

# FATORES DE RISCO PARA ÓBITO POR HANTAVIROSE NO PARANÁ, 1992-2016;

ABORDAGEM VIA UM MODELO DE FRAÇÃO DE CURA.

## Laboratório B

Aluno: Pedro H P Gonçalves

Orientador(a): Dra. Sílvia Emiko Shimakura (UFPR)

# Introdução

## Motivação da Pesquisa

- Hantavirose:
  - Transmitida por ratos do campo.
  - Podem ocorrer por meio de escoriações cutâneas ou mordedura de roedores.
  - Sintomas semelhantes a gripe.
  - Não existe um tratamento específico para as infecções por hantavírus.
  - Entender os fatores de risco para óbito por hantavirose via Análise de Sobrevivência (Modelo de Fração de Cura).
  - Estender o trabalho de Mestrado da Daniele Akemi Arita.

# Introdução

## Análise de Sobrevivência

- Tempo de duração esperado até a ocorrência de um ou mais eventos.
- Área em crescimento na últimas décadas:
  - De 11% em 1979 para 32% em 1989 nos artigos do The New England Journal of Medicine.
- Muito empregada pelas ciências da saúde, biológicas e engenharias.

# Introdução

## Caracterização dos Dados de Sobrevida

- Presença de informações incompletas/parciais.
- Dados comuns poderiam ser explicados por uma regressão simples.
- Principal característica é a presença de censura.
- Tempo de Falha:
  - Tempo até a ocorrência de um evento de interesse (óbito).
- Censura:
  - Observação parcial das respostas.

# Materiais e Métodos

## Conjunto de Dados

- Sistema de Informação de Agravos de Notificação (SINAN).
- Casos confirmados de janeiro de 1992 a junho de 2016.
  - 280 Observações
  - 69 Variáveis
- Software R versão 4.0.4
  - survival.
  - smcure.

# Material e Métodos

República Federativa do Brasil  
Ministério da Saúde

SINAN  
SISTEMA DE INFORMAÇÃO DE AGRAVOS DE NOTIFICAÇÃO  
FICHA DE INVESTIGAÇÃO HANTAVIROSE

Nº

**CASO SUSPEITO:** Paciente com febre, mialgia, cefaléia e insuficiência respiratória aguda de etiologia não determinada, na 1ª semana da doença OU paciente com enfermidade aguda e insuficiência respiratória aguda, com evolução para o óbito na 1ª semana da doença OU paciente com febre, mialgia, cefaléia e exposição a pelo menos uma situação de risco\* até 60 dias anteriores ao início dos sintomas (\*1. atividades de risco para a infecção por hantavírus ou 2. locais com roedores silvestres ou com condições ambientais favoráveis ao seu estabelecimento).

Dados Gerais		Notificação Individual	
1	Tipo de Notificação	2 - Individual	
2	Agravo/doença	HANTAVIROSE	
3	Código (CID10)	A 98.8	
4	UF	5	Município de Notificação
6	Unidade de Saúde (ou outra fonte notificadora)	7	Data dos Primeiros Sinomas
8	Nome do Paciente	9	Data de Nascimento
10	(ou) Idade	11	Sexo M - Masculino F - Feminino I - Ignorado
12	Gestante	13	Raça/Cor
14	Escolaridade	15	Nome da mãe
16	Número do Cartão SUS		

# Materiais e Métodos

## Função de Sobrevivência

- Principal função da Análise de Sobrevivência.
- Probabilidade da falha ocorrer até o tempo  $t$ .

$$S(t) = P(T \geq t)$$

Apresenta uma forma diferente de acordo com o modelo de interesse.

## Estimador de Kaplan-Meier

- Função não paramétrica que utiliza  $S(t)$  em forma de uma função escada com degraus.

$$\hat{S}(t) = P(T > t) = \prod_{j:t_j < t} \left( \frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left( 1 - \frac{d_j}{n_j} \right)$$

# Materiais e Métodos

## Censura

- Tipos de censura
  - Censura tipo 1.
  - Censura tipo 2.
  - Censura Aleatória.
- Função indicadora de censura

$$\delta = \begin{cases} 1, & \text{se } t \text{ é um tempo de falha} \\ 0, & \text{se } t \text{ é um tempo censurado} \end{cases}$$



# Materiais e Métodos

## Modelo de Fração de Cura

- Modelo de fração de cura Semiparamétrico: Tipo especial de modelo de sobrevivência.
- Estimar a taxa de cura do tratamento e a taxa de sobrevivência de pacientes não curados ao mesmo tempo.
- Pode ser expresso como:

$$S_{pop}(t|x, z) = \pi(z)S(t|x) + 1 - \pi(z)$$

, onde  $1 - \pi(z)$  é a probabilidade de um paciente ser curado dependendo de  $z$ .

