

ANÁLISE DE DADOS

MEDIDAS DE DISPERSÃO

Prof. Dr. Daniel Caetano

2020 - 1

Objetivos

- Compreender o que é uma medida de dispersão
- Compreender e calcular a variância e o desvio padrão
- Compreender e calcular o coeficiente de variação de Pearson

- **Atividade da Aula 4 no SAVA!**



Material de Estudo



Material

Acesso ao Material

Apresentação

<http://www.caetano.eng.br/>
(Análise de Dados – Aula 4)

Material Didático

- Probab. e Estatística Aplicada à Engenharia –
Cap. 4

Minha Biblioteca

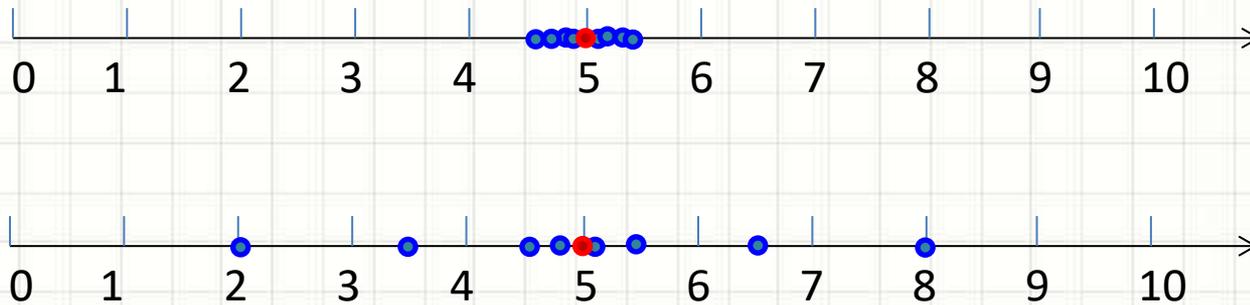
Estatística – Teoria e Aplicações usando MS Excel –
Cap. 3
Estatística Aplicada e Probabilidade para Engenheiros
– Cap. 3, 4 e 7



MEDIDAS DE DISPERSÃO

Medidas de Dispersão

- O que é dispersão?



- Essas amostras são iguais?
 - Mas a média de ambas é 5!
- O que as difere?

Medidas de Dispersão

- Medidas de Dispersão
 - Trazem informação sobre o “espalhamento” de uma amostra
- Medidas mais comuns
 - Variância
 - Desvio Padrão
 - Coeficiente de Dispersão



VARIÂNCIA

Variância de uma Amostra

- Calculada por:

$$Var[x] = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$

- Exemplo:

– Amostra: 2, 4, 6, 6, 8, 10

– Média: $36/6 = 6$

$$Var[x] = \frac{(2 - 6)^2 + (4 - 6)^2 + (6 - 6)^2 + (6 - 6)^2 + (8 - 6)^2 + (10 - 6)^2}{6 - 1}$$

$$Var[x] = \frac{16 + 4 + 0 + 0 + 4 + 16}{5} = \frac{40}{5} = \mathbf{8}$$

Variância de uma População

- Calculada por:

$$Var[x] = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

- Ou, quando agrupado por frequências relativas

$$Var[x] = (x_1 - \bar{x})^2 \cdot fr_1 + (x_2 - \bar{x})^2 \cdot fr_2 + \dots + (x_n - \bar{x})^2 \cdot fr_n$$

Variância de uma População

- Exemplo

Número de Filhos	Frequência Relativa
0	20%
1	45%
2	20%
3	10%
4	5%

- Média?

$$\bar{x} = 0,2 \cdot 0 + 0,45 \cdot 1 + 0,2 \cdot 2 + 0,1 \cdot 3 + 0,05 \cdot 4$$

$$\bar{x} = 0 + 0,45 + 0,4 + 0,3 + 0,2 \quad \bar{x} = \mathbf{1,35}$$

- Variância?

$$Var[x] = (0 - 1,35)^2 \cdot 0,2 + (1 - 1,35)^2 \cdot 0,45 + (2 - 1,35)^2 \cdot 0,2 + (3 - 1,35)^2 \cdot 0,1 + (4 - 1,35)^2 \cdot 0,05$$

$$Var[x] = 0,3645 + 0,055125 + 0,0845 + 0,27225 + 0,351125 = \mathbf{1,1275}$$

Variância

- 4 Propriedades

1. A variância de “variável aleatória” constante será sempre 0.
2. Ao multiplicar valores de uma variável aleatória por uma constante **C**, sua variância fica multiplicada por **C²**.
3. A soma de uma constante **C** à variável aleatória não altera sua variância.
4. A variância da soma ou subtração de duas variáveis aleatórias é igual à soma ou à subtração das respectivas variâncias.



DESVIO PADRÃO

Desvio Padrão

- Variância: “erro quadrático médio”
- Desvio padrão: “erro médio”
 - Como?

$$\sigma_x = \sqrt{Var[x]}$$

- Para populações, usa-se σ
- Para amostras, usa-se S

Desvio Padrão

- Exemplo:

- Amostra: 2, 4, 6, 6, 8, 10

- Média: $36/6 = 6$

- $Var[x] = 8$

$$S_x = \sqrt{Var[x]}$$

$$S_x = \sqrt{8}$$

$$S_x = \mathbf{2,83}$$

- Outro Exemplo:

- Média: 1,35

- Variância: 1,1275

$$\sigma_x = \sqrt{Var[x]}$$

$$\sigma_x = \sqrt{1,1275}$$

$$\sigma_x = \mathbf{1,06}$$

Número de Filhos	Frequência Relativa
0	20%
1	45%
2	20%
3	10%
4	5%

Desvio Padrão

- 3 Propriedades
 1. O desvio padrão de “variável aleatória” constante será sempre 0.
 2. Ao multiplicar valores de uma variável aleatória por uma constante **C**, seu desvio padrão fica multiplicado pela mesma **C**.
 3. A soma de uma constante **C** à variável aleatória não altera seu desvio padrão.

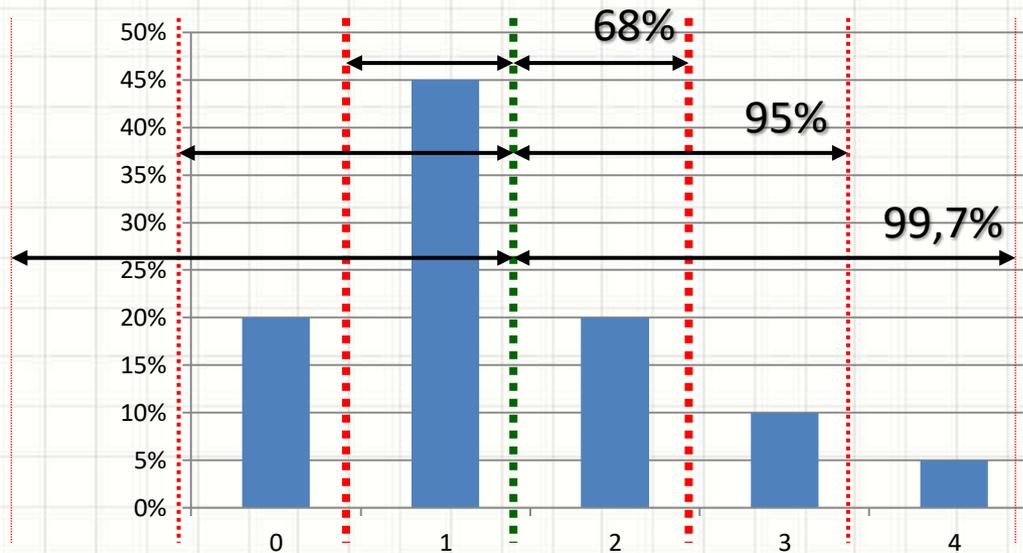
Desvio Padrão

- Interpretação

Número de Filhos	Frequência Relativa
0	20%
1	45%
2	20%
3	10%
4	5%

$$\bar{x} = 1,35$$

$$\sigma_x = 1,06$$





COEFICIENTE DE VARIAÇÃO

Coeficiente de Variação

- Dá uma noção relativa do desvio padrão

$$CV = \frac{\sigma_x}{\bar{x}} \cdot 100\%$$

- Permite classificar a dispersão
 - Baixa dispersão: $0\% \leq CV \leq 15\%$
 - Média dispersão: $15\% \leq CV \leq 30\%$
 - Alta dispersão: $CV > 30\%$

Coeficiente de Variação

- Exemplo, amostra: 2, 4, 6, 6, 8, 10

– Média: $36/6 = 6$

– $\text{Var}[x] = 8$

– $S_x = 2,83$

$$CV = \frac{\sigma_x}{\bar{x}} \cdot 100\% \quad CV = \frac{2,83}{6} \cdot 100\%$$

$$CV = 47,2\% \quad \text{Alta!}$$

- Outro Exemplo:

– Média: 1,35

– $\text{Var}[x]: 1,1275$

– $\sigma_x = 1,06$

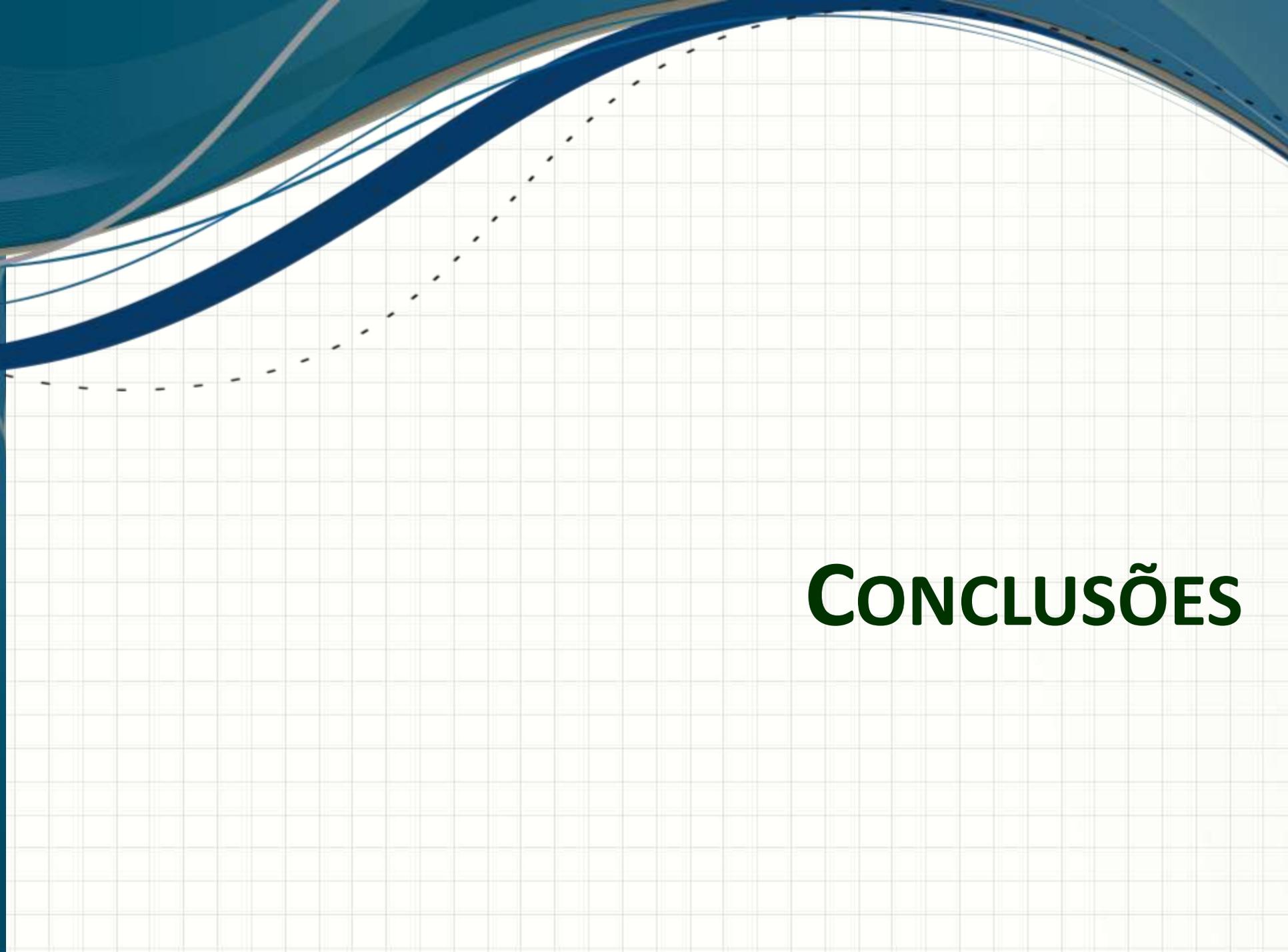
$$CV = \frac{\sigma_x}{\bar{x}} \cdot 100\%$$

$$CV = \frac{1,06}{1,35} \cdot 100\%$$

$$CV = 78,5\%$$

Alta!

