Activity-06
Statsitics and inferencing
Analysis of Theoretical and Simulated Variances
Using Monte-Carlo Simulations

Name: Basil khowaja
ID: bk08432

1 Introduction

In this report, I investigate the estimation of variances of three key variables: X, X_{est} (the estimator of X), and X_{err} (the error term between X and X_{est}). These estimations are done using Monte-Carlo simulations and compared to their respective theoretical variances.

The random variable X is uniformly distributed, $X \sim U[0, T]$, and the conditional random variable $Y|X$ is uniformly distributed over $[X, X + \mu]$. The goal is to compute the variance of the estimator X_{est} using both theoretical and simulated approaches. We are particularly interested in how the results of the simulations align with the theoretical expectations, especially as T and the number of simulations N change.

Monte-Carlo simulations are used to estimate the variances, and these estimates are compared to analytical expressions derived in class for X_{est} and X_{err}. By running the simulations for different values of T, we aim to observe how the accuracy of the estimates is affected by increasing T and how the number of trials N impacts the simulation's convergence to theoretical values.

2 Simulation Setup

To conduct the Monte-Carlo experiment, the following setup was used:

- The random variable $X \sim U[0, T]$ was generated using uniform random numbers in the range $[0, T]$.
- The conditional variable $Y|X \sim U[X, X + \mu]$ was generated by adding a random number scaled by μ to the value of X.
- For each value of T, the estimator X_{est} was computed using the piecewise-

defined formula provided in class:

$$X_{\text{est}} = \begin{cases} \frac{Y}{2} & \text{if } 0 \le Y \le \mu \\ \frac{Y-\mu}{2} & \text{if } \mu \le Y \le T \\ \frac{T+Y-\mu}{2} & \text{if } T \le Y \le T+\mu \end{cases}$$

- The error term $X_{\text{err}} = X - X_{\text{est}}$ was calculated for each trial.
- We used values of $T = [1, 5, 10, 15]$, $\mu = 0.5$, and the number of simulations $N = 1000$.

For each trial, the variances of X, X_{est}, and X_{err} were computed. The simulation results were compared to the theoretical variances calculated using the following formulas:

- **Variance of X (Theoretical)**:

$$\text{Var}(X) = \frac{T^2}{12}$$

- **Variance of X_{est} (Theoretical)**:

$$\text{Var}(X_{\text{est}}) = \left(\frac{u^3}{16T} - \frac{T^2}{4}\right) + \left(\frac{T^2}{12} - \frac{\mu T}{2} + \frac{\mu^2}{4} - \frac{\mu^3}{12T}\right) \text{ for } \mu \le y \le T + \left(-\frac{\mu^2}{4} + \frac{3\mu^3}{16T} - \frac{T^2}{4} - \frac{\mu T}{2}\right)$$

- **Variance of X_{err} (Theoretical)**:

$$\text{Var}(X_{\text{err}}) = T^2 \left(\frac{1}{3} - \ln\left(\frac{4}{3}\right)\right)$$

3 Results

The following table shows a detailed comparison between the theoretical and simulated variances for X, X_{est}, and X_{err} for different values of T:

3.1 Analysis of Results

1. **Variance of X**: - The simulated variances for X closely match the theoretical variances for all values of T. This shows that the Monte-Carlo method is highly accurate for estimating the variance of uniformly distributed random variables, as expected. - As T increases, the variance naturally grows, but the relationship between theoretical and simulated values remains strong.

2. **Variance of X_{est}**: - The simulated variance of the estimator X_{est} shows a larger discrepancy compared to the theoretical value, particularly for larger

T	**Var(X) (Theory)**	**Var(X) (Simulated)**	**Var(X_{est}) (Theory)**	**Var(X_{est}) (Simulated)**	**Var(X_{err}) (Theory)**	**Var(X_{err}) (Simulated)**
1	0.0833	0.0817	0.093	0.0894	0.055	0.0512
5	2.0833	2.134	2.32	1.0747	1.37	0.5749
10	8.3333	8.3272	9.29	3.6988	5.48	2.1007
15	18.75	19.2306	20.85	7.7402	12.35	4.8803

Table 1: Comparison of Theoretical and Simulated Variances for X, X_{est}, and X_{err}

values of T. For $T = 15$, the theoretical variance is over 20.85, while the simulated value is significantly lower at 7.74. - This discrepancy is likely due to the relatively small number of simulations ($N = 1000$) used in the experiment. A higher N would result in better convergence of the simulated variance to the theoretical value.

3. **Variance of X_{err}**: - Similar to X_{est}, the simulated variances of X_{err} are smaller than the theoretical ones, especially for larger T. This suggests that as T increases, the estimation error becomes more variable, and a larger sample size is needed to accurately capture this variability. - The theoretical variance for X_{err} at $T = 15$ is 12.35, while the simulated value is only 4.88, a significant underestimation.

4 Conclusion

The Monte-Carlo simulations conducted for this task yielded close matches between the theoretical and simulated variances for X, confirming the validity of the simulation for this uniform distribution. However, the variances of X_{est} and X_{err} were underestimated, especially for larger values of T. This discrepancy suggests that the number of simulations ($N = 1000$) may be insufficient to capture the true variability of the estimator and error terms at higher T values.

4.1 Impact of Increasing T and N

Increasing T: As T increases, both the theoretical and simulated variances for X_{est} and X_{err} tend to diverge. This is likely because larger values of T introduce more variability into the system, and with a fixed N, the simulation is unable to capture this increased variability accurately.

Increasing N: The results strongly suggest that increasing N (the number of simulations) would improve the accuracy of the simulated variances for both X_{est} and X_{err}. A higher number of simulations would reduce the stochastic noise in the Monte-Carlo process and allow the simulation results to converge more closely to the theoretical values, particularly for larger values of T.

5 MATLAB Code for Monte-Carlo Simulation

The following MATLAB code performs a Monte-Carlo simulation for estimating the variances of X, X_{est}, and X_{err} using different values of T and μ:

```
%% Monte-Carlo simulation for Problem 2

clc; clear;

% Define the range of T and   values
Tall = [1 5 10 15]; % Different values of T
mu_all = [0.1 0.5 1]; % Different values of

% Number of trials (Monte Carlo simulations)
N = 1e5;

% Loop over different values of T and
for jj = 1:length(Tall)
    for kk = 1:length(mu_all)

        % Set the current values of T and
        T = Tall(jj);
        mu = mu_all(kk);

        % Preallocate arrays for X, Y, X_est, and X_err
        X = zeros(1, N);
        Y = zeros(1, N);
        X_est = zeros(1, N);
        X_err = zeros(1, N);

        % Monte Carlo simulation
        for ii = 1:N
            % Generate X from U[0, T]
            X(ii) = rand * T;

            % Generate Y | X from U[X, X +   ]
            Y(ii) = X(ii) + rand * mu; % Y is uniform from X to X +

            % Compute the MSE estimator X_est based on the given Y
            if Y(ii) <= mu
                X_est(ii) = (mu^3 / (16 * T)) - (T^2 / 4); % For 0 <= y <=
            elseif Y(ii) <= T
                X_est(ii) = (T^2 / 12) - (mu * T) / 2 + mu^2 / 4 - (mu^3 / (12 * T));
% For    <= y <= T
            else
                X_est(ii) = - mu^2 / 4 + (3 * mu^3 / (16 * T)) - (T^2 / 4) - (mu * T) / 2;
```

```
% For T < y <= T +
            end

            % Compute the estimation error X_err
            X_err(ii) = X(ii) - X_est(ii);
        end

        % Theoretical Variances
        var_X_theory = T^2 / 12;  % Variance of X

        % Theoretical variance of the estimator X_est based on the three conditions
        if Y(ii) <= mu
            var_Xest_theory = (mu^3 / (16 * T)) - (T^2 / 4);  % For 0 <= y <=
        elseif Y(ii) <= T
            var_Xest_theory = (T^2 / 12) - (mu * T) / 2 + mu^2 / 4 - (mu^3 / (12 * T));
% For    <= y <= T
        else
            var_Xest_theory = - mu^2 / 4 + (3 * mu^3 / (16 * T)) - (T^2 / 4) - (mu * T) / 2;
% For T < y <= T +
        end

        % Theoretical variance of the error X_err
        var_Xerr_theory = var_X_theory - var_Xest_theory;  % Var(Xerr) = Var(X) - Var(X_est)

        % Display results for each combination of T and
        disp('==================================================');
        disp(['Value of T = ' num2str(T) ', Value of    = ' num2str(mu)]);
        disp('==================================================');

        % Variance of X (theoretical and simulated)
        disp(['Variance of X (Theory) = ' num2str(var_X_theory)]);
        disp(['Variance of X (Simulated) = ' num2str(var(X))]);

        % Variance of the estimator X_est (simulated and theoretical)
        disp(['Variance of X_est (Theory) = ' num2str(var_Xest_theory)]);
        disp(['Variance of X_est (Simulated) = ' num2str(var(X_est))]);

        % Variance of the error X_err (simulated and theoretical)
        disp(['Variance of X_err (Theory) = ' num2str(var_Xerr_theory)]);
        disp(['Variance of X_err (Simulated) = ' num2str(var(X_err))]);

        disp('==================================================');
        disp(' ');
    end
```