.

Activity 06
Dr. Shafayat Abrar
CE/MATH 362/322-L1

Course: Statistics and Inferencing
Due Date: Sept. 22, 2023
Due Time: 08:00 PM

Computer Simulation of the Estimator Discussed in the Classroom

In the classroom, we discussed the following estimation problem:

Problem 01: Consider $X \sim \mathcal{U}[0,T]$, and also consider a sensor measurement Y made on X such that, given X, Y is conditionally uniform from 0 to X,

$$Y \mid X \sim \mathcal{U}[0,X]$$

I showed you that the optimal MSE estimator X given the measurement $Y=y$ is

$$\widehat{X} = h(Y) = \mathrm{E}[X \mid Y] = \frac{T-y}{\ln T - \ln y} \tag{1}$$

We also obtained the variances of and $\widehat{X}$ and $\widetilde{X} = X - \widehat{X}$, as follows:

$$\begin{aligned}
\sigma_{\widehat{x}}^2 &= \mathrm{E}\widehat{x}^2 - (\mathrm{E}\widehat{x})^2 \\
&= T^2 \ln\left(\frac{4}{3}\right) - \frac{T^2}{4} \\
&= T^2\left(\ln\left(\frac{4}{3}\right) - \frac{1}{4}\right) & (2) \\
\sigma_{\widetilde{x}}^2 &= \sigma_x^2 - \sigma_{\widehat{x}}^2 \\
&= \frac{T^2}{12} - (\ln(4/3) - 1/4)T^2 \\
&= T^2\left(\frac{1}{3} - \ln\left(\frac{4}{3}\right)\right) & (3)
\end{aligned}$$

In the following, I show how a MATLAB simulation may be designed to verify the above-mentioned findings.

- We conduct Monte-Carlo simulations. In simple words, I mean that we conduct a random experiment N many times, and we take the average of those N outcomes to *numerically* estimate the means and variances of the outcomes.
- We generate a single realization of $X \sim \mathcal{U}[0, T]$ for $T = 1$ in MATLAB as follows:

$$\texttt{T} = \texttt{1};$$
$$\texttt{X} = \texttt{rand} * \texttt{T};$$

- Then, we generate a single realization of $Y \mid X \sim \mathcal{U}[0, X]$ in MATLAB as follows:

$$\texttt{Y} = \texttt{rand} * \texttt{X};$$

- Using the value of `Y`, you can compute the estimator $\widetilde{X}$, as shown in equation (1); denote it as `X_est` in MATLAB. Similarly, compute the estimation error $\widetilde{X} = X - \widehat{X}$ using the MATLAB values `X` and `X_est`, and denote it as `X_err` such that

$$\texttt{X\_err} = \texttt{X} - \texttt{X\_est}$$

- Repeat this experiment N many times and store the values of `X_est` and `X_err`.
- Finally estimate the variances of stored values of `X_est` and `X_err` using the MATLAB command `var`, and compare it with their respective analytical variances, shown earlier in equations (2) and (3), respectively. These analytical and numerical values must be close if N is reasonably large.
- Repeat the experiment for different values of T like $5, 10, 15$, etc.
- For each value of T, choose different values of N like $10^3, 10^4, 10^5$, etc.
- **Record your observations.** Like how the value of N affects the outcomes, whether the values of T have any effect on the accuracy of estimation, and whether the analytical and numerically computed variances are close to each other.
- Consider a data vector `X` containing N elements.

$$X = [X_1, X_2, \cdots, X_N]^T$$

  The MATLAB command `var` computes the variance of a data vector as follows:

  1. It first computes an **estimate** of sample mean of the data vector `X` as follows:

  $$\widehat{X}_{\text{mean}} = \frac{1}{N} \sum_{i=1}^{N} X_i$$

  2. Secondly, it computes an estimate of variance as follows:

  $$\widehat{\text{var}(X)} = \frac{1}{N-1} \sum_{i=1}^{N} \left(X_i - \widehat{X}_{\text{mean}} \right)^2.$$

  3. The reason why we divide the mean and variance estimates by N and $N-1$, respectively, will be explained later in class when we discuss the bias and variance of optimal and non-optimal maximum likelihood estimators.

Sample MATLAB Code for Problem 01:

A sample MATLAB code is shown below:

```
%% Monte-Carlo simulation
clc; clear
Tall = [1 5 10 15];
N = 1e3;
for jj=1:length(Tall)
    T=Tall(jj);
    for ii=1:N
        X(ii)=rand*T;
        Y(ii) = rand*X(ii);
        X_est(ii)=(T-Y(ii))/(log(T)-log(Y(ii)));
        X_err(ii)=X(ii)-X_est(ii);
    end

    disp('==============================================================')
    disp(['Value of T = ' num2str(T)])
    disp('==============================================================')
    disp('')
    disp(['Variance of X (Theory) = ' num2str(T^2/12)])
    disp(['Variance of X (Sim.) = ' num2str(var(X))])
    disp('')
    disp(['Variance of X_est (Theory) = ' num2str(T^2*(log(4/3)-1/4))])
    disp(['Variance of X_est (Sim.) = ' num2str(var(X_est))])
    disp('')
    disp(['Variance of X_err (Theory) = ' num2str(T^2*(1/3-log(4/3)))])
    disp(['Variance of X_err (Sim.) = ' num2str(var(X_err))])
    disp('')
end
```

In Activity 02, you were assigned the following estimation problem:

Problem 02: Consider $X \sim \mathcal{U}[0, T]$, and also Consider a sensor measurement Y made on X such that

$$Y \mid X \sim \mathcal{U}[X, X + \mu]$$

where $\mu \in \mathbb{R}^+$. Therefore, Y conditioned on X is uniformly distributed in the range $[X, X + \mu]$. Although it is not necessary, we may assume that μ is much smaller than T to make the sensor reasonably acceptable.

For the given measurement $Y = y$, you obtained the following MSE estimator:

$$\widehat{X} = h(Y) = \mathrm{E}[X \mid Y] = \begin{cases} \dfrac{y}{2} & \text{for} \quad 0 \leqslant y \leqslant \mu \\[2ex] y - \dfrac{\mu}{2} & \text{for} \quad \mu \leqslant y \leqslant T \\[2ex] \dfrac{T + y - \mu}{2} & \text{for} \quad T \leqslant y \leqslant T + \mu. \end{cases}$$

In Activity 04, you have also obtained the variance of the abovementioned estimator.

Task:

- Select appropriate values of μ and T. Conduct a Monte-Carlo experiment. All you need to do is to make modifications (in the sample code shared in this document) in the evaluation of observation `Y` and the estimator `X_est`.
- Show that the analytical variances of $\widehat{X}$ and $\widetilde{X}$ (as computed in Activity 04 for estimation **problem 02**) and their numerical estimates obtained by the Monte-Carlo experiment match perfectly for sufficiently large number of trials.