Hypothesis
Tests

Example: Student-to-faculty ratio

<u>Question</u>

Is there a difference between the average student-to-faculty ratio between public and private four-year colleges?

<u>Data</u>

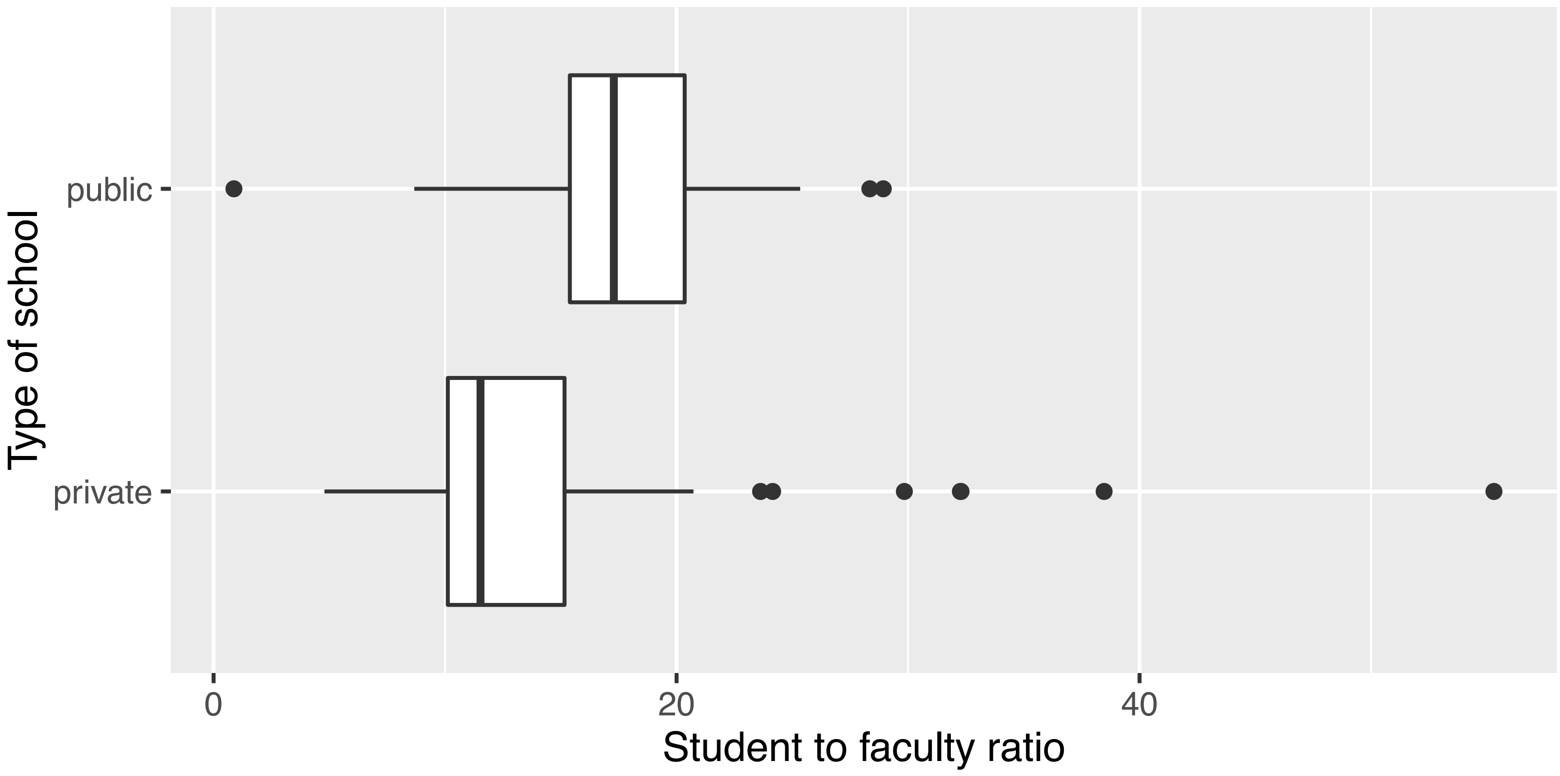
Random sample of 85 private and 57 public four-year colleges

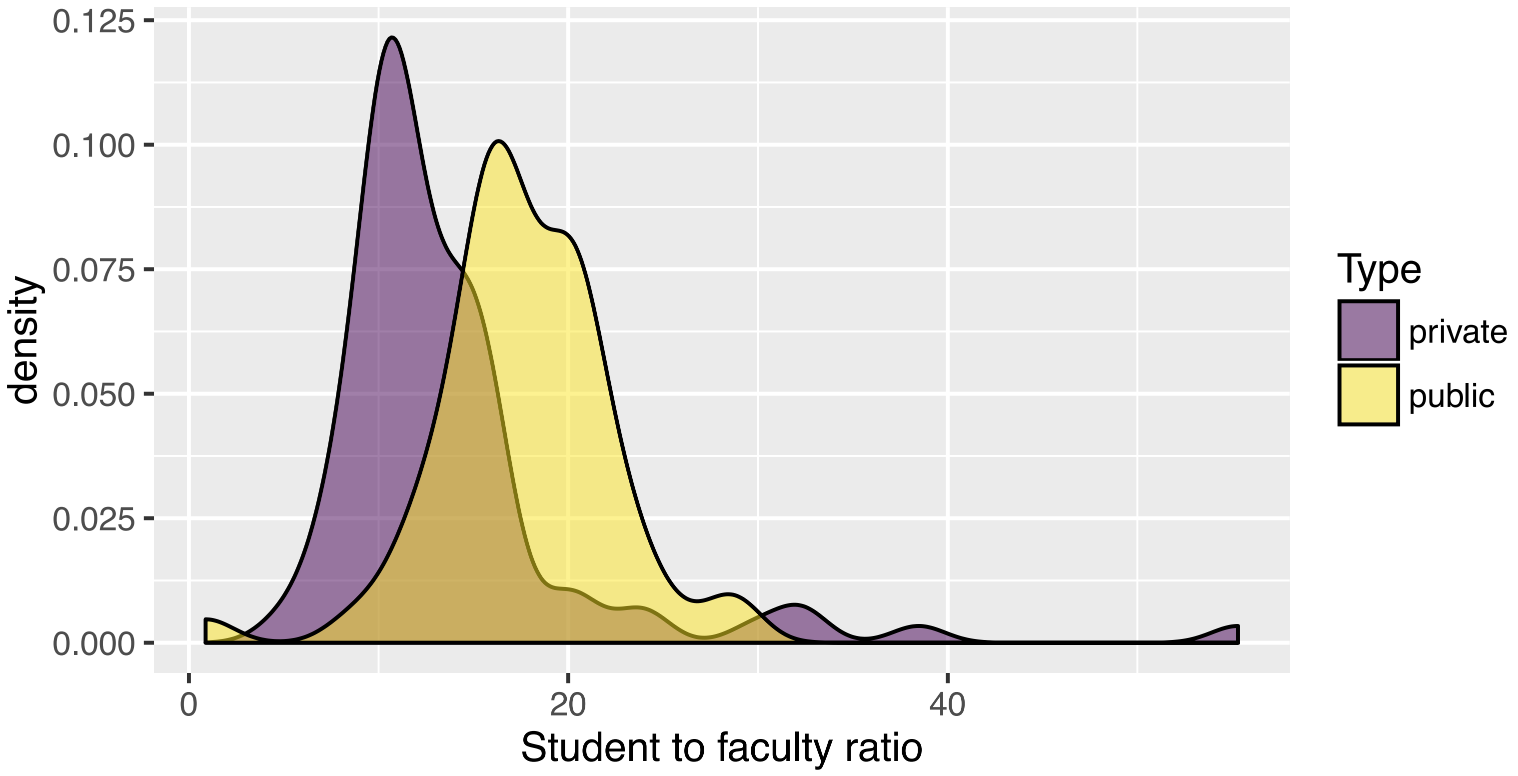
Example: Student-to-faculty ratio

Basic summary statistics from R:

```
# A tibble: 2 × 9
      type   min    Q1 median    Q3   max  mean    SD     n
    <fctr> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <int>
1 private  4.79 10.12   4.79 15.16 55.31 13.84  7.28    85
2  public  0.88 15.39   0.88 20.35 28.93 17.60  4.57    57
```

Example: Student-to-faculty ratio





1. Formulate two competing <u>hypotheses</u> about the population
2. Calculate a <u>test statistic</u> summarizing the relevant information to the claims
3. Look at the <u>behavior of the test statistic</u> assuming that the initial claim is true
4. <u>Compare</u> the observed test statistic to the distribution created in step 3 to see if it is "extreme"

1

Formulate two competing hypotheses about the population

Notation Parameters vs. statistics

The null hypothesis

A claim that the parameter is equal to some value

e.g. prior belief, "no effect", "no difference"

The alternative hypothesis

Opposition to the null: The claim for which we seek evidence

Example: Student-to-faculty ratio

H_0:

H_a:

2

Calculate a test statistic summarizing the relevant information to the claims

Example: Student-to-faculty ratio

```
# A tibble: 2 × 9
      type   min    Q1 median    Q3   max  mean    SD     n
    <fctr> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <int>
1 private  4.79 10.12   4.79 15.16 55.31 13.84  7.28    85
2  public  0.88 15.39   0.88 20.35 28.93 17.60  4.57    57
```

3

Look at the behavior of the test statistic assuming that the initial claim is true

Randomization distribution

A collection of statistics from samples simulated by assuming that the null hypothesis is true

