OCR
Oxford Cambridge and RSA

Wednesday 14 June 2017 – 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

6710721633

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 barcodes.
- 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

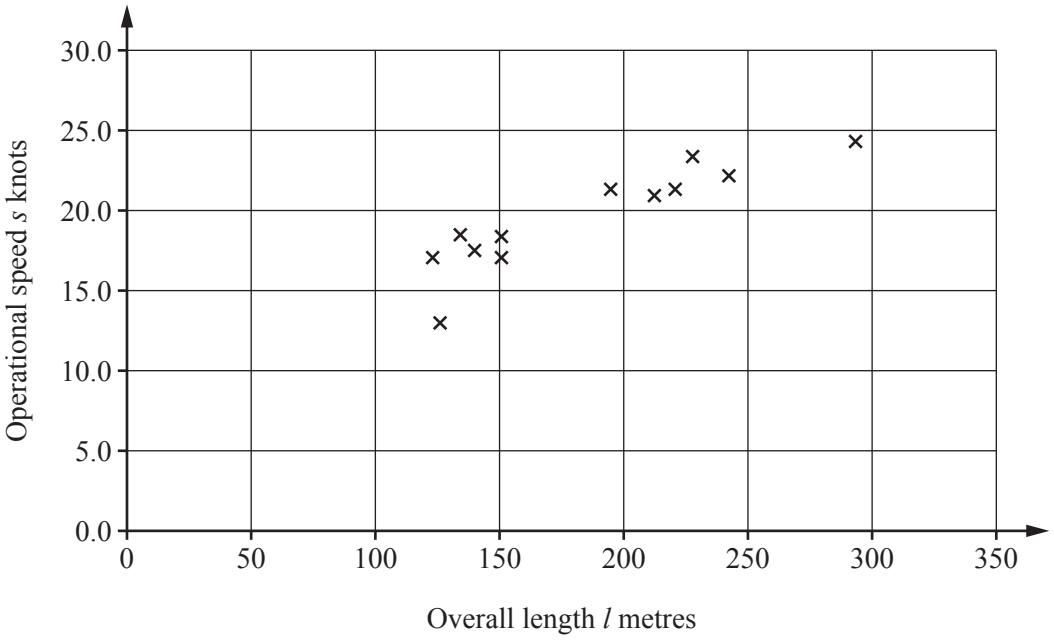
This information is the same on the Printed Answer Book and the Question Paper.

- 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 **8** 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 The scatter diagram below illustrates the overall lengths l metres and the typical operational speeds s knots (nautical miles per hour) of 12 container ships. The length of a ship is one of the factors which determines its typical operational speed.



Summary statistics for these data are as follows:

$n = 12, \quad \Sigma l = 2219, \quad \Sigma s = 234.6, \quad \Sigma l^2 = 443\,867, \quad \Sigma s^2 = 4700.56, \quad \Sigma ls = 45\,149.0.$

(i) State which of the two variables l and s is the independent variable and which is the dependent variable. Briefly explain your answers. **[2]**

(ii) Calculate the equation of the regression line of s on l. **[5]**

(iii) Interpret the coefficient of l in terms of the relationship between overall length and speed in the equation of the regression line found in part **(ii)**. **[2]**

(iv) Calculate the value of the residual for the data point where $l = 126$ and $s = 13.0$. **[3]**

(v) Use the equation of the regression line in part **(ii)** to calculate an estimate of the operational speed of a ship with overall length of 100 metres. Comment on the reliability of this estimate. **[2]**

If the data point in part **(iv)** is removed from the data set, the equation of the new regression line is

$$s = 0.0453l + 11.5.$$

(vi) Recalculate the estimate for an overall length of 100 metres using this new equation. Discuss which of these two estimates you think is likely to be more reliable. **[3]**

2 At a small hospital maternity department, there is an average of 1.3 births per day. Throughout this question, you should assume that births occur independently, at random times, and at a uniform average rate.

