Math 300 Lesson 30 Notes

Conducting Hypothesis Tests

YOUR NAME HERE

July, 2022

Contents

Objectives

1. Correctly use terminology and notation of hypothesis testing.
2. Using the `infer` package, conduct a hypothesis test and interpret the results.
3. Compare and contrast confidence intervals and hypothesis tests.

Reading

Chapter 9.2 - 9.3

Lesson

Work through the learning checks LC 9.1 - LC 9.4.

- Focus on the terminology. There are many new terms.
- Try to write your own code for hypothesis tests and confidence intervals from scratch using the `infer` package.

Libraries

```
library(tidyverse)
library(infer)
library(moderndive)
library(nycflights13)
library(ggplot2movies)
```

LC 9.1 (Objective 2)

(LC 9.1) Why does the following code produce an error? In other words, what about the response and predictor variables make this not a possible computation with the `infer` package?

```
null_distribution_mean <- promotions %>%
  specify(formula = decision ~ gender, success = "promoted") %>%
  hypothesize(null = "independence") %>%
  generate(reps = 1000, type = "permute") %>%
  calculate(stat = "diff in means", order = c("male", "female"))
```

Solution:

LC 9.2 (Objective 1)

(LC 9.2) Why are we relatively confident that the distributions of the sample proportions will be good approximations of the population distributions of promotion proportions for the two genders?

Solution:

LC 9.3 (Objective 1)

(LC 9.3) Using the definition of *p-value*, write in words what the p-value represents for the hypothesis test comparing the promotion rates for males and females.

Solution:

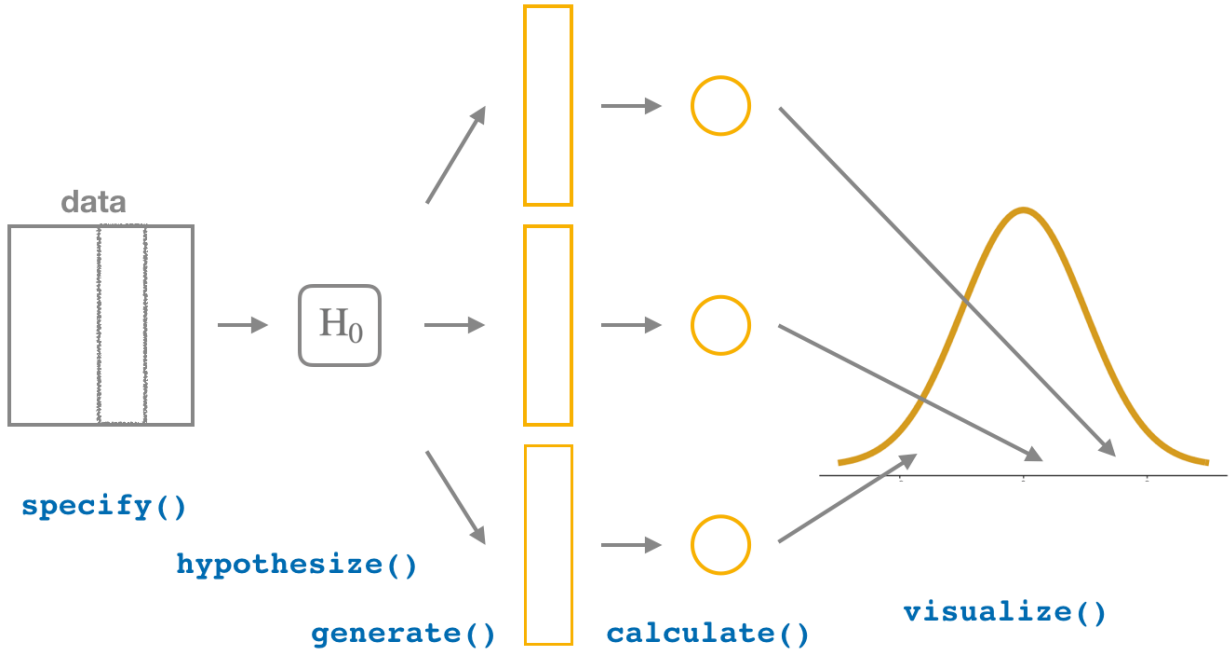
LC 9.4 (Objective 1)

(LC 9.4) Describe in a paragraph how we used Allen Downey's diagram to conclude if a statistical difference existed between the promotion rate of males and females using this study.

Solution:

Using `infer` (Objective 2)

Use `infer` to create a hypothesis and confidence interval for the promotion data set. Try to not use the book but using the following figure and the help menu. The data object is `promotions`.




```
head(promotions)
```

```
## # A tibble: 6 x 3
##      id decision gender
##   <int> <fct>    <fct>
## 1     1 promoted male
## 2     2 promoted male
## 3     3 promoted male
## 4     4 promoted male
## 5     5 promoted male
## 6     6 promoted male
```

- Find null distribtuion.
- Visualize the results.
- Use code to get observed value.
- Get p-value.

Find bootstrap percentile confidence interval.

- Get bootstrap distribution
- Get confidence interval

Documenting software

- File creation date: 2022-07-08
- R version 4.1.3 (2022-03-10)

- `tidyverse` package version: 1.3.1
- `moderndive` package version: 0.5.4
- `infer` package version: 1.0.2
- `nycflights13` package version: 1.0.2
- `ggplot2movies` package version: 0.0.1