, col = "blue")
```



5.) Calculate the p-value

```
p_value <- pnull(obs_stat, null_dist, lower.tail = F )
p_value
```

```
[1] 0.2402
```

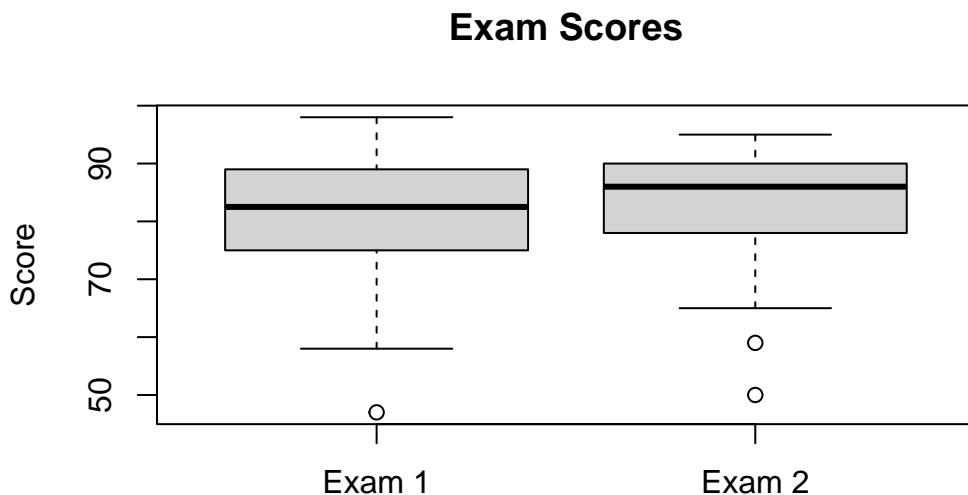
SDS1000/ClassMaterial/practice\_sessions/practice\_session\_09/practice\_09\_answers.qmd  
at main · emeyers/SDS1000

It seems that the sample mean wasn't very far from the null value of 80, which supports the lack of significance in our test.

## Two Sample T Test

a)

```
boxplot(StatGrades$Exam1, StatGrades$Exam2,
        names = c("Exam 1", "Exam 2"),
        main = "Exam Scores",
        ylab = "Score")
```



The median scores appear to be similar for both exams. Exam 1 scores appear to have more variability than exam 2 scores. There are two outliers in the exam 2 scores.

b)

- $H_0 : \mu_1 = \mu_2$
- $H_A : \mu_1 \neq \mu_2$

```
n2 = length(StatGrades$Exam2) # get sample size (n = 50)

mean_exam2 = mean(StatGrades$Exam2) # compute sample mean
sd_exam2 = sd(StatGrades$Exam2)

# Compute standard error for two means
se_two_means = sqrt(sd_exam1^2/n1 + sd_exam2^2/n2)
T_stat_two_means = (mean_exam2 - mean_exam1) / se_two_means
T_stat_two_means
```

[1] 1.026024

Since we are doing a two-sided test, we need to multiply the p-value by 2.

```
2 * pt(T_stat_two_means, df = 50 - 1, lower.tail = F)
```

```
[1] 0.3099189
```

Since our p-value is greater than 0.05, we will fail to reject the null hypothesis. We therefore do not have evidence that this year's students performed differently on the two exams.

c)

- $H_0 : \mu_1 = \mu_{final}$
- $H_A : \mu_1 < \mu_{final}$

Note that we are doing a one-sided test, with the alternative being improvement on the final compared to the first exam.

```
n_final = length(StatGrades$Final) # get sample size (n = 50)

mean_final = mean(StatGrades$Final) # compute sample mean
sd_final = sd(StatGrades$Final)

# Compute standard error for two means
se_one_final = sqrt(sd_exam1^2/n1 + sd_final^2/n_final)
T_stat_one_final = (mean_final - mean_exam1) / se_one_final
T_stat_one_final
```

```
[1] 2.264897
```

```
pt(T_stat_one_final, df = 50 - 1, lower.tail = F)
```

```
[1] 0.01398826
```

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have evidence that student scores improved on the final compared to the first exam.

## Paired T Test

Let  $D$  be the true difference between the exam 1 and exam 2 scores. We will perform the following two-sided test:

- $H_0 : D = 0$
- $H_A : D \neq 0$

```
diff_exams = StatGrades$Exam2 - StatGrades$Exam1
n_diff = length(diff_exams)

mean_diff = mean(diff_exams)
se_diff = sd(diff_exams) / sqrt(n_diff)

T_stat_diff = mean_diff / se_diff

2 * pt(T_stat_diff, df = n_diff - 1, lower.tail = F)
```

```
[1] 0.01411695
```

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have evidence that this year's students performed differently on the two exams.

## Part 2 Permutations tests for one mean and difference between two means:

### Test for the difference between two means with permutations:

The same local statistics instructor believes that students perform the same on the first exam (`Exam1`) and the second exam (`Exam2`). Again, the data is available in the `StatsGrades` data set from the `Lock5Data` library.

- 1.) Create a null distribution and create its histogram.
- 2.) Calculate the p-value

