edexcel

Principal Examiner Feedback

Summer 2016

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 2016
Publications Code 5ST1H_01_1606_ER

GCSE Statistics 2ST01
Principal Examiner Feedback – Higher Paper 1

Introduction

It was pleasing to note the success of students on Q11 which encompassed much of the statistical enquiry cycle, from hypothesis to conclusion. Students are also generally making sensible attempts at comments when asked for interpretations, comparisons or reasons. Poor clarity of expression continues to be a problem for some students however; students should take more care in this respect, re-reading their responses to ensure that examiners are able to award the marks deserved.

It seemed that overall students found this paper quite challenging; there were some topics in particular which had not been tested regularly and for which students had not prepared themselves (eg mutually exclusive and exhaustive events, random response, cluster sampling and formal use of the general addition law.) The whole of the specification content will be tested, not just what can be seen on recent past papers. Insufficient time did not generally appear to be an issue as most questions were usually attempted, although index numbers, standardised scores and formal probability work appeared to be unfamiliar topics for some students.

Questions often demand a reason to support an answer, in which case a mark will usually not be earned without a supporting statistical reason. Calculations must be demonstrated with clear working when the answer is given in the question (such as Q4(a), Q13(b), Q14(a) and Q15(a)(ii)) or when an answer could be guessed without use of any real statistical understanding (such as Q7(g) and Q12(d)).

Students' use of correct statistical language was mixed. For example it has always been common practice that this is expected when comparing distributions (eg 'median' rather than 'average'). It should also be noted that stating individual values is not a comparison in itself; when values are stated there needs to be use of comparative language (eg " ... which is higher than ...").

Report on individual questions

Question 1

Most students were able to propose an appropriate question in part (a) although there was a minority of students who neglected to give closed options. Students should note that a question is not closed unless such options are explicitly stated; eg "Do you prefer cats or dogs" without option boxes remains an open question as it can still be answered in many ways other than just with 'cats' or 'dogs'.

In part (b) a minority of students offered a two part questionnaire rather than a data collection sheet and so did not score. Most students realised, however, that a table of some sort was required and gained at least one mark, usually for enabling gender to be recorded. Although a majority of students scored both marks, where a second mark was not gained it was usually due to a lack of closed options for pet type that made their table not fully fit for purpose.

In part (c) many students realised data type was the issue but often this was not correctly expressed, for example by suggesting that the data was not discrete or was not continuous – these were not sufficient responses. Some students did not realise that data type was an issue; instead they suggested that there were too many variables or pet types to show on a scatter diagram.

Question 2

Interpreting the graph in parts (a) and (b) provided easy marks for most students, although a small number of students had an incorrect limit to their answer in part (a) or read from the wrong line in part (b).

Many students correctly identified the trend from the graph in part (c) as upward/rising, with the most common answers being ‘positive’ or ‘increasing’. It must be noted that ‘positive correlation’ is an incorrect description. Most of the incorrect responses seen, however, described the falls and rises rather than the overall trend. Students are advised that the trend is the overall picture shown by the data ignoring fluctuations along the way and that a trend should be described as upwards (or rising) or downwards (or falling).

Recognising that not all age groups were included was common in part (d) although some students incorrectly thought that older age groups (eg over 55) were missing. Fewer students focussed on the fact that the data was for the UK only. The question asked why Jonathan *is* correct, not why he *might be* correct: the most common incorrect answers suggested, for example, that the data only came from a sample or that some people may not have given their age.

Question 3

Most students were able to gain at least one mark in part (a), although some students misread the question with, for example, 0.2 being a common incorrect answer to part (a)(ii), having missed the word ‘**not**’. It was evident in part (b), however, that few students were familiar with the term ‘mutually exclusive’ or were able to express their understanding clearly. Descriptions showed confusion with a variety of other properties such as exhaustive, independence or equally likely.

In part (c) about half of the students were able to gain the method mark for multiplying two probabilities. A number of these students assumed ‘without replacement’ however, with 3/10 × 2/9 being fairly common. Some students wrote down the correct multiplication but then failed to evaluate it correctly. Other students miscounted the number of odd numbers with 4/10 × 4/10 being a not uncommon method. The most common misconception was adding the probabilities rather than multiplying them. Simply stating $\frac{3}{10}$ was another common incorrect answer.

Question 4

Many students were able to offer an appropriate equivalent calculation for the given size of stratum in part (a). Some students appeared to know how to obtain the given answer, but did not show the full calculation - students are advised that when a question asks them to establish a result full working is required. Poor mathematical statements were not uncommon among those doing the calculation in stages, although this was not penalised this time if an understanding of the stages required was demonstrated.

For their description of random sampling in part (b) most students gained at least one of the marks, usually for numbering or listing the students. Descriptions of matching the selected numbers with the corresponding students were less well expressed but quite common, although many students described the use of numbers they would generate rather than those provided. Less common was the stating of the need to deal with the repeated number. Few students scored full marks for including all three points.

Many students were able to find the correct estimate in part (c), although some students showed poor accuracy when working in stages, sometimes leading to an incorrect answer (often 41). Other students lost a mark by giving the proportion ($\frac{42}{90}$) rather than the number of students. Common errors in working for this part were in multiplying the correct fraction ($\frac{7}{15}$) by 100 or by the total number of students (240).

Question 5

Although many students offered appropriate answers to part (a)(i) it was clear that a fair number of students did not understand what was being asked of them (perhaps not understanding the term 'variable'). These students often described details of how the survey could be conducted. 'Size' was a very common answer but was too vague to gain credit as a variable. The most common answers to gain credit were 'height' and 'weight', usually followed by a correct description in part (a)(ii). Some students managed to gain this mark following a non-creditworthy answer in part (a)(i) if the answer mentioned a variable.

There was a common misconception by students in part (b) that the data was primary, perhaps because they were told the company was carrying out a survey, but ignoring that they were also told that the company would ask hospitals for the data. A good proportion of students, however, realised that it was indeed secondary because the data was already held by the hospitals. Very few students failed to give a reason for their choice. Whatever was their answer to part (b)(i), students were quite adept at gaining at least one mark (often both) for a correct advantage or disadvantage of their chosen data type. Most commonly for secondary data, the advantages stated were 'speed' or 'low cost of collection' whilst disadvantages usually related to reliability of the data. Suggestions that the data was unreliable or inaccurate were condoned, whereas the more correct answer that it was of unknown reliability was seen less often. Those students opting for primary data frequently had 'reliability' as their advantage and being 'time consuming' as their disadvantage. Some students contradicted themselves by suggesting that primary data was more reliable but that it might contain errors. Such contradictory statements do not score any marks.

Question 6

The 'random response' technique for obtaining answers to sensitive questions had not been previously tested and few gained full credit. The context was perhaps the reason why about one in eight students surprisingly failed to estimate how many people from 1000 would score Heads on a fair coin. Many of these students spoilt their answer by stating 500/1000. Just over half of the students correctly predicted 500 but got no further. The remaining students recognised that 60 was a significant figure (many students incorrectly stating it as their final answer) but only a small number of students realised that this was 60 from 500 and hence there would be 120 expected from 1000. Common incorrect answers seen were simply 560/1000 or halving to give 280/1000.

Question 7

The first three parts of this question were mostly answered well with students correctly interrogating the table of data. In part (a) however a small number of students misread the question and described changes for individual operating systems rather than for total sales. Some students found part (c) harder with slightly fewer correct answers. A few students gave numerical answers to parts (b) and (c), again having not correctly read the question.

Most students were able to make reference to rounding in part (d), although it was not always obvious that they understood that it was individual values having been rounded that meant the values did not add to 100%. Common incorrect answers made vague reference to their being errors or suggested it was incomplete data.
The overall response to part (e) was disappointing with too few students considering statistical reasons. Very common responses were the somewhat trite answers, such as pie charts would be quicker or simpler or would show the results better. Those responses gaining credit usually focussed on pie charts being better at showing results in proportion to the whole ('showing

percentages better'), with very few responses focussing on the pie charts being 'comparative', ie representing the different total sales.

