Pearson
Edexcel

Examiners' Report
Principal Examiner Feedback

October 2020

Pearson Edexcel International A Level
In Statistics 3 (WST03)

Paper: 01 Statistics S3

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 2020
Publications Code WST03_01_2010_ER

General

Candidates were generally well prepared for this paper with many strong performances seen. Questions 5(c), 6(b) and 7(c) were the most discriminating parts of the paper. Candidates are encouraged to show all stages of their working. This will allow them to access method marks in a question after mistakes are made. This is particularly the case when working out expected values for a goodness of fit test. Those questions that require reasoning still cause difficulty for many candidates as written expression often lacks sufficient detail.

Report on individual questions

Question 1

This question provided a mixed start to the paper for candidates with some successfully earning full marks but others finding this question challenging. Some candidates confused the sum of 1, 2, 3, … α for the mean in part (a). Others simply found $\mathrm{E}(\bar{X})$ instead of $2\mathrm{E}(\bar{X})$. Candidates were required to show that the expectation was not equal to α in order to satisfy the demand of the question.
Those successful in part (a) often went on to correctly find 11 as the estimate of α in part (b). Others just found the sample mean but went no further. On the whole, candidates' notation was sufficient to make their methods clear, but not always correct.

Question 2

This question was the most successfully answered question on the paper and accessible to candidates at all levels. Part (a) was nearly always correct as candidates displayed confidence finding expected values and calculating the test statistic for a contingency table.
Part (b) was again answered successfully by most candidates. Common errors included: swapping the hypotheses the wrong way round, using 5 degrees of freedom instead of 4 and failing to give the conclusion in context. Some candidates believed this was a test to see if the degree categories were independent of each other rather than independent of the department.

Question 3

Parts (a) and (b) were successfully answered by the majority of candidates with part (c) proving to be more demanding. Most candidates obtained an accurate value for the Spearman's rank correlation coefficient in part (a). Those who did make slips were able to pick up method marks by showing the stages of their working clearly. In part (b) the majority of candidates are aware of the requirement to give their hypotheses in terms of ρ though a few still attempted to give their hypotheses in words. Almost all found the correct critical value and gave their conclusion in context.
In part (c) some candidates decided to explain in words how to find ranks and how to plug numbers into the Spearman's rank formula rather than examining this particular case where they are tied scores. The answer required candidates to explain that average ranks were required and, hence, the full product moment correlation coefficient formula would be necessary here. Written expression remains a challenge for some candidates and just stating 'use pmcc' was not sufficient.

Question 4

Though most candidates made very good attempts at parts (b) and (c), parts (a) and (d) were less well answered.
Part (a) should have been standard for candidates but it was not uncommon to omit the comment about randomly selecting the first student. Some explained that you should randomly select one student from each group of 7 rather than systematically selecting every 7th student.
In part (b), there were many good responses seen with some going above and beyond the demand by completing the entire hypothesis test. The most common mistake here was to use 8 degrees of freedom. Part (c) was again well answered. Most gave the appropriate context in the hypotheses and went on to correctly find the value of the test statistic. Those making errors used 19 as class width instead of 20 and ended up with incorrect expected values. Fewer errors were seen in identifying the critical value for the test and many contextualised conclusions to the hypothesis test were given.
Many failed to engage with the demand of part (d) and were unable to score marks here. The question asked about the use of Luka's tables but many simply explained how to carry out a simple random sample without referencing the tables. In order to use the tables, the students would need to be numbered and random numbers would need to be generated - but many candidates' comments did not make reference to these ideas.

Question 5

Part (a) provided an accessible start to this question with part (c) proving to be one of the most discriminating parts of the entire paper. Most candidates are confident with finding unbiased estimates of population variances as was indeed the case here in part (a).
The hypotheses is part (b) caused some trouble, particularly H_0 which was sometimes given as $\mu_A = \mu_B$. There were a few who used W instead of μ. Most were able to give a correct expression for the standard error and compute accurately the test statistic required. Many concluded the hypothesis correctly by stating that the greengrocer's belief was incorrect.

Though some candidates were able to earn the first mark in part (c) for stating that both sample means could be modelled as being normally distributed, it was rare to see the second mark being scored. A lack of clarity often prevented either or both marks from being scored. Comments such as 'the means are normally distributed', 'sample variance = population variance' were not sufficient to score the marks in (c).

Question 6

This question provided higher achieving candidates the chance to display their knowledge. Though a fair number of candidates did realise that the sample mean would follow a normal distribution, there were many errors seen in the variance. Some candidates attempted to use a Poisson approximation in part (a) to a normal distribution whilst others thought the distribution would have a Poisson distribution itself.
Nearly all candidates found the correct z –value in part (b) and used the correct form for a confidence interval, but fully correct answers were reserved for the most able candidates on this paper. The most common error was to believe that $\sigma = \lambda$ and this led to the frequent incorrect answer of $\lambda = 3.19$.

Part (c) was generally well answered. On some occasions a calculated value of p was used instead of the given value of the confidence interval.

Question 7

The final question of the paper again provided an accessible start but a difficult finish. Part (a) gave rise to the common mistake made by candidates, namely confusing $3C$ with $C_1 + C_2 + C_3$ leading to an incorrect value of the standard deviation. A few candidates gave as their final answer the probability that the total weight was less than 475.8 instead of greater than it.

Similar errors were made in part (b), but on the whole, this part was well answered.

The final part of the paper was one of the most demanding questions. Most were able to combine the distributions to come up with a correct expression for the mean. The variance proved far more difficult and a variety of errors were seen here. $n^2 \times 4 + n \times 4$ and $(n-1)^2 \times 4 + n^2 \times 4$ were common errors. Most were able to standardise using their mean and standard deviation and went on to use the quadratic formula or their calculator to solve the resulting quadratic. Those who did rely solely on the calculator were unable to access the final M mark if their answer was incorrect.

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