THE QUEEN'S UNIVERSITY OF BELFAST

110SOR201

Level 2 Examination

Statistics and Operational Research 201

Probability and Distribution Theory

Wednesday 10 January 2001 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 a minimal set of axioms concerning the probability measure P in a probability space $(\mathcal{S}, \mathcal{F}, \mathrm{P})$. **[3]**

Deduce from the axioms that, if $A, B \in \mathcal{F}$, then

(a) $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$; **[2]**

(b) the probability that exactly one of the events occurs is

$$\mathrm{P}(A) + \mathrm{P}(B) - 2\mathrm{P}(A \cap B).$$ **[4]**

(ii) If $A_1, ..., A_n \in \mathcal{F}$, use the axioms (together with the addition law for two events, which can be derived from the axioms) to show by induction that

$$\mathrm{P}(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} \mathrm{P}(A_i).$$

Write down (without proof) an exact expression for $\mathrm{P}(A_1 \cup \cdots \cup A_n)$ in terms of the probabilities of the events $A_1, ..., A_n$ and their intersections. **[5]**

(iii) In a special promotion, a garage issues a token for every £10 worth of petrol purchased. Each token bears one of 6 symbols, with equal likelihood, and any customer who acquires a complete set of the 6 symbols wins a prize. Find the probability that a customer who acquires 12 tokens on visits to the garage will win a prize. (Do not reduce your answer.) **[6]**

2. (i) Given a probability space $(S, \mathcal{F}, \mathrm{P})$ and an event $B \in \mathcal{F}$ with $\mathrm{P}(B) > 0$, define the conditional probability $\mathrm{P}(A|B)$ for an event $A \in \mathcal{F}$. If $A_1, ..., A_n \in \mathcal{F}$, prove that, under a certain condition (to be carefully stated),

$$\mathrm{P}(A_1 \cap \cdots \cap A_n) = \mathrm{P}(A_n|A_1 \cap \cdots \cap A_{n-1}).\mathrm{P}(A_{n-1}|A_1 \cap \cdots \cap A_{n-2})...\mathrm{P}(A_2|A_1)\mathrm{P}(A_1).$$

How does this simplify if the events $A_1, ..., A_n$ are independent? Show that in this case

$$\mathrm{P}(\overline{A}_1 \cup \cdots \cup \overline{A}_n) = 1 - \prod_{i=1}^{n} \mathrm{P}(A_i).$$ **[7]**

(ii) Suppose that a fair die is thrown repeatedly.

(a) Find the probability that a six is thrown before an odd number is thrown. **[3]**

(b) Let u_n denote the probability that, in the first n throws, an odd number of sixes is obtained. Derive an expression for u_n in terms of u_{n-1}, and show, by induction or otherwise, that

$$u_n = \tfrac{1}{2}\left[1 - \left(\tfrac{2}{3}\right)^n\right], \qquad n \geq 0.$$ **[5]**

(c) Let v_n denote the probability that, during the first n throws, a run of even numbers in 3 successive throws is *not* obtained . By conditioning on the first occurrence of an odd number, derive the recurrence relation

$$v_n = \tfrac{1}{2}v_{n-1} + \tfrac{1}{4}v_{n-2} + \tfrac{1}{8}v_{n-3}, \qquad n \geq 3,$$

and use it to compute v_5. **[5]**

3. (i) The numbers X and Y of male and female customers entering a certain store are independent and Poisson distributed with means λ_1 and λ_2 respectively, i.e.

$$P(X = x) = \frac{\lambda_1^x e^{-\lambda_1}}{x!}, \quad x = 0, 1, \ldots \quad ; \qquad P(Y = y) = \frac{\lambda_2^y e^{-\lambda_2}}{y!}, \quad y = 0, 1, \ldots.$$

Any customer entering the store has a probability p of spending more than £10 on purchases.

(a) Show that N, the total number of people entering the store, is Poisson distributed with mean $\lambda = \lambda_1 + \lambda_2$. **[4]**

(b) Use the result

$$E(Z) = E(E(Z|N))$$

to show that Z, the number of customers spending more than £10, has mean λp. **[2]**

(c) Show that the distribution of Z is Poisson. **[4]**

(ii) A bag contains N balls numbered 1 to N. Balls are drawn at random, one at a time, without replacement.

(a) Let X be the largest number selected after n balls have been withdrawn ($n \leq N$). Find the probability distribution of X. **[2]**

