Statistics &Data Analysis Concepts for Data Science and ML 9

9

Introduction to Hypothesis Testing

Learning Objectives

- **Understand the importance of hypothesis testing in data analysis**
- **Formulate the null and alternate hypothesis about a single population parameter of interest**
- **Develop both one-and two-tailed null and alternate hypotheses and apply correct distribution and test statistic to test these hypotheses**
- **Understand Type I and Type II errors in the context of hypothesis testing and learn how the type I error, or the level of significance (α) is decided**
- **Develop the rejection and non-rejection areas based on the level of significance or the Type I error, α**

Learning Objectives…continued

- Develop and test hypotheses about a single population mean for known and unknown population standard deviation using the normal (z-statistic) and t-distribution
- Develop and test hypotheses about a single population proportion and a single population variance using the normal and chi-square distributions respectively
- Solve problems involving Type II errors

Hypothesis And Hypothesis Testing

A ***hypothesis*** is a statement about a *population parameter*. This statement may come from a claim made by a manufacturer, a mathematical model, a theory, a design specification, etc.

- For example, an automobile manufacturer may claim that they have come up with a new fuel injection system design that will provide an improved average mileage of 35 miles a gallon. In such a case, we may want to test the claim by taking a sample.
- *The claim can formally be written as a hypothesis testing problem and can be tested using a hypothesis testing procedure.*

Test of Hypothesis

- The hypothesis test may involve one or more population parameter of interest such as, a population mean, a population variance or a population proportion.
- *The population parameters of interest in testing hypothesis may be the following:*

Population mean :	μ
Population variance	σ^2
Population standard deviation	σ
Population proportion	p

Hypothesis and Hypothesis Testing

- A ***hypothesis*** is a statement about the population parameter is called a ***hypothesis***.

Many problems require that we decide whether or not a statement about some parameter is true or false. Data are some form of measurements

----

- ***Hypothesis Testing*** *is* the decision making procedure about the statement being true or false is called ***hypothesis testing***.

Hypothesis testing comes under the broad category of Inferential Statistics. It is also known as inference procedure. The two major tools of inferential statistics: estimation and hypothesis testing.

Hypothesis Testing Procedure

Make a statement about a population parameter (state the hypothesis)

Collect the sample data and calculate the sample statistics

Use the information in the sample data and an appropriate test statistic (formula) to decide how likely it is that our hypothesized population parameter is correct

Example

Suppose we make a statement about a population parameter (for example, a population mean, μ) which is a claim from an automobile manufacturer. The manufacturer claims that their new hybrid model will provide 60 miles per gallon on the average because of an improved design. We want to test if this claim is correct. The problem can be stated formally as follows:

$$H_0 : \mu = 60mpg$$

$$H_1 : \mu \neq 60mpg$$

H_0 is known as the null hypothesis and H_1 (also written H_a) is called the alternate hypothesis. The hypothesis written with an 'equal to' sign under the null hypothesis and a 'not equal to' sign under the alternate hypothesis is known as a two-sided or two-tailed test. A hypothesis can also be one-sided or one-tailed.

WHEN TO ACCEPT OR REJECT A HYPOTHESIS?

Suppose that a report suggests that the average salary of mid-level executives in certain industry category is \$120,000. The appropriate hypothesis can be stated as:

$$H_0: \mu = \$120,000$$
$$H_1: \mu \neq \$120,000$$

To verify this claim, we collect sample data. Suppose that a sample of 50 mid-level executives indicated that the average salary is \$130,000. In this case, we would not reject the null hypothesis of \$120,000 and conclude that the report is correct.

On the other hand, suppose we took another sample and the average came out to be \$70,000. In this case, we would be skeptical and probably reject the null hypothesis and conclude that the report is not correct.

Continued…

What if the average salary from a sample suggested the mean of =$115,000. This value is close to the hypothesized value of $120,000 and the decision to accept or reject the hypothesis is difficult in this case. Whether we accept or reject, we cannot be absolutely certain that our decision is correct. Thus, we have to deal with uncertainty in our decision making.

In hypothesis testing, the decision to reject or not to reject a hypothesis is based on a single sample and therefore, there is always a chance of not rejecting a hypothesis that is false, or rejecting a hypothesis that is true. In fact, we always encounter two types of errors in hypothesis testing. These errors are explained in the next slide.

Two types of errors in Hypothesis Testing

$$\text{Type I Error} = \alpha = P\{\text{Reject } H_0 \mid H_0 \text{ is true}\}$$
$$\text{Type II Error} = \beta = P\{\text{Fail to reject } H_0 \mid H_0 \text{ is false}\}$$

The type I error is denoted by a Greek letter α ("alpha"). It is the probability of rejecting a true null hypothesis.

The type II error is denoted by a Greek letter β ("beta") is the probability of not rejecting a null hypothesis when it is false. In a hypothesis testing situation there is always a possibility of making one of these errors.

Basic Concepts of Hypothesis Testing

Suppose that a movable roof is to be constructed using aluminum sheets. The design specification suggests that the average thickness of the sheets should be 0.05 inch. If the average thickness is more than 0.05 inch, the roof may not sustain the extra load. If the average thickness is less than this value, the roof may not be strong enough. You just received a shipment of 1,000 sheets from the manufacturer of these sheets, and would like to determine if the shipment is suitable to construct the roof.

You would be testing the following hypothesis:

$$H_0: \mu = 0.05 inch$$

$$H_1: \mu \neq 0.05 inch$$

A random sample of n=100 sheets, measured the thickness and the sample average thickness was found to be $\bar{x}$=0.0509 inch with a known population standard deviation of σ=0.005 inch.

Does this sample data indicate that the shipment of 1,000 sheets is suitable to construct the roof? or, in other words, what is the chance (probability) of getting a sample mean of 0.0509 inch or more from a population with a mean of 0.05 inch?

This probability will determine if it is reasonable to get a sample mean of x = 0.0509 that might have come from a population with a mean of μ = 0.05. If the probability is low (for example, 15%), we will conclude that the average thickness of the sheets is not 0.05 on the average. If the calculated probability is high (for example, 90% or higher), we will conclude that the sample actually came from a population with a mean of 0.05 inch and we will accept the shipment.

For our example the probability is approximately 7%.

The probability of 7% is small therefore, it could be concluded that there is a less chance that a sample mean of $\bar{x}$=0.0509 in. came from a population with a mean μ=0.5 inch. In this case, the probability that the population would produce such a random sample is too low and we reject the null hypothesis.

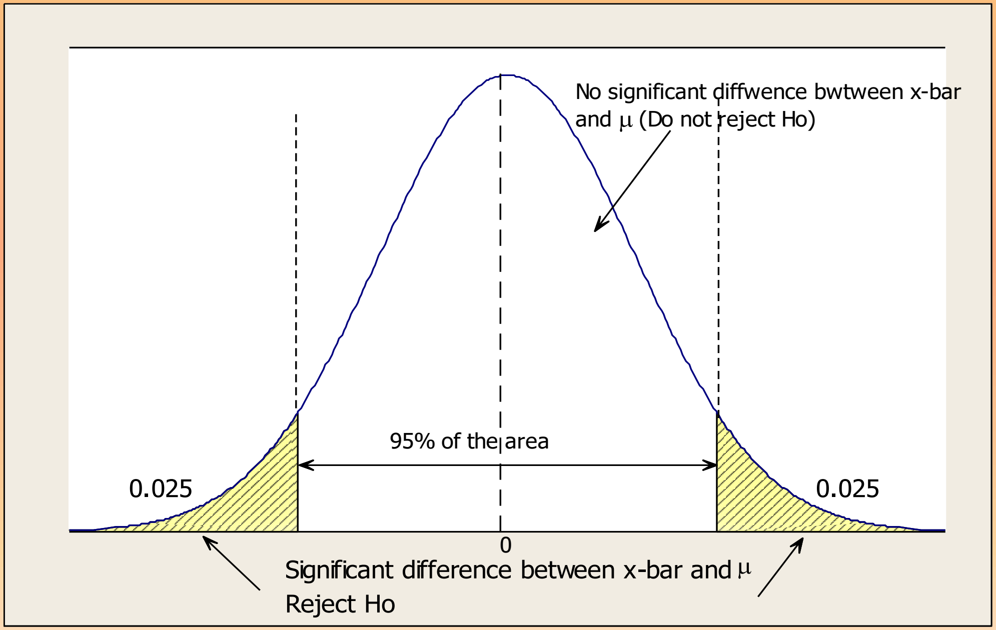
In the above case, we decided to reject the null hypothesis because the probability of 7% was low. What would you do if the calculated probability was 9% or 10% or 15%? Would you still conclude that these values are low? When do you say it is not low?

- ***This leads to an important decision that we need to make before we conduct a hypothesis test. We must decide a minimum acceptable level of probability before we conduct the test. This minimum acceptable level of probability is known as the level of significance and is denoted by alpha (α).***
- ***We must decide α suitable value of a before we conduct the test. The next question is: how do we select an appropriate value of the level of significance, α ? The value of a is decided based on***
  - ***the cost resulting from an incorrect decision, and***
  - ***the level of risk you are willing to take.***
- ***As we've seen earlier, there is always a risk of making an incorrect decision while testing a hypothesis. Whenever you make an incorrect decision, there is a cost or a risk involved. So, we must be careful in deciding the value of α.***

- ***The level of significance, α is the same as the type I error.***
- ***The type I error is the probability of rejecting a true null hypothesis. Since we don't want this to happen, α is set to a low value of 1%, 5% or 10% in general cases.***
- ***If you set the value of α = 5%, it means that there is a 5% chance of making an incorrect decision, and a 95% chance of making a right decision. In a hypothesis testing, there is never a 100% chance of making a right decision. Why?***

What does It mean to Test a Hypothesis at a 5% Level of Significance?

