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.

387 375 382 379 381

(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

Do not write outside the box

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 2

Do not write outside the box

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 3

Do not write outside the box

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

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

where $b > a > 0$.

(a) (i) Prove that $k = b - a$.

[2 marks]

(ii) Write down the value of $E(X)$.

[1 mark]

(iii) Show, by integration, that $E(X^2) = \frac{1}{3}(b^2 + ab + a^2)$.

[3 marks]

(iv) Hence derive a simplified formula for $\text{Var}(X)$.

[2 marks]

(b) Given that $a = 4$ and $\text{Var}(X) = 3$, find the numerical value of $E(X)$.

[3 marks]

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

Do not write outside the box

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 5

Do not write outside the box

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 25 29 16 18 27 19 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

Do not write outside the box

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

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 7

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS