Curso de Pós-Graduação em Ciências Veterinárias - UFRRJ

Métodos Estatísticos

Prof: Wagner Tassinari

wagner.tassinari@ini.fiocruz.br

Probabilidade

Probabilidade

Probabilidade

- Teoria matemática utilizada para se estudar incerteza oriunda de fenômenos de caráter aleatório.
- Base sobre a qual são construídos importantes métodos de inferência estatística.

Alguns conceitos

- **Experimento Determinístico:** conhecidas as condições sob as quais o experimento vai ser realizado podemos indicar, a "*priori*", o resultado que será obtido.
  - **Exemplo:** Na operação $2 + 2$, sabe-se que o resultado é o número 4.
- **Experimento Aleatório:** sabemos os resultados possíveis, mas não podemos dizer qual deles irá ocorrer.
  - **Exemplo:** No lançamento de uma moeda, o resultado poderá ser *cara* ou *coroa*.

Espaço Amostral: é o conjunto de todos os resultados possíveis de um certo experimento aleatório.

Evento: é qualquer subconjunto do espaço amostral.

Características de um experimento aleatório

- Cada experimento poderá ser repetido indefinidamente sob condições inalteradas;
- Podemos descrever os possíveis resultados;
- Há uma regularidade quando o experimento for repetido um grande número de vezes.

Exemplo de um experimento aleatório

1. Experimento Aleatório: Jogar um dado homogêneo e observar o resultado
   - Espaço Amostral: $\{1, 2, 3, 4, 5, 6\}$
   - Evento: {sair o número 6}, {sair um número par}
2. Experimento Aleatório: Uma pessoa faz um exame para saber se tem uma determinada doença
   - Espaço Amostral: {tem a doença, não tem a doença}
   - Evento: {tem a doença}

Como atribuir probabilidades aos eventos ?

1. Através de características teóricas da realização do fenômeno.

- **Exemplo:** Em um dado não viciado temos igual probabilidade para cada uma das faces.

P(sair o número 6) $= 1/6 \rightarrow$ **eventos equiprováveis**

2. Através da frequência relativa.

- **Exemplo:** Suponha a população de uma cidade no ano de 2001 $=$ 100.000 habitantes. No mesmo ano, os casos registrados de uma determinada doença nessa cidade $=$ 20 pacientes.
- Prevalência $=$ P(ter a doença) $=$ freq.relativa $= 20/100000 = 0,0002$

Definição de probabilidade clássica (frequência relativa)

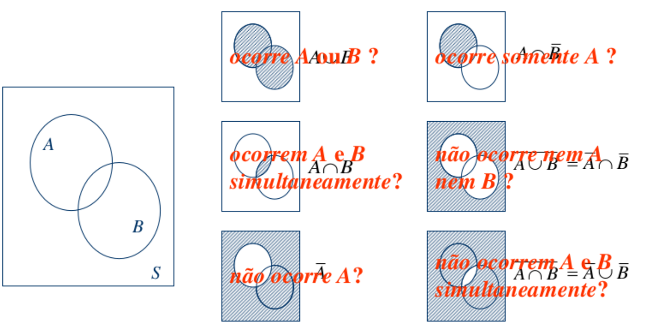
$$P(A) = \frac{k}{n}$$

Sendo:

k: Número de casos favoráveis ao evento A no experimento

n: Número total de casos no experimento

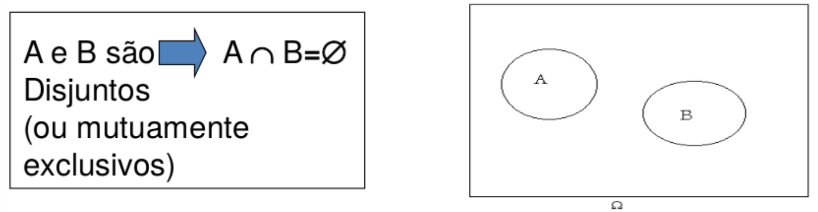
Relembrando conjuntos



Diagramas de Venn

Noções Fundamentais de Probabilidade

- Vamos representar o espaço amostral por Ω e dois eventos quaisquer deste espaço amostral por A e B.
- Definição: Uma função de probabilidade $P(.)$ é denominada probabilidade se satisfaz as condições:
  i. $0 \leq P(A) \leq 1$
  ii. $P(\Omega) = 1$
  iii. $P(A \cap B) = P(A) + P(B)$, se A e B são eventos disjuntos



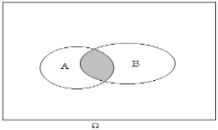
Noções Fundamentais de Probabilidade

Pode-se mostrar que valem também as seguintes relações:
Se A e B são dois eventos quaisquer, então:

$$\text{iv)}\ P(\underset{\substack{\uparrow \\ \text{ou}}}{A \cup B}) = P(A) + P(B) - P(A \cap B)$$

(regra da adição)

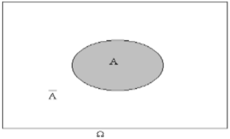
A probabilidade da interseção é contada duas vezes!!



