



Pearson

Examiners' Report Principal Examiner Feedback

Summer 2018

Pearson Edexcel GCSE
In Statistics (2ST01)
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 2018

Publications Code 5ST1H_01_1806_ER

All the material in this publication is copyright

© Pearson Education Ltd 2018

GCSE Statistics 2ST01 Principal Examiner Feedback – Higher Paper 1

Introduction

In general, students appear to have had sufficient time to complete the examination with few questions left blank, including where topics had not been commonly tested before. In the main, students showed an awareness of the need to show clearly their calculations and to work to sensible degrees of accuracy.

Whilst candidates were generally making sensible attempts at comments or reasoning where expected, some were too vague in their answers or presented handwriting and calculations which made it difficult to understand their responses.

Some questions required candidates to consider the suitability of a method. They need to be careful in their choice of vocabulary as it was not uncommon to see suggested that results would be 'inaccurate' rather than 'unreliable'. These do not mean the same thing.

Centres should be aware that the whole of the specification content is expected to be examined, not just what can be seen on past papers.

Reports on Individual Questions

Question 1

Many candidates did not recognise the quota sampling method. Commonly it was wrongly identified as stratified or systematic sampling.

A minority gave a valid advantage referring to ensuring the inclusion of a mix of age or gender. Many were too vague, not mentioning these attributes (e.g. 'includes a mix of people'), or suggested other pragmatic considerations such as ease of use.

Most candidates however gave an appropriate disadvantage, with some suggesting it could be biased or raising concerns about fairly representing the population, but most commonly the mark was earned for practical considerations related to time and cost. Those not scoring here often suggested the sample would not be 'accurate' or would be too small.

Question 2

Parts (a) and (b) were starred for Quality of Written Communication so candidates should have considered their vocabulary carefully. Whilst some correctly identified the type of data as continuous in part (a) most candidates scored just 1 mark, often for identifying it as quantitative or numerical. It was not uncommon for 1 mark instead to be given for recognising the data the pupils recorded was primary. It was also not uncommon for candidates to incorrectly suggest the data was discrete 'because times were to the nearest minute'.

In part (b) it was most common for candidates to gain a mark for suggesting reasons why the data was unreliable (often due to pupils' poor recording of times, or making up the times so it looked like they spent longer on homework), however sometimes they incorrectly said this meant it was not 'accurate' rather than not 'reliable'. Few candidates gained both marks as they did not correctly identify the data that Supul collected as secondary.

Most candidates gave a correct ordered stem and leaf diagram in (c), but it was disappointing to see many failing to provide a usable key (despite the question reminding them to do so). Unordered plots were rare but some candidates made an omission in completing the plot; candidates should be encouraged to check that their stem and leaf diagram has the correct number of leaves.

Candidates were usually successful in parts (d) and (e), sometimes following through from an omission in the stem and leaf diagram. A few used an ordered list rather than the diagram. Those with the stem 'upside down' in (c) had to be more careful counting in for the median with some getting the position wrong.

In part (f) most found the maximum and minimum times but some did not realise they needed to use bounds, simply calculating $64 - 21$ to score no marks. Whilst the correct answer was often found using correct bounds for 2 marks, it was not uncommon to see an incorrect upper bound used (e.g. 64.4 or 64.49) with a correct lower bound of 20.5 so scoring 1 mark only.

Question 3

The great majority of candidates knew that the product of probabilities was required to reach the given answer of 0.16 and so scored the mark. A small number added then thought that 0.8×2 was 0.16.

Again in (b) most knew to multiply, usually giving both required products or $0.24 + 0.24$, but on reaching the result of 0.48 many failed to make an explicit comparison with 0.5 to answer the question. Recognition that 'evens' likelihood meant a probability of 0.5 was required. A few disappointing solutions had just one product reaching 0.24 and said this was an even number.

Part (c) was generally answered well with many candidates finding the expected outcome of 36, and fewer taking the alternative approach of comparing with probabilities. A small number failed to draw a conclusion.

Question 4

Most candidates gained the mark in part (a) although some responses were, at best, poorly expressed as e.g. 'England had more losses' which was condoned in this instance. There is insufficient information given at this point to suggest there were more losses – the question was about the *proportion* of games which were losses. The calculation in part (b) was successful for most although some calculated a percentage, either leaving this as their answer or introducing rounding error for the second stage of the calculation.

Comparison of the size of the draws sector, occasionally referred to as proportion, usually gained candidates a mark here and 2 marks was quite common. Some however tripped on the interpretation that a smaller section representing the same number meant that the total must be larger for England. A few candidates did not score as although they concluded England played more they had no correct reason.

Question 5

Candidates were generally able to extract the required information to answer parts (a) to (c) successfully, although the main cause for not scoring here was using the wrong year. There were a small number with arithmetic errors in (b) and some who could not manage the percentage calculation in (c) having found the correct figures.

