M343/S

The Open University

Third Level Course Examination 1997
Applications of Probability

Wednesday, 22 October, 1997 2.30 – 5.30 pm

Time allowed: 3 hours

This examination is in **TWO** parts. Part I carries 40% of the total available marks and Part II carries 60%.

You should attempt **FOUR** questions from Part I: each question carries 10 marks. You should attempt **THREE** questions from Part II: the questions in this part carry 20 marks each.

Write your answers to Parts I and II in **separate** answer books. Please start each question on a new page, and cross out rough working.

At the end of the examination

Fasten together your answer books for Parts I and II, using the paper fastener provided. Check that you have written your name, personal identifier and examination number on **each** answer book used. **Failure to do so will mean that your work cannot be identified.**

PART I (Questions 1 to 6)

You should attempt **FOUR** *questions from this part of the examination, which carries 40% of the total available marks. Each question carries 10 marks.*

Question 1

In an investigation into whether redwood seedlings are randomly located in an experimental plot, a sample of point-to-nearest-object distances (R-distances) is obtained, and a sample of the same number of object-to-nearest-object distances (S-distances) is obtained. The following data (distances in metres) are collected.

R-distances	10	5	19	2	8	12	17	3	7	12
S-distances	7	2	5	11	1	10	3	7	2	8

Use this information to investigate whether the seedlings may be regarded as being randomly located in the plot. If your investigation suggests that they are not randomly located, then say how you think the seedlings are located.

Question 2

In a Galton-Watson branching process starting with a single individual in generation zero, the offspring random variable X has the probability generating function

$$\Pi(s) = \left(\frac{2}{5-3s}\right)^2.$$

(i) Calculate to four decimal places the probability that the process becomes extinct

(a) by the second generation,

(b) at the third generation.

(ii) Calculate to four decimal places the probability that the process eventually becomes extinct.

Question 3

James is playing *Rout* with his brother. *Rout* is a game involving both skill and luck; no game ever results in a draw; the probability that James wins any particular game is 0.4. The brothers start all square: games are played until one or other of the brothers has won three games more than the other brother. That player is the overall winner of the contest.

(i) Calculate the probability that in any contest James' brother is declared the winner.

(ii) Find the average number of games played in any particular contest.

(iii) James secretly obtains tuition in the skills of *Rout* from his uncle, and the next time he meets his brother in a contest the pair of them may be assumed to be equally matched. What is the expected duration of the contest?

(iv) When James plays *Rout* with his younger sister, the probability that he wins any particular game is 0.8. To give her an opportunity to develop her skills at the game, he agrees to keep playing until she has won three games more than he has. Calculate the probability that she will ever be three games ahead of James.

Question 4

(a) Characterize the following queues according to the standard notation.

(i) Vehicles arrive at random at a filling station where there are two pumps. The time taken to refuel is equally likely to be anything from 2 to 5 minutes.

(ii) Patients arrive every ten minutes at a dental surgery; there is one dentist on duty. The time spent with each patient has an exponential distribution with mean 8 minutes.

(b) Two assistants are on duty in a baker's shop. Customers may be assumed to arrive independently and at random at an average rate of 18 per hour. If an assistant is free, then an arriving customer is served immediately; otherwise a central queue is formed. The service time for each assistant may be assumed to be exponentially distributed with mean 4 minutes.

(i) Calculate the traffic intensity of this queue.

Assume that the queue is in equilibrium.

(ii) Find the proportion of the time that both assistants are idle.

(iii) What proportion of customers arriving at the shop receive immediate attention?

(iv) One day a customer arrives to find both assistants busy but nobody waiting for service. Find the probability that he has to wait at least three minutes for service.

Question 5

The lifetime of members of a stationary population of insects has probability density function

$$f(x) = \frac{2x}{9}, \quad 0 \leq x \leq 3,$$

where x is measured in weeks.

(i) Find the life table function $Q(x)$ for members of the population.

(ii) Find the mean age at which insects die.

