



Associações entre Variáveis Quantitativas - Parte 3

# Aula	40
<input checked="" type="checkbox"/> Preparada	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Revisada	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> Lecionada	<input checked="" type="checkbox"/>

▼ Até aqui vimos que a Correlação:

- Resume o sentido e a força da associação linear...
- Entre duas variáveis quantitativas (digamos x e y)...
- É denotada por r (ou por ρ)...

▼ Propriedades da Correlação:

▼ Valores de r e a Associação:

- Valor positivo de $r \rightarrow$ associação positiva...
- Valor negativo de $r \rightarrow$ associação negativa...

▼ Força da Associação:

- Assume valores entre -1 e 1...
- Quanto mais próximo de ± 1 , mais forte é esta associação...
- Quanto mais próximo de 0, mais fraca é esta associação...

▼ Se $x = y$...

$$\text{Correlação}(x, x) = \rho_{(x,x)} = r_{(x,x)} = \frac{\sum_{i=1}^n [(x_i - \bar{x}) \times (x_i - \bar{x})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \times \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}} = 1$$

▼ Vantagens da Correlação:

▼ A Correlação é adimensional:

- Ela não depende das unidades das variáveis envolvidas.
- Isso significa que a correlação entre x e y pode ser comparada com a correlação entre z e w ...

▼ A Correlação é simétrica:

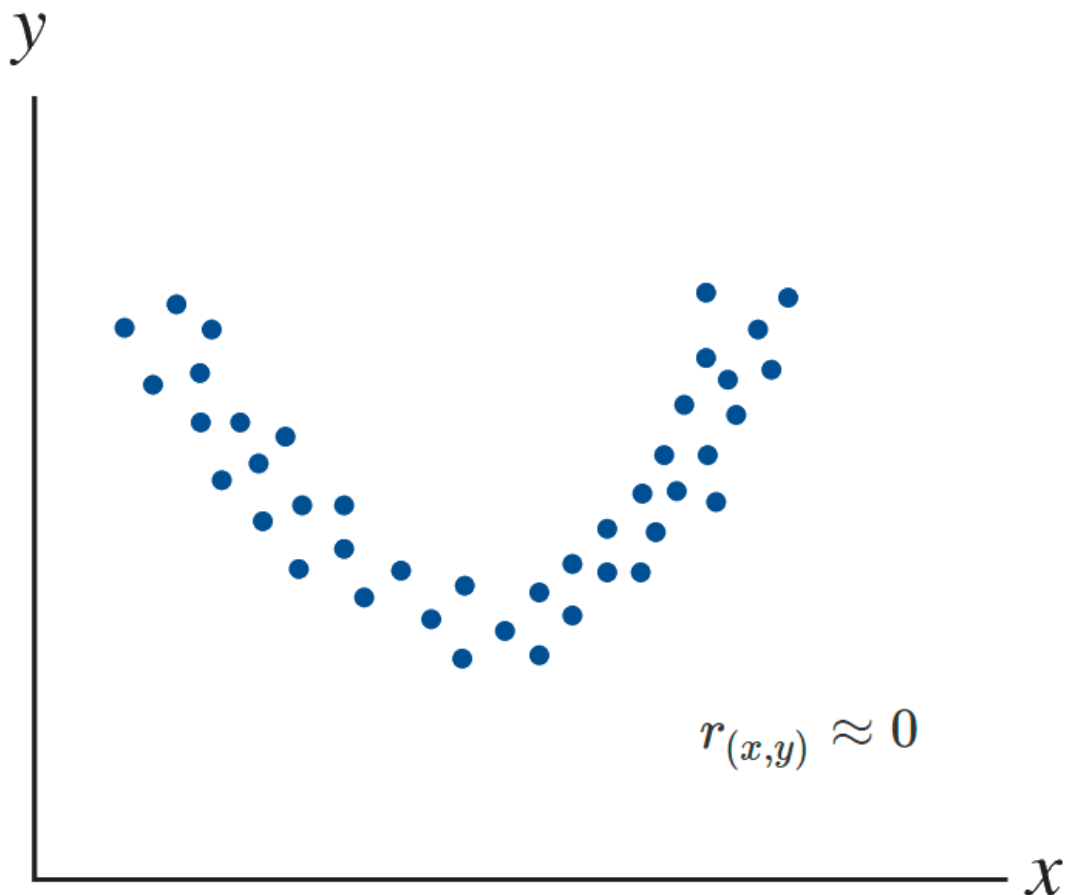
- A correlação entre x e y é igual à correlação entre y e x :

$$\text{Correlação}(x, y) = \rho_{(x,y)} = r_{(x,y)} = \frac{\sum_{i=1}^n [(x_i - \bar{x}) \times (y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \times \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} =$$

$$\text{Correlação}(y, x) = \rho_{(y,x)} = r_{(y,x)} = \frac{\sum_{i=1}^n [(y_i - \bar{y}) \times (x_i - \bar{x})]}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \times \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}}$$

▼ Limitações da Correlação:

- A Correlação é uma medida de **associação linear**. Associações não lineares não são representadas pela Correlação.



- A Correlação não deve ser apenas calculada, mas também plotada!

▼ Lembre-se do Quarteto de Anscombe:

Quarteto de Anscombe

Quarteto de Anscombe são quatro conjuntos de dados que têm estatísticas descritivas quase idênticas (como a média e a variância), mas que têm distribuições muito diferentes e aparências muito distintas quando exibidos graficamente. Cada

↗ https://pt.wikipedia.org/wiki/Quarteto_de_Anscombe

