Practice Session Answers 6

Part 1: Hypothesis Testing Terms and Definitions

Generally speaking, hypothesis testing is a way for researchers to answer questions in a meaningful way. The methodology leverages statistical reasoning and probability to either **reject** or **fail to reject** a statement about the population as a whole. Hypothesis testing in statistics often focuses on questions about population parameters (e.g., μ, π). We do not write hypotheses in terms of the sample statistics (e.g., $\bar{x}$, $\hat{p}$). This aligns with the general theme of the course, which focuses on using samples to better understand the population.

Practice 1:

Please match the following terms with the appropriate definition.

Terms:

1. Null Hypothesis
2. Alternative Hypothesis
3. P-value

Definitions:

a. The probability of observing results at least as extreme as what we observed, assuming the null hypothesis is true.

b. A statement of interest that represents the *status quo*, or that there is no effect or difference.

c. A statement of interest that proposes there is an effect or difference.

Answers

1. Null Hypothesis: A statement of interest that represents the *status quo*, or that there is no effect or difference.

2. Alternative Hypothesis: A statement of interest that proposes there is an effect or difference.

3. P-value: The probability of observing results at least as extreme as what we observed, assuming the null hypothesis is true.

Part 2: Hypothesis Testing for a Single Proportion

Practice 2.1 : Preferred Water Brand

A national sales bureau believes that 30% of consumers prefer Fiji water as their first choice for water. Specifically, they choose it over Aquafina, Sam's Choice, and tap water. The Fiji company claims that this proportion is greater than 30%. Using the `WaterTaste` data set from the Lock5Data library, run a hypothesis test to check this claim. Look at the `help` page for the data set to see which variable is appropriate to use.

```
library(SDS1000)
library(Lock5Data)
data(WaterTaste)
```

Answers

1. **First**: State the null and alternative hypothesis in words

- H_0: $\pi = 0.3$
- H_A: $\pi > 0.3$

2. **Second**: Calculate the observed statistic of interest

```
p_hat_fiji <- get_proportion(WaterTaste$First, "Fiji")
p_hat_fiji
```

```
Fiji
0.41
```

3. **Third**: Create a null distribution in R using `do_it()` and `rflip()`. Plot the null distribution, and add a vertical red line at your point estimate.

```
# Number of questions to simulate
n_water <- 100

# Simulate the null distribution
null_dist_fiji <- do_it(10000) * {
  rflip(n_water, 0.3) / n_water
}
# Another way to create the null distribution
# null_dist_fiji <- do_it(10000) * {
# rflip(n_water, 0.3, report_proportion = T)
#}

# Visualize the null distribution
hist(null_dist_fiji,  breaks = 50,
     main = "Null Distribution",  xlab = "Proportion of Consumers",
     col  = "forestgreen", border= "skyblue1")
abline(v = p_hat_fiji, col = "red")
```
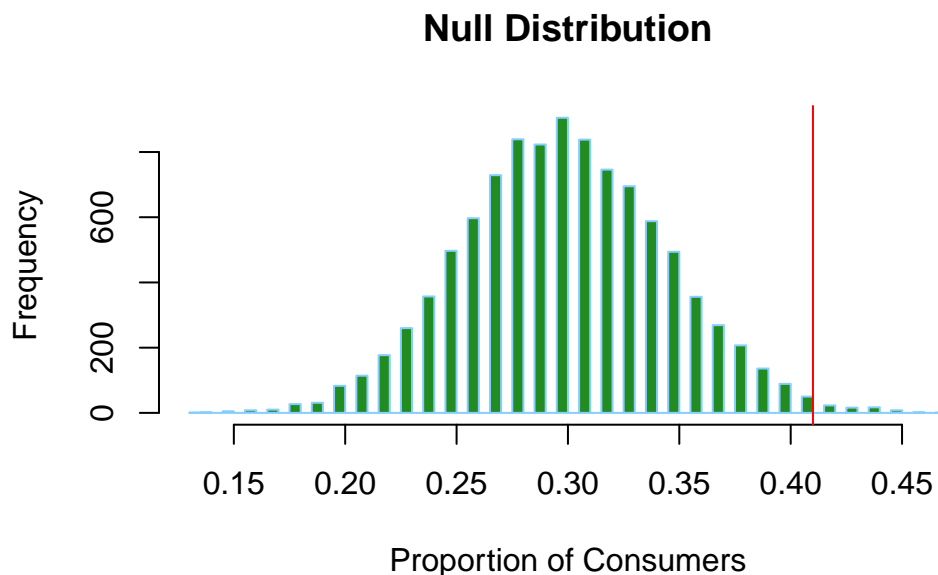
Null Distribution