- $S(t|x)$  é referido como "latência" (não observada).

# Materiais e Métodos

## Modelo de Fração de Cura

- Geralmente utiliza função logito para modelar o efeito de  $z$

$$\pi(z) = \frac{\exp(bz)}{1 + \exp(bz)}$$

- $b$  vetor de parâmetros desconhecidos.
- Permite outras funções de ligação.

# Materiais e Métodos

## Método de Estimação

- Parâmetros estimados via algoritmo EM (algoritmo de maximização de expectativa)
- Função de Verossimilhança:

$$\prod_{i=1}^n [1 - \pi(z_i)]^{1-y_i} \pi(z_i)^{y_i} h(t_i | Y = 1, X_i)^{\delta_i y_i} S(t_i | Y = 1, X_i)^{Y_i}$$

Onde  $\mathbf{O} = (t_i, \delta, z_i, x_i)$  e  $\Theta = (b, \beta, S_0(t))$ .

- Utiliza amostras bootstrap para os erros padrões.
- Permite modelar a parte latente com riscos proporcionais (PH) ou tempo de falha acelerado (AFT).

# Análise Descritiva

## Material final do estudo

- Exclusão de variáveis.
- Todas as observações foram mantidas.

# Análise Descritiva

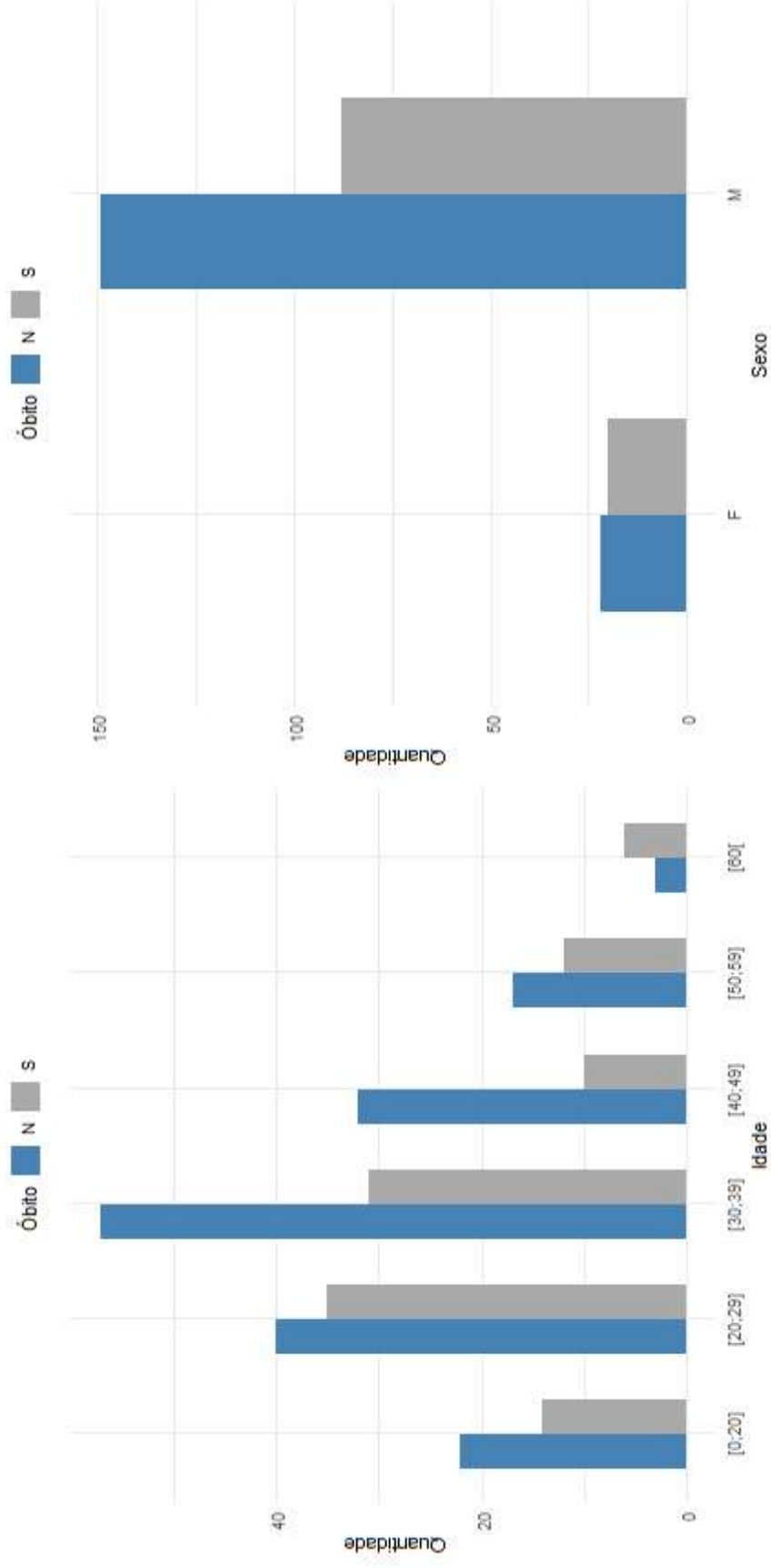
## Material final do estudo

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.
4. **Tontura**: Apresentou tontura.
5. **Cefaleia**: Apresentou cefaleia.
6. **Sangramento Respiratório**: Apresentou sangramento respiratório.
7. **Dispneia**: Apresentou falta de ar.
8. **Hipotensão**: Apresentou problemas com pressão baixa.
9. **Mialgia**: Apresentou dores musculares.
10. **Regional de Saúde**: Regional de saúde na qual o paciente foi atendido.
11. **Sinais Hemorrágicos**: Apresentou sinais hemorrágicos.
12. **Internação**: Paciente foi internado no período em que esteve com a doença.
13. **Diarreia**: Apresentou diarreia.
14. **Respirador Mecânico**: Preciso de respirador mecânico.

