

ESTATÍSTICA PARA SAÚDE COLETIVA

Aula 7

Recados

- Semana que vem inicia-se a apresentação de seminário!
- Apresentações para dia 01 de Outubro

Nomes (Aluno ou Dupla)	Nome do artigo
ALEX SILVA + WILLIAM	Quedas entre idosos brasileiros residentes em áreas urbanas: ELSI-Brasil
DIEGO REIS	Resistência insulínica e sua relação com excesso de peso em adolescentes de uma capital do Nordeste Brasileiro
MARYNA LUZ + MARINA MENOTTI	Problemas emocionais e uso de medicamentos psicotrópicos: uma abordagem da desigualdade racial

Recados

- Próxima aula vai ser curta porque o foco vai ser que vocês apliquem um teste quase do zero.
- Se na próxima aula você conseguir rodar o ultimo teste da lista de exercícios, isso quer dizer que você consegue rodar qualquer testes apresentado nessa disciplina!

Feedback lista 5 e 6

- Revisão das questões com maior erro

Feedback lista 5

Perguntas erradas com frequência

Pergunta

Respostas corretas

Pergunta 4. Quando comparamos intervalos de confiança, existe um problema de aumento do erro por conta da incerteza sobre a estimativa de cada um dos parâmetros (assunto discutido em aula). Contudo, é comum que pessoas façam uma avaliação superficial, comparando os intervalos de confiança de dois ou mais grupos. Note que isso foi sugerido na região "nota" dessa tabela (que foi extraída na íntegra do sistema tabnet.saude.prefeitura.sp.gov.br). Mesmo sabendo desse problema, faça como o sugerido na nota, e compare os intervalos da Regional Centroeste e Leste. Em seguida avalie qual das afirmativas a seguir é VERDADEIRA.

5 / 16

Feedback lista 5

INQUÉRITO DE SAÚDE DO MUNICÍPIO DE SÃO PAULO - ISA CAPITAL 2015 - DIABETES MELLITUS

% Coord Regional de Saúde segundo Coord Regional de Saúde
Período: 2015

Coord Regional de Saúde	% Coord Regional de Saúde	IC _{95%}
TOTAL	100,00	(100,00-100,00) CV=...
Centroeste	13,41	(7,93-18,90) CV=0,2086
Leste	20,65	(13,63-27,66) CV=0,1734
Norte	19,51	(12,35-26,68) CV=0,1874
Sudeste	23,72	(15,71-31,73) CV=0,1723
Sul	22,71	(15,03-30,39) CV=0,1726

Fonte: ISA Capital 2015

Nota:

No ISA Capital, a tabulação e a análise diferem das realizadas para as demais bases de dados disponíveis no Tabnet SMS/SP. Portanto não deixe de ler as **Instruções de Uso**.

Nos resultados do ISA Capital é importante estar atento que na comparação das prevalências devem ser consideradas *diferenças significativas* quando não houver sobreposição dos respectivos intervalos de confiança e *sem diferença* quando um dos intervalos de confiança for parcial ou totalmente englobado pelo outro.

O coeficiente de variação é uma forma de expressar a variabilidade dos dados e é o resultado do quociente do desvio padrão pela média. CV maior do que 0,3 (30%) indica que a dispersão dos dados é muito alta e que a estimativa não tem confiança estatística, mesmo que o valor (n) na célula seja maior que 30.

Feedback lista 5

INQUÉRITO DE SAÚDE DO MUNICÍPIO DE SÃO PAULO - ISA CAPITAL 2015 - DIABETES MELLITUS

% Coord Regional de Saúde segundo Coord Regional de Saúde
Período: 2015

Coord Regional de Saúde	% Coord Regional de Saúde	IC _{95%}
TOTAL	100,00	(100,00-100,00) CV=...
Centroeste	13,41	(7,93-18,90) CV=0,2086
Leste	20,65	(13,63-27,66) CV=0,1734
Norte	19,51	(12,35-26,68) CV=0,1874
Sudeste	23,72	(15,71-31,73) CV=0,1723
Sul	22,71	(15,03-30,39) CV=0,1726

Fonte: ISA Capital 2015

Nota:

No ISA Capital, a tabulação e a análise diferem das realizadas para as demais bases de dados disponíveis no Tabnet SMS/SP. Portanto não deixe de ler as **Instruções de Uso**.

Nos resultados do ISA Capital é importante estar atento que na comparação das prevalências devem ser consideradas *diferenças significativas* quando não houver sobreposição dos respectivos intervalos de confiança e *sem diferença* quando um dos intervalos de confiança for parcial ou totalmente englobado pelo outro.

O coeficiente de variação é uma forma de expressar a variabilidade dos dados e é o resultado do quociente do desvio padrão pela média. CV maior do que 0,3 (30%) indica que a dispersão dos dados é muito alta e que a estimativa não tem confiança estatística, mesmo que o valor (n) na célula seja maior que 30.

☒ Selecione as respostas corretas:

Pergunta 3. Note que foi apresentado o intervalo de confiança nessa tabela.
Sobre essa observação, qual das afirmativas a seguir é VERDADEIRA.

1 pontos

- ☐ A região Leste da cidade de São Paulo, detém exatamente 20,65% da população de pessoas com diabetes (com 100% de certeza).
- ☐ A região Centroeste da cidade de São Paulo, detém exatamente 13,41% da população de pessoas com diabetes (com 95% de certeza).
- ☒ A região Norte da cidade de São Paulo, detém algum valor entre 12,35 e 26,68% da população de pessoas com diabetes (com 95% de certeza). ✓
- ☐ A região Sul da cidade de São Paulo, detém um valor aproximado de 22,71 (com 95% de certeza), e com 100% de certeza apresenta algum valor entre 15,03 e 30,39% da população de pessoas com diabetes.
- ☐ Todas as afirmativas anteriores estão erradas.

 Adicionar feedback da resposta

Concluído

Feedback lista 5

INQUÉRITO DE SAÚDE DO MUNICÍPIO DE SÃO PAULO - ISA CAPITAL 2015 - DIABETES MELLITUS

% Coord Regional de Saúde segundo Coord Regional de Saúde
Período: 2015

Coord Regional de Saúde	% Coord Regional de Saúde	IC _{95%}
TOTAL	100,00	(100,00-100,00) CV=...
Centroeste	13,41	(7,93-18,90) CV=0,2086
Leste	20,65	(13,63-27,66) CV=0,1734
Norte	19,51	(12,35-26,68) CV=0,1874
Sudeste	23,72	(15,71-31,73) CV=0,1723
Sul	22,71	(15,03-30,39) CV=0,1726

Fonte: ISA Capital 2015

Nota:

No ISA Capital, a tabulação e a análise diferem das realizadas para as demais bases de dados disponíveis no Tabnet SMS/SP. Portanto não deixe de ler as **Instruções de Uso**.

Nos resultados do ISA Capital é importante estar atento que na comparação das prevalências devem ser consideradas *diferenças significativas* quando não houver sobreposição dos respectivos intervalos de confiança e *sem diferença* quando um dos intervalos de confiança for parcial ou totalmente englobado pelo outro.

O coeficiente de variação é uma forma de expressar a variabilidade dos dados e é o resultado do quociente do desvio padrão pela média. CV maior do que 0,3 (30%) indica que a dispersão dos dados é muito alta e que a estimativa não tem confiança estatística, mesmo que o valor (n) na célula seja maior que 30.

☒ Selecione as respostas corretas:

Pergunta 4. Quando comparamos intervalos de confiança, existe um problema de aumento do erro por conta da incerteza sobre a estimativa de cada um dos parâmetros (assunto discutido em aula). Contudo, é comum que pessoas façam uma avaliação superficial, comparando os intervalos de confiança de dois ou mais grupos. Note que isso foi sugerido na região "nota" dessa tabela (que foi extraída na íntegra do sistema tabnet.saude.prefeitura.sp.gov.br. Mesmo sabendo desse problema, faça como o sugerido na nota, e compare os intervalos da Regional Centroeste e Leste. Em seguida avalie qual das afirmativas a seguir é VERDADEIRA.

1  pontos

- ☐ A região Leste da cidade de São Paulo, detém uma proporção maior das pessoas com o diabetes mellitus porque a média observada para essa região foi de 20,65%, enquanto a média observada para a região Centroeste foi de 13,41%.
- ☐ A região Leste da cidade de São Paulo, detém uma proporção maior das pessoas com o diabetes mellitus porque a média observada para essa região foi de 20,65%, valor esse maior que o limite superior da região Centroeste (18,90%).
- ☒ Não é possível saber qual das duas regionais apresenta maior proporção de pessoas com o diabetes mellitus, porque há sobreposição dos intervalos de confiança. ✓
- ☐ Todas as afirmativas anteriores estão erradas

 Adicionar feedback da resposta

Feedback lista 5

INQUÉRITO DE SAÚDE DO MUNICÍPIO DE SÃO PAULO - ISA CAPITAL 2015 - DIABETES MELLITUS

% Coord Regional de Saúde segundo Coord Regional de Saúde
Período: 2015

Coord Regional de Saúde	% Coord Regional de Saúde	IC95%
Coord Regional de Saúde	% Coord Regional de Saúde	IC95%
Coord Regional de Saúde	% Coord Regional de Saúde	IC95%
TOTAL	100,00	(100,00-100,00) CV=...
Centroeste	13,41	(7,93-18,90) CV=0,2086
Leste	20,65	(13,63-27,66) CV=0,1734

Nos resultados do ISA Capital é importante estar atento que na comparação das prevalências devem ser consideradas *diferenças significativas* quando não houver sobreposição dos respectivos intervalos de confiança e *sem diferença* quando um dos intervalos de confiança for parcial ou totalmente englobado pelo outro.

