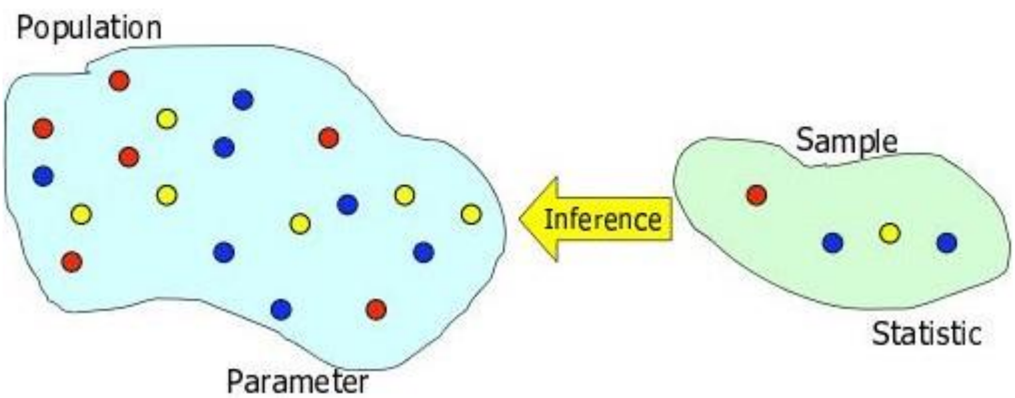
Chapter Eight

8. Estimation and Hypothesis Testing

Statistical inference is the act of generalizing from the data (“sample”) to a larger phenomenon (“population”) with calculated degree of certainty. The act of **generalizing** and deriving statistical judgments is the process of inference.



There are two types of statistical inference:

- Estimation
- Hypotheses Testing

The concepts involved are actually very similar, which we will see in due course. Below, we provide a basic introduction to estimation.

8.1 Introduction to Estimation

One aspect of inferential statistics is **estimation,** which is the process of estimating the value of a parameter from information obtained from a sample. For example, consider the following statements:

- *“Eight percent of the people surveyed in the United States said that they participate in skiing in the winter time.” (IMRE sports).*
- *“Consumers spent an average of $126 for Valentine’s Day this year.” (Nationa1 Retail Federation).*

- *"The average amount spent by a TV Super Bowl viewer is \$63.87." (Retail Advertising and Marketing Association).*

Since the populations from which these values were obtained are large, these values are only *estimates* of the true parameters and are derived from data collected from samples. In other words, inferential statistics uses the sample results to make decisions and draw conclusions about the population from which the sample is drawn.

Statistical Estimation

Definition; In statistics, **estimation** refers to the process by which one makes inferences about a population parameter, based on information obtained from a sample. In inferential statistics, μ is called the true population mean (population parameter). There are many other population parameters, such as the median, mode, proportion, variance, and standard deviation.

Estimate and Estimator

The value(s) assigned to a population parameter based on the value of a sample statistic is called an **estimate**. The sample statistic used to estimate a population parameter is called an **estimator**.

***Example 1*:**

An example of population parameters and their corresponding estimator;

Parameter	Estimator
Population Mean, μ	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Population variance, σ^2	$s^2 = \frac{(x_i - \bar{x})^2}{n-1}$
Population S.D, σ	$s = \sqrt{\frac{(x_i - \bar{x})^2}{n-1}}$
Population proportion, p	$\hat{p} = \frac{x}{n}$

Population parameters can have more than one estimator, but one **best** estimator.

You might ask why other measures of central tendency, such as the median and mode, are not used to estimate the population mean. The reason is that the means of samples vary less than other statistics (such as medians and modes) when many samples are selected from the same population. Therefore, the sample mean is the **best** estimate of the population mean.

Sample measures (i.e., statistics) are used to estimate population measures (i.e., parameters). These statistics are called **estimators.** As previously stated, the sample mean is a **better** estimator of the population mean than the sample median or sample mode.

Example 2

Population mean μ can be estimated by either of: $\bar{x}$ sample mean: $\tilde{x}$ *sample median* or $\hat{x}$ *sample* mod*e*. But to be **best** estimator, the sample statistic should be as close to the true value of the parameter as possible.

8.1.1 Properties of best estimator

An estimator or statistic is considered to be best estimator if it satisfies the following three properties.

- Un-biasedness
- Consistency
- Efficiency

i. Un-biasedness

An unbiased estimator of a population parameter is an estimator whose expected value is equal to that parameter. Formally, an estimator $\widehat{\boldsymbol{\Theta}}$ for parameter θ is said to be unbiased if:

$$E[\widehat{\boldsymbol{\Theta}}] = \theta$$

In other words, an estimator is unbiased if it produces parameter estimates that are on average correct. For instance, the sample mean $\boldsymbol{\bar{x}}$ is an unbiased estimator for the population mean μ, since;

$$E[\bar{\mathrm{x}}] = E[\hat{\mu}] = \mu$$

Example 3

A population of five units has the observations 7, 6, 8, 4 and 10. Random samples of size 3 are drawn, and the possible number of samples is $\binom{5}{3} = 10$ without replacement, as given below.

Sample number	Samples	Means
1	7,6,8	$\bar{x}_1 = 7$
2	7,6,4	$\bar{x}_2 = 17/3$
3	7,6,10	$\bar{x}_3 = 23/3$
4	7,8,4,	$\bar{x}_4 = 19/3$
5	7,8,10	$\bar{x}_5 = 25/3$
6	7,4,10	$\bar{x}_6 = 7$
7	6,8,4	$\bar{x}_7 = 6$
8	6,8,10	$\bar{x}_8 = 8$
9	6,4,10	$\bar{x}_9 = 20/3$
10	8,4,10	$\bar{x}_{10} = 22/3$
	Mean	$\bar{\boldsymbol{x}} = \mathbf{7}$

$$Population\ Mean = \mu = \frac{4+6+7+8+10}{5} = 7 = E[\bar{\boldsymbol{x}}]$$

$$\Rightarrow thesample\ mean\ \bar{\boldsymbol{x}}\ is\ unbiased\ estimatore\ of\ population\ mean\ \mu$$

ii. Consistency

An unbiased estimator is said to be consistent if the difference between the estimator and the target population parameter becomes smaller as we increase the sample size. Formally, an unbiased estimator $\widehat{\boldsymbol{\Theta}}$ for parameter θ is said to be consistent if V $(\widehat{\boldsymbol{\Theta}})$ approaches zero as n → ∞

Note that being unbiased is a precondition for an estimator to be consistent.

Example 4

The variance of the sample mean $\bar{x}$ is $\sigma^2/_n$, which decreases to zero as we increase the sample size n.

Hence, the sample mean is a consistent estimator for μ

For any distribution, sample mean $\bar{\boldsymbol{x}}$ is a consistent estimator of the population mean, sample proportion $\hat{p}$ is a consistent estimator of population proportion P and sample variance s^2 is a consistent estimator of the population variance σ^2.

iii. Efficiency

An estimator is said to be efficient if its value remains the same or stable as the sample size changes. This property is called sampling stability.

The estimator should be a **relatively efficient estimator.** That is, of all the statistics that can be used to estimate a parameter, the relatively efficient estimator has the smallest variance.

In other word, suppose we are given two unbiased estimators for a parameter. Then, we say that the estimator with a smaller variance is more efficient.

Example 5

For a normally distributed population, it can be shown that the sample median is an unbiased estimator for μ. It can also be shown, however, that the sample median has a greater variance than that of the sample mean, for the same sample size. Hence, sample mean $\bar{x}$ is a more efficient estimator than sample median.

iv. Sufficiency

A statistic is said to be a sufficient estimator of the population parameter if it contains all the information in the sample regarding the parameter. In other words, a sufficient statistic utilizes all the information that a given sample can furnish about the parameter.

Example 6

Sample mean is the most sufficient estimator because it takes in to account all the information in the population.

8.1.2 Types of Estimation

In statistics there are two types of estimation these are:

1. Point Estimation and
2. Interval Estimation

1. Point Estimation

A **point estimate** of some population parameter θ is a single numerical value $\hat{\theta}$ of a statistic $\widehat{\boldsymbol{\Theta}}$**.** The statistic $\widehat{\boldsymbol{\Theta}}$ is called the point **estimator**.

As an example, suppose that the random variable *X* is normally distributed with an unknown mean μ. The sample mean is a point estimator of the unknown population mean μ. That is, $\hat{\mu} = \bar{X}$. After the sample has been selected, the numerical value $\bar{x}$ is the point estimate ofμ. Thus, if $x_1 = 25$, $x_2 = 30$, $x_3 = 29, and\ x_4 = 31$, the point estimate of μ is

$$\bar{x} = \frac{25 + 30 + 29 + 31}{4} = 28.75$$

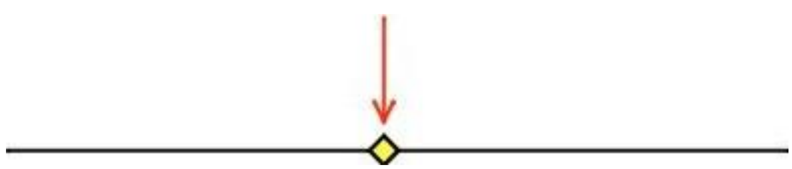
$$\Rightarrow \hat{\mu} = \bar{X} = 28.75$$

Similarly, if the population variance σ^2. is also unknown, a point estimator for σ^2 is the sample variance s^2, and the numerical value $s^2 = 6.9$ calculated from the sample data above is called the point estimate of σ^2.

Reasonable point estimates of these parameters are as follows:

- For, *populations mean* μ the estimate $\boldsymbol{\hat{\mu} = \bar{X},}$ the value of sample mean.
- For $population\ variance\ \ \sigma^2$, the estimate $\boldsymbol{\hat{\sigma}^2 = s^2}$, the value of sample variance.
- For population proportion *p*, the estimate $\boldsymbol{\hat{p} = {}^{x}/_{n}}$, the value of sample proportion, where *x* is the number of items in a random sample of size *n* that belong to the class of interest.

Point estimator



A **point estimator** draws inferences about a population by estimating the value of an unknown parameter using a single value or point.

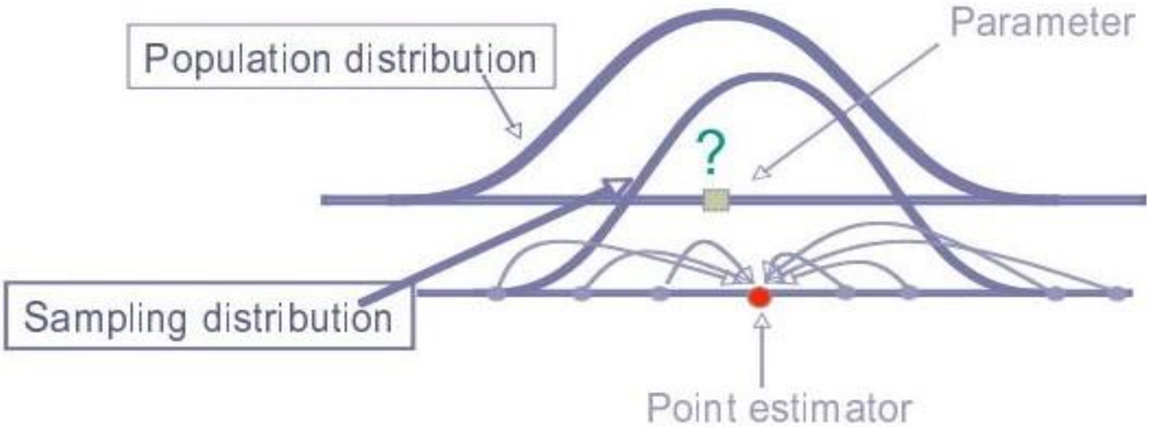


Figure 8.1: point estimator for unknown population parameter

2. Interval Estimation

Although point estimate $\bar{x}$ is a valuable reflection of parameter μ, it provides no information about the precision of the estimate. We ask: How precise is $\bar{x}$ as estimate of μ? How much can we expect any given $\bar{x}$ to vary from μ?

❖ Confidence Intervals

An **interval estimator** draws inferences about a population by estimating the value of an unknown parameter using an interval. Here, we try to construct an interval that "covers" the true population parameter with a specified probability. The range of the interval would depend up on the probability with which the population parameter is expected to fall in the range.

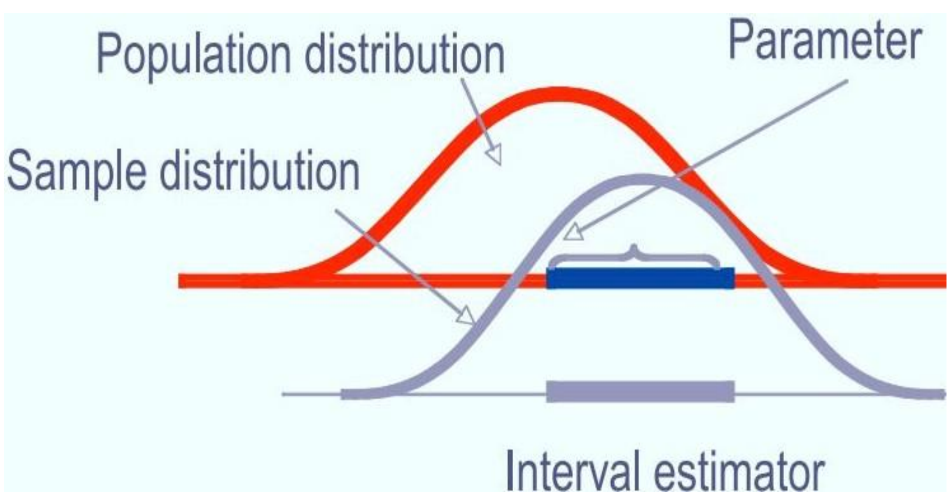


Figure 8.2: interval estimator for unknown population parameter

Given two values T1 and T2, we can always determine the probability that the interval (T1, T2) contains the parameter. In general, $P\,(T1,\,T2) = 1-\alpha$. Where $\alpha =$ the probability that the parameter may not be contained in the interval. Usually called **level of significance**

As an example, suppose we are trying to estimate the mean summer income of students. Then, an interval estimate might say that the (unknown) mean income is between $380 and $420 with probability 0.95.

i.*e. A 100 (1-α) % CI = 95% an **interpretation** also be, we are 95% confident that the average summer income of students to fall in between $380 and $420.*

Statisticians use a **confidence interval** to express the precision and uncertainty associated with a particular sampling method.

A confidence interval consists of three parts.

- A confidence level.
- A statistic.
- A margin of error.

The confidence level describes the uncertainty of a sampling method. The statistic and the margin of error define an interval estimate that describes the precision of the method. The interval estimate of a confidence interval is defined by;

the sample statistic ± margin of error.

For example, suppose we compute an interval estimate of a population parameter. We might describe this interval estimate as a 95% confidence interval. This means that if we used the same sampling method to select different samples and compute different interval estimates, the true population parameter would fall within a range defined by the *sample statistic ± margin of error* 95% of the time.

Confidence intervals are preferred to point estimates, because confidence intervals indicate

- the precision of the estimate and
- the uncertainty of the estimate.

i. Confidence Level

The probability part of a confidence interval is called a **confidence level**. The confidence level describes the likelihood that a particular sampling method will produce a confidence interval that includes the true population parameter.

Here is how to interpret a confidence level. Suppose we collected all possible samples from a given population, and computed confidence intervals for each sample. Some confidence intervals would include the true population parameter; others would not. A 95% confidence level means that 95% of the intervals contain the true population parameter; a 90% confidence level means that 90% of the intervals contain the population parameter; and so on.

ii. A statistic

A statistic means here, it means that any point estimator of a population parameter. *For example* $\bar{x}$.

iii. Margin of Error

In a confidence interval, the range of values above and below the sample statistic is called the **margin of error**.

$$ME = Z\alpha/_2\sigma_{\bar{x}} = Z\alpha/_2\sqrt{\frac{\sigma^2}{n}} = Z\alpha/_2\frac{\sigma}{\sqrt{n}} \text{ , since } \sigma^2{}_{\bar{x}} = \frac{\sigma^2}{n}$$

For example, suppose the local newspaper conducts an election survey and reports that the independent candidate will receive 30% of the vote. The newspaper states that the survey had a 5% margin of error and a confidence level of 95%. These findings result in the following confidence interval: We are 95% confident that the independent candidate will receive between 25% and 35% of the vote.

<u>Note</u>: Many public opinion surveys report interval estimates, but not confidence intervals. They provide the margin of error, but not the confidence level. To clearly interpret survey results you need to know both! We are much more likely to accept survey findings if the confidence level is high (say, 95%) than if it is low (say, 50%).

To gain further insight into μ, we surround the point estimate with a **margin of error**:

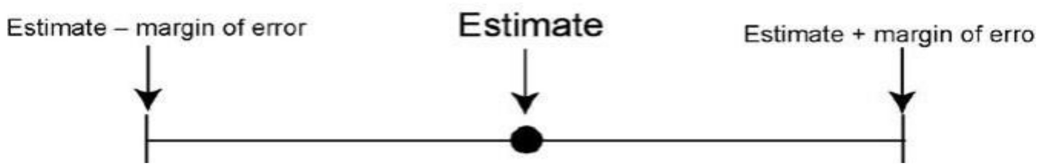


Figure 8.3: margin of an error

This forms a **confidence interval (CI)**. The lower end of the confidence interval is the **lower confidence limit (LCL)**. The upper end is the **upper confidence limit (UCL)**. **Note:** The margin of error is the plus-or-minus wiggle-room drawn around the point estimate; it is equal to half the confidence interval length.

Let 100 (1-α) % represent the **confidence level** of a confidence interval. The α ("alpha") level represents the "lack of confidence" and is the chance the researcher is willing to take in not capturing the value of the parameter.

A 100 (1-α) % CI for μ is given by:

$$\bar{x} \pm ME$$

$$or\ \bar{x} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = \bar{x} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \ \mu \leq \ \bar{x} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

You may use the z/t distribution table given at the of any statistics book to determine z quantiles for various levels of confidence. Here are the common levels of confidence and their associated alpha levels and z quantiles.

Table 8.1: The value of $\boldsymbol{Z_{\alpha/2}}$ *for different level of significance* (α) *and different confidence level*

100 (1−α) %	$\boldsymbol{\alpha}$	$\boldsymbol{Z_{\alpha/2}}$
90%	.10	1.64
95%	.05	1.96
99%	.01	2.58

❖ **Confidence Interval Estimation for population mean (μ)**

Case I: *Sampling from a normally distributed population with known variance $\sigma 2$ (for large or small sample)*

- It is equivalent to say a "Confidence Interval Estimation" when the sample data is *normally distributed.*

Recall that z_α denotes the value of Z for which the area under the standard normal curve to its right equal to α. Analogously $z_{\alpha/2}$ denotes the value of Z for which the area to its right is $\alpha/2$, and $-z_{\alpha/2}$ denotes the value of Z for which the area to its left is $\alpha/2$.

Consider the following figure,

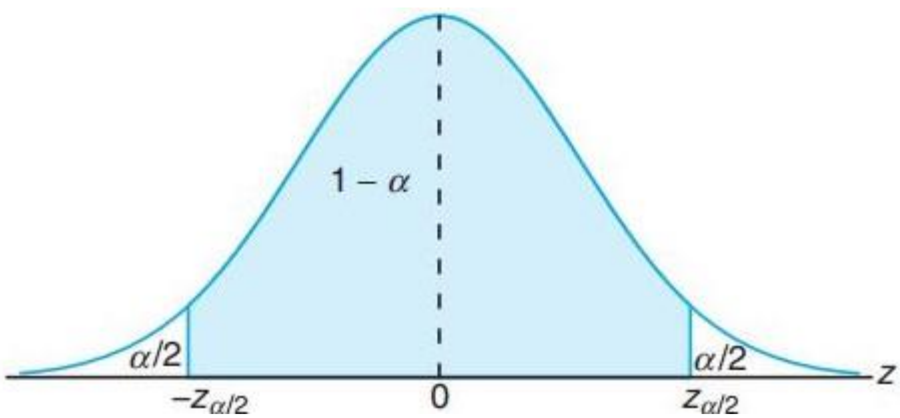


Figure 8.4: P (-zα/2 < Z < zα/2) = 1 - α.

From the above figure we have:

$$P\left[-z_{\alpha/2} < Z < z_{\alpha/2}\right] = 1-\alpha$$

$$\Rightarrow P\left[-z_{\alpha/2} < \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right] = 1-\alpha$$

$$\Rightarrow P\left[-z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}} < \bar{x}-\mu < z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}}\right] = 1-\alpha$$

$$\Rightarrow P\left[\bar{x}-z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}} < \mu < \bar{x}+z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}}\right] = 1-\alpha$$

Thus *a* 100 (1 −α)% confidence interval for the population mean μ is given by:

$$\bar{x} \pm z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}} \text{ or } \bar{x} \pm ME \text{--- } (8.1)$$

Example 7

The average zinc concentration recovered from a sample of measurements taken in 36 different locations in a river is found to be 2.6 grams per milliliter. Find the 95% and 99% confidence intervals for the mean zinc concentration in the river. Assume that the population is normally distributed with standard deviation is 0.3 gram per milliliter.

Soln.

The point estimate of μ is $\bar{x}$ = 2.6. The z-value leaving an area of 0.025 to the right, and therefore an area of 0.975 to the left, is $z_{0.025}$ = 1.96 (from table). Hence, the 95% confidence interval is

$$[\bar{x} - z_{\alpha/2}\,{}^{\sigma}\!/\!_{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2}\,{}^{\sigma}\!/\!_{\sqrt{n}}]$$

$$= [2.6 - 1.96\ {}^{(.3)}\!/\!_{\sqrt{36}} < \mu < 2.6 + 1.96\ {}^{(.3)}\!/\!_{\sqrt{36}}]$$

$$[2.50 < \mu < 2.70]$$

$\Rightarrow$ 95%CI for population mean μ will be [2.50, 2.70]

To find a 99% confidence interval, we find the z-value leaving an area of 0.005 to the right and 0.995 to the left. From Table again, $z_{0.005}$ = 2.58, and the 99% confidence interval is

$$[2.6 - 2.58\ {}^{(.3)}\!/\!_{\sqrt{36}} < \mu < 2.6 + 2.58\ {}^{(.3)}\!/\!_{\sqrt{36}}]$$

Or simply:

$$2.47 < \mu < 2.73.$$

$\Rightarrow$ A 99% CI for μ = [2.47, 2.73]

We now see that a longer interval is required to estimate μ with a higher degree of confidence.

Example: 8

A computer company delivers computers directly to customers who order via the Internet. To reduce inventory cost, the company employs an inventory model. The model requires information about the mean demand during delivery lead time between a central manufacturing facility and local warehouses. Past experience indicates that lead-time demand is normally distributed with a standard deviation of 75 computers per lead time (which is also random). Construct the 95% confidence interval for the mean demand. Demand data for a sample of 25 lead-time periods are given already given. ***Exercise!***

(Ans. the required confidence interval = (340.76, 399.56)).

Case II: *Large sample confidence interval for the population mean μ.*

If the parent distribution of the population from which the sample is drawn is unknown or is not normal but the sample size is large; i.e., the sample size is ≥ 30, then by applying the concept of "**Central Limit Theorem**", a $100(1 - \alpha)\%$ confidence interval for the population mean μ will be given by:

$$\bar{x} \pm z_{\alpha/2}\, {}^{\sigma}\!/\!_{\sqrt{n}} \text{ if } \sigma \text{ is known and } ----------------------(8.2)$$

$$\bar{x} \pm z_{\alpha/2}\, {}^{s}\!/\!_{\sqrt{n}} \text{ if } \sigma \text{ is unknown } ----------------------(8.3)$$

Example 9

A forester wishes to estimate the average number of ''count trees'' per acre (trees larger than a specified size) on a 2,000-acre plantation. She can then use this information to determine the total timber volume for trees in the plantation. A random sample of $n = 50$ one-acre plots is selected and examined. The average (mean) number of count trees per acre is found to be 27.3, with a standard deviation of 12.1. Use the described information and construct 99% confidence interval estimation for μ, (the mean number of count trees per acre for the entire plantation).

Soln.

Given: n = 50, sample mean = 27.3, s = 12.1 confidence level = 99%

We use the general confidence interval with confidence coefficient equal to .99 and a $z_{\alpha/2}$-value equal to 2.58 (see Table 8.1). Substituting into the formula $\bar{x} \pm 2.58\, {}^{\sigma}/_{\sqrt{n}}$ and replacing σ with *s*, we have;

$$\bar{x} \pm 2.58\, {}^{\sigma}/_{\sqrt{n}} = 27.3 \pm 2.58\, {}^{12.1}/_{\sqrt{50}}$$

This corresponds to the confidence interval 27.3 $\pm$ 4.41 that is, the interval from 22.89 to 31.71.

∴ A 100 (1- α) % CI = 99%CI for a population mean $\underline{\mu = [22.89, 31.71]}$

Thus, we are 99% sure that the average number of count trees per acre is between 22.89 and 31.71.

Example 10

A manufacturer claims that his tyres last 20, 000 miles on average. A research organization tests a random sample of 64 tyres and reports an average mileage of 19,200 with a standard deviation of 2,000 miles. Does a 99% confidence interval for the mean life of all tyres produced by this manufacturer support the claim?

Soln.

Given: n = 64, $\bar{x}$ =19,200 *miles and s* = 2,000 *miles and the confidence level = 99%*

Here from confidence level, we can determine the value of α as;

a 100 (1- α) CI = 99%CI

$\Rightarrow 1 - \alpha = .99$

$\Rightarrow \alpha = .01$

$\Rightarrow \frac{\alpha}{2} = 0.005 \Rightarrow \Rightarrow z_{\alpha/2} = 2.58 (From\ table)$

Even if we have no information about the normality assumption of the parent population, but, since n is large (*i e. n* ≥30) we can consider the distribution of the population as a normal distribution by the use of the idea of **central limit theorem**. Hence,

- A 99 percent confidence interval for the mean (μ) will thus be;

$$\bar{x} \pm z_{\alpha/2}\, {}^{s}/_{\sqrt{n}} = 19{,}200 \pm (2.58)\frac{2{,}000}{\sqrt{64}}$$

$$\Rightarrow 19{,}200 \pm 645.0$$

$\Rightarrow$ a 99%CI for population mean μ *will be [18555miles,19845miles]*

Hence, we are 99 percent confident that the true mean mileage is at most 19, 845.0 which is less than the claimed mean 20,000 miles is. Therefore, the claim is not true.

Case III. *Small Sample Confidence interval for the Population Mean: Sampling from a normally distributed population with σ^2unknown and n <30.*

Suppose that the population of interest has a normal distribution with unknown mean $\boldsymbol{\mu}$ and unknown variance $\boldsymbol{\sigma^2}$**.** Assume that a random sample of size n, say $x_1, x_2, \ldots x_n$ is available, and let $\bar{\boldsymbol{x}}$ and $\boldsymbol{s^2}$ be the sample mean and variance, respectively. We wish to construct a two-sided CI on $\boldsymbol{\mu}$. If the variance $\boldsymbol{\sigma^2}$ is known, we know that $z = {(\bar{x}-\mu)}\Big/{\left(\sigma/\sqrt{n}\right)}$ has a standard normal distribution. When $\boldsymbol{\sigma^2}$ is unknown, a logical procedure is to replace $\boldsymbol{\sigma}$ with the sample standard deviation s. The random variable Z now becomes $t = {(\bar{x}-\mu)}\Big/{\left(s/\sqrt{n}\right)}$. A logical question is what effect does replace $\boldsymbol{\sigma}$ by *s* have on the distribution of the random variable *t?* If n is large, the answer to this question is "very little," and we can proceed to use the confidence interval based on the normal distribution from **Case II** above. However, *n* is usually small in most engineering and applied science problems, and in this situation a **t-distribution** must be employed to construct the CI for $\boldsymbol{\mu}$.

❖ **The *t*-Distribution**

Let $x_1, x_2,\ldots, x_n$ be a random sample from a normal distribution with unknown mean $\boldsymbol{\mu}$ and unknown variance $\boldsymbol{\sigma^2}$. The random variable

$$t = \frac{(\bar{x}-\mu)}{\left(s/\sqrt{n}\right)} ---------------------(8.4)$$

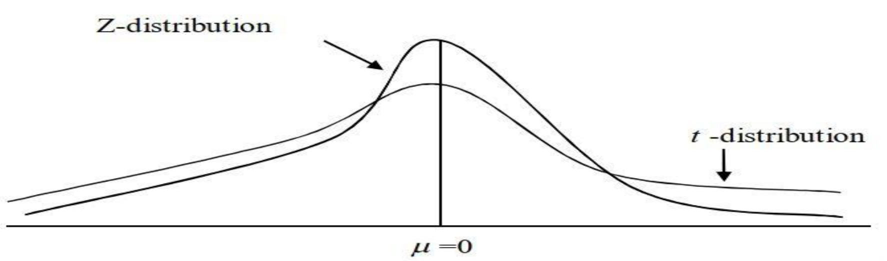
has a *t* distribution with *n* - 1 degrees of freedom.

When:

1. the population from which the sample is selected is (approximately normally distributed),
2. the sample size is small (that is, *n* < 30), and

3. the population standard deviation σ is not known, the normal distribution is replaced by the student *t*- distribution to construct confidence intervals about μ.

The t distribution was developed by W. S. Gosset in 1908 and published under the Pseudonym Student. As a result, the t distribution is also called Student's *t* distribution. The *t* distribution is similar to the normal distribution in some respects. Like the normal distribution curve, the *t* distribution curve is symmetric (bell-shaped) about the mean and it never meets the horizontal axis. The total area under a *t* distribution curve is 1.0 or 100%. However, the *t* distribution curve is flatter than the standard normal distribution curve. In other words, the *t* distribution curve has a lower height and a wider spread (or, we can say, larger standard deviation) than the standard normal distribution. However, as the sample size increases, the *t* distribution approaches the standard normal distribution. The units of a *t* distribution are denoted by *t*. *The* shape of a particular *t* distribution curve depends on the number of degrees of freedom (df).



The *t* Confidence Interval on μ.

It is easy to find a 100 $(1-\alpha)$ % confidence interval on the mean of a normal distribution with unknown variance by proceeding essentially as we did in in the above section (Case I). We know that the distribution of $t = \frac{\bar{x}-\mu_{\bar{x}}}{s/\sqrt{n}}$ is t with $n-1$ degrees of freedom. Letting $\left(t_{\frac{\alpha}{2}}, n-1\right)$be the upper 100 $\frac{\alpha}{2}$ percentage point of the t distribution with $n-1$ degrees of freedom, we may write:

$$P\left(-t_{\frac{\alpha}{2}}, n-1 \leq t \leq t_{\frac{\alpha}{2}}, n-1\right) = 1-\alpha$$

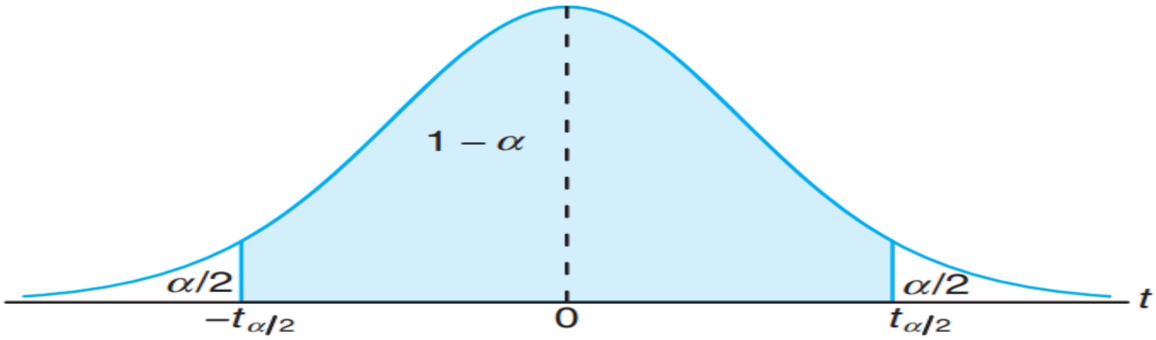


Figure 8.5: P $(-t\alpha_{/2} < t < t\alpha_{/2}) = 1 - \alpha$.

$$P\left(-t_{\frac{\alpha}{2}}, n-1 \leq \frac{\bar{x}-\mu_{\bar{x}}}{s/\sqrt{n}} \leq t_{\frac{\alpha}{2}}, n-1\right) = 1-\alpha$$

Rearranging this last equation yields

$$P\left(\bar{x} - t_{\frac{\alpha}{2}}, n-1\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}}, n-1\frac{s}{\sqrt{n}}\right) = 1-\alpha - - - - - - - - - - - - - -(8.5)$$

This leads to the following definition of the 100 $(1-\alpha)$% two-sided confidence interval on μ.

If $\bar{x}$ and s are the mean and standard deviation of a random sample from a normal distribution with unknown variance $\boldsymbol{\sigma^2}$, a **100 $(1-\alpha)$ % confidence interval on μ** is given by:

$$\bar{x} - t_{\frac{\alpha}{2}}, (n-1)\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}}, (n-1)\frac{s}{\sqrt{n}} - - - - - - - - - - - - - - - - - -(8.6)$$

where $t_{\frac{\alpha}{2}}, n-1$ is the upper $100\frac{\alpha}{2}$ percentage point of the t distribution with $n-1$ degrees of freedom.

<u>**Note:**</u> **One-sided confidence bounds** on the mean of a normal distribution are also of interest and are easy to find. Simply use only the appropriate lower or upper confidence limit from Equation 8-6 and replace $t_{\frac{\alpha}{2}}, n-1$ *by* $t_{\alpha}, n-1$.

Example 11

The contents of seven similar containers of sulfuric acid are 9.8, 10.2, 10.4, 9.8, 10.0, 10.2, and 9.6 liters. Find a 95% confidence interval for the mean contents of all such containers, assuming an approximately normal distribution.

Soln.

The sample mean and standard deviation for the given data are $\bar{x}$ = 10.0 and s = 0.283

$$100(1-\alpha)\%CI = 95\%$$
$$\Rightarrow \alpha = 0.05$$
$$\Rightarrow \frac{\alpha}{2} = 0.025$$

We use student's t-distribution since we are given small sample size and population variance is unknown.

$$\Rightarrow t\alpha_{/2}, n-1 = t_{0.025}, 6 = 2.447\ from\ t - distribution\ table$$

Hence,

95% confidence interval for μ is

$$10.0 - (2.447)\left(\frac{0.283}{\sqrt{7}}\right) < \mu < 10.0 + (2.447)\left(\frac{0.283}{\sqrt{7}}\right).$$

which reduces to 9.74 < μ < 10.26.

Example 12

A random sample of 10 children found that their average growth for the first year was 9.8 inches. Assume the variable is normally distributed and the sample standard deviation is 0.96 inch. Find the 95% confidence interval of the population mean for growth during the first year.

Given: $\bar{x} = 9.8,\quad s = 0.96,\quad n = 10$

Soln.

Since σ is unknown and s must replace it, the t distribution must be used for the confidence interval. Hence, with 9 degrees of freedom $t\alpha_{/2}, 9 = 2.262\ from\ table$. The 95% confidence interval can be found by substituting in the formula.

$$\bar{x} - t_{\frac{\alpha}{2}}, n-1\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\frac{\alpha}{2}}, n-1\frac{s}{\sqrt{n}}$$
$$\Rightarrow 9.8 - 2.262\left(\frac{0.09}{\sqrt{10}}\right) < \mu < 9.8 + 2.262\left(\frac{0.09}{\sqrt{10}}\right)$$
$$\Rightarrow 9.8 - 0.69 < \mu < 9.8 + 0.69$$
$$\Rightarrow 9.11 < \mu < 10.49$$

Therefore, one can be 95% confident that the population mean of the first-year growth is between 9.11 and 10.49 inches.

Example 13

The data represent a random sample of the number of home fires started by candles for the past several years. (Data are from the National Fire Protection Association.) Find the 99% confidence interval for the mean number of home fires started by candles each year. (**Exercise!**)

5460 5900 6090 6310 7160 8440 9930

Ans. $4785.2 < \mu < 9297.6$ ***or*** *the confidence interval will be* $[4785.2,\ 9297.6]$

❖ **Choosing the Sample Size for Estimating μ (To build Interval Estimation for μ)**

The 100(1-α) % confidence interval provides an estimate of the accuracy of our point estimate. If μ is actually the center value of the interval, then $\bar{x}$ estimates μ without error. Most of the time, however, $\bar{x}$ will not be exactly equal to μ and the point estimate will be in error. The size of this error will be the absolute value of the difference between μ and $\bar{x}$, and we can be 100(1 - α) % confident that this difference will not exceed $z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}}$. We can readily see this if we draw a diagram of a hypothetical confidence interval, as in Figure 8.4 below.

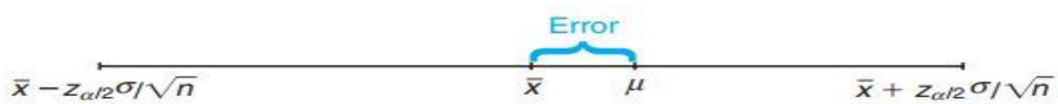


Figure 8.6: error in estimating μ by $\bar{x}$

Theorem 8.1: If $\bar{x}$ is used as an estimate of μ, we can be 100(1 − α) % confident that the error will not exceed $z_{\alpha/2}\,{}^{\sigma}/_{\sqrt{n}}$.

In Example 9 above, we are 95% confident that the sample mean $\bar{x}$ = 2.6 differs from the true mean μ by an amount less than $(1.96)\frac{(.3)}{\sqrt{36}}$ = 0.1 and 99% confident that the difference is less than $(2.58)\frac{(.3)}{\sqrt{36}}$ = 0.13.

Frequently, we wish to know how large a sample is necessary to ensure that the error in estimating μ will be less than a specified amount E. By Theorem 8.1, we must choose n such that $z_{\alpha/2} * {}^{\sigma}/_{\sqrt{n}}$= E. Solving this equation gives the following formula for n.

Theorem 8.2: If $\bar{x}$ is used as an estimate of μ, we can be 100(1 − α) % confident that the error will not exceed a specified amount E when the sample size is

$$n = [\frac{z_{\frac{\alpha}{2}}\sigma}{E}]^2 = \frac{(z_{\frac{\alpha}{2}})^2\sigma^2}{E^2}$$

When solving for the sample size, *n*, we round all fractional values up to the next whole number. By adhering to this principle, we can be sure that our degree of confidence never falls below 100(1 - α)% Strictly speaking, the formula in Theorem 8.2 is applicable only if we know the variance of the population from which we select our sample. Lacking this information, we could take a preliminary sample of size $n \geq 30$ to provide an estimate of σ. Then, using s as an approximation for σ in Theorem 8.2, we could determine approximately how many observations are needed to provide the desired degree of accuracy.

Example 14

How large a sample is required if we want to be 95% confident that our estimate of μ in *Example 9* is off by less than 0.05?

Soln.

The population standard deviation is $\sigma = 0.3$. Then, by Theorem 8.2,

$$n = [\frac{z_{\frac{\alpha}{2}}\sigma}{E}]^2 = \frac{(1.96)^2(.3)^2}{.05^2} = 138.3$$

Therefore, we can be 95% confident that a random sample of size 139 will provide an estimate $\bar{x}$ differing from μ by an amount less than 0.05.

Exercise!

A federal agency has decided to investigate the advertised weight printed on cartons of a certain brand of cereal. The company in question periodically samples cartons of cereal coming off the production line to check their weight. A summary of 1,500 of the weights made available to the agency indicates a mean weight of 11.80 ounces per carton and a standard deviation of .75 ounce. Use this information to determine the number of cereal cartons the federal agency must examine to estimate the average weight of cartons being produced now, using a 99% confidence interval of width .50.

8.2 Hypothesis Testing

8.2.1 Statistical Hypothesis

In the previous session we illustrated how to construct a confidence interval estimate of a parameter from sample data. However, many problems in applied science require that we decide whether to accept or reject a statement about some parameter. The statement is called a **hypothesis**, and the decision-making procedure about the hypothesis is called **hypothesis testing.** This is one of the most useful aspects of statistical inference, since many types of decision-making problems, tests, or experiments in the world can be formulated as hypothesis-testing problems. Furthermore, as we will see, there is a very close connection between hypothesis testing and confidence intervals. Statistical hypothesis testing and confidence interval estimation of parameters are the fundamental methods used at the data analysis stage of a comparative experiment**,** in which the researcher is interested, for example, in comparing the mean of a population to a specified value. These simple comparative experiments are frequently encountered in practice and provide a good foundation for the more complex experimental design problems. In this part we discuss comparative experiments involving a single population, and our focus is on testing hypotheses concerning the parameters of the population. We now give a formal definition of a statistical hypothesis.

Definition: A *statistical hypothesis* is a statement about the parameters of one or more populations.

Hypotheses concerning parameters such as means and proportions can be investigated. There are two specific statistical tests used for hypotheses concerning means: the z test and the t test. This chapter will explain in detail the hypothesis-testing procedure along with the z test and the t test.

- **There are three different methods which used to test hypotheses. These are:**
  1. The Critical Value Method (**The Traditional Approach**)
  2. The P-value Method (**The Modern Approach**)
  3. The Confidence Interval Method

The traditional method will be explained first. It has been used since the hypothesis testing method was formulated. A newer method, called the P-value method, has become popular with the advent of modern computers and high-powered statistical calculators. In this course, we use the first method (critical value method).

- **Hypothesis and types of hypotheses**

A claim (or statement) about a population parameter is called a hypothesis.

Example 15

a) The mean daily profit of a supermarket is 1000 Birr or μ =1000 Birr.

b) The mean time to complete a certain assembly job is less than 2 hours or $\mu < 2$ hours.

There are two types of hypotheses: - these are:

1. ***Null hypothesis:*** A null hypothesis is a claim (or statement) about a population parameter that is assumed to be true until it is declared false. It is denoted by H_o.
   - ✓ It is the hypothesis to be tested.
   - ✓ The hypothesis of equality.
   - ✓ $H_0: \mu = \mu_0$
2. ***Alternative hypothesis:*** An alternative hypothesis is a claim about a population parameter that will be true if the null hypothesis is false. The alternative hypothesis is denoted by H_A or H_a or H_1.
   - ✓ It is also known as research hypothesis.
   - ✓ The hypothesis of inequality.
   - ✓ $H_a: \mu < \mu_0$,
     $H_a: \mu > \mu_0$ *or*
     $H_a: \mu \neq \mu_0$

Note: A claim, however, can be stated as either the null hypothesis or the alternative hypothesis; but the statistical evidence can only support the claim if it is the alternative hypothesis.