Explaining how the data in the table supported the statements in part (d) was more of a problem however. Some were too vague, not referring explicitly to key aspects such as total or all types of travel for (1), or to 'all visits' or all reasons for travel for (2). Some responses did little more than re-express the statement they were trying to justify. When using figures candidates need to realise explicit comparisons are required: e.g. '38 519 is more than half the total' is correct reasoning, whilst '38 519 out of 60 082 travel for holiday' is insufficient as the reasoning remains implicit.

Whilst reference to the wrong year was not allowed, a number referred unnecessarily to all years (which was condoned). Candidates need to be reminded to read a question carefully, especially when it regards extracting information.

Question 6

The majority of candidates showed familiarity with the method to estimate mean and were systematic in arriving at a correct answer to (a). Those who failed to do so were due arithmetic slips, using end points instead of midpoints, or a few scoring no marks for an addition with division by 5. Most were aware in (b) that grouping or using midpoints was the reason for the difference between their estimate and the true value, but a few incorrectly thought it was due to 'rounding' or simply stated that (a) was only an estimate. A few candidates incorrectly referred to midpoints of classes as medians.

Whilst many correctly identified the skew as negative in part (c) only stronger candidates scored 2 marks for also correctly comparing the mean and median. Common incorrect reasoning made no reference to the mean or described its position relative to the quartiles.

Calculations of interquartile range for part (d) were usually correct for the first 2 marks, including for most the correct reading of the scale for the lower quartile values. Most then compared appropriately the interquartile ranges but many fewer were able to offer an appropriate interpretation of this comparison. Some attempted no interpretation or referred only to the 'data' rather than the race times, and hence did not score the final mark.

The final part of the question was starred for QWC so careful reasoning was needed. Many answers here were too vague however suggesting only where one event was more likely than the other (e.g. based on proximity to a median). Few candidates focussed on the key values/ranges on the graph where definite decisions could be made. A minority identified the two key values and usually scored 2 marks but only stronger candidates gave a complete answer realising there were three ranges of times to comment on. It was not too uncommon for candidates who did not recognise the key values to gain 1 mark with the special case for identifying a correct decision for any one time or range of times.

Question 7

It was apparent that candidates were less than familiar with cumulative frequency step polygons, whilst some incorrectly were using a total of 50 teaching sets rather than the correct 48.

The median was usually managed by candidates (and not affected by those using the incorrect 50) but the IQR was more of a problem. A common error here was using an upper quartile of 17, following from the use of 50, although the lower quartile was usually correct and 1 mark was often gained if working was shown. The correct mode was often found, using the largest step, although it was unclear where the various incorrect answers came from.

A common wrong answer in (b) for the largest set size was 22 from the right hand end of the grid. In part (c) inexperience with these diagrams led to many reading the 'wrong end of the riser' at 29 rather than 37. Only the stronger candidates gained both marks, whilst a few gained 1 mark for correctly identifying 37 but either leaving that as their answer or subtracting it from 50.

Question 8

Few candidates correctly identified a time series graph as the most appropriate diagram.

The biggest difference was often found correctly in part (b) although some misread the question to find the largest year on year difference.

Most candidates realised in part (c) that using the data for a prediction would not be suitable, many recognising the data was not recent or that extrapolation would be needed. Common incorrect reasoning referred to fluctuations in the data making it unreliable.

In part (d) a number of candidates failed to score by not reading the question properly; some described how to use the data to extrapolate, and stated that this would be unreliable (having often already answered along these lines in part (c)). Many did however suggest gathering more data but gaining the second mark for identifying a possible problem was less common; some failed to score this final mark by stating the data might be inaccurate (possibly meaning not reliable).

Question 9

It was not too common for both marks to be scored in part (a). Usually a mark was gained for recognising positive correlation, often qualified as weak, although there were a number who said there was no correlation. An appropriate interpretation of the correlation was much less common however. The correct calculation of SRCC usually led to full marks in (b) as candidates recognised the value as positive. (Those who had said weak positive correlation in (a) often concluded this was not supported as 0.56 was quite strong positive correlation, and this was accepted.) Errors in the calculation of SRCC however included using the wrong value for n , or no correct use of the '1 - ...' Some did not attempt the calculation so could not score.

Question 10

Generally candidates were able to suggest a suitable hypothesis in part (a) with only a few not scoring due to writing a question. The majority were then able to identify the two necessary variables, although these were not always clearly expressed. e.g. 'time' on its own for one of the variables was not uncommon but too vague. Incorrect answers included variables such as book type or genre which were not necessary for the investigation.

More poorly attempted was part (c) where the population for the investigation into books was often not well defined, or was often incorrectly specified in terms of the borrowers. Sometimes numeric values were given to each part showing a lack of understanding of the statistical terms. Those referring to the books in (c)(i) did not always specify all books and in (c)(ii) it was clear that 'sampling frame' was not commonly well understood, with instead a sample being described. Only stronger candidates scored 2 marks here.

