AP Statistics **HW Pg. 590 #11.5, 11.8, 11.18**
Mr. Murphy
11.1 Comparing Two Populations or Treatments

Goals:
1. Run a hypothesis test for the **difference in means** between **two <u>independent</u> samples**
2. Construct a confidence interval for the **difference in means** between **two <u>independent</u> samples**

In this section, we consider using sample data to **compare** two population means or two treatment means.

■ Summary of the Two-Sample *t* Test for Comparing Two Population Means

Null hypothesis: H_0: $\mu_1 - \mu_2$ = hypothesized value

Test statistic: $$t = \frac{\bar{x}_1 - \bar{x}_2 - \text{hypothesized value}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

The appropriate number of degrees of freedom for the two-sample *t* test is

$$\text{df} = \frac{(V_1 + V_2)^2}{\frac{V_1^2}{n_1 - 1} + \frac{V_2^2}{n_2 - 1}}$$

where

$$V_1 = \frac{s_1^2}{n_1} \quad \text{and} \quad V_2 = \frac{s_2^2}{n_2}$$

The number of degrees of freedom should be truncated (rounded down) to an integer.

Alternative Hypothesis	***P*-Value**
H_a: $\mu_1 - \mu_2 >$ hypothesized value	Area under appropriate *t* curve to the right of the computed *t*
H_a: $\mu_1 - \mu_2 <$ hypothesized value	Area under appropriate *t* curve to the left of the computed *t*
H_a: $\mu_1 - \mu_2 \neq$ hypothesized value	(1) 2(area to the right of the computed *t*) if *t* is positive, or (2) 2(area to the left of the computed *t*) if *t* is negative

Assumptions: 1. The two samples are *independently selected random samples*.
2. The *sample sizes are large* (in general, 30 or larger), or *the population distributions are (at least approximately) normal.*

Ex1 To assess the impact of oral contraceptive use on bone mineral density (BMD), researchers in Canada carried out a study comparing BMD for women who had used oral contraceptives for at least 3 months to BMD for women who had never used oral contraceptives. Data on BMD (in grams per centimeter) is given below.

Never used oral contraceptives:	0.82	0.94	0.96	1.31	0.94
	1.21	1.26	1.09	1.13	1.14
Used oral contraceptives:	0.94	1.09	0.97	0.98	1.14
	0.85	1.30	0.89	0.87	1.01

The authors of the paper believed that it was reasonable to view the samples used in the study as representatives of the two populations of interest. Use the given information and a significance level of 0.05 to determine whether there is evidence that women who use oral contraceptives have a lower BMD than women who have never used oral contraceptives.

Ex2 First year nursing students enrolled in a science course were randomly assigned to one of two groups. Students in the control group were provided with conventional study–skill guidelines. Students in the treatment group were provided with homework prescriptions based on their identified learning style as well as with the conventional study–skill guidelines. Does the treatment (the homework prescriptions based on learning style) improve science GPA? We will test the relevant hypothesis using a significance level of 0.01.

▪ **Two-Sample *t* Test for Comparing Two Treatments**

When

1. *treatments are randomly assigned* to individuals or objects (or vice versa; i.e., individuals or objects are randomly assigned to treatments), and
2. the *sample sizes are large* (in general, 30 or larger) or the *treatment response distributions are approximately normal*

the two-sample t test can be used to test H_0: $\mu_1 - \mu_2 =$ hypothesized value, where μ_1 and μ_2 represent the mean response for Treatments 1 and 2, respectively.

In this case, these two conditions replace the assumptions previously stated for comparing two population means. The validity of the assumption of normality of the treatment response distributions can be assessed by constructing a normal probability plot or a boxplot of the response values in each sample.

