OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE
MATHEMATICS (MEI) 4766
Statistics 1

QUESTION PAPER

Candidates answer on the Printed Answer Book

OCR Supplied Materials:
- Printed Answer Book 4766
- MEI Examination Formulae and Tables (MF2)

Other Materials Required:
- Scientific or graphical calculator

Friday 18 June 2010
Afternoon

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- Write your name clearly in capital letters, your Centre Number and Candidate Number in the spaces provided on the Printed Answer Book.
- **The questions are on the inserted Question Paper.**
- **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. Pencil may be used for graphs and diagrams only.
- Read each question carefully and make sure that 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 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.

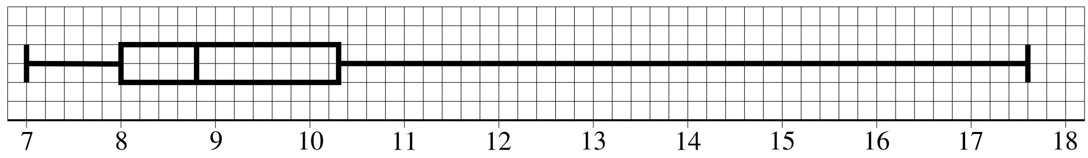
- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- You are advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
- The total number of marks 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 destroyed.

Section A (36 marks)

1 A business analyst collects data about the distribution of hourly wages, in £, of shop-floor workers at a factory. These data are illustrated in the box and whisker plot.



(i) Name the type of skewness of the distribution. **[1]**

(ii) Find the interquartile range and hence show that there are no outliers at the lower end of the distribution, but there is at least one outlier at the upper end. **[5]**

(iii) Suggest possible reasons why this may be the case. **[2]**

2 The probability distribution of the random variable X is given by the formula

$$\mathrm{P}(X = r) = kr(5 - r) \quad \text{for} \quad r = 1, 2, 3, 4.$$

(i) Show that $k = 0.05$. **[2]**

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

3 The lifetimes in hours of 90 components are summarised in the table.

Lifetime (x hours)	$0 < x \leqslant 20$	$20 < x \leqslant 30$	$30 < x \leqslant 50$	$50 < x \leqslant 65$	$65 < x \leqslant 100$
Frequency	24	13	14	21	18

(i) Draw a histogram to illustrate these data. **[5]**

(ii) In which class interval does the median lie? Justify your answer. **[2]**

4 Each packet of *Cruncho* cereal contains one free fridge magnet. There are five different types of fridge magnet to collect. They are distributed, with equal probability, randomly and independently in the packets. Keith is about to start collecting these fridge magnets.

(i) Find the probability that the first 2 packets that Keith buys contain the same type of fridge magnet. **[2]**

(ii) Find the probability that Keith collects all five types of fridge magnet by buying just 5 packets. **[3]**

(iii) Hence find the probability that Keith has to buy more than 5 packets to acquire a complete set. **[1]**

5 A retail analyst records the numbers of loaves of bread of a particular type bought by a sample of shoppers in a supermarket.

Number of loaves	0	1	2	3	4	5
Frequency	37	23	11	3	0	1

(i) Calculate the mean and standard deviation of the numbers of loaves bought per person. **[5]**

(ii) Each loaf costs £1.04. Calculate the mean and standard deviation of the amount spent on loaves per person. **[3]**

Section B (36 marks)

6 A manufacturer produces tiles. On average 10% of the tiles produced are faulty. Faulty tiles occur randomly and independently.

A random sample of 18 tiles is selected.

(i) (*A*) Find the probability that there are exactly 2 faulty tiles in the sample. **[3]**

(*B*) Find the probability that there are more than 2 faulty tiles in the sample. **[3]**

(*C*) Find the expected number of faulty tiles in the sample. **[2]**

A cheaper way of producing the tiles is introduced. The manufacturer believes that this may increase the proportion of faulty tiles. In order to check this, a random sample of 18 tiles produced using the cheaper process is selected and a hypothesis test is carried out.

(ii) (*A*) Write down suitable null and alternative hypotheses for the test.

(*B*) Explain why the alternative hypothesis has the form that it does. **[4]**

(iii) Find the critical region for the test at the 5% level, showing all of your calculations. **[4]**

(iv) In fact there are 4 faulty tiles in the sample. Complete the test, stating your conclusion clearly. **[2]**

7 One train leaves a station each hour. The train is either on time or late. If the train is on time, the probability that the next train is on time is 0.95. If the train is late, the probability that the next train is on time is 0.6. On a particular day, the 09 00 train is on time.

