Properties of Estimators

M. S. Fadali
Professor of Electrical Engineering
University of Nevada, Reno
Outline

• Small sample properties of estimators.
  – Unbiasedness.
  – Efficiency.

• Large sample properties of estimators.
  – Consistency.
  – Asymptotic unbiasedness.
  – Asymptotic efficiency.
  – Asymptotic normality.
Estimator

• **Estimator of a parameter:**
  – Function of the data whose
  – Value assumed close to the parameter value

• **Data:** random sample

• **Estimator:** random variable.

• **Distribution of estimator:** typically unknown and complex.

• **Statistical properties of an estimator.**
Small-sample Properties

$N$ finite or infinite

1. Unbiasedness (mean).
2. Efficiency (variance).
Large-sample Properties

\[ N \to \infty \]

1. Asymptotic unbiasedness.
2. Asymptotic efficiency.
3. Asymptotic normality.
Small-sample properties

**Unbiased Estimator:** correct “on the average”

\[ E\{\hat{\theta}\} = E\{\theta\} \text{ for random } \theta \]

\[ E\{\hat{\theta}\} = \theta \text{ for constant } \theta \]
Example: Sample Mean

- $N$ i.i.d. measurements with population mean $m_z$
- Sample mean is an unbiased estimator

$$
\hat{\theta} = \bar{Z} = \frac{1}{N} \sum_{i=1}^{N} Z_i
$$

$$
E\{\hat{\theta}\} = \frac{1}{N} \sum_{i=1}^{N} E\{z_i\} = m_z = \theta
$$
Mean Squared Error

Standard Error: Standard deviation of the sampling distribution of the estimate $\sigma_{\hat{\theta}}$.

Mean Squared Error $MSE(\hat{\theta}) = E \left\{ (\hat{\theta} - \theta)^2 \right\}$

$$E \left\{ (\hat{\theta} - \theta)^2 \right\} = E\{\hat{\theta}^2\} - 2\theta E\{\hat{\theta}\} + \theta^2$$

$$= \sigma^2_{\hat{\theta}} + E\{\hat{\theta}\}^2 - 2\theta E\{\hat{\theta}\} + \theta^2$$

$$MSE(\hat{\theta}) = \sigma^2_{\hat{\theta}} + (E\{\hat{\theta}\} - \theta)^2 = var(\hat{\theta}) + bias(\hat{\theta})^2$$

Unbiased estimator $E\{\hat{\theta}\} = \theta$: $MSE(\hat{\theta}) = \sigma^2_{\hat{\theta}}$
Efficiency

Efficient Estimator: An estimator $\hat{\theta}^*(N)$ is more efficient than any other estimator if

$$MSE\{\hat{\theta}^*(N)\} \leq MSE\{\hat{\theta}(N)\}, \forall N$$

• For an unbiased estimator, $Var\{\hat{\theta}\} = MSE\{\hat{\theta}\}$

Efficient Unbiased Estimator: An unbiased estimator $\hat{\theta}^*(N)$ is more efficient than any other unbiased estimator if $\sigma^2_{\theta^*} \leq \sigma^2_{\theta}, \forall N$
Example: Sample Mean

\[ \hat{\theta} = \bar{Z} = \frac{1}{N} \sum_{i=1}^{N} z_i \]

\[ X_i \text{ i.i.d.} \Rightarrow var \left\{ \sum_{i=1}^{n} X_i \right\} = \sum_{i=1}^{n} var\{X_i\} = n\sigma_x^2 \]

\[ var\{\hat{\theta}^2\} = var \left\{ \frac{1}{N} \sum_{i=1}^{N} z_i \right\} = \frac{N\sigma_z^2}{N^2} = \frac{\sigma_z^2}{N} \]

\[ E\{\tilde{\theta}^2(k)\} = var\{\hat{\theta}^2\} = \sigma_z^2/N, \quad \tilde{\theta}(N) = \theta - \hat{\theta}(N) \]
Cramer-Rao Inequality

\[ z = \text{col}\{z_1, z_2, \ldots, z_k\} = \text{set of data} \]

- Characterized by the pdf \( f(z; \theta) = f(z) \)
- Variance of an unbiased estimator \( \hat{\theta} \) of a deterministic \( \theta \) is bounded below by

\[
E\{\hat{\theta}^2(k)\} \geq \frac{1}{E\left\{ \left[ \frac{\partial \ln[f(z)]}{\partial \theta} \right]^2 \right\}}, \forall k
\]