O coeficiente de variação é uma forma de expressar a variabilidade dos dados e é o resultado do quociente do desvio padrão pela média. CV maior do que 0,3 (30%) indica que a dispersão dos dados é muito alta e que a estimativa não tem confiança estatística, mesmo que o valor (n) na célula seja maior que 30.

☒ Selecione as respostas corretas:

Pergunta 4. Quando comparamos intervalos de confiança, existe um problema de aumento do erro por conta da incerteza sobre a estimativa de cada um dos

- ☐ mellitus porque a média observada para essa região foi de 20,65%, valor esse maior que o limite superior da região Centroeste (18,90%)
- ☒ Não é possível saber qual das duas regionais apresenta maior proporção de pessoas com o diabetes mellitus, porque há sobreposição dos intervalos de confiança. ✓
- ☐ Todas as afirmativas anteriores estão erradas

 [Adicionar feedback da resposta](#)

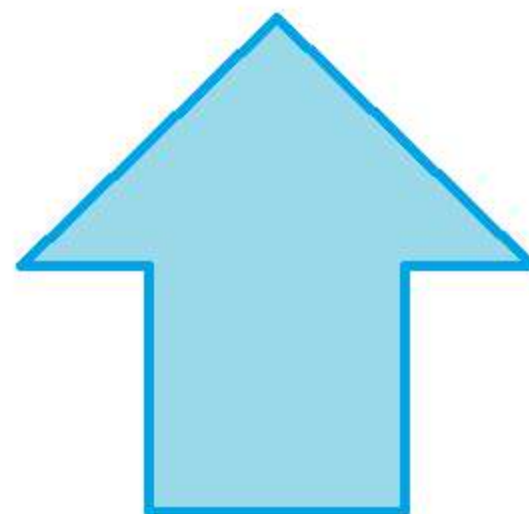
Coord Regional de Saúde	% Coord Regional de Saúde	IC _{95%}
TOTAL	100,00	(100,00-100,00) CV=...
Centroeste	13,41	(7,93-18,90) CV=0,2086
Leste	20,65	(13,63-27,66) CV=0,1734

Centroeste

7,93 ————— **18,90**

Leste

13,63 ————— **27,66**



Feedback lista 5

Para as perguntas 6 e 7, suponha o seguinte cenário. Foi realizado um estudo que buscou avaliar o peso de bebês recém nascidos em diferentes cidades do estado de São Paulo. Os resultados observados desse estudo são apresentados na tabela a seguir:

Cidade	Peso médio (g)	Desvio-padrão
Santos	3000	100
Sorocaba	3000	250
Ribeirão Preto	3000	500

☒ Selecione as respostas corretas:

Pergunta 7. Em qual das cidades foi feito o registro do nascimento do bebê mais pesado (obrigatoriamente)?

 pontos

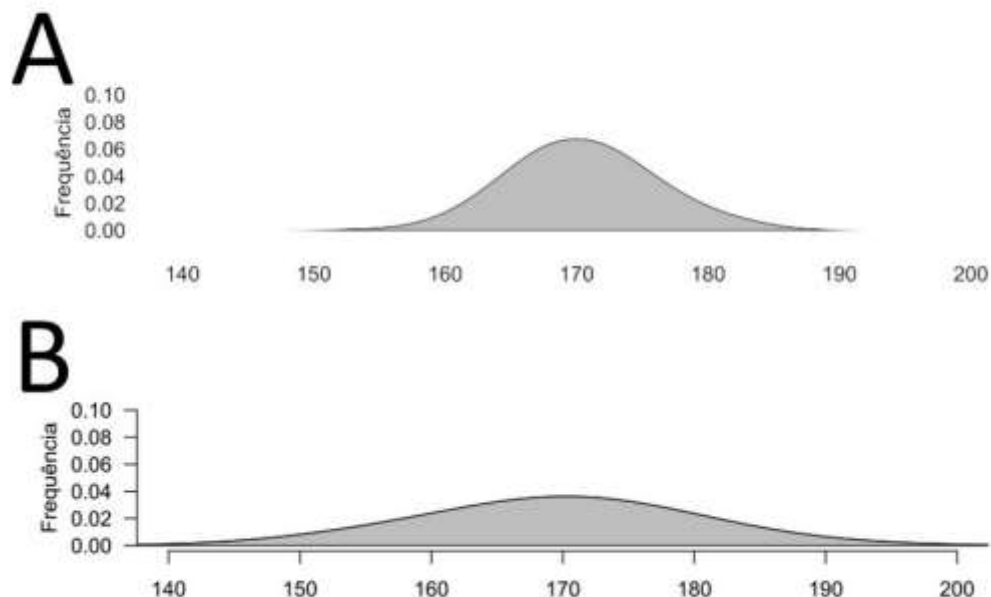
- ☐ Santos
- ☐ Sorocaba
- ☐ Ribeirão Preto
- ☐ A chance de encontrar um bebê com qualquer peso é igual, porque as médias são iguais.
- ☒ Não é possível responder essa pergunta, somente com os dados apresentados. ✓

 Adicionar feedback da resposta

Concluído

Feedback lista 5

Pergunta 5. Observe os gráficos de distribuição de frequência a baixo. Em qual dos casos aparentemente há um menor desvio-padrão?



☒ Selecione as respostas corretas:

Pergunta 7. Em qual das cidades foi feito o registro do nascimento do bebê mais pesado (obrigatoriamente)?

1 pontos

- ☐ Santos
- ☐ Sorocaba
- ☐ Ribeirão Preto
- ☐ A chance de encontrar um bebê com qualquer peso é igual, porque as médias são iguais.

☒ Não é possível responder essa pergunta, somente com os dados apresentados.



 Adicionar feedback da resposta

Concluído

OBS: a figura da pergunta 5 era uma dica para como pensar na pergunta 7



Feedback lista 5

Tabela de contingencia relacionando dados de diabetes e prática de atividade física

	Diabetes		Total
	Sim	Não	
Pratica atividade física	72	179	225
Não pratica atividade física	95	204	325
Total	167	383	550

☒ Selecione as respostas corretas:

Pergunta 9. Probabilidade de uma pessoa praticar atividade física.

  pontos

☐ $72/167 = 0,43$

☐ $72/225 = 0,32$

☐ $72/550 = 13,09$

☐ $167/550 = 30,36$

☒ $225/550 = 0,41$



 [Adicionar feedback da resposta](#)

Concluído


Feedback lista 5

Tabela de contingencia relacionando dados de diabetes e prática de atividade física

	Diabetes		Total
	Sim	Não	
Pratica atividade física	72	179	225
Não pratica atividade física	95	204	325
Total	167	383	550

☒ Selecione as respostas corretas:

Pergunta 9. Probabilidade de uma pessoa praticar atividade física.

 pontos

☐ $72/167 = 0,43$

☐ $72/225 = 0,32$

☐ $72/550 = 13,09$

☐ $167/550 = 30,36$

☒ $225/550 = 0,41$



 Adicionar feedback da resposta

Concluído

Feedback lista 5

Tabela de contingencia relacionando dados de diabetes e prática de atividade física

	Diabetes		Total
	Sim	Não	
Pratica atividade física	72	179	225
Não pratica atividade física	95	204	325
Total	167	383	550

☒ Selecione as respostas corretas:

Pergunta 10. Probabilidade de uma pessoa ter diabetes, sabendo que ela pratica atividade física.

pontos

☐ $72/167 = 0,43$

☐ $225/550 = 0,41$

☒ $72/225 = 0,32$

☐ $72/550 = 13,09$

☐ $167/550 = 30,36$

Adicionar feedback da resposta

Concluido

Feedback lista 5

Tabela de contingencia relacionando dados de diabetes e prática de atividade física

	Diabetes		Total
	Sim	Não	
Pratica atividade física	72	179	225
Não pratica atividade física	95	204	325
Total	167	383	550

Pergunta 11. Probabilidade de uma pessoa praticar atividade física, sabendo que ela tem diabetes.

1 pontos

☒ $72/167=0,43$



☐ $225/550 = 0,41$

☐ $72/225 = 0,32$

☐ $72/550 = 13,09$

☐ $167/550 = 30,36$

 Adicionar feedback da resposta

Concluido

Feedback lista 6

- A lista 6 não teve questões com alta taxa de erros!

Revisão última aula

- H_0 = Hipótese nula
- H_1 = Hipótese alternativa

Revisão última aula

1. Você aplica um teste
2. Observe o valor de p
3. Interprete o resultado
 - **Se o valor de p calculado for maior que o valor que o nível de significância** (margem que você assumiu de erro): você assume que H_0 é verdadeira.
 - **Se o valor de p calculado for menor que o valor que o nível de significância** (margem que você assumiu de erro): você rejeita H_0 e assume que H_1 é verdadeiro.

Revisão última aula

Erro tipo 1



Erro tipo 2



Aula de hoje

- Hoje vamos comparar médias entre 2 grupos
- Exemplo: Qual sexo tem maior expectativa de vida?

O que você precisa saber para escolher sua análise?

