Please write clearly in block capitals.

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

Surname _____

Forename(s) _____

Candidate signature _____

AS
STATISTICS

Unit Statistics 1B

Wednesday 23 May 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.

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.

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

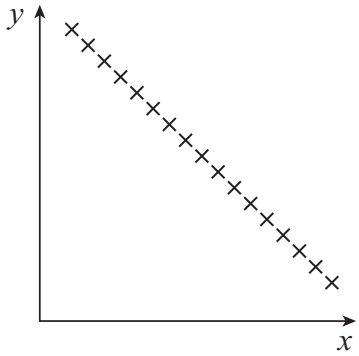
Do not write outside the box

Answer **all** questions.

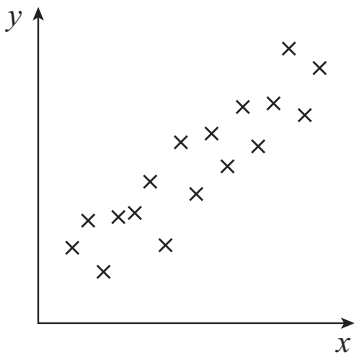
Answer each question in the space provided for that question.

1 (a) Estimate, **without undertaking any calculations**, the value of the product moment correlation coefficient between the variables x and y for **each** of **Scatter Diagrams 1 and 2**.

(i) Scatter Diagram 1



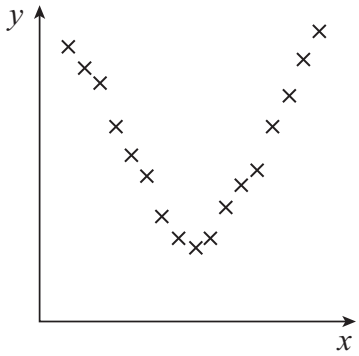
(ii) Scatter Diagram 2



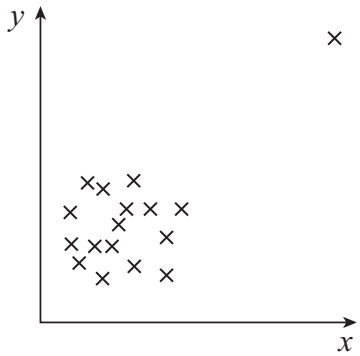
[2 marks]

(b) For **each** of **Scatter Diagrams 3 and 4**, give a reason why a calculation of the product moment correlation coefficient would not be appropriate.

(i) Scatter Diagram 3



(ii) Scatter Diagram 4



[2 marks]

QUESTION PART REFERENCE	**Answer space for Questions 1(a) and 1(b)**

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Questions 1(a) and 1(b)

Question 1 continues on the next page

Do not write outside the box

1 (c) Christine, a trainee biologist, was investigating the characteristics of a particular variety of cucumber.

She selected a sample of 20 cucumbers.

For each cucumber, she measured the length, l centimetres, the maximum diameter, d centimetres, and the weight, w grams.

Christine then calculated values of the product moment correlation coefficients, reporting her results as follows:

(i) 0.481 between l and d;

(ii) -0.866 between l and w;

(iii) 1.046 between d and w.

For **each** of Christine's three calculated values, state whether the value is definitely correct, possibly correct, probably incorrect or definitely incorrect.

[3 marks]

QUESTION PART REFERENCE	**Answer space for Question 1(c)**

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 1(c)

Turn over ▶

Do not write outside the box

2 The table summarises the volume of water, v cubic metres, used during one month by each of the 100 households on a new estate.

Volume of water (v)	**Number of households**
$0 \leqslant v < 2$	1
$2 \leqslant v < 4$	5
$4 \leqslant v < 6$	13
$6 \leqslant v < 8$	18
$8 \leqslant v < 10$	27
$10 \leqslant v < 15$	16
$15 \leqslant v < 20$	12
$20 \leqslant v < 30$	8

(a) Calculate estimates for the mean and the standard deviation of these data.

[4 marks]

(b) For each household, the monthly cost for water used is a fixed cost of £1.85 plus a cost of £1.30 per cubic metre of water used.

For the households on this new estate, calculate, to the nearest 1p, estimates for the mean and the standard deviation of the monthly cost of water used.

[4 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 2

Turn over ▶

Do not write outside the box

3 (a) The weight of cheese, X, in a **large pack** can be modelled by a normal random variable with a mean of 365.0 grams and a standard deviation of 10.0 grams.

Determine the probability that the weight of cheese in a randomly selected large pack is:

(i) less than 377.5 grams;

(ii) more than 367.5 grams;

(iii) between 365.0 grams and 367.5 grams.

[6 marks]

(b) The weight of cheese, Y, in an **extra-large pack** can be modelled by a normal random variable with mean 475.0 grams, unknown standard deviation σ grams and $\mathrm{P}(Y < 450.0) = 0.05$.

(i) Find, to the nearest 0.1 gram, the value of σ.

[3 marks]

(ii) Determine the probability that in a sample of 12 randomly selected extra-large packs:

(A) every pack contains more than 450.0 grams of cheese;

(B) the **mean** weight of cheese per pack is more than 470.0 grams.

