

# Lab 3: Guppies

## Hypothesis Testing

Cori Carver

4-4-2021

### General information

This lab is due February 10th by 11:59 pm. **Upload the R Markdown document and the knitted PDF to Canvas.** The lab is worth 10 points (1 point per question unless otherwise noted). You are welcome and encouraged to talk with classmates and ask for help. However, each student should turn in their own lab assignment and all answers, including all code, needs to be solely your own.

#### Important: some notes on knitting

*Before you knit:* save your R markdown file as something other than Lab 3.Rmd; i.e. append your last name to it. Then you can knit your .Rmd file to a PDF and you won't get it confused with the version you've downloaded from me.

Your .Rmd will not knit if you have any errors in code. After you try to knit your document, a tab will appear in the console below. If you click on the R markdown tab it will tell you whether the knitting worked. If it failed, R usually gives an error referring to a specific line, and usually shows a snippet of the code that made the error. However, it will show only the first bug it encountered when reading through the document, so there may be additional errors. Thus it is an iterative process: knit, fix bug, knit again, fix next bug, and so on until you get your final PDF!

If you have errors you are not able to correct but you still want to knit, you have two options. (1) If you know what line of code produces the error, you can just add a `#` in front of it to comment the line out. (Don't delete the buggy code—that way I can give you feedback). (2) if all else fails, copy this code into the beginning of your .Rmd document, under the line that says `{r setup, include=FALSE}`: `knitr::opts_chunk$set(error = TRUE)`

Note that option 2 works in a pinch but isn't ideal because it doesn't force you to go through the process of iteratively fixing your errors. I suggest it as a Wednesday afternoon last resort.

### Objective

In this lab you will learn how to implement hypothesis testing in R by using the binomial test of proportions. You will also answer several conceptual questions about hypothesis testing.

#### The data: sexual selection and parasite infection in guppies

The guppy (*Poecilia reticulata*) is considered a model organism for studies of sexual selection because female guppies exhibit preference in their choice of male guppies. Females tend to choose males with certain display traits, which are thought to be indicators of fitness. One theory is that females are able to detect which males are resistant to certain parasites, and are more likely to choose those males. This question was investigated in a well-known study by C.E.J. Kennedy, J.A. Endler, S.L. Poynton, and H. McMinn, published in 1987 in

the journal Behavioral Ecology and Sociobiology. Our lab will be loosely based upon this study, which you can feel free (but are not required) to read. A copy is posted on Canvas.

## Experimental Design

In this hypothetical experiment (which I've simplified for the purposes of this lab) the researchers infected 30 male guppies with a non-lethal nematode parasite (*Camallanus cotti*). They also obtained 30 uninfected male guppies which underwent a de-worming treatment. From these male guppies, they randomly drew a pair of males for their experiment. The pair consisted of one randomly selected infected male and one randomly selected uninfected male.

Female guppies were drawn randomly from a population of wild guppies in Trinidad. For each trial, researchers placed one female guppy in a tank with a pair of male guppies. The female was able to swim beyond a partition to enter the side of the tank with either the uninfected male or the infected male. After 20 minutes, researchers recorded which side of the tank the female was on. If the female swam to the uninfected male, a 1 was recorded. If the female swam to the infected male, a 0 was recorded.

This was repeated 25 times, each time using a separate female.

Researchers used this data to investigate whether or not females displayed preference for males based on their infection status. In other words: do females choose uninfected males more often than expected by chance?

**Question 1:** What is the response (dependent) variable in this study? What kind of variable is it?

*Female choice in mate is the dependent variable, and this is a nominal variable.*

**Question 2:** (a) What is the null hypothesis the researchers are testing? (b) What is the alternative hypothesis? Hint: think about whether a one-tailed or two-tailed hypothesis test is appropriate here.

*(a) Because the researchers are testing the theory that female guppies preferentially mate with male guppies that are resistant to certain parasites, the null hypothesis would be that females do not use parasite load to preferentially choose mates. (b) The alternative hypothesis would be that females prefer to mate with uninfected males. This would be a one-tailed hypothesis because the researchers emphasize directionality in their hypothesis of testing this theory.*

## Experimental results

The researchers finished their experiment and their 25 observations yielded the following data: 1, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1.

We'll store this dataset as an object in R. Since there is only one column of data, we can use a vector rather than a dataframe:

```
guppy.trials <- c(1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1)
```

To keep consistent, from here on out let's refer to the event of a female choosing an uninfected male as a "success" since the event is coded as a 1.

