Pearson Edexcel

Mark Scheme (Results)

Summer 2019

Pearson Edexcel GCE In Statistics (9ST0)

Paper 3 Statistics

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 9ST0_03_1906_MS

All the material in this publication is copyright
© Pearson Education Ltd 2019

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)	0.073 + 0.011 = 0.084	B1	1.1	cao read from table
1(b)	$(1 – 0.084)^3$ or use of B(3, 0.084), P(X=0)	M1	1.2	oe, eg 0.916^3 PI
	= 0.769	A1	1.2	awfw 0.768~0.769
	SC: 0.77 seen with no method shown scores M1 A0			
1(c)	mode = 6	B1	1.1	from table
	P(X < M) = 0.5 0.006 + 0.02 + 0.096 + 0.377 = 0.499	M1	1.2	PI correct method adding p, can be 0.011 + 0.073 + 0.417 = 0.501 Check for working on table
	median = 6	A1	1.2	cao
	(2×0.006) + … + (8×0.011)	M1	1.2	PI correct method Check for working on table
	mean = 5.442	A1	1.2	awrt 5.44
	SC: 5.4 with no method shown scores M1 A0			
1(d)	$E(X^2) = (2^2×0.006) + … + (8^2×0.011)$ (= 30.458) $Var(X) = 30.458 – 5.442^2 = 0.842$	M1	1.2	Correct method for Var(X) or awfw 0.842~0.843 seen
	$sd = \sqrt{0.842} = 0.918$	A1	1.2	awrt 0.91~0.95

Question	Scheme	Marks	AO	Notes
1(e)	**Supporting comments:** Most of the data are within 3 sd of the mean. Approximately 2/3 of the data are within 1 sd of the mean. Approximately 95% of the data are within 2 sd of the mean.			oe comparison with proportion of data. mean$\pm$1sd (4.5, 6.4) $\approx$ 60% of data mean$\pm$2sd (3.6, 7.3) $\approx$ 90% of data mean$\pm$3sd (2.7, 8.2) $\approx$ 100% of data
	The mean, median and mode are close so the **dataset** is fairly **symmetrical**.			
		E1	3.1b	One comment in support.
	Comments against: Clutch size is discrete data not continuous. Clutch sizes in different nests may not be independent of one another (due to birds being related or living very close together). There is a *negative skew* to the data so it is not symmetrical.			
	Clutch sizes have finite limits.			oe **or** you cannot have a negative/infinite number of eggs
		E1	3.1b	One comment against
		E1	3.1b	One further comment from either list
				Comments must reference the data specifically, not a general normal distribution.
	Total	**13**		

Question	Scheme	Marks	AO	Notes
2(a)	$d = \dfrac{58 - 47.9}{\sqrt{160.7}} = 0.797$	B1	1.2	awfw 0.795~0.800
2(b)	The t-test shows **significant** evidence that **male** IBS **patients** have a **different/ higher quality of life** than **female** IBS patients on average.	B1	2.1a	oe Correct one or two tailed interpretation of test in context. Allow "people with IBS" instead of patients.
	The Cohen's *d* value shows that this difference in average quality of life is of a medium/large size.	B1dep	2.1a	Standard interpretation of Cohen's $0.5 \leq d \leq 0.8$. Accept medium/large effect size. Dep on correct (a). May be seen in (a).
	Total	**3**		

