Hypothesis tests for a single proportion

Overview

Review of hypothesis tests

Continuation of hypothesis tests for a single proportion

Hypothesis tests for a single proportion in R

More practice of hypothesis tests for a single proportion

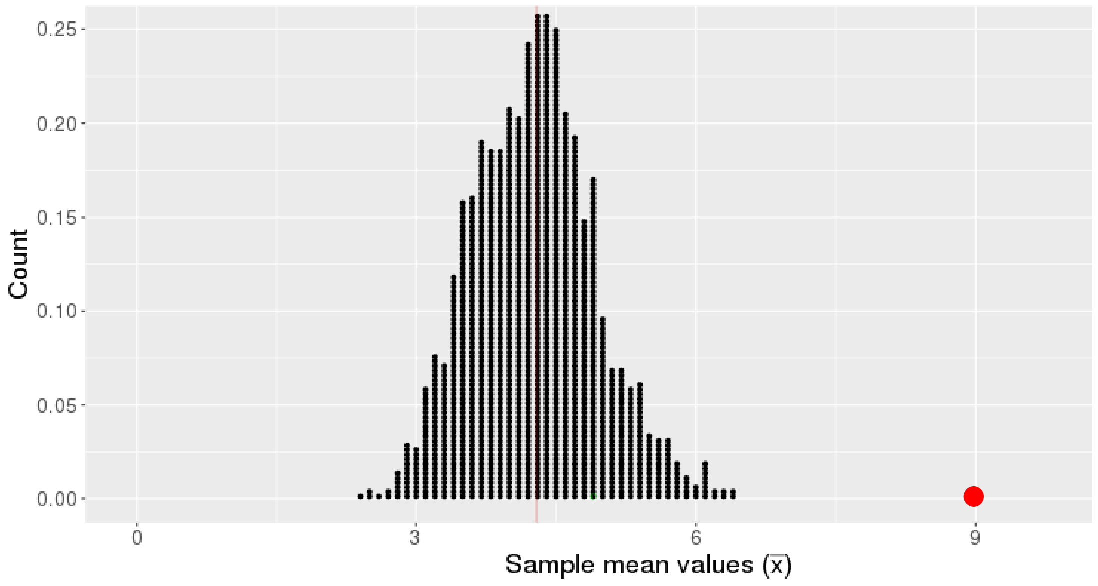
One-sided vs. two-sided tests

Review: The basic logic of hypothesis tests

We start with a claim about a population parameter

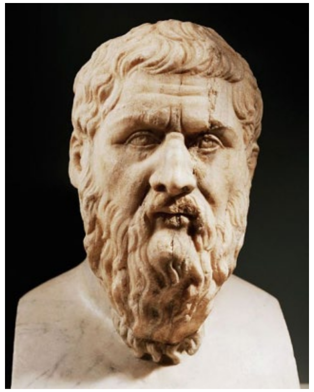
- E.g., $\mu = 4$

This claim implies we should get a certain distribution of statistics



If our observed statistic is highly unlikely, we reject the claim

Five steps of hypothesis testing

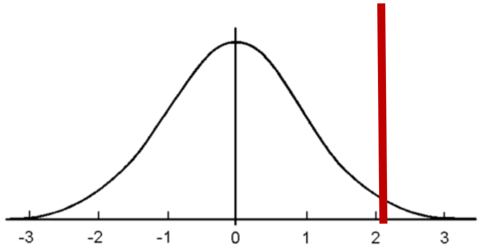


1. State H_0 and H_A
   - Assume Gorgias (H_0) was right

2. Calculate the actual observed statistic

3. Create a distribution of what statistics would look like if Gorgias is right
   - Create the **null distribution** (that is consistent with H_0)

4. Get the probability we would get a statistic more than the observed statistic from the null distribution
   - p-value



5. Make a judgement
   - Assess whether the results are statistically significant

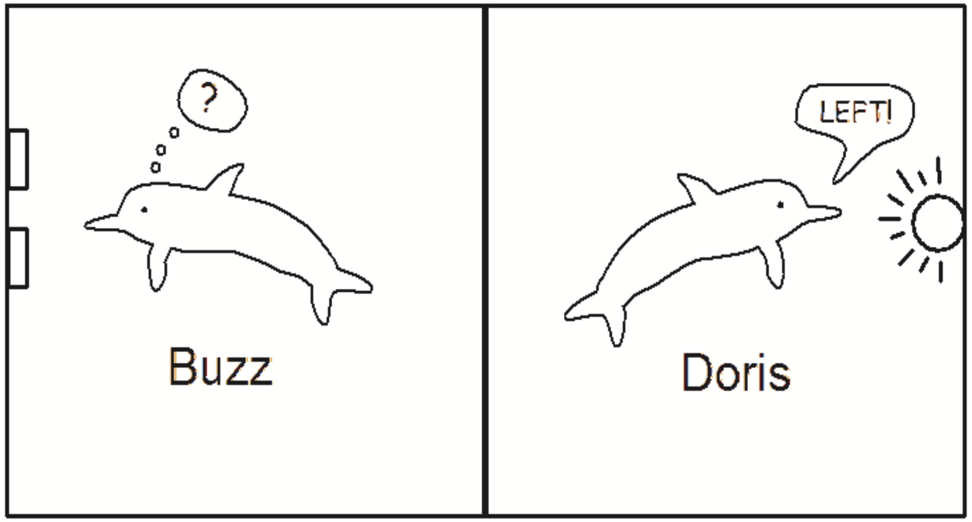
Are dolphins capable of abstract communication?

Dr. Jarvis Bastian is the 1960's wanted to know whether dolphins are capable of abstract communication

He used an old headlight to communicate with two dolphins (Doris and Buzz)

- Stead light = push button on right to get food
- Flashing = push button on the left to get food

The two dolphins were then separated by a barrier



Buzz got 15 out of 16 trials correct

Hypothesis testing in 5 easy steps!

1. State the null hypothesis… and the alternative hypothesis
   - Buzz is just guessing so the results are due to chance: H_0: $\pi = 0.5$
   - Buzz is getting more correct results than expected by chance: H_A: $\pi > 0.5$

2. Calculate the observed statistic
   - Buzz got 15 out of 16 guesses correct, or $\hat{p}$ = .973

3. Create a null distribution that is consistent with the null hypothesis
   - i.e., what statistics would we expect if Buzz was just guessing

4. Examine how likely the observed statistic is to come from the null distribution
   - What is the probability that the dolphins would guess 15 or more correct?
   - The p-value = 0.003

5. Make a judgement
   - If we have a small p-value, this means that $\pi = .5$ is unlikely and so $\pi > .5$
   - i.e., we say our results are 'statistically significant'


¿Qué animal de mar llega
siempre al final?

Do more that 25% of US adults believe in ghosts?

A telephone survey of 1000 randomly selected US adults found that 31% of them say they believe in ghosts. Does this provided evidence that more than 1 in 4 US adults believe in ghosts?

On Canvas: answer the following questions

1. What are the cases here?

2. What is the variable of interest and is it categorical or quantitative?

3. What is the observed statistic - and what symbols should we use to denote it?

4. What is the population parameter we are trying to estimate - and what symbol should we use to denote it?

5. Do you think that more that 25% of US citizens believe in ghosts?

5 steps to null-hypothesis significance testing (NHST)

Let's go through the 5 steps!

1. State null and alternative hypotheses
2. Calculate statistic of interest
3. Create a null distribution
4. Calculate a p-value
5. Assess if there is convincing evidence to reject the null hypothesis

Step 1: State the null and alternative hypotheses

Null Hypotheis (H_0): Claim that there is no effect or no difference

Alternative Hypothesis (H_a): Claim for which we seek significant evidence.

Believing in ghosts study

Q: What is the null hypothesis? (please state it using words)

Q: How would you write it in terms of the population parameter?

H_0: $\pi = 0.25$

Q: What is the alternative hypothesis?

H_A: $\pi > 0.25$

Step 2: Calculate statistic of interest

For the ghost study, what was the observed statistic?

31% percent of Americans believe in ghosts ($\hat{p}$ = .31)

Step 3: Create a null distribution

Q: Please describe what the null distribution is here

Answer: A distribution of statistics ($\hat{p}$'s) consistent with the null hypothesis (H_0: $\pi = 0.25$)

Q: How can we create a null distribution?

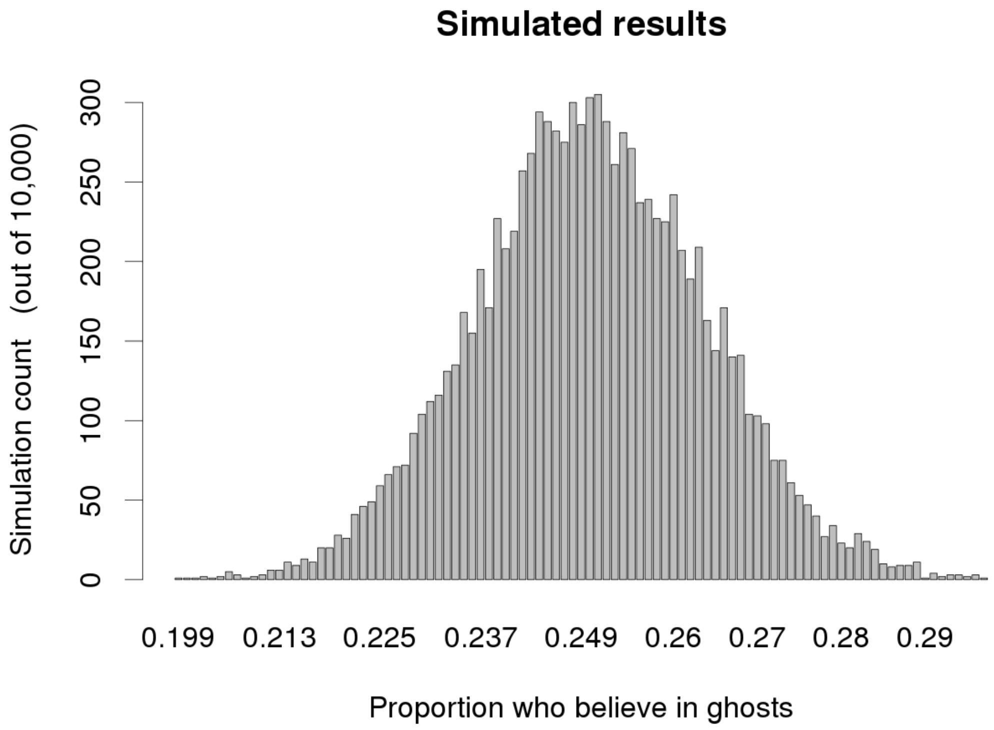
Answer: when making inferences on *proportions* we can simulate flipping coins

Step 3: Create a null distribution

Please answer the following questions for the ghost study

1. How many coins should we flip?
   - A: 1,000 people were surveyed, so we want to flip 1,000 coins to simulate each person's answer
2. What should the probability of heads be on each flip?
   - A: Our null parameter is $\pi = 0.25$ so to be consistent with the null hypothesis the probability of heads on each flip should be 0.25
3. How many simulations should we run?
   - A: 10,000 simulations should be enough to give us a good sense of the statistics we would get if the null hypothesis was true

Step 3: Create a null distribution



A null distribution ($\hat{p}$'s) based on:

- ***10,000 simulations***
- Each simulation consists of flipping 1,000 coins
- With the probability of getting a head on each flip of 0.25

Step 4: Calculate a p-value

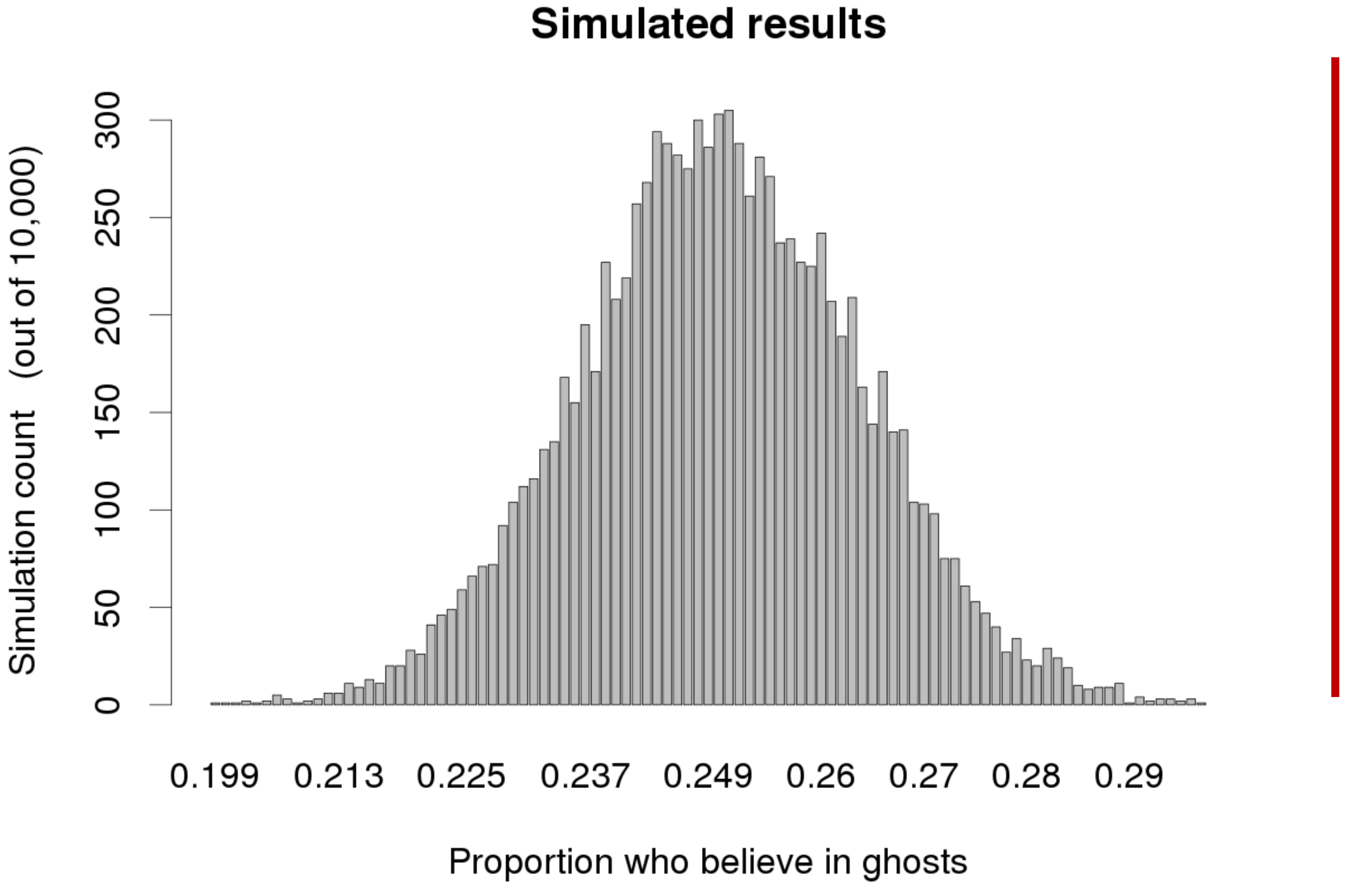
The **p-value** is the probability, when the null hypothesis is true, of obtaining a statistic as extreme or more extreme than the observed statistic

$$Pr(\text{STAT} \geq \text{observed statistic} \mid H_0 = \text{True})$$

The smaller the p-value, the stronger the statistic evidence is against the null hypothesis and in favor of the alternative

Step 4: Calculate a p-value

$\hat{p} = 0.31$



What is the p-value here?

- A: The p-value is close to 0

Step 5a: Assess if results are statistically significant

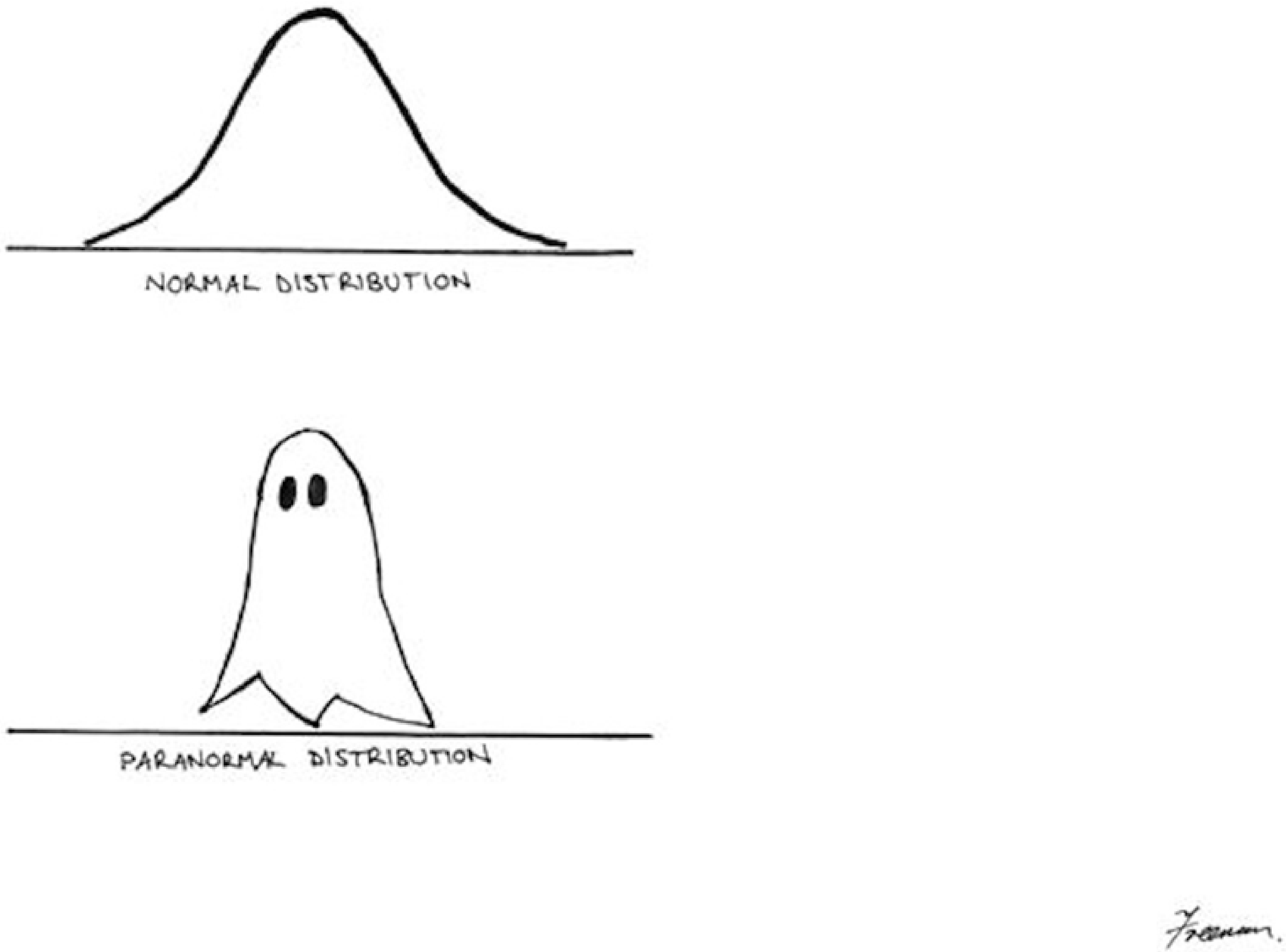
When our observed sample statistic is unlikely to come from the null distribution, we say the sample results are **statistically significant**

- i.e., we have a small p-value

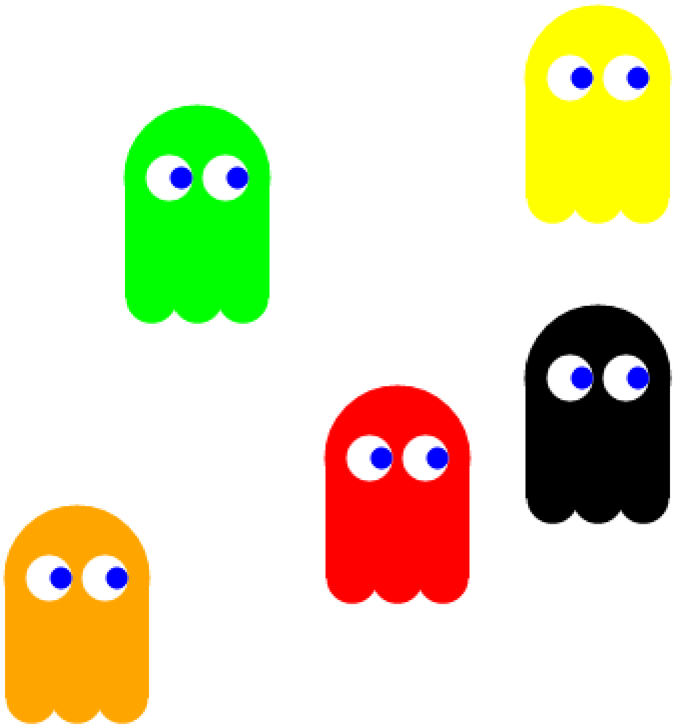
'Statistically significant' results mean we have convincing evidence against H_0 in favor of H_a

Step 5b: Make a decision

Are the results seem statistically significant?


NORMAL DISTRIBUTION
PARANORMAL DISTRIBUTION

Let’s try the ghost example in R…



The amazing woman who can smell Parkinson’s disease

Joy Milne claimed to have the ability to smell whether someone had Parkinson’s disease

To test this claim researchers gave Joy 6 shirts that had been worn by people who had Parkinson’s disease and 6 people who did not

Joy identified 11 out of the 12 shirts correctly

The amazing woman who can smell Parkinson's disease

Please try to complete the following steps to analyze the data:

1. State Null and Alternative in symbols and words

2. Calculate the observed statistic of interest (obs_stat)

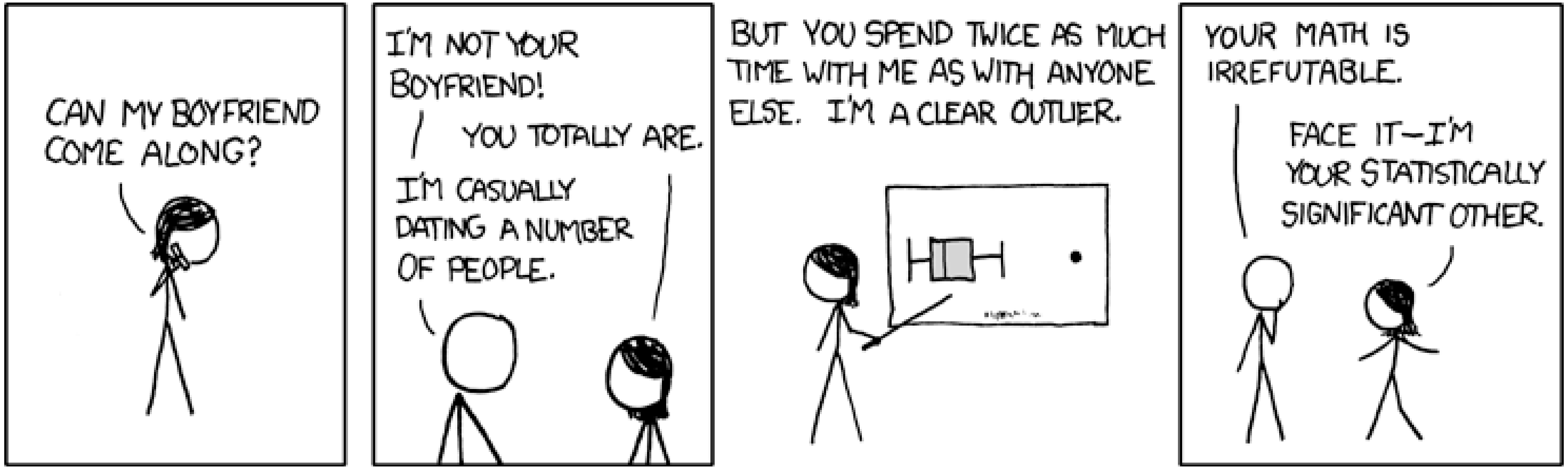
3. Create a null distribution using:

```
null_dist <- do_it(10000) * { rflip_count(num_flips = ... prob = ... ) }
```

4. Calculate a p-value by assessing the probability of getting a statistic as or more extreme than the observed statistic from the null distribution
   - `pnull(obs_stat, null_dist, lower.tail = ...)`

5. Make a decision about whether the results are statistically significant

Bad example?



What is the null hypothesis here?

Are the results statistically significant?

One-sided vs. two-sided

In the examples we have seen, we were just interested if the parameter was **<u>greater</u>** than an hypothesized parameter

H_0: $\pi = 0.25$ H_A: $\pi > 0.25$

In other cases we might not have a directional alternative hypothesis

Testing whether a coin is biased

Suppose we wanted to test what whether Buzz chose the correct food well more ***or less*** than 50% of the time

- e.g., Buzz might not like the food so was avoiding the well with the food

1. Write down the null and alternative hypotheses

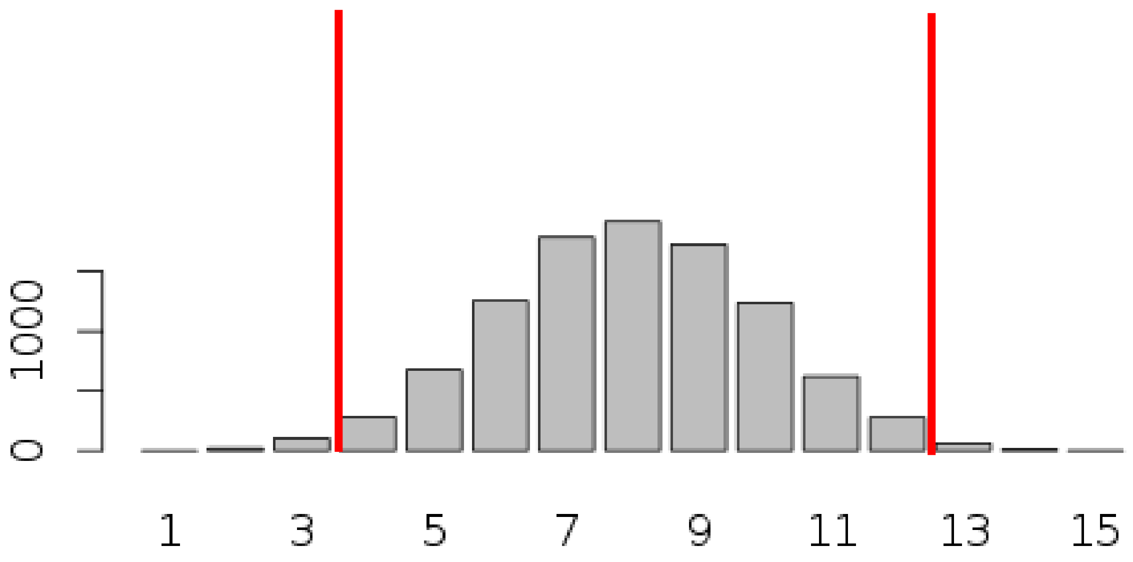
2. Suppose out of the 16 trials, Buzz got the correct 3 times. How would we use a randomized distribution to tell if the coin is biased?

0	0
1	1
2	22
3	105
4	283
5	679
6	1257
7	1786
8	1920
9	1726
10	1238
11	623
12	279
13	63
14	15
15	3
16	0

3. Suppose out of the 16 trials, Buzz got the correct 3 times. How would we use our null to get the p-value?

4. Based on this table, what is the p-value?

0	0
1	1
2	22
3	105
4	283
5	679
6	1257
7	1786
8	1920
9	1726
10	1238
11	623
12	279
13	63
14	15
15	3
16	0



p-value = 209/10000 = .0209

Compare this p-value to we would have gotten if we **expected** Buzz to avoid the food well?

Statement of alternative hypothesis is important

We need to state what you expect before analyzing the data

Our expectation (hypothesis statement) can change the p-value!

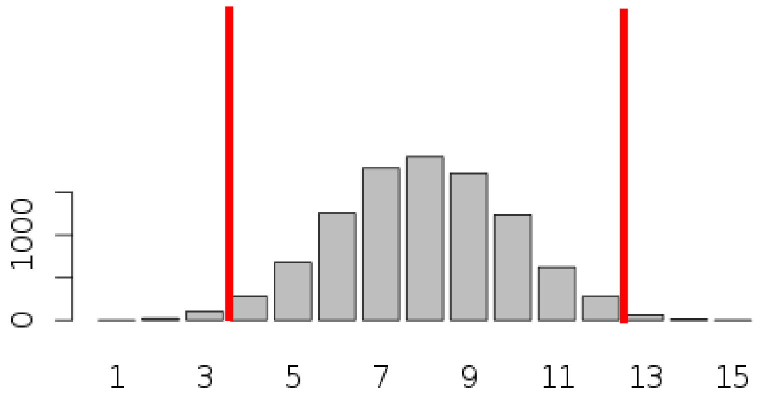
Estimating a p-value from a null distribution

For a one tailed alternative: Find the proportion of statistics in the null distribution that equal or exceed the original statistic in the direction (tail) indicated by the alternative hypothesis

For a two-tailed alternative: Find the proportion of statistics in the null distribution in the tails beyond the observed statistic and 1 - the observed statistic.

- Alternatively, find the proportion of statistics in the null distribution beyond the original statistic in one of the tails, and then double the proportion to account for the other tail

How to estimate two sided p-values in R?




```
null_distribution <-  do_it(10000) * {
            rflip_count(16,  prob = .5)
}

p_left_tail  <-   pnull(3,  null_distribution,  lower.tail = TRUE)
p_right_tail  <- pnull(16 - 3,  null_distribution,  lower.tail = FALSE)

p_value <-  p_right_tail + p_left_tail
```

Paul the Octopus

In the 2010 World Cup, Paul the Octopus (in a German aquarium) became famous for correctly predicting 11 out of 13 soccer games.

Question: is Paul psychic?

Homework 5!