

Friday 20 January 2012 – Afternoon

A2 GCE MATHEMATICS (MEI)

4768 Statistics 3

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:

- Printed Answer Book 4768
- MEI Examination Formulae and Tables (MF2)

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.
- 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) Define simple random sampling. Describe briefly one difficulty associated with simple random sampling. [4]

(b) Freeze-drying is an economically important process used in the production of coffee. It improves the retention of the volatile aroma compounds. In order to maintain the quality of the coffee, technologists need to monitor the drying rate, measured in suitable units, at regular intervals. It is known that, for best results, the mean drying rate should be 70.3 units and anything substantially less than this would be detrimental to the coffee. Recently, a random sample of 12 observations of the drying rate was as follows.

66.0 66.1 59.8 64.0 70.9 71.4 66.9 76.2 65.2 67.9 69.2 68.5

(i) Carry out a test to investigate at the 5% level of significance whether the mean drying rate appears to be less than 70.3. State the distributional assumption that is required for this test. [10]

(ii) Find a 95% confidence interval for the true mean drying rate. [4]

2 In a particular chain of supermarkets, one brand of pasta shapes is sold in small packets and large packets. Small packets have a mean weight of 505 g and a standard deviation of 11 g. Large packets have a mean weight of 1005 g and a standard deviation of 17 g. It is assumed that the weights of packets are Normally distributed and are independent of each other.

(i) Find the probability that a randomly chosen large packet weighs between 995 g and 1020 g. [3]

(ii) Find the probability that the weights of two randomly chosen small packets differ by less than 25 g. [3]

(iii) Find the probability that the total weight of two randomly chosen small packets exceeds the weight of a randomly chosen large packet. [4]

(iv) Find the probability that the weight of one randomly chosen small packet exceeds half the weight of a randomly chosen large packet by at least 5 g. [4]

(v) A different brand of pasta shapes is sold in packets of which the weights are assumed to be Normally distributed with standard deviation 14 g. A random sample of 20 packets of this pasta is found to have a mean weight of 246 g. Find a 95% confidence interval for the population mean weight of these packets. [4]

- 3 (a) A medical researcher is looking into the delay, in years, between first and second myocardial infarctions (heart attacks). The following table shows the results for a random sample of 225 patients.

Delay (years)	0 –	1 –	2 –	3 –	4 – 10
Number of patients	160	40	13	9	3

The mean of this sample is used to construct a model which gives the following expected frequencies.

Delay (years)	0 –	1 –	2 –	3 –	4 – 10
Number of patients	142.23	52.32	19.25	7.08	4.12

Carry out a test, using a 2.5% level of significance, of the goodness of fit of the model to the data. [9]

- (b) A further piece of research compares the incidence of myocardial infarction in men aged 55 to 70 with that in women aged 55 to 70. Incidence is measured by the number of infarctions per 10000 of the population. For a random sample of 8 health authorities across the UK, the following results for the year 2010 were obtained.

Health authority	A	B	C	D	E	F	G	H
Incidence in men	47	56	15	51	45	54	50	32
Incidence in women	36	30	30	47	54	55	27	27

A Wilcoxon paired sample test, using the hypotheses $H_0: m = 0$ and $H_1: m \neq 0$ where m is the population median difference, is to be carried out to investigate whether there is any difference between men and women on the whole.

- (i) Explain why a paired test is being used in this context. [1]
- (ii) Carry out the test using a 10% level of significance. [8]

[Question 4 is printed overleaf.]

- 4 At the school summer fair, one of the games involves throwing darts at a circular dartboard of radius a lying on the ground some distance away. Only darts that land on the board are counted. The distance from the centre of the board to the point where a dart lands is modelled by the random variable R . It is assumed that the probability that a dart lands inside a circle of radius r is proportional to the area of the circle.

(i) By considering $P(R < r)$ show that $F(r)$, the cumulative distribution function of R , is given by

$$F(r) = \begin{cases} 0 & r < 0, \\ \frac{r^2}{a^2} & 0 \leq r \leq a, \\ 1 & r > a. \end{cases} \quad [3]$$

(ii) Find $f(r)$, the probability density function of R . [2]

(iii) Find $E(R)$ and show that $\text{Var}(R) = \frac{a^2}{18}$. [7]

The radius a of the dartboard is 22.5 cm.

(iv) Let \bar{R} denote the mean distance from the centre of the board of a random sample of 100 darts. Write down an approximation to the distribution of \bar{R} . [3]

(v) A random sample of 100 darts is found to give a mean distance of 13.87 cm. Does this cast any doubt on the modelling? [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) 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.