

Principal Examiner Feedback

Summer 2012

GCSE Statistics (2ST01)

Unit 1: 5ST1F_01 Foundation

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 UG033130

All the material in this publication is copyright

© Pearson Education Ltd 2012

GCSE Statistics 2ST01

Principal Examiner Feedback – Foundation Paper 1

Introduction

The vast majority of candidates had time to attempt all questions.

This is the second examination of the new specification which places greater emphasis on making deductions and drawing conclusions. Overall, it was pleasing to see that most centres have adapted their teaching to prepare their candidates for the demands of this specification. However, poor clarity of expression remains an issue for a number of candidates, particularly when making comparisons.

Candidates should be encouraged to show their working as some may have picked up more credit when their answers were incorrect. There was a good standard of diagrams; candidates should take care however with reading correctly the scale on axes. They should also be encouraged to use a ruler when drawing bar charts, histograms, lines of best fit, etc.

With comparison and interpretation, especially where a question is indicated as QWC (marked with an asterisk, *), 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. 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

The first question of the paper was very accessible to the majority of candidates and many scored full marks.

The accuracy of the 33.5 bar in part (a) proved difficult for a minority of candidates and some incorrectly drew this at a height of 35.

Only a small number failed to read the word 'second' in part (b) and incorrectly gave the country with the highest number of internet users, Germany.

Spain was correctly identified in part (c) by nearly all candidates.

Question 2

A small proportion of candidates were able to calculate correct angles for the pie chart. Many would have benefitted from showing the angle calculations in part (a), particularly when they did not have a protractor to accurately complete the diagram. The majority of candidates were able to score one mark for labelling the sections of their pie chart. A small proportion of candidates were able to calculate correct angles for the pie chart. Many would have benefitted from showing the angle calculations in part (a), particularly when they did not have a protractor to accurately complete the diagram. The majority of candidates were able to score one mark for labelling the sections of their pie chart.

Candidates were usually able to offer some sensible comparisons in part (b) between the sectors for the different years but occasionally were unable to articulate this precisely enough in context to score the marks for their intended comparisons.

Question 3

Nearly all candidates successfully completed the two-way table in part (a). Most went on to give a correct answer to (b)(i).

In (b)(ii), $\frac{23}{57}$ or $\frac{1}{23}$ were common errors. When giving probabilities candidates should be reminded that only fractions, decimals or percentages are acceptable, not ratios. Fractions need not be cancelled down, but if doing so the unsimplified fraction should be written first.

It was good to see candidates using figures from the two-way table to support their response in part (c). Most picked up at least 1 mark in part (c), often recognising that there are 'overall more females than males'. Some candidates did not express their conclusion whilst others went on to incorrectly conclude that the two-way table indeed supports the newspaper's claim.

Question 4

Most candidates were unable to come up with scatter as the diagram for representing bivariate data in part (a). The most common non-scoring diagrams were bar chart or line graph. Others offered methods for collecting data such as a tally chart. Again in part (b) most struggled to identify that reaction time is the response variable.

The calculation of the mean in (c) was much more successful. A handful of candidates made calculation errors but were able to gain a method mark for showing their working.

Question 5

Most candidates are well aware of the advantages of a sample and completed part (a) well. A common error continues to occur when candidates attempt to make a converse statement but neglect to use the term 'census'. Quite a few believe a sample is more reliable; others think a census applies to the whole country and not just to Collis.

In (b) only a significant minority of candidates produced correct answers. It is clear that candidates remain unfamiliar with the term 'sampling frame,' many offering sampling methods, such as stratified sample, in its place.

In (c) most candidates identified this was not a good sample and a large proportion of these acknowledged the fact that one street is simply not representative. 'People will not do the questionnaire' was seen as a common incorrect response.

Though clarity of expression varied between candidates, (d)(i) was generally identified as a leading or biased question. In (d)(ii) many candidates correctly expressed that the answer choices overlap. However, some felt that there were not enough options or the options did not include enough variety.

Question 6

Question 6 was accessible to all candidates with most scoring at least three marks here while parts (b) and (e) served to distinguish the more able candidates.

Plotting the points in part (a) and drawing the line of best fit in (c) were generally within tolerance. Some marks were lost due to small inaccuracies resulting from failing to read the scales properly. It was pleasing to see that most candidates used a ruler to draw the line.

The majority of candidates correctly identified positive correlation in (b), but most found it difficult to give an appropriate interpretation. 'Females live longer than males' was the most commonly seen incorrect interpretation of positive correlation.

Part (d) was answered well by nearly all candidates.

