OCR
RECOGNISING ACHIEVEMENT

Thursday 24 May 2012 – Morning

AS 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

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.
- 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.

© OCR 2012 [H/102/2650]
DC (RW) 58420

OCR is an exempt Charity

Section A (36 marks)

1 At a garden centre there is a box containing 50 hyacinth bulbs. Of these, 30 will produce a blue flower and the remaining 20 will produce a red flower. Unfortunately they have become mixed together so that it is not known which of the bulbs will produce a blue flower and which will produce a red flower.

Karen buys 3 of these bulbs.

(i) Find the probability that all 3 of these bulbs will produce blue flowers. **[3]**

(ii) Find the probability that Karen will have at least one flower of each colour from her 3 bulbs. **[3]**

2 An examination paper consists of two sections. Section A has 5 questions and Section B has 9 questions. Candidates are required to answer 6 questions.

(i) In how many different ways can a candidate choose 6 questions, if 3 are from Section A and 3 are from Section B? **[3]**

(ii) Another candidate randomly chooses 6 questions to answer. Find the probability that this candidate chooses 3 questions from each section. **[3]**

3 At a call centre, 85% of callers are put on hold before being connected to an operator. A random sample of 30 callers is selected.

(i) Find the probability that exactly 29 of these callers are put on hold. **[3]**

(ii) Find the probability that at least 29 of these callers are put on hold. **[3]**

(iii) If 10 random samples, each of 30 callers, are selected, find the expected number of samples in which at least 29 callers are put on hold. **[2]**

4 It is known that 8% of the population of a large city use a particular web browser. A researcher wishes to interview some people from the city who use this browser. He selects people at random, one at a time.

(i) Find the probability that the first person that he finds who uses this browser is

(*A*) the third person selected, **[3]**

(*B*) the second or third person selected. **[2]**

(ii) Find the probability that at least one of the first 20 people selected uses this browser. **[3]**

5 A manufacturer produces titanium bicycle frames. The bicycle frames are tested before use and on average 5% of them are found to be faulty. A cheaper manufacturing process is introduced and the manufacturer wishes to check whether the proportion of faulty bicycle frames has increased. A random sample of 18 bicycle frames is selected and it is found that 4 of them are faulty. Carry out a hypothesis test at the 5% significance level to investigate whether the proportion of faulty bicycle frames has increased. **[8]**

© OCR 2012

Section B (36 marks)

6 The engine sizes $x\,\text{cm}^3$ of a sample of 80 cars are summarised in the table below.

Engine size x	$500 \leqslant x \leqslant 1000$	$1000 < x \leqslant 1500$	$1500 < x \leqslant 2000$	$2000 < x \leqslant 3000$	$3000 < x \leqslant 5000$
Frequency	7	22	26	18	7

(i) Draw a histogram to illustrate the distribution. **[5]**

(ii) A student claims that the midrange is $2750\,\text{cm}^3$. Discuss briefly whether he is likely to be correct. **[1]**

(iii) Calculate estimates of the mean and standard deviation of the engine sizes. Explain why your answers are only estimates. **[5]**

(iv) Hence investigate whether there are any outliers in the sample. **[3]**

(v) A vehicle duty of £1000 is proposed for all new cars with engine size greater than $2000\,\text{cm}^3$. Assuming that this sample of cars is representative of all new cars in Britain and that there are 2.5 million new cars registered in Britain each year, calculate an estimate of the total amount of money that this vehicle duty would raise in one year. **[3]**

(vi) Why in practice might your estimate in part **(v)** turn out to be too high? **[1]**

7 Yasmin has 5 coins. One of these coins is biased with P(heads) = 0.6. The other 4 coins are fair. She tosses all 5 coins once and records the number of heads, X.

(i) Show that $P(X = 0) = 0.025$. **[2]**

(ii) Show that $P(X = 1) = 0.1375$. **[4]**

The table shows the probability distribution of X.

r	0	1	2	3	4	5
$P(X=r)$	0.025	0.1375	0.3	0.325	0.175	0.0375

(iii) Draw a vertical line chart to illustrate the probability distribution. **[2]**

(iv) Comment on the skewness of the distribution. **[1]**

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

(vi) Yasmin tosses the 5 coins three times. Find the probability that the total number of heads is 3. **[4]**

© OCR 2012

THERE ARE NO QUESTIONS WRITTEN 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 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 2012

OCR
RECOGNISING ACHIEVEMENT

Thursday 24 May 2012 – Morning

AS 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

Duration: 1 hour 30 minutes

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.

- 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.
- 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.

© OCR 2012 [H/102/2650]
DC (RW/CGW) 58421
OCR is an exempt Charity

Section A (36 marks)

1 (i)	
1 (ii)	

© OCR 2012

2 (i)

2 (ii)

3 (i)

© OCR 2012

3 (ii)

3 (iii)

4 (i) (*A*)

© OCR 2012

4 (i) (*B*)	
4 (ii)	

