Pearson Edexcel

Examiners' Report
Principal Examiner Feedback

January 2022

Pearson Edexcel International A Level

Mathematics in Statistics S2 (WST02)

Paper : WST02/01

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

January 2022
Publications Code WST02_01_2201_ER

General

Students were generally well prepared for the demands of this paper with many strong performances seen. Question 7 proved to be the most challenging and only the most able were able to access this. There were several places where students struggled to translate the context into correct statistical processes/calculations.

Report on individual questions

Question 1

This question was answered well by the vast majority of students and provided an accessible start to the paper.

Part (a) This was well answered with many students attaining full marks; however the common error was to use P $(X \leqslant x)$ rather than P $(X \leqslant x - 1)$ which resulted in an incorrect answer of 8

Part (b) The vast majority of students were able to write or use the correct normal distribution of N(36, 36). A few then went on to standardise with an incorrect standard deviation, usually using their variance in the standardisation calculation. Others lost marks due to an incorrect use of the continuity correction or not using a continuity correction at all.

Part (c) Again the vast majority of students were able to write or use the correct binomial distribution of B(50, 0.1528). The common error in this part was to use $1 - \mathrm{P}(W = 0)$ rather than $1 - \mathrm{P}(W \leqslant 1)$ which resulted in the loss of the last 3 marks.

Question 2

Whilst this question on the whole was answered well by the majority of students, many failed to realise that they were dealing with a uniform distribution and so parts (b) and (c) were often attempted by integration rather than using the formulae associated with this distribution.

Part (a) Many students were able to write down the correct pdf. Those that lost mark usually failed to give a fully correct pdf, often with the '0 otherwise' missing.

Part (b) The vast majority of students were able to state that E(X) = k but many took an integration approach which not only took longer but increased the risk of a numerical error.

Part (c) This was a "show that" question which was well answered. Those students that used $\frac{(\beta-\alpha)^2}{12}$ usually gained full marks. Again too many students took an integration approach which again not only took longer but increased the risk of a numerical mistake.

Part (d) Although there were a substantial number of correct solutions, there were quite a few errors. Those that took the $E(X^2) = Var(X) + E(X)^2$ route occasionally subtracted $E(X)^2$ rather than adding $E(X)^2$, whilst a few simply added E(X) rather than $E(X)^2$. Those that took an integration approach could of course use their integration from part (c), however if it was incorrect led to an incorrect answer in part (d). Some students that had used integration in part (c) started again and repeated the integration in part (d) which again increased the chances of numerical errors.

Question 3

Part (a) This was answered well by the vast majority of students and it is clear that students are well rehearsed in stating the conditions needed for a Poisson distribution.

Part (b) Many students stated the correct two tailed hypotheses needed. Common errors were to use a one tailed test, often H_1: $\lambda < 8$ or to use p instead of λ.

Part (c) Many students were able to identify at least one of the required critical regions, usually $X \leqslant 2$.

Those that showed their working often gained at least 2 marks in this part of the question for showing the correct probabilities needed as stated in the mark scheme. Common errors included giving an upper critical region of $X \geqslant 14$ rather than $X \geqslant 15$ or writing their critical regions as probability statements.

Part (d) This was answered well by the vast majority of students with many scoring full marks. Even those students with incorrect critical regions were able to follow through and give a correct answer.

Part (e) Many students were able to gain the M mark for writing a non-contextual comment that followed through from their critical region, however the correct conclusion in context was more problematic with some students not even attempting this. Those that did, often lost the mark as they failed to refer to the rate of breakdowns or they wrote statements that referred to the Headteachers claim (the Headteacher had not made a claim in this question).

Question 4

This was a long and a challenging question for some students but those students who laid out their work clearly and communicated their reasoning methodically were most successful.

Part (a) This was answered well by many students who were able to give a fully correct sketch of the pdf. A few students lost marks due to missing labels of the key points. Others lost marks as they started their sketch at the origin or drew sketches that did not start or end on the x axis.

Part (b) A mixture of methods were seen. Those that took an area route, either for a trapezium or two triangles and a rectangle usually did so successfully. Those that took an integration route were able to write down accurately the required integration but an occasional error was seen in the integration. As this was a "show that" question the final A mark was not awarded.

Part (c) It is reassuring that many perfect solutions were seen for this type of question. There were the usual slips with the algebra and the arithmetic. The most common problem was the parts of the distribution dealing with $3 < x \leqslant 6$ and $6 < x \leqslant 10$ where many students simply worked out $\int_3^x \frac{1}{6}\,\mathrm{d}x$ and failed to realise that they needed to add F(3) and worked out $\int_6^x \left(\frac{5}{12} - \frac{1}{24}x\right)\mathrm{d}x$ and failed to realise that they needed to add F(6). The later was less of an issue for those students who took an indefinite integration route as they usually integrated and used F(10) = 1 correctly.

