

# MAT02018 - Estatística Descritiva

## Conceitos Básicos

Rodrigo Citton P. dos Reis  
citton.padilha@ufrgs.br

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DEPARTAMENTO DE ESTATÍSTICA

Porto Alegre, 2020



# Introdução

# Introdução

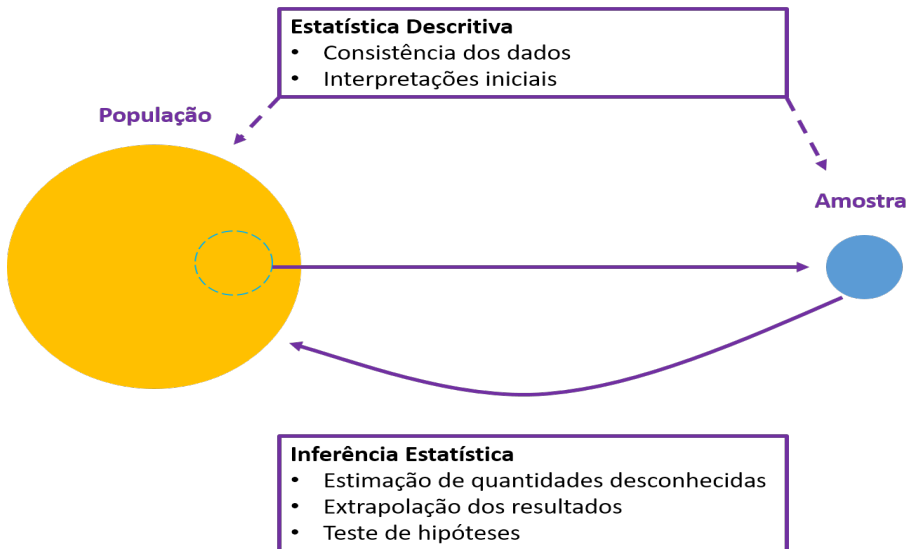
- ▶ Em diversas situações, perguntas são feitas com respeito a algum fenômeno de interesse.

# Conceitos básicos

# O que é a estatística?

- ▶ A **estatística** é a ciência que tem por objetivo orientar a coleta, o resumo, a apresentação, a análise e a interpretação de dados.
  - ▶ Do grego *statistós*, de *statízo*, “estabelecer”, “verificar”, acrescido do sufixo *ica*.
- ▶ Podem ser identificadas duas grandes áreas de atuação desta ciência:
  - ▶ a **estatística descritiva**, envolvida com o resumo e a apresentação dos dados.
  - ▶ a **estatística inferencial**, que ajuda a concluir sobre conjuntos maiores de dados (populações) quando apenas partes desses conjuntos (as amostras) foram estudadas.

# O que é a estatística?



# O que é a estatística?

- ▶ A **Estatística Descritiva** corresponde aos procedimentos relacionados com a coleta, elaboração, tabulação, análise, interpretação e apresentação dos dados.
- ▶ Isto é, inclui as técnicas que dizem respeito à sintetização e à descrição de dados numéricos.
- ▶ Tais métodos tanto podem ser gráficos como envolver análise computacional.

# Unidades experimentais e observacionais

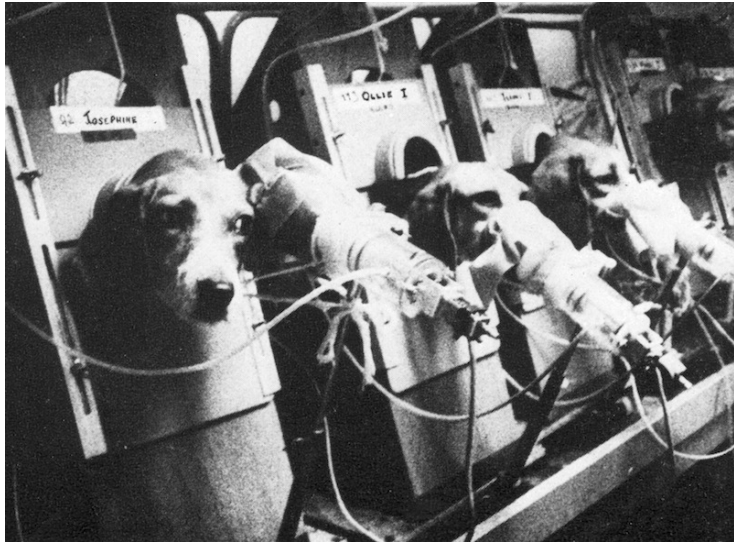
- ▶ **Unidade experimental** ou **unidade de observação** é a menor unidade a fornecer informação.
  - ▶ **Ex:** alunos, pacientes, animais, plantas, carros, hospitais, escolas, cidades, universidades, países, *tweets*, etc.



# Experimento



# Experimento (?)



# Dados e variáveis

## Dados

São as informações obtidas de uma unidade experimental ou observacional.

- ▶ **Ex:** “Vitor tem 25 anos e é fumante”. Os dados são “25 anos” e “fumante”.

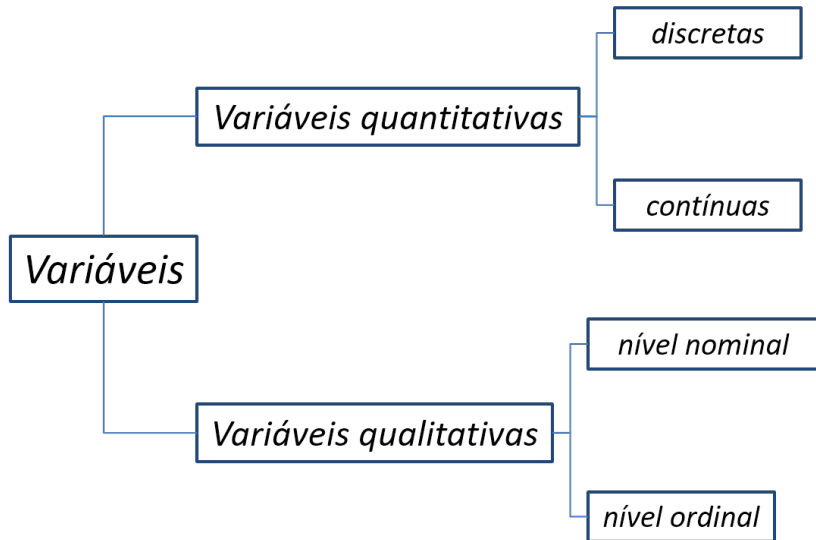
## Variável

É toda característica que, observada em uma unidade (experimental ou observacional), pode variar de um indivíduo para outro.

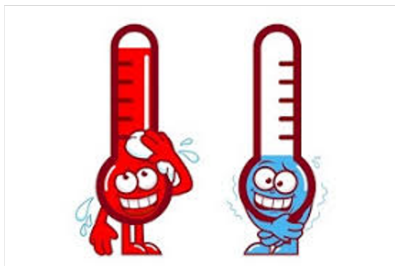
- ▶ **Ex:** idade, sexo, altura, nível de hemoglobina no sangue, espaçamento entre plantas, doses de um medicamento, tipo de medicamento, cultivares, número de caracteres, velocidade da rede, tempo gasto na rede social, nível de monóxido de carbono em emissões do escape de automóveis, etc.

É importante **identificar que tipo de variável** está sendo estudada, uma vez que são recomendados **procedimentos estatísticos diferentes** em cada situação.