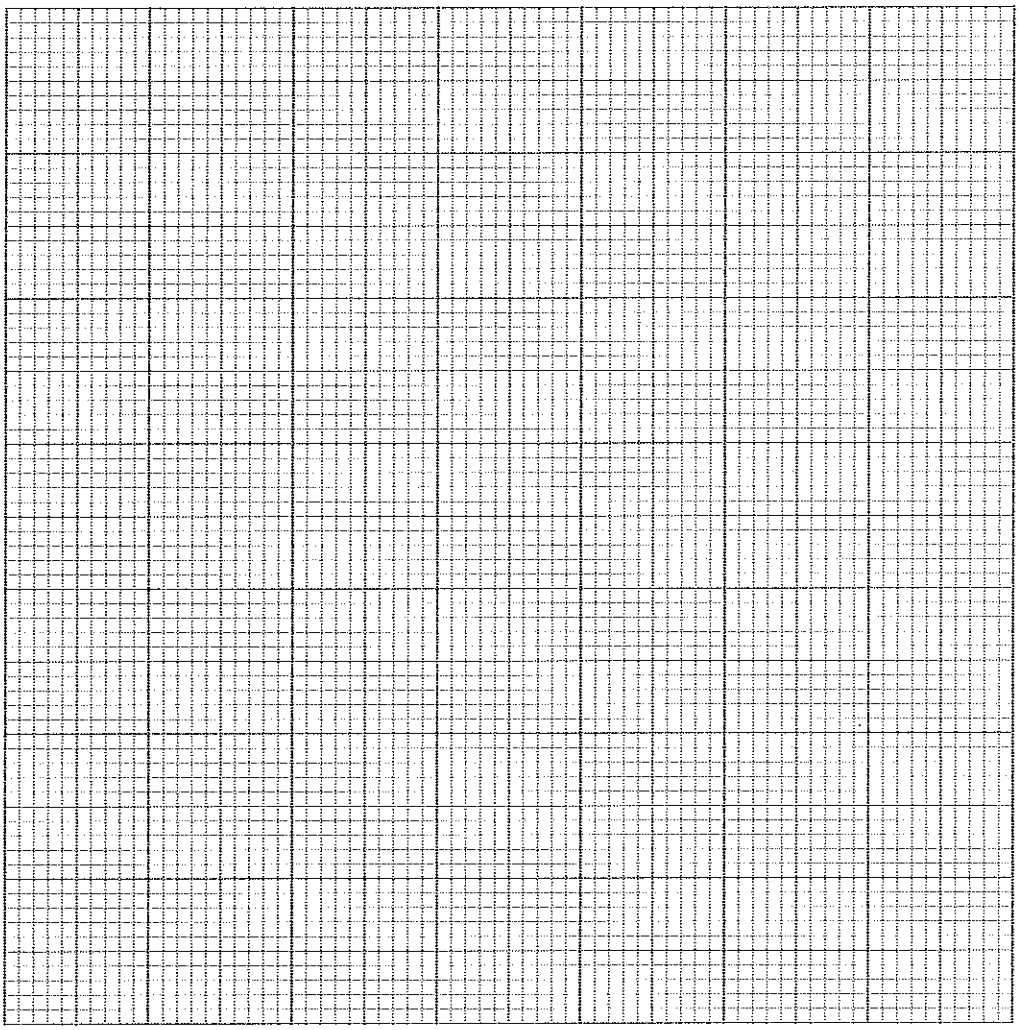
© OCR 2012

5

© OCR 2012

Section B (36 marks)

6 (i)



© OCR 2012

6 (ii)	
6 (iii)	

© OCR 2012

6 (iv)	
6 (v)	
6 (vi)	

© OCR 2012

7 (i)	
7 (ii)	
7 (iii)	

© OCR 2012

7 (iv)	
7 (v)	

© OCR 2012

7 (vi)	

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 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 2012

OCR
RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE

Unit **G241:** Statistics 1 (Z1)

Mark Scheme for June 2012

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 candidates 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 2012

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

Annotations

Annotation in scoris	Meaning
✓and ✖	
BOD	Benefit of doubt
FT	Follow through
ISW	Ignore subsequent working
M0, M1	Method mark awarded 0, 1
A0, A1	Accuracy mark awarded 0, 1
B0, B1	Independent mark awarded 0, 1
SC	Special case
^	Omission sign
MR SC SC this year as MR symbol missing	Misread
Highlighting	
Other abbreviations in mark scheme	**Meaning**
E1	Mark for explaining
U1	Mark for correct units
G1	Mark for a correct feature on a graph
M1 dep*	Method mark dependent on a previous mark, indicated by *
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working
awrt	Answer which rounds to

Subject-specific Marking Instructions

a. Annotations should be used whenever appropriate during your marking.

The A, M and B annotations must be used on your standardisation scripts for responses that are not awarded either 0 or full marks. It is vital that you annotate standardisation scripts fully to show how the marks have been awarded.

For subsequent marking you must make it clear how you have arrived at the mark you have awarded.

b. An element of professional judgement is required in the marking of any written paper. Remember that the mark scheme is designed to assist in marking incorrect solutions. Correct *solutions* leading to correct answers are awarded full marks but work must not be judged on the answer alone, and answers that are given in the question, especially, must be validly obtained; key steps in the working must always be looked at and anything unfamiliar must be investigated thoroughly.

Correct but unfamiliar or unexpected methods are often signalled by a correct result following an *apparently* incorrect method. Such work must be carefully assessed. When a candidate adopts a method which does not correspond to the mark scheme, award marks according to the spirit of the basic scheme; if you are in any doubt whatsoever (especially if several marks or candidates are involved) you should contact your Team Leader.

c. The following types of marks are available.

M
A suitable method has been selected and *applied* in a manner which shows that the method is essentially understood. Method marks are not usually lost for numerical errors, algebraic slips or errors in units. However, it is not usually sufficient for a candidate just to indicate an intention of using some method or just to quote a formula; the formula or idea must be applied to the specific problem in hand, eg by substituting the relevant quantities into the formula. In some cases the nature of the errors allowed for the award of an M mark may be specified.

A
Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. Accuracy marks cannot be given unless the associated Method mark is earned (or implied). Therefore M0 A1 cannot ever be awarded.