Centered around the parameter value specified in H_0

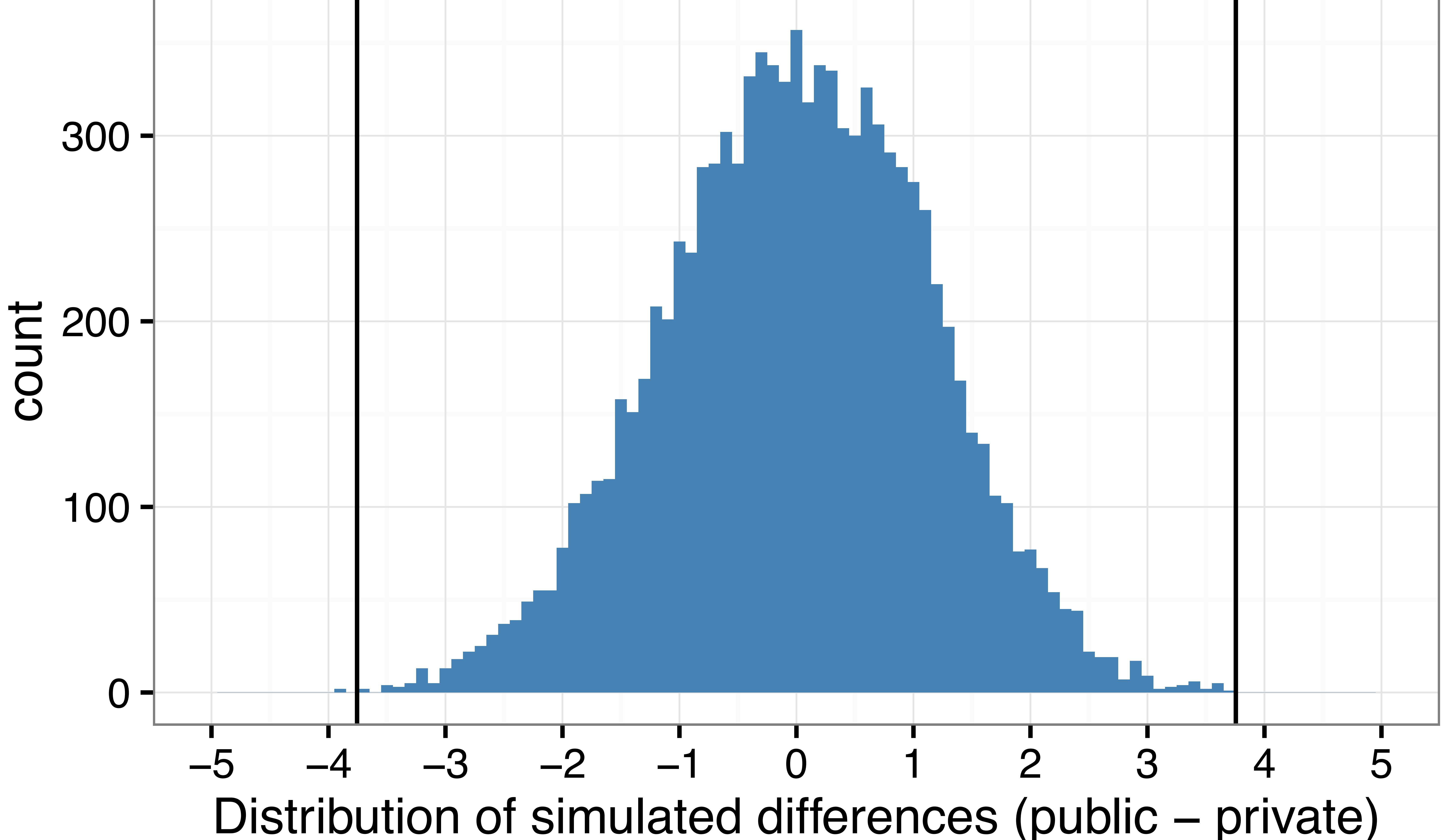
```
permTest(sf_ratio ~ type, data = colleges)

      ** Permutation test **

  Permutation test with alternative: two.sided
  Observed  mean
   private :  13.84482    public :  17.60018
  Observed difference: -3.75535

  Mean of permutation distribution: -0.00192
  Standard error of permutation distribution: 1.13257
  P-value:  6e-04

      *--------------*
```



300
200
100
0
count
−5
−4
−3
−2
−1
0
1
2
3
4
5
Distribution of simulated differences (public − private)

4

Compare the observed test statistic to the distribution created in step 3 to see if it is "extreme"

Quantifying evidence

An observed test statistic is rare if it is "too far out in the tails" of the randomization distribution

p-value:
Proportion of statistics in a randomization distribution that are at least as extreme as the observed test statistic

Interpreting the p-value

The p-value is the chance of obtaining a test statistic at least as extreme as the observed test statistic, <u>if the null hypothesis is true</u>

Strength of evidence

Making decisions

A p-value of 0.05 or below is conventionally called "statistically significant"

A p-value of 0.01 or below is conventionally called "highly statistically significant"

CAUTION: These thresholds are arbitrary

Example: Student-to-faculty ratio

What should we conclude about the difference between the average student-to-faculty ratio between public and private four-year colleges?