Introduction to mathematical Statistics

Cramer-Rao lower bound

Unbiased Estimator of θ, say $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \cdots$

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

Theorem. Cramer-Rao Lower Bound

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

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

Then, we have

$$\text{var}(\hat{\theta}) \geq \frac{1}{n \cdot E\left[\frac{\partial(\ln f(\theta))}{\partial \theta}\right]^2} = \frac{1}{-n \cdot E\left[\frac{\partial^2(\ln f(\theta))}{\partial \theta^2}\right]}$$

Theorem. Properties of the MLE

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

Let $\hat{\theta}$ be the MLE of θ

Then, $\hat{\theta} \overset{n \to \infty}{\to} N(\theta, \frac{1}{n \cdot E\left[\frac{\partial \ln(f(\theta))}{\partial \theta}\right]^2})$

The MLE is asymptotically unbiased and its asymptotic variance : C-R lower bound

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

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. $$L = \prod_{i=1}^{n} f(y_i;p)$$
$$= \prod_{i=1}^{n} p^{y_i}(1-p)^{1-y_i}$$
$$= p^{\sum y_i}(1-p)^{n-\sum y_i}$$
$$l = \ln L = (\sum y_i)\ln p + (n - \sum y_i)\ln(1-p)$$
$$\frac{\partial l}{\partial p} = \frac{\sum y_i}{p} - \frac{n-\sum y_i}{1-p} = 0$$

$$\hat{p} = \frac{\sum_{i=1}^{n} Y_i}{n} \text{ : MLE}$$

2. $$E(\hat{p}) = p$$
$$\text{var}(\hat{p}) = \frac{p(1-p)}{n}$$

3. $$\ln f(y,p) = y\ln p + (1-y)\ln(1-p)$$

$$\frac{\partial \ln f(y,p)}{\partial p} = \frac{y}{p} - \frac{1-y}{1-p}$$
$$\frac{\partial^2 \ln f(y,p)}{\partial p^2} = -\frac{y}{p^2} - \frac{1-y}{(1-p)^2}$$
$$E\left[-\frac{Y}{p^2} - \frac{1-Y}{(1-p)^2}\right] = -\frac{p}{p^2} - \frac{1-p}{(1-p)^2} = -\frac{1}{p(1-p)}$$

C-R lower bound

$$\text{var}(\hat{p}) \ge \frac{1}{-nE\left[\frac{\partial^2 \ln f}{\partial p^2}\right]} = \frac{p(1-p)}{n}$$

The MLE of p is unbiased and its variance = C-R lower bound.

Definition. **Efficient Estimator**

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

Definition. **Best Estimator**

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

Always true

Efficient Estimator ⇌ **Best Estimator**

May not be true

Example 2. If $Y_1, Y_2, \cdots, Y_n$ is a random sample from $f(y,\theta) = \frac{2y}{\theta^2},\ 2 \prec y \prec \theta$, then $\hat{\theta} = \frac{3}{2}\overline{Y}$ is an unbiased estimator for θ.

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

Solution.

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

$$Var(Y_i) = E(Y_i^2) - [E(Y_i)]^2 = \int_0^\theta y^2 \frac{2y}{\theta^2} dy - \left[\int_0^\theta y^2 \frac{2y}{\theta^2} dy\right]^2 = \frac{\theta^2}{18}$$

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

2. C-R lower bound

$$\ln f_Y(y,\theta) = \ln(\frac{2y}{\theta^2}) = \ln 2y - 2\ln\theta$$

$$\frac{\partial \ln f_Y(y,\theta)}{\partial\theta} = -\frac{2}{\theta}$$

$$E\left[\left(\frac{\partial \ln f_Y(y,\theta)}{\partial\theta}\right)^2\right] = E\left(\frac{4}{\theta^2}\right) = \int_0^\theta \frac{4}{\theta^2}\frac{2y}{\theta^2}dy = \frac{4}{\theta^2}$$

$$\therefore \frac{1}{nE\left[\left(\frac{\partial \ln f_Y(y,\theta)}{\partial\theta}\right)^2\right]} = \frac{\theta^2}{4n}$$

The value of 1 is less than the value of 2. But, it is **NOT** a contradiction to the theorem. Because $0 \prec y \prec \theta$, the domain depends on θ. Thus the C-R Theorem doesn't hold for this problem.