

## The handout:

- 1) Discuss concepts as they arise (Boxes)
- 2) Serves as a text for the concepts
- 3) Keep it; review this material

**This practical** introduces all of the main concepts for the statistics portion of BMS2

We will return to these concepts in lecture

# Does average female forearm length differ from average male forearm length?

Differentiate between a Null and Alternative hypothesis.

Differentiate between a 'sample' and a 'population'.

Learn to collect data consistently.

Appreciate the concept of 'statistical significance'.

Explain how statistics uses a 'Null Distribution'

Conduct a Randomization Test to test whether two means differ

Questions?

1) Make an observation (females and males have forearms)

2) Formulate a hypothesis

3) Design experiment

4) Conduct the experiment

5) Analyze the data

6) Formulate conclusion

*(Synthesize results with other studies, and determine next step)*

1) Make an observation (females and males have forearms)

2) Formulate a hypothesis

3) Design experiment

4) Conduct the experiment

5) Analyze the data

6) Formulate conclusion

*(Synthesize results with other studies, and determine next step)*

1) Make an observation (females and males have forearms)

2) Formulate a hypothesis

3) Design experiment

4) Conduct the experiment

5) Analyze the data

6) Formulate conclusion

*(Synthesize results with other studies, and determine next step)*

2) Formulate a hypothesis

# Form a Hypothesis

Often most difficult step in science

Essential to design an experiment properly

You must know what you are testing

## Null hypothesis:

Posits that the factor we test does NOT affect our data

We test the null hypothesis: sufficient evidence to reject it?

## Alternative hypothesis:

Opposite of the Null hypothesis (the factor affects our data)

We normally think in terms of an Alternative hypothesis



Null hypothesis: ????

Alternative hypothesis: ????

3 & 4) Design and conduct experiment

Focus:

Populations vs. Sample

Does average female forearm length  
differ from average male forearm length?

How do we answer this question if we're only  
interested in knowing about people in this room?

Does average female forearm length  
differ from average male forearm length?

How do we answer this question if we're only  
interested in knowing about people in this room?

How do we answer this question if we wish to  
know about...

...students at the University of Edinburgh?

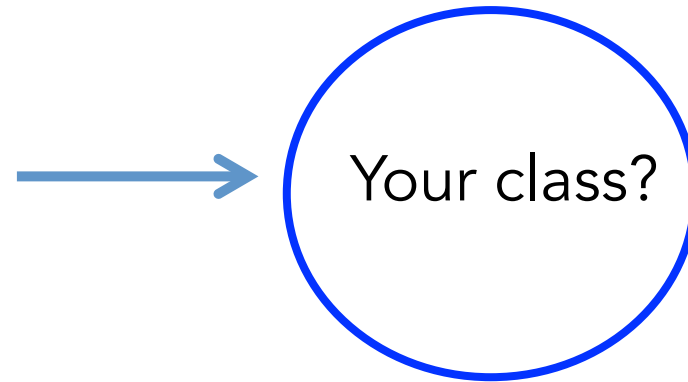
...people in Scotland?

...all humans?

Population



Sample



What we wish to know about

What we are able to study

Population → Sample

To make study feasible

Population ← Sample

How do we do this?  
...this is the point of the  
practical, and statistics

"You" (the class) are our sample

What population do you wish to know about?

3) Collect data  
(and “design experiment”)

