

Mark Scheme (Results)

Summer 2014

Pearson Edexcel GCE in Statistics S1 (6683/01)

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2014
Publications Code UA040117
All the material in this publication is copyright
© Pearson Education Ltd 2014

General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

PEARSON EDEXCEL GCE MATHEMATICS

General Instructions for Marking

- 1. The total number of marks for the paper is 75
- 2. The Edexcel Mathematics mark schemes use the following types of marks:
- M marks: Method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
- A marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
- **B** marks are unconditional accuracy marks (independent of M marks)
- Marks should not be subdivided.

3. Abbreviations

These are some of the traditional marking abbreviations that will appear in the mark schemes.

- bod benefit of doubt
- ft follow through
- the symbol $\sqrt{}$ will be used for correct ft
- cao correct answer only
- cso correct solution only. There must be no errors in this part of the question to obtain this mark
- isw ignore subsequent working
- awrt answers which round to
- SC: special case
- oe or equivalent (and appropriate)
- d... or dep dependent
- indep independent
- dp decimal places
- sf significant figures
- * The answer is printed on the paper or ag- answer given
- L or d... The second mark is dependent on gaining the first mark
- 4. All A marks are 'correct answer only' (cao.), unless shown, for example, as A1 ft to indicate that previous wrong working is to be followed through. After a misread however, the subsequent A marks affected are treated as A ft, but manifestly absurd answers should never be awarded A marks.

- 5. For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.
- 6. If a candidate makes more than one attempt at any question:
 - If all but one attempt is crossed out, mark the attempt which is NOT crossed out.
 - If either all attempts are crossed out or none are crossed out, mark all the attempts and score the highest single attempt.
- 7. Ignore wrong working or incorrect statements following a correct answer.

Question Number	Scheme				
1. (a)	<i>a</i> = 44	These answers may be in or near the	B1		
	<i>b</i> = 76	table	B1		
			(2)		
(b)	55+1.5(55-31)=91 [and $31-1.5(55-31)$	'	M1		
	<u>Penv</u>	<u>ille</u>	D1		
			B1		
	<u> </u>		B1		
			A1		
	0 10 20 30 40 50 60	70 80 90 100 110 120			
	Years		(4)		
(c)	Greenslax: $[Q_2 - Q_1 = 20, Q_3 - Q_2 = 12]$	or $(Q_2 - Q_1) > (Q_3 - Q_2)$] \Rightarrow -ve(skew)	B1		
	Penville: $[Q_2 - Q_1 = 8, Q_3 - Q_2 = 16 \text{ or}$	$(Q_3 - Q_2) > (Q_2 - Q_1) \implies +\text{ve (skew)}$	B1		
	Don't insist on seeing "skew" so just -ve an	d +ve will do. Treat "correlation" as ISW			
	Justification that is consistent		ddB1 (3)		
			Total 9		
	Not		M1		
	A fully correct box plot scores 4/4. If <u>not</u> fully correct apply scheme and need evidence for M1 If two box plots are seen ignore the one for Greenslax. If not on graph paper M1 max for (b)				
(b)	M1 for sight of $55+1.5(55-31)$ or 91 see	n (possibly implied by RH whisker of box r	olot)		
	May be implied by a fully correct box				
	1 st B1 box with whiskers (condone missing median)				
	2 nd B1 25, 31, 39, 55, RH whisker to end at 75 or 91. Two RH whiskers is B0 Accuracy must be to within 0.5 of a square so e.g. lower quartile at 30 or 32 is OK				
	A1 only one outlier plotted at 99. Allow cross to be vertically displaced If the RH whisker goes to 99 (2 nd B0) and A0 even if outlier is identified since we require				
	a horizontal "gap" between RH whisker and outlier.				
(c)	1 st B1 Greenslax – ve (skew) We must l				
	implied by their <u>values</u> but not simply from $Q_3 - Q_2 > Q_2 - Q_1$ If there is just <u>one</u> , unlabelled comment assume Penville. 3^{rd} ddB1 dependent on 1^{st} and 2^{nd} B marks being scored. Justification for <u>both</u> based on: quartiles, median relative to quartiles, or "tail"				
	If only values for $Q_3 - Q_2$ etc are given they should be correct ft for Greenslax and correct for Penville If values for Greenslax imply +ve skew then 1^{st} B0 and 3^{rd} B0				

Question Number	Scheme	Marks		
2	mean = $\frac{60.8 + 20}{1.4}$ or $60.8 = 1.4x - 20$ (o.e.)	M1		
	= 57.7142 awrt 57.7	A1		
	standard deviation = $\frac{6.60}{1.4}$ or $6.60 = 1.4x$	M1		
	= 4.7142 awrt 4.71	A1		
		(4)		
		Total 4		
	Notes			
	1^{st} M1 sub. 60.8 for y into a correct equation. Allow use of x or any other letter or expression for mean			
	1 st A1 for awrt 57.7 or $\frac{404}{7}$ (o.e.). Correct answer only is 2/2			
	2^{nd} M1 sub. 6.60 or 6.6 for y and ignoring the 20 Allow use of x or any other letter or expression for st. dev.			
	$6.60^2 = 1.4^2 x^2$ is M0 until we see them take a square root. 2^{nd} A1 for awrt 4.71 or $\frac{33}{7}$ (o.e.). Correct answer only is 2/2			

Question Number	Scheme			
3 (a)	$r = \frac{31512.5}{\sqrt{42587.5 \times 25187.5}} = 0.962$ awrt 0.962	M1 A1 (2)		
(b)	r is close to 1 or a strong correlation. ["points are close to a straight line" is B0]	B1		
	[Just "positive" correlation is B0] [Use of "relationship" or "skew" not "correlation" is B0]	(1)		
(c)	$b = \frac{31512.5}{42587.5} = 0.739947 = 0.740 \text{ (3 dp)}$ 0.740 (only)	M1 A1cao		
		(2)		
(d)	$a = 1326.25 - (0.7399 \times 2423.75)$ [= -467.2 or awrt -467]	M1		
	So $m = -467 + 0.74v$	A1 (2)		
(e)	b is the <u>money (spent)</u> per <u>visitor</u> . (i.e. definition of a rate in words.)[ignore values] So each 1000 visitors generates an extra £0.74 million <u>or</u> each visitor spends £740 oe	B1 B1ft (2)		
(f)	$m = -467 + 0.74 \times 2500$ m = 1383 (£ million) awrt 1380	M1 A1		
		(2)		
(g)	As 2500 is within the range of the data set <u>or</u> it involves <u>interpolation</u> . The value of money spent is reliable	B1 dB1 (2)		
	Notes			
(a)	M1 for a correct expression for r. Ans only of 0.96 or awrt 0.96 is M1A0 Ans only of 0.962 or awrt 0.962 is M1A1. Do not allow fractions for A1			
(b)	B1 for comment implying strong correlation. (e.g. big/high/clear etc) B0 if $ r > 1$			
(c)	M1 for a correct expression for b (may be implied by 0.74 or better in regression equation)			
	A1 A1 for 0.740 only in (c) or $b = 0.740$ seen elsewhere (M1A0 for $\frac{2521}{3407}$ or awrt 0.74 here)			
(d)	M1 for $1326.25 - (\text{'their } b' \times 2423.75)$ Condone fractions or awrt 1330 for \overline{m} and awrt 2420 for \overline{v} A1 for a correct equation in m and v with $a = \text{awrt} - 467$ and $b = \text{awrt } 0.74$ Condone $\frac{2521}{3407}$ for b and $\frac{-1591740}{3407}$ for a . [Equation in y and x is A0]			
(e)	1 st B1 for a correct definition of the rate in words. Must state or imply "money per visitor"			
	Allow alternative words or symbols e.g. £ or "pounds" for money, "people" for	visitors etc		
	2^{nd} B1ft for a correct numerical rate (ft their value of b)			
	e.g. "each <u>visitor spends</u> £740" is B1B1, "b is the extra <u>money</u> spent per <u>visitor</u> " is B1B0 [no visitor]	/alues]		
	"b is increase of £0.74 million in m as v increases by 1000" is B0B1[£ for money but no "visit "increase in \underline{m} as \underline{v} increases" is B0B0 [Idea of rate but letters not words and no numerical value.	_		
(f)	M1 sub. $v = 2500$ into their equation. Simply substituting 2 500 000 is M0 (unless ac	ljusted eqn)		
	A1 awrt 1380 units (£ and million not required)			
(g)	1 st B1 for 2500 or 2 500 000 or visitors or v is in range. "it" is B0 unless v clearly in	nplied		
	2 nd dB1 for stating it <u>is</u> reliable. Dependent on previous B mark being awarded			
	"both <i>v</i> and <i>m</i> in range" or "1380 in range" is B0 but use ISW so "interpolation since by range" scores B1 for the "interpolation". "Not extrapolation" counts as "interpolation"			

