AQA Maths Statistics 2

Past Paper Pack

2006-2015

General Certificate of Education
January 2006
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Thursday 12 January 2006 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the **blue** 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

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

Answer **all** questions.

1 A study undertaken by Goodhealth Hospital found that the number of patients each month, X, contracting a particular superbug can be modelled by a Poisson distribution with a mean of 1.5 .

(a) (i) Calculate $\mathrm{P}(X = 2)$. *(2 marks)*

(ii) Hence determine the probability that exactly 2 patients will contract this superbug in each of three consecutive months. *(2 marks)*

(b) (i) Write down the distribution of Y, the number of patients contracting this superbug in a given 6-month period. *(1 mark)*

(ii) Find the probability that at least 12 patients will contract this superbug during a given 6-month period. *(2 marks)*

(c) State **two** assumptions implied by the use of a Poisson model for the number of patients contracting this superbug. *(2 marks)*

2 Year 12 students at Newstatus School choose to participate in one of four sports during the Spring term.

The students' choices are summarised in the table.

	Squash	**Badminton**	**Archery**	**Hockey**	**Total**
Male	5	16	30	19	70
Female	4	20	33	53	110
Total	9	36	63	72	180

(a) Use a χ^2 test, at the 5% level of significance, to determine whether the choice of sport is independent of gender. *(10 marks)*

(b) Interpret your result in part (a) as it relates to students choosing hockey. *(2 marks)*

3 The time, T minutes, that parents have to wait before seeing a mathematics teacher at a school parents' evening can be modelled by a normal distribution with mean μ and standard deviation σ.

At a recent parents' evening, a random sample of 9 parents was asked to record the times that they waited before seeing a mathematics teacher.

The times, in minutes, are

5 12 10 8 7 6 9 7 8

(a) Construct a 90% confidence interval for μ. *(7 marks)*

(b) Comment on the headteacher's claim that the mean time that parents have to wait before seeing a mathematics teacher is 5 minutes. *(2 marks)*

4 (a) A random variable X has probability density function defined by

$$f(x) = \begin{cases} k & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

(i) Show that $k = \dfrac{1}{b-a}$. *(1 mark)*

(ii) Prove, using integration, that $E(X) = \frac{1}{2}(a+b)$. *(4 marks)*

(b) The error, X grams, made when a shopkeeper weighs out loose sweets can be modelled by a rectangular distribution with the following probability density function:

$$f(x) = \begin{cases} k & -2 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

(i) Write down the value of the mean, μ, of X. *(1 mark)*

(ii) Evaluate the standard deviation, σ, of X. *(2 marks)*

(iii) Hence find $P\left(X < \dfrac{2-\mu}{\sigma}\right)$. *(3 marks)*

5 The Globe Express agency organises trips to the theatre. The cost, £X, of these trips can be modelled by the following probability distribution:

x	40	45	55	74
P($X = x$)	0.30	0.24	0.36	0.10

(a) Calculate the mean and standard deviation of X. *(4 marks)*

(b) For special celebrity charity performances, Globe Express increases the cost of the trips to £Y, where

$$Y = 10X + 250$$

Determine the mean and standard deviation of Y. *(2 marks)*

6 In previous years, the marks obtained in a French test by students attending Topnotch College have been modelled satisfactorily by a normal distribution with a mean of 65 and a standard deviation of 9.

Teachers in the French department at Topnotch College suspect that this year their students are, on average, underachieving.

In order to investigate this suspicion, the teachers selected a random sample of 35 students to take the French test and found that their mean score was 61.5.

(a) Investigate, at the 5% level of significance, the teachers' suspicion. *(6 marks)*

(b) Explain, in the context of this question, the meaning of a Type I error. *(2 marks)*

7 Engineering work on the railway network causes an increase in the journey time of commuters travelling into work each morning.

The increase in journey time, T hours, is modelled by a continuous random variable with probability density function

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

(a) Show that $E(T) = \frac{8}{15}$. *(3 marks)*

(b) (i) Find the cumulative distribution function, $F(t)$, for $0 \leqslant t \leqslant 1$. *(2 marks)*

(ii) Hence, or otherwise, for a commuter selected at random, find

$$P(\text{mean} < T < \text{median})$$ *(5 marks)*

8 Bottles of sherry nominally contain 1000 millilitres. After the introduction of a new method of filling the bottles, there is a suspicion that the mean volume of sherry in a bottle has changed.

In order to investigate this suspicion, a random sample of 12 bottles of sherry is taken and the volume of sherry in each bottle is measured.

The volumes, in millilitres, of sherry in these bottles are found to be

996	1006	1009	999	1007	1003
998	1010	997	996	1008	1007

Assuming that the volume of sherry in a bottle is normally distributed, investigate, at the 5% level of significance, whether the mean volume of sherry in a bottle differs from 1000 millilitres. *(10 marks)*

END OF QUESTIONS

General Certificate of Education
June 2006
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Wednesday 24 May 2006 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the **blue** 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 The number of A-grades, X, achieved in total by students at Lowkey School in their Mathematics examinations each year can be modelled by a Poisson distribution with a mean of 3.

(a) Determine the probability that, during a 5-year period, students at Lowkey School achieve a total of more than 18 A-grades in their Mathematics examinations. *(3 marks)*

(b) The number of A-grades, Y, achieved in total by students at Lowkey School in their English examinations each year can be modelled by a Poisson distribution with a mean of 7.

(i) Determine the probability that, during a year, students at Lowkey School achieve a total of fewer than 15 A-grades in their Mathematics and English examinations. *(3 marks)*

(ii) What assumption did you make in answering part (b)(i)? *(1 mark)*

2 The weights of lions kept in captivity at Wildcat Safari Park are normally distributed.

The weights, in kilograms, of a random sample of five lions were recorded as

46 48 57 49 54

(a) Construct a 95% confidence interval for the mean weight of lions kept in captivity at Wildcat Safari Park. *(6 marks)*

(b) State the probability that this confidence interval does **not** contain the mean weight of lions kept in captivity at Wildcat Safari Park. *(1 mark)*

3 Morecrest football team always scores at least one goal but never scores more than four goals in each game. The number of goals, R, scored in each game by the team can be modelled by the following probability distribution.

r	1	2	3	4
$\mathrm{P}(R = r)$	$\frac{7}{16}$	$\frac{5}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

(a) Calculate exact values for the mean and variance of R. *(4 marks)*

(b) Next season the team will play 32 games. They expect to win 90% of the games in which they score at least three goals, half of the games in which they score exactly two goals and 20% of the games in which they score exactly one goal.

Find, for next season:

(i) the number of games in which they expect to score at least three goals; *(1 mark)*

(ii) the number of games that they expect to win. *(2 marks)*

4 It is claimed that the area within which a school is situated affects the age profile of the staff employed at that school. In order to investigate this claim, the age profiles of staff employed at two schools with similar academic achievements are compared.

Academia High School, situated in a rural community, employs 120 staff whilst Best Manor Grammar School, situated in an inner-city community, employs 80 staff.

The **percentage** of staff within each age group, for each school, is given in the table.

Age	**Academia High School**	**Best Manor Grammar School**
22–34	17.5	40.0
35–39	60.0	45.0
40–59	22.5	15.0

(a) (i) Form the data into a contingency table suitable for analysis using a χ^2 distribution. *(2 marks)*

(ii) Use a χ^2 test, at the 1% level of significance, to determine whether there is an association between the age profile of the staff employed and the area within which the school is situated. *(9 marks)*

(b) Interpret your result in part (a)(ii) as it relates to the 22–34 age group. *(2 marks)*

Turn over ▶

5 (a) The continuous random variable X follows a rectangular distribution with probability density function defined by

$$f(x) = \begin{cases} \dfrac{1}{b} & 0 \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

(i) Write down $\mathrm{E}(X)$. *(1 mark)*

(ii) Prove, using integration, that

$$\mathrm{Var}(X) = \tfrac{1}{12}b^2$$ *(5 marks)*

(b) At an athletics meeting, the error, in seconds, made in recording the time taken to complete the 10 000 metres race may be modelled by the random variable T, having the probability density function

$$f(t) = \begin{cases} 5 & -0.1 \leqslant t \leqslant 0.1 \\ 0 & \text{otherwise} \end{cases}$$

Calculate $\mathrm{P}(|T| > 0.02)$. *(3 marks)*

6 The lifetime, X hours, of Everwhite camera batteries is normally distributed. The manufacturer claims that the mean lifetime of these batteries is 100 hours.

(a) The members of a photography club suspect that the batteries do not last as long as is claimed by the manufacturer. In order to investigate their suspicion, the members test a random sample of five of these batteries and find the lifetimes, in hours, to be as follows:

85 92 100 95 99

Test the members' suspicion at the 5% level of significance. *(9 marks)*

(b) The manufacturer, believing that the mean lifetime of these batteries has not changed from 100 hours, decides to determine the lifetime, x hours, of each of a random sample of 80 Everwhite camera batteries. The manufacturer obtains the following results, where $\bar{x}$ denotes the sample mean:

$$\sum x = 8080 \quad \text{and} \quad \sum (x - \bar{x})^2 = 6399$$

Test the manufacturer's belief at the 5% level of significance. *(8 marks)*

7 The continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} \frac{1}{5}(2x+1) & 0 \leqslant x \leqslant 1 \\ \frac{1}{15}(4-x)^2 & 1 < x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(2 marks)*

(b) (i) Show that the cumulative distribution function, $F(x)$, for $0 \leqslant x \leqslant 1$ is

$$F(x) = \frac{1}{5}x(x+1)$$ *(3 marks)*

(ii) Hence write down the value of $P(X \leqslant 1)$. *(1 mark)*

(iii) Find the value of x for which $P(X \geqslant x) = \frac{17}{20}$. *(5 marks)*

(iv) Find the lower quartile of the distribution. *(4 marks)*

END OF QUESTIONS

General Certificate of Education
January 2007
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Friday 12 January 2007 9.00 am to 10.30 am

For this paper you must have:
- an 8-page answer book
- the **blue** 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 Alan's journey time, in minutes, to travel home from work each day is known to be normally distributed with mean μ.

Alan records his journey time, in minutes, on a random sample of 8 days as being

36 38 39 40 50 35 36 42

Construct a 95% confidence interval for μ. *(5 marks)*

2 The number of computers, A, bought during one day from the Amplebuy computer store can be modelled by a Poisson distribution with a mean of 3.5.

The number of computers, B, bought during one day from the Bestbuy computer store can be modelled by a Poisson distribution with a mean of 5.0.

(a) (i) Calculate $\mathrm{P}(A = 4)$. *(2 marks)*

(ii) Determine $\mathrm{P}(B \leqslant 6)$. *(1 mark)*

(iii) Find the probability that a total of fewer than 10 computers is bought from these two stores on one particular day. *(3 marks)*

(b) Calculate the probability that a total of fewer than 10 computers is bought from these two stores on at least 4 out of 5 consecutive days. *(3 marks)*

(c) The numbers of computers bought from the Choicebuy computer store over a 10-day period are recorded as

8 12 6 6 9 15 10 8 6 12

(i) Calculate the mean and variance of these data. *(2 marks)*

(ii) State, giving a reason based on your results in part (c)(i), whether or not a Poisson distribution provides a suitable model for these data. *(2 marks)*

3 The handicap committee of a golf club has indicated that the mean score achieved by the club's members in the past was 85.9.

A group of members believes that recent changes to the golf course have led to a change in the mean score achieved by the club's members and decides to investigate this belief.

A random sample of the scores, x, of 100 club members was taken and is summarised by

$$\sum x = 8350 \quad \text{and} \quad \sum (x - \bar{x})^2 = 15\,321$$

where $\bar{x}$ denotes the sample mean.

Test, at the 5% level of significance, the group's belief that the mean score of 85.9 has changed. *(8 marks)*

4 The number of fish, X, caught by Pearl when she goes fishing can be modelled by the following discrete probability distribution:

$\boldsymbol{x}$	1	2	3	4	5	6	$\geqslant 7$
$\mathbf{P}(\boldsymbol{X = x})$	0.01	0.05	0.14	0.30	k	0.12	0

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

(b) Find:

(i) $\mathrm{E}(X)$; *(1 mark)*

(ii) $\mathrm{Var}(X)$. *(3 marks)*

(c) When Pearl sells her fish, she earns a profit, in pounds, given by

$$Y = 5X + 2$$

Find:

(i) $\mathrm{E}(Y)$; *(1 mark)*

(ii) the standard deviation of Y. *(3 marks)*

5 Jasmine's French teacher states that a homework assignment should take, on average, 30 minutes to complete.

Jasmine believes that he is understating the mean time that the assignment takes to complete and so decides to investigate. She records the times, in minutes, that it takes for a random sample of 10 students to complete the French assignment, with the following results:

29	33	36	42	30	28	31	34	37	35

(a) Test, at the 1% level of significance, Jasmine's belief that her French teacher has understated the mean time that it should take to complete the homework assignment. *(7 marks)*

(b) State an assumption that you must make in order for the test used in part (a) to be valid. *(1 mark)*

6 The waiting time, T minutes, before being served at a local newsagents can be modelled by a continuous random variable with probability density function

$$f(t) = \begin{cases} \frac{3}{8}t^2 & 0 \leqslant t < 1 \\ \frac{1}{16}(t+5) & 1 \leqslant t \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) For a customer selected at random, calculate $P(T \geqslant 1)$. *(2 marks)*

(c) (i) Show that the cumulative distribution function for $1 \leqslant t \leqslant 3$ is given by

$$F(t) = \frac{1}{32}(t^2 + 10t - 7)$$ *(5 marks)*

(ii) Hence find the median waiting time. *(4 marks)*

7 A statistics unit is required to determine whether or not there is an association between students' performances in mathematics at Key Stage 3 and at GCE.

A survey of the results of 500 students showed the following information:

		GCE Grade				
		A	**B**	**C**	**Below C**	**Total**
Key Stage 3 Level	**8**	60	55	47	43	205
	7	55	32	31	26	144
	6	40	38	35	38	151
	Total	155	125	113	107	500

(a) Use a χ^2 test at the 10% level of significance to determine whether there is an association between students' performances in mathematics at Key Stage 3 and at GCE. *(9 marks)*

(b) Comment on the number of students who gained a grade A at GCE having gained a level 7 at Key Stage 3. *(1 mark)*

8 The continuous random variable X has the cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0 & x \leqslant -4 \\ \dfrac{x+4}{9} & -4 \leqslant x \leqslant 5 \\ 1 & x \geqslant 5 \end{cases}$$

(a) Determine the probability density function, $\mathrm{f}(x)$, of X. *(2 marks)*

(b) Sketch the graph of f. *(2 marks)*

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

(d) Evaluate the mean and variance of X. *(2 marks)*

END OF QUESTIONS

General Certificate of Education
June 2007
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Tuesday 5 June 2007 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the **blue** 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 Two groups of patients, suffering from the same medical condition, took part in a clinical trial of a new drug. One of the groups was given the drug whilst the other group was given a placebo, a drug that has no physical effect on their medical condition.

The table shows the number of patients in each group and whether or not their condition improved.

	Placebo	**Drug**
Condition improved	20	46
Condition did not improve	55	29

Conduct a χ^2 test, at the 5% level of significance, to determine whether the condition of the patients at the conclusion of the trial is associated with the treatment that they were given. *(10 marks)*

2 The number of telephone calls per day, X, received by Candice may be modelled by a Poisson distribution with mean 3.5 .

The number of e-mails per day, Y, received by Candice may be modelled by a Poisson distribution with mean 6.0 .

(a) For any particular day, find:

(i) $\mathrm{P}(X = 3)$; *(2 marks)*

(ii) $\mathrm{P}(Y \geqslant 5)$. *(2 marks)*

(b) (i) Write down the distribution of T, the total number of telephone calls and e-mails per day received by Candice. *(1 mark)*

(ii) Determine $\mathrm{P}(7 \leqslant T \leqslant 10)$. *(3 marks)*

(iii) Hence calculate the probability that, on each of three consecutive days, Candice will receive a total of at least 7 but at most 10 telephone calls and e-mails. *(2 marks)*

3 David is the professional coach at the golf club where Becki is a member. He claims that, after having a series of lessons with him, the mean number of putts that Becki takes per round of golf will reduce from her present mean of 36.

After having the series of lessons with David, Becki decides to investigate his claim.

She therefore records, for each of a random sample of 50 rounds of golf, the number of putts, x, that she takes to complete the round. Her results are summarised below, where $\bar{x}$ denotes the sample mean.

$$\sum x = 1730 \qquad \text{and} \qquad \sum (x - \bar{x})^2 = 784$$

Using a z-test and the 1% level of significance, investigate David's claim. *(8 marks)*

4 Students are each asked to measure the distance between two points to the nearest tenth of a metre.

(a) Given that the rounding error, X metres, in these measurements has a rectangular distribution, explain why its probability density function is

$$\mathrm{f}(x) = \begin{cases} 10 & -0.05 < x \leqslant 0.05 \\ 0 & \text{otherwise} \end{cases}$$

(3 marks)

(b) Calculate $\mathrm{P}(-0.01 < X < 0.02)$. *(2 marks)*

(c) Find the mean and the standard deviation of X. *(2 marks)*

Turn over for the next question

5 Members of a residents' association are concerned about the speeds of cars travelling through their village. They decide to record the speed, in mph, of each of a random sample of 10 cars travelling through their village, with the following results:

33	27	34	30	48	35	34	33	43	39

(a) Construct a 99% confidence interval for μ, the mean speed of cars travelling through the village, stating any assumption that you make. *(7 marks)*

(b) Comment on the claim that a 30 mph speed limit is being adhered to by most motorists. *(3 marks)*

6 The continuous random variable X has the probability density function given by

$$\mathrm{f}(x) = \begin{cases} 3x^2 & 0 < x \leqslant 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Determine:

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

(ii) $\mathrm{Var}\left(\dfrac{1}{X}\right)$. *(4 marks)*

(b) Hence, or otherwise, find the mean and the variance of $\left(\dfrac{5+2X}{X}\right)$. *(5 marks)*

7 On a multiple choice examination paper, each question has five alternative answers given, only one of which is correct. For each question, candidates gain 4 marks for a correct answer but lose 1 mark for an incorrect answer.

(a) James guesses the answer to each question.

(i) Copy and complete the following table for the probability distribution of X, the number of marks obtained by James for each question.

x	4	-1
$\mathbf{P}(X=x)$		

(1 mark)

(ii) Hence find $\mathrm{E}(X)$. *(2 marks)*

(b) Karen is able to eliminate two of the incorrect answers from the five alternative answers given for each question before guessing the answer from those remaining.

Given that the examination paper contains 24 questions, calculate Karen's expected total mark. *(4 marks)*

8 A jam producer claims that the mean weight of jam in a jar is 230 grams.

(a) A random sample of 8 jars is selected and the weight of jam in each jar is determined. The results, in grams, are

220 228 232 219 221 223 230 229

Assuming that the weight of jam in a jar is normally distributed, test, at the 5% level of significance, the jam producer's claim. *(9 marks)*

(b) It is later discovered that the mean weight of jam in a jar is indeed 230 grams.

Indicate whether a Type I error, a Type II error or neither has occurred in carrying out the hypothesis test in part (a). Give a reason for your answer. *(2 marks)*

END OF QUESTIONS

General Certificate of Education
January 2008
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Friday 11 January 2008 9.00 am to 10.30 am

For this paper you must have:
- an 8-page answer book
- the blue 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 David claims that customers have to queue at a supermarket checkout for more than 5 minutes, on average.

The queuing times, x minutes, of 40 randomly selected customers result in $\bar{x} = 5.5$ and $s^2 = 1.31$.

Investigate, at the 1% level of significance, David's claim. *(6 marks)*

2 A new information technology centre is advertising places on its one-week residential computer courses.

(a) The number of places, X, booked each week on the publishing course may be modelled by a Poisson distribution with a mean of 9.0 .

(i) State the standard deviation of X. *(1 mark)*

(ii) Calculate $\mathrm{P}(6 < X < 12)$. *(3 marks)*

(b) The number of places booked each week on the web design course may be modelled by a Poisson distribution with a mean of 2.5 .

(i) Write down the distribution for T, the **total** number of places booked each week on the publishing and web design courses. *(1 mark)*

(ii) Hence calculate the probability that, during a given week, a total of fewer than 2 places are booked. *(3 marks)*

(c) The number of places booked on the database course during each of a random sample of 10 weeks is as follows:

14	15	8	16	18	4	10	12	15	8

By calculating appropriate numerical measures, state, with a reason, whether or not the Poisson distribution $\mathrm{Po}(12.0)$ could provide a suitable model for the number of places booked each week on the database course. *(3 marks)*

3 (a) The continuous random variable T follows a rectangular distribution with probability density function given by

$$f(t) = \begin{cases} k & -a \leqslant t \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

(i) Express k in terms of a and b. *(1 mark)*

(ii) Prove, using integration, that $\mathrm{E}(T) = \frac{1}{2}(b-a)$. *(4 marks)*

(b) The error, in minutes, made by a commuter when estimating the journey time by train into London may be modelled by the random variable T with probability density function

$$f(t) = \begin{cases} \frac{1}{10} & -4 \leqslant t \leqslant 6 \\ 0 & \text{otherwise} \end{cases}$$

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

(ii) Calculate $\mathrm{P}(T < -3 \text{ or } T > 3)$. *(2 marks)*

4 A speed camera was used to measure the speed, V mph, of John's serves during a tennis singles championship.

For 10 randomly selected serves,

$$\sum v = 1179 \qquad \text{and} \qquad \sum (v - \bar{v})^2 = 1014.9$$

where $\bar{v}$ is the sample mean.

(a) Construct a 99% confidence interval for the mean speed of John's serves at this tennis championship, stating any assumption that you make. *(7 marks)*

(b) Hence comment on John's claim that, at this championship, he consistently served at speeds in excess of 130 mph. *(1 mark)*

5 A discrete random variable X has the probability distribution

$$\mathrm{P}(X = x) = \begin{cases} \dfrac{x}{20} & x = 1, 2, 3, 4, 5 \\[2ex] \dfrac{x}{24} & x = 6 \\[2ex] 0 & \text{otherwise} \end{cases}$$

(a) Calculate $\mathrm{P}(X \geqslant 5)$. *(2 marks)*

(b) (i) Show that $\mathrm{E}\left(\dfrac{1}{X}\right) = \dfrac{7}{24}$. *(2 marks)*

(ii) Hence, or otherwise, show that $\mathrm{Var}\left(\dfrac{1}{X}\right) = 0.036$, correct to three decimal places. *(3 marks)*

(c) Calculate the mean and the variance of A, the area of rectangles having sides of length $X + 3$ and $\dfrac{1}{X}$. *(5 marks)*

6 A survey is carried out in an attempt to determine whether the salary achieved by the age of 30 is associated with having had a university education.

The results of this survey are given in the table.

	Salary < £30 000	**Salary $\geqslant$ £30 000**	**Total**
University education	52	78	130
No university education	63	57	120
Total	115	135	250

(a) Use a χ^2 test, at the 10% level of significance, to determine whether the salary achieved by the age of 30 is associated with having had a university education. *(9 marks)*

(b) What do you understand by a Type I error in this context? *(2 marks)*

7 The waiting time, X minutes, for fans to gain entrance to see an event may be modelled by a continuous random variable having the distribution function defined by

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x & 0 \leqslant x \leqslant 1 \\ \frac{1}{54}(x^3 - 12x^2 + 48x - 10) & 1 \leqslant x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

(a) (i) Sketch the graph of F. *(4 marks)*

(ii) Explain why the value of q_1, the lower quartile of X, is $\frac{1}{2}$. *(2 marks)*

(iii) Show that the upper quartile, q_3, satisfies $1.6 < q_3 < 1.7$. *(3 marks)*

(b) The probability density function of X is defined by

$$\mathrm{f}(x) = \begin{cases} \alpha & 0 \leqslant x \leqslant 1 \\ \beta(x-4)^2 & 1 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(i) Show that the **exact** values of α and β are $\frac{1}{2}$ and $\frac{1}{18}$ respectively. *(5 marks)*

(ii) Hence calculate $\mathrm{E}(X)$. *(5 marks)*

END OF QUESTIONS

General Certificate of Education
June 2008
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Monday 2 June 2008 9.00 am to 10.30 am

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use black ink or black 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.
- Unit Statistics 2B has a **written paper only**.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 It is thought that the incidence of asthma in children is associated with the volume of traffic in the area where they live.

Two surveys of children were conducted: one in an area where the volume of traffic was heavy and the other in an area where the volume of traffic was light.

For each area, the table shows the number of children in the survey who had asthma and the number who did not have asthma.

	Asthma	**No asthma**	**Total**
Heavy traffic	52	58	110
Light traffic	28	62	90
Total	80	120	200

(a) Use a χ^2 test, at the 5% level of significance, to determine whether the incidence of asthma in children is associated with the volume of traffic in the area where they live. *(8 marks)*

(b) Comment on the number of children in the survey who had asthma, given that they lived in an area where the volume of traffic was heavy. *(1 mark)*

2 (a) The number of telephone calls, X, received per hour for Dr Able may be modelled by a Poisson distribution with mean 6.

Determine $\mathrm{P}(X = 8)$. *(2 marks)*

(b) The number of telephone calls, Y, received per hour for Dr Bracken may be modelled by a Poisson distribution with mean λ and standard deviation 3.

(i) Write down the value of λ. *(1 mark)*

(ii) Determine $\mathrm{P}(Y > \lambda)$. *(2 marks)*

(c) (i) Assuming that X and Y are independent Poisson variables, write down the distribution of the **total** number of telephone calls received per hour for Dr Able and Dr Bracken. *(1 mark)*

(ii) Determine the probability that a total of at most 20 telephone calls will be received during any one-hour period. *(1 mark)*

(iii) The total number of telephone calls received during each of 6 one-hour periods is to be recorded. Calculate the probability that a total of at least 21 telephone calls will be received during exactly 4 of these one-hour periods. *(3 marks)*

3 Alan's company produces packets of crisps. The standard deviation of the weight of a packet of crisps is known to be 2.5 grams.

Alan believes that, due to the extra demand on the production line at a busy time of the year, the mean weight of packets of crisps is not equal to the target weight of 34.5 grams.

In an experiment set up to investigate Alan's belief, the weights of a random sample of 50 packets of crisps were recorded. The mean weight of this sample is 35.1 grams.

