STATISTICS

Paper 4040/12
Paper 1

Key messages

It is as just as important to be able to explain and interpret the values of statistical measures as it is to be able to find them accurately. Statistical calculations have their greatest value when used to understand real-life practical situations.

When drawing conclusions about a practical situation following statistical analysis, use should be made of any numerical values found, and of any additional specific information given.

General comments

Once more the overall standard of work involving calculations of a routine nature was good. This was particularly true of the topics of basic statistical measures, crude and standardised rates, and line of best fit. Answers were generally more limited on the explanation and interpretation parts of questions. This was especially true of parts of the question on one of these topics (see **Questions 7(c)**, **(e)** and **(g)** below). Some answers were often of a very general nature, without specific reference to, and use of, information given in the question.

As has been emphasised previously in these reports, candidates of this subject need to appreciate the essential practical nature of Statistics in its application to a wide variety of real-life situations. It is certainly important for a candidate to be able to perform routine calculations efficiently and accurately; but it is just as important to understand why such calculations are being carried out, what they demonstrate, and how they help in providing useful practical information about the situation being considered.

Comments on specific questions

Question 1

On sampling methods, **part (b**) was answered best. Even though the first three schools in the list had been chosen, it was generally understood that, of the methods on offer, only random sampling could have been used. There were far fewer correct answers in **part (a)**, especially **part (a)(i)**, where it was rarely appreciated that all three sampling methods could have produced this outcome. In **part (c)** varied understanding of ‘census’ was displayed. Some candidates thought that only a fraction of the schools would be selected, usually half or a third.

Question 2

Almost all candidates obtained fully correct answers, the relevant angles being measured accurately. Occasionally methods were longer than they needed to be, sometimes in **part (a)** the quantity first being found, and in **part (b)** the percentage first being found.

Question 3

Candidates were generally able to deal with the negative data values in **part (a)** properly, though occasionally negative values for Σx^2 were seen in **part (a)(ii)**. Few fully correct answers to **part (b)** were seen.

Question 4

Some candidates continue to have difficulty with probability. **Part (a)** was answered well, but in **part (b)** the second case was frequently omitted. Many attempts at **part (c)** followed the unnecessarily long direct route, considering the cases of one, then two, then three counterfeit coins. Often, in adopting this approach, cases were omitted, and an incorrect answer obtained. Candidates finding the one case not required, and subtracting from 1, the most economical approach, were almost always successful.

Question 5

Almost all candidates were able to form a clear and accurate table in **part (a)**. Responses to **part (b)** were less satisfactory. The main limitation of many answers was lack of numerical support for conclusions using the numerical values that had been entered into the table. Without these many answers amounted to little more than restatement of the original claims. Such answers were not given credit. Candidates who did cite numerical values often drew the wrong conclusion regarding Flo's claim, referring to absolute values only rather than proportions.

Question 6

Very good understanding was shown of how to use the graph to find the statistical measures in **part (a)**, and many fully correct answers were seen. Answers to **part (b)** varied more in quality. A simple and effective approach adopted by some candidates was to find the percentage of readings in excess of 90 decibels and compare with 7.5 per cent. Other candidates showed limited understanding of the situation, and a few made no response.

Question 7

Once more candidates displayed good computational skills in the calculation of crude and standardised death rates, and it was fairly common for all the computational parts to be answered correctly. The main limitation in such cases was that, where the number of deaths had to found, answers were not given as integers.

Of the non-calculation parts, **part (c)** was answered best, but by no means universally well. It is of course important that the candidate knows the reason why a standardised rate is being found, as well as being able to find it. It appears that this is not always the case. Answers to **parts (e)** and **(g)** were much more limited, and tended to reproduce very general observations on crude and standardised rates, which were not always correct and had little relevance to the question. Very few focused on the essential comparisons: in **part (e)** between the structure of the town's population and that of the standard population; and in **part (g)** on the differences between the crude and standardised rates.

Question 8

The first four parts were answered well. In **part (c)** it was almost universally understood that to calculate a semi-average the data must first be ordered by increasing values of the independent variable. Many accurate answers to **part (d)** were seen, obtained using the averages through which the line of best fit must necessarily pass. It has to be pointed out again that choosing points other than the averages from the drawn line to find its equation, whilst correct, does not lead to the most accurate result.

In **part (e)** candidates who realised that 7.5 g had to be used on the graph or in the equation, and not 15 g, were usually successful. These however were in a minority. General difficulty was apparent in **part (f)** in adapting the given information and analysis so far carried out to a can of drink of the given size. One valid approach seen, though needlessly long, was to calculate the energy content, for such a can size, of all eight drinks.

Question 9

As sometimes seems to be the case with histograms, answers in **part (a)** fell into two categories: those that showed clear understanding of the principle on which a histogram is constructed, and those where column heights were simply drawn at the given class frequencies. In **part (b)** little understanding was shown of the assumption needed to be able to estimate the number of cyclists in part of a class.

Performance on the linear Venn diagram in **parts (c)**, **(d)** and **(e)** was very varied. Many candidates able to interpret such a diagram were able to obtain fully correct answers in **parts (c)** and **(d)**, though sometimes the numerical answers to **parts (d)(ii)** and **(d)(iii)** were the wrong way round. Other candidates still demonstrate limited understanding of what the different regions of such a diagram represent. In **part (d)** in particular, complex calculations were sometimes offered, when all that was required was the extraction of relevant numbers from the diagram and the formation of a simple fraction. Attempts at **part (e)** needed to show recognition that only two intersections on the diagram had to be considered in order to obtain credit.

Question 10

Answers to **part (a)** were almost universally correct. In **part (b)**, reasonably good computational skills were seen, though sometimes answers were in error by small amounts as a result of using incorrect mid-points. There was also some premature rounding from the calculation of the mean when finding the standard deviation, which resulted in loss of accuracy. In questions such as this, a table is given with blank space so that candidates can create columns of their own in the table to order their working and keep it tidy. Candidates are advised to make use of such space when it is made available.

On the frequency polygons, good answers were often seen to **parts (c)** and **(d)**. There were common limitations in answers to **part (e)**. Instead of plotting the points at the mid-class values, some candidates plotted at the upper class limits, and occasionally at the lower class limits. Few also labelled the polygons or provided a key. It was especially important to do this so that, when making the comparisons required in **part (f)**, the people were identified correctly. Answers to **part (f)** were mixed, though many comments, provided they related to the lengths of the messages, were acceptable. As it was given that both sets of data were grouped, it was impossible to say who sent the longest or shortest message.

STATISTICS

Paper 4040/13
Paper 1

Key messages

It is as just as important to be able to explain and interpret the values of statistical measures as it is to be able to find them accurately. Statistical calculations have their greatest value when used to understand real-life practical situations.

When drawing conclusions about a practical situation following statistical analysis, use should be made of any numerical values found, and of any additional specific information given.

General comments

Once more the overall standard of work involving calculations of a routine nature was good. This was particularly true of the topics of basic statistical measures, crude and standardised rates, and line of best fit. Answers were generally more limited on the explanation and interpretation parts of questions. This was especially true of parts of the question on one of these topics (see **Questions 7(c)**, **(e)** and **(g)** below). Some answers were often of a very general nature, without specific reference to, and use of, information given in the question.

As has been emphasised previously in these reports, candidates of this subject need to appreciate the essential practical nature of Statistics in its application to a wide variety of real-life situations. It is certainly important for a candidate to be able to perform routine calculations efficiently and accurately; but it is just as important to understand why such calculations are being carried out, what they demonstrate, and how they help in providing useful practical information about the situation being considered.

Comments on specific questions

Question 1

On sampling methods, **part (b**) was answered best. Even though the first three schools in the list had been chosen, it was generally understood that, of the methods on offer, only random sampling could have been used. There were far fewer correct answers in **part (a)**, especially **part (a)(i)**, where it was rarely appreciated that all three sampling methods could have produced this outcome. In **part (c)** varied understanding of ‘census’ was displayed. Some candidates thought that only a fraction of the schools would be selected, usually half or a third.

Question 2

Almost all candidates obtained fully correct answers, the relevant angles being measured accurately. Occasionally methods were longer than they needed to be, sometimes in **part (a)** the quantity first being found, and in **part (b)** the percentage first being found.

Question 3

Candidates were generally able to deal with the negative data values in **part (a)** properly, though occasionally negative values for Σx^2 were seen in **part (a)(ii)**. Few fully correct answers to **part (b)** were seen.

Question 4

Some candidates continue to have difficulty with probability. **Part (a)** was answered well, but in **part (b)** the second case was frequently omitted. Many attempts at **part (c)** followed the unnecessarily long direct route, considering the cases of one, then two, then three counterfeit coins. Often, in adopting this approach, cases were omitted, and an incorrect answer obtained. Candidates finding the one case not required, and subtracting from 1, the most economical approach, were almost always successful.

