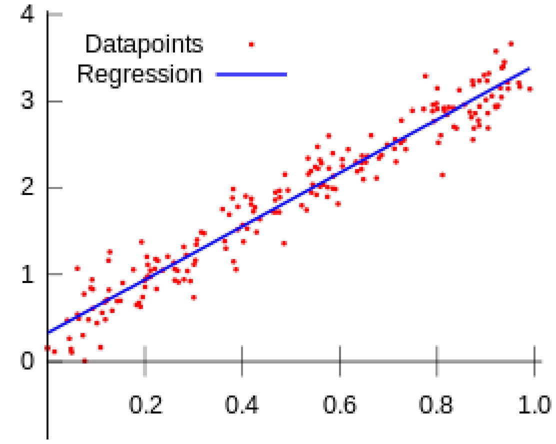
WIKIPÉDIA

Regressão linear

Origem: Wikipédia, a enciclopédia livre.

Em estatística ou econometria, **regressão linear** é uma equação para se estimar a condicional (valor esperado) de uma variável *y*, dados os valores de algumas outras variáveis *x*.[1][2]

A *regressão*, em geral, tem como objectivo tratar de um valor que não se consegue estimar inicialmente.



Exemplo de regressão linear.

A regressão linear é chamada "linear" porque se considera que a relação da resposta às variáveis é uma função linear de alguns parâmetros. Os modelos de regressão que não são uma função linear dos parâmetros se chamam modelos de regressão não-linear. Sendo uma das primeiras formas de análise *regressiva* a ser estudada rigorosamente, e usada extensamente em aplicações práticas. Isso acontece porque modelos que dependem de forma linear dos seus parâmetros desconhecidos, são mais fáceis de ajustar que os modelos não-lineares aos seus parâmetros, e porque as propriedades estatísticas dos estimadores resultantes são fáceis de determinar.[3]

Modelos de regressão linear são frequentemente ajustados usando a abordagem dos mínimos quadrados, mas que também pode ser montada de outras maneiras, tal como minimizando a "falta de ajuste" em alguma outra norma (com menos desvios absolutos de regressão), ou através da minimização de uma penalização da versão dos mínimos quadrados. Por outro lado, a abordagem de mínimos quadrados pode ser utilizado para ajustar a modelos que não são modelos lineares. Assim, embora os termos "mínimos quadrados" e "modelo linear" estejam intimamente ligados, eles não são sinônimos. [*carece de fontes*?]

Índice

Equação da Regressão Linear

Para se estimar o valor esperado, usa-se de uma equação, que determina a relação entre ambas as variáveis.

$y_i = \alpha + \beta \, X_i + \varepsilon_i$

, onde:

y_i: Variável explicada (dependente); representa o que o modelo tentará prever

α: É uma constante, que representa a interceptação da reta com o eixo vertical;

β: Representa a inclinação (coeficiente angular) em relação à variável explicativa;

X_i: Variável explicativa (independente);

ε_i: Representa todos os factores residuais mais os possíveis erros de medição. O seu comportamento é aleatório, devido à natureza dos factores que encerra. Para que essa fórmula possa ser aplicada, os erros devem satisfazer determinadas hipóteses, que são: terem distribuição normal, com a mesma variância σ^2 , independentes e independentes da variável explicativa X, ou seja, i.i.d. (independentes e identicamente distribuídas).

Notação Matricial

A equação acima pode ser reescrita em forma de matriz:

$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$

Onde $\mathbf{y}$ é uma matriz de $n \times 1$ observações, $\mathbf{X}$ é uma matriz de tamanho $n \times p+1$ (sendo a primeira coluna com valores sempre = 1, representando a constante α, e p é a quantidade de variáveis explicativas), $\boldsymbol{\beta}$ é uma matriz de $p+1 \times 1$ variáveis explicativas (sendo que β_0 representa a constante α) e $\boldsymbol{\varepsilon}$ é uma matriz de $n \times 1$ de resíduos.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1p} \\ 1 & X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Estimativa dos fatores α e β

A técnica mais usual para estimativa dos parâmetros α e β é o Método dos mínimos quadrados, mas também podem ser usados:

- Mínimos Quadrados Ponderados
- Mínimos quadrados generalizados

- Máxima verossimilhança
- Regularização de Tikhonov
- Mínimo Desvio absoluto

Ver também

- Método dos mínimos quadrados
- Regressão não linear
- Regressão

Ligações Externas

- SysLinea 0.1.2 (https://sites.google.com/site/mgbfreeware/) : Programa de código aberto com regressão linear e não linear.
- Manual da Regressão Linear (https://www.fm2s.com.br/regressao-linear-economizar-milhoes/)

Referências

1. «Linear regression» (https://lagunita.stanford.edu/c4x/HumanitiesScience/StatLearning/asset/linear_regression.pdf) (PDF) (em inglês). Stanford.edu. Consultado em 10 de julho de 2019
2. «Chapter 9 - Simple linear regression» (http://www.stat.cmu.edu/~hseltman/309/Book/chapter9.pdf) (PDF) (em inglês). Carnegie Mellon University - Statistics & Data Science. Consultado em 10 de julho de 2019
3. http://www.fisica.ufs.br/egsantana/cinematica/regresion/regresion.htm Regressão linear com experimêntos físicos [ligação inativa]

Bibliografia

- REIS, E., *Estatistica Descritiva* (2ª ed.). Lisboa: Edições Sílabo, 1994