

Universidade Federal do Paraná

Setor de Ciências Exatas

Departamento de Estatística

Pedro Henrique Pavan Gonçalves

**FATORES DE RISCO PARA O ÓBITO POR  
HANTAVIROSE NO PARANÁ, 1992-2016,  
UMA ABORDAGEM POR UM MODELO DE  
FRAÇÃO DE CURA.**

Curitiba

2022

Pedro Henrique Pavan Gonçalves

**FATORES DE RISCO PARA O ÓBITO POR  
HANTAVIROSE NO PARANÁ, 1992-2016, UMA  
ABORDAGEM POR UM MODELO DE FRAÇÃO DE  
CURA.**

Trabalho de Conclusão de Curso apresentado à disciplina Laboratório B do Curso de Graduação em Estatística da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientador(a): Silvia Emiko Shimakura

Curitiba  
2022

Dedico Ã fulano...

# Agradecimentos

Agradeço a mim mesmo e a todos que me ajudaram.

Agradeço também a todo mundo que me ajudou e também aos que não me ajudaram, pois eles também me ajudaram.

*Democracia Ã© oportunizar a todos o mesmo ponto de partida.  
Quanto ao ponto de chegada, depende de cada um.*  
Fernando Sabino.

# Resumo

O resumo deve ser escrito em apenas um parágrafo, e deve ser bastante chamativo para fazer com que o leitor tenha interesse em prosseguir com a leitura. Um bom resumo é sucinto e ao mesmo tempo empolgante. O resumo deve conter um pouco de cada parte do texto, mas deve enfatizar aquilo que é novidade e os principais resultados obtidos.

**Palavras-chave:** Palavra-chave 1. Palavra-chave 2.

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>7</b>
<b>2</b>	<b>OBJETIVOS</b>	<b>8</b>
<b>2.1</b>	<b>Objetivo Geral</b>	<b>8</b>
<b>2.2</b>	<b>Objetivos Específicos</b>	<b>8</b>
<b>3</b>	<b>REVISÃO DE LITERATURA</b>	<b>9</b>
<b>4</b>	<b>MATERIAL E MÉTODOS</b>	<b>10</b>
<b>4.1</b>	<b>Material</b>	<b>10</b>
4.1.1	Conjunto de dados	10
4.1.2	Recursos Computacionais	11
<b>4.2</b>	<b>Métodos</b>	<b>11</b>
4.2.1	Análise de Sobrevivência	11
4.2.2	Modelo de Fração de Cura	11
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>12</b>
<b>5.1</b>	<b>Material final do Estudo</b>	<b>12</b>
<b>5.2</b>	<b>Análise Descritiva</b>	<b>13</b>
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>15</b>
	<b>REFERÊNCIAS</b>	<b>16</b>

# 1 Introdução

A hantavirose, zoonose viral aguda, cuja infecção em humanos no Brasil, se apresenta na forma da Síndrome Cardiopulmonar por Hantavírus, apresentou seus primeiros casos registrados em 1993, e desde então diversos outros estados tem sido comfirmados em todas as regiões do país.

Uma zoonose é uma doença infecciosa que saltou de um animal não humano para humanos. Os patógenos zoonóticos podem ser bacterianos, virais ou parasitários, ou podem envolver agentes não convencionais, se espalhando assim para humanos por contato direto ou por meio de alimentos, água ou meio ambiente. Eles representam um grande problema de saúde pública em todo o mundo devido à nossa estreita relação com os animais, seja na agricultura, como companheiros ou no ambiente natural (OMS, 2020).

A transmissão da hantavirose é feita por roedores. O mais comum é que o contágio ocorra diretamente pela inalação de partículas de urina, fezes e saliva de roedores silvestres, não pelo contato com outros humanos infectados. Por isso, os casos da doença costumam ser isolados. Diferentemente dos seres humanos, roedores, como ratos e ratazanas, podem carregar o hantavírus por toda a vida sem adoecer (BBC, 2021).

