STATISTICS

Paper 4040/11

Paper 11

Key Messages

When a question specifies a degree of accuracy for numerical answers, the instruction must be followed for full marks to be obtained.

Where a comment is required following calculation, the comment must relate to the context of the question.

Candidates should have a rough idea of the magnitude of the answer to be expected from a calculation to see that their answer is reasonable.

General comments

The overall standard of work was much the same as last year. Some high marks were obtained, and there were few exceptionally low marks. As is noted regularly in these reports there were again many instances of marks being lost as a result of final answers not being given to the required accuracy in cases where this was stated in the question. It is not clear whether this is a consequence of candidates ignoring the instruction in the question, or not understanding clearly the difference between significant figures and decimal places. Some candidates also continue to struggle in trying to answer those parts of questions which require comment related to the significance of calculated results for the situation in the question. To earn marks in these cases it is essential that answers relate specifically to the practical situation on which the question is based, and not be expressed in abstract or purely mathematical terms.

It ought to be essential practice for a candidate of statistics to observe whether or not the result of a calculation is reasonable. If it can be seen that it is not, then work can be checked to find the error. If the mean age of a group of people is required (see **Question 8** below), it should be obvious that an answer of the order of 125 cannot be correct. There were several instances in the paper where impossible values for the given situation were presented. It would be extremely valuable if candidates could develop the skill of being able to predict an approximate answer to a problem, even before starting the calculation.

Almost all candidates have adapted well to working with the new format for the question paper, now in its third year. Careful thought is given in the preparation of the paper to ensure that plenty of space is provided for giving good answers. When, for example, two full lines are given for the comment part of a question, then that is all the space which is needed to make a good answer. If candidates find themselves writing much more, or wanting to write much more, they should pause to consider if they really have the essential point the question is asking about.

Comments on specific questions

Section A

Question 1

The aim of part **(i)** was to test elementary knowledge of the positions in a distribution of the median and the quantiles. Many candidates obtained the four answers correctly. Others seemed to think that the values of the measures had to be incorporated into calculations somehow in finding answers, rather than recognising, for example, that the median had been identified as 12.3 cm, so that **(a)** was simply asking about the percentage of items in a large distribution which are less than the median. It was in part **(ii)** where three of the values needed to be actually used. These indicate that the median is almost exactly mid-way between the lower quartile and the upper quartile, since (7.9 + 16.5)/2 = 12.2, indicating that the distribution is almost

perfectly symmetrical. Few candidates recognised that it was the relationship between these three quantities that was the key. The answers of those who did would have been improved by using numerical support.

Answers: **(i)(a)** 50%, **(b)** 75%, **(c)** 15%, **(d)** 35%; **(ii)** about the same.

Question 2

In part **(i)** candidates were asked to give just one word for their answers, so were being asked to consider which of the words they had learned about methods of sampling were the most appropriate to the situations described. There were many correct answers to **(a)** and **(b)**, but a fairly common incorrect answer to **(c)** was "random". There were many fully correct answers to part **(ii)**, though where there was confusion it tended to be between the terms "quantitative" and "qualitative" rather than between "discrete" and "continuous".

Answers: **(i)(a)** systematic, **(b)** stratified, **(c)** quota; **(ii)** quantitative and discrete, qualitative and discrete, quantitative and continuous.

Question 3

There were few fully correct answers to this question. Correct answers were most often seen to parts **(i)** and **(iv)**. A fairly common incorrect answer to part **(v)** was 28, which ignored the other 14 people who speak one or both of these languages. Candidates would improve their performance on this type of question if they understood clearly the precise meaning of the numbers in the different parts of a Venn diagram.

Answers: **(i)** 30; **(ii)** 2; **(iii)** 5; **(iv)** 16; **(v)** 42.

Question 4

The method of using an assumed mean is intended to simplify the calculation of these measures, and candidates who used the method obtained the required values quickly and easily. Those who ignored the instruction inevitably had to work with large numbers, and could be given little credit, even though they obtained correct answers. Sometimes responses started off well, using the assumed mean method to find the mean. The power of the method was then wasted by working on the original values to find the standard deviation.

In part **(ii)**, it seemed to be quite well understood that a constant can be added to a set of data without changing the standard deviation. Credit was given even when answers in part **(i)** were incorrect if answers to part **(ii)** followed properly from the previous incorrect answers. Some answers were seen which indicated that the question had not been read properly, because these answers gave the mean and standard deviation of the empty cans.

Answers: **(i)** 450.5, 1.97; **(ii)** 630.5, 1.97.

Question 5

Few candidates were able to earn more than the mark for part **(i)** in this question. The difficulty seems to have been a lack of proper appreciation of the sequence of events in the game. In part **(ii)** a common incorrect answer was 1/8, following from 1/6 × 3/4. This assumed that Carlo would be drawing a disc from the bag, but he would not have been able to do this had he not thrown a 6 on the dice to begin with. There is thus also the possibility from part **(i)** to be added. Similarly in part **(iii)** most candidates did not consider the full sequence of events. Only a few realised that there must be three non-winning turns, not one, before Dean's second turn.

Answers: **(i)** 5/6; **(ii)** 23/24; **(iii)** 12167/331776 or 0.0367.

Question 6

This question was very well done, with many candidates showing their ability to interpret the textual information with ease. Where errors were made it was usually because the 8 and 4, and the 3 and 7 were interchanged. A few candidates also took the women to win three times as many silver medals as bronze medals.

Answers: **(i)** row totals 11, 11; **(ii)** column totals 12, 13; **(iii)** medals for men 2, 3, 7, medals for women 1, 8, 4; **(iv)** men 21 points, women 24 points, women.

Section B

Question 7

The probability question was answered better than last year. Good answers showed clear understanding of the distinction between rooms and people, so that correct denominators were used for the probability fractions. Many obtained good marks in parts **(i)**, **(ii)** and **(iii)**. But very few were able to produce a fully correct solution to part **(iv)**. If these guests are to be staying in the same room they must be chosen from one of the rooms which contain either 2, 3 or 4 guests. The probability of choosing a guest from one of the rooms with two guests is 24/47, but the probability of choosing the second guest from the same room is only 1/46, not 23/46 as many assumed. No doubt many candidates automatically decreased both numerator and denominator like this because it is often done in other problems. More generally it was good to see that very few offered probabilities greater than 1 this year.

Answers: **(i)** 47; **(ii)(a)** 3/25, **(b)** 24/25; **(iii)(a)** 15/47, **(b)** 19/47; **(iv)** 33/1081 or 0.0305.

Question 8

Candidates who knew their formulae, and worked with correct mid-class values, obtained good marks in part **(i)**. Whilst no particular method was specified in the question, candidates should be advised that the method for standard deviation which uses Σfx and Σfx^2 is generally better for computational purposes than that which uses $\Sigma f(x - \text{mean})^2$. The latter needs an accurate value of the mean to be used for accuracy in the final answer, and this necessarily involves working with (often long) decimal numbers. As was mentioned in general comments above, this was one of those questions where, in some cases, candidates should have been aware that the answer they presented could not possibly be correct. Apart from it being unrealistic in this particular situation for the mean age of these people to be 15 or 125, a candidate should know that any set of data ranging from 20 to 80 must have a mean somewhere between these values. Also it was one of those questions where marks were lost because the 3 significant figure instruction was not followed.

