OCR
RECOGNISING ACHIEVEMENT

GCE

Mathematics

Advanced GCE

Unit **4733:** Probability and Statistics 2

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.

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

1		Number all the houses sequentially, or use house numbers Select using random numbers Ignore numbers > 263	B1 B1 B1	**3**	Any mention of using house numbers, or houses, or other numbering. (List can be implied). *Not* random numbering unless correct subsequent method (e.g. sort them numerically) Mention random numbers. *Not* "select numbers randomly". Must be random method. NB: Using 263 × calculator Rand # is biased: B0. But "Ran#(263)" is unbiased. Deal with problem of > 263, *or* repeats. "Select 20 random numbers between 1 and 263": B1B0 *[If this, need to mention repeats to get last B1]* *Example:* "put numbers/house names (etc) into hat and select": B1B0B0
2	α	$\mu = \frac{48+57}{2} = 52.5$ $\Phi^{-1}(0.9332) = 1.5$ $4.5 \div 1.5$ $[\sigma = 3]$	M1 A1 B1 M1		Use symmetry to find μ Obtain $\mu = 52.5$ 1.5 seen, e.g. in 4.5 ÷ 1.5 4.5 ÷ their Φ^{-1}, or 1.645 ÷ their Φ^{-1}, must be +ve, allow cc
	β	$\frac{57-\mu}{\sigma} = 1.5, \frac{48-\mu}{\sigma} = -1.5$ Solve simultaneously: $\mu = 52.5$ $[\sigma = 3]$	M1 A1 B1 A1		$\frac{57-\mu}{\sigma} = z$, $\frac{48-\mu}{\sigma} = -z$ M1 for one, ignoring cc, σ^2, sign or "1 –" errors, RHS must be Φ^{-1} (*not* Φ [e.g. 0.8246 or 0.5267] or 0.0668 or 0.9332); A1 for both completely correct except for value of z. $z = 1.5$ or -1.5 in at least one equation Solve without obvious errors, get $\mu = 52.5$, OK from wrong z [NB: 52.5 from both signs wrong: A0]
		$\mu + \frac{4.5}{1.5} \times 1.645$ $= \mathbf{57.4}(35)$	M1 B1 A1	**7**	$\mu + z\sigma$ *[Their μ and σ, anything recognisable as z]* [expect to see 52.5 + 3×1.645] $z = 1.645$ seen Answer in range [57.4, 57.45], cwo
3		CV $20 - \frac{5}{\sqrt{16}} \times 2.326 = 17.0925$ P(X > 17.0925) $= \Phi\left(\frac{17.0925-15}{5/\sqrt{16}}\right) = \Phi(1.674)$ Answer **0.0471**	M1 B1 A1 M1* A1 dep M1 A1	**7**	Attempt $20 - 5z/\sqrt{16}$, allow SD ↔ var errors, allow $20 \pm 5z/\sqrt{16}$, *not* $20 + 5z/\sqrt{16}$, allow cc 2.326 seen CV a.r.t. 17.1 *[NB: not 17.9075]* Standardise any attempt at a CV (from μ = 20) with 15 and any SD that would have got first M1, allow cc $z = 1.674$ seen or implied, e.g. by $p = 0.047$ or 0.953 or 0.9535, allow anything in range [1.67, 1.68] Probability < 0.5, or > 0.5 if their CV is < 15 Answer, a.r.t. 0.047 [including 0.0465 from CV 17.1] *Notes:* 16 missing: can get M0B1A0M1A0M1A0, or even last two A1's if 16 used then

