STT 200 10-19-20

Note Title 10/19/2009

THE ONLY NEW ITEMS ARE #7 & #6.

PICK UP YOUR SHEET AT FRONT TABLE.

RECALL

95% CI for μ_x: $\bar{x} \pm 1.96 \frac{s_x}{\sqrt{n}}$

claim: P(CI COVERS μ_x) $\approx .95$

INSTEAD, WITHOUT REPL

CI $\bar{x} \pm 1.96 \frac{s_x}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$ FPC

Pop μ_x σ_x $\Rightarrow$ RANDOM $\Rightarrow$ $\bar{x}$ s_x n

WITH $\sqrt{s}$ WITHOUT

MAY NOT BE NORMAL DIST^{D}

$\bar{x} \sim \mu_x$

$s_x \sim \sigma_x$

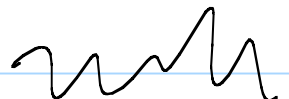
$\frac{s_x}{\sqrt{n}} \sim \frac{\sigma_x}{\sqrt{n}} = \sqrt{\bar{x}}$

T-CI $\bar{x} \pm T \sigma_x/\sqrt{n}$

~~Z~~

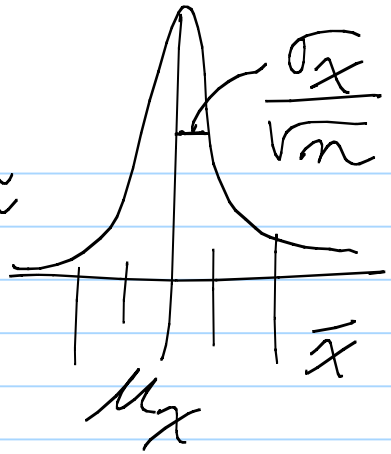
? Z-BASED 95% CI for $\mu_x - \mu_y$

ALSO



Pop μ_x σ_x

BUT ≈



---

7. IMPROVED CI FOR μ_y = MEAN 2010 TAX.

95% CI FOR μ_y IS (USUALLY) $\bar{y} \pm 1.96 \sigma_y/\sqrt{n}$ $y_1 \ldots y_n$

NOTE: CAN LOOK UP x = 2009 TAX. FREE INFORMATION $\bar{y}, \sigma_y$ n

WORKINGS OF METHOD:

"IMPROVED ESTIMATOR" $\bar{y} \ominus (\bar{x} - \mu_x)$ FIX!! POS??

$R \, \sigma_y/\sigma_x$ ← CHANGE

SLOPE OF SIMPLE REGRESSION

IF $\bar{x} > \mu_x$ (KNOWN) THEN IF $R > 0$
, IT SUGGESTS $\bar{y}$ TOO BIG ($\bar{y} > \mu_y$) ALSO.
AND THIS $*$ IS MAKING A DOWNWARD
CORRECTION TO $\bar{y}$ IN THIS CASE.

MECHANICS AND INTUITION ASIDE PERFORMANCE
CLAIM IS YOU ARE ENTITLED TO 95% CI

$$\left(\bar{y} - (\bar{x} - \mu_x) R \frac{\sigma_y}{\sigma_x}\right) \pm 1.96 \frac{\sigma_y}{\sqrt{n}} \sqrt{1 - R^2}$$

NEW ESTIMATION

$\sqrt{1-R^2}$ → 0 IF $R = 1$

IF $\sqrt{1-R^2} = \frac{1}{2}(SAY)$ $\Rightarrow$ $1-R^2 = \frac{1}{4}$

$R = \sqrt{3/4}$

TO GET SAME $\frac{1}{2}$ HAVE
TO TAKE SAMPLE SIZE 4n COSTLY.!!

DATA OF #2.

OUR DATA IS $(x_1, y_1) \cdots (x_{300}, y_{300})$ ASSUME WITH REPL.

FINDING $\bar{x} = 22.43$ $\bar{y} = 26.5$

$\mu_x = 22$ $s_x = 4.3$ $s_y = 5.$

COULD USE BUT WON'T $[[\sigma_x = 4.5]]$ $R(\text{SAMPLE}) = .77$

"IMPROVED ESTIMATOR" $\bar{y} - (\bar{x} - \mu_x) R \frac{\sigma_y}{\sigma_x}$

$= 26.5 - (22.43 - 22) .77 \frac{5}{4.3}$

$= 26.11$ YOU CHECK IT

VS ORDINARY $\bar{y} = 26.5$ (IGNORES INFO IN x)

ASSOCIATED 95% CI FOR μ_y IS

$(\text{IMPR EST OF } \mu_y) \pm 1.96 \frac{\sigma_y}{\sqrt{n}} \sqrt{1 - R^2}$

$26.11 \pm 1.96 \frac{5}{\sqrt{300}} \sqrt{1 - .77^2}$ GOT TO ENCLOSE THE VALUE 26.11

IMPR ESD

#6. IN TEXT READINGS — SEE SYLLABUS.

CI FOR p

SCORE $X = \begin{cases} 1 & \text{DEM} \\ 0 & \text{REP.} \end{cases}$ $\begin{cases} 1 & \text{HAD TREATMENT} \\ 0 & \text{NOT} \end{cases}$

0-1 SCORING

$\Rightarrow \bar{X} =$ SAMPLE PROP^N ~~$\frac{1}{n}$~~ INSTEAD $\sqrt{\frac{\sum (X-\bar{X})^2}{n}} = \sqrt{\bar{X}(1-\bar{X})}$

SO, IN CONTEXT OF 0-1 SCORES:

$\bar{X} =$ SAMPLE PROP^N OF $1^{\text{S}} = \hat{p}$ SAMPLE PROP^N

~~$\frac{1}{n}$~~ RATHER n-DIVISOR VERSION WHICH IS $\sqrt{\hat{p}\hat{q}}$

SO IF 45 OF 100 (MALE IN SAMPLE)

$\bar{X} = 45/100 = \hat{p} = .45$ ~~$\frac{1}{n}$~~ RATHER $\sqrt{.45 \cdot .55}$

$\bar{x} \pm 1.96 \frac{\sigma_x}{\sqrt{n}}$ USE $\bar{x} \pm 1.96 \frac{\sqrt{\bar{x}(1-\bar{x})}}{\sqrt{n}}$

OR (SAME) $\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}\hat{q}}}{\sqrt{n}}$ eg $.45 \pm 1.96 \frac{\sqrt{.45 \cdot .55}}{\sqrt{n}}$

IGNORE IF YOU WISH

THIS CI IS JUST $\bar{x} \pm 1.96 \frac{s_x}{\sqrt{n-1}}$

TRY IT 1 ODD (LAST DIGIT STUDENT #) COUNT 18

0 EVEN

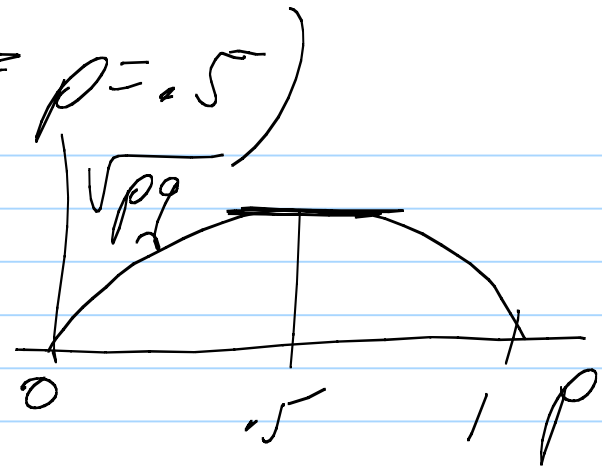
$\hat{p} = 18/52 \approx 1/3$ 52 TOTAL

95% CI FOR p (PRIVATELY BELIEVE $p = .5$)

$$\frac{1}{3} \pm 1.96 \frac{\sqrt{\frac{1}{3}\frac{2}{3}}}{\sqrt{52}}$$

$$\approx \frac{1}{3} \pm 1.96 \left(\frac{1}{2}\right) \frac{1}{\sqrt{52}}$$

$$\frac{1}{3} \pm \frac{1}{7}$$



CLOSE TO COVERING .5 WHICH WE BELIEVE IS TRUE p.

ANOTHER CI FOR p = FRACTION APPROVING BALLOT PROPOSAL.

SAMPLE OF $n = 200$ VOTERS FINDS $116/200$

95% CB FOR POPULATION FRACTION:

$$\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}\hat{q}}}{\sqrt{n}}$$

$$\frac{116}{200} \pm 1.96 \frac{\sqrt{.58 \cdot .42}}{\sqrt{200}} \quad \leftarrow \text{NUM}$$

$$.58 \pm \left(\frac{1}{14}\right)_{.07}$$

$.58 \pm .07$

$.51$ TO $.65$

ETC WORDS ~