STAT 206 28 Jan 21 lecture ①

99.9% CI : (5.6, 31.5) mmHg

$\hat{\Delta}_n$ = 18.6 mmHg

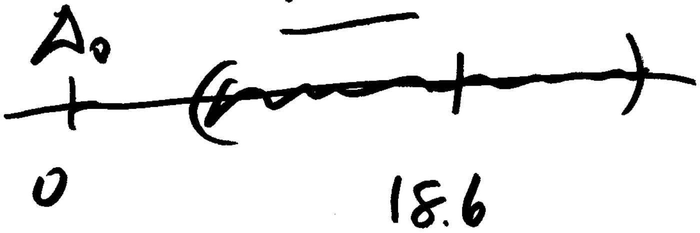
$\widehat{SE}(\hat{\Delta}_n)$ = 2.9 mmHg

Devil's advocate (0) nothing $\boxed{\text{null}}$ (hypothesis) (theory) : (captopril doesn't work on average in P^*)

$\Delta_0 = 0$ mmHg

(the diff. between $\hat{\Delta}_n = 18.6$ & $\Delta_0 = 0$ is due to unlucky random sampling) (this IS a logical possibility)

99.9% CI for Δ



CI approach to statsig judgments.

① build your 99.9% CI | ② see if null value (Δ_0 here) is in or not in the

~~the diff. between~~ $\hat{\Delta}_n$ & Δ_0 is not statsig / is statsig here, $\Delta_0 = 0$ is NOT in 99.9% CI, so:

---

the difference between $\hat{\Delta}_n = 18.6$ mmHg

(a) $\Delta_0 = 0$ mmHg (is) both stat sig

(b) pract sig

this is an inference

about $\underline{P}^*$; if we wish instead to draw

inferences about $\underline{P} = \{$ all British

hypertensive adults in mid-1970s$\}$,

we have to make the further

assumption that $P^* \doteq P$

similar in relevant ways

Neyman (hypothesis testing) & E. Pearson ($\doteq$ 1930)

Fisher (significance testing) ($\doteq$ 1930) there is no explicit alternative hyp.

null hypothesis (I) (devil's advocate)

$\Delta \leq 0$ null hyp. (II):

p^* Δ_0

$(\underline{\underline{\Delta = 0}})$ (Captopril provides no average improvement in p^*) (the diff. between $\hat{\Delta}_n$ & Δ_0 is due to unlucky random sampling)

alternative hypothesis to null (I) hyp

(I) $\Delta \neq 0 \, (\Delta_0)$

$\boxed{\text{(II)} \ \Delta > 0 \, (\Delta_0)}$

(III) $\Delta < 0 \, (\Delta_0)$

(Captopril does provide average improvement in p^*) (the diff. between $\hat{\Delta}_n$ & Δ_0 is probably real (not due to unlucky random sampling))

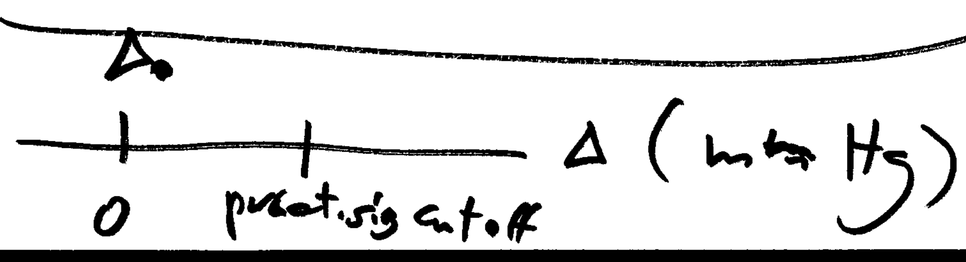
Fisher's logic

1. temporarily pretend null is true
2. construct a distance measure between (how the data came out: random) ② (how data should have come out if null true: fixed)

which one? in ②

3. if distance measure is big (how big?) reject null; if not big, conclude "insufficient evidence to reject null"

this is probabilistic proof by contradiction



what distance measure should we use?

(how data came out) vs. (how data should have come out if null true)

$\hat{\Delta}_n$ vs. 0

$$\frac{\hat{\Delta}_n - \Delta_0}{\widehat{SE}_{IID}(\hat{\Delta}_n)} = \frac{18.6\,\text{mmHg} - 0\,\text{mmHg}}{2.9\,\text{mmHg}} = +6.4$$

$$= \frac{\text{signal}}{\text{noise}} \quad (\text{pure \# unitless})$$

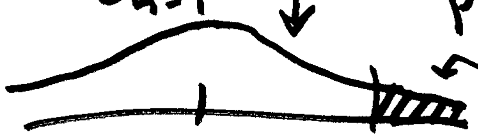
best estimate of Δ: $\hat{\Delta}_n = 18.6$ mmHg

we think Δ is around $\hat{\Delta}_n = 18.6$ mmHg (signal)

give or take about $\widehat{SE}_{IID}(\hat{\Delta}_n) \doteq 2.9$ mmHg (noise)

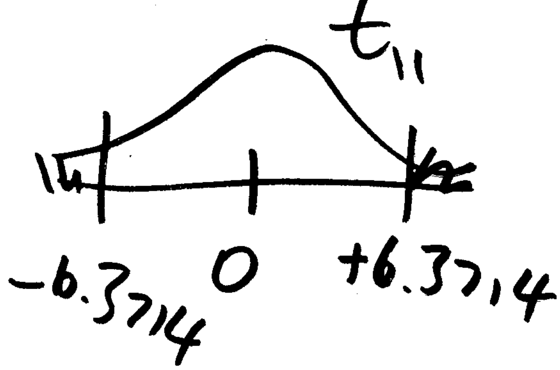
(if pop PDF is Normal) t test

t_{n-1} PDF of $\frac{\hat{\Delta}_n - \Delta_0}{\widehat{SE}_{IID}(\hat{\Delta}_n)}$ = t statistic if null true



0 6.4 p-value

p-value = chance, if null true, of getting data as extreme as, or more extreme than, what you got



if $\frac{\text{signal}}{\text{noise}}$ is big, reject null

$\Rightarrow$ p-value is small

how small? ⊛

2-tailed p-value = 2 · (1-tailed p-value) $\doteq$ 0.000053

alt: $\Delta \neq 0$ (2-sided t-test)

⊛ Fisher: $\hat{\Delta}_n$ is stat sig different from $\Delta_0 = 0$ $\Leftrightarrow$ iff p-value < 0.05