OCR
RECOGNISING ACHIEVEMENT

Thursday 6 June 2013 – Morning

A2 GCE MATHEMATICS (MEI)

4767/01 Statistics 2

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4767/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

INSTRUCTIONS TO CANDIDATES

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

- The Question Paper will be found 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

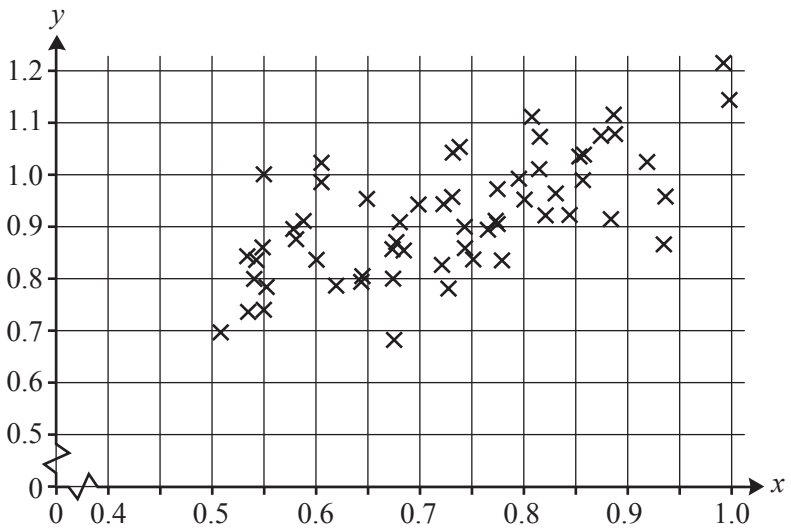
This information is the same on the Printed Answer Book and the Question Paper.

- 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.

1 Salbutamol is a drug used to improve lung function. In a medical trial, a random sample of 60 people with impaired lung function was selected. The forced expiratory volume in one second (FEV1) was measured for each person, both before being given salbutamol and again after a two-week course of the drug. The variables x and y, measured in suitable units, represent FEV1 before and after the two-week course respectively. The data are illustrated in the scatter diagram below, together with the summary statistics for these data.



Summary statistics:

$$n = 60, \quad \Sigma x = 43.62, \quad \Sigma y = 55.15, \quad \Sigma x^2 = 32.68, \quad \Sigma y^2 = 51.44, \quad \Sigma xy = 40.66$$

(i) Calculate the sample product moment correlation coefficient. **[5]**

(ii) Carry out a hypothesis test at the 5% significance level to investigate whether there is positive correlation between FEV1 before and after the course. **[6]**

(iii) State the distributional assumption which is necessary for this test to be valid. State, with a reason, whether the assumption appears to be valid. **[2]**

(iv) Explain the meaning of the term 'significance level'. **[2]**

(v) Calculate the values of the summary statistics if the data point $x = 0.55$, $y = 1.00$ had been incorrectly recorded as $x = 1.00$, $y = 0.55$. **[3]**

2 Suppose that 3% of the population of a large city have red hair.

(i) A random sample of 10 people from the city is selected. Find the probability that there is at least one person with red hair in this sample. **[2]**

A random sample of 60 people from the city is selected. The random variable X represents the number of people in this sample who have red hair.

(ii) Explain why the distribution of X may be approximated by a Poisson distribution. Write down the mean of this Poisson distribution. **[3]**

(iii) Hence find

(A) $\mathrm{P}(X = 2)$, **[2]**

(B) $\mathrm{P}(X > 2)$. **[2]**

(iv) Discuss whether or not it would be appropriate to model X using a Normal approximating distribution. **[2]**

A random sample of 5000 people from the city is selected.

(v) State the exact distribution of the number of people with red hair in the sample. **[2]**

(vi) Use a suitable Normal approximating distribution to find the probability that there are at least 160 people with red hair in the sample. **[5]**

3 The scores, X, in Paper 1 of an English examination have an underlying Normal distribution with mean 76 and standard deviation 12. The scores are reported as integer marks. So, for example, a score for which $75.5 \leqslant X < 76.5$ is reported as 76 marks.

