MATHEMATICS

Paper 9709/12
Pure Mathematics 1

Key messages

Much like in previous exam series, it was clear some candidates had used their calculators when solving equations and evaluating definite integrals. There were a significant number of instances of candidates using calculators in questions on these topics and therefore providing inadequate working in their response, resulting in them being unable to gain full credit.

Candidates would benefit from more time spent practicing questions containing negative values. This was particularly noticeable in responses to **Question 3(a)**, where minus signs were often ignored, and **5(a)**, where the correct value for r was often discounted because it was negative.

When working with decimal numbers, candidates need to work to at least 4 significant figures to ensure accuracy when they round their answer to 3 significant figures. Candidates are reminded that exact decimal answers should not be corrected to 3 significant figures and that improper fractions which are equivalent to integers are not acceptable as final answers.

General comments

The paper was generally found to be accessible for candidates and many very good scripts were seen. Questions involving trigonometry and differentiation were particularly well done by the majority of candidates. Most candidates were able to complete the paper in the available time. Candidates should take more care to ensure that they read a question carefully before beginning a response.

Comments on specific questions

Question 1

For most candidates, this question provided a straightforward introduction to the paper and a good source of marks. Some candidates, however, did not realise that an interval was required and omitted this part of the solution entirely.

Question 2

(a) Most candidates were able to correctly differentiate the given equation to find an expression for $\frac{dy}{dx}$ and understood that they needed to use $x = 1$ to find the gradient at the point P. After differentiating correctly, candidates usually went on to gain full marks.

(b) Successful candidates dealt with the equation arising from $\frac{dy}{dx} = 0$ and arrived at $x = \sqrt[3]{-\frac{5}{4}}$. Some candidates, however, started by differentiating their expression for $\frac{dy}{dx}$ and were unable to gain any marks. A large number of candidates did not provide answers of sufficient accuracy due to premature approximation of decimals.

Question 3

(a) Most candidates produced a fully correct expansion, but there was a minority who attempted to multiply out four brackets, usually unsuccessfully. Poorer attempts at this question did not apply the power of 4 correctly to the first and last terms or incorrectly simplified the other three terms. Not dealing with the minus sign correctly or ignoring it completely was a common error. Candidates are encouraged to write out the full expansion rather than leaving the terms as a list.

(b) Almost all candidates with an expansion selected the two terms needed to produce the x^2 coefficient. A large majority scored both marks.

Question 4

(a) The majority of candidates scored full marks on this question. Candidates usually found AB by the cosine rule, or $\frac{1}{2}AB$ using the trigonometric formulae for right-angled triangles. Correct answers were obtained by finding the arc CD using $r\theta$ and adding this and 8 to AB. Candidates who changed from radians to degrees and worked out a fraction of the circumference often lost accuracy in their final answer.

(b) Many candidates scored full marks on this question. As in part **(a)**, candidates who worked in degrees often lost accuracy in their final answer. The most straightforward method was to subtract the area of sector OCD from the area of the triangle OAB, using $\frac{1}{2}r^2\theta$ and $\frac{1}{2}ab\sin C$ respectively.

Question 5

Many candidates scored no marks for this question because they started by evaluating S_{150} and S_{400}. As highlighted in the General Comments, candidates are advised to read the question carefully before beginning a response.

Candidates who attempted to find the correct summations that needed to be subtracted, or the first and last terms of the new progression, were able to score the majority of the marks available even if they arrived at the wrong answers.

Question 6

(a) Many candidates obtained the correct coordinates in terms of r, but a large number had the coordinates in the wrong order.

(b) Most candidates used their coordinates from part **(a)** and Pythagoras' theorem to arrive at a quadratic equation. Some candidates used a calculator to solve their quadratic equation and, as a result, did not clearly show all the necessary working needed in their response to receive full credit. Almost all candidates correctly rejected the negative solution because r represents a length.

(c) Many candidates realised that the tangent would be perpendicular to the radius and used that information to find the gradient of the radius.

Question 7

(a) Most candidates, sensibly, started with the left-hand side of the given identity and immediately replaced $tan^2\theta$ with $\dfrac{\sin^2\theta}{\cos^2\theta}$. They then either replaced $\cos^2\theta$ with $1 - \sin^2\theta$ and collected the terms into a single fraction, or they converted to a single fraction first. To gain full marks in a *'show that'* question it is necessary for candidates to show each step clearly and to be accurate in their algebra. Candidates who started with the right-hand side of the identity were often able to obtain only the first two marks.

(b) Most candidates correctly used the result from part **(a)** and arrived at a quartic equation which they recognised as a quadratic in $\sin^2\theta$. One of the solutions to the correct quadratic was greater than 1 and could therefore be discounted. The other solution for $\sin^2\theta$ should have produced two values for $\sin\theta$, which many candidates did not realise and consequently did not find all three solutions.