(b) A *match* A_i is said to occur if the i^{th} ball drawn bears the number i. Let

$$I_i = \begin{cases} 1, & \text{if } A_i \text{ occurs} \\ 0, & \text{otherwise,} \end{cases}$$

and let S denote the number of matches obtained by the time the bag is empty. Obtain expressions for $E(I_i)$ and $\text{Var}(I_i)$, and show that

$$\text{Cov}(I_i, I_j) = \frac{1}{N^2(N-1)}, \qquad i \neq j.$$

Hence show that

$$E(S) = \text{Var}(S) = 1.$$ **[8]**

4. (i) Define the P generating function (PGF) $G_X(s)$ of a count random variable X. If X has the geometric distribution

$$\mathrm{P}(X = x) = pq^{x-1}, \qquad x = 1, 2, ...; \quad p + q = 1, \qquad (*)$$

show that

$$G_X(s) = \frac{ps}{1 - qs}, \qquad |qs| < 1.$$ **[4]**

(ii) Consider a sequence of independent Bernoulli trials, each with probability of success p, and let Z be the number of trials required for r successes to occur. Explain why

$$Z = X_1 + X_2 + \cdots + X_r,$$

where $X_1, ..., X_r$ are independent random variables, each with the distribution (*) in part (i). Obtain an expression for $G_Z(s)$; then use it to obtain $\mathrm{E}(Z)$ and to show that

$$\mathrm{P}(Z = z) = \binom{z-1}{r-1} p^r q^{z-r}, \qquad z = r, r+1, ...$$

[*Hint*: $\dfrac{1}{(1-a)^r} = \sum_{i=0}^{\infty} \binom{i+r-1}{i} a^i, \qquad |a| < 1.$] **[7]**

(iii) In a simple branching process, the family sizes are independent and identically distributed random variables, each with mean μ and PGF $G(s)$, X_n denotes the size of the n^{th} generation, and the initial population X_0 is 1.

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

$$\mathrm{E}(X_n) = \mu \mathrm{E}(X_{n-1}) = \mu^n.$$

Define the probability of ultimate extinction, e, and state (without proof) how e can be derived from $G(s)$. Determine e in the case where the family size distribution is

$$\mathrm{P}(C = k) = \begin{cases} \frac{1}{5}, & k = 0 \\ \frac{2}{5}, & k = 1, 2 \\ 0, & \text{otherwise.} \end{cases}$$ **[9]**

5. (i) Balls are randomly distributed, one at a time, among N cells. The system is in state k at time n if exactly k cells are occupied after the n^{th} ball has been distributed. Explain why this system is a homogeneous Markov chain, and give its transition probability matrix. **[4]**

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

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

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

(a) Find $\mathrm{P}(X_0 = 1, X_1 = 2, X_2 = 0)$. **[2]**

(b) Find the absolute probability distribution at time $n = 2$. **[3]**

(c) Say why a unique limiting distribution $\boldsymbol{\pi}$ exists, and determine it. **[6]**

(iii) A finite Markov chain has a set A of absorbing states and a set T of transient states, and its transition probability matrix is $\mathbf{P} = (p_{ij})$. When starting from the transient state i, let f_{ik} denote the probability that the system eventually enters the absorbing state k, and μ_i the mean time for absorption to occur (in any $k \in A$).

Derive a set of linear equations satisfied by the $\{f_{ik}\}$. Write down (without derivation) a set of linear equations for the $\{\mu_i\}$. **[5]**

6. (i) A continuous non-negative random variable X is distributed Gamma(α, λ), with p.d.f.

$$f(x) = \frac{\lambda^\alpha x^{\alpha-1}\text{exp}(-\lambda x)}{\Gamma(\alpha)}, \qquad x \geq 0; \quad \alpha, \lambda > 0,$$

(where the function

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

has the properties

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

(a) Describe how the shape of $f(x)$ depends on the value of α (four cases can be distinguished), and comment briefly on the modelling implications. **[4]**

(b) Obtain an expression for $\mathrm{E}(X^r)$ and derive expressions for the mean μ and variance σ^2: also determine the mode for the case where $\alpha > 1$. Deduce μ and σ^2 for the χ^2 distribution with r degrees of freedom, which has p.d.f.

$$f(x) = \frac{1}{2^{r/2}\Gamma(r/2)}x^{\frac{1}{2}r-1}e^{-\frac{1}{2}x}, \qquad x \geq 0; \quad r \text{ a positive integer.}$$ **[7]**

(ii) If X is a standard Cauchy random variable with p.d.f.

$$f_X(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty,$$

show that $Y = 1/X$ is also a standard Cauchy random variable. **[3]**

(iii) A continuous random variable X, defined over $(-\infty, \infty)$, has c.d.f. $F_X(x)$ and p.d.f. $f_X(x)$. If $Y = X^2$, describe briefly two methods whereby the p.d.f. $f_Y(y)$ may be derived. Using either method, obtain an expression for $f_Y(y)$ in terms of f_X. Hence show that, if $X \sim N(0,1)$, i.e.

