



Mark Scheme

Summer 2025

Pearson Edexcel GCE
In Statistics (9ST0)
Paper 02: Statistical Inference

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

June 2025

Question Paper Log Number 76382

Publications Code 9ST0_01_2506_MS

All the material in this publication is copyright

© Pearson Education Ltd 2025

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)	2.0 represent a 100% increase which is equivalent to doubling	E1	1.2	Or equivalent statement
1(b)	$H_0: \mu = 2.0$ $H_1: \mu \neq 2.0$ $s = \sqrt{\frac{1}{31}(6.5596)}$ $s = 0.46$ (or $s^2 = 0.2116$) $ts = \frac{\frac{70.05}{32} - 2.0}{\frac{0.46}{\sqrt{32}}}$ 2.32 $cv = (\pm) 1.9600$ 2.32 > 1.96 Reject H_0 There is significant evidence to suggest that the population mean likelihood of pothole damage differed from 2.0 during the period 2014 to 2021.	B1 M1 A1 M1 A1 B1 M1 E1dep	1.3 1.2 1.2 1.3 1.3 1.3 2.1b 2.1a	oe Attempt to find either the standard deviation or variance cao Attempt to find test statistic awfw 2.30~2.37 Comparison of their ts and their cv or 0.010 < 0.025 Conclusion in context dep correct cv and ts

Alternative				
1(b)	$H_0: \mu = 2.0$	(B1)		oe
	$H_1: \mu \neq 2.0$			
	$s = \sqrt{\frac{1}{31}(6.5596)}$	(M1)		Attempt to find either the standard deviation or variance
	$s = 0.46$ (or $s^2 = 0.2116$)	(A1)		cao
	standard error = $\frac{0.46}{\sqrt{32}}$	(M1)		Attempt to find test statistic Accept $\frac{0.2116}{32}$
	$cv = (1.84)$ and 2.16	(A1)		cao
	$ts = 2.19$	(B1)		
	$2.19 > 2.16$	(M1)		Comparison of their ts and their cv (does not require the lower cv)
	Reject H_0			
	There is significant evidence to suggest that the population mean likelihood of pothole damage differs from 2.0	(E1dep)		Conclusion in context dep correct cv and ts
		Total	9	

Question	Scheme	Marks	AO	Notes																																												
2(a)	Wilcoxon Rank Sum	M1	2.1a	Used or clearly stated or Mann-Whitney (U test) stated																																												
	H ₀ : $\eta_{18} = \eta_{23}$ H ₁ : $\eta_{18} > \eta_{23}$	B1	1.3	H ₀ : Samples from identical populations H ₁ : Samples from different populations oe																																												
	<table><tr><th colspan="2">2018 Number of Aces</th><th colspan="2">2023 Number of Aces</th></tr><tr><td>2</td><td>12</td><td>10</td><td>1</td></tr><tr><td>3</td><td>14</td><td>19</td><td>4</td></tr><tr><td>6.5</td><td>27</td><td>20</td><td>5</td></tr><tr><td>10</td><td>37</td><td>27</td><td>6.5</td></tr><tr><td>11</td><td>40</td><td>32</td><td>8</td></tr><tr><td>14</td><td>44</td><td>36</td><td>9</td></tr><tr><td>16</td><td>47</td><td>42</td><td>12</td></tr><tr><td>18</td><td>55</td><td>43</td><td>13</td></tr><tr><td>19</td><td>57</td><td>45</td><td>15</td></tr><tr><td>20</td><td>76</td><td>50</td><td>17</td></tr></table>				2018 Number of Aces		2023 Number of Aces		2	12	10	1	3	14	19	4	6.5	27	20	5	10	37	27	6.5	11	40	32	8	14	44	36	9	16	47	42	12	18	55	43	13	19	57	45	15	20	76	50	17
	2018 Number of Aces		2023 Number of Aces																																													
	2	12	10	1																																												
	3	14	19	4																																												
	6.5	27	20	5																																												
	10	37	27	6.5																																												
	11	40	32	8																																												
	14	44	36	9																																												
16	47	42	12																																													
18	55	43	13																																													
19	57	45	15																																													
20	76	50	17																																													
	M1	1.3	Attempt at ranking as one group																																													
	A1	1.3	All ranking correct Or reversed ranks																																													
	T ₁₈ = 119.5 or T ₂₃ = 90.5	M1dep	1.3	Effort to obtain T																																												