Statistical evidence can be used to reject the claim if the claim is the null hypothesis. These facts are important when you are stating the conclusion of a statistical study.

Example 16

A soft drink bottling company's advertisement states that a bottle of its products contains 330 milliliters (ml.). But customers are complaining that the company is under filling its products. To check whether the complaint is true or not, an inspector may test the following hypotheses:

H_o: The average content of a bottle of this product is equals to 330 ml, against, H_a: The average content of a bottle of this product is less than 330ml Or symbolically,

$$H_0: \mu = 330ml$$

$$H_a: \mu < 330ml$$

If the inspector takes a random sample of bottles of this product and finds that the mean content per bottle is much less than 330 ml, then he may conclude that the complaint of the customers is correct.

Example 17

State the null and alternative hypotheses for each conjecture.

A. A researcher thinks that if expectant mothers use vitamin pills, the birth weight of the babies will increase. The average birth weight of the population is 8.6 pounds.

B. An engineer hypothesizes that the mean number of defects can be decreased in a manufacturing process of USB drives by using robots instead of humans for certain tasks. The mean number of defective drives per 1000 is 18.

C. A psychologist feels that playing soft music during a test will change the results of the test. The psychologist is not sure whether the grades will be higher or lower. In the past, the mean of the scores was 73.

Soln.

A. $H_0: \mu = 8.6$ Vs $H_a: \mu > 8.6$
B. $H_0: \mu = 18$ Vs $H_a: \mu < 18$
C. $H_0: \mu = 73$ Vs $H_a: \mu \neq 73$

Hypothesis testing is a procedure for checking the validity of statistical hypothesis. It is the process by which we decide whether the null hypothesis should be rejected or not. The value computed from a sample that is used to determine whether the null hypothesis has to be rejected or not is called a ***test statistic***. Sometimes some known population quantities are used in the calculation of a test statistic alongside sample values.

❖ **Types of errors in statistical hypothesis tests**

Applying a hypothesis test may lead to a wrong conclusion. There are two kinds of possible errors, called type I error and type II error.

a) **Type I error:** 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,

$$\alpha = P(H_0 \text{ is rejected } | H_0 \text{ is true }).$$

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

b) **Type II error:** Type II error occurs when a false null hypothesis is not rejected. The value of β represents the probability of committing a type II error; that is

$$\beta = P(H_0 \text{ is not rejected } | H_0 \text{ is false})$$

The value of $1-\beta$ is called the power of the test. It represents the probability of not making a type II error. 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 , However, we can decrease both α and β simultaneously by increasing the sample size. The following table presents the possible conclusions and errors in performing a test.

Two-way decision process

Decision	Null Hypothesis: True	Null Hypothesis: False
Reject H_0	**Type I error** $\boldsymbol{\alpha}$	correct decision
Accept H_0	correct decision	**Type II error** $\boldsymbol{\beta}$

❖ Type of Tests

Based on the form of the null and alternative hypotheses, we have two types of tests:

- **One-sided**(one tailed) tests and

Where the alternative hypotheses are inequalities type are called one-sided (one-tailed) tests. μ_0 is the hypothesized (assumed) mean.

$$\textbf{Form: } a)\ \begin{cases} H_0: \mu = \mu_0 \\ H_0: \mu < \mu_0 \end{cases} or \begin{cases} H_0: \mu \geq \mu_0 \\ H_0: \mu < \mu_0 \end{cases} \Rightarrow \textbf{\textit{Left tailed test}}$$

$$b)\ \begin{cases} H_0: \mu = \mu_0 \\ H_0: \mu > \mu_0 \end{cases} or \begin{cases} H_0: \mu \leq \mu_0 \\ H_0: \mu > \mu_0 \end{cases} \Rightarrow \textbf{\textit{Right tailed test}}$$

- **Two-sided** (two tailed) test.

In which the critical region (rejection region) includes both large and small values of the test statistic.

$$\textbf{Form: } \begin{cases} H_0: \mu = \mu_0 \\ H_0: \mu \neq \mu_0 \end{cases} \Rightarrow \textbf{\textit{Two tailed test}}$$

❖ General Procedure for Hypothesis Tests

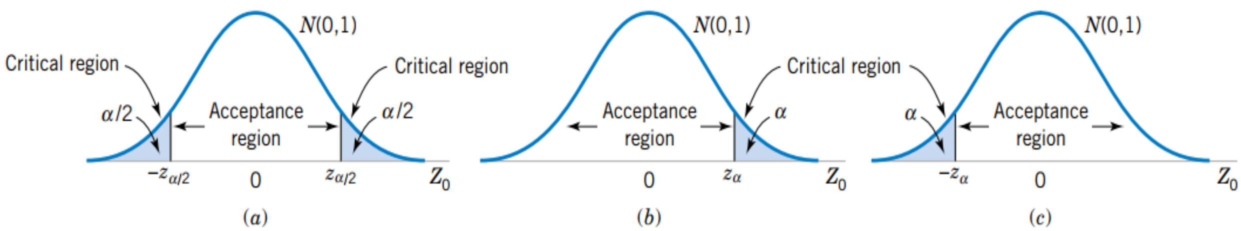
This chapter develops hypothesis-testing procedures for many practical problems. Use of the following sequence of steps in applying hypothesis-testing methodology is recommended.

1. From the problem context, identify the parameter of interest.
2. State the null hypothesis, H_0.
3. Specify an appropriate alternative hypothesis H_a.
4. Choose a significance level α.

5. Determine an appropriate test statistic.
6. State the rejection region for the statistic.
7. Compute any necessary sample quantities, substitute these into the equation for the test statistic, and compute that value.
8. Decide whether or not H_0 should be rejected and report that in the problem context.

The ***level of significance*** is the maximum probability of committing a type I error. This probability is symbolized by α (Greek letter alpha). That is, P (type I error) $=\alpha$.

☞ **Acceptance and Rejection Region**



$$\boldsymbol{Figure}\ 8.7\text{: } The\ distribution\ of\ Z_0(Z_{\text{call}})\ when\ H_0{:}\ \mu=\mu_0\ is\ true,\ with\ critical\ region\ for\ (a) the\ two-sided\ alternative\ H_a{:}\ \mu\neq\mu_0, (b) the\ one-sided\ alternative\ H_a{:}\ \mu>\mu_0, (c) the\ one-sided\ alternative\ H_a{:}\ \mu<\mu_0,$$

8.2.2 Tests on the Mean of a Normal Distribution:

<u>Case I:</u> Hypothesis Tests on the Mean: When Population Variance is known

Suppose we have a random sample of size n (small or large) from a normal population with mean μ and standard deviation, σ where σ is known.

☞ **For the two tailed hypothesis testing**

$$H_0{:}\ \mu=\mu_0$$
$$H_a{:}\ \mu\neq\mu_0$$

Consider figure 8.6 (a): $P\left(-Z\alpha_{/2}<\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}<Z\alpha_{/2}\right)=1-\alpha=acceptance\ region$

$$P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}<-Z\alpha_{/2}\right)=\frac{\alpha}{2} or\ P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}>Z\alpha_{/2}\right)=\frac{\alpha}{2}\ critical\ region$$

$\Rightarrow P\left(\left|\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|>Z\alpha_{/2}\right)=\alpha=critical\ region/rejection\ region$

$$\Rightarrow reject\ H_0\ if\ \left|\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right|>Z\alpha_{/2}$$

Example 18

A producer of an electric bulbs claims that the average life length of its product is 1800 hrs. A sample of 400 bulbs gave mean life length of 1780 hrs. Suppose it's known that life lengths are normally distributed with standard deviation of 200 hrs. Would you support the producers claim at 5% level of significance? Test also that the average life length of the bulbs is less than the producer's claims at 1% level of significance.

Soln.

