ESTIMATING POP. PARAMETERS
POINT EST, MARGIN OF ERROR,
CONFIDENCE INTERVALS.

EXAM #2 TUESDAY

5.4, 6.1-4, 7.4-6, 8.3-6 ~~(7)~~

NORMAL PROB DISTR

~ APPROX. BINOM. R.V.

SAMPLE DISTRIBUTIONS : $\bar{x}$, $\hat{p}$

(CENTRAL LIMIT THM)

$\bar{x} \sim \mu$

$\hat{p} \sim p$

$\bar{x}_1 - \bar{x}_2 \sim \mu_1 - \mu_2$

HYPER GEOMETRIC: 5 GOATS 6 SHEEP, PICK 3 ANIMALS.

FIND PROB 2 GOATS, 1 SHEEP.

$$\frac{C_2^5 \, C_1^6}{C_3^{11}}$$

~ 20 QUESTIONS, FULL PERIOD 2 HOURS 10:30-12:30 PM

OFFICE HOURS MON / WED 1:30-3:30 PM.

PROJECTS.

§8.6 ESTIMATING THE DIFFERENCE BETWEEN POPULATION MEANS.

- BANANAS LEFT ON COUNTER LAST 3 DAYS LONGER THAN BANANAS STORED IN REFRIDGERATOR.

- PLANTS GIVEN FERTILIZER GROWS 5 INCHES TALLER THAN PLANTS NOT GIVEN FERTILIZER.

- PATIENTS THAT RECEIVE MEDICATION X SURVIVE 10 MOS. LONGER THAN PATIENTS THAT DO NOT.

	POPULATION 1	POPULATION 2
MEAN	μ_1	μ_2
STAND. DEV.	σ_1	σ_2
SIZE	N_1	N_2

POPULATION PARAMETERS

$\downarrow$ SELECT A RANDOM SAMPLE. $\downarrow$

	SAMPLE 1	SAMPLE 2
MEAN	$\bar{x}_1$	$\bar{x}_2$
STAND. DEV.	s_1	s_2
SIZE	n_1	n_2

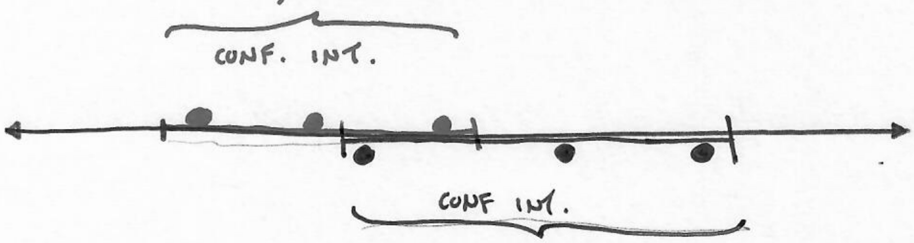
NEW RANDOM VARIABLE: $\bar{X}_1 - \bar{X}_2$ DIFFERENCE OF SAMPLE MEANS.

$\bar{X}_1$ CLT $(n \geq 30)$

$\therefore$ NORMAL DISTR.

$\bar{X}_2$ NORM DISTR

$\mu_1 \approx \bar{x}_1$



$\mu_2 \approx \bar{x}_2$

ESTIMATE $\mu_1 - \mu_2 \approx \bar{x}_1 - \bar{x}_2$ (POINT ESTIMATE)

$\mu_1 - \mu_2 \approx \bar{x}_1 - \bar{x}_2 + \underbrace{\text{MARGIN OF ERROR}}_{1.96\ \text{S.E.}}$

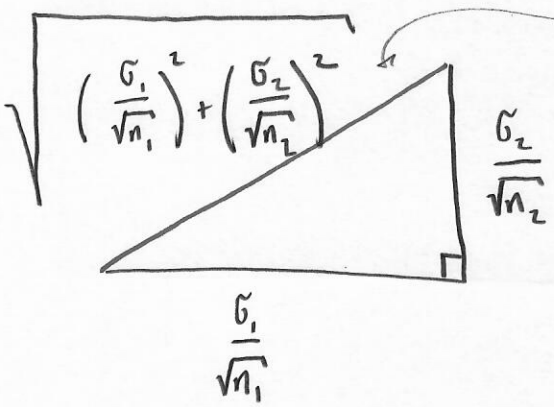
WHAT IS THE STANDARD ERROR FOR

$\bar{X}_1 - \bar{X}_2$?

WHEN INDEPENDENT RANDOM SAMPLES OF SIZE n_1 & n_2 OBSERVATIONS ARE TAKEN FROM POPULATIONS WITH MEANS μ_1 & μ_2 & STANDARD DEVIATIONS σ_1 & σ_2, RESPECTIVELY THE SAMPLING DISTRIBUTION OF THE DIFFERENCE $\bar{x}_1 - \bar{x}_2$ OF SAMPLE MEANS HAS MEAN $= \mu_1 - \mu_2$ (UNBIASED)

AND STANDARD ERROR $S.E. = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$

$$\underset{\left(S.E.\text{ FOR } \bar{x}_1\right)^2}{\dfrac{\sigma_1^2}{n_1}} \qquad \underset{\left(S.E.\text{ FOR } \bar{x}_2\right)^2}{\dfrac{\sigma_2^2}{n_2}}$$



- THE SAMPLING DISTRIBUTION FOR $(\bar{x}_1 - \bar{x}_2)$ IS EXACTLY NORMAL IF THE POPULATIONS ARE NORMALLY DISTRIBUTED.