	$U_{18} = 119.5 - \frac{1}{2}(10)(11) = 64.5$ $U_{23} = 90.5 - \frac{1}{2}(10)(11) = 35.5$ $cv = 28$ (or 70) $35.5 > 28$ (or $64.5 < 70$) Do not reject H_0 There is insufficient evidence at the 5% level that the number of aces has begun to decrease over recent years.	A1 B1 M1 E1dep	1.3 1.3 2.1b 2.1a	Either U correct Either cv correct Comparison of their ts and their cv Conclusion in context dep correct cv and ts
2(b)	Any two from: Only used the top 10 players for each tournament. Only used male players within the sample. Only used one tournament/type of court/location Only used professional players Only used two specific years	E1, E1	3.1a, 3.1a	Any reference to use of one location Identifying a variable that may affect a specific year e.g. weather/illness/injury Any reasonable comment. Comments must be made in context

2(c)	There may have been a faster serve before 1989 but would not have been recorded. Advancements in the technology may mean that earlier serves were not recorded correctly/accurately The technology is used during major tournaments and is not widely used in the amateur sector so there may be an amateur who served a faster serve but wasn't recorded.	E1, E1	3.1a, 3.1a	or comment about errors in the modern technology or comment with regards to variables affecting speed of serve e.g. wind resistance Any relevant comment Comments must be made in context
Total		13		

Question	Scheme	Marks	AO	Notes
3(a)	$H_0: p = 0.564$ $H_1: p < 0.564$	B1	1.3	Must have % symbol if using 56.4
	$X \sim B(120, 0.564)$	B1	1.3	PI Do not accept 56.4% here
	$P(X \leq 58) = 0.0460$ $P(X \leq 59) = 0.0665$	M1	1.3	Use of their binomial distribution to find to values both sides of the significance level using Binomial CD
	$cr: X \leq 58$	A1	1.3	cao
	$57 < 58$ Reject H_0	M1	2.1b	Comparison of their ts and their cv
	There is significant evidence to suggest that the proportion of adults in the UK exhibiting sedentary behaviour has decreased between 2013 and 2017.	E1dep	2.1a	Conclusion in context dep correct cv and ts
3(b)	$P(\text{Type I error}) = 4.60\%$	B1	1.2	awrt 4.6 oe
	Accepting that there is evidence to suggest the proportion of adults exhibiting sedentary behaviour has decreased when in fact it has not.	E1	3.1a	Must have context and reference to lower tail test
3(c)	The power of the hypothesis test would increase .	E1	3.1a	Context not required
Total		9		

Question	Scheme	Marks	AO	Notes
4(a)	$H_0: \mu_F - \mu_M = 0$ $H_1: \mu_F - \mu_M > 0$	B1	1.3	oe both subscripts clearly identified
	$s_p^2 = \frac{(9-1) \times 348 + (11-1) \times 387}{9+11-2}$	M1	1.3	PI Do not award if standard deviation used
	$s_p^2 = 370$	A1	1.3	PI awfw 368~372
	$\bar{f} = 54.8$ $\bar{m} = 33.8$	B1	1.2	awrt 55 and 34
	$ts = \frac{(54.8 - 33.8) - 0}{\sqrt{370 \left(\frac{1}{9} + \frac{1}{11} \right)}}$	M1	1.3	Attempt at test statistic
	2.43	A1	1.3	Correct ts
	$cv = 1.734$	B1	1.3	
	$2.43 > 1.734$ Reject H_0	M1	2.1b	Comparison of their ts and their cv or $0.0128 < 0.05$
	There is significant evidence to suggest that the population mean time taken to back a bag for holiday is greater for the company's female customers than the male customers	E1 dep	2.1a	Conclusion in context dep correct cv and ts

