(1) Lecture # 11 — 12 Feb. 2009

To see whether a single sample $(y_1, \ldots y_n)$ from a pop. w/ a mean μ supports a theory that $\mu = \mu_0$ (theoretical value), build a 95% CI for μ using the t machinary & see if μ is in the interval; if not, data don't support theory @ 95% CI; if so, data do support theory at that level.

- When μ_0 is not in 95% interval, people say difference between theory μ_0 & data ($\bar{y}$) is Statistically significant (statsig); when μ_0 is not in 99% CI, the difference is highly statsig.

Q: Is difference between μ_0 & $\bar{y}$ practically significant (practsig)?

ex. Like asking whether a difference of 0.7°C ($\underset{\bar{y}}{25°} - \underset{\mu_0}{24.3°}$) is large in biological terms (from crab case study from lecture #10)

- Don't know - but q important
- Practical significance = the 1st question

Q: - Is a difference that is statsig real?

Not real: easy to attribute to unlucky random sampling (if in 95% CI?)

Real: hard to attribute to unlucky random sampling = statsig

Statsig (Statistically significant) ⟷ $\boxed{\text{Difference between } \bar{y} \text{ \& } \mu_0 \text{ hard to attribute to unlucky random sampling}}$ ⟷ diff. (probably) real

meaning of CIs — Neyman 1925

(2) Lecture #11 — 12 Feb, 2009

- Back to Crab Case Study:

  95% CI for μ was (24.5°C, 25.69°C

  Does this mean that

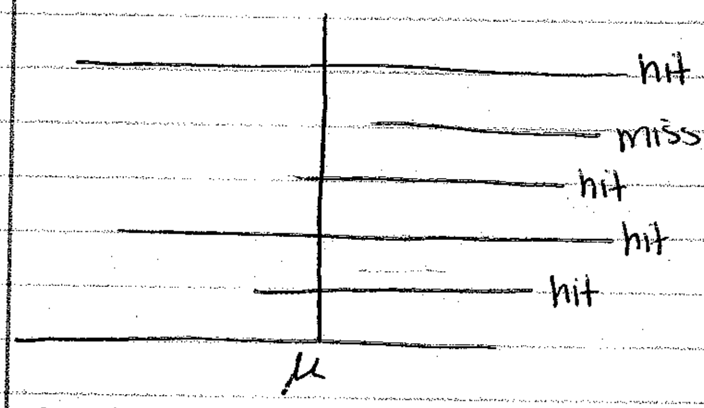
  $$P_F(24.5°C < \mu < 25.6°C) = 95\%$$

  $\uparrow$ relative frequency

  In other words, confidence = probability

  No! μ is a fixed unknown constant that's not changing across repititions

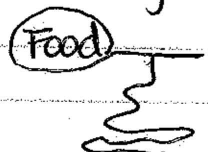
Q: OK, what does the interval mean?



⇐ about 95% of these should be hits

Can't tell if any single CI you build is a hit or miss, because we don't know μ; so our confidence in CIs is in the process of building the intervals, not in any single interval

Case Study Lab Rats



lab rats: 12 animals, 10 turn left to food

$$\begin{matrix} 1\,L \\ 0R \end{matrix} \begin{bmatrix} 1^s \\ 0^s \end{bmatrix} n=12$$

mean $\frac{10}{12} = 83\%$

Is the difference between $\frac{10}{12} = 83.\bar{3}\%$ 50% large?

- large in practical terms? → Huge
- large in statistical terms? ⟹

next pg

CO

pg3 of 5

12 Feb. 2009

(3) Lecture # 11

population

ask biologists → All similar lab rats (relevant)

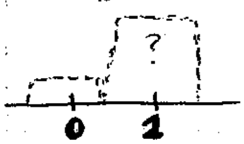
left?

$N = ?$ (big) $\begin{bmatrix} 1^s \\ 0^s \end{bmatrix}$

mean $\mu = p = ?$

SD $\sigma =$

pop. hist.



(actual) like SRS = IID →

hyp IID →

sample

the observed lab rats

y=1
n=0

left?

$\begin{bmatrix} 1^s \\ 0^s \end{bmatrix}$ n=12

mean $\bar{y} = \hat{p} = \frac{10}{12}$

"p hat" $= .83 = 83\%$

$\begin{bmatrix} \ \end{bmatrix}$ n=12

mean $\hat{p} = ?$

ex: $\frac{9}{12} = 75\%$

Imag. Data

possible $\hat{p}$s

$\hat{p}$s

$\begin{bmatrix} 83\% \\ 75\% \\ \vdots \end{bmatrix}$ $M = \infty$

<u>long run mean</u>:

expected value of $\hat{p} = p$

EV of $\hat{p} = E_{IID}(\hat{p}) =$

$E_{IID}(\bar{y}) = \mu = p$

$\boxed{E(\hat{p}) = p}$ math fact

CI: $y \pm t^{.95}_{n-1} \hat{SE}_{IID}(\bar{y})$

^continuous outcome

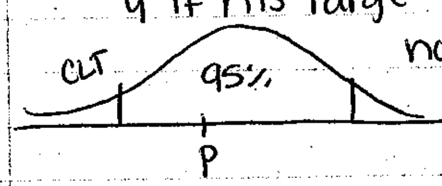
approximate : $\hat{p} \pm 1.96\, \hat{SE}_{IID}(\hat{p})$

