Solution to 6.262 Final Examination 5/21/2009

Solution to Question 1

a) We first solve the steady state equations for the Markov process. As we have seen many times, the p_i, $i \geq 0$ for a birth-death chain are related by $p_{i+1}q_{i+1,i} = p_i q_{i,i+1}$, which in this case is $p_{i+1}\mu = p_i\lambda/(i+1)$. Iterating this equation,

$$p_i = p_{i-1}\frac{\lambda}{\mu i} = p_{i-2}\,\frac{\lambda^2}{\mu^2 i(i-1)} = \cdots = p_0\,\frac{\lambda^i}{\mu^i i!}.$$

Denoting λ/μ by ρ,

$$1 = \sum_{i=0}^{\infty} p_i = p_0\left[\sum_{i=0}^{\infty}\frac{\rho^i}{i!}\right] = p_0 e^{\rho}.$$

Thus,

$$p_0 = e^{-\rho}; \qquad p_i = \frac{\rho^i e^{-\rho}}{i!}.$$

This is a solution to the steady state equations, and we know (Theorem 6.2) that if $\sum_i p_i\nu_i < \infty$, where the ν_i are given in part b, then this solution is unique and the process is positive recurrent. Part b) also shows that $\sum_i p_i\nu_i < \infty$.

To explain p_i as a time average more carefully, each sample point ω of the underlying sample space gives rise to a sample path $\{x(t);\, t \geq 0\}$ of the process $\{X(t);\, t \geq 0\}$. For all of these sample paths except a set of probability 0, p_i is the limiting fraction of time that the process is in state i. That is, for a sample path $x(t)$, let $R_i(t) = 1$ for t such that $x(t) = i$ and let $R_i(t) = 0$ otherwise. Then p_i is the time-average fraction of time in state i if $\lim_{t\to\infty}(1/t)\int_0^t R_i(\tau) = p_i$ for all sample paths except a set of probability zero.

Most students confused time average with limiting ensemble average or limiting time and ensemble average.

b) For the embedded chain, $P_{01} = 1$, and for all $i > 0$,

$$P_{i,i+1} = \frac{\lambda}{\lambda + \mu(i+1)}; \qquad P_{i,i-1} = \frac{\mu(i+1)}{\lambda + \mu(i+1)}.$$

All other transition probabilities are 0. The departure rate from state i is

$$\nu_0 = \lambda; \qquad \nu_i = \mu + \frac{\lambda}{i+1} \quad \text{for all } i > 0.$$

The steady state probabilities π_i for the embedded chain are quite messy; sorry for that. They can be found directly from the transition probabilities, but it is quite a bit easier to calculate them as $\pi_i = \frac{p_i\nu_i}{\sum_{j=0}^{\infty} p_j\nu_j}$. First we calculate $\sum_j p_j\nu_j$ by separating out the $j = 0$ term and then summing separately over the two terms, μ and $\lambda/(j+1)$, of ν_j.

$$\sum_{j=0}^{\infty} p_j\nu_j = e^{-\rho}\lambda + \sum_{j=1}^{\infty} e^{-\rho}\frac{\rho^j\mu}{j!} + \sum_{j=1}^{\infty} e^{-\rho}\frac{\rho^j\lambda}{j!(j+1)}.$$

Substituting $\mu\rho$ for λ and combining the first and third term,

$$\begin{aligned}\sum_{j=0}^{\infty} p_j\nu_j &= \sum_{j=1}^{\infty} e^{-\rho}\frac{\rho^j\mu}{j!} + \sum_{j=0}^{\infty} e^{-\rho}\frac{\rho^{j+1}\mu}{(j+1)!} \\ &= 2\sum_{j=1}^{\infty} e^{-\rho}\frac{\rho^j\mu}{j!} = 2\mu(1-e^{-\rho}).\end{aligned}$$

The steady state probabilities for the embedded chain are then given by

$$\pi_0 = \frac{\rho}{2(e^\rho-1)}; \qquad \pi_i = \frac{\rho^i}{2i!(e^\rho-1)}\left(\frac{\rho}{i+1}+1\right); \quad \text{for } i>1.$$

