CAMBRIDGE INTERNATIONAL EXAMINATIONS

Pre-U Certificate

MARK SCHEME for the May/June 2013 series

1347 MATHEMATICS (STATISTICS WITH PURE MATHEMATICS)

1347/02 Paper 2 (Statistics), maximum raw mark 80

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 Examinations

1 (i)	Mean = 2.068 million Median = 2.035 million Values not very different, so no evidence of skew, fairly symmetric	B1 B1 B1 **[3]**	2.068 or 2068000, accept 2.07 or 2070000 2.035 or 2035000, accept 2.04 or 2040000 Symmetric or equivalent conclusion about the distribution of increases
(ii)	Interquartile range = 2.52 – 1.35 = 1.17 Either: 2.52 + 1.5 × 1.17 = 4.275 Or: (3.81 – 2.52) ÷ 1.17 = 1.103 3.81 is less than 1.5 × IQR above the upper quartile, so it is not an outlier	M1 M1 A1 B1 **[4]**	IQR calculated correctly, or implied from subsequent working 1.5 × their IQR or 1.755 or 1.29 ÷ their IQR A valid and correct calculation Explanation of how this shows that 3.81 is not an outlier
(iii)	The recent increases are all quite small compared with the earlier ones.	B1 **[1]**	Early values larger than later ones
2 (i)	1,2 1,3 1,4 1,5 2,3 2,4 2,5 3,4 3,5 4,5 Total: 3, 4, 5, 6, 7, 8, 9 Prob: 0.1, 0.1, 0.2, 0.2, 0.2, 0.1, 0.1	B1 B1 M1 M1 A1 **[5]**	Listing the ten pairs (allow twenty if given in both orders or as a two-way table), must be seen not implied Total can take values 3,4,5,6,7,8,9 (only) Correct prob (or ft *their* list) for 3 or 9 Correct prob (or ft *their* list) for 5, 6 or 7 All probabilities correct
(ii)	Expectation = 3 × 0.1 + 4 × 0.1 + 5 × 0.2 + ... = 6 $3^2 \times 0.1 + 4^2 \times 0.1 + 5^2 \times 0.2 + ... = 39$ Variance = 39 – 36 = 3 Standard deviation = √3 = 1.732	M1 A1 M1 A1 B1 **[5]**	Or by symmetry, method may be implied 6.0 or 6 39 seen or valid method for variance 3, from valid method seen √3 or 1.73 or better (ft from variance)
3 (i)	P(robin) = 0.87 X = number of children who saw a robin $X \sim B(4, 0.87)$ $P(X < 2) = (0.13)^4 + 4(0.87)(0.13)^3$ = 0.00793117	 B1 M1 A1 **[3]**	 Parameters 4 and 0.87 seen or implied A valid calculation of $P(X < 2)$ or $P(X \leqslant 2)$ 0.0079 or greater accuracy
(ii)	The probability of a robin is 0.87 in the places where these four children live The four gardens are independent 'The gardens are a random sample from the population' ⇒ SR2	B1 B1 **[2]**	p = 0.87 or constant probability Independence / random / not all from the same area

(iii)	H_0: P(see house sparrows) = 0.65 H_1: P(see house sparrows) < 0.65 X = no of families who see house sparrows Assuming H_0: $X \sim B(20, 0.65)$ Using tables, either: $P(X \leqslant 5) = 0.0003$ or cv = 7 0.0003 < 0.01 or 0.0003 < 1% or 7 > 5 Reject H_0 (accept H_1) Evidence supports the claim that the probability of seeing house sparrows in city areas is less than 65%	B1 B1 M1 A1 B1 M1 A1 **[7]**	'Probability of seeing house sparrows' < 0.65 or < 65% Using B(20, 0.65), or implied from subsequent working 0.0003 or 7 from correct method Correct comparison seen Reject H_0 or equivalent, correct for their tail probability or cv Correct conclusion in words
4 (i)	$S_{xx} = 4171 - \frac{(143)^2}{5} = 81.2$ $S_{yy} = 432 - \frac{(46)^2}{5} = 8.8$ $S_{xy} = 1342 - \frac{143 \times 46}{5} = 26.4$ $r = \frac{26.4}{\sqrt{81.2 \times 8.8}} = 0.9876$ (0.988) r is very near 1, so a good fit to (an upward sloping) line	B1 B1 B1 M1 A1 **[5]**	81.2 8.8 26.4 Calculating r from their S_{xx}, S_{yy} and S_{xy} (numerical working or their r value correct to 3 sf or better) Drawing a valid conclusion (confirming that a linear fit is appropriate, as stated in question)
(ii)	$b = \frac{26.4}{81.2} = 0.3251...$ $a = \frac{46}{5} - 0.3251 \times \frac{143}{5}$ $= 9.2 - 0.3251 \times 28.6 = -0.0985$ $y = 0.325x - 0.0985$ $x = 20 \Rightarrow \hat{y} = 6.40$ or 6 Extrapolation beyond range of data Small sample / only based on one sample	M1 M1 A1 M1 A1 B1 B1 **[7]**	Calculating b from *their* S_{xx}, S_{xy} (calculation seen or 0.32 to 0.33) Calculating a from Σx, Σy and *their* b (seen or implied) –0.10 to –0.09 for a (without wrong working) Line correct with coefficients 0.32 to 0.33 and –0.10 to –0.09 used correctly 6.3 to 6.5 (before rounding) or integer 6 Extrapolation Any valid objection
5 (i)	H_0: population mean = 0.2 H_1: population mean > 0.2	B1 B1 **[2]**	'population mean' or μ (but not just 'mean') > 0.2
(ii)	X = reaction time Assuming H_0: $X \sim N(0.2, 0.001)$ $0.2 + 1.645 \times \sqrt{0.001} = 0.252$ Would not expect a reaction time > 0.252 if H_0 is true. If the reaction time is greater than 0.252 it provides evidence that Sam is right and $\mu > 0.2$, otherwise there is insufficient evidence to reject H_0.	 M1 A1 B1 **[3]**	 Use of 1.645 cv = 0.252 Explaining how (*their*) cv is used

