Write your name here

Surname	Other names

Pearson Edexcel
International
Advanced Level

Centre Number | | | | | |

Candidate Number | | | | |

Statistics S1

Advanced/Advanced Subsidiary

Friday 20 January 2017 – Afternoon
Time: 1 hour 30 minutes

Paper Reference
WST01/01

You must have:
Mathematical Formulae and Statistical Tables (Blue)

Total Marks

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided
  – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

P48246A

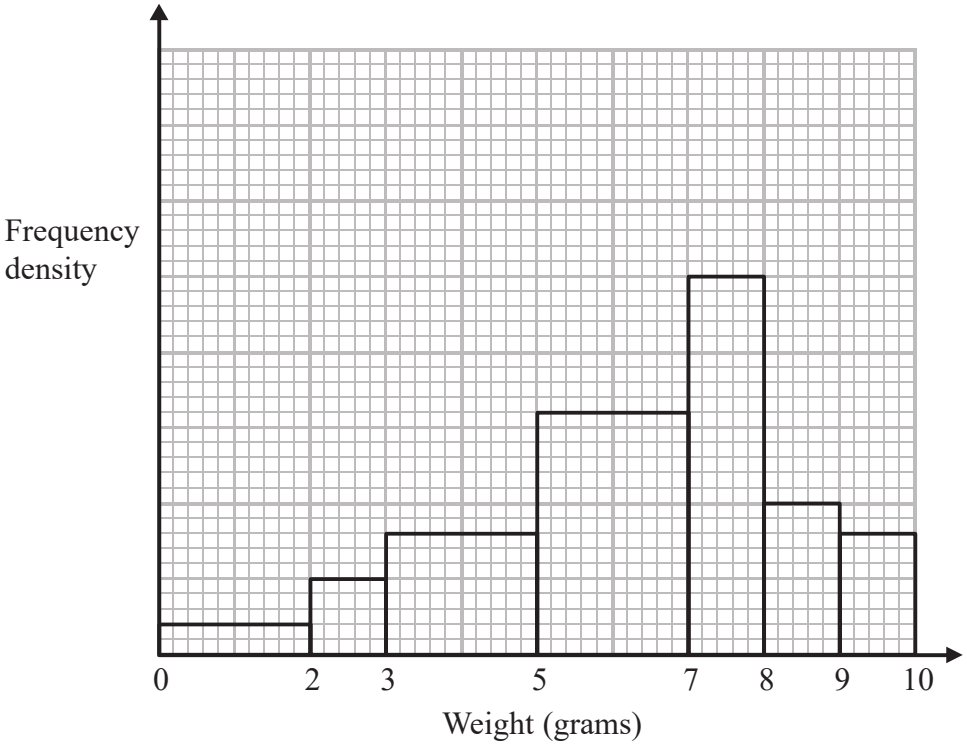
1/1/1/1/1/

P 4 8 2 4 6 A 0 1 2 8

Pearson

Leave blank

1. Ralph records the weights, in grams, of 100 tomatoes. This information is displayed in the histogram below.



Given that 5 of the tomatoes have a weight between 2 and 3 grams,

(a) find the number of tomatoes with a weight between 0 and 2 grams.

(2)

One of the tomatoes is selected at random.

(b) Find the probability that it weighs more than 3 grams.

(2)

(c) Estimate the proportion of the tomatoes with a weight greater than 6.25 grams.

(2)

(d) Using your answer to part (c), explain whether or not the median is greater than 6.25 grams.

(1)

Given that the mean weight of these tomatoes is 6.25 grams and using your answer to part (d),

(e) describe the skewness of the distribution of the weights of these tomatoes. Give a reason for your answer.

(1)

Two of these 100 tomatoes are selected at random.

(f) Estimate the probability that both tomatoes weigh within 0.75 grams of the mean.

(4)

Leave blank

Question 1 continued

Leave blank

Question 1 continued

Leave blank

Question 1 continued

(Total 12 marks)

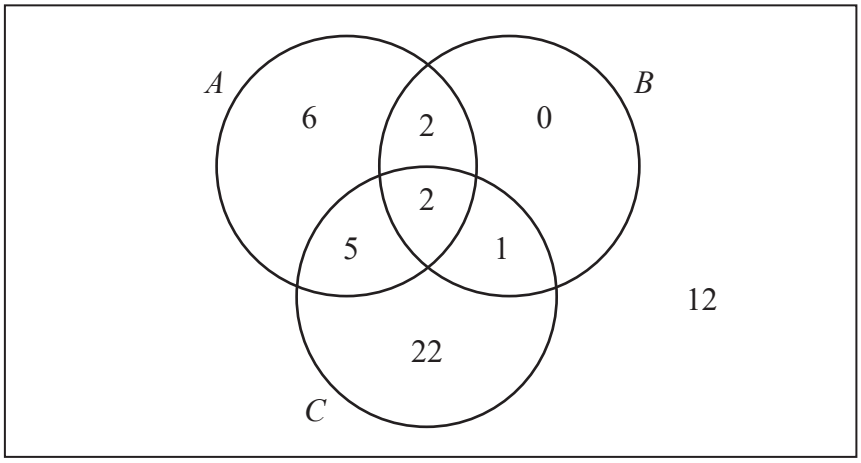
Q1

Leave blank

2. An integer is selected at random from the integers 1 to 50 inclusive.

A is the event that the integer selected is prime.
B is the event that the integer selected ends in a 3
C is the event that the integer selected is greater than 20

The Venn diagram shows the number of integers in each region for the events A, B and C



(a) Describe in words the event $(A \cap B)$

(1)

(b) Write down the probability that the integer selected is prime.

