ISyE 6644 – Spring 2023

Homework #3 — Due Tuesday, February 28

Follow the instructions on the syllabus to submit a single archive containing all Simio files and codes, and a single PDF document containing your discussion and findings. Name the archive as LastName-FirstName.zip. Problems 3–7 require a careful study of Chapter 4 of Law (2015) and Module 3 on stationary processes.

1. Consider Problem 10 from Chapter 5 of the Simio text of KSS. Build a model and conduct an experiment with 20 replications, each over a 250-hour time window with a 50-hour warm-up period. Create (S)MORE plots for

   - the average time-in-system for a part,
   - the average number of parts in the system,
   - the average number of reworks,
   - the maximum number of reworks, and
   - the hourly rate at which good parts are finished (total number of parts produced divided by 200).

   The main challenge will be the collection of statistics for the number of parts that are reworked. A "reworked" part must have been processed by server S1 more than once. The models in Chapter 5 use an entity state variable to store the number of times a part has been processed. For this problem, you need to create a model state variable that accounts for the number of parts undergoing rework and a state statistic to keep track of the value of your model state. The model state can be updated using state assignments and/or add-on processes at the input node of server S1 and the input nodes of the four sinks. In either case, you must be careful to (a) increase the current count of parts undergoing rework only the first time they are reworked and (b) decrease the current count of parts undergoing rework only when a part leaving has actually undergone rework.

   Hint: The creation of Status plots (and other animation features) is discussed extensively in Module 3 of the *SIMIO ONLINE COURSE*. To create a Status plot for the instantaneous number of parts in the system, select the Status Plot from the Animation menu, draw the area of the plot on the canvas with your mouse, and add the Expression Part.Population.NumberInSystem. You can beautify the properties of a plot (title, axes, etc.) by selecting it. You can add the other two plots using the appropriate additional expressions.

   Bonus [3 points]: In addition to the status plots, attach a floor label or a dynamic label text to each entity (part) with the ID of the part (the Simio expression is Entity.ID) and the number of times it has been processed. This topic is discussed in Module 3(IV) of the *SIMIO ONLINE COURSE*.

2. Enhance the model in Problem 1 to add:

   - Calendar-Time-Based failures at Server S1 with uptimes between failures from the exponential distribution with mean 10 hours and repair times from the uniform distribution between 5 and 10 minutes.

- Processing-Count-Based failures at Server S4: after 100 parts are processed, the server requires cleaning that is uniformly distributed between 5 and 10 minutes.

3. Suppose that $\{X_1, X_2, \ldots, X_n\}$ is a finite sample from a stationary process with autocovariance function $\{C_j = \mathrm{Cov}(X_i, X_{i+j}), j = 0, \pm 1, \pm 2, \ldots\}$, so that $C_0 = \mathrm{Var}(X_i) = \sigma_X^2$, and autocorrelation function $\{\rho_j = C_j/\sigma_X^2, j = 0, \pm 1, \pm 2, \ldots\}$. Let $\bar{X}_n$ and S_n^2 be the sample mean and sample variance, respectively. Show that

$$E(S_n^2) = \sigma_X^2 \left[1 - \frac{2}{n-1} \sum_{j=1}^{n-1} (1 - j/n)\rho_j \right]$$

and then use the latter equation and the expression for the $\mathrm{Var}(\bar{X}_n)$ on slide 8 to obtain the relationship

$$E\left(\frac{S_n^2}{n}\right) = \frac{(n/a_n) - 1}{n-1} \mathrm{Var}(\bar{X}_n), \quad \text{where } a_n = 1 + 2\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\rho_j.$$

<u>Hint</u>: Start with the expression for S_n^2, write each $(X_i - \bar{X}_n)^2 = [(X_i - \mu) - (\bar{X}_n - \mu)]^2$, expand the squares inside the summation, and then take expectations...

4. Suppose that the ε_i are i.i.d. random variables from the normal distribution with mean 0 and standard deviation 4. Define the process $\{X_i\}$ by

$$X_i = a + \varepsilon_i + \tfrac{3}{4}\varepsilon_{i-1} + \tfrac{1}{4}\varepsilon_{i-2}, \quad i = 1, 2, \ldots,$$

where a is a constant.

