6.262 Discrete Stochastic Processes Wednesday, May 18, 9:00-12:00 noon, 2011
MIT, Spring 2011

## Solutions to final examination

**Problem 1:** A final exam is started at time 0 for a class of $n$ students. Each student is allowed to work until completing the exam. It is known that each student's time to complete the exam is exponentially distributed with density $\mathsf{f}_X(x) = \lambda e^{-\lambda x}$; $x \geq 0$. The times $X_1, \ldots, X_n$ are IID.

**a)** Let $Z$ be the time at which the last student finishes. Show that $Z$ has a distribution function $\mathsf{F}_Z(z)$ given by $[1 - \exp(-\lambda z]^n$.

**b)** Let $T_1$ be the time at which the first student leaves. Show that the probability density of $T_1$ is given by $n\lambda e^{-n\lambda t}$. For each $i$, $2 \leq i \leq n$, let $T_i$ be the interval from the departure of the $i-1$st student to that of the $i$th. Show that the density of each $T_i$ is exponential and find the parameter of that exponential density. Explain why each $T_i$ is independent. Finally note that $Z = \sum_{i=1}^n T_i$.

**Solution 1a):** Note that $Z \leq t$ if and only if $X_i \leq t$ for each $i$, $1 \leq n$, so

$$\Pr\{Z \leq t\} = \prod_{i=1}^n \Pr\{X_i \leq t\} = [1 - \exp(-\lambda t)]^n$$

**Solution 1b):** You can view $T_1$ as the time of the first arrival out of $n$ Poisson processes each of rate $\lambda$. Thus $T_1$ is exponential with parameter $n\lambda$. More directly yet, $T_1 > t$ if and only if $X_i > t$ for $1 \leq i \leq t$, so $\Pr\{T_1 > t\} = [\exp(-\lambda t)]^n = \exp(-n\lambda t)$. The time $T_2$ is the remaining time until the next student out of the remaining $n-1$ finishes. Because of the memorylessness of the exponential distribution, each of these $n-1$ students has an exponential time to go, so $\Pr\{T_2 > t_2\} = \exp(-(n-1)\lambda t_2)$. Each of these times-to-go is independent of $T_1$, so $T_2$ is independent of $T_1$.

In the same way $T_i$ is exponential wih parameter $(n-i+1)\lambda$ and is independent of the earlier $T_i$'s. Finally the remaining time $T_1$ from the next to last to the last student is exponential with parameter $\lambda$. It is curious and somewhat unintuitive that the expected times between departures increases with the longest time $1/\lambda$ for the last (but of course this is a property of exponential rv's rather than real students).

Note also that $Z = \sum_{i=1}^n T_1$ has the very simple distribution given in part a). This was the reason for this problem, to derive the distribution function of a sum of exponential rv's of harmonically related expected values.

**Problem 2:** A Yule process is a continuous time version of a branching process with the special property that the process never decreases. The process starts at time 0 with one organism. This organism splits into two organisms after a time $Y_1$ with the density $\mathsf{f}_{Y_1}(y) = \lambda e^{-\lambda y}$, $y \geq 0$. Each of these two organisms splits into two more organisms after independent exponentially distributed delays, each with this same density $\lambda e^{-\lambda y}$. In general, each old and new organism continues to split forever after a delay $y$ with this same density $\lambda \exp(-\lambda y)$.

**a)** Let $T_1$ be the time at which the first organism splits, and for each $i > 1$, let $T_i$ be the interval from the $i-1$st splitting until the $i$th. Show that $T_i$ is exponential with parameter $i\lambda$ and explain why the $T_i$ are independent.

**b)** For each $n \geq 1$, let the continuous rv $S_n$ be the time at which the $n$th splitting occurs, *i.e.*, $S_n = T_1 + \cdots + T_n$. Find a simple expression for the distribution function of $S_n$. Hint: look carefully at the solution to parts a) and b) of problem 1.

**c)** Let $X(t)$ be the number of organisms at time $t > 0$. Express the distribution function of $X(t)$ for each $t > 0$ in terms of $S_n$ for each $n$. Show that $X(t)$ is a rv for each $t > 0$ (*i.e.*, show that $X(t)$ is finite WP1).