Quando olhamos os dados relacionados a essa zoonose, de 2007 a 2015, foram notificados 13.181 casos de hantavirose no Brasil, dos quais 8% ( $N = 1,060$ ) foram confirmados e 3,1% ( $N = 410$ ) evoluíram para óbito. Observou-se uma média de 1.465 casos suspeitos notificados por ano, sendo 2008 ( $N=1.148$ ) e 2013 ( $N=1.804$ ) os períodos de menor e maior número de notificações, respectivamente (OLIVEIRA; DUARTE, 2018).

Segundo a OMS, não há nenhum tratamento, cura ou vacina para a infecção. A alta taxa de mortalidade e dificuldade com o tratamento são fatores que juntamente com a não descoberta de um tratamento tem aumentado a preocupação com a doença que desafia as autoridades de saúde pública ao redor do mundo. As alternativas terapêuticas para os indivíduos infectados limitam-se à introdução de medidas de suporte na fase aguda em ambiente hospitalar, preferivelmente em UTIs.

Dadas as circunstâncias, este presente trabalho de conclusão de curso tem como finalidade estender o trabalho realizado pela DANIELE AKEMI ARITA (cite dani), utilizando os mesmos dados e buscando analisar a sobrevida desses pacientes diagnosticados com hantavirose.

Dado o fato de que nem todos indivíduos do estudo experimentaram o evento de interesse (isto é, não foram a óbito devido à contaminação da hantavirose) até o termino do estudo, é proposta a utilização do modelo de sobrevivência com fração de cura apresentado por Corbiere e Joly (2007), para analisar a sobrevida desses pacientes, onde a variável resposta é o tempo decorrido entre a data do primeiro sintoma do paciente e a data em que o mesmo foi levado a óbito.



## 2 Objetivos

### 2.1 Objetivo Geral

O presente estudo tem como objetivo estudar o tempo até a cura ou óbito, de um paciente infectado Hantavirose no Estado do Paraná no período de 1992 a 2016.

### 2.2 Objetivos Específicos

- a. Revisar a literatura no que diz respeito às abordagens propostas para Análise de Sobrevivência;
- b. Revisar os pacotes disponíveis no software R (cite r) para estimação e diagnóstico de uma modelagem usando Análise de Sobrevivência fazendo uso dos pacotes survival (cite survival) e smcure (cite smcure);
- c. Realizar uma análise descritiva dos dados descritos na Seção 3.1.1 para entendimento mais detalhado e consistente das informações;
- d. Com base nos dados obtidos do banco de monitoramento da Secretaria de Estado da Saúde do Paraná (SESA/PR), realizar uma modelagem usando Análise de Sobrevivência em busca de definir os fatos causadores do óbito ou cura de um paciente infectado por Hantavirose.
- e. Apresentar o modelo, discutir os resultados obtidos e tirar conclusões.

## 3 Revisão de Literatura

Digite aqui sua revisão de literatura.

Veja como fazer citações na introdução.

## 4 Material e Métodos

### 4.1 Material

#### 4.1.1 Conjunto de dados

O conjunto de dados composto por 280 indivíduos e 69 variáveis utilizado é proveniente do Sistema de Informação de Agravos de Notificação (SINAN), onde para todos os indivíduos presentes na base, as informações referentes foram retiradas da ficha de investigação, onde cada um dos 69 campos presentes na ficha foram preenchidos manualmente. A ficha de investigação preenchida pelos pacientes foi anexada aos apêndices do trabalho.

A população do estudo compreendeu todos os casos de hantavirose confirmados no estado do Paraná e que apresentaram início dos sintomas dentro do período do estudo (Janeiro de 1992 a Junho de 2016)

Neste trabalho, a variável resposta de interesse tempo (em dias), foi calculada através do tempo entre as datas data de óbito e a data do 1º sintoma do indivíduo conforme a função a seguir.

$$tempo = \text{data óbito} - \text{data primeiro sintoma}$$

Para os indivíduos que não apresentaram data de óbito, o tempo (em dias) foi calculada através do tempo entre as datas data de encerramento e a data do 1º sintoma do indivíduo conforme a função a seguir.

$$tempo = \text{data encerramento} - \text{data primeiro sintoma}$$

Feito isso, os dados foram divididos entre cura e óbito, sendo cura o “indivíduo notificado por serviço de saúde do Estado do Paraná, no período de estudo e que tenha sido confirmado para hantavirose com evolução para cura” e óbito, o “indivíduo notificado por serviço de saúde do Estado do Paraná, no período de estudo e que tenha sido confirmado para hantavirose com evolução para óbito”.

