THE QUEEN'S UNIVERSITY OF BELFAST

110SOR201

Level 2 Examination

Statistics and Operational Research 201

Probability and Distribution Theory

Monday 21 August 2000 9.30 am — 12.30 pm

Examiners Professor R M Loynes
and the internal examiners

Answer **FIVE** questions

All questions carry equal marks

Approximate marks for parts of questions are shown in brackets

Write on both sides of the answer paper

1. (i) State three fundamental axioms concerning P in a probability space $(\mathcal{S}, \mathcal{F}, \mathrm{P})$. Given two events $A, B \in \mathcal{F}$, show from the axioms that the probability that exactly one of the events occurs is

$$\mathrm{P}(A) + \mathrm{P}(B) - 2\mathrm{P}(A \cap B).$$ **[6]**

(ii) Express the probability $\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n)$ in terms of the probabilities of the events $A_1, ..., A_n \in \mathcal{F}$ and their intersections.

Each packet of a certain breakfast cereal contains one of a set of 5 distinct tokens, and a given packet is equally likely to contain any token. Find the probability that a consignment of 10 packets will contain at least one complete set of tokens (do not simplify your answer). **[10]**

(iii) Prove by induction, or otherwise, that for n general events $A_1, ..., A_n \in \mathcal{F}$,

$$\mathrm{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \geq \sum_{i=1}^{n} \mathrm{P}(A_i) - (n-1).$$ **[4]**

2. (i) Prove that if two events $A, B \in \mathcal{F}$ are independent, then A and $\overline{B}$ are independent. **[3]**

(ii) A random number N of fair dice is thrown, where

$$\mathrm{P}(N = n) = 2^{-n}, \qquad n \geq 1.$$

Let S denote the sum of the scores on the dice. Find the probability that

(a) $N = 2$, given $S = 3$; **[4]**

(b) $S = 3$, given N is odd. **[5]**

(iii) In a series of independent games, a player has probabilities $\frac{4}{9}, \frac{1}{3}, \frac{2}{9}$ of scoring $0, 1, 2$ points respectively in any game. The series ends when the player scores 0 in a game, and the scores in individual games are then added to give a total score.

(a) By conditioning on the result of the first game, find the probability that the total score is an odd number. **[4]**

(b) Let p_n denote the probability that the total score is exactly n points. Obtain a recurrence relation for p_n, and give the values of p_0, p_1. **[4]**

3. (i) Suppose that X and Y are independent count random variables. Express the probability distribution of $Z = X + Y$ in terms of the probability distributions of X and Y. **[2]**

(a) If X and Y are Poisson distributed with parameters λ_1 and λ_2 respectively,i.e.

$$\mathrm{P}(X = x) = \frac{{\lambda_1}^x e^{-\lambda_1}}{x!}; \qquad \mathrm{P}(Y = y) = \frac{{\lambda_2}^y e^{-\lambda_2}}{y!},$$

show that Z is also Poisson distributed. **[3]**

(b) If X and Y have the same distribution

$$\mathrm{P}(X = k) = \mathrm{P}(Y = k) = pq^k, \qquad k = 0, 1, \ldots; \quad p + q = 1,$$

show that the conditional distribution $\mathrm{P}(X = x \mid Z = n)$ is uniform. Also, show that

$$\mathrm{P}(X > n) = q^{n+1}$$

and hence, or otherwise, determine $\mathrm{P}(X > Y)$. **[8]**

(ii) Suppose that $(X_1, \ldots, X_k)$ has the multinomial distribution

$$\mathrm{P}(X_1 = x_1, \ldots, X_k = x_k) = \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k}$$

where $\sum_{i=1}^{k} x_i = n, \sum_{i=1}^{k} p_i = 1$ and X_i denotes the number of times outcome i occurs in a sequence of n independent trials. Let

$$I_{ri} = \begin{cases} 1, & \text{if the } r^{\text{th}} \text{ trial results in outcome } i \\ 0, & \text{otherwise.} \end{cases}$$

How are X_i and $I_{1i}, \ldots, I_{ni}$ related? Determine $\mathrm{E}(I_{ri})$ and $\mathrm{Cov}(I_{ri}, I_{rj}), i \neq j$ and deduce that

$$\mathrm{Cov}(X_i, X_j) = -np_i p_j, \quad i \neq j.$$ **[7]**

4. (i) Define the probability generating function (PGF) $G_X(s)$ of a count random variable X, and indicate how $\mathrm{E}(X)$ can be derived from it. If X is binomially distributed with parameters (n, p), show that

$$G_X(s) = (q + ps)^n \qquad (q = 1 - p)$$

and deduce $\mathrm{E}(X)$. **[5]**

(ii) Given a sequence $X_1, X_2, \ldots$ of independent, identically distributed random variables, each with mean μ and PGF $G(s)$, write down (without proof) expressions for the PGFs of the random variables $\sum_{i=1}^{n} X_i$ (where n is known) and $\sum_{i=1}^{N} X_i$ (where N is a further independent count random variable with PGF $G_N(s)$).