Question Number	Scheme				
4 (a)	0.02 B P(JnB) = 0.005 or 1	M1			
	$J = 0.98$ B' P(JnB') = 0.245 or $\frac{49}{200}$				
	0.25 0.03 B P(KAB) = 0.0135 or 27	A1			
	0.97 B' P(KnB') = 0.4365 or 873				
	0.3 $L = 0.05$ B P(LAB) = 0.015 or $\frac{3}{200}$				
	0.3 L 0.05 B P(LAB) = 0.015 or $\frac{3}{200}$ 0.95 B' P(LAB') = 0.285 or $\frac{57}{200}$	(2)			
(b)	0.25×0.98 , = 0.245 (or exact equiv. e.g. $\frac{49}{200}$)	M1A1			
		(2)			
(c)	$0.25 \times 0.02 + 0.45 \times 0.03 + 0.3 \times 0.05$, = 0.0335 (or exact equiv. e.g. $\frac{67}{2000}$)	M1A1			
		(2)			
(d)	$[P(J \cup L \mid B)] = \frac{0.25 \times 0.02 + 0.3 \times 0.05}{0.0335} \qquad \underline{\text{or}} \qquad \frac{0.0335 - 0.45 \times 0.03}{0.0335}$	M1A1ft			
	= 0.5970 awrt 0.597 (or $\frac{40}{67}$ or exact equiv.)	A1			
		(3)			
	Notes				
	Allow fractions or percentages throughout this question				
(a)	Allow 3+6 tree diagram with the 6 correct "end" probs and labels to get 2/2 (1 st , 3 rd , 5 th M1 for (3+6) tree drawn with 0.25, 0.45, 0.02, 0.03, 0.05 on correct branches	gets M1)			
	A1 for 0.3, 0.98, 0.97, 0.95 on the correct branches and labels, condone missing B' s Correct answer only scores full marks for parts (b), (c) and (d) When using "their probability p " for M1 and A1ft they must have 0				
(b)	M1 for $0.25 \times$ 'their 0.98 ' o.e.				
(c)	M1 for $0.25 \times$ their $0.02 + 0.45 \times$ their $0.03 +$ their $0.3 \times$ their 0.05 Condone 1 transcription error. Or $1 - (0.25 \times$ their $0.98 + 0.45 \times$ their $0.97 +$ their $0.3 \times$ their $0.95)$				
(d)	M1 for use of conditional probability with their (c) as denominator. Also exactly 2 products on num' and at least one correct (or correct ft) or their (c) – one of the products from their (c). Ignore an incorrect expression inside their probability statement				
	A1ft for $\frac{0.25 \times \text{their } 0.02 + \text{their } 0.3 \times \text{their } 0.05}{\text{their } (c)} \underbrace{\text{or }}_{\text{their } (c)} \frac{\text{their } (c) - 0.45 \times \text{their } 0.03}{\text{their } (c)} \underbrace{\text{or }}_{\text{their } (c)} \frac{0.02}{\text{their } (c)}$				
	A1 awrt 0.597 or exact fraction e.g. $\frac{40}{67}$				

Question Number	Scheme					
5 (a)	2k + 4k + 6k + k(8 - 2) = 1 (commas instead of + or a table OK if $18k = 1$ seen later)	M1				
	$k = \frac{1}{18} (*)$					
(b)	$[2k + 4k] = \frac{6}{18} = \frac{1}{3}$ (\frac{1}{3} or any exact <u>numerical</u> equivalent)					
(c)	$E(X) = \left(2 \times \frac{1}{9}\right) + \left(4 \times \frac{2}{9}\right) + \left(6 \times \frac{1}{3}\right) + \left(8 \times \frac{1}{3}\right) \underline{\text{or}} \left(2 \times 2k\right) + \left(4 \times 4k\right) + \left(6 \times 6k\right) + \left(8 \times 6k\right)$					
	$= 5\frac{7}{9} \text{(or exact equivalent e.g. } \frac{52}{9} \text{)}$	A1 (2)				
(d)	$E(X^{2}) = \left(4 \times \frac{1}{9}\right) + \left(16 \times \frac{2}{9}\right) + \left(36 \times \frac{1}{3}\right) + \left(64 \times \frac{1}{3}\right) \underline{\text{or}} \left(4 \times 2k\right) + \left(16 \times 4k\right) + \left(36 \times 6k\right) + \left(64 \times 6k\right)$	M1				
	$= 37\frac{1}{3} \text{ (or exact equivalent e.g. } \frac{112}{3})$	A1 (2)				
(e)	Var $(X) = 37\frac{1}{3} - \left(5\frac{7}{9}\right)^2$ [= 3.95 or $\frac{320}{81}$]	M1				
	$Var(3-4X) = 16 \times 3.95$					
	= awrt 63.2 (allow $\frac{5120}{81}$)	A1 (3)				
		Total 10				
(-)	Notes					
(a)	M1 for $2k + 4k + 6k + k(8 - 2) = 1$ A1 for $k = \frac{1}{18}$ NB cso so no incorrect working seen M1 for $2 \times \frac{1}{18} + 4 \times \frac{1}{18} + 6 \times \frac{1}{18} + \frac{1}{18}(8 - 2)$ A1 for =1 and "therefore $k = \frac{1}{18}$ "					
(c)	If in parts (c), (d) and (e) there is a correct expression worthy of M1 but later they incorrectly g and multiply or divide by some number n, then allow the M1 but mark their final answer (A0) Answers only in (b), (c), (d) and (e) score all the marks. (c) M1 for an expression for E(X) with at least 3 correct terms (products) allow use of k e.g. 1					
(d)						
(e)	1 st M1 for $E(X^2) - [E(X)]^2$ ft their answers to (c) and (d). Must see values <u>used</u> correct	ilv.				
2^{nd} M1 for statement " $4^2 \times \text{Var}(X)$ " seen <u>or</u> for $4^2 \times \text{their Var}(X)$ provided their Var Do not allow for $16 \times \text{E}(X^2)$ but can score M0M1 NB condone $-4^2 \times \text{Var}(X)$ if the answer later becomes positive.						
	A1 for exact fraction ($\frac{5120}{81}$ o.e.) or decimal approximation that is awrt 63.2					
	Beware: rounding to 3sf in (c) (5.78) and (d) (37.3) gives 62.3 which could be misread as 63.2					

Question Number	Scheme	Marks			
6 (a)	70 – 80 group - width 0.5 (cm)	B1			
	1.5 cm ² is 10 customers or 3.75cm ² is 25 customers or $0.5c = 3.75$ or $\frac{2.5}{\frac{1}{3}}$				
	70 – 80 group - height 7.5 (cm)				
(b)	Median = $(70) + \frac{13.5}{25} \times 10$ allow $(n+1) = (70) + \frac{14}{25} \times 10$	(3) M1			
	= 75.4 (or if using $(n + 1)$ allow 75.6)	A1 (2)			
(c)		B1			
	$\sigma = \sqrt{\frac{529400}{85} - 76^2}$	M1			
	= 21.2658 ($s = 21.3920$) awrt 21.3	A1 (3)			
(d)	Coeff' of skewness = $\frac{3(76-75.4)}{21.2658}$ = 0.08464 awrt 0.08 (awrt 0.06 for 75.6)	(3) M1 A1			
	There is (very slight) positive skew or the data is almost symmetrical (or both) <u>Any</u> mention of "correlation" is B0				
	Notes				
(a)	B1 for 0.5				
	M1 for one of the given statements <u>or</u> any method where "their width" × "their height" = 3.75 Correct height scores M1A1 independent of width so B0M1A1 is possible.				
(b)	Will for a correct fraction: $+\frac{1}{25} \times 10$ where $k = 13.5$ or 14 for $(n + 1)$ case.				
	NB may work down so look out for (80) $-\frac{11.5}{25} \times 10$ etc Beware: $69.5 + \frac{13.5}{25} \times 11 = 75.44$ (but M0)				
(c)	M1 for a correct expression with square root, ft their mean A1 for awrt 21.3 or, if clearly using <i>s</i> allow awrt 21.4. Must be evaluatedno surds.				
(d)	M1 sub. their values into formula allow use of s but their σ or s must be > 0 A1 for awrt 0.08 but accept 0.085 No fraction B1ft for a correct comment compatible with their coefficient. Allow "symmetrical" for coeff' < 0.25 They may say it is "slightly skew" so omit "positive" but do not allow "negative" if	coef' +ve			
	Condone "strongly" positive skew.				

