Examiners' Report
Principal Examiner Feedback

Summer 2024

Pearson Edexcel GCSE
In Statistics (1ST0)
Higher Paper 1H

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

Summer 2024
Publications Code 1ST0_1H_2406_ER

GCSE (9-1) Statistics – 1ST0
Principal Examiner Feedback – Higher Paper 1

Introduction

General comments

Most candidates responded to the challenges within this paper well and demonstrated understanding of a range of areas of the specification. They were generally confident at completing calculations and diagrams and demonstrated good statistical understanding when asked to interpret these. As seen in previous series, candidates found questions requiring interpretation in context and evaluation of approaches or techniques more slightly more challenging.

Candidates should be encouraged to show full working and set this out clearly so that partial credit can be awarded if a fully correct solution is not obtained. They should also read the question carefully to identify the demand, for example whether an interpretation in context or conclusion is required.

Question 1

This question on interpreting data presented in tables and unemployment rate was attempted by the vast majority of candidates.

Parts (a)(i) and (a)(ii) were answered extremely well with candidates generally able to calculate the unemployment rate using the formula given (in (a)(i)) and draw a conclusion about unemployment rate in the UK based on the figures in the table and their calculated value (in (a)(ii)).

In part (a)(i), a minority of candidates made errors. These included errors with place value, for example 1 290 000 used with 34.04 and the stating of 3.8 with no working which scored no marks as not accurate enough. Those who got part (a)(i) wrong, often followed through the correct interpretation for part (a)(ii).

In part (b), the majority of candidates scored 1 mark only for correctly concluding that Bob was correct together with a partial reason. Most were able to identify that the unemployment rate was lower (with or without calculations) from reviewing the table, matched with giving the correct decision. However, a lesser number of students were able to identify that the total workforce numbers had also decreased or in fact had changed, which was needed to give the complete reason. Where candidates did recognise that the total workforce had reduced this was sometimes misinterpreted as meaning that it was not possible to determine whether or not Bob was correct. Even though the question referenced not using calculations, quite a few students calculated the number of unemployed.

Question 2

In part (a) of this question candidates were required to interpret the grouped frequency table in order to determine the number of Fjords with a length of at least 100 km. Almost all candidates completed this correctly, generally using the most straightforward method of adding the three values. A small minority of candidates gave an incorrect answer of 15 or 17 but did not always show working which meant that method marks could not be awarded. Those who did not read the question carefully most often had 216 as answer from summing 199 and 17 for ‘at most’ 100km.

Part (b)(i) required the calculation of an estimated mean for the length of all the Fjords in the table. This was generally well answered by candidates. Common errors included not using the correct midpoint of the intervals or dividing by the number of groups rather than total frequency. Another incorrect answer was 232/5 which came from the total number of Fjords divided by the number of class intervals in the table.

Candidates were then asked to explain why the answer to (b)(i) was an estimate. Almost all candidates gave a correct response here, using a selection of the responses indicated on the mark scheme. Reference to using the midpoint or not knowing exact values were the most common.

In part (b)(iii), candidates were asked how the accuracy of the answer to part (b)(i) could be improved. This was again well answered. Most who gave a correct answer suggested more groups/narrower groups or use of the raw data. Incorrect answers generally referenced use of more decimal places for accuracy or collecting primary data.

Part (c) asked for a discussion of whether a frequency polygon would be an appropriate diagram to use for the data on Fjord lengths or not. This was less well answered than the previous parts. There were many references to continuous data (linked with appropriate and not appropriate by candidates), but fewer to grouped data. Some candidates were able to describe the process of plotting at midpoints which was sufficient indication that they understood the data was grouped. Very few recognised the that a frequency polygon could be inappropriate due to the large frequency variation. Some candidates suggested alternative diagrams, but this was not what was asked for in the question.

Question 3

In this extended response question, candidates were presented with information about the total populations and population distribution (population pyramids) for France in 2010 and Italy in 2010. They were then presented with two conclusions and asked to assess the validity of these.

Candidates generally did very well on this question. The vast majority were able to gain some marks. A lot of candidates received two marks for successfully identifying 6.6% and 6.8% for the first claim but then many made errors for the second claim. When assessing the second conclusion, many candidates tried to only use the percentages 3.4% and 4.1% rather than calculating the number of males aged 40-44 in France and in Italy. A small minority of candidates made slips in their calculations. If candidates were able to correctly get the first 4 marks, they were likely to also receive the final mark for assessing the two conclusions.

