Examiners' Report

Principal Examiner Feedback

January2020

Pearson Edexcel International GCE In Statistics Mathematics 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

Grade Boundaries

Grade boundaries for all papers can be found on the website at:
https://qualifications.pearson.com/en/support/support-topics/results-certification/grade-boundaries.html

January 2020
Publications Code WST01_01_2001_ER*

General

This was an accessible paper with opportunities for most students to get started in all the questions apart from possibly question 1. It was encouraging to see that some typical types of questions that, in the past, were answered poorly were being tackled more successfully this time: test of independence in 2(d), histograms in 4(a) and standard deviation in 4(d). The questions requiring comments are understandably still quite challenging but again it is encouraging to see fewer "blanks" here than we have encountered in the past.

Report on Individual Questions

Question 1

This short starter question had no structure and many students struggled to find enough equations to solve it fully though over 40% did gain full marks. A common starting point was to write down the equation $-a+b+3c=0.7$, based on E(X) = 1, which was not especially helpful. Of course a little thought would have shown them that E(X) = 1 means that the distribution is symmetric and therefore $a = c$ a result that most successful candidates had to derive from solving their 3 linear equations. The most common error was failing to understand what F(1) = 0.63 meant and realise that F(1) = 0.15 + a + b. Many failed to identify F as the cumulative distribution function but a few realised that if F(1) = 0.63 then c + 0.15 = 0.37 and they could then write down the value of c. A large number of candidates failed to identify the basic equation $a+b+c=0.7$, from using the sum of probabilities, and so they were always one equation short when trying to find their values. It is disappointing that some students are still happy to leave answers for probabilities that are either negative or greater than 1 but this may simply be a recognition of the fact that they have made an error and do not have time or inclination to find out where.

Question 2

Nearly all students were able to score some marks here and as we followed through answers from their Venn diagrams even those who struggled to complete part (a) correctly were able to pick up marks in the later parts of the question. Only around 17% though were able to score full marks.

In part (a) many went straight to a diagram with 3 intersecting circles showing the usual 8 regions but they failed to place zeros in the regions giving $D \cap R$ and so lost the first mark. Those who incorporated this feature into their diagram and gave a "chain" of 3 intersecting circles often found completing the diagram more straightforward. The box and the 9 were usually placed correctly and many had the 10 and 2 correct as well but the other three values caused problems for a number of students. Some had the R circle on its own and there were a number who had a total in their Venn diagram that did not equal 40. In part (b) most knew what mutually exclusive meant but some failed to give a reason and others did not describe the events (D and R) but simply wrote down "dog" and "rabbit" or said that P(D) and P(R) were the mutually exclusive events. A follow through mark in part (c) meant that most

students were able to secure this mark but some failed to add together their two values. The responses to part (d) were encouraging. Most knew what to do and most labelled the relevant probabilities correctly with the most common approach being to show that $\mathrm{P}(D) \times \mathrm{P}(C) \neq \mathrm{P}(D \cap C)$. There were some who tried using an incorrect formula, $\mathrm{P}(D \cup C)$ being a common one, and a few able students who successfully used a conditional probability approach. In part (e) a good number identified that a conditional probability was required and some just scored the method mark for writing the ratio $\frac{\mathrm{P}(R \cap C)}{\mathrm{P}(C)}$ and with a follow through accuracy mark available many were able to score both marks here. Many interpreted part (f) correctly but common errors involved including too many values on their numerator with $\frac{23}{26}$ being a typical incorrect answer.

Question 3

Apart from the last two marks this question was answered well by the majority of students. The mark in part (a) was rarely lost as all 4 numbers were seen correctly placed in the expression. Most went on to write down a correct expression for r in part (b) but some still persist in giving their answer to fewer than 3significant figures. This has been mentioned many times in these reports and is clearly stated in the instructions on the front cover. Responses to part (c) were slightly better than in previous series but some simply stated that the correlation was positive rather than commenting on the "size" of the correlation coefficient or the scatter diagram. Part (d) was a bit different and many simply tried using a single point from the scatter diagram rather than the required values of Σw and Σp given in the question. The formula for b was usually correctly used in part (e) and most knew how to find the intercept but it was common to see a rounded value for b being used and this meant that the candidates were unable to present their final equation with coefficients correct to 3 significant figures as required. A small accuracy error here though usually didn't prevent candidates from achieving the mark in part (f). The final part of the question required some careful thought and engagement with the context. There were a number of attempts to find an estimate of the number of guarantee policies sold under the manager's model but fewer were able to reason that a lower estimated number would give the staff a greater chance of receiving the bonus and therefore the manager's model should be chosen.

