

OXFORD CAMBRIDGE AND RSA EXAMINATIONS
AS GCE
4766
MATHEMATICS (MEI)
Statistics 1
QUESTION PAPER

MONDAY 23 JANUARY 2012: Morning

DURATION: 1 hour 30 minutes

SUITABLE FOR VISUALLY IMPAIRED CANDIDATES

Candidates answer on the Printed Answer Book, or any suitable paper provided by the Centre. The Printed Answer Book may be enlarged by the Centre.

OCR SUPPLIED MATERIALS:

Printed Answer Book 4766

Loose sheet for question 7

MEI Examination Formulae and Tables (MF2)

OTHER MATERIALS REQUIRED:

Scientific or graphical calculator

READ INSTRUCTIONS OVERLEAF

INSTRUCTIONS TO CANDIDATES

- **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.**
- **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.**

INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- **Please contact OCR Copyright should you wish to re-use this document.**

SECTION A (36 marks)

- 1** The mean daily maximum temperatures at a research station over a 12-month period, measured to the nearest degree Celsius, are given below.

Jan	8
Feb	15
Mar	25
Apr	29
May	31
Jun	31
Jul	34
Aug	36
Sep	34
Oct	26
Nov	15
Dec	8

- (i) Construct a sorted stem and leaf diagram to represent these data, taking stem values of 0, 10, [4]
- (ii) Write down the median of these data. [1]
- (iii) The mean of these data is 24.3. Would the mean or the median be a better measure of central tendency of the data? Briefly explain your answer. [2]

2 The hourly wages, £ x , of a random sample of 60 employees working for a company are summarised as follows.

$$n = 60$$

$$\sum x = 759.00$$

$$\sum x^2 = 11\,736.59$$

- (i) Calculate the mean and standard deviation of x . [3]**
- (ii) The workers are offered a wage increase of 2%. Use your answers to part (i) to deduce the new mean and standard deviation of the hourly wages after this increase. [2]**
- (iii) As an alternative the workers are offered a wage increase of 25p per hour. Write down the new mean and standard deviation of the hourly wages after this 25p increase. [2]**

3 Jimmy and Alan are playing a tennis match against each other. The winner of the match is the first player to win three sets. Jimmy won the first set and Alan won the second set. For each of the remaining sets, the probability that Jimmy wins a set is

- 0.7 if he won the previous set,**
- 0.4 if Alan won the previous set.**

It is not possible to draw a set.

- (i) Draw a probability tree diagram to illustrate the possible outcomes for each of the remaining sets. [3]**
- (ii) Find the probability that Alan wins the match. [3]**
- (iii) Find the probability that the match ends after exactly four sets have been played. [2]**

- 4 In a food survey, a large number of people are asked whether they like tomato soup, mushroom soup, both or neither. One of these people is selected at random.
- T is the event that this person likes tomato soup.
 - M is the event that this person likes mushroom soup.

You are given that $P(T) = 0.55$, $P(M) = 0.33$ and $P(T | M) = 0.80$.

- (i) Use this information to show that the events T and M are not independent. [1]
- (ii) Find $P(T \cap M)$. [2]
- (iii) Draw a Venn diagram showing the events T and M , and fill in the probability corresponding to each of the four regions of your diagram. [3]

- 5 A couple plan to have at least one child of each sex, after which they will have no more children. However, if they have four children of one sex, they will have no more children. You should assume that each child is equally likely to be of either sex, and that the sexes of the children are independent. The random variable X represents the total number of girls the couple have.

- (i) Show that $P(X = 1) = \frac{11}{16}$. [3]

The table shows the probability distribution of X .

r	0	1	2	3	4
$P(X = r)$	$\frac{1}{16}$	$\frac{11}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

- (ii) Find $E(X)$ and $\text{Var}(X)$. [5]

SECTION B (36 marks)

- 6 It is known that 25% of students in a particular city are smokers. A random sample of 20 of the students is selected.**
- (i) (A) Find the probability that there are exactly 4 smokers in the sample. [3]**
- (B) Find the probability that there are at least 3 but no more than 6 smokers in the sample. [3]**
- (C) Write down the expected number of smokers in the sample. [1]**

A new health education programme is introduced. This programme aims to reduce the percentage of students in this city who are smokers. After the programme has been running for a year, it is decided to carry out a hypothesis test to assess the effectiveness of the programme. A random sample of 20 students is selected.

- (ii) (A) Write down suitable null and alternative hypotheses for the test. [3]**
- (B) Explain why the alternative hypothesis has the form that it does. [1]**
- (iii) Find the critical region for the test at the 5% level, showing all of your calculations. [4]**
- (iv) In fact there are 3 smokers in the sample. Complete the test, stating your conclusion clearly. [2]**

7 The birth weights of 200 lambs from crossbred sheep are illustrated by the cumulative frequency diagram on Fig. 7.1 on the loose sheet.

- (i) Estimate the percentage of lambs with birth weight over 6 kg. [2]**
- (ii) Estimate the median and interquartile range of the data. [3]**
- (iii) Use your answers to part (ii) to show that there are very few, if any, outliers. Comment briefly on whether any outliers should be disregarded in analysing these data. [4]**

The box and whisker plot on Fig. 7.2 on the loose sheet shows the birth weights of 100 lambs from Welsh Mountain sheep.

- (iv) Use appropriate measures to compare briefly the central tendencies and variations of the weights of the two types of lamb. [4]**
- (v) The weight of the largest Welsh Mountain lamb was originally recorded as 6.5 kg, but then corrected. If this error had not been corrected, how would this have affected your answers to part (iv)? Briefly explain your answer. [2]**
- (vi) One lamb of each type is selected at random. Estimate the probability that the birth weight of both lambs is at least 3.9 kg. [4]**