



Pearson  
Edexcel

Examiners' Report  
Principal Examiner Feedback

June 2023

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](http://www.edexcel.com) or [www.btec.co.uk](http://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](http://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](http://www.pearson.com/uk)

June 2023

Publications Code 1ST0\_1H\_2306\_ER

All the material in this publication is copyright

© Pearson Education Ltd 2023

## **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. 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**

Part (a) of the question required candidates to complete a back-to-back stem and leaf diagram. The majority of candidates were able to construct an ordered stem and leaf correctly; those who struggled often reversed the position of the tens and units. Completing the key for their diagram proved to be challenging with many candidates only displaying a key for the right-hand side of the diagram or attempting to show the 2 different keys but failing to indicate how these related to the two sides of the stem and leaf.

In part (b) candidates were asked to calculate the interquartile range. Candidates usually calculated the IQR correctly using one of the acceptable methods, although there were occasional errors leading to one incorrect value. Some candidates who calculated the positions of the quartiles did not go on to find the related values from the stem and leaf, instead subtracting 6.75 from 20.25, making 13.5 a fairly common incorrect answer.

Part (ci) of the question presented the candidates with conclusions about the relative performance of the team in the 2017 season and the 2018 season and asked them to comment on the validity of these based on the statistics. Candidates were generally able to compare at least one of the medians or the interquartile ranges, and often correctly compared both, however many of them failed to relate this back to the statements and comment on the contextual interpretation. Some candidates struggled when their IQR was incorrect in part b and was lower, as then the medians proved Naomi wrong but the IQRs proved her right. A few candidates also confused themselves with which showed consistency, the median or the IQR.

Part (cii) asked candidates to give a limitation of using the data provided to comment on whether or not the team had improved between the 2017 and 2018 seasons. Many students failed to explain this correctly either resorting to just statistical definitions using averages or sample size, referring to the different amounts of games in each season. Those that wrote in context sometimes

commented on individual players performance or team members changing, which was not a limitation on commenting on how the overall team performance had changed between the two seasons. Some were able to refer correctly to the possibility that the opposition performance had changed, lack of scores for the opposition, or lack of knowledge of the outcome of matches.

Part (d) of the question referred to checking of outliers, by performing calculations, before doing any further processing. Candidates were asked to comment on the appropriateness of this suggestion. This was not answered well, with many answers being vague and some candidates forgetting to conclude whether or not it was appropriate. Most candidates talked about the data, rather than graphs or calculations. Candidates often suggested removal of outliers improved data accuracy with no further information, this did not recognise that these might be genuine data. Correct answers tended to refer to skew or mean calculations, or that it was all 'actual scores' and should be included.

## **Question 2**

In part (a) candidates were asked to decide which of the statistical words given best described weight as a type of data. As would be expected, this was answered very well by candidates.

Part (b) of this question required candidates to complete a frequency table using the information in the histogram. Generally, the question was well answered with most candidates correctly identifying the height of both or either bar. A minority of candidates failed to recognise that the histogram had the same class widths and proceeded to calculate frequency densities which, although correct, would lead to unnecessary calculations and in some cases calculation errors.

Part (c) of this question required candidates to complete the histogram using the table. This question was also generally well answered although most marks were lost for either missing the scale on the vertical axis or for using frequency density values without changing the labelling of the axis to "frequency density". Candidates are expected to be familiar with both histograms with equal width bars (where either frequency or frequency density may be utilised on the vertical axis) and histograms with unequal width bars (where frequency density would be used on the vertical axis).

Part (d) of this question required candidates to provide a contextual interpretation of the negative skew. As with other questions relating to the interpretation of skew, this was not well answered. A common misconception when interpreting the negative skew was to state that most values should be above the median, which is not true, rather than referencing the proportion of values above the mean. Many candidates were incorrectly making comments relating to numbers of players being heavier e.g "more players are heavier". It was often common to see responses which lacked contextual interpretation.

Part (e)(i) of this question required candidates to estimate the mean from the frequency table. The question was generally well answered although there were a number of calculation errors when multiplying frequencies by the mid-interval values, this was especially the case for those candidates who did not use the columns provided to write down their working. A small minority of candidates

divided the total in the frequency column (810) by the number of classes (6). The answer should have raised red flags for candidates (135cm for the height of basketball players) who failed to link the value obtained with its meaning.

Part (e)(ii) of this question required candidates to comment on the mean and was well answered by most candidates who answered part e(i).

### **Question 3**

This question was answered well with the majority of candidates demonstrating a good understanding of time series graphs and trends. There were, however, some candidates who got confused between seasonal trends and the overall trend.

