STATISTICS **4040/23**

Paper 2 **October/November 2017**

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 2017 series for most Cambridge IGCSE[®], Cambridge International A and AS Level components and some Cambridge O Level components.

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 ✓ 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(i)	39 and 34	**1**	B1
1(ii)	Key/labelling on sectional bars (pass, merit, distinction) and labelling on horizontal axis (male, female)	**4**	B1
	12/'39'×100, 16/'39'×100, 11/'39'×100; 19/'34'×100, 4/'34'×100, 11/'34'×100 At least one correct percentage calculation		M1
	31, 41, 28; 56, 12, 32 (**awrt**) At least two correct percentages seen, **ft** their totals from **(i)**		A1✓
	Fully correct bar heights		A1

Question	Answer	Marks	Partial marks
2(i)	(table below)	**4**	B4 for all 5 correct

Qualitative	Discrete quantitative	Continuous quantitative	Not a variable
		✓	
✓			
	✓		
✓			
			✓

(*B3 for 4 correct, B2 for 3 correct, B1 for 2 correct*)

Question	Answer	Marks	Partial marks
2(ii)	19 and 22	**1**	B1
2(iii)	49.5 and 54.5	**1**	B1
3(i)	A pair of frequency polygons drawn for comparison	**4**	B1
	Key/polygons labelled (male, female), vertical axis labelled (number/frequency) and horizontal axis labelled (height (cm))		B1
	Suitable linear scales		B1*
	Correct plots horizontally and vertically		B1dep
3(ii)	Male elephants have a greater shoulder height **oe**	**1**	B1

Question	Answer	Marks	Partial marks
4(a)(i)	Use of P(A∩B) = P(A) × P(B)	**2**	M1
	P(B) = 0.25/0.5 = 0.5 **www**		A1
4(a)(ii)	Obtaining a head/tail when **another** coin is thrown Or obtaining a head/tail when the coin is thrown **again** Or some other independent event with probability of 0.5 e.g. obtaining an even number when a [fair] die is thrown	**1**	B1

Question	Answer	Marks	Partial marks
4(b)	Use of P($C \cup D$) = P(C) + P(D)	**3**	M1
	P($C \cup D$) = 0.62 + 0.21 = 0.83		A1
	P($C \cap D$) = 0		B1

Question	Answer	Marks	Partial marks
5(i)	Houses at equal intervals	**3**	M1
	40/5 [=8] or intervals of 8 seen		M1
	02 10 18 26 34		A1
5(ii)(a)	All even numbered houses/all from same side of road **ft**	**2**	B1✔
	People from just 5 households/people from same household may hold similar opinions		B1
5(ii)(b)	A named sampling method aiming for representation from each side of the road e.g. a sample stratified by side of road, quota – some from each side of road, systematic – odd interval, random	**2**	B1
	of the people [rather than the houses]		B1

Question	Answer	Marks	Partial marks
6(i)(a)	4/25 or 0.16	**1**	B1
6(i)(b)	19/25 or 0.76	**1**	B1
6(i)(c)	2/15 or 0.13[3]	**1**	B1
6(i)(d)	17/25 or 0.68	**1**	B1
6(ii)	10/25 × 9/24 + 6/25 × 5/24 + 9/25 × 8/24 Sum of 3 products of 2 probabilities	**3**	M1
	$m/n \times (m – 1)/(n – 1)$ seen		M1
	192/600 or 8/25 or 0.32 **oe**		A1

Question	Answer	Marks	Partial marks
7(i)	70 000	**1**	B1
7(ii)(a)	Median	**1**	B1
7(ii)(b)	Any values between 60 000 and 80 000 (but not including 80 000)	**2**	B1 B1

Question	Answer	Marks	Partial marks
7(iii)	30th value **and** 90th value (allow 30.25th and 90.75th)	**7**	B1
	Either: lower quartile 20 000 +		M1
	………... ('30' – 25)/33 × 5000 [= 757.5757…]		M1
	Or: upper quartile 30 000 +		
	………… ('90' – 87)/24 × 10 000 [=1250]		
	Lower quartile = 20 760 **awrt**		A1
	Upper quartile = 31 250		A1
	Upper quartile – lower quartile		M1
	10 500 **awrt**		A1
7(iv)	**Either:** 2000/5000 × 33 [= 13.2]	**5**	M1
	'13.2' + 20 + 5		M1
	38		A1
	'38' × $36 + (120 – '38') × $45 [= 5058]		M1
	Or: 3000/5000 × 33 [= 19.8]		
	'19.8' + 29 + 24 + 6 + 3		
	82		
	(120 – '82') × $36 + '82' × $45 [=5058]		
	[$] 5060 **awrt**		A1

