THE QUEEN'S UNIVERSITY OF BELFAST

110SOR201

Level 2 Examination

Statistics and Operational Research 201

Probability and Distribution Theory

Wednesday 18 January 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 basic properties of the probability function P defined on the event space $\mathcal{F}$ of a probabilistic experiment with sample space $\mathcal{S}$. **[4]**

For two events $A, B \in \mathcal{F}$, deduce *from the axioms* that

(a) $\mathrm{P}(A) \le \mathrm{P}(B)$ if $A \subset B$;

(b) $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$. **[5]**

(ii) State (without proof) the generalisation of the result in part (i)(b) to n events $A_1, ..., A_n \in \mathcal{F}$. **[2]**

Two players are each given a well-shuffled standard pack of 52 playing cards. Each player deals out cards from their pack, one by one, until the pack is exhausted. A *match* A_i is said to occur when the i^{th} card dealt out by player 2 is the same (in both suit and denomination) as the i^{th} card dealt out by player 1. Obtain an expression for the probability p_0 that *no* matches occur, and give a convenient approximation to its value. **[6]**

By enumerating 'favourable' outcomes, or otherwise, show that p_1, the probability that exactly *one* match occurs, is approximately equal to p_0. **[3]**

2. (i) (a) Given a probability space $(S, \mathcal{F}, \mathrm{P})$, 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 $A_1, A_2, ..., A_n \in \mathcal{F}$ $(n \ge 3)$, 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

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

(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

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

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

3. (i) The count random variables X and Y are independent and Poisson distributed with parameters λ and μ respectively, i.e.

$$\mathrm{P}(X=k)=\frac{\lambda^k e^{-\lambda}}{k!}, \quad \mathrm{P}(Y=k)=\frac{\mu^k e^{-\mu}}{k!}, \quad k \geq 0.$$

Show that $Z = X + Y$ is Poisson distributed with parameter $(\lambda + \mu)$. Show also that the conditional distribution of X, given that $X + Y = n$, is binomial, and determine the parameters. **[8]**

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

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

(b) A prisoner is trapped in a dark cell containing three doors. Doors 1 and 2 lead to tunnels which return the prisoner to the cell after a travel time of 10 hours and 15 hours respectively: door 3 leads to freedom after 12 hours. If it is assumed that the prisoner will always select doors 1,2,3 with probabilities 0.3, 0.2, 0.5 respectively, what is the expected time for the prisoner to reach freedom? **[3]**

(iii) Cards are drawn at random, one by one, from a standard pack of 52 cards. How many cards would one expect to have to draw in order to obtain a king? **[4]**

[*Hint*: number the other (non-king) cards 1,...,48; then define

$$I_i = \begin{cases} 1, & \text{if card } i \text{ is drawn before any king} \\ 0, & \text{otherwise} \end{cases}$$

and explain why $\mathrm{P}(I_i = 1) = \frac{1}{5}$.]

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

$$p_k \equiv \mathrm{P}(X = k) = pq^k, \qquad k = 0, 1, ...; \ \ p + q = 1, \qquad (*)$$

show that

$$G_X(s) = \frac{p}{1 - qs}, \qquad |qs| < 1,$$

and hence obtain $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **[5]**

(ii) If $X_1, ..., X_n$ are independent count random variables with PGFs $G_1(s), ..., G_n(s)$ respectively, and $X = \sum_{i=1}^{n} X_i$, state how $G_X(s)$ is related to the $\{G_i(s)\}$.
If X is the 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 deduce the value of $\mathrm{P}(X = 13)$. **[5]**

/continued...

(iii) Consider a simple branching process in which the family sizes are i.i.d. random variables, each with mean μ and PGF $G(s)$. Let X_n denote the size of the n^{th} generation, with PGF $G_n(s)$, and suppose that the initial population X_0 is 1.

(a) Quoting the relevant Theorem, explain why

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

and deduce that

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

(b) Define the probability of ultimate extinction, e, and state (without proof) how e can be derived from $G(s)$. If the family size distribution is modified geometric with parameter p (see (*) above), determine e as a function of p. **[4]**

5. (i) If $X_0, X_1, \ldots$ is a sequence of random variables defined on a state space $\{0, 1, \ldots\}$, explain what is meant by saying that $\{X_n : n = 0, 1, \ldots\}$ is a *homogeneous Markov chain* with 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 explain the significance of $p_{ij}^{(n)}$, 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

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{2}{5} & \frac{3}{5} \\ \frac{1}{5} & \frac{2}{5} & \frac{2}{5} \\ \frac{4}{5} & 0 & \frac{1}{5} \end{pmatrix}.$$

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

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

(b) Explain briefly why a limiting distribution $\boldsymbol{\pi}$ exists, and determine it. **[6]**

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

Write down (without proof) sets of linear equations for the $\{f_{ik} : i \in T\}$ and the $\{\mu_i : i \in T\}$. If an absorbing chain with states $\{0, 1, 2, 3\}$ has

$$\mathbf{P} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{3}{5} & 0 & \frac{3}{10} & \frac{1}{10} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

determine f_{02} and μ_0. **[6]**

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

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

(where the function

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

has the properties

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

Obtain an expression for $\mathrm{E}(X^r)$ and deduce expressions for the mean μ, variance σ^2 and coefficient of skewness γ_1 (defined as $\mathrm{E}\{(X-\mu)^3/\sigma^3\}$): also determine the mode for the case where $\alpha > 1$. Comment briefly on the usefulness of the Gamma distribution for modelling data. **[10]**

