AMS 1
Lecture 10
7/11/16

L-161 This has been 1-sample CIs for continuous outcomes and 0/1 outcome.

CIs are 1 way to do statistical inference - the other is called <u>hypothesis tests</u> (Neyman + Pearson 1930) and significance tests (Fisher 1930)

ex. note: hyp/ sig are not as good as CIS

( 95% CI )
↑ 24.4° 25°C 25.6°

24.3° theory probably wrong

| | | |
|---|---|---|
| null hypothesis ($H_0$) | $\mu = 24.3°C$ ($\mu_0$) | Theory Correct $\boxed{1}$ |
| alternative hypothesis ($H_A$) | $\mu \neq 24.3°C$ | Theory Wrong $\boxed{2}$ |

2-sided alternative

$\boxed{1}$ The difference b/w $\mu_0 = 24.3°C$ and $\bar{y} = 25.0°C$ is due to unlucky random sampling (this <u>is</u> a logical possibility.

$\boxed{2}$ No, difference b/w 24.3 and 25.0 is real

<u>Neyman's Logic:</u> (+ Fisher's) try null on for size see if discrepancy b/w

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

is large, if yes, favor alt ("reject null") if not, favor null ("fail to reject null")

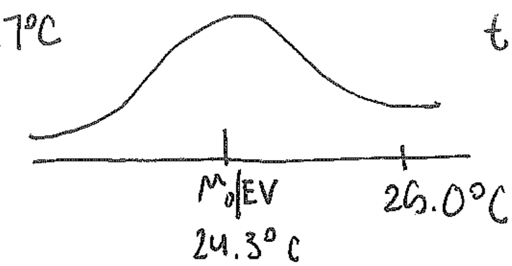
⊛(how data came out):
$\bar{y} = 26.0°$ < vs (⊛⊛) (how should have come out if null true)

$E_{IID}(\bar{y})$ (if null true) $= \mu_0 = 24.3°C$

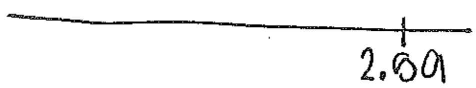
L-163 ex. Since in the example $\mu = 24.3°C = \mu_0$
EV of $\bar{y} = \mu_0 = 24.3°C$

[3] long run histogram of $\bar{y}$ accounting for uncertainty in $\sigma$, <u>if null were true</u>

$\hat{SE} = .27°C$ t curve




Std Units




$\frac{26.0 - 24.3}{.27}$ °C

$\cancel{z} = 2.59$

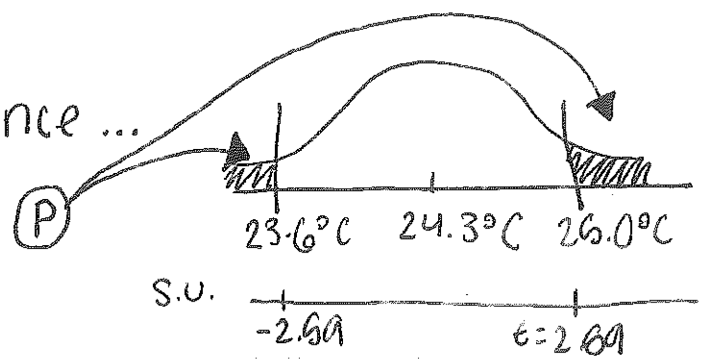
$t = 2.59$

Since this is a t-curve 2.59 seems surprizing...

Work out the chance, if the null true, of getting data as extreme as, or more extreme than what I got

numerical surprise measure = <u>P value</u>

P value = chance ...



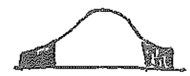

How do we measure discrepancy?

$$\frac{\bar{y}-\mu_0}{s/\sqrt{n}} \leftarrow SD_{\bar{y}} \text{ if null true} = \frac{26.0°C-24.3°C}{.27°C}$$

$\boxed{\text{t-test}}$

$$= \frac{.7°C}{.27°C} = \frac{\text{"signal"}}{\text{"noise"}} = +2.69 = t$$

"the t statistic"

| Alternatives to the Null | | |
|---|---|---|
| $\mu \neq \mu_0$<br>alt$_1$ | two-sided alternative<br>(2-tailed P value, 2 tailed test) |  |
| $\mu > \mu_0$<br>alt$_2$ | one-sided alternative<br>(right side) one tailed Pvalue/test |  |
| $\mu < \mu_0$<br>alt$_3$ | one-sided alternative<br>(left side) |  |

If P-value is small, favor alt, if p is big, favor null.

How "small" is small enough?

$\boxed{\begin{array}{l} \text{if } P \leq 5\% \text{ result is statistically significant} \\ \text{if } P \leq 1\% \text{ result is highly stat sig} \end{array}}$

Hyp testing w/ 2 sided alt, conclusion is identical to that of CI approach } MATH FACT

Frequent use of hypothesis testing:

- null: my theory wrong  alt: my theory right
- Want to reject null ⟷ I want a small p-value; some journals only accept paper if $p \leq 5\%$.

- Nonsense: People can get $P = 8\%$ 2 tailed but can make $P = 4\%$ pretending real alt. was 1-sided

(a) rigid adherence to $p \leq 5\%$ is silly
(b) convinced against null by 1-tailed $p = 4\%$, should be equally convinced by 2 tailed $p = 8\%$

If all you know is P value, can't tell if $\bar{y}$ comes out above or below $\mu_0$

Judge if: ① practical sig
② stat sig

L-172 $\mu = 0 \; \mu_0$  $\bar{y} \pm t^{.95}_{.99} \frac{s}{\sqrt{n}}$  $t = z$

95% CI:  $1.96 \frac{10}{\sqrt{1000}} = .32$
$-1 \pm 0.6$

is stat sig

☆ Stat sig but not prac sig = too much data

# Sample Size Determination
## CI approach

plan: $\bar{y} \pm t_{n-1}^{.95(2)} \frac{s}{\sqrt{n}}$

$95\%\ CI = 100(1-\alpha)\%$ $\quad (\overset{\text{alpha}}{\alpha} = .05)$

theory 1: null $\quad \mu = 32 = \mu_0$
2: alt $\quad \mu = 31.5 = \mu_A$

$$n = \frac{\left(t_{n-1}^{(1-\alpha)(2)}\right)^2 s^2}{(\mu_0 - \mu_A)^2}$$

Always round <u>up</u> with n (the sample size)

[Type I error = false rejection of null = $\alpha$
Type II error = false acceptance of null = $\beta$]

$$n = \frac{\left[t_{n-1}^{(1-\alpha)(?)} + t_{n-1}^{(1-\beta)(1)}\right]^2 \cdot s^2}{(\mu_0 - \mu_A)^2}$$

1 tailed = 1
2 tailed = 2

Convention:
$\alpha$: .05 .01
$\beta$: .1 .2

(See worked out problem L-181)

## Section 5: Two Sample Inf. Problems

Repeated Measures: same variable on n individuals at 2 different pts in time (longitudinal)

| Person# | B Befor | A After | A − B |
|---|---|---|---|
| | | | |