Step-1: formulate the hypothesis

$$H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0$$

Step-2: identify the significance level: $\alpha = 0.05$

Step-3: the parent Population is normal, with $\sigma = 200$ $(known)$, $n = 400$ Then the test statistic will be:

$$Z_{Call} = \left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right|$$

Step 4: the critical region is

$$\left|\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right| > Z_{\alpha/2}$$

i.e. reject H_0 if $\left|\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}\right| > Z_{\alpha/2}$

Step 5: Evaluate the critical region and compare with tabulated value

$$Z_{call} = \left|\frac{1780-1800}{200/\sqrt{400}}\right| = |-2| = 2 \text{ and}$$

$$Z_{tab} = Z_{\alpha/2} = Z_{0.05/2} = Z_{0.025} = 1.96\ (From\ table\ 8.1)$$

Step-6: decide whether or not H_0 should be rejected

Since $Z_{call} > Z_{tab}$, *reject Ho* that is the claim of the producer is not correct.

Exercise! Test also that the average life length of the bulbs is less than the producer's claims at 1% level of significance.

☞ **To test the hypothesis**

$$H_0: \mu = \mu_0$$
$$H_a: \mu > \mu_0$$

where μ_0 is some specified constant.

Test Statistics: - Since the best estimator of μ is $\bar{x}$ the test statistic must be dependent on $\bar{x}$.

We know that $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = Z_{Call} = \boldsymbol{Test\ Statistics}$$

- By considering *fig 8.6 (b)* above, $P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > Z_\alpha\right) = \alpha$

  $\Rightarrow$ the critical region or the rejection rejoin will be when $Z_0 = Z_{call} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} > Z_\alpha \ or\ (Z_{tab})$
- i.e. if $Z_{call} > Z_\alpha\ for\ a\ specified\ level\ of\ significance, we\ can\ reject\ H_0$

Example 19

A random sample of 100 recorded deaths in the United States during the past year showed an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is greater than 70 years? Use a 0.05 level of significance.

<u>Soln.</u>

1. H_0: μ = 70 years.
2. H_a: μ > 70 years.
3. α = 0.05.
4. Critical region: $Z_{call} > Z_{tab}$, where $Z_{tab} = 1.645\ from\ table$

   Computations: $\bar{x} = 71.8\ years, \sigma = 8.9 years\ and\ hence,\ \ Z_{call} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{71.8-70}{8.9/\sqrt{100}} = 2.02$
5. Decision Making and conclusion:

   Since $Z_{call} > Z_{tab}$, reject H_0 and conclude that the mean life span today is greater than 70 years.

- **To test the hypothesis**

Left tailed Hypothesis:

$$H_0: \mu = \mu_0$$
$$H_a: \mu < \mu_0$$

where μ_0 is some specified constant.

Test Statistics: - Since the best estimator of μ is $\bar{x}$ the test statistic must be dependent on $\bar{x}$.

We know that $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = Z_{Call} = \boldsymbol{Test\ Statistic}$$

- By considering *fig 8.6 (b)* above, $P\left(\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < -Z_\alpha\right) = \alpha$

  $\Rightarrow$ the critical region or the rejection rejoin will be when $Z_0 = Z_{call} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} < -Z_\alpha\ or\ (Z_{tab})$

- i.e. if $Z_{call} < -Z_\alpha\ for\ a\ specified\ level\ of\ significance, we\ can\ reject\ H_0$

Example 20

According to the advertisement of car manufacturing company, their cars averaged at least 32 miles per gallon (mpg) in the city. From past records it is known that mileage is normally distributed with a standard deviation of 2.5 mpg. Tests on 16 cars showed that mean mileage in the city is 31.5 mpg. Do the data support the advertisement at the 99 percent confidence level?

Given: $\mu_0 = 32mpg, \sigma = 2.5mpg, \bar{x} = 31.5mpg, \alpha = 0.01\ and\ n = 16.$

<u>Soln.</u>

1. H_0: $\mu \geq$ 32mpg.
2. H_a: $\mu <$ 32mpg.
3. $\alpha = 0.01$.
4. Critical region: $Z_{call} < -Z_{tab}$, where $-Z_{tab} = -2.33\ from\ table$

$$\text{Computations} \qquad Z_{call} = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{31.5-32}{2.5/\sqrt{16}} = -0.8$$

5. Decision Making and conclusion:

   Since $Z_{call} > -Z_{tab}$, *we accept H₀*. That is the mean mileage is at least 32 mpg.

Case II: Hypothesis Tests on the Mean: When Population Variance is Unknown: The *t*-Test

We now consider the case of **hypothesis testing** on the mean of a normal population with **unknown variance** σ^2 *and* ***small sample*** *case* ($n < 30$). The situation is analogous to Section 8-1, where we considered a **confidence interval** on the mean for the same situation. As in that section, the validity of the test procedure we will describe rests on the assumption that the population distribution is at **least approximately normal.** The important result upon which the test procedure relies is that if $X_1, X_{2, \ldots} X_n$ is a random sample from a normal distribution with mean μ and variance σ^2, the random variable

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

has a t distribution with $n - 1$ degrees of freedom. Recall that we used this result in Section 8-1 to devise the *t*-confidence interval for μ. Now consider testing the hypotheses

$$H_0: \mu = \mu_0$$
$$H_a: \mu \neq \mu_0$$

We will use the **test statistic**

$$\boldsymbol{t_0} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

If the null hypothesis is true, t_0 has a t distribution with $n - 1$ degrees of freedom. When we know the distribution of the test statistic when H_0 is true (this is often called the **reference distribution** or the **null distribution**), we can locate the critical region to control the type I error probability at the desired level. In this case we would use the t percentage points $-t\alpha_{/2}, (n-1)$ and $t\alpha_{/2}, (n-1)$ as the boundaries of the critical region so that we would reject $H_0: \mu = \mu_0$ if

$$t_0 > t\alpha_{/2}, (n-1) \; or \; if \;\; t_0 < -t\alpha_{/2}, (n-1)$$

Where t_0 is the observed value of the test statistic. The test procedure is very similar to the test on the mean with known variance described in "***Case I***" above, except that t_0 is used as the test statistic instead of Z_0 and the t_{n-1} distribution is used to define the critical region instead of the standard normal distribution. A summary of the test procedures for both two- and one sided alternative hypotheses follows:

❖ **The one sample t-test Rejection Rule**

Null hypothesis $H_0 \ : \mu = \mu_0$

Test statistic $t_0 = \frac{\bar{x}-\mu}{s/\sqrt{n}}$

Alternative Hypothesis	**Rejection Criteria** Reject H_0 if:
$H_a \ : \mu \neq \mu_0$	$t_0 > t_{\alpha/2}, (n-1) \ \ or \ \ t_0 < -t_{\alpha/2}, (n-1)$
$H_a \ : \mu > \mu_0$	$t_0 > t_\alpha, (n-1)$
$H_a \ : \mu < \mu_0$	$t_0 < -t_\alpha, (n-1)$

Example 21

The increased availability of light materials with high strength has revolutionized the design and manufacture of golf clubs, particularly drivers. Clubs with hollow heads and very thin faces can result in much longer tee shots, especially for players of modest skills. This is due partly to the "spring-like effect" that the thin face imparts to the ball. Firing a golf ball at the head of the club and measuring the ratio of the outgoing velocity of the ball to the incoming velocity can quantify this spring-like effect. The ratio of velocities is called the coefficient of restitution of the club. An experiment was performed in which 15 drivers produced by a particular club maker were selected at random and their coefficients of restitution measured. In the experiment the golf balls were fired from air cannon so that the incoming velocity and spin rate of the ball could be precisely controlled. It is of interest to determine if there is evidence (with 0.05) to support a claim that the mean coefficient of restitution exceeds 0.82. The observations are given as follow:

0.8411	0.8191	0.8182	0.8125	0.8750
0.8580	0.8532	0.8483	0.8276	0.7983
0.8042	0.8730	0.8282	0.8359	0.8660

Soln.

Since the objective of the experimenter is to demonstrate that the mean coefficient of restitution exceeds 0.82, a **one-sided** alternative hypothesis is appropriate.

- ***Step 1:*** The parameter of interest is the mean coefficient of restitution,μ.
- ***Step 2:*** $H_0 \; : \mu = 0.82$
- ***Step 3*:** $H_a \; : \mu > 0.82$. We want to reject H_0 if the mean coefficient of restitution exceeds 0.82.
- ***Step 4:*** $\alpha = 0.05$
- ***Step 5:*** The test statistic is

$$t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

- ***Step 6:*** Reject H_0 if $t_o or\ t_{call} > t_\alpha, (n-1)$

  Where $t_{tab} = t_\alpha, (n-1) = t_{0.05}, 14 = 1.761$

- ***Step 7:*** Computations: Since $\bar{x} = 0.83725, s = 0.02456, \mu_0 = 0.82\ and\ n = 15, we\ have$

$$t_0 = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{0.83725 - 0.82}{0.02456/\sqrt{15}} = 2.72$$

Step 8: Conclusions: Since $t_0 = 2.72 > t_\alpha, (14) = 1.761$, we reject H_0 and conclude at the 0.05 level of significance that the mean coefficient of restitution exceeds 0.82.