(1)

(c) Find $P([A \cup B \cup C]')$

(1)

Given that the integer selected is greater than 20

(d) find the probability that it is prime.

(2)

Using your answers to (b) and (d),

(e) state, with a reason, whether or not the events A and C are statistically independent.

(2)

Given that the integer selected is greater than 20 and prime,

(f) find the probability that it ends in a 3

(2)

Leave blank

Question 2 continued

Leave blank

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

Q2

(Total 9 marks)

Leave blank

3. A scientist measured the salinity of water, x g/kg, and recorded the temperature at which the water froze, y °C, for 12 different water samples. The summary statistics are listed below.

$$\sum x = 504 \qquad \sum y = -27 \qquad \sum x^2 = 22842 \qquad \sum y^2 = 62.98$$

$$\sum xy = -1190.7 \qquad S_{xx} = 1674 \qquad S_{yy} = 2.23$$

(a) Find the mean and variance of the recorded temperatures.

(3)

Priya believes that the higher the salinity of water, the higher the temperature at which the water freezes.

(b) (i) Calculate the product moment correlation coefficient between x and y

(ii) State, with a reason, whether or not this value supports Priya's belief.

(4)

(c) Find the least squares regression line of y on x in the form $y = a + bx$
Give the value of a and the value of b to 3 significant figures.

(4)

(d) Estimate the temperature at which water freezes when the salinity is 32 g/kg

(1)

The coding $w = 1.8y + 32$ is used to convert the recorded temperatures from °C to °F

(e) Find an equation of the least squares regression line of w on x in the form $w = c + dx$

(2)

(f) Find

(i) the variance of the recorded temperatures when converted to °F

(ii) the product moment correlation coefficient between w and x

(3)

Leave blank

Question 3 continued

Leave blank

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 3 continued

Q3

(Total 17 marks)

Leave blank

4. In a game, the number of points scored by a player in the first round is given by the random variable X with probability distribution

x	5	6	7	8
$P(X = x)$	0.13	0.21	0.29	0.37

Find

(a) $E(X)$

(2)

(b) $\text{Var}(X)$

(3)

(c) $\text{Var}(3 - 2X)$

(2)

The number of points scored by a player in the second round is given by the random variable Y and is independent of the number of points scored in the first round.

The random variable Y has probability function

$$P(Y = y) = \frac{1}{4} \qquad \text{for } y = 5, 6, 7, 8$$

(d) Write down the value of $E(Y)$

(1)

(e) Find $P(X = Y)$

(2)

(f) Find the probability that the number of points scored by a player in the first round is greater than the number of points scored by the player in the second round.

(3)

Leave blank

Question 4 continued

Leave blank

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued

(Total 13 marks)

Q4

Leave blank

5. In a survey, people were asked if they use a computer every day.

Of those people under 50 years old, 80% said they use a computer every day.
Of those people aged 50 or more, 55% said they use a computer every day.

The proportion of people in the survey under 50 years old is p

(a) Draw a tree diagram to represent this information.

(2)

In the survey, 70% of all people said they use a computer every day.

(b) Find the value of p

(2)

One person is selected at random. Given that this person uses a computer every day,

(c) find the probability that this person is under 50 years old.

(2)

Leave blank

Question 5 continued

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 5 continued

Q5

(Total 6 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

6. In a factory, a machine is used to fill bags of rice. The weights of bags of rice are modelled using a normal distribution with mean 250 g.

Only 1% of the bags of rice weigh more than 256 g.

(a) Write down the percentage of bags of rice with weights between 244 g and 256 g.

(1)

(b) Find the standard deviation of the weights of the bags of rice.

(3)

An inspection consists of selecting a bag of rice at random and checking if its weight is within 4 g of the mean. If the weight is more than 4 g away from the mean, then a second bag of rice is selected at random and checked. If the weight of each of the 2 bags of rice is more than 4 g away from the mean, then the machine is shut down.

(c) Find the probability that the machine is shut down following an inspection.

(4)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Q6

(Total 8 marks)

Leave blank

7. The discrete random variable X can take only the values 1, 2, 3 and 4. For these values, the probability function is given by

$$P(X = x) = \frac{ax + b}{60} \qquad \text{for } x = 1, 2, 3, 4$$

where a and b are constants.

(a) Show that $5a + 2b = 30$

(2)

Given that $F(3) = \frac{13}{20}$

(b) find the value of a and the value of b

(4)

Given also that $Y = X^2$

(c) find the cumulative distribution function of Y

(4)

Leave blank

Question 7 continued

Leave blank

Question 7 continued

Q7

(Total 10 marks)

TOTAL FOR PAPER: 75 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA