



Cambridge O Level

STATISTICS**4040/22**

Paper 2

October/November 2020

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 2020 series for most Cambridge IGCSE™, Cambridge International A and AS Level and Cambridge Pre-U 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 descriptors 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.

MARK SCHEME NOTES

The following notes are intended to aid interpretation of mark schemes in general, but individual mark schemes may include marks awarded for specific reasons outside the scope of these notes.

Types of mark

M Method marks, awarded for a valid method applied to the problem.

A Accuracy mark, awarded for a correct answer or intermediate step correctly obtained. For accuracy marks to be given, the associated Method mark must be earned or implied.

B Mark for a correct result or statement independent of Method marks.

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 M or B mark is dependent on an earlier, asterisked, mark in the scheme.

The symbol $\surd$ 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.

Abbreviations

AG	answer given on question paper
awrt	answer which rounds to
cao	correct answer only
dep	dependent
ft	follow through after error
oe	or equivalent
SC	special case
soi	seen or implied
www	without wrong working

Question	Answer	Marks	Partial Marks
1	A dual or multiple bar chart with a suitable linear scale	3	M1
	Correct dual bar chart with bars to correct heights: 70, 60; 64, 47; 122, 192.		A1
	Fully labelled bar chart: Number/frequency; Osprey, Eagle, Hawk; 2009, 2019.		B1

Question	Answer	Marks	Partial Marks
2(a)	Five correct words used: Gender – [discrete] qualitative Height – continuous quantitative Shoe size – discrete quantitative (B3 for 4 correct, B2 for 3 correct, B1 for 2 correct)	4	B4
2(b)	Scatter diagram/graph/plot	1	B1
2(c)	Mean	3	B1
	Median		B1
	One from, e.g.: - Median is not influenced by extreme values - Mean takes extreme values into account/uses all the values - Mean is useful for further calculations		B1

Question	Answer	Marks	Partial Marks
3	2, 4, 5, 8, 8 (total of 27, median of 5, mode of 8) (B2 for five numbers with 2 correct features, B1 for five numbers with 1 correct feature)	3	B3

Question	Answer	Marks	Partial Marks
4	Worst = Javelin, from correct reasoning: as this result below the mean [and others above] OR results scaled to a common mean and standard deviation with Javelin $\frac{1}{4}$ of common s.d. below the common mean	5	B1
	($\pm$) 13.31 – 9.63 or ($\pm$) 3.68 or ($\pm$) 32.40 – 30.54 or ($\pm$) 1.86 seen or equivalent reasoning in words		M1*
	($\pm$) (13.31 – 9.63)/1.84 or ($\pm$) (32.40 – 30.54)/3.72 seen or equivalent reasoning in words		M1dep*
	Calculation of scaled scores for Shot put and Discus with common mean and standard deviation or equivalent reasoning in words		M1dep
	Best = Shot put, from fully correct working/reasoning		A1

Question	Answer	Marks	Partial Marks
5(a)	One of: $P(A) + P(B) > 1$ $P(A) + P(B) \neq P(A \cup B)$ $P(A) + P(B) > P(A \cup B)$ $[P(A \cap B) =] P(A) + P(B) - P(A \cup B) \neq 0$	2	M1
	So A and B are not mutually exclusive		A1
5(b)	EITHER $[P(A \cap B) = P(A) + P(B) - P(A \cup B) =] 0.7 + 0.5 - 0.85$	4	(M1*
	[If independent $P(A \cap B) = P(A) \times P(B) =] 0.7 \times 0.5$		M1*
	Clear explanation of method		M1 dep
	Each evaluated to 0.35 and ' A and B are independent'		A1)
	OR $[If independent P(A \cap B) = P(A) \times P(B) =] 0.7 \times 0.5$		(M1*
	[so if independent $P(A \cup B) = P(A) + P(B) - P(A) \times P(B) =] 0.7 + 0.5 - '0.35'$		M1*
	Clear explanation of method		M1 dep
	Correctly evaluated to 0.85 and ' A and B are independent'		A1)

