Examiners' Report
Principal Examiner Feedback

Summer 2023

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

Summer 2023
Publications Code WST03_01_2306_ER

General

Students were generally well prepared for the demands of this paper with many strong performances seen. Q5 was the most discriminating on the paper followed by Q3. Questions involving chi-squared tests remain a strong topic for students at all levels. Students should be advised to comment in the context of the question not only when completing a hypothesis test but when referring to assumptions required to carry out the tests.

Question 1

Part (a) was answered well by the better students. Most were able to give one correct condition, but many failed to give two. Common incorrect answers included a comment about tied ranks or saying that the distributions were normally distributed.

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

Part (c) was again answered well by many students. Same as part (b) some students lost marks for incorrect hypotheses or an incorrect critical value, although this was seen less than part (b).

Question 2

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

In part (a) the vast majority of students were able to calculate all three expected values accurately.

In part (b) most students stated the hypotheses correctly. The common error was giving the null and alternative hypotheses the wrong way round. Other errors seen include stating an incorrect value for the degrees of freedom or failing to contextualise the conclusion.

Question 3

Part (a) caused some students difficulty. Some just repeated the information given in the stem that μ and σ^2 were unknown and scored no marks.

In part (b) many students were able to state or use $\mathrm{E}(S) = a\mathrm{E}(X_1) + b\mathrm{E}(X_2)$ and get to $\mathrm{E}(S) = \frac{46}{35}\mu$

However, some students did not go on to state that this was not equal to μ and so lost the final A mark.

Part (c) was answered well by the vast majority of students.

Part (d) was answered well with many students just correctly writing the required equation in *a* and *b*. The common error here was to forget to equate $a\mathrm{E}(X_1)+b\mathrm{E}(X_2)$ to μ

Part (e) caused some students difficulty and careless algebraic errors were seen. Students should be encouraged in show that questions to ensure they have sufficient working shown.

Question 4

In part (a) those students that realised that they could use the area under the graph usually did so with a high degree of success. Other approaches involved integration and many students were able to do this successfully. Again, as this was a show that question students need to ensure that they provide sufficient working so that all the available marks can be awarded.

Part (b) of this question proved to be very accessible to students. Many were able to correctly state both hypotheses, The expected values were generally calculated correctly with only a few using the incorrect end of the pdf or a discrete uniform distribution. Mostly correct calculations and conclusions in context followed.

Question 5

Parts (a) and (b) were good sources of marks with most students correctly calculating the required confidence intervals. Students should be encouraged to not round the required *z* values as this could affect the final result.

Part (c) proved to be a good discriminator with the most able students able to gain full marks. At the other end of the spectrum students continue to make errors in finding the mean and variance. A common error was that students did not realise that they were dealing with the distribution of sample means which resulted in errors in the calculation of the correct variance. Students generally struggled to identify the regions in which the two confidence intervals overlapped and so made very little progress here. Common error included finding $\mathrm{P}(\overline{x}-\overline{y}>0)$ or $\mathrm{P}(\overline{x}-\overline{y}<0)$. Even better students who found a correct region often lost marks as they failed to realise that there were 2 regions in which they overlapped.

Question 6

Part (a) was answered well by the vast majority of students with many scoring full marks for calculating α and β correctly.

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 H_0, but some still don't provide a contextual conclusion.

Parts (c) and (d) proved to be a challenge to students. Many students in part (c) seemed to think that was necessary to assume a normal distribution and failed to realise that large samples allow the central limit

theorem to be applied. In part (d) 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 7

This question was answered well by the vast majority of students. The probability in part (a) was a straightforward 4 marks for most students. Common error included an incorrect calculation for the variance of the normal distribution.

Again, in part (b) the calculation of the variance caused some students issues often writing $41+\sigma^2$ rather than $41+3\sigma^2$. Most were able to standardise and equate to a correct *z* value but some students had *z* values that were incompatible with their standardisation e.g. $\frac{-9}{\sqrt{41+3\sigma^2}} = 0.7$ and, whilst this led to a correct answer, the final A marks was not awarded.

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