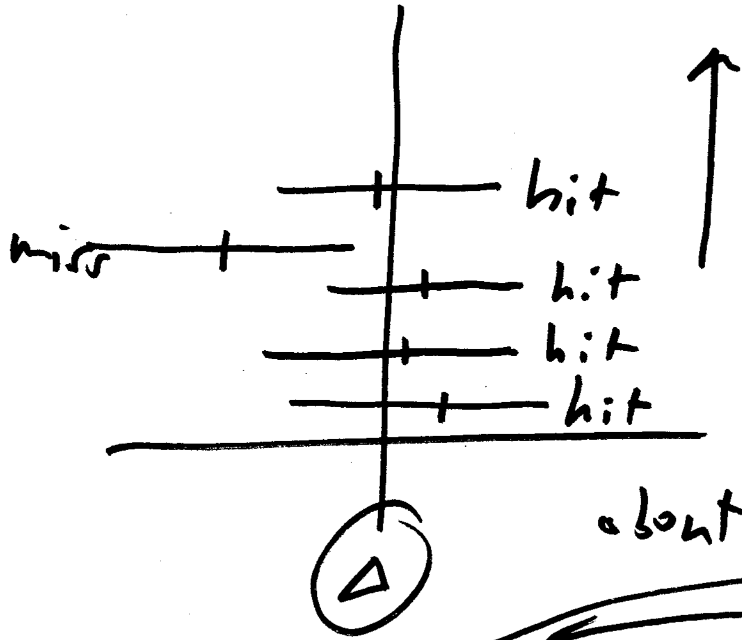
STAT 131
8 Jun 20

THT 3
#2a

DD extra office 1.5-hour session



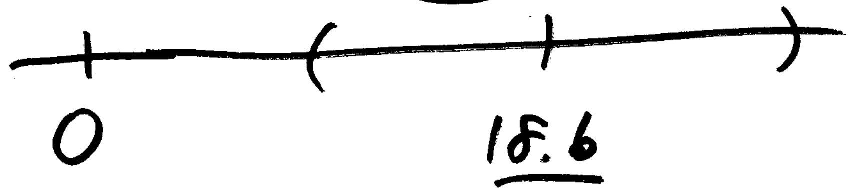
confidence number

$\hat{\Delta} \pm (\quad) \, \hat{SE}(\hat{\Delta})$

about 95% hits

R. A. Fisher (1935) (1)
statistician & ~~geneticist~~

95% CI for Δ



α (no effect on average of Captopril in the population)

~~since 0 is not in 95% CI,~~

the difference between $\hat{\Delta} = 18.6$ mmHg and the average no-effect value of 0 mmHg is statistically significant ↔ is hard to attribute to unlucky random sampling ↔ is probably real

This is a positive finding! we think Captopril works. But

That brings up the possibility of a false positive (worse). Science is eager to avoid false positives. (*)

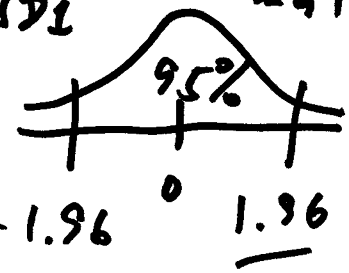
False negative here: we decide Captopril doesn't work when it does.

(*) To decrease chance of a false (+), we can raise the confidence level, e.g., from 95% to (e.g.) 99.9% (99.0, 99.5, 99.7)

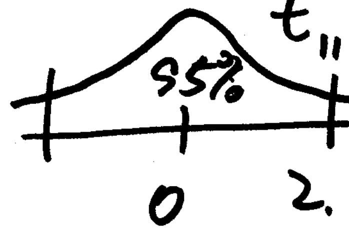
$\hat{SE}$ $\frac{s_D}{\sqrt{n}}$

PDF of $\bar{D}$ →

cheating: standard normal



PDF $\frac{\bar{D} - \Delta}{s_D / \sqrt{n}}$



t_{11} curve (10% wider)

n = 12

~~"Student"~~

W. Gosset ← n small

(1908 ←)

$$\hat{\Delta} \pm \boxed{\begin{matrix} 95\% \\ (z \\ \#) \end{matrix}} \cdot \hat{SE}(\hat{\Delta})$$

