Randomization tests continued
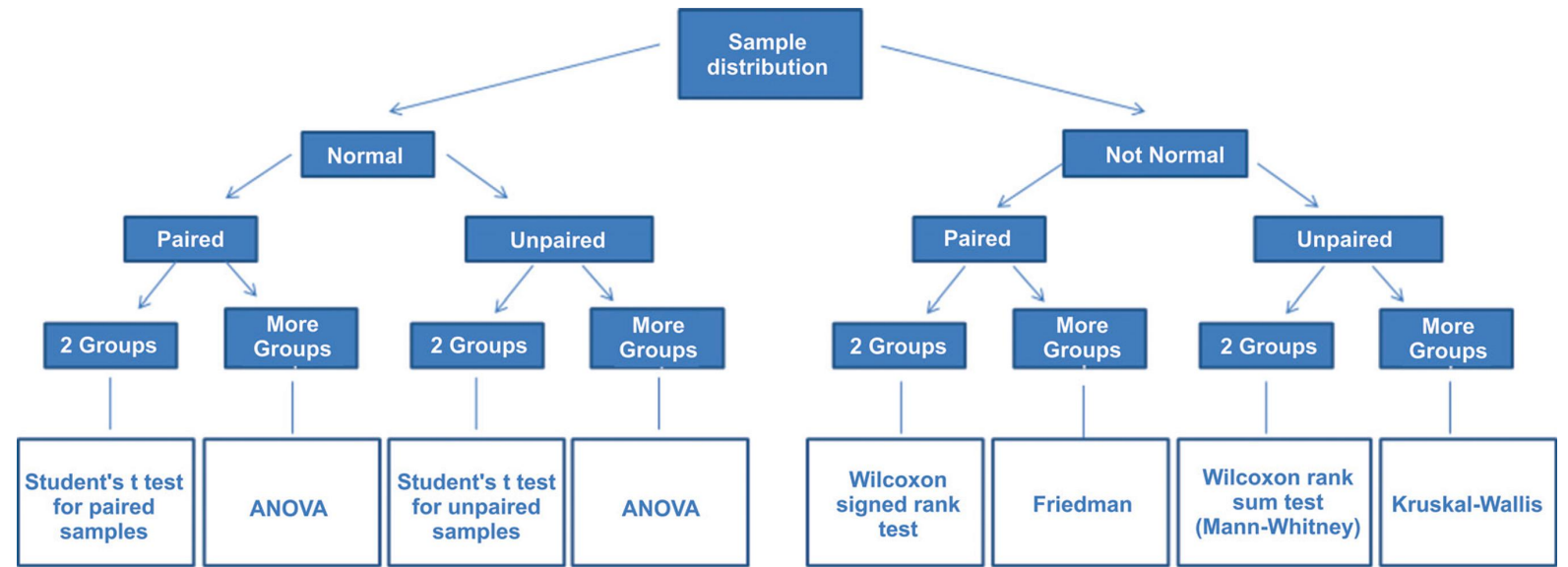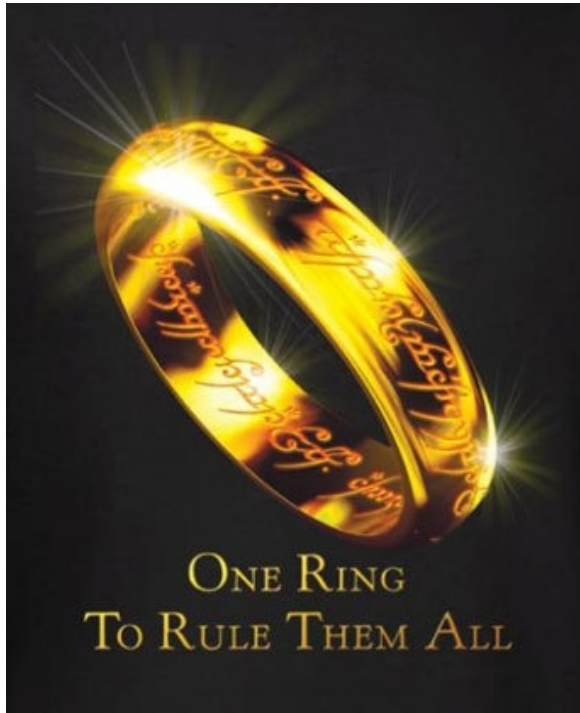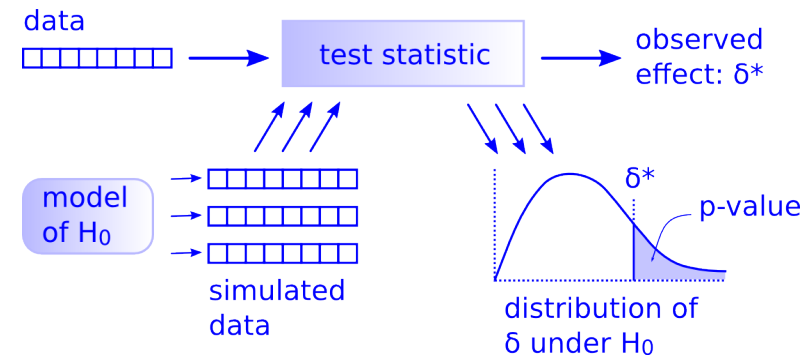
Overview

Randomization tests for two means

Randomization tests for more than two means

Theories of hypothesis testing (if there is time)

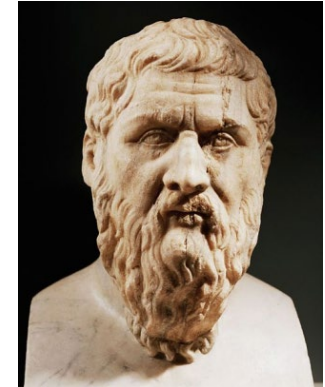# The big picture: There is only one hypothesis test!
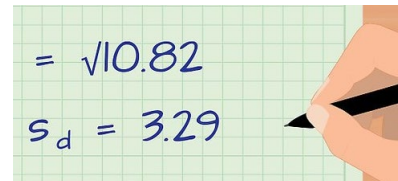

Just need to follow 5 steps!

Five steps of hypothesis testing

1. State H_0 and H_A
   - Assume Gorgias (H_0) was right
   - α = .05 of the time he will be right, but we will say he is wrong

2. Calculate the actual observed statistic
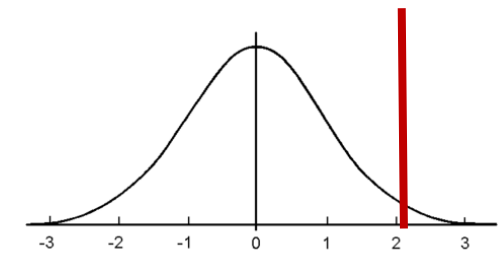
$$= \sqrt{10.82}$$
$$s_d = 3.29$$

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

Review: hypothesis test for a single proportion

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 shirts by people who did not.

Joy identified 11 out of the 12 shirts correctly.

Step 1: state the null and alternative hypotheses
- H_0: $\pi = 0.5$
- H_A: $\pi > 0.5$ ← H_0 and H_A need to be mutually exclusive

Review: hypothesis test for a single proportion

We can run a hypothesis test for a single proportion in R using:

obs_stat <- 11/12   # Step 2: calculate the observed statistic

flip_sims_prop <- rbinom(10000, 12, .5)/12   # Step 3: create null distribution

p_value <- sum(flip_sims_prop >= obs_stat)/length(flip_sims)   # Step 4: p-value

p-value is   0.0029   Step 5: Should we reject H_0?

Do you really believe Joy can smell Parkinson's disease?

Her Incredible Sense Of Smell Is Helping Scientists Find New Ways To Diagnose Disease

March 23, 2020 · 4:45 PM ET

Observational and experimental studies

An **experiment** is a study in which the researcher actively controls one or more of the explanatory variables

- Allows one to get at questions of **causation**!

An **observational study** is a study in which the researcher does not actively control the value of any variable but simply observes the values as they naturally exist

Question: Is the smelling Parkinson's disease study experimental or observational?
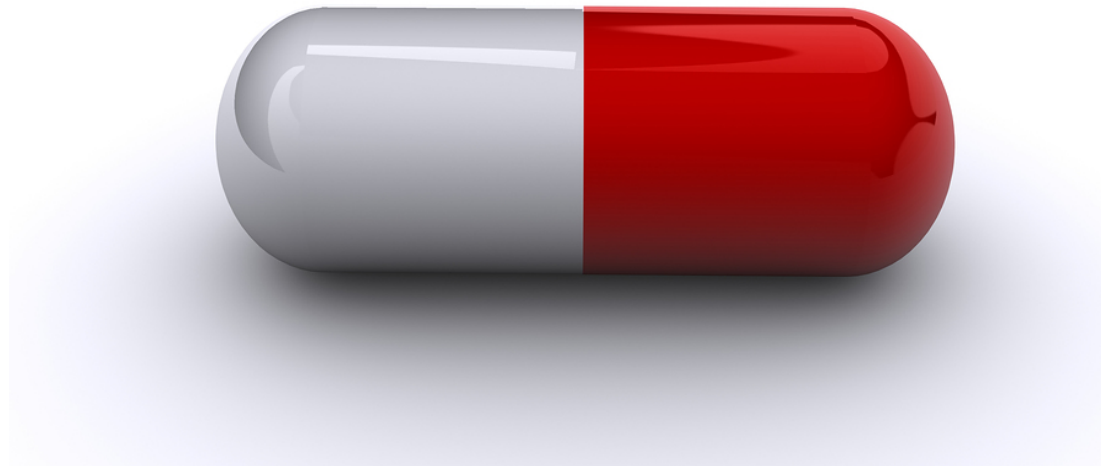
Hypothesis tests comparing 2 means

