edexcel

Examiners' Report

Summer 2014

Pearson Edexcel International Advanced Level in Statistics S3 (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 2014
Publications Code IA040146

Mathematics Unit Statistics 3

Specification WST03/01

General Introduction

This paper proved accessible to much of the candidature. The work on hypothesis tests was generally very good with most students using population parameters, where appropriate, and giving their conclusions in context.

Report on Individual Questions

Question 1

This question discriminated across students of all abilities. In Q01(a) nearly all students wrote down 168 followed by 8, with only a few incorrectly writing 8 followed by 63.

In Q01(b)(i) many students realised that they would need to sample every 6^{th} member in the membership book, but a number did not explain that the first member of their sample should be selected at random from the first 6 members in the book.

In Q01(b)(ii) only a few students were able to articulate that as the list is not sorted by gender then a systematic sample taken from it may not give a sample that represents the proportions of males and females in the club.

A number of students in Q01(c) failed to mention the need to label the males and females separately but most did state the need to select their samples using random numbers.

In Q01(d) a number of students did not realise that the question required advantages for using a stratified sample rather than a quota sample. So common responses such as "a stratified sample can reflect the population structure" or "estimates can be obtained from each of the strata" did not score any marks. Most students who were successful in answering Q01(d) usually stated that "a stratified sample is not biased".

Question 2

A significant number of students did not attempt Q02(c). In Q02(a), the majority of students were able to show that $\mathrm{E}(\bar{X}) = \frac{3\alpha}{2}$, but a significant number did not then state this result implied $\bar{X}$ was a biased estimator of α or that $\frac{3\alpha}{2} \neq \alpha$. The majority of students, however, correctly deduced a bias of $\frac{\alpha}{2}$.

In Q02(b) whilst many students wrote down the correct $k = \frac{2}{3}$, a significant number incorrectly stated $k = \frac{3}{2}$.

Only a minority of students in Q02(c), used the data given to find a sample mean of 8, and estimated the highest value of X by applying $2\alpha + 3$ to give $2\left(\frac{2}{3}(8)\right) + 3 = \frac{41}{3}$. Some students, however, used the highest value of 13 in the data set and found $2(13) + 3 = 29$ as their estimate of the highest value.

Question 3

This question was well answered with the majority of students scoring full marks. Most students stated their hypotheses correctly in terms of population parameters, calculated the test statistic correctly, made an appropriate comparison and gave their conclusion in context. Common errors in this question included writing hypotheses in words, writing hypotheses using $\bar{x}_A$ and $\bar{x}_B$, or incorrectly applying the standard error as $\sqrt{\frac{35}{80}+\frac{28}{100}}$ instead of $\sqrt{\frac{35^2}{80}+\frac{28^2}{100}}$.

Question 4

This question was well answered although only a minority of students were successful in answering Q04(d). In Q04(a), nearly all students were able to calculate r_s, although a small minority made errors in ranking or forgetting to subtract $\frac{6\sum d^2}{n(n^2-1)}$ from 1 when evaluating their answer.

There were a good number of fully correct solutions to Q04(b) and Q04(c). Common errors in Q04(b) included writing hypotheses in words and not in terms of ρ, or not writing their conclusion in context.

Most students who were successful in answering Q04(d) stated that the data was not from a bivariate normal distribution. Few appreciated that the results of their hypothesis testing in Q04(b) and Q04(c) meant that there could be positive non-linear correlation between blood pressure and weight. This is because the product moment correlation coefficient, *r*, was measuring the degree of a linear relationship and because Spearman's coefficient used ranks.

Question 5

This question was well answered with the majority of students scoring full marks. Q05(a) was a good source of marks for nearly all the students. The only common error seen was in the statement of the hypotheses and the conclusion, which were sometimes incomplete.

In Q05(b) many students either stated that a 0.5% significance level would change the conclusion to Q05(a), or stated that this would mean that "there is no association between drink preferred and gender". Some students did not justify their statement by writing down $\chi^2_2(0.005) = 10.597$, whilst others found an incorrect critical value from their tables.

Question 6

This question was well answered. In Q06(a) some students failed to provide sufficient evidence in order to show that the probability of a randomly selected task being failed is 0.3. Some did not show where the 300 came from in $\frac{300}{1000}$, whilst others did show that the denominator was 8×125, with only the most careful students explaining how both numerator and denominator were found and securing both marks.

Q06(b) was well done although some did not give their answers to 2 decimal places. Few students incorrectly found *s* by multiplying 125 by $\mathrm{P}(X = 6)$, even though, *s*, on the table, corresponded with 6 or more failed tasks.

The test in Q06(c) was also done well, but a number of students included 0.3 in their hypotheses even though this was incompatible with their correct treatment of the degrees of freedom. The pooling of the last two classes was usually carried out and the calculation of the test statistic was often correct. A few, however, attempted to pool the first two classes in the mistaken belief that observed values were required to be at least 5. Most students correctly subtracted 2 from the number of classes after pooling to find their degrees of freedom, although some subtracted 1.

Most students who were successful in completing their test in Q06(c) usually used the result of their test to give a correct explanation in Q06(d).

Question 7

This question was well answered. Most students were successful in answering Q07(a). The most common error was to simplify $\mathrm{Var}(4X - 3Y)$ incorrectly to give either $16\mathrm{Var}(X) - 9\mathrm{Var}(Y)$, $4\mathrm{Var}(X) + 3\mathrm{Var}(Y)$ or $4\mathrm{Var}(X) - 3\mathrm{Var}(Y)$.

Again, many students were successful in answering Q07(b). The majority started by finding $A - C \sim \mathrm{N}(5, 27 + \sigma^2)$, although a few incorrectly found $\mathrm{Var}(A - C)$ as either $27 - \sigma^2$ or $81 + \sigma^2$. Although the majority of students wrote down a *z*-value equation such as $\frac{-5}{\sqrt{27 + \sigma^2}} = -0.8416$, a significant minority struggled at this stage with some replacing -0.8416 with either 0.8416 or 0.5793 (which is $\Phi(0.2)$). Some students, who did not use the percentage points table, wrote $z = -0.84$ or -0.85 and lost a mark.

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 Edinburgh Gate, Harlow, Essex CM20 2JE