

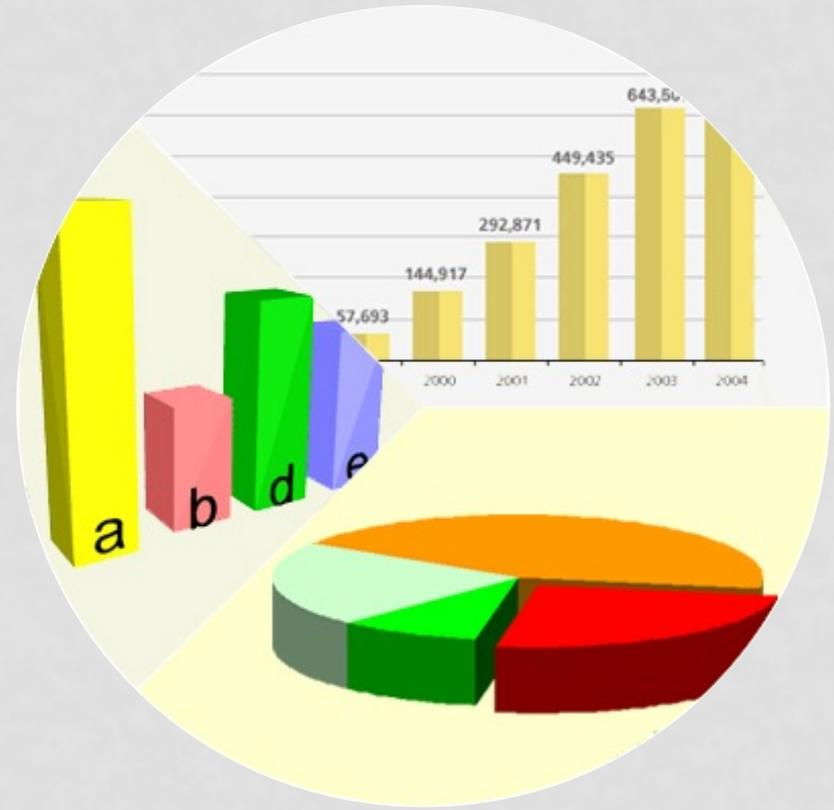
VARIÁVEIS E PROCEDIMENTOS ESTATÍSTICOS

THAIS CRISTINA CHAVES



OBJETIVOS DA AULA

- Conceitos Básicos de Estatística
- População e Amostragem
- Natureza da Variável
- Estatística Descritiva
- Distribuição Normal
- Estatística Inferencial
 - Paramétrica
 - Não Paramétrica
- Interpretação dos Resultados



CONCEITOS

- O que é estatística?
 - Obter conclusões a partir de dados observados
 - Explora a Coleta, organização, apresentação, análise e interpretação dos dados
 - Bioestatística → Estatística em Ciências Biológicas e da Saúde
- Pra que serve?
 - Para afirmar a viabilidade da análise ou das comparações que deseja fazer
 - Para dar credibilidade aos dados encontrados
 - Para dar veracidade às conclusões da pesquisa

População

- Conjunto referência em uma investigação
- Apresentam características em comum que se pretenda estudar
- Pode ser finita ou infinita, concreta ou abstrata.

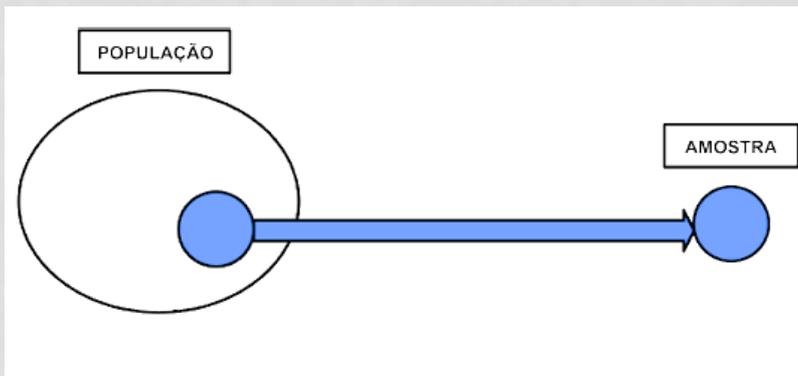


EXEMPLOS:

- **Alunos da USP de Ribeirão Preto**
- **Cadeiras pertencentes a FMRP**
- **Mulheres com osteopenia**
- **Pacientes com distrofia muscular de Duchenne**

AMOSTRA

- Subconjunto representativo de uma população
- Deve ter as mesmas características da população
- Maior a amostra maior → Representar a população
- Finita
- Economia e rapidez dos resultados



EXEMPLO:

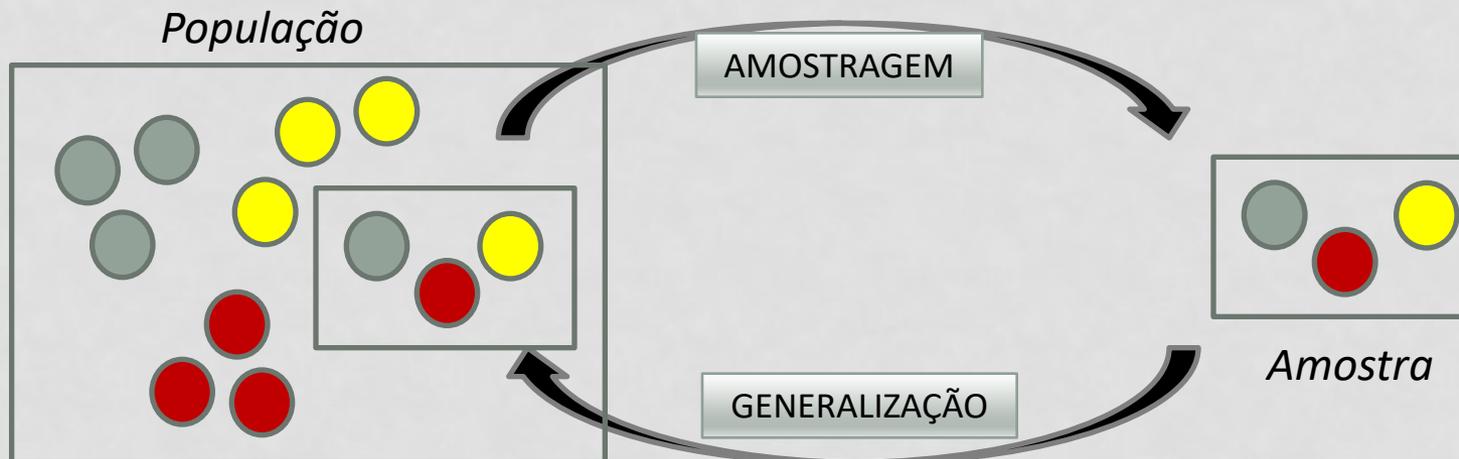
Alunos USP Ribeirão Preto (População)