In part (a) of the question candidates were asked to write a suitable hypothesis for the investigation. This was generally answered well with the majority of candidates giving a suitable hypothesis. It was pleasing to note that there were few candidates giving a question rather than a hypothesis. A common error was to give a hypothesis that was not related to the investigation and data collection described.

Part (b) required candidates to identify and interpret one example of seasonal trend shown in the time series graph. The majority of candidates were able to correctly identify an example of seasonal trend and often immediately related this to registrations for the interpretation aspect. Only a small number of candidates identified the seasonal trend but did not link this to the context given. Candidate should be encouraged to use the quarters to refer to seasonal trend from time series graphs and take care over any attempt to translate to seasons or months. A common error was to attempt to describe the overall trend.

In part (c) candidates were asked to explain why using 4-point moving averages was appropriate. Some candidates showed good understanding of the use of moving averages with reference to removal of seasonal variation and allowing us to see the trend; others were able to give a reason for using 4-points in their moving averages. The most common correct answer was referencing that the data was in 4 quarters.

Part (d) required candidates to compare the seasonal trend shown for motorcycles and the seasonal trend shown for cars. There were a good proportion of candidates who were able to correctly compare these, in some cases referencing both the greatest and lowest values. Some candidates only commented on one of the two vehicle types – motorcycles or cars – rather than giving the required comparison. A common incorrect answer was to attempt to describe the changes quarter by quarter along the time series.

The plotting of moving averages and drawing of a trend line as required in part (e) was attempted by the vast majority of candidates. There were a good number of candidates who correctly plotted the moving averages and added an acceptable trend line. Where plotting was attempted, most candidates were able to plot at least one correctly, but often made errors with the horizontal plotting.

Trend lines were often omitted or points incorrectly joined from quarter to quarter rather than an overall trend being shown as a single ruled line.

The final part of the question asked candidates to describe and interpret the trend in the number of cars first registered in the UK from 2017 to 2019. This was generally answered well by the majority of candidates. Only a small number of responses were seen where the description of the overall trend was given, but the interpretation omitted. A common error was to give a step-by-step description of individual changes over the time period rather than the overall trend.

#### **Question 4**

This question presented candidates with a plan for an investigation including a description of how Maciek planned to collect data and how he planned to process and present the data. The candidates were asked to discuss whether these plans were appropriate.

It was pleasing to note that the vast majority of candidates made a significant attempt to answer this extended response question. Many candidates were able to gain some marks, but candidates were limited with their marks due to giving statements which were too simple and not justifying their reasoning clearly or specifically enough. It appeared that some candidates did not realise that 5 marks required 5 statements and, in a small number of cases, did not recognise that they should comment on both the collection of data and the processing and presenting of data.

Common correct observations from candidates included identifying that it was not appropriate for students to record own data as they may lie, that the axes for the scatter diagrams were reversed, that it was not appropriate to record age as it was not being used and that it was not appropriate to extrapolate.

Candidates who commented on the 50 athletes often discussed this as a census rather than referencing the suitability of sample size. They generally did not identify that some athletes might be missing and that to make it better, athletes outside the club would be beneficial.

Some candidates referred to multiple jumps being taken, and the mean being used as a potential improvement for data collection.

Candidates often did not realise the benefit recording gender, dismissing it like age. Therefore, they struggled to reference the benefits of drawing the two separate scatter graphs. Many students referred to scatter graphs being useful, with a line of best fit, but failed to refer to its use to notice correlation.

#### **Question 5**

Part (a) of this question which required candidates to explain why SRCC may not be appropriate when looking for an association between the ranks given to competitors in two different competitions, was answered correctly by a minority of candidates. Correct responses frequently made reference to the fact that

there may not have been the same entrants to both competitions or the same number of entrants in both. Less frequent was a reference to the data not being bivariate. The most common response was an incorrect one, "these are different competitions which have nothing in common". Other incorrect answers included, "correlation does not imply causation" or reference to "Mr Smith being biased since he judges both competitions". A minority of candidates did not attempt this question.

Part (b) of this question required candidates to calculate Spearman's Rank Correlation Coefficient and provide a statistical and contextual interpretation of the coefficient. The formulae sheet provided the formula to calculate the coefficient, so candidates were required to substitute the correct values. This part of the question was answered relatively well. Errors in the calculation included forgetting to subtract from 1 or by using 342 in an otherwise correct formula after successfully finding 34. Other errors from those attempting to use the correct formula included having negative values for  $d^2$  in the table, for example incorrectly squaring -3 as -9. A small minority of candidates squared the individual ranks rather than finding differences and squaring these.

The candidates were generally able to give at least partial interpretation of their value of the correlation coefficient either by giving the type of correlation (positive if their value was correct) or by interpreting their correlation coefficient in context. Many were able to give both the type of correlation and a contextual interpretation of this. In some cases, candidates had incorrectly obtained a value for Spearman Rank Correlation Coefficient which was below -1 or above 1 which should indicate that an error has been made, but often this was ignored and an attempt to interpret the value made.

## **Question 6**

In part (a) of this question candidates were asked to calculate a chain base index number in order to complete the table. This was answered well by the majority of candidates. The most common error was where the incorrect base was used – most commonly 1001.4 which was the price for January and would have represented an index number, omitting the requirement for chain base.

In (b)(i) candidates were asked to calculate the geometric mean of the 5 chain base index numbers. Candidates either knew how to calculate the geometric mean, or incorrectly reverted to the arithmetic mean. A good proportion of candidates were able to correctly calculate the geometric mean.

The interpretation of the geometric mean in context (part b(ii)) was attempted by the vast majority of candidates. Most candidates were able to give a partially correct answer, referring to the 'increase' but few gave the fully correct description including the reference to "1.4% per month". Common errors were to reference the increase as being over the 6-month period, or per year, or made reference just to the data "monthly average price".

## Question 7

This question required candidates to use the Petersen capture-recapture method to calculate an estimate of the total amount of burbot fish in Tolsona Lake (part a) and then comment on a student's plan for capture-recapture (part b).

In part (a), many candidates accurately calculated the correct number in the population using the correct fractions. A common error was failing to round to the nearest integer when giving the final answer. There was confusion with how to get from the fractions to a value for  $n$ , as many multiplied by 683, making 172 a common incorrect response.

In part (b), the students were given a plan for performing capture recapture to estimate the number of frogs and asked to review the appropriateness. Most candidates made some progress with responding to this, with a good number of detailed and accurate answers being seen. The most common correct observation candidates referred to was the stick on tags falling off in some way. Many candidates identified the small sample size. It was also common to see comments on the potential change in population size. Reference to time period was also common, although some simply said it was too long or too short, without any justification, or giving incorrect justification such as 'not long enough for frogs to mix back in'. Few candidates commented on the need for the second sample to be collected at the same place or the fact that the second sample was planned to be collected at the same time.

## Question 8

Part (a) of this question required candidates to complete a probability tree. The vast majority of candidates answered this question correctly with only a small minority proceeding the use the probabilities for advanced drivers ("0.16" and "0.84") for not advanced drivers instead.

Part (b) of this question required candidates to calculate conditional probability. This was very poorly answered with only a small minority of candidates able to deal with this conditional probability question. The majority proceeded to calculate the probability of advanced drivers that do not exceed the speed limit ( $0.18 \times 0.84$ ), although an answer of 0.18 was also frequently seen. Another common incorrect response was  $3/14$ , from  $0.18/0.84$ . A number of candidates failed to attempt this part.

Part (c) and (d) of the question related to the calculation and interpretation of relative risk. This proved challenging for candidates.

Part (c)(i) of this question required candidates to calculate the relative risk for not advanced drivers compared to advanced drivers. This question was answered correctly by some students while others calculated the risk for advanced drivers compared to not advanced drivers or multiplied the probabilities instead. A common incorrect response was 7.4, coming from  $(0.82 \times 0.26)/(0.18 \times 0.16)$ . An answer of 0.1, from  $0.26 - 0.16$ , also appeared frequently.

Part (c)(ii) of this question required candidates to interpret the relative risk. This was answered reasonably well by those candidates who produced the correct answer of 1.625 in (ci) and also by those who answered 7.4 in (ci). However, some misinterpretation was seen with both these responses. Some candidates believed a relative risk of 1.625 was equivalent to a 162.5% increased risk while others thought 7.4 was equivalent to an increase in risk of 7.4% or 740%. In some cases interpretations were not very accurate either because the value of the relative risk was omitted. Finally, candidates who had obtained values below 1, for example 0.6, incorrectly interpreted their value as a higher risk or exceeding the speed limit.

Part (d) of this question required candidates to comment on the appropriateness of conclusions with statistical reasons. This was answered poorly. Candidates generally did not consider the proportion of not newly qualified drivers exceeding the speed limit as a possible reason for the decrease in the relative risk.

### **Question 9**

In part (a) nearly all candidates used the formula correctly to achieve the correct value for the skew. In a small number of cases candidates incorrectly changed their value to be positive when giving their final answer.

