Chapter III

Testing Hypotheses

R. 3. 1. (***Introduction***)
A statistical hypothesis is an assumption about a population parameter. This assumption may or may not be true.
The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are consistent with the statistical hypothesis, the hypothesis is accepted; if not, it is rejected.

This chapter discusses how to make such tests of hypotheses about the population mean, μ, and the population proportion, P.

D. 3. 1. (***Null Hypothesis***)
A *null hypothesis* is a statement about a population parameter that is assumed to be true until it is declared false.
The null hypothesis will be denoted by H_0.

D. 3. 2. (***Alternative Hypothesis***)
An *alternative hypothesis* is a statement about the population parameter that will be true if the null hypothesis is false.
The alternative hypothesis will be denoted by H_1.

R. 3. 2.
In some applications it may not be obvious how the null and alternative hypotheses should be formulated. Care must be taken to structure the hypotheses appropriately so that the hypothesis testing conclusion provides the information the researcher or decision maker wants.
Guidelines for establishing the null and alternative hypotheses will be given for three types of situation in which hypothesis testing procedures are commonly employed:

1. Testing Research Hypotheses
As a general guide, a research hypothesis should be stated as the *alternative hypothesis*.

2. Testing the Validity of a Claim
In any situation that involves testing the validity of a claim, the null hypothesis is generally based on the assumption that the claim is true. The alternative hypothesis is then formulated so that the rejection of the null hypotheses will provide statistical evidence that the stated assumption is incorrect. Action to correct the claim should be considered whenever the null hypothesis is rejected.

3. Testing in Decision-Making
In testing research hypotheses or testing the validity of a claim, action is taken if the hull hypothesis is rejected. In many instances, however, action must be taken both when the null hypothesis cannot be rejected and when it can be rejected. In general, this type of situation occurs when a decision-maker must choose between two courses of action, one associated with the null hypothesis and another associated with the alternative hypothesis.

D. 3. 3. (***Type I Error***)
A *type I error* occurs when a true null hypothesis is rejected. The value of α represents the probability of committing this type of error, that is,

$$\alpha = P\left(H_0 \text{ is rejected/ } H_0 \text{ is true}\right)$$

The value of α represents *significance level* of the test.

R. 3. 3.
Although any value can be assigned to α, the commonly used values of α are 0.01, 0.025, and 0.10. Usually the value assigned to α does not exceed 0.10 (or 10%).

D. 3. 4. (***Type II Error***)
A *type II error* occurs when a false null hypothesis is not rejected. The value of β represents the probability of committing this type of error, that is,

$$\beta = P\left(H_0 \text{ is not rejected/ } H_0 \text{ is false}\right)$$

The value of $1-\beta$ is called *power of the test*. It represents the probability of not making a type II error.

R. 3. 4.
The two types of errors that occur in tests of hypotheses depend on each other. We cannot lower the values of α and β simultaneously for a test of hypothesis for a fixed sample size. Lowering the value of α will raise the value of β, and lowering the value of β will raise the value of α.

R. 3. 5. (***Summary of Error Types***)

		Actual Situation	
		H_0 is true	H_0 is false
Decision	Do not reject H_0	Correct decision	Type II or β error
	Reject H_0	Type I or α error	Correct decision

Ex. 3. 1. (***Justice System - Trial***)

	H_0 true Defendant Innocent	H_0 False Defendant Guilty
Fail to Reject Presumption of Innocence (Not Guilty Verdict)	Correct	Type II Error
Reject Presumption of Innocence (Guilty Verdict)	Type I Error	Correct

D. 3. 5. (***Tails of the Test***)
A *two-tailed test* has rejection regions in both tails, a *left-tailed test* has the rejection region in the left tail, and a *right-tailed test* has the rejection region in the right tail of the distribution curve.

R. 3. 6.
Whether a test is two-tailed or one-tailed is determined by the sign in the alternative hypothesis. If the alternative hypothesis has a not equal to ($\neq$) sign, it is a two tailed test.