Pacientes que não apresentaram data de óbito foram considerados como censura.

### 4.1.2 Recursos Computacionais

O software escolhido para a condução do estudo é o software livre R Core Team (2021), que será utilizado como ferramenta para a análise exploratória, bem como ajustar os modelos. Os pacotes survival Therneau (2022) e smcure Cai Yubo Zou e Zhang (2022) serão utilizados no ajuste dos modelos de Análise de Sobrevida.

## 4.2 Métodos

### 4.2.1 Análise de Sobrevida

### 4.2.2 Modelo de Fração de Cura

## 5 Resultados e Discussão

### 5.1 Material final do Estudo

A princípio foi realizada uma análise descritiva dos dados visando um maior entendimento da base.

A base de dados original era composta por 280 observações e 69 variáveis, que são elas os 69 campos presentes na ficha de investigação preenchida pelos pacientes.

Após a importação dos dados para o R, logo de cara foi possível notar uma grande presença de dados *missing*, ou seja, uma grande quantidade de dados faltantes para as variáveis. Tendo em vista esse problema, foi decidido que as variáveis que apresentassem dados faltantes seriam retiradas, fazendo assim com que restassem apenas 14 variáveis para a análise. A seguir temos uma breve apresentação das variáveis:

1. Tempo: Tempo decorrido do primeiro sintoma do paciente até o óbito ou perda do acompanhamento.
2. Idade: Idade do paciente.
3. Sexo: Sexo do paciente (homem ou mulher).
4. Tontura: Apresentou tontura (sim ou não).
5. Cefaleia: Apresentou cefaleia (sim ou não).
6. Sangramento Respiratório: Apresentou sangramento respiratório (sim ou não).
7. Dispneia: Apresentou falta de ar (sim ou não).
8. Hipotensão: Apresentou problemas com pressão baixa (sim ou não).
9. Mialgia: Apresentou dores musculares (sim ou não).
10. Regional de Saúde: Regional de saúde na qual o paciente foi atendido, podendo ser União da vitória, Guarapuava, Irati e Outros (sim ou não).
11. Sinais Hemorrágicos: Apresentou sinais hemorrágicos (sim ou não).
12. Internação: Paciente foi internado no período em que esteve com a doença (sim ou não).
13. Diarreia: Apresentou diarreia (sim ou não).
14. Ureia e Creatina: Aumento de Ureia e Creatina (sim ou não).

Para fins de visualização e melhor entendimento da variável, entendeu-se que seria adequado categorizar a variável *idade*, entre indivíduos de 1 a 20 anos de idade, 20 a 29, 30 a 39, 40 a 49, 50 a 59 e com idade superior a 60 anos.

Conforme já foi discutido anteriormente na seção xxxx, em análise de sobrevivência, as censuras são compostas por indivíduos do estudo que ainda não experimentaram o evento final, portanto, além das 14 variáveis escolhidas que sobraram para a continuação do estudo, ainda foi adicionada a variável de censura, onde o indivíduo que apresentou censura é 1 e 0 para o indivíduo não censurado.

## 5.2 Análise Descritiva

Inicialmente, foi realizada uma análise descritiva dos dados apresentados na Seção 5.1 a fim de analisar a quantidade de indivíduos em cada categoria das variáveis.

Na figura 1, conseguimos analisar a distribuição e comportamento das variáveis idade e sexo, além das quantidades de censura para cada uma.

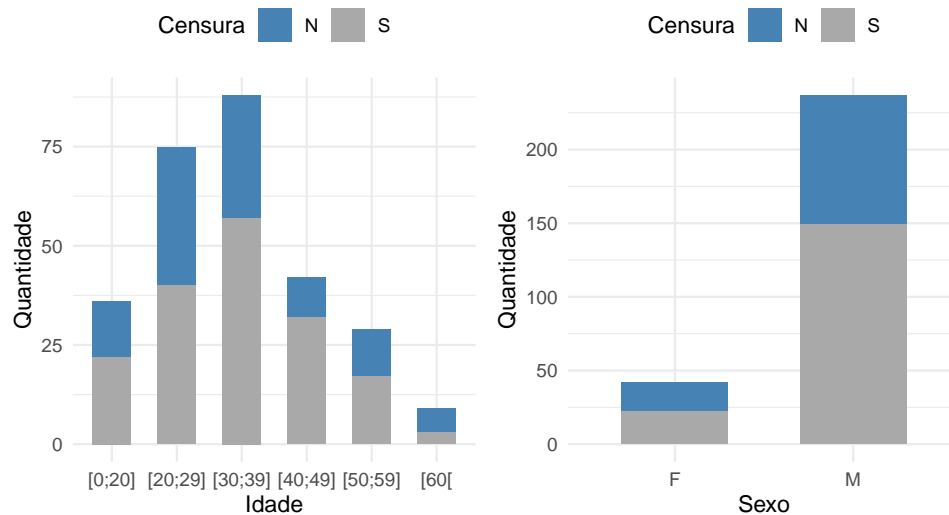


Figura 1 – Distribuição das variáveis Idade e Sexo

Dos 280 indivíduos presentes no estudo, 85,0% (238/280) eram do sexo masculino e os outros 15% (42/280) do sexo feminino, no entanto o sexo feminino apresentou maior letalidade se comparada ao sexo masculino, o que pode ser decorrente da baixa quantidade de observações presentes para o sexo feminino.

Em relação a idade, a média dos indivíduos presentes é de 33 anos que também equivale a mediana dos dados. O indivíduo com a menor idade possui 2 anos, enquanto o indivíduo com a maior idade possui 80.

Entre as categorias da variável idade, 32% (88/280) estão presentes na categoria de indivíduos de 30 a 39 anos, e apenas 3% (9/280) dos indivíduos do estudo presentes na categoria de 60 ou mais anos de idade.

Quanto a variável resposta tempo, conforme a figura 2, cerca de 88% (250/280) dos indivíduos apresentam um tempo igual ou inferior a 22 dias. Um ponto interessante de se observar é o fato de que para um tempo superior a 22 dias, apenas 2 dos 30 indivíduos apresentam censura, ou seja, experimentaram o evento de interesse (óbito).

Outro ponto de atenção é a alta incidência de censura para os tempos iniciais. Cerca de 94% das censuras registradas aconteceram para tempos iguais ou inferiores à 10, e apenas 1 censura entre os indivíduos com tempo superior a 40.

Para a regional de saúde em que ocorreu o atendimento, olhando a figura 3 União da vitória com 33% (93/280), foi a regional de saúde com o maior número de ocorrências, e Irati com 16% (46/280) a com o menor número de ocorrências registradas.