Question Number	Scheme			
7 (a)	The random variable $H \sim \text{height of females}$ $P(H > 170) = P\left(Z > \frac{170 - 160}{8}\right) [= P(Z > 1.25)]$			
	=1-0.8944			
	= 0.1056 (calc 0.1056498) awrt 0.106 (accept 10.6%)	A1 (3)		
(b)	$P(H > 180) = P(Z > \frac{180 - 160}{8})$ [=1-0.9938]	M1		
	= 0.0062 (calc 0.006209) awrt 0.0062 or $\frac{31}{5000}$	A1		
	$[P(H>180 H>170)] = \frac{0.0062}{0.1056}$	M1		
	= 0.0587 (calc 0.0587760) awrt 0.0587 or 0.0588	A1 (4)		
(c)	$P(H > h H > 170) (= 0.5)$ or $\frac{P(H > h)}{P(H > 170)} (= 0.5)$	M1		
	$[P(H > h)] = 0.5 \times "0.1056" = 0.0528 \text{ (calc } 0.0528249) or } [P(H < h)] = 0.9472$	A1ft		
	$\frac{h-160}{8} = 1.62 \text{(calc } 1.6180592)$			
	h = awrt 173 cm awrt 173			
	Notes			
(a)	1 st M1 for attempt at standardising with 170, 160 and 8. Allow \pm i.e. for $\pm \frac{170-160}{8}$			
	2^{nd} M1 for attempting $1 - p$ where $0.8 . Correct answer only 3/3$			
(b)	1 st M1 for standardising with 180, 160 and 8 1 st A1 for 0.0062 seen, maybe seen as part of another expression/calculation.			
	2^{nd} M1 using conditional probability with denom = their (a) and num < their denom. <u>Values</u> needed. 2^{nd} A1 for awrt 0.0587 <u>or</u> 0.0588. Condone 5.87% or 5.88% or $\frac{31}{528}$ Correct answer only 4/4			
(c)	ot.			
	1^{st} A1ft for $[P(H > h)] = 0.5 \times \text{their}(a)$ Award M1A1ft for correct evaluation of $0.5 \times \text{their}(a)$ or sight of 0.0528 or better			
	2^{nd} M1 for attempt to standardise (\pm) with 160 and 8 and set equal to $\pm z$ value (1.56 < $ z $ < 1.68)			
	B1 for $(z =)$ awrt ± 1.62 (seen)			
	2 nd A1 for awrt 173 but dependent on <u>both</u> M marks.			

Question Number	Scheme				Marks
8 (a)	[P(A) = 1 - 0.18 - 0.22] = 0.6 (or exact equivalent)			B1	
					(1)
(b)	$P(A \cup B) = "0.6" + 0.22 = 0.82$ (or exact equivalent)			B1ft	
					(1)
(c)					M1
	$\frac{x}{x+0.22} = 0.6$	$P(B) \times [1 - 0.6] = 0$.22	Find $P(B)$	
	x = 0.6x + 0.132	Use $P(A \cap B) = P(A \cap B)$	$A \mid B)P(B)$	Use $P(B)P(A) = P(A \cap B)$	dM1
	0.4x = 0.132	$P(A \cap B) = 0.6 \times 0.$	55	$P(A \cap B) = 0.6 \times 0.55$	divii
		x = 0.33 (o	r exact equivalent)		A1cso
(1)	D(D) 0.55				(3)
(d)	P(B) = 0.55	506	on stating D(A)	$\mathbf{D}(A \mathbf{D}) = 0.61$	3.61
	$P(B) \times P(A) = 0.53$ $= 0.33$		or stating P(A)	= P(A B) [= 0.6]	M1
	$P(B) \times P(A) = P(A)$		$\mathbf{or}\ \mathrm{P}(A) = \mathrm{P}(A A)$	R)	Alcso
	therefore (statistic	<i>'</i>		istically) independent	(2)
	(((((((((((((((((((Total 7	
	Notes				
(b)	B1ft for their (a) + 0.22 or 1 - $P(A' \cap B')$ Do not ft their (a) if it is > 0.78				
	NB 3 versions for (c). Check carefully that Ms are genuinely scored.				
				(B) = 0.55 check it is <u>derived</u>	
(c)	1 st M1 for a corr	ect equation for x e	$g. \frac{x}{x + 0.22} = 0.6$	or a correctly derived equation	on for P(B)
	2^{nd} dM1 for solving to get in form $kx = L$ or correct use of P(B) to find P(A \cap B) [2 nd or 3 rd ver] or P(A \cap B) = P(B) - 0.22				
	A1cso for 0.33 Dep. on both Ms and no incorrect working seen.				
(d)	M1 for finding $P(B) \times P(A) = 0.33$ (values needed) or stating $P(A) = P(A B)$ (= 0.6 not needed)				
	A1cso for a correct statement: $P(B) \times P(A) = P(A \cap B)$ or $P(A) = P(A B)$ and stating independent				dependent
	NB The M1 in (d) using $P(A \cap B)$ requires $P(B) = 0.55$			-	
	There is no ft of an incorrect $P(B)$ Full marks in (d) is OK even if $0/3$ in (c)			B	
	{This Venn diagram may be helpful.}			18	