edexcel

Principal Examiner Feedback

Summer 2014

Pearson Edexcel GCSE
In Statistics
5ST1H_01 (Higher)

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 2014
Publications Code UG040113

GCSE Statistics 5ST1H
Principal Examiner Feedback – Higher Paper 1

Introduction

It was pleasing to note that candidates are perhaps improving in using the correct statistical language, which is expected when comparing distributions. Some however are unaware that these comparisons should be using a correct average, measure of spread and direction of skew. It should also be noted that stating individual values (e.g. IQR of ages of teachers in the USA and UK) is not a comparison in itself; when values are stated there needs to be use of comparative language (e.g. “which is higher than ...”).

Commonly questions demand a reason to support an answer. Especially when there are only two possibilities a mark will usually not be earned without a supporting reason. Drawing conclusions from statistical diagrams needs to make clear reference to a feature of the diagram which provides the evidence. For example in question 9(b) candidates often simply stated that the chart showed the percentage for USA was higher than for UK, which was just restating the hypothesis.

There are some topics in particular which were not well attempted including index numbers, working with seasonal variation, systematic sampling, and the more advanced probability work. Most questions were usually attempted, although sometimes with parts left blank, suggesting that insufficient time was not an issue.

Interpretation skills are a key part of this specification and whilst candidates in general are attempting to make sensible comments, poor clarity of handwriting and poor clarity of expression hinder many. They should be advised to take more care in this respect to ensure that examiners are able to award the marks deserved.

There was some evidence of a lack of care in reading questions. For example in 12(b) many stated problems that might occur rather than assumptions being made, giving the converse of what was expected.

Report on individual questions

Question 1

The first question of the paper was very accessible with the vast majority of candidates gaining at least one mark, although many fewer managed full marks. Most commonly correct reference was made to the graph being 3D, but often a second comment was made making the same point. Less commonly candidates made reference to the vertical scale not starting from zero and a few mentioned that only certain months were included. Comments regarding the months were often poorly expressed or incorrect. For example some thought there were not equal gaps between the months shown, whilst others thought it unfair that 2012 appeared once compared with three bars for 2011. Credit was not given to criticism such as poor labelling or large jumps in the vertical scale.

Question 2

Candidates are quite familiar with certain aspects of data collection such as reasons for taking a sample rather than a census, and recognising a leading question; many scored well here.

In part (a) many were less successful, being unable to show a clear understanding of a population with many simply stating for example ‘the customers’. These are sampling units; the key feature of a population is that it must include **all** sampling units. Some misunderstood the question completely and described taking a sample. (e.g. ‘10 customers from each office’).

Part (b) and particularly (c) were perhaps the best answered parts, with standard responses being well known. Most candidates state that a sample will be quicker and cheaper, whilst a census will include the opinions of all. A not uncommon incomplete answer given in (b) was that ‘more varied results’ or ‘more data’ would be obtained. These did not score.

Many candidates failed to recognise the demand in part (d) and chose to describe how to take a random sample rather than explain what is meant by a random sample. It was not uncommon to see stated ‘a sample that is taken randomly’.

Part (e) generally scored well for candidates, usually recognising that the stated question was biased. Some gave this same reason twice (e.g. leading and biased). Some stated ‘misleading’ rather than leading, which is not the same, and did not score. There were many candidates who incorrectly seemed to think it was a closed question, believing it could only garner a yes/no answer. It is perhaps not clearly known why a closed question is more useful to make the analysis of results manageable. Other responses which did not score made reference to not knowing how much the sandwiches were, or that value for money is subjective.

In part (f) candidates were more successful with the disadvantage than the advantage, often making reference to interviews being expensive or time consuming. Credit was also not uncommon for stating that interviewees may feel under pressure to give the 'expected answer' rather than their own true opinion; this was often poorly expressed however, sometimes stated in terms of 'lying'. Acceptable advantages were less common, the best and most popular being that questions could be explained. Poor responses for the advantage suggested that interviewees were less likely to lie to someone's face, that their facial expressions could be read, or that follow-up questions could be asked. (Follow-up questions are also possible in a questionnaire.)

