

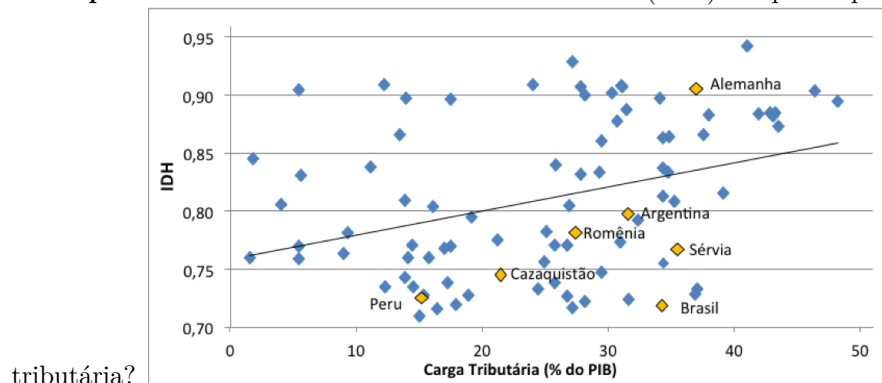


Plano Aula 07 e 08

Medidas de Associação

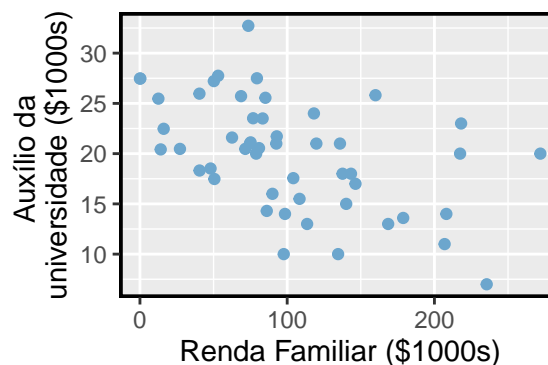
Agora, nosso interesse será analisar o relacionamento entre **duas variáveis numéricas** de interesse.

- **Exemplo 1:** O Índice de Desenvolvimento Humano (IDH) em países pode estar associado à carga tributária?



Artigo de 2012: <https://carodinho.blogfolha.uol.com.br/2012/12/14/pagamento-de-impostos-no-brasil-e-um-investimento-sem-retorno/>

- **Exemplo 2:** O valor do auxílio estudantil oferecido por uma universidade pode estar relacionado com a renda familiar dos estudantes?



Associação entre Variáveis Quantitativas (Bussab e Morettin - seção 4.5)

Para duas variáveis quantitativas também podemos estar interessados em verificar se existe associação (relação) entre elas.

Gráfico de Dispersão

- ... cont. exemplo 1:

Como resumir a informação do gráfico acima em um só número?



Coeficiente de correlação (linear) (de Pearson)

Relembrando sobre covariância em probabilidade:

Definição (**covariância**): Sejam X e Y duas v.a. então $Cov(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$.

Definição (**covariância amostral**): Dados n pares de valores observados $(x_1, y_1), \dots, (x_n, y_n)$ (de duas v.a. X e Y), chamaremos de covariância amostral entre X e Y a expressão

$$cov(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n}$$

.

Assim, “padronizamos” a covariância para obtemos o coeficiente $corr(X, Y) = \frac{cov(X, Y)}{s_x \cdot s_y}$, $-1 \leq corr(X, Y) \leq 1$.

Definição (**coeficiente de correlação**): Dados n pares de valores observados $(x_1, y_1), \dots, (x_n, y_n)$ (de duas v.a. X e Y), chamaremos de covariância amostral entre X e Y a expressão

$$corr(X, Y) = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})}{s_x} \cdot \frac{(y_i - \bar{y})}{s_y}.$$

Em Estatística Geral 2 veremos como usar os valores de uma amostra observada para testar se existe correlação (associação)?

Correlação espúria

Causalidade e correlação

Ler slides e ver vídeos da semana 4.

Continuar lista de exercícios 1-2.
