OCR

Oxford Cambridge and RSA

Friday 16 June 2017 – Afternoon

A2 GCE MATHEMATICS (MEI)

4768/01 Statistics 3

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4768/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

6864217991

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 **16** 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 food manufacturer produces baby food, which should not contain more than 30 mg of salt per jar on average. For quality control purposes the food manufacturer tests a random sample of jars every week.

In a particular week, the amounts of salt, x mg, in a random sample of 16 jars are measured. The results are summarised as follows:

$$\Sigma x = 492, \qquad \Sigma x^2 = 15\,186.$$

(i) (*A*) Why is a test based on the Normal distribution not appropriate in this case? **[2]**

(*B*) Carry out a t test, at the 5% significance level, to test whether the mean amount of salt per jar exceeds 30 mg. You may assume that all the conditions required for the t test are fulfilled. **[10]**

(ii) Construct a 95% confidence interval for the true mean amount of salt per jar. **[4]**

(iii) The marketing director says that there is a 95% chance that the true mean amount of salt lies in this interval. Explain what is wrong with the marketing director's statement, and write an improved statement interpreting the meaning of a 95% confidence interval. **[2]**

2 **(i)** In a dance contest, judges award each competitor a mark between 1.0 and 10.0, inclusive. Marks are given to one decimal place. There is some concern that Judge 1 awards higher marks on the whole than Judge 2. The marks given by those two judges, for a random sample of 8 competitors, are as follows.

Competitor	A	B	C	D	E	F	G	H
Judge 1	9.9	3.4	8.1	4.0	7.2	4.7	4.2	3.8
Judge 2	7.4	5.7	6.5	8.1	4.2	1.6	3.4	6.0

(*A*) Explain why a t test might not be appropriate in this case. **[1]**

(*B*) Carry out an appropriate test, at the 5% significance level, to test whether Judge 1 awards higher marks on the whole than Judge 2. **[10]**

(ii) In a different round of the contest, the judges were instructed to award only integer marks between 3 and 10 inclusive. One of the organisers believes that the eight possible marks are equally likely to be awarded. To check this he obtains the following random sample of 80 marks awarded.

Mark	3	4	5	6	7	8	9	10
Frequency	5	6	10	9	14	16	14	6

Carry out a goodness of fit test, with a significance level of 10%, to investigate the organiser's belief. **[8]**

3 The random variable X has the following probability density function, $\mathrm{f}(x)$.

$$\mathrm{f}(x) = \begin{cases} \frac{1}{108}x^2(6-x) & \text{for } 0 \leqslant x \leqslant 6, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Sketch the probability density function. **[2]**

(ii) Find the mode of X. **[2]**

(iii) Find the mean of X and show that the standard deviation of X is $\frac{6}{5}$. **[8]**

(iv) Let $\overline{X}$ be the mean of a random sample of 50 observations of X. Find $\mathrm{P}(\overline{X} > 4)$.

Why did you need to use the Central Limit Theorem to find this probability? **[6]**

4 A fishmonger sells two types of fish, mackerel and trout. The weights of fish are Normally distributed, with means and standard deviations shown in the table below.

Fish	Mean weight (kg)	Standard deviation (kg)
Mackerel	0.468	0.067
Trout	0.395	0.093

(i) Find the probability that a randomly chosen mackerel weighs more than 0.5 kg. **[3]**

(ii) Find the probability that a randomly chosen mackerel weighs less than a randomly chosen trout. **[4]**

(iii) Mackerel costs £3.50 per kg and trout £4.00 per kg. Tim buys one mackerel and two trout, chosen randomly. Find the probability that he pays more than £5. **[4]**

(iv) The fishmonger offers a discount for buying 10 or more mackerel. The discounted price is £w per kg.

(A) Let £D be the discounted price of 10 mackerel. Find, in terms of w, the mean and standard deviation of D. **[2]**

(B) The probability that, with the discount, 10 mackerel cost less than £14 should not be greater than 0.1. Find the smallest possible value of w. **[4]**

END OF QUESTION PAPER

OCR
Oxford Cambridge and RSA

Friday 16 June 2017 – Afternoon

A2 GCE MATHEMATICS (MEI)

4768/01 Statistics 3

PRINTED ANSWER BOOK

Candidates answer on this Printed Answer Book.

