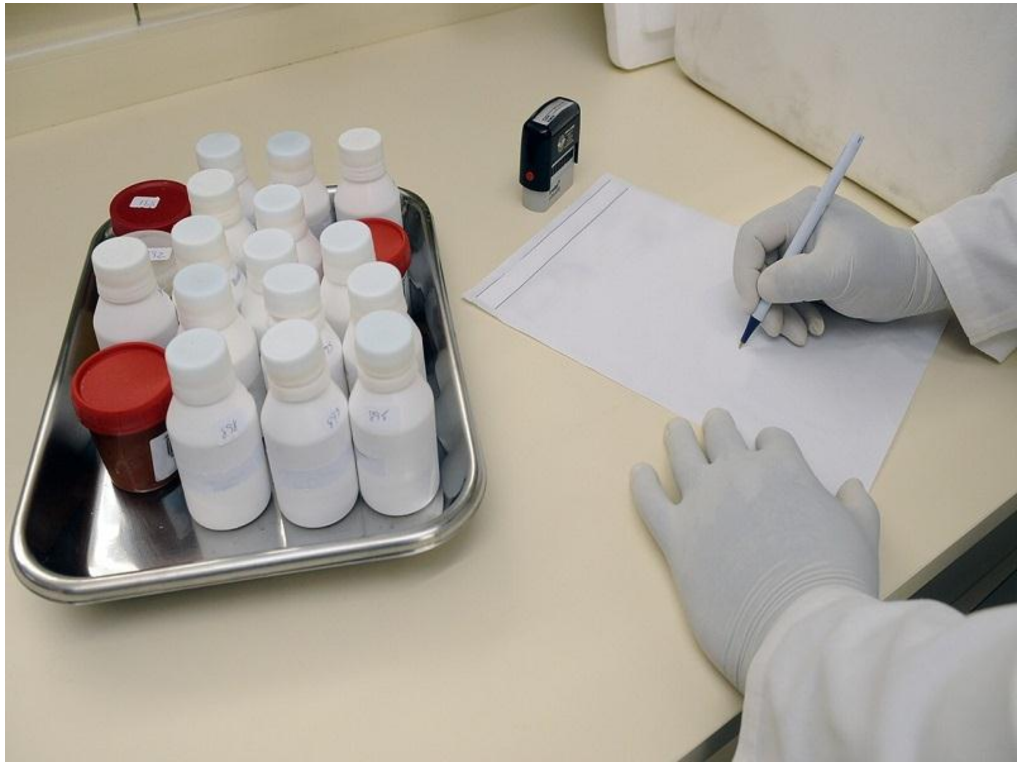
Hypothesis tests for two or more means



Overview

Worksheet 8 questions

Review of hypothesis testing for a single proportion

Hypothesis tests for two means

Hypothesis tests for more than 2 means

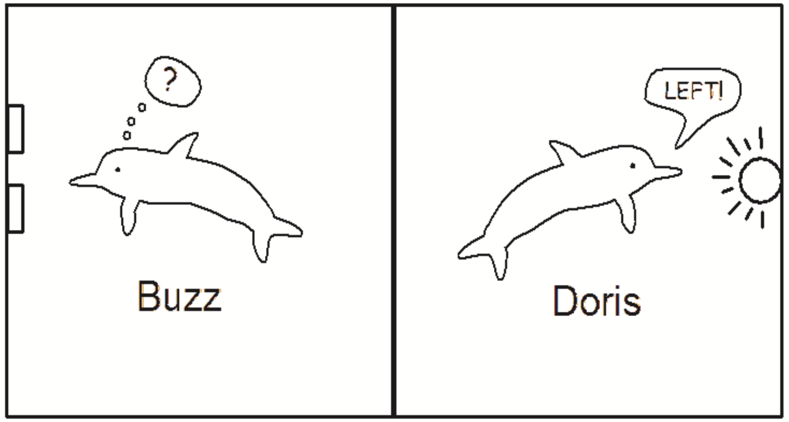
Worksheet 8 questions?

