CAMBRIDGE
International Education

Cambridge O Level

STATISTICS **4040/22**

Paper 2 **October/November 2024**

MARK SCHEME

Maximum Mark: 100

Published

This mark scheme is published as an aid to teachers and candidates, to indicate the requirements of the examination. It shows the basis on which Examiners were instructed to award marks. It does not indicate the details of the discussions that took place at an Examiners' meeting before marking began, which would have considered the acceptability of alternative answers.

Mark schemes should be read in conjunction with the question paper and the Principal Examiner Report for Teachers.

Cambridge International will not enter into discussions about these mark schemes.

Cambridge International is publishing the mark schemes for the October/November 2024 series for most Cambridge IGCSE, Cambridge International A and AS Level components, and some Cambridge O Level components.

This document consists of **11** printed pages.

Generic Marking Principles

These general marking principles must be applied by all examiners when marking candidate answers. They should be applied alongside the specific content of the mark scheme or generic level descriptions for a question. Each question paper and mark scheme will also comply with these marking principles.

GENERIC MARKING PRINCIPLE 1:

Marks must be awarded in line with:

- the specific content of the mark scheme or the generic level descriptors for the question
- the specific skills defined in the mark scheme or in the generic level descriptors for the question
- the standard of response required by a candidate as exemplified by the standardisation scripts.

GENERIC MARKING PRINCIPLE 2:

Marks awarded are always **whole marks** (not half marks, or other fractions).

GENERIC MARKING PRINCIPLE 3:

Marks must be awarded **positively**:

- marks are awarded for correct/valid answers, as defined in the mark scheme. However, credit is given for valid answers which go beyond the scope of the syllabus and mark scheme, referring to your Team Leader as appropriate
- marks are awarded when candidates clearly demonstrate what they know and can do
- marks are not deducted for errors
- marks are not deducted for omissions
- answers should only be judged on the quality of spelling, punctuation and grammar when these features are specifically assessed by the question as indicated by the mark scheme. The meaning, however, should be unambiguous.

GENERIC MARKING PRINCIPLE 4:

Rules must be applied consistently, e.g. in situations where candidates have not followed instructions or in the application of generic level descriptors.

GENERIC MARKING PRINCIPLE 5:

Marks should be awarded using the full range of marks defined in the mark scheme for the question (however; the use of the full mark range may be limited according to the quality of the candidate responses seen).

GENERIC MARKING PRINCIPLE 6:

Marks awarded are based solely on the requirements as defined in the mark scheme. Marks should not be awarded with grade thresholds or grade descriptors in mind.

NOTES FOR MARKERS

Awarding marks

M marks are for method and are not lost for purely numerical errors.
A marks are for accuracy and depend on a correct method.
B marks are independent of method.

Once an acceptable answer is seen, ignore subsequent working, except where such working illustrates a conceptual misunderstanding.

No response

If there is any attempt at a solution award 0 marks not NR. '-' or '?' constitute no attempt at a solution.

Abbreviations

AG answer given (on question paper)
awrt answer which rounds to
cao correct answer only
dep mark depends on earlier, asterisked (*), mark
ft follow through (from earlier error)
oe or equivalent
SC special case
soi seen or implied
nfww not from wrong working

Annotations

Where a candidate's correct final answer is both clear and clearly identified (given on the answer line, encircled, underlined etc.), it is not necessary to annotate that item; nor is it necessary to annotate when there is No Response.

Where there is a response that scores 0, either SEEN should be used, or some other annotation(s) to indicate why no marks can be awarded.

Partial credit should be indicated with an annotation at the point at which that mark has been earned.

The highlighter should be used anywhere it is helpful to clarify the marking.

	Correct item
	Incorrect item
M1	M0, M1, M2 Method marks awarded
A1	A0, A1, A2 Accuracy marks awarded
B1	B0, B1, B2 Independent marks awarded
SC	Special case
FT	Correct follow through
ISW	Ignore subsequent working
BOD	Benefit of doubt
MR	Misread
TE	Transcription error
^	Essential element of answer/working missing
XP	Incorrect process
SEEN	Working seen but no credit awarded; blank page checked
Highlight	Use anywhere it is helpful to clarify the marking

Question	Answer	Marks	Partial Marks
1(a)	<table><tr><th></th><th>Not a variable</th><th>Qualitative</th><th>Discrete quantitative</th><th>Continuous quantitative</th></tr><tr><td>Favourite sport</td><td></td><td>✓</td><td></td><td></td></tr><tr><td>No. of people</td><td>✓</td><td></td><td></td><td></td></tr><tr><td>Distance travelled</td><td></td><td></td><td></td><td>✓</td></tr></table> ***B2*** *for 2 correct* ***B1*** *for 1 correct*	**3**	B3
1(b)	Lcb = 7 **and** Ucb = 9	**1**	B1
1(c)	Lcb = 6.5 **and** Ucb = 8.5	**1**	B1

Question	Answer	Marks	Partial Marks
2(a)	P($A \cap B$) = 0	**3**	B1
	Use of P($A \cup B$) = P(A) + P(B) = 0.35 + 0.2		M1
	P($A \cup B$) = 0.55		A1
2(b)(i)	Use of P($A \cap C$) = P(A) + P(C) – P($A \cup C$) = 0.35 + 0.4 – 0.52	**2**	M1
	0.23		A1
2(b)(ii)	**EITHER** [If independent P($A \cap C$) = P(A) × P(C)] = 0.35 × 0.4	**3**	M1*
	Comparing *their* 0.23 with 0.35 × 0.4		M1dep
	OR [If independent P($A \cup C$) = P(A) + P(C) – P(A) × P(C) =] 0.35 + 0.4 – 0.35 × 0.4		(M1*
	Comparing 0.52 with 0.35 + 0.4 – 0.35 × 0.4		M1dep)
	So A and C are not independent		A1

Question	Answer	Marks	Partial Marks
3(a)	[number of women =] 80 – 52 or 28 **or** [number of men =] 60 – 42 or 18 **or** [number under 50 =] 0.4(0.75×80 + 0.5×60) or 36 **or** [number 50 and over =] 0.2(0.25 × 80 + 0.5 × 60) or 10	**2**	M1
	28 + 18 = 46 **or** 36 + 10 = 46		A1
3(b)	25% [of women are under 30]	**2**	M1
	25% × 80 = 20		A1
3(c)	x% × 80 + y% × 60 **oe or** 25% × w + 50% × m	**3**	M1
	50% of '50'		M1
	25		A1

Question	Answer	Marks	Partial Marks
4(a)	**EITHER** 10 × 94 **or** 5 × 103 **or** 3 × 101 **or** 2 × x	**4**	M1*
	10 × 94 + 5 × 103 + 3 × 101 + 2 × x		M1dep*
	'(10 × 94 + 5 × 103 + 3 × 101 + 2 × x)'/(10 + 5 + 3 + 2) = 98.1		M1dep
	OR 98.1 × (10 + 5 + 3 + 2)		(M1*
	'1962' ─ (10 × 94 + 5 × 103 + 3 × 101)		M1dep*
	'204' ÷ 2		M1dep)
	102		A1
4(b)	• [Overall] prices/costs decreased • By 1.9% • Between 2022 and 2023 ***B1*** *for 2 correct from the 3 above*	**2**	B2
4(c)	A change/increase/decrease in **an amount** e.g. amount of equipment, distance travelled, amount of advertising	**1**	B1

Question	Answer	Marks	Partial Marks
5(a)	Prizes of 2, 3 and 4	**4**	B1
	At least one correct method from 5/6 × 3/4 5/6 × 1/4 + 1/6 × 3/4 1/6 × 1/4		M1
	At least two correct from P(2) = 5/8, P(3) = 1/3, P(4) = 1/24 **oe**		A1
	3 probabilities that sum to 1		B1
5(b)	**EITHER** Σ 'prize' × 'prob' 2 × 15/24 + 3 × 8/24 + 4 × 1/24	**4**	M1*
	15 × Σ 'prize' × 'prob' **or** ±(Σ 'prize' × 'prob' – 2.50)		M1dep*
	±(15 × Σ 'prize' × 'prob' – 15 × 2.50)		M1dep
	OR ±Σ ('prize' – 2.5) × 'prob' ±(– 0.5 × 15/24 + 0.5 × 8/24 + 1.5 × 1/24)		(M2*
	15 × ±Σ ('prize' – 2.5) × 'prob'		M1dep)
	($)1.25 loss		A1

Question	Answer	Marks	Partial Marks
6(a)	±(11.4 – 12.1)/0.35 **or** ±(23.3 – 24.3)/0.8 ***M1*** *for* ±(11.4 – 12.1) **or** ±(23.3 – 24.3)	**4**	M2
	(±) –2 **and** (±) –1.25		A1
	100 m **and** scaled scores of –2 and –1.25		A1
6(b)	52.6	**2**	B1
	1.6		B1

Question	Answer	Marks	Partial Marks
7(a)	<table><tr><th>*Frequency*</th></tr><tr><td></td></tr><tr><td></td></tr><tr><td>25</td></tr><tr><td>34</td></tr><tr><td></td></tr><tr><td></td></tr></table> <table><tr><th>*C frequency*</th></tr><tr><td></td></tr><tr><td></td></tr><tr><td></td></tr><tr><td></td></tr><tr><td>76</td></tr><tr><td>76</td></tr></table> ***B2*** *for 3 correct frequencies/cumulative frequencies* ***B1*** *for 1 or 2 correct frequencies/cumulative frequencies*	**3**	B3
7(b)	Correct labels: frequency/number [of sea cucumbers] **and** length, cm	**4**	B1
	At least 4 correct midpoints plotted from 4, 12, 20, 28, 36 and 44		M1
	At least 4 of their frequencies plotted from 0, 7, '25', '34', 10 and 0		M1
	6 correct plots joined with straight line segments **ft**		A1ft
7(c)	'76' ÷ 2 **or** 38th value	**4**	M1*
	EITHER 24 +		M1dep
	'6'/'34' × 8		M1dep
	OR 32 –		(M1dep
	'28'/'34' × 8		M1dep)
	OR (('38' – 32) × 32 + (66 – '38') × 24)/34		(M2dep)
	25.4		A1

Question	Answer	Marks	Partial Marks
7(d)	**ALT 1** Length less than 23 [cm]	**4**	B1
	EITHER 7/8 × '25'		M1
	7 + a fraction of '25'		M1
	OR 1/8 × '25'		(M1
	32 – a fraction of '25'		M1)
	OR (7 × 32 + 1 × 7)/8		(M2)
	ALT 2 Change to width intervals $0 \leqslant w < 4$, $4 \leqslant w < 8$, $8 \leqslant w < 12$, …		(B1
	3.5/4 × '25'		M1
	7 + a fraction of '25'		M1)
	29		A1

Question	Answer	Marks	Partial Marks
8(a)	A disadvantage of an open question, e.g. • May get too many different responses/difficult to analyse • Questions/responses can be misinterpreted	**1**	B1
8(b)	Census	**1**	B1
8(c)	An advantage of using a sample, e.g. • Quicker • Cheaper • Easier to collect data/easier to analyse data	**2**	B1
	A disadvantage of using a sample, e.g. • May not be representative • May be biased • Some sections of the population may be missed		B1
8(d)	A systematic sample (any interval size)	**3**	M1*
	With intervals of 11		M1dep
	2, 13, 24, 35, 46, 57		A1

Question	Answer	Marks	Partial Marks
8(e)	Sample has 4 FT and 2 PT	**3**	M1
	(24 + 12)/66 × 6 [=3.2… or 3] **or** (20 + 10)/66 × 6 [=2.7… or 3] **oe**		M1
	36/66 × 6 ≈ 3 **and** 30/66 × 6 ≈ 3 **and** so sample is not representative **ft**		A1ft
8(f)	24/66 × 11 = 4	**1**	B1
8(g)	11, 12 and 13 ***B2*** *for 1 extra* ***or*** *1 omission* ***B1*** *for 1 extra and 1 omission* ***or*** *2 extras* ***or*** *2 omissions*	**3**	B3

Question	Answer	Marks	Partial Marks
9(a)	It is necessary when an **even**-pointed moving average value is taken	**2**	B1
	so that the centred moving average values coincide with the readings		B1
9(b)	4.45, 4.3, 4.175, 4.1, 4 and 3.875 correctly positioned in the table ***B1*** *for 4 or more correct values seen anywhere*	**2**	B2
9(c)	±(3.6 – '4.45') **or** ±(3.4 – '4') **or** ±(360000 – '445000') **or** ±(340000 – '400000')	**3**	M1*
	±(Sum of two differences)/2 (–0.85 + –0.6)/2		M1dep
	–72500		A1
9(d)	[On average] they are [72500 dollars] below the trend line/moving average values/[annual] average **ft**	**1**	B1ft
9(e)	6 correct plots ***B1ft*** *for 5 correct plots or* ***B1*** *for 6 correct plots horizontally* ***or B1ft*** *for 6 correct plots vertically*	**3**	B2ft
	An appropriate trend line through their moving average values		B1
9(f)	Reading from their trend line at 2024 Q3 + their seasonal component	**3**	M1*
	Dealing correctly with 100000		M1dep
	222500 to 252500 *If first M1 only,* ***SC1*** *for answer in range 2.225 to 2.525*		A1

<table>
<tr><th>Question</th><th>Answer</th><th>Marks</th><th>Partial Marks</th></tr>
<tr><td>10(a)</td><td>1/6 oe</td><td>1</td><td>B1</td></tr>
<tr><td rowspan="2">10(b)(i)</td><td>0 × 4 + 1 × 5 + 2 × 8 + 3 × 7 or 0 + 5 + 16 + 21</td><td rowspan="2">2</td><td>M1</td></tr>
<tr><td>42</td><td>A1</td></tr>
<tr><td rowspan="2">10(b)(ii)</td><td>(2 × 8)/‘42’</td><td rowspan="2">2</td><td>M1</td></tr>
<tr><td>8/21 oe</td><td>A1</td></tr>
<tr><td>10(c)</td><td><table><tr><td>‘42’</td><td></td></tr><tr><td>24</td><td>20</td></tr><tr><td>‘1.75’</td><td></td></tr><tr><td>3</td><td></td></tr></table>
B3ft for 4 correct, B2ft for 3 correct, B1ft for 2 correct</td><td>4</td><td>B4ft</td></tr>
<tr><td rowspan="2">10(d)</td><td>A correct reference to the range, e.g.
• the range supports the claim
• the range [for the western section] is bigger
• the range is the only relevant measure available
• the range is not sufficient because…</td><td rowspan="2">2</td><td>B1ft</td></tr>
<tr><td>A disadvantage of the range, e.g.
• it is affected by extreme values
• it only takes the largest and smallest values into account
• the interquartile range/standard deviation would be better
• we do not know the interquartile range/standard deviation</td><td>B1</td></tr>
<tr><td rowspan="2">10(e)</td><td>(‘42’ + 43)/(‘24’ + ‘20’) or (‘1.75’ × ‘24’ + 2.15 × ‘20’)/(‘24’ + ‘20’)</td><td rowspan="2">2</td><td>M1</td></tr>
<tr><td>1.93</td><td>A1</td></tr>
<tr><td rowspan="3">10(f)</td><td>B</td><td rowspan="3">3</td><td>B1</td></tr>
<tr><td>A</td><td>B1</td></tr>
<tr><td>A</td><td>B1</td></tr>
</table>

