Cramer-Rao lower bound: an example

Suppose that $\underline{X} = (X)$, a single observation from $Bin(m, p)$, where m is *known*. The pmf is given by

$$f(x;p) = \binom{m}{x} p^x (1-p)^{m-x} \quad \text{where} \quad x = 0, 1, ..., m.$$

Note that the *range* of X depends on m, but *not* on the unknown parameter p. Also, the sample size is $n = 1$.

Cramer-Rao lower bound

Since the range of X does not depend on the unknown parameter p which we wish to estimate, we can proceed to compute and use the Cramer-Rao lower bound for unbiased estimators:

$$\log f(x;p) = \log \binom{m}{x} + x \log p + (m-x)\log(1-p)$$

$$\frac{\partial}{\partial p} \log f(x;p) = \frac{x}{p} - \frac{m-x}{1-p} = \frac{x-mp}{p(1-p)}$$

$$\left(\frac{\partial}{\partial p} \log f(x;p)\right)^2 = \frac{(x-mp)^2}{p^2(1-p)^2}.$$

Thus,

$$E\left(\left(\frac{\partial}{\partial p} \log f(X;p)\right)^2\right) = \frac{E(X-mp)^2}{p^2(1-p)^2} = \frac{Var(X)}{p^2(1-p)^2} = \frac{m}{p(1-p)}.$$

It follows that for any *unbiased* estimator, $g(\underline{X})$, for p, we have

$$Var(g(\underline{X}) \geq \frac{1}{1 \cdot \frac{m}{p(1-p)}} = \frac{p(1-p)}{m}.$$

Alternatively, we can compute the Cramer-Rao lower bound as follows:

$$\frac{\partial^2}{\partial p^2} \log f(x;p) = \frac{\partial}{\partial p}\left(\frac{\partial}{\partial p} \log f(x;p)\right) = \frac{\partial}{\partial p}\left(\frac{x}{p} - \frac{m-x}{1-p}\right) = \frac{-x}{p^2} - \frac{(m-x)}{(1-p)^2}.$$

Thus,

$$E\left(\frac{\partial^2}{\partial p^2} \log f(X;p)\right) = \frac{-E(X)}{p^2} - \frac{(m-E(X))}{(1-p)^2} = \frac{-mp}{p^2} - \frac{(m-mp)}{(1-p)^2} = \frac{-m}{p(1-p)}.$$

It follows that the Cramer-Rao lower bound is given by

$$\frac{1}{-nE\left(\frac{\partial^2}{\partial p^2} \log f(X;p)\right)} = \frac{1}{-1 \cdot \frac{-m}{p(1-p)}} = \frac{p(1-p)}{m}$$

as above.

Comparing estimators

Consider the estimator $g_1(\underline{X}) = \frac{X}{m}$.

$$E(g_1(\underline{X})) = \frac{E(X)}{m} = \frac{mp}{m} = p$$

so $g_1(\underline{X})$ is an *unbiased* estimator of p. Is it the most efficient unbiased estimator for p? To answer this question, we compute the variance of g_1 and compare it to the Cramer-Rao lower bound which we calculated above.

$$Var(g_1(\underline{X})) = Var(\frac{X}{m}) = \frac{Var(X)}{m^2} = \frac{mp(1-p)}{m^2} = \frac{p(1-p)}{m}.$$

Since $Var(g_1)$ equals the Cramer-Rao lower bound, we can conclude that $g_1(\underline{X})$ is the most efficient *unbiased* estimator for p.

Now consider the estimator $g_2(\underline{X}) = \frac{X+1}{m+2}$.

$$E(g_2(\underline{X})) = \frac{E(X)+1}{m+2} = \frac{mp+1}{m+2} \neq p \quad \text{(except when } p = 1/2\text{)}.$$

So g_2 is a biased estimator with

$$bias(g_2) = E(g_2(\underline{X})) - p = \frac{mp+1}{m+2} - p = \frac{1-2p}{m+2}.$$

To compare the performance of g_2 with the performance of g_1, we must first compute the mean square error of g_2:

$$Var(g_2(\underline{X})) = Var(\frac{X+1}{m+2}) = \frac{Var(X+1)}{(m+2)^2} = \frac{mp(1-p)}{(m+2)^2}.$$

Thus,

$$MSE(g_2) = Var(g_2)+(bias(g_2))^2 = \frac{mp(1-p)}{(m+2)^2}+\frac{(1-2p)^2}{(m+2)^2} = \frac{1}{(m+2)^2}(1+(m-4)p-(m-4)p^2).$$

We can compare the (relative) efficiency of g_1 and g_2 by comparing the *graphs* of $MSE(g_1)$ (which is just the variance of g_1) and $MSE(g_2)$ as *functions of p.*

Exercise: Fix $m = 10$ and sketch the graphs of $MSE(g_1)$ and $MSE(g_2)$ as *functions of p.* Also, determine the values of p for which g_2 is more efficient than g_1.