Your notes

GCSE Statistics: revision notes

Scatter diagrams and correlation

Contents

1. Scatter diagrams
2. Correlation isn't causation
3. Spearman's rank correlation coefficient
4. Pearson's product moment correlation coefficient
5. Key terms
6. Quick questions
7. What to do next

Scatter diagrams

- A **scatter diagram** plots bivariate data, with the explanatory variable on the horizontal axis.
- **Positive correlation**: as one variable increases, the other increases. **Negative correlation**: as one increases, the other decreases. No correlation: no pattern.
- Correlation can be strong (points close to a line) or weak (points scattered).

Correlation isn't causation

- Two variables can be correlated without one causing the other. Ice cream sales and drownings both rise in summer: a third (lurking) variable, temperature, affects both.
- Only a well-designed experiment can show that one variable causes a change in another.

Spearman's rank correlation coefficient

- Rank each set of data, find the difference in ranks (d) for each item, then use $r_s = 1 - 6\Sigma d^2 \div (n(n^2 - 1))$.
- r_s is between –1 and +1. +1 means the rankings agree perfectly, –1 means they are exactly opposite, and 0 means no agreement.
- Example: 5 items with $\Sigma d^2 = 4$. $6 \times 4 = 24$ and $5 \times 24 = 120$, so $r_s = 1 - 24 \div 120 = 0.8$.

Pearson's product moment correlation coefficient

- The **PMCC** (r) measures the strength of linear correlation, also from –1 to +1. You interpret it rather than calculate it by hand: 0.9 is strong positive, –0.3 is weak negative.

© LycoLearn · lycolearn.co.uk · Free GCSE and KS3 revision

- Spearman's is used for ranked data or when the relationship isn't a straight line; the **PMCC** is used for linear relationships.

Key terms

Scatter diagram	A graph plotting pairs of values for bivariate data.
Positive correlation	As one variable increases, the other increases.
Negative correlation	As one variable increases, the other decreases.
Causation	When a change in one variable directly causes a change in another.
Spearman's rank correlation coefficient	A measure of agreement between two sets of rankings.
PMCC	A measure of the strength of linear correlation.
Lurking variable	A third variable that affects both variables being studied.

Quick questions

Cover the answers and test yourself: remembering is what makes it stick.

1. What does positive correlation mean?
2. What does negative correlation mean?
3. Ice cream sales and drownings are correlated. Why doesn't one cause the other?
4. What is the range of possible values for Spearman's coefficient?
5. $n = 5$ and $\Sigma d^2 = 4$. What is Spearman's rank correlation coefficient?
6. A PMCC of -0.92 shows what?
7. A PMCC of 0.1 shows what?
8. Which variable goes on the horizontal axis of a scatter diagram?

Answers

1. As one variable increases, the other increases
2. As one variable increases, the other decreases
3. Temperature affects both
4. -1 to $+1$
5. 0.8
6. Strong negative correlation
7. Very weak or no correlation
8. The explanatory variable

What to do next

Practise 10 questions on this topic

Try 3 exam questions, with mark schemes

Read the lesson, one step at a time

Free revision for GCSE and KS3 at lycolearn.co.uk. Some notes are for one exam board only, and say so in brackets.

Your notes