In part **(ii)**, answers were mixed, with good candidates showing clear understanding that class widths had to be taken into consideration when finding the heights of the rectangles to be drawn. Others simply drew the rectangles of heights equal to the frequencies, even though it should have been realised that this must be incorrect as one of the heights then went beyond the given grid. In part **(iii)** credit was still given for a frequency polygon formed of straight lines connecting the mid-points of the tops of the rectangles drawn in part **(ii)**, even if the histogram was incorrect.

It was not enough in part **(iv)** to say simply that the mean had decreased and the standard deviation had increased. The values found had to be be interpreted in the context of the situation. Good answers stated that a greater proportion of younger people (or a smaller proportion of older people) booked holidays than previously, and that the ages of people booking holidays were more varied than previously. It is not possible to conclude that the number of bookings had changed as some seemed to think.

Answers: **(i)** 43.6, 13.8; **(ii)** heights of rectangles to be drawn 7, 8, 14, 7, 2.5.

Question 9

This question was generally well answered, and many earned 13 of the 16 marks available. As has been pointed out before in these reports, good answers to this type of question give some indication on the graph (for example with lines drawn and labelled) of how the required information is being found. Credit can then be given for method, even if the answer is incorrect. An incorrect answer with no indication of how it has been obtained cannot be awarded marks.

Some candidates made the error of using mid-class values instead of upper class boundaries when drawing the cumulative frequency curve in part **(iv)**. But it was in part **(v)** where marks were usually lost. A well drawn curve leads to the conclusion that, after training, there were 13 salespeople who spent more than 12.5 hours travelling (values of 12 or 14 salespeople were also accepted for graphs drawn slightly differently). To find the required percentage reduction the base for the percentage calculation needs to be the number who originally spent more than 12.5 hours travelling (the candidate's answer to part **(ii)**), not 40 as many candidates assumed.

Answers: **(i)(a)** 11.4 – 11.7, **(b)** 7.3 – 7.9, **(c)** 16.8 – 17.5; **(ii)** 17; **(iii)** 7, 20, 33, 38, 40; **(v)** 18% or 24% or 29%.

Question 10

A good proportion of candidates answered the first three parts well, with accurately plotted points and accurately calculated averages, leading to a good line of best fit. For others the fact that *y* decreased as *x* increased resulted in a common error, it being assumed that the smallest values of *x* always had to be paired with the smallest values of *y*. This error meant that the location of the plotted averages on the grid, and the line subsequently drawn through them, bore no relationship whatsoever to the pattern of the plotted data. The line had a positive gradient when clearly the trend of the data indicated the gradient should be negative. Here a little reflection on the part of the candidate would have indicated that something was wrong. There were a few instances, in good answers, of candidates making the error, understanding it was an error, and then correcting it.

In part **(iv)** Examiners checked the value following from the candidate's equation in **(iii)**. Provided the answer given followed accurately, and the gradient of the line was negative, the marks were awarded. Here again there were many instances of unrealistic answers, it being very common to see the price of the car stated as less than $10. Candidates need to look at how a scale is labelled as well as the numbers on the scale.

In parts **(v)** and **(vi)**, to gain credit, it was necessary to reference answers to this practical situation, rather than just use purely mathematical language. In part **(v)** good responses said that this was the price of a brand new car, whereas weak responses said only that it was the intercept. Similarly in part **(vi)** good responses went beyond saying that the gradient should be negative, pointing out that the price of the car should decrease with time, not increase as the equation indicates.

Answers: **(i)** (5.5 or 5.6, 8.2), (3.1, 11.2), (8.0, 5.2); **(iii)** line through at least two of points plotted in **(ii)**, m = -1.30 to -1.20, c = 14.8 to 15.3.

Question 11

Answers to the first three parts showed good general understanding of the calculation of these rates. Marks only tended to be lost in parts **(i)** and **(iii)** through not following the instruction on two decimal places. The correct rate was almost always chosen for the calculations in part **(iv)**. Good understanding was also shown in part **(v)** that Riva has the healthiest environment because its standardised death rate is lower than that of the other two towns.

In part **(iv)** it is found that Riva experienced the largest number of deaths, and in part **(v)** that Riva has the healthiest environment. The task in part **(vi)** is to explain this apparent contradiction. Some answers were needlessly long, and seemed to be bringing in ideas, not always relevant, from other questions worked in the past. Good answers were very brief. There could be more elderly people living in Riva (though speculative, this was awarded the mark). The best answer, because it uses evidence available in the question, is that Riva has the largest population. It is only to be expected that the largest population will experience the largest number of deaths.

Answers: **(i)** 8.90; **(ii)** 2.5, 4.4, 11.25, 36; **(iii)** 9.905 or 9.91; **(iv)** Riva 160, Techno 144, Riva.

STATISTICS

Paper 4040/12

Paper 12

Key Messages

When a question specifies a degree of accuracy for numerical answers, the instruction must be followed for full marks to be obtained.

Where a comment is required following calculation, the comment must relate to the context of the question.

Candidates should have a rough idea of the magnitude of the answer to be expected from a calculation to see that their answer is reasonable.

General comments

The overall standard of work was much the same as last year. Some high marks were obtained, and there were few exceptionally low marks. As is noted regularly in these reports there were again many instances of marks being lost as a result of final answers not being given to the required accuracy in cases where this was stated in the question. It is not clear whether this is a consequence of candidates ignoring the instruction in the question, or not understanding clearly the difference between significant figures and decimal places. Some candidates also continue to struggle in trying to answer those parts of questions which require comment related to the significance of calculated results for the situation in the question. To earn marks in these cases it is essential that answers relate specifically to the practical situation on which the question is based, and not be expressed in abstract or purely mathematical terms.

It ought to be essential practice for a candidate of statistics to observe whether or not the result of a calculation is reasonable. If it can be seen that it is not, then work can be checked to find the error. If the mean age of a group of people is required (see **Question 8** below), it should be obvious that an answer of the order of 125 cannot be correct. There were several instances in the paper where impossible values for the given situation were presented. It would be extremely valuable if candidates could develop the skill of being able to predict an approximate answer to a problem, even before starting the calculation.

Almost all candidates have adapted well to working with the new format for the question paper, now in its third year. Careful thought is given in the preparation of the paper to ensure that plenty of space is provided for giving good answers. When, for example, two full lines are given for the comment part of a question, then that is all the space which is needed to make a good answer. If candidates find themselves writing much more, or wanting to write much more, they should pause to consider if they really have the essential point the question is asking about.

Comments on specific questions

Section A

Question 1

The aim of part **(i)** was to test elementary knowledge of the positions in a distribution of the median and the quantiles. Many candidates obtained the four answers correctly. Others seemed to think that the values of the measures had to be incorporated into calculations somehow in finding answers, rather than recognising, for example, that the median had been identified as 12.3 cm, so that **(a)** was simply asking about the percentage of items in a large distribution which are less than the median. It was in part **(ii)** where three of the values needed to be actually used. These indicate that the median is almost exactly mid-way between the lower quartile and the upper quartile, since (7.9 + 16.5)/2 = 12.2, indicating that the distribution is almost

perfectly symmetrical. Few candidates recognised that it was the relationship between these three quantities that was the key. The answers of those who did would have been improved by using numerical support.

Answers: **(i)(a)** 50%, **(b)** 75%, **(c)** 15%, **(d)** 35%; **(ii)** about the same.

Question 2

In part **(i)** candidates were asked to give just one word for their answers, so were being asked to consider which of the words they had learned about methods of sampling were the most appropriate to the situations described. There were many correct answers to **(a)** and **(b)**, but a fairly common incorrect answer to **(c)** was "random". There were many fully correct answers to part **(ii)**, though where there was confusion it tended to be between the terms "quantitative" and "qualitative" rather than between "discrete" and "continuous".

Answers: **(i)(a)** systematic, **(b)** stratified, **(c)** quota; **(ii)** quantitative and discrete, qualitative and discrete, quantitative and continuous.

Question 3

There were few fully correct answers to this question. Correct answers were most often seen to parts **(i)** and **(iv)**. A fairly common incorrect answer to part **(v)** was 28, which ignored the other 14 people who speak one or both of these languages. Candidates would improve their performance on this type of question if they understood clearly the precise meaning of the numbers in the different parts of a Venn diagram.

Answers: **(i)** 30; **(ii)** 2; **(iii)** 5; **(iv)** 16; **(v)** 42.

Question 4

The method of using an assumed mean is intended to simplify the calculation of these measures, and candidates who used the method obtained the required values quickly and easily. Those who ignored the instruction inevitably had to work with large numbers, and could be given little credit, even though they obtained correct answers. Sometimes responses started off well, using the assumed mean method to find the mean. The power of the method was then wasted by working on the original values to find the standard deviation.

In part **(ii)**, it seemed to be quite well understood that a constant can be added to a set of data without changing the standard deviation. Credit was given even when answers in part **(i)** were incorrect if answers to part **(ii)** followed properly from the previous incorrect answers. Some answers were seen which indicated that the question had not been read properly, because these answers gave the mean and standard deviation of the empty cans.

Answers: **(i)** 450.5, 1.97; **(ii)** 630.5, 1.97.

Question 5

Few candidates were able to earn more than the mark for part **(i)** in this question. The difficulty seems to have been a lack of proper appreciation of the sequence of events in the game. In part **(ii)** a common incorrect answer was 1/8, following from 1/6 × 3/4. This assumed that Carlo would be drawing a disc from the bag, but he would not have been able to do this had he not thrown a 6 on the dice to begin with. There is thus also the possibility from part **(i)** to be added. Similarly in part **(iii)** most candidates did not consider the full sequence of events. Only a few realised that there must be three non-winning turns, not one, before Dean's second turn.

Answers: **(i)** 5/6; **(ii)** 23/24; **(iii)** 12167/331776 or 0.0367.

Question 6

This question was very well done, with many candidates showing their ability to interpret the textual information with ease. Where errors were made it was usually because the 8 and 4, and the 3 and 7 were interchanged. A few candidates also took the women to win three times as many silver medals as bronze medals.

Answers: **(i)** row totals 11, 11; **(ii)** column totals 12, 13; **(iii)** medals for men 2, 3, 7, medals for women 1, 8, 4; **(iv)** men 21 points, women 24 points, women.

Section B

Question 7

The probability question was answered better than last year. Good answers showed clear understanding of the distinction between rooms and people, so that correct denominators were used for the probability fractions. Many obtained good marks in parts **(i)**, **(ii)** and **(iii)**. But very few were able to produce a fully correct solution to part **(iv)**. If these guests are to be staying in the same room they must be chosen from one of the rooms which contain either 2, 3 or 4 guests. The probability of choosing a guest from one of the rooms with two guests is 24/47, but the probability of choosing the second guest from the same room is only 1/46, not 23/46 as many assumed. No doubt many candidates automatically decreased both numerator and denominator like this because it is often done in other problems. More generally it was good to see that very few offered probabilities greater than 1 this year.

Answers: **(i)** 47; **(ii)(a)** 3/25, **(b)** 24/25; **(iii)(a)** 15/47, **(b)** 19/47; **(iv)** 33/1081 or 0.0305.

Question 8

Candidates who knew their formulae, and worked with correct mid-class values, obtained good marks in part **(i)**. Whilst no particular method was specified in the question, candidates should be advised that the method for standard deviation which uses Σfx and Σfx^2 is generally better for computational purposes than that which uses $\Sigma f(x - \text{mean})^2$. The latter needs an accurate value of the mean to be used for accuracy in the final answer, and this necessarily involves working with (often long) decimal numbers. As was mentioned in general comments above, this was one of those questions where, in some cases, candidates should have been aware that the answer they presented could not possibly be correct. Apart from it being unrealistic in this particular situation for the mean age of these people to be 15 or 125, a candidate should know that any set of data ranging from 20 to 80 must have a mean somewhere between these values. Also it was one of those questions where marks were lost because the 3 significant figure instruction was not followed.

In part **(ii)**, answers were mixed, with good candidates showing clear understanding that class widths had to be taken into consideration when finding the heights of the rectangles to be drawn. Others simply drew the rectangles of heights equal to the frequencies, even though it should have been realised that this must be incorrect as one of the heights then went beyond the given grid. In part **(iii)** credit was still given for a frequency polygon formed of straight lines connecting the mid-points of the tops of the rectangles drawn in part **(ii)**, even if the histogram was incorrect.

It was not enough in part **(iv)** to say simply that the mean had decreased and the standard deviation had increased. The values found had to be be interpreted in the context of the situation. Good answers stated that a greater proportion of younger people (or a smaller proportion of older people) booked holidays than previously, and that the ages of people booking holidays were more varied than previously. It is not possible to conclude that the number of bookings had changed as some seemed to think.

Answers: **(i)** 43.6, 13.8; **(ii)** heights of rectangles to be drawn 7, 8, 14, 7, 2.5.

Question 9

This question was generally well answered, and many earned 13 of the 16 marks available. As has been pointed out before in these reports, good answers to this type of question give some indication on the graph (for example with lines drawn and labelled) of how the required information is being found. Credit can then be given for method, even if the answer is incorrect. An incorrect answer with no indication of how it has been obtained cannot be awarded marks.

Some candidates made the error of using mid-class values instead of upper class boundaries when drawing the cumulative frequency curve in part **(iv)**. But it was in part **(v)** where marks were usually lost. A well drawn curve leads to the conclusion that, after training, there were 13 salespeople who spent more than 12.5 hours travelling (values of 12 or 14 salespeople were also accepted for graphs drawn slightly differently). To find the required percentage reduction the base for the percentage calculation needs to be the number who originally spent more than 12.5 hours travelling (the candidate's answer to part **(ii)**), not 40 as many candidates assumed.

Answers: **(i)(a)** 11.4 – 11.7, **(b)** 7.3 – 7.9, **(c)** 16.8 – 17.5; **(ii)** 17; **(iii)** 7, 20, 33, 38, 40; **(v)** 18% or 24% or 29%.

Question 10

A good proportion of candidates answered the first three parts well, with accurately plotted points and accurately calculated averages, leading to a good line of best fit. For others the fact that *y* decreased as *x* increased resulted in a common error, it being assumed that the smallest values of *x* always had to be paired with the smallest values of *y*. This error meant that the location of the plotted averages on the grid, and the line subsequently drawn through them, bore no relationship whatsoever to the pattern of the plotted data. The line had a positive gradient when clearly the trend of the data indicated the gradient should be negative. Here a little reflection on the part of the candidate would have indicated that something was wrong. There were a few instances, in good answers, of candidates making the error, understanding it was an error, and then correcting it.

In part **(iv)** Examiners checked the value following from the candidate's equation in **(iii)**. Provided the answer given followed accurately, and the gradient of the line was negative, the marks were awarded. Here again there were many instances of unrealistic answers, it being very common to see the price of the car stated as less than $10. Candidates need to look at how a scale is labelled as well as the numbers on the scale.

In parts **(v)** and **(vi)**, to gain credit, it was necessary to reference answers to this practical situation, rather than just use purely mathematical language. In part **(v)** good responses said that this was the price of a brand new car, whereas weak responses said only that it was the intercept. Similarly in part **(vi)** good responses went beyond saying that the gradient should be negative, pointing out that the price of the car should decrease with time, not increase as the equation indicates.

Answers: **(i)** (5.5 or 5.6, 8.2), (3.1, 11.2), (8.0, 5.2); **(iii)** line through at least two of points plotted in **(ii)**, m = -1.30 to -1.20, c = 14.8 to 15.3.

Question 11

Answers to the first three parts showed good general understanding of the calculation of these rates. Marks only tended to be lost in parts **(i)** and **(iii)** through not following the instruction on two decimal places. The correct rate was almost always chosen for the calculations in part **(iv)**. Good understanding was also shown in part **(v)** that Riva has the healthiest environment because its standardised death rate is lower than that of the other two towns.

In part **(iv)** it is found that Riva experienced the largest number of deaths, and in part **(v)** that Riva has the healthiest environment. The task in part **(vi)** is to explain this apparent contradiction. Some answers were needlessly long, and seemed to be bringing in ideas, not always relevant, from other questions worked in the past. Good answers were very brief. There could be more elderly people living in Riva (though speculative, this was awarded the mark). The best answer, because it uses evidence available in the question, is that Riva has the largest population. It is only to be expected that the largest population will experience the largest number of deaths.

Answers: **(i)** 8.90; **(ii)** 2.5, 4.4, 11.25, 36; **(iii)** 9.905 or 9.91; **(iv)** Riva 160, Techno 144, Riva.

STATISTICS

Paper 4040/13

Paper 13

Key messages

- For questions requiring statistical or interpretive comments to be made, candidates must ensure that their comments are not overly general but relate specifically to the context of the question.
- Candidates should ensure that questions are read carefully to avoid misinterpretation of what is being asked for.

General comments

There was an improvement this year in final answers being stated to the required level of accuracy. Calculations were, as usual, the customary source of the majority of marks for many candidates, and a number of standard procedures have clearly been learned very thoroughly. However, the problems caused by the calculation of the standard deviation of a grouped frequency distribution remain a cause for concern. Although not exactly simple, this is both a basic and important calculation and needs to be mastered.

Candidates should ensure that plotted points are marked by a clearly visible 'x', and not by a faint dot. Also, candidates should avoid overwriting rough pencil working with their final answers; the pencil work should be rubbed out or the final answers written separately. Both of these aspects have in some cases made it difficult for Examiners to award marks due to a lack of clarity.

Comments on specific questions

Section A

Question 1

Most candidates scored very well on this question, but a few were clearly totally unaware of the most basic of statistical measures, the range being seen given as a measure of average and the mean as a measure of dispersion.

Answers: **(i)** Median 6, mode 5; **(ii)** inter-quartile range 3.

Question 2

For a question to which the answers were totally comments, this was answered relatively well. In parts **(a)** and **(b)** many candidates were able to present a sufficiently suitable comment to earn at least one of the two marks available. What was particularly good, however, was the number of candidates who, in part **(c)**, not only knew the difference between bias and error, but were also able to express the difference relatively clearly in words.

Question 3

A straightforward question on the use of pictograms, which was answered very well by the vast majority of candidates.

Answers: **(i)** 20; **(ii)** 20; **(iii)** 7; **(iv)** 190/2701.

Question 4

This question was not answered anywhere near as well as questions on table completion have been in previous years, it being rare to find an answer scoring full marks. There did not appear to be any one particular part of the question which was the cause of the majority of problems.

Answers: Completed table

13	8	5	3
6	4	2	0
3	1	0	0
2	0	0	0

Question 5

Most candidates coped perfectly well with the information about the different categories being given as fractions of an unknown total, rather than as frequencies. Pie charts presented were generally well-drawn and accurate. Most candidates also knew the correct method to apply to answer part **(ii)**.

Answers: **(i)** Angles (in degrees) 120, 90, 75, 75; **(ii)** 6.12 cm.

Question 6

Most candidates appeared to be fully aware of the principle of area being proportional to frequency and amongst such candidates the main cause of loss of marks was simple arithmetic error. What was very good, however, was the number of candidates who interpreted the expression 'at most 35' correctly. In previous years such expressions have often caused problems. There are still a few candidates using height as proportional to frequency, and as a result they could score only a maximum possible mark of 1 for this question (the mark for interpreting 'at most 35' correctly).

Answers: **(i)** 45; **(ii)** 122; **(iii)** Rectangle of height 6.

Section B

Question 7

The different parts of this question tested different aspects of probability, and it was not unusual to find scripts in which part of the question had been answered correctly but another part not at all well. The most common error in part **(a)** was failure to realise that if the game were to reach the stage of a competitor having a second turn, then both must have failed to win on their first turn, and for probabilities to take this into account. Part **(b)** was probably the best-answered of the three parts, many candidates handling correctly the difference between the with- and without-replacement situations, and that between the same-colour and different-colour situations. Many errors in part **(c)** stemmed from failure to read the definition of event *B* sufficiently carefully.

Answers: **(a)(i)** 11/72, **(ii)** 25/216; **(b)(i)** 25/144, **(ii)** 35/72, **(iii)** 5/33, **(iv)** 35/66; **(c)(ii)** 0.22, 0.9, **(iii)** 45/57.

Question 8

Three of the five parts of this question were answered very well, the other two extremely poorly by almost all candidates. Parts **(i)** and **(iii)** required very standard use and interpretation of a cumulative frequency curve and part **(iv)** a calculation for which the question clearly described the method. Most candidates scored well on these parts. In part **(ii)** very few candidates seemed able to interpret correctly what was meant by 'the central 40% of throws', that is those between the cumulative 30% and 70% values, and hence were unable to attempt an answer. In part **(v)** very few apprehended the comment being looked for: that the curves did illustrate the expected situation, and that males would throw further than females.

Answers: **(i)** 19 to 20; **(ii)** 23 m, 30.5 m; **(iii)** 16 to 17, 19 to 20, 15 to 16; **(iv)** 24.2%.

Question 9

There were very few completely correct answers for part **(a)**, and from the reasons given there was a clear indication that the cause of this was that the question had not been read sufficiently carefully by many candidates. The question referred to marriage rates, whereas most answers were clearly discussing either the number or percentage of the population in each age group who were actually married. Part **(b)** of the question followed the pattern of many years of answers related to death rates. Calculations were frequently totally correct. Over the years an increasing percentage of candidates have answered correctly that the 'relative chance of survival' is indicated by a standardised death rate, but the mark scheme for part **(b)(iv)** was designed so that marks could only be scored if a candidate could explain why this is so. As a result, hardly any marks were scored. Answers which came anywhere near to scoring the available mark in part **(b)(v)** were extremely rare.

Answers: **(a)(i)** 0, **(ii)** 12, **(iii)** 5; **(b)(i)** P = 15, Q = 20, R = 1500, **(ii)** 13.57 per thousand, **(iii)** 14.06 per thousand, **(iv)** Town B.

