Mark Scheme (Results)

October 2024

Pearson Edexcel International Advanced Level
In Statistics 1 (WST01) Paper 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.edexcel.com or 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

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.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

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 and can be used if you are using the annotation facility on ePEN:
   - bod – benefit of doubt
   - ft – follow through
     - the symbol $\sqrt{}$ ill 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
   - ⌐ 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. If you are using the annotation facility on ePEN, indicate this action by 'MR' in the body of the script.

6. If a candidate makes more than one attempt at any question:

   a) If all but one attempt is crossed out, mark the attempt which is NOT crossed out.

   b) 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.

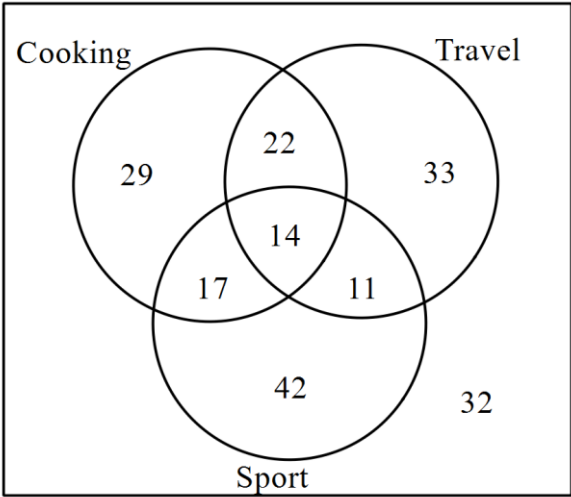### Special notes for marking Statistics exams (for AAs only)

- Any correct method should gain credit. If you cannot see how to apply the mark scheme but believe the method to be correct then please send to review.

- For method marks, we generally allow or condone a slip or transcription error if these are seen in an expression. We do not, however, condone or allow these errors in accuracy marks.

Question Number	Scheme	Marks
1 (a)	[Mode =] 137	B1
		(1)
(b)	$a = 106$ $b = 129$ $c = 126$	B1 B1 B1
		(3)
(c) (i)	$\left[\text{mean} = \dfrac{4016}{31} = \right] 129.5$	B1
(ii)	[Standard deviation =] $\sqrt{\dfrac{525056}{31} - \left(\dfrac{4016}{31}\right)^2}$ or $\sqrt{\dfrac{31}{30}\left(\dfrac{525056}{31} - \left(\dfrac{4016}{31}\right)^2\right)}$	M1
	$= 12.4$ or 12.6	A1
		(3)
(d)	$\dfrac{'129.5' - '137'}{'12.4'} [= -0.6]$	M1
	Negative [skew]	A1ft
		(2)
(e)	A correct difference of the average or a correct difference of the spread e.g. On average action films **run for longer** than comedy films as the **median** is greater **129 > 117**	B1ft
		(1)
Notes		**Total 10**
(a)	**B1**	Cao Must be seen in part (a)
(b)	**B1**	$a = 106$ Must be attached to a (Condone Q_1/lower quartile for a)
	B1	$b = 129$ Must be attached to b (Condone Q_2/median for b)
	B1	$c = 126$ Must be attached to c (Condone Q_3/upper quartile for c)
(c) (i)	**B1**	awrt 130
(ii)	**M1**	For a correct method (including the square root) to find the standard deviation. Ft their mean May be implied by awrt 12.4 or awrt 12.9 or awrt 12.6 if sample standard deviation is calculated
	A1	awrt 12.4 or awrt 12.6 if sample standard deviation is calculated Correct answer only scores 2/2
(d)	**M1**	For substitution of the mode, the mean and the standard deviation into the expression. Ft their mode, their mean and their standard deviation. Allow awrt −0.6 if no working shown.
	A1ft	A correct interpretation ft their expression. Ignore any reference to correlation
(e)	**B1ft**	for a correct comment, **referring to length**, with reference to a **correctly named statistic.** Must **include the correct figures** compared. Ft their values. Ignore any reference to skew

The following table may be useful (but not an exhaustive list of possible answers)

Statistic	A	C	Comment
Median	129	117	On average action films **run for longer** as the **median** is greater **129 > 117**
Mean	129.5	114.5	On average action films have a **greater running time** as the **mean** is greater **129.5 > 114.5**
Mode	137	127	On average comedy films are **shorter in length** as the **mode** is less **127 < 137**
IQR	17	20	Comedy films have a **greater spread of running times** as the **IQR** is greater **20 > 17**
Range	57	39	Action films have a **larger variation of running times** as the **range** is greater **57 > 39**
SD	12.4	11.9	Comedy films show **less variability in the length** of films as the **SD** is less **11.9 < 12.4**
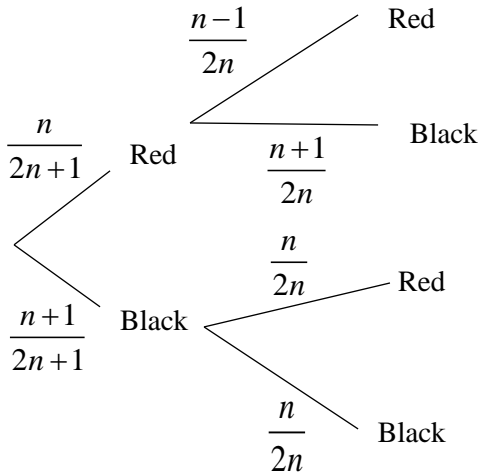
Question Number		Scheme	Marks
2 (a) (i)		$S_{yy} = 81938.5 - \dfrac{2015^2}{50}\ [= 734]\ *$	B1*
(ii)		$r = \dfrac{219.55}{\sqrt{734 \times 72.25}} = 0.95338...$ awrt 0.953	M1 A1
			(3)
(b)		e.g. [In general] the longer the rabbit the greater the weight	B1ft
			(1)
(c)		Consistent/Yes as r/PMCC is close to 1	B1ft
			(1)
(d)		$b = \dfrac{219.55}{734} = 0.2991...$	M1 A1
		$a = \left(\dfrac{125}{50}\right) - {}'b'\left(\dfrac{2015}{50}\right)[= -9.554...]$	M1
		$w = -9.55 + 0.299\,y$	A1
			(4)
(e)		$'-9.55' + '0.299' \times 45 = 3.905$ awrt 3.91	M1 A1ft
			(2)
		Notes	**Total 11**
(a)(i)	**B1***	Answer is given so a correct numerical expression and no incorrect working seen	
(ii)	**M1**	For use of $\dfrac{S_{yw}}{\sqrt{S_{yy} \times S_{ww}}}$ May be implied by awrt 0.953	
	A1	awrt 0.953	
(b)	**B1ft**	A correct interpretation ft their r value (provided that $|r| < 1$) e.g as length/y increases the weight/w increases Ignore any figures quoted. Do not accept comments about correlation on their own	
(c)	**B1ft**	A correct statement with a correct reason ft their r value (provided that $|r| < 1$) Allow $0.953 \approx 1$ Allow 'my value' to imply r	
(d)	**M1**	A correct method to find the gradient (May be implied by awrt 0.299 or $\frac{4391}{14680}$)	
	A1	awrt 0.299 (Condone awrt 0.3 if M1 scored) May be implied by a correct gradient in the regression line.	
	M1	A correct method to find the intercept ft their b (May be implied by awrt –9.55)	
	A1	For $w = (\text{awrt}) -9.55 + (\text{awrt})0.299\,y$ Must be seen in part (d)	
(e)	**M1**	For substitution of 45 into their regression equation	
	A1ft	For awrt 3.91 or ft their regression line, provided their final answer is > 0. If regression line is not correct, then you will need to check their answer	

Question Number		Scheme	Marks
3 (a)		Cooking 29, 22, 33 Travel, 14, 17, 11, 42, 32 Sport	B1 B1 B1
			(3)
(b) (i)		$\dfrac{'14'}{200}$	B1ft (1)
(ii)		$\dfrac{'33'+'11'+'42'+'32'}{200}=\dfrac{118}{200}$ or $\dfrac{200-'29'-'22'-'14'-'17'}{200}=\dfrac{118}{200}$	M1 A1
			(2)
(c)		$\dfrac{'14'+'11'}{'14'+'11'+'22'+'33'}=\dfrac{25}{80}$	M1 A1
			(2)
Notes			**Total 8**
(a)	**B3**	For a fully correct Venn diagram (B2 for at least 6 numbers in the correct place on the Venn diagram) (B1 for at least 3 numbers in the correct place on the Venn diagram) Treat blanks on the diagram as zero Condone correct probabilities instead of frequencies e.g. $^{22}\!/_{200}$ oe or 0.11 as 22	
(b) (i)	**B1ft**	For $\dfrac{'14'}{200}$ oe ft their Venn diagram	
(ii)	**M1**	$\dfrac{'33'+'11'+'42'+'32'}{200}$ or $\dfrac{200-'29'-'22'-'14'-'17'}{200}$ ft their Venn diagram. May be implied by $\dfrac{118}{200}$ oe	
	A1	For $\dfrac{118}{200}$ oe	
(c)	**M1**	$\dfrac{n}{'14'+'11'+'22'+'33'}$ provided the answer gives a probability and $0 < n < 80$ ft the Venn diagram for the denominator or $\dfrac{m}{'0.4'}$ where $0 < m < 0.4$ ft the Venn diagram for the denominator	$\dfrac{'n'}{80}$ provided $0 < n < 80$ ft the Venn diagram for numerator or $\dfrac{'m'}{0.4}$ where $0 < m < 0.4$ ft the Venn diagram for numerator
	A1	For $\dfrac{25}{80}$ oe Allow 0.313	

