

Please check the examination details below before entering your candidate information

Candidate surname

Other names

Pearson Edexcel
International
Advanced Level

Centre Number

--	--	--	--	--

Candidate Number

--	--	--	--

Wednesday 12 June 2019

Morning (Time: 1 hour 30 minutes)

Paper Reference **WST01/01**

Mathematics

**International Advanced Subsidiary/Advanced Level
Statistics S1**

You must have:

Mathematical Formulae and Statistical Tables (Blue), calculator

Total Marks

Candidates may use any calculator permitted 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).
- **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.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 6 questions in this question paper. 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.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

Turn over ►

P56169A

©2019 Pearson Education Ltd.

1/1/1/

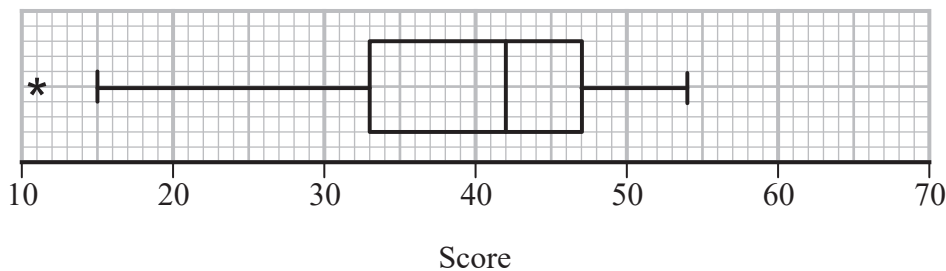



Pearson

2. Chi wanted to summarise the scores of the 39 competitors in a village quiz. He started to produce the following stem and leaf diagram

Score	Key: 2 5 is a score of 25
1 1 5 8 9	
2 0 2 5 8 9	
3 3 5 5 7 8 9 ...	

He did not complete the stem and leaf diagram but instead produced the following box plot.



Chi defined an outlier as a value that is

$$\text{greater than } Q_3 + 1.5 \times (Q_3 - Q_1)$$

or

$$\text{less than } Q_1 - 1.5 \times (Q_3 - Q_1)$$

(a) Find

- (i) the interquartile range
- (ii) the range.

(2)

(b) Describe, giving a reason, the skewness of the distribution of scores.

(2)

Albert and Beth asked for their scores to be checked.

Albert's score was changed from 25 to 37

Beth's score was changed from 54 to 60

(c) On the grid on page 5, draw an updated box plot.
Show clearly any calculations that you used.

(7)

Some of the competitors complained that the questions were biased towards the younger generation. The product moment correlation coefficient between the age of the competitors and their score in the quiz is -0.187

(d) State, giving a reason, whether or not the complaint is supported by this statistic.

(2)



3. A certain disease occurs in a population in 2 mutually exclusive types.

It is difficult to diagnose people with type A of the disease and there is an unknown proportion p of the population with type A.

It is easier to diagnose people with type B of the disease and it is known that 2% of the population have type B.

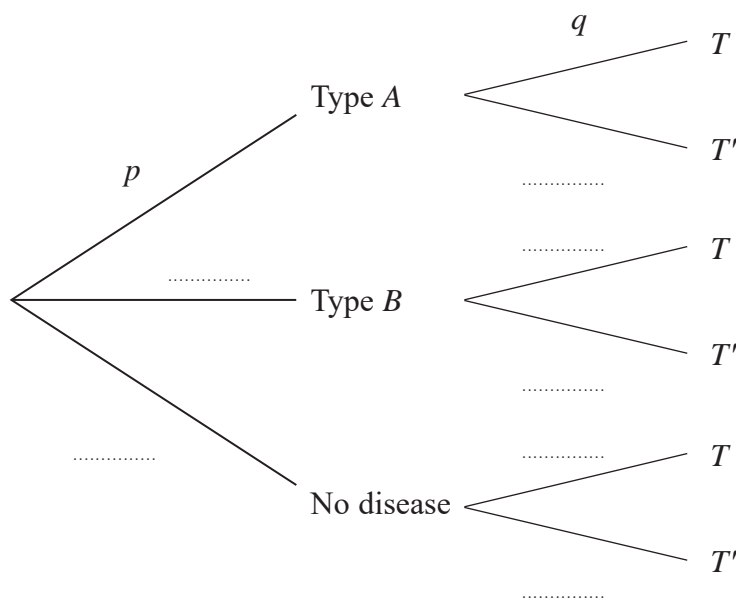
A test has been developed to help diagnose whether or not a person has the disease. The event T represents a positive result on the test. After a large-scale trial of the test, the following information was obtained.

For a person with type B of the disease the probability of a positive test result is 0.96

For a person who does not have the disease the probability of a positive test result is 0.05

For a person with type A of the disease the probability of a positive test result is q

(a) Complete the tree diagram.



(2)

The probability of a randomly selected person having a positive test result is 0.169

For a person with a positive test result, the probability that they do not have the

disease is $\frac{41}{169}$

(b) Find the value of p and the value of q .

(7)

A doctor is about to see a person who she knows does not have type B of the disease but does have a positive test result.

(c) (i) Find the probability that this person has type A of the disease.

(3)

(ii) State, giving a reason, whether or not the doctor will find the test useful.

(1)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA



6. *Ranpose* hospital offers services to a large number of clinics that refer patients to a range of hospitals.

The manager at *Ranpose* hospital took a random sample of 16 clinics and recorded

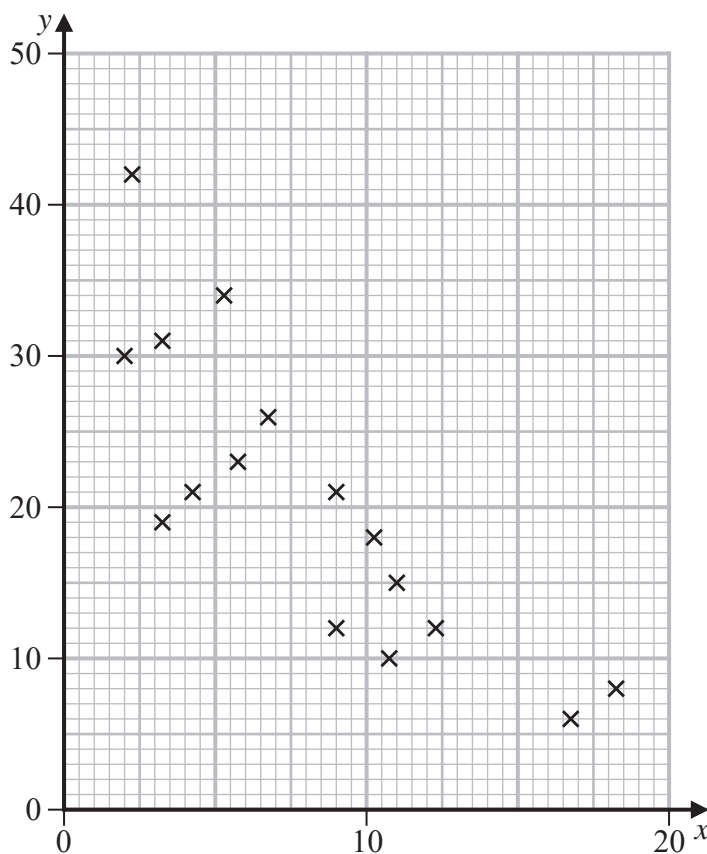
- the distance, x km, of the clinic from *Ranpose* hospital
- the percentage, y %, of the referrals from the clinic who attend *Ranpose* hospital.

The data are summarised as

$$\bar{x} = 8.1 \quad \bar{y} = 20.5 \quad \sum y^2 = 8266 \quad S_{xx} = 368.16 \quad S_{xy} = -630.9$$

- (a) Find the product moment correlation coefficient for these data. (4)
- (b) Give an interpretation of your correlation coefficient. (1)

The manager at *Ranpose* hospital believes that there may be a linear relationship between the distance of a clinic from the hospital and the percentage of the referrals who attend the hospital. She drew the following scatter diagram for these data.



- (c) State, giving a reason, whether or not these data support the manager's belief. (1)

Question 6 continues on page 22

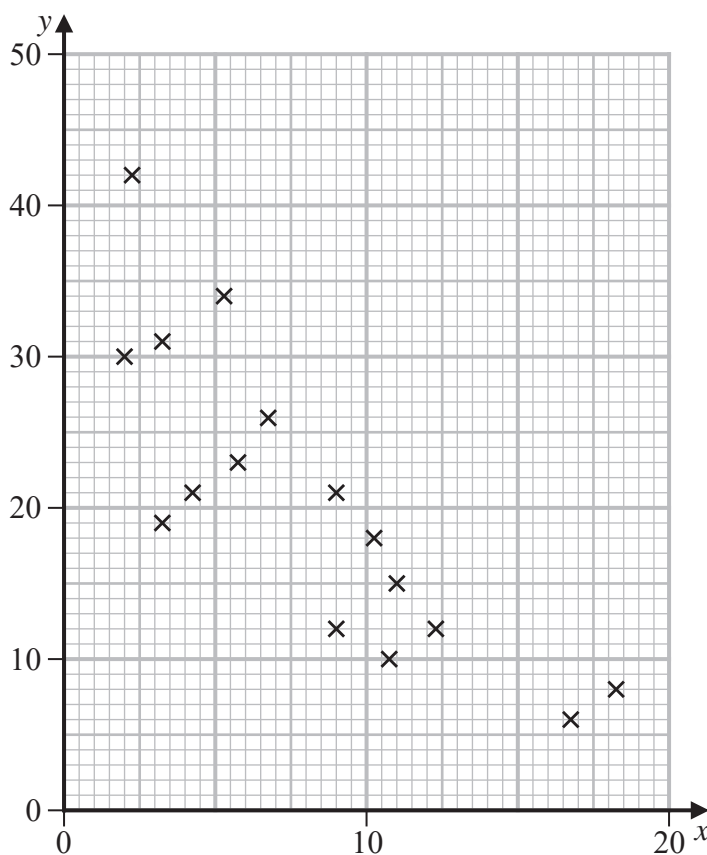


Question 6 continued

[The summary data and the scatter diagram are repeated below.]

The data are summarised as

$$\bar{x} = 8.1 \quad \bar{y} = 20.5 \quad \sum y^2 = 8266 \quad S_{xx} = 368.16 \quad S_{xy} = -630.9$$



- (d) Find the equation of the regression line of y on x , giving your answer in the form $y = a + bx$ (4)

- (e) Give an interpretation of the gradient of your regression line. (1)

- (f) Draw your regression line on the scatter diagram. (1)

The manager believes that *Ranpose* hospital should be attracting an “above average” percentage of referrals from clinics that are less than 5 km from the hospital. She proposes to target one clinic with some extra publicity about the services *Ranpose* offers.

- (g) On the scatter diagram circle the point representing the clinic she should target. (1)



