90 POINTS $\qquad \dfrac{\#\text{ POINTS}}{90} = $ PERCENTAGE

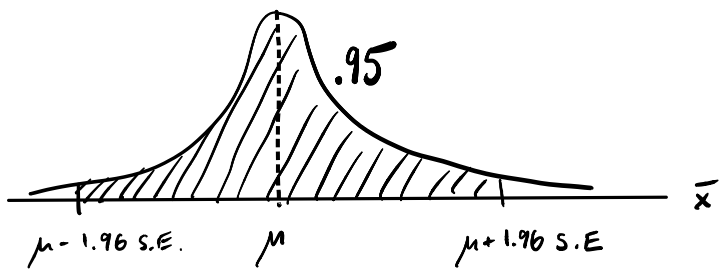
§ 8.5 INTERVAL ESTIMATION

CENTRAL LIMIT THEOREM:

WHEN SAMPLE SIZE $n \geq 30$ (OR IF POPULATION HAS NORMAL DISTRIBUTION)
THEN THE SAMPLE DISTRIBUTION OF THE SAMPLE MEAN $\bar{X}$ IS
APPROX. NORMALLY DISTRIBUTED WITH MEAN μ (SAME AS POP.)
AND STANDARD ERROR $S.E. = \dfrac{\sigma}{\sqrt{n}} \approx \dfrac{s}{\sqrt{n}}$

WE USE SAMPLE STAND. DEV. TO APPROXIMATE THE POPULATION STANDARD DEVIATION.

P(SAMPLE MEAN $\bar{X}$ IS WITHIN 1.96 S.E. OF POP MEAN μ) = .95



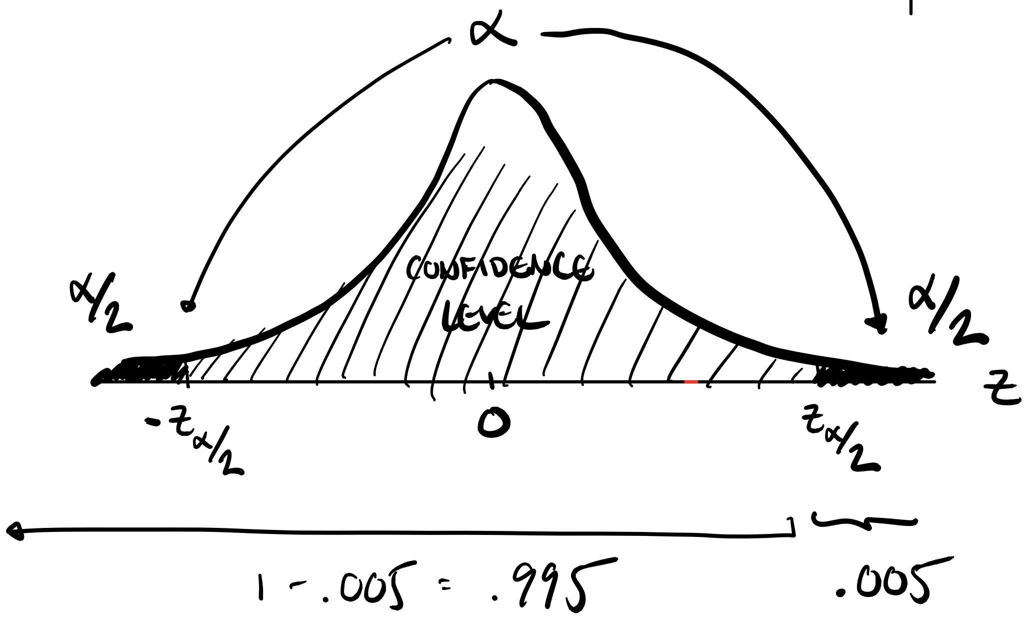
P(POP MEAN μ IS WITHIN 1.96 S.E. OF SAMPLE MEAN $\bar{X}$) = .95

95% PROB μ IS IN THIS INTERVAL

$\Rightarrow$ 5% PROB μ IS NOT IN THIS INTERVAL

CONFIDENCE INTERVALS

CONFIDENCE LEVEL (PROB THAT μ IS IN THE INTERVAL ESTIMATE)	α = PROB THAT μ IS NOT IN THE INTERVAL ESTIMATE	$\alpha/2$	$z_{\alpha/2}$
.95	.05	.025	1.96
.90	.10	.05	1.645
.98	.02	.01	2.33
.99	.01	.005	2.58



