UNIVERSITY OF LONDON

IMPERIAL COLLEGE LONDON

BSc and MSci EXAMINATIONS (MATHEMATICS)
MAY–JUNE 2003

This paper is also taken for the relevant examination for the Associateship.

M2S1 PROBABILITY AND STATISTICS II

DATE: Friday, 16th May 2003 TIME: 2 pm – 4 pm

Credit will be given for all questions attempted but extra credit will be given for complete or nearly complete answers.

Calculators may not be used. A formula sheet is given on pages 7 & 8.

1. *a)* Consider the discrete random variable N with range $\mathbb{N} \equiv \{0, 1, 2, 3, 4\}$ and probability mass function (pmf) given by

$$f_N(n) = \begin{cases} p(1-p)^n & n = 0, 1, 2, 3; \\ (1-p)^n & n = 4, \end{cases}$$

for some constant p, $0 < p < 1$.

i) Show that this is a valid pmf.

ii) Find G_N, the probability generating function (pgf) of N.
Hence deduce that, if $p = 1/2$, then

$$G_N(t) = \frac{1}{2-t} + \left(\frac{1-t}{2-t}\right)\left(\frac{t}{2}\right)^4 .$$

b) Suppose that the random variable N in *a)* with $p = 1/2$ is the number of sets of traffic lights at which a car must stop on a daily trip to work. Suppose also that the length of each traffic light stop is a random variable having a *Gamma* distribution with parameters α and β. It can be assumed that the lengths of stops at different traffic lights are independent random variables. Finally, let S be the continuous random variable corresponding to the **total** time spent stopped at traffic lights during a daily journey.

i) Find the expectation of N.

ii) **Conditional** on $N = n$, find the expectation of S.

iii) Find the **unconditional** expectation of S.

2. *a)* Consider the continuous random variable X having an *Exponential* distribution with parameter λ.

i) Prove the **lack of memory** property for X, that is, for $0 < s < t$,

$$P\left[X > t \,|\, X > s\right] = P\left[X > t - s\right].$$

ii) Find the probability density function (pdf) of random variable Y defined by

$$Y = X + t_0,$$

for some $t_0 > 0$.

iii) Find the distribution of random variable W defined by

$$W = X^{1/\alpha},$$

for some $\alpha > 0$.

iv) Consider the random variable V that has a *Pareto* distribution on $\mathbb{V} \equiv \mathbb{R}^+$ with parameters θ and α. Show that the lack of memory property does not hold for V.

b) The *hazard function*, h_T, for a positive, continuous random variable T is defined by

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)},$$

for $t > 0$.

i) Show that the hazard function for W (from *a) iii)*), h_W, is constant if and only if $\alpha = 1$, is monotone increasing if $\alpha < 1$ and is monotone decreasing if $\alpha > 1$.

ii) Comment on the nature of the hazard function for V (from *a) iv)*), h_V.

iii) Consider the joint distribution of continuous random variables U_1 and U_2 defined by joint pdf

$$f_{U_1,U_2}(u_1, u_2) = \alpha u_2^{\alpha-1},$$

if $0 < u_2 < (-\log u_1)^{1/\alpha} < \infty, \quad u_1 > 0$ and is zero otherwise, for some $\alpha > 0$.

Find the marginal distribution of U_2.

3. *a)* The joint distribution of the continuous random variables X and Y is defined by the joint probability density function

$$f_{X,Y}(x,y) = k_1 \sin(x+y) \qquad (x,y) \in \mathbb{X}^{(2)},$$

and zero otherwise, where $\mathbb{X}^{(2)} \equiv \{(x,y) : 0 \leq x \leq \pi/2, 0 \leq y \leq \pi/2\}$.

i) Show that $k_1 = 1/2$.

ii) Find the probability $P\left[2X < Y\right]$.

iii) Are X and Y independent ? Justify your answer.

b) *i)* Find the moment generating function of the marginal distribution of X, $M_X(t)$.

ii) Find the expectation of X.

c) Suppose now that X and Y have a joint cumulative distribution function given by

$$F_{X,Y}(x,y) = k_2 \sin x \sin y,$$

where X and Y have the same range $\mathbb{X}^{(2)}$ as before.

Find the covariance of X and Y.

4. *a)* Suppose that X and Y are independent continuous random variables where $X \sim N(0,1)$ and Y is strictly positive (so that $\mathbb{Y} \equiv \mathbb{R}^+$).
Let $T = XY$.

i) Prove that, for $r = 1, 2, ...$

$$E_{f_T}[T^r] = E_{f_X}[X^r] \, E_{f_Y}[Y^r] \, .$$

ii) Show that the cumulative distribution function of T can be written

$$F_T(t) = \int_0^{\infty} \int_{-\infty}^{t/y} f_X(x) \, f_Y(y) \, \mathrm{d}x \, \mathrm{d}y.$$

iii) Show that

$$f_T(t) = \int_0^{\infty} f_X\left(\frac{t}{y}\right) \frac{1}{y} f_Y(y) \, \mathrm{d}y.$$

b) *i)* Suppose that $X \sim N(0,1)$, and that Y is a random variable with probability density function (pdf)

$$f_Y(y) = \frac{2\beta^{\alpha}}{\Gamma(\alpha)} \left(\frac{1}{y^2}\right)^{\alpha + \frac{1}{2}} \exp\left\{-\frac{\beta}{y^2}\right\} \qquad y > 0,$$

for constant $s, \alpha, \beta > 0$, where $\Gamma(\cdot)$ is the Gamma Function.

