UNIVERSITY COLLEGE LONDON

University of London

EXAMINATION FOR INTERNAL STUDENTS

For The Following Qualifications:-

B.Sc. *M.Sci.*

Mathematics M252: Probability and Statistics

COURSE CODE	:	**MATHM252**
UNIT VALUE	:	**0.50**
DATE	:	**19-MAY-03**
TIME	:	**14.30**
TIME ALLOWED	:	**2 Hours**

03-C0957-3-100

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. Let A and B be two events with $P(A) = \frac{1}{4}$ and $P(B|A) = \frac{1}{3}$.

   (a) Write down $P(A \cap B)$, and hence deduce that $P(A \cup B) = \frac{1}{6} + P(B)$.

   (b) Explain why, in this case, $\frac{1}{4} \leqslant P(A \cup B) \leqslant 1$. State the range of values within which $P(B)$ must lie.

   (c) If $P(A \cup B) = \frac{1}{2}$, state whether A and B are (i) disjoint (ii) independent.

2. A box contains 13 black discs and 7 red ones. 6 discs are drawn from the box at random, without replacement.

   (a) Find the probability that

      (i) 3 black and 3 red discs are drawn.

      (ii) The 6 discs are all the same colour.

      (iii) The 5th disc drawn is the third red disc drawn.

   (b) Suppose that all the discs are now replaced in the box and the experiment (i.e. the drawing of 6 discs) is repeatedly carried out until 3 red and 3 black discs are obtained. Find the probability that

      (i) Exactly 3 repetitions will be required.

      (ii) At most 5 repetitions will be required.

      (iii) Between 10 and 20 (inclusive) repetitions will be required.

   Calculate the median number of repetitions required.

3. (a) There are 31 days in the month of May. In a certain part of the world, each day of the month is dry with probability 0.6, independently of all other days. Any day which is not dry is classified as 'wet', and experiences some non-zero rainfall amount.

   Find the probability that

   (i) There are 10 wet days in the month.

   (ii) There are 10 wet days in the month, but none of these occur in the first week.

   (b) On wet days, the amount of rain which falls (in mm) is a continuous random variable with density

$$f(x) = \begin{cases} \frac{x}{25}e^{-x/5} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

   EITHER by naming this distribution, stating the values of its parameters and quoting the appropriate results, OR directly from the definition of expectation, find the mean rainfall amount on a wet day. Deduce that the mean rainfall during May is 124mm. Explain your reasoning carefully here, if necessary stating any assumptions which you are making.

   (c) Find the probability that the rainfall on a wet day will exceed 20mm.

4. (a) Use Bayes' Theorem to express $P(B|A)$ in terms of $P(A|B)$, $P(A)$ and $P(B)$.

   (b) Suppose that a particular species of plant occurs in two forms, type I and type II, and that in a certain region 25% of these plants are type I. For type I plants the lengths (in mm) of the stems have a normal distribution with mean 50 and variance 9, and for type II the lengths are normally distributed with mean 60 and variance 25. A plant is classified as type I if its stem length is less than 55m, and as type II otherwise.

   (i) Find the probability that a type I plant is classified as type I, and the probability that a type II plant is classified as type II.

   (ii) What is the probability that a randomly-selected plant is misclassified?

   (iii) What is the probability that a plant is actually of type I if it has been classified as type I?

   (c) Now suppose that plants are classified as type I if the stem length is less than αmm, and type II otherwise. What value of α ensures that the misclassification probabilities are the same for both plant types?

5. The following data are recorded weight increases (in grams) of babies over a period of one week. The first group of 10 babies is used as control, and the second group of 14 babies has received an anti-allergic treatment.

Control:	-60	-45	-30	0	0	0	10	10	10	35	75			
Treated:	-75	-75	-50	-50	-50	-45	-30	-30	-25	-10	0	10	25	100

(a) Assume that the two groups of measurements can be regarded as independent samples from populations with means μ_1 and μ_2, and variances σ_1^2 and σ_2^2, respectively. Calculate estimates of μ_1, μ_2, σ_1^2 and σ_2^2.

(b) Test, at the 95% level, the hypothesis that the $\sigma_1^2 = \sigma_2^2$. State your conclusions clearly.

(c) Assuming that $\sigma_1^2 = \sigma_2^2$, calculate a 95% confidence interval for $\mu_1 - \mu_2$.

(d) What do these analyses tell you about the effectiveness of the new treatment? State any additional assumptions you have made during the analyses, and comment on their validity.

6. (a) Let X be a Poisson random variable with mean m. Show that $E\left[e^{aX}\right] = \exp\left[m\left(e^a - 1\right)\right]$, for any constant a.

(b) Sheep are prone to parasite infection. Suppose the number of parasites living on each animal follows a Poisson distribution with mean μ. If the animal has one or more parasites, it is said to be 'infected'. A farmer wants to know the proportion, p, of her flock that is infected. She takes n animals from the flock and counts the number of parasites on each of them. She chooses the animals in such a way that the counts can be regarded as independent.

(i) Let $X_1, \ldots, X_n$ denote the parasite counts, and let $S_n = \sum_{i=1}^{n} X_i$. The farmer notices that $p = 1 - e^{-\mu}$ and that $n^{-1}S_n$ is an unbiased estimator of μ. She therefore proposes to use $T = 1 - \exp\left[-n^{-1}S_n\right]$ as an estimator of p.
By stating the distribution of S_n and applying the result from part (a), or otherwise, show that T is a biased estimator of p. Also, show that as $n \to \infty$, $E(T) \to p$.

(ii) Let Y be the total number of infected animals in the sample. State the distribution of Y, and show that Y/n is an unbiased estimator of p. State the standard error of this estimator (in terms of p).

(iii) If you were asked to choose between the estimators T and Y/n, what considerations would influence your decision? If necessary, state clearly any additional calculations that would be required (but do not attempt to carry out any such calculations).