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}, \qquad -\infty < x < \infty,$$

then $Y \sim \chi^2(1)$. **[6]**

7. (i) Let X and Y be independent continuous random variables, with p.d.f.s

$$f_X(x) = \frac{1}{x^2}; \quad x \geq 1 \quad : \quad f_Y(y) = \frac{1}{y^2}, \quad y \geq 1,$$

and let $U = XY, \quad V = X/Y.$ Show that the joint p.d.f. of U, V is

$$f_{U,V}(u,v) = \frac{1}{2u^2 v}, \qquad \frac{1}{u} \leq v \leq u, \quad u \geq 1,$$

and derive the marginal p.d.f. of V. **[7]**

(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. **[8]**

(iii) Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the order statistic random variables associated with random samples of size n from a distribution with p.d.f. $f(x)$ and c.d.f. $F(x)$, where $-\infty < x < \infty$, and let $f_{(i)}(x)$ denote the p.d.f. of $X_{(i)}$. Outline one proof of the expression

$$f_{(i)}(x) = \frac{n!}{(i-1)!(n-i)!}\{F(x)\}^{i-1}\{1-F(x)\}^{n-i} f(x), \quad -\infty < x < \infty.$$

For a random sample of size 3 from the uniform distribution with

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

find the probability that the median $X_{(2)}$ is less than $\frac{1}{3}$. **[5]**

8. (i) Define the moment generating function (MGF) $M_X(\theta)$ of a continuous random variable X in terms of its p.d.f. $f_X(x), -\infty < x < \infty$. State how moments of $f_X(x)$ about the origin can be derived from M_X. **[3]**

If $X \sim \chi^2_r$, i.e.

$$f_X(x) = \frac{1}{2^{r/2}\Gamma(r/2)} x^{\frac{1}{2}r-1} e^{-\frac{1}{2}x}, \qquad x \geq 0; \quad r \text{ a positive integer},$$

show that

$$M_X(\theta) = (1-2\theta)^{-r/2}, \qquad \theta < 1/2, \qquad (**)$$

and hence derive E(X) and Var(X).

[*Note*: see Question 6 for the definition and properties of $\Gamma(p)$. You may require the expansion

$$(1-x)^{-q} = 1 + qx + \frac{q(q+1)}{2!}x^2 + \cdots; \quad |x| < 1, q > 0.]$$ **[6]**

(ii) If X and Y are independent continuous random variables, show how $M_{X+Y}(\theta)$ is related to $M_X(\theta)$ and $M_Y(\theta)$, and state the generalisation to n independent random variables $X_1, ..., X_n$. **[3]**

The random variables $Z_1, ..., Z_n$ are independent and

$$Z_i \sim N(0,1), \qquad i = 1, ..., n.$$

Show that

$$M_{Z_i^2}(\theta) = (1-2\theta)^{-1/2}, \qquad i = 1, ..., n; \quad \theta < 1/2.$$

If

$$V_n = Z_1^2 + \cdots + Z_n^2,$$

obtain an expression for the MGF of V_n, and, using the result (**) in part (i), deduce the distribution of V_n. **[5]**

By appeal to the central limit theorem, deduce a convenient approximation to the distribution of V_n when n is large. **[3]**

9. (i) Explain what is meant by the assertion that a counting process $\{N(t), t \geq 0\}$ has *independent* and *stationary* (or *time-homogeneous*) increments.

State the conditions which characterise a Poisson process with rate λ. For such a process, state (without proof) the probability distributions of

(a) $N(u+t) - N(u)$ for $u \geq 0, t > 0$;

(b) $X_n = T_n - T_{n-1}$ for $n \geq 1$, where T_n is the time at which the nth event after time $t = 0 = T_0$ occurs;

(c) $T_n - T_{n-r}$ for $r \geq 1, n \geq r$. **[8]**

(ii) In a birth and death process $\{X(t),\ \ t \geq 0\}$, the transition probability functions $\{p_{ij}(t)\}$ are such that, for small δt,

$$p_{ij}(\delta t) = \begin{cases} \alpha_i \delta t + o(\delta t), & i \geq 0, j = i+1 \\ \beta_i \delta t + o(\delta t), & i \geq 1, j = i-1 \\ 1 - (\alpha_i + \beta_i)\delta t + o(\delta t), & i \geq 0, j = i \\ o(\delta t), & \text{otherwise,} \end{cases}$$

where $\beta_0 \equiv 0$. Show that the probabilities $p_n(t) \equiv \mathrm{P}(X(t) = n)$ satisfy the equations

$$\begin{aligned} \frac{dp_n(t)}{dt} &= -(\alpha_n + \beta_n)p_n(t) + \alpha_{n-1}p_{n-1}(t) + \beta_{n+1}p_{n+1}(t), \quad n \geq 1 \\ \frac{dp_0(t)}{dt} &= -\alpha_0 p_0(t) + \beta_1 p_1(t). \end{aligned}$$

Assuming that $\{\alpha_n\}$ and $\{\beta_n\}$ are such that a steady-state solution $\{\pi_m : m = 0, 1, ...\}$ exists, show that

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

and deduce that

$$\pi_0 = \left(1 + \frac{\alpha_0}{\beta_1} + \frac{\alpha_0\alpha_1}{\beta_1\beta_2} + \cdots\right)^{-1}.$$ **[9]**

(iii) Consider a single-server queue with discouragement, in which, if n is the number of customers in the system, the arrival and service rates are respectively

$$\lambda_n = \frac{\lambda}{n+1}, \qquad n \geq 0$$

and

$$\mu_n = \mu, \qquad n \geq 1.$$

Using the results in part (ii), show that the steady-state distribution is Poisson with parameter λ/μ. **[3]**