**d)** Find $\mathsf{E}[X(t)]$ for each $t > 0$.

**e)** Is $\{X(t);\, t \geq 0\}$ a countable-state Markov process? Explain carefully. If so describe the embedded Markov chain and identify each state as positive recurrent, null recurrent, or transient.

**f)** Is $\{X(t);\, t \geq 0\}$ an irreducible countable-state Markov process?

**g)** Is $\{S_n;\, n \geq 1\}$ either a martingale or a submartingale?

**h)** Now suppose the births of a Yule process (viewing the original organism as being born at time 0) constitute the input to a queueing process (each birth enters the queue at its time of birth). The queue has a single server and whenever $n$ organisms are in the system (queue plus server), the time to completion of service for the given organism is exponential with rate $\mu_n$. This means that if a new organism enters the queue while the server is busy, the service rate changes according to the new number in the system. Let $Z(t)$ be the number in the system at time $t$.

h(i): Is $\{Z(t);\, t \geq 0\}$ a countable-state Markov process?

h(ii): Is $Z(t)$ a rv for each $t > 0$ no matter what the $\mu_n$ are?

**Solution 2a)** After the $i-1$st splitting, there are $i$ organisms. The time until each splits is exponential with parameter $\lambda$, so the time until the first of them splits is exponential with paramter $i\lambda$. For each organism, the time until the next split is exponential, independent of how it has been since the last split, and thus the time $T_i$ is independent of $T_1, \dots, T_{i-1}$.

**Solution 2b)** $S_n = T_1 + \cdots T_n$ where $T_i$; $1 \le i \le n$ are independent and exponential with parameters $\lambda, 2\lambda, \dots, n\lambda$. This has the same distribution as the rv $Z$ in problem 1, and thus $\Pr\{S_n \le t\} = [1 - e^{-\lambda t}]^n$.

**Solution 2c)** The number of splittings up to time $t$ is one less than the number of organisms at time $t$, so the event $\{S_n \le t\}$ is the same as the event $\{X(t) \ge n+1\} = \{X(t) > n\}$. Thus

$$\Pr\{X(t) > n\} = \Pr\{S_n \le t\} = [1 - e^{-\lambda t}]^n$$

For each finite $t$, $1 - e^{-\lambda t} < 1$, so $\lim_{n\to\infty}[1 - e^{-\lambda t}]^n = 0$. Since $X(t)$ is nonnegative and its complementary distribution function goes to 0 with increasing $n$, it is a rv.

**Solution 2d)** By integrating the complementary distribution function over $n$ for a given $t$, we get

$$\mathsf{E}[X(t)] \;=\; 1 + [1 - e^{-\lambda t}] + [1 - e^{-\lambda t}]^2 + \cdots \;=\; \frac{1}{1 - [1 - e^{-\lambda t}]} \;=\; e^{\lambda t}$$

**Solution 2e)** The sample space for $\{X(t);\, t \ge 0\}$ is countable (*i.e.*, the positive integers). Given that $X(t) = n$ for given $t$, the time to the next state change is exponential with parameter $(n-1)\lambda$, independent of the states at all previous times. Thus $\{X(t);\, t \ge 0\}$ is a countable state Markov process.

The embedded chain for $\{X(t);\, t \ge 0\}$ has a single transition of probability 1 from each state $n$ to $n+1$. All states of the embedded chain are thus transient.

**Solution 2f)** This Markov process is not irreducible since not all states communicate with each other. In fact no two states communicate.

**Solution 2g)** The sequence $\{S_n;\, n \ge 1\}$ is a submartingale since $\mathsf{E}[|S_n|] < \infty$ for all $n \ge 1$ and $\mathsf{E}[S_n \mid S_{n-1}, \dots, S_1] \ge S_{n-1}$. This follows because $S_n = S_{n-1} + T_n$ where $T_n \ge 0$ is the time from the $n-1$st splitting until the $n$th. This fact does not appear to be very useful.

**Solution 2h)** There were two possible interpretations here. In the first, which most people followed, an organism dies at the end of its service, so that the state of the queue equals the number of organisms. In this case, the transition rate from state $i$ to $i+1$ is $i\lambda$, transition times are exponential, and we have a Markov process. The state space for $\{Z(t);\, t \ge 0\}$ is non-negative valued and thus countable. Also $\lim_{n\to\infty} \Pr\{Z(t) > n\} \le \Pr\{X(t) > n\}$ so $Z(t)$ is a rv for each $t$. For the other interpretation, organisms continue to split after service, and then the number $i$ of organisms in the queue no longer determines the rate of transitions from 'state' $i$ to $i+1$ and the process is not Markov.

The first interpretation is of interest since it allows us to generate a family of Markov processes for which the 'steady state' $p_i$ exist but $\sum p_i\nu_i = \infty$ and the state is always finite WP1.

**Problem 3:** A random walk $\{S_n;\, n \geq 1\}$, with $S_n = \sum_{i=1}^{n} X_i$ , has the following probability density for each $X_i$

$$\mathsf{f}_X(x) = \begin{cases} \frac{e^{-x}}{e-e^{-1}} \,; & -1 \leq x \leq 1 \\ \\ = 0 \,; & \text{elsewhere.} \end{cases}$$

**a)** Find the values of $r$ for which $g(r) = \mathsf{E}\left[\exp(rX)\right] = 1$. Hint: these values turn out to be integers.

**b)** Let $P_\alpha$ be the probability that the random walk ever crosses a threshold at some given $\alpha > 0$. Use the Wald identity to find an upper bound to $P_\alpha$ of the form $P_\alpha \leq e^{-\alpha A}$ where $A$ is a constant that does not depend on $\alpha$. Evaluate $A$. Hint: you may assume that the Wald identity applies to a single threshold at $\alpha > 0$ without any lower threshold or assume another threshold at some $\beta < 0$.

**c)** Use the Wald identity to find a lower bound to $P_\alpha$ of the form $P_\alpha \geq Be^{-\alpha A}$ where $A$ is the same as in part b) and $B > 0$ is a constant that does not depend on $\alpha$. Hint: Keep it simple — you are not being asked to find the tightest possible such bound. If you use 2 thresholds, find your lower bound in the limit $\beta \to -\infty$.

**Solution 3a:** We integrate to find $g(r)$,

$$g(r) \;=\; \int_{-1}^{1} \frac{e^{-x}e^{rx}}{e - e^{-1}}\,dx \;=\; \frac{e^{r-1} - e^{-(r-1)}}{(r-1)(e - e^{-1})}$$

We know that $g(0) = 1$ for any rv, and using the hint, we see pretty easily that $g(2) = 1$. Since $g(r)$ is convex, it can have at most 2 solutions to $g(r) = 1$, so $r = 0$ and $r = 2$ are the only solutions. Note that $r = 2$ is the $r^*$ of all the exponential bounds on a positive threshold crossing of a rv with negative expectation.

**Solution 3b:** Consider a stopping rule where stopping occurs when $S_n$ first exceeds $\alpha$. The Wald identity then says that $\mathsf{E}\left[\exp(rS_J - J\gamma(r))\right] = 1$. This applies for any $r$ such that $\gamma(r) = \ln(g(r)$ exists. In particular, at $r = 2$, $\gamma(r) = 0$ so the Wald identity reduces to $\mathsf{E}\left[\exp(2S_J\right] = 1$. We can write this as

$$\Pr\{S_J \geq \alpha\}\,\mathsf{E}\left[\exp(2S_J) \mid S_J \geq \alpha\right] = 1. \tag{1}$$

Lower bounding $S_J$ by $\alpha$ for the conditioning above,

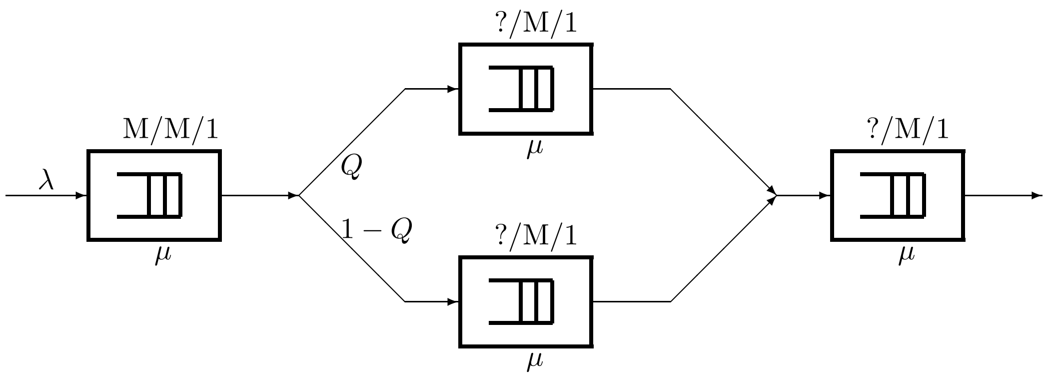
$$\Pr\{S_J \geq \alpha\} \leq e^{-2\alpha}$$

**Solution 3c:** The simplest lower bound comes from observing that, since $X_i \in [-1, +1]$, it is not possible for $S_J$ to exceed $\alpha$ by more than 1. Thus upper bounding $S_J$ by $\alpha + 1$, we get

$$\Pr\{S_J \geq \alpha\} \geq \exp(-2(\alpha + 1) = e^{-2}e^{-2\alpha}$$

**Problem 4:** A queueing system has four queues in the configuration shown. Each queue is identical, with a single server with IID exponential service times, each with density $\mu e^{-\mu x}$. The service times are IID both within each queue and between each queue.

The left most queue (queue 1, say) is $M/M/1$ with an input which is a Poisson process of rate $\lambda < \mu$. Assume that this process started at time $-\infty$. The inputs to the other queues are indicated in the figure below. More specifically, each output from queue 1 is switched to one of the intermediate queues, say queues 2 and 3. Each output goes to queue 2 with probability Q and to queue 3 with probabiity 1-Q. These switching decisions are independent of the inputs and outputs from queue 1. The outputs from queues 2 and 3 are them combined and pass into queue 4.




**a)** Describe the output process from queue 1. That is, describe whether it is a renewal process, and if so, whether it is a Poisson process. At what rate do customers leave queue 1? What can you say about the relation between the outputs and the inputs of queue 1?

**b)** Describe the input processes to queues 2 and 3. Are they Poisson, and if so, of what rate? What is the relationship between the input process to queue 2 and that to queue 3?

**c)** Describe the output processes from queues 2 and 3. Are they Poisson, and what is the relationship between them? What is the rate at which customers leave each of these queues?

**d)** Describe the input and output process for queue 4 (whether they are Poisson, what their rates are, and whether they are independent)?

**e)** Find the expected delay through the entire system for those customers going through queue 2.

**Solution 4a:** From Burke's theorem, the output process from queue 1 is a Poisson process of rate $\lambda$. For any $t$, the output epochs before $t$ are independent of the input epochs after $t$. Similarly, by reversibility, the output epochs after $t$ are independent of the input epochs before $t$.

**Solution 4b:** The switching simply splits one Poisson process into two independent Poisson processes. Thus the input to queue 2 is a Poisson process of rate $Q\lambda$ and the input process to queue 3 is a Poisson process of rate $(1-Q)\lambda$.

**Solution 4c:** By Burke's theorem, again, the outputs from Queue 2 and that from Queue 3 are Poisson. The two are independent of each other since the inputs are independent.

**Solution 4d:** When these two independent Poisson processes from queues 2 and 3 are combined, we get a Poisson process of rate $\lambda$. Thus Queue 4 is also an M/M/1 queue and its output is Poisson at rate $\lambda$ and the input and output bear the same relationship as that in queue 1.

**Solution 4e:** The expected delay is the sum of the expected delays through the 3 queues, *i.e.*, $\frac{1}{\mu-\lambda} + \frac{1}{\mu-Q\lambda} + \frac{1}{\mu-\lambda}$. One can go one step further here and claim that the 3 waiting times are independent of each other.

6.262 Discrete Stochastic Processes
Spring 2011