8.4. Suppose that Y_1, Y_2, Y_3 denote a random sample from an exponential distribution with density function

$$f(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta}, & y > 0 \\ 0 & \text{else} \end{cases}$$

Consider the following five estimators of θ:

$\hat{\theta}_1 = Y_1$, $\hat{\theta}_2 = \frac{Y_1 + Y_2}{2}$ $\hat{\theta}_3 = \frac{Y_1 + 2Y_2}{3}$ $\hat{\theta}_4 = \min(Y_1, Y_2, Y_3)$

$\hat{\theta}_5 = \bar{Y}$.

a. Which of these estimators are unbiased?

$Y \sim Exp(\theta)$ $E[Y] = \theta$ $V(Y) = \theta^2$

$E[\hat{\theta}_1] = E[\hat{\theta}_2] = E[\hat{\theta}_3] = E[\hat{\theta}_5] = \theta$

$g_{(1)}(y) = n(1 - F(y))^{n-1} f(y) = 3(e^{-y/\theta})^2 \frac{e^{-y/\theta}}{\theta} = \frac{3e^{-3y/\theta}}{\theta}$

$F(y) = 1 - e^{-y/\theta}$

$Y_{(1)} \sim Exp(\frac{\theta}{3})$ $E[Y_{(1)}] = \frac{\theta}{3} = E[\hat{\theta}_4]$

$\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ and $\hat{\theta}_5$ are unbiased

b. Among the unbiased estimators, which has the smallest variance?

$V(\hat{\theta}_1) = \theta^2$ $V(\hat{\theta}_2) = \frac{\theta^2}{2}$ $V(\hat{\theta}_3) = \theta^2 \frac{(1+4)}{9} = \frac{5}{9}\theta^2$

$V(\hat{\theta}_5) = \frac{\theta^2}{3}$

Exercise

Let $Y_1, \ldots, Y_n$ be a random sample from a uniform distribution $(0, \theta)$

(a) Find a constant a such that $\hat{\theta}_1 = a\bar{Y}$ is an unbiased estimator of θ

(b) Find a constant b such that $\hat{\theta}_2 = b \max(Y_1 \ldots Y_n)$ is an unbiased estimator of θ

(c) Find a constant b such that $\hat{\theta}_3 = c \min(Y_1, \ldots Y_n)$ is an unbiased estimator of θ.

(d) Find the MSE of $\hat{\theta}_1$, $\hat{\theta}_2$ and $\hat{\theta}_3$. Which estimator has the smallest MSE?

(a) $f(y) = \begin{cases} \frac{1}{\theta} & \text{if } 0 < y < \theta \\ 0 & \text{else} \end{cases}$ $\quad E[Y] = \frac{\theta}{2}$. $\quad Var(Y) = \frac{\theta^2}{12}$ $\quad F(y) = \frac{y}{\theta}$

$E[a\bar{Y}] = aE[Y] = a\frac{\theta}{2} = \theta \qquad a = 2$

$Var(2\bar{Y}) = 4 Var(\bar{Y}) = \frac{4}{n} Var(Y) = \frac{4}{n}\frac{\theta^2}{12} = \frac{\theta^2}{3n}.$

(b) $f_{Y_{(n)}}(y) = n(F(y))^{n-1} f(y) = n\left(\frac{y}{\theta}\right)^{n-1}\frac{1}{\theta} = \frac{ny^{n-1}}{\theta^n}$ if $0 < y < \theta$

$\theta = E[bY_{(n)}] = b\int_0^\theta y \frac{ny^{n-1}}{\theta^n} dy = b\frac{n}{n+1}\theta \quad$ So $\quad b = \frac{n+1}{n}$

$E\left[\left(\frac{n+1}{n}Y_{(n)}\right)^2\right] = \frac{(n+1)^2}{n^2}\int_0^\theta y^2 \frac{ny^{n-1}}{\theta^n} dy = \frac{(n+1)^2}{n^2}\frac{n\theta^2}{n+2} = \frac{(n+1)^2}{n(n+2)}\theta^2$

$$MSE[\hat{\theta}_2] = Var\left(\frac{n+1}{n} Y_{(n)}\right) = \frac{(n+1)^2}{n(n+2)}\theta^2 - \theta^2 = \frac{\theta^2}{n(n+2)}$$

(c) $$f_{Y_{(1)}}(y) = n\left(1-F(y)\right)^{n-1} f(y) = n\left(1-\frac{y}{\theta}\right)^{n-1}\frac{1}{\theta} = \frac{n(\theta-y)^{n-1}}{\theta^n}, \text{ if } 0<y<\theta$$

$$E[Y_{(1)}] = \int_0^\theta y\,\frac{n(\theta-y)^{n-1}}{\theta^n}\,dy = \int_0^\theta (\theta-x)\,\frac{n x^{n-1}}{\theta^n}\,dx = \int_0^\theta \left(\frac{n x^{n-1}}{\theta^{n-1}} - \frac{n x^n}{\theta^n}\right)dx$$