Suppose that each X_i has the following distribution:

$$\mathrm{P}(X = k) = \begin{cases} pq, & k = 0, 2 \\ p^2 + q^2, & k = 1 \\ 0, & \text{otherwise} \end{cases} \quad ; \qquad p + q = 1.$$

Find the PGF of $S = \sum_{i=1}^{n} X_i$ and deduce that S has the same distribution as $U + V$, where U and V are independent binomially distributed random variables with parameters (n, p) and (n, q) respectively. **[6]**

(iii) Let X_n denote the size of the n^{th} generation in a branching process in which the family sizes are independent and identically distributed random variables, each with mean μ and PGF $G(s)$, and suppose that $X_0 = 1$.

(a) Explain why $G_n(s)$, the PGF of X_n, satisfies the recurrence relation

$$G_n(s) = G_{n-1}(G(s)), \qquad n \geq 1,$$

and deduce that

$$\mathrm{E}(X_n) = \mu^n, \qquad n \geq 1.$$ **[6]**

(b) It can be shown that

$$e \equiv \lim_{n\to\infty} \mathrm{P}(X_n = 0)$$

is the smallest non-negative root of the equation $e = G(e)$. Using the properties of G, deduce that

$$e = 1 \quad \text{if and only if} \quad \mu \leq 1.$$ **[3]**

5. (i) In a population of constant size N, each individual is of type A or not. Let $X_n = i$ if i individuals are of type A at time $n \quad (n = 0, 1, 2, \ldots)$. During each unit time interval an individual is chosen at random from the population and replaced by a new individual : the probability of the new individual being of type A is i/N when there were i individuals of type A in the population at the beginning of the time interval.

Explain why this system is a homogeneous Markov chain, and give its transition probability matrix. Classify the states of the chain. **[8]**

(ii) A homogeneous Markov chain $\{X_n : n = 0, 1, \ldots\}$ has states $\{0, 1, 2\}$ and transition probability matrix

$$\boldsymbol{P} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

At time $n = 0$, the system is equally likely to be in state 0 or state 1.

Quoting any standard results used, determine

(a) $\mathrm{P}(X_2 = 1)$;

(b) $\lim_{n\to\infty} \mathrm{P}(X_n = 1)$. **[9]**

(iii) In a finite absorbing Markov chain with transition probability matrix $\boldsymbol{P} = (p_{ij})$, let T and A denote the sets of transient and absorbing states respectively, and f_{ik} the probability of eventual absorption in state $k \in A$ if starting from state $i \in T$. Write down (and briefly justify) a set of linear equations satisfied by the absorption probabilities $\{f_{ik} : i \in T\}$. **[3]**

6. (i) The lifetime X of a certain device has c.d.f.

$$F(x) = 1 - e^{-\lambda x^2}, \qquad x \geq 0, \quad \lambda > 0.$$

Derive the p.d.f. of X, $f(x)$, and determine its mean, variance and mode.
Also determine the hazard rate function $r(x) = f(x)/[1-F(x)]$, and briefly explain its significance.
(*Hint*: the function

$$\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt, \qquad p > 0,$$

has the properties

$$\Gamma(p+1) = p\Gamma(p) : \quad \Gamma(1/2) = \sqrt{\pi} : \quad \Gamma(n+1) = n!, \quad n \text{ integer} \geq 0.)$$

[10]

(ii) If the random variable X is uniformly distributed over $[0, 1]$, i.e. its p.d.f. is

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise,} \end{cases}$$

find the p.d.f. and c.d.f. of

$$Y = X^\alpha,$$

where α may be positive or negative. Sketch the p.d.f. of Y when $\alpha = -1, \frac{1}{2}, 2$.
[10]

7. (i) Let X and Y be independent continuous random variables, each distributed uniformly on $[0,1]$, and let

$$U = XY, V = X.$$

Show that the joint p.d.f. of U, V is

$$f_{U,V}(u,v) = \begin{cases} 1/v, & 0 \le u \le v \le 1 \\ 0, & \text{otherwise,} \end{cases}$$

and deduce the p.d.f. of U. **[5]**

(ii) Let $Z_1, ..., Z_n$ be independent $N(0,1)$ random variables, and let

$$\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{Z},$$

where $\boldsymbol{Y}, \boldsymbol{Z}$ are $(n \times 1)$ vectors with components $(Y_1, ..., Y_n), (Z_1, ..., Z_n)$ respectively, and $\boldsymbol{C}$ is an orthogonal $(n \times n)$ matrix, so that

