
Unbiased Estimator of $\theta$, say $\overline{\hat{\theta}}_1, \overline{\hat{\theta}}_2, \overline{\hat{\theta}}_3$ ...

It could be really difficult for us to compare $\text{Var}(\overline{\hat{\theta}}_1)$ when there are many of them.

**Theorem. Cramer-Rao Lower Bound**

Let $Y_1, Y_2, ..., Y_n$ be a random sample from a population with p.d.f. $f(y; \theta)$. Let $\overline{\hat{\theta}} = h(y_1, y_2, ..., y_n)$ be an unbiased estimator of $\theta$.

Given some regularity conditions (continuous differentiable etc.) and the domain of $f(y_i; \theta)$ does not depend on $\theta$. Then, we have

$$\text{Var}(\overline{\hat{\theta}}) \geq \frac{1}{nE[(\frac{\partial \text{Inf}(\theta)}{\partial \theta})^2]} = \frac{1}{-nE[\frac{\partial^2 \text{Inf}(\theta)}{\partial \theta^2}]}$$

**Theorem. Properties of the MLE**

Let $Y_i \overset{i.i.d.}{\sim} f(y; \theta), i = 1, 2, ..., n$

Let $\hat{\theta}$ be the MLE of $\theta$, then

$$\hat{\theta} \overset{n \to \infty}{\sim} N\left(\theta, \frac{1}{nE[(\frac{\partial \text{Inf}(\theta)}{\partial \theta})^2]}\right)$$

The MLE is asymptotically unbiased and its asymptotic variance: C-R lower bound
Harald Cramér (left) was born in Stockholm, Sweden on September 25, 1893, and died there on October 25, 1985. (wiki)

Calyampudi Radhakrishna Rao (right), FRS known as C R Rao (born 10 September 1920) is an Indian statistician. He is professor emeritus at Penn State University and Research Professor at the University at Buffalo. Rao was awarded the US National Medal of Science in 2002. (wiki)

Example 1. Let $Y_1, Y_2, \ldots, Y_n \sim \text{i.i.d. \ Bernoulli}(p)$

1. MLE of $p$?
2. What are the mean and variance of the MLE of $p$?
3. What is the Cramer-Rao lower bound for an unbiased estimator of $p$?

Solution. $P(Y = y) = f(y; p) = p^y(1 - p)^{1-y}, y = 0,1$

1. 
\[
    L = \prod_{i=1}^{n} f(y_i; p) = \prod_{i=1}^{n} [p^{y_i}(1 - p)^{1-y_i}] = p^{\sum y_i} (1 - p)^{n - \sum y_i}
\]

\[
l = \ln L = \left( \sum y_i \right) \ln p + (n - \sum y_i) \ln (1 - p)
\]

Solving:
\[
    \frac{dl}{dp} = \frac{\sum y_i}{p} - \frac{n - \sum y_i}{1 - p} = 0,
\]

we have the MLE:
\[
    \hat{p} = \frac{\sum_{i=1}^{n} y_i}{n}
\]

2.
\[
    E(\hat{p}) = p, \ Var(\hat{p}) = \frac{p(1 - p)}{n}
\]

3.
\[
    \ln f(y, p) = y \ln p + (1 - y) \ln (1 - p)
\]
\[
    \frac{\partial \ln f(y, p)}{\partial p} = \frac{y}{p} - \frac{1 - y}{1 - p}
\]
\[
    \frac{\partial^2 \ln f(y, p)}{\partial p^2} = -\frac{y}{p^2} - \frac{1 - y}{(1 - p)^2}
\]
Define $\hat{p}$ as the MLE of $p$.

Thus, the MLE of $p$ is unbiased and its variance = C-R lower bound.

**Definition. Efficient Estimator**

If $\hat{\theta}$ is an unbiased estimator of $\theta$ and its variance = C-R lower bound, then $\hat{\theta}$ is an efficient estimator of $\theta$.

**Definition. Best Estimator**

If $\hat{\theta}$ is an unbiased estimator of $\theta$ and $\text{var}(\hat{\theta}) \leq \text{var}(\tilde{\theta})$ for all unbiased estimator $\tilde{\theta}$, then $\hat{\theta}$ is a best estimator for $\theta$.