**Question 3:** Calculate the estimated proportion of successes, or  $\hat{p}$  (p-hat). Do this using functions in R, such as `sum()` and `length()`. Write out your answer.

```
propSuccesses <- sum(guppy.trials) / length(guppy.trials)
```

*The proportion of successes is calculated by dividing the number of successes (denoted in code by `sum(guppy.trials)` because successes are 1's and failures are 0's) by the sample size (denoted `length()` of the `guppy.trials` vector).*

**Question 4:** Calculate standard error of  $\hat{p}$  (p-hat). The formula is in your lecture notes. Do the calculation in R and show your code below. Write out your answer

```
SEsuccesses <- sqrt((propSuccesses * (1-propSuccesses))/length(guppy.trials))
x <- sum(guppy.trials)
n <- length(guppy.trials)
SEsuccesses
```

```
## [1] 0.08979978
```

The standard error of  $p$ -hat is 0.08979978 because SE is equivalent to the square root of  $p$ -hat (`propSuccesses`) multiplied by  $(1 - p\text{-hat})$ , divided by the sample size (denoted by the length of `guppy.trials`).

**Question 5:** If the null hypothesis is true, what is the probability of observing the exact number of successes we observed? Hint: use the formula given in lecture (rather than R's `dbinom` function), and to save time, you may want to Google how to do a factorial (!) in R. Show your code. Write out your answer

```
nChooseX <- ((factorial(n)) / (factorial(x) * factorial(n - x)))
p <- 0.5
PrX <- nChooseX * (p^x) * (1-p)^(n-x)
PrX
```

```
## [1] 0.01432598
```

If the null hypothesis is true and female mate preference has no association with male infection status, then  $p = 0.5$  and the probability of observing the exact number of successes we observed is 0.01432598.

### Testing the hypothesis: P-values

In this experiment, we had 18 successes out of 25 trials. Does this show evidence of preference for uninfected males? To answer this question, we need to know the chance of observing these data (or any data more extreme) if females truly had no preference. In other words: what is the chance of getting 18/25, or any probability more extreme, if females really don't care about infection status? This probability is your P-value.

Let's test the hypothesis by calculating a P-value from the observed data. We will do this by comparing the observed data to a null distribution of what the data might look like if females really had no preference.

Using the formula you employed in question 5, we can determine the probability of observing a given number of "successes" under the null hypothesis. If we calculate the probability of all possible observations, we have a *null distribution*. Luckily you don't have to do this by hand; R has some handy functions to do it for you.

### Calculating the null distribution

To make a null distribution, we need to know the null hypothesis (in this case, a probability). If you wrote this in a full sentence in question 2, you'll now need to convert it into a probability. For example, a null hypothesis about picking cards might be that you are *equally likely to pick diamonds as any other suit*. Phrasing this quantitatively, you could say:  $p(\text{diamonds}) = 1/4$ .

Assign the null probability as a variable in R:

```
p.null.diamond <- 1/4
```

**Question 6:** (1/2 point) Create a variable, `p.null.success` that represents your null hypothesis about the probability of a female choosing an uninfected male.

```
p.null.success <- 0.5
```

You have already calculated the probability of getting 18 successes out of 25. Now we need to calculate the probabilities of all the other possible outcomes. There are 26 possible outcomes: 0 successes, 1, 2, 3... up to 25 successes.

We will now create a vector listing all the possible outcomes:

```
observations <- seq(0,25)
observations
```

```
## [1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
## [26] 25
```

Next, the `dbinom` function calculates the probability of a given observation, given a null probability. Let's calculate the probability of having 18 successes out of 25 trials, under the null hypothesis.

```
prob.18 <- dbinom(x = 18, size = 25, prob = p.null.success)
prob.18
```

```
## [1] 0.01432598
```

The `x=` argument gives the number of “successes”. The `size=` argument gives the number of observations or trials, and the `prob=` is the proportion you expect if the null hypothesis is true. The `dbinom` function uses the same formula that you used in question 5. Does your probability match? It should!

To get the full null distribution, we must calculate the probability for each of the 26 possible observations. To do this, use the `dbinom` function above, but replace the 18 in `x = 18` with your *vector* of possible observations, `observations`, so that R can calculate the probability of all 26 possibilities.