   (a) Find the marginal mean and variance of X_i.
   (b) Compute the $\mathrm{Cov}(X_i, X_{i+j})$ for all i and $j = \pm 1, \pm 2, \ldots$. Argue that the process $\{X_i\}$ is weakly stationary by showing that the $\mathrm{Cov}(X_i, X_{i+j})$ depends only on the lag j.
   (c) Is the process $\{X_i\}$ (strictly) stationary? Give a concise answer.
   (d) Find the autocorrelation function $\rho_j = \mathrm{Corr}(X_i, X_{i+j})$, $j = 0, \pm 1, \pm 2, \ldots$.
   (e) Find an expression for the $\mathrm{Var}(\bar{X}_n)$ using parts (b) and (d).
   (f) Compute the asymptotic variance $\sigma^2 = \lim_{n\to\infty} n\mathrm{Var}(\bar{X}_n)$.
   (g) Use Problem 3 above to compare $E(S_n^2/n)$ and $\mathrm{Var}(\bar{X}_n)$.

5. Suppose that $\{X_i : i \geq 0\}$ is a stationary Gaussian AR(1) sequence defined by the recursion $X_i = \mu + \phi(X_{i-1} - \mu) + \varepsilon_i$, where $\phi \in (-1, 1)$, $X_0 \sim N(\mu, 1)$ and the errors ε_i are i.i.d. from the normal distribution with mean 0 and variance $1 - \phi^2$. Clearly, the marginal distribution of the process is $N(\mu, 1)$. Recall that we obtained the autocorrelation function of this process in class (the constant μ plays no role), and showed that the asymptotic variance parameter is $\sigma^2 = \lim_{n\to\infty} n\mathrm{Var}(\bar{X}_n) = (1+\phi)/(1-\phi)$.

   Set $\mu = 0$ and $\phi = 0.9$. For $n = 1, 2, \ldots$ define the random functions

$$Z_n(t) \equiv \frac{\lfloor nt \rfloor (\bar{X}_{\lfloor nt \rfloor} - \mu)}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad t \in [0, 1].$$

(a) We will try to demonstrate that as $n \to \infty$, the random function $Z_n(t)$ converges to a standard Brownian motion process on the interval $[0, 1]$. (Of course, the Brownian motion is defined for $t \in [0, \infty)$, but we will focus on its portion on $[0, 1]$.) For $n = 10^5$ simulate and plot the values $Z_n(1/n), Z_n(2/n), \ldots, Z_n(1)$.

(b) By the central limit theorem for stationary sequences, the random variables $Z_n(1)$ converge to the standard normal distribution as $n \to \infty$. To evaluate the convergence to the Brownian motion at another value of t, conduct 400 independent replications of the experiment; each experiment will produce a realizations of $Z_n(1/2)$. Then assess the fit of the normal distribution with mean 0 and variance $1/2$ to the dataset of the 400 realizations of $Z_n(1/2)$ using a histogram or a formal goodness-of-fit test such as the Anderson–Darling goodness-of-fit test with level of significance 0.05.

6. Consider the Gaussian process in Problem 5 with $\mu = 10$ and $\phi = 0.9$. Let's conduct the following three experiments pretending that you don't know the true mean (μ) of the process. I strongly recommend using a Python/R script or a computer language.

   (a) Simulate 400 independent sample paths, each comprised of $n = 10^4$ observations. For each sample path, compute the 95% CI

$$\bar{X}_n \pm t_{n-1,0.975} \frac{S_n}{\sqrt{n}}, \tag{1}$$

   where $\bar{X}_n$ and S_n^2 are the sample mean and sample variance of the respective dataset (I did not list the index for the dataset to reduce notation). Report the fraction of the CIs that contain μ.

   (b) Repeat the experiment in part (a) with $n = 10^5$ observations. Does the fraction of the CIs (1) that contain μ increase compared to the outcome of part (a)? Give a plausible explanation. Hint: use Problem 3.

   (c) Simulate 400 independent sample paths, each comprised of $n = 10^4$ observations. For each sample path, compute the 95% CI

$$\bar{X}_n \pm z_{0.975} \frac{\sigma}{\sqrt{n}}.$$

   Report the fraction of the CIs that contain μ.

   Explain the outcomes of the experiments in parts (a)–(c) above.

7. Repeat the experiments in Problem 6(a)–(b) above with $\mu = 10$ and $\phi = -0.9$. Explain the outcome of these experiments against those in Problem 6(a)–(b).