a) **Design:** Within your rows, discuss how to measure forearm length

**Discuss:** decide on common approach (measurements must be consistent)

b) Measure forearm length of nearest mm

c) Enter your data



Import the data

Describe the data



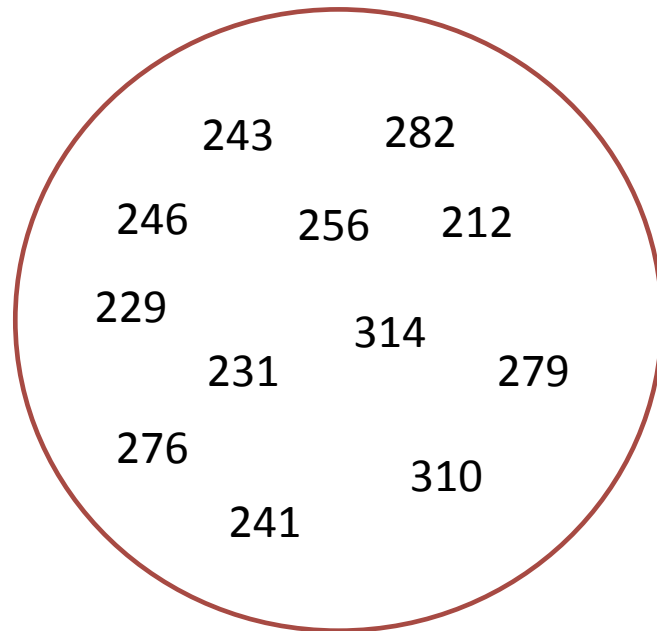
5) Analyze the data

How can we use our sample to test whether mean female and male forearm lengths differ?

Population ← Sample

If mean female forearm length  $\neq$  mean male forearm length,  
then we can think of females and males coming from  
different populations

