UNIVERSITY OF CAMBRIDGE

M. Phil. in STATISTICAL SCIENCE

Thursday, 28 May, 2009 9:00 am to 11:00 am

INTRODUCTION TO PROBABILITY

Attempt no more than ***THREE*** *questions.*

There are ***FIVE*** *questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

(a) State and prove the central limit theorem. [You may use without proof the fact that

$$\lim_{n\to\infty}(1+c_n/n)^n = e^c$$

if $(c_n)_{n\geqslant 0}$ is a sequence of complex numbers such that $c_n \to c$ as $n \to \infty$, and any other result of the course].

(b) Let $(X_1, \ldots, X_n)$ be independent and identically distributed exponential random variables with parameter $\lambda > 0$. Let $M = \min(X_1, \ldots, X_n)$. Show that the distribution of M is that of an exponential random variable, with a parameter that you should identify.

(c) Let $(X_1, \ldots, X_n)$ be as in (b). Define $M_1, \ldots, M_n$ inductively by saying $M_1 = M$, and for $1 \leqslant i \leqslant n-1$:

$$M_{i+1} = \min\{X_j : 1 \leqslant j \leqslant n, X_j > M_i\}.$$

Show that if $1 \leqslant i \leqslant n$, M_i has the same distribution as

$$\frac{Y_n}{n} + \frac{Y_{n-1}}{n-1} + \ldots + \frac{Y_{n-i+1}}{n-i+1}.$$

where $Y_1, \ldots, Y_n$ are also independent and identically distributed exponential random variables with parameter λ. [*Hint:* Reason by induction. Use the memoryless property of exponential random variables together with (b).]

Deduce the mean, variance and Laplace transform of

$$M^* = \max(X_1, \ldots, X_n).$$

2 Let $(\theta_1, \theta_2, \ldots)$ be a sequence of independent and identically distributed random variables such that

$$\mathbb{P}(\theta_1 = 1) = 1/2; \quad \mathbb{P}(\theta_1 = -1) = 1/2.$$

Fix $X_0 = x \in \mathbb{R}$ an arbitrary nonrandom number, and for any $a > 0$, define a sequence of random variables $(X_1, X_2, \ldots)$ through the following recursive equation:

$$X_{n+1} = aX_n + \theta_{n+1}, \quad n \geqslant 0.$$

(a) Assume that $a = 1$ and $x = 0$. This process has been studied in class: what is its name, and what can you say about the asymptotic behaviour of X_n as $n \to \infty$?

(b) Assume now that $a > 0$ and $x \in \mathbb{R}$. Proceeding by induction, compute $\mathbb{E}(X_n)$.

(c) Assume that $a > 1$. Show that if the absolute value of the starting point $|x|$ is chosen large enough, then $\lim_{n\to\infty} |X_n| = \infty$, almost surely.

[*Hint:* Prove that if $|x|$ is large enough, the sequence X_n is monotone almost surely.]

Conclude that $\lim_{n\to\infty} |X_n| = \infty$ almost surely, no matter what the starting point x is.

3

(a) Define the notion of *martingale*. What does it mean to say that a martingale is bounded in L^2? State and prove the optional stopping theorem for bounded stopping times.

(b) Let $(X_n, n \geqslant 0)$ be a Markov chain on a finite state space S. Let D be a subset of S, and let

$$T = \inf\{n \geqslant 0 : X_n \in D\}.$$

Assume that $T < \infty$, $\mathbb{P}_x$-almost surely for all $x \in S$, and that there is a function $f : S \to \mathbb{R}$ and a number $\delta > 0$ such that for all $x \in S$, and for all $n \geqslant 0$,

$$\mathbb{E}\,(f(X_{n+1})|\, X_n = x) = f(x) - \delta\,,$$

and for all $x \in D$, $f(x) = 0$. Show that the process $Y_n = f(X_n) + n\delta$, $n \geqslant 0$, defines a martingale in an appropriate filtration, and use this to show that

$$\mathbb{E}_x\,(T) = \frac{f(x)}{\delta}$$

for all starting points $x \in S$.

(c) Fix $q \in (1/2, 1)$ and $N \geqslant 2$. Let $X = (X_n, n \geqslant 0)$ be a Markov chain on

$$S = \{\ldots, -1, 0, 1, \ldots, N-1, N\},$$

such that if $x \in S$, $p(x, x-1) = 1 - p(x, x+1) = q$ if $x < N$, and $p(x, x-1) = 1$ if $x = N$. You may assume without proof that X is transient.

Let $T = \inf\{n \geqslant 0 : X_n = 0\}$ be the hitting time of 0. Show that for every $x \in \{1, 2, \ldots\}$,

$$\mathbb{E}_N\,(T) = \frac{N + 2q - 2}{2q - 1}.$$

[*Hint:* consider the function f defined by $f(x) = x$ for $x < N$, and $f(N) = N + c$, and apply the result from (b) by choosing a suitable c.]

4

(a) Let $(S_n, n \geqslant 0)$ be a simple random walk on $\mathbb{Z}$. Let $a, b \in \mathbb{N}$ with $a, b > 0$, and for all $x \in \mathbb{Z}$, let $T_x = \inf\{n \geqslant 0 : S_n = x\}$. Prove that

$$\mathbb{P}_0(T_{-a} < T_b) = \frac{b}{a+b}$$

and that

$$\mathbb{E}_0(T_{-a} \wedge T_b) = ab.$$

(You may use freely in your proof that S_n and $S_n^2 - n$ are martingales in the canonical filtration, as well as any other result of the course you wish.)

(b) Fix $N \geqslant 1$ and let $\lambda_1, \dots, \lambda_{N-1}$ be $(N-1)$ fixed positive numbers. Let $(X_t, t \geqslant 0)$ be a Markov chain in continuous time on $S = \{0, \dots, N\}$, whose generator is given by the matrix $Q = (q_{i,j})_{0 \leqslant i,j \leqslant N}$, where

$$q_{i,i} = -\lambda_i, \quad q_{i,i+1} = q_{i,i-1} = \frac{\lambda_i}{2}, \quad 1 \leqslant i \leqslant N-1,$$

and all the other entries are 0. Assuming that the chain starts at $X_0 = i \in S$, let J_i be the amount of time that X stays at i before jumping to a different state. What is the distribution of J_i? What can you say about X_{J_i}, the next state visited by the chain? (Distinguish between the cases $0 < i < N$ and the boundary cases $i = 0$ or $i = N$.)

Show that with probability 1, eventually the chain gets absorbed either at state $i = 0$ or $i = N$. Let E be the event that the chain gets absorbed at 0. What is $\mathbb{P}_i(E)$, for $i \in S$?

(c) Let $(E_1, E_2, \dots,)$ be independent and identically distributed random variables with exponential distribution with rate $\lambda > 0$, and let K be an independent random variable with geometric distribution with parameter $0 < p < 1$. Compute $\mathbb{E}(S)$, where S is the random variable defined by:

$$S = \sum_{n=1}^{K} E_n. \tag{1}$$

Fix $0 < i < N$, and let I be the total time spent by the chain $(X_t, t \geqslant 0)$ of part (b) at state i. Show that

$$\mathbb{E}_i(I) = \frac{2i(N-i)}{N\lambda_i}.$$

[*Hint:* Show that I may be expressed as a random variable of the form of (**??**), and use appropriate versions of the result in (a) to identify the parameters].

5 Let $\lambda > 0$ and let Z be a Poisson random variable with parameter λ.

(a) Compute the probability generating function $\mathbb{E}(s^Z)$ of Z, for $s \in [0, 1]$.

Give a definition of the Poisson process $(N_t, t \geqslant 0)$ with rate $\lambda > 0$. Show that the law of N_t for some fixed time $t > 0$ is Poisson with some parameter which you should specify.

Deduce that if $s < t$, then $N_t - N_s$ is a Poisson random variable with mean $\lambda(t - s)$.

(b) Let $(X_i, i \geqslant 1)$ be independent and identically distributed random variable with a fixed given distribution such that $X_1 \geqslant 0$ almost surely, which are independent from Z. For $q \geqslant 0$, we denote by $\phi(q) = \mathbb{E}(e^{-qX_1})$ the Laplace transform of the distribution of X_1. Compute the Laplace transform of the random variable Y defined by:

$$Y = \sum_{i=1}^{Z} X_i.$$

Express your answer in terms of the function ϕ. How can you use your answer to compute $\mathbb{E}(Y)$? Find another method to compute this expectation and check that the two methods give the same result.

(c) Let $(N_t, t \geqslant 0)$ be a Poisson process with rate λ. Let

$$Y_t = \sum_{i=1}^{N_t} X_i,$$

where $(X_1, \ldots)$ are as in (b). Show that for all $q \geqslant 0$, the process $(Z_t, t \geqslant 0)$ defined by:

$$Z_t = \exp\left(-qY_t - \lambda t\phi(q) + \lambda t\right)$$

is a martingale, in the sense that $\mathbb{E}(|Z_t|) < \infty$ for all $t \geqslant 0$ and for all $s < t$,

$$\mathbb{E}(Z_t|\mathcal{F}_s) = Z_s$$

where $\mathcal{F}_s = \sigma(Y_u, u \leqslant s)$.

[*Hint:* show that Z_t/Z_s is independent of $\mathcal{F}_s$.]

END OF PAPER