Please check the examination detai	ls below before ent	ering your candidate information
Candidate surname		Other names
Pearson	Centre Number	Candidate Number
Edexcel GCE		
Wednesday 1	2 June	2019
Morning (Time: 1 hour 30 minute:	s) Paper R	deference 6683/01
Statistics S1		
Advanced/Advanced Su	ubsidiary	
You must have: Mathematical Formulae and Stati	istical Tables (Pi	nk)

Candidates may use any calculator allowed by Pearson regulations.

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 quide 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.

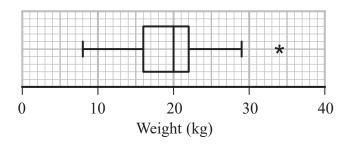
Turn over ▶







1. The box plot summarises the weights of luggage for a group of tourists at an airport.



(a) Write down the range.

(1)

A quarter of the tourists had to pay an extra charge because their luggage was too heavy.

(b) Write down an estimate of the maximum weight that is allowed before having to pay the extra charge.

(1)

Four tourists with luggage weighing 10 kg, 17 kg, 21 kg and 25 kg join the group.

(c) State what, if any, changes this will make to the box plot. Give a reason for your answer.

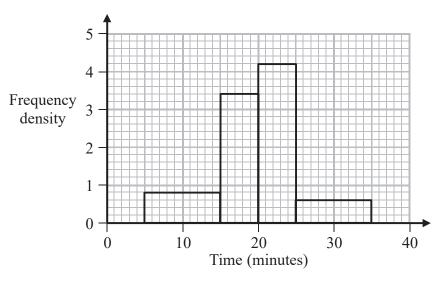
(1)



Question 1 continued		blank
	Question 1 continued	
(Total 3 marks)		Q1
	(Total 3 marks)	



2. Peta recorded the times, t minutes, a group of children took to swim 300 metres. She summarised these times in the following histogram.



(a) Calculate the number of children who took part in the swim.

(2)

Adam used the histogram to estimate the mean and standard deviation of the times taken by the children to complete the swim.

(b) Find Adam's estimate of

(i) the mean,

(2)

(ii) the standard deviation.

(3)

Adam used linear interpolation to estimate the median time taken by the children to complete the swim.

(c) Find Adam's estimate of the median.

(2)

Peta also calculated the mean, standard deviation and median of the times taken by the children but she used each child's actual time taken to complete the swim. She obtained a mean time of 20.8 minutes, a standard deviation of 5.51 minutes and a median time of 20.5 minutes.

(d) Explain an assumption Adam made about these data that has led him to get different answers to Peta.

(1)

Adam and Peta each calculate a coefficient of skewness by using their statistics in the formula

$$skewness = \frac{3(mean - median)}{standard deviation}$$

- (e) (i) Calculate the coefficients of skewness found by Adam and Peta.
 - (ii) Suggest how Peta could improve her histogram to describe the data more accurately.



Question 2 continued	blank



estion 2 continued		
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_
		_

Question 2 continued	
	Q2
(Total 12 marks)	



3. The 3 events A, B and C have probabilities

$$P(A) = 0.6$$

$$P(B) = 0.8$$

$$P(C) = 0.25$$

The events A and B are independent, the events A and C are mutually exclusive and $P(A \cup B \cup C) = 1$

- (a) Find (i) $P(A \cap B)$
 - (ii) $P(A \cap C)$

(3)

(b) Draw a Venn diagram to show the events A, B and C and all the associated probabilities.

(7)

(c) Find $P(B \mid C')$

(3)

(d) State, giving a reason, whether or not the events *B* and *C* are independent.



estion 3 continued	
Selon 5 continued	



Question 3 continue	d		

Question 3 continued	Leave blank
	Q3
(Total 15 marks)	
(Total 13 marks)	



- 4. The journey time for Charlie to travel to the train station is normally distributed with mean 25 minutes and standard deviation 6 minutes. If Charlie's journey time is less than 20 minutes, he will have time to buy a coffee before boarding his train.
 - (a) Find the probability that, on a randomly selected day, Charlie has time to buy a coffee.

(3)

In a particular week, Charlie travels to the station on 2 days.

(b) Find the probability that he has time to buy a coffee on just one of these days.

(3)

On 1% of days, Charlie's journey time is longer than m minutes.

(c) Find the value of m.

(3)

A new road is built and this should make Charlie's journey time to the station more reliable.

(d) State, giving a reason, what effect you would expect this to have on the standard deviation of Charlie's journey times to the station.

(1)

After the new road was built, Charlie's journey time to the station is still modelled by a normal distribution with mean 25 minutes but the standard deviation, σ , has changed. In the first 20 journeys to the station using this new road, Charlie took longer than 30 minutes on 3 occasions.

(e) Estimate the value of	σ .
---------------------------	------------

(4)





Question 4 continued	ł
Question 4 Continued	
	1



Question 4 continued		

Question 4 continued	blank
Question 4 continued	
	Q4
(Total 14 marks)	



5. The random variable *X*, with the following probability distribution, represents the score when a 4-sided die is rolled.

x	1	2	3	4
P(X=x)	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

(a) Write down the name of this probability distribution.

(1)

The die is rolled 3 times.

- (b) Find the probability that
 - (i) the score is no more than 2 on each of the 3 rolls of the die,

(2)

(ii) the score on each of the 3 rolls is different.

(3)

The die is now rolled twice. The random variable X_1 represents the score on the first roll and the random variable X_2 represents the score on the second roll. The random variable M is the maximum of X_1 and X_2

(c) Complete the table below to show the values of M

Values of M

		, , ,	011/1	
X_1 X_2	1	2	3	4
1		2		
2	2	2	3	
3				
4				

(2)

(d) Hence find the probability distribution for M

(2)

- (e) Find the exact value of
 - (i) E(M)

(2)

(ii) Var(M)

(3)

(f) Find $P(X_1 \le 2 | M = 4)$

estion 5 continued	



Question 5 continued	

Question 5 continued	Leave blank
	Q5
(Total 17 marks)	
(10millimins)	



6.	8 students studying Spanish and records their vocabulary, v words, and the n of years, y, that they have been studying Spanish. Figure 1, opposite, shows a sket the scatter diagram for these data.		
	(a) State, giving a reason, whether or not a linear regression model might be suitable to model the relationship between v and y.(1)		

Question 6 continued

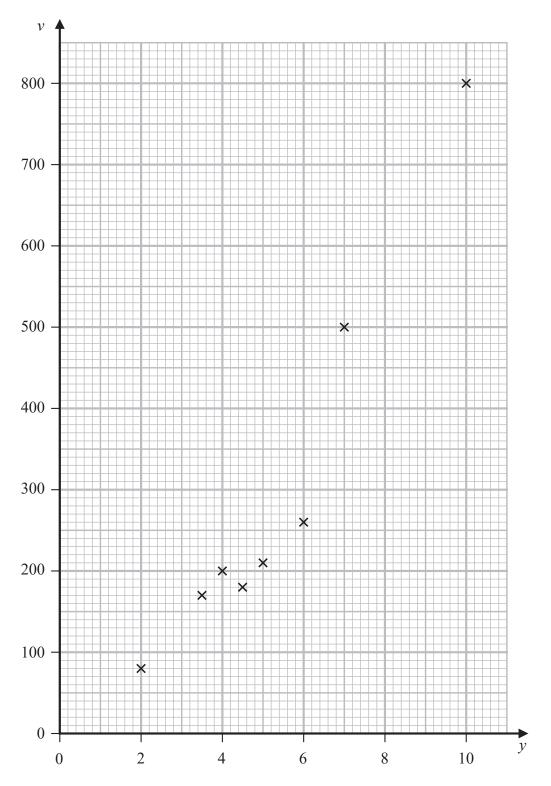


Figure 1

(Question 6 continues on page 22)



Question 6 continued

Maria's data is summarised as

$$\sum y = 42$$
 $\sum v = 2400$ $S_{vv} = 389400$ $S_{yy} = 42$ $\sum vy = 16475$

(b) Calculate the value of the product moment correlation coefficient for these data.

(3)

(c) State, giving a reason, whether or not this result is consistent with your conclusion in part (a).

(1)

Maria wishes to estimate the vocabulary of a student who has been learning Spanish for 5 years.

(d) By calculating a suitable regression line, find an estimate for the vocabulary of this student.

(6)

(e) Give an interpretation of the gradient of your regression line.

(1)

Frederic, who has studied Spanish for 2 years at the college, was part of this sample. His vocabulary is 80 words.

(f) With reference to this information and the scatter diagram, suggest how Maria's model for the relationship between *v* and *y* could be improved and sketch a suitable alternative model.





uestion 6 continued	



uestion 6 continued		blank
acstron o continued		
		Q6
	(Total 14 marks)	