Comparisons Involving Means

- Hypothesis Tests about the Difference between the Means of Two Populations: Independent Samples
- Inferences about the Difference between the Means of Two Populations: Matched Samples

Hypothesis Tests About the Difference between the Means of Two Populations: Independent Samples

- **Hypotheses**

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

$$H_0: \mu_1 - \mu_2 = 0 \qquad H_a: \mu_1 - \mu_2 < 0$$

$$H_0: \mu_1 - \mu_2 = 0 \qquad H_a: \mu_1 - \mu_2 \neq 0$$

- **Test Statistic**

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 / n_1 + \sigma_2^2 / n_2}}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2 (1/n_1 + 1/n_2)}}$$

Example: ABC, Inc.

- **Hypothesis Tests About the Difference between the Means of 2 Populations: Large-Sample Case**
- ABC, Inc. is a manufacturer of golf equipment and has developed a new golf ball that has been designed to provide "extra distance." In a test of driving distance using a mechanical driving device, a sample of ABC golf balls was compared with a sample of golf balls made by XYZ, Ltd., a competitor. The sample statistics appear on the next slide.

Example: ABC, Inc.

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Large-Sample Case**
  - **Sample Statistics**

	Sample #1	**Sample #2**
	ABC, Inc.	**XYZ, Ltd.**
Sample Size	n_1 = 120 balls	n_2 = 80 balls
Mean	$\bar{x}_1$ = 235 yards	$\bar{x}_2$ = 218 yards
Standard Deviation	σ_1 = 15 yards	σ_2 = 20 yards

Example: ABC, Inc.

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Large-Sample Case**

  Can we conclude, using a .01 level of significance, that the mean driving distance of ABC, Inc. golf balls is greater than the mean driving distance of XYZ, Ltd. golf balls?

Example: ABC, Inc.

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Large-Sample Case**

  μ_1 = mean distance for the population of ABC, Inc. golf balls

  μ_2 = mean distance for the population of XYZ, Ltd. golf balls

  - **Hypotheses** H_0: $\mu_1 - \mu_2 = 0$

    H_a: $\mu_1 - \mu_2 > 0$

Example: ABC, Inc.

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Large-Sample Case**
  - **Decision Rule**

Reject H_0 if p-value $< \alpha$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(235 - 218) - 0}{\sqrt{\frac{(15)^2}{120} + \frac{(20)^2}{80}}} = \frac{17}{2.62} = 6.49$$

Since p-value = $P(z > 6.49) \approx 0.000 < 0.01 = \alpha$, we reject H_0.

Example: ABC, Inc.

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Large-Sample Case**
  - **Conclusion**

Reject H_0. We are at least 99% confident that the mean driving distance of ABC, Inc. golf balls is greater than the mean driving distance of XYZ, Ltd. golf balls.

Example: Ali Motors

Ali Motors of Detroit has developed a new automobile known as the M car. 12 M cars and 8 J cars (from Japan) were road tested to compare miles-per- gallon (mpg) performance. The sample data is below.

	Sample #1 **<u>M Cars</u>**	Sample #2 **<u>J Cars</u>**
Sample Size	n_1 = 12 cars	n_2 = 8 cars
Mean	$\bar{x}_1$ = 29.8 mpg	$\bar{x}_2$ = 27.3 mpg
Standard Deviation	s_1 = 2.56 mpg	s_2 = 1.81 mpg

Example: Ali Motors

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Small-Sample Case**

Can we conclude, using a .05 level of significance, that the miles-per-gallon (mpg) performance of M cars is greater than the miles-per-gallon performance of J cars?

Example: Ali Motors

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Small-Sample Case**

  μ_1 = mean mpg for the population of M cars

  μ_2 = mean mpg for the population of J cars

  - Hypotheses H_0: $\mu_1 - \mu_2 = 0$

    H_a: $\mu_1 - \mu_2 > 0$

Example: Ali Motors

- **Hypothesis Tests About the Difference Between the Means of 2 Populations: Small-Sample Case**
  - Decision Rule

    Reject H_0 if p-value < α

    (a = .05, d.f. = 18)

  - Test Statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2(1/n_1 + 1/n_2)}}$$

where:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Example: Ali Motors

Estimate of $\mu_1 - \mu_2 = \bar{x}_1 - \bar{x}_2 = 29.8 - 27.3 = 2.5$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{11(2.56)^2 + 7(1.81)^2}{12 + 8 - 2} = 5.28$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2(1/n_1 + 1/n_2)}} = \frac{2.5 - 0.0}{\sqrt{5.28(\frac{1}{12} + \frac{1}{8})}} = 2.38$$

Since p-value = $P(t_{18} > 2.38) \approx 0.014 < 0.05 = \alpha$, we reject H_0.

We are 95% confident that the miles-per-gallon (mpg) performance of M cars is greater than the miles-per-gallon performance of J cars.

Inference About the Difference between the Means of 2 Populations: Matched Samples

- With a matched-sample design each sampled item provides a pair of data values.
- The matched-sample design can be referred to as blocking.
- This design often leads to a smaller sampling error than the independent-sample design because variation between sampled items is eliminated as a source of sampling error.

Example: Express Deliveries

- **Inference About the Difference between the Means of Two Populations: Matched Samples**

  A Chicago-based firm has documents that must be quickly distributed to district offices throughout the U.S. The firm must decide between two delivery services, UPX (United Parcel Express) and INTEX (International Express), to transport its documents. In testing the delivery times of the two services, the firm sent two reports to a random sample of ten district offices with one report carried by UPX and the other report carried by INTEX.

  Do the data that follow indicate a difference in mean delivery times for the two services?

Example: Express Deliveries

	Delivery Time (Hours)		
District Office	UPX	INTEX	Difference
Seattle	32	25	7
Los Angeles	30	24	6
Boston	19	15	4
Cleveland	16	15	1
New York	15	13	2
Houston	18	15	3
Atlanta	14	15	-1
St. Louis	10	8	2
Milwaukee	7	9	-2
Denver	16	11	5

Example: Express Deliveries

- **Inference About the Difference between the Means of Two Populations: Matched Samples**

  Let μ_d = the mean of the difference values for the two delivery services for the population of district offices

  - Hypotheses

    $$H_0\colon \mu_d = 0, \quad H_a\colon \mu_d \neq 0$$

Example: Express Deliveries

- **Inference About the Difference between the Means of Two Populations: Matched Samples**
  - **Decision Rule**

    Assuming the population of difference values is approximately normally distributed, the t distribution with n - 1 degrees of freedom applies. With $\alpha = .05$, and df = 10-1 = 9, reject H_0 if p-value < α

Example: Express Deliveries

- **Inference About the Difference between the Means of Two Populations: Matched Samples**

$$\bar{d} = \frac{\sum d_i}{n} = \frac{(7+6+\ldots+5)}{10} = 2.7$$

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{76.1}{9}} = 2.9$$

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{2.7 - 0}{2.9/\sqrt{10}} = 2.94$$

P- value = P(t_9 > 2.94) = 2(0.0082) = 0.0164 < 0.05 = α

Example: Express Deliveries

- **Inference About the Difference between the Means of Two Populations: Matched Samples**
  - **Conclusion**

Reject H_0.

There is a significant difference between the mean delivery times for the two services.