MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Fall 2007 — 6.436J/15.085J

Final exam, 1:30–4:30pm, (180 mins/100 pts) — 12/19/07

Problem 1: (24 points)
During the time interval $[0, t]$, men and women arrive according to independent Poisson processes with parameters λ_1 and λ_2, respectively. With the exception of part (e), just provide answers (possibly based on your intuitive understanding)—justifications are not required.

(a) (3 pts.) Let $[a, b]$ be an interval contained in $[0, t]$. Give a formula for the probability that the total number of male arrivals during the interval $[a, b]$ is equal to 7.

(b) Out of all the people who arrived during $[0, t]$, we select one at random, with each one being equally likely to be selected.

- (i) (3 pts.) Write an expression for the probability that the selected person is male.
- (ii) (3 pts.) Suppose that the randomly selected person tells us that he/she arrived at a particular time τ. What is the conditional probability that this person is male?
- (iii) (3 pts.) Write an expression (as simple as you can) for the expected time at which the selected person arrived.

(c) (4 pts.) Suppose that $0 < a < b < t$. Let N_1 be the number of male arrivals during $[0, b]$. Let N_2 be the number of female arrivals during $[a, t]$. What is the probability mass function of $N_1 + N_2$?

(d) (4 pts.) Suppose that in (c) above we are told that $N_1 + N_2 = 10$. Find the conditional variance of N_1, given this information.

(e) (4 pts.) Find a good approximation for the probability of the event

$$\{\text{the number of arriving men during } [0, t] \text{ is at least } \lambda_1 t\},$$

when t is large, and justify the approximation.

Solution: (a)

$$e^{-\lambda_1(b-a)}\frac{(\lambda_1(b-a))^7}{7!}$$

(b)(i) $\lambda_1/(\lambda_1 + \lambda_2)$
(ii) Since the time a person has arrived is independent of whether he was classified into male or female, the answer is the same as in (i).

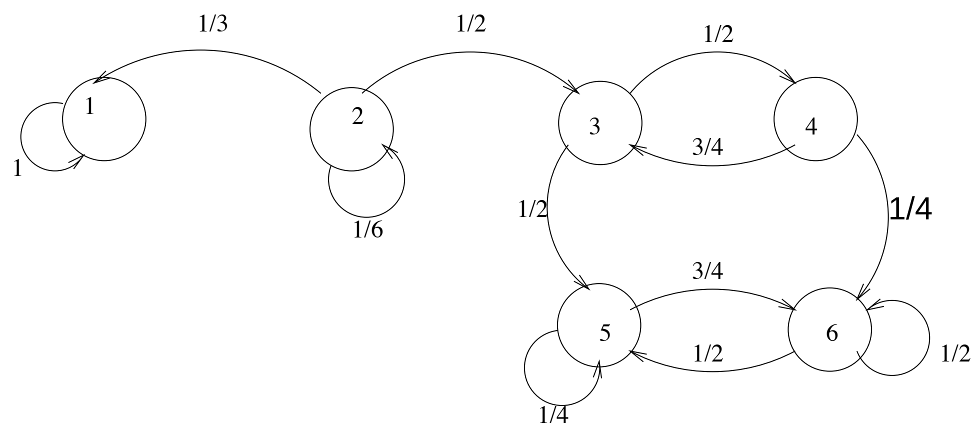
(iii) The distribution of a randomly selected arrival is uniform over $[0, t]$. Its expectation is $t/2$.
(c) Male and female arrivals are independent processes, and the sum of independent Poisson random variables is again Poisson. The answer is Poisson with parameter $\lambda_1 b + \lambda_2(t-a)$.
(d) Since each arrival $N_1 + N_2$ independently comes from N_1 with probability

$$p = \frac{\lambda_1 b}{\lambda_1 b + \lambda_2(t-a)},$$

the distribution of N_1 conditional on $N_1 + N_2 = 10$ is binomial with parameters $n = 10$ and p. Its variance is $10p(1-p)$.
(e) Suppose t is integer. Then, $N([0,t)) = \sum_{i=0}^{t-1} N([i, i+1)]$. The random variables $N([i, i+1))$ are iid with finite variance of λ_1. Then, applying the central limit theorem approximation, $N([0,t)) \approx N(\lambda(t-1), \lambda(t-1))$ and the probability that it is above its mean is approximately $1/2$, by the symmetry of the normal distribution.

If t is not an integer we can make a similar argument by defining $\Delta t = t/\lfloor t \rfloor$, where $\lfloor t \rfloor$ is the largest integer smaller than t. Then Δt is between 1 and 2, and $N([0,t)) = \sum_{i=0}^{\lfloor t \rfloor} N([i\Delta t, (i+1)\Delta t))$, and the same argument as before applies.