Question 3

The vast majority of candidates found this question very straightforward with very many scoring full marks. There were a number of candidates in part (c) who wrote comments about the number of people in Richmond being higher rather than the percentage. Although this was condoned in this instance, candidates need to be more aware of whether information relates to figures or proportions/percentages.

Question 4

Index numbers is a topic common with Foundation tier but was not well attempted by the Higher tier candidates. Fewer than half were able to successfully complete the calculation in (a); many divided the given figures the wrong way and some simply subtracted to find the increase as £145. Some with the correct calculation unfortunately gave their answer as £ or % losing the accuracy mark.
Part (b) was marginally more successful with many recognising that both prices had gone up, but full marks here was only likely for those who were correct in (a). This part was marked out for QWC so complete answers should have been given. The most common missing aspect from answers was showing an understanding of what an index number means in terms of a percentage change. (e.g. that there was a 20% rise for 3-bed properties). A not uncommon incorrect comment suggested that 3-bed properties were more expensive than 2-bed. Although this may well be true no evidence was given for this as we only had the index number for 3-bed properties.

Question 5

Many candidates scored well in parts (a) to (c) although about one in five made an error in (b) or (c), extracting information from the wrong column in the table. Part (d) was not an unusual demand for this type of question but only about two thirds of candidates scored the mark. Some gave vague reference to calculation errors or that perhaps not all spending had been recorded.
Although some candidates failed to extract the correct information from the table it was most common to see a first stage of working using correct values, gaining one mark. Unfortunately a disappointing few were then able to follow this through to a correct percentage increase. There were some candidates who obtained the answer inaccurately to only 2sf but failed to show any working, and so did not gain full credit. Some used the correct table figures but divided the wrong way.

Question 6

For Higher tier a surprising minority (more than one in five) were unable to identify the modal class from the table. Two fifths were unable to calculate the mean estimate in part (b), with the common errors being dividing either Σfx or the total frequency by 6. Some ignored the stated Σfx value choosing instead to find their own. A few candidates correctly worked out the mean estimate but then incorrectly wrote the class that it fell in (81 to 160) as their final answer. Commonly candidates were able to explain why it would be an estimate with reference to using midpoints or the data being grouped. Incorrect answers often suggested rounding error or that the answer should be a whole number.
In part (c) only a minority of candidates recognised that using narrower class widths is likely to lead to a more accurate estimate. Arguments were given suggesting that the value might either increase or decrease.
Candidates were not given an explicit trigger to consider skew so not too many made reference in part (d) to the data being skewed or having extreme values. Quite common were simple definition statements such as the median giving the middle value, or that the median would not be a decimal. Some candidates appeared unfamiliar with the term 'central tendency'.

Question 7

This question tested the selection of appropriate statistical techniques, which is key to statistical investigation. A small majority of candidates were able to correctly select three from five although they found parts (b) and (c) more challenging. It may be they were not sufficiently familiar with the vocabulary 'central tendency' and 'dispersion' in part (b). A few candidates in (a) think a line of best fit is used to see if there is a relationship between variables; this is a common misconception in controlled assessment too. A line of best fit is used if there is evidence of a relationship.

Question 8

Part (a) was well answered with very few candidates not scoring. Most were able to come up with a hypothesis relating to age or time of day and safety, although some would have needed further hypotheses to make them useful in practice. There were few questions given and some statements of 'aims', both of which do not score as hypotheses.

Reasons for a pilot study in part (b) were not well known on the whole. Too often reasons were far too vague: e.g. 'to look for mistakes' or 'to make changes'. Others referred to things such as checking for bias or spelling errors, which of course can all be checked by proof reading without the need for a pilot study. Candidates should be specific about the problems being checked for. The best comments that achieved a mark were for checking that questions could be understood, and that they produced the information needed. There were a number of candidates who clearly had no idea of a pilot study as they answered with the advantages of taking a sample. (e.g quicker, cheaper, etc.)

Most candidates gained a mark in part (c) for the recognition of it being a bad sampling frame. Less common were marks for a correct reason but they were still achieved by many, usually for realising that not everyone is in the phone book. It was clear that some candidates were not familiar with the idea of a telephone directory. Many of those not scoring well made references to the phone not being answered, not realising that the question was about a sampling frame being appropriate, rather than executing a sample.

Question 9

Part (a) was well answered with most candidates scoring both marks. The most common mistake was to use the wrong bar, coming up with 13% for the UK, or for poor accuracy in reading the scales.

Although most candidates had the correct conclusions in part (b), it was indicated for QWC and so required clear statistical reasoning from the charts. A common answer to not score for each hypothesis gave a correct conclusion but then repeated the hypothesis (e.g. 'the chart shows the USA percentage for housing is higher than that of the UK') without stating the feature of the chart which showed this. Although referring to the size of the bar sections would have been easiest, most who used the chart referred to percentages read from it. If done with sufficient accuracy this was enough to gain the marks.

Part (c) was answered less well with many candidates not realising that only having percentages meant that actual figures were not known. Some suggested that Tobi was correct although they had been told he was not. Others suggested that there may be food items included in the 'other' category. There were also those who incorrectly thought the relative population sizes were relevant.

Question 10

A good number of candidates were able to gain 2 marks for this part of the question. The most common reason for gaining only 1 was poor accuracy reading from the scales, either identifying the median as 40 or upper quartile as 49. A common error which was not able to gain any marks was to give the values as 32, 42 and 52, dividing the interval between 22 and 62 into equal parts. Usually the values were correctly used to produce a box plot within tolerance which scored full marks in (b). A small number of candidates misread the scale.

Part (c) was a very standard demand being a key part of statistical analysis. Candidates should be familiar with comparing a correct average, a measure of spread and the direction of skew. Whilst they were often using the correct statistical language, which is required, most commonly missed out was either reference to skew or to a measure of spread. Some had the incorrect direction of skew for their box plot or failed to mention USA being symmetrical. Nearly all drew a correct conclusion about the Headteacher's view with most making correct reference to medians. Some candidates perhaps feel it necessary to use all the lines in the answer space, which is not so. Many made additional reference to other values such as maximum age or individual quartiles, which do not generally score when comparing
distributions.

Question 11

Most candidates in part (a) were able to agree with Rupert and many of these commented correctly that there was a strong and/or positive correlation shown in the scatter graph to support that conclusion. Some failed to use the correct statistical language referring to 'relationship' rather than correlation. There were some who felt there was no relationship due to the two points which were separate from the others.
The plotting of the point for the Daily Mirror in part (b) was badly done by a surprisingly large number of candidates. There were many incorrect answers showing an obvious difficulty in putting these large numbers on the graph, with many misinterpreting the scales and a few who transposed the axes.
In (c) part (i) whilst many felt that the point followed the same trend of positive correlation they commonly missed the significance that it was well out of line compared with the other points. There was a minority correctly identifying it as an anomaly however. In part (ii) many confused comments were seen, some referring to an increase in sales. Some only referred to the Daily Mirror sales falling without comparing with the other papers, and so did not score. Of those who did identify that all sales had fallen, only a small minority recognised that the Mirror had fallen by significantly more than most, to gain both marks.

Question 12

It was evident that many candidates are not familiar with the capture-recapture method for estimating population size. Many candidates either left out part (a) completely or calculated $20-4+30=46$ assuming that all the fish had been caught over the two days. Little more than a third of candidates were able to complete the calculation correctly.
In part (b) many candidates described the method of calculation rather than the underlying assumptions. It was common for two or more comments to refer to there being no change in the population e.g. 'no fish joined the canal, no fish left the canal, none were born' etc, gaining just one mark. A constant population was the most common assumption to gain credit. Unfortunately there were a number of candidates who did not read the question carefully, who seemed to state problems rather than assumptions, and so did not score (e.g. 'some fish may have died' rather than 'no fish died'.)

Question 13

Very disappointingly few candidates knew how to find or use seasonal variation, with two thirds failing to score a mark on this question. Most commonly candidates found the average of the three values for Quarter 2, whilst some in part (b) simply stated the trend line value. A few in (b) seemed to guess a value, often about 660; this was out of the acceptable answer range but the calculation was required to be seen on this question anyway. Those who knew what seasonal variation meant commonly scored all 4 marks.

Question 14

Part (a) was the most likely to score on this question as many candidates understood that a census involved all fireworks, and so commented that there would be nothing left to sell. Just under half did not manage this however with some comments showing a clear confusion with the national census. Others failed by simply commenting that it would be expensive or time-consuming, missing the main issue of using up all of the stock.

In part (b) it was clear that very few candidates were familiar with the idea of a systematic sample. Some may have misinterpreted the demand of the question and set out to describe an alternative to taking a systematic sample. Commonly they correctly found that a sample size of 30 was needed but often went on to describe how to take a random sample. The most common mark to be scored in part (i) was for stating that every 100^{th} firework would be selected, but some incorrectly stated every 30^{th}. It was far less common for candidates to recognise that a random starting point was needed. Of those who showed some understanding of a systematic sample not many were able to recognise a disadvantage of the method in part (ii). A few realised there could be an underlying pattern which the sampling might coincide with, but most commonly a mark was gained for recognising that the method was not random. Typical answers not to score said that it was too expensive or that the sample size was too small. (This is a criticism of the sample size not the sampling method.)

Question 15

It was common to see correct answers or at least 1 mark being scored for the tree diagram in part (a). By far the most common errors were to have 0.4 and 0.47 in place of the 0.95 and 0.92 (thinking that the 0.05 and 0.4 had to add up to the 0.45 on the 1^{st} branch, etc.)

Part (b) was answered well by about half of those with a correct tree. Where 'exact' answers are found candidates should be encouraged not to round their answers: 0.07 here without a more accurate intermediate answer lost the final accuracy mark. The most common error was adding the two end branch probabilities to get 0.13

Very few candidates were able to recognise or deal with the conditional probability in part (c). The most common error was to just do a calculation along the top route in the tree to find the probability of an egg being free range and not suitable for sale. Another common error was to find a ratio using 0.05 and $(0.05+0.08)$ to get the answer $\frac{5}{13}$.

Question 16

For part (a) many students failed to find the frequency densities correctly but calculated $f \times x$ instead. Others simply constructed a bar chart for the given frequencies. There were a number of successfully completed tables of frequency density although this did not necessarily transfer to correct or completed histograms, as candidates either struggled to deal with the scale for the histogram or were poor with plotting accuracy. There were lots of unlabelled graphs whilst many incorrectly had frequency for the y axis.Only about one in six managed full marks for the histogram.

There was a roughly 50:50 split between those who could and could not carry out the standard deviation calculation for part (b). Although the formula is provided on page 2, not uncommon were the use of an incorrect denominator (often 6 or Σfx) or a failure to square root. There were also a number left blank. It was clear in part (c) that fewer than one in four candidates knew the ±2 standard deviations property of a normal distribution. Some used 1sd or 3sd instead. There were a number left blank. Those with correct boundaries in part (i) often did not refer to these in part (ii), instead taking the easier line of commenting on the shape or symmetry of the histogram. Some candidates were able to score the final mark in this way without the correct boundaries being found in part (i).

Question 17

Most candidates were able to score the mark for part (a) but little over a half could gain further credit in this question. The question discriminated well for those who were able to go further as there was a fairly even spread of candidates gaining the remaining marks. Part (b) was next most successful although common errors seen here were 0.4×5 or 0.6^5. In part(c) a number were able to identify the correct term but either had 0.4 and 0.6 transposed or failed to use the factor of 10. Some were clearly lost by this stage attempting to substitute into the complete expansion or just writing down the correct term in p and q. Part (d) was least successful although some managed to collect 1 mark for finding a further individual probability term. Only one in eight succeeded in achieving full marks on this question.

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 Edinburgh Gate, Harlow, Essex CM20 2JE