Examiners' Report
Principal Examiner Feedback

Summer 2022

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

GCSE (9 – 1) Statistics – 1ST0
Principal Examiner Feedback – Foundation Paper 2

General comments
It was pleasing to see students performed well on questions requiring standard techniques such as tally charts (Q01), probability scale (Q03) and two- way tables (Q04). It was also very satisfying to see them attempt the extended response questions; however they would benefit from more practice on this type of question as the quality of responses was variable. Like previous exam sessions students were still less familiar with certain topics e.g. Risk and index numbers.

It is worth mentioning here that centres must impress upon their students the importance of legible handwriting as a few responses were very difficult to read. Students were generally able to attempt the whole paper within the time allowed, although some centres would benefit from reminding students about the time allocation and exam technique. There was a significant amount of evidence to suggest answers were not checked afterwards.

Question 1

The tally chart in part (a) was completed accurately by the vast majority of students. A small number of students clearly appeared not to understand how to tally and instead wrote the numbers on the tally column and a similar number did not include the zero for four, although we did condone this.
In part (b) the vast majority of students were able to state that the mode was '1' although a significant minority made the error of stating that 5 was the mode as there were two frequencies which had a value of 5 or by saying the mode was 9 as this was the largest frequency.

In part (c) most students either scored zero or full marks, though some students scored one mark for writing correct reasoning but did not read the question carefully, instead saying that it didn't support their opinion. Most students scored full marks for supporting their opinion and giving the reason that highest frequency was 1.

Many students were able to give a suitable reason in part (d) with 'only collects data on one morning' and 'small sample size' being the most common answers. Some students argued that the sample wasn't reliable because it showed the scheme wasn't working.

Question 2

In parts (a), (b) and (c) virtually every student achieved the marks for writing down the correct values extracted from the table. Where marks were lost in part (c) it was to only through giving one aspect.

In parts (d) students were asked to describe the relationship. Students who were unsuccessful in writing down a correct answer usually wrote down a reference to one age group. Centres are advised that misunderstanding the demand in the question is a frequent source of error. In this case the demand word 'describe' requires the students to write down a description of the relationship.

Part (e) was successfully answered by students by either stating the figures for both males and females or by giving a comparison.

Question 3

In part (a) the question required students to demonstrate understanding of the probability scale. Students were generally able to place all four letters in the correct position. When incorrect answers were given, it was often because they had placed the letters C and D in the incorrect position.

Part (b) was generally answered well, students often stipulated 'flip it a certain amount of times and do a tally' o.e. Where students failed to score, a common reason was because they had stated just 'do it' a certain a number of times without reference to throwing/flipping. Some students failed to understand that the coin was biased and still stated 'it will be 50/50' or similar. Others included ideas such as a sample space diagrams or probability trees to show theoretical probabilities, again failing to grasp the experimental nature of the question. Maybe the reading of the question was too rushed in these instances.

Question 4

The two-way table was completed accurately by the vast majority of students in part (a).

In part (b) students were asked to work out the size of the angle for electric cars for male drivers. Students who were successful on this question were able to show the required calculation (or equivalent), although lots of incorrect rounding – 21.6 rounded to 21 were seen. A significant number of responses were placed on the pie chart, with no working out shown in the answer space and a significant proportion stated that they had used a protractor, which they classed as their working out.

Part (c) was the first of the extended response questions on this paper worth 5 marks. Students were asked to use the information in the pie charts to assess the two conclusions. The majority of students who tried this question gained between 2 and 4 marks. Many students lost marks by not rejecting the first conclusion and/or accepting the second conclusion. Many were successful in identifying petrol to be the most common, but many failed to give supporting evidence for the second conclusion as they would often refer to the values in the table even though the question had asked them to use the information in the pie charts. The majority of the students failed to make any comment on the reliability of the conclusions as they had forgotten. Centres need to

remind students to go back and read over questions to ensure that they have answered all parts.

Question 5

In part (a) far too many students answered using the phrase 'a sample that is taken at random' or even 'a random sample that is taken simply. They needed to provide an explanation of the word random as applied to a sample, so the word random cannot be used in the explanation. The correct explanation was 'everyone has an equal chance of being picked', but very few students were able to score this mark. Many also described in detail the process to create a random sample (give everyone a number and then generate a random number).

