OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G241

Statistics 1 (Z1)

QUESTION PAPER

Candidates answer on the Printed Answer Book

OCR Supplied Materials:
- Printed Answer Book G241
- 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

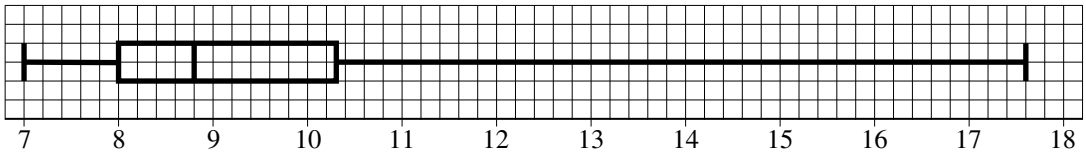
- Do not send this Question Paper for marking; it should be retained in the centre or destroyed.

© OCR 2010 [A/100/3618]

OCR is an exempt Charity

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

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

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

(ii) Find $E(X)$ and $Var(X)$. **[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]**

© OCR 2010

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]**

© OCR 2010

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, is given to all schools that receive assessment material 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 2010

OCR
RECOGNISING ACHIEVEMENT

ADVANCED SUBSIDIARY GCE

MEI STATISTICS

G241

Statistics 1 (Z1)

PRINTED ANSWER BOOK

Candidates answer on this Printed Answer Book

OCR Supplied Materials:
- Question Paper G241 (inserted)
- MEI Examination Formulae and Tables (MF2)

Other Materials Required:
- Scientific or graphical calculator

Friday 18 June 2010
Afternoon

Duration: 1 hour 30 minutes

* G 2 4 1 *

Candidate Forename		Candidate Surname	

Centre Number						Candidate Number				

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.

© OCR 2010 [A/100/3618]
OCR is an exempt Charity

Section A (36 marks)

1 (i)	
1 (ii)	
1 (iii)	

