Centre Number		Candidate Number	
Surname	MR BARTONS		
Other Names	SOLUTIONS		
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 Subsidiary Examination
June 2015

Mathematics MS1B

Unit Statistics 1B

Wednesday 3 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.
- Unit Statistics 1B has a **written paper only**.

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

Do not write outside the box

Answer **all** questions.

Answer each question in the space provided for that question.

1 The table shows the annual gas consumption, x kWh, and the annual electricity consumption, y kWh, for a sample of 10 bungalows of similar size and occupancy.

x	21 371	18 521	15 222	17 312	19 854	23 561	20 738	22 111	17 897	24 523
y	2281	2327	2221	2378	2787	2856	3078	2647	2566	2559

$$S_{xx} = 76\,581\,640 \qquad S_{yy} = 694\,250 \qquad S_{xy} = 3\,629\,670$$

(a) Calculate the value of r_{xy}, the product moment correlation coefficient between x and y. **[2 marks]**

(b) Interpret your value of r_{xy} in the context of this question. **[2 marks]**

QUESTION PART REFERENCE

Answer space for question 1

1 a) From formula book : $r_{xy} = \dfrac{S_{xy}}{\sqrt{S_{xx} \times S_{yy}}}$

$$= \frac{3\,629\,670}{\sqrt{76581640 \times 694250}}$$

$$= 0.497791\ldots$$

b) Moderate positive linear correlation between gas and electricity consumption

2 The table summarises the diameters, d millimetres, of a random sample of 60 new cricket balls to be used in junior cricket.

Diameter (d mm)	Number of cricket balls	MID POINT
$65 < d \leqslant 66$	5	65.5
$66 < d \leqslant 67$	9	66.5
$67 < d \leqslant 68$	12	67.5
$68 < d \leqslant 69$	15	68.5
$69 < d \leqslant 70$	10	69.5
$70 < d \leqslant 71$	7	70.5
$71 < d \leqslant 72$	2	71.5
Total	**60**	

(a) Calculate estimates of the mean and the variance of these 60 diameters.

[4 marks]

(b) David, a retired professional cricketer, requests that the values calculated in part **(a)** are expressed in inches, rather than in millimetres.

Given that 1 inch is equivalent to approximately 25.4 mm, calculate new values for the mean and the variance in response to David's request.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 2

2 a) For midpoints, see table.

using calculator to enter data →

$\Sigma fx = 4095$

Mean $(\bar{x}) = 68.25$

Sample variance $(s^2) = 1.56903^2$

$= 2.461864...$

b) Linear coding:

Mean $= \dfrac{68.25}{25.4} = 2.68700...$

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 2

VARIANCE $= \dfrac{2.46186...}{25.4^2} = 0.0038159...$

Turn over ▶

Do not write outside the box

3 A ferry sails once each day from port D to port A. The ferry departs from D on time or late but never early. However, the ferry can arrive at A early, on time or late.

The **probabilities** for some combined events of departing from D and arriving at A are shown in the table below.

(a) Complete the table.

[2 marks]

(b) Write down the probability that, on a particular day, the ferry:

(i) both departs and arrives on time;

(ii) departs late.

[2 marks]

(c) Find the probability that, on a particular day, the ferry:

(i) arrives late, given that it departed late;

(ii) does **not** arrive late, given that it departed on time.

[5 marks]

(d) On three particular days, the ferry departs from port D on time.

Find the probability that, on these three days, the ferry arrives at port A early once, on time once and late once. Give your answer to three decimal places.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 3

(a)

		Arrive at A			
		Early	**On time**	**Late**	**Total**
Depart from D	**On time**	0.16	0.56	0.08	0.80
	Late	0.06	0.09	0.05	0.20
	Total	0.22	0.65	0.13	1.00

b) i) 0.56 } From Table
ii) 0.2

c) i) $\frac{0.05}{0.20} = 0.25$ $\left[\frac{P(A_L \cap D_L)}{P(D_L)}\right]$

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 3

ii) $\frac{0.16 + 0.56}{0.8} = \frac{0.72}{0.8} = 0.9$

$$\left[\frac{P(A_L \cap D_T)}{P(D_T)}\right]$$

d) We know (are given) the ferry departs on time

$\rightarrow \left[\frac{0.16}{0.8} \times \frac{0.56}{0.8} \times \frac{0.08}{0.8}\right] \times 3!$

Number of ways of arranging 3 different objects,

$= 0.084$

Do not write outside the box

4 Stephan is a roofing contractor who is often required to replace loose ridge tiles on house roofs. In order to help him to quote more accurately the prices for such jobs in the future, he records, for each of 11 recently repaired roofs, the number of ridge tiles replaced, x_i, and the time taken, y_i hours. His results are shown in the table.

Roof (i)	1	2	3	4	5	6	7	8	9	10	11
x_i	8	11	14	14	16	20	22	22	25	27	30
y_i	5.0	5.2	6.3	7.2	8.0	8.8	10.6	11.0	11.8	12.1	13.0

(a) The pairs of data values for roofs 1 to 7 are plotted on the scatter diagram shown on the opposite page.

Plot the 4 pairs of data values for roofs 8 to 11 on the scatter diagram.

[2 marks]

(b) (i) Calculate the equation of the least squares regression line of y_i on x_i, and draw your line on the scatter diagram.

[6 marks]

(ii) Interpret your values for the gradient and for the intercept of this regression line.

[3 marks]

(c) Estimate the time that it would take Stephan to replace 15 loose ridge tiles on a house roof.

[1 mark]

(d) Given that r_i denotes the residual for the point representing roof i:

(i) calculate the value of r_6;

[2 marks]

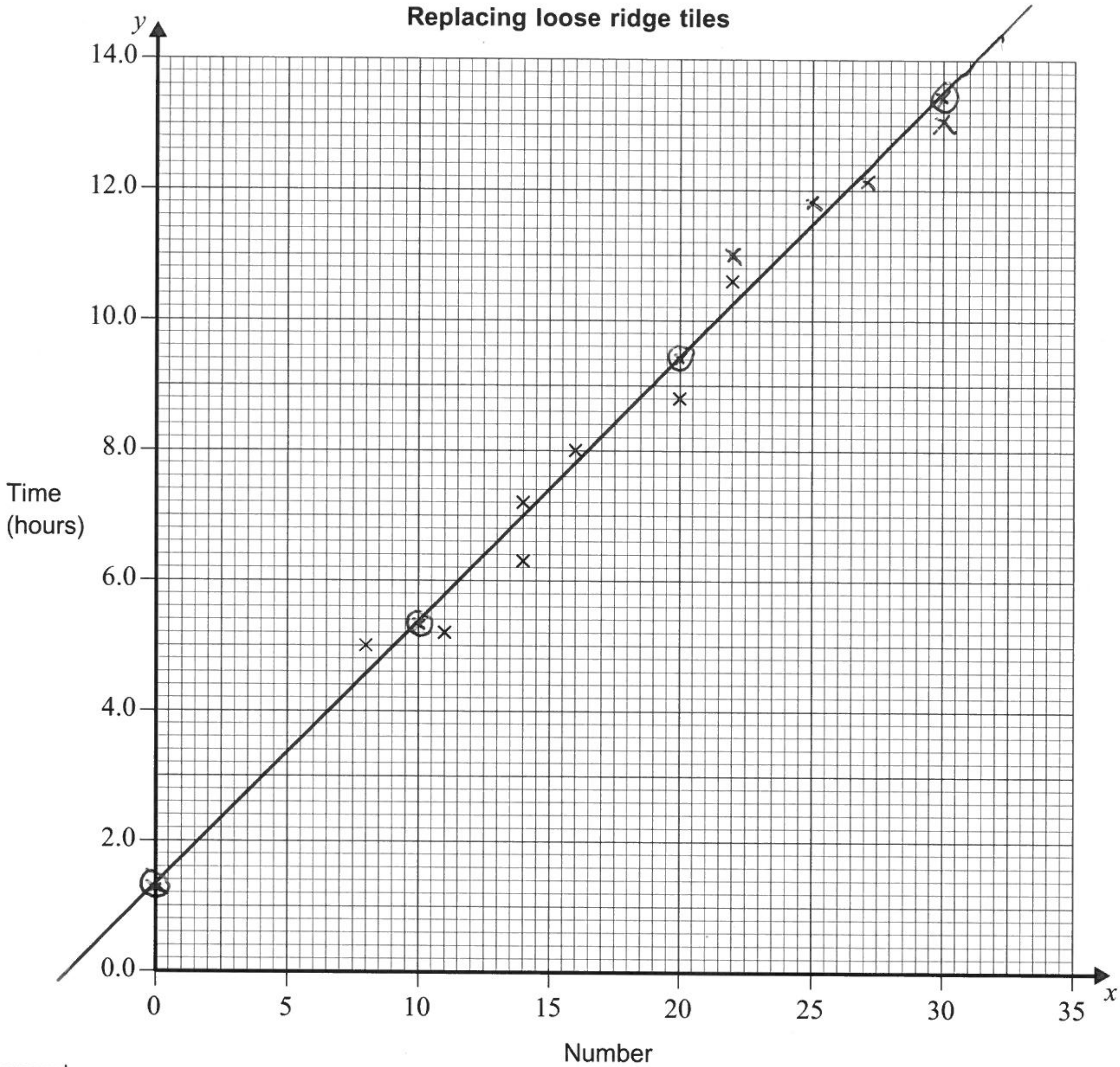
(ii) state why the value of $\sum_{i=1}^{11} r_i$ gives no useful information about the connection between the number of ridge tiles replaced and the time taken.

