QUESTION 1

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the Exp (λ) distribution. Consider the following estimators for $\theta = 1/\lambda$: $\widehat{\theta_1} = (1/n) \sum_{i=1}^n X_i$ and $\widehat{\theta_2} = (1/(n+1)) \sum_{i=1}^n X_i$.

(i) Find the biases of $\widehat{\theta_1}$ and $\widehat{\theta_2}$.

(ii) Find the variances of $\widehat{\theta_1}$ and $\widehat{\theta_2}$.

(iii) Find the mean squared errors of $\widehat{\theta_1}$ and $\widehat{\theta_2}$.

(iv) Which of the two estimators ($\widehat{\theta_1}$ or $\widehat{\theta_2}$) is better and why?

SOLUTIONS TO QUESTION 1

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the Exp (λ) distribution. Consider the following estimators for $\theta = 1/\lambda$: $\widehat{\theta_1} = (1/n) \sum_{i=1}^n X_i$ and $\widehat{\theta_2} = (1/(n+1)) \sum_{i=1}^n X_i$.

(i) The bias of $\widehat{\theta_1}$ is

$$\begin{aligned} E\left(\widehat{\theta_1}\right) - \theta &= E\left(\frac{1}{n}\sum_{i=1}^n X_i\right) - \theta \\ &= \frac{1}{n}\sum_{i=1}^n E\left(X_i\right) - \theta \\ &= \frac{1}{n}\sum_{i=1}^n \theta - \theta \\ &= \theta - \theta \\ &= 0. \end{aligned}$$

The bias of $\widehat{\theta_2}$ is

$$\begin{aligned} E\left(\widehat{\theta_2}\right) - \theta &= E\left(\frac{1}{n+1}\sum_{i=1}^n X_i\right) - \theta \\ &= \frac{1}{n+1}\sum_{i=1}^n E\left(X_i\right) - \theta \\ &= \frac{1}{n+1}\sum_{i=1}^n \theta - \theta \\ &= \frac{n\theta}{n+1} - \theta \\ &= -\frac{\theta}{n+1}. \end{aligned}$$

(ii) The variance of $\widehat{\theta_1}$ is

$$\begin{aligned} Var\left(\widehat{\theta_1}\right) &= Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n Var\left(X_i\right) \\ &= \frac{1}{n^2}\sum_{i=1}^n \theta^2 \\ &= \frac{\theta^2}{n}. \end{aligned}$$

The variance of $\widehat{\theta_2}$ is

$$Var\left(\widehat{\theta_2}\right) = Var\left(\frac{1}{n+1}\sum_{i=1}^n X_i\right)$$

$$
\begin{aligned}
&= \frac{1}{(n+1)^2}\sum_{i=1}^{n} Var\left(X_i\right) \\
&= \frac{1}{(n+1)^2}\sum_{i=1}^{n} \theta^2 \\
&= \frac{n\theta^2}{(n+1)^2}.
\end{aligned}
$$

(iii) The mean squared error of $\widehat{\theta_1}$ is

$$MSE\left(\widehat{\theta_1}\right) = \frac{\theta^2}{n}.$$

The mean squared error of $\widehat{\theta_2}$ is

$$MSE\left(\widehat{\theta_2}\right) = \frac{n\theta^2}{(n+1)^2} + \left(\frac{\theta}{n+1}\right)^2 = \frac{\theta^2}{n+1}.$$

(iv) In terms of bias, $\widehat{\theta_1}$ is unbiased and $\widehat{\theta_2}$ is biased (however, $\widehat{\theta_2}$ is asymptotically unbiased). So, one would prefer $\widehat{\theta_1}$ if bias is the important issue.

In terms of mean squared error, $\widehat{\theta_2}$ has better efficiency (however, both estimators are consistent). So, one would prefer $\widehat{\theta_2}$ if efficiency is the important issue.

QUESTION 2

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with the common probability mass function (pmf):

$$p(x) = \theta(1-\theta)^{x-1}$$

for $x = 1, 2, \ldots$ and $0 < \theta < 1$.

(i) Write down the likelihood function of θ.

(ii) Find the maximum likelihood estimator (mle) of θ.

(iii) Find the mle of $\psi = 1/\theta$.

(iv) Determine the bias, variance and the mean squared error of the mle of ψ.

(v) Is the mle of ψ unbiased? Is it consistent?

SOLUTIONS TO QUESTION 2

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with the common probability mass function (pmf):

$$p(x) = \theta(1-\theta)^{x-1}$$

for $x = 1, 2, \ldots$ and $0 < \theta < 1$. This pmf corresponds to the geometric distribution, so $E(X_i) = 1/\theta$ and $Var(X_i) = (1-\theta)/\theta^2$.

