



Probability & Statistics Workbook

Hypothesis testing

INFERENCEAL STATISTICS AND HYPOTHESES

- 1. A current pain reliever has an 85 % success rate of treating pain. A company develops a new pain reliever and wants to show that its success rate of treating pain is better than the current option. Decide if the hypothesis statement would require a population proportion or a population mean, then set up the statistical hypothesis statements for the situation.

- 2. A research study on people who quit smoking wants to show that the average number of attempts to quit before a smoker is successful is less than 3.5 attempts. How should they set up their hypothesis statements?

- 3. A factory creates a small metal cylindrical part that later becomes part of a car engine. Because of variations in the process of manufacturing, the diameters are not always identical. The machine was calibrated to create cylinders with an average diameter of $\frac{1}{16}$ of an inch. During a periodic inspection, it became clear that further investigation was needed to determine whether or not the machine responsible for making the part needed recalibration. Write statistical hypothesis statements.

- 4. A marketing study for a clothing company concluded that the mean percentage increase in sales could potentially be over 17 % for creating a



clothing line that focused on lime green and polka dots. Which hypothesis statements do they need to write in order to test their theory?

■ 5. A food company wants to ensure that less than 0.0001 % of its product is contaminated. Which hypothesis statements will it write if it wants to test for this?

■ 6. A new medication is being developed to prevent heart worms in dogs, and the developer wants it to work better than the current medication. The current medication prevents heart worms at a rate of 75 % . What hypothesis statements should they write if they want to test whether or not the new medication works better than the existing one?



SIGNIFICANCE LEVEL AND TYPE I AND II ERRORS

- 1. You're running a statistical test on a new pharmaceutical drug. The stakes are high, because the side effects of the drug could potentially be serious, or even fatal. If you want to reduce the Type I and Type II error rates as low as possible to avoid rejecting the null when it's true or accepting the null when it's false, what should you do when you take the sample for your test?

- 2. If the probability of making a Type II error in a statistical test is 5 %, what is the power of the test?

- 3. On average, professional golfers make 75 % of putts within 5 feet. One golfer believes he does better than this, and wants to use a statistical test to see whether or not he's correct. Unbeknownst to him, in actuality this golfer makes 7 out of 10 of these kinds of putts. When he takes a sample of his putts, he finds $\hat{p} = 0.92$. What kind of error might he be in danger of making?

- 4. The average age of a guest at an amusement park is 15 years old. One amusement park believes the average age of their guests is younger than this, and wants to use a statistical test to see whether or not they're correct. Unbeknownst to them, in actuality the average guest age at this



particular amusement park is 12 years old. When they take a sample of his guests, they find $\bar{x} = 16$ years. What kind of error might they be in danger of making?

■ 5. Of all political donations, 70 % come from corporations and lobbies, not from individual citizens. One politician believes he receives less than 70 % of his own donations from corporations and lobbies, and wants to use a statistical test to see whether or not he's correct. Unbeknownst to him, in actuality the proportion of his donations that come from corporations and lobbies is 65 %. When he takes a sample of his donations that come from corporations and lobbies, he finds $\hat{p} = 0.72$. What kind of error might he be in danger of making?

■ 6. A coffee shop owner believes that he sells 500 cups of coffee each day, on average, and he wants to test this assumption. The truth is, he actually sells fewer than 500 cups each day. He takes a random sample of 10 days and records the number of cups he sells each of those days. What kind of error is the coffee shop owner in danger of making?

Day	1	2	3	4	5	6	7	8	9	10
Cups sold	488	502	496	506	492	489	510	511	506	500



TEST STATISTICS FOR ONE- AND TWO-TAILED TESTS

- 1. A local high school states that its students perform much better than average on a state exam. The average score for all high school students in the state is 106 points. A sample of 256 students at this particular school had an average test score of 129 points with a sample standard deviation of 26.8. Choose and calculate the appropriate test statistic.
- 2. A dietician is looking into the claim at a local restaurant that the number of calories in its portion sizes is lower than the national average. The national average is 1,500 calories per meal. She samples 35 meals at the restaurant and finds they contain an average of 1,250 calories per meal with a sample standard deviation of 350.2. Choose and calculate the appropriate test statistic.
- 3. In a recent survey, 567 out of a 768 randomly selected dog owners said they used a kennel that was run by their veterinary office to board their dogs while they were away on vacation. The study would like to make a conclusion that the majority (more than 50 %) of dog owners use a kennel run by their veterinary office when the owners go on vacation. Choose and calculate the appropriate test statistic.



■ 4. You want to open a day care center, so you take a random sample of 500 households in your town with children under preschool age, and find that 243 of them were using a family member to care for those children. You want to determine if, at a statistically significant level, fewer than half of households in your town are using a family member to care for the kids.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.

■ 5. The highest allowable amount of bromate in drinking water is 0.0100 mg/L^2 . A survey of a city's water quality took 50 water samples in random locations around the city and found an average of 0.0102 mg/L^2 of bromate with a sample standard deviation of 0.0025 mg/L . The survey committee is interested in testing if the amount of bromate found in the water samples is higher than the allowable amount at a statistically significant level.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.



■ 6. A farmer reads a study that states: The average weight of a day-old chick upon hatching is $\mu_0 = 38.60$ grams with a population standard deviation of $\sigma = 5.7$ grams. The farmer wants to see if her day-old chicks have the same average. She takes a simple random sample of 60 of her day-old chicks and finds their average weight is $\bar{x} = 39.1$ grams.