B
Mark for a correct result or statement independent of Method marks.

E
A given result is to be established or a result has to be explained. This usually requires more working or explanation than the establishment of an unknown result.

Unless otherwise indicated, marks once gained cannot subsequently be lost, eg wrong working following a correct form of answer is ignored. Sometimes this is reinforced in the mark scheme by the abbreviation isw. However, this would not apply to a case where a candidate passes through the correct answer as part of a wrong argument.

d. When a part of a question has two or more 'method' steps, the M marks are in principle independent unless the scheme specifically says otherwise; and similarly where there are several B marks allocated. (The notation 'dep *' is used to indicate that a particular mark is dependent on an earlier, asterisked, mark in the scheme.) Of course, in practice it may happen that when a candidate has once gone wrong in a part of a question, the work from there on is worthless so that no more marks can sensibly be given. On the other hand, when two or more steps are successfully run together by the candidate, the earlier marks are implied and full credit must be given.

e. The abbreviation ft implies that the A or B mark indicated is allowed for work correctly following on from previously incorrect results. Otherwise, A and B marks are given for correct work only — differences in notation are of course permitted. A (accuracy) marks are not given for answers obtained from incorrect working. When A or B marks are awarded for work at an intermediate stage of a solution, there may be various alternatives that are equally acceptable. In such cases, exactly what is acceptable will be detailed in the mark scheme rationale. If this is not the case please consult your Team Leader.

Sometimes the answer to one part of a question is used in a later part of the same question. In this case, A marks will often be 'follow through'. In such cases you must ensure that you refer back to the answer of the previous part question even if this is not shown within the image zone. You may find it easier to mark follow through questions candidate-by-candidate rather than question-by-question.

f. Wrong or missing units in an answer should not lead to the loss of a mark unless the scheme specifically indicates otherwise. Candidates are expected to give numerical answers to an appropriate degree of accuracy, with 3 significant figures often being the norm. Small variations in the degree of accuracy to which an answer is given (eg 2 or 4 significant figures where 3 is expected) should not normally be penalised, while answers which are grossly over- or under-specified should normally result in the loss of a mark. The situation regarding any particular cases where the accuracy of the answer may be a marking issue should be detailed in the mark scheme rationale. If in doubt, contact your Team Leader.

g. Rules for replaced work

If a candidate attempts a question more than once, and indicates which attempt he/she wishes to be marked, then examiners should do as the candidate requests.

If there are two or more attempts at a question which have not been crossed out, examiners should mark what appears to be the last (complete) attempt and ignore the others.

NB Follow these maths-specific instructions rather than those in the assessor handbook.

h. For a *genuine* misreading (of numbers or symbols) which is such that the object and the difficulty of the question remain unaltered, mark according to the scheme but following through from the candidate's data. A penalty is then applied; 1 mark is generally appropriate, though this may differ for some units. This is achieved by withholding one A mark in the question.

Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

Question			Answer	Marks	Guidance	
1	(i)		P(All blue) $= \frac{30}{50} \times \frac{29}{49} \times \frac{28}{48} = 0.2071$	M1	For $\frac{30}{50} \times$ (as part of a triple product)	$(30/50)^3 = 0.216$ scores M1M0A0 $\frac{k}{50} \times \frac{(k-1)}{49} \times \frac{(k-2)}{48}$ for values of k other than 30 scores M1M0A0 Zero for binomial unless simplifies to $(3/5)^3$
				M1	For product of other two fractions	Correct working but then multiplied or divided by some factor scores M1M0A0
			OR $\binom{30}{3} / \binom{50}{3} = 4060/19600 = 29/140 = 0.2071$ M2 for the complete method	A1 **[3]**	CAO SC2 for P(All red) = 0.0582	Accept 0.21 with working and 0.207 without working Allow unsimplified fraction as final answer 24360/117600 oe
1	(ii)		P(All red) $= \frac{20}{50} \times \frac{19}{49} \times \frac{18}{48} = 0.0582$ or $\binom{20}{3} / \binom{50}{3} = 0.0582$	M1	For P(All red)	SC2 for $1 - (30/50)^3 - (20/50)^3 = 1 - 0.216 - 0.064 = 0.72$, providing consistent with (i) . If not consistent with (i) M0M0A0
			P(At least one of each colour) $= 1 - (0.2071 + 0.0582) = 0.7347$ or $1 - \left(\frac{29}{140} + \frac{57}{980}\right) = 1 - \frac{260}{980} = 1 - \frac{13}{49} = \frac{36}{49}$	M1	For $1 - (0.2071 + 0.0582)$	
				A1 **[3]**	CAO	Allow 0.73 with working Allow unsimplified fraction as final answer 86400/117600 oe
			OR P(2b,1r)+P(1b,2r)	(M1)	For either $\frac{30}{50} \times \frac{29}{49} \times \frac{20}{48}$ or $\frac{20}{50} \times \frac{19}{49} \times \frac{30}{48}$	Allow M1 for $3 \times (30/50)^2 \times (20/50)$ or $3 \times (30/50) \times (20/50)^2$ and second M1 for sum of both if = 0.72 If not consistent with (i) M0M0A0