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n} Z_i^2.$$

Show that $Y_1, ..., Y_n$ are independent $N(0,1)$ random variables. Then, by choosing the first row of $\boldsymbol{C}$ to be $(\frac{1}{\sqrt{n}}, ..., \frac{1}{\sqrt{n}})$, show that, for a random sample from $N(0,1)$, the sample mean $\overline{Z}$ and the sample variance S^2 are independent random variables. State the distributions of $\overline{Z}$ and $(n-1)S^2$. **[9]**

(iii) Let $X_{(n)}$ denote the largest observation in a random sample of size n from a distribution with p.d.f. $f(x)$ and c.d.f. $F(x)$, $-\infty < x < \infty$. By considering the c.d.f. of $X_{(n)}$, or otherwise, obtain an expression for $f_{(n)}(x)$, the p.d.f. of $X_{(n)}$. For the uniform distribution with

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise,} \end{cases}$$

show that

$$\text{Var}(X_{(n)}) = \frac{n}{(n+1)^2(n+2)}.$$

[6]

8. (i) Define the moment generating function (MGF) $M_X(\theta)$ of a continuous random variable X, and explain why it is so called. Write down (without proof)

(a) the MGF of $aX + b$ in terms of $M_X(\theta)$;

(b) the MGF of the sum of n independent random variables $X_1, \ldots, X_n$ in terms of their individual MGFs. **[5]**

(ii) If $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$, show that

$$M_Z(\theta) = \exp(\tfrac{1}{2}\theta^2)$$

and deduce the MGF of X. **[5]**

(iii) If X is uniformly distributed on [0,1], i.e.

$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{otherwise,} \end{cases}$$

show that

$$M_X(\theta) = (e^\theta - 1)/\theta$$

and hence find the mean and variance of X. **[4]**

(iv) If $X_i \ \ (i = 1, \ldots, n)$ are independent random variables, each uniformly distributed on [0,1], write down the MGF of

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

and show that, for large n,

$$\log_e M_{\overline{X}}(\theta) = \tfrac{1}{2}\theta + \tfrac{1}{24}\theta^2 \cdot \tfrac{1}{n} + o(\tfrac{1}{n}).$$

Hence find an approximation to the distribution of $\overline{X}$ when n is large.
(*Hint*: $\log_e(1+a) = a - \frac{1}{2}a^2 + \frac{1}{3}a^3 - \cdots$ provided $|a| < 1$). **[6]**

9. (i) State the conditions under which a counting process $\{N(t), t \geq 0\}$ is a *Poisson process* with rate λ. Show that in a Poisson process the probabilities

$$p_n(t) = \mathrm{P}[N(t) = n], \qquad n = 0, 1, ...$$

satisfy the equations

$$\begin{aligned} \frac{dp_0(t)}{dt} &= -\lambda p_0(t) \\ \frac{dp_n(t)}{dt} &= \lambda p_{n-1}(t) - \lambda p_n(t), \qquad n = 1, 2, ... \end{aligned}$$

and prove by induction, or otherwise, that

$$p_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, \qquad n = 0, 1, ...$$ **[11]**

(ii) (a) Consider a stationary Markov process $\{X(t), t \geq 0\}$ with transition probability functions $\{p_{ij}(t)\}$. What form does $p_{ij}(\delta t)$ take for a birth-and-death process with birth rates $\{\alpha_i : i = 0, 1, ...\}$ and death rates $\{\beta_i : i = 1, 2, ...\}$? Draw the corresponding transition rate diagram. **[3]**

(b) For the birth-and-death process in (a), it can be shown that the probabilities

$$p_n(t) = \mathrm{P}(X(t) = n), \quad n = 0, 1, ...$$

satisfy the equations

$$\frac{dp_n(t)}{dt} = -(\alpha_n + \beta_n)p_n(t) + \alpha_{n-1}p_{n-1}(t) + \beta_{n+1}p_{n+1}(t), \quad n = 0, 1, ...$$

where $\alpha_{-1} = \beta_0 = 0$. Deduce that if a steady-state distribution $\{\pi_m : m = 0, 1, \ldots\}$ exists, then

$$\alpha_m \pi_m = \beta_{m+1}\pi_{m+1}, \quad m = 0, 1, \ldots$$

In a simple (M/M/1) queue, arrivals follow a Poisson process with rate λ and the service time is negative exponential with mean μ^{-1}, where $\lambda < \mu$. Show that

$$\pi_n = (1 - \rho)\rho^n, \qquad n = 0, 1, ...,$$

where $\rho = \lambda/\mu$. **[6]**