

Examiners' Report/
Principal Examiner Feedback

October 2016

Pearson Edexcel International A-Level
Statistics 1
(WST01)

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 2016

Publications Code WST01_01_1610_ER

All the material in this publication is copyright

© Pearson Education Ltd 2016

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>

Introduction

The paper seemed to work very well with almost every question providing an opportunity for all students to get started and the later parts providing discrimination for the top grades. Questions 1, 3 and 5 proved to be good discriminators, but there were plenty of fully correct solutions to these as well, whilst question 2 was probably the most accessible with a mean score of nearly 12 and over 40% of students scoring full marks.

Comments on individual questions

Question 1

Apart from the large minority who failed to make any progress here, there were some good solutions to parts (a) and (b) but fewer scored well in part (c). Some made the process more difficult by standardising and finding a value for $\frac{a}{\sigma}$ but they often eventually found the correct answers in (a) and (b). Some less able students simply evaluated $P(Z < 0.35)$ in (a) and they made little progress. In part (c), students needed first to recognise the conditional probability and then identify the numerator as their answer to part (b) and students who achieved these two stages often went on to score full marks for the question.

Question 2

Nearly all students answered part (a) correctly and most were able to find $E(X)$ successfully too. Forming a correct second equation in a and b proved more troublesome with no use of brackets leading to $-4b$ and a surprising number forgetting to square the 3. It was encouraging to see far fewer students dividing an otherwise correct expression for $E(X^2)$ by 5 but disappointing that a sizeable minority still thought that the formula for $E(X^2) = \sum xP(X = x)^2$. Most ended up with 2 linear equations in a and b and could solve these correctly, though some seemed not to worry when one of their values was negative.

Part (d) was answered well, with only a few failing to square their $E(X)$, but parts (e) and (f) proved more challenging. In part (e), some simply found $P(X > 0)$ whilst others solved the inequality correctly but then failed to identify the relevant values of X . Some tried to find the probability distribution of Y and this often helped them in (e) but they frequently came unstuck in (f) when trying to find $E(Y^2)$ on their way to $\text{Var}(Y)$. Those who used the $E(aX + b)$ and $\text{Var}(aX + b)$ formulae usually answered part (f) very successfully.

Question 3

Drawing the Venn diagram presented some challenges here with some drawing 3 intersecting circles and failing to label the “empty” regions with zeros and others struggling to calculate the number buying just a scarf, but only a very small number of the students failed to make any progress on this question. In part (b) many were able to secure the mark in part (i) from their Venn diagram and there were many correct answers to part (ii) as well suggesting that far more students now have a working understanding of conditional probability than used to be the case a few years ago. Part (iii) attracted some good responses with most students giving the correct, labelled probabilities, and clearly showing that the two events were independent. Most chose to do this by demonstrating that $P(S) \times P(C) = P(S \cap C)$ rather than using $P(S)$ and $P(S | C)$ found in parts (i) and (ii) and stating that the events are independent because these two probabilities are equal. The final part proved quite challenging with many choosing “gloves” because $P(G \cap S) > P(C \cap S)$ and others calculating $P(C | S)$ but a reasonable number of the students did compare the correct conditional probabilities and made a correct deduction.

Question 4

These calculations around correlation and regression equations were handled confidently by most students and the majority of them achieved 11 or more marks. Parts (a), (b) and (c) were usually answered correctly, although a few still persist in giving their correlation coefficient to only 2 significant figures instead of the standard 3sf required. Some were unsure about part (d) and did not realise that it was the fact that the correlation coefficient was close to 1 that was important not that it was simply positive. In part (e), we did not give the required form of the answer and this meant that some students were unsure which way around to use the means of h and f when calculating the intercept. In part (f) a good number standardised correctly but only a small minority realised the need to consider $-3 < E < 3$ and most simply found $P(E < 3)$.

Question 5

This proved to be the most challenging question with some failing to make any progress and less than 10% securing full marks. Part (a) was a fairly standard question requiring the students to set up and solve 2 linear simultaneous equations. Problems arose when using $P(A < 388) = 0.001$ as many students failed to realise that their z value should be negative and equations of the form $388 - \mu = 3.10\sigma$ or $388 - \mu = 3.0902\sigma$ were quite common. Sadly, few seemed to realise that an error must have occurred when they found that the value for σ was negative and students should be encouraged to watch out for these simple indicators that mistakes have been made. Part (b) was not answered well by the majority of the students. About 20% of them made some progress and realised that an expected value for a random variable X was required. Unfortunately many of these failed to assign the correct signs to the values of X and others could not match the values with the appropriate probabilities.

Question 6

This question was answered well with nearly all of the students scoring over 50% of the available marks. Often the final question on the paper has a number of blank responses and it was encouraging to see that almost all of the students found something they could do in this final question. Parts (a) and (b) were answered very well and most calculated the limits for the outliers in part (c). Sometimes these were not identified at this stage but we could occasionally award the marks if they were seen on their box plot. The box plots were usually correct with the main loss of marks arising from missing outliers. Nearly everyone could calculate the mean correctly but there was still plenty of confusion over the standard deviation with many finding S_{xx} instead of $\sqrt{\frac{S_{xx}}{n}}$.

The calculation in part (f) was an unfamiliar one and was often misinterpreted with many students calculating $2.7 \times$ their standard deviation but failing to add or subtract this value from the mean. In part (g), some students suggested skewness may be a reason for rejecting a normal distribution but they sometimes described their answer as having positive skewness when it should have been negative. A common response was simply to say that the data was discrete not continuous but, as this was not based on any of their calculations, it did not score the mark.

