Hypothesis tests for more than two means

	5	3	2		7			8
6		1	5					2
2			9	1	3		5	
7	1	4	6	9	2			
	2						6	
			4	5	1	2	9	7
	6		3	2	5			9
1					6	3		4
8			1		9	6	7	

Overview

Hypothesis tests for two means continued

Hypothesis test for more than two means

Announcement: Midterm exam

Thursday October 23rd during regular class time

- Exam is on paper

If you have accommodations, please schedule the exam with SAS

A practice exam (last year's exam) has been posted

Midterm exam “cheat sheet”

You are allowed an exam “cheat sheet”

One page, double sided, that contains **only code and equations**

- No code comments allowed

Cheat sheet must be on a regular 8.5 x 11 piece of paper

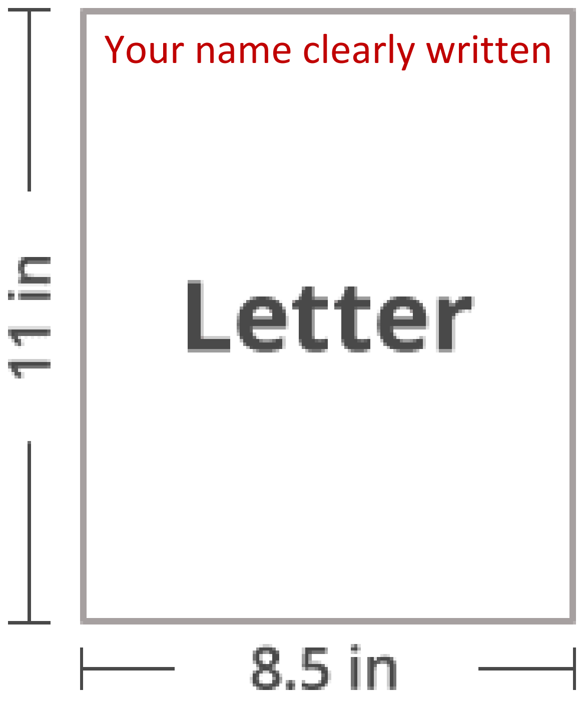
- Your name on the upper left of both sides of the paper

Recommend making a typed list of all functions discussed in class and on the homework

- This will be useful beyond the exam

You must turn in your cheat sheet with the exam

- Failure to do so will result in a 20 point deduction



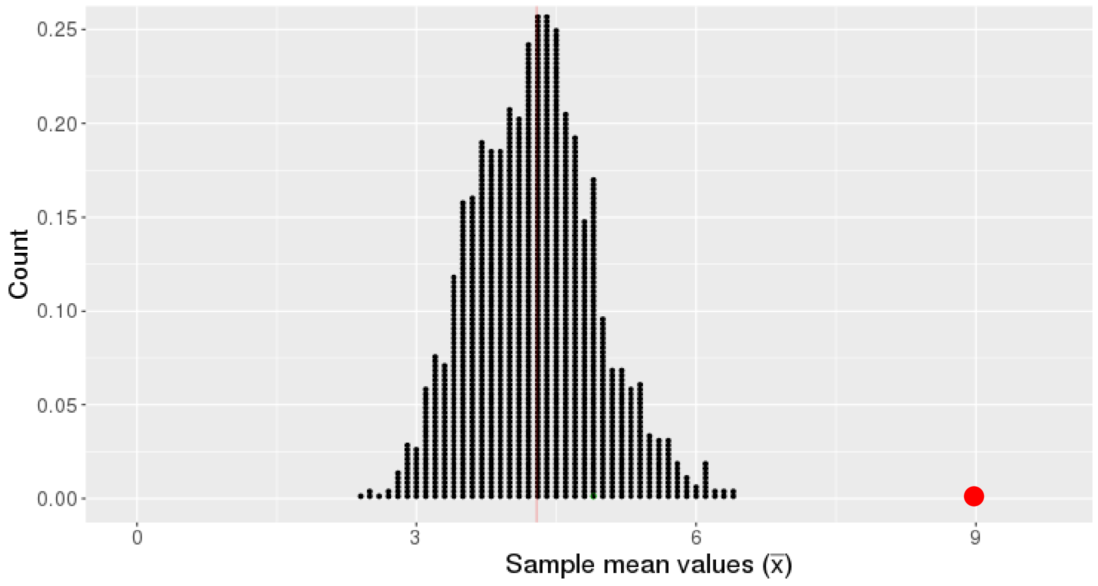
Hypothesis tests for comparing two means continued