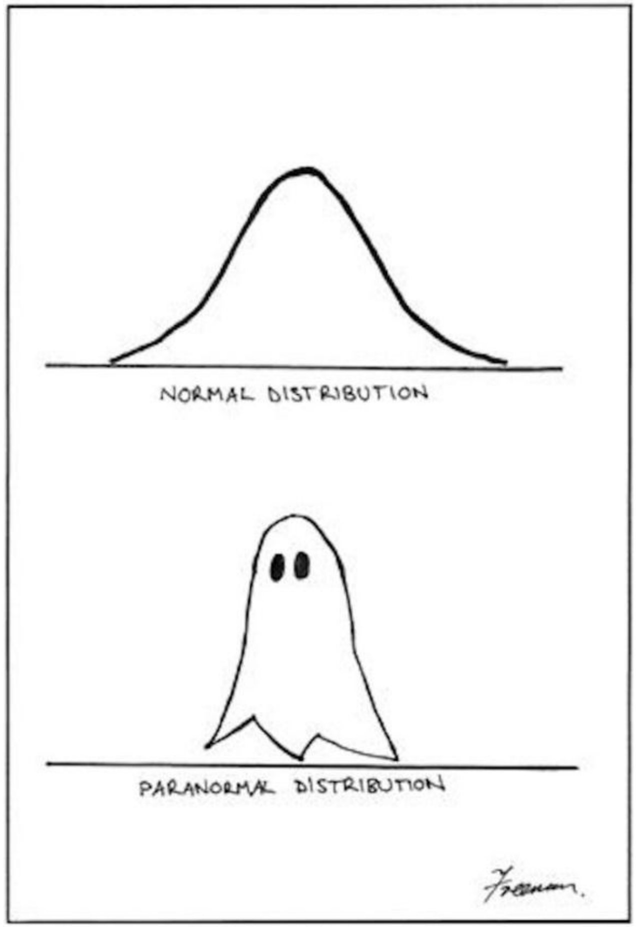
How did the Lock5 questions go?

Worksheet 8 question: Is Paul psychic?

Why/why not?

Review: hypothesis tests for a single proportion





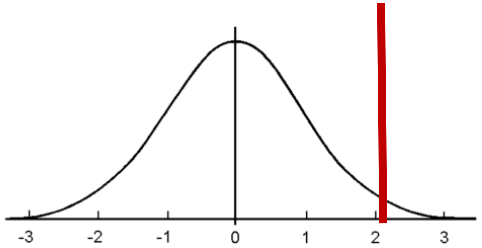
[Movie 1] [Movie 2]

Five steps of hypothesis testing

1. Assume innocence: H_0 is true
   - State H_0 and H_A

2. Gather evidence
   - Calculate the observed statistic

3. Create a distribution of what evidence would look like if H_0 is true
   - Null distribution



4. Assess the probability that the observed evidence would come from the null distribution
   - p-value

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

Adult persistence of head-turning asymmetry

Background:

- Most people are right handed, right eye dominant, etc.
- Biologists have suggested that human embryos tend to turn their heads to the right as well.

German bio-psychologist Onur Güntürkün conjectured that this tendency manifests itself in other ways, so he studies which ways people turn their heads when they kiss.

Adult persistence of head-turning asymmetry

He and his researchers observed kissing couples in public places and noted whether the couple leaned their heads to the right or left.

They observed 124 couples, ages 13-70 years.

Adult persistence of head-turning asymmetry

Please write down answers to these questions:

1. What are the observational units?

2. What are the variables (categorical or quantitative)?

3. What is Onur’s conjecture? How would you state the null and alternative hypothesis in words and in symbols?

brief communications

Adult persistence of head-turning asymmetry

A neonatal right-side preference makes a surprising romantic reappearance later in life.

