CS 536 : Estimation Problems

16:198:536

Uniform Estimators

Let $X_1, X_2, \ldots, X_n$ 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

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

We want to consider the question of which estimator is better. If $\hat{L}$ is meant to be an estimator for L, we define the **mean squared error** to be

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

the expected square discrepancy between the estimator and the thing it is supposed to be estimating.

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

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

Note that an estimator might have no bias, but huge variance, or no variance (constant), but significant bias - the MSE summarizes these two sources of 'error' in an estimator.

2) Compute the bias of $\hat{L}_{\text{MOM}}$ and $\hat{L}_{\text{MLE}}$. In general, $\hat{L}_{\text{MLE}}$ consistently underestimates L - why? *Hint: What is the pdf for $\hat{L}_{MLE}$?*

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

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

5) Experimentally verify your computations in the following way: Taking $n = 100$ and $L = 10$,
   - For $j = 1, \ldots, 1000$:
     * Simulate $X_1^j, \ldots, X_n^j$ and compute values for $\hat{L}^j_{\text{MOM}}$ and $\hat{L}^j_{\text{MLE}}$
   - 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 $\hat{L}_{\text{MLE}}$, 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 $\mathbb{P}\left(\hat{L}_{\text{MLE}} < L - \epsilon\right)$ 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

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

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