Lecture 20 (Ch. 8)

Last time:

Example Data says: \( n = 64, \bar{x}_{\text{obs}} = 34.4, \); \( s = 1.1 \).

Does the data provide evidence to support \( \mu_x \geq 34 \)?

So, the a priori belief is \( \mu_x < 34 \).

Let's assume the belief is true.

Formally, Let \( H_0: \mu_x < 34 \). Assume \( H_0 = \text{true} \).

Evidence, from data, against (contrary to) This assumption:

\[
p\text{-value} = P\left( \bar{x} > \bar{x}_{\text{obs}} \mid \mu_x = 34 \right)
\]

\[= P\left( \frac{\bar{x} - \mu_x}{\sigma/\sqrt{n}} > \frac{\bar{x}_{\text{obs}} - \mu_x}{s/\sqrt{n}} \right)
\]

\[= P\left( t > t_{\text{obs}} \right) \approx 0.0025
\]

This \( p\text{-value} \) is small!

It's really easy to misinterpret this small \( p\text{-value} \). But the best way to think of it is that it suggests our assumption was bad, and so, we should reject it.

In short: \( \text{small } p\text{-value} \Rightarrow \text{Reject } H_0 \).

What determines "small"? \( \alpha \) does = 1 - conf. level.

(*) Some skeptical students will want to compute the \( p\text{-value} \) of \( \bar{x} > \bar{x}_{\text{obs}} \) if \( \mu_x = 34 \), and add to that the \( p\text{-value} \) of \( \bar{x} > \bar{x}_{\text{obs}} \) if \( \mu_x = 33.9 \), and add to that the \( p\text{-value} \) when \( \mu_x = 33.8 \) etc. The problem with that logic is that Technically, \( P\left( \bar{x} > \bar{x}_{\text{obs}} \mid \mu_x = \text{anything} \right) \) is not well defined, because "\( \mu_x = \text{anything} \)" is not a random thing. (But see the FYI fig. below, anyway!)
Here is the generalization of the above example:

1) Decide the pop. parameter being tested
   
2) Set up $H_0$ (Null hyp.) and $H_1$ (Alternative hyp.)
   \[
   \begin{align*}
   H_0 : & \quad \mu > \mu_0 \quad | \quad \mu = \mu_0 \quad | \quad \mu < \mu_0 \\
   H_1 : & \quad \mu < \mu_0 \quad | \quad \mu > \mu_0 \quad | \quad \mu \neq \mu_0
   \end{align*}
   \]
   Sufficient to test equality in $H_0$.

4) Choose appropriate statistic.

5) Assume $H_0 = \text{TRUE}$.

6) Compute test statistic for observed data/samples.

7) Find p-value, i.e., prob of getting a random test statistic more extreme (contrary to $H_0$) than the observed one,
   
   \[
   \text{p-value} = \Pr(\bar{X} < \bar{X}_{\text{obs}}) \quad | \quad \Pr(\bar{X} > \bar{X}_{\text{obs}}) \quad | \quad \Pr(\bar{X} = \bar{X}_{\text{obs}}) = \ldots (\text{see below})
   \]

   Mnemonic:
   - Left area
   - Right area
   - Sum of tail areas

8) If p-value < $\alpha$, reject $H_0$ in favor of $H_1$; else cannot reject $H_0$ in favor of $H_1$. 
Note: If p-value < α, then reject $H_0$ in favor of $H_1$.

In English: There is evidence from data in favor of $H_1$ ... (against $H_0$).

Else, cannot reject $H_0$ in favor of $H_1$.
(Not the same thing as "Accept $H_0".)

In English: There is no evidence from data in favor of $H_1$ ... (against $H_0$).
(Not the same thing as "There is evidence for $H_0".)

We cannot accept $H_0$, because we assumed it was true!
When p-value is large, there is simply no evidence from data for anything - not for $H_1$, not for $H_0$!

As I said, there are lots of similarities between the C.I. and the p-value approach, but the differences are very important. For example, there is no $t^*$ (or $z^*$) here: from Tables.

But there is $t_{obs}$ (or $z_{obs}$) from data: $z^* \neq z_{obs, \alpha}$, $t^* \neq t_{obs, \alpha}$.
Now, let's repeat our example, but now following the general procedure.

Data says: \( n = 64, \ x_{\text{obs}} = 34.4, \ S = 0.1. \)

Does the data provide evidence to support \( \mu > 34 \)?

1) The parameter of interest: \( \mu \) (dropping subscript \( x \).)

2,3) \( H_0: \mu \leq 34 \) (or \( \mu = 34 \)) \( H_0 = 34 \)

\( H_1: \mu > 34 \)

Setting up \( H_0/H_1 \) is the hardest part of these problems. See below for guidance.

4) Appropriate test statistic: \( z, t \)

5) Assume \( H_0 = \text{True} \) (i.e. set \( \mu = 34 \))

6) Compute statistic assuming \( H_0 = \text{True} \) (i.e. \( \mu = \mu_0 \)) \( t_{\text{obs}} = \frac{34.4 - 34}{1/\sqrt{64}} = 2.91 \)

Contrary to \( H_0 \) (i.e. in the direction of \( H_1 \))

7) \( p\)-value = \( p_{\text{obs}}(\bar{x} \geq x_{\text{obs}}) = p_{\text{obs}}(t > 2.91) \approx 0.0025 \)

8) Conclusion: \( \alpha = .05 \), \( p\)-value < \( \alpha \).

