Confidence intervals
(Neyman 1937)

Random variable
way of writing down
this diagram:

$\sum (X_i - \mu_0) \overset{\text{IID}}{\sim} \{E(X_i) = \mu \cap \text{Cov}(X_i) = 0\}$

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 $\mu$.
The inequalities:

\[ \mu \leq \mu + 1.96 \times SE(\bar{X}) \]

\[ \mu - 1.96 \times SE(\bar{X}) \leq \bar{X} \leq \mu + 1.96 \times SE(\bar{X}) \]

\[ SE(\bar{X}) = \frac{s}{\sqrt{n}} \]

\[ \nu \approx 1.96 \]

\[ E(\bar{X}) = \frac{\nu}{n} \]

\[ SE(\bar{X}) = \frac{s}{\sqrt{n}} \]

\[ P(\bar{X} - 1.96 \times SE(\bar{X}) \leq \bar{X} \leq \mu + 1.96 \times SE(\bar{X}) \] is close to a normal distribution for CLT to yield enough samples.
Neyman now says, "Trust me: let's pretend that this is a probability statement about \( \mu \) (when in fact it's a probability statement about \( \bar{x} \))."

He proposes \( (\bar{x}_n - 1.96 \text{SE}, \bar{x}_n + 1.96 \text{SE}) \) as what he calls a 95% confidence interval for \( \mu \). \[
\text{SE} = \frac{\sigma}{\sqrt{n}}
\]

Complication: what if \( \sigma \) is not known?

Obvious thing to try: Since \( \sigma \) is known,

is a good way to learn about \( \mu \),
it should also be a good way to learn about \( \sigma \): let's estimate \( \sigma \) by \( \bar{x}_n \).
\[ \bar{x}_n = \overline{\bar{x}_n} = \frac{\bar{x}}{\sqrt{n}} \]
\( \text{So, Neyman's 95% CI becomes} \)
\[ \left( \bar{x}_n - 1.96 \frac{s}{\sqrt{n}}, \bar{x}_n + 1.96 \frac{s}{\sqrt{n}} \right) \]

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
\( s = s \) when in fact \( s \) is random and would come out a bit differently if we took another IID sample of size \( n \).

\[ \text{we need to pay a price for having estimated } s, \text{ and} \]
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 \( \bar{X} \) is Normal, we should get our confidence numbers not from the Normal curve but for what we now call a member...
of the family of \( t \) distributions indexed by a quantity called the degree of freedom of \( s^2 \) as an estimate of \( \sigma^2 \): because our estimate of \( \sigma^2 \) is centered at \( \bar{x} \), the sum of \( s \) squares \( \sum (x_i - \bar{x})^2 \) has \( n \) terms in it, but only \( (n-1) \) of them are free to vary, because \( \sum (x_i - \bar{x}) = 0 \).

So the right degrees of freedom for the \( t \) curve we need to use in this problem is \( (n-1) \).
As \( n \to \infty \), intuitively we don't need to pay as big a price for pretending \( n \to \infty \), so the \( t \) curves \( \to \) standard normal curve as \( n \to \infty \).

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 \( 100(1-d)\% \) CI the right confidence \# is \( t^{1-d/2}_{n-1} \).
The $(1-\alpha/2)$ quantile of the $t$ distribution with $(n-1)$ degrees of freedom is $t_{n-1}^{-1}(1-\alpha/2)$.

For $\mu$ is

\[
\bar{x}_n \pm t_{n-1}^{-1}(1-\alpha/2) \frac{s}{\sqrt{n}} \quad \star
\]

**Problem 2**

$\bar{x}_n = 18.6 \text{ mm Hg}$

$n = 12$

$s = 10.1$

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

$18.6 \pm 2.201 \frac{10.1}{\sqrt{12}}$

$18.6 \pm 6.4 \text{ mm Hg}$

$12.2 \leq \bar{x}_n \leq 25.0$
Process \rightarrow \text{Outcome}

F

\int \text{frequentist}

\frac{122}{n} = 25.0

122 \pm 1.96 \times \frac{25.0}{\sqrt{n}}

\frac{10}{9} = 1.11

\frac{122}{18.6} = 6.59

\frac{12}{25.0} = 0.48

\frac{100}{50} \% = 4\%

\frac{9}{m} = \Delta

\frac{10}{120} = 0.083

\text{95\% confidence interval for m = \Delta under our assumptions}
I think $\mu = \mu_0$ is not statistically significant, or we (captopril not useful, e.g.).

$m_0$ is in your CI; if not, declare difference between $\bar{m}_n$ and $m_0$ statistically significant.

If $m_0$ is in, not (statistically).

At 95% confidence level, probably real. Diff.

Large in statistical terms, hard to attribute to unlucky sampling.
2nd (different) question: Is the difference large in practical terms? (practically significant)

A₁: (clinical) expert judgment

A₂: relative improvement:

No hard & fast rule

Pinch of salt:

25% improvement is not meaningful

166.8 - 185.3 = -10% decline
99.9% CI
18.6 ± 4.437 \frac{10.1}{\sqrt{12}}

99.9% CI for \mu

0 \quad 5.7 \quad 18.6 \quad 31.5

statistical at 99.9% level

turning \ \text{bias} = 0 \ \text{but}

\text{bias} \leq 5.7 \ \rightarrow \ \text{no lower statistics}

\downarrow \quad \text{set}

\text{publish} \rightarrow \ \text{tenure}

\downarrow \quad \text{find lots of}

\text{statistics difference}
hypothesis:
null: $\mu = 0$
alt: $\mu \neq 0$

Fisher:
$p \leq 5\%$

null value is not in 95% CI $\Rightarrow$ reject

$Z = \frac{10.1}{SE} = \frac{10.1}{\sqrt{10/5n}}$

PDF of $Z$ if null true:
-15.9 0 15.9

$p$-value $= 0\%$

$15.9 - 0 \geq +6$

Fisher $= 5\%$

were

replicable

stats

0.5\%