ASSESSMENT and
QUALIFICATIONS
ALLIANCE

General Certificate of Education

Mathematics – Statistics

SPECIMEN UNITS AND MARK SCHEMES

ADVANCED SUBSIDIARY MATHEMATICS (5361)
ADVANCED SUBSIDIARY PURE MATHEMATICS (5366)
ADVANCED SUBSIDIARY FURTHER MATHEMATICS (5371)

ADVANCED MATHEMATICS (6361)
ADVANCED PURE MATHEMATICS (6366)
ADVANCED FURTHER MATHEMATICS (6371)

General Certificate of Education
Specimen Unit
Advanced Subsidiary Examination

MATHEMATICS **MS1A**
Unit Statistics 1A

In addition to this paper you will require:
- an 8-page answer book;
- the AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed: 1 hour 15 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS1A.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The maximum mark for this paper is 60.
- Mark allocations are shown in brackets.

Advice
- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

1 Ten per cent of coloured beads used in costume jewellery are orange.

(a) Find the probability that in a string of 40 beads, 4 or fewer beads are orange. *(3 marks)*

(b) Calculate the probability that in a string of 35 beads, exactly 2 beads are orange. *(3 marks)*

(c) State **one** assumption that you have made in answering parts (a) and (b). *(1 mark)*

2 The weights of bags of red gravel may be modelled by a normal distribution with mean 25.8 kg and standard deviation 0.5 kg.

(a) Determine the probability that a randomly selected bag of red gravel will weigh less than 25 kg. *(3 marks)*

(b) Determine, to two decimal places, the weight exceeded by 10% of bags. *(4 marks)*

3 (a) A sample of people, who commute regularly from a town in Surrey into London, was asked for an estimate of the time taken on their most recent journey. The replies are summarised below.

Time (minutes)	**Frequency**
35 -	12
45 -	54
55 -	68
65 -	41
85 - 105	23

Calculate estimates of the mean and the standard deviation of these times. *(5 marks)*

(b) A sample of people who commute regularly from a town in Essex into London was also asked for an estimate of the time taken on their most recent journey. Their replies had a mean of 64 minutes and a standard deviation of 21 minutes. Compare, briefly, the journey times estimated by commuters from the two towns. *(2 marks)*

(c) Give **two** reasons why the data presented in parts (a) and (b) may not adequately represent typical commuting times from the two towns. *(2 marks)*

4 A cricket team meets for fielding practice. One exercise consists of a cricket ball being thrown at different heights, speeds and angles to one side of a fielder who tries to catch it using one hand.

Each member of the team attempts 25 catches with each hand. The number of successful catches are given in the following table.

Fielder	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**
Left hand	11	13	9	17	21	16	14	8	19	19	20
Right hand	18	17	20	22	14	19	21	15	10	24	23

(a) Calculate the value of the product moment correlation between the number of catches with the left hand and the number of catches with the right hand. *(3 marks)*

(b) Comment on the performance of fielders **E** and **I**. *(2 marks)*

(c) When fielders **E** and **I** are omitted from the calculation, the value of the product moment correlation coefficient between the number of left-handed catches and the number of right-handed catches is 0.812, correct to three decimal places. Comment on this value and the value you calculated in part (a) *(2 marks)*

5 Pencils produced on a certain machine have lengths, in millimetres, which are distributed with a mean of μ and a standard deviation of 3. A random sample of 90 pencils was taken and the length of each pencil measured. The mean length was found to be 178.5 millimetres.

(a) Construct a 99% confidence interval for μ. *(5 marks)*

(b) State why, in answering part (a), it is not necessary to assume that the length of pencils are normally distributed. *(2 marks)*

TURN OVER FOR THE NEXT QUESTION

Turn over ►

6 Last year the employees of a firm either received no pay rise, a small pay rise or a large pay rise. The following table shows the number in each category, classified by whether they were weekly paid or monthly paid.

	No pay rise	**Small pay rise**	**Large pay rise**
Weekly Paid	25	85	5
Monthly paid	4	8	23

A tax inspector decides to investigate the tax affairs of an employee selected at random.

D is the event that a weekly paid employee is selected.
E is the event that an employee who received no pay rise is selected.
E' is the event not E.

(a) Find the value of:

(i) $\mathrm{P}(D)$;

(ii) $\mathrm{P}(D \mid E)$;

(iii) $\mathrm{P}(D \cap E')$. *(5 marks)*

(b) The tax inspector now decides to select three employees. Find the probability that they are all weekly paid if:

(i) one is selected at random from those who had no pay rise, one from those who had a small pay rise and one from those who had a large pay rise; *(3 marks)*

(ii) they are selected at random (without replacement) from all the employees of the firm. *(2 marks)*

7 **[**A sheet of graph paper is provided for use in this question.**]**

Andrew (**A**), Charles (**C**) and Edward (**E**) are employed by the Palace Hotel. Each is responsible for one floor of the building and their duties include cleaning the bedrooms. The number of bedrooms occupied on each floor varies from day to day.

The following table shows 10 observations of the number, x, of bedrooms to be cleaned and the time taken, y minutes, to carry out the cleaning. The employee carrying out the cleaning is also indicated.

Employee	**A**	**C**	**E**	**E**	**C**	**A**	**A**	**E**	**C**	**C**
x	8	22	12	24	19	14	22	16	10	21
y	110	211	132	257	184	165	248	171	97	196

(a) Plot a scatter diagram of the data. Identify the employee by labelling each point. *(3 marks)*

(b) Calculate the equation of the regression line of y on x. Draw the line on your scatter diagram. *(6 marks)*

(c) Calculate the residuals for the three observations when Andrew did the cleaning. *(3 marks)*

(d) Comment on the times taken by Andrew to carry out his cleaning. *(1 mark)*

END OF QUESTIONS

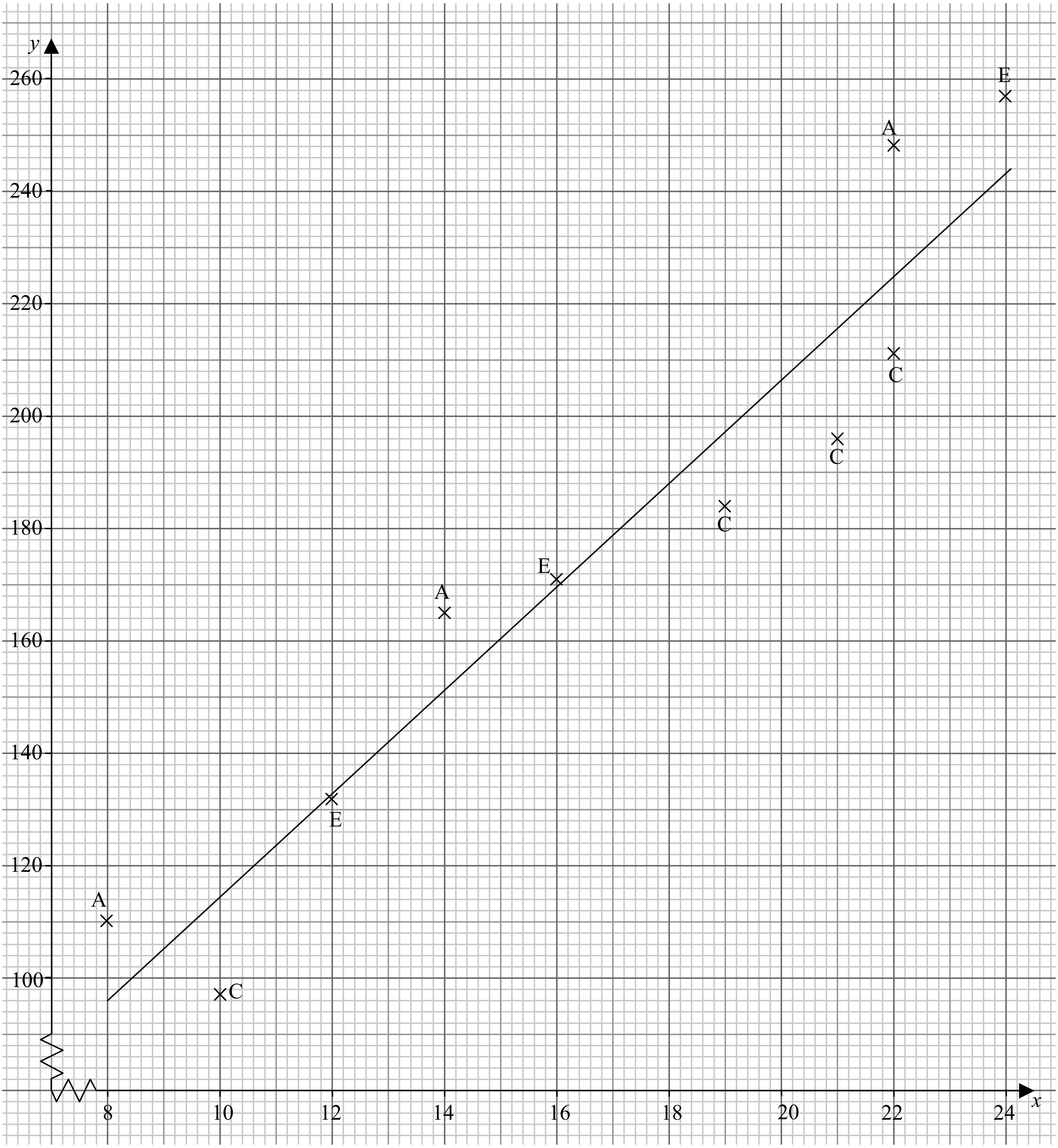
Question	Solution	Marks	Total	Comments
1(a)	Binomial $n = 40$ $p = 0.1$	B1B1		
	P(4 or fewer) = 0.629	B1	3	
(b)	$P(2) = (35 \times 34/2) \times 0.1^2 \times 0.9^{33}$	B1M1		
	= 0.184	A1	3	0.183 – 0.184
(c)	Beads selected randomly/independently	E1	1	
	Total		**7**	
2(a)	$z = (25 - 25.8)/0.5 = -1.6$	M1		
	Probability less than 25kg = 1 – 0.94520	M1		
	= 0.0548	A1	3	
(b)	z = 1.2816	B1		
	Weight exceeded by 10% of bags	M1m1		
	$25.8 + 1.2816 \times 0.5 = 26.44$	A1	4	
	Total		**7**	
3(a)	Class mid-mark / Frequency 40 / 12 50 / 54 60 / 68 75 / 41 95 / 23	M1		Allow m1A1 for mean and s.d. if method shown. 63.2 (63.1 – 63.3)
	$\bar{x} = 63.2$ $s = 15.2$	A2 A2	5	15.2 (15.0 – 15.3)
(b)	Journeys from Surrey have similar duration, on average, but are less variable than those from Essex.	E1 E1	2	
(c)	People asked may not be representative. Times are estimated not measured.	E1 E1	2	Or any other sensible comments e.g. journey time not defined , weather conditions may be extreme etc
	Total		**9**	

