Examiners' Report
Principal Examiner Feedback

October 2024

Pearson Edexcel International Advanced Level
In Statistics S2 (WST02) 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

October 2024
Publications Code WST02_01_2410_ER

WST02 PE Report October 2024

General Introduction

Overall, this paper worked as expected with students of all ability levels able to demonstrate their knowledge. Good attempts were seen at all questions with higher achieving students using correct mathematical notation and answering questions in context.

Report on Individual Questions

Question 1

Part (a) was routine start to the paper. Many fully correct responses were seen, but some students incorrectly used a cumulative probability.

Part (b) saw mixed success. Those who identified that a Po(6) distribution was required here usually answered this part correctly. Many thought a B(5, 0.3) was appropriate, but these were still able to gain a method mark by writing a correct probability statement.

Again in part (c), there was mixed success. Those who identified Po(56) generally went on to answer this part correctly using a correct continuity correction. Many tried to approximate a binomial distribution and these were able to access method marks for standardising and for attempting an appropriate continuity correction.

Question 2

Most students answered part (a) correctly including the correct parameters for the Binomial distribution. The most common reason that students did not earn the mark here was to write ‘binomial’ on its own. A few U[0,25] responses were seen, and an occasional Po(5).

Part (b)(i) was generally well answered for the majority of students. Of correct answers to part (ii) most opted for the main scheme, but a significant number of students used the Special Case method. Some students failed to state E(M) and lost the accuracy mark. Some students found E(X) explicitly but failed to use expectation in the other statements. Of incorrect answers the most common was to solve $5X - 25 = 0$ and state $X = 5$. Less common but still significant was $5\times5 - 25 = 0$. These answers usually had no expectation statement at any stage.

Many correct answers were seen in part (c) and even those who made a slip with the inequality were still able to earn the second method mark here.

Part (d) was again well answered. Some students used an $(n - 1)$ method and scored in the Special Case. A few students used the Normal distribution after a correct initial statement. There were a noticeable number of non-responses to this part.

Question 3

Part (a) was generally answered well. The big clue in the question, "enter at a mean **rate** of 7 every 10 minutes" was missed by a few students; some suggested binomial. Most students knew this was a Poisson but some failed to give the parameter.

Most students knew the conditions for Poisson distribution to be an appropriate model in part (b) and were able to give two different conditions although the context was not always present. A few students said there was a constant probability rather than rate. Many students had the right idea and wrote *singly* or *independently* or *randomly* or a *constant rate*, but it should be stressed that context is key here.

Part (c) was usually correct although some students didn't use a letter at all in their hypotheses and some used x or p. Almost all got the alternative hypothesis as $\lambda \neq 7$ rather than an inequality.

In part (d), most students knew how to find a critical region and made a good attempt at this. The most common error was to use 0.03 as the probability for each tail region rather than 0.015. Some students went straight from the general statements for the lower and upper tails to the critical region, giving no specific probabilities. Most students identified that the test was two-tailed but there were a few who only found the upper tail.

Most students know the difference between the target significance value and the actual one and most followed through correctly from their critical region to score both marks in part (e).

In part (f), most students stated that 12 was not in the critical region. Some however, tried to calculate P($X \geq 12$) and therefore did not make any statement about 12 and their critical region. The context was not always given or was not specific enough (simply mentioning that there was no evidence that the rate was unchanged was not enough). Some students mentioned a claim being accepted or refuted although no claim had been made in the question. Of those not scoring both marks it was usually that there was no context.

Question 4

Part (a) many students were able to set up 2 correct equations and was often implied by a = 12 and b = 72. Few students went about it in a long and convoluting manner by expanding $\frac{(b-a)^2}{12} = 300$ and then solving the 3 term quadratic.

Part (b) was answered particularly well and most students scored full marks. Method marks were generally scored here for a correct follow through for their a and their b values. A few students lost the accuracy mark due to the usual slips with the algebra and the arithmetic.