R. 3. 7.
The following table summarises the discussion about the relation between the signs in H_0 and H_1:

	Two-Tailed Test	Left-Tailed Test	Right-Tailed Test
Sign in the null hypothesis H_0	$=$	$=$ or $\geq$	$=$ or $\leq$
Sign in the alternative hypothesis H_1	$\neq$	$<$	$>$
Rejection region	In both tails	In the left tail	In the right tail

R. 3. 8. (***Test Procedures***)
In this chapter we will use the following two procedures to make tests of hypothesis:

1. ***The $p-$value approach***
   Under this procedure, we calculate what is called the $p-$value for the observed value of the sample statistic. If we have a predetermined significance level, then we compare the $p-$value with this significance level and make a decision (p stands for probability.)

2. ***The critical-value approach***
   In this approach, we find the critical value(s) from a table (such as the normal distribution or the t distribution table) and find the value of the test statistic for the observed value of the sample statistic. Then we compare these two values and make a decision.

R. 3. 9. (***Hypothesis Tests About μ : σ Known***)
We distinguish three possible cases:

Case I. If the following conditions are fulfilled:

1. The sample size is small (i.e., $n < 30$)
2. The population from which the sample is selected is normally distributed,

then we use the normal distribution to perform a test of hypothesis about μ because from Chapter 1 the sampling distribution of $\bar{x}$ is normal with its mean equal to μ and the standard deviation equal to $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ assuming that $\frac{n}{N} \leq 0.05$.

Case II. If the following condition is fulfilled:

The sample size is large (i.e. $n \geq 30$),

then, again, we use the normal distribution to perform a test of hypothesis about μ because from Chapter 1, due to the central limit theorem, the distribution of $\bar{x}$ is (approximately normal with its mean equal to μ and the standard deviation equal to $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ assuming that $\frac{n}{N} \leq 0.05$.

Case III. If the following two conditions are fulfilled:

1. The sample size is small (i.e., $n < 30$)
2. The population from which the sample is selected is not normally distributed or its distribution is unknown),

then we use a nonparametric method to perform a test of hypothesis about μ. Such a procedure is not covered in this chapter.

D. 3. 6. **(** $p-$ ***Value*****)**
The $p-value$ is the smallest significance level at which the null hypothesis is rejected.

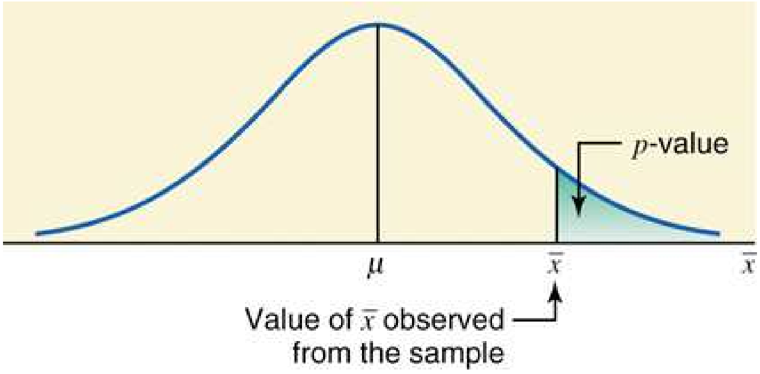
R. 3. 10. **(*****The*** $p-$ ***Value Approach*****)**
Using the $p-$value approach, we reject the null hypothesis if

$$p-value < \alpha$$

and we do not reject the null hypothesis if

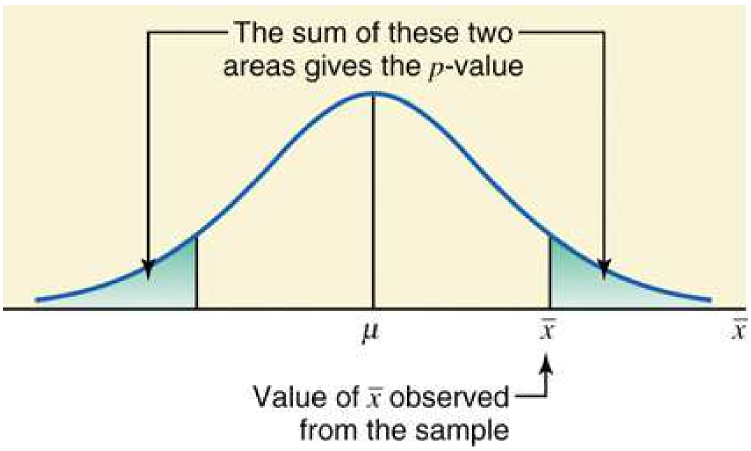
$$p-value \geq \alpha$$

For a one-tailed test, the $p-$value is given by the area in the tail of the sampling distribution curve beyond the observed value of the sample statistic. The following figure shows the $p-$value for a right-tailed test about μ.



For a left-tailed test, the $p-$value will be the area in the lower tail of the sampling distribution curve to the left of the observed value of $\bar{x}$.

For a two-tailed test, the $p-$value is twice the area in the tail of the sampling distribution curve beyond the observed value of the sample statistic. The following figure shows the $p-$value for a two-tailed test. Each of the areas in the two tails gives one-half the $p-$value.



Then we find the area under the tail of the normal distribution curve beyond this value of z. This area gives the $p-$value or one-half of the $p-$value depending on whether it is a one-tailed test or a two-tailed test.

R. 3. 11. (***Steps to Perform a Test of Hypothesis Using the $p-$Value Approach)***

1. State the null and alternative hypothesis.
2. Select the distribution to use.

   When using the normal distribution *the value of* z for $\bar{x}$ for a test of hypothesis about μ is computed as follows:

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \qquad \text{where} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.$$

The value of z calculated for $\bar{x}$ using this formula is also called *observed value of* z.

3. Calculate the $p-v$alue.

4. Make a decision:

   Using the $p-$value approach, we reject the null hypothesis if

$$p - value < \alpha$$

   and we do not reject the null hypothesis if

$$p - value \geq \alpha$$

<u>Ex. 3. 2.</u>
A consumer advocacy group suspects that a local supermarket's 10-ounce packet of cheddar cheese actually weighs less than 10 ounces. The group took a random sample of 30 such packages and found the mean weight for the sample was 9.965 ounces. The population follows a normal distribution with the population standard deviation of 0.15 ounces.
Test the consumer advocacy group's suspicion at the significance level 0.01.

Solution:

$$n = 30, \quad \bar{x} = 9.965, \quad \sigma = 0.15, \quad \alpha = 0.01.$$

1.

$$H_0: \mu = 10; \quad H_1: \mu < 10. \quad .$$

2. Use the normal distribution.

3.

$$z = \frac{9.965 - 10}{\frac{0.15}{\sqrt{30}}} \approx -1.28,$$

$$p - value = 1 - 0.8997 = 0.1003.$$

4.

$$0.1003 > 0.01.$$

$\therefore$ We do not reject H_0.

<u>Ex. 3. 3.</u>
The manufacturer of a certain brand of car batteries claims that mean life of these batteries is 45 months. A consumer protection agency that wants to check this claim took a random sample of 24 such batteries and found that the mean life for this sample is 43.05 months. The lives of all such batteries have a normal distribution with the population standard deviation of 4.5 months.

Test the manufacturer's claim at the 0.025 degree of significance.

Solution:

$$n = 24, \quad \bar{x} = 43.05, \quad \sigma = 4.5, \quad \alpha = 0.025.$$

1.
$H_0: \mu = 45; \quad H_1: \mu < 45.$.

2. Use the normal distribution.

3.
$$z = \frac{43.05 - 45}{\frac{4.5}{\sqrt{24}}} \approx -2.12,$$

$$p - value = 1 - 0.9830 = 0.0170.$$

4.
$0.017 < 0.025.$

$\therefore$ We reject H_0.

Ex. 3. 4.
In the past, the mean running time for a certain type of flashlight battery has been 8.5 hours. The manufacturer has introduced a change in the production method which he hopes has increased the mean running time. The mean running time for a random sample of 40 light bulbs was 8.7 hours.
Do the data provide sufficient evidence to conclude that the mean running time of all light bulbs has increased from the previous mean of 8.5 hours?
Perform the appropriate hypothesis test using a significance level of 0.05. Assume that the standard deviation of the running time of all light bulbs is 0.5 hours.

