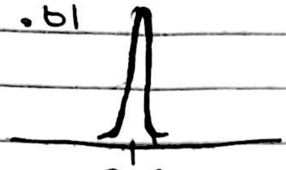
Stat 7 - class 13 11-07

As $n\uparrow$, $SE(\bar{y})\downarrow$ but only at a $\sqrt{n}$ rate: to cut the SE in half you need to quadruple the sample size

SE .01



3.8

hist of $\bar{y}$, $n=400$

similar to sampled crabs in all relevant ways

L-139 pop — all crabs of same species

body Temp

$N=?$ big

like SRS → IID

sample — the observed crabs

body temp

25.8
|
25.4

$n=25$

mean $\bar{y} = 25.0°C$

SD $s = 1.34°C$

sample hist.

imag data

25.0
24.8
:
:

$M \to \infty$

possible values of $\bar{y}$

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

long run SD = $SE_{\bar{y}}$

long run hist

L-143 pop μ

pop σ

pop hist

pop ⟷ broadest scope of valid generalizability outward from the data set

probability

pop general whole —deduction→ sample specific part

←induction

Statistical inference

each number in pop dataset is around $\mu \doteq \bar{y}$ give or take about $\sigma \doteq s$

inferential summarry

pop	unknown pop quantity of main interest	μ = pop mean of body temp at equilibrium
sample	estimate of μ	$\bar{y} = 25.0\,^\circ C$
imag data	give or take for $\bar{y}$ as est. of μ	$\hat{SE}(\bar{y}) \doteq 0.27\,^\circ C$
	95% CI for μ	

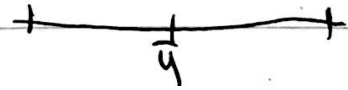
on basis of data, we think that μ is around $\bar{y} = 25.0\,^\circ C$, give or take about $\hat{SE}(\bar{y}) = 0.27\,^\circ C$

each measurement in sample data set is around $\bar{y} = 25.0\,^\circ C$ give or take about $s = 1.34\,^\circ C$

each number in the imaginary data set is ($\bar{y}$) around μ , give or take about $\frac{\sigma}{\sqrt{n}}$

estimated EV of $\bar{y} = \epsilon_{IID}(\bar{y}) = \frac{\sigma}{\sqrt{n}}$

how $\hat{SE}(\bar{y}) = \frac{s}{\sqrt{n}} = \frac{1.34}{\sqrt{23}} \doteq 0.27$

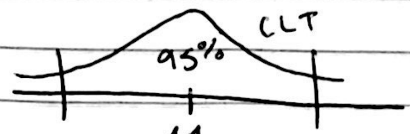


$\boxed{95\%}$ confidence level

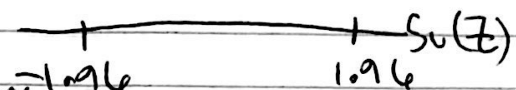
confidence interval for μ

long run histogram of $\bar{y}$

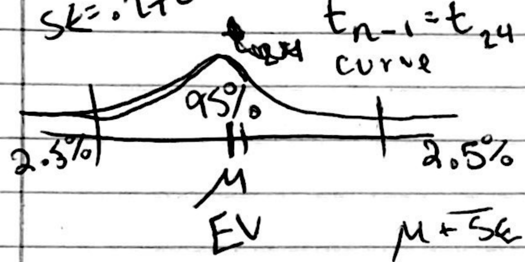
$\hat{SE} = 0.27°C$



if σ known



$SE = .27°C$



long-run hist of $\bar{y}$ accounting for uncertainty about σ

L-142

$$P_F(\mu - 2.064\hat{SE} < \bar{y} < \mu + 2.064\hat{SE}) = 95\%$$

$$P_F(\bar{y} - 2.064\hat{SE} < \mu < \bar{y} + 2.064\hat{SE}) = 95\%$$

95% CI for μ $\qquad \bar{y} \pm t \frac{s}{\sqrt{n}}$

sample mean $\pm$ (t #) (estimated SE)