STATISTICS

Paper 4040/12
Paper 1

Key messages

A question should always be read carefully to appreciate fully the practical situation described and the nature of the given data.

Use should be made of all the information given in a question. Given data should not be ignored.

It is as just as important to be able to explain and interpret the results of calculations as it is to be able to perform them accurately. Statistical analysis only has any real value when it is applied to a real-life practical situation.

General comments

The overall standard of work involving calculations of a routine nature was good. This was particularly true of the topics of pie charts, crude and standardised rates, grouped frequency distributions, and line of best fit. Answers were generally much more limited on the explanation and interpretation parts of questions. It was clear from some responses that few candidates understood properly the factors affecting the results of their calculations (see **Question 9** below).

There was evidence on this paper of candidates not reading carefully enough and understanding properly the given information in a question (see **Questions 2** and **4** below). There was one instance of it having sometimes been ignored totally (see **Question 9** below).

The need for students of this subject to appreciate the practical nature of Statistics in its application to a wide variety of real-life situations cannot be emphasised strongly enough. 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 what the results of such calculations demonstrate, and their implications when applied to the practical context of the question.

Comments on specific questions

Question 1

Answers generally indicated sound knowledge of the nature of stratified sampling and quota sampling. There was more confusion over the type of sampling method which is free from bias, and limited knowledge of the term ‘census’.

Answers: **(a)** stratified **(b)** random **(c)** census **(d)** quota

Question 2

There were few fully correct answers to this question, because many candidates clearly did not understand the nature of the given data. Thus it was quite common for the data to be seen, incorrectly, as two groups: the six non-zero values, and the two zeros. Such candidates commonly divided the sum of the six non-zero values by six in part **(a)**, and gave answers of 0 in part **(b)** and 14 in part **(c)**. The strongest answers, showing good understanding of the interconnecting parts, found the mean of the eight given numbers in part **(a)**, subtracted this from 7 in part **(b)**, then multiplied this by 8 in part **(c)**. Longer methods were seen and credited.

Answers: **(a)(i)** 2.75 **(ii)** 4.25 **(b)** 34

Question 3

Performance on the Venn diagram question was again very varied. Many candidates still demonstrate limited understanding of what the different regions of a Venn diagram represent. Correct answers were seen most often in parts **(a)(i)**, **(a)(ii)**, and **(b)**.

Answers: **(a)(i)** 32 **(ii)** 19 **(iii)** 13 **(iv)** 3 **(b)** the number of superior, smoking, mountain view rooms

Question 4

Those candidates who had read the question carefully, and understood what the given measures indicate in a distribution, were able to write down quickly, by inspection, the four correct answers to part **(a)**. Others seem to have regarded the numbers in the table as a data set, and carried out complex calculations with no meaning.

There was no uniquely acceptable answer to part **(b)**. An example is given below. The key idea required was that the median is unaffected by extreme values and these are unknown. Credit was also given to answers which pointed to the extremely large value of the upper quartile for Person B, making it probable that the total for Person B was probably larger (though by no means certain).

Answers: **(a)(i)(a)** 75% **(i)(b)** 60% **(a)(ii)(a)** 15% **(ii)(b)** 75%
(b) disagree, because there could be extremely large values for either person making their total larger, and these would not affect the median

Question 5

Reasonable understanding of the two-way table was shown, and many candidates obtained all the marks in parts **(a)** and **(b)**. There were fewer correct answers to part **(c)**, with some candidates presenting a frequency distribution for the number of outdoor courts only. Either choice was acceptable in part **(d)**, but it had to be well justified in order to obtain credit. The most serious limitation of many answers was that the reason offered was simply a description of what the chosen table displayed, for example, 'the first table, because it shows the number of indoor and outdoor courts'. The strongest answers made some comparison, pointing out the advantages of one table over the other. An example is given below. Vague, insubstantial reasons such as 'easy to interpret', 'easy to understand' were not accepted.

Answers: **(a)(i)** 7 **(ii)** 13 **(b)** 79 **(c)** additional values of x 3, 4, 6, 8, 9, 10, 12 corresponding values of f 1, 2, 4, 3, 2, 1, 1 **(d)** first, it contains more information than the second, and the second can be obtained from the first if required, but the first cannot be obtained from the second

Question 6

Almost all candidates showed excellent understanding of how to read a pie chart in parts **(a)** and **(b)**. Widespread knowledge was also shown of how to size comparative pie charts relating to different sample sizes in part **(c)**.

Answers: **(a)** 550 **(b)** 1350 **(c)** 5.2 cm

Question 7

Almost all candidates answered part **(a)** correctly. There were few fully correct answers to part **(b)**, though partial credit was given for any indication in an answer that choices were made without replacement. There was also partial credit for recognising that, for the situation described in the question, two greens or two blues had to be chosen, and that three further choices had to be made, after the initial blue. Some attempts included in the working the choice made in part **(a)**, producing products of four probabilities instead of three. Whilst this approach is perfectly valid, it involves more work and the use of conditional probability for correct completion. Very few successful attempts using this approach were seen.

Answers: **(a)** 4/7 **(b)** 4/15

Question 8

There were many fully correct answers to parts **(a)** and **(b)(i)**. Some answers to part **(b)(ii)** omitted to express the cumulative frequency corresponding to the given value as a percentage of 60. Very few acceptable answers to part **(c)** were seen. Even those candidates who recognised that the crucial observation related to the steepness of the graphs did not identify their central portions. Part **(d)** was well answered, but in part **(e)** a common error was to read the graph for 2800 Calories rather than 2200 Calories.

In all parts of the question the best responses showed very clearly, with lines drawn on the graph, how the various answers had been found. With such responses marks for method were possible, even if the graph had been read incorrectly.

Answers: **(a)(i)** 2620 **(ii)** 400 **(b)(i)** 2560 **(ii)** 65 **(c)** the central part of the graph for the women is steeper than the central part of the graph for the men **(d)** 39, 51 **(e)** 55

Question 9

Candidates generally continued to perform very well in the computation of crude and standardised rates, and there were many fully correct answers to parts **(a)**, **(b)** and **(c)**. Performance on part **(d)** was much more limited.

It was clear from answers to part **(d)** that very few candidates understood properly why in specific cases a crude rate and a standardised rate might be substantially different. Many answers recited, in very general terms, how crude rates and standardised rates are calculated, with no reference to the specific numerical values in the question. Very few focussed on the relative proportions of the different job groups within this population and how these differed from the corresponding proportions within the standard population.

Answers to part **(e)** were mixed: those candidates who used the numbers in the second table to calculate three probabilities were almost always successful; those who ignored this information, and offered a general opinion having done no calculations, earned no credit.

Answers: **(a)** 261.4 **(b)** 0, 60, 350, 100, 320 **(c)** 240.7 **(d)** full-time firefighters, which have the highest injury rate, make up a higher proportion in the population (about 62%) than in the standard population (55%) **(e)** respective probabilities are 0.132, 0.143, 0.125 so part-time firefighter

Question 10

Very good computational skills were seen in part **(a)** with many candidates producing fully correct answers. There was however 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.

As often seems to be the case with histograms, answers in part **(b)** fell into two categories: those that showed clear understanding of the principle on which a histogram is constructed, and those drawn simply with column heights equal to the given frequencies.

Answers to part **(c)** referring vaguely to 'grouped data' were not considered to be precise enough. Specific reference to the lack of knowledge of the masses of individual turtles was needed for credit.

Only modest appreciation was shown of the simple calculation required in part **(d)**. Many candidates attempted an irrelevant interpolation calculation based on a class in the second table only.

Answers: **(a)** 19.2, 1.75 **(b)** rectangle heights 2.5, 4, 14, 7.5
(c) the masses of individual turtles, both when collected and when released, are unknown **(d)** 12

Question 11

Almost all candidates understood in part **(b)** that the points plotted in part **(a)** showed negative correlation, although some felt that it was not a strong correlation. Answers to parts **(c)** and **(d)** were generally very good, with a correctly calculated upper semi-average, and a well explained method in finding the line of best fit. Most candidates used the averages in finding the line of best fit, a more accurate method than using points chosen from the drawn line, even though the latter is a valid method. The line was almost always well used in part **(e)** to produce the correct answer.

Answers to part **(f)** were more varied. Candidates who were able to extend what they had done in part **(e)** to the 7 letters and 13 letters of the second and third words, and form three probabilities, were usually successful.

Answers: **(b)** strong and negative **(c)** (9.6, 13) **(d)** $y = -1.54x + 27.8$ **(e)** 20 **(f)** 0.155

STATISTICS

Paper 4040/13
Paper 1

Key messages

A question should always be read carefully to appreciate fully the practical situation described and the nature of the given data.

Use should be made of all the information given in a question. Given data should not be ignored.

It is as just as important to be able to explain and interpret the results of calculations as it is to be able to perform them accurately. Statistical analysis only has any real value when it is applied to a real-life practical situation.

General comments

The overall standard of work involving calculations of a routine nature was good. This was particularly true of the topics of pie charts, crude and standardised rates, grouped frequency distributions, and line of best fit. Answers were generally much more limited on the explanation and interpretation parts of questions. It was clear from some responses that few candidates understood properly the factors affecting the results of their calculations (see **Question 9** below).

There was evidence on this paper of candidates not reading carefully enough and understanding properly the given information in a question (see **Questions 2** and **4** below). There was one instance of it having sometimes been ignored totally (see **Question 9** below).

The need for students of this subject to appreciate the practical nature of Statistics in its application to a wide variety of real-life situations cannot be emphasised strongly enough. 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 what the results of such calculations demonstrate, and their implications when applied to the practical context of the question.

Comments on specific questions

Question 1

Answers generally indicated sound knowledge of the nature of stratified sampling and quota sampling. There was more confusion over the type of sampling method which is free from bias, and limited knowledge of the term ‘census’.

Answers: **(a)** stratified **(b)** random **(c)** census **(d)** quota

Question 2

There were few fully correct answers to this question, because many candidates clearly did not understand the nature of the given data. Thus it was quite common for the data to be seen, incorrectly, as two groups: the six non-zero values, and the two zeros. Such candidates commonly divided the sum of the six non-zero values by six in part **(a)**, and gave answers of 0 in part **(b)** and 14 in part **(c)**. The strongest answers, showing good understanding of the interconnecting parts, found the mean of the eight given numbers in part **(a)**, subtracted this from 7 in part **(b)**, then multiplied this by 8 in part **(c)**. Longer methods were seen and credited.