(i) Briefly explain the meaning of each of the three terms 'independently', 'random' and 'uniform', in the context of births at the maternity department. **[3]**

X represents the number of births at the hospital on a randomly chosen day.

(ii) State the distribution of X and also the variance of X. **[2]**

(iii) Find $P(X > 3)$. **[2]**

(iv) Find the probability that there are exactly 3 births in a period of 3 days. **[2]**

There is an average of 0.4 home births per **week** in the area served by the hospital.

(v) Find the probability that the total number of births in a week (at home and in hospital) is at least 10. **[2]**

(vi) Use a suitable approximating distribution to find the probability that there is a total of at least 50 births in a period of 4 weeks. **[5]**

(vii) How realistic do you think the assumption of independence made at the start of this question is? **[2]**

3 The random variable X represents the weight, in grams, of a particular type of chocolate bar. It is known that X is Normally distributed with mean 50.7 and variance 0.72. On the wrapper it states that the bar weighs 50 grams.

(i) Find the proportion of these chocolate bars that actually weigh at least 50 grams. **[3]**

(ii) A quality control manager wishes to increase this proportion to 95%.

(A) Find the required value of the mean if the variance remains unchanged. **[4]**

(B) Find the required value of the variance if the mean remains unchanged. **[2]**

The weights of another type of chocolate bar are also Normally distributed. On the wrapper it states that the bar weighs 25 grams. It is known that 99% of these bars weigh at least 25.0 grams and 75% of them weigh at least 25.4 grams.

(iii) Find the probability that one of these bars weighs at least 26.0 grams. **[6]**

(iv) One bar of the first type (with the original mean and standard deviation) and 2 bars of the second type are selected at random. Find the probability that at least one of the bars has a weight less than that stated on its wrapper. **[2]**

4 **(a)** In an investigation into dietary supplements, a random sample of 200 adults was selected. Each of them was asked whether or not they regularly take dietary supplements. The 200 adults were categorised as 'Male under 50', 'Male 50 or older', 'Female under 50', 'Female 50 or older'. Their answers to the question are summarised in the table below.

	Yes	No
Male under 50	13	33
Male 50 or older	18	31
Female under 50	24	37
Female 50 or older	24	20

(i) Write down null and alternative hypotheses for a test to examine whether there is any association between category of adult and whether or not they regularly take dietary supplements. **[1]**

The expected frequencies under the null hypothesis for the usual χ^2 test are shown in the table below.

Expected frequency	Yes	No
Male under 50	18.17	27.83
Male 50 or older	19.36	29.65
Female under 50	24.10	36.91
Female 50 or older	17.38	26.62

(ii) Verify the expected frequency of 18.17 for Male under 50 answering Yes. **[2]**

The contributions to the test statistic are shown in the table below.

Contribution	Yes	No
Male under 50	1.4710	0.9604
Male 50 or older	0.0949	0.0619
Female under 50	0.0004	0.0002
Female 50 or older	2.5215	1.6463

(iii) Verify the contribution to the test statistic of 1.4710 for Male under 50 answering Yes. **[2]**

(iv) Given that the total of the contributions is 6.757, correct to 3 decimal places, complete the test at the 10% significance level. **[4]**

(v) For each category of adult, comment briefly on how their taking of dietary supplements compares with what would be expected if there were no association. **[3]**

(b) The breaking strength of a particular type of rope is known to be Normally distributed with mean 562 kg and a standard deviation of 27.4 kg when the rope is dry. A researcher believes that the breaking strength may be different when the rope is wet. He selects a random sample of 12 pieces of wet rope and finds that the mean of their breaking strengths is 547 kg. Carry out a test at the 10% significance level to investigate the researcher's belief. You may assume that the breaking strengths of wet ropes of this type are still Normally distributed with standard deviation 27.4 kg. **[8]**

END OF QUESTION PAPER

BLANK PAGE

BLANK PAGE

BLANK PAGE

OCR
Oxford Cambridge and RSA

Wednesday 14 June 2017 – 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 barcodes.
- 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 **8** pages. Any blank pages are indicated.

