OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 2641

Probability & Statistics 1

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.

1 The random variable X has a $B(18, 0.4)$ distribution. Using the tables of cumulative binomial probabilities, and giving your answers correct to 3 significant figures, find

(i) $P(X \leqslant 7)$, [1]

(ii) $P(X > 6)$, [2]

(iii) $P(8 < X \leqslant 14)$. [3]

2 A student carried out a statistics survey in a supermarket. At a checkout she recorded how many items each customer had in their shopping basket. The table below gives her results.

Number of items in the basket	4	5	6	7	8	9
Number of customers with that number of items	2	7	9	11	2	5

(i) Calculate the mean and standard deviation of the number of items the customers had in their shopping baskets. [5]

(ii) At the checkout each customer was given two free items. State the mean and the standard deviation of the total number of items the customers now had. [2]

3 A discrete random variable X has the probability distribution given in the following table. It is given that $E(X) = 0.95$.

x	-1	0	1	2	3
$P(X = x)$	a	b	0.1	0.3	0.2

(i) Find the probability that X is greater than $E(X)$. [1]

(ii) Find the values of a and b. [4]

(iii) Calculate $\text{Var}(X)$. [3]

4 A doctor conducted a survey about the ages of patients who visited his surgery. He chose a day at random and recorded the age, in completed years, of each patient visiting the surgery on that particular day. The results are given in the table below.

Age in completed years	0–9	10–19	20–29	30–39	40–59	60–99
Frequency	19	6	7	11	13	24

(i) The doctor decided to draw a histogram of the data. He calculated that the frequency density for the 10–19 class was 0.6. Calculate the other frequency densities, and draw a histogram to represent the data. [5]

(ii) Later the doctor decided that picking just one day at random was a bad idea, so he conducted a second survey. This time patients were chosen at random from a list of all the patients who had visited the surgery in the last six months. The ages, in completed years, of the selected patients are recorded below.

3	4	4	7	13	18	24	26	28	31
34	35	35	36	42	47	49	57	59	61
64	64	67	72	73	78	81	81	83	86

Draw a stem-and-leaf diagram to show the data, and state one advantage which the stem-and-leaf diagram has when compared with a histogram. [4]

5 Seven men and five women have been nominated to serve on a committee. The committee consists of four members who are to be chosen from the seven men and five women.

(i) In how many different ways can the committee be chosen? [2]

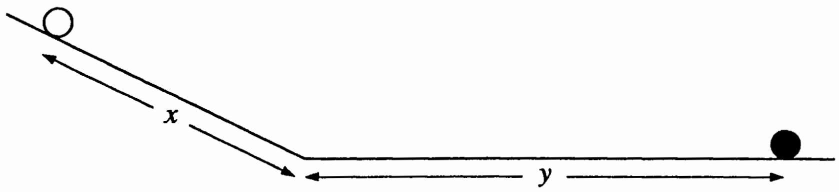
(ii) In how many of these ways will the committee consist of two men and two women? [4]

(iii) Assuming that each choice of four members is equally likely, find the probability that the committee will contain exactly two men. [2]

[Questions 6 and 7 are printed overleaf.]

[Turn over

6



A student conducted an experiment in which she rolled a ball from a point on an inclined plane. The ball rolled down the inclined plane and then continued to roll along a horizontal plane (see diagram). The student measured the distance, x cm, up the inclined plane to the ball's initial position. She also measured the distance, y cm, travelled by the ball along the horizontal plane before it stopped. She repeated the experiment to give 10 pairs of readings altogether. The values of x were increased at regular intervals of 5 cm starting at 40 cm. For each repetition of the experiment she tried to roll the ball down the plane with the same initial speed. The results are given in the table below.

x	40	45	50	55	60	65	70	75	80	85
y	58	69	70	73	88	91	98	108	109	125

$[n = 10,\ \Sigma x = 625,\ \Sigma y = 889,\ \Sigma x^2 = 41\,125,\ \Sigma y^2 = 83\,153,\ \Sigma xy = 58\,440.]$

(i) Calculate the value of the product moment correlation coefficient for the data. [2]

(ii) The student's teacher suggested that, instead of using the original x and y values, she should transform the data using the equations $u = \frac{x-40}{5}$ and $v = y - 50$, and then calculate the product moment correlation coefficient for the transformed data. Explain what relationship, if any, the product moment correlation coefficient for the transformed data would have with the product moment correlation coefficient calculated in part (i). [1]

The student wishes to estimate the value of x for a ball which rolls a horizontal distance of $y = 100$ cm.

(iii) Calculate the equation of an appropriate regression line and use it to estimate the value of x when $y = 100$. [5]

(iv) Give a reason for your choice of regression line. [1]

7 Adill and Beth are playing a game. Adill throws a fair die three times. The number of sixes that he obtains is denoted by A. Beth throws a fair coin repeatedly. The number of throws up to and including the first throw on which the coin lands head upwards is denoted by B.