Question 10

Many parts of this question required solutions of the standard line-of-best-fit type and were generally well answered, but others, which were slightly more unusual, were not. Graphical work was usually good. Hardly any candidates scored the available mark in part **(iii)** because, although they referred to the values needing to be split according to their ascending order, they did not mention explicitly that this should be done in terms of the *X*-variable. Then, in part **(iv)**, although the question had previously indicated how the points should be split for calculation of the semi-averages, some candidates ignored this information completely and used the first four and last four pairs of values in the table. In obtaining the equation of the line of best fit, a few candidates based their calculations on pairs of values in the original table, rather than on the averages they had calculated. Some candidates, both those who had obtained the correct values of *m* and *c* and those who had not, missed out on the final mark in part **(vi)** through not giving the equation in the required form. Candidates should note that, whatever their values of *m* and *c* may be, if they use them in the required form of the equation the mark available for the equation will score. In part **(vii)(a)** some simply listed those points which lay exactly on the line, rather than commenting on how well the line fitted the points as a whole. Some of the presented answers to part **(vii)(b)** gave the clear impression that candidates did not know what was meant by an 'independent variable', and the question in part **(vii)(c)** was therefore correspondingly meaningless to some.

Answers: **(ii)** (8.1875, 53); **(iv)** (5.5, 41.5), (10.875, 64.5); **(vi)** *m* in range 4.1 to 4.4, *c* in range 17 to 19; **(vii)(c)** E F G H.

Question 11

The first three parts and part **(v)** of this question were generally answered very well indeed. It was particularly good that, in part **(ii)**, there was much correct work in relating stated class limits to class mid-points. In part **(iv)** very few indeed realised that the comment being looked for referred to one calculation using actual values but the other only class mid-points. In part **(v)** most deduced correctly that the 'true value' referred to was the result they had obtained in part **(i)**. It continues to be a cause for concern that so few candidates are able to estimate the standard deviation of a grouped frequency distribution correctly. The work on this topic in part **(vi)** was possibly slightly better than in recent years, but still left a lot to be desired.

Answers: **(i)** 37.14; **(ii)** First column 32 33 41 42, Second column 27 37 38 47, Third column 2 10 8 4; **(iii)** 37.06; **(v)** 0.215%; **(vi)** 3.56.

STATISTICS

Paper 4040/21

Paper 21

Key message

The best scripts typically included work which was well structured and demonstrated clear statistical thinking. Candidates scoring the highest marks showed clear methods on the numerical questions and precision and detail on those questions which required definitions and explanations.

General comments

In the best scripts candidates set out clear solutions showing evidence of the methods they had used at each stage of the problem. This is particularly important in multistage problems such as those found in **Questions 7(iii), (iv), (v), 11(v) and (vi)**. In these questions it was possible to award several marks for method, if it was sufficiently clear, even when numerical slips or errors had been made at an earlier stage in the problem.

In questions requiring an explanation, sufficient detail is required for full marks to be achieved (see below, **Question 2** and the last parts of **Questions 3, 5, 8 and 10,** for examples of this).

In general many candidates found parts of the ***Section B*** questions to be more testing than last year, with **Question 9**, on probability, proving to be particularly so.

Candidates generally coped well with the inclusion on the paper of 'change charts', a topic on the syllabus which has not appeared in recent years. The accuracy and labelling of the chart produced was generally very good.

Fewer marks were lost, than has been the case in the past, from candidates not providing a solution to the required degree of accuracy, although candidates were generally better at providing an answer to a specified number of decimal places than to the nearest cent (see below **Question 7** and **Question 11**).

Question 7, on probability and expectation, proved to be the least popular ***Section B*** question, although it was often well done by those who attempted it. **Questions 10** and **11**, on sampling and linear interpolation, proved, as last year, to be the most popular ***Section B*** questions, although many candidates struggled with the parts towards the end of **Question 11** in particular, and **Question 10** was not as well done as the equivalent question last year.

Comments on specific questions

Section A

Question 1

The majority of candidates answered most parts of this question correctly. Those that caused the most difficulty were parts **(iii)** and **(vi)**, with some candidates thinking that the range was not a suitable measure of dispersion for these data, and many thinking that the value 0 must be incorrect.

Answers: **(i)** False; **(ii)** True; **(iii)** True; **(iv)** False; **(v)** True; **(vi)** False.

Question 2

Candidates seemed familiar with this topic but many found it difficult to define sufficiently precisely the terms in part **(a)**. For example, in part **(a)(ii)** some candidates simply repeated the wording of the question, stating that seasonal variation is 'variation which occurs in a particular season' without expressing that the patterns observed repeat at regular intervals of time.

In part **(b)** many candidates seemed to remember examples of variables met in this context but were unable to specify them precisely enough. For example, 'ice cream' or 'winter clothes' rather than 'sales of ice cream' or 'sales of winter clothes' were common answers seen. With these (and other) variables many candidates stated their season in part **(b)(ii)** as 'summer' or 'winter' without specifying the exact length of the season. Those who did specify their season as lasting 3 months or 4 months, for example, were often then confused in part **(b)(iii)** as to whether the 'odd' or 'even' criterion for centring referred to the number of months in a season or the number of seasons in a year.

Question 3

The majority of candidates gained all the marks in the first three parts of this question. Where marks were dropped in part **(i)**, the main error seen was where the working shown led to +10.1% rather than -10.1% and the minus sign was then added with no explanation. In part **(ii)** some candidates lost the mark because of incorrect rounding from their calculator display.

In part **(iv)**, whilst many candidates were able to provide a suitable advantage of displaying data in a change chart, it was rare to see the correct disadvantage of 'the loss of original data' expressed.

Answers: **(ii)** –2.2, –6.0, +4.0.

Question 4

Again many candidates were able to score well on this question. The most common error in part **(i)** was that candidates thought that the explanation for *Y* being below the trend line was that the trend is decreasing. The most common incorrect answer to part **(ii)** was –6. In part **(iii)** some candidates simply took a reading of 29 from the graph and omitted to add 12 to this value.

Answers: **(ii)** 6; **(iii)** 41.

Question 5

For many candidates the first three parts of this question were straightforward. The most common error in parts **(i)** and **(ii)** was to give the general term of 'comparative bar chart' for one of the two charts. In part **(iii)** marks were sometimes lost for incomplete answers such as simply '*x*-axis' rather than 'label on the *x*-axis', although most candidates were able to score at least one of the two marks available.

In part **(iv)**, as with **Question 3**, candidates found stating an advantage of the diagram easier than stating a disadvantage, although sometimes the advantage was not stated in terms of the context of the question. The disadvantage that this diagram does not show overall totals was often missing.

Answers: **(i)** Sectional or component bar chart; **(ii)** Dual or multiple bar chart.

Question 6

Parts **(i)** and **(ii)** were well answered by the majority of candidates. Occasionally some candidates forgot to put their answers to part **(i)** in a suitable table.

Part **(iii)**, however, was far more difficult and most candidates felt that the explanation lay with the fact that one candidate had marks either side of the means whilst the other's were both above the class means. The correct reason however was that the impact of the scaling had a greater effect on the Statistics marks because the standard deviation for this subject had been enlarged by a greater amount.

Answers: **(i)** 82, 46, 66, 66; **(ii)** 83, 78, 128, 132.

Section B

Question 7

This was not a popular choice of question. Many candidates answered part **(i)** correctly but then decided that this was the question to omit. Amongst the work of those who chose to persist with this question there were many excellent solutions.

Common mistakes were the inclusion, in part **(iii)**, of wrong sequences such as 5-6-6 or 6-5-5 and a misunderstanding of the information given in part **(iv)** that the probability of failing to win a prize was 5/6. Some candidates read this as the probability of winning. Others applied the 1/6 to both the $2 and the $8 prizes in their expectation calculation. Most candidates realised that their expected payout needed to be compared with $1 to get the expected gain or loss. Some candidates did not observe the instruction in parts **(iv)** and **(v)** to give the answer to the nearest cent.

Answers: **(ii)** 125/216; **(iii)** 5/216; **(iv)** loss of 53 cents; **(v)** $5.68.

Question 8

This was quite a popular choice of question, although marks were lost by many at both the beginning and end of the question.

In part **(i)** marks were lost because the word 'expenditure' was often used instead of 'price/cost'. The difference between these terms seems to cause problems for many. In part **(ii)** quite a large number of candidates tried to calculate a weighted rather than an unweighted index. Some candidates lost the final mark as a result of not giving the answer to the required degree of accuracy, although not as many as in previous years. Similarly the accuracy mark was lost by some in part **(iii)**, although in general the method for achieving the weighted index was understood by the majority of candidates, and more were successful with part **(iii)** than part **(ii)**. Most candidates were successful with parts **(iv)** and **(vi),** and attempted a correct calculation in part **(v).** Some of these answers were not of the correct magnitude, for example an estimate of $803 rather than $803 000. Part **(vii)** was an example of a question where even the most able candidates rarely gave a full explanation worthy of both the available marks. Some candidates correctly identified the unweighted index as being too high and the fuel price relative as being the likely cause, but the full explanation, stating that it had been given equal weighting when in fact it had the lowest weight, was often missing.

Answers: **(ii)** 114.6; **(iii)** 110.0; **(v)** $836 000 or $837 000, $803 000.

Question 9

Marks varied considerably on this question with some candidates finding it very difficult.

Answers to part **(a)(i)** were often correct but **(a)(ii)** was often incorrect, with the most common error being 12/51. It was rare to see a correct answer to part **(a)(iii)** with many candidates simply describing how they had got their answers to parts **(i)** and **(ii)**. Many candidates did not explain that the existence of the ace of hearts meant that the events were not mutually exclusive in part **(a)(iv)**; indeed some candidates stated that such a card did not exist. There was generally more success with part **(b)** of this question and in particular **(b)(ii)** where many candidates seemed to know how to show that events were independent. In **(b)(i)** some candidates seemed to think the fact that P(A) + P(B) = 1 showed that the events were mutually exclusive. In **(b)(iii)** and **(b)(iv)** the required quantities were often calculated correctly, but only the most able candidates were able to interpret in words what these quantities represented. Common incorrect answers either did not refer to 'probability' or were simply an explanation in words of the formula, such as 'the probability of A plus B minus the intersection'.

Answers: **(a)(i)** 1/13; **(ii)** 1/13; **(b)(iii)** 0.79, the probability of A or B or both occurring; **(iv)** 0.58, the probability of A or B but not both occurring.

Question 10

This question on sampling proved to be popular, with most candidates selecting it.

The vast majority were successful with part **(i)**. Many scored the mark in part **(ii)** by rejecting stratified sampling. Those who suggested other sampling methods struggled to give a valid reason for their choice. Parts **(iii)(a)** and **(b)**, requiring selection of a simple random sample, were successfully completed by most candidates. Many, however, struggled with the systematic sample in part **(iv)**. In part **(v)** most candidates were able to calculate the make-up of the stratified sample, although errors did sometimes occur in its selection. The final part of this question was another example of a question where even the most able candidates often did not provide a full answer. Most candidates correctly pointed out that the simple random sample consisted only of books on rail transport, but it was rare for candidates to suggest that this would only be a concern if the survey was related to different types of transport.

Answers: **(i)** 7; **(iii)(a)** 20, 21, 30, 25, 38, 35, 39; **(iv)(a)** 00, 09; **(b)** 06; **(c)** 16, 26, 36, 46, 56, 66; **(v)(a)** Rail 4, Road 2, Sea 1; **(b)** 02, 33, 10, 36, 45, 55, 60.

Question 11

This question was attempted by almost all candidates, although many struggled with parts **(ii)** and **(vi)** in particular.

In part **(i)** many candidates correctly mentioned the existence of extreme values in the data, although some referred only to the fact that the mean takes account of all the values without pointing out that extreme values were present. Many candidates struggled with part **(ii)**, often giving an answer that is not a measure of dispersion. Parts **(iv)** and **(v)** caused more difficulties than similar questions have in the past, possibly because of the context of these questions. Many candidates were however able to score at least some of the marks by performing part of the method correctly. As with **Question 7**, some answers were required to the nearest cent, and some candidates did not give their answers to this degree of accuracy. In part **(vi)** more candidates were successful with part **(a)** than part **(b)**, with many correctly attempting to multiply their answer to part **(v)** by 1.2 and add 5 in **(a)**, but incorrectly leaving the standard deviation alone in **(b)** rather than multiplying it by 1.2.

Answers: **(iii)** 85, 205, 430, 565, 670, 690; **(iv)** 16.7%; **(v)** $60.24 (or $60.00 or $60.48); **(vi)(a)** $77.29 (or $77.00 or $77.58); **(b)** $31.25.

STATISTICS

Paper 4040/22

Paper 22

Key message

The best scripts typically included work which was well structured and demonstrated clear statistical thinking. Candidates scoring the highest marks showed clear methods on the numerical questions and precision and detail on those questions which required definitions and explanations.

General comments

In the best scripts candidates set out clear solutions showing evidence of the methods they had used at each stage of the problem. This is particularly important in multistage problems such as those found in **Questions 7(iii), (iv), (v), 11(v) and (vi)**. In these questions it was possible to award several marks for method, if it was sufficiently clear, even when numerical slips or errors had been made at an earlier stage in the problem.

In questions requiring an explanation, sufficient detail is required for full marks to be achieved (see below, **Question 2** and the last parts of **Questions 3, 5, 8 and 10,** for examples of this).

In general many candidates found parts of the ***Section B*** questions to be more testing than last year, with **Question 9**, on probability, proving to be particularly so.

Candidates generally coped well with the inclusion on the paper of 'change charts', a topic on the syllabus which has not appeared in recent years. The accuracy and labelling of the chart produced was generally very good.

Fewer marks were lost, than has been the case in the past, from candidates not providing a solution to the required degree of accuracy, although candidates were generally better at providing an answer to a specified number of decimal places than to the nearest cent (see below **Question 7** and **Question 11**).

Question 7, on probability and expectation, proved to be the least popular ***Section B*** question, although it was often well done by those who attempted it. **Questions 10** and **11**, on sampling and linear interpolation, proved, as last year, to be the most popular ***Section B*** questions, although many candidates struggled with the parts towards the end of **Question 11** in particular, and **Question 10** was not as well done as the equivalent question last year.

Comments on specific questions

Section A

Question 1

The majority of candidates answered most parts of this question correctly. Those that caused the most difficulty were parts **(iii)** and **(vi)**, with some candidates thinking that the range was not a suitable measure of dispersion for these data, and many thinking that the value 0 must be incorrect.

Answers: **(i)** False; **(ii)** True; **(iii)** True; **(iv)** False; **(v)** True; **(vi)** False.

Question 2

Candidates seemed familiar with this topic but many found it difficult to define sufficiently precisely the terms in part **(a)**. For example, in part **(a)(ii)** some candidates simply repeated the wording of the question, stating that seasonal variation is 'variation which occurs in a particular season' without expressing that the patterns observed repeat at regular intervals of time.

