(1) Lecture #10 Statistical Inference

hypokolemia (probability)

population		sample
[S]	$\xrightarrow{\text{IID}}$	[S]
mean μ: 3.8		mean $\bar{y}$ = ?

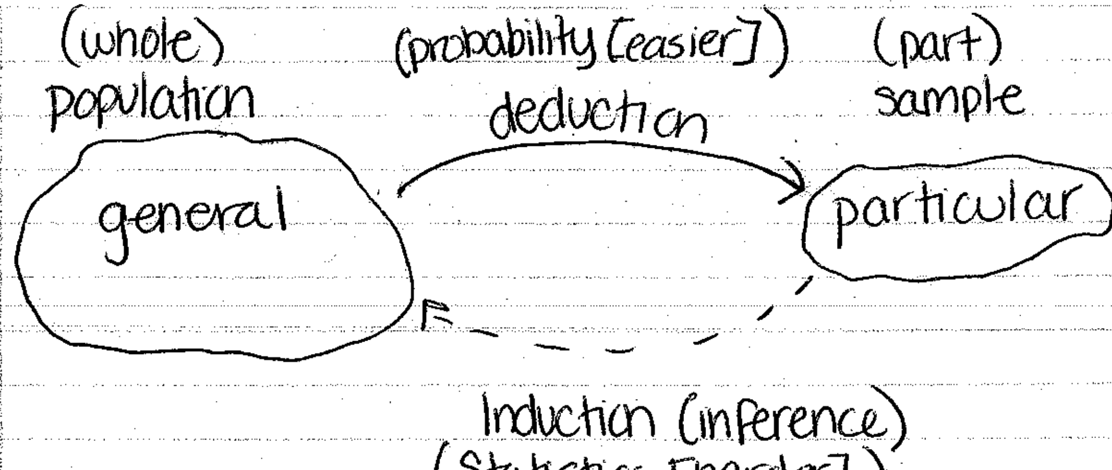
* We know the population story but don't know the sample. DEDUCTIVE REASONING (from the whole part)

Intertidal Crab Case Study

population	sample
	the observed crabs
	temp
[S]	[S] n=25
mean μ: ?	mean $\bar{y}$ = 25.0

* Here we know the sample & must go back to the population. INDUCTIVE REASONING

(inference & statistical)



10 Feb, 2009

(2) Lecture #10

Intertidal Crab Case Study

population	Sample	Imag. Data
All crabs similar to those sampled	the obs. crabs	possible $\bar{y}$s

Temp $\left[\; \{ \; \right] N = ?$ (big) $\xrightarrow[= 110 \text{ (actual)}]{\text{like SRS}}$ Temp $\left[\; \{ \; \right] n = 25$

$\bar{y}$: $\begin{bmatrix} 25.0 \\ 25.2 \\ \vdots \end{bmatrix}$ $M = \infty$

mean $\mu = ?$
SD $\sigma = ?$

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

pop. histogram ?

sample hist.

hypothetical 110 $\rightarrow [\;\;] n = 25$

mean $\bar{y}$?
(ex: 25.2)

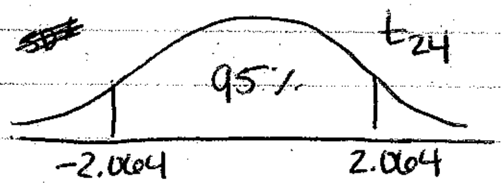
long run mean →
expected value of $\bar{y} = \mu$

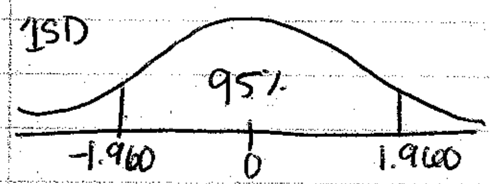
$EV \text{ of } \bar{y} = EV_{110}(\bar{y}) = \mu$

long run SD
standard error of $\bar{y}$

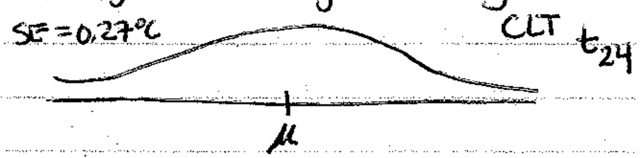
$SE \text{ of } \bar{y} = SE_{110}(\bar{y}) = \sigma/\sqrt{n}$

$\frac{\cancel{\sigma} s}{\sqrt{n}} = \widehat{SE}_{110}(\bar{y}) = \frac{1.34°C}{\sqrt{25}} = 0.27°C$

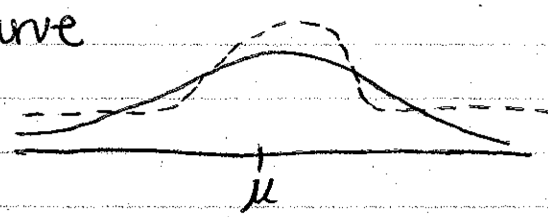




long run histogram of $\bar{y}$



1908 W.S. Gossett "student"
1915 ← guess → Fisher show correct + curve



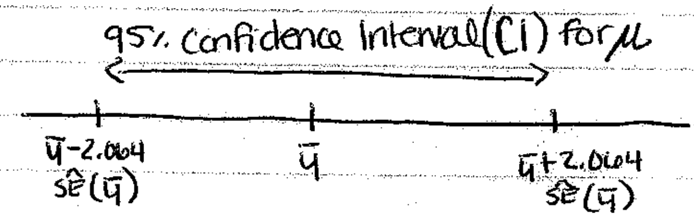
- more uncertainty, long run histogram of $\bar{y}$, accounting for uncertainty in σ
  (t machinary)

(3) Lecture #10 — 10 Feb, 2009

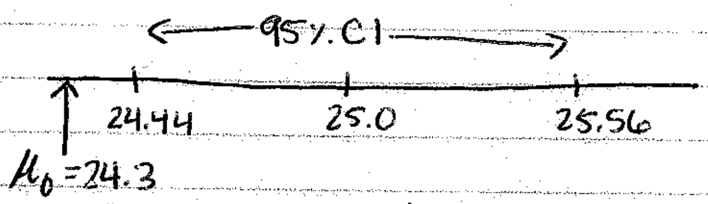
Inferential summary

population	unknown (pop) quantity of interest	μ = pop. mean temp. at which intertidal crabs of this species would equilibrate
sample	estimate of μ	$y = 25.0°C$
imaginary data	Give or take for $\bar{y}$ as est. of μ	$\hat{SE}(\bar{y}) = \frac{s}{\sqrt{n}} = 0.27°C$
	95% CI for μ	$y \pm (t_{n-1}^{0.95}) \frac{s}{\sqrt{n}} = (24.4, 25.6°C)$

late 1920s
Jerzy Neyman



$$\bar{y} \pm (t_{n-1}^{0.95}) \underbrace{\hat{SE}(\bar{y})}_{\frac{s}{\sqrt{n}}}$$



↑ theoretical value

- Look & see if μ_0 is within the 95% CI or not. If it is not there, the data does not support theory @ 95% confidence interval level; if it <u>is</u> in there, data do support it at that level.

- Jargon: The difference between $\mu_0 = 24.3°C$ & $\bar{y} = 25°C$ <u>is</u> statistically significant (large in statistical terms) because 24.3°C is not in 95% CI.