Question 8

(a) The majority of candidates used correct expressions for the second term and sum to infinity of a geometric progression. Elimination of *a* to form a quadratic equation in *r* was the more straightforward approach. The question asked candidates to find a single common ratio and so they needed to reject $r = \frac{3}{2}$ as it was invalid for a convergent progression. Unfortunately, a significant majority rejected $r = -\frac{1}{2}$, possibly because it is negative.

(b) Some candidates chose to subtract the sum of the first 9 terms from the sum to infinity, while others found a new first term (the 10th term of the original progression) and used that to find the new sum to infinity. Negative signs were often omitted. As the answer to this part was exact, this meant that candidates who gave the answer as a decimal and rounded it to three significant figures were unable to obtain the final mark.

Question 9

(a) Most candidates successfully evaluated the second derivative when $x = \frac{1}{2}$ and stated that the stationary point is a minimum because the second derivative is positive. Some candidates did not justify and conclude that the second derivative is positive and consequently were unable to gain full marks.

(b) Successful candidates began by integrating the second derivative to find $\frac{dy}{dx}$, and used the fact that $\frac{dy}{dx} = 0$ when $x = \frac{1}{2}$ to find the constant of integration. They then integrated again to find an equation for *y*, substituted $x = \frac{1}{2}$ and $y = 9$ to find the second constant of integration and hence found the equation of the curve.

A significant number of candidates incorrectly used $\frac{dy}{dx} = 9$ after the first integration, whilst others ignored the requirement for a constant of integration at that stage altogether.

Question 10

(a) Most candidates found this question straightforward and obtained full marks, although a significant minority missed the factor of $\frac{1}{3}$ obtained from integrating $(3x+4)^{\frac{1}{2}}$. It was necessary to show the substitution of both limits to justify a full and complete method was being used. Some candidates incorrectly assumed that using the lower limit of 0 after integration was unnecessary, overlooking the fact that this gave a non-zero value.

(b) Only the stronger candidates found this question straightforward. Finding the equation of the tangent at (7, 0) was required to find the area of triangle *OPQ*, either from integration of the tangent or from using its *x* and *y* intercepts. Some candidates omitted the step of subtracting their answer to part **(a)**. A more common error was that some candidates did not realise that finding $\frac{dy}{dx}$ was initially essential, and further progress was significantly hindered as a result.

Question 11

(a) It was clear that the majority of candidates understood the concept of composite functions and were able to correctly combine the given functions. However, in many candidates' responses, algebraic errors were present which resulted in incorrect answers or insufficient steps being shown in this *'show that'* question.

(b) In questions of this nature, where multiple transformations are to be applied, candidates are encouraged to apply them one at a time. This allows the order in which the candidate applies the transformations to be clearly expressed in their solution.

(c) This part of the question proved to be very demanding for most candidates. The format of their final answer to part **(b)** made this part more or less straightforward when looking for the maximum value of $h(x)$. Successful candidates' answers to part **(b)** contained $-8k(x-2)^2$ which is non-positive, so the rest of the function determined the highest value of the range. Other starting points could eventually result in the same inequality but only after significantly more work that required algebraic accuracy.

MATHEMATICS

Paper 9709/22
Pure Mathematics 2

There were too few candidates for a meaningful report to be produced.

MATHEMATICS

Paper 9709/32
Pure Mathematics 3

Key messages

Candidates need to:
- know all the log laws and how to apply these
- gauge the method that best suits a particular problem before commencing, as often numerous different approaches are possible
- ensure any formulae are used correctly such that that the remainder of the question can be completed without issue. For example, in **Question 4**, candidates did not always begin their response by stating the correct expansion for $\tan(x - 60°)$.
- set their work out such that one line follows another, as opposed to statements being disconnectedly organised across the answer space. A disorganised approach is more likely to lead to issues such as sign errors, the accidental omission of terms or omitting constants from outside bracketed expressions. All of these were common errors in **Question 10(b)** and **Question 11**.

Since candidates need to show working to justify their results, calculators should not be used in complex number questions, nor where an exact answer is required, such as in **Question 10(b)** and **Question 11**. Candidates need to understand that the calculator should be restricted to simple work, such as solving a quadratic equation or a pair of linear equations.

General comments

The standard of work on this paper was variable, with some of the short initial questions proving more challenging for candidates than usual; in particular, **Question (2)**, **Question 3(a)** and **Question 3(b)**. However, it was good to see many candidates showing all their working for some of the longer response questions, even if it contained errors; for example, **Question 2** and **Question 11**. Whilst different candidates will always display different levels of working, a good gauge of what is deemed adequate working can be found in past question papers and their mark respective schemes.