There are many forms for this answer. One sanity check is to observe that the embedded chain probabilities do not change if λ and μ are both multiplied by the same constant, and thus the π_i must be a function of ρ alone. Another sanity check is to observe that in the limit $\rho \to 0$, the embedded chain is dominated by an alternation between states 0 and 1, so that in this limit $\pi_0 = \pi_1 = 1/2$.

c) The time-average interval between visits to state j is $\overline{W}_j = 1/(p_j\nu_j)$. This is explained in detail in section 6.2.3 of the class notes, but the essence of the result is that for renewals at successive entries to state j, p_j must be the ratio of the expected time $1/\nu_j$ spent in state j to the overall expected renewal interval $\overline{W}_j$. Thus $\overline{W}_j = 1/(\nu_j\rho_j)$.

$$\overline{W}_0 = \frac{e^\rho}{\lambda}; \qquad \overline{W}_j = \frac{(j+1)!\,e^\rho}{\rho^j[\lambda+(j+1)\mu]}.$$

The time-average number of state transitions per visit to state j is $\overline{T}_{jj} = 1/\pi_j$. This is proven in Chapter 5, but the more interesting way to see it here is to use the same argument as used for $\overline{W}_j$ above. That is, consider the embedded Markov chain as a discrete-time process with one unit of time per transition. Then π_j is the ratio of the unit time spent on a visit to j to the expected number of transitions per visit to j. Thus

$$\overline{T}_{00} = \frac{2(e^\rho-1)}{\rho}; \qquad \overline{T}_{jj} = \frac{2(j+1)!\,(e^\rho-1)}{\rho^j(\rho+j+1)} \quad \text{for } j \geq .$$

d) On each successive entry to state 0, the embedded chain is in state 0 and the duration in that state is a uniform rv independent of everything except the fact of its existence due to entering state 0. This duration in state 0 is finite and IID between successive visits to 0. All subsequent states and durations are independent of those after each subsequent visit to state 0 and the time to return to state 0 is finite with probability 1 from the same argument used for the Markov process. Thus successive returns to state 0 form a renewal process. The same argument works for exits from state 0.

For successive entries to some arbitrary given state j, the argument is slightly more complex, but it also forces us to understand the issue better. The argument that the epochs of successive entries to a given state in a Markov process form a renewal process does not depend at all on the exponential holding times. It depends only on the independence

of the holding times and their expected values. Thus changing one holding time from an exponential to a uniform rv of the same expected value changes neither the expected number of transitions per unit time nor the expected time between entries to a given state. Thus successive visits to any given state form renewal epochs.

e) As explained in part d, the expected time between returns to any given state j is the same as in part c. Since the embedded Markov chain has not changed at all, the expected number of transisions between visits to any given state j is also given in part c.

f) The modified process is not a Markov process. Essentially, the fact that the holding time in state 0 is no longer exponential, and thus not memoryless, means that if $X(t) = 0$, the time at which state 0 is entered provides information beyond knowing $X(\tau)$ at some given $\tau < t$

As an example, suppose that $\mu << \lambda$ so we can ignore multiple visits to state 0 within an interval of length $2/\lambda$. We then see that $\mathrm{P}\{X(t) = 0 \mid X(t - 1/\lambda) = 0, X(t - 2/\lambda) = 0\} = 0$ since there is 0 probability of a visit to state 0 lasting for more than time $2/\lambda$. On the other hand,

$$\mathrm{P}\{X(t) = 0 \mid X(t - 1/\lambda) = 0, X(t - 2/\lambda = 1\} > 0,$$

In fact, this latter probability is $1/8$ in the limit of large t and small μ.

Solution to Question 2

a) First consider the case where $S_N \geq \alpha$. As supplied by the hint, conditioning on $S_N \geq \alpha$ implies that $S_{N-1} < \alpha$ and thus $S_N = S_{N-1} + Y_N$ (i.e. $I_N = 1$). Now let Z be the amount by which $S_{N-1} + Y_N$ exceeds the threshold, that is,

$$|S_N| = \alpha + Z\,, \quad \text{where} \quad Z = |S_N| - \alpha = |S_{N-1} + Y_N| - \alpha.$$

Let $S_{N-1} = s$ for some $s \in (-\alpha, \alpha)$ and note that

$$\begin{aligned} \mathrm{P}(Z \geq z \mid S_{N-1} = s, S_N \geq \alpha) &= \mathrm{P}(s + Y_N - \alpha \geq z \mid S_{N-1} = s, S_N \geq \alpha) \\ &= \mathrm{P}(Y_N \geq z - s + \alpha \mid S_{N-1} = s, S_N \geq \alpha). \end{aligned}$$

But, for any $s \in (-\alpha, \alpha)$,

$$\{S_{N-1} = s, S_N \geq \alpha\} = \{S_{N-1} = s, Y_N \geq \alpha - s\}.$$

Furthermore, Y_N is independent of $S_{N-1} = Y_1 + \ldots + Y_{N-1}$. Thus,

$$\mathrm{P}(Z \geq z \mid S_{N-1} = s, S_N \geq \alpha) = \mathrm{P}(Y_N \geq z - s + \alpha \mid Y_N \geq \alpha - s) = e^{-\lambda z},$$

where the last equality follows by the fact that Y_N is exponentially distributed with rate λ. (Recall the memoryless property of the exponential distribution.) In other words, conditioned on $S_N \geq \alpha$ and $S_{N-1} = s$, the overshoot Z is exponentially distributed with rate α. Since the result holds for any $s \in (-\alpha, \alpha)$, it follows that conditioned on $S_N \geq \alpha$, Z is exponentially distributed with rate λ.

Similarly, for $S_N \leq -\alpha$, we write $S_N = -\alpha - Z$, where Z is distributed as above.

b) Let $q_\alpha = \mathrm{P}(S_N \geq \alpha)$. It follows that $\mathrm{P}(S_N \leq -\alpha) = 1 - q_\alpha$. By the Total Expectation Theorem from 6.041,

$$\begin{aligned} \mathrm{E}(e^{rS_N}) &= \mathrm{E}(e^{rS_N} \mid S_N \geq \alpha) q_\alpha + \mathrm{E}(e^{rS_N} \mid S_N \leq -\alpha)(1 - q_\alpha) \\ &= \mathrm{E}(e^{r(\alpha + Z)}) q_\alpha + \mathrm{E}(e^{r(-\alpha - Z)})(1 - q_\alpha) \\ &= e^{r\alpha} \frac{\lambda}{\lambda - r} q_\alpha + e^{-r\alpha} \frac{\lambda}{\lambda + r}(1 - q_\alpha), \end{aligned}$$

where $\frac{\lambda}{\lambda - r}$ is the MGF associated with Z and valid for $r < \lambda$, and $\frac{\lambda}{\lambda + r}$ is the MGF associated with $-Z$ and valid for $r > -\lambda$. It follows that the above expression for $\mathrm{E}(e^{rS_N})$ is valid for $|r| < \lambda$.

c) By the Total Expectation Theorem, $\mathrm{E}(e^{rI_1Y_1}) = \mathrm{E}(e^{rI_1Y_1} \mid I_1 = 1)p + \mathrm{E}(e^{rI_1Y_1} \mid I_1 = -1)(1 - p)$. Since I_1 is independent of Y_1, it follows that

$$\mathrm{E}(e^{rI_1Y_1}) = \mathrm{E}(e^{rY_1})p + \mathrm{E}(e^{-rY_1})(1 - p) = p\frac{\lambda}{\lambda - r} + (1 - p)\frac{\lambda}{\lambda + r},$$

from which we obtain that $\mathrm{E}(e^{rI_1Y_1}) = 1 \iff r \in \{0, \lambda(1 - 2p)\}$. Note that for $p \in (0, 1/2)$, the two values are distinct and satisfy $|r| < \lambda$.