Answers: **(a)(i)** 2.75 **(ii)** 4.25 **(b)** 34

Question 3

Performance on the Venn diagram question was again very varied. Many candidates still demonstrate limited understanding of what the different regions of a Venn diagram represent. Correct answers were seen most often in parts **(a)(i)**, **(a)(ii)**, and **(b)**.

Answers: **(a)(i)** 32 **(ii)** 19 **(iii)** 13 **(iv)** 3 **(b)** the number of superior, smoking, mountain view rooms

Question 4

Those candidates who had read the question carefully, and understood what the given measures indicate in a distribution, were able to write down quickly, by inspection, the four correct answers to part **(a)**. Others seem to have regarded the numbers in the table as a data set, and carried out complex calculations with no meaning.

There was no uniquely acceptable answer to part **(b)**. An example is given below. The key idea required was that the median is unaffected by extreme values and these are unknown. Credit was also given to answers which pointed to the extremely large value of the upper quartile for Person B, making it probable that the total for Person B was probably larger (though by no means certain).

Answers: **(a)(i)(a)** 75% **(i)(b)** 60% **(a)(ii)(a)** 15% **(ii)(b)** 75%
(b) disagree, because there could be extremely large values for either person making their total larger, and these would not affect the median

Question 5

Reasonable understanding of the two-way table was shown, and many candidates obtained all the marks in parts **(a)** and **(b)**. There were fewer correct answers to part **(c)**, with some candidates presenting a frequency distribution for the number of outdoor courts only. Either choice was acceptable in part **(d)**, but it had to be well justified in order to obtain credit. The most serious limitation of many answers was that the reason offered was simply a description of what the chosen table displayed, for example, 'the first table, because it shows the number of indoor and outdoor courts'. The strongest answers made some comparison, pointing out the advantages of one table over the other. An example is given below. Vague, insubstantial reasons such as 'easy to interpret', 'easy to understand' were not accepted.

Answers: **(a)(i)** 7 **(ii)** 13 **(b)** 79 **(c)** additional values of x 3, 4, 6, 8, 9, 10, 12 corresponding values of f 1, 2, 4, 3, 2, 1, 1 **(d)** first, it contains more information than the second, and the second can be obtained from the first if required, but the first cannot be obtained from the second

Question 6

Almost all candidates showed excellent understanding of how to read a pie chart in parts **(a)** and **(b)**. Widespread knowledge was also shown of how to size comparative pie charts relating to different sample sizes in part **(c)**.

Answers: **(a)** 550 **(b)** 1350 **(c)** 5.2 cm

Question 7

Almost all candidates answered part **(a)** correctly. There were few fully correct answers to part **(b)**, though partial credit was given for any indication in an answer that choices were made without replacement. There was also partial credit for recognising that, for the situation described in the question, two greens or two blues had to be chosen, and that three further choices had to be made, after the initial blue. Some attempts included in the working the choice made in part **(a)**, producing products of four probabilities instead of three. Whilst this approach is perfectly valid, it involves more work and the use of conditional probability for correct completion. Very few successful attempts using this approach were seen.

Answers: **(a)** 4/7 **(b)** 4/15

Question 8

There were many fully correct answers to parts **(a)** and **(b)(i)**. Some answers to part **(b)(ii)** omitted to express the cumulative frequency corresponding to the given value as a percentage of 60. Very few acceptable answers to part **(c)** were seen. Even those candidates who recognised that the crucial observation related to the steepness of the graphs did not identify their central portions. Part **(d)** was well answered, but in part **(e)** a common error was to read the graph for 2800 Calories rather than 2200 Calories.

In all parts of the question the best responses showed very clearly, with lines drawn on the graph, how the various answers had been found. With such responses marks for method were possible, even if the graph had been read incorrectly.

Answers: **(a)(i)** 2620 **(ii)** 400 **(b)(i)** 2560 **(ii)** 65 **(c)** the central part of the graph for the women is steeper than the central part of the graph for the men **(d)** 39, 51 **(e)** 55

Question 9

Candidates generally continued to perform very well in the computation of crude and standardised rates, and there were many fully correct answers to parts **(a)**, **(b)** and **(c)**. Performance on part **(d)** was much more limited.

It was clear from answers to part **(d)** that very few candidates understood properly why in specific cases a crude rate and a standardised rate might be substantially different. Many answers recited, in very general terms, how crude rates and standardised rates are calculated, with no reference to the specific numerical values in the question. Very few focussed on the relative proportions of the different job groups within this population and how these differed from the corresponding proportions within the standard population.

Answers to part **(e)** were mixed: those candidates who used the numbers in the second table to calculate three probabilities were almost always successful; those who ignored this information, and offered a general opinion having done no calculations, earned no credit.

Answers: **(a)** 261.4 **(b)** 0, 60, 350, 100, 320 **(c)** 240.7 **(d)** full-time firefighters, which have the highest injury rate, make up a higher proportion in the population (about 62%) than in the standard population (55%) **(e)** respective probabilities are 0.132, 0.143, 0.125 so part-time firefighter

Question 10

Very good computational skills were seen in part **(a)** with many candidates producing fully correct answers. There was however 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.

As often seems to be the case with histograms, answers in part **(b)** fell into two categories: those that showed clear understanding of the principle on which a histogram is constructed, and those drawn simply with column heights equal to the given frequencies.

Answers to part **(c)** referring vaguely to ‘grouped data’ were not considered to be precise enough. Specific reference to the lack of knowledge of the masses of individual turtles was needed for credit.

Only modest appreciation was shown of the simple calculation required in part **(d)**. Many candidates attempted an irrelevant interpolation calculation based on a class in the second table only.

Answers: **(a)** 19.2, 1.75 **(b)** rectangle heights 2.5, 4, 14, 7.5
(c) the masses of individual turtles, both when collected and when released, are unknown **(d)** 12

Question 11

Almost all candidates understood in part **(b)** that the points plotted in part **(a)** showed negative correlation, although some felt that it was not a strong correlation. Answers to parts **(c)** and **(d)** were generally very good, with a correctly calculated upper semi-average, and a well explained method in finding the line of best fit. Most candidates used the averages in finding the line of best fit, a more accurate method than using points chosen from the drawn line, even though the latter is a valid method. The line was almost always well used in part **(e)** to produce the correct answer.

Answers to part **(f)** were more varied. Candidates who were able to extend what they had done in part **(e)** to the 7 letters and 13 letters of the second and third words, and form three probabilities, were usually successful.

Answers: **(b)** strong and negative **(c)** (9.6, 13) **(d)** $y = -1.54x + 27.8$ **(e)** 20 **(f)** 0.155

STATISTICS

Paper 4040/22
Paper 2

Key message

This examination requires candidates to be able to produce appropriate statistical diagrams, to make statistical calculations and to interpret these diagrams and calculations in context. Candidates should have an appreciation of the advantages and disadvantages of particular statistical diagrams and calculations, and thus be able to make judgments about which would be the most appropriate to use in specific circumstances. The best statistical diagrams should be accurately drawn, taking care to use scales correctly, and have clearly labelled axes, including units, where appropriate. Candidates scoring the highest marks in the numerical calculations will provide clear indications of the methods they have used in logical and clearly presented solutions. In questions that require interpretation of those statistical techniques, the most successful responses will include detailed explanations in the context of the situation presented.

General comments

This was the first year in which the new syllabus was examined. The curriculum content has been updated and the optional questions removed so that candidates answer all of the questions on the paper. The examination continues to include shorter questions at the start of the paper followed by four longer questions.

Some of the new content, namely stem-and-leaf and box-and-whisker diagrams, formed the basis of one of the longer answer questions on this paper, **Question 8**. The responses to this question reflected those of the rest of this paper, and previous papers, in the following ways:

1. The advantage that one particular statistical diagram has over another was not always known.
2. The numerical part of the question, namely identifying the quartiles from the stem-and-leaf diagram, was usually well done, with the key being used correctly.
3. The statistical diagrams, in this case a pair of box-and-whisker diagrams, were accurately drawn with scales used correctly, but labelling of the axis was often missing.
4. The interpretation of these diagrams was sometimes incorrect or lacked sufficient detail.

This pattern, of questions assessing interpretation and communication (AO2) proving more difficult than those assessing knowledge and techniques (AO1), was seen throughout the paper.

Comments on specific questions

Question 1

In **part (a)**, it was necessary to find upper and lower class boundaries of age in completed years. This proved difficult for many candidates, with common errors being 9.5 for the lower boundary and/or 13.5 for the upper boundary.

Most candidates did **part (b)** well, scoring at least two of the three available marks. The most common error was for ‘discrete’ to be missing from the description of the data that will be collected from Question B. A few candidates described the method of data collection, by referring to open/closed questions, rather than providing a description of the data itself.

Answers: **(a)** 10, 14; **(ii)** Question B will give discrete quantitative data, Question C will give qualitative data.

Question 2

It was pleasing to see so many successful attempts at **part (a)**, with most candidates correctly using the result, $P(X \cup Y) = P(X) + P(Y)$, for mutually exclusive events.

Part (b) was much more difficult, as it was necessary to use the result for independent events, $P(X \cap Y) = P(X) \times P(Y) = q \times \frac{3}{5}$, combined with the general result $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$. This led to an equation in which the unknown q appeared twice. Most candidates who reached this equation solved it successfully. A large number tried to find a numerical value for $P(X \cap Y)$ before inserting it into $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$, often resorting to using $P(X) = \frac{1}{10}$ (from part **(a)**), in order to achieve this.

Answers: **(a)** $\frac{1}{10}$; **(b)** $\frac{1}{4}$.

Question 3

Fully correct responses were seen from the most able candidates, with many partially correct responses also seen. For a fully correct solution, it was necessary to add together the probabilities of all three ways in which 2 men and 1 woman could be selected and the probability of the one way in which 3 men and 0 women could be selected. The most common errors were for the option of selecting 3 men to be missing or for only one way of selecting 2 men and 1 woman to be considered. The probabilities of each option were usually found correctly, with the majority of candidates correctly reducing the denominator by one each time in their products of three probabilities.

Answer: $\frac{11}{60}$.

Question 4

Most candidates found the correct sample stratified by age group in **part (a)**. A few selected samples in which the number of patients aged 50 or under was the same as the number aged over 50, rather than correctly considering the proportions in the whole population of 90 patients.

In **part (b)** the population was further divided into the department in which the patient was staying, and it was necessary to comment on how representative the sample from **(a)** was when considering the departments only. Space was made available on the question paper for the necessary calculations, although many candidates chose not to use this space for that purpose. Those with a correct sample in part **(a)** should have found that their sample was split between the departments A, B and C in the ratio 2 : 2 : 2. They should then have calculated that the population was split between those departments in the ratio 1 : 2 : 3 and hence that the sample was not representative in terms of the department. Some partially correct responses were seen, the most common being that the ratio 2 : 2 : 2, above, was found correctly, but no further calculations were seen as these candidates incorrectly believed that showing that each department was represented equally meant that they had shown the sample to be representative of the departments.

Candidates found **part (c)** difficult, as it involved looking back to the original purpose of the survey before a decision could be made. Either option of stratifying by age group or by department could be considered correct, so long as the candidate referred back to the original purpose of the survey. An example of a good answer is '(in order to be representative of care received) Pedro should stratify by department because patients in the same department are likely to receive similar care'.

Answers: **(a)** 78, 06, 47, 13, 64, 51.

Question 5

Good responses were seen in **parts (a)**, **(b)** and **(c)**. Many candidates showed their working clearly and were therefore able to gain method marks for partially correct responses. For some candidates there seemed to be some confusion about which values were the actual ones and which were the standardised ones.

It was rare to see a fully correct response to **part (d)**, suggesting that many candidates could calculate standardised scores without necessarily understanding the purpose of doing so.

Correct responses were often seen in **part (e)**, but an incorrect answer of 50 and 22 was quite common.

Answers: **(a)** 24; **(b)** 25; **(c)** 44; **(d)** Dancing, as it has the highest standardised score; **(e)** 50 and 12.

Question 6

This question on expectation was correctly answered by many candidates and, in particular, fully correct answers were usually seen to **parts (a)** and **(c)**. Most candidates were able to use their answer to part **(a)** correctly in the solution to **part (b)**, although a few incorrectly multiplied together the probabilities that each type of tomato would produce fruit.

The most difficult part of this question was **part (d)**, although many fully correct solutions were seen. The most common error was an incomplete solution of 12, the number of Ruby Red seeds that must produce fruit. It was necessary to then divide this by 0.75 to obtain the number of Ruby Red seeds that should be put in the packet.

Answers: **(a)** $\frac{3}{5}$; **(b)** 0.78; **(c)** 39; **(d)** 16.

Question 7

This time series question was the first of the longer questions on the paper.

In **part (a)**, many fully correct responses were seen, but a significant number gave the values of *n* as 4 and 5, respectively. Most were then able to make a correct decision for their values of *n* regarding whether or not centring was necessary, with correct reasons provided for these choices.

Correct results were usually seen in **part (b)**. Some fully correct solutions were seen in **part (c)**, but more often they were either partially correct or incorrect. Some of the partially correct results seen came from subtracting the profits from the moving averages, rather than the other way around, or from finding just one correct difference rather than averaging 3 correct differences. Some incorrect methods seen involved subtracting the 3-point moving totals rather than the 3-point moving averages from the profits or simply averaging the three Jan – Apr profits.

Plots were usually accurate in **part (d)**, with appropriate trend lines drawn. A few candidates incorrectly joined up their plots rather than producing a trend line. The attempt at a trend line from some candidates was simply a line joining the first and last plots, which, in this case, did not produce a suitable trend line. Most candidates correctly identified that the profit was increasing in **part (e)**, giving their answer in context.

Those candidates who had produced a correct response to **part (c)** were usually able to use this value and their trend line to answer **part (f)** correctly. The most common error was for a reading simply to be taken from the trend line with no use made of the answer to **part (c)**. In **part (g)**, only a few of the most able candidates spotted that the points that they had plotted in **part (d)** looked more like they would lie on a curve than a straight line.

Answers: **(a)** $n = 3$, centring is not necessary as *n* is odd; $n = 4$, centring is necessary as *n* is even; **(b)** 36 100 and 12 030; **(c)** –1700; **(e)** Profits are increasing; **(f)** 11 500; **(g)** Points appear to lie on a curve rather than a straight line.

Question 8

The advantage that a stem-and-leaf diagram has over a box-and-whisker diagram was often given incorrectly in **part (a)**. Common errors seen were to state that it is ‘easier to draw’ or that it ‘has a key’. Advantages of a particular diagram should be given in terms of their benefits when it comes to interpreting them, rather than benefits in terms of ease of drawing. In this case, an example of a correct response is ‘in a stem-and-leaf diagram the original data is not lost’.

There were many fully correct answers to **part (b),** with the vast majority getting the least and greatest values correct. It was particularly impressive to see that the key had been used correctly in the vast majority of scripts with, for example, the least value for Eastpool being given correctly as 0.3 rather than simply 3.

Appropriate scales were usually seen in **part (c)**, together with accurately drawn diagrams. Labels explaining that the numbers on the axis represented hours of sunshine were often missing, however. A small number of candidates did not appear to know what a box-and-whisker diagram was, leaving this question blank.

The interpretation of the box-and-whisker diagrams in **part (d)** was sometimes incorrect. Some candidates appeared to think that a larger box indicated more sunshine, with ‘more sunshine in Eastpool’ being a common wrong answer. It was also fairly common for just one comparison, or two identical comparisons, to be made. For two marks, comparisons, in context, of both the central tendency and the dispersion was expected. For example, ‘there was more sunshine in Westsea and more varied amounts of sunshine in Eastpool’ would be a good answer. Many correct answers were seen to **part (e),** although a few candidates thought that the length of the diagram represented the mean or the median.