(i) Find the probability that a candidate's reported mark is 76. **[4]**

(ii) Find the probability that a candidate's reported mark is at least 80. **[3]**

(iii) Three candidates are chosen at random. Find the probability that exactly one of these three candidates' reported marks is at least 80. **[2]**

The proportion of candidates who receive an A* grade (the highest grade) must not exceed 10% but should be as close as possible to 10%.

(iv) Find the lowest reported mark that should be awarded an A* grade. **[5]**

The scores in Paper 2 of the examination have an underlying Normal distribution with mean μ and standard deviation 12.

(v) Given that 20% of candidates receive a reported mark of 50 or less, find the value of μ. **[4]**

4 An art gallery is holding an exhibition. A random sample of 150 visitors to the exhibition is selected. The visitors are asked which of four artists they prefer. Their preferences, classified according to whether the visitor is female or male, are given in the table.

		Artist preferred			
		Monet	Renoir	Degas	Cézanne
Sex	Male	8	25	18	19
	Female	18	35	10	17

(i) Carry out a test at the 10% significance level to examine whether there is any association between artist preferred and sex of visitor. Your working should include a table showing the contributions of each cell to the test statistic. **[12]**

(ii) For each artist, comment briefly on how the preferences of each sex compare with what would be expected if there were no association. **[6]**

OCR
RECOGNISING ACHIEVEMENT

Thursday 6 June 2013 – Morning

A2 GCE MATHEMATICS (MEI)

4767/01 Statistics 2

PRINTED ANSWER BOOK

Candidates answer on this Printed Answer Book.

OCR supplied materials:
- Question Paper 4767/01 (inserted)
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

* 4 7 6 7 0 1 *

* 4 7 1 5 8 6 0 6 1 3 *

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.

1 (i)	

1 (ii)	
1 (iii)	

1 (iv)	
1 (v)	

2 (i)

2 (ii)

2 (iii) (*A*)

2 (iii) (*B*)

2 (iv)	
2 (v)	
2 (vi)	

3 (i)

3 (ii)

3 (iii)

3 (iv)	

3 (v)	

4 (i)	
	(answer space continued on next page)

4 (i)	**(continued)**
4 (ii)	
	(answer space continued on next page)

4 (ii)	**(continued)**

OCR
RECOGNISING ACHIEVEMENT

GCE

Mathematics (MEI)

Advanced GCE

Unit **4767:** Statistics 2

Mark Scheme for June 2013

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, Cambridge Nationals, Cambridge Technicals, 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.

1. **Annotations and abbreviations**

Annotation in scoris	Meaning
 and 	
	Benefit of doubt
	Follow through
	Ignore subsequent working
 ,	Method mark awarded 0, 1
 ,	Accuracy mark awarded 0, 1
 ,	Independent mark awarded 0, 1
	Special case
	Omission sign
	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

12. **Subject-specific Marking Instructions for GCE Mathematics (MEI) Statistics strand**

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. 3 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 probabilities from Normal tables, we generally expect *some* evidence of interpolation and so quotation to 4 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, especially if the calculated value of a test statistic is nowhere near any of these values. Sensible discretion *must* be exercised in such cases.

Discretion must also be exercised in the case 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, loss of an accuracy ("A") mark is likely to be appropriate; but if 4 significant figures are given, this should not normally be penalised. Likewise, answers which are slightly deviant from what is expected in a very minor manner (for example a Normal probability given, after an attempt at interpolation, as 0.6418 whereas 0.6417 was expected) should not be penalised. However, answers which are *grossly* over- or under-specified should normally result in the loss of a mark. This includes cases such as, for example, insistence that the value of a test statistic is (say) 2.128888446667

merely because that is the value that happened to come off the candidate's calculator. Note that this applies to answers that are given as final stages of calculations; intermediate working should usually be carried out, and quoted, to a greater degree of accuracy to avoid the danger of premature approximation.
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. Genuine misreading (of numbers or symbols, occasionally even of text) occurs. If this results in the object and/or difficulty of the question being considerably changed, it is likely that all the marks for that question, or section of the question, will be lost. However, misreads are often such that the object and/or difficulty remain substantially unaltered; these cases are considered below.

