Applied Probability Qualifying Review Exam Questions
May, 2009
time allowed: approximately 1/2 hr per question

1. Consider a continuous-time Markov chain $\{X(t), t \geq 0\}$ with stationary probabilities $\{P_i, i \geq 0\}$. Denote the transition rates by $\{q_{ij}\}$ and the rate of leaving i by $\nu_i = \sum_{j \neq i} q_{ij}$. Let T denote the first time that the chain has been in state 0 for τ consecutive time units. Find $E[T \mid X(0) = 0]$.

   Solution: Define F to be the first exit time from 0, and R to be the first return time to 0. All the following calculations are conditional on $X(0) = 0$. We find that

   $$\begin{aligned} E[T] &= E\big[E[T \mid F]\big] \\ &= \tau P(F > \tau) + \int_0^\tau E[T|F = x]\nu_0 e^{-\nu_0 x}\, dx. \end{aligned}$$

   Now, $E[R] = (P_0\nu_0)^{-1}$ and $E[F] = \nu_0^{-1}$, so $E[R - F] = \frac{1}{\nu_0}\left(\frac{1}{P_0} - 1\right)$. Making use of the Markov property, we then have

   $$\begin{aligned} E[T] &= \tau e^{-\nu_0\tau} + \int_0^\tau \left(x + \frac{1}{\nu_0}\left(\frac{1}{P_0} - 1\right) + E[T]\right)\nu_0 e^{-\nu_0 x}\, dx. \\ &= \tau e^{-\nu_0\tau} + \left(1 - e^{-\nu_0\tau}\right)\{x + E[T]\} + \int_0^\tau x\nu_0 e^{-\nu_0 x}\, dx. \end{aligned}$$

   Rearranging and integrating by parts,

   $$E[T] = \tau + (e^{\nu_0\tau} - 1)\frac{1}{\nu_0}\left(\frac{1}{P_0} - 1\right) - \tau + \frac{1}{\nu_0}(e^{\nu_0\tau} - 1) \qquad = \frac{1}{\nu_0 P_0}(e^{\nu_0\tau} - 1).$$

2. Let $B(t)$ be a standard Brownian motion started at $b > 0$, i.e., $B(0) = b$ and $B(t) - B(s) \sim N\big(0, |t - s|\big)$. Let T be the first time that $B(t)$ hits zero, i.e., $T = \min\{t : B(t) = 0\}$. Brownian motion with absorption at zero is $Y(t) = B(t \wedge T)$, so $Y(t)$ follows the path of the Brownian motion until the first visit to zero and then stays there forever. Calculate the transition probablity density function $p(x, y, t)$ of $Y(t)$, defined such that

   $$P\Big(y_0 < Y(s + t) < y_1 \mid Y_s = x\Big) = \int_{y_0}^{y_1} p(x, y, t)\, dy$$

   for $0 < y_0 < y_1$.

   Solution: From the reflection principle,

   $$\begin{aligned} P[B(t + s) > y, \min_{s<u<t} B(u) < 0 \mid B(s) = x] &= P[B(t + s) < -y \mid B(s) = x] \\ &= \Phi(-(x + y)/\sqrt{t}). \end{aligned}$$

We can then calculate

$$
\begin{aligned}
&P[B(t+s) > y, \min_{s<u<t} B(u) > 0 \mid B(s) = x] \\
&\quad = P[B(t+s) > y \mid B(s) = x] - P[B(t+s) > y, \min_{s<u<t} B(u) < 0 \mid B(s) = x] \\
&\quad = \Phi((x-y)/\sqrt{(t)}) - \Phi(-(x+y)/\sqrt{t}).
\end{aligned}
$$

The required density is

$$
\begin{aligned}
p(x,y,t) &= -\frac{d}{dy} P[B(t+s) > y, \min_{s<u<t} B(u) > 0 \mid B(s) = x] \\
&= \frac{1}{\sqrt{t}} \Big[\phi((x-y)/\sqrt{t}) - \phi((x+y)/\sqrt{t}). \Big]
\end{aligned}
$$

3. This question studies a simple version of the Wright-Fisher model, which is fundamental to the study of population genetics. Consider a population containing N copies of a gene that can each be one of two types, A or a. (Since humans have two copies of each gene, this would correspond to the N copies of a gene belonging to $N/2$ individuals.) We model the number of genes ot type A in successive generations, supposing that each generation has the same fixed number N of copies of the gene. Specifically, let X_n be the number of genes of type A in the nth generation, so X_n takes values in $\{0, 1, \ldots, N\}$. Suppose a model of reproduction in which the population of genes at time $n+1$ is obtained by drawing N times with replacement from the population at time n. Thus, X_n is a Markov chain with transition probabilities given by

$$
p(i,j) = \binom{N}{j} \Big(\frac{i}{N}\Big)^j \Big(1 - \frac{i}{N}\Big)^{N-j}.
$$

(a) Show that $Y_n = X_n(N - X_n)/(1 - 1/N)^n$ is a martingale.

(b) Hence, show that

$$
(N-1) \leq \frac{x(N-x)(1-1/N)^n}{P(0 < X_n < N \mid X(0) = x)} \leq \frac{N^2}{4}.
$$

(This gives a bound on the chance that either type a or A has spread through the entire population by time n.)