Alunos Terapia Ocupacional e Fisioterapia (Amostra)

AMOSTRAGEM

- Técnica para escolher a amostra
- Garantia de que qualquer elemento da população tenha a mesma probabilidade de ser escolhido
- Boa amostragem → representatividade da amostra



AMOSTRAGEM

- Probabilística

Aleatória (casual)

Proporcional estratificada

⋮

- Não probabilística

De conveniência (a esmo)

Intencional

⋮

Amostragem Probabilística

- Todos os elementos da população tem probabilidade de serem escolhidos para compor a amostra.
- **Aleatória (casual, randômica)**
 - Cada elemento tem a MESMA probabilidade de ser escolhido (Ex: sorteio)
- **Proporcional estratificada**
 - A população é dividida em estratos e a amostragem é proporcional a ela

EXEMPLO:

AMOSTRAGEM PROBABILÍSTICA ALEATÓRIA:



POPULAÇÃO: estudantes de medicina FMRP

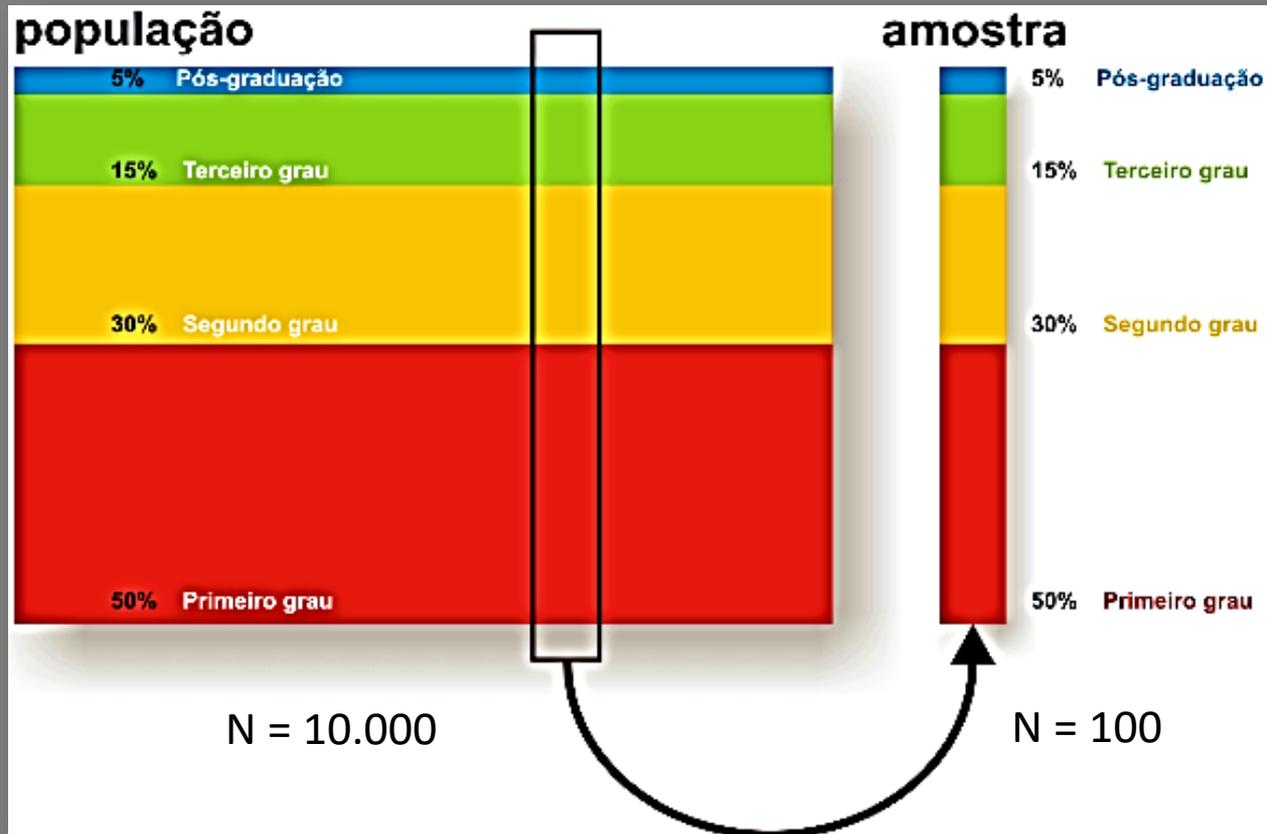
- Relação dos elementos da população
- Enumerá-los de 1 até N (tamanho da população)
- Determinar o número de elementos sorteados
- Selecionar os elementos por mecanismo aleatório

Em uma população composta por 50% de mulheres e 50% de homens, a amostragem probabilística aleatória formou uma amostra com 65% de mulheres e 35% de homens

A amostra continua aleatória, mas não é representativa

EXEMPLO:

AMOSTRAGEM PROBABILÍSTICA ESTRATIFICADA



Tipos de Variáveis

Quantitativas

Discretas
Contínuas

Qualitativas

Ordinais
Nominais

QUANTITATIVAS

- Resultantes de uma contagem ou mensuração
- Valores numéricos
- **Variáveis discretas**
 - Valores inteiros. Geralmente são resultado de contagem
 - Ex: número de filhos, número de pacientes
- **Variáveis contínuas**
 - Valores em uma escala contínua. Mensurados por equipamentos
 - Ex: peso, altura, tempo, pressão arterial

