

**Freeport AP Statistics**

## Chapter 10 : Comparing Two Populations of Groups

**10.1** Comparing Two Populations**OBJECTIVE(S):**

- Students will learn how to describe the characteristics of the sampling distribution of  $\hat{p}_1 - \hat{p}_2$ .
- Students will learn how to calculate probabilities using the sampling distribution  $\hat{p}_1 - \hat{p}_2$ .
- Students will learn how to determine whether the conditions for performing inference are met.
- Students will learn how to construct and interpret a confidence interval to compare two proportions.
- Students will learn how to perform a significance test to compare two proportions.
- Students will learn how to interpret the results of inference procedures in a randomized experiment.

**The Sampling Distribution of  $\hat{p}_1 - \hat{p}_2$** 

- **Shape**
  - **Center**
  - **Spread**
1. Who is more likely to binge drink – male or female college students? The Harvard School of Public Health surveys random samples of male and female undergraduates at four-year colleges and universities about whether they have engaged in binge drinking.
    - a. Is this a problem about comparing means or comparing proportions? Explain.

- b. What type of study design is being used to produce data?
  
2. A bank wants to know which of two incentive plans will most increase the use of its credit cards. It offers each incentive to a group of current credit card customers, determined at random, and compares the amount charged during the following six months.
  - a. Is this a problem about comparing means or comparing proportions? Explain.
  
  - b. What type of study design is being used to produce data?
  
3. A researcher reports that 80% of high school graduates but only 40% of high school dropouts would pass a basic literacy test. Assume that the researcher's claim is true. Suppose we give a basic literacy test to a random sample of 60 high school graduates and a separate random sample of 75 high school dropouts.
  - a. What is the shape of the sampling distribution of  $\hat{p}_G - \hat{p}_D$ ? Why?
  
  - b. Find the mean of the sampling distribution. Show work.
  
  - c. Find the standard deviation of the sampling distribution. Show work.

Standard Error of the statistic  $\hat{p}_1 - \hat{p}_2$ :

**Two-Sample z Interval for a Difference between Two Proportions**

- **Random**
  
  
  - **Normal**
  
  
  - **Independent**
4. Explain why the conditions for using two-sample z procedures to perform inference about  $p_1 - p_2$  are not met in the following settings.
- a. A study of injuries to in-line skater used data from the national Electronic Injury Surveillance System, which collects data from a random sample of hospital emergency rooms. The researchers interviewed 161 people who came to emergency rooms with injuries from in-line skating. Wrist injuries (mostly fractures) were the most common. The interviews found that 53 people were wearing wrist guards and 6 of these had wrist injuries. Of the 108 who did not wear wrist guards, 45 had wrist injuries.

- b. We don't like to find broken crackers when we open the package. How can makers reduce breaking? One idea is to microwave the crackers for 30 seconds right after baking them. Breaks start as hairline cracks called "checking." Randomly assign 65 newly baked crackers to the microwave and another 65 to a control group that is not micro waved. After one day, none of the microwave group and 16 of the control group show checking.
5. Many news organizations conduct polls asking adults in the United States if they approve of the job the president is doing. How did President Obama's approval rating change from October 2012 to October 2013? According to a Gallup poll of 1500 randomly selected U.S. adults on October 2-4, 2012, 52% approved of Obama's job performance. A Gallup poll of 1500 randomly selected U.S. adults on October 5-7, 2013, showed that 46% approved of Obama's job performance.
- a. Use the results of these polls to construct and interpret a 90% confidence interval for the change in Obama's approval rating among all U.S. adults.

*Parameters of interest*

*Are the conditions met?*

- **Random:**
  
- **Normal:**
  
  
  
  
  
  
  
  
  
  
- **Independent:**

**CHAPTER 10**

*Name of the interval*

*Interval*

*Conclusion*

- b. Based on your interval, is there convincing evidence that Obama's job approval rating changed between October 2012 and October 2013?

6. Is rap music more popular among young blacks than among young whites? A sample survey compared 634 randomly chosen blacks aged 15 to 25 with 567 randomly selected whites in the same age group. It found that 368 of the blacks and 130 of the whites listened to rap music every day. Construct and interpret a 95% confidence interval for the difference between the proportions of black and white young people who listen to rap every day.