$\theta - y = x$

$$= \theta - \frac{n}{n+1}\theta = \frac{\theta}{n+1}$$

$$\theta = E[\hat{\theta}] = E[c\,Y_{(1)}] = \frac{c\theta}{n+1} \qquad c = n+1$$

$$E[Y_{(1)}^2] = \int_0^\theta y^2\,\frac{n(\theta-y)^{n-1}}{\theta^n}\,dy = \int_0^\theta (\theta-x)^2\,\frac{n x^{n-1}}{\theta^n}\,dx = \int_0^\theta \left(\frac{n x^{n-1}}{\theta^{n-2}} - \frac{2n x^n}{\theta^{n-1}} + \frac{n x^{n+1}}{\theta^n}\right)dx$$

$$= \theta - \frac{2n\theta}{n+1} + \frac{n}{n+2}\theta = \frac{\left((n+1)(n+2) - 2n(n+2) + n(n+1)\right)\theta^2}{(n+1)(n+2)}$$

$$= \frac{(n^2+3n+2-2n-4n+n^2+n)\theta^2}{(n+1)(n+2)} = \frac{(2n^2-2n+2)\theta^2}{(n+1)(n+2)}$$

$$Var(Y_{(n)}) = \frac{2(n^2-n+1)\theta^2}{(n+1)(n+2)} - \frac{\theta^2}{(n+1)^2} = \frac{\theta^2\left((2n^2-2n+2)(n+1) - (n+2)\right)}{(n+1)^2(n+2)} = \frac{\theta^2(2n^3-n)}{(n+1)^2(n+1)}$$

$$MSE(\hat{\theta}) = Var\left((n+1)Y_{(1)}\right) = \frac{\theta^2 n(2n^2-1)}{n+1}$$

How to find the mean and the MSE of an estimator

Common estimators are, $\frac{1}{n}\sum_{j=1}^{n} Y_j$, $\frac{1}{n-1}\sum_{j=1}^{n}(Y_j-\bar{Y})^2$, $\frac{1}{n}\sum_{j=1}^{n} Y_j^2$

, $\frac{1}{n}\sum_{j=1}^{n} Y_j^K$, $\min(Y_1 \ldots Y_n)$, $\max(Y_1 \ldots Y_n)$

We have that

$$E\left[\frac{1}{n}\sum_{j=1}^{n} Y_j^K\right] = E[Y^K]$$

$$Var\left(\frac{1}{n}\sum_{j=1}^{n} Y_j^K\right) = Var(Y^K) = E[Y^{2K}] - \left(E[Y^K]\right)^2$$

So, the MSE of $\frac{1}{n}\sum_{j=1}^{K} Y_j^K$ as an estimator of θ is

$$MSE(\theta) = E_\theta[Y^{2K}] - \left(E_\theta[Y^K]\right)^2 + \left(E_\theta[Y^K] - \theta\right)^2$$

To find the MSE of $\min(Y_1 \ldots Y_n)$ and $\max(Y_1 \ldots Y_n)$ we use their pdf's.

Order statistics

Let $Y_1 \ldots Y_n$ be iid r.v from the df F (and pdf f)

We denote the ordered r.v's $Y_1 \ldots Y_n$, by $Y_{(1)}, \ldots, Y_{(n)}$

where $Y_{(1)} \le Y_{(2)} \le \ldots \le Y_{(n)}$.

$Y_{(1)}, \ldots, Y_{(n)}$ are called the order statistics

$Y_{(1)} = \min(Y_1, Y_2, \ldots, Y_n)$ is the first order statistic

$Y_{(n)} = \max(Y_1, Y_2, \ldots, Y_n)$ is the n-th order statistic.

Distribution of $Y_{(1)}$

$$F_{Y_{(1)}}(y) = P(Y_{(1)} \le y) = P(\min_{1\le i\le n} Y_i \le y) = 1 - P(\min_{1\le i\le n} Y_i > y)$$

$$= 1 - P\left(\bigcap_{i=1}^{n} (Y_i > y)\right) = 1 - \prod_{i=1}^{n} P(Y_i > y) = 1 - (1 - F_Y(y))^n$$

If Y has a continuous distribution with pdf $f(y)$

$$f_{Y_{(1)}}(y) = n(1 - F_Y(y))^{n-1} f_Y(y).$$

Distribution of $Y_{(n)}$

$$F_{Y_{(n)}}(y) = P(Y_{(n)} \le y) = P(\max_{1\le i\le n} Y_i \le y) = P\left(\bigcap_{i=1}^{n} (Y_i \le y)\right)$$

$$= \prod_{i=1}^{n} P(Y_i \le y) = \prod_{i=1}^{n} F(y) = (F(y))^n$$

If Y has a continuous distribution

$$f_{Y_{(n)}}(y) = n(F_Y(y))^{n-1} f_Y(y)$$