Note Title 3/22/2010

STT200 3pm 3-22-10

REINFORCE CI $\hat{p} \pm z \sqrt{\hat{p}\hat{q}}/\sqrt{n}$ $np \geq 10$, $nq \geq 10$

lot of pop

NO UNIV. RECIPE FOR n (LARGE) $\rightarrow$ CI $\bar{x} \pm z\, s/\sqrt{n}$

$$s = \sqrt{\frac{(x_1-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n-1}}$$

CI (all n) $\bar{x} \pm t\, s/\sqrt{n}$ DF = n-1

n FROM REDUCTION IN CI WIDTH DUE TO n

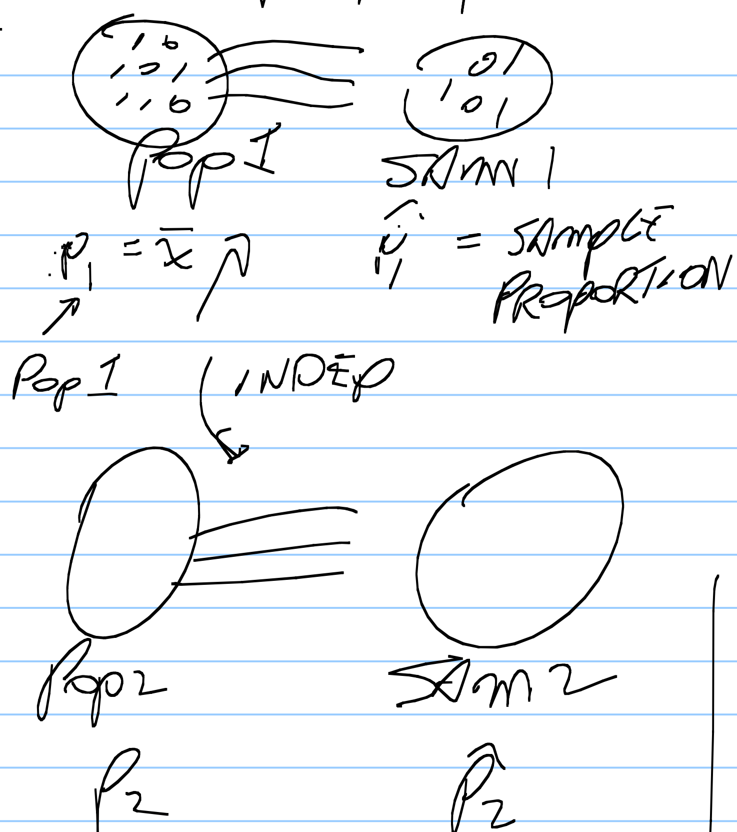
Ch 24 CI + TESTS FOR DIFF OF TWO POPULATION MEANS (INCL $\hat{p}$ CASE WHICH IS $\bar{x}$ FOR SCORES $\begin{cases} 1 & \text{HAVE CHARACTERISTIC} \\ 0 & \text{NOT} \end{cases}$

BEFORE BEGINNING RECALL X, Y INDEP. $\Rightarrow$ Var(X $\pm$ Y) = Var X + Var Y

CI for $p_1 - p_2$

MSU U.G.

J.M. U.G.



$$\hat{p}_1 \pm z \sqrt{\hat{p}_1 \hat{q}_1} / \sqrt{n_1} \qquad \text{CI FOR } p_1$$

AROUND $(.95)^2$ CHANCE BOTH 95% CI COVER THEIR RESPECTIVE p_i $i = 1, 2$

$$\hat{p}_2 \pm z \sqrt{\hat{p}_2 \hat{q}_2} / \sqrt{n_2} \qquad \text{CI FOR } p_2$$

OBJECTIVE IS CI FOR $p_1 - p_2$

$$\underbrace{\hat{p}_1 - \hat{p}_2}_{\text{R. VARIABLE}} \pm z \sqrt{\hat{p}_1 \hat{q}_1 / n_1 + \hat{p}_2 \hat{q}_2 / n_2} \quad \text{— CI}$$

WHOSE SD IS $\sqrt{p_1 q_1 / n_1 + p_2 q_2 / n_2}$ DON'T KNOW THIS

1.

$N_1 = 5{,}000$

TRACT 1

$n_1 = 60$

(LARGE)

FIND $\bar{x}_1 = 99.3$

$s_1 = \cancel{28.4}$ 0.3 PERHAPS

WITH-REPL? WITHOUT REPL?

