

ESTATÍSTICA PARA SAÚDE COLETIVA

Aula 11

Seminários de hoje

Nomes (Aluno ou Dupla)	Nome do artigo
ALEX SILVA + WILLIAM	Quedas entre idosos brasileiros residentes em áreas urbanas: ELSI-Brasil
ELZA DA SILVA	Carga de trabalho de enfermagem em unidade de terapia intensiva destinada a pacientes com queimaduras
LETICIA GOMES	Clima organizacional e trabalho em equipe na estratégia saúde da família
THAIS RIBEIRO	Estudo quanti-qualitativo sobre amamentação exclusiva por gestantes de alto risco

Próxima aula

Nomes (Aluno ou Dupla)	Data seminário	Nome do artigo
CLAUDIA MARTINS	08/10/2020	Análise Espacial dos Casos de COVID-19 e leitos de terapia intensiva no estado do Ceará, Brasil
CLAUDINEIA SOARES TORRES	08/10/2020	Estimativas de impacto do Covid-19 na mortalidade de idosos institucionalizados no Brasil
PRISCILA PRATES	08/10/2020	Promoção comercial ilegal de produtos que competem com o aleitamento materno

Revisão + feedback lista 9 e 10

ANOVA

Revisão

Variável resposta



Variável preditora

Revisão (comparação de médias)

Variável resposta

Variável preditora

Expectativa de vida



Renda, Sexo, UF

1 variável preditora

Variável resposta

Variável preditora

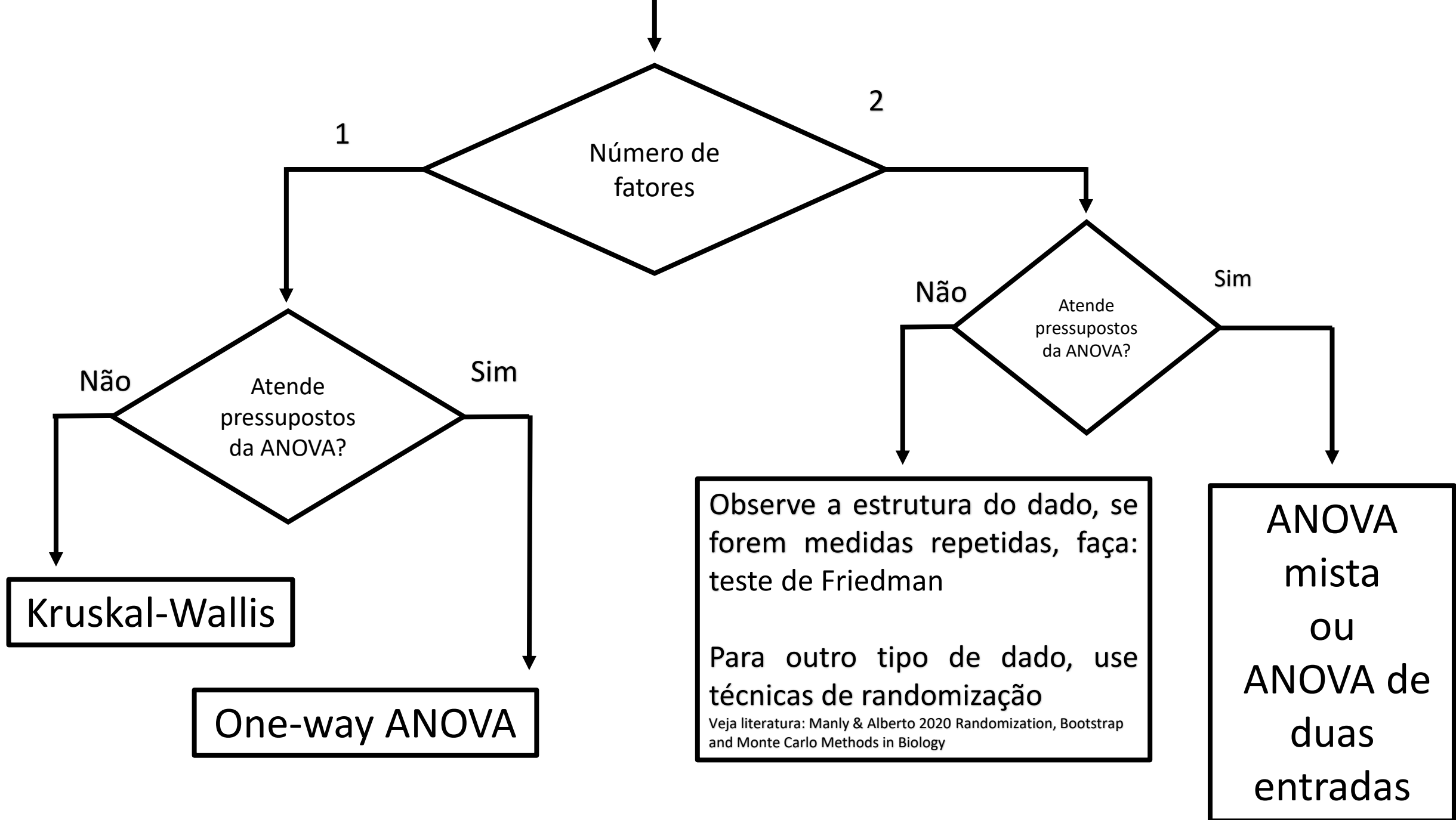
Expectativa de vida ← UF de residência

2 variáveis preditoras

Variável resposta


Variáveis preditoras


Expectativa de vida ← UF de residência + Sexo



Lista 9

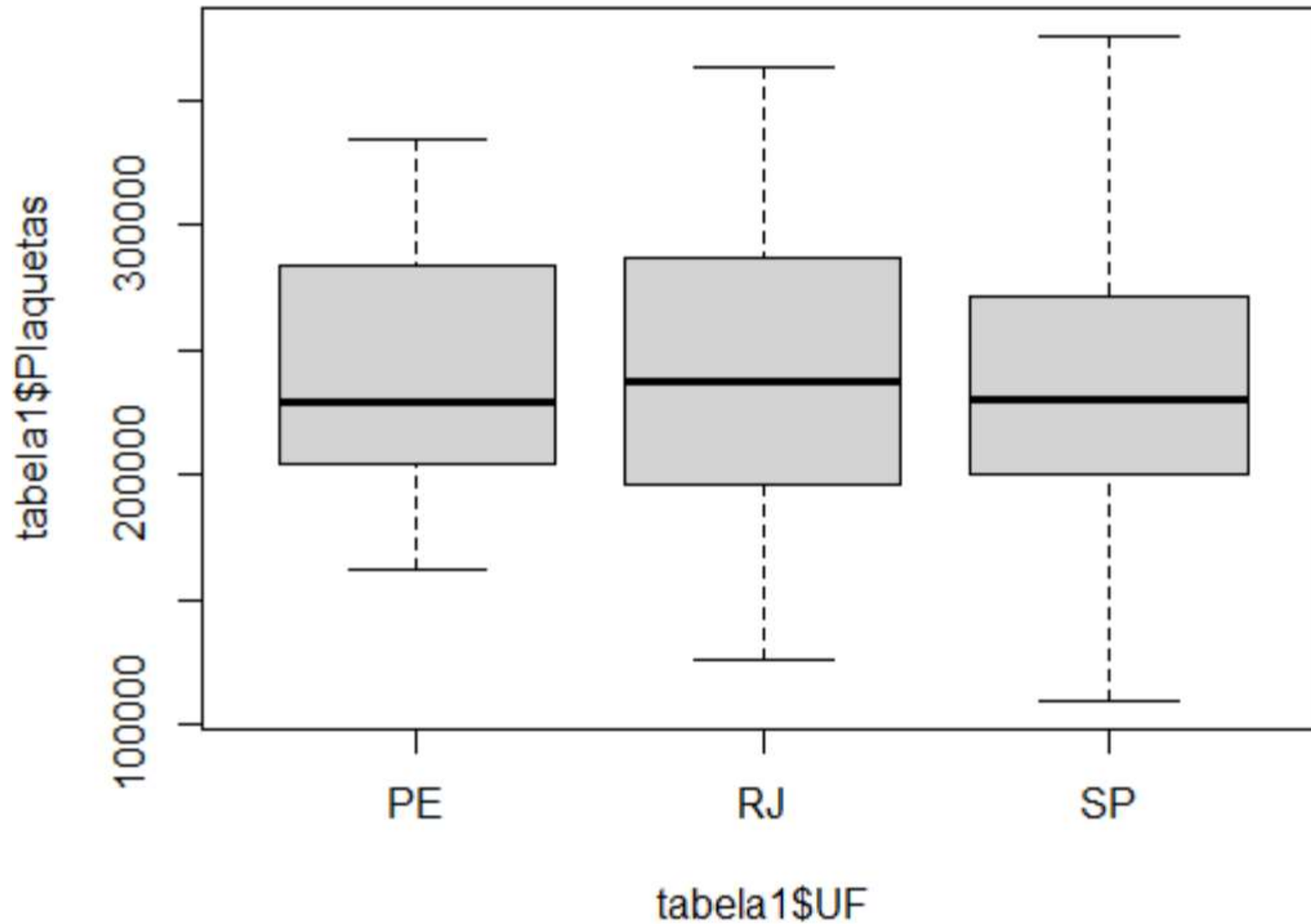
Pergunta 6. Existe evidências de que há diferença na média de plaquetas (Coluna Plaquetas) entre diferentes estados (Coluna UF)? Assuma um nível de significância de 5%. Assuma que os pressupostos da ANOVA foram atendidos.

