Confidence Intervals

Hypothesis testing

Confidence Intervals

$\widehat{\text{estimate}} \pm \underbrace{t_c \times \widehat{SE}}_{\text{error bounds}}$

$\bar{y} \pm t_c \times \widehat{SEM}$ .. $\hat{\delta} \pm t_c \times \widehat{SED}$

size

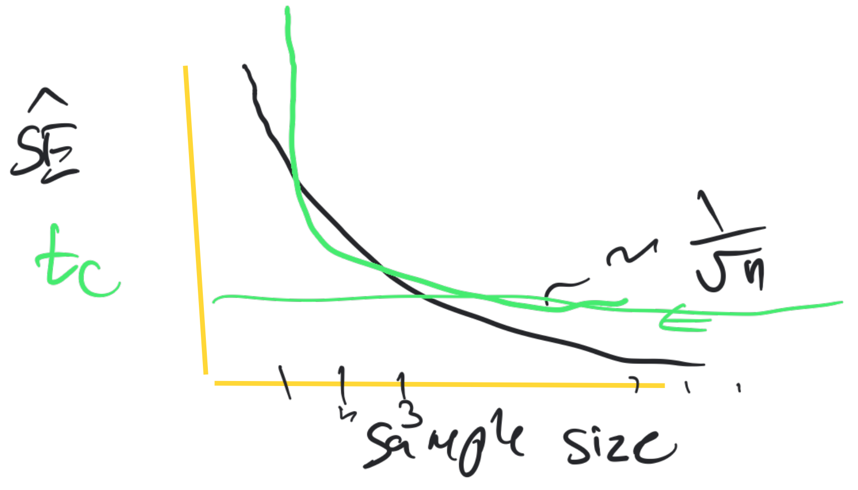
95% $\Rightarrow \alpha = 0.05$

$t_c = t_{\frac{\alpha}{2}, df} = qt(\frac{\alpha}{2}, df, \text{low.tail}=F)$ < α - coverage; df - sample size - 2

$\widehat{SE} = \widehat{SED} = \sqrt{\frac{s_2^2}{n_2} + \frac{s_1^2}{n_1}}$

- s^2's - experimental design, - reduce σ^2
- n's - sample size

$s^2 \rightarrow \hat{\sigma}^2 \rightarrow \sigma^2$



Diminishing returns

Sample size	Int. size	
2	9σ	
		> 5.6x
4	1.6σ	
		> 1.9x
8	0.8σ	
		> 1.6x
16	0.5σ	

$Y_i \sim N(\mu, \sigma^2)$

Is the reference pop'n Normal?

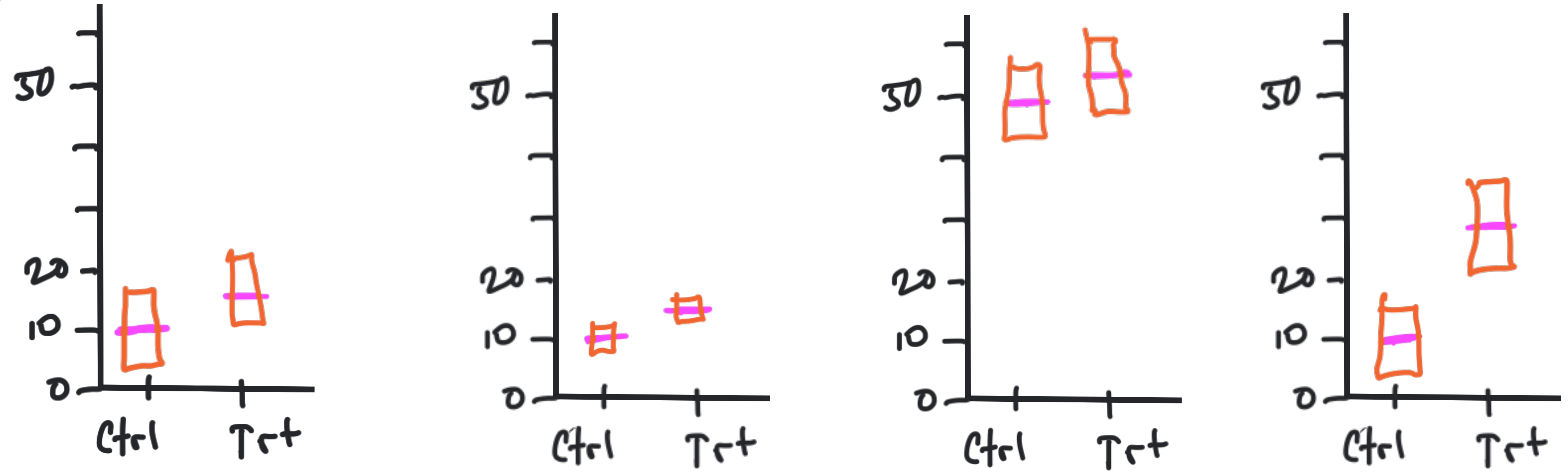
1) many are.
2) $\dfrac{\bar{y}-\mu}{SEM}$ or $\dfrac{\hat{\delta}-\delta}{SED} \sim N(0,1)$ if n is large

