Oxford Cambridge and RSA

GCE

Statistics (MEI)

Advanced Subsidiary GCE AS **H132**

OCR Report to Centres June 2014

CONTENTS

Advanced Subsidiary GCE Statistics (H132)

OCR REPORT TO CENTRES

G241 Statistics 1

General Comments:

The majority of candidates coped well with this paper. A good number of candidates scored at least 60 marks out of 72 and there were quite a number who achieved full marks. There was no evidence of candidates being unable to complete the paper in the allocated time. As in previous years, only a small minority of candidates attempted parts of questions in answer sections intended for a different question/part and most candidates had adequate space in the answer booklet without having to use additional sheets.

Surprisingly many candidates seemed to cope better on the topics which are not part of GCSE than they did on Question 1, which is was a very standard GCSE topic. Candidates performed rather better on the conditional probability question, than in the past, although this topic still causes difficulties for many. The majority of candidates found Q4(ii) very difficult, with the many scoring at most 1 mark out of 3. In Question 5, many candidates did not provide a convincing explanation of why k = 0.09, with quite a number substituting k = 0.09 into the given formula and trying to show that the sum of the probabilities was 1. This was only given credit if there was very convincing working. The earlier parts of Question 7 on the binomial distribution and hypothesis testing was fairly well answered, with many candidates defining the hypotheses correctly, and also carrying out the hypothesis test correctly. In the last part of this question, candidates often found $P(X \le 0)$ for n = 3 but omitted $P(X \le 0)$ for n = 2, and so only scored one mark out of three. Most candidates supported their numerical answers with appropriate working, but when written explanations were required, as in Q6(v), the poor handwriting and in some cases the poor use of English of some candidates made it difficult to determine what they were trying to say.

Fortunately, rather fewer candidates lost marks due to over specification of some of their answers, than in past sessions. The message, repeated in every examiners' report, warning against this seems to be getting through. A number of candidates, did however over specify some of their answers, particularly in Q6(ii), where candidates often gave an answer of 63.416, some adding 'to 3dp', which they thought was appropriate accuracy. Of course it is the number of significant figures rather than the number of decimal places that is important, and giving an estimated mean to 5 significant figures is not sensible and so attracted a penalty.

Comments on Individual Questions:

Question No. 1(i)

Many candidates gained full credit. A common error which resulted in the loss of 2 marks was to plot the correct height but at mid-points. Only a few used the lower class boundaries. Some candidates drew cumulative frequency bars and a small number just plotted frequency against midpoints. Some candidates forgot to label their axes or more often omitted the word "cumulative" on their vertical axis.

Question No. 1(ii)

This part was very well answered with many candidates picking up the follow through marks for correctly identifying the median and quartiles from their mid-point plotted graph.

Question No. 2(i)

The vast majority of candidates were able to correctly construct the tree diagram although it did appear that quite a few needed two attempts (it looked as though there made been some rubbing out under the final version). Only a very small number of candidates omitted any of the required labels or mixed up some of the probabilities, but these candidates were able to gain follow through marks in subsequent parts of the question. A few candidates omitted the middle set of branches, or added extra sets following 'Accept' or 'Reject'.

Question No. 2(ii)

This was generally very well answered.

Question No. 2(iii)

Candidates found this part much more difficult and many gave an answer of 0.096, which is simply the probability that a candidate for the job is retested at least once and accepted, so not a conditional probability at all. This scored zero unless it was as the numerator of a fraction. Other candidates did have a fraction with the correct denominator but their numerator was incorrect.

Question No. 3(i)

The majority of candidates who scored this mark showed that $P(L \cap R) = 0.099 \neq P(L) \times P(R) = 0.033$. Very few candidates gave the simplest explanation which is that $P(L|R) \neq P(L)$. For the former, candidates had to quote the correct probabilities, but for the latter the symbolic representation was adequate, as the probabilities were given in the question.

Question No. 3(ii)

There were three common answers here. The majority correctly obtained 0.099, but some candidates multiplied the wrong probabilities together to obtain 0.033 or 0.0675. Brief working was generally given both for the correct and the incorrect answers

Question No. 3(iii)

Most candidates gained full credit here, often from a follow through of a wrong answer to part (ii). Some candidates failed to subtract $P(L \cap R)$ away from P(L) and P(R) and but were still able to score one mark for the two labelled circles.