Question			Answer	Marks	Guidance	
			$= 3\times\frac{30}{50}\times\frac{29}{49}\times\frac{20}{48}+3\times\frac{20}{50}\times\frac{19}{49}\times\frac{30}{48}$	(M1)	For sum of both or for 3× either	NB M2 also for $\frac{30}{50}\times\frac{20}{49}\left(\times\frac{48}{48}\right)$ even if not multiplied by 3
			$= 3 \times 0.1480 + 3 \times 0.0969 = 0.7347$	(A1)	CAO	Allow 0.73 or better with working
			OR Either $\binom{30}{2}\times\binom{20}{1}/\binom{50}{3}$ or $\binom{30}{1}\times\binom{20}{2}/\binom{50}{3}$	(M1)		
				(M1)	For sum of both	
				(A1)	CAO	
2	**(i)**		$^9C_3 \times {}^5C_3 = 84 \times 10 = 840$	M1	For either $^9C_3$ or $^5C_3$	Zero for permutations
				M1	For product of both correct combinations	
				A1	CAO	
				[3]		
2	**(ii)**		Total number of ways of answering 6 from 14 is $^{14}C_6 = 3003$	M1	For $^{14}C_6$ seen in part (ii)	
			Probability $= \frac{840}{3003} = \frac{40}{143} = 0.27972 = 0.280$	M1	For their 840/ 3003 or their 840/$^{14}C_6$	
				A1	FT their 840	Allow full marks for unsimplified fractional answers
				[3]		
			OR $^6C_3 \times 5/14 \times 4/13 \times 3/12 \times 9/11 \times 8/10 \times 7/9 = 0.280$	(M1)	For product of fractions	SC1 for $^6C_3 \times (5/14)^3 \times (9/14)^3 = 0.2420$
				(M1)	For $^6C_3\times$ correct product	
				(A1)		

Question			Answer	Marks	Guidance	
3	(i)		$X \sim B(30, 0.85)$ $P(X = 29) = \binom{30}{29} \times 0.85^{29} \times 0.15^{1} = 30 \times 0.0013466 = 0.0404$	M1 M1 A1 **[3]**	For $0.85^{29} \times 0.15^{1} = 0.0013466$ For $\binom{30}{29} \times p^{29} \times q^{1}$ CAO	With $p + q = 1$ Allow 0.04 www If further working (EG P(X=29) –P(X=28)) give M2A0
3	(ii)		$P(X = 30) = 0.85^{30} = 0.0076$ $P(X \geq 29) = 0.0404 + 0.0076 = 0.0480$	M1 M1 A1 **[3]**	For 0.85^{30} For P(X = 29) + P(X = 30) (not necessarily correct, but both attempts at binomial, including coefficient in (i)) CAO	Allow eg 0.04+0.0076=0.0476 Allow 0.05 with working
3	(iii)		Expected number = 10 × 0.0480 = 0.480	M1 A1 **[2]**	For 10 × their **(ii)** FT their **(ii)** but if answer to (ii) leads to a whole number for (iii) give M1A0	**provided (ii) between 0 and 1** Do not allow answer rounded to 0 or 1.

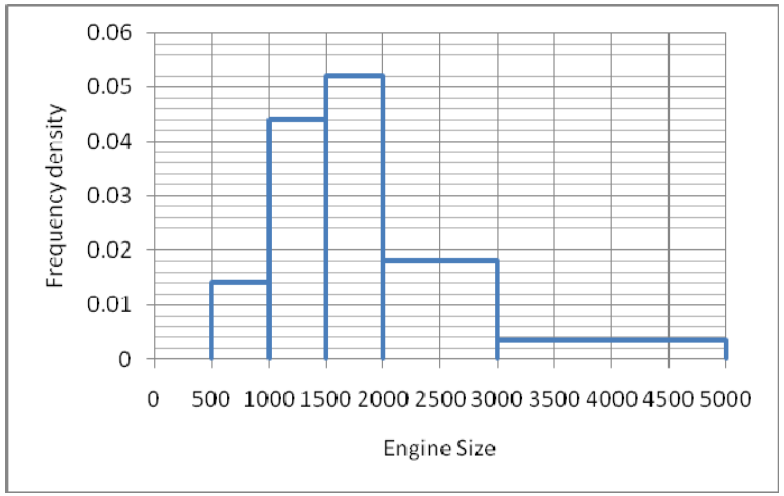
Question			Answer	Marks	Guidance	
4	(i)	(*A*)	P(third selected) = $0.92^2 \times 0.08 = 0.0677$ Or = 1058/15625	M1 M1 A1 **[3]**	For 0.92^2 For $p^2 \times q$ CAO SC1 for 'without replacement' method 92/100×91/99×8/98 =0.0690	With $p + q = 1$ With no extra terms Allow 0.068 but not 0.067 nor 0.07
4	(i)	(*B*)	P (second) + P(third) = (0.92 × 0.08) + ($0.92^2 \times 0.08$) = 0.0736 + 0.0677 = 0.1413 = 2208/15625	M1 A1 **[2]**	For 0.92 × 0.08 FT their 0.0677 SC1 for answer of 0.143 from 'without replacement' method	With no extra terms Allow 0.141 to 0.142 and allow 0.14 with working
4	(ii)		P(At least one of first 20) = 1 - P(None of first 20) = $1 - 0.92^{20} = 1 - 0.1887 = 0.8113$	M1 M1 A1 **[3]**	0.92^{20} $1 - 0.92^{20}$ CAO	Accept answer of 0.81 or better from P(1) + P(2) + ... , or SC2 if all correct working shown but wrong answer No marks for 'without replacement' method' Allow 0.81 with working but not 0.812