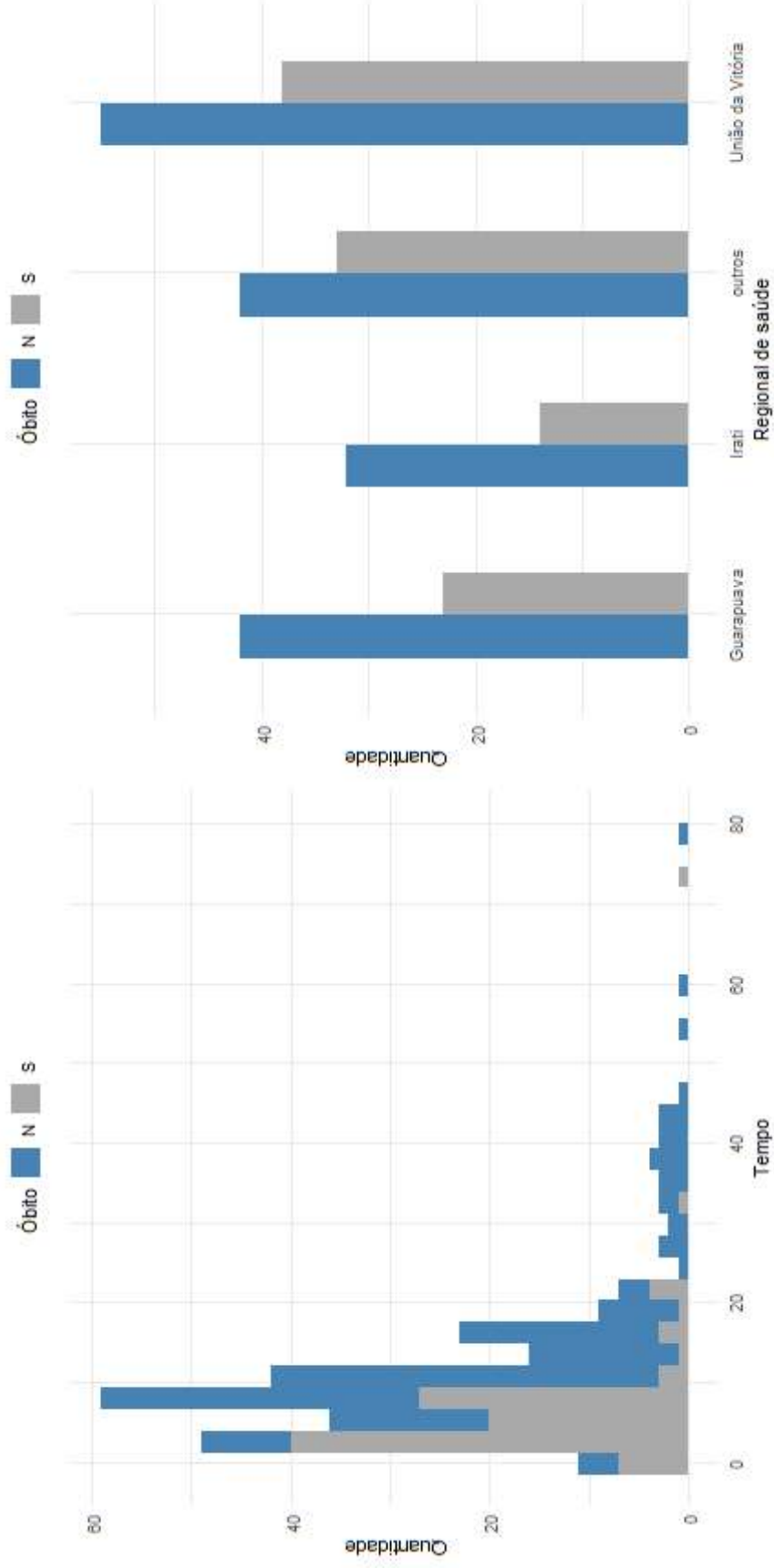
# Análise Descritiva

- Idade:
  - Entre 0 até 19 anos
  - Entre 20 até 29 anos
  - Entre 30 a 39 anos
  - Entre 40 a 49 anos
  - Entre 50 a 59 anos
  - Maior ou igual a 60
- Regional de Saúde:
  - Guarapuava
  - Irati
  - Outros
  - União da Vitória
- Sexo:
  - Masculino
  - Feminino