The logic of hypothesis tests…

We start with a claim about a population parameter

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

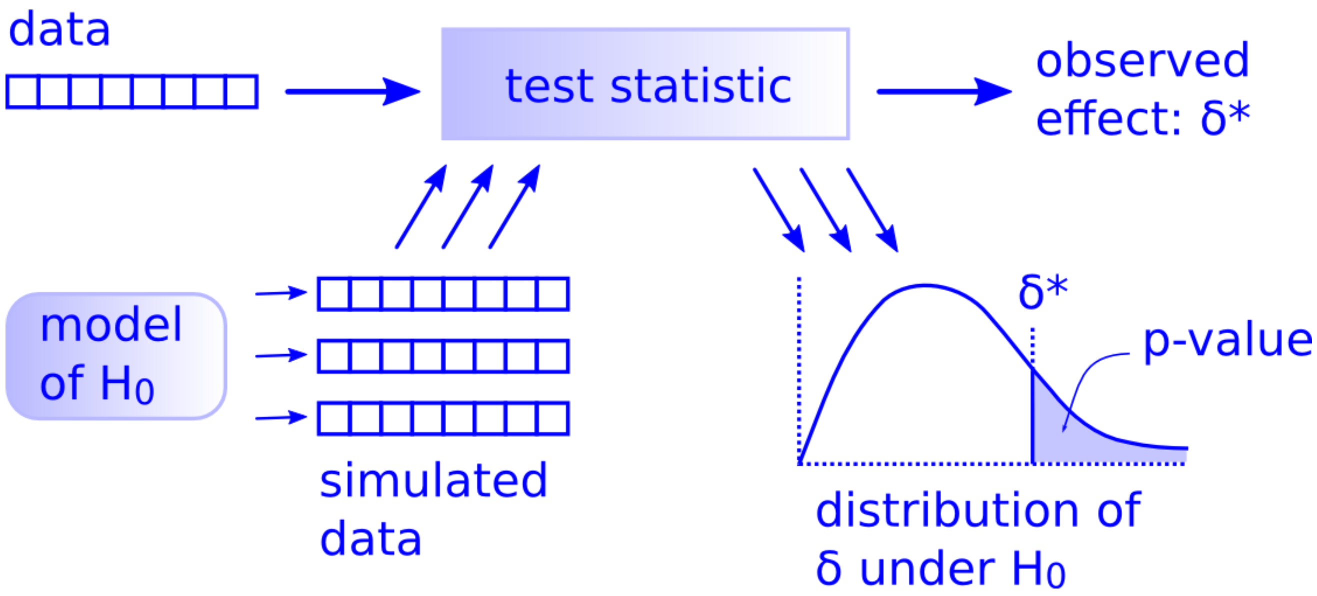
This claim implies we should get a certain distribution of statistics

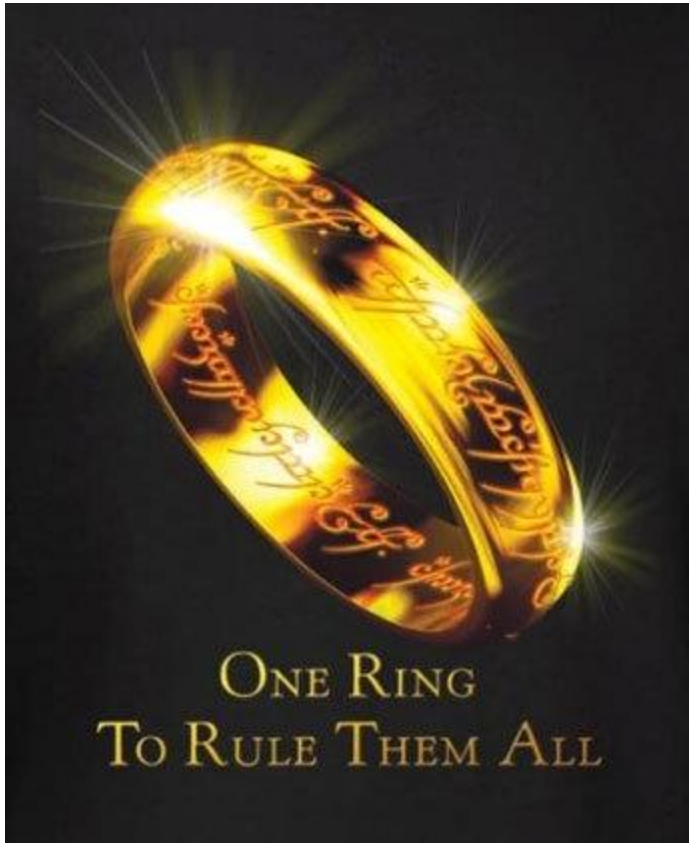


If our observed statistic is highly unlikely, we reject the claim

The logic of hypothesis tests…

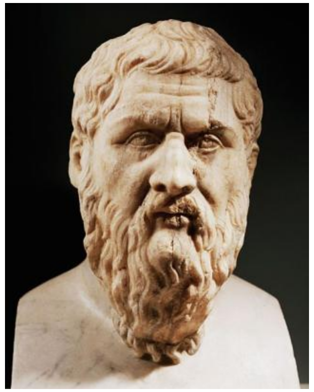
There is only one hypothesis test!





Just follow the 5 hypothesis tests steps!

Five steps of hypothesis testing

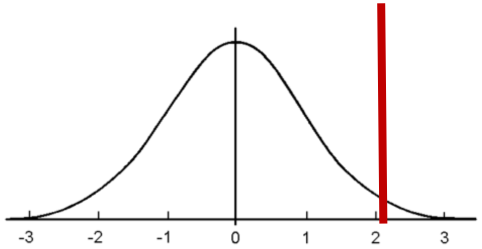


1. State H_0 and H_A
    - Assume Gorgias (H_0) was right

2. Calculate the actual observed statistic

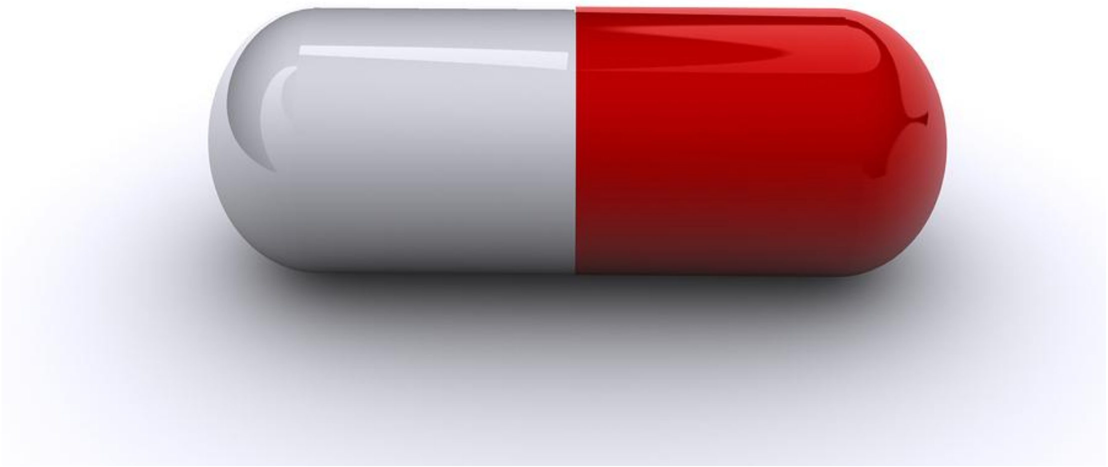
3. Create a **null distribution** of statistics that are consistent with H_0
    - i.e., a distribution of statistics that we would expect if Gorgias is right