d) For $p < 1/2$, part c) yields that for $r^* = \lambda(1-2p)$, $g_{I_1Y_1}(r^*) = 1$. By Lemma 7.1 in the notes, N is a stopping rule and applying Wald's identity yields $\mathrm{E}(e^{r^*S_N}) = 1$. It then follows from b) that

$$1 = e^{r^*\alpha}\frac{\lambda}{\lambda - r^*}q_\alpha + e^{-r^*\alpha}\frac{\lambda}{\lambda + r^*}(1-q_\alpha),$$

and thus

$$q_\alpha = \frac{1 - e^{-r^*\alpha}\frac{\lambda}{\lambda+r^*}}{e^{r^*\alpha}\frac{\lambda}{\lambda-r^*} - e^{-r^*\alpha}\frac{\lambda}{\lambda+r^*}}.$$

e) Note that $\mathrm{E}(S_N) = q_\alpha(\alpha+1/\lambda) + (-\alpha-1/\lambda)(1-q_\alpha) = (\alpha+1/\lambda)(2q_\alpha - 1)$ and $\mathrm{E}(I_1Y_1) = p/\lambda - (1-p)/\lambda = (2p-1)/\lambda$. By Wald's equality, $\mathrm{E}(N)\mathrm{E}(I_1Y_1) = \mathrm{E}(S_N)$, and therefore

$$\mathrm{E}(N) = \frac{(\alpha+1/\lambda)(2q_\alpha-1)}{\mathrm{E}(I_1Y_1)} = \frac{(\alpha+1/\lambda)(2q_\alpha-1)}{(2p-1)/\lambda} = \frac{(\alpha\lambda+1)(2q_\alpha-1)}{2p-1},$$

where q_α was found in part d).

f) Here, it suffices to note that by the symmetry in the problem, $\mathrm{E}(N)$ for $p > 1/2$ must be the same as $\mathrm{E}(N)$ for $p < 1/2$ and is therefore given as in d). Another way to arrive at this conclusion is to replace all p by $1-p$ and all q_α by $1-q_\alpha$ in the expression in d).

g) For $p = 0$, $q_\alpha = 0$ and $\mathrm{E}(S_N) = -\alpha - 1/\lambda$. By Wald's equality, $\mathrm{E}(N)(-1/\lambda) = -\alpha - 1/\lambda$. Thus, $\mathrm{E}(N) = \alpha\lambda + 1$. The same answer applies to $p = 1$ by symmetry.

Note that the $p \in \{0,1\}$ scenario reduces to a Poisson process with rate λ. In particular, $\mathrm{E}(N)$ can be computed without any recourse to Wald's equality or even stopping rules. It suffices to note that $N = n$ if and only if there are exactly $n-1$ Poisson arrivals in the segment $[0,\alpha]$. Letting K be the number of Poisson arrivals in $[0,\alpha]$, it follows that

$$\mathrm{E}(N) = \sum_{n=0}^{\infty} n\mathrm{P}(N=n) = \sum_{n=0}^{\infty} n\mathrm{P}(K=n-1) = \sum_{k=1}^{\infty}(k+1)\mathrm{P}(K=k) = \mathrm{E}(K)+1.$$

Since K has the Poisson distribution with mean $\alpha\lambda$, the result follows.

f) (Bonus) From Wald's identity, since $\mathrm{E}(Y_1I_1) = 0$, we have $\mathrm{E}(S_N^2) = \mathrm{E}(N)\mathrm{var}(I_1Y_1)$. By symmetry, $q_\alpha = 1/2$. Thus,

$$\mathrm{E}(S_N^2) = \frac{1}{2}\mathrm{E}((\alpha+Z_\alpha)^2) + \frac{1}{2}\mathrm{E}((-\alpha-Z_{-\alpha})^2) = \mathrm{E}((\alpha+Z_\alpha)^2) = \alpha^2 + 2\frac{\alpha}{\lambda} + \frac{2}{\lambda^2}.$$

Since

$$\mathrm{var}(I_1Y_1) = \mathrm{E}(I_1^2Y_1^2) - 0 = \mathrm{E}(Y_1^2) = \frac{2}{\lambda^2},$$

it follows that

$$\mathrm{E}(N) = \frac{\alpha^2 + 2\frac{\alpha}{\lambda} + \frac{2}{\lambda^2}}{\frac{2}{\lambda^2}} = (\lambda\alpha/2)^2 + \lambda\alpha + 1 = \left(\frac{\lambda\alpha}{2}+1\right)^2.$$

Solution to Question 3

a) We will associate the nth term X_n in the branching process $\{X_n;\, n \geq 0\}$ with the nth state X_n of a Markov chain $\{X_n; n \geq 0\}$. This has a countable state space consisting of the integers $0, 1, \ldots$. The initial state X_0 is 1.

