OCR

RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 2644

Probability & Statistics 4

Friday **18 JANUARY 2002** Afternoon 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.

[Turn over

1 The events A and B are such that

$$P(A) = 0.40, \quad P(B) = 0.25 \quad \text{and} \quad P(A \cap B) = 0.10.$$

(i) Find $P(A \cup B)$. [1]

It is also given that the event C is such that $P(C) = 0.20$ and $P(B \cap C) = 0.12$. The events A and C are mutually exclusive.

(ii) Find $P(A \cup B \cup C)$. [3]

2 The 567 Members of Parliament (MPs) elected for constituencies in England and in Wales in a General Election were either Labour, Conservative, Liberal Democrat or Welsh Nationalist. The table below shows the number in each category.

	Labour	Conservative	Liberal Democrat	Welsh Nationalist
England	328	165	34	0
Wales	34	0	2	4

An MP is chosen at random from this group.

E is the event that an MP for an English constituency is chosen.

D is the event that a Liberal Democrat MP is chosen.

(i) Find $P(E)$. [1]

(ii) Find $P(E \mid D)$. [1]

(iii) State, with a reason, whether E and D are independent events. [1]

F is the event that a woman MP is chosen.

(iv) Given that 108 of the 567 MPs are women, and that $P(E \cap F) = 0.183$, correct to 3 decimal places, find $P(E \mid F)$. State what $P(E \mid F)$ represents. [4]

3 X is a discrete random variable taking values 0, 1, 2, 3, ... The probability generating function of X is given by

$$G_X(t) = \frac{at}{b - t}$$

where a and b are constants. Given that $E(X) = 3$,

(i) find the values of a and b, [4]

(ii) find $P(X = 5)$. [3]

4 A population has mean μ and variance σ^2. Two independent observations, X_1 and X_2, are taken from the population. In order to estimate the value of μ, the following three estimators are being considered:

$$U_1 = \frac{X_1 + X_2}{2}, \qquad U_2 = \frac{X_1 + 2X_2}{2}, \qquad U_3 = \frac{4X_1 - X_2}{3}.$$

(i) Find the means of these three estimators and hence state which two of them are unbiased. [4]

(ii) Find the variance of each of the two unbiased estimators. [3]

(iii) Hence state, giving a reason, which of the two unbiased estimators is more efficient. [1]

5 The random variable X has probability density function given by

$$f(x) = \begin{cases} e^{-x} & x \geqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Show that the moment generating function of X is given by

$$M_X(t) = \frac{1}{1-t}.$$ [3]

The random variable Y has the same distribution as X and is independent of X.

(ii) Write down the moment generating function of $X + Y$. [1]

(iii) The random variable W has probability density function given by

$$g(w) = \begin{cases} we^{-w} & w \geqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

Show that W has the same moment generating function as $X + Y$. [5]

[Questions 6 and 7 are printed overleaf.]

6 For each year group that enters a Sixth Form College, the College collects data from each student on whether their parents have had a University education and, at a later date, whether the student goes on to University or not.

A statistician produced the bivariate probability distribution given below as a model for this data. For a randomly chosen student, X is the number of parents who went to University, and takes values 0, 1 or 2. The random variable Y takes the value 1 if the student went on to University; otherwise it takes the value 0.

		Values of Y	
		0	1
Values of X	0	0.13	0.37
	1	0.05	0.25
	2	0.02	0.18

(i) Find the marginal distributions of X and Y. [2]

(ii) Show that $\mathrm{E}(X) = 0.70$. Hence estimate, for a year group of 100 students, the number of parents of these students who had been to University. (You may assume that no two students had any parents in common.) [2]

(iii) Find the conditional distribution of X given $Y = 1$. Hence estimate what proportion of the students who went on to University had at least one parent who had been to University. [3]

(iv) Calculate the covariance of X and Y. State any conclusion you can draw about the random variables X and Y. [5]

7 A Tourist Information Authority wished to find out whether the holiday industry in their local area had had a good trading month in August 2000. At the end of the month they sent a questionnaire to a random sample of 10 of the local hotels asking for the figures of room occupancy in August 2000 and also in August 1999. The figures are given in the table below.

Hotel	A	B	C	D	E	F	G	H	I	J
Room occupancy in 1999	53	62	82	115	125	118	159	234	442	643
Room occupancy in 2000	44	68	87	103	110	119	142	231	428	624

(i) Test whether there was a decrease in room occupancy from 1999 to 2000,

(a) using a sign test at the 5% significance level, [6]

(b) using a Wilcoxon signed-rank test at the 5% significance level. [5]

(ii) Describe what features of the data lead to the different results of the two tests. [2]

S4 jan 02

1. **(i)** $P(A\cup B)=0.4+0.25-0.1=0.55$

   (ii) $P(A\cup B\cup C)=0.4+0.25+0.2-0.1-0.12-0+0=0.63$

---

2. **(i)** $P(E)=\dfrac{527}{567}=0.929$

   (iii) $P(E|D)=\dfrac{34}{36}=0.944$

   (iii) They are not independent because otherwise $P(E|D)$ would be the same as $P(E)$

   (iv) $P(E|F)=\dfrac{P(E\cap F)}{P(F)}=\dfrac{0.183}{\frac{108}{567}}=0.96$

   It's the probability that a female MP chosen at random will be English.

