

# GCE

## Mathematics (MEI)

Advanced GCE

Unit 4768: Statistics 3

### Mark Scheme for June 2011

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of pupils of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, OCR Nationals, 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.

© OCR 2011

Any enquiries about publications should be addressed to:

OCR Publications PO Box 5050 Annesley NOTTINGHAM NG15 0DL

Telephone:0870 770 6622Facsimile:01223 552610E-mail:publications@ocr.org.uk

Q1				
(i)	<ul> <li><i>t</i> test might be used because</li> <li>population variance is unknown</li> <li>background population is Normal</li> </ul>	E1 E1	Allow "sample is small" as an alternative.	2
(ii)	H <sub>0</sub> : $\mu = 15.3$ H <sub>1</sub> : $\mu < 15.3$ where $\mu$ is the mean of Gerry's times.	B1 B1	Both hypotheses. Hypotheses in words only must include "population". Do NOT allow " $\overline{X} =$ " or similar unless $\overline{X}$ is clearly and explicitly stated to be a <u>population</u> mean. For adequate verbal definition. Allow	
	where $\mu$ is the mean of Gerry's times.	DI	absence of "population" if correct notation $\mu$ is used.	
	$\overline{x} = 14.987$ $s_{n-1} = 0.4567(5)$	B1	$s_n = 0.4333$ but do <u>NOT</u> allow this here or in construction of test statistic, but FT from there.	
	Test statistic is $\frac{14.987 - 15.3}{0.45675}$	M1	Allow c's $\overline{x}$ and/or $s_{n-1}$ . Allow alternative: 15.3 + (c's -1.833) $\times \frac{0.45675}{\sqrt{10}}$ (= 15.035) for subsequent comparison with $\overline{x}$ . (Or $\overline{x}$ - (c's -1.833) $\times \frac{0.45675}{\sqrt{10}}$	
	= -2.167(0).	A1	(= 15.252) for comparison with 15.3.) c.a.o. but ft from here in any case if wrong. Use of $\mu - \overline{x}$ scores M1A0, but ft.	
	Refer to $t_9$ . Single-tailed 5% point is $-1.833$ .	M1 A1	No ft from here if wrong. Must be minus 1.833 unless absolute values are being compared. No ft from here if wrong. P(t < -2.167(0)) = 0.0292.	
	Significant. Seems that Gerry's times have been reduced on average.	A1 A1	ft only c's test statistic. ft only c's test statistic. Conclusion in context to include "average" o.e.	9
(iii)	A 5% significance level means that the probability of rejecting $H_0$ given that it is true is 0.05. Decreasing the significance level would make it less	E1		
	likely that a true $H_0$ would be rejected. Evidence for rejecting $H_0$ would need to be stronger.	E1 E1	Or equivalent. Allow answers that relate to the context of the question.	3
(iv)	CI is given by 14.987 ± 2.262	M1 B1	ZERO/4 if not same distribution as test. Same wrong distribution scores maximum M1B0M1A0. Recovery to <i>t</i> <sub>9</sub> is OK.	
	$\times \frac{0.45675}{\sqrt{10}}$	M1		
	$= 14.987 \pm 0.3267 = (14.66(0), 15.31(3))$	Al	c.a.o. Must be expressed as an interval.	4 18
				10

