Announcements

- Nothing besides Lab 6 will be due next Monday
- Test 1 is on Friday!
  - Covers Ch 1, Ch 5.1–5.3, Ch 6.1 and 6.3, and A.1–A.2
  - Specific test objectives posted on Campuswire
  - You get a 3x5 inch index card to handwrite whatever you like on one side
  - You will want/need a basic calculator. Phones will not suffice. I have some you can borrow, email me if you'll want to use one.
  - If you haven't done them, you do have the practice problems from each chapter
  - Poll in a moment on if you'd rather I spend my time finishing the grading or putting out some sample questions from each chapter
- Read A.3 (Random Variables) for next Monday
- Polling: `rembold-class.ddns.net`

My Time

I have a limited amount of time before test day. How would you rather I spend it? Which would be most useful to you?

A) Finish grading the assignments and labs

B) Write problems and solutions to a few example problems from each chapter

Warm Up

A poll by the National Sleep Foundation found that college students average about 7 hours of sleep per night. Suppose you were to poll Willamette students about their sleeping habits with the intent of investigating whether they sleep on average less than 7 hours per night. What statement best describes the null hypothesis in this case?

A) Willamette students sleep on average less than 7 hours per night.

B) Willamette students sleep an average of about 7 hours per night.

C) Willamette students sleep on average more than 7 hours per night.

D) Willamette students need to sleep more.

Warm Up

A poll by the National Sleep Foundation found that college students average about 7 hours of sleep per night. Suppose you were to poll Willamette students about their sleeping habits with the intent of investigating whether they sleep on average less than 7 hours per night. What statement best describes the null hypothesis in this case?

A) Willamette students sleep on average less than 7 hours per night.

B) Willamette students sleep an average of about 7 hours per night.

C) Willamette students sleep on average more than 7 hours per night.

D) Willamette students need to sleep more.

When Nature Calls

Saw last class how simulation could give us the odds of getting a certain result assuming the null hypothesis was true

Example

How many trials in each simulation were equal or larger than the 29.2% difference we saw in the experiment?

- 1000 sims: 4/1000 = 0.4%
- 10000 sims: 34/10000 = 0.34%

When Nature Calls

Saw last class how simulation could give us the odds of getting a certain result assuming the null hypothesis was true

Example

How many trials in each simulation were equal or larger than the 29.2% difference we saw in the experiment?

- 1000 sims: 4/1000 = 0.4%
- 10000 sims: 34/10000 = 0.34%

p-value

The ***p-value*** is the probability of observing data at least as favorable to the alternative hypothesis as our data, assuming the null hypothesis is true.

This is (or isn't) Significant!

We need a way to quantify when to reject the null hypothesis in favor of the alternative.

- Look for when the p-value is below some threshold
- Typically called the ***significance level***
- Denoted with an α

This is (or isn't) Significant!

We need a way to quantify when to reject the null hypothesis in favor of the alternative.

- Look for when the p-value is below some threshold
- Typically called the ***significance level***
- Denoted with an α

Statistical Significance

We say the data provides ***statistically significant*** evidence against the null hypothesis if the p-value is less that the significance level, usually with

$$\alpha = 0.05 = 5\%$$

but not always!

Mistakes Happen

- Just because something is statistically unlikely doesn't mean that it won't happen
- Mistakes can be made when accepting or rejecting hypotheses
- Always choose between two options:
  - Failing to reject null hypothesis
  - Rejecting null hypothesis
- Two errors we could make, one for each option

Type 1 Error

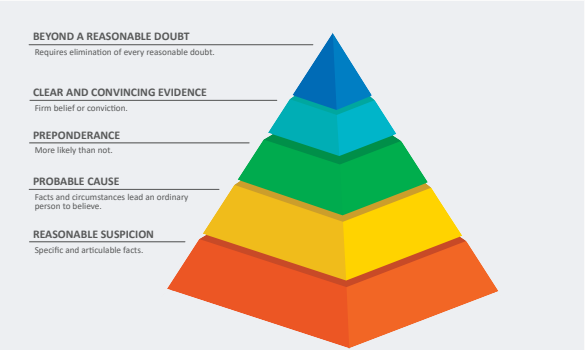
- A ***Type 1 Error*** is rejecting the null hypothesis when it was actually true
- Akin to falsely imprisoning an innocent person in court
- The evidence convinced us when it apparently should not have
- To not make Type 1 Errors, should set a very low significance level
    - Demanding only the most rigorous and convincing evidence to convict

Type 2 Error

- A ***Type 2 Error*** is failing to reject the null hypothesis when it was in fact false.
- Akin to declaring a guilty person innocent in court
- Threshold of evidence was maybe higher than could reasonably be obtained
- To not make Type 2 Errors, should set a higher significance level
  - Lower the burden of proof

A balance point

- Sometimes it is imperative that you don't commit a Type 1 Error
  - Adjust significance level down
- Sometimes it is imperative that you don't commit a Type 2 Error
  - Adjust significance level up
- The significance level corresponds to the chance of you making a Type 1 Error



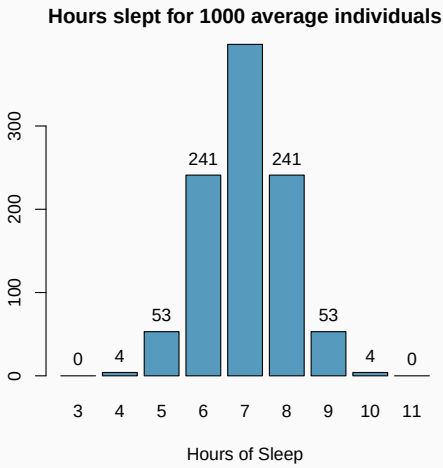
Understanding Check

Suppose a 1000 simulations of average individual sleep habits resulted in the histogram to the right. Your survey of Willamette students results in an average sleep time of 5 hours. Based on this, do you reject the null hypothesis that Willamette students sleep an average of 7 hours per night at the $\alpha = 0.05$ threshold?

A) Yes, reject the null hypothesis

B) No, fail to reject the null hypothesis

C) There is not enough data to tell



And then there were two

- What if instead the research question was: “Do the data provide convincing evidence that the average amount of sleep Willamette students get is different than the average?”
- Then the alternative hypothesis would be different:

$$H_0 : \mu = 7$$
$$H_A : \mu \neq 7$$

And then there were two

- What if instead the research question was: “Do the data provide convincing evidence that the average amount of sleep Willamette students get is different than the average?”
- Then the alternative hypothesis would be different:

$$H_0 : \mu = 7$$
$$H_A : \mu \neq 7$$

- Now need to check odds of being both smaller and greater

Two Tail Tests

- In principle, since normal data is symmetric, if we know one p-value, the other would be the same
- Can just double the p-value of a single tail to get the two tail p-value
- Comparison to the threshold works the same as one tail tests

The Choice

- When in doubt, default to a two-tail test
  - Unless you have convincing reasons to believe only a single direction is valid
- Definitely do not switch between test types after viewing your results!
  - Doing so will without a doubt drive up the chance of you making a Type 1 Error

Test Preparation

I've left the remainder of class available for you to work on book problems, ask questions, or whatever would be most valuable to you to ensure you are ready for the test on Friday.

Or you can leave early, whatever is best for you.