

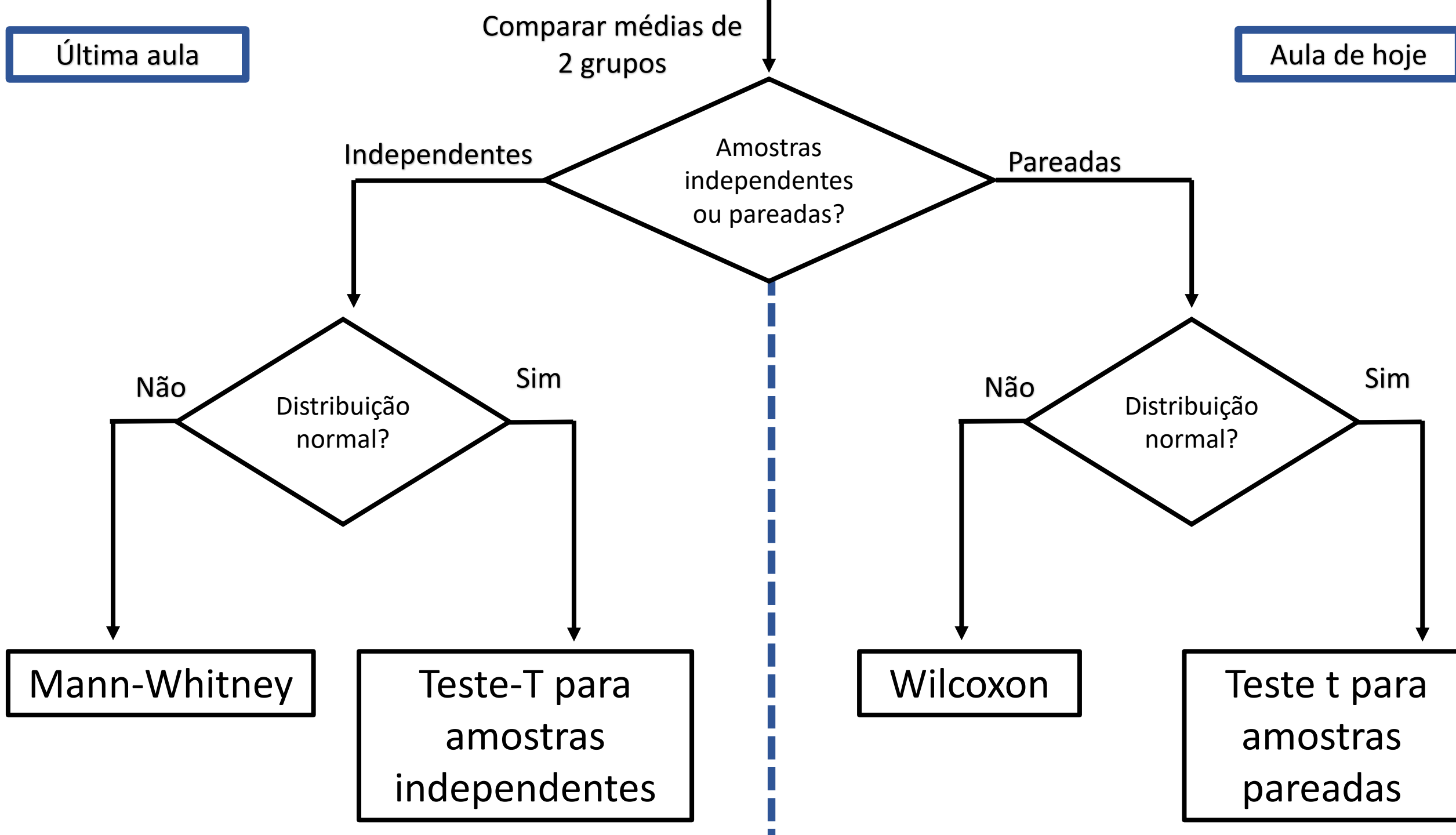
ESTATÍSTICA PARA SAÚDE COLETIVA

Aula 8

Última aula

Comparar médias de
2 grupos

Aula de hoje



Revisão última aula

Variável resposta

Variável preditora

Expectativa de vida



Sexo

Em outras palavras, sexo determina a expectativa de vida

Revisão última aula

Contínua

Categórica

Expectativa de vida

Sexo

Em outras palavras, sexo determina a expectativa de vida

Aula de hoje

Variável resposta

Variável preditora

Glicose no sangue ← antes e depois da insulina

Em outras palavras, ao tomar insulina houve mudança nos níveis de glicose sanguínea?

Aula de hoje

Contínua

Categórica

Glicose no sangue ← antes e depois da insulina

Em outras palavras, ao tomar insulina houve mudança nos níveis de glicose sanguínea?