A preference in humans for turning the head to the right, rather than to the left, during the final weeks of gestation and for the first six months after birth[1,2] constitutes one of the earliest examples of behavioural asymmetry and is thought to influence the subsequent development of perceptual and motor preferences by increasing visual orientation to the right side[3,4]. Here I show that twice as many adults turn their heads to the right as to the left when kissing, indicating that this head-motor bias persists into adulthood. My finding may be linked to other forms of sidedness (for example, favouring the right foot, ear or

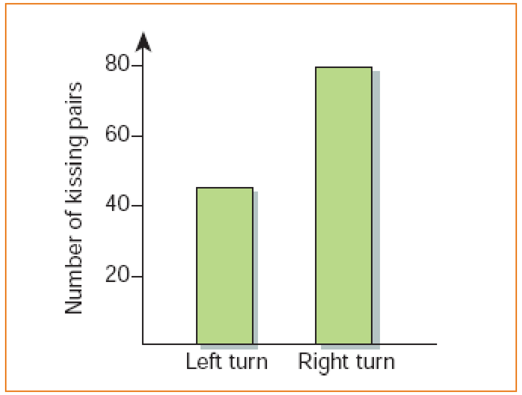


Figure 1 The number of couples who turn their heads to the right rather than to the left when kissing predominates by almost 2:1 (64.5%: 35.5%; $n = 124$ couples).

Of the 124 couples observed, 80 leaned their heads to the right while kissing

- Run a hypothesis test by going through the 5 steps….

Adult persistence of head-turning asymmetry

1. $H_0: \pi = 0.5$ $H_A: \pi > 0.5$
2. $\hat{p} = 80/124 = .64$

10,000 124 0.5

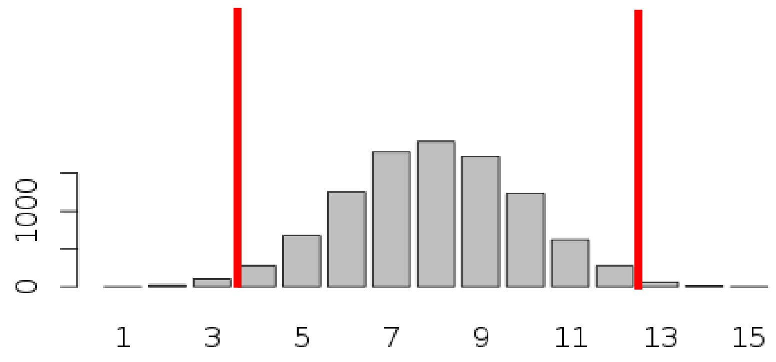
3. null_dist <- rbinom(num_sims, size, prob)/size
4. p_value <- sum(null_dist >= obs_stat)/num_sims # p-value = 0.0007
5. Decision?

One-tailed vs. two-tailed

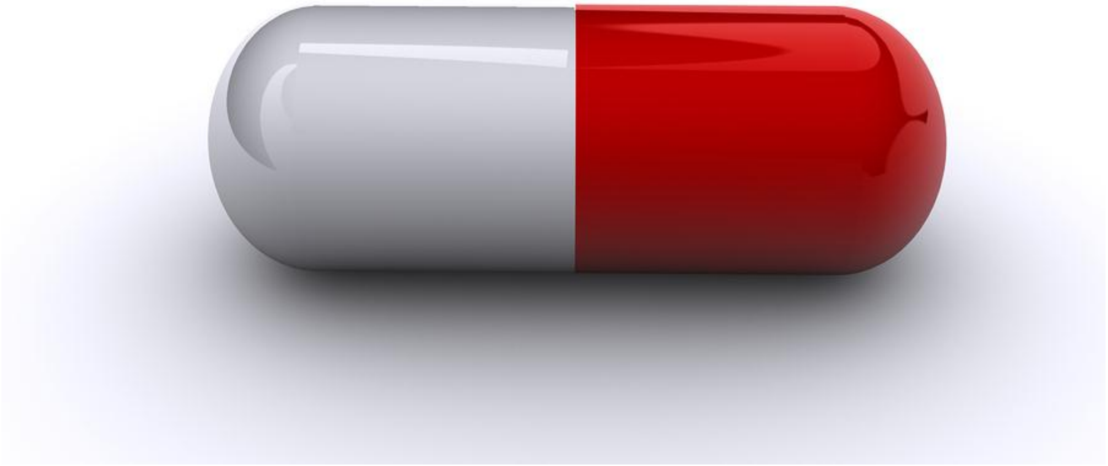
In two-tailed the parameter could be more or less extreme than the hypothesized null parameter

$H_0: \pi = 0.25$ $\qquad\qquad$ $H_A: \pi \neq 0.25$

<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



Hypothesis tests for comparing two means



Question: Is this pill effective?

Testing whether a pill is effective

How would we design a study?

What would the cases and variables be?

What would the statistic of interest be?

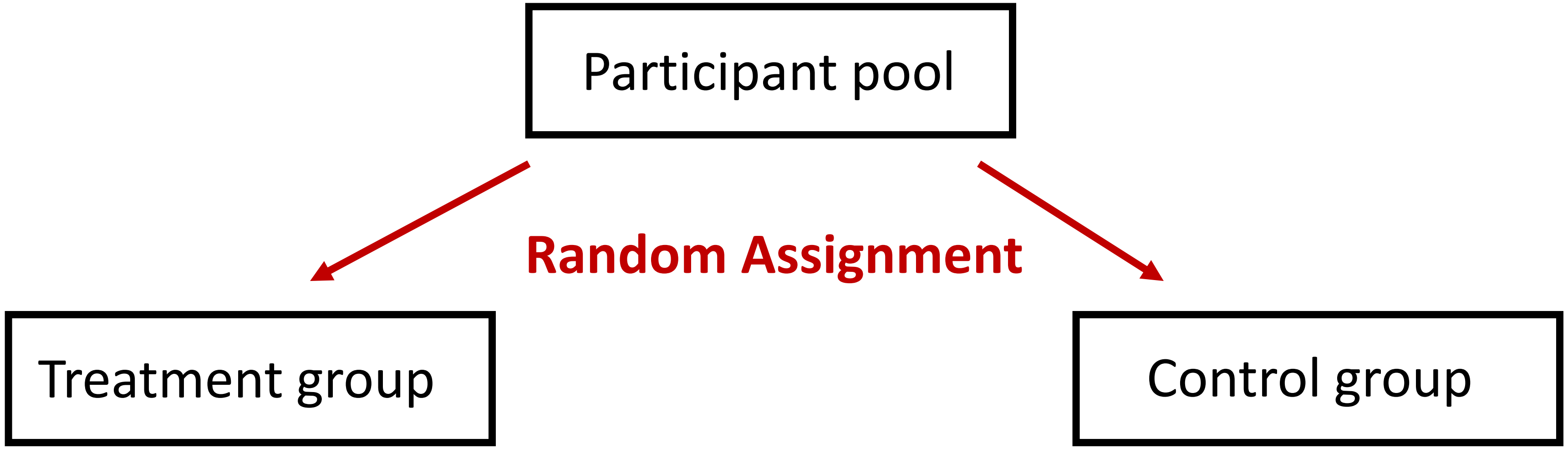
What are the null and alternative hypotheses?

- 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

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

Example: Does calcium reduce blood pressure?

A randomized by Lyle et al (1987) comparative experiment investigated whether calcium lowered blood pressure in African-American men

- 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?

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

3) What is step 3?

3. Create the null distribution!

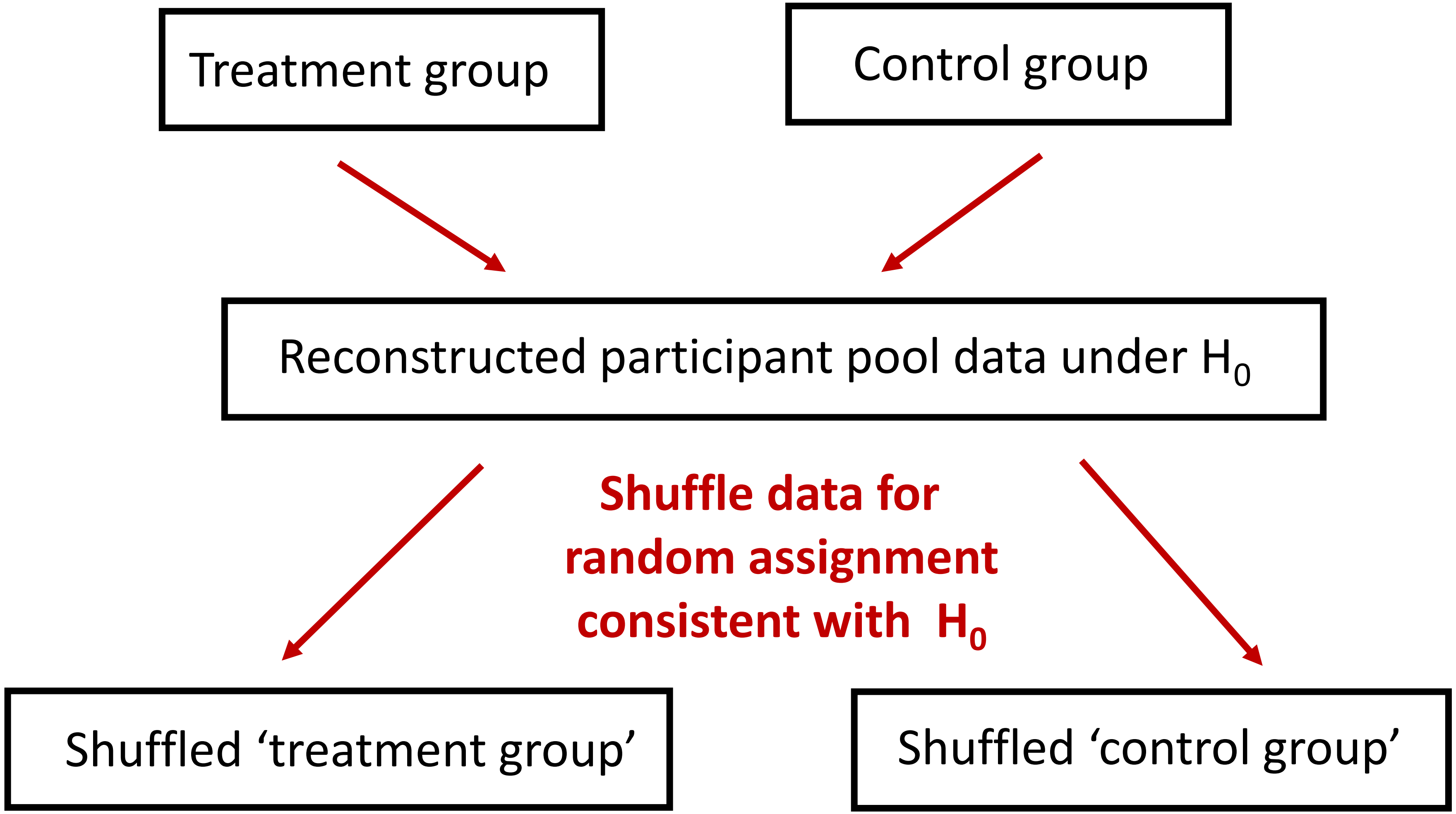
How could we create the null distribution?

Need to generate data consistent with H_0: $\mu_{Treatment} - \mu_{Control} = 0$

- i.e., we need fake $\bar{x}_{Effect}$ that are consistent with H_0

Any ideas how we could do this?

3. Create the null distribution!



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

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 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

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

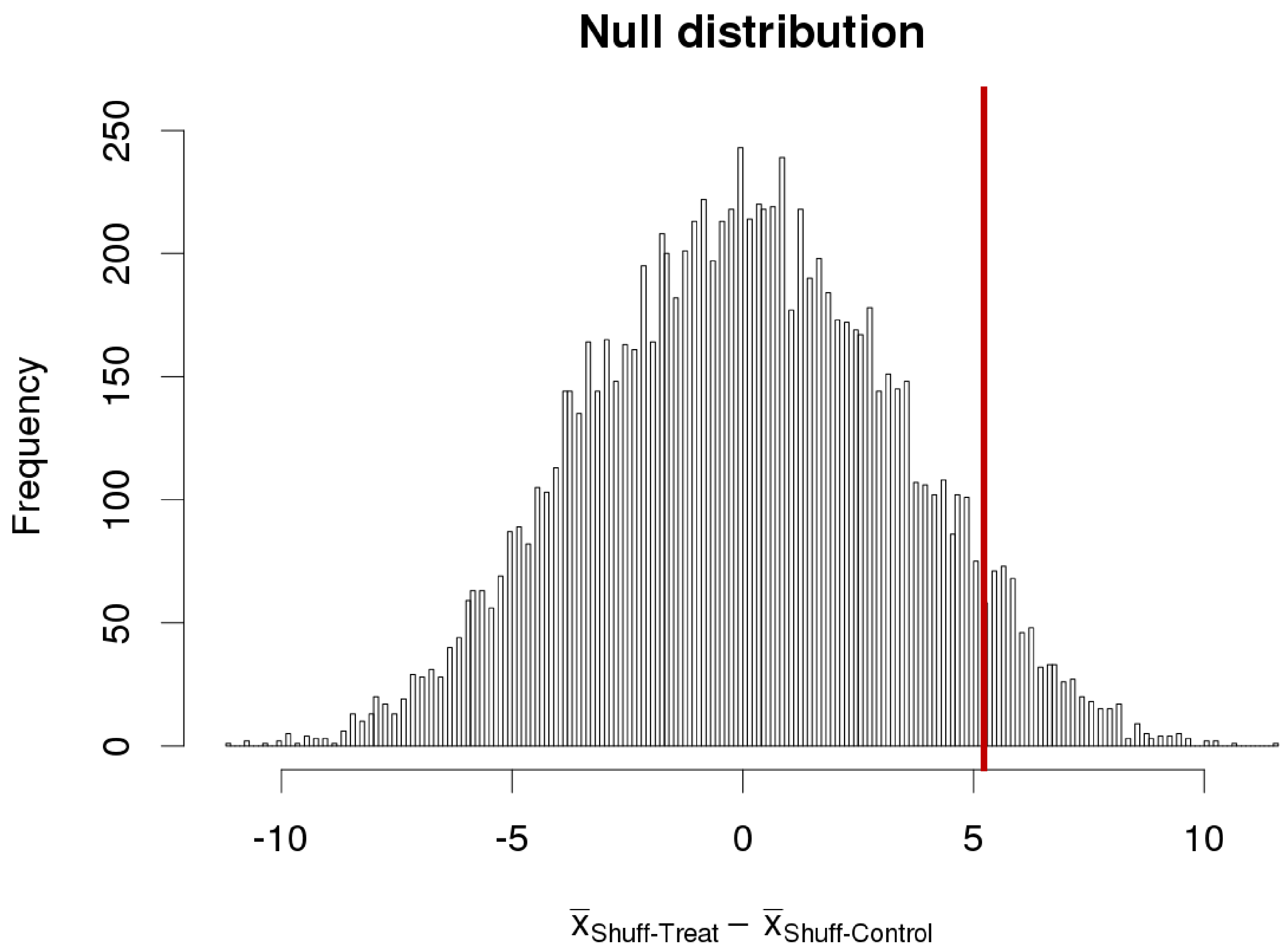
```
null_distribution <-  NULL
for (i in 1:10000) {

    # shuffle data
    shuff_data <- sample(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
    null_distribution[i] <- mean(shuff_treat) - mean(shuff_control)

}
```



hist(null_distribution, nclass = 200)

Next step?

4. Calculate the p-value

```
# 8)  Calculate the p-value
    > p_value <- sum(null_distribution >= obs_stat)/10000
```

p-value = .064

Next step?

5. Are the results statistically significant?

?

What should we do?

More/larger studies!

Worksheet 9

Worksheet 9 is due at 11:59pm on Sunday November 11th

```
> source('/home/shared/intro_stats/cs206_functions.R')
> get_worksheet(9)
```