ECG790C - Problem Set 1 - Spring 2009

1. Suppose you have data that is recorded sequentially over time but the times are random. Write a MATLAB function that averages the data over time "bins" of length `Delta` starting at time `t0`.

   For example, suppose `times=[1;3;6;7;8]` and `x=[2 4;3 1;6 2;1 1;3 2]`, with `Delta=2` and `t0=0`.

   The function
   `[t,y]=timeavg(times,x,t0,Delta)`
   should return
   `t=[2;4;6;8;10]`
   and
   `y=[2 4;3 1;NaN NaN;3.5 1.5;3 2]`
   Note: bin i is the average of all observations for which $t(i) - \Delta \leq \text{time} < t(i)$, and that, if there are no observations in a given interval, the function returns a row of NaNs for that interval.

   Be sure to document your function and put in error checking to handle exceptional cases in a reasonable way.

2. Consider the function

   $$C(u,v,a) = \frac{a}{\ln\left(e^{a/u} + e^{a/v} - e^{a}\right)}$$

   where $u, v \in [0,1]$ and $a \in (0, \infty)$. The limiting cases are

   $$C(u,v,0) = \frac{uv}{u+v-uv}$$

   and $C(u,v,\infty) = \min(u,v)$. Also $C(0,v,a) = C(u,0,a) = 0$, $C(1,v,a) = v$ and $C(u,1,a) = u$. Write a MATLAB function that accepts vectors of u and v and scalar a and returns vectors of C. It should be written to handle limiting cases of u, v and a and to address potential overflow and other numerical problems. To do this you will need to write the function in a mathematically equivalent form. Test your code on a grid of values for u, v and a. Provide a script file that conducts the tests you perform.

3. A triangular probability distribution is defined with three parameters a, b and c. A random variable with a triangular distribution has domain $[a, c]$ and the most likely value is b (where the parameters must satisfy $a \leq b \leq c$).

   The PDF associated with this distribution.

$$f(x) = \begin{cases} \frac{2(x-a)}{(c-a)(b-a)} & x \leq b \\ \frac{2(c-x)}{(c-a)(c-b)} & x \geq b \end{cases}$$

   The name come from the fact that the PDF looks like a triangle with vertices $(a, 0)$, $(c, 0)$ and $(b, 2/(c-a))$.

   Write a MATLAB procedure that generates random numbers for this distribution. It should have the calling syntax

```
function x=randtri(n,a,b,c)
```

   and should return an $n \times 1$ vector.

   Hint: find the CDF associated with this distribution and use the inverse CDF method.

4. Using Monte Carlo and quasi-Monte Carlo integration, estimate the expectation of $f(X) = 1/(1 + X^2)$ where X is exponentially distributed with CDF $F(x) = 1 - \exp(-\alpha x)$ for $x \geq 0$. Compute an estimate using 100, 500,1000 and 10000 replicates using $\alpha = 2$. Discuss.

5. Consider the integral $\int_1^x \frac{1}{z} dz$ which, of course, is equal to $\ln(x)$. Using a change of variables, the integral is equal to

$$\int_0^1 \frac{x-1}{1 + u(x-1)} du$$

   Using this form, the integral can be approximated by generating a random sample of uniformly distributed values of u and taking the sample mean of the integrand, i.e., using

$$\frac{1}{n} \sum_{i=1}^{n} \frac{x-1}{1 + u_i(x-1)}$$

   It is possible that the Monte Carlo approximation can be improved by using variance reduction methods. Specifically use as control variates $u - 1/2$ and $(u - 1/2)^2 - 1/12$. Examine how much this improves the approximation by comparing the standard errors.