9.5: Full Hypothesis Test Examples

Exercise 9.5.1

Jeffrey, as an eight-year old, **established a mean time of 16.43 seconds** for swimming the 25-yard freestyle, with a **standard deviation of 0.8 seconds**. His dad, Frank, thought that Jeffrey could swim the 25-yard freestyle faster using goggles. Frank bought Jeffrey a new pair of expensive goggles and timed Jeffrey for **15 25-yard freestyle swims**. For the 15 swims, **Jeffrey's mean time was 16 seconds. Frank thought that the goggles helped Jeffrey to swim faster than the 16.43 seconds.** Conduct a hypothesis test using a preset α = 0.05. Assume that the swim times for the 25-yard freestyle are normal.

Answer

Set up the Hypothesis Test:

Since the problem is about a mean, this is a **test of a single population mean**.

$H_0 : \mu = 16.43 \quad H_a : \mu < 16.43$

For Jeffrey to swim faster, his time will be less than 16.43 seconds. The "<" tells you this is left-tailed.

Determine the distribution needed:

Random variable: $\bar{X} =$ the mean time to swim the 25-yard freestyle.

Distribution for the test: $\bar{X}$ is normal (population **standard deviation** is known: $\sigma = 0.8$)

$\bar{X} \sim N\left(\mu, \frac{\sigma_X}{\sqrt{n}}\right)$ Therefore, $\bar{X} \sim N\left(16.43, \frac{0.8}{\sqrt{15}}\right)$

$\mu = 16.43$ comes from H_0 and not the data. $\sigma = 0.8$, and $n = 15$.

Calculate the p-value using the normal distribution for a mean:

$p\text{-value} = P(\bar{x} < 16) = 0.0187$ where the sample mean in the problem is given as 16.

$p\text{-value} = 0.0187$(This is called the **actual level of significance**.) The p-value is the area to the left of the sample mean is given as 16.

Figure 9.5.1

$\mu = 16.43$ comes from H_0. Our assumption is $\mu = 16.43$.

Interpretation of the p-value: If H_0 is true, there is a 0.0187 probability (1.87%)that Jeffrey's mean time to swim the 25-yard freestyle is 16 seconds or less. Because a 1.87%chance is small, the mean time of 16 seconds or less is unlikely to have happened randomly. It is a rare event.

Compare α and the p-value:

$$\alpha = 0.05 p\text{-value } = 0.0187 \alpha > p\text{-value} \tag{9.5.1}$$

Make a decision: Since $\alpha > p$ -value, reject H_0.

This indicates that you reject the null hypothesis that the mean time to swim the 25-yard freestyle is at least 16.43 seconds.

Conclusion: At the 5% significance level, there is sufficient evidence that Jeffrey's mean time to swim the 25-yard freestyle is less than 16.43 seconds. Thus, based on the sample data, we conclude that Jeffrey swims faster using the new goggles.

The Type I and Type II errors for this problem are as follows:

The Type I error is to conclude that Jeffrey swims the 25-yard freestyle, on average, in less than 16.43 seconds when, in fact, he actually swims the 25-yard freestyle, on average, in at least 16.43 seconds. (Reject the null hypothesis when the null hypothesis is true.)

The Type II error is that there is not evidence to conclude that Jeffrey swims the 25-yard freestyle, on average, in less than 16.43 seconds when, in fact, he actually does swim the 25-yard free-style, on average, in less than 16.43 seconds. (Do not reject the null hypothesis when the null hypothesis is false.)

Try It 9.5.1

The mean throwing distance of a football for Marco, a high school quarterback, is 40 yards, with a standard deviation of two yards. The team coach tells Marco to adjust his grip to get more distance. The coach records the distances for 20 throws. For the 20 throws, Marco's mean distance was 45 yards. The coach thought the different grip helped Marco throw farther than 40 yards. Conduct a hypothesis test using a preset $\alpha = 0.05$. Assume the throw distances for footballs are normal.

First, determine what type of test this is, set up the hypothesis test, find the *p*-value, sketch the graph, and state your conclusion.

Exercise 9.5.2

Jasmine has just begun her new job on the sales force of a very competitive company. In a sample of 16 sales calls it was found that she closed the contract for an average value of 108 dollars with a standard deviation of 12 dollars. Company policy requires that new members of the sales force must exceed an average of $100 per contract during the trial employment period. Can we conclude that Jasmine has met this requirement at the significance level of 95%?