Question	Scheme	Marks	AO	Notes
3(a)	**Possible comments** **(not exhaustive)**			Must see **relative search interest**. Condone **search** or **interest** throughout.
	Relative search interest is **higher** for **boots** than for **sandals**.			May use numerical comparison of maxima/minima
	Relative search interest in footwear appears to have **increased** over this time period.			Must refer to search/interest increasing
	Relative search interest in footwear follows a **cyclical/seasonal/annual** pattern.			**or** boots/sandals **or** peaks/troughs are roughly 365/350 days apart
	When relative search interest in **boots** is **high**, relative search interest in **sandals** is **low**.			**or** vice versa oe
	There is greater variation in the relative search interest for boots than sandals			
	The **largest** relative search interest shown **in boots** was at **around day 650**.			Question asks for four *different* comments.
	The **largest** relative search interest shown **in sandals** was at **around day 870**.			Hence two marks only available for highest/ lowest comments.
	The **smallest** relative search interest shown **in boots** was at **around day 150**.			
	The **smallest** relative search interest shown **in sandals** was at **around day 280**.			**or** 670/1000
		E1, E1, E1, E1	2.1a	Any four relevant and distinct comments **about the graph provided only**. Do not allow

				inference such as “people search for sandals more in the summer” except as support for correct comments.

Question	Scheme	Marks	AO	Notes
3(b)	**Possible comments (not exhaustive)** The numbers/volumes/frequencies of search interest (into boots and sandals).			Frequency information
	The times of the year rather than just the days.			**or** months/seasons etc
	Which year(s) the data covers.			
	Weather information.			
	Gender split.			
	Whether the searches only returned footwear-related sites.			oe
		B1, B1	3.1a	Any two correct relevant comments
	Total	**6**		

Question	Scheme	Marks	AO	Notes
4(a)	**Possible assumptions** Radioactive particles are detected at a constant average rate. Radioactive particles are detected at random. Radioactive **particles** are detected independently **from one another**. Radioactive particles are detected **singly**. Radioactive particle **counts** have (almost) no upper limit. Radioactive particle **counts** have approximately equal mean and variance.	B1, B1, B1	2.1a	allow Radiation is detected/emitted/occurs at random/randomly. oe **or** Counts/particles occur singly Any three relevant and distinct comments **all in correct context**.
4(b)	84 ÷ 60 mean = 1.4 sd = √1.4 = 1.18	M1 A1 A1	1.2 1.2 1.2	rescaling cao awrt 1.18
4(c)	P(X < 4) = P(X ≤ 3) = 0.946	M1 A1	1.2 1.2	 awrt 0.946
4(d)	Exponential distribution with parameter / λ = 1.4 **or** mean = 1/1.4 = 0.714	B1 B1dep	2.1b 2.1b	Choosing correct dist. **or** $[F(x) =]1 - e^{-\lambda x}$ **or** $[f(x) =]\lambda e^{-\lambda x}$ Dep on previous B1

Question	Scheme	Marks	AO	Notes
4(e)	$P(X < 2) = 1 - e^{-1.4 \times 2}$	M1	1.2	oe PI Condone use of $\lambda = 0.714$
	= 0.939	A1	1.2	awrt 0.939
4(f)	$\left[\frac{1}{1.4} =\right]$ 0.714 (seconds)	B1	1.2	awrt 0.714 or $\frac{5}{7}$
	Total	**13**		

Question	Scheme	Marks	AO	Notes
5(a)	A signifies dose without cardboard. B signifies dose with cardboard.			
	$H_0: \mu_A - \mu_B = 2$ $H_1: \mu_A - \mu_B > 2$	B1	1.3	oe both, subscripts defined.
	$s_p^2 = \dfrac{(5-1)0.1^2 + (5-1)0.08^2}{5+5-2}$	M1	1.3	PI formula all correct
	$= 0.0082$	A1	1.3	PI awrt 0.0082 **or** s_p = awfw 0.0905~0.0906
	$ts = \dfrac{(13-9.99)-2}{\sqrt{0.0082\left(\frac{1}{5}+\frac{1}{5}\right)}}$	M1	1.3	PI Numerator or denominator correct
		M1dep	1.3	PI ts formula correct dep previous M1 and sp or sp^2 formula attempted
	= 17.6	A1	1.3	awrt
	cv = 2.896 or $p = 5.46\times10^{-8} < 0.01$	B1	1.3	**cv** awfw ±(2.89~2.9) **or** correct **p-value compared with 1%** p-value awrt 5.5×10^{-8} implies M1A1M1m1A1
	(17.6 > 2.896) Reject H_0	A1dep	2.1b	PI correct conclusion, dep previous A1, B1
	There is significant evidence that the **cardboard/cover** has	E1dep	2.1a	Conclusion correct and in context

