THE QUEEN'S UNIVERSITY OF BELFAST

110SOR201

Level 2 Examination

Statistics and Operational Research 201

Probability and Distribution Theory

Wednesday 15 August 2001 14.30 pm — 17.30 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) (a) State a set of fundamental axioms concerning the probability measure P in a probability space $(\mathcal{S}, \mathcal{F}, \mathrm{P})$. Indicate (without proof) any redundancy in the stated axioms. **[4]**

(b) Using the axioms, together with the addition law of probability for two events (which can be derived from the axioms), show by induction that, for n events $A_1, ..., A_n \in \mathcal{F}$,

$$\mathrm{P}(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^{n} \mathrm{P}(A_i).$$ **[5]**

(c) 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. **[2]**

(ii) A janitor hangs n keys, numbered $1, ..., n$, at random on n similarly numbered hooks, one key to each hook. Explaining your reasoning carefully, obtain an expression for the probability that no key is hung on a hook with the same number, and deduce a good approximation to this probability when n is large. **[9]**

2. (i) (a) Given a probability space $(\mathcal{S}, \mathcal{F}, \mathrm{P})$, explain what is meant by the assertion that two events $A, B \in \mathcal{F}$ are *independent*. Show that, if A, B are independent, then so too are the complementary events $\overline{A}, \overline{B}$. **[4]**

(b) For events $A_1, A_2, ..., A_n \in \mathcal{F} \quad (n \geq 3)$, explain the distinction between the property of *pairwise independence* and that of *mutual* (or *complete*) *independence*. **[3]**

(ii) State carefully, and prove, the *law of total probability* (or partition rule). **[4]**

(iii) A biased coin is such that the probability of getting a head in a single toss is p. Suppose that the coin is tossed n times.

(a) Let u_n denote the probability that an even number of heads is obtained (0 being regarded as an even number). Obtain a recurrence relation for u_n and show, by induction or otherwise, that

$$u_n = \tfrac{1}{2}[1 + (1 - 2p)^n], \quad n \geq 1.$$ **[5]**

(b) Let v_n denote the probability that two successive heads are *not* obtained, and define the events

T_i: first tail obtained on the i^{th} toss $(i = 1, 2, ...)$.

By conditioning on the $\{T_i\}$, or otherwise, show that

$$v_n = (1 - p)v_{n-1} + p(1 - p)v_{n-2}, \quad n \geq 2,$$

and indicate how v_n can be determined for given n and p. **[4]**

3. (i) The discrete random variables X, Y are independent, and each has the geometric distribution with parameter p, i.e.

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

(a) Determine the distribution of the random variable $Z=X+Y$. **[4]**

(b) Let $V=\max(X,Y)$. By first considering $\mathrm{P}(V\leq v)$, or otherwise, determine the distribution $\mathrm{P}(V=v), v=1,2,...$ **[4]**

(ii) (a) Let (X,Y) be discrete random variables with joint probability function $\{\mathrm{P}(X=x,Y=y) : x=x_1,x_2,...;y=y_1,y_2,...\}$. Define $\mathrm{E}(X|Y=y_j)$, and prove that

$$\mathrm{E}[\mathrm{E}(X|Y)]=\mathrm{E}(X).$$ **[5]**

(b) Consider a sequence of independent Bernoulli trials, each with probability of success p. Let

$$X_r = \text{number of trials required to obtain } r \text{ successes;}$$
$$Y = \begin{cases} 1 & \text{if the first trial yields a success} \\ 0 & \text{if the first trial yields a failure.} \end{cases}$$

Explain why

$$\mathrm{E}(X_r|Y=0)=\mathrm{E}(X_r)+1, \qquad r\geq 1,$$

and give a similar relation for $\mathrm{E}(X_r|Y=1)$. Hence obtain a simple recurrence relation for $\mathrm{E}(X_r)$ and deduce that $\mathrm{E}(X_r)=\dfrac{r}{p}$. **[4]**

(iii) A supermarket issues N different types of prize coupons to customers: each coupon issued is equally likely to be one of the N types. Suppose that a customer has collected n coupons. Let

$$X_i = \begin{cases} 1, & \text{if there is at least one type } i \text{ coupon in the set} \\ 0, & \text{otherwise;} \end{cases}$$
$$X = \text{number of \textit{different} types of coupon in the set.}$$

Find $\mathrm{E}(X_i)$ and hence $\mathrm{E}(X)$. **[3]**

4. (i) Define the P generating function (PGF) $G_X(s)$ of a count random variable X. If $G_X(s)$ is known, indicate how $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ can be found. If $Y = a + bX$, express the PGF of Y in terms of G_X. **[5]**

(ii) If $X = \sum_{i=1}^{n} X_i$, where the $\{X_i\}$ are independent count random variables, state how $G_X(s)$ is related to the PGFs $G_1(s), \ldots, G_n(s)$ of $X_1, \ldots, X_n$.

Let X be the total score obtained in 3 rolls of a fair die. Show that

$$G_X(s) = \frac{s^3(1-s^6)^3}{6^3(1-s)^3}$$

and derive the value of $\mathrm{P}(X = 14)$. **[6]**

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

(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) Given a sequence of random variables $X_0, X_1, \ldots$ defined on a state space $\{0, 1, \ldots\}$, explain what is meant by the assertion that $\{X_n : n = 0, 1, \ldots\}$ is a *homogeneous Markov chain*, and define the transition probability matrix $\mathbf{P}$.

Show that

$$\mathbf{p}^{(n)} = \mathbf{p}^{(0)}\mathbf{P}^n,$$

where $\mathbf{p}^{(r)}$ denotes the row vector $(\mathrm{P}(X_r = 0), \mathrm{P}(X_r = 1), \ldots)$, and that

$$\mathrm{P}(X_n = j | X_0 = i) = p_{ij}^{(n)},$$

the (i, j) element of $\mathbf{P}^n$. **[6]**

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

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

At time $n = 0$, the system is equally likely to be in states 0, 1 or 2.

(a) Find $\mathrm{P}(X_2 = 2)$. **[3]**

(b) Quote a theoretical result which confirms that a limiting distribution $\boldsymbol{\pi}$ exists in this case, and determine $\boldsymbol{\pi}$. **[6]**

(iii) An absorbing Markov chain has states $\{0, 1, 2, 3, 4\}$ and transition P matrix

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

Let f_{i1} denote the probability that the system eventually enters the absorbing state 1, given that it started in the transient state i. Write down (without proof) a set of equations for $\{f_{i1}\}$, and hence determine f_{01}. **[5]**

6. (i) Define the median of a continuous random variable X with p.d.f. $f(x)$, $-\infty < x < \infty$. If $f(x)$ is symmetrical about $x = a$, i.e.

$$f(a+y) = f(a-y), \qquad y > 0,$$

show that the median is a. **[4]**

(ii) Suppose that $Z \sim N(0,1)$, with p.d.f.

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

and let $V = Z^2$. Show that $V \sim \chi^2(1)$.
Determine $\mathrm{E}(V^2)$ and deduce the fourth moment about the mean of $N(\mu, \sigma^2)$. **[9]**
[*Note*: the p.d.f. for the $\chi^2(r)$ distribution is

$$f_V(v) = \frac{1}{2^{r/2}\Gamma(r/2)} v^{r/2-1} e^{-v/2}, \qquad 0 \le v < \infty, r \text{ a positive integer},$$

and

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

has the properties

$$\Gamma(p+1) = p\Gamma(p) : \quad \Gamma(1/2) = \sqrt{\pi}. \qquad]$$

(iii) If $X \sim$ beta(a,b), with p.d.f.

$$f_X(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \qquad 0 \le x \le 1; \quad a, b > 0,$$

obtain expressions for $\mathrm{E}(X)$ and $\mathrm{Var}(X)$, and show that $Y = 1 - X$ is also beta distributed. **[7]**

[*Note*: $\quad B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.]

7. (i) The random variables X,Y are independent and have the same negative exponential distribution:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise;} \end{cases} \qquad f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that the random variables $U = \dfrac{Y}{X}, V = X + Y$ are independent.

