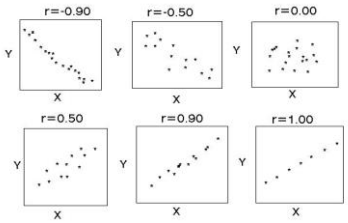
Weeks 1 and 2
Term 3
2021

HAWKER COLLEGE
Engage | Inspire | Achieve

Learning Brief SMM4: Bivariate Data – Correlation Coefficient

Goals



By the end of this fortnight you should be able to:
• Display, summarise, and describe relationships in bivariate data
• Identify dependent (response) and independent (explanatory) variable
• Calculate and use r and r^2 to interpret the strength of bivariate relationship between two variables. Comment on the reliability of the predicted values using r^2
• Understand the underlying theory behind least-squares to fit straight line to bivariate data
• Calculate the equation of a least-squares regression line by hand and by using a CAS calculator
• Use the equation of the line to 'predict' data values for a given value of x (Interpolate and Extrapolate)

Theoretical Components

Chapter 2 and 3 of Quest Further Maths 12 (Jacplus e-text – pdf on Google Classrooms):

•Read through Section 2E on Scatter Plots. Study examples 8 & 9.
•Read through Section 2F on Correlation Coefficient. Study example 10.
•Read through Section 2G on Calculating r and r^2. Study examples 11 & 12.

Make your notes on the following key concepts:
• Dependent/Independent Variables
• Back to Back Stem Plots
• Parallel Box plots
• Scatter Plots
• r and r^2

Minimising Least Square error:
https://youtu.be/6OvhLPS7rj4
(Probably no need to worry about the proof)

•Read through Section 3C on Fitting Straight Line – Least Squares Method. Study and make notes on Examples 3 & 4.
•Read through Section on the use of CAS to approximate the regression equation (Section 3C).
•Read through Section 3D on Interpolation and Extrapolation. Study examples 5 - 7.

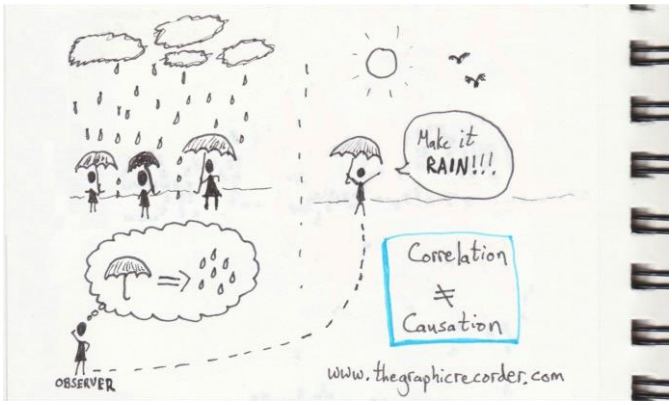
Practical Components

Do the following questions. Organise your solutions neatly in your exercise book:
EX 2E: 2, 3, 4
EX 2F: 1 - 4
EX 2G: ALL.
EX 3C: ALL.
EX 3D: Many
Use Class Pad and show working.

Investigation&Quiz

Quiz: mathspace.co task:

- Bivariate Data



Investigation:
Complete the investigation on the following page

INVESTIGATION

Copy and complete the table to estimate the parameters of linear regression (i.e. constant and the coefficient in y=ax + b). Then use the totals and the formulae given to work out **'a', 'b', r and r^2**. Check your answers on CAS. The first two columns give the values for age (x, in years) and systolic blood pressure (y, in mmHg) for 15 women.

X	Y	X^2	Y^2	XY
42	130	1764	16900	5460
46	115			
42	148			
71	100			
80	156			
74	162			
70	151			
80	156			
85	162			
72	158			
64	155			
81	160			
41	125			
61	150			
75	165			
Total 984	Total 2193			

Formulae:

$$a = \frac{n\sum XY - \sum X \sum Y}{n\sum X^2 - (\sum X)^2}$$

$$b = \bar{Y} - a\bar{X}$$

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{n}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{n}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{n}\right]}}$$

Use the formula you have found to predict the systolic blood pressure for a 59 year old woman.

The coefficient of determination (r^2) provides a measure of how well the linear rule linking the two variables (x and y) predicts the value of y when we are given the value of x. Comment on the r^2 found in the about example and the predictability of the linear model found.