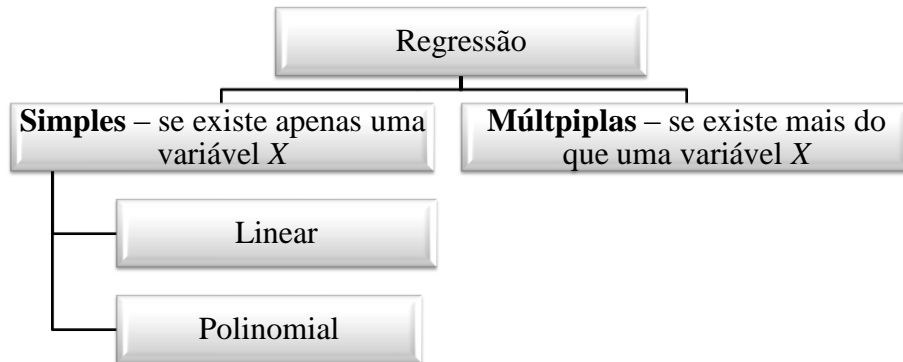


## REGRESSÃO

Quantifica a relação entre uma dada variável (em geral chamada de variável *explicada* ou *dependente* –  $Y$ ) e outra ou mais variáveis (em geral chamadas de variáveis *explicativas* ou *independentes* –  $X$ )



## OBJETIVO

1. Determinar a relação existente entre uma variável qualquer de interesse, dependente, e outra característica independente, tomadas juntas;
2. Prever o valor de  $Y$  para um dado conjunto de  $X$ ;
3. Examinar se alguma das variáveis  $X$  tem efeito significativo em  $Y$ .

Tal relação é expressa por uma função matemática (equação de regressão), onde se diz que a variável dependente ( $Y$ ) é uma função da variável independente ( $X$ )

## REGRESSÃO LINEAR SIMPLES

$$Y = a + bX \text{ ou } Y = b_0 + b_{01}X$$

onde:

$\hat{Y}$  = estimativa da variável dependente;

a ou  $b_0$  = intercepção no eixo dos Y, ou seja, o valor de Y quando X = 0;

b ou  $b_1$  = coeficiente angular da reta, que determina a declividade da mesma e expressa o valor de Y para X = 0;

X = variável independente

## ESTIMATIVA DOS PARÂMETROS DE REGRESSÃO

As estimativas dos parâmetros a e b são obtidas pelas fórmulas:

$$a = \bar{Y} - b\bar{X}$$

$$b = \frac{S_{XY}}{S_{XX}}$$

$$S_{YY} = \sum Y^2 - N\bar{Y}^2$$

ou

$$S_{YY} = \sum Y^2 - \frac{(\sum Y)^2}{N}$$

$$S_{XY} = \sum XY - N\bar{X}\bar{Y}$$

ou

$$S_{XY} = \sum XY - \frac{(\sum X)(\sum Y)}{N}$$

$$S_{XX} = \sum X^2 - N\bar{X}^2$$

ou

$$S_{XX} = \sum X^2 - \frac{(\sum X)^2}{N}$$

Então podemos dizer que:

$$b = \frac{S_{XY}}{S_{XX}} \quad \text{é} \quad b = \frac{\sum XY - n\bar{X}\bar{Y}}{\sum X^2 - n\bar{X}^2} \quad \text{ou} \quad b = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum X^2 - \frac{(\sum X)^2}{N}}$$

### ***ANÁLISE DE VARIÂNCIA PARA O MODELO DE REGRESSÃO LINEAR SIMPLES***

Ela consiste em compor a soma dos quadrados totais (SQT) em duas partes: soma dos quadrados explicados ou da regressão (SQE) e soma dos quadrados residuais (SQR).

O propósito de se apresentar a tabela é testar a significância de b.

#### ***O procedimento de análise de variância é o seguinte***

Fonte de Variação	GL	Soma dos quadrados	Quadrados médios	F
Regressão ou Explicados (x)	1	SQE = bS <sub>XY</sub>	QME = SQE/GLE	QME/QMR
Resíduo	N - 2	SQR = SQT - SQE	QMR = SQR/GLR	
Total	N - 1	SQT = S <sub>YY</sub>		

Onde

$$SQT = S_{YY} = \sum Y^2 - \frac{(\sum Y)^2}{N}$$

$$SQR = SQT - SQE$$

$$SQE = bS_{XY} = b \left( \sum XY - \frac{(\sum X)(\sum Y)}{N} \right) = \frac{\left[ \sum XY - \frac{(\sum X)(\sum Y)}{N} \right]^2}{\sum X^2 - \frac{(\sum X)^2}{N}}$$

$$F \text{ Calculado para a Regressão} = \frac{\text{Quadrado Médio da Regressão}}{\text{Quadrado Médio do Resíduo}}$$

### DETERMINAÇÃO DO COEFICIENTE DE DETERMINAÇÃO

$$R_{XY}^2 = \frac{SQE}{SQT} = \frac{SQT - SQR}{SQT} = \frac{S_{XY}^2}{S_{XX}S_{YY}} = \frac{bS_{XY}}{S_{YY}}$$

$$R^2 = \frac{\hat{b} \sum XY - \frac{(\sum X)(\sum Y)}{N}}{\sum Y^2 - \frac{(\sum Y)^2}{N}}$$