QUALITATIVAS

Características que não possuem valores quantitativos

Variáveis ordinais

- Ordem, uma hierarquia entre as categorias
- Ex: escolaridade (1° , 2° , 3° graus), estágio da doença (inicial, intermediário, terminal)

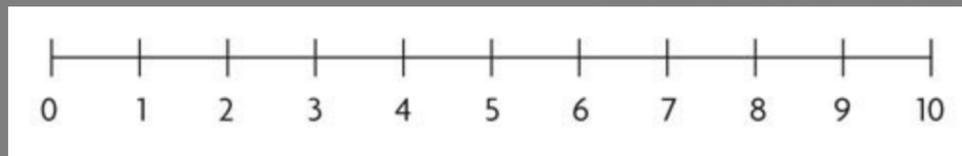
Variáveis nominais

- Identificar as categorias, sem ordenamento nem hierarquia
- Ex: sexo, cor dos olhos, fumante/não fumante, doente/sadio

EXEMPLO:

COMO SÃO CLASSIFICADAS ESSAS VARIÁVEIS?

IMC	Classificação
<18,5	Abaixo do peso ideal
18,5 - 25	Peso normal
25 - 30	Excesso de peso
30 - 35	Obesidade (grau I)
35 - 40	Obesidade (grau II)
> 40	Obesidade (grau III)



Sem dor

*Pior dor
possível*

EXEMPLO:

E AS MEDIDAS OBTIDAS POR GONIOMETRIA?



EXEMPLO:

E NA ESCALA DOS TESTES MANUAIS DE FUNÇÃO MUSCULAR?

Descrição da força muscular	Graduação do teste
Não se sente contração nenhuma	0 (Nula)
Pode-se sentir o enrijecimento muscular, mas não se consegue produzir o movimento	1 (Traço)
Produz movimento a favor da gravidade	2 (Fraco)
Produz movimento contra a gravidade	3 (Regular)
Produz movimento contra resistência externa e contra a gravidade	4 (Bom)
É capaz de superar a resistência aplicada	5 (Normal)

QUIZ !

1. Amostra de conveniência é aquela escolhida de forma intencional pelo pesquisador ()
2. Amostra probabilística aleatória é aquela em que todos os elementos da população não têm a mesma chance de serem escolhidos ()
3. Variáveis contínuas são aquelas que assumem valores inteiros ()
4. Variáveis nominais são aquelas que identificam categorias, com ordem hierárquica ()

QUIZ !

1. Amostra de conveniência é aquela escolhida de forma intencional pelo pesquisador (V)
2. Amostra probabilística aleatória é aquela em que todos os elementos da população não têm a mesma chance de serem escolhidos (**F**)
3. Variáveis contínuas são aquelas que assumem valores inteiros (**F**)
4. Variáveis nominais são aquelas que identificam categorias, com ordem hierárquica (V)

**COMO ANALISAR ESTATISTICAMENTE
AS VARIÁVEIS?**

Ramos da estatística

- **Estatística descritiva ou dedutiva**
 - Métodos para descrever, organizar e RESUMIR os dados
- **Estatística inferencial ou indutiva**
 - Visa inferir sobre características de uma POPULAÇÃO a partir de uma parte dela (AMOSTRA)
 - Utiliza-se a probabilidade, pois existe a INCERTEZA do resultado.

Estadística descritiva

Medidas que buscam sumarizar as informações sobre um conjunto de dados

Medidas de tendência central

Média

Mediana

Moda

Medidas de dispersão

Desvio médio

Variância

Desvio Padrão

⋮

MEDIDAS DE TENDÊNCIA CENTRAL

MÉDIA (μ, x)

MÉDIA
ARITMÉTICA DE
UM CONJUNTO
DE DADOS

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

SOMA
DOS
VALORES

QUANTIDADE
DE VALORES

Média

- Indicada para distribuições simétricas
- Medida descritiva mais utilizada
- Utiliza todos os pontos de dados
- **Fortemente influenciada por valores extremos (outliers)**
- **Em distribuições assimétricas pode dar uma informação distorcida**

EXEMPLO:

CALCULAR A MÉDIA DA IDADE DOS IDOSOS INTERNADOS POR QUEDAS NO MÊS DE SETEMBRO NA CIDADE DE RIBEIRÃO PRETO:

65

70

84

90

67

76

81

73

72

82

$$\bar{X} = 760/10 = 76 \text{ anos}$$

MEDIANA

- Valor que ocupa a posição central de um conjunto de valores dispostos em ordem crescente
- Indica que 50% dos valores estão abaixo e 50% estão acima da mediana.
- Apropriada em distribuições assimétricas
- **Não é afetada** por eventuais valores discrepantes (**outliers**) existentes no conjunto

EXEMPLO:

CALCULAR A MEDIANA DA IDADE DOS IDOSOS INTERNADOS POR QUEDAS NO MÊS DE SETEMBRO NO ESTADO DE SÃO PAULO:

65

70

84

90

67

76

81

73

72

1° Passo: ordenar os dados em ordem crescente

65 67 70 72 73 76 81 84 90

2° Passo: identificar o valor central (número do meio)

65 67 70 72 **73** 76 81 84 90

E se o número total de valores for par?