Alternative				
4(a)	$H_0: \mu_F - \mu_M = 0$	(B1)		oe both
	$H_1: \mu_F - \mu_M > 0$			subscripts clearly identified
	$s_p^2 = \frac{(9-1) \times 348 + (11-1) \times 387}{9+11-2}$	(M1)		PI
	$s_p^2 = 370$	(A1)		Do not award if standard deviation used
	$\bar{f} = 54.8$	(B1)		PI
	$\bar{m} = 33.8$			awfw 368~372
	$cv = 0 + 1.734 \times \sqrt{370 \left(\frac{1}{9} + \frac{1}{11} \right)}$	(M1)		awrt 55 and 34
	15.0	(A1)		Attempt at critical value
	$ts = 21.0$	(B1)		Correct cv
	$21.0 > 15.0$	(M1)		awrt 21.0
	Reject H_0			Comparison of their ts and their cv
	There is significant evidence to suggest that the population mean time taken to back a bag for holiday is greater for the company's female customers than the male customers	(E1 dep)		Conclusion in context dep correct cv and ts

4(b)	The population distribution time to pack a bag follows a normal distribution for females and males	E1	3.1b	Must have context
	The population variance for time to pack a bag has a common variance for females and males	E1	3.1b	Context not required Accept standard deviation
4(c)	The results are only for one flight and therefore cannot be considered to be random The customers are only for a long-haul flight so conclusion cannot be generalised for all types of flight People may not be able to recall the specific time spent packing their bags The customers on a flight might be in couples/families/groups so the samples might not be independent	E1, E1, E1	3.1a, 3.1a, 3.1a	Comment about only one airline sampled or people may lie Any reasonable comment. Comments must be made in context
Total		14		

Question	Scheme	Marks	AO	Notes
5(a)	Total probability of 1 for the whole day and 8 working hours within which he can be on walkabout this would make a probability for a particular hour to be $\frac{1}{8}$	B1	1.2	Comment that the total probability would be 1
	If the timings are to be unpredictable, they must be equally likely to occur across the working day	E1	1.2	Comment on equally distributed implying uniform
5(b)	H_0 : That a uniform distribution is a good fit	B1	1.3	oe
	H_1 : That a uniform distribution is not a good fit			
	Expected values = 75	M1	1.3	PI
	$\frac{(89-75)^2}{75} + \dots + \frac{(65-75)^2}{75}$	M1	1.3	Method for contributions correct
	= 8.24	A1	1.3	Give mark for one correct value seen
	$cv = 14.067$	B1	1.3	awfw 7.9 ~ 8.5
	8.24 < 14.067 Do not reject H_0	M1	2.1b	Comparison of their ts and their cv or 0.312 > 0.05
	There is insufficient evidence to suggest that a uniform distribution is not a good fit for the times of Matt's walkabouts	E1dep	2.1a	Conclusion in context dep correct cv and ts

5(c)	H_0 : Andy's and Matt's decisions agree	B1	1.3	oe Allow $p = 0.5$ or $\eta_d = 0$
	H_1 : Andy's and Matt's decisions differ			
	$X \sim B(30, 0.5)$	B1	1.3	PI
	$P(X \leq 8) = 0.00806$	B1	1.2	Uses their binomial and T^+/T^- to find a probability
	$0.00806 > 0.005$ Do not reject H_0	M1	2.1b	Comparison of their p -value and 0.005
	There is insufficient evidence to suggest that there is a difference in Andy's and Matt's judgements	E1dep	2.1a	Conclusion in context dep on correct p -value
5(d)	From Matt's recorded start times it was shown that his walkabouts were unpredictable for the workers	E1ft	2.1b	Comment with regards to their conclusion in (b) in context
	There is no significant disagreement between Matt and Andy's quality assurance tests	E1ft	2.1b	Comment with regards to their conclusion in (c) in context
	Target 1 and 2 were both met	B1ft	2.1b	Clear statement that both targets have been met
	Since Matt has met both targets, he should be allowed to maintain his promotion	B1ft	2.1b	Comment about maintaining his promotion based on at least 1 of the targets
		E1	2.1b	Style appropriate to general public. No difficult statistical vocabulary. No oversimplification.
Total		19		