Investigate Alan's belief at the 5% level of significance. *(6 marks)*

4 The delay, in hours, of certain flights from Australia may be modelled by the continuous random variable T, with probability density function

$$\mathrm{f}(t) = \begin{cases} \frac{2}{15}t & 0 \leqslant t \leqslant 3 \\ 1 - \frac{1}{5}t & 3 \leqslant t \leqslant 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Calculate:

(i) $\mathrm{P}(T \leqslant 2)$; *(2 marks)*

(ii) $\mathrm{P}(2 < T < 4)$. *(3 marks)*

(c) Determine $\mathrm{E}(T)$. *(4 marks)*

5 The weight of fat in a digestive biscuit is known to be normally distributed.

Pat conducted an experiment in which she measured the weight of fat, x grams, in each of a random sample of 10 digestive biscuits, with the following results:

$$\sum x = 31.9 \qquad \text{and} \qquad \sum (x - \bar{x})^2 = 1.849$$

(a) (i) Construct a 99% confidence interval for the mean weight of fat in digestive biscuits. *(5 marks)*

(ii) Comment on a claim that the mean weight of fat in digestive biscuits is 3.5 grams. *(2 marks)*

(b) If 200 such 99% confidence intervals were constructed, how many would you expect **not** to contain the population mean? *(1 mark)*

6 The management of the Wellfit gym claims that the mean cholesterol level of those members who have held membership of the gym for more than one year is 3.8.

A local doctor believes that the management's claim is too low and investigates by measuring the cholesterol levels of a random sample of 7 such members of the Wellfit gym, with the following results:

4.2 4.3 3.9 3.8 3.6 4.8 4.1

Is there evidence, at the 5% level of significance, to justify the doctor's belief that the mean cholesterol level is greater than the management's claim? State any assumption that you make. *(8 marks)*

7 (a) The number of text messages, N, sent by Peter each month on his mobile phone never exceeds 40.

When $0 \leqslant N \leqslant 10$, he is charged for 5 messages.
When $10 < N \leqslant 20$, he is charged for 15 messages.
When $20 < N \leqslant 30$, he is charged for 25 messages.
When $30 < N \leqslant 40$, he is charged for 35 messages.

The number of text messages, Y, that Peter is charged for each month has the following probability distribution:

y	5	15	25	35
$\mathbf{P}(Y = y)$	0.1	0.2	0.3	0.4

(i) Calculate the mean and the standard deviation of Y. *(4 marks)*

(ii) The Goodtime phone company makes a total charge for text messages, C pence, each month given by

$$C = 10Y + 5$$

Calculate $\mathrm{E}(C)$. *(1 mark)*

(b) The number of text messages, X, sent by Joanne each month on her mobile phone is such that

$$\mathrm{E}(X) = 8.35 \qquad \text{and} \qquad \mathrm{E}(X^2) = 75.25$$

The Newtime phone company makes a total charge for text messages, T pence, each month given by

$$T = 0.4X + 250$$

Calculate $\mathrm{Var}(T)$. *(4 marks)*

8 The continuous random variable X has cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0 & x < -1 \\ \dfrac{x+1}{k+1} & -1 \leqslant x \leqslant k \\ 1 & x > k \end{cases}$$

where k is a positive constant.

(a) Find, in terms of k, an expression for $\mathrm{P}(X < 0)$. *(2 marks)*

(b) Determine an expression, in terms of k, for the lower quartile, q_1. *(3 marks)*

(c) Show that the probability density function of X is defined by

$$\mathrm{f}(x) = \begin{cases} \dfrac{1}{k+1} & -1 \leqslant x \leqslant k \\ 0 & \text{otherwise} \end{cases}$$ *(2 marks)*

(d) Given that $k = 11$:

(i) sketch the graph of f; *(2 marks)*

(ii) determine $\mathrm{E}(X)$ and $\mathrm{Var}(X)$; *(2 marks)*

(iii) show that $\mathrm{P}(q_1 < X < \mathrm{E}(X)) = 0.25$. *(2 marks)*

END OF QUESTIONS

General Certificate of Education
January 2009
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Thursday 29 January 2009 9.00 am to 10.30 am

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use black ink or black 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 Fortune High School gave its students a wider choice of subjects to study. The table shows the number of students, of each gender, who chose to study each of the additional subjects during the school year 2007/08.

	Bulgarian	**Climate Change**	**Finance**	**Polish**
Male	7	31	25	40
Female	2	24	22	19

Assuming that these data form a random sample, use a χ^2 test, at the 10% level of significance, to test whether the choice of these subjects is independent of gender.

(11 marks)

2 A group of estate agents in a particular area claimed that, after the introduction of a new search procedure at the Land Registry, the mean completion time for the purchase of a house in the area had not changed from 8 weeks.

(a) A random sample of 9 house purchases in the area revealed that their completion times, in weeks, were as follows.

6 7 10 12 9 11 7 8 14

Assuming that completion times in the area are normally distributed with standard deviation 2.5 weeks, test, at the 5% level of significance, the group's claim. *(7 marks)*

(b) It was subsequently discovered that, after the introduction of the new search procedure at the Land Registry, the mean completion time for the purchase of a house in the area remained at 8 weeks.

Indicate whether a Type I error, a Type II error or neither has occurred in carrying out your hypothesis test in part (a). Give a reason for your answer. *(2 marks)*

3 Joe owns two garages, Acefit and Bestjob, each specialising in the fitting of the latest satellite navigation device.

The daily demand, X, for the device at Acefit garage may be modelled by a Poisson distribution with mean 3.6.

The daily demand, Y, for the device at Bestjob garage may be modelled by a Poisson distribution with mean 4.4.

(a) Calculate:

(i) $\mathrm{P}(X \leqslant 3)$; *(1 mark)*

(ii) $\mathrm{P}(Y = 5)$. *(2 marks)*

(b) The total daily demand for the device at Joe's two garages is denoted by T.

(i) Write down the distribution of T, stating any assumption that you make. *(2 marks)*

(ii) Determine $\mathrm{P}(6 < T < 12)$. *(3 marks)*

(iii) Calculate the probability that the total demand for the device will exceed 14 on each of two consecutive days. Give your answer to one significant figure. *(4 marks)*

(iv) Determine the minimum number of devices that Joe should have in stock if he is to meet his total demand on at least 99% of days. *(2 marks)*

Turn over for the next question

4 The continuous random variable X has the cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0 & x < -c \\ \dfrac{x+c}{4c} & -c \leqslant x \leqslant 3c \\ 1 & x > 3c \end{cases}$$

where c is a positive constant.

(a) Determine $\mathrm{P}\left(-\dfrac{3c}{4} < X < \dfrac{3c}{4}\right)$. *(2 marks)*

(b) Show that the probability density function, $\mathrm{f}(x)$, of X is

$$\mathrm{f}(x) = \begin{cases} \dfrac{1}{4c} & -c \leqslant x \leqslant 3c \\ 0 & \text{otherwise} \end{cases}$$

(2 marks)

(c) Hence, or otherwise, find expressions, in terms of c, for:

(i) $\mathrm{E}(X)$; *(1 mark)*

(ii) $\mathrm{Var}(X)$. *(1 mark)*

5 Jane, who supplies fruit to a jam manufacturer, knows that the weight of fruit in boxes that she sends to the manufacturer can be modelled by a normal distribution with unknown mean, μ grams, and unknown standard deviation, σ grams.

Jane selects a random sample of 16 boxes and, using the t-distribution, calculates correctly that a 98% confidence interval for μ is (70.65, 80.35).

(a) (i) Show that the sample mean is 75.5 grams. *(1 mark)*

(ii) Find the width of the confidence interval. *(1 mark)*

(iii) Calculate an estimate of the standard error of the mean. *(3 marks)*

(iv) Hence, or otherwise, show that an unbiased estimate of σ^2 is 55.6, correct to three significant figures. *(2 marks)*

(b) Jane decides that the width of the 98% confidence interval is too large.

Construct a 95% confidence interval for μ, based on her sample of 16 boxes. *(2 marks)*

(c) Jane is informed that the manufacturer would prefer the confidence interval to have a width of at most 5 grams.

(i) Write down a confidence interval for μ, again based on Jane's sample of 16 boxes, which has a width of 5 grams. *(1 mark)*

(ii) Determine the percentage confidence level for your interval in part (c)(i). *(3 marks)*

Turn over for the next question

6 A small supermarket has a total of four checkouts, at least one of which is always staffed. The probability distribution for R, the number of checkouts that are staffed at any given time, is

$$\mathrm{P}(R=r)=\begin{cases}\dfrac{2}{3}\left(\dfrac{1}{3}\right)^{r-1} & r=1,\,2,\,3\\ k & r=4\end{cases}$$

(a) Show that $k=\frac{1}{27}$. *(2 marks)*

(b) Find the probability that, at any given time, there will be at least 3 checkouts that are staffed. *(1 mark)*

(c) It is suggested that the total number of customers, C, that can be served at the checkouts per hour may be modelled by

$$C=27R+5$$

Find:

(i) $\mathrm{E}(C)$; *(3 marks)*

(ii) the standard deviation of C. *(4 marks)*

7 The continuous random variable X has the probability density function given by

$$\mathrm{f}(x)=\begin{cases}\dfrac{1}{16}x^3 & 0\leqslant x\leqslant 2\\ \dfrac{1}{6}(5-x) & 2\leqslant x\leqslant 5\\ 0 & \text{otherwise}\end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Prove that the cumulative distribution function of X for $2\leqslant x\leqslant 5$ can be written in the form

$$\mathrm{F}(x)=1-\frac{1}{12}(5-x)^2$$

(4 marks)

(c) Hence, or otherwise, determine $\mathrm{P}(X\geqslant 3\mid X\leqslant 4)$. *(5 marks)*

END OF QUESTIONS

General Certificate of Education
June 2009
Advanced Level Examination

MATHEMATICS **MS2B**
Unit Statistics 2B

Monday 15 June 2009 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed: 1 hour 30 minutes

Instructions
- Use black ink or black 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.
- Show all necessary working; 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.
- The marks for questions are shown in brackets.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 A machine fills bottles with bleach. The volume, in millilitres, of bleach dispensed by the machine into a bottle may be modelled by a normal distribution with mean μ and standard deviation 8.

A recent inspection indicated that the value of μ was 768. Yvonne, the machine's operator, claims that this value has not subsequently changed.

Zara, the quality control supervisor, records the volume of bleach in each of a random sample of 18 bottles filled by the machine and calculates their mean to be 764.8 ml.

Test, at the 5% level of significance, Yvonne's claim that the mean volume of bleach dispensed by the machine has not changed from 768 ml. *(6 marks)*

2 John works from home. The number of business letters, X, that he receives on a weekday may be modelled by a Poisson distribution with mean 5.0.

The number of private letters, Y, that he receives on a weekday may be modelled by a Poisson distribution with mean 1.5.

(a) Find, for a given weekday:

(i) $\mathrm{P}(X < 4)$; *(2 marks)*

(ii) $\mathrm{P}(Y = 4)$. *(2 marks)*

(b) (i) Assuming that X and Y are independent random variables, determine the probability that, on a given weekday, John receives a **total** of more than 5 business and private letters. *(3 marks)*

(ii) Hence calculate the probability that John receives a **total** of more than 5 business and private letters on at least 7 out of 8 given weekdays. *(3 marks)*

(c) The numbers of letters received by John's neighbour, Brenda, on 10 consecutive weekdays are

15	8	14	7	6	8	2	8	9	3

(i) Calculate the mean and the variance of these data. *(2 marks)*

(ii) State, giving a reason based on your answers to part (c)(i), whether or not a Poisson distribution might provide a suitable model for the number of letters received by Brenda on a weekday. *(2 marks)*

3 A sample survey, conducted to determine the attitudes of residents to a proposed reorganisation of local schools, gave the following results.

		Against reorganisation	**Not against reorganisation**
Age of resident	**16–17**	9	2
	18–21	17	10
	22–49	115	90
	50–65	41	34
	Over 65	3	4

Use a χ^2 test, at the 5% level of significance, to determine whether there is an association between the ages of residents and their attitudes to the proposed reorganisation of local schools. *(12 marks)*

4 The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{2} & 0 \leqslant x \leqslant 1 \\ \frac{3-x}{4} & 1 \leqslant x \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Explain why the value of η, the median of X, is 1. *(2 marks)*

(c) Show that the value of μ, the mean of X, is $\frac{13}{12}$. *(4 marks)*

(d) Find $P(X < 3\mu - \eta)$. *(3 marks)*

Turn over for the next question

5 Joanne has 10 identically-shaped discs, of which 1 is blue, 2 are green, 3 are yellow and 4 are red. She places the 10 discs in a bag and asks her friend David to play a game by selecting, at random and without replacement, two discs from the bag.

(a) Show that:

(i) the probability that the two discs selected are the same colour is $\frac{2}{9}$; *(2 marks)*

(ii) the probability that exactly one of the two discs selected is blue is $\frac{1}{5}$. *(2 marks)*

(b) Using the discs, Joanne plays the game with David, under the following conditions:

If the two discs selected by David are the same colour, she will pay him 135p.
If exactly one of the two discs selected by David is blue, she will pay him 145p.
Otherwise David will pay Joanne 45p.

