Mark Scheme (Results)

Summer 2023

Pearson Edexcel GCE
In Statistics (9ST0)
Paper 03: Statistics In Practice

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

Summer 2023
Publications Code 9ST0_03_2306_MS

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 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)	p=0.2	B1	1.1	
1(b)	$1 \times 0.4 + 2 \times 0.2 + 3 \times 0.2 + 4 \times 0.2 = 2.2$	B1*	1.1	AG
1(c)	$var(X) = 1.36$	B1	1.1	
1(d)	$\frac{0.4}{0.2 + 0.4}$	M1	1.2	
	$\frac{2}{3}$	A1	1.2	
1(e)	(Regardless of the distribution of X,) when n is large…	E1	1.3	n≥ 30
	the distribution of $\bar{X}$ is approximately normal…	E1	1.3	
	with mean equal to mean of X and variance = $\frac{\text{var}(X)}{n}$	E1	1.3	
1(f)	$\bar{X} \sim N\left(2.2, \frac{1.36}{60}\right)$	B1	1.2	Seen or used
	$P(\bar{X} > 2.25) = 0.3699$	A1	1.2	awrt 0.37
	Total	**10**		

Question	Scheme	Marks	AO	Notes
2(a)	**Possible comments (not exhaustive)**			
	Elias is just looking at the difference of the percentage points.			
	Brett expresses the difference in the percentage points as a percentage.			
	Both the calculations are correct.			
	Brett's statement may be misleading.			**or** Elias's statement is easier to understand.
		E1, E1	3.1a, 3.1a	E1 for each sensible comment
2(b)	Most likely answers: • Denmark • Germany • Luxembourg • Netherlands • Sweden • Norway • Switzerland	B1	2.1a	**Accept** other countries above EU average in both categories • Belgium • Estonia • Ireland • France • Finland • Iceland
	This country would be useful to advertise in, as a high percentage of the population use the internet and have made a recent online purchase.	E1	2.1b	SC Malta shows all people using internet have purchased online B1E1

2(c)	**Possible explanations (not exhaustive)**			
	The population of a chosen country could be very small, with not many customers to advertise to.			
	Advertising costs could vary greatly by country.			
	It may not be practical to ship his products to all countries.			
	His products may not be popular in every country.			
	Pandemic might change habits			
		E1, E1, E1	3.1a, 3.1a, 3.1a	E1 for each sensible explanation (max E3)
2(d)	**Possible comments (not exhaustive)**			
	The percentage of the population who have used the internet in the last 12 months…			
	Age category			
	…is roughly the same for males and females.			**or** slightly higher for males.
	…is highest in the 25-34 age category.			**or** lowest in the 65-74 category.
	…decreases with age after the 25-34 age category.			

	Education category			
	…is highest in the 'High education' category.			**or** lowest in the 'Low education' category.
	…increases with level of education.			
	Employment category			
	…is highest for employed people [and students].			**or** lowest for retired people.
		E1, E1, E1	1.1, 1.1, 1.1	E1 for each sensible explanation (max E3)
		Total	**10**	

Question	Scheme	Marks	AO	Notes
3(a)	Two-factor…	B1	1.3	
	ANOVA	B1	1.3	
3(b)	**Possible assumptions**			
	Lifespans are normally distributed The distributions from which the samples are taken have equal variances There is no interaction between the factors (no reason to believe any particular material works any better or worse at any particular temperature.	B1, B1	3.1a, 3.1a	Context needed Max 2 marks
3(c)	H_0: $\mu_A = \mu_B = \mu_C = \mu_D$ H_1: At least two means differ	B1	2.1a	oe
	$SS_T = \Sigma\Sigma\, x_{ij}^2 - \frac{T^2}{mn}$ $= 831.7 - \frac{88^2}{12}$ $= 186.36$ ...	M1	1.3	PI SS Total awrt 186.4
	$SS_R = \Sigma\, \frac{R_i^2}{n_i} - \frac{T^2}{mn}$ $SS_R = \frac{19.6^2}{3} + \frac{18.7^2}{3} + \frac{25.6^2}{3} + \frac{24.1^2}{3} - \frac{88^2}{12}$ $= 11.34$	M1	1.3	PI SS between materials
	$SS_C = \Sigma\, \frac{C_i^2}{m_i} - \frac{T^2}{mn}$ $SS_C = \frac{47^2}{4} + \frac{31^2}{4} + \frac{10^2}{4} - \frac{88^2}{12}$ $= 172.16$ ...	M1	1.3	PI SS between temperatures

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td></td><td>SS_E = 186.37-11.34-172.17
=2.86</td><td>M1dep</td><td>1.3</td><td>PI
Dep previous two Ms (not negative)</td></tr>
<tr><td></td><td><table>
<tr><th>Source</th><th>Sums of Squares</th><th>Degrees of freedom</th><th>Mean square</th><th>F ratio</th></tr>
<tr><td>Rows</td><td>11.34</td><td>3</td><td>3.78</td><td>7.93</td></tr>
<tr><td>Columns</td><td>172.17</td><td>2</td><td>86.083</td><td>180.59</td></tr>
<tr><td>Error</td><td>2.86</td><td>6</td><td>0.4767</td><td></td></tr>
<tr><td>Total</td><td>186.37</td><td>11</td><td></td><td></td></tr>
</table></td><td></td><td></td><td></td></tr>
<tr><td></td><td></td><td>B1</td><td>1.3</td><td>df correct, 3 and 6</td></tr>
<tr><td></td><td></td><td>M1</td><td>1.3</td><td>PI
MS=SS/df for between materials and error</td></tr>
<tr><td></td><td>$F = \frac{3.78}{0.4767} = 7.93$</td><td>A1</td><td>1.3</td><td>awrt 7.93
or
p = 0.0165
awfw 0.016 ~ 0.017</td></tr>
<tr><td></td><td></td><td></td><td></td><td>Correct table implies all previous marks</td></tr>
<tr><td></td><td>Critical value $F_{3,6}(0.05) = 4.757$</td><td>B1</td><td>1.3</td><td></td></tr>
<tr><td></td><td>7.93 > cv so reject H_0

There is significant evidence to suggest at least two average battery lifetimes differ for the four materials used.</td><td>M1

E1 dep</td><td>2.1b

2.1a</td><td>OR
p=0.0165<0.05
PLUS correct conclusion

Conclusion in context
Condone “difference between average lifetimes”

Dep on correct cv and ts (p-value and sig)</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
3(d)	It is recommended material C or D is used	B1	2.1a	Condone just material C or material D recommended SC if Ho not rejected then "it doesn't matter so pick the cheapest"
3(e)	Identification of temperatures as the blocking factor	B1	1.3	PI
	F test for difference between temperatures gives *F* =180.6 which is highly significant (5% cv = 5.143)	M1	1.3	For attempting *F* test between temperatures
		A1	1.3	awrt 181 Condone for M1 MS between temperatures (86.083) is the largest (ie temperatures are the largest source of variation in lifespans) **or** uses only the totals/means for each temperature **or** p=4.36×10^{-6}
	So there seems to be a difference between at least two average lifespans for the three **temperatures** **or** Blocking factor was effective	E1 dep	3.1a	Dep some analysis completed
	Total	**20**		

Question	Scheme	Marks	AO	Notes
4(a)(i)	Binomial distribution	B1	2.1b	
	(The number of days fixed, and) the probability of snow is likely to be similar in all days of January	E1	2.1b	
4(a)(ii)	Poisson distribution	B1	2.1b	
	The visitors could be reasonably expected to come with a constant average rate	E1	2.1b	
4(a)(iii)	Exponential distribution	B1	2.1b	
	Visits to a website are likely to occur at a constant average rate and therefore follow a Poisson distribution	E1	2.1b	
	The time between consecutive Poisson events follows the exponential distribution	E1	2.1b	

Question	Scheme	Marks	AO	Notes
4(b)	**Binomial model** Unlikely to be independent if it snows on one day to the next as it may be more likely to snow in periods of cold weather **Poisson model** If two people are involved in an accident together they will not attend independently **Exponential model** Visitors are unlikely to visit at a constant rate across the whole day If something goes viral on the website clicks are unlikely to be independent Any sensible reason attached to model	E1dep	3.1a	List not exhaustive **Dep on student using correct model in (a)**
	Total	**8**		

Question	Scheme	Marks	AO	Notes
5(a)	As participants who have a slower individual time will probably have a slower average time	E1	2.1a	
5(b)	-0.374 + (1.564 × 9)	M1	1.2	PI
	$y = 13.70$ seconds	A1	1.2	
5(c)	It is impossible to have an individual time of 0 seconds. It is impossible to have a negative average time.	E1	2.1a	Either
5(d)	When a competitor's individual time increases by 1 second their average time increases by 1.564 seconds.	E1	2.1a	
5(e)	**Query** the database to produce a report	E1	1.1	
	Join the tables together	E1	1.1	
	Using the **unique ID**	E1	1.1	

It is possible to see a SQL query here, e.g.

