Stochastic Simulation (MIE1613H) - Homework 1

Due: February 3, 2021

- 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.
- **You should fully and clearly explain your answers**. Full marks will only be given to correct solutions that are fully and clearly explained.
- To make it easier for us to check your solutions, set the random seed to 1 in all simulations using `np.random.seed(1)`.

Problem 1. **(20 Pts.)** Assume that X is uniformly distributed in $[2, 8]$ (X is a continuous random variable). We are interested in computing $\theta = E[(X-4)^+]$. (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.

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

(a) What is the average number of functional components until time $T = 1000$ based on one replication of the simulation?

(b) We say that the system is available provided that there is at least one functional component. Denote by $A(t)$ a process that takes value 1 if the system is available and 0 otherwise. Then,

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

is the average system availability from time 0 to T. Modify your simulation model to estimate the average system availability until $T = 1000$ based on one replication of the simulation.

(c) Obtain the above estimates until $T = 3000$ and compare them with the estimates from part (a) and (b). Summarize your observation in one sentence.

HINT: To get the simulation to stop at time T use the functional version of the code and create a new event called EndSimulation, with a variable NextEndSimulation as part of the event calendar. Note that you also need to modify the logic of the model in the failure and repair event functions.

Problem 3. (30 Pts.) Modify the TTF simulation 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 when $N = 2, 3$, 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. (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 samples 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.