Proportion of Consumers

4. **Fourth**: Calculate the p-value (probability of obtaining a result as or more extreme than what we observed) using the `pnull()` function.

```
one_sided_pv <- pnull(p_hat_fiji, null_dist_fiji, lower.tail = F)
one_sided_pv
```

```
[1] 0.0117
```

5. **Fifth**: Report the p-value. Make a decision on whether the results are statistically significant, and state your conclusion.

Since our p-value is less than the standard 0.05, we will reject the null hypothesis. We therefore have evidence that over 30% of consumers would select Fuji water as their first choice of water.

Practice 2.2: AP Multiple Choice

Does the answer choice "C" occur **less** frequently than expected? Answer this question with a hypothesis test. Use the `APMultipleChoice` data set from the `Lock5Data` library. Make sure to follow all 5 steps of hypothesis testing. *Hint:* when defining the null hypothesis, consider how likely each answer choice would be if answers were selected at random.

```
library(SDS1000)
library(Lock5Data)
data(APMultipleChoice)
```

Answers

1. **First**: State the null and alternative hypothesis in words

- $H_0 : \pi = 0.2$

- $H_A : \pi < 0.2$

2. **Second**: Calculate the observed statistic of interest

```
p_hat_ap <- get_proportion(APMultipleChoice$Answer, "C")
p_hat_ap
```

```
     C
0.1975
```

3. **Third**: Create a null distribution in R using `do_it()` and `rflip()`. Plot the null distribution, and add a vertical red line at the observed number of questions that had 'C' as the answer. *Note*: there were 400 multiple-choice questions in the data set.

```
# Number of questions to simulate
n_questions <- 400

# Simulate the null distribution
null_dist_ap <- do_it(10000) * {
  rflip(n_questions, 0.2) / n_questions
}

# Visualize the null distribution
hist(null_dist_ap, breaks = 50,
     main = "Null Distribution of Choice 'C' in 400 Questions",
     xlab = "Proportion of C's", ylab = "Frequencies")
abline(v = p_hat_ap, col = "red")
```

Null Distribution of Choice 'C' in 400 Questions
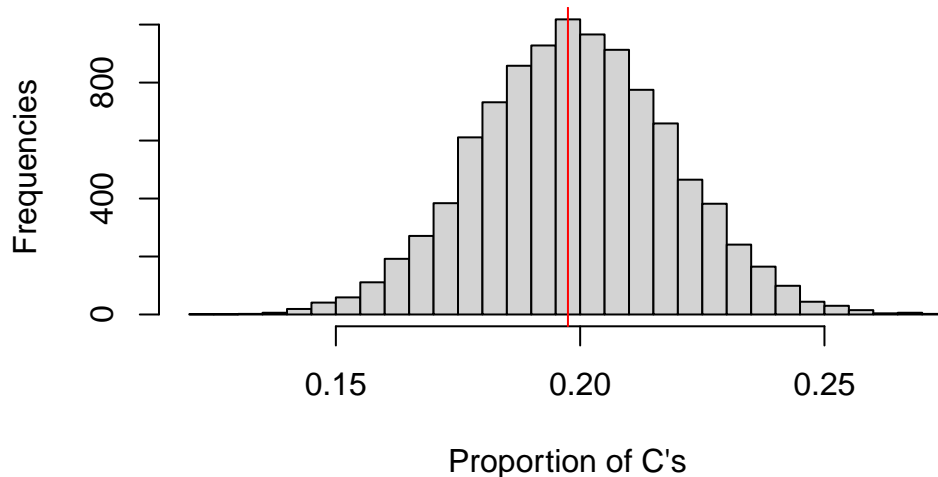
4. **Fourth**: Calculate the p-value (probability of obtaining a result as or more extreme than what we observed) using the `pnull()` function.

```
one_sided_pv <- pnull(p_hat_ap, null_dist_ap, lower.tail = T)
one_sided_pv
```

```
[1] 0.4719
```

5. **Fifth**: Report the p-value. Make a decision on whether the results are statistically significant, and state your conclusion.

