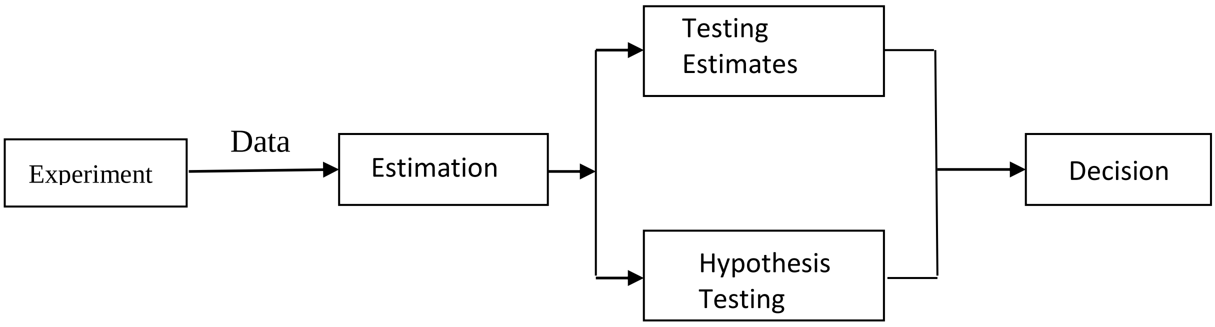
## 8. Statistical Intervals for Single Sample:

Process of making decisions:




**Estimation:** Using an Estimator to obtain estimates of population parameters based on sample data.

**Inferences:** Expressing an estimate with a certain probability statement; or making a hypothesis about a parameter and testing the hypothesis for acceptance or rejection. The result of the test is in terms of a certain probability of being correct.

**Estimators:**

- Point Estimates: - Single valued estimates are considered.
  -Express probability of its having a certain value.
- Interval Estimators: Range of values between which the parameter will lie with a given probability is considered.

**Confidence Interval:** Interval Estimator

It is an interval estimate for a population parameter. The length of the interval indicates precision of the estimate. We have high confidence that the interval contains the unknown population parameter.

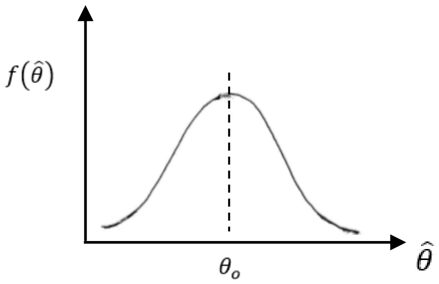
**Confidence coefficient:** (1-$\alpha$) [ = 90%, **95%** or 99%]

Proportion of cases (%) that the parameter under consideration lies within the interval estimator. [usually 95%].

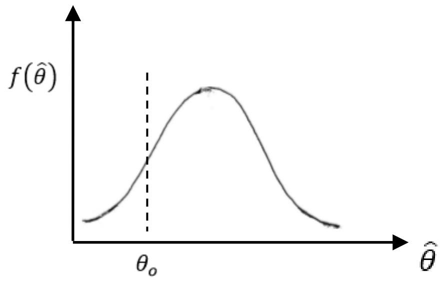
**Properties of Estimators:**

Quality of an estimator is judged by the distribution of the estimates (sampling distribution)

- Estimator should be unbiased:



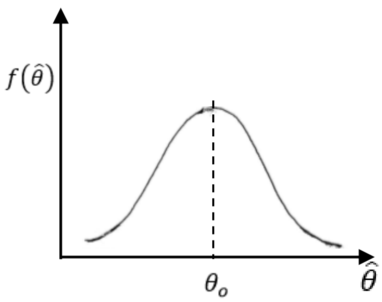

Unbiased



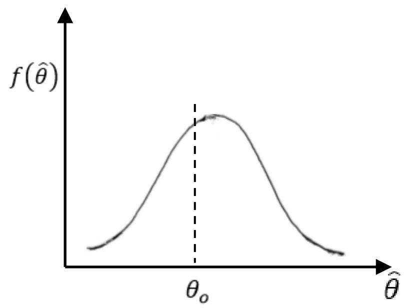

Biased

- Estimator should be consistent:

  Spread of distribution of estimate should be as small as possible. It is inversely proportional to the sample size.




More accurate




Less Accurate

$\hat{\theta}$: *Sampling Distribution of Estimator*; $\theta_o$ *Population Parameter*

**Confidence Interval on the Mean of a Normal Distribution: known $\sigma^2$ and unknown μ.**

$\bar{X}$ : is Normal then $Z = \dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ is also normal.