Amostras pareadas

Amostras pareadas

- Amostras pareadas/inter-relacionadas são aquelas obtidas a partir da diferença entre dois valores-interrelacionados, podem ter origem em:
 - Medições repetidas
 - Amostras combinadas

Amostras pareadas

- Quando uma mesma unidade amostral foi medida duas vezes no decorrer da pesquisa
 - A ideia é testar se houve ou não diferença em alguma coisa antes e depois de alguma intervenção
 - Exemplos:
 - Testar se o conhecimento geral sobre estatística mudou de antes e depois do curso
 - Algum parâmetro da saúde foi avaliado em um mesmo indivíduo com ou sem a aplicação de um medicamento
- Amostras são dadas em pares, exemplos:
 - Participantes de uma pesquisa que foram entrevistados em duplas.
 - Avaliar se existem diferenças na altura de homens e mulheres, amostrando irmãos de sexo oposto

Exemplos de adaptações de perguntas da Lista um que podem ser respondidas com essa metodologia

- 1) Aplicação de insulina de fabricação de laboratório altera a glicemia sanguínea? (Anderson)
- 2) Implementar serviço público de atendimento psicológico implica em melhoria da qualidade de vida? (medido pelo IDH) (William)
- 3) A expectativa de vida é maior em países que adotam práticas de universalização da saúde pública? (Tainara Pradella)

Exemplos de adaptações de perguntas: identificar variáveis

- Aplicação de insulina de fabricação de laboratório altera a glicemia sanguínea?
 - Glicemia:
 - variável resposta
 - quantitativas contínua
 - Aplicação de insulina:
 - variável preditora
 - qualitativa (Sim ou Não)
 - Amostras
 - Uma mesma pessoa terá dados glicemia coletados no estado “sem insulina” e no estado “com insulina”

Exemplos de adaptações de perguntas: identificar variáveis

- Implementar serviço público de atendimento psicológico implica em melhoria da **qualidade de vida(medido pelo IDH)?**
 - Qualidade de vida (IDH):
 - variável resposta
 - quantitativas contínua
 - Implementação de serviço público de atendimento psicológico:
 - variável preditora
 - qualitativa (Sim ou Não)
 - Amostras
 - Uma mesmo país terá dados IDH antes e depois de implementar o serviço publico de atendimento psicológico (imagine que você dê um intervalo de 20 anos entre uma observação e outra).

Exemplos de adaptações de perguntas: identificar variáveis

- A expectativa de vida é maior em países que adotam práticas de universalização da saúde pública ?
 - Expectativa de vida
 - variável resposta
 - quantitativas contínua
 - Antes e depois de adotar práticas de universalização da saúde pública
- OBS: nesse caso seria interessante controlar no numero de anos que o país implementou a pratica. Uma possibilidade seria incluir no teste apenas países que implementaram a prática a mais de 30 anos.

Desenho experimental

- Para comparar valores de média em 2 momentos, você deve usar logica da aula anterior, só que agora em vez de amostrar grupos diferentes você está amostrando em momentos diferentes.

Exemplos:

- a. Coletar dados relacionado a quantidade de glicemia sanguínea em pessoas antes aplicação da insulina, e 10 minutos depois da aplicação da insulina.
- b. Buscar dados de países que implementaram serviço público de atendimento psicológico, e comparar o valor de IDH antes da implementação, e 20 anos depois da implementação.
- c. Obter dados de expectativa de vida países antes e depois da adoção práticas de universalização da saúde pública

Hipóteses

- Teste t pareado
 - H_0 : A média da diferença entre os grupos não difere de zero
 - H_1 : A média da diferença entre os grupos difere de zero
- Wilcoxon
 - H_0 : A mediana da diferença entre os grupos não difere de zero
 - H_1 : A mediana da diferença entre os grupos difere de zero

Como fica uma tabela desse tipo de dado?

(Medições repetidas)

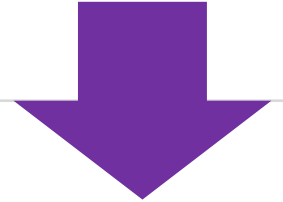
Identificação do voluntário	Numero de horas de sono quando não toma café	Numero de horas de sono quando toma café
Aline	7,9	7,62
Paulo	8,91	7,71
Fernanda	7,83	7,36
José	9,11	7,13
Juliana	7,9	6,75
Antônio	8,17	8,49
Maria	7,92	7,05
João	8,04	7,37
Ana	7,21	7,35
Francisco	7,78	6,71

Faça uma tabela com 3 colunas repetidas		
Identificação do dado	Observação 1	Observação 2
Identificação do voluntário	Numero de horas de sono quando não toma café	Numero de horas de sono quando toma café
Aline	7,9	7,62
Paulo	8,91	7,71
Fernanda	7,83	7,36
José	9,11	7,13
Juliana	7,9	6,75
Antônio	8,17	8,49
Maria	7,92	7,05
João	8,04	7,37
Ana	7,21	7,35
Francisco	7,78	6,71

NOTA

- Aqui o foco é avaliar se existe diferença na média dos grupos

Note que nós estamos interessados no dado de diferença



Identificação do voluntário	Numero de horas de sono quando não toma café	Numero de horas de sono quando toma café	Diferença
Aline	7,9	7,62	0,28
Paulo	8,91	7,71	1,2
Fernanda	7,83	7,36	0,47
José	9,11	7,13	1,98
Juliana	7,9	6,75	1,15
Antônio	8,17	8,49	-0,32
Maria	7,92	7,05	0,87
João	8,04	7,37	0,67
Ana	7,21	7,35	-0,14
Francisco	7,78	6,71	1,07

Como fica uma tabela desse tipo de dado?
(Amostras combinadas)

Familia	Irmão	Irmã	Diferença (grupo 1- grupo 2)
Silva	160	155	5
Souza	163	167	-4
Costa	172	173	-1
Santos	180	175	5
Oliveira	159	161	-2
Pereira	165	160	5
Rodrigues	175	165	10
Almeida	175	169	6
Nascimento	159	152	7
Lima	170	163	7

NOTA 2

- Os pressupostos (normalidade e homogeneidade) são os mesmos do teste t não pareado. Mas em vez de olhar para os dados originais, olhamos apenas para os valores de diferença!

Teste uni ou bicaudal?

Existe evidências de que existe direção na comparação?

Teste bicaudal

- Avalia se existe alguma diferença entre os grupos comparados.
- Nesse caso, você testa:
 - Existe diferença entre o grupo A e grupo B?
 - Se houver diferença, pode ser tanto que A é maior como menor que B

Teste unicaudal

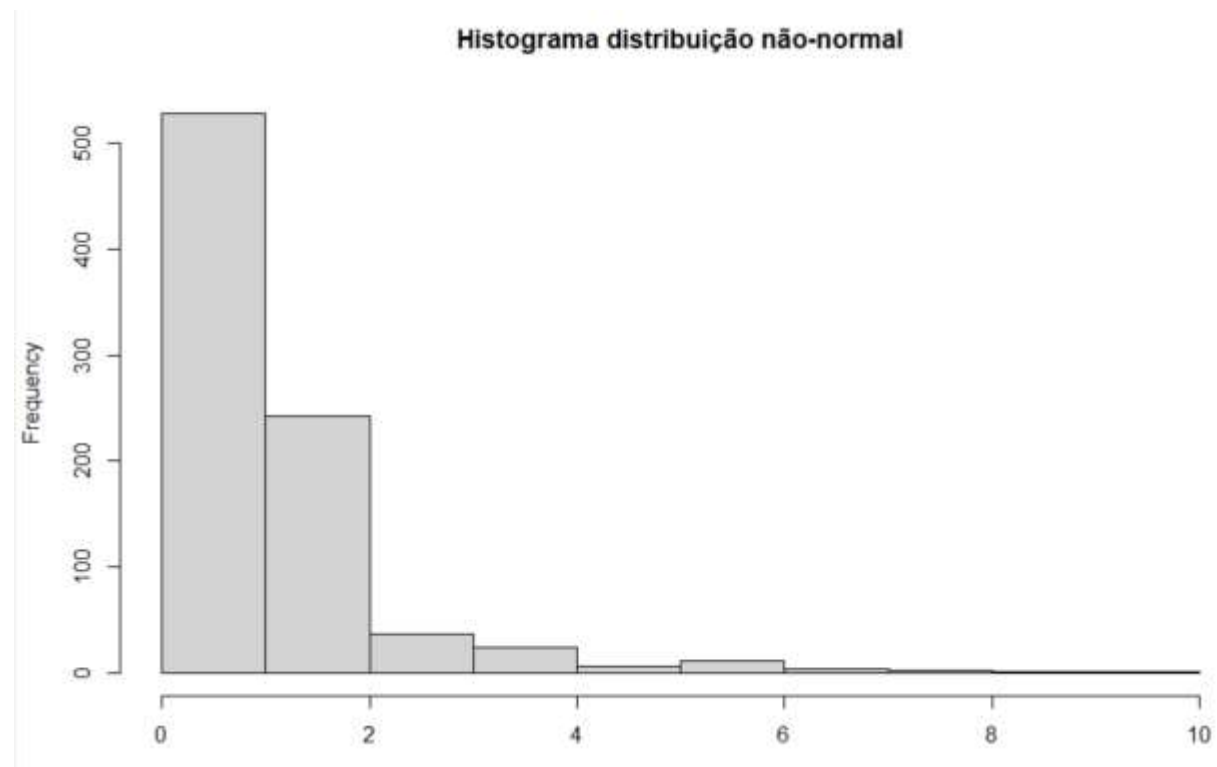
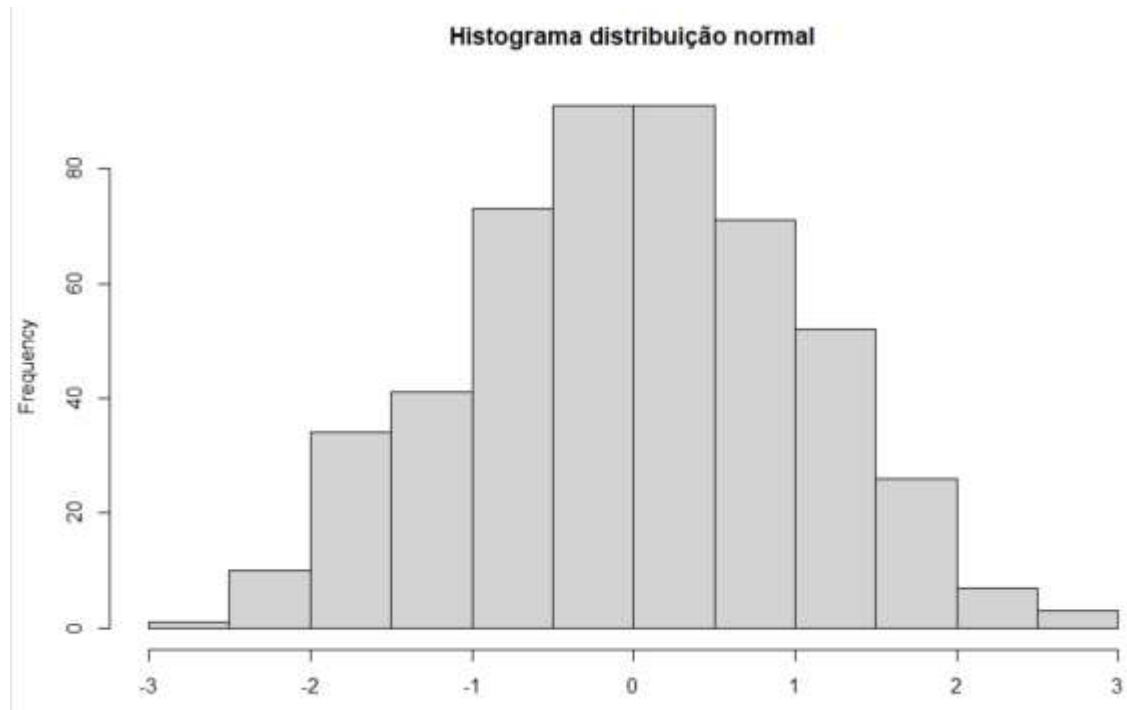
- Avalia se um grupo é maior(ou menor) que o outro.
- Nesse caso, você testa especificamente:
 - Grupo A é maior que o grupo B?
 - Grupo A é menor que o grupo B?

O que você vai precisar verificar
antes de rodar o teste

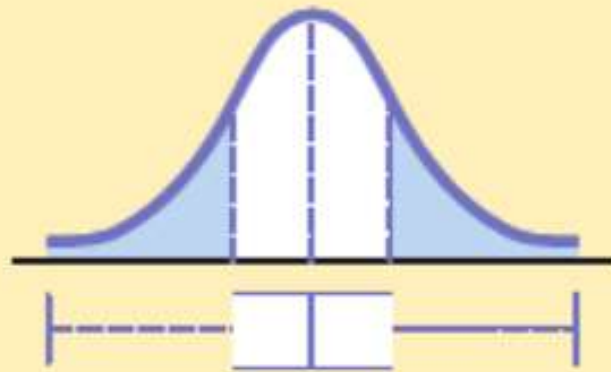
O que você vai precisar saber o dia de hoje?

- Para comparar médias de amostras dependentes você vai ter que saber se:
 - Amostras são pareadas ou independentes?
 - Dados tem distribuição normal?