In part (f), perhaps because they had a choice of data to use (sales or market share), students did not always find the calculation of the pie chart angle as straightforward as expected. Whilst there were many correct answers, some students incorrectly gained an apparently correct answer by summing the sales and market share and dividing by the sum of total sales and total market share but correct' answers from incorrect working do not gain credit. Other errors seen here were inverted fractions or correct fractions ×100 rather than ×360.

Very few students were able to carry out the calculation required for the radius of the comparative pie chart in part (g). Most commonly, students simply scaled the radius rather than the area by the fraction of total sales, giving a common incorrect answer of 7.3 cm. Many students thought the radius should still be 5 cm whilst some students stated 6 cm with no working – this was insufficient to demonstrate understanding and so could not gain full credit.

Question 8

Cluster sampling has not been previously tested and only a quarter of students recognised it by name in part (b). Most students were able to gain credit in part (a), however, by suggesting a sensible advantage or (less often) a disadvantage of the described method. It seemed many students were giving stock advantages and disadvantages of sampling in general compared to a census, rather than for cluster sampling in particular as a method. Being quicker or cheaper were common correct advantages (as fewer offices would need surveying) but many students incorrectly thought it was random and so fair to all employees, having focussed on the random selection of the offices. This showed an underlying lack of understanding of the method. Not being representative was common as an acceptable disadvantage, although it was usually not clear that the student understood why. Common incorrect disadvantages, which illustrated the lack of understanding of cluster sampling, were that this method would be time consuming and expensive. This contradicted the advantage that it is quicker and cheaper to survey only a smaller number of offices.

Students are far more familiar with the advantages and disadvantages of questionnaires and face to face interviews and so performed much better in part (c), often gaining both marks. Unfortunately, rather than considering more than one aspect, some students gave both sides of just one aspect and so only gained one mark. The most common aspects considered (in order) were the 'honesty' or reliability of answers, the time taken, the issue of questions not being understood and the response rate. 'Honesty' (not the best choice of words but most commonly used by students) continues to be misunderstood by many who think people are 'less likely to lie' face to face, rather than realising the issue is to do with candour – ie feeling able to offer their true answer. A few students failed to score by pointing out the aspects (eg expense and honesty of response) that should be considered without stating how these related to the two options proposed.

Question 9

Index numbers continues to be a problem topic for many students with about half failing to score a mark in each of parts (a) and (b). There were a number of blank responses for this question. Although some students failed to multiply their correct fraction by 100 in part (a) the most common error was simply finding the difference in insurance costs. Some students spoilt a correct calculation by incorrectly applying £ or % to their answer.

In part (b), those students who forgot to multiply by 100 could still gain credit but students often used 2010 as the base year – not understanding that 'chain base' means the previous year is always used as the base year. Some students recovered with their interpretation of an index number (eg by recognising a fall) but as this was a QWC question the interpretation needed to be complete to score full marks. (Not many students included the size of the percentage change and particularly the time period.) Some students interpreted an index under 100 as a negative value. Students not attempting index numbers often used subtraction to find a change in cost, which did not score.

Question 10

Whilst most students were able to read off the median from the box plot in part (a) fewer students were able to find the IQR. The most common incorrect answer was to give the range (43).

In answering part (b) it was disappointing that the majority of students seemed to take no notice of the data presented in the box plot when making their choice. If they had, the outliers and skew should have made the rejection of option 2 straightforward as mean and standard deviation would be distorted by these features. Instead many students seemed to opt for text book answers that mean and standard deviation would better represent the data as it included all values. Of those students choosing the correct option 1, some incorrectly just stated that it would be simpler or would not involve decimals. This was QWC so some missed a mark by referring to ‘extreme values’ rather than using the correct vocabulary of ‘outliers’. Very few students mentioned skewness.

Comparing distributions is essential in statistics and students are now quite familiar with this. Hence many students scored well in part (c), provided they remembered to use the correct vocabulary. (Fewer students now seem to lose credit by, for example, referring to ‘average’ rather than ‘median’.) Only a small number of students used a mixed pair of measures (eg mean and IQR), scoring a maximum of one mark or listed values without making an explicit comparison.