1. Set up the hypothesis statements.
2. Check that the conditions for normality are met.
3. State the type of test: upper-tailed, lower-tailed, or two-tailed.
4. Calculate the test statistic using the appropriate formula.



THE P-VALUE AND REJECTING THE NULL

- 1. A medical trial is conducted to test whether or not a new medicine reduces total cholesterol, when the national average is 230 mg/dL with a standard deviation of 16 mg/dL. The trial takes a simple random sample of 223 adults who take the new medicine, and finds $\bar{x} = 227$ mg/dL with standard error $\sigma_{\bar{x}} = 14$ mg/dL. What can the trial conclude at a significance level of $\alpha = 0.01$?
- 2. The national average length of pregnancy is 283.6 days with a population standard deviation of 10.5 days. A hospital wants to know if the average length of a pregnancy at their hospital deviates from the national average. They use a sample of 9,411 births at the hospital to calculate a test statistic of $z = -1.60$. Set up the hypothesis statements and find the p -value.
- 3. The highest allowable amount of bromate in drinking water is 0.0100 (mg/L)^2 . A survey of a city's water quality took 31 water samples in random locations around the city and used the data to calculate a test statistic of $t = 2.04$. The city wants to know if the amount of bromate in their drinking water is too high. Set up the hypothesis statements and determine the type of test, then find the p -value.



■ 4. A paint company produces glow in the dark paint with an advertised glow time of 15 min. A painter is interested in finding out if the product behaves worse than advertised. She sets up her hypothesis statements as $H_0 : \mu \geq 15$ and $H_a : \mu < 15$, then calculates a test statistic of $z = -2.30$. What would be the conclusions of her hypothesis test at significance levels of $\alpha = 0.05$, $\alpha = 0.01$, and $\alpha = 0.001$?

■ 5. An article reports that the average wasted time by an employee is 125 minutes every day. A manager takes a small random sample of 16 employees and monitors their wasted time, calculating that average wasted time for her employees is 118 minutes with a standard deviation of 28.7 minutes. She wants to know if 118 minutes is below average at a significance level of $\alpha = 0.05$. She assumes the population is normally distributed.

1. State the population parameter and whether you should use a t -test or z -test statistic.
2. Check that the conditions for performing the statistical test are met.
3. Set up the hypothesis statements.
4. State the type of test: upper-tailed, lower-tailed, or two-tailed.
5. Calculate the test statistic using the appropriate formula.
6. Calculate the p -value.



7. Compare the p -value to the significance level and draw a conclusion.

■ 6. We want to test if college students take fewer than 5 years to graduate, on average, so we take a simple random sample of 36 students and record their years to graduate. For the sample, $\bar{x} = 4.9$ and $s = 0.5$. What can we conclude at 90% confidence?



HYPOTHESIS TESTING FOR THE POPULATION PROPORTION

- 1. A large electric company claims that at least 80 % of the company's 1,000,000 customers are very satisfied. Using a simple random sample, 100 customers were surveyed and 73 % of the participants were very satisfied. Based on these results, should we use a one- or two- tailed test, and should we accept or reject the company's hypothesis? Assume a significance level of 0.05.

- 2. A university is conducting a statistical test to determine whether the percentage of its students who live on its campus is above the national average of 64 % . They've calculated the test statistic to be $z = 1.40$. Set up hypothesis statements and find the p -value.

- 3. A report claims that 60 % of American families take fewer than 6 months to purchase a home, from the time they start looking to the time they make their first offer. A realtor wants to know if her clients purchase at the same rate, so she takes a simple random sample of 50 of her clients and finds a sample proportion $\hat{p} = 0.64$ and standard error $\sigma_{\hat{p}} = 0.05$. What can she conclude with 90 % confidence?

- 4. A gambler wins 48 % of the hands he plays, but he feels like he's on a losing streak recently, winning fewer hands than normal. He takes a



random sample of 40 of his recent hands, and finds the proportion of winning hands in the sample to be $\hat{p} = 0.45$ with $\sigma_{\hat{p}} = 0.02$. What can he conclude with 90 % confidence?

■ 5. A study claims that the proportion of new homeowners who purchase an internet subscription plan is 0.92. You take a random sample of 140 new homeowners to test this claim, and find $\hat{p} = 0.9$ with $\sigma_{\hat{p}} = 0.015$. What can you conclude at a significance level of $\alpha = 0.05$?

■ 6. A recent study reported that the 15.3 % of patients who are admitted to the hospital with a heart attack die within 30 days of admission. The same study reported that 16.7 % of the 3,153 patients who went to the hospital with a heart attack died within 30 days of admission when the lead cardiologist was away.

Is there enough evidence to conclude that the percentage of patients who die when the lead cardiologist is away is any different than when they're present? Make conclusions at significance levels of $\alpha = 0.05$ and $\alpha = 0.01$.

1. State the population parameter and whether you should use a t -test or z -test statistic.
2. Check that the conditions for performing the statistical test are met.
3. Set up the hypothesis statements.



4. State the type of test: upper-tailed, lower-tailed, or two-tailed.
5. Calculate the test statistic using the appropriate formula.
6. Calculate the p -value.
7. Compare the p -value to the significance level and draw a conclusion.