(i) The likelihood function of θ is

$$L(\theta) = \prod_{i=1}^{n} \left\{ \theta(1-\theta)^{X_i - 1} \right\} = \theta^n (1-\theta)^{\sum_{i=1}^n X_i - n}.$$

(ii) The log likelihood function of θ is

$$\log L(\theta) = n \log \theta + \left(\sum_{i=1}^{n} X_i - n \right) \log(1-\theta).$$

The first and second derivatives of this with respect to θ are

$$\frac{d \log L(\theta)}{d\theta} = \frac{n}{\theta} - \frac{\sum_{i=1}^n X_i - n}{1-\theta}$$

and

$$\frac{d^2 \log L(\theta)}{d\theta^2} = -\frac{n}{\theta^2} - \frac{\sum_{i=1}^n X_i - n}{(1-\theta)^2},$$

respectively. Note that $d \log L(\theta)/d\theta = 0$ if $\theta = n / \sum_{i=1}^n X_i$ and that $d^2 \log L(\theta)/d\theta^2 < 0$ for all $0 < \theta < 1$. So, it follows that $\widehat{\theta} = n / \sum_{i=1}^n X_i$ is the mle of θ.

(iii) By the invariance principle, the mle of $\psi = 1/\theta$ is $\widehat{\psi} = (1/n) \sum_{i=1}^n X_i$.

(iv) The bias of $\widehat{\psi}$ is

$$\begin{aligned}
E\left(\widehat{\psi}\right) - \psi &= E\left(\frac{1}{n} \sum_{i=1}^{n} X_i \right) - \psi \\
&= \frac{1}{n} \sum_{i=1}^{n} E\left(X_i\right) - \psi \\
&= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\theta} - \psi \\
&= \psi - \psi \\
&= 0.
\end{aligned}$$

The variance of $\widehat{\psi}$ is

$$
\begin{aligned}
Var\left(\widehat{\psi}\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} Var\left(X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \frac{1-\theta}{\theta^2} \\
&= \frac{1-\theta}{n\theta^2} \\
&= \frac{\psi^2-\psi}{n}.
\end{aligned}
$$

The mean squared error of $\widehat{\psi}$ is

$$
MSE\left(\widehat{\psi}\right) = \frac{\psi^2-\psi}{n}.
$$

(v) The mle of ψ is unbiased and consistent.

QUESTION 3

Let X and Y be uncorrelated random variables. Suppose that X has mean 2θ and variance 4. Suppose that Y has mean θ and variance 2. The parameter θ is unknown.

(i) Compute the bias and mean squared error for each of the following estimators of θ: $\widehat{\theta_1} = (1/4)X + (1/2)Y$ and $\widehat{\theta_2} = X - Y$.

(ii) Which of the two estimators ($\widehat{\theta_1}$ or $\widehat{\theta_2}$) is better and why?

(iii) Verify that the estimator $\widehat{\theta}_c = (c/2)X + (1-c)Y$ is unbiased. Find the value of c which minimizes Var $(\widehat{\theta}_c)$.

SOLUTIONS TO QUESTION 3

Let X and Y be uncorrelated random variables. Suppose that X has mean 2θ and variance 4. Suppose that Y has mean θ and variance 2. The parameter θ is unknown.

(i) The biases and mean squared errors of $\widehat{\theta}_1 = (1/4)X + (1/2)Y$ and $\widehat{\theta}_2 = X - Y$ are:

$$\begin{aligned} Bias\left(\widehat{\theta}_1\right) &= E\left(\widehat{\theta}_1\right) - \theta \\ &= E\left(\frac{X}{4} + \frac{Y}{2}\right) - \theta \\ &= \frac{E(X)}{4} + \frac{E(Y)}{2} - \theta \\ &= \frac{2\theta}{4} + \frac{\theta}{2} - \theta \\ &= 0, \end{aligned}$$

$$\begin{aligned} Bias\left(\widehat{\theta}_2\right) &= E\left(\widehat{\theta}_2\right) - \theta \\ &= E\left(X - Y\right) - \theta \\ &= E(X) - E(Y) - \theta \\ &= 2\theta - \theta - \theta \\ &= 0, \end{aligned}$$

$$\begin{aligned} MSE\left(\widehat{\theta}_1\right) &= Var\left(\widehat{\theta}_1\right) \\ &= Var\left(\frac{X}{4} + \frac{Y}{2}\right) \\ &= \frac{Var(X)}{16} + \frac{Var(Y)}{4} \\ &= \frac{4}{16} + \frac{2}{4} \\ &= \frac{3}{4}, \end{aligned}$$

and

$$\begin{aligned} MSE\left(\widehat{\theta}_2\right) &= Var\left(\widehat{\theta}_2\right) \\ &= Var\left(X - Y\right) \\ &= Var(X) + Var(Y) \\ &= 4 + 2 \\ &= 6. \end{aligned}$$

(ii) Both $\widehat{\theta}_1$ and $\widehat{\theta}_2$ are unbiased. The MSE of $\widehat{\theta}_1$ is smaller than the MSE of $\widehat{\theta}_2$. So, we prefer $\widehat{\theta}_1$.