In part **(b)** many candidates seemed to remember examples of variables met in this context but were unable to specify them precisely enough. For example, 'ice cream' or 'winter clothes' rather than 'sales of ice cream' or 'sales of winter clothes' were common answers seen. With these (and other) variables many candidates stated their season in part **(b)(ii)** as 'summer' or 'winter' without specifying the exact length of the season. Those who did specify their season as lasting 3 months or 4 months, for example, were often then confused in part **(b)(iii)** as to whether the 'odd' or 'even' criterion for centring referred to the number of months in a season or the number of seasons in a year.

Question 3

The majority of candidates gained all the marks in the first three parts of this question. Where marks were dropped in part **(i)**, the main error seen was where the working shown led to +10.1% rather than -10.1% and the minus sign was then added with no explanation. In part **(ii)** some candidates lost the mark because of incorrect rounding from their calculator display.

In part **(iv)**, whilst many candidates were able to provide a suitable advantage of displaying data in a change chart, it was rare to see the correct disadvantage of 'the loss of original data' expressed.

Answers: **(ii)** –2.2, –6.0, +4.0.

Question 4

Again many candidates were able to score well on this question. The most common error in part **(i)** was that candidates thought that the explanation for *Y* being below the trend line was that the trend is decreasing. The most common incorrect answer to part **(ii)** was –6. In part **(iii)** some candidates simply took a reading of 29 from the graph and omitted to add 12 to this value.

Answers: **(ii)** 6; **(iii)** 41.

Question 5

For many candidates the first three parts of this question were straightforward. The most common error in parts **(i)** and **(ii)** was to give the general term of 'comparative bar chart' for one of the two charts. In part **(iii)** marks were sometimes lost for incomplete answers such as simply '*x*-axis' rather than 'label on the *x*-axis', although most candidates were able to score at least one of the two marks available.

In part **(iv)**, as with **Question 3**, candidates found stating an advantage of the diagram easier than stating a disadvantage, although sometimes the advantage was not stated in terms of the context of the question. The disadvantage that this diagram does not show overall totals was often missing.

Answers: **(i)** Sectional or component bar chart; **(ii)** Dual or multiple bar chart.

Question 6

Parts **(i)** and **(ii)** were well answered by the majority of candidates. Occasionally some candidates forgot to put their answers to part **(i)** in a suitable table.

Part **(iii)**, however, was far more difficult and most candidates felt that the explanation lay with the fact that one candidate had marks either side of the means whilst the other's were both above the class means. The correct reason however was that the impact of the scaling had a greater effect on the Statistics marks because the standard deviation for this subject had been enlarged by a greater amount.

Answers: **(i)** 82, 46, 66, 66; **(ii)** 83, 78, 128, 132.

Section B

Question 7

This was not a popular choice of question. Many candidates answered part **(i)** correctly but then decided that this was the question to omit. Amongst the work of those who chose to persist with this question there were many excellent solutions.

Common mistakes were the inclusion, in part **(iii)**, of wrong sequences such as 5-6-6 or 6-5-5 and a misunderstanding of the information given in part **(iv)** that the probability of failing to win a prize was 5/6. Some candidates read this as the probability of winning. Others applied the 1/6 to both the $2 and the $8 prizes in their expectation calculation. Most candidates realised that their expected payout needed to be compared with $1 to get the expected gain or loss. Some candidates did not observe the instruction in parts **(iv)** and **(v)** to give the answer to the nearest cent.

Answers: **(ii)** 125/216; **(iii)** 5/216; **(iv)** loss of 53 cents; **(v)** $5.68.

Question 8

This was quite a popular choice of question, although marks were lost by many at both the beginning and end of the question.

In part **(i)** marks were lost because the word 'expenditure' was often used instead of 'price/cost'. The difference between these terms seems to cause problems for many. In part **(ii)** quite a large number of candidates tried to calculate a weighted rather than an unweighted index. Some candidates lost the final mark as a result of not giving the answer to the required degree of accuracy, although not as many as in previous years. Similarly the accuracy mark was lost by some in part **(iii)**, although in general the method for achieving the weighted index was understood by the majority of candidates, and more were successful with part **(iii)** than part **(ii)**. Most candidates were successful with parts **(iv)** and **(vi),** and attempted a correct calculation in part **(v).** Some of these answers were not of the correct magnitude, for example an estimate of $803 rather than $803 000. Part **(vii)** was an example of a question where even the most able candidates rarely gave a full explanation worthy of both the available marks. Some candidates correctly identified the unweighted index as being too high and the fuel price relative as being the likely cause, but the full explanation, stating that it had been given equal weighting when in fact it had the lowest weight, was often missing.

Answers: **(ii)** 114.6; **(iii)** 110.0; **(v)** $836 000 or $837 000, $803 000.

Question 9

Marks varied considerably on this question with some candidates finding it very difficult.

Answers to part **(a)(i)** were often correct but **(a)(ii)** was often incorrect, with the most common error being 12/51. It was rare to see a correct answer to part **(a)(iii)** with many candidates simply describing how they had got their answers to parts **(i)** and **(ii)**. Many candidates did not explain that the existence of the ace of hearts meant that the events were not mutually exclusive in part **(a)(iv)**; indeed some candidates stated that such a card did not exist. There was generally more success with part **(b)** of this question and in particular **(b)(ii)** where many candidates seemed to know how to show that events were independent. In **(b)(i)** some candidates seemed to think the fact that P(A) + P(B) = 1 showed that the events were mutually exclusive. In **(b)(iii)** and **(b)(iv)** the required quantities were often calculated correctly, but only the most able candidates were able to interpret in words what these quantities represented. Common incorrect answers either did not refer to 'probability' or were simply an explanation in words of the formula, such as 'the probability of A plus B minus the intersection'.

Answers: **(a)(i)** 1/13; **(ii)** 1/13; **(b)(iii)** 0.79, the probability of A or B or both occurring; **(iv)** 0.58, the probability of A or B but not both occurring.

Question 10

This question on sampling proved to be popular, with most candidates selecting it.

The vast majority were successful with part **(i)**. Many scored the mark in part **(ii)** by rejecting stratified sampling. Those who suggested other sampling methods struggled to give a valid reason for their choice. Parts **(iii)(a)** and **(b)**, requiring selection of a simple random sample, were successfully completed by most candidates. Many, however, struggled with the systematic sample in part **(iv)**. In part **(v)** most candidates were able to calculate the make-up of the stratified sample, although errors did sometimes occur in its selection. The final part of this question was another example of a question where even the most able candidates often did not provide a full answer. Most candidates correctly pointed out that the simple random sample consisted only of books on rail transport, but it was rare for candidates to suggest that this would only be a concern if the survey was related to different types of transport.

Answers: **(i)** 7; **(iii)(a)** 20, 21, 30, 25, 38, 35, 39; **(iv)(a)** 00, 09; **(b)** 06; **(c)** 16, 26, 36, 46, 56, 66; **(v)(a)** Rail 4, Road 2, Sea 1; **(b)** 02, 33, 10, 36, 45, 55, 60.

Question 11

This question was attempted by almost all candidates, although many struggled with parts **(ii)** and **(vi)** in particular.

In part **(i)** many candidates correctly mentioned the existence of extreme values in the data, although some referred only to the fact that the mean takes account of all the values without pointing out that extreme values were present. Many candidates struggled with part **(ii)**, often giving an answer that is not a measure of dispersion. Parts **(iv)** and **(v)** caused more difficulties than similar questions have in the past, possibly because of the context of these questions. Many candidates were however able to score at least some of the marks by performing part of the method correctly. As with **Question 7**, some answers were required to the nearest cent, and some candidates did not give their answers to this degree of accuracy. In part **(vi)** more candidates were successful with part **(a)** than part **(b)**, with many correctly attempting to multiply their answer to part **(v)** by 1.2 and add 5 in **(a)**, but incorrectly leaving the standard deviation alone in **(b)** rather than multiplying it by 1.2.

Answers: **(iii)** 85, 205, 430, 565, 670, 690; **(iv)** 16.7%; **(v)** $60.24 (or $60.00 or $60.48); **(vi)(a)** $77.29 (or $77.00 or $77.58); **(b)** $31.25.

STATISTICS

Paper 4040/23

Paper 23

Key messages

- For questions requiring statistical or interpretive comments to be made, candidates must ensure that their comments are not overly general but relate specifically to the context of the question.
- Candidates should ensure that questions are read carefully to avoid misinterpretation of what is being asked for.

General comments

It is good that Centres/candidates do seem to be addressing something which has been commented on frequently in these reports in previous years, the matter of the level of accuracy to which final answers are stated, particularly when such accuracy is specified in a question. The general standard of comments was on the whole of a similar standard to last year, although there are a few areas where incorrect comments are still being made. Numerical work once again tended to be good.

As always appears to be the case when a question on the topic of expectation appears among the optional questions in ***Section B*** of a paper, it was by far the least popular and least well-answered of the questions in that Section.

Candidates should ensure that plotted points are marked by a clearly visible 'x', and not by a faint dot. Also, candidates should avoid overwriting rough pencil working with their final answers; the pencil work should be rubbed out or the final answers written separately. Both of these aspects have in some cases made it difficult for Examiners to award marks due to a lack of clarity.

Comments on specific questions

Section A

Question 1

Most candidates applied correctly the procedure for calculating the new mean, but a surprising number did not realise that the median and/or mode remained unchanged. 'New', incorrect values for the median were particularly common.

Answers: Mean = 11, median = 11.5, mode = 13.

Question 2

Most candidates obtained the correct numerical results in part **(i)**, although some then allocated the values to the incorrect workforce categories. In part **(ii)** a large number did not appreciate that because the only information available was in the form of percentages, it was impossible to tell whether or not the given statement was true.

Answers: **(i)** Executive 55, clerical 115, manual 80.

Question 3

Given that this topic had not been examined for many years, the answers submitted to this question were commendably good, both the diagrammatic work in part **(i)** and the comments in part **(ii)**.

Question 4

Although questions on this topic have appeared very regularly in recent years, there is clearly still considerable confusion in the minds of many candidates as regards the difference between mutual exclusivity and independence of events. Candidates who clearly knew the difference and the relevant basic probability formulae scored heavily on this question.

Answers: **(ii)** 0.5; **(iii)(a)** 0.25, **(b)** 0.55.

Question 5

Only a minority of candidates appreciated that the key to answering this question was the scaling relationship between the different variables. A few attempted lengthy calculations, even though numerical frequencies were not given. The small amount of working space given should have alerted candidates to the fact that such calculations were not required (or indeed possible) to answer the question successfully.

Answers: **(i)(a)** (m + 4), s, **(b)** 4m, 4s; **(ii)** 15.36, 2.24.

Question 6

Part **(i)** of this question was generally answered very well indeed, and there were two particularly commendable features of the answers submitted. Firstly, the common error in the selection of a systematic sample which was extremely prevalent only a few years ago, and mentioned explicitly in the report on last year's paper, now appears to have been considerably eliminated. Secondly, the fact that it was not possible to select a stratified sample which exactly matched the population proportionally was successfully resolved by most candidates. However, hardly any candidates were able to answer part **(ii)** correctly, suggesting that this is a use of random number tables which needs attention.

Answers: **(i)(a)** 15, 37, 50, 02, 49; **(b)(i)** 11; **(ii)** 23, 35, 47, 59; **(c)** 2 male 3 female.

Section B

Question 7

As is almost always the case with questions on the topic of index numbers, the numerical parts of the question were answered very well, but the parts requiring relevant comment much less so. The latter category in this question constituted the first two parts and the last. The only obvious conclusion about such answers is that many candidates know what they have to do to answer a question, but have no clear understanding of why they are doing what they are doing, or of what their answers 'mean'. A price relative, as the name implies, is the ratio of two prices, and therefore gives an indication of the percentage by which one price is greater or smaller than another. The correct answer to part **(i)** therefore was that the maintenance price had decreased by 4% between the two years. Any answer referring to the increase or decrease of a price relative or expenditure is automatically incorrect. In answering part **(viii)**, as to why the estimate obtained might be inaccurate, some candidates again produced 'bookwork' comments which did not contain any context and which therefore could score no marks. A few candidates are still giving 'inflation' as a reason for the possible inaccuracy, clearly not appreciating what 'inflation' means. In simple terms inflation is reflected in the change in prices, something which has been included in the calculation of the estimate. The possible inaccuracy might arise through changes in the quantities 'consumed', e.g. in this case the quantity of fuel used or the number of staff employed, quantities having been used to obtain weightings which may no longer be accurate.

Answers: **(iii)** 117; **(iv)** 107; **(v)** 110; **(vi)** 114.4; **(vii)** $7344.

Question 8

As was mentioned in the General Comments, this question was by far the least popular in ***Section B***, being attempted by fewer than one third of all candidates, and not well answered at all by most of those who did attempt it. Even those few candidates whose answers scored the majority of marks obtained hardly any at all in the final part. Almost all other candidates who attempted the question either abandoned it after the first few parts, or, at that stage, crossed out their answer and answered the other four ***Section B*** questions. The only conclusion which it is possible to draw is that it is the concept of expectation which causes candidates difficulty, because the standard of probability work required in this question was similar to that required in **Question 10**, a question which was in general answered very well.

Answers: **(i)(a)** 1/9, **(b)** 4/9, **(c)** $1.33; **(ii)(a)** 1/3, 2/9, 4/9, **(b)** 3, 2, 0; **(iii)** 52/81.

Question 9

The vast majority of candidates scored all the marks available for parts up to and including part **(b)(iv)**, but very few thereafter. In part **(b)(v)** any answer related to the fact that calculations were being based on a grouped frequency distribution would have scored the available mark, but almost none did. In part **(b)(vi)** there were a few general 'bookwork' comments about extreme values which did not answer the question, but again almost none referring to the actual shape of the given distribution in which there actually were extreme values. Strangely, although part **(b)(vii)** required the use of a procedure which has been examined frequently in previous years, and applied successfully by many candidates, this particular question produced hardly any correct solutions at all.

Answers: **(b)(ii)** 1, 15, 45, 74, 95, 107, 114, 117, 120; **(iii)** 25 – 29; **(iv)** 27.1 (or 27.2); **(vii)** 9.

Question 10

After the slight dip in the marks obtained by even the best candidates on the pure probability question last year, this question reverted to the experience of previous years, with the best candidates scoring full marks, a majority scoring more than half marks, and even the weakest scoring a few marks. The most frequent cause of loss of marks was candidates not appreciating that the sampling was without replacement.

Answers: **(i)(a)** 0.179, **(b)** 0.536, **(c)** 0.583; **(ii)(a)** 0.119, **(b)** 0.508, **(c)** 0.254; **(iii)** 0.023.

Question 11

This was a very standard question on the topic of moving averages and most candidates scored well on it. Comments were generally correct. Probably the most frequent cause of loss of marks was neglecting to apply the relevant seasonal component to the reading from the graph in part **(vii)**, even though the question explicitly mentioned doing this.

Answers: **(iv)** 221.7, 225.7, 232, 235, 240, 243.3, 247; **(vi)** 20.5.