Question 5

Almost all candidates were able to form a clear and accurate table in **part (a)**. Responses to **part (b)** were less satisfactory. The main limitation of many answers was lack of numerical support for conclusions using the numerical values that had been entered into the table. Without these many answers amounted to little more than restatement of the original claims. Such answers were not given credit. Candidates who did cite numerical values often drew the wrong conclusion regarding Flo's claim, referring to absolute values only rather than proportions.

Question 6

Very good understanding was shown of how to use the graph to find the statistical measures in **part (a)**, and many fully correct answers were seen. Answers to **part (b)** varied more in quality. A simple and effective approach adopted by some candidates was to find the percentage of readings in excess of 90 decibels and compare with 7.5 per cent. Other candidates showed limited understanding of the situation, and a few made no response.

Question 7

Once more candidates displayed good computational skills in the calculation of crude and standardised death rates, and it was fairly common for all the computational parts to be answered correctly. The main limitation in such cases was that, where the number of deaths had to found, answers were not given as integers.

Of the non-calculation parts, **part (c)** was answered best, but by no means universally well. It is of course important that the candidate knows the reason why a standardised rate is being found, as well as being able to find it. It appears that this is not always the case. Answers to **parts (e)** and **(g)** were much more limited, and tended to reproduce very general observations on crude and standardised rates, which were not always correct and had little relevance to the question. Very few focused on the essential comparisons: in **part (e)** between the structure of the town's population and that of the standard population; and in **part (g)** on the differences between the crude and standardised rates.

Question 8

The first four parts were answered well. In **part (c)** it was almost universally understood that to calculate a semi-average the data must first be ordered by increasing values of the independent variable. Many accurate answers to **part (d)** were seen, obtained using the averages through which the line of best fit must necessarily pass. It has to be pointed out again that choosing points other than the averages from the drawn line to find its equation, whilst correct, does not lead to the most accurate result.

In **part (e)** candidates who realised that 7.5 g had to be used on the graph or in the equation, and not 15 g, were usually successful. These however were in a minority. General difficulty was apparent in **part (f)** in adapting the given information and analysis so far carried out to a can of drink of the given size. One valid approach seen, though needlessly long, was to calculate the energy content, for such a can size, of all eight drinks.

Question 9

As sometimes seems to be the case with histograms, answers in **part (a)** fell into two categories: those that showed clear understanding of the principle on which a histogram is constructed, and those where column heights were simply drawn at the given class frequencies. In **part (b)** little understanding was shown of the assumption needed to be able to estimate the number of cyclists in part of a class.

Performance on the linear Venn diagram in **parts (c)**, **(d)** and **(e)** was very varied. Many candidates able to interpret such a diagram were able to obtain fully correct answers in **parts (c)** and **(d)**, though sometimes the numerical answers to **parts (d)(ii)** and **(d)(iii)** were the wrong way round. Other candidates still demonstrate limited understanding of what the different regions of such a diagram represent. In **part (d)** in particular, complex calculations were sometimes offered, when all that was required was the extraction of relevant numbers from the diagram and the formation of a simple fraction. Attempts at **part (e)** needed to show recognition that only two intersections on the diagram had to be considered in order to obtain credit.

Question 10

Answers to **part (a)** were almost universally correct. In **part (b)**, reasonably good computational skills were seen, though sometimes answers were in error by small amounts as a result of using incorrect mid-points. There was also some premature rounding from the calculation of the mean when finding the standard deviation, which resulted in loss of accuracy. In questions such as this, a table is given with blank space so that candidates can create columns of their own in the table to order their working and keep it tidy. Candidates are advised to make use of such space when it is made available.

On the frequency polygons, good answers were often seen to **parts (c)** and **(d)**. There were common limitations in answers to **part (e)**. Instead of plotting the points at the mid-class values, some candidates plotted at the upper class limits, and occasionally at the lower class limits. Few also labelled the polygons or provided a key. It was especially important to do this so that, when making the comparisons required in **part (f)**, the people were identified correctly. Answers to **part (f)** were mixed, though many comments, provided they related to the lengths of the messages, were acceptable. As it was given that both sets of data were grouped, it was impossible to say who sent the longest or shortest message.

STATISTICS

Paper 4040/22
Paper 2

Key message

In this examination candidates need to demonstrate the ability to make sensible decisions about the most appropriate statistical techniques to employ, as well as being able to employ those techniques accurately. They also need to be able to interpret the results of their statistical analysis in the context of the situations presented. Thus the full cycle of statistical enquiry is explored within this examination, from planning and data collection, to presentation and analysis, and finally to the interpretation of that analysis.