Problem 2: (23 points)
Consider the discrete-time Markov chain shown in the figure.



(a) (3 pts.) What are the recurrent classes?

(b) (5 pts.) Assume that $X_0 = 2$. For each recurrent class, compute the probability that the process eventually enters this class.

(c) (5 pts.) Find $\lim_{n\to\infty} \mathbf{P}(X_n = 5 \mid X_0 = 2)$.

(d) (5 pts.) Given that $X_0 = 2$, find the expected time until a recurrent state is reached.

(e) (5 pts.) Find the probability $\mathbf{P}(X_{n-1} = 5 \mid X_n = 5)$, in the limit of large n.

Solution:

(a) The recurrent classes are $\{1\}$ and $\{5, 6\}$.

(b) Let a_i denote the probability of absorption into State 1 starting from state i. Then, it is clear that $a_1 = 1$ and $a_3 = 0$. We also have

$$a_2 = \frac{1}{3}a_1 + \frac{1}{6}a_2 + \frac{1}{2}a_3 = \frac{1}{3} + \frac{1}{6}a_2.$$

It follows that $a_2 = 2/5$. This also implies that the probability of absorption in the class $\{5, 6\}$ starting from state 2 is $3/5$.

(c) With probability $3/5$ the chain enters the recurrent class $\{5, 6\}$. Once in the class, $P(X_n = 5)$ will approach π_5 of the Markov chain composed only of states 5 and 6, which we can determine with the aid of the birth-death equation $\pi_5(3/4) = \pi_6(1/2)$, which yield $\pi_5 = 2/5$. The final answer is $(3/5) \cdot (2/5) = 6/25$.

(d) Let t_i be the expected time to enter a recurrent class conditioned on being at state i. Then the equations are:

$$\begin{aligned} t_1 = t_5 = t_6 &= 0 \\ t_2 &= 1 + \frac{1}{3}t_1 + \frac{1}{6}t_2 + \frac{1}{2}t_3 \\ t_3 &= 1 + \frac{1}{2}t_4 + \frac{1}{2}t_5 \\ t_4 &= 1 + \frac{3}{4}t_3 + \frac{1}{4}t_6 \end{aligned}$$

which have the solution of $t_2 = 66/25$.

(e)

$$\lim_n P(X_{n-1} = 5 | X_n = 5) = \lim_n \frac{P(X_{n-1} = 5)P(X_n = 5 | X_{n-1} = 5)}{P(X_n = 5)} = \frac{\pi_5(1/4)}{\pi_5} = \frac{1}{4}.$$

Problem 3: (13 points)
The number of people that enter a pizzeria in a period of 15 minutes is a (nonnegative integer) random variable K with known moment generating function $M_K(s) = \mathbf{E}[e^{sK}]$. Each person who comes in buys a pizza. There are n types of pizzas, and each person is equally likely to choose any type of pizza, independently of what anyone else chooses.

Give a formula, in terms of $M_K(\cdot)$, for the expected number of different types of pizzas ordered.

Solution: Let D be the number of types of pizza the chef has to prepare, and let M be the number of people to enter the pizzeria. Let $X_1, \ldots, X_n$ be the respective indicator variables of each pizza. Thus if at least one person orders pizza type i, then $X_i = 1$, otherwise $X_i = 0$. Note that $D = X_1 + \cdots + X_n$. Thus we have:

$$\begin{aligned}
\mathbf{E}[D] &= \mathbf{E}[\mathbf{E}[D|M]] \\
&= \mathbf{E}[\mathbf{E}[X_1 + \cdots + X_n|M]] \\
&= n \cdot \mathbf{E}[\mathbf{E}[X_i|M] \\
&= n \cdot \mathbf{E}\Big[1 - \Big(\frac{n-1}{n}\Big)^M\Big] \\
&= n - n \cdot \mathbf{E}\Big[\Big(\frac{n-1}{n}\Big)^M\Big] \\
\text{(letting } s = \log((n-1)/n)) &= n - n \cdot \mathbf{E}[e^{sM}] \\
&= n - n \cdot M_K\big(\log((n-1)/n)\big).
\end{aligned}$$

Problem 4: (13 points)
Let S be the set of arrival times in a Poisson process on $\mathbb{R}$ (i.e., a process that has been running forever), with rate λ. Each arrival time in S is displaced by a random amount. The random displacement associated with each element of S is a random variable that takes values in a finite set. We assume that the random displacements associated with different arrivals are independent and identically distributed. Show that the resulting process (i.e., the process whose arrival times are the displaced points) is a Poisson process with rate λ. (We expect a proof consisting of a verbal argument, using known properties of Poisson processes; formulas are not needed.)

