Statistics 150 (Stochastic Processes): Midterm Exam, Spring 2009. J. Pitman, U.C. Berkeley.

1. A sequence of random variables $X_1, X_2, \ldots$, each with two possible values 0 and 1, is such that $\mathbb{P}(X_1 = 1) = p_1$ and for each $n \geq 1$

$$\mathbb{P}(X_{n+1} = 1 \,|\, X_1, \ldots, X_n) = (1 - \theta_n)p_1 + \theta_n S_n/n$$

where (θ_n) is a sequence of parameters with $0 \leq \theta_n \leq 1$, and $S_n := X_1 + \cdots + X_n$. Find and prove a formula for $\mathbb{P}(X_n = 1)$ in terms of p_1 and $\theta_1, \theta_2, \ldots$.

2. Consider a Markov chain (X_n) with transition matrix P. For $1 \leq m < n$ find and explain a formula for the conditional distribution of X_m given $X_0 = i$ and $X_n = k$ in terms of the entries of appropriate matrix powers of P.

3. Consider a $p \uparrow, 1 - p \downarrow$ walk (S_n) started at $S_0 = a$ and run until the time T when it first hits either 0 or b for some positive integers $0 \leq a \leq b$. Assuming $p \neq 1/2$, justify an application of Wald's identity to derive a simple formula for $\mathbb{E}(T \,|\, S_0 = a)$ in terms of p and the known solution of the gambler's ruin problem for an unfair coin, denote it

$$h(p, a, b) := \mathbb{P}(S_T = b \,|\, S_0 = a).$$

Note: You are *not* asked to provide or derive the formula for the hitting probability $h(p, a, b)$: you are asked to express $\mathbb{E}(T \,|\, S_0 = a)$ in terms of these probabilities.

4. Let $p_0, p_1, \ldots$ be a probability distribution on non-negative integers with mean $\mu := \sum_n np_n$, and let S_n be a Markov chain with transition probabilities $P(i, j) = p_{j-i+1}$ for $0 \leq i - 1 \leq j$ and $P(0, 0) = 1$. Let

$$f_{ij} := \mathbb{P}(S_n = j \text{ for some } n \geq 1 \,|\, S_0 = i)$$

Assume that $\mu \leq 1$. It then follows from results derived in class that $f_{i,j} = 1$ for all $i > j \geq 0$, and you can assume this to be true. Derive a system of equations satisfied by the f_{ij} for $0 < i \leq j$ which allow computation of these f_{ij} from the p_k. In particular, use these equations for $j = 2, 3$ to give an explicit formula for $f_{1,2}$ and for $f_{1,3}$.

5. In a branching process started with one individual, each individual has a geometrically distributed number of children, with probability $p(1 - p)^i$ for i children for $i = 0, 1, \ldots$. Find the probability of eventual extinction of the branching process in terms of p, as explicitly as possible.

6. Consider a simple nearest neighbour random walk on m points arranged around the circumference of a circle, at each step moving either one step clockwise or one step counterclockwise with equal probability. Let C_m be the *cover time*, that is the number of steps until every point has been visited at least once. Express C_m in terms of $\max_{1 \leq k \leq n} S_k$ and $\min_{1 \leq k \leq n} S_k$ for a simple random walk S_n on the integers instead of points around a circle, and deduce a formula for $\mathbb{E}(C_m)$.

7. In a simple population genetics model for neutral evolution with a fixed total population size N, let X_n represent the number of individuals in the population with a particular genetic characteristic, so $0 \leq X_n \leq N$. The model supposes that given $X_0, \ldots, X_n$, the distribution of X_{n+1} is binomial with parameters N and $p = X_n/N$. Show that for $0 \leq k \leq N$,

$$\lim_{n \to \infty} \mathbb{P}(X_n = N \,|\, X_0 = k) = k/N.$$