1  pontos

- ☒ Existe (sim) evidências de que há diferença. 
- ☐ Não existe evidencias de que há diferença .
- ☐ Não é possível responder essa pergunta com os dados apresentados.
- ☐ Não consegui responder essa perguntar porque não entendi como aplicar o teste.
- ☐ Não consegui responder essa perguntar porque não tenho computador disponível para executar o teste.
- ☐ Não sei/Não quero responder essa pergunta.

 [Adicionar feedback da resposta](#)

Lista 9 – Pergunta 6



Lista 9 – Pergunta 6

```
> summary(modelo)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
tabela1\$UF	2	74	36.99	3.092	0.0462 *
Residuals	578	6914	11.96		

Resultado: P menor que 0.05
Conclusão: H_1 é verdadeira

Hipóteses testadas

H_0 : Não há diferenças entre as médias de todos os grupos comparados

H_1 : Existe diferença de pelo menos uma das médias comparadas


Portanto:


Existe evidências de ao menos 1 estado apresenta pessoas com número de plaquetas média diferente dos demais estados.

OBS: esse padrão poderia ser explicado por uso de equipamentos diferentes para fazer a contagem de plaquetas

Lista 9

Pergunta 7. Existe evidências de que há diferença no nível de insulina (Coluna Insulina) dependendo do estado (Coluna UF)? Assuma um nível de significância de 5%. Assuma que os pressupostos da ANOVA não foram atendidos.

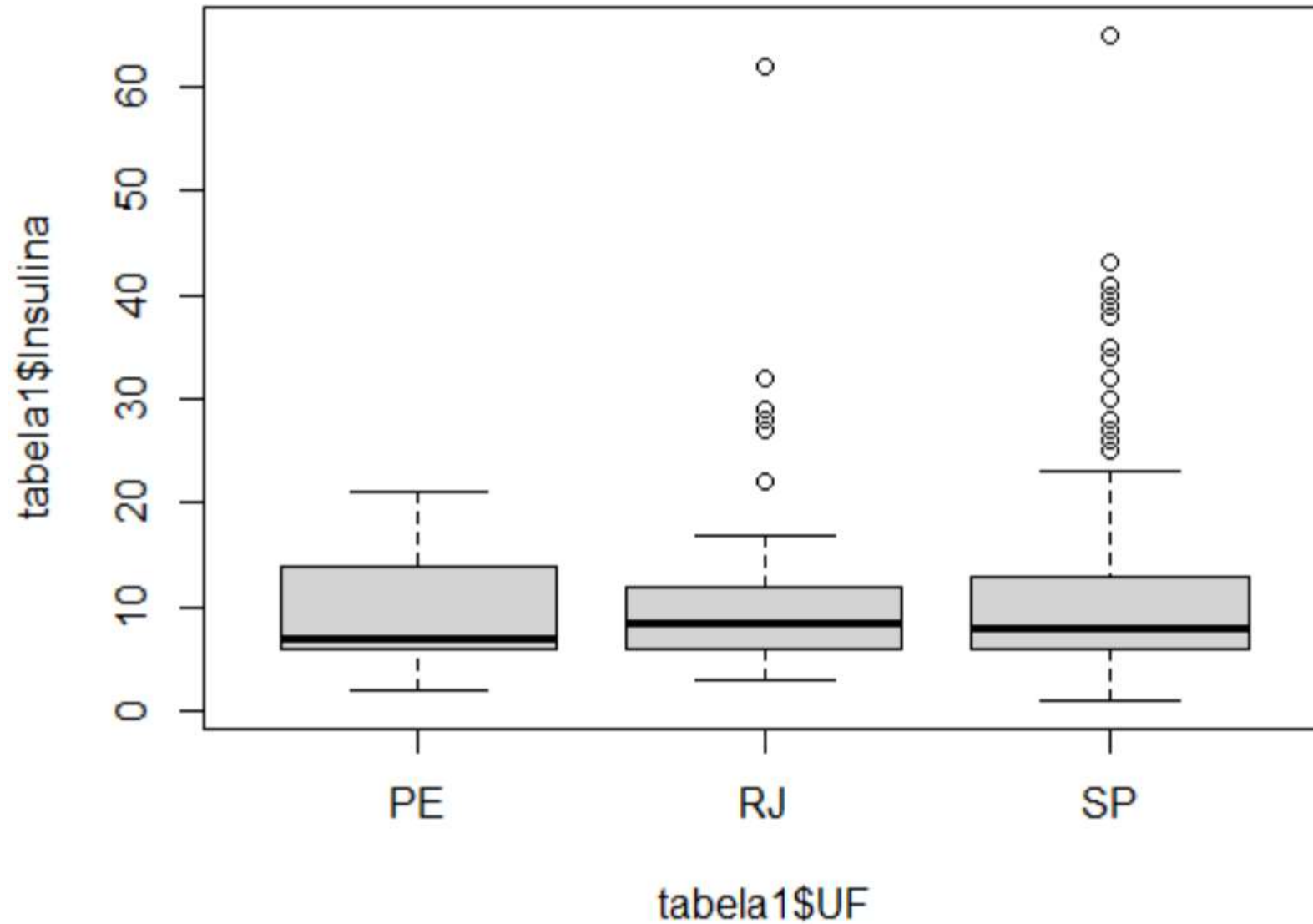
1  pontos

- ☐ Existe (sim) evidências de que há diferença.
- ☒ Não existe evidencias de que há diferença . 
- ☐ Não é possível responder essa pergunta com os dados apresentados.
- ☐ Não consegui responder essa perguntar porque não entendi como aplicar o teste.
- ☐ Não consegui responder essa perguntar porque não tenho computador disponível para executar o teste.
- ☐ Não sei/Não quero responder essa pergunta.



Adicionar feedback da resposta

Lista 9 – Pergunta 7



Lista 9 – Pergunta 7