Answer

1. $H_0 : \mu \leq 100$

$H_a : \mu > 100$

The null and alternative hypothesis are for the parameter μ because the number of dollars of the contracts is a continuous random variable. Also, this is a one-tailed test because the company has only an interested if the number of dollars per contact is below a particular number not "too high" a number. This can be thought of as making a claim that the requirement is being met and thus the claim is in the alternative hypothesis.

2. Test statistic: $t_c = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{108 - 100}{\left(\frac{12}{\sqrt{16}}\right)} = 2.67$

3. Critical value: $t_a = 1.753$ with $\mathrm{n} - 1$ degrees of freedom $= 15$

The test statistic is a Student's t because the sample size is below 30; therefore, we cannot use the normal distribution. Comparing the calculated value of the test statistic and the critical value of $t\,(t_a)$ t a 5\% significance level, we see that the calculated value is in the tail of the distribution. Thus, we conclude that 108 dollars per contract is significantly larger than the hypothesized value of 100 and hus we cannot accept the null hypothesis. There is evidence that supports Jasmine's performance meets company standards.

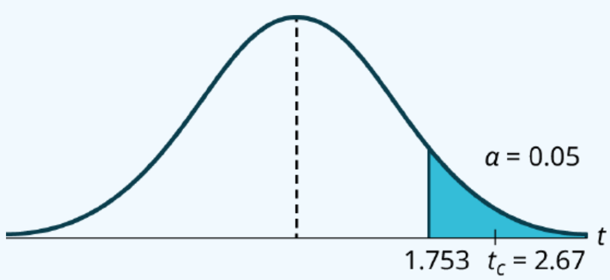

Figure 9.5.2

Try It 9.5.2

It is believed that a stock price for a particular company will grow at a rate of \$5 per week with a standard deviation of \$1. An investor believes the stock won't grow as quickly. The changes in stock price is recorded for ten weeks and are as follows: \$4, \$3, \$2, \$3, \$1, \$7, \$2, \$1, \$1, \$2. Perform a hypothesis test using a 5% level of significance. State the null and alternative hypotheses, state your conclusion, and identify the Type I errors.

Exercise 9.5.3

Problem

A manufacturer of salad dressings uses machines to dispense liquid ingredients into bottles that move along a filling line. The machine that dispenses salad dressings is working properly when 8 ounces are dispensed. Suppose that the average amount dispensed in a particular sample of 35 bottles is 7.91 ounces with a variance of 0.03 ounces squared, s^2. Is there evidence that the machine should be stopped and production wait for repairs? The lost production from a shutdown is potentially so great that management feels that the level of significance in the analysis should be 99%.

Again we will follow the steps in our analysis of this problem.

Answer

STEP 1: Set the Null and Alternative Hypothesis. The random variable is the quantity of fluid placed in the bottles. This is a continuous random variable and the parameter we are interested in is the mean. Our hypothesis therefore is about the mean. In this case we are concerned that the machine is not filling properly. From what we are told it does not matter if the machine is over-filling or under-filling, both seem to be an equally bad error. This tells us that this is a two-tailed test: if the machine is malfunctioning it will be shutdown regardless if it is from over-filling or under-filling. The null and alternative hypotheses are thus:

$$\begin{aligned} H_0 &: \mu = 8 \\ H_a &: \mu \neq 8 \end{aligned} \tag{9.5.2}$$

STEP 2: Decide the level of significance and draw the graph showing the critical value.

This problem has already set the level of significance at 99%. The decision seems an appropriate one and shows the thought process when setting the significance level. Management wants to be very certain, as certain as probability will allow, that they are not shutting down a machine that is not in need of repair. To draw the distribution and the critical value, we need to know which distribution to use. Because this is a continuous random variable and we are interested in the mean, and the sample size is greater than 30 , the appropriate distribution is the normal distribution and the relevant critical value is 2.575 from the normal table or the t-table at 0.005 column and infinite degrees of freedom. We draw the graph and mark these points.

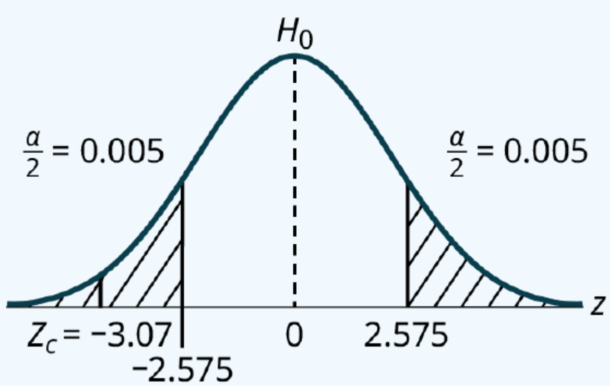

Figure 9.5.3

STEP 3: Calculate sample parameters and the test statistic. The sample parameters are provided, the sample mean is 7.91 and the sample variance is .03 and the sample size is 35 . We need to note that the sample variance was provided not the sample standard deviation, which is what we need for the formula. Remembering that the standard deviation is simply the square root of the variance, we therefore know the sample standard deviation, s , is 0.173 . With this information we calculate the test statistic as -3.07, and mark it on the graph.

$$Z_c=\frac{\bar{x}-\mu_0}{s/\sqrt{n}}=\frac{7.91-8}{.173/\sqrt{35}}=-3.07 \tag{9.5.3}$$

STEP 4: Compare test statistic and the critical values Now we compare the test statistic and the critical value by placing the test statistic on the graph. We see that the test statistic is in the tail, decidedly greater than the critical value of 2.575 . We note that even the very small difference between the hypothesized value and the sample value is still a large number of standard deviations. The sample mean is only 0.08 ounces different from the required level of 8 ounces, but it is 3 plus standard deviations away and thus we cannot accept the null hypothesis.

STEP 5: Reach a Conclusion
Three standard deviations of a test statistic will guarantee that the test will fail. The probability that anything is within three standard deviations is almost zero. Actually it is 0.0026 on the normal distribution, which is certainly almost zero in a practical sense. Our formal conclusion would be " At a 99% level of significance we cannot accept the hypothesis that the sample mean came from a distribution with a mean of 8 ounces" Or less formally, and getting to the point, "At a 99\% level of significance we conclude that the machine is under filling the bottles and is in need of repair".

Try It $9.5.3$

A company records the mean time of employees working in a day. The mean comes out to be 475 minutes, with a standard deviation of 45 minutes. A manager recorded times of 20 employees. The times of working were (frequencies are in parentheses) 460(3); 465(2); 470(3); 475(1); 480(6); 485(3); 490(2).

Conduct a hypothesis test using a 2.5% level of significance to determine if the mean time is *more than 475*.

Hypothesis Test for Proportions

Just as there were confidence intervals for proportions, or more formally, the population parameter p of the binomial distribution, there is the ability to test hypotheses concerning p.

The population parameter for the binomial is p. The estimated value (point estimate) for p is p' where $p'=x/n$, x is the number of successes in the sample and n is the sample size.

When you perform a hypothesis test of a population proportion p, you take a simple random sample from the population. The conditions for a binomial distribution must be met, which are: there are a certain number n of independent trials meaning random sampling, the outcomes of any trial are binary, success or failure, and each trial has the same probability of a success p. The shape of the binomial distribution needs to be similar to the shape of the normal distribution. To ensure this, the quantities np' and nq' must both be greater than five ($np'>5$ and $nq'>5$). In this case the binomial distribution of a sample (estimated) proportion can be approximated by the normal distribution with $\mu=\mathrm{np}$ and $\sigma=\sqrt{\mathrm{npq}}$. Remember that $q=1-p$. There is no distribution that can correct for this small sample bias and thus if these conditions are not met we simply cannot test the hypothesis with the data available at that time. We met this condition when we first were estimating confidence intervals for p.

Again, we begin with the standardizing formula modified because this is the distribution of a binomial.

$$Z=\frac{\mathrm{p}'-p}{\sqrt{\frac{\mathrm{pq}}{n}}} \tag{9.5.4}$$

Substituting p_0, the hypothesized value of p, we have:

$$Z_c=\frac{\mathrm{p}'-p_0}{\sqrt{\frac{p_0q_0}{n}}} \tag{9.5.5}$$

This is the test statistic for testing hypothesized values of p, where the null and alternative hypotheses take one of the following forms:

Table 9.5.1

Two-tailed test	One-tailed test	One-tailed test
$H_0 : p = p_0$	$H_0 : p \leq p_0$	$H_0 : p \geq p_0$
$H_a : p \neq p_0$	$H_a : p > p_0$	$H_a : p < p_0$

The decision rule stated above applies here also: if the calculated value of Z_c shows that the sample proportion is "too many" standard deviations from the hypothesized proportion, the null hypothesis cannot be accepted. The decision as to what is "too many" is pre-determined by the analyst depending on the level of significance required in the test.

? Exercise 9.5.4

Problem

The mortgage department of a large bank is interested in the nature of loans of first-time borrowers. This information will be used to tailor their marketing strategy. They believe that 50% of first-time borrowers take out smaller loans than other borrowers. They perform a hypothesis test to determine if the percentage is **the same or different from 50%**. They sample **100 first-time borrowers** and find **53** of these loans are smaller that the other borrowers. For the hypothesis test, they choose a 5% level of significance.

Answer

STEP 1: Set the null and alternative hypothesis.

$H_0 : p = 0.50 \quad H_a : p \neq 0.50$

The words "is the same or different from" tell you this is a two-tailed test. The Type I and Type II errors are as follows: The Type I error is to conclude that the proportion of borrowers is different from 50% when, in fact, the proportion is actually 50%. (Reject the null hypothesis when the null hypothesis is true). The Type II error is there is not enough evidence to conclude that the proportion of first time borrowers differs from 50% when, in fact, the proportion does differ from 50%. (You fail to reject the null hypothesis when the null hypothesis is false.)

STEP 2: Decide the level of significance and draw the graph showing the critical value
The level of significance has been set by the problem at the 5% level. Because this is two-tailed test one-half of the alpha value will be in the upper tail and one-half in the lower tail as shown on the graph. The critical value for the normal distribution at the 95% level of confidence is 1.96 . This can easily be found on the student's t-table at the very bottom at infinite degrees of freedom remembering that at infinity the t-distribution is the normal distribution. Of course the value can also be found on the normal table but you have go looking for one-half of 95 (0.475) inside the body of the table and then read out to the sides and top for the number of standard deviations.

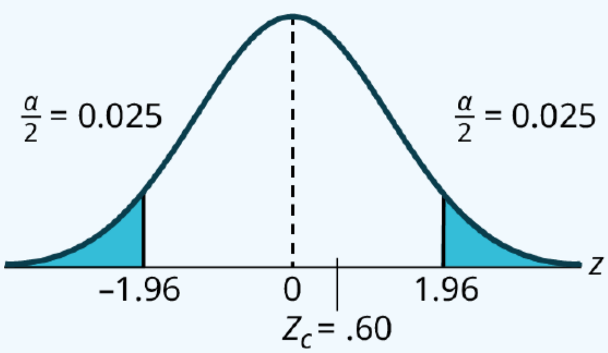

Figure 9.5.4

STEP 3: Calculate the sample parameters and critical value of the test statistic.
The test statistic is a normal distribution, Z, for testing proportions and is:

$$Z=\frac{\mathrm{p}'-p_0}{\sqrt{\frac{p_0q_0}{n}}}=\frac{.53-.50}{\sqrt{\frac{.5(.5)}{100}}}=0.60 \tag{9.5.6}$$

For this case, the sample of 100 found 53 of these loans were smaller than those of other borrowers. The sample proportion, $\mathrm{p}'=53/100=0.53$ The test question, therefore, is : "Is 0.53 significantly different from .50 ?" Putting these values into the formula for the test statistic we find that 0.53 is only 0.60 standard deviations away from .50. This is barely off of the mean of the standard normal distribution of zero. There is virtually no difference from the sample proportion and the hypothesized proportion in terms of standard deviations.

STEP 4: Compare the test statistic and the critical value.
The calculated value is well within the critical values of ± 1.96 standard deviations and thus we cannot reject the null hypothesis. To reject the null hypothesis we need significant evident of difference between the hypothesized value and the sample value. In this case the sample value is very nearly the same as the hypothesized value measured in terms of standard deviations.

STEP 5: Reach a conclusion
The formal conclusion would be "At a 5% level of significance we cannot reject the null hypothesis that 50% of first-time borrowers take out smaller loans than other borrowers." Notice the length to which the conclusion goes to include all of the conditions that are attached to the conclusion. Statisticians, for all the criticism they receive, are careful to be very specific even when this seems trivial. Statisticians cannot say more than they know, and the data constrain the conclusion to be within the metes and bounds of the data.

Try It 9.5.4

A teacher believes that 85% of students in the class will want to go on a field trip to the local zoo. The teacher performs a hypothesis test to determine if the percentage is the same or different from 85%. The teacher samples 50 students and 39 reply that they would want to go to the zoo. For the hypothesis test, use a 1% level of significance.

Exercise 9.5.5

Suppose a consumer group suspects that the proportion of households that have three or more cell phones is 30%. A cell phone company has reason to believe that the proportion is not 30%. Before they start a big advertising campaign, they conduct a hypothesis test. Their marketing people survey 150 households with the result that 43 of the households have three or more cell phones.

