Note Title 10/28/2009

# STT 200 10-28-09

TODAY ① WALK THRU CI $\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

MOE

POPULATION = THOSE PRESENT.

$N = 30$. (PRESENT)

RANDOM SAMPLE FOR PURPOSE OF EST'G $p$ = FRACTION OF POP'N WHO ARE JUNIOR OR SENIOR.

[ $p = \frac{20}{30} = 2/3$ NOT USUALLY SEEN. ]

DRAW RANDOM SAMPLE OF POPULATION WITHOUT REPLACEMENT SAMPLE OF $n = 11$

SAMPLE CONSISTS OF THOSE HAVING LAST DIGIT OF STUDENT # = 0 or 1 or 2.

1. HAVE NOW 11 PEOPLE WHO REPRESENT A WITHOUT REPL EQ PROB RANDOM SAMPLE OF THE POPULATION OF $N=30$. $n=11$

2. WE FIND 4 OF 11 ARE JUNIOR OR SENIOR. SO PARAMETER $p = 20/30$ (POPN) IS IN THIS CASE ESTIMATED FROM THE SAMPLE AS $\hat{p} = 4/11$. (CHANCE $\hat{p}$ (RANDOM) $= p$ SLIGHT)

3. A 95% CI PRESCRIPTION $\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}$

THIS EVALUATES TO $4/11 \pm 1.96 \sqrt{\frac{\frac{4}{11}\frac{7}{11}}{11}}$ EITHER THIS COVERS $p$ OR NOT

4. $P\left(p \text{ IN } \hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\right) \approx .95$ AS $n \to \infty$ IN WITH-REPL SAMPLING.

5. SINCE WE SAMPLED WITHOUT-REPL THE FPC SHOULD BE USED:

$$\hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$4/11 \pm \underbrace{1.96 \sqrt{\frac{4/11 \cdot 7/11}{11}} \sqrt{\frac{30-11}{30-1}}}_{MOE}$$

6. CLAIM: $P\left(p \text{ IN } \hat{p} \pm 1.96 \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right) \sim .95$

(REQUIRE $N \sim \infty$ $n \sim \infty$ - SUBTLE)

---

NOTE: HAVE TO PAY ATTENTION TO ~~THE~~ ASSUMPTIONS & THE NATURE OF SAMPLING.

(2) PER EXAM 3 (YOU'LL SEE IN PREP I WILL POST)

SURE — "MECHANICAL ACTS"

ALSO — ASSUMPTIONS. RANDOM SAMPLE

$n \to \infty$

(3) CURRENT MATERIAL WILL INCLUDE STRATIFIED SAMPLING. (POST-STRATIFIED SAMPLING).

IDEA.!! PERHAPS POP$^N$ HAS ~40% W/ POST SECONDARY ED

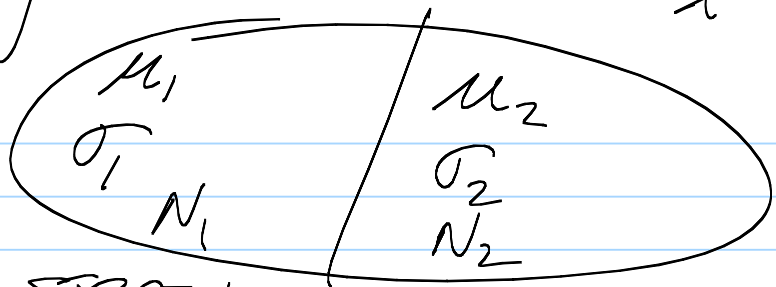
~~BUT RANDOM SAMPLE HAS ~10%~~ "

THEN WE SUSPECT FOR SCORE

$x$ = INCOME THAT $\overline{X}$ IS BIASED DOWN, ???

Pop N

$x$ = INCOME




$\rightarrow$ SAMPLE $n$ FROM $N = N_1 + N_2$

STRATUM 1 | STRATUM 2 $\quad K = \#$ STRATA $= 2$

GRAD HS OR LESS | POST-SECONDARY

(SUB POPULATION 1) (SUBPOPULATION 2)

[ USUALLY JUST SAMPLE $n$ (WITH REPL) FROM ENTIRE POP (IGNORING STRATUM) $P(\mu_x \text{ IN } \bar{x} \pm 1.96\ \sigma_x/\sqrt{n}) \sim .95$ ]

NOTE $\bar{x} = \frac{n_1}{n}\bar{x}_1 + \frac{n_2}{n}\bar{x}_2$ (BECAUSE $\cancel{\frac{n_1}{n}}\frac{x_{11}+\dots+x_{n_1 1}}{\cancel{n_1}} + \cancel{\frac{n_2}{n}}\frac{x_{21}\dots x_{2n_2}}{\cancel{n_2}}$

# OF ≤ HS IN SAMPLE

AVG INCOME OF ≤ HS GROUP IN SAMPLE

BUT COMPARE W/ $\mu_x = \frac{N_1}{N}\mu_1 + \frac{N_2}{N}\mu_2$

OVER WHOLE POPN

IF YOU KNOW $\frac{N_1}{N}$ THIS SUGGESTS USING

(INSTEAD OF $\bar{x}$) $\bar{x}^*$ DEFINED $= \frac{N_1}{N}\bar{x}_1 + \frac{N_2}{N}\bar{x}_2$

