Lecture 19 (Ch. 7)

Last time we learned about the CLT which gives us the distr. of sample means:

$$\bar{x} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right)$$

μ_x = pop. mean (of x)
σ_x = " std. dev. (of x)

From that, we can get probs like in the last example:

E.g. Suppose a sample of size 25 yields $\bar{x}_{obs} = 3$, $s_{obs} = 1$. If the population is $N(\mu = 2, \sigma = 1)$, what's the prob. of getting an even larger sample mean?

$$\text{prob}(\bar{x} > \bar{x}_{obs}) = $$

$$\text{prob}(\bar{x} > 3) = \text{prob}\Big(\overset{\frac{\bar{x}-\mu_x}{\sigma_x/\sqrt{n}}}{\overset{\parallel}{z}} > \frac{3-2}{1/\sqrt{25}}\Big) = \text{prob}(z > 5) \approx 0!$$

In English, if we know pop. mean and std dev., then we can compute the prob. of a sample mean being in some range. That's useful because that prob can tell us whether the single sample mean we have observed is *typical* or not.

So, how can we use all this to infer μ_x (or σ_x)?

1) Build Confidence Intervals for μ_x (Ch. 7)
2) Test hypotheses about μ_x (Ch. 8)

pop. param.

Notation & terminology:

$\bar{x}$ (sample mean) is a point estimate of μ_x (pop. mean)
s (" std. dev.) " " " " σ_x (" std. dev.)
p (" prop.) " " " π_x (" prop.)
n (" size) is NOT related to pop. size. ← for us $= \infty$

Later, I may drop the x subscript.

Let's do the C.I. for pop. mean μ_x, first.

In words, we know that $\sigma_{\bar{x}}$ tells us how far a typical sample mean is from μ_x.

In Math, CLT: $\bar{x} \to z = \dfrac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \in N(0,1)$.

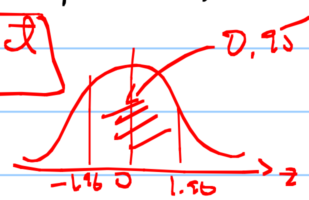
$\sigma/\sqrt{n}$

So, we can compute prob($a < z < b$) = blah

The trick is to look at special values of a, b, blah.

self-evident fact

$$P(-1.96 < z < 1.96) = 0.95$$



$$P\left(-1.96 < \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} < 1.96\right) = 0.95$$

$\mu_{\bar{x}} \leftarrow \mu_x$, $\sigma_{\bar{x}} \leftarrow \sigma_x/\sqrt{n}$

$$P\left(-1.96 \frac{\sigma_x}{\sqrt{n}} < \bar{x} - \mu_x < +1.96\, \sigma_x/\sqrt{n}\right) = 0.95$$

"Confidence level"

$$\bar{x} - 1.96 \frac{\sigma_x}{\sqrt{n}} < \mu_x < \bar{x} + 1.96 \frac{\sigma_x}{\sqrt{n}}$$

No Prob.

pop. mean. Slowly, I may drop the x!

$$\therefore \text{95\% C.I. for } \mu_x: \quad \bar{x} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}$$

approximate with sample std. dev.

This is a random C.I, because $\bar{x}$ is random (how else would it have a sampling dist?!)

The (observed) 95% C.I. for μ_x is $\boxed{\bar{x}_{obs} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}}$

Later, I may drop saying "observed"!

E.g. Suppose a sample of size 25 yields $\bar{x}_{obs} = 3$, $s_{obs} = 1$.
What can we say about the pop. mean?

prev. eg.

Suppose pop is Normal $(\mu_x = 2, \sigma_x = 1)$. What's the prob of getting an even larger sample mean? (Ch. 8)

$\text{prob}(\bar{x} > x_{obs}) =$

$$\text{prob}(\bar{x} > 3) = \text{prob}\left(z > \frac{3-2}{1/\sqrt{25}}\right) = \text{prob}(z > 5) \approx 0!$$

$\bar{x} > 3$ is unlikely, if $\mu_x = 2$.

(observed) 95% C.I. for μ_x : $\bar{x}_{obs} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}$ — estimate with s_x

$$3 \pm 1.96 \frac{1}{\sqrt{25}} = 3 \pm .392 = (2.6, 3.4)$$

→ We can be 95% confident that the true mean is in here.

