CSci 426/526 — Simulation — Fall 2009
Midterm Exam

This exam is due at the beginning of class (11:00m) on Tuesday, November 10 and will not be accepted late without prior approval. Graduate students must work all four problems; undergraduate students must work problem 1 and two of the remaining three problems.

You may not seek or receive assistance on this test from any person other than the TA (for coding questions) or me. Failure to observe this rule will be treated as an honor code violation.

Problem 1: [20 points] Mr. K is the manager of a new Snowball Express sno-cone stand which will be opening for business. The stand is operated by a single attendant who is behind a window. People come up to the window, tell the attendant which flavor sno-cone they want, wait for the attendant to fill their order, then pay and leave. When there is more than one customer at the window, a line forms so that customers are served in the order that they arrived. People always buy one and only one sno-cone at \$3 each. The cost of materials to make sno-cones is negligible, so Mr. K considers all of the money taken in to be profit, except that the attendant has to be paid from the money received.
After interviewing several applicants for the attendant's position, Mr. K has narrowed his choice to two people: Fran and Bill. Fran can complete a transaction (take the order, make the sno-cone, take payment and make change) in 20 seconds on the average. Bill can perform the same job in 30 seconds, on the average. Fran is faster than Bill but demands \$12 per hour, whereas Bill will work for \$6 per hour.
Market research has revealed that, on the average, Mr. K can expect one customer per minute to come up to the window. The research also shows that if a customer comes to the window and there are already N people in line, then the customer will promptly turn about in a huff and storm off to Ben and Jerry's.
Studies of real sno-cone stands have further shown that, even though the average customer arrival rate is one per minute, usually the amount of time between customers is less than that. It's the occasional lull in business that makes the average time between customers longer than that which is usually observed. The same is true of the time it takes to make sno-cones. Thus, assume that the time between customer arrivals and the time it takes a person to make a sno-cone are exponentially distributed random values.
(a) For both Fran and Bill, and based on 1 000 000 served customers, construct a table of the estimated steady-state probability that a customer will go to Ben and Jerry's when the number of people in line N is 1, 2, 3, 4, and 5. (b) Construct similar tables if the time to serve a customer for Fran is changed to be *Uniform*(5, 35) and the time to serve a customer for Bill changed to be *Uniform*(15,45). (c) Comment on how the probability of rejection (i.e., the probability to go to Ben and Jerry's) depends on the service process. (d) Assuming exponentially distributed service times for both Fran and Bill, if N is equal to 3, which person is the more cost-effective to hire: Fran or Bill? Provide an analytic solution. (e) Discuss what you did to convince yourself that your results are correct.

Problem 2: [20 points] The two–player game of tic tac toe is played on a 3×3 board where players "X" and "O" take turns writing their character on the board. A player wins if she has three of her characters in a line. The game ends when a player wins or there are no empty spaces on the board. Player "X" always goes first. Some example games:

X	O	X
O	X	
O		X

X wins

O		X
O	X	
O		X

O wins

O	X	X
X	X	O
O	O	X

Stalemate

Assume both players play "at random".

- Write a Monte–Carlo simulation to estimate the probability that player "X" wins, player "O" wins, and nobody wins.

- Explain your interpretation of playing "at random".
- Compute the axiomatic probabilities and compare with your simulation results.

Problem 3: [20 points] Suppose you play the "win in any order" version of Pick-3 (see exercise 6.1.7) each day for 365 days. You start with \$365 and it costs you \$1 each day to play. Let the discrete random variable X represent the amount of money you will have after 365 days, including the \$80 wins, if any. The possible values of X are 0, 80, 160, 240, ... Use Monte Carlo simulation to estimate (*i*) the pdf of X; (*ii*) the expected amount of money you will have after 365 days; (*iii*) the probability that you are a "winner" (i.e., have more than \$365 dollars) after 365 days. Comment on the value of playing this game as a long-term investment strategy.

Problem 4: [20 points] Autocorrelation is used as a statistical measure of the relationship between a random variable and itself. Consider a stationary time series of random variables $\{X_n\}$, where $n = 0, \ldots, \infty$, in discrete time. The autocorrelation function (ACF) $\rho_X(k)$ shows the value of the correlation coefficient for different time lags $k > 0$:

$$\rho_X(k) = \rho_{X_t, X_{t+k}} = \frac{E[(X_t - \mu)(X_{t+k} - \mu)]}{\delta^2},$$

where μ is the mean and δ^2 is the common variance of $\{X_n\}$. The argument k is called the lag and denotes the time separation between the occurrences X_t and X_{t+k}. The values of $\rho_X(k)$ may range from -1 to 1. If $\rho_X(k) = 0$, then there is no autocorrelation at lag k.

1. Plot the autocorrelation of the data set that is made available on the class web page. Using the one-pass version of the autocorrelation formula plot the autocorrelation functions for intermediate $n = 1000$ and $n = 10000$, with n been the index of elements in the dataset. Compare with the ACF functions of the entire dataset. Comment.
2. Provide the density of the dataset using `cdh`, as well as its mean and C.V.
3. Do trace-driven simulation and plot the wait times of a single serve queue where the service process is driven by the dataset and the arrival process is exponential. Your plot should provide wait times for utilization levels from 0.1 to 0.9.
4. Now shuffle the data set (you may use *any* shuffling algorithm you want) and repeat steps 1 to 3. Comment on differences and similarities of your results.

–E. Smirni 10-18-2009