

# Algumas medidas de associação

## 1. Dados qualitativos *vs* dados qualitativos

- Nominal *vs* nominal ou ordinal

- Coeficiente de contingência de Pearson. Toma valores entre 0 e 1, mas nunca atinge o valor 1. O valor máximo que pode atingir é  $\sqrt{\frac{m-1}{m}}$ , onde  $m$  é o mínimo entre o número de linhas e o número de colunas da tabela. O valor 0 corresponde a ausência de associação entre as variáveis, valores próximos de zero correspondem a fraca associação e valores mais elevados correspondem a associação mais forte.
- Coeficiente de contingência  $V$  de Cramer. Toma valores entre 0 e 1. O valor 0 corresponde a ausência de associação entre as variáveis, valores próximos de zero correspondem a fraca associação e valores mais próximos de 1 correspondem a associação mais forte.

- Ordinal *vs* ordinal

- Coeficiente  $tau-b$  de Kendall. Pode tomar valores entre  $-1$  e  $1$  mas os valores  $-1$  e  $1$  só são atingidos em tabelas em que o número de linhas é igual ao número de colunas. Valores próximos de  $1$  ou de  $-1$  indicam forte associação. Se o sinal é positivo, então o “crescimento” de uma das variáveis é acompanhado do “crescimento” da outra; se o sinal é negativo, então o “crescimento” de uma das variáveis é acompanhado do “decréscimento” da outra.  
Valores próximos de  $0$  indicam fraca associação.

## 2. Dados quantitativos *vs* dados quantitativos

- Coeficiente de correlação de Pearson. Mede a relação linear entre as variáveis. Pode tomar valores entre  $-1$  e  $1$ . A relação entre as variáveis está tão mais próxima da linear quanto mais próximo de  $1$  ou de  $-1$  estiver o valor do coeficiente de correlação. Se o sinal é positivo, então o crescimento dos valores de uma das variáveis é acompanhado do crescimento dos valores da outra; se o sinal é negativo, então o crescimento dos valores de uma das variáveis é acompanhado do decréscimento dos valores da outra.

Se o coeficiente de correlação está próximo de  $0$  é de excluir a existência de uma relação linear entre as variáveis.

## 3. Dados qualitativos *vs* dados quantitativos

- Coeficiente  $Eta$ . Para o cálculo deste coeficiente, a variável qualitativa deve ser codificada numericamente.  $Eta$  pode tomar valores entre  $0$  e  $1$ . Valores próximos de  $1$  indicam forte associação e valores próximos de  $0$  indicam fraca associação. Como habitualmente, o valor  $0$  indica total ausência de associação.