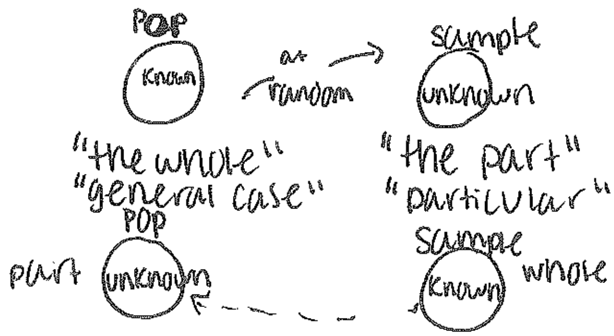
AMS7
Lecture 9
7/8/16

Crab case study continued:

Lets compare theoretical mean 24.3°C and experimental sample 25.0°C.



whole → part "probability" → Inductive Reasoning

sample → pop Deductive Reasoning

called "Statistical Inference"

Probability is easier than Statistical Inference.

Theory says $\mu_0 = 24.3°C$

Inferential Summary

population →	unknown pop summary of main interest	μ = pop mean equilibration T °C
sample →	estimate of μ (around)	$\bar{y} = 25.0°C$
imag data set	give or take $\bar{y}$ as estimate of μ	$\hat{SE}(\bar{y}) = .27°C$
	Confidence Interval for μ	24.4°C , 25.6°C

I think that μ is around $\bar{y} = 25.0°C$ give or take around $\hat{SE}(\bar{y}) = .27°C$, and my confidence 24.4°C to 25.6°C

Now let's build

imag data

$\boxed{\text{all possible } \bar{y}\text{s}}$

$$\begin{bmatrix} \bar{y} \\ \bar{y}_1 \\ \bar{y}_M \end{bmatrix} \quad M \to \infty$$

hypothetical IID

$$\begin{array}{l} \bar{y} = \\ \bar{y}_1 = \\ \dots \end{array} \begin{bmatrix} 25.0 \\ 24.8 \\ \vdots \end{bmatrix} \quad M \to \infty$$

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

long run SD (SE $\bar{y}$) $= \frac{\sigma}{\sqrt{n}} = \frac{s}{\sqrt{n}} = .27°C$

long run hist = SE = .27, μ

So lets fill in those blanks ↑

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

$SE_{IID}(\bar{y}) = \frac{\sigma}{\sqrt{n}}$

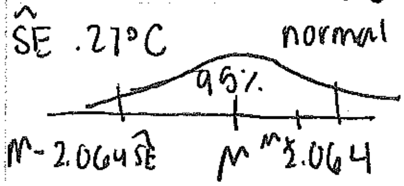
MATH FACTS

$\boxed{\text{CLT}}$

but lets get a number...

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

this is the "give or take" → $= 0.27°C$

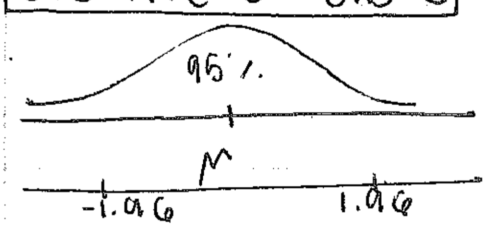


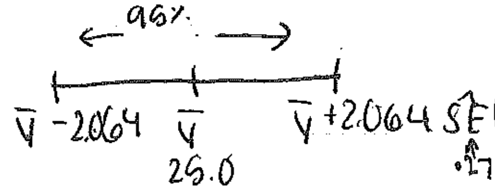
long run hist. $\bar{y}$

t curve

n-1 degrees of freedom as s as estimate of σ

use the t-table





$$\boxed{\bar{Y} \pm (t_{n-1,.95}) \frac{s}{\sqrt{n}}}$$

Neyman (1931)

interval of confidence
is a 95% confidence interval for μ

24.4°C, 25.6°C

"The 95% interval"

Theory value for μ was 24.3°C. Since 24.3°C is $\boxed{\text{not}}$ in the 95% interval for μ, we conclude that the data set does $\boxed{\text{not}}$ support theory at 95% confidence level.

When μ_0 is outside 95% interval, people say the difference b/w $\bar{Y}$ and μ_0 is statistically significant (statsig).

Significance: ① are 24.3°C + 25.0°C practically significantly different
- if big enough to matter practically
- more important, harder to judge

② are 24.3°C + 25.0°C stat sig different (μ_0 in 95% interval → not statsig)

Highly statsig → 99%

CI are about the process of building intervals, not exact numbers.

Case Study: Maze w/ Rats

1 = Left $n = 12$ rats

0 = Right   mean $10/12 = 83\%$

Is it stat sig? Lets check (build interval)

(See L-157 for data sets)

note: $\bar{y} = \hat{p} = 10/12$

SD of 0/1 (pop w/ only 2 values):

(larger value) - (smaller value) $\cdot \sqrt{\underset{p}{\text{prob large}} \cdot \underset{1-p}{\text{prob small}}}$

$\boxed{\text{With pop } 0/1 \quad \sigma = \sqrt{p(1-p)}}$ MATH FACT

Then build inferential summary table (see L-158)

	p = pop% that would turn left

imag data [ ]

EV $\hat{p}$?

SE $\hat{p}$?

hist ?

} L-159

New variables in this problem...

$$SE_{IID}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$$

Estimated SE use $p = \hat{p}$ → 11%

$\hat{p} = 83\%$ $\widehat{SE}(\hat{p}) = 11\%$ 95% interval:

$\hat{p} \pm 1.96\ \widehat{SE}(\hat{p})$

$83 \pm 2(11\%)$

61   83   105%

(61%, 105%) → 100%