General comments

In terms of planning and data collection, most candidates demonstrated that they knew the correct statistical language to describe data in **Question 2a**. They were usually able to select a stratified random sample in **Questions 6a** and **6b**, but when considering whether or not that sample was representative, in **Question 6c**, full details of the necessary calculations were often missing.

Presentation of data was examined with a dual bar chart in **Question 1** and a box-and-whisker diagram in **Question 8d**. Most candidates drew the correct dual bar, and did so accurately, but the labelling was sometimes missing. A scale was provided for the box-and-whisker diagram, and most candidates were able to use it correctly. What proved to be more challenging was making decisions about which statistical techniques to use in **Questions 2b,** and **2c**. In **Question 11b** it was pleasing to see that most candidates were able to communicate well the purposes of finding moving average values.

Questions that involved probability and statistical calculations were usually well presented, but in **Question 4** some candidates provided answers with no justification and in **Question 5** some of the steps in the proof were missing. In **Questions 10b** and **10d**, for example, it was pleasing to see clear working shown so that part-marks could be awarded for solutions that were not fully correct.

Finally, most candidates interpreted the weighted aggregate cost index correctly in **Question 7c**, although some referred to an increase in the expenditure rather than an increase in the prices. In **Question 11f** the interpretation was usually given in context, but some candidates incorrectly stated that the increase in rainfall that they had observed was taking place every quarter rather than over time. Other questions requiring interpretation of data, such as **Question 8b**, proved to be challenging for many candidates.

Comments on specific questions

Question 1

Accurately produced dual bar charts were often seen, with pairs of bars for each type of bird allowing easy comparison of the numbers seen in October 2009 and the numbers seen in October 2019. The most common error was a missing label of ‘number seen’ or ‘frequency’ on the appropriate axis. Weaker responses sometimes did not start the axis representing frequency at zero, resulting in misleading diagrams where bar heights could not be compared meaningfully.

Question 2

Most candidates used correct statistical language in **part (a)** to describe the type of data being collected. The most common error was for shoe size to be described as continuous rather than discrete. Some candidates correctly identified each piece of information being collected as either qualitative or quantitative, but did not describe sufficiently fully by including, for the quantitative data, whether each was continuous or discrete.

Parts (b) and **(c)** proved to be difficult for many candidates, with a wide variety of incorrect statistical diagrams being named in **part (b)**. Examples of incorrect statistical diagrams included box-and-whisker diagrams, stem-and-leaf diagrams and a range of bar charts, which do not allow for an individual's height and shoe size to be illustrated so that a relationship between the paired, bivariate data can be investigated.

Part (c) asked for statistical measures, but many candidates named statistical diagrams or sampling methods instead. Those that gave statistical measures usually named at least one correctly, often the mean, but then gave a measure of spread or indeed two measures of spread. Those that gave correct answers of the mean and the median were usually able to give a correct advantage of one over the other.

Question 3

Many candidates were successful with this question. The most common errors were answers that satisfied two rather than all three of the features, such as 3, 3, 5, 8 and 8, which has the required mean and median but does not have the single mode of 8.

Question 4

This question was less structured than similar questions in the past, and required candidates to decide on an appropriate strategy to allow for meaningful comparison of Onalenna's results. Those that correctly standardised her results to a common mean and standard deviation usually chose a mean of 0 and a standard deviation of 1, making the calculations straightforward. Those that chose other common means and standard deviations usually did so successfully, but were more likely to make arithmetic slips. Some candidates only got as far as finding deviations from the mean, without a meaningful inclusion of the standard deviation in their calculation. Those candidates that did not take the route of performing a calculation could achieve full marks with convincing reasoning involving the distance between each result and the relevant mean, and then describing this difference in terms of the size of the relevant standard deviation. Such descriptions often only went as far as a comparison with the mean and some candidates simply gave two final answers with no attempt at a justification.

Question 5

This question required clear reasoning to be provided and this was often missing, particularly in **part (a)**. A common answer seen that did not provide sufficient reasoning was '$P(A) + P(B) = 1.2$, so A and B are not mutually exclusive'. Using this approach, the result from the addition of the probabilities, 1.2, needed to be stated to be greater than 1 or not equal to $P(A \cup B)$ prior to the conclusion that A and B are not mutually exclusive. An alternative approach was to calculate $P(A \cap B)$ and show that this was not equal to zero. This calculation needed to be done using the formula $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ and not by assuming independence and using $P(A) \times P(B)$, as was often seen.

