

Parte II – Inferência estatística

4. Correlação

Ana Maria Lopez Calvo de Feijoo

SciELO Books / SciELO Livros / SciELO Libros

FEIJOO, AMLC. Correlação. In: *A pesquisa e a estatística na psicologia e na educação* [online]. Rio de Janeiro: Centro Edelstein de Pesquisas Sociais, 2010, pp. 70-77. ISBN: 978-85-7982-048-9. Available from SciELO Books <<http://books.scielo.org>>.



All the contents of this chapter, except where otherwise noted, is licensed under a Creative Commons Attribution-Non Commercial-ShareAlike 3.0 Unported.

Todo o conteúdo deste capítulo, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição - Uso Não Comercial - Partilha nos Mesmos Termos 3.0 Não adaptada.

Todo el contenido de este capítulo, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-NoComercial-CompartirIgual 3.0 Unported.

4. CORRELAÇÃO

Em estatística descritiva, foram estudados, especialmente, os métodos de cálculo das medidas estatísticas destinadas a representar de maneira fidedigna a execução de um indivíduo ou de um grupo no tocante a determinada característica.

Frequentemente, porém, é de importância maior examinar a relação entre duas variáveis do que mediar a execução de cada uma por si. Nesse aspecto pode-se falar em correlação, que estuda a relação concomitante entre duas ou mais variáveis. Quanto mais aumenta a concomitância, maior correlação existe entre as variáveis.

O coeficiente de correlação é um índice que expressa o grau de associação entre duas ou mais variáveis. Para se atingir este coeficiente, é preciso medir as duas variáveis. O primeiro passo é conseguir, na mesma amostra, duas medidas dos respectivos indivíduos: uma medida de X e outra de Y, as quais serão emparelhadas.

Considerações gerais

O coeficiente de correlação é expresso por um índice numérico que varia de -1 a +1. Onde:

- 1: Correlação negativa, inversa e perfeita.

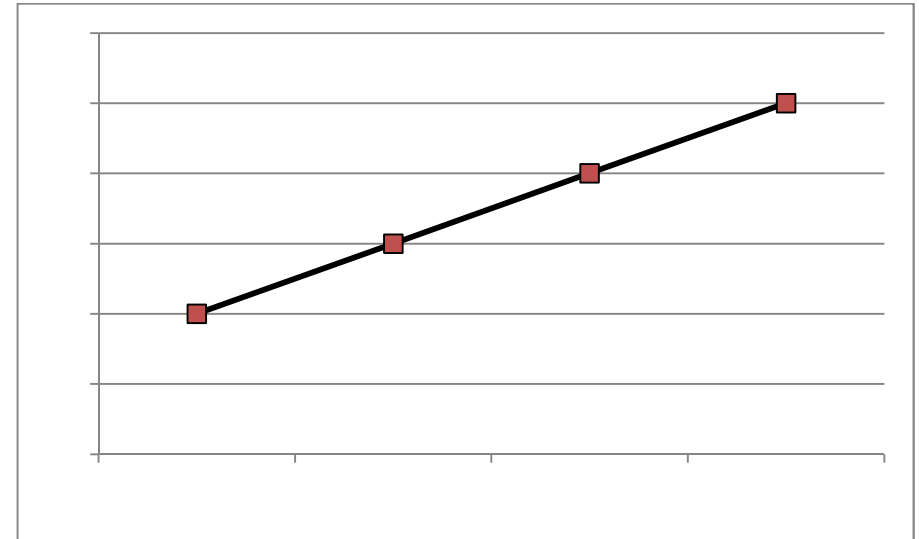
+ 1: Correlação positiva, direta e perfeita.

Aplicações do coeficiente de correlação

Utiliza-se o cálculo da correlação quando se deseja recorrer à prova de hipóteses e na construção de testes psicológicos na análise de itens, validade e fidedignidade, e fazer previsões em termos de seleção de pessoal.

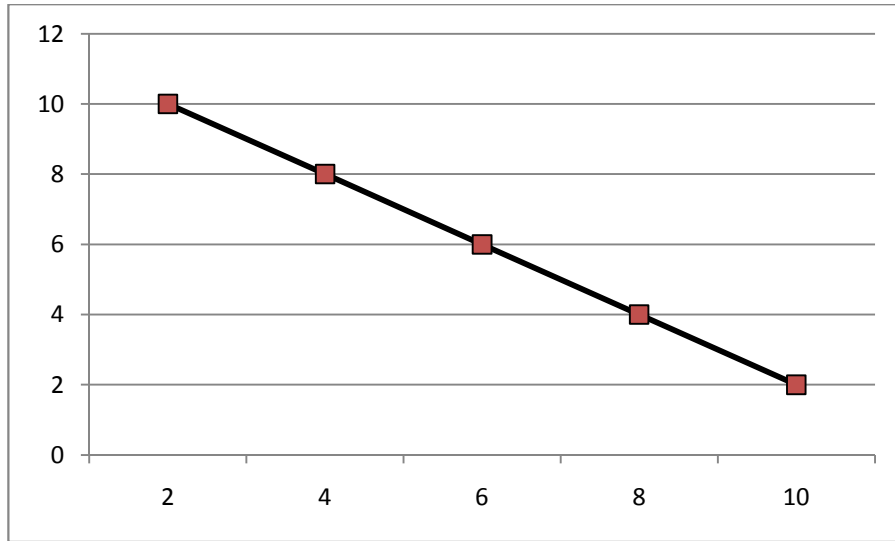
Representação gráfica

O coeficiente de correlação é expresso graficamente da seguinte forma:



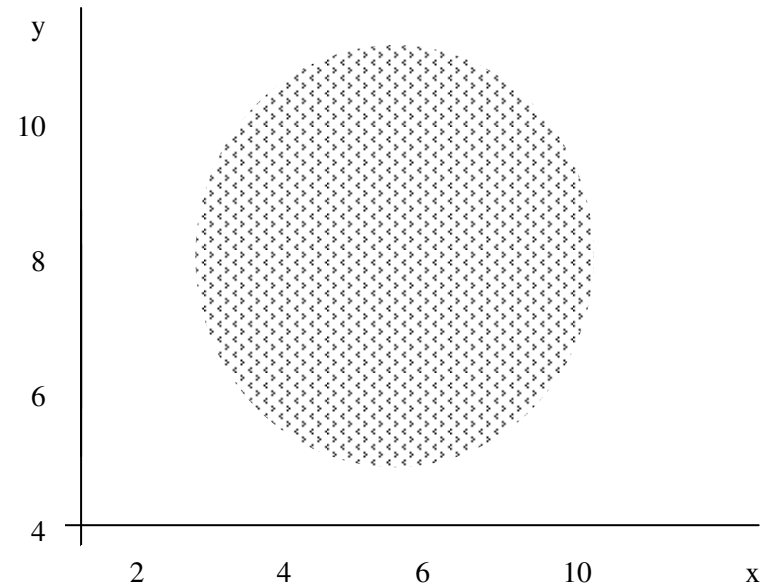
Positiva, direta e perfeita $r = +1$

Ind.	x	y
A	2	4
B	4	6
C	6	8
D	10	10



Negativa, inversa e perfeita $r = -1$

Ind.	x	y
A	2	10
B	4	8
C	6	6
D	8	4
E	10	2



Ausência de correlação $r = 0$

Derivações do coeficiente de correlação

1 – *Predição*: sabendo-se a relação entre duas variáveis, pode-se prever uma a partir da outra.

Na predição faz-se necessária a equação de regressão.

Ou então:

Coefficiente de determinação

É a variância independente:

$$K^2 = 1,00 - r^2$$

Coeficiente de alienação

Importante para verificar o erro de prognóstico:

$$K = \sqrt{1 - r^2}$$

O erro de prognóstico é maior quanto maior é o valor de K.

K = medida de ausência de relação entre duas variáveis X e Y.

Quando K = 1,00, r = zero.

Quanto mais exatos forem os prognósticos que podemos fazer, tanto melhor será o valor preditivo.

Se a correlação entre um teste e um critério é zero, o valor de nossos prognósticos será nulo. Quando a correlação entre o teste e o critério é perfeita não se comete nenhum erro de predição.

Quando a correlação entre o teste e o critério diminui de 1 a 0, o coeficiente de alienação aumenta rapidamente.

2 – Pesquisa: controle no processo de investigação – averiguar leis.

É necessário, para tanto, partindo do objetivo, formular: H_0 e H_1 , estabelecer os graus de liberdade (N-2) e o teste de significância do R de Pearson – Razão t.

Calcula-se, então:

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

Compara-se o t calculado e o t tabelado:

t cal > t tab – aceita-se H_1

t cal < t tab – aceita-se H_0

3 – Construção de instrumentos de medida psicológicos

– Fidedignidade – correlacionam-se:

teste – reteste

consistência interna

r = teste/acerto/erro

0,60 – dados razoáveis

0,80 – fidedignidade boa

– Validade – correlacionam-se os escores do teste com os escores do critério:

Acima de 0,40

Na interpretação do coeficiente de correlação, deve ser considerada a finalidade com a qual o r foi calculado.

Na escolha do coeficiente de correlação deve-se considerar alguns critérios a fim de que se encontre um resultado confiável, a saber:

- nível de mensuração
- tipos de variáveis
- normalidade da distribuição
- homocedasticidade (linha de regressão)
- linearidade

Coeficiente de correlação de pearson

Exigências:

- distribuição normal (ver assimetria e curtose em amostra < 30)
- medida intervalar (nas duas variáveis)
- variáveis contínuas
- variância aproximada – homocedasticidade

$$v = \frac{\sum \chi^2}{N}$$

- linearidade

Fórmulas:

$$1) r = \frac{\sum zx.zy}{N}$$

(quando N é pequeno)

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

Exemplo:

Ind.	X	Y	XY	X ²	Y ²
A	2	4	8	4	16
B	4	6	24	16	36
C	6	8	48	36	64
D	8	10	80	64	100
E	10	12	120	100	144
Σ	30	40	280	220	360

$$r_s = 1 - \frac{6(96)}{10(10^2 - 1)} = 0,42$$

Obs.: No caso de postos empatados, devemos somar esses postos (o “valor” numérico deles) e dividir pelo número de empates.

Coefficiente de correlação spearman-brown

Exigências:

- poucos dados
- linearidade
- nível de medida original
- amostra aleatória

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

r_s = coeficiente de correlação de postos

D = diferença entre postos (relativa ao mesmo sujeito em ambas as variáveis)

N = número de respondentes

Exemplo:

Ind.	X	Y	D	D ²
1	6°	4°	2	4
2	4°	1°	3	9
3	7°	9°	-2	4
4	2°	3°	-1	1
5	3°	5°	-2	4
6	1°	6°	-5	25
7	8°	2°	6	36
8	5°	8°	-3	9
9	10°	10°	0	0
10	9°	7°	2	2
Σ	-	-	-	96