Pearson Edexcel

Examiners' Report

Principal Examiner Feedback

October 2021

Pearson Edexcel International A Level

Mathematics in Statistics S1 (WST01)

Paper : WST01/01

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 candidates at: www.pearson.com/uk

October 2021
Publications Code WST01_01_2110_ER

General

There were opportunities on this paper in all the questions for all students to make some progress but questions 2(e), 3 (e) and (f) and 6(d) proved to be more challenging. The questions requiring a comment or explanation in words were often not answered very well and sometimes not even attempted. A general comment, which applies to other questions as well, is that many students do not really understand that evidence has to be given for 'show that' questions.

Report on Individual Questions

Question 1

In part (a) whilst most students wrote the correct answer of 0.5 a significant minority gave their answer in terms of the unknowns p and q rather than find the actual probability. The instruction "find" means that some calculation needs to be done.

In part (b) many students lost marks due to lack of clear labelling, particularly of $\mathrm{P}(A \cap B)$
Labelling clearly what the probabilities are they have found is required since they could represent other areas. For example, in this case both $\mathrm{P}(A' \cap B)$ and P(A) equal 0.3
There was the usual confusion between events being independent and events being mutually exclusive highlighted by statements such as 'they are not independent as they overlap'
Others just guessed and made a statement without any backing evidence.

Part(c) was well answered. Those with algebraic answers in (a) and (b) usually realised they needed to use $p + q = 0.2$ to find the required values.

Some highly unusual positions of the circle were seen in part (d) but most answered this part correctly. The main errors were drawing a circle overlapping A or a circle that had part of it outside the rectangle.

Question 2

In part (a) most students showed enough working to gain the mark.

Part (b) was also generally well answered. The most common error made was using $\left(\frac{273}{12}\right)^2$ in the calculation of Spp and giving their answer to an inappropriate degree of accuracy. Candidates should give their final answers to a minimum of 3 significant figures unless stated otherwise in the question.

The calculations for the regression line in part (c) were usually carried out well and an encouraging number arriving at a correct equation in x and p. The gradient was mostly correct although a few had the fraction the wrong way up. In calculating the intercept errors were made with students using 273 & 93 instead of the means. In a few cases *a* and *b* were found but the final equation of the line was not stated.

Part (d) was poorly answered. The questions asked for an interpretation which indicates that a comment in context is required. The main error made by those who gave a reason in context was not recognising that a single rise in the number of employees led to a rise in the amount spent on paper of $156 and not $1.56

Part (e) was more challenging with many students unable to attempt the question. For those who did attempt the question the most common error was using 93 rather than $\frac{93}{12}$. These students then compounded their error by showing that they didn't know how to find the percentage.

<u>Question 3</u>

This question proved to be accessible for most students with many gaining the majority of the marks for parts (a) - (d).

Very few incorrect answers were seen in part (a). In part (b) the main errors, since the answer was given, were making up two ages (not the UQ and LQ) which differed by 16 as their 'working'. A few responses presented the positions of Q_1 and Q_3 as the quartile ages, but of course did not obtain an answer of 16. A small number merely put $Q_3 - Q_1 = 16$ without giving the values of the quartiles.

In part (c) most students gained the correct limits of 21 and 85. The question requires students to show that there are 3 outliers and having gained the correct limits some did not list the 3 outliers in this part in order to show there are 3outliers. A few made errors due to misunderstanding the definitions given for outliers, a mistake that has frequently occurred on similar questions in past examinations.

Part (d) was another opportunity for students to score highly with many gaining full marks. It was pleasing to see that very few drew 2 tails on each side of the box. Plotting errors tended to be with plotting of Q_2 or plotting the outlier of 13 at 18.

Part (e) proved to be challenging for students where few comments referring to the distribution of ages were seen. Others didn't include any values for the statistics they used and others commented on the similarities.

The vast majority of students gained no marks in part (f). Of those who gave a response some gave examples of actual ages rather than a range of ages. In (i) many stated that the granddaughter's age is an outlier and in (ii) the most common response, gaining M1A0 was saying Anja was above upper quartile.

However, there were many scripts where part (f) was left blank.

Question 4

Students who completed the tree diagram in part (a) by adding the remaining branches had a high success rate with this question. Unfortunately, a significant number of students did not add any branches and were therefore far less successful.

In part (b) it was usually those with a 'full' tree diagram who recognised the possible cases added together the products of 3 probabilities. As the students were asked to show a probability some more imaginative students came up with their own products that made $\frac{1}{10}$ rather than trying to establish where the given answer came from.

Part (c) discriminated the most able students and was clearly the most challenging part of this question. Conditional probability is a topic that many find difficult. The most common error was using a product of 2 probabilities rather than 3 in the numerator.

The majority of students knew what was required in part (d). The most common error was to miss out the value $x = 0$. This was often accompanied by the correct probabilities for $x = 1$, 2 and 3. Had students checked whether the sum of their probabilities equalled 1 they may have realised that they had missed zero out.

In part (e) finding the expectation seemed to be a well understood technique with the majority of students demonstrating their knowledge using their probability distribution.

Question 5

In part (a) the majority of students gained the mark although many calculated E(X) rather than using symmetry to write down the correct answer.

Part (b) was a good source of marks for those who were happy to work in terms of p and q. Others came back to this part after solving part (c) and substituted their values of p and q to try and show that $u = \frac{4}{15}$.

It should be noted that information given later in a question should not be used in earlier parts.

Pleasingly most students know the equation for Var(X) and were able to score the first two marks for $162q + 50r = 37$. Those who realised that $q + r = \frac{11}{30}$ were then able to continue on to find the correct values of q and r.

Part (d) was another more challenging question which many did not attempt. Those that did manage to find the correct values for the length of the line OP usually went on to score at least 5 marks. The last mark was often lost by failing to combine the two values of 13 and 15. Some failed to transfer skills from pure maths and were unable to find the lengths of the lines and some just used the values of Y.

Question 6

This question involves the use of the normal distribution and students are welcome to use calculators to find the equivalent values to the tables. When a question asks you to show a probability full working must be shown. In part (a) many students wrote down the standardisation followed by 0.15 missing out the intermediate step and the accurate answer. Some students simply wrote down 0.1500… with no working at all. Students who used tables tended to show all the necessary steps and gave an intermediate answer to the required degree of accuracy.

In part (b) the majority were able to standardise correctly and set equal to a z-value in the required range. The main errors were

- not using an inaccurate value such as 1.64 instead of 1.6449, Students should use the Percentage points table to find the z value when appropriate.
- using the wrong sign for 1.6449 appropriate to their standardisation giving 34.1, a value higher than the upper limit in part (a)

In (c)(i) many students did not realise that the ratios only applied to the middle 80% of the data, thus failing to obtain 0.55 and hence standardising with the wrong z-value. Those that did standardise with a correct z-value often did not show how they arrived at 25.8

In part (c) (ii) students tended to be far more successful in part (ii) mainly due to the fact that it was $P(H < m) = 0.25$ could be obtained in the correct way of 0.05 + 0.20 or for those who thought the ratios referred to the whole curve and getting $P(H < m) = 2/8$. Those who did lose marks after this were because they used $z = 0.68$ instead of 0.67 or using the wrong sign for their standardisation.

Students found part (d) quite challenging with few students successfully completing this part of the question. The main error was using the smaller end of the measurement of the shelf height rather than the upper end. Some forgot that there were multiple shelves of each height and others forgot to add the correct width of the wood. Few diagrams were seen but would probably have led to a few more fully correct solutions.

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