# Examiners' Report/ Principal Examiner Feedback

Summer 2013

GCE Statistics S2 (6684)
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 UA036998

**Statistics S2 (6684)**

## Introduction

The paper proved accessible to the majority of candidates and there was little evidence of there not being enough time to complete the paper. There were the usual arithmetic and algebraic errors, but candidates work was better presented than on previous occasions making it easier to follow. The main errors were to write down the probability corresponding to an adjacent position (one place either side, up or down) of the required answer or using the wrong ‘block’, for example looking up in B(30, 0.25) instead of B(25, 0.5).

## Report on individual questions

### Question 1

The opening question, testing the sampling distribution of the median, proved to be both challenging and a good discriminator for candidates. Some responses contained lengthy lists of many, if not all, 27 possible permutations of the possible samples. However, some candidates were able to express their probabilities concisely.

In part (a), many candidates were able to obtain at least one of the two marks available, but common errors included failing to give all three permutations of the triples (1, 5, 5) and (2, 5, 5) when asked to list all possible samples to include extra triples such as (1, 5, 2) in the mistaken belief that this has a median of 5 because it has been written as the “middle number” in the triple.

In part (b), a large proportion of candidates were able to obtain each of the probabilities (0.3)3 with either (0.5)(0.3)2 or (0.2)(0.3)2, though the factors of 3 for the latter two cases were often absent.

Part (c) was more testing. The most efficient approach was to calculate P(M =1) and then use P(M =2) = 1 – P(M =1) – P(M =5). However, much more often, candidates tried to calculate P(M =2) directly, some with success, but often their calculation omitted either some or all of the 6 permutations of (1, 2, 5). Some weaker candidates made attempts at the sampling distribution of the mean.

It was also interesting to notice that a significant proportion of candidates who obtained the correct probability of 0.216 in part (b), either did not include it in their sampling distribution in part (c) or gave a different value in part (c). It was also apparent that the vast majority of candidates do not check their work as their the three probabilities they gave did not add up to 1.

**Question 2**

This was a relatively straightforward question, for which many candidates were awarded full marks.

In part (a), nearly all candidates correctly identified the Poisson distribution, and were generally able to accurately use its probability function However, although a small minority were able to write down a correct expression for $P(X=1)$, they were unable to accurately evaluate it on their calculators.

The responses to part (b) were usually correct, with the main errors being writing $P(X > 2)$ as
$1 - P(X \leq 1)$ instead of $1 - P(X \leq 2)$ and reading the tables incorrectly.

The normal approximation to a Poisson distribution was familiar to virtually all candidates in part (c) The main errors were:
(i) standardising with 90.5 or 90 instead of 89.5, causing a loss of 3 marks;
(ii) using an incorrect value of 56.25 for the variance.

**Question 3**

A large majority of candidates were successful in part (a), the only significant error seen being to write $P(X > 10) = 1 - P(X \leq 9)$ or $1 - P(X \leq 11)$. Occasionally, candidates lost marks through using incorrect tables e.g. Po(1) or using Po(6.5) having written Po(7).

Many candidates found part (b) to be very challenging, with a small minority offering no response. There were a number of elegant solutions to this part of the question, but more often the candidate's working was not clear or not consistent. Competent use of inequalities ought to be an area of concern for some candidates, e.g. $P(X < d) < 0.05$ was not uncommon. Nonetheless a significant proportion of candidates managed to state that $n = 12$ (or $n = 13$), sometimes with little supporting working. Some candidates spoilt an otherwise correct solution by concluding that $n = 11$.

It was pleasing to see that the hypothesis test in part (c) was generally answered well with most candidates achieving full marks or just losing the final A1. The most common errors were: some candidates used either p, instead of $\lambda$ or $\mu$, or no parameter at all when stating their hypotheses; an incorrect application of the continuity correction, using 36.5 instead of 35.5; in the contextualised conclusion, omitting the key words rate / average number.

**Question 4**

In part (a) a large majority of candidates were able to obtain E(X) = $\frac{5b}{2}$ , although E(X) = $\frac{4b-b}{2} = \frac{3b}{2}$ was occasionally seen.

Despite the clear instruction in part (b) to 'use integration' to show that Var(X) = $\frac{3b^2}{4}$ , a significant minority quoted and used the formula $\frac{(b-a)^2}{12}$ , securing precisely 0 marks.

Of those who did attempt the integration, there was a mixed response with some candidates providing a perfect solution, whilst others contained at least one of the following common errors: an incorrect expression for f(x) was used; sometimes [E(X)]2 was not subtracted from E(X2); b was sometimes used for the variable of integration as well as a constant.

This was condoned unless candidates cancelled $\int_b^{4b} \frac{b^2}{3b}\,\mathrm{d}x$ to $\int_b^{4b} \frac{b}{3}\,\mathrm{d}x$ prior to integration, which led to a forfeiture of all the 3 available marks.