Many candidates were able to score at least some of the marks in **part (b)**, but clear explanations of the method were sometimes missing. As with **part (a)** there were alternative approaches that could be taken. Those that calculated $P(A \cap B)$ in two different ways needed to equate each to $P(A \cap B)$ prior to reaching a conclusion, and this level of detail in the proof was not always present. Those that assumed independence and calculated $P(A \cup B)$ using the formula $P(A) + P(B) - P(A) \times P(B)$, again needed to provide a clear explanation of the method they were using. Some candidates did not make a final conclusion following correct work and some used incomplete notation such as $(A \cap B)$ rather than $P(A \cap B)$.

Question 6

Most candidates correctly calculated the number of lower school boys that should be in the sample in **part (a)** and then went on to find a correct stratified sample in **part (b)**. Some candidates selected a sample that was stratified by gender only or by age only, and some chose a simple random sample with no stratification.

In **part (c)**, most candidates correctly identified which of the given sample represented lower school boys. What was often missing, however, was a calculation to show the number of lower school boys that there should be in a sample of size eight. Those that made the correct calculation usually realised that by rounding to the nearest integer the sample provided correctly represented the lower school boys (or was as representative as it could be).

Question 7

Fully correct answers were usually seen in **parts (a)** and **(b)**. The interpretation in **part (c)** was sometimes partially correct, with a correct percentage increase often quoted, but with either the dates between which this increase had taken place missing or the increase described as an increase in the expenditure rather than an increase in the prices. Weaker responses sometimes omitted the per cent sign from their answer.

Some good responses were seen in **part (d)**, but many candidates referred incorrectly to a change in the prices or costs, which have already been accounted for in the price relatives. Appropriate answers seen were usually those connected to a change in the number of workers or a change in the amount of materials. Some candidates gave answers that did not refer sufficiently well to the context of the problem, such as 'a new category may be introduced' or 'the consumption may change'.

Question 8

In **part (a)** most candidates correctly found the required range and interquartile range. In **part (a)(iii)** many candidates correctly interpreted a comparison of the medians by stating that the head circumferences of the babies at 3 months old were generally larger than those at birth. The comparison of the interquartile ranges proved more challenging, as very few candidates commented that there was very little difference in the variation of head circumferences between the babies at birth and at 3 months old.

An answer of either 'correct' or 'not correct' was acceptable in **part (b)**, but needed to be accompanied by an appropriate explanation. This proved very challenging for the majority of candidates. Many simply stated that she was correct because each circumference represented that of the lower quartile baby, without referring to the fact that the baby in the lower quartile position may not be the same baby at birth and at 3 months old, and she may therefore be incorrect. Alternatively, a candidate could answer that the distributions are a similar shape and thus the babies' rank order of size may not have changed and she may therefore be correct.

Many fully correct answers were seen to **parts (c)** and **(d)**. In **part (d)** a few candidates made errors in their plots using the given scale: in particular, the plot for the highest value of 46.2 was sometimes seen at 46.4.

Candidates were generally more successful with the probability in **part (e)(i)** than with that in **part (e)(ii)**, with common errors being $\frac{1}{19}$ in **part (i)** and $\frac{3}{19}$ in **part (ii)**.

In **part (f)** it was impressive to see a large number of candidates correctly adjusting the median for the babies at 1 year old and leaving the interquartile range alone.

Question 9

Whilst some candidates scored full marks in this question, for many candidates question 9 proved to be the most difficult on the paper. In **part (a)(i)** some candidates produced probability distributions in which the probabilities did not sum to one. It was quite common to see values for the probabilities that were actually the product of attempted probabilities and prizes. For candidates with incorrect probability distributions in **part (i)**, it was very important that clear working was shown in **part (a)(ii)**, and this was not always present.

In **part (b)(i)** most candidates correctly used the fact that the probabilities sum to one to find *p*. In **part (b)(ii)** many candidates omitted to consider that a prize of $0 and $2 could be obtained in two ways, and thus an incorrect answer of 9/32 was quite common. As with **part (ii)**, many part-marks were seen in **part (b)(iii)**. Many candidates began correctly by multiplying the prizes by their probabilities, but often errors were made thereafter, and those with fully correct working did not always state whether the value they had obtained represented a profit or a loss for the owner.

Question 10

Almost all candidates correctly found the cumulative frequencies in **part (a)**.

In **part (b)** most candidates correctly found 10 per cent of the 120 mangoes and worked out which class interval the largest mango from the smallest 10 per cent of mangoes would lie. A small number of candidates found the lower quartile mass and some found the mass of the smallest mango from the largest 10 per cent of masses. As the solution using linear interpolation is an estimate, 3 significant figures were requested, but some candidates gave an answer with a greater degree of accuracy than was sensible in the context of this question.

Part (c) proved challenging for many, although some fully correct solutions were seen and some solutions in which the only error was to not consider the remaining mangoes, once those for making chutney had been removed. Others correctly divided the classes containing the smallest and largest medium mangoes correctly, but made errors thereafter.

In **part (d)** partially correct solutions were often seen, with the most common error being the omission of multiplying a product of four probabilities by the 4 ways in which three large mangoes and one small mango can be arranged. Also, some candidates omitted to take into account the fact that the mangoes, once selected, would not be replaced.

Question 11

In **part (a)** many candidates gave a description of the trend rather than the seasonal variation. In describing the seasonal variation it was necessary to describe one complete season from the data presented. Some candidates made vague comments about the values varying rather than describing how the rainfall varies over one season in this town.

Responses to **part (b)** were good, with most candidates giving at least one correct purpose. Some candidates gave two purposes that were effectively the same, such as ‘to find the trend’ and ‘to draw a trend line’.

Some weaker candidates continued in **part (c)** to give general reasons for finding moving average values rather than explaining why it was necessary to centre in this case. Most however provided a correct explanation.

Not all candidates who had correctly explained the reason for centring the moving average values went on to centre them in **part (d)**, and some found four correct values, but did not position them correctly in the table. However, it was pleasing to see many fully correct answers to this question, which was not as structured as some similar questions have been in the past.

Values tended to be accurately plotted in **part (e)** with appropriate trend lines drawn, with the most common error being the plots incorrectly positioned horizontally.

Most candidates gave an explanation in context in **part (f)**, but some said that the rainfall was increasing every quarter rather than increasing over time.

Some candidates did not use the fact that the sum of the seasonal components is zero to find the seasonal component in **part (g)**; however, many were able to use their answer to that part together with their line of best fit to make a sensible estimate in **part (h)**.

STATISTICS

Paper 4040/23
Paper 2

Key message

In this examination candidates need to demonstrate the ability to make sensible decisions about the most appropriate statistical techniques to employ, as well as being able to employ those techniques accurately. They also need to be able to interpret the results of their statistical analysis in the context of the situations presented. Thus the full cycle of statistical enquiry is explored within this examination, from planning and data collection, to presentation and analysis, and finally to the interpretation of that analysis.

General comments

In terms of planning and data collection, most candidates demonstrated that they knew the correct statistical language to describe data in **Question 2a**. They were usually able to select a stratified random sample in **Questions 6a** and **6b**, but when considering whether or not that sample was representative, in **Question 6c**, full details of the necessary calculations were often missing.

Presentation of data was examined with a dual bar chart in **Question 1** and a box-and-whisker diagram in **Question 8d**. Most candidates drew the correct dual bar, and did so accurately, but the labelling was sometimes missing. A scale was provided for the box-and-whisker diagram, and most candidates were able to use it correctly. What proved to be more challenging was making decisions about which statistical techniques to use in **Questions 2b,** and **2c**. In **Question 11b** it was pleasing to see that most candidates were able to communicate well the purposes of finding moving average values.

Questions that involved probability and statistical calculations were usually well presented, but in **Question 4** some candidates provided answers with no justification and in **Question 5** some of the steps in the proof were missing. In **Questions 10b** and **10d**, for example, it was pleasing to see clear working shown so that part-marks could be awarded for solutions that were not fully correct.

Finally, most candidates interpreted the weighted aggregate cost index correctly in **Question 7c**, although some referred to an increase in the expenditure rather than an increase in the prices. In **Question 11f** the interpretation was usually given in context, but some candidates incorrectly stated that the increase in rainfall that they had observed was taking place every quarter rather than over time. Other questions requiring interpretation of data, such as **Question 8b**, proved to be challenging for many candidates.

Comments on specific questions

Question 1

Accurately produced dual bar charts were often seen, with pairs of bars for each type of bird allowing easy comparison of the numbers seen in October 2009 and the numbers seen in October 2019. The most common error was a missing label of ‘number seen’ or ‘frequency’ on the appropriate axis. Weaker responses sometimes did not start the axis representing frequency at zero, resulting in misleading diagrams where bar heights could not be compared meaningfully.

Question 2

Most candidates used correct statistical language in **part (a)** to describe the type of data being collected. The most common error was for shoe size to be described as continuous rather than discrete. Some candidates correctly identified each piece of information being collected as either qualitative or quantitative, but did not describe sufficiently fully by including, for the quantitative data, whether each was continuous or discrete.