(ii) A continuous random variable X has p.d.f.

$$f_X(x) = \begin{cases} \frac{3}{8}(1+x)^2, & -1 \leq x \leq 1 \\ 0, & \text{elsewhere.} \end{cases}$$

Using any suitable method, determine the p.d.f. of $Y = X^2$. **[6]**

(iii) The life-time X of a certain device has the 2-parameter Weibull distribution, with c.d.f.

$$F(x) = 1 - \exp\left\{-\left(\frac{x}{b}\right)^c\right\}, \quad x \geq 0.$$

Determine the p.d.f. $f(x)$ and the hazard rate function $r(x) = f(x)/\{1-F(x)\}$. What is the significance of $r(x)$ and how does its behaviour depend on the value of c? **[4]**

7. (i) Suppose that the independent random variables X,Y are each exponentially distributed with parameter λ, i.e. X has the p.d.f.

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

with a similar expression for $f_Y(y)$. Show that the random variables

$$U = \frac{Y}{X}, \quad V = X + Y$$

are independent, and that the distribution of V is Gamma(2,λ). What is the distribution of U? **[8]**

[*Note*: see Question 6(i) for the definition of the Gamma distribution.]

/continued...

(ii) Let $X_{(1)}, ..., X_{(n)}$ 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. $F(x)$, $-\infty < x < \infty$.

(a) Show that the p.d.f. of the smallest observation, $X_{(1)}$, is

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

Hence determine $\mathrm{E}(X_{(1)})$ and $\mathrm{Var}(X_{(1)})$ in the case of sampling from a uniform distribution with

$$f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$$ **[7]**

[*Note*: $\displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}dx = B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a, b > 0,$
where the Gamma function is defined in Question 6(i).]

(b) It may be shown that the joint p.d.f. of $X_{(1)}, X_{(n)}$ is

$$f_{(1),(n)}(x,y) = n(n-1)\{F(y) - F(x)\}^{n-2} f(x) f(y), \quad -\infty < x < y < \infty.$$

By introducing a suitable transformation, show that, when sampling is from the uniform distribution in (a), the sample range R has the p.d.f.

$$f_R(r) = n(n-1)r^{n-2}(1-r), \quad 0 \le r \le 1.$$ **[5]**

8. (i) (a) Define the moment generating function (MGF) $M_X(\theta)$ of a continuous random variable X in terms of its p.d.f. $f_X(x), -\infty < x < \infty$, and indicate concisely two ways in which moments of $f_X(x)$ about the origin may be derived from $M_X(\theta)$. Express the MGF of $Y = aX + b$ in terms of M_X. **[4]**

(b) If X is exponentially distributed with parameter λ (see Question 7(i)), show that

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

and hence determine $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. **[5]**

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

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

and deduce $M_X(\theta)$, where $X \sim N(\mu, \sigma^2)$. **[5]**

(b) If $X_1, ..., X_n$ are independent random variables, and $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, ..., n$, use MGFs to prove that, if $W = \sum_{i=1}^n a_i X_i$, then

$$W \sim N\left(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2\right)$$ **[3]**

(iii) State the central limit theorem. **[3]**

9. (i) In a Poisson process with rate λ, let $N(t)$ be the number of events occurring in the time interval $(0,t]$ and W_r the time interval from $t=0$ to the r^{th} event.

(a) Write down an expression for $\mathrm{P}[N(t_2)-N(t_1)=n]$, where $t_1<t_2$, commenting briefly on how t_1 and t_2 enter the expression. **[2]**

(b) By using the fact that $W_r \leq w$ if and only if $N(w) \geq r$, or otherwise, show that W_r is Gamma distributed with parameters (r,λ), i.e. its p.d.f. is

$$f_{W_r}(w)=\frac{\lambda^r w^{r-1}e^{-\lambda w}}{(r-1)!}, \qquad w \geq 0.$$ **[5]**

(ii) (a) State the assumptions concerning the transition probabilities $p_{ij}(t)$ which characterise a birth-and-death process $\{X(t), t \geq 0\}$ with birth rates $\{\alpha_i : i = 0,1,...\}$ and death rates $\{\beta_i : i = 1,2,...\}$, and draw the corresponding transition rate diagram.

It can be shown that, under these assumptions, the probabilities

$$p_n(t)=\mathrm{P}(X(t)=n)$$

satisfy the equations

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

where, for convenience, $\alpha_{-1}=\beta_0=0$ have been added. Assuming that $\{\alpha_n\}$ and $\{\beta_n\}$ are such as to ensure the existence of a steady-state distribution $\{\pi_m : m=0,1,\ldots\}$, show that

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

and deduce that

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

(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

$$\pi_n=\rho^n(1-\rho)(1-\rho^{N+1})^{-1}, \quad 0 \leq n \leq N,$$

where $\rho=\lambda/\mu$. **[4]**