SELECT firstName, surname, uniqueID, bestTime FROM competitors, times WHERE competitors.uniqueID = times.uniqueID

Or a similar version with INNER JOIN

Sight of a SQL statement should be awarded the first mark.

A sensible INNER JOIN or WHERE statement would award the second and third if unique ID field was mentioned.

Total	**8**

Question	Scheme	Marks	AO	Notes
6(a)	($\bar{x}$ = 3.25 s = 1.2 n = 15)			PI
	$3.25 \pm 2.1448 \times \frac{1.2}{\sqrt{15}}$	M1	1.3	Use of $\frac{1.2}{\sqrt{15}}$
		B1	1.3	PI t_{14} = 2.14(5) or z=1.96 used awrt SC (2.64, 3.86) if z=1.96 used scores M1B1A0
	CI=(2.59, 3.91)	A1	1.3	SC 15 df used t = 2.131 (2.58, 3.91) M1B1A0
6(b)	Number of penalty corners awarded per game is normally distributed.	E1	2.1a	Context required
6(c)	2.976 is within the CI…	B1ft	2.1b	oe comparison of 2.976 with their CI
	…so there is not enough evidence to suggest Douglas' team were awarded a different number of penalties per game, on average, than those awarded at the 1998 Field Hockey World Cup.	E1dep	2.1a	oe Dep on previous B1
6(d)	$\frac{114}{114+136} = 0.456$	B1	2.1a	AG
6(e)	Because $P(X \leq 10) = 0.03024$ and $0.03024 < 0.05$	E1	1.3	
	But $P(X \leq 11) = 0.06361$ and is $0.06361 > 0.05$	E1	1.3	
6(f)	H_0: $\pi = 0.456$ H_1: $\pi < 0.456$			Not needed, given before (d) but may be seen

Question	Scheme	Marks	AO	Notes
	$7< 10$, Reject H_0	M1	2.1b	Or $0.00149 < 0.05$
	Evidence to support Douglas's belief (his team score fewer goals from penalty corners than expected)	E1dep	2.1a	Dependent on M1
6(g)	You can test the same hypotheses again without further calculation	E1	1.3	
6(h)	H_0 is rejected in error so…	E1	1.3	Explanation mark Implied by correct contextual answer
	…concluding, for his team, penalty corners resulted in fewer goals than expected when actually they result in at least as many as expected.	E1	2.1a	Context mark
6(i)	0.03024	A1ft	1.3	ft their (e) awrt 0.03
6(j)	X~B(35,0.15)	M1	2.1a	
	P(X>10)=0.01098	A1	1.2	SC P(X≥10)=0.0292 scores M1A1A0
	Power = 1-0.01098=0.98902	A1	1.2	awrt 0.99
6(k)	**Constant probability** is unlikely to be valid because… …different players may have different chance of scoring and it's	E1	3.1a	Constant probability

Question	Scheme	Marks	AO	Notes
	unlikely the same player is always involved. …different teams may be better at defending	E1	2.1a	Any sensible reason
	Independence is unlikely to be valid because… …decisions of tactics for taking a penalty corner may be impacted by the previous penalty corners. …confidence and form may be impacted by scoring/missing previous penalty corner.	E1	3.1a	Independence
		E1	2.1a	Any sensible reason
6(l)	Douglas's team might not have specialist corner takers. **or** Data from the 1998 Field Hockey World Cup predates the change in rule but Douglas's team's data is after the rule change.	E1	3.1b	oe
	So the conclusion may not be valid because Douglas's team's record of scoring from penalties is not being compared with the same circumstances (in terms of specialist players being able to try to score goals from penalties).	E1	3.1b	
	Total	**24**		

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