The simple rule is that *all* method ("M") marks [and of course all independent ("B") marks] remain accessible but at least some accuracy ("A") marks do not. It is difficult to legislate in an overall sense beyond this global statement because misreads, even when the object and/or difficulty remains unchanged, can vary greatly in their effects. For example, a misread of 1.02 as 10.2 (perhaps as a quoted value of a sample mean) may well be catastrophic; whereas a misread of 1.6748 as 1.6746 may have so slight an effect as to be almost unnoticeable in the candidate's work.

A misread should normally attract *some* penalty, though this would often be only 1 mark and should rarely if ever be more than 2. Commonly in sections of questions where there is a numerical answer either at the end of the section or to be obtained and commented on (eg the value of a test statistic), this answer will have an "A" mark that may actually be designated as "cao" [correct answer only]. This should be interpreted *strictly* – if the misread has led to failure to obtain this value, then this "A" mark must be withheld even if all method marks have been earned. It will also often be the case that such a mark is implicitly "cao" even if not explicitly designated as such.

On the other hand, we commonly allow "fresh starts" within a question or part of question. For example, a follow-through of the candidate's value of a test statistic is generally allowed (and often explicitly stated as such within the marking scheme), so that the candidate may exhibit knowledge of how to compare it with a critical value and draw conclusions. Such "fresh starts" are not affected by any earlier misreads.
A misread may be of a symbol rather than a number – for example, an algebraic symbol in a mathematical expression. Such

misreads are more likely to bring about a considerable change in the object and/or difficulty of the question; but, if they do not, they should be treated as far as possible in the same way as numerical misreads, *mutatis mutandis*. This also applied to misreads of text, which are fairly rare but can cause major problems in fair marking.

The situation regarding any particular cases that arise while you are marking for which you feel you need detailed guidance should be discussed with your Team Leader.

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

Question			Answer	Marks	Guidance	
1	(i)		**EITHER:**			
			$S_{xy} = \sum xy - \frac{1}{n}\sum x \sum y = 40.66 - \frac{1}{60} \times 43.62 \times 55.15$	M1*	For method for S_{xy}	
			$= 0.56595$			
			$S_{xx} = \sum x^2 - \frac{1}{n}\left(\sum x\right)^2 = 32.68 - \frac{1}{60} \times 43.62^2$	M1*	For method for at least one of S_{xx} or S_{yy}	
			$= 0.96826$			
			$S_{yy} = \sum y^2 - \frac{1}{n}\left(\sum y\right)^2 = 51.44 - \frac{1}{60} \times 55.15^2$	A1	For at least one of S_{xy}, S_{xx} or S_{yy} (to 2 sf) Note Allow 0.57322 for S_{xy} and 0.76634 for S_{yy} from rounding mean of y to 0.919.	
			$= 0.74796$			
			$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{0.56595}{\sqrt{0.96826 \times 0.74796}} = 0.665$	M1 dep* A1	For structure of r For answer rounding to 0.66 or 0.67	
			OR:	**[5]**		
			$\text{cov}(x,y) = \frac{\sum xy}{n} - \bar{x}\bar{y} = 40.66/60 - (43.62/60 \times 55.15/60)$	M1*	For method for cov (x,y)	
			$= 0.0094325$			
			$\text{rmsd}(x) = \sqrt{\frac{S_{xx}}{n}} = \sqrt{(0.96826/60)} = \sqrt{0.016137\ldots} = 0.1270$	M1*	For method for at least one msd or rmsd	
			$\text{rmsd}(y) = \sqrt{\frac{S_{yy}}{n}} = \sqrt{(0.74796/60)} = \sqrt{0.012466} = 0.1117$	A1	For at least one of cov (x,y), msd or rmsd correct (to 2 sf)	
			$r = \frac{\text{cov}(x,y)}{\text{rmsd}(x)\text{rmsd}(y)} = \frac{0.0094325}{0.1270 \times 0.1117} = 0.665$	M1 dep* A1	For structure of r For answer rounding to0.66 or 0.67	
					Methods mixed – max M0M1A1M0A0	
				[5]		