Question	Answer	Marks	Partial Marks
6(a)	2	1	B1
6(b)	010, 523, 421, 148, 260, 325 (B2 for 1 independent error/omission only, B1 for exactly 2 independent errors/omissions) SC B2 for stratified sample by age only: 010, 523, 496, 148, 260, 052 If B0 then SC B1 for calculation of correct strata sizes seen 2:2:1:1	3	B3
6(c)	Sample contains 3 lower school boys or 010, 082, 154 indicated	3	B1*
	$192/576 \times 8$ oe		M1
	$192/576 \times 8$ [= 2.6... or 2.7] compared with 3, so yes sample correctly represents lower school boys (or is as representative as it can be)		A1 dep

Question	Answer	Marks	Partial Marks
7(a)	$185.64/182 \times 100$ or $(185.64 - 182)/182 \times 100$ [+ 100] or $(185.64 - 182)/182 \times 100 + 100$	2	M1
	102		A1
7(b)	$9 \times '102' + 6 \times 107 + 2 \times 98 + 3 \times 104$	3	M1*
	$\div (9 + 6 + 2 + 3)$		M1dep
	103.4 or 103		A1
7(c)	All three of the following: - the cost/price has 'increased' - by '3.4%' - between 2016 and 2019 (B1 for two correct points above)	2	B2ft
7(d)	One from, e.g.: - The number of workers may have changed - The number of hours worked (on average) may have changed - The amount of materials may have changed - The amount of electricity/phone calls etc. may have changed	1	B1

Question	Answer	Marks	Partial Marks
8(a)(i)	37.8 and 32.4 seen or ['highest' – 'lowest'] = 37·... – 32·...	2	M1
	5.4		A1
8(a)(ii)	41.6 and 39.4 seen or ['UQ' – 'LQ'] = 41·... – 39·...	2	M1
	2.2		A1
8(a)(iii)	The head circumferences have increased oe	2	B1
	The variation in the head circumferences has not really changed/has decreased slightly oe		B1
8(b)	EITHER No because the baby at the LQ position at 3 months old might not have been at the LQ position at birth OR Yes because the distributions are a similar shape suggesting that the babies rank order of size may not have changed	1	B1
8(c)	Median = 43.8, LQ = 42.6, UQ = 45.0, IQR = 2.4 All 4 correct (B3ft for 3 correct, B2ft for 2 correct, B1ft for 1 correct (allow ft for iqr))	4	B4
8(d)	Box-and-whisker with their median, LQ and UQ correct	2	B1ft
	Whiskers to 41.0 and 46.2		B1
8(e)(i)	2/19	1	B1
8(e)(ii)	3/8	1	B1
8(f)	Median = 46.3, IQR = 2.4	1	B1ft

Question	Answer	Marks	Partial Marks
9(a)(i)	$P(0) = \frac{1}{2}$	3	B1
	$\frac{1}{2} \times \frac{1}{6}$		M1
	$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/12$		A1
9(a)(ii)	$\Sigma (\text{prize} \times \text{'probability'})$	2	M1
	[\$]1.75		A1
9(b)(i)	Use of $\Sigma \text{ prob} = 1$	2	M1
	$\frac{1}{16}$ or 0.0625		A1
9(b)(ii)	$\frac{1}{4} \times \frac{1}{8} (\times 2)$	3	M1
	$\frac{1}{2} \times \frac{1}{2}$		M1
	$\frac{5}{16}$ or 0.3125 oe If 0 marks SC B1 for evidence of the 3 ways to win a total of \$2: [\$]1, [\$]1 and [\$]0, [\$]2 and [\$]2, [\$]0		A1
9(b)(iii)	$20(1 - (0 \times 1/4 + 1 \times 1/2 + 2 \times 1/8 + 3 \times '1/16' + 4 \times '1/16'))$ $\Sigma (\text{prize} \times \text{'prob'})$	4	M1*
	$1 - \Sigma (\text{prize} \times \text{'prob'})$ or $20 \times \Sigma (\text{prize} \times \text{'prob'})$		M1dep*
	$20 \times (1 - \Sigma (\text{prize} \times \text{'prob'}))$ oe		M1dep
	$20 \times (1 - \Sigma (\text{prize} \times \text{'prob'}))$ oe Loss of \$3.75		A1