(i) Illustrate the possible outcomes for the 10 00, 11 00 and 12 00 trains on a probability tree diagram. **[4]**

(ii) Find the probability that

(*A*) all three of these trains are on time, **[2]**

(*B*) just one of these three trains is on time, **[4]**

(*C*) the 12 00 train is on time. **[4]**

(iii) Given that the 12 00 train is on time, find the probability that the 10 00 train is also on time. **[4]**

3 (i) **Write any calculations on page 5.**

OCR
RECOGNISING ACHIEVEMENT

GCE

Mathematics (MEI)

Advanced Subsidiary GCE **4766**

Statistics 1

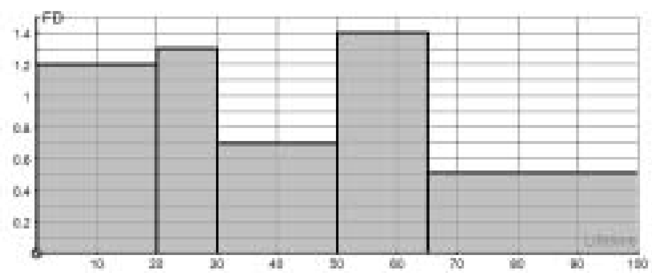
Mark Scheme for June 2010

Oxford Cambridge and RSA Examinations

Q1 (i)	Positive skewness	B1	**1**
(ii)	Inter-quartile range = 10.3 – 8.0 = 2.3 Lower limit 8.0 – 1.5 × 2.3 = 4.55 Upper limit 10.3 + 1.5 × 2.3 = 13.75 Lowest value is 7 so no outliers at lower end Highest value is 17.6 so at least one outlier at upper end.	B1 M1 for 8.0 – 1.5 × 2.3 M1 for 10.3 + 1.5 × 2.3 A1 A1	**5**
(iii)	Any suitable answers Eg minimum wage means no very low values Highest wage earner may be a supervisor or manager or specialist worker or more highly trained worker	E1 one comment relating to low earners E1 one comment relating to high earners	**2**
		TOTAL	**8**
Q2 (i)	$4k + 6k + 6k + 4k = 1$ $20k = 1$ $k = 0.05$	M1 A1 **NB Answer given**	**2**
(ii)	$E(X) = 1\times0.2 + 2\times0.3 + 3\times0.3 + 4\times0.2 = 2.5$ (or by inspection) $E(X^2) = 1\times0.2 + 4\times0.3 + 9\times0.3 + 16\times0.2 = 7.3$ $\text{Var}(X) = 7.3 - 2.5^2 = 1.05$	M1 for Σrp (at least 3 terms correct) A1 CAO M1 for Σr^2p (at least 3 terms correct) M1dep for – their $E(X)^2$ A1 FT their E(X) provided $\text{Var}(X) > 0$	**5**
		TOTAL	**7**

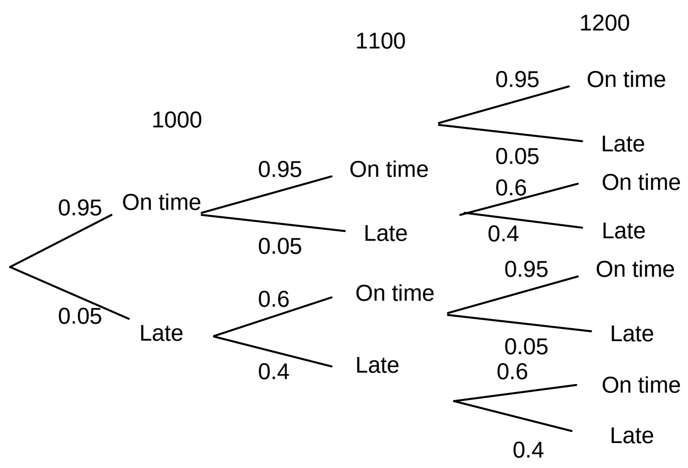
Q3 (i)

Lifetime (x hours)	Frequency	Width	FD
$0 < x \le 20$	24	20	1.2
$20 < x \le 30$	13	10	1.3
$30 < x \le 50$	14	20	0.7
$50 < x \le 65$	21	15	1.4
$65 < x \le 100$	18	35	0.51



M1 for fds A1 CAO Accept any suitable unit for fd such as eg freq per 10 hours. L1 linear scales on both axes and label on vert axis W1 width of bars H1 height of bars	**5**

