Solutions to Assignment #15

1. Suppose that when the radius of a disc in the plane is measured, an error is made that has a normal$(0, \sigma^2)$ distribution. If n independent measurements are made, find an unbiased estimator for the area of the disc. Is this the best unbiased estimator for the area? Assume that σ^2 is known.

   ***Solution**:* Let A denote the area of the disc and $R_1, R_2, \ldots, R_n$ denote n independent measurements of the radius of the disc. By the information given in the problem, we may assume that

   $$R_i = \sqrt{\frac{A}{\pi}} + E_i, \quad \text{for } i = 1, 2, \ldots, n,$$

   where $E_1, E_2, \ldots, E_n$ are iid normal$(0, \sigma^2)$ random variables. It then follows that $R_1, R_2, \ldots, R_n$ are normal(μ, σ^2) random variables, where $\mu = \sqrt{\dfrac{A}{\pi}}$. Then, the sample mean, $\overline{R}_n$, is an unbiased estimator for μ. It is the best unbiased estimator for μ in the sense that

   $$\text{var}(\overline{R}_n) = \frac{\sigma^2}{n}$$

   is the Crámer–Rao lower bound. To see why this is the case, compute the information function

   $$I(\mu) = -E\left(\frac{\partial^2}{\partial \mu^2} \ln f(R \mid \mu, \sigma^2)\right),$$

   where

   $$f(r \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-(r-\mu)^2/2\sigma^2}, \quad \text{for } r \in \mathbb{R},$$

   so that

   $$\ln f(r \mid \mu, \sigma^2) = -\frac{1}{2\sigma^2}\ (r-\mu)^2 - \ln(\sqrt{2\pi}\ \sigma).$$

   Thus,

   $$\frac{\partial}{\partial \mu} \ln f(r \mid \mu, \sigma^2) = \frac{1}{\sigma^2}(r-\mu),$$

   and

   $$\frac{\partial^2}{\partial \mu^2} \ln f(r \mid \mu, \sigma^2) = -\frac{1}{\sigma^2}.$$

We then have that

$$I(\mu) = -E\left(-\frac{1}{\sigma^2}\right) = \frac{1}{\sigma^2}.$$

Consequently, the Crámer–Rao lower bound is

$$\frac{1}{nI(\mu)} = \frac{\sigma^2}{n},$$

which is attained by the variance of the sample mean, $\overline{R}_n$. Hence, $\overline{R}_n$ provides and unbiased estimator of $\sqrt{\frac{A}{\pi}}$ with the lowest possible MSE. Thus, the formula $\pi(\overline{R}_n)^2$ provides the best unbiased estimator for the area, A, of the disc. □

2. Let $X_1, X_2, \ldots, X_n$ be iid Bernoulli(p) random variables. Show that the MLE for p is an efficient estimator.

***Solution**:* The MLE for p is the sample proportion $\widehat{p} = \overline{X}_n$. Thus, $\widehat{p}$ is also and unbiased estimator for p. The variance of this estimator is

$$\text{var}(\widehat{p}) = \frac{p(1-p)}{n}.$$

To see that this in the Crámer–Rao lower bound, we compute the information

$$I(p) = -E\left(\frac{\partial^2}{\partial p^2}\ln f(x \mid \; p)\right),$$

where

$$f(x \mid p) = p^x(1-p)^{1-x}, \quad \text{for } x = 0, 1.$$

Then,

$$\ln f(x \mid p) = x\ln p + (1-x)\ln(1-p),$$

$$\frac{\partial}{\partial p}\ln f(x \mid p) = \frac{x}{p} - \frac{(1-x)}{1-p},$$

and

$$\frac{\partial^2}{\partial p^2}\ln f(x \mid p) = -\frac{x}{p^2} - \frac{(1-x)}{(1-p)^2}.$$

Thus,

$$\begin{aligned} I(p) &= -E\left(-\frac{X}{p^2} - \frac{(1-X)}{(1-p)^2}\right) \\ &= \frac{1}{p} + \frac{1}{1-p} \\ &= \frac{1}{p(1-p)}. \end{aligned}$$

Consequently, the Crámer–Rao lower bound is

$$\frac{1}{nI(p)} = \frac{p(1-p)}{n},$$

which is attained by $\text{var}(\widehat{p})$. Hence, $\widehat{p}$ is an efficient estimator of p. □

3. Let $X_1, X_2, \ldots, X_n$ be iid exponential(β) random variables, and define

$$Y = \min\{X_1, X_2, \ldots, X_n\}.$$

Find an unbiased estimator, W, based only on Y. Compute $\text{var}(W)$ and compare it to the variance of the sample mean, $\overline{X}_n$. Which of W or $\overline{X}_n$ is a more efficient estimator?

***Solution**:* The common distribution function of the X_is is

$$f_X(x \mid \beta) = \begin{cases} \dfrac{1}{\beta}\, e^{-x/\beta} & \text{if } \ x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Thus, the common cdf if

$$F_X(x \mid \beta) = \begin{cases} 1 - e^{-x/\beta} & \text{if } \ x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

To find the distribution function of $Y \min\{X_1, X_2, \ldots, X_n\}$, we first compute the cdf

$$\begin{aligned} F_Y(y \mid \beta) &= \text{P}(Y \leqslant y) \\ &= 1 - \text{P}(Y > y) \\ &= 1 - \text{P}(X_1 > y, X_2 > y, \ldots, X_n > y) \\ &= 1 - \text{P}(X_1 > y) \cdot \text{P}(X_2 > y) \cdots \text{P}(X_n > y), \end{aligned}$$

where we have used the independence of the X_is. Consequently, using the assumption that the X_is are identically distributed, we obtain

$$\begin{aligned} F_Y(y \mid \beta) &= 1 - [\mathrm{P}(X > y)]^n \\ &= 1 - [1 - \mathrm{P}(X \leqslant y)]^n \\ &= 1 - [1 - F_X(y \mid \beta)]^n. \end{aligned}$$

Thus, differentiating with respect to y we have that

$$f_Y(y \mid \beta) = n[1 - F_X(y \mid \beta)]^{n-1} f_X(y),$$

where we have used the Chain Rule. It then follows that

$$f_Y(y \mid \beta) = \begin{cases} \dfrac{n}{\beta}\, e^{-ny/\beta} & \text{if } x > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Consequently, the expected value of Y is

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f_Y(y \mid \beta)\ \mathrm{d}y \\ &= \int_0^{\infty} y\ \frac{n}{\beta} e^{-ny/\beta}\ \mathrm{d}y \\ &= \frac{\beta}{n}. \end{aligned}$$

We then have that $E(nY) = \beta$. Thus, if we set $W = nY$, we see that W is an unbiased estimator of β.

Observe that $f_Y(y \mid \beta)$ is the pdf of an exponential(β/n) distribution. It then follows that

$$\mathrm{var}(Y) = \frac{\beta^2}{n^2}.$$

Therefore,

$$\mathrm{var}(W) = \mathrm{var}(nY) = n^2 \mathrm{var}(Y) = \beta^2.$$

On the other hand, $\overline{X}_n$ is also an unbiased estimator of β. However,

$$\mathrm{var}(\overline{X}_n) = \frac{\beta^2}{n} < \beta^2$$

for $n > 1$. We then have that $\mathrm{var}(\overline{X}_n) < \mathrm{var}(W)$ and therefore $\overline{X}_n$ is more efficient than W. □

4. Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal(μ, σ^2) distribution. Prove that the sample mean, $\overline{X}_n$, is an efficient estimator of μ for every known $\sigma^2 > 0$.

   Solution: The information function is

$$I(\mu) = -E\left(\frac{\partial^2}{\partial\mu^2} \ln f(X \mid \mu, \sigma^2)\right),$$

   where

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\ \sigma}\ e^{-(x-\mu)^2/2\sigma^2}, \quad \text{for } x \in \mathbb{R},$$

   so that

$$\ln f(x \mid \mu, \sigma^2) = -\frac{1}{2\sigma^2}\ (x-\mu)^2 - \ln(\sqrt{2\pi}\ \sigma).$$

   Thus,

$$\frac{\partial}{\partial\mu} \ln f(x \mid \mu, \sigma^2) = \frac{1}{\sigma^2}(x-\mu),$$

   and

$$\frac{\partial^2}{\partial\mu^2} \ln f(x \mid \mu, \sigma^2) = -\frac{1}{\sigma^2}.$$

   We then have that

$$I(\mu) = -E\left(-\frac{1}{\sigma^2}\right) = \frac{1}{\sigma^2}.$$

   Consequently, the Crámer–Rao lower bound is

$$\frac{1}{nI(\mu)} = \frac{\sigma^2}{n},$$

   which is attained by the variance of the sample mean, $\overline{X}_n$. Hence, $\overline{X}_n$ is an efficient estimator of μ for every known $\sigma^2 > 0$. □

5. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a uniform distribution over the interval $[0, \theta]$ for some parameter $\theta > 0$.

   Let $Y = \max\{X_1, X_2, \ldots, X_n\}$ and define $W = \dfrac{n+1}{n} Y$. Compute the variance W. Is W an efficient estimator of θ?

***Solution**:* We saw in Problem 5 of Assignment 14 that the pfd of Y is given by

$$f_Y(y \mid \theta) = \begin{cases} \dfrac{ny^{n-1}}{\theta^n} & \text{if } 0 \leqslant y \leqslant \theta; \\ \\ 0 & \text{otherwise,} \end{cases}$$

and that

$$E_\theta(Y) = \frac{n}{n+1}\,\theta.$$

It then follows that

$$E(W) = E\left(\frac{n+1}{n}Y\right) = \frac{n+1}{n}E(Y) = \theta.$$

Hence, W is an unbiased estimator of θ.
To find the variance of W, we first compute the variance of Y:

$$\text{var}(Y) = E(Y^2) - [E(Y)]^2,$$

where

$$\begin{aligned} E(Y^2) &= \int_{-\infty}^{\infty} y^2 f_Y(y \mid \theta)\ \mathrm{d}y \\ &= \int_0^\theta y^2\ \frac{ny^{n-1}}{\theta^n}\ \mathrm{d}y \\ &= \frac{n}{\theta^n}\int_0^\theta y^{n+1}\ \mathrm{d}y \\ &= \frac{n}{n+2}\,\theta^2. \end{aligned}$$

Therefore,

$$\begin{aligned} \text{var}(Y) &= \frac{n}{n+2}\,\theta^2 - \left[\frac{n}{n+1}\,\theta\right]^2 \\ &= \frac{n}{n+2}\,\theta^2 - \frac{n^2}{(n+1)^2}\,\theta^2 \\ &= \frac{n}{(n+2)(n+1)^2}\,\theta^2. \end{aligned}$$

We then have that

$$\begin{aligned} \text{var}(W) &= \text{var}\left(\frac{n+1}{n}Y\right) \\ &= \frac{(n+1)^2}{n^2}\text{var}\,(Y) \\ &= \frac{(n+1)^2}{n^2}\;\frac{n}{(n+2)(n+1)^2}\;\theta^2 \\ &= \frac{1}{n(n+2)}\;\theta^2. \end{aligned}$$

To see if W is an efficient estimator, we compute the information

$$I(\theta) = \text{var}_\theta\left(\frac{\partial}{\partial\theta}\ln(f(X \mid \theta))\right) = E_\theta\left(\left[\frac{\partial}{\partial\theta}\ln(f(X \mid \theta))\right]^2\right),$$

where

$$f(x \mid \theta) = \begin{cases} \dfrac{1}{\theta} & \text{if } 0 \leqslant x \leqslant \theta; \\ 0 & \text{otherwise,} \end{cases}$$

so that

$$\ln(f(x \mid \theta)) = \begin{cases} -\ln\theta & \text{if } 0 \leqslant x \leqslant \theta; \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\frac{\partial}{\partial\theta}\ln(f(x \mid \theta)) = \begin{cases} -\dfrac{1}{\theta} & \text{if } 0 \leqslant x \leqslant \theta; \\ 0 & \text{otherwise.} \end{cases}$$

Thus,

$$I(\theta) = \int_0^\theta \left(-\frac{1}{\theta}\right)^2 \frac{1}{\theta}\;\mathrm{d}x = \frac{1}{\theta^2}.$$

We then see that the Crámer–Rao lower bound is

$$\frac{1}{nI(\theta)} = \frac{\theta^2}{n}.$$

Note that this is larger than $\text{var}(W) = \dfrac{1}{n(n+2)}\,\theta^2$. Thus, the Crámer–Rao inequality does not apply to this situation. To see why this is so, note that for any function g of x,

$$\begin{aligned}
\frac{d}{d\theta}\int_{-\infty}^{\infty} g(x)f(x \mid \theta)\ \mathrm{d}x &= \frac{d}{d\theta}\int_0^\theta g(x)\ \frac{1}{\theta}\ \ \mathrm{d}x \\
&= \frac{d}{d\theta}\left(\frac{1}{\theta}\ \int_0^\theta g(x)\ \ \mathrm{d}x\right) \\
&= \frac{g(\theta)}{\theta} + \int_0^\theta g(x)\left(-\frac{1}{\theta^2}\right)\ \ \mathrm{d}x,
\end{aligned}$$

where we have used the Product Rule and the Fundamental Theorem of Calculus. On the other hand

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\partial}{\partial\theta}\left(g(x)f(x \mid \theta)\right)\ \mathrm{d}x &= \frac{d}{d\theta}\int_0^\theta g(x)\ \frac{1}{\theta}\ \ \mathrm{d}x \\
&= \int_0^\theta g(x)\left(-\frac{1}{\theta^2}\right)\ \ \mathrm{d}x.
\end{aligned}$$

Thus, differentiation and integration can be interchanged if and only if

$$\frac{g(\theta)}{\theta} = 0 \quad \text{for all} \quad \theta.$$

□