

Examiners' Report/
Principal Examiner Feedback

Summer 2013

GCE Statistics S3 (6691)
Paper 01

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from 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 2013

Publications Code UA037010

All the material in this publication is copyright

© Pearson Education Ltd 2013

Statistics S3 (6691)

Introduction

The paper was accessible to most candidates and questions 1 (on χ^2) and 2 (on Spearman's rank correlation) were answered very well. The test for a difference between two means in question 6(a) and the combined sample in 7(b) proved more discriminating. Most candidates had a good grasp of the topics on S3 but the quality of their written communication was sometimes quite poor and lacked sufficient precision to secure the marks.

Report on individual questions

Question 1

This proved to be a friendly opening to the paper with very few failing to show sufficient working and many scoring full marks. Some lost marks for the hypotheses either through laziness (simply stating “no association” for the null hypothesis is not sufficient as we want to see the variables under consideration being mentioned) or for stating them the wrong way around. The calculations were usually correct but some mistakes occurred in the degrees of freedom and the conclusion was not always given in context.

Question 2

The calculation of r_s was, as usual, completed very well with only small minorities making slips in ranking or forgetting the “1 – “ when evaluating their answer. There were a good number of fully correct responses to part (b) too although some still fail to state their hypotheses in terms of ρ . A number of candidates did not ensure that their critical value matched their alternative hypothesis and of course there were the usual crop of conclusions that did not include a reference to the context.

Question 3

Most candidates had a stab at part (a) but their responses were sometimes rather too brief or vague. Many quoted standard responses from textbooks or past mark schemes and, although a suitable pair of such answers was acceptable on this occasion candidate should really be aiming to address the question with reference to the context in the question. For example the access to a database means that a sampling frame is readily available.

The candidates were on much more familiar ground in part (b) and most scored full marks although a few failed to round 33.7 and 4.3 to suitable integers.

In part (c) there was some confusion with systematic sampling here and others failed to mention the need for a suitable sampling frame for each course although many did score the mark for using random numbers for the selection from each course.

Question 4

In part (a) the mean was usually correct but some candidates were not sure how to find the value of s^2 . Most knew how to start part (b) and used the normal distribution correctly to find the value of a or b . However, rather than using the fact that the sum of the expected frequencies must add up to 50, they used the normal distribution again to find the other value and then did not obtain a pair of values that added to 21.15. The most frequent error in part (c) was to include the 5.49 or the 6.88 in the hypotheses but apart from this most knew how to calculate the test statistic and the final 4 marks were often scored too.

Question 5

Most candidates knew the difference between these two parts but some failed to use clear notation and a variable $X = L - 3S$ was often referred to in both parts but used as $L - (S_1 + S_2 + S_3)$ in part (a). Many lost the first mark in part (a) for this deficiency but often went on to score all the remaining marks. Candidates should be encouraged to define their new variables carefully and then clearly state the distribution and the required probability.

Question 6

The hypotheses in part (a) caused several problems. Many missed the 1 in the null hypothesis and it was not always possible to determine whether their alternative hypothesis was correct either: a simple statement $\mu_1 - \mu_2 > 1$ is no use unless it is clear what variable 1 refers to. Candidates should ensure that they clearly define their variables in questions of this type.

There was some confusion between standard deviation and variance when calculating the standard error for the test statistic but many did achieve this only to lose marks when finding the numerator by failing to subtract the “1”.

In part (b) many candidates simply quoted textbook definitions of the Central Limit Theorem (CLT) without referring to this situation. The examiners were looking for a comment that mentioned that because the sample size is large, the CLT can be applied and the mean yields of tomatoes can be assumed to follow a normal distribution. Most candidates know that CLT enabled one to assume a normal distribution but some failed to mention the word “mean” or refer to \bar{X}_{new} .

Question 7

Most candidates answered part (a) correctly with only a small minority failing to use the unbiased estimate for variance. Part (b) though discriminated very well. The majority could deal with the mean and obtained an answer rounding to 179 (although a few took the sum or the average of 4.16 and 4.55) but only the very best candidates could find the combined value of s^2 . Some simply added the two values of 0.411 and 0.25, others tried a more sophisticated weighted average such as that used in calculating the pooled estimate of variance in S4. Many candidates knew how to obtain an estimate of the standard error using their combined value of s but few reached an answer rounding to 0.0862 as required.

The final part marked a return to more familiar territory for most and many gained full marks here and even those who had failed to combine the means correctly would often score the M1 and B1.

Grade Boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:

<http://www.edexcel.com/iwant to/Pages/grade-boundaries.aspx>

Pearson Education Limited. Registered company number 872828
with its registered office at Edinburgh Gate, Harlow, Essex CM20 2JE

Ofqual



Llywodraeth Cynulliad Cymru
Welsh Assembly Government

