Principal Examiner Feedback

Summer 2013

GCSE Statistics (5ST1F)
Paper 01

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from 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

GCSE Statistics 5ST1F
Principal Examiner Feedback – Foundation Paper 01

Introduction

Candidates on the whole seemed to find the paper generally accessible although a number left the final question blank; this may well have been due to the challenging topic rather than running out of time.

Interpretation skills show signs of improvement but this aspect remains a challenge for some. Poor clarity of handwriting and especially poor clarity of expression are again an issue for a number of candidates; they should be advised to take more care in this respect to ensure that examiners are able to award the marks deserved.

Often a question demands a reason to support an answer. Especially when there are only two possibilities a mark may not be earned without a supporting reason. There were occasions however, such as in Q5(b), where some candidates provided reasoning but failed to state a conclusion. Candidates would be advised to re-read a question after writing their response to make sure they have answered it.

Candidates must be encouraged to show their working as some may have picked up more credit when their answers were incorrect (e.g. questions 6(c), 7(d), 10(a)). This includes drawing lines on diagrams to show where values have been read off. The standard of drawing diagrams was often quite pleasing although a minority drew freehand when a ruler was appropriate (e.g. when drawing box and whisker diagrams).

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. zoo deer and wild deer 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 …"). 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 – this was apparently not appreciated by many candidates on some questions.

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.

Most candidates produced a table with two or three columns or rows. As a data collection sheet 'tally' as a heading would have been best, although some gained

credit with 'frequency' or 'number of cars'. Some candidates failed to label and so lost a mark. A very few candidates only gave four of the colours in their table.

Some candidates drew axes for a bar chart which is not really appropriate for collecting data.

In (b) many candidates identified an appropriate diagram. Common incorrect answers included diagrams only appropriate for numeric data, or tally chart which is for collecting data not representing. A small number seemed to be describing a shape they would use (presumably for a pictogram) but without saying pictogram.

Question 2

This question was generally well answered with extremely few incorrect answers in (a) or (b). Although not always well expressed, most candidates were also able to make a correct comment in (c), often recognising France as the exception. The main flaws were that some candidates referred only to individual countries rather than the overall trend, or they only mentioned changes without saying 'increase'.

Question 3

This was often answered well with many candidates choosing the correct three words. Some were less successful in placing the cross in an appropriate position on the probability scale. They should think carefully about the position rather than for example just going half way between 0 and 0.5 for the unlikely event of winning a lottery.

Question 4

Nearly all candidates successfully completed the two-way table in part (a), with about two thirds going on to give a correct answer in (b)(i). A small number appeared to be giving a conditional probability of $\frac{7}{19}$ or $\frac{7}{12}$. It should be noted that conditional probability is not tested at Foundation tier.

In (b)(ii) far fewer candidates were successful with $\frac{5}{30}$ being a very common incorrect answer. Each part of a question needs to be read more carefully.

A small number of candidates answered with a number of children rather than a probability or gave an answer in ratio form. It must be noted that probability answers must be given as fractions or decimals (or percentages) only.

Question 5

More than two thirds of candidates correctly completed the tree diagram for part (a). However tree diagrams are often difficult for some candidates; common incorrect answers were putting $\frac{1}{6}$ on each branch or sometimes $\frac{1}{4}$.

90% of candidates scored just one of the two available marks in part (b). Usually this was by agreeing with the statement but either referring to a six being unlikely on just one die, or by having the common misconception that the likelihood of double six was $\frac{2}{12}$ (as they would need the two sixes from twelve available numbers as they perceived it). To score the second mark a small number of candidates were able to give a coherent argument along the lines of six being unlikely on each die, but only the strongest candidates used the more rigorous approach of evaluating the product of probabilities from the tree diagram to get $\frac{1}{36}$.

Question 6

The majority of candidates were able to find the correct supermarket in (a).

Whilst many performed a correct calculation in (b) very few realised the table gave figures in £'000s and gave a fully correct answer; this was condoned however. There were arithmetic errors resulting in answers such as 1.89 instead of 1890(000), which again demonstrated a lack of recognition of the significance of the figures in this real data problem.

