3Y03 - PROBABILITY AND STATISTICS FOR ENGINEERING

WS19 Lecture 22

Testable Material for Midterm 2

All relevant material (see Lecture Schedule/Suggested Problems) from:

- 4. Continuous Random Variables ...
- 5. Joint Probability Distributions
- 6. Descriptive Statistics
- 7. Point Estimation of Parameters ...

---

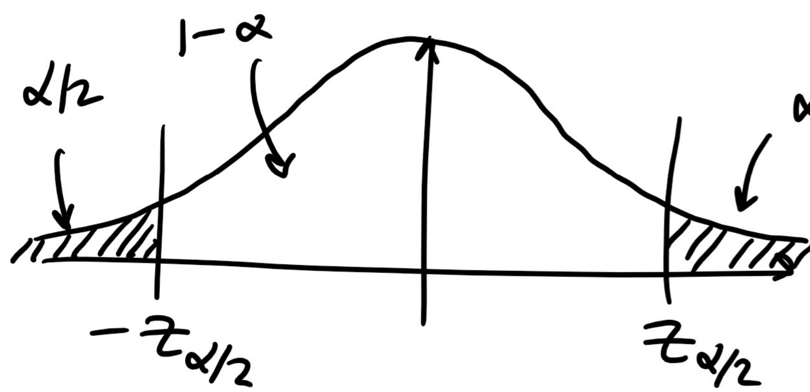
8. $100(1-\alpha)\%$ Confidence Intervals - interested in θ

- use estimate $\hat{\theta}$ (i.e. data) to find (L, R) with $P(L < \theta < R) = 1-\alpha$.

C.I. for mean μ of Normal pop., Variance known (σ^2)

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$



First $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$

$$\Rightarrow P\left(-z_{\alpha/2} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\Big(\overbrace{\bar{X} - z_{\alpha/2}\left(\tfrac{\sigma}{\sqrt{n}}\right)}^{L} < \mu < \underbrace{\bar{X} + z_{\alpha/2}\left(\tfrac{\sigma}{\sqrt{n}}\right)}_{R}\Big) = 1-\alpha$$

$$\frac{\alpha}{2} = P(Z > z_{\alpha/2})$$

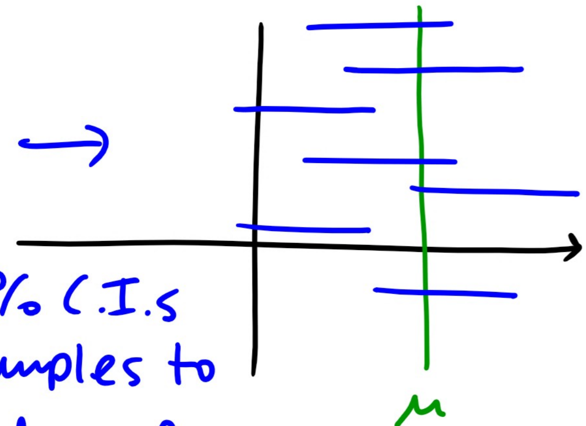
$$1 - \cancel{z_{\alpha/2}}\ \tfrac{\alpha}{2} = P(Z \le z_{\alpha/2})$$

A $100(1-\alpha)\%$ C.I. for mean μ of normal pop. variance σ^2 known is

$$\left(\bar{x} - z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right), \; \bar{x} + z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)\right)$$

or $$\bar{x} \pm \boxed{z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)} \leftarrow \text{Margin of Error (ME)}$$

Expect $100(1-\alpha)\%$ of the $100(1-\alpha)\%$ C.I.s for different samples to contain true value of μ



<u>Example</u> Find a 93% C.I. for the mean of a Normal pop. where $\sigma = 5.7$ & $n = 10$.

<u>Solution</u> Need $\bar{x}$ & $ME = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$

$$= z_{\alpha/2}\left(\frac{5.7}{\sqrt{10}}\right)$$

where $100(1-\alpha) = 93$

$\Rightarrow \quad 1-\alpha = 0.93$

$\Rightarrow \quad \alpha = 0.07$

$\Rightarrow \quad \alpha/2 = 0.035$

Want $z_{\alpha/2}$ i.e. need to find z with $P(Z < z) = 1 - \frac{\alpha}{2}$

$= 1 - 0.035 = 0.965.$

Look up 0.965 in standard normal table : $z_{\frac{\alpha}{2}} = 1.81$

So $ME = 1.81\left(\frac{5.7}{\sqrt{10}}\right) = 3.26$

So C.I. : $\bar{x} \pm 3.26$.

Or can turn this around: demand certain ME

or interval length = $2 \times ME$.

How big a sample do we need in previous setup to have a 93% C.I. with $\underbrace{ME = 0.3}_{\text{This really means } ME \leq 0.3}$?

$$ME = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

$$\underbrace{z_{\alpha/2}}_{1.81}\underbrace{\left(\frac{\sigma}{\sqrt{n}}\right)}_{\frac{5.7}{\sqrt{n}}} \leq 0.3$$

$$\Rightarrow \frac{(0.3)(1.81)}{5.7} \geq \frac{1}{\sqrt{n}}$$

$$\Rightarrow \frac{5.7}{(0.3)(1.81)} \leq \sqrt{n}$$

$$\Rightarrow n \geq \left(\frac{5.7}{(0.3)(1.81)}\right)^2 = 110.191$$

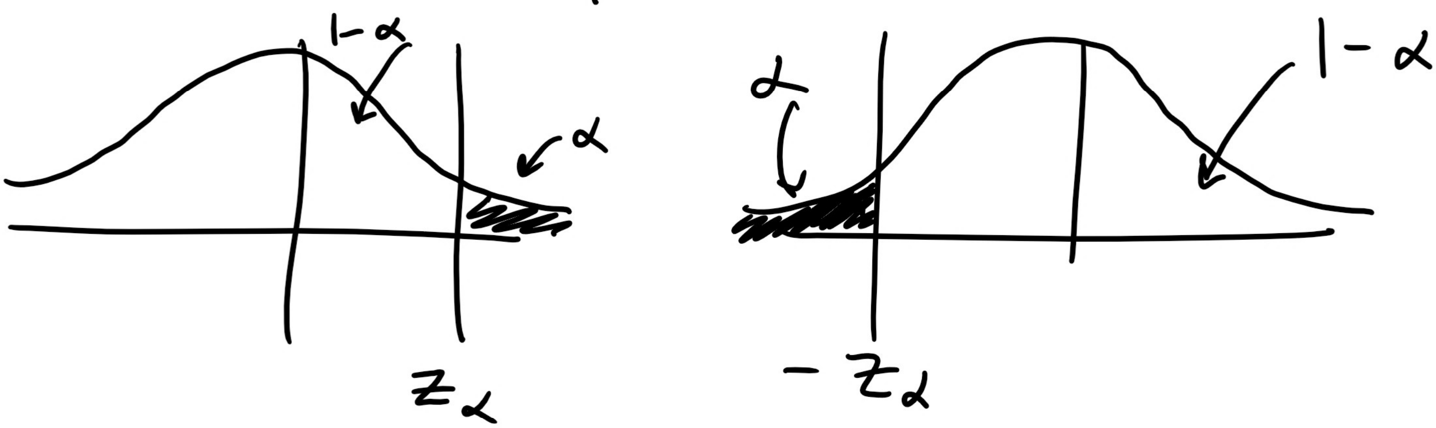
Choose smallest whole # n satisfying this

i.e. 111.

Sometimes we call C.I. a 2-sided C.I. to emphasise difference from:

<u>One-sided confidence bounds</u>

Same idea but asymmetric
$\rightarrow$ all "α prob." in one tail



So a <u>$100(1-\alpha)\%$ upper confidence bound</u> for mean μ, normal pop., variance σ^2 known

is $\bar{x} + z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right)$

& <u>lower</u> : $\bar{x} - z_\alpha\left(\frac{\sigma}{\sqrt{n}}\right)$

<u>C.I. for large samples, variance known</u>

$\hookrightarrow n \geq 30$

If underlying distr. not necessarily normal, Central Limit Theorem says if $n \geq 30$ ~~&~~ then $\bar{X} \approx N(\mu, \frac{\sigma^2}{n})$

So then $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ & all the same formulas as above work.

What if σ^2 <u>not</u> known?

<u>C.I. for mean, variance not known, large sample</u>

If n large ($n \geq 40$), substituting S for σ still gives normal:
$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim N(0,1).$$

Then proceed using same formulas as above substituting s for σ everywhere:

e.g. 2-sided C.I.: $\bar{x} \pm z_{\alpha/2}\left(\dfrac{s}{\sqrt{n}}\right)$.

1-sided bounds: $\bar{x} + z_\alpha\left(\dfrac{s}{\sqrt{n}}\right)$, $\bar{x} - z_\alpha\left(\dfrac{s}{\sqrt{n}}\right)$

<u>C.I. for mean μ, sample NOT large, variance NOT known</u>

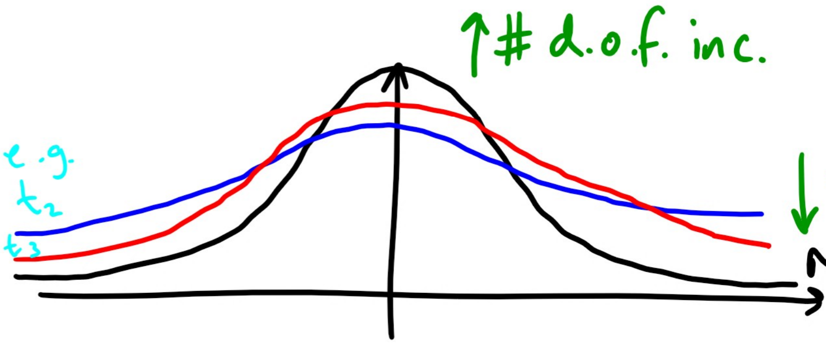
We have to assume <u>something</u>: that underlying distr. is normal $N(\mu, \sigma^2)$ (so $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ but <u>don't know</u> σ !!!)

Again: $\dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ replaced by $\dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ $\leftarrow$ but now cannot assume normal if n small

$\dfrac{\bar{X}-\mu}{S/\sqrt{n}} =: T$ has a (student) t-distribution with (n-1) degrees of freedom

AKA t_{n-1}-distribution

- Depends on n (only)
- Different t-distr. for each # of "degrees of freedom"
- Table values (not integration)!



As # d.o.f. inc. converge to N(0, 1).

To find a $100(1-\alpha)\%$ C.I. for mean μ (normal pop., variance unknown, sample size small)

Now find $t_{\frac{\alpha}{2}, n-1}$ with

T has a t_{n-1}-distr.

$$P\left(-t_{\frac{\alpha}{2}, n-1} < T < t_{\frac{\alpha}{2}, n-1}\right) = 1-\alpha$$

So $\boxed{\text{C.I. is } \bar{x} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}}$.

Where $\bar{x}$, s found using data &

$t_{\frac{\alpha}{2}, n-1}$ found using t-table

Look up value in row $n-1$ T.B.C.

column $\alpha/2$.