Q2											
(i)											
~ /		No. particles	0	1	2		3	4	5	2	
		Obs fr	4	7	1(	)	20	17			
		Prob'y	0.0150	0.0630	0.13	322	0.1852	0.1944			
		Expfr	1.50	6.30	13.	22	18.52	19.44			
		Contrib to $X^2$	(4.1667)	(0.0778)	0.78	343	0.1183	0.3063			
		Combined		1 80 128						_	
	0.18	28 + 0.7843 + 0. 313 + 0.6676 + 0. 884(5)		063 + 0.1082	3 +	M1 M1 A1 M1 M1 A1	× 100 All con Merge Calcul	first 2 cell ation of $X^2$	ed freque		
		Poisson model fit Poisson model do		e data.		B1 B1		any refere t accept "d		ne parameter. odel" oe.	
	Refer to	$\chi_6^2$ .				M1	wrong	correct df ly grouped vise, no ft	table ar	nd ft.	
	Upper 10	0% point is 10.64				A1	No ft f		f wrong.	$(\chi_7^2 = 12.02)$	
	Not signi Evidence	ificant. e suggests that the	e model fits	the data.		A1 A1	ft only ft only	c's test sta	atistic. atistic. D	o not accept	12
(ii)	H <sub>0</sub> : $m = 1$ where $m$	15 $H_1: m > 15$ is the population		meter( in µr	n).	B1 B1	Adequ	Accept hypate definition		in words. to include	
	Given W	$M_{-} = 53 (: W_{+} = 1)$	57)								
	Refer to tables of Wilcoxon paired (/single sample) statistic for $n = 20$ .			M1	No ft f	rom here i	f wrong.				
	Lower 59	% point is 60 (or	upper is 150	) if $W_+$ used	).	A1	i.e. a 1 wrong	-tail test. N	No ft froi	m here if	
		significant.				A1		c's test sta			
		e suggests that the re than 15 μm.	e median dia	ameter appe	ars	A1		c's test sta t to include		conclusion in ge" o.e.	6
											18

Q3				
(i) (A)	G(X)	M1 A1 A1	Increasing curve, through (0, 0), in first quadrant only. Asymptotic behaviour. Asymptote labelled; condone absence of axis labels.	3
(B)	For the UQ G(u) = 0.75 $\therefore \left(1 + \frac{u}{200}\right)^{-2} = \frac{1}{4}  \therefore u = 200$ For the LQ G(l) = 0.25	M1 A1	Use of G( <i>x</i> ) for either quartile. c.a.o.	
	$\therefore \left(1 + \frac{l}{200}\right)^{-2} = \frac{3}{4}  \therefore l = 200 \left(\frac{2}{\sqrt{3}} - 1\right) = 30.94$	A1	c.a.o.	
	$\therefore IQR = 200 - 30.94 = 169(.06)$ For an outlier $x > UQ + 1.5 \times IQR = 200 + 1.5 \times 169$ =453(.58) $\approx$ 454 (nearest hour)	M1 M1 E1	UQ - LQ $UQ + 1.5 \times IQR$ . Answer given; must be obtained genuinely.	6
(ii) (A)	$F(x) = \int_{0}^{x} \frac{1}{200} e^{\frac{-t}{200}} dt$	M1	Correct integral, including limits (which may be implied subsequently).	
	$= \left[ -e^{\frac{-t}{200}} \right]_{0}^{x} = \left( -e^{\frac{-x}{200}} \right) - \left( -e^{\frac{-0}{200}} \right) = 1 - e^{\frac{-x}{200}}$	A1 E1	Correctly integrated. Limits used. Answer given; must be shown convincingly. Condone the omission of $x < 0$ part. Allow use of "+ c" with $F(0) = 0$ .	3
(B)	P(X > 50) = 1 - F(50)	M1	Use of $1 - F(x)$	
	$= e^{\frac{-50}{200}} = e^{-0.25}$	E1	Answer given: must be convincing. (= 0.7788(0))	2
(C)	$P(X > 400) = e^{\frac{-400}{200}} = 0.1353(35)$	B1	Accept any form.	
	$P(X > 450) = e^{\frac{-450}{200}} = 0.1053(99)$ $P(X > 450   X > 400) = \frac{P(X > 450)}{P(X > 400)}$	B1 M1	Accept any form. Conditional probability. Not $P(X > 50) \times P(X > 400)$ unless <u>clearly</u> justified.	
	$=\frac{e^{\frac{-450}{200}}}{e^{\frac{-400}{200}}}=e^{\frac{-50}{200}}=e^{-0.25} (=0.7788)$	A1	Accept division of decimals, 3dp or better. Accept a.w.r.t. 0.778 or 0.779.	4
				18

#### Mark Scheme

$\begin{array}{c c c c c c c c c c c c c c c c c c c $	-	$(10, 0.4^2),  D \sim N(35, 3.5^2)$	1 - 4 - 1	1 1	
$\begin{array}{c c c c c c c c c c c c c c c c c c c $					
$\begin{array}{ c c c c c } \hline & & & & & & & & & & & & & & & & & & $	(i) $P(C < $	$9.5) = P\left(Z < \frac{9.5 - 10}{0.4} = -1.25\right)$			
$\sigma^{2} = 3.5^{2} + (0.4^{2} + 0.4^{2} + 0.4^{2} + 0.4^{2}) = 12.89)$ Want P(D > S) = P(D - S > 0) $= 1 - \Phi\left(\frac{0 - (-5)}{3.59} = 1.39(27)\right)$ $= 1 - 0.9182 = 0.0818$ A1 c.a.o. $A1 c.a.o. A3 (iii) New (D - S) = (D \times 1.3) - (C_{1} + + C_{5}) \sim N(-4.5, \sigma^{2} = (3.5^{2} \times 1.3^{2}) + (0.4^{2} + + 0.4^{2}) = 21.5025)$ Again want P(D > S) = P(D - S > 0) $= 1 - \Phi\left(\frac{0 - (-4.5)}{4.637} = 0.9704\right)$ $= 1 - 0.8341 = 0.1659$ A1 $CI \text{ is given by } 9.73 \pm \frac{1.96}{\times \frac{0.4}{\sqrt{12}}}$ $= 9.73 \pm 0.2263 = (9.50(37), 9.95(63))$ Since 10 lies above this interval, it seems that the cheeses are underweight. In repeated sampling, 95% of all confidence intervals $F = 1 - \frac{1}{2} + \frac{1}$		= 1 - 0.8944 = 0.1056	A1	c.a.o.	3
$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	(ii) $D-S$	$= D - (C_1 + C_2 + C_3 + C_4) \sim N(-5,$	B1	Mean. Accept $+5$ for $S - D$ .	
Accept S - D < 0. This mark could be awarded in (iii) if not earned here. $= 1 - \Phi\left(\frac{0 - (-5)}{3.59} = 1.39(27)\right)$ A1Accept S - D < 0. This mark could be awarded in (iii) if not earned here.(iii) $New (D - S) = (D \times 1.3) - (C_1 + + C_5) \sim N(-4.5,$ $\sigma^2 = (3.5^2 \times 1.3^2) + (0.4^2 + + 0.4^2) = 21.5025)$ B1Mean. Accept +4.5 for S - D. Correct use of $\times 1.3^2$ for variance. 	$\sigma^2$ =	$= 3.5^{2} + (0.4^{2} + 0.4^{2} + 0.4^{2} + 0.4^{2}) = 12.89)$	B1	Variance. Accept sd (= 3.590).	
$= 1 - 0.9182 = 0.0818$ A1       c.a.o.       4         (iii) $New (D-S) = (D \times 1.3) - (C_1 + + C_5) \sim N(-4.5, \\ \sigma^2 = (3.5^2 \times 1.3^2) + (0.4^2 + + 0.4^2) = 21.5025)$ B1       Mean. Accept +4.5 for $S - D$ .       Correct use of $\times 1.3^2$ for variance.       Mathematic there.       Correct use of $\times 1.3^2$ for variance.       Correct use of $\times 1.3^2$ for vari	Want I	P(D > S) = P(D - S > 0)	M1	Accept S – D $<$ 0. This mark could be awarded in (iii) if	
$= 1 - 0.9182 = 0.0818$ A1       c.a.o.       4         (iii) $New (D-S) = (D \times 1.3) - (C_1 + + C_5) \sim N(-4.5, \\ \sigma^2 = (3.5^2 \times 1.3^2) + (0.4^2 + + 0.4^2) = 21.5025)$ B1       Mean. Accept +4.5 for $S - D$ .       Correct use of $\times 1.3^2$ for variance.       Mathematic there.       Correct use of $\times 1.3^2$ for variance.       Correct use of $\times 1.3^2$ for vari		$= 1 - \Phi\left(\frac{0 - (-5)}{359} = 1.39(27)\right)$			
$\sigma^{2} = (3.5^{2} \times 1.3^{2}) + (0.4^{2} + + 0.4^{2}) = 21.5025)$ $MI$ $Again want P(D > S) = P(D - S > 0)$ $= 1 - \Phi\left(\frac{0 - (-4.5)}{4.637} = 0.9704\right)$ $= 1 - 0.8341 = 0.1659$ $A1$ $C.a.o.$ $MI$ $CI is given by 9.73 \pm 1.96$ $\times \frac{0.4}{\sqrt{12}}$ $= 9.73 \pm 0.2263 = (9.50(37), 9.95(63))$ $A1$ $C.a.o.$		( ••••• )	A1	c.a.o.	4
MI Again want $P(D > S) = P(D - S > 0)$ MI AlCorrect use of $\times 1.3^2$ for variance. c.a.o. Accept sd (= 4.637)Again want $P(D > S) = P(D - S > 0)$ Or $S - D < 0$ . 	(iii) New (	$D-S$ = ( $D \times 1.3$ ) – ( $C_1$ + + $C_5$ ) ~ N(-4.5,	B1	Mean. Accept +4.5 for $S - D$ .	
$\begin{bmatrix} 1 & - \Phi \left( \frac{0 - (-4.5)}{4.637} = 0.9704 \right) \\ = 1 - 0.8341 = 0.1659 \end{bmatrix} = 1 - 0.8341 = 0.1659 \\ A1 \\ (iv) \\ CI is given by 9.73 \pm \\ 1.96 \\ \times \frac{0.4}{\sqrt{12}} \\ = 9.73 \pm 0.2263 = (9.50(37), 9.95(63)) \\ Since 10 lies above this interval, it seems that the cheeses are underweight. \\ In repeated sampling, \\ 95\% of all confidence intervals \end{bmatrix} \begin{bmatrix} M1 \\ B1 \\ M1 \\ CI is given by 9.73 \pm \\ CI is given$	$\sigma^2$ =	$= (3.5^2 \times 1.3^2) + (0.4^2 + + 0.4^2) = 21.5025)$			
$= 1 - 0.8341 = 0.1659$ A1c.a.o.4(iv)CI is given by $9.73 \pm$ $1.96$ $\times \frac{0.4}{\sqrt{12}}$ M1 B1 M11.96 seen. M11.96 seen. M11.96 seen. M1 $= 9.73 \pm 0.2263 = (9.50(37), 9.95(63))$ A1c.a.o. Must be expressed as an interval. E1E1Ft c's interval.Since 10 lies above this interval, it seems that the cheeses are underweight.E1Ft c's interval.E1	Again	want $P(D > S) = P(D - S > 0)$		M1 for formulation in (ii) available	
(iv)CI is given by $9.73 \pm$ $1.96$ M1 B1 M11.96 seen. $\times \frac{0.4}{\sqrt{12}}$ M1 B1 M11.96 seen. $= 9.73 \pm 0.2263 = (9.50(37), 9.95(63))$ A1 C.a.o. Must be expressed as an interval.Since 10 lies above this interval, it seems that the cheeses are underweight.E1In repeated sampling, 95% of all confidence intervalsE1		$= 1 - \Phi\left(\frac{0 - (-4.5)}{4.637} = 0.9704\right)$			
1.96B11.96 seen. $\times \frac{0.4}{\sqrt{12}}$ M11.96 seen.= 9.73 ± 0.2263 = (9.50(37), 9.95(63))A1c.a.o. Must be expressed as an interval.Since 10 lies above this interval, it seems that the cheeses are underweight.E1Ft c's interval.In repeated sampling, 95% of all confidence intervalsE1E1		= 1 - 0.8341 = 0.1659	A1	c.a.o.	4
$= 9.73 \pm 0.2263 = (9.50(37), 9.95(63))$ A1c.a.o. Must be expressed as an interval.Since 10 lies above this interval, it seems that the cheeses are underweight.E1Ft c's interval.In repeated sampling, 95% of all confidence intervalsE1E1	(iv) CI is g	$1.96 \times \frac{0.4}{\sqrt{1-100}}$	B1	1.96 seen.	
cheeses are underweight.       In repeated sampling,       95% of all confidence intervals       E1	= 9	-	A1	c.a.o. Must be expressed as an interval.	
95% of all confidence intervals E1			E1	Ft c's interval.	
constructed in this way will contain the true mean.E17			E1		
			E1		7

OCR (Oxford Cambridge and RSA Examinations) 1 Hills Road Cambridge CB1 2EU

**OCR Customer Contact Centre** 

#### 14 – 19 Qualifications (General)

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