```
> kruskal.test(tabela1$Insulina, tabela1$UF)
```

```
Kruskal-Wallis rank sum test
```

```
data: tabela1$Insulina and tabela1$UF
```

```
Kruskal-Wallis chi-squared = 0.4302, df = 2, p-value = 0.8065
```

Resultado: P maior que 0.05

Conclusão: H_0 é verdadeira

Hipóteses testadas

H_0 : Não há diferenças entre as medianas de todos os grupos comparados

H_1 : Existe diferença de pelo menos uma das medianas comparadas

Portanto:

Não existe evidências de que há diferença na insulina quanto ao estado de origem do paciente.

Chi-quadrado de aderência

Revisão: Chi-quadrado de aderência

- Esse teste avalia se as observações estão distribuídos entre diferentes classes de forma semelhante a de uma expectativa prévia.

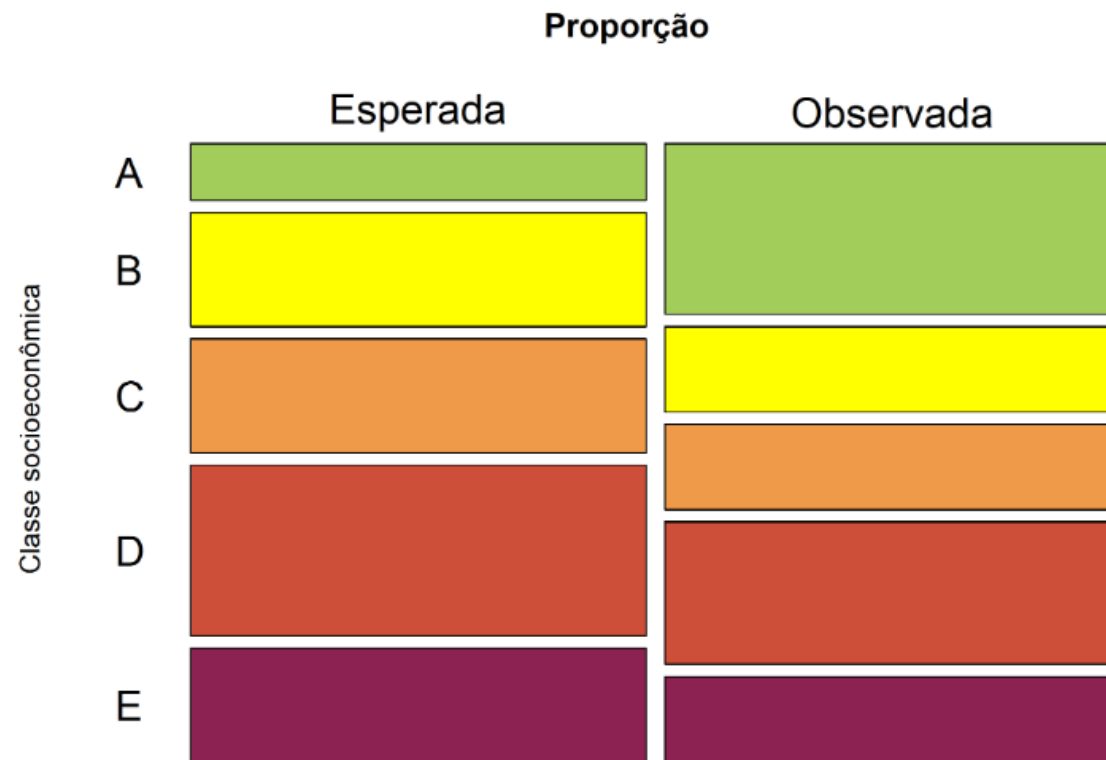




Figura. Mosaic plot representado as proporções esperadas vs. observadas

Lista 10

Pergunta 5. Existe evidências de que familiares de profissionais da saúde (Coluna Familiares) foram infectados por COVID numa razão diferente da distribuição de frequência de funcionários de um dado hospital (Coluna NumeroFuncionarios)? Assuma um nível de significância de 5%. Dados: "DadosHospital.xlsx"

1  pontos

- ☐ Existe (sim) evidências de que há diferença na distribuição de frequências de familiares infectados, da qual seria esperada se a distribuição fosse igual a distribuição de funcionários de um dado hospital.
- ☒ Não existe evidências de que há diferença na distribuição de frequências de familiares infectados, da qual seria esperada se a distribuição fosse igual a distribuição de funcionários de um dado hospital. 
- ☐ Não é possível responder essa pergunta com os dados apresentados.
- ☐ Não consegui responder essa pergunta porque não entendi como aplicar o teste.
- ☐ Não consegui responder essa pergunta porque não tenho computador disponível para executar o teste.
- ☐ Não sei/Não quero responder essa pergunta.

 [Adicionar feedback da resposta](#)

Concluído

Lista 10

Familiar da profissão

Proporção

Esperada

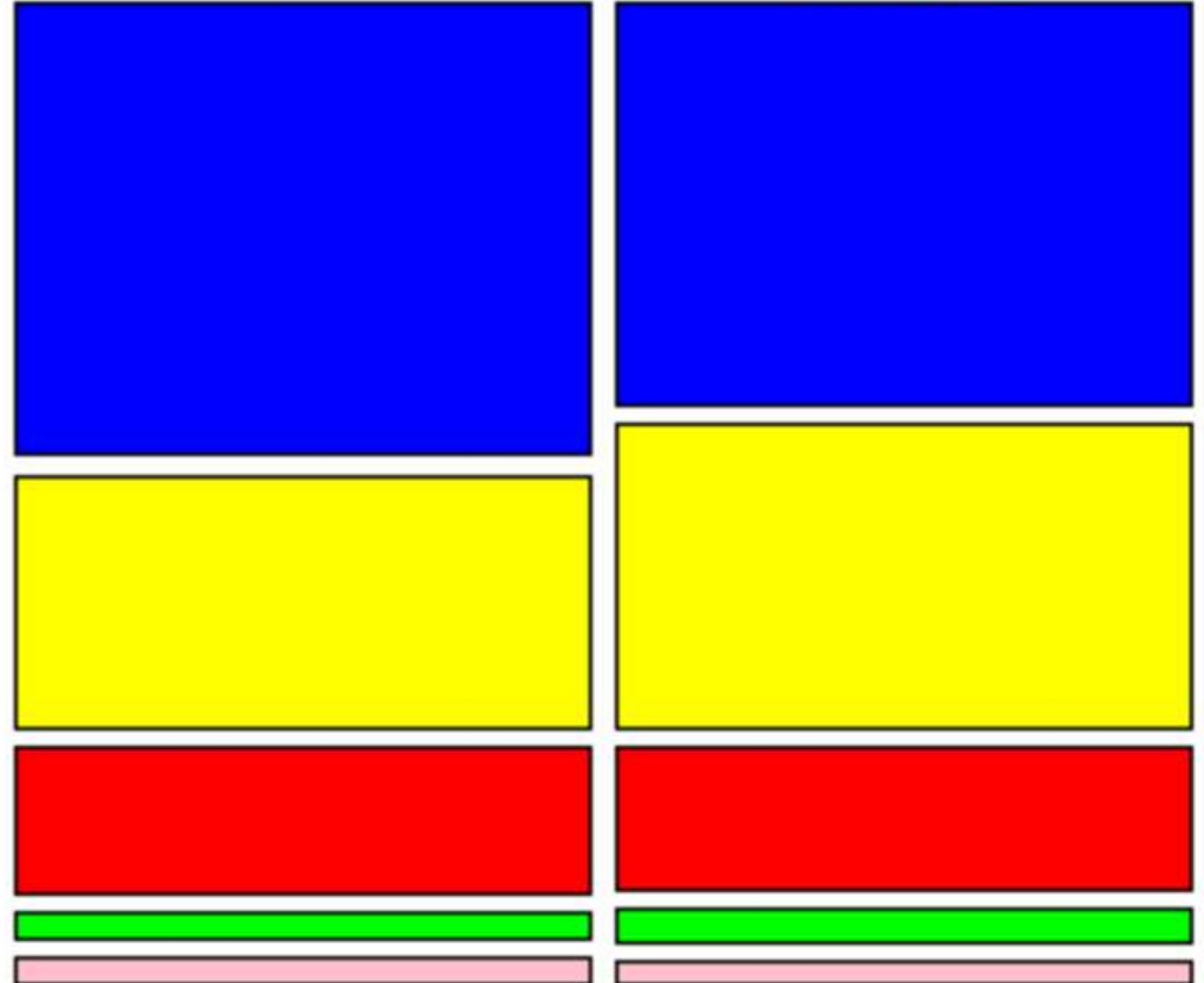
Observada

Auxiliar Enfermagem

Enfermeiro

Médico

Nutricionista
Psicólogo



Chi-squared test for given probabilities

```
data: tabela1$Familiares
```

```
X-squared = 3.9945, df = 4, p-value = 0.4067
```

Resultado: P maior que 0.05

Conclusão: H_0 é verdadeira

Hipóteses testadas

H_0 : A distribuição das frequências observas NÃO difere das frequências esperadas


H_1 : A distribuição das frequências observas difere (SIM) das frequências esperadas

Portanto:

Não existe evidências de que há diferença na distribuição de frequências de familiares infectados, da qual seria esperada se a distribuição fosse igual a distribuição de funcionários de um dado hospital.

Lista 10

Pergunta 6. Existe evidências de que há diferença a proporção de pessoas que pratica atividade física (Coluna PraticaAtividade) está distribuída em uma mesma frequência dos registros de estado civil (Coluna ProporcaoIBGE)? Assuma um nível de significância de 5%. Dados: "DadosEstadoCivil.xlsx"

1  pontos

- ☐ Existe (sim) evidências de que há diferença na frequência de pessoas que praticam atividade física, da qual seria esperada se a distribuição fosse igual a frequência de registros de pessoas em cada categoria, dada pelo IBGE.
- ☒ Não existe evidências de que há diferença na frequência de pessoas que praticam atividade física, da qual seria esperada se a distribuição fosse igual a frequência de registros de pessoas em cada categoria, dada pelo IBGE. ✓
- ☐ Não é possível responder essa pergunta com os dados apresentados.
- ☐ Não consegui responder essa perguntar porque não entendi como aplicar o teste.
- ☐ Não consegui responder essa perguntar porque não tenho computador disponível para executar o teste.
- ☐ Não sei/Não quero responder essa pergunta.

 [Adicionar feedback da resposta](#)

Concluído

Lista 10

		Proporção	
		Esperada	Observada
PraticaAtividade	Casado		
	Desquitado		
	Divorciado		
	Viuvo		
	Solteiro		

Chi-squared test for given probabilities

