Spearman’s Rank Correlation Test

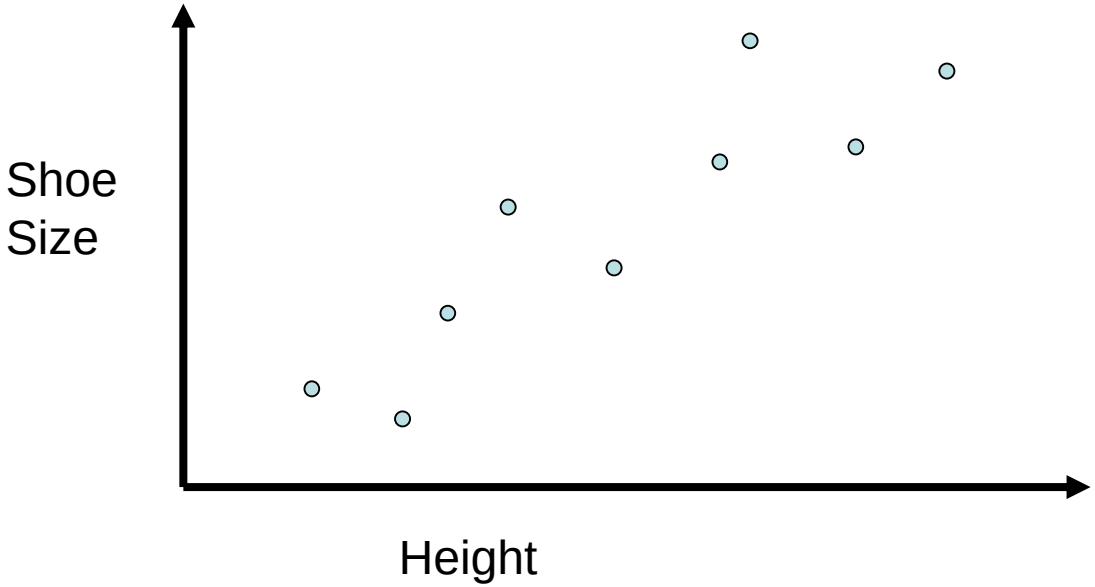
A statistical method of testing the correlation between two variables.

What is a correlation?

- A relationship between one set of data and another.
  - Eg. An environmental factor such as light intensity and the population of a certain organism.

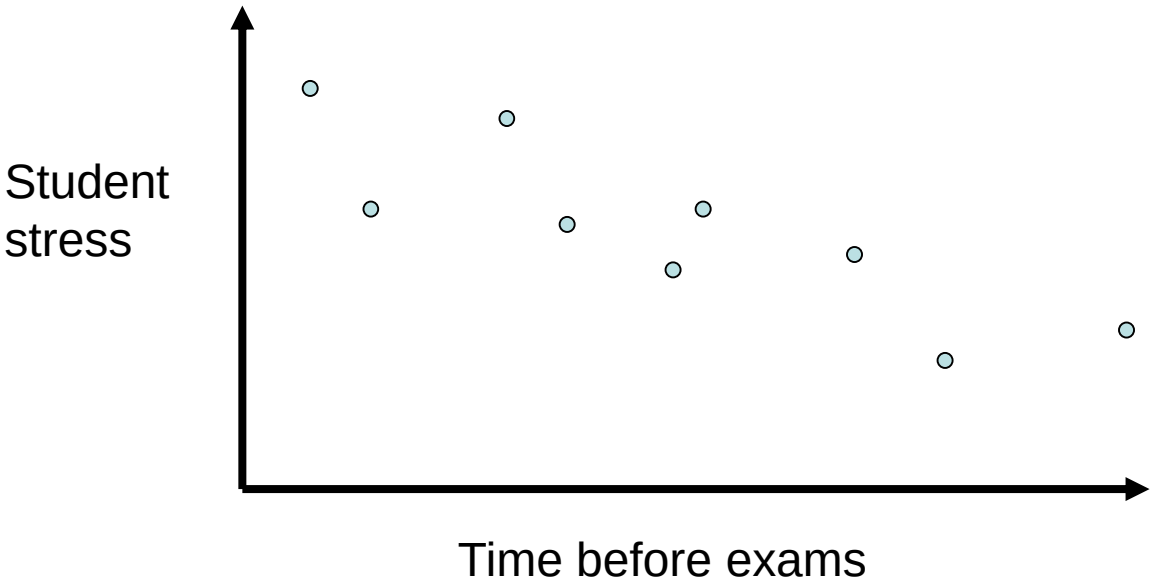
What types of correlation are there?

- Positive correlation
  - Person's height and shoe size.
  - Amount of rain and sale of umbrellas.



What types of correlation are there?

- Negative correlation
  - Winter temperatures and electricity bills.
  - Time left before an exam and student stress.