```

n=50
library(SDS1000)
combined_data <- c(StatGrades$Exam1, StatGrades$Exam2)
null_dist <- do_it (10000) * {
  shuff_data <- shuffle(combined_data)
  shuff_grp1 <- shuff_data[1:50]
  shuff_grp2 <- shuff_data[51:100]
  mean(shuff_grp2) - mean(shuff_grp1)
}

```

```

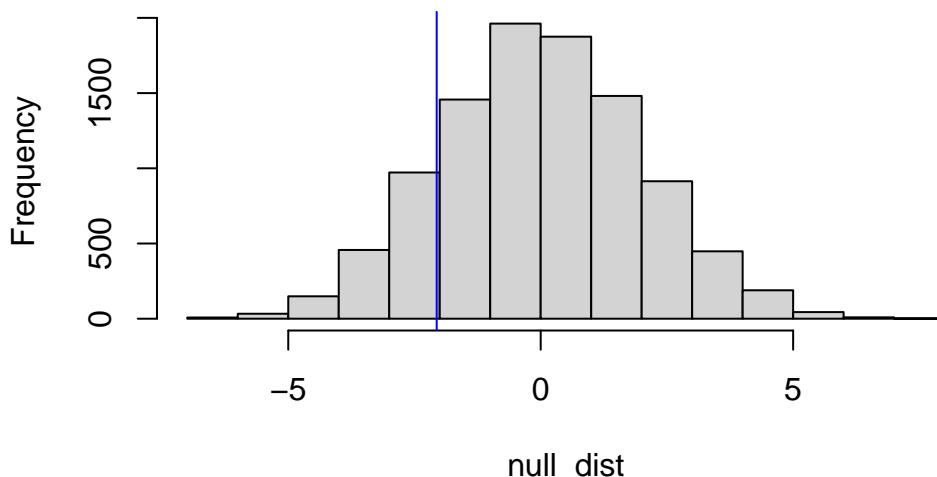
hist(null_dist)

obs_stat <- mean_exam1 -mean_exam2

abline(v = obs_stat, col = "blue")

```

**Histogram of null\_dist**



4.) Calculate p-value

```

p_value <- pnull(obs_stat, null_dist, lower.tail = T)
p_value

```

[1] 0.1559

```
2*p_value
```

```
[1] 0.3118
```

```
# 0.3072
```

### Test for one mean:

(Let go back to he the question of testing one mean).

According to a local statistics instructor, the average score on his first exam of the semester (`Exam1`) is historically around 80%. The data is available in the `StatsGrades` data set from the `Lock5Data` library. let Create a bootstrap distribution with a mean of the hypothesised value

1.) Calculate the mean of the data.

```
mean_exam1 <- mean(StatGrades$Exam1) # compute sample mean
```

2.) Subtract this mean from the data.

```
subtract_mean <- StatGrades$Exam1 - mean_exam1
```

3.) Add the hypothesized value to the data.

```
modified_data <- subtract_mean + 80  
mean(modified_data )
```

```
[1] 80
```

4.)Now create a bootstrap distribution with the last modified data you have created in the previous step.

```
n=50  
my_sample <- modified_data  
  
boot_sample <- sample(my_sample, n, replace = TRUE) #sample with replacement  
null_dist <- do_it(10000) * {  
curr_boot <- sample(my_sample, n, replace = TRUE)
```

```

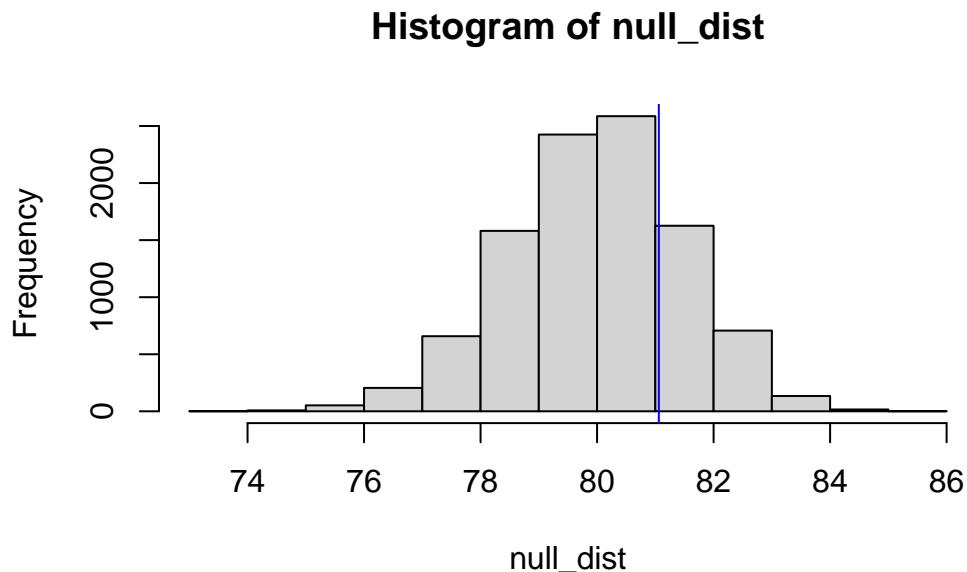
mean(curr_boot)
}

hist( null_dist )

obs_stat <- mean_exam1

abline(v = obs_stat, col = "blue")

```



5.) Calculate the p-value

```

p_value <- pnull(obs_stat, null_dist, lower.tail = F )
p_value

```

[1] 0.2399