

Problem Set II - Solutions

Content: Order Statistics and Estimation

Question 1

Suppose that the length of time Y it takes a worker to complete a certain task has the probability density function given by

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

where θ is a positive constant that represents the minimum time until task completion. Let $Y_1, Y_2, \dots, Y_n$ denote a random sample of completion times from this distribution. Find

- The density function for $Y_{(1)} = \min(Y_1, Y_2, \dots, Y_n)$
- $E(Y_{(1)})$

Solution: The distribution function of Y is

$$F(y) = P(Y \leq y) = \int_{\theta}^y e^{-(t-\theta)} dt = 1 - e^{-(y-\theta)}, y > \theta.$$

- $g_{(1)}(y) = n [e^{-(y-\theta)}]^{n-1} e^{-(y-\theta)} = n e^{-n(y-\theta)}, y > \theta.$
- $E(Y_{(1)}) = \frac{1}{n} + \theta.$

Question 2

Let $Y_1, Y_2, \dots, Y_n$ be independent, uniformly distributed random variables on the interval $[0, \theta]$. Find the

- Probability distribution function of $Y_{(n)} = \max(Y_1, Y_2, \dots, Y_n)$
- Density function of Y_n
- Mean and variance of Y_n

Solution:

- $G_{(n)}(y) = (y/\theta)^n, 0 \leq y \leq \theta.$
- $g_{(n)}(y) = G'_{(n)}(y) = ny^{n-1}/\theta^n, 0 \leq y \leq \theta.$
- It is easily shown that $E(Y_{(n)}) = \frac{n}{n+1}\theta, V(Y_{(n)}) = \frac{n}{(n+1)^2(n+2)}\theta^2.$

Question 3

Let $Y_1, Y_2, \dots, Y_n$ denote a random sample of size n from a population whose density is given by. Find

$$f(y) = \begin{cases} \alpha y^{\alpha-1}/\theta^\alpha, & 0 \leq y \leq \theta \\ 0, & \text{elsewhere} \end{cases}$$

where $\alpha > 0$ is a known, fixed value but θ is unknown. Consider the estimator $\hat{\theta} = \max(Y_1, Y_2, \dots, Y_n).$

- Show that $\hat{\theta}$ is a biased estimator for $\theta.$
- Find a multiple of $\hat{\theta}$ that is an unbiased estimator of $\theta.$
- Derive $MSE(\hat{\theta}).$

Solution:

- $E(\hat{\theta}) = E(Y_{(n)}) = E(Y) = (\frac{n\alpha}{n\alpha+1})\theta$, so that the estimator is biased.
- The unbiased estimator is $(\frac{n\alpha+1}{n\alpha})\hat{\theta} = (\frac{n\alpha+1}{n\alpha})Y_{(n)}$
- $MSE(Y_{(n)}) = E((Y_{(n)} - \theta))^2 = E(Y_{(n)}^2) - 2\theta E(Y_{(n)}) + \theta^2 = \frac{2}{(n\alpha+1)(n\alpha+2)}\theta^2.$

Question 4

Let $X_1, X_2, \dots, X_n$ denote a random sample from a Uniform $(0, \theta)$ distribution with $\theta > 0$ as the unknown parameter. Let $\bar{X}$ denote the sample mean.

- Is $\bar{X}$ an unbiased estimator for θ ?
- Find an unbiased estimator for $\theta.$
- Find the variance of the estimator in the part b.

Solution:

- We know that the sample mean of the uniform distribution can be founded with this formula: $\bar{X} = (a + b)/2$. Here, if we can apply this formula to our interval $(0, \theta)$, we can conclude that $\bar{X} = \theta/2$.
 $E(\bar{X}) = \theta/2$ and this is not equal to the θ . Thus, we can conclude that $\bar{X}$ is not an unbiased estimator.
It is a biased estimator.

- b. If we think an estimator of $2\bar{X}$, we can see that it is an unbiased estimator for θ . $E(2\bar{X}) = \theta/2$
- c. $var(2\bar{X}) = 4var(\bar{X}) = \frac{4var(X)}{n} = 4\frac{\sigma^2}{12n}$

Question 5

Suppose that $\hat{\theta}$ is an estimator for parameter θ and $E(\hat{\theta}) = a\theta + b$ for some nonzero constants a and b

- a. In terms of a , b and θ , what is $B(\hat{\theta})$?
- b. Find a function $\hat{\theta}^*$, say $\hat{\theta}^*$ that is an unbiased estimator for θ .
- c. Consider $\hat{\theta}^*$ you found in b. Express $MSE(\hat{\theta}^*)$ as a function of $V(\hat{\theta})$.
- d. Give an example of a value of a for which $MSE(\hat{\theta}^*) < MSE(\hat{\theta})$ and $MSE(\hat{\theta}^*) > MSE(\hat{\theta})$.

Solution:

- a. The bias of a point estimator $\hat{\theta}$ is given by $B(\hat{\theta}) = E(\hat{\theta}) - \theta$, so $B(\hat{\theta}) = a\theta + b - \theta = (a - 1)\theta + b$
- b. Let $\hat{\theta}^* = (\hat{\theta} - b)/a$. We must have $E(\hat{\theta}^*) = \theta$ if it is unbiased. $E(\hat{\theta}^*) = (E(\hat{\theta}) - b)/a$ so we proved that it is unbiased.
- c. We already found that $E(\hat{\theta}^*) = \theta$ and $V(\hat{\theta}^*) = V[(\hat{\theta} - b)/a] = V(\hat{\theta})/a^2$. Then $MSE(\hat{\theta}^*) = V(\hat{\theta}^*) = V(\hat{\theta})/a^2$.
- d. $MSE(\hat{\theta}) = V(\hat{\theta}) + B(\hat{\theta})^2 = V(\hat{\theta}) + [(a-1)\theta + b]^2$. A sufficiently large value of a will force $MSE(\hat{\theta}^*) < MSE(\hat{\theta})$ for example when $a=10$. Similarly, a sufficiently small value of a will make $MSE(\hat{\theta}^*) > MSE(\hat{\theta})$. $E(\theta)$ for example when $a = 0.5$ and $b = 0$.

