OCR

RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 2644

Probability & Statistics 4

Thursday **30 MAY 2002** Morning 1 hour 20 minutes

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF8)

TIME 1 hour 20 minutes

INSTRUCTIONS TO CANDIDATES

- Write your Name, Centre Number and Candidate Number in the spaces provided on the answer booklet.
- Answer **all** the questions.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.
- You are permitted to use a graphic calculator in this paper.

INFORMATION FOR CANDIDATES

- The number of marks is given in brackets [ ] at the end of each question or part question.
- The total number of marks for this paper is 60.
- Questions carrying smaller numbers of marks are printed earlier in the paper, and questions carrying larger numbers of marks later in the paper.
- **You are reminded of the need for clear presentation in your answers.**

This question paper consists of 4 printed pages.

1 A continuous random variable X has moment generating function given by

$$\mathrm{M}_X(t) = \frac{9}{(3-t)^2}.$$

Find the mean and variance of X. [4]

2 A random sample of 12 drivers was taken. Each driver was asked to test-drive two different cars, A and B, and to give each car a mark out of 10. The marks given by the drivers are shown in the table.

Driver	1	2	3	4	5	6	7	8	9	10	11	12
Marks for car A	9	5	8	8	7	7	6	7	5	9	8	7
Marks for car B	5	9	6	5	6	5	5	8	7	4	5	6

Use a sign test, at the 10% significance level, to test whether car A would be preferred by the majority of drivers. [6]

3 The continuous random variable X has probability density function given by

$$\mathrm{f}(x) = \begin{cases} kx & 0 \leqslant x \leqslant a, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a constant and the value of the parameter a is unknown.

(i) Show that $k = \dfrac{2}{a^2}$. [2]

The random variable U is defined by $U = \frac{3}{2}X$.

(ii) Show that U is an unbiased estimator of a. [3]

(iii) Find, in terms of a, the variance of U. [3]

4 A gambler has two six-sided dice. One of these dice is an unbiased die for which the probability of a six is $\frac{1}{6}$. The other die is a biased die for which the probability of a six is $\frac{1}{3}$. The gambler chooses one of these dice at random and rolls it.

(i) Find the probability that this roll results in a six. [2]

(ii) Given that this roll of the die has resulted in a six, find the probability that it was the unbiased die that was rolled. [3]

(iii) Given that this roll of the die has resulted in a six, find the probability that when the gambler rolls the same die again a second six is obtained. [4]

5 A University's Department of Computing is interested in whether students who have passed A level Mathematics perform better in Computing examinations that those who have not.

A random sample of 19 students was taken from those students who took a particular first year Computing examination. This sample included 12 students who have passed A level Mathematics and 7 students who have not. The marks gained in the Computing examination were as follows:

Students who have passed A level Mathematics: 27, 34, 39, 41, 45, 47, 55, 59, 66, 75, 78, 86.

Students who have not passed A level Mathematics: 17, 21, 28, 35, 37, 54, 64.

Use a suitable non-parametric test to determine if there is evidence, at the 5% significance level, that students who have passed A level Mathematics gain a higher average mark than students who have not passed A level Mathematics. (A normal approximation may be used.) [10]

6 The joint probability distribution of the discrete random variables X and Y is given by

$$\mathrm{P}(X = x \text{ and } Y = y) = \frac{x + 2y}{18} \quad \text{for } (x, y) = (-1, 2), (-1, 3), (0, 2), (0, 3).$$

(i) Copy and complete the joint probability distribution table given below. [1]

		x	
		-1	0
y	2		
	3		

(ii) Show that $\mathrm{E}(X) = -\frac{4}{9}$ and find $\mathrm{Var}(X)$. [3]

It is given that $\mathrm{E}(Y) = \frac{47}{18}$ and $\mathrm{Var}(Y) = \frac{77}{324}$.

(iii) Find $\mathrm{Cov}(X, Y)$ and state, with a reason, whether X and Y are independent. [4]

(iv) Find $\mathrm{Var}(X + Y)$. [2]

[Question 7 is printed overleaf.]

7 The random variable X has a geometric distribution with parameter p.

(i) Show that the probability generating function $\mathrm{G}_X(t)$ of X is given by

$$\mathrm{G}_X(t) = \frac{pt}{1 - t(1-p)}.$$ [3]

(ii) Hence show that $\mathrm{E}(X) = \frac{1}{p}$ and that $\mathrm{Var}(X) = \frac{1-p}{p^2}$. [5]

A child has 4 fair, six-sided dice, one white, one yellow, one blue and one red.

(iii) The child rolls the white die repeatedly until the die shows a six. The number of rolls up to and including the roll on which the white die first shows a six is denoted by W. Write down an expression for $\mathrm{G}_W(t)$. [1]

(iv) The child then repeats this process with the yellow die, then with the blue die and then with the red die. By finding an appropriate probability generating function, find the probability that the total number of rolls of the four dice, up to and including the roll on which the red die first shows a six, is exactly 24. [4]

S4 jun 02

1. Since $M_X(t) = 1 + \mu t + \frac{\mu^2+\sigma^2}{2!}t^2 + \cdots$
let's expand the MGF:
$= \left(1 - \frac{t}{3}\right)^{-2}$
$= 1 + \frac{2}{3}t + \frac{1}{3}t^2 + \cdots$
so
$\mu = \frac{2}{3}$ and $\sigma^2 = \frac{2}{9}$

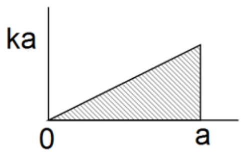
---

2. This will be a paired sample sign test.
H_0: there is no preference of car
H_1: car A is preferred by the majority of drivers.
Signs are: $+ - + + + + + - - + + +$

Under H_0, the number, X, of $+$ signs is distributed $B(12, \frac{1}{2})$
and $P(X \geq 9) = 1 - 0.9270 = 0.073$
so 9 is within the 10% tail and we conclude that there is significant evidence to suggest that drivers generally prefer car A.

---

3. (i)



$\frac{1}{2} \times a \times ka = 1$, so
$k = \frac{2}{a^2}$

(ii)
$E\left(\frac{3}{2}X\right) = \frac{3}{2}E(X)$
$= \frac{3}{2}\int_0^a x\frac{2}{a^2}xdx$
$= \frac{3}{a^2}\left[\frac{x^3}{3}\right]_0^a$
$= a$
So $\frac{3}{2}X$ is an unbiased estimator for a

(iii)
$V(U) = V\left(\frac{3}{2}X\right) = \frac{9}{4}V(X)$

$$= \frac{9}{4}\left(\int_0^a x^2 . \frac{2}{a^2} x dx - \frac{4}{9}a^2\right)$$

$$= \frac{1}{8}a^2$$

---

4. (i) $P(6) = \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{1}{3} = \frac{1}{4}$

(ii) ($\frac{2}{3}$: trivial or:)

$$P(biased\ |6) = \frac{P(biased\ and\ 6)}{P(6)}$$

$$\frac{\frac{1}{2} \times \frac{1}{3}}{\frac{1}{4}}$$

$$= \frac{2}{3}$$

(iii)

$$\frac{P(6,6)}{P(1st\ is\ 6)}$$

$$= \frac{\frac{1}{2} \times \frac{1}{6} \times \frac{1}{6} + \frac{1}{2} \times \frac{1}{3} \times \frac{1}{3}}{\frac{1}{4}}$$

$$= \frac{5}{18}$$

5.

H_0: the two samples come from populations with identical distributions
H_1: the population of students who passed A level Maths have a higher median mark

Putting all the marks in rank order:

17	21	27	28	34	35	37	39	41	45	47	54	55	59	64	66	75	78	86
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Y	Y	X	Y	X	Y	Y	X	X	X	X	Y	X	X	Y	X	X	X	X

Sum of Y ranks= 47, and $m(n + m + 1) - 47 = 7 \times 20 - 47 = 93$.
We take the smaller, i.e. 47.
Qu. suggests using Normal approx. (although m, n should be greater than 10)
Mean $= \frac{1}{2} \times 7 \times 20 = 70$
Variance= $\frac{1}{12} \times 7 \times 12 \times 20 = 140$
So $P(\sum r_Y \leq 47) \approx \Phi\left(\frac{47.5-70}{\sqrt{140}}\right)$
$= \Phi(-1.9015..)$
0.0286.
There is significant evidence that students with A level Maths score higher marks than others.

6. (i)

		x	
		-1	**0**
y	**2**	$\frac{1}{6}$	$\frac{2}{9}$
	3	$\frac{5}{18}$	$\frac{1}{3}$

(ii)
$$E(X) = -1 \times \frac{4}{9} + 0 \times \frac{5}{9} = -\frac{4}{9}$$
$$V(X) = (-1)^2 \times \frac{4}{9} + 0 - \left(-\frac{4}{9}\right)^2 = \frac{20}{81}$$

(iii)
$$Cov(X,Y) = E(XY) - E(X)E(Y)$$
$$= -2 \times \frac{1}{6} - 3 \times \frac{5}{18} - \left(-\frac{4}{9}\right)\left(\frac{47}{18}\right)$$
$$= -\frac{1}{162}$$
X and Y cannot be independent otherwise $Cov(X,Y)$ would be 0

(iv)
$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$
$$= \frac{20}{81} + \frac{77}{324} - \frac{1}{81} = \frac{153}{324}$$

---

7. (i)
$$G_X(t) = \sum_1^\infty q^{i-1} p t^i$$
$$= \frac{p}{q} \sum_1^\infty (qt)^i$$
$$= \frac{p}{q} \times \frac{qt}{1-qt}$$
$$= \frac{pt}{1-t(1-p)}$$

(ii)
$$G_X'(t) = \frac{p}{(1-qt)^2}$$
and
$$G_X''(t) = \frac{2pq}{(1-qt)^3}$$
$$E(X) = G_X'(1) = \frac{p}{p^2} = \frac{1}{p}$$
and
$$V(X) = G_X''(1) + G_X'(1) - G_X'(1)^2$$
$$= \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2}$$

$$= \frac{1-p}{p^2}$$

(iii)

$$G_W(t) = \frac{\frac{1}{6}t}{1 - \frac{5}{6}t}$$

(iv) Defining variables in the obvious way, the number of rolls up to and including when the red die shows a 6 is $W + Y + B + R$

$$G_{W+Y+B+R}(t) = G_W(t) \times G_Y(t) \times G_B(t) \times G_R(t)$$

$$= \frac{t^4}{6^4}\left(1 - \frac{5}{6}t\right)^{-4}$$

Required probability is the coefficient of the t^{24} term

The relevant term is

$$\frac{t^4}{6^4} \times \frac{(-4)(-5)(-6)\dots(-23)}{20!}\left(-\frac{5}{6}t\right)^{20}$$

and probability is

$$= \frac{21 \times 22 \times 23 \times 5^{20}}{6^{25}}$$

(or 0.0356)