

REGRESSÃO LINEAR

ASPECTOS GERAIS

= calcular a **expressão matemática** que relaciona duas variáveis x e y:

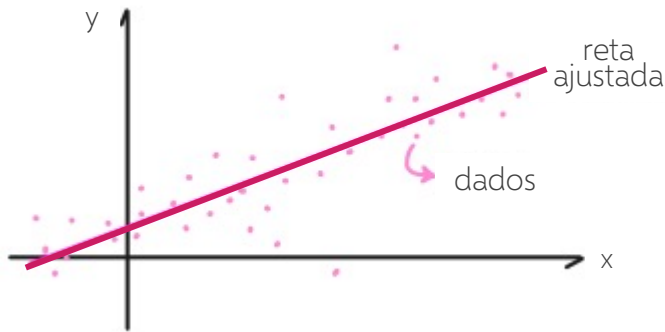
$$Y = p + m \cdot x$$

p = coeficiente linear

m = coeficiente angular

$$m = \frac{\Delta y}{\Delta x}$$

Exemplo:



COEFICIENTE DE CORRELAÇÃO LINEAR DE PEARSON

$$r = \frac{\sum[(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]}{\sqrt{\sum[(X_i - \bar{X})^2 \cdot (Y_i - \bar{Y})^2]}}$$

Mede a força da relação linear entre 2 variáveis x e y

$-1 \leq r \leq 1$:

- Mais próximo de -1 ou 1 → Correlação mais forte
- Mais próximo de zero → Correlação mais fraca (não há relação linear entre as variáveis)

Em questões, é útil usar:

$$\sum[(X_i - \bar{X}) \cdot (Y_i - \bar{Y})] = \sum(X_i \cdot Y_i) - n \cdot \bar{X} \cdot \bar{Y}$$

MODELO ESTATÍSTICO

$$Y_i = \alpha + \beta X_i + \mu_i$$

x = variável independente

y = variável dependente

n = número de pares de valores observados

μ_i = erro ou desvio

$\alpha + \beta X_i$ = componente de y que varia linearmente com x

RETA DE REGRESSÃO ESTIMADA

$$\hat{Y} = a + bX_i \quad e = Y_i - \hat{Y}_i \rightarrow \text{desvio}$$

MÉTODO DOS MÍNIMOS QUADRADOS

= determina as estimativas de a e b dos parâmetros **minimizando a soma dos quadrados** dos desvios

Após calcular b, basta substituir na equação $\hat{Y} = a + b\bar{X}$ para determinar a

$$b = \frac{\sum[(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]}{\sum(X_i - \bar{X})^2}$$

RETA QUE PASSA PELA ORIGEM

• $\alpha = 0$

$$Y_i = \beta X_i + \varepsilon_i$$

$$\beta = \frac{\sum X \cdot Y}{\sum X^2}$$

ANÁLISE DA VARIÂNCIA DA REGRESSÃO

regressão linear

$$SQT = SQM + SQR$$

Soma dos quadrados total

Soma dos quadrados do modelo

Soma dos quadrados dos resíduos

FÓRMULAS IMPORTANTES

$$SQM = b \cdot \sum [(X_i - \bar{X}) \cdot (Y_i - \bar{Y})]$$

Soma dos quadrados do modelo de regressão

ou

$$SQM = b^2 \cdot \sum (X_i - \bar{X})^2$$

$$SQR = \sum (Y_i - \bar{Y})^2$$

É a parte do desvio total não explicada pelo modelo de regressão

TABELA DA ANOVA



FONTE DE VARIAÇÃO	GRAUS DE LIBERDADE (gl)	SOMA DE QUADRADOS	QUADRADOS MÉDIOS	F
MODELO	1	SQM	$QMM = \frac{SQM}{1}$	$F_{teste} = \frac{SMM}{QMR}$
RESÍDUOS	n-2	SQR	$QMR = \frac{SQR}{n-2}$	
TOTAL	n-1	SQT	$QMT = \frac{SQT}{n-1}$	

n = tamanho da amostra

(QMR = Estimativa de variância σ^2 residual)

COEFICIENTES



- Coeficiente de **correlação**:

$$R = \sqrt{\frac{SQM}{SQT}}$$

- Coeficiente de **determinação**:

Exprime a proporção da variação total de y que é explicada pela reta de regressão

$$R^2 = 1 - \frac{SQR}{SQT}$$

$$(0 \leq R^2 \leq 1)$$