Section 10.2
Comparing Two Means

The Sampling Distribution of the Difference Between Sample Means

Choose an SRS of size n_1 from Population 1 with mean μ_1 and standard deviation σ_1 and an independent SRS of size n_2 from Population 2 with mean μ_2 and standard deviation σ_2.

Shape When the population distributions are Normal, the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is approximately Normal. In other cases, the sampling distribution will be approximately Normal if the sample sizes are large enough ($n_1 \geq 30, n_2 \geq 30$).

Center The mean of the sampling distribution is $\mu_1 - \mu_2$. That is, the difference in sample means is an unbiased estimator of the difference in population means.

Spread The standard deviation of the sampling distribution of $\bar{x}_1 - \bar{x}_2$ is

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

as long as each sample is no more than 10% of its population (10% condition).

Feb 8-11:27 AM

The Two-Sample *t* Statistic cont.

If the Normal condition is met, we standardize the observed difference to obtain a *t* statistic that tells us how far the observed difference is from its mean in standard deviation units:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

0

The two-sample *t* statistic has approximately a *t* distribution. We can use technology to determine degrees of freedom OR we can use a conservative approach, using the smaller of $n_1 - 1$ *and* $n_2 - 1$ for the degrees of freedom.

Feb 8-12:16 PM

Conditions for Performing Inference About $\mu_1 - \mu_2$

Random: The data come from two independent random samples or from two groups in a randomized experiment.

10%: When sampling without replacement, check that $n_1 \leq 0.10N_1$ and $n_2 \leq 0.10N_2$.

Normal/Large Sample: Both population distributions (or the true distributions of responses to the two treatments) are Normal or both sample sizes are large($n_1 \geq 30$ and $n_2 \geq 30$). If either population (treatment) distribution has unknown shape and the corresponding sample size is less than 30, use a graph of the sample data to assess the Normality of the population (treatment) distribution. Do not use two-sample *t* procedures if the graph shows strong skewness or outliers.

Feb 8-12:17 PM

Confidence Intervals for $\mu_1 - \mu_2$

Two-Sample *t* Interval for a Difference Between Two Means

When the conditions are met, an approximate *C*% confidence interval for $\mu_1 - \mu_2$ is

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

On the calculator 2-sample t interval (not pooled)

Here, t^* is the critical value with C% of its area between $-t^*$ and t^* for the *t* distribution with degrees of freedom using either Option 1 (technology) or Option 2 (the smaller of $n_1 - 1$ and $n_2 - 1$).

Feb 8-12:19 PM

Example 1: A fast-food restaurant uses an automated filling machine to pour its soft drinks. The machine has different settings for small, medium, and large drink cups. According to the machine's manufacturer, when the large setting is chosen, the amount of liquid *L* dispensed by the machine follows a Normal distribution with mean 27 ounces and standard deviation 0.8 ounces. When the medium setting is chosen, the amount of liquid *M* dispensed follows a Normal distribution with mean 17 ounces and standard deviation 0.5 ounces. To test the manufacturer's claim, the restaurant manager measures the amount of liquid in each of 20 cups filled with the large setting and 25 cups filled with the medium setting. Let $\bar{x}_L - \bar{x}_M$ be the difference in the sample mean amount of liquid under the two settings.

a) What is the shape of the sampling distribution of $\bar{x}_L - \bar{x}_M$? Why?

The sampling distribution of $\bar{x}_L - \bar{x}_m$ is approximately normal because both population distributions are normal.

b) Find the mean of the sampling distribution. Show your work.

The mean is $\mu_{\bar{x}_L - \bar{x}_M} = 27 - 17 = 10$ ounces

c) Find the standard deviation of the sampling distribution. Show your work.

The standard deviations is

$$\sigma_{\bar{x}_L - \bar{x}_M} = \sqrt{\frac{\sigma_L^2}{n_L} + \frac{\sigma_M^2}{n_M}} = \sqrt{\frac{(0.80)^2}{20} + \frac{(0.50)^2}{25}} = 0.205 \text{ ounces.}$$

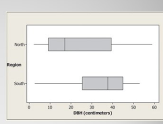
Note that we do not need to check the 10% condition because we are not sampling without replacement from a finite population.

Feb 8-12:10 PM

Example 2: The Wade Tract Preserve in Georgia is an old-growth forest of longleaf pines that has survived in a relatively undisturbed state for hundreds of years. One question of interest to foresters who study the area is "How do the sizes of longleaf pine trees in the northern and southern halves of the forest compare?" To find out, researchers took random samples of 30 trees from each half and measured the diameter at breast height (DBH) in centimeters. Here are comparative boxplots of the data and summary statistics from Minitab.

Descriptive Statistics: North, South

Variable	N	Mean	StDev
North	30	23.70	17.50
South	30	34.53	14.26



b) Construct and interpret a 90% confidence interval for the difference in the mean DBH for longleaf pines in the northern and southern halves of the Wade Tract Preserve.

State: Our parameters of interest are μ_1 = the true mean DBH of all trees in the southern half of the forest and μ_2 = the true mean DBH of all trees in the northern half of the forest. We want to estimate the difference $\mu_1 - \mu_2$ at a 90% confidence level.

Plan: We should use a two-sample *t* interval for $\mu_1 - \mu_2$ if the conditions are satisfied.

✓ **Random** The data come from a random samples of 30 trees each from the northern and southern halves of the forest.

✓**10%:** Because sampling without replacement was used, there have to be at least 10(30) = 300 trees in each half of the forest. This is fairly safe to assume.

Feb 8-12:20 PM

✓ **Normal** The boxplots give us reason to believe that the population distributions of DBH measurements may not be Normal. However, since both sample sizes are at least 30, we are safe using *t* procedures.

Do: Since the conditions are satisfied, we can construct a two-sample *t* interval for the difference $\mu_1 - \mu_2$. We'll use the conservative df = 30 – 1 = 29.

From the minitab output, $\bar{x}_1 = 34.53$, $s_1 = 14.26$, $n_1 = 30$, $\bar{x}_2 = 23.70$, $s_2 = 17.50$, and $n_2 = 30$. We'll use the conservative df = the smaller of $n_1 - 1$ and $n_2 - 1$, which is 29. For a 90% confidence level the critical value from **Table B** is $t^* = 1.699$. So a 90% confidence interval for $\mu_1 - \mu_2$ is

Upper-tail probability *p*

df	.10	.05	.025
28	1.313	1.701	2.048
29	1.311	1.699	2.045
30	1.310	1.697	2.042
	80%	90%	95%

Confidence level *C*

$$(\bar{x}_1 - \bar{x}_2) \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (34.53 - 23.70) \pm 1.699 \sqrt{\frac{14.26^2}{30} + \frac{17.50^2}{30}}$$

$$= 10.83 \pm 7.00$$

$$= (3.83, 17.83)$$

calculator's 2-SampTIntfeature to compute the confidence interval on the AP® exam. Be sure to name the procedure(two-sample *t* interval) and to give the interval (3.9362, 17.724) and df (55.728) as part of the "Do" step.

Feb 8-12:22 PM

CONCLUDE: We are 90% confident that the interval from 3.9362 to 17.724 centimeters captures the difference in the actual mean DBH of the southern trees and the actual mean DBH of the northern trees.

AP EXAM TIP: The formula for the two-sample *t* interval for $\mu_1 - \mu_2$ often leads to calculation errors by students. As a result, we recommend using the calculator's 2-SampTIntfeature to compute the confidence interval on the AP® exam. Be sure to name the procedure(two-sample *t* interval) and to give the interval(3.9362, 17.724) and df (55.728) as part of the "Do" step.

Feb 8-12:23 PM

CHECK YOUR UNDERSTANDING

The U.S. Department of Agriculture (USDA) conducted a survey to estimate the average price of wheat in July and in September of the same year. Independent random samples of wheat producers were selected for each of the two months. Here are summary statistics on the reported price of wheat from the selected producers, in dollars per bushel:[23]

Month	n	$\bar{x}$	s_x
July	90	\$2.95	\$0.22
September	45	\$3.61	\$0.19

Construct and interpret a 99% confidence interval for the difference in the true mean wheat price in July and in September.

Correct Answer

S: μ_1 = the true mean price of wheat in July and μ_2 = the true mean price of wheat in September. *P*: Two-sample *t* interval for $\mu_1 - \mu_2$. Random: Independent random samples. 10%: $n_1 = 90 < 10\%$ of all wheat producers in July and $n_2 = 45 < 10\%$ of all wheat producers in September. Normal/Large Sample: $n_1 = 90 \geq 30$ and $n_2 = 45 \geq 30$. *D*: Using df = 40, (– 0.759, –0.561). Using df = 100.45, (–0.756, –0.564). *C*: We are 99% confident that the interval from 0.756 to 0.564 captures the true difference in mean wheat prices in July and September.

Feb 8-9:20 PM

Example 3: Does increasing the amount of calcium in our diet reduce blood pressure? Examination of a large sample of people revealed a relationship between calcium intake and blood pressure. The relationship was strongest for black men. Such observational studies do not establish causation. Researchers therefore designed a randomized comparative experiment. The subjects were 21 healthy black men who volunteered to take part in the experiment. They were randomly assigned to two groups: 10 of the men received a calcium supplement for 12 weeks, while the control group of 11 men received a placebo pill that looked identical. The experiment was double blind. The response variable is the decrease in systolic (top number) blood pressure for a subject after 12 weeks, in millimeters of mercury. An increase appears as a negative number. Here are the data:

Group 1 (calcium):	7	−4	18	17	−3	−5	1	10	11	−2	
Group 2 (placebo):	−1	12	−1	−3	3	−5	5	2	−11	−1	−3

a) Do the data provide convincing evidence that a calcium supplement reduces blood pressure more than a placebo? Carry out an appropriate test to support your answer.

Calcium > Placebo

State: We want to perform a test of

$$H_0: \mu_1 - \mu_2 = 0$$
$$H_a: \mu_1 - \mu_2 > 0$$

where μ_1 = the true mean decrease in systolic blood pressure for healthy black men like the ones in this study who take a calcium supplement, and μ_2 = the true mean decrease in systolic blood pressure for healthy black men like the ones in this study who take a placebo. We will use $\alpha = 0.05$.

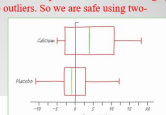
Feb 8-12:24 PM

Plan: If conditions are met, we will carry out a two-sample *t* test for $\mu_1 - \mu_2$.

Random The 21 subjects were randomly assigned to the two treatments.

10%: Don't need to check because there was no sampling.

Normal: With such small sample sizes, we need to graph the data to see if it's reasonable to believe that the actual distributions of differences in blood pressure when taking calcium or placebo are Normal. The figure below shows hand sketches of calculator boxplots for these data. The graphs show no strong skewness and no outliers. So we are safe using two-sample *t* procedures.



Group	Treatment	n	$\bar{x}$	s_x
1	Calcium	10	5.000	8.743
2	Placebo	11	−0.273	5.901

Do: Since the conditions are satisfied, we can perform a two-sample *t* test for the difference $\mu_1 - \mu_2$.

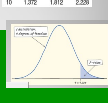
Test statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{[5.000 - (-0.273)] - 0}{\sqrt{\frac{8.743^2}{10} + \frac{5.901^2}{11}}} = 1.604$$

Upper-tail probability *p*

df	.10	.05	.025
8	1.397	1.860	2.306
9	1.383	1.833	2.262
10	1.372	1.812	2.228

***P*-value** Using the conservative df = 10 – 1 = 9, we can use Table B to show that the *P*-value is between 0.05 and 0.10.



Using technology: The calculator's 2-SampTTest gives $t = 1.60$ and *P*-value = 0.0644 using df = 15.59.

Feb 8-12:25 PM

Conclude: Fail to reject H_0. Since the *P*-value, 0.0644, is greater than $\alpha = 0.05$, the experiment does not provide convincing evidence that the true mean decrease in systolic blood pressure is higher for men like these who take calcium than for men like these who take a placebo.

b) Interpret the *P*-value you got in part (a) in the context of this experiment.

Assuming $H_0: \mu_1 - \mu_2 = 0$ is true, the probability of getting a difference in mean blood pressure reduction for the two groups (calcium − placebo) of 5.273 or greater just by the chance involved in the random assignment is 0.0644.

Feb 8-12:27 PM

CHECK YOUR UNDERSTANDING

How quickly do synthetic fabrics such as polyester decay in landfills? A researcher buried polyester strips in the soil for different lengths of time, then dug up the strips and measured the force required to break them. Breaking strength is easy to measure and is a good indicator of decay. Lower strength means the fabric has decayed.

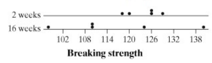
For one part of the study, the researcher buried 10 strips of polyester fabric in well-drained soil in the summer. The strips were randomly assigned to two groups: 5 of them were buried for 2 weeks and the other 5 were buried for 16 weeks. Here are the breaking strengths in pounds:[25]

Group 1 (2 weeks):	118	126	126	120	129
Group 2 (16 weeks):	124	98	110	140	110

Do the data give convincing evidence that polyester decays more in 16 weeks than in 2 weeks?

16 > 2

S: $H_0: \mu_1 - \mu_2 = 0$ versus $H_a: \mu_1 - \mu_2 > 0$, where μ_1 is the true mean breaking strength for polyester fabric buried for 2 weeks and μ_2 is the true mean breaking strength for polyester fabric buried for 16 weeks. *P*: Two-sample *t* test. Random: Two groups in a randomized experiment. Normal/Large Sample: The dotplots below show no strong skewness or outliers in either group.



D: $\bar{x}_1 = 123.8$, $s_1 = 4.60$, $\bar{x}_2 = 116.4$, $s_2 = 16.09$. $t = 0.989$. Using df = 4, the *P*-value is between 0.15 and 0.20. Using df = 4.65, *P*-value = 0.1857. *C*: Because the *P*-value of $0.1857 > \alpha = 0.05$, we fail to reject H_0. We do not have convincing evidence that the true mean breaking strength of polyester fabric that is buried for 2 weeks is greater than the true mean breaking strength for polyester fabric that is buried for 16 weeks.

Feb 8-9:21 PM

Example 4: In each of the following settings, decide whether you should use paired *t* procedures or two-sample *t* procedures to perform inference. Explain your choice.

a) To test the wear characteristics of two tire brands, A and B, one Brand A tire is mounted on one side of each car in the rear, while a Brand B tire is mounted on the other side. Which side gets which brand is determined by flipping a coin.

Paired *t* procedures. This is a matched pairs experiment, with the two treatments (Brand A and Brand B) being randomly assigned to the rear pair of wheels on each car.

b) Can listening to music while working increase productivity? Twenty factory workers agree to take part in a study to investigate this question. Researchers randomly assign 10 workers to do a repetitive task while listening to music and the other 10 workers to do the task in silence.

Two-sample *t* procedures. The data are being produced using two distinct groups of workers in a randomized experiment.

Feb 8-12:29 PM

Feb 8-9:21 PM