

Análise de sobrevivência e sua aplicação nas ciências sociais

Juliana Betini Fachini Gomes
Leylannne Nayra Figueira de Alencar
Departamento de Estatística

Brasília - 2018

COMO CHEGUEI NA CIÊNCIA POLÍTICA...

- Professor Denilson Bandeira Coêlho
- Márcio Corrêa de Mello
- Pedro Lucas de Moura Palotti
- Parceria formalizada com a co-orientação de doutorado do Márcio
- Hoje!

- Difusão de Políticas entre governos nacionais e subnacionais
- Segundo Coêlho, Cavalcante e Turgeon (2016):

"O interesse central em estudos de difusão é saber o que causa a adoção de um novo programa por outros governos. Por que alguns governos adotam determinadas políticas públicas enquanto outros governos não adotam? Por que certas inovações de políticas se disseminam? O que explica a velocidade do processo de difusão?"

PROGRAMA SAÚDE DA FAMÍLIA

- Adoção do Programa Saúde da Família (PSF)
- Histórico:
 - projeto baseado no Programa de Agentes Comunitários de Saúde (PACS) criado em 1991 pelo Ministério da Saúde
 - Objetivo: era reduzir a mortalidade infantil e materna, especialmente nas regiões Norte e Nordeste
 - o PACS se desenvolveu nos níveis estadual e municipal, consolidando-se em 1994
 - o sucesso do PACS incentivou a formulação do PSF em 1994

PROGRAMA SAÚDE DA FAMÍLIA

- Nossos objetivos:
 - estudar o que causa a adoção do Programa Saúde da Família
 - quais fatores de natureza política, institucional e regional influenciaram os governos municipais aderirem ou não ao PSF

ANÁLISE DE SOBREVIVÊNCIA

"Análise de sobrevivência é uma área da Ciência Estatística que reúne técnicas e métodos para estudar dados relacionados ao tempo decorrido até a ocorrência de determinado fenômeno de interesse. Geralmente, em análise de sobrevivência também tem-se o interesse em verificar quais fatores biológicos, sociais, econômicos e políticos podem influenciar na ocorrência do fenômeno de interesse."

ANÁLISE DE SOBREVIVÊNCIA

- nas Engenharias - "Análise de Confiabilidade"
- nas Ciências Sociais - "Análise do Histórico de Eventos - Event History Analysis".

TEMPO DE FALHA

A variável resposta é definida como o tempo até a ocorrência de um evento de interesse. Em análise de sobrevivência esse tempo é denominado **tempo de falha**.

Para definir o tempo de falha é preciso:

- 1 fixar o tempo de início do estudo,
- 2 a escala de medida a ser utilizada e
- 3 estabelecer o evento de interesse, que é conhecido como falha.

Outra importante definição em Análise de Sobrevida é a **censura**.

- **Censura** refere-se à perda de informação decorrente de não se ter observado a data de ocorrência do desfecho, resultando em observações parciais ou incompletas.
- As observações censuradas devem ser incluídas na análise dos dados, pois mesmo sendo incompletas, essas observações fornecem informações sobre o tempo de vida de objetos e indivíduos e a omissão das censuras pode acarretar conclusões viciadas na análise estatística.

CARACTERÍSTICAS DOS DADOS DE SOBREVIVÊNCIA

Os dados de sobrevivência são, em geral, compostos pelas seguintes informações: **tempo de falha**, **censura** e **variáveis explicativas ou covariáveis**.

- **As variáveis explicativas ou covariáveis** - podem influenciar no tempo de sobrevivência ou de censura. Exemplos: idade, sexo, salário, escolaridade, região do município, ano eleitoral, alinhamento partidário proximidade com o governo Estadual ou Federal, entre outros fatores.

