COEFICIENTE DE CORRELAÇÃO LINEAR (r)

É o estudo do relacionamento entre duas ou mais variáveis.

A correlação

Simples – Se o estudo tratar apenas de duas variáveis

Múltpiplas – Se envolver mais do que duas variáveis

OBJETIVO

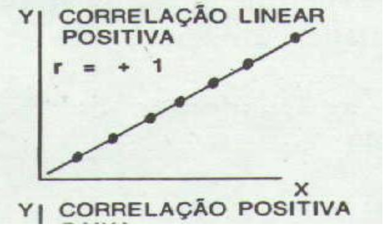
Determinar quanto uma equação linear, ou de outra espécie, descreve ou explica a relação entre as variáveis e fornece um número que resume o grau de relacionamento linear entre as duas variáveis.

VALORES DO r

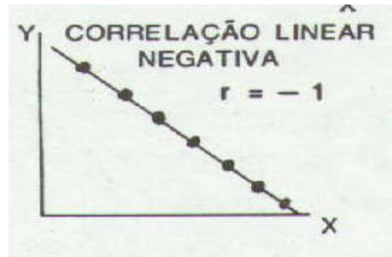
O coeficiente de correlação (r) fica situado no intervalo entre – 1 e 1, ou seja:

$$-1 \geq r \leq 1$$

Um coeficiente de +1, indica uma correlação linear máxima positiva perfeita. Neste caso, as duas variáveis serão exatamente iguais.



Um coeficiente de correlação de –1, indica correlação linear máxima perfeita negativa, com os resultados padronizados exatamente iguais em valores absolutos, diferindo apenas no sinal.



Uma correlação de +1 ou –1 é raramente observado.

Um coeficiente de correlação “0”, significa que não existe um relacionamento linear entre as duas variáveis, isto é, não existe correlação.

<u>DETERMINAÇÃO DO COEFICIENTE DE CORRELAÇÃO (r)</u>

$$r = \frac{\sum XY - \frac{\left(\sum X\right)\left(\sum Y\right)}{N}}{\sqrt{\left[\sum X^2 - \frac{\left(\sum X\right)^2}{N}\right]\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{N}\right]}}$$

TESTE PARA PROVAR SE O *r* É SIGNIFICATIVO (DIFERENTE DE ZERO)

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}}$$

onde:
t = valor calculado do teste t com N – 2 graus de liberdade;
r = estimativa do coeficiente de correlação;
N = número de observações.

Para saber a significância compara o valor do t calculado com um valor do t tabelado a 5% de significância.

Frederico Antonio Loureiro Soares
Econometria

Graus de Liberdade	10%	5%	2%	1%	0,5%	0,1%
1	6,31	12,71	31,82	63,66	127,32	636,62
2	2,92	4,30	6,97	9,92	14,09	31,60
3	2,35	3,18	4,54	5,84	7,45	12,94
4	2,13	2,78	3,75	4,60	5,60	8,61
5	2,02	2,57	3,37	4,03	4,77	6,86
6	1,94	2,45	3,14	3,71	4,32	5,96
7	1,90	2,36	3,10	3,50	4,03	5,41
8	1,86	2,31	2,90	3,36	3,83	5,04
9	1,83	2,26	2,82	3,25	3,69	4,78
10	1,81	2,23	2,76	3,17	3,58	4,59
11	1,80	2,20	2,72	3,11	3,50	4,44
12	1,78	2,18	2,68	3,06	3,43	4,32
13	1,77	2,16	2,65	3,01	3,37	4,22
14	1,76	2,14	2,62	2,98	3,33	4,14
15	1,75	2,13	2,60	2,95	3,29	4,07
16	1,75	2,12	2,58	2,92	3,25	4,02

Se $|t| < t_t$, conclui-se que não existe correlação, ou seja, que a correlação é "0" e não existe um relacionamento linear entre as duas variáveis,;

Se $|t| \geq t_t$, conclui-se que existe correlação, ou seja, indica uma correlação linear máxima positiva ou negativa perfeita e neste caso, as duas variáveis serão exatamente iguais

- **Exemplos**

1) Considere os dados abaixo para testar se existe ou não correlação linear entre a quantidade de feijão produzido (X) e o trato cultural expresso pelo número de capinas (Y)

Tabela 2. Valores das variáveis X (tratos culturais, em números de capinas durante o ciclo) e Y (quantidade de feijão produzido, em kg ha^{-1})

X	Y
2	48
4	56
5	64
6	60
8	72

Frederico Antonio Loureiro Soares
Econometria

X	Y	XY	X^2	Y^2
2	48	96	4	2304
4	56	224	16	3136
5	64	320	25	4096
6	60	360	36	3600
8	72	576	64	5184
25	300	1576	145	18320

$$r = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sqrt{\left[\sum X^2 - \frac{(\sum X)^2}{N}\right]\left[\sum Y^2 - \frac{(\sum Y)^2}{N}\right]}} = \frac{1576 - \frac{(25)(300)}{5}}{\sqrt{\left[145 - \frac{25)^2}{5}\right]\left[18320 - \frac{(300)^2}{5}\right]}} =$$

$$\frac{1576 - 1500}{\sqrt{[145-125][18320-18000]}} = \frac{76}{\sqrt{20x320}} = \frac{76}{80} = 0{,}95$$

Teste "t":

$$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}} = \frac{0{,}95\sqrt{5-2}}{\sqrt{1-(0{,}95)^2}} = \frac{0{,}95\sqrt{3}}{\sqrt{1-0{,}9025}} =$$

$$\frac{0{,}95x1{,}732}{\sqrt{0{,}0975}} = \frac{1{,}6454}{0{,}3122} \cong 5{,}27$$

O valor tabelado de t com 3 GL. e a 5% de significância, considerando um teste bilateral é: 3,182.

Com estes valores pode-se afirmar, com 5% de significância, que as duas variáveis possuem um relacionamento linear.