Exponential Models

In AS Statistics, you saw how regression lines can be used to model a linear relationship between two variables. Sometimes experimental data does not fit a linear model but still shows a clear pattern. We can use logarithms and coding to examine trends with non-linear data.

If data can be modelled with a relationship $\boldsymbol{y = ax^n}$, for constants a and n, then taking logs of both sides gives:

$$\boldsymbol{\log y = \log a + n \log x}$$

If you then plot a graph of $\log y$ against $\log x$, you get a straight-line graph with gradient n and intercept $\log a$.

We can describe this as **coding the data** using $Y = \log y$ and $X = \log x$, and then plotting Y against X.

If data is modelled with an **exponential** relationship $\boldsymbol{y = kb^x}$, for constants k and b, then logs of both sides gives:

$$\boldsymbol{\log y = \log k + x \log b}$$

If you then plot a graph of $\log y$ against x, you get a straight-line graph with gradient $\log b$ and intercept k.

We can describe this as **coding the data** using $Y = \log y$ and $X = x$ and then plotting Y against X.

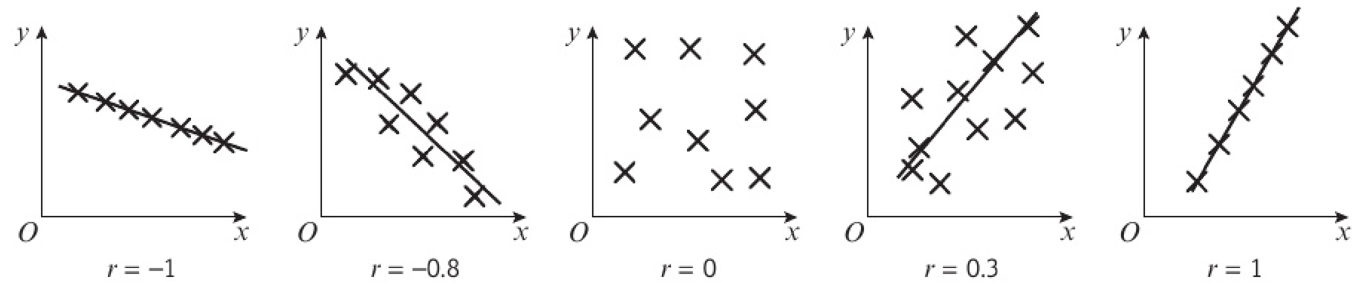
Product Moment Correlation Coefficient

Previously you have only had to identify whether a linear correlation between two variables is **positive** or **negative** by inspecting the data (or, more often, a scatter graph displaying the data).

Now, we will see how you can quantify the strength and type of linear correlation between two variables by calculating the **product moment correlation coefficient**.

The product moment correlation coefficient describes the linear correlation between two variables, by giving a value between -1 and 1, and is usually denoted with the letter r

- If $r = 1$, there is a perfect positive linear correlation (the points lie on a straight line with a positive gradient
- If $r = -1$, there is a perfect negative linear correlation (the points lie on a straight line with a negative gradient
- The closer the value of r is to 1 or –1, the stronger the positive or negative correlation, respectively
- If $r = 0$ (or is close to 0), there is no linear correlation, although there may still be some other relationship



You need to be able to calculate the PMCC for bivariate data using your calculator.

Hypothesis Testing

You can use a hypothesis test to determine whether the product moment correlation coefficient, r, for a sample of bivariate data indicates that there is likely to be a linear relationship within the whole population.

As the PMCC for a **sample** is denoted r, so the PMCC for a **whole population** is denoted ρ, the Greek letter *rho*.

To test for positive or negative linear correlation – whether or not the population PMCC ρ is greater than or less than zero – we can use a **one-tailed test**:

$$\boldsymbol{H_0:\ \rho = 0 \qquad H_1:\ \rho > 0 \ \text{ or } \ \rho < 0}$$

To test for any correlation – whether or not the population PMCC ρ is not equal to zero, we use a **two-tailed test**:

$$\boldsymbol{H_0:\ \rho = 0 \qquad H_1:\ \rho \neq 0}$$

The **critical values** you need to compare with your sample PMCC r are given in a table of values in the formula book.

You have to look up the critical value for your **sample size** and **significance level**. The value given is an **absolute** value, so when testing for negative correlation you should take the negative of the printed value.

As the null hypothesis is always no correlation, the value of r is more significant the further it is from zero:

For $H_1: \rho > 0$, if r is closer to 1 than the critical value, then the result is significant and you should reject H_0

For $H_1: \rho < 0$, if r is closer to -1 than the negative critical value, the result is significant and you should reject H_0

For $H_1: \rho \neq 0$, the result is significant if either of the above cases is true. **Don't forget to split the significance level!**