Question No. 4(i)

This was generally well answered but those candidates who did struggle with this question often still managed to score the first mark for $^{16}/_{30}$ multiplied by another probability. There were very few over specified answers seen. A very small minority of candidates mixed up boys and girls but still gained SC2. Rather fewer candidates used the combinations method than the probability method, but those who did were usually successful.

Question No. 4(ii)

This part was found to be rather difficult. The most successful method was to add together the probabilities of 'no boys' and 'no girls' then take the sum from 1. However, a significant number of candidates took each probability from 1 and then multiplied the resulting answers, which only scored one mark. Those considering the three possibilities 1g3b, 2g2b, 3g1b, often either omitted the coefficients of 4, 6 and 4 altogether or got at least one of them wrong, usually the middle coefficient, replacing 6 with either 4 or 5.

Question No. 5(i)

A surprising number of candidates could not cope with the algebra required for this part, and whilst credit was given for the substitution method (if all working was shown) it is not a suitable method at this level. A significant number of candidates omitted the summation equal to 1 and so could only gain one mark out of three if their table was correct. A small number of candidates forgot to include the table.

Question No. 5(ii)

This part was very well answered by the vast majority of candidates with many scoring all 5 marks. Solutions were well laid out, formulae quoted, and correct values for E(X) and Var(X) obtained. It is very pleasing to note that very few candidates made the mistake of dividing by 5, as was more frequently seen in the past. Fortunately most candidates used the $E(X)^2$ – $E(X^2)$ method rather than the alternative – these latter often making calculation errors. A number of candidates had wrong probabilities. If their probabilities added to 1 they could still score three marks, but if not only two marks. Candidates should be advised always to check that their probabilities do actually add up to 1 in probability distribution questions.

Question No. 6(i)

Most candidates found the frequency densities correctly. They usually then went on to draw the axes correctly although a few failed to start the frequency density scale at zero or to label the axes. A few candidates used inequalities on the horizontal axis, which attracted a penalty of one mark. The choice of scales on the vertical axis was not always ideal, and this left some candidates vulnerable to drawing the heights at incorrect positions. In particular the height of the first bar was frequently incorrectly plotted at 0.5 rather than 0.55.

Question No. 6(ii)

The calculation of the mean of the grouped data was in most cases accurately performed using correct mid-points. The calculation of the standard deviation was less well executed. Whilst there were many correct solutions seen, some forgot to factor in the frequencies and worked with Σx^2 rather than Σfx^2. Over specification of either or both of the answers caused some candidates to lose one mark.

Question No. 6(iii)

Most candidates scored at least the first two marks. However many omitted the fact that there were definitely no outliers at the top end of the data and/or stated that there were definitely some outliers present at the bottom end, thus missing the final mark.

Question No. 6(iv)

This was generally very well answered.

Question No. 6(v)

For this type of question candidates should be taught to discuss ‘average’ and ‘variation’. Simply stating for example that the mean of A is lower than the mean of B does not attract any credit.

Question No. 7(i)a

This was generally very well answered.

Question No. 7(i)b

Although most candidates answered this correctly, some gave $P(X \leq 12)$ rather than $P(X \leq 11)$, and some found the required probability but then subtracted it from 1.

Question No. 7(ii)

Most candidates wrote down the correct hypotheses using the correct notation. It is encouraging to report that rather more candidates gave a correct definition of p than was the case in previous years.

Question No. 7(iii)

Those candidates who calculated $P(X \leq 13)$ were generally more successful than those using a critical region method. Those who used the latter method often got the critical region wrong, thereby losing credit. In general conclusions were given more clearly than in previous sessions, although not always in context. There was also rather less use of point probabilities than in the past.

Question No. 7(iv)

Many candidates, despite having answered the previous part correctly, reverted to point probabilities in this part, using their calculator to find P(X = 33). This of course gained no credit. Others made a correct comparison (33 < 35) but were not always sure what this meant in the context of the test.

Question No. 7(v)

Most candidates who knew how to tackle this question wrote down 'for n = 3, P(X = 0) = 0.0034 < 0.01'. However many did not then justify their answer by writing down P(X = 0) for n = 2 and thus only gained one mark. There were very few successful attempts using logarithms.

OCR (Oxford Cambridge and RSA Examinations)
1 Hills Road
Cambridge
CB1 2EU

OCR Customer Contact Centre

Education and Learning
Telephone: 01223 553998
Facsimile: 01223 552627
Email: general.qualifications@ocr.org.uk

www.ocr.org.uk

For staff training purposes and as part of our quality assurance programme your call may be recorded or monitored