UNIVERSITY OF SURREY©

B. Sc. Undergraduate Programmes in Mathematical Studies
M. Math. Undergraduate Programmes in Mathematical Studies

Level HE2 Examination

Module MS237 MATHEMATICAL STATISTICS

Time allowed – 2 hours

Spring Semester 2006

Attempt THREE questions. If any candidate attempts more than THREE questions only the best THREE solutions will be taken into account.
A formula sheet will be provided.
Cambridge Statistical Tables will be provided.

Question 1

(a) (i) Define the probability generating function for a random variable X. Using this definition obtain an expression for $E[X]$ and for $\text{Var}[X]$. [5]

(ii) Also obtain an expression for the probability generating function for the sum of two independent random variables X_1 and X_2. [3]

(b) Let X_1 and X_2 be independently distributed Poisson random variables with expectations λ_1 and λ_2 respectively.

(i) Derive the probability generating functions of X_1 and X_2.
Hence, or otherwise, show that $Z = X_1 + X_2$ has a Poisson distribution with expectation $\lambda_1 + \lambda_2$. [7]

(ii) Show that if k, n are positive integers such that $k \leq n$ then

$$P\{(X_1 = k) \cap (Z = n)\} = P(X_1 = k)P(X_2 = n - k).$$

[3]

(iii) Hence, or otherwise, show that the conditional probability

$$P(X_1 = k \,|\, Z = n)$$

is the probability of k successes in n Bernouilli trials, where each trial has probability of success $\lambda_1/(\lambda_1 + \lambda_2)$. [7]

Question 2
Random variables X and Y have the joint probability density function

$$\begin{aligned} f_{X,Y}(x,y) &= kx \qquad x^2 < y < x, \quad 0 < x < 1, \quad 0 < y < 1; \\ &= 0 \qquad \text{elsewhere.} \end{aligned}$$

(a) Sketch the region in which $f_{X,Y}(x,y)$ is non-zero and show that $k = 12$. [5]

(b) Show that the marginal density function of X is

$$f_X(x) = 12x^2(1-x) \qquad 0 < x < 1.$$

[3]

(c) Find the marginal density function of Y. Also, state whether X and Y are independent, giving a reason for your answer. [5]

(d) Find non-zero values for c_1 and c_2 such that $c_1E[X] + c_2E[Y] = 0$. [5]

(e) Show that the conditional density function of Y given $X = x$ is

$$f(y|X=x) = \frac{1}{x(1-x)} \qquad x^2 < y < x, \quad 0 < x < 1, \quad 0 < y < 1.$$

[3]

(f) Calculate $E[Y]$ and $E[E[Y|X]]$ and comment on your answer. [4]

Question 3

(a) (i) Prove the Cauchy-Schwartz inequality:

$$\{E[XY]\}^2 \leq E[X^2]\, E[Y^2]$$

for any two random variables X and Y that have finite variances. [6]

(ii) Use the Cauchy-Schwartz inequality to prove that

$$-1 \leq \mathrm{Corr}(X, Y) \leq 1$$

for any two random variables X and Y with finite variances. [5]

(iii) Let X be a random variable with variance σ^2 and let $Y = cX$, where c is a real constant. Show that $\{E[XY]\}^2$ attains the upper bound $E[X^2]\, E[Y^2]$ in this case. [4]

(b) (i) Outline *briefly* the necessary steps required to establish Chebyshev's inequality

$$\mathrm{pr}(|Y - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

where $a > 0$ and μ and σ are the mean and standard deviation of Y respectively. [4]

(ii) Suppose that Y is a discrete random variable with probability mass function given in the following table:

y	-3	-1	1	3
$p(y)$	$\frac{3}{32}$	$\frac{13}{32}$	$\frac{13}{32}$	$\frac{3}{32}$

Evaluate $\mathrm{pr}(|Y - \mu| \geq \sqrt{\frac{5}{2}})$ for this distribution. How does this probability compare with the upper bound given by Chebyshev's inequality? [6]

Question 4

(a) The random variable X has a Beta distribution with parameters a and b, that is, $X \sim \text{Beta}(a, b)$.

(i) Show that

$$E[X(1-X)] = \frac{ab}{(a+b)(a+b+1)}.$$

[4]

(ii) Assuming that $b > 1$, show that

$$E\left[\frac{X}{1-X}\right] = \frac{a}{b-1}.$$

[4]

(b) The random variable Y has an F distribution on m and n degrees of freedom, that is, $Y \sim F(m, n)$.

(i) Write down the probability density function of Y and hence show that

$$W = \frac{\frac{mY}{n}}{1 + \frac{mY}{n}}$$

has a $\text{Beta}(\frac{m}{2}, \frac{n}{2})$ distribution. [9]

(ii) By expressing Y in terms of W and using the results of part (a), show that $E[Y] = \frac{n}{n-2}$ for $n > 2$. [4]

(iii) Without carrying out any further calculations, explain how a similar technique could be used to obtain the variance of Y. [4]

Question 5

Let $X_1, \ldots, X_n$ be independent normally distributed random variables, $X_i \sim N(\mu, \sigma^2)$. Write $\bar{X} = n^{-1} \sum_i X_i$ and $S^2 = (n-1)^{-1} \sum_i (X_i - \bar{X})^2$, and let $\underline{X}$ denote the $n \times 1$ vector $\underline{X} = (X_1, \ldots, X_n)^T$.

(a) Show that $\underline{X}$ has the multivariate $N_n(\mu 1_n, \sigma^2 I_n)$ distribution, where $\underline{1_n}$ is defined as the $n \times 1$ vector all of whose elements equal unity. [6]

(b) Let the vector $\underline{Y} = C(\underline{X} - \mu \underline{1_n})/\sigma$, where C is an orthogonal matrix, i.e. a $n \times n$ matrix such that $CC^T = I_n$.

Show that $\underline{Y}$ has the multivariate $N_n(0, I_n)$ distribution, stating clearly the theorem that you require. [5]

(c) Assume that the first row of C consists of the vector $\underline{1_n}^T/\sqrt{n}$.

Show that $Y_1 = \sqrt{n}(\bar{X} - \mu)/\sigma$ and evaluate its distribution. Hence or otherwise, show that $\bar{X}$ has the univariate $N(\mu, \sigma^2/n)$ distribution. [4]

(d) Show that

$$\sum_{i=2}^{n} Y_i^2 \quad \text{has a } \chi^2_{n-1} \text{ distribution.}$$

stating clearly the theorem that you require. Hence, or otherwise, find the distribution of $(n-1)S^2/\sigma^2$. [3]

(e) Explain why $\bar{X}$ and S^2 are independent random variables. [3]

(f) Comment **briefly** on the distribution of

$$T = \frac{\sqrt{n}(\bar{X} - \mu)}{s},$$

stating clearly the theorem that you require. [4]

INTERNAL EXAMINER: J.D. Godolphin
EXTERNAL EXAMINER: W. Krzanowski