

### Objetivos da estatística

Fazer inferências sobre a população com base em dados amostrais

Planejar o experimento e o tamanho da amostra para que tais inferências tenham uma base confiável

### Variáveis

O que é?

É uma característica da população

### Tipos Variáveis

Dependentes: Efeitos são esperados

Independentes: Efeitos que queremos medir

### Variáveis Qualitativas

Nominais: não existe ordenação entre as categorias

Ordinais: existe uma ordenação entre as categorias

### Variáveis Quantitativas

Discretas: geralmente são o resultado de contagens

Contínuas: características mensuráveis que assumem valores em uma escala contínua (na reta real)

### Tipos de Gráficos

Colunas e Barras

Histograma

Setores

Linhas

Diagrama de Ordenadas

### Medidas Descritivas - Objetivo

Sumarizar conjuntos de dados quanto a:

Centralidade -> Medidas de localização

Variabilidade -> Medidas de dispersão

### Medidas de Dispersão

Variância  $(s^2) = \sum[(x_i - \bar{x})^2] / n - 1$

Desvio  $(x_i - \bar{x})$

Desvio Médio  $\sum[(x_i - \bar{x})^2] / n$

Amplitude: maior valor - menor valor

CV  $s / \bar{x}$

### Medidas de Localização

Moda: Valor que mais se repete

Média  $\sum x_i / n$

Mediana: Valor médio quando ordenado

### Intervalo de Variação

Combinação entre medidas de localização e dispersão

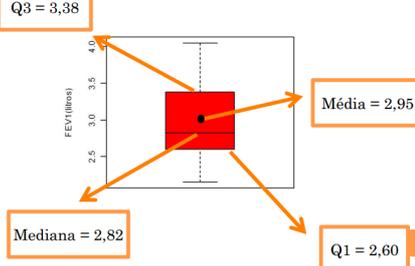
### Lei de Chebychev

510g de Vitamina C, desvio padrão de 3g

75%  $(1 - 1/2^2)$  possuem entre 504g e 516g

### Boxplot

#### BOXPLOT - EXEMPLO I



### Razão de Verossimilhança Positiva e Negativa

**Positiva** É uma razão entre a probabilidade de um teste ser positivo, dado que existe a doença, e a probabilidade de um teste ser positivo, dado que não existe a doença.

**Negativa** É uma razão entre a probabilidade de um teste ser negativo, dado que existe a doença, e a probabilidade de um teste ser negativo, dado que não existe a doença.

### Cálculo do Intervalo de Confiança

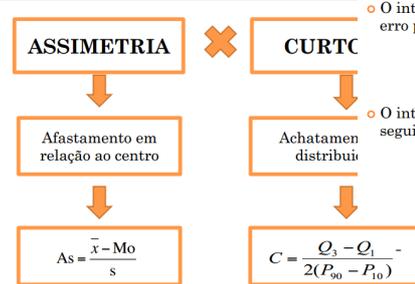
O intervalo de confiança é construído a partir do erro padrão.

$$EP = \sqrt{\frac{(1-p)p}{n}}$$

O intervalo de confiança de 95% é estimado pela seguinte fórmula:

$$p \pm 1,96 EP$$

### Assimetria x Curtose



### Sensibilidade e Especificidade

#### SENSIBILIDADE E ESPECIFICIDADE

	D	ND	
T+	a	b	Sensibilidade = $\frac{a}{a+c}$ Especificidade = $\frac{d}{b+d}$
T-	c	d	

- Sensibilidade: Probabilidade de um teste ser positivo, dado que existe a doença.
- Especificidade: Probabilidade de um teste ser negativo, dado que não existe a doença.

### Função Discreta de Probabilidade

Função que atribui a cada valor da variável aleatória sua probabilidade.

$$P(X = x_i) = p(x_i), i = 1, 2, \dots$$

$$0 \leq p_i \leq 1 \text{ e } \sum p_i = 1$$



By malandro123

Published 25th September, 2015.

Last updated 13th May, 2016.

Page 1 of 2.

Sponsored by [CrosswordCheats.com](http://CrosswordCheats.com)

Learn to solve cryptic crosswords!

<http://crosswordcheats.com>

### Distribuição Normal

Mais importante distribuição de probabilidade, aplicada a inúmeros fenômenos e utilizada para o desenvolvimento teórico da Estatística.

Seja X uma variável aleatória. X terá distribuição normal se:

$$f(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

Sendo:  $\mu$  = média da distribuição  
 $\sigma$  = desvio-padrão da distribuição  
 $e = 3,1416$   
 $e = 2,7$   
 $e - \infty < X < + \infty$

X tem distribuição Normal com média  $\mu$  e variância  $\sigma^2$ :  $X \sim N(\mu, \sigma^2)$ .

### Distribuição Normal Padrão

Para calcular probabilidades  $P(X \leq x)$ , é necessário integrar  $f(X)$  para diferentes valores de  $\mu$  e  $\sigma$ . A solução é transformar a variável X na variável Z, dada por:

$$Z_i = \frac{X_i - \mu}{\sigma}$$

A distribuição de probabilidades associada à variável Z denomina-se distribuição normal padrão:

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

Sendo: média da distribuição = 0  
 variância = 1  
 $e - \infty < Z < + \infty$

Z tem distribuição Normal com média 0 e variância 1:  $Z \sim N(0,1)$ .

### Exemplo

As alturas dos alunos de determinada escola são normalmente distribuídas com média 1,60 m e desvio-padrão 0,30 m. Encontre a probabilidade de um aluno medir:

- entre 1,50 m e 1,80 m -  $P(1,50 \leq X \leq 1,80)$ ;
- mais de 1,75 m -  $P(X > 1,75)$
- menos de 1,48 m -  $P(X < 1,48)$

Qual intervalo compreende 90% da população? -  $P(a \leq X \leq b = 90\%)$

- $P(1,50 \leq X \leq 1,80) = P(1,50 - 1,60, 0,30 \leq Z \leq 1,80 - 1,60, 0,30) = P(-0,33 \leq Z \leq 0,67)$

Pela tabela:  $P = 1 - (0,3707 + 0,2514) = 0,3779$

- $P(X > 1,75) = P(Z > 1,75 - 1,60, 0,30) = P(Z > 0,50)$

Pela tabela:  $P = 0,3085$

- $P(X < 1,48) = P(Z < 1,48 - 1,60, 0,30) = P(Z < -0,40)$

Pela tabela:  $P = 0,3446$

$P(a \leq X \leq b = 90\%) = P(a - 1,60, 0,30 \leq Z \leq b - 1,60, 0,30 = 90\%) \rightarrow a - 1,60, 0,30 = -Z_{15\%}$  e  $b - 1,60, 0,30 = Z_{15\%}$

$a - 1,60, 0,30 = -1,64$  e  $b - 1,60, 0,30 = 1,64 \rightarrow a = 1,11$  m e  $b = 2,09$  m

$\rightarrow (1,11 \leq X \leq 2,09)$



By malandro123

Published 25th September, 2015.

Last updated 13th May, 2016.

Page 2 of 2.

Sponsored by [CrosswordCheats.com](http://CrosswordCheats.com)

Learn to solve cryptic crosswords!

<http://crosswordcheats.com>