Since our p-value is greater than the standard 0.05, we will fail to reject the null hypothesis. We therefore have evidence that the answer choice 'C' is occurring less than the expected.

Part 3: Hypothesis Testing for Difference in Two Means

Practice 3.1 : Caffeine Taps

A sample of college students were asked to tap their fingers at a rapid rate. The sample was then divided at random into two groups of ten students each. Each student drank the equivalent of about two cups of coffee, which included about 200 mg of caffeine for the students in one group but was decaffeinated coffee for the second group. After a two hour period, each student was tested to measure finger tapping rate (taps per minute). The goal of the experiment was to determine whether caffeine produces an increase in the average tap rate.

1. Write the `null hypothesis` and `alternative hypothesis` in words and in symbols.

- Null: There is no difference in the average taping rate between the caffeinated and decaffeinated groups

- Alternative: The average taping rate is greater for the caffeinated group than for the decaffeinated group

- H_0: $\mu_c - \mu_{nc} = 0$
- H_A: $\mu_c - \mu_{nc} > 0$

2. Create a boxplot to describe tap rates for `Caffeine` versus `No Caffeine`.

```
library(Lock5Data)
library(SDS1000)
data(CaffeineTaps)
boxplot(data = CaffeineTaps, Taps ~ Group)
```
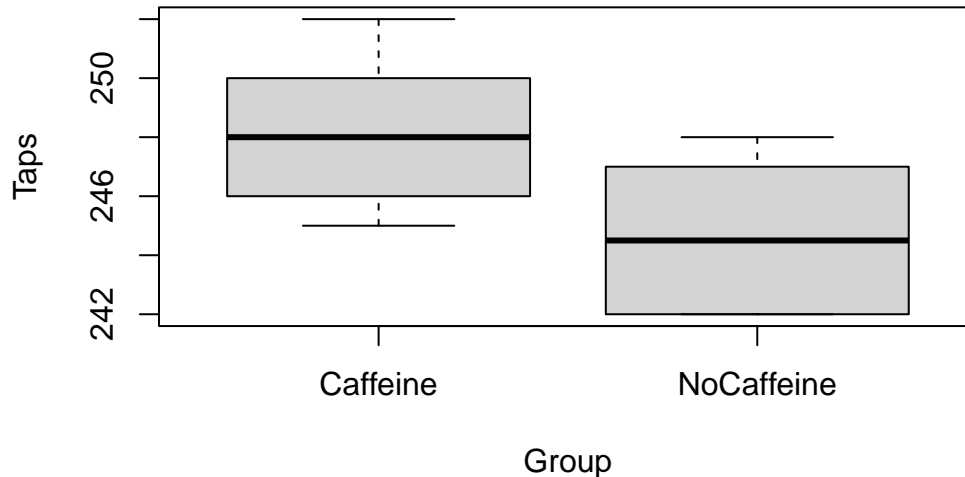

3. Find some favorite statistics to visualize the number of taps for the `Caffeine` and `No Caffeine` group. You might find the function: `mosaic::favstats` useful. *Note*: you can search online for this function arguments.

```
mosaic::favstats(Taps ~ Group, data = CaffeineTaps)
```

```
Registered S3 method overwritten by 'mosaic':
  method                                 from
  fortify.SpatialPolygonsDataFrame ggplot2

      Group min    Q1 median    Q3 max  mean       sd  n missing
1   Caffeine 245 246.5  248.0 250.00 252 248.3 2.213594 10       0
2 NoCaffeine 242 242.5  244.5 246.75 248 244.8 2.394438 10       0
```

4. Subset the data `CaffeineTaps` to two groups: `Caffeine` and `NoCaffeine`.

```
Caffeine <- subset(CaffeineTaps$Taps, CaffeineTaps$Group == "Caffeine")
NoCaffeine <-  subset(CaffeineTaps$Taps, CaffeineTaps$Group == "NoCaffeine")
```

