GCE

Further Mathematics A

Y542/01: Statistics

A Level

Mark Scheme for June 2023

OCR (Oxford Cambridge and RSA) is a leading UK awarding body, providing a wide range of qualifications to meet the needs of candidates of all ages and abilities. OCR qualifications include AS/A Levels, Diplomas, GCSEs, Cambridge Nationals, Cambridge Technicals, 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 2023

MARKING INSTRUCTIONS

PREPARATION FOR MARKING
RM ASSESSOR

1. Make sure that you have accessed and completed the relevant training packages for on-screen marking: *RM Assessor Online Training*; *OCR Essential Guide to Marking*.

2. Make sure that you have read and understood the mark scheme and the question paper for this unit. These are posted on the RM Cambridge Assessment Support Portal http://www.rm.com/support/ca

3. Log-in to RM Assessor and mark the **required number** of practice responses ("scripts") and the **number of required** standardisation responses.

MARKING

1. Mark strictly to the mark scheme.

2. Marks awarded must relate directly to the marking criteria.

3. The schedule of dates is very important. It is essential that you meet the RM Assessor 50% and 100% (traditional 40% Batch 1 and 100% Batch 2) deadlines. If you experience problems, you must contact your Team Leader (Supervisor) without delay.

4. **Annotations**

Annotation	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
BP	Blank Page
Seen	
Highlighting	

Other abbreviations in mark scheme	Meaning
dep*	Mark dependent on a previous mark, indicated by *. The * may be omitted if only one previous M mark
cao	Correct answer only
oe	Or equivalent
rot	Rounded or truncated
soi	Seen or implied
www	Without wrong working
AG	Answer given
awrt	Anything which rounds to
BC	By Calculator
DR	This question included the instruction: In this question you must show detailed reasoning.

5. **Subject Specific Marking Instructions**

a. Annotations must be used during your marking. For a response awarded zero (or full) marks a single appropriate annotation (cross, tick, M0 or ^) is sufficient, but not required.

For responses that are not awarded either 0 or full marks, you must make it clear how you have arrived at the mark you have awarded and all responses must have enough annotation for a reviewer to decide if the mark awarded is correct without having to mark it independently.

It is vital that you annotate standardisation scripts fully to show how the marks have been awarded.

Award NR (No Response)
- if there is nothing written at all in the answer space and no attempt elsewhere in the script
- OR if there is a comment which does not in any way relate to the question (e.g. 'can't do', 'don't know')
- OR if there is a mark (e.g. a dash, a question mark, a picture) which isn't an attempt at the question.

Note: Award 0 marks only for an attempt that earns no credit (including copying out the question).

If a candidate uses the answer space for one question to answer another, for example using the space for 8(b) to answer 8(a), then give benefit of doubt unless it is ambiguous for which part it is intended.

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 always 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, escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.

If you are in any doubt whatsoever 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, e.g. 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 method mark may usually be implied by a correct answer unless the question includes the DR statement, the command words "Determine" or "Show that", or some other indication that the method must be given explicitly.

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.

Unless otherwise indicated, marks once gained cannot subsequently be lost, e.g. 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, what is acceptable will be detailed in the mark scheme. If this is not the case please, escalate the question to your Team Leader who will decide on a course of action with the Principal Examiner.

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. We are usually quite flexible about the accuracy to which the final answer is expressed; over-specification is usually only penalised where the scheme explicitly says so.
   - When a value is given in the paper only accept an answer correct to at least as many significant figures as the given value.
   - When a value is not given in the paper accept any answer that agrees with the correct value to 3 s.f. unless a different level of accuracy has been asked for in the question, or the mark scheme specifies an acceptable range.
     NB for Specification B (MEI) the rubric is not specific about the level of accuracy required, so this statement reads “2 s.f”.

