CSE 107
Lab Assignment 5

In this project you will construct the CDF of a random variable $Z = Y_1 + Y_2$, where both Y_1 and Y_2 are exponential random variables with parameter $\lambda = 1$. You will present your CDF as a table similar in form to the Standard Normal CDF table posted on the class webpage. We will show in lecture that Z has a well-known distribution called the *Erlang* distribution. Here you will construct your CDF by performing 20,000 trials of the following simple experiment. Obtain samples Y_1 and Y_2 of the exponential distribution, then compute the sum $Z = Y_1 + Y_2$. Using these trials, compute the relative frequencies of the events $\{Z \le z\}$ for all $z \in \{0.0, 0.1, 0.2, 0.3, \ldots \ldots, 9.7, 9.8, 9.9\}$, which are 100 equally spaced points in the range 0.0 to 9.9, at distance 0.1 apart.

Your relative frequencies will constitute your estimates of $F_Z(z) = P(Z \le z)$. The format of your table will resemble the following

```
Sum of Exponentials CDF:

          .0       .1       .2       .3       .4       .5       .6       .7       .8       .9
       ----------------------------------------------------------------------------------------
  0.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  1.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  2.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  3.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  4.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  5.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  6.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  7.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  8.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
  9.0 | .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000    .0000
```

where of course $.0000$ is replaced in each case with an appropriate relative frequency. The intersection of the row labeled r and the column labeled c gives your estimate of $F_Z(r + c)$. So for instance, $F_Z(4.7)$ is found in row 4.0 and column $.7$.

Everything is clear except for one thing. How can you "sample" from the exponential distribution? What does that even mean? We know the PDF and CDF of the exponential random variable, but how do those functions enable you to obtain random values from that random variable? Most computer languages have built in libraries allowing you to sample from a uniform continuous distribution on the interval $[0, 1)$, i.e. pick a "random" real number x in the range $0 \le x < 1$. Usually this function is called something like $\mathrm{random}()$ or $\mathrm{rand}()$. Using such a function you can flip coins with various weights, which is all we've done up to now, but uniform is not exponential. Some languages, like MATLAB or R, may have a built-in capability for sampling from non-uniform distributions, but C, C++, Java and Python do not. So how do we transform a uniform random variable into an exponential?

We seek a function $g: \mathbb{R} \to \mathbb{R}$ such that if X is uniform on $[0, 1)$, then $Y = g(X)$ is exponential with parameter $\lambda = 1$. The function that does this is

$$g(x) = \ln\left(\frac{1}{1 - x}\right).$$

In other words, if X is uniform on $[0, 1)$, then $Y = g(X)$ is exponential with parameter $\lambda = 1$. Thus we can use our built-in random number generator to obtain a number x in the range $0 \le x < 1$, then plug it into $g(x)$ to get a random number y in the range $0 \le y < \infty$. It is an interesting and useful exercise to prove that the above function works as advertised, which we leave to the reader. More generally, it can be shown that the function

$$g(x) = \frac{1}{\lambda}\ln\left(\frac{1}{1-x}\right)$$

transforms a uniform random variable X on $[0,1)$ into an exponential random variable $Y = g(X)$ with parameter λ. Hint: assume that X has CDF

$$F_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \le x < 1 \\ 1 & \text{if } x \ge 1 \end{cases}$$

Then show that $Y = g(X)$ has CDF

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ 1 - e^{-\lambda y} & \text{if } y \ge 0 \end{cases}$$

With this piece of the puzzle in place, this project is in fact very straight forward. There is still much to do though, so please start early, and get help when you need it.