In part (b) very few students scored all three marks on this question, often stating the same bullet point two or three times: e.g. they might not remember to bring it back, they should do it in the health centre, do it online. Most students who scored did so with 'only one day' and/or 'might not return it'. Few students recognised that a questionnaire was a good way to gather opinions, they were more likely to question the motives of the manager in asking biased/leading questions to get the desired results. Many thought the doctors would be a better population to take a sample from. Centres should encourage students to look at the total number of marks available for the question as this will give a hint on what is required, they needed to provide three different reasons.

Students found part (c) challenging where they were asked why it was not a good question after they had been told it was an open question. Many students scored no marks for stating not a closed question, a leading question or simply repeating the question by stating that it is an open question. Many students also gave their own opinion on the question as if they were completing the questionnaire, for example, '5 minutes isn't enough because patients may need longer depending on why they are there'. The most common correct response involved 'too many different answers,' but some just stated 'different answers.' There was the impression that not all students were reading the full question and/or fully understanding the text. A good number of students who scored the mark actually submitted multiple correct responses.

A good number of students answered part (d) well, scoring both marks. Where students lost marks they either failed to include answer boxes at all, or they overlapped and/or failed to include units. Students occasionally included the tick boxes but failed to write a question or, more frequently, repeated the question asked in part (c) (perhaps not understanding the difference between closed vs open or not reading the question carefully enough).

Question 6

This question was the second extended response question on the paper and was worth a total of 6 marks. Three marks were available for reference to the data collection and three marks for the analysis of results. This question was generally poorly answered, with many students not willing to write a long response and it was often left blank. There were a number of responses which only commented on the data collection method and made no reference to the analysis of results. However, the vast majority of students who attempted this question managed to pick up at least two marks for data collection by stating that the sample was for only one day and then by stating that this could be improved by collecting data from other times. The most common mistake for data collection is that ' people may not like to disclose their spending as it is private' or ' you should ask for permission before recording their data'. In terms of the analysis of results, many students did not make any reference to this part or simply agreed that the mean and a bar chart was a good idea but did not attempt to make any critique. Those who scored marks on the analysis part of this question did state that the bar chart was inappropriate, although many incorrectly suggested a pie chart or a scatter graph as an alternative. Those who scored close to full marks or full marks identified that the mean can be affected by outliers and offered the median as an alternative.

Students should take heed of the number of marks available and the number of bullet points that they could comment upon and aim to develop their answers.

Question 7

Where this question was attempted (and it generally was) part (a) was usually answered well. Nearly all students completed 0.6 and 0.4 but a minority failed to realise that a set of branches sum to 1 and once 0.5 was entered on the top branch of the second set of pairs of branches the 'remaining' 0.5 was shared arbitrarily between the remaining three branches, rather than realising each branch was 0.5 or 1 minus 0.5.

Once part (a) had been completed (often successfully) it was disappointing to see so many students answering 0.9 on (b), or the follow through sum of their branches from (a), rather than 0.2, or their follow through product.

Part (c) was very challenging for the majority of students. Most who scored marks here correctly pointed out that 1 (day) out of five = 0.2 (or converted their decimal in part (b) and their fraction in part (c) into 20%). Very few used the product of their theoretical probability from (b) and the number of trials to find the expected outcome. A small number of students who scored well overall got part (c) incorrect because they used 7 days not 5, again failing to read the question carefully.

Question 8

In parts (a) and (b) of this question required students to calculate the crude birth rate and the number of births using the formula that had been provided. The majority of students were able to substitute values into the formula but failed to recognise that the population figure in the table had been given in thousands, which meant both their answers were incorrect and meant they gave the two incorrect answers of 11803 in part (a) and 4.77 in part (b), although we did condone this for 1 mark in both parts. They found using the formula to find the number of birth rates more of a challenge as they needed to rearrange the formula. Incorrect responses included not dividing the answer by 1000, not rounding the final answer to an integer (so losing the final accuracy mark) or just dividing the total population by the crude birth rate. There were a large proportion of students who didn't attempt this second part of the question.

Part (c) of this question was found to be the most accessible and very few students scored no marks and where they did it was usually for just listing figures. Where students did only score one mark it was often for only referring to one aspect e.g. crude birth rate is higher.

Question 9

This topic of absolute risk and relative risk proved a challenge for the majority of foundation students with the majority not understanding relative risk. More able students could calculate the absolute risk in part (a). In part (b) students needed to calculate the relative risk for June compared with April. Very few students were able to calculate the answer, most leaving it blank, there were a wide range of answers but the most common incorrect answer was 0.13 or 13/30.

Students had a very limited understanding of relative risk and could not explain in part (c) what the relative risk of 0.8 showed. Many wrote that 0.8 means 80% chance of rain or purely stated that it meant it would rain less without explaining that it was because it was less than 1.

Question 10

This question was the first of the common questions on the paper (also on 2H). At this stage in the paper, some Foundation students had given up, but considering that the box plot for basketball players was drawn on the grid, this should have given a good hint on how to proceed. It was therefore surprising to see so many blank responses. In part (a) the majority of students who attempted the question scored at least one mark, for drawing the general shape of a box plot with a correct value, usually 170 and 182, as these were given. Many students scored two marks for making a slip on one of the values, usually just being one square away from the correct value and this was often when a student didn't use a ruler. A significant proportion scored full marks.

In part (b) the majority of students scored at least one mark, usually for a correct interpretation that basketballers are taller. Interpretation of the range or interquartile range was rare and skew was very rare. Some students compared the quartiles or minimum and maximum values which do not score any marks. Some students listed values instead of making comparisons which scores no marks. Very few students scored full marks, usually not making a third correct comparison.

In part (c) many students scored full marks usually using reasoning involving secondary data or sample size. Many students scored one mark for unreliable with an attempt at a reason.

Question 11

This was the second of the common questions, again it was common to see that many of the students had given up at this point and left blank responses. In part (a) many students missed the word urban, leading to irrelevant answers. Many students had failed to read the question carefully and had the hypothesis 'people living in urban areas live longer' which we did condone. Some wrote the hypothesis as question which scores no marks.

In part (b) many answers had right idea but missed the key language of the two sets being paired or bivariate or talking about how a scatter diagram can 'show the pattern' but not mentioning correlation.

Very few students understood what an explanatory variable was in part (c); stating 'countries' or even naming a particular country was a popular response. Even students who got a mark for urban population, most didn't get the second mark for correct reasoning.

In part (d) where asked to give a 'statistical' reason whether the scatter graph supported their hypothesis in (a). A large number of students scored 1 mark giving a description of positive correlation but missed the key language of 'positive correlation'.

Part (e) was the most successful part in this question. Many students plotted an appropriate line of best fit through the mean point graph. Where marks were lost it was to not have their line going through the mean point or lines were not splitting the data evenly. Many students again failed to use a ruler and centres are advised to remined their students that the right equipment is needed to draw graphs accurately.

Students had no idea how to interpret the given gradient in part (f). This question was mostly left blank or gave an answer of 'positive correlation.' Some students even attempted to work out the gradient of their line of best fit even though a gradient had been given in the question.

Part (g) was quite well answered; many got that South Africa 'did not fit'. The most successful route was probably plotting the point and then saying it was an anomaly. A few said just 'it's too low' and didn't say they were talking about the life expectancy.

Question 12

The final question on the paper was using index numbers, which foundation students found challenging. In part (a) students needed to use the index numbers to work out an estimate for 2015. A significant proportion of the students calculated the (correct) value for the incorrect year, showing the correct method, but resulting in a wrong answer. Many multiplied by 106.9, forgetting to convert to the correct decimal equivalent multiplier.

Some students used a correct method to establish the correct amount to add to the £4650 (£320.85), but this gave them no marks as they needed to sum the two values. The word 'estimate' in the question confused some students and caused a disappointing number of students to calculate a 7% increase, rather than the 6.9%.

It was very rare to award more than one mark on part (b) coming from a 'correct conclusion with attempt at reason'. This question was worth 3 marks and of those that did understand the question it was infrequent that more than two marks were awarded as students didn't give two reasons.

A lot of students incorrectly thought that 110.8 − 106.9 = 3.9 was the correct answer that was being asked for, again not looking at the number of marks that were available and thinking if they stated Thomas was correct and copied out the question that they would get all three marks.

Summary

Based on their performance in this paper, students are offered the following advice:
- check how many marks are available for each question and ensure that you have given enough points in your answers to gain the marks.
- practice development of extended response questions, laying out answers in bullet points and ensure that you consider all parts of the question.
- use a ruler to draw graphs.
- revise absolute and relative risk.
- revise index numbers.
- if a question asks if you support a conclusion ensure that you give a conclusion in your answer and not just a reason.
- read each question fully and carefully before and after answering to ensure you haven't missed out any details. E.g. on Q04(c) many students failed consider how the data collection method affects the reliability of the conclusions.