Hypothesis Tests for Two Populations

Objective 1: Determining the Appropriate Formula and Calculator Function

	Formula for Test Statistic	Calculator Function for Hypothesis Test
Two Population Proportions	$z_0 = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $\hat{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$	`2-PropZTest`
Two Population Means – Dependent Samples	$t_0 = \dfrac{\bar{d}}{s_d / \sqrt{n}}$	`T-Test`
Two Population Means – Independent Samples	$t_0 = \dfrac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	`2-SampTTest`
Two Population Standard Deviations	$F_0 = \dfrac{s_1^2}{s_2^2}$	`2-SampFTest`

Example: A statistics professor at Cypress College wanted to see if students were performing better in the online statistics class, rather than in the face-to-face statistics classes. To test this, the professor randomly selected 40 students who had taken the online statistics class and found that their mean final grade was a 77.8 with a standard deviation of 12.3. The professor also randomly selected 40 students who had taken the face-to-face statistics class and found that their mean grade was a 73.4 with a standard deviation of 8.5. Determine the appropriate formula and calculator function that would be used to test the professor's claim, at the $\alpha = 0.05$ level of significance, that the mean final grade for students taking the online statistics class is higher than the mean final grade for students taking the face-to-face statistics class.

Example: Over the recent years, it has become somewhat common for pet owners to speak with their pets on the phone. A veterinarian feels that a smaller proportion of cat owners speak with their pet on their phone when compared to dog owners. To test this, the veterinarian randomly surveyed 150 cat owners and found that 42 of them speak with their cat on the phone. Then the veterinarian randomly selected 200 dog owners and found that 68 of them speak with their dog on the phone. Determine the appropriate formula and calculator function that would be used to test the veterinarian's claim at the $\alpha = 0.10$ level of significance.

Example: A doctor wanted to determine if there was a difference in the variability of resting heart rates of people who do not exercise compared to those who do exercise. To test this, the doctor randomly selected 65 people who exercise regularly and found that the mean resting heart rate for these people was 72 beats per minute (bpm) with a standard deviation of 5.8 bpm. The doctor also randomly selected 50 people who do not exercise regularly and found that the mean resting heart rate for these people was 89 bpm with a standard deviation of 7.2 bpm. Determine the appropriate formula and calculator function that would be used to test the claim that there is a difference in the standard deviation resting heart rates for those who do not exercise compared to those who do exercise at the $\alpha = 0.01$ level of significance. Note that the resting pulse rates for both groups are normally distributed.

Example: A doctor was conducting a study to test whether marijuana was effective in reducing pain for patients with chronic back pain. The doctor randomly selected 9 patients with chronic back pain and asked them to rate their level of pain on a scale from 1 to 10 (with a higher score indicating more pain). Then, he had the 9 patients smoke marijuana everyday for a week and asked them to rate their level of pain on a scale from 1 to 10. The results are given below. Determine the appropriate formula and calculator function that would be used to test the doctor's claim that marijuana is effective in reducing pain for patients with chronic back pain at the $\alpha = 0.05$ level of significance.

Patient	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Pain Level Before Smoking Marijuana	6.8	7.6	9.2	8.5	6.7	8.4	9.5	7.3	8.2
Pain Level After Smoking Marijuana	5.2	6.8	9.2	8.5	7.3	8.8	9.2	7.6	7.3

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A college football coach was interested in whether the college's strength development class increased his players' maximum lift (in pounds) on the bench press exercise. He asked four of his players to participate in a study. The amount of weight they could each lift was recorded before they took the strength development class. After completing the class, the amount of weight they could each lift was again measured. The data is given in the table below. Determine the appropriate formula and calculator function that would be used to test the coach's claim that the college's strength development class is increasing the players' maximum lift on the bench press, at the $\alpha = 0.10$ level of significance.

	Player 1	Player 2	Player 3	Player 4
Amount of Weight Lifted Before the Class	205	241	338	368
Amount of Weight Lifted After the Class	295	252	330	360

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A mathematics professor was interested in seeing if calculus students have more consistent final exam grades when compared to statistics students. To test this, the professor randomly selected 10 calculus students and found that the mean final exam grade was a 73.5 with a standard deviation of 13.6 and he randomly selected 15 statistics students and found that the mean final exam grade was 77.8 with a standard deviation of 17.2. Determine the appropriate formula and calculator function that would be used to test the professor's claim that the standard deviation for the final exam grade for calculus students is lower than the standard deviation for the final exam grade for statistics students. Note that the grades for both groups are normally distributed.

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A professor wanted to investigate if there was a difference in the proportion of males and females that text during class. To test this, the professor randomly selected 200 female students and found that 105 of them text during their classes. The professor then randomly selected 225 male students and found that 120 of them text during their classes. Determine the appropriate formula and calculator function that would be used to test the professor's claim at the $\alpha = 0.05$ level of significance.

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: The average entry-level salaries for people with mechanical engineering degrees and electrical engineering degrees are believed to be approximately the same. However, a recruiting office thinks that the mean mechanical engineering salary is actually lower than the mean electrical engineering salary. The recruiting office randomly surveys 50 entry level mechanical engineers and finds that their mean salary is $46,100 with a standard deviation of $3,450. The office also randomly selects 60 entry level electrical engineers and finds that their mean salary is $46,700 with a standard deviation of $4,210. Determine the appropriate formula and calculator function that would be used to test the office's claim at the $\alpha = 0.01$ level of significance.

Objective 2: Constructing the Hypotheses

	Two Tailed Hypotheses	Left Tailed Hypotheses	Right Tailed Hypotheses
Two Population Proportions	$H_0: p_1 = p_2$ $H_1: p_1 \neq p_2$	$H_0: p_1 = p_2$ $H_1: p_1 < p_2$	$H_0: p_1 = p_2$ $H_1: p_1 > p_2$
Two Population Means – Dependent Samples	$H_0: \mu_d = 0$ $H_1: \mu_d \neq 0$	$H_0: \mu_d = 0$ $H_1: \mu_d < 0$	$H_0: \mu_d = 0$ $H_1: \mu_d > 0$
Two Population Means – Independent Samples	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$
Two Population Standard Deviations	$H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 \neq \sigma_2$	$H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 < \sigma_2$	$H_0: \sigma_1 = \sigma_2$ $H_1: \sigma_1 > \sigma_2$

Example: A statistics professor at Cypress College wanted to see if students were performing better in the online statistics class, rather than in the face-to-face statistics classes. To test this, the professor randomly selected 40 students who had taken the online statistics class and found that their mean final grade was a 77.8 with a standard deviation of 12.3. The professor also randomly selected 40 students who had taken the face-to-face statistics class and found that their mean grade was a 73.4 with a standard deviation of 8.5. Set up the appropriate hypotheses for the professor's claim, at the $\alpha = 0.05$ level of significance, that the mean final grade for students taking the online statistics class is higher than the mean final grade for students taking the face-to-face statistics class. Let population 1 represent online students and population 2 represent face-to-face students.

Example: Over the recent years, it has become somewhat common for pet owners to speak with their pets on the phone. A veterinarian feels that a smaller proportion of cat owners speak with their pet on their phone when compared to dog owners. To test this, the veterinarian randomly surveyed 150 cat owners and found that 42 of them speak with their cat on the phone. Then the veterinarian randomly selected 200 dog owners and found that 68 of them speak with their dog on the phone. Set up the appropriate hypotheses for the veterinarian's claim at the $\alpha = 0.10$ level of significance. Let population 1 represent cat owners and let population 2 represent dog owners.

Example: A doctor wanted to determine if there was a difference in the variability of resting heart rates of people who do not exercise compared to those who do exercise. To test this, the doctor randomly selected 65 people who exercise regularly and found that the mean resting heart rate for these people was 72 beats per minute (bpm) with a standard deviation of 5.8 bpm. The doctor also randomly selected 50 people who do not exercise regularly and found that the mean resting heart rate for these people to be 89 bpm with a standard deviation of 7.2 bpm. Set up the appropriate hypotheses for the claim that there is a difference in the standard deviation resting heart rates for those who do not exercise compared to those who do exercise at the $\alpha = 0.01$ level of significance. Note that the resting pulse rates for both groups are normally distributed. Let population 1 represent those who do exercise regularly and population 2 represent those who do not exercise regularly.

Example: A doctor was conducting a study to test whether marijuana was effective in reducing pain for patients with chronic back pain. The doctor randomly selected 9 patients with chronic back pain and asked them to rate their level of pain on a scale from 1 to 10 (with a higher score indicating more pain). Then, he had the 9 patients smoke marijuana every day for a week and asked them to rate their level of pain on a scale from 1 to 10. The results are given below. Set up the appropriate hypotheses for the doctor's claim that marijuana is effective in reducing pain for patients with chronic back pain at the $\alpha = 0.05$ level of significance. Let L1 = Pain Level Before Smoking, L2 = Pain Level After Smoking, and L3 = L1 – L2.

Patient	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
Pain Level Before Smoking Marijuana	6.8	7.6	9.2	8.5	6.7	8.4	9.5	7.3	8.2
Pain Level After Smoking Marijuana	5.2	6.8	9.2	8.5	7.3	8.8	9.2	7.6	7.3

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A college football coach was interested in whether the college's strength development class increased his players' maximum lift (in pounds) on the bench press exercise. He asked four of his players to participate in a study. The amount of weight they could each lift was recorded before they took the strength development class. After completing the class, the amount of weight they could each lift was again measured. The data is given in the table below. Set up the appropriate hypotheses for the coach's claim that the college's strength development class is increasing the players' maximum lift on the bench press, at the $\alpha = 0.10$ level of significance. Let L1 = Amount of Weight Lifted Before Class, L2 = Amount of Weight Lifted After Class, and L3 = L1 – L2.

	Player 1	**Player 2**	**Player 3**	**Player 4**
Amount of Weight Lifted Before the Class	205	241	338	368
Amount of Weight Lifted After the Class	295	252	330	360

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A mathematics professor was interested in seeing if calculus students have more consistent final exam grades when compared to statistics students. To test this, the professor randomly selected 10 calculus students and found that the mean final exam grade was a 73.5 with a standard deviation of 13.6 and he randomly selected 15 statistics students and found that the mean final exam grade was 77.8 with a standard deviation of 17.2. Set up the appropriate hypotheses for the professor's claim that the standard deviation for the final exam grade for calculus students is lower than the standard deviation for the final exam grade for statistics students. Note that the grades for both groups are normally distributed. Let population 1 be calculus students and population 2 be statistics students.

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: A professor wanted to investigate if there was a difference in the proportion of males and females that text during class. To test this, the professor randomly selected 200 female students and found that 105 of them text during their classes. The professor then randomly selected 225 male students and found that 120 of them text during their classes. Set up the appropriate hypotheses for the professor's claim at the $\alpha = 0.05$ level of significance.

Pause the video, and try this problem. Resume the video to show the answer to the problem.

Example: The average entry-level salaries for people with mechanical engineering degrees and electrical engineering degrees are believed to be approximately the same. However, a recruiting office thinks that the mean mechanical engineering salary is actually higher than the mean electrical engineering salary. The recruiting office randomly surveys 50 entry level mechanical engineers and finds that their mean salary is \$47,100 with a standard deviation of \$3,450. The office also randomly selects 60 entry level electrical engineers and finds that their mean salary is \$46,700 with a standard deviation of \$4,210. Set up the appropriate hypotheses for the office's claim at the $\alpha = 0.01$ level of significance. Let population 1 be those with mechanical engineering degrees and population 2 be those with electrical engineering degrees.