Mark Scheme (Results)

Summer 2019

Pearson Edexcel GCE In Statistics

Paper 8ST0_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

Summer 2019
Publications Code 8ST0_01_1906_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.

EDEXCEL GCE Statistics

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 $\surd$ 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)
- dep – dependent
- indep – independent
- dp decimal places
- sf significant figures
- * The answer is printed on the paper
- ⌐ 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.

Question	Scheme	Marks	AO	Notes
1(a)(i)	The area under the graph represents total probability = 1 base of rectangle is 18 – 8 = 10 so height of rectangle is 1 ÷ 10 = 0.1	M1 A1	1.1 1.1	Meaning of area under graph PI OE
1(a)(ii)	(The distribution is) symmetrical	E1	1.2	
1(b)	0	B1	1.2	OE Zero
1(c)	$(18 - 15) \times 0.1 = 0.3$	B1	1.1	OE correct area calculation
1(d)	P(After 3pm $\cap$ before 5:30pm) $= 2.5 \times 0.1 = 0.25$ P(Before 5:30pm | after 3pm) $= 0.25 \div 0.3$ $= \frac{5}{6} = 0.833$	M1 A1	1.1 1.1	Attempt at area of intersection or division by 0.3 cao Either
	Alternative 5 of the 6 half-hour periods between 3pm and 6pm will be OK. $= \frac{5}{6} = 0.833$	(M1) (A1)		PI OE method or explanation. cao Either
	Total	**7**		

Question	Scheme	Marks	AO	Notes
2(a)	Wilcoxon rank-sum test	B1	2.1a	Stated Condone Mann-Whitney (U-test)
	H_0: Samples come from identical populations H_1: Samples do not come from identical populations	B1	1.3	OE both or using population medians 1-tailed
	Standard advice rank: 1, 3, 4.5, 6, 7.5, 10.5, 13, 13, 16 18, 16, 14.5, 13, 11.5, 8.5, 6, 6, 3 **New therapy rank**: 2, 4.5, 7.5, 9, 10.5, 13, 15, 17, 18 17, 14.5, 11.5, 10, 8.5, 6, 4, 2, 1			
		M1	1.3	attempt at ranking (ignore tied ranks)
		A1	1.3	all ranking correct (including ties)
	T = 74.5 T = 96.5	M1dep	1.3	Effort to obtain T
	U = 74.5 – ½ (9)(10) = 29.5 U = 96.5 – ½ (9)(10) = 51.5	A1	1.3	Either U correct
	(5% 1-tailed) c.v. = 21 (or 60)	B1	1.3	Either cv correct
	29.5 > 21 or 60 > 51.5 Accept H_0 **or** There is no sig. evidence (that the new therapy led to a greater improvement in average memory score, in the population)	A1dep	2.1b	Correct comparison in same tail and correct conclusion. Dep on **both** correct ts and cv. Context **not** required.
				SC Paired Wilcoxon Max B1 hypotheses, M1 differences

Question	Scheme	Marks	AO	Notes
2(a) continued				SC Paired Sign Max B1 hypotheses, M1 signs
2(b)	The new therapy did result in a higher average/rank sum/U value for memory improvement score than the standard GP advice. However, this result was not statistically significant at the 5% level. Nevertheless, the results do indicate that a larger/longer study might demonstrate significant beneficial effects of the new therapy.	E1 E1 E1 E1dep	2.1a 2.1b 3.1a 2.1b	Consider data and compare groups. Result of hypothesis test. Larger/longer study. Clear statements, suitable for target audience. Dependent on at least one E statement above correct, context used & formal style.
	Total	**12**		

Question	Scheme	Marks	AO	Notes
3	Number of purple anemones X ~ B(3000, 0.68)			
	Approximate to Y ~ N(2040, 652.8)	M1	2.1b	Normal approx. stated or clearly used
		A1	1.2	Mean = 2040 (condone 960)
		A1	1.2	Variance = awrt 653 or SD = awrt 25.5~25.6
	P(X < 2000) ≈ P(Y < 1999.5) P(X≤1999)	M1	1.2	PI Use of 1999.5 (or 2000) if mean=2040, 1000 or 1000.5 if mean=960 z =awfw -1.56 ~ -1.59
	= 0.0565	A1	1.2	awfw 0.056 ~ 0.059 Actual: 0.05646837… **Binomial prob**: awrt 0.0569 Actual: 0.05686… scores up to M0A0A0M1A1
	Total	**5**		