   Follow through should be used so that only one mark in any question is lost for each distinct accuracy error.
   Candidates using a value of 9.80, 9.81 or 10 for g should usually be penalised for any final accuracy marks which do not agree to the value found with 9.8 which is given in the rubric.

g. Rules for replaced work and multiple attempts:
   - If one attempt is clearly indicated as the one to mark, or only one is left uncrossed out, then mark that attempt and ignore the others.
   - If more than one attempt is left not crossed out, then mark the last attempt unless it only repeats part of the first attempt or is substantially less complete.
   - if a candidate crosses out all of their attempts, the assessor should attempt to mark the crossed out answer(s) as above and award marks appropriately.

h. For a genuine misreading (of numbers or symbols) which is such that the object and the difficulty of the question remain unaltered, mark according to the scheme but following through from the candidate’s data. A penalty is then applied; 1 mark is generally appropriate, though this may differ for some units. This is achieved by withholding one A or B mark in the question. Marks designated as cao may be awarded as long as there are no other errors. If a candidate corrects the misread in a later part, do not continue to follow through. Note that a miscopy of the candidate’s own working is not a misread but an accuracy error.

i. If a calculator is used, some answers may be obtained with little or no working visible. Allow full marks for correct answers, provided that there is nothing in the wording of the question specifying that analytical methods are required such as the bold “In this question you must show detailed reasoning”, or the command words “Show” or “Determine”. Where an answer is wrong but there is some evidence of method, allow appropriate method marks. Wrong answers with no supporting method score zero. If in doubt, consult your Team Leader.

j. If in any case the scheme operates with considerable unfairness consult your Team Leader.

Question		Answer	Marks	AO	Guidance
1	(a)	A geometric distribution should still be a good model as the number of books is large/can be taken to be infinite/etc.	B1 [1]	3.3	OE. Allow "population is large" but *not* "n is large" or "sample is large". "Likely to finish long before lecturer gets to the end": B1 "Only a good model if the books are selected with replacement": B0
	(b)(i)	Whether or not one book refers to the topic is independent of whether another does *or* the probability that one book refers to the topic is the same for any book *or* arrangement/selection of books is random	B1 [1]	3.3	One correct modelling assumption, stated in context (e.g. "each book is independent", "the probability of finding the book is constant", etc) Allow any number of correct statements but B0 if any definitely incorrect or 'scattergun' statements seen, e.g. *not* "a book must either refer to the topic or not"
	(b)(ii)	E.g. arrangement/selection of the books not independent of whether they refer to the topic, *or* books containing the topic may be grouped together, *or* lecturer chooses books by author, etc	B1 [1]	3.3	Reason why it might not be valid. Needn't be very plausible but must show correct understanding of the concept. *Not* "sampling without replacement means that the probabilities change" (part **(a)** focusses on the fact that, although the [conditional] probabilities change, they do so by only a very small amount and so the model is not invalidated)
	(c)	$1 - (1 - 0.05)^n > 0.9$	M1*	3.3	Correct expression soi, allow $n - 1$, allow any of $>, \geq, =$
		$n > \log_{0.95}(0.1)$ *or* $\log(0.1) \div \log(0.95)$ etc	depM1	2.1	Correct solution method, e.g. use logarithms, ignore inequality (T&I: 44 or 46 gets M1M1A0)
		$n > 44.(8)9$, so minimum 45 books	A1 [3]	3.4	45 only, can be from T&I (command word is "find"). *Not* from 0.95^{n-1} (giving $n - 1 > 44.9$) but allow "$n < 44.9$ so 45" BOD. *Not* $n > 45$. SC: N(20, 380) giving 45: B2 SC: 45 following wrong working and first M1 not gained: B1
	(d)	3 in first 39, then 1 $\Rightarrow {}^{39}C_3 \times 0.95^{36} \times 0.05^4$	M1	3.1b	Allow if any indication of B(39, 0.05) used (but NB: 0.0901 from B(40, 0.05) is M0A0).
		$= 0.009\,01(215)$	A1 [2]	1.1	Correct to 3sf, www