(iii) The bias of $\widehat{\theta}_c$ is

$$\begin{aligned}
E\left(\widehat{\theta}_c\right) - \theta &= E\left(\frac{c}{2}X + (1-c)Y\right) - \theta \\
&= \frac{c}{2}E(X) + (1-c)E(Y) - \theta \\
&= \frac{c}{2}2\theta + (1-c)\theta - \theta \\
&= 0,
\end{aligned}$$

so $\widehat{\theta}_c$ is unbiased.

The variance of $\widehat{\theta}_c$ is

$$\begin{aligned}
Var\left(\widehat{\theta}_c\right) &= Var\left(\frac{c}{2}X + (1-c)Y\right) \\
&= \frac{c^2}{4}Var(X) + (1-c)^2 Var(Y) \\
&= \frac{c^2}{4}4 + 2(1-c)^2 \\
&= c^2 + 2(1-c)^2.
\end{aligned}$$

Let $g(c) = c^2 + 2(1-c)^2$. Them $g^{'}(c) = 6c - 4 = 0$ if $c = 2/3$. Also $g^{''}(c) = 6 > 0$. So, $c = 2/3$ minimizes the variance of $\widehat{\theta}_c$.

QUESTION 4

Consider the linear regression model with zero intercept:

$$Y_i = \beta X_i + e_i$$

for $i = 1, 2, \ldots, n$, where $e_1, e_2, \ldots, e_n$ are independent and identical normal random variables with zero mean and variance σ^2 assumed known. Moreover, suppose $X_1, X_2, \ldots, X_n$ are known constants.

(i) Write down the likelihood function of β.

(ii) Derive the maximum likelihood estimator of β.

(iii) Find the bias of the estimator in part (ii). Is the estimator unbiased?

(iv) Find the mean square error of the estimator in part (ii).

(v) Find the exact distribution of the estimator in part (ii).

SOLUTIONS TO QUESTION 4

Consider the linear regression model with zero intercept:

$$Y_i = \beta X_i + e_i$$

for $i = 1, 2, \ldots, n$, where $e_1, e_2, \ldots, e_n$ are independent and identical normal random variables with zero mean and variance σ^2 assumed known. Moreover, suppose $X_1, X_2, \ldots, X_n$ are known constants.

(i) The likelihood function of β is

$$L(\beta) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left\{-\frac{e_i^2}{2\sigma^2}\right\} = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left\{-\frac{(Y_i - \beta X_i)^2}{2\sigma^2}\right\}.$$

(ii) The log likelihood function of β is

$$\log L(\beta) = -\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta X_i)^2 .$$

The normal equation is:

$$\frac{\partial \log L(\beta)}{\partial \beta} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta X_i)\, X_i = 0.$$

Solving this equation gives

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

This is an mle since

$$\frac{\partial^2 \log L(\beta)}{\partial \beta^2} = -\frac{1}{\sigma^2}\sum_{i=1}^{n} X_i^2 < 0.$$

(iii) The bias of $\widehat{\beta}$ is

$$\begin{aligned} E\widehat{\beta} - \beta &= E\frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} - \beta \\ &= \frac{\sum_{i=1}^{n} X_i E(Y_i)}{\sum_{i=1}^{n} X_i^2} - \beta \\ &= \frac{\sum_{i=1}^{n} X_i \beta X_i}{\sum_{i=1}^{n} X_i^2} - \beta \\ &= 0, \end{aligned}$$

so $\widehat{\beta}$ is indeed unbiased.

(iv) The variance of $\widehat{\beta}$ is

$$\begin{aligned}
Var\widehat{\beta} &= Var\frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} \\
&= \frac{\sum_{i=1}^{n} X_i^2 Var\left(Y_i\right)}{\left(\sum_{i=1}^{n} X_i^2\right)^2} \\
&= \frac{\sum_{i=1}^{n} X_i^2 \sigma^2}{\left(\sum_{i=1}^{n} X_i^2\right)^2} \\
&= \frac{\sigma^2}{\sum_{i=1}^{n} X_i^2}.
\end{aligned}$$

(v) The estimator is a linear combination of independent normal random variable, so it is also normal with mean β and variance $\frac{\sigma^2}{\sum_{i=1}^{n} X_i^2}$.

QUESTION 5

Suppose X_1 and X_2 are independent Uniform$[-\theta, \theta]$ random variables. Let $\widehat{\theta_1} = 3 \min\left(|X_1|, |X_2|\right)$ and $\widehat{\theta_2} = 3 \max\left(X_1, X_2\right)$ denote possible estimators of θ.

(i) Derive the bias and mean squared error of $\widehat{\theta_1}$;

(ii) Derive the bias and mean squared error of $\widehat{\theta_2}$;

(iii) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to bias and why?

