OCR
Oxford Cambridge and RSA

Friday 17 June 2016 – 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

6367236276

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 the Printed Answer Book. 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 game consists of 20 rounds. Each round is denoted as either a starter, middle or final round. The times taken for each round are independently and Normally distributed with the following parameters (given in seconds).

Type of round	Mean	Standard deviation
Starter	200	15
Middle	220	25
Final	250	20

The game consists of 4 starter, 12 middle and 4 final rounds. Find the probability that

(i) the mean time per round for the 4 final rounds will exceed 260 seconds, **[3]**

(ii) all 20 rounds will be completed in a total time of 75 minutes or less, **[5]**

(iii) the 12 middle rounds will take at least 3.5 times as long in total as the 4 starter rounds, **[5]**

(iv) the mean time per round for the 12 middle rounds will be at least 25 seconds less than the mean time per round for the 4 final rounds. **[5]**

2 **(a)** A genetic model involving body colour and eye colour of fruit flies predicts that offspring will consist of four phenotypes in the ratio 9 : 3 : 3 : 1.

A random sample of 200 such offspring is taken. Their phenotypes are found to be as follows.

Phenotype	Brown body Red eye	Brown body Brown eye	Black body Red eye	Black body Brown eye
Frequency	125	37	32	6
Relative proportion from model	9	3	3	1

Carry out a test, using a 2.5% level of significance, of the goodness of fit of the genetic model to these data. **[9]**

(b) The median length of European fruit flies is 2.5 mm. South American fruit flies are believed to be larger than European fruit flies. A random sample of 12 South American fruit flies is taken. The flies are found to have the following lengths (in mm).

1.7 1.4 3.1 3.5 3.8 4.2 2.2 2.9 4.4 2.6 3.9 3.2

Carry out a Wilcoxon signed rank test, using a 5% level of significance, to test this belief. **[9]**

3 The random variable X has the following probability density function:

$$f(x) = \begin{cases} k(1-x^2) & -1 \leqslant x \leqslant 1 \\ 0 & \text{elsewhere,} \end{cases}$$

where k is a positive constant.

(i) Calculate the value of k. **[3]**

(ii) Sketch the probability density function. **[3]**

(iii) Calculate Var(X). **[3]**

(iv) Find a cubic equation satisfied by the upper quartile q, and hence verify that $q = 0.35$ to 2 decimal places. **[5]**

(v) A random sample of 40 values of X is taken. Using a suitable approximating distribution, calculate the probability that the mean of these values is greater than 0.125. Justify your choice of distribution. **[4]**

4 An insurance company is investigating a new system designed to reduce the average time taken to process claim forms. The company has decided to use 10 experienced employees to process claims using the old system and the new system.

Two procedures for comparing the systems are proposed.

Procedure A There are two sets of claim forms, set 1 and set 2. Each contains the same number of forms. Each employee processes set 1 on the old system and set 2 on the new system. The times taken are compared.

Procedure B There is just one set of claim forms which each employee processes firstly on the old system and then on the new system. The times taken are compared.

(i) State one weakness of each of these procedures. **[2]**

In fact a third procedure which avoids these two weaknesses is adopted. In this procedure each employee is given a randomly selected set of claim forms. Each set contains the same number of forms. The employees each process their set of claim forms on both systems. The times taken, in minutes, are shown in the table.

Employee	1	2	3	4	5	6	7	8	9	10
Old system	40.5	42.9	52.8	51.7	77.2	66.7	65.2	49.2	55.6	58.3
New system	39.2	40.7	50.6	50.7	71.4	70.5	71.1	47.7	52.1	55.5

(ii) Carry out a paired t test at the 5% level of significance to investigate whether the mean length of time taken to process a set of forms has reduced using the new system. **[10]**

(iii) State fully the usual conditions for a paired t test. **[3]**

(iv) Construct a 99% confidence interval for the mean reduction in time taken to process a set of forms using the new system. **[3]**

END OF QUESTION PAPER

OCR
Oxford Cambridge and RSA

Friday 17 June 2016 – 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 the Printed Answer Book. 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.

1 (i)	
1 (ii)	

1 (iii)

1 (iv)

2 (a)	

2 (b)

3 (i)	
3 (ii)	

3 (iii)	

3 (iv)	

3 (v)	

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

4 (ii)	**(continued)**
4 (iii)	
4 (iv)	

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 2016

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

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.

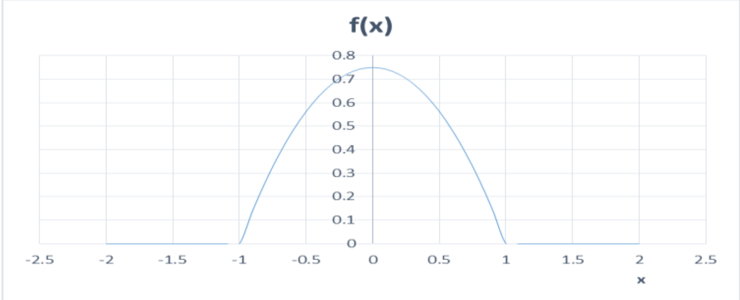
Note that a miscopy of the candidate's own working is not a misread but an accuracy error.

Question			Answer	Marks	Guidance
1	i		$F \sim N(250, 20^2)$ $P(\bar{F_4} > 260) = P\left(Z > \frac{260-250}{10}\right)$ $= P(Z > 1)$ $= 0.1587$	M1 M1 A1 [3]	standardisation including division by $\sqrt{n}$ correct tail (probability < 0.5) cao (to 3 or 4 sf)
1	ii		$(F_1 + F_2 + F_3 + F_4) = F' \sim N(1000,1600)$ $(M_1 + M_2 + .. + M_{12}) = M' \sim N(2640,7500)$ $(S_1 + S_2 + S_3 + S_4) = S' \sim N(800,900)$ $\rightarrow T \sim N(4440, 100^2)$ $P(T < 4500) = P\left(Z < \frac{4500-4440}{100}\right) = P(Z < 0.6)$ $= 0.7258$	M1 A1 B1 M1 A1 [5]	for variances: at least one of 4×15^2 etc. seen allow '4×15 + 12×25 + 4×20', but not 4^2 etc. for 10,000 (or 2.778 in minutes) for 4440 (or 74 in minutes) correct tail (probability > 0.5) and √(their variance) art 0.726 (given to 3 or 4 sf)
1	iii		Looking for $P\left((M' - 3.5S') > 0\right)$ $[M' \sim N(2640,7500)]$ $3.5S' \sim N(2800,11025)$ $(M' - 3.5S') \sim N(-160,18525)$ $= P\left(Z > \frac{160}{\sqrt{18525}}\right) = P(Z > 1.1755) = 0.1199$	M1 M1 B1,A1 A1 [5]	interpret the question correctly; e.g. '12M – 3.5×4S' or '12M > 3.5×4S' seen their Var(S')×12.25 mean and variance cao (0.1198 to 0.120)
1	iv		Looking for $P\left((\bar{F_4} - \bar{M_{12}}) > 25\right)$ $\bar{M_{12}} \sim N\left(220, \frac{625}{12}\right)$ $\bar{F_4} \sim N(250,100)$ $(\bar{F_4} - \bar{M_{12}}) \sim N(30,152.08)$ $P\left(Z > \frac{25-30}{\sqrt{152.08}}\right) = P(Z > -0.4054) = 0.6574$	M1 M1 A1 B1 A1 [5]	interpret the question correctly; e.g. $P(\bar{F_4} > \bar{M_{12}} + 25)$ seen variance: at least one of $\frac{25^2}{12}$ or $\frac{20^2}{4}$ seen correct variance correct mean answer rounds to 0.657 or 0.658

