

More Hypotheses to Test

Stat 145

For each scenario below.

1. Convert the question into a null and alternative hypothesis about a parameter or parameters. State this carefully using both words and symbols.
2. Check the data set to see if the information is there to do the test the way you had planned. If not, adjust your hypotheses accordingly. (You may have had a perfectly reasonable plan, but it didn't match the plan of the people conducting the study, you won't have data matching your plan.)
3. Compute a number or make graph (or both) from your data to get some sense for the data before you do your test. Make a ball park estimate for the p-value.
4. Use R to create the randomization distribution.
5. Use your randomization distribution to compute a p-value. (Are you doing a 1-tailed test or a 2-tailed test? Why?)

How did the p-value compare to your estimate?

6. Interpret your p-value.

What conclusion can we draw? Do you find the result surprising or confirming?

Cocaine Addiction

Question: Does taking Lithium help cocaine addicts avoid relapse?

Data: The data are summarized in the table below.

Treatment	Relapse	No Relapse
Lithium	18	6
Placebo	20	4

We can create a data set like this with the following R code (which is faster than typing it all into Excel or something like that).

```
CocaineStudy <-  
  bind_rows(  
    do(18) * tibble(treatment = "Lithium", result = "Relapse"),  
    do( 6) * tibble(treatment = "Lithium", result = "No Relapse"),  
    do(20) * tibble(treatment = "Placebo", result = "Relapse"),  
    do( 4) * tibble(treatment = "Placebo", result = "No Relapse")  
  )
```

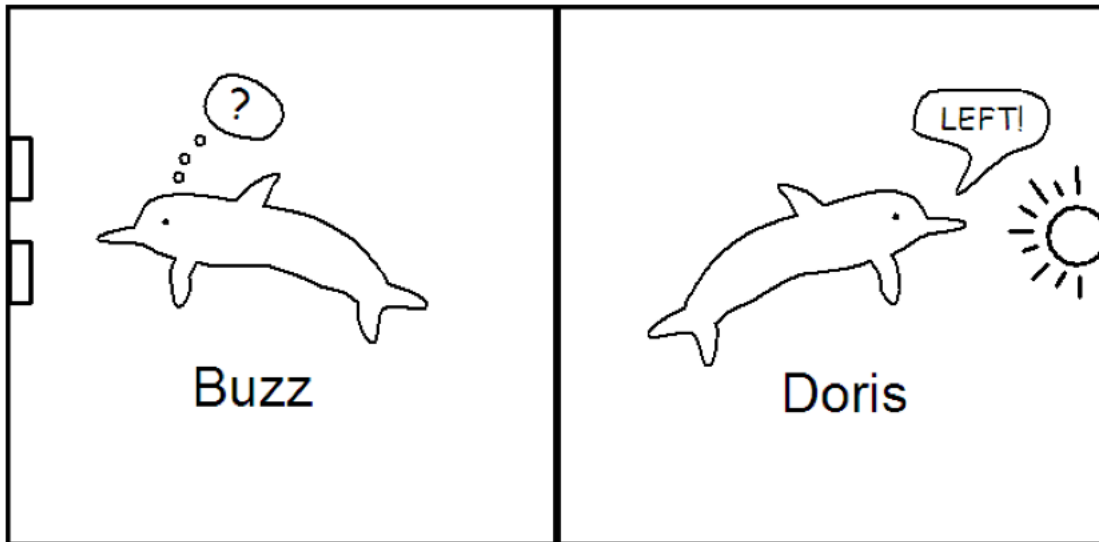
```
View(CocaineStudy) # Don't put this line into your R markdown!
```

Doris and Buzz

Question: Can dolphins communicate?

Setup: Two dolphins, Doris and Buzz, were trained to learn that they can get food by pushing one of two buttons depending on whether a light is on or off.

Later, Doris and Buzz were placed in the same tank, but separated by a curtain. The light was on the side with Doris and the buttons on the side with Buzz.



Data Set: Buzz pushed the correct button in 15 times in 16 attempts.

Smiles and Leniency

Question: Will college disciplinary panels be more lenient if you smile in the photo that is attached to your case paperwork?

Data: Smiles (in Lock5withR)

Pulse and Sex

Question: Does resting pulse in healthy young adults differ by sex?

Data: BodyTemp50 (in Lock5withR)

Doris and Buzz, Part 2

In a repetition of the Doris and Buzz experiment, a wooden separator was used instead of a curtain.

Data Set 2: Buzz pushed the correct button 16 times in 28 attempts.

Hugo

The German game *Mitternachtsparty* (Midnight party) has an important “character” named Hugo, a ghost who climbs the stairs out of the cellar and then chases players around a balcony while they try to duck into unoccupied rooms. Players roll dice and move the number rolled, unless they roll Hugo, in which case Hugo moves instead.



The first time I played this with my kids, I did some quick calculations about the average squares moved per turn for players and for Hugo, and used those calculations to optimize the placement of my pieces. (Rule number one at our house: Dad plays to win.) It didn't go so well for me. I was expecting to see Hugo rolled one time in six, but it seemed like we got way too many Hugos (and I got clobbered).

Curious about the die and my apparent bad luck, I rolled the die 50 times. Hugo came up 16 times. Is that enough evidence to be suspicious that the die is not fair?

Overview

1. One way to think about a hypothesis test is that we have to make a decision based on the data, much like a jury makes a decision in a court of law based on the evidence.
 - a. What are the two decisions a jury can make?
 - b. When we conduct a hypotheses test, what are the two decisions we can make? How do we decide which one to make?
2. When we conduct a hypothesis test, there are two types of mistakes we can make. What are they? (They have very boring names: Type I error and Type II error. Call me over to find out which is which.)
3. How can we reduce how often we make type I error? Why does this make it easier to make a type II error?
4. What is a “significance level”? What letter do we usually use for significance level?
5. What does the phrase “statistically significant mean”?
6. What are the 4 steps for conducting a hypothesis test?
7. We have studied three kinds of hypothesis tests. Every example we have seen has been one of these three kinds.
 - a. What are the three kinds
 - b. How do you create a randomization distribution for each kind?
8. What is the difference between a 1-tailed test and 2-tailed test? How do we decide which to do? [Note: 1-sided and 2-sided are other terms for the same thing.]
9. If you make a histogram of a null distribution, what features do you expect to see? Why?
10. What is a p-value and how do we use it to weigh the evidence provided by the data?