- Quais e quantas são suas variáveis preditoras e variáveis resposta
- Qual é o tipo de variável envolvida
 - Identifique se você tem variáveis categóricas (grupos) ou contínuas (números)

Aula de hoje

Variável resposta

Variável preditora

Expectativa de vida



Sexo

Em outras palavras, sexo determina a expectativa de vida

Aula de hoje

Contínua

Categórica

Expectativa de vida

Sexo

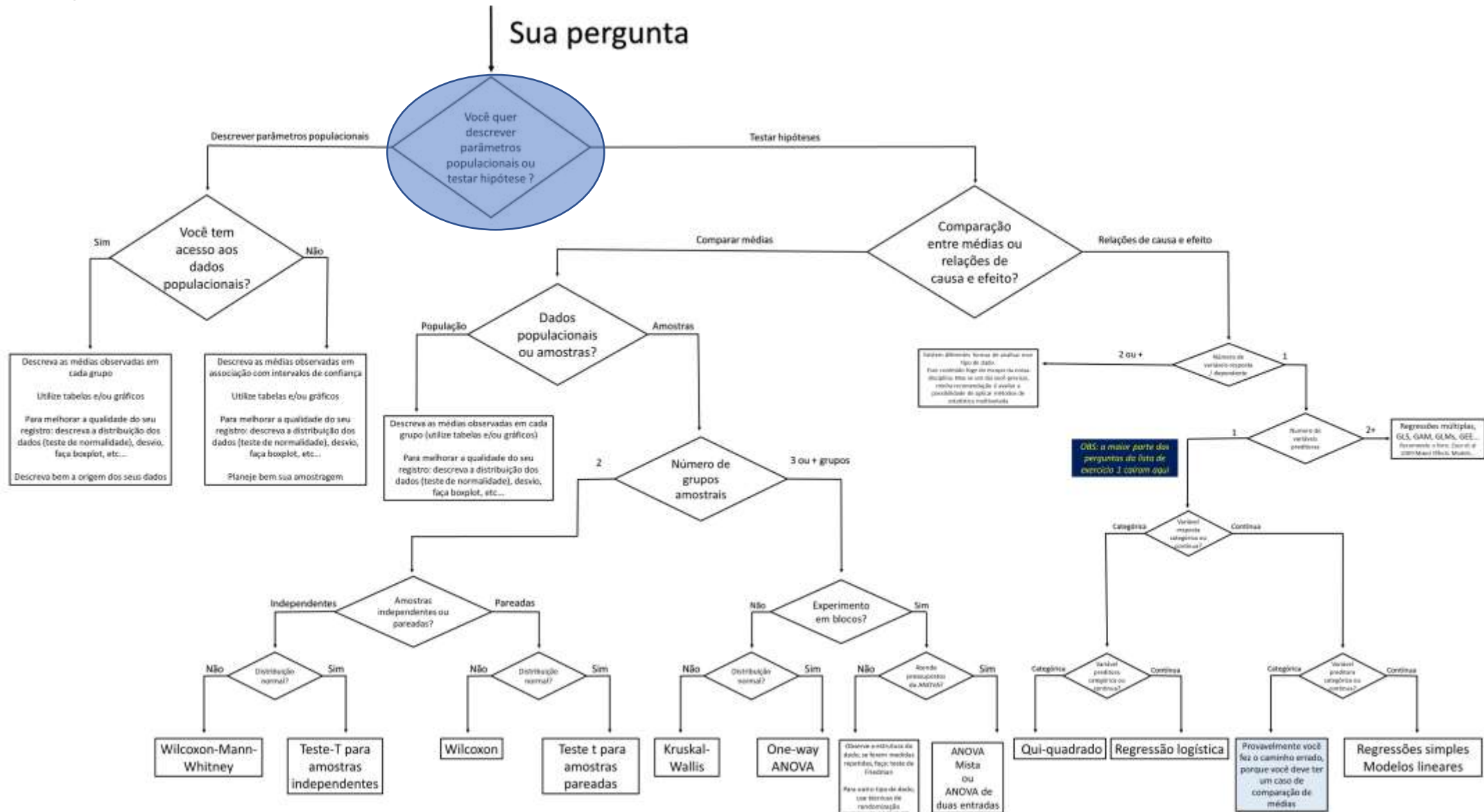
Em outras palavras, sexo determina a expectativa de vida

- Nesse caso, a expectativa de vida de qualquer ser humano, nada mais é que a média de expectativa de vida de outros seres humanos parecidos com ele (seja em relação a sexo ou a país de residência, condição socioeconômica, etc...)
- No fim, isso se trona uma comparação de médias.

