Pearson Edexcel

Examiners' Report
Principal Examiner Feedback

October 2022

Pearson Edexcel International Advanced Level In Statistics S1 (WST01) Paper 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

October 2022
Publications Code WST01_01_ER_2210

General

There were opportunities on this paper in all the questions for all students to make some progress but questions 2(e), 3(c), 3(d), 5(c) and 7(f) 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 a full method should be given for 'show that' questions.

Report on individual questions

Question 1

(a) It was pleasing to see that many candidates knew a suitable method to work out the required area. The majority of those who were unsuccessful worked with frequency densities and did not realise that if they had the same frequency density the bars would have the same height. Others had an extra multiple or divisor resulting in answers of 16.5 or 37.125.

(b) In this part those who worked with frequency density were far more successful than those who worked with area. The most common errors were to use the frequencies with no scaling or not using given height of 10 cm.

(c) In (i) it was disappointing to see that some candidates were unsure of how to find the quartiles with a minority giving quartiles which were outside the range of the data. Students would be advised to check their answers to see if they make sense. In (ii) some candidates lost marks by stopping once they found Q3 and did not go on to calculate the IQR. Students are advised to.

(d) There were quite a few candidates who used other methods to calculate skewness ignoring the question which asked them specifically to use the measure given in the question. Many of those who had incorrect answers in (c) were able to gain this mark by substituting their values with 45 into this formula. Other students did the calculation but did not give a description of the skewness.

Question 2

(a) Most responses were fully correct, although a disappointing number of students lost marks for only stating their values of S_{tt} and S_{ct} to the nearest integer or the value of S_{tt} to 2 decimal places. The question clearly stated that exact values were required. The occasional incorrect method seen was $S_{tt} = 82873 - \left(\frac{1361}{40}\right)^2$ rather than $S_{tt} = 82873 - \left(\frac{1361^2}{40}\right)$.

(b) This part was well answered with the majority of students scoring both marks. A small number of them lost the final mark by not showing any answer correct to at least three significant figures. The requirement for this level of accuracy is given on the instructions on the front of the paper and has been required for many examination series.

(c) Non contextual responses, scoring zero, were extremely common; the majority only commenting on the value of *r* showing strong positive correlation with no attempt to explain what this actually means in terms of the two variables, ticket sales and production costs. Others incorrectly interpreted 0.865 as the gradient of a regression line, stating as the value of c increases by 1 (million), *p* increases by 0.865 (million).

(d) The key to part was to understand that, because this was a 'show that' question it is essential for students to give clear and complete working if they are to score full marks. This topic is well understood but far too often the required detail was missing from the students answers. The final mark was often lost if we did not see the actual values of S_{ct} and S_{cc} in the calculation for b. We also required a more accurate value for either b or a than the values of 0.976 and −5.84 which are given in the question.

(e) Most students were able to gain the method mark in (e) for substituting 90 into their regression equation. Very many lost the accuracy mark by giving incomplete or incorrect units for their answer. We needed to see both the £ sign and the 'millions' for this mark to be scored. A small number incorrectly substituted $c = 90\,000\,000$ other substituted 90 million for t instead of c.

(f) This part was quite challenging for many students. The most common errors seen here were substituting t with just 0.8 or with $0.2c$ and occasionally there were errors in dealing with the direction of the inequality sign either in the working or in the final answer.

Question 3

Overall, this question proved slightly more challenging with a few students making no attempt at the question.

(a) Those students who attempted this part were usually able to gain some marks. A common misunderstanding was for students to substitute −1.2 into the given equation rather than first calculating the value of $\bar{x}$.

(b) It was disappointing to see that less than 30% of students used a correct formula to find the σ_x or σ_x^2. The most common errors were $\sqrt{5.1 - \left(\frac{-1.2}{8}\right)^2}$ or $\sqrt{\frac{5.1}{8} - \left(\frac{-1.2^2}{8}\right)}$. Even fewer students realised that once they found σ_x or σ_x^2 that they needed to multiply by 2 or 4 respectively. The common errors were to multiply the standard deviation by 4 or multiply the variance by 2 or in some cases students multiplied by $\frac{1}{2}$ or $\frac{1}{4}$.

(c) This part was challenging for many students. In (i) a common incorrect solution was an answer suggesting a rather small squirrel of length 1.2cm, from looking at the difference between $\Sigma x = -1.2$ and $\Sigma x = 0$ not realising the body length is represented by b and that x is the coded length. In (ii) less than 35% of students used a correct method to find $\sum x^2$. Of those who were able to find $\sum x^2$, many divided by the original 8 instead of the new number of squirrels. Another occasional misunderstanding had students stating that adding does not change or affect the standard deviation and so gave the same answer here as they had in part (b).

Question 4

By far the most common total scores for this question were 1 or 4. The vast majority scored the first mark with well recognised notation enabling them to find F(6) and F(7). Students who understood that the values obtained from the cumulative distribution function then had to be used to find the three probabilities P(W = 6, 7 and 8) usually found them accurately and went on to successfully calculate E(W) and score all four marks. The most common misunderstanding was simply to use the cumulative probability values instead of the individual probabilities in the attempted calculation of E(W).

Question 5

(a) This was the standard normal distribution question and was generally well answered with the vast majority scoring all three marks. A significant number of students lost the third mark by subtracting from one, the value obtained from the normal tables. In recent examination series this basic question had been designed such that the subtraction from one was required. This suggests that some students were somewhat programmed to carry out this final subtraction with no thought about the area of the distribution which was relevant. A simple diagram would have helped many to avoid this error.

(b) Students who identified the correct probability statement $P(W < k) = 0.85$ usually went on to score well. The most common error was the use of an inaccurate z-value of 1.04. Students must be advised that the use of the 'percentage points' table, rather than main normal distribution table, gives the four decimal place accuracy that is expected here.

(c) At the beginning of part (c) many students were able to score the first mark for correctly finding the probability of a fruit weighing less than 66g but what followed was often very confused with only the most able scoring any further marks.

Question 6

(a) This was particularly well answered with the vast majority of candidates being able to score all three marks here. Those that struggled, struggled with (ii) often giving an answer of 0.12 or 0.13.

(b) This proved to be challenging for many candidates. The most able candidates were able to score all 6 marks, but many struggled to form all three equations required to find q and r. The equation most commonly missed was the one using $\sum p = 1$ to gain the third equation.

(c) Most students knew what was required here and used their value of r to form the starting equation. The most common error was omitting 0.13.

Question 7

(a) This was very poorly scored overall with around 40% of students gaining the mark. Whilst there was a mixture of incorrect answers, the most common incorrect answers were 0 or 0.1.

(b) This was well done with the majority of the candidates showing the full method required.

(c) This was well answered. The main error was forgetting to square E(X) before subtracting.

(d) This was well answered. The most common errors were not squaring the (–3) and/or adding 4. A number of candidates multiplied by 4^2 rather than 3^2 which is something not often seen in previous years.

(e) Few students gained full marks in this part of the question, with many omitting one of the products or incorrectly multiplying their answer by 2 at the end.

(f) This part was challenging for the majority of students with many making no attempt. Where students understood the requirements and calculated four products full marks were normally achieved. Common errors were having just the 2 products 0 and 50 or identifying all products but using combinations and stating the probabilities as $n/9$.

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