Question	Scheme	Marks	AO	Notes																				
6(a)	$H_0: \mu_H = \mu_T = \mu_M$ H_1 : At least two of the means differ	B1	1.3	oe or $\mu = \mu_i$ for $i = H, T, M$																				
	$SS_T = 6450.32 - \frac{272.6^2}{12}$ $= 257.76$	M1	1.3	SS_T method PI																				
	$SS_R =$ $\frac{75.9^2}{3} + \frac{74.2^2}{3} + \frac{60.0^2}{3} + \frac{62.5^2}{3} - \frac{272.6^2}{12}$ $= 65.003$	M1	1.3	SS_R method PI																				
	$SS_C =$ $\frac{95.0^2}{4} + \frac{70.9^2}{4} + \frac{106.7^2}{4} - \frac{272.6^2}{12}$ $= 166.61$	M1	1.3	SS_C method PI																				
	<table><tr><td></td><td>SS</td><td>Dof</td><td>MS</td></tr><tr><td>Between Colours</td><td>65.003</td><td>3</td><td>21.67</td></tr><tr><td>Between Font style</td><td>166.61</td><td>2</td><td>83.31</td></tr><tr><td>Error</td><td>26.142</td><td>6</td><td>4.357</td></tr><tr><td>Total</td><td>257.76</td><td>11</td><td></td></tr></table>		SS	Dof	MS	Between Colours	65.003	3	21.67	Between Font style	166.61	2	83.31	Error	26.142	6	4.357	Total	257.76	11		M1	1.3	SS_E method No negative SS values.
		SS	Dof	MS																				
	Between Colours	65.003	3	21.67																				
	Between Font style	166.61	2	83.31																				
	Error	26.142	6	4.357																				
	Total	257.76	11																					
	B1	1.3	PI Dof given																					
	M1	1.3	PI MS divide SS by Dof Must attempt to calculate all 3 MS values																					
$F = \frac{83.31}{4.357} = 19.1$	A1ft	1.3	PI F method Dep on all previous M marks but ft on Dof values																					
			or p -value = 0.00250																					
$cv F_6^2 = 5.143$	B1	1.3	Obtaining critical value																					

	$19.1 > 5.143$ Reject H_0 There is significant evidence to suggest that mean perceived expense differs for the font styles	M1 E1dep	2.1b 2.1a	Comparison of their ts and their cv or $0.00250 < 0.05$ Conclusion in context dep correct cv and ts
6(b)	$\bar{h} = 23.75$ $\bar{t} = 17.725$ $\bar{m} = 26.675$ The restaurant should use a traditional font as this was the menu which had the lowest mean and therefore appeared to be the least expensive	B1 E1	1.2 2.1a	Must be in context Could also discuss that modern and handwritten fonts had higher means Must refer to means in their response to obtain this mark can be implied by calculation of mean
6(c)	$\frac{21.67}{4.357} = 4.974$ $4.974 > 4.757$ The use of colour of the menu as a blocking factor was effective	B1ft M1 E1dep	1.3 1.3 3.1a	awrt 5.0 or $p\text{-value} = 0.0459$ ft on Dof values Comparison of their F for between colours and their F_6^3 ft on Dof values or $0.0459 < 0.05$ Dependent on previous 2 marks and the correct comparison for their values
Total		16		