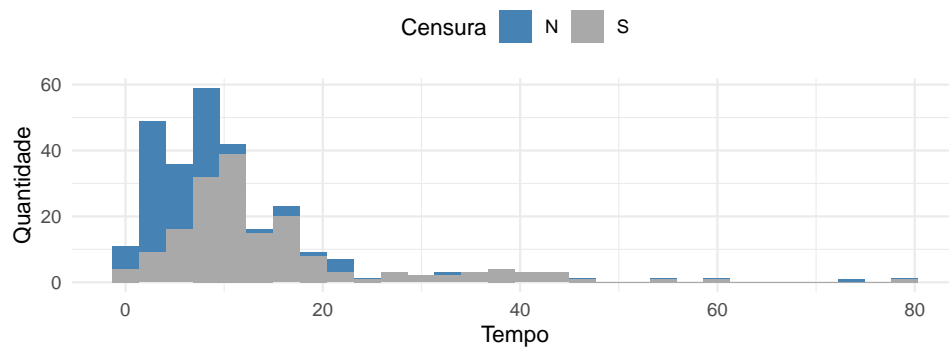


Figura 2 – Distribuição da variável tempo

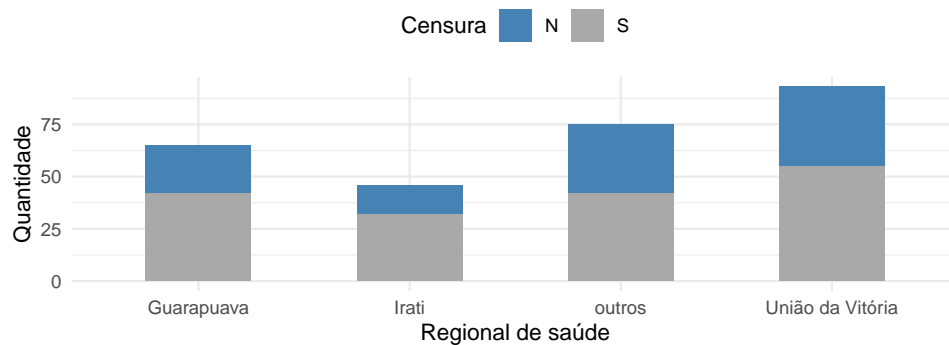


Figura 3 – Distribuição da variável regional de saúde

Quanto as variáveis relacionadas a sintomas, olhando a tabela 1 conseguimos as frequências absolutas e relativas (percentuais) em cada uma das categorias, bem como os percentuais de falhas (apresentar o evento de interesse) e censura.

O sintoma que esteve mais presente entre os infectados foi a febre, presente em 262 dos 280 indivíduos, seguido da cefaleia em 242. O sintoma que se mostrou menos presente

Tabela 1 – Classes gramaticais das palavras presentes na fala dos indivíduos, considerando os grupos de interesse.

..Sintomas	Freq.Absoluta	Freq.Relativa	Censura	Falha
Tontura	1	1	1	1
Cefaleia	2	2	1	1
Sangramento RespiratÃ³rio	3	3	1	1
Dispineia	3	3	1	1
HipotensÃ£o	3	3	1	1
Mialgia	3	3	1	1
Sinais HemorrÃ¡gicos	3	3	1	1
Diarreia	3	3	1	1
Ureia e Creatina	3	3	1	1

## 6 Considerações Finais

Apresente as considerações finais (ou conclusões) do trabalho.



## Referências

BBC. 2021. Disponível em: <<https://www.bbc.com/portuguese/brasil-57245848>>.

CAI YUBO ZOU, Y. P. C.; ZHANG, J. *smcure: Fit Semiparametric Mixture Cure Models*. [S.l.], 2022. R package version 2.1. Disponível em: <<https://CRAN.R-project.org/package=smcure>>.

CORBIERE, F.; JOLY, P. A sas macro for parametric and semiparametric mixture cure models. *Computer Methods and Programs in Biomedicine*, Elsevier Ireland Ltd, v. 85, p. 173–180, 2007. ISSN 01692607. Disponível em: <[https://www.researchgate.net/publication/6641628\\_A\\_SAS\\_macro\\_for\\_parametric\\_and\\_semiparametric\\_mixture\\_cure\\_models](https://www.researchgate.net/publication/6641628_A_SAS_macro_for_parametric_and_semiparametric_mixture_cure_models)>.

OLIVEIRA, S. V. D.; DUARTE, E. C. Magnitude and distribution of deaths due to hantavirus in brazil, 2007-2015. *Epidemiol. Serv. Saude*, v. 27, p. 11, 2018.

OMS. 2020. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/zoonoses>>.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.

THERNEAU, T. M. Survival analysis [r package survival version 3.4-0]. Comprehensive R Archive Network (CRAN), 8 2022. Disponível em: <<https://CRAN.R-project.org/package=survival>>.