[5 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 3

Turn over ▶

Do not write outside the box

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 Large bags of *Luckidips* contain exactly 50 chocolates. Each chocolate has the same shape and is wrapped with the same silver foil.

The type of chocolate coating and the type of centre of the 50 chocolates in each bag are as follows.

		Coating		
		Milk	**White**	**Dark**
Centre	**Soft**	22	8	0
	Hard	6	6	8

(a) Munir selects at random a chocolate from a bag of 50 *Luckidips*.

Calculate the probability that his selected chocolate has:

(i) either a hard centre or a white coating or both;

(ii) either a soft centre or a milk coating but **not** both;

(iii) a soft centre, given that it has a milk coating.

[4 marks]

(b) Ning selects at random, without replacement, **four** chocolates from a second bag of 50 *Luckidips*.

Calculate the probability that in her selected chocolates:

(i) none have both a dark coating and a soft centre;

(ii) exactly two have a milk coating;

(iii) at least one has both a milk coating and a soft centre.

[8 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 4

Turn over ▶

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

Turn over ▶

Do not write outside the box

5 Ethylene glycol, commonly known as antifreeze, is added to water in a vehicle's radiator to lower the freezing point of the resultant fluid.

The table shows the percentage, x, of antifreeze by volume in the fluid and the fluid's corresponding freezing point, y ºC.

x	0.0	5.0	10.0	15.0	20.0	25.0	30.0
y	0.0	–2.1	–4.1	–6.3	–8.3	–11.9	–14.7

(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.

[3 marks]

(ii) Hence interpret, in context, the value you obtained for the line's gradient.

[2 marks]

(c) (i) Calculate the value of the residual for the point $(15, -6.3)$.

[2 marks]

(ii) Hence, given that the residual for the point $(30, -14.7)$ is -0.65, correct to two significant figures, find the sum of the remaining 5 residuals.

[2 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 5

Turn over ▶

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

Turn over ▶

Do not write outside the box

6 Each of the four major blood groups is split into positive and negative blood types.

The percentage of the population of Brazil with each blood type is shown in the table.

Blood group	**O**		**A**		**B**		**AB**	
Blood type	**O+**	**O–**	**A+**	**A–**	**B+**	**B–**	**AB+**	**AB–**
Percentage	36	9	34	8	8	2	2	1

Estimate the probability that, in Brazil, a random sample of:

(a) 10 people contains exactly 3 with blood type A+;

[2 marks]

(b) 20 people contains fewer than 3 with blood group B;

[1 mark]

(c) 30 people contains more than 10 with blood group O;

[2 marks]

(d) 40 people contains at most 20 with blood groups other than A and AB;

[3 marks]

(e) 50 people contains at least 6 but at most 12 with blood types that are negative.

[4 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 6

Turn over ▶

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 6

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 6

Turn over ▶

Do not write outside the box

7 A local authority has established that the time taken to collect recyclable household waste from 'bags & boxes' in a particular area has a mean of 390 minutes and a standard deviation of 25 minutes.

In order to assess the effect on the mean collection time of a change from '**bags** & boxes' to '**bins** & boxes', a sample of 34 collections was timed.

The **total time** for the 34 collections of 'bins & boxes' was $13\,668$ minutes.

(a) (i) Assuming that the value of the standard deviation remained unchanged at 25 minutes, construct a 90% confidence interval for the mean time taken when collecting from 'bins & boxes'. Give the limits to the nearest minute.

[5 marks]

(ii) Hence comment on a claim that the change to 'bins & boxes' has increased the mean time taken.

[2 marks]

(iii) If, instead of a 90% confidence interval in part **(a)(i)**, an 80% confidence interval had been calculated, state, with a reason, whether this would have changed your answer to part **(a)(ii)**.

[2 marks]

(b) In answering part **(a)(i)**:

(i) state an assumption needed about the sample of 34 collections;

(ii) explain fully why no assumption was needed about the distribution of collection times.

[3 marks]

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

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 7

Turn over ▶

Do not write outside the box

QUESTION PART REFERENCE	Answer space for Question 7

Do not write outside the box

QUESTION PART REFE)RENCE	Answer space for Question 7

END OF QUESTIONS

Do not write outside the box

There are no questions printed on this page

DO NOT WRITE ON THIS PAGE
ANSWER IN THE SPACES PROVIDED

Copyright information

For confidentiality purposes, from the November 2015 examination series, acknowledgements of third party copyright material will be published in a separate booklet rather than including them on the examination paper or support materials. This booklet is published after each examination series and is available for free download from www.aqa.org.uk after the live examination series.

Permission to reproduce all copyright material has been applied for. In some cases, efforts to contact copyright-holders may have been unsuccessful and AQA will be happy to rectify any omissions of acknowledgements. If you have any queries please contact the Copyright Team, AQA, Stag Hill House, Guildford, GU2 7XJ.

Copyright © 2018 AQA and its licensors. All rights reserved.