	decreased the **radiation/dose** by **more than 2**μSv/h on average.			Dependent on test all correct. Not definite in conclusion
	Alternative A signifies dose without cardboard. B signifies dose with cardboard. H_0: $\mu_A - \mu_B = 2$ H_0: $\mu_A - \mu_B > 2$	(B1)		oe both, subscripts defined.
	$s_p^2 = \frac{(5-1)0.1^2 + (5-1)0.08^2}{5+5-2}$	(M1)		PI formula all correct
	= 0.0082	(A1)		PI awrt 0.0082 **or** s_p = awfw 0.0905~0.0906
	cv = 2 +	(M1)		
	$t \times \sqrt{0.0082\left(\frac{1}{5}+\frac{1}{5}\right)}$	(M1)		
	= 2.166	(A1)		
	Use of $t = 2.896$	(B1)		awfw ±(2.89~2.9)
	ts (= 13-9.99) = 3.01 > 2.166	(A1)		Comparison of 3.01 with cv
	There is significant evidence that the **cardboard/cover** has **decreased** the **radiation/dose** by **more than 2**μSv/h on average.	E1dep	2.1a	Conclusion correct and in context Dependent on test all correct. Not definite in conclusion

	SC1: Incorrect unpooled t-test or z-test used $\frac{(13-9.99)-2}{\sqrt{\frac{0.1^2}{5}+\frac{0.08^2}{5}}}=17.6$ Check not **SC2** can score B1 M0 A0 M1 M1 A1 B1 A0 E0				
	SC2: Correct unpooled t-test $\frac{(13-9.99)-2}{\sqrt{\frac{0.1^2}{5}+\frac{0.08^2}{5}}}=17.6$ **and** $p=9.22\times10^{-8}$ can score B1 M1 A1 M1 M1 A1 B1 A1 E1				cao

Question	Scheme	Marks	AO	Notes
5(b)	**These first two points** The test in part (a) showed that a dry cardboard cover 4mm thick over (contaminated) soil can reduce the dose of radiation from the dangerous level of 13μSv/h to less than 11μSv/h.	E1 E1	2.1a 2.1b	reference to **cardboard** and **soil** and **reduce radiation/dose** Numerical evidence **or** by more than **2** μSv/h **or** below **11.4** μSv/h
	Any two practical considerations **(list not exhaustive)**			
	However, this was only a test of dry cardboard. If covering soil outside it would be likely to get wet.			Wet cardboard not tested.
	The test was only of 4mm thick cardboard. Cardboard might not always be available in this thickness.			Other thicknesses of cardboard not tested.
	There is no indication from this test that the cardboard would continue to provide an effective shield over a length of time.			Length of time of protection not tested.
	The assumptions of the t-test might not have been met so the results of the test may not be reliable.			Violation of assumptions.
	There was no indication that the same soil samples were tested both with and without the cardboard. So the results could be biased.			Possible bias in the sampling.
	Comment on the cheapness of cardboard.			
	Comment on availability of cardboard.			

