Estimation and Hypothesis Testing

Md. Pear Hossain
Lecturer of Statistics
BSMRSTU, Bangladesh

Estimation

Definition: Estimation is a procedure by which a numerical value or values are assigned to a population **parameter** based on the information collected from a **sample**.

Example.

1. An auto mobile company may want to estimate the mean fuel consumption for a particular model of a car.
   This is an example of estimating true population mean μ.
2. The University may want to find the proportion of all students who are in engineering faculty.
   This is an illustration of estimating the true population proportion, p.

Estimate and estimator

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

The sample statistic used to estimate a population parameter is called an **estimator**.

Example.

Suppose the manager of auto mobile company takes a sample of 40 cars of that particular car and find that the mean fuel consumption, $\bar{x}$, is 60 liters. If he assigns this value to the population mean, then 60 liters is called the **estimate** of μ.

The sample mean $\bar{x}$ is an estimator of the population mean.

Steps in estimation procedure

The estimation procedure involves the following steps

1. Select a sample.
2. Collect the required information from the members of the sample.
3. Calculate the value of the sample statistic.
4. Assign value (s) to the corresponding population parameter.

Two types of estimate (cont…..

1. A Point Estimate
2. An Interval Estimate

Definition
Point Estimate The value of a sample statistic that is used to estimate a population parameter is called a *point estimate*.

For the example, suppose the Census Bureau takes a sample of 10,000 households and determines that the mean housing expenditure per month, $\bar{x}$, for this sample is $1970. Then, using $\bar{x}$ as a point estimate of μ , the Bureau can state that the mean housing expenditure per month, μ, for all households is about $1970. Thus,

Point estimate of a population parameter = Value of the corresponding sample statistic

Two types of estimate

1. A Point Estimate
2. **An Interval Estimate**

Definition
Interval Estimation In *interval estimation*, an interval is constructed around the point estimate, and it is stated that this interval is likely to contain the corresponding population parameter.

For the example about the mean housing expenditure, instead of saying that the mean housing expenditure per month for all households is $1970, we may obtain an interval by subtracting a number from $1970 and adding the same number to $1970. Then we state that this interval contains the population mean, μ.

The question arises

What number should we subtract from and add to a point estimate to obtain an interval estimate?

Point estimate $\pm$ Margin of error

Definition

Confidence Level and Confidence Interval Each interval is constructed with regard to a given *confidence level* and is called a *confidence interval*. The confidence interval is given as

Point estimate $\pm$ Margin of error

The confidence level associated with a confidence interval states how much confidence we have that this interval contains the true population parameter. The confidence level is denoted by $(1-\alpha)100\%$.

Estimation of a Population Mean: σ^2 Known

Confidence Interval for μ

The $(1-\alpha)100\%$ *confidence interval* for μ is

$$\bar{x} \pm \frac{z\sigma}{\sqrt{n}}$$

Where The value of z used here is obtained from the standard normal distribution table for the given confidence level.

Margin of Error

The margin of error for the estimate for μ , denoted by E, is the quantity that is subtracted from and added to the value of $\bar{x}$ to obtain a confidence interval for μ. Thus,

$$E = \frac{z\sigma}{\sqrt{n}}$$

Example

A publishing company has just published a new college textbook. Before the company decides the price at which to sell this textbook, it wants to know the average price of all such textbooks in the market. The research department at the company took a sample of 25 comparable textbooks and collected information on their prices. This information produced a mean price of $145 for this sample. It is known that the standard deviation of the prices of all such textbooks is $35 and the population of such prices is normal.

(a) What is the point estimate of the mean price of all such college textbooks?

(b) Construct a 90% confidence interval for the mean price of all such college textbooks.

Solution: We are given that $n = 25, \bar{x} = \$145$ and $\sigma = \$35$

(a) The point estimate of the mean price of all such college textbooks is $145; that is,

$$Point\ estiamte\ of\ \mu = \bar{x} = \$\mathbf{145}$$

(b) For 90% confidence level the z value is $z = 1.65$. Hence, 90% confidence interval for the mean price of all such college textbooks

$$\bar{x} \pm \frac{z\sigma}{\sqrt{n}} = (\$133.45, \quad \$156.55)$$

Thus, we are 90% confident that the mean price of all such college textbooks is between $133.45 and $156.55.

In the above estimate, $11.55 is called the margin of error.

When to use what

1. When σ^2 is unknown and n is small $(n < 30)$, then use t statistic with $(n-1)$ d.f.
2. When σ^2 is unknown and n is large $(n \geq 30)$, then use z statistic with $N(0,1)$.
3. When σ^2 is known (n is small or large), then use z statistic with $N(0,1)$.