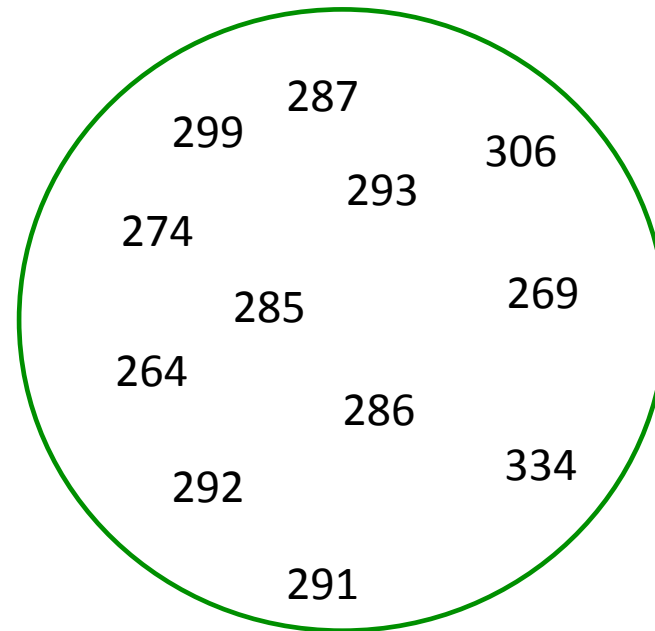
Female (1000's)...only 12 shown



Pop. Mean = 270mm

Sampled females = 229, 279, 243, 246  
Sample mean = 249

Male (1000's)... only 12 shown



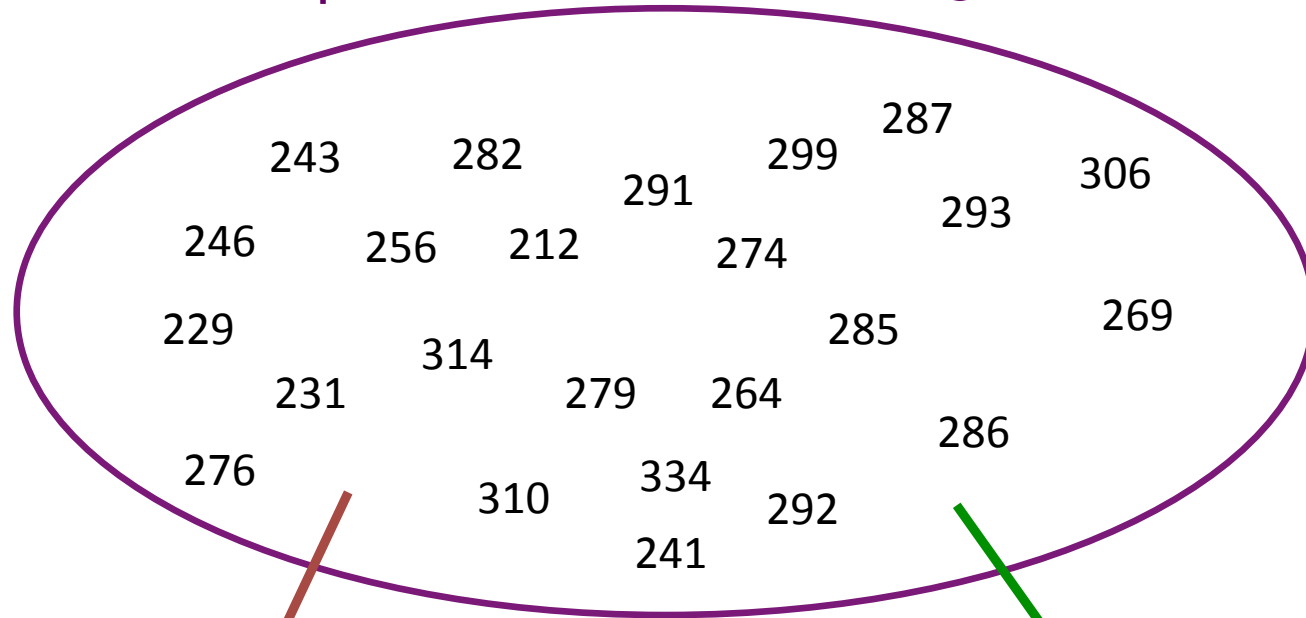
Pop. Mean = 300mm

Sampled males = 285, 291, 334, 293  
Sample mean = 301

**Note:** due to random chance, the sample means do not equal the  
population means      Population ← Sample

But, by random chance, we could likewise obtain what look like different mean forearm lengths for females vs. males when their means are actually the same (i.e., they come from 1 population)

### Population of forearm lengths



Pop. Mean = 285mm

By random chance, we select females, whose average forearm length is less than that of males:

Sampled females = 229, 279, 243, 246  
Sample mean = 249

Sampled males = 285, 291, 334, 293  
Sample mean = 301

How can we use our samples to determine whether female and male forearm length likely differ (come from 2 populations of forearm lengths), or not (come from 1 population)?

Read, discuss and answer questions in Box 3



Is our observed difference likely to arise by chance  
if mean forearm length does not differ for females and males?  
i.e., we imagine that females and males are 1 population

Sampled females = 229, 279, 243, 246    Sampled males = 285, 291, 334, 293  
Sample mean = 249    Sample mean = 301

Observed difference between sample means: 52

### Random Differences:

Female = 229 293 334 243    Male = 285 291 279 246    Difference = 0.5

Female = 279 334 243 246    Male = 291 293 285 229    Difference = 1

Female = 246 285 291 334    Male = 229 279 243 293    Difference = 28

Female = 285 293 243 291    Male = 229 334 246 279    Difference = 6

Etc...

**\*\*How does the observed difference compare  
to the random differences?\*\***

Do 5 times: Randomize your data, and calculate the difference between the (random) female and male forearm lengths