	Sensible comments considering practical physical considerations.	E1, E1	3.1a	eg • Flammability • Durability • Flexibility • Weather • Time constraints Do **not** accept ‘small sample’
	Total	**13**		

Question	Scheme	Marks	AO	Notes
6(a)	A signifies incorrect claim submitted **after** change in procedure B signifies incorrect claim submitted **before** change in procedure			
	$H_0: \pi_A = \pi_B$ $H_1: \pi_A \neq \pi_B$	B1	1.3	oe condone use of p or population proportion. both hypotheses correct.
	$\hat{p} = \frac{43+19}{522+315} = \frac{62}{837} = \frac{2}{27}$	M1	1.3	PI Attempt at $\hat{p}$ (= 0.0741)
	Test statistic	M1	1.3	PI Attempt at formula with numerator correct, ignore signs
	$= \dfrac{\frac{43}{522} - \frac{19}{315}}{\sqrt{\frac{2}{27} \times \left(1 - \frac{2}{27}\right) \times \left(\frac{1}{522} + \frac{1}{315}\right)}}$	M1	1.3	PI $\frac{1}{522} + \frac{1}{315}$ seen
		M1ft	1.3	PI Denominator of correct form with their $\hat{p}$ used
	ts = 1.18	A1	1.3	awrt 1.18 ignore sign Note p = awrt 0.238 implies M1M1M1M1A1
	1.18 < 1.96 **or** 0.119 > 0.025	A1	2.1b	Comparison of cv with ts in same tail **or** p = awrt 0.238 compared to 0.05
	(Accept H_0.) There is **no** significant evidence to suggest the proportion of expenses	E1dep	2.1a	Conclusion correct, in context. Dependent on test all correct.

	claims that are **rejected** because they are **incorrectly submitted** has **changed**.			Not definite in conclusion. **Rejected** or **incorrectly submitted**.

Question	Scheme	Marks	AO	Notes
6(b)	These were not random samples of claims made under the old and new procedures. One person may have submitted more than one claim. Therefore claims may not have been independent of each other. The two samples of claims were made at different times of the year This might cause a bias/confound in the results. The probability that someone makes a claim might be different at different times of the year. The error rate might be different when the procedure is new than when people have got used to it. There may have been staff changes between the two periods. New employees are likely to make more errors. Andrew only had information for one department. The change in procedure might have affected other departments differently.			Not random samples Violates assumption of independence. Possible bias due to time of year/special events. Not constant p
		E1, E1, E1	3.1a	Any three of these comments **in context**

Question	Scheme	Marks	AO	Notes
6(c)	Andrew should use his spreadsheet of employees to select a random sample (from each grade) by using a random number generator function or RAND() or RANDBETWEEN()	B1	1.1	oe Random sample eg using an appropriate spreadsheet function **or** Randomly select clusters of employees by (eg) manager
	Andrew should find out what proportion of staff are in each grade and stratify his sample by grade. **or** Andrew should find out what proportion of claims are made by each grade and stratify his sample by grade.	B1	1.1	Stratify by grade.
	Andrew should contact his sample of employees directly to ensure a response within the timeframe.	B1	1.1	Indication of active rather than passive data collection: emailing/phoning etc.
	If Andrew does not get responses from all of his sample in time to do the analysis, he should randomly sample further employees in the appropriate grade and approach them. **or** Andrew could initially contact more employees than he requires, then use the completed responses as a new sampling frame. **or** Andrew could make the survey easy to complete by using a web form.	B1	1.1	Sensible contingency for non-response/non-contact within the time-scale **or** Sensible suggestion for additional action to ensure sample obtained by deadline
				NOTE: All marks are independent
	Total	**15**		

<table>
<tr><th>Question</th><th>Scheme</th><th>Marks</th><th>AO</th><th>Notes</th></tr>
<tr><td>7(a)</td><td>H$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6$
H_1: at least two of the means differ</td><td>B1</td><td>1.3</td><td>or $\mu_i = \mu$
for $i = 1 \text{ to } 6$
oe</td></tr>
<tr><td></td><td>T = 574
$SS_T = 13962 - \frac{574^2}{26} = 1289.846$</td><td>M1</td><td>1.3</td><td>SS_T method
PI</td></tr>
<tr><td></td><td>Time of day SS_B working
$\frac{109^2}{3} + \frac{89^2}{3} + \frac{61^2}{4} + \frac{91^2}{5} + \frac{135^2}{7} + \frac{89^2}{4}$
= 13770.938

