GCE

Mathematics (MEI)

Advanced GCE

Unit **4769:** Statistics 4

Mark Scheme for June 2011

Oxford Cambridge and RSA Examinations

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.

Any enquiries about publications should be addressed to:

OCR Publications
PO Box 5050
Annesley
NOTTINGHAM
NG15 0DL

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

4769 June 2011 Qu 1

$\mathrm{f}(x) = \frac{1}{\sqrt{2\pi\theta}} \mathrm{e}^{-x^2/2\theta}$ [N(0, θ)]	
(i) $L = \frac{1}{\sqrt{2\pi\theta}} \mathrm{e}^{-x_1^2/2\theta} \cdot \frac{1}{\sqrt{2\pi\theta}} \mathrm{e}^{-x_2^2/2\theta} \cdot \ldots \cdot \frac{1}{\sqrt{2\pi\theta}} \mathrm{e}^{-x_n^2/2\theta}$	M1 product form A1 fully correct
$\left[= (2\pi\theta)^{-n/2} \mathrm{e}^{-\Sigma x_i^2/2\theta} \right]$	Note. This A1 mark and the next five A1 marks depend on *all* preceding M marks having been earned.
$\ln L = -\frac{n}{2} \ln(2\pi\theta) - \frac{1}{2\theta} \sum x_i^2$	M1 for ln L A1 fully correct
$\frac{\mathrm{d} \ln L}{\mathrm{d}\theta} = -\frac{n}{2} \cdot \frac{1}{\theta} + \frac{1}{2\theta^2} \sum x_i^2$	M1 for differentiating A1, A1 for each term
$\frac{\mathrm{d} \ln L}{\mathrm{d}\theta} = 0$ gives $\frac{n}{2\hat{\theta}} = \frac{1}{2\hat{\theta}^2} \sum x_i^2$	M1 A1
i.e. $\hat{\theta} = \frac{1}{n} \sum x_i^2$	A1
Check this is a maximum. Eg:	M1
$\frac{\mathrm{d}^2 \ln L}{\mathrm{d}\theta^2} = \frac{n}{2} \cdot \frac{1}{\theta^2} - \frac{1}{\theta^3} \sum x_i^2$	A1
which, for $\theta = \hat{\theta}$, is $\frac{n}{2\hat{\theta}^2} - \frac{n}{\hat{\theta}^2} = -\frac{n}{2\hat{\theta}^2} < 0$.	A1 for expression involving $\hat{\theta}$ A1 for showing < 0 **[14]**
(ii) First consider E(X^2) = Var(X) + {E(X)}^{2} = θ + 0	M1 A1
$\therefore \mathrm{E}(\hat{\theta}) = \frac{1}{n}(\theta + \theta + \ldots + \theta) = \theta$	A1
i.e. $\hat{\theta}$ is unbiased.	A1 **[4]**
(iii) Here $\hat{\theta} = 10$ and Est Var$(\hat{\theta})$ = 2 × 10^2/100 = 2	B1, B1
Approximate confidence interval is given by $10 \pm 1.96\sqrt{2} = 10 \pm 2.77$, i.e. it is (7.23, 12.77).	M1 centred at 10 B1 1.96 M1 Use of √2 A1 c.a.o. Final interval **[6]**

4769 June 2011 Qu 2

(i) $n = 2$ $\quad \mathrm{f}(x) = \frac{1}{2}\mathrm{e}^{-x/2}$	
$\mathrm{M}(\theta) = \mathrm{E}(\mathrm{e}^{\theta X}) = \int_0^\infty \frac{1}{2}\mathrm{e}^{-x(\frac{1}{2}-\theta)}\mathrm{d}x$	A1 Any equivalent form
$= \frac{1}{2}\left[\frac{e^{-x(\frac{1}{2}-\theta)}}{-(\frac{1}{2}-\theta)}\right]_0^\infty$ **[A1]** $= \frac{\frac{1}{2}}{\frac{1}{2}-\theta}$ **[A1]** $= (1-2\theta)^{-1}$ **[A1]**	A1, A1, A1 for each expression, as shown, **beware printed answer**
$n = 4$ $\quad \mathrm{f}(x) = \frac{1}{4}x\mathrm{e}^{-x/2}$	
$\mathrm{M}(\theta) = \int_0^\infty \frac{1}{4}x\mathrm{e}^{-x(\frac{1}{2}-\theta)}\mathrm{d}x$	M1 for attempt to integrate this by parts
$= \frac{1}{4}\left\{\left[\frac{x\mathrm{e}^{-x(\frac{1}{2}-\theta)}}{-(\frac{1}{2}-\theta)}\right]_0^\infty \textbf{[A1]} - \int_0^\infty \frac{\mathrm{e}^{-x(\frac{1}{2}-\theta)}}{-(\frac{1}{2}-\theta)}\mathrm{d}x \textbf{ [A1]}\right\}$	A1, A1 for each component, as shown
$= \frac{1}{4}\left\{[0-0] \textbf{ [A1]} + \frac{1}{\frac{1}{2}-\theta}.2(1-2\theta)^{-1} \textbf{ [A1]}\right\}$	A1, A1 for each component, as shown
$= \frac{1}{2}\frac{1}{\frac{1}{2}(1-2\theta)}(1-2\theta)^{-1} = (1-2\theta)^{-2}$	A1 for final answer, **beware printed answer**
	[10]
(ii) Mean = M'(0) $\quad \mathrm{M}'(\theta) = -2(-\frac{n}{2})(1-2\theta)^{-\frac{n}{2}-1} = n(1-2\theta)^{-\frac{n}{2}-1}$	M1 A1
∴ mean = n	A1
Variance = M''(0) – {M'(0)}^{2}	
$\mathrm{M}''(\theta) = n(-\frac{n}{2}-1)(-2)(1-2\theta)^{-\frac{n}{2}-2} = n(n+2)(1-2\theta)^{-\frac{n}{2}-2}$	M1 A1
∴ M''(0) = $n(n + 2)$	A1
∴ variance = $n(n + 2) - n^2 = 2n$	A1
[**Note.** This part of the question may also be done by expanding the mgf.]	**[7]**

Solution continued on next page

4769 June 2011 Qu 2 **continued**

(iii)	By convolution theorem, $$M_W(\theta) = \left\{(1-2\theta)^{-\frac{1}{2}}\right\}^k = (1-2\theta)^{-k/2}.$$ This is the mgf of χ_k^2, so (by uniqueness of mgfs) $W \sim \chi_k^2$.	M1 B1 M1 B1 **[4]**
(iv)	$W \sim \chi_{100}^2$ has mean 100, variance 200. Can regard W as the sum of a large "random sample" of χ_1^2 variates. $$\therefore \mathrm{P}\left(\chi_{100}^2 < 118.5\right) \approx \mathrm{P}\left(\mathrm{N}(0,1) < \frac{118.5-100}{\sqrt{200}} = 1.308\right)$$ = 0.9045.	M1 for use of N(0,1) A1 c.a.o. for 1.308 A1 c.a.o. **[3]**