VOLTANDO AO NOSSO ESTUDO:

- **tempo de falha:** tempo até governos municipais adotarem o Programa Saúde da Família
 - tempo inicial=1997
 - tempo final=2011
 - escala de medida= dias
- **tempo de censura:** tempo até governos municipais Não adotarem o Programa Saúde da Família devido ao término do estudo
- **variáveis explicativas ou covariáveis:**
 - número efetivo de partidos (NEP), margem de vitória, alinhamento federal, ano eleitoral, ideologia, percentual de ganhador, competição política, porte, arrecadação própria do município, taxa de urbanização, % da população até 17 anos, % da população até 65 anos, capacidade fiscal, região, vizinho, região metropolitana (RM),

VARIÁVEIS EXPLICATIVAS

- As variáveis:
 - número efetivo de partidos,
 - margem de vitória,
 - percentual de ganhador,
 - ideologia,
 - arrecadação própria do município,
 - taxa de urbanização,
 - % da população até 17 anos,
 - % da população até 65 anos

forem medidas nos anos eleitorais de 1996, 2000, 2004, 2008

- Criação de novas variáveis
- 4137 municípios

COMO REALIZAR A ANÁLISE DOS DADOS?

- Dados contínuos: Modelos de Regressão $\left\{ \begin{array}{l} \text{de Cox} \\ \text{Paramétrico} \end{array} \right.$
- Dados discretos: Modelos de Regressão $\left\{ \begin{array}{l} \text{de Cox} \\ \text{Paramétrico} \end{array} \right.$
- Qualidade do ajuste do modelo

Modelo de Regressão Paramétrico

ANÁLISE DOS DADOS

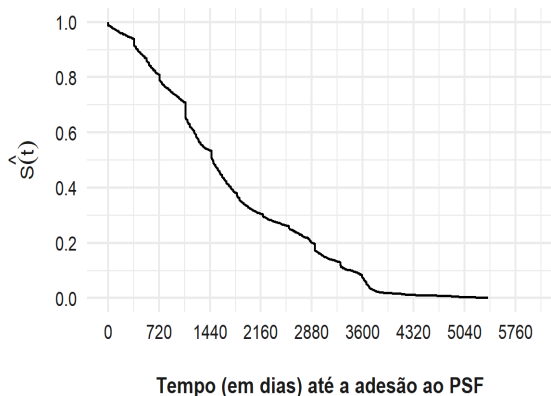


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier para os dados do PSF.

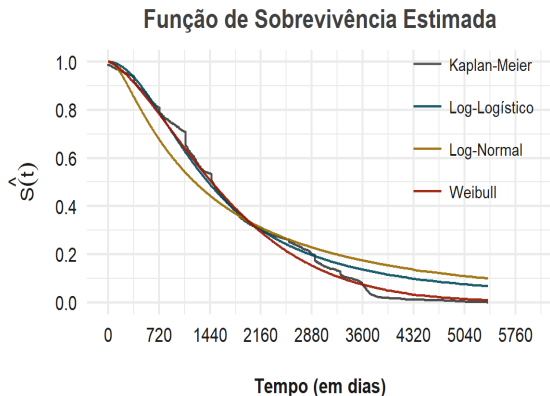


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier e funções de sobrevivência estimativa por diferentes modelos para os dados do PSF.

ANÁLISE DOS DADOS

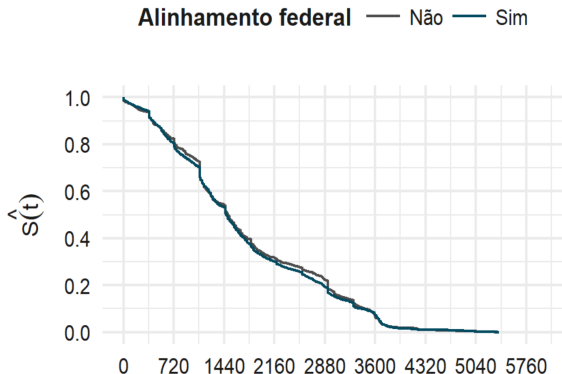


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier considerando a variável Alinhamento Federal para os dados do PSF.

ANÁLISE DOS DADOS

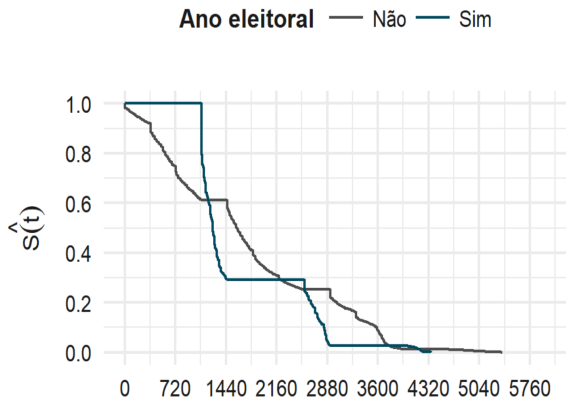


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier considerando a variável Ano Eleitoral para os dados do PSF.

ANÁLISE DOS DADOS

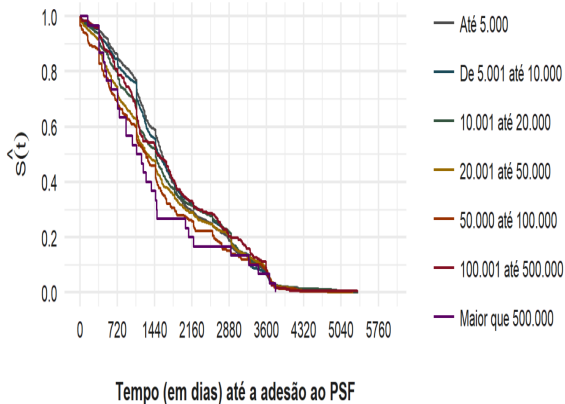


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier considerando a variável Porte do município para os dados do PSF.

ANÁLISE DOS DADOS

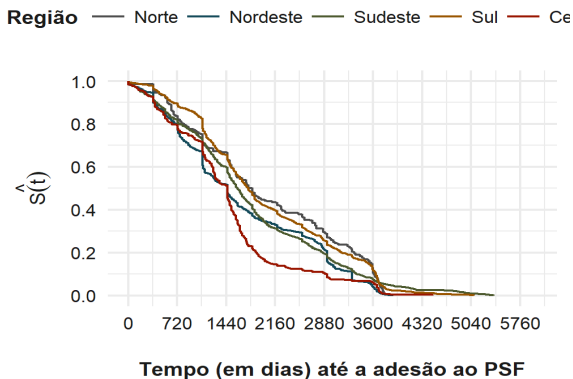


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier considerando a variável Região que o município pertence para os dados do PSF.

ANÁLISE DOS DADOS

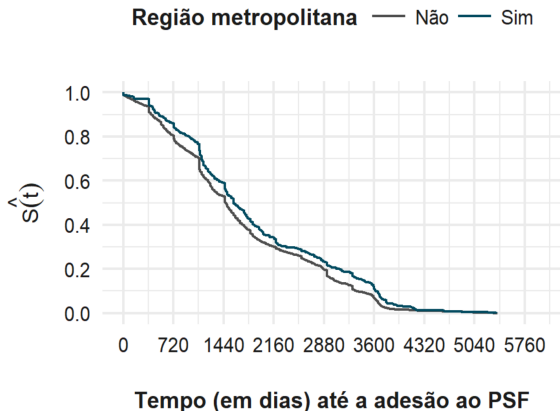


FIGURA : Função de sobrevivência estimativa pelo método de Kaplan-Meier considerando a variável Região Metropolitana para os dados do PSF.

ANÁLISE DOS DADOS

TABELA : Estimativas de máxima verossimilhança para o Modelo de regressão Weibull completo, $\hat{\gamma} = 1,7050$

Variável	Estimativa	Erro padrão	P-valor
Intercepto	10,179	0,240	0,000
NEP	0,027	0,023	0,232
Margem de Vitória	-0,057	0,067	0,401
Percent_ganhador	0,112	0,101	0,266
Alinhamento federal	-0,069	0,022	0,002
Ano eleitoral	0,008	0,022	0,712
Ideologia	0,009	0,005	0,067
População até 17 anos	-5,525	0,292	<0,001
População acima de 65 anos	-0,930	0,804	0,248
Vizinho	-0,002	0,001	<0,001
Região Metropolitana	0,137	0,036	<0,001
Arrecadação própria	0,0003	0,000	0,012
Taxa de Urbanização	0,004	0,051	0,941
Capacidade fiscal média	-0,922	0,193	<0,001
Nordeste	-0,399	0,045	<0,001
Sudeste	-0,764	0,048	<0,001
Sul	-0,762	0,052	<0,001
Centro-Oeste	-0,902	0,053	<0,001
Porte populacional abaixo de 5 mil	0,225	0,135	0,095
Porte populacional entre 5 mil e 10 mil	0,312	0,133	0,019
Porte populacional entre 10 mil e 20 mil	0,290	0,131	0,026
Porte populacional entre 20 mil e 50 mil	0,244	0,128	0,057
Porte populacional entre 50 e 100 mil	0,163	0,129	0,207
Porte populacional entre 100 mil e 500 mil	0,223	0,127	0,078

