Estimation

Estimation in statistics refers to the process of inferring or approximating the value of a population parameter based on sample data. Since it is often impractical or impossible to gather data from an entire population, estimation provides a means to make inferences using a smaller, representative sample. There are two main types of estimation:

1. Point Estimation

Point estimation involves providing a single value as the estimate of an unknown population parameter. For example, the sample mean (denoted as $\bar{x}$) is often used as a point estimate of the population mean (μ).

The key properties of good point estimators include:

- Unbiasedness : The expected value of the estimator equals the true population parameter.
- Consistency : As the sample size increases, the estimator becomes more accurate.
- Efficiency : Among all unbiased estimators, the one with the smallest variance is preferred.
- Sufficiency : The estimator uses all available information in the data relevant to the parameter being estimated.

2. Interval Estimation

Interval estimation involves calculating a range of values, called a confidence interval, within which the population parameter is likely to lie. A confidence interval is expressed as:

Estimate ±Margin of Error

CI (Confidence Interval population mean**):** $\bar{x} \pm z \frac{\sigma}{\sqrt{n}}$

CI (Confidence Interval population mean) : $\hat{p} - z\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} < p < \hat{p} + z\sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$

Where, $\hat{p} = \frac{x}{n}$ and x = number of success and n = the size of the sample

Applications of Estimation:

Estimation is widely used in various fields, such as economics, biology, and social sciences, to make informed decisions based on data. Examples include:

- Estimating the average income of a population based on a sample survey.
- Determining the proportion of voters who support a particular candidate using sample data.
- Estimating the effect of a drug in clinical trials by comparing sample outcomes.

In practice, statistical software is often used to compute both point estimates and confidence intervals, enabling more precise and reliable decision-making.

Q1

The time taken in minutes for a certain daily train journey has a normal distribution with standard deviation 5.8 . For a random sample of 20 days the journey times were noted and the mean journey time was found to be 81.5 minutes.

(a) Calculate a 98% confidence interval for the population mean journey time. [3]

Q2

The widths, w cm, of a random sample of 150 leaves of a certain kind were measured. The sample mean of w was found to be 3.12 cm.

Using this sample, an approximate 95% confidence interval for the population mean of the widths in centimetres was found to be [3.01, 3.23].

(a) Calculate an estimate of the population standard deviation. [3]

Q3

A student wishes to estimate the proportion, p, of students at her college who have exactly one brother. She surveys a random sample of 50 students at her college and finds that 18 of them have exactly one brother. She calculates an approximate α% confidence interval for p and finds that the lower limit of the confidence interval is 0.244 correct to 3 significant figures.

Find α correct to the nearest integer. [4]