Summary of key points

1 If $y = ax^n$ for constants a and n then $\log y = \log a + n \log x$

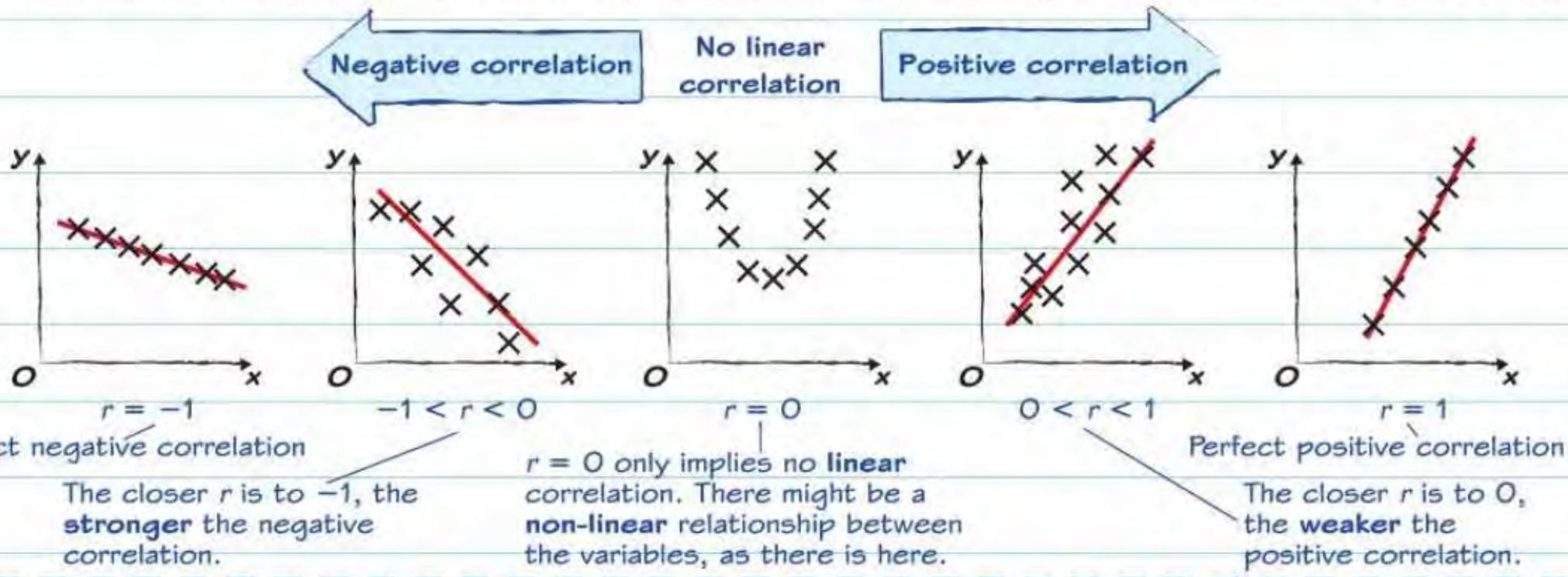
2 If $y = kb^x$ for constants k and b then $\log y = \log k + x \log b$

3 The **product moment correlation coefficient** describes the linear correlation between two variables. It can take values between -1 and 1.

Had a look ☐ Nearly there ☐ Nailed it! ☐

Measuring correlation

You can use the **product moment correlation coefficient (PMCC)** to measure the strength of the linear correlation between two variables. You need to be able to use your calculator to find the PMCC.



Worked example

An engineer is using a wind tunnel to investigate the effects of air speed on drag over an aircraft wing. She records her results in a table.

Air speed, a (mph)	20	40	80	120	160
Drag, d (N)	80	260	1670	5400	12700

The engineer believes the data can be modelled by a relationship of the form $d = Pa^k$ for constants P and k. She codes the data using $x = \log a$ and $y = \log d$.

(a) Calculate the product moment correlation coefficient of the coded data. **(3 marks)**

x	1.301	1.602	1.903	2.079	2.204
y	1.903	2.415	3.223	3.732	4.104

$r = 0.994$

The equation of the regression line of y on x is found to be $y = 2.470x - 1.415$

(b) Estimate the values of P and k in the engineer's model, and comment on the validity of this model. **(4 marks)**

$\log d = 2.470 \log a - 1.415$

$\log d = \log a^{2.470} - \log 10^{1.415}$

$\log d = \log \dfrac{a^{2.470}}{10^{1.145}}$

$d = 10^{-1.415} a^{2.470}$

So $P = 10^{-1.415} = 0.0385$ and $k = 2.470$

The PMCC is close to 1 ($r = 0.994$), which suggests this model is valid.

Non-linear models

You can use **logarithms** to test for non-linear relationships between two variables x and y. You need to know about two different forms:

1. If $y = ax^n$ for constants a and n then $\log y = \log a + n \log x$

You need to use the coding $Y = \log y$ and $X = \log x$ to obtain a linear relationship.

2. If $y = kb^x$ for constants k and b then $\log y = \log k + x \log b$

You need to use the coding $Y = \log y$ and $X = x$ to obtain a linear relationship.

You can revise the techniques needed to determine these relationships on page 53.

Use your calculator to find the product moment correlation coefficient. You might have to select the type of model you are using. Because the PMCC measures **linear** correlation, you should choose a model of the form $y = a + bx$. The letter r is usually used to denote the PMCC.

Now try this

The daily maximum relative humidity (%) and the daily mean visibility (km) recorded in Leeming for the first 8 days in May 2015 were recorded in the large data set.

Humidity (%)	99	94	95	94	99	87	88
Visibility (km)	25	16	14	20	15	24	32

Calculate the product moment correlation coefficient of the data, and describe the nature of the linear relationship (if any).

Answer **(2 marks)**