Exercise!

The sodium content of thirty 300-gram boxes of organic corn flakes was determined. The data (in milligrams) are as follows:

131.15	130.69	130.91	129.54	129.64	128.77	130.72	128.33	128.24	129.65
130.14	129.29	128.71	129.00	129.39	130.42	129.53	130.12	129.78	130.92
131.15	130.69	130.91	129.54	129.64	128.77	130.72	128.33	128.24	129.65

Can you support a claim that the mean sodium content of this brand of cornflakes is 130milligrams? Use $\alpha = 0.05$

8.2.3 When we have a large sample size, non-normal population and σ^2 *may known/ unknown*.

We have developed the test procedure for the null hypothesis $H_0 : \mu = \mu_0$ by assuming that the population is normally distributed and that σ^2 is known in most case. In many if not most practical situations σ^2 will be unknown. Furthermore, we may not be certain that the population is well modeled by a normal distribution. In these situations if n is large (say $n \geq 30$) the sample standard deviation s can be substituted for σ in the test procedures with little effect. Thus, while we have given a test for the mean of a normal distribution with known σ^2, it can be easily converted into a **large sample test procedure for unknown** σ^2 that is valid regardless of the form of the distribution of the population. This large-sample test relies on the central limit theorem just as the large sample confidence interval on that was presented in the previous sessions did.

☞ **We wish to test**

$$H_0 : \mu = \mu_0 \;\; V_s \;\; H_a : \mu \neq \mu_0 \quad \text{or} \quad H_0 : \mu = \mu_0 \;\; V_s \;\; H_a : \mu > \mu_0 \quad \text{or} \quad H_0 : \mu = \mu_0 \;\; V_s \;\; H_a : \mu < \mu_0$$

By central limit theorem, if n is large $(n \geq 30)$ sample mean, $\bar{x}$ is approximately normal.

i.e. $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

The critical regions in all these tests in this case are the same as in "***Case I***" above. If σ is unknown, estimate it by sample standard deviation, s

Where, $s = \sqrt{\frac{\sum(x_i-\bar{x})^2}{n-1}}$ *and* $z = \frac{\bar{x}-\mu_{\bar{x}}}{s/\sqrt{n}} \sim N(0,1)$

Example 22

A company has a computer system that can process at most 1200 bills per hour. A new system is tested which processes an average of 1260 bills per hour with a standard deviation of 215 bills in a sample of 40 hours. Test if the new system is significantly better than old one at the 5% level of significance.

<u>Soln.</u>

1. H_0: μ ≤1200
   H_a: μ > 1200
2. α = 0.05.
3. The population is non- normal, σ is unknown

   But, n = 40 ⇒ large sample size, $s = 215, and\ \bar{x} = 1260$

   Then by central limit theorem, the test statistic will be:

   $$Z = \frac{\bar{x}-\mu}{s/\sqrt{n}} \Rightarrow Z_{call} = \frac{1260-1200}{215/\sqrt{40}} = 1.76$$

4. Critical region: $Z_{call} > Z_{tab}$, where $Z_{tab} = Z_{\alpha} = Z_{0.05} = 1.645\ from\ table$
5. Decision Making and conclusion:

Since $Z_{call} > Z_{tab}$, reject H_0 that is we conclude that the new system represents an improvement over the old system at the $\alpha = 0.05$ level of significance.

Exercise!

Medical researchers have developed a new artificial heart constructed primarily of titanium and plastic. The heart will last and operate almost indefinitely once it is implanted in the patient's body, but the battery pack needs to be recharged about every four hours. A random sample of 50 battery packs is selected and subjected to a life test. The average life of these batteries is 4.05 hours with standard deviation 0.2 hour. Is there an evidence to support the claims that mean battery life is beneath 4 hours? Use $\boldsymbol{\alpha = 0.1}$

❖ P-Value Method for Hypothesis Testing on μ.

Statisticians usually test hypotheses at the common α levels of 0.05 or 0.01 and sometimes at 0.10. Recall that the choice of the level depends on the seriousness of the type I error. Besides listing α value, many computer statistical packages give a *P*-value for hypothesis tests. The P-value (or probability value) is the probability of getting a sample statistic (such as the mean) or a more extreme sample statistic in the direction of the alternative hypothesis when the null hypothesis is true.

In other words, the *P*-value is the actual area under the standard normal distribution curve (or other curve, depending on what statistical test is being used) representing the probability of a particular sample statistic or a more extreme sample statistic occurring if the null hypothesis is true.

For example, suppose that an alternative hypothesis is H_a: $\mu > 50$ and the mean of a sample is $\bar{X}$ =52. If the computer printed a P-value of 0.0356 for a statistical test, then the probability of getting a sample mean of 52 or greater is 0.0356 if the true population mean is 50 (for the given sample size and standard deviation). The relationship between the P-value and the α value can be explained in this manner. For P = 0.0356, the null hypothesis would be rejected at α = 0.05 but not at α = 0.01.

When the hypothesis test is two-tailed, the area in one tail must be doubled. For a two-tailed test, if α is 0.05 and the area in one tail is 0.0356, the P-value will be 2(0.0356) = 0.0712. That is, the null hypothesis should not be rejected at α = 0.05, since 0.0712 is greater than 0.05. In summary, then, if the P-value is less than α , reject the null hypothesis. If the P-value is greater than α , do not reject the null hypothesis.

☞ Decision Rule When Using a P-Value

- *If P- value $\leq \alpha$ reject the null hypothesis*
- *If P- value $> \alpha$ do not reject the null hypothesis*

☞ Procedure Table

Solving Hypothesis-Testing Problems (P-Value Method)

Step 1 State the hypotheses and identify the claim.

Step 2 Compute the test value.

Step 3 Find the P-value.

Step 4 Make the decision.

Step 5 Summarize the results

Example 23

It is claimed that automobiles are driven on average more than 20,000 kilometers per year. To test this claim, 100 randomly selected automobile owners are asked to keep a record of the kilometers they travel. Would you agree with this claim if the random sample showed an average of 23,500 kilometers and a standard deviation of 3900 kilometers? Use a P-value in your conclusion.

Soln.

- ***Step 1.*** State the hypotheses and identify the claim.

$$H_0: \mu \leq 20{,}000 \; kms \; and$$

$$H_a: \mu > 20{,}000 \; kms(claim).$$

- ***Step 2.*** Compute the test statistic

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{s/\sqrt{n}} = \frac{23{,}500 - 20{,}000}{3900/\sqrt{100}} = 8.97$$

- ***Step 3:*** Find the P-value.

The P-value for the right-tailed test is $P(x > 20{,}000) = P\,(z > 8.97) = 1 - P\,(z \leq 8.97) = 1 - 0.9999 = 0.0001$.

- ***Step 4.*** Make the decision.

The decision is to reject the null hypothesis, since the P-value is less than 0.05 or 0.01.

- ***Step 5.*** Summarize the results.

Therefore, the claim is true, that is $\mu > 20{,}000$ *kilometers*.

Example 24

The mean lifetime of a sample of 100 fluorescent light bulbs produced by a company is computed to be1570 hours with a standard deviation of 120 hours. If μ is the mean lifetime of all the bulbs produced by the company, test the hypothesis μ =1600 hours against the alternative hypothesis $\mu \neq$ 1600 hours, using a level of significance of $\alpha = 0.05$.

Soln.

Step 1. State the hypotheses and identify the claim.

H_0: μ = 1600 hours *and*

H_a: $\mu \neq$1600 hours (*claim*).

- ***Step 2.*** Compute the test statistic

$$Z = \frac{\bar{x} - \mu_{\bar{x}}}{s/\sqrt{n}} = \frac{1570 - 1600}{120/\sqrt{100}} = -2.50$$

- ***Step 3:*** Find the P-value.

  The P-value for the two-tailed test is $P(x \leq -2.50) + P(z > 2.50) = 2P(Z \leq -2.50) = 2P(Z \geq 2.50) = 0.0124$

- ***Step 4.*** Make the decision.

  The decision is to reject the null hypothesis, since the P-value is less than 0.05.

- ***Step 5.*** Summarize the results.

  There is no enough evidence to accept the null hypothesis.

Exercise!

Label on a large can of Hilltop Coffee states average weight of coffee contained in all cans it produces is 3 pounds of coffee. A coffee drinker association claims average weight is less than 3 pounds of coffee, $\mu < 3$. Suppose a random sample of 30 cans has an average weight of $\bar{x} = 2.95$ pounds and standard deviation of $s = 0.18$. Does data support coffee drinker association's claim at $\alpha = 0.05$?