Comments on specific questions

Question 1

Many candidates were able to gain full marks by re-arranging to obtain $1 - e^{-2x} = e^{-3}$ before solving an equation of the form $e^{-2x} = a$ to find the solution. Some candidates did not obtain the last accuracy mark due to quoting the answer to three decimal places instead of the required four. A significant number of candidates struggled with this question, demonstrating a lack of knowledge of the rules of logarithms and how to apply them when solving equations. A common incorrect first step was to split the $\ln(1 - e^{-2x})$ term into $\ln 1 - \ln e^{-2x}$.

Question 2

Candidates were generally able to gain at least some of the marks available on this question but relatively few gained full marks. Candidates recognised that they had to differentiate an implicit function, and many could differentiate xy^2 correctly. However, many found it challenging to differentiate $\ln(x + 2y)$ correctly and therefore, unless they found the correct value for y at $x = 0$, gained no further marks. Of those who differentiated the function correctly, many did not achieve full marks due to mistakes when rearranging the terms to find $\frac{dy}{dx}$. A few candidates rearranged the original function to get $x+2y=e^{1-xy^2}$, which is of course still an implicit equation. A few candidates were able to gain marks following this rearrangement.

Question 3

(a) This proved to be a challenging question and relatively few candidates gained full marks. Many were able to identify the centre and obtained something of the form $|z \pm (-2+i)| \le 3$. A few candidates omitted the modulus signs, a number stated and equation instead of an inequality and some had the incorrect radius. Many did not state a second inequality at all. A common mistake was to use $\text{Im}(z) \le i$.

(b) This part also proved challenging for candidates, with many not attempting a solution at all and only very few gaining full marks. Of those that did score full marks, most found the coordinates of the correct point and then calculated the distance. An alternative method was to use Pythagoras' Theorem twice.

Question 4

This question was answered well, with many candidates either gaining full marks or just missing the final accuracy mark from stating the incorrect angles or including too many values for x in their answer. Common mistakes were to replace $2\cot x$ with $\frac{1}{2\tan x}$ or to use an incorrect trigonometric identity. As with **Question 2**, some candidates found it challenging to rearrange the terms to obtain the correct quadratic equation.

Question 5

Most candidates were able to achieve full marks on this question. Candidates were able to square $x + iy$ correctly and then equate the real and imaginary parts to -4 and $6\sqrt{5}$ respectively. They were then able to solve the simultaneous equations, finding the correct quartic and solving for the correct values for x and y. Those candidates who did not achieve full marks either did not give two solutions or gave $\sqrt{5}-3i$ rather than $-\sqrt{5}-3i$.

Question 6

Of those candidates who knew how to integrate $\sin^2 2\theta$, most went on to gain full marks. Some did not obtain the final accuracy mark as they did not fully simplify to obtain the final equation for x in terms of θ. Many candidates knew how to separate the variables correctly and were able to attempt an integration of $\frac{1}{\frac{1}{5}x+1}$, but a significant number of these candidates were unsure how to integrate $\sin^2 2\theta$ and so were unable to gain more than two marks.

Question 7

(a) The majority of candidates correctly differentiated the given expression and set this equal to 0. However, there were some who incorrectly expressed the equation of the curve in terms of p instead of x before differentiating. Most candidates knew it was then necessary to divide their equation by $\cos 2x$ to get an equation in terms of $\tan 2x$, although there were a few candidates who attempted to write $\sin 2x$ and $\cos 2x$ in terms of a single angle instead. Many candidates went on to correctly obtain the given equation; it should be noted that, in a 'show that' question, every step of the solution should be clearly shown to gain the final mark.

(b) This question was usually done well with a variety of approaches seen. Most candidates approached this by obtaining an equation equal to 0, substituting 0.5 and 0.7 for x, showing that the equation was positive for one value and negative for the other, therefore showing a root lies between these two values. Other candidates substituted 0.5 and 0.7 for p into $\frac{1}{2}\tan^{-1}\left(\frac{3}{2p}\right)$ and showed that the first was greater than 0.5 and the second was less than 0.7. Occasionally, candidates went back to their expression for the gradient in part **(a)** and successfully showed that this was positive and negative for these two values of x.

(c) This question was extremely well attempted, and candidates usually gained full marks. However, some candidates used an incorrect iterative formula, were working in degrees rather than radians or did not give the required number of decimal places in their solutions.

Question 8

(a) Most candidates understood that to show that the lines are skew they need to show both that the two lines do not intersect and that the two lines are not parallel. Although the correct method was often seen, there were many arithmetic errors in attempts to show that the two lines do not intersect. Far less common was a correct method to show that the two lines were not parallel. Many solutions did not attempt to show this at all, with another large proportion of solutions showing insufficient detail in proving that the lines are not parallel. Candidates need to clearly show that they understand which part of the line equation corresponds to the direction vector. Candidates also need to be aware that just stating that the two direction vectors are not equal is not enough to show that they are not parallel. Following the identification of the direction vectors, it must be clearly shown or stated that these two vectors are not scalar multiples of each other.

(b) This question was very well attempted, with correct answers usually seen in degrees rather than radians. Candidates who were not so successful in this part tended not to be using the direction vectors of the lines but instead using the position vectors of the two points found in part **(a)**.

Question 9

(a) This question was usually done very well, but as in the previous question, there were many arithmetic errors in the attempts to find the values of a and b. Candidates who substituted $x = 3$ into $p(x)$ and $p'(x)$ and equated to 0 and 72 tended to be more successful than those who divided $p(x)$ and $p'(x)$ by $(x - 3)$ and went on to equate the remainders to 0 and 72.

(b) There were some candidates who clearly used their calculator to solve the resulting cubic equation found in part **(a)**. Candidates are reminded that the rubric clearly states they must show all necessary working clearly; no marks will be given for unsupported answers from a calculator. Candidates needed to show division of the cubic by $x - 3$, which was the most popular method seen, or to equate the cubic to a quadratic factor multiplied by $x - 3$ and then find this resulting quadratic factor. There were many candidates who, having found the correct quadratic factor, did not go on to factorise this.

(c) Very few candidates were fully successful in this part, perhaps because it relied on correct answers in both parts **(a)** and **(b)**. Candidates are expected to know the shape of a cubic graph, although it was rare to see a sketch which would have helped them to find the correct inequalities for x in this part.

Question 10

(a) This question was usually well attempted. Again, there were often arithmetic errors seen in these solutions, but those candidates who knew the correct form of partial fractions required in this case usually went on to find correct values for all three coefficients.

(b) Fully correct solutions depended on the correct coefficients being found in part **(a)**, although even with one or more incorrect coefficient, candidates could still gain all but one of the marks available. Integrating $\frac{-3}{1+x}$ caused very few problems, apart from the occasional sign or constant error. Integrating $\frac{-4x+6}{4+x^2}$ caused more problems, usually as candidates did not realise they needed to split this into two integrable fractions. Many candidates realised the correct form for the integral of $\frac{6}{4+x^2}$ but tried to give an answer for this which included the $-4x$ in the numerator also. The final challenge was to give the answer in the required form; many candidates, having reached

$\frac{3}{4}\pi + \ln\frac{1}{108}$, did not rewrite this as $\frac{3}{4}\pi - \ln 108$. Another basic error was to often see $3\ln 3$ written as $\ln 9$.

Question 11

Almost all candidates knew they needed to use integration by parts. Some candidates, having correctly integrated by parts once, did not then realise they needed to integrate by parts a second time and did not appear to recognise the structure of this second problem. Candidates who understood they needed to integrate by parts twice usually did so successfully, aside from the occasional sign or coefficient error. Presenting work clearly and carefully organising work written in different areas of the page may lead to fewer of this type of error.

MATHEMATICS

Paper 9709/42
Mechanics

Key messages

- When answering questions involving any system of forces, a well annotated force diagram could help candidates to make sure that they include all relevant terms with correct signs when forming either an equilibrium situation or a Newton's Law equation. Such a diagram would have been particularly useful here in **Questions 4** and **6**.

- It was noted some candidates do not substitute given values in a question but have an algebraic expression. The only indication that values have been used is in a final answer. If the final answer is incorrect, examiners have no indication that correct values have been used in what appears to be a correct algebraic expression. It would be good practice to have at least an expression with all values substituted just in case some error is made in using a calculator correctly.

- In questions where the request is to 'show that', sufficient detail needs to be given so that examiners can be certain the result has been shown and not just quoted.

General comments

Many candidates were well prepared for the demand of this paper, with many questions attempted well.

Candidates at all levels were able to show their knowledge of the subject. **Questions 1**, **3(a)**, **4(a)**, **5(a)** and **5(b)** were found to be the most accessible questions whilst **Questions 4(b)** and **5(c)** proved to be the most challenging.

One of the rubric points on the front cover of the question paper is to give non-exact numerical answers correct to 3 significant figures and it was noted that almost all candidates followed this instruction. However, some answers on this paper are exact and so should not be approximated, notably the answer to **2(b)** is 17.25 seconds, but a final answer of 17.3 seconds was often seen.

Comments on specific questions

Question 1

This question was answered well by many candidates. The most common approach being to resolve forces in two directions and using the resulting equations to solve for θ and X.

Question 2

(a) Most candidates looked at the final stage and used the equation $v^2 = u^2 + 2as$ with $v = 0$, $a = -2$ and $s = 16$ to find V. Some tried to use the earlier stages of the motion but often failed to successfully find V.

(b) The majority of candidates found the correct total time. Some who failed to do this had errors in finding the acceleration in the first stage of the motion. Those who had found an incorrect V in part **2(a)** were still able to score some marks here.

Question 3

(a) Most found the speed the with correct use of the equation $P = Fv$. Occasionally, errors were made with the number of zeros in either the power or the driving force but the majority of candidates scored well here.

(b) This question caused some confusion with candidates finding, incorrectly, the speed of $\frac{300}{50}$ ms^{-1} for the upward motion even though the question stated that the aircraft ascends while maintaining the same speed. Also, some included a KE term in their energy equation even though this was the same at the beginning and end of the motion.

Question 4

(a) This question was answered well by the majority of candidates, with only a minority of candidates having a sign error in one of the equations of motion.

(b) Candidates found this to be one of the most challenging questions on the paper. Most found the speed of particle *B* at the time when particle *A* reached the ground. The majority of candidates then thought that *B* still had $1.2–x$ metres to travel to reach the highest point. However, as *A* has moved a distance x metres downwards, then B is now $2x$ metres above the ground and so is a distance $1.2 – 2x$ metres from its highest point. Examiners commented that a correct answer was rarely seen.

Question 5

(a) This question was answered well by the majority of candidates. It is necessary to use the principle of conservation of momentum for the collision between particle *P* and particle *Q*.

(b) Again, this part was well done by most. The method used was to apply the principle that the linear momentum of Q can be equated to the linear momentum of the coalesced particle QR. This gave the required result of 0.75 ms^{-1} which was obtained by the majority of candidates.

(c) This request was found difficult by a significant number of candidates. The most common error seen by examiners was for the use of constant acceleration formulas even though none of the particles were accelerating. The majority of candidates found the time taken for particle *Q* to reach particle *R*. Also, many found the distance between *P* and *Q* at the instant the collision between *Q* and *R* took place. However, few candidates completed the solution beyond this point.

Question 6

(a) Overall, this was well done by most candidates. The common errors seen were incorrect signs when setting up the equation of motion, or the omission of the acceleration due to gravity in the weight component term.

(b) Many candidates made a reasonable attempt at this request by forming two equations for when the particle was on the point of moving down and up the plane. However, without a relevant diagram, a significant number of candidates had a sign error with either the 10 N or the weight component so had a wrong value of μ.

Question 7

The concept of variable acceleration being related to calculus rather than constant acceleration continues to be well understood, so this question was a good source of marks for most the candidates.

(a) The majority of candidates were able to differentiate and integrate correctly and go on to use the given information at $t = 1$ to produce two correct equations. The majority then solved these simultaneous equations correctly to show the given values. However, a minority of candidates showed no working and just stated the given values and hence did not earn full credit.

(b) The majority of candidates correctly found the times at which $a = 0$ and $v = 0$ and used these correctly to find the required distance. The main error seen by examiners was to find the difference between the two times found and use this to find a distance.

MATHEMATICS

Paper 9709/52
Probability & Statistics 1

Key messages

Candidates should be aware of the need to communicate their method clearly. Simply stating values often does not provide sufficient evidence of the calculation undertaken, especially when there are errors earlier in the solution. The use of algebra to communicate processes is anticipated at this level and enables candidates to review their method effectively. When errors are corrected, candidates would be well advised to cross through and replace the term. It is extremely difficult to interpret accurately terms that are overwritten.

Candidates should state non-exact answers to three significant figures. Exact answers must be stated exactly. Candidates should have a clear understanding of how significant figures work for decimal values less than 1. It is important that candidates realise the need to work to at least four significant figures throughout to justify a final value to three significant figures. Many candidates rounded prematurely in normal approximation questions, which caused them to identify incorrect values from the normal distribution tables.

It is an inefficient use of time to convert an exact fractional value to an inexact decimal equivalent, as there is no requirement for probabilities to be stated as a decimal.

There was a noticeable improvement in the interpretation of success criteria, which is encouraging as this is an essential skill for this component. Candidates are well advised to include this within their preparation.

General comments

Although many well-structured responses were seen, some candidates made it difficult to follow their solutions by not using the response space in a clear manner. The best solutions often included some simple notation to clarify the process that was being used.

The use of simple sketches and diagrams can help to clarify both context and information provided. These were often seen in successful solutions. The lack of accuracy in constructing the cumulative frequency graph in **Question 3** was disappointing, as data values were often not explicitly stated and the scales chosen for axes did not use the grid effectively.

Sufficient time seems to have been available for candidates to complete all the work they were able to, although some candidates may not have managed their time effectively. A few candidates did not appear to have prepared well for some topics, in particular when more than one technique was required within a solution. Many good solutions were seen for **Questions 1** and **4**. The context in **Questions 5** and **6** was found to be challenging for many.

Comments on specific questions

Question 1

This probability question was accessible to the majority of candidates. Some tree diagrams were seen, but most solutions used a method involving listing scenarios.

(a) Many good solutions involving the clear identification of which of the coins was biased were seen. Weaker solutions simply stated the unsimplified calculation, which was acceptable as the coins could be identified in the context by the relevant probability, providing the coins were in a consistent order. If coins were not in a consistent order, there needed to be clear identification of the scenarios because of the 'show' requirement.

(b) The majority of solutions were fully correct, with the probability distribution table clearly shown and with some supporting work. Very few candidates had probabilities that did not add to 1. The most common error was to omit $P(X = 0)$ from the table. A small number of candidates presented decimal answers, and these were not always the accurate conversion of a stated fraction.

(c) Candidates who had an appropriate probability distribution table were usually able to access this standard process to calculate the variance. The best solutions stated an unsimplified variance calculation and then used a calculator efficiently to evaluate. Some weaker candidates also calculated $E(X)$ which was stated in the question, not always accurately. It is important that candidates are aware that problems involving syllabus content do require clear supporting work, and simply stating the variance value will not gain full credit. It is important that non-exact answers are given to 3 significant figures, as 0.57 was seen frequently, not always with a more accurate value present.

Question 2

Almost all candidates attempted this standard problem involving discrete random variables. The interpretation of the success criteria was more accurate than in previous sessions.

(a) Most candidates found the context of this part challenging. The actual selection of the 3 random customers was not considered by the majority of candidates, who assumed that the order was not relevant and simply multiplied the 3 probabilities. The best solutions identified that the 3 customers could be ordered in 3! ways and multiplied appropriately. Weaker solutions used a listing approach to identify the ways the customers could be selected. A common error was to multiply by 3, although no reasoning was presented to justify this.

(b) The majority of candidates correctly identified the binomial distribution, correctly interpreting the success criteria and using the efficient approach of $1 – P(10,11,12)$. The very small number of candidates who used the alternative approach were generally successful. The most common error was to use $P(F) = 0.70$ throughout the method. A small number of candidates omitted brackets within their calculation, which is not acceptable at this level

(c) This question was a standard textbook example of the use of the normal approximation for the binomial distribution. It was surprising that the question was not attempted by 7 per cent of the candidates. Good solutions clearly identified the mean and variance when calculated, substituted these appropriately into the normal standardisation formula, using a continuity correction as the data was discrete, and found the required probability area – often with a simple sketch of the normal distribution curve to help interpret the success criteria. A small number of candidates stated that $\sigma = npq$ which is not correct. Again, almost all candidates interpreted the success criteria appropriately with few candidates finding the complement of the required probability.

Question 3

(a) The quality of the cumulative frequency graphs seen was disappointing. The best solutions used a scale which enabled points to be plotted accurately, for example 1 cm = 20 for the cumulative frequency axis. A large number of solutions did not state the cumulative frequencies, and then used a scale where it was not possible to identify the values accurately. As in all data representations, the axes do need to be fully labelled, including units where appropriate. Good candidates did interpret the classes appropriately, e.g. 5 – 9 was equivalent to $4.5 < l \leq 9.5$, and plotted values at the upper boundary. Many candidates plotted at the integer value stated and weaker candidates at class mid-values. Candidates would be well advised to use + or × when plotting points for clarity. Where graphs were joined to the length axis, it was frequently at (0, 0) rather than (4.5, 0) as shown in the data table. Encouragingly, very few graphs using line segments were noted.

(b) It was encouraging that more candidates were clearly indicating their use of the graph as required in the question than in previous years. The best candidates used a ruler to draw appropriate lines to read accurately from their cumulative frequency graph. The most common error, as in previous years, was misinterpreting the success criteria and finding the length of the 38th percentile rather than the 62nd percentile.

(c) Many good solutions to this grouped mean problem were seen. The best identified the midpoints, often by the data table at the start of the question, showed an appropriate calculation and stated their final answer accurately. Even though the question requires an estimated mean to be found, as the final answer was an 'exact value', it was inappropriate to round to 3 significant figures. There were occasional errors with the mid-value of the 30 – 39 class and a small number of candidates appear to have assumed the first class was 1 – 9. The use of a 'table' approach to this question is acceptable, as it is an efficient way of communicating the process. Where candidates do all their work by the original data table, they do need to ensure that their final answer is presented in the appropriate part of the question.

Question 4

The context of this probability question was accessibly by the majority of candidates. Solutions which used either a tree diagram, or a logical way of listing scenarios, were often the most successful. Almost all candidates recognised that the context was without replacement.