Question	Solution	Marks	Total	Comments
4(a)	0.0477	B3	3	0.047 – 0.048 allow M2 A1 if method shown
(b)	E and I held more catches with left than with right hand - all others held more with right than left.	E1 E1	2	
(c)	Correlation coefficient of 0.812 suggests that those who caught a lot of catches with one hand also caught a lot of catches with the other. When E and I (possibly left handers) are included the correlation coefficient of 0.0477 suggests no association between the number of catches with each hand.	E1 E1	2	
	Total		**7**	
5(a)	99% confidence interval for mean $178.5 \pm 2.5758 \times 3/\sqrt{90}$ 178.5 ± 0.8145 177.69 – 179.31	B1M1 m2 A1	5	
(b)	Sample is large. Sample mean may be assumed to be Normally distributed by Central Limit Theorem.	E1 E1	2	
	Total		**7**	
6(a)(i)	115/150 = 0.767	B1	1	acf
(ii)	25/29 = 0.862	M1A1	2	acf
(iii)	90/150 = 0.6	M1A1	2	acf
(b)(i)	$25/29 \times 85/93 \times 5/28 = 0.141$	M1 M1A1	3	0.14 – 0.141
(ii)	$115/150 \times 114/149 \times 113/148 = 0.448$	M1 A1	2	
	Total		**10**	

Question	Solution	Marks	Total	Comments
7(a)	See graph on next page	M1 A1 B1	3	
(b)	$y = 22.8 + 9.19x$	B2 B2		22.7 – 22.8 9.18 – 9.2 Allow M1 A1 for a and b if method shown
	$x = 8$ $y = 96.3$ $x = 23$ $y = 234.1$	M1A1	6	+ line on graph
(c)	Residuals $110 - 22.77 - 9.186 \times 8 = 13.7$ $165 - 22.77 - 9.186 \times 14 = 13.6$ $248 - 22.77 - 9.186 \times 22 = 23.1$	M1 A1 A1	 3	M1 method - ignore sign, allow read from graph A1 one correct - ignore sign 13.7 (13 – 14) 13.6 (13 – 14) 23.1 (22 – 24) A1 all correct, including sign
(d)	Andrew appears to be slowest (all residuals positive / all times longer than predicted by regression line)	E1	1	
	Total		**13**	
	TOTAL		**60**	

MS1A (cont)

Graph for Question 7



General Certificate of Education
Specimen Unit
Advanced Subsidiary Examination

MATHEMATICS **MS1B**
Unit Statistics 1B

In addition to this paper you will require:
- an 8-page answer book;
- the AQA booklet of formulae and statistical tables;
- a sheet of graph paper for use in Question 6;
- a ruler.

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS1B.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The maximum mark for this paper is 75.
- Mark allocations are shown in brackets.

Advice
- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

1 Jeremy sells a magazine which is produced in order to raise money for homeless people. The probability of making a sale is 0.09 for each person he approaches.

(a) Given that he approaches 40 people, find the probability that he will make:

(i) 2 or fewer sales; *(3 marks)*

(ii) more than 5 sales. *(2 marks)*

(b) Find the probability that he will make two sales given that he approaches 16 people. *(3 marks)*

(c) State **one** assumption you have made in answering parts (a) and (b). *(1 mark)*

2 (a) A sample of people, who commute regularly from a town in Surrey into London, was asked for an estimate of the time taken on their most recent journey. The replies are summarised below.

Time (minutes)	**Frequency**
35-	12
45-	54
55-	68
65-	41
85-105	23

Calculate estimates of the mean and the standard deviation of these times. *(5 marks)*

(b) A sample of people who commute regularly from a town in Essex into London was also asked for an estimate of the time taken on their most recent journey. Their answers had a mean of 64 minutes and a standard deviation of 21 minutes. Compare, briefly, the journey times estimated by commuters from the two towns. *(2 marks)*

(c) Give **two** reasons why the data presented in parts (a) and (b) may not adequately represent typical commuting times from the two towns. *(2 marks)*

3 A cricket team meets for fielding practice. One exercise consists of a cricket ball being thrown at different heights, speeds and angles to one side of a fielder who tries to catch it one handed.

Each member of the team attempts 25 catches with each hand. The number of successful catches are given in the following table.

Fielder	**A**	**B**	**C**	**D**	**E**	**G**	**H**	**I**	**J**	**K**	**L**
Left hand	11	13	9	17	21	16	14	8	19	19	20
Right hand	18	17	20	22	14	19	21	15	10	24	23

(a) Calculate the value of the product moment correlation between the number of catches with the left hand and the number of catches with the right hand. *(3 marks)*

(b) Comment on the performance of fielders **E** and **J**. *(2 marks)*

(c) When fielders **E** and **J** are omitted from the calculation, the value of the product moment correlation coefficient between the number of left-handed and the number of right-handed catches is 0.812, correct to three decimal places. Comment on this value and the value you calculated in part (a) *(2 marks)*

4 The weights of the contents of jars of honey may be assumed to be normally distributed with the standard deviation 3.1 grams. The weights of the contents, in grams, of a random sample of eight jars were as follows:

458 450 457 456 460 459 458 456

(a) Calculate a 95% confidence interval for the mean weight of the contents of all jars. *(6 marks)*

(b) On each jar it states "Contents 454 grams". Comment on this statement using the given sample and your results in part (a). *(3 marks)*

TURN OVER FOR THE NEXT QUESTION

Turn over ►

5 Last year the employees of a firm either received no pay rise, a small pay rise or a large pay rise. The following table shows the number in each category, classified by whether they were weekly paid or monthly paid.

	No pay rise	**Small pay rise**	**Large pay rise**
Weekly Paid	25	85	5
Monthly paid	4	8	23

A tax inspector decides to investigate the tax affairs of an employee selected at random.

D is the event that a weekly paid employee is selected.
E is the event that an employee who received no pay rise is selected.
E' is the event "not E".

(a) Find the value of:

(i) P (D);

(ii) P ($D \mid E$) ;

(iii) P ($D \cap E'$). *(5 marks)*

(b) The tax inspector now decides to select three employees. Find the probability that they are all weekly paid if:

(i) one is selected at random from those who had no pay rise, one from those who had a small pay rise and one from those who had a large pay rise; *(3 marks)*

(ii) they are selected at random (without replacement) from all the employees of the firm. *(2 marks)*

6 **[**A sheet of graph paper is provided for use in this question.**]**

Andrew (**A**), Charles (**C**) and Edward (**E**) are employed by the Palace Hotel. Each is responsible for one floor of the building and their duties include cleaning the bedrooms. The number of bedrooms occupied on each floor varies from day to day.

The following table shows 10 observations of the number, x, of bedrooms to be cleaned and the time taken, y minutes, to carry out the cleaning. The employee carrying out the cleaning is also indicated.

Employee	**A**	**C**	**E**	**E**	**C**	**A**	**A**	**E**	**C**	**C**
x	8	22	12	24	19	14	22	16	10	21
y	110	211	132	257	184	165	248	171	97	196

(a) Plot a scatter diagram of the data. Identify the employee by labelling each point. *(3 marks)*

(b) Calculate the equation of the regression line of y on x. Draw the line on your scatter diagram. *(6 marks)*

(c) Use your regression equation to estimate the time which would be taken to clean 18 bedrooms. *(1 mark)*

(d) Calculate the residuals for the three observations when Andrew did the cleaning. *(3 marks)*

(e) Modify your estimate in part (c), given that the 18 bedrooms are to be cleaned by Andrew. *(2 marks)*

7 A gas supplier maintains a team of engineers who are available to deal with leaks reported by customers. Most reported leaks can be dealt with fairly quickly but some require a long time. The time (excluding travelling time), X, taken to deal with reported leaks is found to have a mean of 65 minutes and a standard deviation of 60 minutes.

(a) Assuming that the times may be modelled by a normal distribution, find the probability that it will take:

(i) more than 185 minutes to deal with a reported leak; *(3 marks)*

(ii) between 50 minutes and 125 minutes to deal with a reported leak. *(4 marks)*

(b) The mean of the times taken to deal with each of a random sample of 90 leaks is denoted by $\overline{X}$.

(i) State the distribution of $\overline{X}$. *(3 marks)*

(ii) Find the probability that $\overline{X}$ is less than 70 minutes. *(2 marks)*

(c) A statistician consulted by the gas supplier stated that, as the times had a mean of 65 minutes and a standard deviation of 60 minutes, the normal distribution would not provide an adequate model.