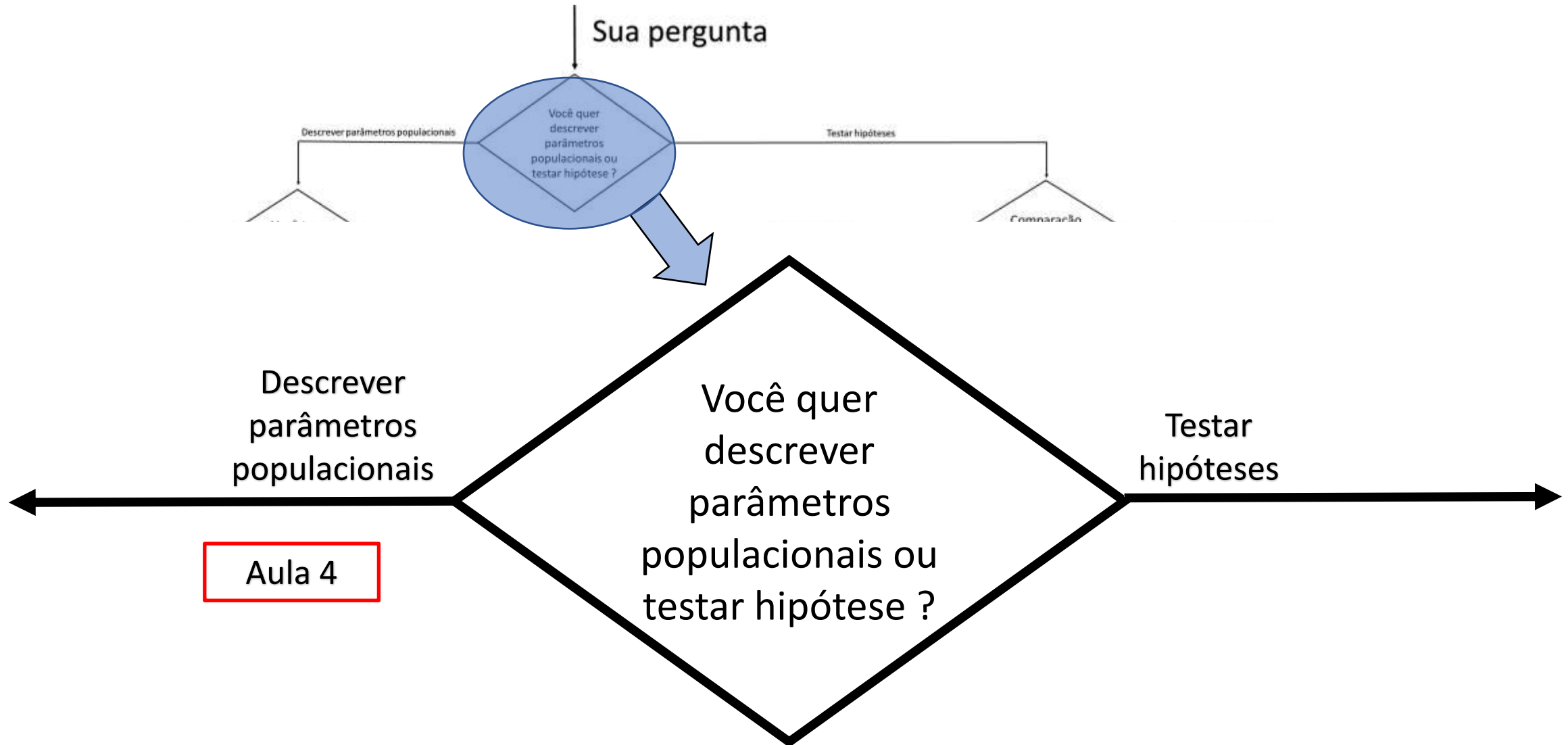
Mapa para seleção do método estatístico



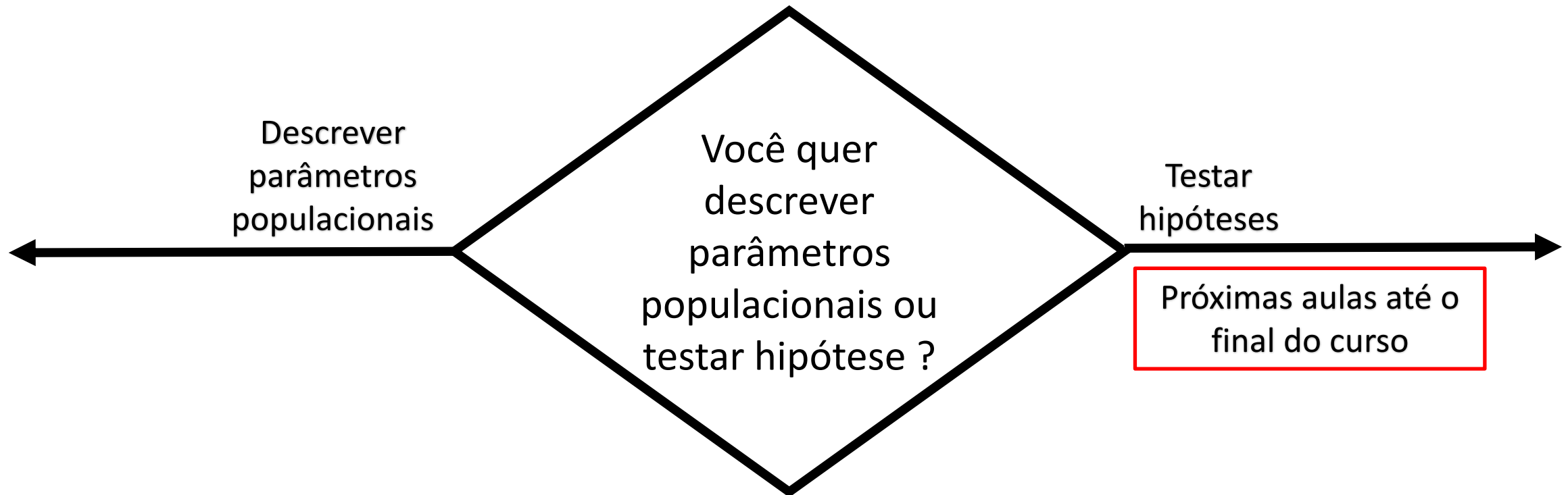
Tópicos dessa semana



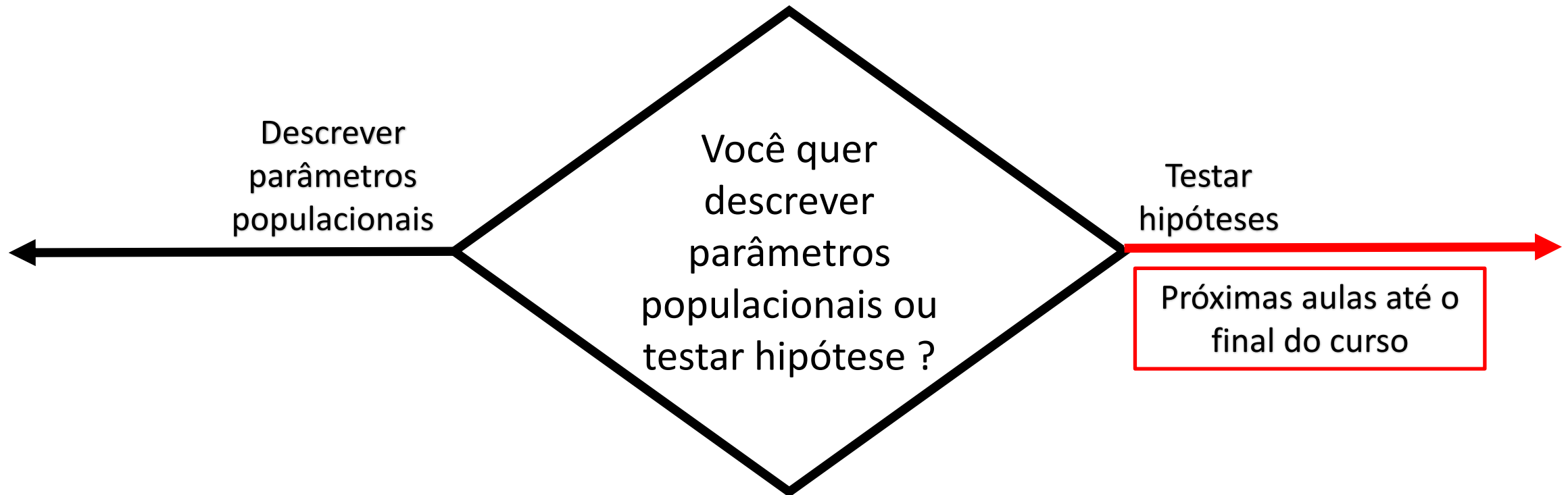
Tópicos dessa semana



Tópicos dessa semana



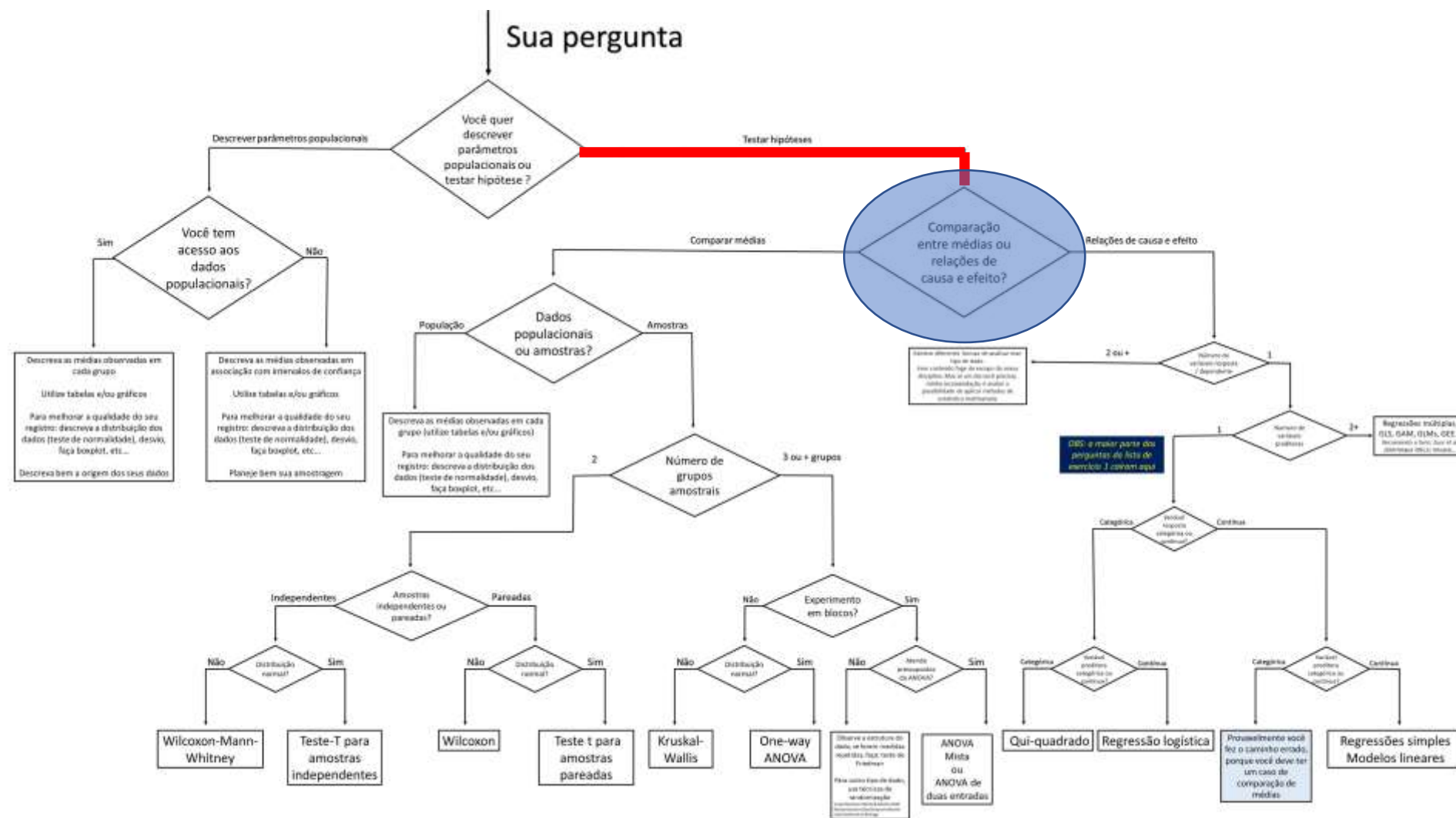
Tópicos dessa semana



Tópicos dessa semana



Tópicos dessa semana



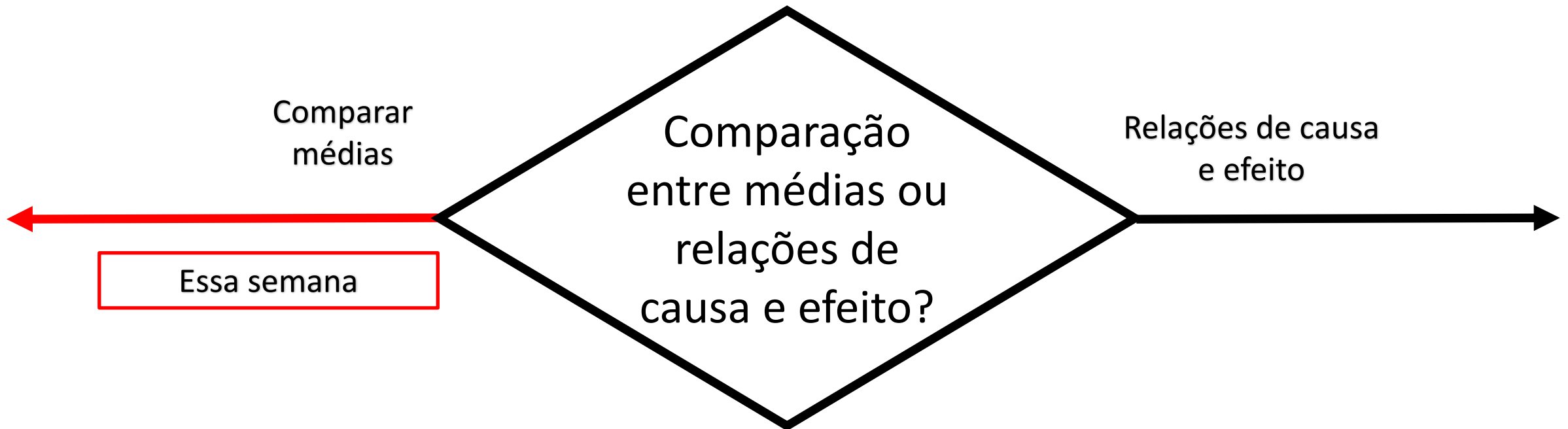
Tópicos dessa semana



Tópicos dessa semana



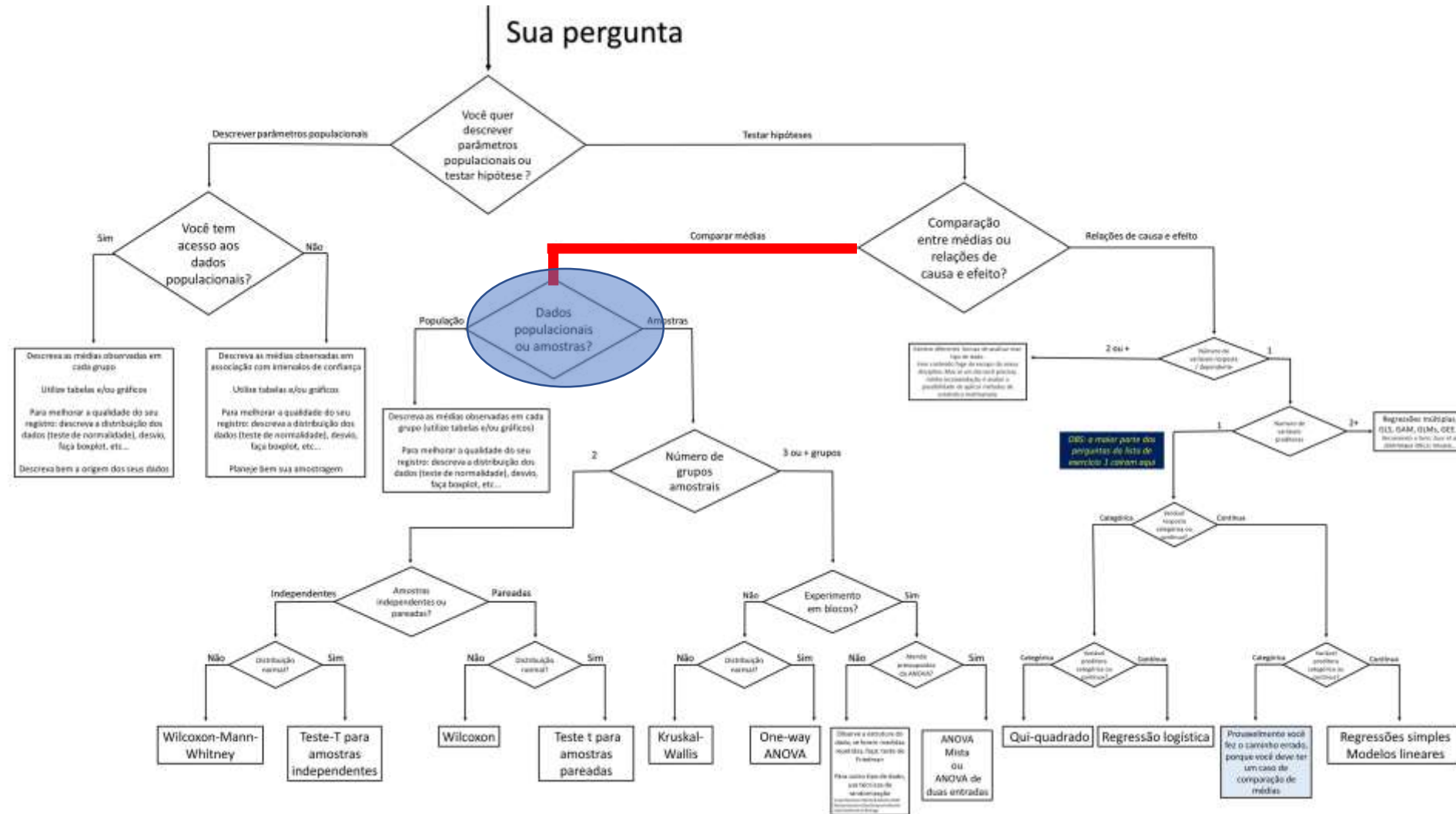
Tópicos dessa semana



Tópicos dessa semana



Tópicos dessa semana



Tópicos dessa semana



Tópicos dessa semana



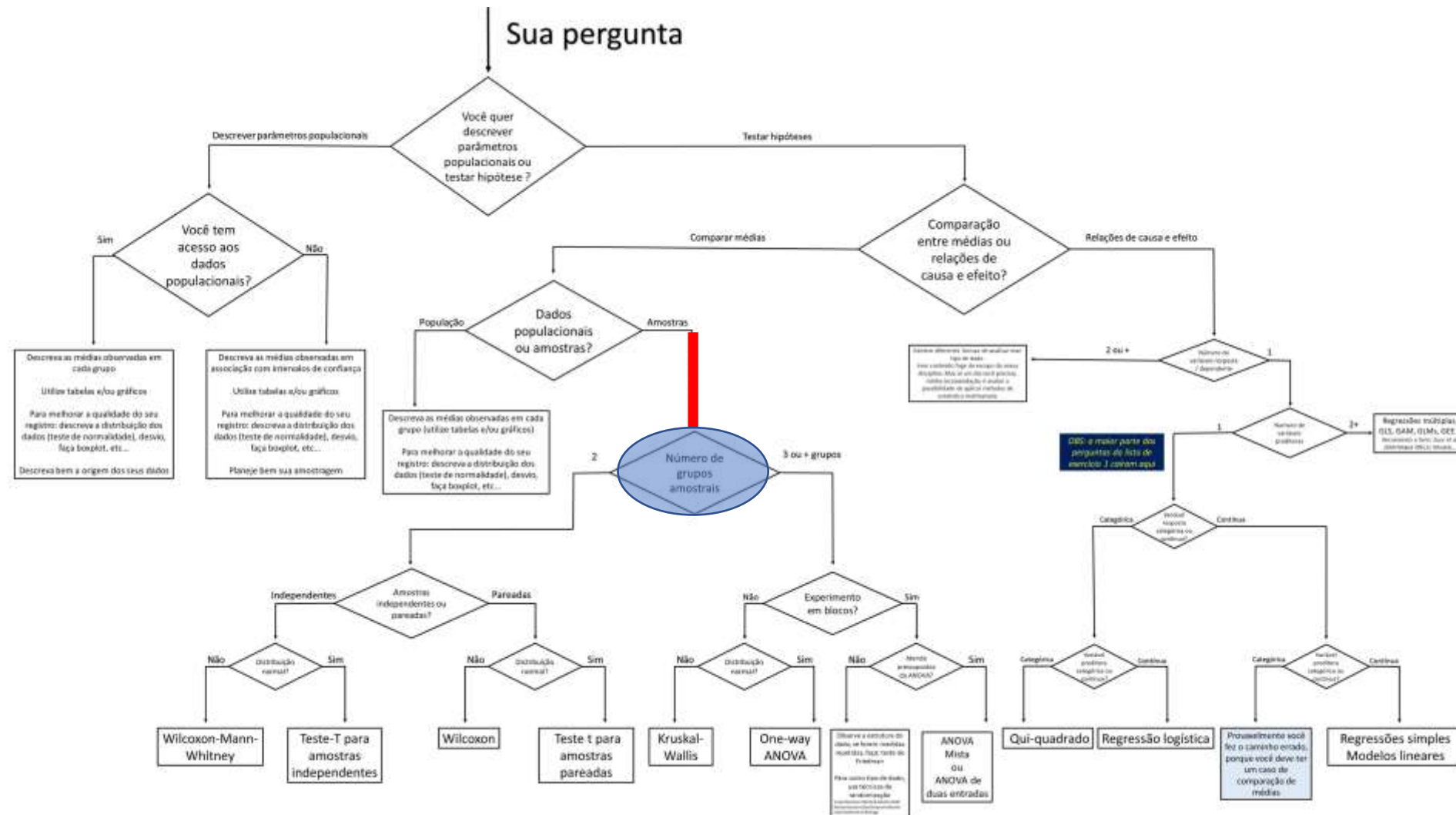
Tópicos dessa semana



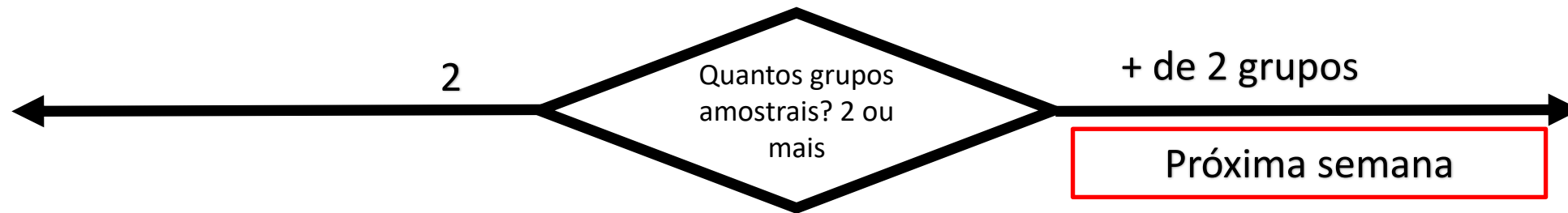
Tópicos dessa semana



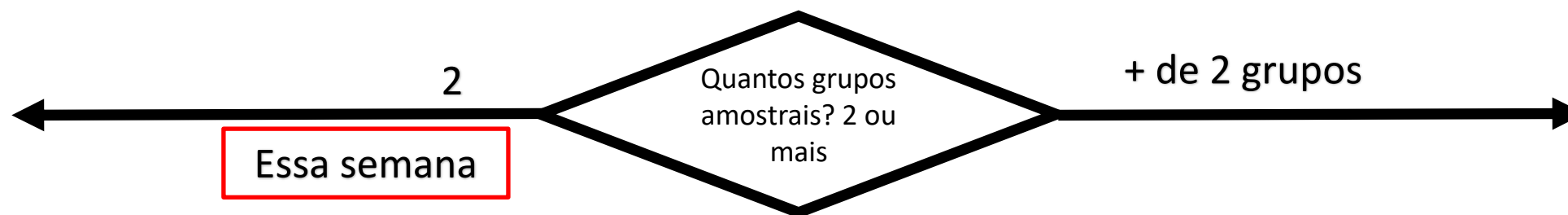
Tópicos dessa semana