(i) When a game is played, X is the amount, in pence, **won by David**. Construct the probability distribution for X, in the form of a table. *(2 marks)*

(ii) Show that $\mathrm{E}(X) = 33$. *(2 marks)*

(c) Joanne modifies the game so that the amount per game, Y pence, that **she wins** may be modelled by

$$Y = 104 - 3X$$

(i) Determine how much Joanne would expect to win if the game is played 100 times. *(3 marks)*

(ii) Calculate the standard deviation of Y, giving your answer to the nearest 1p. *(4 marks)*

6 Bishen believes that the mean weight of boxes of black peppercorns is 45 grams. Abi, thinking that this is not the case, weighs, in grams, a random sample of 8 boxes of black peppercorns, with the following results.

44 44 43 46 42 40 43 46

(a) (i) Construct a 95% confidence interval for the mean weight of boxes of black peppercorns, stating any assumption that you make. *(6 marks)*

(ii) Comment on Bishen's belief. *(2 marks)*

(b) (i) Abi claims that the mean weight of boxes of black peppercorns is less than 45 grams. Test this claim at the 5% level of significance. *(6 marks)*

(ii) If Bishen's belief **is** true, state, with a reason, what type of error, if any, may have occurred when conclusions to the test in part (b)(i) were drawn. *(2 marks)*

END OF QUESTIONS

General Certificate of Education
Advanced Level Examination
January 2010

Mathematics MS2B

Unit Statistics 2B

Wednesday 20 January 2010 1.30 pm to 3.00 pm

For this paper you must have:
- an 8-page answer book
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black 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.
- Show all necessary working; 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 marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions.

1 Roger claims that, on average, his journey time from home to work each day is greater than 45 minutes.

The times, x minutes, of 30 randomly selected journeys result in $\bar{x} = 45.8$ and $s^2 = 4.8$.

Investigate Roger's claim at the 1% level of significance. *(5 marks)*

2 The error, in minutes, made by Paul in estimating the time that he takes to complete a college assignment may be modelled by the random variable T with probability density function

$$\mathrm{f}(t) = \begin{cases} \dfrac{1}{30} & -5 \leqslant t \leqslant 25 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find:

(i) $\mathrm{E}(T)$; *(1 mark)*

(ii) $\mathrm{Var}(T)$. *(1 mark)*

(b) Calculate the probability that Paul will make an error of magnitude at least 2 minutes when estimating the time that he takes to complete a given assignment. *(3 marks)*

3 Lorraine bought a new golf club. She then practised with this club by using it to hit golf balls on a golf range.

After several such practice sessions, she believed that there had been no change from 190 metres in the mean distance that she had achieved when using her old club.

To investigate this belief, she measured, at her next practice session, the distance, x metres, of each of a random sample of 10 shots with her new club. Her results gave

$$\sum x = 1840 \quad \text{and} \quad \sum (x - \bar{x})^2 = 1240$$

Investigate Lorraine's belief at the 2% level of significance, stating any assumption that you make. *(7 marks)*

4 Julie, a driving instructor, believes that the first-time performances of her students in their driving tests are associated with their ages.

Julie's records of her students' first-time performances in their driving tests are shown in the table.

Age	**Pass**	**Fail**
17–18	28	20
19–30	2	14
31–39	12	33
40–60	6	5

(a) Use a χ^2 test at the 1% level of significance to investigate Julie's belief. *(9 marks)*

(b) Interpret your result in part (a) as it relates to the 17–18 age group. *(1 mark)*

5 (a) In a remote African village, it is known that 70 per cent of the villagers have a particular blood disorder. A medical research student selects 25 of the villagers at random.

Using a binomial distribution, calculate the probability that more than 15 of these 25 villagers have this blood disorder. *(3 marks)*

(b) (i) In towns and cities in Asia, the number of people who have this blood disorder may be modelled by a Poisson distribution with a mean of 2.6 per 100 000 people.

A town in Asia with a population of 100 000 is selected. Determine the probability that at most 5 people have this blood disorder. *(1 mark)*

(ii) In towns and cities in South America, the number of people who have this blood disorder may be modelled by a Poisson distribution with a mean of 49 per **million** people.

A town in South America with a population of 100 000 is selected. Calculate the probability that exactly 10 people have this blood disorder. *(3 marks)*

(iii) The random variable T denotes the **total** number of people in the two selected towns who have this blood disorder.

Write down the distribution of T and hence determine $\mathrm{P}(T > 16)$. *(3 marks)*

Turn over ▶

6 (a) Ali has a bag of 10 balls, of which 5 are red and 5 are blue. He asks Ben to select 5 of these balls from the bag at random.

The probability distribution of X, the number of red balls that Ben selects, is given in **Table 1**.

Table 1

x	0	1	2	3	4	5
$\mathbf{P}(X = x)$	$\frac{1}{252}$	$\frac{25}{252}$	$\frac{100}{252}$	a	$\frac{25}{252}$	$\frac{1}{252}$

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

(ii) Hence write down the value of $\mathrm{E}(X)$. *(1 mark)*

(iii) Determine the standard deviation of X. *(5 marks)*

(b) Ali decides to play a game with Joanne using the same 10 balls. Joanne is asked to select 2 balls from the bag at random.

Ali agrees to pay Joanne 90p if the two balls that she selects are the same colour, but nothing if they are different colours. Joanne pays 50p to play the game.

The probability distribution of Y, the number of red balls that Joanne selects, is given in **Table 2**.

Table 2

y	0	1	2
$\mathbf{P}(Y = y)$	$\frac{2}{9}$	$\frac{5}{9}$	$\frac{2}{9}$

(i) Determine whether Joanne can expect to make a profit or a loss from playing the game once.

(ii) Hence calculate the expected size of this profit or loss after Joanne has played the game 100 times. *(3 marks)*

7 Jim, a mathematics teacher, knows that the marks, X, achieved by his students can be modelled by a normal distribution with unknown mean μ and unknown variance σ^2.

Jim selects 12 students at random and from their marks he calculates that $\bar{x} = 64.8$ and $s^2 = 93.0$.

(a) (i) An estimate for the standard error of the sample mean is d. Show that $d^2 = 7.75$. *(2 marks)*

(ii) Construct an 80% confidence interval for μ. *(3 marks)*

(b) (i) Write down a confidence interval for μ, based on Jim's sample of 12 students, which has a width of 10 marks. *(1 mark)*

(ii) Determine the percentage confidence level for the interval found in part (b)(i). *(4 marks)*

8 The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} \frac{1}{2}(x^2 + 1) & 0 \leqslant x \leqslant 1 \\ (x-2)^2 & 1 \leqslant x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Calculate $P(X \leqslant 1)$. *(3 marks)*

(c) Show that $E(X^2) = \dfrac{4}{5}$. *(5 marks)*

(d) (i) Given that $E(X) = \dfrac{19}{24}$ and that $\mathrm{Var}(X) = \dfrac{499}{k}$, find the numerical value of k. *(3 marks)*

(ii) Find $E(5X^2 + 24X - 3)$. *(2 marks)*

(iii) Find $\mathrm{Var}(12X - 5)$. *(2 marks)*

END OF QUESTIONS

Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
TOTAL	

General Certificate of Education
Advanced Level Examination
June 2010

Mathematics MS2B

Unit Statistics 2B

Friday 18 June 2010 1.30 pm to 3.00 pm

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

Answer **all** questions in the spaces provided.

1 Judith, the village postmistress, believes that, since moving the post office counter into the local pharmacy, the mean daily number of customers that she serves has increased from 79.

In order to investigate her belief, she counts the number of customers that she serves on 12 randomly selected days, with the following results.

88 81 84 89 90 77 72 80 82 81 75 85

Stating a necessary distributional assumption, test Judith's belief at the 5% level of significance. *(9 marks)*

QUESTION PART REFERENCE

2 It is claimed that a new drug is effective in the prevention of sickness in holiday-makers. A sample of 100 holiday-makers was surveyed, with the following results.

	Sickness	No sickness	Total
Drug taken	24	56	80
No drug taken	11	9	20
Total	35	65	100

Assuming that the 100 holiday-makers are a random sample, use a χ^2 test, at the 5% level of significance, to investigate the claim. *(8 marks)*

QUESTION PART REFERENCE

3 The continuous random variable X has a rectangular distribution defined by

$$f(x) = \begin{cases} k & -3k \leqslant x \leqslant k \\ 0 & \text{otherwise} \end{cases}$$

(a) (i) Sketch the graph of f. *(2 marks)*

(ii) Hence show that $k = \frac{1}{2}$. *(2 marks)*

(b) Find the **exact** numerical values for the mean and the standard deviation of X. *(3 marks)*

(c) (i) Find $P\left(X \geqslant -\frac{1}{4}\right)$. *(2 marks)*

(ii) Write down the value of $P\left(X \neq -\frac{1}{4}\right)$. *(1 mark)*

QUESTION PART REFERENCE

4 The error, X °C, made in measuring a patient's temperature at a local doctors' surgery may be modelled by a normal distribution with mean μ and standard deviation σ .

The errors, x °C, made in measuring the temperature of each of a random sample of 10 patients are summarised below.

$$\sum x = 0.35 \quad \text{and} \quad \sum (x - \bar{x})^2 = 0.12705$$

Construct a 99% confidence interval for μ , giving the limits to three decimal places.

(5 marks)

QUESTION PART REFERENCE

Do not write outside the box

5 The number of telephone calls received, during an 8-hour period, by an IT company that request an urgent visit by an engineer may be modelled by a Poisson distribution with a mean of 7.

(a) Determine the probability that, during a given 8-hour period, the number of telephone calls received that request an urgent visit by an engineer is:

(i) at most 5; *(1 mark)*

(ii) exactly 7; *(2 marks)*

(iii) at least 5 but fewer than 10. *(3 marks)*

(b) Write down the distribution for the number of telephone calls received each hour that request an urgent visit by an engineer. *(1 mark)*

(c) The IT company has 4 engineers available for urgent visits and it may be assumed that each of these engineers takes exactly 1 hour for each such visit.

At 10 am on a particular day, all 4 engineers are available for urgent visits.

(i) State the maximum possible number of telephone calls received between 10 am and 11 am that request an urgent visit and for which an engineer is immediately available. *(1 mark)*

(ii) Calculate the probability that at 11 am an engineer will **not** be immediately available to make an urgent visit. *(4 marks)*

(d) Give a reason why a Poisson distribution may not be a suitable model for the number of telephone calls per hour received by the IT company that request an urgent visit by an engineer. *(1 mark)*

QUESTION PART REFERENCE

6 (a) The number of strokes, R, taken by the members of Duffers Golf Club to complete the first hole may be modelled by the following discrete probability distribution.

r	$\leqslant 2$	3	4	5	6	7	8	$\geqslant 9$
$\mathbf{P}(\boldsymbol{R} = \boldsymbol{r})$	0	0.1	0.2	0.3	0.25	0.1	0.05	0

(i) Determine the probability that a member, selected at random, takes at least 5 strokes to complete the first hole. *(1 mark)*

(ii) Calculate $\mathrm{E}(R)$. *(2 marks)*

(iii) Show that $\mathrm{Var}(R) = 1.66$. *(4 marks)*

(b) The number of strokes, S, taken by the members of Duffers Golf Club to complete the second hole may be modelled by the following discrete probability distribution.

s	$\leqslant 2$	3	4	5	6	7	8	$\geqslant 9$
$\mathbf{P}(\boldsymbol{S} = \boldsymbol{s})$	0	0.15	0.4	0.3	0.1	0.03	0.02	0

Assuming that R and S are independent:

(i) show that $\mathrm{P}(R + S \leqslant 8) = 0.24$; *(5 marks)*

(ii) calculate the probability that, when 5 members are selected at random, at least 4 of them complete the first two holes in fewer than 9 strokes; *(3 marks)*

(iii) calculate $\mathrm{P}(R = 4 \mid R + S \leqslant 8)$. *(3 marks)*

QUESTION PART REFERENCE

7 The random variable X has probability density function defined by

$$f(x) = \begin{cases} \frac{1}{2} & 0 \leqslant x \leqslant 1 \\ \frac{1}{18}(x-4)^2 & 1 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) State values for the median and the lower quartile of X. *(2 marks)*

(b) Show that, for $1 \leqslant x \leqslant 4$, the cumulative distribution function, $F(x)$, of X is given by

$$F(x) = 1 + \frac{1}{54}(x-4)^3$$

(You may assume that $\int (x-4)^2 \, dx = \frac{1}{3}(x-4)^3 + c$.) *(4 marks)*

(c) Determine $P(2 \leqslant X \leqslant 3)$. *(2 marks)*

(d) (i) Show that q, the upper quartile of X, satisfies the equation $(q-4)^3 = -13.5$. *(3 marks)*

(ii) Hence evaluate q to three decimal places. *(1 mark)*

QUESTION PART REFERENCE

General Certificate of Education
Advanced Level Examination
January 2011

Mathematics MS2B

Unit Statistics 2B

Wednesday 26 January 2011 1.30 pm to 3.00 pm

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

1 A factory produces bottles of brown sauce and bottles of tomato sauce.

(a) The content, Y grams, of a bottle of brown sauce is normally distributed with mean μ_Y and variance 4.

A quality control inspection found that the mean content, $\bar{y}$ grams, of a random sample of 16 bottles of brown sauce was 450.

Construct a 95% confidence interval for μ_Y. *(3 marks)*

(b) The content, X grams, of a bottle of tomato sauce is normally distributed with mean μ_X and variance σ^2.