5. Compute the observed statistic (mean difference of tap number for the two groups).

```
diff_caffeine <- mean(Caffeine) - mean(NoCaffeine)
```

6. Create null hypothesis distribution

a. Shuffle the two groups of `Caffeine` and `No Caffeine` into two samples, and find the mean difference of the two shuffled samples.

```
combined_sample <- c(Caffeine, NoCaffeine)
```

b. Create the Null hypothesis Distribution using `do_it()` function.

```
null_dist <- do_it(10000) * {
  shuffled_sample <- sample(combined_sample)

  shuff_caff <- shuffled_sample[1:10]
  shuff_nocaff <- shuffled_sample[11:20]

  shuff_stat <- mean(shuff_caff) - mean(shuff_nocaff)
}
```

c. Plot a `histogram` of the null distribution and show the `line` of the `observed mean difference` using the `abline()` function.

```
hist(null_dist ,
     xlab = "Difference in Mean Tap Rate", main = "Null distribution")
abline(v = diff_caffeine, col = "blue")
```

Null distribution
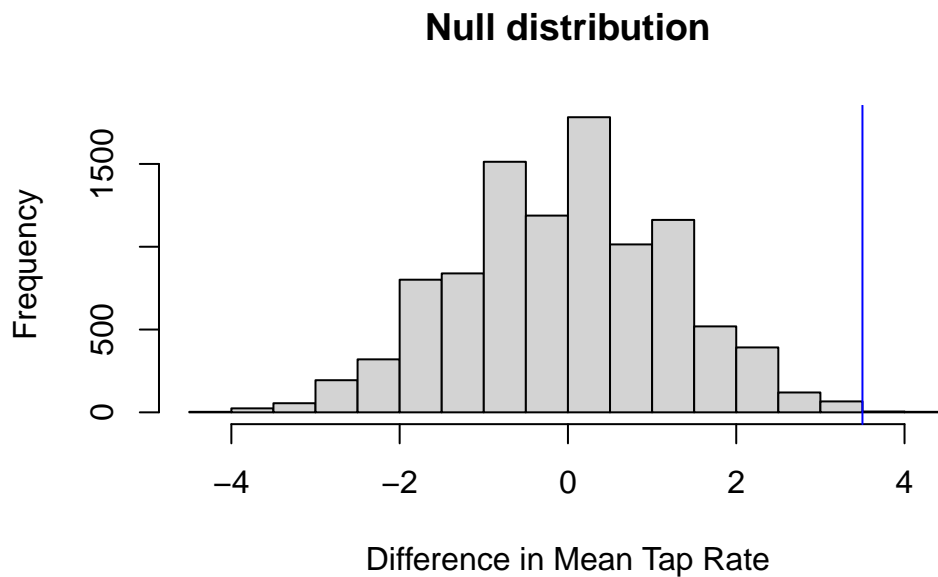
7. Calculate p-value

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

```
[1] 0.0019
```

8. Make a decision and state your conclusion:

Since our p-value is less than 0.05, we will reject the null hypothesis. We therefore have enough evidence that drinking caffeinated coffee produces an increase in the average tap rate compared to those that drink decaffeinated coffee.

Part 4: Extra Practices

Practice 4.1: Exploring the `pnull()` Function

In this exercise, we will practice using some basic R coding to calculate a p-value without using the `pnull()` function.

1. Consider null distribution and point estimate from the `AP Multiple Choice` exercise. Using the `<=` operator, write a single line of code that will show which values in the null distribution are **less than** the point estimate. Run this code and observe what the values look like. Save this result to a vector. *Hint:* don't over think this part. Just write the inequality as you would in math class.

```
true_false_total <- null_dist_ap <= p_hat_ap
```

2. Using your vector from above, count how many values from the null distribution are **less than** the point estimate. Use the `sum()` function for this. *Hint:* the `==` operator can be used within the `sum()` function to compare values.

```
sum_true <- sum(true_false_total == T)
```

3. Using your result from part 2.), find the proportion of observations in the null distributions that are **less than** the point estimate. Compare to the number from `pnull()` earlier.

```
prop_obs_lessthan_phat <- (sum_true) / 10000
prop_obs_lessthan_phat
```

```
[1] 0.4719
```

Practice 4.2: UFO Shapes

UFOs, or Unidentified Flying Objects have become commonplace in media and popular culture. Some witnesses claim that the objects have a cylindrical shape, while others believe they are more spherical. Load the data `ufo_data.csv` using the function `read.csv`.

Run a hypothesis test to check if the claim that UFOs are spherical is different from 0.75.

Answers:

1. State the null and alternative hypothesis in words:

- $H_0 : \pi = 0.75$
- $H_A : \pi \neq 0.75$

2. Calculate the observed statistic of interest

```r
library(SDS1000)
library(Lock5Data)

ufo_data <- read.csv("ufo_data.csv")

p_hat_sphere <- get_proportion(ufo_data$shape, "sphere")
p_hat_sphere
```

```
   sphere
0.8076462
```

3. Create a null distribution in R using `do_it()` and `rflip()`. Plot the null distribution, and add a vertical red line at your observed total number of spheres.

```r
# Number of questions to simulate
n_ufos <- 6670

# Simulate the null distribution
null_dist_sphere <- do_it(10000) * {
  rflip(n_ufos, 0.75) / n_ufos
}

# Visualize the null distribution
hist(null_dist_sphere, breaks = 50,
     main = "Null Distribution of Proportion of Spherical UFOs",
     xlab = "Proportion of Spheres", xlim = c(0.72, 0.82))
abline(v = p_hat_sphere, col = "red")
```

Null Distribution of Proportion of Spherical UFOs
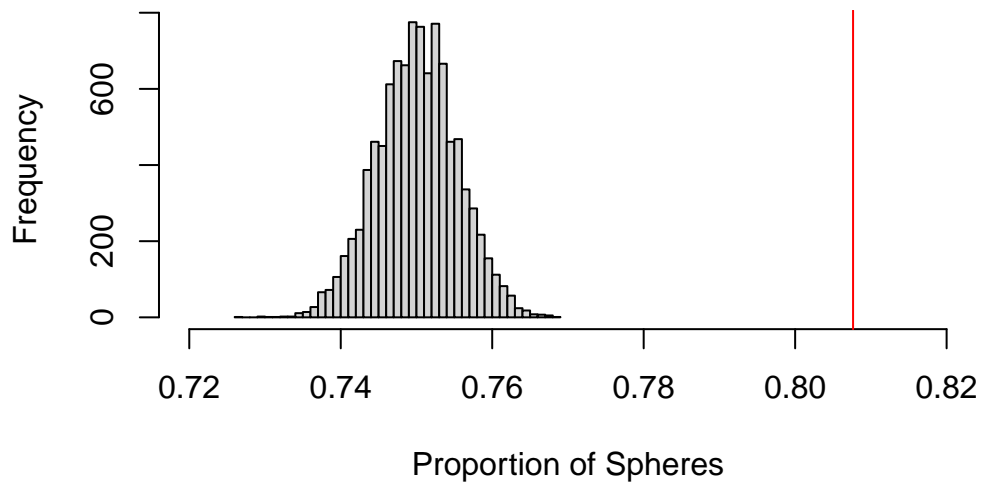
4. Calculate the p-value (probability of obtaining a result as or more extreme than what we observed) using the `pnull()` function.

```
# find the margin between the null_value and the obs_value
null_val <- 0.75
margin_null_obs <- abs(null_val- p_hat_sphere)
margin_null_obs
```

```
    sphere
0.05764618
```

```
# Find the upper tail of the p- value:
upper <- pnull(null_val+margin_null_obs , null_dist_sphere, lower.tail = F)
upper
```

```
[1] 0
```

```
# Find the lower tail  of the p- value:
lower <- pnull(null_val-margin_null_obs, null_dist_sphere, lower.tail = T)
lower
```

```
[1] 0
```

```
pv <- lower+ upper
pv
```

```
[1] 0
```

5. Report the p-value. Make a decision on whether the results are statistically significant, and state your conclusion.

Since our p-value is less than the standard 0.05, we will reject the null hypothesis. We therefore conclude that the data provides enough evidence that UFO sightings in spherical shape is different from 0.75 .