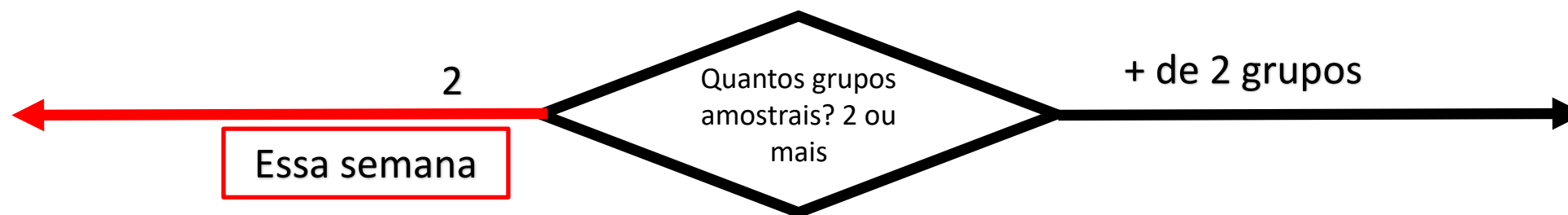
Tópicos dessa semana



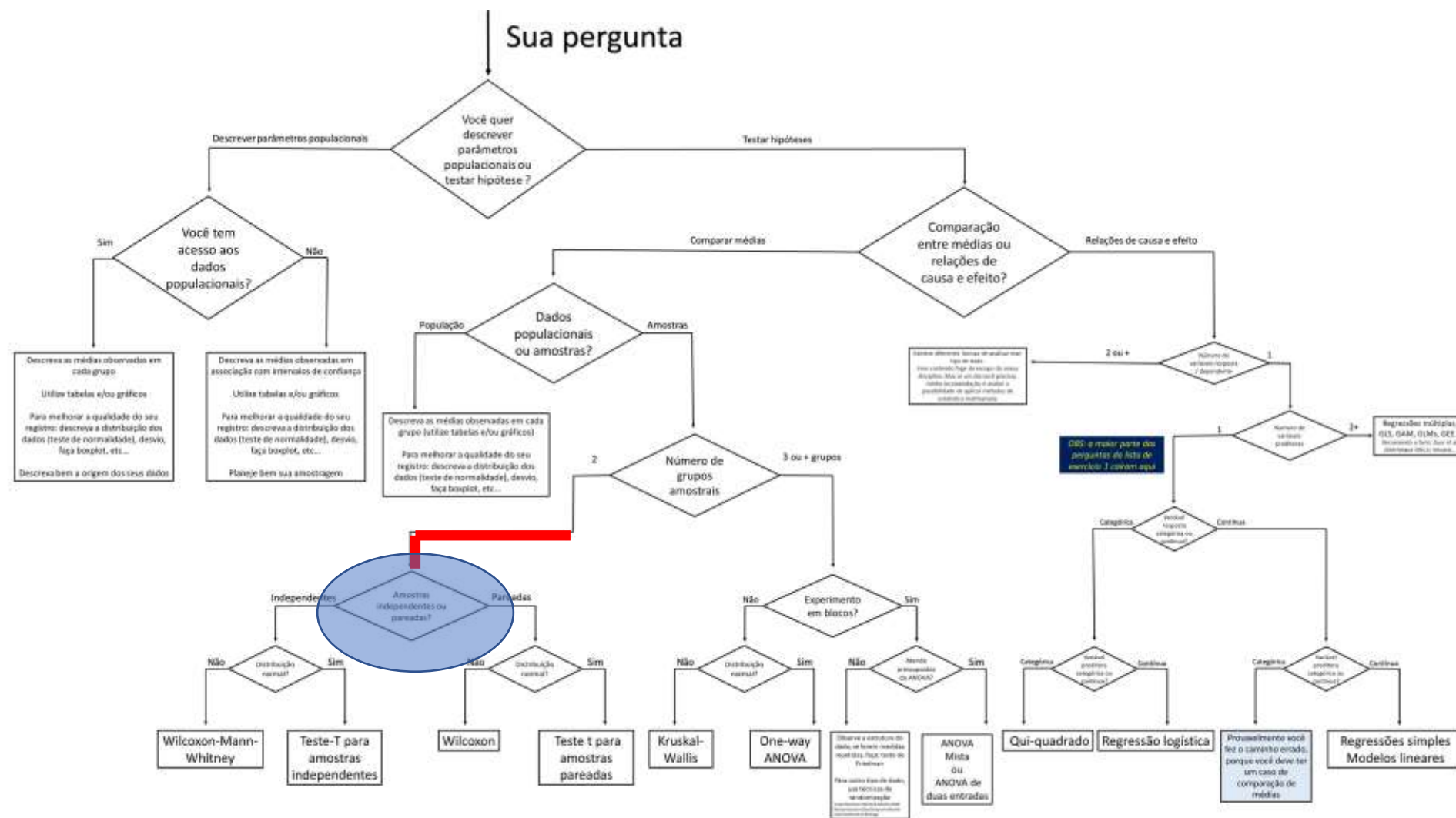
Tópicos dessa semana



Tópicos dessa semana



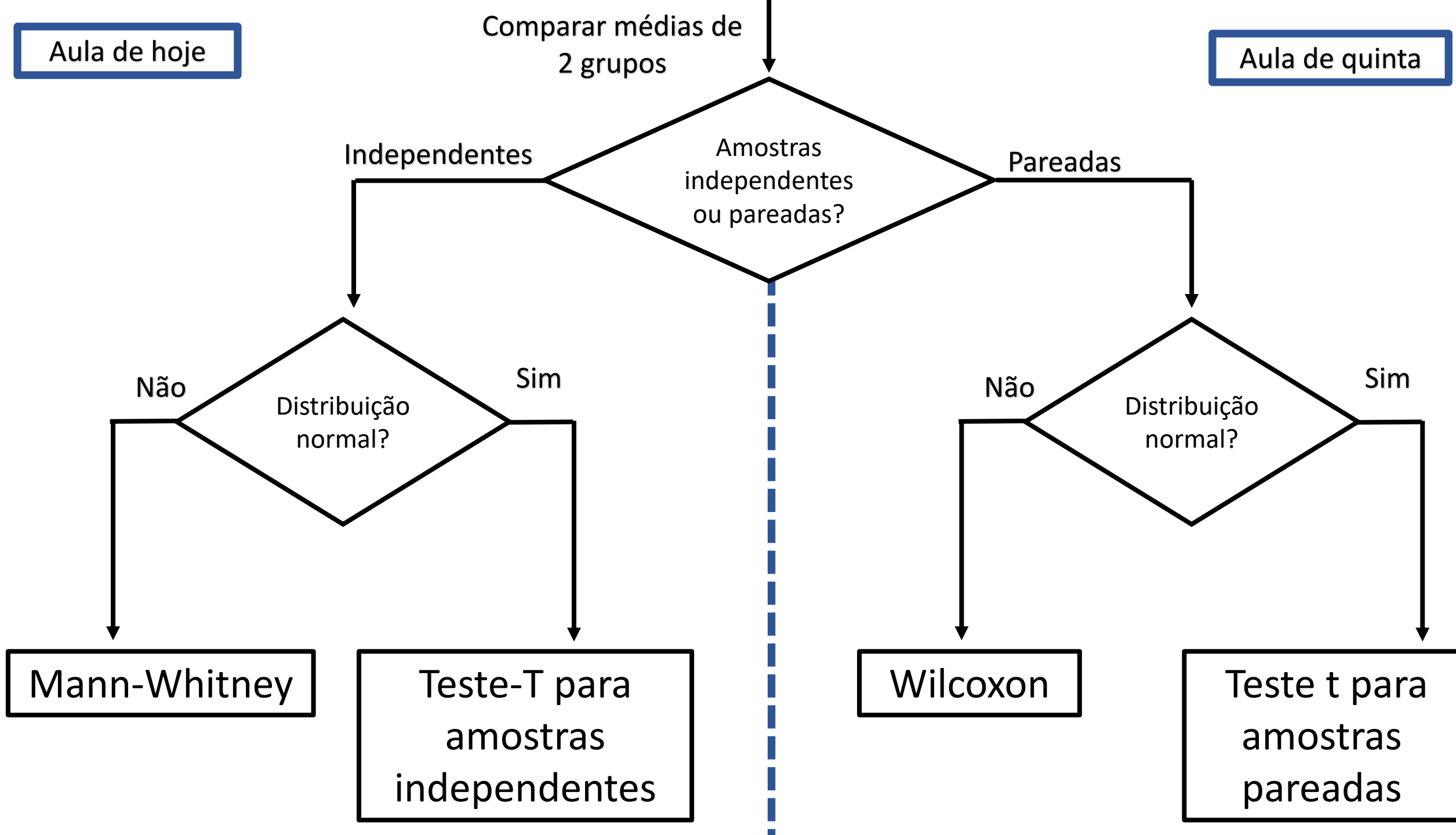
Tópicos dessa semana



Aula de hoje

Comparar médias de
2 grupos

Aula de quinta



Histórico do teste-t

- O teste t de Student foi criado por William Sealy Gosset (químico e matemático)
- O teste foi usado para a quantidade de leveduras no processo de fermentação da cerveja Guinness
- O pseudônimo Student foi usado porque os funcionários da marca de cerveja eram proibidos de divulgar suas descobertas



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

- 1) Existe diferença da nota escolar média de crianças que receberam leite materno 6 meses (ou mais) e das demais? (Jozielda)
- 2) Níveis plasmáticos de cortisol (medida de estresse) diferem entre pessoas que residem em cidades grandes (+de 500 mil habitantes) de residentes de cidades menores? (Maryna Luz)
- 3) O número de dias para cicatrização de uma ferida é menor entre os pacientes que rebem troca diária do curativo? (Thayla)
- 4) Cidades que implementam medidas de proteção a pessoas em situação de rua, apresentam menores taxas de homicídio que cidades que não implementam essas medidas? (William)

Exemplos de adaptações de perguntas: identificar variáveis

- Existe diferença da nota escolar média de crianças que receberam leite materno 6 meses (ou mais) e das demais?
 - Nota escolar:
 - variável resposta
 - quantitativas contínua (número de 0-100%)
 - Receber leite materno por 6 meses:
 - variável preditora
 - qualitativa (Sim ou Não)