Question	Answer	Marks	Partial Marks
10(a)	7, 20, 35, 65, 92, 112, 120	1	B1
10(b)	EITHER $0.1 \times 120 [= 12]$	4	(M1
	190 +		M1
	$(12 - 7)/13 \times (225 - 190)$		M1
	203 cao SC <i>If the smallest mass of the largest 10% found, SC M1 for $0.9 \times 120 [=108]$ and SC A1 for 324 cao</i>		A1)
	OR $0.1 \times 120 [= 12]$		(M1
	$(5 \times 225 + 8 \times 190)/13$ (M1 for $(a \times 225 + b \times 190)/13$ or $(5 \times c + 8 \times d)/13$, $c - d = 35$)		M2
	203 cao		A1)
10(c)	$(240 - 225)/(250 - 225) \times 15 [= 9]$ oe OR $(250 - 240)/(250 - 225) \times 15 [= 6]$ oe	5	M1
	Small = $9 + 13 + 7 - 12 = 17$		A1
	$(290 - 270)/(300 - 270) \times 27 [= 18]$ oe OR $(300 - 290)/(300 - 270) \times 27 [= 9]$ oe		M1
	Medium = $15 - 9 + 30 + 18 = 54$ Large = $8 + 20 + (27 - 18) = 37$		A1
	OR $108 - 54 - 17 = 37$		A1ft
10(d)	$37/108 \times 36/107 \times 35/106 \times 17/105 \times 4$ Product of 4 probabilities $\times 4$	4	M1
	'37' and '17' seen in numerator of product of 4 probs		M1
	Without replacement understood, $n(n-1)(n-2)(n-3)$ seen in denominator or '37' ('37' - 1)('37' - 2) seen in numerator		M1
	$1258/51039$ or $0.0246(478\dots)$ oe		A1

Question	Answer	Marks	Partial Marks
11(a)	[Rainfall] highest in Q1 [and Q4] and lowest in Q3 [and Q2]	1	B1
11(b)	Any two from - To eliminate/reduce the variation - To find the trend - To find seasonal components - To make predictions (B1 for one correct purpose)	2	B2
11(c)	Any acceptable explanation, e.g.: - Because 4 is even - It is an even-pointed moving average - So that the moving average values will coincide with the original data/times	1	B1
11(d)	EITHER	5	(M1
	Sum of 4 consecutive values soi 498, 506, 518, 530, 550		
	Mean of 4 consecutive values soi 124.5, 126.5, 129.5, 132.5, 137.5		M1
	Mean of 2 consecutive moving average values		M1
	125.5, 128, 131, 135		A1
	Their 4 values in table lined up with 2017 Q3 to 2018 Q2		B1)
	OR		(M1
	Sum of 4 consecutive values soi 498, 506, 518, 530, 550		
	Sum of 2 consecutive sums above soi 1004, 1024, 1048, 1080		M1
	Sum of 8 values above $\div 8$		M1
	125.5, 128, 131, 135		A1
	Their 4 values in table lined up with 2017 Q3 to 2018 Q2		B1)
11(e)	Plots correct horizontally	3	B1
	Their 4 (or 5) plots correct vertically		B1ft
	Appropriate trend line		B1ft
11(f)	On average rainfall is increasing	1	B1ft
11(g)	60	1	B1

Question	Answer	Marks	Partial Marks
11(h)	'155' + '60'	2	M1
	215		A1ft