Deducing the value of Var(3 – 2X) is a routine calculation on S1, but it proved too demanding for a significant minority of S2 candidates in part (c).

In part (d) the distribution function was correctly obtained by a majority of candidates, with the two most common errors being:
forgetting to subtract F(1), yielding x / 3 on the middle line;
carelessness with the inequalities on the 1st and / or 3rd lines,
e.g. x < 0 or x > 1, x > 5.

The median was generally given correctly as $\frac{5b}{2}$ , usually using the distribution function obtained in part (b). Those who obtained various incorrect answers may have benefited from an appeal to the symmetry of the continuous uniform distribution.

**Question 5**

Again, many exemplary responses to this question with a high percentage of candidates gaining full marks. In finding values for a and b candidates used a number of different methods e.g. finding two linear equations, F(1) = 0 and F(2) = 1 and solving, or using F(2) – F(1) = 1 which gave the value of a and candidates then used F(1) = 0 and F(2) = 1 to find the value of b. Less successful candidates put the value of a in F(2) – F(1) = 1 again and often came up with a value b = 0 or 1. Candidates who struggled to find values for a and b sometimes used an alternative method i.e. $\int_1^x \frac{3}{10}\left(x^2+2x-2\right)$ , which they saw in part (b) to get $\frac{x^3}{10}+\frac{3x^2}{10}-\frac{3}{5}x+\frac{1}{5}$ and used the coefficient of x to give the value for a and the constant for b. No marks were awarded in such cases as the equations but candidates were not penalised for using these values in other parts of the question.

Part (b) was generally well answered. Candidates who were able to obtain values for a and b in part (a) often used the expression given for f(x) and 'derived' a value for a = -0.6 which they then used in F(x) before differentiating and then factorising successfully.

A high percentage of candidates answered part(c) well. The majority of candidates used xf(x), with the f(x) given in part b), and successfully integrated to get the correct answer. A small percentage of candidates lost marks through integrating f(x) or forgetting to multiply through by $\frac{3}{10}$ to get the final answer.

In part (d), the majority of candidates used the method of finding values for F(1.425) and F(1.435), making reference to F(Q1) = 0.25 and making a statement that '0.25 lies between F(1.425) and F(1.435)'. Candidates were less successful in finding the value of the lower quartile when solving F(x) = 0.25 as the majority were unable to solve the cubic equation found. Marks were also lost for quoting incorrect values for F(1.425) and/or F(1.435) without showing substitution. Common incorrect final statements given were 'The lower quartile lies between F(1.425) and F(1.435) or ‘The lower quartile lies between these values’ without stating which values. A small number of candidates referred to the median.

**Question 6**

For part (a) the majority of candidates should have been in possession of the required knowledge and skills. However, there were occasions when reading the question did not appear to have been included in these and full marks were achieved by far fewer candidates than expected. Reading the question carefully before attempting it is vital and failure to do so was penalised here. The question asked for the probability in each tail but the probability 0.9861, corresponding to $P(X \leq 9)$, was often seen instead of the related probability $P(X \geq 10) = 1 - 0.9861 = 0.0139$ , required the by the question. There were still a few candidates using incorrect notation for critical regions: $P(X \leq 1)$, for example, is not a critical region: it is a probability.

The only serious problem with part (b) concerned the conclusions. Some candidates made an effort, but did not include all the required detail to establish a conclusion in context. Other candidates did not even attempt any conclusion, in context or otherwise: they finished their work with a probability. It is important for candidates to refer to all parts of the question posed i.e. the changes to the process, reducing the percentage and the defective articles. Many candidates referred to the latter two comments but missed the first.

**Question 7**

A significant proportion of candidates obtained full marks on parts (a), (b) and (d). The main exceptions were to refer to the Poisson distribution featured in part (a) and use it in part (b). Attempts to use the Normal distribution appeared occasionally in part (d).

Part (c) proved to be a challenge for many candidates. A significant number failed to make a serious attempt at solving the problem. The next most commonly appearing error was made by those candidates who believed that the cumulative Binomial table should be used and gave an answer of n = 30, the closest number to the true answer that appears in the table.

Many candidates adopted a formal approach: $0.9^n < 0.05$ . On the other hand, there were candidates who adopted a less rigorous approach, writing $0.9^n = 0.05$ . They were then able to write $n = 28.4$ , and conclude, on the grounds of common sense, that the answer must be 29. It is pleasing to report that some candidates were able to handle the inequality signs correctly (reversing the inequality when dividing by a negative). However, some did not: they arrived (as the result of incorrect work) at $n < 28.4$ . Many of these candidates, unfortunately, also used 'common sense' to conclude that the answer must be 29, despite the fact that an answer of 29 is not consistent with the statement n < 28.4. In this case the final A mark could not be awarded.

**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