Find the pdf of $T = XY$.

ii) The *kurtosis* of a random variable Z, κ_Z, is defined by

$$\kappa_Z = \frac{E_{f_Z}\left[(Z - E_{f_Z}[Z])^4\right]}{\{Var_{f_Z}[Z]\}^2}.$$

Show, using properties of expectations and variances, that if $X \sim N(0,1)$ and $T = XY$ as in *a)* then

$$\kappa_X \leq \kappa_T.$$

5. *a)* *i)* Suppose that $X \sim Poisson(\lambda)$. By considering the limiting moment generating function of an appropriately standardized variable show that, for large λ, X is approximately normally distributed with parameters to be identified.

ii) Suppose that $X_1, .., X_n$, are independent and identically distributed $Poisson(\lambda)$ random variables.
Find either the method of moments estimator or the maximum likelihood estimator of λ derived from $X_1, .., X_n$.
Denote this estimator M_n.

iii) Show that M_n converges in probability to λ, that is,

$$M_n \xrightarrow{p} \lambda.$$

Hint:

$$P\left[|M_n - \lambda| < \varepsilon\right] = P\left[|nM_n - n\lambda| < n\varepsilon\right].$$

Recall that the Chebychev inequality states that for any random variable Z with *finite expectation μ and variance σ^2*

$$P\left[|Z - \mu| < \varepsilon\right] \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$

You may quote without proof any relevant properties of Poisson random variables.

b) *i)* Suppose that the independent random variables $X_1, .., X_n$ have a uniform distribution on $\mathbb{X} \equiv (0, \theta)$, and let

$$Y_n = \max\{X_1, .., X_n\}.$$

Show that

$$\sqrt{Y_n} \xrightarrow{p} \sqrt{\theta}.$$

Hint: you need only consider $P\left[\left|\sqrt{Y_n} - \sqrt{\theta}\right| < \varepsilon\right]$ *for* $0 < \varepsilon < \sqrt{\theta}$.

ii) Let the sequence of random variables $X_1, .., X_n$ on range $\mathbb{X} \equiv [0, 2\pi]$ have probability density functions given by

$$f_{X_n}(x) = \frac{1}{2\pi}\left[1 + \cos(nx)\right] \qquad 0 \leq x \leq 2\pi,$$

and zero otherwise.

Deduce whether or not the sequence $X_1, .., X_n$ converges in distribution, and if it does, find the limiting distribution.

DISCRETE DISTRIBUTIONS

	RANGE $\mathbb{X}$	PARAMETERS	MASS FUNCTION f_X	CDF F_X	$\mathrm{E}_{f_X}[X]$	$\mathrm{Var}_{f_X}[X]$	MGF M_X
$Bernoulli(\theta)$	$\{0,1\}$	$\theta \in (0,1)$	$\theta^x(1-\theta)^{1-x}$		θ	$\theta(1-\theta)$	$1-\theta+\theta e^t$
$Binomial(n,\theta)$	$\{0,1,...,n\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n}{x}\theta^x(1-\theta)^{n-x}$		$n\theta$	$n\theta(1-\theta)$	$(1-\theta+\theta e^t)^n$
$Poisson(\lambda)$	$\{0,1,2,...\}$	$\lambda \in \mathbb{R}^+$	$\frac{e^{-\lambda}\lambda^x}{x!}$		λ	λ	$\exp\{\lambda(e^t-1)\}$
$Geometric(\theta)$	$\{1,2,...\}$	$\theta \in (0,1)$	$(1-\theta)^{x-1}\theta$	$1-(1-\theta)^x$	$\frac{1}{\theta}$	$\frac{(1-\theta)}{\theta^2}$	$\frac{\theta e^t}{1-e^t(1-\theta)}$
$NegBinomial(n,\theta)$	$\{n,n+1,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{x-1}{n-1}\theta^n(1-\theta)^{x-n}$		$\frac{n}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta e^t}{1-e^t(1-\theta)}\right)^n$
or	$\{0,1,2,...\}$	$n \in \mathbb{Z}^+, \theta \in (0,1)$	$\binom{n+x-1}{x}\theta^n(1-\theta)^x$		$\frac{n(1-\theta)}{\theta}$	$\frac{n(1-\theta)}{\theta^2}$	$\left(\frac{\theta}{1-e^t(1-\theta)}\right)^n$

For **CONTINUOUS** distributions (given on Page 8), define the **GAMMA FUNCTION**

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

and the LOCATION/SCALE transformation $Y = \mu + \sigma X$ gives

$$f_Y(y) = f_X\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} \qquad F_Y(y) = F_X\left(\frac{y-\mu}{\sigma}\right) \qquad M_Y(t) = e^{\mu t}M_X(\sigma t) \qquad \mathrm{E}_{f_Y}[Y] = \mu+\sigma\mathrm{E}_{f_X}[X] \qquad \mathrm{Var}_{f_Y}[Y] = \sigma^2\mathrm{Var}_{f_X}[X]$$

					ν
					(if $\alpha > 2$)
$Beta(\alpha, \beta)$	$(0, 1)$	$\alpha, \beta \in \mathbb{R}^+$	$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$	$\frac{\alpha}{\alpha+\beta}$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$