Question			Answer	Marks	Guidance	
5			Let p = probability that a randomly selected frame is faulty	B1	For definition of p in context Minimum needed for B1 is p = probability that frame/bike is faulty. Do not allow is p = probability that it is faulty Allow p = P(frame faulty) Definition of p must include word probability (or chance or proportion or percentage or likelihood but NOT possibility). Preferably as a separate comment. However can be at end of H_0 as long as it is a clear definition ' p = the probability that frame is faulty, NOT just a sentence 'probability is 0.05' Do NOT allow 'p = the probability that faulty frames have increased'	
			H_0: $p = 0.05$	B1	H_0 : p(frame faulty) = 0.05, H_1 : p(frame faulty) > 0.05 gets B0B1B1 Allow p=5%, allow θ or π and ρ but not x. However allow any single symbol <u>if defined</u> Allow $H_0 = p$=0.05, Allow $H_0 : p = {}^1/_{20}$ Do not allow H_0 : P(X=x) = 0.05, H_1 : P(X=x) > 0.05 Do not allow H_0: =0.05, =5%, P(0.05), p(0052), p(x)=0.05, x=0.05 (unless x correctly defined as a probability) Do not allow H_1:$p \geq 0.05$, Do not allow H_0 and H_1 reversed Allow NH and AH in place of H_0 and H_1 For hypotheses given in words allow Maximum B0B1B1 Hypotheses in words must include probability (or chance or proportion or percentage) and the figure 0.05 oe.	
			H_1: $p > 0.05$	B1		
			P($X \geq 4$)	B1	For notation P($X \geq 4$) or 1- P($X \leq 3$) This mark may be implied by 0.0109 as long as no incorrect notation.	No further marks if point probs used - P($X = 4$) = 0.0094 DO NOT FT wrong H_1 But if H_1 is $p \geq 0.05$ allow the rest of the marks if earned so max 7/8
			= 1- P($X \leq 3$) = 1 – 0.9891 = 0.0109	B1*	For 0.0109, indep of previous mark	Or for 1 – 0.9891

Question	Answer	Marks	Guidance	
	0.0109 < 0.05	M1* dep	For comparison with 5%	
	So reject H_0	A1*	or significant or 'accept H_1'	
	There is evidence to suggest that the proportion of faulty frames has increased.	E1* Dep on A1	Must include 'sufficient evidence' or something similar such as 'to suggest that' ie an element of doubt for E1. 'Sufficient evidence' or similar can be seen in the either the A mark or the E mark.	
		[8]		
	OR *Critical region method:*			No marks if CR not justified
	Let $X \sim B(18, 0.05)$			Do not insist on correct notation
	$P(X \geq 3) = 1 - P(X \leq 2) = 1 - 0.9419 = 0.0581 > 5\%$	(B1)	For 0.0581	as candidates have to work out two probabilities for full marks
	$P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0.9891 = 0.0109 < 5\%$	(B1)	For 0.0109	
		(M1)	For at least one correct comparison with 5%	
	So critical region is {4,5,6,7,8,9,10,11,12,13,14,15,16,17,18}	(A1)	CAO for critical region and significant oe	Condone {4,5 ..}, $X \geq 4$, oe but not $P(X \geq 4)$
	4 lies in the critical region, so significant,			
	There is evidence to suggest that the proportion of faulty frames has increased.	(E1)		

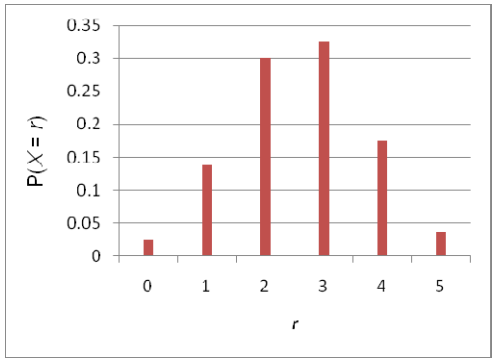
Question	Answer	Marks	Guidance
6 (i)		M1	At least 4 fds correct for M1 M1 can be also be gained from freq per 1000 – 14, 44, 52, 18, 3.5 (at least 4 correct) and A1 for all correct or freq per 500 - 7, 22, 26, 9, 1.75 Accept any suitable unit for fd, eg freq per 1000, BUT NOT FD per 1000 Allow fds correct to at least three dp If fd not explicitly given, M1 A1 can be gained from all heights correct (within one square) on histogram (and M1A0 if at least 4 correct) Allow restart with correct heights if given fd wrong
		A1	For fd's all correct
		G1(L1)	linear scales on both axes and label on vertical axis Label required on vert axis **IN RELATION** to first M1 mark ie fd or frequency density or if relevant freq/1000, etc (NOT fd/1000, but allow fd×1000, etc) Accept f/w or f/cw (freq/width or freq/class width) Ignore horizontal label and allow horizontal scale to start at 500 Can also be gained from an accurate key
	INCORRECT DIAGRAMS: Frequency diagrams can get M0, A0, G0, G1, G0 MAXIMUM Thus frequency density = frequency × width, frequency/midpoint etc gets MAX M0A0G0G1G0 **Frequency polygons MAX M1A1G0G0G0**	G1(W1)	Width of bars Must be drawn at 500, 1000etc NOT 499.5 or 500.5 etc NO GAPS ALLOWED Must have linear scale. No inequality labels on their own such as 500≤S<1000, etc but allow if a clear horizontal linear scale is also given.