```
data: tabela2$PraticaAtividade
```

```
X-squared = 0.10918, df = 4, p-value = 0.9986
```

Resultado: P maior que 0.05

Conclusão: H_0 é verdadeira

Hipóteses testadas

H_0 : Não há associação entre as variáveis

H_1 : Há (SIM) associação entre as variáveis

Portanto:

Não existe evidências de que há diferença na frequência de pessoas que praticam atividade física, da qual seria esperada se a distribuição fosse igual a frequência de registros de pessoas em cada categoria, dada pelo IBGE.

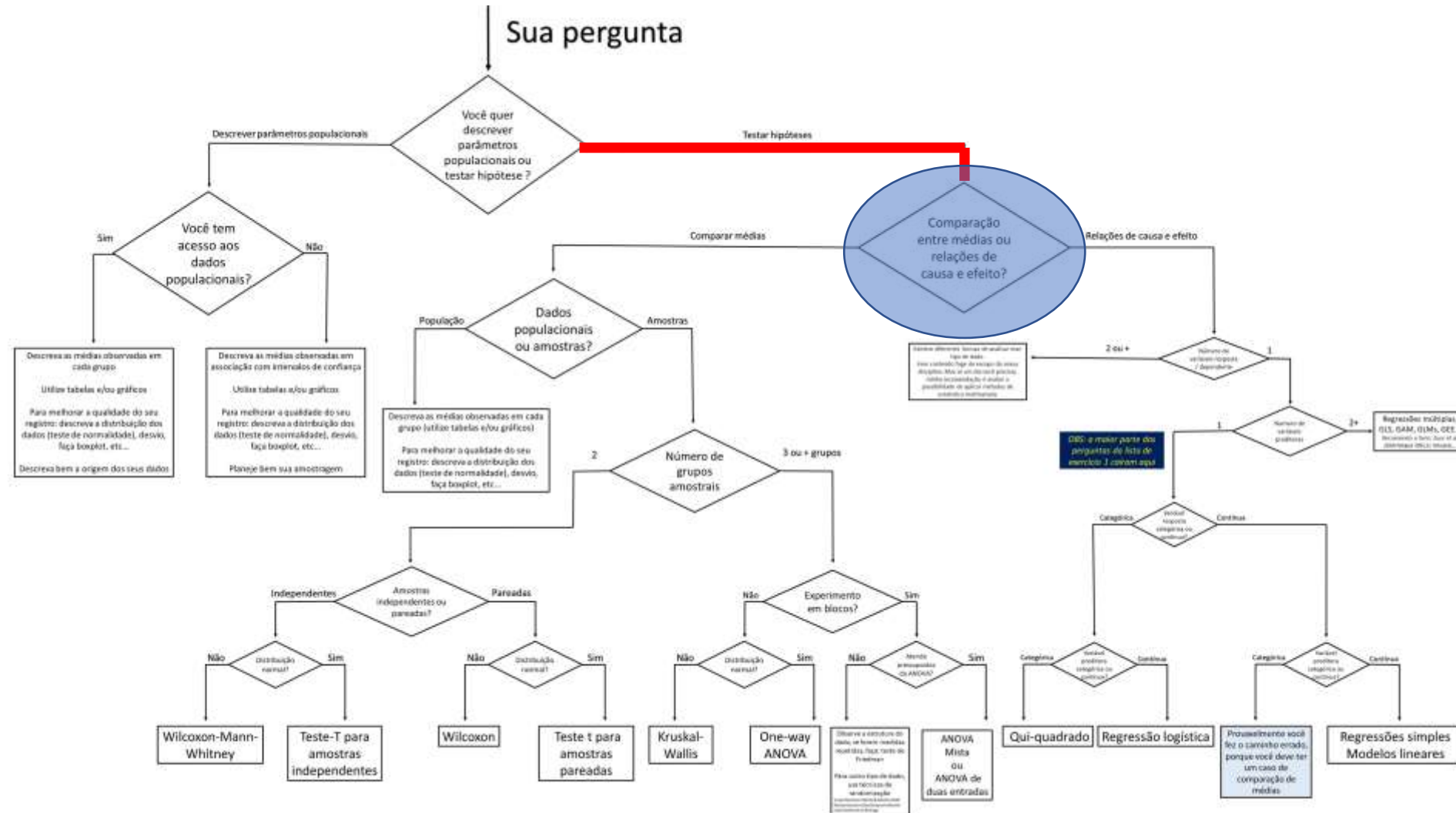
Aula de hoje

Teste de qui-quadrado de independência

Mapa para seleção do método estatístico



Teste de hipóteses



Seleção do método estatístico



Seleção do método estatístico



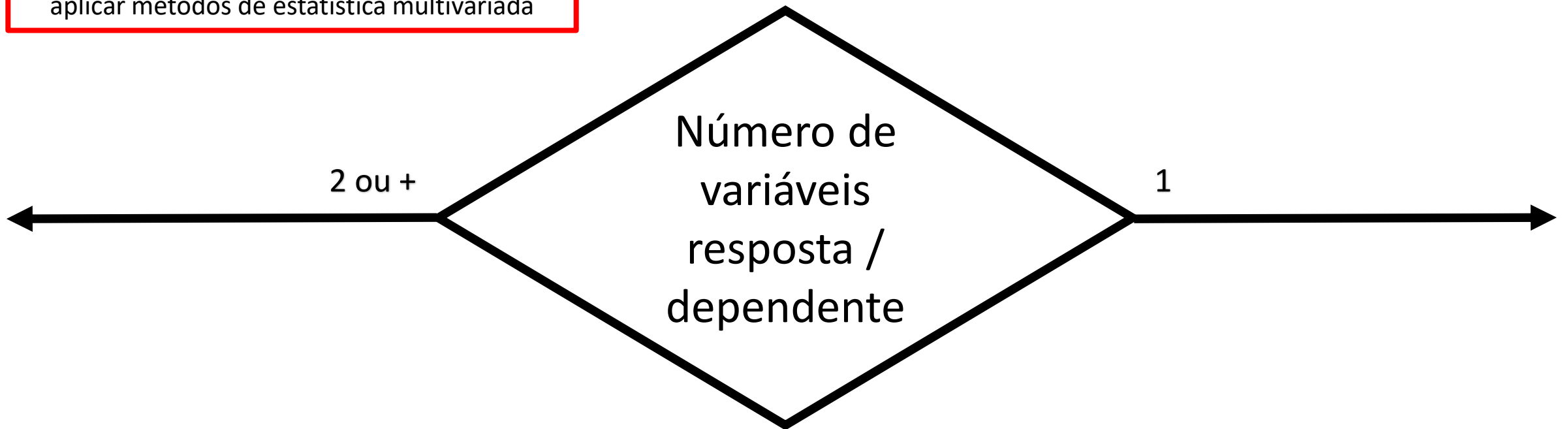
Seleção do método estatístico



Seleção do método estatístico

Existem diferentes formas de analisar esse tipo de dado.

Esse conteúdo foge do escopo da nossa disciplina. Mas se um dia você precisar, minha recomendação é avaliar a possibilidade de aplicar métodos de estatística multivariada



Ideia geral de uma análise multivariada

Variável resposta

Expectativa de
vida, nível de
estresse, nota no
enem, etc...

Variável preditora

Grau de
escolaridade, UF,
Renda familiar



Então, nessa disciplina só vamos trabalhar com 1 variável resposta

Variável resposta

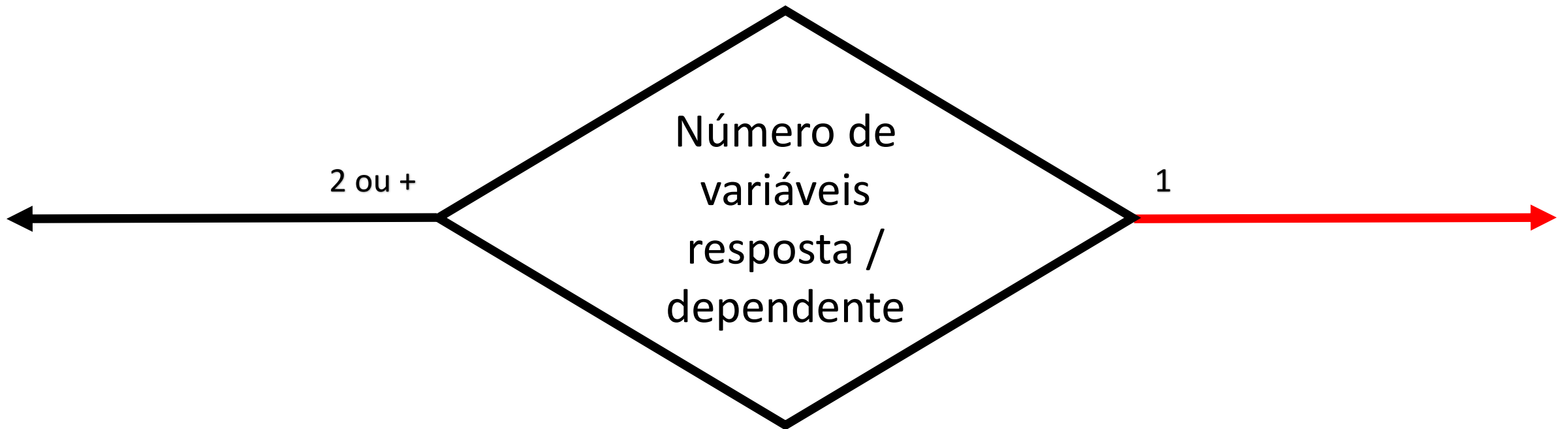
Nessa disciplina
aqui sempre será 1
única variável

Variável preditora

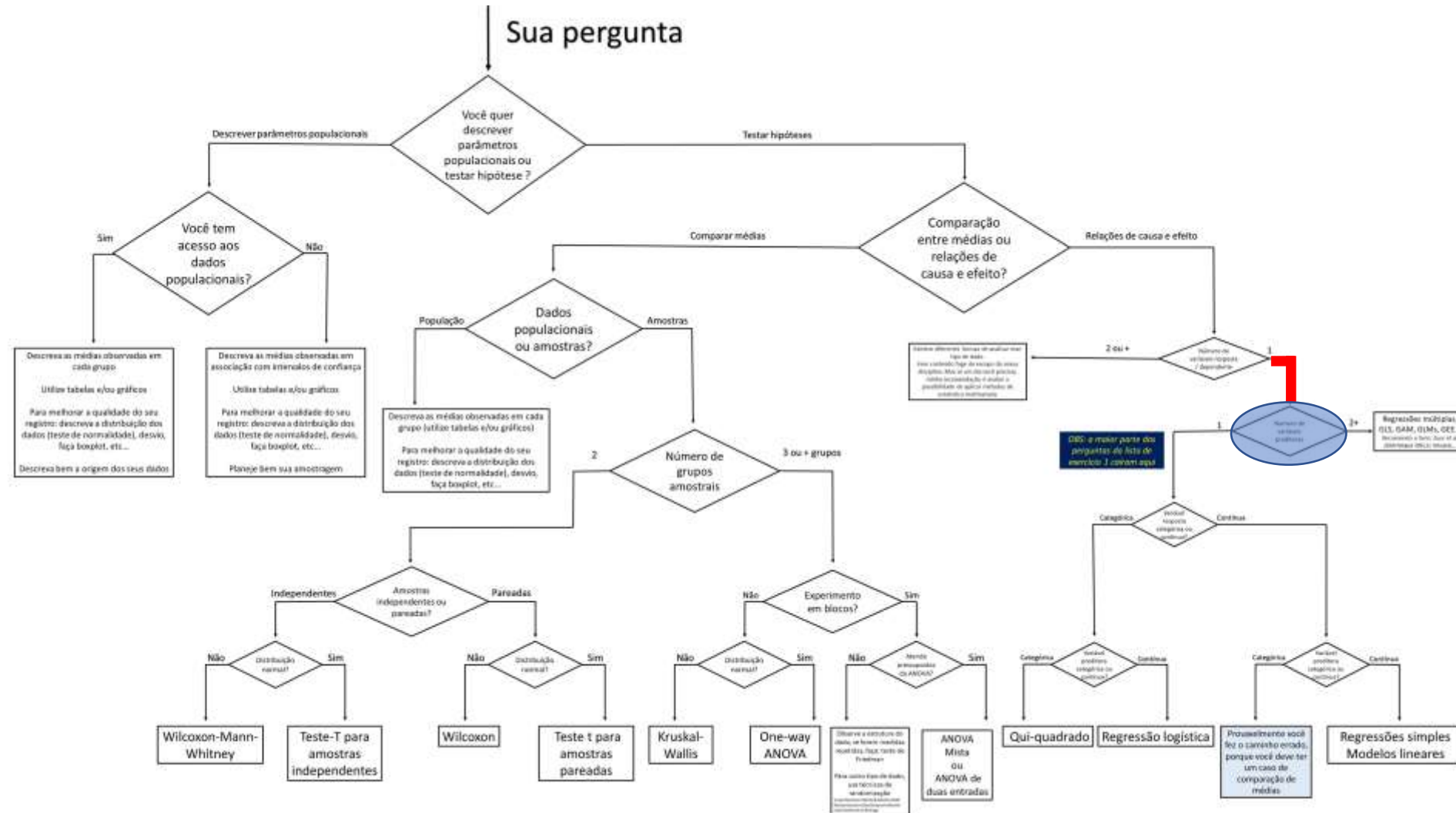
Desse lado pode ter
1 ou mais variáveis



Seleção do método estatístico



Seleção do método estatístico



Seleção do método estatístico

Regressões múltiplas, GLS,
GAM, GLM, GLMM, GEE...
Recomendo o livro: Zuur et al
2009 Mixed Effects Models...

2 ou +

Numero de
variáveis
preditoras

1

Assunto da última aula

Aqui temos

Variável resposta

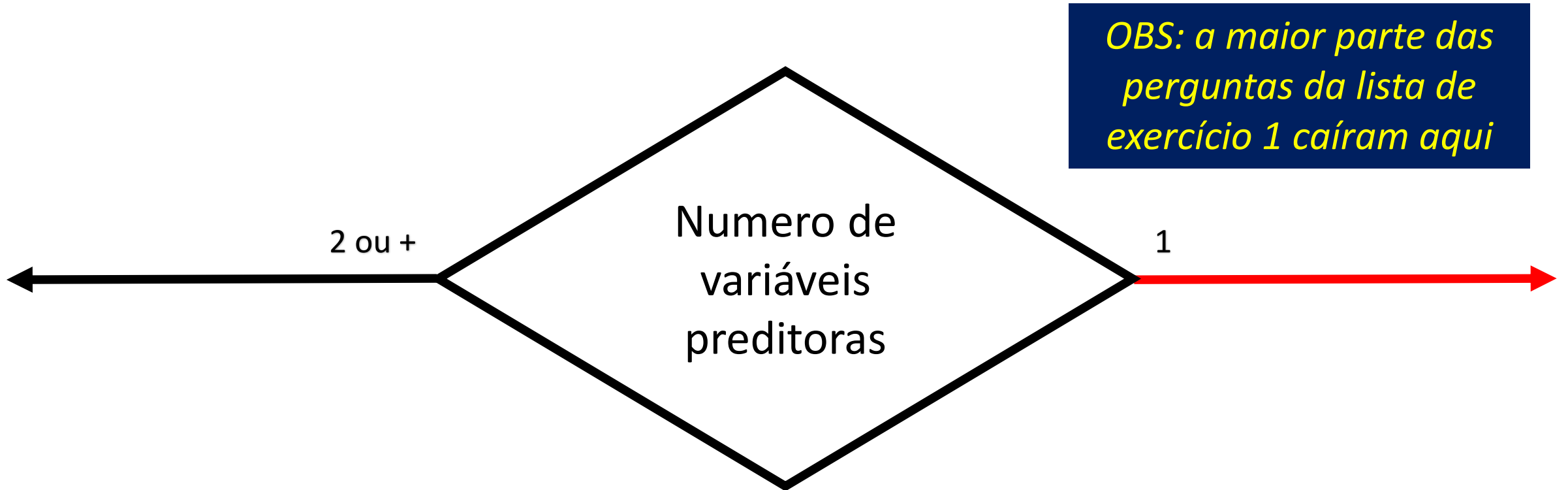
Variável preditora

Nessa disciplina aqui
sempre será 1 única
variável

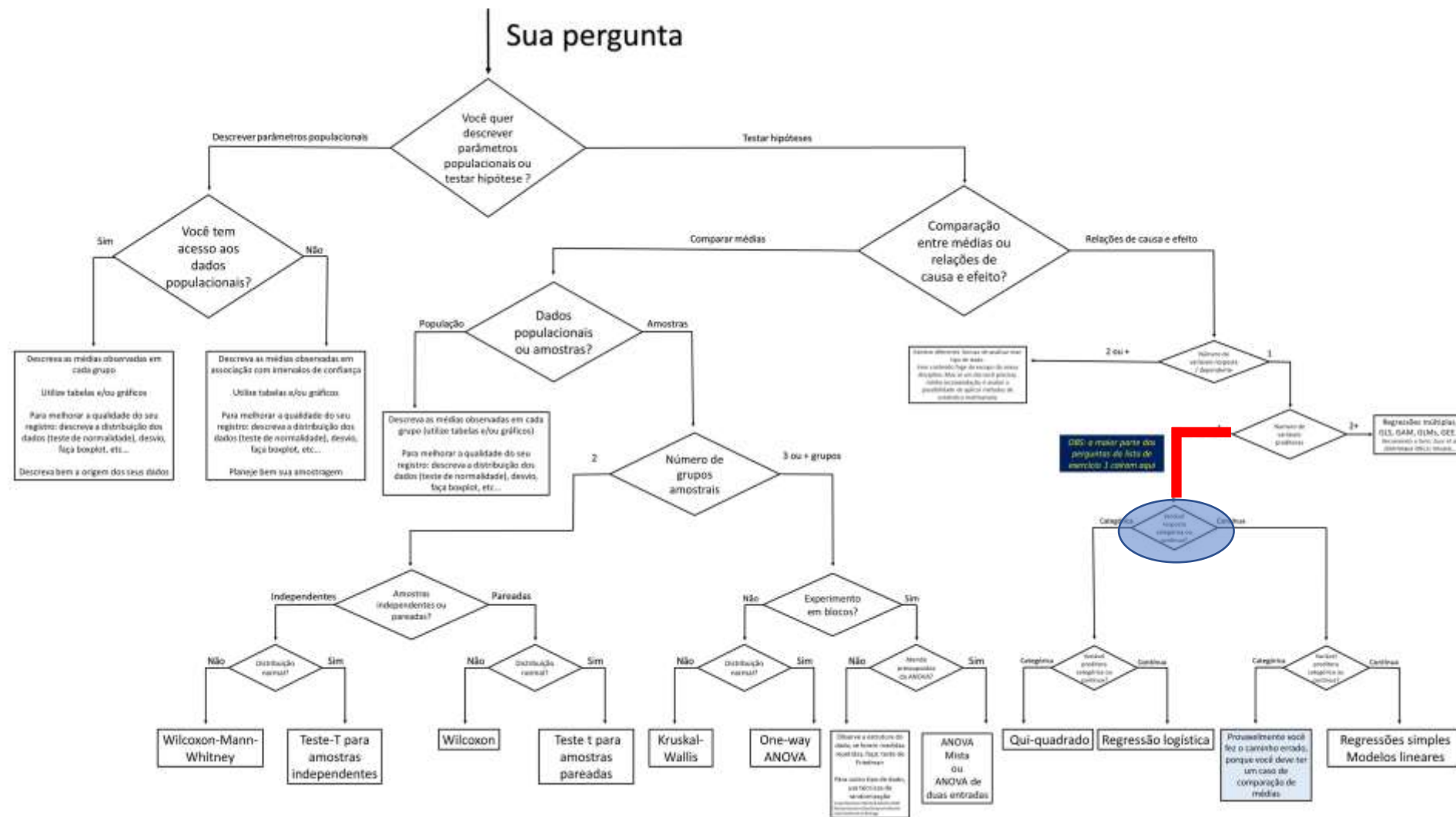
2 ou + variáveis que
pode ser categóricas
ou continuas



Seleção do método estatístico



Seleção do método estatístico



Aqui temos

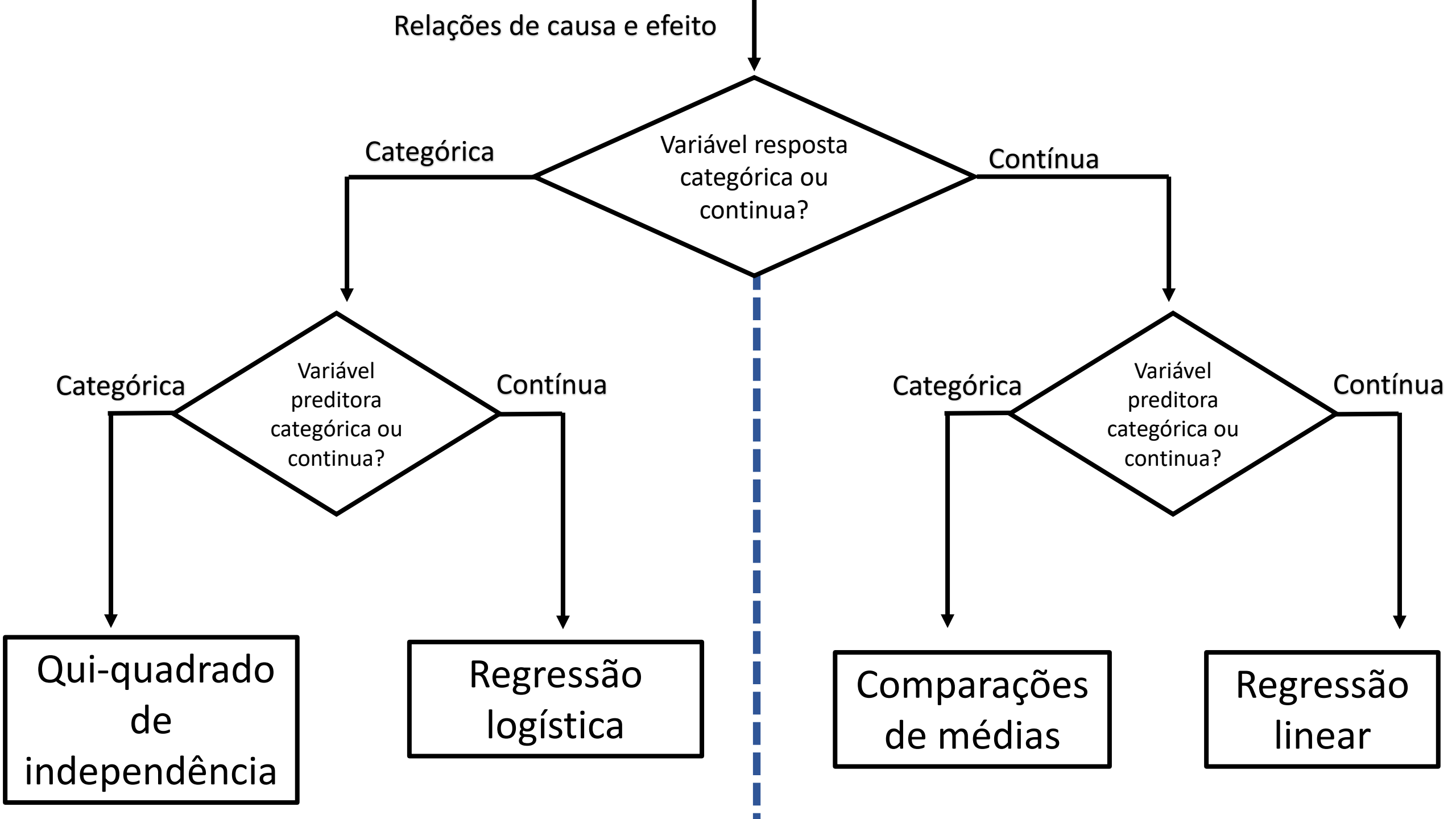
Variável resposta

Variável preditora

1 variável que pode
ser categórica ou
continua

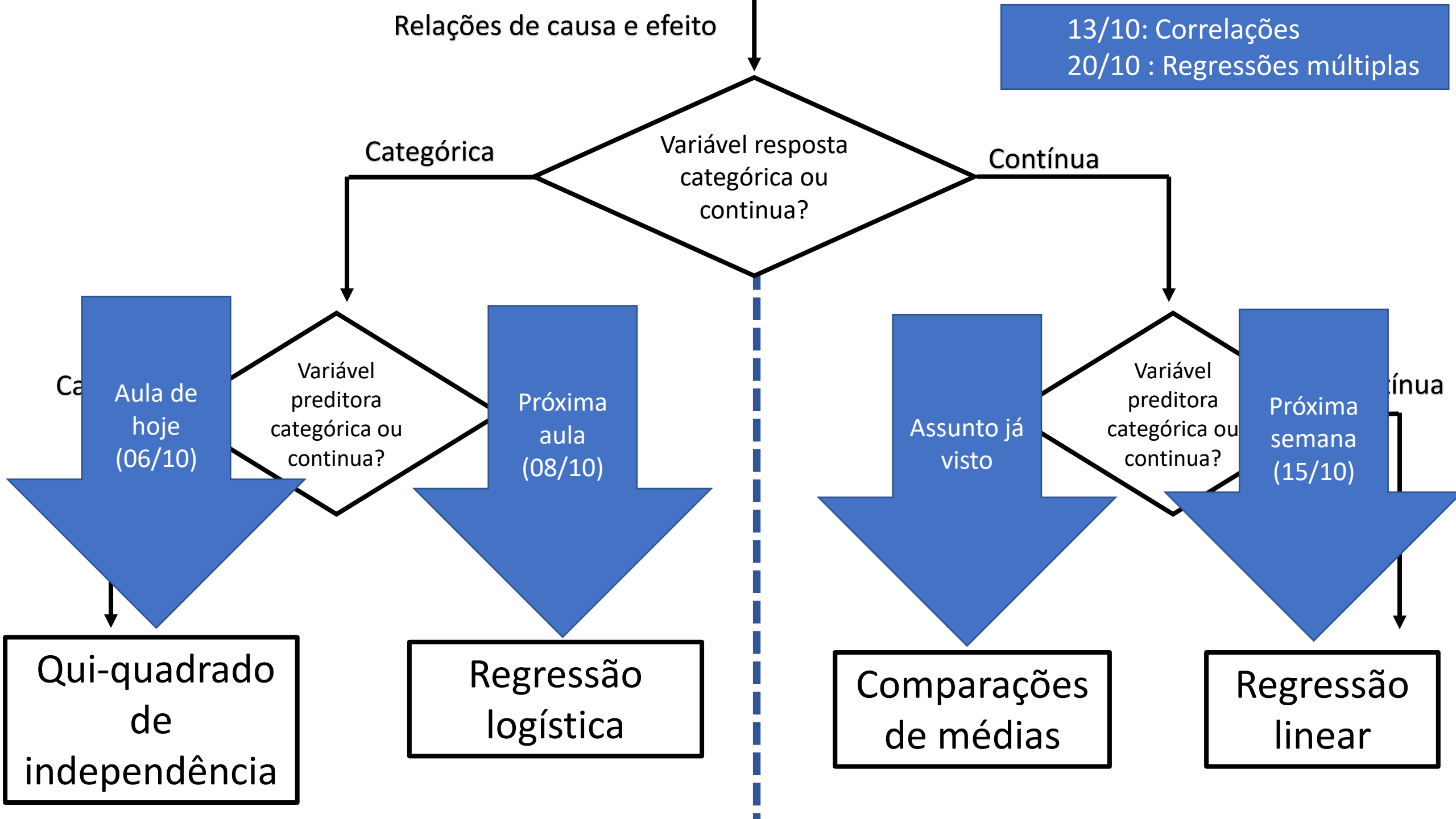
1 variável que pode
ser categórica ou
continua





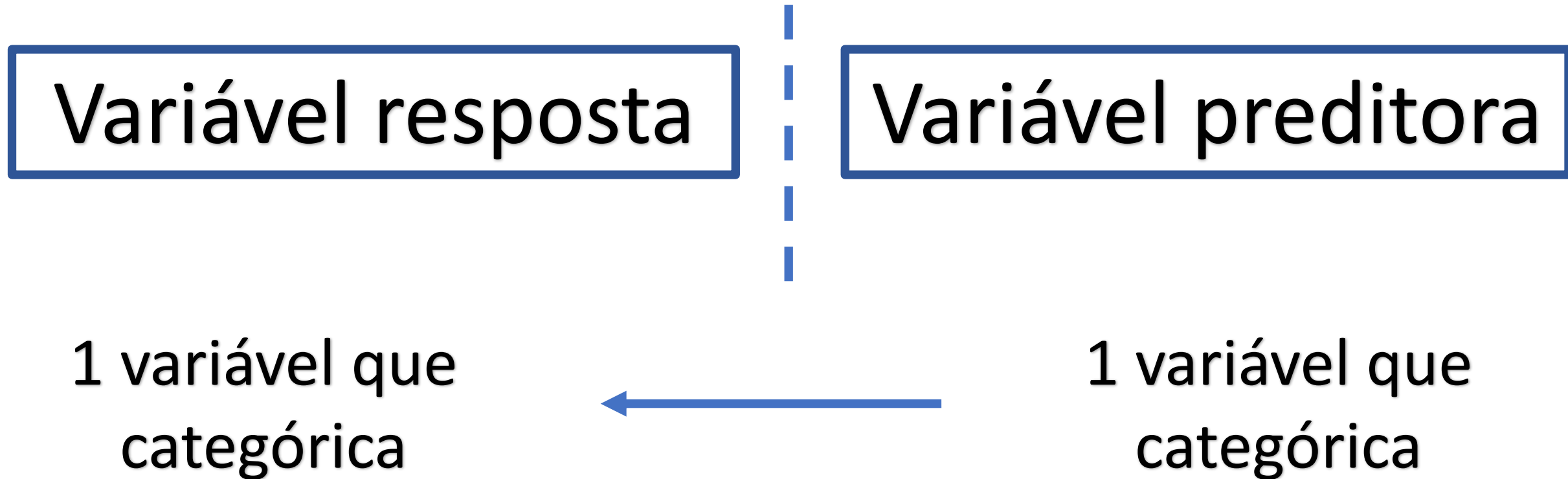
Relações de causa e efeito

