



Pearson
Edexcel

Mark Scheme (Results)

Summer 2018

Pearson Edexcel GCE

In Statistics (8ST0) Paper 02

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 2018

Publications Code 8ST0_02_1806_MS

All the material in this publication is copyright

© Pearson Education Ltd 2018

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.

Question	Scheme	Marks	AO	Notes
1(a)	Disproportional...	E1	1.1	Seen anywhere
	...stratified (sampling)	E1	1.1	Seen anywhere Subtract second E1 for any additional terms except: <ul style="list-style-type: none"> • random • unbiased
1(b)	Unrestricted...	E1	1.1	Seen anywhere
	...random (sampling)	E1	1.1	Seen anywhere Accept 'simple random sampling with replacement' for E2 Subtract second E1 for any additional terms except: <ul style="list-style-type: none"> • unbiased
Total		4		

Question	Scheme	Marks	AO	Notes
2	<p>Possible criticisms</p> <ul style="list-style-type: none"> • The bar length is inconsistent with the numbering (for prison officers and school teachers). • All (or most) pay has gone down in the bar chart between 2005 and 2015 whereas pay has continued to rise since 2010 in the (line) graph. • We don't know whether inflation has been included. • 'CPI' has not been defined. • Units on vertical axis unclear. • The horizontal scale is unclear. • Not all public-sector workers are included (e.g. MPs) • It is difficult to read values from the line graph. • The data in one table starts at 2005, and the other starts at 2010. • It is not clear what date in 2005 and 2015 the bar data is from 	<p>E1, E1, E1</p> <p>E1</p>	<p>3.1a, 3.1a, 3.1a</p> <p>2.1a</p>	<p>Clear explanation that pay has decreased in the bar chart.</p> <p>Clear explanation that pay has increased in the line graph.</p> <p>or 'index not defined'</p> <p>Do not accept 'vertical axis does not start at 0'</p> <p>Accept 'cannot read off exact values'.</p> <p>or 'one table ends in 2015 and the other ends in 2017'</p> <p>Each bullet point scores E1 (max E3)</p> <p>Explanations in context</p>

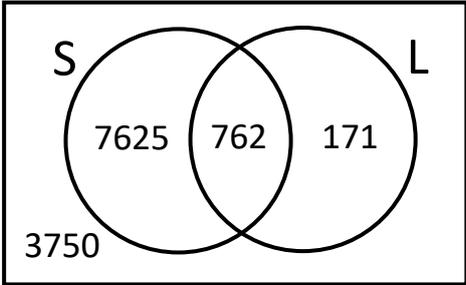
	Special case Insufficient information given regarding numbers used in Figure 2.	(E1)		This solution scores max E1
		Total	4	

Question	Scheme				Marks	AO	Notes							
3(a)	$H_0: \eta_M = \eta_F$ $H_1: \eta_M < \eta_F$				B1	1.3	oe [H ₀ : male population median = female population median] [H ₀ : samples taken from identical populations] [H ₀ : female frogs are not longer than male frogs, on average]							
	M	60	58	55			62	67	53	55	67	54	55	56
	F	57	72	77			86							
	M (rank)	9	8	4			10	11.5	1	4	11.5	2	4	6
	F (rank)	7	13	14			15							
	$T_M = 9 + 8 + \dots + 6 = 71$ or $T_F = 7 + 13 + 14 + 15 = 49$ $U_M = 71 - \frac{1}{2}(11 \times 12) = 5$ or $U_F = 49 - \frac{1}{2}(4 \times 5) = 39$ ts = 5 (or 39) (1-tail test, $\alpha = 0.05$) cv = 9 (or 35)				M1	1.3	Separating measurements into two groups.							
					A1	1.3	All ranks correct for either group.							
					M1	1.3	Attempt at rank total for either group.							
					M1	1.3	Attempt at either U Must include '11×12' or '4×5'							
					A1	1.3	cao							
					B1	1.3	cao							

	$5 < 9$ (or $39 > 35$) so reject H_0 .				B1ft	2.1b	Comparison of their sensible ts and sensible cv in same tail																																																		
	Conclude that there is evidence (at the 5% significance level) that the population median body length for <u>female</u> frogs is <u>longer</u> than that of <u>male</u> frogs.				E1dep	2.1a	Must be in context. Should not be definite in conclusion. Dep on both correct ts and cv.																																																		
	or Conclude that there is evidence (at the 5% significance level) to support Ramesh's belief.																																																								
	<i>Alternative (rank descending)</i> $H_0: \eta_M = \eta_F$ $H_1: \eta_M < \eta_F$				(B1)	(1.3)	oe [H_0 : male population median = female population median] [H_0 : samples taken from identical populations]																																																		
<table border="1"> <tr> <td>M</td> <td>60</td> <td>58</td> <td>55</td> <td>62</td> <td>67</td> <td>53</td> <td>55</td> <td>67</td> <td>54</td> <td>55</td> <td>56</td> </tr> <tr> <td>F</td> <td>57</td> <td>72</td> <td>77</td> <td>86</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>M (rank)</td> <td>7</td> <td>8</td> <td>12</td> <td>6</td> <td>4.5</td> <td>15</td> <td>12</td> <td>4.5</td> <td>14</td> <td>12</td> <td>10</td> </tr> <tr> <td>F (rank)</td> <td>9</td> <td>3</td> <td>2</td> <td>1</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> </table>				M	60	58	55	62	67	53	55	67	54	55	56	F	57	72	77	86								M (rank)	7	8	12	6	4.5	15	12	4.5	14	12	10	F (rank)	9	3	2	1								(M1)	(1.3)	Separating measurements into two groups, and attempt at ranks.			
M	60	58	55	62	67	53	55	67	54	55	56																																														
F	57	72	77	86																																																					
M (rank)	7	8	12	6	4.5	15	12	4.5	14	12	10																																														
F (rank)	9	3	2	1																																																					
				(A1)	(1.3)	All ranks correct for either group																																																			

4(c)	Possible explanations There is another group of (brighter) stars which were not in Figure 3 (the first scatter diagram). The distribution is not bivariate normal. The small sample is clearly not representative of the population. An outlier was removed, but this is not an outlier in the larger sample.	E1	3.1b	oe Accept ‘the relationship is nonlinear’
4(d)	Separate into different <u>star types</u>and produce <u>different regression lines</u> (or curves) for each star type.	E1	3.1a	oe
	Total	8		

Question	Scheme	Marks	AO	Notes
5(a)(i)	$P(S) = \frac{8387}{12308}$ or $\frac{16774}{24616}$ or 0.681	B1	1.2	oe awfw 0.68~0.682 Actual: 0.681426714
5(a)(ii)	$P(L') = \frac{12308 - 933}{12308}$	M1	1.2	12308 - 933 or 10749 + 626 or 11375 seen
	$= \frac{11375}{12308}$ or $\frac{22750}{24616}$ or 0.924	A1	1.2	oe awfw 0.92~0.925 Actual: 0.9241956...
	<i>Alternative</i> $P(L) = \frac{933}{12308} = 0.076$	(M1)	(1.2)	PI
	$P(L') = 1 - 0.076$ $= \frac{11375}{12308}$ or $\frac{22750}{24616}$ or 0.924	(A1)	(1.2)	oe awfw 0.92~0.925 Actual: 0.9241956...
	SC $\frac{10749}{12308}$ or awfw 0.87~0.874 scores M1			

<p>5(b)</p>		<p>B1</p> <p>B1</p> <p>B1</p>	<p>1.1</p> <p>1.1</p> <p>1.1</p>	<p>Outer box drawn</p> <p>Two overlapping circles</p> <p>Circles labelled S & L</p> <p>At least two figures correct and in correct region</p> <p>All four figures correct.</p> <p>SC: Venn diagram with probabilities scores max B2</p> <p>S: 0.6195</p> <p>$S \cap L$: 0.0619</p> <p>L: 0.0139</p> <p>Ext: 0.3047</p> <p>SC:</p> <p>-B1 for S/L interchanged</p>
<p>5(c)</p>	$P(S L') = \frac{P(S \cap L')}{P(L')} = \frac{n(S \cap L')}{n(L')}$ $= \frac{7625}{11375} = \frac{15250}{22750} = \frac{61}{91} = 0.670$	<p>M1</p> <p>A1</p>	<p>1.2</p> <p>1.2</p>	<p>PI</p> <p>Evidence of multiplication rule used (prob or freq)</p> <p>Denominator = 11375, 22750, or 91 seen scores M1</p> <p>oe</p> <p>awfw 0.67~0.671</p> <p>Actual: 0.67032967</p>
<p>SC:</p> <p>$P(L' S)$ calculated correctly = $\frac{7625}{8387} = 0.909$ scores M1</p>				