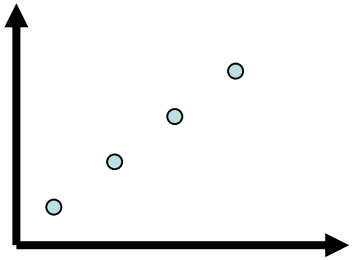


What types of correlation are there?

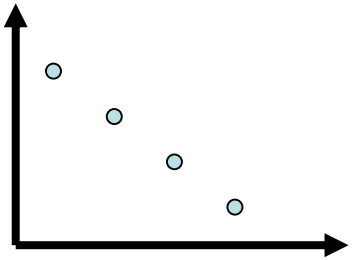
- No correlation
  - Amount of spinach eaten and lottery wins.
  - Science GCSE results and hair length.

Correlation Coefficients

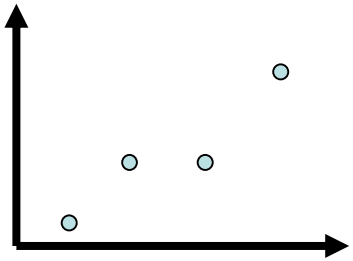
Range from -1 to +1



Coefficient = 1
perfect positive correlation

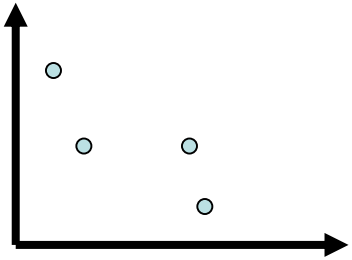


Coefficient = -1
perfect negative correlation



Coefficient = 0.5
positive correlation

some



Coefficient = -0.5
some negative correlation

Making a Prediction

- With any scientific test you should make a prediction.
- Statistical tests use Null Hypothesis and Alternative Hypothesis.

- Null Hypothesis – there is no correlation.

- Alternative Hypothesis – there is a correlation.

Calculating Correlation Coefficients

Counting the number of aphids and ladybirds at 10 different m^2 sites

Site	No of aphids	No of Ladybirds
A	13	2
B	16	5
C	13	1
D	5	1
E	49	5
F	23	5
G	33	4
H	8	2
I	40	7
J	7	0

Calculating Correlation Coefficients

Rank each variable, taking care with tied scores.

Site	No of aphids	No of Ladybirds	R_1	R_2
A	13	2	4.5	4.5
B	16	5	6	8
C	13	1	4.5	2.5
D	5	1	1	2.5
E	49	5	10	8
F	23	5	7	8
G	33	4	8	6
H	8	2	3	4.5
I	40	7	9	10
J	7	0	2	1

Calculating Correlation Coefficients

Calculate the difference between the ranks.

Site	No of aphids	No of Ladybirds	R_1	R_2	Difference (R_1-R_2)
A	13	2	4.5	4.5	0
B	16	5	6	8	-2
C	13	1	4.5	2.5	2
D	5	1	1	2.5	-1.5
E	49	5	10	8	2
F	23	5	7	8	-1
G	33	4	8	6	2
H	8	2	3	4.5	-1.5
I	40	7	9	10	-1
J	7	0	2	1	1

Calculating Correlation Coefficients

Square the difference between the ranks and sum the squares.

Site	No of aphids	No of Ladybirds	R_1	R_2	D	D^2
A	13	2	4.5	4.5	0	0
B	16	5	6	8	-2	4
C	13	1	4.5	2.5	2	4
D	5	1	1	2.5	-1.5	2.25
E	49	5	10	8	2	4
F	23	5	7	8	-1	1
G	33	4	8	6	2	4
H	8	2	3	4.5	-1.5	2.25
I	40	7	9	10	-1	1
J	7	0	2	1	1	1
					Sum	23.5

Calculating Correlation Coefficients

The final step involves a formula:

$$r_s = 1 - \frac{6\sum D^2}{n(n^2-1)}$$

r_s = Spearman Correlation Coefficient.

$\sum D^2$ = Sum of all squared differences.

n = Number of pairs of data analysed.

$$r_s = 1 - \frac{6*23.5}{10*99} = 1 - \frac{141}{990} = 0.86$$

Positive r_s = positive correlation

Negative r_s = negative correlation

Significance

- Now we need to know how significant this result is.
  - Could this have happened purely by chance or is the correlation real?

Significance

- Look up the r_s value in published statistical tables.
- Compare the result to the listed critical value when n=10.
- If the result (ignoring the sign) is greater than the critical value then reject the Null Hypothesis and accept the Alternative Hypothesis.
- If the result is less than the critical value then accept the Null Hypothesis.

Significance

- Critical values commonly list 95% significance.

- So, if your result is over the critical value we can say with 95% certainty that the correlation is real.
  - Or that there is a less than 5% probability that the data have been produced by chance alone.