Very few candidates showed working in part (c) but a large number gained credit for at least one correct answer, often both. Typical errors were adding for all supermarkets or finding the average rather than total. These demonstrated a lack of care in reading the question. Sensible comments were usually made for their values in answering (d). This usually made reference to the increase in market share although when figures were quoted they were sometimes not correct.

Question 7

It was clear that a minority of candidates did not know how to use the stem and leaf diagram.

In part (a) the stem plot was often completed correctly or was left unchanged. The mode was often correctly identified but incorrect answers seen were 31 (the highest value) or '2 and 3' presumably as the leaves 2 and 3 appeared in the stem an equal number of times. The median of 12 was often correctly identified (or 6 if they had not added to the stem plot), but a common incorrect answer was 2, just writing the leaf.

The mean calculation in (d) was the least successful of the averages (just over half scored full marks) with errors such as incorrect addition, finding the total without dividing, or dividing by 9 or 10. Using only the original nine data (75÷9) was not uncommon. Those candidates correctly summing the 11 data often divided by 10, possibly because one of the figures was zero. Some candidates got the three averages mixed up. A few got all the correct answers but not necessarily in the right order.

In answering part (e) 'mode' was the most popular candidate choice, with the reasoning that it had occurred most often so was most likely to happen again. This is not appropriate as it is most unrepresentative for this data. The flaw in the common incorrect reasoning is that the number caught the next day is not a random selection from the observations so far. Of the candidates who correctly stated mean as the

most appropriate average very few gave a good reason, instead stating how it was calculated.

Question 8

This question was accessible to all candidates. Nearly all correctly identified the anomaly in part (a) and commonly the line of best fit in (b) was appropriate. It was pleasing to see that most candidates used a ruler to draw the line. There were a small number who ignored the instruction to only use the other seven points, sometimes with unusual results. It should be noted that lines of best fit are expected to be straight, so those drawing curves or freehand lines usually do not gain credit.

The majority of candidates scored 1 mark in (c) for correctly identifying positive correlation, or fewer for giving a correct interpretation, but not too many did both to score 2 marks. 'The line slopes up' is not interpretation but was commonly seen. Most candidates correctly attempted estimates in parts (d) and (e), although some marks were lost due to inaccuracies resulting from failing to read the scales properly.

For (e) the line should have been extended to obtain the estimate. Candidates should be encouraged to draw lines on the graph to show where they are reading values. The terms 'interpolation' and 'extrapolation' were rarely seen in candidates' reasons.

Acceptable equivalent descriptions of these were infrequent and often poor. In (d) it was common to see responses similar to 'it is close to other points' or 'because it follows the pattern' or even 'because I used the line of best fit', none of which will gain credit as they do not sufficiently express that it lies within the given data set. In (e) many gained a mark for stating their value was unreliable but could not say why. Some incorrectly thought it would be reliable, again often stating that it was from their line of best fit or followed the pattern.

Question 9

Part (a) of this question was accessible to all candidates with very many scoring three marks and most scoring at least one.

In part (b) QWC was assessed, but candidates' responses usually fell short as they did not use correct statistical language. Comparison of distributions is a common demand in statistics. A large number of candidates focussed solely on contextual comparisons and commented on the eating habits of deer but did not mention even one statistical term. 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'/'range' (not spread). It should be noted also that IQR/range are statistical measures (i.e. values), not a description of the box length etc, so should be compared using e.g. 'bigger' or 'larger' rather than 'wider'. Some did not gain marks because they just listed values, e.g. medians, without making a comparison. Those mentioning skew mostly identified it correctly.

Despite having four marks, which should have indicated to candidates how many comments were required, many made only one or two brief comments.

Question 10

Whilst nearly half of candidates found the correct value in (a), gaining 3 marks, those who did not usually scored zero as working was commonly not shown. Some of these might otherwise have gained method marks if only minor slips were made. There were a small number who used values from the final column. Plotting the final moving average was more of an issue; either the horizontal position was not correct or it was not attempted, even when they had found the correct value.

A large number of candidates correctly identified the 3rd quarter for (b), but some referred specifically to 2007 so losing the mark.

It was apparent in (c) that a number of candidates do not understand the distinction between trend and seasonal variation, giving a description of the latter. A large number of candidates commented on the fewer visitors over time, gaining a mark for interpretation, but describing it as a 'falling trend' was uncommon and 2 marks were rarely awarded. It should be noted that a simple comparison of 2009 with 2007 is not equivalent to describing a trend.

Question 11

In part (a) many appropriate advantages were seen, most often scoring one mark for suggesting that a questionnaire would be quicker. Only a small number gave the more complete answer that it can be a quicker way to collect a large number of responses. Other popular correct answers were that people can take their time or that it is cheaper to not pay interviewers. A disadvantage was more difficult for some candidates, the most popular (but not too frequent) correct answer being about non-response. Common answers which were not accepted were that people were more likely to make up answers, that it would cost more, or that detailed answers could not be given. Some candidates did not read the question carefully and gave two advantages rather than one advantage and one disadvantage.

Part (b) was answered quite well with most candidates gaining at least 1 mark. Bias was most commonly commented upon, sometimes twice in different ways but less so the lack of response boxes. Some incorrectly stated that it was a closed question. Candidates who did not score often said people might not know about the theatre or how much was spent, or commented on whether it was actually the council's money.

Most candidates were able to write an unbiased question for part (c), but often response boxes were omitted or were overlapping. Although 'incomplete' boxes were not penalized, some choices of box labeling was poor, e.g. '£1 to £2, £3 to £5, over £6'. Candidates need to label clearly; use of inequality signs is best discouraged as they are often misused (e.g. '£1<£2') and so do not score. A minority failed to score as their question did not relate to the cost of parking as requested.

Many candidates know quite well the standard advantages of sampling and over half gained at least one mark in part (d). However, a surprising number misunderstood the question and gave advantages of a pilot study instead.

In part (e) many candidates thought this was a good sample, often stating 'because it is random'. These candidates had missed that selecting from the telephone directory is not a random sample of the population. But a majority of candidates realized it was not a good sample, with many giving the small sample size as a reason. Not all

reached a conclusion. It was apparent that a number of candidates were unfamiliar with a telephone directory stating that it would only include businesses.

Question 12

Many candidates did not recognise that 'simple random sampling' is a specific technique and so described other methods, often selecting every 20th or one from each class. The poorest examples demonstrated no use of statistical method, such as choosing the best behaved or asking for volunteers.

This was a QWC question so a clear description was required. Responses which did gain at least one mark usually referred to picking names from a hat, or less often to numbering the students. The second mark was also often scored, for a more complete answer of selecting eight from the hat or using eight random numbers. Some candidates unnecessarily gave a detailed description of how to get a random number from their calculator. The final mark was gained less often as candidates did not generally make the explicit link between numbers/names drawn and the selection of boys for the sample/event.

Question 13

Many candidates are aware that a hypothesis needs to be in the form of a statement, although over a quarter incorrectly stated a question in answer to part (a), or gave a description of what they would do.

Control groups is not an easy topic for candidates at this level and few demonstrated a strong understanding of the concept in part (b). Some did suggest in part (b)(i) that it was to make a comparison but this was often poorly expressed. A common incorrect response was 'to avoid damaging all the crops if it went wrong'; some gave the impression that the control group would have the new fertilizer, which is incorrect.

The most commonly scored mark was in part (b)(ii) for stating that two crops needed to be grown, one with the new fertilizer and one with the old/no fertilizer. However, few then stated that other conditions should be kept the same. There was a small number of very eloquent answers however.

Question 14

Foundation candidates usually find index numbers a demanding topic and this was no exception with very few scoring more than 1 mark for the question. In part (a) about half of candidates stated that the mean earnings in 2004 were higher than in 2000, but very few stated that it had gone up by 12%. Those who did usually scored all marks in (b) too.

Part (b) proved beyond most candidates with few gaining marks. Although there were some good complete answers, most did some incorrect calculation such as dividing by 123; many simply skipped this question. There were those who multiplied by 123 but failed to then divide by 100, seemingly unaware that their answer was unusually large. The simplest correct calculation was 1.23 x 14 000, but various other correct methods were seen.

Grade Boundaries

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