Pearson
Edexcel

Examiners' Report
Principal Examiner Feedback

Summer 2022

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 2022
Publications Code WST03_01_2206_ER

Students were generally well prepared for the demands of this paper although they found this more challenging than previous papers in the series. The calculation of Spearman's rank correlation coefficient continues to be well attempted by most students as is the calculation of probabilities from combinations of independent normal distributions. Question 4 was the most challenging on the paper and discriminated the most able students. Students should continue to take care when completing a hypothesis test to use correct notation, where appropriate, in their hypotheses.

Question 1

Question 1 provided a successful start to the paper for many students as they confidently displayed knowledge of Spearman's rank correlation coefficient and the associated hypothesis test.

In part (a) many were able to explain how to deal with tied rankings. However too many referred to adding 0.5 to both ranks which is an incorrect answer.

The attempt at rankings in part (b) was generally well done, though some used reverse rankings which caused some problems later on in the question. Most were able to score full marks for obtaining a correct value of the coefficient. A few ranked televised tournaments won and found the difference between an incorrect pair of rankings.

In part (c), correct notation was generally used and a one-tailed test was virtually always selected. Most were able to state the correct critical value and successfully concluded that there was sufficient evidence of a positive correlation between rank and total tournaments won.

Part (d) was again well answered with the correct critical value and significance level stated. Those that lost marks usually failed to give one or the other of the 2 values needed

Question 2

This question was accessible to many. The unbiased estimates of the mean and the variance in part (a) were nearly always correct. A few found the variance and so lost 2 marks.

Many were able set up correct hypotheses in part (b) generally using the correct notation and attempting a two-tailed test. A few however gave hypotheses in words or a one-tailed test. Mistakes in calculating the standard error included neglecting to square the 2.2. Most attempted the test statistic using their standard error but too many lost marks as they used the wrong standard error altogether. Those who found the correct test statistics virtually always compared it with a correct critical value and gave an appropriate conclusion in context.

Part (c) was poorly attempted with very few referring to both samples needing to large.

Question 3

This was again a very accessible question on the paper with many students scoring full marks.

In part (a) many knew that mean of the sample was the mid-point of the confidence interval. Some overcomplicated the calculation by using simultaneous equations, but most arrived at the correct result.

In part (b) most were able to work backwards from the given confidence interval in to show the correct value of σ . The z-value of 1.96 was almost always correctly selected though some incorrectly used 1.6449..

Again, part (c) was generally well answered and good attempts were seen. The first 2 marks were generally awarded as students were able to give a correct z value. In many cases students went on to calculate a correct probability of 0.0401. In a number of cases, the last two marks were not scored when students gave a final answer of 96% which neglected that there are two ends of a confidence interval.

Part (d) again was generally answered well, and a z-value of 1.96 was almost always correctly selected. The common error here was failing to use the square root of n.

Question 4

This was the most demanding question on the paper with many students making little or no progress whatsoever.

In part (a) a significant number of students thought that this was an example of a Poisson distribution, which then cost them the vast majority of marks available in this question. This was a little surprising as time is a continuous variable.

In part (b), this was answered well by those that had a correct distribution in part (a) but those that had an incorrect distribution usually lost all 3 marks.

Part (c) allow for those students that had done part (a) wrong and many were able to score the 3 M marks. Common error here was to not divide the variance by 46. Many standardised correctly for one of two ends but too many failed to do both ends of the required probability and so the final 2 marks were not awarded.

In part (d) more relevant accuracy is needed when commenting on aspects of the Central Limit Theorem. There were too many students that made reference to $s^2 = \sigma^2$ and in this case scored no marks.

Question 5

In part (a) most responses indicated that μ and σ were the reason that the expression was not a statistics but not all stated that these were unknown values.

In part (b) students struggled with the definition of an unbiased estimator and the common incorrect response seen was bias = 0

In part (c) although candidates showed an understanding that they were to show that $E(U_1) = \mu$ and $E(U_2) = \mu$ quite often the notation used was generally poor.

In part (d) although there was a good understanding that the coefficients of X_1 and X_2 would be squared when finding the variance of a statistic, the –2 was not always squared to give 4 and the 4 in the denominator of U_2 was not always squared. As the efficiency of an estimator was explained in the question a reason why U_2 was the most efficient was usually given.

Question 6

Part (a) of this question was accessible to most candidates but pert (b) proved to more challenging.

In part (a) most were able to combine the incidences of men and women to derive the correct parameters for the Normal distribution. Students generally standardised correctly and went on to calculate a correct probability.

In part (b) the initial strategy was the downfall of most students, with too many failing to realise that the maximum needed to be calculated from the most likely heaviest people, namely the men. Many tried some combination of men and women but the did not proceed any further. Those that did reach a Normal distribution of $N(80x, 100x)$ occasionally faltered by standardising using $10x$ rather than $10\sqrt{x}$ and were unable to form a 3TQ. However, most were able to set an equation = 1.96

Question 7

This question had students fall into 2 camps, those that could and those that could not. Those that could often scored full marks but those that could not scored poorly.

In part (a) Most attempted to write the hypotheses using the notation that they were testing whether the die was biased or not. However too many gave the null and alternative hypotheses the wrong way round.

In part (b) those that found the expected value as x usually scored full marks but those that did not usually led to complicated and often messy algebra, although many were able to gain the method marks for their attempt at the chi squared test statistic. Most found the correct degrees of freedom and hence the correct critical value.

In (ii) many scored at least 1 mark as they followed through their answer to (i) and multiplied by 6

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