Hypothesis Testing

→ Explain the idea behind a hypothesis test.

→ Identify the main parts of a hypothesis test:
  1. 
  2. 
  3. 
  4. 

Hypothesis Testing: #1 The hypotheses

→ (#1) There are two types of hypotheses:
  1. notation: 
  2. notation: 

→ In a single sample hypothesis test for the mean what are we comparing our parameter of interest to?

Example: CD4 counts (con't) Suppose that the researcher wanted to see if the patients in his clinic were on average "different from" than the AIDS category (200 cells/uL) defined by the US government

→ Find the appropriate null and alternative hypotheses. Write these in English and using statistical notation.

→ What is the difference between a one- and two-tailed hypothesis test? Explain.

→ What are the general forms of the hypotheses for the single sample hypothesis test for the mean?

Ho: 
Ha: 
Ha: 
Ha:
Why is the null hypothesis always equal to some value and the alternative hypothesis changes (depending on the main question of interest?)

Why do we use the notation for the population parameter rather than the sample statistic in the wording of the hypotheses?

Hypothesis Testing: #2 Test Statistic

(#2) A test statistic measures:

$$t_s = \frac{\bar{x} - \text{hyp value}}{s/\sqrt{n}}$$

where: the sample mean and sd are from a random sample; and the sample size is large ($n \geq 30$) or the population distribution is approximately normal

If there were absolutely NO difference between the sample mean and the hypothesized value what would the value of $t_s$ be?

If there were absolutely NO difference between the sample mean and the hypothesized value what would we probably conclude about the null hypothesis?

If the test statistic is close to 0, this shows

If the test statistic is far from 0, this shows

With respect to the formula for $t_s$ why do the two scenarios above this make sense? Explain.
Example: CD4 count (cont)

Descriptive Statistics: CD4

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>N*</th>
<th>Mean</th>
<th>SE Mean</th>
<th>StDev</th>
<th>Minimum</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4</td>
<td>25</td>
<td>0</td>
<td>321.4</td>
<td>14.8</td>
<td>73.8</td>
<td>208.0</td>
<td>261.5</td>
<td>325.0</td>
<td>394.0</td>
<td>449.0</td>
</tr>
</tbody>
</table>

→ Calculate the test statistic.

→ Did we meet our assumptions?
  1) Random?
  2a) Large sample?
  or 2b) Approximately normal?

→ What does the value of $t_s$ mean? Why?

→ From the value of $t_s$ above does it seem like this supports or goes against $H_0$? Explain.

Hypothesis Testing: #3 P-value

→ The p-value is the area…

→ Mark the area on the t distribution that corresponds to a p-value

Definition (p. 419): The p-value (aka observed significance level) is a measure of inconsistency between the hypothesized value for a population characteristic and the observed sample. It is the probability, assuming that $H_0$ is true, of obtaining a test statistic value at least as inconsistent with $H_0$ as what actually resulted.
→ A large p-value implies...? What is large?

→ A small p-value implies...? What is small?

→ The significance level of a hypothesis test is denoted by what?

→ Why is it important to select your significance level before you collect and analyze your data?

→ What are the rules for making a decision?

*Example: CD4 counts (cont’*)

→ Find the p-value that corresponds to the results of the CD4 comparison to 200 cells/μL

Hypothesis Testing: #4 Conclusion

→ Suppose the researchers had set $\alpha = 0.05$, what would our quick conclusion be?

→ Give a conclusion in the context of this example.

→ If we calculated the difference in mean CD4 compared to 200 to be 121.4, why are we concluding that there is no true difference? 121.4 isn’t 0, explain.
Example: CD4 counts (cont’) Could we have said that these patients are significantly *healthier* that AIDS?

→ Could the researchers carried this out as a one-tailed test? In other words would this have been a legitimate research questions?

→ What would have changed in the hypothesis test?

Hypothesis Testing and your HW

→ What are the four parts of a hypothesis test?

→ Do you need to include all 4 parts on your HW even if the problem doesn’t specifically ask for them?

→ How can you identify a hypothesis test problem?

→ What are the important parts of a hypothesis test conclusion?

1.
2.
3.
4.
5.
Hypothesis Tests for Proportions

→ What are the similarities between hypothesis tests for the mean and hypothesis tests for proportions?

→ What are the differences between hypothesis tests for the mean and hypothesis tests for proportions?

→ What are the general forms of the hypotheses for the test of proportions?

\[ \text{Ho:} \]
\[ \text{Ha:} \]
\[ \text{Ha:} \]
\[ \text{Ha:} \]

\[ z_\alpha = \frac{p - \text{hyp value}}{(\text{hyp value})(1 - \text{hyp value})} \sqrt{n} \]

Where: \( p \) is from a random sample; \( n(\text{hyp value}) \geq 10 \) and \( n(1-\text{hyp value}) \geq 10 \); and the sample size is no more than 10% of the population size

Example: Obstructive sleep apnea is a disorder that causes a person to stop breathing momentarily and then awaken briefly. These sleep interruptions, which may cause hundreds of times in a night, can drastically reduce the quality of rest and cause fatigue during waking hours. Researchers studied 159 randomly selected commercial truck drivers and found that 43 of the suffered from obstructive sleep apnea. Suppose that the researchers believe that 25\% of the general public suffer from obstructive sleep apnea. They think that the data will suggest that the truckers have higher rates of sleep apnea than the general public? Test this theory using \( \alpha = 0.01 \).

→ Is this a one- or two-tailed test? Explain why.

→ How do you know that this is a hypothesis test and not a confidence interval problem?

→ Carry out the hypothesis test above.
More on the significance level $\alpha$

*Example:* $\alpha = 0.05$

$\rightarrow$ The chance of rejecting $H_0$, if $H_0$ is true is:

$\rightarrow$ Why (above)?

$\rightarrow$ When we are analyzing one data set in real life at the 0.05 level and our conclusion is to reject $H_0$ there are two possible scenarios:

1. 
2. 

**Type I and Type II Errors**

$\rightarrow$ A type I error is:

$\rightarrow$ $P$(type I error) =

$\rightarrow$ A type II Error is:

$\rightarrow$ $P$(type II error) =

<table>
<thead>
<tr>
<th>Reality</th>
<th>Ho True</th>
<th>Ho False</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision</td>
<td>FTReject Ho</td>
<td>Reject Ho</td>
</tr>
</tbody>
</table>

ANALOGY: Think of a car with a car alarm being broken into.

$\rightarrow$ If the alarm goes off for no reason

$\rightarrow$ If the car gets broken into and the alarm doesn't go off
Example: Measuring pollution in a lake. Say the EPA institutes a rule that companies near bodies of water must test their pollution output. If the company doesn’t find any statistical significance in their results, they may continue their current practices.

→ Ho:
→ Ha:
→ Which error would be worse? Why?

Example: Drug Treatments. Say a doctor would like to study a new highly toxic drug treatment for cancer patients. There are many risks and side effects of the new drug, but would be of benefit if the proportion of patients responding is greater than 50%.

→ Ho:
→ Ha:
→ Which error would be worse? Why?

→ Which error are we protecting ourselves from when we decide on a significance level for our test?

→ The chance of rejecting Ho when it is false is called:

→ Power =

→ We can help control the power of a test by