Part (d) As the correct answer related part (c) many students scored the M mark in this question. Students should be encouraged to show their working as this question required the use of $1 - \mathrm{F}\left(\frac{61}{12}\right)$ and simply writing this gained no credit. Whilst examiners checked that the answer given came from using the students 3rd line of the cdf it would be advisable to encourage students to show their working so it is clear in what they are doing. The common error was to use the 4th line of the cdf instead of the 3rd line.

Part (e) This proved to be more problematic for students. Those that gained marks usually did so by comparing part (d) with 0.5. Others correctly calculate the median and compared this to the mean. However some students then went on to give an incorrect answer (negative skew). Too many students tried to compare the mode to mean/median and lost both marks. Very few students used the sketch in part (a) to answer this question.

Question 5

Part (a) This was answered well by most students as they could identify the appropriate distribution of B(n, 0.045). Common error was to write Po($0.045n$)

Part (b) Quite often this mark was not awarded due to the lack of context needed for this mark. 'Independent' on its own without reference to 'applicants' was often seen.

Part (c) This was answered well by many students and full marks was often awarded.

Part (d) Many students were able to state that n needed to be large and p needed to be small. However too many assumed that $np < 10$ was sufficient for the 2 marks and usually scored no marks.

Part (e) This part of the question proved to be more difficult than the other parts. Many students were able to define the correct hypotheses. Common errors included either setting up a one tailed test ($H_1 : p < 0.75$) or writing the hypotheses in terms of μ rather than p. Many students defined the correct distribution of N(72, 18) but the common error was to write or use N(72,72). Most students knew that they needed to apply a continuity correction but too many standardised with 67 or 66.5. Those that did not use or used an incorrect continuity correct scored no further mark in this question. The students that used the correct continuity correction usually scored the next 2 A marks. At this point many students were able to give a non-contextual conclusion and a conclusion in context. Some students that gave a correct non-contextual statement went on to state 'Jaymini's claim was incorrect' and so lost the final A mark. A few students took a critical region approach to this question and generally they were less successful than those who took a probability approach. Only a few students attempted the alternative approach of using the probability that an applicant failed to become a pilot.

Question 6

Part (a) This was poorly done with very few students scoring the one mark. Those students who had learnt the standard definition of a sampling distribution fared better than those who used their own understanding of the term because they were less likely to leave out vital elements of the definition such as 'all'.

Part (b) It was pleasing to see so many completely correct solutions to this part of the question. At the other extreme, there were a small number of students who made no progress beyond establishing the probabilities of each size of breakfast separately. Then in the middle there were students who adopted a correct strategy but failed to master all the details. The biggest single problem was P(T = 14) where students had to find 2 × SL and MM size breakfasts. Some students worked out these two probabilities but did not add the probabilities together. Some students omitted the 2 × when finding P(T = 13) and P(T

= 15). Some students did not work out the totals when presenting their probability distribution. Generally, the students who obtained full marks, achieved this by being more efficiently than others. Students are encouraged to show their methods in how they obtain their probabilities and should check their probabilities to see if they add up to 1.

Part (c) Many students were able to score at least the M mark as they knew how to calculate E(T) even if they had lost marks in part (b). Some students used a correct method to find the mean and then divided their working by 5 losing both marks.

Question 7

This question was not answered well by the vast majority of students and only the most able students were able to score full marks, but, this was rarely seen.

Part (a) Those that knew how to answer this question often scored both marks. However too many failed to realise that $P(L \geqslant 4.5)$ implied $P(A \geqslant 20.25)$ and so score no marks. Common error was to assume that L was distributed $\mathrm{U}(\sqrt{10}, \sqrt{30})$ which led to writing $\frac{\sqrt{30} - 4.5}{\sqrt{30} - \sqrt{10}} = 0.422$ which is an incorrect method.

Part (b) This was problematic for students with the vast majority of students not knowing where to start. A few students scored B1 as they recognised that they needed to work with the variable A and were able to state that $E(L^2) = 20$. A few students were then able to score the final M mark as they stated a value for E(L) and correctly used $\mathrm{Var}(L) = E(L^2) - E(L)^2$ ft their E(L). A few students that knew how to answer this part occasionally lost the final A mark due to rounding early when finding E(L). Again the common error here was to assume that L was distributed $\mathrm{U}(\sqrt{10}, \sqrt{30})$ and $\frac{1}{12}\left(\sqrt{30} - \sqrt{20}\right)^2 = 0.447$ was often seen as an incorrect method.

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