(b) Show that V has the Erlang$(2, \lambda)$ distribution.
[*Note*: the p.d.f. for the Erlang(n, λ) distribution is

$$f(x) = \frac{\lambda^n x^{n-1} \text{exp}(-\lambda x)}{(n-1)!}, \qquad x \geq 0; \quad \lambda > 0, n \text{ integer} \geq 1.]$$

(c) Find the distribution of U. **[8]**

(ii) Suppose that X and Y are independent continuous random variables with p.d.f.s $f_X(x)$ and $f_Y(y)$ respectively. By considering a suitable bivariate transformation, show that the p.d.f. of $U = \dfrac{Y}{X}$ can be expressed as

$$f_U(u) = \int_{-\infty}^{\infty} f_X(v) f_Y(uv)|v|dv.$$

If X and Y are both uniformly distributed on $[0, 1]$, deduce that

$$f_U(u) = \begin{cases} \dfrac{1}{2}, & 0 \leq u \leq 1 \\ \dfrac{1}{2u^2}, & 1 \leq u < \infty. \end{cases}$$ **[6]**

(iii) Let $X_{(1)}, X_{(2)}, ..., X_{(n)}$ 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. $F(x)$, $-\infty < x < \infty$. Show that the p.d.f. of $X_{(n)}$ is

$$f_{(n)}(x) = n\{F(x)\}^{n-1} f(x), \qquad -\infty < x < \infty.$$

Then indicate *briefly* how the argument you have used can be extended to yield the p.d.f. of $X_{(i)} (i = 1, ..., n)$ and, as an illustration, show that

$$f_{(n-1)}(x) = n(n-1)\{F(x)\}^{n-2}\{1 - F(x)\}f(x), \qquad -\infty < x < \infty.$$ **[6]**

8. (i) Define the moment generating function (MGF) $M_X(\theta)$ of a continuous random variable X, and state how the MGF of $Y = a + bX$ is related to M_X. Indicate concisely two methods whereby moments of X about the origin can be derived from $M_X(\theta)$. How can these procedures be modified to yield (directly) moments about the mean $\mathrm{E}(X)$? **[5]**

(ii) If $Z \sim N(0, 1)$, show that

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

and hence obtain $M_X(\theta)$, where $X = \mu + \sigma Z$. Use $M_X(\theta)$ to verify that $E(X) = \mu$ and $\mathrm{Var}(X) = \sigma^2$. **[7]**

(iii) Show that if the distribution of X is negative exponential with parameter λ (see Question 7(i) for definition), then

$$M_X(\theta) = \frac{\lambda}{\lambda - \theta}, \qquad \theta < \lambda.$$

Hence obtain the MGF of

$$S_n = \sum_{i=1}^{n} X_i,$$

where $X_i (i = 1, \ldots, n)$ are independent random variables, each exponentially distributed with parameter λ.

Show that

$$Z_n = \frac{S_n - n/\lambda}{\sqrt{n}/\lambda}$$

is asymptotically normally distributed with mean 0 and variance 1.

(*Hint*:

$$\lim_{n\to\infty} \left(1 + \frac{a}{n} + o(1/n)\right)^n = e^a \text{ for fixed } a. \qquad)$$ **[8]**

9. (i) State the assumptions which characterize a counting process $\{N(t), t \geq 0\}$ as a *Poisson process* with rate λ.

Show from the assumptions that the probabilities

$$p_n(t) = \text{P}[N(t) = n]$$

satisfy the equations

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

Indicate *briefly* how these equations can be solved and quote the resulting expression for $p_n(t)$. Also state (without proof) the distribution of the inter-event times

$$X_n = T_n - T_{n-1},$$

where

$$T_n = \text{inf}\{t : N(t) = n\}.$$ **[11]**

(ii) (a) For a 'birth and death' process $\{X(t), \ t \geq 0\}$ with 'birth' rates $\{\alpha_i; i = 0, 1, \ldots\}$ and 'death' rates $\{\beta_i; i = 1, 2, \ldots\}$, it can be shown that the probabilities

$$p_n(t) = \text{P}[X(t) = n], \qquad n = 0, 1, \ldots$$

satisfy the equations

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

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

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

and deduce that

$$\pi_0 = \left(1 + \frac{\alpha_0}{\beta_1} + \frac{\alpha_0\alpha_1}{\beta_1\beta_2} + \cdots\right)^{-1}.$$ **[5]**

(b) Consider a single-server queueing system in which the service time is negative exponential with mean μ^{-1} and customer arrivals form a Poisson process with rate λ, except that any customer arriving when there are already N customers in the system leaves without joining the queue. Show that the steady-state distribution of the number of customers in the system is

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right) \left\{1 - \left(\frac{\lambda}{\mu}\right)^{N+1}\right\}^{-1}, \quad 0 \leq n \leq N.$$

Indicate briefly how your discussion would be affected if the single server were replaced by c similar servers working independently in parallel. **[4]**