OCR
RECOGNISING ACHIEVEMENT

ADVANCED GCE

MATHEMATICS (MEI) **4768**

Statistics 3

Candidates answer on the answer booklet.

OCR supplied materials:
- 8 page answer booklet (sent with general stationery)
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Thursday 23 June 2011
Morning

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.
- 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

- The number of marks is given in brackets **[ ]** at the end of each question or part question.
- 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**.
- This document consists of **4** pages. Any blank pages are indicated.

© OCR 2011 [R/102/2658]
RP–0I23

OCR is an exempt Charity

1 Gerry runs 5000-metre races for his local athletics club. His coach has been monitoring his practice times for several months and he believes that they can be modelled using a Normal distribution with mean 15.3 minutes. The coach suggests that Gerry should try running with a pacemaker in order to see if this can improve his times. Subsequently a random sample of 10 of Gerry's times with the pacemaker is collected to see if any reduction has been achieved. The sample of times (in minutes) is as follows.

14.86 15.00 15.62 14.44 15.27 15.64 14.58 14.30 15.08 15.08

(i) Why might a t test be used for these data? **[2]**

(ii) Using a 5% significance level, carry out the test to see whether, on average, Gerry's times have been reduced. **[9]**

(iii) What is meant by 'a 5% significance level'? What would be the consequence of decreasing the significance level? **[3]**

(iv) Find a 95% confidence interval for the true mean of Gerry's times using a pacemaker. **[4]**

2 Scientists researching into the chemical composition of dust in space collect specimens using a specially designed spacecraft. The craft collects the particles of dust in trays that are made up of a large array of cells containing aerogel. The aerogel traps the particles that penetrate into the cells.

(i) For a random sample of 100 cells, the number of particles of dust in each cell was counted, giving the following results.

Number of particles	0	1	2	3	4	5	6	7	8	9	10+
Frequency	4	7	10	20	17	15	10	9	5	3	0

It is thought that the number of particles collected in each cell can be modelled using the distribution Poisson(4.2) since 4.2 is the sample mean for these data.

Some of the calculations for a χ^2 test are shown below. The cells for 8, 9 and 10+ particles have been combined.

Number of particles	5	6	7	8+
Observed frequency	15	10	9	8
Expected frequency	16.33	11.44	6.86	6.39
Contribution to X^2	0.1083	0.1813	0.6676	0.4056

Complete the calculations and carry out the test using a 10% significance level to see whether the number of particles per cell may be modelled in this way. **[12]**

(ii) The diameters of the dust particles are believed to be distributed symmetrically about a median of 15 micrometres (μm). For a random sample of 20 particles, the sum of the signed ranks of the diameters of the particles smaller than 15 μm (W_-) is found to be 53. Test at the 5% level of significance whether the median diameter appears to be more than 15 μm. **[6]**

3 The time, in hours, until an electronic component fails is represented by the random variable X. In this question two models for X are proposed.

(i) In one model, X has cumulative distribution function

$$\mathrm{G}(x) = \begin{cases} 0 & x \leqslant 0, \\ 1 - \left(1 + \dfrac{x}{200}\right)^{-2} & x > 0. \end{cases}$$

(A) Sketch $\mathrm{G}(x)$. **[3]**

(B) Find the interquartile range for this model. Hence show that a lifetime of more than 454 hours (to the nearest hour) would be classed as an outlier. **[6]**

(ii) In the alternative model, X has probability density function

$$\mathrm{f}(x) = \begin{cases} \frac{1}{200}\mathrm{e}^{-\frac{1}{200}x} & x > 0, \\ 0 & \text{elsewhere.} \end{cases}$$

(A) For this model show that the cumulative distribution function of X is

$$\mathrm{F}(x) = \begin{cases} 0 & x \leqslant 0, \\ 1 - \mathrm{e}^{-\frac{1}{200}x} & x > 0. \end{cases}$$ **[3]**

(B) Show that $\mathrm{P}(X > 50) = \mathrm{e}^{-0.25}$. **[2]**

(C) It is observed that a particular component is still working after 400 hours. Find the conditional probability that it will still be working after a further 50 hours (i.e. after a total of 450 hours) given that it is still working after 400 hours. **[4]**

4 The weights of Avonley Blue cheeses made by a small producer are found to be Normally distributed with mean 10 kg and standard deviation 0.4 kg.

(i) Find the probability that a randomly chosen cheese weighs less than 9.5 kg. **[3]**

One particular shop orders four Avonley Blue cheeses each week from the producer. From experience, the shopkeeper knows that the weekly demand from customers for Avonley Blue cheese is Normally distributed with mean 35 kg and standard deviation 3.5 kg. In the interests of food hygiene, no cheese is kept by the shopkeeper from one week to the next.

(ii) Find the probability that, in a randomly chosen week, demand from customers for Avonley Blue will exceed the supply. **[4]**

Following a campaign to promote Avonley Blue cheese, the shopkeeper finds that the weekly demand for it has increased by 30% (i.e. the mean and standard deviation are both increased by 30%). Therefore the shopkeeper increases his weekly order by one cheese.

(iii) Find the probability that, in a randomly chosen week, demand will now exceed supply. **[4]**

(iv) Following complaints, the cheese producer decides to check the mean weight of the Avonley Blue cheeses. For a random sample of 12 cheeses, she finds that the mean weight is 9.73 kg. Assuming that the population standard deviation of the weights is still 0.4 kg, find a 95% confidence interval for the true mean weight of the cheeses and comment on the result. Explain what is meant by a 95% confidence interval. **[7]**

PMT

THERE ARE NO QUESTIONS PRINTED ON THIS PAGE

OCR
RECOGNISING ACHIEVEMENT

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.

© OCR 2011