Examiners' Report

Principal Examiner Feedback

January 2024

Pearson Edexcel International Advanced Level in Statistics S3 (WST03) 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

January 2024
Publications Code WST03_01_2401_ER

WST03 PE Report January 2024

General Introduction

Overall, this paper allowed all students to demonstrate their ability and knowledge of the WST03 specification. Q6(c) was the most discriminating on the paper followed by Q7(c). Questions involving chi-squared tests remain a strong topic for students at all levels. In particular questions that ask a student to show something is true require all the steps in the working to be shown.

Report on Individual Questions

Question 1

This question was accessible to many students with the majority scoring full marks.

Most students stated the hypotheses correctly. The common error was giving the null and alternative hypotheses the wrong way round. The vast majority of students were able to calculate all the expected values accurately. Many then went on to calculate a correct value for $X^2 = \sum \frac{(O-E)^2}{E}$ or $\sum \frac{O^2}{E} - 150$. Other errors seen include stating an incorrect value for the degrees of freedom or failing to contextualise the conclusion.

Question 2

Part (a) was generally answered well by many students, with many scoring at least one mark. However, some students incorrectly seemed to think that systematic sampling required dividing the population into 80 groups and then selecting one employee from each group.

Part (b) again was generally answered well by many students, with many scoring at least one mark. A common error was to give 53.75 for London and 31.25 for Edinburgh. Students should have realised that these values should have been rounded to a whole number.

Question 3

Part (a) was answered well by many students. Some students lost marks for either stating their hypotheses in words or not in terms of rho. The vast majority of students were able to calculate the product moment correlation coefficient. Occasionally incorrect critical values were stated, which was usually the value for a one-tailed test rather the required two-tailed test critical

value. The contextual conclusions were generally well written but occasionally failed to give the required context of 'tea consumption' and 'population'.

Part (b) was answered well by many students, with many scoring full marks. Occasionally there were errors in the rankings or in the calculation of $\sum d^2$, but it was clear that students knew how to calculate a Spearman's rank correlation coefficient.

Part (c) was again answered well by many students. Same as part (a) some students lost marks for incorrect hypotheses or an incorrect critical value, although this was seen less than part (b). The contextual conclusions were generally well written but occasionally failed to give the required context of 'positive', 'tea consumption' and 'population'.

Question 4

Part (a) was a nice easy start to the question with many students showing that they were able to calculate the mean of the given frequency table.

Part (b) and part (c) were answered well by the vast majority of students. It was clear that students knew how to use the Poisson distribution to calculate expected frequencies.

Part (c) of this question proved to be accessible to students. Many were able to correctly state both hypotheses, however some students lost this mark as they thought they needed to state Po(1.75) being a good fit rather than just Poisson being a good fit. Many students realised that the last 2 cells needed to be combined. Mostly correct calculations and conclusions in context followed.

Question 5

In part (a) many students were able to correctly state the hypotheses in terms of μ. A few however lost marks as they used different letters to represent History and Maths and it was not clear which was History and which was Maths Many students were able to find a correct test statistic, but a common error was to use an incorrect standard error. Some students lost marks due to failing to provide the critical value to at least 4 decimal places, often stating $z = 1.645$ instead of 1.6449 or better. Most students made the correct decision to not reject H_0, but some still didn't provide a contextual conclusion.

Part (b) proved to be a challenge to students. Many students failed to realise that large samples allows the central limit theorem to be applied. Many stated that we need to assume that $s^2 = \sigma^2$ but missed the fact that this was needed in both groups. A common incorrect answer was that the samples needed to be independent.

Question 6

A variety of approaches was seen in part (a). The most common approach used the given information to calculate σ and then this was used to calculate the required 99% confidence interval. Some students lost marks due to failing to use a z value to at least 4 decimal places, often using $z = 2.576$ instead of 2.5758 or better.

Part (b) was answered well by the vast majority of students with many scoring full marks for calculating unbiased estimates of the mean and variance.

Part (c) proved to be a good discriminator with only the most able students gaining full marks. Many students were able to score the 1st M mark as they could calculate the required combined sum, but for many this was the only mark scored. Many students struggled to find a combined sum of squares. Even those that did find this then struggled to find the required combined s^2. A common error seen was based on students assuming that the two samples were normally distributed and therefore tried to combine the mean and variance of these normal distributions. Only the most able understood that $\frac{s}{\sqrt{n}}$ was required for the standard error of the combined sample.

Question 7

Part (a) was a good source of marks for many students. Common errors included $a = 4 \times 180 - 330$ (which gave an incorrect answer of 390) and $b = 4.5^2 \times 4 + 6.7^2$ (which gave an incorrect answer of 125.89).

Part (b) was generally answered well. Most students were able to find E(X) = 6 and Var(X) = 110.5. Errors usually occurred when finding the variance. Most students were able to standardise using their mean and standard deviation, however a few used the variance rather than the standard deviation and so lost the final 2 marks.

Part (c) proved to be a good discriminator with only the most able students gaining full marks. Many students failed to realise that they needed to write a single distribution using $\overline{S} = \frac{S_1 + S_2 + S_3}{3}$. Many students gained the 2nd M1 as they correctly identified E(T) = 0 but the variance proved to be more problematic and the common error here was to give Var(T) = 27, which comes from assuming S_1 and $\overline{S}$ are independent. Most students were able to standardise using their mean and standard deviation, however a few used the variance rather than the standard deviation and so lost the final 2 marks.

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