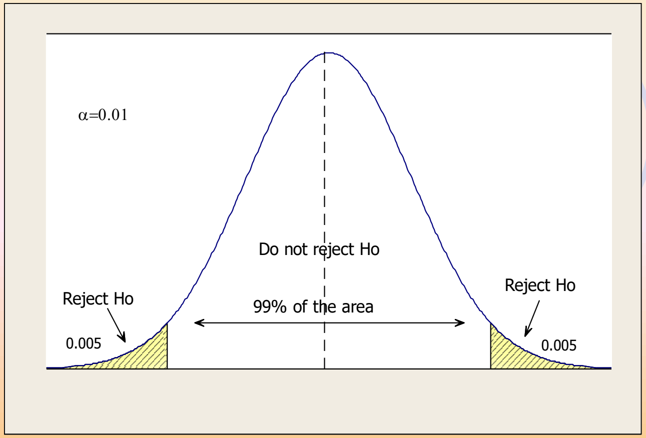
Suppose we want to test a hypothesis about a population mean and the sample size is large (n>= 30) so that the sample mean follows a normal distribution. If the level of significance α is set at 5%, it means that we will reject the null hypothesis if the difference between the sample statistic (in this case, $\bar{x}$) and the hypothesized population mean μ is so large that it would occur on the average only 5 or fewer times in every 100 samples. See Figure 9.3.

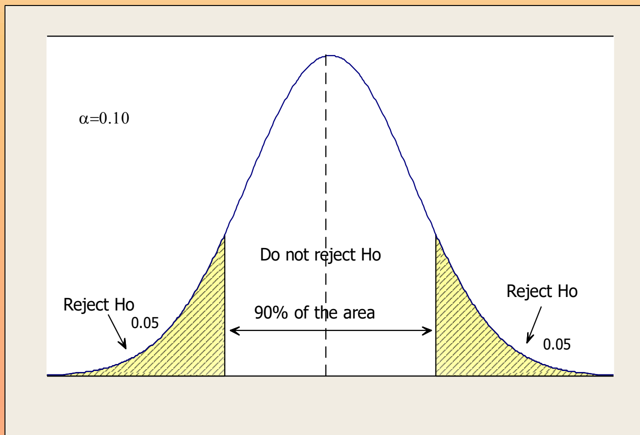


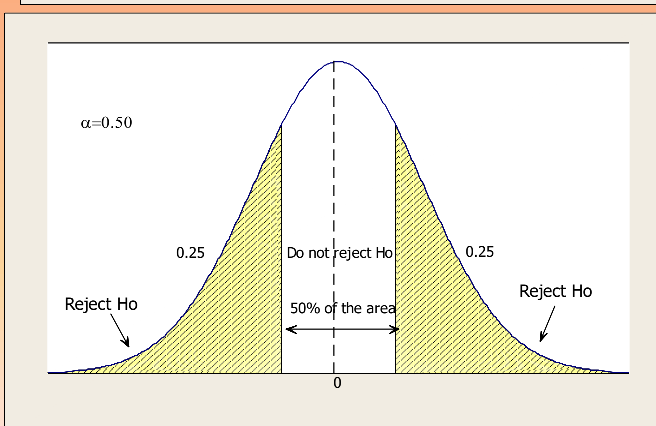
If the sample statistic falls in the non-rejection area (non-shaded area in Figure 9.3), it does not prove that our null hypothesis (Ho) is true, it simply means that we don't have statistical evidence to reject it. Thus, accepting the null hypothesis means that there is no sufficient evidence to reject it. This is why we don't say 'accept the null hypothesis (Ho)', we simply say, "do not reject Ho."

Selecting a Value of Significance Level (α)

- ***Once the significance level, α is selected, the areas of rejection and non-rejection are determined.***
- ***The value of α is selected by the analyst and it must be determined before conducting the test.***
- ***There is no universal or single level of significance used. A significance level α of 1% and 5% are widely used in general cases but, the hypothesis can be tested at any level of significance.***
- ***In general, higher the significance level, higher is the probability of rejecting a null hypothesis. (see the next slide)***







Increasing α makes the "do not reject" area quite small.

With the acceptance level this small, we will frequently reject a null hypothesis when it is true (high α), and we will rarely accept a null hypothesis when it is false (low β).

In order to get a low β, we must put up with a high α.

To deal with this, we need to decide an appropriate level of significance by examining the cost and risk involved in both types of errors.

In general, it is desirable to have both type I and type II errors small. In some exceptional cases it may be required to have a very small β. This can be achieved by selecting a high α.

Relationship between Type I and Type II Error

We can reject a null hypothesis when it is true (Type I Error) and we can fail to reject a null hypothesis when it is false (type II error). These errors can also be stated in symbolic form as shown below.

$$\alpha = P\text{ (type I error)} = P\{\text{Reject Ho} \mid \text{Ho is true}\}$$

$$\beta = P\text{ (type II error)} = P\{\text{Fail to reject Ho} \mid \text{Ho is false}\}$$

Sometimes it is more convenient to work with the power of the test. The power of the test is defined as follows:

$$\text{Power} = 1 - \beta = P\{\text{Reject Ho} \mid \text{Ho is false}\}$$

The power of the test is the probability that a false null hypothesis is correctly rejected.

The type I error, α is selected by the analyst. Increasing the type I error α will decrease the type II error β and decreasing the type I error, α will increase the type II error, β.

Types of Hypothesis Test

A two-tailed test or a two-sided test

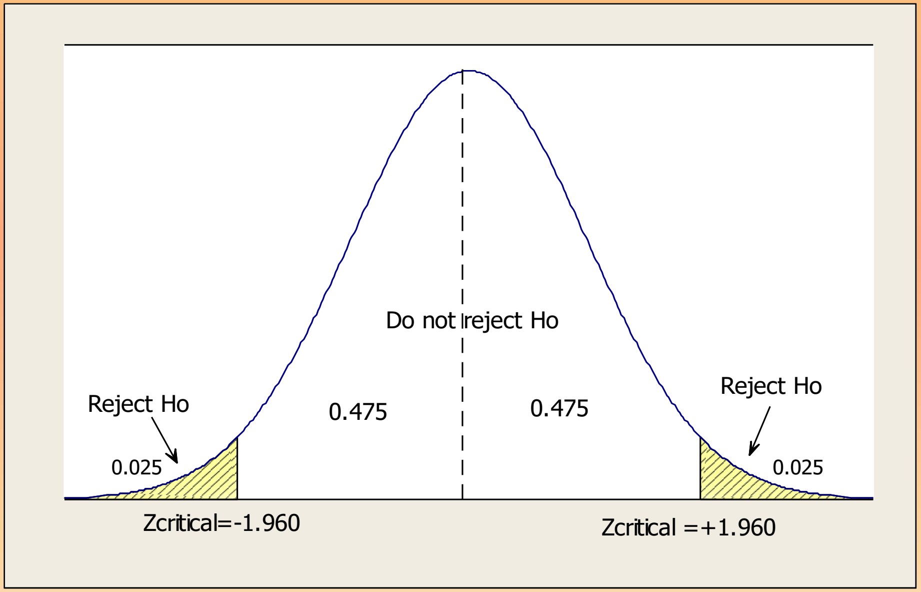
The null and the alternate hypothesis in a two-tailed test is written as

$$H_0 : \mu = 200$$

$$H_1 : \mu \neq 200$$

There are two rejection regions in a two-tailed test. In this type of test, the null hypothesis is rejected when the sample mean is significantly higher or lower than the hypothesized population mean m.

The acceptance and rejection regions for a two-sided test with α =0.05 is shown in Figure 9.6. In a two-tailed test, α is divided into two halves. For a test with α=0.05, each tail contains 2.5% of the area leading to two critical points, one in each tail.



A one-tailed test (or, a one-sided test)

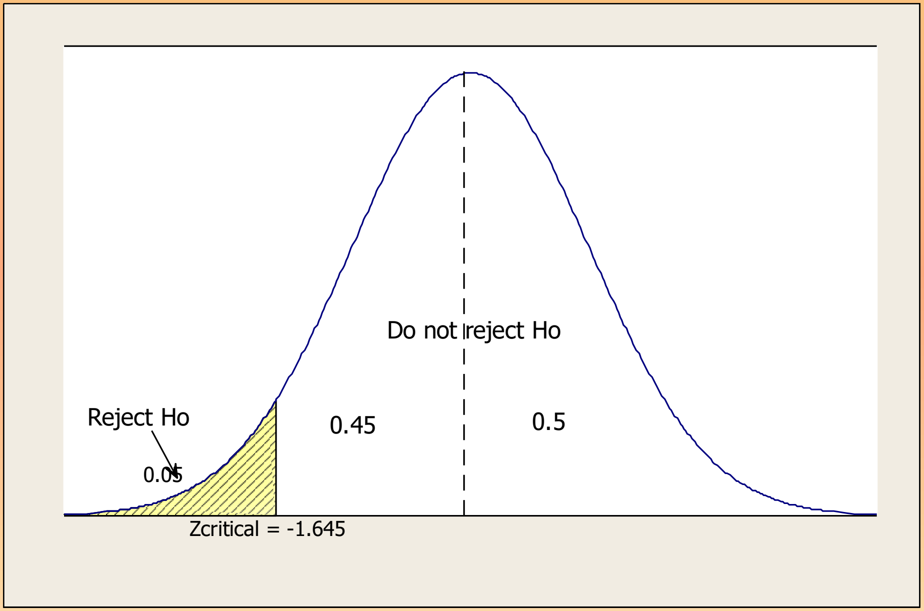
A left-tailed or a left-sided test

A right-tailed or a right-sided test

A one-tailed test can be a lower or left-tailed and a upper right-tailed test.