Lancashire was chosen by a high percentage of candidates in part (e). Less successful was giving an acceptable reason as to why the estimate is more reliable. The terms 'interpolation' and 'extrapolation' were rarely seen. It is common to see responses similar to 'Lancashire is closer to other points' which will not score the mark as it does not sufficiently express that it lies within the given data set.

Question 7

Most candidates appreciate that a hypothesis must be given as a statement rather than a question

Part (b) was far less successful with most giving incomplete responses by failing to include 'all'. Others gave sample sizes (e.g. '50 boys and 50 girls' or '20 students from each year group').

In part (c) the data from the 'favourite vegetable' variable was generally identified as qualitative, however most confused continuous for discrete data. Some candidates now clearly understand one of the reasons for conducting a pilot study in part (d), whilst a significant number are still unclear as to what one is. Some attempted to give advantages associated with samples such as 'quick' or 'easy'.

Part (e)(i) was generally answered well (although not always expressed well) referring to the question being open or lacking response boxes. A common answer to gain no credit here was given by candidates who expressed that some people may not like vegetables. In (e)(ii) many were able to offer a better question with response boxes. Others changed the question to 'what vegetable do you like?' and lost the mark by not including any answer choices.

It is apparent that the calculation of the sample size in part (f) was too difficult for the majority of candidates with many opting to simply divide the 40 pupils evenly among the three groups. Another common mistake was to multiply the correct fraction by 100 instead of 40.

Question 8

Most candidates were able to read the values correctly from the table in this question and extract the correct answer in (a)(i) and (a)(ii).

Part (b) had a lower success rate with some opting to give their own reason rather than stating the correct reason from the table.

Question 9

The majority of candidates completed the choropleth map correctly. A significant number of candidates, however, did not attempt this question at all.

In part (b) many picked the single square with the most tiles rather than the area. The question asked candidates to use the choropleth map but a large number did not refer back to the map when stating their reason. Reasons given should be statistical reasons.

Question 10

Most candidates were able to describe the trend in part (a) as decreasing. Those referring to individual years did not score the mark (e.g. 'it goes up in year 7 then goes down after that').

Many attempted to give reasons why the viewing figures were decreasing rather than stating the problem with the vertical axis in part (b).

A variety of answers were seen in part (c) including the calculation of the range instead of the interquartile range. Some simply stated the upper quartile.

Part (d) was assessed QWC, so answers had to use correct statistical language to gain credit. Despite having four marks, which should have indicated to candidates how many comments were required, many made only one or two comments. It was evident that the context of the question confused many as candidates often referred to pupils in Year 9 and Year 10, rather than the 9th and 10th years of Big Brother. The most common reasons for not scoring marks include using 'spread' in place of range/IQR, comparing minimums/maximums or stating values without actually comparing them. It was pleasing to see a good number of three and four mark responses.

Question 11

For those who were able to complete the cumulative frequency table in part (a) correctly, many went on to score at least 2 marks for drawing the diagram in part (b). Candidates still incorrectly plot their points over class mid-points. Some attempted to draw a line of best fit through the plotted points whereas others did not attempt to join up the points at all. Many histograms were also seen.

In (c) candidates should indicate clearly on the cumulative frequency diagram how they are estimating the median. A lack of caution when reading the horizontal scale meant that some candidates lost a mark here. For those with an answer to part (c), the majority gave a correct response in part (d).

Question 12

Most candidates were able to place the letters on the correct side of 0.5 in part (a); however it was common for some to believe that 'Getting at least one Tail when two fair coins are thrown' was almost certain.

The majority of candidates chose the correct method for estimating the probability in part (b).

The fact that all probabilities sum to 1 was understood by nearly all candidates as most scored the mark in (c)(i). A good proportion also went on to score both marks in (c)(ii), but a few multiplied the required probabilities instead of adding them.

Very few candidates answered (c)(iii) correctly, with most multiplying 0.3 by 2 instead of by itself. Some that knew the correct method unfortunately lost a mark for evaluating 0.3×0.3 as 0.9

Question 13

It was good to see that most candidates made an attempt at the final question on the paper. Most were able to complete the histogram correctly in part (a); however several forgot to label the axes.

Part (b) was met with varied success. Quite a lot calculated fxw in the table, either correctly or consistently, but then failed to use these for calculating the mean. Division by 5 was seen as the most common error. Answers of 105 without working could not score marks as this answer could have been obtained from incorrect methods and candidates must be encouraged to show their working even when using a calculator.

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 UG033130 Summer 2012

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

Pearson Education Limited. Registered company number 872828
with its registered office at Edinburgh Gate, Harlow, Essex CM20 2JE

Ofqual



Llywodraeth Cynulliad Cymru
Welsh Assembly Government