1 (i)	
1 (ii)	

1 (iii)	
1 (iv)	
1 (v)	
1 (vi)	

2 (i)

2 (ii)

2 (iii)

2 (iv)

2 (v)

2 (vi)

2 (vii)

3 (i)

3 (ii)(*A*)

(answer space continued on next page)

3 (ii)(*A*)	**(continued)**
3 (ii)(*B*)	

3 (iii)	
3 (iv)	

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

4 (a)(iv)	
4 (a)(v)	

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 2017

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.

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

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)		l independent & s dependent **l is independent** since the values of l are not subject to random variation, **or** values of l are controlled/pre-determined/set/chosen (by the manufacturer/researcher) **or s is dependent** since the values of s are subject to random variation.	B1 B1 **[2]**	relevant comment regarding l **or** s do not accept "….changed"
1	(ii)		$\bar{s}$ = 234.6/12 (=19.55), $\bar{l}$ = 2219/12 (=184.917) $b = \frac{S_{ls}}{S_{ll}} = \frac{45149.0-(2219\times234.6/12)}{443867-2219^2/12} = \frac{1767.55}{33536.9} = 0.05270$ OR $b = \frac{45149.0/12-(19.55\times184.917)}{443867/12-184.917^2} = \frac{147.296}{2794.74} = 0.05270$ hence least squares regression line is: $s - \bar{s} = b(l - \bar{l})$ $\Rightarrow$ $s - 19.55 = 0.05270\,(l - 184.917)$ $\Rightarrow$ $s = 0.0527\,l + 9.80$ (accept $s = 0.05270\,l + 9.804$, $s = 0.05270\,l + 9.805$)	B1 M1 A1 M1 A1 **[5]**	for $\bar{s}$ and $\bar{l}$ seen (or can be implied by correct value of b) for attempt at gradient (b) with correct structure. See additional notes on 'structure'. for 0.0527. Allow 0.053 for equation of line with their $b>0$, $\bar{l}$ & $\bar{s}$ FT for complete equation in terms of s and l. Accept equation in terms of x and y only if defined as length and speed respectively. Allow $s = 0.053\,l + 9.8$ www See additional note RE over-specification
	(iii)		The coefficient of l is the additional speed resulting from an increase of 1 metre in length	B1 B1 **[2]**	for connecting **increase** in l increase in s for relating to **unit** increase in length.
	(iv)		$l = 126 \Rightarrow$ predicted speed = 0.0527 × 126 + 9.80 (= 16.4) Residual = 13.0 – 16.4 = – 3.4 (or – 3.44 or – 3.45)	M1 M1 A1 **[3]**	for prediction FT their equation for a subtraction involving 13.0 and their prediction, either way round SOI. FT only 13.0 – their prediction. See additional note RE over-specification

Question			Answer	Marks	Guidance
	(v)		$0.0527 \times 100 + 9.80 = 15.1$ Might not be reliable as extrapolation	B1 B1 **[2]**	FT their equation. See additional note RE over-specification not reliable and extrapolation oe seen
	(vi)		$0.0453 \times 100 + 11.5 = 16.0$ 3sf The point where $l = 126$ and s = 13.0 may be an error and as such it might be better to use the second line which does not involve it. On the other hand the first model may be better as it uses all the available data.	B1 B1 B1 **[3]**	Allow 16 or 16.03 Allow "outlier" or equivalent for "error" or this point might suggest that a curve might be a better model in which case the first model would be better.
2	(i)		'Independently' means that the occurrence of one birth does not affect the **probability** of another birth occurring. 'Random' means that births occur with no particular **pattern**. 'uniform' means that the **average rate** of births is **constant** or the **average over any given time period** is **constant.**	B1 B1 B1 **[3]**	must be in context and include 'probability' or 'chance' but do not allow "the probability of a birth does not affect the probability of another" must be in context. Allow 'not predictable' do not allow "no particular order" must be in context
2	(ii)		$X \sim$ Poisson (1.3) **Variance** = 1.3	B1 B1 **[2]**	allow X ~ Po(1.3) and X ~ P(1.3) and Poisson with $\lambda = 1.3$ must include 1.3 but do **not** allow Po(1.3,1.3) for variance = 1.3, allow $\sigma^2 = 1.3$, do not allow $\lambda = 1.3$
2	(iii)		From tables $P(X > 3) = 1 - P(X \leq 3)$ $= 1 - 0.9569$ $= 0.0431$	 M1 A1 **[2]**	 Attempting $1 - P(X \leq 3)$ e.g. for 1 – 0.9463 (using $\lambda = 1.4$) or 1 – 0.6248 (using $\lambda = 3.1$) See additional note RE over-specification