Engine size	Frequency	Group width	Frequency density
$500 \leq x \leq 1000$	7	500	0.014
$1000 < x \leq 1500$	22	500	0.044
$1500 < x \leq 2000$	26	500	0.052
$2000 < x \leq 3000$	18	1000	0.018
$3000 < x \leq 5000$	7	2000	0.0035



Question			Answer	Marks	Guidance
				G1(H1) **[5]**	Height of bars FT of heights *dep* on at least 3 heights correct and all must agree with their fds If fds not given and one height is wrong then max M1A0G1G1G0 – visual check only (within one square) –no need to measure precisely
6	**(ii)**		Do not know exact highest and lowest values so cannot tell what the midrange is. **OR** No and a counterexample to show it may not be 2750 **OR** (500 + 5000) / 2 = 2750. But very unlikely to be absolutely correct but probably close to the true value. Some element of doubt needed. Allow 'Likely to be correct'	E1 **[1]**	Allow comment such as 'Highest value could be 5000 and lowest could be 500 therefore midrange could be 2750' NO mark if incorrect calculation Sight of 1750 AND 3000 (min and max of midrange) scores E1
6	**(iii)**		Mean = $\frac{(750\times 7)+(1250\times 22)+(1750\times 26)+(2500\times 18)+(4000\times 7)}{80}$ $=\frac{151250}{80}=1891$ $\Sigma x^2 f=(750^2\times 7)+(1250^2\times 22)+(1750^2\times 26)+(2500^2\times 18)+(4000^2\times 7)$ = 3937500 + 34375000 + 79625000 + 112500000 + 112000000 =342437500 $Sxx=342437500-\frac{151250^2}{80}=56480469$ $s=\sqrt{\frac{56480469}{79}}=\sqrt{714943}=846$ Only an estimate since the data are grouped.	M1 A1 M1 A1 E1 indep **[5]**	For midpoints (at least 3 correct) No marks for mean or sd unless using midpoints Answer must **NOT** be left as improper fraction CAO Accept correct answers for mean (1890 or 1891) and sd (850 or 846 or 845.5) from calculator even if eg wrong S_{xx} given For sum of at least 3 correct multiples fx^2 Allow M1 for anything which rounds to 342400000 Only penalise once in part (iii) for over specification, even if mean and standard deviation both over specified. Allow SC1 for RMSD 840.2 or 840 from calculator Or for any mention of midpoints or 'don't have actual data' or 'data are not exact' oe

Question			Answer	Marks	Guidance
6	(iv)		$\bar{x} - 2s = 1891 - (2 \times 846) = 199$ Allow 200	M1	For either. FT any positive mean and their positive sd/rmsd for M1 Only follow through numerical values, not variables such as *s*, so if a candidate does not find *s* but then writes here 'limit is 40.76+ 2 × standard deviation', do NOT award M1 No marks in (iv) unless using $\bar{x} + 2s$ or $\bar{x} - 2s$
			$\bar{x} + 2s = 1891 + (2 \times 846) = 3583$ Allow 3580 or 3600	A1	For both (FT) Do **NOT** penalise over specification here as it is not the final answer
			So there are probably some outliers	E1 **[3]**	Must include an element of doubt Dep on upper limit in range 3000 – 5000 Allow comments such as 'any value over 3583 is an outlier' Ignore comments about possible outliers at lower end.
6	(v)		Number of cars over 2000 cm^3 = 25/80 × 2.5 million = 781250 So duty raised = 781250 × £1000 = £781 million	M1 M1 indep A1 **[3]**	For 25/80× 2.5 million or (18+7) /80× 2.5 million For something × £1000 even if this is the first step CAO NB £781250000 is over specified so only 2/3
6	(vi)		Because the numbers of cars sold with engine size greater than 2000 cm^3 might be reduced due to the additional duty.	E1 **[1]**	Allow any other reasonable suggestion Condone 'sample may not be representative' Allow 'sample is not of **NEW** cars'

Question			Answer	Marks	Guidance
7	(i)		$P(X = 0) = 0.4 \times 0.5^4 = 0.025$ **NB ANSWER GIVEN**	M1 A1 **[2]**	For 0.5^4
7	(ii)		$P(X = 1) = (0.6 \times 0.5^4) + (4 \times 0.4 \times 0.5 \times 0.5^3)$ $= 0.0375 + 0.1 = 0.1375$ **NB ANSWER GIVEN**	M1* M1* M1* dep A1 **[4]**	For 0.6×0.5^4 seen as a single term (not multiplied or divided by anything) For $4 \times 0.4 \times 0.5^4$ Allow 4×0.025 Watch out for incorrect methods such as (0.4/4) 0.1 **MUST** be justified For sum of both , dep on both M1's
7	(iii)			G1 G1 **[2]**	For labelled linear scales on both axes Dep on attempt at vertical line chart. Accept P on vertical axis For heights – visual check only but last bar taller than first and fifth taller than second and fourth taller than third. Lines must be thin (gap width > line width). All correct. Zero if vertical scale not linear Everything correct but joined up tops G0G1 MAX Everything correct but f poly G0G1 MAX Everything correct but bar chart G0G1 MAX Curve only (no vertical lines) gets G0G0 Best fit line G0G0 Allow transposed diagram