In (d) part (i) clear explanations were not common; a variety of answers were often vague or recognised the different book types but suggested an equal number from each was needed. Not uncommon was that there were a large number of books (presumably recognising the need for sampling rather than stratified sampling). Calculations for (d) part (ii) were usually correct however with nearly all recognising the need to round to an integer answer. Premature rounding when the calculation was done in stages was an issue with a small number.

Candidates were told in part (e) to include no further calculations. However the majority proceeded to *describe* further calculations, without using figures, for the other strata. This alone could at best score the 3rd mark for recognising a need to sample separately from each stratum. Often candidates made reference to random selection or random numbers for the 1st mark, but less common was gaining the mark for clearly corresponding the random number selected to a book number. Whilst some candidates managed to gain all three marks, clear coherent answers that covered the three aspects on the mark scheme were quite rare. It was clear that some candidates were giving descriptions of simple random or selective sampling, which could score at most 2 marks.

Question 11

This proved an unfamiliar topic for many candidates which is not well understood. For example in part (a) an incorrect suggestion of using 10 rather than 0 was not uncommon and showed a lack of understanding of the purpose of the numbers in the simulation. Many candidates concluded the choice was not sensible. A number of correct answers were seen however, showing in various ways the link between the number of numbers assigned and the percentages.

It was disappointing in (b) to see a number of candidates not read the question and total the tallies without including the final 10 results. Those who did include the given results were commonly correct, scoring at least 1 of the marks.

Least well understood was the reason for repeating the simulation, required in part (c). Some incorrectly focussed on achieving results closer to the original percentages or suggested there may be more than 80 customers. Whilst responses were often vague, most commonly a mark was gained for reference to spotting anomalies or finding an average for each.

It was very rare to see a complete answer that recognised there would be a distribution of results from the repeated simulations, which in turn would help the manager decide how many of each dish to be prepared for.

Question 12

Most were able to show a correct calculation for the given frequency density in part (a). Many were then able to go on to find correct heights of bars for the histogram in (b), commonly scoring at least 2 marks: 1 for at least two correct heights and the second either for correct boundaries or all heights correct. The most common reasons for not scoring full marks were candidates failing to label the axes appropriately (common errors were 'number of patients' on the vertical axis or 'class width' on the horizontal), or for having incorrect scaling on the horizontal axis. This incorrect scaling often led to incorrect bar widths also.

Scoring 1 mark in part (c) was unusual as those who understood they needed to add a quarter of the penultimate class frequency to that of the final class typically reached the correct answer. Many realised they needed to add to the final class frequency but did not know what. (A few used a fifth rather than a quarter of the penultimate class frequency.) Strictly the answer should have been scaled up by 10 000 referring to the table; many did not do this but were not penalised.

Question 13

There was a general lack of familiarity among candidates with weighted index numbers as indicated by part (c). However many were able to establish the connection between weightings and running costs for at least 1 mark in part (a). This was done in a variety of ways, usually involving a calculation: recognising the weightings as percentages was quite common, as was finding a common scaling of 800. Using index numbers to find a new cost is a common demand: many were familiar with this and able to reach the correct answer in part (b).

Finding the weighted average of the index numbers for part (c) was not often seen however. There were some left blank. Of those who found the correct weighted index number most took the 'long route' of first evaluating the total of new costs using each index number. Usually they remembered to not include a % sign on the correct 109, and made an appropriate comparison to answer the question for full marks. More frequent were responses that attempted to assess the manager's assertion without using index numbers, as was required in the question: often candidates found the percentage rise or compared the cash rise with a 10% rise, to score at most 1 mark for the special case.

Question 14

Familiarity with the probability distributions was not strong as this was not particularly well answered, with 4 marks being very rare. Correctly naming the distributions however often gained candidates 1 or 2 marks. Picking the normal distribution for (i) was the most common to score but very few identified it as a continuous variable. Common wrong answers included that it was likely to be the same each day as it was the same journey. Binomial was a common incorrect answer to (ii) but for those correctly opting for the discrete uniform distribution very few said that each digit was equally likely.

Question 15

Although some candidates struggled with this question, scoring 0 (with a small number leaving it blank), it was quite well attempted for the final question on the paper. Despite 'without replacement' being clearly stated some either ignored this or did not understand its significance, scoring at most 1 mark in each part. However, a correct product was often seen in part (a) but not always correctly evaluated. (Some lost accuracy by converting crudely to decimals.)

Those scoring full marks in (a) were often also successful in part (b). Common was considering six pairings rather than three (e.g. using red, not red), sometimes with a tree diagram, and then evaluating correct products. Some failed to consider e.g. RW as well as WR, so scored at most the 2nd mark. A minority attempted the alternative working, often correctly, but some of these forgot to subtract from 1. Arithmetic errors (sometimes due to using rounded decimals again) led to the final mark being lost for a number of candidates.

Grade Boundaries

Grade boundaries for this, and all other papers, can be found on the website on this link:

<http://www.edexcel.com/iwantto/Pages/grade-boundaries.aspx>

