edexcel

Principal Examiner Feedback

Summer 2012

GCSE Statistics (2ST01)

Unit 1: 5ST1H_01 Higher

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the world's leading learning company. 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 for our BTEC qualifications.
Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

If you have any subject specific questions about this specification that require the help of a subject specialist, you can speak directly to the subject team at Pearson.
Their contact details can be found on this link: www.edexcel.com/teachingservices.

You can also use our online Ask the Expert service at www.edexcel.com/ask. You will need an Edexcel username and password to access this service.

Pearson: helping people progress, everywhere
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 2012
Publications Code UG033133

GCSE Statistics 2ST01
Principal Examiner Feedback – Higher Paper 1

Introduction

Candidates generally had time to attempt all questions.

As the second paper of this specification the emphasis on interpretation remained a challenge for some. However, candidates on the whole seemed better prepared for this aspect, finding the paper generally more accessible. Poor clarity of handwriting and especially poor clarity of expression are an issue for a number of candidates. Often candidates are asked to give reasons for their answers, i.e. to explain why they have reached their stated conclusion, but this instruction is sometimes misunderstood.

Candidates should be encouraged to show their working. This includes drawing lines on diagrams to show where values have been read off. A minority drew freehand when a ruler was appropriate

A small number of candidates incorrectly give probabilities as a ratio. Only fractions, decimals or percentages are allowable.

With comparison and interpretation, especially where a question is indicated as QWC (marked with *), candidates should be aware that correct statistical language is expected. When comparing distributions this should be using a correct average, measure of spread and direction of skew. It should also be noted that stating values (e.g. males' and females' medians) is not a comparison in itself; when values are stated there needs to be use of comparative language (e.g. " ... which is larger than ...").

Particular care should be taken when making comparisons as to whether numbers or proportions/percentages are known. Where more than one mark is available for a question, candidates should be aware that the number of marks generally indicates the number of comments expected.

Report on individual questions

Question 1

This proved a straightforward start to the paper for most candidates with many scoring full marks.

Question 2

A generally very accessible question with most candidates able to use the key correctly to complete the choropleth map in part (a).

In part (b) the best answers referred to 'the top right' of the grid or to the north east corner. Those referring to a list of squares were allowed credit, although this is not ideal in describing an 'area'. One common answer which did not score here, referred to the single square F1 (as it had the highest number of tiles); this was not allowed as an 'area' and did not make use of the choropleth map. Whilst most correctly gave a reason why they had reached their conclusion, using the key (either referring to darker shaded or '16-24' squares), there were some who decided to suggest a reason why more tiles were in that area. The most common disallowed answer referred to a building having been there.

Question 3

On the whole this was a successful question for many candidates with the exception of part (b).

Most were able to give at least one appropriate advantage of using a sample in (a), usually related to being quicker or cheaper. Wrong answers which did occur suggested that a sample would be more reliable/less biased or simply described a sample or census. A small number of answers seemed to describe advantages of a pilot study.

Very few candidates seem to be familiar with the idea of a sampling frame; common wrong answers to (b) were to suggest sampling methods, or a suitable population. Whilst referring to all the adults in Collis is a step towards the correct answer, the idea of a *list* is key in describing a sampling frame. A telephone directory was suggested by some: although this is a list it is not acceptable as it is not a complete list of *all* adult residents.

The majority of candidates in (c) correctly concluded that the sampling method was not good, giving at least one appropriate reason typically focussing on it being only one street. Some suggested reasons for and against it being a good sampling method without reaching a conclusion.

In (d) most identified that (i) was a leading question (sometimes using their own words) and that (ii) contained overlaps. The most common answers not to score referred vaguely to poor wording in (i), and suggested that the "£2 or more" option was too wide in (ii). Very few picked up on the lack of a time frame in (ii) which was also a flaw.

