

**ADVANCED SUBSIDIARY GCE**

**4732/01**

**MATHEMATICS**

Probability & Statistics 1

**FRIDAY 23 MAY 2008**

Morning

Time: 1 hour 30 minutes

**Additional materials (enclosed):** None

**Additional materials (required):**

Answer Booklet (8 pages)

List of Formulae (MF1)

**INSTRUCTIONS TO CANDIDATES**

- Write your name in capital letters, your Centre Number and Candidate Number in the spaces provided on the Answer Booklet.
- Read each question carefully and make sure you know what you have to do before starting your answer.
- 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.

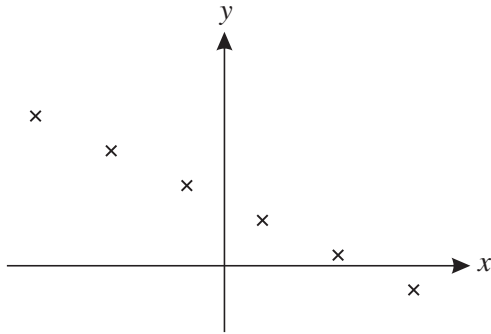
**INFORMATION FOR CANDIDATES**

- 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**.
- **You are reminded of the need for clear presentation in your answers.**

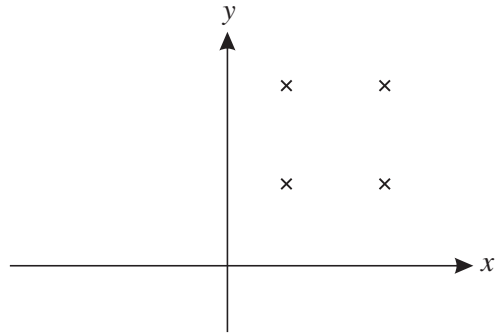
This document consists of **6** printed pages and **2** blank pages.

- 1 (i) State the value of the product moment correlation coefficient for each of the following scatter diagrams. [2]

(a)



(b)



- (ii) Calculate the value of Spearman's rank correlation coefficient for the following data. [5]

$x$	3.8	4.1	4.5	5.3
$y$	1.4	0.8	0.7	1.2

- 2 A class consists of 7 students from Ashville and 8 from Bewton. A committee of 5 students is chosen at random from the class.

(i) Find the probability that 2 students from Ashville and 3 from Bewton are chosen. [3]

(ii) In fact 2 students from Ashville and 3 from Bewton are chosen. In order to watch a video, all 5 committee members sit in a row. In how many different orders can they sit if no two students from Bewton sit next to each other? [2]

- 3 (i) A random variable  $X$  has the distribution  $B(8, 0.55)$ . Find

(a)  $P(X < 7)$ , [1]

(b)  $P(X = 5)$ , [2]

(c)  $P(3 \leq X < 6)$ . [3]

(ii) A random variable  $Y$  has the distribution  $B(10, \frac{5}{12})$ . Find

(a)  $P(Y = 2)$ , [2]

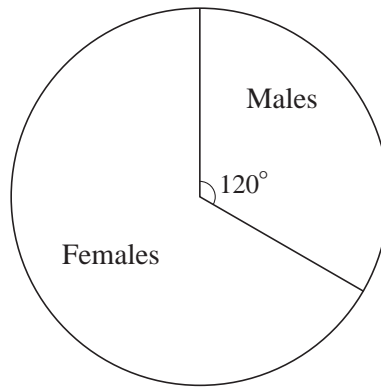
(b)  $\text{Var}(Y)$ . [1]

- 4 At a fairground stall, on each turn a player receives prize money with the following probabilities.