Question			Answer	Marks	Guidance
2	(iv)		$\lambda = 3 \times 1.3 = 3.9$ P(3 births) $= \dfrac{e^{-3.9}3.9^3}{3!} = 0.2001$ Or from tables P(3 births) $= 0.4532 - 0.2531 = 0.2001$	B1 B1 **[2]**	for mean For 0.2001 Allow 0.200, 0.20, 0.2 www See additional note RE over-specification
2	(v)		$\lambda = 7 \times 1.3 + 0.4 = 9.5$ From tables P($X \geq 10$) $= 1 - P(X \leq 9) = 1 - 0.5218 = 0.4782$	B1 B1 **[2]**	for mean for 0.4782 or 0.478www or 0.48www See additional note RE over-specification
2	(vi)		Normal approx. to the Poisson, $X \sim N(38, 38)$ P($X \geq 50$) $= P\left(Z \geq \dfrac{49.5-38}{\sqrt{38}}\right)$ $= 1 - 0.9690$ $= 0.0310$	B1 B1 B1 M1 A1 **[5]**	for Normal approximation (SOI) for correct parameters (SOI) continuity correction i.e. 49.5 for correct structure of Normal probability calculation cao **(**Do not FT wrong or omitted CC) (answer from calculator = 0.031052 so accept 0.0310 or 0.0311) Allow 0.031www See additional note RE over-specification
2	(vii)		This assumption is not fully valid as there will be some multiple births but the proportion of multiple births is fairly small so it is not totally unreasonable	B1 B1 **[2]**	e.g. twins, triplets, …
3	(i)		P($X \geq 50$) $= P\left(Z > \dfrac{50-50.7}{\sqrt{0.72}}\right) = P(Z > -0.825)$ $= \Phi(0.825)$ $= 0.7953$	M1 M1 A1 **[3]**	For standardizing. M0 for using "continuity corrections" e.g. 49.5, 49, 51, and/or $\sigma = 0.72$ used. Condone numerator reversed. For correct tail Cao allow 0.795 www See additional note RE over-specification

Question			Answer	Marks	Guidance
3	(ii)	(*A*)	P(weight > 50) = 0.95 so P($Z > \frac{50-\mu}{\sqrt{0.72}}$) = 0.95		
			$\Phi^{-1}(0.95) = -1.645$	B1	For ±1.645
			$\frac{50-\mu}{\sqrt{0.72}} = -1.645$	M1*	For equation for μ as seen or equivalent with their negative *z*-value. See additional note. Allow M1* if "continuity correction" and/or σ = 0.72 used and penalised in part (i). NOTE σ = 0.8485 (allow 0.85 or better)
			$\mu = 50 + 1.645 \times \sqrt{0.72} = 51.395...$	M1dep*	for rearranging to find μ cao allow 51.40
			= 51.4	A1	See additional note RE over-specification
				[4]	
3	(ii)	(*B*)	P(weight > 50) = 0.95 so P($Z > \frac{50-50.7}{\sigma}$) = 0.95		
			$\frac{50-50.7}{\sigma} = -1.645$	M1	for equation as seen or equivalent allow M1 if "continuity correction" has been used and already penalised in part (i) or part(ii)
			$\sigma = \frac{50-50.7}{-1.645} = 0.4255...$		
			Var = $0.4255^2 = 0.181$	A1	for 0.181 or 0.1811 or 0.18www NOTE M0 A0 for 0.181 from $(-0.4255)^2$ See additional note RE over-specification
				[2]	