Solution:

$$n = 40, \quad \bar{x} = 8.7, \quad \sigma = 0.5, \quad \alpha = 0.05.$$

1.
$H_0: \mu = 8.5; \quad H_1: \mu > 8.5.$.

2. Use the normal distribution.

3.
$$z = \frac{8.7 - 8.5}{\frac{0.5}{\sqrt{40}}} \approx 2.53,$$

$$p - value = 1 - 0.9943 = 0.0057.$$

4.

$$0.0057 < 0.05.$$

$\therefore$ We reject H_0. At the 5% significance level, the data provide sufficient evidence to conclude that the mean running time of all light bulbs has increased from the previous mean of 8.5 hours.

R. 3. 12. (***The Critical Value Approach***)
This is the traditional or classical approach. In this procedure, we have a predetermined value of the significance level α. The value of α gives the total area of the rejection region(s). First we find the critical value(s) of z from the normal distribution table. Then we find the value of the test statistic z for the observed value of the sample statistic $\bar{x}$. Finally we compare these two values and make a decision.

R. 3. 13. (***Steps to Perform a Test of Hypothesis with the Critical-Value Approach)***

1. State the null and alternative hypothesis.
2. Select the distribution to use.
3. Determine the rejection and nonrejection regions.
4. Calculate the value of the test statistic.

   In tests of hypothesis about μ using the normal distribution, the random variable

   $$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} \qquad \text{where} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

   is called the *test statistic.*
   (The test statistic can be defined as a rule that is used to make the decision whether or not to reject the null hypothesis.)

5. Make a decision.

Ex. 3. 5. (***See: Ex. 3. 2.***)
Solution:

$$n = 30, \qquad \bar{x} = 9.965, \qquad \sigma = 0.15, \qquad \alpha = 0.01.$$

1.

$$H_0:\ \mu = 10; \qquad H_1:\ \mu < 10. \qquad .$$

2. Use the normal distribution.

3.

$$z_{-0.01} = -2.326$$

4.

$$z = \frac{9.965 - 10}{\frac{0.15}{\sqrt{30}}} \approx -1.28$$

5.

$$-1.28 > -2.326.$$

$\therefore$ We do not reject H_0.

Ex. 3. 6. (***See: Ex. 3. 3.***)
Solution:

$$n = 24, \quad \bar{x} = 43.05, \quad \sigma = 4.5, \quad \alpha = 0.025.$$

1.

$$H_0: \mu = 45; \quad H_1: \mu < 45. \qquad .$$

2. Use the normal distribution.

3.

$$z_{-0.025} = -1.96$$

4.

$$z = \frac{43.05 - 45}{\frac{4.5}{\sqrt{24}}} \approx -2.12$$

5. $-2.12 < -1.96$

$\therefore$ We reject H_0.

Ex. 3. 7. (***See: Ex. 3. 4.***)
Solution:

$$n = 40, \quad \bar{x} = 8.7, \quad \sigma = 0.5, \quad \alpha = 0.05.$$

1.

$$H_0: \mu = 8.5; \quad H_1: \mu > 8.5. \qquad .$$

2. Use the normal distribution.

3.

$$z_{0.05} = 1.645$$

4.

$$z = \frac{8.7 - 8.5}{\frac{0.5}{\sqrt{40}}} \approx 2.53.$$

5.

$$2.53 > 1.645$$

$\therefore$ We reject H_0. At the 5% significance level, the data provide sufficient evidence to conclude that the mean running time of all light bulbs has increased from the previous mean of 8.5 hours.

<u>R. 3. 14.</u> (*Hypothesis Tests About* μ *:* σ *Not Known*)
We distinguish three possible cases:

Case I. If the following conditions are fulfilled:

1. The sample size is small (i.e., $n < 30$)
2. The population from which the sample is selected is normally distributed,

then we use the t distribution to perform a test of hypothesis about μ.

Case II. If the following condition is fulfilled:

The sample size is large (i.e., $n \geq 30$),

then, again, we use the t distribution to perform a test of hypothesis about.