Enter the absolute value of your 5 differences into the spreadsheet at the front

...We should have ~200 randomizations



Import the collective random data

Make a histogram of the random data



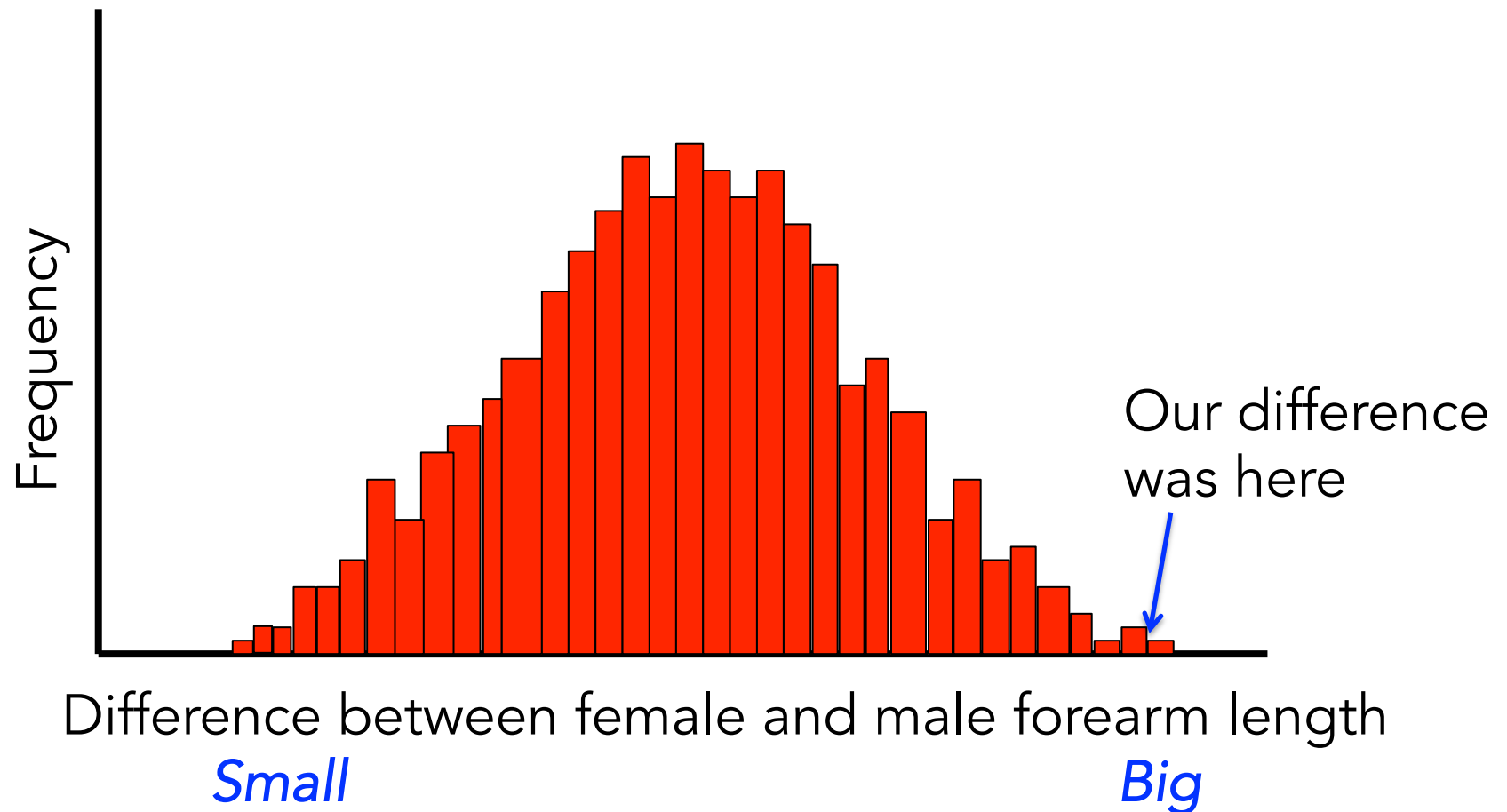
What do the axes represent?

What does this histogram represent?

It is known as a "Null Distribution"

Q: why would it be called a Null distribution?