Question			Answer	Marks	Guidance	
1	(ii)		H_0: $\rho = 0$ H_1: $\rho > 0$ (one-tailed test)	B1	For H_0, H_1 in symbols. Hypotheses in words must refer to population. Do not allow alternative symbols unless clearly defined as the population correlation coefficient.	
			where ρ is the population correlation coefficient	B1	For defining ρ. Condone omission of "population" if correct notation ρ is used, but if ρ is defined as the **sample** correlation coefficient then award **B0**. Allow "ρ is the pmcc".	
			For $n = 60$, 5% critical value = 0.2144	B1	For critical value	
			Since 0.665 > 0.2144, the result is significant.	M1	For sensible comparison leading to a conclusion provided that $\|r\| < 1$. The comparison can be in the form of a diagram as long as it is clear and unambiguous. Sensible comparison: e.g. 0.665 > 0.2144 is 'sensible' whereas 0.665 > - 0.2144 is 'not sensible'. Reversed inequality sign e.g. 0.665 < 0.2144 etc. gets max M1 A0.	
			Thus we have sufficient evidence to reject H_0	A1	For reject H_0 o.e. FT their r and critical value from 5% 1-tail column.	
			There is sufficient evidence at the 5% level to **suggest** that there is **positive** correlation between FEV1 before and after the two-week course.	E1	For correct, **non-assertive** conclusion in context (allow 'x and y' for context). E0 if H_0 and H_1 not stated, reversed or mention a value other than zero for ρ in H_0.	
				[6]		

Question			Answer	Marks	Guidance	
1	(iii)		The underlying population must have a bivariate Normal distribution. Yes, since the scatter diagram appears to have a roughly elliptical shape.	B1 E1 **[2]**	Condone "bivariate Normal distribution", "underlying bivariate Normal distribution", but **do not allow** "the **data** have a bivariate Normal distribution" Condone 'oval' or suitable diagram	
1	(iv)		The significance level is the probability of rejecting the null hypothesis when in fact it is true.	E1* E1dep* **[2]**	For "probability of rejecting H_0" or "probability of a significant result". For "when H_0 is true"	
1	(v)		$\sum x = 43.62 + 0.45 = 44.07$ $\sum y = 55.15 - 0.45 = 54.70$ $\sum xy = 40.66$ $\sum x^2 = 32.68 + 1 - 0.55^2 = 33.3775$ $\sum y^2 = 51.44 - 1 + 0.55^2 = 50.7425$	B1 B1 B1 **[3]**	For $\sum x$ or $\sum y$ or $\sum xy$ For $\sum x^2$ or $\sum y^2$ (to 2 dp) For all correct (ignore n)	
2	(i)		P(At least one has red hair) = $1 - 0.97^{10}$ = 0.263	M1 A1 **[2]**	M1 for $1 - 0.97^{10}$ Allow 0.26	
2	(ii)		(Because X is binomially distributed), n is large and p is small. Mean = 1.8	E1 E1 B1 **[3]**	Allow "sample is large" for n is large Allow "$np < 10$" or "mean ≈ variance" for "p is small" Do not allow "the probability is small"	

Question			Answer	Marks	Guidance	
2	(iii)	(*A*)	$P(X = 2) = e^{-1.8}\frac{1.8^2}{2!} = 0.2678$ OR = 0.7306 – 0.4628 = 0.2678	M1 A1 **[2]**	For calculation for $P(X = 2)$ FT their mean. Allow answer to 3sf.	
2	(iii)	(*B*)	$P(X > 2) = 1 – P(X \leq 2) = 1 – 0.7306$ = 0.2694	M1 A1 **[2]**	$1 – P(X \leq 2)$ used. e.g.$1 – P(X \leq 2) = 1 – 0.4628$ gets M0 CAO	
2	(iv)		The mean ($np = 1.8$) is too small It is not appropriate to use a Normal approximation	E1* E1dep* **[2]**	For "mean is too small" or "mean < 10" For "not appropriate". Do not allow "p is too small".	
2	(v)		Binomial(5000, 0.03)	B1* B1dep* **[2]**	For binomial, or B(,) For parameters	
2	(vi)		Mean $5000 \times 0.03 = 150$ Variance $= 5000 \times 0.03 \times 0.97 = 145.5$ Using Normal approx. to the binomial, $X \sim N(150, 145.5)$ $P(X \geq 160) = P\left(Z \geq \frac{159.5 – 150}{\sqrt{145.5}}\right)$ $= P(Z > 0.7876) = 1 – \Phi(0.7876) = 1 – 0.7846$ = 0.215 (to 3 sig.fig.)	B1 B1 B1 M1 A1 **[5]**	For mean (soi) For variance (soi) For continuity corr. For probability using correct tail and structure (condone omission of/incorrect c.c.) CAO, (Do not FT wrong or omitted CC) Allow 0.2155. Do not allow 0.216	

