3Y03 - 3J04 PROBABILITY & STATISTICS FOR (CIVIL) ENGINEERING

C01 - Lecture 23

Last time — CONFIDENCE INTERVALS

$100(1-\alpha)\%$ C.I. for θ : (L, R) where $P(L < \theta < R) = 1-\alpha$.

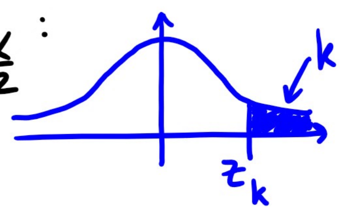
Normal Population, known variance σ^2 :

$100(1-\alpha)\%$ C.I. for μ given by $\boxed{\bar{x} \pm \underbrace{z_{\frac{\alpha}{2}}\left(\frac{\sigma}{\sqrt{n}}\right)}_{\text{ME}}}$

ME : Margin of Error

n = sample size

$z_{\frac{\alpha}{2}}$:



$P(Z < z_k) = 1-k$

Exercise How big do we need our sample to be to have a $100(1-\alpha)\%$ C.I. with ME = 0.3 when σ is known?

Solution

$$0.3 \geq \text{ME} = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

← given σ

↑ given

↑ Find from table

Then solve for n.

If we set $100(1-\alpha) = 95\%$ and $\sigma = 1.2$, then

$\alpha = 0.05$

$\Rightarrow \frac{\alpha}{2} = 0.025$

$z_{\frac{\alpha}{2}} = 1.96$ so solve :

$$0.3 = \frac{1.96 \times 1.2}{\sqrt{n}}$$

i.e. $$n = \left(\frac{1.96 \times 1.2}{0.3}\right)^2 = 61.46$$

So we would need $n \geqslant 62$ for this 0.3 margin of error. (We would likely get under 0.3 with 62, but we would be over 0.3 with $n = 61$.)

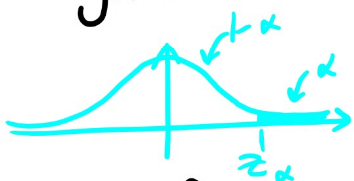
(<u>Note</u> : more data points $\Rightarrow$ smaller margin of error.)

One-sided confidence bounds for μ

Same idea but asymmetric : now all α is in one tail (left or right)

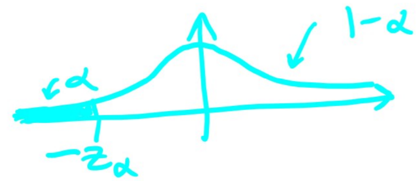
So for $100(1-\alpha)\%$ <u>upper confidence bound</u> for μ we find z_α to give $\bar{x} + z_\alpha \left(\frac{\sigma}{\sqrt{n}}\right)$

$P(Z < z_\alpha) = 1 - \alpha$



& $100(1-\alpha)\%$ <u>lower confidence bound</u> for μ is

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



C.I. for large samples (variance known)

Even if underlying distribution is not (known to be) Normal, Central Limit Theorem says that if n is large ($\geqslant 30$), then $\dfrac{\bar{X} - \mu}{(\sigma/\sqrt{n})} \sim N(0,1)$ so to get a C.I. for μ we can use same formulas as above.

(or for upper & lower confidence bounds).

But what if σ^2 <u>not</u> known?

<u>CI for μ, variance unknown</u>

<u>FACT</u> If n large, it turns out that estimating σ^2 by S^2 (so substituting s for σ in above) has little effect.

More precisely

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

<u>when n large.</u>

Consequently, <u>if</u> n is large we can use almost same formulas as before (but with S in place of σ) ~~i.e.~~ e.g.

$100(1-\alpha)\%$ C.I. for μ is

$$\left(\bar{x} - z_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right),\ \bar{x} + z_{\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)\right), \text{ where}$$

$\bar{x}$ and s are computed using sample $\{x_1, \ldots, x_n\}$.

UP TO HERE FOR TEST 2.

C.I. for mean μ when variance unknown AND sample size small.

Have to assume something: we assume that the underlying distribution (of X_is) is Nomal.

(Don't know μ, don't know σ^2.)

Again look at $T = \dfrac{\bar{X} - \mu}{(S/\sqrt{n})}$. Now DO NOT Know this is $N(0,1)$ (& in general it won't be)

But T, when n is small, has a

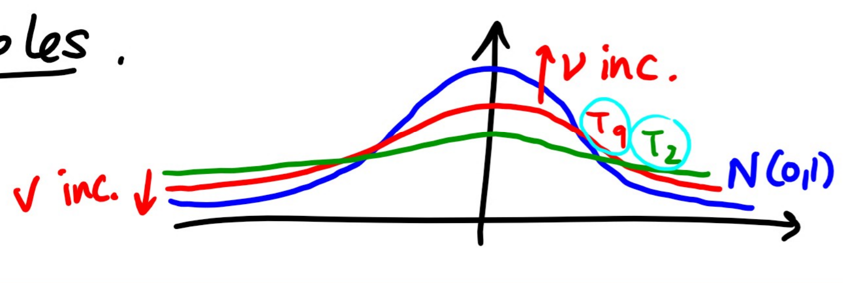
(Student) t distribution "with $(n-1)$ degrees of freedom"

In general ν degrees of freedom:

For interest pdf $f(x) = \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}} \left(\dfrac{x^2}{\nu} + 1\right)^{-\frac{(\nu+1)}{2}}$

Just as with normal distribution, we don't work with pdf, we use tables.

As $\nu \to \infty$, limit is $N(0,1)$

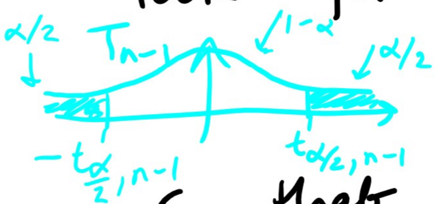


We said that $T = \dfrac{\bar{X} - \mu}{(S/\sqrt{n})}$ has t distribution "with $n-1$ degrees of freedom" i.e. with $\nu = n-1$

So underlying distribution of T depends on n but not on μ or σ (or on S).

To find $(100)(1-\alpha)\%$ C.I. for μ, again look for symmetric interval (t distr. is symmetric)

— look for $t_{\frac{\alpha}{2},n-1}$ with $P(-t_{\frac{\alpha}{2},n-1} < T < t_{\frac{\alpha}{2},n-1}) = 1-\alpha$

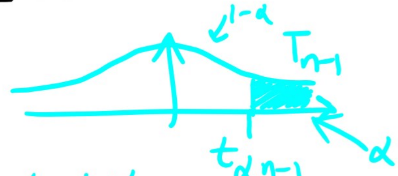


so that $P\left(\bar{x} - t_{\frac{\alpha}{2},n-1}\left(\frac{S}{\sqrt{n}}\right) < \mu < \bar{x} + t_{\frac{\alpha}{2},n-1}\left(\frac{S}{\sqrt{n}}\right)\right) = 1-\alpha$

& so C.I. for μ given by $\bar{x} \pm t_{\frac{\alpha}{2},n-1}\left(\frac{s}{\sqrt{n}}\right)$ where $\bar{x}$ & s computed using data set & $t_{\frac{\alpha}{2},n-1}$ is found using t-table.

Can also make one-sided $100(1-\alpha)\%$ confidence bounds:

upper: $\bar{x} + t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right)$



lower: $\bar{x} - t_{\alpha,n-1}\left(\frac{s}{\sqrt{n}}\right)$.