Cambridge International Examinations
General Certificate of Education Ordinary Level

STATISTICS	**4040/02**
Paper 2	**For Examination from 2018**
SPECIMEN MARK SCHEME	
	2 hours 15 minutes

MAXIMUM MARK: 100

This document consists of **7** printed pages and **1** blank page.

CAMBRIDGE
International Examinations

[Turn over

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 notation '**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.

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
1(a)	sectional/component/composite bar chart	**B1**
1(b)	discrete quantitative	**B1**
1(c)	3 as numerator or 11 as denominator of probability	**M1**
	3/11 or **awrt** 27% *(accept '3 out of 11' or '3 in 11')* *(S.C. B1 for 3:8 if ratio notation used (do not accept any other ratio))*	**A1**

Question	Answer	Marks
2(a)(i)	use of $P(A) \times P(B)$	**M1**
	0.03	**A1**
2(a)(ii)	use of $P(A) + P(B) - P(A \cap B)$	**M1**
	0.32	**A1**
2(b)(i)	0	**B1**
2(b)(ii)	0.64	**B1**

Question	Answer	Marks
3(a)	stay the same	**B1**
	decrease	**B1**
3(b)	105% of 12000 + 200	**M1**
	12800	**A1**
	105% of 1000 only	**M1**
	1050	**A1**

Question	Answer	Marks
4(a)	P(cards same/Ariana wins) = $\frac{1}{4}$ **oe**	**B1**
	P(cards different/Bella wins) = $\frac{3}{4}$ **oe ft** *(1 – their P(cards same))*	**B1**
	$\frac{1}{4} \times 3 = \frac{3}{4} \times 1$, show expected winnings equal so it is a fair game *(allow follow through of their probabilities (with conclusion of 'fair' or 'unfair' as appropriate))*	**B1**
4(b)	P(4 or less/Ariana wins) = 6/16 = 3/8 or P(5 or more/Bella wins) = 10/16 = 5/8	**M1**
	5 × *their* 3/8 = *x* × *their* 5/8	**M1**
	3	**A1**

Question	Answer	Marks
5(a)	biased/not representative or too small	**B1**
	more detail about why biased such as will contain people who want/catch a bus at 7 am *(e.g. workers, school children...)*	**B1**
5(b)	15, 08, 00, 31, 52, 47 *(–1 each independent error/omission)*	**B2**
5(c)(i)	2, 3, 1	**B1**
5(c)(ii)	the different age groups are likely to want buses at different times *(some mention of the potential connection between people's ages and when they might want a bus)*	**B1**

Question	Answer	Marks
6(a)	$(+/-)\ 25 \times -9 + 9 \times 21 + (0 \times 2)$	**M1**
	–36	**A1**
	–36/36 leading to 8.59 or 1 min early *(do not accept –1)*	**A1**
6(b)	correct formula for s.d. or var * *(even if e.g. 8.59 used for mean, may just be the formula)*	**M1**
	their number of mins late used in the formula **dep** *(this must not be a time and all other values must be correct)*	**M1**
	11.9 (from a correct answer to **(a)**)	**A1**

Question	Answer	Marks
7(a)	15 correct values *(may be in wrong order)*	**M1**
	attempt to order *(allow one error/omission)*	**M1**

	15 correctly ordered values, equally spaced
	
	<pre> 9 8 5 │ 14
	     8 7 5 4 2 2 0 │ 15
	              9 4 3 1 │ 16
	                       1 │ 17
	                         │ 18</pre>
		A1

7(b)	box-and-whisker diagram with: whiskers extending to 158 and 184	**B1**
	median at 170 marked correctly	**B1**
	lower quartile at 164, and upper quartile at 177, marked correctly *(if 0 scored, 1 mark for correct values of LQ, median and UQ seen)*	**B1**
7(c)	boys taller/boys have bigger median **oe**	**B1**
	spread/range/IQR similar	**B1**