Question			Answer	Marks	Guidance	
3	(i)		$P(Y = 76) = P\left(\frac{75.5-76}{12} \le Z \le \frac{76.5-76}{12}\right)$ $= P(-0.04166\ldots < Z < 0.04166\ldots)$ $= \Phi(0.04166\ldots) - (1 - \Phi(0.04166\ldots))$ $= 2 \times \Phi(0.04166\ldots) - 1$ $= 2 \times 0.5167 - 1$ $= 0.0334$	B1 M1 M1 A1 **[4]**	For one correct continuity correction used For standardizing For correctly structured probability calculation. CAO inc use of diff tables. Allow 0.0330 – 0.0340 www.	
3	(ii)		$P(Y \ge 80) = P\left(Z \ge \frac{79.5-76}{12}\right)$ $= P(Z > 0.2917) = 1 - \Phi(0.2917)$ $= 1 - 0.6148 = 0.3852 = 0.385$ to 3 sig fig	B1 M1 A1 **[3]**	For correct cc used For correct structure CAO do not allow 0.386	
3	(iii)		$3 \times 0.3852 \times 0.6148^2 = 0.4368$	M1 A1 **[2]**	$3 \times$ their $p \times (1 -$ their $p)^2$ FT their p. Allow 2sf if working seen.	

Question			Answer	Marks	Guidance	
3	(iv)		**EITHER:** P(Score $\geq k$) = 0.1			
			$\Phi^{-1}(0.9) = 1.282$	B1	For 1.282	
			$\frac{k-76}{12} = 1.282$	M1	Allow $k - 0.5$ used for k. Positive z used.	
			$k = 76 + (1.282 \times 12) = 91.38$ or $k = 76 + 0.5 + (1.282 \times 12) = 91.88$	A1	For 91.38 or 91.88	
			91.38 > 90.5 or 91.88 > 91	M1	Relevant comparison (e.g. diagram)	www
			so lowest reported mark = 92	A1		
			OR Trial and improvement method	M1	M1 for attempt to find P(Mark ≥ integer)	
			P(Mark ≥ 91) = P(Score ≥ 90.5) = 0.1135	A1	A1 for 0.1135	
			P(Mark ≥ 92) = P(Score ≥ 91.5) = 0.0982	A1	A1 for 0.0982	
			P(Mark ≥ 91) > 10% and P(Mark ≥ 92) < 10%	M1	M1 for comparisons	www
			so lowest reported mark = 92	A1		
				[5]		
3	(v)		P($Y \leq 50$) = 0.2			
			$P(Z \leq \frac{50.5-\mu}{12}) = 0.2$	B1	For 50.5 used	
			$\frac{50.5-\mu}{12} = \Phi^{-1}(0.2) = -0.8416$	B1	For –0.8416. Condone – 0.842 Condone 0.8416 if numerator reversed.	
				M1	For structure.	
			$\mu = 50.5 + (12 \times 0.8416) = 60.6$	A1	CAO	
				[4]		

<table>
<tr><th colspan="3">Question</th><th>Answer</th><th>Marks</th><th>Guidance</th><th></th></tr>
<tr><td>4</td><td>(i)</td><td></td><td>H_0: no association between sex and artist preferred
H_1: some association between sex and artist preferred

<table><tr><td>EXPECTED</td><td>Monet</td><td>Renoir</td><td>Degas</td><td>Cézanne</td></tr><tr><td>Male</td><td>12.13</td><td>28</td><td>13.07</td><td>16.8</td></tr><tr><td>Female</td><td>13.87</td><td>32</td><td>14.93</td><td>19.2</td></tr></table>
<table><tr><td>CONTRIB’N</td><td>Monet</td><td>Renoir</td><td>Degas</td><td>Cézanne</td></tr><tr><td>Male</td><td>1.4081</td><td>0.3214</td><td>1.8626</td><td>0.2881</td></tr><tr><td>Female</td><td>1.2321</td><td>0.2813</td><td>1.6298</td><td>0.2521</td></tr></table>
$X^2 = 7.28$
Refer to χ_3^2

Critical value at 10% level = 6.251

Result is significant

There is evidence to suggest that there is some association between sex and artist preferred

NB if H_0 H_1 reversed, or ‘correlation’ mentioned, do not award first B1 or final E1</td><td>B1

M1
A2

M1
A2

B1
B1

B1

B1

E1

[12]</td><td>For both hypotheses in context

For expected values
(to 2 dp where appropriate)
(allow A1 for at least one row or column correct)

For valid attempt at $(O–E)^2/E$
For all correct (to 2 dp) and presented in a table or clear list. (Allow A1 for at least one row or column correct)

Allow 7.27
for 3 deg of f

CAO for cv No FT from here if wrong or omitted, unless p-value used instead
FT their X^2

For correct (FT their X^2), non-assertive conclusion, in context.</td><td>

NB:These three marks cannot be implied by a correct final value of X^2

www

B1 for p-value = 0.0636</td></tr>
</table>

Question			Answer	Marks	Guidance	
4	(ii)		Monet: **More females and fewer males than expected** prefer Monet, as indicated by l**arge contribution(s)** (of 1.4081 and 1.2321). Renoir: Preferences are much **as expected**, as indicated by s**mall contributions.** Degas: **Fewer females and more males than expected** prefer Degas, as indicated by **large contribution(s)** (of 1.8626 and 1.6298). Cézanne: Preferences are much **as expected**, as indicated by s**mall contributions**.	E1* E1dep* E1 E1* depE1* E1 **[6]**	FT their table of contributions	NB MAX 3/6 for answers not referring to contributions (explicitly or implicitly). SC1 Renoir and Cézanne have correct comments for both but without referring to contributions

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

RECOGNISING ACHIEVEMENT

GCE

Mathematics (MEI)

Advanced GCE **A2 7895-8**

Advanced Subsidiary GCE **AS 3895-8**

OCR Report to Centres

June 2013

4767 Statistics 2

General Comments

Candidates appeared to have enough time to complete the question paper and seemed comfortable with the level of difficulty. The majority attempted all parts of all questions. A very small number made little or no attempt at any of the questions. In general, solutions were well-presented and easy to follow. Candidates continue to improve their techniques in tackling hypothesis tests and the use of non-assertive conclusions is becoming quite commonplace. Over-specified answers are also becoming less prevalent.

Comments on Individual Questions

1(i) Most candidates coped very well with this question. A few candidates did not calculate intermediate values and ran the risk of losing marks if they made an error in calculating the final value. Some candidates used a rounded value for the mean of *y* throughout, showing little awareness of rounding error.

1(ii) This part was well-answered by many, with the customary form of the hypotheses using the appropriate symbol, ρ, being seen. Having successfully looked up the appropriate critical value, many went on to demonstrate the comparison with the sample value before rejecting the null hypothesis and concluding in an appropriate, *non-assertive* manner.

1(iii) Many candidates recognised the requirement for the underlying bivariate Normal distribution and knew what to look for in the shape of the scatter diagram. The main confusion concerned the words “bivariate Normal”, with “normal bivariate” regularly seen along with a host of poor spellings.

1(iv) Many candidates found this difficult. In trying to explain that the significance level was the probability of rejecting the null hypothesis when it is true, many ended-up writing meaningless phrases.

1(v) Most coped well with this part of the question. Even the few who were confused as to which point needed changing gained credit for realising there would be no change in Σxy.

2(i) Generally well done. The most common error was to use a Poisson distribution either directly or from tables. Another common error involved calculating the probability of ’one person’ having red hair, rather than ’at least one’.

2(ii) Most responses seen were entirely correct, with occasional omissions of one or more of the three requirements. Some candidates reeled out general reasons why a Poisson distribution might apply rather than explaining why the binomial distribution could be approximated using a Poisson distribution. Ambiguous phrases such as “the probability is small” were not given credit.

2(iii)*A* Very few incorrect answers seen.

2(iii)*B* Again, very few incorrect answers seen. Some candidates calculated $P(X \geq 2)$ rather than $P(X > 2)$.