ANÁLISE DOS DADOS

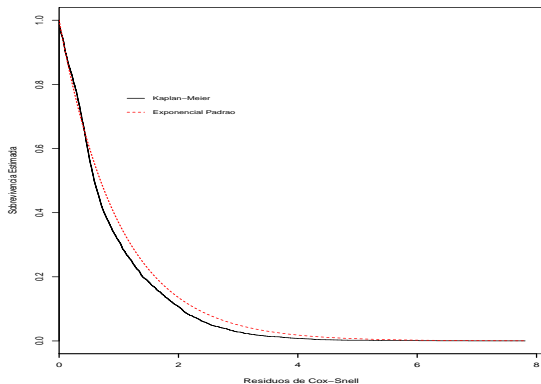


FIGURA : Resíduo de Cox-Snell do Modelo de Regressão Weibull completo para os dados do PSF.

ANÁLISE DOS DADOS

TABELA : Estimativas de máxima verossimilhança para o Modelo de regressão Weibull, $\hat{\gamma} = 1,7046$

Variável	Estimativa	Erro padrão	P-valor
Intercepto	10,197	0,169	0,000
Alinhamento federal	-0,071	0,022	0,001
Ideologia	0,008	0,005	0,0868
População até 17 anos	-5,309	0,202	<0,001
Vizinho	-0,002	0,001	<0,001
Região Metropolitana	0,144	0,035	<0,001
Arrecadação própria	0,0002	0,000	0,0131
Capacidade fiscal média	-0,875	0,186	<0,001
Nordeste	-0,407	0,044	<0,001
Sudeste	-0,762	0,048	<0,001
Sul	-0,759	0,051	<0,001
Centro-Oeste	-0,889	0,052	<0,001
Porte populacional abaixo de 5 mil	0,176	0,129	0,174
Porte populacional entre 5 mil e 10 mil	0,266	0,127	0,0367
Porte populacional entre 10 mil e 20 mil	0,244	0,126	0,0521
Porte populacional entre 20 mil e 50 mil	0,201	0,123	0,104
Porte populacional entre 50 e 100 mil	0,125	0,125	0,315
Porte populacional entre 100 mil e 500 mil	0,190	0,123	0,122

ANÁLISE DOS DADOS

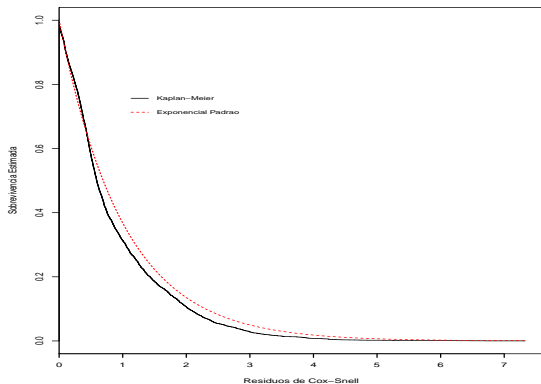


FIGURA : Resíduo de Cox-Snell do Modelo de Regressão Weibull Final para os dados do PSF.

- Melhorar o modelo apresentado

- Estudar mais de um fenômeno de difusão sob diferentes abordagens metodológicas

Obrigada!

REFERÊNCIAS

- COELHO, D. B.; CAVALCANTE, P.; TURGEON, M. Mecanismos de difusão de políticas sociais no Brasil: uma análise do Programa Saúde da Família. *Revista de Sociologia e Política*, v.24, p. 145-165, 2016.
- COLOSIMO, E. A.; GIOLO, S. R. Análise de Sobrevivência Aplicada. Edgard Blucher, São Paulo, 2006.
- KALBFLEISH, J. D.; PRENTICE, R. L. The Statistical Analysis of Failure Time Data. John Wiley & Sons, New York, 2002.
- LAWLESS, J. F. *Statistical Methods and Models for Lifetime Data*. John Wiley & Sons, New York, 1982.