Question 4

This question proved very accessible with most students once again scoring some marks and nearly half scoring 10 or more. In part (a) most found the correct width but the height still proved challenging for some although there were more successful attempts here than we have sometimes seen. Most students knew how to carry out the linear interpolation in part (b) but a number lost marks because they rounded up $\frac{98}{4} = 24.5$ to 25. Part (c) was answered quite well with most carrying out the correct comparison but a number concluding that this suggested the skewness was positive not negative. Some may have been thrown by part (e);

the question was designed to illustrate that these "tests" for skewness are only indicators and not hard and fast rules so this question may form a useful basis for discussion of this topic in class. Most answered part (d) (i) correctly and many had a correct formula in (ii) although forgetting the square root or using S_{ww} as the variance were common errors. A number of candidates lost the accuracy mark because they used their rounded value for the mean rather than the full value on their calculator. A small number did not use the values for Σfx and Σfx^2 given in the question and invariably made errors in attempting to calculate these. Part (e) was answered very well and even those with only minor accuracy errors in part (d) were able to gain the accuracy mark here too. Part (f) proved very challenging and many simply omitted it. A few identified that the median would now be 3.25 thus representing an increase but very few explained that Σfx would increase with many simply commenting that the mid-point had increased without explaining that because there were more than 9 in the upper group this meant that Σfx increased and therefore the estimate of the mean would increase.

Question 5

Part (a) was a familiar style of question and it was encouraging to see around 80% of the candidates securing all 3 marks here. Most chose to standardise the value of 7 although a few preferred to use 13 and most were successful. There are still a number of students who are unsure what to do with the probability from the tables (we have said many times before that a small diagram would probably help them) and some left the answer as 0.6915 rather than 1 – 0.6915 as required. Part (b) required the students to "work backwards" but it was not quite as straightforward as sometimes. The first step was to realise that $P(X < 10 + k) = 0.80$ and then the problem was fairly standard. The most common value of z used was 0.84 though and a number of students seem to be unaware of the percentage points tables which would give the more accurate value of 0.8416. Of course some weaker students set their standardised expression equal to a probability rather than a z value and this scored no marks.

Part (c) proved to be a very effective discriminator allowing most students a chance to make a start but only around 10% were able to complete it successfully. Many were able to write down an expression for the area of the rectangle and often they went on to use the given value of 40 and solve a suitable quadratic equation correctly. Some never seemed to appreciate that they were trying to solve a quadratic inequality and others were convinced that X could not be negative and so it was rare to see the correct pair of probabilities being attempted. A number did try and find $P(X > 8)$, and scored a mark for attempting to standardise, but finding the two correct probabilities of 0.6293 and 0.0062 and then adding these together was only achieved by the best students.

Question 6

Despite there only being 6 questions some students seemed to be running out of time or stamina at this point. Fortunately the first two parts were very accessible and most attempted these but then their solutions quickly faded out. Parts (a) and (b) were answered very well and most scored full marks here. The rest of the question required a careful grasp of the variables being considered. In part (c) many simply used E(S) rather than considering Serena's profit. Those who did realise that 260P($S = 5$) was required often forgot the entry fee and failed to subtract 10. In part (d) it appeared that some students were not familiar with the idea of a "knock-out" tournament, however there were plenty of good explanations for the p^2 though some did not make it clear that Roger must win the first two rounds and then lose the third match and did not secure the 2nd mark for the $(1 - p)$ term. Some used a tree diagram but without labels or a written commentary this was not clear enough to secure the marks. Those who persevered to part (e) would often score the B1 mark for giving the correct values (though 0 was sometimes omitted) and usually there were at least 3 correct probabilities but P($R = 0$) was often given as just p and P($R = 5$) was often given as $p^5(1 - p)$. The final part was not popular with students or the examiners! Some gained a couple of marks for considering Roger's expected profit using the same approach as in part (c) but a number went straight to comparing P($R = 5$) with P($S = 5$). Many simply equated these but a few used P($R = 5$) > P($S = 5$) with very few using $\geqslant$. A final answer of 0.59 was rarely seen, and those using equality would lose the final mark for an answer of 0.58, but a small number of able students did complete the question correctly.

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