edexcel

Examiners' Report/
Principal Examiner Feedback

January 2013

GCE Statistics S1 (6683) Paper 01

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the world's leading learning company. 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 for our BTEC qualifications.

Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

If you have any subject specific questions about this specification that require the help of a subject specialist, you can speak directly to the subject team at Pearson. Their contact details can be found on this link: www.edexcel.com/teachingservices.

You can also use our online Ask the Expert service at www.edexcel.com/ask. You will need an Edexcel username and password to access this service.

Pearson: helping people progress, everywhere

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

January 2013
Publications Code UA034848

General Introduction

The paper proved to be accessible to candidates of all abilities. Some candidates are still unsure about the cumulative distribution function and Q2(b) posed some difficulties. In the final part of question 4, few candidates identified the conditional probability in the context of normal distributions but the conditional probability in the more familiar context of question 7 was answered quite well. Calculating the height of a bar in a histogram still causes problems in Q5 (a) and many did not calculate the final probability in question 6 correctly. Overall, the standard of work was high and most candidates were clearly well prepared for this paper.

Comments on individual questions

Question 1

Q1 (a) proved an accessible opening to the paper and nearly all the candidates answered this correctly. Most also knew how to find *r* in part (b) but a number still gave their final answer as -0.71 rather than the 3 significant figures requested. In Q1(c) many identified that the required correlation would be positive and gave a simple argument based on the context, although a few linked the variables *v* and *g* via the third variable *t*. A number of candidates seemed to misread this part of the question and gave a description of the negative correlation between *t* and *v*.

Question 2

Although over 40% of the candidates scored full marks on this question there were a large number who were still confused about the cumulative distribution function F(*x*). In Q2(a) many used F(1) + F(2) + F(3) = 1 and arrived at $k = \frac{4}{3}$ and others did not give sufficient working to secure both marks. A "show that" question requires some clear explanation and in this case we needed to see an explicit reference to F(3) = 1 as well as the solution of the equation $27 + k = 40$. Due to little understanding of the meaning of F(*x*) many couldn't find the correct probability distribution in Q2(b) but there were plenty of correct answers seen often given in a table. Q2(c) was answered well and the Var ($aX + b$) formula was known and used correctly by a good number of candidates.

Question 3

Q3(a) was answered very well and only a handful of candidates did not secure the 4 marks here. Most knew how to find the equation of the regression line but sometimes candidates failed to use a sufficiently accurate value of b to ensure that their value of a was accurate to three significant figures and they therefore lost the final accuracy mark. A growing number of candidates are giving their coefficients as fractions, presumably because their calculators are set in this mode. Whilst such answers were accepted, they are not as useful as coefficients of a regression line and arguably not really appropriate in this branch of statistics.

Q3(c) was answered very well but in Q3(d) some candidates' responses were vague: a comment that "it is reliable because it is in the range" was not accepted because "it" does not clearly refer to the temperature. Some candidates used the technical terms of "interpolation" or "not extrapolation" correctly and these were accepted.

Question 4

A small minority were still unsure whether the final answer was 0.9641 or 1 – 0.9641. There were 3 common sources of error in Q4(b). Some candidates simply set their standardised expression equal to 0.1 or 0.5398 and lost all 3 marks. Others realised that the standardised expression should be set equal to a z value but did not use the percentage points table and lost a mark. The final problem was choosing the correct sign on their z value and a number of answers of 119 were seen. Some candidates gave an answer of 80.776… from their calculators and gained all 3 marks.

Q4(c) was not answered well and most attempts did not notice the usual prompt (the wording "given that…") and thus did not attempt a conditional probability. Common solutions were simply a calculation of $P(L > 133)$ or $P(127 < L < 133)$.