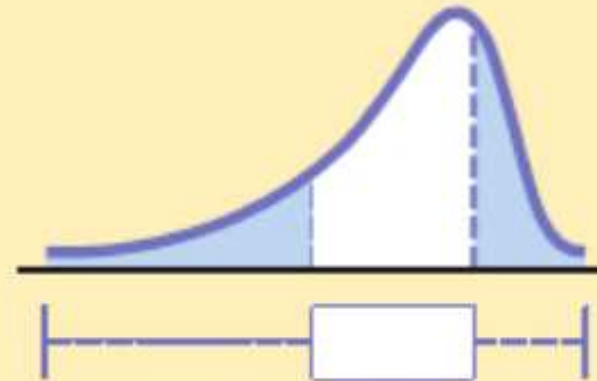
Normalidade



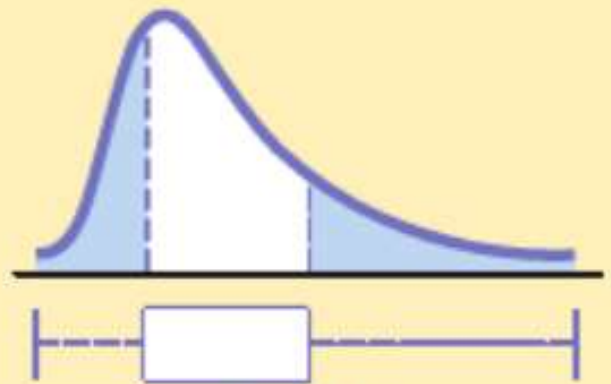
Normalidade



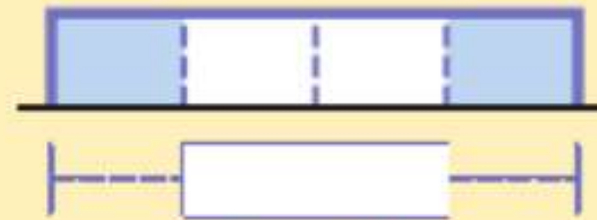
Simétrico



Assimétrico à esquerda



Assimétrico à direita

