Lecture 20 (Ch. 8)

WARNING: Things will get very wordy!

We have built lots of CI's. They are good for 2 things:
1) Convey uncertainty [Reliability] 2) Yes/No decision making

→ Suppose a 2-sided 99% CI for μ_x is $[1.1, 2.3]$
Someone believes $\mu_x = 0$, $\mu_x = 1.5$, $\mu = 3$
Reject or [Not Reject ≠ Accept]? Reject cannot reject Reject.

→ Suppose a 2-sided 99% CI for $\mu_1 - \mu_2$ is $[1.1, 2.3]$
Someone believes $\mu_1 = \mu_2$. Reject

→ Suppose a 2-sided 99% CI for $\mu_1 - \mu_2$ is $[-1.1, 2.3]$
Someone believes $\mu_1 = \mu_2$. Cannot reject (≠ Accept)

Statistical Definition of Belief : A statement not based on data. So, technically, a belief is an assumption

If decision-making (i.e. Reject or Not-reject) is the final goal of your study, then the machinery of computing C.I. can be massaged to form a more direct response. The revised methodology is called often called hypothesis testing (but, don't forget that we can test hypotheses with CIs, too). So, there are 3 hypothesis testing methods: 1) Using CIs, 2) using p-values (common approach), and 3) the Rejection Region approach; better for something called power. We are going to do the p-value approach.

The logic of the p-value methodology is very tricky!
It requires assuming a statement/belief about a pop. parameter, and then checking to see if evidence from data contradict the assumption. Recall that, without knowing pop parameters we cannot compute probs.

The question one asks is of the form "Does data provide sufficient evidence contrary to the assumption/belief?"
If Yes, then we reject the assumption/belief.
If No, then we cannot reject the assumption/belief, i.e., we just don't know!
Notice that "Cannot reject blah" is NOT the same as "Accept blah"!

One can also ask "Does the data provide sufficient evidence in support of some claim (but this time the claim is based on data, ie. the opposite of the assumption/belief)?"

The reason we have to worry about this level of detail is this: There is no such thing as evidence for something that you have already assumed true! Evidence can only come from data, and we can only use evidence to reject an assumption (not to support it). Evidence (from data) cannot support an assumption/belief.

Example: Data says: $n = 64$, $\bar{x}_{obs} = 34.4$, $S = 1.1$.

Does the data provide evidence to support $\mu_x > 34$?

FYI: Technically, we need a lower conf. bound (LCB) for questions like This. But we're skipping Those This quarter. So, let's use our 2-sided CI:

⇒ **CI approach**: 95% CI for μ_x: $34.4 \pm 2 \dfrac{1.1}{\sqrt{64}} = [34.1, 34.7]$ — μ_x in $[34.1, 34.7]$

$2 \leftarrow$ qt(.05/2, 64-1)

We are 95% confident That $\mu_x \in [34.1, 34.7]$

There is evidence That $\mu_x > 34$.

I.e. If someone claims $\mu_x < 34$, Then we can Reject The claim.

⇒ **A different way** of arriving at That conclusion.

This time: Assume $\mu_x < 34$, and find evidence (from data) to The contrary.

Another way of saying This: Let $H_0: \mu_x < 34$, Then assume $H_0 =$ True

H_0 called "Null hypothesis"

Q: What's contrary?

A: Really large $\bar{x}$'s are contrary to $\mu_x < 34$. So, here is a measure of evidence (from data) contrary to $\mu_x < 34$: $pr(\bar{x} > \bar{x}_{obs} \mid \text{if } \mu_x < 34)$.

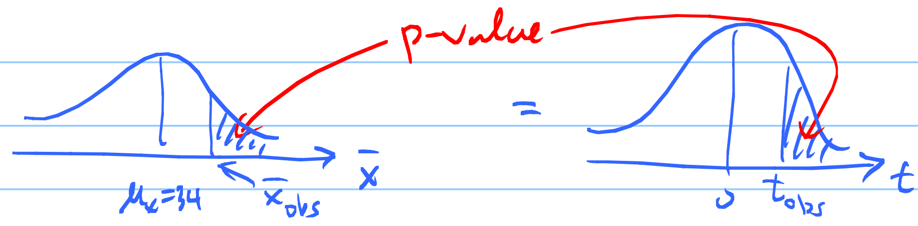
⇒ Let's start by computing That prob, if $\mu_x = 34$ (The worse-case scenario for H_0)

$$\underbrace{prob(\bar{x} > \bar{x}_{obs} \mid \mu_x = 34)}_{\text{one type of "p-value"}} = prob\Big(\underbrace{\frac{\bar{x} - \mu_x}{S/\sqrt{n}}}_{t} > \underbrace{\frac{\bar{x}_{obs} - \mu_x}{S/\sqrt{n}}}_{t_{obs}} \;\Big|\; \mu_x = 34\Big)$$

$$t_{obs} = \frac{\bar{x}_{obs} - \mu_x}{S/\sqrt{n}} = \frac{34.4 - 34}{1.1/\sqrt{64}} = 2.91$$

1 - pt(2.91, 63)

$= prob(t > 2.91) \simeq .0025$ $\doteq$ very small.