(i) Explain the reason for the statistician's statement. *(2 marks)*

(ii) Give a reason why, despite the statistician's statement, your answer to part (b)(ii) is still valid. *(2 marks)*

END OF QUESTIONS

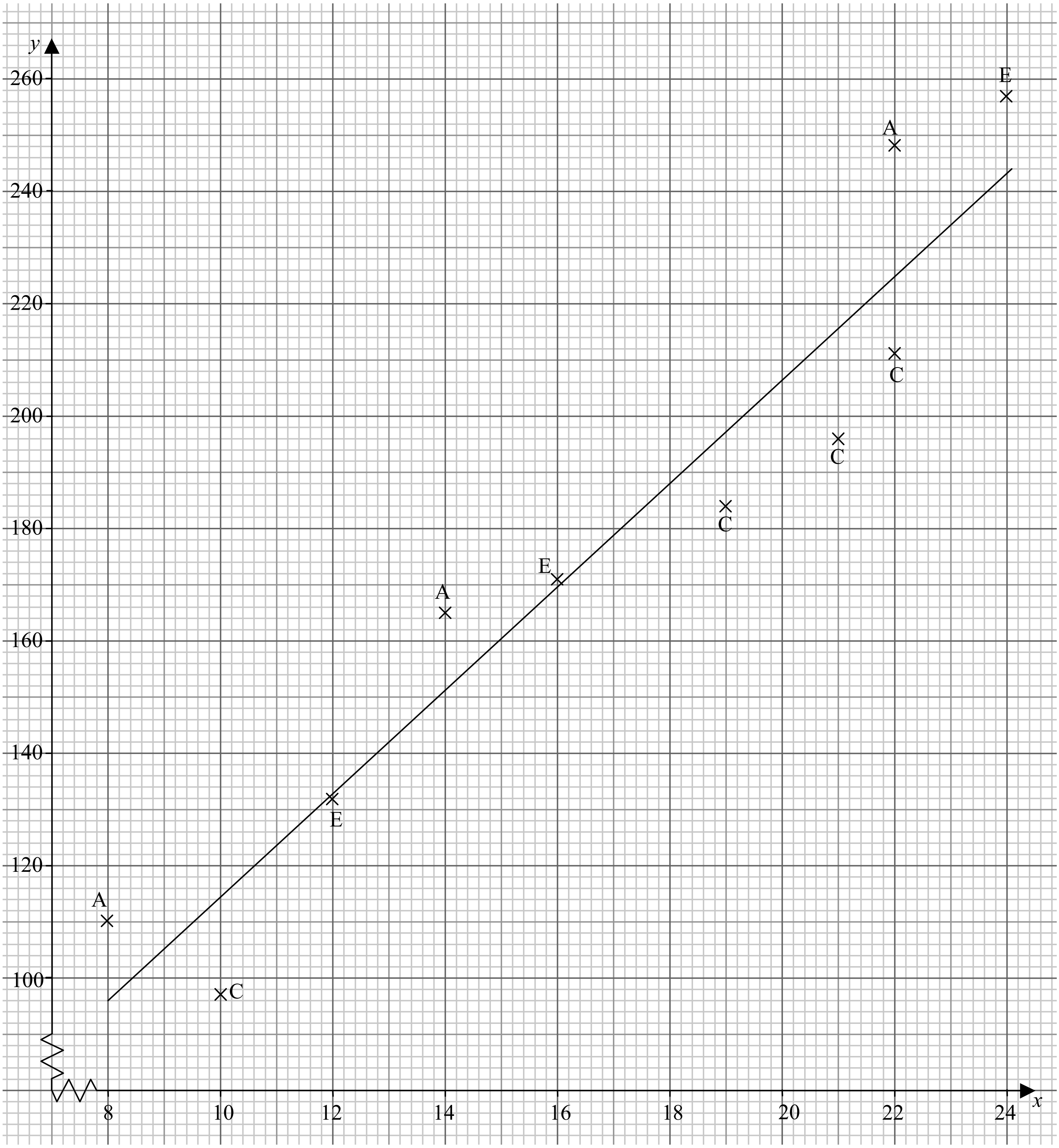
MS1B Specimen

Question	Solution	Marks	Total	Comments
1(a)(i)	Binomial $n = 40$ $p = 0.09$	B1B1		
	P(2 or fewer) = 0.2894	B1	3	0.289 - 0.29
(ii)	P(>5) = 1– P(5 or fewer)	M1		
	= 1– 0.8535 = 0.1465	A1	2	0.146 - 0.147
(b)	$P(2) = (16\times15/2) \times 0.09^2 \times 0.91^{14}$	B1M1		
	= 0.260	A1	3	0.259 – 0.26
(c)	probabilities independent/people selected at random/equivalent	E1	1	
	Total		**9**	
2(a)	Class mid-mark Frequency 40 12 50 54 60 68 75 41 95 23	M1		Allow m1A1 for mean and s.d. if method shown.
	$x = 63.2$ $s = 15.2$	A2A2	5	63.2 (63.1 – 63.3) 15.2 (15.0 - 15.3)
(b)	Journeys from Surrey have similar duration, on average, but are less variable than those from Essex.	E1 E1	2	
(c)	People asked may not be representative. Times are estimated not measured.	E1 E1	2	Or any other sensible comments e.g. journey time not defined , weather conditions may be extreme etc
	Total		**9**	
3(a)	0.0477	B3	3	0.047 – 0.048 allow M2A1 if method shown
(b)	**E** and **J** held more catches with left than with right hand - all others held more with right than left.	E1 E1	2	
(c)	Correlation coefficient of 0.812 suggests that those who caught a lot of catches with one hand also caught a lot of catches with the other. When **E** and **J** (possibly left handers) are included the correlation coefficient of 0.0477 suggests no association between the number of catches with each hand.	E1 E1	2	
	Total		**7**	

Question	Solution	Marks	Total	Comments
4(a)	$\bar{x}$ = 456.75	B1		
	95% confidence interval for mean	B1M1		
	456.75 $\pm$ 1.96 $\times$3.1/$\sqrt{8}$	M2		
	456.75 $\pm$ 2.15			
	454.60 – 458.90	A1	6	
(b)	The confidence interval provides evidence that the mean contents are greater than 454 grams. However the sample shows that some jars will contain less than 454 grams.	E1 E1 E1	3	E1 confidence interval refers to **mean** contents E1 evidence mean >454 E1 some individual contents <454
	Total		**9**	
5(a)(i)	115/150 = 0.767	B1	1	acf
(ii)	25/29 = 0.862	M1A1	2	acf
(iii)	90/150 = 0.6	M1A1	2	acf
(b)(i)	25/29 $\times$ 85/93 $\times$ 5/28 = 0.141	M1 M1A1	3	0.14 - 0.141
(ii)	115/150 $\times$ 114/149 $\times$ 113/148 = 0.448	M1 A1	2	
	Total		**10**	
6(a)	See graph on next page	M1 A1 B1	3	
(b)	y = 22.8 + 9.19x	B2B2		22.7 – 22.8 9.18 – 9.2 Allow M1A1 for a and b if method shown
	x = 8 y = 96.3 x = 23 y = 234.1	M1A1	6	+ line on graph
(c)	188	B1	1	188 – 188.3, allow 190
(d)	Residuals 110 – 22.77 – 9.186 $\times$ 8 =13.7 165 – 22.77 —9.186 $\times$ 14 = 13.6 248 – 22.77 – 9.186 $\times$ 22 = 23.1	M1 A1 A1	3	M1 method - ignore sign, allow read from graph A1 one correct - ignore sign 13.7 (13 – 14) 13.6 (13 – 14) 23.1 (22 – 24) A1 all correct, including sign
(e)	188 + 17 = 205	M1 A1	2	Any sensible method 201 – 211
	Total		**15**	

MS1B (cont)

Graph for Question 6



Question	Solution	Marks	Total	Comments
7(a)(i)	$z = \frac{(185-65)}{60} = 2.0$	M1		
	P(X > 185) = 1 − 0.97725	M1		
	= 0.02275	A1	3	0.0227 − 0.023
(ii)	$z_1 = \frac{(50-65)}{60} = -0.25$	M1		
	$z_2 = \frac{(125-65)}{60} = 1.0$	m1		
	P(50 < X < 125)=	M1		
	0.84134 − (1 − 0.59871) = 0.440	A1	4	
(b)(i)	Normal, mean 65, s.d. 60/√90 = 6.32	B1 B1 B1	3	normal may be implied in (b)(ii)
(ii)	$z = \frac{(70-65)}{\frac{60}{\sqrt{90}}} = 0.7906$	M1		
	Probability mean of 90 less than 70 is 0.785	A1	2	0.785 − 0.786
(c)(i)	Mean is only a little more than one standard deviation above zero. For normal this implies substantial proportion of times would be negative.	E1		
	This is impossible so model must be inadequate.	E1	2	
(ii)	Mean of large sample will be approximately normally distributed even	E1		
	if parent distribution is not.	E1	2	
	Total		**16**	
	TOTAL		**75**	

General Certificate of Education
Specimen Unit
Advanced Level Examination

MATHEMATICS **MS2A**
Unit Statistics 2A

In addition to this paper you will require:
• an 8-page answer book;
• the AQA booklet of formulae and statistical tables.
You may use a graphics calculator.

Time allowed: 1 hour 15 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS2A.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The maximum mark for this paper is 60.
- Mark allocations are shown in brackets.

Advice
- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

1 For her 21st birthday present, Joanne wishes to have a course of driving lessons. In order to select the better driving school available in her area, she decides to compare the recent performances of participants taking lessons at two driving schools P and Q.

These performances are tabulated below.

	School P	**School Q**
Pass	100	120
Fail	36	24

Use a χ^2 test, at the 10% level of significance, to determine whether there is an association between the performance of participants and driving school. *(9 marks)*

2 The number of vehicles arriving at a toll bridge during a 5-minute period can be modelled by a Poisson distribution with mean 3.6.

(a) State the value for the standard deviation of the number of vehicles arriving at a toll bridge during a 5-minute period. *(1 mark)*

(b) Find:

(i) the probability that at least 3 vehicles arrive in a 5-minute period; *(3 marks)*

(ii) the probability that at least 3 vehicles arrive in each of three successive 5-minute periods. *(2 marks)*

(c) Show that the probability that **no** vehicles arrive in a 10-minute period is 0.0007, correct to four decimal places. *(2 marks)*

3 At a cinema, the time, T minutes, that customers have to wait in order to collect their tickets has the following probability density function.

$$f(t) = \begin{cases} \dfrac{t^2}{18} & 0 \leqslant t < 3 \\ \dfrac{1}{4}(5-t) & 3 \leqslant t \leqslant 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Write down the value of $\mathrm{P}(T = 4)$. *(1 mark)*

(b) Show that the median waiting time is 3 minutes. *(2 marks)*

(c) Find the probability that customers have to wait for less than 4 minutes in order to collect their tickets. *(4 marks)*

(d) Calculate the mean time that customers have to wait in order to collect their tickets. *(4 marks)*

4 The random variable X has the following distribution.

x	2	4	8
P($X = x$)	$\frac{3}{8}$	$\frac{7}{16}$	$\frac{3}{16}$

(a) Calculate values for $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. *(3 marks)*

(b) A rectangle has sides of length X and $\left(2 + \dfrac{64}{X}\right)$.

(i) Find values for the mean and variance of the **area** of the rectangle. *(4 marks)*

(ii) By tabulating the distribution for $Y = X + \dfrac{64}{X}$, or otherwise, show that $\mathrm{E}(Y) = 24.5$. *(2 marks)*

(iii) Hence find the mean value for the **perimeter** of the rectangle. *(3 marks)*

Turn over ►

5 Charles is an athlete specialising in throwing the javelin. During practice, he throws a javelin the following distances, in metres.

40.3	39.8	41.6	42.8	39.0	38.6	40.8	41.1

(a) Calculate unbiased estimates of the mean and the variance of the distance thrown. *(2 marks)*

(b) (i) Hence calculate a 95% confidence interval for the mean distance thrown. *(5 marks)*

(ii) State **two** assumptions that you make in order to do this calculation. *(2 marks)*

6 A random variable X is normally distributed with mean μ and variance 0.64 .

The null hypothesis $H_0 : \mu = 50$ is to be tested against the alternative hypothesis $H_1 : \mu \neq 50$, using the 5% level of significance.

The mean, $\overline{X}$, of a random sample of 40 observations of X is to be used as the test statistic.

(a) Write down the distribution of $\overline{X}$ assuming H_0 is true. *(3 marks)*

(b) Explain what is meant by:

(i) a Type I error;

(ii) a Type II error. *(2 marks)*

(c) Write down the probability of a Type I error. *(1 mark)*

(d) Calculate the acceptance region for X, giving the limits to two decimal places. *(5 marks)*

END OF QUESTIONS

MS2A Specimen

Question	Solution	Marks	Total	Comments
1	H_0 : no association	B1		
	<table><tr><td>100</td><td>120</td><td>220</td></tr><tr><td>36</td><td>24</td><td>60</td></tr><tr><td>136</td><td>144</td><td>280</td></tr></table>	B1		
	Calculate E_i : $\frac{220}{280}\times 136 = 106.9$ $\frac{220}{280}\times 144 = 113.1$ $\frac{60}{280}\times 136 = 29.1$ $\frac{60}{280}\times 144 = 30.9$	M1		Any correct method
	$\alpha = \lvert O_i - E_i \rvert - 0.5$ α^2 $\frac{\alpha^2}{E_i}$	M1		Use of $\sum \frac{(O-E)^2}{E}$
	6.357 40.4 0.3782 0.3572 1.3867 1.3097	M1		Attempted use of Yates' correction
	total = 3.432	A1		awfw 3.4 to 3.5
	$\nu = 1$ $\chi^2(1) = 2.706$	B1 B1		
	Reject H_0 at the 10% level			
	Evidence at the 10% level to suggest an association between performance and the driving school selected.	E1		Any correct interpretation
	Total		**9**	

MS2A (cont)

Question	Solution	Marks	Total	Comments
2(a)	Standard deviation = $\sqrt{3.6} = 1.90$	B1	1	1.897
(b)(i)	$\lambda = 3.6$ for a 5-minute period Let Y = no. of vehicles arriving in a 5-minute period then $Y \sim \mathrm{P}_0(3.6)$ and $\mathrm{P}(Y \leqslant 2) = 0.3027$	B1		
	$\mathrm{P}(Y \geqslant 3) = 1 - \mathrm{P}(Y \leqslant 2)$ $= 1 - 0.3027$ $= 0.6973$	M1		
	$= 0.697$	A1	3	
(ii)	p = 0.6973 P(at least 3 arrive in 3 succ. 5 min) $= (0.6973)^3$	M1		for $p(i)^3$
	$= 0.339046$ $= 0.339$	A1	2	
(c)	let X = no. vehicles arriving in a 10-minute period then $X \sim \mathrm{P}_0(7.2)$	B1		
	and $\mathrm{P}(X = 0) = e^{-7.2}$ $= 0.0007$	B1	2	**ag**
	Total		**7**	

MS2A (cont)

Question	Solution	Marks	Total	Comments
3(a)	$\mathrm{P}(T=4)=0$	B1	1	
(b)	Median = 3 $\Rightarrow \mathrm{P}(T \le 3)=\mathrm{P}(T>3)=0.5$			
	$\mathrm{P}(T>3)=\frac{1}{2}\times f(3)\times(5-3)$ $\frac{1}{2}\times 0.5\times 2$	M1		**or**
	$=0.5$	A1	2	$\int_0^3 \frac{t^2}{18}\mathrm{d}t=\left(\frac{t^3}{54}\right)_0^3=\frac{1}{2}$
(c)	$\mathrm{P}(T<4)=1-\mathrm{P}(T\ge 4)$	M1		
	$=1-\frac{1}{2}\times 1\times\frac{1}{4}$ $=1-\frac{1}{8}$	M1A1		
	$=\frac{7}{8}$	A1	4	
(d)	$\mathrm{E}(T)=\int_0^3 \frac{t^3}{18}\mathrm{d}t+\frac{1}{4}\int_3^5 t(5-t)\mathrm{d}t$	M1		
	$=\left[\frac{t^4}{72}\right]_0^3+\left[\frac{5t^2}{8}-\frac{t^3}{12}\right]_3^5$	m1		Integration attempted
	$=1\frac{1}{8}+\left[\left(\frac{125}{8}-\frac{125}{12}\right)-\left(\frac{45}{8}-\frac{9}{4}\right)\right]$	M1		
	$=2\frac{23}{24}$	A1	4	(2.9583)
	Total		**11**	

MS2A (cont)

Question	Solution	Marks	Total	Comments
4 (a)	$\mathrm{E}(X)=2\times\frac{3}{8}+4\times\frac{7}{16}+8\times\frac{3}{16}$	B1		
	$=4$			
	$\mathrm{E}(X^2)=4\times\frac{3}{8}+16\times\frac{7}{16}+64\times\frac{3}{16}$			
	$=20.5$			
	$\mathrm{Var}(X)=\mathrm{E}(X^2)-\mathrm{E}^2(X)$			
	$=20.5-4^2$	M1		
	$=4.5$	A1	3	
(b)(i)	Area = $A=X\left(2+\dfrac{64}{X}\right)=2X+64$	B1		
	$\mathrm{E}(A)=\mathrm{E}(2X+64)=2\mathrm{E}(X)+64$			
	$=2\times4+64$			
	$=72$	B1		
	$\mathrm{Var}(A)=\mathrm{Var}(2X+64)$	M1		
	$=4\mathrm{Var}(X)$			
	$=4\times4.5$			
	$=18$	A1	4	
(ii)	$Y=X+\dfrac{64}{X}$			
	Y: 34, 20, 16; P(Y=y): $\frac{3}{8}$, $\frac{7}{16}$, $\frac{3}{16}$	B1		
	$\mathrm{E}(Y)=34\times\dfrac{3}{8}+20\times\dfrac{7}{16}+16\times\dfrac{3}{16}$			
	$=24.5$	B1	2	
(iii)	$P=2Y+4$	B1		
	$\mathrm{E}(P)=2\mathrm{E}(Y)+4$	M1		
	$=2\times24.5+4$			
	$=53$	A1	3	
	Total		**12**	

MS2A (cont)

Question	Solution	Marks	Total	Comments
5(a)	$\sum = 324$ and $\sum x^2 = 13135.34$			
	$\hat{\mu} = \bar{x} = \dfrac{\sum x}{n} = \dfrac{324}{8} = 40.5$	B1		
	$s^2 = \dfrac{1}{7} \times \left(13135.34 - \dfrac{324^2}{8}\right)$			
	$= 1.91$	B1	2	awfw 1.90 to 1.91 ($s = 1.38$)
(b)(i)	Confidence Interval for μ			
	$= \bar{x} \pm \dfrac{t \times s}{\sqrt{n}}$	M1		
	Degrees of freedom $= v = 8 - 1 = 7$	B1		
	$t_7(0.975) = 2.365$	B1		
	$\therefore$ CI for μ is			
	$40.5 \times \pm \dfrac{2.365\sqrt{1.90571}}{\sqrt{8}}$	A1		
	$= 40.5 \pm 1.154$	A1	5	
	= (39.3, 41.7)			accept (39.3/4, 41.6/7)
(ii)	distances are normally distributed distances are independent	B1B1	2	accept random sample
	Total		**10**	

MS2A (cont)

Question	Solution	Marks	Total	Comments
6(a)	$\overline{X} \sim \mathrm{N}\left(50, \frac{0.64}{40}\right) \sim \mathrm{N}(50, 0.016)$	B1 M1 A1	3	for Normal and 50 idea of $\frac{\sigma^2}{n}$; correct
(b)(i)	Type I error: Reject H_0 when H_0 true	B1		or equivalent
(ii)	Type II error: Accepting H_0 when H_0 incorrect	B1	2	or equivalent
(c)	$\mathrm{P}(\text{Type I error}) = 0.05$	B1	1	
(d)	Acceptance region: $z = \pm 1.96$	B1		
	$\frac{\sigma}{\sqrt{n}} = 0.1265$	B1		
	$50 \pm 1.96 \times \sqrt{0.016}$	M1 A1		
	$= (49.75, 50.25)$	A1	5	
	Total		**11**	
	TOTAL		**60**	

General Certificate of Education
Specimen Unit
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

In addition to this paper you will require: • an 8-page answer book; • the AQA booklet of formulae and statistical tables. You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS2B.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The maximum mark for this paper is 75.
- Mark allocations are shown in brackets.

Advice
- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

1 The number of strikes per game, obtained by a tenpin bowler, can be modelled by a Poisson distribution with mean λ. For the first game played, $\lambda = 0.2$ and, for each subsequent game played, $\lambda = 1.1$.

A match consists of three consecutive games.

(a) Write down the distribution of T, the total number of strikes obtained by a tenpin bowler in a match. *(2 marks)*

(b) Write down the value of $\mathrm{Var}(T)$. *(1 mark)*

(c) Find the probability that, in a match, the tenpin bowler will obtain a total of between 3 and 5 strikes, inclusive. *(3 marks)*

2 Charles is an athlete specialising in throwing the javelin. During practice, he throws a javelin the following distances, in metres.