Question		Answer	Marks	AO	Guidance
2	(a)	$h = 0.322p + 1.64$ (1.636) or $h = \frac{975}{596} + \frac{961}{2980}p$	B1 B1 [2]	1.1 1.1	a in range [1.63, 1.64] or b in range [0.322, 0.323] All correct including h and p, but allow if a and b correct and numerical or sign error made in writing out final equation. NB: *not* $b = 1.66$ SC: $h = 32.2p + 164$: B1
	(b)	New equation is $h = 0.0322p + 0.1635$ *or* "both a and b divided by 10"	B1ft [1]	2.2a	Both their coefficients divided by 10, ignore letters
	(c)	17.8 hundred	B1 [1]	1.1	In range [17.7, 17.8] hundred, *or* in range [1770, 1780], *or* 1.8 thousand. *Not* just 17.8 or 1.8.
	(d)	Fair correlation only (so only fairly reliable) 50 is in range of data e.g. audiences may be different for a charity event, *or* attendance depends on causal factors, *or* not based on a random sample, *or* sample size small, etc Overall not very reliable	B1 B1 B1 B1 [4]	1.1 1.1 2.4 2.3	Any comment based on size of r, allow comparison with CV but *not* with b, allow 'association' State that 50 is in the data range, but *not just* "50 is not one of the data values" Any reasonable relevant comment based on context or sampling, but *not* "correlation not causality", or similar rote statement. *Final* nuanced conclusion between "fairly reliable" and "not reliable" inclusive, based on at least two or three of the above, but if they use 0.642 > CV (or other wrong statement) this cannot count towards the 'two or three'. Not just separate assessments for each statement [Comparison with critical values is not relevant. The issue is *not* "is there any correlation?" (i.e., is $\rho = 0$?) but how strong is the correlation (is ρ close to 1?), and hypothesis tests don't tell you this.]

Exemplar

A	0.642 quite high so fairly reliable; 50 is in data range so reliable; charity event not typical so less reliable *(maximum for 3/4 – i.e., don't give final B1 for successively, e.g., "fairly reliable, more reliable, less reliable")*	B1B1B1B0

Question		Answer	Marks	AO	Guidance
3	(a)	$E(W) = 6$, $Var(W) = 10$, $E(V) = 3$, $Var(V) = 2$	B1	1.1	All four, can be implied by subsequent working
		$6m + 3n = 0$	B1ft	2.1	Ft on their means
		$10m^2 + 2n^2 = 1$	B1ft	2.1	Ft on their variances
		$10m^2 + 2.(4m^2) = 1 \Rightarrow m = \frac{\sqrt{2}}{6}$ (or $\frac{1}{\sqrt{18}}$)	M1	1.1	Solve simultaneously to get one letter, allow $n > 0$ if found first
		$n = -\frac{\sqrt{2}}{3}$ (or $-\frac{2}{\sqrt{18}}$)	A1 [5]	2.2a	Both correct ($n = -2m$), any exact form (*not* ISW on both signs as question requires $m > 0$)
	(b)	No, as the number of independent random variables being averaged is not big enough for the Central Limit Theorem to apply.	B1 [1]	2.4	State "no", with correct reason, e.g. "3 is not big enough" or "sample is not big enough" but *not* "n is not big enough" unless explicit that $n = 3$. Allow "no as $3 < 25$", but if a number is quoted it must be 25 or 30. No wrong extras.

Question		Answer	Marks	AO	Guidance
4	(a)	E.g. ratings are arbitrary, *or* population not bivariate normal, *or r* measures only linear correlation	B1 [1]	2.4	*Not* "data are ratings not scores" unless explained further. *Not* "population may not be normal", nor *just* "looking for agreement not correlation"
	(b)	I 5 4 3 2 1 II 4 5 2 3 1 $\Sigma d^2 = 4$ $r_s = 1 - \frac{6\Sigma d^2}{5(5^2-1)} = 0.8$	B1 B1 M1 A1 [4]	1.1 1.1 1.2 1.1	Or with rankings reversed (2 1 4 3 5) 4 stated or implied Correct formula used, with reasonable attempt at Σd^2 using rankings 0.8 or 4/5 only
	(c)	H_0: $\rho_s = 0$, H_1: $\rho_s > 0$, where ρ_s is the population rank correlation coefficient between the rankings given by the two magazines, *or* H_0: no correlation between population rankings, H_1: positive correlation 0.8 < 0.9 so do not reject H_0. Insufficient evidence of agreement between magazines' opinions.	B2 B1ft B1ft [4]	1.1 2.5 1.1 2.2b	One error, e.g. two-tailed or ρ_s not defined, *or* verbal: B1 If verbal, must include clearly one-sided H_1 for B2 Allow B1 for H_0: $\rho = 0$, H_1: $\rho > 0$. Same marks also for using r or r_s For B2 must include either context or "population" (or both). Allow any of correlation, association, agreement. Compare with 0.9 and do not reject, FT on their r_s if M1 gained in **(b)**. FT on CV 1.0 if hypotheses two-tailed. Allow "Accept H_1", etc Contextualised, not too assertive, FT as above. *Not* "evidence of no agreement", but allow from CV 1.0. Needs hypotheses right way round
	(d)	Magazine III 2 1 4 3 5	B1ft [1]	3.1b	Reverse of rankings of II used in part **(b)** (= 6 – their rankings of II) NB: "4 5 2 3 1" could come from ranking high-to-low.

Exemplars

(a)

A	Takes into account magnitudes of scores *(wrong)*	B0
B	Takes into account differences in ranks *(wrong)*	B0
C	Not enough data to use PMCC	B0
D	One magazine may have a stricter standard and Spearman would eliminate this *(scaling would not affect PMCC)*	B0
E	Spearman's Rank will not be as much affected by uncertainties in measurements/anomalous results *(too vague?)*	B0
F	Eliminates the magnitude of the scores *(too vague)*	B0
G	Testing association rather than a linear relationship *(generally allow any answer that says SRCC doesn't need linearity)*	B1
H	Looking for agreement between opinions rather than between numerical values, so association not correlation	B1
I	The numbers are just opinions	B1

(c)

J	H_0: no correlation between magazine's ratings, H_1: positive correlation	B2
K	H_0: no correlation, $\rho_s = 0$, H_1: positive correlation, $\rho_s > 0$ *(neither the verbal nor the symbolic statement scores B2)*	B1
L	H_0: $r = 0$, H_1: $r > 0$, where r is the correlation coefficient between the rankings given by the magazines	B2
M	H_0: no agreement between population rankings, H_1: there is agreement *(2-tailed so B1B0)* $0.8 < 1.0$ so do not reject H_0. *(FT on 2-tailed)* Insufficient evidence of no agreement between magazines' opinions. *(FT)*	B1B0 B1 B1
N	H_0: no positive agreement between population rankings, H_1: positive agreement *(1-tailed so B1B1 – allow this H_0)* $0.8 < 1.0$ so do not reject H_0. *(NO FT on 1-tailed if 0.9 not used)* Insufficient evidence of no agreement between magazines' opinions. *(NO FT)*	B1B1 B0 B0
O	H_0: no agreement between population rankings, H_1: there is agreement *(2-tailed so B1B0)* $0.8 < 0.9$ so do not reject H_0. *(allow, even though inconsistent)* Significance evidence of agreement between magazines' opinions. *(can't assert that H_0 is correct)*	B1B0 B1 B0

Question		Answer	Marks	AO	Guidance
5	(a)	H_0: distributions of ages for urban and rural areas are identical; H_1: distributions differ as to median *Or* H_0: $m_d = 0$, H_1: $m_d > 0$, where m_d is the median of population differences of age of marriage for men in urban and rural areas *or* H_0: population median ages in urban & rural areas are equal, H_1: median age higher in urban areas	**B1**	1.1	If "distributions identical", allow any type of average. If not "distributions identical" then must use median. Allow use of m_u and m_r, provided they are defined as median, or entirely verbal (see below) but need some context
		Rankings of rural ages are 1, 2, 4, 7, 9, 11	**M1**	3.1b	Find rankings of rural ages within whole sample, can be implied by 34
		$R_m = 34$ and $m(m + n + 1) - R_m = 56$	**B1**	1.1	Consider $m(m + n + 1) - R_m$ (if omitted, can get all other marks)
		$W = 34$	**A1**	1.1	Correct W, allow $p = 0.0876$ from N(45, 60) with cc
		CV 31	**B1**	1.1	Correct CV used, allow 31 or 31.76 from N(45, 60)
		34 > 31 so do not reject H_0, *or* 0.0876 > 0.05	**M1ft**	1.1	Comparison and correct first conclusion, needs correct method for R_m. FT on their 34, but not 56, and on 29 instead of 31 but no other CV except FT on wrong m or n (e.g. $m = 5$, CV 23 which gets B1M1B0A1B0M1A1)
		Insufficient evidence that average ages are higher	**A1ft** **[7]**	2.2b	Contextualised, not over-assertive, not "significant evidence that average ages are the same". Same ft. Allow "different".
If from valid method their $W \le 31$, so that conclusion changes, FT for **(a)** ("significant evidence that average ages are different") and *consult PE* if **(b)** is problematic					
	(b)	1 + 2 + 4 + 7 + 8 + 10 = 32	**M1**	3.1b	Attempt to change total to 32 (or 30 if CV 29 used), e.g. 8 seen
			A1	1.1	Find age that makes rank sum of 32 (or 31 if 29 used – can't get 30??) (M1A1 can be implied by answer of 20/0, allow 19/12)
	OR	T&I: Test at least two new ages and find R_m for each 19/10 or 20/0 or 19/12 stated	**M1** **A1**		
		11th becomes 7th so least age of Mr X is 19/10	**A1** **[3]**	3.2a	(old 11th becomes new 7th, old 7th becomes new 8th) *Condone* absence of working.

A	H_0: median in urban & rural areas are equal, H_1: median higher in urban areas *(no context – needs "ages" as a minimum)*	B0
B	H_0: average ages in urban & rural areas are equal, H_1: average age higher in urban areas *(minimum context, but "average" not median")*	B0
C	H_0: median ages in urban & rural areas are equal, H_1: median age higher in urban areas *(condone omission of "population")*	B1
D	H_0: median of differences in ages is 0, H_1: greater than 0 *(BOD for their meaning "urban minus rural" rather than "rural minus urban")*	B1
E	H_0: $m_u = m_r$, H_1: $m_u > m_r$ where m_u and m_r are the median ages of marriage for men in urban and rural districts respectively	B1

Question		Answer	Marks	AO	Guidance
6		CDF of X is (F(x) =) $\frac{\pi + x}{2\pi}$	M1 A1	2.1 1.1	Attempt to find CDF of X, e.g. $x/2\pi$ Correct
		P($Y < y$) = P(sin $X < y$) (= P($X < \sin^{-1} y$))	M1	2.1	Allow for F($\sin^{-1} y$), and allow inequality errors
		$+ \frac{\sin^{-1} y}{2\pi}$	M1	2.1	Method for dealing with ranges
	OR	P($Y < y$) = P(sin $X < y$) = F($\sin^{-1} y$) where F is CDF of X Method for dealing with ranges + ½	M1 M1 M1 A1		 OE
		$\frac{\pi + 2\sin^{-1} y}{2\pi}$	A1	2.2a	Any equivalent form, WWW
		Range of Y is $-1 \le y \le 1$	B1	1.1	Stated. NB $-\pi \le \sin^{-1} y \le \pi$ is insufficient, same letter as in function
		$F(y) = \begin{cases} 0 & y < -1, \\ \frac{\pi + 2\sin^{-1} y}{2\pi} & -1 \le y \le 1, \\ 1 & y > 1. \end{cases}$	A1ft [7]	1.2	Correct including 0 and 1, with ranges, but allow = in wrong places, FT on their $\frac{\pi + 2\sin^{-1} y}{2\pi}$. Must be in terms of y (if F(y), or x if F(x), etc) (*not* $-\pi < \sin^{-1} y < \pi$ etc), depends on second M1

Typically, $\frac{\sin^{-1} y}{2\pi}$ can get M1, A0, M1, M0A0, B1 (unlikely), A0. $\frac{\sin^{-1} y}{\pi}$ can get M1A0 M1 M1 A0 B1 A0

$\frac{\pi + \sin^{-1} y}{2\pi}$ can get M1A1, M1M0A0, B1A1

Correct but wrong ranges loses final B1A1

Question		Answer	Marks	AO	Guidance
7	(a)	$\hat{\mu} = \bar{x} = 9.9$	B1	1.1	9.9 seen anywhere
		$s^2 = \frac{4271.40}{40} - 9.9^2$ [= 8.775]	M1	2.1	Correct method for biased or unbiased estimate used
		$\hat{\sigma}^2 = \frac{40}{39} \times 8.775$	M1	1.2	×40/39, or divisor 39, seen anywhere
		= 9	A1	1.1	9 only (or 3 if clear)
		$9.9 \pm z\sqrt{\frac{9}{40}}$	M1	3.1b	*Their* 9.9 and 9. 40 and numerical z must be seen or implied. Allow √ errors
		z = 2.576, allow 2.58	B1	1.1	Explicit, or implied by correct answer only (*not* for just "$Z_{0.995}$")
		9.9 ± 1.22 = (8.68, 11.12)	A1 [7]	1.1	Both, needs 40/39 oe. Allow awrt 8.68, and awrt 11.1. Use of 8.775 gives 8.69: B1M1 M0A0 M1B1 A0 Use of 2.236 gives (8.80, 11.0): 5/7.
	(b)	$\bar{X} \pm 2.576\sqrt{10/50}$ $(= \bar{X} \pm 1.152)$	M1*	3.1b	Find new confidence interval in terms of $\bar{X}$ (not μ), needn't use 50 here 9.9 ± 1.152 or (8.75, 11.05) on their own are M0, but see SCs below
		$P(\bar{X} - 1.152 < \mu + 1.6 < \bar{X} + 1.152)$ $= P\left(\frac{0.448}{\sqrt{0.2}} < \frac{(\bar{X} - \mu)}{\sqrt{0.2}} < \frac{2.752}{\sqrt{0.2}}\right)$	depM1	1.1	Obtain at least LH inequality, 50 needed here, allow √ errors
		= Φ(6.15) – Φ(1.002) = 0.158(2)	A1 [3]	1.1	Awrt 0.158 (No marks *just* for finding new interval, or 10.348 & 12.472) If RH inequality not considered *at all*, withhold final A1
		SC: *Explicitly* take specific value of μ, e.g. 0 or 9.9 (must be μ not $\bar{X}$, and *not* just $\mu \sim$ N(9.9, 0.2): first M1* as above; e.g. $P(0.448 < \bar{X} < 2.752)$ from $\mu = 0$ or $P(10.348 < \bar{X} < 12.472)$ from $\mu = 9.9$ depM1; 0.158(2) A1			
		SC: 9.9 ± 1.152 = (8.748, 11.05); $\mu \sim$ N(9.9, 0.2) M0 so far. P(8.748 – 1.6 < μ < 11.052 – 1.6) M1 = 0.158 A1 max 2/3			

Question		Answer	Marks	AO	Guidance
8	(a)	1.6	B1 [1]	1.1	
	(b)	1.94	B1 [1]	1.1	
	(c)	These are fairly but not very close, so the Poisson model may be good but there is some doubt.	B1ft [1]	3.5a	Nuanced conclusion, not just yes/no, based on whether mean and variance are *close,* not just 1.6 ≠ 1.94. FT on their *s* but must be nuanced
	(d)	The expected frequencies for 4 and ≥ 5 are both < 5	B1 [1]	2.4	State "two cells have expected frequencies < 5", or specify at least one cell with expected frequency < 5. Not general statement, nor "some are ..."
	(e)	Po(1.6)	B1	3.3	Stated, or implied by formula
		$P(R = 2) = e^{-1.6}\frac{1.6^2}{2!}$	M1	1.1	Correct formula used, or *at least* 0.25843 or 15.5056… seen, SC: 60 × 0.258 *or* 60 × 0.2584 without working for prob: M1A0
		Expected frequency is 60 × 0.25843 = 15.506 **AG**	A1	3.4	Correctly obtain 15.506, e.g. 15.5056 seen **AG** so full working necessary
		$(15.506 – 12)^2/15.506 = 0.793$ **AG**	B1 [4]	1.1	Independent, but need to see formula used **AG**
	(f)	H_0: Population of number of calls has a Poisson distribution. H_1: it doesn't	B1	1.1	Needs general Poisson and *not* Po(1.6). *Not* "evidence that ...", *not* "data follows Poisson distribution", allow "data fits Poisson distribution"
		$\nu = 2$ CV is 5.991	M1 A1	1.1 1.1	Allow $\nu = 3$ for M1 5.991, allow either 5.991 or 7.815 if hypotheses mention 1.6
		(Their) 6.99 > 5.991 so reject H_0 (*or* Accept H_1)	M1ft	1.1	Allow if 6.99 wrong or not explicit; FT on 7.815 but no other CVs.
		There is significant evidence that the data is not well modelled by a Poisson distribution	A1ft [5]	2.2b	Contextualised, not over-assertive, needs 5.99(1) seen. If 7.815 used can get A1 with conclusion reversed: insufficient evidence that it doesn't fit. Needs hypotheses right way round for A1 but can be poorly stated
	(g)	$a + 2b = 18$, $a + b = 13$, $a = 12$, $a – b = 9$, $a – 2b = 8$	M1	3.1b	Consider at least three of these (e.g. $a = 12$ and two other equations), allow from ×60 (e.g. $a + 2b = 1080$), allow sum of probs = 1 for one eqn
		e.g. $(a, b) = (12, 3)$ [→ 18, 15, 12, 9, 6]; (11, 3) [→ 18, 13, 12, 9, 4], and (12, 3) is better	B1	3.4	Consider at least two different possibilities and select one, allow from ×60 (or, if equations used are $a + 2b = 18$, $a = 12$, $a – 2b = 8$, check third eqn)
		E.g. a must be 12 and $b = 3$ gives a close approximation to the frequencies	A1 [3]	3.5c	$a = 12$ only, $b = 1$ or 2 or 3, with any relevant reason. SC: if M0B0, give SC B1 for (12, 3) or (12, 2)

Exemplars

(f)	A	H_0: Number of calls has a Poisson distribution. H_1: it doesn't √	**B1**
		$\nu = 3$ √ CV is 7.815 **X**	**M1** **A0**
		6.99 < 7.815 so accept H_0 **FT** √	**M1ft**
		There is significant evidence that the data has a Poisson distribution **X** *(could get this last A1 if conclusion was "insufficient evidence that it doesn't have a Poisson distribution")*	**A0** **[3/5]**
	B	H_0: Number of calls fits Po(1.6). H_1: it doesn't **X**	**B0**
		$\nu = 3$ √ CV is 7.815 **FT** √	**M1** **A1**
		6.99 < 7.815 so do not reject H_0 **FT** √	**M1ft**
		There is insufficient evidence that the data does not have a Poisson distribution √ **BOD** *(ignoring the 1.6)*	**A1ft** **[4/5]**
	C	H_0: Number of calls fits Po(1.6). H_1: it doesn't **X**	**B0**
		$\nu = 2$ √ CV is 5.991 √ *(allow this even though it's not consistent with their hypotheses)*	**M1** **A1**
		6.99 > 5.991 so accept H_1 √	**M1ft**
		There is significant evidence that the data does not fit Po(1.6) √ **BOD** *(again ignoring the 1.6)*	**A1** **[4/5]**
	D	H_0: Data doesn't fit Poisson distribution. H_1: it does **X**	**B0**
		$\nu = 2$ √ CV is 5.99 √	**M1** **A1**
		6.99 > 5.99 so reject H_0 √	**M1ft**
		There is insufficient evidence that the data does not have a Poisson distribution **X** *(can't get final A1 if hypotheses wrong way round)*	**A0** **[3/5]**

Need to get in touch?

If you ever have any questions about OCR qualifications or services (including administration, logistics and teaching) please feel free to get in touch with our customer support centre.

Call us on

01223 553998

Alternatively, you can email us on

support@ocr.org.uk

For more information visit

ocr.org.uk/qualifications/resource-finder

ocr.org.uk

Twitter/ocrexams

/ocrexams

/company/ocr

/ocrexams

CAMBRIDGE
UNIVERSITY PRESS & ASSESSMENT

OCR is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored. © OCR 2023 Oxford Cambridge and RSA Examinations is a Company Limited by Guarantee. Registered in England. Registered office The Triangle Building, Shaftesbury Road, Cambridge, CB2 8EA.

Registered company number 3484466. OCR is an exempt charity.

OCR operates academic and vocational qualifications regulated by Ofqual, Qualifications Wales and CCEA as listed in their qualifications registers including A Levels, GCSEs, Cambridge Technicals and Cambridge Nationals.

OCR provides resources to help you deliver our qualifications. These resources do not represent any particular teaching method we expect you to use. We update our resources regularly and aim to make sure content is accurate but please check the OCR website so that you have the most up-to-date version. OCR cannot be held responsible for any errors or omissions in these resources.

Though we make every effort to check our resources, there may be contradictions between published support and the specification, so it is important that you always use information in the latest specification. We indicate any specification changes within the document itself, change the version number and provide a summary of the changes. If you do notice a discrepancy between the specification and a resource, please contact us.

Whether you already offer OCR qualifications, are new to OCR or are thinking about switching, you can request more information using our Expression of Interest form.

Please get in touch if you want to discuss the accessibility of resources we offer to support you in delivering our qualifications.