3	**(iii)**		$P(Y > 25) = 0.99 \Rightarrow P\left(Z > \frac{25-\mu}{\sigma}\right) = 0.99$ $\Rightarrow \frac{25-\mu}{\sigma} = \Phi^{-1}(0.99) = -2.326 \quad \Rightarrow 25 = \mu - 2.326\sigma$ $P(Y > 25.4) = 0.75 \Rightarrow P\left(Z > \frac{25.4-\mu}{\sigma}\right) = 0.75$ $\Rightarrow \frac{25.4-\mu}{\sigma} = \Phi^{-1}(0.75) = -0.6745 \quad \Rightarrow 25.4 = \mu - 0.6745\sigma$ $1.6515\sigma = 0.4 \quad \sigma = 0.2422\ldots$ $\mu = 25 + 2.326 \times 0.2422\ldots \quad \mu = 25.563\ldots$ P(Weighs > 26.0) = $P\left(Z > \frac{26.0-25.563}{0.2422}\right) = 1 - \Phi(1.804) = 1 - 0.9644 = 0.0356$	B1 M1 A1 A1 A1 A1 **[6]**	for ±2.326 or ±0.6745 seen For obtaining two equations in terms of mean, standard deviation and their *z*- values (but not *z* = 0.99 or *z* = 0.75 or e.g.1 – 2.326) in any form equivalent to these. for at least one equation correct cao for answers in the range 0.0345 to 0.036 See additional note RE over-specification
3	**(iv)**		$1 - (0.7953 \times 0.99^2) = 1 - 0.7795$ $= 0.2205$	M1 A1 **[2]**	or equivalent FT their 3(i) allow 0.221www or 0.22www
4	**(a)**	**(i)**	H_0: no association between category of adult and taking dietary supplements. H_1: some association between category of adult and taking dietary supplements.	B1 **[1]**	Hypotheses must refer to 'association' and be in context. Allow hypotheses appropriately worded in terms of independence.
		(ii)	Expected frequency = (46 × 79)/200 = 18.17 AG	M1 A1 **[2]**	attempt at row total × column total/grand total 46, 79 and 200 used correctly **and 18.17 seen** **NB Answer given**

		(iii)	Contribution = $(13 - 18.17)^2 / 18.17$ = 1.4710 AG	M1 A1 **[2]**	for valid attempt at $(O-E)^2/E$ 13 and 18.17 used correctly **and 1.4710** or better **seen** **NB Answer given**
		(iv)	Refer to χ^2_3 Critical value at 10% level = 6.251 (6.757 > 6.251 so result is) significant There is sufficient evidence to **suggest/support** association between category of adult and taking dietary supplements NB if H_0 H_1 reversed do not award first B1 or final A1	B1 B1 M1 A1 **[4]**	for 3 degrees of freedom seen (e.g. in subscript) for 6.251 - **No further marks from here if wrong or omitted** for 'significant' or 'Accept H_1' or 'Reject H_0' seen For **non-assertive** conclusion in context. Do not allow 'relationship' or 'correlation' in place of 'association'
		(v)	**Large** contribution for **males under 50** suggests that there are fewer than expected saying **yes**. **Large** contribution for **females 50 or older** suggests that there are more than expected saying **yes**. **Small** contributions for the **other two groups** show that numbers are much as expected.	B1 B1 B1 **[3]**	or **large** contribution for **males under 50** suggests that there are more than expected saying **no**. NB if both comments are provided they must both be correct for B1 or **large** contribution for **females 50 or older** suggests that there are fewer than expected saying **no**. NB if both comments are provided they must both be correct for B1 Do not accept e.g. a few less/more Special Case – if sizes of contributions are not mentioned but comments are otherwise correct award SC1 Comments about what should have been observed (e.g. there should have been more males under 50 saying yes) get 0/3..