Answers: **(b)** 0.3, 1.7, 3.3, 5.0, 7.1; 2.8, 4.3, 5.2, 6.1, 7.3; **(e)(i)** Range; **(ii)** Interquartile range.

Question 9

Fully correct answers were often seen in **parts (a)**, **(b)** and **(c)**. In **part (a),** a small number of candidates simply subtracted the percentages and gave an answer of 114.

Errors in the interpretation of the weighted aggregate cost index were quite common in **part (d)**. Many candidates correctly stated the percentage increase of 11.3%, but stated that this was an increase in the ‘expenditure’ rather than in the ‘costs’ or ‘prices’. This increase can only be expressed as an increase in the expenditure if the proviso ‘assuming that the quantities remain the same’ is also included. The dates between which this change had taken place were sometimes missing, and this is required for a fully correct answer (see below).

Most candidates were able to estimate correctly the expenditure in **part (e)**. The explanations required in **part (f)** were more difficult, but were often given correctly by the more able candidates. An example of a fully correct answer is ‘Gouta’s explanation is not correct because this increase in costs has already been accounted for in the calculation of the estimate. Suma’s explanation is correct because a change in quantities has not been accounted for in the estimate’.

Answers: **(a)** 113; **(b)** 104, 108, 97; **(c)** 111.3; **(d)** costs increased by 11.3% between 2015 and 2017**;**
(e) $390.

Question 10

There were some good attempts at **part (a)**, although in some cases the presentation was not as good as usual, which made the logic of the solutions sometimes difficult to follow. A comparison of the morning and afternoon mean masses was usually correct, and given in the context of the problem.

Many candidates correctly used the formula for the standard deviation in **part (b)**, although some approximated prematurely and ended up with an incorrect final answer. Some incorrectly squared 28.29, the sum of the squares of the masses, and others squared rather than square rooted the variance.

In **part (c)** an advantage of the interquartile range as a measure of dispersion was required. As with the advantage of a stem-and-leaf diagram, earlier, some candidates focused on the relative ease of calculation compared with that of the standard deviation, rather than an advantage in terms of how it can be used. A correct answer was to say, for example, that the interquartile range would not be distorted by extreme values.

It was pleasing to see so many fully correct answers to **part (d)**, although a surprising number of candidates recalculated the lower quartile, perhaps missing that this information had been provided under the table.

The more able candidates were successful with this final part of this question. A common error was for the number of fish between 0.1 kg and 0.2 kg (14) to be calculated, but then for no further progress to be made. Another common incorrect answer seen was 26, where the 6 had effectively been added to 14 twice; this occurred when candidates, having taken a ratios approach to finding *n*, such as $\frac{(n-6)}{(27-6)} = \frac{(0.2 - 0.1)}{(0.25-0.1)}$, solved for *n* but then added a further 6. Some other candidates also appeared to be mixing methods, working from both the upper and the lower boundary. So for example, the number of fish between 0.2 kg and 0.25 kg (7) was added to 6 rather than subtracted from 27.

Answers: **(a)** 0.202, fish sold in the morning generally have a greater mass than fish sold in the afternoon; **(b)** 0.332; **(d)** 0.647, 0.44; **(e)** 20.

STATISTICS

Paper 4040/23
Paper 2

Key message

This examination requires candidates to be able to produce appropriate statistical diagrams, to make statistical calculations and to interpret these diagrams and calculations in context. Candidates should have an appreciation of the advantages and disadvantages of particular statistical diagrams and calculations, and thus be able to make judgments about which would be the most appropriate to use in specific circumstances. The best statistical diagrams should be accurately drawn, taking care to use scales correctly, and have clearly labelled axes, including units, where appropriate. Candidates scoring the highest marks in the numerical calculations will provide clear indications of the methods they have used in logical and clearly presented solutions. In questions that require interpretation of those statistical techniques, the most successful responses will include detailed explanations in the context of the situation presented.

General comments

This was the first year in which the new syllabus was examined. The curriculum content has been updated and the optional questions removed so that candidates answer all of the questions on the paper. The examination continues to include shorter questions at the start of the paper followed by four longer questions.

Some of the new content, namely stem-and-leaf and box-and-whisker diagrams, formed the basis of one of the longer answer questions on this paper, **Question 8**. The responses to this question reflected those of the rest of this paper, and previous papers, in the following ways:

1. The advantage that one particular statistical diagram has over another was not always known.
2. The numerical part of the question, namely identifying the quartiles from the stem-and-leaf diagram, was usually well done, with the key being used correctly.
3. The statistical diagrams, in this case a pair of box-and-whisker diagrams, were accurately drawn with scales used correctly, but labelling of the axis was often missing.
4. The interpretation of these diagrams was sometimes incorrect or lacked sufficient detail.

This pattern, of questions assessing interpretation and communication (AO2) proving more difficult than those assessing knowledge and techniques (AO1), was seen throughout the paper.

Comments on specific questions

Question 1

In **part (a)**, it was necessary to find upper and lower class boundaries of age in completed years. This proved difficult for many candidates, with common errors being 9.5 for the lower boundary and/or 13.5 for the upper boundary.

Most candidates did **part (b)** well, scoring at least two of the three available marks. The most common error was for ‘discrete’ to be missing from the description of the data that will be collected from Question B. A few candidates described the method of data collection, by referring to open/closed questions, rather than providing a description of the data itself.

Answers: **(a)** 10, 14; **(ii)** Question B will give discrete quantitative data, Question C will give qualitative data.

Question 2

It was pleasing to see so many successful attempts at **part (a)**, with most candidates correctly using the result, $P(X \cup Y) = P(X) + P(Y)$, for mutually exclusive events.

Part (b) was much more difficult, as it was necessary to use the result for independent events, $P(X \cap Y) = P(X) \times P(Y) = q \times \frac{3}{5}$, combined with the general result $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$. This led to an equation in which the unknown q appeared twice. Most candidates who reached this equation solved it successfully. A large number tried to find a numerical value for $P(X \cap Y)$ before inserting it into $P(X \cup Y) = P(X) + P(Y) - P(X \cap Y)$, often resorting to using $P(X) = \frac{1}{10}$ (from part **(a)**), in order to achieve this.

Answers: **(a)** $\frac{1}{10}$; **(b)** $\frac{1}{4}$.

Question 3

Fully correct responses were seen from the most able candidates, with many partially correct responses also seen. For a fully correct solution, it was necessary to add together the probabilities of all three ways in which 2 men and 1 woman could be selected and the probability of the one way in which 3 men and 0 women could be selected. The most common errors were for the option of selecting 3 men to be missing or for only one way of selecting 2 men and 1 woman to be considered. The probabilities of each option were usually found correctly, with the majority of candidates correctly reducing the denominator by one each time in their products of three probabilities.

Answer: $\frac{11}{60}$.

Question 4

Most candidates found the correct sample stratified by age group in **part (a)**. A few selected samples in which the number of patients aged 50 or under was the same as the number aged over 50, rather than correctly considering the proportions in the whole population of 90 patients.

In **part (b)** the population was further divided into the department in which the patient was staying, and it was necessary to comment on how representative the sample from **(a)** was when considering the departments only. Space was made available on the question paper for the necessary calculations, although many candidates chose not to use this space for that purpose. Those with a correct sample in part **(a)** should have found that their sample was split between the departments A, B and C in the ratio 2 : 2 : 2. They should then have calculated that the population was split between those departments in the ratio 1 : 2 : 3 and hence that the sample was not representative in terms of the department. Some partially correct responses were seen, the most common being that the ratio 2 : 2 : 2, above, was found correctly, but no further calculations were seen as these candidates incorrectly believed that showing that each department was represented equally meant that they had shown the sample to be representative of the departments.

Candidates found **part (c)** difficult, as it involved looking back to the original purpose of the survey before a decision could be made. Either option of stratifying by age group or by department could be considered correct, so long as the candidate referred back to the original purpose of the survey. An example of a good answer is '(in order to be representative of care received) Pedro should stratify by department because patients in the same department are likely to receive similar care'.

Answers: **(a)** 78, 06, 47, 13, 64, 51.

Question 5

Good responses were seen in **parts (a)**, **(b)** and **(c)**. Many candidates showed their working clearly and were therefore able to gain method marks for partially correct responses. For some candidates there seemed to be some confusion about which values were the actual ones and which were the standardised ones.

It was rare to see a fully correct response to **part (d)**, suggesting that many candidates could calculate standardised scores without necessarily understanding the purpose of doing so.

Correct responses were often seen in **part (e)**, but an incorrect answer of 50 and 22 was quite common.

Answers: **(a)** 24; **(b)** 25; **(c)** 44; **(d)** Dancing, as it has the highest standardised score; **(e)** 50 and 12.

Question 6

This question on expectation was correctly answered by many candidates and, in particular, fully correct answers were usually seen to **parts (a)** and **(c)**. Most candidates were able to use their answer to part **(a)** correctly in the solution to **part (b)**, although a few incorrectly multiplied together the probabilities that each type of tomato would produce fruit.

The most difficult part of this question was **part (d)**, although many fully correct solutions were seen. The most common error was an incomplete solution of 12, the number of Ruby Red seeds that must produce fruit. It was necessary to then divide this by 0.75 to obtain the number of Ruby Red seeds that should be put in the packet.

Answers: **(a)** $\frac{3}{5}$; **(b)** 0.78; **(c)** 39; **(d)** 16.

Question 7

This time series question was the first of the longer questions on the paper.

In **part (a)**, many fully correct responses were seen, but a significant number gave the values of n as 4 and 5, respectively. Most were then able to make a correct decision for their values of n regarding whether or not centring was necessary, with correct reasons provided for these choices.

Correct results were usually seen in **part (b)**. Some fully correct solutions were seen in **part (c)**, but more often they were either partially correct or incorrect. Some of the partially correct results seen came from subtracting the profits from the moving averages, rather than the other way around, or from finding just one correct difference rather than averaging 3 correct differences. Some incorrect methods seen involved subtracting the 3-point moving totals rather than the 3-point moving averages from the profits or simply averaging the three Jan – Apr profits.

Plots were usually accurate in **part (d)**, with appropriate trend lines drawn. A few candidates incorrectly joined up their plots rather than producing a trend line. The attempt at a trend line from some candidates was simply a line joining the first and last plots, which, in this case, did not produce a suitable trend line. Most candidates correctly identified that the profit was increasing in **part (e)**, giving their answer in context.