USING $\bar{x}$ IS LIKE USING $\frac{n_1}{n}$ $\frac{n_2}{n}$

SO UPSHOT IS ANOTHER 95% CI FOR $\mu_x$

$$\bar{x}^* \pm 1.96\sqrt{\sum_{i=1}^{2}\left(\frac{N_i}{N}\right)^2 \frac{\sigma_i^2}{n_i}}$$

$w_i = N_i/N$

NOTE $\overline{x^*} = \sum_{i=1}^{K} w_i \bar{x}_i$

$\bar{x}_i$ = AVG INCOME FOR SAMPLE INDIVIDUALS IN STRATUM $i = 1, 2$

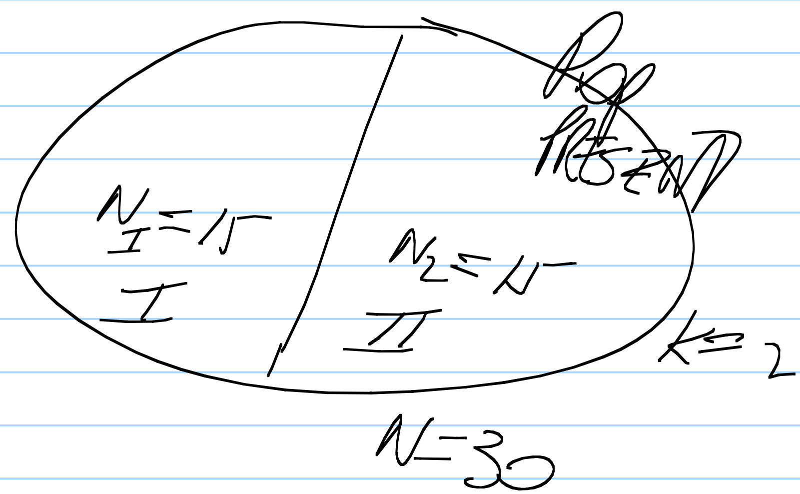
$s_i$ = SAMPLE SD OF INCOMES OF SAMPLE INDIVIDUALS IN STRATUM $i$

EXAMPLE

$x$ INCOME = MONEY ON HAND

$n = 11$. STRATUM 1: WOMEN

STRATUM 2: MEN




| | | | | | |
|---|---|---|---|---|---|
| \$ $x$ INDEX 1 | 100 | 0 | | | |
| \$ $x$ INDEX 2 | 30 | 51.27 | 15 | 20 | |

WOMEN

STRATUM 1

$$s_1 = \sqrt{\frac{\sum^2 (x_{1i}-\bar{x}_1)^2}{2-1}} = \sqrt{\frac{(100-50)^2+(0-50)^2}{2-1}} = 1.4(50)$$

$$\bar{x}_1 = \frac{100+0}{2} = 50 \qquad n_1 = 2$$

$$s_2 = \sqrt{\frac{(30-29)^2 + \sim + (20-29)^2}{4-1}}$$

MEN

$$\bar{x}_2 = \frac{30+5/.\,27+15+20}{4} \sim 29$$

$$\cancel{\bar{x} \pm 1.96\, s_x/\sqrt{n}} \qquad \frac{N_1}{N} = \frac{15}{30}$$

$$\frac{N_1}{N}\bar{x}_1 + \frac{N_2}{N}\bar{x}_2 = \frac{1}{2}\bar{x}_1 + \frac{1}{2}\bar{x}_2 = \frac{1}{2}50 + \frac{1}{2}29$$

POST STRATIFIED ESTIMATOR

POST STRATIFIED ESTIMATE

EST'D MOE FOR POST-STRAT EST IS

$$1.96\sqrt{\frac{(1.4\,50)^2}{2} + \frac{???}{4}}$$

$s_1^2 = 2(50)^2$

CHECK IT ABOVE

$n_1 = 2$

SAMPLE WOMEN

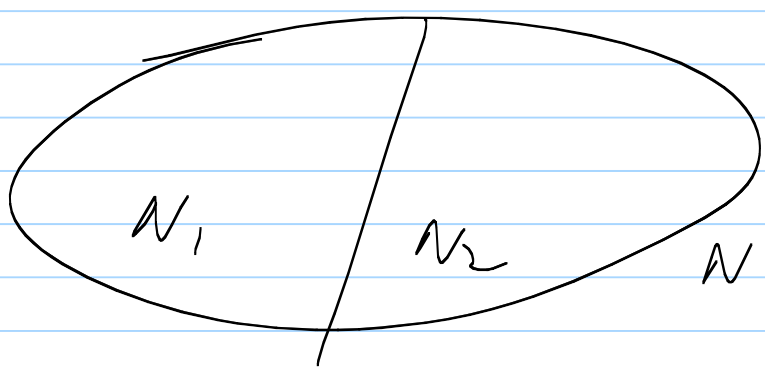
CAUTION!

$n_1 = 2$ $n_2 = 4$

FAR TOO SMALL!!

THIS WAS JUST A WALK-THRU OF CALCULATIONS

WHAT THEN IS STRATIFICATION?




$N = N_1 + N_2$

SAMPLE $\frac{n}{2}$ RANDOM FROM WOMEN (BECAUSE

$$\bar{x} \pm 1.96 \sqrt{\sum_{i=1}^{K=2} \left(\frac{N_i}{N}\right)^2 s_i^2 / n_i}$$

WORKS OUT

SAME AS $\frac{N_1}{N_2}\bar{x}_1 + \frac{N_2}{N}\bar{x}_2$ SINCE $\frac{n_1}{n} = \frac{N_1}{N}$