Solution: Since each point is perturbed independently of all the others, we can consider the perturbation as follows: Let the perturbation values be $\{v_1, \ldots, v_m\}$, which occur with respective probabilities p_i. Then consider splitting our d-dimensional Poisson process according to the probabilities p_i, into processes $N_1, N_2, \ldots, N_m$. By the results on splitting Poisson processes, these will be independent, and process N_i will be a Poisson process on $\mathbb{R}$ of rate λp_i Now shift all the points split into the i^{th} process, N_i, by the same value v_i. Clearly this translated version of N_i is a Poisson process, and it is independent of all the other Poisson processes, and hence independent of the translated versions as well. Therefore, once we merge the shifted processes, again we have a Poisson process of the same rate.

Problem 5: (10 points)
Let $\{X_n\}$ be a sequence of nonnegative random variables such that $\lim_{n\to\infty} X_n = 0$, almost surely. For the following statements, answer (together with a brief justification) whether it is: (i) always true, (ii) always false; (iii) sometimes true and sometimes false,

(a) (5 pts.) $\lim_{n\to\infty} \mathbf{P}(X_n > 0) = 0$.

(b) (5 pts.) For every $\epsilon > 0$, $\lim_{n\to\infty} \mathbf{P}(X_n > \epsilon) = 0$.

Solution: part (a) is sometimes true, sometimes false. Consider $X_n = 1/n$ with probability 1, in which case (a) is false; and consider $X_n = 0$ with probability 1, in which case (a) is true.

Part (b) is always true because $P(X_n > \epsilon) \leq P(|X_n - 0| > \epsilon)$, and since convergence almost everywhere implies convergence in probability, the last expression approaches 0 as n goes to infinity.

Problem 6: (17 points)
Let $\{X_n\}$ be a sequence of random variables defined on the same probability space.

(a) (4 pts.) Suppose that $\lim_{n\to\infty} \mathbf{E}[|X_n|] = 0$. Show that X_n converges to zero, in probability.

(b) (5 pts.) Suppose that X_n converges to zero, in probability, and that for some constant c, we have $|X_n| \leq c$, for all n, with probability 1. Show that

$$\lim_{n\to\infty} \mathbf{E}[|X_n|] = 0.$$

(c) Suppose that each X_n can only take the values 0 and 1 and, that $\mathbf{P}(X_n = 1) = 1/n$.

(i) (4 pts.) Given an example in which we **have** almost sure convergence of X_n to 0.

(ii) (4 pts.) Given an example in which we **do not have** almost sure convergence of X_n to 0.

Solution: For part (a), observe that Markov's inequality implies

$$P(|X_n - 0| \geq \epsilon) \leq \frac{\mathbf{E}[|X_n|]}{\epsilon},$$

so that if $\mathbf{E}[|X_n|]$ approaches 0, we have that X_n approaches 0 in probability.

For part (b), fix $\epsilon > 0$ and define a new random variable X_n^ϵ as follows. We have $X_n^\epsilon = \epsilon$ whenever $|X_n| \leq \epsilon$, and $X_n^\epsilon = c$ whenever $|X_n| > \epsilon$. Then, it is always true that $|X_n| \leq X_n^\epsilon$ and therefore

$$\mathbf{E}[|X_n|] \leq \mathbf{E}[X_n^\epsilon] = \epsilon P(|X_n| \leq \epsilon) + cP(|X_n| > \epsilon)$$

Taking limits as n goes to infinity, we get

$$\lim_n \mathbf{E}[|X_n|] \leq \epsilon,$$

and since this holds for all $\epsilon > 0$, we get $\lim_n \mathbf{E}[|X_n|] = 0$.

For part (c).i, we can generate a uniform random variable on $[0, 1]$ and declare that $X_n = 1$ if the outcome is at most $1/n$, and 0 otherwise. It immediately follows X_n

is binary with $P(X_n = 1) = 1/n$. Now suppose that the uniform random variable generated the value x with $x > 0$. Then eventually $1/n$ is smaller than x, and $X_n = 0$ after this point. Since the outcome is positive with probability 1 (the probability of getting 0 is 0), it follows that X_n approaches 0 almost surely.

For part (c).ii, we can take X_n to be independent. Since

$$\sum_{i=1}^{\infty} P(X_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

and the events $\{X_n = 1\}$ are independent, the Borel-Cantelli lemma implies that $X_n = 1$ occurs infinitely often with probability 1. It follows that X_n does not converge to 0 almost surely.

6.436J / 15.085J Fundamentals of Probability
Fall 2008