AQA

Please write clearly in block capitals.

Centre number | | | | | | Candidate number | | | | |

Surname _____

Forename(s) _____

Candidate signature _____

AS MATHEMATICS

Unit Statistics 1B

Wednesday 6 June 2018 Morning Time allowed: 1 hour 30 minutes

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

You may use a graphics calculator.

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.

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

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.

Answer **all** questions.

Answer each question in the space provided for that question.

1 Bernie recorded the durations, to the nearest minute, of 23 telephone calls made from his landline. The recorded durations, in numerical order, are as follows.

5 5 5 5 10 10 10 13 14 15 16 17 18 19 20 21 22 22 23 30 30 35 95

(a) For these data, give a reason why:

(i) the mode is **not** a suitable measure of average;

(ii) the range is **not** a suitable measure of spread.

[2 marks]

(b) Determine values for the median and the interquartile range of the 23 durations.

[2 marks]

(c) Calculate values for the mean and the standard deviation of the 23 durations.

[3 marks]

(d) Give **two** reasons why the measures determined in part **(b)** might be more appropriate than those calculated in part **(c)** for summarising the 23 durations.

[2 marks]

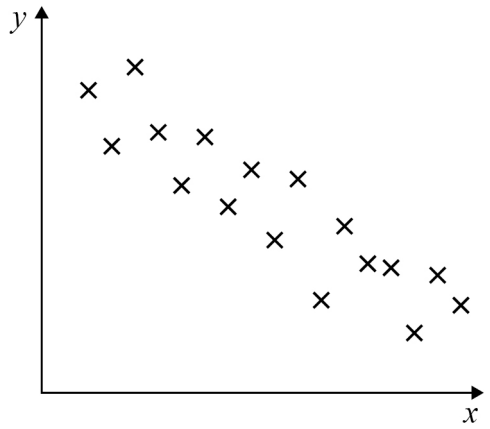
QUESTION PART REFERENCE

Answer space for question 1

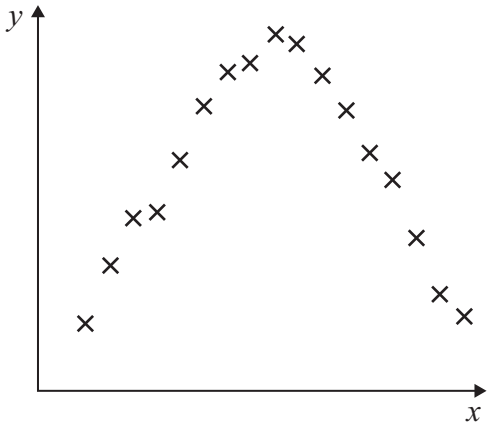
QUESTION PART REFERENCE	**Answer space for question 1**

2 (a) Estimate, **without undertaking any calculations**, the value of the product moment correlation coefficient between the variables x and y for each of the two scatter diagrams.

(i)



(ii)



[2 marks]

(b) Sarah, the manager of a supermarket, recorded the time, t minutes, that each of a sample of 10 customers using shopping trolleys spent in the supermarket, together with the cost, £c, of the items they purchased. The results are shown in the table.

t	15	23	38	32	19	32	27	31	23	50
c	42.97	35.28	56.89	65.75	50.05	22.18	17.36	46.25	77.86	45.41

(i) Calculate, to three significant figures, the value of the product moment correlation coefficient, r, between the variables t and c.

[3 marks]

(ii) Assuming that the sample is representative of all customers using shopping trolleys at the supermarket, interpret, in context, the value for r.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 2

QUESTION PART REFERENCE	**Answer space for question 2**

3 (a) The mass of peanuts, X, in a large tub can be modelled by a normal random variable with a mean of 730 grams and a standard deviation of 20 grams.

Determine the probability that the mass of peanuts in a randomly selected large tub is:

(i) more than 757 grams;

[3 marks]

(ii) between 706 grams and 730 grams.

[2 marks]

(b) The mass of cashew nuts, Y, in a small tub can be modelled by a normal random variable with a mean of μ grams and a standard deviation of σ grams.

(i) Given that $\mathrm{P}(Y > 320) = 0.99$ and $\mathrm{P}(Y > 350) = 0.50$, find, to the nearest gram, values for μ and σ.

[4 marks]

(ii) Hence find the value of w such that $\mathrm{P}(\mu - w < Y < \mu + w) = 0.80$.

[3 marks]

QUESTION PART REFERENCE	**Answer space for question 3**

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

QUESTION PART REFERENCE

Answer space for question 3

4 Part of an investigation into the yield of tomato plants involved an analysis of the effect of different concentrations of potassium in a plant's liquid feed.

Thirteen tomato plants were grown under identical conditions except for a different concentration, x milligrams per litre, of potassium in each plant's liquid feed. The resultant yield of tomatoes, y kilograms, from each plant was measured. The results are shown in the table.

Plant	A	B	C	D	E	F	G	H	I	J	K	L	M
x	0	10	20	30	40	50	60	70	80	90	100	110	120
y	0.31	0.69	0.97	1.28	1.34	1.99	2.44	2.45	2.94	3.61	3.96	4.03	4.54

(a) State why the least squares regression line of x **on** y would **not** be appropriate for these data.

