Hypothesis tests for a single proportion

(belated)

Halloween edition...

Overview

Review and continuation of hypothesis tests for a single proportion

One-sided vs. two-sided tests

Announcement: No class on Tuesday

I will be away at the Society for Neuroscience conference

Class will resume next Thursday (11/8)

TA office hours on Tuesday (11/6) during class time

Statistical tests

A **statistical test** uses data from a sample to assess a claim about a population

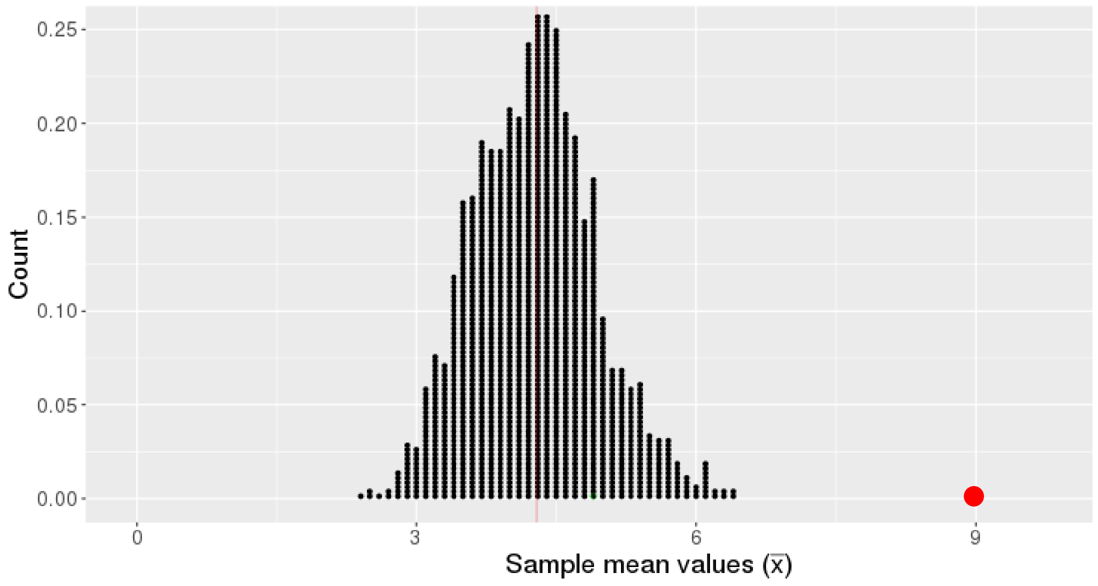
Basic hypothesis test logic

We start with a claim about a population parameter

- E.g., $\mu = 4.2$

This claim implies we should get a certain distribution of statistics

- The null distribution



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

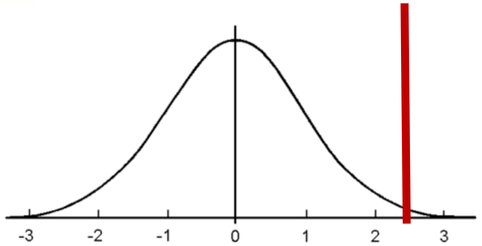
There are 5 steps involved in running a statistical test…

Trial metaphor for the 5 steps

1. State null and alternative hypotheses
    - Assume innocence (H_0)
    - State what guilt means (H_A)

2. Calculate statistic of interest
    - Gather evidence

3. Create a null distribution
    - Describe what an innocent person looks like



4. Calculate a p-value
    - See if the evidence is compatible with an innocent person's behavior

5. Assess if results are statistically significant
    - Make your innocent/guilty judgment

Statistical tests

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?

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

1. What is the null hypothesis?

2. How would you write it in terms of the population parameter?

H_0: $\pi = 0.25$

3. 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

How can we create a null distribution?

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

```
rbinom(num_sims, size, prob)
```

num_sims: the number of simulations run
size: the number of trials on each simulation
- i.e., the sample size (n)

prob: the probability of success on each trial

How do we run simulate 500 simulations of 16 trials with the probability of heads is 0.5?

How do we do run 10,000 simulation of 1,000 trials where the probability of heads is 0.25?

Step 3: Create a null distribution

For our ghost example, we need to:

- Flip a biased coin Pr(head) = .25, 1000 times
- Count the number of heads
- Repeat 10,000 times

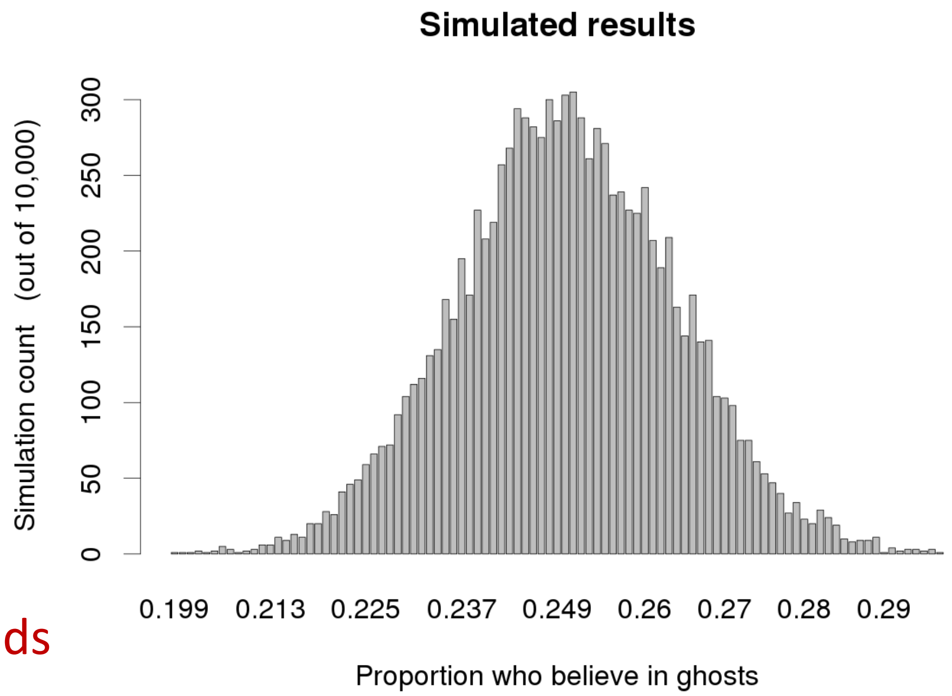
```
# numbers from 0 to 1000
null_dist <- rbinom(10000, 1000, .25)
```

Number of simulations

Number of trials on each simulation

Probability of heads on each trial

```
# convert to proportions: numbers from 0 to 1
null_dist <- null_dist/size
```



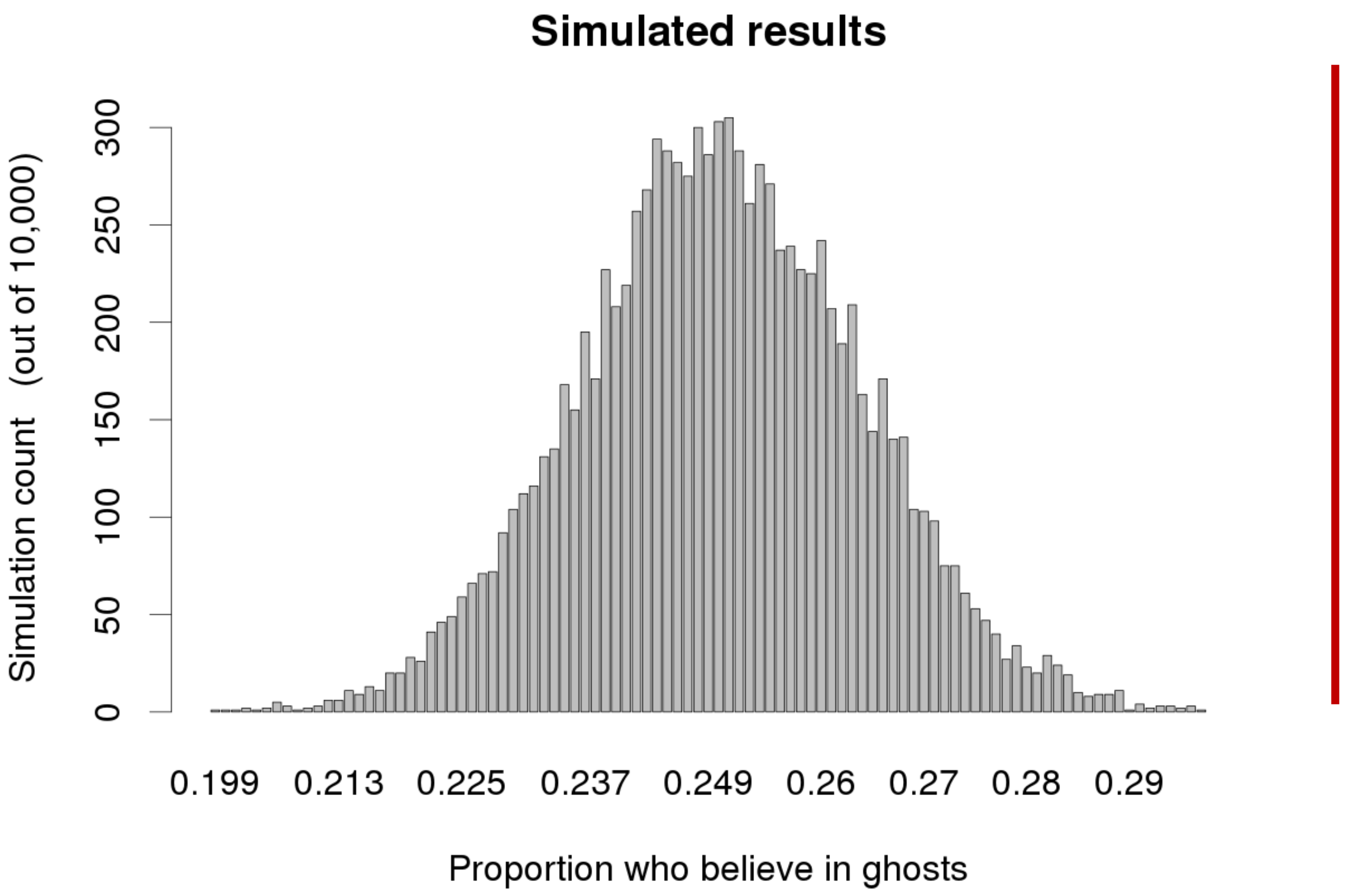
Step 4: Calculate a p-value

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

$$Pr(\text{Stat} > \text{observed statistic} \mid H_0 = \text{True})$$

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



p-value is close to 0

pval <- sum(null_dist >= .31)/10000

Step 5a: Assess if results are statistically significant

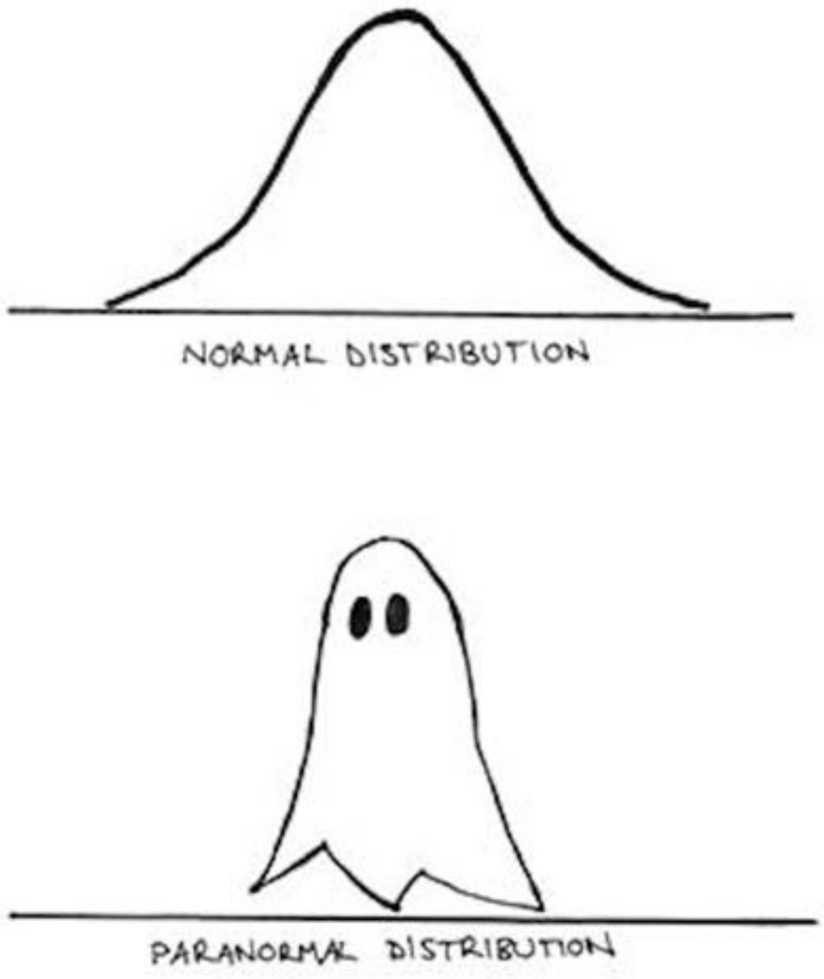
When our observed sample statistic is unlikely to come from the null distribution, we say the sample results are **statistically significant**

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

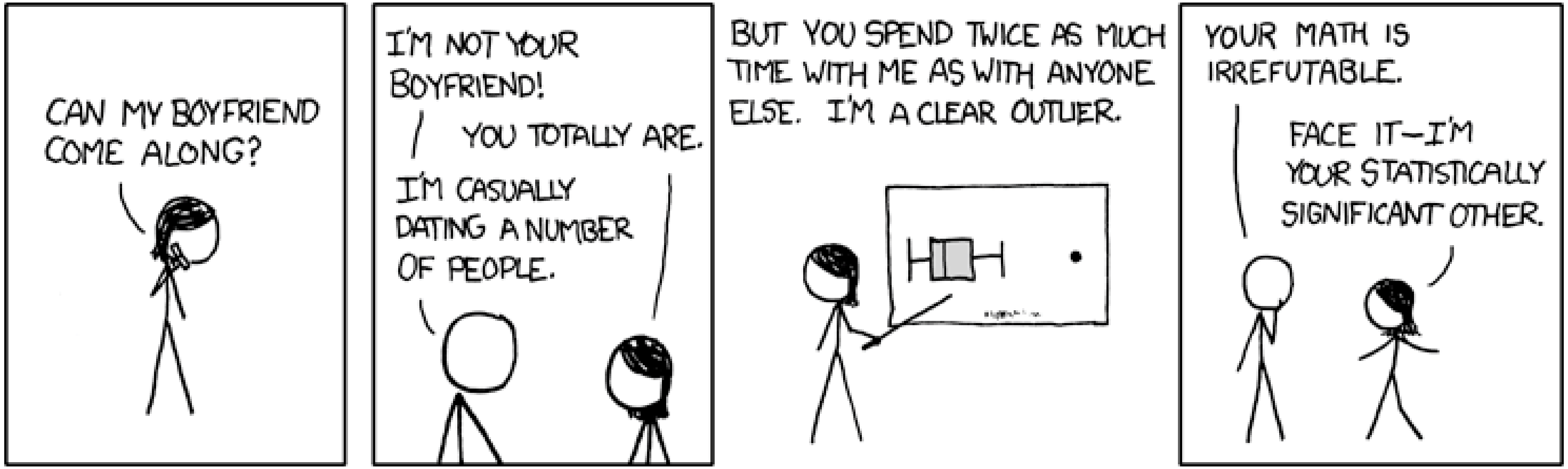
'Statistically significant' results mean we have convincing evidence against H_0 in favor of H_a

Step 5b: Make a decision

Do the results seem statistically significant?


NORMAL DISTRIBUTION
PARANORMAL DISTRIBUTION

Bad example?



What is the null hypothesis here?

Are the results statistically significant?

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

Work in pairs 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. Use the flip_sims <- rbinom(num_sims, size, prob) to create a null distribution

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
    - p_value <- sum(flip_sims >= obs_stat)/num_sims

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

One-tailed vs. two tailed

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

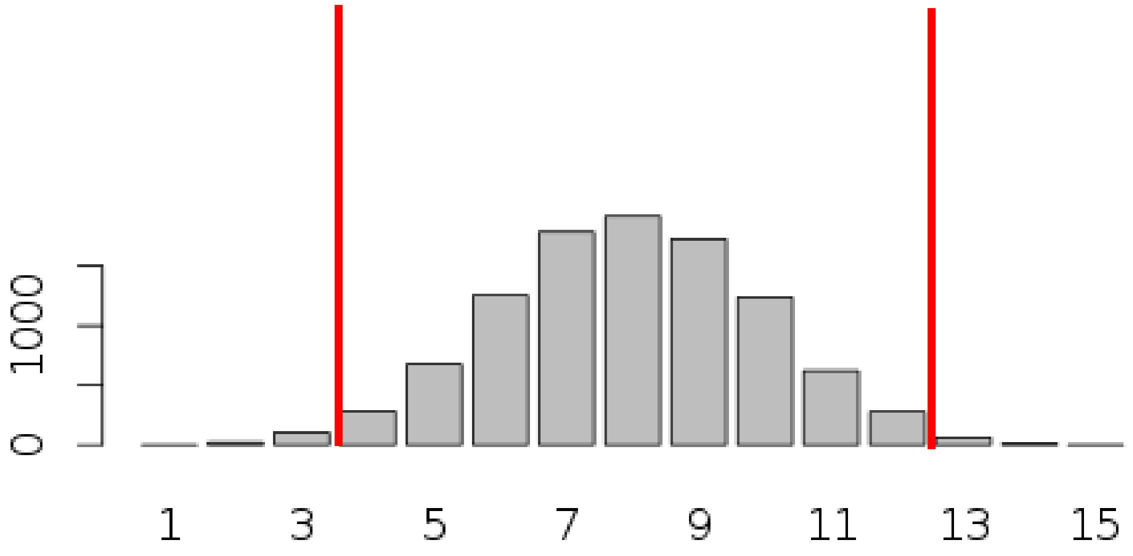
2. Suppose out of the 16 trials, Buzz got the correct 3 times. How would we use a randomized distribution to tell if the coin is biased?

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

2. Suppose out of the 16 trials, Buzz got the correct 3 times. How would we use our randomized distribution to tell?

3. 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

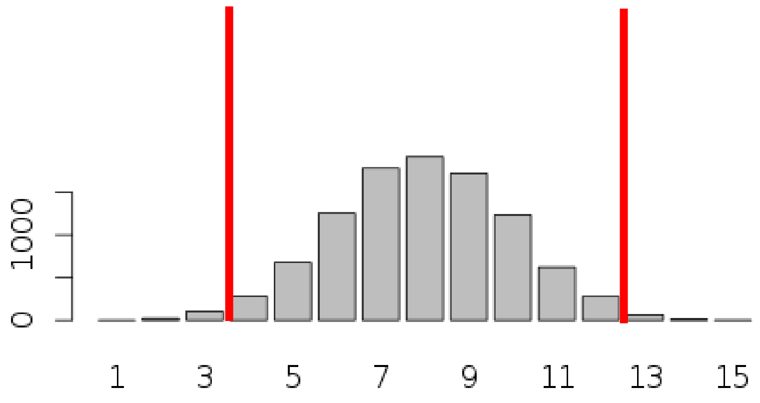
Estimating a p-value from a randomized distribution

<u>For a one tailed alternative</u>: Find the proportion of randomized samples that equal or exceed the original statistic in the direction (tail) indicated by the alternative hypothesis

<u>For a two-tailed alternative</u>: Find the proportion of randomization samples in the tails beyond the observed statistic and 1 - the observed statistic

- Alternatively, find the proportion of randomization samples in the smaller tail at or beyond the original statistic and then double the proportion to account for the other tail

How to estimate two sided p-values in R?




```
null_dist <- rbinom(10000, 16, .5)     # numbers 0 to 16
null_dist <- null_dist/16    # convert to p̂ numbers from 0 to 1

num_right_tail  <-  sum(null_dist >= 13/16)
num_left_tail  <-  sum(null_dist <= 3/16)

p_value <- (num_right_tail + num_left_tail)/10000
```

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?

Worksheet 8!

You know the drill…

No class on Tuesday – I'm away at a conference

Worksheet 8 is due at 11:59pm on Sunday November 4th

```
> source('/home/shared/intro_stats/cs206_functions.R')
> get_worksheet(8)
```