Question Number		Scheme	Marks
4 (a)		$X \sim N(170,16^2)$	
		$P(X > 190) = P\left(Z > \dfrac{190-170}{16}\right)\left[= P(Z > 1.25)\right]$	M1
		$[= 1 - 0.8944] = 0.1056$ awrt 0.106	A1
			(2)
(b)		$P(X > d) = 0.9$	
		$\dfrac{d-170}{16} = -1.2816$ or $\dfrac{170-d}{16} = 1.2816$ (Calc value $\pm 1.28155...$)	M1 A1
		$d = 149.494...$ awrt 149	dA1
			(3)
		Notes	**Total 5**
		NB correct answers with no working scores no marks	
(a)	**M1**	For standardising using 190, 170 and 16 (May be implied by 1.25)	
	A1	awrt 0.106 Do not ISW	
(b)	**M1**	For standardising and setting = z value, where $1 < |z| < 2$	
	A1	A fully correct standardisation = a correct compatible z value to 4 dp or better.	
	dA1	Dependent on previous A1 149.5 or awrt 149	

Question Number	Scheme	Marks
5 (a)	[Time is] continuous	B1
		(1)
(b)	16 photographers = 160 small squares or 64 photographers = 640 small squares or $\dfrac{16}{160}=0.1$ or $\dfrac{64}{640}=0.1$ or $\dfrac{160}{16}=10$ or $\dfrac{640}{64}=10$ Frequency density = 1.6 or Correct scale on the frequency density axis	M1
	$\dfrac{x}{240}=\dfrac{16}{160}$ oe or $\dfrac{(20-12)}{10}\times16+\dfrac{(24-20)}{5}\times14$	M1
	= 24	A1
		(3)
(c)	**Using n** $[Q_2=]20+\dfrac{(32-21)}{35-21}\times(25-20)$ oe or $25-\dfrac{35-32}{35-21}\times(25-20)$ oe $\qquad$ **Using $n+1$** $[Q_2=]\,20+\dfrac{(32.5-21)}{35-21}\times(25-20)$ oe or $25-\dfrac{35-32.5}{35-21}\times(25-20)$ oe	M1
	= awrt 23.9 $\qquad\qquad$ = awrt 24.1	A1
		(2)
(d)	Mean = Median or Mean $\approx$ Median	M1
	e.g. Appropriate decision. Consistent with expectation for a normal distribution.	A1ft
		(2)
	Notes	**Total 8**
(a)	**B1**	Allow not discrete
(b)	**M1**	For establishing a ratio between photographers and area or calculating frequency density (may be implied by 2nd M1)
	M1	For a correct ratio or expression using areas for photographers from 12 to 24
	A1	Cao
(c)	**M1**	For a correct method to find median using either n or $n+1$
	A1	awrt 23.9 or awrt 24.1 if using $n+1$
(d)	**M1**	For a correct comment about mean and median ft their median Allow mean is close to median to imply mean $\approx$ median Ignore any comments made about the shape of the histogram
	A1ft	For a correct compatible comment about Charlie's decision ft their median If Mean = Median or Mean $\approx$ Median, then the decision should be that a normal distribution is suitable [due to symmetry] If Mean < Median or Mean > Median or mean $\neq$ median, then the decision should be that a normal distribution is not suitable [due to the skew in the data]