Número de Filhos	Frequência Relativa
0	20%
1	45%
2	20%
3	10%
4	5%

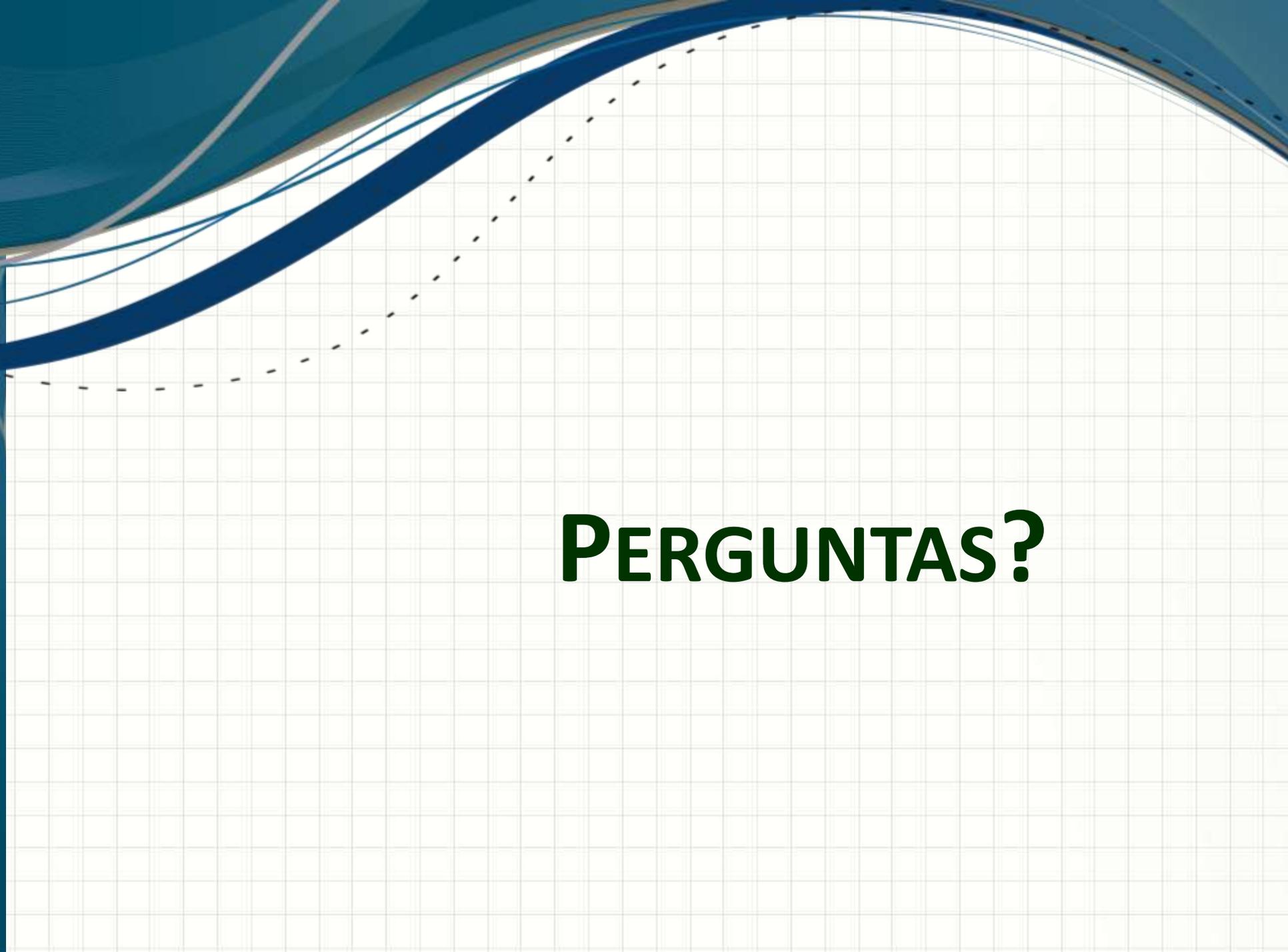


CONCLUSÕES

Resumo

- O que são medidas de dispersão
 - Variância e Desvio Padrão
 - Coeficiente de Variação
-

- Probabilidade...
 - O que é?
 - Espaço amostral?
 - Evento aleatório?



PERGUNTAS?



EXERCÍCIOS