Lecture 8: Cramer–Rao Bound

An important result concerning the minimum variance attainable by an unbiased estimator follows.

**Result 1 (Cramer–Rao Lower Bound)**

Let $X$ be a random variable (or vector) with pdf/pmf $f_X(x; \theta)$, and let $W$ be an unbiased estimator for $\theta$ (with $E|W| < \infty$). Suppose the order of differentiation and integration can be interchanged as follows:

\[
\frac{d}{d\theta} \int W(x) f(x; \theta_0) dx = \int W(x) \frac{\partial}{\partial \theta} f(x; \theta_0) dx \quad \text{and} \quad \frac{d}{d\theta} \int f_X(x; \theta_0) dx = \int \frac{\partial f_X}{\partial \theta}(x; \theta_0) dx.
\]

Then

\[
V(W) \geq \frac{1}{E \left[ \frac{\partial \ln f}{\partial \theta}(X; \theta_0) \right]^2}.
\]

**Proof:** Note that the square of the covariance of two random variables is less than or equal to the product of the variances (that is the same as saying that the correlation coefficient is less than or equal to one in absolute value):

\[
\text{Cov}^2(S, W) \leq V(S) \cdot V(W).
\]

Now let us take $S = \frac{\partial \ln f}{\partial \theta}(X; \theta)$. First consider the expectation of $S$, known as the score function:

\[
1 = \int f_X(x; \theta) dx.
\]

So,

\[
0 = \frac{\partial}{\partial \theta} \int f_X(x; \theta) dx.
\]

Assuming we can change the order of differentiation and integration, we get

\[
0 = \int \frac{\partial f_X}{\partial \theta}(x; \theta) dx
\]

\[
= \int \frac{\partial \ln f_X}{\partial \theta}(x; \theta) \cdot f_X(x; \theta) dx.
\]
\[ E[\frac{\partial \ln f_X(x; \theta)}{\partial \theta}] = E[S] = 0. \]

Therefore the covariance of \( W \) and \( S \) is the expectation of the product of \( S \) and \( W \):

\[ E[SW] = \int W \frac{\partial \ln f_X(x; \theta)}{\partial \theta} f_X(x; \theta) dx = \int W \frac{\partial f_X(x; \theta)}{\partial \theta} dx \]

\[ = \frac{\partial}{\partial \theta} \int_x W f(x; \theta) dx = \frac{\partial}{\partial \theta} \theta = 1. \]

So

\[ 1 \leq V(W) \cdot V(S), \]

implying

\[ V(W) \geq 1/V(S) = 1/E[S^2]. \]

Of course finding a lower bound for the variance is not so hard. Zero is a lower bound that applies with no conditions attached. The interest in the Cramer–Rao bound stems largely from the fact that in many cases the bound can actually be reached; there are often estimators with variance equal to the bound.

**Example**

Suppose \( X \) has an exponential distribution with mean \( \mu \). Consider the estimator \( \hat{\mu} = X \). This estimator is unbiased with variance \( \mu^2 \). To calculate the Cramer–Rao Bound, consider the log of the density \( (f(x; \mu) = \frac{1}{\mu} \exp(-\frac{x}{\mu}))\):

\[ \ln f(x; \mu) = -\ln(\mu) - x/\mu. \]

The derivative of the log of the density, the score function is

\[ \frac{\partial}{\partial \mu} \ln f(x; \mu) = -1/\mu + x/\mu^2 = (x - \mu)/\mu^2. \]

Clearly this has expectation zero. The variance of the score is

\[ E \left[ \frac{\partial}{\partial \mu} \ln f(x; \mu) \right]^2 = E(X - \mu)^2/\mu^4 = 1/\mu^2, \]
and so the Cramer–Rao bound is $\mu^2$. This is the variance of the unbiased estimator $\hat{\mu}$ suggested, so that estimator is the minimum variance unbiased estimator.

A corollary of the Cramer–Rao bound is the following result for $N$ iid random variables.

**Result 2** Let $X_1, \ldots, X_N$ be iid random variable with common pdf/pmf $f_X(x; \theta)$, and let $W$ be an unbiased estimator for $\theta$. Then

$$V(W) \geq \frac{1}{N \cdot E\left[\frac{\partial \ln f}{\partial \theta}(x; \theta)\right]^2}.$$ 

**Example**

Suppose $X_1, \ldots, X_N$ are independent with normal distributions with mean $\mu$ and known variance $\sigma^2$. The obvious estimator for the mean is the sample average $\bar{x}$ with variance $\sigma^2/N$. Consider the log of the density function:

$$\ln f_X(x; \mu) = -\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(x - \mu)^2.$$ 

The score function is

$$\frac{\partial}{\partial \mu} \ln f_X(x; \mu) = \frac{1}{\sigma^2}(x - \mu).$$ 

Again the score clearly has expectation zero, and the variance is $\sigma^2/\sigma^4 = 1/\sigma^2$, and therefore the Cramer–Rao bound for the single observation case is $\sigma^2$, and the Cramer–Rao bound for the $N$ observation case is $\sigma^2/N$.

**Example**

Finally let us consider an example where the Cramer–Rao bound does not apply. Recall that in the proof we have to be able to reverse the order of integration and differentiation. That does not work if the argument of the function enters in the bounds of the integral. Suppose $X$ has a uniform distribution on the interval from zero to $\theta$. The log of the density function is

$$\ln f_X(x; \theta) = -\ln \theta.$$
The derivative is
\[ \frac{\partial}{\partial \theta} \ln f_X(x; \theta) = -1/\theta. \]

Note that this clearly does not have expectation zero, which is a property we used in the proof of the Cramer–Rao bound. Nevertheless, let us ignore this and proceed with the calculation. The expectation of the square is $1/\theta^2$, and the Cramer–Rao bound is equal to $\theta^2$. Now consider the estimator $2X$. It clearly is unbiased. The variance is $\theta^2 \cdot 4/12 = \theta^2/3$, lower than the supposed Cramer–Rao bound that does not apply in this case.