(iv) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to mean squared error and why?

SOLUTIONS TO QUESTION 5

(i) Let $Z = \min(|X_1|, |X_2|)$. Then

$$\begin{aligned} F_Z(z) &= \Pr[\min(|X_1|, |X_2|) < z] \\ &= 1 - \Pr[\min(|X_1|, |X_2|) > z] \\ &= 1 - \Pr[|X_1| > z, |X_2| > z] \\ &= 1 - [\Pr(|X| > z)]^2 \\ &= 1 - [1 - \Pr(|X| < z)]^2 \\ &= 1 - \left[1 - \frac{z}{\theta}\right]^2 \\ &= 1 - \frac{(\theta - z)^2}{\theta^2} \end{aligned}$$

and

$$f_Z(z) = \frac{2(\theta - z)}{\theta^2}$$

and

$$E(Z) = \int_0^\theta \frac{2z(\theta - z)}{\theta^2} = \frac{2}{\theta^2}\left[\frac{\theta z^2}{2} - \frac{z^3}{3}\right]_0^\theta = \frac{\theta}{3}$$

and

$$E\left(Z^2\right) = \int_0^\theta \frac{2z^2(\theta - z)}{\theta^2} dz = \frac{2}{\theta^2}\left[\frac{\theta z^3}{3} - \frac{z^4}{4}\right]_0^\theta = \frac{\theta^2}{6}.$$

So, $Bias\left(\widehat{\theta_1}\right) = 0$ and $MSE\left(\widehat{\theta_1}\right) = \theta^2/2$.

(ii) Let $Z = \max(X_1, X_2)$. Then

$$\begin{aligned} F_Z(z) &= \Pr[\max(X_1, X_2) < z] \\ &= \Pr[X_1 < z, X_2 < z] \\ &= [\Pr(X < z)]^2 \\ &= \left[\frac{z + \theta}{2\theta}\right]^2 \end{aligned}$$

and

$$f_Z(z) = \frac{z + \theta}{2\theta^2}$$

and

$$E(Z) = \int_{-\theta}^\theta z\frac{z + \theta}{2\theta^2} dz = \left[\frac{z^3}{6\theta^2} + \frac{z^2}{4\theta}\right]_{-\theta}^\theta = \frac{\theta}{3}$$

and

$$E\left(Z^2\right) = \int_{-\theta}^{\theta} z^2 \frac{z+\theta}{2\theta^2} dz = \left[\frac{z^4}{8\theta^2} + \frac{z^3}{6\theta}\right]_{-\theta}^{\theta} = \frac{\theta^2}{3}.$$

So, $Bias\left(\widehat{\theta_2}\right) = 0$ and $MSE\left(\widehat{\theta_2}\right) = 2\theta^2$.

(iii) Both estimators are equally good with respect to bias.

(iv) $\widehat{\theta_1}$ has smaller MSE.

QUESTION 6

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function $f(x) = a^{-1}x^{a^{-1}-1}$, $0 < x < 1$, where $a > 0$ is an unknown parameter.

(a) Write down the likelihood function of a.

(b) Show that the maximum likelihood estimator of a is $\widehat{a} = -\dfrac{1}{n}\sum_{i=1}^{n} \log X_i$.

(c) Derive the expected value of $\widehat{a}$ in part (b). You may use the fact that $\int_0^1 x^\alpha \log x dx = -(\alpha+1)^{-2}$ without proof.

(d) Derive the variance of $\widehat{a}$ in part (b). You may use the fact that $\int_0^1 x^\alpha (\log x)^2 dx = 2(\alpha+1)^{-3}$ without proof.

(e) Show that $\widehat{a}$ is an unbiased and a consistent estimator for a.

(f) Find the maximum likelihood estimator of $\Pr(X < 0.5)$, where X has the probability density function $f(x) = a^{-1}x^{a^{-1}-1}$, $0 < x < 1$. Justify your answer.

SOLUTIONS TO QUESTION 6

(a) The likelihood function is

$$L(a) = a^{-n} \left(\prod_{i=1}^{n} x_i \right)^{1/a-1}.$$

(b) The log likelihood function is

$$\log L(a) = -n \log a + \left(a^{-1} - 1 \right) \sum_{i=1}^{n} \log x_i.$$

The derivative of $\log L$ with respect to a is

$$\frac{d \log L(a)}{da} = -\frac{n}{a} - a^{-2} \sum_{i=1}^{n} \log x_i.$$

Setting this to zero and solving for a, we obtain

$$\widehat{a} = -\frac{1}{n} \sum_{i=1}^{n} \log x_i.$$

This is indeed an MLE since

$$\begin{aligned} \frac{d^2 \log L(a)}{da^2} &= \frac{n}{a^2} + 2a^{-3} \sum_{i=1}^{n} \log x_i \\ &= \frac{n}{\widehat{a}^2} - 2n\widehat{a}^{-2} \\ &< 0 \end{aligned}$$

at $a = \widehat{a}$.

