



**ADVANCED SUBSIDIARY GCE UNIT
MATHEMATICS**

4732/01

Probability & Statistics 1

TUESDAY 5 JUNE 2007

Afternoon

Time: 1 hour 30 minutes

Additional Materials: Answer Booklet (8 pages)
List of Formulae (MF1)

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 graphical 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 72.

ADVICE TO CANDIDATES

- Read each question carefully and make sure you know what you have to do before starting your answer.
- **You are reminded of the need for clear presentation in your answers.**

This document consists of **6** printed pages and **2** blank pages.

2

- 1 The table shows the probability distribution for a random variable X .

x	0	1	2	3
$P(X = x)$	0.1	0.2	0.3	0.4

Calculate $E(X)$ and $\text{Var}(X)$. [5]

- 2 Two judges each placed skaters from five countries in rank order.

Position	1st	2nd	3rd	4th	5th
Judge 1	UK	France	Russia	Poland	Canada
Judge 2	Russia	Canada	France	UK	Poland

Calculate Spearman's rank correlation coefficient, r_s , for the two judges' rankings. [5]

- 3 (i) How many different teams of 7 people can be chosen, without regard to order, from a squad of 15? [2]

(ii) The squad consists of 6 forwards and 9 defenders. How many different teams containing 3 forwards and 4 defenders can be chosen? [2]

- 4 A bag contains 6 white discs and 4 blue discs. Discs are removed at random, one at a time, **without** replacement.

(i) Find the probability that

(a) the second disc is blue, given that the first disc was blue, [1]

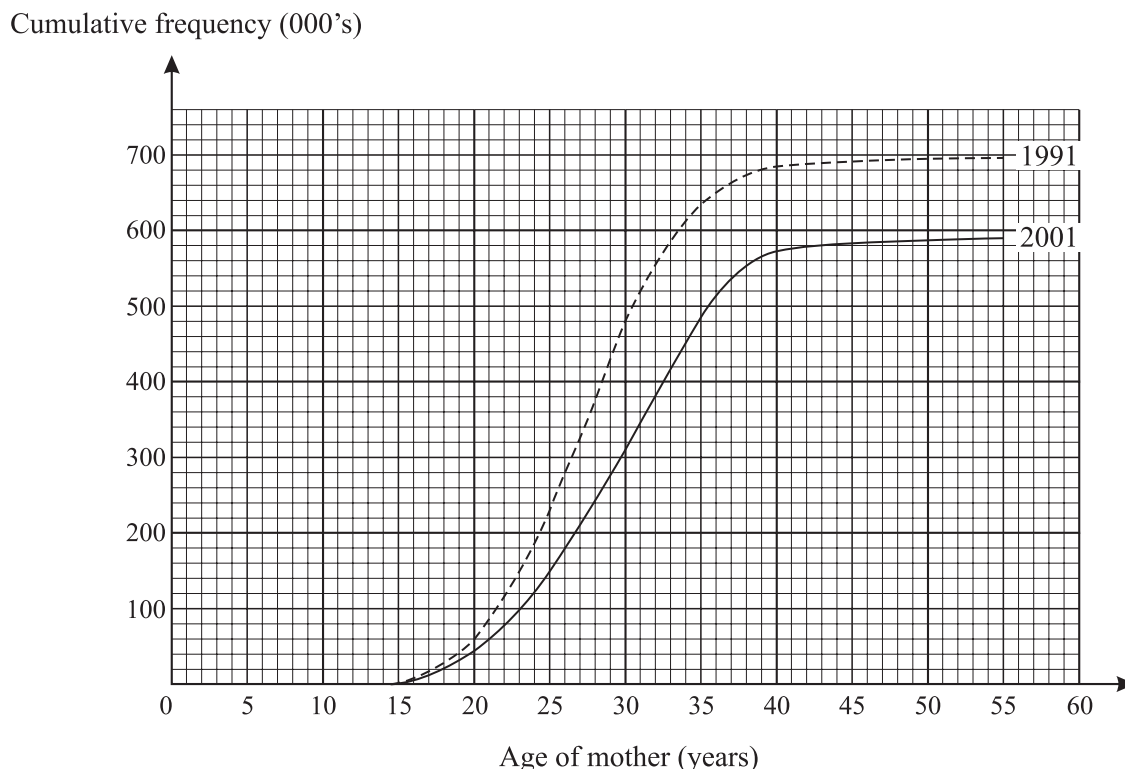
(b) the second disc is blue, [3]

(c) the third disc is blue, given that the first disc was blue. [3]

(ii) The random variable X is the number of discs which are removed up to and including the first blue disc. State whether the variable X has a geometric distribution. Explain your answer briefly. [1]

3

- 5 The numbers of births, in thousands, to mothers of different ages in England and Wales, in 1991 and 2001 are illustrated by the cumulative frequency curves.



- (i) In which of these two years were there more births? How many more births were there in this year? [2]
- (ii) The following quantities were estimated from the diagram.

Year	Median age (years)	Interquartile range (years)	Proportion of mothers giving birth aged below 25	Proportion of mothers giving birth aged 35 or above
1991	27.5	7.3	33%	9%
2001				18%

- (a) Find the values missing from the table. [5]
- (b) Did the women who gave birth in 2001 tend to be younger or older or about the same age as the women who gave birth in 1991? Using the table and your values from part (a), give two reasons for your answer. [3]

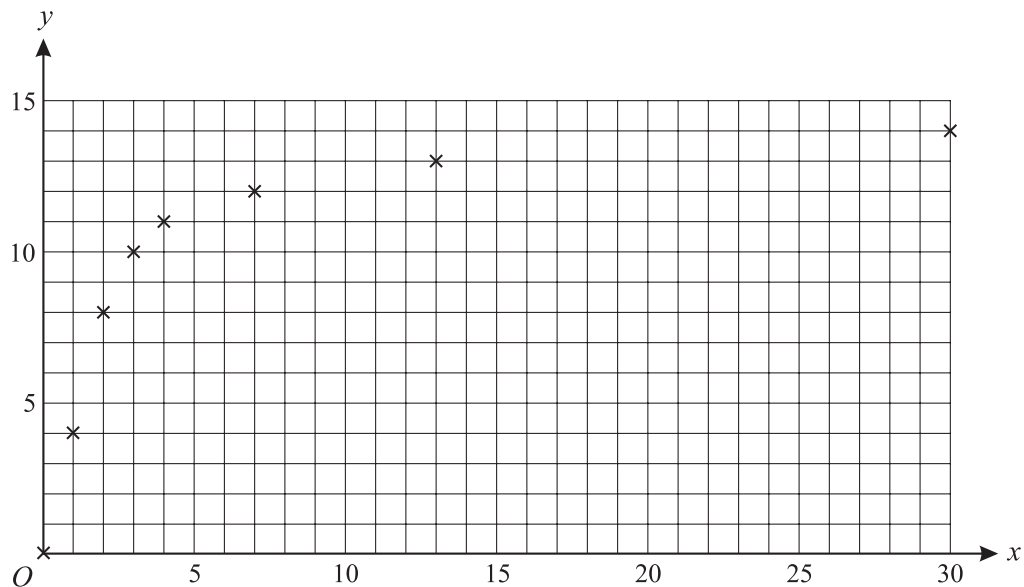
4

- 6 A machine with artificial intelligence is designed to improve its efficiency rating with practice. The table shows the values of the efficiency rating, y , after the machine has carried out its task various numbers of times, x .

x	0	1	2	3	4	7	13	30
y	0	4	8	10	11	12	13	14

$[n = 8, \Sigma x = 60, \Sigma y = 72, \Sigma x^2 = 1148, \Sigma y^2 = 810, \Sigma xy = 767.]$

These data are illustrated in the scatter diagram.



- (i) (a) Calculate the value of r , the product moment correlation coefficient. [3]
- (b) Without calculation, state with a reason the value of r_s , Spearman's rank correlation coefficient. [2]
- (ii) A researcher suggests that the data for $x = 0$ and $x = 1$ should be ignored. Without calculation, state with a reason what effect this would have on the value of
 - (a) r , [2]
 - (b) r_s . [2]
- (iii) Use the diagram to estimate the value of y when $x = 29$. [1]
- (iv) Jack finds the equation of the regression line of y on x for all the data, and uses it to estimate the value of y when $x = 29$. Without calculation, state with a reason whether this estimate or the one found in part (iii) will be the more reliable. [2]

5

7 On average, 25% of the packets of a certain kind of soup contain a voucher. Kim buys one packet of soup each week for 12 weeks. The number of vouchers she obtains is denoted by X .

(i) State two conditions needed for X to be modelled by the distribution $B(12, 0.25)$. [2]

In the rest of this question you should assume that these conditions are satisfied.

(ii) Find $P(X \leq 6)$. [2]

In order to claim a free gift, 7 vouchers are needed.

(iii) Find the probability that Kim will be able to claim a free gift at some time during the 12 weeks. [1]

(iv) Find the probability that Kim will be able to claim a free gift in the 12th week but not before. [4]

8 (i) A biased coin is thrown twice. The probability that it shows heads both times is 0.04. Find the probability that it shows tails both times. [3]

(ii) Another coin is biased so that the probability that it shows heads on any throw is p . The probability that the coin shows heads exactly once in two throws is 0.42. Find the two possible values of p . [5]

9 (i) A random variable X has the distribution $\text{Geo}(\frac{1}{5})$. Find

(a) $E(X)$, [2]

(b) $P(X = 4)$, [2]

(c) $P(X > 4)$. [2]

(ii) A random variable Y has the distribution $\text{Geo}(p)$, and $q = 1 - p$.

(a) Show that $P(Y \text{ is odd}) = p + q^2p + q^4p + \dots$. [1]

(b) Use the formula for the sum to infinity of a geometric progression to show that

$$P(Y \text{ is odd}) = \frac{1}{1+q}. \quad [4]$$

6

BLANK PAGE

BLANK PAGE

Permission to reproduce items where third-party owned material protected by copyright is included has been sought and cleared where possible. Every reasonable effort has been made by the publisher (OCR) to trace copyright holders, but if any items requiring clearance have unwittingly been included, the publisher will be pleased to make amends at the earliest possible opportunity.

OCR is part of the Cambridge Assessment Group. Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.