For each $n \geq 1$, $X_{n+1} = \sum_{i=1}^{X_n} Y_{i,n}$. Thus X_{n+1} is the sum of X_n IID rv's $\{Y_{1,n}\}$ which are given to be independent of all rv's $Y_{i,k}$ for $k < n$. Thus, conditional on X_n (which determines the number of elements in the sum), X_{n+1} is independent of all the rv's $Y_{i,k}$ for $k < n$, and thus independent of $X_0, X_1, \ldots, X_{n-1}$.

To show that $\{X_n; n \geq 0\}$ is a Markov chain, it remains to show that the transition probabilities $\mathrm{P}\{X_{n+1} \mid X_n\}$ do not depend on n. However, $X_{n+1} = \sum_{i=1}^{X_n} Y_{i,n}$ is simply the sum of X_n IID rv's whose distribution is the same for all n. Thus

$$P_{ij} = \mathrm{P}\left\{\sum_{\ell=1}^{i} Y_\ell = j\right\},$$

where the Y_ℓ are IID with the same distribution as the $Y_{i,n}$.

b) The function $P_{ij}(n)$ is the probability that the chain reaches state j for the first time in exactly j trials starting in state i (this is called $f_{ij}(n)$ in the text). It follows that $F_{ij}(n) = \sum_{k=1}^{n} P_{ij}(k)$ is the probability that that the chain reaches state j on any of the first n trials. Thus $F_{ij}(\infty)$ is the probability of ever reaching state j, and $F_{i0}(\infty)$ is the probability of ever reaching state 0, *i.e.*, the probability of extinction starting in state i. Since the process is started in state 1, the probability of extinction is $p = F_{10}(\infty)$.

Next note that the process is extinguished immediately at the first trial if $Y_{1,1} = 0$. Thus $P_{10}(1) = \mathrm{P}\{Y = 0\} > 0$. It follows that

$$F_{1,0}(\infty) = \sum_k P_{1,0}(k) > \mathrm{P}\{Y = 0\} > 0$$

Finally, if $X_n = i$ for some n, then the offspring of each of these i individuals will die out with probabiity p. Since the offspring of the different individuals at time n are independent of each other, the probability that all i of these families will die out is

$$F_{i0}(\infty) = F_{10}^i(\infty) = p^i > 0.$$

c) Since $\mathrm{P}\{Y = 0\} > 0$, state 0 is the single member of a recurrent class and all other states are transient. This follows since all states have paths to 0 but no paths back.

For the major case of interest, where $\mathrm{P}\{Y = 1\} > 0$ and $\mathrm{P}\{Y \geq 2\} > 0$, there is a path from state 1 to arbitrarily high numbered states, and for each i, there is a transition from i to $i - 1$. One way this transition occurs is when one of the i individuals dies out and the others all have one descendant. Thus in this case, all the transient states are in a single class.

You were not intended to analyze the classification of the transient states further, but the classification is as follows: If the possible values for Y have a greatest common denominator,

say d, greater than 1, then there is a class of transient states consisting of all multiples of d, and each other transient state lies in a singleton class. If the possible values for Y have a greatest common denominator equal to 1, than all transient states lie in the same class.

d) Let $\mathcal{E}$ be the event that extinction occurs. Thus $\mathrm{P}\{\mathcal{E}\} = p$. Also, from part b,

$$\mathrm{P}\{\mathcal{E} \mid X_n = i\} = F_{i0}(\infty) = p^i$$

Thus

$$\begin{aligned} p &= \mathrm{P}\{\mathcal{E}\} = \sum_{i=0}^{\infty} \mathrm{P}\{X_n = i\}\, \mathrm{P}\{\mathcal{E} \mid X_n = i\} \\ &= \sum_{i=0}^{\infty} \mathrm{P}\{X_n = i\}\, p^i = \mathsf{E}\left[p^{X_n}\right] \end{aligned}$$

This is a nice proof that $\mathsf{E}\left[p^{X_n}\right] = p$ for each n, but how does one go about guessing that the answer is p in order to construct the proof above? The first part, $\mathsf{E}\left[p^{X_n}\right] = \sum_i \mathrm{P}\{X_n{=}i\}\, p^i$ is a straightforward way to start the problem. Then associating p^i with $\mathrm{P}\{\mathcal{E} \mid X_n = i\}$ is a natural association given part b and given some understanding that this association is fundamental to understanding branching processes.

