# 1. Introduction

Welcome to the series of e-learning modules on Estimate. In this module we are going cover the basic concepts of estimate, role of estimates, properties, types of estimates, biased estimates and standard error.

By the end of this session, you will be able to explain:

- Estimates
- Types of estimates
- Properties of estimates
- Biased estimates
- Role of Estimates

- One of the objectives of sampling theory is to extract maximum information about the population by drawing a sample from the population and to get the best estimates of the parameters of the population.
- The Estimate of parameters using sampling theory needs to be such that the estimated value of the parameter lies almost closer to the actual value of the parameter.

Everyone makes estimates.
When you are ready to cross a street, you estimate the speed of any car that is approaching, the distance between you and that car, and your own speed.
Having made these quick estimates you decide whether to wait, walk or run.
As educated citizens and professionals one will be able to make more useful estimates by applying the proper techniques.

We shall now look at some applications of estimate in our daily life.

- Managers must make quick estimates that may affect their organizations as seriously as the outcome of your decision as to whether to cross a street or not.
- Managers make estimates because in all but the most trivial decisions they must make rational decisions without complete information and with a great deal of uncertainty about what the figure will bring.
- A credit manager will estimate whether a purchaser will eventually pay his bills
  The credit manager attempts to estimate the credit worthiness of prospective customers from a sample of their past payment habits.
- Prospective home buyers make estimates concerning the behavior of interest rates in the mortgage market
  The home buyer attempts to estimate the future course of interest rates by observing the current behavior of these ratesmin the Mortgage market.
  In each case somebody is trying to infer something about a population from information taken from a sample.
- The concept of probability theory forms the foundation for statistical inference.

This branch of statistics is concerned with using probability concepts to deal with uncertainty in decision making.
Statistical Inference is based on estimation and hypothesis testing.
We shall be making inferences about the characteristics of the populations from information contained in these samples.

# 2. Meaning of Estimate

Estimation is the process by which sample data are used to indicate the value of an unknown quantity in a population.
As Estimate is:

- A specific observed value of a statistic
- An indication of the value of an unknown quantity based on observed data
- A particular value of an estimator that is obtained from a particular sample of data and used to indicate the value of a parameter

For instance, suppose the manager of a shop wanted to know the mean expenditure of customers in his shop in the last year. He could use an estimate of the population mean by calculating the mean of a representative sample of customers. If this value was found to be $25, then $25 would be his estimate.

Now, lets suppose that we calculate the mean odometer reading or mileage from a sample of used taxis and find it to be 98,000 miles. If we use this specific value to estimate the mileage for a whole fleet of used taxis, the value 98,000 would be an estimate.

Methods available enable us to estimate with reasonable accuracy the population proportion and the population mean.
To calculate the exact proportion or the exact mean would be an impossible goal.
Even so, we will be able to make an estimate, make a statement about the error and will probably accompany this estimate and implement some controls to avoid as much of the errors as possible.
As decision makers we will be forced to at times to rely on blind hunches.
An estimate of the population parameters can be expressed in two ways:
1)Point Estimate
2) Interval Estimate

Point estimate.
A point estimate of a population parameter is a single value of a statistic.
For example, the value of the sample mean is a point estimate of the population mean μ.
Similarly, the value of the sample proportion *p* is a point estimate of the population proportion *P*.

If while watching the first members of a football team come into a field you say "Why, I bet their line must average 250 pounds" You have made a point estimate.
Similarly, a department head would make a point estimate if she said " Our current data indicate that this course will have 350 students in the fall"

In this example if you are only told that her point estimate of enrolment is wrong, you do not know how wrong it is, and you cannot be certain of the estimates reliability.
If you learn that it is off by only 10 students you would accept 350 students as a good estimate of future enrolment.
Therefore point estimate is much more useful if it is accompanied by an estimate of the error that might be involved.

An interval estimate is defined by two numbers, between which a population parameter is said to lie.
An interval estimate is a range of values used to estimate a population parameter.
It indicates an error in two ways.
1) By the extent of its range
2) By the probability of the true population parameter lying within the range
For example, a less than x less than b is an interval estimate of the population mean 'mew'.
It indicates that the population mean is greater than a but less than b.
In this case a department head would say something like:
"I estimate that the true enrolment in this course in the fall will be between 330 and 380. It is very likely that the exact enrolment will fall within this interval."
Here she has better idea of reliability of her estimate.

# 3. Point Estimates and Interval Estimates

Point Estimates vs. Interval Estimates
A point estimate doesn’t tell us how close the estimate is likely to be to the parameter, while an interval estimate is more useful. It incorporates a margin of error which helps us to gauge the accuracy of the point estimate.
A point estimate is a single number that is our best guess for the parameter while an interval estimate is an interval of number within which the parameter value is believed to fall.

The level of confidence we want to establish is given by the number alpha, which is the probability that a point estimate will not fall in a confidence range.

The lower the alpha, the more confident we want to be - e.g. alpha of five percent indicates we want to be ninety five percent confident; one percent alpha indicates ninety nine percent confidence.

Here is a look at this table which illustrates several populations, population parameters and its estimates .

Properties of an Estimat**e**
The four desirable properties of an estimate are that they are unbiased, efficient , consistent and sufficient:

1. Unbiased - The expected value (mean) of the estimate's sampling distribution is equal to the underlying population parameter; that is, there is no upward or downward bias.
2. Efficiency - While there are many unbiased estimates of the same parameter, the most efficient has a sampling distribution with the smallest variance. Good estimate has smaller standard error than other estimates. Falls closer than other estimates to parameter. Sample mean has a smaller standard error than sample median

3. Consistency - Larger sample sizes tend to produce more accurate estimates; that is, the sample parameter converges on the population parameter.