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

Hypothesis tests for comparing two means

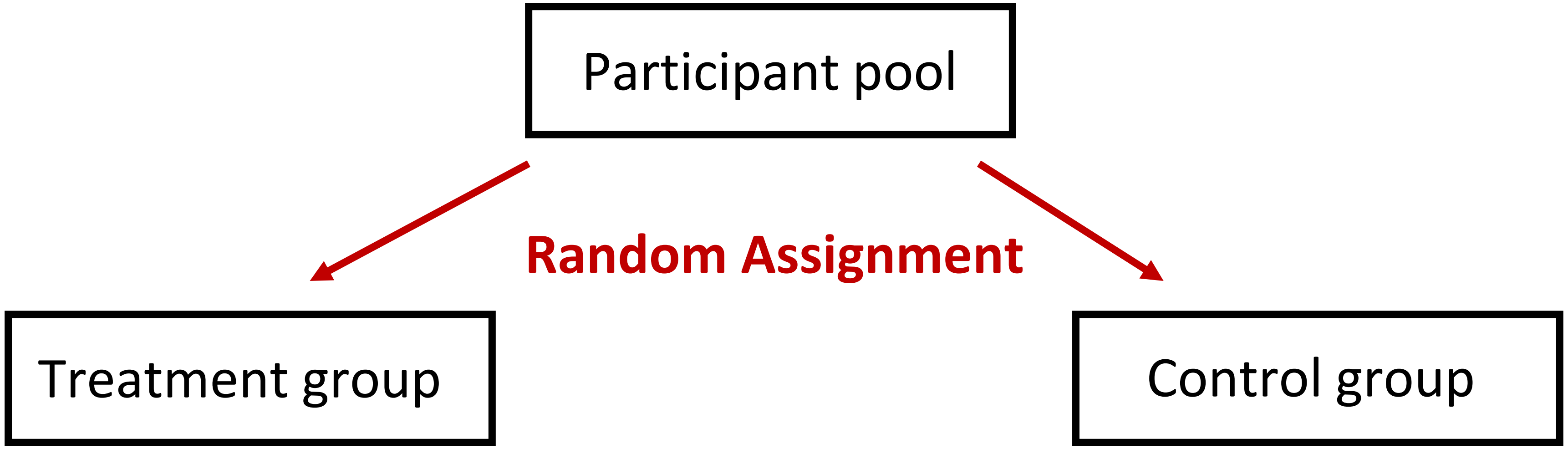


Question: Is this pill effective?

Experimental design: randomized controlled trial

Take a group of participant and ***randomly assign***:

- Half to a *treatment group* where they get the pill
- Half in a *control group* where they get a fake pill (placebo)
- See if there is more improvement in the treatment group compared to the control group



Example: Does calcium reduce blood pressure?

A randomized controlled trial by Lyle et al (1987) investigated whether calcium lowered blood pressure

- A treatment group of 10 men received a calcium supplement for 12 weeks
- A control group of 11 men received a placebo during the same period

The blood pressure of these men was taken before and after the 12 weeks of the study

Hypothesis tests for differences in two group means

1. State the null and alternative hypothesis

H_0: $\mu_{Treatment} = \mu_{Control}$ or $\mu_{Treatment} - \mu_{Control} = 0$

H_A: $\mu_{Treatment} > \mu_{Control}$ or $\mu_{Treatment} - \mu_{Control} > 0$

2. Write down the statistic of interest using appropriate symbols

$\bar{x}_{Effect} = \bar{x}_{Treatment} - \bar{x}_{Control}$

Does calcium reduce blood pressure?

Treatment data (n = 10):

Begin	107	110	123	129	112	111	107	112	136	102
End	100	114	105	112	115	116	106	102	125	104
Decrease	**7**	**-4**	**18**	**17**	**-3**	**-5**	**1**	**10**	**11**	**-2**

Control data (n = 11):

Begin	123	109	112	102	98	114	119	112	110	117	130
End	124	97	113	105	95	119	114	114	121	118	133
Decrease	**-1**	**12**	**-1**	**-3**	**3**	**-5**	**5**	**2**	**-11**	**-1**	**-3**

2. What is the observed statistic of interest?

- $\bar{x}_{Effect} = 5 - -.2727 = 5.273$

3. What is step 3?

3. Create the null distribution!

Treatment group

Control group

Reconstructed participant pool data under H_0

Shuffle data for random assignment consistent with H_0

Shuffled 'treatment group'

Shuffled 'control group'

One null distribution statistic: $\overline{x}_{Shuff\_Treatment} - \overline{x}_{Shuff\_control}$

3. Create a null distribution

1. Combine data from both groups

2. Shuffle data

3. Randomly select 10 points to be the ‘shuffled’ treatment group

4. Take the remaining points to the ‘shuffled’ control group

5. Compute the statistic of interest on these ‘shuffled’ groups

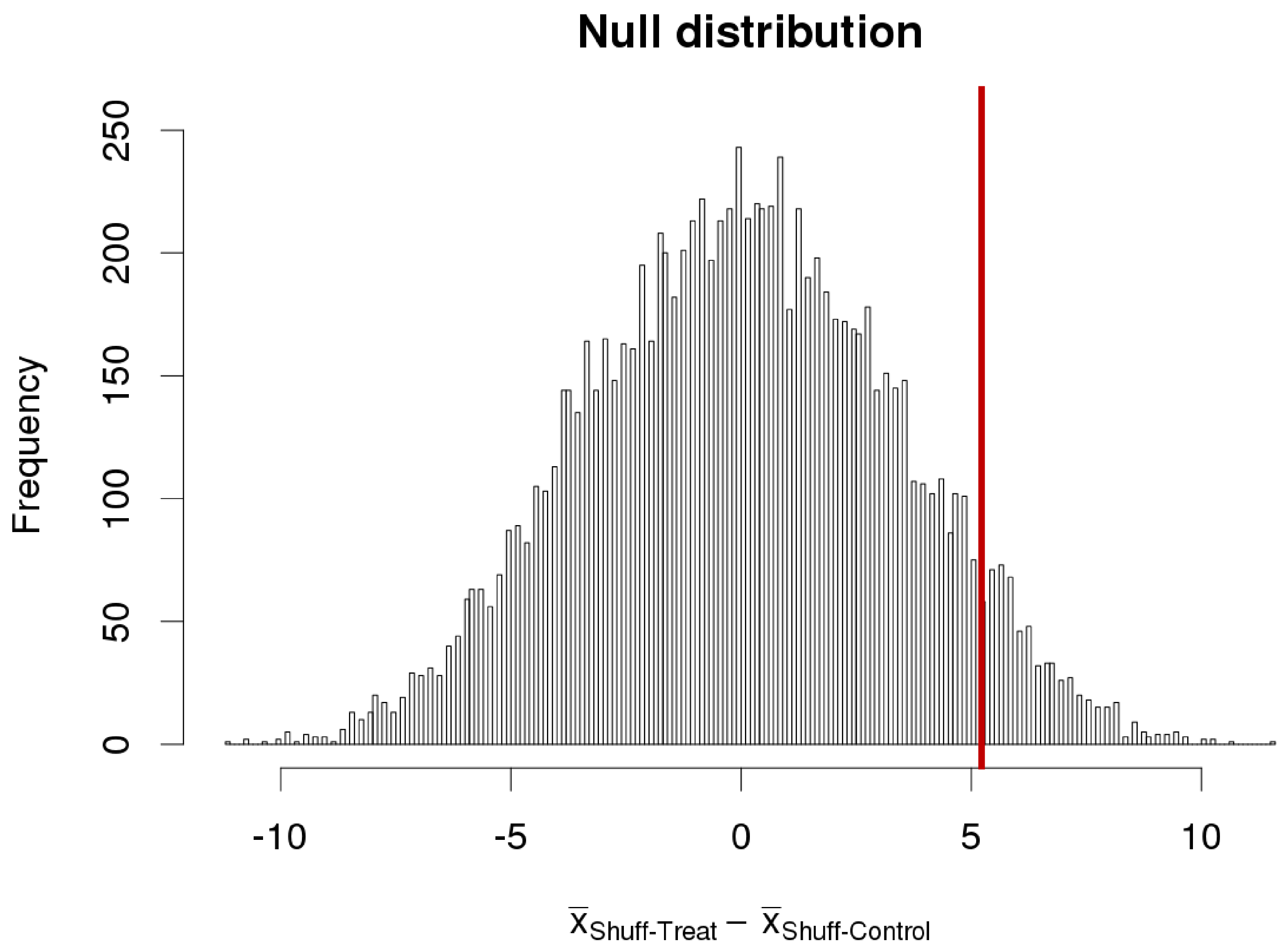
6. Repeat 10,000 times to get a null distribution

3. Creating a null distribution in R

```
# the data from the calcium study
treat <- c(7, -4, 18, 17, -3, -5,  1, 10, 11, -2)
control <- c(-1,  12,  -1,  -3,   3,  -5,   5,   2, -11,  -1,  -3)

# observed statistic
obs_stat <- mean(treat) - mean(control)

#  Combine data from both groups
combined_data <- c(treat, control)
```

3. Creating a null distribution in R

```
null_distribution <-  do_it(10000) * {

    # shuffle data
    shuff_data <- shuffle(combined_data)

    # create fake treatment and control groups
    shuff_treat   <-  shuff_data[1:10]
    shuff_control  <-  shuff_data[11:21]

    # save the statistic of interest
    mean(shuff_treat) - mean(shuff_control)

}
```



hist(null_distribution, breaks = 200)

Next step?

4. Calculate the p-value

```
# Calculate the p-value
    p_value <- pnull(obs_stat, null_distribution, lower.tail = FALSE)
```

p-value = .064

Next step?

5. Are the results statistically significant?

?

Let’s try it in R…

More practice of hypothesis test for two means

Question: Are husbands older than wives?

To address this question, we can use data from a sample of 100 marriage licenses from St. Lawrence County, NY that give the ages of husbands and wives

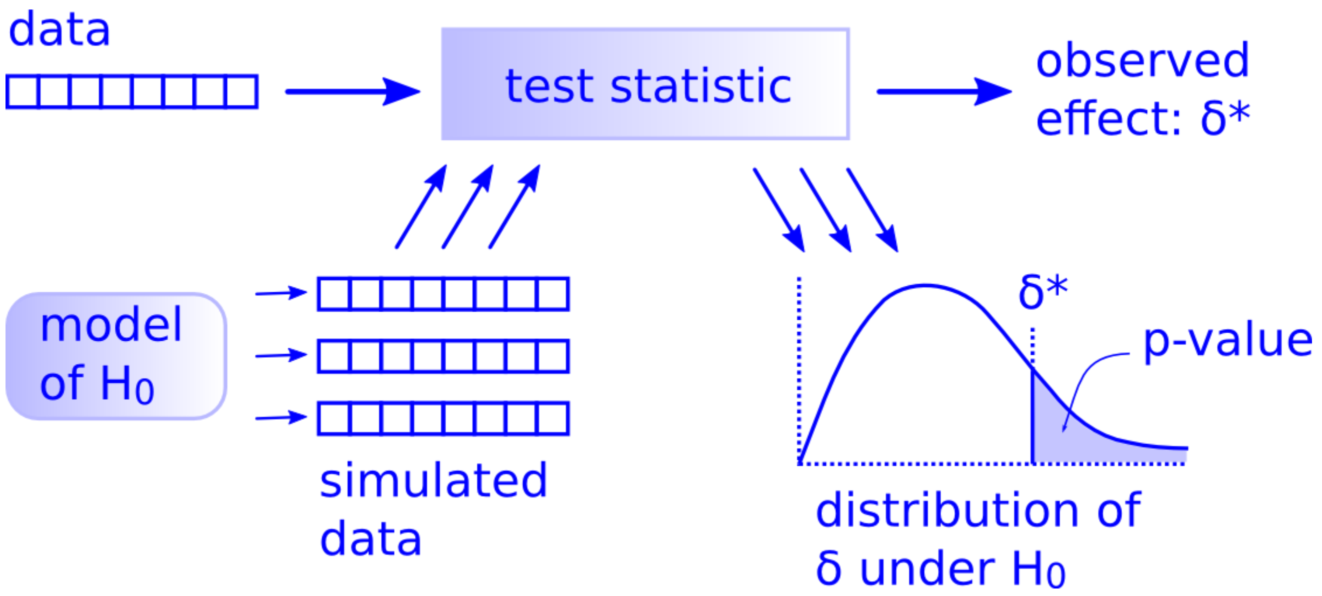
Please run the 5 steps of hypothesis testing to assess if the mean age of husbands is greater than the mean age of wives

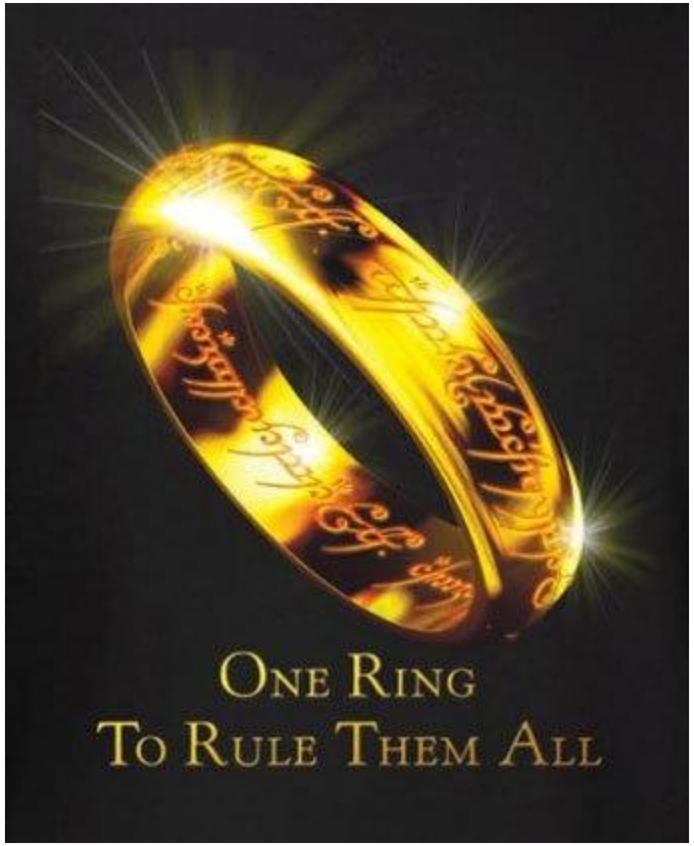
Let’s try it in R!

Hypothesis tests for more than two means

The logic of hypothesis tests…

There is only one hypothesis test!





Just follow the 5 hypothesis tests steps!

Comparing more than two means

A group of Hope College students wanted to see if there was an association between a student's major and the time it takes to complete a small Sudoku-like puzzle

	5	3	2		7			8
6		1	5					2
2			9	1	3		5	
7	1	4	6	9	2			
	2						6	
			4	5	1	2	9	7
	6		3	2	5			9
1					6	3		4
8			1		9	6	7	

Comparing more than two means

A group of Hope College students wanted to see if there was an association between a student's major and the time it takes to complete a small Sudoku-like puzzle

They grouped majors into four categories

- Applied science (as)
- Natural science (ns)
- Social science (ss)
- Arts/humanities (ah)

What is the first step of hypothesis testing?