*Parameters of interest*

*Are the conditions met?*

- **Random:**

- **Normal:**

- **Independent:**

*Name of the interval*

*Interval*

*Conclusion*

7. The elderly fear crime more than younger people, even though they are less likely to be victims of crime. One study recruited separate random samples of 56 black women and 63 black men over the age of 65 from Atlantic City, New Jersey. Of the women, 27 said they “felt vulnerable” to crime; 46 of the men said this.
  - a. Construct and interpret a 90% confidence interval for the difference in population proportions (men minus women).

*Parameters of interest*

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

**CHAPTER 10**

*Name of the interval*

*Interval*

**Conclusion**

- b. Does your interval from part a. give convincing evidence of a difference between the population proportions? Explain.



**Pooled (or Combined) Sample Proportion –**

**Two-Sample  $z$  Test for the Difference between Two Proportions**

*Conditions:*

- **Random**
- **Normal**
- **Independent**

8. Are teenagers going deaf? In a study of 3000 randomly selected teenagers in 1988-1994, 15% showed some hearing loss. In a similar study of 1800 teenagers in 2005-2006, 19.5% showed some hearing loss. (These data were reported in *Arizona Daily Star*, August 18, 2010.)
- a. Do these data give convincing evidence that the proportion of all teens with hearing loss has increased?

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

*Name of the test*

*Test of statistic*

*Obtain a P-value.*

*P-value =*

*Make a decision about null*

*State your conclusion ( $H_a$  in context of the problem)*

- b. Between the two studies, Apple introduced the iPod. If the results of the test are statistically significant, can we blame iPods for the increased hearing loss in teenagers?

9. In an effort to reduce health care costs, General Motors sponsored a study to help employees stop smoking. In the study, half of the subjects were randomly assigned to receive up to \$750 for quitting smoking for a year while the other half were simply encouraged to use traditional methods to stop smoking. None of the 878 volunteers knew that there was a financial incentive when they signed up. At the end of one year, 15% of those in the financial rewards group had quit smoking while only 5% in the traditional group had quit smoking. Do the results of this study give convincing evidence that a financial incentive helps people quit smoking? (These data are reported in *Arizona Daily Star*, February 11, 2009.)

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

**CHAPTER 10**

*Name of the test*

*Test of statistic*

*Obtain a P-value.*

*P-value =*

*Make a decision about null*

*State your conclusion ( $H_a$  in context of the problem)*

10. Aspirin prevents blood from clotting and so helps prevent strokes. The Second European Stroke Prevention Study asked whether adding another anticlotting drug, named dipyridamole, would be more effective for patients who had already had a stroke. Here are the data on strokes and deaths during the two years of the study:

	<b>Number of patients</b>	<b>Number of strokes</b>
Aspirin alone	1649	206
Aspirin + dipyridamole	1650	157

The study was a randomized comparative experiment.

- a. Is there convincing evidence at the  $\alpha = 0.05$  level that adding dipyridamole helps reduce the risk of stroke?

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**
  
- **Normal:**

- **Independent:**

*Name of the test*

*Test of statistic*

*Obtain a P-value.*

*P-value =*

*Make a decision about null*

State your conclusion ( $H_a$  in context of the problem)

- b. Describe a Type I and a Type II error in this setting. Which is more serious? Explain.

DAY 3

---



**Freeport AP Statistics**

Chapter 10 : Comparing Two Populations of Groups

**10.2 Comparing Two Means****OBJECTIVE(S):**

- Students will learn how to describe the characteristics of the sampling distribution of  $\bar{x}_1 - \bar{x}_2$ .
- Students will learn how to calculate probabilities using the sampling distribution of  $\bar{x}_1 - \bar{x}_2$ .
- Students will learn how to determine whether the conditions for performing inference are met.
- Students will learn how to use two-sample  $t$  procedures to compare two means based on summary statistics.
- Students will learn how to use two-sample  $t$  procedures to compare two means from raw data.
- Students will learn how to interpret standard computer output for two-sample  $t$  procedures.
- Students will learn how to perform a significance test to compare two means.
- Students will learn how to interpret the results of inference procedures in a randomized experiment.
- Students will learn how to determine the proper inference procedure to use in a given setting.

