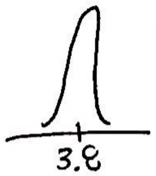
SQUARE ROOT LAW, INFERENTIAL SUMMARY, CONFIDENCE INTERVALS

11/7/19

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

SE .01



hist at $\bar{y}$
$n=400$

similar to sampled crabs in all relevant ways

L-139 pop: all crabs of same species

sample: the observed crabs body temp

imag. data

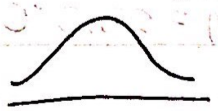
possible values of $\bar{y}$

pop		sample		imag. data
Body Temp, $N=?$ big	like SRS $\xrightarrow{}$ IID	25.8 … 25.4	$n=25$	25.0, 24.8, ⋮ ; M ∞

mean $\bar{y} = 25.0°C$
SD $s = 1.34°C$

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

L-143 pop μ
pop σ



sample hist

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

long run hist.

pop $\longleftrightarrow$ broadest scope of valid generalizability outward from the data set probability

population $\xrightarrow{\text{deduction}}$ sample

general whole $\xleftarrow{\text{induction}}$ specific part

statistical interference

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

inferential summary

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

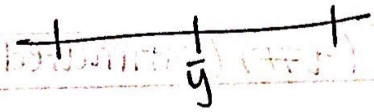
on basis of data, we think that μ is around $\bar{y} = 25.0°C$, give or take about $\hat{SE}(\bar{y}) = 0.27°C$

each measurement in ~~a~~ sample data set is around $\bar{y} = 25.0°C$ give or take about $s = 1.34°C$

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

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

how $\hat{SE}(\bar{y}) = \frac{s}{\sqrt{n}} = \frac{1.34}{\sqrt{25}} = 0.27$

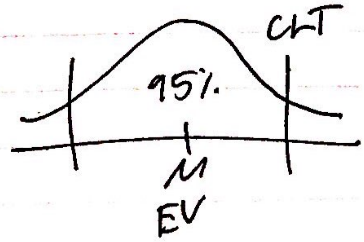


$\boxed{95\%}$ confidence level

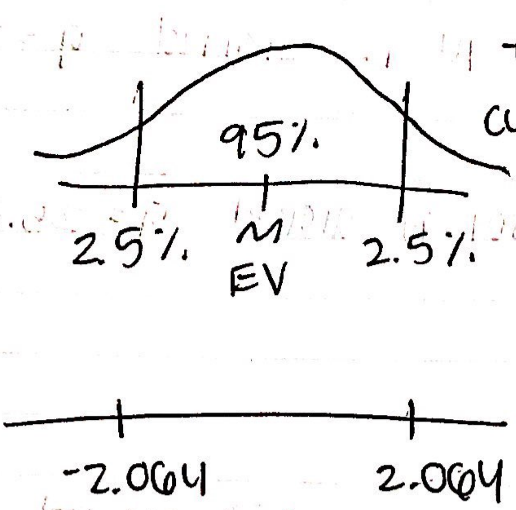
confidence interval for μ

long run histogram of $\bar{y}$

$\hat{SE} = 0.27°C$



if σ known



$t_{n-1} = t_{24}$ curve

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

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

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

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

sample mean ± (t#) (estimated SE)