this time: statistical inference:

next time: interval estimation

read: LN pp. L-37 ↓ L-162

AMS7 12 May 17

all official note-taker notes are now on course web page

new due date for midterm at canvas: Sunday night 11.59pm 14 May

new due date for homework 3: probably Mon 22 May

today: LN p. L-137 →

to define the pop. data set with as much valid generality as possible, answer this

Q: what is the broadest scope of valid generalizability outward from my data set?

pop = what sample would have looked like if you had done your experiment on a

pop. | stat. model | sample | imag.
--- | --- | --- | ---

all intertidal crabs (like sampled crabs in all relevant ways)

the observed intertidal crabs

(rep. sampling) all possible $\bar{y}$'s

body temp.

$N = ?$ (big)

?
?
?
?
?

like SRS (actual)

$\doteq$ IID

body temp.

$n = 25$

mean $\bar{y} = 25.0°C$

SD $s = 1.34°C$

$\bar{y}$'s

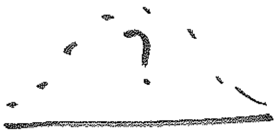
25.0°C

25.2

$M = \infty$

mean $\mu = ?$

SD $\sigma = ?$



pop. hist.

hyp. IID

SD 1.34

25.0°C

sample hist.

$n = 25$

mean $\bar{y} = ?$

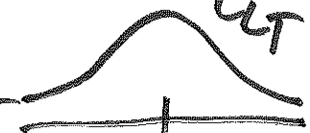
ex. 25.2°C

long run mean $\bar{y}$ = EV of $\bar{y}$ = $\mu$

est. long run SD $\bar{y}$ = $\hat{SE}$ of $\bar{y}$ = $\frac{s}{\sqrt{n}} \doteq 0.27°C$

$\hat{SE} = 0.27°C$

long run hist



CLT

$\mu$

$\hat{SE}$ 0.27°C

CLT

$\mu$

long run hist of $\bar{y}$, $\sigma$ unknown

not quite

SE $\frac{\sigma}{\sqrt{25}}$

CLT

$\mu$

long run hist of $\bar{y}$, $\sigma$ known

correct

William S. Gosset | Guinness Brewery (Dublin)

(1908)

lot more individuals than you did

N

n

(hypokalemia)

pop

whole

general

known

probability model

deduction

easier

sample

part

particular

unknown

random SRS

pop

unknown

statistical model

induction

= (statistical) inference

harder

sample

known

(intertidal crabs)

# inferential summary



| | |
|---|---|
| unknown pop. quantity of main interest | $\mu$ = pop. mean body temp. after equilibration |
| estimate of $\mu$ | $\bar{y} = 25.0°C$ |
| give or take for $\bar{y}$ as ~~est.~~ est. of $\mu$ | $\hat{SE}(\bar{y}) = 0.27°C$ |
| 95% interval for $\mu$ (CI) | $\bar{y} \pm 2.064 \hat{SE}(\bar{y}) = (24.5, 25.6)°C$ |

← sample ←
↑ ← ↓

math fact

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

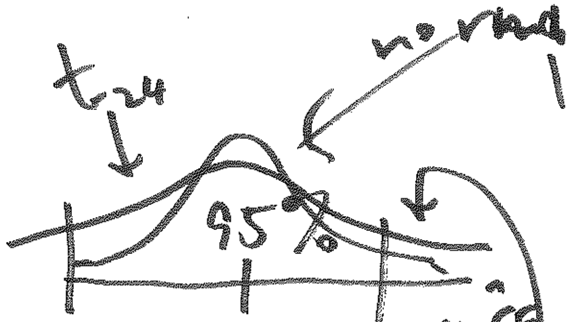
reader p. R-22

$$SE \text{ of } \bar{Y} = \boxed{SE_{IID}(\bar{y}) = \frac{\sigma}{\sqrt{n}}}$$

$$\text{estimated } SE \text{ of } \bar{y} = \hat{SE}_{IID}(\bar{y}) = \frac{s}{\sqrt{n}}$$

$$= \frac{1.34°C}{\sqrt{25}} = \frac{1.34°C}{5} \doteq 0.27°C$$

(estimate of) "SE hat"




long run hist of $\bar{T}$, accounting for uncertainty in $\sigma$

$\mu \pm 2.064\, SE$

L-14

$t$ curve is close to normal if $n$ large

24 with $(n-1)$ degrees of freedom

1.96 normal