Examiners' Report

Principal Examiner Feedback

November 2021

Pearson Edexcel GCE

In Mathematics (9MA0)

Paper 31 Statistics

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by 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

November 2021
Publications Code 9MA0_31_2111_ER

Introduction

The entry for this paper was relatively small and it was clear that we were not seeing the full range of abilities this time. Students are becoming more adept at using their calculators to find probabilities and standard statistics and there is evidence of progress in their familiarisation with the large data set where a good knowledge of the variables and their units would yield some easy marks. There are still areas of weakness in handling of notation for example and in interpreting results from calculations using the context of the question.

Comments on individual questions

Question 1

In part (a) many scored the mark by referring to the possibility of bias being introduced by the interviewer or the lack of randomness when using quota sampling. The most common answers that did not score the mark referred to cost or time failing to appreciate that quota sampling is generally quicker and therefore cheaper to implement than simple random sampling. Some thought that quota sampling required a sampling frame which suggested they were confused about the different types of sampling. Part (b) was generally answered very well with most stating the distribution, and scoring the method mark, and then evaluating part (i) correctly on their calculator. Part (ii) met with less success as many could not deal with the inequality or were unsure how to use their calculator correctly. Some seemed to confuse significant figures with decimal places: the front of the exam paper instructs then to use 3 significant figures but some only gave an answer of 0.022 here and lost this mark. Part (c) was usually answered correctly and many went on to use the correct binomial distribution in part (d) though some found interpreting "fewer than 3" a problem and a common incorrect answer was 0.924 (from P($X \leqslant 3$). Overall though this proved a good start to the paper and over 40% scored 6 or 7 marks here.

Question 2

Most scored the mark in part (a) by identifying that the correlation was negative. When we ask students to "describe" we are simply requiring a non-contextual description. We will ask them to "interpret" when we require a contextual description. In this case we allowed students to describe the correlation in terms of the variables so answers such as "as x increases y decreases" were accepted. Some thought the correlation was strong and others weak but at this stage all we wanted was the negative description. Part (b) caused some problems with some stating that the scatter diagram showed negative correlation but failing to say whether or not this was compatible with Marc's suggestion. Some students felt that there was no correlation shown by the scatter diagram and did not reconsider this even after correctly answering part (d).

Most answered part (c) correctly though a number "lost" the minus sign between their calculator and writing down the answer on the page. Usually there were sufficient figures given for us to award the mark when the minus sign was included. In part (d) most attempted the hypotheses and the majority used ρ. Some didn't have "= 0" for their null hypothesis with alternative values or inequalities being used and of course some used r or simply wrote the hypotheses in words. We allowed ($\pm$) for the critical value provided the actual value was compatible with their alternative hypothesis and many achieved this mark but few scored the final mark as the conclusion was not related to the context of the question.

Question 3

Although this question involved familiarisation with the large data set nearly 70% scored 3 or more marks here. Many students knew the units used for Daily Mean Pressure (or wrote "Pascals" which was allowed). The coding caused problems for many in part (b) finding $\bar{y}$ was easy enough but then adding the 1010 seemed to elude a great number of the candidates. Another frequently seen error was to add 1010 to Σy leading to the common wrong answer of 40.8. Part (c) was the most successful part with many correct answers seen; sometimes the effect of coding was clearly stated but often, I suspect, it was simply ignored and the correct answer obtained anyway. Occasionally students felt they should add the 1010 on after reaching the correct value of 12.1 and this just lost the first method mark for this part. The final part proved challenging. There was evidence that students knew the relative locations of the 3 sites and often they deduced that it was high pressure and therefore the winds circulated clockwise. The major problem though was that students did not seem to appreciate that Cardinal Wind Direction in the large data set gives the direction from which the wind comes. This led to an incorrect allocation of directions in the table, some students were clearly confused by this and felt that the circulation must be anticlockwise and then a correct allocation of wind directions could, of course, be deduced but this lost the first mark. A small percentage of students though did achieve full marks for this question.

Question 4

This proved to be a good source of marks for most candidates with only 2% failing to score and over 18% scoring full marks. Most students navigated parts (a) and (b) successfully but a few missed the information that "No students read all three magazines" and wasted a lot of time trying to find two equations in p and q. The conditional probability in part (c) proved to be a challenge for some; there were those who effectively used P($E \mid S$) and others were miscopying values from their Venn diagram. Those who failed to make any progress in (i) were often able to pick up a mark for a correct follow through when finding the value of t. Part (d) was the least successful part of the question often due to poor use of the notation. There were a curiously large number of students who confused events with the probability of events; for example we saw a number of cases of students writing $S \cap E' = 0.48$ instead of P($S \cap E'$) = 0.48. Another source of errors was confusing the symbols $\cap$ and $\cup$ in their expressions. There was a demand to show clear working here and this meant that we needed to see probabilities correctly labelled and an appropriate test clearly shown and a number of attempts lacked these features but there were plenty of good solutions to this part too.

Question 5

Most students used their calculators in part (a) with reasonable success but a frequent error was to use 0.1 instead of 0.01 and without any working being shown there was no opportunity to award the method mark. Part (b) was a straightforward probability calculation designed to be completed on a calculator as the specification requires and so it was disappointing that nearly 25% of the candidates scored no marks at all on this question. Identifying the conditional probability in part (c) defeated many students and some of those who seemed to be attempting to use a conditional probability were unable to use the correct notation to write their expression. Those who did start with a correct conditional probability and moved onto a ratio of probabilities often had a correct denominator but a common error was to have P($F > 160$) as the numerator rather than P($160 < F < 175$). Part (d) was a familiar hypothesis test and many students scored something here. The hypotheses were usually stated in terms of μ and most of the time 166.5 rather than 164.6 was used. Selecting the correct model caused a number of problems: some used 164.6 as the mean and others used 7.4 (or occasionally 6.1) as the standard deviation, simply stating the correct model will score the method mark here but a number of students didn't show us the model they were using and we had to award this mark by implication from correct working. Most students successfully used the probability approach (as we would recommend) with a few attempting a critical region or using a z value and comparing with 1.6449 from the tables. The final mark required a correct interpretation of their calculation in context and some students simply stopped after stating that the result was significant and failed to score this final mark.

Question 6

There is a requirement in the specification for questions occasionally to draw together knowledge from different areas of 9MA0 and this was a question aiming to do that.

Many students realised that to start the problem they could use the fact that the sum of the probabilities must equal 1 but fewer than half of the students could move beyond that mark. Those who realised they had a sum of logarithms were often able to arrive at the equation $abc = 36$ but few could devise a strategy to deduce the values of a, b and c. Some stated that all of the values had to be greater than 1 though it was not always clear that they deduced this from the given bullet points in the question. Around a quarter of the students found the correct set of values with varying degrees of explanation on the way. Part (b) was often left blank though the first mark was available without values for a, b and c provided brackets were used correctly for the expression. Nearly 10% managed to score full marks for the question.

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom