

MARK SCHEME for the October/November 2014 series

4040 STATISTICS

4040/13

Paper 1, maximum raw mark 100

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 will not enter into discussions about these mark schemes.

Cambridge is publishing the mark schemes for the October/November 2014 series for most Cambridge IGCSE[®], Cambridge International A and AS Level components and some Cambridge O Level components.

Page 2	Mark Scheme	Syllabus	Paper
	Cambridge O Level – October/November 2014	4040	13

- 1 8 is the mode
The value which occurs most frequently. M1
A1
- 9 is the median
Obtained by arranging the values in ascending or descending order and selecting the 'middle' one. M1
A1
- 11 is the (Arithmetic) mean
Obtained by summing the numbers and then dividing by 13. M1
A1
- 2 (i) X is discrete
Because it only takes integer values (or equivalent comment) B1*
B1dep
- (ii) 0 and 4 (B1 for each) B2
- (iii)
- | | | | | | | | | |
|-----------|---|---|----|----|---|---|---|---|
| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Frequency | 0 | 5 | 15 | 10 | 0 | 7 | 6 | 7 |
- (–1 each independent error) B2
- 3 (a) Similar in that both would sample proportionately from the different age groups. B1
In stratified sampling interviewers would be given a list of specific people to interview, in quota sampling the interviewer selects the individuals. B1
- (b) (i) Because the last page of a chapter is less likely than all other others to be filled with words, B1
the sample is likely to be biased. B1
- (ii) A systematic sample is a form of random sampling B1
and so unless there is some pattern in the pages which matches the sampling interval the sample will be unbiased. B1
- 4 (i) 0 8 18 35 46 50 (all correct) B1
- (ii) All points plotted correctly both horizontally and vertically B1✓
Plotted points connected by a suitable smooth curve B1
- (iii) (a) Correct reading from graph of a point between cum. freqs. 12 and 13 B1✓
- (b) Clear attempt to use appropriate point on the graph and any valid method to find the required percentage. M1
14%–16% A1

Page 3	Mark Scheme	Syllabus	Paper
	Cambridge O Level – October/November 2014	4040	13

- 5 (i) Advantage: it shows actual amounts of wood. B1
Disadvantage: it only shows information about individual sizes. B1
- (ii) The total amount of wood of all sizes produced. B1
- (iii) Pie chart B1
Sectional (component) bar chart B1
- (iv) Change chart B1
- 6 (i) Attempt to sum the values in the diagram and subtract the total from 70. M1
5 A1
- (ii) None of the people in the sample speak all three languages. B1
- (iii) (a) No, because this person will still only speak two languages. B1
(b) Yes, because the person now speaks all three languages. B1
(c) No, as this person only speaks one of the three languages. B1

Page 4	Mark Scheme	Syllabus	Paper
	Cambridge O Level – October/November 2014	4040	13

- 7 (a) Sight of $3/7$ used B1
 EITHER 1 – sum of two two-factor products M1
 $1 - [(4/7 \times 1/5) + (3/7 \times 1/9)]$ A1
 $88/105$ A1
- OR Sight of $4/5$ and $8/9$ used M1
 $(4/7 \times 4/5) + (3/7 \times 8/9)$ A1
 $88/105$ A1
- (b) (i) EITHER $3/7 \times 2/6 \times 1/5$ OR $1/7 \times 1/6 \times 1/5 \times 3!$ M1
 $1/35$ A1
- (ii) Any appreciation of the fact that it is irrelevant which two are the brother and sister. B1
- EITHER $1/7 \times 1/6 (\times 1) \times 3!$ OR $5/7 \times 1/6 \times 1/5 \times 3!$ M1
 $1/7$ A1
- (c) (i) Clear attempt at both two blue and two white M1
 $(2/8 \times 3/8) + (6/8 \times 5/8)$ A1
 $9/16$ A1
- (ii) Given first balls were the same colour, $P(\text{both were blue}) = 1/6$,
 $P(\text{both were white}) = 5/6$ B1
- Attempt to add probabilities relating to whether first balls were blue or white M1
- $(1/6)[(3/9 \times 5/7) + (6/9 \times 2/7)] + (5/6)[(2/9 \times 4/7) + (7/9 \times 3/7)]$ A1
 $86/189 = 0.455$ A1

8

(1)	(2)	(3)	(4)	(5)	(6)
<i>Time (x)</i> (minutes)	<i>Frequency (f)</i>	<i>Mid-pts (m)</i>	<i>y</i>	<i>fy</i>	<i>fy²</i>
0 – under 30	6	15	–12	–72	864
30 – under 35	11	32.5	–5	–55	275
35 – under 40	4	37.5	–3	–12	36
40 – under 50	40	45	0	0	0
50 – under 60	26	55	4	104	416
60 – under 70	14	65	8	112	896
70 – under 100	4	85	16	64	1024
TOTAL	105			141	3511

- (i) Mid-points correct B1
- (ii) Values of *y* found correctly M1
y values correct A1
- (iii) *fy* values found correctly M1
- (iv) *fy²* values found correctly M1
- (v) Summations correct A1
- (vi) Use of their values in a correct method for mean of *y* M1
Mean of *y* = 1.34 A1
- (vii) Use of their values in a correct formula for variance or s.d. of *y* M1
s.d. of *y* = 5.62 A1
- (viii) (a) (Their *y* mean \times 2.5) + 45 M1
48.4 A1✓
- (b) (Their *y* s.d. \times 2.5) only M1
14.1 A1✓
- (ix) The distribution is reasonably symmetrical with relatively M1
few extreme values, (or similar comment), A1
and so the s.d. is preferable to the IQR.

Page 6	Mark Scheme	Syllabus	Paper
	Cambridge O Level – October/November 2014	4040	13

- 9 (i) 36 32 in second and third cells B1
Any appreciation of area being proportional to frequency M1
24 28 in first and last cells A1
21 18 22 19 in remaining cells A1
- (ii) Correct classes, 15–17, 17–19 etc. M1
Correct frequencies 24 68 80 28 A1✓
Their results presented in a suitable table B1
- (iii) Four rectangles of equal width M1
Vertical axis correctly annotated M1
Rectangles of correct heights A1✓
- (iv) Use of 'diagonal line' on histogram or equivalent numerical method seen M1
19.35 cm A1✓
- (v) Proportions of first and last classes found correctly M1
Total cakes which can be sold found correctly M1
Percentage expressed correctly M1
84% A1
- 10 (i) (3×7) or $(3 \times 7000)/1000$ or equivalent seen **AG** B1
- (ii) Total deaths $25 + 21 + 47 + 83 (= 176)$ M1
Total population $4500 + 7000 + 6000 + 7000 (= 24500)$ M1
 $CDR = (\text{Total deaths} / \text{Total population}) \times 1000$ M1
 $= 7.18$ A1
- (iii) $(\text{Deaths/Population}) \times 1000$ seen for any age group (or can be implied by one correct result) M1
5.56 7.83 11.86 all correct A1
- (iv) Rate \times SP% seen for any age group (or can be implied by one correct result) M1
Attempt to sum results for all age groups M1
 $5.56 \times 0.2 + 3 \times 0.35 + 7.83 \times 0.25 + 11.86 \times 0.2$ A1✓
6.49 A1
- (v) Rate \times SP% added for four groups M1
7.90 A1
- (vi) Any valid comment relating to the towns having different age structures B1
- (vii) Because the SDR is lower M1
Eastbury has the healthier environment. A1✓

Page 7	Mark Scheme	Syllabus	Paper
	Cambridge O Level – October/November 2014	4040	13

11 (i)	Correct plots (–1 each error) Correct labels	B2 B1
(ii)	(37.5,104.5) (B1 each coordinate) Correct plot	B2 B1✓
(iii)	Correct SA plots (B1 for each) Line of best fit through at least two averages	B2 B1
(iv)	A and B results are both approximately linear. C results are completely inconsistent.	B1 B1
(v)	Correct plot	B1
(vi)	Experienced technician’s result totally consistent with those of B, suggesting that B’s observations are accurate.	B1 B1
(vii)	Line drawn through results of B and the experienced technician	B1
(viii)	135 kg, with clear indication value found from use of the revised line	B1✓