4	(i)		M1 A1 **2**	Positive parabola, all above axis. [Don't worry about being pointed unless extreme.] Correct place, touches x-axis, not beyond the limits suggested by their axes, symmetric ends, not too straight
	(ii)	$\frac{3}{16}\int_0^4 x^2(x-2)^2\,dx$ $=\frac{3}{16}\left[\frac{x^5}{5}-x^4+4\frac{x^3}{3}\right]_0^4$ $[=6\frac{2}{5}]$ $\sigma^2 = 6\frac{2}{5}-2^2$ $=2\frac{2}{5}$	M1 M1 B1 B1 A1 **5**	Attempt $\int x^2 f(x)dx$, limits 0 and 4 Method for integration, e.g. multiply out *[indept]* [Or use $\sigma^2 = \frac{3}{16}\int_0^4 (x-2)^4\,dx$] Correct indefinite integral, limits not needed, e.g. parts: $\frac{3}{16}\left[\frac{x^2(x-2)^3}{3}-\frac{x(x-2)^4}{6}+\frac{(x-2)^5}{30}\right]$ Subtract 2^2 Final answer 2.4, any equivalent exact form, cwo
	(iii)	No because x represents a value taken by the random variable *[not an event that "occurs"]*	B1 **1**	Show clear understanding that x is a value of X. Usual misunderstanding is "X is an event that may or may not occur, depending on x". However: SR: Allow B1 for answer clearly indicating that probabilities higher where curve higher, *or* clearly stating that all probabilities are effectively zero. E.g.: "Agree as area under graph [or "f(x)"] increases", or "minimum at 2" B1 "True only between 0 and 4": B0 unless explanation Mention of variance etc: 0. "Agree because the graph shows this": B0
5	(i)	H_0: $p = 0.4$; H_1: $p < 0.4$ B(10, 0.4)	B1B1 M1	Both: B2. Allow π. One error, B1, but x or r: 0. ***SEE NOTES AT START AND END*** B(10, 0.4) stated or implied, e.g. N(4, 2.4) [*P(=1) [=0.0404] or P(≥ 1) [=0.9940] or P(<1) [=0.0060] or Poisson or normal, or RH tail for CR, gets no more marks in (i)*]
	α	P(≤ 1) = 0.0464 < 0.05 so reject H_0	A1 A1	This probability or 0.9536 only Explicit comparison with 0.05, or 0.9536 with 0.95
	β	CR is ≤ 1 and compare 1 Probability of this is 0.0464	A1 A1	Comparison needn't be explicit in this method This probability needs to be seen
		Reject H_0. Significant evidence that % who book with travel agents reduced	M1 A1√ **7**	Correct method, √, comparison and first conclusion Interpreted in context, "evidence that" or equiv needed, √ on numbers
	(ii)	Can't deduce cause-and-effect	B1 **1**	Equivalent comment, regardless of answer to (i). Ignore wrong answer if right answer seen "Other factors haven't been considered" B1 "Sample is small" , or "test may be wrong" B0

6	(i)	$H_0: \mu = 24.3$; $H_1: \mu \neq 24.3$	B1B1	Both: B2. 1 error, B1, but *t*, *x* etc: B0 ***SEE NOTES AT START AND END***
		$\bar{t} = 26.28$	B1	26.28 seen or implied
		$\hat{\sigma}^2 = \frac{50}{49}\left[\frac{36602.17}{50} - 26.28^2\right]$	M1	Correct formula for biased estimate [= 41.405]
			M1	Multiply by 50/49
				[Single formula: M2, or give M1 if wrong but 49 divisor seen]
		$= 42.25$	A1	42.25 or 6.5 seen or implied
	α	$z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154$	M1	Standardise their $\bar{t}$ with 24.3, √50, allow sign/√/cc errors, their variance
			A1	2.15(4) *or p* in range [0.0153, 0.0158], *not* –2.154 unless 0.015(6) subsequently used, *not* 1-tail
		< 2.576	A1	Compare *z* with ±2.576, or $p > 0.005$, or $2p$ with 0.01, *not* from $\mu = 26.28$
	β	CV $24.3 + 2.576 \times \sqrt{\frac{42.25}{50}}$	M1	24.3 + *zs*/√50, allow cc, √ errors, allow ± but not – only. Not 26.28 – *zs*/√50
			A1	*z* = 2.576, *not* from $\mu = 26.28$ or 50 omitted, *not* from 1-tail
		= 26.67 and 26.28 < 26.67	A1	Correct CV, ✓ on *z*, and compare sample mean
		Do not reject H_0. Insufficient evidence of a change in maximum daily temperature.	M1	Conclusion, ✓, needs method, like-with-like, 50, *not* from $\mu = 26.28$, *doesn't* need correct *z*
			A1✓ **11**	Contextualised, recognise uncertainty, ✓ on numbers
				NB: Clear evidence of $\mu = 26.28$: can't get last 4 marks. *See exemplars γ and δ*
	(ii)	*n* is large	B1 **1**	This answer ***only*** or "*n* > number" where number ≥ 29, *not* both this and "distribution unknown". But "*n* is large so we can approximate even though we don't know the distribution" is B1 "Possible as *n* = 50" B0.
7	(i)	Po(11)	M1	Po(11) stated or clearly implied
		$1 - P(\leq r) = 0.854$ gives $r = 14$	M1	Find 1 – 0.146 in tables, e.g. answer 14 [RH tail, e.g. "7", or single value only: max M1M0A0]
		so $n = 15$	A1 **3**	$n = 15$ only, allow "≥ 15"
	(ii)	Po(44) ≈ N(44, 44)	M1	Normal, mean attempted 2.2×20
			A1	Both parameters 44, allow var = √44 or 44^2
		$\Phi\left(\frac{37.5 - 44}{\sqrt{44}}\right) = \Phi(-0.980)$	M1	Standardise, their 44, allow cc,√ errors, e.g. ans 0.283 or 0.2036 or 0.4411, *not* ÷20
			A1	√ and cc both correct
		= **0.1635**	A1 **5**	Answer in range [0.163, 0.164]
	(iii)	B(40, 0.146)	M1	B(40, 0.146) stated or implied, e.g. by Po(5.84)
		≈ N(5.84, 4.98736)	M1	Normal, attempt at mean = *np* [Poisson etc, or exact binomial (0.22132): no more marks]
			A1	Both parameters correct [Poisson(5.84) → N(5.84, 5.84): M0A0]
		$1 - \Phi\left(\frac{7.5 - 5.84}{\sqrt{4.98736}}\right) = 1 - \Phi(0.7433)$	M1	Standardise with their *np* and *npq*, allow cc, √ errors, e.g. ans 0.3838 or 0.302 or 0.370
			A1	cc and √ both correct
		= **0.2286**	A1 **6**	Answer in range [0.228, 0.229]
				SC: B(40, 0.854) ≈ N(34.16, 4.98736): can get full marks, but if *R* > 7 used, max 3

8	(a) (i)	Several calls may all refer to the same incident	B1	**1**	Any reason showing correct understanding of "independent", but not just "singly" or equivalent. Ignore extra condition(s) unless clearly wrong in which case B0. Not "*fires*" independent. "Fires might spread" B0
	(ii)	Calls occur at constant average rate	B1		This condition only, allow "average" omitted, *not* "constant *probability*", *not* "random" unless clearly correct interpretation follows. No third condition unless fully justified by subsequent answer. Need contextualising *somewhere* in this part.
		E.g. No, because incidents are less/more common at night	B1	**2**	Any comment (with either yes or no) showing correct understanding, but "Fires might not occur at constant average rate" is not enough (gets B1 B0) "Different rates at different times of year": B0
	(b) (i)	$1-\left(1+2.74+\frac{2.74^2}{2!}\right)e^{-2.74}$	M1		Formula for any one correct Poisson probability for $r \geq 1$ [1 – (0.06457 + 0.17692 + 0.24238)]
			M1		Correct overall formula, allow 1 error (e.g. 1 term extra or missing or no "1 –")
		= 0.516(1)	A1	**3**	Answer, a.r.t. 0.516 [Interpolation (0.51604) or no working: B0 or B3]
	(ii)	$(e^{-2}\times 1)(e^{-3}\times 3)+(e^{-2}\times 2)(e^{-3}\times 1)$	M1		Two correct terms multiplied, or all 4 bits seen, e.g. .1353×.1494 + .2707×.4979 = 0.0202 + 0.0135
			A1		Correct expression
		= 0.0337	A1	**3**	Answer, a.r.t. 0.0337
	(iii)	$(e^{-\lambda}\times 1)(e^{-\mu}\times\mu)+(e^{-\lambda}\times\lambda)(e^{-\mu}\times 1)$	M1		Correct algebraic expression *[Ignore 1! throughout]*
		$= e^{-\lambda}\times e^{-\mu}(\lambda+\mu)$	M1		Take out factor of $e^{-\lambda}\times e^{-\mu}$ or equivalent essential step
		$= e^{-(\lambda+\mu)}(\lambda+\mu)$	A1		Correctly obtain exact answer [allow $e^{-\lambda-\mu}(\lambda+\mu)$]
		$= \mathrm{P}(T=1)$	A1	**4**	All correct, and write down correct formula for P(T = 1) [NB: T needed] Allow working towards middle
					SR: $\lambda = 2$, $\mu = 3$: Can get M1M1A1A0 if e^{-2} and e^{-3} retained. As soon as decimal approximations seen, no more marks.

Specific examples for question 5(i)

α	H_0: $p = 0.4$; H_1: $p < 0.4$ N(4, 2.4) P(≤ 1) = 0.0533 > 0.05 So do not reject H_0. Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A0 M0 **3**	δ	H_0: $p = 0.4$; H_1: $p < 0.4$ B(10, 0.4) P(≥ 1) = 0.9939 > 0.95 So reject H_0 Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A0 A0 M0 A0 **3**
β	H_0: $p = 0.4$; H_1: $p < 0.4$ B(10, 0.4) "P(= 1) = 0.0464" *[allow this]* < 0.05 So reject H_0 Insufficient evidence that % who book with travel agents reduced	B1B1 M1 A1 A1 M1 A0 **6**	ε	H_0: $p = 0.4$; H_1: $p \neq 0.4$ *[two-tailed]* B(10, 0.4) "P(= 1) = 0.0464" > 0.025 So do not reject H_0 Insufficient evidence that % who book with travel agents reduced	B1B0 M1 A1 A0 M1 A1 **5**
γ	H_0: $p = 0.4$; H_1: $p < 0.4$ B(10, 0.4) P(= 1) = 0.0404 *[look out for this]* < 0.05 so reject H_0 Significant evidence that % who book with travel agents reduced	B1B1 M1 A0 A0 M0 A0 **3**	ζ	H_0: $p = 0.4$; H_1: $p < 0.4$ B(10, 0.4) P(= 1) = 0.0464 *[no explicit comparison]* So reject H_0. Significant evidence that % who book with travel agents reduced	B1B1 M1 A1 A0 M1 A1 **6**

Specific examples for question 6(i)

α	H_0: $\bar{t} = 24.3$; H_1: $\bar{t} \neq 24.3$ *[wrong symbol]* $\bar{t}$ not seen explicitly *[implied by ...]* $\hat{\sigma}^2 = \left[\frac{36602.17}{50} - 26.28^2\right] = 41.405$ *[biased est]* $z = \frac{26.28 - 24.3}{\sqrt{41.405/50}} = 2.1758$ < 2.576 Accept H_0, maximum temp unchanged *[over-assertive, otherwise A1]*	B0B0 B1 M1 M0 A0 M1 A0 A1 M1A0 **5**	δ	$H_0 = 24.3$; $H_1 \neq 24.3$ *[missing symbol]* $\bar{t} = 26.28$ $\hat{\sigma}^2 = ... = 42.25$ $z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154$ *[loses 1]* > -2.576 Insufficient evidence to reject H_0. No change in maximum daily temperature. *[OK]*	B1 only B1 M1M1 A1 M1 A0 A1 M1 A1 **9**
β	H_0: $\mu = 26.28$; H_1: $\mu \neq 26.28$ *[WRONG]* $\bar{t} = 24.3$ *[explicitly]* $\hat{\sigma}^2 = ... = 42.25$ $z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154$ *[allow this – BOD]* < 2.576 Accept H_0. Insufficient evidence of a change in maximum daily temperature.	B0B0 B0 M1M1 A1 M1 A1 A1 M1 A1 **8**	ε	H_0: $\mu = 24.3$; H_1: $\mu > 24.3$ *[one-tail]* $\bar{t} = 26.28$ $\hat{\sigma}^2 = ... = 42.25$ $z = \frac{26.28 - 24.3}{\sqrt{42.25/50}} = 2.154$ < 2.326 Accept H_0. Insufficient evidence of a change in maximum daily temperature.	B1B0 B1 M1M1 A1 M1 A1 A0 M1 A1 **9**
γ	H_0: $\mu = 26.28$; H_1: $\mu \neq 26.28$ *[WRONG]* $\bar{t}$ not seen separately *[implied]* $\hat{\sigma}^2 = ... = 42.25$ $z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154$ *[DON'T allow this]* > -2.576 Accept H_0. Insufficient evidence of a change in maximum daily temperature.	B0B0 B1 M1M1 A1 M1 A0 A0 M0 A0 **5**	ζ	$z = \frac{24.3 - 26.28}{\sqrt{42.25/50}} = -2.154$ *but then…* So $p = 0.0156 > 0.005$ *[OK here]* Accept H_0. Insufficient evidence of a change in maximum daily temperature.	M1 A1 A1 M1 A1 **(11)**
			η	$z = \frac{26.28 - 24.3}{\sqrt{42.25}} = 0.3046$ *[no √50]* < 2.576 Accept H_0. Insufficient evidence of a change in maximum daily temperature.	M0 A0 A0 M0 A0 **(6)**

The following guidance notes are provided.

1 Standardisation using the normal distribution.

(a) In *stating* parameters of normal distributions, don't worry about the difference between σ and σ^2, so allow N(9, 16) or N(9, 4^2) or N(9, 4). When *calculating* $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$, the following mistakes are accuracy mistakes and not method mistakes so can generally score M1A0:
confusion of σ with σ^2 or √σ ; *n* versus √*n* ; wrong or no continuity corrections.

(b) Use of $\frac{\mu-\bar{x}}{\sigma}$ instead of $\frac{\bar{x}-\mu}{\sigma}$ is not penalised if it leads to a correct probability, but if the candidate is using a *z*-value in a hypothesis test, an answer of $z = -2.15$ when it ought to be 2.15 is an accuracy error and loses the relevant A1. When finding μ or σ^2 from probabilities, some candidates are taught to use $\frac{\mu-\bar{x}}{\sigma}$ whenever $\mu > \bar{x}$; provided the signs are consistent this gains full marks.

(c) Some candidates are taught to calculate, for example, P($X > 5$) from N(9, 16) by calculating instead P($X < 13$). This is a correct method, though it looks very strange the first time you see it.

(d) When calculating normal approximations to binomial or Poisson, use of the wrong, or no, continuity correction generally loses the last two marks: A0 A0.

2 Conclusions to hypothesis tests. There are generally 2 marks for these.

(a) In order to gain M1, candidates must not only say the correct "Reject/do not reject H_0" but have done the whole test in essence correctly apart from numerical errors. In other words, they must have compared their *p* value with a critical *p* value or other "like-with-like" (e.g. *not* say 0.0234 with 1.96), using the correct tail (e.g. *not* –2.61 with +2.576), and the working should in general have accuracy errors only. Thus miscalculation of *z*, comparison with 1.645 instead of 1.96, or using *n* instead of √*n*, or omission of a continuity correction when it is necessary, are all accuracy errors and the candidate can still gain the last M1 A1. Omission of √*n* where it is necessary is a method mistake and so gets M0. In hypothesis tests using discrete distributions, use of P(≤ 12) or P(> 12) or P(= 12) when it should be P(≥ 12) is a method mistake and usually loses all the final marks in a question.

(b) The A1 mark is for interpreting the answer *in the context of the question*, and *without over-assertiveness*. Thus "The mean number of applicants has increased" is over-assertive and gets A0 (although we allow "There is sufficient evidence to reject H_0. The mean number of applicants has increased", A1), and "There is sufficient evidence that the mean has increased" is not contextualised, so that too is A0.

(c) A wrong statement such as $-2.61 > -2.576$ generally gets B0 for comparison but can get the subsequent M1A1. Otherwise:

(d) If there is a self-contradiction, award M1 only if "Reject/Accept H_0" is consistent with their comparison. Thus if, say, we had $z = 2.61 > z_{crit} = 2.576$:
"Reject H_0, there is insufficient evidence that the mean number of … has changed" is M1A0.
but "Do not reject H_0, there is evidence that the mean number of … has changed" is M0A0.

(e) We don't usually worry about differences between "Reject H_0" and "Accept H_1" etc.

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