We define interval for μ as:

$$l \le \mu \le u$$

*l* and *u* are end points computed from sample data. Therefore, they are also random variables: L and U. *l* and *u* are the lower and upper confidence limits, such that

$$P\{L \le \mu \le U\} = 1-\alpha \qquad\qquad 0 \le \alpha \le 1$$

This means that there is a probability of (1- $\alpha$) of selecting a sample for which the interval will contain the true value of $\mu$. **(1- $\alpha$)** is known as the confidence coefficient.

The Z values that correspond to the probability of (1- $\alpha$) are defined as

$$P\left\{-Z_{\alpha/2} \le Z \le +Z_{\alpha/2}\right\} = 1-\alpha$$

$$P\left\{-Z_{\alpha/2} \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le +Z_{\alpha/2}\right\} = 1-\alpha$$

$$P\left\{\bar{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}\right\} = 1-\alpha$$

**Definition of Confidence Interval. known $\sigma^2$ and unknown μ:**

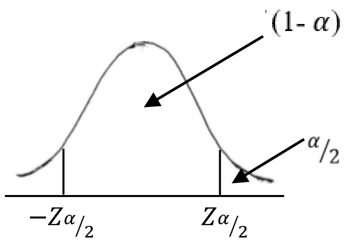
A 100 $(1-\alpha)$ % confidence interval for $\mu$ is given by:

$$\underbrace{\bar{X} - Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}}_{\text{Lower limit } l} \le \mu \le \underbrace{\bar{X} + Z_{\alpha/2}.\frac{\sigma}{\sqrt{n}}}_{\text{upper limit } u}$$

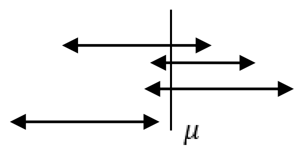
n: sample size $\alpha/2$

$\bar{x}$ : sample mean

$Z_{\alpha/2}$ = upper 100 $\alpha/2$ percentage point of standard normal distribution




This means that if an infinite number of random samples are collected and a 100 $(1-\alpha)$ % confidence interval for $\mu$ for each sample, then 100 $(1-\alpha)$ % of these intervals will contain the true value of $\mu$. In other words, the observed interval [*l*,*u*] brackets the true value of $\mu$ with confidence of $100(1-\alpha)$. Which means that the method used to obtain the interval [*l*,*u*] yields correct statements 100 $(1-\alpha)$ % of the time.



Example:

Experiment on measurements of impact energy (J) on specimens of A238 Steel cut at 6ºC.

64.1 64.7 64.5 64.6 64.5 64.3 64.6 64.8 64.2 64.3

Assume impact energy is normally distributed with $\sigma$ =1 J. Find 95% CI for $\mu$, mean impact energy.

$$\bar{x} - z\alpha/_2 . \frac{\sigma}{\sqrt{n}} \leq \ \mu \ \leq \ \bar{x} + z\alpha/_2 . \frac{\sigma}{\sqrt{n}}$$

$$\bar{x} = \frac{\sum \bar{x}_i}{n} \ = \ 64.46 \qquad n = 10$$

$$100\ \alpha = 5\% \ \rightarrow \ \alpha/_2 = 0.025$$

$$z\alpha/_2 = z_{0.025} \ = 1.96$$

95% CI:

$$64.46 - 1.96 . \frac{1}{\sqrt{10}} \leq \ \mu \ \leq \ 64.46 + 1.96 . \frac{1}{\sqrt{10}}$$

$$63.84 \leq \ \mu \ \leq \ 65.08$$

**Confidence Level and Precision of Estimation:**

The length of Confidence Interval is:

$$2 \left( Z\alpha/_2 . \frac{\sigma}{\sqrt{n}} \right)$$

It is a measure of precision, and it is $\propto \frac{\sigma}{\sqrt{n}}$ .

**Choice of sample size:**

The error $E = \ |\bar{x} - \ \mu| \ should\ be \ \leq \left( Z\alpha/_2 . \frac{\sigma}{\sqrt{n}} \right)$ with confidence 100 (1-α).

$$\therefore n = \frac{\left( Z\alpha/_2 \ \sigma \right)^2}{E}$$ [n to be rounded up]

The error $|\bar{x} - \ \mu|$ will not exceed E. In this case 2E is the resulting confidence interval.