OCR
RECOGNISING ACHIEVEMENT

OXFORD CAMBRIDGE AND RSA EXAMINATIONS

Advanced Subsidiary General Certificate of Education
Advanced General Certificate of Education

MATHEMATICS 4732

Probability and Statistics 1

Specimen Paper

Additional materials:
Answer booklet
Graph paper
List of Formulae (MF 1)

TIME 1 hour 30 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 72.
- 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.

Registered Charity Number: 1066969
[Turn over

1 Janet and John wanted to compare their daily journey times to work, so they each kept a record of their journey times for a few weeks.

(i) Janet's daily journey times, x minutes, for a period of 25 days, were summarised by $\Sigma x = 2120$ and $\Sigma x^2 = 180\,044$. Calculate the mean and standard deviation of Janet's journey times. [3]

(ii) John's journey times had a mean of 79.7 minutes and a standard deviation of 6.22 minutes. Describe briefly, in everyday terms, how Janet and John's journey times compare. [2]

2 Two independent assessors awarded marks to each of 5 projects. The results were as shown in the table.

Project	*A*	*B*	*C*	*D*	*E*
First assessor	38	91	62	83	61
Second assessor	56	84	41	85	62

(i) Calculate Spearman's rank correlation coefficient for the data. [5]

(ii) Show, by sketching a suitable scatter diagram, how two assessors might have assessed 5 projects in such a way that Spearman's rank correlation coefficient for their marks was $+1$ while the product moment correlation coefficient for their marks was not $+1$. (Your scatter diagram need not be drawn accurately to scale.) [2]

3 Five friends, Ali, Bev, Carla, Don and Ed, stand in a line for a photograph.

(i) How many different possible arrangements are there if Ali, Bev and Carla stand next to each other? [2]

(ii) How many different possible arrangements are there if none of Ali, Bev and Carla stand next to each other? [3]

(iii) If all possible arrangements are equally likely, find the probability that two of Ali, Bev and Carla are next to each other, but the third is not next to either of the other two. [3]

4 Each packet of the breakfast cereal Fizz contains one plastic toy animal. There are five different animals in the set, and the cereal manufacturers use equal numbers of each. Without opening a packet it is impossible to tell which animal it contains. A family has already collected four different animals at the start of a year and they now need to collect an elephant to complete their set. The family is interested in how many packets they will need to buy before they complete their set.

(i) Name an appropriate distribution with which to model this situation. State the value(s) of any parameter(s) of the distribution, and state also any assumption(s) needed for the distribution to be a valid model. [3]

(ii) Find the probability that the family will complete their set with the third packet they buy after the start of the year. [2]

(iii) Find the probability that, in order to complete their collection, the family will need to buy more than 4 packets after the start of the year. [3]

5 A sixth-form class consists of 7 girls and 5 boys. Three students from the class are chosen at random. The number of boys chosen is denoted by the random variable X. Show that

(i) $\mathrm{P}(X = 0) = \frac{7}{44}$, [2]

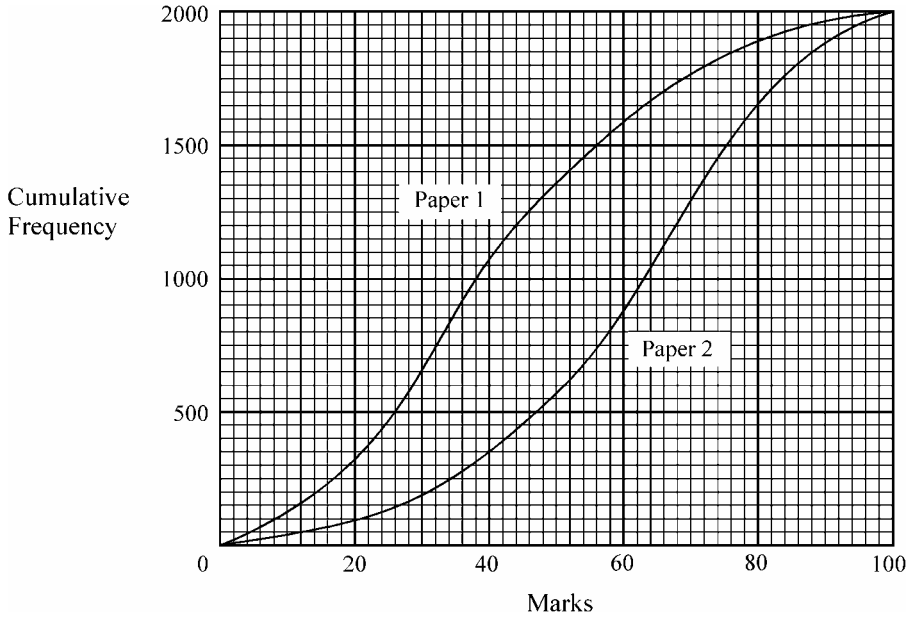
(ii) $\mathrm{P}(X = 2) = \frac{7}{22}$. [3]

The complete probability distribution of X is shown in the following table.

x	0	1	2	3
$\mathrm{P}(X = x)$	$\frac{7}{44}$	$\frac{21}{44}$	$\frac{7}{22}$	$\frac{1}{22}$

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

6



The diagram shows the cumulative frequency graphs for the marks scored by the candidates in an examination. The 2000 candidates each took two papers; the upper curve shows the distribution of marks on paper 1 and the lower curve shows the distribution on paper 2. The maximum mark on each paper was 100.

(i) Use the diagram to estimate the median mark for each of paper 1 and paper 2. [3]

(ii) State with a reason which of the two papers you think was the easier one. [2]

(iii) To achieve grade A on paper 1 candidates had to score 66 marks out of 100. What mark on paper 2 gives equal proportions of candidates achieving grade A on the two papers? What is this proportion? [4]

(iv) The candidates' marks for the two papers could also be illustrated by means of a pair of box-and whisker plots. Give two brief comments comparing the usefulness of cumulative frequency graphs and box-and-whisker plots for representing the data. [2]

7 Items from a production line are examined for any defects. The probability that any item will be found to be defective is 0.15, independently of all other items.

(i) A batch of 16 items is inspected. Using tables of cumulative binomial probabilities, or otherwise, find the probability that

(a) at least 4 items in the batch are defective, [2]

(b) exactly 4 items in the batch are defective. [2]

(ii) Five batches, each containing 16 items, are taken.

(a) Find the probability that at most 2 of these 5 batches contain at least 4 defective items. [4]

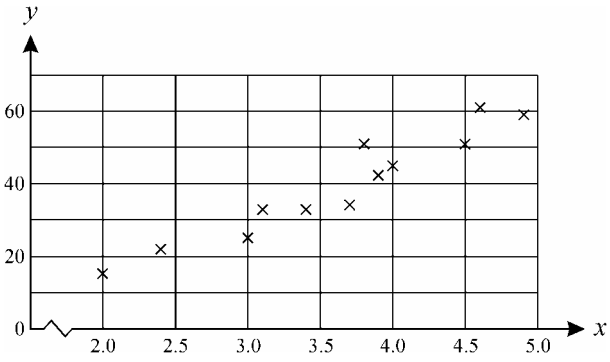
(b) Find the expected number of batches that contain at least 4 defective items. [2]

8 An experiment was conducted to see whether there was any relationship between the maximum tidal current, $y\,\text{cm}\,\text{s}^{-1}$, and the tidal range, x metres, at a particular marine location. [The *tidal range* is the difference between the height of high tide and the height of low tide.] Readings were taken over a period of 12 days, and the results are shown in the following table.

x	2.0	2.4	3.0	3.1	3.4	3.7	3.8	3.9	4.0	4.5	4.6	4.9
y	15.2	22.0	25.2	33.0	33.1	34.2	51.0	42.3	45.0	50.7	61.0	59.2

$[\Sigma x = 43.3, \Sigma y = 471.9, \Sigma x^2 = 164.69, \Sigma y^2 = 20\,915.75, \Sigma xy = 1837.78.]$

The scatter diagram below illustrates the data.



(i) Calculate the product moment correlation coefficient for the data, and comment briefly on your answer with reference to the appearance of the scatter diagram. [4]

(ii) Calculate the equation of the regression line of maximum tidal current on tidal range. [4]

(iii) Estimate the maximum tidal current on a day when the tidal range is 4.2 m, and comment briefly on how reliable you consider your estimate is likely to be. [3]

(iv) It is suggested that the equation found in part **(ii)** could be used to predict the maximum tidal current on a day when the tidal range is 15 m. Comment briefly on the validity of this suggestion. [2]