Examiners' Report/
Principal Examiner Feedback

Summer 2013

GCE Statistics S3 (6691)
Paper 01R

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 UA037012

Statistics S3 (6691R)

Introduction

The paper was accessible to most candidates and question 3 on Spearman's rank correlation was answered very well. There was only one question on χ^2 tests (question 4) and although the first three marks were quite accessible the goodness of fit test (and the bulk of the marks) were not answered well and weaker candidates lost out on what is often a good source of marks for them on this paper. Questions 5 and 6 also proved more discriminating but enabled the stronger candidates to shine. Most candidates had a good grasp of the topics on S3 but the quality of their written communication was sometimes quite poor and sometimes lacked sufficient precision to secure the marks.

Report on Individual Questions

Question 1

There were three stages required here and many candidates missed at least one of them. A number of candidates 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 and the calculations (45 males and 15 females) were usually seen although a few only mentioned one group.

Question 2

Although many candidates clearly defined their variable there were some who wrote $X \sim \mathrm{N}\left(40, \frac{9}{n}\right)$ but clearly meant $\bar{X}$ as their standardization indicated. Most used a z value of 1.6449 but a number used an incorrect inequality and ended up with $n <$ 6.087…and gave a final answer of 6.

Question 3

Part (a) was answered very well and most scored full marks too for the hypothesis test in part (b) but a few failed to give their hypotheses in terms of ρ. Part (c) was answered well but then in part (d) few knew how to interpret these results. A handful realized that the product moment correlation coefficient (pmcc) was measuring the degree of a linear relationship and because Spearman's used ranks a non-linear relationship between the variables could be present. A few more mentioned that a bivariate normal distribution is needed to use pmcc but this assumption is not required for Spearman's.

Question 4

The χ^2 tests were only examined in this one question on this paper and a number of candidates failed to identify the fairly straightforward goodness of fit test in part (d). The first three parts were answered well and most scored full marks although a few mis-read the tables and thought α was 1 and some gave α as 0.025. In part (d) the common error was to simply calculate the test statistic for a test of association between hair colour and eye colour rather than using the column totals from Table 1 to conduct the goodness of fit test on the hair colour ratios. Those who did attempt the correct test could usually calculate the expected frequencies and often completed the test accurately.

Question 5

In part (a) most were able to find a correct expression for $\bar{x}$ and use the z values 1.6449 and 2.3263. There was some strange notation used for the standard error but most used their expressions correctly to find an accurate confidence interval.

In part (b) some failed to use the relationship between diameter and circumference of a circle and simply wrote down the same confidence interval but the answer to part (c) was often correct.

Question 6

In part (a) many candidates calculated Var(X) = 3 but few gave $\bar{X} \sim \mathrm{N}\left(a+2, \frac{3}{50}\right)$ with the majority having a mean of 17.2 rather than a + 2. Many found a correct confidence interval for a + 2 but few appreciated this was what they had obtained and only the better candidates found the correct confidence interval for a by subtracting 2 from each of their previous limits.

Question 7

In part (a) it was not always possible to determine whether their alternative hypothesis was correct: a simple statement $\mu_1 - \mu_2 > 0$ 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 in future. Apart from this though most candidates could calculate a correct test statistics and interpret it appropriately.

In part (b) most candidates know that Central Limit Theorem enabled one to assume a normal distribution but some failed to mention the word “mean” or refer to $\bar{X}$.
The assumption in part (c) was often correct but some simply said “independence” but this was implied by virtue of the fact that the samples were both random and the second sample was chosen from a different part of the field.

Question 8

There were many good answers to this question although few scored full marks. In part (a) some used one chicken egg and one duck egg rather than 2 duck eggs and a number calculated the incorrect area sometimes it was only one of the “tails” and sometimes it was the “central region”. Part (b) was usually answered very well with a new variable based on $\frac{4}{5}D$ and C being a popular approach. The final part caused some problems. A number of candidates did not use a variable for the sum of 6 duck eggs and a box and a second variable for a sum of 6 chicken eggs and a box and so when they found the difference in these variables their variances often only included one box or no boxes. Those who did define their variables carefully usually arrived at a fully correct solution with few difficulties.

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

Ofqual

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

cea
Rewarding Learning

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