MATH 5763 Homework 5 Due Fri, 3/4/11

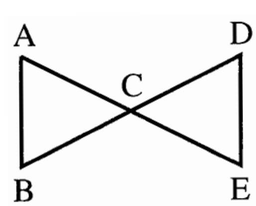
Problem 1. Consider an irreducible positive recurrent (discrete-time discrete-state space) Markov chain $\{X_n\}_{n=0}^{\infty}$, and assume that the initial state $X_0 = i$. Let $N_n(i)$ be the number of visits to state i in the first n trials, and $T_m(i)$ denote the number of trials until the mth visit to state i. Justify the relationship $\mathbb{P}(T_m(i) \geq n) = \mathbb{P}(N_n(i) \leq m)$. (Just give a convincing explanation in a couple of sentences.)

Problem 2. A *(finite, simple, undirected) graph* G is a finite collection of vertices V and a collection of edges E where each edge connects two different vertices, and any two vertices are connected by at most one edge. We write $v_1 \sim v_2$ if the vertices v_1 and v_2 are *adjacent*, i.e., if there is an edge connecting v_1 and v_2.

Consider the Markov chain whose states are the vertices of the graph. At each time interval, the chain chooses a new state randomly from among the states adjacent to the current state, with equal probability. In other words, if at time n the chain is in state i, at time $n+1$ it can be at any of the states adjacent to i, with equal probabilities. (Note that the Markov chain is not allowed to be in the same state in two consecutive times, i.e., the diagonal elements of the transition matrix are all zero.) This Markov chain is called *simple random walk on the graph.* Let $|E|$ be the total number of edges.

(a) Show that the stationary distribution of a simple Markov chain on a graph is given by $\pi_i = \dfrac{d_i}{2|E|}$, where d_i is the *degree* of the vertex i, i.e., the number of edges incident to the vertex i. In the figure in part (b) below, $d_{\mathrm{A}} = d_{\mathrm{B}} = d_{\mathrm{D}} = d_{\mathrm{E}} = 2$, and $d_{\mathrm{C}} = 4$.
Hint: Just check that $\boldsymbol{\pi}$ satisfies $\boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$ and $\sum_{i \in V} \pi_i = 1$.

(b) Write the transition probability matrix $\mathbf{P}$ for the "bowtie" graph in the picture below.



(c) Find the stationary distribution $\boldsymbol{\pi}$ of the simple random walk on the graph from (b) by using the result from (a).
Remark: To make sure that you understood all definitions, check by hand that $\boldsymbol{\pi}$ is a left eigenvector of $\mathbf{P}$ with eigenvalue 1.

(d) Find the expected time of first return to each of the five states of the "bowtie" graph.
Hint: One of the theorems discussed in class will make this part of the problem extremely easy.

Problem 3. Consider a simple 1-dimensional random walk with probability p of moving up and probability $q = 1 - p$ of moving down.

(a) For a state i, argue that the probability $p_{ii}^{(2n)} = \mathbb{P}(X_{2n} = i \,|X_0 = i)$ of revisiting this state after exactly $2n$ steps is equal to

$$p_{ii}^{(2n)} = \binom{2n}{n} p^n q^n$$

(the state i may have also been revisited before the $(2n)$th step). Note that $p_{ii}^{(2n+1)}$ (the probability of revisiting after an odd number of steps) is zero.

(b) Using Stirling's formula,

$$n! \approx \sqrt{2\pi}\, n^{n+\frac{1}{2}}\, e^{-n} \qquad \text{for large } n \ ,$$

show that

$$p_{ii}^{(2n)} \approx \frac{(4pq)^n}{\sqrt{\pi n}} \ .$$

(c) Using the result from part (b), show that for $p = q = \frac{1}{2}$,

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty \ ,$$

therefore in this case the simple random walk is persistent.

Remark: In parts (c) and (d), treat the Stirling's approximation as exact for large enough values of n (whether a series converges or diverges depends on how its terms for large n behave).

(d) Using the result from part (b), prove that, for $p \neq \frac{1}{2}$, the simple random walk is transient, i.e., that

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty \ .$$

Problem 4. The *probability generating function* (p.g.f.) of a random variable N taking only values in $\mathbb{Z}_+ = \{0, 1, 2, 3, \ldots\}$ is defined as

$$G_N(s) = \mathbb{E}[s^N] = \sum_n p_n s^n \ ,$$

provided the right-hand side exists (here $p_n = \mathbb{P}(N = n)$). Prove the following properties of the p.g.f.'s:

(a) $G_N(1) = 1$;

(b) $G'_N(1) = \mathbb{E}[N]$;

(c) $G''_N(1) = \mathbb{E}[N^2] - \mathbb{E}[N]$;

(d) $\operatorname{Var} N = G''_N(1) + G'_N(1) - [G'_N(1)]^2$;

(e) if N_1, N_2, ..., N_r are i.i.d. random variables taking values in $\mathbb{Z}_+$, and $Y = N_1 + \cdots + N_r$, then

$$G_Y(x) = [G_N(x)]^r \ .$$

Please write explicitly where you use each of the assumptions.

Problem 5. A simple random walk starts at position 1 and returns to position 1 at time $2n$ (i.e., $X_0 = 1$, $X_{2n} = 1$). If we use the pictorial representation of a random walk in the "time-position" plane (in which the random walk is represented as as a connected graph like in Problem 4 of Homework 4), then we can say that the random walk starts at the point $(0, 1)$ and goes to $(2n, 1)$ after $2n$ steps.

(a) Using the reflection principle from Problem 4 of Homework 4, show that there are $\dfrac{(2n)!}{n!(n+1)!}$ different paths between the points $(0, 1)$ and $(2n, 1)$ that do not ever revisit the origin (by "origin" we mean "position 0").

(b) What is the probability that the walk starting at $(0, 1)$ ends at $(2n, 1)$ after $2n$ steps without ever visiting the origin, assuming that the random walk is symmetric?

(c) Show that the probability that the first visit to the origin occurs at time $2n + 1$ is

$$p_n = \frac{1}{2^{2n+1}} \frac{(2n)!}{n!(n+1)!} \ .$$

Problem 6. Let X be a geometric random variable with parameter $p \in (0, 1)$, i.e., with p.m.f.

$$p_X(k) = \mathbb{P}(X = k) = (1 - p)^{k-1}\, p \ , \qquad k \in \mathbb{N} = \{1, 2, 3, \ldots\} \ .$$

(a) From the p.m.f. of X, find the probability that $X > n$ for $n \in \mathbb{N}$.

(b) Using your result from (a), show that the geometric random variable is *memoryless*, i.e., that it satisfies the property $\mathbb{P}(X > n + m | X > n) = \mathbb{P}(X > m)$ for any $n \in \mathbb{N}$ and $m \in \mathbb{N}$.

(c) Recall the meaning of the geometric random variables, and explain the reason for calling the property you proved in part (b) "memorylessness" (i.e., lack of memory).

(d) Can you explain the the reason for the memorylessness of geometric random variables intuitively, without any math?