A left-tailed or a left-sided test

In a left-tailed test, there is one area of rejection (in the left tail of the distribution) and the entire α value is on one side. For a test with α=0.05, the left tail contains 5% of the area leading to one critical point. The decision areas for a left-tailed test are shown

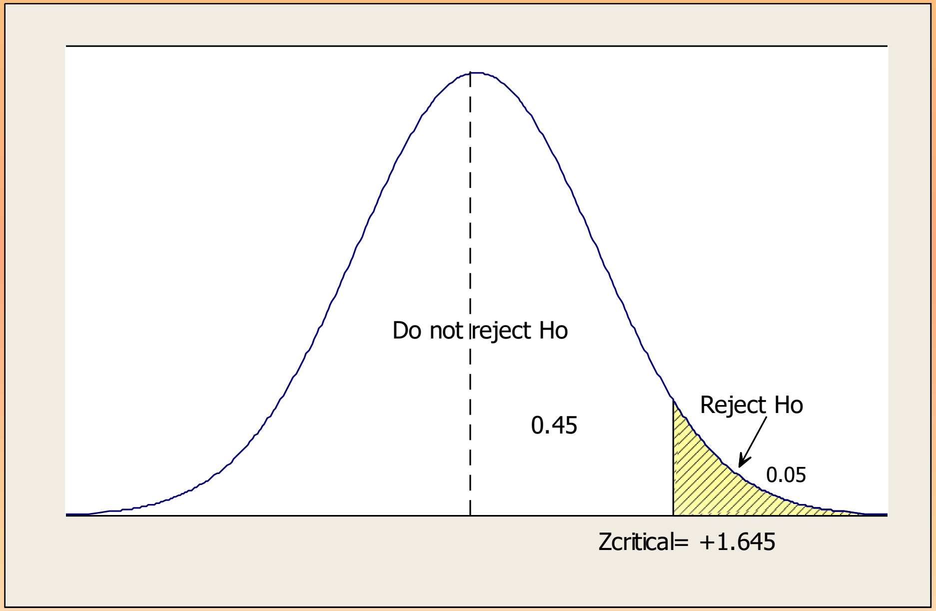


A right-tailed or a right-sided test

A ***right -tailed*** test is written as:

$$H_0 : \mu \leq 200$$
$$H_1 : \mu > 200$$

In a right-tailed test, there is one area of rejection (in the right tail of the distribution) and the entire α value is on one side. For a test with α=0.05, the right tail contains 5% of the area leading to one critical point. The decision areas for a right-tailed test are shown



Note that the null hypothesis in a one-tailed or a two-tailed test is always written with an "equal to" sign. It is either "equal to", "less-than or equal to," or "greater-than or equal to" sign.

Formulating the correct Hypothesis

A hypothesis test can be formulated as a one-sided or a two- sided test. If we have a one-sided test, it can be a left-sided or a right-sided test.

Example of a Left-sided test

Suppose a tire manufacturer claims that the average mileage provided by a certain type of tire is at least 60,000 miles. A product research group has received some complains in the past and wants to check if the average mileage is below 60,000 miles. They would formulate their hypothesis as

---

$$H_0 : \mu \geq 60000$$

$$H_1 : \mu < 60000$$

The hypothesis is formulated as shown above because the average of at least 60,000 miles is $\mu \geq 60{,}000$ and the opposite of this statement is $\mu < 60{,}000$ miles. Since the null hypothesis is written with an 'equal to' sign, and this is the claim made by the manufacturer, the statement $\mu \geq 60{,}000$ is written under the null hypothesis, and the statement $\mu < 60{,}000$ is written under the alternate hypothesis.

Formulating the correct Hypothesis…cont.

The alternate hypothesis is also known as the research hypothesis. If you are trying to establish certain hypothesis, then it should be written as the alternate hypothesis.

The null hypothesis is where the claim or theory is. Therefore, rejecting a null hypothesis is a strong statement. This is the reason that the conclusion of a hypothesis test is stated as "reject the null hypothesis" or "do not reject the null hypothesis."

Example of a Right -sided test:

A car manufacturer has made a significant improvement in the fuel injection system that is expected to provide an improved gas mileage. The average mileage before the modification was 24 miles or less. The research group expects that the modified system will provide significant improvement in the gas mileage. The group would like to test the following hypothesis to show the improvement:

$$H_0 : \mu \leq 24$$
$$H_1 : \mu > 24$$

Formulating the correct Hypothesis…cont.

Example of a Two-sided test

A robot welder in an assembly line takes on average 1.5 minutes to finish a welding job. If the average time taken to finish the job is higher or lower than 1.5 min. it will disrupt other activities along the production line. Since there has been too much variation in the time it takes to perform the welding job by the robot, the line supervisor wants to take a sample to check if the average time taken by the robot is significantly higher or lower than the average of 1.5 minutes. The supervisor would be testing the following hypothesis:

$$H_0 : \mu = 1.5$$

$$H_1 : \mu \neq 1.5$$

Note that it is a two-sided test. The null hypothesis will be rejected if the sample mean is significantly higher or lower than the hypothesized population average of 1.5 min.

Important Note

It is important to remember that the hypotheses are always statements about the population. The value of the population parameter specified in the null hypothesis is determined in the following ways:

- ***It may come from the past experience or the knowledge of the process or from previous experimentation. The objective of the hypothesis testing then is to determine if the process parameter or the experimental condition has changed.***

- ***The value specified in the null hypothesis may be determined from some mathematical model or theory regarding the process under study. In this case the objective of hypothesis testing is to verify the theory or the model.***

- ***The value in the hypothesis may result from design or engineering specifications or from contractual obligations. In such cases, the objective of the test is to determine if the sample data conforms to specification.***

- ***Finally, the hypothesis may be stated based on some type of claim. The objective of the hypothesis testing is then to verify the claim.***

TESTING A SINGLE POPULATION MEAN

Testing a population mean involves testing a one-sided or a two-sided test. The hypothesis is stated as:

$H_0 : \mu = \mu_0$ $H_1 : \mu \neq \mu_0$ ***Two-tailed or Two-sided Test***	$H_0 : \mu \geq \mu_0$ $H_1 : \mu < \mu_0$ ***Left-tailed or Left-sided Test***	$H_0 : \mu \leq \mu_0$ $H_1 : \mu > \mu_0$ **Right –Tailed or *ight-sided Test***

Case (1): Testing a single mean with known variance or known population standard deviation σ and large sample. In this case, the sample mean follows a Normal distribution and the test statistic is given as follows:

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

TESTING A SINGLE POPULATION MEAN…continued

Case (2): Testing a single mean with unknown variance or unknown population standard deviation σ and large sample. In this case, the sample mean follows a Normal distribution and the test statistic is given by

$$z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Case (3): Testing a single mean with unknown variance or known population standard deviation σ and small (n<30) sample. In this case, the sample mean, follows a t- distribution and the test statistic is given by

$$t_{n-1} = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

EXAMPLE ON FORMULATING HYPOTHESES

Example 1

A machine is used to fill plastic bottles with a net volume of 32 ounces. A sample of bottles is periodically selected and weighed to determine whether under filling or over filling is occurring. If the sample data leads to a conclusion of under filling or overfilling; the production line must be stopped and the machine should be adjusted. Suppose a hypothesis is to be tested that the mean volume is 32 ounces.

Solution:

(a) Formulate the appropriate null and alternate hypothesis.

$$H_0 : \mu = 32oz.$$
$$H_1 : \mu \neq 32oz.$$

(b) Comment on the conclusion and the decision when the null hypothesis cannot be rejected.

There is no evidence that the mean weight is different from 32 oz. The machine does not need adjustment.

(c) Comment on the conclusion and the decision when the null hypothesis can be rejected.

There is evidence that the mean wt. is different from 32 oz. The machine needs adjustment.

Example 2

People in management positions spend an average of 150 minutes per day reading and responding to e-mails. A researcher believes that this time is more than 150 minutes. A sample of individuals from management was selected by the researcher to verify the claim.

(a) Write the appropriate hypothesis to be tested.

$$H_0 : \mu \leq 150\,\text{min}.$$
$$H_1 : \mu > 150\,\text{min}.$$

(b) What is Type I and Type II error in this situation and what is the consequence of making each error?