65 67 70 72 **73 76** 81 82 84 90

Mediana = $73 + 76 / 2 = 74,5$ anos

OUTLIER

MEAN

MEDIAN

MODA

OBSERVAÇÃO MAIS FREQUENTE

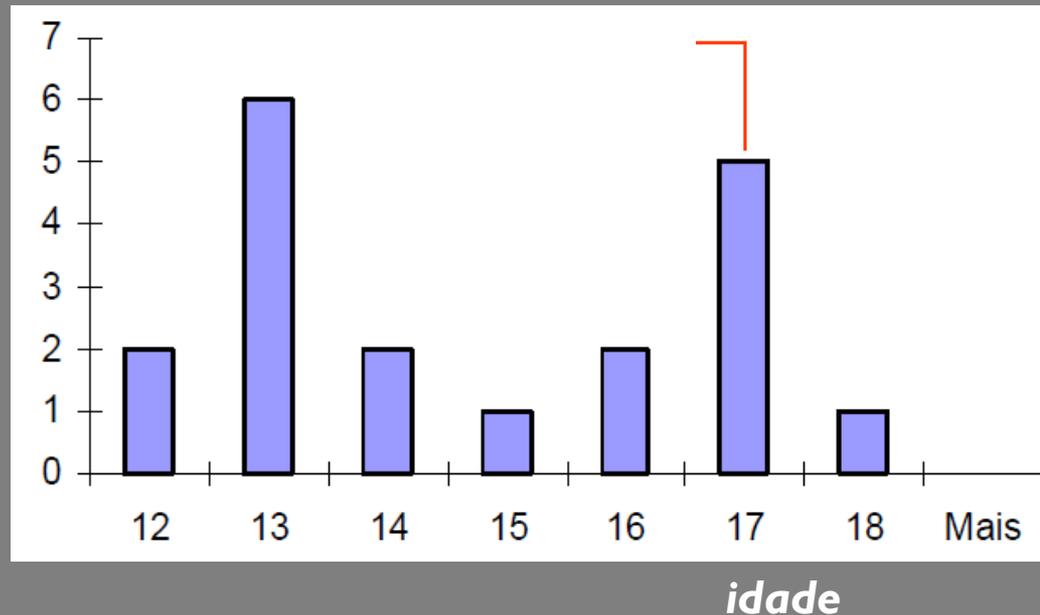
- Indica o valor ou categoria mais provável
- Pode não existir
- Pode não ser única no conjunto de dados
 - Bimodal, trimodal, plurimodal
- Não é afetada por valores discrepantes

EXEMPLO:

CALCULAR A MODA DA IDADE DOS PACIENTES DENTRO DE UMA CLÍNICA ODONTOLÓGICA:

17 13 17 18 13 12 13 17 16 13 12 17 13 15 13 16 17

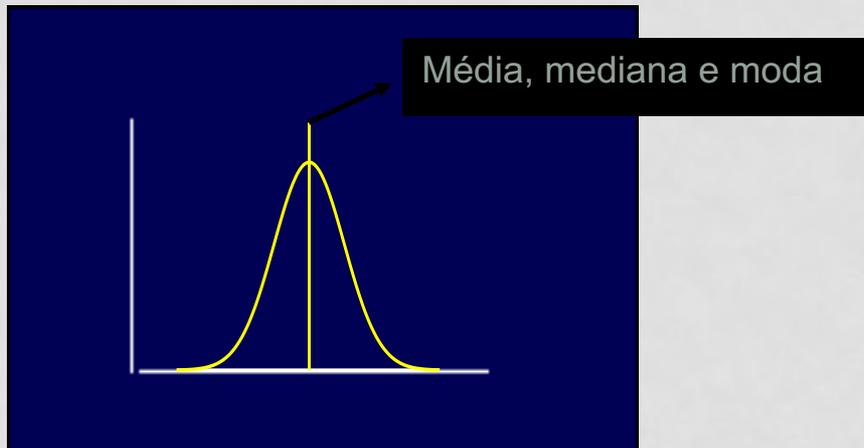
Número de indivíduos



A moda dos pacientes na clínica odontológica é 13 anos.

MEDIDAS DE TENDÊNCIA CENTRAL

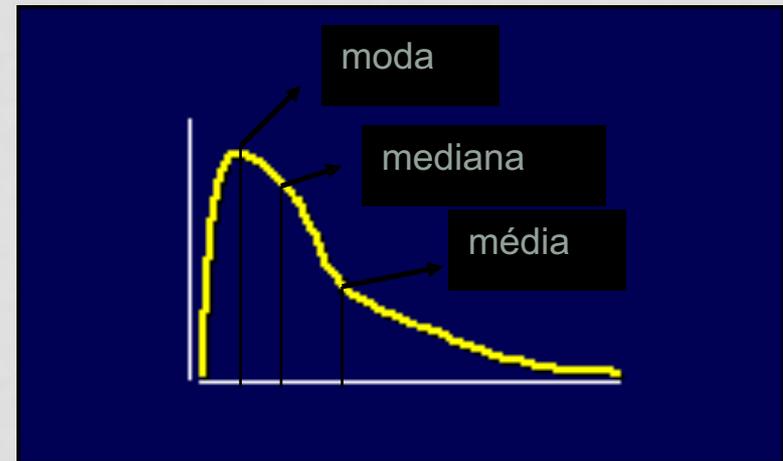
1



Distribuição Simétrica

60 – 70 anos = 10 indivíduos
70 – 80 anos = 10 indivíduos
80 – 90 anos = 10 indivíduos

2



**Distribuição com
Assimetria Positiva**

60 – 70 anos = 2 indivíduos
70 – 80 anos = 2 indivíduos
80 – 90 anos = 26 indivíduos

Qual medida de tendência central que melhor representa o grupo 2?

MEDIANA

	Média	Mediana	Moda
Definição	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	Valor do meio	Valor mais frequente
Existência	Sempre existe	Sempre existe	Pode não existir. Pode haver mais de uma.
Leva em conta todos os valores?	Sim	Não	Não
Afetada por valores discrepantes?	Sim	Não	Não
Vantagens	Usada em muitos métodos estatísticos	Menos sensível a valores discrepantes	Apropriada para dados qualitativos

MEDIDAS DE DISPERSÃO

- Medidas de tendência central → Não representam a variabilidade dos dados
- Permite a dimensão de valores discrepantes
- Diferença (dispersão, variabilidade) que existe entre a média e os outros valores do conjunto

Medidas de dispersão

Desvio médio

Variância

Desvio Padrão

⋮

Desvio médio

É a diferença de cada valor do conjunto em relação à média (\bar{x})

IDADE (X)	MÉDIA (\bar{x})	$ x - \bar{x} $
38	50,25	12,25
40	50,25	10,25
33	50,25	17,25
54	50,25	3,75
SOMA		43,5

DESVIO MÉDIO = $43,5/4 = 10,875$

VARIÂNCIA

É a soma dos quadrados dos desvios em relação a média (\bar{x}) dividida pelo número de observações menos 1.