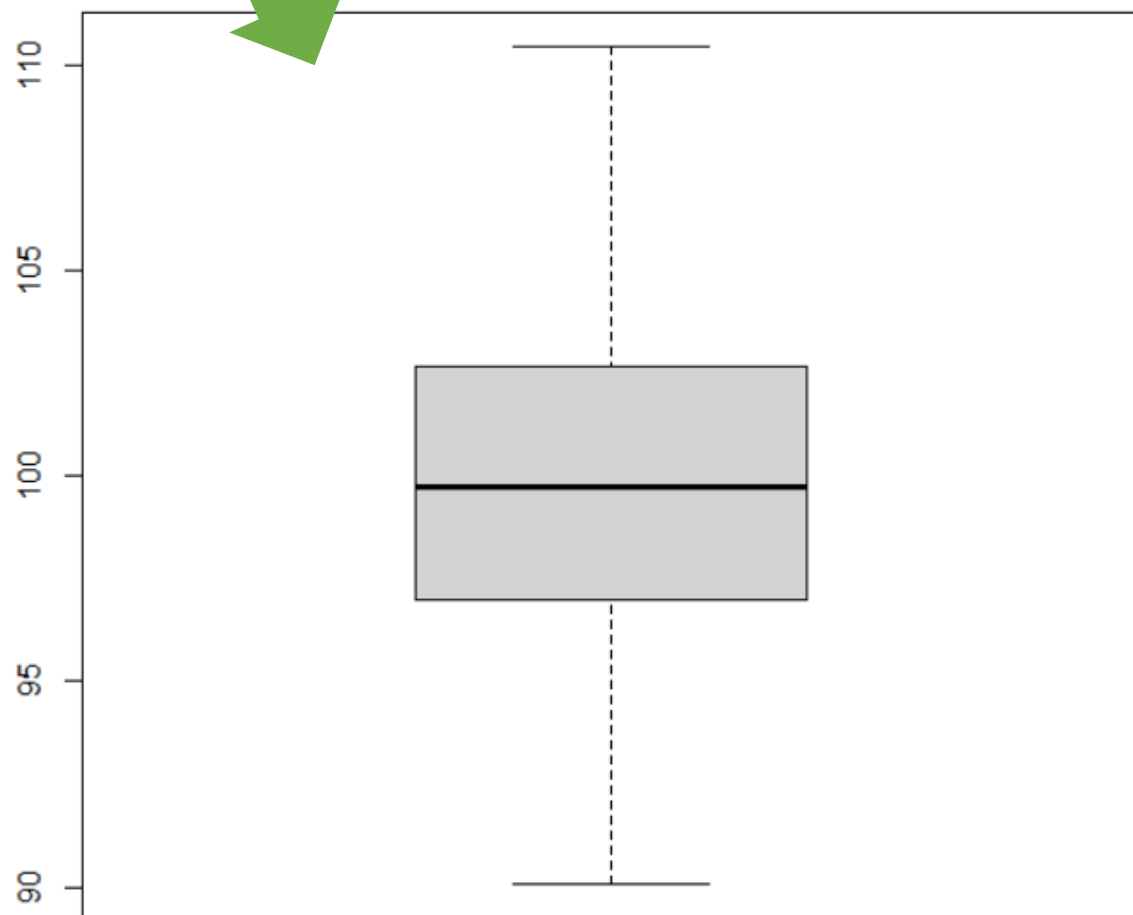


Distribuição retangular

Normalidade

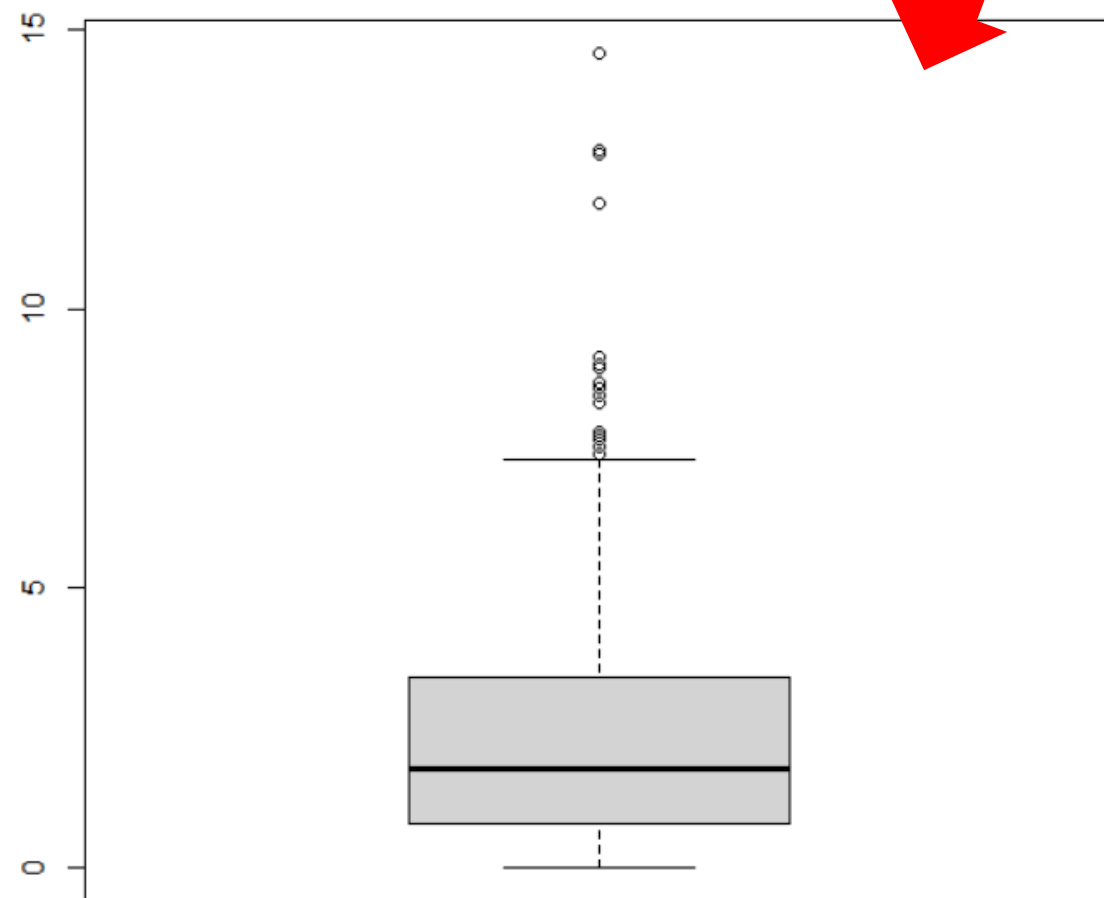
Indício de que os dados são normais

Dados normais



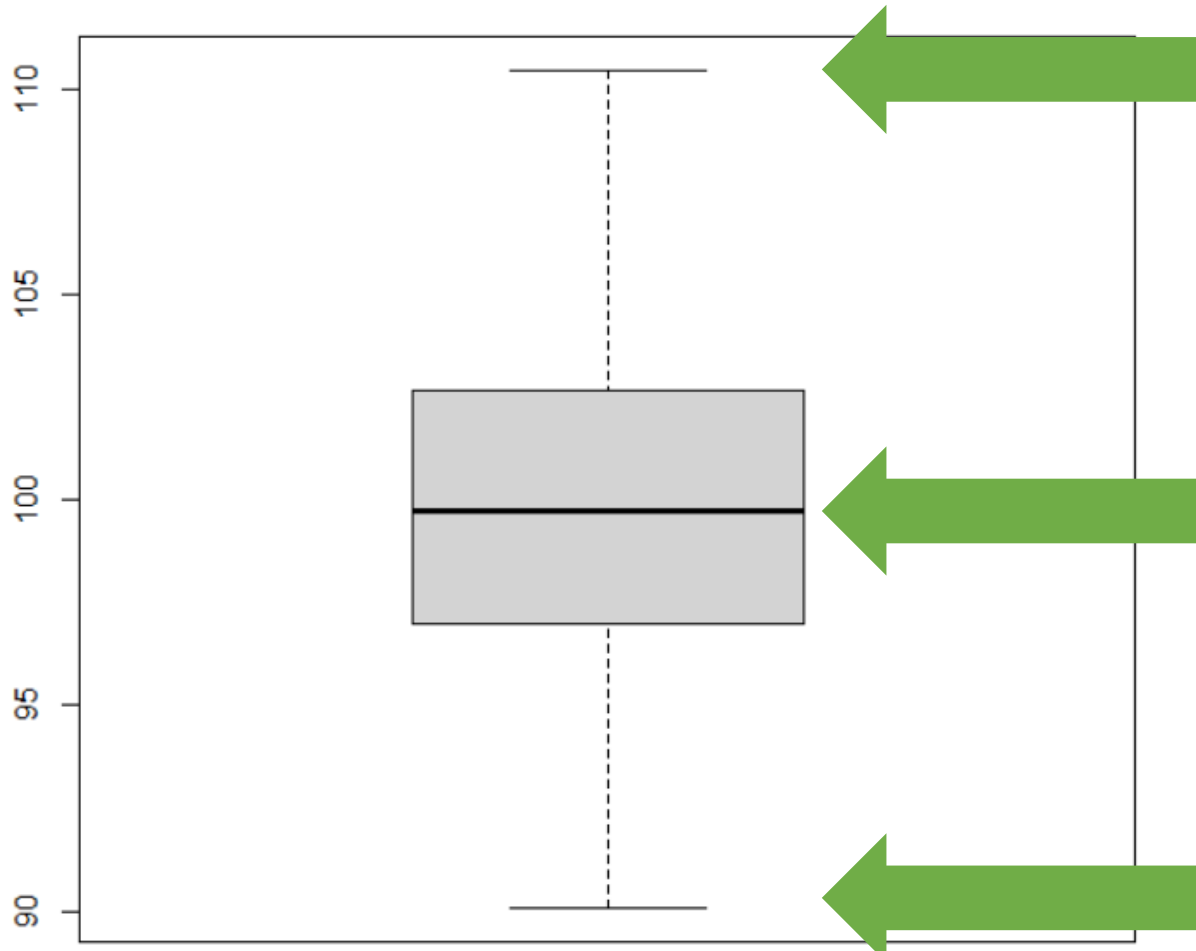
Indício de que os dados não são normais

Dados não-normais



Normalidade

Dados normais



Simetria no tamanho das cauda (a distância da caixa cinza até a extremidade de cima e de baixo é igual)

Caixa cinza está no meio do box-plot

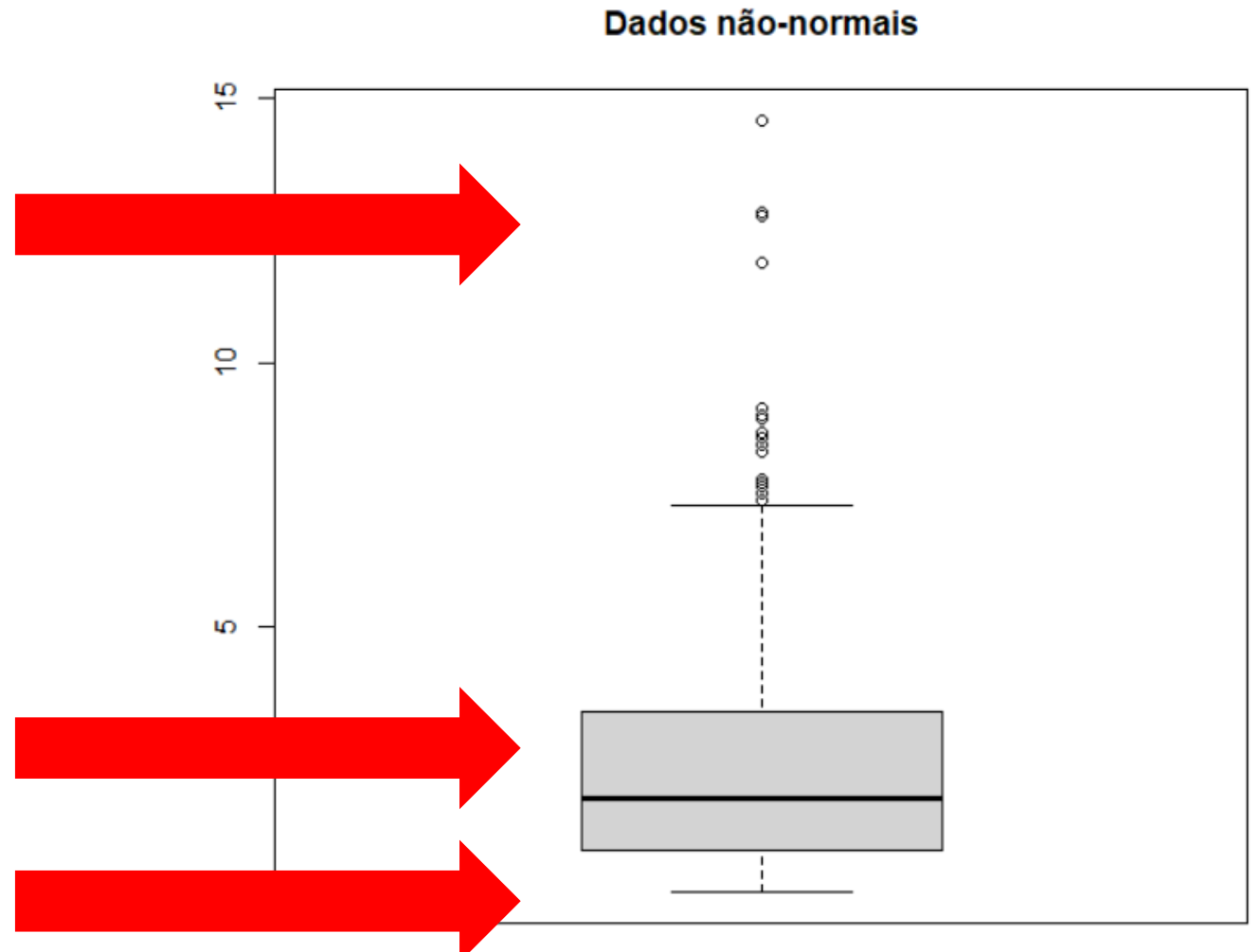
Simetria no tamanho das cauda (a distância da caixa cinza até a extremidade de cima e de baixo é igual)

Normalidade

Evidências de outliers (iremos falar mais disso na próxima aula)

Caixa cinza não está no meio do box-plot

Não há simetria no tamanho das cauda



Vamos começar nosso teste!

- Abra o Script dessa aula

Prática – 1 Existe diferença no número de horas de sono médio de mulheres e homens?

- Tabela “Dados para aula 7 e 8.xlsx”
- Variável preditora: Antes e depois de experimento
- Variável resposta: Vitamina B12
- Tipo de dado:
 - Amostras pareadas/inter-relacionadas
 - Diferença com distribuição normal
- Teste T
 - Usado para compara 2 medias
 - H_0 = a média da diferença é igual a zero
 - H_1 = a média da diferença é diferente de zero

Interpretação dos resultados

Resultado: P maior que 0.05
Conclusão: H_0 é verdadeira

Paired t-test

```
data:  tabela1$B12A and tabela1$B12D
t = 0.48583, df = 99, p-value = 0.6282
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -15.82191  26.08191
sample estimates:
mean of the differences
          5.13
```

Prática – 2 Existe diferença no número de horas de sono médio de mulheres e homens?

- Tabela “Dados para aula 7 e 8.xlsx”
- Variável preditora: Antes e depois de experimento
- Variável resposta: Vitamina B6
- Tipo de dado:
 - Amostras pareadas/inter-relacionadas
 - Diferença com distribuição normal
- Wilcoxon
 - Usado para compara 2 mediana
 - H_0 = a mediana da diferença é igual a zero
 - H_1 = a mediana da diferença é diferente de zero

Interpretação dos resultados

Welch Two Sample t-test

Resultado: P menor que 0.05
Conclusão: H_1 é verdadeira

```
data:  tabela1$Altura by tabela1$Sexo
t = -7.9585, df = 71.643, p-value = 1.889e-11
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -16.197265  -9.707935
sample estimates:
mean in group F mean in group M
    157.2442         170.1968
```

OBS: $1.889e-11 = 1.889 \times 10^{-11} = 0.0000000000001889$