Question	Scheme	Marks	AO	Notes
6(a)(i)	[Let M represent number of claims made by motorcycle riders] $E(M) = 0.223$	B1	1.2	awfw 0.22~0.224 Actual: 0.2226
6(a)(ii)	$E(M^2) = 0.304$	M1	1.2	Clear attempt at finding $E(M^2)$ PI
	$\text{Var}(M) = E(M^2) - (E(M))^2$ $= 0.254$	A1	1.2	awfw 0.25~0.255 Actual: 0.2544492
SC sd=0.504 scores M1				
6(b)	[Let C represent number of claims made by car drivers] $\text{Var}(C) = 0.486^2 = 0.236$	B1ft	2.1b	awrt
	or $\sigma_M = \sqrt{0.254} = 0.504$			awrt Actual: 0.5044296
	$\mu_M = 0.223$ or $E(C) = 0.146$	B1ft	2.1b	Recognition that mean = expected value Must be convinced
The motorcycle riders have more claims, on average, than car drivers and the spread is roughly the same.		E1dep	2.1a	Both statements needed Must see <u>average</u> and <u>spread</u> oe Dep on both B1 marks

<p>6(c)</p>	<p>Possible statements (not exhaustive)</p> <p>The text just says ‘average’. This may not necessarily be the mean.</p> <p>Responses in the 2016 questionnaire may include claims in 2015.</p> <p>People may not want to admit if they’ve had lots of claims.</p> <p>The (car) sample may not be representative, as only certain people would be happy to answer a questionnaire with (at least) 18 questions.</p> <p>The (car) sample may not be representative as it only includes people who have visited the website.</p> <p>The motorcycles are only from a single company, so there may be bias (e.g. low risk insurance companies).</p>	<p>E1, E1, E1</p>	<p>3.1b, 3.1b, 3.1b</p>	<p>or people may have lied</p> <p>E1 for each sensible statement (Max E3)</p>
<p>Total</p>		<p>9</p>		

Question	Scheme	Marks	AO	Notes																				
7(a)	[Let X represent shares at close on 23 rd , and Y represent shares at close on 24 th]																							
	$H_0: \eta_X = \eta_Y$ $H_1: \eta_X > \eta_Y$	B1	1.3	oe [H_0 : population median share price at close on 23/6 = population median share price at close on 24/6] [H_0 : Population median difference=0] [H_0 : Share prices at close on 24/6 were not lower than close on 23/6, on average]																				
	<table border="1"> <tr><td>Vodafone</td><td>+</td></tr> <tr><td>Worldpay</td><td>-</td></tr> <tr><td>Sky</td><td>-</td></tr> <tr><td>TUI AG</td><td>-</td></tr> <tr><td>St. James's Place</td><td>-</td></tr> <tr><td>Direct Line</td><td>-</td></tr> <tr><td>SSE</td><td>-</td></tr> <tr><td>Schroders</td><td>-</td></tr> <tr><td>Smiths Group</td><td>-</td></tr> <tr><td>Royal Dutch Shell</td><td>+</td></tr> </table>	Vodafone	+	Worldpay	-	Sky	-	TUI AG	-	St. James's Place	-	Direct Line	-	SSE	-	Schroders	-	Smiths Group	-	Royal Dutch Shell	+	M1	1.3	Attempt to calculate signs of differences (or signed differences) between the correct two columns.
	Vodafone	+																						
	Worldpay	-																						
	Sky	-																						
	TUI AG	-																						
	St. James's Place	-																						
	Direct Line	-																						
	SSE	-																						
Schroders	-																							
Smiths Group	-																							
Royal Dutch Shell	+																							
(Under H_0), $S \sim B(10, 0.5)$ $P(S \geq 8) = 0.0547$	M1	1.3	PI awfw 0.0546~0.055 Actual: 0.0546875																					
or $P(S \leq 2) = 0.0547$	A1	1.2	Alt: Critical region {9,10} or {0,1} with justification																					