Question 11

This was a good question for many students with more than half scoring at least 7 marks. Whilst most students were able to state an appropriate hypothesis, a small number failed to mention medals or included an incorrect causation for the given scenario (eg that the wealth of a country depends on the number of medals won). Few students made the mistake of posing a question.

Plotting the three missing points was done in part (b) without problem by most students but some plots were out by a whole square or more. (It should be noted that the usual expected tolerance is half a square.)

Ranking the number of medals was usually correct but there were arithmetic or sign errors seen in evaluating the squared differences. Students should show the total of their d^2 column to help ensure follow through method marks are appropriately awarded. Most students showed the substitution into the Spearman’s formula, although some errors were seen with incorrect values of n or the failure to subtract from 1. Despite arithmetic slips many students still picked up at least three of the four marks. Few students used reversed ranks.

The conclusion in part (d) was QWC so students were expected to state the correlation observed as well as to give a correct conclusion for their hypothesis in part (a), thus completing the statistical cycle. A number of students stated that there was correlation without giving the direction while other students failed to refer back to their hypothesis in part (a). Some students tried to base conclusions on the evidence of one or two individual countries, which is not acceptable, showing a lack of understanding of the significance of their Spearman’s result. The few students who used reversed ranks found it difficult to reach a correct conclusion from their negative rank correlation result, incorrectly rejecting their correct hypothesis.

Question 12

In this question students were most successful in describing the seasonal variation shown by the time series graph and so tended to score best in part (b), gaining at least one mark. Some students did however describe the trend instead or describe the rises and falls rather that the specific seasonal variations. Fewer students were successful in part (a), suggesting ‘greater accuracy’ or referring to four years as the reason, although many students correctly focussed on the ‘4-point’ realising that this was to do with the four seasons or quarters in a year.

Finding and interpreting the gradient of the trend line in part (c) was probably the least successfully attempted part of the question. Most students took little or no account of the graph scales even when a sensible size triangle was drawn and so did not produce a useful value. ‘Counting squares’ typically led to 1/7, whilst working with very small triangles often gained method marks but gave an inaccurate answer outside of the accepted range. A small number of students attempted an equation for the line without identifying its gradient. When the gradient was found correctly very few students scored full marks as part (c)(ii) required a full interpretation, being QWC. Many students simply stated it was positive.

Students were a little more successful in part (d), even if just for finding the trend line value. As it is easy to ‘guess’ a value in the required answer range, (some students found an answer in the range by simply following the pattern of Quarter 1 values), a correct method was required to be seen in order to gain credit. The better students made an attempt at finding the mean seasonal variation for Quarter 1; unfortunately it was then not always known what to do with the value – some students stated it as their final answer, whilst others added it to their trend line value having found it as a positive value when strictly it should have been negative. It was not then noticed that this gave a value above the trend line instead of below. Students performed less well on this topic than in previous years.

Question 13

Being asked for the distribution shape in part (a) rather than for its skewness required students to think more; many students failed to realise that the answer was to do with skewness and, instead, described it as decreasing or triangular. Some students realised that they were being asked for skewness but described it as negative. Only one in six students gave the correct description of ‘positive skew’.

Parts (b) and (c) often illustrated a reasonable understanding of histograms, with many students able to demonstrate working for the given answer in part (b) (often simply as 4/5) and to find at least one correct frequency for part (c). Errors seen in part (c) were often arithmetic slips or including extra classes in error (usually the first).

There was limited evidence of correct attempts at the interpolation needed for part (d), although many students were able to identify the correct class or more commonly suggest a value within it, gaining one mark. A number of students did not attempt this part or gave an answer with no working.

Many students redeemed themselves in part (e) with appropriate reasoning as to why the London results should not be used for predictions in New York. They usually cited differing conditions (sometimes different length which was condoned) or different competitors. There were some students who gave reasoned arguments both for and against without reaching a conclusion, whilst a number of students gained one mark for suggesting it would be sensible, usually due to it being the same distance.

Question 14

It was apparent that many students were unaware of standardised scores with nearly half gaining no marks (with a number of blank responses), whereas over 30% of students gained full marks.