Question Number		Scheme	Marks
6(a)		$[10k=1\Rightarrow]\,k=0.1$	B1
			(1)
(b)		e.g. $P(X=1)=0.1$ and $P(X=2)\big[=F(2)-F(1)\big]=0.1$	B1
		e.g. $P(X=3)\big[=F(3)-F(2)\big]=0.1$	M1
		<table><tr><td>X</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td></tr><tr><td>$P(X=x)$</td><td>0.1</td><td>0.1</td><td>0.1</td><td>0.2</td><td>0.2</td><td>0.3</td></tr></table>	A1
			(3)
(c)		$a+a+a+b+b+b+0.11+0.05=1\ \ [3a+3b=0.84]$	M1
		$a+2a+3a+4b+5b+6b+0.77+0.4=4.02\ \ [\Rightarrow 6a+15b=2.85]$	M1
		e.g. $9a=1.35\Rightarrow a=0.15*$	A1*
			(3)
(d)		$b=0.13$	
		$E(Y^2)=1^2\times0.15+2^2\times0.15+3^2\times0.15+4^2\times'0.13'+5^2\times'0.13'+6^2\times'0.13'$	M1
		$+7^2\times0.11+8^2\times0.05=[20.7]*$	A1*
			(2)
(e)		$\big[Var(Y)=\big]20.7-4.02^2[=4.5396]$	M1
		$Var(5-2Y)=4Var(Y)=4\times'4.5396'=18.1584$ \hfill awrt 18.2	M1 A1
			(3)
(f)		$'0.1'\times0.15+'0.1'\times0.15=0.03$	M1 A1
			(2)
		Notes	**Total 14**
(a)	**B1**	For $k=0.1$ oe	
(b)	**B1**	For correct use of F(x) to find 2 probabilities May be implied by two correct probabilities	
	M1	For correct use of F(x) to find one other probability. May be implied by one other correct probability	
	A1	For a fully correct probability distribution. Need not be in a table but probabilities must be attached to the correct X values	
(c)	**M1**	For use of the sum of the probabilities $=1$ to form a linear equation in a and b (May be implied by $3a+3b=0.84$)	
	M1	For use of $\sum y\times P(Y=y)=4.02$ to form a linear equation in a and b (May be implied by $6a+15b=2.85$)	
	A1*	Answer is given so there must be a correct line between the 2 equations and the given answer	
(d)	**M1**	For finding $E(Y^2)$ with their b (At least 4 correct terms). Values for a and b must be substituted into their expression for $E(Y^2)$	
	A1*	For a fully correct expression (no incorrect working seen) eg. $14\times0.15+77\times0.13+49\times0.11+64\times0.05$ Allow $0.15+0.6+1.35+2.08+3.25+4.68+5.39+3.2$	
(e)	**M1**	For a correct expression for $Var(Y)$. May be implied by awrt 4.54 Allow $E(5-2Y)=-\frac{76}{25}$ oe and $E((5-2Y)^2)=\frac{137}{5}$	
	M1	For use of $4Var(Y)$ ft their $Var(Y)$ provided $Var(Y)$ is not 20.7 or 4.02 Do not allow $5+4Var(Y)$ Allow $Var(5-2Y)=\frac{137}{5}-\left(-\frac{76}{25}\right)^2$ ft their $E(5-2Y)$ and their $E((5-2Y)^2)$	
	A1	awrt 18.2	
(f)	**M1**	For use of $P(X=1)\times P(Y=2)+P(X=2)\times P(Y=1)$ ft their X distribution probabilities	
	A1	Cao	

Question Number	Scheme	Marks
7 (a)	$\dfrac{n}{2n+1}$ and $\dfrac{n+1}{2n+1}$ in the correct places on the tree diagram	B1
	$\dfrac{n-1}{2n}$ and $\dfrac{n+1}{2n}$ in the correct places on the tree diagram	B1
	$\dfrac{n}{2n}$ and $\dfrac{n}{2n}$ in the correct places on the tree diagram	B1
		(3)
(b)	$'\dfrac{n}{2n+1}' \times '\dfrac{n+1}{2n}' + '\dfrac{n+1}{2n+1}' \times '\dfrac{n}{2n}'$	M1
	$= \dfrac{2n(n+1)}{2n(2n+1)} = \dfrac{n+1}{2n+1}$ *	A1*
		(2)
(c)	$\dfrac{n+1}{2n+1} = \dfrac{25}{49} \Rightarrow n = 24$ So 49 counters in the box	M1 A1
		(2)
(d)	$\dfrac{'\dfrac{25}{49}' \times '\dfrac{24}{48}'}{\dfrac{25}{49}} = \dfrac{1}{2}$	M1 A1
		(2)
Notes		**Total 9**