Therefore, \( \mu > 34 \)

Data provide sufficient evidence to reject \( H_0 \) in favor of \( H_1 \).

"In English": Data provide suff. evidence in favor of \( \mu > 34 \).

\[ \text{At } \alpha = 0.001, \ \text{p-value } > \alpha \] \( \mu < 34 \) \( \mu > 34 \)

Therefore, we cannot reject \( H_0 \) in favor of \( H_1 \).

There is no support for \( \mu > 34 \).

"In English": Note: This conclusion is NOT the same thing as 
"There is support for \( \mu < 34 \." All we can say is that we cannot reject \( \mu < 34 \).
The hardest part of hypothesis testing is setting-up H0/H1 correctly. Here is some guidance:

Four ways I go about for deciding what H0/H1 should be:

1) Don't assume what DATA are supposed to test.
   The question asks "Does data provide evidence for claim X?" Meanwhile, the hypothesis testing procedure begins by assuming whatever you put under H0 is True. So, it makes no logical sense to assume X is true, even before data. So, put the complement/opposite of X under H0.

2) Ask yourself what statement you should be left with if there is NO DATA at all. The answer to that question tells you what H0 should be. Then, the complement of that goes under H1.
   The data provide evidence for H1 (against H0), because of the way the whole procedure is set-up. Then, if the evidence is weak (eg when there is no data at all), then the procedure leaves you with H0, as it should. In our example, if there is no data at all, then we should not reject the belief that \( \mu < 34 \), and so, H0 should be \( \mu < 34 \).

3) Some problems ask you to test some prior belief (i.e., some claim based on something other than data). Then that belief should go under H0.

4) Another way of deciding on H0/H1 will be discussed later, when we learn the meaning of alpha, and Type I and Type II errors.

Further comments:

H0 and H1 are statements about some pop. param, and so, they have no probability.

The p-value is the quantity that represents the evidence from data against H0, in favor of H1. But note that smaller p-value means more evidence against H0 (in favor of H1). This is so because we are giving the benefit of our doubt to H0; so, if H0 is true, and the prob of getting data more extreme than the observed data is large, then there is no evidence for rejecting H0.

If we cannot reject H0 in favor of H1, then we don't know anything. Not rejecting H0 is not the same thing as accepting H0. Making that mistake of interpreting the lack of evidence for H1 as support for H0 is the source of much confusion in science.

In general, we cannot accept a belief about an unknown pop. parameter (eg. \( \mu < 34 \)). All we can to is either reject it, or not, based on evidence from data. And that evidence comes through the p-value; the mathematical way to see this is to note that the p-value is a conditional prob. i.e. it already assumes H0 is true.
In prev. example, we had \( n = 64, \ \bar{x}_{obs} = 34.4, \ \sigma_{obs} = 1.1 \) and asked “Does data provide evidence to support \( \mu > 34 \)?” Thus

\[ H_0: \mu = 34 \quad \text{Note that } H_1 \text{ plays an important role.} \]
\[ H_1: \mu > 34 \]

\[ t = \frac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}} = \frac{34.4 - 34}{1.1 / \sqrt{64}} = 2.91 \]

\[ df = 64 - 1 \]

\[ p-value = \Pr(t > 2.91) = \Pr(z > 2.91) = 0.0025, \]

\[ L = \frac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}} = \frac{34.4 - 34}{1.1 / \sqrt{64}} = 2.91 \]

At \( \alpha = 0.05 \), \( p-value < \alpha \). So, reject \( H_0 (\mu = 34) \) in favor of \( H_1 (\mu > 34) \).

"In English:" There is evidence to support \( \mu > 34 \).

It is tempting to say the above conclusion \( (\text{at } \alpha = 0.05) \), that \( \mu > 34 \), is obvious and trivial. After all the sample gave \( \bar{x}_{obs} = 34.4 \), which is greater than 34 already.

It’s NOT obvious! Suppose the sample/data gave \( \bar{x}_{obs} = 34.1 \), i.e. still larger than 34. Then

\[ t = \frac{34.1 - 34}{1.1 / \sqrt{64}} = 0.73 \Rightarrow p-value = \Pr(t > 0.73) = 0.24 \]

This p-value is larger than any reasonable \( \alpha \). So, we cannot reject \( H_0 \) in favor of \( H_1 \), even though the obs-sample mean is bigger than 34. 34.1 is larger than 34, but just not enough (in units of standard error, \( \frac{5}{\sqrt{n}} \)) to justify rejecting \( H_0 (\mu < 34) \) in favor of \( H_1 (\mu > 34) \).
Study this page carefully.
I’ll go over it on Mon. anyway.

There are many ways to rephrase the statement/question in a problem. Here are some of them:

\[ \alpha = 0.05 \]

Data Says: \( n = 64 \), \( \bar{x} = 34.4 \), \( s = 1.1 \)

\[ \Rightarrow t_{obs} = \frac{34.4 - 34}{1.1/\sqrt{64}} = 2.91 \]

**Does data support \( \mu > 34 \)?**

\( H_0: \mu < 34 \)
\( p-value = \text{prob}(\bar{x} \geq \bar{x}_{obs} | \mu = 34) = \text{prob}(t > t_{obs}) \)
\( H_1: \mu > 34 \)
\( = \text{prob}(t > 2.91) = 0.0025 < \alpha \)

\( \therefore \text{Reject } H_0 (\mu < 34) \text{ in favor of } H_1 (\mu > 34). \)

\( \therefore \text{Data does support } \mu > 34. \)

**Does data support \( \mu < 34 \)?**

\( H_0: \mu \geq 34 \)
\( p-value = \text{prob}(\bar{x} < \bar{x}_{obs} | \mu = 34) = \text{prob}(t < t_{obs}) \)
\( H_1: \mu < 34 \)
\( = \text{prob}(t < 2.91) = 1 - \text{prob}(t > 2.91) = 0.978 > \alpha \)

\( \therefore \text{Cannot reject } H_0 (\mu > 34) \text{ in favor of } H_1 (\mu < 34). \)

\( \therefore \text{Data does not support } \mu < 34. \)

**Does data contradict \( \mu > 34 ?**

\( H_0: \mu \geq 34 \)
\( p-value = \text{prob}(\bar{x} < \bar{x}_{obs} | \mu = 34) = \text{prob}(t < t_{obs}) \)
\( H_1: \mu < 34 \)
\( = \text{prob}(t < 2.91) = 1 - \text{prob}(t > 2.91) = 0.978 > \alpha \)

\( \therefore \text{Cannot reject } H_0 (\mu > 34) \text{ in favor of } H_1 (\mu < 34). \)

\( \therefore \text{Data does not contradict } \mu > 34. \)

**Does data contradict \( \mu < 34 ?**

\( H_0: \mu \geq 34 \)
\( p-value = \text{prob}(\bar{x} \geq \bar{x}_{obs} | \mu = 34) = \text{prob}(t > t_{obs}) \)
\( H_1: \mu > 34 \)
\( = \text{prob}(t > 2.91) = 0.0025 < \alpha \)

\( \therefore \text{Reject } H_0 (\mu < 34) \text{ in favor of } H_1 (\mu > 34). \)

\( \therefore \text{Data does contradict } \mu < 34. \)
If $\mathcal{H}_0: \mu \leq \mu_0$, then

1. **$H_0: \mu = \mu_0$**
2. $H_0: \mu = \mu_0$

why is $p$-value
The right area?

This is a bit wrong/misleading
but helpful anyway!

So, if $\mathcal{H}_0: \mu \leq \mu_0$, then

\[ P-value = \text{prob}(\bar{x} > \bar{x}_{obs}) \]

\[ \text{mnemonic} = \text{right area} \]

If $H_1: \mu > \mu_0$
Toothpaste tubes may be wasteful because there is always some amount of toothpaste that one cannot extract. To find out how much toothpaste is wasted, 5 discarded tubes are selected, cut open, and the amount of remaining toothpaste is recorded. The data are: 0.52, 0.65, 0.46, 0.50, 0.37 (in ounces). Is there evidence that the true average of the wasted toothpaste is less than 0.55 ounces? Apply the hypothesis testing procedure as follows:

a) what is the pop. parameter being tested? Write the symbol for it, AND explain it in words.
b) Restate the question as "Does data provide evidence ---"
c) which of the following pairs of hypotheses is appropriate?

$$H_0 : \mu_x \leq 0.55 \quad H_0 : \mu_x > 0.55 \quad H_0 : \mu_x = 0.55$$
$$H_1 : \mu_x > 0.55 \quad H_1 : \mu_x < 0.55 \quad H_1 : \mu_x \neq 0.55$$

d) In our procedure we must assume $H_0 = \text{True}$. Assuming $H_0 = \text{T}$, what is the "worse" value that $\mu_x$ can take? Hint: values in the direction of $H_1$ are "worse" for $H_0$.

e) Assuming the "worse-case" scenario of part d, compute the p-value. Hint; remember that the p-value measures evidence against (contrary to) $H_0$, or in favor of $H_1$; Use Table VI.

f) Is The p-value you have computed small (less than 0.05) or large (larger than 0.05)?
g) Based on your answer to part f should you reject $H_0$ in favor of $H_1$?
h) What is the conclusion (In English)?

I Suppose you are asked if there is evidence that $\mu > \bar{x}_{obs}$?

a) Set-up the appropriate $H_0/\text{H}_1$,
b) Compute the p-value.

II Suppose you are asked if there is evidence that $\mu > \bar{x}_{obs} - 1.645 \frac{\sigma}{\sqrt{n}}$?

FYI: The right-hand side is the 95% LCB (which we are skipping).

c) Set-up the appropriate $H_0/\text{H}_1$, d) Compute the p-value. Use $pr(t > 1.645) = 0.05$, i.e. $df = n-1 = 80$

Look at the soln later to see the moral of this hw.