



**ADVANCED GCE  
MATHEMATICS**

Probability & Statistics 3

**4734**

Candidates answer on the answer booklet.

**OCR supplied materials:**

- 8 page answer booklet (sent with general stationery)
- List of Formulae (MF1)

**Other materials required:**

- Scientific or graphical calculator

**Wednesday 19 January 2011  
Afternoon**

**Duration:** 1 hour 30 minutes



**INSTRUCTIONS TO CANDIDATES**

- Write your name, centre number and candidate number in the spaces provided on the answer booklet. Please write clearly and in capital letters.
- 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.
- 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 scientific or 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.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- This document consists of **4** pages. Any blank pages are indicated.

- 1 A random variable has a normal distribution with unknown mean  $\mu$  and known standard deviation 0.19. In order to estimate  $\mu$  a random sample of five observations of the random variable was taken. The values were as follows.

5.44    4.93    5.12    5.36    5.40

Using these five values, calculate,

(i) an estimate of  $\mu$ , [1]

(ii) a 95% confidence interval for  $\mu$ . [4]

- 2 In a Year 8 internal examination in a large school the Geography marks,  $G$ , and Mathematics marks,  $M$ , had means and standard deviations as follows.

|     | Mean  | Standard deviation |
|-----|-------|--------------------|
| $G$ | 36.42 | 6.87               |
| $M$ | 42.65 | 10.25              |

Assuming that  $G$  and  $M$  have independent normal distributions, find the probability that a randomly chosen Geography candidate scores at least 10 marks more than a randomly chosen Mathematics candidate. Do not use a continuity correction. [5]

- 3 The continuous random variable  $T$  has probability density function given by

$$f(t) = \begin{cases} 0 & t < 0, \\ \frac{a}{e} & 0 \leq t < 2, \\ ae^{-\frac{1}{2}t} & t \geq 2, \end{cases}$$

where  $a$  is a positive constant.

(i) Show that  $a = \frac{1}{4}e$ . [3]

(ii) Find the upper quartile of  $T$ . [4]

- 4 A study in 1981 investigated the effect of water fluoridation on children's dental health. In a town with fluoridation, 61 out of a random sample of 107 children showed signs of increased tooth decay after six months. In a town without fluoridation the corresponding number was 106 out of a random sample of 143 children. The population proportions of children with increased tooth decay are denoted by  $p_1$  and  $p_2$  for the towns with fluoridation and without fluoridation respectively. A test is carried out of the null hypothesis  $p_1 = p_2$  against the alternative hypothesis  $p_1 < p_2$ . Find the smallest significance level at which the null hypothesis is rejected. [7]

5 An experiment with hybrid corn resulted in yellow kernels and purple kernels. Of a random sample of 90 kernels, 18 were yellow and 72 were purple.

- (i) Calculate an approximate 90% confidence interval for the proportion of yellow kernels produced in all such experiments. [4]
- (ii) Deduce an approximate 90% confidence interval for the proportion of purple kernels produced in all such experiments. [1]
- (iii) Explain what is meant by a 90% confidence interval for a population proportion. [2]
- (iv) Mendel’s theory of inheritance predicts that 25% of all such kernels will be yellow. State, giving a reason, whether or not your calculations support the theory. [2]

6 The continuous random variable  $X$  has (cumulative) distribution function given by

$$F(x) = \begin{cases} 0 & x < \frac{1}{2}, \\ \frac{2x - 1}{x + 1} & \frac{1}{2} \leq x \leq 2, \\ 1 & x > 2. \end{cases}$$

- (i) Given that  $Y = \frac{1}{X}$ , find the (cumulative) distribution function of  $Y$ , and deduce that  $Y$  and  $X$  have identical distributions. [6]
- (ii) Find  $E(X + 1)$  and deduce the value of  $E\left(\frac{1}{X}\right)$ . [6]

7 (i) When should Yates’ correction be applied when carrying out a  $\chi^2$  test? [1]

Two vaccines against typhoid fever,  $A$  and  $B$ , were tested on a total of 700 people in Nepal during a particular year. The vaccines were allocated randomly and whether or not typhoid had developed was noted during the following year. The results are shown in the table.

|                         | Vaccines |     |
|-------------------------|----------|-----|
|                         | $A$      | $B$ |
| Developed typhoid       | 19       | 4   |
| Did not develop typhoid | 310      | 367 |

- (ii) Carry out a suitable  $\chi^2$  test at the 1% significance level to determine whether the outcome depends on the vaccine used. Comment on the result. [10]

[Question 8 is printed overleaf.]

- 8 (i) State circumstances under which it would be necessary to calculate a pooled estimate of variance when carrying out a two-sample hypothesis test. [1]
- (ii) An investigation into whether passive smoking affects lung capacity considered a random sample of 20 children whose parents did not smoke and a random sample of 22 children whose parents did smoke. None of the children themselves smoked. The lung capacity, in litres, of each child was measured and the results are summarised as follows.

For the children whose parents did not smoke:  $n_1 = 20$ ,  $\Sigma x_1 = 42.4$  and  $\Sigma x_1^2 = 90.43$ .

For the children whose parents did smoke:  $n_2 = 22$ ,  $\Sigma x_2 = 42.5$  and  $\Sigma x_2^2 = 82.93$ .

The means of the two populations are denoted by  $\mu_1$  and  $\mu_2$  respectively.

- (a) State conditions for which a  $t$ -test would be appropriate for testing whether  $\mu_1$  exceeds  $\mu_2$ . [1]
- (b) Assuming the conditions are valid, carry out the test at the 1% significance level and comment on the result. [11]
- (c) Calculate a 99% confidence interval for  $\mu_1 - \mu_2$ . [3]

**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.