A quality control inspection found that the content, x grams, of a random sample of 9 bottles of tomato sauce was summarised by

$$\sum x = 4950 \quad \text{and} \quad \sum (x - \bar{x})^2 = 334$$

(i) Construct a 90% confidence interval for μ_X. *(5 marks)*

(ii) Holly, the supervisor at the factory, claims that the mean content of a bottle of tomato sauce is 545 grams.

Comment, with a justification, on Holly's claim. State the level of significance on which your conclusion is based. *(3 marks)*

2 It is claimed that the way in which students voted at a particular general election was independent of their gender.

In order to investigate this claim, 480 male and 540 female students who voted at this general election were surveyed. These students may be regarded as a random sample.

The **percentages** of males and females who voted for the different parties are recorded in the table.

	Conservative	**Labour**	**Liberal Democrat**	**Other parties**
Male	32.5	30	25	12.5
Female	40	25	20	15

(a) Complete the contingency table below. *(2 marks)*

(b) Hence determine, at the 1% level of significance, whether the way in which students voted at this general election was independent of their gender. *(9 marks)*

	Conservative	**Labour**	**Liberal Democrat**	**Other parties**	**Total**
Male					480
Female					540
Total					1020

---

3 Lucy is the captain of her school's cricket team.

The number of catches, X, taken by Lucy during any particular cricket match may be modelled by a Poisson distribution with mean 0.6.

The number of run-outs, Y, effected by Lucy during any particular cricket match may be modelled by a Poisson distribution with mean 0.15.

(a) Find:

(i) $\mathrm{P}(X \leqslant 1)$; *(1 mark)*

(ii) $\mathrm{P}(X \leqslant 1$ **and** $Y \geqslant 1)$. *(4 marks)*

(b) State the assumption that you made in answering part **(a)(ii)**. *(1 mark)*

(c) During a particular season, Lucy plays in 16 cricket matches.

(i) Calculate the probability that the number of catches taken by Lucy during this **season** is exactly 10. *(2 marks)*

(ii) Determine the probability that the **total** number of catches taken and run-outs effected by Lucy during this **season** is at least 15. *(3 marks)*

---

4 (a) A red biased tetrahedral die is rolled. The number, X, on the face on which it lands has the probability distribution given by

x	1	2	3	4
$P(X=x)$	0.2	0.1	0.4	0.3

(i) Calculate $E(X)$ and $Var(X)$. *(3 marks)*

(ii) The red die is now rolled **three** times. The random variable S is the **sum** of the three numbers obtained.

Find $E(S)$ and $Var(S)$. *(2 marks)*

(b) A blue biased tetrahedral die is rolled. The number, Y, on the face on which it lands has the probability distribution given by

$$P(Y=y) = \begin{cases} \dfrac{y}{20} & y = 1, 2 \text{ and } 3 \\ \dfrac{7}{10} & y = 4 \end{cases}$$

The random variable T is the value obtained when the number on the face on which it lands is **multiplied** by 3.

Calculate $E(T)$ and $Var(T)$. *(6 marks)*

(c) Calculate:

(i) $P(X > 1)$; *(1 mark)*

(ii) $P(X + T \leqslant 9$ **and** $X > 1)$; *(4 marks)*

(iii) $P(X + T \leqslant 9 \mid X > 1)$. *(2 marks)*

---

5 In 2001, the mean height of students at the end of their final year at Bright Hope Secondary School was 165 centimetres.

In 2010, David and James selected a random sample of 100 students who were at the end of their final year at this school. They recorded these students' heights, x centimetres, and found that $\bar{x} = 167.1$ and $s^2 = 101.2$.

To investigate the claim that the mean height had increased since 2001, David and James each correctly conducted a hypothesis test. They used the same null hypothesis and the same alternative hypothesis. However, David used a 5% level of significance whilst James used a 1% level of significance.

(a) (i) Write down the null and alternative hypotheses that both David and James used. *(1 mark)*

(ii) Determine the outcome of each of the two hypothesis tests, giving each conclusion in context. *(6 marks)*

(iii) State why both David and James made use of the Central Limit Theorem in their hypothesis tests. *(1 mark)*

(b) It was later found that, in 2010, the mean height of students at the end of their final year at Bright Hope Secondary School was actually 165 centimetres.

Giving a reason for your answer in each case, determine whether a Type I error or a Type II error or neither was made in the hypothesis test conducted by:

(i) David;

(ii) James. *(4 marks)*

6 The continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} \frac{3}{8}x^2 & 0 \leqslant x \leqslant \frac{1}{2} \\ \frac{3}{32} & \frac{1}{2} \leqslant x \leqslant 11 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Show that:

(i) $P\left(X \geqslant 8\frac{1}{3}\right) = \frac{1}{4}$;

(ii) $P(X \geqslant 3) = \frac{3}{4}$. *(3 marks)*

(c) Hence write down the **exact** value of:

(i) the interquartile range of X;

(ii) the median, m, of X. *(3 marks)*

(d) Find the **exact** value of $P(X < m \mid X \geqslant 3)$. *(3 marks)*

General Certificate of Education
Advanced Level Examination
June 2011

Mathematics MS2B

Unit Statistics 2B

Monday 20 June 2011 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.

1 The number of cars passing a speed camera on a main road between 9.30 am and 11.30 am may be modelled by a Poisson distribution with a mean rate of 2.6 per minute.

(a) (i) Write down the distribution of X, the number of cars passing the speed camera during a 5-minute interval between 9.30 am and 11.30 am. *(1 mark)*

(ii) Determine $\mathrm{P}(X = 20)$. *(2 marks)*

(iii) Determine $\mathrm{P}(6 \leqslant X \leqslant 18)$. *(3 marks)*

(b) Give **two** reasons why a Poisson distribution with mean 2.6 may not be a suitable model for the number of cars passing the speed camera during a 1-minute interval between 8.00 am and 9.30 am on weekdays. *(2 marks)*

(c) When n cars pass the speed camera, the number of cars, Y, that exceed 60 mph may be modelled by the distribution $\mathrm{B}(n, 0.2)$.

Given that $n = 20$, determine $\mathrm{P}(Y \geqslant 5)$. *(2 marks)*

(d) Stating a necessary assumption, calculate the probability that, during a given 5-minute interval between 9.30 am and 11.30 am, exactly 20 cars pass the speed camera of which at least 5 are exceeding 60 mph. *(3 marks)*

---

2 (a) The continuous random variable X has a rectangular distribution defined by the probability density function

$$\mathrm{f}(x) = \begin{cases} 0.01\pi & u \leqslant x \leqslant 11u \\ 0 & \text{otherwise} \end{cases}$$

where u is a constant.

(i) Show that $u = \dfrac{10}{\pi}$. *(2 marks)*

(ii) Using the formulae for the mean and the variance of a rectangular distribution, find, in terms of π, values for $\mathrm{E}(X)$ and $\mathrm{Var}(X)$. *(2 marks)*

(iii) Calculate **exact** values for the mean and the variance of the circumferences of circles having diameters of length $\left(X + \dfrac{10}{\pi}\right)$. *(4 marks)*

(b) A machine produces circular discs which have an area of $Y\,\mathrm{cm}^2$. The distribution of Y has mean μ and variance 25.

A random sample of 100 such discs is selected. The mean area of the discs in this sample is calculated to be $40.5\,\mathrm{cm}^2$.

Calculate a 95% confidence interval for μ. *(3 marks)*

3 Emily believed that the performances of 16-year-old students in their GCSEs are associated with the schools that they attend. To investigate her belief, Emily collected data on the GCSE results for 2010 from four schools in her area.

The table shows Emily's collected data, denoted by O_i, together with the corresponding expected frequencies, E_i, necessary for a χ^2 test.

	$\geqslant$5 GCSEs		**1 $\leqslant$ GCSEs < 5**		**No GCSEs**	
	O_i	E_i	O_i	E_i	O_i	E_i
Jolliffe College for the Arts	187	193.15	93	90.62	30	26.23
Volpe Science Academy	175	184.43	97	86.52	24	25.05
Radok Music School	183	183.81	78	86.23	34	24.96
Bailey Language School	265	248.61	112	116.63	22	33.76

Emily used these values to correctly conduct a χ^2 test at the 1% level of significance.

(a) State the null hypothesis that Emily used. *(1 mark)*

(b) Find the value of the test statistic, X^2, giving your answer to one decimal place. *(3 marks)*

(c) State, in context, the conclusion that Emily should reach based on the results of her χ^2 test. *(3 marks)*

(d) Make **one** comment on the GCSE performances of 16-year-old students attending Bailey Language School. *(1 mark)*

(e) Emily's friend, Joanna, used the same data to correctly conduct a χ^2 test using the 10% level of significance.

State, with justification, the conclusion that Joanna should reach. *(2 marks)*

Turn over ▶

4 A discrete random variable X has the probability distribution

$$P(X = x) = \begin{cases} \dfrac{3x}{40} & x = 1, 2, 3, 4 \\ \dfrac{x}{20} & x = 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Calculate $E(X)$. *(2 marks)*

(b) Show that:

(i) $E\left(\dfrac{1}{X}\right) = \dfrac{7}{20}$; *(2 marks)*

(ii) $\mathrm{Var}\left(\dfrac{1}{X}\right) = \dfrac{7}{160}$. *(4 marks)*

(c) The discrete random variable Y is such that $Y = \dfrac{40}{X}$.

Calculate:

(i) $P(Y < 20)$; *(3 marks)*

(ii) $P(X < 4 \mid Y < 20)$. *(3 marks)*

---

5 (a) The lifetime of a new 16-watt energy-saving light bulb may be modelled by a normal random variable with standard deviation 640 hours. A random sample of 25 bulbs, taken by the manufacturer from this distribution, has a mean lifetime of 19 700 hours.

Carry out a hypothesis test, at the 1% level of significance, to determine whether the mean lifetime has changed from 20 000 hours. *(6 marks)*

(b) The lifetime of a new 11-watt energy-saving light bulb may be modelled by a normal random variable with mean μ hours and standard deviation σ hours.

The manufacturer claims that the mean lifetime of these energy-saving bulbs is 10 000 hours. Christine, from a consumer organisation, believes that this is an overestimate.

To investigate her belief, she carries out a hypothesis test at the 5% level of significance based on the null hypothesis $H_0 : \mu = 10\,000$.

(i) State the alternative hypothesis that should be used by Christine in this test. *(1 mark)*

(ii) From the lifetimes of a random sample of 16 bulbs, Christine finds that $s = 500$ hours.

Determine the range of values for the sample mean which would lead Christine **not** to reject her null hypothesis. *(5 marks)*

(iii) It was later revealed that $\mu = 10\,000$.

State which type of error, if any, was made by Christine if she concluded that her null hypothesis should **not** be rejected. *(1 mark)*

---

6 The continuous random variable X has the probability density function defined by

$$f(x) = \begin{cases} \frac{3}{8}(x^2 + 1) & 0 \leqslant x \leqslant 1 \\ \frac{1}{4}(5 - 2x) & 1 \leqslant x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) The cumulative distribution function of X is denoted by $F(x)$.

Show that, for $0 \leqslant x \leqslant 1$,

$$F(x) = \frac{1}{8}x(x^2 + 3)$$ *(3 marks)*

(b) Hence, or otherwise, verify that the median value of X is 1. *(2 marks)*

(c) Show that the upper quartile, q, satisfies the equation $q^2 - 5q + 5 = 0$ and hence that $q = \frac{1}{2}(5 - \sqrt{5})$. *(5 marks)*

(d) Calculate the **exact** value of $P(q < X < 1.5)$. *(4 marks)*

END OF QUESTIONS

General Certificate of Education
Advanced Level Examination
January 2012

Mathematics MS2B

Unit Statistics 2B

Wednesday 25 January 2012 1.30 pm to 3.00 pm

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer the questions in the spaces provided. Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

1 Josephine accurately measures the widths of A4 sheets of paper and then rounds the widths to the nearest 0.1 cm. The rounding error, X centimetres, follows a rectangular distribution.

A randomly selected A4 sheet of paper is measured to be 21.1 cm in width.

(a) Write down the limits between which the true width of this A4 sheet of paper lies. *(1 mark)*

(b) Write down the value of $\mathrm{E}(X)$ and determine the **exact** value of the standard deviation of X. *(3 marks)*

(c) Calculate $\mathrm{P}(-0.01 \leqslant X \leqslant 0.03)$. *(1 mark)*

---

2 (a) A particular bowling club has a large number of members. Their ages may be modelled by a normal random variable, X, with standard deviation 7.5 years.

On 30 June 2010, Ted, the club secretary, concerned about the ageing membership, selected a random sample of 16 members and calculated their mean age to be 65.0 years.

(i) Carry out a hypothesis test, at the 5% level of significance, to determine whether the mean age of the club's members has changed from its value of 61.4 years on 30 June 2000. *(6 marks)*

(ii) Comment on the likely number of members who were under the age of 25 on 30 June 2010, giving a numerical reason for your answer. *(1 mark)*

(b) During 2011, in an attempt to encourage greater participation in the sport, the club ran a recruitment drive.

After the recruitment drive, the ages of members of the bowling club may be modelled by a normal random variable, Y years, with mean μ and standard deviation σ. The ages, y years, of a random sample of 12 such members are summarised below.

$$\sum y = 702 \quad \text{and} \quad \sum (y - \bar{y})^2 = 88.25$$

(i) Construct a 90% confidence interval for μ, giving the limits to one decimal place. *(5 marks)*

(ii) Use your confidence interval to state, with a reason, whether the recruitment drive lowered the average age of the club's members. *(1 mark)*

3 (a) **Table 1** contains the observed frequencies, a, b, c and d, relating to the two attributes, X and Y, required to perform a χ^2 test.

Table 1

	Y	**Not *Y***	**Total**
X	a	b	m
Not *X*	c	d	n
Total	p	q	N

(i) Write down, in terms of m, n, p, q and N, expressions for the 4 expected frequencies corresponding to a, b, c and d. *(2 marks)*

(ii) Hence prove that the sum of the expected frequencies is N. *(3 marks)*

(b) Andy, a tennis player, wishes to investigate the possible effect of wind conditions on the results of his matches. The results of his matches for the 2011 season are represented in **Table 2**.

Table 2

	Windy	**Not windy**	**Total**
Won	15	18	33
Lost	12	5	17
Total	27	23	50

Conduct a χ^2 test, at the 10% level of significance, to investigate whether there is an association between Andy's results and wind conditions. *(8 marks)*

4 (a) A discrete random variable X has a probability function defined by

$$\mathrm{P}(X = x) = \frac{\mathrm{e}^{-\lambda}\,\lambda^x}{x!} \quad \text{for } x = 0, 1, 2, 3, 4, \ldots\ldots$$

(i) State the name of the distribution of X. *(1 mark)*

(ii) Write down, in terms of λ, expressions for $\mathrm{E}(3X - 1)$ and $\mathrm{Var}(3X - 1)$. *(2 marks)*

(iii) Write down an expression for $\mathrm{P}(X = x + 1)$, and hence show that

$$\mathrm{P}(X = x + 1) = \frac{\lambda}{x + 1}\mathrm{P}(X = x)$$ *(3 marks)*

(b) The number of cars and the number of coaches passing a certain road junction may be modelled by independent Poisson distributions.

(i) On a winter morning, an average of 500 cars per hour and an average of 10 coaches per hour pass this junction.

Determine the probability that a total of at least 10 such vehicles pass this junction during a particular 1-minute interval on a winter morning. *(3 marks)*

(ii) On a summer morning, an average of 836 cars per hour and an average of 22 coaches per hour pass this junction.

Calculate the probability that a total of at most 3 such vehicles pass this junction during a particular 1-minute interval on a summer morning. Give your answer to two significant figures. *(3 marks)*

5 (a) Joshua plays a game in which he repeatedly tosses an unbiased coin. His game concludes when he obtains either a head or 5 tails in succession.

The random variable N denotes the number of tosses of his coin required to conclude a game.

By completing **Table 3** below, calculate $\mathrm{E}(N)$. *(4 marks)*

(b) Joshua's sister, Ruth, plays a separate game in which she repeatedly tosses a coin that is **biased** in such a way that the probability of a head in a single toss of her coin is $\frac{1}{4}$. Her game also concludes when she obtains either a head or 5 tails in succession.

The random variable M denotes the number of tosses of her coin required to conclude her game.

Complete **Table 4** below. *(3 marks)*

(c) (i) Joshua and Ruth play their games simultaneously. Calculate the probability that Joshua and Ruth will conclude their games in an equal number of tosses of their coins. *(5 marks)*

(ii) Joshua and Ruth play their games simultaneously on 3 occasions. Calculate the probability that, on at least 2 of these occasions, their games will be concluded in an equal number of tosses of their coins. Give your answer to three decimal places. *(4 marks)*

Table 3

n	1	2	3	4	5
$\mathrm{P}(N = n)$			$\frac{1}{8}$		$\frac{1}{16}$

Table 4

m	1	2	3	4	5
$\mathrm{P}(M = m)$	$\frac{1}{4}$	$\frac{3}{16}$			

6 The random variable X has probability density function defined by

$$\mathrm{f}(x) = \begin{cases} \frac{1}{40}(x+7) & 1 \leqslant x \leqslant 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(2 marks)*

(b) Find the **exact** value of $\mathrm{E}(X)$. *(3 marks)*

(c) Prove that the distribution function F, for $1 \leqslant x \leqslant 5$, is defined by

$$\mathrm{F}(x) = \frac{1}{80}(x+15)(x-1)$$ *(4 marks)*

(d) Hence, or otherwise:

(i) find $\mathrm{P}(2.5 \leqslant X \leqslant 4.5)$; *(2 marks)*

(ii) show that the median, m, of X satisfies the equation $m^2 + 14m - 55 = 0$. *(3 marks)*

(e) Calculate the value of the median of X, giving your answer to three decimal places. *(2 marks)*

General Certificate of Education
Advanced Level Examination
June 2012

Mathematics MS2B

Unit Statistics 2B

Thursday 21 June 2012 1.30 pm to 3.00 pm

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

1 At the start of the 2012 season, the ages of the members of the Warwickshire Acorns Cricket Club could be modelled by a normal random variable, X years, with mean μ and standard deviation σ.

The ages, x years, of a random sample of 15 such members are summarised below.

$$\sum x = 546 \quad \text{and} \quad \sum (x - \bar{x})^2 = 1407.6$$

(a) Construct a 98% confidence interval for μ, giving the limits to one decimal place. *(6 marks)*

(b) At the start of the 2005 season, the mean age of the members was 40.0 years.

Use your confidence interval constructed in part **(a)** to indicate, with a reason, whether the mean age had changed. *(2 marks)*

---

2 The times taken to complete a round of golf at Slowpace Golf Club may be modelled by a random variable with mean μ hours and standard deviation 1.1 hours.

Julian claims that, on average, the time taken to complete a round of golf at Slowpace Golf Club is greater than 4 hours.

The times of 40 randomly selected completed rounds of golf at Slowpace Golf Club result in a mean of 4.2 hours.

(a) Investigate Julian's claim at the 5% level of significance. *(6 marks)*

(b) If the actual mean time taken to complete a round of golf at Slowpace Golf Club is 4.5 hours, determine whether a Type I error, a Type II error or neither was made in the test conducted in part **(a)**. Give a reason for your answer. *(2 marks)*

---

3 The continuous random variable X has a cumulative distribution function defined by

$$\mathrm{F}(x) = \begin{cases} 0 & x < -5 \\ \dfrac{x+5}{20} & -5 \leqslant x \leqslant 15 \\ 1 & x > 15 \end{cases}$$

(a) Show that, for $-5 \leqslant x \leqslant 15$, the probability density function, $\mathrm{f}(x)$, of X is given by $\mathrm{f}(x) = \frac{1}{20}$. *(1 mark)*

(b) Find:

(i) $P(X \geqslant 7)$; *(1 mark)*

(ii) $P(X \neq 7)$; *(1 mark)*

(iii) $E(X)$; *(1 mark)*

(iv) $E(3X^2)$. *(3 marks)*

---

4 A house has a total of five bedrooms, at least one of which is always rented.

The probability distribution for R, the number of bedrooms that are rented at any given time, is given by

$$P(R = r) = \begin{cases} 0.5 & r = 1 \\ 0.4(0.6)^{r-1} & r = 2, 3, 4 \\ 0.0296 & r = 5 \end{cases}$$

(a) Complete the table below. *(2 marks)*

(b) Find the probability that fewer than 3 bedrooms are **not** rented at any given time. *(3 marks)*

(c) (i) Find the value of $E(R)$. *(2 marks)*

(ii) Show that $E(R^2) = 4.8784$ and hence find the value of $\text{Var}(R)$. *(3 marks)*

(d) Bedrooms are rented on a monthly basis.

The monthly income, £M, from renting bedrooms in the house may be modelled by

$$M = 1250R - 282$$

Find the mean and the standard deviation of M. *(3 marks)*

r	1	2	3	4	5
$P(R = r)$	0.5				0.0296

5 (a) The number of **minor** accidents occurring each year at RapidNut engineering company may be modelled by the random variable X having a Poisson distribution with mean 8.5 .

Determine the probability that, in any particular year, there are:

(i) at least 9 minor accidents; *(2 marks)*

(ii) more than 5 but fewer than 10 minor accidents. *(3 marks)*

(b) The number of **major** accidents occurring each year at RapidNut engineering company may be modelled by the random variable Y having a Poisson distribution with mean 1.5 .

Calculate the probability that, in any particular year, there are fewer than 2 major accidents. *(2 marks)*

(c) The **total** number of minor and major accidents occurring each year at RapidNut engineering company may be modelled by the random variable T having the probability distribution

$$P(T = t) = \begin{cases} \dfrac{e^{-\lambda}\lambda^t}{t!} & t = 0, 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Assuming that the number of minor accidents is independent of the number of major accidents:

(i) state the value of λ ; *(1 mark)*

(ii) determine $P(T > 16)$; *(2 marks)*

(iii) calculate the probability that there will be a total of more than 16 accidents in each of at least two out of three years, giving your answer to four decimal places. *(3 marks)*

6 Fiona, a lecturer in a school of engineering, believes that there is an association between the class of degree obtained by her students and the grades that they had achieved in A-level Mathematics.

In order to investigate her belief, she collected the relevant data on the performances of a random sample of 200 recent graduates who had achieved grades A or B in A-level Mathematics. These data are tabulated below.

		Class of degree				
		1	**2(i)**	**2(ii)**	**3**	**Total**
A-level grade	**A**	20	36	22	2	80
	B	9	55	48	8	120
	Total	29	91	70	10	200

(a) Conduct a χ^2 test, at the 1% level of significance, to determine whether Fiona's belief is justified. *(9 marks)*

(b) Make **two** comments on the degree performance of those students in this sample who achieved a grade B in A-level Mathematics. *(2 marks)*

---

7 A continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} \frac{1}{6}(4 - x) & 1 \leqslant x \leqslant 3 \\ \frac{1}{6} & 3 \leqslant x \leqslant 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Draw the graph of f on the grid on page 6. *(2 marks)*

(b) Prove that the mean of X is $2\frac{5}{9}$. *(4 marks)*

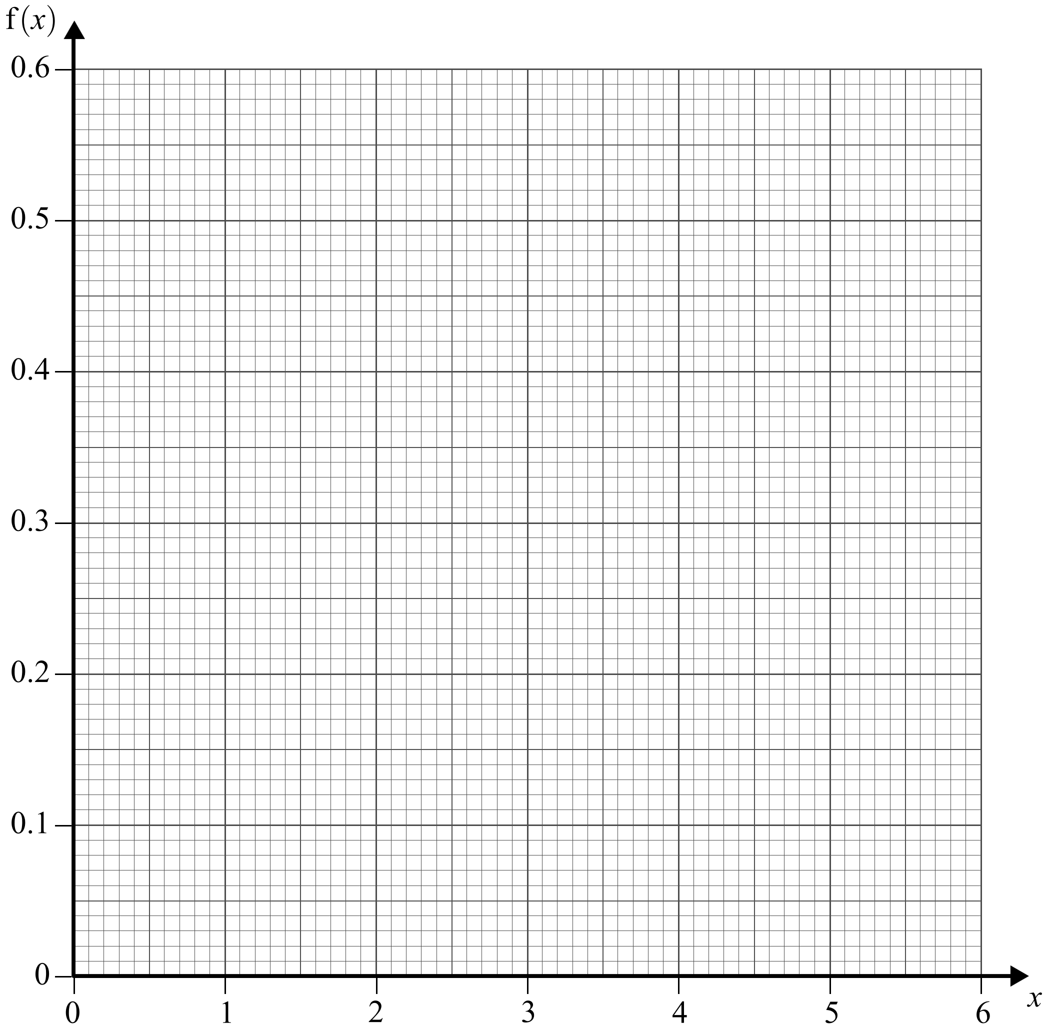
(c) Calculate the **exact** value of:

(i) $P(X > 2.5)$; *(2 marks)*

(ii) $P(1.5 < X < 4.5)$; *(3 marks)*

(iii) $P(X > 2.5$ **and** $1.5 < X < 4.5)$; *(2 marks)*

(iv) $P(X > 2.5 \mid 1.5 < X < 4.5)$. *(2 marks)*


f(x)
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
x

General Certificate of Education
Advanced Level Examination
January 2013

Mathematics MS2B

Unit Statistics 2B

Monday 28 January 2013 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

1 Dimitra is an athlete who competes in 400 m races. The times, in seconds, for her first six races of the 2012 season were

54.86 53.09 53.75 52.88 51.97 51.81

(a) Assuming that these data form a random sample from a normal distribution, construct a 95% confidence interval for the mean time of Dimitra's races in the 2012 season, giving the limits to two decimal places. *(5 marks)*

(b) For the 2011 season, Dimitra's mean time for her races was 53.41 seconds. After her first six races of the 2012 season, her coach claimed that the data showed that she would be more successful in races during the 2012 season than during the 2011 season. Make **two** comments about the coach's claim. *(2 marks)*

---

2 A large estate agency would like all the properties that it handles to be sold within three months. A manager wants to know whether the type of property affects the time taken to sell it. The data for a random sample of properties sold are tabulated below.

	Type of property				
	Flat	**Terraced**	**Semi-detached**	**Detached**	**Total**
Sold within three months	4	34	28	18	84
Sold in more than three months	9	18	8	6	41
Total	13	52	36	24	125

(a) Conduct a χ^2-test, at the 10% level of significance, to determine whether there is an association between the type of property and the time taken to sell it. Explain why it is necessary to combine two columns before carrying out this test. *(10 marks)*

(b) The manager plans to spend extra money on advertising for one type of property in an attempt to increase the number sold within three months. Explain why the manager might choose:

(i) terraced properties;

(ii) flats. *(2 marks)*

3 A large office block is busy during the five weekdays, Monday to Friday, and less busy during the two weekend days, Saturday and Sunday. The block is illuminated by fluorescent light tubes which frequently fail and must be replaced with new tubes by John, the caretaker.

The number of fluorescent tubes that fail on a particular weekday can be modelled by a Poisson distribution with mean 1.5 .

The number of fluorescent tubes that fail on a particular weekend day can be modelled by a Poisson distribution with mean 0.5 .

(a) Find the probability that:

(i) on one particular Monday, exactly 3 fluorescent light tubes fail; *(2 marks)*

(ii) during the two days of a weekend, more than 1 fluorescent light tube fails; *(2 marks)*

(iii) during a complete seven-day week, fewer than 10 fluorescent light tubes fail. *(4 marks)*

(b) John keeps a supply of new fluorescent light tubes. More new tubes are delivered every Monday morning to replace those that he has used during the previous week. John wants the probability that he runs out of new tubes before the next Monday morning to be less than 1 per cent. Find the minimum number of new tubes that he should have available on a Monday morning. *(2 marks)*

(c) Give a reason why a Poisson distribution with mean 0.375 is unlikely to provide a satisfactory model for the number of fluorescent light tubes that fail between 1 am and 7 am on a weekday. *(1 mark)*

---

4 A continuous random variable X has probability density function defined by

$$\mathrm{f}(x) = \begin{cases} kx^2 & 0 \leqslant x \leqslant 3 \\ 9k & 3 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f. *(3 marks)*

(b) Show that the value of k is $\frac{1}{18}$. *(4 marks)*

(c) (i) Write down the median value of X.

(ii) Calculate the value of the lower quartile of X. *(4 marks)*

5 Aiden takes his car to a garage for its MOT test. The probability that his car will need to have X tyres replaced is shown in the table.

x	0	1	2	3	4
$\mathbf{P}(X = x)$	0.1	0.35	0.25	0.2	0.1

(a) Show that the mean of X is 1.85 and calculate the variance of X. *(4 marks)*

(b) The charge for the MOT test is £c and the cost of **each** new tyre is £n. The total amount that Aiden must pay the garage is £T.

(i) Express T in terms of c, n and X. *(1 mark)*

(ii) Hence, using your results from part **(a)**, find expressions for $\mathrm{E}(T)$ and $\mathrm{Var}(T)$. *(4 marks)*

---

6 The time, in weeks, that a patient must wait to be given an appointment in Holmsoon Hospital may be modelled by a random variable T having the cumulative distribution function

$$\mathrm{F}(t) = \begin{cases} 0 & t < 0 \\ \dfrac{t^3}{216} & 0 \leqslant t \leqslant 6 \\ 1 & t > 6 \end{cases}$$

(a) Find, to the nearest day, the time within which 90 per cent of patients will have been given an appointment. *(3 marks)*

(b) Find the probability density function of T for all values of t. *(3 marks)*

(c) Calculate the mean and the variance of T. *(6 marks)*

(d) Calculate the probability that the time that a patient must wait to be given an appointment is more than one standard deviation above the mean. *(4 marks)*

7 A factory produces 3-litre bottles of mineral water. The volume of water in a bottle has previously had a mean value of 3020 millilitres. Following a stoppage for maintenance, the volume of water, x millilitres, in each of a random sample of 100 bottles is measured and the following data obtained, where $y = x - 3000$.

$$\sum y = 1847.0 \qquad\qquad \sum (y - \bar{y})^2 = 6336.00$$

(a) Carry out a hypothesis test, at the 5% significance level, to investigate whether the mean volume of water in a bottle has changed. *(8 marks)*

(b) Subsequent measurements establish that the mean volume of water in a bottle produced by the factory after the stoppage is 3020 millilitres. State whether a Type I error, a Type II error or no error was made when carrying out the test in part **(a)**. *(1 mark)*

General Certificate of Education
Advanced Level Examination
June 2013

Mathematics MS2B

Unit Statistics 2B

Thursday 13 June 2013 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

1 Gemma, a biologist, studies guillemots, which are a species of seabird. She has found that the weight of an adult guillemot may be modelled by a normal distribution with mean μ grams. During 2012, she measured the weight, x grams, of each of a random sample of 9 adult guillemots and obtained the following results.

$$\sum x = 8532 \qquad \text{and} \qquad \sum (x - \bar{x})^2 = 38\,538$$

(a) Construct a 98% confidence interval for μ based on these data. *(5 marks)*

(b) The corresponding confidence interval for μ obtained by Gemma based on a random sample of 9 adult guillemots measured during 2011 was (927, 1063), correct to the nearest gram.

(i) Find the mean weight of guillemots in this sample. *(1 mark)*

(ii) Studies of some other species of seabird have suggested that their mean weights were less in 2012 than in 2011. Comment on whether Gemma's two confidence intervals provide evidence that the mean weight of guillemots was less in 2012 than in 2011. *(2 marks)*

---

2 A town council wanted residents to apply for grants that were available for home insulation. In a trial, a random sample of 200 residents was encouraged, either in a letter or by a phone call, to apply for the grants. The outcomes are shown in the table.

	Applied for grant	**Did not apply for grant**	**Total**
Letter	30	130	160
Phone call	14	26	40
Total	44	156	200

(a) The council believed that a phone call was more effective than a letter in encouraging people to apply for a grant. Use a χ^2-test to investigate this belief at the 5% significance level. *(8 marks)*

(b) After the trial, all the residents in the town were encouraged, either in a letter or by a phone call, to apply for the grants. It was found that there was no association between the method of encouragement and the outcome. State, with a reason, whether a Type I error, a Type II error or neither occurred in carrying out the test in part **(a)**. *(2 marks)*

3 Mehreen lives a 2-minute walk away from a tram stop. Trams run every 10 minutes into the city centre, taking 20 minutes to get there. Every morning, Mehreen leaves her house, walks to the tram stop and catches the first tram that arrives. When she arrives at the city centre, she then has a 5-minute walk to her office.

The total time, T minutes, for Mehreen's journey from house to office may be modelled by a rectangular distribution with probability density function

$$\mathrm{f}(t) = \begin{cases} 0.1 & a \leqslant t \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

(a) (i) Explain why $a = 27$.

(ii) State the value of b. *(3 marks)*

(b) Find the values of $\mathrm{E}(T)$ and $\mathrm{Var}(T)$. *(2 marks)*

(c) Find the probability that the time for Mehreen's journey is within 5 minutes of half an hour. *(2 marks)*

---

4 Gamma-ray bursts (GRBs) are pulses of gamma rays lasting a few seconds, which are produced by explosions in distant galaxies. They are detected by satellites in orbit around Earth. One particular satellite detects GRBs at a constant average rate of 3.5 per week (7 days).

You may assume that the detection of GRBs by this satellite may be modelled by a Poisson distribution.

(a) Find the probability that the satellite detects:

(i) exactly 4 GRBs during one particular week; *(2 marks)*

(ii) at least 2 GRBs on one particular **day**; *(3 marks)*

(iii) more than 10 GRBs but fewer than 20 GRBs during the 28 days of February 2013. *(3 marks)*

(b) Give **one** reason, apart from the constant average rate, why it is likely that the detection of GRBs by this satellite may be modelled by a Poisson distribution. *(1 mark)*

5 In a computer game, players try to collect five treasures. The number of treasures that Isaac collects in one play of the game is represented by the discrete random variable X.

The probability distribution of X is defined by

$$P(X = x) = \begin{cases} \dfrac{1}{x+2} & x = 1, 2, 3, 4 \\ k & x = 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) (i) Show that $k = \frac{1}{20}$. *(2 marks)*

(ii) Calculate the value of $E(X)$. *(2 marks)*

(iii) Show that $\text{Var}(X) = 1.5275$. *(3 marks)*

(iv) Find the probability that Isaac collects more than 2 treasures. *(2 marks)*

(b) The number of points that Isaac scores for collecting treasures is Y where

$$Y = 100X - 50$$

Calculate the mean and the standard deviation of Y. *(4 marks)*

---

6 A supermarket buys pears from a local supplier. The supermarket requires the mean weight of the pears to be at least 175 grams. William, the fresh-produce manager at the supermarket, suspects that the latest batch of pears delivered does not meet this requirement.

(a) William weighs a random sample of 6 pears, obtaining the following weights, in grams.

160.6 155.4 181.3 176.2 162.3 172.8

Previous batches of pears have had weights that could be modelled by a normal distribution with standard deviation 9.4 grams. Assuming that this still applies, show that a hypothesis test at the 5% level of significance supports William's suspicion. *(7 marks)*

(b) William then weighs a random sample of 20 pears. The mean of this sample is 169.4 grams and $s = 11.2$ grams, where s^2 is an unbiased estimate of the population variance.

Assuming that the population from which this sample is taken has a normal distribution but with unknown standard deviation, test William's suspicion at the **1%** level of significance. *(5 marks)*

(c) Give a reason why the probability of a Type I error occurring was smaller when conducting the test in part **(b)** than when conducting the test in part **(a)**. *(1 mark)*

7 A continuous random variable X has the probability density function defined by

$$f(x) = \begin{cases} x^2 & 0 \leqslant x \leqslant 1 \\ \frac{1}{3}(5 - 2x) & 1 \leqslant x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

(a) Sketch the graph of f on the axes below. *(3 marks)*

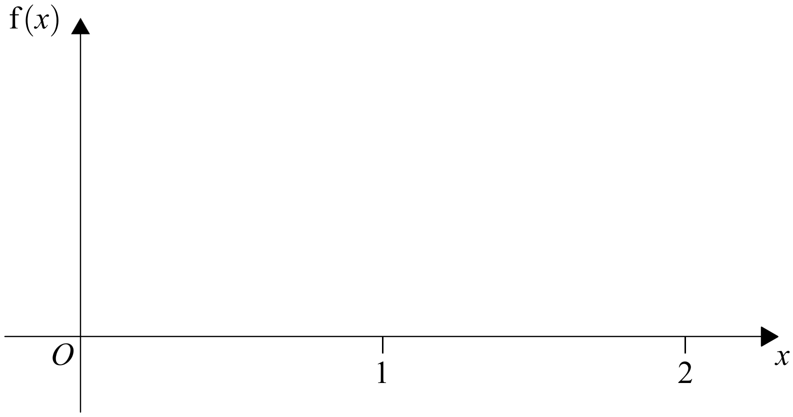
(b) (i) Find the cumulative distribution function, F, for $0 \leqslant x \leqslant 1$. *(2 marks)*

(ii) Hence, or otherwise, find the value of the lower quartile of X. *(2 marks)*

(c) (i) Show that the cumulative distribution function for $1 \leqslant x \leqslant 2$ is defined by

$$F(x) = \frac{1}{3}(5x - x^2 - 3)$$ *(4 marks)*

(ii) Hence, or otherwise, find the value of the upper quartile of X. *(4 marks)*



Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
TOTAL	

AQA

General Certificate of Education
Advanced Level Examination
June 2014

Mathematics MS2B

Unit Statistics 2B

Monday 16 June 2014 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

Answer **all** questions.

Answer each question in the space provided for that question.

1 Vanya collected five samples of air and measured the carbon dioxide content of each sample, in parts per million by volume (ppmv). The results were as follows.

$$387 \quad 375 \quad 382 \quad 379 \quad 381$$

(a) Assuming that these data form a random sample from a normal distribution with mean μ ppmv, construct a 90% confidence interval for μ.

[6 marks]

(b) Vanya repeated her sampling procedure on each of 30 days and, for each day's results, a 90% confidence interval for μ was constructed.

On how many of these 30 days would you expect μ to lie outside that day's confidence interval?

[1 mark]

QUESTION PART REFERENCE

Answer space for question 1

2 A large multinational company recruits employees from all four countries in the UK. For a sample of 250 recruits, the **percentages** of males and females from each of the countries are shown in **Table 1**.

Table 1

	England	Scotland	Wales	Northern Ireland
Male	22.8	17.6	10.8	6.8
Female	15.6	17.2	7.6	1.6

(a) Add the frequencies to the contingency table, **Table 2**, below.

[2 marks]

(b) Carry out a χ^2-test at the 10% significance level to investigate whether there is an association between country and gender of recruits.

[8 marks]

(c) By comparing observed and expected values, make **one** comment about the distribution of **female** recruits.

[1 mark]

QUESTION PART REFERENCE

Answer space for question 2

Table 2

	England	Scotland	Wales	Northern Ireland	Total
Male					145
Female					105
Total					250

3 A box contains a large number of pea pods. The number of peas in a pod may be modelled by the random variable X. The probability distribution of X is tabulated below.

x	2 or fewer	3	4	5	6	7	8 or more
$\mathbf{P}(X=x)$	0	0.1	0.2	a	0.3	b	0

(a) Two pods are picked randomly from the box. Find the probability that the number of peas in **each** pod is at most 4.

[2 marks]

(b) It is given that $\mathrm{E}(X) = 5.1$.

(i) Determine the values of a and b.

[4 marks]

(ii) Hence show that $\mathrm{Var}(X) = 1.29$.

[2 marks]

(iii) Some children play a game with the pods, randomly picking a pod and scoring points depending on the number of peas in the pod. For each pod picked, the number of points scored, N, is found by doubling the number of peas in the pod and then subtracting 5.

Find the mean and the standard deviation of N.

[3 marks]

QUESTION PART REFERENCE	Answer space for question 3

4 A continuous random variable X has a probability density function defined by

$$f(x) = \begin{cases} \dfrac{1}{k} & a \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

where $b > a > 0$.

(a) (i) Prove that $k = b - a$.

[2 marks]

(ii) Write down the value of $\mathrm{E}(X)$.

[1 mark]

(iii) Show, by integration, that $\mathrm{E}(X^2) = \frac{1}{3}(b^2 + ab + a^2)$.

[3 marks]

(iv) Hence derive a simplified formula for $\mathrm{Var}(X)$.

[2 marks]

(b) Given that $a = 4$ and $\mathrm{Var}(X) = 3$, find the numerical value of $\mathrm{E}(X)$.

[3 marks]

QUESTION PART REFERENCE

Answer space for question 4

5 Peter, a geologist, is studying pebbles on a beach. He uses a frame, called a quadrat, to enclose an area of the beach. He then counts the number of quartz pebbles, X, within the quadrat. He repeats this procedure 40 times, obtaining the following summarised data.

$$\sum x = 128 \quad \text{and} \quad \sum (x - \bar{x})^2 = 126.4$$

Peter believes that the distribution of X can be modelled by a Poisson distribution with $\lambda = 3.2$.

(a) Use the summarised data to support Peter's belief.

[3 marks]

(b) Using Peter's model, calculate the probability that:

(i) a single placing of the quadrat contains more than 5 quartz pebbles;

[2 marks]

(ii) a single placing of the quadrat contains at least 3 quartz pebbles but fewer than 8 quartz pebbles;

[3 marks]

(iii) when the quadrat is placed **twice**, at least one placing contains no quartz pebbles.

[3 marks]

(c) Peter also models the number of **flint** pebbles enclosed by the quadrat by a Poisson distribution with mean 5. He assumes that the number of flint pebbles enclosed by the quadrat is independent of the number of quartz pebbles enclosed by the quadrat.

Using Peter's models, calculate the probability that a single placing of the quadrat contains a **total** of either 9 or 10 pebbles which are quartz or flint.

[3 marks]

QUESTION PART REFERENCE	**Answer space for question 5**

6 South Riding Alarms (SRA) maintains household burglar-alarm systems. The company aims to carry out an annual service of a system in a mean time of 20 minutes. Technicians who carry out an annual service must record the times at which they start and finish the service.

(a) Gary is employed as a technician by SRA and his manager, Rajul, calculates the times taken for 8 annual services carried out by Gary. The results, in minutes, are as follows.

$$24 \quad 25 \quad 29 \quad 16 \quad 18 \quad 27 \quad 19 \quad 23$$

Assume that these times may be regarded as a random sample from a normal distribution.

Carry out a hypothesis test, at the 10% significance level, to examine whether the mean time for an annual service carried out by Gary is 20 minutes.

[8 marks]

(b) Rajul suspects that Gary may be taking longer than 20 minutes on average to carry out an annual service. Rajul therefore calculates the times taken for 100 annual services carried out by Gary.

Assume that these times may also be regarded as a random sample from a normal distribution but with a standard deviation of 4.6 minutes.

Find the highest value of the sample mean which would **not** support Rajul's suspicion at the 5% significance level. Give your answer to two decimal places.

[4 marks]

QUESTION PART REFERENCE	**Answer space for question 6**

7 A continuous random variable X has the probability density function defined by

$$f(x) = \begin{cases} \frac{4}{5}x & 0 \leqslant x \leqslant 1 \\ \frac{1}{20}(x-3)(3x-11) & 1 \leqslant x \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find $P(X < 1)$.

[2 marks]

(b) (i) Show that, for $1 \leqslant x \leqslant 3$, the cumulative distribution function, $F(x)$, is given by

$$F(x) = \frac{1}{20}(x^3 - 10x^2 + 33x - 16)$$

[4 marks]

(ii) Hence verify that the median value of X lies between 1.13 and 1.14.

[3 marks]

QUESTION PART REFERENCE

Answer space for question 7

Centre Number						Candidate Number				
Surname										
Other Names										
Candidate Signature										

For Examiner's Use	
Examiner's Initials	
Question	Mark
1	
2	
3	
4	
5	
6	
7	
TOTAL	

General Certificate of Education
Advanced Level Examination
June 2015

Mathematics MS2B

Unit Statistics 2B

Friday 12 June 2015 9.00 am to 10.30 am

For this paper you must have:
- the blue AQA booklet of formulae and statistical tables.

You may use a graphics calculator.

Time allowed
- 1 hour 30 minutes

Instructions
- Use black ink or black ball-point pen. Pencil should only be used for drawing.
- Fill in the boxes at the top of this page.
- Answer **all** questions.
- Write the question part reference (eg (a), (b)(i) etc) in the left-hand margin.
- You must answer each question in the space provided for that question. If you require extra space, use an AQA supplementary answer book; do **not** use the space provided for a different question.
- Do not write outside the box around each page.
- Show all necessary working; otherwise marks for method may be lost.
- Do all rough work in this book. Cross through any work that you do not want to be marked.
- The **final** answer to questions requiring the use of tables or calculators should normally be given to three significant figures.

Information
- The marks for questions are shown in brackets.
- The maximum mark for this paper is 75.

Advice
- Unless stated otherwise, you may quote formulae, without proof, from the booklet.
- You do not necessarily need to use all the space provided.

Answer **all** questions.

Answer each question in the space provided for that question.

1 In a survey of the tideline along a beach, plastic bottles were found at a constant average rate of 280 per kilometre, and drinks cans were found at a constant average rate of 220 per kilometre. It may be assumed that these objects were distributed randomly and independently.

Calculate the probability that:

(a) a $10\,\mathrm{m}$ length of tideline along this beach contains no more than 5 plastic bottles;

[2 marks]

(b) a $20\,\mathrm{m}$ length of tideline along this beach contains exactly 2 drinks cans;

[3 marks]

(c) a $30\,\mathrm{m}$ length of tideline along this beach contains a **total** of at least 12 but fewer than 18 of these two types of object.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 1

2 The continuous random variable X has probability density function defined by

$$f(x) = \begin{cases} \dfrac{1}{k} & a \leqslant x \leqslant b \\ 0 & \text{otherwise} \end{cases}$$

(a) Write down, in terms of a and b, the value of k.

[1 mark]

(b) (i) Given that $E(X) = 1$ and $\text{Var}(X) = 3$, find the values of a and b.

[4 marks]

(ii) Four independent values of X are taken. Find the probability that exactly one of these four values is negative.

[3 marks]

QUESTION PART REFERENCE

Answer space for question 2

3 A machine fills bags with frozen peas. Measurements taken over several weeks have shown that the standard deviation of the weights of the filled bags of peas has been 2.2 grams.

Following maintenance on the machine, a quality control inspector selected 8 bags of peas. The weights, in grams, of the bags were

$$910.4 \quad 908.7 \quad 907.2 \quad 913.2 \quad 905.6 \quad 911.1 \quad 909.5 \quad 907.9$$

It may be assumed that the bags constitute a random sample from a normal distribution.

(a) Giving the limits to **four** significant figures, calculate a 95% confidence interval for the mean weight of a bag of frozen peas filled by the machine following the maintenance:

(i) assuming that the standard deviation of the weights of the bags of peas is known to be 2.2 grams;

[4 marks]

(ii) assuming that the standard deviation of the weights of the bags of peas may no longer be 2.2 grams.

[4 marks]

(b) The weight printed on the bags of peas is 907 grams. One of the inspector's concerns is that bags should not be underweight.

Make **two** comments about this concern with regard to the data and your calculated confidence intervals.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 3

4 Wellgrove village has a main road running through it that has a $40\ \text{mph}$ speed limit. The villagers were concerned that many vehicles travelled too fast through the village, and so they set up a device for measuring the speed of vehicles on this main road. This device indicated that the mean speed of vehicles travelling through Wellgrove was $44.1\ \text{mph}$.

In an attempt to reduce the mean speed of vehicles travelling through Wellgrove, life-size photographs of a police officer were erected next to the road on the approaches to the village. The speed, $X\ \text{mph}$, of a sample of 100 vehicles was then measured and the following data obtained.

$$\sum x = 4327.0 \qquad \sum (x - \bar{x})^2 = 925.71$$

(a) State an assumption that must be made about the sample in order to carry out a hypothesis test to investigate whether the desired reduction in mean speed had occurred.

[1 mark]

(b) Given that the assumption that you stated in part **(a)** is valid, carry out such a test, using the 1% level of significance.

[8 marks]

(c) Explain, in the context of this question, the meaning of:

(i) a Type I error;

(ii) a Type II error.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

5 In a particular town, a survey was conducted on a sample of 200 residents aged 41 years to 50 years. The survey questioned these residents to discover the age at which they had left full-time education and the greatest rate of income tax that they were paying at the time of the survey.

The summarised data obtained from the survey are shown in the table.

Greatest rate of income tax paid	Age when leaving education (years)			
	16 or less	17 or 18	19 or more	Total
Zero	32	3	4	39
Basic	102	12	17	131
Higher	17	5	8	30
Total	151	20	29	200

(a) Use a χ^2-test, at the 5% level of significance, to investigate whether there is an association between age when leaving education and greatest rate of income tax paid.

[9 marks]

(b) It is believed that residents of this town who had left education at a later age were more likely to be paying the higher rate of income tax. Comment on this belief.

[1 mark]

QUESTION PART REFERENCE

Answer space for question 5

6 The continuous random variable X has the cumulative distribution function

$$\mathrm{F}(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2}x - \frac{1}{16}x^2 & 0 \leqslant x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

(a) Find the probability that X lies between 0.4 and 0.8 .

[2 marks]

(b) Show that the probability density function, $\mathrm{f}(x)$, of X is given by

$$\mathrm{f}(x) = \begin{cases} \frac{1}{2} - \frac{1}{8}x & 0 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

[1 mark]

(c) (i) Find the value of $\mathrm{E}(X)$.

[3 marks]

(ii) Show that $\mathrm{Var}(X) = \frac{8}{9}$.

[4 marks]

(d) The continuous random variable Y is defined by

$$Y = 3X - 2$$

Find the values of $\mathrm{E}(Y)$ and $\mathrm{Var}(Y)$.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 6

7 Each week, a newsagent stocks 5 copies of the magazine *Statistics Weekly*. A regular customer always buys **one** copy. The demand for **additional** copies may be modelled by a Poisson distribution with mean 2.

The number of copies sold in a week, X, has the probability distribution shown in the table, where probabilities are stated correct to three decimal places.

x	1	2	3	4	5
$P(X=x)$	0.135	0.271	0.271	a	b

(a) Show that, correct to three decimal places, the values of a and b are 0.180 and 0.143 respectively.

[3 marks]

(b) Find the values of $E(X)$ and $E(X^2)$, showing the calculations needed to obtain these values, and hence calculate the standard deviation of X.

[5 marks]

(c) The newsagent makes a profit of £1 on each copy of *Statistics Weekly* that is sold and loses 50 p on each copy that is not sold. Find the mean weekly profit for the newsagent from sales of this magazine.

[2 marks]

(d) Assuming that the weekly demand remains the same, show that the mean weekly profit from sales of *Statistics Weekly* will be greater if the newsagent stocks only 4 copies.

[5 marks]

QUESTION PART REFERENCE	Answer space for question 7