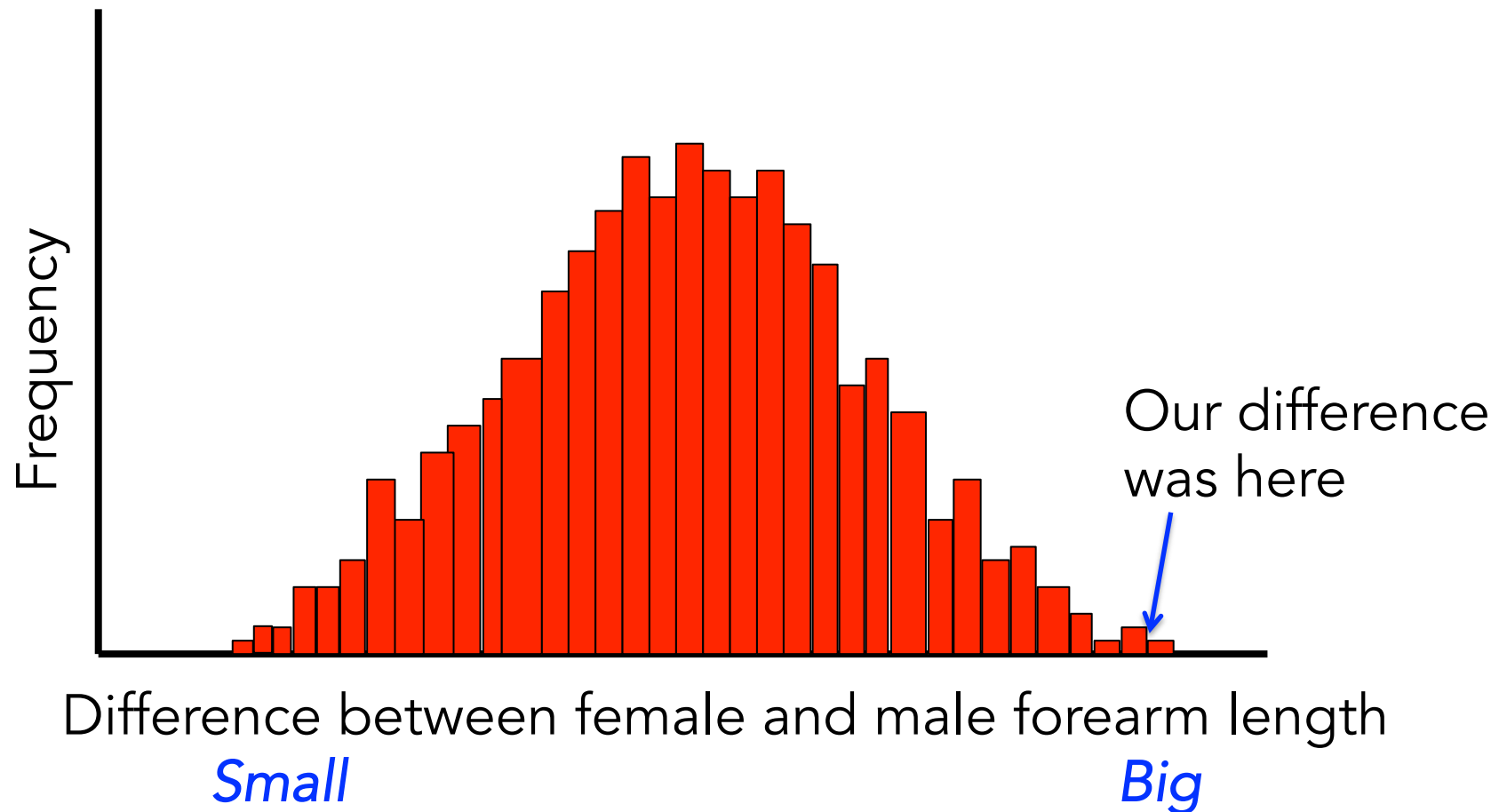
Imagine that we have 200 observations



Would an outcome like this suggest that our observed difference was likely to arise by chance?



Imagine that we have 200 observations



No: this outcome suggests that, if females and males have equal average forearm lengths, we're unlikely to get our observed outcome by chance

Suggests a better explanation: female and male average forearm lengths are not equal; *we consider rejecting the Null hypothesis*

## Describe the Null Distribution

1) Answer Q's 5.9 and 5.10 after Box 4

How can we decide, more objectively, whether to reject the Null hypothesis?

2) Determine the fraction of the Null distribution that is more extreme than the observed difference: Complete up to Box 5



P-value: The proportion of the Null distribution that lies beyond your observed difference.

Gives the probability of observing the data (or data with an even more extreme difference) due to random chance.

By convention, if  $p < 0.05$ , we reject the Null hypothesis  
( $p=0.05$  means 5% of Null distribution is more extreme than our observed outcome)

P-value: The proportion of the Null distribution that lies beyond your observed difference.

Gives the probability of observing the data (or data with an even more extreme difference) due to random chance.

By convention, if  $p < 0.05$ , we reject the Null hypothesis  
( $p=0.05$  means 5% of Null distribution is more extreme than our observed outcome)

CAUTION!!!! The fact that  $p \neq 0$  tells us that there is still a chance that we obtained our outcome by chance, so that the Null hypothesis is actually correct (and rejecting it is wrong).

Given our data, what do we conclude?

# Does average female forearm length differ from average male forearm length?

Differentiate between a Null and Alternative hypothesis.

Differentiate between a 'sample' and a 'population'.

Learn to collect data consistently.

Appreciate the concept of 'statistical significance'.

Explain how statistics uses a 'Null Distribution'

Conduct a Randomization Test to test whether two means differ

Complete the feedback form before leaving