Parts (b) and **(c)** proved to be difficult for many candidates, with a wide variety of incorrect statistical diagrams being named in **part (b)**. Examples of incorrect statistical diagrams included box-and-whisker diagrams, stem-and-leaf diagrams and a range of bar charts, which do not allow for an individual's height and shoe size to be illustrated so that a relationship between the paired, bivariate data can be investigated.

Part (c) asked for statistical measures, but many candidates named statistical diagrams or sampling methods instead. Those that gave statistical measures usually named at least one correctly, often the mean, but then gave a measure of spread or indeed two measures of spread. Those that gave correct answers of the mean and the median were usually able to give a correct advantage of one over the other.

Question 3

Many candidates were successful with this question. The most common errors were answers that satisfied two rather than all three of the features, such as 3, 3, 5, 8 and 8, which has the required mean and median but does not have the single mode of 8.

Question 4

This question was less structured than similar questions in the past, and required candidates to decide on an appropriate strategy to allow for meaningful comparison of Onalenna's results. Those that correctly standardised her results to a common mean and standard deviation usually chose a mean of 0 and a standard deviation of 1, making the calculations straightforward. Those that chose other common means and standard deviations usually did so successfully, but were more likely to make arithmetic slips. Some candidates only got as far as finding deviations from the mean, without a meaningful inclusion of the standard deviation in their calculation. Those candidates that did not take the route of performing a calculation could achieve full marks with convincing reasoning involving the distance between each result and the relevant mean, and then describing this difference in terms of the size of the relevant standard deviation. Such descriptions often only went as far as a comparison with the mean and some candidates simply gave two final answers with no attempt at a justification.

Question 5

This question required clear reasoning to be provided and this was often missing, particularly in **part (a)**. A common answer seen that did not provide sufficient reasoning was '$P(A) + P(B) = 1.2$, so A and B are not mutually exclusive'. Using this approach, the result from the addition of the probabilities, 1.2, needed to be stated to be greater than 1 or not equal to $P(A \cup B)$ prior to the conclusion that A and B are not mutually exclusive. An alternative approach was to calculate $P(A \cap B)$ and show that this was not equal to zero. This calculation needed to be done using the formula $P(A \cap B) = P(A) + P(B) - P(A \cup B)$ and not by assuming independence and using $P(A) \times P(B)$, as was often seen.

Many candidates were able to score at least some of the marks in **part (b)**, but clear explanations of the method were sometimes missing. As with **part (a)** there were alternative approaches that could be taken. Those that calculated $P(A \cap B)$ in two different ways needed to equate each to $P(A \cap B)$ prior to reaching a conclusion, and this level of detail in the proof was not always present. Those that assumed independence and calculated $P(A \cup B)$ using the formula $P(A) + P(B) - P(A) \times P(B)$, again needed to provide a clear explanation of the method they were using. Some candidates did not make a final conclusion following correct work and some used incomplete notation such as $(A \cap B)$ rather than $P(A \cap B)$.

Question 6

Most candidates correctly calculated the number of lower school boys that should be in the sample in **part (a)** and then went on to find a correct stratified sample in **part (b)**. Some candidates selected a sample that was stratified by gender only or by age only, and some chose a simple random sample with no stratification.

In **part (c)**, most candidates correctly identified which of the given sample represented lower school boys. What was often missing, however, was a calculation to show the number of lower school boys that there should be in a sample of size eight. Those that made the correct calculation usually realised that by rounding to the nearest integer the sample provided correctly represented the lower school boys (or was as representative as it could be).

Question 7

Fully correct answers were usually seen in **parts (a)** and **(b)**. The interpretation in **part (c)** was sometimes partially correct, with a correct percentage increase often quoted, but with either the dates between which this increase had taken place missing or the increase described as an increase in the expenditure rather than an increase in the prices. Weaker responses sometimes omitted the per cent sign from their answer.

Some good responses were seen in **part (d)**, but many candidates referred incorrectly to a change in the prices or costs, which have already been accounted for in the price relatives. Appropriate answers seen were usually those connected to a change in the number of workers or a change in the amount of materials. Some candidates gave answers that did not refer sufficiently well to the context of the problem, such as 'a new category may be introduced' or 'the consumption may change'.

