

Lab D: Hypothesis Testing and P-values

Toews, Math 160, Spring 2017

March 8, 2017

Objective:

Use R Studio to explore hypothesis testing, and learn how to construct randomization samples.

Self-assessment:

Include answers to the following two questions in your lab book. Write the answers after you have finished the other lab activities.

1. Explain how a randomization sample is different from a bootstrap sample.
2. Write a short paragraph in which you explain to a non-statistically minded friend what a P-value is, and what role P-values play in formal hypothesis testing.

Introduction

In this lab we investigate the following question: *do dogs and their owners tend to resemble one another?* We will see how to use data to test the claim that they do. Using data to draw a conclusion about a claim is called **statistical inference**. This lab walks you through the process of statistical inference.

Due by Friday, March 10:

1. Your <YourFirstName_PartnersFirstName>_labD.R script, in your Dropbox. (Eg. “bob_ann_labC.R”)
2. Your Lab Notebook.

Activities

Getting Organized:

Please make a “LabD” directory on your machine and copy the data file “DogOwner.Rda” into the directory.

Next, open RStudio, navigate to your LabD directory, and set this as your working directory. Open a new .R file called <YourFirstName_PartnersFirstName>_labD.R in which to enter your commands.

Getting to know the data

Load up the data via

```
load('DogOwner.Rda')
```

and peek at the data to see what it looks like.

Pause for reflection #1: *What is this data set about? Look both at the data itself, and also the description file. Briefly summarize what the data represents. How large is the data set?*

Let's rename the data to make it easier to access. We'll call it `data`:

```
data = DogOwner$Match
```

Let's imagine our wacky social-theorist friend believes that dogs and owners tend to resemble one another. Our task is to see to what extent the data supports this claim. Our basic procedure will be as follows:

1. write down null and alternative hypotheses
2. form a point estimate \hat{p} with the data
3. use what's called a "randomization sample" to see how unlikely our value of \hat{p} is if the null hypothesis is actually true.
4. depending on how unlikely it is, either reject the null hypothesis, or fail to reject it.

We'll take these steps one by one.

Writing down the null and alternative hypotheses

The first step is write down hypotheses.

Pause for reflection #2: *Write down an appropriate null and alternative hypothesis for our particular claim. Justify your choice.*

Forming the point estimate

The next step is to form a point estimate. Use the data to form a point estimate \hat{p} for the percentage of dog owners that were correctly paired with their dog. (Note: you'll need to look at the data and figure out how to do this. You can do this by hand, if you wish, but another solution is to use the `as.numeric` command to turn the vector of yes/no data into numerical data.)

Pause for reflection #3: *Write down what your \hat{p} is, and what you did to calculate it. What would you expect \hat{p} to be if the null hypothesis were true? Does your intuition tell you that this value of \hat{p} you got is likely or unlikely given the null hypothesis?*

Forming a randomization sample

We now need to calculate the probability of getting a \hat{p} "as or more extreme" than the one we got, assuming the null hypothesis is true. To do this, we'll do the following: first, assume the null is true, i.e. that there is "no connection" between how an owner looks and how his/her dog looks. In this case, the people tasked with assigning dogs to owners should have a 50/50 chance of getting it right, i.e. $p = .5$. Under this assumption, we'll simulate an experiment in which we ask a person to assign dogs to each of 25 owners, and calculate the \hat{p} for the simulation. We'll do this a bunch of times, and keep track of all our \hat{p} 's. At the end of the day, we'll ask what percentage of these statistics were "as extreme or more extreme" than the \hat{p} we got. We'll take this percentage to be our P -value, i.e. the probability of getting such an extreme result under the assumption that H_0 is true.

Pause for Reflection #3.5 *What does "as or more extreme" mean in the context of your alternative hypothesis? Write down the specific ranges of numbers that qualify as "as or more extreme". Does your range depend on whether your alternative is one-sided or two-sided?*

The following code performs the simulation described above. Read it carefully, and try to understand every line (discuss with your lab partner):

```

n_owners = 25 # there are 25 owners in this sample
n_simulations=100 # number of times we'll run the matching simulation
all_phats = rep(0,n_simulations) # a vector of 0s to store  $\hat{p}$  for each simulation
for (idx in seq(n_simulations)) # for each simulation...
{
  thissamp = sample(c(0,1),n_owners,replace=TRUE) # ...generate a random vector of 0's and
                                                    # 1's, 1 for "match", 0 for "no match"
  all_phats[idx] = sum(thissamp)/n_owners # calculate and store  $\hat{p}$ 
}

```

Pause for Reflection #4 Form a histogram of the results. Looking at the histogram, estimate what percentage of the data is “as extreme or more extreme” than the result you got. (In other words, estimate a P -value.)

Another way of estimating P -values is to sort the simulation statistics, and figure out exactly what percentage of these statistics is “as extreme or more extreme” than what the data is.

Pause for Reflection #5 Use the `sort` command in R to order the vector `all_phats` and count exactly how many are as or more extreme than the \hat{p} given by our data. Use this count to form a P -value. (How do you think you should do this?)

Drawing Inferences

Our final task is to draw an inference about our claim. Most people reject a null hypothesis if the P -value of the data is less than .05.

Pause for Reflection #6 Use your results above to draw an appropriate conclusion regarding the question “do dog owners resemble their pets.” Make sure to use the phrase “at the .05 significance level” in your answer.