13/10: Correlações
20/10 : Regressões múltiplas



Qui-quadrado de independência

Qui-quadrado de independência



- OBS: Esse foi o método estatístico com maior numero de pessoas que enviaram propostas que poderiam ser respondidas com esse métodos

Tabelas de contingência

Distribuição de registros de obesidade dado o sexo dos pacientes

Sexo	Obesidade		Total	Frequência obesidade
	Sim	Não		
Masculino	30	70	100	30%
Feminino	20	80	100	20%
Total	50	150	200	25%

Pergunta: A probabilidade de registro de obesidade depende do sexo do paciente?

Sexo	Obesidade		Total	Frequência obesidade
	Sim	Não		
Masculino	30	70	100	30%
Feminino	20	80	100	20%
Total	50	150	200	25%

Pergunta: A probabilidade de registro de obesidade depende do sexo do paciente?

Sexo	Obesidade		Total	Frequência obesidade
	Sim	Não		
Masculino	30	70	100	30%
Feminino	20	80	100	20%
Total	50	150	200	25%

Em outras palavras, existe diferença nas proporções observadas nas colunas, dependendo das linhas?

30/100 é diferente de 20/100 ?

Sexo	Obesidade		Total	Frequência obesidade
	Sim	Não		
Masculino	30	70	100	30%
Feminino	20	80	100	20%
Total	50	150	200	25%

ÚLTIMA AULA

- Na última aula comparamos se os registros de obesidade estavam distribuídos entre diferentes classes sociais, da mesma forma como as pessoas de toda população estão distribuídas entre as classes sociais

A área de cada caixa representa a proporção dos dados para cada grupo

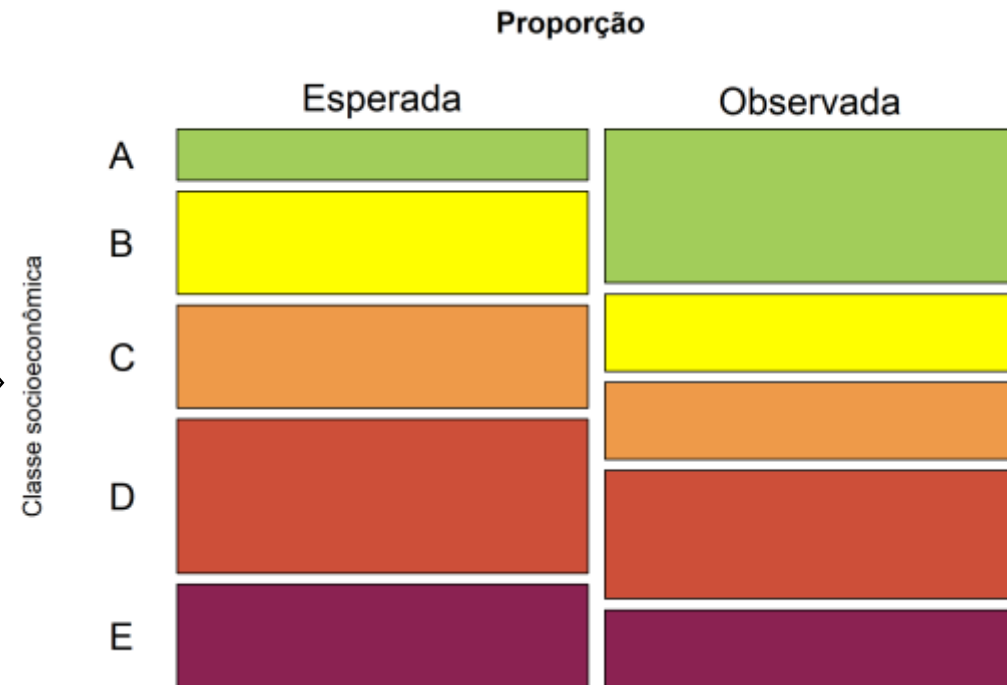


Figura. Mosaic plot representado as proporções esperadas vs. observadas

AULA DE HOJE

- Na aula de hoje, o procedimento por trás do teste será usar a combinação das observações, para estabelecer uma frequência esperada, em seguida o teste avalia se a distribuição de cada grupo difere da frequência observada

A área de cada caixa representa a proporção dos dados para cada grupo

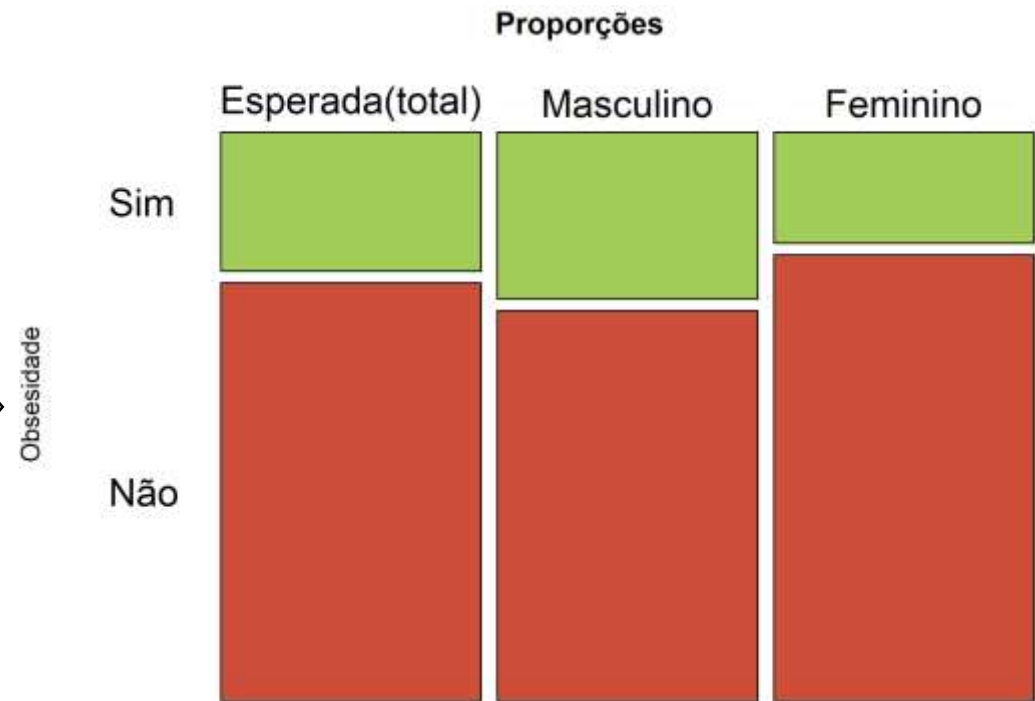
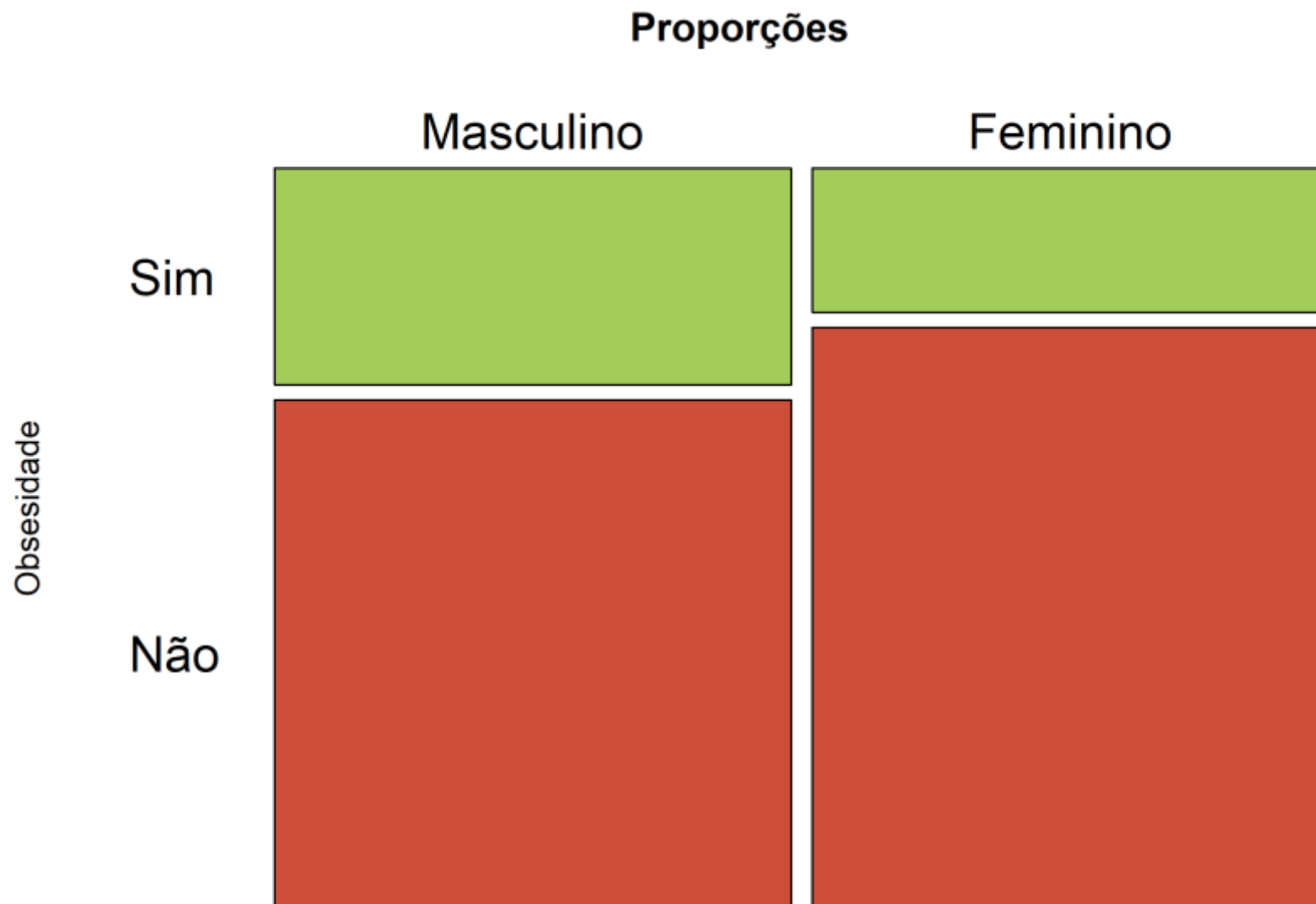


Figura. Mosaic plot representado as proporções esperadas vs. observadas

Na prática nós só vamos observar essas proporções



Hipóteses

- H_0 : Não há associação entre as variáveis
- H_1 : Há (SIM) associação entre as variáveis

No exemplo de hoje, as hipóteses são:

- H_0 : A probabilidade de relatar obesidade NÃO depende do sexo do paciente
- H_1 : A probabilidade de relatar obesidade depende (SIM) do sexo do paciente

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

- 1) Usar a máscara para sair na rua, reduz o risco de contrair COVID-19? (Adriana Barros, Jozielda, Thayla)
- 2) O uso de Ivermectina reduz o risco de morte por covid-19? (Claudia, Rafaela Reimberg)
- 3) O tratamento fornecido pelo SUS para tuberculose reduz o risco de morte do paciente? (Ana Lucia)
- 4) Escovar o dente antes de dormir, reduzem o risco de contrair cárie? (Claudineia)
- 5) O uso de remédios psiquiátricos altera a chance de pacientes desempregados encontrarem um emprego? (Marina Menotti)
- 6) Aferir a temperatura para detectar o COVID-19 é realmente eficaz? (Leticia)

Exemplos de adaptações de perguntas: identificar variáveis

Saia de casa de máscara?	Pegou COVID?		Total	Frequência de COVID
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Exemplos de adaptações de perguntas: identificar variáveis

Usou Ivermectina?	Morte?		Total	Frequência de morte
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Exemplos de adaptações de perguntas: identificar variáveis

Recebeu tratamento do SUS ?	Morte?		Total	Frequência de morte
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Exemplos de adaptações de perguntas: identificar variáveis

Escova o dente antes de dormir?	Pegou cárie?		Total	Frequência de cárie
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Exemplos de adaptações de perguntas: identificar variáveis

Uso de remédios psiquiátricos?	<u>Arrumou um emprego?</u>		Total	Frequência de emprego
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Exemplos de adaptações de perguntas: identificar variáveis

Detectou COVID?	<u>A pessoa estava infectada?</u>		Total	Frequência de emprego
	Sim	Não		
Sim	30	70	100	30%
Não	20	80	100	20%
Total	50	150	200	25%

Desenho experimental

Ideia geral: Você precisa amostrar dados de forma que todas as células dessa tabela estejam preenchidas com dados

Grupos	Fenômeno estudado		Total	Frequência do fenômeno
	Sim	Não		
Grupo 1	?	?	-	-%
Grupo 2	?	?	-	-%
Total	-	-	-	-%

Desenho experimental

- a. Acompanhar pessoas que usam, e pessoas que não usam, máscara para sair na rua, e avaliar a proporção de cada uma delas que pegou e que não pegou COVID depois de 1 mês.
- b. Acompanhar pacientes em tratamento do COVID, que receberam, e não que não receberam, tratamento com Ivermectina, e avaliar a proporção óbitos em cada grupo.
- c. Acompanhar pacientes com Tuberculose que recebem, e que não receberam, tratamento fornecido pelo SUS, e avaliar a proporção óbitos em cada grupo.