OCR supplied materials:
- Question Paper 4768/01 (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 **16** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

1 (i) (*A*)	
1 (i) (*B*)	
	(answer space continued on next page)

1 (i) (*B*)	**(continued)**
1 (ii)	
1 (iii)	

2 (i) (*A*)	
2 (i) (*B*)	
	(answer space continued on next page)

2 (i) (*B*)	**(continued)**

2 (ii)	

3 (i)	
3 (ii)	

3 (iii)	
	(answer space continued on next page)

3 (iii)	**(continued)**
3 (iv)	

4 (i)	
4 (ii)	

4 (iii)

4 (iv) (*A*)	
4 (iv) (*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 **4768**: Statistics 3

Advanced GCE

Mark Scheme for June 2017

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

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)	(*A*)	Population standard deviation unknown. Small sample.	M1 A1 **[2]**	Allow n < 30. Just "small sample" scores M0A0 Needs "population" for M1, but "unknown variance and small sample" is SC B1 SC "Population not normal" B1
1	(i)	(*B*)	$H_0: \mu = 30$ $H_1: \mu > 30$ where μ(mg) is the mean amount of salt per jar. $\bar{x} = 30.75$ $s^2 = 3.8$ (**OR** $s = 1.949$) Test statistic is $\frac{30.75-30}{\sqrt{\frac{3.8}{16}}}$ $= 1.54$ Using t_{15}. One-tail 5% point is 1.753 (**OR** $P(t_{15} > 1.54) = 0.0722$) 1.753 > 1.54 (**OR** $0.0722 > 0.05$) Not significant. Insufficient evidence that the mean amount of salt per jar is greater than 30(mg).	B1 B1 B1 B1 M1 A1 M1 A1 M1 A1 ft **[10]**	both hypotheses. Not using $\bar{x}$ Adequate verbal definition. If not using μ must say "population" ft c's mean and sd. no ft if wrong. no ft if wrong. Explicit comparison seen. consistent with c's mean and sd. If the comparison is not explicit (e.g. the two numbers written close to each other) can get SC B1 for the correct conclusion.
1	**(ii)**		CI is $30.75 \pm$ 2.131 $\times \sqrt{\frac{3.8}{16}}$ $= 30.75 \pm 0.4873 = (29.7, 31.8)$	M1 B1 M1 A1 **[4]**	Using c's mean using c's sd c.a.o. Must be expressed as an interval. 3 or 4 sf only.

1	(iii)		The (population) mean is fixed, we can't talk about the probability of it being in a given interval. 95% of the confidence intervals created by repeated sampling will contain the true mean.	B1 B1 **[2]**	"It either is or isn't" is ok	

<table>
<tr><td>2</td><td>(i)</td><td>(A)</td><td>The (population) distribution of differences of marks is (not known to be) normal.</td><td>B1
[1]</td><td>Must refer to differences (not just “the underlying population”</td><td></td></tr>
<tr><td>2</td><td>(i)</td><td>(B)</td><td>HH_0: $m = 0$, H_1: $m > 0$
where m is the population median of the difference between Judge 1’s and Judge 2’s marks.
<table><tr><td>difference</td><td>2.5</td><td>-2.3</td><td>1.6</td><td>-4.1</td><td>3.0</td><td>3.1</td><td>0.8</td><td>-2.2</td></tr><tr><td>rank</td><td>5</td><td>4</td><td>2</td><td>8</td><td>6</td><td>7</td><td>1</td><td>3</td></tr></table>
$W_- = 4 + 8 + 3 = 15$
Refer to Wilcoxon tables with n=8.
5% critical value is 5.
15 < 5 so result not significant

Insufficient evidence to suggest that judge 1 awards higher marks on the whole.</td><td>B1
B1

M1
M1
A1

A1
M1
A1
M1

A1

[10]</td><td>Both hypotheses
For adequate definition of m, in context; must refer to <u>population</u> median.
(not “difference in medians”
For differences.
For ranks.
ft from here if ranks wrong.

or $W_+ = 21$
no ft if wrong
or 31 no ft if wrong
ft c’s test statistic. Must compare this or W_+ to 31. No ft if comparing wrong tail.
ft c’s test statistic. Must include ‘evidence’ and ‘on the whole’ or oe.</td><td></td></tr>
<tr><td>2</td><td>(ii)</td><td></td><td>H_0: Judges awarded the same number of each mark.
H_1: Judges did not award the same number of each mark.
<table><tr><td>observed</td><td>5</td><td>6</td><td>10</td><td>9</td><td>14</td><td>16</td><td>14</td><td>6</td></tr><tr><td>expected</td><td>10</td><td>10</td><td>10</td><td>10</td><td>10</td><td>10</td><td>10</td><td>10</td></tr></table>
$\chi^2 = 2.5 + 1.6 + 0 + 0.1 + 1.6 + 3.6 + 1.6 + 1.6$
$= 12.6$
Refer to χ^2_7
The 10% critical value is 12.02.
12.6 > 12.02 so significant
There is sufficient evidence that judges have not been awarding the same number of each mark.</td><td>B1

B1

M1

A1
M1
A1
M1
A1

[8]</td><td>Both hypotheses. Must be in correct context. Allow ‘uniform distribution’ or ‘in equal proportions’.
‘Model fits data’ or ‘belief is justified’ is ok.
Do not accept ‘data fits model’ oe’.

For expected frequencies.

Calculation of χ^2. (if 12.6 not seen, must see evidence of calculation)
cao.
No ft if wrong.
No ft if wrong.
ft their test statistic
Must be in context and mention ‘evidence’. (‘organiser’s belief’ is sufficient context)</td><td></td></tr>
</table>

3	(i)		Negative cubic through the origin, positive x-intercept Only the part between 0 and 6 shown; gradient 0 at the origin and correct shape between max and 6.	M1 A1 **[2]**	A0 if there is a point of infelction between 4 and 6.	
3	(ii)		$f'(x) = \frac{4x-x^2}{36} = 0$ Mode is $x = 4$	M1 A1 **[2]**	Needs both attempt at differentiation and =0. No need to justify this is max. SC B1 for '*x* = 4 is the maximum point on the graph'.	
3	(iii)		$E(X) = \int_0^6 \frac{1}{108}x^3(6-x)\,dx$ $= \left[\frac{x^4}{72} - \frac{x^5}{540}\right]_0^6$ $= \frac{18}{5}$ (= 3.6) $Var(X) = \int_0^6 \frac{1}{108}x^4(6-x)\,dx - \left(\frac{18}{5}\right)^2$ Integral $= \left[\frac{x^5}{90} - \frac{x^6}{648}\right]_0^6 \left(= \frac{72}{5}\right)$ $Var(X) = \frac{72}{5} - \left(\frac{18}{5}\right)^2 = \frac{36}{25}$ sd $\left(= \sqrt{\frac{36}{25}}\right) = \frac{6}{5}$	M1 A1 A1 M1 B1 A1 A1 ft B1 **[8]**	Limits needed somewhere. For correct integration. Can be implied by correct answer. c.a.o. (3.6 seen implies all three marks) For the integral, needs limits. For -(their mean)^2 Correct integrated expression. Using their mean www	
3	(iv)		$\bar{X}$ is approximately normal mean = 3.6 sd = $\frac{1.2}{\sqrt{50}}$ (= 0.1697) or var = 0.028798 $P(\bar{X} > 4) = 1 - P\left(Z < \frac{4-3.6}{0.1697} = 2.357\right)$ $= 0.00921$ The distribution of *X* is not normal.	M1 B1 ft B1 M1 A1 B1 **[6]**	s.o.i. Allow c's mean from (ii) For standardising using their mean, sd or var divided by $\sqrt{50}$. Requires 1- (Note: can't get this M1 if no $\sqrt{50}$) c.a.o. (ans 0.0092 from tables is ok, 2 to 4 sf)	