**The Sampling Distribution of  $\bar{x}_1 - \bar{x}_2$** 

- **Shape**
  
- **Center**
  
- **Spread**

11. A potato chip manufacturer buys potatoes from two different suppliers, Riderwood Farms and Camberley, Inc. The weights of potatoes from Riderwood Farms are approximately Normally distributed with a mean of 175 grams and a standard deviation of 25 grams. The weights of potatoes from Camberley are approximately Normally distributed with a mean of 180 grams and a standard deviation of 30 grams. When shipments arrive at the factory, inspectors randomly select a sample of 20 potatoes from each shipment and weigh them. They are surprised when the average weight of the potatoes in the sample from Riderwood Farms  $\bar{x}_r$  is higher than the average weight of the potatoes in the sample from Camberley  $\bar{x}_c$ .

Describe the shape, center, and spread of the sampling distribution  $\bar{x}_c - \bar{x}_r$ .

12. Determine whether or not the conditions for using two-sample t procedures are met.
- How do the numbers of people living in households in the United Kingdom (U.K.) and South Africa compare? To help answer this question, we used Census At School's random data selector to choose independent samples of 50 students from each country. See **p. 654** for the dot plot of the household sizes reported by the students in the survey.

- b. Mary was interested in comparing the mean word length in articles from a medical journal and an airline's in-flight magazine. She counted the number of letters in the first 200 words of an article in the medical journal and in the first 100 words of an article in the airline magazine. Mary then used Minitab statistical software to produce the histograms are shown on **p. 655**.

**Standard Error of  $\bar{x}_1 - \bar{x}_2$ :**

**Two-Sample  $t$  Statistic:**

**Two-Sample  $t$  Interval for a Difference between Two Means**

- **Random**
  
  
  
  
  
  
  
  
  
  
- **Normal**
  
  
  
  
  
  
  
  
  
  
- **Independent**

13. Do plastic bags from Target or plastic bags from Giant Eagle hold more weight? A group of AP Statistics students decided to investigate by filling a random sample of 5 bags from each store with common grocery items until the bags ripped. Then they weighed the contents of items in each bag to determine its capacity. Here are their results, in grams:

<b>Target:</b>	12,572	13,999	11,215	15,447	10,896
<b>Giant Eagle:</b>	9552	10,896	6983	8767	9972

- a. Construct and interpret a 99% confidence interval for the difference in mean capacity of plastic grocery bags from Target and Giant Eagle.

*Parameters of interest*

*Are the conditions met?*

- **Random:**

- **Normal:**

- **Independent:**

**CHAPTER 10**

*Name of the interval*

*Interval*

**Conclusion**

- b. Does your interval provide convincing evidence that there is a difference in the mean capacity between the two stores?

14. Mia and Josi want to know if generic chocolate chip cookies have as many chocolate chips, on average, as Chips Ahoy. To investigate, they randomly select 10 packages of Chips Ahoy and 10 packages of Great Value cookies and randomly select 1 cookie from each pack. They they carefully broke apart each cookie and counted the number of chocolate chips. Here are their results:

Chips Ahoy	17	19	21	16	17	18	20	21	17	18
Great Value	22	20	14	17	21	22	15	19	26	18

- a. Construct and interpret a 99% confidence interval for the difference in the mean number of chocolate chips in Chips Ahoy and Great Value cookies.

*Parameters of interest*

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

*Name of the interval*

*Interval*

## Conclusion

- b. Does the interval provide convincing evidence that there is a difference in the mean number of chocolate chips?
15. As the Hispanic population in the United States has grown, businesses have tried to understand what Hispanics like. One study interviewed a random sample of customers leaving a bank. Customers were classified as Hispanic if they preferred to be interviewed in Spanish or as Anglo if they preferred English. Each customer rated the importance of several aspects of bank service on a 10-point scale. Here are summary results for the importance of “reliability” (the accuracy of account records and so on):

<b>Group</b>	<b><i>n</i></b>	<b><math>\bar{x}</math></b>	<b><math>s_x</math></b>
Anglo	92	6.37	0.60
Hispanic	86	5.91	0.93

- a. The distribution of reliability ratings in each group is not Normal. A graph of the data reveals no outliers. The use of two-sample  $t$  procedures is justified. Why?

- b. Construct and interpret a 95% confidence interval for the difference between the mean ratings of the importance of reliability for Anglo and Hispanic bank customers.

*Parameters of interest*

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

Name of the interval

Interval

Conclusion



- c. Interpret the 95% confidence level in the context of this study?