e)

i) Recall that X_n for $n \geq 1$ is the sum of the number of offspring of each individual in X_{n-1}. Thus $\mathsf{E}[X_n \mid X_{n-1}] = \overline{Y} X_{n-1}$. Since $\overline{Y} \neq 1$, we see that $\{X_n;\, n \geq 0\}$ is not a martingale.

ii) Parts ii and iii are very similar, with the constant $\overline{Y}$ used in one and the constant p used in the other, so we first calculate $\mathsf{E}\left[z^{X_n} \mid z^{X_{n-1}}, \ldots, z^{X_0}\right]$ for arbitrary z. The calculation is similar to that in part d.

$$\mathsf{E}\left[z^{X_n} \mid z^{X_{n-1}}, \ldots, z^{X_0}\right] = \mathsf{E}\left[z^{X_n} \mid X_{n-1}\right]$$

Letting j be the sample value of X_{n-1}, $X_n = \sum_{\ell=1}^{j} Y_\ell$ and $z^{X_n} = \prod_{\ell=1}^{j} a^{Y_\ell}$. These rv's are independent, and thus

$$\mathsf{E}\left[z^{X_n} \mid X_{n-1} = j\right] = \prod_{\ell=1}^{j} \mathsf{E}\left[z^{Y_\ell}\right] = \left[\mathsf{E}\left[z^{Y}\right]\right]^j$$

Thus $\{z^{X_n};\, n \geq 0\}$ is a martingale if and only if $z = \mathsf{E}\left[z^Y\right]$. Now $\mathsf{E}\left[z^Y\right] = \sum_i \mathrm{P}\{Y = i\}\, z^i$ is the z-transform, say $h(z)$, of the PDF of Y. It is a convex function of z and thus there can be at most two values of z satisfying this equation. The first is $z = 1$ and another, as can be seen from part d with $n = 1$, is $z = p$. By assumption, $\overline{Y} \neq 1$, and we now show that $\overline{Y} \neq p$. First, since $p \leq 1$, $\overline{Y} \neq p$ if $\overline{Y} > 1$. Second if $\overline{Y} < 1$, then $p = 1$, so in both cases, $\overline{Y} \neq p$. Thus the process in part ii is not a martingale.

iii) It was shown in part ii that $\{p^{X_n};\, n \geq 0\}$ is a martingale.

f (Bonus) If $\overline{Y} = 1$, then $\{X_n;\, n \geq 0\}$ is a martingale (by the fact that the scaled proess $\{X_n/\overline{Y}^n;\, n \geq 0\}$ is a martingale. Thus it is also a submartingale and we can use the Kolmogorov submartingale inequality. This says that for any positive integer m,

$$\mathrm{P}\left\{\max_{0\leq i\leq m} X_i \geq 2\right\} \leq \mathsf{E}\left[X_m\right]/2.$$

Since the X_n also form a martingale $\mathsf{E}\left[X_m\right] = \mathsf{E}\left[X_0\right] = 1$ for all m. Thus we can go to the limit $m \to \infty$, and get the upper bound

$$\mathrm{P}\left\{\sup_{i<\infty} X_i \geq 2\right\} \leq 1/2$$

If Y is binary with equiprobable values 0 and 2, then the process reaches the threshold on trial 1 with probability 1/2 and is extinguished on trial 1 with probability 1/2.

g (Bonus) On trial 1, the process is extinguished with probability q_0, crosses the threshold with probability $q_{\geq 2} = 1 - q_0 - q_1$ and lives to see another day with probability q_1. On trial 2, given $Y = 1$ on the first trial, the same thing happens. The threshold is crossed on trial 2 with marginal probability $q_1 q_{\geq 2}$. Extending the argument forever,

$$\mathrm{P}\left\{\sup_{n\geq 1} X_n \geq 2\right\} = (q_{\geq 2})[1 + q_1 + q_1^2 + \cdots] = \frac{q_{\geq 2}}{1 - q_1} = \frac{1 - q_0 - q_1}{1 - q_1}$$

6.262 Discrete Stochastic Processes
Spring 2011