UNIVERSITY COLLEGE LONDON

EXAMINATION FOR INTERNAL STUDENTS

MODULE CODE : **MATH7501**

MODULE NAME : **Probability and Statistics**

DATE : **30-Apr-07**

TIME : **10:00**

TIME ALLOWED : **2 Hours 0 Minutes**

2006/07-MATH7501A-001-EXAM-110

TURN OVER

All questions may be attempted but only marks obtained on the best four *solutions will count.*
The use of an electronic calculator is *permitted in this examination.*

New Cambridge Statistical Tables are provided.

1. (a) Let X be a discrete random variable taking values in the (finite or infinite) set $\{x_i : i = 1, 2, \ldots\}$. Let $p(\cdot)$ denote the probability mass function of X, so that $P(X = x_i) = p(x_i)$ for all i.
   (i) Write down the definition of the expected value of X, $\mathrm{E}(X)$. State clearly any conditions required for this expected value to exist.
   (ii) Let $\phi : \mathbb{R} \to \mathbb{R}$ be a function. Write down the definition of the expected value of $\phi(X)$.
   (iii) Prove that if a and b are constants, $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$.

   (b) How does the definition in part (a)(i) change if X is continuous rather than discrete?

   (c) Let X be any random variable whose expected value exists and is equal to μ. The variance of X is defined as $\mathrm{Var}(X) = \mathrm{E}(X - \mu)^2$.
   (i) Show that $\mathrm{Var}(X)$ can equivalently be written as $E(X^2) - \mu^2$.
   (ii) Show that if a and b are constants, $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.

2. (a) State and prove the Law of Total Probability, taking care to define any notation that you use.

   (b) Following a civil war in a region of political instability, a United Nations peace-keeping force is sent to clear the area of land mines. A proportion p of mines in the area are inert; the rest will explode if stepped on or driven over. One way to find out whether a mine is inert is to fire a bullet at it from a safe distance: an inert mine will never explode, but a 'live' one will explode with probability α, say, independently each time it is hit.

   Suppose a UN troop finds a mine, and fires bullets at it repeatedly from a safe distance.
   (i) What is the probability that after being hit n times, the mine has not exploded?
   (ii) The troop has hit the mine n times, and it has not exploded. Given this information, what is the probability that it will remain unexploded after the next hit? What is the limiting value of this probability as $n \to \infty$?

3. (a) Show that if $\alpha > 0$ and $\beta > 0$ are both constants, the function

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1}e^{-\alpha x^{\beta}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

is a valid probability density function. Find the corresponding distribution function.

(b) A distribution with the density given in part (a) is called a *Weibull distribution* with parameters α and β. Show that if X has a Weibull distribution with parameters α and β, and $Y = X^{\gamma}$ for some $\gamma > 0$, then Y also has a Weibull distribution. Give the parameters of this distribution.

4. The concentration of dissolved organic carbon (DOC) in water samples taken from a particular lake is approximately normally distributed, with mean 6 parts per million (ppm) and standard deviation 1.5ppm.

(a) What is the probability that a given water sample will have a DOC concentration in excess of 8ppm?

(b) What concentration of DOC is exceeded with probability 0.01 in any given sample?

(c) Among five samples from the lake (assumed to have been taken independently), what is the probability that at least two have DOC concentrations in excess of 8ppm?

(d) Suppose now that 20 samples are taken from the lake, and that exactly four of these are known to have DOC concentrations in excess of 8ppm. Five of the 20 samples are randomly selected for further testing. What is the probability that at least two of these five samples have DOC concentrations in excess of 8ppm?

5. 15 students on a statistical computing course were each given two tests. The tests were designed by different instructors. The percentage marks obtained by the students on each test were as follows:

Student	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Test 1	51	53	64	56	53	83	77	72	77	78	70	72	56	80	90
Test 2	72	54	72	40	46	90	68	58	90	40	58	74	40	78	86

(a) Compute the differences between the two test scores for each student, and prepare a stem-and-leaf plot to display the distribution of these differences.

(b) Test, at the 95% level and using a 2-tailed test, the hypothesis that the two sets of marks are drawn from distributions with the same mean. State your conclusions clearly. Also, state clearly any assumptions you make, and if possible comment on the validity of these assumptions.

6. (a) Suppose $X_1, \ldots, X_n$ are independent and identically distributed random variables, each coming from a probability distribution with some parameter θ. Let $T_n = T_n(X_1, \ldots, X_n)$ be a function of the $\{X_i\}$.

   (i) What is meant by saying that T_n is an unbiased estimator of θ?

   (ii) What is meant by saying that T_n is a consistent estimator of θ?

   (b) Let $X_1, \ldots, X_n$ be independent and identically distributed, each coming from some distribution (not necessarily normal) with mean μ and variance σ^2.

   (i) Show that the sample mean

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

   is an unbiased estimator of μ, and show that the variance of $\overline{X}$ is σ^2/n.

   (ii) Show that the sample variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$$

   is an unbiased estimator of σ^2.