4. Sufficiency: If an estimate consists of sufficient information about the population parameter being estimated then estimate is said to be sufficient.

The precision of any estimate made from a sample depends both on the method by which the estimate is calculated from the sample data and on the plan of sampling. To save space we sometimes write as: “the precision of the sample mean”, without specifically mentioning the other fundamental factors.

A Method of estimation is called consistent if the estimate becomes exactly equal to the population value. When n=N , as in, the sample size equals to the population size, that is when the sample consists of the whole population, Consistency is a desirable property of estimates.
On the other hand inconsistent estimates are not necessarily useless since it may give

satisfactory precision when n is small compared to N.

# 4. Unbiased Estimate and Biased Estimate

Unbiased Estimate:

The estimates are denoted by the symbol cap. A Good estimate has sampling distribution centred at the parameter. An estimate with this property is unbiased.
Let teeta be the parameter of the population which is a function of the population values. Then Teeta cap is said to be an unbiased estimate of the parameter Teeta ; if Expected value of teeta cap equals to teeta, it implies that, Expected value of teeta cap minus teeta equals to zero. Unbiased estimates determine the tendency, on the average, for the estimates to assume values close to the parameters of interest.

Biased Estimate
Let teeta be the parameter which is a function of the population values, Let teeta cap be an estimate of the population parameter teeta. Then teeta cap is said to be a biased estimate of the parameter teeta if Expected value of teeta cap is not equal to teeta OR
That is, Expected value of teeta cap minus teeta is not equal to zero, which is denoted by B teeta cap.

Then the estimated value is not close to the actual value of the parameter.Biased estimate lies away from the actual value of the parameter it is trying to estimate. The error which arises when estimating a quantity is known as Bias.

Here is a look at an example of biased and unbiased estimates.
The police decide to estimate the average speed of drivers using the fast lane of the motorway and consider how it can be done.

The first method considered is to tail the cars using police patrol cars and record their speeds as being the same as that of the police car. The result of this is likely to produce a biased result as any driver exceeding the speed limit will slow down on seeing a police car behind them.

The other method to do estimate the speed would be where in the police decides to use an unmarked car for their investigation using a speed gun operated by a constable.
This would result in an unbiased method of measuring speed, but is imprecise compared to using a calibrated speedometer to take the measurement.

When an estimate p cap is systematically skewed away from the true parameter *p*, it is considered to be a biased estimate of the parameter.
Let us suppose that an analyst wishes to determine the percentage of defective items which are produced by a factory over the course of a week.

Since the factory produces thousands of items per week, the analyst takes a sample three hundred items and observes that fifteen of these are defective. Based on these results, the analyst computes the *statistic* as , p cap equals to fifteen divided by three hundred which is equal to zero point zero five, as an estimate of the *parameter p* , or true proportion of defective items in the entire population.

In the factory example discussed, if the true percentage of defective items was known to be eight percent, then our sampling distribution would be biased in the direction of estimating too few defective items. An *unbiased estimate* will have a sampling distribution whose mean is equal to the true value of the parameter.

From the sample, a value is calculated which serves as a point estimate for the population parameter of interest.
The best estimate of the population percentage, capital P, is the sample percentage, lowercase p.

B, The best estimate of the unknown population mean, ‘mew’ , is the sample mean, where x bar equals to summation x divided by n.

This estimate of ‘mew’ is often written and referred to as 'mew hat'.

C, the best estimate of the unknown population standard deviation, sigma , is the sample standard deviation s, where:
S is equal to the square root of summation of x minus x bar whole square divided by n minus one. This is an unbiased estimate of population standard deviation sigma.
Here we have to note that, S is equal to square root of summation of x minus x bar whole square divided by n is not used as an unbiased estimate of standard deviation as it underestimates the value of standard deviation.

# 5. Examples

Example : Here is another example.

An Accountant at a very small supermarket wishes to obtain some information about all the invoices sent out to its account customers. In order to obtain an estimate of this information, a sample of twenty invoices were randomly selected from the whole population of invoices. Use the results below obtained from them to obtain point estimates for invoices whose value of invoices are given in dollars as 32, 22,33,41,38,25,26,30,38,38,31,38,43,29,22,25,24,43, 32 and 42.

What we will calculate here will be, one, the percentage of invoices in the population, P, which exceed forty dollars. Two, the population mean 'mew' and three, the population standard deviation, sigma.

In the solution, first we shall calculate the percentage of invoices in the population, P, which exceed forty dollars.

Where lowercase p is equal to the number of observations which exceeds forty dollars in a sample of twenty multiplied by 100.

Which is equal to 4 into 100 divided by 20 is equal to twenty percent. Therefore the estimate of P is, p cap is equal to twenty percent.

Next to find the estimate of the population mean, which is 'mew hat' is equal to summation of xi divided by 20, which is equal to 652 by 20 which is 32 point 6 dollars.

Next, we will now look at calculating the estimate of population standard deviation, s, where we know that, s is equal to the square root of summation of x minus x bar whole square divided by n minus one,Which is equal to square root of 972 point 8 divided by 19, which is equal to 7 point one six. Hence an estimate of population standard deviation, sigma hat is equal to 7 point one six dollars.

In short here are a few points to remember.Both the precision and credibility of the estimates improve with the increasing quality and quantity of a sample. A procedure of "guessing" properties of the population from which data are collected is known as a statistical estimation. An estimate of an unknown parameter is a value that represents a "guess" of the properties of the population.And finally, the aim of statistical estimation is to provide a best estimate for the population under study, satisfying almost all the properties of a good estimate without any bias.

Here's a summary of our learning in this session:

- Statistical definition of an estimate
- Types of estimates
- Difference between point and interval estimates

- Properties of estimates
- Biased estimates
- Illustrative example to obtain the estimates