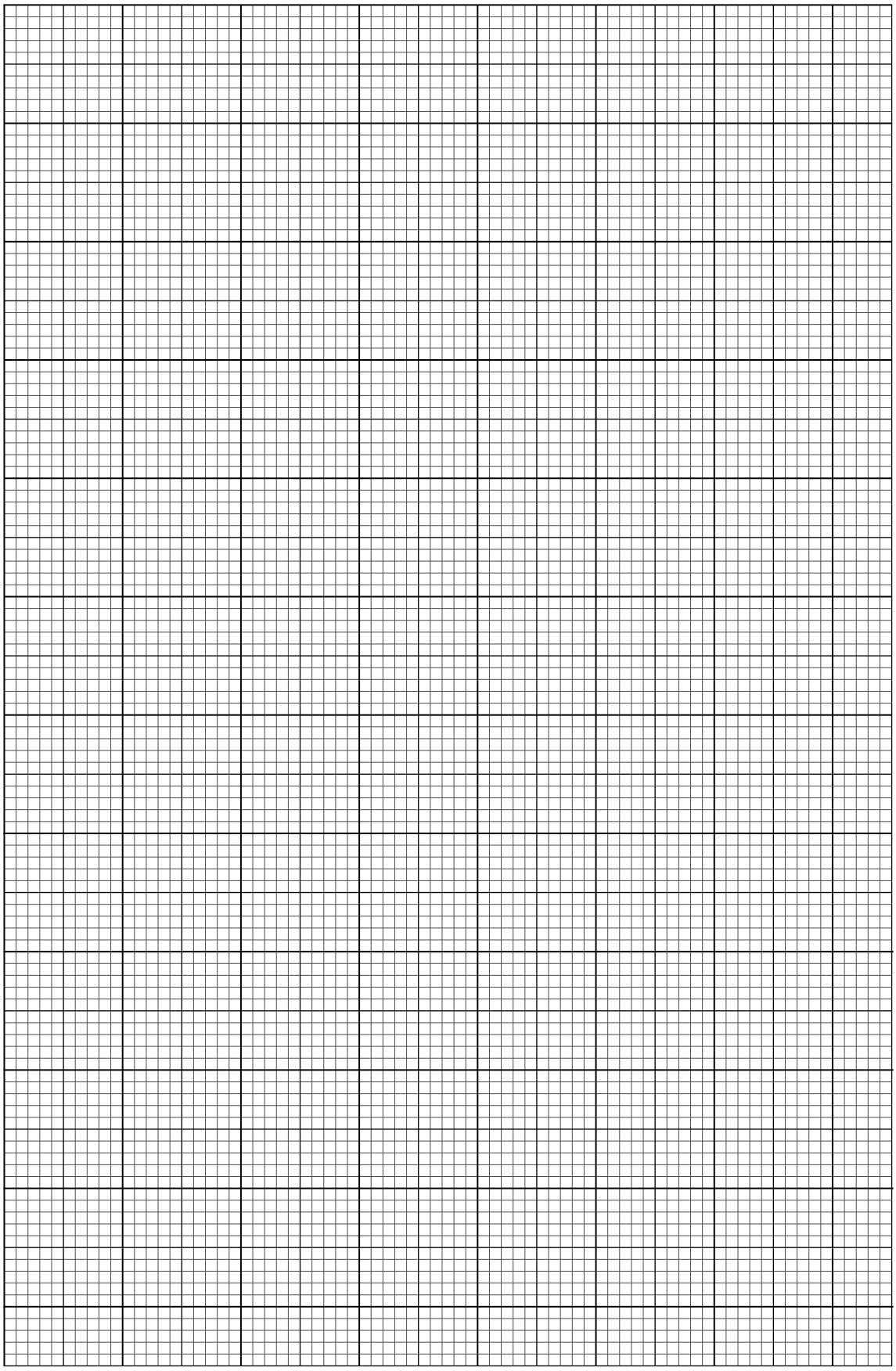
© OCR 2010

2 (i)

2 (ii)

© OCR 2010

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



© OCR 2010

3 (i)	**Write your calculations for 3 (i) here.**
3 (ii)	

© OCR 2010

4 (i)	
4 (ii)	
4 (iii)	

© OCR 2010

5 (i)	
5 (ii)	

© OCR 2010

Section B (36 marks)

6 (i) (*A*)	
6 (i) (*B*)	
6 (i) (*C*)	

© OCR 2010

6 (ii) (*A*)	
6 (ii) (*B*)	
6 (iii)	
6 (iv)	

© OCR 2010

7 (i)	

© OCR 2010

7 (ii) (*A*)

7 (ii) (*B*)

7 (ii) (*C*)

© OCR 2010

7 (iii)	

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, is given to all schools that receive assessment material 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 2010

OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE **G241**

Statistics 1 (Z1)

Mark Scheme for June 2010

Oxford Cambridge and RSA Examinations

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of pupils of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, OCR Nationals, Functional Skills, Key Skills, Entry Level qualifications, NVQs and vocational qualifications in areas such as IT, business, languages, teaching/training, administration and secretarial skills.

It is also responsible for developing new specifications to meet national requirements and the needs of students and teachers. OCR is a not-for-profit organisation; any surplus made is invested back into the establishment to help towards the development of qualifications and support which keep pace with the changing needs of today's society.

This mark scheme is published as an aid to teachers and students, to indicate the requirements of the examination. It shows the basis on which marks were awarded by Examiners. It does not indicate the details of the discussions which took place at an Examiners' meeting before marking commenced.

All Examiners are instructed that alternative correct answers and unexpected approaches in candidates' scripts must be given marks that fairly reflect the relevant knowledge and skills demonstrated.

Mark schemes should be read in conjunction with the published question papers and the Report on the Examination.

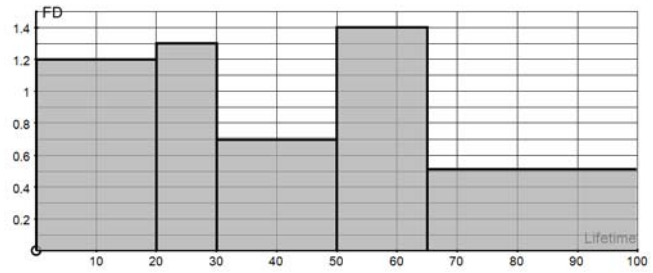
OCR will not enter into any discussion or correspondence in connection with this mark scheme.

© OCR 2010

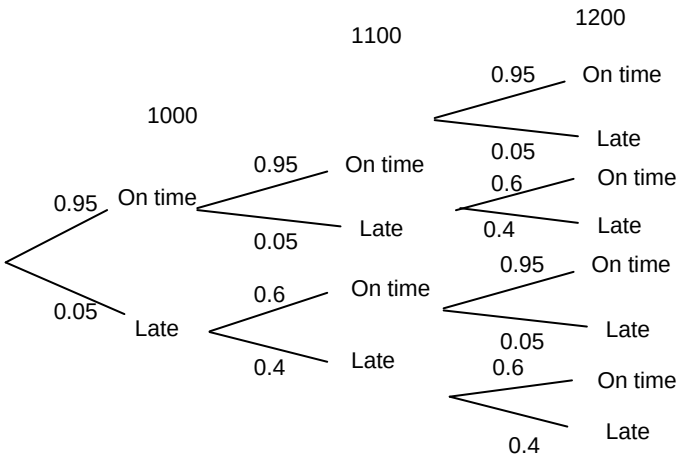
Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

Telephone: 0870 770 6622
Facsimile: 01223 552610
E-mail: publications@ocr.org.uk

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**
(iií)	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) $E(X^2) = 1\times0.2+4\times0.3+9\times0.3+16\times0.2 = 7.3$ $Var(X) = 7.3 – 2.5^2 = 1.05$	M1 for Σrp (at least 3 terms correct) A1 CAO M1 for $\Sigma r^2 p$ (at least 3 terms correct) M1dep for – their E(X)² A1 FT their E(X) provided Var(X) > 0	**5**
		TOTAL	**7**
Q3 (i)	<table><tr><th>Lifetime (x hours)</th><th>Frequency</th><th>Width</th><th>FD</th></tr><tr><td>0 < x ≤ 20</td><td>24</td><td>20</td><td>1.2</td></tr><tr><td>20 < x ≤ 30</td><td>13</td><td>10</td><td>1.3</td></tr><tr><td>30 < x ≤ 50</td><td>14</td><td>20</td><td>0.7</td></tr><tr><td>50 < x ≤ 65</td><td>21</td><td>15</td><td>1.4</td></tr><tr><td>65 < x ≤ 100</td><td>18</td><td>35</td><td>0.51</td></tr></table> 	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($X \leq 2$) M1 *dep* for 1-P($X \leq 2$) A1 CAO	**3** **3**

	(*C*) E(X) = np = 18 × 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) = $\frac{0.95\times0.95\times0.95 + 0.95\times0.05\times0.6}{0.926375} = \frac{0.885875}{0.926375} = 0.9563$	M1 either term of numerator M1 full numerator M1 denominator A1 CAO	**4**
		Total	**18**

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

14 – 19 Qualifications (General)
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored

Oxford Cambridge and RSA Examinations
is a Company Limited by Guarantee
Registered in England
Registered Office; 1 Hills Road, Cambridge, CB1 2EU
Registered Company Number: 3484466
OCR is an exempt Charity

OCR (Oxford Cambridge and RSA Examinations)
Head office
Telephone: 01223 552552
Facsimile: 01223 552553

© OCR 2010

Chief Examiner's Report

The Principal Examiners' reports that follow discuss the candidates' performances on the individual modules. There is one matter that should be discussed in a general way as it applies to all the statistics modules. This is in respect of arithmetical accuracy in intermediate working and in quotation of final answers.

Most candidates are sensible in their arithmetical work, but there is some unease as to exactly what level of accuracy the examiners are expecting. There is no general answer to this! The standard rubric for all the papers sums the situation up by including "final answers should be given to a degree of accuracy appropriate to the context". Three significant figures may often be the norm for this, but this always needs to be considered in the context of the problem in hand. For example, in quoting from Normal tables, *some* evidence of interpolation is generally expected and so quotation to four decimal places will often be appropriate. But even this does not always apply – quotations of the standard critical points for significance tests such as 1.96, 1.645, 2.576 (maybe even 2.58 – but not 2.57) will commonly suffice.

Talking now in general terms, the examiners always exercise sensible discretion in cases of small variations in the degree of accuracy to which an answer is given. For example, if 3 significant figures are expected (either because of an explicit instruction or because the general context of a problem demands it) but only 2 are given, a candidate is likely to lose an Accuracy mark; but if 4 significant figures are given, there would normally be no penalty. Likewise, answers which are slightly deviant from what is expected in a very minor manner are not penalised (for example, a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected). However, there are increasing numbers of cases where candidates give answers which are *grossly* over- or under-specified, such as insistence that the value of a test statistic is (say) 2.128888446667 merely because that is the value that happens to come off the candidate's calculator. **Such gross over-specification indicates a lack of appreciation of the nature of statistical work and, with effect from the January 2011 examinations, will be penalised by withholding of associated Accuracy marks.**

Candidates must however always be aware of the dangers of premature rounding if there are several steps in a calculation. If, say, a final answer is desired that is correct to 3 decimal places, this can in no way be guaranteed if only 3 decimal places are used in intermediate steps; indeed, it may not be safe to carry out the intermediate work even to 4 decimal places. The issue of over-specification may arise for the final answer but not for intermediate stages of the working.

It is worth repeating that most candidates act sensibly in all these respects, but it is hoped that this note may help those who are perhaps a little less confident in how to proceed.

G241 Statistics 1

General comments

The level of difficulty of the paper appeared to be entirely appropriate for the candidates with a good range of marks obtained. Question 4 proved to be the most challenging question on the paper and question 7 the easiest. Very low scores were rare and very few candidates seemed totally unprepared. There were, on the other hand, a good number of almost completely or completely correct scripts. There seemed to be no trouble in completing the paper within the time allowed.

Most candidates supported their numerical answers with appropriate explanations and working although some rounding errors were noted particularly in questions 5 and 6. Arithmetic accuracy was generally good although there is still evidence of candidates not being proficient or sensible in their use of calculators. In particular the simplest method of doing question 5(i) is by use of the statistical functions on a calculator, but few candidates used this approach. Amongst some candidates, there was evidence of incorrect use of point probabilities instead of tail probabilities in question 6 and of a totally wrong method to establish outliers in question 1.

The scripts were invariably well presented and legible with the use of a pre-printed answer book not appearing to constrict candidates' work; most candidates were able to answer in the space provided in the answer book, and only a few used additional sheets.

Comments on individual questions

1) (i) Most candidates gave the correct answer of positive skewness although a few thought that the skewness was negative; the occasional response of ‘skewed to the right’ was not acceptable.

1) (ii) The answer of 2.3 for the IQR was obtained by most candidates. Wrong answers included (17.6 – 7)/4 = 2.65 and errors in stating the value of the upper quartile. Many candidates made mistakes in finding the boundaries for outliers with the use of median ± 1.5 × IQR being very common. Those who used the quartiles occasionally combined the values with multiples of 1, 2 or even 3 of the IQR. The use of the limits to establish the presence of outliers, or otherwise, was good although a number of candidates used a value of 18 rather than 17.6. This error was treated generously. Some candidates curiously tried to argue in terms of the standard deviation.

1) (iii) There were many sensible and complete answers with the most common including ‘an error in the data’, ‘no lower outlier due to the minimum wage’ and ‘the outlier being a manager or supervisor’. Some candidates only gave one reason or just concentrated on one end of the data. A very few candidates just repeated the information about outliers given in part (ii).

2) (i) This was very often correct but a number of candidates stopped when they had worked out the first two terms. Some candidates tried to sum the terms without k or the k became an afterthought after the summation was completed.

2) (ii) The calculation of E(X) and Var(X) was well executed on the whole. There were still some candidates who mistakenly divided E(X) and/or Var(X) by 4. Some forgot to square E(X) in the calculation of the variance. A few candidates thought that they could attempt this question without using any probabilities.

3) (i) Most candidates knew that they had to find frequency density and on the whole were very successful. Occasionally-seen errors were attempts to multiply frequency by width or divide by mid-interval or divide width by frequency. However the most common error was to use the given frequencies as the heights. Labelling was not always successful, and although a pleasing number of candidates knew that the label should be frequency density, some gave it simply as frequency, and those using a non-unitary class width as standard often had difficulty indicating this correctly on the graph. The vertical linear scale was usually correct (and sensible!). On the horizontal scale the majority of candidates were able to get the width of the bars correct, but a number of candidates thought that they should number their scale with inequalities rather than giving a correct linear scale. Very few candidates mistakenly left gaps between the bars. Use of rulers is to be encouraged to produce a clearer diagram.

3) (ii) This was usually well done although some candidates seemed to think the frequency was 100. Although candidates should have been trying to find the 45.5th value many were looking for the 45th value; this error was not penalised. Many candidates failed to indicate that it was the 45th value not just 45 that was in the correct interval.

4) (i) This was the least well done question on the whole paper. The majority of candidates had 0.2 × 0.2 in some form as their answer.

4) (ii) This part was done slightly better, but even so the correct answer of 0.0384 was fairly unusual. Common errors included $(^1/_5)^5$ and $^1/_{5!}$ whilst some candidates tried to use some form of binomial probability.

4) (iii) Most candidates managed to subtract their answer to part (ii) from 1 although some made arithmetical errors whilst others did not attempt this part at all.

5) (i) This question was answered more successfully than in the past. There were many wholly correct solutions, usually showing full working but occasionally by use of calculator. The vast majority of candidates found the mean correctly, although a number of incorrect answers were seen including $^{75}/_{59}$, $^{59}/_6$ or $^{59}/_5$. Some lost a mark because of inappropriate rounding of their answer. Many candidates found the standard deviation correctly but there was a wide variety of wrong methods including finding $(fx)^2$ or xf^2 instead of fx^2. A few candidates correctly found fx^2 but then forgot to subtract $(\sum x)^2/n$ or used 59, 58 or 6 rather than 75 as the value of n. Only a few candidates divided by 75 and thus found the root mean square deviation and only a few forgot to square root their variance. Candidates who obtained ridiculously large answers often did not seem to realise that their answers could not possibly be correct.

5) (ii) Most candidates found the new mean successfully. However many stated that the standard deviation would not change. Units were often missing or only given for the mean. A number of candidates gave the new mean as 0.82p rather than 82 pence or £0.82. Some candidates did not realise that they could just multiply their answers to part (i) by 1.04 and instead multiplied the numbers of loaves by 1.04 then recalculated the new mean and standard deviation.

6) (i)(A) The vast majority of candidates found the correct value of 0.2385, with most preferring to use a binomial expression rather than tables. Occasionally an answer of $P(X \leq 2) = 0.7338$ was seen.

6) (i)(B) Candidates were less successful in this part, with mistakes occurring due to rounding errors when using the point probability approach, the omission of a term such as P(X = 0) when using point probabilities, misuse of tables, or answers such as 1 – 0.2835, 1 – 0.9018 or 1 – 0.4503 rather than 1 – 0.7338.

6) (i)(C) This was very well answered although a significant number of candidates rounded to 2 or even 1, losing a mark.

6 (ii) Many candidates constructed the hypotheses correctly although a few used "equals" for H_1. The main loss of marks came from poor notation such as P(X) = 0.1, H_0 = 0.1, X=0.1, P(0.1), etc. However many candidates still failed to define p as the probability that a randomly selected tile is faulty. Virtually all adequately explained why H_1 took the form it did.

6) (iii) Few candidates confidently scored full marks in this part. Some candidates often had little idea as to where to begin; other candidates used point probabilities and even those who used the correct probabilities of 0.0982 and 0.0282 with a comparison of 0.05 often started the critical region at 4 rather than 5. Occasionally the critical region was given as 'from 5 to 8' rather than 'from 5 to 18'. Some candidates failed to show necessary working; an answer along the lines of 'the first value in tables above 0.95 is 4, so critical region is 5 to 18' did not score full marks. A very small number of candidates thought their comparison should be with 0.025 rather than 0.05. As has been stressed in past examiners' reports, candidates must quote specific probabilities in finding critical regions and then explicitly compare these probabilities to the significance level. If they do not do this they may not get any marks.

Although it is given in the mark scheme, it is worth repeating here the recommended method for comparing the probabilities with the significance level. Candidates should find the two upper tail (in this case) cumulative probabilities which straddle the significance level.

P($X \geq 4$) = 1 – P($X \leq 3$) = 1 – 0.9018 or 0.0982 > 5%

P($X \geq 5$) = 1 – P($X \leq 4$) = 1 – 0.9718 or 0.0282 < 5%

6) (iv) This was poorly answered with very few using their critical region and stating that 4 was outside it. Most successful answers started again with 0.0982>0.05; often those using this approach wrote down 0.0282<0.05 and wrongly rejected H_0. The other major error was to finish by saying 'insufficient evidence to reject H_0' and then making no reference to the context. The use of point probabilities was again frequent, even by some candidates who had successfully used cumulative probabilities in part (iii).

7) (i) This question provoked a variety of responses. Those who read the question carefully often gained high marks but there were many candidates who got off on the wrong foot by either having too many sets of branches on their tree diagram e.g. (0900, 1000, 1100, 1200) or having the 1000 branches labelled with probabilities of 0.5 instead of the correct 0.95 and 0.05. Some candidates omitted some or all of the labels for 'on time' and 'late'.

7) (ii)(A) Whether using their tree diagram or not, this was well answered and most candidates gained both of the marks.

7) (ii)(B) Most candidates were able to trace their way through the tree diagram to achieve the correct response of 0.0325. A generous follow through was in place for those who may have made an error in one or more of their probabilities.

7) (ii)(C) This was again well answered by most candidates, and once again a generous follow through was in place. Candidates should be reminded that total accuracy in intermediate working is important when dealing with probabilities e.g. 0.857375 should not be rounded to 0.86. The cumulative effect of 4 prematurely rounded probabilities caused some candidates to have their final answer outside the required range.

7) (iii) Many candidates realised they had to evaluate the 1000, 1100 and 1200 on time and the 1000 on time, 1100 late and 1200 on time to reach 0.885875 or its equivalent on follow through. Often they stopped at this point, not realising the conditional probability requirement of the question. The more discerning candidates used their answer from part (ii) (C) to complete the question successfully.