EE 325: Probability and Random Processes

Programming Assignment 2

Submission deadline on Moodle: 1159pm, Thursday, 12 September

1. Recall the 'capture-release-recapture' problem: Catch m fish, mark them and release them back into the lake. Allow the fish to mix well and then you catch m fish. Of these p are those that were marked before. Assume that the actual fish population in the lakes is n and has not changed between the catches. Let $P_{m,p}(n)$ be the probability of the event (for a fixed p recatches out of m) coming from n fish in the lake. Generate a plot for $P_{m,p}(n)$ as a function of n for the following values of m and $p : m = 100$ and $p = 10, 20, 50, 75$. For each of these p, use the plots to estimate (educated guess) the actual value of n, i.e., what is the best guess for n if $m = 100$ and you catch p of the marked fish after mixing them up. Call these four estimates $\hat{n}_1, \ldots, \hat{n}_4$. You define your notion of "best guess." Do not search, THINK!

2. Consider the following discrete time system. Packets arrive randomly at at router to be transmitted on a link. The time between successive packets are independent geometric random variables with parameter λ. Packet transmission times are also random and have a geometric distribution with parameter μ. Only one packet can be transmitted at any time and packets that arrive during the transmission of a previously arrived packet wait in buffer memory. There is infinte memory and can accommodate any number of packets. This can be simulated as follows. At the beginning of each second, if there is a packet in the buffer, it leaves with probability μ. And a new packet is added to the queue with probability λ. Simulate this queue for 1,000,000 time steps. For $n = 0, \ldots 50$, plot $p(n)$, the fraction of time that there are are n packets in the queue. Also find the time average of the number of packets in the memory. Use $\lambda = 0.3$ and $\mu = 0.4$.

3. Now extend the program from the previous part to simulate 10,000 queues in simultaneously parallel. When you stop the simulation after 100,000 time steps you have 10,000 values for the number of packets in the system. Use this data to plot $p(n)$ the fraction of queues that have n packets in the system and calculate the sample average from this 10,000 samples.

4. A jury of N members is to be constituted to decide on a complaint. In the population from which the jury has to be selected, each member makes the correct (fair) decision with probability $(0.5 + c)$, $0.05 \leq c \leq 0.25$. The majority rule is applied, i.e., all members vote yes/no and the majority vote is the decision of the panel. If the probability of the correct decision has to be at least 0.75, what combinations of c and N are feasible. Discretize c in steps of 0.01. You can assume N to be odd numbered integers. Repeat if the requirement is to be correct 90% of the time. Provide a short discussion of your findings. Submit the plots for the combinations of c and N for the two correctness requirements.

Note that there is typically a cost involved with the choices of c and N. A juror with a higher c is both rare and expensive. Similarly higher N makes the logistics of managing them complex. Suggest suitable cost functions that depend on c and N and comment on the right combination of c and N for each of the two preceding correctness requirements.