Solution:

$$\begin{aligned}
E[Y_{n+1} \mid X_n = x] &= \left(1 - \frac{1}{N}\right)^{-(n+1)} \left[NE[X_{n+1} \mid X_n = x] - E[X_{n+1}^2 \mid X_n = x]\right] \\
&= \left(1 - \frac{1}{N}\right)^{-(n+1)} \left[NE[Z] - \{Var(Z) + E[Z]^2\}\right] \\
&\quad \text{where } Z \sim Binomial(N, x/N) \\
&= \left(1 - \frac{1}{N}\right)^{-(n+1)} \left[Nx - N\frac{x}{N}\left(1 - \frac{x}{N}\right) - x^2\right] \\
&= \left(1 - \frac{1}{N}\right)^{-(n+1)} \left[Nx\left(1 - \frac{1}{N}\right) - x^2\left(1 - \frac{1}{N}\right)\right] \\
&= \left(1 - \frac{1}{N}\right)^{-(n)} [x(N - x)] \\
&= Y_n.
\end{aligned}$$

It follows from the above calculation that $E[Y_{n+1} \mid X_1, \ldots, X_n] = Y_n$ and hence, since $\{Y_1, \ldots, Y_n\}$ is a function of $\{X_1, \ldots, X_n\}$, that $E[Y_{n+1} \mid Y_1, \ldots, Y_n] = Y_n$. Y_n is therefore a martingale, completing (a). For (b), notice that $\{0 < X_n < N|X(0) = x\} = \{Y_n > 0|X(0) = x\}$. Also, if $Y_n > 0$ then

$$(N-1)\Big(1 - 1/N\Big)^{-n} \leq Y_n \leq \frac{N^2}{4}\Big(1 - 1/N\Big)^{-n}, \tag{1}$$

since X_n takes values in $0, \ldots, N$. The martingale property gives $E[Y_n|X(0) = x] = x(N - x)$, and so the identity $E[Y_n] = E[Y_n|Y_n > 0]P(Y_n > 0)$ can be rewritten as

$$\frac{x(N-x)}{P(0 < X_n < N \mid X(0) = x)} = E[Y_n|Y_n > 0]$$

Subsituting in (1) gives the required bound.

4. Let S_n be a simple nearest-neighbor random walk on the integers (i.e., $P\{S_{n+1}{=}S_n + 1\} = P\{S_{n+1}{=}S_n - 1\} = 1/2$), started at $S_0 = 1$. Define T to be the time of the first visit to the origin, i.e., $T = \min\{n : S_n = 0\}$. Define $Z_0 = 1$ and

$$Z_k = \sum_{n=0}^{T-1} \mathbf{1}\{S_n = k \text{ and } S_{n+1} = k + 1\}.$$

Thus, Z_k is the number of times that the random walk S_n crosses from k to $k + 1$ before first visiting 0.

(a) Show that the sequence $\{Z_k, k \geq 0\}$ is a Galton-Watson branching process, and identify the offspring distribution as a geometric distribution.

Hint: recall that X_n is a Galton-Watson process if it can be written as $X_{n+1} = \sum_{m=1}^{X_n} C_{nm}$ where $\{C_{nm}\}$ is an independent, identically distributed collection of non-negative integer-valued random variables.

(b) The probability generating function of Z_k is $\phi_k(s) = \sum_{j=0}^{\infty} s^j P(Z_k = j)$. Show how to write $\phi_k(s)$ in terms of $\phi_1(s)$, and use part (a) to find an explicit expression for $\phi_1(s)$.

Solution: (a). Let $U_{k1}, U_{k2}, \dots$ be the times of the crossings from k to $k+1$. These are regeneration times for the process $\{S_n, n \geq 0\}$. Defining N_{ki} to be the number of crossings from $k+1$ to $k+2$ in the time interval $[U_{ki}, U_{k(i+1)}]$, it follows that $N_{k1}, N_{k2}, \dots$ are independent and identically distributed. Letting $A_k = \{j : S_j = k\}$, we see that $\{N_{k1}, N_{k2}, \dots\}$ depend only on $\{S_{n+1} - S_n : n \in A_{k+1}\}$ and therefore $\{N_{k1}, N_{k2}, \dots\}$ is independent of $\{N_{\ell,m} : \ell \neq k\}$. Consequently,

$$Z_{k+1} = \sum_{j=1}^{Z_k} N_{kj}$$

is a Galton-Watson process. Now,

$$\begin{aligned}
&P(N_{kj} = n \mid N_{kj} \geq n) \\
&\quad = P[n\text{th upcrossing from } k+1 \text{ is the last before a downcrossing to } k] \\
&\quad = 1/2,
\end{aligned}$$

so $N_{kj} + 1$ is distributed as Geometric(1/2).
(b). Write $\phi(s) = \phi_1(s) = E[s^{Z_1}]$. Then,

$$\begin{aligned}
\phi_{k+1}(s) &= E[s^{Z_{k+1}}] \\
&= E[s^{\sum_{j=1}^{Z_k} N_{kj}}] \\
&= E\Big[E[s^{N_{k1}}]^{Z_k}\Big] \\
&= \phi_k(\phi(s)).
\end{aligned}$$

So, $\phi_k(s) = \phi(\phi(\dots\phi(s)))$, iterating k times. Here, $\phi(s) = \sum_{j=0}^{\infty} s^j (1/2)^{j+1} = 1/(2 - s)$.