

CS 536 : Estimation Problems

Uniform Estimators

Let X_1, X_2, \dots, 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

$$\begin{aligned}\hat{L}_{\text{MOM}} &= 2\bar{X}_n \\ \hat{L}_{\text{MLE}} &= \max_{i=1, \dots, n} X_i.\end{aligned}\tag{1}$$

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

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

Note: the answers to homework zero may also be useful here.

1) Show that in general, $\text{MSE}(\hat{\theta}) = \text{bias}(\hat{\theta})^2 + \text{var}(\hat{\theta})$, where var is the variance, and bias is given by

$$\text{bias}(\hat{\theta}) = \theta - \mathbb{E}[\hat{\theta}].\tag{3}$$

Solution:

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

2) Show that \hat{L}_{MOM} is *unbiased*, but that \hat{L}_{MLE} has bias. In general, \hat{L}_{MLE} consistently underestimates L - why?

Solution:

$$\begin{aligned}\text{bias}(\hat{L}_{\text{MOM}}) &= L - \mathbb{E}[\hat{L}_{\text{MOM}}] \\ &= L - \mathbb{E}[2\bar{X}_n] \\ &= L - 2\mathbb{E}[\bar{X}_n] \\ &= L - 2\mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] \\ &= L - \frac{2}{n} \sum_{i=1}^n \mathbb{E}[X_i] \\ &= L - \frac{2}{n} n \mathbb{E}[X] \\ &= L - 2\mathbb{E}[X] \\ &= 0\end{aligned}$$

Thus \hat{L}_{MOM} is unbiased.

The cdf of MLE:

$$F(x) = \mathbb{P}\left(\max_{i=1, \dots, n} X_i \leq x\right) = \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) = \mathbb{P}(X_1 \leq x) \dots \mathbb{P}(X_n \leq x) = \left(\frac{x}{L}\right)^n.$$

Thus, the pdf of MLE is:

$$f(x) = F'(x) = \frac{n \cdot x^{n-1}}{L^n}.$$

Therefore,

$$\begin{aligned}\mathbb{E}[\hat{L}_{\text{MLE}}] &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_0^L x \frac{n \cdot x^{n-1}}{L^n} dx \\ &= \int_0^L n \left(\frac{x}{L}\right)^n dx \\ &= \frac{n}{n+1} \cdot \frac{x^{n+1}}{L^n} \Big|_0^L \\ &= \frac{n}{n+1} \cdot \frac{L^{n+1}}{L^n} \\ &= \frac{n}{n+1} L.\end{aligned}$$

We have

$$\text{bias}(\hat{L}_{\text{MLE}}) = L - \mathbb{E}[\hat{L}_{\text{MLE}}] = L - \frac{n}{n+1} L = \frac{1}{n+1} L \neq 0.$$

Thus, \hat{L}_{MLE} has bias.

$$\mathbb{P}(\hat{L}_{\text{MLE}} \geq L) = \mathbb{P}(\max_{i=1, \dots, n} X_i \geq L) = 0.$$

Thus, \hat{L}_{MLE} consistently underestimates L .

3) Compute the variance of \hat{L}_{MOM} and \hat{L}_{MLE} .

Solution:

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mathbb{E}(X))^2] = \mathbb{E}[(X - \frac{X}{2})^2] = \frac{1}{4} \mathbb{E}[X^2] = \frac{1}{4} \int_{-\infty}^{\infty} x^2 f(x) dx = \frac{1}{4} \int_0^L x^2 \frac{1}{L} dx \\ &= \frac{1}{4} \frac{1}{3} \frac{1}{L} x^3 \Big|_0^L = \frac{1}{12L} L^3 = \frac{L^2}{12} \end{aligned}$$

$$\text{Var}(\hat{L}_{\text{MOM}}) = \text{Var}(2\bar{X}_n) = 4\text{Var}(\bar{X}_n) = 4\text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{4}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{4}{n^2} n \text{Var}(X) = \frac{L^2}{3n}$$

$$\text{Var}(\hat{L}_{\text{MLE}}) = \text{Var}\left(\max_{i=1, \dots, n} X_i\right) = \text{Var}(X) = \frac{L^2}{12}$$

4) Which one is the better estimator, i.e., which one has the smaller mean squared error?

$$\begin{aligned} \text{MSE}(\hat{L}_{\text{MOM}}) &= \text{bias}(\hat{L}_{\text{MOM}})^2 + \text{var}(\hat{L}_{\text{MOM}}) = \frac{L^2}{3n} \\ \text{MSE}(\hat{L}_{\text{MLE}}) &= \text{bias}(\hat{L}_{\text{MLE}})^2 + \text{var}(\hat{L}_{\text{MLE}}) = \left(\frac{1}{n+1}L\right)^2 + \frac{L^2}{12} \end{aligned}$$

5) Experimentally verify your computations in the following way: Taking $n = 100$ and $L = 10$,

- For $j = 1, \dots, 1000$:
- Simulate X_1^j, \dots, X_n^j and compute values for \hat{L}_{MOM}^j and \hat{L}_{MLE}^j
- For $n = 100, L = 10$, simulate X_1, \dots, X_n , and compute values for \hat{L}_{MOM} and \hat{L}_{MLE} .
- Estimate the mean squared error for each population of estimator values.
- How do these estimated MSEs compare to your theoretical MSEs?

6) You should have shown that \hat{L}_{MLE} , while biased, has a smaller error over all. Why? The mathematical justification for it is above, but is there an explanation for this?

7) Find $\mathbb{P}(\hat{L}_{\text{MLE}} < L - \epsilon)$ as a function of L, ϵ, n . Estimate how many samples I would need to be sure that my estimate was within ϵ with probability at least δ .

8) Show that

$$\hat{L} = \left(\frac{n+1}{n}\right) \max_{i=1, \dots, n} X_i, \quad (4)$$

is an unbiased estimator, and has a smaller MSE still.