



**Thursday 6 June 2013 – Morning**

**A2 GCE MATHEMATICS**

**4734/01** Probability & Statistics 3

**QUESTION PAPER**

Candidates answer on the Printed Answer Book.

**OCR supplied materials:**

- Printed Answer Book 4734/01
- List of Formulae (MF1)

**Other materials required:**

- Scientific or graphical calculator

**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. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Do **not** write in the bar codes.
- You are permitted to use a scientific or graphical calculator in this paper.
- 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.

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

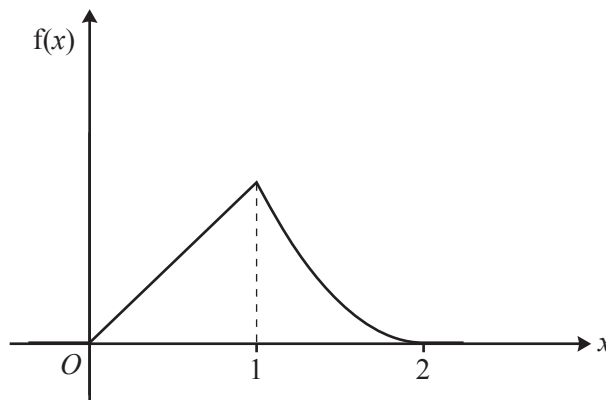
**INSTRUCTION TO EXAMS OFFICER/INVIGILATOR**

- 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.

2

- 1 The blood-test procedure at a clinic is that a person arrives, takes a numbered ticket and waits for that number to be called. The waiting times between the numbers called have independent normal distributions with mean 3.5 minutes and standard deviation 0.9 minutes. My ticket is number 39 and as I take my ticket number 1 is being called, so that I have to wait for 38 numbers to be called. Find the probability that I will have to wait between 120 minutes and 140 minutes. [6]
- 2 In order to estimate the total number of rabbits in a certain region, a random sample of 500 rabbits is captured, marked and released. After two days a random sample of 250 rabbits is captured and 24 are found to be marked. It may be assumed that there is no change in the population during the two days.
- (i) Estimate the total number of rabbits in the region. [2]
  - (ii) Calculate an approximate 95% confidence interval for the population proportion of marked rabbits. [4]
  - (iii) Using your answer to part (ii), estimate a 95% confidence interval for the total number of rabbits in the region. [3]

3



The continuous random variable  $X$  has probability density function given by

$$f(x) = \begin{cases} ax & 0 < x \leq 1, \\ b(2 - x)^2 & 1 < x \leq 2, \\ 0 & \text{otherwise,} \end{cases}$$

where  $a$  and  $b$  are constants. The graph is shown in the above diagram.

- (i) Find the values of  $a$  and  $b$ . [5]
  - (ii) Find the value of  $E\left(\frac{1}{X}\right)$ . [3]
- 4 A new computer was bought by a local council to search council records and was tested by an employee. She searched a random sample of 500 records and the sample mean search time was found to be 2.18 milliseconds and an unbiased estimate of variance was  $1.58^2$  milliseconds<sup>2</sup>.
- (i) Calculate a 98% confidence interval for the population mean search time  $\mu$  milliseconds. [3]
  - (ii) It is required to obtain a sample mean time that differs from  $\mu$  by less than 0.05 milliseconds with probability 0.95. Estimate the sample size required. [4]
  - (iii) State why it is unnecessary for the validity of your calculations that search time has a normal distribution. [1]

## 3

- 5 The continuous random variable  $Y$  has probability density function given by

$$f(y) = \begin{cases} \ln(y) & 1 \leq y \leq e, \\ 0 & \text{otherwise.} \end{cases}$$

- (i) Verify that this is a valid probability density function. [4]

- (ii) Show that the (cumulative) distribution function of  $Y$  is given by

$$F(y) = \begin{cases} 0 & y < 1, \\ y \ln y - y + 1 & 1 \leq y \leq e, \\ 1 & \text{otherwise.} \end{cases} \quad [3]$$

- (iii) Verify that the upper quartile of  $Y$  lies in the interval  $[2.45, 2.46]$ . [2]

- (iv) Find the (cumulative) distribution function of  $X$  where  $X = \ln Y$ . [4]

- 6 A random sample of 80 students who had all studied Biology, Chemistry and Art at a college was each asked which they enjoyed most. The results, classified according to gender, are given in the table.

		Subject		
		Biology	Chemistry	Art
Gender	Male	13	4	11
	Female	37	8	7

It is required to carry out a test of independence between subject most enjoyed and gender at the  $2\frac{1}{2}\%$  significance level.

- (i) Calculate the expected values for the cells. [3]

- (ii) Explain why it is necessary to combine cells, and choose a suitable combination. [2]

- (iii) Carry out the test. [8]

- 7 Two machines  $A$  and  $B$  both pack cartons in a factory. The mean packing times are compared by timing the packing of 10 randomly chosen cartons from machine  $A$  and 8 randomly chosen cartons from machine  $B$ . The times,  $t$  seconds, taken to pack these cartons are summarised below.

	Sample size	$\Sigma t$	$\Sigma t^2$
Machine $A$	10	221.4	4920.9
Machine $B$	8	199.2	4980.3

The packing times have independent normal distributions.

- (i) Stating a necessary assumption, carry out a test, at the 1% significance level, of whether the population mean packing times differ for the two machines. [12]

- (ii) Find the largest possible value of the constant  $c$  for which there is evidence at the 1% significance level that  $\mu_B - \mu_A > c$ , where  $\mu_B$  and  $\mu_A$  denote the respective population mean packing times in seconds. [3]

**THERE ARE NO QUESTIONS PRINTED ON THIS PAGE.**



**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](http://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.