Pearson Edexcel

Examiners' Report
Principal Examiner Feedback

November 2021

Pearson Edexcel GCSE
In Statistics (1ST0) Foundation Tier
Paper 2F

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 candidates at: www.pearson.com/uk

November 2021
Publications Code 1ST0_2F_2111_ER

GCSE (9 – 1) Statistics 1ST0

Principal Examiner feedback – Foundation Paper 2

Introduction

General Comments

This report is based on a relatively small number of entries due to the exceptional circumstances of this series. Attempts were made at most questions showing candidates were prepared for the full demand of the specification.

Question 1

Part (a) was generally correct. Of incorrect responses, 'likely' was given most often.

Part (b) was virtually always correct.

Good responses were also seen in part (c), though some candidates gave numerical responses rather than the letters that were required. Part (i) had a slightly higher success rate than part (ii), but on the whole this question provided a positive start to the paper.

Question 2

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

In part (b) some candidates gave worded responses rather than writing down the numerical probability.

It was pleasing to see that most candidates gave two correct comparisons in part (c).

Question 3

Many candidates were able to give a suitable reason in part (a) with 'only collects data on one day of the week' being the most common answer.

Surprisingly only about 1/3 of candidates identified the correct type of data in part (b). Many opted for continuous.

Many identified at least one problem with the completed table in part (c). Comments referring to about 10 not being exact and the missing time period were common.

Candidates were confident with parts (d) and (e) as most were able to complete the bar chart accurately and state the type of art demanded.

Question 4

Many correct answers were seen in part (a) as 'primary' data is well-known by most candidates.

There was some confusion in part (b) as some mistook a data collection sheet for a question on a questionnaire. Others gave an example of a diagram that could be used to represent this information.

Part (c) discriminated more able candidates on this paper as this part required reasoning as well as the correct answer. More successful candidates recognised that this data set included an outlier and therefore the median was more appropriate.

Though many candidates were able to calculate the range in part (d), very few were able to interpret the comparison of ranges in part (e). Many thought that a higher range meant that more people lived in the houses rather than understanding it shows variability.

Question 5

Many strong responses were given to this question. Most candidates understood the need to use the key when completing the composite bar chart. On some occasions, the completed bar only went to a height of 350 showing a lack of understanding of the word 'composite'.

For those completing the composite bar chart, parts (b) and (c) were very well answered.

Question 6

Part (a) was answered correctly by every candidate who attempted this paper.

Many were able to go on to give two comparisons in part (b) as well making this one of the most successful questions on the entire paper.

Question 7

This was another accessible question with many candidates scoring well here.

Parts (a), (b) and (c) were well answered.

Many referred to the position in the table when comparing the performance of Wales and Scotland in part (d). Others totalled the number of matches won. Many correct responses were seen here.

The mode was stated successfully by the vast majority of candidates in part (e).

Mixed responses in part (e) with many not fully reflecting upon what the stem and leaf diagram would look like if it were drawn.

Question 8

Part (a) had a very high success rate with candidates demonstrating well their knowledge of secondary data.

Candidates found part (b) challenging and many struggled to score more than 1 or 2 marks here. Those showing their calculations tended to do well as it made the comparisons clear. Many did not take advantage of the working space and simply offered generic answers without reference to any summary statistics. Candidates do find the open response questions more difficult, but they should pay attention to the clues offered in the question to help structure their responses so that they can gain more marks on these types of questions. Only the most able candidates were able to calculate the interquartile range from the list accurately.

Question 9

At this stage of the paper, stronger candidates were able to show their ability and make good progress on this question. Some overcomplicated part (a) attempting a calculation rather than identifying and writing down the probability from the tree diagram.

Part (b) was even less successful with many candidates unable to process their answer to part (a).

Only the most able candidates gave a complete calculation consisting of the sum of two products. The number of marks available here should be a clue to candidates that a single calculation is not sufficient.

Question 10

The calculation of a mean from a grouped frequency table saw the usual errors including division by 5 rather than division by 35. There were also those candidates who simply found the mean frequency rather than the mean distance. The units ‘miles’ were given on the answer line and should serve as a reminder to candidates that their answer should make sense in the context of the question.

Part (b) was equally challenging. A large number of candidates pointed out that the data for 2001 was out of date – but this was not relevant since we are looking to compare data with 2001. Very few were able to give reasoned answers here.

Question 11

This question on choropleth maps was generally well attempted. Most were able to give a suitable reason in part (a) about the suitability of the diagram.

Most began to describe why the choropleth map supports the conclusion, but many answers were incomplete. Here we needed full statistical reasoning which referred directly to the shading on the choropleth map. Many simply agreed with Tim's conclusion but made little or no reference to the map.

Question 12

Though many candidates are able to describe the correlation, only a small minority were able to interpret it in the context of the question.

In part (b), many candidates did not appreciate the connection between the double mean point and the line of best fit despite the question explicitly stating the need for the line to pass through the point. There were a surprising number of lines drawn with a positive gradient.

Candidates struggled to express the idea of 'extrapolation' in part (c) and on the whole, the final part of this question was not well answered.

Question 13

For this open response question, it was rare to see candidates scoring more than one mark. Most scored for recognising that the hypothesis should not be given in the form of a question. On some occasions, candidates went on to explain that having the participants record their own data may not be reliable.

Candidates 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. Most only used up one or two lines of answer space. This is one area that needs further attention.

Question 14

The final question of the paper was one of the most challenging. Most are unfamiliar with the details of a systematic sample and it was rare to see candidates accessing the marks in parts (a) and (b). Of those who did identify that every kth member of the population needed to be selected, many thought that every 50^{th} plate, rather than a total of 50 plates, would suffice.

Though a fair number of candidates were able to correctly identify the class that contained the median, not a single candidate was able to make a valid attempt at using linear interpolation to calculate the median. A difficult question for candidates at this level, and indeed at higher level too. It was pleasing to see candidates persevering all the way to the end and very few blank responses were seen.

Summary

Based on their performance on this paper, candidates should:

- develop methods for calculating more complex summary statistics (interquartile range)
- understand that the range describes the variability of the variable being studied
- give statistical interpretations in the context of the question
- practice development of extended response questions by focusing on details mentioned in the question

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