$P($ μ IS WITHIN $z_{\alpha/2}$ S.E. OF $\overline{x}$ $)$ = CONFIDENCE LEVEL

8.33 Acid Rain Acid rain, caused by the reaction of certain air pollutants with rainwater, is a growing problem in the United States. Pure rain falling through clean air registers a pH value of 5.7 (pH is a measure of acidity: 0 is acid; 14 is alkaline). Suppose water samples from 40 rainfalls are analyzed for pH, and $\bar{x}$ and s are equal to 3.7 and .5, respectively. Find a 99% confidence interval for the mean pH in rainfall and interpret the interval. What assumption must be made for the confidence interval to be valid?

SAMPLE DATA

$\bar{x} = 3.7$

$s = .5$

$S.E = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$

$= \frac{.5}{\sqrt{40}}$

ESTIMATE μ (MEAN PH RAINFALL)

CONFIDENCE LEVEL 99% $\Rightarrow$ $\alpha = .01$, $\alpha/2 = .005$, $z_{\alpha/2} = 2.58$

$\bar{x} - z_{\alpha/2}\, S.E.$ $\qquad \bar{x} \qquad$ $\bar{x} + z_{\alpha/2}\, S.E.$

$3.7 - 2.58\left(\frac{.5}{\sqrt{40}}\right)$ $\qquad 3.7 \qquad$ $3.7 + 2.58\left(\frac{.5}{\sqrt{40}}\right)$

3.4960 $\qquad\qquad$ 3.9040

$$\boxed{\text{CONFIDENCE INTERVAL: } \left[\bar{x} - z_{\alpha/2}\, S.E.\,,\ \bar{x} + z_{\alpha/2}\, S.E.\right]}$$

99% CONF. INT. $[3.4960, 3.9040]$

$P(3.4960 \le \mu \le 3.9040) = .99$

ASSUMPTION: RAINFALLS IN SAMPLE MUST BE RANDOMLY SELECTED.

If this experiment were to be repeated over and over many times, and a confidence interval is generated each time, just like this, then approximately 99% of those confidence intervals would contain the true population mean.

8.35 **Hamburger Meat** The meat department of a local supermarket chain packages ground beef using meat trays of two sizes: one designed to hold approximately 1 pound of meat, and one that holds approximately 3 pounds. A random sample of 35 packages in the smaller meat trays produced weight measurements with an average of 1.01 pounds and a standard deviation of .18 pound.

a. Construct a 99% confidence interval for the average weight of all packages sold in the smaller meat trays by this supermarket chain.

b. What does the phrase "99% confident" mean?

c. Suppose that the quality control department of this supermarket chain intends that the amount of ground beef in the smaller trays should be 1 pound on average. Should the confidence interval in part a concern the quality control department? Explain.

SAMPLE DATA: $\bar{x} = 1.01$ $n = 35$

$s = .18$

$$S.E = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}} = \frac{.18}{\sqrt{35}}$$

CONFIDENCE INTERVAL: $\left[\bar{x} - z_{\alpha/2}\, S.E\ ,\ \bar{x} + z_{\alpha/2}\, S.E.\right]$

CONFIDENCE LEVEL (PROB THAT μ IS IN THE INTERVAL ESTIMATE)	α = PROB THAT μ IS NOT IN THE INTERVAL ESTIMATE	$\alpha/2$	$z_{\alpha/2}$
.95	.05	.025	1.96
.90	.10	.05	1.645
.98	.02	.01	2.33
.99	.01	.005	2.58

99% CONFIDENCE INTERVAL: $\left[1.01 - 2.58\left(\frac{.18}{\sqrt{35}}\right),\ 1.01 + 2.58\left(\frac{.18}{\sqrt{35}}\right)\right]$

$= [.9315,\ 1.0885]$

SINCE 1 IS IN THE INTERVAL, THIS DOES NOT SUGGEST THAT THE AVERAGE IS NOT 1.

If this experiment were to be repeated over and over many times, and a confidence interval is generated each time, just like this, then approximately 99% of those confidence intervals would contain the true population mean.

90% CONFIDENCE INTERVAL: $\left[1.01 - 1.645\left(\frac{.18}{\sqrt{35}}\right),\ 1.01 + 1.645\left(\frac{.18}{\sqrt{35}}\right)\right]$

$= [.9599,\ 1.0601]$

IF 99% CONF. INT. $[.8873,\ .9871]$ NOT IN THIS INTERVAL.

THEN IT WOULD BE UNLIKELY THAT $\mu = 1$

$< 1\%$ PROBABILITY.

IF 99% CONF. INT. $[1.0231 \;,\; 1.0924]$ NOT IN THIS INTERVAL.

THEN IT WOULD BE UNLIKELY THAT $\mu = 1$

$< 1\%$ PROBABILITY.

$$S.E. = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{n}}$$

CONFIDENCE INTERVALS FOR POPULATION PROPORTIONS

$\frac{x}{n}$ ← # SUCCESSES IN SAMPLE OF SIZE n
← SAMPLE SIZE.

SAMPLE PROPORTIONS $\hat{p}$ ARE APPROXIMATELY NORMALLY DISTRIBUTED WITH MEAN p (POPULATION PROPORTION) & STANDARD ERROR

$$S.E. = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\hat{q} = 1 - \hat{p}$$

(AS LONG AS $n \geq 30$)

CONFIDENCE LEVEL (PROB THAT μ IS IN THE INTERVAL ESTIMATE)	α = PROB THAT μ IS NOT IN THE INTERVAL ESTIMATE	$\alpha/2$	$z_{\alpha/2}$
.95	.05	.025	1.96
.90	.10	.05	1.645
.98	.02	.01	2.33
.99	.01	.005	2.58

$$\boxed{\text{CONFIDENCE INTERVAL FOR } p: \left[\hat{p} - z_{\alpha/2}\, S.E. \;,\; \hat{p} + z_{\alpha/2}\, S.E.\right]}$$

8.40 Gonna' Vote? How likely are you to vote in the next national election? In a survey by *Pew Research,*[10] fully 77% of the registered Republican voters are *absolutely* going to vote this year while only 65% of Democrats are *absolutely* going to vote in the next election. The sample consisted of 469 registered Republicans, 490 registered Democrats, and 480 registered Independents.

a. Construct a 98% confidence interval for the proportion of registered Republicans who say they are *absolutely* going to vote in the next election. If a Republican senator predicts that at least 85% of registered Republicans will *absolutely* vote in the next election, is this figure realistic?

b. Construct a 99% confidence interval for the proportion of registered Democrats who say they are *absolutely* going to vote in the next election.

(a) $n = 469 \quad \hat{p} = .77 \quad \hat{q} = .23$

98% CONF. INT $\Rightarrow z_{\alpha/2} = 2.33$

$$S.E. = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(.77)(.23)}{469}}$$

98% CONF. INT: $\left[.77 - 2.33\sqrt{\frac{(.77)(.23)}{469}} \;,\; .77 + 2.33\sqrt{\frac{(.77)(.23)}{469}}\right]$

$$= [.7247 \;,\; .8153]$$

If this experiment were to be repeated over and over many times, and a confidence interval is generated each time, just like this, then approximately 99% of those confidence intervals would contain the true population proportion.

$P(p \text{ IS NOT IN } [.7247, .8153]) = .02$

SO THIS IS NOT REALISTIC.

8.36 Same-Sex Marriage The results of a *CBS News Poll* concerning views on same-sex marriage and gay rights given in Exercise 7.68 showed that of $n = 1082$ adults, 40% favored legal marriage, 30% favored civil unions, and 25% believed there should be no legal recognition.[7] The poll reported a margin of error of plus or minus 3%.

a. Construct a 90% confidence interval for the proportion of adults who favor the "legal marriage" position.

b. Construct a 90% confidence interval for the proportion of adults who favor the "civil unions" position.

c. How did the researchers calculate the margin of error for this survey? Confirm that their margin of error is correct.

CONFIDENCE LEVEL (PROB THAT μ IS IN THE INTERVAL ESTIMATE)	α = PROB THAT μ IS NOT IN THE INTERVAL ESTIMATE	$\alpha/2$	$z_{\alpha/2}$
.95	.05	.025	1.96
.90	.10	.05	1.645
.98	.02	.01	2.33
.99	.01	.005	2.58

$$S.E. = \sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\boxed{\text{CONFIDENCE INTERVAL FOR } p: \left[\hat{p} - z_{\alpha/2}\, S.E.,\ \hat{p} + z_{\alpha/2}\, S.E.\right]}$$

(a) GIVEN: $n = 1082$, $\hat{p} = .40$, $\hat{q} = .60$

$$\therefore S.E. \approx \sqrt{\frac{(.4)(.6)}{1082}}$$

$$90\% \text{ CONF. INT.}: \left[.4 - 1.645\sqrt{\frac{(.4)(.6)}{1082}},\ .4 + 1.645\sqrt{\frac{(.4)(.6)}{1082}}\right]$$

$$[.3755, .4245]$$

§ 8.6 ESTIMATING THE DIFFERENCE BETWEEN 2 POPULATION MEANS

e.g. Plants given fertilizer grow 5 inches taller than plant not given fertilizer. *

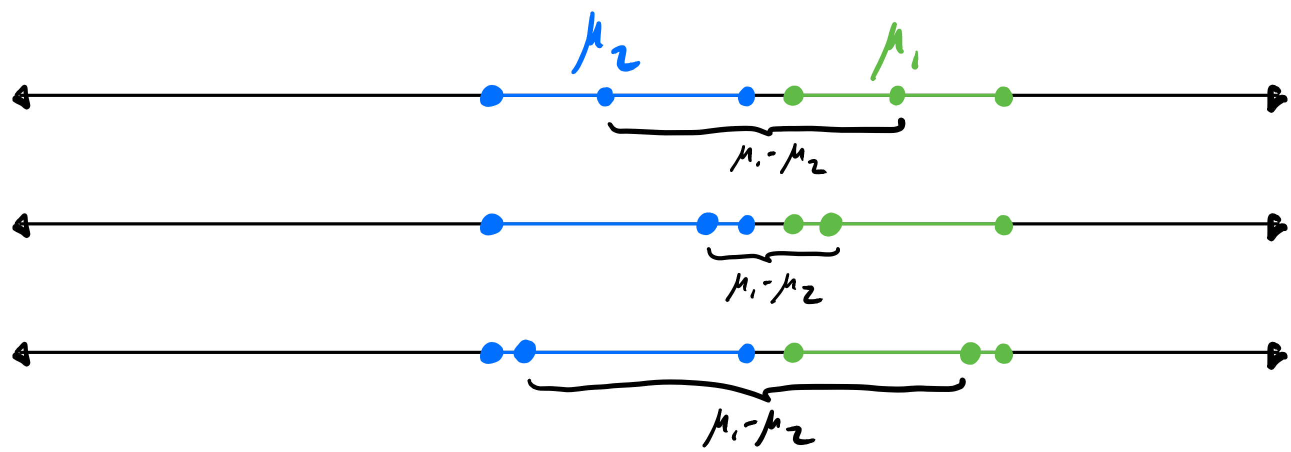
Clients following exercise program X lost 5 pounds more than clients following exercise program Y.

City dwellers watch 5 hours less TV per week than non-city dwellers.

Etc.

* TWO POPULATIONS: POP 1: PLANTS GIVEN FERTILIZER ⟶ 99% CONF INT μ_1

POP 2: PLANTS NOT GIVEN FERTILIZER ⟶ 99% CONF INT. μ_2



	POPULATION 1	POPULATION 2	SAMPLE 1	SAMPLE 2
MEAN	μ_1	μ_2	$\bar{x}_1$	$\bar{x}_2$
STAND. DEV.	σ_1	σ_2	s_1	s_2
SIZE	N_1	N_2	n_1	n_2

NOW WE WANT TO ESTIMATE $(\mu_1 - \mu_2)$ (DIFF. OF POP. MEANS).

C.L.T. $\Rightarrow$ $(\bar{X}_1 - \bar{X}_2)$ IS AN UNBIASED ESTIMATOR.

Furthermore, the sampling distribution of the differences $\bar{X}_1 - \bar{X}_2$ HAS

MEAN $\mu_1 - \mu_2$ & STANDARD ERROR $S.E. = \sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}} \approx \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}$.