4	(i)		$P(M > 0.5) = 1 - P\left(Z < \frac{0.5-0.468}{0.067} = 0.4776\right)$ $= 0.316$	M1 A1 A1 **[3]**	For standardising. Award here or elsewhere. For 0.4776. Can be implied by correct 0.316. answer between 0.316 and 0.3165, 3 or 4 sf.	
4	(ii)		Require $P(M - T) < 0$ $M - T \sim N(0.073, 0.0131\ )$ $P\left(Z < \frac{0-0.073}{0.1146} = -0.637\right) = 0.262$	M1 B1 B1 A1 **[4]**	Mean Variance. Accept sd (=0.1146) c.a.o, 3 or 4 sf (0.263 comes from early rounding and is A0)	
4	(iii)		Cost $C = 3.50M + 4.00T_1 + 4.00T_2$ $\sim N(4.798, 0.332)$ $P(C > 5) = 1 - P\left(Z < \frac{5-4.798}{0.576} = 0.351\right)$ $= 0.363$	M1 B1 B1 A1 **[4]**	Recognising that the two T's are different (if incorrect variance, need an explicit calculation showing that T_1 and T_2 were used rather than 2T) Mean Variance. Accept sd (0.576) c.a.o, 3 or 4 sf.	
4	(iv)	(*A*)	mean = $4.68w$ sd = $0.067\sqrt{10}w$ ($= w\sqrt{0.00489}$ = $0.212w$)	B1 B1 **[2]**		
4	(iv)	(*B*)	$\frac{14-4.68w}{0.067\sqrt{10}w} \le -1.282$ $14 - 4.68w \le -0.272w$ $w \ge 3.18$	M1 B1 M1 A1 **[4]**	Standardising with their mean and sd (allow variance) -1.282 seen (must be the correct sign, so can be 4.68w-14 and +1.282; allow = Attempting to solve, some working needed, but can be implied by correct answer; allow =. 3 or 4 sf; allow w = 3.18	

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

4768 Statistics 3

General Comments:

The standard of candidates on this paper was again very high. Almost all candidates attempted every question and it was clear that they were well prepared for this unit. Statistical tests and calculations were generally performed accurately and presented with sufficient detail and clarity. Candidates should be reminded to round their answers to an appropriate degree of accuracy; in particular, in most situations, an accuracy of more than four significant figures is unlikely to be justified.

In hypothesis tests most candidates are writing their hypotheses and conclusions correctly. The most common error is to forget to make it clear that they are referring to a *population* parameter. A more subtle mistake arises in conclusions, which should state whether there is evidence to reject H_0, rather than evidence to accept it (for example, 'There is insufficient evidence that the mean exceeds 30mg' rather than 'There is evidence that the mean does not exceed 30mg'). A good guiding principle is to use the wording of the question when writing down the conclusion.

Although mathematical presentation and communication were generally very good, there was often lack of clarity in notation for random variables. In particular, when working with combinations of random variables, candidates should be careful to distinguish between, for example, $2T$ and $T_1 + T_2$. Not only would this improve the clarity of their presentation, but it would also help them find the correct variance in questions such as Q.4iii and Q.4iv.

Several of the questions on this paper tested candidates' understanding of conditions and assumptions used in their calculations. The assumptions required for the use of various statistical tests were generally well known, although not always clearly expressed. Many candidates were able to reproduce standard statements accurately, for example to explain the meaning of a confidence interval. The most challenging concept seems to be the distribution of the sample mean. Many candidates seem to be forgetting that, if the underlying distribution is Normal, then the sample mean will also be Normally distributed, regardless of the sample size (and hence the Central Limit Theorem would not be required). In general, candidates seem very confident in stating the required assumptions for the application of various theorems or test, but not so good at recognising when those assumptions are not met. They could benefit from seeing more situations in which those theorems do not apply.

Many candidates are using calculators effectively and appropriately to find values from statistical distributions. For those who are confident with it, technology could be utilised further in teaching to develop deeper understanding of the concepts such as the distribution of the sample mean and confidence intervals.

Comments on Individual Questions:

Question No. 1

This question required using the t-distribution to conduct a hypothesis test and construct a confidence interval.