40.3 39.8 41.6 42.8 39.0 38.6 40.8 41.1

(a) Calculate unbiased estimates of the mean and the variance of the distance thrown. *(2 marks)*

(b) (i) Hence calculate a 95% confidence interval for the mean distance thrown. *(5 marks)*

(ii) State **two** assumptions that you need to make in order to do this calculation. *(2 marks)*

3 The results of a recent police survey of traffic travelling on motorways produced information about the genders of drivers and the speeds, S miles per hour, of their vehicles, as tabulated below.

	$S \leqslant 70$	$70 < S \leqslant 90$	$S > 90$
Male	17	40	70
Female	30	25	18

Investigate, at the 1% level of significance, the claim that there is no association between the gender of the driver and the speed of the car. *(11 marks)*

4 The continuous random variable X has a rectangular distribution with the following probability density function, where k is a constant.

$$\mathrm{f}(x)=\begin{cases}\frac{1}{6} & 6\leqslant x\leqslant k\\ 0 & \text{otherwise}\end{cases}$$

(a) State the value of k. *(1 mark)*

(b) Hence find values for the mean, μ, and the variance, σ^2, of X. *(2 marks)*

(c) Determine $\mathrm{P}(X > \mu + \sigma)$. *(2 marks)*

5 At a cinema, the time, T minutes, that customers have to wait in order to collect their tickets has the following probability density function.

$$\mathrm{f}(t)=\begin{cases}\dfrac{t^2}{18} & 0\leqslant t < 3\\ \dfrac{1}{4}(5-t) & 3\leqslant t\leqslant 5\\ 0 & \text{otherwise}\end{cases}$$

(a) Write down the value of $\mathrm{P}(T = 4)$. *(1 mark)*

(b) Show that the median waiting time is 3 minutes. *(2 marks)*

(c) Find the probability that customers have to wait for less than 4 minutes in order to collect their tickets. *(4 marks)*

(d) Calculate the mean time that customers have to wait in order to collect their tickets. *(4 marks)*

TURN OVER FOR THE NEXT QUESTION

Turn over ▶

6 The random variable X has the following probability distribution.

$\boldsymbol{x}$	2	4	8
$\mathbf{P}(\boldsymbol{X} = \boldsymbol{x})$	$\frac{3}{8}$	$\frac{7}{16}$	$\frac{3}{16}$

(a) Calculate values for $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. *(3 marks)*

(b) A rectangle has sides of length X and $\left(2 + \dfrac{64}{X}\right)$.

(i) Find values for the mean and variance of the **area** of the rectangle. *(4 marks)*

(ii) By tabulating the distribution for $Y = X + \dfrac{64}{X}$, or otherwise, show that $E(Y) = 24.5$. *(2 marks)*

(iii) Hence find the mean value for the **perimeter** of the rectangle. *(3 marks)*

7 An instrument for measuring the speed of passing motorists is tested by a police force.

A car is driven at known speeds along a straight road.

The error (*speed recorded by the instrument – actual speed of the car*) is observed on eight occasions with the following results in metres per second.

4.2 –2.8 3.7 –5.9 0.2 6.4 4.1 –1.9

Assuming these data to be a random sample from a normal distribution, investigate, at the 5% level of significance, whether the instrument is biased (i.e. whether the mean error differs from zero). *(10 marks)*

8 A random variable X is normally distributed with mean μ and variance 0.64.

The null hypothesis, $H_0 : \mu = 50$, is to be tested against the alternative hypothesis, $H_1 : \mu \neq 50$, using the 5% level of significance.

The mean, $\overline{X}$, of a random sample of 40 observations of X is to be used as the test statistic.

(a) Write down the distribution of $\overline{X}$ assuming H_0 is true. *(3 marks)*

(b) Explain what is meant by:

(i) a Type I error;

(ii) a Type II error. *(2 marks)*

(c) Write down the probability of a Type I error. *(1 mark)*

(d) Calculate the acceptance region for $\overline{X}$, giving the limits to two decimal places. *(5 marks)*

END OF QUESTIONS

MS2B Specimen

Question	Solution	Marks	Total	Comments
1(a)	X = number of strikes per game then: $X_1 \sim \text{Po}(0.2)$ $X_2 \sim \text{Po}(1.1)$ $X_3 \sim \text{Po}(1.1)$ $\therefore\ T = X_1 + X_2 + X_3 \sim \text{Po}(2.4)$	M1 A1	2	
(b)	Var (T) = 2.4	B1	1	
(c)	$\text{P}(3 \le T \le 5) = \text{P}(T = 3 \text{ or } 4 \text{ or } 5)$ $= \text{P}(T \le 5) - \text{P}(T \le 2)$ $= 0.9643 - 0.5697$ $= 0.3946$	M1 M1 A1	3	awfw 0.394 to 0.395
	Total		**6**	
2(a)	$\sum = 324$ and $\sum x^2 = 13135.34$ $\hat{\mu} = \bar{x} = \frac{\sum x}{n} = \frac{324}{8} = 40.5$ $s^2 = \frac{1}{7} \times \left(13135.34 - \frac{324^2}{8}\right)$ $= 1.91$	B1 B1	2	awfw 1.90 to 1.91 (s = 1.38)
(b)(i)	Confidence Interval for μ $= \bar{x} \pm \frac{t \times s}{\sqrt{n}}$ Degres of freedom = $\nu = 8 - 1 = 7$ $t_7(0.975) = 2.365$ $\therefore$ CI for μ is $= 40.5 \pm \frac{1.154\sqrt{1.90571}}{\sqrt{8}}$ $= 40.5 \pm 1.154$ $= (39.3, 41.7)$	M1 B1 B1 A1 A1	5	Use of Accept (39.3/4 , 41.6/7)
(ii)	distances are normally distributed distances are independent	B1 B1	2	Accept random sample
	Total		**9**	

MS2B (cont)

Question	Solution	Marks	Total	Comments
3	H_0: Gender of driver and speed of car are independent	B1		
	H_1: Gender of driver and speed of car are associated	B1		
	$\nu = (3 - 1)(2 - 1) = 2$	B1		
	O_i E_i $\dfrac{(O_i - E_i)^2}{E_i}$	M1		
	17 29.8 5.5284 (5.4980)	A2		
	40 41.3 0.0394 (0.0409)			
	70 55.9 3.5679 (3.5565)			
	30 17.2 9.6178 (9.5256)			
	25 23.7 0.0685 (0.0713)			
	18 32.1 6.2072 (6.1935)	M1		
	$\chi^2_{\text{calc}} = 25.03$ (24.89)	A1		awfw 24.5 to 25.5
	$\chi^2_{1\%}(2) = 9.210$	B1		
	$\therefore$ reject H_0	B1		
	There is (very strong) evidence at the 1% level to suggest an association between the gender of the driver and the speed of the car.	E1	11	
	Total		**11**	

MS2B (cont)

Question	Solution	Marks	Total	Comments
4 (a)	$k = 12$	B1	1	
(b)	$\mu = 9$	B1		
	$\sigma^2 = 3$	B1	2	
(c)	$\mathrm{P}(X > \mu + \sigma)$			
	$= \mathrm{P}\left(X > 9 + \sqrt{3}\right)$			
	$= \dfrac{12 - 10.732}{6}$	M1		$\dfrac{3-\sqrt{3}}{6}$
	$= 0.211$	A1	2	
	Total		**5**	
5(a)	$\mathrm{P}(T = 4) = 0$	B1	1	
(b)	Median = 3			
	$\Rightarrow \mathrm{P}(T \le 3) = \mathrm{P}(T > 3) = 0.5$			
	$\mathrm{P}(T > 3) = \tfrac{1}{2} \times \mathrm{f}(3) \times (5-3)$	M1		or
	$= \tfrac{1}{2} \times 0.5 \times 2$			$\int_0^3 \dfrac{t^2}{18}\,\mathrm{d}t = \left[\dfrac{t^3}{54}\right]_0^3 = \dfrac{1}{2}$
	$= 0.5$	A1	2	
(c)	$\mathrm{P}(T < 4) = 1 - \mathrm{P}(T \ge 4)$	M1		
	$= 1 - \dfrac{1}{2} \times 1 \times \dfrac{1}{4}$	M1 A1		
	$= 1 - \dfrac{1}{8}$			
	$= \dfrac{7}{8}$	A1	4	
(d)	$\mathrm{E}(T) = \int_0^3 \dfrac{t^3}{18}\,\mathrm{d}t + \tfrac{1}{4}\int_3^5 t(5-t)\,\mathrm{d}t$	M1		
	$= \left[\dfrac{t^4}{72}\right]_0^3 + \left[\dfrac{5t^2}{8} - \dfrac{t^3}{12}\right]_3^5$	m1		Integration attempted
	$= 1\tfrac{1}{8} + \left[\left(\tfrac{125}{8} - \tfrac{125}{12}\right) - \left(\tfrac{45}{8} - \tfrac{9}{4}\right)\right]$	M1		
	$= 2\tfrac{23}{24}$	A1	4	(2.9583)
	Total		**11**	

MS2B (cont)

Question	Solution	Marks	Total	Comments
6 (a)	$\mathrm{E}(X) = 2\times\frac{3}{8} + 4\times\frac{7}{16} + 8\times\frac{3}{16}$			
	$= 4$	B1		
	$\mathrm{E}(X^2) = 4\times\frac{3}{8} + 16\times\frac{7}{16} + 64\times\frac{3}{16}$			
	$= 20.5$			
	$\mathrm{Var}\,(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X)$	M1		
	$= 20.5 - 4^2$			
	$= 4.5$	A1	3	
(b)(i)	$\text{Area} = A = X\left(2 + \frac{64}{X}\right) = 2X + 64$	B1		
	$\mathrm{E}(A) = \mathrm{E}(2X + 64) = 2\mathrm{E}(X) + 64$			
	$= 2\times 4 + 64$			
	$= 72$	B1		
	$\mathrm{Var}(A) = \mathrm{Var}(2X + 64)$			
	$= 4\mathrm{Var}(X)$	M1		
	$= 4\times 4.5$			
	$= 18$	A1	4	
(ii)	$Y = X + \frac{64}{X}$			
	y: 34, 20, 16; $\mathrm{P}(Y = y)$: $\frac{3}{8}$, $\frac{7}{16}$, $\frac{3}{16}$	B1		
	$E(Y) = 34\times\frac{3}{8} + 20\times\frac{7}{16} + 16\times\frac{3}{16}$			
	$= 24.5$	B1	2	
(iii)	$P = 2Y + 4$	B1		
	$\mathrm{E}(P) = 2\mathrm{E}(Y) + 4$	M1		
	$= 2\times 24.5 + 4$	A1	3	
	$= 53$			
	Total		**12**	