Question 4

The vast majority of candidates were able to use the histogram to correctly complete the table (part a). Where errors were seen these included incorrect inequality symbols in giving the interval.

Part (b) of the question asked for an estimate of the number of runners that took less than or equal to 23 minutes to complete the race. There were a significant proportion of candidates who were not able to identify the calculation required to find this number.

In part (c), the skew was generally correctly indicated as being positive. However, this too often was followed by no attempt at interpretation, an attempt that described negative skew or a simplistic description of the times with the highest frequencies. Incorrect responses along the lines of 'less than half the times were above the median' were frequently seen. Most candidates struggled to express the fact that the times were mainly at

the lower end of the distribution, in statistical language, and as a result, lost marks for being too vague.

Question 5

In the first part of this question, (a)(i), candidates were asked to name the sampling technique described. This was answered very poorly by the majority of candidates; the most common answer was to incorrectly identify quota sampling or judgement sampling; aside from that, candidates gave a full range of other sampling techniques from the specification. Some candidates were able to identify this as either cluster sampling or, more commonly, convenience sampling or opportunity sampling.

Part (a)(ii) asked for two reasons why the sampling technique described (cluster sampling) may not be appropriate. This was well answered with almost all candidates able to give at least one reason and many able to give two. The most common correct responses were 'biased' and 'not representative' with 'not random' being a less commonly given reason. Some candidates attempted to describe reasons why this was inappropriate in context, but the resulting explanations were often unclear.

Candidates were asked to design a closed question for use in a questionnaire in part (b). There were two key aspects that needed to be included – an unbiased question about length of time with a timeframe and at least three non-overlapping, exhaustive response boxes with units. The majority of candidates were able to write a suitable question. However, omission of the timeframe for this was a common error. A minority of candidates responded with a series of unrelated questions like "are you willing to help the company?".

In terms of the response boxes, the majority of candidates made an attempt at these, with most opting for discrete lists rather than inequality notation. A significant proportion of candidates would add an "other box" or signify that a number was to be written in a box. Candidates were generally successful at giving non-overlapping response boxes (although those that ran into difficulty here were those who had used inequalities). The exhaustive aspect was where a greater proportion of candidates made errors - forgetting an option for 0 hours was common and more so than candidates forgetting an open interval at the other end.

In part (c), candidates were asked to suggest strategies which could have reduced the number of non-responses to the questionnaire. This was generally answered well with many candidates able to give at least one appropriate suggestion. The most commonly given suggestion was a face-to-face interview, followed by incentives and making questions as easy/quick as possible to answer. Some candidates did suggest compulsion, but this was rare. Generally, responses concentrated on the sampling technique used rather than the question that was asked. An extremely common, but incorrect response, was to perform a "pilot survey" to gauge the quality of the survey.

Candidates were asked to describe in detail how to take a 10% systematic sample of the employees in the factory (part (d)). The majority of candidates were able to give at least a

partial description of systematic sampling, although some described stratified sampling or simple random sampling instead.

Where candidates did attempt a description of how to take a systematic sample few secured all three marks. Most candidates were able to state that the employees needed numbering or listing. Some candidates were able to state that every tenth employee needed sampling, however, the far more common response was to sample every 6th employee, and it looked like they'd come to this conclusion by finding 10% of 60. Achieving the final mark, for the random starting point, was extremely rare; most candidates seemed to forget that the starting point needed to be random, or they would just state to repeatedly sample using a random number between 1 and n. The second example response given in the mark scheme - to use the spreadsheet - was very rare.

In part (e) candidates were told that a face-to-face interview was to be used and asked how this would improve the quality of responses compared to the questionnaire. This was well answered by a majority of candidates; "questions can be explained" and "follow up questions can be asked" were common correct responses. Where incorrect responses were seen these often referenced the response rate rather than the quality of the response. References to an increased or decreased likelihood of lying were also frequently seen.

Question 6

This question focussed on interpreting a time series graph, calculating mean seasonal variation and using this to make a prediction. Candidates generally attempted all parts of the question with parts (a) and (b) being well answered by the majority and later parts proving more challenging.

Part (a) required candidates to describe and interpret the trend shown in the time series graph. This part was generally well answered with most scoring full marks, or at least one mark for stating a rising trend. A few students did get muddled between 'trend' and 'seasonal variation', particularly interpreting trend using seasonal variation resulting in them repeating themselves for 6a and 6b. A few incorrectly described this as a positive correlation.

In part (b) candidates were asked to discuss any seasonal variation shown in the graph and interpret one in context. The majority of candidates were able to identify that quarter 1 was the lowest each year and quarter 3 was the greatest each year, more often than not with a reference to the number of visitors to Canada from the UK being included for the contextual aspect. Some candidates stated that quarter 2 and quarter 3 had higher numbers of visitors and quarter 1 and quarter 4 had lower numbers of visitors which was accepted. Incorrect responses often referred to 'increasing in 3rd and 4th quarter' but not talking about 'highest' or describing quarter by quarter changes.

In part (c)(i), only a minority of candidates were able to correctly find the mean seasonal variation for Quarter 1. Where a correct process was attempted, there were often errors seen in reading the graph to find the required seasonal variations. Others worked with

positive rather than negative seasonal variations (even though q1s were below trend line) which could gain partial credit. The most common error was taking the quarter 1 values from the time series graph and then working out the mean rather than using the trend line to find seasonal variations.

In (c)(ii), candidates were asked to interpret their answer to (i). This question was almost universally attempted, but not well answered. Where candidates had a fully correct answer to (ii) they were often able to correctly interpret this. Candidates who got a positive answer for 6(c)(i) often misinterpreted it and treated it as a negative for this question part. Many incorrectly commented that their answer was the average number of visitors to Canada or that this was the number of visitors to Canada.

Using the answer to (c)(i) to make a prediction of the number of visitors to Canada from the UK in Quarter 1 of 2020, as required in part (d)(i), was answered well by those who had an attempt at mean seasonal variation in (c)(i). Many were able to recover here from an incorrect sign in part (c)(i), as they realised it was below the trend line by looking at the graph. A common error was to subtract the mean seasonal variation from 200 rather than 210, seemingly due to incorrect reading from the graph.

Most candidates made a comment on the reliability of using their answer to (c)(i) to make a prediction of the number of visitors to Canada to the UK in Quarter 1 of 2025. Candidates were, in the main, able to establish that it was an unreliable prediction. The most common reason seen was extrapolation and this was closely followed by suggestion that the trend not continuing.

Question 7

Part (a) asked for an explanation of why the website 'Office for National Statistics' would give reliable data. A good proportion of the candidates were able to give a correct reason with common correct responses referring to 'government data, 'trusted' or implying large amounts of data. Where unacceptable responses were seen this was generally where candidates repeated parts of the question such as stating 'reliable', 'its official data', 'contains statistics' or 'national statistics'.

Part (b) asked for an explanation of why the suggested hypothesis was not appropriate. This question was very well answered with the majority of candidates indicating that a hypothesis should not be written as a question or that should be written as a statement or both. A very small minority of candidates proceeded to comment on the suitability of the question for the research purpose (for example, "the question is sensitive because asks about age").

Candidates found discussing the suitability of pie charts and comparative pie charts for comparing the proportion of brides in each age group for 2003 and 2013 challenging (part (c)). Only around half of candidates were able to make any correct comments on the suitability of the two potential types of diagram. Only a minority were able to give a full discussion, identifying that pie charts were suitable to show proportions, but that

comparative pie charts would additionally represent the different total populations of brides in the two years.

Many candidates were able to indicate that the difference in population size could be shown by using comparative pie charts, however they often did not indicate that all pie charts would be suitable to represent proportion. Often candidates did not indicate whether they believed pie charts or comparative pie charts would be more suitable. The most complete answers acknowledged strengths in both suggestions whilst still selecting one as the preferred option.

A few candidates were dismissive of pie charts altogether or wanted to suggest an alternative diagram. Quite a few just said comparative pie charts were better because you can compare which was not an acceptable response. Others offered generic key/colour coding or other recommendations for how to draw pie charts or comparative pie charts.

Part (d) required candidates to calculate the radius of one of two comparative pie charts given the radius of the other. Fewer of the candidates were able to correctly find the required radius. Candidates struggled to apply the formula correctly, substituting incorrectly, failing to square the known radius, and/or failing to square root the final answer or getting the fraction upside down. Some rounding errors were seen. A common incorrect answer was 3.57.

In the final part of the question (part (e)), candidates were asked what additional information should be included with the comparative pie charts to help non-statisticians make comparisons. The question was generally well answered, with candidates correctly identifying some appropriate additional information in a slightly majority of responses. The common correct answers included saying to add the total population or percentages. Some candidates focussed on providing information to the audience of what the relative size of the pie charts represented, which was also an acceptable response. Incorrect responses generally focused on minor changes such as adding labels, colouring sectors, avoiding 3D pie charts, add a key or state the angles.