Question 4

Nearly all candidates were able to find the cumulative frequencies with only a few making arithmetic errors. A small minority simply transferred the frequencies to the new table. Typically they were then successful in at least plotting the heights correctly in (b). Not all positioned points at upper class boundaries; the most common error was to plot at mid-points. Whilst some bars or 'lines of best fit' were in evidence and some stopped after plotting points, the majority joined with a sensible curve. Candidates should avoid drawing double lines. Some joined point to point with a ruler which is also acceptable for a c-f graph.

Those with a c-f graph attempt were usually able to read off an appropriate answer for (c) although very few drew lines to show where they were reading off; drawing lines should be encouraged. There were some without a c-f graph who simply wrote "30" for their answer with no justification, probably because it was the boundary between the two central classes - this gained no credit. Most were able to make a correct simple statement based on their answer to (c).

Question 5

It was pleasing to see many correct, ordered stem plots in part (a); not many left it unordered. The most common error was in missing out one or more values. There was space to draw an unordered plot first but many missed out this stage of working. This might be responsible for some of the errors but certainly resulted in some unclear or unreadable diagrams. Candidates should be encouraged to have equal spacing between leaves, thus maintaining the dual functionality of a stem plot, in that the shape of it enables it to act as a bar chart.

Using their correct stem plots, most were able to identify the correct median in (b).

Many correctly calculated the mean for (c).

Some totalled incorrectly but showed sufficient working, including the division by 16, to score 1 mark. The effect of the new weight on the median and mean in part (d) was generally quite well known although some may not have gained the mark in (ii) for an incomplete answer. The mean does increase but is the same to 1 d.p. in this case. Some candidates who unnecessarily did a new calculation only worked to 1 d.p. and incorrectly concluded that the mean is unaffected. For those correctly identifying that mean would increase the reason was sometime poor, e.g. simply referring to a bigger total to divide. A few candidates confused the effects on the median and mean of including the new weight.

Most were able to identify the five children referred to in (c) but many did not read the question carefully and gave the incorrect answer of 5/16. A very small number wrote their probability as a ratio which is not acceptable.

Question 6

It was pleasing to note in (a) that many candidates understand what a hypothesis is, but there are still a significant few who incorrectly either gave a question or stated a plan. Some failed to realise that the *length* of unemployment was key to the survey, instead stating that, e.g. “there are more males unemployed than females”. A smaller number incorrectly thought the data could be paired, suggesting that e.g. “the longer that males are unemployed, the longer females are unemployed”. (The only correct conclusion that could be reached in (c) from these errors would be that “it is not possible to say”.) The best answers were often the ones worded most simply, e.g. “men are unemployed for longer than women”.

The diagram in (b) was generally well completed with many gaining full marks. Of the few who did not score, a small number left the graph blank whilst others failed to use cumulative percentages, instead plotting the given percentages.

Reaching a correct conclusion in (c) from the graph was often done well if they had a sensible hypothesis in (a). Whilst some gained credit for re-stating (or refuting) their hypothesis from (a), the best answers simply stated that it was e.g. “(shown to be) correct”. Less successful was the reasoning, where many candidates referred to *numbers* of males or females unemployed in each category, e.g. “more males than females are unemployed for longer periods” - nothing is known about *numbers* as we only know *percentages* for each gender.

Question 7

Whilst there were many correct answers in (a) some candidates may have wrongly assumed that each word should be used only once, meaning that one of their answers, usually to part (iii), was wrong. Most correctly identified ‘speed of serve’ as being continuous.

The majority of candidates were able to rank ‘position in world’, with a minority using reversed ranks. Many were able to go on to find squared differences in ranks and then go on to substitute correctly into the formula. Arithmetic mistakes were not uncommon including negative squares of negatives, or incorrect addition. These errors led to the final accuracy mark being lost. Some candidates subtracted using the wrong columns, losing method marks, and leading to impossible values of r_s. Only occasional errors were made in the use of the formula, such as omitting the 6 or the subtraction from 1. Answers were usually rounded to appropriate accuracy.