2(iv) The least well answered part of the question. Many candidates chose to discuss, individually, the size of n and p rather than the size of np, failing to note that np was too small to use a normal approximation.

2(v) Whilst this was mostly done correctly, many candidates mistook this either for a Poisson distribution with mean 150, leading them to lose marks in part 2(vi) with the wrong variance, or for a N(150, 145.5) distribution.

2(vi) A variety of errors cropped up amongst the many completely correct answers to this part. These included using the wrong variance, not using a continuity correction, using a wrong continuity correction and using the wrong tail of the distribution (i.e. giving Φ(0.7876) as the answer).

3(i) Most candidates obtained a correct answer. A small but significant number did not use one or both of the continuity corrections. Most used the difference column of the Standard Normal table correctly to provide suitably accurate answers. A relatively small number struggled with the structure of the calculation.

3(ii) Many candidates answered correctly, but a common mistake was to omit or to provide an incorrect continuity correction. A relatively small number did not use the difference column correctly.

3(iii) Most candidates knew and applied the method correctly but many were dependent on the FT to gain the 2 marks. A small number omitted the x3 from their binomial calculation.

3(iv) 1.282 was identified by the majority of candidates who went on to set up a correct equation and arrive at 91.38 or 91.88. Many of these gave 92% as the final answer but many others gave 91%. Others rearranged incorrectly and arrived at 60.6 for the first calculation. Few candidates demonstrated a proper understanding of the requirement of this question.

3(v) In this part, the continuity correction was omitted, or an incorrect value was used, by many candidates. Many used +0.8416 leading to 60.0992 which was a common answer. The issue of over-specification was most apparent in this part of the question.

4(i) Most candidates were able to give two correct hypotheses. Mistakes were rare, but included reversed hypotheses, hypotheses with no context and hypotheses which included correlation. Most candidates knew how to calculate the expected values, but a significant number did so with insufficient accuracy, either by rounding inaccurately or simply by getting one or two of the calculations wrong. The same was also true of the table of contributions, although there were a few more candidates who did not know how to calculate contributions. Almost all candidates knew to calculate χ^2, and that there were 3 degrees of freedom. Some candidates obtained the critical value for the wrong significance level or used the wrong row in the table. Some candidates thought that, as the test statistic was greater than the critical value, this was not a significant result and so accepted H_0. In the conclusion, many candidates knew what was required, but there were two reasonably common errors. The first of these was an over-assertive conclusion such as “there is some association between sex and artist preferred”. The second main error was a lack of context in the conclusion.

4(ii) This was poorly answered by most candidates. Very few seemed to understand the importance of the contributions and did not comment on them. Some candidates mentioned the contribution, but did not comment on the magnitude. For example “More females than expected preferred Monet, as indicated by a contribution of 1.408”. The word ’large’ was required to indicate the special feature of the contribution. A similar problem occurred with Renoir and Cezanne where the important feature of the contributions was that they were small. With small contributions the response should be that the observed frequencies are much as expected. Many candidates tried to ‘sit on the fence’ and state that the frequencies were much as expected but slightly more (or less). This led to a loss of marks. Poorer responses were characterised by two features. One was a lack of clarity. It was often unclear whether the candidate was saying the observed frequency was greater than the expected frequency, or vice-versa. The second feature was a comparison of male and female figures rather than of observed and expected values.

For a description of how UMS marks are calculated see:
www.ocr.org.uk/learners/ums_results.html

Unit level raw mark and UMS grade boundaries June 2013 series
AS GCE / Advanced GCE / AS GCE Double Award / Advanced GCE Double Award

GCE Mathematics (MEI)