**Two-Sample  $t$  Test for the Difference between Two Means**

*Conditions:*

- **Random**
  
- **Normal**
  
- **Independent**

16. In commercials for Bounty paper towels, the manufacturer claims that they are the “quicker picker-upper.” But are they also the stronger picker upper? Two AP Statistics students, Wesley and Maverick, decided to find out. They selected a random sample of 30 Bounty paper towels and a random sample of 30 generic paper towels and measured their strength when wet. To do this, they uniformly soaked each paper towel with 4 ounces of water, held two opposite edges of the paper towel, and counted how many quarters each paper towel could hold until ripping, alternating brands. Here are their results:

---

<b>Bounty:</b>	106	111	106	120	103	112	115	125
	116	120	126	125	116	117	114	118
	126	120	115	116	121	113	111	128
	124	125	127	123	115	114		

<b>Generic:</b>	77	103	89	79	88	86	100	90	81
	84	84	96	87	79	90	86	88	81
	91	94	90	89	85	83	89	84	90
	100	94	87						

- 
- a. Use a significance test to determine whether there is convincing evidence that wet Bounty paper towels can hold more weight, on average, than wet generic paper towels can.

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

Name of the test

Test of statistic

Obtain a  $P$ -value.

$P$ -value =

Make a decision about null

State your conclusion ( $H_a$  in context of the problem)

- b. Interpret the  $P$ -value from part a. in the context of the question.

17. Some doctors have begun to use medical magnets to treat patients with chronic pain. Scientists wondered whether this type of therapy really worked. So they designed an experiment to find out. Fifty patients with chronic pain were recruited for the study. A doctor identified a painful site on each patient and asked him or her to rate the pain on a scale from 0 (mild pain) to 10 (severe pain). Then, the doctor selected a sealed envelope containing a magnet from a box that contained both active and inactive magnets. The chosen magnet was applied to the site of the pain for 45 minutes. After treatment, each patient was again asked to rate the level of pain from 0 to 10.

In all, 29 patients were given active magnets and 21 patients received inactive magnets. All but 1 of the patients rated their initial pain as an 8, 9, or 10. So scientists decided to focus on patients' final pain ratings. Here they are, grouped by the type of magnet used:

---

**Active:** 0 4 7 0 4 2 5 5 3 2 2 2 3 4 6 4 3 0 2 0 4 4 5 9 10 10 10 10  
7

**Inactive:** 4 7 5 8 8 6 8 10 10 6 10 8 10 10 10 10 9 9 10 10 9

---

Did the active magnets help reduce pain? Give appropriate statistical evidence to support your answer.

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

Name of the test

Test of statistic

Obtain a  $P$ -value.

$P$ -value =

Make a decision about null

State your conclusion ( $H_a$  in context of the problem)

18. Poisoning by the pesticide DDT causes convulsions in humans and other mammals. Researchers seek to understand how the convulsions are caused. In a randomized comparative experiment, they compared 6 white rats poisoned with DDT with a control group of 6 unpoisoned rats. Electrical measurements of nerve activity are the main clue to the nature of DDT poisoning. When a nerve is stimulated, its electrical response shows a sharp spike followed by a much smaller second spike. The researchers measured the height of the second spike as a percent of the first spike when a nerve in the rat's leg was stimulated. For the poisoned rats the results were

12.207      16.869      25.050      22.429      8.456      20.589

The control group data were

11.074      9.686      12.064      9.351      8.182      6.642

- a. Do these data provide convincing evidence that DDT affects the mean height of the second spike's electrical response? Carry out a significance test to help answer this question.

*Parameters of interest*

*Hypothesis*

$H_0 :$

$H_a :$

Are the conditions met?

- **Random:**

- **Normal:**

- **Independent:**

Name of the test

Test of statistic

Obtain a  $P$ -value.

$P$ -value =

Make a decision about null

State your conclusion ( $H_a$  in context of the problem)



- b. Interpret the  $P$ -value from part a. in the context of this study.
19. 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.
- To compare the average weight gain of pigs fed two different rations, nine pairs of pigs were used. The pigs in each pair were littermates. A coin toss was used to decide which pig in each pair got Ration A and which got Ration B.
  - A random sample of college professors is taken. We wish to compare the average salaries of male and female teachers.
  - To test the effects of a new fertilizer, 100 plots retreated with the new fertilizer, and 100 plots are treated with another fertilizer. A computer's random number generator is used to determine which plots get which fertilizer.