In part (a) some students carried out their calculations in stages (sometimes with poorly expressed mathematics) but usually leading to the stated answer. When successful in part (a), most students then realised that for part (b) a standardised score was needed (even though this demand had not been stated) and this was usually found correctly. A small number of students then interpreted incorrectly stating that 0.9 being closer to zero (or closer to 1) meant that Physics was best. Common incorrect attempts at part (b) simply compared raw test scores which led to the wrong answer of Physics.

Most of the students who had been successful in parts (a) and (b) were usually then able to reverse the calculation to obtain a correct answer in part (c), although a small number of students made a sign error in reaching the incorrect answer of 51.

Question 15

Students found the calculations in part (a) more challenging than usual although many were able to find the mean correctly. This was often the only part attempted successfully, with later parts often left blank. The most common error in part (a)(i) was to count the repeated value once only but still divide by 12 leading to the incorrect answer of 2.57 As the value of $\sum x^2$ was not given in part (a)(ii) many students failed to realise they needed to find it. Some students mistakenly used $\left(\sum x\right)^2$ for $\sum x^2$ in the formula for standard deviation whilst other students chose to take the longer route of finding $\sum (x-\bar{x})^2$, often with limited success. Those students substituting correctly into the formula did not always evaluate to more than one significant figure. It should be noted that when asked to demonstrate a stated answer is correct to a certain degree of accuracy it is necessary to first evaluate the answer to at least one more figure to demonstrate the correct rounding.

In part (b) the reference to 2 standard deviations prompted some students to state 95% without working, presuming a link to a normal distribution. Some students *misread* the question and evaluated limits using just 1 standard deviation. Those students who correctly found the limits were often unaware how they could then use them to identify that just one value from the list was beyond these values. A comparison with 95% was required in part (c) as there was a clear demand to use their percentage from (b), but this was not commonly seen. Many answers attempted to refer to other properties of a normal distribution (eg by making reference to symmetry or a bell-shape) and did not score.

Question 16

Very few students were able to interpret the Venn diagram correctly in this question. In part (a) students appeared unfamiliar with the term ‘exhaustive’ with recognition that the probabilities added to 1 being rare. A common error was confusion with ‘mutually exclusive’ by making reference to the intersection.

The misreading of the Venn continued in part (b) where the majority of students used 0.3 and 0.5 as P(X) and P(Y), omitting to include the intersection in each case. Those with an incorrect product could still have addressed part (b)(ii) but it was very rare for any student to consider the connection between $P(X) \times P(Y)$ and $P(X \cap Y)$ in considering independence. It is clear that most students were not aware of the conditions for the independence of two events. Many responses again showed confusion with ‘mutually exclusive’, making reference to the intersection. Some students gave the explanation for exhaustive events to this question, ie probabilities add up to 1.

Students’ responses to part (c) demonstrated little awareness of the general addition law and a lack of familiarity with formal probability notation. Those students attempting to use the law sometimes had a product of P(A) and P(B) instead of a sum, whilst those attempting to use a Venn diagram often incorrectly labelled it with 0.6 and 0.5 thereby taking no account of the 0.25 in the intersection. Following a correct Venn diagram some students failed then to add the intersection, giving 0.6 as their answer. A further common incorrect response was to subtract $P(A \cap B)$ (given) from 1 to get 0.75 as their answer.

Formal probability work seems to be unfamiliar to many students, with a large number of students scoring no marks on this question.

Summary

Based on their performance on this paper, students are offered the following advice:

- read each question fully and carefully before attempting to answer it and check that answers make sense in the context of the question
- show clear working to support the final answer and when necessary give a clear decision as well as the reasons
- write down probabilities as fractions, decimals or percentages and understand the terms and notation of formal probability
- understand and use the correct terminology especially when making comparisons
- ensure that when asked to make a comparison that two or more things are actually compared and a conclusion given
- be both precise and explicit in comparisons of distributions
- for a ‘Show that …’ style question, show all intermediate stages in the calculations not just the substitution stage
- revise all the specification content and understand what the formulae on the Formulae sheet mean

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

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