The big picture: There is only one hypothesis test!
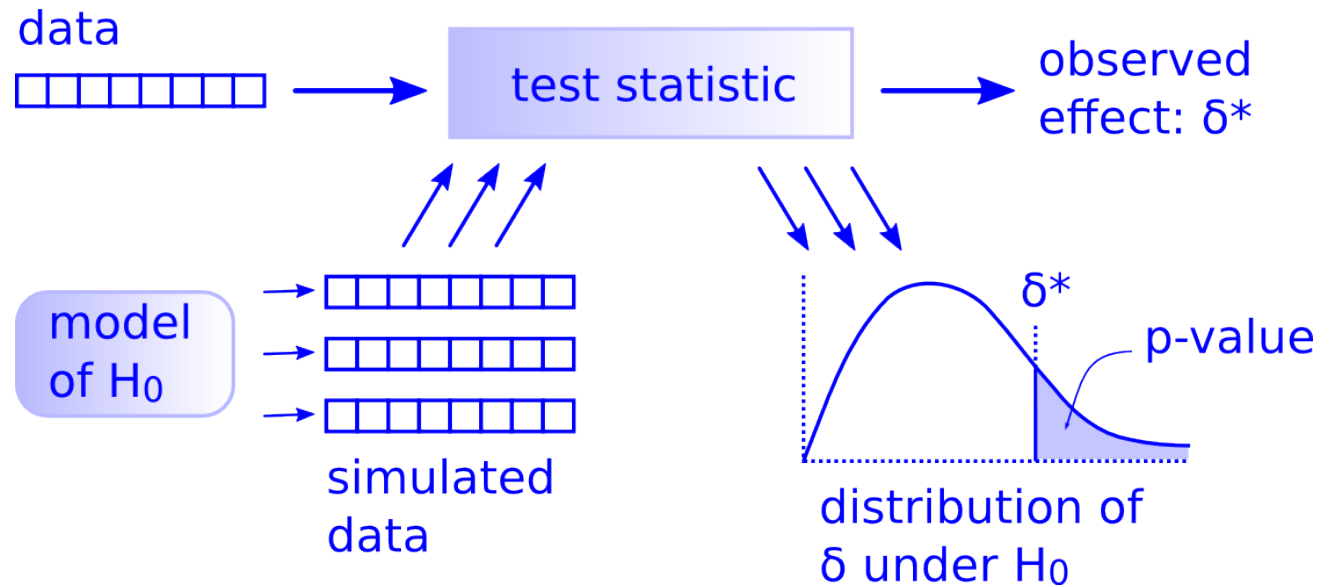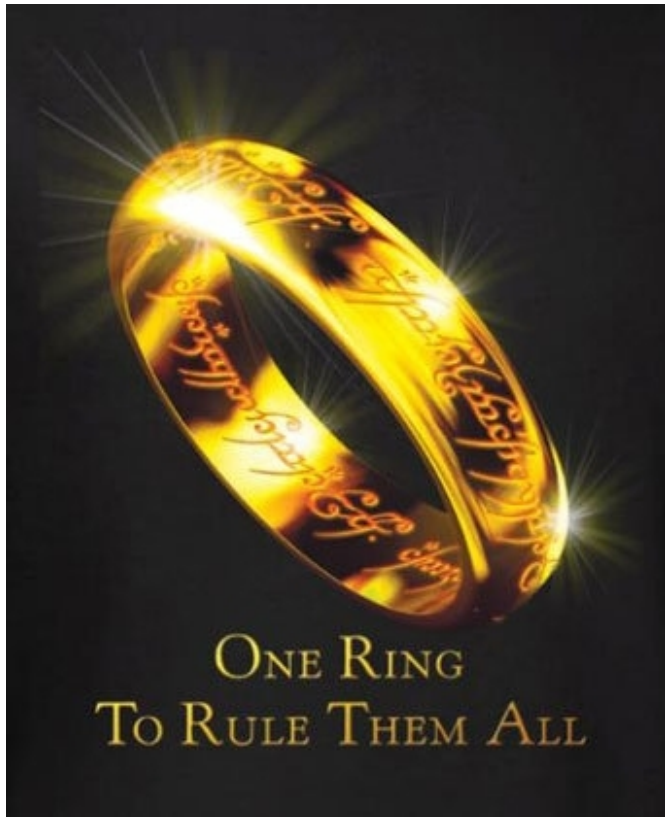
Just need to follow 5 steps!

Hypothesis tests for comparing two means

Question: Is this pill effective?

Testing whether a pill is effective

1. How would we design a study?

2. What would the cases and variables be?

3. What would the parameter and statistic of interest be?

4. What are the null and alternative hypotheses?
   - Assume we are looking for differences in means between the groups

Experimental design

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
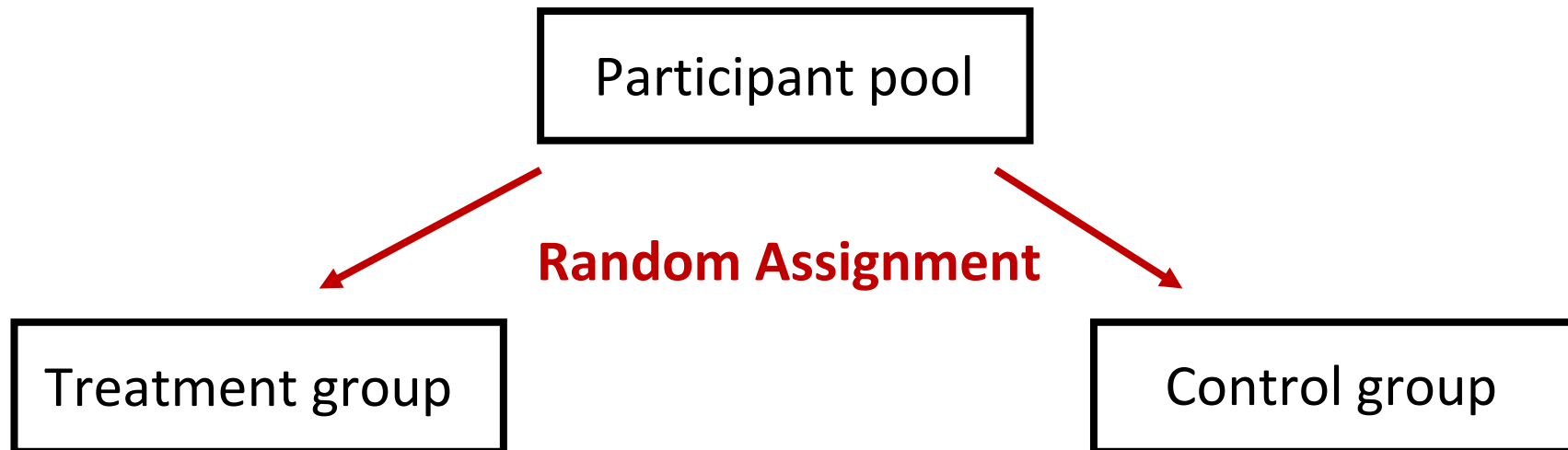
Hypothesis tests for differences in two group means

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. Calculate statistic of interest
   - $\overline{x}_{Effect} = \overline{x}_{Treatment} - \overline{x}_{Control}$

Example: Does calcium reduce blood pressure?

A randomized comparative experiment 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

1. What are the null and alternative hypotheses?
- H_0: $\mu_{Treatment} = \mu_{Control}$ or $\mu_{Treatment} - \mu_{Control} = 0$
- H_A: $\mu_{Treatment} > \mu_{Control}$ or $\mu_{Treatment} - \mu_{Control} > 0$
  - i.e., a <u>greater decrease</u> in blood pressure after taking calcium

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?

- $\overline{x}_{Effect}$ = 5 - -.2727 = 5.273
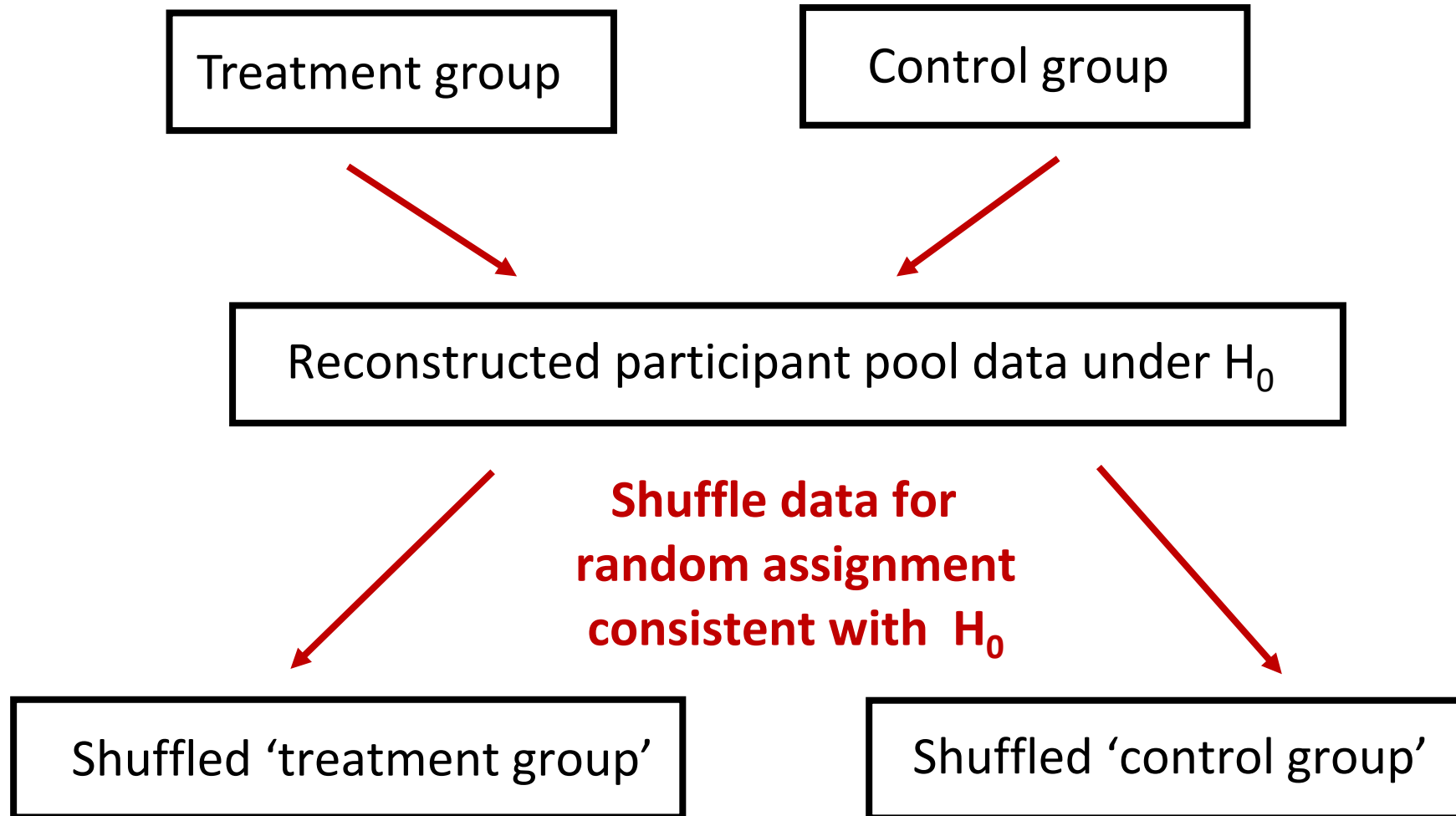
3. What is step 3?

3. Create the null distribution!

How could we create the null distribution?

Need to generate data consistent with H_0: $\mu_{Treatment} - \mu_{Control} = 0$
- i.e., we need fake $\overline{x}_{Effect}$ that are consistent with H_0

Any ideas how we could do this?

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 'null' treatment group

4. Take the remaining 11 points to the 'null' control group

5. Compute the statistic of interest on these 'null' groups

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

Let's try the rest of the hypothesis test in R…

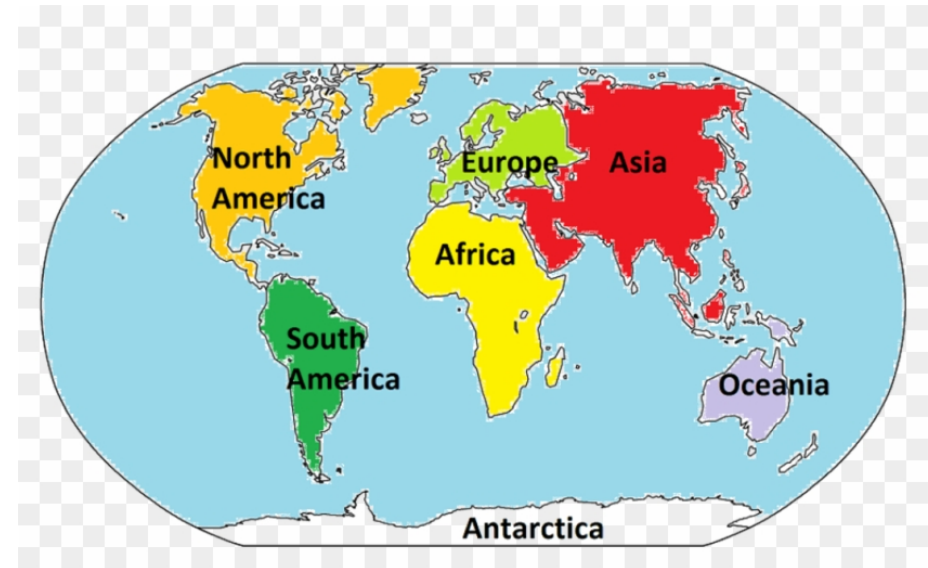# Hypothesis test for comparing more than two means

The big picture: There is only one hypothesis test!
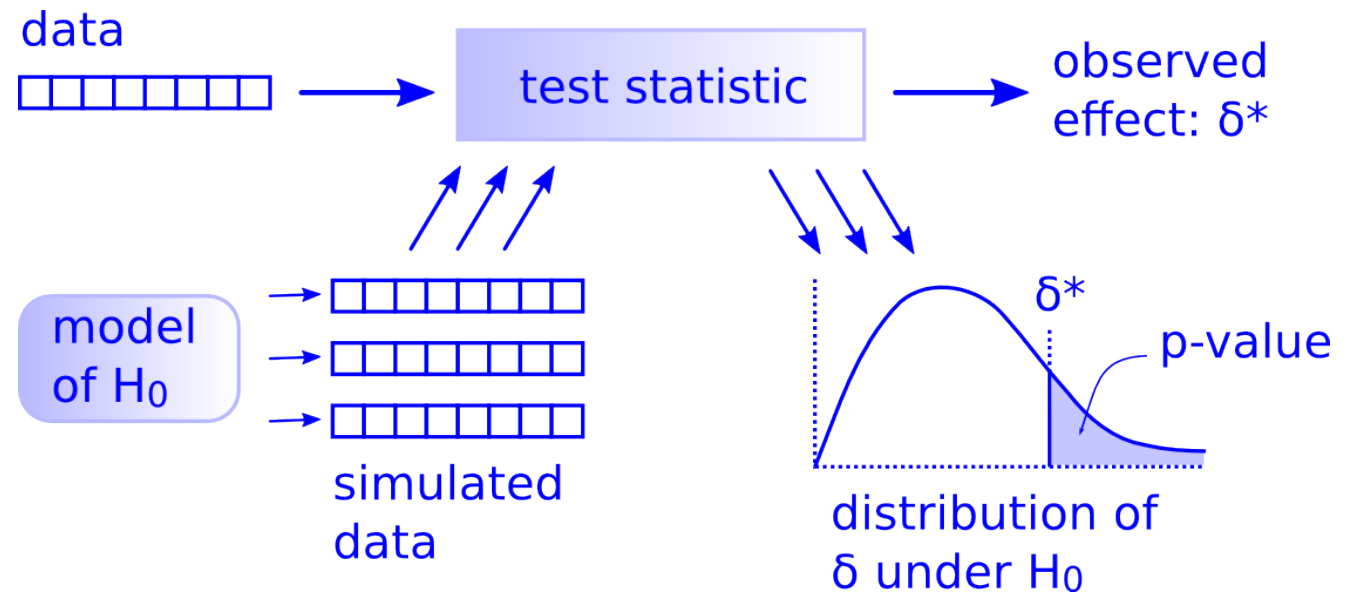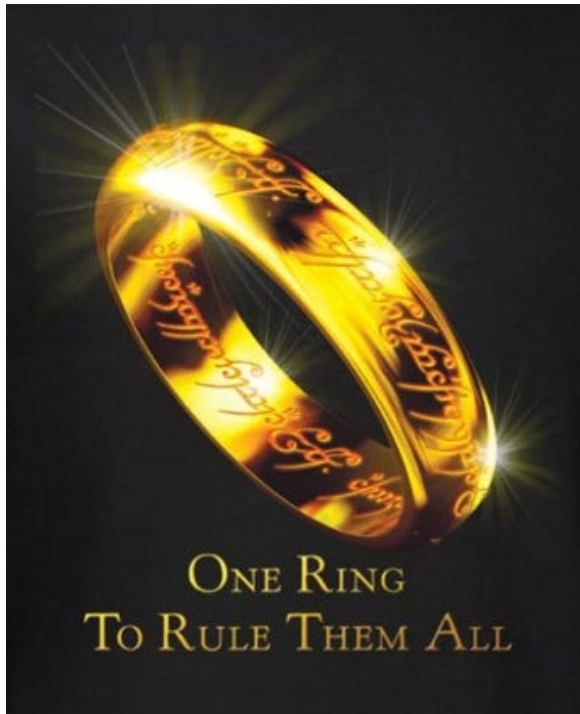
Just need to follow 5 steps!

Comparing more than two means

Let's examine the beer consumption in different continents!

Analysis inspired by:
- Minitab blog article
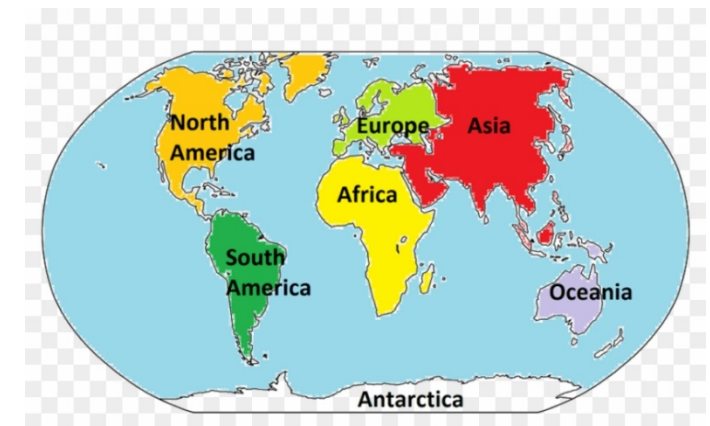- Five thirty eight analysis

Question: Does the average beer consumption in countries different depending on the continent?

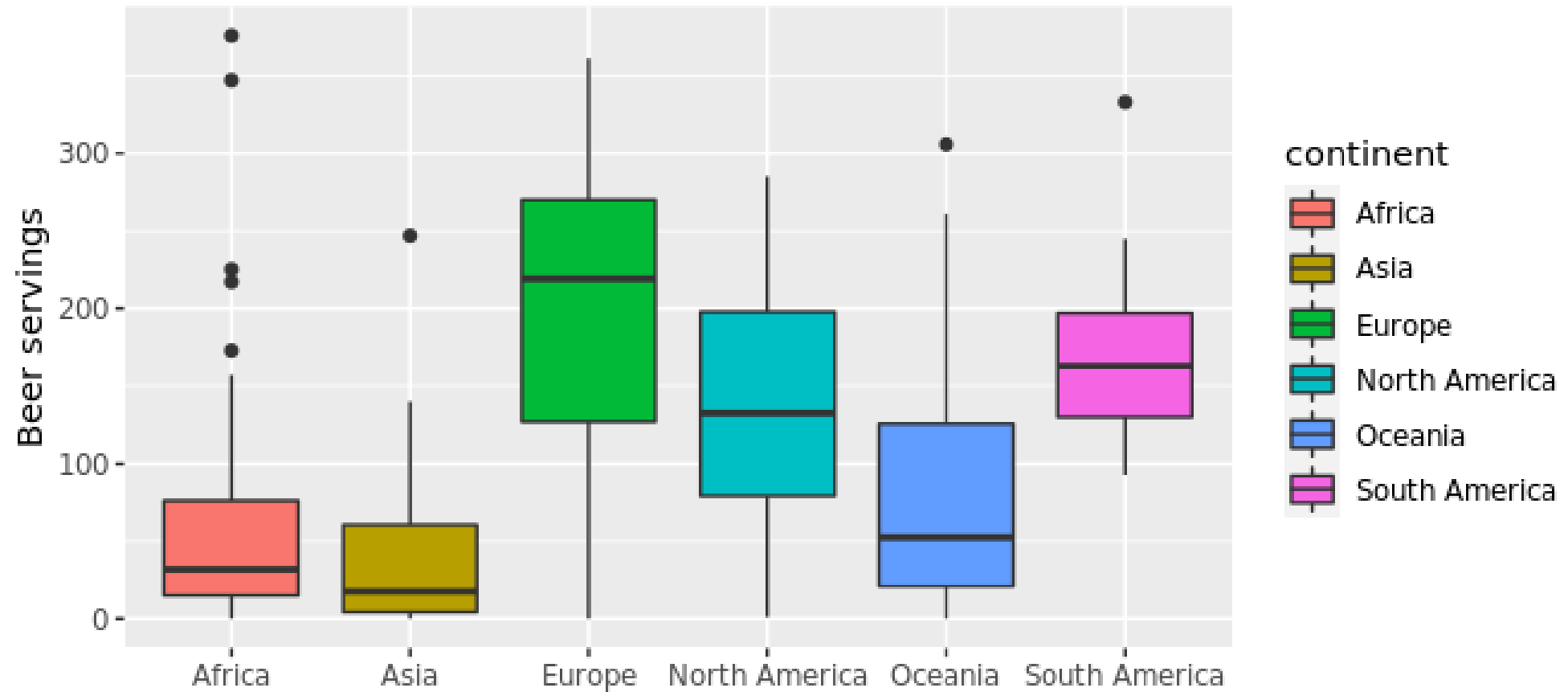1. State the null and alternative hypotheses!

H_0: $\mu_{Asia} = \mu_{Europe} = \mu_{Africa} = \mu_{North-America} = \mu_{South-America} = \mu_{Oceania}$

H_A: $\mu_i \neq \mu_j$ for at least one pair of fields of continents

What should we do next?

Plot of the beer consumption in different continents

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 $\overline{x}$ - min $\overline{x}$

2. Mean absolute difference (MAD):

   $(|\overline{x}_{Africa} - \overline{x}_{Asia}| + |\overline{x}_{Africa} - \overline{x}_{Europe}| + ... + |\overline{x}_{Oceania} - \overline{x}_{South-America}|)/15$

