# Probabilidade e Estatística

**Professores:**

Otton Teixeira da Silveira Filho

Regina Célia Paula Leal Toledo

# Lembrando ...

→ Funções discretas de probabilidade

→ Modelos discretos

Uniforme
Bernoulli
Binomial
Geométrico
Poisson
Hipergeométrico

X

# Aula 6

Professores:

*Otton Teixeira da Silveira Filho*
*Regina Célia Paula Leal Toledo*

# Medidas de resumo (para variáveis aleatórias discretas)

Conteúdo:

6.1 Medidas de posição
- média
- mediana
- moda

6.2 Exemplo

6.3 Medidas de dispersão
- variância
- desvio padrão

$$\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$$\sigma_x^2 \qquad \bar{Y}=\frac{\sum_{i=1}^{n} y_i}{n}$$

$$A\cap B=\varnothing$$

# 6. Medidas de resumo para variáveis aleatórias discretas

Descrição do comportamento das variáveis aleatórias discretas feitas através

função de probabilidade

Medida utilizada para resumir o comportamento das variáveis aleatórias discretas

# 6. Medidas de resumo para variáveis aleatórias discretas

Descrição do comportamento das variáveis aleatórias discretas feitas através

↓

função de probabilidade

↕

Medida utilizada para resumir o comportamento das variáveis aleatórias discretas

## 6.1 Medidas de posição (tendência central)

medidas mais comuns:

- média
- mediana
- moda

| X | $x_1$ | $x_2$ | ... | $x_k$ |
|---|---|---|---|---|
| $p_i$ | $p_1$ | $p_2$ | ... | $p_k$ |

A **média** ou valor esperado ou esperança de uma variável **X** é dada pela expressão:

$$E(X) = \sum_{i=1}^{k} x_i p_i$$

| X | $x_1$ | $x_2$ | ... | $x_k$ |
|---|---|---|---|---|
| $p_i$ | $p_1$ | $p_2$ | ... | $p_k$ |

A **média** ou valor esperado ou esperança de uma variável **X** é dada pela expressão:

$$E(X) = \sum_{i=1}^{k} x_i p_i$$

Outras notações utilizadas para E(**X**)

$\mu_x$ ou simplesmente $\mu$

## Exemplo:

Considere a variável aleatória **X** e a função discreta de probabilidade a ela associada $p_i$

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

A média ou esperança ou valor esperado é dado por:

## Exemplo:

Considere a variável aleatória **X** e a função discreta de probabilidade a ela associada $p_i$

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

A média ou esperança ou valor esperado é dado por:

$$\mu = \sum_{i=1}^{4} x_i p_i = -5 \times 0,3 + 10 \times 0,2 + 15 \times 0,4 + 20 \times 0,1$$

$$\boxed{\mu = 8,5}$$

A mediana é o valor que satisfaz às seguintes condições:

$$P(\mathbf{X} \geq Md) \geq 0,5$$

e

$$P(\mathbf{X} \leq Md) \geq 0,5$$

A mediana é o valor que satisfaz às seguintes condições:

$$P(\mathbf{X} \geq Md) \geq 0,5$$

e

$$P(\mathbf{X} \leq Md) \geq 0,5$$

**Exemplo:**

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

A mediana é o valor que satisfaz às seguintes condições:

$$P(\mathbf{X} \geq Md) \geq 0,5$$

e

$$P(\mathbf{X} \leq Md) \geq 0,5$$

**Exemplo:**

$$\begin{array}{c|cccc} X & -5 & 10 & 15 & 20 \\ \hline p_i & \underbrace{0,3 \quad 0,2}_{0,5} & & \underbrace{0,4 \quad 0,1}_{0,5} & \end{array}$$

A mediana pode ser qualquer valor entre 10 e 15.