Always true

**Efficient Estimator** $\Rightarrow$ **Best Estimator**

May not be true

**Example 2.** If $Y_1, Y_2, ..., Y_n$ is a random sample from $f(y; \theta) = \frac{2y}{\theta^2}, 0 < y < \theta$, then $\hat{\theta} = \frac{3}{2} \bar{Y}$ is an unbiased estimator for $\theta$.

Compute 1. $\text{Var}(\hat{\theta})$ and 2. C-R lower bound for $f_Y(y; \theta)$

**Solution.**

1. $\text{Var}(\hat{\theta}) = \text{Var}\left(\frac{3}{2} \bar{Y}\right) = \frac{9}{4} \text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} Y_i\right) = \frac{9}{4n^2} \sum_{i=1}^{n} \text{Var}(Y_i)$

$\text{Var}(Y_i) = E(Y_i^2) - [E(Y_i)]^2 = \int_{0}^{\theta} y^2 \frac{2y}{\theta^2} dy - \left[ \int_{0}^{\theta} y^2 \frac{2y}{\theta^2} dy \right]^2 = \frac{\theta^2}{18}$

Therefore, $\text{Var}(\hat{\theta}) = \frac{9}{4n^2} \frac{n\theta^2}{18} = \frac{\theta^2}{8n}$

2. C-R lower bound

$\ln f_Y(y; \theta) = \ln \left(\frac{2y}{\theta^2}\right) = \ln 2y - 2\ln \theta$
The value of 1 is less than the value of 2. But, it is NOT a contradiction to the theorem. Because the domain, \(0 < y < \theta\), depends on \(\theta\). Thus the C-R Theorem doesn’t hold for this problem.

**Example 3.** Let \(X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma^2)\), be a random sample from the normal population where both \(\mu\) and \(\sigma^2\) are unknown. Please derive

1. The maximum likelihood estimators for \(\mu\) and \(\sigma^2\).
2. The best estimator for \(\mu\) assuming that \(\sigma^2\) is known.

**Solution:**

(1) MLEs for \(\mu\) and \(\sigma^2\). The Likelihood function is:

\[
L = \prod_{i=1}^{n} f(x_i; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2}
\]

\[
\ln L = (-n) \ln(\sqrt{2\pi\sigma^2}) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i-\mu)^2
\]

\[
\frac{\partial \ln L}{\partial \mu} = 2 \sum_{i=1}^{n} (x_i - \mu) = 0 \Rightarrow \hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n}
\]

\[
\frac{\partial \ln L}{\partial \sigma^2} = \left(\frac{-n}{2}\right) \frac{1}{\sigma^2} + \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{4\sigma^4} = 0 \Rightarrow \hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (x_i - \hat{\mu})^2}{n} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}
\]

(2) Use the Cramer-Rao lower bound:
\[ f_X(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]
\[ \ln f = \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(x-\mu)^2}{2\sigma^2} \]
\[ \frac{d\ln f}{d\mu} = \frac{(x-\mu)}{\sigma^2} \]
\[ \frac{d^2\ln f}{d\mu^2} = -\frac{1}{\sigma^2} \]

Hence, the C-R lower bound of variance is \( \frac{\sigma^2}{n} \).

Since the normal pdf satisfies all the regularity conditions for the C-R lower bound theorem to hold, and since

\[ X \sim N \left( \mu, \frac{\sigma^2}{n} \right) \]

The variance of \( X \) equals to the C-R lower bound, and thus this unbiased estimator is an efficient estimator for \( \mu \), and thus it is also a best estimator for \( \mu \).

**Dear students:** This part of our study is very important (*will be in the midterm & final).

Please study the corresponding chapters/sections for our latest four lectures (lectures 6, 7, 8, 9):

Section 4.2 – Sampling from normal populations

Section 4.3 – Order Statistics

Section 5.1 – for an overview of Point Estimators

Section 5.2.1 – MOME

Section 5.2.2 – MLE

Section 5.3.1 – Unbiased Estimators

Note: Please refer to this lecture note (rather than the textbook) for the Theorem on the Cramer-Rao Lower Bound, and the concepts of Efficient Estimator and Best Estimator (also called the Minimum Variance Unbiased Estimator or Uniformly Minimum Variance Unbiased Estimator, and abbreviated as **MVUE** or **UMVUE**).
Midterm will be held in class on Tuesday, March 10