LEVEL 3
CERTIFICATE IN MATHEMATICAL STUDIES

1350/2A Paper 2A Statistical Techniques
Report on the Examination

1350
June 2022

Version: 1.0

Further copies of this Report are available from aqa.org.uk

General

There was a reasonable response to most question parts, with nearly all students attempting each question. In particular, students showed good understanding of:

- Using calculators, including the statistical functions
- Calculating probabilities from a normal distribution
- Calculating percentages
- Working with ratios
- Simple random sampling
- Suggesting improvements to graphs

Areas that students found more challenging included:

- Finding a combined mean from two or more different samples (see questions **2e** and **5c**)
- Calculating a confidence interval
- Dealing with an outlier when calculating a regression line
- The theory of normal distribution (notation and the standardised normal distribution)
- Using the inverse normal distribution
- Interpretation of written information and critical analysis

A minority of students gave written answers that were poorly presented, with handwriting on the qualitative responses sometimes difficult to read. There were also some issues with students writing outside of the prescribed answer space, or attempting to write a solution in a space too small for it to fit into. Students should make use of the additional answer space where necessary.

Question 1

This question was generally well answered. The majority of students selected the correct ratio of 11 : 5 in part **(a)**, with the most common error being to select 5 : 11.

In part **(b)**, most students were able to calculate the correct total of 185 spaces not including visitors, but there was some confusion as to whether or not this satisfied the council's requirements, with some students concluding that they had been met. A significant minority of students used one space per property rather than one per bedroom for properties up to 3-bed, leading to an incorrect total of 105.

In part **(c)**, most students started by correctly finding 23% of 80, but some lost a mark by rounding this down to 8 two-bedroom flats instead of rounding up to 9. A small number of students misinterpreted the council's requirements and started by finding 23% of 40. Nearly all students were able to gain at least 1 mark on this part of the question by ensuring that the total number of properties was still 80.

Question 2

In part **(a)**, most students were able to correctly identify two valid improvements to the bar chart. However some students lost marks by making criticisms rather than suggesting improvements, or by suggesting a different type of chart.

Part **(b)** discriminated well.. For the *Morning Record*, some misinterpreted the statement as saying that the electricity generated by renewables exceeded that generated by fossil fuels. Others used the wrong value from the preliminary material, using the overall total of 236 794 GWh as their denominator instead of the 110 221 GWh generated by fossil fuels.

For the *Daily Bulletin Review*, a wide range of different approaches was seen, with many students demonstrating good proportional reasoning and showing sufficient detail in their calculations to score all 3 marks. Of those that were unable to complete this part of the question, most scored 1 mark for using the preliminary material to work out that 44 314 GWh was generated by 'other renewables'.

In part **(c)**, around half of students scored full marks, with most of the others neglecting to convert the cost from pence into pounds, hence obtaining an incorrect value of £60.3 billion. Another common error was to divide the total amount of electricity by 14.4 rather than multiplying.

A variety of approaches was seen for part **(d)**, with most students obtaining the correct answer. However, a significant minority could not apply a correct method for finding a reverse percentage and scored no marks. Some students were unable to identify the correct figure of 26.9% from the preliminary material.

Part **(e)** was generally not well answered, with many students focusing on the fact that the four percentages added up to a value greater than 100%. Others gave descriptions of Bobby's method rather than giving a critical analysis.

Question 3

Only a small proportion of students were able to identify both the correct mean and standard deviation in this question.

Question 4

Most students selected the correct z–value of 1.64 in part **(a)**, with 0.9 and 1.28 featuring about equally in the incorrect answers.

Part **(b)** showed quite a big disparity between students. Many showed a good knowledge of this topic and used their z–value to find a correct confidence interval. Some showed correct use of a formula, while others were able to gain full marks without showing a method by using the statistical functions on their calculators. However, a significant number of students were unable to recall the method and scored no marks.

In part **(c)**, it was necessary to compare the claimed mean value of 120 to the confidence interval found in **(b)**. it was not uncommon to see statements such as "it is outside the interval", which was condoned, but an explicit comparison between the value 120 and the upper limit of the confidence interval would have been clearer. Some students referred to their interval without making any comparison to 120, which scored no marks.

Question 5

Parts **(a)** and **(b)** were both very well answered, with the majority of students picking out the correct dice scores from the data and identifying Carly's sample with a correct reason.

However, only the most successful students were able to score highly in parts **(c)** and **(d)**. Most failed to appreciate that the calculation needed to be weighted because of the different sample sizes. Many students attempted to add the point estimates and divide by 4, or similar, which scored no marks. Students who started a correct approach, for example working out 10×3.6 to find the total of Ali's dice scores, scored at least one mark.

For part **(d)**, students needed to appreciate that a fair dice would be expected to produce a mean score of 3.5 and compare that value to their point estimate in part **(c)**. Most students just stated that the dice must be fair or unbiased, scoring no marks. Some had the right idea, but thought that the mean score on a fair dice should be 3.

Question 6

The vast majority of students correctly stated the value of -1 in part **(a)**.

Most students scored both marks in part **(b)**, although a few lost a mark as they did not comment on both the type (positive) and the strength (weak) of the correlation. A significant minority scored no marks because they did not calculate a value for the product moment correlation coefficient and just based their answer on looking at the values in the table.

Question 7

Responses to part **(a)** showed that the standard notation for identifying a Normal distribution was not well known. A minority of students correctly stated the notation form but many students wrote down values or formulae that did not relate to what was being asked. Others had the correct basic structure, but lost the mark because they omitted the 'N' outside the bracket, or did not square the given value for the standard deviation.

Parts **(b)** and **(c)** were a good source of marks for many students, with some standardising first to find a z–value, and others using a calculator to obtain the probability directly. Common errors included:

- Having the inequality in the wrong direction, for example finding $\mathrm{P}(X > 2)$ instead of $\mathrm{P}(X < 2)$
- Using the wrong sign for the z–value, for example -1.355 instead of 1.355, which gave the same incorrect answer as above
- Confusing heights with a standardised variable, so finding $\mathrm{P}(Z < 2)$ instead of $\mathrm{P}(X < 2)$

Part **(d)** required use of the inverse normal distribution, which students generally found more challenging. Some students set up the equation $\frac{k-1.58}{0.31} = 0.2533$, using the wrong sign for the z–value. Another common error was to start with $\frac{k-1.58}{0.31} = 0.72575$, confusing the z–value with a probability. Each of these scored 1 method mark.

A small proportion of students made no attempt at parts **(b)** to **(d)** of this question.

Question 8

The vast majority of students correctly identified the outlier in part **(a)**.

In part **(b)**, most students were able to use the statistical function on a calculator to find values for a and b. However, many of these did not score any marks because they did not input the correct pairs of data into the calculator. Common errors were to include the outlier or to input the absent students' scores as zero rather than excluding them from the calculation. It should also be noted that no marks were awarded for just writing down the values of a and b; it was necessary to state a linear equation relating y and x.

In part **(c)**, it was expected that students would substitute the given marks into their regression equation and use a calculator to find estimates for the missing marks. Students who did this were usually successful in obtaining the mark for student H. More errors were seen in the calculation for Student L, as some students had x and y the wrong way round, while others set up the equation $49 = a + bx$ but were unable to correctly rearrange this to make x the subject

Some students chose to draw a regression line on to the graph and used this to estimate the missing marks. No credit was given for working on the graph, but students could still score marks if their estimates were sufficiently accurate (correct to the nearest whole number). Students struggled to achieve this level of accuracy, so those taking this approach were generally less successful than those working with the equation.

Part **(d)** discriminated well between students, with most being able to score at least one mark, and a significant number achieving full marks. This was a multi-step question which required students to show working at each stage of their calculations. There were a few possible methods, with some very clearly presented solutions seen. The most common approach was to find the 12 separate values of t, use these to find a value for m and then calculate $0.8m$ and $1.1m$ to find the required grade boundaries.

Some students found the means of x and y separately, which on its own gained no credit. Others did not include the outlier and the two estimated marks in their calculations, despite the direction in the question to do so. Some solutions were seen which identified the correct 4 students (C, E, F, L) but could not be awarded all the marks, either because the method was incorrect or because the working shown was insufficient to justify the answer.

Mark Ranges and Award of Grades

Grade boundaries and cumulative percentage grades are available on the Results Statistics page of the AQA Website.