A mediana é o valor que satisfaz às seguintes condições:

$$P(\mathbf{X} \geq Md) \geq 0,5$$

e

$$P(\mathbf{X} \leq Md) \geq 0,5$$

**Exemplo:**

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

$$\underbrace{0{,}3 \quad 0{,}2}_{0{,}5} \qquad \underbrace{0{,}4 \quad 0{,}1}_{0{,}5}$$

A mediana pode ser qualquer valor entre 10 e 15.
Adotaremos a mediana igual a média, ou seja:

$$\boxed{Md = \frac{10+15}{2} = 12,5}$$

## Observações:

- como no exemplo, às vezes as desigualdades que definem a mediana são verificadas em qualquer valor de um determinado intervalo. Nesse caso tomamos a mediana como o ponto médio do intervalo.

## Observações:

- como no exemplo, às vezes as desigualdades que definem a mediana são verificadas em qualquer valor de um determinado intervalo. Nesse caso tomamos a mediana como o ponto médio do intervalo.

- nem a mediana nem a média precisam ser valores assumidos pela variável aleatória. No exemplo:

$$\mu = 8,5$$

$$Md = 12,5$$

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

A moda (Mo) é o valor (ou valores) da variável que tem a maior probabilidade de ocorrência, ou seja:

$$P(X = Mo) = \max(p_1, p_2, \ldots, p_k)$$

A moda (Mo) é o valor (ou valores) da variável que tem a maior probabilidade de ocorrência, ou seja:

$$P(X = Mo) = \max(p_1, p_2, \ldots, p_k)$$

**Exemplo:**

| $X$ | -5 | 10 | 15 | 20 |
|---|---|---|---|---|
| $p_i$ | 0,3 | 0,2 | 0,4 | 0,1 |

*Mo*

## Outras Observações:

Como nas medidas de resumo para um conjunto de dados tem-se:

- a multiplicação das variáveis aleatórias por uma constante fará com que as medidas de posição fiquem multiplicadas por essa constante;

- a adição de uma constante as variáveis aleatórias fará com que as medidas de posição fiquem acrescidas dessa constante;



Fundação CECIERJ Consórcio cederj

## Outro Exemplo:

Calcular a média, a mediana e a moda.

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} x_i p_i = 2 \times 0,1 + 5 \times 0,3 + 8 \times 0,2 + 15 \times 0,2 + 20 \times 0,2 =$$

$$\mu = 10,3$$

Mediana: $Md = 8$

Moda: $Mo = 5$

## Outro Exemplo:

Calcular a média, a mediana e a moda.

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} x_i p_i = 2\times 0,1 + 5\times 0,3 + 8\times 0,2 + 15\times 0,2 + 20\times 0,2 =$$

$$\mu = 10,3$$

Mediana: $Md = 8$

Moda: $Mo = 5$

# Outro Exemplo:

Calcular a média, a mediana e a moda.

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} x_i p_i = 2\times 0,1 + 5\times 0,3 + 8\times 0,2 + 15\times 0,2 + 20\times 0,2 =$$

$$\mu = 10,3$$

Mediana: $Md = 8$

Moda: $Mo = 5$

# Outro Exemplo:

Calcular a média, a mediana e a moda.

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} x_i p_i = 2\times 0,1 + 5\times 0,3 + 8\times 0,2 + 15\times 0,2 + 20\times 0,2 =$$

$$\mu = 10,3$$

Mediana: $Md = 8$

Moda: $Mo = 5$

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

**Média:**

$$\mu_X = 10,3$$

**Mediana:**

$Md(X) = 8$

**Moda:**

$Mo(X) = 5$

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Se estamos interessados na grandeza ***Y*** = 5 ***X*** -10, temos:

Média:

$$\mu_X = 10,3$$

Mediana:

$Md(X) = 8$

Moda:

$Mo(X) = 5$

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Se estamos interessados na grandeza ***Y*** = 5 ***X*** -10, temos:

Média:

$$\mu_X = 10,3$$

$$\boxed{\mu_Y = 5\mu_X - 10 = 41,5}$$

Mediana:

*Md(X) = 8*

Moda:

*Mo(X) = 5*

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Se estamos interessados na grandeza ***Y*** = 5 ***X*** -10, temos:

**Média:**

$\mu_X = 10,3$

$$\boxed{\mu_Y = 5\mu_X - 10 = 41,5}$$

**Mediana:**

$Md(X) = 8$

$$\boxed{Md(\mathbf{Y}) = 5Md(\mathbf{X}) - 10 = 30}$$

**Moda:**

$Mo(X) = 5$

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Se estamos interessados na grandeza ***Y*** = 5 ***X*** -10, temos:

**Média**:

$\mu_X = 10,3$

$$\boxed{\mu_Y = 5\mu_X - 10 = 41,5}$$

**Mediana**:

*Md(X) = 8*

$$\boxed{Md(\mathbf{Y}) = 5Md(\mathbf{X}) - 10 = 30}$$

**Moda**:

*Mo(X) = 5*

$$\boxed{Mo(\mathbf{Y}) = 5Mo(\mathbf{X}) - 10 = 15}$$

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

$Y=5X-10$

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

$Y=5X-10$

| $Y$ | 0 | 15 | 30 | 65 | 90 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

*Y=5X-10*

| $Y$ | 0 | 15 | 30 | 65 | 90 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} y_i p_i = 0 \times 0,1 + 15 \times 0,3 + 30 \times 0,2 + 65 \times 0,2 + 90 \times 0,2 = 41,5$$

Fundação CECIERJ Consórcio cederj

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| X | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

*Y=5X-10*

| Y | 0 | 15 | 30 | 65 | 90 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} y_i p_i = 0\times 0,1 + 15\times 0,3 + 30\times 0,2 + 65\times 0,2 + 90\times 0,2 = 41,5$$

Mediana: *Md = 30* (onde $P(Y \geq 30) \geq \frac{1}{2}$ e $P(Y \leq 30) \geq \frac{1}{2}$ )

Ou, fazendo pela tabela: ***Y*** = 5 ***X*** -10, temos:

| X | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

*Y=5X-10*

| Y | 0 | 15 | 30 | 65 | 90 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} y_i p_i = 0\times 0,1 + 15\times 0,3 + 30\times 0,2 + 65\times 0,2 + 90\times 0,2 = 41,5$$

Mediana: *Md = 30* (onde $P(Y \geq 30) \geq \frac{1}{2}$ e $P(Y \leq 30) \geq \frac{1}{2}$ )

Moda: *Mo = 15*

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média: $\mu_X = 10,3$

Mediana: $Md(X) = 8$

Moda: $Mo(X) = 5$

| $X$ | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média: $\mu_X = 10,3 \longrightarrow \mu_Y = 5\mu_X - 10 = 41,5$

Mediana: $Md(X) = 8 \longrightarrow Md(\mathbf{Y}) = 5Md(\mathbf{X}) - 10 = 30$

Moda: $Mo(X) = 5 \longrightarrow Mo(\mathbf{Y}) = 5Mo(\mathbf{X}) - 10 = 15$

Fundação CECIERJ Consórcio cederj

| X | 2 | 5 | 8 | 15 | 20 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média: $\mu_X = 10,3 \longrightarrow \mu_Y = 5\mu_X - 10 = 41,5$

Mediana: $Md(X) = 8 \longrightarrow Md(\mathbf{Y}) = 5Md(\mathbf{X}) - 10 = 30$

Moda: $Mo(X) = 5 \longrightarrow Mo(\mathbf{Y}) = 5Mo(\mathbf{X}) - 10 = 15$

| Y | 0 | 15 | 30 | 65 | 90 |
|---|---|---|---|---|---|
| $p_i$ | 0,1 | 0,3 | 0,2 | 0,2 | 0,2 |