(c) The expected value is

$$\begin{aligned} E(\widehat{a}) &= -\frac{1}{n} \sum_{i=1}^{n} E(\log x_i) \\ &= -\frac{1}{na} \sum_{i=1}^{n} \int_0^1 \log x x^{1/a-1} dx \\ &= \frac{1}{na} \sum_{i=1}^{n} a^2 \\ &= a. \end{aligned}$$

(d) The variance is

$$Var(\widehat{a}) = \frac{1}{n^2} \sum_{i=1}^{n} Var(\log x_i)$$

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{E\left[(\log x_i)^2\right]-a^2\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}E\left[(\log x_i)^2\right]-\frac{a^2}{n} \\
&= \frac{1}{n^2 a}\sum_{i=1}^{n}\int_0^1 (\log x)^2 x^{1/a-1}dx-\frac{a^2}{n} \\
&= \frac{1}{n^2 a}\sum_{i=1}^{n}2a^3-\frac{a^2}{n} \\
&= \frac{a^2}{n}
\end{aligned}
$$

(e) Bias $(\widehat{a}) = 0$ and MSE $(\widehat{a}) \to 0$, so the estimator is unbiased and consistent.

(f) Note that

$$
\Pr(X < 0.5) = \int_0^{0.5} a^{-1}x^{a^{-1}-1}dx = \left[x^{a^{-1}}\right]_0^{0.5} = 0.5^{1/a},
$$

which is a one-to-one function of a for $a > 0$. By the invariance principle, its maximum likelihood estimator is $0.5^{n/\left(\sum_{i=1}^{n}\log x_i\right)}$.

QUESTION 7

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function $f(x) = \frac{1}{2a} \exp\left(-\frac{|x|}{a}\right)$, $-\infty < x < \infty$, where $a > 0$ is an unknown parameter.

(a) Write down the likelihood function of a.

(b) Show that the maximum likelihood estimator of a is $\widehat{a} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} |X_i|$.

(c) Derive the expected value of $\widehat{a}$ in part (b).

(d) Derive the variance of $\widehat{a}$ in part (b).

(e) Show that $\widehat{a}$ is an unbiased and consistent estimator for a.

SOLUTIONS TO QUESTION 7

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function $f(x) = \frac{1}{2a} \exp\left(-\frac{|x|}{a}\right)$, $-\infty < x < \infty$, where $a > 0$ is an unknown parameter.

(a) The likelihood function of a is

$$L(a) = \prod_{i=1}^{n} \left[\frac{1}{2a} \exp\left(-\frac{\mid X_i \mid}{a} \right) \right] = \frac{1}{(2a)^n} \exp\left(-\frac{1}{a} \sum_{i=1}^{n} \mid X_i \mid \right).$$

(b) The log likelihood function of a is

$$\log L(a) = -n \log(2a) - \frac{1}{a} \sum_{i=1}^{n} \mid X_i \mid .$$

The derivative of $\log L$ with respect to a is

$$\frac{d \log L(a)}{da} = -\frac{n}{a} + \frac{1}{a^2} \sum_{i=1}^{n} \mid X_i \mid .$$

Setting this to zero and solving for a, we obtain

$$\widehat{a} = \frac{1}{n} \sum_{i=1}^{n} \mid X_i \mid .$$

This is a maximum likelihood estimator of a since

$$\begin{aligned}
\left. \frac{d^2 \log L(a)}{da^2} \right|_{a=\widehat{a}} &= \frac{n}{\widehat{a}^2} - \frac{2}{\widehat{a}^3} \sum_{i=1}^{n} \mid X_i \mid \\
&= \frac{n}{\widehat{a}^2} \left[1 - \frac{2}{n\widehat{a}} \sum_{i=1}^{n} \mid X_i \mid \right] \\
&= \frac{n}{\widehat{a}^2} [1 - 2] \\
&< 0.
\end{aligned}$$

(c) The expected value of $\widehat{a}$ is

$$\begin{aligned}
E(\widehat{a}) &= \frac{1}{n} \sum_{i=1}^{n} E[\mid X_i \mid] \\
&= \frac{1}{2na} \sum_{i=1}^{n} \int_{-\infty}^{\infty} \mid x \mid \exp\left(-\frac{\mid x \mid}{a} \right) dx \\
&= \frac{1}{na} \sum_{i=1}^{n} \int_{0}^{\infty} x \exp\left(-\frac{x}{a} \right) dx \\
&= \frac{a}{n} \sum_{i=1}^{n} \int_{0}^{\infty} y \exp(-y)\, dy
\end{aligned}$$

$$
\begin{aligned}
&= \frac{a}{n}\sum_{i=1}^{n}\Gamma(2) \\
&= \frac{a}{n}\sum_{i=1}^{n}1 \\
&= a.
\end{aligned}
$$