3. F statistic:

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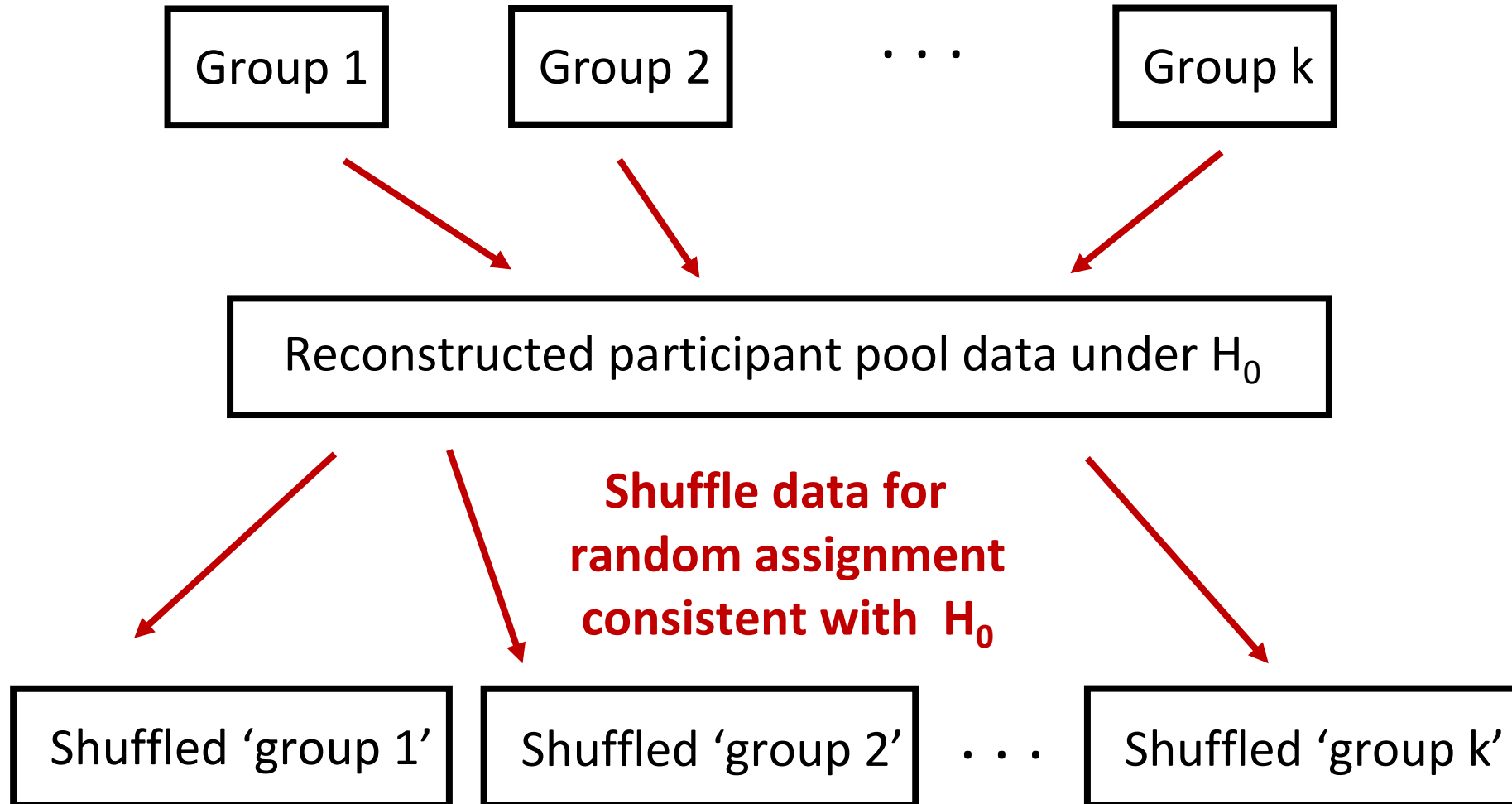
Using the MAD statistic

Mean absolute difference (MAD):

$$(|\bar{x}_{Africa} - \bar{x}_{Asia}| + |\bar{x}_{Africa} - \bar{x}_{Europe}| + \ldots + |\bar{x}_{Oceania} - \bar{x}_{South-America}|)/15$$
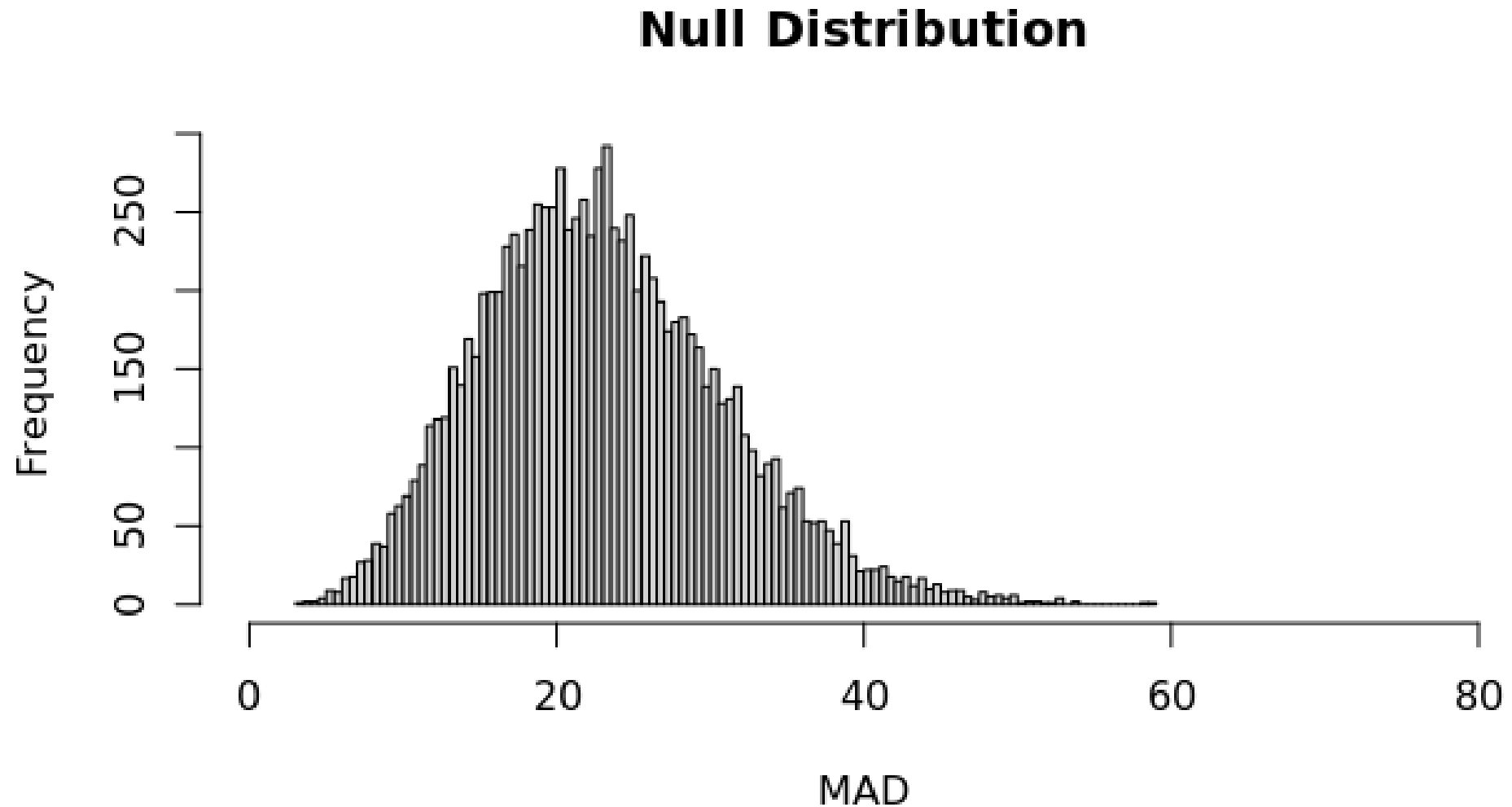
Observed statistic value = 78.86