Case III. If the following two conditions are fulfilled:

1. The sample size is small (i.e., $n < 30$)
2. The population from which the sample is selected is not normally distributed or its distribution is unknown),

then we use a nonparametric method to perform a test of hypothesis about μ. Such a procedure is not covered in this chapter.

<u>R. 3. 15.</u> (*Test Statistic*)

The value of the *test statistic* t for the sample mean $\bar{x}$ is computed as

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} \quad \text{where} \quad s_{\bar{x}} = \frac{s}{\sqrt{n}}.$$

The value of t calculated for $\bar{x}$ by using this formula is also called the *observed value* of t.

<u>R. 3. 16.</u>
The two procedures, the $p-$approach and the critical-value approach, are also used to test hypotheses about μ when σ is not known.

The steps used in these procedures are the same as in the case of σ being known. The only difference is that we will be using the t distribution in place of the normal distribution.

<u>Ex. 3. 7.</u>
The manager of a restaurant in a large city claims that waiters working in all restaurants in this city earn an average of \$150 or more in tips per week. A random sample of 25 waiters selected from restaurants of this city yielded a mean of \$139 in tips per week with a standard

deviation of \$28. Assume that the weekly tips for all waiters in this city have a normal distribution.
Using the 1% significance level, can you conclude that the manager's claim is true?

Solution:

$$n = 25, \quad \bar{x} = 139, \quad s = 28, \quad \alpha = 0.01.$$

1.

$$H_0: \mu \geq 150; \qquad H_1: \mu < 150.$$

2.
Use *t* distribution.

3.

$$t = \frac{139 - 150}{\frac{28}{\sqrt{25}}} = -1.964, \qquad p - \text{value}=0.0306.$$

4.

$$p - \text{value}=0.0306 \geq 0.01.$$

$\because$ Do not reject H_0.

Ex. 3. 8.
A company produces car batteries. The company claims that its top-of-the-line batteries are good, on average, for at least 65 months. A consumer protection agency tested 45 such batteries to check this claim. It found that the mean life of these 45 batteries is 63.4 months and the standard deviation is 3 months.
Find the $p-$value for the test that the mean life of all such batteries is less than 65 months. What will your conclusion be if the significance level is 2.5%.

Solution:

$$n = 45, \quad \bar{x} = 63.4, \quad s = 3, \quad \alpha = 0.025$$

1.

$$H_0: \mu \geq 65; \qquad H_1: \mu < 65.$$

2.
Use *t* distribution.

3.

$$t = \frac{63.4 - 65}{\frac{3}{\sqrt{45}}} = -3.578, \qquad p - \text{value} = 0.00045.$$

4.

$$p - \text{value} = 0.00045 < 0.025 = \alpha$$

$\because$ We reject H_0 and conclude that the mean life of such batteries is less than 65 months.

R. 3. 17. (***The Critical Value Approach***)
In this procedure, we have a predetermined value of the significance level α. The value of α gives the total area of the rejection region(s). First we find the critical value(s) of t from the t distribution table for the given degree of freedom and the significance level. Then we find the value of the test statistic t for the observed value of the sample statistic $\bar{x}$. Finally we compare these two values and make a decision.
Remember, if the test is one-tailed, there is only one critical value of t and it is obtained by using the value of α, which gives the area in the left or right tail of the t distribution curve depending on whether the test is left-tailed or right-tailed, respectively. However, if the test is two-tailed, there are two critical values of t and they are obtained by using $\alpha/2$ area in each tail of the t distribution curve. The value of the test statistic t is obtained as mentioned before.

Ex. 3. 9. (***See: Ex. 3. 7.***)
Solution:

$$n = 25, \quad \bar{x} = 139, \quad s = 28, \quad \alpha = 0.01.$$

1.

$$H_0: \mu \geq 150; \qquad H_1: \mu < 150.$$

2.
Use t distribution.

3.

$$t_{24;0.01} = -2.492$$

4.

$$t = \frac{139 - 150}{\frac{28}{\sqrt{25}}} = -1.964.$$

5.

$$-1.964 > -2.492.$$

$\because$ Do not reject H_0.