In part (c) although there were a substantial number of correct solutions, there were quite a few errors. Some students used the part (a), a = 12 and b = 72, values in the calculation. In this part three different approaches were seen with many students obtaining a correct answer. The majority used $\frac{L}{4} - \left(\frac{42-L}{4}\right) > 2$ approach with L~U[21, 42] and obtained the correct answer $\frac{17}{21}$ or equivalent. A common incorrect method was using $\frac{L}{4} - \left(\frac{42-L}{4}\right) > 2$ with L~U[0, 42] arriving at $\frac{17}{42}$.

Question 5

Most students chose to use a + C method in showing the given result in part (a). There were occasional sign errors in finding the constant which were then forced into the correct form. There were many elegant and concise responses with the first method, though a noticeable number used poor notation with x rather than different variable. This was condoned. It was common for students to expand the bracket for integration rather than factorising.

It was pleasing to see so many complete correct answers to part (b) of this question. Again a + C method was used more often. There were occasional errors evaluating F(2) or F(3), or in using a correct evaluation accurately in finding F(x). A few students did not find or use F(2) or F(3) and therefore unable to access any of the method marks. A significant minority of students lost the B mark by omitting both the lines, or omitting F(x) = 0 for $x < 1$ or F(x) = 0 for $x > 4$.

When answered, part (c) was usually accurate. Most students substituted in their expressions for F(x) but a few evaluated the integration from scratch. A number of students substituted 3.1 into F(x) for the interval (2,3] which gave an answer very close to the correct answer, but obviously did not score for the correct method.

Question 6

In part (a) most students gained full marks or very few. Virtually all students were able to go from the ratios given for bags A and B to correct probabilities. The most common error was to not realise that the combinations 1,2,5 and 2,2,5 each gave rise to two probabilities. Other mistakes were to multiply each probability by 2 or 3 or even 6. A few students got the correct probabilities but didn't give the totals, thus missing out on the 2nd B mark. Although not often stated explicitly, most students used the correct probability for picking each coin. Most students wrote their sampling distribution in a table which was helpful, especially if the table was drawn up first with the working following underneath.

On the whole a good attempt was made in part (b). A few students worked out the mean values rather than the medians. Those that did use the median correctly identified the only two values (occasionally 1 was given with probability 0 which was acceptable). Although this was well answered in general, the usual errors for this question surfaced. Failing to recognise that there were only two possible values for the median was most common, and assuming that the three coins could be freely picked from either bag meant that medians of impossible combinations were calculated, e.g. median of (1, 1, 1) = 1. Some students managed to earn the second M by find their P(M = 5) from 1 – 'P(M = 2)'.

Question 7

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

Part (a) 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 work. Most students worked with the area under the curve $\frac{1}{2} \times 8 \times 4a = 1$, with only a minority using the integration $\int_0^4 ax\,dx = 0.5$ to obtain $a = \frac{1}{16}$.

In part (b) many did not use the big hint from the graph that the distribution was symmetric to write the value of $b = -\frac{1}{16}$. Instead some used simultaneous equations to work out the value of b and c with occasional arithmetic slips seen.

In part (c) again, many fail to use the graph that the distribution was symmetric to identify $\mathrm{E}(X) = 4$. Very long, and often inaccurate, attempts to find the mean through integration were seen. Students should be encouraged to show all their working including the substitution of limits in questions like these that stated, 'Use algebraic integration to find …'. Many relied on the calculator for too many steps. Since the answer is given, we need to see use of $\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{8}{3}$ with values substituted for the final accuracy mark.

In part(d) many strong attempts were made, but only the most able students went on to complete the question with the correct values for lower quartile and upper quartile. Some took many stages of working to get to the upper quartile without using the given symmetrical distribution.

In part (e) many correct responses were seen, mostly using the correct probability calculated awrt 0.65. Other students successfully used the alternative approach of comparing ‘2.83’ > ‘2.37’ and ‘5.16’ < ‘5.63’. Some attempted normal distribution calculations whilst others left this part blank.

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