Para o evento complementar vale a relação:

$$\text{v)}\ P(\overline{A}) = 1 - P(A)$$

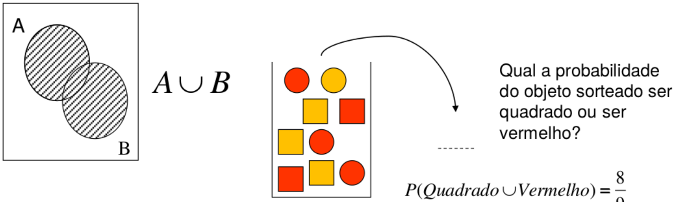
$\overline{A}$ é o complementar de A ⇒ $A \cap \overline{A} = \varnothing$ **e** $A \cup \overline{A} = \Omega$



vi. $P(\emptyset) = 0$

Para o evento vazio $\emptyset$, a probabilidade é zero.

- **Exemplo 1:** Supondo a urna abaixo, qual a probabilidade do objeto sorteado ser quadrado ou vermelho ?

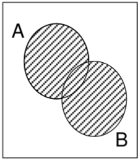


$$P(Quadrado \cup Vermelho) = \frac{8}{9}$$

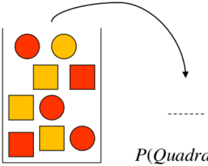
$$P(Quadrado \cup Vermelho) = P(Quadrado) + P(Vermelho)$$

$$\cancel{= \frac{5}{9} + \frac{5}{9} = \frac{10}{9} > 1?}$$

- **Exemplo 1:** Supondo a urna abaixo, qual a probabilidade do objeto sorteado ser quadrado ou vermelho ?



$A \cup B$



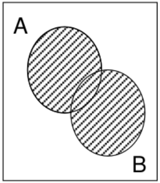
Qual a probabilidade do objeto selecionado ser quadrado ou ser vermelho?

$$P(Quadrado \cup Vermelho) = \frac{8}{9}$$

$$P(Quadrado \cup Vermelho) = P(Quadrado) + P(Vermelho) - P(Quadrado \cap Vermelho)$$

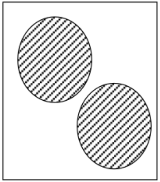
$$= \frac{5}{9} + \frac{5}{9} - \frac{2}{9} = \frac{8}{9}$$

- Resumindo:



$$A \cup B$$

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$



$$P(A \cap B) = 0 \quad \therefore \quad P(A \cup B) = P(A) + P(B)$$

(eventos mutuamente exclusivos)

- **Exemplo 2:** Calcule a probabilidade de sair ímpar no lançamento de um dado
- A: Sair ímpar
- $A = 1, 3, 5$

$$P(A) = \frac{k}{n} = \frac{3}{6} = \frac{1}{2}$$

- **Exemplo 3:** Retira-se uma carta de um baralho. Determine a probabilidade de ocorrer:

a. Dama $\rightarrow P(Dama) = \dfrac{4}{52}$
b. Carta de copas ou espada $\rightarrow P(Dama \cup Espada) = \dfrac{26}{52}$
c. Rei ou carta de espada
   i. A: Sair rei
   ii. B: Sair uma carta de espada

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

Probabilidade Condicional

- Muitas vezes o fato de um evento ocorrer influencia na probabilidade da ocorrência de um outro evento. Neste caso, definimos:
- Dados dois eventos A e B , a probabilidade condicional de B dado que A ocorreu é representada por $P(B|A)$:

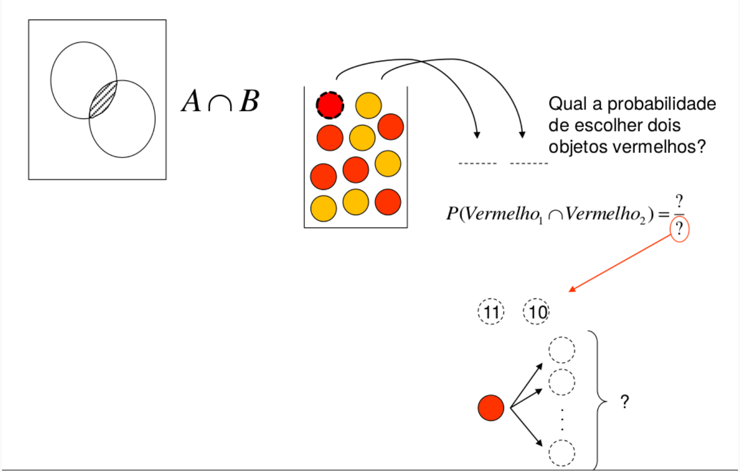
$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

- Que pode ser escrito como (pela regra do produto):

$$P(A \cap B) = P(A).P(B|A)$$

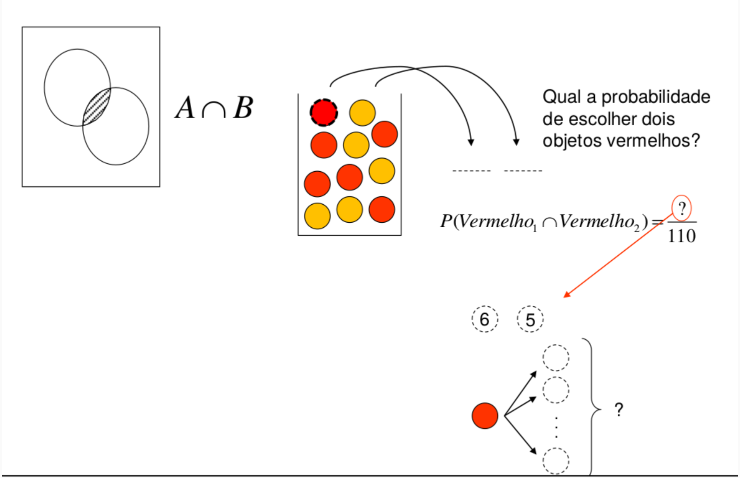
Probabilidade Condicional

- **Exemplo de Probabilidade condicional**



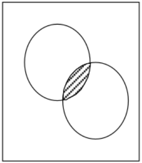
Probabilidade Condicional

- **Exemplo de Probabilidade condicional**

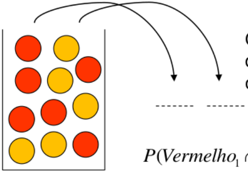


Probabilidade Condicional

- **Exemplo de Probabilidade condicional**



$A \cap B$



Qual a probabilidade de escolher dois objetos vermelhos?

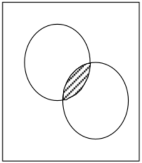
$$P(Vermelho_1 \cap Vermelho_2) = \frac{30}{110}$$

$$\frac{6}{11}\frac{5}{10} = \frac{6.5}{11.10}$$

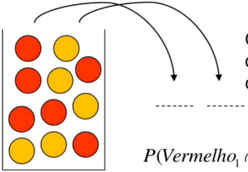
$P(Vermelho_2 \text{ sabendo que } Vermelho_1)$

$P(Vermelho_2 / Vermelho_1)$

- **Exemplo de Probabilidade condicional**



$A \cap B$



Qual a probabilidade de escolher dois objetos vermelhos?

$$P(Vermelho_1 \cap Vermelho_2) = \frac{30}{110}$$

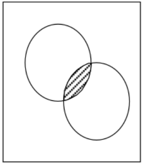
$$\frac{6}{11}\frac{5}{10} = \frac{6.5}{11.10}$$

$$P(Vermelho_2 \text{ sabendo que } Vermelho_1)$$

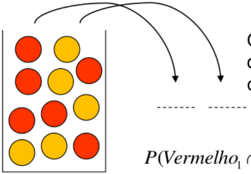
$$P(Vermelho_2 / Vermelho_1)$$

Probabilidade Condicional

- **Exemplo de Probabilidade condicional**



$A \cap B$



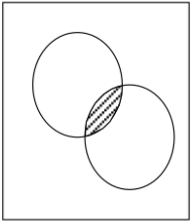
Qual a probabilidade de escolher dois objetos vermelhos?

$$P(Vermelho_1 \cap Vermelho_2) = \frac{30}{110}$$

$$P(Vermelho_1 \cap Vermelho_2) = P(Vermelho_1).P(Vermelho_2 / Vermelho_1)$$
$$= \frac{6}{11}.\frac{5}{10} = \frac{30}{110}$$

Probabilidade Condicional

- **Resumindo**



$A \cap B$

$$\boxed{\begin{aligned} P(A \cap B) &= P(A).P(B/A) \\ &= P(B).P(A/B) \end{aligned}}$$

Probabilidade Condicional

- **Exemplo 4:** Raio X do tórax para detectar tuberculose

	Tuberculose		
Raio X	Não	Sim	Total
Negativo	1739	8	1747
Positivo	51	22	73
Total	1790	30	1820

a)

$$P(\text{um paciente escolhido ao acaso ter tuberculose}) = \frac{30}{1820} = 0{,}016$$

b)

$$P(\text{ter tuberculose} \mid \text{raio X positivo}) = \frac{P(\text{ter tuberculose} \cap \text{raio X positivo})}{P(\text{raio X positivo})} = \frac{\frac{22}{1820}}{\frac{73}{1820}} = \frac{22}{73} = 0{,}301$$

Independência

- Dois eventos A e B são independentes se a informação da ocorrência ou não de A não altera a probabilidade de ocorrência de B. Isto é,

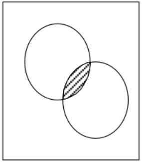
$$P(B|A) = P(B)$$

- ou ainda de forma equivalente (pela regra do produto):

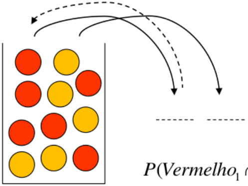
$$P(A \cap B) = P(A).P(B)$$

Independência

- **Exemplo de Independência**



$A \cap B$



Qual a probabilidade de escolher dois objetos vermelhos com reposição?

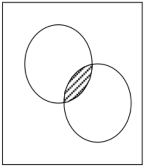
$$P(Vermelho_1 \cap Vermelho_2) = \frac{?}{?}$$

$$P(Vermelho_1 \cap Vermelho_2) = \frac{6}{11} . \frac{6}{11}$$ (eventos independentes)

$$= P(Vermelho_1) . P(Vermelho_2)$$

Independência

- **Resumindo**



$$A \cap B$$

$$\boxed{\begin{aligned} P(A \cap B) &= P(A).P(B/A) \\ &= P(B).P(A/B) \end{aligned}}$$

$$P(A/B) = P(A) \ e \ P(B/A) = P(B) \quad \therefore \quad P(A \cap B) = P(A).P(B)$$

(eventos independentes)

Independência

Exemplo 5:

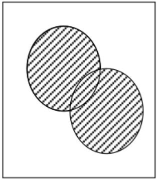
	Doença		
Exposição	Sim	Não	Total
Sim	40	60	100
Não	60	90	150
Total	100	150	250

$$P(\text{um paciente ter a doença}) = \frac{100}{250} = 0{,}4$$

$$P(\text{ter a doença} \mid \text{houve exposição}) = \frac{40}{100} = 0{,}4$$

"Não há associação entre a doença e a exposição"

Probabilidade - Resumo

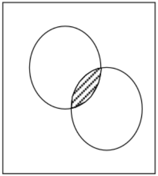


$A \cup B$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A \cup B) = P(A) + P(B)$$

eventos mutuamente exclusivos

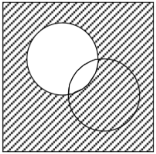


$A \cap B$

$$P(A \cap B) = P(A).P(B/A) = P(B).P(A/B)$$

$$P(A \cap B) = P(A).P(B)$$

eventos independentes



$\bar{A}$

$$P(\bar{A}) = 1 - P(A)$$

Medidas de qualidade de um teste (Sensibilidade e Especificidade)

- Exemplo de Lind & Singer (1986): qualidade da tomografia computadorizada para o diagnóstico de metástase de carcinoma de fígado. (Comparado com laparotomia, um teste padrão ouro). Suponha que nessa população a prevalência (p) de metástase de carcinoma de fígado é de 2%.

Exemplo de Lind & Singer (1986):

Tabela 1: Resultados da tomografia computadorizada em 150 pacientes.

Metástase	Tomografia computadorizada		Total
	Positivo (T_+)	Negativo (T_-)	
Presente (D_+)	52	15	67
Ausente (D_-)	9	74	83
Total	61	89	150

Eventos:

T_+: o resultado do teste é positivo

T_-: o resultado do teste é negativo

D_+: o indivíduo é portador da doença

D_-: o indivíduo não é portador da doença.

