Oxford Cambridge and RSA

Wednesday 15 June 2016 – 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

5996471967

INSTRUCTIONS TO CANDIDATES

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

- The Question Paper will be found inside 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.** If additional space is required, you should use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- 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 A researcher believes that there may be negative association between the quantity of fertiliser used and the percentage of the population who live in rural areas in different countries. The data below show the percentage of the population who live in rural areas and the fertiliser use measured in kg per hectare, for a random sample of 11 countries.

Percentage of population	33	6	58	35	81	69	61	7	74	71	17
Fertiliser use	76	44	6	68	3	10	7	176	5	137	157

(i) Draw a scatter diagram to illustrate the data. **[3]**

(ii) Explain why it might not be valid to carry out a test based on the product moment correlation coefficient in this case. **[2]**

(iii) Calculate the value of Spearman's rank correlation coefficient. **[5]**

(iv) Carry out a hypothesis test at the 1% significance level to investigate the researcher's belief. **[6]**

(v) Explain the meaning of '1% significance level'. **[1]**

(vi) In order to carry out a test based on Spearman's rank correlation coefficient, what modelling assumptions, if any, are required about the underlying distribution? **[1]**

2 When a genetic sequence of plant DNA is given a dose of radiation, some of the genes may mutate. The probability that a gene mutates is 0.012. Mutations occur randomly and independently.

(i) Explain the meanings of the terms 'randomly' and 'independently' in this context. **[2]**

A short stretch of DNA containing 20 genes is given a dose of radiation.

(ii) Find the probability that exactly 1 out of the 20 genes mutates. **[2]**

A longer stretch of DNA containing 500 genes is given a dose of radiation.

(iii) Explain why a Poisson distribution is an appropriate approximating distribution for the number of genes that mutate. **[2]**

(iv) Use this Poisson distribution to find the probability that there are

(*A*) exactly two genes that mutate, **[3]**

(*B*) at least two genes that mutate. **[2]**

A third stretch of DNA containing 50 000 genes is given a dose of radiation.

(v) Use a suitable approximating distribution to find the probability that there are at least 650 genes that mutate. **[5]**

3 Many types of computer have cooling fans. The random variable X represents the lifetime in hours of a particular model of cooling fan. X is Normally distributed with mean 50 600 and standard deviation 3400.

(i) Find $P(50\,000 < X < 55\,000)$. **[3]**

(ii) The manufacturers claim that at least 95% of these fans last longer than 45 000 hours. Is this claim valid? **[3]**

(iii) Find the value of h for which 99.9% of these fans last h hours or more. **[3]**

(iv) The random variable Y represents the lifetime in hours of a different model of cooling fan. Y is Normally distributed with mean μ and standard deviation σ. It is known that $P(Y < 60\,000) = 0.6$ and $P(Y > 50\,000) = 0.9$. Find the values of μ and σ. **[5]**

(v) Sketch the distributions of lifetimes for both types of cooling fan on a single diagram. **[4]**

4 **(a)** A random sample of 80 GCSE students was selected to take part in an investigation into whether attitudes to mathematics differ between girls and boys. The students were asked if they agreed with the statement 'Mathematics is one of my favourite subjects'. They were given three options 'Agree', 'Disagree', 'Neither agree nor disagree'. The results, classified according to sex, are summarised in the table below.

	Agree	Disagree	Neither
Male	17	13	8
Female	12	11	19

The contributions to the test statistic for the usual χ^2 test are shown in the table below.

	Agree	Disagree	Neither
Male	0.7550	0.2246	1.8153
Female	0.6831	0.2032	1.6424

(i) Calculate the expected frequency for females who agree. Verify the corresponding contribution, 0.6831, to the test statistic. **[3]**

(ii) Carry out the test at the 5% level of significance. **[6]**

(b) The level of radioactivity in limpets (a type of shellfish) in the sea near to a nuclear power station is regularly monitored. Over a period of years it has been found that the level (measured in suitable units) is Normally distributed with mean 5.64. Following an incident at the power station, a researcher suspects that the mean level of radioactivity in limpets may have increased. The researcher selects a random sample of 60 limpets. Their levels of radioactivity, x (measured in the same units), are summarised as follows.

$$\sum x = 373 \qquad \sum x^2 = 2498$$

Carry out a test at the 5% significance level to investigate the researcher's belief. **[11]**

END OF QUESTION PAPER

OCR
Oxford Cambridge and RSA

Wednesday 15 June 2016 – 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

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 inside 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**. If additional space is required, you should use the lined page(s) at the end of this booklet. The question number(s) must be clearly shown.
- 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.