Central limit theorm.

3) use our data to check...
   - histogram
   - qq plots
   - ~~test for Normality~~

$\boxed{n_c = n_t = 32}$

A B C D



	A	B	C	D
δ	5	5	5	20
σ	8	2	8	8
SED	2	0.5	2	2
$\frac{\delta}{\mu_c}$	$= 5/10 = 0.5$	0.5	$5/50 = 0.1$	2
$\frac{\delta}{SED}$	$= \frac{5}{2} = 2.5$	$\frac{5}{0.5} = 10$	$\frac{5}{2} = \cancel{2.5}$	$\frac{20}{2} = 10$

- which effect is largest?
- Which effect is most important?
- which effect is most significant?

t-test

p-value

Hypothesis testing

Confidence Interval

- bounds on error
- how good is our estimate?

1) estimate: $\hat{\delta} = \bar{y}_2 - \bar{y}_1$

2) $\widehat{SED}$: $\sqrt{\frac{s_2^2}{n_2} + \frac{s_1^2}{n_1}}$

$\hat{\delta} \pm t_c \times \widehat{SED}$

qt function

α, df

T-test

- could the TRUE effect be = 0?

Null hypothesis

$H_0: \delta = 0$

Statistic

$\hat{t} = \frac{(\hat{\delta} - 0)}{\widehat{SED}}$ — $\frac{\text{observed effect}}{\text{typical error}}$

$|\hat{t}| > 1$ — effect is bigger than a typical error

- bigger $\hat{t}$, more confident $\delta \neq 0$

P-value: $P(|t| > |\hat{t}|)$ — our data when δ is unknown

t from random expt ~~~~ when $\delta = 0$

T-test

calculate $\frac{\hat{\delta}}{SE\hat{\delta}} = \hat{t}$

bigger $|\hat{t}|$ $\Rightarrow$ smaller p-value

$\Rightarrow$ less plausible that $\delta = 0$

H_0 is TRUE

What do we do w/ T-test?

1) Making a decision Y/N ⟷ 2) weighing evidence

both

- more / less significant

Weigh consequences of being wrong

H_0	Accept	Reject — Declare $\delta \neq 0$
$\delta = 0$ TRUE	✓	X Type I error — false positive α
False $\delta \neq 0$	Type II error	✓

└ δ might be 0

Lost opportunity

false negative

β

α = Prob of Reject when $\delta = 0$

β = Prob of Accept when $\delta \neq 0$

Decision strategy

1) choose α $\Rightarrow$ calculate t_c — $t_{\alpha/2, df}$ z t

if $|\hat{t}| > t_c$ $\Rightarrow$ Reject H_0 $\quad \delta \neq 0$

if $|\hat{t}| < t_c$ $\Rightarrow$ Accept H_0. $\quad \delta$ could be 0.

2) calculate pvalue.

$p \leq \alpha$ $\Rightarrow$ Reject

$p > \alpha$ $\Rightarrow$ accept