Question	Answer	Marks
8(a)	there are extreme values *(if these are specified they must be correctly referring to the large masses)*/positive skew	**M1**
	so the median would be better	**A1**
8(b)	12, 43, 72, 86, 94, 98, 100	**B1**
8(c)	250 and 450	**B1**
8(d)	**Solution 1** 25th or 25.25^{th} value	M1
	sight of 25 – 12 (= 13) or 25.25 – 12 (= 13.25)	M1
	250 +	M1
	... 13/31 × 200 or 13.25/31 × 200 (333.87 or 335.48) *(dependent on an attempt at both LQ and UQ)*	M1
	UQ – LQ *(but not 75 – 25)*	M1
	480 or 495 **awrt**	A1
	OR **Solution 2** 75th or 75.75^{th} value	M1
	sight of 75 – 72 (= 3) or 75.75 – 72 (= 3.75)	M1
	750 +	M1
	... 3/14 × 300 or 3.75/14 × 300 (814.29 or 830.36) *(dependent on an attempt at both LQ and UQ)*	M1
	UQ – LQ *(but not 75 – 25)*	M1
	480 or 495 **awrt**	A1
	Available marks for **Solution 1 OR Solution 2**	**6**
8(e)	they are less varied **oe**	**B1**
8(f)	200/300 × 29 (= $19.\dot{3}$)	**M1**
	(some fraction of 29) + 12 + 31 *(some working must be seen for the fraction of 29 if the value is anything other than 19(.3...))*	**M1**
	62($.\dot{3}$)	**A1**

Question	Answer	Marks
9(a)	x = 168	**B1**
	y = 1308	**B1**
	z = 218 **ft** *(from their y)*	**B1**
9(b)	because 6 is even	**B1**
	centring is necessary so that the centred values coincide with original data plots	**B1**
9(c)	213.5, 215, 215.5, 217, 219.25 (or 219.3), 221.25 (or 221.3) *(B1 for 4 correct)*	**B2**
	6 values in the correct places in the table	**B1**
9(d)	plots at correct points vertically **ft**	**B1**
	plots at correct points horizontally	**B1**
	suitable trend line	**B1**
9(e)	number of patients increasing	**B1**
9(f)	attempt at reading from graph (even at wrong place) – 11.25	**M1**
	correct reading from *their* graph – 11.25, correctly evaluated and rounded to nearest whole number **ft**	**A1**

Question	Answer	Marks
10(a)(i)	256/240 × 100	**M1**
	107	**A1**
10(a)(ii)	the cost/price has increased	**B1**
	by 10%	**B1**
	between 2012 and 2014/over the two years	**B1**
10(a)(iii)	98	**B1**
10(b)(i)	by using the **expenditure**/total money spent on each category **oe**	**B1**
10(b)(ii)	(7 × *their* 107 + 2 × 110 + 5 × *their* 98) *****	**M1**
	... ÷ 14 **dep**	**M1**
	104	**A1**
10(b)(iii)	5760 × *their* **(b)(ii)**/100 *(their **(b)(ii)** may be an earlier unrounded version)*	**M1**
	5990 **ft** *(ft answer must be to nearest dollar)*	**A1**
10(b)(iv)	two correct reasons e.g.: the quantity of equipment used may have changed the number of employees may have changed the number of hours worked by employees may have changed *(award one mark for each correct answer up to a maximum of two marks)*	**B2**

Question	Answer	Marks
11(a)(i)	60/300 = 1/5 **oe**	**B1**
11(a)(ii)	160/300 = 8/15 **oe**	**B1**
11(a)(iii)	230/300 = 23/30 **oe**	**B1**
11(a)(iv)	40/140 = 2/7 **oe**	**B1**
11(b)	$100/300 \times 99/299 + 130/300 \times 129/299 + 70/300 \times 69/299$ $n/m \times (n - 1)/(m - 1)$ seen somewhere	**M1**
	the sum of three pairs of products of probabilities *(if they consider males and females separately make sure all cases have been considered)*	**M1**
	315/897 or 105/299 **oe** or 0.35 or better	**A1**
11(c)	$(100/300 \times 130/299 \times 70/298) \times 6$ product of three probabilities $\times 6$	**M1**
	denominator of $300 \times 299 \times 298$	**M1**
	54600/267306 or 700/3427 **oe** or 0.20 or better	**A1**
11(d)	4, 3, 6	**B1**
11(e)(i)	$4 \times 100/(4 \times 100 + 3 \times 130 + 6 \times 70)$ numerator of *their* 4×100 *(must be a probability)*	**M1**
	denominator of *their* 4×100 + *their* 3×130 + *their* 6×70 *(must be a probability)*	**M1**
	40/121 **oe** or 0.33 better	**A1**
11(e)(ii)	$2 \times (100 + 70)$	**M1**
	34/121 **oe** or 0.28 or better	**A1**

BLANK PAGE