Exemplos de adaptações de perguntas: identificar variáveis

- **Níveis plasmáticos de cortisol** (medida de estresse) diferem entre pessoas que **residem em cidades grandes (+de 500 mil habitantes)** de **residentes de cidades menores**?
 - Níveis plasmáticos de cortisol
 - variável resposta
 - quantitativas contínua
 - Ser residem de uma cidade grande
 - variável preditora
 - qualitativa (Sim ou Não)

Exemplos de adaptações de perguntas: identificar variáveis

- O número de dias para cicatrização de uma ferida é menor entre os pacientes que **rebem troca diária do curativo**?
 - O número de dias para cicatrização de uma ferida
 - variável resposta
 - quantitativas contínua
 - Receber troca diária do curativo
 - variável preditora
 - qualitativa (Sim ou Não)

Exemplos de adaptações de perguntas: identificar variáveis

- Cidades que **implementam medidas de proteção a pessoas em situação de rua**, apresentam menores **taxas de homicídio** que cidades que não implementam essas medidas?
 - Taxas de homicídio:
 - variável resposta
 - quantitativas contínua
 - Implementar medidas de proteção a pessoas em situação de rua
 - variável preditora
 - qualitativa (Sim ou Não)

Desenho experimental

- Para comparar valores de média de 2 grupos, você deve usar alguma técnica para obter dados de cada um dos grupos que estão comparados. Exemplos:
 - a. Registrar a nota de 50 crianças que receberam aleitamento por mais de 6 meses, e registrar a nota 50 que não receberam aleitamento por mais de 6 meses.
 - b. Registrar o nível plasmáticos de cortisol em pessoas que habitam cidades grande e saber o nível de pessoas em cidades pequenas
 - c. Registro de numero de dias que levou para cicatrizar uma ferida de pacientes que tiveram troca diária de curativo e que não tiveram troca diária
 - d. Registro de taxa de homicídio em cidades que adotaram ou não medidas de proteção a pessoas em situação de rua

Hipóteses

Teste - T

- H_0 : Não há diferenças entre as médias comparadas
- H_1 : Há diferença entre as médias comparadas

Mann-Whitney

- H_0 : Não há diferenças entre as medianas dos grupos comparadas
- H_1 : Há diferença entre as medianas dos grupos comparadas

Como fica uma tabela desse tipo de dado?

Nome	Sexo	Altura
Aline	Feminino	160
Ana	Feminino	163
Antônio	Masculino	172
Fernanda	Feminino	180
Francisco	Masculino	185
João	Masculino	165
José	Masculino	160
Juliana	Feminino	175
Maria	Feminino	159
Paulo	Masculino	170

Identificação do
dado

Variável categórica

Variável numérica

Qual a unidade desse tipo de dado?

Nome	Sexo	Altura
Aline	Feminino	160
Ana	Feminino	163
Antônio	Masculino	172
Fernanda	Feminino	180
Francisco	Masculino	185
João	Masculino	165
José	Masculino	160
Juliana	Feminino	175
Maria	Feminino	159
Paulo	Masculino	170

Amostras independentes

Amostras independentes

- Neste caso, você coletou dados uma única vez de cada unidade amostral.
 - Exemplo. Você mediu o nível de estresse de cada participante da pesquisa uma única vez
- Cada medida feita em um indivíduo é independente de todas as outras medidas.
 - Exemplo. Os eventos que influenciaram o estresse das pessoas de cidade grande é independente dos eventos que influenciaram o estresse das pessoas do interior.
- Espera-se que a única coisa que influência no parâmetro que você está avaliando seja a variável preditora (e o erro inerente a qualquer estudo)
 - Exemplo. Esperasse que as diferenças entre níveis de estresse de pessoas de cidade grande e interior, seja explicado pelo fato da pessoa viver na cidade grande vs. viver na cidade pequena.

Pressupostos do teste-t

- Nós vamos avaliar pressupostos para executar um teste t.
- Se os dados atenderem todos os pressupostos você usa o teste t.
- Se os dados não atenderem os pressupostos você usa o teste de Mann-Whitney ou Wilcoxon (próxima aula)

O que você vai precisar saber o dia de hoje?

- Para os testes de hoje você vai precisar saber se:
 - Amostras são pareadas ou independentes?
 - Dados tem distribuição normal?
 - As variâncias são iguais?

Combinando as possibilidades

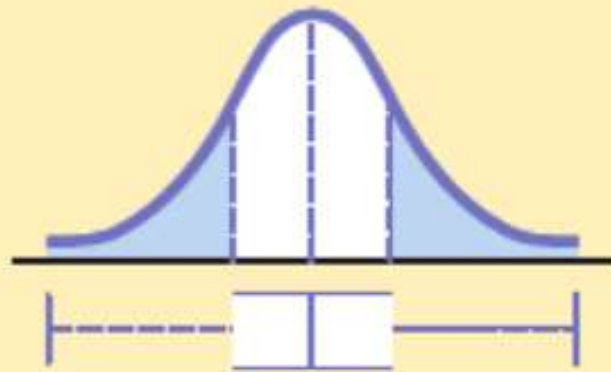
Testes da aula de hoje		
Amostras	Distribuição dos dados	Variâncias
Independentes	Normal	Iguais
Independentes	Normal	Diferentes
Independentes	Não-normal	

Próxima aula	
Amostra	Distribuição dos dados
Pareadas	Normal
Pareadas	Não-normal

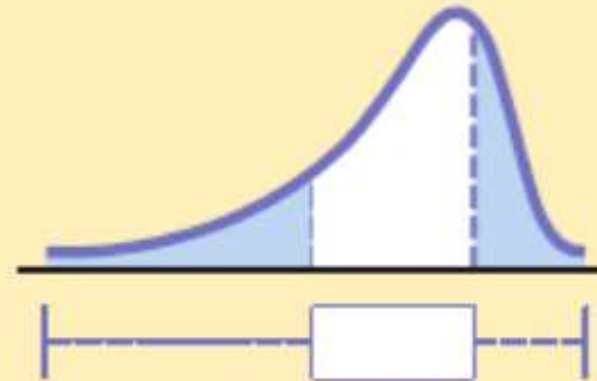
Normalidade

- Existem varias formas de avaliar normalidade nos dados
- Opção 1: inspeção visual (histogramas, boxplot, dotchart)
- Opção 2: testes estatístico (Shapiro-Wilk)

