vc389 **Homework 1: Estimation Problems** CS 536

Let $X_1, X_2, ..., X_n$ be i.i.d. random variables, uniformly distributed on [0, L] (i.e., with density 1/L on this interval). In the posted notes on estimation, it is shown that the method of moments and maximum likelihood estimators for L are given by

$$\begin{aligned} \hat{L_{mom}} &= 2\bar{X}_n \\ \hat{L_{mle}} &= \max_{i=1,...,n} X_i \end{aligned} \tag{1}$$

We want to consider the question of which estimator is better. Recall the definition of the mean squared error of an estimator as

$$MSE(\hat{L}) = \mathbb{E}[(\hat{L} - L)^2] \tag{2}$$

Problem 1

Squaring the equation of bias,

$$\begin{aligned} bias(\hat{\theta})^2 &= (\theta - \mathbb{E}[\hat{\theta}])^2 \\ &= \theta^2 + \mathbb{E}[\hat{\theta}]^2 - 2\theta\,\mathbb{E}[\hat{\theta}] \end{aligned}$$

Expanding (2),

$$\begin{aligned} MSE(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \theta)^2] \\ &= \mathbb{E}[\hat{\theta}^2 + \theta^2 - 2\theta\hat{\theta}] \\ &= \mathbb{E}[\hat{\theta}^2] + \theta^2 - 2\theta\,\mathbb{E}[\hat{\theta}] \\ &= bias(\hat{\theta})^2 + \mathbb{E}[\hat{\theta}^2] - \mathbb{E}[\hat{\theta}]^2 \end{aligned} \tag{3}$$

Now, we know that

$$var(\hat{\theta}) = \mathbb{E}[\hat{\theta}^2] - \mathbb{E}[\hat{\theta}]^2 \tag{4}$$

Using 4 in 3,

$$MSE(\hat{\theta}) = bias(\hat{\theta})^2 + var(\hat{\theta})$$

Hence, proved.

Problem 2

For calculating bias through $\hat{L_{mom}}$, using (1),

$$\begin{aligned} bias(\hat{L_{mom}}) &= L - \mathbb{E}[2\hat{L}] \\ &= L - 2\,\mathbb{E}[\hat{L}] \\ &= L - \cancel{2}\frac{L}{\cancel{2}} \\ &= 0 \end{aligned} \tag{5}$$

Since the bias is zero, it shows that $\hat{L_{mom}}$ is unbiased.

For calculating bias through $\hat{L_{mle}}$, let $Y = \max_{i=1,...,n} X_i$.

Utilizing the fact that it is a uniform distribution, we will first find the p.d.f. of this distribution and then differentiate it to find its expected value,

$$\begin{aligned}
P(Y \leq x) &= P(\max_{i=1,...,n} X_i) \\
&= P(X_1 \leq x, X_2 \leq x, ...X_n \leq x) \\
&= P(X_1 \leq x)P(X_2 \leq x)...P(X_n \leq x) \\
&= P(X \leq x)^n \\
&= (\frac{x}{L})^n \\
f(x) &= \frac{n}{L}(\frac{x}{L})^{n-1} \\
\mathbb{E}[Y] &= \int_0^L xf(x)dx \\
&= \int_0^L x\frac{n}{L}(\frac{x}{L})^{n-1}dx \\
&= \frac{n}{L^n}\frac{L^{n+1}}{n+1} \\
&= \frac{n}{n+1}L
\end{aligned} \tag{6}$$

using (6),

$$\begin{aligned}
bias(\hat{L_{mle}}) &= L - \frac{n}{n+1}L \\
&= \frac{nL + L - nL}{n+1} \\
&= \frac{L}{n+1}
\end{aligned} \tag{7}$$

$\hat{L_{mle}}$ has bias since it is not zero. Since, n will always be greater than 1, $\hat{L_{mle}}$ will always be less than L, hence underestimating it.

Problem 3

For calculating variance through $\hat{L_{mom}}$,

$$\begin{aligned}
var(\hat{L_{mom}}) &= var(2\bar{X}) = 4 * var(\bar{X}) \\
&= 4 * var(\frac{1}{n}\sum X_i) = \frac{4}{n^2} * \sum var(X_i) \\
&= \frac{4}{n^2}\frac{L^2}{12} \\
&= \frac{L^2}{3n}
\end{aligned} \tag{8}$$

For calculating variance through $\hat{L_{mle}}$, let $Y = \max_{i=1,...,n} X_i$. Using p.d.f. and expected value from 6,

$$\begin{aligned} var(\hat{L_{mle}}) &= var(\max_{i=1,...,n} X_i) = var(Y) \\ &= \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 \\ \mathbb{E}[Y^2] &= \int_0^L x^2 f(x)dx \\ &= \int_0^L \frac{n}{L^n} x^{n+1} dx \\ &= \frac{nL^{n+2}}{(n+2)L^n} \\ &= \frac{nL^2}{n+2} \end{aligned} \tag{9}$$

Using 9,

$$\begin{aligned} var(\hat{L_{mle}}) &= \frac{nL^2}{n+2} - (\frac{n}{n+1})^2 L^2 \\ &= \frac{\cancel{n^3L^2} + \cancel{2n^2L^2} + nL^2 - \cancel{n^3L^2} - \cancel{2n^2L^2}}{n+2(n+1)^2} \\ &= \frac{nL^2}{(n+2)(n+1)^2} \end{aligned} \tag{10}$$

Problem 4

To compute which has lower MSE value, first let us calculate each estimation's MSE value. Using 5 and 8,

$$MSE(\hat{L_{mom}}) = 0 + \frac{L^2}{3n} = \frac{L^2}{3n} \tag{11}$$

Using 7 and 10

$$\begin{aligned} MSE(\hat{L_{mle}}) &= \frac{L^2}{(n+1)^2} + \frac{nL^2}{n+2(n+1)^2} \\ &= \frac{nL2 + 2L^2 + nL^2}{(n+1)^2(n+2)} \\ &= \frac{2L^2\cancel{(n+1)}}{(n+1)^{\cancel{2}}(n+2)} = \frac{2L^2}{(n+1)(n+2)} \end{aligned} \tag{12}$$

Subtracting 11 and 12, we can get to know which one is smaller

$$\begin{aligned} MSE(\hat{L_{mom}}) - MSE(\hat{L_{mle}}) &= \frac{L^2}{3n} - \frac{2L^2}{(n+1)(n+2)} \\ &= \frac{L^2(n-1)(n-2)}{3n(n+1)(n+2)} \end{aligned}$$

For any value of $n \geq 2$, we observe that $MSE(\hat{L_{mle}}) < MSE(\hat{L_{mom}})$. Hence, $\hat{L_{mle}}$ is a better estimator as it has a lower MSE value.

Problem 5

The code has been attached alongside this PDF.

After estimating the parameters through M.O.M and M.L.E, we find that the difference between the true MSE (found in Problem 4) and estimated MSE is very small.

Through the code, we also establish practically that $MSE(\hat{L_{mle}}) < MSE(\hat{L_{mom}})$, thereby validating our solution to Problem 4.

Problem 6

Although there is bias in $\hat{L_{mle}}$, we see that the variance of $\hat{L_{mom}}$ is way greater than $\hat{L_{mle}}$ when n increases. This leads us to the problem of bias-variance trade-off. We see $\hat{L_{mle}}$ has a better bias-variance trade-off to reduce the MSE value than $\hat{L_{mom}}$. $\hat{L_{mom}}$ estimation can be regarded as a low bias - high variance estimation.

Problem 7

From 6, we know that $P(Y \leq x) = (\frac{x}{L})^n$, where $Y = \max_{i=1,...,n} X_i$ So,

$$\begin{aligned} P(\hat{L_{mle}} \leq L - \epsilon) &= (\frac{L-\epsilon}{L})^n \\ &= (1 - \frac{\epsilon}{L})^n \end{aligned} \tag{13}$$

According to the problem, we now want to estimate how many samples would be needed to be sure that the estimate is within ϵ with probability at least δ

$$\begin{aligned} P(L - \hat{L_{mle}} < \epsilon) > \delta \\ 1 - P(L - \hat{L_{mle}} > \epsilon) > \delta \\ P(L - \hat{L_{mle}}) < 1 - \delta \end{aligned} \tag{14}$$

Using 13 and 14, and taking log

$$\begin{aligned} (1 - \frac{\epsilon}{L})^n &< 1 - \delta \\ n \log(1 - \frac{\epsilon}{L}) &< \log(1 - \delta) \\ n &> -\frac{\log(1-\delta)}{\log(1 - \frac{\epsilon}{L})} \\ n &> \frac{\log(\frac{1}{1-\delta})}{\log(1 - \frac{\epsilon}{L})} \end{aligned}$$

One point to notice is that $\log(1 - \delta$ is negative, hence in the above solution $<$ changes to $>$.

Problem 8

Given,

$$\hat{L} = (\frac{n}{n+1}) \max_{i=1,...,n} X_i$$

Bias for this estimator is, using 6

$$\begin{aligned} bias(\hat{L}) &= L - \mathbb{E}[\hat{L}] \\ &= L - (\frac{n+1}{n}) \, \mathbb{E}[\max_{i=1,...,n} X_i] \\ &= L - \frac{\cancel{n}}{\cancel{n+1}} \frac{\cancel{n+1}}{\cancel{n}} L = 0 \end{aligned} \tag{15}$$

Variance for this estimator is, using 9

$$\begin{aligned} var(\hat{L}) &= var((\frac{n+1}{n}) \max_{i=1,...,n} X_i) \\ &= (\frac{n+1}{n})^2 * var(\max_{i=1,...,n} X_i) \\ &= \frac{\cancel{(n+1)^2} \cancel{n} L^2}{n^{\cancel{2}} \cancel{(n+1)^2} (n+2)} \\ &= \frac{L^2}{n(n+2)} \end{aligned} \tag{16}$$

Now, calculating MSE through above two equations,

$$MSE(\hat{L}) = \cancel{bias(\hat{L})^2} + var(\hat{L})$$

$$MSE(\hat{L}) = \frac{L^2}{n(n+2)}$$

Comparing to the $MSE(\hat{L_{mle}})$, we see that $MSE(\hat{L})$ is lower as the value of n increases.