# Análise Descritiva



# Análise Descritiva



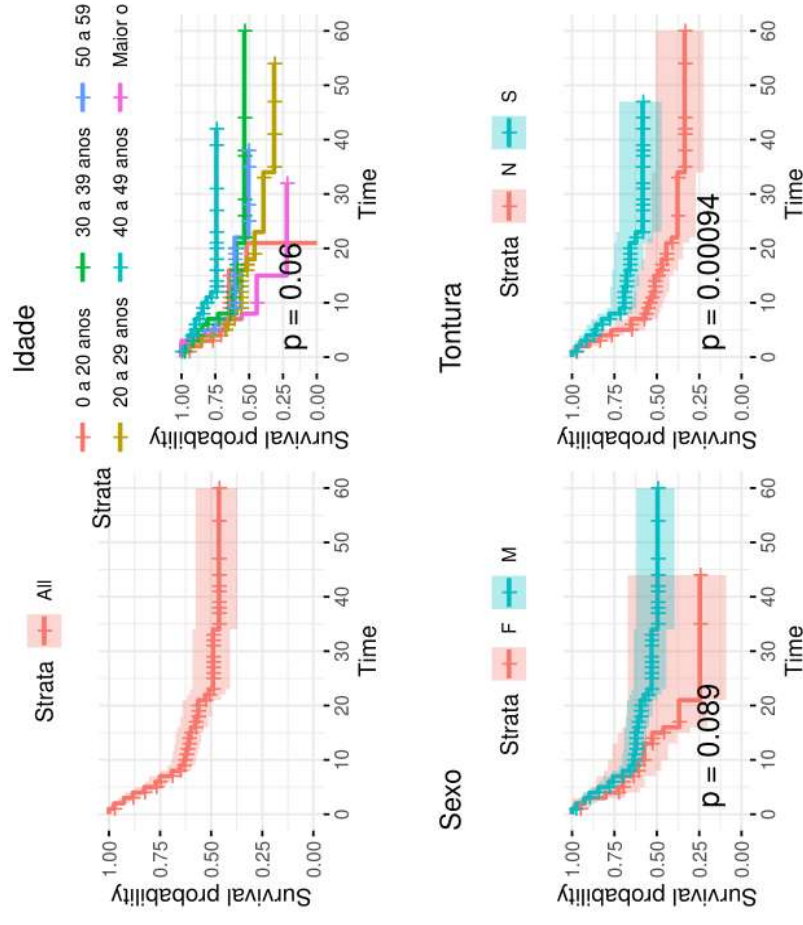


# Análise Descritiva

Variável	Freq.Absoluta	Freq.Relativa	Censura	Falha
Tontura	162	57,50%	49	113
Cefaleia	242	86,43%	87	155
Sangramento Respiratório	123	43,93%	78	45
Dispineia	199	71,07%	92	107
Hipotensão	148	52,86%	94	54
Mialgia	229	81,43%	82	147
Sinais Hemorrágicos	31	11,07%	17	14
Diarreia	72	25,36%	27	45
Respirador Mecânico	118	32,50%	73	45