Sudoku by field

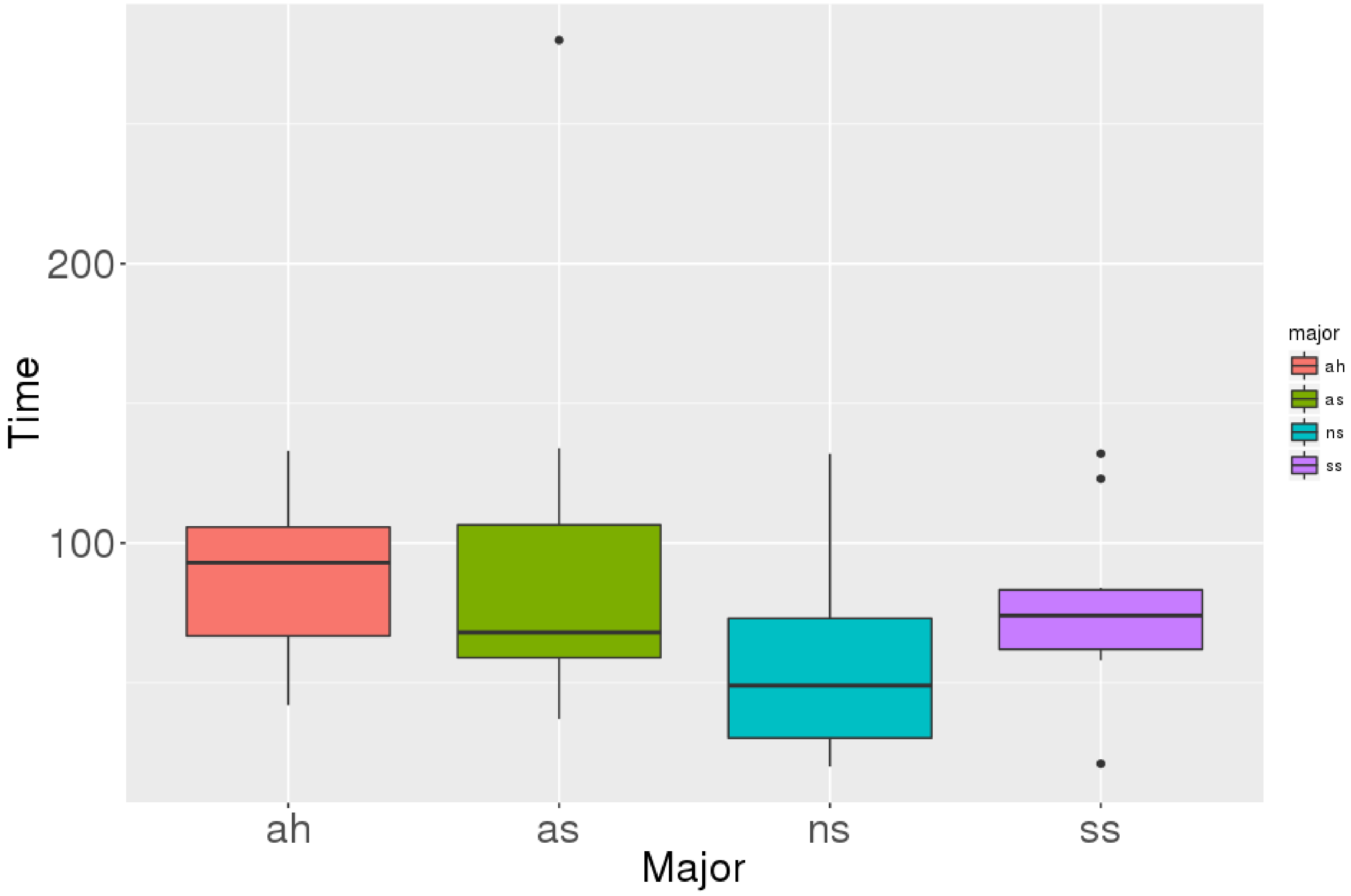
1. State the null and alternative hypotheses!

H_0: $\mu_{as} = \mu_{ns} = \mu_{ss} = \mu_{ah}$
H_A: $\mu_i \neq \mu_j$ for one pair of fields of study

What should we do next?

Let's plot the data first…

Step 2a: Plot of completion time by major



What should we do next?

Sudoku by field

1. State the null and alternative hypotheses!

H_0: $\mu_{as} = \mu_{ns} = \mu_{ss} = \mu_{ah}$
H_A: $\mu_i \neq \mu_j$ for one pair of fields of study

Thoughts on the statistic of interest?

Comparing multiple means

There are many possible statistics we could use. A few choices are:

1. Group range statistic:

   max $\bar{x}$ - min $\bar{x}$

2. Mean absolute difference (MAD):

   $(|\bar{x}_{as} - \bar{x}_{ns}| + |\bar{x}_{as} - \bar{x}_{ss}| + |\bar{x}_{as} - \bar{x}_{ah}| + |\bar{x}_{ns} - \bar{x}_{ss}| + |\bar{x}_{ns} - \bar{x}_{ah}| + |\bar{x}_{ss} - \bar{x}_{ah}|)/6$

3. F statistic:

$$F = \frac{\text{between-group variability}}{\text{within-group variability}} = \frac{\frac{1}{K-1}\sum_{i=1}^{K} n_i(\bar{x}_i - \bar{x}_{tot})^2}{\frac{1}{N-K}\sum_{i=1}^{K}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2}$$