Whilst either order of ranking was allowed in (c), those with reversed ranks had more problems correctly interpreting their negative correlation in part (d). Most were able to gain credit for identifying their correlation as positive or negative or weak, but those with incorrect high values for r_s often referred to strong correlation without mentioning positive or negative. There was potential for confusion in the interpretation with the better/higher ‘position in world’ being the smaller number; those who attempted an interpretation usually managed to make sensible comments.

Question 8

Most candidates were able to pick up both marks in (a) for two sensible disadvantages of using a census, usually relating to the time or cost aspect.
A few candidates gave very brief answers to (b) despite it being worth 5 marks. Most gained at least 1 mark but only the best candidates scored all 5. The majority realised that stratified sampling would be most appropriate and usually named this method. Most commonly gained was the mark for taking proportionate numbers from each year group, although some incorrectly stated that the same *number* of students should be surveyed from each year group. The question asked for a detailed description: some gave detailed calculations of strata size for every year group but then said no more, so failed to address key aspects of the sampling method.

The next most commonly gained mark was for the randomness aspect. Usually candidates referred to random sampling within each year group or referred to using random numbers. Although 'names in a hat' was allowed to imply the random selection aspect, this is not appropriate for selections from a large group and did not contribute to the award of QWC marks. Some suggested systematic sampling within year groups, but this is not necessarily random.

Numbering of students was the most likely aspect for candidates to omit in their descriptions. Correct statistical language was expected for QWC to allow all marks to be awarded.

Some candidates had confused answers, stating stratified sampling but then describing a systematic or random sample.

Question 9

Candidates were helped by the first value being given in the table and many correctly completed the table in (a). The most common error was due to an incorrect class width, often the second, or one of the last two.

There were some errors in transferring frequency densities to the histogram in (b) but most candidates scored well here. An incorrect width for the first or last bar, extending to the edge of the graph paper was sometimes seen. Use of a ruler would have improved some candidates' diagrams.

In part (c) the award of 1 mark was common for suggesting that young drivers had more accidents or were more likely to have an accident. However, reasons given often explored why they might have more accidents, and so did not gain the other mark. Candidates should be explaining the feature of the histogram which led them to their conclusion. Few referred to the skew in the diagram; instead the mark was often gained for stating that young drivers had the highest frequency density or highest bar. Reference to frequency or area of bars is incorrect here.

Question 10

Although working was often not shown, candidates generally were successful in both finding and plotting the moving averages in (a). The most common mistake seen was in plotting the values as a continuation of the time series points rather than the moving average points.

The moving average points were quite close to a straight line which meant that many candidates were able to draw an appropriate trend line. A common reason for not gaining the mark in (b) was drawing a freehand line or curve, or for joining point to point. Trend lines should be ruled. Most candidates were able to describe the downward trend for part (c) even if no other marks were gained in the question. Some referred to negative *correlation* which is not appropriate on a time series graph.

Only a minority of candidates made a good attempt at parts (d) and (e) which required the trend line to be extended. Some of these made accuracy errors in reading seasonal variations from the graph but most full attempts at seasonal variation scored both marks in (d). A common error was to simply average the seasonal values for quarter 3.

Those with correct attempts at (d) were usually also successful in part (e), although sufficient working to justify the final answer was not always seen. One or two candidates managed to pick up a single mark in (e) for finding the trend line value from their extended graph.

Question 11

Parts (a) and (b) were often correctly interpreted and answered well although there was some inaccuracy in reading the time scale for part (a).

While most candidates were able to identify the cross on the box plot as an outlier (or anomaly) less were able to show why for (c)(ii). A few attempted *explanations*, e.g. “it was a very slow runner”, or was “far from the other values”, or just left it blank. Many gained 1 mark for correctly finding the IQR. Fewer then correctly added this to the upper quartile, the most common error being to add to 7.4 (the end of the whisker). Those who correctly found the outlier boundary of 7.5 did not always then indicate why the cross was an outlier, i.e. that it was above this value, or that any value above 7.5 was an outlier.