Question	Answer	Marks	Partial marks
8(i)	1200 × 0.12, 600 × 0.4, 20 × 1.2 one correct product	**3**	M1
	144 : 240 : 24 **oe**		A1
	6 : 10 : 1		A1
8(ii)	Leaflets: 103	**5**	B1
	Phone calls: 0.38/0.4 [×100] or 0.02/0.4 [×100] **oe**		M1
	95		A1
	Petrol: 1.26/1.2 [×100] or 0.06/1.2 [×100] **oe**		M1
	105		A1
8(iii)(a)	'6' × '103' + '10' × '95' + '1' × '105'	**3**	M1
	÷ ('6' + '10' + '1')		M1
	98.4 **cao** (must be to 1 dp)		A1

Question	Answer	Marks	Partial marks
8(iii)(b)	[Costs/prices] reduced	**3**	B1✓
	by 1.6% **awrt**		B1✓
	between this year and last year/since last year/over the year		B1
8(iv)	2 **in context** reasons e.g.: • Number of leaflets may have changed/increased/decreased • Number (of minutes) of phone calls may have changed/increased/decreased • Number of litres of petrol may have changed/increased/decreased/she may travel more/less/change her car [affecting petrol consumption] • Another category, such as e.g. 'online', may be introduced	**2**	B1 B1

Question	Answer	Marks	Partial marks
9(i)	1 – 0.8 [= 0.2]	**4**	M1
	0.8 × 0.1		M1
	0.2 × 0.7		M1
	0.8 × 0.1 + 0.2 × 0.7 = 0.22 **AG**		A1
9(ii)	**Either:** 1 – 0.22 [= 0.78]	**3**	M1*
	0.22 × 14.50 + '0.78' × 16.50		M1dep
	Or: 0.22 × (–)2 [= (–)0.44]		
	16.50 – '0.44'		
	[$]16.06		A1
9(iii)	Number of days late = 11	**2**	B1
	Expected earnings = 11 × 14.50 + 39 × 16.50 = [$]803 **or** 50 × '16.06' = [$]803 **ft**		B1✓
9(iv)	y × (1 – 0.22) + (y – 3) × 0.22 = '16.06' **or** y – 0.22 × 3 = '16.06' **oe** *(Attempt at expected earnings (involving an unknown) = '16.06')*	**4**	M1
	A correct LHS above		M1
	Fully correct equation above		A1
	[$]16.72		A1

Question	Answer	Marks	Partial marks
9(v)	**Either:** $0.16 \times 0.78 + 0.22 \times 0.84 + 0.16 \times 0.22$ **Or:** $1 – 0.78 \times 0.84$ *(At least one correct product seen ($\pm$))*	**3**	M1
	Fully correct expression		M1
	0.3448 or 0.345 or 431/1250 **oe**		A1

Question	Answer	Marks	Partial marks
10(i)	[Generally] quicker on first circuit **oe**	**2**	B1
	Less varied on first circuit **oe**		B1
10(ii)	$(57.1 – 52.3)/3.2$ [$= (Z_1 – 0)/1$] **or** $(63.6 – 57.6)/4.8$ [$= (Z_2 – 0)/1$]	**3**	M1
	1.5 **and** 1.25		A1
	[Zara performed better] in the second circuit as her scaled time is lower **oe ft**		B1
10(iii)(a)	Attempt at mid-points 220, 260, 300 (at least one correct, allow +/- 0.5)	**9**	M1
	Subtraction of assumed mean from their mid-points [-40, 0, 40]		M1
	$\Sigma f'x'$ [= 280]		M1*
	$\Sigma fx/50$		M1dep
	5.6 or 280/50 **oe**		A1
	265.6 or 266		A1
	$\Sigma f'x'^2$ [= 27200]		M1*
	Use of correct formula for variance or standard deviation		M1dep
	22.6 **awrt** (from correct use of assumed mean)		A1
10(iii)(b)	Data is grouped/mid-points used/we do not know the distribution within classes /large classes/the actual values are not known	**1**	B1
10(iii)(c)	More classes/smaller class widths	**1**	B1

Question	Answer	Marks	Partial marks
11(i)(a)	Pattern is likely to repeat every 5 days/5 days is one complete cycle/5 days in this school week	**1**	B1
11(i)(b)	Moving average values will coincide with original data/original time (or B1 for *n* is odd/values are already centred)	**2**	B2
11(i)(c)	Totals: 1269, 1275, 1280, 1288, 1297, 1305 1 correct total (may be implied)	**3**	M1
	253.8, 255, 256, 257.6, 259.4, 261 2 correct moving averages seen		A1
	All correct and in correct positions in table		A1

Question	Answer	Marks	Partial marks
11(ii)	281 – '253.8' [=27.2] , 289 – '261' [=28] one correct difference (allow ±)	**3**	M1
	Sum of 2 differences ÷ 2		M1
	27.6		A1
11(iii)	6 correct plots vertically **ft**	**3**	B1✓
	6 correct plots horizontally		B1
	Suitable trend line		B1✓
11(iv)	A reading from the trend line + '27.6'	**2**	M1
	296 (whole number) **ft** their 27.6 and accurate reading from their trend line		A1✓
11(v)(a)	Increasing **oe**	**1**	B1
11(v)(b)	Any plausible explanation with correct associated judgement e.g. No as there will be an upper limit (number of pupils in the school)	**1**	B1