Using the MAD statistic

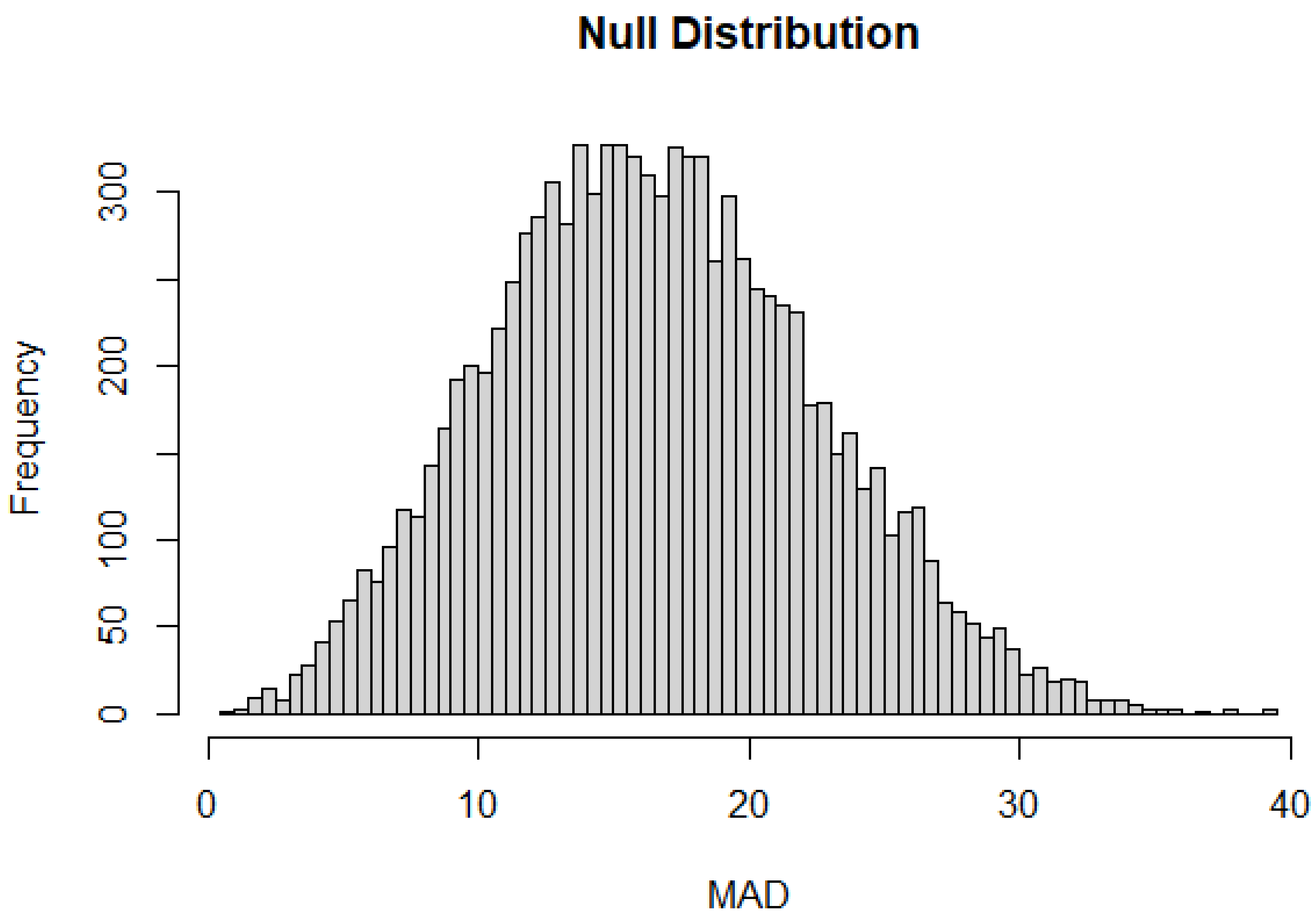
<u>Mean absolute difference (MAD):</u>

$(|\bar{x}_{as} - \bar{x}_{ns}| + |\bar{x}_{as} - \bar{x}_{ss}| + |\bar{x}_{as} - \bar{x}_{ah}| + |\bar{x}_{ns} - \bar{x}_{ss}| + |\bar{x}_{ns} - \bar{x}_{ah}| + |\bar{x}_{ss} - \bar{x}_{ah}|)/6$

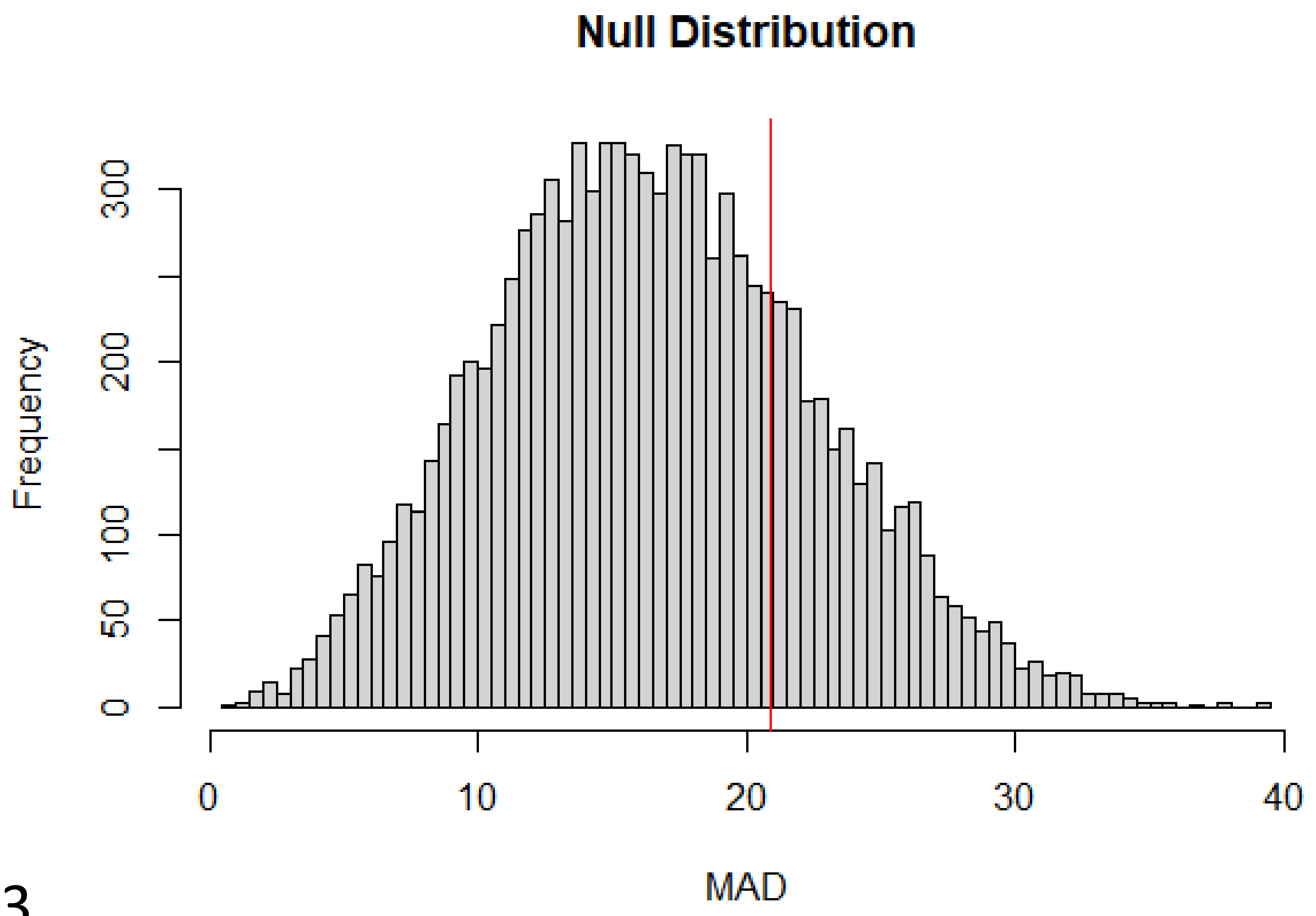
Observed statistic value = 20.88

How can we create the null distribution?

Null distribution



P-value



p-value = .233

Conclusions?

Hypothesis tests for more than two means in R

Step 1: null and alternative hypotheses…

H_0: $\mu_{as} = \mu_{ns} = \mu_{ss} = \mu_{ah}$

H_A: $\mu_i \neq \mu_j$ for one pair of fields of study

	5	3	2		7			8
6		1	5					2
2			9	1	3		5	
7	1	4	6	9	2			
	2						6	
			4	5	1	2	9	7
	6		3	2	5			9
1					6	3		4
8			1		9	6	7	

Let’s try this analysis in R…

```r
# get the data
sudoku_data <- read.table("MajorPuzzle.txt", header = TRUE)

# Extract vectors from the data frame  (how do we do this?)
completion_times <- sudoku_data$time
majors <- sudoku_data$major
```

Visualize the data

How can we visualize the data?

```
# We can create side-by-side boxplots using
boxplot(completion_times ~ majors,
        xlab = "Major",   ylab = "Time (s)")
```

Calculating the statistic of interest

We can get the MAD statistic using the get_MAD_stat() function

get_MAD_stat(data_vector, grouping_vector)

- data_vector: a vector of quantitative data
- grouping_vector: a vector indicating which group the quantitative data is in

Can you get the MAD statistic for the sudoku data?

```
obs_stat <-  get_MAD_stat(completion_time, major)
```

Creating the null distribution

Q: How could we create one point in a null distribution?

- A: Shuffle the grouping_vector (major vector) and calculate the MAD statistic

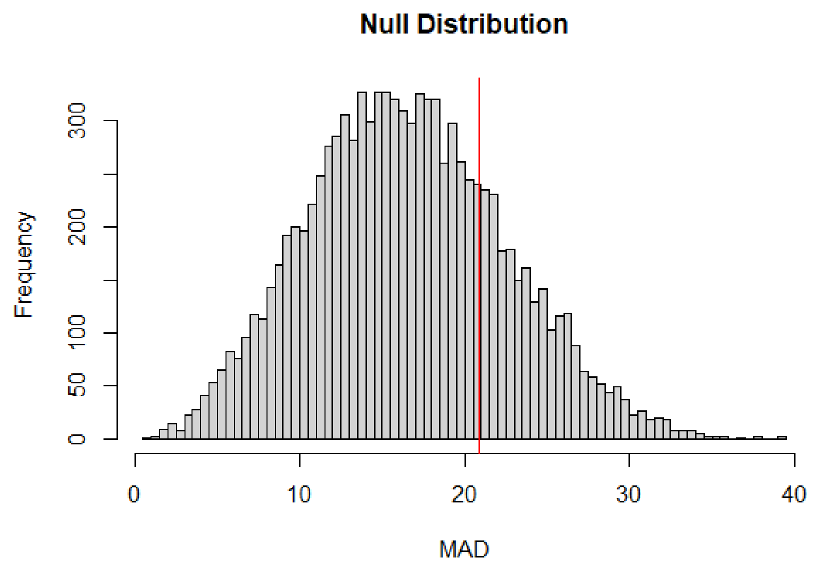
Q: How can we do this in R?

```
shuffled_majors <- shuffle(major)

get_MAD_stat(completion_time, shuffled_majors)
```

Creating the null distribution

Q: How can we create a full null distribution?

```
null_dist <- do_it(10000) * {
    shuffled_majors <- shuffle(majors)
    get_MAD_stat(completion_times, shuffled_majors)
}

# visualize the null distribution
hist(null_dist, breaks = 200)
abline(v = obs_stat, col = "red")
```



Steps 4 and 5

Let's try it in R!

Q: What do we do next and how do we do it?

- A: We get the p-value

```
pnull(obs_stat, null_dist, lower.tail = FALSE)
```