Part (i)A asked why a test based on Normal distribution might not be appropriate in this case. Many candidates seem to associate a t-test with small samples, forgetting that, if the underlying distribution is Normal and the population variance is known, the sample mean would follow a Normal distribution even for a small sample.

The t-test in part (i)(B) was generally done well, with the hypotheses and conclusions clearly stated in a majority of cases. The most common mistake was to use 2.131 as the critical value (which is for the 2-tail test at the 5% significance level).

In part (ii) candidates were generally able to construct the confidence interval correctly, with a very small number using 1.753 (which is for a 90% confidence interval) or 1.96 (from Normal distribution). However, a significant number lost a mark for giving the final answer to 5 or 6 significant figures; this level of accuracy is unjustified.

Part (iii) asked for an interpretation of a 95% confidence interval. Many candidates were able to give a correct interpretation, but a significantly smaller number managed to explain clearly why the proposed statement was incorrect. Many seemed to be saying that 95% of samples would have a mean in "this interval" (seemingly referring to the particular interval calculated in part (ii)). Candidates should be encouraged to understand that the population mean is fixed and the confidence interval changes with each sample.

Question No. 2

Part (i) of this question involved a Wilcoxon test. Candidates were first asked why a t-test was not appropriate. Many knew that this was to do with the distribution of the underlying population, but few seemed aware that only the differences, rather than the scores themselves, needed to come from a Normal distribution. The test itself was generally conducted accurately. However, many marks were lost for the hypotheses and the conclusion, which did not always clearly refer to the population median difference (some candidates talked about the difference between the scores, which is not precise enough).

Part (ii) required a chi squared test and was generally very well done. The simplicity of the expected frequencies meant that there were very few accuracy errors, and a vast majority used the correct critical values. The main loss of marks came from the conclusions, which were not always given in context. ('The model of equal probability is not appropriate' was not considered sufficient context.) The hypotheses were sometimes stated in a generic form ('The model fits the data'); this was not penalised, although a more specific statement of what "the model" is would be preferable. Candidates should be reminded that 'Data fits the model' is not a correct statement of the null hypothesis.

Question No. 3

The topic of continuous probability distributions seems to be very well understood and this question produced many correct answers. The sketches in part (i) did not always clearly show the zero gradient at the origin and so many candidates scored only 1 out of the 2 marks.

In part (ii) most scored full marks. A minority thought that the mode was f(4) rather than 4. Calculator solutions needed to clearly refer to the maximum point on the graph rather than just giving the answer; candidates should be reminded that calculator answers are inexact and should therefore be given to an appropriate degree of accuracy (in this case, *not* to the nearest integer).

Part (iii) was probably the best done question on the whole paper. There was plenty of detail shown and there were very few mistakes.

In part (iv) there were very many correct calculations, the most common mistake being the confusion between the standard deviation and the variance. The reasons for requiring the Central Limit Theorem did not seem to be very well understood: many candidates stated that it was needed 'because the sample was large'. Another common misconception is that, for a large sample, the population distribution becomes Normal. Some also seem to believe that the CLT states that the variance of the sample mean needs to be divided by n; this is in fact the case regardless of the distribution and the sample size.

Question No. 4

This question was about linear combinations of Normal random variables. The first two parts were generally very well done.

Part (iii) proved more challenging; many candidates considered $M + 2T$ instead of $M + T_1 + T_2$. (Some wrote $M + 2T$ but then produced the correct variance; they were able to score full marks.) Candidates should be advised to write out their combinations of random variables carefully.

In part (iv)A many candidates seemed to confuse the discounted price with the discount, resulting in incorrect answers; they were still able to score 3 out of the 4 marks in part (iv)B. Many also lost one mark for writing down the variance instead of the standard deviation. In part (iv)B common mistakes were to use the wrong tail of the distribution (for example, using 1.282 instead of -1.282), and to give answers to an inappropriate degree of accuracy (anything other than 2 decimal places, as the value represented the amount of money).

OCR
Oxford Cambridge and RSA

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

			Max Mark	a	b	c	d	e	u
		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