# Tipos de variáveis



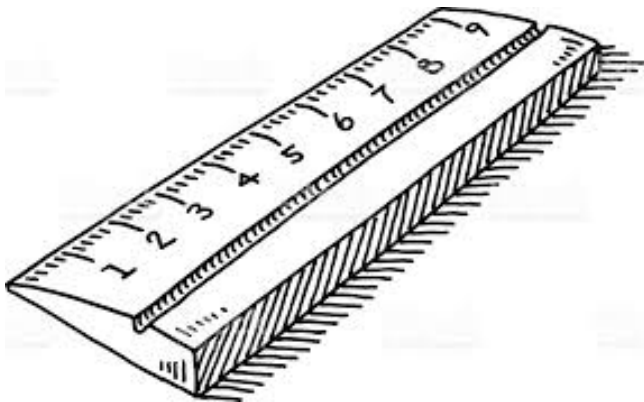
# Variáveis quantitativas



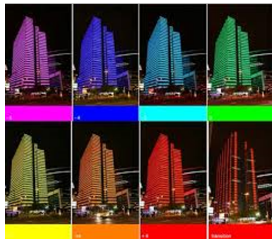
# Variáveis quantitativas discretas



# Variáveis quantitativas contínuas



# Variáveis qualitativas





# Variáveis qualitativas ordinais



# Variáveis qualitativas nominais



# Exemplos (1)



# Exemplos (1)

## Variáveis quantitativas

- ▶ 3 andares
- ▶ 14,85 metros de altura

## Variáveis qualitativas

- ▶ Multicolorido
- ▶ Cheira “bem”

## Exemplos (2)



# Exemplos (2)

## Variáveis quantitativas

- ▶ 4 integrantes
- ▶ 56 anos

## Variáveis qualitativas

- ▶ Inglaterra
- ▶ Rock

# População

- ▶ **População** ou **universo**: esse termo é usado em estatística com um sentido bem mais amplo do que na linguagem coloquial.
  - ▶ É entendido aqui como **todo conjunto de unidades experimentais** (ou observacionais) que apresenta uma ou mais características **em comum**.
- ▶ **Ex**: a população de colegiais de oito anos de Belo Horizonte.
  - ▶ Estes colegiais têm em comum a idade e o local onde vivem.

# Censo e amostra

- ▶ Quando o estudo é realizado com toda a população de interesse, chamemos este estudo de **censo**.
- ▶ Por motivos de tempo, custo, logística, entre outros, geralmente não é possível realizar um censo.
  - ▶ Nestes casos, estudamos apenas uma parcela da população, que chamamos de **amostra**.



# Amostra

- ▶ **Amostra** é qualquer fração de uma população.
  - ▶ Como sua finalidade é representar a população, deseja-se que a amostra escolhida apresente as mesmas características da população de origem, isto é, que seja uma amostra “**representativa**” ou “**não-tendenciosa**”.
- ▶ Tanto o número de indivíduos selecionados para a amostra quanto a técnica de seleção são extremamente importantes para que os resultados obtidos no estudo sejam generalizados para a população.

# Amostra representativa



- ▶ Ver a discussão sobre **representatividade da amostra** na apresentação do **Prof. Chris Fonnesebeck**.

# Amostragem

- ▶ A seleção da amostra pode ser feita de várias maneiras.
- ▶ Esta dependerá:
  - ▶ Do grau de conhecimento que temos da população.
  - ▶ Da quantidade de recursos disponíveis.
- ▶ A seleção da amostra tenta fornecer um subconjunto de valores o **mais parecido possível** com a população que lhe dá origem.
  - ▶ **Amostra representativa** da população.

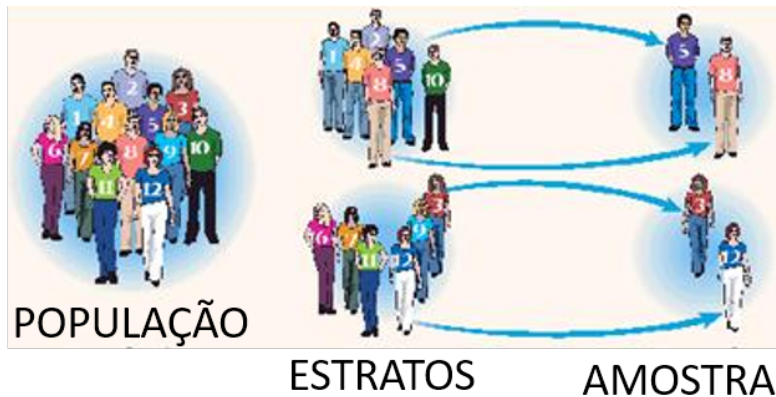
# Amostra aleatória simples

- ▶ A amostragem mais usada é a **amostra casual simples** (ou aleatória simples).
  - ▶ Os indivíduos (unidades) da amostra são selecionados ao acaso, **com** ou **sem reposição**.



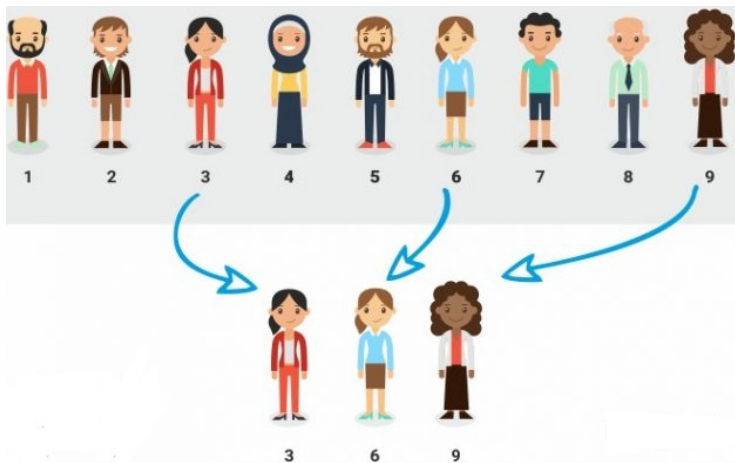
# Amostra estratificada

- ▶ Eventualmente, se tivermos informações adicionais a respeito da população de interesse, podemos utilizar outros esquemas de amostragem mais sofisticados.
  - ▶ **Amostragem estratificada**



# Amostra sistemática

- Em outros casos, pode existir uma relação numerada dos itens da população que nos permitiria utilizar a chamada **amostragem sistemática** em que selecionamos os indivíduos de forma pré-determinada.



# Amostragem

- ▶ Outros esquemas de amostragem poderiam ser citados e todos fazem parte da chamada **teoria da amostragem**, cujos detalhes não serão aprofundados.

# Parâmetros, estatísticas e estimativas

- ▶ **Parâmetro** é um valor que resume, na população, a informação relativa a uma variável.
  - ▶ **Ex:** média populacional, prevalência populacional, coeficiente de variação populacional, taxa de mortalidade populacional, etc.
- ▶ **Estatística** (além de ser o nome da ciência/área do conhecimento) é a denominação dada a uma quantidade, calculada com base nos elementos de uma amostra, que descreve a informação contida nesse conjunto de dados.
  - ▶ **Ex:** A média, a porcentagem, o desvio padrão, o coeficiente de correlação, calculados em uma amostra, são estatísticas.



# Parâmetros, estatísticas e estimativas

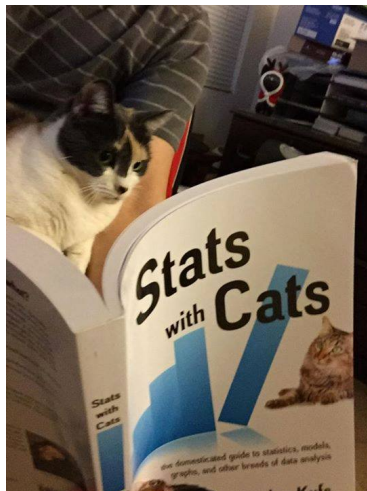
- ▶ Os parâmetros são difíceis de se obter, pois implicam o estudo de toda a população e costumam ser substituídos por valores calculados em amostras representativas da população-alvo.
  - ▶ Se tivesse sido examinada uma amostra de 50 estudantes matriculados na disciplina MAT02219, e 40% fossem do sexo masculino, esse valor constituiria uma estimativa do parâmetro “percentagem de homens matriculados naquela disciplina”.
- ▶ A **estimativa** é um valor numérico de uma estatística, usado para realizar inferências sobre o parâmetro.
  - ▶ Da mesma forma, o valor numérico da média para a estatura desses 50 alunos, digamos 173 cm, é uma estimativa para a média de altura populacional.

# Próxima aula

- ▶ Organização dos dados

## Para casa

- Liste exemplos de variáveis que você utiliza (ou conhece) e classifique-as de acordo com os tipos de variáveis que foram apresentados em aula.



# Por hoje é só! Bons estudos!