# Análise Descritiva

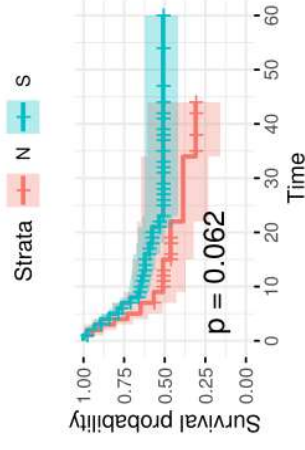
## Estimador de Kaplan-Meier



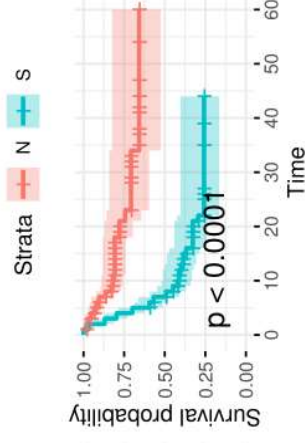
# Análise Descritiva

## Estimador de Kaplan-Meier

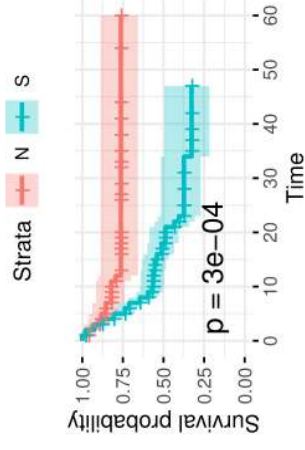
Cefaleia



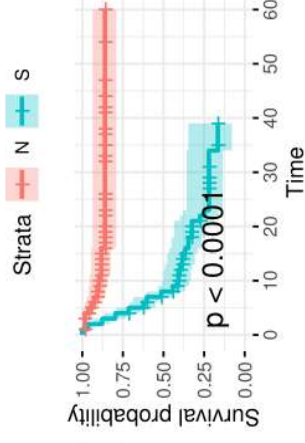
Sangramento Respiratório



Dispneia

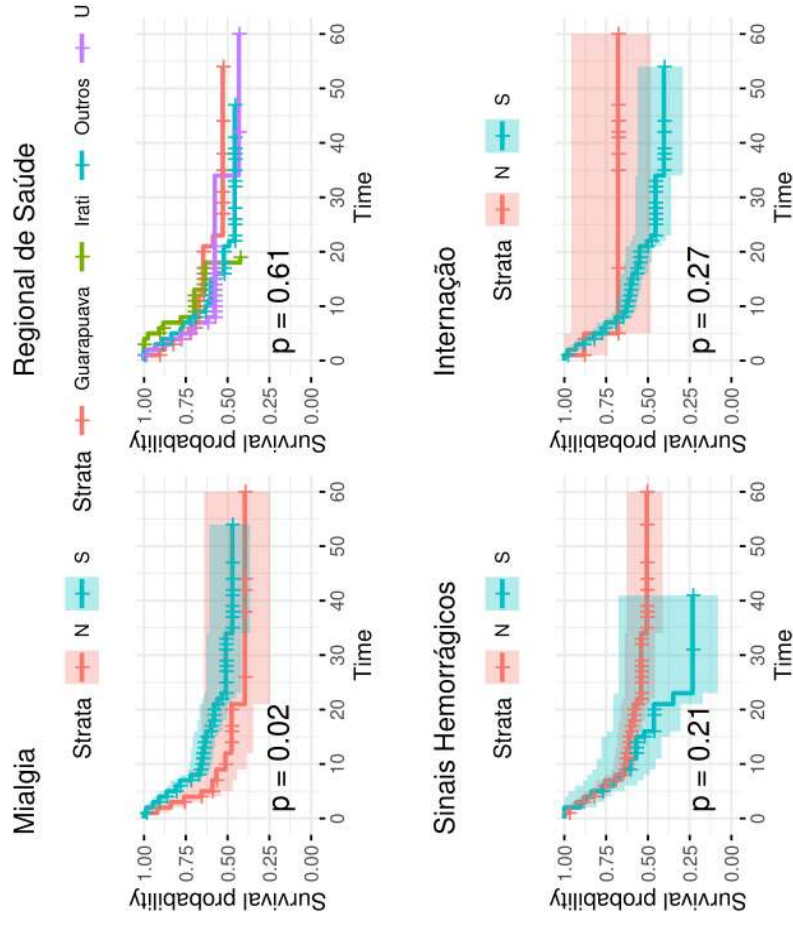


Hipotensão



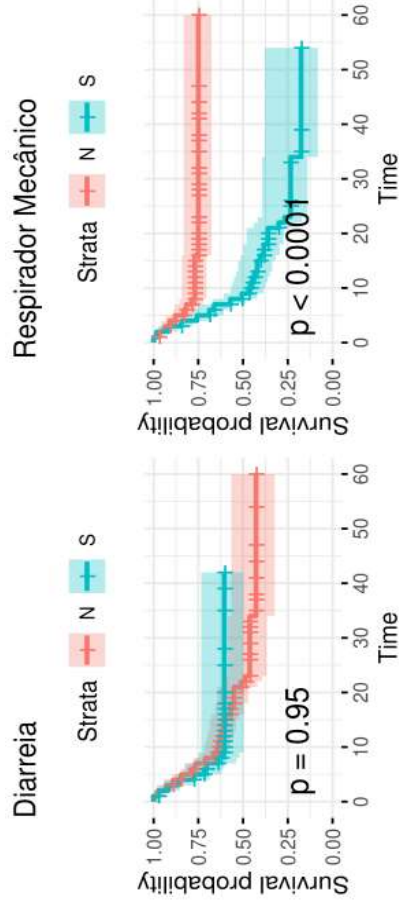
# Análise Descritiva

## Estimador de Kaplan-Meier



# Análise Descritiva

## Estimador de Kaplan-Meier



# Análise Descritiva

## Estimador de Kaplan-Meier

- Resumo dos valores do teste log-rank

Variável	p-valor
Idade	0,06
Sexo	0,089
Tontura	0,0001
Cefaleia	0,062
Sangramento Respiratório	<0,0001
Dispneia	<0,0001
Hipotensão	<0,0001
Mialgia	0,02
Regional de Saúde	0,61
Sinais Hemorrágicos	0,21
Internação	0,27
Diarreia	0,95
Respirador mecânico	<0,0001

# Modelo de Fração de Cura

- Utilização do pacote *smcure*.
- Criação de variáveis dummy.
- Dois ajustes foram realizados:
  - 1º Ajuste: Um modelo para cada variável.
  - 2º Ajuste: Apenas com as variáveis que se mostraram significativas no modelo múltiplo.
- Modelo de Riscos Proporcionais para modelar a parte latente.