Question			Answer	Marks	Guidance	
2	a		H_0: The (genetic) model fits the data. H_1: The (genetic) model does not fit the data Observed 125 37 32 6 Expected 112.5 37.5 37.5 12.5 Cont's 1.3889 0.0067 0.8067 3.38 $X^2 = 5.582$ Degrees of freedom = 3 Critical value = 9.348 5.582 < 9.348 → cannot reject H_0 The data give no reason to doubt the genetic model	B1 B1 M1 A1 A1 B1 B1 M1 E1 [9]	Both hypotheses; Not 'data fits model' Expected values correct use of $(O\text{-}E)^2/E$ (at least one correct) all correct to 3dp where appropriate cao (3sf or 4sf) no FT if wrong (can be implied by 9.348) no FT if wrong FT their X^2 Do not accept data fits model; but 'Evidence suggests that model first the data' is fine	
2	b		$H_0: m = 2.5$ $H_1: m > 2.5$ where m is the population median length (of South American fruit flies)	B1 B1	both hypotheses definition including median, population, and context (If given in words: B1 for mentioning median 2.5, B1 for context)	

Observation	-2.5	rank
1.7	-0.8	**6**
1.4	-1.1	**8**
3.1	0.6	4
3.5	1.0	7
3.8	1.3	9
4.2	1.7	11
2.2	-0.3	**2**
2.9	0.4	3
4.4	1.9	12
2.6	0.1	1
3.9	1.4	10
3.2	0.7	5

Answer	Marks	Guidance
(table above)	M1	subtract 2.5
	M1	ranking
	A1	all ranks correct
$W_- = 16, W_+ = 62$	B1	for either, cao
($n = 12$), Critical value = 17	B1	allow 61 if compared to 62. No FT if wrong
($16 < 17$ →) reject H_0		

Question			Answer	Marks	Guidance	
			Suggests population median length of South American fruit flies exceeds 2.5cm	M1 A1 [9]	FT their W including median (or 'on average') and context	
3	i		$k\int_{-1}^{1}(1-x^2)\mathrm{d}x = 1$ $\left(\to k\left[x-\frac{x^3}{3}\right]_{-1}^{1} = 1\right)$ $\to \frac{4k}{3} = 1$ $\to k = \frac{3}{4}$	M1 M1 A1 [3]	Correct integral including limits (const)×k = 1 cao	
3	ii			B1 B1 B1 [3]	General shape between -1 and +1 Axes labelled with scales and intercepts (FT their k) Nothing outside $\lvert x\rvert < 1$	
3	iii		$\mathrm{E}(X) = 0 \to \mathrm{V}(X) = \mathrm{E}(X^2)$ $\mathrm{V}(X) = \frac{3}{4}\int_{-1}^{1}(x^2 - x^4)\mathrm{d}x = \frac{3}{4}\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_{-1}^{1}$ $= \frac{1}{5}$	B1 M1 A1 [3]	for $\mathrm{E}(X) = 0$ for correct integral including limits cao (ignore mistakes in working)	

Question		Answer	Marks	Guidance
3	iv	$\frac{3}{4}\int_0^q(1-x^2)\mathrm{d}x=\frac{1}{4}$	M1	Correct limits and equality
		integration $=\frac{3}{4}\left[x-\frac{x^3}{3}\right]$	B1	f/t their *k*
		$\rightarrow q-\frac{q^3}{3}=\frac{1}{3}$ or $\rightarrow q^3-3q+1=0$	A1	any correct simplified (3-term) cubic
		$g(0.345)=0.006$ $g(0.355)=-0.02$	M1	(allow correct alternative) If solving using calculator: state all three solutions
		Change of sign $\rightarrow 0.345<q<0.355$ So upper quartile $=0.35$ to 2 dp	E1 [5]	must be explained clearly If solving by calculator: explain why only one works
3	v	$\sum X_i>5\rightarrow \overline{X_{40}}>0.125n$ large and so can use Central Limit Theorem	B1	both (or $\bar{X}$ normal)
		$\overline{X_{40}}\sim\mathrm{N}\left(0,\frac{0.2}{40}\right)$	M1	for $\frac{\mathrm{Var}}{40}$ or $\frac{\mathrm{Var}}{\sqrt{40}}$
			A1	correct mean and variance (ft any positive variance from iii)
		$\mathrm{P}(\overline{X_{40}}>5)=\mathrm{P}\left(Z>\frac{0.125-0}{0.0707}\right)=\mathrm{P}(Z>1.768)$ $\mathrm{P}(\overline{X_{40}}>0.125)=\mathrm{P}\left(Z>\frac{0.125-0}{0.0707}\right)=\mathrm{P}(Z>1.768)$ $=0.0385$	A1 [4]	cao 0.0385 or 0.0386
4	i	A) One set of claim forms could be more difficult to process. B) A form would be more familiar on second processing.	B1 B1 [2]	Allow suitable alternatives

Question			Answer	Marks	Guidance	
4	ii		Values of d 1.3, 2.2, 2.2, 1.0, 5.8, -3.8, -5.9, 1.5, 3.5, 2.8 $\bar{d} = 1.06, s = 3.4378$ $H_0: \mu_D = 0$ $H_1: \mu_D > 0$ Where μ_D is the population mean difference between the times taken to process claims using the old system and the new system $t = \frac{1.06-0}{3.4378/\sqrt{10}} = 0.975$ 9 degrees of freedom Critical value = 1.833 0.975 < 1.833 so cannot reject H_0 Insufficient evidence to suggest that the mean time for processing forms has reduced using the new system	M1 A1 B1 B1 M1 A1 B1 B1 M1 A1 [10]	calculating differences (at least 3 correct) both. Allow $s^2 = 11.818$ Do not allow $s_n = 3.2613$ or $s_n^2 = 10.636$ hypotheses (allow <0 if consistent) definition including difference, mean, and context. Allow other symbols only if they are defined as the population mean difference. Hypotheses in words must include 'population'. Not 'difference of means' including $\sqrt{10}$. FT their s or s_n cao no FT if wrong no FT if wrong sensible comparison using their test statistic if the previous M1 awarded. including mean and context FT; not assertive but accept 'Evidence suggests mean time has reduced'.	
4	iii		Differences should be: Normally distributed in population With unknown variance Sample (of differences) must be random	 B1 B1 B1 [3]	 or 'underlying distribution of differences is normal' NB: candidates may say the sample should be small. This is incorrect, but should be ignored for the purposes of marking. Also ignore 'paired'	
4	iv		$1.06 \pm \frac{3.25(3.4378)}{\sqrt{10}}$ $= (-2.473, 4.593)$	M1 B1 A1 [3]	for their $\bar{d}$ and $s/\sqrt{10}$ in correct position for 3.25 cao SC: Answers from calculator with no working 3sf or 4sf gets 3/3, >4sf gets 1/3 (-4.593, 2.473) gets 2/3	

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

4768 Statistics 3

General Comments:

This paper was generally very well answered. A vast majority of candidates attempted all the questions and were able to show what they could do.

The solutions to the hypothesis testing questions (2 and 4) were correctly structured, with both hypotheses and conclusions stated clearly and in the context of the question. Occasionally they lacked sufficient context, or the conclusions were overly assertive, but it was pleasing to see that such cases were in a minority. Learners should be reminded that it is good practice to state the hypotheses before carrying out any calculations, even though marks in the examination are awarded for the hypotheses seen anywhere within the solution.

Generally the solutions contained sufficient detail to make the methods clear, even when graphical calculators were used to find, for example, probabilities from the Normal distribution.

The notation for random variables was sometimes unclear, both in Question 1 and when explaining the Central Limit Theorem in Question 3(v). Insisting on clearer notation here might have helped candidates avoid confusion between a random variable and the corresponding sample mean, and enable them to score more marks.

It was pleasing to see that most candidates attempted to answer the 'wordy' questions. There were many good explanations, but sometimes the sentences were not clearly structured which made it difficult to follow the ideas. Bad handwriting on occasion made it difficult to award marks.

Candidates should be reminded to give their answers to an appropriate degree of accuracy, which is generally three or four significant figures. Over-specifying answers can lead to a loss of marks.

Comments on Individual Questions:

Question No. 1

This question was about combinations of Normal variables. Calculations using the Normal distribution were generally done well, with sufficient detail and to an appropriate accuracy.

The random variables in question were not always clearly defined. This was not penalised in itself, but often led to the wrong variance and hence the wrong answer. There were two common mistakes: using the variable instead of its sample mean (and thus failing to divide the variance by n); and writing, for example, $4M$ when what was meant was $M_1 + M_2 + M_3 + M_4$ (which would lead to the variance $16\sigma^2$ instead of $4\sigma^2$). Teachers are therefore advised to insist on correct notation for random variables in this type of question.

In part (i) a large number of candidates divided by 20 instead of 10 in the standardisation. In parts (ii) and (iii) careless notation, as described above, often led to the incorrect variance. Where this was avoided, part (ii) in particular was very well done. In part (iv) many either misinterpreted the question ($M – F – 25$ was often seen) or forgot that it was about the means.

Question No. 2

This question consisted of two parts, the first requiring a chi-squared test and the second a Wilcoxon test. Both parts were very well done overall, with the calculations mostly being correct, hypothesis and conclusions being given in sufficient detail, and critical values correctly identified.

In part (a), it was good to see that the hypotheses were generally stated correctly (in the past we saw 'data fits model' much more often). Candidates are expected to understand that 'result insignificant' means that we have insufficient evidence to reject H_0; although this was not always clearly stated, a majority managed to score full marks by correctly interpreting the conclusion in context.

In part (b), the most common loss of marks was for forgetting to include 'population' in the definition of the parameter in the hypotheses, and concluding that the South American fruit flies were larger than the European ones without referring to the 'median' or 'average' length. A more subtle error was to omit the value 2.5 from the hypotheses (stating only that 'the median length of the South American flies is larger than the median length of the European ones'); the value 2.5 is in fact needed in testing the hypotheses. However, it was good to see that most hypotheses included a definition of the parameter and that the conclusion was nearly always contextualised.

Question No. 3

Early parts of this question were usually well done. In part (ii) a few candidates drew a triangle instead of a curve, and some parabolas had "tails" (looking more like a normal distribution curve).

In part (iii), since E(X) = 0, we expected to see '$– 0^2$ ' or '$– \mathrm{E}(X)^2$ ' in the calculation of the variance.

In part (iv) most candidates reached the correct 3-term cubic, although some stopped too soon. Some candidates equated an integral with limits 0 and q to 0.25 which is correct, but it was not clear that they were using the symmetry of the pdf. For the final part, most were content to demonstrate that substituting 0.35 gave an approximately correct right-hand side of the equation, rather than showing that the value of q was 0.35 to two decimal places. The latter requires looking for a sign change between 0.345 and 0.355.

Part (v) asked for an application of the Central Limit Theorem. The numerical answer was mostly correct, but the justification for the use of the Normal distribution revealed some misconceptions about the CLT. Many candidates seem to think that, when a sample is large, X itself can be approximated by a Normal distribution. Many correct references to the CLT forgot to mention the large sample, which is required for the application of the Theorem to be valid.

Question No. 4

In part (i) most candidates made at least one sensible comment. However, the language was often unclear so it was difficult to tell exactly what point was being made. For Procedure A there needed to be some mention of the two systems, rather than just the two sets of forms. Procedure B weakness was generally well explained, although some candidates resorted to the need for a larger sample.

Part (ii) was largely well done. In stating the hypotheses quite a few described μ as the difference between the two sets of times, omitting the word 'mean'. A minority described μ as the 'difference of the means', which is incorrect for a paired test. Again, the final conclusion often failed to mention 'average' reduction in time.

Part (iii) revealed that the conditions for using a t-test are not very well understood. Most knew that something needed to be Normally distributed, but many simply said 'the data' or 'the underlying variable', not specifying that it is in fact the population of differences. Most knew that the sample variance should be unknown. Many candidates stated a requirement for a small sample, which is not in fact necessary. On the other hand, the requirement of a random sample was often forgotten.

Finally, part (iv) asked for a confidence interval for the mean reduction in time. Some candidates used one of the single samples instead of the differences, and it was quite common to see the interval for the increase, rather than the reduction. Quite a high proportion used a wrong t value. Overall, however, most candidates knew the correct method and scored at least one mark on this question.

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