

OCR

Oxford Cambridge and RSA

Tuesday 9 June 2015 – Morning

A2 GCE MATHEMATICS (MEI)

4768/01 Statistics 3

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:

- Printed Answer Book 4768/01
- 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 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) A stratified sample of pupils at secondary schools in a particular local authority is to be chosen in order to collect information on absenteeism. In the local authority there are 4 secondary schools, A, B, C and D, with 1310, 1453, 843 and 1110 pupils respectively.
- (i) How many pupils should be chosen from each school in a stratified sample of 500 so that each school is represented proportionally? [3]
- (ii) Suggest two possible criteria for stratification other than by school. [2]
- (iii) State one advantage of choosing a stratified sample. [1]
- (b) At a large secondary school, the median number of half days absent per pupil per year (based on several years' records) was known to be 23. Last year the school carried out a drive to lower the number of absences. A random sample of 12 pupils had been absent for the following numbers of half days during the year.

14 10 15 13 35 9 24 19 30 26 29 8

A Wilcoxon single sample test is to be carried out to see if the drive has been successful.

- (i) Why might a Wilcoxon test be appropriate? [1]
- (ii) What distributional assumption is needed for the test? [1]
- (iii) Carry out the test, using a 5% level of significance. [10]
- 2 The distribution of the random variable X is thought to be well modelled by the following probability density function:

$$f(x) = \begin{cases} k(1+x) & \text{for } 0 \leq x < 5, \\ 0 & \text{elsewhere,} \end{cases}$$

where k is a positive constant.

- (i) Find the value of k . [3]
- (ii) Show that $P(a \leq X < a+1) = \frac{1}{35}(2a+3)$ for $0 \leq a \leq 4$. [2]

A random sample of 50 observations of X is summarised as follows.

x	$0 \leq x < 1$	$1 \leq x < 2$	$2 \leq x < 3$	$3 \leq x < 4$	$4 \leq x < 5$
Frequency	1	5	7	20	17

- (iii) Test at the 10% level of significance whether the distribution of X is well modelled by $f(x)$. [10]
- (iv) With reference to your calculations in part (iii) discuss briefly the outcome of the test. [2]

- 3 In agricultural research the oil content, as a percentage of the whole grain, of a cereal can be measured using near infra-red spectroscopy. An investigation into the effect of a particular treatment on the oil content of a certain cereal is being carried out. A sample of 10 plots of land is chosen and each plot is divided in half. In one half of each plot the cereal is grown with the treatment and in the other half the cereal is grown without the treatment. Subsequently the percentage oil content of the cereal for each half of each plot is measured and the results are as follows.

Plot	A	B	C	D	E	F	G	H	I	J
With treatment	41.1	44.3	42.4	48.2	52.4	54.6	35.9	33.6	51.1	47.0
Without treatment	42.5	37.7	42.1	32.4	42.7	41.5	36.9	31.7	52.6	41.2

A paired t test with a 5% level of significance is to be used to see if the treatment appears to make any difference to the mean percentage oil content of the cereal.

- (i) Explain what is meant by a 5% level of significance in a hypothesis test. [2]
 - (ii) State the conditions necessary for the test to be carried out. [3]
 - (iii) Assuming the conditions stated in part (ii) are met, carry out the test. [10]
 - (iv) Find a 90% confidence interval for the population mean difference in the percentage oil content with and without the treatment. [4]
- 4 Paul has been trying a new route to work in the mornings. He collects a large random sample of times, in minutes, and calculates a 95% confidence interval for the population mean time by this route. The confidence interval is (45.369, 47.231) and the sample variance is 20.3.

- (i) Explain what is meant by a 95% confidence interval for a population mean. [1]
- (ii) Calculate the sample mean and the sample size. [4]

Paul reverts to his usual route and the time, in minutes, to travel to work each morning is modelled by a random variable which is Normally distributed with mean 41.3 and variance 11.7. The time, in minutes, for Paul to travel home each evening is modelled by a random variable which is Normally distributed with mean 44.8 and variance 14.2. In the rest of this question all journeys are by Paul's usual route and may be assumed to be independent of each other.

- (iii) Calculate the probability that, on a randomly chosen day, Paul's total travelling time will be less than 90 minutes. [3]
- (iv) Calculate the probability that, on a randomly chosen day, the time for Paul to travel home will be more than 5 minutes longer than the time to travel to work. [4]
- (v) Calculate the probability that, in a randomly chosen five-day week, the mean time for Paul to travel to work on Monday and Tuesday will be more than 3 minutes longer than his mean time to travel to work on Wednesday, Thursday and Friday. [6]

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.