Note that we have actually made it to our goal of being able to say something about a pop. mean, from a sample. Review how we needed everything we've done since Ch. 1. Go and celebrate!

BUT, there are other ways of interpreting CIs, and that's where things get difficult again, and unfortunately, the interpretations are very important.

So, if I decide on the 0.95 confidence level, then I can say something about the interval wherein μ_x resides.

Note: NOT "prob of μ_x" !

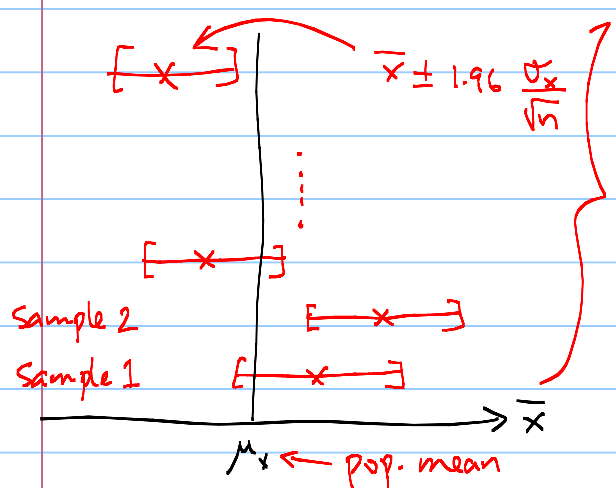
Only confidence that μ_x is in some Conf. Int.

There is a way of squeezing "probability" into the conclusions, but it has to pertain to the random C.I.

We are 95% confident that the pop. mean is in the interval $\boxed{\bar{x}_{obs} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}}$.

Equivalent interpretations of C.I.

There is a 95% prob that a random sample will yield a C.I. $\left(\bar{x} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}\right)$ that covers μ_x.



→ 95% of these intervals cover μ_x.

→ I.e. the prob. that a random C.I. $\left(\bar{x} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}\right)$ will include μ_x is 0.95.

→ If you want to say something directly about μ_x, use "confidence" Not prob.

C I's are all about coverage;
a 95% C.I. for μ_x is designed to cover μ_x in 95% of samples.

For The above example: (Observed) 95% CI: (2.6, 3.4)

→ We can be 95% Confident That The True mean is in here.

→ There is a 95% prob. That a random 95% CI will cover μ_x.

Note that if we Think of $n=25$ as "large," Then we don't even have to assume That The pop. (whose mean we have inferred) is Normal

Q1: Suppose we have computed (or observed) a 95% C.I. for μ_x: $(-2, 1)$. Which of The following statements is Correct.

A) There is a 95% prob. That μ_x is between -2 and $+1$.

B) There is a 95% prob That a random C.I. is $(-2, +1)$.

C) There is a 95% prob That The obs. $\bar{x}$ is in $(-2, +1)$.

D) There is a 95% prob That a random $\bar{x}$ is in $(-2, +1)$.

$\boxed{\text{E)}}$ None of The above.

A) μ_x has no prob.

B) The prob of a random continuous Thing being any specific Thing is zero!

C) We know $\bar{x}_{obs}$! It has no prob.

D) The correct statement would be "... is in $\mu_x \pm 1.96 \frac{\sigma_x}{\sqrt{n}}$"

See how below.

hw_lect 19_1 Rephrasing 7.16

a) The problem tells us That the sample std. dev., s, has a Normal distr. with parameters σ_x and $\sigma_x/\sqrt{2n}$, where σ_x is The pop. std. dev. Now, follow The procedure we have developed, starting from a "self-evident fact" to develop a C.I. formula for σ_x.

Note: in This way of writing The problem, we avoid The notational complexity of μ_s and σ_s in 7.16.

b) Then compute The observed 95% C.I for The true/population standard deviation for The data in problem 7.15.

c) Interpret in 2 ways.

hw_lect 19-2 Explain why The correct re-statement in part d) in The clicker question above is correct, ie. That The prob of a random sample mean to be within $\mu_x \pm 1.96 \frac{\sigma_x}{\sqrt{n}}$ is .95.