4769 June 2011 Qu 3

(i)	Type I error: rejecting null hypothesis **[B1]** when it is true **[B1]** Type II error: accepting null hypothesis **[B1]** when it is false **[B1]** OC: P(accepting null hypothesis **[B1]** as a function of the parameter under investigation **[B1]**) Power: P(rejecting null hypothesis **[B1]** as a function of the parameter under investigation **[B1]**)	8 separate B1 marks for components of answer, as shown Allow B1 out of 2 for P(...) Allow B1 out of 2 for P(...) P(Type II error | the true value of the parameter) scores B1+B1 P(Type I error | the true value of the parameter) scores B1+B1. "1 – OC" as definition scores zero. **[8]**
(ii)	$X \sim \mathrm{N}(\mu, 25)$ $\quad H_0: \mu = 94 \quad H_1: \mu > 94$ We require $0.02 = \mathrm{P}\left(\text{reject } \mathrm{H}_0 \mid \mu = 94\right) = \mathrm{P}\left(\bar{X} > c \mid \mu = 94\right)$ $= \mathrm{P}\left(\mathrm{N}(94, 25/n) > c\right) = \mathrm{P}\left(\mathrm{N}(0,1) > \dfrac{c-94}{5/\sqrt{n}}\right)$ $\therefore \dfrac{c-94}{5/\sqrt{n}} = 2.054$ We also require $0.95 = \mathrm{P}\left(\text{reject } \mathrm{H}_0 \mid \mu = 97\right)$ $= \mathrm{P}\left(\mathrm{N}(97, 25/n) > c\right) = \mathrm{P}\left(\mathrm{N}(0,1) > \dfrac{c-97}{5/\sqrt{n}}\right)$ $\therefore \dfrac{c-97}{5/\sqrt{n}} = -1.645$ $\therefore$ we have $c = 94 + \dfrac{10.27}{\sqrt{n}}$ and $c = 97 - \dfrac{8.225}{\sqrt{n}}$ Attempt to solve; c = 95.666 [allow 95.7 or awrt] $\sqrt{n}$ = 6.165, n = 38.01 Take n as "next integer up" from candidate's value	M1 M1 for first expression M1 for standardising B1 for 2.054 M1 for first expression M1 for standardising B1 for –1.645 M1 two equations A1 both correct (FT any previous errors) M1 A1 c.a.o. A1 c.a.o. A1 **[13]**
(iii)	Power function: step function from 0 with step marked at 94 to height marked as 1	G1 G1 G1 Zero out of 3 if step is wrong way round. **[3]**

4769 June 2011 Qu 4

(a)	Each E2 in this part is available as E2, E1, E0.	
(i)	Description of situation where randomised blocks would be suitable, ie one extraneous factor (eg stream down one side of a field).	E2
	Explanation of why RB is suitable (the design allows the extraneous factor to be "taken out "separately).	E2
	Explanation of why LS is not appropriate (eg: there is only one extraneous factor; LS would be unnecessarily complicated; not enough degrees of freedom would remain for a sensible estimate of experimental error).	E2
(ii)	Description of situation where Latin square would be suitable, ie two extraneous factors (and all with same number of levels) (eg streams down two sides of a field).	E2
	Explanation of why LS is suitable (the design allows the extraneous factors to be "taken out "separately).	E2
	Explanation of why RB is not appropriate (RB cannot cope with two extraneous factors).	E2 **[12]**
(b)	Totals are 56.5 57.4 60.6 82.3 from samples of sizes 4 3 5 4	
	Grand total 256.8 "Correction factor" CF = $256.8^2/16 = 4121.64$	
	Total SS = 4471.92 – CF = 350.28	M1 for attempt to form three sums of squares.
	Between treatments SS = $\frac{56.5^2}{4}+\frac{57.4^2}{3}+\frac{60.6^2}{5}+\frac{82.3^2}{4}-\text{CF}$	M1 for correct method for any two.
	= 4324.1103 – CF = 202.47	
	Residual SS (by subtraction) = 350.28 – 202.47 = 147.81	A1 if each calculated SS is correct.

Source of variation	SS	df	MS **[M1]**	MS ratio **[M1]**
Between treatments	202.47	3 **[B1]**	67.49	5.47(92) **[A1 cao]**
Residual	147.81	12 **[B1]**	12.3175	
Total	350.28	15		

5 marks within the table, as shown

Refer MS ratio to $F_{3,12}$.	M1 No FT if wrong
Upper 5% point is 3.49.	A1 No FT if wrong
Significant.	E1
Seems the effects of the treatments are not all the same.	E1 **[12]**

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