3236 Statistical Theory

3/28/23

Confidence intervals.

Assume that X_i is a random sample from a Normal population with μ and σ^2 unknown.

$$\sigma' = \left(\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{X})^2 \right)^{1/2}$$

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

$$U = \frac{N^{1/2} (\bar{X}_n - \mu)}{\sigma'}$$

U is a T r.v. with $N-1$ d.o.f.

Find c such that

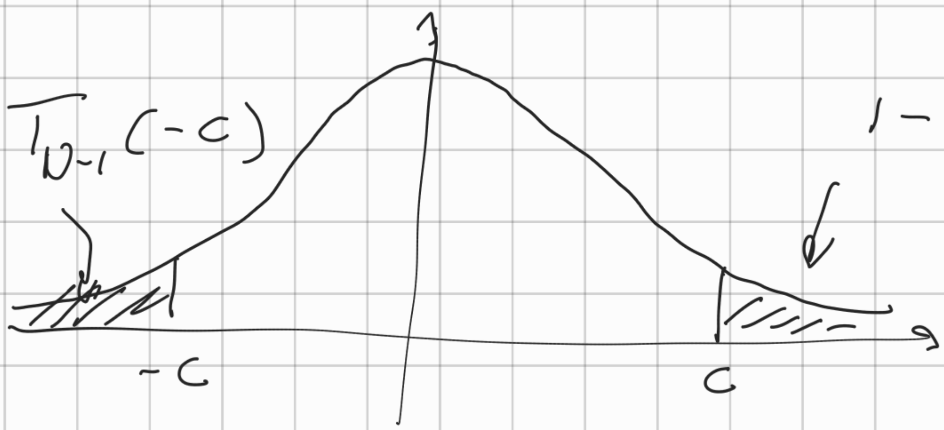
$$\mathbb{P}(-c \leq U \leq c) \geq \gamma$$

$$\Downarrow$$

$$\mathbb{P}\left(\bar{X} - \frac{c\,\hat{\sigma}'}{\sqrt{N}} \le \mu \le \bar{X} + \frac{c\,\hat{\sigma}'}{\sqrt{N}}\right) \geq \gamma$$

$$\mathbb{P}\left(-c \le V \le c\right) = T_{N-1}(c) - T_{N-1}(-c)$$

$T_{N-1}(x)$ is the c.d.f of a T r.v. with $N-1$ d.o.f.



$$T_{N-1}(-c) = 1 - T_{N-1}(c)$$

$$\gamma = 2\,T_{N-1}(c) - 1$$

$$\Downarrow$$

$$\mathbb{P}\left(-c \le V \le c\right) = \gamma$$

$$c = T_{N-1}^{-1}\left((1+\gamma)/2\right) = t_{N-1, \gamma/2}$$

for N large

$$t_{N,\gamma} \longrightarrow z_\gamma$$

With probability γ

$$\bar{X} - T_{N-1}^{-1}((1+\gamma)/2)\frac{\hat{\sigma}'}{\sqrt{N}} \le \mu \le \bar{X} + T_{N-1}^{-1}((1+\gamma)/2)\frac{\hat{\sigma}'}{\sqrt{N}}$$

coefficient γ confidence interval for μ.

100γ percent confidence interval

$\gamma = 0.95,\ 0.98,\ 0.99$

Def: Confidence Interval.

If we have two statistics $A(\underline{X})$ and $B(\underline{X})$ such that

$$\mathbb{P}\left(A(\underline{X}) \le g(\underline{\theta}) \le B(\underline{X})\right) \ge \gamma$$

for every value of $\underline{\theta}$

Then

$$[\, A(\underline{X}) \,,\, B(\underline{X}) \,]$$

is a coeff. γ conf. inter for $g(\underline{\theta})$. If we have

$$\mathbb{P}\left(A(\underline{X}) \le g(\underline{\theta}) \le B(\underline{X})\right) \underline{\underline{=}} \gamma$$

exact conf. inter.