IDADE (X)	MÉDIA (\bar{X})	(X - \bar{X}) ²
38	50,25	(-12,25) ² = 150,06
40	50,25	(-10,25) ² = 105,06
33	50,25	(-17,25) ² = 297,56
54	50,25	(3,75) ² = 14,06
SOMA		566,74

VARIÂNCIA = 566,74 / 4 - 1 = 188,91

DESVIO PADRÃO

- Raiz quadrada da variância

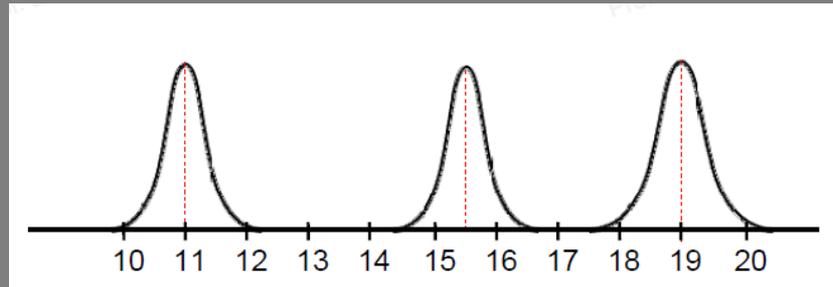
Se a variância = 188,91

$$\sigma = \sqrt{188,91} = 13,74$$

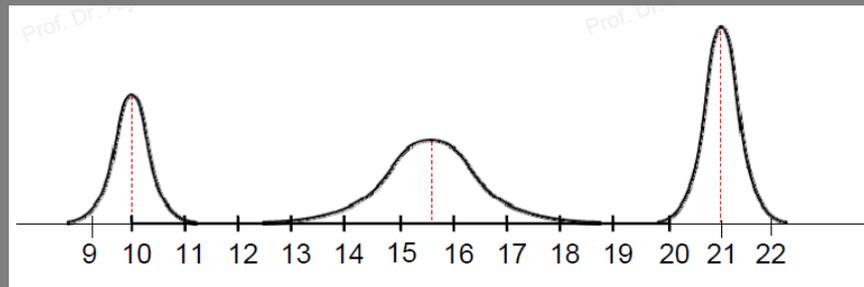
- Quanto maior o desvio padrão, maior a variância e mais a curva se afasta da média (curva normal)
- Quanto menor o desvio padrão, maior a homogeneidade da amostra

EXEMPLO:

CURVAS COM DIFERENTES MÉDIAS E O MESMO DESVIO PADRÃO



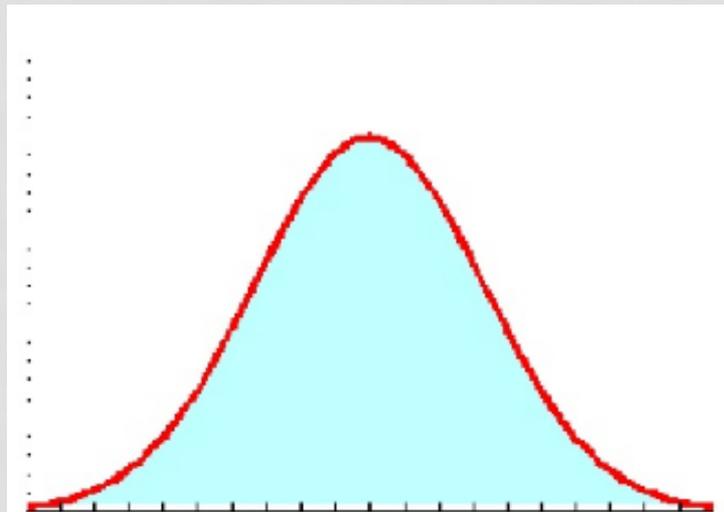
CURVAS COM DIFERENTES MÉDIAS E DESVIOS PADRÃO



QUAL CURVA TEM MAIOR DESVIO PADRÃO?

Distribuição normal

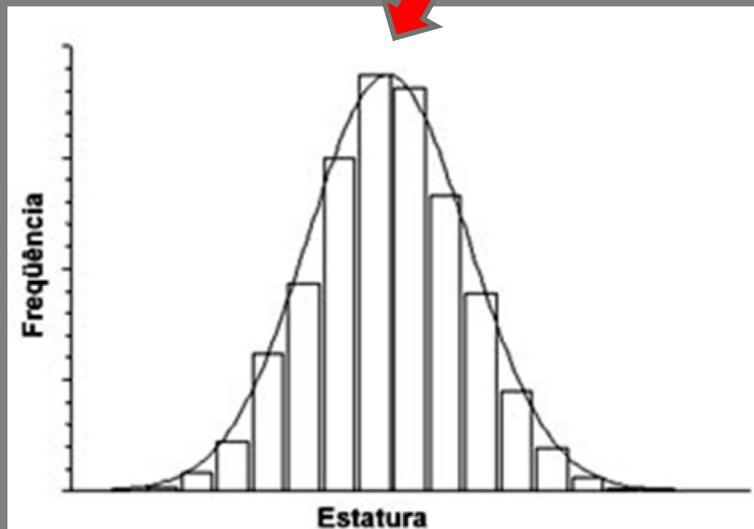
- Uma das distribuições fundamentais da teoria estatística
- É a distribuição contínua de uma variável aleatória X
- Seu gráfico é chamado de **CURVA NORMAL**



EXEMPLO:

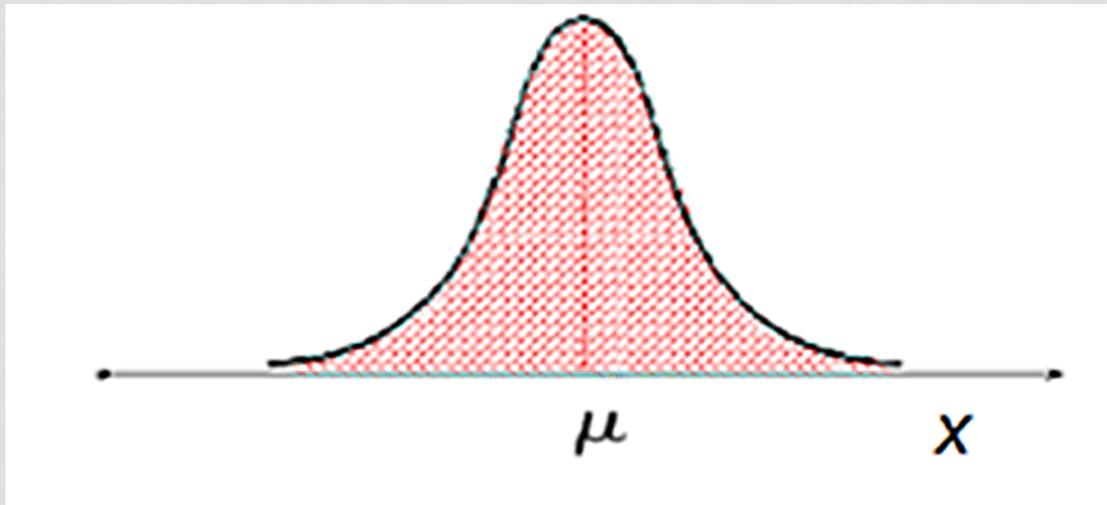
DISTRIBUIÇÃO DAS ESTATURAS (cm) EM CRIANÇAS DE 3 ANOS DE IDADE

MÉDIA



CURVA DE DISTRIBUIÇÃO NORMAL DOS DADOS

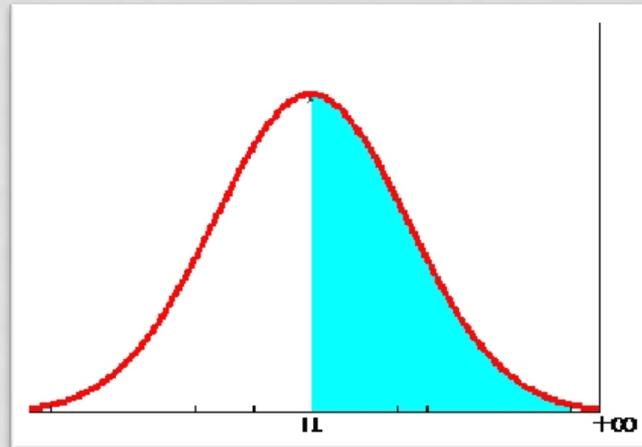
- Tem forma de sino e é simétrica em torno da média
- A área total da curva é **1** ou **100%** do $-\infty$ até $+\infty$
- À medida que a curva se afasta da média, aproxima-se cada vez mais do eixo X mas nunca o toca.



CURVA DE DISTRIBUIÇÃO NORMAL DOS DADOS

- Probabilidade \rightarrow área entre a curva e o eixo das abcissas

A probabilidade de ocorrer um evento entre $x=\mu$ e $x=+\infty$ é 0,5 ou 50%



CURVA NORMAL - FINALIDADE

- Muitos testes estatísticos se baseiam na suposição de que os dados observados na amostra são provenientes de uma população com distribuição normal
- Saber se a variável estudada na amostra tem uma distribuição normal é necessário para a realização de várias análises estatísticas
 - Teste de Shapiro Wilks: amostras menores ($n < 30$)
 - Teste de Kolmogorov-Smirnov ($n \geq 30$)

Força Muscular Respiratória e Mobilidade Torácica em Crianças e Adolescentes com Leucemia Aguda e Escolares Saudáveis

Respiratory Muscle Strength and Thoracic Mobility in Children and Adolescents with Acute Leukemia and Healthy School Students

Fuerza Muscular Respiratoria y Movilidad Torácica en Niños y Adolescentes con Leucemia Aguda y Escolares Sanos

Kátia Myllene Costa Oliveira¹, Thalita Medeiros Fernandes de Macêdo², Raíssa de Oliveira Borja³, Rafaela Andrade do Nascimento⁴, Wilson Cleto de Medeiros Filho⁵, Tânia Fernandes Campos⁶, Diana Amélia de Freitas⁷, Karla Morganna Pereira Pinto de Mendonça⁸

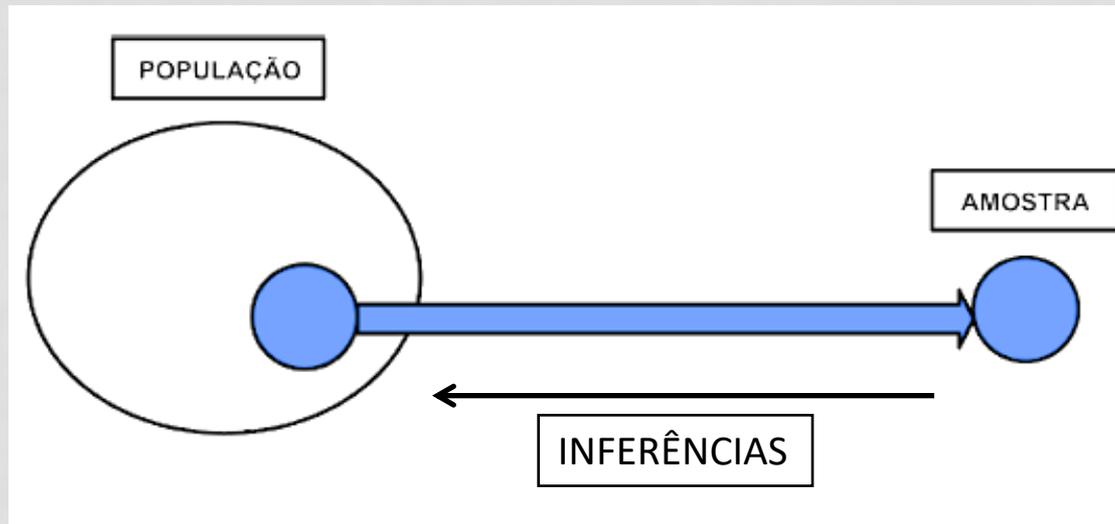
Tendo em vista os sinais e sintomas da leucemia aguda e os efeitos colaterais provenientes de seu tratamento, o presente estudo teve como propósito comparar a força muscular respiratória e a mobilidade da caixa torácica entre crianças e adolescentes com leucemia aguda na fase de manutenção do tratamento quimioterápico e indivíduos saudáveis, avaliando se essas variáveis sofrem influência do tratamento proposto para esses pacientes.