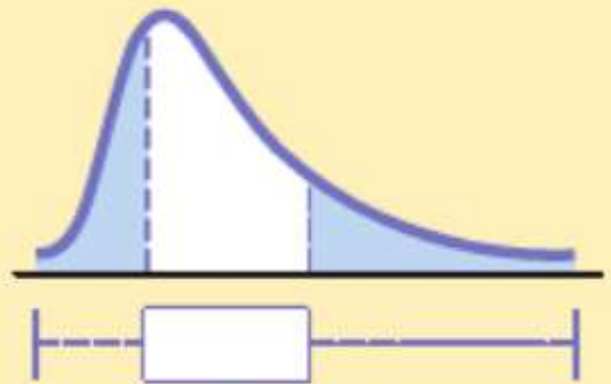
Padrões gerais de distribuição



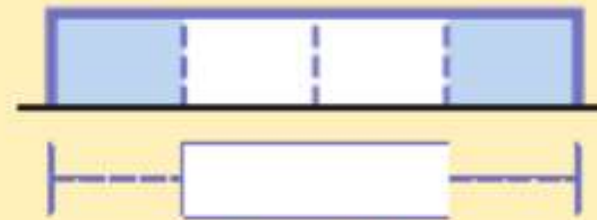
Simétrico



Assimétrico à esquerda



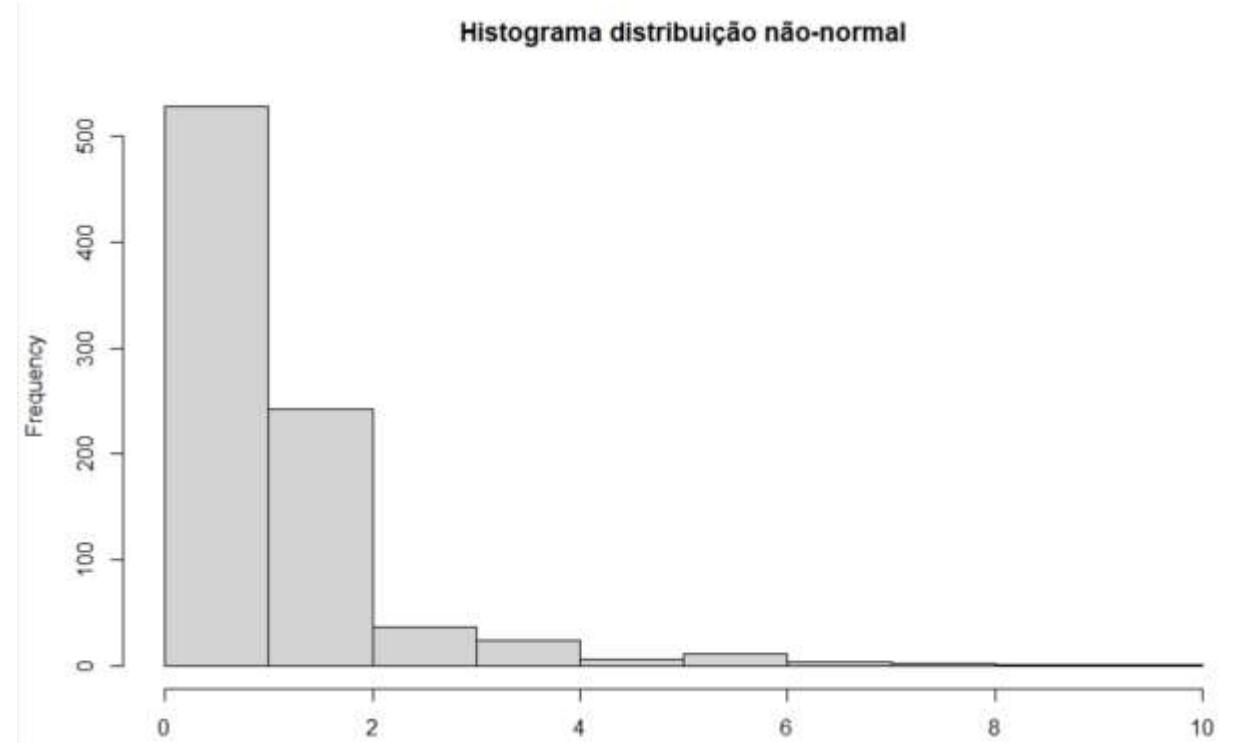
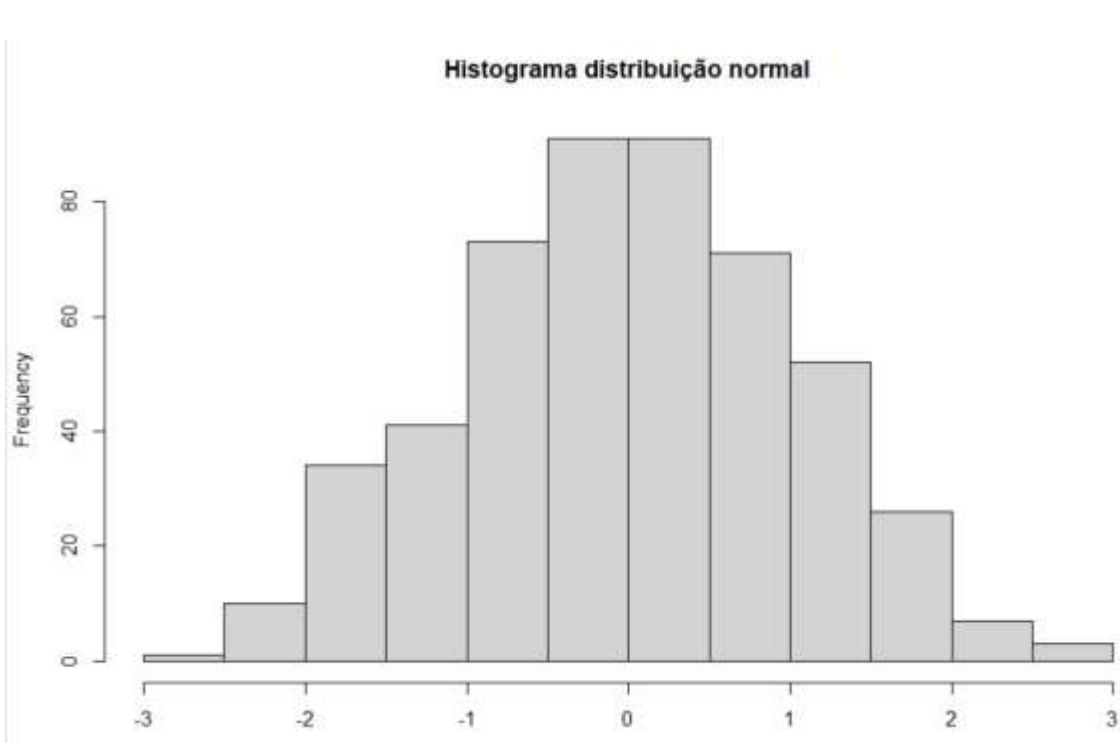
Assimétrico à direita



Distribuição retangular

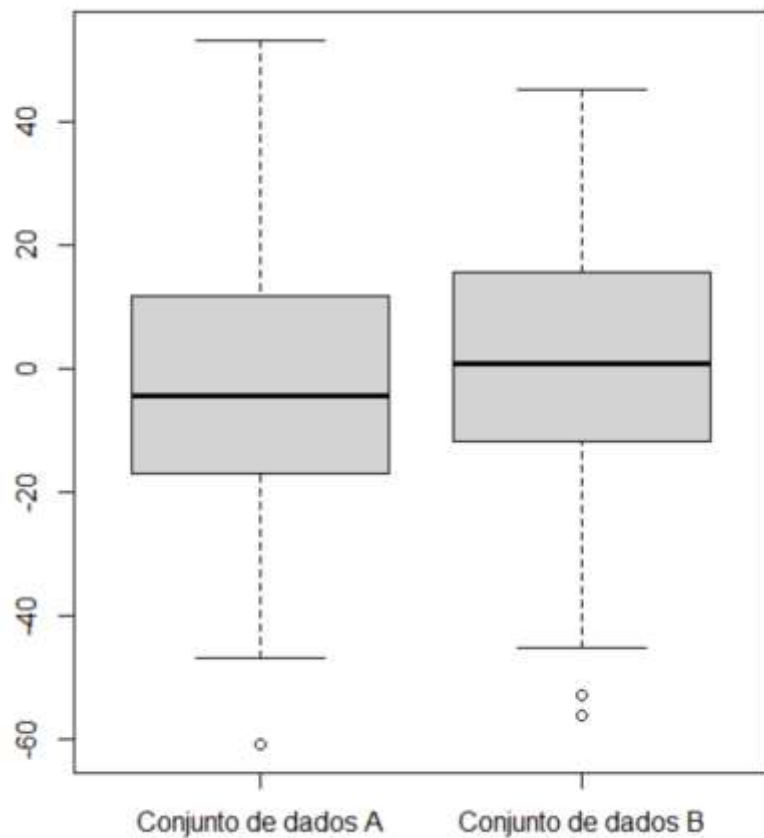
Normalidade

- Visto na atividade prática da ultima aula

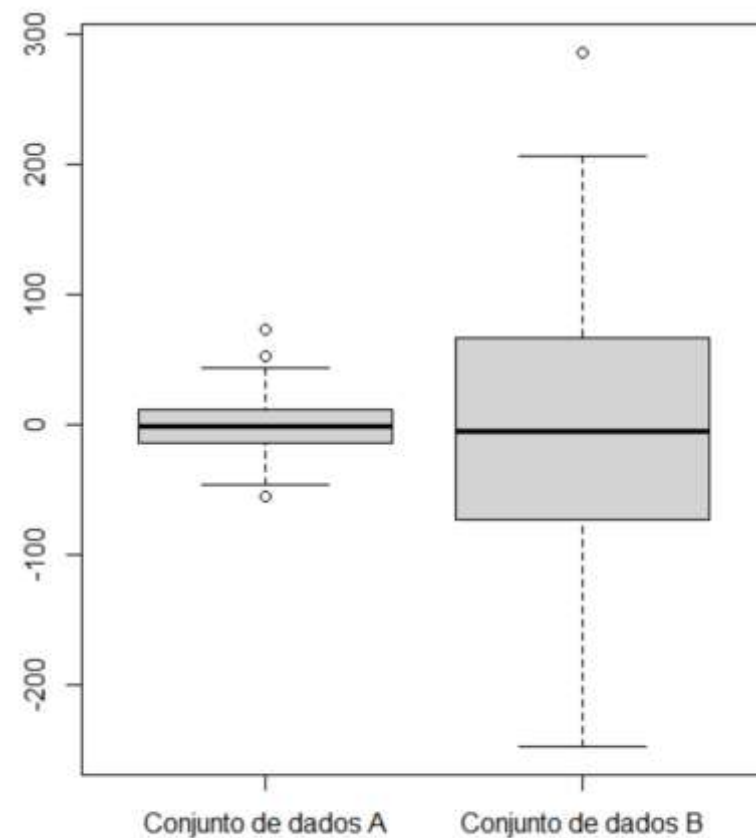


Avaliar se há diferença na homogeneidade de variâncias

Padrão que sugere variâncias iguais



Padrão que sugere variâncias diferentes



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: sexo
- Variável resposta: numero de horas do sono
- Tipo de dado:
 - Amostras independentes
 - Dados com distribuição normal
 - Variâncias iguais
- Teste T
 - Usado para compara 2 medias

Interpretação dos resultados

Welch Two Sample t-test

Resultado: P maior que 0.05
Conclusão: H_0 é verdadeira

```
data: dados_simulados[, 1] by Sexo
t = -1.636, df = 96.699, p-value = 0.1051
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.52049733  0.05014333
sample estimates:
mean in group F mean in group M
      7.854535      8.089712
```

Prática – 2 Existe diferença na altura de mulheres e homens?

- Tabela “Dados para aula 7 e 8.xlsx”
- Variável preditora: sexo
- Variável resposta: altura
- Tipo de dado:
 - Amostras independentes
 - Dados com distribuição normal
 - Variâncias diferentes
- Teste T com variâncias diferentes
 - Usado para compara 2 medias

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

Prática – 3 Existe diferença no peso de mulheres e homens?

- Tabela “Dados para aula 7 e 8.xlsx”
- Variável preditora: sexo
- Variável resposta: peso
- Tipo de dado:
 - Amostras independentes
 - Dados com distribuição não-normal
 - Variâncias diferentes
- Teste de Wilcoxon-Mann-Whitneys
 - Usado para compara 2 medianas
- OBS: Teste T de variância agrupada (sinônimo de iguais) pode ser usado mesmo quando os dados não tenham distribuição normal, desde que você tenha um N grande (maior que 30) para cada grupo (Pagina 346, Levine et al 2012. Estatística: teoria e aplicações)

Interpretação dos resultados

Resultado: P menor que 0.05
Conclusão: H_1 é verdadeira

wilcoxon rank sum test with continuity correction

data: tabela1\$Peso by tabela1\$Sexo

w = 796, p-value = 0.00177

alternative hypothesis: true location shift is not equal to 0