Answer

Here is an abbreviate version of the system to solve hypothesis tests applied to a test on a proportions.

$$\begin{gathered}H_0: p=0.3\\ H_a: p\neq 0.3\\ n=150\\ \mathrm{p}'=\frac{x}{n}=\frac{43}{150}=0.287\\ Z_c=\frac{\mathrm{p}'-p_0}{\sqrt{\frac{p_0q_0}{n}}}=\frac{0.287-0.3}{\sqrt{\frac{3(.7)}{150}}}=0.347\end{gathered} \tag{9.5.7}$$

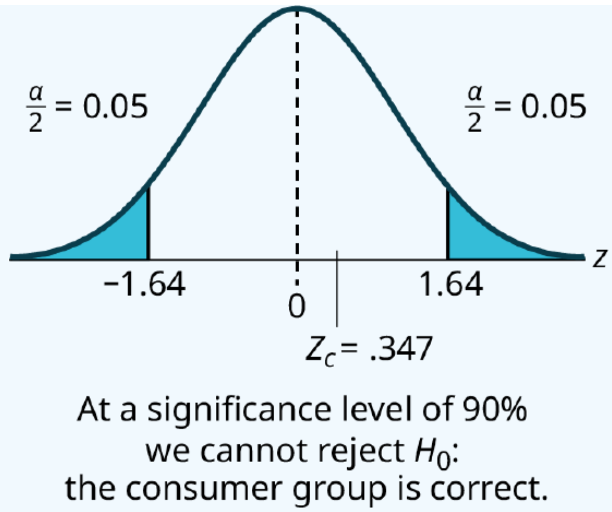

Figure 9.5.5

Try It 9.5.5

Marketers believe that 92% of adults in the United States own a cell phone. A cell phone manufacturer believes that number is actually lower. 200 American adults are surveyed, of which, 174 report having cell phones. Use a 5% level of significance. State the null and alternative hypothesis, find the *p*-value, state your conclusion, and identify the Type I and Type II errors.

Exercise 9.5.6

The National Institute of Standards and Technology provides exact data on conductivity properties of materials. Following are conductivity measurements for 11 randomly selected pieces of a particular type of glass.

1.11; 1.07; 1.11; 1.07; 1.12; 1.08; .98; .98; 1.02; .95; .95
Is there convincing evidence that the average conductivity of this type of glass is greater than one? Use a significance level of 0.05.

Answer

Let's follow a four-step process to answer this statistical question.

1. **State the Question**: We need to determine if, at a 0.05 significance level, the average conductivity of the selected glass is greater than one. Our hypotheses will be

   a. $H_0 : \mu \leq 1$
   b. $H_a : \mu > 1$

2. **Plan**: We are testing a sample mean without a known population standard deviation with less than 30 observations. Therefore, we need to use a Student's-t distribution. Assume the underlying population is normal.

3. **Do the calculations and draw the graph.**

4. **State the Conclusions**: We cannot accept the null hypothesis. It is reasonable to state that the data supports the claim that the average conductivity level is greater than one.

Try It 9.5.6

The boiling point of a specific liquid is measured for 15 samples, and the boiling points are obtained as follows:

205; 206; 206; 202; 199; 194; 197; 198; 198; 201; 201; 202; 207; 211; 205

Is there convincing evidence that the average boiling point is greater than 200? Use a significance level of 0.1. Assume the population is normal.

Exercise 9.5.7

In a study of 420,019 cell phone users, 172 of the subjects developed brain cancer. Test the claim that cell phone users developed brain cancer at a greater rate than that for non-cell phone users (the rate of brain cancer for non-cell phone users is 0.0340%). Since this is a critical issue, use a 0.005 significance level. Explain why the significance level should be so low in terms of a Type I error.

Answer

1. We need to conduct a hypothesis test on the claimed cancer rate. Our hypotheses will be
   a. $H_0: p \leq 0.00034$
   b. $H_a: p > 0.00034$

   If we commit a Type I error, we are essentially accepting a false claim. Since the claim describes cancer-causing environments, we want to minimize the chances of incorrectly identifying causes of cancer.

2. We will be testing a sample proportion with $x = 172$ and $n = 420,019$. The sample is sufficiently large because we have $np' = 420,019(0.00034) = 142.8$, $nq' = 420,019(0.99966) = 419,876.2$, two independent outcomes, and a fixed probability of success $p' = 0.00034$. Thus we will be able to generalize our results to the population.

Try It 9.5.7

In a study of 390,000 moisturizer users, 138 of the subjects developed skin diseases. Test the claim that moisturizer users developed skin diseases at a greater rate than that for non-moisturizer users (the rate of skin diseases for non-moisturizer users is 0.041%). Since this is a critical issue, use a 0.005 significance level. Explain why the significance level should be so low in terms of a Type I error.