Turn over

1 (i)

1 (ii)

1 (iii)	
1 (iv)	

1 (v)	
1 (vi)	
2 (i)	
2 (ii)	
2 (iii)	

2 (iv) (*A*)	
2 (iv) (*B*)	
2 (v)	

3 (i)	
3 (ii)	
3 (iii)	

3 (iv)	
3 (v)	

4 (a) (i)	
4 (a) (ii)	

4 (b)	

ADDITIONAL ANSWER SPACE

If additional space is required, you should use the following lined page(s). The question number(s) must be clearly shown in the margin(s).

OCR
Oxford Cambridge and RSA

GCE

Mathematics (MEI)

Unit **4767**: Statistics 2

Advanced GCE

Mark Scheme for June 2016

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 ✗	
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	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
ito	In terms of

2. 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.

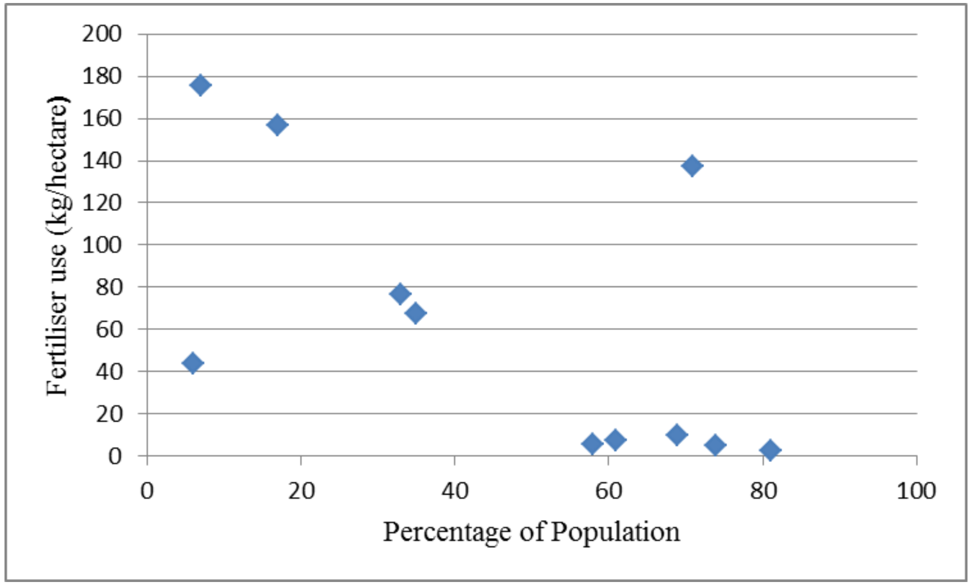
A misread should normally attract *some* penalty, though this would often be only 1 mark and should rarely if ever be more than 2. Commonly in sections of questions where there is a numerical answer either at the end of the section or to be obtained and commented on (eg the value of a test statistic), 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)			G1 G2,1,0 [3]	For suitably labelled axes. Condone absence of scale here. G2 for 11 points correctly plotted relative to a suitable **linear** scale. G1 if 9 or 10 correctly plotted. G0 if 3 or more incorrectly plotted/omitted Allow axes interchanged
1	(ii)		(The points in the scatter diagram) do not appear to be roughly elliptical. The **population** may not have a bivariate Normal distribution.	E1 E1 [2]	For "not elliptical". For not **underlying** bivariate Normal. Do not allow "the data" in place of population/underlying. Allow "data is not **from** a bivariate Normal distribution". Do not allow Normal bivariate ….
1	(iii)		(see table below)	M1 M1	For ranking (allow ranks reversed) **NB No ranking scores 0/5** For d^2