(a)	**B1**	For $\dfrac{n}{2n+1}$ and $\dfrac{n+1}{2n+1}$ in the correct places on the tree diagram
	B1	For $\dfrac{n-1}{2n}$ and $\dfrac{n+1}{2n}$ in the correct places on the tree diagram
	B1	For $\dfrac{n}{2n}$ and $\dfrac{n}{2n}$ in the correct places on the tree diagram. Allow $\dfrac{1}{2}$ for $\dfrac{n}{2n}$ in both places
(b)	**M1**	For use of P(Red) × P(Black) + P(Black) × P(Red) ft their tree diagram
	A1*	Answer is given so no incorrect working can be seen. Must have at least one correct line of working between M1 and the given answer.
(c)	**M1**	For solving to find $n = 24$
	A1	Cao
(d)	**M1**	For a correct ratio ft their n and their tree diagram Allow a correct ratio in terms of n e.g. $\dfrac{\dfrac{n+1}{2n+1} \times \dfrac{n}{2n}}{\dfrac{n+1}{2n+1}}$ oe ft their tree diagram for the numerator
	A1	For 0.5 (Working must be seen)

Question Number	Scheme	Marks
8 (a)	$\dfrac{162-\mu}{\sigma}=-1.2816$ (Calculator gives $-1.28155\ldots$) or $\quad\dfrac{175-\mu}{\sigma}=1.04$ (Calculator gives $1.03987\ldots$)	M1 A1 A1
	$\mu-1.2816\sigma=162$	
	$\mu+1.04\sigma=175$	
	$2.3216\sigma=13$	dM1
	$\sigma=5.5995\ldots\quad$ awrt 5.6 $\qquad \mu=169.1764\ldots\quad$ awrt 169	A1
		(5)

(b)	$Q_1=208.26$ or $Q_3-Q_1=13.48$	B1
	$221.74+1.5(221.74-'Q_1')[=241.96]$ or $'Q_1'-1.5(221.74-'Q_1')[=188.04]$	M1

Probability of an outlier	**1 – Probability of not an outlier**	
$\left[P(B>'241.96')=\right]$ $P\left(Z>\dfrac{'241.96'-215}{10}\right)[=P(Z>2.70)]$ or $\left[P(B<'188.04')=\right]$ $P\left(Z<\dfrac{'188.04'-215}{10}\right)[=P(Z<-2.70)]$ or	$\left[P('188.04'<B<'241.96')=\right]$ $P\left(\dfrac{'188.04'-215}{10}<Z<\dfrac{'241.96'-215}{10}\right)$ $[=P(-2.70)<Z<P(2.70)]$	dM1
$=0.0035$ (Calculator gives $0.0034883\ldots$)	$=0.993$ (Calculator gives $0.99298\ldots$)	
P(Outlier) $=2\times'0.0035'$	P(Outlier) $=1-'0.993'$	M1
$=0.007$ (Calculator gives $0.006976\ldots$) awrt 0.007	$=0.007$ (Calculator gives $0.007017\ldots$) awrt 0.007	A1
		(5)

		Notes	**Total 10**
(a)	**M1**	For standardising with μ and σ and setting = to a z value with $|z|>1$	
	A1	For one correct equation in any form with correct z value as given or better	
	A1	For a 2nd correct equation in any form allow 2dp or better for the z value	
	dM1	Dependent on previous M1. For solving their 2 linear simultaneous equations (Can be implied by both correct answers) If answers are incorrect then we need to see evidence of correct working	
	A1	For $\mu=$ awrt 169 and $\sigma=$ awrt 5.6	
(b)	**B1**	For $Q_1=208.26$ or $Q_3-Q_1=13.48$ Do not accept rounded values	
	M1	For a correct method for finding 1 outlier limit, ft their Q_1 or their IQR. You will need to check that these are correct if no working shown.	
	dM1	Dependent on previous M1. For standardising using their limit(s), 215 and 10 (allow $\pm$) May be implied by $\pm$ awrt 2.70 or awrt 0.0035 or awrt 0.993 or final answer of awrt 0.007	
	M1	If using the LHS of the MS: for multiplying their probability by 2 or If using the RHS of the MS: for 1 – their probability May be implied by awrt 0.007	
	A1	awrt 0.007	