Sensibilidade(s) do Teste

- **Exemplo 6:** Qual é a probabilidade do teste ser positivo sabendo que o paciente está doente, ou seja, $P(T_+|D_+)$?
- Se o paciente está doente: o espaço amostral agora está limitado somente aos pacientes doentes (D_+):

	Positivo (T_+)	Negativo (T_-)	Total
Presente (D_+)	52	15	67

- E essa probabilidade pode ser calculada por: $\frac{52}{67} = 0,78$

$$P(T_+|D_+) = \frac{P(T_+ \cap D_+)}{P(D_+)} = 0,78$$

- Essa probabilidade é chamada de sensibilidade(s) do teste.

Especificidade(s) do Teste

- **Exemplo 6:** Qual é a probabilidade do teste ser negativo dado que o paciente não está doente, $P(T_-|D_-)$?

$$P(T_-|D_-) = \frac{P(T_- \cap D_-)}{P(D_-)} = \frac{74/150}{83/150} = 0,89$$

Outras medidas de interesse são:

Falso negativo (1-s)	$P(T_-\|D_+)$
Falso positivo (1-e)	$P(T_+\|D_-)$
Valor preditivo positivo (VPP)	$P(D_+\|T_+)$
Valor preditivo negativo (VPN)	$P(D_-\|T_-)$

Valores Preditivos

- Valor Preditivo Positivo

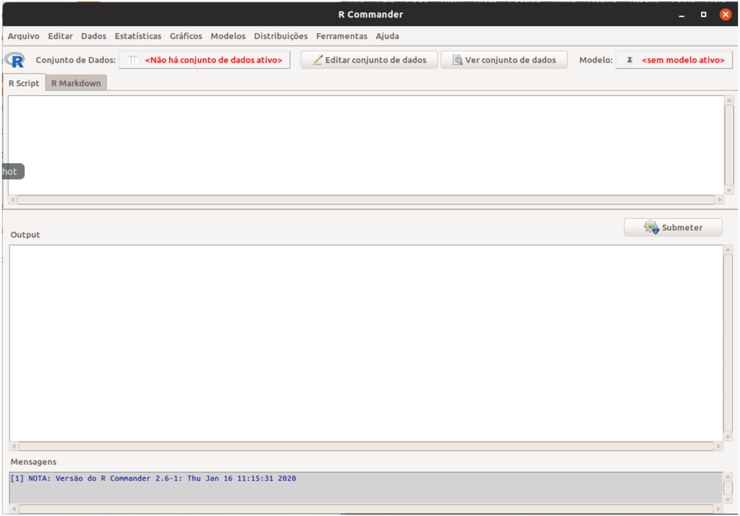
$$VPP = P(D_+|T_+) = \frac{P(D_+ \cap T_+)}{P(T_+)} = \frac{52/150}{61/150} = 0,852$$

- Valor Preditivo Negativo

$$VPN = P(D_-|T_-) = \frac{P(D_- \cap T_-)}{P(T_-)} = \frac{74/150}{89/150} = 0,831$$

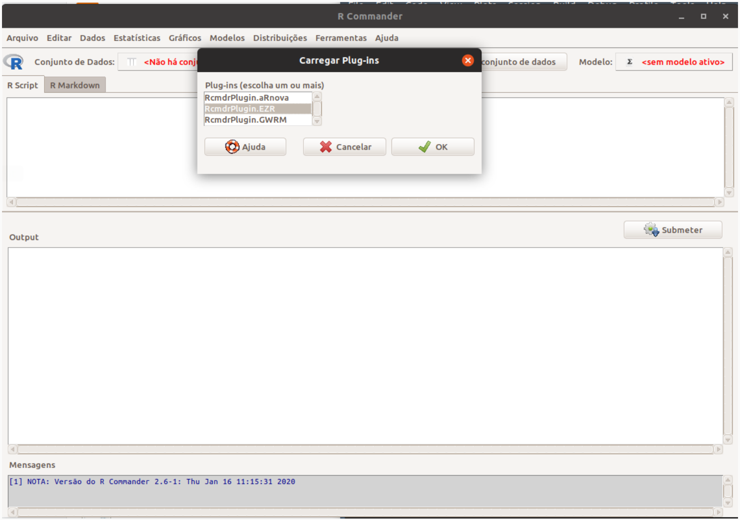
Solução do exemplo utilizando o plugin “RcmdrPlugin.EZR”

- Primeiro abra o Rcommander



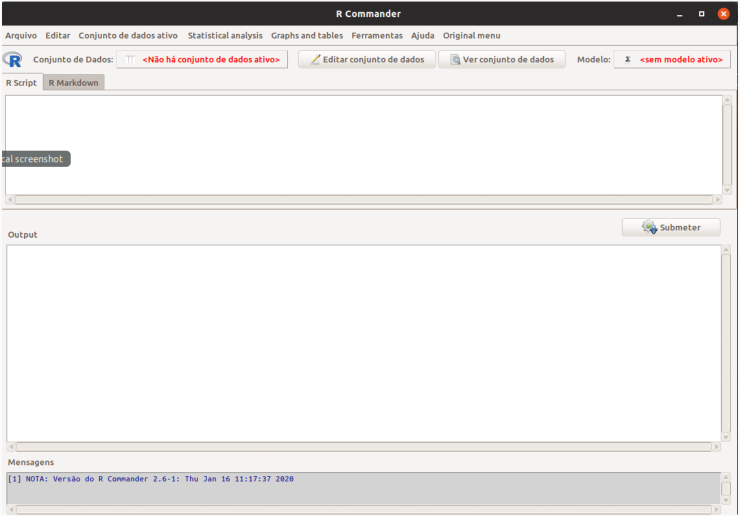
Solução do exemplo utilizando o plugin “RcmdrPlugin.EZR”

- Segundo carregue o plugin RcmdrPlugin.EZR



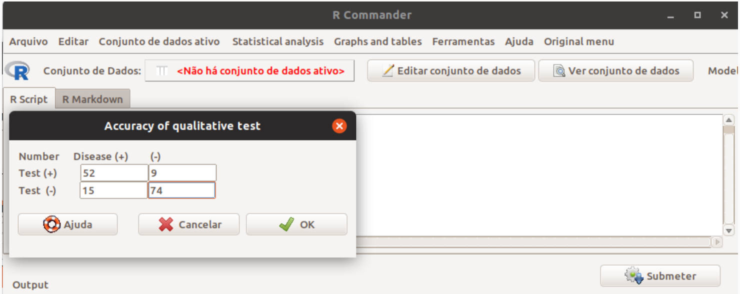
Solução do exemplo utilizando o plugin “RcmdrPlugin.EZR”

- O Rcommander será reinicializado com o RcmdrPlugin.EZR

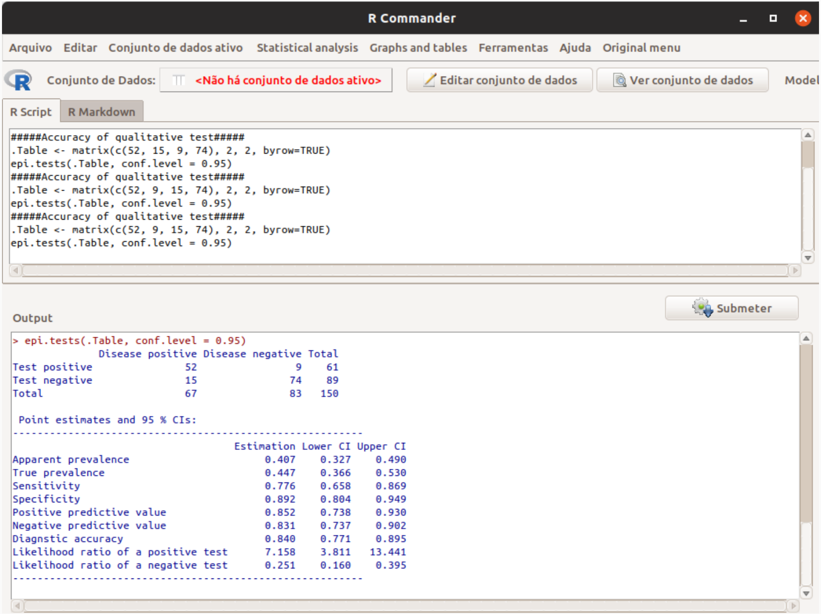


Solução do exemplo utilizando o plugin “RcmdrPlugin.EZR”

- Rcommander → Statistical analysis → Accuracy of diagnostic test → Accuracy of qualitative test



Solução do exemplo utilizando o plugin “RcmdrPlugin.EZR”



Exercício:

Os dados abaixo se referem ao exame de Raio X do tórax para detectar tuberculose.

	Tuberculose		
Raio X	Não	Sim	Total
Negativo	1739	8	1747
Positivo	51	22	73
Total	1790	30	1820

Determine a sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo.

Solução utilizando o plugin “RcmdrPlugin.EZR”