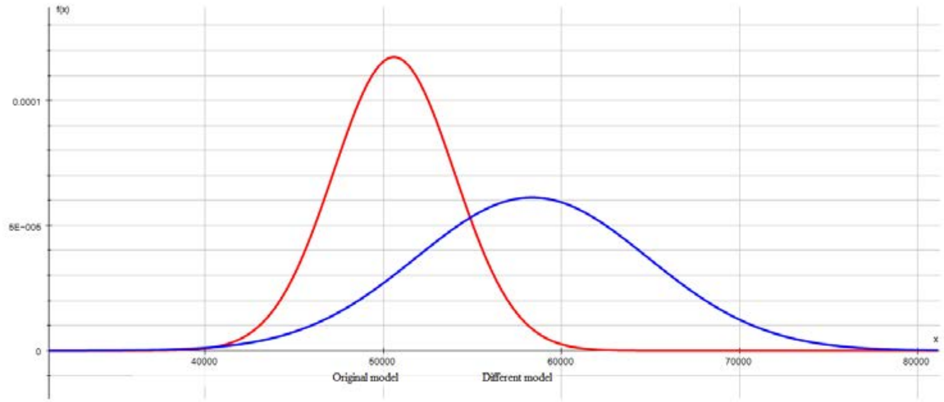
Percentage	33	6	58	35	81	69	61	7	74	71	17
Fertiliser use	76	44	6	68	3	10	7	176	5	137	157
Rank percentage	4	1	6	5	11	8	7	2	10	9	3
Rank Fertiliser	8	6	3	7	1	5	4	11	2	9	10
d	4	5	-3	2	-10	-3	-3	9	-8	0	7
d^2	16	25	9	4	100	9	9	81	64	0	49

Question			Answer	Marks	Guidance
			$\Sigma d^2 = 366$	A1	For Σd^2 (May be embedded in the calculation)
			$r_s = 1 - \dfrac{6\sum d^2}{n(n^2-1)} = 1 - \dfrac{6\times 366}{11\times 120} = 1 - \dfrac{2196}{1320} = 1 - 1.6636$	M1	For method for r_s
			= - 0.664 (to 3 s.f.) [*allow* -0.66 to 2 s.f. or -73/110]	A1 **[5]**	FT their Σd^2 provided $-1 < r_s < 0$¸ and ranking used. **NB No ranking scores 0/5**
1	**(iv)**		H_0: no association between percentage of population living in rural areas and fertiliser use (in the population of countries)	B1	For null hypothesis in context **NB** H_0 H_1 <u>not</u> ito ρ.
			H_1: **negative** association between percentage of population living in rural areas and fertiliser use (in the population of countries)	B1	For alternative hypothesis in context. Context needed in at least one of the hypotheses.
				B1	For **population of countries** or **underlying population**.
			One tail test critical value at 1% level is -0.7091	B1	For ±0.7091 **No further marks from here if incorrect**.
			Since - 0.664 > -0.7091 [or 0.664 < 0.7901] there is …	M1	For sensible comparison of their "- 0.664" with ±0.7091 **seen**, leading to conclusion, only if $-1 <$ their $r_s < 0$.
			…insufficient evidence to reject H_0. There is insufficient evidence to suggest that there is **negative** association between percentage of population living in rural areas and fertiliser use (in the population of countries)	A1 **[6]**	for not significant, oe, and correct conclusion in context. FT their r_s with correct cv.

Question			Answer	Marks	Guidance	
1	(v)		It means that the probability of rejecting H_0 given that it is correct is 1% o.e.	E1 **[1]**	Allow "the probability of a false positive is 1%", "the probability of a Type I Error is 1%". Do not allow "It means that the probability rejecting H_0 when it should have been accepted is 1%"	
1	(vi)		None	E1 **[1]**		
2	(i)		'Randomly' means that **mutations** occur with no (predictable) pattern. 'Independently' means that the occurrence of one **mutation** does not affect the **probability** of another mutation occurring.	E1 E1 **[2]**	In context. Allow "not predictable" Must include 'probability' and context. Allow "chance". If not indicated, assume first comment relates to randomness.	
2	(ii)		P(Exactly one) = $\binom{30}{29} \times 0.85^{29} \times 0.15^1$ $^{20}C_1 \times 0.012^1 \times 0.988^{19}$ = 0.1908	M1 A1 **[2]**	For correct structure i.e. $20p(1-p)^{19}$ Allow 0.191. Allow 0.19 www.	
2	(iii)		Because the number of mutating genes/X is binomially distributed n is large and p is small.	E1 E1 **[2]**	Allow B(500, 0.012) or B(n, p). Allow the sample is large & $np \approx np(1-p)$ or np not too large. Condone suitable numerical ranges – e.g. $n > 30$, $p < 0.1$ Do not allow "the number is large and probability is small". Allow "probability of success/a gene mutating is small" for p is small	
2	(iv)	(*A*)	$\lambda = 500 \times 0.012 = 6$ P(2 mutations) = $e^{-0.85}\frac{0.85^1}{1!}$ [0.0620 – 0.0174 from tables] = 0.0446	B1 M1 A1 **[3]**	For mean Correct structure for P(= 2) using Poisson pdf or tables. CAO Allow 0.04462 or 0.045www	

Question			Answer	Marks	Guidance
2	(iv)	(*B*)	From tables P(At least two) = 1 – P(≤1) = 1 – 0.0174 = 0.9826	 M1 A1 **[2]**	 For using 1 – P(≤ 1) using their mean. CAO Allow 0.983. Allow 0.98 www.
2	(v)		Mean 50000 × 0.012 = 600 , Var = 50000 × 0.012 × 0.988 = 592.8 Using Normal approx. to the binomial, $X \sim N(600, 592.8)$ $P(X \geq 650) = P\left(Z \leq \frac{30.5-25.5}{\sqrt{25.5}}\right)$ $= P(Z > 2.033) = 1 - \Phi(2.033) = 1 - 0.9789$ = 0.0211	 B1 B1 B1 M1 A1 **[5]**	 For Normal approximation (SOI). For correct parameters (SOI). For 649.5 For standardisation and probability calculation using correct tail. CAO (Allow answer from calculator 0.0210)
2	(v)		**Alternative solution using Normal approx. to Poisson** Mean 100 × 6 = 600 Using Normal approx. to the Poisson, $X \sim N(600, 600)$ $P(X \geq 650) = P\left(Z \leq \frac{30.5-25.5}{\sqrt{25.5}}\right)$ $= P(Z > 2.021) = 1 - \Phi(2.021) = 1 - 0.9783$ = 0.0217	 B1 B1 B1 M1 A1 **[5]**	 For Normal approximation (SOI). For correct parameters (SOI). For 649.5 For standardisation and probability calculation using correct tail. CAO (Allow answer from calculator 0.0216)
3	(i)		$P(50000 < X < 55000) =$ $P\left(Z \geq \frac{750-751.4}{2.5}\right)\left(\frac{50000-50600}{3400} < Z < \frac{55000-50600}{3400}\right)$ $= P(-0.176 < Z < 1.294) = \Phi(1.294) - (1 - \Phi(0.176)) = 0.9022 - 1 + 0.5699$ = 0.4721	M1 M1 A1 **[3]**	For standardising both. SOI. Penalise erroneous continuity corrections and wrong sd. Condone numerator(s) reversed. For correct structure Φ(positive *z*) - Φ(negative *z*) CAO including use of difference tables (Answer from calculator 0.4722 and from tables interpolated 0.4723)

Question			Answer	Marks	Guidance	
3	(ii)		P(X > 45000) = P$\left(Z \geq \frac{750-751.4}{2.5}\right)$$\left(Z > \frac{45000-50600}{3400}\right)$ = P(Z > -1.647)	B1*	For -1.647 or -Φ^{-1}(0.95) = -1.645 or 1.647 seen with P(X < 56200) or numerator reversed	
			= Φ(1.647) = 0.9502	B1*	For 0.9502 or 45007 or 0.0498, or B1 for -1.645 if B1 for -1.647 already awarded.	
			0.9502 > 95% so agree with claim	depE1* [3]	**For comparison seen** e.g. -1.647 < -1.645 or 0.0498 < 0.05 or 1.647 > 1.645 or 95% last longer than 45007 hours, **and correct conclusion**. Dependent on B1, B1 awarded	
3	(iii)		From tables Φ^{-1} (0.999) = 3.09	B1	±3.09 seen	
			$\frac{h-50600}{3400} = -3.09$	M1	For equation as seen with their negative z-value	
			$k = 50600 - (3.09 \times 3400)$ = 40100 www	A1 [3]	CAO Allow 40094, 40090	