4	**(b)**		H_0: $\mu = 562$ H_1: $\mu \neq 562$	B1	For both correct **NB** for **H_1: $\mu < 562$** award maxB0B1M1*A1B1(for – 1.282)depM0*A0A0
			Where μ denotes the **mean** breaking strength of **wet rope** of this type (in the population)	B1	For definition of μ in context. Do not allow any other symbols unless clearly defined as population mean.
			Test statistic $= \frac{547-562}{27.4/\sqrt{12}}$	M1*	must include $\sqrt{12}$ with numerator as seen
			$= \frac{-15}{7.910} = -1.896$	A1	cao for – 1.896
			Lower 5% level 2 tailed critical value of z = –1.645	B1	For 1.645 or – 1.645. No further marks from here if B0 awarded
			–1.896 < –1.645 so the result is significant.	M1dep*	For sensible comparison leading to a conclusion. Must be –1.645 unless it is clear that absolute values are being used.
			There is sufficient evidence to reject H0	A1	for correct conclusion
			There is sufficient evidence to suggest that the **mean** breaking strength of **wet** rope of this type is different (not equal to 562)	A1	for **non-assertive** conclusion in words in context. A0 for "…mean of wet rope has **changed**"
					FT candidate's test statistic only if both M marks earned
				[8]	See additional notes regarding alternative methods and sensible comparisons.

Additional notes Re Q4(b)

Critical Value Method	c.v. = 562 – 1.645×27.4/√12	gets M1* B1
	= 548.99 or 549.0 or 549	gets A1 (replacing A1 for -1.896)
	547 < 548.99 with a conclusion	gets M1dep* then final A1 A1 still available
	NB if H_1: μ < 562 award maxB0B1M1*A1(for 551.9)B1(for - 1.282 used correctly) depM0*A0A0	
Probability Method	P(Z < –1.896) = 0.0289 or 0.029	gets B1 (replacing B1 for ±1.645)
	0.0289 < 0.05 with conclusion	gets M1dep* then final A1 A1 still available
	NB if H_1: μ < 562 used award maxB0B1M1*A1B1(for 0.029)depM0*A0A0	

Additional Note RE Over-specification

A0 or B0 for final answers given correct to 5sf or more potentially in Q1ii (final A1), Q1iv (final A1), Q1v (first B1), Q2iii(final A1), Q2iv (final B1), Q2v (final B1), Q2vi (final A1), Q3i (A1), Q3iiA (A1), Q3iiB (A1), Q3iii final A1), Q3iv (A1).

NOTE do not penalise over-specification more than twice in any single question or more than 4 times in a paper.

Additional Notes on Correct Structure in Q1(ii)

Equivalent calculations for finding *b* are allowed. For example use of $12S_{ls}/12S_{ll}$ is allowed. However, where these are mixed we award M0. e.g. use of $12S_{ls}/S_{ll}$ would earn M0. For M1 to be awarded, the structure of the calculation must be numerically equivalent to the one provided – NOTE if it is believed that the candidate has made an error in transcription of a number (for example using 2119 instead of 2219) we can allow M1 BOD if the structure is otherwise correct.

Additional Notes for Q3ii*A*

M1* is for forming a suitable equation using their *z*-value but it must be reasonably clear that the value used is a *z*-value – for example we do not allow 0.05 or 0.95 to be treated as *z*-values here. The M1dep* can be awarded if the candidate correctly rearranges their equation to find μ. Hence, use of an incorrect *z*-value could earn max B0M1*M1dep*A0.

If z = +1.645 is used then award B1 only to give 1/4 unless the numerator of the equation is reversed in which case the remaining marks are available.

Additional Notes on Sensible Comparisons

In Q4 (b) Neither – 1.896 < 0.05 nor 0.0289 < 1.645 are considered sensible as each compares a z-value with a probability.
Inequality sign reversed, e.g. – 1.896 > – 1.645, gets M0A0A0.
Comparing a negative with a positive *z*-value, e.g. – 1.896 < 1.645, gets M0A0A0.

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 nicotine content 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 Q4a iv a conclusion simply stating that "the evidence suggests that there is association" gets A0 as this does not refer to the context.