Prize money	£0.00	£0.50	£5.00
Probability	$\frac{17}{20}$	$\frac{1}{10}$	$\frac{1}{20}$

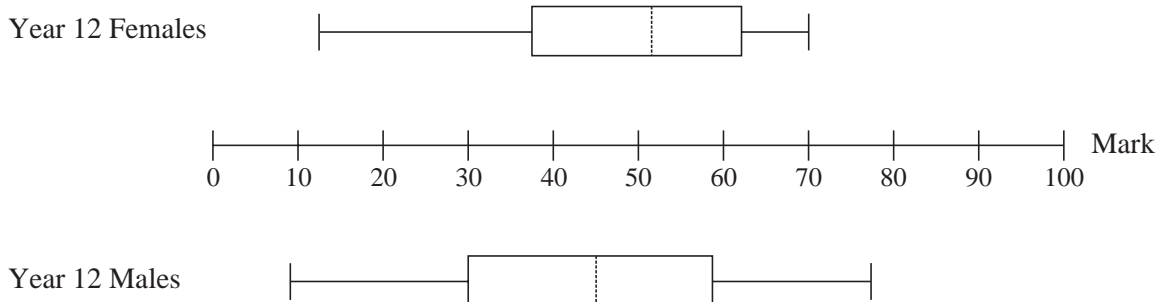
- (i) Find the probability that a player who has two turns will receive a total of £5.50 in prize money. [3]
- (ii) The stall-holder wishes to make a profit of 20p per turn on average. Calculate the amount the stall-holder should charge for each turn. [4]
- 5 (i) A bag contains 12 red discs and 10 black discs. Two discs are removed at random, without replacement. Find the probability that both discs are red. [2]
- (ii) Another bag contains 7 green discs and 8 blue discs. Three discs are removed at random, without replacement. Find the probability that exactly two of these discs are green. [3]
- (iii) A third bag contains 45 discs, each of which is either yellow or brown. Two discs are removed at random, without replacement. The probability that both discs are yellow is  $\frac{1}{15}$ . Find the number of yellow discs which were in the bag at first. [4]

- 6 (i) The numbers of males and females in Year 12 at a school are illustrated in the pie chart. The number of males in Year 12 is 128.



Year 12

- (a) Find the number of females in Year 12. [1]
- (b) On a corresponding pie chart for Year 13, the angle of the sector representing males is  $150^\circ$ . Explain why this does not necessarily mean that the number of males in Year 13 is more than 128. [1]
- (ii) All the Year 12 students took a General Studies examination. The results are illustrated in the box-and-whisker plots.



- (a) One student said “The Year 12 pie chart shows that there are more females than males, but the box-and-whisker plots show that there are more males than females.”  
Comment on this statement. [1]
- (b) Give two comparisons between the overall performance of the females and the males in the General Studies examination. [2]
- (c) Give one advantage and one disadvantage of using box-and-whisker plots rather than histograms to display the results. [2]
- (iii) The mean mark for 102 of the male students was 51. The mean mark for the remaining 26 male students was 59. Calculate the mean mark for all 128 male students. [3]

7 Once each year, Paula enters a lottery for a place in an annual marathon. Each time she enters the lottery, the probability of her obtaining a place is 0.3. Find the probability that

(i) the first time she obtains a place is on her 4th attempt, [3]

(ii) she does not obtain a place on any of her first 6 attempts, [2]

(iii) she needs fewer than 10 attempts to obtain a place, [3]

(iv) she obtains a place exactly twice in her first 5 attempts. [3]

8 A city council attempted to reduce traffic congestion by introducing a congestion charge. The charge was set at £4.00 for the first year and was then increased by £2.00 each year. For each of the first eight years, the council recorded the average number of vehicles entering the city centre per day. The results are shown in the table.

Charge, £ $x$	4	6	8	10	12	14	16	18
Average number of vehicles per day, $y$ million	2.4	2.5	2.2	2.3	2.0	1.8	1.7	1.5

$$[n = 8, \Sigma x = 88, \Sigma y = 16.4, \Sigma x^2 = 1136, \Sigma y^2 = 34.52, \Sigma xy = 168.6.]$$

(i) Calculate the product moment correlation coefficient for these data. [3]

(ii) Explain why  $x$  is the independent variable. [1]

(iii) Calculate the equation of the regression line of  $y$  on  $x$ . [4]

(iv) (a) Use your equation to estimate the average number of vehicles which will enter the city centre per day when the congestion charge is raised to £20.00. [2]

(b) Comment on the reliability of your estimate. [2]

(v) The council wishes to estimate the congestion charge required to reduce the average number of vehicles entering the city per day to 1.0 million. Assuming that a reliable estimate can be made by extrapolation, state whether they should use the regression line of  $y$  on  $x$  or the regression line of  $x$  on  $y$ . Give a reason for your answer. [2]

**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.