Question 5

In Q5(a) there is still some uncertainty about the concept of histograms. Most candidates could find the width of the required bar but a number still failed to consider frequency densities to find the height. Some attempts were on the right lines but used the interval [240, 320); others assumed height was linked to frequency alone and 4.5 cm was a common incorrect answer. There has been a clear improvement in the use of interpolation to calculate the median in recent sessions and many had $\frac{10}{22}\times 80$ (with only a small minority using 10.5 instead of 10). The common error was to use 239.5 rather than 240 as the class boundary but there were many fully correct answers seen.

The calculation of the mean in Q5(c) was usually correct but the standard deviation still causes difficulties for some who failed to use a correct formula. Q5(d) was straightforward although some failed to secure the second mark because they did not give a description (e.g. “positive”) of their skewness.

In Q5(e) the insertion of a normal calculation in this context seemed to fox some candidates but many were able to carry out this routine calculation correctly. However the interpretation in Q5(f) was not answered so well. Some drew the connection between their answer to Q5(e) and the proportion in the table and deduced that a normal distribution was reasonable (over this range). Slightly more focussed on the fact that in Q5(d) they had stated that the distribution was skewed and therefore a normal distribution was probably not a suitable model (overall) but few were able to marry these statements up and secure both marks.

Question 6

Q6(a) was answered well and most gave the distribution in a table with values 1, 3 and 5 and probabilities of $\frac{1}{3}$ for each. Labelling each face as 1, 1, 3, 3, 5, 5 was quite common too, usually accompanied by probabilities of $\frac{1}{6}$ for each face, but some gave probabilities of $\frac{1}{3}$ for each face and others had values of 1, 2 and 3.

Q6(b) was not answered well with many candidates only having one of the two words "discrete" and "uniform" in their answer. An answer of "discrete random distribution" was quite common but some candidates did not know the correct terms, describing it as an "equal" distribution.

Very few candidates used the symmetry of the distribution to "write down" the answer to Q6(c) but they were still able to gain the mark by calculating E(*B*). Some candidates could not answer part (d) correctly but in part (e) although a number found $\mathrm{E}(R^2) = \frac{34}{3}$ some thought this was Var(*R*), others only subtracted 3, despite having previously quoted a correct formula, and some rounded their answer to 2.33 without first giving the exact value. In the final part of the question many candidates did not seem to understand what was required. Some drew complete tree diagrams but then did not realise that Avisha always chose the die that gave her the best chance of winning. Others identified which die she should choose without taking account of the probability of 0.5 for the outcomes of the spin of the coin. Nevertheless a small number of candidates did complete this part of the question successfully.

Question 7

Q7(a) was answered very well with only a small minority adding $\mathrm{P}(A \cap B)$ rather than subtracting it. In Q7(b) although most could quote a correct formula, some thought that $\mathrm{P}(A' \cap B') = 1 - \mathrm{P}(A \cap B)$ and then had a numerator greater than their denominator. Others assumed independence and used $\mathrm{P}(A' \cap B') = \mathrm{P}(A') \times \mathrm{P}(B')$. Q7(c) was usually correct although a small number of candidates confused independence with mutually exclusive and gave an answer of zero.

Most candidates gave a correct arrangement of the events in Q7 (d) and few omitted the "box". There were a number of errors in the calculations of the other probabilities, in particular the probability for *B* only, but a large number of fully correct diagrams were seen. Q7(e) was usually answered quite well with many demonstrating a correct understanding of the method by using the values from their Venn diagram and others avoiding any problems with the Venn diagram by using $1 - \mathrm{P}(B \cup C) = 1 - [0.45 + 0.20 - 0.09]$ or occasionally $\mathrm{P}(B') \times \mathrm{P}(C')$.

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

Further copies of this publication are available from
Edexcel Publications, Adamsway, Mansfield, Notts, NG18 4FN

Telephone 01623 467467
Fax 01623 450481
Email publication.orders@edexcel.com

Order Code UA034848 January 2013

For more information on Edexcel qualifications, please visit
www.edexcel.com/quals

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

Rewarding Learning