Type I error = α = P $\{$Reject H_o H_o is true$\}$
(Probability of rejecting a true null hypothesis)

Type II error = β = P $\{$Fail to reject $H_o | H_o$ is false$\}$
(Probability of not rejecting a false null hypothesis)

Type I error in this situation would be to conclude that the average time is greater than 150 min. when it is not. Type II error would be to conclude that the average time is at most 150 min. when it is greater than 150 min.

Example 3

A company sells orange juice from concentrate and claims that their juice contains an average of 25 grams of total carbohydrate or less. Suppose a sample of 35 cans selected randomly is to be tested to verify the carbohydrate content.

(a) Write the appropriate hypothesis that could be used to test the claim.

$$H_0 : \mu \leq 25 grams$$
$$H_1 : \mu > 25 grams$$

(b) State and explain the Type I and Type II error in this situation.

Type I error: Reject Ho: $\mu \leq 25$ gm. and conclude that the average carbohydrate content is more than 25 gm. ($\mu > 25$ gm). [Reject the claim when it is true].

Type II error would be to conclude that the carbohydrate content is at most 25 gm. when it is not. That is, the product does not meet the label specification [accept the claim when it is false].

THE HYPOTHESIS TESTING STEPS

A hypothesis can be tested using several methods. The steps below are for testing hypothesis using the z-value approach or the classical approach. These steps are applicable for testing a single mean when the population variance is known.

z-value Approach for Testing a Hypothesis

1. State the null and alternate hypothesis.
2. Determine the sample size or use the given sample size.
3. Determine the appropriate level of significance (a) or use the given a.
4. Select the appropriate distribution and test statistic to perform the test.
5. Based on step 3, find the critical point or points and the area or areas of rejection. Show the critical point(s) and the area or areas of rejection and non-rejection using a sketch.
6. Write the decision rule
7. Use the test data (sample data) and find the value of the test statistic.
8. Find out if the value of the test statistic is in rejection or non-rejection region; make appropriate decision and state your conclusion in terms of the problem.

Example 4: Illustrating the steps of hypothesis testing

In a production line, automated machines are used to fill beverage cans with an average fill volume of 16 ounces. If the mean weight falls above or below this figure, the production line must be stopped and some remedial action be taken. A quality control inspector samples 30 cans every hour; opens and weighs the content, and tests the appropriate hypothesis and makes a decision whether to shut down the line for making adjustments. Write the appropriate hypothesis to be tested in this situation and perform the hypothesis test. A significance level of α=0.05 is selected for the test. The sample results indicate a sample mean of 16.32 oz. and the standard deviation is assumed to be 0.8 oz.

Solution: For this problem, the given data are:

$$n = 30, \alpha = 0.05$$

$$\sigma = 0.8, \bar{x} = 16.32$$

1. State the null and alternate hypothesis

$$H_0 : \mu = 16$$

$$H_1 : \mu \neq 16$$

Note that this is a two-sided test.

Example 4…continued

2. Determine the sample size or use the given sample size.
   The given sample size is n=30 (large sample)

3. Determine the appropriate level of significance (α) or use the given α.
   The given level of significance, $\alpha = 0.05$

4. Select the appropriate distribution and test statistic to perform the test

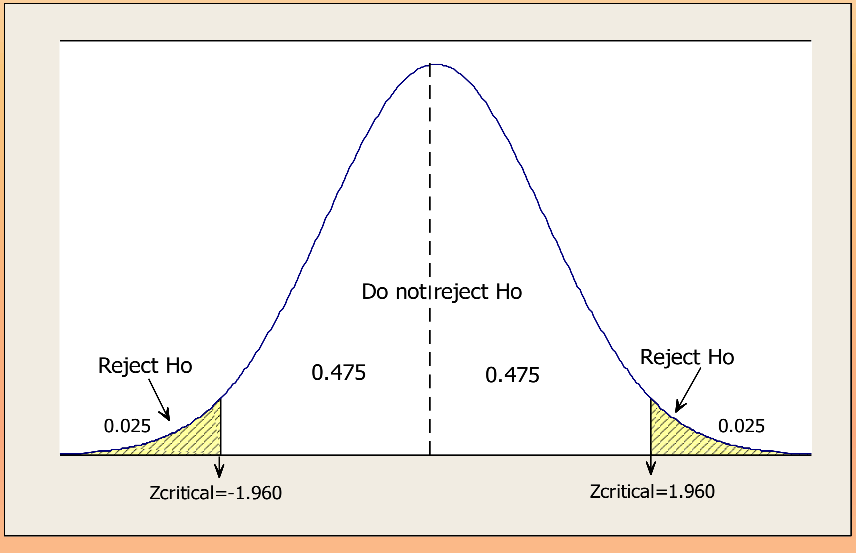
The sample size is large and the population standard deviation is known, therefore, use normal distribution with the following test-statistic

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

5. Based on step 3, find the critical point or points and the area or areas of rejection. Show the critical point(s) and the area or areas of rejection and non-rejection on a sketch

Example 4...continued

This is a two-sided test. The level of significance, $\alpha = 0.05$ must be split into two halves for a two-tailed test with each tail area 0.025. The critical point (z-critical value) for an area of 0.475 is 1.96 from the normal or z- table. The sketch is shown below.



6. Write the decision rule

Reject Ho if $z > 1.96$
or, if $z < -1.96$

Example 4…continued

7. Use the test data (sample data) and find the value of the test statistic.

$$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{16.32 - 16}{0.8 / \sqrt{30}} = 2.19$$

8. *Find out if the value of the test statistic is in rejection or non rejection region; make appropriate decision and state your conclusion in the context of the problem.*

z = 2.19 > Zcritical =1.96 therefore, reject Ho

The fill volume is different from 16 oz., the production line must be stopped and remedial action be taken.

Example 5

Suppose that a sample of 40 individuals was taken to study the time the professionals in management positions spent reading and responding to e-mails every day. The sample provided an average e-mail reading and responding time of 160.5 min with a standard deviation of 25 minutes. Assuming a level of significance of 2%, test the hypothesis that the average time is greater than 150 minutes.

(a) Test the appropriate hypothesis. ➤

$$H_0 : \mu \le 150 \text{ min}.$$
$$H_1 : \mu > 150 \text{ min}.$$

(Right-sided test)

Given data

$$\bar{x} = 160.5$$

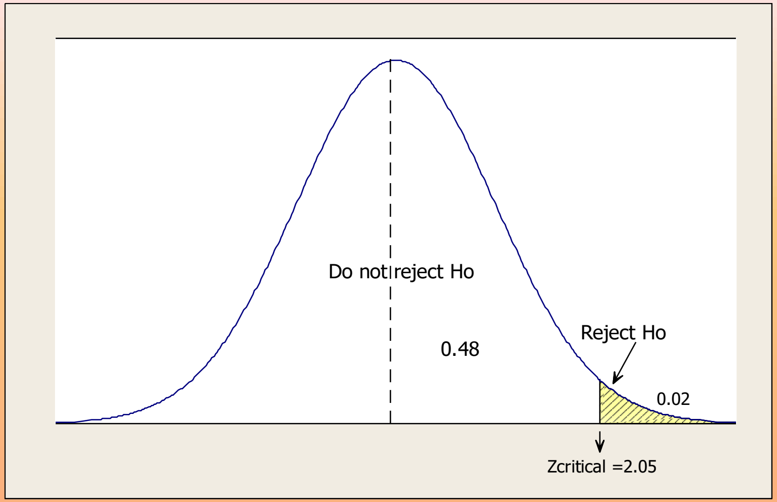
n = 40 s = 25 min. α = 0.02

(b) What distribution and test statistics should be used to test this hypothesis? ➤ Since the sample size (n) is large and s unknown, the test statistics is given by

$$z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Example 5…continued

(c) Critical value: see the figure below.



(d) Write the decision rule for the test.

Reject Ho if z > 2.05

(e) Find the test statistics value and test the hypothesis. What is your conclusion?

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{160.5 - 150}{25/\sqrt{40}} = 2.66$$

Since, z= 2.65 > z critical-value (= 2.05) therefore, reject Ho (the average e-mail reading time is greater than 150 minutes).

Different ways of Testing a Hypothesis

A hypothesis can be tested using the following methods:

(1) z-value approach or the classical approach
(2) p-value approach
(3) Critical value approach
(4) Confidence interval approach

Example 6

Suppose we want to test the following hypothesis

$$H_0 : \mu = 15$$
$$H_1 : \mu \neq 15$$

A sample size of n=50 provides a sample mean of 14.2 with a sample standard deviation of 5. The type I error or the level of significance is 2%. The data for the problem: n = 50, $\bar{x}$ = 14.2, s = 5, α = 0.02
Use the four methods above to test his hypothesis.