**Fisher information matrix:**

\[
E\left\{ \left[ \frac{\partial \ln[f(z)]}{\partial \theta} \right] \left[ \frac{\partial \ln[f(z)]}{\partial \theta} \right]^T \right\}
\]
Other expressions of Cramer-Rao

\[ E\{\tilde{\theta}^2(k)\} \geq \frac{1}{\int_{-\infty}^{\infty} \left[ \frac{\partial f(z)}{\partial \theta} \right]^2 \frac{1}{f(z)} dz} , \forall k \]

\[ dz = dz_1 \, dz_2 \ldots dz_k \]

\[ E\{\tilde{\theta}^2(k)\} \geq -\frac{1}{E \left\{ \frac{\partial^2 \ln[f(z)]}{\partial \theta^2} \right\}} , \forall k \]

- Proof requires that the derivatives exist and be absolutely integrable.
Example: Sample Mean

- Normal pdf: $C_z = \sigma^2 I_N$, \( \det(C_z) = \sigma^{2N} \)

\[
f_X(x) = \frac{1}{[2\pi]^{N/2}\sqrt{\det(C_z)}} \exp\left\{-\frac{1}{2\sigma^2}(z - m_z 1)^T(z - m_z 1)\right\}
\]

\[
z = [z_1 \ldots z_N]^T = \text{set of data}
\]

\[
\theta = m_z
\]

- Cramer-Rao lower bound

\[
E\{\tilde{\theta}^2(k)\} \geq -1/E\{\partial^2 \ln[f(z)]/\partial \theta^2\}
\]
CRLB for Sample Mean

\[-\ln[f(z)]\]

\[= \frac{1}{2\sigma^2} (z - m_z \mathbf{1})^T (z - m_z \mathbf{1}) + \frac{N}{2} \ln(2\pi\sigma^2)\]

\[= \frac{1}{2\sigma^2} (z^T z - 2m_z \mathbf{1}^T z + m_z^2 \mathbf{1}^T \mathbf{1}) + \cdots\]

\[-E\left\{\partial^2 \ln[f(z)]/\partial \theta^2\right\} = \frac{2N}{2\sigma^2}, \quad \mathbf{1}^T \mathbf{1} = N\]

Cramer-Rao lower bound \(= \sigma^2 / N = E\{\tilde{\theta}^2(k)\}\)

Efficient estimator for any variance \(\sigma^2 > 0\).
Consistent Estimator

\[ P\left[\|\theta - \hat{\theta}(N)\| > \epsilon\right] \to 0 \text{ as } N \to \infty \]

\( \hat{\theta}(N) = \) estimate based on \( N \) data points.

- Convergence in probability.
- Using a lot of data tends to give a better estimate.
Example: Sample Mean

\[ \theta = m_z \quad \hat{\theta} = \bar{z} = \frac{1}{N} \sum_{i=1}^{N} z_i \]

\[ E\{\hat{\theta}\} = \theta, \quad \text{var}\{\hat{\theta}\} = \frac{\sigma^2}{N} \]

- Chebychev Inequality

\[ P[|\hat{\theta} - E\{\hat{\theta}\}| \geq \epsilon] \leq \frac{\text{var}\{\hat{\theta}\}}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2} \]

- Convergence in probability (unbiased)

\[ P[|\hat{\theta}(N) - \theta| > \epsilon] \to 0 \text{ as } N \to \infty \]
Other Asymptotic Properties

- **Asymptotic unbiasedness**
  \[
  \lim_{N \to \infty} E\{\hat{\theta}(N)\} = E\{\theta\}, \quad \theta \text{ random}
  \]
  \[
  \lim_{N \to \infty} E\{\hat{\theta}(N)\} = \theta, \quad \theta \text{ constant}
  \]

- **Asymptotic efficiency** of consistent estimator: more efficient than any other consistent estimator (approaches the Cramer-Rao lower bound as \(N \to \infty\)).

- **Asymptotic normality**: converges in distribution to a normal distribution.
Example: Sample Variance

• Asymptotically Unbiased:

\[
\hat{\theta} = \frac{1}{N} \sum_{i=1}^{N} (z_i - \bar{z})^2 = \left( \frac{1}{N} \sum_{i=1}^{N} z_i^2 \right) - \bar{z}^2
\]

\[
E\{\hat{\theta}\} = \left( \frac{1}{N} \sum_{i=1}^{N} E\{z_i^2\} \right) - E\{\bar{z}^2\}
\]

• Use: (i) \( E\{z^2\} = \sigma^2 + m_z^2 \)

(ii) \( var\{\bar{z}\} = \sigma^2 / N, \; E\{\bar{z}\} = m_z \)

\[
E\{\hat{\theta}\} = \sigma^2 + m_z^2 - \left( \frac{\sigma^2}{N} + m_z^2 \right) = \frac{N - 1}{N} \sigma^2
\]
References