Group	Sample Size	Sample Mean	Sample SD
Control	100	2.75	0.85
Treatment	103	3.28	0.70

Ex3 Does talking elevate blood pressure, contributing to the tendency for blood pressure to be higher when measured in a doctor's office than when measured in a less stressful environment? Patients with high blood pressure were randomly assigned to one of two groups. Those in the first group (the talking group) were asked questions about their medical history and about the sources of stress in their lives in the minutes before their blood pressure was measured. Those in the second group (the counting group) were asked to count aloud from 1 to 100 four times before their blood pressure was measured. The following data values for diastolic blood pressure (in millimeters Hg) are consistent with the summary quantities appearing in the paper:

▪ **Two-Sample *t* Confidence Interval for the Difference Between Two Population or Treatment Means**

The general formula for a confidence interval for $\mu_1 - \mu_2$ when

1. the two samples are *independently chosen random samples*, and
2. the *sample sizes are both large* (in general, $n_1 \geq 30$ and $n_2 \geq 30$) or the *population distributions are approximately normal* is

$$\bar{x}_1 - \bar{x}_2 \pm (t \text{ critical value})\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The *t* critical value is based on

$$\text{df} = \frac{(V_1 + V_2)^2}{\dfrac{V_1^2}{n_1 - 1} + \dfrac{V_2^2}{n_2 - 1}}$$

where

$$V_1 = \frac{s_1^2}{n_1} \quad \text{and} \quad V_2 = \frac{s_2^2}{n_2}$$

The number of degrees of freedom should be truncated (rounded down) to an integer. The *t* critical values for the usual confidence levels are given in Appendix Table 3.

Talking	104	110	107	112	108	103	108	118
	$n_1 = 8$		$\bar{x}_1 = 108.75$			$s_1 = 4.74$		
Counting	110	96	103	98	100	109	97	105
	$n_2 = 8$		$\bar{x}_2 = 102.25$			$s_2 = 5.39$		

Construct a 95% C.I. to estimate $\mu_1 - \mu_2$.

Checkpoint
Multiple Choice

1. Suppose we compare the data on two independent and random samples, *A* and *B*, and come up with the following data: $n_A = n_B = 10,\ \bar{x}_A = 25, s_A = 3.21,\ \bar{x}_B = 22.2,\ s_B = 3.09$. Find a 99% confidence interval for the difference between the means. Assume the samples are independent.

(a) (-1.26, 6.86)
(b) (-1.78, 7.38)
(c) (-1.67, 7.27)
(d) (-0.83, 6.43)
(e) None of the above

2. A random sample of 32 games is chosen for a professional basketball team, team A, and their results are recorded. The team averages 88 points per game with a standard deviation of 8. The same is done for a second team, team B, with this team averaging 90 points per game with a standard deviation of 6. A 95% confidence interval is constructed for the difference in points scored per game between the two teams. What do the results of the confidence interval show?

(a) We can be 95% confident that on average team A scores between 1.54 and 5.54 more points per game than team B.
(b) We can be 95% confident that on average team A scores between 1.54 and 5.54 fewer points per game than team B.
(c) We can be 95% confident that on average team A scores between 1.54 points fewer than and 5.54 points more than team B.
(d) We can be 95% confident that on average team A scores between 5.54 points fewer than and 1.54 points more than team B.
(e) The conditions necessary to find a 95% confidence interval have not been met.

3. Two professors, A and B, got into an argument about who grades tougher. Professor A insisted that his grades were lower than those for Professor B. In order to test this theory, each professor took a random sample of 25 student grades and conducted a test of significance. The graphical displays showed each grade distribution was approximately normal. The results are recorded below.

H_0 : Population mean of Professor A *equals that* of Professor B
H_a : Population mean of Professor A *less than that* of Professor B

		Professor A	**Professor B**	
Count:		25	25	
Mean:		79	82	
Std dev:		6	4	
Std error:		1.2	0.8	

Student's t: -2.08 DF: 41.844 P-value: 0.022

Which of the following conclusions is/are supported by the results of the significance test?

I. At the $\alpha = 0.05$ level, we have evidence to show that every student in Professor A's class scored lower than every student in Professor B's class.

II. If there were no difference in grades between the two professors, then we could get results as extreme as those from the samples approximately 2.2% of the time.

III.The test results are not valid, since the conditions necessary to perform the test were not met.

(a) I only
(b) II only
(c) III only
(d) I and II
(e) II and III