# Modelo de Fração de Cura

- 1º Ajuste

Parâmetro	Categoria	Probabilidade de óbito			Tempo de falha		
		Coefficiente	Erro Padrão	Valor-p	Coefficiente	Erro Padrão	Valor-p
Idade	-	-0,005	0,156	0,768	-0,122	0,009	0,19
Sexo	Masculino	-1,915	0,612	0,001	-0,173	0,612	0,001
Tontura	Sim	-1,13	0,439	0,001	-0,438	0,239	0,067
Cefaleia	Sim	-1,395	1,262	0,268	-0,253	0,322	0,432
Sangramento Respiratório	Sim	1,997	0,491	0,001	1,092	0,322	0,001
Dispneia	Sim	2,923	0,462	0,001	0,053	0,362	0,883
Hipotensão	Sim	6,673	0,576	0,001	0,402	0,298	0,177
Mialgia	Sim	-0,189	0,478	0,001	-0,498	0,338	0,14
Regional de Saúde	Irati	0,983	0,651	0,001	-0,954	0,505	0,058
-	Outros	0,242	0,523	0,001	-0,88	0,49	0,072
-	União da Vitória	2,297	0,626	0,001	-0,79	0,581	0,174
Sinais Hemorrágicos	Sim	4,469	2,566	0,001	-0,045	0,342	0,896
Internação	Sim	2,489	0,76	0,001	-1,531	0,67	0,224
Diarreia	Sim	-1,636	0,391	0,001	0,943	0,29	0,001
Respirador Mecânico	Sim	7,394	0,585	0,001	-0,741	0,27	0,006



# Modelo de Fração de Cura

- 2º Ajuste (Probabilidade de Óbito):

Parâmetro	Categoria	Coefficiente
Sexo	Masculino	-1,358
Tontura	Sim	-0,886
Sangramento Respiratório	Sim	2,299
Sinais Hemorrágicos	Sim	1,187
Cefaleia	Sim	-0,499

- Probabilidade de óbito pode ser dada por:

$$1 - \pi(z) = 1 - \frac{\exp(\mathbf{bz})}{1 + \exp(\mathbf{bz})} = 0,082$$

- 2º Ajuste (Tempo de Falha):

Parâmetro	Categoria	Coefficiente
Sexo	Masculino	-0,002
Sinais Hemorrágicos	Sim	-0,349
Respirador Mecânico	Sim	0,048
Cefaleia	Sim	-0,020
Hipotensão	Sim	1,595

# Considerações finais

- Muita dificuldade para tratativa e usabilidade das informações.
- Alta quantidade de dados missing.
- Muitos problemas com a utilização do pacote **smcure**.
- Resultados do modelo de Fração de Cura foram satisfatórios. Variáveis semelhantes com o trabalho apresentado pela Daniele Akemi Arita.
- Servir de base para estudos posteriores com dados mais contundentes.

# Referências

- Cai C, Zou Y, Peng Y, Zhang J. smcure: an R-package for estimating semiparametric mixture cure models. *Comput Methods Programs Biomed.* 2012;108(3):1255-1260. doi:10.1016/j.cmpb.2012.08.013.
- Zhang J, Peng Y. Semiparametric Estimation Methods for the Accelerated Failure Time Mixture Cure Model. *J Korean Stat Soc.* 2012;41(3):415-422. doi:10.1016/j.jkss.2012.01.003.
- Berkson J, Gage R. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association.* 1952;47:501–515.
- Boag J. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society Series B* 1949;11(1):15–53.
- Li C-S, Taylor JMG. A semi-parametric accelerated failure time cure model. *Statist. Med.* 2002;21(21):3235–3247.

# Referências

- ARITA, D. A. Survival of persons with hantavirus infection diagnosed in parana state, brazil. Cadernos de Saude Publica, Fundacao Oswaldo Cruz, v. 35, 2019. ISSN 16784464.
- BBC. 2021. Disponível em: <https://www.bbc.com/portuguese/brasil-57245848>.
- BOTELHO, F.; SILVA, C.; CRUZ, F. Artigos de revisão epidemiologia explicada-análise de sobrevivência. 2009. Disponível em: .
- KRUGER, D. H. et al. Hantaviruses—globally emerging pathogens. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology, J Clin Virol, v. 64, p. 128–136, 3 2015. ISSN 1873-5967. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25453325/>.
- 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.

# Referências

- COLOSIMO, E. A. et al. Análise de sobrevivência aplicada. Revista Entreteses, p. 77, 2006. Disponível em: <http://cursodegestaoelideranca.paginas.ufsc.br/files/2016/03/Apostila-Orienta%CC83o-ao-TCC.pdf>.
- TEAM, R. D. C. Download R-4.0.4 for Windows. The R-project for statistical computing. 2021. Disponível em: <https://cran.r-project.org/bin/windows/base/>.
- 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>.

Fim

Obrigado.