(ii)	Median lies in third class interval ($30 < x \leq 50$) Median = 45.5th lifetime (which lies beyond 37 but not as far as 51)	B1 CAO E1 *dep* on B1	**2**
		TOTAL	**7**
Q4 (i)	$1 \times \frac{1}{5} = \frac{1}{5}$	M1 A1	**2**
(ii)	$1 \times \frac{4}{5} \times \frac{3}{5} \times \frac{2}{5} \times \frac{1}{5} = \frac{24}{625} = 0.0384$	M1 For $1 \times \frac{4}{5} \times$ *or just* $\frac{4}{5} \times$ M1 *dep* for fully correct product A1	**3**
(iii)	1 – 0.0384 = 0.9616 or 601/625	B1	**1**
		TOTAL	**6**
Q5 (i)	Mean = $\frac{0 \times 37 + 1 \times 23 + 2 \times 11 + 3 \times 3 + 4 \times 0 + 5 \times 1}{75} = \frac{59}{75} = 0.787$ S_{xx} = $0^2 \times 37 + 1^2 \times 23 + 2^2 \times 11 + 3^2 \times 3 + 4^2 \times 0 + 5^2 \times 1 - \frac{59^2}{75} = 72.59$ s = $\sqrt{\frac{72.59}{74}} = 0.99$	M1 A1 M1 for Σfx^2 s.o.i. M1 *dep* for good attempt at S_{xx} BUT NOTE M1M0 if their $S_{xx} < 0$ A1 CAO	**5**
(ii)	New mean = 0.787 × £1.04 = £0.818 or 81.8 pence New s = 0.99 × £1.04 = £1.03 or 103 pence	B1 ft their mean B1 ft their s B1 for correct units *dep* on at least 1 correct (ft)	**3**
		TOTAL	**8**
	Section B		
Q6 (i)	X ~ B(18, 0.1) (***A***) P(2 faulty tiles) $= \binom{18}{2} \times 0.1^2 \times 0.9^{16} = 0.2835$ OR from tables $0.7338 - 0.4503 = 0.2835$ (***B***) P(More than 2 faulty tiles) $= 1 - 0.7338 = 0.2662$	M1 $0.1^2 \times 0.9^{16}$ M1 $\binom{18}{2} \times p^2 q^{16}$ A1 CAO OR: M2 for 0.7338 – 0.4503 A1 CAO M1 P(X≤2) M1 *dep* for 1-P(X≤2) A1 CAO	**3** **3**

	(*C*) E(X) = $np = 18 \times 0.1 = 1.8$	M1 for product 18×0.1 A1 CAO	**2**
(ii)	(*A*) Let p = probability that a randomly selected tile is faulty H_0: $p = 0.1$ H_1: $p > 0.1$	B1 for definition of p in context B1 for H_0 B1 for H_1	**3**
	(*B*) H_1 has this form as the manufacturer believes that the number of faulty tiles may <u>increase</u>.	E1	**1**
(iii)	Let $X \sim B(18, 0.1)$ $P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.9018 = 0.0982 > 5\%$ $P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.9718 = 0.0282 < 5\%$ So critical region is {5,6,7,8,9,10,11,12,13,14,15,16,17,18}	B1 for 0.0982 B1 for 0.0282 M1 for at least one comparison with 5% A1 CAO for critical region *dep* on M1 and at least one B1	**4**
(iv)	4 does not lie in the critical region, (so there is insufficient evidence to reject the null hypothesis and we conclude that there is not enough evidence to suggest that the number of faulty tiles has increased.	M1 for comparison A1 for conclusion **in context**	**2**
		TOTAL	**18**
Q7 **(i)**		G1 first set of branches G1 *indep* second set of branches G1 *indep* third set of branches G1 labels	**4**

(ii)	(A) P(all on time) = $0.95^3 = 0.8574$	M1 for 0.95^3 A1 CAO	**2**
	(B) P(just one on time) = $0.95\times0.05\times0.4 + 0.05\times0.6\times0.05 + 0.05\times0.4\times0.6$ $= 0.019 + 0.0015 + 0.012 = 0.0325$	M1 first term M1 second term M1 third term A1 CAO	**4**
	(C) P(1200 is on time) = $0.95\times0.95\times0.95 + 0.95\times0.05\times0.6 + 0.05\times0.6\times0.95 + 0.05\times0.4\times0.6 = 0.857375+0.0285+0.0285+0.012 = 0.926375$	M1 any two terms M1 third term M1 fourth term A1 CAO	**4**
(iii)	P(1000 on time given 1200 on time) = P(1000 on time and 1200 on time) / P(1200 on time) = $\dfrac{0.95\times0.95\times0.95+0.95\times0.05\times0.6}{0.926375} = \dfrac{0.885875}{0.926375} = 0.9563$	M1 either term of numerator M1 full numerator M1 denominator A1 CAO	**4**
		Total	**18**