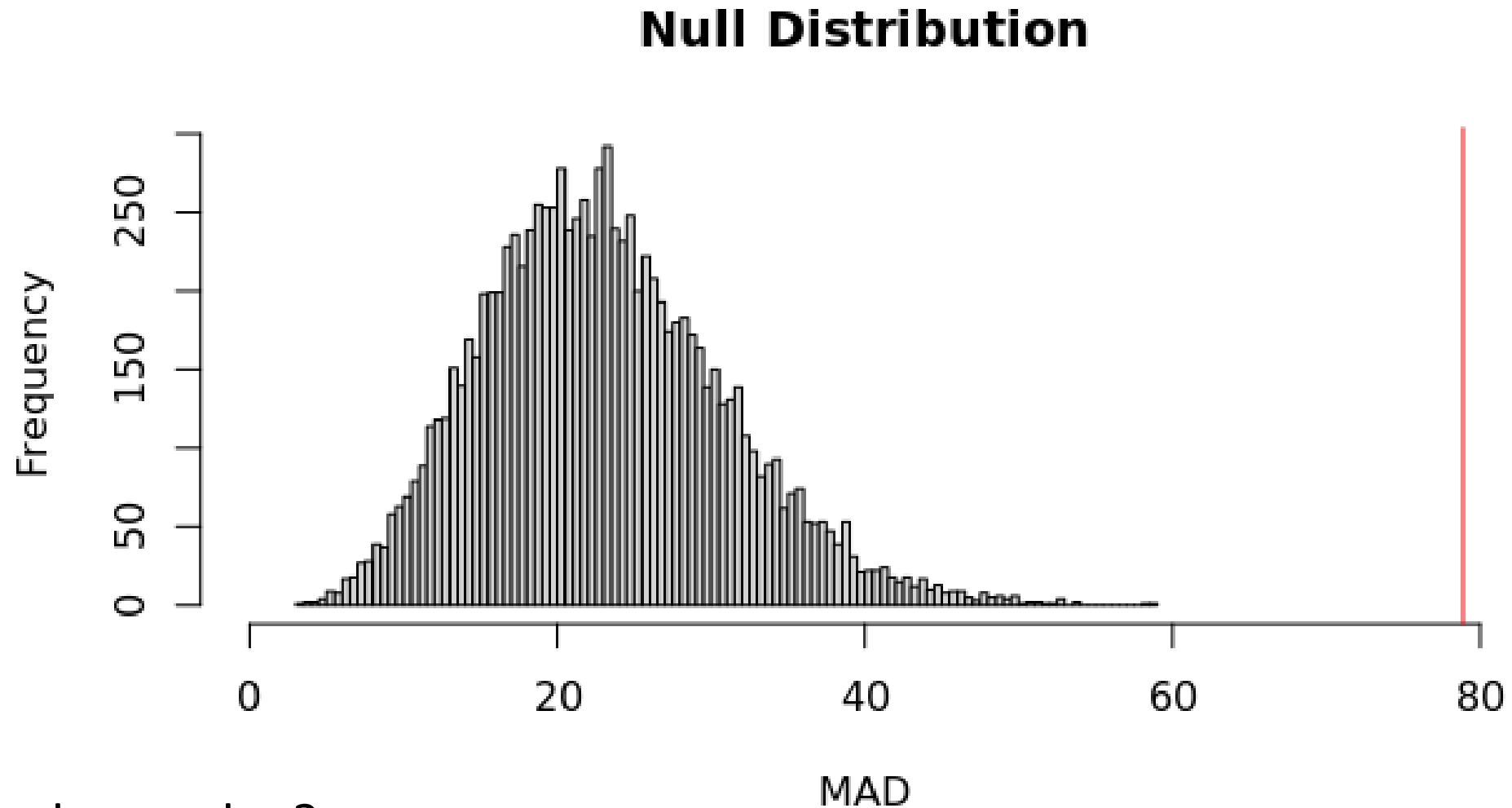How can we create the null distribution?

3. Create the null distribution!

Compute statistics from shuffled groups

3. Create the null distribution!

4. Calculate the p-value

Null Distribution

What is the p-value?

Conclusions?

"TRUST ME, YOU CAN DANCE."
- BEER

Let's try it in R...