Those candidates who had produced a correct response to **part (c)** were usually able to use this value and their trend line to answer **part (f)** correctly. The most common error was for a reading simply to be taken from the trend line with no use made of the answer to **part (c)**. In **part (g)**, only a few of the most able candidates spotted that the points that they had plotted in **part (d)** looked more like they would lie on a curve than a straight line.

Answers: **(a)** $n = 3$, centring is not necessary as n is odd; $n = 4$, centring is necessary as n is even; **(b)** 36 100 and 12 030; **(c)** –1700; **(e)** Profits are increasing; **(f)** 11 500; **(g)** Points appear to lie on a curve rather than a straight line.

Question 8

The advantage that a stem-and-leaf diagram has over a box-and-whisker diagram was often given incorrectly in **part (a)**. Common errors seen were to state that it is 'easier to draw' or that it 'has a key'. Advantages of a particular diagram should be given in terms of their benefits when it comes to interpreting them, rather than benefits in terms of ease of drawing. In this case, an example of a correct response is 'in a stem-and-leaf diagram the original data is not lost'.

There were many fully correct answers to **part (b),** with the vast majority getting the least and greatest values correct. It was particularly impressive to see that the key had been used correctly in the vast majority of scripts with, for example, the least value for Eastpool being given correctly as 0.3 rather than simply 3.

Appropriate scales were usually seen in **part (c)**, together with accurately drawn diagrams. Labels explaining that the numbers on the axis represented hours of sunshine were often missing, however. A small number of candidates did not appear to know what a box-and-whisker diagram was, leaving this question blank.

The interpretation of the box-and-whisker diagrams in **part (d)** was sometimes incorrect. Some candidates appeared to think that a larger box indicated more sunshine, with 'more sunshine in Eastpool' being a common wrong answer. It was also fairly common for just one comparison, or two identical comparisons, to be made. For two marks, comparisons, in context, of both the central tendency and the dispersion was expected. For example, 'there was more sunshine in Westsea and more varied amounts of sunshine in Eastpool' would be a good answer. Many correct answers were seen to **part (e),** although a few candidates thought that the length of the diagram represented the mean or the median.

Answers: **(b)** 0.3, 1.7, 3.3, 5.0, 7.1; 2.8, 4.3, 5.2, 6.1, 7.3; **(e)(i)** Range; **(ii)** Interquartile range.

Question 9

Fully correct answers were often seen in **parts (a)**, **(b)** and **(c)**. In **part (a),** a small number of candidates simply subtracted the percentages and gave an answer of 114.

Errors in the interpretation of the weighted aggregate cost index were quite common in **part (d)**. Many candidates correctly stated the percentage increase of 11.3%, but stated that this was an increase in the 'expenditure' rather than in the 'costs' or 'prices'. This increase can only be expressed as an increase in the expenditure if the proviso 'assuming that the quantities remain the same' is also included. The dates between which this change had taken place were sometimes missing, and this is required for a fully correct answer (see below).

Most candidates were able to estimate correctly the expenditure in **part (e)**. The explanations required in **part (f)** were more difficult, but were often given correctly by the more able candidates. An example of a fully correct answer is 'Gouta's explanation is not correct because this increase in costs has already been accounted for in the calculation of the estimate. Suma's explanation is correct because a change in quantities has not been accounted for in the estimate'.

Answers: **(a)** 113; **(b)** 104, 108, 97; **(c)** 111.3; **(d)** costs increased by 11.3% between 2015 and 2017**;**
(e) $390.

Question 10

There were some good attempts at **part (a)**, although in some cases the presentation was not as good as usual, which made the logic of the solutions sometimes difficult to follow. A comparison of the morning and afternoon mean masses was usually correct, and given in the context of the problem.

Many candidates correctly used the formula for the standard deviation in **part (b)**, although some approximated prematurely and ended up with an incorrect final answer. Some incorrectly squared 28.29, the sum of the squares of the masses, and others squared rather than square rooted the variance.

In **part (c)** an advantage of the interquartile range as a measure of dispersion was required. As with the advantage of a stem-and-leaf diagram, earlier, some candidates focused on the relative ease of calculation compared with that of the standard deviation, rather than an advantage in terms of how it can be used. A correct answer was to say, for example, that the interquartile range would not be distorted by extreme values.

It was pleasing to see so many fully correct answers to **part (d)**, although a surprising number of candidates recalculated the lower quartile, perhaps missing that this information had been provided under the table.

The more able candidates were successful with this final part of this question. A common error was for the number of fish between 0.1 kg and 0.2 kg (14) to be calculated, but then for no further progress to be made. Another common incorrect answer seen was 26, where the 6 had effectively been added to 14 twice; this occurred when candidates, having taken a ratios approach to finding *n*, such as $\frac{(n-6)}{(27-6)} = \frac{(0.2 - 0.1)}{(0.25-0.1)}$, solved for *n* but then added a further 6. Some other candidates also appeared to be mixing methods, working from both the upper and the lower boundary. So for example, the number of fish between 0.2 kg and 0.25 kg (7) was added to 6 rather than subtracted from 27.

Answers: **(a)** 0.202, fish sold in the morning generally have a greater mass than fish sold in the afternoon; **(b)** 0.332; **(d)** 0.647, 0.44; **(e)** 20.