Question			Answer	Marks	Guidance
7	(iv)		'Negative' or 'very slight negative'	E1 **[1]**	E0 for symmetrical but E1 for (very slight) negative skewness even if also mention symmetrical Ignore any reference to unimodal
7	(v)		$E(X) = (0\times0.025) + (1\times0.1375) + (2\times0.3) + (3\times0.325) + (4\times0.175) + (5\times0.0375)$ $= 2.6$ $E(X^2) = (0\times0.025) + (1\times0.1375) + (4\times0.3) + (9\times0.325) + 16\times0.175) + (25\times0.0375) = 0 + 0.1375 + 1.2 + 2.925 + 2.8 + 0.9375 = 8$ $\text{Var}(X) = 8 - 2.6^2$ $= 1.24$	M1 A1 M1* M1* dep A1 **[5]**	For Σrp (at least 3 terms correct) CAO For $\Sigma r^2 p$ (at least 3 terms correct) for – their E(X)² FT their E(X) provided Var(X) > 0 USE of $E(X-\mu)^2$ gets M1 for attempt at $(x-\mu)^2$ should see $(-2.6)^2$, $(-1.6)^2$, $(-0.6)^2$, 0.4^2, 1.4^2, 2.4^2 (if $E(X)$ correct but FT their $E(X)$) (all 5 correct for M1), then M1 for $\Sigma p(x-\mu)^2$ (at least 3 terms correct) Division by 5 or other spurious value at end gives max M1A1M1M1A0, or M1A0M1M1A0 if $E(X)$ also divided by 5. Unsupported correct answers get 5 marks.
7	(vi)		$P(\text{Total of } 3) = (3\times0.325\times0.025^2) + (6\times0.3\times0.1375\times0.025) + 0.1375^3 = 3\times0.000203 + 6\times0.001031 + 0.002600 =$ $0.000609 + 0.006188 + 0.002600 = 0.00940$ $(= 3\times13/64000 + 6\times33/32000 + 1331/512000)$	M1 M1 M1 A1 **[4]**	For decimal part of first term 0.325×0.025^2 For decimal part of second term $0.3\times0.1375\times0.025$ For third term – ignore extra coefficient All M marks above depend on triple probability products CAO: AWRT 0.0094. Allow 0.009 with working.

NOTE RE OVER-SPECIFICATION OF ANSWERS

If answers are grossly over-specified, deduct the final answer mark in every case. Probabilities should also be rounded to a sensible degree of accuracy. In general final non probability answers should not be given to more than 4 significant figures. Allow probabilities given to 5 sig fig. In general accept answers which are correct to 3 significant figures when given to 4 or 5 significant figures.
If answer given as a fraction and as an over-specified decimal – ignore decimal and mark fraction.

ADDITIONAL NOTES RE Q5

Comparison with 95% method
If 95% seen anywhere then
B1 for P($X \leq 3$)
B1 for 0.9891
M1* for comparison with 95% dep on B1
A1* for significant oe
E1*

Smallest critical region method:

Either:
Smallest critical region that 4 could fall into is {4,5,6,7,8,9,10,11,12,13,14,15, 16, 17, 18} gets B1 and has size 0.0109 gets B1, This is < 5% gets M1*, A1*, E1* as per scheme
NB These marks only awarded if 4 used, not other values.

Use of k method with no probabilities quoted:

P($X \geq 3$) = 1 – P($X \leq 2$) > 5%
P($X \geq 4$) = 1 – P($X \leq 3$) < 5%
These may be seen in terms of k or n.
Either $k = 4$ or $k – 1 = 3$ so $k = 4$ gets SC1
so CR is {4,5,6,7,8,9,10,11,12,13,14,15, 16, 17, 18} gets another SC1and conclusion gets another SC1

Use of k method with one probability quoted:

1 - 0.9891 < 5% or 0.0109 < 5% gets B0B1M1
P($X \leq k$ - 1) = P($X \leq 3$)
so $k – 1 = 3$ so $k = 4$ (or just $k = 8$)
so CR is {4,5,6,7,8,9,10,11,12,13,14,15, 16, 17, 18} and conclusion gets A1E1

Two tailed test done but with correct H_1: $p>0.05$
Hyp gets max B1B1B1
if compare with 5% ignore work on lower tail and mark upper tail as per scheme but withhold A1E1
if compare with 2.5% no marks B0B0M0A0E0

Line diagram method
B1 for squiggly line between 3 and 4 or on 4 exclusively (ie just one line), B1*dep* for arrow pointing to right, M1 0.0109 seen on diagram from squiggly line or from 4, A1E1 for correct conclusion

Bar chart method
B1 for line clearly on boundary between 3 and 4 or within 4 block exclusively (ie just one line), B1*dep* for arrow pointing to right, M1 0.0109 seen on diagram from boundary line or from 8, A1E1 for correct conclusion.

Using P(Not faulty) method
H_0: $p=0.95$, H_1: $p<0.95$ where p represents the prob that a frame is faulty gets B1B1B1.
$P(X\leq 14)=0.0109 < 5\%$ So significant, etc gets B1B1M1A1E1

NB
If H_0: $p=0.5$, H_1: $p>0.5$, etc seen, but then revert to 0.05 in working allow marks for correct subsequent working. However if 0.5 used consistently throughout, then max B1 for definition of p and possibly B1 for notation $P(X \geq 4)$.

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

OCR Customer Contact Centre

Education and Learning
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 2012

RECOGNISING ACHIEVEMENT

GCE

Statistics (MEI)

Advanced Subsidiary GCE **AS H132**

OCR Report to Centres

June 2012

Oxford Cambridge and RSA Examinations

G241 Statistics 1

General Comments

The level of difficulty of the paper appeared to be appropriate and there was no evidence of candidates being unable to complete the paper in the allocated time. Many candidates were well prepared but a number struggled to handle even the more standard parts of questions. Most candidates supported their numerical answers with appropriate explanations and working.

It is pleasing to report that the hypothesis test question was generally answered very well, with most candidates not only giving their hypotheses in terms of *p* but also defining *p* as the probability of a bike frame being faulty. Most candidates also included an element of doubt in their conclusion saying eg. 'There is sufficient evidence to suggest that the proportion of faulty frames has increased'. Unfortunately most candidates lost marks due to over specification of some of their answers, despite recent examiners' reports warning against this. Particular examples are given in the comments on 6(iii) and 6(v) below.

