SNS COLLEGE OF TECHNOLOGY

An Autonomous Institution
Coimbatore-35

DEPARTMENT OF ARTIFICIAL INTELLIGENCE & DATA SCIENCE

23ADT202 – FUNDAMENTALS OF DATA SCIENCE AND ANALYTICS

II YEAR IV SEM

UNIT II – STANDARD ERROR OF ESTIMATE

STANDARD ERROR OF ESTIMATE

EMPATHY:

- Standard Error is the measure of the variability of a sample statistic used to estimate the variability of a population.

- Standard Error is important in dealing with sample statistics, such as sample mean, sample proportion, etc.

- Sample Error Formula is used to determine the accuracy of a sample that reflects a population.

- The standard error formula is the discrepancy between the sample mean and the population mean.

What is Standard Error?

The term "sample" in statistics refers to a specific set of information that is generated.

The data we obtained on the height of some people in a locality, for example, maybe the sample

A population is a collection of people from which we draw a sample.

There are several ways to define a population, and we must always be clear about what constitutes a population.

This collection necessitates a large number of calculations.

FODS/Standard Error of Estimate/N. Padmashri/SNSCT

Term	Represents
Population	Entire dataset
Sample	Part of the population
Population Mean (μ)	True average of population
Sample Mean ($\bar{x}$)	Estimated average from sample
Population SD (σ)	True variability of population
Sample SD (s)	Estimated variability from sample

2. Sample Mean vs Population Mean

Aspect	Population Mean (μ)	Sample Mean (x̄)
Symbol	μ (mu)	x̄ (x-bar)
Definition	Average of **all population values**	Average of **sample values**
Formula	$\mu = \Sigma X / N$	$\bar{x} = \Sigma x / n$
Usage	True average (usually unknown)	Estimate of the population mean
Example	Average height of all citizens in a country	Average height of selected 500 citizens

3. Sample Standard Deviation vs Population Standard Deviation

Aspect	Population SD (σ)	Sample SD (s)
Symbol	σ (sigma)	s
Definition	Measures spread of **entire population data**	Measures spread of **sample data**
Formula	$\sigma = \sqrt{\left(\Sigma(X - \mu)^2 / N \right)}$	$s = \sqrt{\left(\Sigma(x - \bar{x})^2 / (n - 1) \right)}$
Denominator	N	n − 1
Reason	Exact dispersion	Uses $n - 1$ to correct bias (Bessel's correction)
Usage	When full population data is available	When data is from a sample

Population Standard Deviation

$$\sigma = \sqrt{\frac{\Sigma(X_i - \mu)^2}{N}}$$

Sample Standard Deviation

$$S = \sqrt{\frac{\Sigma(X_i - \bar{x})^2}{N-1}}$$

Standard Error(SE) Formula

The SE formula is used to determine the reliability of a sampling that represents a population. The sample mean that differs from the provided population and is expressed as:

$$SE = S/\sqrt{(n)}$$

where,

S is Standard Deviation of Data
n is Number of Observations

Standard Error of Estimate (SEE)

- Standard Error Estimate is use to find the accuracy of prediction of any event. Its abbreviation is SEE.
- Standard Error Estimate (SEE) is also called the Sum of Sqaures Error. SEE is the square root of average squared deviation.

Standard Error of Estimate(SEE) formula is discussed below,

$$SEE = \sqrt{[\Sigma(x_i - \mu)/(n - 2)]}$$

where,

xi is values of Data
μ is Mean Value of Data
n is Sample Size

How to Calculate Standard Error?

Various steps to calculate the standard error are added below,

Step 1: Find the number of measurement(n) and find the sample mean(μ).

Step 2: Find the variation of sample values from the mean value.

Step 3: Find the square of the deviation and find their sum. $\Sigma(xi - \mu)^2$

Step 4: Divide the sum from step 3 by (n-1).

Step 5: Take the square root of the number, i.e. standard deviation (σ).

Step 6: Divide the standard deviation by the square root of measurement(n).

Standard Error Solved Examples

Example 1: Find the standard error for the sample data: 1, 2, 3, 4, 5.

Solution:
Mean of Given Data = (1+2+3+4+5)/5

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

μ = 15/5

μ = 3

Standard Deviation = $\sqrt{((1-3)^2 + (2-3)^2 + (3-3)^2 + (4-3)^2 + (5-3)^2)/(5-1)}$

$\sigma = \sqrt{((4+1+0+1+4)/4)}$

$\sigma = \sqrt{(10/4)}$

σ = 1.5

SE = 1.5/$\sqrt{5}$

SE = 0.67

Example 2: Find the standard error for the sample data: 2, 3, 4, 5, 6.

Example 3: Find the standard error for the sample data: 10, 20, 30, 40, 45.

Example 4: Find the standard error for the sample data: 2, 6, 9, 5.

Example 5: Find the standard error for the sample data: 5, 8, 10, 12.