Question		Answer	Marks	Guidance
3	(iv)	$P(Y < 60000) = 0.6 \Rightarrow P\left(Z < \frac{60000-\mu}{\sigma}\right) = 0.6$ $\Rightarrow \frac{60000-\mu}{\sigma} = \Phi^{-1}(0.6) = 0.2533$ $\Rightarrow 60000 = \mu + 0.2533\sigma$ $P(Y > 50000) = 0.9 \Rightarrow P\left(Z > \frac{50000-\mu}{\sigma}\right) = 0.9$ $\Rightarrow \frac{50000-\mu}{\sigma} = \Phi^{-1}(0.1) = -1.282$ $\Rightarrow 50000 = \mu - 1.282\sigma$ $1.5353\sigma = 10000$ $\sigma = 6513$ $\Rightarrow \mu = 50000 + (1.282 \times 6513) = 58350$	B1 M1 A1 A1 A1 **[5]**	For ±0.2533 or ±1.282 seen For an equation ito μ, σ, *z* and *y* formed. NB using *z* = ±0.2533 with *y* = 60 000 or ±1.282 with *y* = 50 000 For two correct equations seen. CAO Allow 6510, 6515 CAO Allow 58400
3	(v)		G1 G1 G1 G1 **[4]**	For two Normal shapes including attempt at asymptotic behaviour with horizontal axis at each of the four ends. Penalise clear asymmetry. For means, shown explicitly or by scale **on a single diagram.** If shown explicitly, the positions must be consistent with horizontal scale if present. FT part (iv). For greater width (variance) for Different model. FT part (iv). For lower max height for Different model. FT part (iv) If not labelled assume the larger mean represents Different model. FT part(iv).

Question			Answer	Marks	Guidance	
4	**(a)**	**(i)**	Expected frequency = 42/80 × 29 = 15.225 Contribution = $(12 – 15.225)^2$ / 15.225 (= 0.6831 AG)	B1 M1 A1 **[3]**	for 15.225 For valid attempt at $(O\text{-}E)^2/E$ leading to correct answer. ***NB Answer given***	
4		**(ii)**	H_0: no association between sex and attitude to Mathematics. H_1: some association between sex and attitude to Mathematics. Test statistic X^2 = 5.3236 Refer to χ_2^2 Critical value at 5% level = 5.991	B1 B1 B1 B1	For correct hypotheses in context (with context seen in at least one hypothesis). NB if H_0 H_1 reversed do not award first B1or final A1. Allow hypotheses expressed in terms of independence, and in context. Allow 5.324 or 5.32 Allow “2 degrees of freedom” or ν = 2 seen. No further marks from here if wrong or omitted.	
			(5.3236 < 5.991 so result is) not significant There is insufficient evidence **to suggest** that there is association <u>between sex and attitude to Mathematics</u>	M1 A1 **[6]**	For not significant oe. FT their test statistic. Allow ‘Accept H_0’ or ‘Reject H_1’ For **non-assertive** conclusion <u>in context</u> FT their test statistic. Do not allow “relationship” or “correlation” for “association”.	

Question			Answer	Marks	Guidance
4	(b)		$\bar{x} = 373/60 = 6.217$	B1	Allow 6.22
			$s = \sqrt{\dfrac{2498-(373)^2/60}{59}} = \sqrt{\dfrac{179.183}{59}}$	M1	For correctly structured calculation (divisor = 59) for the sample standard deviation or variance.
			$= \sqrt{3.0370} = 1.743$	A1	Allow answers which round to 1.74
			H_0: $\mu = 5.64$; H_1 : $\mu > 5.64$	B1	For both hypotheses correct.
			Where μ denotes the **mean** radioactivity level in (the population of) limpets	B1	For definition of μ in context. Do not allow other symbols unless clearly defined as population mean.
			Test statistic $= \dfrac{6.217-5.64}{1.743/\sqrt{60}} = \dfrac{0.5767}{0.2250}$ $\dfrac{4.995-5.0}{0.0072/\sqrt{8}} = -\dfrac{0.005}{0.002546} = -1.964$	M1*	Structure of test statistic using their sd and mean. Must include correct use of √60. Do not condone numerator reversed.
			= 2.563	A1	Allow answers between 2.56 and 2.57 inclusive.
			Upper 5% level 1 tailed critical value of z = 1.645	B1	For 1.645 No further marks from here if wrong.
			2.563 > 1.645 The result is…	depM1*	For sensible comparison leading to a conclusion (even if incorrect). FT their test statistic.
			…significant. There is sufficient evidence to reject H_0	A1	Correct conclusion. FT their test statistic.
			There is sufficient evidence **to suggest** that the **mean** level of radioactivity has increased.	A1 **[11]**	For correct **non-assertive** conclusion in words in context. FT their test statistic.

