

BUSA 603

Module 4 Supplement

A Deeper Dive on Hypothesis Testing

Brian Weikel

brian.weikel@franklin.edu



Business Analytics
Franklin University

Spring 2024

An Introduction to Validating Conjectures

Hypothesis-testing procedures enable us to test the validity of some conjecture or claim by using sample data. As we have seen, hypothesis testing is foundational to A/B testing.

Hypothesis testing complements point and interval estimation procedures.

The process begins with an investigator forming a hypothesis about the nature of some population. We clearly state this hypothesis as involving two options, and then we select one option based on the results of a statistic computed from a random sample of data.

The following are business examples of hypothesis-testing problems:

1. **Scotts** Miracle-Gro claims that, on average, its largest water soluble fertilizer item weighs at least 5.5 pounds, and thus does not weigh less than 5.5 pounds. The company can test this claim by collecting a random sample of packages, determining the weight of each one, and computing the sample mean package weight from the data.
2. **Hitachi Astemo** in Sunbury, Ohio, develops and manufactures suspension systems for motorcycles, shock absorbers, power steering systems and powertrain pumps for automobiles. Suppose it wishes to monitor its manufacturing process to ensure that the diameter of shock absorbers meets engineering tolerance specifications. It could obtain random samples every (say) 4 hours from the production line and use them to determine if standards are being maintained.

3. Suppose the [Troy-Bilt](#) company's R&D organization in Valley City, Ohio, has developed a new, energy-efficient lawn mower engine. It claims that the engine will run continuously for more than 5 hours (300 minutes) on a single gallon of regular gasoline. (The leading brand lawnmower engine runs for 300 minutes on 1 gallon of gasoline.) They could select a simple random sample of (say) 50 engines to determine the validity of the claim.
4. The marketing department of the [Goodyear Tire & Rubber Company](#), headquartered in Akron, Ohio, believes that their Facebook digital display ads are producing an incremental 1% sales lift in tires sold in their [Goodyear Auto Service Centers](#) and a 0.5% lift in all other Goodyear retailers. To determine the validity of this claim, they have engaged with Facebook to conduct a [Conversion Lift Test](#).

These examples indicate a standard procedure. We state a hypothesis about some population parameter and then collect sample data to test the validity of our hypothesis.

The hypothesis testing process is analogous to a criminal jury trial. A person charged with a crime is either innocent or guilty. In a jury trial in the U.S., Canada, Mexico, United Kingdom, France, India and many other countries we initially assume that the accused is innocent, and the jury will decide that **a person is guilty** only if there is very strong evidence against the presumption of innocence. The criminal jury trial process for choosing between guilt and innocence has the following characteristics:

- ▶ Rigorous procedures or rules for presenting and evaluating evidence.
- ▶ A judge to enforce the rules.
- ▶ A decision process that assumes innocence unless there is evidence to prove guilt beyond a reasonable doubt.

This process will fail to convict some people who are, in fact, guilty. But if a person's innocence is rejected and the person is found guilty, we have strong evidence that the person is guilty.

A general framework to test hypotheses.

- ▶ First we need to define two hypotheses, or “alternatives”, that cover all possible outcomes.
- ▶ Then, by using statistics computed from random samples, we select one of the two hypotheses.
- ▶ Since these statistics have a sampling distribution, our decision is made in the face of random variation. Thus, clear decision rules are needed for choosing between the two alternatives.
- ▶ The sample statistics cannot in general be used to absolutely “prove” that one of the two alternatives is correct. However, we can find that one of the alternatives has a very small probability of being correct. Thus as a result we would select the other alternative.
- ▶ This approach is the fundamental decision-making process used in scientific research. The term **counterfactual testing** is commonly used to define this decision process.

Suppose you are a marketing analytics consultant for a strategic management company, such as the [Boston Consulting Group](#). In addition, you and your team are assigned to support the needs of many clients across different verticals, such as the [MileOne Auto Group](#), [Nestlé](#) and the [Kroger Company](#).

A few moments ago the managing partner called you into her office and said that its vitally important that you marshal the team to address three business questions she was sent this morning. The requests are on the next three slides. She said that responses to each request must be returned by noon EST today. As you leave her office, she proudly exclaims the [company's motto](#), “Unlocking the potential of those who advance the world!”¹

¹For the first request, the Module 4 lecture notes about Bayes theorem should be helpful. You and your team may find the material in the Appendix sections useful in addressing requests 2 and 3.

Request 1: The MileOne VP of Marketing knows from past experience that 15% of the people who come into one of their showrooms and talks to a salesperson will eventually purchase an automobile.

The VP and her team executed a marketing campaign where there was a free dinner with a salesperson for all people who agreed to listen to a 2 hour sales presentation. In addition, the campaign provided 2 free tickets to either a [Philadelphia Phillies](#) or [Washington National](#) game, or a [Philadelphia Union](#) or [D.C. United](#) match. The tickets were distributed after the sales presentation and dinner.

The campaign was conducted for 6 months, and 10% of the people who purchased an automobile had a free dinner. In addition, 8% of the people who did not purchase an automobile had a free dinner.

Question 1 Does accepting the free dinner lead to a larger probability of purchasing a new automobile?

Question 2 What is the probability that a person who does not accept a free dinner will purchase an automobile?

Also interpret the results and suggest next steps that MileOne could pursue.

Request 2: Based on product ingredient changes during the last year, the Nestlé Moist & Meaty Brand Manager is asking the consulting firm to conduct an analysis to determine if store-level 6-month sales of 216-ounce packages of Moist & Meaty have increased. For the first and second fiscal quarters of last year, average weekly store sales were 2,400 units.

A random sample of sales data from 134 stores has been obtained for this year's first and second fiscal quarters, where the sample mean is 3,593 units. The sample standard deviation is 4,919 units. Conduct a test that the population mean is 2,400 using a significance level of 0.01.

Also find the p -value.

Request 3: Due to inflationary pressures and changes in the price per unit of many SKUs, the Kroger VP of Pricing, Promotion and Execution wants to know if shoppers are sensitive to the prices of items sold in a supermarket.

A random sample of 802 shoppers was obtained, and 378 of those supermarket shoppers were able to state the correct price of an item immediately after putting it into their cart.

Test at the 10% level the null hypothesis that at least one-half of all shoppers are able to state the correct price. Also find the p -value of the test and provide guidance to the the VP on how this result may affect the business.

Appendix - Hypothesis Tests for a Population Mean, Standard Deviation Unknown

In this section we consider the same form of hypothesis tests discussed in the previous section. The only difference is that the population variance is unknown; thus, we must use tests based on the Student's t distribution.

In a previous section we introduced the Student's t distribution and showed its application for developing confidence intervals.

Recall that the Student's t distribution depends on the degrees of freedom for computing the sample variance, $\nu = n - 1$.

In addition, the Student's t distribution becomes close to the normal distribution as the sample size increases. Thus, for sample sizes greater than 100 the normal probability distribution can be used to approximate the Student's t distribution.

Suppose there is a random sample of n observations from a normal distribution, $n > 30$, or the sample is approximately normal. Let μ be the mean and variance σ^2 , where σ^2 is unknown. Given the sample standard deviation s , a test with significance level α for either of the following null hypotheses,

$$H_0: \mu = \mu_0 \quad \text{or} \quad H_0: \mu \leq \mu_0,$$

against the right-tailed composite alternative

$$H_1: \mu > \mu_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{n-1, \alpha}, \quad (1)$$

where $t_{n-1, \alpha}$ is the number for which $P(T > t_{n-1, \alpha}) = \alpha$, and $T \sim t_{n-1}$.