[1 mark]

QUESTION PART REFERENCE	Answer space for question 4
a)	see graph
b)	i) From calculator:
	$a = 1.30186...$ $\longrightarrow$ $y = 1.302 + 0.405x$
	$b = 0.40517...$

Do not write outside the box

Answer space for question 4



QUESTION PART REFERENCE

To plot, use table of values

x	0	10	20	30
y	1.302	5.352	9.402	13.452

(see ⊗ on diagram)

Turn over ▶

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 4

ii) GRADIENT: Every extra tile takes an extra 0.405 hours

INTERCEPT: The time taken to replace 0 tiles (or maybe to get set up) is 1.302 hours.

c) $y = 1.302 + 0.405(15)$
$= 7.37934\ldots$ hours

d) i) Residual = Actual − Estimated
$= 8.8 - [1.302 + 0.405(20)]$
$= -0.60517\ldots$

ii) The value will always be zero as positive & negative cancel each other out.

5 (a) Wooden lawn edging is supplied in $1.8\,\text{m}$ length rolls. The actual length, X metres, of a roll may be modelled by a normal distribution with mean 1.81 and standard deviation 0.08.

Determine the probability that a randomly selected roll has length:

(i) less than $1.90\,\text{m}$;

(ii) greater than $1.85\,\text{m}$;

(iii) between $1.81\,\text{m}$ and $1.85\,\text{m}$.

[6 marks]

(b) Plastic lawn edging is supplied in $9\,\text{m}$ length rolls. The actual length, Y metres, of a roll may be modelled by a normal distribution with mean μ and standard deviation σ.

An analysis of a batch of rolls, selected at random, showed that

$$\mathrm{P}(Y < 9.25) = 0.88$$

(i) Use this probability to find the value of z such that

$$9.25 - \mu = z \times \sigma$$

where z is a value of $Z \sim \mathrm{N}(0, 1)$.

[2 marks]

(ii) Given also that

$$\mathrm{P}(Y > 8.75) = 0.975$$

find values for μ and σ.

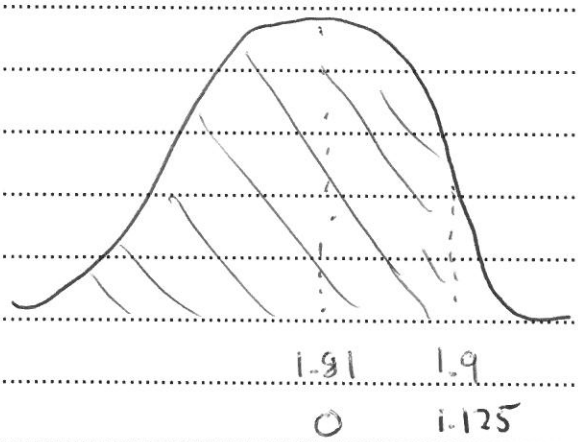
[4 marks]