---

3. **(i)** $G(1)=1$ (total probability), so

   $$\frac{a}{b-1}=1$$

   $$G'(t)=\frac{ab}{(b-t)^2}$$

   and $G'(1)=E(X)$, so

   $$\frac{ab}{(b-1)^2}=3$$

   Solving simultaneously,

   $$a=\frac{1}{2},b=\frac{3}{2}$$

   (ii) Expanding the p.g.f.,

   $$G_X(t)=\frac{1}{2}t\times\frac{2}{3}\times\left(1-\frac{2}{3}t\right)^{-1}$$

   $$=\frac{t}{3}\left(1+\frac{2}{3}t+\cdots+\frac{(-1)(-2)(-3)(-4)}{4!}\left(-\frac{2}{3}t\right)^4+\cdots\right)$$

   So

   $$p_5=\frac{16}{243}$$

---

4. **(i)** The expectations are $\mu,\frac{3}{2}\mu,\mu$

   so U_1 and U_3 are unbiased.

(ii) $V(U_1) = \frac{1}{4}V(X_1 + X_2) = \frac{1}{2}\sigma^2$

$V(U_3) = \frac{1}{9}V(4X_1 - X_2) = \frac{17}{9}\sigma^2$

(iii) U_1 is the more efficient because it has smaller variance.

---

5. **(i)**

$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx}e^{-x}dx = \int_0^\infty e^{-x(1-t)}dx$$
$$= -\frac{1}{1-t}\left[e^{-x(1-t)}\right]_0^\infty$$
$$= \frac{1}{1-t}$$

(ii)

$$M_{X+Y}(t) = \frac{1}{(1-t)^2}$$

(iii)

$$M_W(t) = E(e^{tW}) = \int_0^\infty e^{tw}we^{-w}\,dw$$
$$= \int_0^\infty w.e^{-w(1-t)}\,dw$$
$$= \left[-\frac{1}{1-t}we^{-w(1-t)}\right]_0^\infty - \int_0^\infty -\frac{1}{1-t}e^{-w(1-t)}\,dw$$

$$= 0 + \frac{1}{1-t}M_X(t) = \frac{1}{(1-t)^2} = M_{X+Y}(t)$$

---

6. **(i)**

x_i	0	1	2
$P(X = x_i)$	0.5	0.3	0.2

y_i	0	1
$P(Y = y)$	0.2	0.8

(ii) $E(X) = 0 \times 0.5 + 1 \times 0.3 + 2 \times 0.2 = 0.7$
$E(X_1 + X_2 + \cdots + X_{100}) = 100 \times E(X) = 70$
i.e. expected number of parents is 70.

(iii)

x_i	0	1	2
$P(X = x_i \mid Y = 1)$	$\frac{37}{80}$	$\frac{25}{80}$	$\frac{18}{80}$

Question means $P(X \geq 1|Y = 1)$
which is $\frac{43}{80} = 54\%$

(iv) $Cov(X, Y) = E(XY) - \mu_X\mu_Y$
$= 1 \times 0.25 + 2 \times 0.18 - 0.7 \times 0.8$
$= 0.05$
Firstly, they are not independent. Secondly, there is a mild positive correlation between number of parents who went to uni and whether the child also goes.

---

7. **(i)**

(a) H_0: room occupancy was the same in 2000 as in 1999
H_1: room occupancy decreased

signs are: $- \; + \; + \; - \; - \; + \; - \; - \; - \; -$
Under H_0 the number, X, of – signs is distributed $B(10, \frac{1}{2})$
$P(X \geq 7) = 1 - 0.8281 = 0.172$
and is not in the 5% tail. So we conclude that there isn't sufficient evidence at the 5% level to suppose that occupancy has fallen.

(b)
H_0: the paired data are drawn from the same population
H_1: they are not

d_i	-9	6	5	-12	-15	1	-17	-3	-14	-19
signed ranks	-5	$+4$	$+3$	-6	-8	$+1$	-9	-2	-7	-10

$P = 8$, and $Q = -47$, so $T = 8$
With $n = 10$, the largest value of T which leads to rejection of H_0 (from tables) is 10.
8 is deeper in the tail, so reject H_0. There is significant evidence of a reduction in the occupancy in 2000.

Although there are a few increases leading to an 'acceptable' number of $+$'s in the sign test, they are very small increases and the decreases are much bigger, leading to many high negative rankings and a consequent finding of a reduction when the amounts of decrease are taken into account.