Additional Notes on Sensible Comparisons

e.g. In Q4 (b) Neither 2.563 > 0.05 nor 0.0052 < 2.326 are considered sensible as each compares a z-value with a probability.
- 2.563 < 1.645 is not considered to be sensible.
For 2.563 < 1.645 leading to a conclusion, award M0 A0.

Additional Notes on Conclusions to Hypothesis Tests

The following are examples of conclusions which are considered too assertive.

There is sufficient evidence to reject H_0 and **conclude** that…
"there is a positive association between…" or
"there seems to be evidence that there is a positive association between…" or
"the mean level of radioactivity is greater …."
"there doesn't appear to be association between…"

Also note that final conclusions **must refer to H_1 in context** for the final mark to be given.
e.g. In Q4 (a) part (ii), a conclusion just stating that "there is insufficient evidence to suggest that there is an association" gets A0 as this does not refer to the context.

Additional Notes on Alternative Methods in Q4 (b)

Critical value method	cv = 5.64 + 1.645 × 1.743÷ √60	gets M1* (structure) FT their sd. B1 for 1.645 used (otherwise B0M0A0A0)
	= 6.01	gets A1 (replacing the A1 for 2.563)
	6.217 > 6.01	gets depM1* if a conclusion is made, FT their mean only if 1.645 used. Then A1, A1 available as before.
Probability Method	P(Z > 2.563) = 0.0052	gets B1 for value rounding to 0.005 which replaces the B1 for 1.645 (otherwise B0depM0*A0A0).
	0.005 < 0.05	gets depM1* if a conclusion is made only if B1for 0.005 has been awarded. Then A1, A1 available as before.
		NOTE Condone B1 for 0.995 obtained from P(Z < 2.563) only if compared with 0.95 at which point the final depM1*A1A1are available. B0depM0*A0A0 if 0.995 obtained from P(z > -2.563).

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

4767 Statistics 2

General Comments:

The vast majority of candidates appeared to be well prepared for this examination. The overall performance was very good and the average score is once again very high. In hypothesis tests, most candidates provided appropriately worded hypotheses and conclusions. Most candidates were able to complete required calculations correctly and with suitable working provided. Over-specified answers were present though many candidates managed to choose suitable degrees of accuracy for their final answers. Most of the candidates with access to more advanced calculators managed to provide sufficient detail in their solutions to be awarded full credit.

Comments on Individual Questions:

Question No. 1

(i) The scatter diagram was well drawn by many, with some choosing more manageable scales than others.

(ii) The key points concerning the absence of any discernible elliptical shape and the corresponding questioning of the underlying bivariate Normal population were handled well by many. Though most candidates managed to comment on the lack of an elliptical spread of points, often with poor spelling of ‘ellipse’, there were still many candidates who struggled to differentiate between data and population.

(iii) Many achieved full marks here. Only a few candidates reversed the ranking of one of the sets of values. Errors tended to involve mistakes in ranking, in adding the squares of the differences or in rounding the final answer. Very few candidates failed to rank the data.

(iv) This part proved to be challenging for many candidates, especially in the identification of the underlying population of countries involved. Inappropriately worded hypotheses, using ‘correlation’ in place of ‘association’, were seen in a few cases. Many realised the one-sided nature of the test which investigated the possible negative association and went on to use the appropriate critical value to obtain a suitable conclusion in context. Few candidates expressed the conclusion in terms of the null hypothesis.

(v) There were many correct answers to this part of the question, either in terms of incorrectly rejecting H_0 or one of the equivalents. Many answers referring to ‘accuracy’ or ‘reliability’ were seen.

(vi) Few candidates appeared to understand that for this test no modelling assumptions about the underlying distribution are required.

Question No. 2

This proved to be a very straightforward question with most candidates scoring high marks.

(i) Many candidates succeeded in defining ‘random’ and ‘independent’ but many failed to define independence in terms of probability.

(ii) Well answered. A minority of candidates used a Poisson calculation here.

