# 4.2 Correlation & Regression
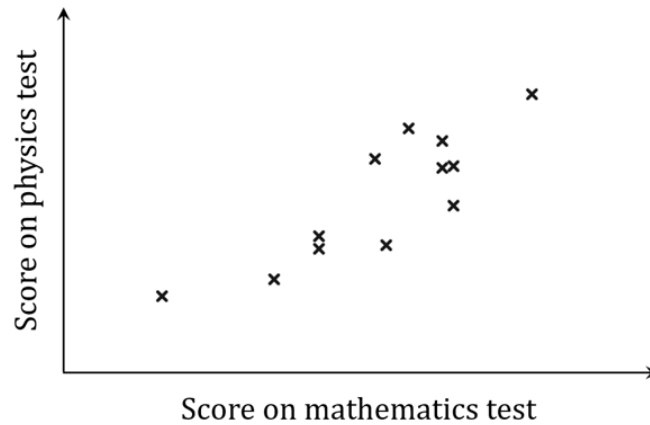
## Question Paper

| Course | DP IB Maths |
|---|---|
| Section | 4. Statistics & Probability |
| Topic | 4.2 Correlation & Regression |
| Difficulty | Medium |

**Time allowed:** 60

**Score:** /44

**Percentage:** /100

## Question 1a

A teacher collected the maths and physics test scores of a number of students and drew a scatter diagram to represent this data.
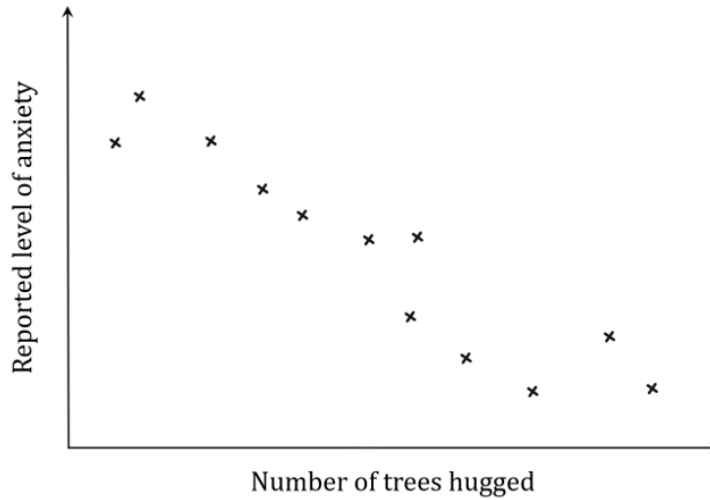


(a) Describe the correlation shown by the scatter diagram, and interpret the correlation in context.

**[2 marks]**

Question 1b

An alternative therapist collected data on his clients' reported levels of anxiety as well as the number of trees they had hugged in the course of therapy. He drew a scatter diagram to represent this data.



(b) Describe the correlation shown by the scatter diagram, and interpret the correlation in context.

[2 marks]

## Question 2a

Jennifer sells cups of tea at her shop and has noticed that she sells more tea on cooler days. On five different days, she records the maximum daily temperature, $T$, measured in degrees Celsius, and the number of cups of teas sold, $C$. The results are shown in the following table.

| Maximum Daily Temperature, *T*. | 3 | 5 | 8 | 9 | 12 |
|---|---|---|---|---|---|
| Cups of tea sold, *C*. | 37 | 34 | 33 | 26 | 21 |

The relationship between $T$ and $C$ can be modelled by the regression line of $C$ on $T$ with equation $C = aT + b$.

(a)  (i)  Find the value of $a$ and the value of $b$.

(ii)  Write down the value of Pearson's product-moment correlation coefficient, $r$.

**[4 marks]**

## Question 2b

(b) Use your regression equation from part (a)(i) to estimate the number of teas that Jennifer will sell on a day when the maximum temperature is 11°C.

**[2 marks]**

## Question 2c

(c) Being sure to consider the result from part (a)(ii) in your answer, state how confident you would be in your estimate from part (b).

**[2 marks]**

## Question 3a

The following table shows the mean height, $y$ cm, of primary school children who are age $x$ years old.

| Age, $x$ years | 6.25 | 7.35 | 8.5 | 9.25 | 10.75 |
|---|---|---|---|---|---|
| Mean Height, $y$ cm | 115 | 121 | 129 | 136 | 140 |

The relationship between $x$ and $y$ can be modelled by the regression line of $y$ on $x$ with equation $y = ax + b$.

