UNIVERSITY OF SURREY©

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

Level HE1 Examination

Module MS125 Proof, Probability and Experiment: Probability part

Time allowed – 2 hours

Spring Semester 2008

Attempt FOUR questions. If any candidate attempts more than FOUR questions only the best FOUR solutions will be taken into account.
Cambridge Statistical Tables and a formula sheet will be provided.
Calculators may be used.

Question 1

(a) An insurance company has 10,000 policy holders. Each policy holder is classified as: young or old; male or female; and married or single.
Of these policy holders, 4500 are young, 6000 are male, and 4000 are married. The policy holders can also be classified as 2500 young males, 1500 married males, and 2000 young married persons. Finally, 1000 of the policyholders are young married males.

(i) How many of the policy holders are females? [1]

(ii) How many of the policy holders are young or married? [4]

(iii) How many of the policy holders are old, female, and single? [6]

(b) (i) A committee of size 5 is to be selected from a group of 6 mathematics students and 9 engineering students. If the selection is made randomly, what is the probability that the committee consists of 3 mathematics students and 2 engineering students?
[Give your answer correct to four decimal places] [4]

(ii) Now suppose that a representative committee of size 5 is to be selected from 600 mathematics students and 900 engineering students.

i. Write down an expression for the probability that the committee consists of 3 mathematics students and 2 engineering students and evaluate this probability.
[Give your answer correct to four decimal places] [4]

ii. Write down a binomial random variable that can be used to approximate the probability that exactly r of the 5 committee members are mathematics students. Use this random variable to estimate the probability that exactly 3 the committee members are mathematics students. Based on this binomial random variable, what is the expected number of mathematics students in the committee?
[Give your answer correct to four decimal places] [6]

Question 2

(a) A doctor is studying the relationship between blood pressure and heartbeat abnormalities in her patients. She tests a random sample of her patients and notes their blood pressures (high, low, or normal) and their heartbeats (regular or irregular). She finds that: 25% have high blood pressure; 15% have low blood pressure; and 60% have normal blood pressure. Of those with high blood pressure, one-fourth have an irregular heartbeat. Of those with low blood pressure, one-fifth have an irregular heartbeat. Of those with normal blood pressure, one-eighth have an irregular heartbeat.

(i) What is the probability, that a patient chosen at random has irregular heartbeat? [4]

(ii) If we find a patient with regular heartbeat, what is the probability that the patient has high blood pressure? [5]

(b) A random variable X has distribution given by probability density function

$$p(x) = c\left(\frac{1}{2}\right)^x, \quad x = 1, 2, ..,$$

(i) Determine the constant c. [3]

(ii) Show that the probability generating function $G_x(s)$ for this distribution is

$$G_x(s) = \frac{s}{2-s}$$

. [4]

(iii) Using $G_x(s)$ from part (ii), evaluate the expectation $E[X]$. [4]

(iv) Using $G_x(s)$ from part (ii), evaluate the variance $Var[X]$. [5]

Question 3

(a) Let X be a random variable with probability density function $f(x) = cx^2$, $0 < x < 1$.

(i) Find the value of the constant c. [2]

(ii) Find the mean $E[X]$ and variance $Var[X]$. [5]

(iii) Find the probability $P(X > 0.25|X < 0.75)$. [4]

(iv) If the random variable Y is defined by: $Y = -3X + 2$, calculate $E[Y]$ and $Var[Y]$. [2]

(b) (i) $x_1, \ldots, x_n$ is a random sample of size n from a random variable with probability density function:

$$f(x) = \theta^2 x \exp(-\theta x), x > 0$$

$$\theta > 1$$

Find $\hat{\theta}$, the maximum likelihood estimator of θ. [8]

(ii)) If the random variable X has the probability density function given in part (i), show that

$$P(a < X < b) = e^{-a\theta}(1 + a\theta) - e^{-b\theta}(1 + b\theta)$$

[4]

Question 4
Suppose that X_1, X_2 and X_3 are independent random variables which each take values 1 and 2 with probability $\frac{1}{2}$.

Let $Y_1 = X_1X_2X_3$ and $Y_2 = X_1X_2 - X_3$.

(a) Calculate the joint probability density function $f_{X,Y}(y_1, y_2)$ for the pair (Y_1, Y_2) and display the joint probabilities in an appropriate table. [6]

(b) Find the marginal probability density functions for Y_1 and Y_2. [4]

(c) Calculate the means and variances of Y_1 and Y_2. [3]

(d) Calculate the covariance of Y_1 and Y_2. [4]

(e) Find the conditional distribution of Y_1 given $Y_2 = 0$. [4]

(f) Find the conditional mean of Y_1 given $Y_2 = 0$. [2]

(g) State, with justification, whether or not Y_1 and Y_2 are independent. [2]

Question 5

(a) On the basis of the exam scores in the following table, test whether there are gender-associated differences in mathematical ability: [7]

	Boys	Girls	Total
Score: 70-85	50	60	110
Score: 85-100	50	40	90
Total	100	100	200

(b) (i) The rainfall at a certain site is a random variable X, which may be assumed to be distributed as $N(\mu_X, \sigma^2)$ with μ_X unknown and $\sigma = 3$ inches. For the past 10 years, the following rainfall figures have been collected:

30.5, 34.1, 27.9, 29.4, 35.0, 26.9, 30.2, 28.3, 31.7, 25.8.

Test the hypothesis $H_0 : \mu_X = 30$ against the alternative $H_1 : \mu_X < 30$. [7]

(ii) Now, consider that we also have data from another site, where the rainfall is a random variable Y, which may be assumed to be distributed as $N(\mu_Y, \sigma^2)$ with μ_Y unknown and $\sigma = 3$ inches. For a different 10 year period, the following rainfall figures have been collected:

24.4, 20.9, 25.6, 25.8, 23.2, 31.3, 26.1, 31.8, 28.3, 24.3.

i. Test the hypothesis $H_0 : \mu_X = \mu_Y$ against the alternative $H_1 : \mu_X \neq \mu_Y$. [8]

ii. Would the decision about accepting or rejecting the null hypothesis be different if the alternative hypothesis is $H_1 : \mu_X > \mu_Y$? [3]

Question 6

(a) Clinical trials involving 12 patients were carried out to compare the effect of placebo and pronethalol on angina. The number of attacks suffered by each patient under each treatment is given below:

Patient ID	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
Placebo	71	323	8	14	23	34	79	60	2	3	17	7
Pronethalol	29	348	1	7	16	25	65	41	0	0	15	2

Carry out a statistical test to determine if there is any evidence to indicate that the treatment with Pronethalol is effective. [12]

(b) A worker travels to work by bus every day. The following table gives the worker's waiting time at the bus stop to the nearest minute on 320 working days.

Minutes to wait for next bus	1	2	3	4	5	6	7	8	9	10
Frequency	25	30	30	35	40	40	35	35	30	20

Test the hypothesis that the data fit a discrete uniform distribution $U(1, 10)$. [13]

INTERNAL EXAMINER: R. Retkute