Pearson

Examiners' Report
October 2025

International Advanced Level in Statistics S1

WST01/01A

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 2025
Question Paper Log Number P87446A
Publication Code WST01_01A_2510_ER
All the material in this publication is copyright
© Pearson Education Ltd

General Introduction

Overall, this paper allowed all learners to demonstrate their ability and knowledge of the WST01 specification. Learners should be advised to read questions carefully as marks were dropped carelessly for not meeting the required demands. Questions asking learners to prove something require all steps to be shown. Questions that state 'using standardisation' means that the standardisation is required to be shown and if not shown then marks will not be awarded. Learners would be advised to take note of the instructions on the front of the paper in particular the one that says "Inexact answers should be given to three significant figures unless otherwise stated"

Question 1

Part (a) and Part (b)

The numerical parts of the question were generally well answered but premature rounding and answers with an accuracy of less than the required 3 significant figures lost marks in parts (a) and (b). Where comments were required, learners were less successful.

Part (c)

In (c) many did not interpret the correlation coefficient in the context of the question.

Part (d)

In (d) some thought a negative value for the correlation coefficient meant the value did not support the researcher's belief. Learners should note that fractions were not accepted for coefficients of the regression line in (e).

Part (f)

Whilst most answered f(i) correctly by substituting 40 into their regression equation, a large majority lost the mark in f(ii) by failing to realise that this was interpolation or saying 'it' was in the range without specifying age.

Question 2

Part (a) and Part (b)

Most learners answered parts (a) and (b) successfully. There were a few arithmetic errors, but most could form equations for b or c and usually the correct values were found.

Part (c)

Part (c) was generally answered well too although a number wrote down $P(X < 4)$.

Part (d) and Part (e)

Parts (d) and (e) were only answered correctly by the stronger responses and then not always in the most efficient manner.

In part (d) many tried listing all 9 cases and invariably either forgot to include cases like 1, 3 and 3, 1 or they missed the "doubles" such as 1, 1. Finding the probability that X is odd and squaring it was rarely seen.

The major problem in part (e) was that learners did not realise that a conditional probability was required and those that did sometimes missed one of the probabilities on the numerator such as one of the 1, 5 pair or 3, 3.

Question 3

There were accessible marks for all learners on this question with parts (a) and (d) discriminating the most able.

Part (a)

Part (a) was the part of the question that was often answered incorrectly with the most common incorrect answer being 33. Learners did not show understanding that 75% exceed the lower quartile.

Part (b)

Most learners answered part (b) well although there were a few, who, having correctly written down the values of Q_1 & Q_3 then failed to subtract them to find the interquartile range.

Part (c)

There were many correct box plots seen in part (c), with most calculating the correct upper and lower boundaries of 38 and 22 and then usually marking in the two correct outliers. There are still some responses which demonstrated difficulty in interpreting the boundary instructions for identifying the outliers and even some who did the correct calculation but then marked in 2 or 3 outliers at the lower end. It was clear from some responses that learners did not know if 22 itself was an outlier (being on the boundary) and sometimes plotted their minimum value as 23. A significant number of learners drew two upper whiskers and this was penalised.

Part (d)

Most learners demonstrated good knowledge of how to use quartiles to justify the skewness of the distribution. It was very common for learners only to identify that **Eastyou** was symmetrical with correct justification in part (d). They did not read the question properly and said nothing about **Westyou**. A handful spent

time calculating the mean for **Eastyou** and then used this as a justification for (slight) positive skew.

Question 4

Learners who draw sketches tend to perform better on normal distribution questions and this was indeed the case here.

Part (a)

Part (a) was usually very well answered. For those who only scored partial marks, the two main errors were truncating and using $z = 1.3$ instead of the more accurate $z = 1.33$ and also not realising that subtraction from one was required. This would have been more obvious to learners if they had drawn a sketch.

Part (b)

There was a good standard of responses seen in part (b), although some learners lost the final accuracy mark because they did not express their answer correct to the four significant figures that the question demanded. A loss of accuracy was also noted by those who used a rounded standardised z-value of ± 0.25 instead of the value 0.2533 which should be obtained from the percentage points table. A quite common misunderstanding was to use a probability value, particularly 0.4 as standardised z-value to solve for the minimum distance.

Part (c)

Many did not attempt part (c) of the question. Often solutions tended to be jumbled as responses struggled with realising that conditional probability was a key concept. A probability of 0.5 leading to a z-value of 0 was used regularly here.

Part (d)

In part (d), many learners thought the only requirement here was for a jump of greater than 4.1 metres, not realising that it needed a learner to firstly qualify for having a second jump. This meant that there were many final answers of 0.0918 with no attempt to multiply by 0.4 whilst others calculated 0.0918^2. Not many realised they could use their answer from part (a) and hence started again to calculate P($X > 4.1$).

Question 5

Part (a) and Part (b)

This probability question had an accessible start and a challenging end. Some drew a tree diagram to get a better sense of the question, and this generally enabled them to complete parts (a) and (b) correctly. A surprising number of learners evaluated $\frac{1}{20} \times \frac{1}{20}$ as $\frac{1}{40}$ in part (a).

In part (b), not everyone appreciated that there were two possible ways that this could be done.

Part (c)

Most realised that 3 products were required in part (c) though equivalent products were not always given.

Part (d)

Those who realised the most efficient approach to part (d) was to find the probability that all 4 of the beads were yellow and take this away from 1 were most successful. When learners attempted to find all of the different ways that at least

1 green bead could be selected, they often failed to find all of the 15 favourable outcomes and inevitably scored no marks here.

Question 6

Most responses made some initial progress with this question but only the more able scored 10 or more marks here.

Part (a) and Part (b)

A correct expression for E(X) was usually obtained in part (a) but the statement about Var(X) led many learners to find a similar expression, in terms of a and b for E(X^2) and often these two linear equations (both expressions mysteriously set equal to 0) were solved. Those who did find a correct 2nd equation based on the sum of the probabilities equalling 1 usually assumed E(X) = 0 and went on to solve their equations successfully.

Learners should note that where a specific demand is made "write down the value of E(X)" we will expect to see an answer given for this and so a number of learners lost the mark for (b)(i).

Part (c)

Those who had genuine probabilities for a and b (there were plenty with negative values) could usually find Var(X) and multiply this by 9 to answer part (c).

Part (d)

Many learners attempted (d) and often a correct, or correct follow through, answer was seen for (i), however, part (ii) proved more challenging and often the marks were not obtained.

Question 7

This question revealed that some learners do not have a very thorough understanding of the concept of standard deviation.

Part (a)

Almost all learners could find the mean in part (a) but the standard deviation caught many out. Some simply found $\sqrt{63840}$, others attempted to "unpack" $\sum(x-\bar{x})^2$ to find $\sum x^2$ but there were often errors involved. It was encouraging to see some though who realised that $\sigma = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$ and easily achieved the 2 marks.

Part (b)

Most were able to recover in part (b) and used the mean and median to determine positive skewness.

Part (c)

Part (c) was generally answered well suggesting that learners have a confident understanding of the mean and the mechanics of calculating it.

Part (d)

Many left part (d) blank or simply said that the standard deviation would increase or decrease with no attempt at giving a suitable reason. Correct answers based on a consideration of variance as the average deviation squared from the mean were rare but impressive, others were able to calculate the new value of $\sum(x-\bar{x})^2$ and hence show that the standard deviation decreased.