THE SAMPLING DISTRIBUTION OF $(\bar{x}_1 - \bar{x}_2)$ IS EXACTLY NORMAL IF THE POPULATIONS ARE NORMALLY DISTRIBUTED, AND APPROXIMATELY NORMAL IF BOTH n_1 & $n_2 \geq 30$.

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$:

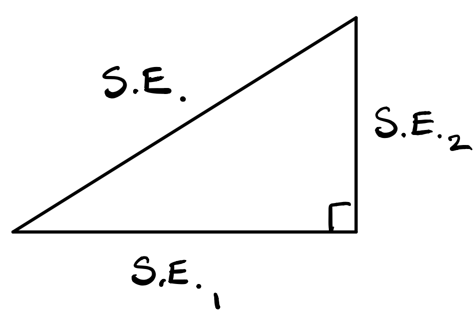
$$\left[(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\, S.E. \ ,\ (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\, S.E. \right]$$

$$S.E.\ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

NOTE: $S.E._1 = \dfrac{s_1}{\sqrt{n}}$ $\qquad$ $S.E._2 = \dfrac{s_2}{\sqrt{n}}$

$$S.E. = \sqrt{S.E._1^2 + S.E._2^2}$$

$$S.E.^2 = S.E._1^2 + S.E._2^2$$



EXAMPLE 8.9

The wearing qualities of two types of automobile tires were compared by road-testing samples of $n_1 = n_2 = 100$ tires for each type and recording the number of miles until wearout, defined as a specific amount of tire wear. The test results are given in Table 8.4. Estimate $(\mu_1 - \mu_2)$, the difference in mean miles to wearout, using a 99% confidence interval. Is there a difference in the average wearing quality for the two types of tires?

TABLE 8.4 Sample Data Summary for Two Types of Tires

Tire 1	Tire 2
$\bar{x}_1 = 26{,}400$ miles	$\bar{x}_2 = 25{,}100$ miles
$s_1^2 = 1{,}440{,}000$	$s_2^2 = 1{,}960{,}000$

CONFIDENCE INTERVAL FOR $\mu_1 - \mu_2$:

$$\left[(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2}\ S.E.\ ,\ (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2}\ S.E. \right], \quad S.E.\ \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \approx \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

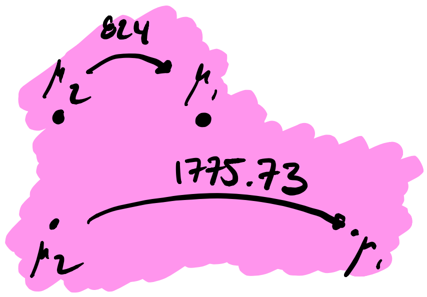
$$\bar{x}_1 - \bar{x}_2 = 26{,}400 - 25{,}100 = 1{,}300$$

$$S.E. \approx \sqrt{\frac{1{,}440{,}000}{100} + \frac{1{,}960{,}000}{100}} = 184.39$$

99% CONF. INT. : $\left[1{,}300 - 2.58(184.39)\ ,\ 1{,}300 + 2.58(184.39) \right]$

$= [824.27\ ,\ 1775.73]$ ← IS THERE A DIFFERENCE?

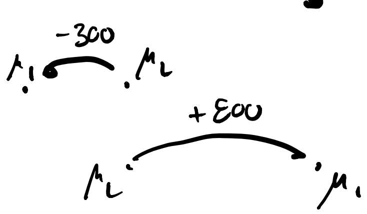
$\mu_1 - \mu_2$ 99% SURE



$$824.27 < \mu_1 - \mu_2 < 1775.73$$

$$\mu_2 + 824.27 < \mu_1 < \mu_2 + 1775.73$$

$[-300\ ,\ 800]$



$\mu_1 < \mu_2$ ← $\mu_1 - \mu_2 < 0$

$\mu_1 > \mu_2$ ← $\mu_1 - \mu_2 > 0$

$\mu_1 = \mu_2$ ← $\mu_1 - \mu_2 = 0$

NOTE: IF CONF. INT. FOR DIFF. OF POP. MEANS CONTAINS 0, THEN THERE MAY BE NO DIFFERENCE BETWEEN POPULATIONS.

IF IT DOES <u>NOT</u> CONTAIN 0, THEN SAMPLE DATA SUGGESTS THAT THERE <u>IS</u> A DIFFERENCE BETWEEN THE POPULATION MEANS