Comparison of distributions is a common demand in statistics. Some candidates focussed solely on contextual comparisons which could score at most 1 mark in part (d). Comparisons need to be made using measures of central tendency and measures of dispersion as well as skew. Credit is not gained for comparing other values such as maximum values or individual quartiles. Correct descriptors are required, so for box plots we need to see *median* (not mean or average) and *IQR* or *range* (not spread). Candidates’ descriptive skills were often poor; e.g. “more distributed”, “evenly spread” or “equal skew” have little statistical meaning. It should be noted also that IQR is a statistical measure (i.e. a value), not a description of the box length, so should be compared using e.g. *bigger* or *larger* rather than *wider*. Some did not gain marks because they listed values, e.g. medians, without making a comparison. Those mentioning skew mostly identified it correctly but some incorrectly had the female box plot as negative skew.

Question 12

In part (a) most candidates were able to show why the mean was 14 but many had more problems with the standard deviation. The formula is given on page 2, but some were unable to select the appropriate form. As the answer was given, sufficient working needed to be shown either by working to more than 1 d.p. or by simplifying the square root before giving the answer. A common error was to divide 4220 by 280 and then square root (which gives 3.88).

Standardised score was often poorly known, with many either guessing at a calculation or leaving (b) blank. There was a minority of sensible attempts with many of these getting the correct answer, but some subtracted in the wrong order to obtain an answer with the incorrect sign.

Most with answers to (b) were then able to reach a correct conclusion in (c) but reasons were often insufficient or wrong. Many candidates did not score as they simply referred to how close the values were to the mean or to zero, without indicating that one score was *higher* than the other - this is the key, irrespective of signs.

Question 13

Most candidates correctly identified the Lake District as having the larger population in (a) and usually made reference to the pie chart being larger in some way. Although reference to the different radii was allowed in this part, candidates should be aware that it is the *area* that represents population size (i.e. frequency); only a minority made reference to area. It was apparent that a small number of candidates were unaware of comparative pie charts as they said the different sizes were irrelevant or that it was misleading. Some candidates failed to read the question carefully, focussing on the individual age groups being similar rather than considering the whole population.

Nearly all candidates identified the correct age group in (b). The most common incorrect answer was 80+, where candidates wrote the largest, i.e. widest or oldest, age group from the key.

In (c) it was again common for candidates to correctly identify the Lake District as having the larger number of people (in the given category), but a common error was to simply refer to the equal proportions without answering the question. A more thorough reason (i.e. reference to *area*, which is key in the comparison) was expected in this part, and most often was not seen; it was not sufficient to re-state that the Lake District pie chart was larger.

Question 14

Candidates often find Venn diagrams difficult, and despite being given the figure at the centre in part (a) they still found it a challenge here. It was common to see the remaining overlap values entered correctly for the first mark, but using subtraction to find those who played only one instrument caused a problem for many. Of those who correctly identified the single instrument pianists or violinists some failed to go on to find the oboe-only musicians, not realising there was information they had still not used (the total of 100).

The most common (and incorrect) answer to part (b) was $\frac{30}{100}$, ignoring the 32 in the diagram. Clearly some candidates misunderstood the significance of the word ‘only’, not distinguishing between ‘play piano and oboe’ and ‘play piano and oboe only’.

In part (c) there was only a minority of candidates who recognised the question as being about conditional probability, so using a denominator of 84 not 100. Of those who did have a correct denominator it was very common to see the same problem as in part (b), excluding the 32 and giving a wrong answer of $\frac{12}{84}$.
An incorrect numerator of 12 was more common than an incorrect denominator.

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

Further copies of this publication are available from
Edexcel Publications, Adamsway, Mansfield, Notts, NG18 4FN

Telephone 01623 467467
Fax 01623 450481
Email publication.orders@edexcel.com
Order Code UG033133 Summer 2012

For more information on Edexcel qualifications, please visit
www.edexcel.com/quals

Ofqual

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Rewarding Learning