- d. Acompanhar pessoas que escovam, e que não escovam, o dente antes de dormir, e avaliar a proporção de pessoas com cárie em cada um dos grupos.
- e. Acompanhar pessoas que usam de remédios psiquiátricos, e que não usam, e registrar a proporção destes que conseguiram arrumar um emprego depois de um intervalo de tempo.
- f. Requerir que baseado em medidas de temperatura por termômetro digital, um profissional avaliem se um grupo de pessoas está com covid ou não. Em seguida contrastar os resultados observados com os registros de COVID por outro método, tal como PCR.

Como fica uma tabela desse tipo de dado?

Nome	Usava máscara?	Pegou COVID?
Lorena	Não	Não
Livia	Não	Sim
Maria Luiza	Não	Não
Cecilia	Não	Sim
Eloa	Não	Sim
Giovanna	Não	Sim
Maria Clara	Sim	Não
Maria Eduarda	Sim	Sim
Mariana	Sim	Não
Lara	Sim	Não
Beatriz	Sim	Não
Antonella	Sim	Sim

Identificação do dado

Agrupamento 1

Agrupamento 2

Nome	Usava máscara	Pegou COVID?
Lorena	Não	Não
Livia	Não	Sim
Maria Luiza	Não	Não
Cecilia	Não	Sim
Eloa	Não	Sim
Giovanna	Não	Sim
Maria Clara	Sim	Não
Maria Eduarda	Sim	Sim
Mariana	Sim	Não
Lara	Sim	Não
Beatriz	Sim	Não
Antonella	Sim	Sim

Pré-requisitos

- Existem divergências (Levine et al 2012)
 - Nenhuma célula de frequência esperada apresentar valor inferior à 0,5 (Lewontin & Felsenstein 1965 Biometrics)
 - Não mais que 20% das célula de frequência esperada apresentar valor inferior à 5 (Dixon & Massey-Jr 1983 Introduction to statistical analysis)
 - Solução conciliatória: nenhuma célula de frequência esperada apresentar valor inferior à 1 (Levine et al 2012)
- Minha sugestão, se no seu estudo houver uma célula de observações esperadas for menor do que 5, avalie a viabilidade de:
 - Aumentar seu tamanho de amostra
 - Consolidar duas ou mais categorias em apenas uma
 - Utilizar métodos alternativos (ex. teste exato de Fisher ou Barnard)

Dimensões das tabelas

Tabelas de contingência 2x2

Sexo	Gostou do atendimento do hospitalar?		Total	Frequência de avaliações positivas
	Sim	Não		
Masculino	30	70	100	30%
Feminino	20	80	100	20%
Total	50	150	200	25%

Tabelas de contingência de outras dimensões

Sexo	Como avalia a qualidade do atendimento do hospital?				Total
	Excelente	Bom	Regular	Ruim	
Masculino	10	15	30	20	75
Feminino	15	10	20	30	75
Total	25	25	50	50	150

Prática

Prática – 1 O diagnostico de COVID está associado com o sexo?

- Tabela “Dados Fleury.xlsx”
 - OBS: dados reais
- Covid19PCR: Diganostico de COVID por PCR.
 - Respostas possíveis: positivo ou negativo
- Sexo: Sexo biológico do paciente.
 - Respostas possíveis: Masculino ou feminino

Dados

Diagnóstico	Sexo		Total	Frequência de mulheres
	Feminino	Masculino		
Negativo	291	190	481	60,5%
Positivo	61	39	100	61,0%
Total	352	229	581	60,58%

Negativo

Positivo

F



M



Interpretação dos resultados

Pearson's Chi-squared test with Yates' continuity correction

data: TabCont
X-squared = 2.772e-30, df = 1, p-value = 1

Resultado: P maior que 0.05
Conclusão: H_0 é verdadeira

Hipóteses testadas

H_0 : Não há associação entre as variáveis
 H_1 : Há (SIM) associação entre as variáveis

Prática – 2 O risco de pegar COVID é influenciado pelo uso de mascara?

- Tabela “TabelaContingencia1.xlsx”
 - OBS: dados hipotéticos
- COVID: Diagnóstico de COVID por PCR.
 - Respostas possíveis: positivo ou negativo
- Mascara: a pessoa usava mascara para sair de casa?
 - Respostas possíveis: Usava mascara ou Não usava mascara

Dados

Diagnóstico COVID	Usava máscara?		Total	Frequência de COVID
	Sim	Não		
Positivo	40	60	100	40%
Negativo	60	40	100	60%
Total	100	100	200	50%

Negativo

Positivo

Nao usava mascara

Usava mascara



Interpretação dos resultados

Pearson's Chi-squared test

data: TabCont
X-squared = 13.934, df = 2, p-value = 0.0009423

Resultado: P menor que 0.05

Conclusão: H_1 é verdadeira

Hipóteses testadas

H_0 : Não há associação entre as variáveis

H_1 : Há (SIM) associação entre as variáveis

Prática – 3 A percepção de qualidade do atendimento depende do sexo do paciente?

- Tabela “TabelaContingencia2.xlsx”
 - OBS: dados hipotéticos
- Sexo: Sexo biológico do paciente.
 - Respostas possíveis: Masculino ou feminino
- QualidadeAtendimento: avaliação do paciente sobre a qualidade do atendimento
 - Respostas possíveis: Excelente, Bom, Regular ou Ruim

Dados

Como avalia a qualidade do atendimento do hospital?					
Sexo	Excelente	Bom	Regular	Ruim	Total
Feminino	40	30	48	32	150
Masculino	15	13	9	13	50
Total	55	43	57	45	200

F

M

Bom

Excelente

Regular

Ruim



Interpretação dos resultados

Pearson's Chi-squared test

```
data: TabCont  
X-squared = 3.7213, df = 3, p-value = 0.2932
```

Resultado: P maior que 0.05
Conclusão: H_0 é verdadeira

Hipóteses testadas

H_0 : Não há associação entre as variáveis

H_1 : Há (SIM) associação entre as variáveis