(a)  (i)    Find the value of $a$ and the value of $b$.

     (ii)   Write down the value of Pearson's product-moment correlation coefficient, $r$.

**[4 marks]**

## Question 3b

(b) Use your regression equation from part (a)(i) to estimate the height of a child aged 9 years old.

[2 marks]

## Question 3c

(c) Explain why it is not appropriate to use the regression equation to estimate the age of a child who is 133 cm tall.

[1 mark]

## Question 4a

Rebecca, a regular jogger, ran the "Thao Dien Loop" on 7 consecutive days. The following table shows the distance, $x$ km, that she ran and the corresponding number of calories, $y$, that she was able to burn during the run.

| Distance ($x$) | 2 | 5 | 6 | 7 | 10 | 12 | 14 |
|---|---|---|---|---|---|---|---|
| Calories ($y$) | 180 | 315 | 365 | 435 | 619 | 830 | 871 |

The number of calories burnt during a run is dependent upon on the length of the run. The relationship between $x$ and $y$ can be modelled by the regression line of $y$ on $x$ with equation $y = ax + b$.

(a)  (i)   Find the value of $a$ and the value of $b$.

(ii)   Write down the value of Pearson's product-moment correlation coefficient, $r$.

**[4 marks]**

## Question 4b

(b) Interpret, in the context of the question, the value of $a$ found in part (a)(i).

**[1 mark]**

## Question 4c

On the 8th day, Rebecca is only able to run for 8 kilometres.

(c) Use the result from part (a)(i) to estimate the number of calories Rebecca will lose.

**[2 marks]**

## Question 4d

(d) Comment on the validity of using the result from part (a)(i) answer part (c).

**[1 mark]**

## Question 5a

The percentage of people who are willing to get a particular vaccine is dependent on their age. The following table shows the age, $A$ years old, and the corresponding percentage of people, $V$, that are willing to receive a vaccine for 6 different ages.

| Age, ($A$) | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|
| Percentage of willing people, ($V$) | 57 | 59 | 61 | 62 | 68 | 75 |

The relationship between $A$ and $V$ can be modelled by the regression line of $V$ on $A$ with equation $V = aA + b$.

(a)  (i)   Find the value of $a$ and the value of $b$.

   (ii)   Write down the value of Pearson's product-moment correlation coefficient, $r$.

[4 marks]

## Question 5b

(b) Interpret, in the context of the question, the value of $a$ found in part (a)(i).

[1 mark]

## Question 5c

(c) Use the result from part (a)(i) to estimate the percentage of people aged 95 years old who are in willing to receive a vaccine.

[2 marks]

## Question 5d

(d) Comment on the validity of using the result from part (a)(i) to answer part (c).

[1 mark]

## Question 6a

The price, $P$, of an airline ticket is dependent on the distance, $D$ km, between two cities. The table below shows the airfares in US dollars from Prague in the Czech Republic, to eight different destinations in Europe.

| Distance ($D$) | 885 | 340 | 835 | 330 | 1270 | 295 | 650 | 1930 |
|---|---|---|---|---|---|---|---|---|
| Price ($P$) | 99 | 50 | 90 | 45 | 119 | 42.5 | 59 | 139 |

Let $L_1$ be the regression line of $P$ on $D$. The equation of the line $L_1$ can be written in the form $P = aD + b$.

Let $L_2$ be the regression line of $D$ on $P$. The equation of the line $L_2$ can be written in the form $D = cP + d$.

(a)  (i)   Find the value of $a$ and the value of $b$.

(ii)   Find the value of $c$ and the value of $d$.

[3 marks]

## Question 6b

(b) Write down the value of Pearson's product-moment correlation coefficient, $r$.

[2 marks]

## Question 6c

(c) Use the result from part (a)(i) to estimate the price of an airline ticket for a flight from Prague to a destination that is 2635 km away.

[2 marks]

## Question 6d

The lines $L_1$ and $L_2$ both pass through the same point with coordinates $(p, q)$.

(d) Find the value of $p$ and the value of $q$.

[2 marks]