① **Unbiasedness:**

A statistic T is said to be an unbiased estimator of a parametric function $\gamma(\theta)$ if $E_\theta(T) = \gamma(\theta)$, whatever be the true value of θ.

Eg: ① Let, 'p' be the probability of getting head in a single toss of a coin. To estimate 'p', let the coin be tossed 'n' times & let X be the no. of heads obtained in these 'n' tosses. Then, $X \sim B_{in}(n, p)$ as such, $E(X) = np \Rightarrow E(\frac{X}{n}) = p$. $= p$

Hence, the sample proportion of heads $\frac{X}{n}$ is an unbiased estimator of 'p'.

Eg: ② Let, $X_1, X_2, \ldots, X_n$ be a random sample from a $N(\mu, \sigma^2)$ distribution.

Then, $E(X_i) = \mu, \forall i$; $V(X_i) = \sigma^2, \forall i$.

Also, $X_1, X_2, \ldots, X_n$ are independent random variables.

Now, $E(\bar{X}) = E(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{n\mu}{n} = \mu$.

Thus we have, $(n+1)$ unbiased estimators of μ. Thus, the problem is to choose the best among them. Hence, we need a 2nd criteria.

(14) **Remark :**

① If there exists two unbiased estimators of a certain parameter, one may construct an infinite no. of unbiased estimators of the parameter θ.

Suppose, T_1 & T_2 be both unbiased for θ, ie $E(T_1) = E(T_2) = \theta$.

Then, for any real 'a', then, an estimator of 'θ' can be given by

$T^* = aT_1 + (1-a)T_2$. Clearly, $E(T^*) = \theta$.

② Absard unbiased estimator :

There may exist an unbiased estimator of a certain positive parameter, such that, we may get occationally a negative ~~unbiase~~ unbiased estimator.

eg: Suppose, we have a single observation drawn from a Poisson dist^n with parameter 'λ'. Then, let us consider a parameter $(-2)^x$.

Now, $E[(-2)^X] = \sum_{x=0}^{\infty} (-2)^x \frac{e^{-\lambda} \cdot \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(-2\lambda)^x}{x!} = e^{-\lambda} \cdot e^{-2\lambda} = e^{-3\lambda}$.

$\therefore (-2)^X$ is an unbiased estimator of e^{-3X}. Note that, here, $e^{-3X} > 0$ but $(-2)^X < 0$.

Result: Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from a distribution with mean μ & variance σ^2. Then, $s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is a biased estimator of σ^2. Hence suggest an unbiased estimator of σ^2.

Proof: ~~Let~~ Since, $X_1, X_2, \ldots, X_n$ be a random sample drawn from a distribution with mean μ & variance σ^2, we have,

i) $E(X_i) = \mu, \forall i$; ii) $V(X_i) = \sigma^2, \forall i$; iii) $X_1, X_2, \ldots, X_n$ are independent r.v.'s.

Now, $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\mu = \frac{n\mu}{n} = \mu$.

$\therefore \bar{X}$ is an unbiased estimator of μ.

$V(\bar{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i)$ [$\because X_i$'s are mutually independent]

$= \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$.

$$= \frac{\sigma^2}{n}$$

Again, $E(s^2) = E[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2] = E[\frac{1}{n}\sum x_i^2 - \bar{x}^2] = \frac{1}{n}\sum_{i=1}^{n} E(x_i^2) - E(\bar{x}^2)$

$$= \frac{1}{n}\sum_{i=1}^{n}\{V(x_i) + (E(x_i))^2\} - \{V(\bar{x}) + (E(\bar{x}))^2\}$$

$$= \frac{1}{n}\sum_{i=1}^{n}(\sigma^2+\mu^2) - (\frac{\sigma^2}{n}+\mu^2) = \sigma^2+\mu^2-\frac{\sigma^2}{n}-\mu^2 = (1-\frac{1}{n})\sigma^2 \neq \sigma^2$$

Hence, s^2 is a biased estimator of σ^2.

But, $E(s^2) = (1-\frac{1}{n})\sigma^2$ ie, $E(s^2) = \frac{n-1}{n}\sigma^2$ ie $\frac{n}{n-1}E(s^2) = \sigma^2$

ie $E(\frac{n}{n-1}s^2) = \sigma^2$ ie $E(\frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2] = \sigma^2$ or, $E[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2] = \sigma^2$

As such, $\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$ is an unbiased estimator of σ^2.