(a) Almost all candidates identified clearly the three possible scenarios that fulfilled the success criteria, stated the unsimplified calculations and evaluated accurately. Candidates should be aware that there is no benefit in converting an accurate fraction answer to an inexact decimal answer. The most common error was not evaluating the expressions accurately. A small number of candidates misinterpreted the question and found the probability of picking one white, one black and one silver car.

(b) The context was found more challenging by candidates for this question, who often failed to appreciate that there were different arrangements of the outcomes required. Good solutions often used a listing approach to identify the possible outcomes, with tree diagrams being present occasionally. The appropriate probability calculations were then stated and evaluated accurately. A significant number of solutions assumed that there was only one way of arranging each scenario. Another common error was to assume that the number of arrangements for WBB would be the same for WWB and WBS, not recognising the impact of the repeated colour in the first value. A small number of candidates only calculated the total of possible outcomes, rather than finding the required probability.

Question 5

Some candidates appeared to have assumed that this was a normal distribution question and did not appreciate that parts **(c)** and **(d)** linked to the geometric distribution. There were an unexpected number of arithmetical errors in this question.

(a) The best solutions often included a simple sketch of the normal distribution curve to help interpret the success criteria. Almost all candidates were able to apply the normal standardisation formula accurately at least once. A few candidates assumed that the probability area would be symmetrical and used this fact to inaccurately calculate their probability area.

(b) This standard question required a variable in the standardisation formula to be found. Candidates should be aware that if the probability is a 'critical value' then the appropriate z-value from the critical value section of the normal tables must be used, as this is a component skill. The most frequent error was to equate the standardisation formula to a probability, 0.5398 being the most common value seen. Some candidates introduced a continuity correction of 0.5, which would be inappropriate for a continuity correction in this context, if weight was not a continuous variable.

(c) The context of the question changed in the body of the paper to being a geometric distribution. As such, this question is to simply calculate P(5), which is a textbook skill. Good solutions stated the unsimplified expression before evaluating. Some candidates misinterpreted the success criteria and calculated P(6). Approximately 10 per cent of candidates did not attempt this question, which was unexpected.

(d) Approximately 15 per cent of candidates did not attempt this basic geometric distribution question. The most successful approach was to calculate P(1,2,3,4), stating each unsimplified term and then evaluating the sum. More able students often attempted the more efficient $1 - 0.9^4$ process, but failed to interpret the success criteria appropriately, with $1 - 0.9^5$ the most common error.

Question 6

The context of this permutation and combination question appeared to be more accessible. Candidates who identified the scenarios they were considering in a clear and logical manner were often more successful. A significant number of candidates made little or no progress with many parts of the question.

(a) This was a standard textbook style question, which was successfully answered by the majority of candidates. The best solutions stated clearly the unevaluated expression. The most common error was to omit one or more of the repeated colours from the denominator. A small number of candidates did not state the exact value before rounding to 3 significant figures.

(b) Both standard approaches were seen successfully completed by many candidates. The best solutions did provide some explanation to explain the calculation, often in the form of a simple diagram where the 3 blue books are treated as a single unit.

(c) Some good solutions were seen, but many candidates failed to appreciate that the repeated colours were still present and needed to eliminated using values on the denominator. Again, solutions where a possible acceptable solution was illustrated, e.g. ^ Y ^ ^ ^ ^ Y ^ ^ ^, were frequently more successful in identifying all possible arrangements. Weaker solutions often included additional factors for arranging the Ys, which should then have been eliminated as the Ys were repeated and not distinguishable.

(d) A good start was made by many candidates in identifying the possible scenarios that fulfilled the success criteria. Good solutions then used combinations to determine the number of possible selections available. Some weaker solutions used a ‘counting’ approach for how the books could be selected, but this was often inaccurate. A significant number of candidates made no attempt, while some weaker candidates gained credit for using the appropriate process to sum values for the possible scenarios.

MATHEMATICS

Paper 9709/62
Probability & Statistics 2

Key messages

- In all questions, sufficient method must be shown to justify answers.
- When the answer is given in the question, candidates need to show, convincingly, their full and clear working leading to the answer.
- Candidates should be encouraged to read questions carefully and to reread the question when they have completed their solution to ensure that the question has been fully answered.
- Answers should be given to 3 significant figure accuracy, unless stated otherwise. This means keeping 4 significant figures or more in the working leading to the answer.
- All working should be done in the correct question space of the answer booklet. If a candidate does not have enough space to complete their answer, they should use the additional page and label it clearly with the question number.
- Conclusions to Hypothesis tests should be written in context and with a level of uncertainty in the language used.

General comments

This was a well attempted paper, with many candidates performing well, although it must also be noted that there were some candidates who were not fully prepared for the demands of the examination. Questions that were well attempted were **Questions 1**, **2(a)**, **3(a)** and **6(b)** whereas **Questions 2(c)** and **3(c)** proved to be more demanding.

Timing did not appear to be an issue.

The following comments highlight particular common errors made but equally there were some very good, accurate answers seen as well.

Comments on specific questions

Question 1

This was a good, well answered, opening question for the majority of candidates. There were occasional errors seen when calculating the variance (for example subtracting 16 and 9 rather than adding) and some candidates found the wrong probability area (i.e., 1 – 0.579 rather than 0.579).

Question 2

(a) This part was well attempted; candidates knew the formulae to use to find the unbiased estimates of E(T) and Var(T) and there was very little confusion shown between the alternative formulae for the variance.

(b) The hypothesis test was reasonably well attempted, with the null and alternative hypotheses generally correctly stated. The most common error noted was the omission of $\sqrt{60}$ in the denominator when standardising, and premature approximation of 14.44 to 14.4 caused a loss of accuracy for some candidates. After standardising to find z, a clear comparison was required and a common error for those who compared z values was to compare with 2.326 rather than 2.054. It was important that the conclusion was written as a non-definite statement and in the context of the question. For example, ‘The mean time has decreased’ would not be acceptable as it is a definite statement, and ‘There is sufficient evidence to suggest that the mean has decreased’ would also not be acceptable as there is no context.

(c) This part was not well answered and showed a basic lack of understanding of the Central Limit Theorem by a large number of candidates. There were some candidates who realised that the underlying distribution was not given, but failed to qualify that this was the ‘population’ distribution; the distribution being referred to must be clear, so merely saying ‘the distribution is unknown’ is not precise enough.

Question 3

(a) This question was well attempted with most candidates correctly calculating 1- P(0, 1, 2). The most common errors were to calculate P(0, 1, 2) rather than 1– P(0, 1, 2) or to calculate 1– P(0, 1, 2, 3). It was important that the full Poisson expression was shown in order to fully justify the answer; candidates are advised to show all necessary working.

(b) Many candidates correctly used $\lambda = 4.5$, though some used $\lambda = 1.5$. P(3, 4, 5) was generally calculated with or without the correct value for λ. Again, the full Poisson expression was required.

(c) Many candidates found this question demanding. N(1.5 n, 1.5 n) was required, followed by standardising (with a continuity correction) and equating to 1.761 to form an equation in n. Errors included using an incorrect Normal distribution, equating the standardised expression to 0.0391 rather than 1.761, omitting or using an incorrect continuity correction, and dividing the denominator by $\sqrt{n}$. The equation formed should have been a quadratic; algebraic errors in rearranging this equation were commonly seen. Only one value of n was valid.

Question 4

(a) (i) Many candidates realised that the area under the pdf was equal to 1. The area could be found by finding the area of a trapezium, a rectangle plus a triangle, or a larger rectangle minus a triangle, or even by finding the equation of the line and using integration. All methods were seen, though the final method was not generally successful and was not an efficient method here. It was important that all working was shown, particularly as this was a ‘show that’ question. There were some candidates who did not know how to approach the question.

(ii) The equation of the straight line was found by many candidates and most knew that E(X) was found by integrating xf (x). Errors included an incorrect equation of the line, incorrect integration and incorrect limits.

(b) Many candidates set up a correct expression, but errors were made with limits; most candidates integrated cos t between –c and +c and equated to $\frac{1}{2}$, others integrated between 0 and c and equated to $\frac{1}{4}$ and other similar correct expressions were seen. However, some candidates were unable to match the correct limits with the correct area. After integrating and substituting limits many candidates reached a correct expression involving sin(c) and sin(–c) but were unable to simplify this expression to find c. A few candidates incorrectly stated that c was 30.

Question 5

(a) Many candidates correctly calculated $P(X \leq 3)$ using B(35, 0.2). It was important that the full Binomial expression was shown here. An acceptable significance level was then required, and this was often successfully given. In general candidates made a good attempt at this question.

(b) Many candidates realised that the probability of a Type I error was 0.0605 as calculated in **5(a)**.

(c) Finding the probability of a Type II error was not quite so well done. Many candidates correctly used B(35, 0.05) and used this to find the required probability, but some attempted calculations using B(35, 0.2) or a Normal distribution. Questions on Type II errors have not always been answered well in the past, so it was pleasing to note a good number of correct attempts here.

Question 6

(a) There were mixed responses here. Many candidates correctly stated that they did not agree as the sample would be biased or not be representative of all students in the school. However, some candidates thought the choice was good because these were the students who knew most about the sports facilities failing to note that the investigation was to discover the views of all students in the school.

(b) This was well attempted. Most candidates successful found the required confidence interval. A common error was to use an incorrect z value, and some candidates centred the interval around 45 instead of 45/60.

(c) A reasonable attempt was made by many candidates here, though a common error, after successfully finding 1.288 (from 2.576 divided by 2), was to give an answer of 90.12%.