Part (b) of the question required candidates to compare in context the distribution of male handspans and the distribution of female handspans from the results that had been presented. This was answered well by candidates with the comparison of average and comparison of spread the most common marks awarded. Many candidates could describe if the skew was positive or negative, but then could not interpret it. A common error was to omit the contextual interpretation of the comparisons. Those candidates who were most successful in giving a complete answer worked through the different elements to compare one by one including paired comparisons and contextual interpretations for each. Some failed to make comparisons here though and merely stated the values without saying which was bigger/smaller.

In part (c) candidates were asked to explain why a weighted mean was appropriate in this case. Some candidates correctly identified that there were differing numbers of males and females in the results presented, however many did not show understanding of why a weighted mean is used and wrote about accuracy, reliability, or distribution.

### **Question 10**

Part (a) asked candidates to interpret in context the gradient of a regression line. Candidates tended to state that the number was the gradient and did not realise they were being asked to define the rate or said incorrectly 19.61 was the weight of the snakes. A small number of those that attempted to interpret the gradient incorrectly stated that it was the increase in length per 1g increase of weight.

Part (b) was also not answered well. Many candidates misinterpreted the data as representing the growth of a snake and the intercept as a start value - discussing a negative weight. Candidates struggled to understand that the

gradient given in the equation related to the rate of increase, therefore were unable to suitably explain and compare the relationships between snout-vent length and body weight for the different snakes. Where candidates were able to give a partially correct answer this was often just the first mark for identifying how the increase in snout-vent length effected the weight.

Part (c) of the question related to how the researchers could check whether a normal distribution was appropriate using histograms (part i) and using averages and standard deviation (part ii).

In part (c)(i) there were a good number of candidates who were able to correctly identify that the histogram could be used to check whether the distribution fitted a bell-shaped curve. Others were able to refer to identifying whether the data was symmetrical or checking for skew. Some candidates referred to comparing to a normal distribution but gave no details of what they would actually look for in the histogram. Other incorrect responses referred to averages, standard deviation, or box plots.

Candidates found (c)(i) challenging. Where candidates were able to make correct comments the most popular answer was to describe the proportions of data that should fall within 1/2/3 standard deviations of the mean, including correct percentages. There were some candidates who gave a partial description of this referencing the percentages and standard deviation, but failing to refer to this being the proportions within that many standard deviations of the mean. Some candidates stated all three statements relating to standard deviation and percentage away from the mean, but did not make reference to the mean, median and mode being equal. Very few candidates were able to fully articulate the link between normal distribution and the mean, median and mode. Only a few of the candidates referred to using the formula and calculating a skew value.

## **Question 11**

Part (a) of this question required candidates to comment of the appropriateness of the random response technique. The majority of candidates were able to comment that it was a sensitive question, honesty and about making staff more likely to answer it. The most common incorrect answers stated inaccuracy or referenced randomness, the latter showing a lack of understanding in random response technique which had been confused with the use of random sampling.

Part (b) of this question required candidates to comment on the appropriateness of the given random response question stating statistical reasons. Common incorrect answers said that the question was confusing, and that people would lie or change their mind about the number selected. There were also responses that referenced not being able to distinguish between those answering the questions with 'A' based on the even number and those who had answered the question highlighting a lack of familiarity with the random response technique. Some candidates suggested using a spinner or flipping a coin however there were only a few that were able to reference that the numbers are not randomly generated. Identifying that the numbers were not randomly generated was more commonly seen as a comment than referencing the lack of timeframe or definition of personal errands.

Part (c) of this question related to stratified sampling. Candidates were asked to find the minimum total sample size required for there to be at least 15 people from each strata included. The correct answer of 394 was rarely seen as a result of candidates not realising that to get at least 15 as a minimum then 14.5 should be used. Answers of 407 and 408 were more frequently seen as a result of using 15 as opposed to 14.5, often with candidates arriving there through a trial and improvement approach, this approach generally gaining partial credit for 407 or 407.(4...). Many candidates attempted to find frequencies for each cell separately but failed to score any marks as they used 15 and made subsequent errors in calculations or premature rounding.

## **Summary**

Based on their performance on this paper, students should:

- 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.
- Practise analysing plans for data collection and the subsequent diagrams and calculations together with giving statistical reasons or against the approaches suggested.
- Develop their understanding of the meaning of skew.
- Practice calculating geometric means for a range of different contexts and giving interpretations of these.
- Develop their understanding of conditional probability and practice a range of questions.
- Develop their understanding of relative risk.
- Develop their understanding of regression equations and the interpretation of the gradient of these.
- Develop their understanding of the random response technique.