- THE SAMPLING DISTRIBUTION FOR $(\bar{x}_1 - \bar{x}_2)$ IS APPROXIMATELY NORMAL IF BOTH n_1 & $n_2 \geq 30$ ($n_1 \geq 30$ AND $n_2 \geq 30$)

IN SUMMARY: $\bar{X}_1 - \bar{X}_2$ IS APPROX. NORMALLY DISTR.

WITH MEAN $\mu_1 - \mu_2$

S.E. $\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$$\approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$Z = \frac{\text{R.V.} - \text{MEAN}}{\text{STAND DEV.}}$

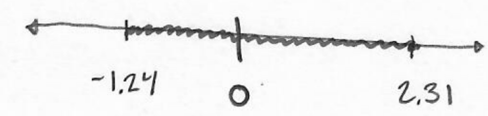
THE STATISTIC $Z \approx \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ HAS (APPROX) STANDARD NORMAL DISTRIBUTION.

POINT ESTIMATE: $\mu_1 - \mu_2 \approx \bar{X}_1 - \bar{X}_2 \pm 1.96 \underbrace{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}_{\text{S.E.}}$

CONFIDENCE INT. $\mu_1 - \mu_2 \approx \underbrace{\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

WE CONCLUDE THAT THERE IS LIKELY A DIFFERENCE BETWEEN THE POPULATION MEANS IF 0 IS NOT CONTAINED IN THIS INTERVAL.

e.g. CONFIDENCE INT. FOR $\mu_1 - \mu_2$:



·) $\mu_1 - \mu_2 = -1.24$

$$\mu_1 = \mu_2 - 1.24 \Rightarrow \mu_1 > \mu_2 \quad \text{(DIFFERENCE)}$$

·) $\mu_1 - \mu_2 = 2.31$

$$\mu_1 = \mu_2 + 2.31 \Rightarrow \mu_2 > \mu_1 \quad \text{(DIFFERENCE)}$$

·) $\mu_1 - \mu_2 = 0 \Rightarrow \mu_1 = \mu_2$ (NO DIFFERENCE)

8.42

$n_1 = 64$ $\qquad$ $n_2 = 64$

$\bar{X}_1 = 2.9$ $\qquad$ $\bar{X}_2 = 5.1$

$S_1 = \sqrt{S_1^2} = \sqrt{.83}$ $\qquad$ $S_2 = \sqrt{S_2^2} = \sqrt{1.67}$

(a)
$$\mu_1 - \mu_2 \approx \bar{X}_1 - \bar{X}_2 \pm (1.645)\underbrace{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}_{\text{S.E.}}$$

$$\mu_1 - \mu_2 \approx 2.9 - 5.1 \pm 1.645\sqrt{\frac{.83}{64} + \frac{1.67}{64}}$$

$$-2.2 \pm .3251 \rightarrow [-2.5251, -1.8749]$$

90% OF INTERVALS COMPUTED IN THIS WAY

REPEAT THE EXPERIMENT & GENERATE 90% CONF. INTERVAL.

WILL CONTAIN $\mu_1 - \mu_2$ (TRUE VALUE OF DIFF. OF POP. MEANS)

(b)

$$\mu_1 - \mu_2 \approx \bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\, S.E.$$

$$S.E. = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$\mu_1 - \mu_2 \approx 2.9 - 5.1 \pm 2.58\sqrt{\frac{.83}{64} + \frac{1.67}{64}}$$

$$\approx -2.2 \pm .5099$$

DOES NOT CONTAIN 0

YES, WE CAN CONCLUDE WITH 99% CONFIDENCE THAT THE DIFFERENCE IS BETWEEN -2.7099 & -1.6901, IN PARTICULAR THE DIFFERENCE IS NEGATIVE, AND $\neq 0$.

$$\mu_1 - \mu_2 \approx \bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\approx 2.4 - 3.1 \pm 1.645 \sqrt{\frac{1.44}{100} + \frac{2.64}{100}}$$

$$\mu_1 - \mu_2 \approx -.7 \pm 1.645\,(.2020)$$

$$\underbrace{[-1.0323\ ,\ -.3677}\ ]$$

$$-1.0323 \le \mu_1 - \mu_2 \le -.3677$$

$$\mu_2 - 1.0323 \le \mu_1 \le \mu_2 - .3677$$

POP 1 HAS SMALLER MEAN THAN POP 2

(WITH 90% CONFIDENCE)

$$\mu_1 - \mu_2 \approx \bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\approx 15 - 23 \pm 2.58\sqrt{\frac{4^2}{30} + \frac{10^2}{40}}$$

$$\approx -8 \pm 4.4934$$

$$\underbrace{[-12.4934, -3.5066]}$$

DOES NOT CONTAIN 0

$\therefore$ CONCLUDE THERE IS A DIFFERENCE

AND $\mu_1 - \mu_2 < 0$

$\underline{\mu_1 < \mu_2}$

CONFIDENCE INT.	$z_{\alpha/2}$
99%	2.58
98%	2.33
95%	1.96
90%	1.645

IF THE CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$ CONTAINS 0,

THE IT IS POSSIBLE THAT $\mu_1 - \mu_2 = 0$

i.e. $\mu_1 = \mu_2$

i.e. NO DIFFERENCE BETWEEN POPULATIONS.

SO IN THIS CASE WE CANNOT THAT THERE IS A DIFFERENCE BETWEEN THE POPULATIONS.