**Question 7:** (1/2 point) Using the `dbinom` function, create a vector, `null.probabilities`, giving the probability of each observation.

```
null.probabilities <- dbinom(x=observations, size=25, prob=p.null.success)
null.probabilities
```

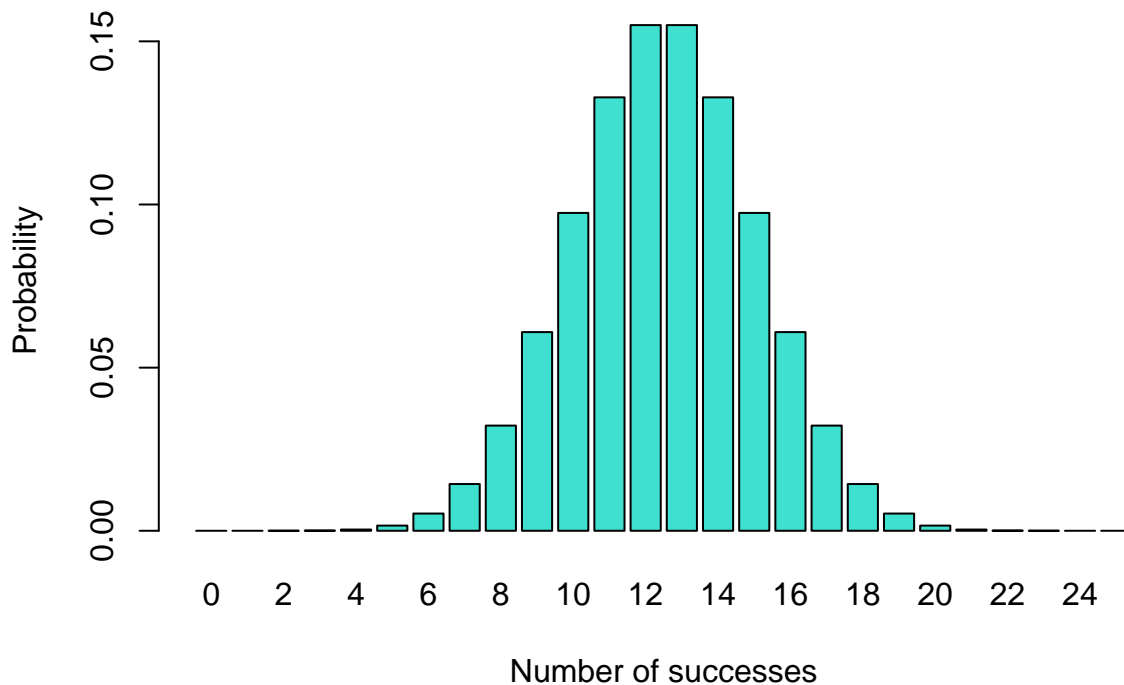
```
## [1] 2.980232e-08 7.450581e-07 8.940697e-06 6.854534e-05 3.769994e-04
## [6] 1.583397e-03 5.277991e-03 1.432598e-02 3.223345e-02 6.088540e-02
## [11] 9.741664e-02 1.328409e-01 1.549810e-01 1.549810e-01 1.328409e-01
## [16] 9.741664e-02 6.088540e-02 3.223345e-02 1.432598e-02 5.277991e-03
## [21] 1.583397e-03 3.769994e-04 6.854534e-05 8.940697e-06 7.450581e-07
## [26] 2.980232e-08
```

Now, let's make a dataframe and plot the data:

```
# create a dataframe
null.distribution <- data.frame(observations, null.probabilities)
# set column names
colnames(null.distribution)<-c("Observation", "Probability")
head(null.distribution)
```

```
## Observation Probability
## 1          0 2.980232e-08
## 2          1 7.450581e-07
## 3          2 8.940697e-06
## 4          3 6.854534e-05
## 5          4 3.769994e-04
## 6          5 1.583397e-03
```

```
barplot(height = null.probabilities, names = observations,
        ylab = "Probability", xlab = "Number of successes", col="turquoise")
```



This bar plot gives the probability of a given observations IF your null hypothesis is true. This is the **null distribution**.

Notice that the null distribution sums to 1, but each individual observation is pretty unlikely:

```
sum(null.distribution$Probability) # all the probabilities added together
```

```
## [1] 1
```

```
# probability of one specific outcome, say 3 out of 25 successes:
```

```
null.distribution[observations == 3, ] # a very small number!
```

```
##      Observation      Probability
```

```
## 4              3 6.854534e-05
```

### Calculating the P-value

Given an observation of 18 successes out of 25 trials, we want to know if this is sufficient evidence to reject the null hypothesis.