$SS_B = 13770.938 - \frac{574^2}{26}$
= 1098.784</td><td>M1</td><td>1.3</td><td>SS_B method
PI</td></tr>
<tr><td></td><td><table>
<tr><th></th><th>SS</th><th>df</th><th>MS</th></tr>
<tr><td>Times of day</td><td>1098.78</td><td>5</td><td>219.76</td></tr>
<tr><td>Error</td><td>191.06</td><td>20</td><td>9.55</td></tr>
<tr><td>Total</td><td>1289.85</td><td>25</td><td></td></tr>
</table></td><td>M1dep</td><td>1.3</td><td>PI
SS_E method, no negative SS values.
Dep one previous M1</td></tr>
<tr><td></td><td></td><td>B1</td><td>1.3</td><td>PI
df 5 and 20 in correct row</td></tr>
<tr><td></td><td></td><td>M1dep</td><td>1.3</td><td>PI
MS divide SS by df
dep all previous M marks but ft on df values</td></tr>
<tr><td></td><td>$F = \frac{219.76}{9.55}$</td><td>M1dep</td><td>1.3</td><td>PI
F method
dep all previous M marks but ft on df values</td></tr>
<tr><td></td><td>= 23.0</td><td>A1</td><td>1.3</td><td>awfw F = 21.9~23.5
Note $p = 1.16\times10^{-7}$ implies M1M1M1B1M1M1A1</td></tr>
</table>

	cv $F_{5, 20} = 2.71$ (< 23) **or** $1.16\times10^{-7} < 0.05$	A1	2.1b	Correct awrt cv 2.71 or comparison awrt p = 1.16×10^{-7} with 5%.
	(Reject H_0) There is significant evidence that **mean** healing **times** for **at least two** of the different times of day that burns occurred have **different** healing times on average.	E1dep	2.1a	Correct conclusion in context dep test fully correct. Not definite in conclusion.
	Means: 00:00 – 03:59 = 36.3; 04:00 – 07:59 = 29.7; 08:00 – 11:59 = 15.3; 12:00 – 15:59 = 18.2; 16:00 – 19:59 = 19.3; 20:00 – 23:59 = 22.3	M1	2.1b	*Attempt* at calculation of mean average heal times, at least two correct to 3 s.f.
	(It appears that) burns incurred at night take longer to heal than those that occur during the day.	E1	2.1a	Correct comparison of specific times, in context. Accept 'different' instead of 'longer'.
	Specifically, burns incurred between 0:00 and 03:59 take longer to heal than those incurred between 08:00 and 11:59.	E1dep	2.1b	Clear identification of correct time-intervals. Dep test correct & M1 for averages. (Both E marks can be gained for this statement alone.) Must see 'longer' oe.

Means	
00:00 – 03:59	36.3
04:00 – 07:59	29.7
08:00 – 11:59	15.3
12:00 – 15:59	18.2
16:00 – 19:59	19.3
20:00 – 23:59	22.3

Question	Scheme	Marks	AO	Notes
7(b)	Secondary data was used, so they did not have any control over: • the method of recording the times to recovery meaning the data may not be reliable, **or** • monitoring the degree of healing.	E1	3.1a	Secondary data Lack of control of process by investigators, implying unreliability of data; any sensible context.
	There might be unknown bias inherent in the sampling method used or unknown confounding factors that influenced the result.	E1	3.1a	Sensible reason for possible bias or confounding factors. Could be specific e.g. severity or type of burn, age of patient, location of treatment etc.
7(c)	The researchers should take a larger random sample.	E1	2.1b	Larger **or** random sample
	They should control the collection of the recovery times (and other factors about the burn victims) themselves. **or** Ensure data on other factors relevant to burn healing are included.	E1	2.1b	Use primary data Allow a mention of using data on specific factors. eg severity, type of burn, age of patient, location of treatment etc.
	Total	**17**		

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