Pearson
Edexcel

Examiners' Report
Principal Examiner Feedback

Summer 2023

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

Summer 2023
Publications Code WST01_01_2306_ER

General introduction

Overall, this paper allowed all students to demonstrate their ability and knowledge of the WST01 specification. Students should be advised to read questions carefully as marks were dropped carelessly for not meeting the required demands. In particular questions that ask a student to show something is true require all the steps in the working to be shown. If asked to use standardisation then the standardisation should be shown. Students 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"

Report on Individual Question

Question 1

Generally, this question was quite well answered.

(a) Many students used the frequency density of 6 from $\frac{60}{10}$ for the 10 - 20 group, and marked this on the histogram. They then found the frequency of the other bars successfully, so found 28 easily.

Otherwise, finding 0.4 people per square or using the ratio of 150 : 60 = 70 : x were the other 2 methods that were used well. Students who used other methods often found 16 but not 12, therefore losing the final mark. Some correctly found the area of the bars between 22 and 45 to be 70, but did not then multiply by 0.4

(b) Most students recognised that they needed the 137th value but for those that did not gain full marks it was because they either went no further or just identified (correctly) that the 137th value was in the 5 - 10 group, but did not go on to use interpolation.

(c) The most common error was to use the class widths in the calculation rather than the mid-point.

Question 2

Parts (a) and (b) were generally well answered. In part (a) the most common mistakes tended to be in the calculation of S_{WW} where students had written $\left(6089.12 - \frac{297.8}{15}\right)^2$ rather than the perfect expression. In part (b) most students were able to gain the method mark. Losses of the

A mark usually occurred due to either having previously obtained an incorrect S_{WW} or giving their answer to 2 decimal places rather than 3 significant figures.

(c) The majority of students gained the two method marks for this part, invariably losing the final A mark because they had not given "b" or later "a" to a greater accuracy than the values given in question. It was rare to see 4.771 or better or -16.66 or better. When a question asks you to show that an equation is true students should be working to a greater degree of accuracy in the calculation process.

(d) Many students knew what was required of them in this question. Their most common error was not to include the units or giving both the length and the weight in centimetres. Only occasionally did a student simply quote gradient as being 4.77

(e) while many students were able to give a generic answer such as " extrapolation" or "the weight would be negative" or "it must be an outlier as it was a long way from the mean" few gave a calculation to back up their statement. The most common correct answer which gained marks was calculating that the weight would be –7.16g for a tail length of 2 cm.

Parts (f) and (g) were very poorly answered. In part (f) many students made no attempt at all while others tried to apply the coding to their answer part (b) or recalculated the value using the coded values of x and y

few students made any attempt to answer part (g). For those that adopted the correct approach, many made an error rearranging the code expression for y to $w = 2y + 5$

Question 3

Part (a) was answered well with nearly all students being able to find $\bar{x}$ correctly. Most students realised they then needed $\bar{x} + 600$

Only a minority of students added 600 before dividing by 81 or left the mean of x as a final answer.

In Part (b) the majority of students were able to gain M1 for finding the variance or the standard deviation correctly. This was often labelled as the variance of x or had no labelling at all. Very few students clearly signposted that it was the variance of L they had found so failed to receive the accuracy mark.

In parts (c) the most common incorrect answer was 40.5 found by dividing 81 by 2. Few students understand the context of the question gave the maximum number of salmon as 40

In part(d) students were well rehearsed in what was required of them. The IQR was nearly always found correctly as well as both limits. A few students lost the final A mark because they

didn't explicitly show a comparison with 7700 and 1600 or give a conclusion that but there were no outliers.

Question 4

Most students were able to score well on this question with a far greater proportion of fully correct solutions seen, than on other questions. As parts (c) and (d) could be completed without using their values in the tree diagram it was not unusual to find students failing to score in part (a) and then go on to obtain full marks in parts (c) and (d).

In part (a) the majority of students were able to complete the tree diagram. For those who made an error the most common was to subtract 0.02 from 0.3 to obtain 0.28 (along with 0.41 and 0.19). It was surprising how many students showed a lack of appreciation that all sets of branches should sum to 1.

