THE QUEEN'S UNIVERSITY OF BELFAST

110SOR201

Level 2 Examination

Statistics and Operational Research 201

Probability and Distribution Theory

Wednesday 14 August 2002 2.30 pm — 5.30 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) State (without proof) the generalized addition law (or inclusion-exclusion principle) for n events $A_1, A_2, \ldots, A_n \in \mathcal{F}$.

   If the digits $1, 2, \ldots, 6$ are written in random order, show that the probability that no digit is in its natural position is

   $$\left[\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!}\right] = \frac{265}{720}.$$

   Deduce the number of permutations of the letters of the word RANDOM which have no letter in its correct position. Then, by enumerating 'favourable' permutations, find the probability that, when the letters of the word RANDOMIZE are written in random order, exactly 3 letters are in their correct positions. **[11]**

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 $A_1, A_2, ..., A_n \in \mathcal{F} \quad (n \geq 3)$, 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 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 14 hours and 10 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 0.5, 0.2, 0.3 respectively, what is the expected time for the prisoner to reach freedom? **[3]**

(iii) Each wrapper on a certain chocolate bar is equally likely to bear one of 4 special symbols. Introducing suitable indicator variables, show that the expected number of distinct symbols on a set of n wrappers is $4\left[1-\left(\frac{3}{4}\right)^n\right]$. **[4]**

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]**

/continued...

(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) 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*. Introduce the transition probability matrix $\mathbf{P}$ and prove that

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

where $\mathbf{p}^{(r)}$ denotes the row vector $(\mathrm{P}(X_r = 0), \mathrm{P}(X_r = 1), \ldots)$. **[5]**

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

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

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

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

(b) Find $\mathrm{P}(X_2 = 1)$ and $\mathrm{P}(X_2 = 2)$. **[3]**

(c) By appeal to Markov's Theorem, explain why a limiting distribution $\boldsymbol{\pi}$ exists, and determine it. **[5]**

(iii) Classify the states $\{0, 1, 2, 3, 4, 5\}$ of a Markov chain with

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{pmatrix}.$$ **[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).

(b) Obtain an expression for E(X^r) and derive expressions for the mean μ, variance σ^2 and coefficient of skewness $\gamma_1 = \text{E}(\{X-\mu\}^3)/\sigma^3$. Deduce μ, σ^2 and γ_1 for the χ^2 distribution with r degrees of freedom, which has p.d.f.

$$f(x) = \frac{1}{2^{r/2}\Gamma(\frac{1}{2}r)}x^{\frac{1}{2}r-1}e^{-\frac{1}{2}x}, \qquad x \geq 0.$$ **[11]**

(ii) Let $Z \sim N(0,1)$, i.e.

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

Show that $V = Z^2$ has the χ^2 distribution with 1 degree of freedom. **[5]**

(iii) The 2-parameter Weibull distribution has the c.d.f.

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

Determine the p.d.f. $f(x)$ and the hazard rate function $r(x) = f(x)/\{1 - F(x)\}$, discussing briefly how the behaviour of $r(x)$ depends on the value of the parameter c. **[4]**

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 Gamma$(2, \lambda)$ distribution (see Question 6(i) for the Gamma p.d.f.).

(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) (a) Define the moment generating function (MGF)$M_X(\theta)$ of a continuous random variable X in terms of its p.d.f. $f(x), -\infty < x < \infty$, and state how the MGF may be used to obtain moments about the origin. Express the MGF of $Y = aX + b$ in terms of M_X. **[4]**

(b) If X and Y are independent random variables, show that $M_{X+Y}(\theta)$ can be expressd in terms of $M_X(\theta)$ and $M_Y(\theta)$. **[2]**

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

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

If $X \sim N(\mu, \sigma^2)$, deduce $M_X(\theta)$ and use it to verify that

$$\text{E}(X) = \mu, \text{Var}(X) = \sigma^2.$$ **[7]**

(b) If $X_1, \ldots, X_n$ are independent random variables, and

$$X_i \sim N(\mu_i, \sigma_i^2), \qquad i = 1, \ldots, n,$$

use MGFs to prove that

$$W = \sum_{i=1}^{n} a_i X_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2).$$

Deduce the distribution of the sample mean $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ in the case where all $X_i \sim N(\mu, \sigma^2)$. **[4]**

(iii) State the central limit theorem. **[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 $\{X(t), \quad t \geq 0\}$, 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]**