Question 8

The majority of candidates received one mark for part (a) of this question. They could identify that sample means were more consistent but then would fail to give a suitable reason.

In part (b) candidates were asked to plot two sample means on the given quality control chart. The majority of candidates answered this well, plotting both points correctly. The most common errors were mistaking 219.05 for 219.5 in attempting to plot sample 7.

In part (c) candidates were asked to complete the quality control chart and determine what actions, if any, should have been taken after sample 6 and sample 7. If candidates plotted the action and warning lines correctly, they generally went on to gain full marks. It was common for candidates to draw lines but not label them or to draw lines but not in the

correct place. Often lines were drawn at 1 and 2 standard deviations rather than at 2 and 3. Where incorrect lines were plotted candidates were often able to reference at least one correct action with a reason for either sample 6 or sample 7. Marks were sometimes missed by candidates not referring to the samples specifically in their actions or by being too vague in their justifications, others described very general actions but did not consider the two samples given. A significant proportion of candidates struggled on this question, not plotting lines and not being able to comment on the actions that should be taken. Very occasionally, students picked up marks without drawing the lines by explaining where they would be.

Part (d) of the question asked candidates to identify two other types of quality control chart that could be drawn from a given list. A minority of the candidates were able to identify range and median.

Question 9

The first part of this question required candidates to complete a probability tree diagram. The majority of candidates attempted this part of the question and most were able to make at least some progress. Candidates were frequently able to correctly identifying at least the probability of being of time/late on Monday as 0.75 and 0.25 respectively and placing the correct value for the probability of being late on Tuesday (0.3). However, candidates found more difficulty in calculating the probability of being late on Tuesday [P(B)], given the probability of being late on Monday and Tuesday [P(A∩B)] and the probability of being late on Monday [P(A)]. The most common misconception here was to use 0.6 (the value of P(A∩B)) for P(B). A very small minority of candidates also failed to recognize that no calculations were required to place the probability of being late on Tuesday on the probability tree and used the same method used to calculate the probability of being on time on Tuesday (P(A∩B)/P(B) to calculate it.

In part (b), candidates were required to calculate the probability of two independent events happening at least three out of four times. This was generally answered either very well or incorrectly.

A minority of candidates identified the binomial distribution of the events involved (on time/late) and correctly applied the formula to obtain the correct answer. Some candidates who identified the need to use the binomial distribution only worked with the probability of being late three times, with the answer to ${}^{4}\mathrm{C}_{3} \times 0.6^{3} \times 0.4$ (omitting 0.6^{4}) appearing as the most common incorrect answer when attempting to apply the formula. Others who did recognise that this was a binomial distribution did not recognise that the value of p for their calculations was given in part (a), but could gain credit for correct method to use the binomial distribution with the incorrect value of p.

A significant proportion of candidates failed to identify that the binomial distribution was appropriate. In these cases, they often multiplied or added probabilities from the tree diagram in (a). It was not uncommon for answers greater than 1 to be seen which should have indicated to candidates that an error had been made in their method.

Question 10

This question was attempted by most candidates with a slight majority able to make some progress. A small number of candidates were able to give a fully correct response with a similar proportion gaining 5 marks. Candidates were able to accurately calculate the standardised score for Dominic more successfully than the time for Kai. In general, there was a lack of working seen for this question which meant that candidates were relying on not making errors in their calculations.

Having calculated Dominic's standardised score, many could not accurately explain what it meant in the context of time and running a race. A minority of candidates referred incorrectly to being closer to 1 or -1 or to 0. Others compared the standardised scores correctly, $-0.5 < -0.2$, but followed with the incorrect conclusion. Some made a comparison but did not include a conclusion or gave a conclusion without a comparison of standardised scores.

Candidates who were able to calculate Kai's time generally went on to calculate the difference of 2.7 seconds, but then struggled to justify that Kai was correct in his conclusion, with some stating he was wrong.

Summary

Based on their performance on this paper, students should:

- show working for statistical calculations;
- practise writing clear explanations, bearing in mind exactly what is asked in the question and what evidence you should give to support your answer;
- practise interpreting statistical calculations in the context of the question;
- ensure that they can recognise and name different sampling techniques;
- ensure that they can describe how to take different types of sampling;
- practise giving statistical reasons or against the diagrams suggested;
- practise calculating mean seasonal variation from a time series graph and using this to make predictions;
- develop their understanding of control charts and the actions to be taken for samples falling in different parts of these charts;
- practise using the binomial distribution to calculate probabilities;
- practise working with and interpreting standardised scores in a variety of contexts.

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