(d) The variance of $\widehat{a}$ is

$$
\begin{aligned}
\mathrm{Var}\left(\widehat{a}\right) &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\left[\mid X_i \mid\right] \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{E\left[X_i^2\right] - E^2\left[\mid X_i \mid\right]\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{\frac{1}{2a}\int_{-\infty}^{\infty} x^2 \exp\left(-\frac{\mid x \mid}{a}\right) dx - \frac{1}{4a^2}\left[\int_{-\infty}^{\infty} \mid x \mid \exp\left(-\frac{\mid x \mid}{a}\right) dx\right]^2\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{\frac{1}{a}\int_{0}^{\infty} x^2 \exp\left(-\frac{x}{a}\right) dx - \frac{1}{a^2}\left[\int_{0}^{\infty} x \exp\left(-\frac{x}{a}\right) dx\right]^2\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{a^2\int_{0}^{\infty} y^2 \exp\left(-y\right) dy - a^2\left[\int_{0}^{\infty} y \exp\left(-y\right) dy\right]^2\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\left\{a^2\Gamma(3) - a^2\left[\Gamma(2)\right]^2\right\} \\
&= \frac{1}{n^2}\sum_{i=1}^{n}a^2 \\
&= \frac{a^2}{n}.
\end{aligned}
$$

(e) The bias $\left(\widehat{a}\right) = 0$ and MSE $\left(\widehat{a}\right) = \frac{a^2}{n}$, so $\widehat{a}$ is an unbiased and a consistent estimator for a.

QUESTION 8

Suppose X_1 and X_2 are independent Exp$(1/\theta)$ and Uniform $[0, \theta]$ random variables. Let $\widehat{\theta} = aX_1 + bX_2$ denote a class of estimators of θ, where a and b are constants.

(i) Determine the bias of $\widehat{\theta}$;

(ii) Determine the variance of $\widehat{\theta}$;

(iii) Determine the mean squared error of $\widehat{\theta}$;

(iv) Determine the condition involving a and b such that $\widehat{\theta}$ is unbiased for θ;

(v) Determine the value of a such that $\widehat{\theta}$ is unbiased for θ and has the smallest variance.

SOLUTIONS TO QUESTION 8

Suppose X_1 and X_2 are independent Exp$(1/\theta)$ and Uniform $[0, \theta]$ random variables. Let $\widehat{\theta} = aX_1 + bX_2$ denote a class of estimators of θ, where a and b are constants.

(i) The bias of $\widehat{\theta}$ is

$$\begin{aligned}
\text{Bias}\left(\widehat{\theta}\right) &= E\left(\widehat{\theta}\right) - \theta \\
&= aE\left(X_1\right) + bE\left(X_2\right) - \theta \\
&= \frac{a}{\theta}\int_0^{+\infty} x \exp\left(-\frac{x}{\theta}\right) dx + b\int_0^{\theta} \frac{x}{\theta} dx - \theta \\
&= a\theta \int_0^{+\infty} y \exp\left(-y\right) dx + \frac{b}{\theta}\left[\frac{x^2}{2}\right]_0^{\theta} - \theta \\
&= a\theta \int_0^{+\infty} y \exp\left(-y\right) dx + \frac{b}{\theta}\left[\frac{\theta^2}{2} - 0\right] - \theta \\
&= \left(a + \frac{b}{2} - 1\right)\theta.
\end{aligned}$$

(ii) The variance of $\widehat{\theta}$ is

$$\begin{aligned}
\text{Var}\left(\widehat{\theta}\right) &= a^2\text{Var}\left(X_1\right) + b^2\text{Var}\left(X_2\right) \\
&= a^2\left[\frac{1}{\theta}\int_0^{+\infty} x^2 \exp\left(-\frac{x}{\theta}\right) dx - \theta^2\right] + b^2\left[\frac{1}{\theta}\int_0^{\theta} x^2 dx - \frac{\theta^2}{4}\right] \\
&= a^2\left[\theta^2\int_0^{+\infty} y^2 \exp\left(-y\right) dy - \theta^2\right] + b^2\left\{\frac{1}{\theta}\left[\frac{x^3}{3}\right]_0^{\theta} - \frac{\theta^2}{4}\right\} \\
&= a^2\left[2\theta^2 - \theta^2\right] + b^2\left\{\frac{\theta^2}{3} - \frac{\theta^2}{4}\right\} \\
&= \left(a^2 + \frac{b^2}{12}\right)\theta^2.
\end{aligned}$$

(iii) The mean squared error of $\widehat{\theta}$ is

$$\begin{aligned}
\text{MSE}\left(\widehat{\theta}\right) &= \left(a^2 + \frac{b^2}{12}\right)\theta^2 + \left(a + \frac{b}{2} - 1\right)^2\theta^2 \\
&= \left[a^2 + \frac{b^2}{12} + \left(a + \frac{b}{2} - 1\right)^2\right]\theta^2 \\
&= \left(2a^2 + \frac{b^2}{3} + ab - 2a - b + 1\right)\theta^2.
\end{aligned}$$

(iv) $\widehat{\theta}$ is unbiased if $a + \frac{b}{2} = 1$. In other words, $b = 2(1 - a)$.

