Example

Let $X_1, \dots, X_v$ r.s. of distribution with mean value μ and variance σ^2. If $T = c \cdot \sum_{i=1}^{v} (X_i - \bar{X})^2$, find c so that T is an unbiased estimator (u.e) of σ^2

(From previous example: if $X_1, \dots, X_v$ r.s. of $N(\mu, \sigma^2)$ then $\frac{(v-1)s^2}{\sigma^2} \sim \chi^2_{v-1}$, so $E\left[\frac{(v-1)s^2}{\sigma^2}\right] = v-1 \Rightarrow$
$\Rightarrow E[s^2] = E\left[\frac{1}{v-1}\sum_{i=1}^{v}(X_i - \bar{X})^2\right] = \sigma^2$, hence $c = 1/v-1$)

$$E(T) = E\left[c\sum_{i=1}^{v}(X_i - \bar{X})^2\right] = c \cdot E\left[\sum_{i=1}^{v} X_i^2 + v\bar{X}^2 - 2\bar{X}\sum_{i=1}^{v} X_i\right] =$$
$$= cE\left[\sum_{i=1}^{v} X_i^2 - v\bar{X}^2\right] = c\left(\sum_{i=1}^{v} E[X_i^2] - vE(\bar{X}^2)\right) =$$

$$= c \cdot \left(\sum_{i=1}^{v}\left(V(X_i) + E^2(X_i)\right) - v\left(V(\bar{X}) - E^2(\bar{X})\right)\right) =$$
$$= c \cdot \left(\sum_{i=1}^{v}(\mu^2 + \sigma^2) - v\left(\mu^2 + \frac{\sigma^2}{v}\right)\right) = c\left(v\mu^2 + v\sigma^2 - v\mu^2 - \sigma^2\right) =$$
$$= c \cdot (v-1)\sigma^2$$

so for T to be an u.e. of σ^2, we need $c = 1/v-1$

Definition (Bias)

Let $U = U(\underline{x})$ be an estimator of θ. We define as **bias** the quantity $b(U) = E(U) - \theta$

Generally: if $U(\underline{X})$ estimator of $g(\theta)$, then $b(U) = E(U) - g(\theta)$

Note

Obviously if U is an u.e of θ, then $b(U) = 0$

Definition (Estimation errors)

We can choose to calculate the error as:

- $|U-\theta|$
- $[U-\theta]^2 \rightsquigarrow$ Square error
- $E[(U-\theta)^2] \rightsquigarrow$ Mean square error (MSE)

Note

- $E[(U-\theta)^2] = V(U-\theta) - (E(U-\theta))^2 = V(U) + (E(U)-\theta)^2 =$
  $= V(U) - b^2(U)$ (alternative formula for MSE)
- For unbiased estimators $b^2(U)=0$ so MSE is $V(U)$.

Example

Let $X_1, \dots, X_v$ r.s. of $f(x;\ \theta_1=\mu,\ \theta_2=\sigma^2)$

- $M_2 = \frac{1}{v}\sum_{i=1}^{v}(X_i-\bar{X})^2$

We proved that $E[\sum_{i=1}^{v}(X_i-\bar{X})^2] = (v-1)\sigma^2$, so $E[M_2] = \frac{v-1}{v}\cdot\sigma^2$ and hence M_2 is not an u.e. of σ^2

$b(M_2) = E(M_2) - \sigma^2 = \frac{v-1}{v}\cdot\sigma^2 - \sigma^2 = -\frac{\sigma^2}{v}$

- Obviously $S^2 = \frac{1}{v-1}\sum_{i=1}^{v}(X_i-\bar{X})^2$ u.e. of σ^2

Definition (Sufficiency)

If $X_1, \dots, X_v$ belong to a group $F = \{f(x;\underline{\theta}) : \underline{\theta} \in \Theta\}$ we attempt to "condense" all information of the sample $\underline{X}$ about $\underline{\theta}$ in an s.f. $\underline{T} = \underline{T}(\underline{x})$, which for the above reason is called sufficient. Formally:

(For discrete r.v): A statistic $T = T(\underline{X})$ is called sufficient for the parameter $\theta \in \Theta$ if the probability $P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n \mid T = t)$ is independent from θ (θ is not included in its formula or domain)

Note

Generally a statistic the distribution of which is independent of θ is called an Ancillary Statistic