Examiners' Report

Principal Examiner Feedback

Summer 2019

Pearson Edexcel International A Level

In Statistics (WST01/01)

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

Summer 2019
Publications Code WST01_01_1906_ER

Statistics S1 (WST01) June 2019

General introduction

The paper proved quite challenging but there were opportunities for all candidates to score in each question. Q1(c)(ii), Q2(c), Q3(c)(i), Q4(c) and the comment parts of Q6 proved quite discriminating whereas the start of question Q5 was answered very well.

Report on Individual Question

Question 1

Part (a) proved to be a very accessible opening to the paper and most candidates found the mean correctly and the correct formula for variance was often used. The main errors were caused by some confusion between standard deviation and variance and rounding to only 2 significant figures rather than the standard 3 significant figures we require. Changing the units from metres to centimetres was dealt with well for the mean but the standard deviation caused problems with many giving an answer of 1.35 and some thinking there was no change "because standard deviation is not affected by coding", clearly an understanding of coding is still a little shaky. In part (c) there were many clear and correct arguments to explain that the mean would be unchanged but only the very best candidates gave a convincing argument for the change in the standard deviation. Few realised that standard deviation represents an average of the deviations of values from the mean and that if the new values were greater than 1 standard deviation from the mean there would be an increase in the standard deviation. A little more emphasis on the concept and meaning of standard deviation rather than just the calculation of this statistic would help candidates in future.

Question 2

Again the first part of this question was answered well and most found the correct interquartile range but a when finding the range many ignored the outlier and a common incorrect answer was 139. In part (b) most gave a correct comparison using the quartiles but sometimes a correct comparison was followed by "positive skew". Some candidates gave a reason involving the mean but this value was not available and so no marks could be scored. Part (c) proved to be quite challenging. The major problem was that few candidates realised that the lower quartile would change and increase to 35. Most of those who spotted this realised that this would affect the interquartile range and so they would need to re-calculate for outliers. Those who did identify these features were often able to complete the new box plot correctly but most others usually just obtained a mark for drawing a box plot with an unchanged median. Part (d) did not depend on part (c) and it was encouraging to see many collecting marks here despite missing out on the previous part. The key was the weakness of the correlation not the fact that it is negative and those who realised this often added a suitable comment about the complaint not being supported.

Question 3

Part (a) was answered well although a few candidates lost a mark for mixing probabilities and percentages on their tree diagram. In part (b) most were able to form an equation based around P(T) but many struggled to form a suitable equation using the conditional probability of $\frac{41}{169}$, even when their equation led to a value of p outside the range (0, 1) they did not attempt to find and correct their

error. Those who did obtain both correct equations could usually find the correct values for p and q. The conditional probability required for (c) (i) was rarely identified and only the best candidates scored all 3 marks here. Part (ii) was beyond the reach of many but those who did have an increased probability could often gain the mark.

Question 4

Part (a) was answered very well by the majority of candidates and this straightforward application of the normal distribution is understood quite well now. There were the usual errors that came from rounding the z value to 2sf rather than 3, which gave an answer outside the acceptable range, and there was still a little uncertainty about whether or not to subtract the answer from the tables from 1. A simple sketch would help resolve this easily. In part (b) there were many correct solutions seen too the main problems here were a failure to use the percentage points table and a z value of 1.6449 and some gave an incorrect sign on their z value which yielded the common incorrect answer of 436 g, again a simple diagram should have alerted such candidates to their error. Part (c) was much more challenging and only the very best candidates made progress here. The common error was to miss the conditional probability and simply give an answer of 0.9082 from P(W > 450) a few candidates did realise that a conditional probability was required but a ratio of $\frac{0.9082}{0.95}$ was a common error. Part (d), of course, did not depend on part (c) and it was encouraging to see candidates still attempting this part. Many recognised the pattern $p^4(1-p)$ but they usually forgot to multiply by 5 and some gave an answer of the form $0.8\times0.95+0.2\times0.05$ whilst others thought that the normal distribution still had to be used.

Question 5

Part (a) was a friendly starter here and most answered this correctly. Good progress was usually made in part (b) as well with most knowing the correct method but some still thought that Var(X) = E(X^2) and others forgot to square E(X) before subtracting. There were some who made simple errors when solving their equation in p but most used the sum of probabilities correctly to find a value for q and, provided their value for p was a genuine probability, gained the follow through mark in part (c). Part (d) was quite challenging and only a few were able to clearly list all the cases for Amar to win the game with a common error being to include the case of a score of zero. Those who did find the correct list sometimes multiplied the probabilities of X > 0 together rather than adding these. There was little success in part (d) with many attempts focussing on the fact that E(X) > 0 rather than the critical issue that P(Amar wins) was less than 0.5.

Question 6

There were plenty of marks available here for routine calculations and these were answered very well but the parts requiring comment or interpretation were often lacking the context we require.

Most answered part (a) correctly although a few used $\bar{y}$ instead of $\sum y$ and some gave their correlation coefficient to only 2 sf accuracy rather than the usual 3 we require. In part (b) many failed to use the context of the question and answers mentioning the percentage of referrals and the distance of the clinic from the hospital were rare. Sometimes the response in (c) would have been ideal for part (b) but many here failed to mention that the correlation was strong or use the scatter diagram to comment on the points lying fairly close to a straight line. We tried to make it perfectly clear to candidates that question 6 continued after part (c) but a few, either missed this or were short of time and did not attempt parts (d) to (g). Part (d) was answered very well though with most candidates scoring full marks here. The

interpretation of the gradient proved challenging, despite a similar demand being in previous papers, and a few mixed up their variables and wrote that a the distance increased by 1.7 km for each extra percentage referral. Most attempted part (f) but some were clearly just drawing a line “by eye” rather than using their, often correct, equation in part (d). Only a few candidates were able to identify the correct point in part (g) and those who did were often scoring well on the whole question.

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