The table in (ii):

y	34	20	16
$\mathrm{P}(Y = y)$	$\frac{3}{8}$	$\frac{7}{16}$	$\frac{3}{16}$

MS2B (cont)

Question	Solution	Marks	Total	Comments
7	$\bar{x} = \dfrac{\sum x}{n} = \dfrac{8}{8} = 1$	B1		
	$s^2 = \frac{8}{7} \times 15.925$	M1		
	$s = 4.266$	A1		
	$H_0 : \mu = 0$	B1		
	$H_1 : \mu \neq 0$	B1		
	$t = \dfrac{1-0}{\frac{4.266}{\sqrt{8}}} = 0.663$	M1 A1		
	$t_7 = \pm 2.365$	M1 A1		
	Accept H_0 No significant evidence of bias	A1	10	
	Total		**10**	
8(a)	$\bar{X} \sim N\left(50, \dfrac{0.64}{40}\right) \sim N(50, 0.016)$	B1		For Normal & 50
		M1		Idea of $\frac{\sigma^2}{n}$; correct
		A1	3	
(b)(i)	Type I error: Reject H_0 when H_0 true	B1		oe
(ii)	Type II error: Accepting H_0 when H_0 incorrect	B1	2	oe
(c)	P(Type I error) = 0.05	B1	1	
(d)	Acceptance region: $z = \pm 1.96$	B1		
	$\dfrac{\sigma}{\sqrt{n}} = 0.1265$	B1		
	$50 \pm 1.96 \times \sqrt{0.016}$	M1 A1		
	= (49.75, 50.25)	A1	5	
	Total		**11**	
	TOTAL		**75**	

General Certificate of Education
Specimen Unit
Advanced Level Examination

MATHEMATICS **MS03**
Unit Statistics 3

In addition to this paper you will require:
• an 8-page answer book;
• the AQA booklet of formulae and statistical tables.
You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions

- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS03.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information

- The maximum mark for this paper is 75.
- Mark allocations are shown in brackets.

Advice

- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

1 In a local government election, 40% of adults voted. The random variable V denotes the number who voted out of a random sample of 200 adults.

(a) (i) Write down the probability distribution for V. *(1 mark)*

(ii) Write down an appropriate distributional approximation for V. *(1 mark)*

(iii) Find an approximate value for $\mathrm{P}(V > 90)$.. *(3 marks)*

(b) Only 5% of those who voted in the election voted for candidate A.

(i) Find the probability that an adult, chosen at random, voted for candidate A. *(1 mark)*

(ii) Using a suitable distributional approximation, find the probability that, out of a random sample of 200 adults, more than 5 voted for candidate A. *(3 marks)*

2 Over a period of one year, a greengrocer sells tomatoes at six different prices (x pence per kilogram). He calculates the average number of kilograms, y, sold per day at each of the six different prices.

His results are as follows.

x	99	105	110	120	125	130
y	97	88	80	62	68	59

(a) Calculate the value of the product moment correlation coefficient. *(3 marks)*

(b) (i) Assuming that these data are a random sample from a bivariate distribution with correlation coefficient ρ, investigate, at the 1% level of significance, the hypothesis that $\rho < 0$. *(4 marks)*

(ii) Interpret your result in the context of this question. *(1 mark)*

3 To travel to school in the morning, Sam either catches the bus, cycles or gets a lift with her father. The probabilities for these three modes of travel are 0.3, 0.3 and 0.4, respectively.

The probability that Sam is late for school is:

0.4 when she catches the bus;

0.5 when she cycles;

0.1 when she gets a lift with her father.

(a) Find the probability that, on a particular morning, Sam cycles to school and arrives late. *(2 marks)*

(b) Sam is late for school one morning. Find the probability that she cycled to school that morning. *(4 marks)*

4 (a) The random variable X has a Poisson distribution with parameter λ. Prove that the mean of X is λ. *(3 marks)*

(b) The daily number of missed appointments at a dental surgery can be modelled by a Poisson distribution with mean 1.8.

In an attempt to reduce this figure, the surgery publicised the introduction of fines for missed appointments.

(i) During the first 5 days after the fines were introduced, there was a total of 6 missed appointments. Use a test with the 10% level of significance to show that there is insufficient evidence to conclude that the introduction of fines has succeeded in its aim. *(5 marks)*

(ii) A dentist at the surgery suggested that the number of missed appointments should be monitored over a much longer period of time. Subsequently, there was a total of 297 missed appointments during 180 days. Test, again at the 10% level of significance, the claim that the mean daily number of missed appointments is now less than 1.8. *(8 marks)*

TURN OVER FOR NEXT QUESTION

Turn over▶

5 The random variable L has mean 65 and variance 16. The random variable S has mean 50 and variance 9. Variables L and S are independent and the random variable D is given by

$$D = 4L - 5S$$

(a) (i) Show that D has mean 10. *(2 marks)*

(ii) Find the value of the variance of D. *(2 marks)*

(b) The weights of large eggs are normally distributed with mean 65 grams and standard deviation 4 grams. The weights of standard eggs are normally distributed with mean 50 grams and standard deviation 3 grams.

One large egg and one standard egg are chosen at random. Find the probability that the weight of the standard egg is more than $\frac{4}{5}$ of the weight of the large egg. *(5 marks)*

6 It is claimed that women are faster than men at solving anagrams. To investigate the claim, the same list of anagrams was given to random samples of 110 men and 125 women under identical experimental conditions. The time, in minutes, taken by each person to solve the anagrams was recorded.

The results are summarised in the table below, together with known values for the population standard deviations.

	Sample Size	**Sample Mean**	**Population Standard Deviation**
Men (M)	110	15.4 minutes	2.1 minutes
Women (F)	125	14.9 minutes	1.8 minutes

(a) A hypothesis test is to be applied to these data to determine whether, on average, men take longer than women to solve anagrams.

Stating null and alternative hypotheses concerning the population means μ_M and μ_F, and using the 5% level of significance, find the critical region of this test.

Hence determine whether the evidence supports the claim that women are faster than men at solving anagrams. *(8 marks)*

(b) If the alternative hypothesis states that $\mu_M - \mu_F = 0.3$, find the power of the test. *(5 marks)*

7 A company is introducing a new brand of toothpaste. The company's marketing department is investigating people's opinions of the new brand by giving out free samples and asking people to compare the new brand with their usual brand.

(a) From a random sample of 300 people, 129 preferred the new brand of toothpaste to their usual brand.

(i) Calculate an approximate 98% confidence interval for the proportion of people who prefer the new brand to their usual brand. *(6 marks)*

(ii) Comment on the evidence for claiming that the majority of people prefer the new brand. *(1 mark)*

(b) In a further survey, 200 people aged under 25 and 250 people aged 25 or over were asked their opinion. The results are summarised below.

Age (years)	**Sample size**	**Proportion who preferred new brand**
Under 25	200	0.47
25 or over	250	0.38

(i) Calculate an approximate 98% confidence interval for the difference between the proportions of people aged under 25 and those aged 25 or over who prefer the new brand of toothpaste. *(6 marks)*

(ii) Hence determine whether people aged under 25 are more likely than those aged 25 or over to prefer the new brand of toothpaste. *(1 mark)*

END OF QUESTIONS

MS03 Specimen

Question	Solution	Marks	Total	Comments
1(a)(i)	$V \sim \mathrm{B}(200,\ 0.4)$	B1	1	
(ii)	$V \sim \mathrm{N}(80, 48)$	B1	1	
(iii)	$\mathrm{P}(V > 90) = \mathrm{P}\left(z > \dfrac{90.5-80}{\sqrt{48}}\right)$	M1		
	$= \mathrm{P}(z > 1.52)$	A1		Condone no continuity correction (1.44)
	$= 1 - 0.93574$			
	$= 0.0643$ (3 s f)	A1	3	0.064 to 0.066 ; cao
(b)(i)	P(voted for A)			
	$= 0.4 \times 0.05 = 0.02$	B1	1	cao
(ii)	$X \sim \mathrm{B}(200, 0.02) \approx \mathrm{Poisson}(4)$	B1		Use of Poisson
	$\mathrm{P}(X > 5) = 1 - \mathrm{P}(X \le 5)$	M1		
	$= 0.215$	A1	3	
	Total		**9**	
2(a)	$r = -0.962$	B3	3	awfw -0.962 to -0.961 Allow M2 A1 if method shown
(b)(i)	H_0: $\rho = 0$ H_1: $\rho < 0$	B1		Both
	$n = 6$, $\alpha = 1\%$			
	Critical value of coefficient = -0.8822	B1		Accept positive value
	$-0.962 < -0.8822$	M1		
	Reject H_0 at 1% level.	A1	4	
	Evidence suggests that $\rho < 0$			
(ii)	Fewer tomatoes are sold when price is higher.	E1	1	
	Total		**8**	

MS03 (cont)

Question	Solution	Marks	Total	Comments
3(a)	P(cycles and late) = 0.3×0.5 $= 0.15$	M1 A1	2	
(b)	Denoting events by: A: Sam catches the bus, B: Sam cycles, C: Sam gets a lift, L: Sam is late for school, P(B|L) = $\dfrac{P(L\|B)\times P(B)}{P(L\|A)\times P(A)+P(L\|B)\times P(B)+P(L\|C)\times P(C)}$ $= \dfrac{0.5\times0.3}{(0.4\times0.3)+(0.5\times0.3)+(0.1\times0.4)}$ $= \dfrac{0.15}{0.12+0.15+0.04}$ $= \dfrac{0.15}{0.31} = 0.484$	M1A1 A1 A1	 4	Use of Bayes'theorem with correct numerator and sum of 3 probabilities in denominator. All four terms correct awrt
	Total		**6**	
4(a)	$E(X) = \sum_{r=0}^{\infty} r\times P(X=r)$ $= \sum_{r=1}^{\infty} r\times \dfrac{e^{-\lambda}\lambda^r}{r!}$ $= \lambda e^{-\lambda}\sum_{n=0}^{\infty}\dfrac{\lambda^n}{n!}$ $= \lambda e^{-\lambda}e^{\lambda} = \lambda$	 M1 A1 A1	 3	 Attempt at summation of products of r and P(r). where $n = r - 1$ Convincing use of series to prove result.
(b)(i)	H_0: Mean rate unchanged (or $\mu = 9$) H_1: Mean rate decreased (or $\mu < 9$) Under H_0, $X \sim$ Poisson(9) $P(X \le 6) = 0.2068$ 0.2068 > 10% so insufficient evidence of decrease.	 B1 B1 M1A1 E1	 5	 Both; or equivalent Awrt 0.207 Must show comparison
(ii)	If rate is unchanged at 1.8 per day, $Y \sim$ Poisson(324) $\approx$ N(324, 324) $P(Y \le 297) = P\left(Z \le \dfrac{297-324}{\sqrt{324}}\right)$ $= P(Z \le -1.5)$ $= 1 - 0.93319$ $= 0.0668$ (3 s.f.) 0.067 < 10% Evidence suggests that the mean daily rate is now less than 1.8.	M1A1 M1 A1 A1 A1 M1 A1	 8	Use of normal approximation with correct mean & variance or, with continuity correction $P(Z \le -1.472)$ = 0.071 (awrt) awrt 0.067 Accept calculation of sample z-value and comparison with −1.28
	Total		**16**	

Question	Solution	Marks	Total	Comments
5(a)(i)	$\text{E}(D) = 4\text{E}(L) - 5\text{E}(S)$	M1		
	$= 260 - 250$	A1		
	$= 10$		2	
(ii)	$\text{Var}(D) = 4^2\text{Var}(L) + 5^2\text{Var}(S)$	M1		
	$= 256 + 225$			
	$= 481$	A1		
			2	
(b)	$L \sim \text{N}(65, 4^2)$; $S \sim \text{N}(50, 3^2)$			
	$\text{P}\left(S > \frac{4}{5}L\right) = \text{P}\left(S - \frac{4}{5}L > 0\right)$	M1		
	$= \text{P}(5S - 4L > 0) = \text{P}(D < 0)$	A1		
	$D \sim \text{N}(10, 481)$			
	$\text{P}(D < 0) = \text{P}\left(Z < \frac{0-10}{\sqrt{481}}\right)$	M1		
	$= \text{P}(Z < -0.456)$	A1√		√ on part (a) (ii)
	$= 0.322$ to 0.327	A1	5	
	Total		**9**	
6(a)	H_0: $\mu_M = \mu_F$			
	H_1: $\mu_M > \mu_F$	B1		both
	Under H_0, using central limit theorem,			
	$\overline{X}_M - \overline{X}_F \sim \text{N}\left(0, \frac{2.1^2}{110} + \frac{1.8^2}{125}\right)$	M1		
	$= \text{N}(0, 0.0660)$	A1		may be implied
	One-tailed test at 5% level so			
	$z = 1.6449$	B1		
	Critical region is			
	$\overline{X}_M - \overline{X}_F > 1.6449 \times \sqrt{0.0660}$	M1		
	ie. $\overline{X}_M - \overline{X}_F > 0.423$	A1		awfw 0.42 to 0.43
	From the sample data, $\bar{x}_M - \bar{x}_F = 0.5$	B1		
	Reject H_0. The evidence supports the claim.	A1√	8	√ on critical region
(b)	H_1: $\mu_M - \mu_F = 0.3$			
	Under H_1, $\overline{X}_M - \overline{X}_F \sim \text{N}(0.3, 0.0660)$	B1		may be implied
	Power of test =			
	$\text{P}\left(\overline{X}_M - \overline{X}_F > 0.423 \mid \text{H}_1\right)$	M1		
	$= \text{P}\left(Z > \frac{0.423 - 0.3}{\sqrt{0.0660}}\right)$	M1 A1√		√ on critical value
	$= 1 - \Phi(0.479) = 0.316$	A1	5	awfw 0.31 to 0.32
	Total		**13**	

MS03 (cont)

Question	Solution	Marks	Total	Comments
7(a)(i)	$\hat{p} = \frac{129}{300} = 0.43$	B1		cao
	98% confidence interval so $z = 2.3263$	B1		
	Standard error = $\sqrt{\frac{0.43 \times 0.57}{300}}$	M1 A1		
	98% confidence limits are $0.43 \pm 2.3263\sqrt{\frac{0.43 \times 0.57}{300}}$	M1		
	giving (0.363 to 0.364, 0.496 to 0.497)	A1√	6	√ on $\hat{p}$ and se
(ii)	Confidence interval lies below 50% so the claim is not supported.	B1	1	
(b)(i)	Denoting proportions of younger and older people by p_X and p_Y respectively, $\hat{p}_X - \hat{p}_Y = 0.09$	B1		cao
	Estimated standard error for $p_X - p_Y$ is $\sqrt{\frac{0.47 \times 0.53}{200} + \frac{0.38 \times 0.62}{250}}$	M1 A1		
	$= \sqrt{0.0021879}$	A1		
	98% confidence limits for $p_X - p_Y$: $0.09 \pm 2.3263 \times \sqrt{0.0021879}$ giving (-0.0188, 0.199)	M1 A1	6	
(ii)	Zero lies within confidence interval so not enough evidence for larger proportion.	E1	1	
	Total		**14**	
	TOTAL		**75**	

General Certificate of Education
Specimen Unit
Advanced Level Examination

MATHEMATICS **MS04**
Unit Statistics 4

In addition to this paper you will require:
• an 8-page answer book;
• the AQA booklet of formulae and statistical tables.
You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use blue or black ink or ball-point pen. Pencil should only be used for drawing.
- Write the information required on the front of your answer book. The *Examining Body* for this paper is AQA. The *Paper Reference* is MS04.
- Answer **all** questions.
- All necessary working should be shown; otherwise marks for method may be lost.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The maximum mark for this paper is 75.
- Mark allocations are shown in brackets.

Advice
- Unless stated otherwise, formulae may be quoted, without proof, from the booklet.

Answer **all** questions.

1 When an automatic bottling machine is working properly, the volume of liquid delivered into each bottle can be assumed to be normally distributed with standard deviation 0.5 ml. On a certain day, the Quality Control Manager believes that the standard deviation is larger than this. He therefore measures the contents of six randomly selected bottles with the following results.

Contents (ml)	204.1	202.8	205.0	203.6	203.9	202.0

Investigate the Quality Control Manager's belief at the 10% level of significance. *(9 marks)*

2 The times, T minutes, between successive telephone calls to a small business can be assumed to be exponentially distributed with mean 10.

(a) (i) Write down, in full, the probability density function of T. *(2 marks)*

(ii) Using integration, obtain an expression, valid for $t \geq 0$, for the distribution function of T. *(3 marks)*

(b) A call is received at 11.00 am.

Find the probability that the next call is received:

(i) before 11.15 am; *(2 marks)*

(ii) between 11.10 am and 11.20 am. *(2 marks)*

3 Independent observations X_1, X_2 and X_3 are made on the random variable X which has mean μ and variance σ^2.

(a) Determine which of the following random variables are unbiased estimators for μ.

$$R = \frac{1}{3}X_1 + \frac{1}{2}X_2 + \frac{1}{6}X_3$$

$$S = \frac{3}{5}X_1 + \frac{3}{10}X_2 + \frac{1}{5}X_3$$

$$T = \frac{3}{4}X_1 + \frac{1}{2}X_2 - \frac{1}{4}X_3$$

(5 marks)

(b) Calculate the variance of those estimators which are unbiased. *(3 marks)*

(b) Hence state, giving a reason, which of these estimators is the best unbiased estimator for μ. *(1 mark)*

4 Eight patients have their pulse rates, in beats per minute, measured before and after receiving a certain treatment. The results, including the decrease in pulse rate for each patient, are given in the table.

Patient	***A***	***B***	***C***	***D***	***E***	***F***	***G***	***H***
Before	73.2	69.1	76.9	75.6	68.8	79.5	69.9	74.3
After	70.8	65.7	77.5	69.2	71.1	77.6	67.9	70.4
Decrease	2.4	3.4	–0.6	6.4	–2.3	1.9	2.0	3.9

Using a t-test, investigate, at the 5% significance level, whether this treatment decreases, on average, patients' pulse rates. *(9 marks)*

5 A six-sided die has the numbers 1, 2, 3, 4, 5 and 6 on its faces. The random variable R denotes the number of independent throws of the die needed to first obtain a score of either 5 or 6.

Assuming that the die is unbiased, the probability model for R is:

$$\mathrm{P}(R = r) = \begin{cases} \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^{r-1} & r = 1,\ 2,\ 3, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

The frequency table below shows the results of 243 observations of R.

Number of throws, r	1	2	3	4	5	$\geqslant 6$
Frequency, f	97	58	29	19	12	28

Using a χ^2 goodness of fit test and the 5% level of significance, test the hypothesis that the die is unbiased. *(11 marks)*

Turn over ►

6 The random variable X follows the probability distribution

$$\mathrm{P}(X = r) = q^{r-1}p \quad \text{for } r = 1,2\ldots\ldots$$

where $p + q = 1$.

(a) Prove that:

(i) $\mathrm{E}(X) = \dfrac{1}{p}$; *(3 marks)*

(ii) $\mathrm{Var}(X) = \dfrac{q}{p^2}$. *(5 marks)*

(b) In the case when $p = 0.4$, find the value of n such that

$$\mathrm{P}\,(X > n) < 0.001$$

(4 marks)

7 Radha, a gardener, buys 12 tomato plants: 6 of variety X and 6 of variety Y. She plants these in her greenhouse and she keeps a record of the total yield, in kilograms, from each plant. She obtains the following results.

Variety *X*	3.7	4.9	2.8	5.3	4.6	3.3
Variety *Y*	5.8	4.1	2.9	6.3	3.8	5.0

You may assume that these are random samples from populations distributed $\mathrm{N}(\mu_X, \sigma_X^{\ 2})$ and $\mathrm{N}(\mu_Y, \sigma_Y^{\ 2})$ respectively.

(a) Test, at the 5% significance level, the hypotheses

$$\mathrm{H}_0 : \sigma_X^{\ 2} = \sigma_Y^{\ 2} \text{ against } \mathrm{H}_1 : \sigma_X^{\ 2} \neq \sigma_Y^{\ 2}.$$

(7 marks)

(b) (i) Obtain a 95% confidence interval for $\mu_X - \mu_Y$. *(7 marks)*

(ii) State, giving a reason, whether your result in part (b)(i) indicates a difference between the mean yields of the two varieties. *(2 marks)*

END OF QUESTIONS

MS04 Specimen

Question	Solution	Marks	Total	Comments
1	$H_0 : \sigma = 0.5$ or $\sigma^2 = 0.25$	B1		both
	$H_1 : \sigma > 0.5$ or $\sigma^2 > 0.25$			
	Significance level, $\alpha = 0.10$ (10%)			
	Degrees of freedom, $v = 6 - 1 = \underline{\mathbf{5}}$	B1		cao
	Critical value, $\chi^2 = 9.236$	B1		awfw 9.23 to 9.24
	$\sum x = 1221.4$ and $\sum x^2 = 248641.82$			
	$(n-1)s^2 = \sum x^2 - \frac{\left(\sum x^2\right)}{n} =$	M1		use of, or equivalent accept n rather than $(n-1)$
	5.4933	A1		awfw 5.49 to 5.50 $s^2 = 1.0986$ $s = 1.0482$ $\sigma^2 = 0.91556$ $\sigma = 0.95685$
	$\chi^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{5.4933}{0.25} =$	M1 A1√		use of, or equivalent √ on calculation
	21.97	A1		awfw 21.9 to 22.0
	∴ evidence, at 10% level, to support manager's belief	A1√	9	√ on χ^2 and CV
	Total		**9**	
2 (a) (i)	$f(t) = \begin{cases} \frac{1}{10}e^{-\frac{t}{10}} & t > 0 \\ 0 & t < 0 \end{cases}$	B1		cao
	$t > 0$			
	$f(t) = 0$ $t < 0$	B1	2	cao 0 andboth ranges
(ii)	$F(t) = \int_0^t \frac{1}{10} e^{-\frac{x}{10}} \, dx =$	M1		use of $\int f(x)dx$ condone use of t
	$\left[-e^{-\frac{x}{10}} \right]_0^t = 1 - e^{-\frac{t}{10}}$	A1 A1	3	integral & limits cao
(b) (i)	Require P (T<15) = F (15) =	M1		Use of F(15), or integration
	0.776 to 0.777	A1	2	awfw
(ii)	Require P (10< T< 20) = F(20)– F (10) =	M1		use of difference of Fs or equivalent
	= 0.86466 – 0.63212 = 0.232 to 0.233	A1	2	awfw
	Total		**9**	

MS04 (cont)

Question	Solution	Marks	Total	Comments
3 (a)	$\mathrm{E}(R) = \mathrm{E}\left(\frac{1}{3}X_1 + \frac{1}{2}X_2 + \frac{1}{6}X_3\right)$			
	$= \frac{1}{3}\mathrm{E}(X_1) + \frac{1}{2}\mathrm{E}(X_2) + \frac{1}{6}\mathrm{E}(X_3)$	M1		Used on any one of *R*, *S*, *T*
	$= \frac{1}{3}\mu + \frac{1}{2}\mu + \frac{1}{6}\mu$	A1		
	$\mathrm{E}(S) = \frac{3}{5}\mu + \frac{3}{10}\mu + \frac{1}{5}\mu$	A1		
	$= \frac{11}{10}\mu$			
	$\mathrm{E}(T) = \frac{3}{4}\mu + \frac{1}{2}\mu - \frac{1}{4}\mu$	A1		
	$\therefore R$ and T are unbiased	A1√	5	
(b)	$\mathrm{Var}(R) = \frac{1}{9}\mathrm{Var}(X_1) + \frac{1}{4}\mathrm{Var}(X_2) + \frac{1}{36}\mathrm{Var}(X_3)$	M1		Use of any answer from (a)
	$= \frac{1}{9}\sigma^2 + \frac{1}{4}\sigma^2 + \frac{1}{36}\sigma^2$			
	$= \frac{7}{18}\sigma^2$	A1		
	$\mathrm{Var}(T) = \frac{9}{16}\sigma^2 + \frac{1}{4}\sigma^2 + \frac{1}{16}\sigma^2$			
	$= \frac{7}{8}\sigma^2$	A1	3	
(c)	*R* best (smaller variance)	E1	1	
	Total		**9**	
4	$H_0 : \mu_0 = 0 \qquad H_1 : \mu_D > 0$	B1		Both
	$\nu = 8 - 1 = 7$	B1		cao
	$t_{crit} = 1.895$	B1		awfe 1.89 to 1.90
	$\sum d = 17.1 \qquad \sum d^2 = 86.75$			
	$\bar{d} = \frac{17.1}{8} = 2.1375$	B1		awfw 2.13 to 2.14
	$S_d^2 = \frac{1}{7}\left(86.75 - \frac{17.1^2}{8}\right) = 7.17125$	B1		awrt 7.17 $(S_d = 2.67/8)$
	$t_{calc} = \dfrac{2.1375 - 0}{\sqrt{\frac{7.17125}{8}}}$	M1 A1√		use of √ on $\bar{x}$ and S
	$= 2.25$ to 2.26	A1		awfw, accept 2.25 or 2.26
	$\therefore$ evidence, at 5% level, that mean pulse rate decreases	A1√	9	√ on t_{calc} and t_{crit}
	Total		**9**	

MS04 (cont)

Question	Solution	Marks	Total	Comments
5	H_0: biased $\quad H_1$: unbiased	B1		At least H_0
	$\nu = 6 - 1 = 5$	B1		cao
	$\chi^2_{crit} = 11.070$	B1		awrt 11.1

r	o	$P(R = r)$	e	$\frac{(o-e)^2}{e}$	Marks	Comments
1	97	$\frac{1}{3}$	81	3.16	M1	Probs 1 to 5
2	58	$\frac{2}{9}$	54	0.30	M1	Prob $r \geq 6\left(\sum p = 1\right)$
6	29	$\frac{4}{27}$	36	1.36	M1	$e = 243p$
4	19	$\frac{8}{81}$	24	1.04		
5	12	$\frac{16}{243}$	16	1.00	A2,1	2 for all 6 e awrt integer
≥ 6	28	$\frac{32}{243}$	32	0.50		1 for 4 or 5 e awrt integer
$\sum$	243	1	243	7.36		

Question	Solution	Marks	Total	Comments
	$\chi^2_{calc} = \sum \frac{(o-e)^2}{e} = 7.25$ to 7.50	M1 A1		use of awfw
	$\therefore$ no evidence, at 5% level, that die is biased	A1$\sqrt{}$	11	$\sqrt{}$ on χ^2_{calc} and χ^2_{crit}
	Total		**11**	

MS04 (cont)

Question	Solution	Marks	Total	Comments
6(a)	$\mathrm{E}(x) = 1.\mathrm{p} + 2\,\mathrm{pq}^2 + ...)$	M1		Accept Methods
	$= \mathrm{p} = \frac{p}{p^2}\ (1 + 2\mathrm{q} + 3\mathrm{q}^2 + ...)$			
	$= \mathrm{p}\ (1-\mathrm{q})^{-2}$	M1		Using M.G.F.or P.G. F., which may be quoted, without proof, if known.
		A1	3	
	$= \frac{1}{p}$			
(ii)	$E(x^2) = 1.^2 p + 2^2 . pg + 3^2\ pg^2 +$			
	$= \mathrm{p}\ (1 + 4g + 9q^2 + ...)$			
	$= \mathrm{p}\ \{(1 + 3q + 6q^2 + ...)$	M1		
	$+ q\,(1 + 3q + 6q^2 + ...)$	A1		
	$= \mathrm{p}\ \{(1-\mathrm{q})^{-3} + \mathrm{q}(1-\mathrm{q})^{-3}\ \}$	M1		
	$\mathrm{Var}\ (\times) = \frac{1}{p^2} + \frac{1}{p^2}$			
	$= \frac{q}{p2}$	A1	5	
(b)	$\frac{0.4(1-0.6^n)}{(1-0.6)} > 0.999$	M1		use of M2 if written down immediately
	$0.6^n < 0.001$	M1		
	$n > \frac{\log 0.001}{\log 0.6}\ (= 13.5)$	M1		
	14 trials required	A1	4	
	Total		**12**	

Question	Solution	Marks	Total	Comments
7	$\sum x = 24.6 \quad \sum x^2 = 105.68$			
	$\sum y = 27.9 \quad \sum y^2 = 137.99$			
(a)	$H_0 : \sigma_x^2 = \sigma_y^2 \quad H_1 : \sigma_x^2 \neq \sigma_y^2$			
	$\nu_1 = \nu_2 = 5$	B1		cao both
	$F_{crit} = 7.15$	B1		cao (accept $7.15^{-1} = 0.140$)
	$S_x^2 = \frac{1}{5}\left(105.68 - \frac{24.6^2}{6}\right) = 0.964$	B1		awfw 0.960 to 0.970
	$S_y^2 = \frac{1}{5}\left(137.99 - \frac{27.9^2}{6}\right) = 1.651$	B1		awfw 1.650 to 1.652 accept 1.65
	$F_{calc} = \frac{1.651}{0.964} = 1.71 \text{ to } 1.72$	M1 A1		use of awfw
	$\therefore$ No evidence, at 5% level, to reject equal variances	A1√	7	√ on F_{calc} and F_{crit}
(b)(i)	$\hat{\sigma}^2 = \frac{5 \times 0.964 + 5 \times 1.651}{6 + 6 - 2}$	M1		use of
	$=1.305$ to 1.315	A1		awfw
	$\bar{x} - \bar{y} = -0.55$	B1		cao
	$\nu = 6 + 6 - 2 = 10$	B1		cao
	$t_{10}(0.975) = 2.228$	B1		awfw 2.22 to 2.23
	CI is			
	$-0.55 \pm 2.228\sqrt{1.3075}\sqrt{\frac{1}{6} + \frac{1}{6}}$	M1		
	(-2.03 to –2.01, 0.91 to 0.93)	A1	7	awfw
(ii)	No, since the CI contains zero	B1√		√ on CI
		E1√	2	
	Total		**16**	
	TOTAL		**75**	