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

4767 Statistics 2

General Comments:

The paper proved to be accessible yet provided suitably challenging questions to enable differentiation between those gaining high marks. Most candidates provided well-structured responses to the hypothesis test questions which included appropriately worded conclusions. Many candidates lost marks through over-specification of final answers. Candidates coped well with probability calculations with which most provided clear unambiguous working. Questions requiring interpretation of results or explanation of terms proved to be challenging for many; with such questions examiners look for concise, statistically focused comments.

Comments on Individual Questions:

Question 1.

(i) Most candidates identified independent and dependent variables correctly. Many commented upon the pre-determined nature of the independent variable whilst far fewer successful comments contained references to random variation.

(ii) There were many completely correct calculations of the regression line *s* on *l,* with the main errors being in over-specification of the values in the final equation or in the use of *x* and *y* instead of *l* and *s*.

(iii) Many candidates commented that an increase in *l* led to an increase in *s*, but far fewer understood that that the coefficient of *l* was the increase in *s* per unit length. A few candidates did not appear to understand the word coefficient, taking it to mean that they needed to rearrange their equation making *l* the subject.

(iv) The correct value of *s* was obtained in many cases. The subsequent subtraction to find the residual was carried out the wrong way round by many candidates leaving them with a positive rather than a negative residual. Over-specification of the final answer was penalised.

(v) Most candidates correctly calculated the required estimate, but many candidates did not understand the difference between interpolation and extrapolation. Over-specification of the estimate was penalised.

(vi) Most candidates managed to recalculate the estimate successfully using the new equation and provide a suitable comment in favour of one of their two estimates. Of the successful attempts, answers relating to the removal of the outlier were seen more frequently than answers relating to using all available information. Few candidates referred to both.

Question 2.

(i) Reasonably well answered though a few candidates neglected to answer in context.

(ii) Well answered. Some candidates neglected to provide a parameter for their distribution of X.

(iii) Well answered. A few candidates mistakenly thought that $P(X \geq 3)$ was equivalent to $1 - P(X \leq 2)$.

(iv) Well answered.

(v) Well answered.

(vi) In general this was well answered. Some candidates used incorrect parameters. Most applied the correct continuity correction with only a few inappropriate or missing ones seen. The structure of the probability calculation was mostly correct.

(vii) Many students commented that the existence of multiple births would make the assumption of independence invalid. Very few candidates who recognised that there could be multiple births commented further on their likelihood and how this might relate to the assumption of independence.

Question 3.

(i) Generally well done with most candidates obtaining the correct answer. A few used the wrong tail of the normal distribution. Few made the mistake of applying a 'continuity correction' of 49.5 or 50.5. The use of 0.72 as standard deviation instead of variance was not uncommon.

(ii)(*A*) Generally well done, although the use of a *z*-value leading to a value of μ below 50 was not uncommon.

(ii)(*B*) Generally well answered. Some candidates used +1.645 in place of –1.645 and were happy to work with a negative standard deviation. A few candidates did not finish off by squaring their value of σ and so missed the final mark.

(iii) Despite the relative complexity of this part of the question, there were many good solutions. The question required z values of -2.326 and -0.6745 and this was interpreted well by successful candidates. Some candidates used wrong signs with their z-values, thus forming equations leading to a negative standard deviation. Premature rounding leading to an inaccurate final answer was fairly common.

(iv) This question proved to be too difficult for the majority of candidates. Of the successful attempts, those using $1 - P(X = 0)$ were more successful than those attempting $P(X = 1) + P(X = 2) + P(X = 3)$.

Question 4.

a(i) Well answered.

a(ii) Well answered.

a(iii) Well answered.

a(iv) Well answered. A few candidates misinterpreted the result as not significant. A small number of candidates did not state the number of degrees of freedom.

a(v) Some successful, concise but sufficiently detailed answers were seen for this part of the question. Many candidates made no reference to contributions to the test statistic or simply quoted values without interpreting their magnitude. Some candidates were unclear as to whether their comments related to "yes" or "no" answers. Some candidates recognised that the small contributions indicate that the results are as expected but then go on to say "slightly more than expected" or "slightly fewer than expected".

(b) Generally well answered. Often one or two marks were lost through incomplete/poorly worded conclusions or incomplete definitions of μ. Some candidates attempted a one-tailed test and were penalised. Some candidates calculated a positive test statistic despite the observed value lying below the mean and did not make it clear why they were doing this; a diagram showing the correct test statistic with its symmetrical equivalent or statement that absolute values are being used would have sufficed. A few provided an incorrect critical

value, often -1.282, or made inappropriate comparisons such as 0.029 < 0.1 for the probability method.

Unit level raw mark and UMS grade boundaries June 2017 series

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

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 (AS)	Raw	72	63	58	53	49	45	0
			UMS	100	80	70	60	50	40	0
4752	01	C2 – MEI Concepts for advanced mathematics (AS)	Raw	72	55	49	44	39	34	0
			UMS	100	80	70	60	50	40	0
4753	01	(C3) MEI Methods for Advanced Mathematics with Coursework: Written Paper	Raw	72	54	49	45	41	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	67	61	55	49	43	0
			UMS	100	80	70	60	50	40	0
4755	01	FP1 – MEI Further concepts for advanced mathematics (AS)	Raw	72	57	52	47	42	38	0
			UMS	100	80	70	60	50	40	0
4756	01	FP2 – MEI Further methods for advanced mathematics (A2)	Raw	72	65	58	52	46	40	0
			UMS	100	80	70	60	50	40	0
4757	01	FP3 – MEI Further applications of advanced mathematics (A2)	Raw	72	64	56	48	41	34	0
			UMS	100	80	70	60	50	40	0
4758	01	(DE) MEI Differential Equations with Coursework: Written Paper	Raw	72	63	56	50	44	37	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	57	49	41	34	27	0
			UMS	100	80	70	60	50	40	0
4762	01	M2 – MEI Mechanics 2 (A2)	Raw	72	56	48	41	34	27	0
			UMS	100	80	70	60	50	40	0
4763	01	M3 – MEI Mechanics 3 (A2)	Raw	72	58	50	43	36	29	0
			UMS	100	80	70	60	50	40	0
4764	01	M4 – MEI Mechanics 4 (A2)	Raw	72	53	45	38	31	24	0
			UMS	100	80	70	60	50	40	0
4766	01	S1 – MEI Statistics 1 (AS)	Raw	72	61	55	49	43	37	0
			UMS	100	80	70	60	50	40	0
4767	01	S2 – MEI Statistics 2 (A2)	Raw	72	56	50	45	40	35	0
			UMS	100	80	70	60	50	40	0
4768	01	S3 – MEI Statistics 3 (A2)	Raw	72	63	57	51	46	41	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	52	46	41	36	31	0
			UMS	100	80	70	60	50	40	0
4772	01	D2 – MEI Decision mathematics 2 (A2)	Raw	72	53	48	43	39	35	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	58	53	48	43	37	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	48	41	34	27	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	61	55	49	43	37	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	61	55	49	43	37	0
		UMS	100	80	70	60	50	40	0
G246	01 Decision 1 MEI	Raw	72	52	46	41	36	31	0
		UMS	100	80	70	60	50	40	0

Level 3 Certificate and FSMQ raw mark grade boundaries June 2017 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 2017							
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	54	47	40	34	28	0
H866	02	Critical maths	Raw	60*	48	42	36	30	24	0
		*Component 02 is weighted to give marks out of 72	Overall	144	112	97	83	70	57	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	54	47	40	34	28	0
H867	02	Statistical problem solving	Raw	60*	41	36	31	27	23	0
		*Component 02 is weighted to give marks out of 72	Overall	144	103	90	77	66	56	0

Advanced Free Standing Mathematics Qualification (FSMQ)

				Max Mark	a	b	c	d	e	u
6993	01	Additional Mathematics	Raw	100	72	63	55	47	39	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