Assignment #5 (Due May 31)

1. (Machine reliability) A system needs N working machines to be fully operational; if the system is functioning with fewer than N working machines it is considered to be in a “degraded” state. To guard against system degradation, additional machines are kept available as spares. Whenever a machine stops, it is immediately replaced by a spare and is sent to the repair facility, which consists of two repairmen who each repair machines one at a time. If N machines are running upon the completion of a repair, then the repaired machine becomes available as a spare to be used when the need arises; if less than N machines are running upon completion of repair, the repaired machine is immediately put into use. The lowest-numbered stopped machine is always selected for repair and, similarly, the lowest-numbered spare machine is always selected to be put into use. Whenever both repairmen are idle and a machine stops, both repairmen are equally likely to initiate the repair. Successive repairs by repairman i (i = 1,2) are iid according to a distribution function G_i, and the successive times from when a machine starts running until the next failure of the machine are iid according to a distribution function H. (Only running machines, and not spares, are subject to failure.) There are initially M (> N) functional machines, with machines 1 through N running and machines N+1 through M available as spares. Let \( X(t) = (R_1(t),\ldots,R_M(t),L_1(t),L_2(t)) \), where

\[
R_j(t) = \begin{cases} 
2 & \text{if machine } j \text{ is running at time } t \\
1 & \text{if machine } j \text{ is available as a spare at time } t \\
0 & \text{if machine } j \text{ is awaiting or undergoing repair at time } t 
\end{cases}
\]

for 1 ≤ j ≤ M

\[
L_i(t) = \begin{cases} 
 j & \text{if repairman } i \text{ is repairing machine } j \text{ at time } t \\
0 & \text{if repairman } i \text{ is idle at time } t 
\end{cases}
\]

for i = 1,2.

a) Specify \( \{X(t): t ≥ 0\} \) as a GSMP with event set \( E = \{e_1,\ldots,e_M,e_{M+1},e_{M+2}\} \), where \( e_j = \text{“failure of machine } j\text{”} \) for 1 ≤ j ≤ M and \( e_{M+i} = \text{“end of repair by repairman } i\text{”} \) for i=1,2. (Assume that events never occur simultaneously.)

b) Now suppose that N = 3, M = 6, that the distribution function H is exponential with mean 3, and that the distribution functions G_1 and G_2 are each symmetric triangular on \([0,4/3]\). Write a program to simulate this system. Use an event-list data structure when implementing the clock-setting mechanism (you may use or adapt code from the class web site). Using your program, compute a 95% confidence interval for the following quantities:

i) The interquartile range for the amount of time that repairman 1 is busy during the interval [0,100]. Use m = 5 sections of k = 100 observations each.

ii) The long-run fraction of time that machine 1 is running (i.e., is in state 2 as per the GSMP definition). Use the regenerative method with regeneration points equal to successive entrances into the state \( s = (2,2,2,1,1,1,0,0) \) and base your estimate on 500 cycles. Use the
Taylor-Series approach for computing the ratio estimator and the confidence interval. Explain why successive entrances to states indeed form a sequence of regeneration points.

iii) Same performance characteristic as part (i), but now estimate the standard deviation (SD) instead of the interquartile range. Use the Taylor-series approach based on 500 runs. Based on the output of this simulation, give the approximate number of runs needed to estimate the SD to within \( \pm 2\% \) with 95\% probability.

c) Define a sequence of regeneration points for a modified version of the system in which the distribution H is not exponential and there is only one repairman.

2. Consider a simulation of a database system. Denote by \( U \) the average processing time for a database query over the course of a day, and by \( V \) the fraction (in \%) of database queries during the day that access a particular type of data (say, image data). Suppose that we obtain the following 10 samples of \( U \) and \( V \), based on 10 i.i.d. simulation replications.

<table>
<thead>
<tr>
<th>Run #:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U ):</td>
<td>10</td>
<td>20</td>
<td>15</td>
<td>30</td>
<td>5</td>
<td>16</td>
<td>2</td>
<td>12</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>( V ):</td>
<td>22</td>
<td>43</td>
<td>29</td>
<td>62</td>
<td>10</td>
<td>35</td>
<td>5</td>
<td>43</td>
<td>40</td>
<td>10</td>
</tr>
</tbody>
</table>

Compute a 95\% confidence interval for the coefficient of correlation between \( U \) and \( V \), i.e.,

\[
\text{Corr}(U,V) = \frac{\text{Cov}(U,V)}{\sqrt{\text{Var}(U)\text{Var}(V)}}
\]

a) Using the Taylor-Series method

b) Using the jackknife method

(Hint: write \( \text{Corr}(U, V) = g(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) \) for an appropriate function \( g \) where \( \mu_1 = E[U] \), \( \mu_2 = E[V] \), \( \mu_3 = E[U^2] \), \( \mu_4 = E[V^2] \), and \( \mu_5 = E[UV] \).

3. (Accelerating regenerations) Consider a discrete-time Markov chain \( \{X_n : n \geq 0\} \) with state space \( S = \{1, 2, \ldots, N\} \), where \( N \) is large. Suppose that the transition probabilities satisfy \( p(i, j) > 0 \) for each \( i \) and \( j \).

a) Suppose that we can find positive numbers \( p_1, p_2, \ldots, p_N \) such that the state-transition probabilities satisfy \( p(i, j) > p_j \) for each \( i \) and \( j \). Briefly describe a method for simulating the chain that yields a sequence of regeneration points which are usually more frequent than regeneration points based on entrances or exits from a fixed state. Specify these alternative regeneration points.\( [\text{Hint: write } p(i, j) = bq(j) + (1-b)r(i, j), \text{ where } b = \sum_{j=1}^{N} p_j, q(j) = p_j / b, \text{ and } r(i, j) = (p(i, j) - p_j) / (1-b).] \)
b) For a fixed state i, do the successive times \( N_1, N_2, \ldots \) at which the chain jumps from state i (to some other state) form a sequence of regeneration points? Briefly justify your answer.

c) Do the successive times \( N_1 + 1, N_2 + 1, \ldots \) decompose sample paths of the chain into cycles that are independent and/or identically distributed? Briefly justify your answer.

4. Suppose that \((X(t): t \geq 0)\) is a GSMP having regeneration points \( T_0 = 0, T_1, T_2, \ldots \) and that rewards accrue continuously at rate \( q(s) \) whenever the current state is \( s \in S \). Also suppose that the performance measure of interest is the \( \beta \)-discounted reward \( r \), defined as

\[
\begin{align*}
r &= E \left[ \int_0^\infty e^{-\beta u} q(X(u)) \, du \right].
\end{align*}
\]

a) Re-express the reward \( r \) in the form \( r = E[X]/E[Y] \), where each of \( X \) and \( Y \) is a random variable that is defined in terms of the first regenerative cycle of the GSMP (just like the random variables \( Y_1 \) and \( \tau_i \) in the standard regenerative method). Hint: write

\[
\int_0^\infty e^{-\beta u} q(X(u)) \, du = \int_0^{T_i} e^{-\beta u} q(X(u)) \, du + e^{-\beta T_i} \int_{T_i}^\infty e^{-(\beta - \beta^-) u} q(X(u)) \, du.
\]

b) Based on the above representation, we can use essentially the regenerative method to estimate \( r \), based on \( n \) cycles. In the \( i^{th} \) cycle (where \( 1 \leq i \leq n \)), we obtain a pair of observations \((X_i, Y_i)\). Give explicit formulas for \( X_i \) and \( Y_i \).