Unbiased estimator[1]

For any parameter θ, if you can find an estimator, denoted by $\hat{\theta}$, such that

$$E(\hat{\theta}) = \theta$$

we call $\hat{\theta}$ an unbiased estimator of θ.

If

$$E(\hat{\theta}) \neq \theta$$

we say $\hat{\theta}$ has bias

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$

Example. Suppose $Y_1, ..., Y_n$ are random samples from Y with $EY = \mu$ and $Var(Y) = \sigma^2$. Then

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \qquad S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

are unbiased estimators of μ and σ^2 respectively, while

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

is not unbiased, with bias

$$E\Big\{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2\Big\} - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{1}{n}\sigma^2$$

Example. For linear regression model

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

We can get the unbiased estimators for β_0 and β_1 easily.

Remarks[2]: "Unbias" is an theoretical concept. It has no practical meaning. It does tell you how good is the estimator. A good way to measure the efficiency of an estimator is by

$$E(\hat{\theta} - \theta)^2 = Var(\hat{\theta}) + [E\hat{\theta} - \theta]^2 = Var(\hat{\theta}) + [Bias(\hat{\theta})]^2$$

[1] This materials will not be included directly in the test or examination

[2] If you find this difficult, you can ignore it