Illustrating the Decision Rule for a Right-Tailed H_1

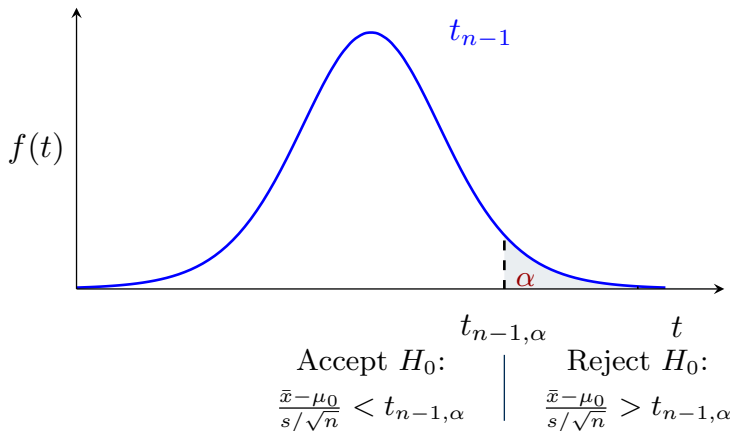


Figure 12: The Decision Rule for a Right-Tailed H_1 for a given Significance Level α and Unknown σ^2

The p -values for these tests are computed in the same way as we did for tests with known variance except that the Student's t value is substituted for the normal Z value.

The p -value for the null hypothesis $H_0: \mu = \mu_0$ tested against the right-tailed alternative hypothesis $H_1: \mu > \mu_0$ is

$$p\text{-value} = P\left(\frac{\bar{x} - \mu_0}{s/\sqrt{n}} \geq t_{n-1,p} | H_0: \mu = \mu_0\right), \quad (2)$$

where $t_{n-1,p}$ is the t_{n-1} distribution value associated with the smallest significance level at which the null hypothesis can be rejected.

To obtain the p -value we often need to interpolate in a t table or use a computer package.

Suppose there is a random sample of n observations from a normal distribution, $n > 30$, or the sample is approximately normal. Let μ be the mean and variance σ^2 , where σ^2 is unknown. Given the sample standard deviation s , a test with significance level α for either of the following null hypotheses,

$$H_0: \mu = \mu_0 \quad \text{or} \quad H_0: \mu \geq \mu_0,$$

against the left-tailed composite alternative

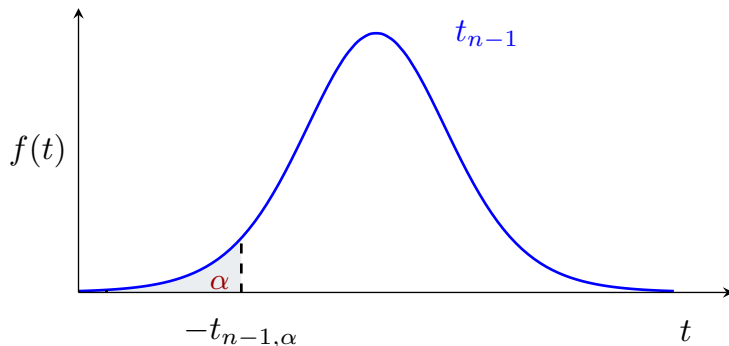
$$H_1: \mu < \mu_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{n-1,\alpha}, \quad (3)$$

where $-t_{n-1,\alpha}$ is the number for which $P(T < -t_{n-1,\alpha}) = \alpha$, and $T \sim t_{n-1}$.

Illustrating the Decision Rule for a Left-Tailed H_1



Reject H_0 :		Accept H_0 :
$\frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{n-1, \alpha}$		$\frac{\bar{x} - \mu_0}{s/\sqrt{n}} > -t_{n-1, \alpha}$

Figure 13: The Decision Rule for a Left-Tailed H_1 for a given Significance Level α and Unknown σ^2

The p -value for the null hypothesis $H_0: \mu = \mu_0$ tested against the left-tailed alternative hypothesis $H_1: \mu < \mu_0$ is

$$p\text{-value} = P\left(\frac{\bar{x} - \mu_0}{s/\sqrt{n}} \leq -t_{n-1,p} | H_0: \mu = \mu_0\right), \quad (4)$$

where $-t_{n-1,p}$ is the t_{n-1} distribution value associated with the smallest significance level at which the null hypothesis can be rejected.

Suppose there is a random sample of n observations from a normal distribution, $n > 30$, or the sample is approximately normal. Let μ be the mean and variance σ^2 , where σ^2 is unknown. Given the sample standard deviation s , a test with significance level α for the null hypothesis

$$H_0: \mu = \mu_0$$

against the two-tailed composite alternative

$$H_1: \mu \neq \mu_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{n-1, \alpha/2} \quad \text{or} \quad \frac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{n-1, \alpha/2}, \quad (5)$$

where $t_{n-1, \alpha/2}$ is the number for which $P(T > t_{n-1, \alpha/2}) = \alpha/2$, and $T \sim t_{n-1}$.

Illustrating the Decision Rule for a Two-Tailed H_1

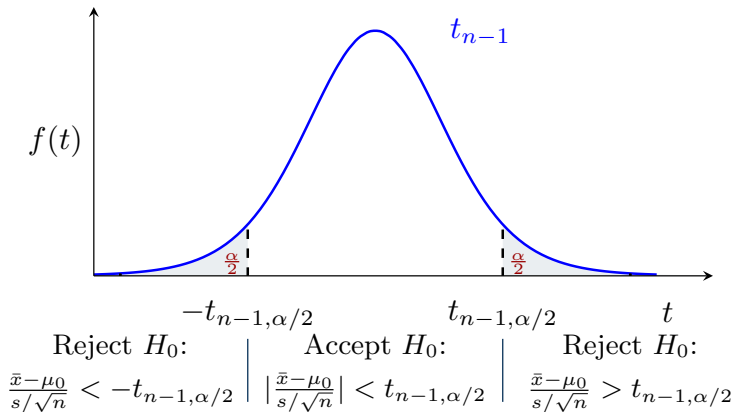


Figure 14: The Decision Rule for a Two-Tailed H_1 for a given Significance Level α and Unknown σ^2

p -value for a Two-Tailed H_1 , σ Unknown

The p -value for the null hypothesis $H_0: \mu = \mu_0$ tested against the two-tailed alternative hypothesis $H_1: \mu \neq \mu_0$ is

$$p\text{-value} = 2 \times P\left(\left|\frac{\bar{x} - \mu_0}{s/\sqrt{n}}\right| \geq t_{n-1, p/2} | H_0: \mu = \mu_0\right), \quad (6)$$

where $t_{n-1, p/2}$ is the t_{n-1} distribution value associated with the smallest significance level at which the null hypothesis can be rejected.

Check to be sure the assumptions are satisfied. If they are, then proceed with the following steps.

1. State the null and alternate hypotheses. A simple null hypothesis is of the form $H_0: \mu = \mu_0$. For a simple null hypothesis, an alternate hypothesis can be stated in one of three ways:
 - ▶ Right-tailed - $H_1: \mu > \mu_0$.
 - ▶ Left-tailed - $H_1: \mu < \mu_0$.
 - ▶ Two-tailed - $H_1: \mu \neq \mu_0$.
2. Choose a significance level α , and find the critical value or values using Student's t distribution with $n - 1$ degrees of freedom.
3. Compute the test statistic, $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$.

4. Determine whether to reject H_0 , as follows:
 - ▶ Right-tailed - $H_1: \mu > \mu_0$; Reject H_0 if $t > t_{n-1, \alpha}$.
 - ▶ Left-tailed - $H_1: \mu < \mu_0$; Reject H_0 if $t < -t_{n-1, \alpha}$.
 - ▶ Two-tailed - $H_1: \mu \neq \mu_0$; Reject H_0 if $t < -t_{n-1, \alpha/2}$ or $t > t_{n-1, \alpha/2}$.
5. Find the p -value of the test.
6. State a conclusion.

Appendix - Hypothesis Tests for Population Proportions

Another important set of business and economics problems involves population proportions.

Business executives are interested in the percent market share for their products, and government officials are interested in the percentage of people that support a proposed new program.

Inference about the population proportion based on sample proportions is an important application of hypothesis testing.

Notation:

- ▶ p is the population proportion of individuals who are in a specified category. p_0 is the population proportion specified by the null hypothesis H_0 .
- ▶ x is the number of sample individuals who are in the specified category.
- ▶ n is the sample size.
- ▶ \hat{p} is the sample proportion of individuals who are in the specified category: $\hat{p} = x/n$. Frequently \hat{p} is referred to as the “observed proportion of successes.”

Assumptions for large sample hypothesis testing of p .

- ▶ We have a simple random sample.
- ▶ The population is at least 20 times as large as the sample.
- ▶ The items in the population are divided into two categories.
- ▶ The values np_0 and $n(1 - p_0)$ are both at least 10.

Suppose that the assumptions mentioned earlier in this section are met. If the observed sample proportion is \hat{p} , then a test with significance level α for either of the following null hypotheses,

$$H_0: p = p_0 \quad \text{or} \quad H_0: p \leq p_0,$$

against the right-tailed composite alternative

$$H_1: p > p_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_\alpha, \quad (7)$$

where z_α is the number for which $P(Z > z_\alpha) = \alpha$, and Z is a standard normal random variable.

Illustrating the Decision Rule for a Right-Tailed H_1

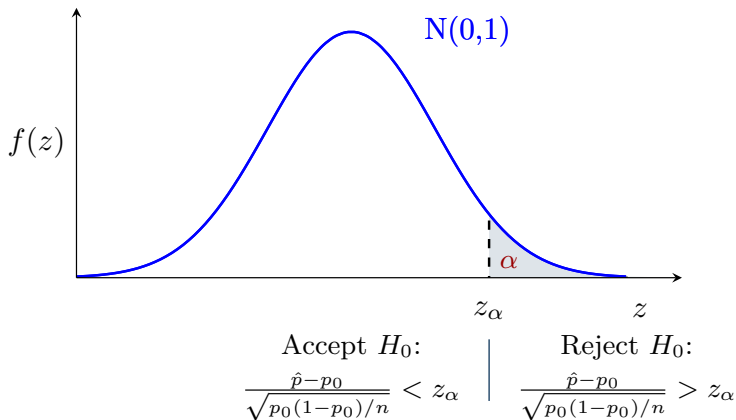


Figure 15: The Decision Rule for Testing p with a Right-Tailed H_1 for a Given Significance Level α

The p -value for the null hypothesis $H_0: p = p_0$ tested against the right-tailed alternative hypothesis $H_1: p > p_0$ is

$$p\text{-value} = P\left(\frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \geq z_p \mid H_0: p = p_0\right), \quad (8)$$

where z_p is the standard normal value associated with the smallest significance level at which the null hypothesis can be rejected.

Suppose that the assumptions mentioned earlier in this section are met. If the observed sample proportion is \hat{p} , then a test with significance level α for either of the following null hypotheses,

$$H_0: p = p_0 \quad \text{or} \quad H_0: p \geq p_0,$$

against the left-tailed composite alternative

$$H_1: p < p_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} < -z_\alpha, \quad (9)$$

where $-z_\alpha$ is the number for which $P(Z < -z_\alpha) = \alpha$, and Z is a standard normal random variable.

Illustrating the Decision Rule for a Left-Tailed H_1

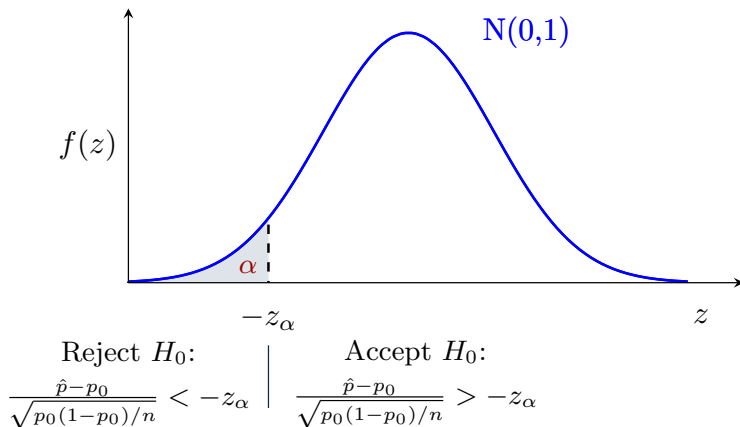


Figure 16: The Decision Rule for Testing p with a Left-Tailed H_1 for a given Significance Level α

The p -value for the null hypothesis $H_0: p = p_0$ tested against the left-tailed alternative hypothesis $H_1: p < p_0$ is

$$p\text{-value} = P\left(\frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \leq -z_p \mid H_0: p = p_0\right), \quad (10)$$

where $-z_p$ is the standard normal value associated with the smallest significance level at which the null hypothesis can be rejected.

Suppose that the assumptions mentioned earlier in this section are met. If the observed sample proportion is \hat{p} , then a test with significance level α for the null hypothesis,

$$H_0: p = p_0$$

against the two-tailed composite alternative

$$H_1: p \neq p_0,$$

is obtained by the decision rule

$$\text{Reject } H_0 \text{ if } \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} < -z_{\alpha/2} \quad \text{or} \quad \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} > z_{\alpha/2}, \quad (11)$$

where $z_{\alpha/2}$ is the number for which $P(Z > z_{\alpha/2}) = \alpha/2$, and Z is a standard normal random variable.

Illustrating the Decision Rule for a Two-Tailed H_1

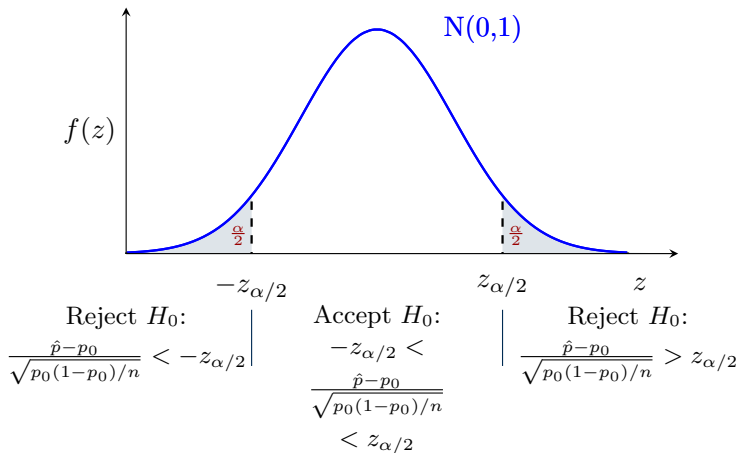


Figure 17: The Decision Rule for Testing p with a Two-Tailed H_1 for a given Significance Level α

The p -value for the null hypothesis $H_0: p = p_0$ tested against the two-tailed alternative hypothesis $H_1: p \neq p_0$ is

$$p\text{-value} = 2 \times P\left(\left|\frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}\right| \geq z_{p/2} | H_0: p = p_0\right), \quad (12)$$

where $z_{p/2}$ is the standard normal value associated with the smallest significance level at which the null hypothesis can be rejected.

Check to be sure the assumptions are satisfied. If they are, then proceed with the following steps.

1. State the null and alternate hypotheses. A simple null hypothesis is of the form $H_0: p = p_0$. For a simple null hypothesis, an alternate hypothesis can be stated in one of three ways:
 - ▶ Right-tailed - $H_1: p > p_0$.
 - ▶ Left-tailed - $H_1: p < p_0$.
 - ▶ Two-tailed - $H_1: p \neq p_0$.
2. Choose a significance level α , and find the critical value or values using the standard normal distribution.
3. Compute the test statistic, $z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$.

4. Determine whether to reject H_0 , as follows:
 - ▶ Right-tailed - $H_1: p > p_0$; Reject H_0 if $z > z_\alpha$.
 - ▶ Left-tailed - $H_1: p < p_0$; Reject H_0 if $z < -z_\alpha$.
 - ▶ Two-tailed - $H_1: p \neq p_0$; Reject H_0 if $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$.
5. Find the p -value of the test.
6. State a conclusion.