(iii) Find the mean age of insects in the population.

(iv) What proportion of insects in the population are less than a week old?

(v) What proportion of insects live for less than a week?

Question 6

The lifetime T (in years) of a heating element in an electric fire may be described in terms of its hazard function

$$h(t) = \frac{t}{16}, \quad t \geq 0.$$

(i) Find the survivor function and hence identify the distribution of the random variable T.

(ii) Calculate the mean lifetime and the median lifetime of these elements.

(iii) What proportion of elements last for more than twice the mean lifetime?

(iv) Are this type of elements NBU, NWU or neither of these? Justify your answer.

PART II (Questions 7 to 12)

You should attempt **THREE** *questions from this part of the examination, which carries 60% of the total available marks. Each question carries 20 marks: the mark allocation for each part of each question is shown beside each part thus:* [4].

Question 7

(a) (i) Give one example of a random phenomenon which could reasonably be modelled as a non-homogeneous Poisson process. Draw a rough sketch of the rate parameter $\lambda(t)$ for your example. [2]

(ii) Give one example of a random phenomenon which could reasonably be modelled as a compound Poisson process. Explain briefly why a compound Poisson process is an appropriate model. [2]

(b) Vehicles arrive at a census checkpoint according to a Poisson process at an average rate of 45 vehicles per hour.

(i) Find the probability that no vehicle arrives during a five-minute period. [2]

Of the vehicles arriving at the checkpoint, 12% are lorries and 4% are motorcycles.

(ii) What is the mean time between the arrival of successive motorcycles at the checkpoint? [2]

(iii) What is the distribution of the number of lorries that arrive at the checkpoint in a twenty-minute period? Calculate the probability that at least three lorries arrive in a twenty-minute period. [5]

(iv) Calculate the probability that ten vehicles arrive in a twenty-minute period, exactly three of which are lorries. [5]

(v) Ten vehicles arrive at the checkpoint between 11 am and 11.20 am one day. What is the probability that exactly three of them are lorries? [2]

Question 8

A Markov chain model has been proposed to describe the way in which the male wage-earners in a family change in occupational status from generation to generation. Data on the occupational status of men in a region of the United Kingdom and of their fathers was used to estimate the matrix $\mathbf{P}$ of transition probabilities given below.

$$\mathbf{P} = \begin{matrix} \mathit{1} \\ \mathit{2} \\ \mathit{3} \end{matrix} \begin{pmatrix} 0.4 & 0.5 & 0.1 \\ 0.1 & 0.6 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$$

The three occupational levels used in the model were as follows.

Level 1: Upper (professional and managerial)

Level 2: Middle (non-manual and skilled manual)

Level 3: Lower (semi-skilled and unskilled manual)

(i) What assumptions about changes in occupational status are made by using a Markov chain model? Comment briefly on whether these assumptions are likely to be valid. [5]

(ii) A man is in the upper occupational level (Level 1). According to the model, what is the probability that his son and grandson are both in the middle occupational level? What is the probability that his grandson is in the lower occupational level? [4]

(iii) According to the model, in the long run what proportion of male wage-earners will be in each occupational level? [7]

(iv) The proportions of male employees of a large company in each of the three occupational levels (Upper, Middle and Lower) are 0.1, 0.3, 0.6 respectively. Assuming that each employee has one son, what proportions of their sons will be in each occupational level? What proportion of their grandsons will be in each occupational level? [4]

Question 9

(a) Give three reasons why the simple birth-death process is not a good model for an evolving human population. [3]

(b) The death rate in a simple birth-death process is $\nu = 2\beta$, where β is the birth rate. A partial differential equation for the probability generating function $\Pi(s,t)$ of $X(t)$, the number of individuals alive at time t in the simple birth-death process, is

$$\frac{\partial \Pi}{\partial t} = \beta(1-s)(2-s)\frac{\partial \Pi}{\partial s}.$$

(i) Rewrite this equation in Lagrange form and identify the functions f, g and h. [2]

(ii) Write down the auxiliary equations. One solution is $c_1 = \Pi$; show that the other solution may be written

$$c_2 = \frac{2-s}{1-s}e^{\beta t}.$$ [4]

Hint: *You might find the following identity useful*:

$$\frac{1}{(1-s)(2-s)} = \frac{1}{1-s} - \frac{1}{2-s}.$$

(iii) Write down the general solution for the partial differential equation. [1]

(iv) Suppose that at time 0 there are four individuals alive in the population. Find the particular solution for $\Pi(s,t)$ in this case. [4]

(v) Show that the population is certain to die out eventually. [2]

(vi) If there are four individuals alive at time 0, find the probability that the population will attain size 7 before it dies out. [4]

Question 10

(a) Identify the main differences between the simple epidemic model and the general epidemic model. [4]

(b) In a family of four, two parents and twin children, the two parents are stricken with an infectious disease. The spread of the disease through the family can be modelled as a stochastic general epidemic. The infectious contact rate is $\beta = 6$. Anyone who is infected remains infectious for a time which is exponentially distributed with mean $\frac{1}{2}$.

(i) Show that the epidemic parameter ρ is equal to 1. [1]

(ii) Find the probability distribution of the number of members of the family who never catch the disease. [7]

In fact, the whole family of four ends up catching the disease. The disease is not serious and the twins return to school while still infectious. There are 29 other children in their primary school class, none of whom has had the disease. The infectious contact rate in this environment is $\beta = 5$.

(iii) Show that the epidemic parameter in this environment is equal to 12. [1]

(iv) According to a stochastic model, what approximately is the probability of a major outbreak of the disease in the class? [2]

(v) Using a deterministic model, calculate the number of children in the class who do not catch the disease. [5]

Question 11

A particular skin condition can be thought of as due to a single recessive allele **a**. Doubly recessive individuals (genotype **aa**) have the skin condition; doubly dominant individuals (genotype **AA**) are normal; heterozygotes (genotype **Aa**) are carriers. In a population where the skin condition occurs, the gene frequencies are stable and the genes **A** and **a** occur in the proportions **A** : **a** = 0.8 : 0.2.

(i) What proportion of individuals (a) are normal, (b) are carriers, (c) have the skin condition? [4]

(ii) Given that an individual does not have the skin condition, what is the probability that (a) he is normal; (b) he is a carrier? [3]

(iii) Assuming that random mating occurs, find the probability that the child of parents who do not have the skin condition will (a) be normal, (b) be a carrier. (c) have the skin condition. [8]

(iv) Find the probability that the parents of a child who is a carrier are both themselves carriers. [5]

Question 12

(a) Describe briefly an example of a situation where a Brownian bridge might be an appropriate model. [2]

(b) Two children leave their home at 2 pm to play along the bank of a stream which runs past their home. Their distance upstream from home t hours after leaving it is denoted by $X(t)$, and may be modelled as an ordinary Brownian motion $\{X(t); t \geq 0\}$ with diffusion coefficient $\sigma^2 = 800$ (metres)^2 per hour.

  (i) Find the probability that two hours after leaving home they are more than 50 metres away from home. [3]

  (ii) At 3.30 pm, the children are observed by a footbridge 60 metres upstream from their home.

    (a) Find the probability that at 4 pm they will be more than 50 metres upstream from their home.

    (b) Find the probability that at 3 pm they were within 30 metres of their home. [8]

  (iii) Given only that they left home at 2 pm, find the probability that when they first reach the footbridge it is after 4 pm. [3]

  (iv) It is required to simulate the distance from home of the children every half hour from 2 pm until 4 pm on an expedition along the bank of the stream. Using the random numbers from normal distributions on page 43 of *Neave*, starting at the beginning of the twentieth row and working across (1.2816. 0.4435, ...), simulate their position to the nearest metre at half-hourly intervals. [4]

[END OF QUESTION PAPER]