Ex. 3. 10. (***See: Ex. 3. 8.***)

$$n = 45, \quad \bar{x} = 63.4, \quad s = 3, \quad \alpha = 0.025$$

1.

$$H_0:\ \mu \geq 65;\qquad H_1:\ \mu < 65.$$

2.
Use t distribution.

3.

$$t_{44;0.025} = -2.015$$

4.

$$t = \frac{63.4 - 65}{\frac{3}{\sqrt{45}}} = -3.578$$

5.

$$-3.578 < -2.015.$$

$\because$ Reject H_0.

<u>R. 3. 18.</u> (*Tests about a Population Proportion*)
The following section presents the procedures to perform tests of hypothesis about the population proportion, P, for large samples.

<u>R. 3. 19.</u> (*Test Statistic*)
The value of the *test statistic* z for the sample proportion, p, is computed as

$$z = \frac{p - P}{\sigma_p} \quad \text{wehere} \quad \sigma_p = \sqrt{\frac{P \cdot (1 - P)}{n}}.$$

The value of P used in this formula is the one used in the null hypothesis. The value of z calculated for p using the above formula is called the *observed value of* z.

<u>R. 3. 20.</u> (*The* $p-$*Value Approach*)
To use the $p-$value approach to perform a test of hypothesis about P, we will use the same four steps that we described in R. 3. 11.

<u>Ex. 3. 11.</u>
A large gas company claims that 80 percent of its 1000000 customers are very satisfied with the service they receive. To test this claim, a consumer advocacy group surveyed 100 customers. Among the sampled customers, 73 percent say they are very satisfied.
Based on these findings, can we reject the company's claim? Use a 0.05 level of significance.

Solution:

$$n = 100, \quad p = 0.73, \quad \alpha = 0.05, \quad P = 0.80.$$

1.

$H_0: P = 0.80$; $\qquad\qquad H_1: P \neq 0.80.$

2.

$n.P = 100 \cdot 0.80 > 5 \quad \wedge \quad n \cdot (1 - P) = 100 \cdot 0.20 > 5$

Use the normal distribution.

3.

$$z = \frac{p - P}{\sigma_p} = \frac{0.73 - 0.80}{\sqrt{\frac{0.80 \cdot 0.20}{100}}} = -1.75 \,. \quad p - value = 2 \cdot (1 - 0.9599) = 2 \cdot 0.0401 = 0.0802$$

4.

$p - value - 0.08 > 0.05.$

$\because$ Do not reject H_0.

<u>R. 3. 21.</u> (*The Critical Value Approach*)
In this procedure, we have a predetermined value of the significance level α. The value of α gives the total area of the rejection region(s). First we find the critical value(s) of z from the normal distribution table for the given significance level. Then we find the value of the test statistic z for the observed value of the sample statistic p. Finally we compare these two values and make a decision.

Remember, if the test is one-tailed, there is only one critical value of z and it is obtained by using the value of α, which gives the area in the left or right tail of the normal distribution curve depending on whether the test is left-tailed or right-tailed, respectively. However, if the test is two-tailed, there are two critical values of z and they are obtained by using $\alpha/2$ area in each tail of the normal distribution curve. The value of the test statistic z is obtained as mentioned before.

<u>Ex. 3. 12.</u> (*See: Ex. 3. 11.*).
Solution:

$$n = 100, \quad p = 0.73, \quad \alpha = 0.05, \quad P = 0.80.$$

1.

$H_0: P = 0.80$; $\qquad\qquad H_1: P \neq 0.80.$

2.

$n.P = 100 \cdot 0.80 > 5 \quad \wedge \quad n \cdot (1 - P) = 100 \cdot 0.20 > 5$

Use the normal distribution.

3.

$$z_c = \pm 1.960$$

4.

$$z = \frac{p - P}{\sigma_p} = \frac{0.73 - 0.80}{\sqrt{\frac{0.80 \cdot 0.20}{100}}} = -1.75 \ .$$

5.

$$-1.75 > -1.96 \ \wedge \ 1.75 < 1.96 \ .$$

$\because$ Do not reject H_0.

(Last updated: 08.06.12)