ANÁLISE ESTATÍSTICA

Os dados foram analisados através do *software* SPSS 15.0 (*Statistical Package for the Social Science*) atribuindo-se o nível de significância de 5%. A prova de Kolmogorov-Smirnov foi aplicada para avaliar a normalidade dos dados da amostra. Segundo a prova citada, observou-se que apenas as variáveis idade, expansão axilar e expansão xifoide não apresentaram distribuição normal. Para a comparação das variáveis com distribuição normal entre os grupos A e B, foi utilizado o teste t' student não pareado, enquanto as demais variáveis foram analisadas através da aplicação do teste de Mann-Whitney. Foi feita a estatística descritiva da amostra, com base na média e desvio-padrão, para todas as variáveis dos grupos A e B.

Estatística inferencial

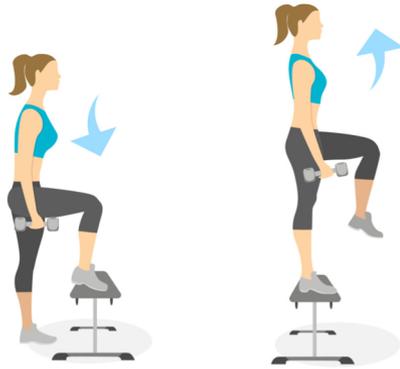
- Visa inferir sobre características de uma POPULAÇÃO a partir da AMOSTRA



Testes de hipóteses

- Verificar se as diferenças entre os valores das amostras são devido ao acaso (variabilidade natural) ou se são decorrentes de outros fatores
- Hipótese nula (**H0**):
 - Não diferença entre as amostras
- Hipótese alternativa (**H1**)
 - Propõe que existe diferença entre as amostras e que a hipótese nula (**H0**) deve ser rejeitada

Grupo exercícios



Grupo acompanhado sem exercício



Hipóteses do estudo:

H0 = A dor vai melhorar em ambos os grupos

H1 (alternativa) = **A dor no grupo submetido ao exercício vai melhorar mais que no grupo que não fez nenhuma intervenção**

Grupo exercícios



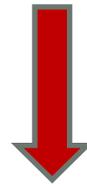
Pré-tto: Intensidade de Dor 8

Grupo acompanhado sem exercício



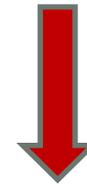
Pré-tto: Intensidade de Dor 8

Se a hipótese **nula** for confirmada

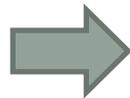


Pós-tto:
Intensidade de Dor 4

Se a hipótese **alternativa** for confirmada

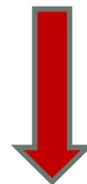


Pós-tto:
Intensidade de Dor 4



=

≠



Pós-tto:
Intensidade de Dor 4

Pós-tto:
Intensidade de Dor 8

Estatística inferencial

TESTES ESTATÍSTICOS

Utilizados para fazer inferências sobre uma população através da amostra

- **Testes Paramétricos**

Teste t de Student

ANOVA (Teste F)

⋮

- **Testes Não Paramétricos**

Teste de Mann-Whitney

Teste de Kruskal-Wallis

⋮

Testes paramétricos

- Partem do pressuposto de que a amostra é proveniente de uma população com distribuição normal
- Sendo a amostra representativa da população, ela também precisa ter uma distribuição normal dos seus dados (curva simétrica)
- Homogeneidade das variâncias
- Testes mais rigorosos
- Análises mais eficientes

Testes Paramétricos

Teste t de Student

Comparar 2 grupos

EX:

Pacientes

Saudáveis

ANOVA (Teste F)

Comparar 3 ou mais grupos

EX:

Alunos

1°

ano

Alunos

2°

ano

Alunos

3°

ano

Testes não paramétricos

- Não exigem populações distribuídas em uma curva normal
- Sem homogeneidade da variância
- Testes menos rigorosos que os paramétricos

Testes Não Paramétricos

Teste de Mann-Whitney

Comparar 2 grupos

EX:

Pacientes

Saudáveis

Teste de Kruskal-Wallis

Comparar 3 ou mais grupos

EX:

Alunos

1°

ano

Alunos

2°

ano

Alunos

3°

ano

RESUMO:

	Testes Paramétricos	Testes Não Paramétricos
Distribuição dos dados	Curva normal	Qualquer uma
Variância assumida	Homogênea	Qualquer uma
Medida utilizada	Média	Mediana
Comparar 2 grupos	Teste t de Student	Mann-Whitney
Comparar 3 grupos	ANOVA	Kruskal-Wallis

COMO INTERPRETAR OS RESULTADOS DE UMA
INFERÊNCIA ESTATÍSTICA?

VALOR p

- O valor da probabilidade de se obter o efeito observado, dado que a hipótese nula é verdadeira, é chamado de p-valor
- p-valor : estatística utilizada para sintetizar o resultado de um teste de hipóteses.
- Se o nível de significância for 5%, o p-valor tem que ser menor que 5% ($p < 0,05$)

($p < 0,05$) \Longrightarrow **REJEITO H_0**

($p > 0,05$) \Longrightarrow **ACEITO H_0**

Grupo exercícios



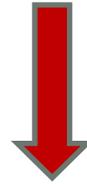
Pré-tto: Intensidade de Dor 8

Grupo acompanhado sem exercício



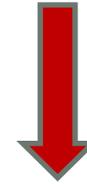
Pré-tto: Intensidade de Dor 8

Se a hipótese **nula** for confirmada

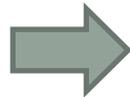


Pós-tto:
Intensidade de Dor 4

Se a hipótese **alternativa** for confirmada

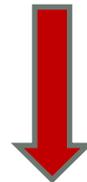


Pós-tto:
Intensidade de Dor 4



=
Aceito hipótese nula

≠
Rejeito a hipótese nula



Pós-tto:
Intensidade de Dor 4

Pós-tto:
Intensidade de Dor 8

VALOR DE CORTE DE 5%

- Valor de de probabilidade para aceitar ou rejeitar a H0

- Quando $P < 0.05$ = Ex: $p=0.03 = 3\%$

$G1 \neq G2$

Há 3% de chance de que a H0 seja verdadeira

- Quando $P > 0.05$ = Ex: $p=0.2 = 20\%$

$G1 = G2$

Há 20% de chance de que a H0 seja verdadeira

A pergunta que não quer calar é: porque $p = 0.05$?



EXEMPLO:

p-valor ($p < 0,05$)	H ₀	Diferença entre os grupos
0,06	Aceito	Não
0,01	Rejeito	Sim
0,85	Aceito	Não
0,04	Rejeito	Sim
0,09	Aceito	Não
0,05	Rejeito	Sim

Efeitos de orientações fisioterapêuticas sobre a ansiedade de pacientes submetidos à cirurgia de revascularização miocárdica

Effects of physiotherapeutic instructions on anxiety of CABG patients

Aline GARBOSSA¹, Emilia MALDANER¹, Daiana Moreira MORTARI¹, Janaina BIASP², Camila Pereira LEGUISAMO³

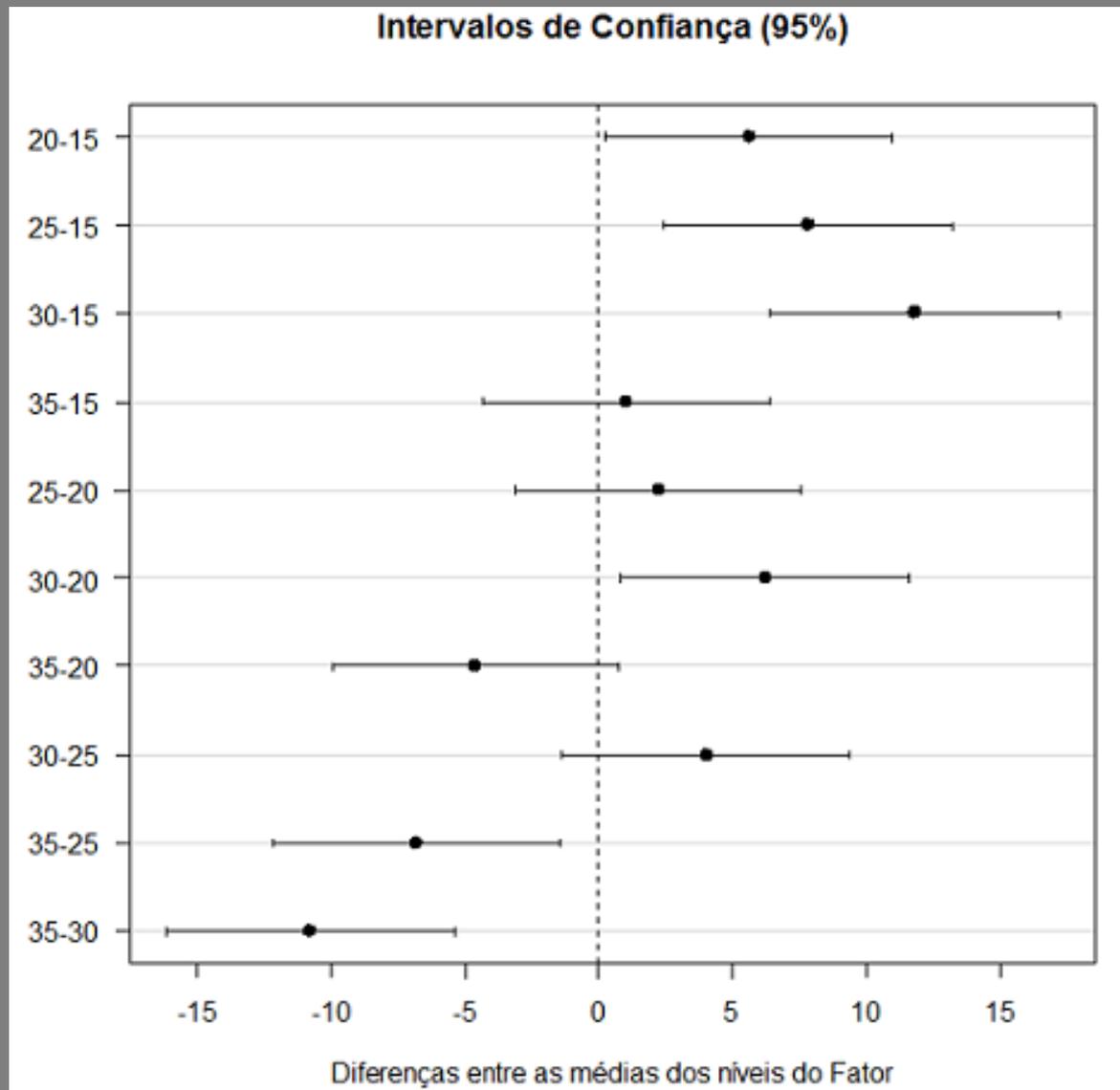
Objetivo: O objetivo deste estudo é verificar os efeitos de orientações fisioterapêuticas sobre o nível de ansiedade em pacientes submetidos à cirurgia de revascularização do miocárdio.

Resultados: Observaram-se escores de ansiedade mais baixos nos pacientes que receberam a intervenção antes da cirurgia ($9,6 \pm 7,2$ versus $13,4 \pm 5,9$, $P=0,02$). No grupo controle, a diferença entre os escores de ansiedade antes e após a cirurgia foi significativa ($P=0,003$). Os indivíduos do sexo feminino apresentaram-se mais ansiosos no pré-operatório

Intervalo de Confiança- IC

- Inferência estatística complementar ao teste de hipóteses
- Extrapolação dos dados da amostra para a população (valor mínimo e máximo)
- Importância Clínica
- Estipula em 95% de certeza que a média populacional se encontra neste intervalo

Relação entre IC e valor p



EXEMPLO:

Tabela 1. Risco Relativo (RR) relacionado aos níveis de incapacidade do NDI entre os grupos M e MC

NDI	M (RR)*	MC (RR)*
Nenhuma	20,000 (3,283;121,84)	0,050 (0,008; 0,305)
Leve	7,818 (2,215; 27,600)	0,128 (0,036;0,452)
Moderada	4,000 (1,096;14,600)	0,250 (0,068;0,912)
Severa	Ref.	Ref.
Completa	**	**

RR: Risco Relativo; * $p < 0.001$. ** Nível de severidade não foi considerada na análise

OBRIGADA !