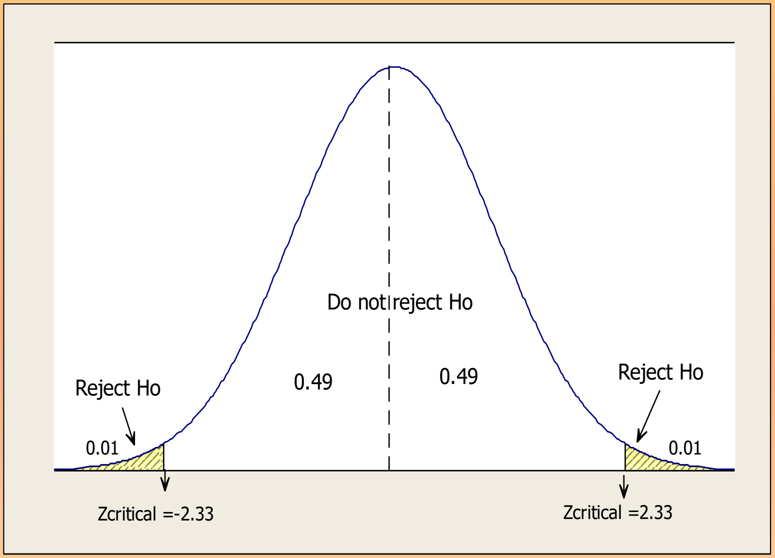
Solution: The problem is solved below using 4 different methods.

Method (1) – z-value approach

The sample size n is large and the population standard deviation, σ is unknown; the ***test statistics*** is given by

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

Critical values:



Decision Rule:

Reject Ho if z > 2.33
or, if z < -2.33

Test Statistics value:

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{14.2 - 15}{5/\sqrt{50}} = -1.13$$

Since, z=-1.13> Zcritical = -2.33, do not reject Ho

or, -2.33 ≤ Z = -1.13 ≤ 2.33; do not reject Ho.

Method (2): *p* - value approach

The p-value or, the probability value approach. This approach compares a probability to the given probability.

Hypothesis

$$H_0 : \mu = 15$$
$$H_1 : \mu \neq 15$$

The data for the problem: n=50, $\bar{x}$ = 14.2, s = 5, α = 0.02

Decision Rule for p-value approach

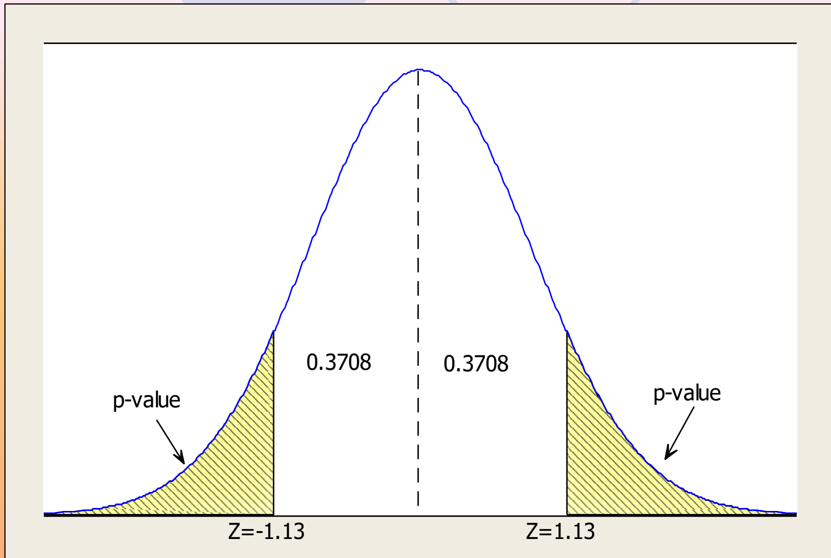
If $p \geq \alpha$, do not reject H_o

If $p < \alpha$; reject H_o

First; using the appropriate test statistic formula, calculate the test statistic value (this is also shown in method 1 above).

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{14.2 - 15}{5/\sqrt{50}} = -1.13$$

This test statistic value of z = -1.13 will be converted into a probability that we call p-value (see the next slide)



Area corresponding to = 1.13 is 0.3708 (from -table).

Probability of $z > 1.13 = 0.5 - 0.3708$
$=0.1292$

Probability of $z < -1.13 = 0.5 - 0.3708$
$= 0.1292$

For a two-sided test, the p value is the sum of the above two values, that is, 0.1292+0.1292 = 0.2584.

Since, $p=0.2584 > \alpha = 0.02$; do not reject Ho.

Method (3) : Critical value approach

in the critical value approach, we find the critical value of $\bar{x}$ (or, $\bar{x}_c$) and compare to the sample value of $\bar{x}$. To test the hypothesis, the x_c can be calculated using:

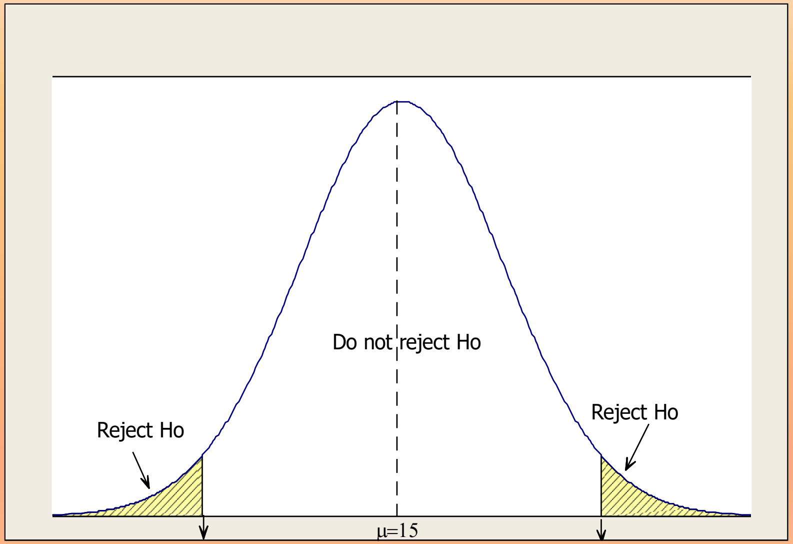
$$\bar{x}_c = \mu \pm z_{critical}\left(\frac{s}{\sqrt{n}}\right)$$

where, Zcritical is the critical point from z-table (see method 1). The critical value is calculated as shown below:

$$\bar{x}_c = 15 \pm (2.33)\frac{5}{\sqrt{50}}$$

$$= 15 \pm 1.65$$

$$= 13.35 \text{ to } 16.65$$



$\bar{x}_c = 13.35$ $\bar{x}_c = 16.65$

The decision rule for the problem:

Reject Ho if $\bar{x} > 16.65$

or, if $\bar{x} < 13.35$

Since, $\bar{x} = 14.2 > \bar{x}_c = 13.35$; do not reject Ho.

Method (4): Confidence Interval Approach

The decision rule using confidence interval approach is given by:

Reject H_o if the hypothesized value is outside of the confidence interval.

The confidence Interval for this problem can be calculated as shown below.

$$\bar{x} \pm z_{critical}\left(\frac{s}{\sqrt{n}}\right) = 14.2 \pm 2.33\left(\frac{5}{\sqrt{50}}\right)$$

or, 14.2 ±1.65

or, 12.55 to 15.85

or, a 98% confidence interval

$$12.55 \le \mu \le 15.85$$

The hypothesized value of $\mu = 15$ is included in the above interval; do not reject Ho.

Example 7 (one-sided test)

A marketing company claims that the mean waiting time of customers calling their call center with technical questions is 1.5 minutes or less. If a longer time is required, additional call attendants will be provided. A sample of 35 calls indicated a sample mean of 1.7 minutes with a standard deviation of 0.4 minutes. Does the call center need additional attendants to justify the claim? The type I error is 0.01.

Solution:

$$H_0 : \mu \leq 1.5 \min.$$

$$H_1 : \mu > 1.5 \min.$$

n=35, $\bar{x}$ =1.7 min., s=0.4 min., α=0.01

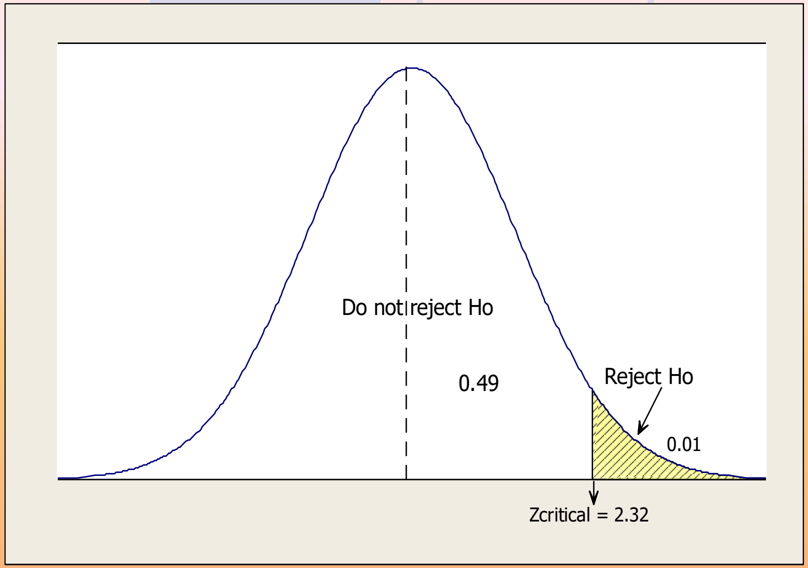
Note that this is a right-sided test. The problem is solved using 4 methods discussed earlier.

Method (1) – z-value approach

The sample size n is large and σ unknown; the test statistic is given by

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

The critical point



Note: z=2.32 is the critical point for a level of significance of 0.01 from the z-table.

Decision rule:

Reject Ho if z >2.32

Calculate the test statistic value as shown

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1.7 - 1.5}{0.4/\sqrt{35}} = 2.96$$

Since, z =2.96 > Zcritical =2.32; reject Ho. The mean time is greater than 1.5 minutes.

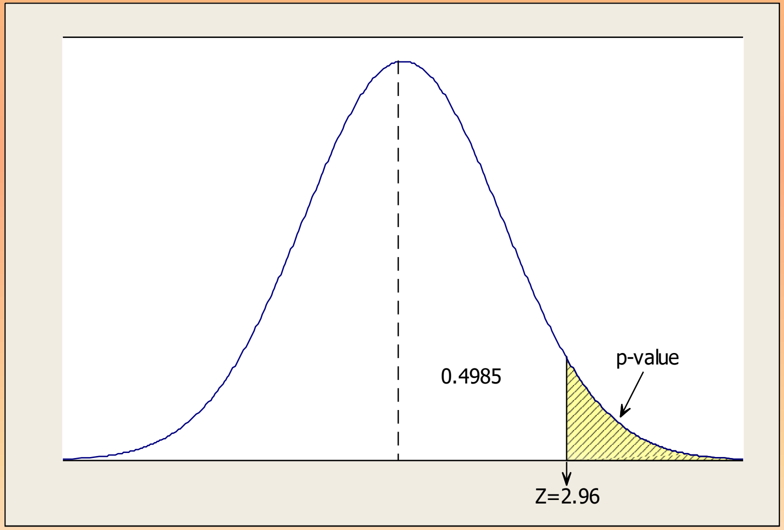
Method (2): *p* - value approach

Decision rule: if $p \geq \alpha$, do not reject H_o.
if $p < \alpha$, reject H_o.

To calculate the p-value, first calculate the value using the appropriate test statistic value. The calculation is shown below.

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1.7 - 1.5}{0.4/\sqrt{35}} = 2.96$$

Next, convert the z value calculated to p-value (see the figure below).



Area corresponding to z =2.96 is 0.4985 (from the z -table)

Probability of z> 2.96= 0.5-0.4985=0.0015

This is the p-value; the area of shaded region in the figure above

Since, p = 0.0015 < $\alpha = 0.01$; reject Ho

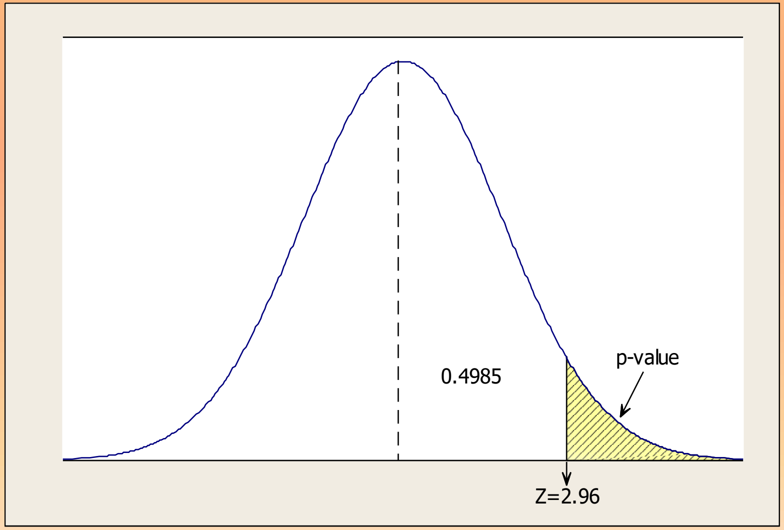
Method (2): *p* - value approach

Decision rule: if $p \geq \alpha$, do not reject H_o.
if $p < \alpha$, reject H_o.

To calculate the p-value, first calculate the value using the appropriate test statistic value. The calculation is shown below.

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1.7 - 1.5}{0.4/\sqrt{35}} = 2.96$$

Next, convert the z value calculated to p-value (see the figure below).



Area corresponding to z =2.96 is 0.4985 (from the z -table)

Probability of z> 2.96= 0.5-0.4985=0.0015

This is the p-value; the area of shaded region in the figure above

Since, $p = 0.0015 < \alpha = 0.01$; reject Ho

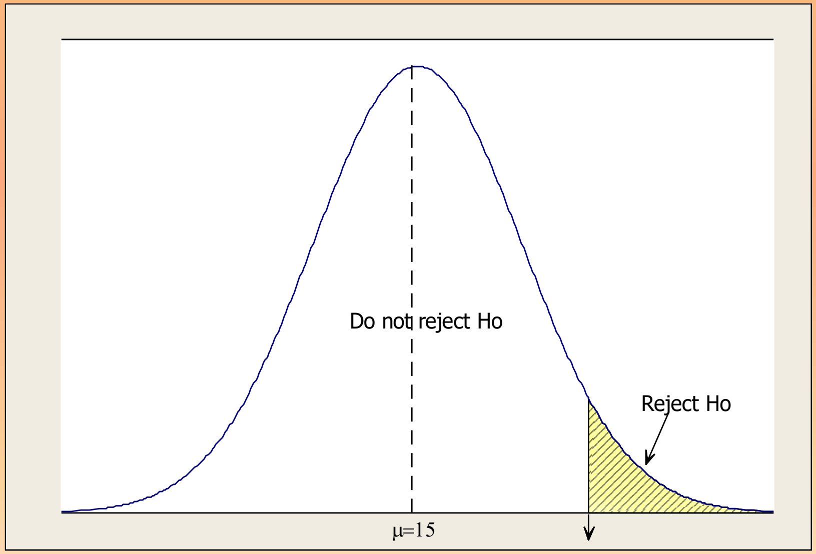
Method (3) : Critical value approach

Calculate $\bar{x}_c$ and compare it to $\bar{x}$ from the sample data. In this case, $\bar{x}_c$ is given by

$$\bar{x}_c = \mu + z_{critical}\left(\frac{s}{\sqrt{n}}\right)$$

$$\bar{x}_c = 1.5 + (2.32)\left(\frac{0.4}{\sqrt{35}}\right) = 1.675$$

$\bar{x}_c$ = 1.657 minute.



$\bar{x}_c = 1.657$

In the figure, the upper critical value, $\bar{x}_c$=16.57. The decision rule is given by

Reject Ho if $\bar{x}$ > 1.657 min.

Since, $\bar{x}$ = 1.7 min. > $\bar{x}_c$ =1.657 min; reject Ho.

Method (4): Confidence Interval Approach

Decision rule for the confidence interval approach is given by:

Reject Ho if the hypothesized value is not included in the interval.

The confidence interval formula for this case

$$\bar{x} \pm z_{critical}\left(\frac{s}{\sqrt{n}}\right) = 1.7 \pm (2.32)\left(\frac{0.4}{\sqrt{35}}\right)$$

or, 1.7 ± 1.57

or, 1.543 to 1.857 min.

or, the confidence interval

$$1.543 \leq \mu \leq 1.857$$

The hypothesized value of μ = 1.5 is outside of the confidence interval above therefore; reject Ho.

[Note: To find the confidence interval for a one-sided test, one way is to double the α.This will not affect the critical point. For this problem, α is doubled and a 98% confidence interval is calculated].

TESTING A MEAN (SMALL SAMPLE USING t-DISTRIBUTION)

Example 8

The mean time to complete an assembly job by an industrial robot is 90 sec. A larger or smaller time will disrupt the assembly line and the line may have to shut down. The production engineer takes data every three hours to determine if the robot is working properly. A sample of 25 jobs provided an average time of 84.50 sec. with a standard deviation of 14.50 sec. Test the appropriate hypothesis using 5% level of significance.

Solution:

$$H_0 : \mu = 90$$

$$H_1 : \mu \neq 90$$

Data for the problem:

n = 25

$\bar{x}$ = 84.50

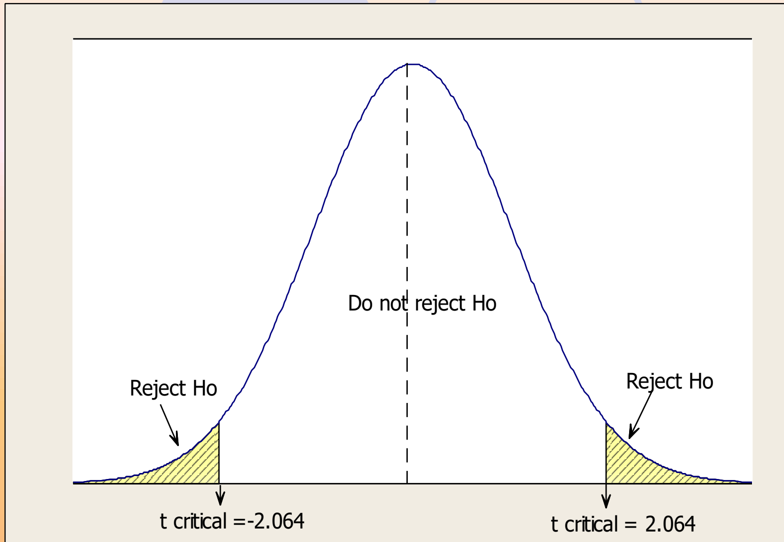
s = 14.50 α= 0.05

Method (1) : t-value approach_ Comparing t-critical value to t-test statistic

Calculate the test statistic value of t and compare it to the critical point obtained from the t-table. The critical point from the t-table

$$t_{n-1,\frac{\alpha}{2}} = t_{24,0.025} = 2.064$$

The critical points and the decision rules are sown in the next slide.