			Max Mark	a	b	c	d	e	u
4751/01	(C1) MEI Introduction to Advanced Mathematics	Raw	72	62	56	51	46	41	0
		UMS	100	80	70	60	50	40	0
4752/01	(C2) MEI Concepts for Advanced Mathematics	Raw	72	54	48	43	38	33	0
		UMS	100	80	70	60	50	40	0
4753/01	(C3) MEI Methods for Advanced Mathematics with Coursework: Written Paper	Raw	72	58	52	46	40	33	0
4753/02	(C3) MEI Methods for Advanced Mathematics with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4753/82	(C3) MEI Methods for Advanced Mathematics with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4753	(C3) MEI Methods for Advanced Mathematics with Coursework	UMS	100	80	70	60	50	40	0
4754/01	(C4) MEI Applications of Advanced Mathematics	Raw	90	66	59	53	47	41	0
		UMS	100	80	70	60	50	40	0
4755/01	(FP1) MEI Further Concepts for Advanced Mathematics	Raw	72	63	57	51	45	40	0
		UMS	100	80	70	60	50	40	0
4756/01	(FP2) MEI Further Methods for Advanced Mathematics	Raw	72	61	54	48	42	36	0
		UMS	100	80	70	60	50	40	0
4757/01	(FP3) MEI Further Applications of Advanced Mathematics	Raw	72	60	52	44	36	28	0
		UMS	100	80	70	60	50	40	0
4758/01	(DE) MEI Differential Equations with Coursework: Written Paper	Raw	72	62	56	51	46	40	0
4758/02	(DE) MEI Differential Equations with Coursework: Coursework	Raw	18	15	13	11	9	8	0
4758/82	(DE) MEI Differential Equations with Coursework: Carried Forward Coursework Mark	Raw	18	15	13	11	9	8	0
4758	(DE) MEI Differential Equations with Coursework	UMS	100	80	70	60	50	40	0
4761/01	(M1) MEI Mechanics 1	Raw	72	57	49	41	33	25	0
		UMS	100	80	70	60	50	40	0
4762/01	(M2) MEI Mechanics 2	Raw	72	50	43	36	29	22	0
		UMS	100	80	70	60	50	40	0
4763/01	(M3) MEI Mechanics 3	Raw	72	64	56	48	41	34	0
		UMS	100	80	70	60	50	40	0
4764/01	(M4) MEI Mechanics 4	Raw	72	56	49	42	35	29	0
		UMS	100	80	70	60	50	40	0
4766/01	(S1) MEI Statistics 1	Raw	72	55	48	41	35	29	0
		UMS	100	80	70	60	50	40	0
4767/01	(S2) MEI Statistics 2	Raw	72	58	52	46	41	36	0
		UMS	100	80	70	60	50	40	0
4768/01	(S3) MEI Statistics 3	Raw	72	61	55	49	44	39	0
		UMS	100	80	70	60	50	40	0
4769/01	(S4) MEI Statistics 4	Raw	72	56	49	42	35	28	0
		UMS	100	80	70	60	50	40	0
4771/01	(D1) MEI Decision Mathematics 1	Raw	72	58	52	46	40	35	0
		UMS	100	80	70	60	50	40	0
4772/01	(D2) MEI Decision Mathematics 2	Raw	72	58	52	46	41	36	0
		UMS	100	80	70	60	50	40	0
4773/01	(DC) MEI Decision Mathematics Computation	Raw	72	46	40	34	29	24	0
		UMS	100	80	70	60	50	40	0
4776/01	(NM) MEI Numerical Methods with Coursework: Written Paper	Raw	72	56	50	44	38	31	0
4776/02	(NM) MEI Numerical Methods with Coursework: Coursework	Raw	18	14	12	10	8	7	0
4776/82	(NM) MEI Numerical Methods with Coursework: Carried Forward Coursework Mark	Raw	18	14	12	10	8	7	0
4776	(NM) MEI Numerical Methods with Coursework	UMS	100	80	70	60	50	40	0
4777/01	(NC) MEI Numerical Computation	Raw	72	55	47	39	32	25	0
		UMS	100	80	70	60	50	40	0
4798/01	(FPT) Further Pure Mathematics with Technology	Raw	72	57	49	41	33	26	0
		UMS	100	80	70	60	50	40	0

GCE Statistics (MEI)

			Max Mark	a	b	c	d	e	u
G241/01	(Z1) Statistics 1	Raw	72	55	48	41	35	29	0
		UMS	100	80	70	60	50	40	0
G242/01	(Z2) Statistics 2	Raw	72	55	48	41	34	27	0
		UMS	100	80	70	60	50	40	0
G243/01	(Z3) Statistics 3	Raw	72	56	48	41	34	27	0
		UMS	100	80	70	60	50	40	0