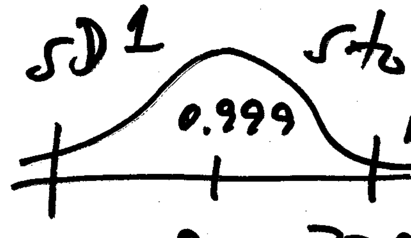
✓ ✓

↑

want this # to increase to account for our cheating

as n ↑ t curves → normal curve

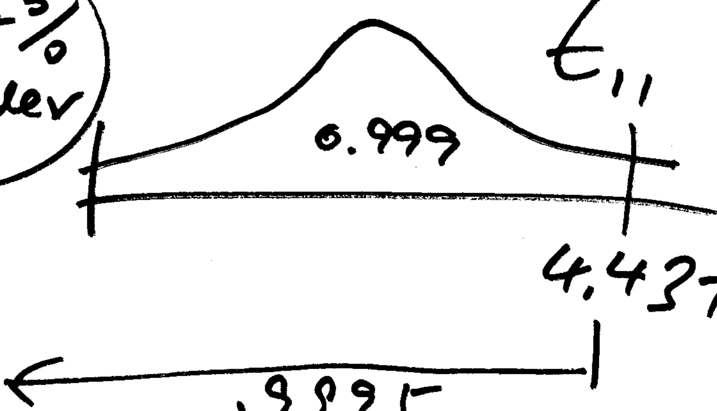
$$s_D = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (d_i - \bar{d})^2}$$



standard normal

35% wider

$\frac{4.437}{3.2905}$

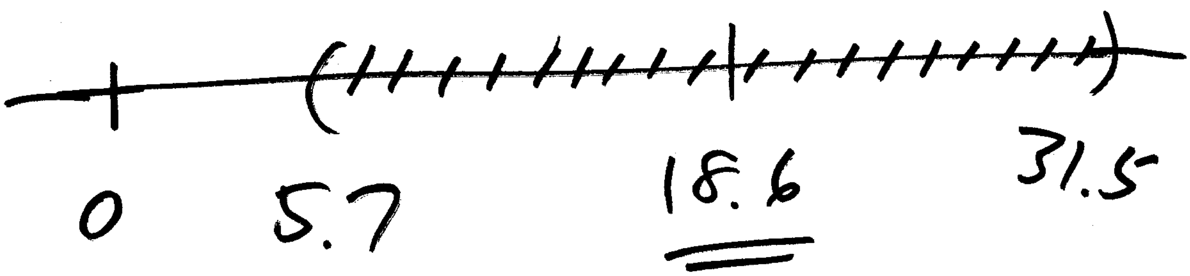


100(1−α)% CI for Δ with small n

$$\hat{\Delta} \pm \left(t_{n-1}^{-1}\right)^{1-\frac{\alpha}{2}} \cdot \frac{s_D}{\sqrt{n}}$$

$$18.6 \pm (4.437)\underbrace{\left(\frac{10.1}{\sqrt{12}}\right)}_{2.9}$$

99.9% CI for Δ

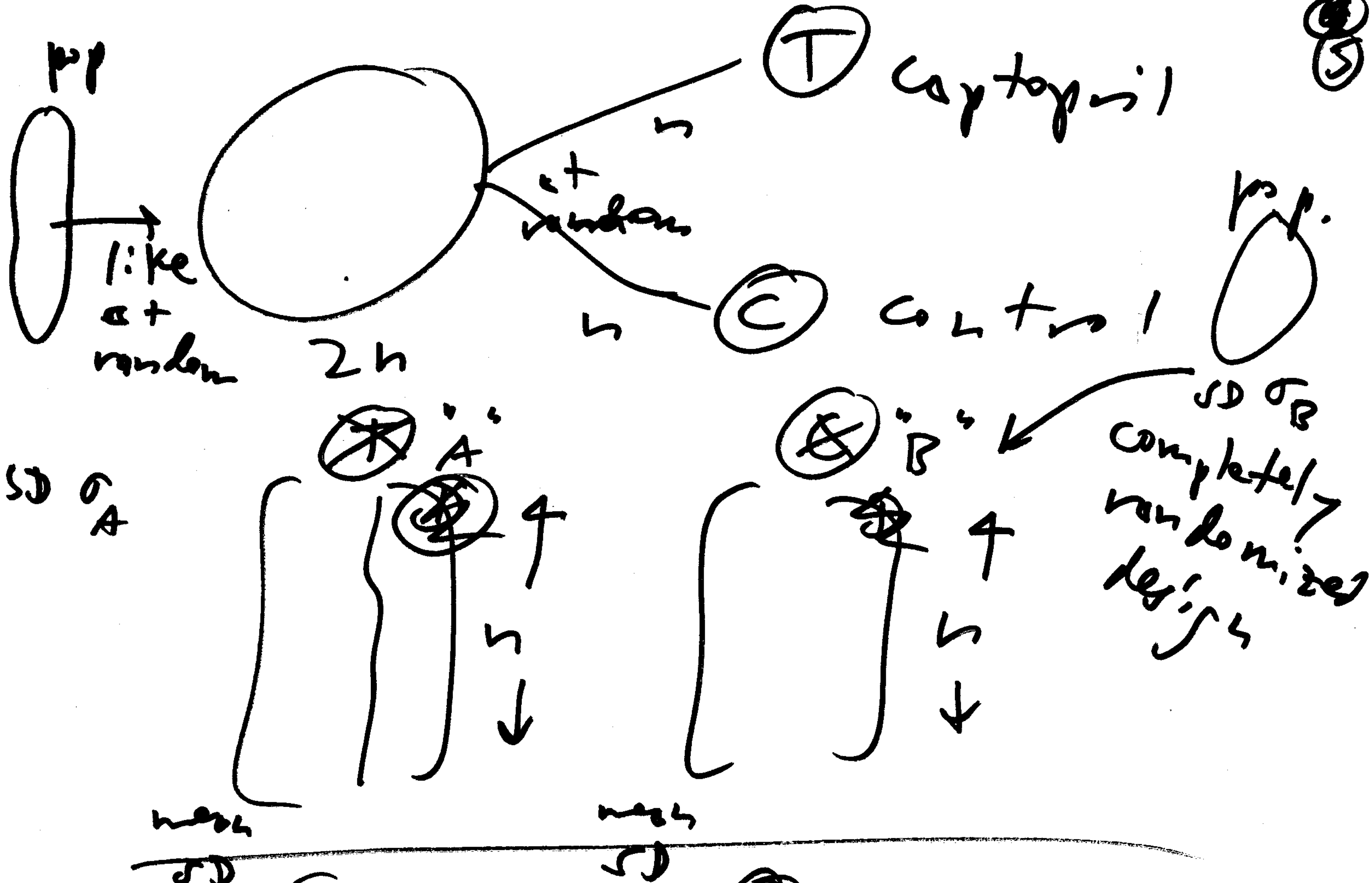


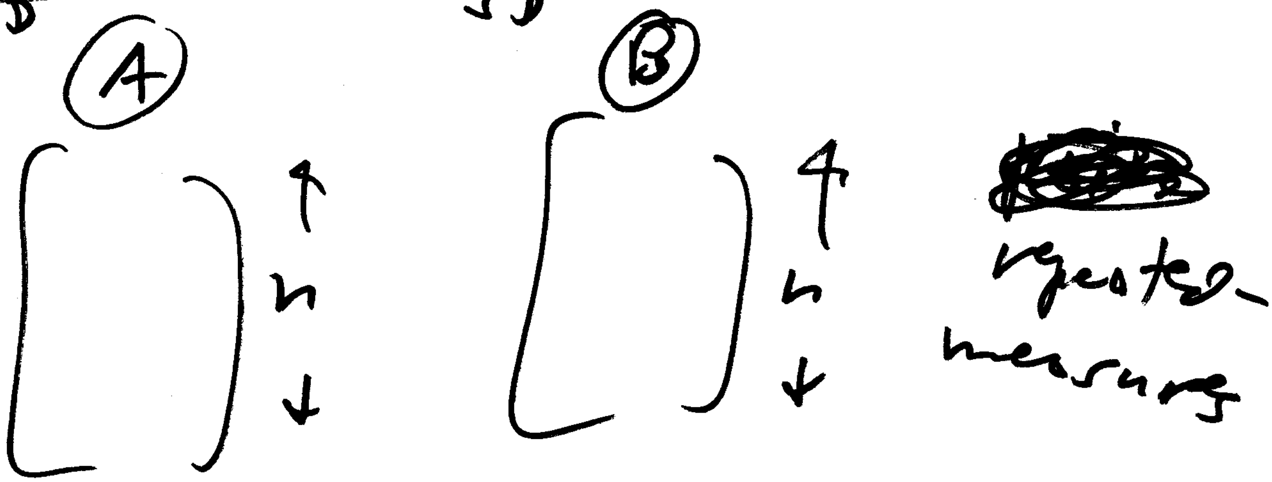
↑ null hypothesis value for Δ $= \Delta_0$

the diff. between $18.6 = \hat{\Delta}$ (data) & 0 (no average effect in pop.)

is stat sig at 99.9% conf. level

$\leftrightarrow$





$$V_{CR}(\bar{D}_n) = V_{\underline{CR}}\left(\bar{B}_n - \bar{A}_n\right)$$

$$\overset{\text{I}}{=} V_{CR}(\bar{B}_n) + (-1)^2 V_{CR}(\bar{A}_n)$$

$$B_i \overset{IID}{\sim} \left\{ \begin{matrix} E(B_i) = \mu_B \\ V(B_i) = \sigma_B^2 \end{matrix} \right\} \qquad \bar{B}_n = \frac{1}{n} \sum_{i=1}^{n} B_i$$

$$V_{CR}(\bar{B}_n) = \frac{\sigma_B^2}{n}$$

$$A_j \overset{IID}{\sim} \left\{ \begin{matrix} E(A_j) = \mu_A \\ V(A_j) = \sigma_A^2 \end{matrix} \right.$$

$$\bar{A}_n = \frac{1}{n} \sum_{i=1}^{n} A_j$$

$$V_{CR}(\bar{A}_n) = \frac{\sigma_A^2}{n}$$

$$\text{so } V_{CR}(\bar{D}_n) = \frac{\sigma_A^2 + \sigma_B^2}{n} \checkmark$$

$$V_{RM}(\bar{D}_n) = V_{RM}\left[\frac{1}{n} \sum_{i=1}^{n} (B_i - A_i) \right]$$

$$= \frac{1}{n^2} V_{RM}\left[\sum_{i=1}^{n} \left(\boxed{B_i - A_i} \right) \right]$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} V_{RM}(B_i - A_i)$$

$$V_{RM}(B_i - A_i) =$$

$$V_{RM}(B_i) + V_{RM}(-A_i)$$

$$+ 2\, C_{RM}(B_i, -A_i)$$

$$= V_{RM}(B_i) + (-1)^2 V_{RM}(A_i)$$

$$- 2\, C_{RM}(\underline{\underline{B_i, A_i}})$$

$$\frac{C_{RM}(A_i, B_i)}{SD(A_i) \cdot SD(B_i)} = \rho_{AB}$$

$$V_{RM}(\bar{D}_n) = \frac{1}{n^2}\sum_{i=1}^{n} V_{RM}(B_i - A_i)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\Big[\overbrace{V(A_i)}^{\sigma_A^2} + \overbrace{V(B_i)}^{\sigma_B^2} - 2\rho_{AB}\,\sigma_A\cdot\sigma_B\Big]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left(\sigma_A^2 + \sigma_B^2 - 2\rho_{AB}\,\sigma_A\sigma_B\right)$$

$$= \frac{n\left(\text{———}\right)}{n^2} = \frac{\sigma_A^2 + \sigma_B^2 - 2\rho_{AB}\,\sigma_A\sigma_B}{n}$$ ✓

$$\hat{V}_{RM}(\bar{D}_n) = \frac{s_A^2 + s_B^2 - 2\overbrace{\hat{\rho}_{AB}}^{r_{AB}}\,s_A s_B}{n}$$

$$\hat{\rho}_{AB} = 0.8084$$