(v) If $\widehat{\theta}$ is unbiased then its variance is

$$\left[a^2 + \frac{(1-a)^2}{3}\right] \theta^2.$$

We need to minimize this as a function of a. Let $g(a) = a^2 + \frac{(1-a)^2}{3}$. The first order derivative is $g^{'}(a) = 2a - \frac{2(1-a)}{3}$. Setting the derivative to zero, we obtain $a = \frac{1}{4}$. The second order derivative is $g^{''}(a) = 2 + \frac{2}{3} > 0$. So, $g(a) = a^2 + \frac{(1-a)^2}{3}$ attains its minimum at $a = \frac{1}{4}$. Hence, the estimator with minimum variance is $\frac{1}{4}X_1 + \frac{3}{2}X_2$.

QUESTION 9

Suppose $X_1, \ldots, X_n$ are independent Uniform$[0, \theta]$ random variables. Let $\widehat{\theta_1} = \frac{2(X_1 + \cdots + X_n)}{n}$ and $\widehat{\theta_2} = \max(X_1, \ldots, X_n)$ denote possible estimators of θ.

(i) Derive the bias and mean squared error of $\widehat{\theta_1}$;

(ii) Derive the bias and mean squared error of $\widehat{\theta_2}$;

(iii) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to bias and why?

(iv) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to mean squared error and why?

SOLUTIONS TO QUESTION 9

Suppose $X_1, \ldots, X_n$ are independent Uniform$[0, \theta]$ random variables. Let $\widehat{\theta_1} = \frac{2(X_1+\cdots+X_n)}{n}$ and $\widehat{\theta_2} = \max(X_1, \ldots, X_n)$ denote possible estimators of θ.

(i) The bias and mean squared error of $\widehat{\theta_1}$ are

$$\begin{aligned}
\text{bias}\left(\widehat{\theta_1}\right) &= E\left(\widehat{\theta_1}\right) - \theta \\
&= \frac{2}{n} E\left(X_1 + \cdots + X_n\right) - \theta \\
&= \frac{2}{n}\left[E\left(X_1\right) + \cdots + E\left(X_n\right)\right] - \theta \\
&= \frac{2}{n}\left[\frac{\theta}{2} + \cdots + \frac{\theta}{2}\right] - \theta \\
&= \theta - \theta \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\text{MSE}\left(\widehat{\theta_1}\right) &= \text{Var}\left(\widehat{\theta_1}\right) \\
&= \frac{4}{n^2}\text{Var}\left(X_1 + \cdots + X_n\right) \\
&= \frac{4}{n^2}\left[\text{Var}\left(X_1\right) + \cdots + \text{Var}\left(X_n\right)\right] \\
&= \frac{4}{n^2}\left[\frac{\theta^2}{12} + \cdots + \frac{\theta^2}{12}\right] \\
&= \frac{\theta^2}{3n}.
\end{aligned}$$

(ii) Let $Z = \widehat{\theta_2}$. The cdf and the pdf of Z are

$$\begin{aligned}
F_Z(z) &= \Pr\left(\max\left(X_1, \ldots, X_n\right) \leq z\right) \\
&= \Pr\left(X_1 \leq z, \ldots, X_n \leq z\right) \\
&= \Pr\left(X_1 \leq z\right) \cdots \Pr\left(X_n \leq z\right) \\
&= \frac{z}{\theta} \cdots \frac{z}{\theta} \\
&= \frac{z^n}{\theta^n}
\end{aligned}$$

and

$$f_Z(z) = \frac{n z^{n-1}}{\theta^n}.$$

So, the bias and mean squared error of $\widehat{\theta_2}$ are

$$\begin{aligned}
\text{bias}\left(\widehat{\theta_2}\right) &= E(Z) - \theta \\
&= \frac{n}{\theta^n}\int_0^\theta z^n dz - \theta \\
&= \frac{n}{\theta^n}\left[\frac{z^{n+1}}{n+1}\right]_0^\theta - \theta \\
&= \frac{n}{\theta^n}\left[\frac{\theta^{n+1}}{n+1} - 0\right] - \theta \\
&= \frac{n\theta}{n+1} - \theta \\
&= -\frac{\theta}{n+1}
\end{aligned}$$

and

$$\begin{aligned}
\text{MSE}\left(\widehat{\theta_1}\right) &= \text{Var}(Z) + \left(-\frac{\theta}{n+1}\right)^2 \\
&= E\left(Z^2\right) - E^2(Z) + \frac{\theta^2}{(n+1)^2} \\
&= E\left(Z^2\right) - \frac{n^2\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} \\
&= \frac{n}{\theta^n}\int_0^\theta z^{n+1} dz - \frac{n^2\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} \\
&= \frac{n}{\theta^n}\left[\frac{z^{n+2}}{n+2}\right]_0^\theta - \frac{n^2\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} \\
&= \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} \\
&= \frac{2\theta^2}{(n+1)(n+2)}.
\end{aligned}$$

