

Mark Scheme (Results)

Summer 2023

Pearson Edexcel GCE In Statistics (9ST0) Paper 01: Data and Probability

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 <u>www.edexcel.com</u> or <u>www.btec.co.uk</u>. Alternatively, you can get in touch with us using the details on our contact us page at <u>www.edexcel.com/contactus</u>.

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 2023 Publications Code 9ST0_01_2306_MS All the material in this publication is copyright © Pearson Education Ltd 2023

General Marking Guidance

Total marks

The total number of marks for the paper is 80.

Mark types

The Edexcel Statistics mark schemes use the following types of marks:

- M Method marks, awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
- A Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
- **B Unconditional accuracy** marks are independent of M marks
- E Explanation marks

NOTE: Marks should not be subdivided.

Abbreviations

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

- ft follow through
- PI possibly implied
- cao correct answer only
- cso correct solution only (There must be no errors in this part of the question)
- awrt answers which round to
- awfw answers which fall within (a given range)
- SC special case
- nms no method shown
- oe or equivalent
- dep dependent (on a given mark or objective)
- dp decimal places
- sf significant figures
- ***** The answer is printed on the paper

Further notes

- 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.
- All A marks are 'correct answer only' (cao), unless shown, for example, as A1ft to indicate that previous wrong working is to be followed through.
- All M marks are 'possibly implied' (PI) unless specifically stated otherwise in the 'Notes' column.
- After a **misread**, the subsequent A marks affected are treated as A1ft, but manifestly absurd answers should never be awarded A marks.
- **Crossed out** work should be marked UNLESS the candidate has replaced it with an alternative response.
- If **two solutions** are given, each should be marked, and the resultant mark should be the mean of the two marks, rounded down to the nearest integer if needed.

Question	Scheme	Marks	AO	Notes
1(a)(i)	Advantages (not exhaustive)			
	The response rate is likely to be good.			
	Customers will be able to feed back as soon as the meal is finished.			Quick feedback
		E1	3.1a	Any sensible advantage
1(a)(ii)	Disadvantages (not exhaustive)			
	It will be expensive to pay a staff member to be there permanently.			
	The restaurant may not have the layout to station someone at the exit.			
	Only asks customers using the main exit			
	Customers may not like being accosted on the way out.			
	Customers may overhear and repeat each other's answers			
		E1	3.1a	Any sensible disadvantage

Question	Scheme	Marks	AO	Notes
1(b)(i)	Advantages (not exhaustive)			
	They are contacting everyone.			
	There will be minimal ongoing costs after the initial start-up.			or relatively cheap
	Provides online data in an easy to use form			
		E1	3.1a	Any sensible advantage
1(b)(ii)	Disadvantages (not exhaustive)			
	The response rate is likely to be low.			
	Lots of customers do not take receipts away.			oe
	Lots of customers do not look at the back of receipts.			
	Customers may not have access to the internet			
	Only the person who pays will have the receipt and therefore the survey			
		E1	3.1a	Any sensible disadvantage

Question	Scheme	Marks	AO	Notes
1(c)(i)	Advantages (not exhaustive)			
	The company may be able to contact the customers in the future for feedback or marketing purposes.			
	Quick to get the survey out.			
	Very cheap.			
	Easy to set up.			
	Provides online data in an easy to use form			
		E1	3.1a	Any sensible advantage
1(c)(ii)	Disadvantages (not exhaustive)			
	Lots of customers may not want to give their email addresses.			
	Asking for email addresses may make customers unhappy.			
	There may be data protection issues.			or GDPR oe
	It will take longer to take orders in the restaurant.			
	Only the person who pays will have the receipt and therefore the survey			
		E1	3.1a	Any sensible disadvantage

Question	Scheme	Marks	AO	Notes
1(d)	Method 2, as it is both cheap and inconvenient to customers as you aren't asked for your email address Method 3, is cheap, and not inconvenient to customers as you can't lose your receipt			
		E1	2.1b	Method 2 or 3 must be chosen as these minimise expense, and a comment must also be made about inconvenience

Question	Sch	eme	Marks	AO	Notes
1(e)	Possible suggesti (not exhaustive)	ons			
	Offer some free for the survey.	ood if people take			
	Enter survey-taker draw.	rs into a prize			
	Make sure the sur concise.	vey is clear and			
	Tell customers that only take 2 minuted	•			
	Send survey multi	ple times			
			E1, E1	3.1a, 3.1a	E1 for each sensible suggestion (max E2)
				J.1a	Not dep on (d)
		Total	9		

Question	Sche	me	Marks	AO	Notes
2(a)	33%		B1	1.1	awfw 30~35%
2(b)	3	B1	1.1		
2(c)	The nations could	be colour-coded.			
	A key could be add	led			
			E1	1.1	Any sensible suggestion
2(d)	Possible criticisms (not exhaustive)	5			
	There is no key/leg having nation label				
	No consistent label axis (e.g., every oth	0			
	No label on (x) axis	8.			
	No label on (y) axis	8			
	There is no line for (y) axis.	the proportion			
	No (y) axis on the l	eft side.			
	Use of ++* symbol explanation	s with no			
	No source				
	y-axis goes				
			E1, E1, E1, E1	3.1a, 3.1a, 3,1a, 3.1a	E1 for each sensible criticism (max E4)
		Total	7		·

Question	Scheme	Marks	AO	Notes
3(a)	R	B1	1.1	сао
				Cells add to 9
		M1	1.1	or Cells add to 16
				or Cells add to 69

Question	Scheme	Marks	AO	Notes
	R 0 7 60 V	A1	1.1	All three correct
				Cells add to 16
		M1	1.1	or Cells add to 74
				or Cells add to 304

Question	Scheme	Marks	AO	Notes
		A1	1.1	All three correct
	R V 91	B1	1.1	cao
	$ \begin{array}{c c} R & 0 & 0 & 5 \\ \hline 7 & 9 & 60 \\ \hline 228 & 91 \\ \end{array} $			Fully correct diagram [scores full marks]

Question	Scheme	Marks	AO	Notes
	SC R 16 9 74 16 9 69 304 400			Scores max B1M1A0M1A0B0 Remove 1 mark for each small slip
3(b)(i)	$P(R) = \frac{16}{400} = \frac{1}{25} = 0.04$	B1	1.2	Any of these, oe Condone SC $\frac{50}{400} = \frac{1}{8} = 0.125$ or $\frac{50}{897} = 0.0557$
3(b)(ii)	$P(S' \cap V) = \frac{7 + 228}{400}$	M1	1.2	PI P($S' \cap V$) or correct numerator seen Condone SC 16 + 304
	$=\frac{235}{400}=\frac{47}{80}=0.588 \text{ (3sf)}$	A1	1.2	awfw 0.587~0.588 Condone SC $\frac{320}{400} = \frac{4}{5} = 0.8$ or $\frac{320}{897} = 0.357$
3(b)(iii)	$P(R' \cap S' \cap V') = \frac{91}{400} = 0.228 \text{ (3sf)}$	B1	1.2	awfw 0.227~0.228 Condone SC $\frac{400}{897} = 0.446$

Question	Scheme	Marks	AO	Notes
3(b)(iv)	P(R S)			
	$=\frac{9}{74}=0.122$ (3sf)	M1	1.2	PI Numerator or denominator correct or correct use of $P(R S) = \frac{P(R \cap S)}{P(S)}$
		A1	1.2	Fully correct Condone SC $\frac{18}{161} = 0.112$
3(c)	P(R) = 0.04 P(R S) = 0.122	M1	2.1b	PI oe P(<i>R</i>) and P(<i>R</i> <i>S</i>) considered
	$P(R) \neq P(R S)$ so <i>R</i> and <i>S</i> are not statistically independent.	A1	2.1b	Correct comparison and conclusion.
	Alternative			
	$P(S) = \frac{74}{400} = \frac{37}{200} = 0.185$ $P(R \cap S) = \frac{9}{400} = 0.0225$ $P(R) \times P(S) = \frac{37}{5000} = 0.0074$	(M1)		PI Attempt at finding $P(R \cap S)$ and $P(R) \times P(S)$
	$P(R) \times P(S) \neq P(R \cap S)$ so <i>R</i> and <i>S</i> are not statistically independent.	(A1ft)		Correct comparison and conclusion. Dep on correct calculations in M1 ft their (a) and (b)(i)

Question	Scheme		Marks	AO	Notes
3 (d)	Possible reasons	(not exhaustive)			
Larger screens ha		ve more space for			
	More pixels need fit.	a larger screen to			
	More expensive to more likely to hav and higher resolut	ve bigger screens			
			E1	2.1a	Any sensible reason linking higher resolution and larger screen
	1	Total	15		1

Question	Scheme	Marks	AO	Notes
4(a)	$x = \frac{1}{15}$	B1	1.1	or awrt 0.067
4(b)(i)	P(X > 10)			
	= P(10 < X < 15)			
	$= 5 \times \frac{1}{15} = \frac{1}{3}$	B1	1.1	oe or awrt 0.33
4(b)(ii)	$\left(\frac{1}{3}\right)^3$	M1	1.2	PI Their (i) cubed
	$=\frac{1}{27}$	A1	1.2	oe or awrt 0.037
4(c)(i)	[X = Number of Zombies spawning in the next 5 seconds]			
	$[X \sim Po(\lambda)]$			
	$\lambda = 1$	B1	1.2	
	$P(X \ge 1) = 1 - P(X = 0)$	M1	1.2	РІ
	= 0.632 (3sf)	A1	1.2	awrt
	Alternative			
	[X = Time before next zombie spawn (seconds)]			
	$[X \sim \operatorname{Exp}(\lambda)]$			
	$\lambda = \frac{1}{5} = 0.2$	(B1)		oe
	$P(X \le 5) = 1 - e^{-0.2 \times 5}$	(M1)		PI Correct use of exponential formula or 0.368 seen Condone $\lambda = 5$ Condone one small slip
	= 0.632 (3sf)	(A1)		

Question	Scheme	Marks	AO	Notes
4(c)(ii)	$X \sim Po(2)$ $P(X = 0)$	M1	1.2	PI
	= 0.135 (3sf)	A1	1.2	awrt
	Alternative			
	$P(X > 10) = 1 - P(X \le 10)$			
	$= 1 - (1 - e^{-0.2 \times 10})$ $= e^{-0.2 \times 10}$	(M1)		PI
	= 0.135 (3sf)	(A1)		
4(c)(iii)	[Y = Number of zombies spawning in the next minute (60 seconds)]			
	$Y \sim Po(\lambda)$	B1	2.1b	PI Poisson distribution stated or clearly used
	Mean: $\lambda = 60 \times 0.2 = 12$	B1	1.2	PI
$P(Y \le 10)$		M1	1.2	PI Attempt to find this probability
	= 0.347 (3sf)	A1	1.2	awrt
4 (d)	From tables or calculator, use $\lambda = 6$ [per minute]	B1	1.2	
4(e)	Exponential mean $=\frac{1}{6}$ minute	B1	1.2	Or equivalent in seconds
	1	Tatal	15	l

Total 15

Question	Scheme			Ma	arks	AO		Notes			
5(a)	Foreign Born	1	2	2.5]	10	11	15.5	18	24.5	27
	English Literacy	90.5	94.3	91.7	9	8.7	98.3	98	98.9	96.5	96.4
						Ν	/ 11	1.1	correct both va	t one po ly locate riables) ne ±0.1	
						N	А1	1.1	located		ectly
	Printing B	rror							Condoi	ne ±0.1	
	 For wit Eng 0 for to 1 For lite 	reign bo h 19 (le glish lite or foreig =-0.165 reign bo	ads to reeracy ree gn born= 6) rn 18 ha	placed =0.624) corded a =18 (lea as Engli (leads to	as ds sh						
5(b)	<i>r</i> = 0.616	(3sf)				I	41	1.2	awfw 0	0.59~0.6	4
5(c)	Possible r	easons									
	The data is	s not lin	ear.								
	There appe groups.	ear to be	two dif	fferent							
]	E1	3.1a			

Question	Scheme	Marks	AO	Notes
5(d)	The same data, when considered at different levels	E1	3.1b	Mention that the same data is used or Mention of different levels
	is producing both positive and negative values of r .	E1	3.1b	Mention of positive and negative <i>r</i> or Mention of extreme difference in values (oe)

Question	Scheme	Marks	AO	Notes
5(e)	Foreign-born people moving to the USA are likely to move to regions with lots of urban areas (where there are more jobs).			Linking foreign-born with a third variable Examples Urbanness Availability of jobs Money/prosperity Better schools Nicer area Others of their nationality in area
	People in urban areas are generally better educated than those in rural areas.			Linking education (or literacy skills) with same variable or Linking education (or literacy skills) with one of above variables if first E1 not awarded
	So, we would expect regions with lots of foreign-born population to have a high literacy.			Association between foreign-born and English literacy due to a common third variable
	However, at an individual level, people who are foreign-born are more likely to have another language as their first language, so their English literacy is likely to be poorer.			Implication that English may not be first language
	An individual can't have a foreign born %, they are either foreign born or not			
	The grouped data is likely to have fewer outliers			Or anomalies
	Looking at individuals would have a larger sample size with more data points			
	9 regions might not cover the whole of the USA while individual data does			

Question	Sch	eme	Marks	AO	Notes
	Literacy rate of A population in an a figures				
	May not have data for all adults in a region				
			E1, E1, E1, E1	2.1a, 2.1a, 2.1a, 2.1a, 2.1a	One mark per comment Not exhaustive
		Total	10		

Question	Scheme	Marks	AO	Notes				
6(a)								
	Type of asbestos							
	$\begin{array}{c} 0.015 \\ W \end{array} (0.972 \times 0.015 = 0.014) \\ \end{array}$							
	0.972 0.98	5 R	,					
	0.018 B	71 R	(0.018	$3 \times 0.971 = 0.0175)$				
	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$							
	0.847 <i>R</i> ′							
		A1	1.2	Any correct pair of conditional values given (0.015/0.985 after 0.972, 0.971/0.029 after 0.018 or 0.153/0.847 after 0.010)				
		A1	1.2	Completely correct				
6(b)	$0.972 \times 0.015 + 0.018 \times 0.971$			Or 0.0146 + 0.0175 + 0.00153				
	$0.972 \times 0.015 \pm 0.018 \times 0.971 + 0.01 \times 0.153$	M1		May be seen on diagram				
	=0.033588	A1		Allow one small slip awrt 0.0336				

Question	Sch	eme	Marks	AO	Notes
6(c)	$P(B R) = \frac{P(B \cap R)}{P(R)}$	B1	1.2		
	=	0.018×0.971			
	0.972×0.015	+ 0.018 × 0.971 +	-0.01×0	.153	I
			B1	1.2	Numerator correct or awrt 0.0175 seen anywhere (including on tree diagram)
			B1	1.2	Denominator correct or is 'their (b)'
	= 0.520 (3sf)		A1	1.2	awfw 0.52~0.55
6(d)	Possible reasons	(not exhaustive)			
	Blue asbestos may remove (so it is re- often).				
	Blue asbestos may dangerous (so it is often).				
			E1	2.1a	Any sensible reason
6(e)	No, I disagree wit	h Riddhi's	E1dep	3.1b	dep on good attempt at explanation ft their (a)
	as about 52% of removals are of bl	E1ft	3.1b	Reference to 0.520 ft their (a)	
	so lots of safety equipment for blue asbestos will be required.		E1	3.1b	Relating back to safety equipment correctly.
		Total	12		

Question	Scheme	Marks	AO	Notes	
7(a)	$\mu \approx 0.971$	B1	2.1b		
	The normal distribution is symmetrical, so the mean equals the median.	E1	2.1b		
7(b)	Approximately $\frac{2}{3}$ of data lies within $\mu \pm \sigma$	E1	2.1b	oe	
	$\dots \text{so } \mu + \sigma = S_5$	M1	2.1b	PI oe for example $\mu - \sigma = S_1$	
	$\sigma \approx 1.000 - 0.971$ $= 0.029$	A1*dep	2.1b	or 0.971 – 0.942 dep on E1	
	SC 1.000 - 0.971 = 0.029 or $0.971 - 0.942 = 0.029$ seen with no explanation scores E0M1A0				
7(c)(i)	$[M = 2D:4D \text{ ratio of a randomly} \\ \text{selected male adult}] \\ [M \sim N(0.952, 0.034^2)]$				
	P(M > 1) = 0.0790 (3sf)	B1	1.2	awrt 0.079	
7(c)(ii)	P(0.9 < M < 1) = 0.858 (3sf)	B1	1.2	awrt 0.86	

Question	Sch	eme	Marks	AO	Notes
7(d)	[F = 2D:4D ratio] selected female ac				
	[F~N(0.971,0.02	(9 ²)]			
	[<i>D</i> = Difference b ratio of a randoml and female adult]				
	D=F-M		M1	2.1b	PI or D = M - F
			M1	2.1b	PI Normal distribution used for difference
	<i>D</i> ~N(0.019,0.0447 ²)			1.2	PI $\mu = \pm 0.019$ cao PI
			B1	1.2	$\sigma = 0.0447 \text{ (awrt)}$ or $\sigma^2 = 0.00200 \text{ (awrt)}$
	P(D > 0) = 0.66	5 (3sf)	A1	1.2	awfw 0.664~0.665
L		Total	12		

Pearson Education Limited. Registered company number 872828 with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom