

The Pearson product-moment correlation coefficient

A worked example

A downstream survey in a river appeared to show that there was a negative correlation between water velocity and cross-sectional area. Measurements were taken at 15 points between the source and the mouth of the river.

Null hypothesis: There is no correlation between water velocity and cross-sectional area.

Alternative hypothesis: There is a negative correlation between water velocity and cross-sectional area.

Observation no	A: x values velocity ms ⁻¹	B x- \bar{x}	C (x- \bar{x}) ²	D: y values Cross-sectional area m ²	E y- \bar{y}	F (y- \bar{y}) ²	G (x- \bar{x}) ² (y- \bar{y}) ²
1	.335	0.164	0.027	.075	-0.575	0.330	-0.094
2	.368	0.197	0.039	.095	-0.555	0.308	-0.109
3	.265	0.094	0.009	.112	-0.538	0.289	-0.050
4	.246	0.075	0.006	.150	-0.500	0.250	-0.037
5	.201	0.030	0.001	.165	-0.485	0.235	-0.014
6	.214	0.043	0.002	.197	-0.453	0.205	-0.019
7	.196	0.025	0.001	.320	-0.330	0.109	-0.008
8	.168	-0.003	0.000	.368	-0.282	0.079	0.001
9	.125	-0.046	0.002	.485	-0.165	0.027	0.008
10	.110	-0.061	0.004	.658	0.008	0.000	-0.001
11	.099	-0.072	0.005	.890	0.240	0.058	-0.017
12	.085	-0.086	0.007	1.025	0.375	0.141	-0.032
13	.065	-0.106	0.011	1.350	0.700	0.490	-0.074
14	.045	-0.126	0.016	1.875	1.225	1.501	-0.155
15	.046	-0.125	0.016	1.982	1.332	1.775	-0.167
Sums of columns	2.568	0	0.145	9.747	0	5.797	-0.771
No. of observations	15						
Means of x and y	0.171			0.649			
	A	B	C	D	E	F	G

<i>For the top row of the formula</i>	What it is	Actual value	Row no.
$\Sigma(x - \bar{x})(y - \bar{y})$	(sum of figures in column G)	-0.771	1
<i>For the bottom row of the formula</i>	What it is	Actual value	2
$\Sigma(x - \bar{x})^2$	(sum of figures in column C)	0.145	3
$\Sigma(y - \bar{y})^2$	(sum of figures in column F)	5.797	4
$\Sigma(x - \bar{x})^2 \Sigma(y - \bar{y})^2$	(multiply the sum of figures in column C by the sum of figures in column F)	0.841	5
$\sqrt{\Sigma(x - \bar{x})^2 \Sigma(y - \bar{y})^2}$	(Calculate the square root of the value above)	0.916	6
$r_{xy} =$	(divide the value in row 1 by the value in row 6): the answer should lie between +1 and -1	-0.841	

The Pearson product-moment correlation coefficient

A worked example

Significance of the result

Now we need to compare your result with a table of significant values of r at the 5% probability level

N = number of pairs of data	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
degrees of freedom = N-2			1	2	3	4	5	6	7	8	9	10	11	12	13
Critical value of r			.997	.950	.878	.811	.754	.707	.666	.632	.602	.576	.553	.532	.514

Calculated value of $r_{xy} = -0.841$

Critical value of r from table above = 0.514

Since calculated value is **greater than** the critical value (ignoring the sign) then the correlation is significant at the 5% probability level, so we can reject the null hypothesis and accept the alternative hypothesis that there is a negative correlation between water velocity and cross-sectional area.