(iii) $\widehat{\theta_1}$ is better with respect to bias since bias $\widehat{\theta_1} = 0$ and bias $\widehat{\theta_2} \neq 0$.

(iv) $\widehat{\theta_2}$ is better with respect to mean squared error since

$$\begin{aligned}
& \frac{2\theta^2}{(n+1)(n+2)} \leq \frac{\theta^2}{3n} \\
\Leftrightarrow\ & 6n \leq (n+1)(n+2) \\
\Leftrightarrow\ & 6n \leq n^2 + 3n + 2 \\
\Leftrightarrow\ & 0 \leq n^2 - 3n + 2 \\
\Leftrightarrow\ & 0 \leq (n-1)(n-2).
\end{aligned}$$

Both $\widehat{\theta_1}$ and $\widehat{\theta_2}$ have equal mean squared errors when $n = 1, 2$.

QUESTION 10

Let $X_1, X_2, \dots, X_n$ be a random sample from the uniform $[0, \theta]$ distribution, where θ is unknown.

(i) Find the expected value and variance of the estimator $\widehat{\theta} = 2\overline{X}$, where $\overline{X} = (1/n)\sum_{i=1}^{n} X_i$.

(ii) Find the expected value of the estimator $\max(X_1, X_2, \dots, X_n)$, i.e. the largest observation.

(iii) Find the constant c such that $\widetilde{\theta} = c\max(X_1, X_2, \dots, X_n)$ is an unbiased estimator of θ. Also find the variance of $\widetilde{\theta}$.

(iv) Compare the mean square errors of $\widehat{\theta}$ and $\widetilde{\theta}$ and comment.

SOLUTIONS TO QUESTION 10

Let $X_1, X_2, \ldots, X_n$ be a random sample from the uniform $[0, \theta]$ distribution, where θ is unknown.

(i) The expected value of $\widehat{\theta}$ is:

$$\begin{aligned} E\left(\widehat{\theta}\right) &= 2E\left(\overline{X}\right) \\ &= (2/n)E\left(X_1 + X_2 + \cdots + X_n\right) \\ &= (2/n)n(\theta/2) \\ &= \theta. \end{aligned}$$

The variance of $\widehat{\theta}$ is:

$$\begin{aligned} Var\left(\widehat{\theta}\right) &= 4Var\left(\overline{X}\right) \\ &= \left(4/n^2\right) Var\left(X_1 + X_2 + \cdots + X_n\right) \\ &= \left(4/n^2\right) n \left(\theta^2/12\right) \\ &= \theta^2/(3n). \end{aligned}$$

(ii) Let $M = \max(X_1, X_2, \ldots, X_n)$. The cdf of M is

$$\begin{aligned} F_M(m) &= \Pr\left(X_1 \leq m, X_2 \leq m, \ldots, X_n \leq m\right) \\ &= \Pr\left(X_1 \leq m\right) \Pr\left(X_2 \leq m\right) \cdots \Pr\left(X_n \leq m\right) \\ &= \left(\Pr\left(X_1 \leq m\right)\right)^m \\ &= (m/\theta)^n \end{aligned}$$

and so its pdf is nm^{n-1}/θ^n. The expected value of M is

$$E(M) = n \int_0^\theta m^n dm/\theta^n = n\theta/(n+1).$$

(iii) Take $c = (n+1)/n$. Then

$$E(cM) = cE(M) = ((n+1)/n)n\theta/(n+1) = \theta.$$

So, $\widetilde{\theta} = (n+1)M/n$ is an unbiased estimator of θ.

The variance of M is

$$E\left(M^2\right) - n^2\theta^2/(n+1)^2 = n \int_0^\theta m^{n+1} dm/\theta^n - n^2\theta^2/(n+1)^2 = n\theta^2/\left((n+2)(n+1)^2\right).$$

So, the mean squared error of $\widetilde{\theta}$ is $(n+1)^2 Var(M)/n^2 = \theta^2/(n(n+2))$.

(iv) The mean square error of $\widehat{\theta}$ is $\theta^2/(3n)$.

The mean squared error of $\widetilde{\theta}$ is $\theta^2/(n(n+2))$.

The estimator $\widetilde{\theta}$ has the smaller mean squared error. So, it should be preferred.