Decision rule:

Reject H_o if $t_{n-1} > 2.064$

or, if $t_{n-1} < -2.064$

Test statistic value

$$t_{n-1} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{84.50 - 90}{14.50/\sqrt{25}} = -1.89$$

$t_{n-1} = -1.89 > t_{crit} = -2.064$; do not reject H_o

Method (2) : Critical value approach

Determine $\bar{x}_c$ and compare it to $\bar{x}$ from the sample data. The $\bar{x}_c$ for this case is determined using

$$\bar{x}_c = \mu \pm t_{n-1,\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$$

Where, $t_{n-1,\frac{\alpha}{2}} = t_{24,0.025} = 2.064$

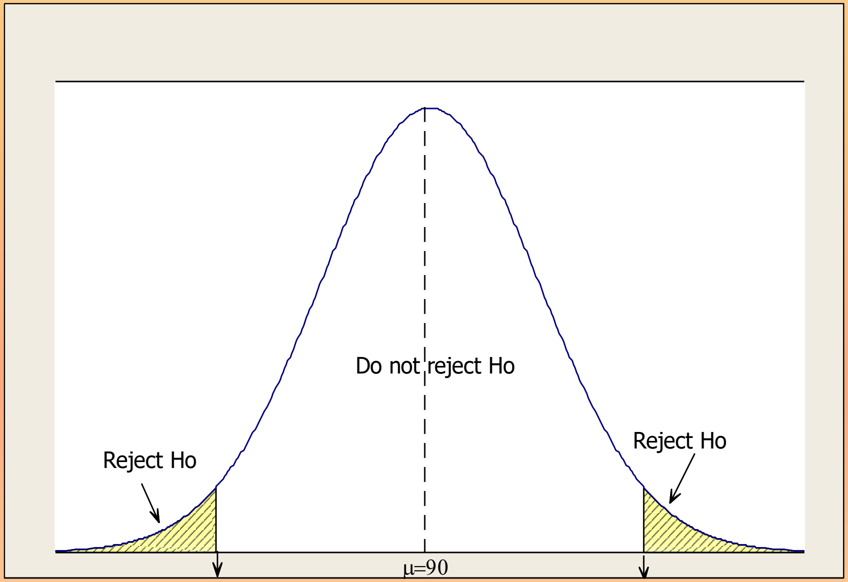
Substituting the values

$$\bar{x}_c = 90 \pm (2.064)\left(\frac{14.50}{\sqrt{25}}\right)$$

$$= 90 \pm 5.99$$

or, 84.01 to 95.99

These critical values along with the decision rule are shown in figure below.



$\bar{x}_c = 84.01$ $\bar{x}_c = 95.99$

Decision rule:

Reject Ho if $\bar{x} > 95.99$

or, if $\bar{x} < 84.01$

Since, $\bar{x} = 84.50 > \bar{x}_c = 84.01$; do not reject Ho.

Method (3): Confidence Interval Approach

Decision rule:

Reject H_o if the hypothesized value is outside of the confidence interval.

The confidence interval for this test is given by

$$\bar{x} \pm t_{n-1,\frac{\alpha}{2}}\left(\frac{s}{\sqrt{n}}\right)$$

$$\bar{x} \pm t_{24,0.05}\left(\frac{s}{\sqrt{n}}\right) = 84.50 \pm (2.064)\left(\frac{14.50}{\sqrt{25}}\right)$$

or, 84.50 ± 5.99

or, 78.51 to 90.49

or, $78.51 \le \mu \le 90.49$

The hypothesized value of μ = 90.0 is included in the above interval; do not reject Ho.

TESTING A SINGLE POPULATION PROPORTION (LARGE SAMPLE)

In this section, we discuss the hypothesis test concerning proportions. A proportion is ratio or fraction, or percentage that indicates the part of the population or sample having a particular trait of interest.

Some examples of proportion are:

(1) a software company claims that its manufacturing simulation software has 12% of the market share;
(2) a public policy department of a large university wants to study the difference in proportion between male and female unemployment rate;
(3) a manufacturing company wants to determine the proportion of defective items produced by its assembly line.

The population proportion will be denoted by “p” whereas, the sample proportion will be denoted by $\overline{p}$.

The basis for the test of proportion is the Binomial distribution. In this section, we will consider the sample size to be large. If the sample size is large the Binomial distribution can be approximated by a Normal distribution and

$$np \geq 5, and$$
$$n(1-p) \geq 5$$

[where, n=sample size, p=probability of success]; In testing a proportion, we will use large sample case so that normal approximation holds

Hypothesis and Test Statistic for Testing Proportions

$H_0 : p = p_0$ $H_1 : p \neq p_0$ ***Two-sided Test***	$H_0 : p \geq p_0$ $H_1 : p < p_0$ ***Left-sided Test***	$H_0 : p \leq p_0$ $H_1 : p > p_0$ ***Right-sided Test***

Test Statistics:

$$z = \frac{\bar{p} - p}{\sqrt{\frac{p\ (1-p\ )}{n}}}$$

$\bar{p}$ = sample proportion, and
p = population proportion or, the hypothesized value.

Example 8: Testing a single population proportion

Let p be the fraction of defective items produced by a certain company. Historical data indicate that 4% of the components manufactured by the company are defective (p=0.04). A new method is proposed that is believed to decrease this percent of defects. Under the new method a sample of n = 500 item had 16 defects. We want to test the hypothesis that the percent of defective items is different from 4% at a 5% level of significance.

The hypothesis:

$$H_0 : p = 0.04$$

$$H_1 : p \neq 0.04$$

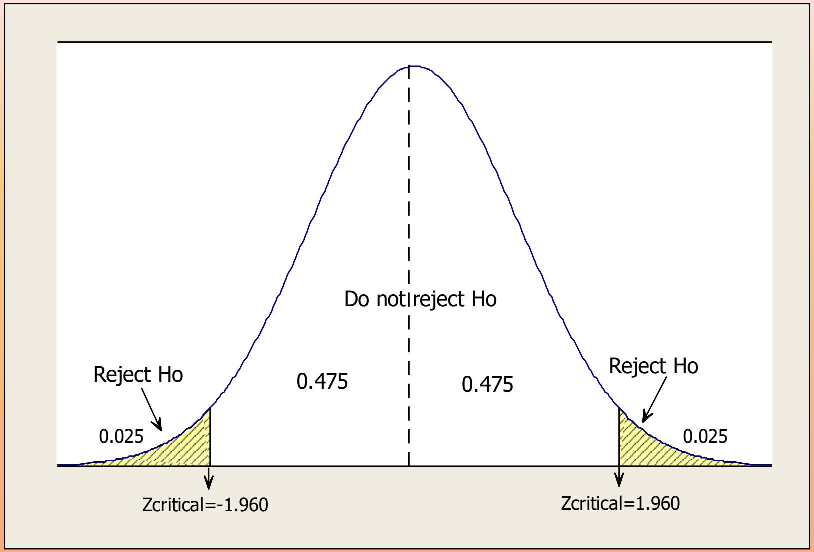
Data: n = 500 x = 16 α = 0.05

The sample proportion can be calculated as

$$\bar{p} = \frac{x}{n} = \frac{16}{500} = 0.032$$

Method (1) – z-value approach

This is a two-sided test. The critical points and the areas of rejection and non-rejection are shown below.



Decision rule:

Reject Ho if > 1.96

or, if < -1.96

Test statistics and its value

$$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

$$= \frac{0.032 - 0.04}{\sqrt{\frac{0.04(1-0.04)}{500}}} = -0.91$$

z = -0.91 > z-critical = -1.96; do not reject Ho.

Method (2): *p* - value approach

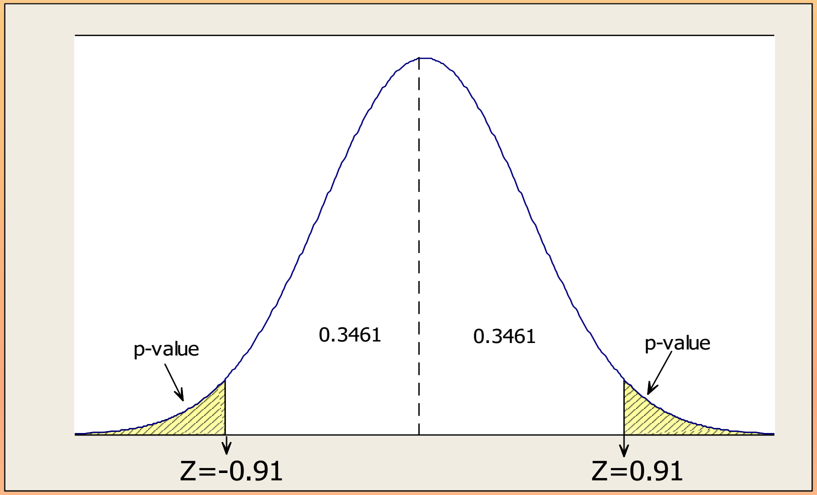
Decision rule:

If p ≥ α, do not reject Ho.

If p < α reject Ho.

To calculate the p-value, calculate the test statistics value first

$$z = \frac{\bar{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.032 - 0.04}{\sqrt{\frac{(0.04)(1-0.04)}{500}}} = -0.91$$



Area corresponding to z = 0.91 is 0.3186 (from the z -table)

Probability of z > + 0.91 = 0.5 - 0.3186 = 0.1814

Probability of z < - 0.91 = 0.5 - 0.3186 = 0.1814

In a two-sided test, the p-value is the sum of areas on both sides.

p -value = 0.1814 +0.1814= 0.3628

Since, p = 0.3628 > α = 0.05; do not reject Ho.

Method (3) : Critical value approach

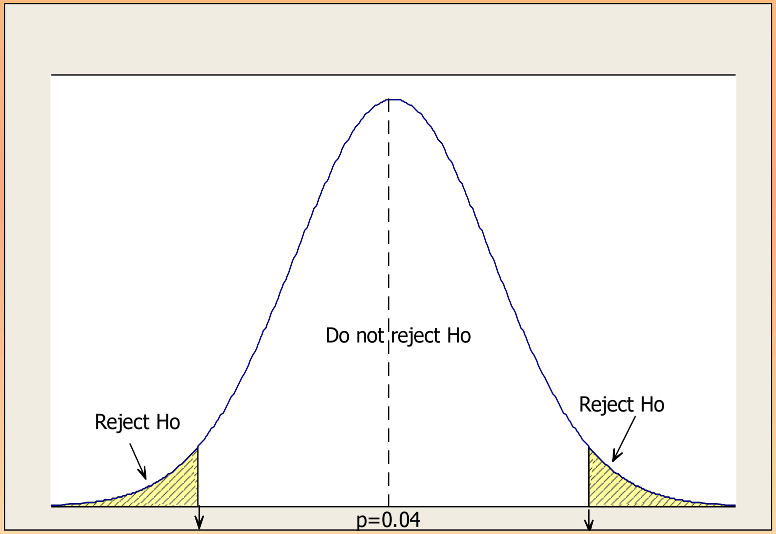
Calculate $\overline{p}_c$ and compare it to the sample proportion, $\overline{p}$

$$\overline{p}_c = p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}$$

$$= 0.04 \pm (1.96)\sqrt{\frac{(0.04)(1-0.04)}{500}}$$

$$= 0.04 \pm 0.02$$

$$= 0.02, 0.06$$



$\overline{p}_c = 0.02$ $\overline{p}_c = 0.06$

Decision rule: Reject Ho if $\overline{p} > 0.06$
or, if $\overline{p} < 0.02$

Since,

$$\overline{p} = 0.032 < \overline{p}_c = 0.06$$

Do not reject H_o.

Method (4): Confidence Interval Approach

Confidence interval formula for this case:

$$\bar{p} \pm Z_{\alpha/2}\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$

Decision rule: Reject H_o if the hypothesized value is not included in the interval.

The confidence interval using the above formula is shown below.

$$0.032 \pm (1.96)\sqrt{\frac{0.032(1-0.032)}{500}}$$

$$= 0.032 \pm 0.015$$

or, 0.032 ± 0.015

or, 0.017 to 0.047

$$0.017 \le p \le 0.047$$

Since the hypothesized value of p = 0.04 is included in the interval; do not reject H_o.

Testing a Single Mean and a Single Proportion: Selecting the Right Test and Right Procedure

Hypothesis Tests for a Single Mean

$H_0 : \mu = \mu_0$
$H_1 : \mu \neq \mu_0$
Two -tailed or Two-sided Test

$H_0 : \mu \geq \mu_0$
$H_1 : \mu < \mu_0$
Left-tailed or Left-sided Test

$H_0 : \mu \leq \mu_0$
$H_1 : \mu > \mu_0$
Right-tailed or Right-sided Test

Case (1) : Sample size (n) large, s known: the sample mean follows a Normal distribution and the test statistics is given by

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Case (2) : Sample size (n) large, s unknown: the sample mean follows a Normal distribution and the test statistics is given by

$$z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Case (3) : Sample size (n) small, s unknown: the sample mean follows a t- distribution and the test statistics is given by

$$t_{n-1} = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Test Methods

***(1) Z-value approach**: calculate the z-value using the test statistics formula and compare it to $Z_{critical}$.*

Decision rules:

Two-sided test	Right-sided test	Left-sided test
Reject H_0 if $Z > Z_{critical}$ or, if $Z < z_{critical}$	Reject H_0 if $Z > Z_{critical}$	Reject H_0 if $Z < Z_{critical}$

Note: If you are using a t-distribution then $z_{critical}$ is replaced by $t_{critical}$ and z is replaced by t_{n-1}.

***(2) p-value approach**: determine the p-value and compare it to the level of significance a. The decision rule is given by:*

Decision rule:

If $p = \alpha$; do not reject H_0
If $p < \alpha$; reject H_0

***(3) Critical value approach:** calculate the critical value of $\bar{x}$ (known as $\bar{x}_c$) and compare it to the $\bar{x}$ of the sample data. Decision rule*

Two-sided test	Right-sided test	Left-sided test
Reject H_0 if $\bar{x} > \bar{x}_c$ or, if $\bar{x} < \bar{x}_c$	Reject H_0 if $\bar{x} > \bar{x}_c$	Reject H_0 if $\bar{x} < \bar{x}_c$

$\bar{x}_c$ is calculated using the following formulas:

$\bar{x}_c = \mu \pm z \frac{\sigma}{\sqrt{n}}$ *For a two-sided test*

$\bar{x}_c = \mu + z \frac{\sigma}{\sqrt{n}}$ *For a right-sided test*

$\bar{x}_c = \mu - z \frac{\sigma}{\sqrt{n}}$ *For a left-sided test*

If you are using t-distribution then z in the above formula, Z is replaced by t_{n-1} and appropriate α , and the σ is replaced by s.

***(4) Confidence interval approach**: Calculate the confidence intervals using appropriate formulas.*
Decision rule: Reject H_0 if the hypothesized value (m) is outside the calculated confidence interval.

Confidence interval formulas:

$\bar{x} \pm z \frac{\sigma}{\sqrt{n}}$ *When n is large and σ is known*

$\bar{x} \pm z \frac{s}{\sqrt{n}}$ *When n is large and σ is unknown*

$\bar{x} \pm t_{n-1, \alpha/2} \frac{s}{\sqrt{n}}$ *When n is small and σ is unknown*

Calculation of type II error β and the power of the test (1-β) See the examples.Relationship between type I error α and β [See the examples]
IV. Sample size for one-sided test

$$n = \frac{(z_\alpha + z_\beta)^2 \sigma^2}{(\mu_0 - \mu_a)^2}$$

Chapter 9 : Introduction to Hypothesis Testing - Flow Chart (1)

Testing a Single Mean and a Single Proportion: Selecting the Right Test and Right Procedure ...continued

Testing a Single Proportion

Hypothesis:

Two-tailed or Two-sided Test	Left-tailed or Left-sided Test	Right-tailed or Right-sided Test
$H_0 : p = p_0$ $H_1 : p \neq p_0$	$H_0 : p \geq p_0$ $H_1 : p < p_0$	$H_0 : p \leq p_0$ $H_1 : p > p_0$

Test Statistics:
If the sample size n is large, the sample proportion $\bar{p}$ follows a Normal distribution and the test statistics is given by

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

$\bar{p}$ *= sample proportion = x /n* p_0 *= population proportion (or, the hypothesized value)*

Test Methods

(1) ***Z-value approach****: calculate the z-value using the test statistics formula and compare it to $Z_{critical}$.*

Decision rule:

Two-sided test	Right-sided test	Left-sided test
Reject H_0 if $Z > Z_{critical}$ or, if $Z < z_{critical}$	Reject H_0 if $Z > Z_{critical}$	Reject H_0 if $Z < Z_{critical}$

(2) p-value approach: determine the p-value and compare it to the level of significance α. The decision rule is given by:

Decision rule:

If $p \geq \alpha$;   do not reject H_0
If $p < \alpha$;   reject H_0

(3) ***Critical value approach****: calculate the critical value of $\bar{p}$ ($\bar{p}_c$) and compare it to the $\bar{p}$ of the sample data.*

Decision rule:

Two-sided test	Right-sided test	Left-sided test
Reject H_0 if $\bar{p} > \bar{p}_c$ or, if $\bar{p} < \bar{p}_c$	Reject H_0 if $\bar{p} > \bar{p}_c$	Reject H_0 if $\bar{p} < \bar{p}_c$

$\bar{p}_c$ *is calculated using the following formulas:*

$$\bar{p}_c = p_0 \pm z\sqrt{\frac{p_0(1 - p_0)}{n}}$$

(For a two-sided test)

Note: for a right sided test, the ± sign in the above formula is replaced by a (+) sign, and for a left-sided test the ± sign is replaced by a negative (-). Also, note that in a two-sided test, $\alpha = \alpha/2$ and for a one sided test, $\alpha = \alpha$.

(4) Confidence interval approach: Calculate the confidence intervals using appropriate formulas.

Decision rule:
Reject H_0 if the hypothesized value (p_0) is outside of the calculated confidence interval.

Confidence interval formula:

$$\bar{p} \pm z\sqrt{\frac{\bar{p}(1 - \bar{p})}{n}}$$

Chapter 9: Hypothesis Testing - Flow Chart (2)