The comments below on individual questions are based on candidate responses to both 4766 and G241.

Comments on Individual Questions

1) (i) The majority of candidates gained full marks in this part. However a significant number completed this question as if the bulbs were replaced before each choice. This was given 1 mark out of the 3 available.

1) (ii) Once again the majority of candidates gained at least 2 marks out of 3. Those who answered by finding the P(2 blue and one red) and adding it to P(2 red and one blue) were in the majority, but were less successful than those who found 1 – (P(3 red) + P(3 blue)). This was due to the omission of the 3 possible arrangements of each probability.

2) (i) Again the majority of candidates gained full marks. A fairly common error was to add rather than multiply ${}^9C_3$ and ${}^5C_3$. A small number of candidates tried to use permutations rather than combinations.

2) (ii) Many candidates gained full credit for dividing their answer to part (i), correct or not, by 3003. Those who did not see the connection with part (i) did not fare so well, and even if they found the correct product of fractions they rarely multiplied this by ${}^6C_3$.

3) (i) This question was very well answered, with most candidates scoring all 3 marks.

3) (ii) Many fully correct responses were seen, although a number of candidates calculated P(X<29) or P(X≤29) rather than P(X≥29).

3) (iii) Many candidates gained full credit here, even if as a follow through from their answer to part (ii). A common error was to multiply their answer to part (ii) by 30 or 300 instead of by 10. A number of candidates also rounded their answer to a whole number, thereby losing the second mark.

4) (i)(*A*) Most candidates scored full marks, but a significant number scored zero. Candidates needed to multiply 0.92^2 by 0.08, but a significant number simply worked out 0.08^3, which gained no credit.

4) (i)(*B*) In this part candidates needed to multiply 0.92 by 0.08 then add this product to their answer to part (i). This was often achieved successfully, but a number of candidates gave their answer to 6 significant figures, thus losing the second mark.

4) (ii) Many fully correct responses were seen. Most other candidates gained no credit as they had no real idea of what was required.

5) Many candidates were awarded at least 7 out of the 8 marks available. The hypotheses were generally correct and well defined but a minority of candidates still omitted a definition of *p*. Only a small number of candidates incorrectly used point probabilities. Many candidates used the first method in the scheme, usually successfully. A smaller number used the critical region method, again fairly successfully, but a number thought that the critical region started at 3. Some candidates who used the critical region method, failed to justify their critical region. In this case they were only eligible for the first 3 marks for the hypotheses.

6) (i) On the whole, this question was answered well, with a very high proportion of candidates calculating the frequency densities correctly. Of those candidates who did not calculate the FD correctly, most achieved 1 mark for the correct widths. There were very few inequality labels on the x axis. However, candidates should be reminded that they need to label the vertical axis. Drawing of the bars was done well although a few candidates struggled to draw a bar of height 0.0035.

6) (ii) Many candidates thought that the calculation involved subtraction rather than addition and even when the calculation was correct, there was often no element of doubt to their conclusion.

6) (iii) On the whole this question was very well answered. It was extremely common to award 4 marks in total, due to the over specification of answers. Many candidates gave the exact answer 1890.625, but an element of sensible rounding, to say 1891 or even 1890, was looked for. A significant number did not find the standard deviation correctly, sometimes giving the root mean square deviation or calculating $(fx)^2$ rather than fx^2. The explanation mark was very well answered.

6) (iv) Again this was also very well answered. Even candidates who had made errors in the previous part usually gained follow through marks. Most candidates knew that the limits for outliers were mean ± 2 standard deviations. A number of candidates did not include an element of doubt in their conclusion about the number of outliers and thus were not awarded the final mark.

6) (v) The modal mark in this question was 2 out of 3, scored by approximately 2/3 of the candidature. This was due to candidates over specifying their answer, giving it as 781250000 rather than for example 780000000. Candidates who were unsure how to do this part nevertheless usually gained a method mark for multiplying by 1000.

6) (vi) Where candidates achieved this mark, they often realised that the duty would reduce the sales of larger cars. They also achieved this mark where they stated that the sample may not be representative, although this needed to be very clearly stated for the mark to be awarded. A number of candidates erroneously stated that people would refuse to pay the duty.

7) (i) This part was very well answered.

7) (ii) Again many fully correct responses were seen. Many other candidates scored 1 mark out of 4 for finding $0.6 \times 0.5^4 = 0.0375$. Some candidates multiplied 0.6×0.5^4 or 0.4×0.5^4 or indeed both by ${}^5C_1$ rather than by ${}^4C_1$.

7) (iii) Approximately half of candidates scored full marks in this part. However many lost one or both marks for a number of errors – a non linear vertical scale, one or both labels missing, heights incorrect (particularly the final height), or less often a frequency polygon or a point plot. Candidates should be advised to use a ruler in questions such as this.

7) (iv) Again approximately half of candidates scored the mark here. The most popular answer was 'slightly negative', but some said positive skew or symmetrical and/or unimodal.

7) (v) There were very many correct answers, with even the weakest candidates often scoring full marks. Arithmetic errors were common often because of writing the probabilities incorrectly as eg 0.375 rather than 0.0375. Only a few candidates left the variance as 8 or did not square E(X). Very few incorrectly divided by 5, unlike in previous sessions.

7) (vi) Candidates needed to have a very good understanding of probability to gain marks in this part. However, some got 1, 2 or 3 products of probabilities correct but very few had the coefficients correct.