Examiners' Report

Principal Examiner Feedback

October 2021

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

October 2021
Publications Code WST02_01_2110_ER

General

Overall, students were well prepared for this paper and, in most cases, were able to attempt all questions. Some students found it challenging to turn contextualised problems into correct statistical processes/calculations in parts 1(c), 2(ii), 4(e) and 5(b). Most persevered and good attempts were made at the final question on the paper.

Report on Individual Questions

Question 1

(a) This was a strong start for many with virtually all students finding the correct probability from the Binomial distribution. Most gave the answer to the required level of accuracy, though on the rare occasion answers were given to 2 significant figures.

(b) More errors were seen here as some students struggled to deal with the inequality, the most common being $P(F \leqslant 14) - P(F \leqslant 8)$. Another common incorrect method was to find $P(F \leqslant 14) + P(F \geqslant 8)$.

(c) This part was the most difficult part of this question with some students making little or no progress. Some tried to find $P(F < 70)$. Others were able to proceed to find $F > 10$ but made an error by then calculating $1 - P(F \leqslant 9)$.

(d) The final part of the question was accessible to students at all levels with many complete correct responses seen. The hypotheses were nearly always correct using correct notation. The Normal distribution was generally correct, though on some occasions the variance was also found to be 70, apparently confusing this with a Poisson approximation. Though most went on to attempt a continuity correction, 86.5 was commonly used. The vast majority of students completed the hypothesis test in context by concluding that the manager's belief was supported.

Question 2

Students displayed a sound knowledge of the uniform distribution in this question, but succinct solutions were not commonly seen.

(i)

(a) A quick sketch would have aided many in this part as students would have been able to use symmetry to find the required probability. Many responses saw students complete many lines of working before arriving at the correct answer.

(b) Those who were able to work out the value of a and b made good progress in this part of the question. Nearly all used the fact that $\frac{a+b}{2}=11$ to make some progress.

(ii)

(a) Students struggled with the context here and a large number wrote down $S \sim U[0, 45]$ not realising that the shorter piece couldn't be longer than 22.5.

(b) Method marks were generally scored here for a correct follow through expression for $P(S < 12)$ from their distribution in part (a).

(c) The probability calculated in part (b) was almost always used correctly here with most earning method marks. Some used rounded values and did not achieve the required accuracy. On rare occasions other distributions were used, such as Poisson, and hence scored no marks.

Question 3

(a) The first part of this question proved to be the most challenging part for many. There were many longwinded attempts including integration/differentiation before some realised the need to equate the two lines of the cumulative distribution function at $x = 1$.

(b) A better attempt at part (b) was seen as most were more comfortable using the fact that the cdf = 1 at the highest value of x. Those who attempted differentiation here often made no progress.

(c) Some interesting attempts at part (c) were seen with some opting to differentiate and then integrate between $x = 2.25$ and $x = 3$. On the whole, this was generally well attempted.

(d)(i) Most students realised the need to differentiate to sketch the probability density function (and often this had already been done earlier in the question). Sketches varied, but, of those who attempted it, most gave the correct shape. On some occasions sketch was incorrectly drawn with a mode at $x = 1$.

(ii) Virtually all students arriving at the second derivative correctly equated it to zero and found the correct mode, $x = 2$. Some students just wrote down $x = 2$ without justification and this was unable to earn marks as the question required full supporting working to be shown.

Question 4

Though good attempts were made at earlier parts of this question, question 4 proved to be one of the most discriminating questions on the paper, with part (e) discerning the most able of all students.

(a) Most used their calculators or the tables to find the correct probability in this part. Of those who attempted to use the formula some mixed up the 8 and the 6.
(b) A correct probability statement was often seen in this part earning the method mark. Most common mistakes were to not apply $n - 1$ and answers of 10 or 12 were common.

(c) This part was more problematic with many believing $m = 3$ (rather than realising that $\lambda = 3$).

(d) Most students were able to identify that Po(3) was the correct distribution here, but many went on to calculate $P(Y = 1)$ or $P(Y > 1)$.

(e) This one of the most discriminating parts of the paper. As is often the case, the conditional probability was not identified and the majority of students simply calculated $P(W = 18)$. Some believed that independence applied and calculated $\dfrac{P(Y = 1) \times P(W = 18)}{P(W = 18)}$.

(f) For those who persevered in this question, part (f) was well attempted and answered. Some carried out the entire test, rather than stating the critical region. Other errors included using Po(6) or more commonly, finding a two-tailed critical region from Po(9). Some left this part blank perhaps giving up after struggling with part (e).

Question 5

Some very strong performances were seen here with nearly half of students scoring full marks. As is usual with sampling distributions, a fair number also make no progress with nearly a quarter of students scoring 0 marks.

(a) Of those attempting this part, virtually all scored the mark for showing the correct multiplication of probabilities.

(b) Those who read the question carefully made good progress here as they understood how Alif's score was calculated. A surprising number of students attempted the sampling distribution of the sum of the two scores without taking into account the 4 and the 2. Those who listed all possible outcomes generally grouped them by score and successfully calculated the associated probabilities. Common errors saw students omitting one of the scores or not checking that their probabilities summed to 1.

(c) Most went on to find the expected value from their distribution in part (b). Some did not fully understand what this meant as the mode was sometimes stated as the answer and on rarer occasions the arithmetic mean of the scores was given.

Question 6

This whole question was accessible to many students with parts (e) and (f) serving as discriminators at the top end. Roughly 10 percent of students went on to achieve full marks here.

(a) The graph was generally clearly drawn with relevant scales on the axes. Some responses incorrectly had the graph connected to the horizontal axis whilst others showed no labels at all.

(b) Many did not use the big hint from the graph that the distribution was symmetric and very long, and often inaccurate, attempts to find the mean through integration were seen. Even more surprising was a lack of understanding of how to calculate the variance, with $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \mathrm{E}(Y)$ often seen. Another common error was to state $\mathrm{Var}(2Y - 3) = 2\mathrm{Var}(Y)$

(c) Though most students made an accurate attempt here, some did not give sufficient detail skipping an important line of working leading to the given answer. As the answer was given, it is important for students to show all stages of the working to arrive at this line of the cumulative distribution function.

(d) Many good attempts were seen here and a generally good understanding of the relationship between the probability density function and the cumulative frequency distribution was shown. Obtaining the third line was the most challenging part and the most common error was to see $\mathrm{F}(y) = \int_1^t \frac{3}{14}\,\mathrm{d}y$ without including the area up to 1 (adding F(1)). Many slips in notation were seen with students swapping x and y frequently and hence losing the final mark.

(e) Most students were able to recognise that the third line of the F(y) was needed to find the 30th percentile. Some did not follow the instruction to give the exact value and lost the accuracy mark. Others left this part blank indicating a lack of familiarity with the vocabulary of percentiles.

(f) One of the most challenging parts of the paper for many. Although the denominator of F(3) was often seen correct, many did not know how to deal and the numerator. F(3) – F(5/4) was often seen.

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