QUESTION PART REFERENCE	Answer space for question 5

5 a) i) $X \sim N(1.81, 0.08^2)$

$P(X < 1.9)$

$= P\left(Z < \dfrac{1.9 - 1.81}{0.08}\right)$

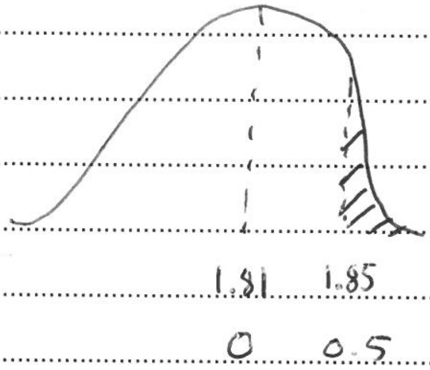


$= P(Z < 1.125)$ $\boxed{X}$

$= 0.86971$ $\boxed{Z}$

QUESTION PART REFERENCE	Answer space for question 5

ii) $P(X > 1.85)$



$= P\left(Z > \dfrac{1.85 - 1.81}{0.08}\right)$ $\boxed{X}$

$\boxed{Z}$

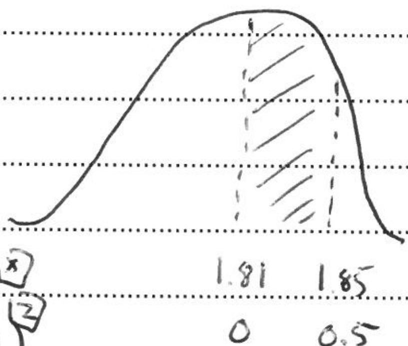
$P(Z > 0.5)$

$= 1 - P(Z < 0.5)$

$= 1 - 0.69146$

$= 0.30854$

iii) $P(1.81 < X < 1.85)$



$= P(X < 1.85) - P(X < 1.81)$

$\boxed{X}$

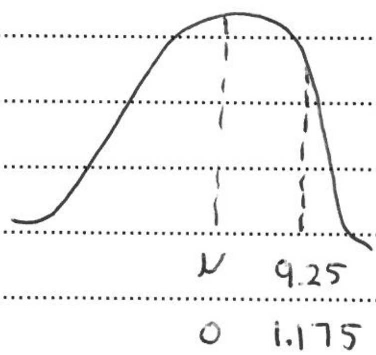
$\boxed{Z}$

$= P(Z < 0.5) - P(Z < 0)$

$= 0.69146 - 0.5$

$= 0.19146$

b) $Y \sim N(\mu, \sigma^2)$



i) $P(Y < 9.25) = 0.88$

$\boxed{Y}$

$P(Z < 1.175) = 0.88$ $\boxed{Z}$

(from Tables)

Turn over ▶

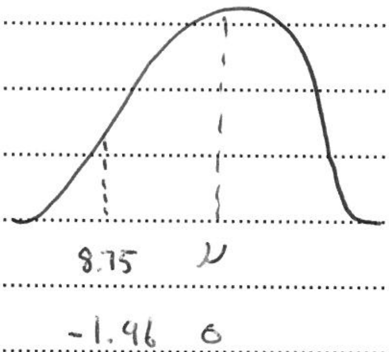
QUESTION PART REFERENCE	Answer space for question 5

Standardise : $\dfrac{9.25 - \mu}{\sigma} = 1.175$

$\rightarrow \quad 9.25 - \mu = 1.175\sigma$

ii)

$P(Y > 8.75) = 0.975$



$P(Z < 1.96) = \cancel{0.025}\ 0.975 \quad \boxed{4}$

$\rightarrow P(Z > -1.96) = 0.975 \quad \boxed{2}$

Standardise : $\dfrac{8.75 - \mu}{\sigma} = -1.96$

$8.75 - \mu = -1.96\sigma$

① $9.25 \boxed{- \mu} = 1.175\sigma$

② $8.75 \boxed{- \mu} = -1.96\sigma$

$-$

$0.5 = 3.315\sigma$

$\rightarrow \sigma = 0.15949...$

using ① $9.25 - \mu = 1.175(0.15949)$

$\rightarrow \mu = 9.06260...$

Do not write outside the box

6 (a) In a particular country, 35 per cent of the population is estimated to have at least one mobile phone.

A sample of 40 people is selected from the population.

Use the distribution $\mathrm{B}(40, 0.35)$ to estimate the probability that the number of people in the sample that have at least one mobile phone is:

(i) at most 15;

(ii) more than 10;

(iii) more than 12 but fewer than 18;

(iv) exactly equal to the mean of the distribution.

