This time: point & interval estimation for means & proportions

Next time: sample size determination; testing

Read: LN pp. L 160 - L 173 Today: LN pp. L-137 →

- If your regular disc. sec. is on Friday, go to section (will help with the midterm), turn in midterm by midnight Fri. in box outside Baskin 357c.
- make a copy of your midterm, put your name on <u>every</u> page.

4: Statistical Inference

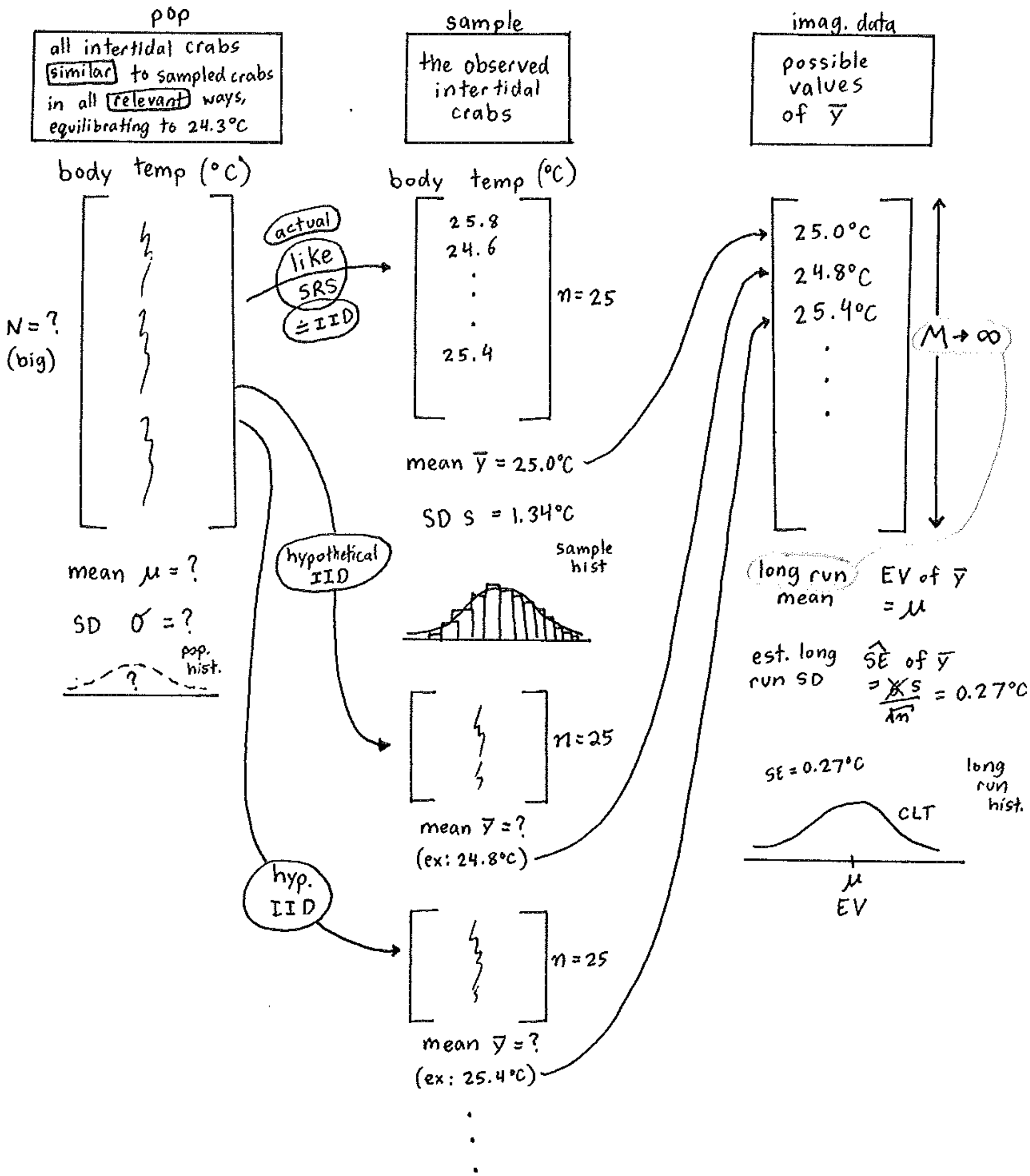
$\boxed{\text{Q1}}$ Is the difference between $\bar{y} = 25.0°C$ & theory temp $24.3°C$ large in <u>practical terms</u> (practically significant)?

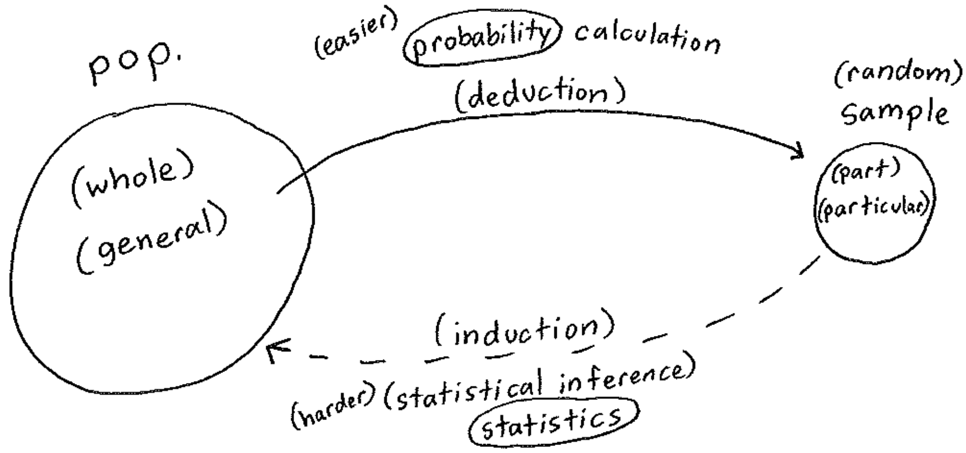
$\boxed{\text{A1}}$ Not clear on biological grounds, but:

$$\frac{25.0°C - 24.3°C}{24.3°C} = \frac{+0.7°C}{24.3°C} = \text{about a } 3\% \text{ increase}$$

hard to tell.

model on next page


pop
all intertidal crabs similar to sampled crabs in all relevant ways, equilibrating to 24.3°C
body temp (°C)
N = ? (big)
mean μ = ?
SD σ = ?
pop. hist.
?
actual
like
SRS
≐ IID
sample
the observed intertidal crabs
body temp (°C)
25.8
24.6
25.4
n = 25
mean ȳ = 25.0°C
SD s = 1.34°C
sample hist
hypothetical IID
n = 25
mean ȳ = ?
(ex: 24.8°C)
hyp. IID
n = 25
mean ȳ = ?
(ex: 25.4°C)
imag. data
possible values of ȳ
25.0°C
24.8°C
25.4°C
M → ∞
long run mean
EV of ȳ = μ
est. long run SD
SE^ of ȳ = σ s / √n = 0.27°C
SE = 0.27°C
long run hist.
CLT
μ
EV



4 rows
2 columns
art w/unknown

Inferential Summary

(pop.)	unknown (pop.) quantity of main interest	μ = mean body temp (°C) of all pop. crabs if equilibrated to 24.3°C
(sample)	(point) estimate	$\bar{y} = 25.0$°C
(imag. data)	give or take for $\bar{y}$ as an estimate of μ	$\widehat{SE}(\bar{y}) = \frac{s}{\sqrt{n}} \doteq 0.27$°C
(imag. data)	95% interval for μ	$\bar{y} \pm 2\,\widehat{SE}(\bar{y})$ $= (24.4, 25.5)$

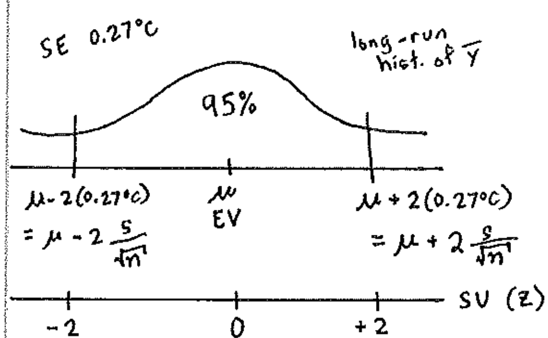
EV of $\bar{y}$ = $E_{IID}(\bar{y}) = \mu$

SE of $\bar{y}$ = $\boxed{SE_{IID}(\bar{y}) = \frac{\sigma}{\sqrt{n}}}$ most important formula

SE = noise/uncertainty

estimated SE of $\bar{y}$ $\boxed{\widehat{SE}_{IID}(\bar{y}) = \frac{s}{\sqrt{n}}} = \frac{1.34°C}{\sqrt{25}} = 0.27°C$

"hat"



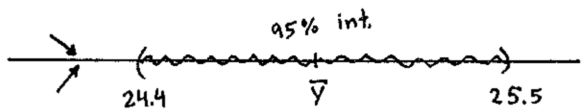
about 95% of the time, $\bar{y}$ will fall within 2 $\widehat{SE}$ of μ.

Therefore $\boxed{\bar{y} \pm 2\widehat{SE}(\bar{y})}$ is a good interval guess for μ (μ is (95%) highly likely to be in that interval).

Jerzy (Jerry) Neyman - great statistician: a 95% (confidence) interval for μ (1927)

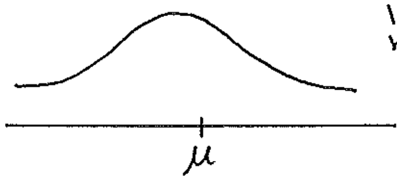
$25.0°C \pm \underbrace{2(0.27°C)}_{0.54°C}$ approx.

(24.4°C, 25.5°C)



Since theory value (24.3) is *not in* 95% int, data do not support theory.

$\widehat{SE}$ $0.27°C = \frac{s}{\sqrt{n}}$

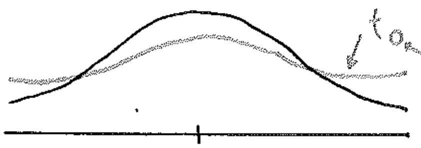


approx. longrun hist. of $\bar{y}$

we cheated by replacing σ with s

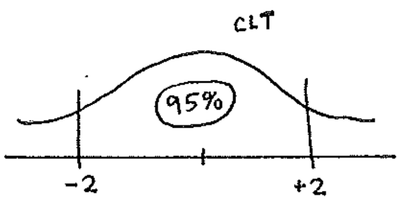
Correct Curve aka student's t curve

William Gossett (1908) brewer (& data analyst) at guiness in Dublin. "student" = fake name



t_0 ← degrees of freedom (d.f.)

with n obs. use $\boxed{t_{n-1} \text{(d.f.)}}$



actually: -1.96 +1.96

$25.0°C \pm 2\ \widehat{SE}$

$\bar{y} \pm \underbrace{2.064}_{t_{n-1}} \widehat{SE}$ (from chart)