



CENTRO DE CIÊNCIA E TECNOLOGIA  
LABORATÓRIO DE CIÊNCIAS MATEMÁTICAS  
UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE

**Curso:** Ciência da Computação **Disciplina:** Estatística e Probabilidade

**Data:** 24./06./2024

## Estatísticas Gerais

Tamanho da amostra:  $n$

Média amostral:  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$   $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$

Quadrado das diferenças:

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2 = \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2 / n$$

$$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n y_i^2 - n\bar{y}^2 = \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2 / n$$

$$S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - n\bar{x}\bar{y} \\ = \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i) / n$$

Variância amostral:  $S_x^2 = \frac{1}{n-1} S_{xx}$   $S_y^2 = \frac{1}{n-1} S_{yy}$

Covariância amostral:  $Cov_{xy} = \frac{1}{n-1} S_{xy}$

Desvio padrão amostral:  $S_x = \sqrt{\frac{1}{n-1} S_{xx}}$   $S_y = \sqrt{\frac{1}{n-1} S_{yy}}$

## Regressão e Correlação

### Coefficiente de Determinação

$$SQE = \sum (y_i - \hat{y}_i)^2 = \sum y_i^2 - \hat{\beta}_0 \sum y_i - \hat{\beta}_1 \sum x_i y_i$$

$$SQT = \sum (y_i - \bar{y})^2 = S_{yy}$$

$$r^2 = 1 - \frac{SQE}{SQT}$$

### Coefficiente de Correlação

$$r = \frac{S_{xy}}{\sqrt{S_{xx}} \sqrt{S_{yy}}}$$

### Reta de Regressão

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$