Média: $\mu_X = 41,5$

Mediana: $Md = 30$

Moda: $Mo = 15$

Medidas de tendência central: para um conjunto de dados e para variáveis aleatórias discretas

| | Conjunto de dados | Variável aleatória |
|---|---|---|
| valores | $\begin{array}{c\|cccc} X & x_1 & x_2 & \ldots & x_k \\ \hline n_i & n_1 & n_2 & \ldots & n_k \end{array}$ | $\begin{array}{c\|cccc} X & x_1 & x_2 & \ldots & x_k \\ \hline p_i & p_1 & p_2 & \ldots & p_k \end{array}$ |
| média | $\overline{x}_{obs} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$ | $\mu_X = \sum_{i=1}^{k} x_i p_i$ |
| mediana | $md_{obs}$: valor central | *Md*: $P(X \geq Md) \geq 0,5$ e $P(X \leq Md) \geq 0,5$ |
| moda | $mo_{obs}$: valor com maior freqüência | *Mo* : valor com maior probabilidade |

Medidas de tendência central: para um conjunto de dados e para variáveis aleatórias discretas

| | Conjunto de dados | Variável aleatória |
|---|---|---|
| valores | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}$ | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline p_i & p_1 & p_2 & \dots & p_k \end{array}$ |
| média | $\overline{x}_{obs} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$ | $\mu_X = \sum_{i=1}^{k} x_i p_i$ |
| mediana | $md_{obs}$ : valor central | *Md:* $P(X \geq Md) \geq 0,5$ e $P(X \leq Md) \geq 0,5$ |
| moda | $mo_{obs}$ : valor com maior freqüência | *Mo* : valor com maior probabilidade |

## Medidas de tendência central: para um conjunto de dados e para variáveis aleatórias discretas

| | Conjunto de dados | Variável aleatória |
|---|---|---|
| valores | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}$ | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline p_i & p_1 & p_2 & \dots & p_k \end{array}$ |
| média | $\overline{x}_{obs} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$ | $\mu_X = \sum_{i=1}^{k} x_i p_i$ |
| mediana | $md_{obs}$ : valor central | *Md:* $P(X \geq Md) \geq 0,5$ e $P(X \leq Md) \geq 0,5$ |
| moda | $mo_{obs}$ : valor com maior freqüência | *Mo* : valor com maior probabilidade |

Medidas de tendência central: para um conjunto de dados e para variáveis aleatórias discretas

| | Conjunto de dados | Variável aleatória |
|---|---|---|
| valores | $\begin{array}{c\|cccc} X & x_1 & x_2 & \ldots & x_k \\ \hline n_i & n_1 & n_2 & \ldots & n_k \end{array}$ | $\begin{array}{c\|cccc} X & x_1 & x_2 & \ldots & x_k \\ \hline p_i & p_1 & p_2 & \ldots & p_k \end{array}$ |
| média | $\bar{x}_{obs} = \frac{1}{n}\sum_{i=1}^{k} n_i x_i$ | $\mu_X = \sum_{i=1}^{k} x_i p_i$ |
| mediana | $md_{obs}$ : valor central | *Md:* $P(X \geq Md) \geq 0,5$ e $P(X \leq Md) \geq 0,5$ |
| moda | $mo_{obs}$ : valor com maior freqüência | *Mo* : valor com maior probabilidade |

# 6.2 Exemplo:

Uma pequena cirurgia dentária pode ser realizada por 3 diferentes técnicas. Foi feito um estudo para avaliar a recuperação, em dias (variável $X_i$ ), para cada uma das 3 técnicas, obtendo:

| $X_1$ | 0 | 4 | 5 | 6 | 10 |
|---|---|---|---|---|---|
| $p_i$ | 0,2 | 0,2 | 0,2 | 0,2 | 0,2 |

| $X_2$ | 1 | 5 | 9 |
|---|---|---|---|
| $p_i$ | 1/3 | 1/3 | 1/3 |

| $X_3$ | 4 | 5 | 6 |
|---|---|---|---|
| $p_i$ | 0,3 | 0,4 | 0,3 |

Encontrar as medidas de tendência central.

# 6.2 Exemplo:

Técnica 1

| $X_1$ | 0 | 4 | 5 | 6 | 10 |
|---|---|---|---|---|---|
| $p_i$ | 0,2 | 0,2 | 0,2 | 0,2 | 0,2 |

Média:

$$\mu = \sum_{i=1}^{5} x_i p_i = 0 \times 0,2 + 4 \times 0,2 + 5 \times 0,2 + 6 \times 0,2 + 10 \times 0,2$$

$$\mu = 5$$

Mediana: *Md* = 5

Moda: *Mo multimodal!*

# 6.2 Exemplo:

Técnica 2

| $X_2$ | 1 | 5 | 9 |
|---|---|---|---|
| $p_i$ | 1/3 | 1/3 | 1/3 |

Média:

$$\mu = \sum_{i=1}^{3} x_i p_i = 1 \times \frac{1}{3} + 5 \times \frac{1}{3} + 9 \times \frac{1}{3}$$

$$\mu = 5$$

Mediana: *Md* = 5

Moda: *Mo multimodal!*

Fundação CECIERJ Consórcio cederj

# 6.2 Exemplo:

Técnica 3

| $X_3$ | 4 | 5 | 6 |
|---|---|---|---|
| $p_i$ | 0,3 | 0,4 | 0,3 |

Média:

$$\mu = \sum_{i=1}^{3} x_i p_i = 4 \times 0,3 + 5 \times 0,4 + 6 \times 0,3$$

$$\mu = 5$$

Mediana: $Md = 5$

Moda: $Mo = 5$

# 6.2 Exemplo:

Técnica 1

| $X_1$ | 0 | 4 | 5 | 6 | 10 |
|---|---|---|---|---|---|
| $p_i$ | 0,2 | 0,2 | 0,2 | 0,2 | 0,2 |

Técnica 2

| $X_2$ | 1 | 5 | 9 |
|---|---|---|---|
| $p_i$ | 1/3 | 1/3 | 1/3 |

Técnica 3

| $X_3$ | 4 | 5 | 6 |
|---|---|---|---|
| $p_i$ | 0,3 | 0,4 | 0,3 |

| | Média | Mediana: | Moda: |
|---|---|---|---|
| Técnica 1 | $\mu = 5$ | *Md* = 5 | *Multimodal* |
| Técnica 2 | $\mu = 5$ | *Md* = 5 | *Multimodal* |
| Técnica 3 | $\mu = 5$ | *Md* = 5 | *Mo* = 5 |

# 6.3 Medidas de dispersão

## Variância de uma variável aleatória discreta $(Var(X) = \sigma^2 = \sigma_X{}^2)$

Seja ***X*** uma variável aleatória com $P(X_i = x_i) = p_i, i = 1, 2, ..., k$ e média μ. A variância de ***X*** é o somatório dos desvios, relativos à média, elevados ao quadrado e ponderados pela respectiva probabilidade

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i$$

# 6.3 Medidas de dispersão

## Variância de uma variável aleatória discreta $(Var(X) = \sigma^2 = \sigma_X{}^2)$

Seja ***X*** uma variável aleatória com $P(X_i = x_i) = p_i \; i = 1, 2, ..., k$ e média μ. A variância de ***X*** é o somatório dos desvios, relativos à média, elevados ao quadrado e ponderados pela respectiva probabilidade

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 \, p_i$$

Que pode também ser definida como o valor esperado do desvio ao quadrado:

$$Var(X) = \sigma^2 = E[(X - \mu)^2]$$

# 6.3 Medidas de dispersão

## Variância de uma variável aleatória discreta $(Var(X) = \sigma^2 = \sigma_X{}^2)$

Seja ***X*** uma variável aleatória com $P(X_i = x_i) = p_i\, i = 1, 2, ..., k$ e média μ. A variância de ***X*** é o somatório dos desvios, relativos à média, elevados ao quadrado e ponderados pela respectiva probabilidade

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i$$

Que pode também ser definida como o valor esperado do desvio ao quadrado:

$$Var(X) = \sigma^2 = E[(X - \mu)^2]$$

Ou... (fica para você mostrar...):

$$Var(X) = \sigma^2 = E(X^2) - \mu^2 = \sum_{i=1}^{k} p_i x_i{}^2 - \mu^2$$

# Desvio padrão de uma variável aleatória discreta

Definido com a raiz quadrada da variância.

$$\sigma = \sqrt{\sum_{i=1}^{k} (x_i - \mu)^2 p_i}$$

ou

$$\sigma = \sqrt{E[(X - \mu)^2]}$$

## Voltando ao exemplo....

Uma pequena cirurgia dentária pode ser realizada por 3 diferentes técnicas. Foi feito um estudo para avaliar a recuperação, em dias (variável $X_i$ ), para cada uma das 3 técnicas, obtendo:

| $X_1$ | 0 | 4 | 5 | 6 | 10 |
|---|---|---|---|---|---|
| $p_i$ | 0,2 | 0,2 | 0,2 | 0,2 | 0,2 |

| $X_2$ | 1 | 5 | 9 |
|---|---|---|---|
| $p_i$ | 1/3 | 1/3 | 1/3 |

| $X_3$ | 4 | 5 | 6 |
|---|---|---|---|
| $p_i$ | 0,3 | 0,4 | 0,3 |

Encontrar as medidas de dispersão

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i \quad \text{com } \mu = 5$$

Técnica 1

$$\sigma^2 = (0-5)^2 \times 0,2 + (4-5)^2 \times 0,2 + (5-5)^2 \times 0,2 + (6-5)^2 \times 0,2 + (10-5)^2 \times 0,2$$

$$\boxed{\sigma^2 = 10,40}$$

$$Var(X) = \sum_{i=1}^{k} (x_i - \mu)^2 p_i \quad \text{com } \mu = 5$$

Técnica 1

$$\sigma^2 = (0-5)^2 \times 0,2 + (4-5)^2 \times 0,2 + (5-5)^2 \times 0,2 + (6-5)^2 \times 0,2 + (10-5)^2 \times 0,2$$

$$\boxed{\sigma^2 = 10,40}$$

Técnica 2

$$\sigma^2 = (1-5)^2 \times \frac{1}{3} + (5-5)^2 \times \frac{1}{3} + (9-5)^2 \times \frac{1}{3}$$

$$\boxed{\sigma^2 = 10,67}$$

$$Var(X)=\sum_{i=1}^{k}(x_i-\mu)^2 p_i \quad \text{com } \mu = 5$$

## Técnica 1

$$\sigma^2=(0-5)^2\times 0,2+(4-5)^2\times 0,2+(5-5)^2\times 0,2+(6-5)^2\times 0,2+(10-5)^2\times 0,2$$

$$\boxed{\sigma^2=10,40}$$

## Técnica 2

$$\sigma^2=(1-5)^2\times\frac{1}{3}+(5-5)^2\times\frac{1}{3}+(9-5)^2\times\frac{1}{3}$$

$$\boxed{\sigma^2=10,67}$$

## Técnica 3

$$\sigma^2=(4-5)^2\times 0,3+(5-5)^2\times 0,4+(6-5)^2\times 0,3$$

$$\boxed{\sigma^2=0,60}$$

# Mais um exemplo

Cacular o valor esperado (média) e a variância da distribuição de Bernoulli

| $X$ | 0 | 1 |
|---|---|---|
| $p_i$ | 1-p | p |

## Mais um exemplo

Cacular o valor esperado (média) e a variância da distribuição de Bernoulli

| $X$ | 0 | 1 |
|---|---|---|
| $p_i$ | 1-p | p |

A média (E(**X**) ou $\mu$)

$$\mu = 0 \times (1 - p) + 1 \times p = p$$

# Mais um exemplo

Cacular o valor esperado (média) e a variância da distribuição de Bernoulli

| $X$ | 0 | 1 |
|---|---|---|
| $p_i$ | 1-p | p |

A média (E(**X**) ou $\mu$)

$$\mu = 0 \times (1 - p) + 1 \times p = p$$

E a variância...

$$\sigma^2 = (0-p)^2 \times (1-p) + (1-p)^2 \times p$$

## Mais um exemplo

Cacular o valor esperado (média) e a variância da distribuição de Bernoulli

| $X$ | 0 | 1 |
|---|---|---|
| $p_i$ | $1-p$ | $p$ |

A média (E(**X**) ou $\mu$)

$$\mu = 0 \text{ x } (1 - p) + 1 \text{ x } p = p$$

E a variância...

$$\sigma^2 = (0\text{-}p)^2 \text{ x } (1\text{-}p) + (1\text{-}p)^2 \text{ x } p$$
$$\sigma^2 = p^2 - p^3 + (1 - 2p + p^2) \text{ x } p$$

# Mais um exemplo

Cacular o valor esperado (média) e a variância da distribuição de Bernoulli

| $X$ | $0$ | $1$ |
|---|---|---|
| $p_i$ | $1-p$ | $p$ |

A média (E(**X**) ou $\mu$)

$$\mu = 0 \text{ x } (1 - p) + 1 \text{ x } p = p$$

E a variância...

$$\sigma^2 = (0-p)^2 \text{ x } (1-p) + (1-p)^2 \text{ x } p$$
$$\sigma^2 = p^2 - p^3 + (1 - 2p + p^2) \text{ x } p$$
$$\sigma^2 = p^2 - p^3 + p - 2p^2 + p^3$$

## Mais um exemplo

Cacular o valor esperado (média) e a variância da distri-buição de Bernoulli

| $X$ | 0 | 1 |
|---|---|---|
| $p_i$ | 1-p | p |

A média (E(**X**) ou $\mu$)

$$\mu = 0 \times (1 - p) + 1 \times p = p$$

E a variância...

$$\sigma^2 = (0-p)^2 \times (1-p) + (1-p)^2 \times p$$
$$\sigma^2 = p^2 - p^3 + (1 - 2p + p^2) \times p$$
$$\sigma^2 = p^2 - p^3 + p - 2p^2 + p^3$$
$$\sigma^2 = p - p^2 = p(1-p)$$

# Valor esperado e variância de modelos discretos

| Variável discreta | Valor esperado | Variância |
|---|---|---|
| Uniforme (1,k) | $\frac{1+k}{2}$ | $\frac{k^2-1}{12}$ |
| Bernoulli (p) | $p$ | $p(1-p)$ |
| Binomial (n,p) | $np$ | $np(1-p)$ |
| Geométrico (p) | $\frac{1-p}{p}$ | $\frac{1-p}{p^2}$ |
| Poisson ($\lambda$) | $\lambda$ | $\lambda$ |
| Hipergeométrica (n,m,r) | $\frac{rm}{n}$ | $\frac{rm(n-m)(n-r)}{n^2(n-1)}$ |

# Medidas de dispersão: para um conjunto de dados e para variáveis aleatórias discretas

| | Conjunto de dados | Variável aleatória |
|---|---|---|
| Valores | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline n_i & n_1 & n_2 & \dots & n_k \end{array}$ | $\begin{array}{c\|cccc} X & x_1 & x_2 & \dots & x_k \\ \hline p_i & p_1 & p_2 & \dots & p_k \end{array}$ |
| Variância | $var_{obs} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_{obs})^2$ | $Var(X) = \sum_{i=1}^{k}(x_i - \mu)^2 p_i$ |
| Variância (outra opção) | $var_{obs} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}_{obs}^2$ | $Var(X) = \sum_{i=1}^{k} p_i x_i^2 - \mu^2$ |

# Lembrando...vimos na aula 6: "Medidas de resumo (para variáveis aleatórias discretas)"

Medidas de posição (tendência central)
- média
- mediana
- moda

Medidas de dispersão
- variância
- desvio padrão

Pergunta: e se as variáveis que temos interesse assumem valores aleatórios mas pertencem ao algum intervalo de números reais (variáveis contínuas)? → Aulas 8 e 9

# Aula 6

Professores:

*Otton Teixeira da Silveira Filho*
*Regina Célia Paula Leal Toledo*

# Medidas de resumo (para variáveis aleatórias discretas)

Conteúdo:

6.1 Medidas de posição
- média
- mediana
- moda

6.2 Exemplo

6.3 Medidas de dispersão
- variância
- desvio padrão

$$\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\sigma_X^2 \qquad \bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$A \cap B = \varnothing$$