(iii)	P(Type II error) = P($X \leqslant 0.252$) where $X \sim N(0.24, 0.001)$ = P($Z \leqslant 0.379$) = 0.6477	M1 M1 A1 **[3]**	$X \sim N(0.24, 0.001)$ used or implied P($X \leqslant$ or < cv) 0.647 to 0.648 as final answer without wrong working, cao
6 (i)	H_0: No association between attractiveness of men and their wives H_1: Some association	B1 **[1]**	H_0: No association (allow independence) H_1: Association, or equivalent
(ii)	$\frac{18}{50} \times \frac{16}{50} \times 50$ or $18 \times 16 \div 50 = 5.76$	M1 A1 **[2]**	18 × 16, 288, 0.32, 0.36 or 0.1152 seen Correct method (answer given in question)
(iii)	$\frac{(10-5.76)^2}{5.76} = \frac{4.24^2}{5.76} = \frac{17.9776}{5.76} = 3.12$	B1 B1 **[2]**	$(10-5.76)^2$ ÷ 5.76
(iv)	$\nu = (3-1)(3-1) = 4$ degrees of freedom Critical value = 13.28 $38.24 > 13.28 \Rightarrow$ reject H_0 (accept H_1) More attractive men usually have more attractive wives and less attractive men have less attractive wives.	M1 A1 M1 A1 **[4]**	4 df 13.28 Reject H_0 or equivalent, correct for their cv A correct conclusion in words
7	Weight loss 8 15 17 10 21 24 32 6 19 46 Rank 2 4 5 3 7 8 9 1 6 10 Type N N N N N N N E E E Rank sum E = 17 ($R_m = 17$) $m(m+n+1) - R_m = 16$ $W = 16$ H_0: same distribution $\alpha = 5\%$ H_1: different distributions two-sided $m = 3, n = 7 \Rightarrow$ cv = 7 $16 > 7 \Rightarrow$ accept H_0 Insufficient evidence to claim that there is a difference in the weight losses.	B1 M1 M1 A1 B1 M1 A1 **[7]**	Ranks correct (at least for E) $R_m = 17$ $m(m+n+1) - R_m$ $W = 16$, cao cv = 7 Accept H_0 (ft *their* W and cv) Correct conclusion in words, from correct working for *their* W and c.

8 **(i)**	Estimate μ using $\bar{x} = \frac{2144}{16} = 134$ Estimate σ^2 using $s^2 = \frac{8640}{15} = 576$	B1 M1 A1 **[3]**	134 8640 ÷ either 15 or 16, 540, 24 or 23.2 576
(ii)	Population distribution is normal and unknown population variance Small sample	B1 B1 **[2]**	Normal and σ^2 unknown Sample size
(iii)	15 degrees of freedom $\Rightarrow t = 2.131$ $134 \pm 2.131 \times \sqrt{(576/16)}$ $= 134 \pm 2.131 \times 6$ $= (121.2, 146.8)$	M1 A1 B1 M1 A1 **[5]**	$n - 1 = 15$ 2.131 576/16 seen or implied Correct method for *their* t or z value, condone omission of 16 Confidence limits correct (3 sf or better)
(iv)	$\bar{X} \sim N(145, 100)$ $P(\bar{X} < 125) = P(Z < \frac{125 - 145}{10})$ $P(Z < -2) = 1 - 0.9772$ $= 0.0228$	B1 M1 M1 A1 **[4]**	N(145, 100) Correct method for their distribution Z value –2 or +2 0.0228