Question 6

Suppose $S^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1}$ $S_1^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n}$ Then S_1^2 is a biased estimator of σ^2 , but S^2 is an unbiased estimator of the same parameter. If we sample from a normal population,

- a. Find $V(S^2)$ and show that $V(S^2) > V(S_1^2)$
- b. Which is a better estimator $V(S^2)$ or $V(S_1^2)$ (Hint: compare mean square errors).
- c. $V(S^2)$ and $V(S_1^2)$ are two estimators of σ^2 of the form $S^2 = c \sum_{i=1}^n (Y_i - \bar{Y})^2$. What value for c yields the estimator σ^2 with the smallest mean square error among all estimators of the form $S^2 = c \sum_{i=1}^n (Y_i - \bar{Y})^2$.

Solution:

- a. $V(S_1^2) = V(\frac{n-1}{n})S^2 = \frac{2(n-1)\sigma^4}{n^2}$
 $V(S_1^2) = V(\frac{n-1}{n})S^2 = (\frac{n-1}{n})^2 V(S^2) < V(S^2)$ since $\frac{n-1}{n} < 1$.

b. Since S^2 is unbiased,

$$\text{MSE}(S^2) = \text{V}(S^2) = \frac{2\sigma^4}{n-1}. \text{ Similarly,}$$

$$\text{MSE}(S_1^2) = \text{V}(S_1^2) + (B(S_1^2))^2 = \frac{2(n-1)\sigma^4}{n^2} + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2 = \frac{2(n-1)\sigma^4}{n^2}$$

By considering the ratio of these two MSEs, it can be seen that S_1^2 has the smaller MSE and thus possibly a better estimator.

c. Define the estimator $\hat{\sigma}^2 = c \sum_{i=1}^n (Y_i - \bar{Y})^2$.

Therefore, $E(\hat{\sigma}^2) = c(n-1)\sigma^2$ and $\text{V}(\hat{\sigma}^2) = 2c^2(n-1)\sigma^4$ so that

$$\text{MSE}(\hat{\sigma}^2) = 2c^2(n-1)\sigma^4 + [c(n-1)\sigma^2 - \sigma^2]^2$$

Minimizing this with respect to c , we find the smallest MSE occurs when $c = 1/(n+1)$

Question 7

Suppose that $Y_1, Y_2, \dots, Y_n$ are iid exponential observations with mean θ . Derive an unbiased estimator for $T(\theta) = 1/\theta$

Solution: Since $E(\bar{Y}) = \theta$, one's intuition might suggest to try $1/\bar{Y}$ as an estimator for $1/\theta$. First, note that

$$E\left(\frac{1}{\bar{Y}}\right) = E\left(\frac{n}{\sum_{i=1}^n Y_i}\right) = nE\left(\frac{1}{T}\right)$$

where $T = \sum_{i=1}^n Y_i$. Recall that $Y_1, Y_2, \dots, Y_n$ iid exponential (θ) $\implies T \sim \text{gamma}(n, \theta)$, so therefore

$$\begin{aligned} E\left(\frac{1}{\bar{Y}}\right) &= nE\left(\frac{1}{T}\right) = n \int_{t=0}^{\infty} \frac{1}{t} \underbrace{\frac{1}{\Gamma(n)\theta^n} t^{n-1} e^{-t/\theta} dt}_{\text{gamma}(n, \theta) \text{ pdf}} \\ &= \frac{n}{\Gamma(n)\theta^n} \underbrace{\int_{t=0}^{\infty} t^{(n-1)-1} e^{-t/\theta} dt}_{=\Gamma(n-1)\theta^{n-1}} \\ &= \frac{n\Gamma(n-1)\theta^{n-1}}{\Gamma(n)\theta^n} = \frac{n\Gamma(n-1)}{(n-1)\Gamma(n-1)\theta} = \left(\frac{n}{n-1}\right) \frac{1}{\theta}. \end{aligned}$$

This shows that $1/\bar{Y}$ is a biased estimator of $\tau(\theta) = 1/\theta$. However,

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

This shows that

$$\widehat{\tau(\theta)} = \frac{n-1}{n\bar{Y}}$$

is an unbiased estimator of $\tau(\theta) = 1/\theta$.

Question 8

Let $Y_1, Y_2, \dots, Y_n$ denote a random sample of size n from a population whose density is given by

$$f(y) = \begin{cases} 3\beta^3 y^{-4}, & \beta \leq y \\ 0, & \text{elsewhere} \end{cases}$$

where $\beta > 0$ is unknown. Consider the estimator $\hat{\beta} = \min(Y_1, Y_2, \dots, Y_n)$.

a. Derive the bias of the estimator $\hat{\beta}$.

b. Derive $MSE(\hat{\beta})$

Solution: $E(Y) = (3/2)\beta$, $E(Y_1^2) = 3 * \beta^2$

$$g_1(y) = 3n\beta^{3n}y^{-(3n+1)}, y \geq \beta$$

$$E(Y_{(1)}) = (\frac{3n}{3n-1})\beta \text{ and } E(Y_{(1)}^2) = (\frac{3n}{3n-2})\beta^2$$

$$\text{With } \hat{\beta} = Y_{(1)}, B(\hat{\beta}) = (\frac{3n}{3n-1})\beta - \beta = (\frac{1}{3n-1})\beta$$