

**ADVANCED GCE
MATHEMATICS (MEI)**

Statistics 2

4767

QUESTION PAPER

Candidates answer on the printed answer book.

OCR supplied materials:

- Printed answer book 4767
- MEI Examination Formulae and Tables (MF2)

Other materials required:

- Scientific or graphical calculator

**Monday 24 January 2011
Morning**

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the printed answer book and the question paper.

- The question paper will be found in the centre of the printed answer book.
- Write your name, centre number and candidate number in the spaces provided on the printed answer book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the printed answer book.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- Use black ink. Pencil may be used for graphs and diagrams only.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Answer **all** the questions.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Final answers should be given to a degree of accuracy appropriate to the context.

INFORMATION FOR CANDIDATES

This information is the same on the printed answer book and the question paper.

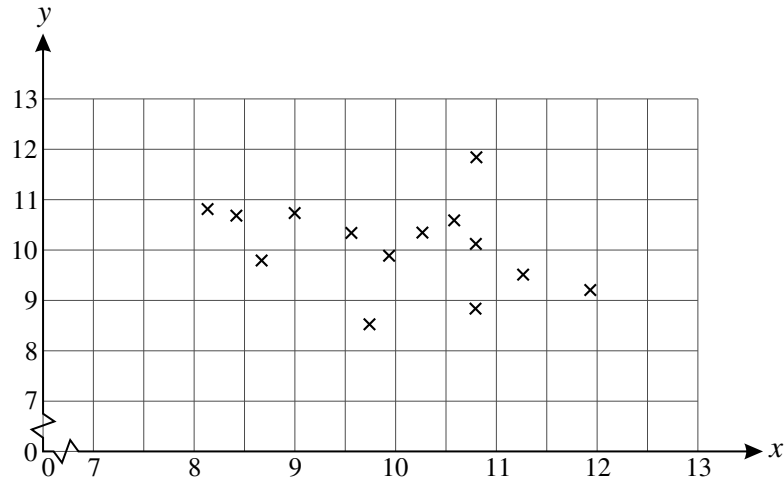
- The number of marks is given in brackets [] at the end of each question or part question on the question paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for this paper is **72**.
- The printed answer book consists of **12** pages. The question paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER / INVIGILATOR

- Do not send this question paper for marking; it should be retained in the centre or destroyed.

- 1 The scatter diagram below shows the birth rates x , and death rates y , measured in standard units, in a random sample of 14 countries in a particular year. Summary statistics for the data are as follows.

$$\Sigma x = 139.8 \quad \Sigma y = 140.4 \quad \Sigma x^2 = 1411.66 \quad \Sigma y^2 = 1417.88 \quad \Sigma xy = 1398.56 \quad n = 14$$



- (i) Calculate the sample product moment correlation coefficient. [5]
- (ii) Carry out a hypothesis test at the 5% significance level to determine whether there is any correlation between birth rates and death rates. [6]
- (iii) State the distributional assumption which is necessary for this test to be valid. Explain briefly in the light of the scatter diagram why it appears that the assumption may be valid. [2]
- (iv) The values of x and y for another country in that year are 14.4 and 7.8 respectively. If these values are included, the value of the sample product moment correlation coefficient is -0.5694 . Explain why this one observation causes such a large change to the value of the sample product moment correlation coefficient. Discuss whether this brings the validity of the test into question. [4]

- 2 A student is investigating the numbers of sultanas in a particular brand of biscuit. The data in the table show the numbers of sultanas in a random sample of 50 of these biscuits.

Number of sultanas	0	1	2	3	4	5	>5
Frequency	8	15	12	9	4	2	0

- (i) Show that the sample mean is 1.84 and calculate the sample variance. [3]
- (ii) Explain why these results support a suggestion that a Poisson distribution may be a suitable model for the distribution of the numbers of sultanas in this brand of biscuit. [1]

For the remainder of the question you should assume that a Poisson distribution with mean 1.84 is a suitable model for the distribution of the numbers of sultanas in these biscuits.

- (iii) Find the probability of
- (A) no sultanas in a biscuit,
- (B) at least two sultanas in a biscuit. [5]
- (iv) Show that the probability that there are at least 10 sultanas in total in a packet containing 5 biscuits is 0.4389. [3]
- (v) Six packets each containing 5 biscuits are selected at random. Find the probability that exactly 2 of the six packets contain at least 10 sultanas. [3]
- (vi) Sixty packets each containing 5 biscuits are selected at random. Use a suitable approximating distribution to find the probability that more than half of the sixty packets contain at least 10 sultanas. [5]

- 3 The random variable X represents the reaction times, in milliseconds, of men in a driving simulator. X is Normally distributed with mean 355 and standard deviation 52.

- (i) Find
- (A) $P(X < 325)$,
- (B) $P(300 < X < 400)$. [6]
- (ii) Find the value of k for which $P(X < k) = 0.2$. [3]

It is thought that women may have a different mean reaction time from men. In order to test this, a random sample of 25 women is selected. The mean reaction time of these women in the driving simulator is 344 milliseconds. You may assume that women's reaction times are also Normally distributed with standard deviation 52 milliseconds. A hypothesis test is carried out to investigate whether women have a different mean reaction time from men.

- (iii) Carry out the test at the 5% significance level. [8]

- 4 A researcher is investigating the sizes of pebbles at various locations in a river. Three sites in the river are chosen and each pebble sampled at each site is classified as large, medium or small. The results are as follows.

		Site			Row totals
		A	B	C	
Pebble size	Large	15	12	10	37
	Medium	28	17	45	90
	Small	47	33	36	116
Column totals		90	62	91	243

- (i) Carry out a test at the 5% significance level to examine whether there is any association between pebble size and site. Your working should include a table of the contributions of each cell to the test statistic. [12]
- (ii) By referring to each site, comment briefly on how the size of the pebbles compares with what would be expected if there were no association. You should support your answers by referring to your table of contributions. [6]

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series. If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity. For queries or further information please contact the Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

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.