



Plano Aula 27 e 28

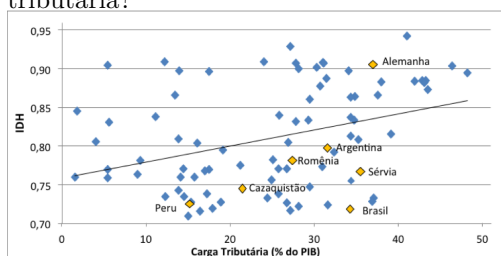
Markus Stein

Análise de Regressão

Agora, nosso interesse será em estimar uma relação **linear** entre **duas variáveis numéricas** de interesse.

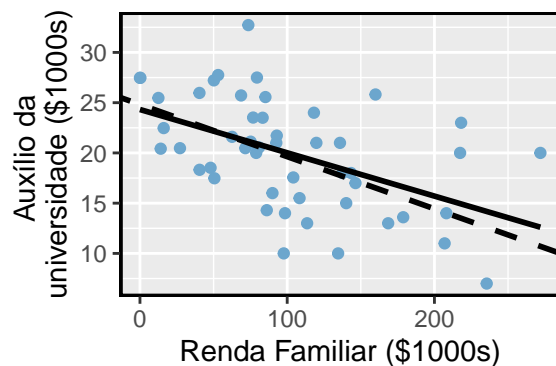
Regressão Linear Simples (Bussab e Morettin - capítulo 16)

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



Artigo de 2012: <https://carodineiro.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?



Estimação dos parâmetros (Bussab e Morettin - seção 16.2)

Relembrando sobre esperança condicional em probabilidade...



Modelo populacional

sejam X e Y duas v.a. queremos estimar a esperança condicional de Y em função de (dado que) $X = x$,

$$E(Y|X = x) = \alpha + \beta \cdot x,$$

ou seja, queremos estimar os parâmetros α e β .

- Para uma amostra de tamanho n podemos escrever que cada observação (x_i, y_i) , para $i = 1, \dots, n$, segue o modelo

$$y_i = \alpha + \beta \cdot x_i + e_i.$$

- chamamos e_i de erro amostral e assumimos que:
 - $E(e_i) = 0$,
 - $Var(e_i) = \sigma^2$, para todo $i, j = 1, \dots, n$,
 - $Cov(e_i, e_j) = 0$ para $i \neq j$.
- Assim $E(y_i) = \alpha + \beta \cdot x_i$, $Var(y_i) = \sigma^2$ e $Cov(y_i, y_j) = 0$.
 - o parâmetro σ^2 também precisa ser estimado.

Método dos Mínimos Quadrados (Ordinários)

Para estimar α e β podemos pensar em minimizar os erros e_i , ou

$$SQ(\alpha, \beta) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - (\alpha + \beta \cdot x_i)]^2$$

- Derivando $SQ(\alpha, \beta)$ em relação a α e β , igualando a zero e resolvendo o sistema de equações temos
 - $\hat{\alpha} = ???$ e $\hat{\beta} = ???$.
- Reta estimada (modelo ajustado): $\hat{y}_i = \hat{\alpha} + \hat{\beta} \cdot x_i$
 - Interpretação de $\hat{\alpha}$ e $\hat{\beta}$;
 - **Prever**, para um dado valor $X = x$, quanto esperamos observar o valor de Y ?

Coeficiente de determinação R^2 (Bussab e Morettin - seção 16.3)

Intervalos de Confiança e Testes de hipóteses (Bussab e Morettin - seção 16.4)

Para α , β e σ^2 . (suposições???)

Se adicionalmente assumimos $e_i \sim Normal$, então $y_i \sim Normal(\alpha + \beta \cdot x, \sigma^2)$.

- Também $\hat{\alpha} \sim Normal$ e $\hat{\beta} \sim Normal$.



Previsão e predição (Bussab e Morettin - seção 16.4.4)

- para o valor esperado $E(Y_i|x_i) = y_i$;
- para uma futura (nova) observação y_f .

Correlação espúria

Causalidade e correlação

Ler slides das aulas 27 e 28

Continuar exercícios da lista 3-3

Fazer avaliação pontual 2 da área 3

REFERÊNCIA EXTRA

Página ‘Probabilidade e Estatística (EaD)’ da UFRGS

- Capítulo 7 - Introdução à Regressão Linear