[1 mark]

(b) (i) Calculate the equation of the least squares regression line of y **on** x.

[4 marks]

(ii) Hence interpret, in context, the values for the line's intercept and gradient.

[3 marks]

(c) (i) Calculate the value of the residual for the point representing plant H.

[2 marks]

(ii) Hence, given that the residual for the point representing plant G is $+0.09$, find the sum of the remaining 11 residuals.

[2 marks]

(d) During a follow-up investigation, the following data were collected.

Plant	N	O
x	150	200
y	4.63	4.89

Give a **general reason** and a **specific reason based on numerical support** why your equation calculated in part **(b)(i)** should not be used to estimate the yield of a tomato plant when given between $150\,\text{mg/l}$ and $200\ \text{mg/l}$ of potassium in the liquid feed.

[2 marks]

QUESTION PART REFERENCE

Answer space for question 4

QUESTION PART REFERENCE	**Answer space for question 4**

QUESTION PART REFERENCE	**Answer space for question 4**

QUESTION PART REFERENCE

Answer space for question 4

5 (a) For **each** of the three variables described below, state whether the distribution $\mathrm{B}(n, p)$ is an appropriate model.
If such a model is appropriate, give values for n and p.
If such a model is **not** appropriate, give a reason why.

(i) Variable U denotes the number of scores of 'five or six' when an unbiased six-sided die is rolled 20 times.

(ii) Variable V denotes the number of tosses of an unbiased coin until exactly 5 heads have been observed.

(iii) Variable W denotes the number of yellow highlighter pens in a random sample of 5 pens, selected without replacement from a box containing 50 highlighter pens, of which exactly 10 are yellow.

[5 marks]

(b) On a particular island, with an adult population of many thousands, 15 per cent of men and 10 per cent of women are left-handed.

Use an appropriate binomial distribution in each case to estimate the probability that:

(i) a sample of 25 men contains at most 3 who are left-handed;

(ii) a sample of 40 women contains at least 2 but at most 6 who are left-handed;

(iii) a sample of 50 women contains more than 40 who are **not** left-handed.

[7 marks]

QUESTION PART REFERENCE	**Answer space for question 5**

QUESTION PART REFERENCE	**Answer space for question 5**

QUESTION PART REFERENCE	**Answer space for question 5**

QUESTION PART REFERENCE

Answer space for question 5

6 A company produces two types of disinfectant, *Pine* and *Lemon*. For each type of disinfectant, the volume in a 1-litre bottle can be modelled by a normal distribution with a standard deviation of 36 millilitres.

(a) The mean volume of *Pine* disinfectant in a random sample of **fifty** 1-litre bottles was found to be 1075 millilitres.

(i) Construct a 95% confidence interval for the mean volume, μ millilitres, of *Pine* disinfectant in 1-litre bottles. Give the limits to the nearest millilitre.

[4 marks]

(ii) Hence comment, with justification, on a claim that $\mu = 1072$ millilitres.

[2 marks]

(b) Based on a random sample of $\boldsymbol{n}$ 1-litre bottles of *Lemon* disinfectant, a $\mathbf{98\%}$ confidence interval for the mean volume was calculated as $(1060, 1080)$ millilitres.

Calculate the value of n.

[4 marks]

QUESTION PART REFERENCE	**Answer space for question 6**

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

QUESTION PART REFERENCE

Answer space for question 6

7 (a) The events L and M are such that

$$\mathrm{P}(L) = 0.55 \quad \text{and} \quad \mathrm{P}(M) = 0.28 \, .$$

Write down the value of:

(i) $\mathrm{P}(L \cap M)$ if L and M are independent;

(ii) $\mathrm{P}(L \cup M)$ if L and M are mutually exclusive;

(iii) $\mathrm{P}(L \cup M)$ if L and M are independent.

[3 marks]

(b) Rhonda, Samantha and Tracy are members of a club which meets every Wednesday.

At any Wednesday meeting, Rhonda's attendance, event R, has probability 0.94, Samantha's attendance, event S, has probability 0.88, and Tracy's attendance, event T, has probability 0.76. The events R, S and T are independent.

For these three members, calculate the probability that, on a particular Wednesday:

(i) all of them attend the meeting;

[1 mark]

(ii) exactly one of them attends the meeting;

[2 marks]

(iii) at least two of them attend the meeting.

[2 marks]

(iv) Ursula, a neighbour of Tracy, is also a member of the club. At any Wednesday meeting, Ursula's attendance, event U, is independent of events R and S but $\mathrm{P}(U|T) = 0.96$ and $\mathrm{P}(U|T') = 0.48$.

For Rhonda, Samantha, Tracy and Ursula, calculate the probability that, on a particular Wednesday:

(A) all of them attend the meeting;

(B) none of them attend the meeting.

[3 marks]

QUESTION PART REFERENCE	**Answer space for question 7**

QUESTION PART REFERENCE

Answer space for question 7

QUESTION PART REFERENCE

Answer space for question 7

END OF QUESTIONS