Examiners' Report
Principal Examiner Feedback

January 2025

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 2025
Publications Code WST03_01_2501_ER

WST03 PE Report January 2025

General Introduction

Overall, this paper allowed all students to demonstrate their ability and knowledge of the WST03 specification. It was pleasing to see some very strong performances on questions and clear structure to some solutions. However, on a number of questions, students did seem to misunderstand what was being asked of them and proceeded down a more anticipated route, as opposed to appreciating the subtle differences between these questions on this paper and maybe ones they have seen on previous past papers. Q7 was the most discriminating on the paper. 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.

Question 1

This question tested the topic of Spearman's rank and testing for positive correlation. Most completed this well, although students should realise that by showing their working fully, they stand to minimise the loss of marks due to small slips.

In part (a), students were usually successful in achieving the coefficient of 0.406 (3sf). Occasionally there were errors in the ranking or in the calculation of $\sum d^2$, but it was clear that students knew how to calculate a Pearman's rank correlation coefficient.

Part (b) was answered well by students. Some students lost marks for writing the hypotheses in words or not in terms of rho. Occasionally incorrect critical values were stated. The contextual conclusions were generally well written but occasionally failed to give the required context of 'positive correlation', 'position' and 'attendance'.

Question 2

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) caused students some difficulty, as it required an explanation. Those that used a calculation to explain why r was 1.203, usually were more successful than those that tried a written explanation. Often the key phrase 'expected frequency' was missing.

Again, part (c) caused students some difficulty, as it required an explanation. Like part (b) often the key phrase 'expected frequency' was missing. Often this part was left blank.

Part (d) of this question proved to be accessible to students. Many were able to correctly state correct hypotheses, with only a few giving hypotheses the wrong way round or missing the key word 'Poisson'. Most errors occurred when stating the degrees of freedom. Even though part (c) had told students that expected frequencies had been combined, too many students subtracted 1 constraint (rather than 2) and so had an incorrect answer for the degrees of freedom. As the mark scheme allowed for follow through from their degrees of freedom, many students were able to use tables to give a correct critical value. The contextual conclusions were generally well written but occasionally failed to give the required context of 'Poisson'.

Question 3

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

In part (b) many students were able to correctly state the hypotheses in terms of μ. 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 reject H0, but some still didn't provide a contextual conclusion.

Part (c) 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 but missed the fact that this was needed in both groups. A common incorrect answer was that the samples needed to be independent.

Question 4

Part (a) asked students to show that the standard error was 0.0124. Whilst many students were able to do this, too many students failed to show sufficient working to gain full marks. Students should be encouraged to show all their steps in their working to ensure that they gain full marks. Some students used rounded z values, e.g. 2.6/2.58/2.576 rather than 2.5758 and so lost 2 marks. Students should be encouraged, where necessary, to use 4 decimal places with z values.

Part (b) was generally answered well as they could use the given standard error (from part (a)) to calculate a 95% confidence interval for μ. As the z value used was 1.96, then there were less issues with rounded z values. Occasionally students lost marks as they gave their final answer to 2 decimal places rather than the required 3 significant figures. A few students also lost marks as they failed to give their answer as a confidence interval.

In part (c), a variety of approaches was seen. Many were able to set up a correct equation to calculate the required confidence interval, with many scoring the first 2 marks by arriving at a correct value for z (awrt 1.61) At this point some students were unable to make any further progress. For those that did many scored the next M mark as they arrived at one of the probabilities stated in the mark scheme. Only the better students then realised that this probability needed to be multiplied by 2. Too often the final answer mark was withheld due to students using rounded values in their calculations, which resulted in an incorrect final answer.

Question 5

Part (a) asked students to give one advantage and one disadvantage of using quota sampling. The advantage of quota sampling caused students more problems than the disadvantage. Common incorrect answers for the advantage included reference to ‘quicker’, ‘cheaper’ or ‘easier’. In this question these were not acceptable as quota sampling was not quicker, cheaper or easier than simple random sampling. A few students confused quota sampling with stratified sampling and so gave an advantage relating to stratified sampling and so no credit was given.

Many students gained the mark for the disadvantage as they often gave answers that referred to bias or non-random sampling.

Part (b) was answered well, with many students scoring full marks for correctly calculating the required expected frequencies.

Part (c) provided students with an opportunity to score several 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 the 2 contributions to the χ^2 value. Many then added to the given 4.549 to arrive at a correct value for $\sum\frac{(O-E)^2}{E}$

Other errors seen include stating an incorrect value for the degrees of freedom or failing to contextualise the conclusion.

Question 6

In this question students fell into 2 categories, those that understood the concept of estimators or those that did not. For those that did not, very few marks were scored and for those that did full marks was often seen.

Part (a) asked students to show that $\bar{X}$ was an unbiased estimator. Unfortunately, too many students failed to realise that they needed to state that $3a+6 \neq a$ and so lost the final A mark. Again, it is worth encouraging students to show sufficient working to gain full marks in show that questions.

For those students that scored marks in part (a), part (b) was often answered well, as students were able to find the bias ($2a + 6$).

In part (c) many students scored full marks. However, for those students that had a poor understanding of estimators it was not unusual for 0 or 1 mark to be awarded. If 1 mark was awarded it was usually awarded for writing $c(3a+6)+d=a$

In part (d) many students scored full marks. However, for those students that had a poor understanding of estimators it was not unusual for 0 marks to be awarded. A few students lost the final mark as they gave incorrect subsequent working following a correct answer.

Question 7

This question was on the distribution of linear combinations of independent normal random variables. It was clear that some students were confident with most of this topic, although part (c), provided the most challenge for many so it was rarely answered correctly by most.

In part (a), it was typical for students to score full marks. They were confident in finding the mean and the variance which was required for the scenario and mostly students were able to standardise correctly. Common errors included finding an incorrect variance, usually 0.0073 which came from $3^2 \times 0.01 + 4^2 \times 0.02$ rather than the required $3 \times 0.01 + 4 \times 0.02$. However, most of these students scored both method marks.

In part (b) many students were able to access the first 3 marks, but only the better students were able to score full mark. Many were able to set up a correct normal distribution and most were able to standardise correctly. However, only the better students were able to recognise that this probability needed to be multiplied by 2. Even when they did multiply by 2 a few lost the final mark as they went on to subtract from 1 and therefore spoilt a correct answer.

Part (c) proved to be a good discriminator with only the most able students gaining full marks. Many failed to identify the correct normal distribution required. Whilst some students were able to recognise that the mean was 0, the variance of the normal distribution caused the majority of students an issue. Many were able to standardise and set equal to *z* value, but the incorrect variance often led to students to an equation which was not a quadratic, and so no further marks were available. Only the best students were obtained the correct 3 term quadratic and those that did usually scored full marks. However, a few lost the final A mark as they failed to reject, the obvious silly answer, –101.

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