Stochastic Simulation (MIE1613H) - Homework 1

Due: February 1, 2023

- Submit your homework on Quercus as a single PDF file by the deadline. Late submissions are penalized 10% each day the homework is late.
- At the top of your homework include your **name, student number, department, and program**.
- You may discuss the assignment with other students in general terms, but each student must solve the problems, write the code and the solutions individually.
- The simulation models must be programmed in Python. You must include both the source code (including comments to make it easy to follow) and the output of the simulation in your submission.
- To make it easier for us to check your solutions, set the random seed to 1 in all simulations using `np.random.seed(1)`.
- **It is very important to explain your answers and solution approach clearly and not just provide the code/results**. Full marks will only be given to correct solutions that are fully and clearly explained.

Problem 1. **(20 Pts.)** Assume that X is continuous and uniformly distributed in $[2, 10]$ (X is a continuous random variable). We are interested in computing $\theta = E[(X - 5)^+]$. (Note: $a^+ = \text{Max}(a, 0)$, i.e, if $a < 0$ then $a^+ = 0$ and if $a \geq 0$ then $a^+ = a$.)

(a) Compute θ exactly using the definition of expected value. **HINT:** Recall that a continuous Uniform random variable in $[a, b]$ has density function,

$$f(x) = \frac{1}{b-a},$$

for $x \in [a, b]$ and 0 otherwise.

(b) Estimate θ using Monte Carlo simulation and provide a 95% confidence interval for your estimate. **Note**: Use the `np.random.random()` method in Python and transform it to a sample of X. You may NOT use other built-in methods of Python to generate the samples.

(c) Create a plot that demonstrates the convergence of the Monte Carlo estimate to the exact value as the number of samples increases.

Problem 2. **(25 Pts.)** In the original TTF example we simulated the system until the time of first failure. Modify the simulation model to simulate the system for a given fixed number of days denoted by T. Assume that all other inputs and assumptions are the same as in the original example.

(a) We say that the system is fully functional provided that both components (active and spare) are functional. Denote by $A(t)$ a process that takes value 1 if the system is fully functional at time t and 0 otherwise. Then,

$$\bar{A}(T) = \frac{1}{T}\int_0^T A(t)dt,$$

is the fraction of time the system is fully functional between 0 and T. Modify your simulation model to estimate $\bar{A}(T)$ until $T = 1000$ based on one replication of the simulation.

(b) Estimate $\bar{A}(T)$ for $T = 2000$ and $T = 4000$ (again using a single replication) and compare the values with the estimate from part (a). Summarize your observation in one sentence.

Problem 3. **(25 Pts.)** Modify the TTF simulation from the lecture so that instead of 2 components there can be any number of components. The number of components N should be an input of the simulation model. As before, assume that one component is active and the rest are kept as spares. Also, only one component can be repaired at a time. Run your simulation for 1000 replications and report a 95% confidence interval for the expected time to failure of the system for $N = 2, 3$, and 4.

Problem 4. (5 Pts.) The standard error of an estimator is defined as the standard deviation of that estimator. In the lecture, we introduced the sample mean $\bar{X}_n = (1/n)\sum_{i=1}^n X_i$ as an estimator of $E[X]$ where X_i's are iid samples of the random variable X. What is the standard error of the estimator $\bar{X}_n$? Assume that the standard deviation of X is σ.

Problem 5. (10 Pts.) Consider random variables X, Y and constants a, b. Using the properties of expected value and definitions of covariance and variance covered in the second lecture, show that the following statements are true (make sure to clearly explain your derivations and specify the properties you use in each step):

(a) $Var(aX + b) = a^2 Var(X)$

(b) $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$.

Problem 6. (15 Pts.) Assume that $\{X_i; i \geq 1\}$ is a sequence of independent uniform random variables between $(0, 1)$, i.e., $X_i \sim \text{Unif}(0, 1)$ for all $i \geq 1$. Consider the sample average of n such random variables,

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i.$$

(a) Simulate 1000 samples of $\bar{X}_1, \bar{X}_2, \bar{X}_{30}, \bar{X}_{500}$ and plot a histogram of the samples for each case (i.e. 4 histograms in total). What happens to the mean and variability of the sample means as n increases? Relate your observations to the Central Limit Theorem discussed in the class. **Note**: You can plot a histogram in Python using the `plt.hist` method of the `matplotlib.pyplot` library.