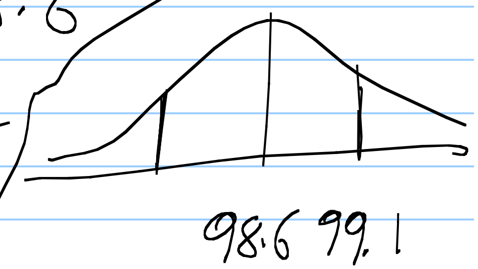
$N_2 = 16{,}000$

TRACT 2

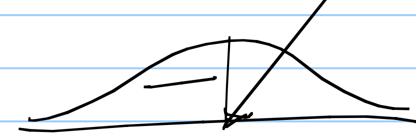
$n_2 = 100$

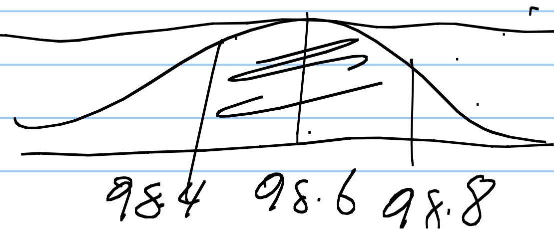
FIND. $\bar{x}_2 = 98.6$

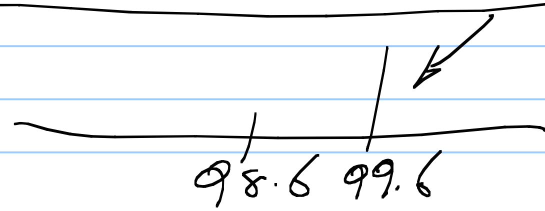
$s_2 = .2$



IF POP2 IS NORMAL 98.6







W/ ABOVE CI FOR $\mu_1 - \mu_2$ $= (99.3 - 98.6) \pm z \sqrt{\underbrace{\frac{(.3)^2}{60} + \frac{(.2)^2}{100}}_{\text{VARIANCE}}}$

[ ] TIGHTER IS BEST.

[ ]

[ ]

POINT EST

$\sqrt{\frac{.3^2}{100} + \frac{.2^2}{60}}$

$\sqrt{\frac{2000}{100} + \frac{50}{10}}$ vs $\sqrt{\frac{2000}{10} + \frac{50}{100}}$

SMALLER LARGER.

RULE IS - GIVEN A CHOICE WE'D WANT THE LARGER SAMPLE TO GO ON THE POPULATION HAVING THE LARGER VARIANCE - (ACTUALLY BETWEEN IS BEST -)

JUST DID #1C.

1d. RE-DO IF SAMPLES WERE WITH-OUT REPL.

$$(99.3-98.6) \pm z \sqrt{\frac{.3^2}{60}\frac{5000-60}{5000-1} + \frac{.2^2}{100}\frac{16000-100}{16000-1}}$$

$$\boxed{\frac{\sigma}{\sqrt{n}} FPC = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$$

1e. WAS THE ABOVE CI COVERED $\mu_1 - \mu_2$?
WHAT ARE THEY ? LET'S SUPPOSE $\mu_1 = \mu_2 = 98.6$.
THEN $\mu_1 - \mu_2 = 0$. CHECK IT!

CI from 1c

CI from 1d CHECK

1f. If $\mu_1 = \mu_2$ (SAME) WHAT IS PROB
95% CI for $\mu_1 - \mu_2$ WILL COVER 0 ?
CHANCE 95% CI for $\mu_1 - \mu_2$ COVERS $\mu_1 - \mu_2$ IS ?

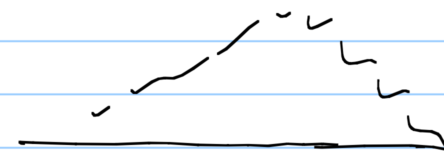
1g.- THERE IS A t ANALOGUE OF CI FOR $\mu_1 - \mu_2$.

BIG NEWS FOR

eg COMPARING MED w PLACEBO

MED1 MED2 etc

APPLICABLE WHEN Pop 1 AND Pop 2 ARE EACH NORMAL.



$$\bar{x}_1 - \bar{x}_2 \pm t_{DF} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

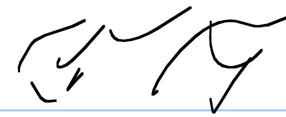
NO FPC

SAME AS z-BASED CI

$$\left.\begin{array}{l} DF1 = n_1 - 1 \\ DF2 = n_2 - 1 \end{array}\right\} ??$$

I WILL SUPPLY THE APPLICABLE DF.

2. 0-1 DATA SETUP



ANT COLONY

TYPE 1 44 of 120 SAMPLE TYPE 1

TYPE 2 38 of 160 " " 2.

SOIL TYPE 1 SOIL TYPE 2

TWO POPULATIONS

SCORE $x = $ 1 LATE DEVELOPMENT COLONY

0 NOT

95% z - CI FOR $p_1 - p_2$

$$\hat{p}_1 = \frac{44}{120} \quad \hat{p}_2 = \frac{38}{160}$$

$$\hat{p}_1 - \hat{p}_2 \pm z \sqrt{\hat{p}_1 \hat{q}_1 / n_1 + \hat{p}_2 \hat{q}_2 / n_2}$$

$$\frac{44}{120} - \frac{38}{160} \pm 1.96 \sqrt{\frac{\frac{44}{120}\frac{76}{120}}{120} + \frac{\frac{38}{160}\frac{122}{160}}{160}}$$

ANALOGUE OF $\sqrt{\hat{p}\hat{q}}/\sqrt{n} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$