Examiners' Report

January 2015

Pearson Edexcel International Advanced Level in Statistics Mathematics S2 (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 students at: www.pearson.com/uk

January 2015
Publications Code IA040681

General introduction

Overall, candidates were well prepared for this paper and, in most cases, were able to attempt all questions. Candidates should be encouraged to write down all stages of their working even when using a calculator or the statistical tables. Some method marks in longer, multi-step questions (5a and 7) were lost due to insufficient working being shown. Questions 1(c), 3(e), 6(d) and 7 were the most demanding parts of this paper.

Report on Individual Questions

Question 1

The opening question on this paper was both accessible, (parts (a) and (d)), and discriminating, (parts (b) and (c)).

Generally candidates answered part (a) well. Most recognised that they needed to use Po(3.2) and knew to calculate P(X = 3). A significant minority of candidates gave an answer of 0.22 so lost the final accuracy mark. Some candidates were able to identify Po(1.6) in part (b) but even then many were unable to calculate the required probability as P($Y \geq 1$). The most common mistake in part (b) was to calculate P(Y = 1).

Part (c) was the most challenging part of this question as most candidates failed to understand that a conditional probability was required here. Many were able to calculate P(X = 1) and P(X = 3) and multiply these together, though some did so with inconsistent values of lambda. Of those candidates who gave a correct expression for the conditional probability, some lost the final accuracy mark due to premature rounding.

Many were able to identify the correct normal approximation in part (d), N(72, 72), and work out that the number of cars caught speeding had to exceed $83\frac{1}{3}$. The continuity correction here caused some confusion and a variety of attempts were seen, notably 82.8, 83.8 or 84.5. A few took the route of finding a normal approximation for the amount of fines collected in 90 days and incorrectly used N(4320, 4320).

Question 2

Candidates had a lot of success with this question as nearly 50% achieved at least 10 of the 11 available marks here. Part (a) was generally well answered though on some occasions candidates only found F(4) or misunderstood P($X > 4$) to be equivalent to P($X \leq 3$) and went on to find 1 – F(3). Correct answers to part (b) were seen less often. Here candidates often gave one of two common incorrect answers, either 0, by finding P(X=4), or $\frac{4}{5}$, by confusing the continuous distribution for a discrete one. Part (c) caused little difficulty and the majority of candidates correctly specified the density function with the required limits.

Those who recognised that this was a continuous uniform distribution often went on to give efficient solutions to parts (d), (e) and (f). Many relied on integration to get them through these parts and had a high success rate in parts (d) and (e). In part (f), the most common error was for candidates to use $E(X)^2$ rather than $[E(X)]^2$. Some thought the formula $\text{Var}(aX + b) = a^2\text{Var}(X)$ applied here and attempted $9[E(X)]^2$.

Question 3

There were some good attempts at this question despite candidates finding the sampling distribution a challenging topic on the specification. It was clear that some candidates had made an effort to learn their definitions and a pleasing number of correct responses in parts (a) and (b) were seen. Sometimes a keyword was omitted from a definition- e.g. ‘only’ in part (a) and ‘statistic’ in part (b). A significant proportion of candidates simply left these parts blank.

In part (c) the calculation of the mean and the variance was well done. Use of a rounded value for the mean sometimes resulted in the loss of the final accuracy mark for the variance. Nearly all candidates came up with at least 2 of the required samples and most went on to give all possible samples in part (d). The final and most challenging part of the question had a mixed response. Some oversimplified the question by assigning a probability of $\frac{1}{8}$ to each possible sample. Those who did calculate the 4 required probabilities correctly sometimes neglected to give the 4 associated values of the means.

Question 4

Nearly $\frac{1}{3}$ of candidates scored full marks on this question. In part (a), many recognised that they needed to use Po(6), but were unable to correctly interpret the inequality as F(6) – F(4). Part (b) elicited a good response with many completely correct solutions.

The hypotheses were usually stated correctly, although a small percentage of candidates used p instead of λ as their population parameter, and a few candidates used H_1: $\lambda \leq 9$ or H_1: $\lambda \neq 9$. Nearly all candidates used Po(9), and whilst the large majority calculated the value of the test statistic $P(X \leq 4)$, it was not uncommon to see $P(X = 4)$ calculated. Correct conclusions were seen often and most of these were in the context of the question, i.e. ‘no evidence that the mean number of <u>accidents</u> has <u>reduced</u>’.

Question 5

This proved to be a demanding question for all but many persevered and made good progress with just over 25% achieving full marks. Many candidates used the given information that $E(X) = \frac{17}{12}$ and correctly multiplied $xf(x)$ prior to integrating, but often an elementary algebraic error prevented candidates from reaching a correct equation. Typically the use of the limit –1 caused some slips.

Given the unstructured nature of the question, a significant minority of candidates could not form a second equation using the fact that the two integrals needed to sum to 1. Algebraic errors were less frequent with those who did form this equation, though some candidates failed to integrate $3k$ before substituting in the limits of 3 and 2.

Nearly all candidates who obtained 2 equations in a and k, made an attempt at solving their simultaneous equations though the algebra here was too challenging for some.
In part (b), the correct mode of 2 was not seen too often. The best responses provided a sketch of f(x). $f(2) = \frac{5}{9}$ was a common incorrect response alongside 0 and 3.

Question 6

Parts (a), (b) and (c) of this question provided a good source of marks for most candidates. In part (a) most calculated the exact value required using the calculator. A significant minority of those who attempted to use the tables struggled to obtain the correct difference of F(5) – F(4), with F(6) – F(5) regularly seen.

Part (b) was extremely well answered although a small number of candidates failed to distinguish between the standard deviation and the variance of X.

Attempts at the hypothesis test in part (c) were generally good, with many completely correct solutions. The hypotheses were nearly always correctly stated, and although B(20,0.3) was nearly always used, some candidates calculated an incorrect test statistic such as $P(X = 8)$ or $P(X \geq 9)$. Conclusions were nearly always accurately stated and most were in context.

Part (d) proved to be extremely demanding for candidates, with a number of blank responses. Although a significant minority were able to formulate a correct algebraic inequality for n, relatively few could conduct a systematic table search. The use of a sample size of $2n$ rather than n also led to a loss of the final accuracy mark for some candidates who managed to reach "n" = 10. Candidates using the tables should be encouraged to write down the values used to reach their conclusion. A significant proportion of candidates who stated that $n = 5$ provided little or no supporting evidence.

Question 7

This question produced a wide range of responses with ¼ of candidates making no progress and ¼ of candidates scoring at least 7 out of 8 marks. Most candidates were able to make a reasonable attempt, with B(n, 0.2) being recognised together with a normal approximation.

Candidates were generally able to standardise using their mean, although a continuity correction was frequently missing and sometimes a variance was used instead of a standard deviation.

Usually, the candidate's z-value was equated to 2, and an attempt was made to simplify the resulting equation. Various methods were used to solve this equation. Those whose used a substitution of the form $x = \sqrt{n}$ tended to be more successful than those who first squared their equation. Even when the correct roots to the quadratic equation were obtained, a large number of candidates selected the wrong one or incorrectly stated that 106.3 should lead to an answer of 107 questions.

Grade Boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:
http://www.edexcel.com/iwantto/Pages/grade-boundaries.aspx

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