



Tuesday 10 June 2014 – Morning

A2 GCE MATHEMATICS (MEI)

4767/01 Statistics 2

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:

- Printed Answer Book 4767/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:

- Scientific or graphical calculator

Duration: 1 hour 30 minutes

MODIFIED LANGUAGE

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside 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. HB 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.

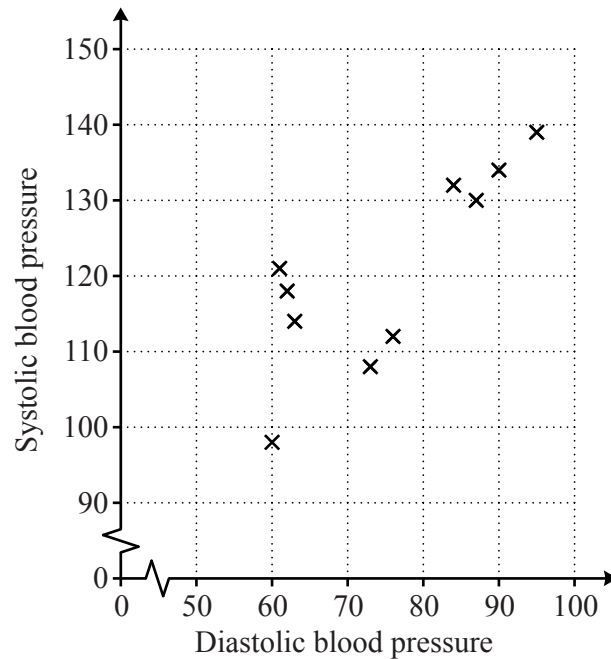
- The number of marks is given in brackets [] at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks for 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 recycled. Please contact OCR Copyright should you wish to re-use this document.

- 1 A medical student is investigating the claim that young adults with high diastolic blood pressure tend to have high systolic blood pressure as well. The student measures the diastolic and systolic blood pressures of a random sample of ten young adults. The data are shown in the table and illustrated in the scatter diagram.

Diastolic blood pressure	60	61	62	63	73	76	84	87	90	95
Systolic blood pressure	98	121	118	114	108	112	132	130	134	139



- (i) Calculate the value of Spearman's rank correlation coefficient for these data. [5]
- (ii) Carry out a hypothesis test at the 5% significance level to examine whether there is positive association between diastolic blood pressure and systolic blood pressure in the population of young adults. [6]
- (iii) In the light of the scatter diagram, explain why it might not be valid to carry out a test based on the product moment correlation coefficient. [2]

The product moment correlation coefficient between the diastolic and systolic blood pressures of a random sample of 10 athletes is 0.707.

- (iv) Carry out a hypothesis test at the 1% significance level to investigate whether there appears to be positive correlation between these two variables in the population of athletes. You may assume that in this case such a test is valid. [5]

- 2 When a particular type of aluminium sheet is manufactured, defects (faults) occur randomly, independently and at a constant average rate of 1.7 defects per square metre.

(i) Explain the meaning of the term ‘independently’ and name the distribution that models this situation. [2]

(ii) Find the probability that there are exactly 2 defects in a sheet of area 1 square metre. [2]

(iii) Find the probability that there are exactly 12 defects in a sheet of area 7 square metres. [3]

In another type of aluminium sheet, defects occur randomly, independently and at a constant average rate of 0.8 defects per square metre.

(iv) A large box is made from 2 square metres of the first type of sheet and 2 square metres of the second type of sheet, chosen independently. Show that the probability that there are at least 8 defects altogether in the box is 0.1334. [3]

A random sample of 100 of these boxes is selected.

(v) State the exact distribution of the number of boxes which have at least 8 defects. [2]

(vi) Use a suitable approximating distribution to find the probability that there are at least 20 boxes in the sample which have at least 8 defects. [5]

- 3 The wing lengths of native English male blackbirds, measured in mm, are Normally distributed with mean 130.5 and variance 11.84.

(i) Find the probability that a randomly selected native English male blackbird has a wing length greater than 135 mm. [3]

(ii) Given that 1% of native English male blackbirds have wing length more than k mm, find the value of k . [3]

(iii) Find the probability that a randomly selected native English male blackbird has a wing length which is 131 mm correct to the nearest millimetre. [3]

It is suspected that Scandinavian male blackbirds have, on average, longer wings than native English male blackbirds. A random sample of 20 Scandinavian male blackbirds has mean wing length 132.4 mm. You may assume that wing lengths in this population are Normally distributed with variance 11.84 mm^2 .

(iv) Carry out an appropriate hypothesis test, at the 5% significance level. [8]

(v) Discuss briefly one advantage and one disadvantage of using a 10% significance level rather than a 5% significance level in hypothesis testing in general. [2]

- 4 A researcher at a large company thinks that there may be some relationship between the numbers of working days lost due to illness per year and the ages of the workers in the company. The researcher selects a random sample of 190 workers. The ages of the workers and numbers of days lost for a period of 1 year are summarised below.

		Working days lost		
		0 to 4	5 to 9	10 or more
Age	Under 35	31	27	4
	35 to 50	28	32	8
	Over 50	16	28	16

- (i) Carry out a test at the 1% significance level to investigate whether the researcher's belief appears to be true. Your working should include a table showing the contributions of each cell to the test statistic. **[10]**
- (ii) For the 'Over 50' age group, comment briefly on how the working days lost compare with what would be expected if there were no association. **[3]**
- (iii) A student decides to reclassify the 'working days lost' into two groups, '0 to 4' and '5 or more'. The age groups are unchanged. The test statistic with this classification is 7.08. Carry out the test at the 1% level with this new classification, using the same hypotheses as for the original test. **[3]**
- (iv) Comment on the results of the two tests. **[2]**

END OF QUESTION PAPER



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.