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>4(a)</td><td>She must pool/merge/collate colour classes

because the expected frequencies for these classes are all < 5</td><td>B1

E1</td><td>3.1a

3.1a</td><td>Pool/merge/collate some colour classes

classes with E < 5</td></tr>
<tr><td rowspan="5">4(b)</td><td>H_0: There is no association (between the sex of a cat and its colour)
H_1: There is an association (between the sex of a cat and its colour)</td><td>B1</td><td>1.3</td><td>both hypotheses required</td></tr>
<tr><td colspan="4">O, E pooled
<table><tr><th></th><th>Black and white</th><th>Black</th><th>Tabby</th><th>Ginger</th><th>Other</th></tr><tr><td>Female</td><td>26, 28.88</td><td>21, 16.99</td><td>17, 15.85</td><td>4, 8.49</td><td>9, 6.78</td></tr><tr><td>Male</td><td>25, 22.13</td><td>9 ,13.01</td><td>11, 12.15</td><td>11, 6.51</td><td>3, 5.22</td></tr></table>(Some other pooling methods allowable provided E>5.)</td></tr>
<tr><td>Pooling attempted</td><td>M1

A1</td><td>1.2

1.2</td><td>condone under- & over-pooling

All pooling correct</td></tr>
<tr><td colspan="4">Contrib.
<table><tr><th></th><th>Black and white</th><th>Black</th><th>Tabby</th><th>Ginger</th><th>Other</th></tr><tr><td>Female</td><td>0.29</td><td>0.95</td><td>0.08</td><td>2.37~2.38</td><td>0.72~0.73</td></tr><tr><td>Male</td><td>0.37</td><td>1.24</td><td>0.11</td><td>3.10</td><td>0.93~0.94</td></tr></table></td></tr>
<tr><td>$\chi^2 = \frac{(26-28.88)^2}{28.88} + \frac{(21-16.99)^2}{16.99} + ...$

t.s $\chi^2 = 10.2$

($\nu = 4$) c.v $\chi^2 = 9.488$
or p = 0.0378 < 0.05

Reject H_0. There is significant evidence to suggest that there is an association between the sex of a cat and its colour.</td><td>M1

A1

A1

E1dep</td><td>1.3

1.3

2.1b

2.1a</td><td>PI method for contributions at least one correct.

awrt 10.2

awrt c.v. = 9.5
or awrt p = 0.04 and comparison with 5%

Must be in context and include element of doubt. Dependent on correct cv and ts.</td></tr>
</table>

Question	Scheme	Marks	AO	Notes
4(c)	There are far more male ginger cats (and fewer female ginger cats) than would be expected. or Ginger cats are more likely to be male than female. or Male cats are more likely than female cats to be ginger.	B1 E1	2.1a 2.1a	Identify "ginger" as the key colour. oe correct statement.
4(d)	There might be bias due to the fact that all of the cats are from one geographic area. There might be bias due to the fact that all of the cats have been stray/unwanted/abandoned/rescued.	E1 E1	3.1a 3.1a	geographical bias cat status bias
4(e)	No, (Katherine selected her data for convenience so) we cannot assume it was a random sample of cats.	E1	3.1a	Not reasonable because not a random sample of cats.
	Total	**14**		

Question	Scheme	Marks	AO	Notes
5(a)(i)	$\frac{95}{145} = \frac{19}{29} = 0.655$	B1	1.2	OE fraction form or awrt 0.655
5(a)(ii)	$\frac{8}{145} = 0.0552$	B1	1.2	OE fraction form or awrt 0.055
5(b)	$P(B_1) = \frac{68}{145} = 0.469$ $P(H) = \frac{26}{145} = 0.179$	B1	1.2	**Both** in OE fraction form or awrt 0.469 and 0.179
	$P(B_1) \times P(H)$ $\approx 0.0841 \neq P(B_1 \cap H)$	M1ft	2.1b	Use of multiplication law with answer to part (a)(ii) and figures above.
	not independent	A1cso	2.1b	dependent on correct calculations seen and part (a)(ii) correct and awrt 0.084 above
	Alternative 1 $P(B_1) = \frac{68}{145} = 0.469$ $P(B_1 \mid H) = \frac{8}{26} = \frac{4}{13} = 0.308$	(B1)		**Both** in OE fraction form or awrt 0.469 and 0.308
	$P(B_1) \neq P(B_1 \mid H)$	(M1ft)		OE using figures above
	not independent	(A1cso)		dependent on correct calculations seen
	Alternative 2 $P(H) = \frac{26}{145} = 0.179$ $P(H \mid B_1) = \frac{8}{68} = \frac{2}{17} = 0.118$	(B1)		**Both** in OE fraction form or awrt 0.179 and 0.118
	$P(H) \neq P(H \mid B_1)$	(M1ft)		OE using figures above
	not independent	(A1cso)		dependent on correct calculations seen

Question	Scheme	Marks	AO	Notes
5(c)	$26 \div 50$	M1	1.1	PI use of 26 or 50
	$= \frac{13}{25} = 0.52$	A1	1.2	either form
5(d)	The **probability** that a (tortoise) burrow is potentially occupied (has not been abandoned) given that it is in a flatwood area (habitat).	E1	2.1a	oe
5(e)	$(P(B_1 \cup B_2 \mid F) = 0.52)$			
	$P(B_1 \cup B_2 \mid H) = \frac{20}{26} = 0.769$	B1	1.2	awrt 0.769
	$P(B_1 \cup B_2 \mid S) = \frac{51}{69} = 0.739$	B1	1.2	awrt 0.739.
	(A potentially occupied burrow is most likely to be found in a) **hardwood** (habitat).	E1cso	2.1b	dependent on all correct figures in decimal form
	Total	**11**		

Question	Scheme	Marks	AO	Notes
6(a)	(A scatter diagram allows Femi to) Check for any a pattern within the data **or** See if there is an effect to be analysed **or** Check that the data are not just randomly scattered	E1	1.1	OE
6(b)	Femi wants to show that a lack of education and training affects prevalence of youth offending. **or** more resources put into education and training	E1	2.1b	OE identifying aims of research leads to seeking evidence of x causing y.
6(c)	Femi started with the conclusion of his research and is trying to find evidence to support it. **or** Femi has decided what the conclusion of his research should be before starting. **or** may be other factors **or** dependency may be the other way round **or** a correlation doesn't mean there is a causation	E1	3.1a	OE identification of the poor practice of **agenda driven design**.
6(d)	No NEETs implies 48 first time youth offenders per 10,000 of the population. 1% less (more) NEETs implies 21.2 fewer (more) first time youth offenders per 10,000 of the population.	E1 E1	2.1a 2.1a	Interpretation of a in context Interpretation of b in context

Question	Scheme	Marks	AO	Notes
6(e)	From diagram, largest y ≈ 250, **or** use x = 9 in regression equation to get y = 238.8 a-value implies that this could be decreased to 48. That would be a reduction of $\frac{250-48}{250}=0.808$ **or** $\frac{238.8-48}{238.8}=0.798$ (approximately 80%.)	M1 A1	1.1 1.2	Identification of c250 or calculation of awrt 239 from regression equation with x = 9 **and** use of 48. OE (may be seen on diagram) Equivalent calculations acceptable e.g. showing 20% remain

Question	Scheme	Marks	AO	Notes
6(f)	Femi's statements imply that not being in education etc. causes crime but the relationship found might be caused by some influence external to them both.	E1	3.1b	no evidence of causation
	Femi can't make any conclusions about education funding because he does not have any data about it.	E1	3.1b	no evidence on funding
	First time offending does not necessarily equate to not having a future.	E1	3.1b	no evidence implying future outcomes
	There is no direct evidence that the NEETs are the young offenders.	E1	3.1b	no direct link between NEETs and offending
	The approximately 80% reduction from 250 to 48 would not apply to all London boroughs.	(E1)		incorrect interpretation of b
	NEETs might have been given opportunities but chose not to take them.	(E1)		misinterpretation of NEETs
	There are other forms of education and training available to young people than sixth forms and colleges.	(E1)		limited scope of interpretation of NEETs
	Only looking at 16-18 year olds	(E1)		
	Statement concludes all sixth forms need more funding, but his previous comment was just about London	(E1)		
	Additional funding may not result in more students going to college	(E1)		
	Extrapolation used in first statement	(E1)		
				Any four relevant comments. List not exhaustive.
	Total	**11**		

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