[9 marks]

(b) In the same country, 70 per cent of households have a landline telephone connection.

A sample of 50 households is selected from all households in the country.

Stating a necessary condition regarding this selection, estimate the probability that fewer than 30 households have a landline telephone connection.

[4 marks]

QUESTION PART REFERENCE

Answer space for question 6

6 a) $X \sim B(40, 0.35)$

i) $P(X \leq 15) = 0.6946$

ii) $P(X > 10)$

$= 1 - P(X \leq 10)$

$= 1 - 0.1215 = 0.8785$

iii) $P(12 < X < 18)$

Need: 13, 14, 15, 16, 17

$= P(X \leq 17) - P(X \leq 12)$

$= 0.8761 - 0.3143$

$= 0.5618$

QUESTION PART REFERENCE	Answer space for question 6

iv) Mean = np = 40×0.35 = 14

Need $P(X = 14)$

$= {}^{40}C_{14} \times 0.35^{14} \times 0.65^{26}$

$= 0.1313\ldots$

b) $Y \sim B(50, 0.7)$

Assumption: Selection of households is RANDOM!

$P(Y < 30) \rightarrow$ need $0, 1, 2, 3, \ldots 28, 29$

Need to think of failures as probability > 0.5

success (Y)	0	1	2	...	28	29
failure (Y')	50	49	48	...	22	21

$Y' \sim B(50, 0.3)$

Need: $P(Y' \geq 21) = 1 - P(Y' \leq 20)$

$= 1 - 0.9522$

$= 0.0478$

Do not write outside the box

7 (a) The weight of a sack of mixed dog biscuits can be modelled by a normal distribution with a mean of 10.15 kg and a standard deviation of 0.3 kg.

A pet shop purchases 12 such sacks that can be considered to be a random sample.

Calculate the probability that the mean weight of the 12 sacks is less than 10 kg.

[4 marks]

(b) The weight of dry cat food in a pouch can also be modelled by a normal distribution.

The contents, x grams, of each of a random sample of 40 pouches were weighed. Subsequent analysis of these weights gave

$$\bar{x} = 304.6 \quad \text{and} \quad s = 5.37$$

(i) Construct a 99% confidence interval for the mean weight of dry cat food in a pouch. Give the limits to one decimal place.

[4 marks]

(ii) Comment, with justification, on **each** of the following two claims.

Claim 1: The mean weight of dry cat food in a pouch is more than 300 grams.

Claim 2: All pouches contain more than 300 grams of dry cat food.

[4 marks]

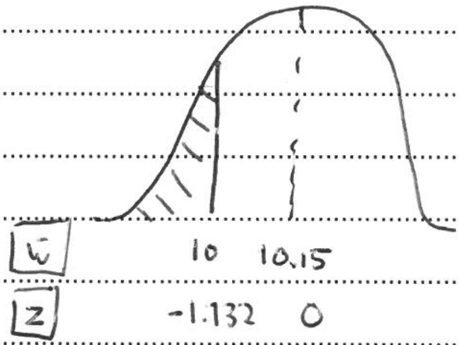
QUESTION PART REFERENCE	Answer space for question 7

7 a) $W \sim N(10.15, 0.3^2)$

$\bar{W} \sim N(10.15, 0.3^2/12)$

$P(\bar{W} < 10)$

$= P\left(Z < \dfrac{10 - 10.15}{0.3/\sqrt{12}}\right)$



$= P(Z < -1.732)$

$= 1 - P(Z < 1.732)$

$= 1 - 0.95837 = 0.04163$

Do not write outside the box

QUESTION PART REFERENCE

Answer space for question 7

b) i) $\bar{x} = 304.6$

$s = 5.37$

$n = 40$

99% CI $= 2.5758$

$\rightarrow \mu = 304.6 \pm 2.5758 \times 5.37/\sqrt{40}$

$= 304.6 \pm 2.187...$

$= (302.\cancel{3}4, \ 306.8)$

ii) $\boxed{\text{CLAIM 1}}$ 300 is below the lower bound of the confidence interval.

Therefore ACCEPT the claim

$\boxed{\text{CLAIM 2}}$ If normally distributed, we would expect SOME data to be within 1 sd of the mean.

$304.6 - 5.37 = 299.23$

This suggests there IS data below 300, so the claim IS likely to be FALSE