Remark: when the sample size is very large, ie when $n\to\infty$, then $\frac{1}{n}\simeq 0$. Hence, in this case, $E(s^2)\simeq\sigma^2$ & thus s^2 may itself be considered as an unbiased ~~es~~ estimator of σ^2.

Bias

Definition *For observations* $X = (X_1, X_2, \ldots, X_n)$ *based on a distribution having parameter value* θ, *and for* $d(X)$ *an estimator for* $h(\theta)$, *the* **bias** *is the mean of the difference* $d(X) - h(\theta)$, *i.e.,*

$$b_d(\theta) = E_\theta d(X) - h(\theta).$$

If $b_d(\theta) = 0$ *for all values of the parameter, then* $d(X)$ *is called an* **unbiased estimator.** *Any estimator that is not unbiased is called* **biased.**

(5) Best Linear Unbiased Estimator (BLUE):

A statistic 'T' is said to be the BLUE of a parametric function $\gamma(\theta)$ if

i) $E_\theta(T) = \gamma(\theta)$, whatever be the true value of θ.

ii) 'T' is of the form $\sum_{i=1}^{n} a_i x_i$, where a_i's are constants & x_i's are sample observations.

iii) $V_\theta(T) \leq V_\theta(T')$, where T' is any other linear unbiased estimator of $\gamma(\theta)$.

Result: Let $X_1, X_2, \cdots, X_n$ be a random sample drawn from a distribution with mean μ & variance σ^2. Then the sample mean $\bar{X}$ will be the BLUE for the population mean μ.

Proof: Let, $T = \sum_{i=1}^{n} a_i X_i$ be the BLUE for μ, where a_i's are constants, $\forall i = 1(1)n$.

Then, to show that $T = \bar{X}$, ie $\sum_{i=1}^{n} a_i X_i = \frac{1}{n} \sum_{i=1}^{n} X_i$ ie $a_i = \frac{1}{n}$, $\forall i = 1(1)n$.

Since, 'T' is unbiased for μ, we have, $E(T) = \mu$.

ie $E\left[\sum_{i=1}^{n} a_i X_i\right] = \mu$

ie. $\sum_{i=1}^{n} a_i E(X_i) = \mu$.

ie $\sum_{i=1}^{n} a_i \mu = \mu$. $\quad \therefore \sum_{i=1}^{n} a_i = 1 \cdots (1)$.

Again, $V(T) = V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i)$ [X_i's being mutually independent]

$= \sigma^2 \sum_{i=1}^{n} a_i^2 \cdots (2)$.

To determine a_i's optimally, we minimize $V(T)$ w.r.t. a_i's subject to the condition $\sum_{i=1}^{n} a_i = 1$. In other words, we have to minimize $\sigma^2 \sum_{i=1}^{n} a_i^2$ subject to the condition $\sum_{i=1}^{n} a_i = 1$.

condition $\sum_{i=1} a_i = 1$.

This is equivalent to minimize $L = \sigma^2 \sum_{i=1}^{n} a_i^2 + \lambda(\sum_{i=1}^{n} a_i - 1)$ unconditionally w.r.t. a_i's & λ, where, λ is an unknown constant called Lagrange's multiplier.

Now, for any i,

$$\frac{\partial L}{\partial a_i} = 0 \Rightarrow \sigma^2 \cdot 2a_i + \lambda = 0 \quad \cdots (3)$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow \sum_{i=1}^{n} a_i = 1 \quad \cdots (4)$$

Taking sum over 'i' in both sides of (3) we get,

$$2\sigma^2 \sum_{i=1}^{n} a_i + n\lambda = 0$$

$$\Rightarrow 2\sigma^2 + n\lambda = 0 \quad [\text{using } (4)].$$

$$\Rightarrow \lambda = -\frac{2\sigma^2}{n} \quad \cdots (5)$$

Substituting the value of λ from (5) into (3) we get,

$$\sigma^2 \cdot 2a_i - \frac{2\sigma^2}{n} = 0 \Rightarrow a_i = \frac{1}{n}, \ \forall i=1(1)n.$$

Hence, the proof. [Proved].