(iii) Though most candidates identified the usual ‘*n* is large and *p* is small’, very few explicitly related these values to the binomial distribution.

(iv) (*A*) Well answered.

(*B*) Well answered.

(v) Well answered. Common errors tended to involve either use of an incorrect standard deviation or omission of the required continuity correction.

Question No. 3

There were many good responses to this question. Spurious continuity corrections were rarely seen. It helps when candidates provide sketches for questions involving the Normal distribution.

(i) Well answered. Errors caused by lack of accuracy reading Normal tables were seen fairly regularly. Most candidates used the correct probability structure with their z values.

(ii) This was well done on the whole though many candidates did not provide the required comparison to justify their conclusion. In most cases the working provided was clear – diagrams were helpful to examiners in conveying the candidates' intentions – often more successfully than their wording.

(iii) Many correctly identified the z value of -3.09 and went on to find the appropriate value for h, rounded to a suitable level of accuracy.

(iv) There were some pleasing attempts at this question, marred only by an inappropriate degree of accuracy for the final answers. It was good to see that once the initial equations had been established with the correct z values, many could still solve the simultaneous equations. A few candidates failed to identify, for the given probabilities, the z values needed for the simultaneous equations.

(v) This was answered well, though many candidates could have made a greater effort to include symmetry in their sketches and to pay more attention to the asymptotic nature. Spurious labelling of axes was seen but only rarely.

Question No. 4

Most candidates scored well on this question. Pleasingly, overly-assertive conclusions to the hypotheses tests were rarely seen.

(a) (i) Most candidates found this to be very easy.

(ii) Well answered. Some candidates failed to word their hypotheses and conclusion in terms of 'association'. Most stated the correct number of degrees of freedom and critical value. Most candidates were able to finish off with appropriately worded conclusions.

(b) With a little more background work to be done, finding sample mean and standard deviation, candidates found this part of the question more difficult than part (a). The calculation of sample standard deviation caused problems for many. Issues with premature rounding of sample mean and standard deviation, leading to inaccuracy in the calculation of the test statistic, were quite common. A small number of candidates did not express their hypotheses in terms of μ. Definitions of μ as 'sample mean' were, thankfully, rare. Given that most candidates provided the correct alternative hypothesis it was disappointing to see many working with a negative test statistic and critical value – these were deemed inappropriate and thus penalised. It was again pleasing to see candidates taking care to word conclusions in an appropriately non-assertive manner.

Oxford Cambridge and RSA

GCE Mathematics (MEI)