Question 8

In **part (a)** most candidates correctly found the required range and interquartile range. In **part (a)(iii)** many candidates correctly interpreted a comparison of the medians by stating that the head circumferences of the babies at 3 months old were generally larger than those at birth. The comparison of the interquartile ranges proved more challenging, as very few candidates commented that there was very little difference in the variation of head circumferences between the babies at birth and at 3 months old.

An answer of either 'correct' or 'not correct' was acceptable in **part (b)**, but needed to be accompanied by an appropriate explanation. This proved very challenging for the majority of candidates. Many simply stated that she was correct because each circumference represented that of the lower quartile baby, without referring to the fact that the baby in the lower quartile position may not be the same baby at birth and at 3 months old, and she may therefore be incorrect. Alternatively, a candidate could answer that the distributions are a similar shape and thus the babies' rank order of size may not have changed and she may therefore be correct.

Many fully correct answers were seen to **parts (c)** and **(d)**. In **part (d)** a few candidates made errors in their plots using the given scale: in particular, the plot for the highest value of 46.2 was sometimes seen at 46.4.

Candidates were generally more successful with the probability in **part (e)(i)** than with that in **part (e)(ii)**, with common errors being $\frac{1}{19}$ in **part (i)** and $\frac{3}{19}$ in **part (ii)**.

In **part (f)** it was impressive to see a large number of candidates correctly adjusting the median for the babies at 1 year old and leaving the interquartile range alone.

Question 9

Whilst some candidates scored full marks in this question, for many candidates question 9 proved to be the most difficult on the paper. In **part (a)(i)** some candidates produced probability distributions in which the probabilities did not sum to one. It was quite common to see values for the probabilities that were actually the product of attempted probabilities and prizes. For candidates with incorrect probability distributions in **part (i)**, it was very important that clear working was shown in **part (a)(ii)**, and this was not always present.

In **part (b)(i)** most candidates correctly used the fact that the probabilities sum to one to find *p*. In **part (b)(ii)** many candidates omitted to consider that a prize of \$0 and \$2 could be obtained in two ways, and thus an incorrect answer of 9/32 was quite common. As with **part (ii)**, many part-marks were seen in **part (b)(iii)**. Many candidates began correctly by multiplying the prizes by their probabilities, but often errors were made thereafter, and those with fully correct working did not always state whether the value they had obtained represented a profit or a loss for the owner.

Question 10

Almost all candidates correctly found the cumulative frequencies in **part (a)**.

In **part (b)** most candidates correctly found 10 per cent of the 120 mangoes and worked out which class interval the largest mango from the smallest 10 per cent of mangoes would lie. A small number of candidates found the lower quartile mass and some found the mass of the smallest mango from the largest 10 per cent of masses. As the solution using linear interpolation is an estimate, 3 significant figures were requested, but some candidates gave an answer with a greater degree of accuracy than was sensible in the context of this question.

Part (c) proved challenging for many, although some fully correct solutions were seen and some solutions in which the only error was to not consider the remaining mangoes, once those for making chutney had been removed. Others correctly divided the classes containing the smallest and largest medium mangoes correctly, but made errors thereafter.

In **part (d)** partially correct solutions were often seen, with the most common error being the omission of multiplying a product of four probabilities by the 4 ways in which three large mangoes and one small mango can be arranged. Also, some candidates omitted to take into account the fact that the mangoes, once selected, would not be replaced.

Question 11

In **part (a)** many candidates gave a description of the trend rather than the seasonal variation. In describing the seasonal variation it was necessary to describe one complete season from the data presented. Some candidates made vague comments about the values varying rather than describing how the rainfall varies over one season in this town.

Responses to **part (b)** were good, with most candidates giving at least one correct purpose. Some candidates gave two purposes that were effectively the same, such as ‘to find the trend’ and ‘to draw a trend line’.

Some weaker candidates continued in **part (c)** to give general reasons for finding moving average values rather than explaining why it was necessary to centre in this case. Most however provided a correct explanation.

Not all candidates who had correctly explained the reason for centring the moving average values went on to centre them in **part (d)**, and some found four correct values, but did not position them correctly in the table. However, it was pleasing to see many fully correct answers to this question, which was not as structured as some similar questions have been in the past.

Values tended to be accurately plotted in **part (e)** with appropriate trend lines drawn, with the most common error being the plots incorrectly positioned horizontally.

Most candidates gave an explanation in context in **part (f)**, but some said that the rainfall was increasing every quarter rather than increasing over time.

Some candidates did not use the fact that the sum of the seasonal components is zero to find the seasonal component in **part (g)**; however, many were able to use their answer to that part together with their line of best fit to make a sensible estimate in **part (h)**.