Assignment 3

1. Let X be a random sample for some process, such as the process in Exercise 3. Let X_k, for $k = 1, \cdots, n$, is a collection of independent samples of X, which means independently simulated sample paths. Let B be a set of paths we consider "bad". We want to estimate $p = \Pr(X \in B)$. For example, we might consider a path bad if $N_{\max} > 60$ (say) in Exercise 3b. In this case B is the set of such paths, so $X \in B$ means $N_{\max} > 60$. Let $Y_k = \mathbf{1}_B(X)(X_k)$. This is the *indicator function*

$$\mathbf{1}_B(X)(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B \end{cases} .$$

This may also be called the *characteristic function* of the set B and written $\chi_B(x)$. [*Warning*. Probabilists use the term "characteristic function" to mean something else.] We want to use simulation to estimate p. For this exercise, assume p is very small (maybe much less than 1%). A *hit* is a simulated sample path X_k with $X_k \in B$ (so $Y_k = 1$). Let N be the (random) number of hits in n independent simulations. We estimate p using the natural estimator

$$\widehat{p} = \frac{N}{n} .$$

   (a) Show that the expected number of hits is np. Show that $\widehat{p}$ is an *unbiased* estimator, which means

$$\mathrm{E}[\,\widehat{p}\,] = p .$$

   (b) Show that $\mathrm{var}(Y) \approx p$, to high accuracy if p is small. Show that, to high accuracy when p is small,

$$\mathrm{var}(\,\widehat{p}\,) \approx \frac{p}{n} .$$

   (c) Let $\sigma_{\widehat{p}}$ be the standard deviation of $\widehat{p}$. Show that, to high accuracy when p is small, the relative accuracy of $\widehat{p}$ is

$$\frac{\sigma_{\widehat{p}}}{p} \approx \frac{1}{\sqrt{\text{expected number of hits}}} .$$

   Explain why this is called relative accuracy of the estimator.

   (d) Show that if $p = 1\%$ and you want to estimate p with a relative accuracy of about 1%, you need about ten thousand hits which requires about a million independent simulations.

2. A *power law* if a function of the form $f(x) = x^\alpha$, for some exponent α. The *tail* of the PDF of a one component random variable is the behavior of its PDF $f(x)$ for large x. A random variable X has a *power law tail* if $f(x) \approx x^{-\beta}$ for large x and some $\beta > 1$. The exponent $\beta = -1$ is not allowed because "the tail has infinite mass". The tail is the part of $f(x)$ for $x > x_0$, and $\int_{x_0}^\infty x^{-1}\, dx = \infty$. [Fat tailed distributions are important, for example, in risk analysis. Power law tails, possibly with powers of $\log(x)$, are called *Pareto* tails.]

   (a) A simple PDF with a power law tail is

   $$f(x) = \begin{cases} \frac{1}{Z}(x+a)^{-\beta} & \text{if } x > 0 \\ 0 & \text{if } x < 0 \ . \end{cases} \tag{1}$$

   Use the relation

   $$\int_0^\infty f(x)\, dx = 1$$

   to find a formula for Z in terms of exponent β and the offset a.

   (b) Find a formula for $\mathrm{E}_{a,\beta}[X]$. For what values of β is this finite? For what values of β is $\mathrm{var}(X)$ finite? [This shows that fat-tailed distributions can have infinite mean, or finite mean and infinite variance.] For a fixed β, what is the range of $\mu_\beta(a) = \mathrm{E}_{a,\beta}[X]$? *Hint.* What happens to μ when $a \to 0$ and $a \to \infty$?

   (c) Find a formula for the CDF, $F(x) = \Pr(X \le x)$.

   (d) Use the formula for $F(x)$ to write an inverse CDF method sampler to create i.i.d. samples $X_k \sim f$. This should be a Python method that takes as arguments n, a, β, an instance of the random number generator class and returns a one index `numpy` array of n independent samples. Use the histogram method of Assignment 2 to verify that this sampler gives samples from the correct distribution. Make a plot with a histogram of samples and the target density. Choose bin size small enough and n large enough that the empirical histogram and target density agree to "plotting accuracy" (the dots of the empirical histogram lie on the curve of the target density). Choose β not large so that the tails are not small. When you do that, you will have to restrict the histogram range because of a few *outlier* samples that are very large.

3. Modify the event-driven simulation code to make the decay time S follow the shifted power law density (1). [This is a model of a web server processing (serving) requests. Some of the requests take a long time to process even though most are simple.] The arrival should continue to be a Poisson process with rate parameter λ.

   (a) Modify the posted code for event-driven simulation to use the fat tailed a, β distribution for the decay time. Modify the code to create a method that produces m independent *sample paths* (functions $n(t)$ for $0 \leq t \leq T_{\max}$). Make plots with more than one (but not too many) sample paths in a single plot. Keeping $\lambda = 20$ and $T_{\max} = 50$, experiment with a and β values to see what trajectories look like. If the equilibrium n is about 40, does the system approach equilibrium more or less quickly depending on a and β? Do a rather large number of simulations yourself but pick just a few to hand in.

(b) Let $N_{\max} = \max \{n(t) \mid 0 \leq t \leq T_{\max}\}$ be the largest number of things in the system up to time $T_{\max}$. Make a histogram of $N_{\max}$ with enough samples to find the 1% and .1% quantiles reasonably accurately. Since $N_{\max}$ is an integer, it is natural to take the bin size equal to one. Repeat this for several interesting parameter combinations (a and β) that show different behavior while keeping an equilibrium n with expected value around 40. You can estimate this equilibrium expected value by eye using graphs of a few sample paths rather than using a quantitative estimator.