				Max Mark	a	b	c	d	e	u
4751	01	C1 – MEI Introduction to advanced mathematics (AS)	Raw	72	63	57	52	47	42	0
			UMS	100	80	70	60	50	40	0
4752	01	C2 – MEI Concepts for advanced mathematics (AS)	Raw	72	56	49	42	35	29	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	47	42	36	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
			UMS	100	80	70	60	50	40	0
4754	01	C4 – MEI Applications of advanced mathematics (A2)	Raw	90	64	57	51	45	39	0
			UMS	100	80	70	60	50	40	0
4755	01	FP1 – MEI Further concepts for advanced mathematics (AS)	Raw	72	59	53	48	43	38	0
			UMS	100	80	70	60	50	40	0
4756	01	FP2 – MEI Further methods for advanced mathematics (A2)	Raw	72	60	54	48	43	38	0
			UMS	100	80	70	60	50	40	0
4757	01	FP3 – MEI Further applications of advanced mathematics (A2)	Raw	72	60	54	49	44	39	0
			UMS	100	80	70	60	50	40	0
4758	01	(DE) MEI Differential Equations with Coursework: Written Paper	Raw	72	67	61	55	49	43	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
			UMS	100	80	70	60	50	40	0
4761	01	M1 – MEI Mechanics 1 (AS)	Raw	72	58	50	43	36	29	0
			UMS	100	80	70	60	50	40	0
4762	01	M2 – MEI Mechanics 2 (A2)	Raw	72	59	53	47	41	36	0
			UMS	100	80	70	60	50	40	0
4763	01	M3 – MEI Mechanics 3 (A2)	Raw	72	60	53	46	40	34	0
			UMS	100	80	70	60	50	40	0
4764	01	M4 – MEI Mechanics 4 (A2)	Raw	72	55	48	41	34	27	0
			UMS	100	80	70	60	50	40	0
4766	01	S1 – MEI Statistics 1 (AS)	Raw	72	59	52	46	40	34	0
			UMS	100	80	70	60	50	40	0
4767	01	S2 – MEI Statistics 2 (A2)	Raw	72	60	55	50	45	40	0
			UMS	100	80	70	60	50	40	0
4768	01	S3 – MEI Statistics 3 (A2)	Raw	72	60	54	48	42	37	0
			UMS	100	80	70	60	50	40	0
4769	01	S4 – MEI Statistics 4 (A2)	Raw	72	56	49	42	35	28	0
			UMS	100	80	70	60	50	40	0
4771	01	D1 – MEI Decision mathematics 1 (AS)	Raw	72	48	43	38	34	30	0
			UMS	100	80	70	60	50	40	0
4772	01	D2 – MEI Decision mathematics 2 (A2)	Raw	72	55	50	45	40	36	0
			UMS	100	80	70	60	50	40	0
4773	01	DC – MEI Decision mathematics computation (A2)	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	55	49	44	39	33	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
			UMS	100	80	70	60	50	40	0
4777	01	NC – MEI Numerical computation (A2)	Raw	72	55	47	39	32	25	0
			UMS	100	80	70	60	50	40	0
4798	01	FPT - Further pure mathematics with technology (A2)	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 Statistics 1 MEI (Z1)	Raw	72	59	52	46	40	34	0
		UMS	100	80	70	60	50	40	0
G242	01 Statistics 2 MEI (Z2)	Raw	72	55	48	41	34	27	0
		UMS	100	80	70	60	50	40	0
G243	01 Statistics 3 MEI (Z3)	Raw	72	56	48	41	34	27	0
		UMS	100	80	70	60	50	40	0

GCE Quantitative Methods (MEI)

			Max Mark	a	b	c	d	e	u
G244	01 Introduction to Quantitative Methods MEI	Raw	72	58	50	43	36	28	0
G244	02 Introduction to Quantitative Methods MEI	Raw	18	14	12	10	8	7	0
		UMS	100	80	70	60	50	40	0
G245	01 Statistics 1 MEI	Raw	72	59	52	46	40	34	0
		UMS	100	80	70	60	50	40	0
G246	01 Decision 1 MEI	Raw	72	48	43	38	34	30	0
		UMS	100	80	70	60	50	40	0

Level 3 Certificate and FSMQ raw mark grade boundaries June 2016 series

For more information about results and grade calculations, see www.ocr.org.uk/ocr-for/learners-and-parents/getting-your-results

Level 3 Certificate Mathematics for Engineering

			Max Mark	a*	a	b	c	d	e	u
H860	01 Mathematics for Engineering		This unit has no entries in June 2016							
H860	02 Mathematics for Engineering									

Level 3 Certificate Mathematical Techniques and Applications for Engineers

			Max Mark	a*	a	b	c	d	e	u
H865	01 Component 1	Raw	60	48	42	36	30	24	18	0

Level 3 Certificate Mathematics - Quantitative Reasoning (MEI) (GQ Reform)

			Max Mark	a	b	c	d	e	u
H866	01 Introduction to quantitative reasoning	Raw	72	55	47	39	31	23	0
H866	02 Critical maths	Raw	60	47	41	35	29	23	0
		Overall	132	111	96	81	66	51	0

Level 3 Certificate Mathematics - Quantitive Problem Solving (MEI) (GQ Reform)

			Max Mark	a	b	c	d	e	u
H867	01 Introduction to quantitative reasoning	Raw	72	55	47	39	31	23	0
H867	02 Statistical problem solving	Raw	60	40	34	28	23	18	0
		Overall	132	103	88	73	59	45	0

Advanced Free Standing Mathematics Qualification (FSMQ)

			Max Mark	a	b	c	d	e	u
6993	01 Additional Mathematics	Raw	100	59	51	44	37	30	0

Intermediate Free Standing Mathematics Qualification (FSMQ)

			Max Mark	a	b	c	d	e	u
6989	01 Foundations of Advanced Mathematics (MEI)	Raw	40	35	30	25	20	16	0

Version	Details of change
1.1	Correction to Overall grade boundaries for H866
	Correction to Overall grade boundaries for H867