The P-value is *not* the probability of observing 18 successes under the null hypothesis, but rather, the probability of observing a result *as or more extreme* than 18. What observations would be “as or more extreme”?

If we are observing coin tosses and we had 9 heads out of 12 tosses, observations AS or more extreme include: 9, 10, 11, or 12 (the “right tail” of the null distribution) heads OR 0, 1, 2, or 3 heads (the “left tail” of the null distribution).

**Question 8:** What observations are as or more extreme than 18/25? Use the bar chart above to help you visualize this.

Observations that are as or more extreme than 18/25 include 18/25, 19/25, 20/25, 21/25, 22/25, 23/25, 24/25 and 25/25 OR 0/25, 1/25, 2/25, 3/25, 4/25, 5/25, 6/25, and 7/25.

To get the P-value, we’ll sum all the probabilities in our dataframe that correspond to observations as or more extreme than 18/25. Think of this as summing the areas of the bar chart corresponding to the tails of the distribution. Switch out the 2 and 5 in the code below to get the P-value for the guppy data.

```
P.value <- sum(null.probabilities[observations <= 7] + null.probabilities[observations >= 18])
P.value
```

```
## [1] 0.04328525
```

**Question 9:** Based on your observed data, is female guppy mate choice influenced by parasite infection? Make sure to refer to the null hypothesis and interpret the P-value in your answer.

*Yes, female guppy mate choice is influenced by parasite infection. We must refute the null hypothesis (that females do not use parasite load to preferentially choose mates) because the probability of observing some as or more extreme as 18/25 is so low ( $p = 0.0432 < 0.05$ ).*

### Using R's built-in function

As demonstrated in class, there is a much easier way to calculate the P-value for a binomial test. The example code below tests whether an observation of 3 heads out of a sample of 10 coin tosses deviates from the null expectation of 50% heads:

```
binom.test(3, n = 10, p = 0.5, alternative = "two.sided")
```

```
##
## Exact binomial test
##
## data: 3 and 10
## number of successes = 3, number of trials = 10, p-value = 0.3438
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.06673951 0.65245285
## sample estimates:
## probability of success
## 0.3
```

Use the `binom.test` function, modifying it as you see fit, to get a P-value for the guppy data. Make sure it matches your P-value above. If not, go back and troubleshoot the P-value you got by hand!

```
binom.test(18, n = 25, p = 0.5, alternative = "two.sided")
```

```
##
## Exact binomial test
##
## data: 18 and 25
## number of successes = 18, number of trials = 25, p-value = 0.04329
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.5061232 0.8792833
## sample estimates:
## probability of success
## 0.72
```

### The importance of sample size

Now, say the researchers only tested 7 female guppies, instead of 25. The observed data are: 0, 1, 0, 1, 1, 1, 1.

**Question 10:** Calculate  $\hat{p}$  (p-hat) and its standard error with the new sample. Again, don't count up the number of successes in your head, but create a vector with the observations and use R's embedded functions.

```
guppy.smallerTrials <- c(0, 1, 0, 1, 1, 1, 1)
propSucc <- sum(guppy.smallerTrials) / length(guppy.smallerTrials)
```

```
SEsucc <- sqrt((propSucc * (1-propSucc))/length(guppy.smallerTrials))
SEsucc
```

```
## [1] 0.1707469
```

*The p-hat for this smaller sample size is 0.7142857, with a SE of 0.1707469.*

How does this p-hat compare to the observed p-hat from Question 3?

**Question 11:** Now calculate and interpret the P-value from the new set of 7 observations. You may use the `binom.test()` function. Why does the P-value change drastically, even though p-hat is similar to our first set of observations?

```
binom.test(5, n = 7, p = 0.5, alternative = "two.sided")
```

```
##
## Exact binomial test
##
## data: 5 and 7
## number of successes = 5, number of trials = 7, p-value = 0.4531
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.2904209 0.9633074
## sample estimates:
## probability of success
## 0.7142857
```

*With a much smaller sample size the p-value becomes 0.4531. The p-value changes drastically despite a similar p-hat to our first observation because a smaller sample size gives us much larger confidence intervals and as sample size increases the impact of stochastic error is reduced.*