Parts (b) and (c) were well answered, even by those students who had incorrect values on their tree diagram.

In part (d) it was pleasing to see that students seemed more comfortable with the concept of conditional probability, and it was rare to find a solution which did not involve a ratio of probabilities. Typical misconceptions involved simply using 0.25 rather than 0.25×0.06 or using $0.12 = 0.02 + 0.04 + 0.06$ instead of their answer to part (c).

Question 5

This question was assessing a students' understanding of a discrete probability distribution.

Parts (a) and (b) were routine questions. The majority of students could correctly form a set of probabilities that sum to one, leading to a fully correct proof in part (a) and in part (b) correctly added P(2) + P(3) + P(4). The most common errors were adding additional probabilities such as P(1) +…+ P(4) or omitting probabilities and finding P(2) + P(3) only.

Part (c) is a fairly standard question on the WST papers where students are using the $\sum y.\mathrm{P}(Y = y)$ to find the expectation. As seen earlier in the question most students demonstrated the correct method and correct solution.

In part (d) most students were able to either form a set of values for X or define an inequality in terms of Y. Students who then realised that they needed to find $\mathrm{P}(Y = 5) + \mathrm{P}(Y = 6)$ generally went on to gain full marks.

Students found part (e) more challenging. The two methods, finding the $\mathrm{Var}(Y)$ and scaling or finding the $\mathrm{Var}(X)$ for the new distribution were used fairly equally. The common incorrect responses were finding $\mathrm{Var}(Y)$ and failing to link this to $\mathrm{Var}(X)$ or incorrectly defining $\mathrm{Var}(X)$ as $\mathrm{E}[X]^2 - \mathrm{E}[X]$.

Question 6

Part(a) was an easy start to the questions with most students able to correctly calculate the required probability. The most common incorrect answer came when students interpreting $\mathrm{P}(A \cup C)$ as $\mathrm{P}(A \text{ and } C)$ leading them to calculate $\mathrm{P}(A) \times P(C)$ instead of adding the two probabilities.

Whilst many students were able to fully justify the given result in part (b), one of the main errors seen was when they did not provide the necessary amount of detail required for a proof which is essential in any 'show that' question. The first method mark required students to explicitly state that $0.3 \times 0.1 = 0.03$ or write the expression $\dfrac{\mathrm{P}(A \cap B)}{0.1} = 0.3$ and not just to state that $\mathrm{P}(A \cap B) = 0.03$ as this was often gained from the given answer.

Attempts at this part (c) of the question were extremely varied. Whilst a number of students produced completely correct Venn diagrams, some drew their circles incorrectly. The most common error was not having all three circles intersecting (often the circle representing *C* was drawn separately, or inside the circle representing *A*). Other common errors included: omission of a single probability (usually the outer section) or leaving sections blank instead of inserting zeros in the version of the diagram where all three circles intersected each other.

The more common diagram used was the one where the three circles had intersections of *A* with *B* and *B* with *C* which and this often scored higher than those who used 3 intersecting circles.

Question 7

This question demanded careful reading and accurate working.

Part (a)(i) was a show that question and those students who heeded the advice given in the stem "use standardisation" usually went on to gain full marks. As the answer is given simply writing the given answer after the standardisation will not gain full marks. An extra step is needed whether tables or a calculator were used. Those who used tables were generally more successful in showing the appropriate working. Those who used calculators often just stated the given answer. In (ii) many students had the wrong sign on their *z* score.

Part (b) was found to be challenging for many students with many calculating $5 \times p$ or $\frac{p}{5}$ rather than p^5. Those who used the correct method almost always chose the correct value of p.

In part(c) the two most common errors were not using the accurate z score from the points table or using the wrong sign with the z value. It is fine to use a calculator to find the z value but it must have the same or greater degree of accuracy than the value given in the tables.

It was pleasing to see that students rarely equated their standardisation to a probability rather than a z value.

Few students sketch the normal distribution and would be advised to as those who do generally avoid mistakes of ‘sign’ in their equations.

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