Pearson

Examiners' Report
Principal Examiner Feedback

Summer 2017

Pearson Edexcel GCE Mathematics

Statistics S3 (6691)

edexcel

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 2017
Publications Code 6691_01_1706_ER

Statistics 3 (6691) – Principal Examiner's report

General introduction

Overall this proved to be an accessible paper, giving the best students the opportunity to score the majority of the marks available. The required methods were well known by a significant number of students. Hypothesis testing questions were answered well with good use of context. However many students are still confused about whether they have proved the null hypothesis to be true or have not shown the alternative hypothesis to be true for a particular test, Some students lost marks needlessly by failing to show intermediate working in the longer questions.

Question 1

Part (a) made a reasonable start for many students. However, there were those who clearly understood the term 'stratified sampling', but missed the point of the question. The question required an advantage of stratified sampling. For example, some students stated that the proportions of each job role are the same in the sample as in the whole population. This explains how a stratified sample is created but not why.

Part (b) was answered very well by the majority of students.

There were many excellent answers to part (c) and typical answers in this category were succinct and precise. All the required detail was often contained in two short sentences. In contrast, there were some lengthy paragraphs from other students that did not always score both marks. A minority of students wrote vague responses. The question required some indication of how a random sample was to be achieved, using either random number tables or a random number generator on a calculator or a computer.

Question 2

Part (a) was answered very well, but many students ignored the instruction to state their hypotheses clearly for part (b). Not as many students' minds were as focussed on this aspect of the question as hoped. Two options were possible, and many successful instances were seen of both. While 'continuous uniform distribution"' was seen frequently, other students missed out either 'continuous' or 'uniform'. Other students wrote 'discrete uniform distribution'. Some students successfully employed the frequently useful technique of restating the words in the question; 'the spinner is equally likely to come to rest at any point between 0 and 360 degrees'. Unfortunately, there were, however, a few students who reversed their hypotheses by using this phrase for their alternative hypothesis.

Question 3

Most students found part (a) straightforward scoring full marks with arithmetic slips or incorrect use of formula rarely seen. Pleasingly, reversed ranks were rarely seen. The most common error from a number of students was to 'code' the letters (A=1, B=2, etc) and use these as ranks, leading to a wholly incorrect solution.

In part (b), there were many good answers here with most students able to identify the correct critical value from the tables and compare with their value to reject the null hypothsis. Most had correct hypotheses with very few in terms of *r* or simply stating them in words. The most common mark to be missed was for the final conclusion lacking some context e.g. no referenced to Judges, or referencing correlation but omitting 'positive'.

Part (c) was usually correct stating in various ways that the training was not effective.

Question 4

Part (a) was fully correct for the majority of students who were able to state their hypotheses using the correct context and key words. However, some are still incorrectly writing there is an 'association'. All but the very weakest of students were able to calculate the correct test statistic, but students should be advised to check every calculation thoroughly as incorrect answers were generally down to a misplaced digit in one of their calculations.

Part (c) was more of a challenge. Those who attempted it generally managed to get to a correct test statistic, but were then unsure how to proceed or simply stated 1% was the lowest significance level with little understanding of how to check this. If a student set off on a route with O and E the same, this method usually resulted in no marks being awarded.

Question 5

Most students were able to gain the marks for part (a). Very few found the sample variance without multiplying by $(n+1)/n$.

In part (b) it was most common for students to gain the first three marks here for a correct attempt at the confidence interval, with very few not using 1.96 for the z-value. A small number lost an accuracy mark for using their s instead of the stated standard deviation of 10. It was very common however for the CI not be stated in terms of the arrival time as required in the question, with most leaving it in terms of minutes late, hence missing the final mark in part (i). Part (ii) was not well answered with few giving assumptions about the bus arrival times. Often they just referred to 'values' being independent with no context. A number of students did not understand what they were being asked in this context and stated various textbook answers referring to the Central Limit Theorem.

Part (c) was the least successful part of the question with only a minority explicitly stating that the stated arrival time was in the interval. Some trivially stated that the mean was in the interval. Some students focussed solely on the bottom end of the interval being negative or below the stated time, without stating the time was contained in the interval. Many students felt that the most significant feature was that the time was near the lower limit so that even if they did say it was in the interval they typically decided that Paul was correct and that the bus was more often late i.e. they did not demonstrate an understanding of the meaning of a confidence interval.

Question 6

The challenge for part (a) was interpreting the second sentence in the question and forming the correct hypotheses. Some students omitted the 8 from one or both of the hypotheses. The calculation of the test statistic was invariably performed correctly. Some students lost the final mark as their conclusion was incomplete. These students should follow the example of those successful students who repeated the same words as the question. The most common reason for the loss of this final mark was, however, incorrect hypotheses. Most students were familiar with the Central Limit Theorem in part (b). The first mark was often awarded. The second, however, was earned more rarely. A few important details were required. Students frequently omitted the word 'approximately'. Mention of both sample means was essential: students often referred simply to 'the sample mean'.

Question 7

This proved to be a discriminating question, especially part (c).

Parts (a) and (b) were primarily testing whether students could distinguish between '3 randomly chosen medium bags' and '3 times the weight of a randomly chosen medium bag'. A number of students made this distinction by identifying and using variances of $3Var(M)$ and $(3^2)Var(M)$ respectively in their working. It was not unusual for students to misinterpret the wording in the question and wholly or partly use the technique required in part (b) in part (a).

Part (c) was an interesting challenge. Students were confronted with one single sentence densely packed with both information and instructions. However, a number of excellent solutions were seen. There were some students that provided unsatisfactory responses to the relatively easier parts (a) and (b), but who were able to earn full marks in part (c) with a clear, concise solution.

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