0-1 outcomes $83\% \pm 2(11\%)$

truncate @ 100%

61% 83% 100% 105%

<u>long run histogram</u> of $\hat{p} =$ $\bar{y}$ if n is large



SU

-1.96 1.96

→ truncate

$p_0 = 5\%$ 61% 83% 100% 105%

<u>long run SD</u>:

est. standard error of $\hat{p} = 11\%$

$\hat{SE}$ of $\hat{p} = \hat{SE}_{IID}(\hat{p}) =$ (estimated)

$\hat{SE}_{IID}(\bar{y}) = \hat{\sigma}/\sqrt{n}$

earlier math fact: if the pop has only 2 values in it,

$\sigma = \begin{pmatrix} \text{larger} - \text{smaller} \\ \text{value} \quad\ \ \text{value} \end{pmatrix} \sqrt{\begin{pmatrix}\text{proportion of} \\ \text{larger values}\end{pmatrix}\begin{pmatrix}\text{proportion of} \\ \text{smaller values}\end{pmatrix}}$

So here; larger value = 1

smaller value = 0

∴ SD of a 0-1 pop is

$\boxed{\sigma = \sqrt{p(1-p)}}$ math fact with 0-1 pop.

So: $SE_{IID}(\hat{p}) = \frac{\sqrt{p(1-p)}}{\sqrt{n}} = \sqrt{\frac{p(1-p)}{n}}$

$= \sqrt{\frac{(.83)(.07)}{12}} = 0.108 = 10.8\% = 11\%$

★ is not in 95% CI, so the

★ difference between 83%

★ ($\hat{p}$) & 50% is stat sig. (is probably real)

(4), Lecture #11 12 Feb. 2009

Inferential Summary

pop.	Unknown quantity of interest	p = pop % of lab animals that would turn left (toward food)
sample	estimate of p	$\hat{p} = 83\%$
imag. data set	give or take for $\hat{p}$ as est of p	$\hat{SE}(\hat{p}) = 11\%$
imag. data set	95% CI for p	$\hat{p} \pm 1.96 \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (61\%, 100\%)$ truncated at 0% & 100%

CI^s are only 1 way to do statistical inference - the other is called hypothesis tests (Neyman & Pearson, 1930^s) & significance tests (Fisher 1930^s)

" Draper will argue hyp./sig. tests are not as good as CI^s but we need to know both.

① null hypothesis (H_0)	$\mu = \overset{\mu_0}{24.3°C}$	" theory correct
② alternative hypothesis (H_A)	$\mu \neq 24.3°C$	" theory wrong

① The difference between $\mu_0 = 24.3°C$ & $\bar{y} = 25.0°C$ is due to unlucky random sampling
(this is a logical possibility)

② No, the difference between 24.3 & 25° is <u>real</u>

Neyman's logic (& Fishers): try null on for size, see if discrepancy between

$*$ (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")

(5) Lecture #11

*(how data came out): $\bar{y} = 25.0°C$ $\boxed{vs}$

**(how data should have come out if null true

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

pop	sample	imaginary data
all relevant crabs	the obs. crabs	possible $\bar{y}$s

pop: temp

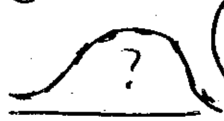
N=? (big) [|]

μ_0 = mean μ = 24.3°C

↳ if null true

SD = σ = ?

pop. hist



actual like IID →

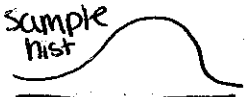
sample: temp

[{] n=25

mean $\bar{y}$ = 25

SD s = 1.34

sample hist



hyp IID →

[{] n=25

mean $\bar{y}$ = ?

ex: 24.5

imaginary data: $\bar{y}$s

[25.0
24.5
⋮]

long run mean

EV of $\bar{y}$ $\mu_0 = 24.3°C$

est. long run SD

SE of $\bar{y}$ $= \frac{s}{\sqrt{n}} = .27°C$

how measure discrepancy?

$\frac{\bar{y} - \mu_0}{\left(\text{SE of } \bar{y} \text{ if null is true}\right)} =$ $\quad \frac{s}{\sqrt{n}} \rightarrow$ (SE of $\bar{y}$ if null is true)

$\frac{25°C - 24.3°C}{0.27°C}$

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

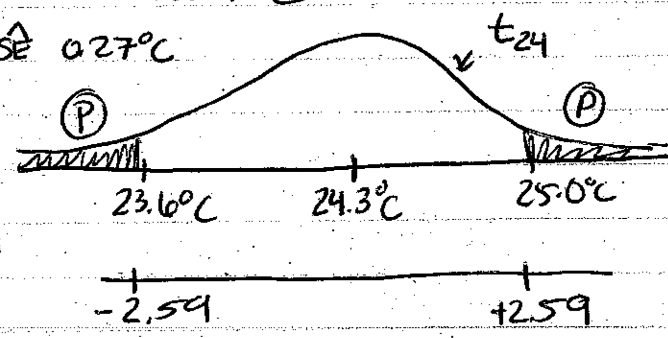
$+2.59 = t$ "the t statistic"

long run hist if null true

accounting for uncertainty in σ

$\frac{25.0° - 24.3°C}{0.27°C} = +2.59 = t$

$\hat{SE}$ 0.27°C



Chance if null true, of getting data as extreme as, or more extreme than, what I got = numerical surprise measure = P value

Continue on Lecture #12 17 Feb 2009