In the case of normal sample we have

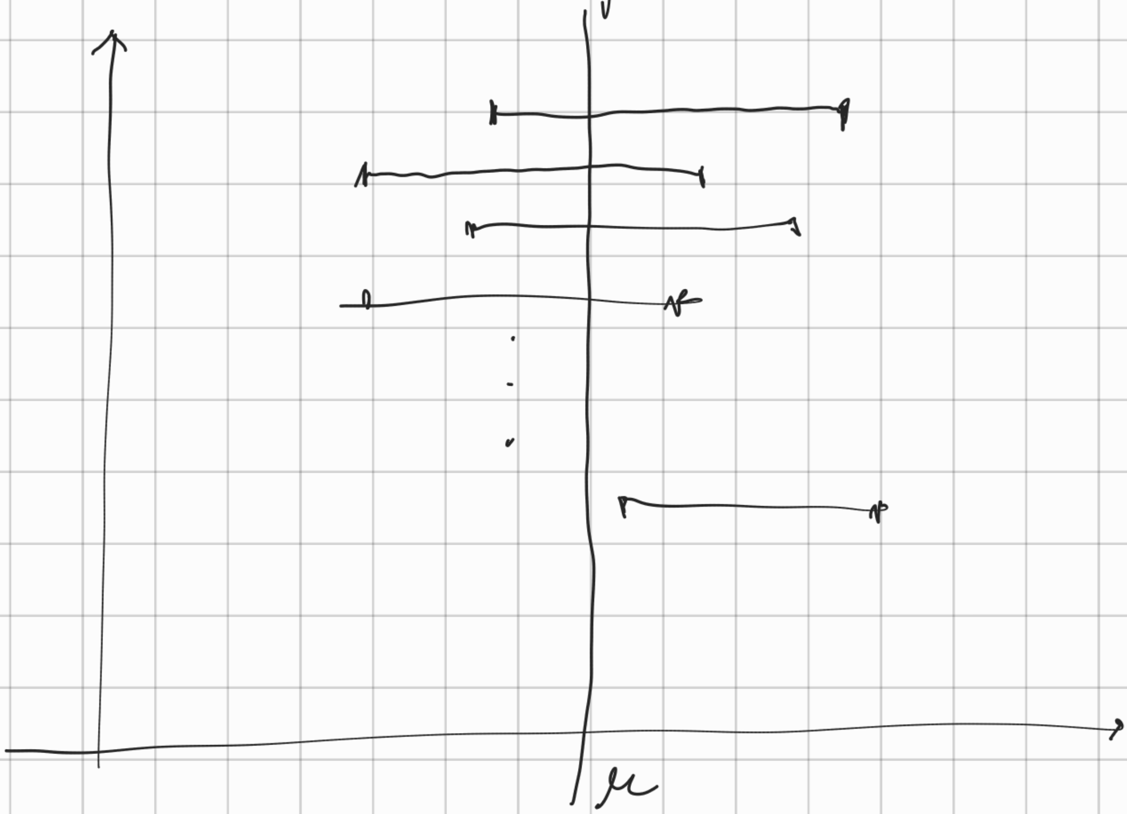
$$\underline{\theta} = (\mu, \sigma) \qquad g(\underline{\theta}) = \mu$$

$$A(\underline{X}) = \bar{X} - T^{-1}_{N-1}\left((1+\gamma)/2\right) \frac{\hat{\sigma}'}{\sqrt{N}}$$

$$B(\underline{X}) = \bar{X} + T^{-1}_{N-1}\left((1+\gamma)/2\right) \frac{\hat{\sigma}'}{\sqrt{N}}$$

Exact.

Run a sample. 95% C.I.



$x_i \qquad i = 1 \ldots N$

$$\bar{x} = \frac{1}{N} \sum_i x_i \qquad s' = \left(\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2 \right)^{1/2}$$

$$\bar{x} - T_{N-1}^{-1}((1+\gamma)/2) \frac{s'}{\sqrt{N}} \leq \mu \leq \bar{x} + T_{N-1}^{-1}((1+\gamma)/2) \frac{s'}{\sqrt{N}}$$

If N is large enough

$$T_{N-1}^{-1}((1+\gamma)/2) \to \Phi^{-1}((1+\gamma)/2)$$

What if X_i are not Normal?

$\mathbb{E}(X_i) = \mu$ $\quad\quad$ $Var(X_i) = \sigma^2$ ← Known

If N is small, ...

If N is large, we can use the CLT.

$$\bar{X} = \frac{1}{N} \sum_i X_i \simeq \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right)$$

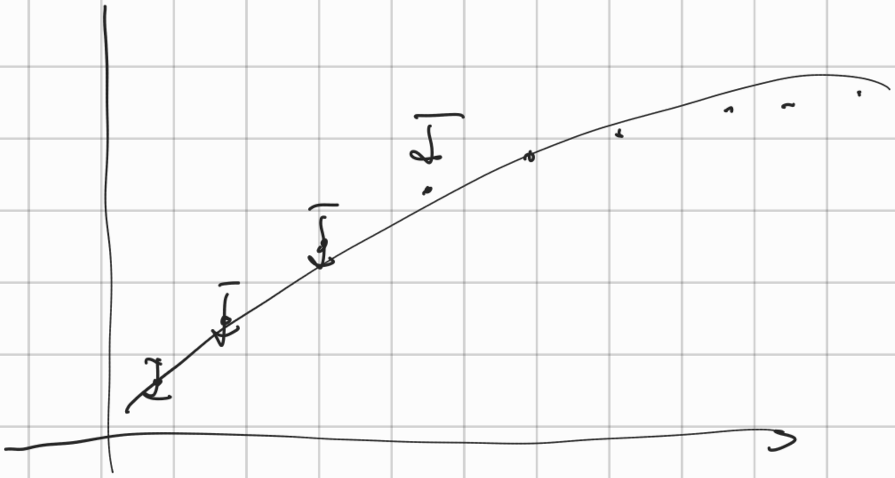
$$\mathbb{P}\left(\bar{X} - \phi^{-1}\left((1+\gamma)/2\right)\frac{\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + \phi^{-1}\left((1+\gamma)/2\right)\frac{\sigma}{\sqrt{N}}\right) \simeq \gamma$$

$N > 40$ mean N large.

$\gamma = 0.95$ $\quad \Phi^{-1}\left((1+\gamma)/2\right) = 1.96$

$\frac{}{2}$

$$\bar{x} - 1.96 \frac{\sigma}{\sqrt{N}} \le \mu \le \bar{x} + 1.96 \frac{\sigma}{\sqrt{N}}$$



σ

X_i are not Normal

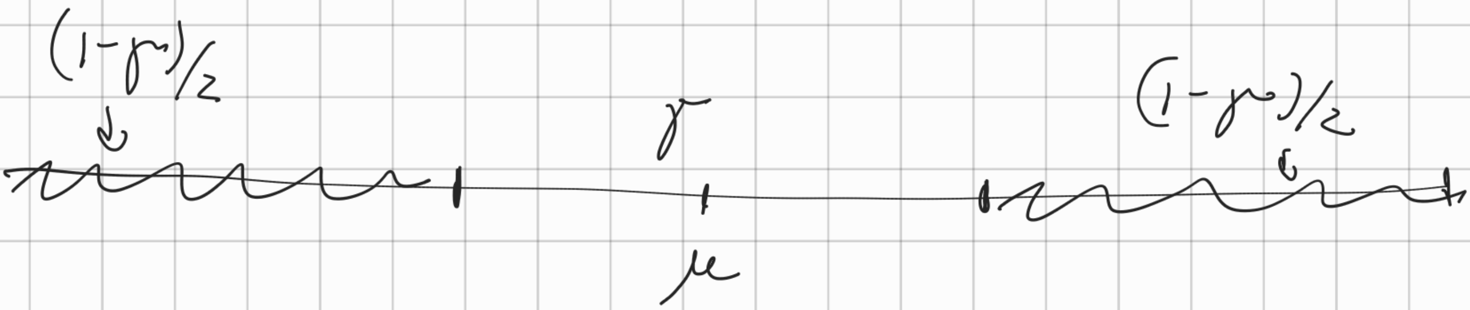
both σ and μ are

unknown.

$$\frac{N^{1/2}(\bar{X} - \mu)}{\hat{\sigma}'} \xrightarrow{d} N(0,1)$$

$N \to \infty$

$$\mathbb{P}\left(\bar{X} - \Phi^{-1}((1+\gamma)/2)\frac{\hat{\sigma}'}{\sqrt{n}} \le \mu \le \bar{X} + \Phi^{-1}((1+\gamma)/2)\frac{\hat{\sigma}'}{\sqrt{n}} \right) \approx \gamma$$



Coeff. γ

Upper confidence bound

$$\mathbb{P}\left(g(\underline{\theta}) \le B(\underline{X}) \right) \ge \gamma$$

Lower confidence bound

$$\mathbb{P}\left(A(\underline{X}) \le g(\underline{\theta}) \right) \ge \gamma$$

If X_i are Normal

$$\mathbb{P}\left(\mu \le \bar{X} + T_{N-1}^{-1}(\gamma) \frac{\hat{\sigma}'}{\sqrt{N}} \right) = \gamma$$

$$\underline{X} = X_1, X_2, \ldots X_N$$

$V(\underline{X}, \underline{\theta})$ whose distribution does not depend on $\underline{\theta}$.

$$V(\underline{X}, \underline{\theta}) = \frac{N^{1/2}(\bar{X} - \mu)}{\hat{\sigma}'} \sim T_{N-1} \quad \text{Normal}$$

$V(\underline{X}, \underline{\theta})$ pivotal quantity

There exists r

$$r(V(\underline{X}, \underline{\theta}), \underline{X}) = g(\underline{\theta})$$

Fix

$$A(\underline{X}) = r(G^{-1}(\gamma_1), \underline{X})$$

$$B(\underline{X}) = r\left(G^{-1}(\gamma_2), \underline{X}\right)$$

where $\gamma_2 > \gamma_1$ $\quad \gamma_2 - \gamma_1 = \gamma$

and G is the c.d.f. of $V(\underline{X}, \underline{\theta})$. Then we have

$$\mathbb{P}\left(A(\underline{X}) \leq g(\underline{\theta}) \leq B(\underline{X})\right) = \gamma$$

---

if X_i are exp. with par μ $\quad \mu X_i$ is exp par 1.

$\mu \sum_i X_i = \mu T_0$ has $\Gamma(N, 1)$