Recall: we can estimate a *95% confidence interval* for an unknown population mean μ by using the sample mean $\overline{X}$ and the population standard deviation σ:

$$\overline{X} \pm 1.96 \cdot SE$$

or equivalently

$$\overline{X} \pm 1.96 \cdot \frac{\sigma}{\sqrt{N}}$$

What if we are not given σ? Can we use our estimate $\hat{\sigma} = \sqrt{\dfrac{SS}{N-1}}$?

Well, yes – sort of – but we have to adjust the 1.96 part.

Why?

- 1.96 is used because for a normal distribution, 95% of sample means fall between $-1.96 \cdot SE$ and $1.96 \cdot SE$. This assumes σ is known/given.
- if estimating σ via $\hat{\sigma}$, we get a t-distribution for the distribution of sample means. The exact shape of this distribution *depends on sample size*

In light of this, let's define a *generalized* confidence interval:

$$\overline{X} \pm t^*_{df} \cdot SE$$

where

1. the exact value of t^*_{df} depends on sample size
   - defined as the value of t which leaves 5% of distribution in the tails (both tails combined) – sometimes called the *critical value* of the t-distribution
2. the formula for SE depends on design:
   - for single group (or paired samples), we have $SE = \dfrac{\hat{\sigma}}{\sqrt{N}}$
   - for independent samples, we have $SE = \hat{\sigma}_p \sqrt{\dfrac{1}{N_1} + \dfrac{1}{N_2}}$

Example 1 (single group design): A sample of 25 people is given a treatment. After treatment, we find $\overline{X} = 22.2$ with $SS = 384$. Construct a 95% confidence interval for μ, the population mean for the treatment group.

For independent groups designs, the goal is to estimate the "mean difference" $\mu_1 - \mu_2$. The resulting CI is

$$(\overline{X}_1 - \overline{X}_2) \pm t^*_{df} \cdot \hat{\sigma}_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}$$

Example 2: Does watching educational TV as a kid predict better high school grades?

Educational TV	No Educational TV
$N_1 = 10$	$N_2 = 10$
$\overline{X}_1 = 93$	$\overline{X}_2 = 85$
$SS_1 = 200$	$SS_2 = 160$

Compute a 95% confidence interval for the mean difference $\mu_1 - \mu_2$.

.