Let's Think! If we assume μ_x is as big as it can get according to the claim, i.e. $\mu_x = 34$, then the prob. of getting $\bar{x}$ larger than $\bar{x}_{obs}$ is very small. That is a lot of evidence contrary to $\mu_x = 34$, because that prob already assumes $\mu_x = 34$. This is a tricky point, so read it again!

Lesson 1: smaller p-value $\Rightarrow$ more evidence against $\mu_x < 34$.

very strange! Just give it time!

So, for our example, the conclusion is there is evidence from data to reject $\mu_x < 34$ (ie. the same conclusion as the CI method), all because the above prob (an example of a p-value) is small!

Q: Now, what if we relax $\mu_x = 34$ into $\mu_x < 34$?

A: Repeat the above calculation,

if $\mu_x < 34$, $\Rightarrow$ t_{obs} = larger $\Rightarrow$ p-value = smaller.

So, if prob($\bar{x} > \bar{x}_{obs} \mid \mu_x = 34$) = small,

we can reject $\mu_x < 34$, not just $\mu_x = 34$.

Blue Note

In English: if data reject $\mu_x = 34$, then $\mu_x < 34$ is also rejected

Lesson 2: It is sufficient to test $\mu_x = 34$

Don't forget this 34 came from the belief (not data)

FYI

Some skeptical students will want to compute the prob of $\bar{x} > \bar{x}_{obs}$ if $\mu_x = 34$, and add to that the prob. of $\bar{x} > \bar{x}_{obs}$ if $\mu_x = 33.9$ and add to that the prob when $\mu_x = 33.8$ etc. The problem with that logic is that technically $pr(\bar{x} > \bar{x}_{obs} \mid \mu_x = \text{anything})$ is not well defined, because "$\mu_x = \text{anything}$" is not a random thing. (But see the FYI fig. in the next lecture, anyway).

Q Who decides what's a "sufficiently small" value for p-value?

A You do! This "threshold probability" is labeled α

It's called significance level. $(= 1 -$ conf. level$)$

.05 sign. level = 95% Conf. level.

Some common values are .05, .01, .001

In summary: If p-value $< \alpha$, then reject H_0
else cannot reject H_0.

Next time, we will formalize things a bit more.

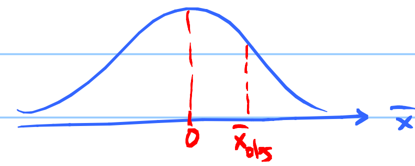
As I said, there are lots of similarities between the C.I and the p-value approach, but the differences are very important.

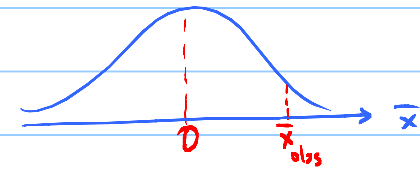
For example there is no t^* (or z^*) here. from Tables.

But there is t_{obs} (or z_{obs}) from data $\boxed{z^* \neq z_{obs},\ t^* \neq t_{obs}}$

hw-lect20-1

Consider a problem wherein someone claims that $\mu_x < 0$, and you want to see if data provide evidence to its contrary. Consider the following 2 hypothetical $\bar{x}_{obs}$ values.

I) 

II) 

a) Forget p-values and all that! Which of the 2 situations (I or II) provides more evidence (from data) against $\mu_x < 0$?

b) Now, we would like to find the p-values for these 2 cases, but I haven't given you any numbers for $\bar{x}_{obs}$. Instead, shade the p-values as areas under those 2 figs.

c) Which case (I or II) has the <u>smaller</u> p-value?

hw_lect20_2

Consider the following sample observations: 2781, 2900, 3013, 2856, and 2888.
Suppose we want to test whether there is evidence contrary to the belief that mu < 3000.
a) Compute the observed 95% 2-sided confidence interval (CI) for mu.
b) Based on the above CI, is there evidence that mu is greater than 3000?
c) Write the appropriate null hypotheses.
d) Compute the p-value, recalling that it measures evidence from data contrary to the null hypothesis.
e) At alpha=0.05, state the conclusion "In English" (i.e., is there evidence that mu is greater than 3000?)