edexcel

Examiners' Report/ Principal Examiner Feedback

Summer 2016

Pearson Edexcel International A-Level Statistics 3 (WST03/01)

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 2016
Publications Code WST03_01_1606_ER

General introduction

Questions 1 – 5 and 6(a) on this paper proved accessible to all candidates but questions 6(b), 7 and 8 provided some discrimination for the top grades. Candidates must take care when completing a hypothesis test to use correct notation and context where appropriate. Candidates are encouraged to show their methods as much as possible even when calculations are carried out on calculators.

Report on individual questions

Question 1

This question provided a gentle start to the paper with virtually all candidates scoring at least 7 of the 9 marks available. The calculation of the Spearman's rank correlation coefficient is a standard procedure for candidates at this level and only a few slips or careless mistakes were seen. Part (b) was handled almost as well with the majority of candidates using correct notation to set up the hypotheses, including a one-tailed alternative hypothesis. Most went on to select the appropriate critical value and conclude the test in context. Some candidates, however, lost the final mark by stating 'do not reject H_0' without going any further.

Question 2

This question was also well attempted as half of the candidates scored 11 or 12 marks. In part (a) most were able to set up correct contextual hypotheses in words though some confused independence and dependence and gave the hypotheses the wrong way round. The instruction to give expected frequencies to 2 decimal places was generally heeded which meant most were able to calculate the χ^2 test statistic to the required degree of accuracy. On the whole, candidates chose the correct critical value by correctly determining the degrees of freedom. Most conclusions were given in context and it was common to see full marks gained in part (a).

Answers to part (b) were mixed. Many candidates only considered the observed frequencies to make their decision (rather than comparing these with the expected frequencies). The most common incorrect answer given was 'centre *B* as it had the highest number of passes.'

Question 3

The success rate on this question was slightly lower but one quarter of candidates did earn full marks. It was clear that many candidates have an understanding of the advantages and disadvantages of different sampling methods, but many relied on 'textbook' responses rather than applying their knowledge to the given situation. This was evident in part (a) where answers such as 'not random' and 'not accurate' were seen. In part (b) the most common error was to omit to mention the need for obtaining/generating a numbered list of employees. The calculation of the number of employees to select from each office was generally done accurately. Part (c) of this question was answered most successfully as context was not required.

Question 4

Most candidates made a good attempt at this question, particularly part (a). Here most set up hypotheses in terms of population parameters using correct notation and made it

sufficiently clear which mean applied to the adults and which applied to the children. The calculation of the standard error was done correctly by the majority of candidates with only a few forgetting to square s_c and s_a or neglecting to take the square root of the expression. The final conclusion was usually given in context though some responses did not make reference to the 'task'. The responses to parts (b) and (c) were mixed. In (b) many did not make it clear that the Central Limit Theorem meant that **both** $\overline{X}_C$ and $\overline{X}_A$ are approximately normally distributed. Though there were a number of blank responses seen in part (c), those who did attempt this part were able to correctly identify $s^2 \simeq \sigma^2$.

Question 5

At this stage of the paper full marks were gained by only the most able candidates. In attempting to set up the hypotheses, many candidates stated that the continuous uniform was a suitable distribution to model the data but forgot to specify the interval for which it was suitable. Despite that, there were many correct attempts at calculating the χ^2 test statistic (though there were also a significant number of incorrect attempts which treated the 5 class widths equally). An increasing number of candidates are using the easier calculation method, $\sum \frac{O^2}{E} - 450$, to calculate the test statistic. Few errors in finding the degrees of freedom were seen, but there were a number of slips made by candidates using 10% or 0.1% instead of 1% level of significance. The test was almost always carried out correctly, but, once again, too many candidates lost the final mark by failing to give their answer in context.

Question 6

Many good attempts were seen in part (a) despite the fact that candidates needed to work backwards and despite the algebra involved. It was quite pleasing to see a large number of correct attempts at an expression for Var(*W*). In attempting to standardise one common mistake was to use the variance another was to simplify $\sqrt{36+16\sigma^2}$ to $6+4\sigma$. At this level, most recognise the requirement to set their standardisation equal to a sufficiently accurate (4dp or better) *z*–value. A significant minority of candidates lost the final two marks in this part by solving for σ^2 instead of σ.

Though a good start was made to part (b) with many correctly calculating E(*B*) and Var(*B*), the final 4 marks of this part discriminated the most able candidates. Surprisingly at this level, candidates still show a lack of understanding of conditional probability. Many attempts at subtraction were used rather than applying a correct ratio method. These attempts generally only scored 1 of the final 4 marks. Those candidates using correct notation and sketching diagrams were more likely to achieve full marks on this question.

Question 7

This proved to be the second most challenging question on the paper with parts (c) and (d) being the most testing. Part (a) was dealt with easily by the vast majority of

candidates; though there are still a large number who could not calculate the unbiased estimate of the variance (often trying to use incorrectly stated formulae). Part (b) received mostly correct answers with most being able to identify that μ prevented it from being a statistic.

Parts (c) and (d) were far more demanding and many candidates made little or no progress in these parts. Many lengthy, overcomplicated calculations were seen, but it was the most concise responses that tended to score the marks in (c). It was pleasing to see some candidates persevere in (d) to show an understanding of bias by subtracting σ^2 from their answer to (c).

Question 8

Question 8 was the most demanding question on the paper with 60% of candidates scoring 2 or fewer of the 7 available marks. Of those that made a valid attempt at part (a), many did correctly identify that $\bar{X}$ would be approximately normally distributed with a mean of 5. The variance was more difficult to obtain and most neglected to divide by 50 in their attempts.

In part (b), use of (±)1.96 was the only mark likely to be scored. Many here attempted to calculate a confidence interval rather than a critical region and most did not use their variance from part (a).

Grade Boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:
http://www.edexcel.com/iwantto/Pages/grade-boundaries.aspx

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