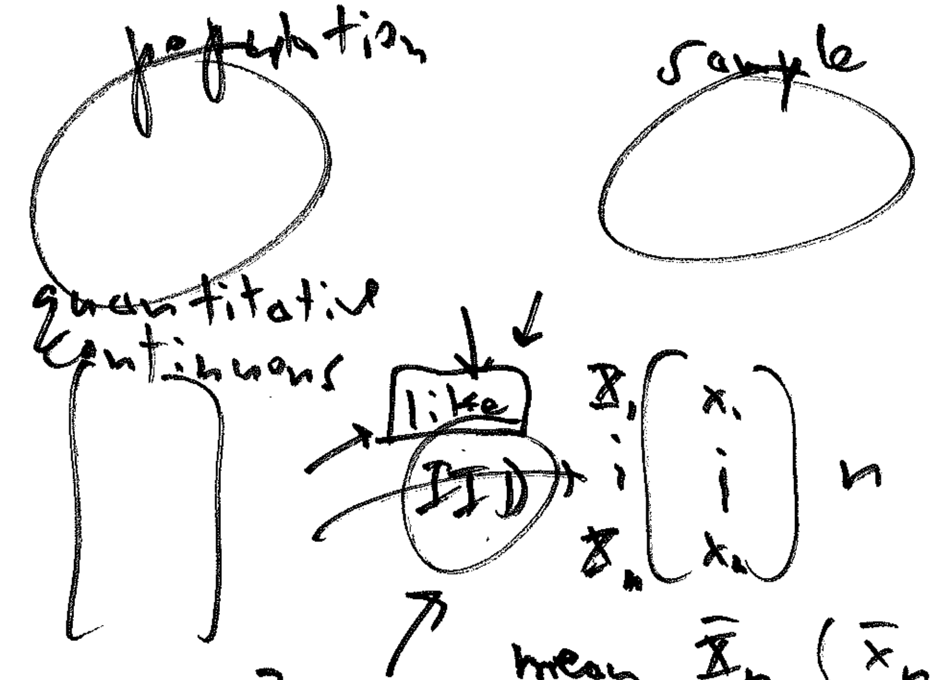
This time: confidence intervals, case studies

AMS 131
6 Jun 19

1

Confidence Intervals (Neyman 1937)



mean $\mu = ?$
SD $\sigma < \infty$ (assume)

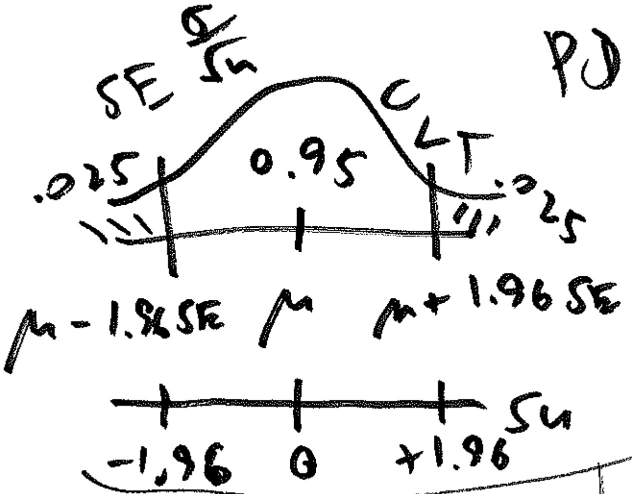
mean $\bar{X}_n$ $(\bar{x}_n)$
SD $s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$

Random variables way of writing down this diagram:

$$(X_i \mid \mu \, \sigma) \overset{IID}{\sim} \begin{cases} E(X_i) = \mu \\ SD(X_i) = \sigma \end{cases}$$

$(i = 1, \dots, n)$

This is a statistical inference problem, not a probability problem, because we know $\bar{X}_n$ (sample) and we're interested in quantifying our uncertainty about μ (pop)



suppose that n is big enough for CLT to yield a Normal distribution for $\bar{X}_n$

$E(\bar{X}_n) = \mu$

$V(\bar{X}_n) = \frac{\sigma^2}{n}$

$SD(\bar{X}_n) = SE(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}$

This picture says that (*)

$$P(\mu - 1.96\,SE \le \bar{X}_n \le \mu + 1.96\,SE) = 0.95$$

fixed unknown

random

Neyman's confidence trick

rearrange the inequalities

$$\mu - 1.96\,SE \le \bar{X}_n \overset{\text{iff}}{\longleftrightarrow} \mu \le \bar{X}_n + 1.96\,SE$$

$$\bar{X}_n \le \mu + 1.96\,SE \overset{\text{iff}}{\longleftrightarrow} \bar{X}_n - 1.96\,SE \le \mu$$

so (*) becomes

random

fixed unknown

$$0.95 = P(\bar{X}_n - 1.96\,SE \le \mu \le \bar{X} + 1.96\,SE)$$

Neyman now says, "Trust me: let's pretend that this is a probability statement about μ (when in fact it's a probability statement about $\bar{X}_n$)."

He proposes $(\bar{X}_n - 1.96\,SE, \bar{X}_n + 1.96\,SE)$ as what he calls a 95% <u>confidence interval</u> (CI) for μ.

$$SE = SE(\bar{X}_n) = \frac{\sigma}{\sqrt{n}}$$

Complication: what if σ is not known?

Obvious thing to try: Since IID sampling is a good way to learn about μ, it should also be a good way to learn about σ: let's estimate σ (pop. SD) by s (sample SD).

$\hat{SE} = \hat{SE}(\bar{X}_n) = \frac{\cancel{\sigma}\, s}{\sqrt{n}}$ ~~(^ = "estimated")~~

So Neyman's 95% CI for μ becomes

$$\left(\bar{X}_n - 1.96 \frac{s}{\sqrt{n}}, \bar{X}_n + 1.96 \frac{s}{\sqrt{n}} \right),$$ which

can also be written $\bar{X}_n \pm 1.96 \frac{s}{\sqrt{n}}$.

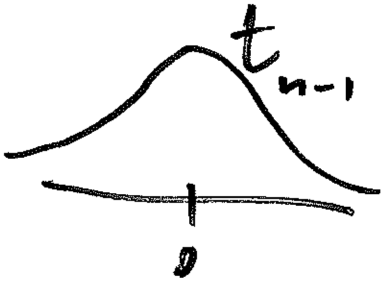
Your something-for-nothing bell should be ringing: we just pretended that $\sigma = s$, when in fact s is random and would come out a bit differently if we took another IID sample of size n ~~[illegible]~~ we need to pay a price for having estimated σ, and intuitively that price should go up as $n \downarrow$ (less data = more uncertainty)

Small-sample correction

$$\bar{X}_n \pm 1.96 \frac{s}{\sqrt{n}}$$

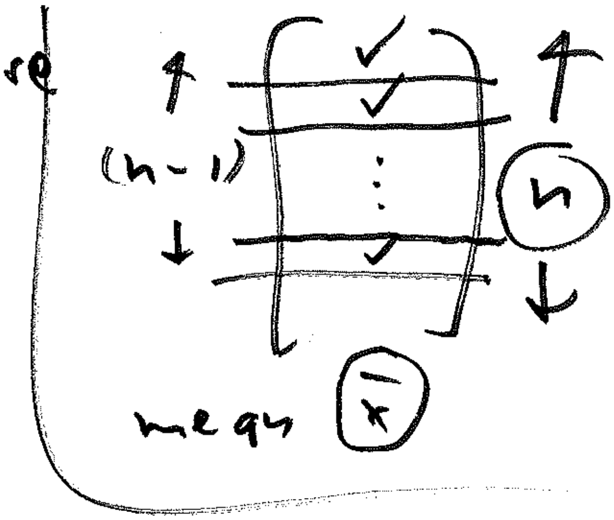
This is called the confidence number, which we got from the Normal curve when we were pretending that $\sigma = s$.

William Gosset ("Student") (1908) showed that, if the population PDF of X_i is Normal, we should get our confidence numbers not from the Normal curve but from what we now call a member



t_{n-1} PDF of $\dfrac{\bar{X}_n - \mu}{s/\sqrt{n}}$

of the family of t distributions, indexed by a quantity called the <u>degrees of freedom</u> of s^2 as an estimate of σ^2: because our estimate of σ^2 is centered at $\bar{x}$, the sum of squares

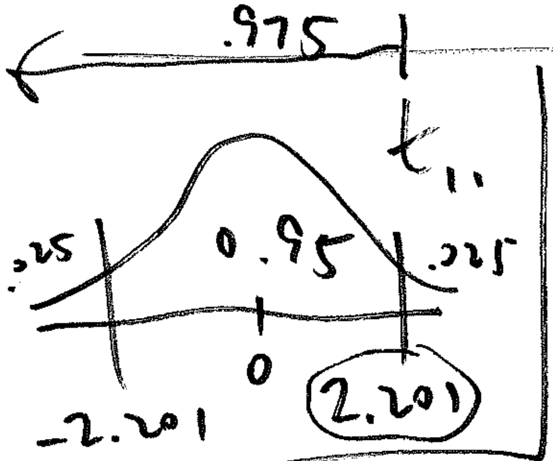


$\sum_{i=1}^{n} (x_i - \bar{x})^2$ has n terms in it but only $(n-1)$ of them are free to vary, because $\sum_{i=1}^{n} (x_i - \bar{x}) = 0$.

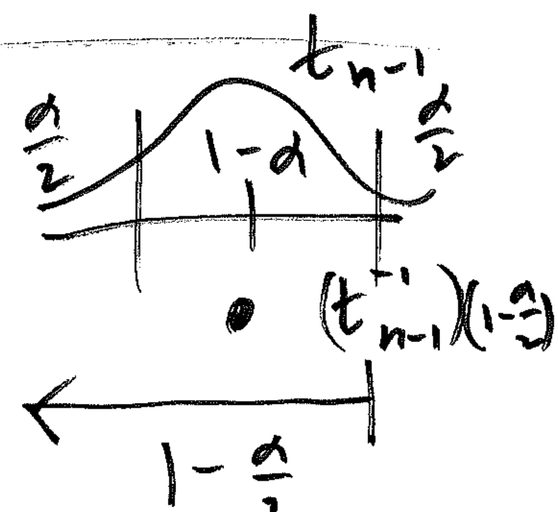
So the right degrees of freedom for the t curve we need to use in this problem is $(n-1)$.

⑦

As $n\uparrow$, intuitively we don't need to pay as big a price for pretending that $\sigma = s$, so the t_n curves approach the standard Normal curve as $n\uparrow$



In THT 3 problem 2, $n = 12$ so we want the t curve with 11 degrees of freedom: the right confidence number is 2.201 instead of 1.96 (12% larger). In general to get a $100(1-\alpha)\%$ CI the right confidence # is $t^{-1}_{n-1}\left(1-\frac{\alpha}{2}\right)$,



$t^{-1}_{n-1}\left(1-\frac{\alpha}{2}\right)$ = the $\left(1-\frac{\alpha}{2}\right)$ quantile of the t distribution with $(n-1)$ degrees of freedom, so Neyman's $100(1-\alpha)\%$ CI for μ is

$$\boxed{\bar{X}_n \pm t^{-1}_{n-1}\left(1-\frac{\alpha}{2}\right)\frac{s}{\sqrt{n}}} \quad (*)$$

THT 3 problem 2

$\bar{X}_n = 18.6$ mmHg

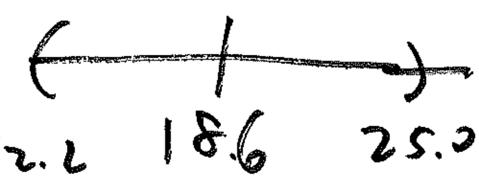
$n = 12$

$s = 10.1$

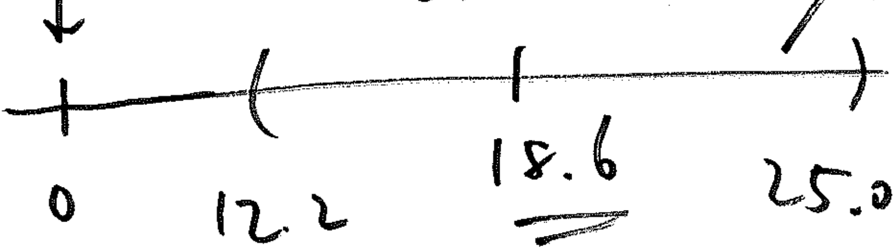
$t^{-1}_{11}(0.975) = 2.201$

$$(*) = \underset{\text{mmHg}}{18.6} \pm \underbrace{2.201\,\frac{10.1}{\sqrt{12}}}_{6.4\text{ mmHg}}$$

mmHg



95% CI for $\mu = \Delta$



under our assumptions (9), we are 95% confident that $\mu = \Delta$ is between 12.2 and 25.0.

$Y_n = a + b + e_n$

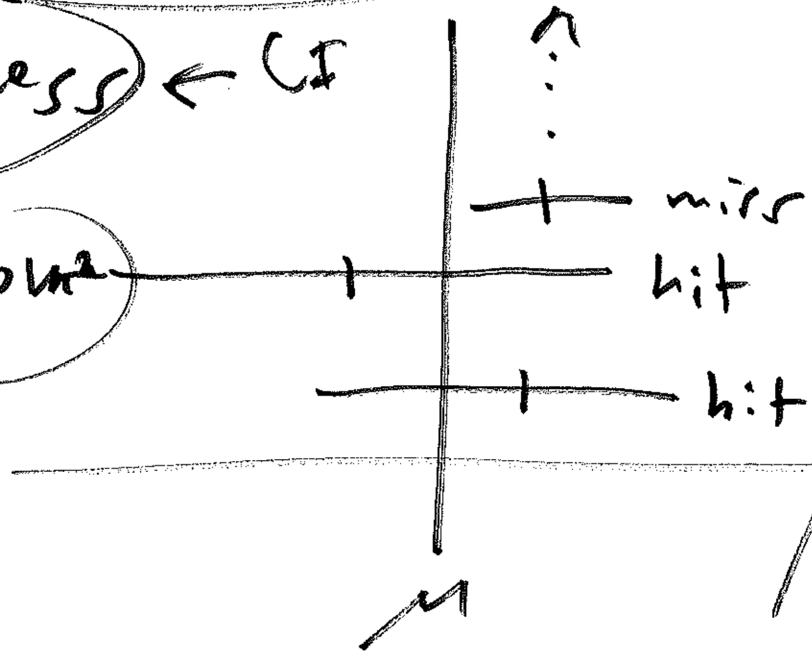
Q:) Does this mean that

$P(12.2 < \mu < 25.0) = \cancel{0.95}$?

F ← frequentist

process ← CI

outcome



Neyman: about 95% of these CIs will be hits devil's advocate

I think $\mu = 0 = \mu_0$ "null" ← no diff.

statistical significance

See if devil's advocate μ_0 is in your CI: if not, declare difference between $\bar{X}_n$ and μ_0 statistically significant; if μ_0 is in, not (stat sig).

on ave (captopril not useful on ave.)

at 95% confidence level

diff. stat sig

probably real

large in statistical terms, so hard to attribute to unlucky sampling

(11)

2nd (different) question Is

the difference large in practical terms? (practically significant) (practice)

A1: best (clinical) expert judgment

A2: relative improvement:

$$\frac{\overset{T}{166.8} - \overset{C}{185.3}}{\underset{C}{185.3}} \doteq -10\%$$

decline

no hard & fast rule

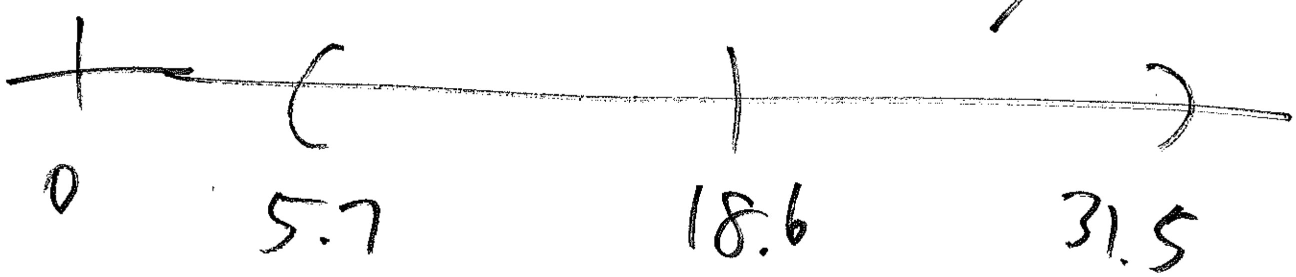
pinch of salt:

25% improvements often practically meaningful

99.9% CI

$$18.6 \pm \underbrace{\boxed{4.437}\,\frac{10.1}{\sqrt{12}}}_{12.9}$$

99.9% CI for μ



stat sig at 99.9% level

assuming bias = 0, but

bias $\geq$ 5.7 → no longer stat sig

publish a lot → get tenure | find lots of stat sig differences hypotheses

null : $\mu = 0$

alt : $\mu \neq 0$

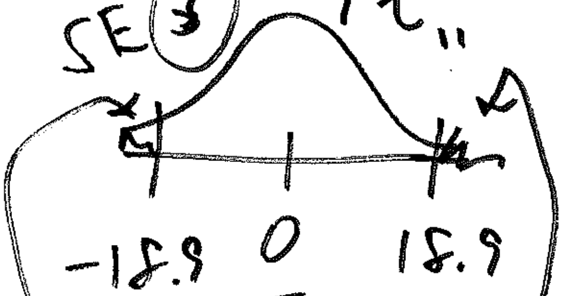
Fisher

$\boxed{p \le 5\%}$

$\Updownarrow$

null value is not in 95% CI $\Leftrightarrow$ statsig

$SE(\bar{x}) = \frac{10.1}{\sqrt{12}}$ "t_{11}" PDF of $\bar{x}$ if null true



p-value $\doteq$ 0%

$$\frac{18.9 - 0}{10/\sqrt{12}} \doteq +6$$

Fisher statsig 5%

more replicable statsig 0.5%