(i) State the distribution of A, giving the values of any parameters. [2]

(ii) State the distribution of B, giving the values of any parameters. [2]

(iii) Find $P(A = 2)$. [2]

(iv) Find $P(B > 2)$. [2]

(v) Find $P(A = B)$. [5]

OCR Prob. & Stats 1 January 2002

1 $X \sim B(18,\ 0\cdot 4)$

$$\text{p}(X \le 7) = 0\cdot 5634... = \mathbf{0\cdot 563} \qquad (3\text{ s.f.})$$ [1]

$$\text{p}(X > 7) = 1 - \text{p}(X \le 6) = 1 - 0\cdot 37427... = \mathbf{0\cdot 626} \qquad (3\text{ s.f.})$$ [2]

$$\text{p}(8 < X \le 14) = \text{p}(X \le 14) - \text{p}(X \le 8) = 0\cdot 99978... - 0\cdot 73684... = \mathbf{0\cdot 263} \qquad (3\text{ s.f.})$$ [3]

2

$$\bar{x} = \frac{\sum x}{n} = \frac{235}{36} = 6\cdot 52\dot{7} = \mathbf{6\cdot 53} \qquad (3\text{ s.f.})$$

$$s = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2} = \sqrt{\frac{1063}{36} - 6\cdot 52\dot{7}^2} = 1\cdot 3841... = \mathbf{1\cdot 38} \qquad (3\text{ s.f.})$$ [5]

new mean = **8.53** new s.d. = **1.38** [2]

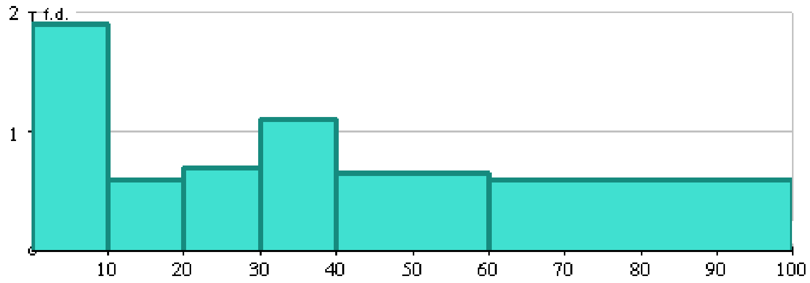
3 $\text{p}(X > \text{E}[X]) = \text{p}(X > 0.95) = 0.1 + 0.3 + 0.2 = \mathbf{0.6}$ [1]

$$\left.\begin{array}{lll} \text{E}[X] = 0.95 & \Rightarrow & -a + 1.3 = 0.95 \\ \sum x_i p_i = 1 & \Rightarrow & a + b + 0.6 = 1 \end{array}\right\} \qquad \boldsymbol{a = 0.35 \quad b = 0.05}$$ [4]

$$\text{Var}[X] = \sum x_i^2 p_i - \text{E}[X]^2 = 0.35 + 0 + 0.1 + 1.2 + 1.8 - 0.95^2 = \mathbf{2.5475}$$ [3]

4

age	0 – 9	10 – 19	20 – 29	30 – 39	40 – 59	60 - 69
freq. density	1.9	0.6	0.7	1.1	0.65	0.6



[5]

8	1 1 3 6
7	2 3 8
6	1 4 4 7
5	7 9
4	2 7 9
3	1 4 5 5 6
2	4 6 8
1	3 8
0	3 4 4 7

Key: | 1 | 8 represents 18 yrs

State one advantage of the stem plot over the histogram

(e.g. the stem plot preserves the original, individual data) [4]

5 No. of different committees $= {}^{12}C_4 = \mathbf{495}$ [2]

No. of different committees with 2 men & 2 women $= {}^{7}C_2 \times {}^{5}C_2 = \mathbf{210}$ [4]

$$\text{p}(2 \text{ men}) = \frac{210}{495} = \mathbf{\frac{14}{33}}$$

[2]

6

$$r = \frac{\frac{1}{n}\sum xy - \overline{x}\,\overline{y}}{\sqrt{\frac{1}{n}\sum x^2 - \overline{x}^2}\sqrt{\frac{1}{n}\sum y^2 - \overline{y}^2}} = \frac{\frac{58440}{10} - 62.5\times 88.9}{\sqrt{\frac{41125}{10} - 62.5^2}\sqrt{\frac{83153}{10} - 88.9^2}} = \frac{287.75}{14.361...\times 20.3} = \mathbf{0.987}$$ [2]

The linear transformations of x and y will have no effect on the correlation coefficient. [1]

Regression line of y on x

$$y - \overline{y} = \frac{\frac{1}{n}\sum xy - \overline{x}\,\overline{y}}{\frac{1}{n}\sum x^2 - \overline{x}^2}(x - \overline{x})$$

$$x - 88.9 = \frac{287.75}{206.25}(y - 62.5)$$

$$\boldsymbol{y = 1.395x + 1.703}$$

for y=100 $\hat{x} = \mathbf{70.5}$ [5]

The y on x line is the only option since **x is a controlled variable**. [1]

7

$A \sim \mathrm{B}\left(3, \frac{1}{6}\right)$ $B \sim \mathrm{Geo}\left(\frac{1}{2}\right)$ [2], [2]

$$\mathrm{p}(A = 2) = {}^3C_2\left(\tfrac{1}{6}\right)^2\left(\tfrac{5}{6}\right) = \mathbf{\tfrac{5}{72}}$$ [2]

$$\mathrm{p}(B > 2) = \mathrm{p}(\text{first two throws are tails}) = \left(\tfrac{1}{2}\right)^2 = \mathbf{\tfrac{1}{4}}$$ [2]

$$\mathrm{p}(A = B) = \mathrm{p}\left(1,\ 2\ or\ 3\right) = \left({}^3C_1\left(\tfrac{1}{6}\right)\left(\tfrac{5}{6}\right)^2\right)\left(\tfrac{1}{2}\right) + \left({}^3C_2\left(\tfrac{1}{6}\right)^2\left(\tfrac{5}{6}\right)\right)\left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{6}\right)^3\left(\tfrac{1}{2}\right)^3 = \tfrac{331}{1728} = \mathbf{0.192}$$ [5]

[Total 60]