

# EPI10 - Análise de Sobrevida

## Apresentações

Rodrigo Citton P. dos Reis  
citton.padilha@ufrgs.br

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE MEDICINA  
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

Porto Alegre, 2021



# O professor

Olá!



# Olá!

- ▶ Desde outubro de 2017 eu sou Professor do Departamento de Estatística e faço parte do Corpo Docente do Programa de Pós Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul (UFRGS). Além disso, eu atuo como pesquisador no Estudo Longitudinal de Saúde do Adulto (ELSA-Brasil).
- ▶ Eu me formei Bacharel em Estatística pelo Departamento de Estatística da UFRGS em 2007, e Mestre (2010) e Doutor (2014) em Estatística pelo Programa de Pós Graduação em Estatística da Universidade Federal de Minas Gerais.
- ▶ A minha dissertação de mestrado, intitulada *Técnicas estatísticas para avaliação de novos marcadores de risco: aplicações envolvendo o Modelo de Cox*, foi orientada pelos Professores Enrico A. Colosimo e Maria do Carmo P. Nunes.

# Olá!

- ▶ A minha tese de doutorado, intitulada *Análise hierárquica de múltiplos sistemas reparáveis*, foi orientada pelos Professores Enrico A. Colosimo e Gustavo L. Gilardoni.
- ▶ Os meus interesses de pesquisa são Inferência causal em epidemiologia, Análise de mediação, Modelos de predição de risco e Análise de sobrevida.
- ▶ Em estatística aplicada eu tenho interesse na epidemiologia do Diabetes Mellitus.

# A disciplina

# Objetivos

- ▶ Apresentar métodos básicos de análise de sobrevida, de programas estatísticos para realização de análises e interpretação de resultados.
- ▶ Proporcionar aos alunos conhecimentos e ferramentas para avaliação crítica de métodos e resultados em publicações.

# Organização

- ▶ **Disciplina:** Análise de Sobrevida
- ▶ **Turma:** U
- ▶ **Modalidade:** Ensino remoto emergencial ([Moodle](#))
- ▶ **Professor:** Rodrigo Citton Padilha dos Reis
  - ▶ e-mail: [citton.padilha@ufrgs.br](mailto:citton.padilha@ufrgs.br) ou [rodrigocpdosreis@gmail.com](mailto:rodrigocpdosreis@gmail.com)
  - ▶ Sala: B215 do Instituto de Matemática e Estatística

# Aulas e material didático

- ▶ **Aulas** (teóricas e práticas)
  - ▶ Exposição e **discussão** dos conteúdos
    - ▶ Faremos leituras semanais de artigos e capítulos de livros
  - ▶ Exemplos
- ▶ **Notas de aula**
  - ▶ Slides
  - ▶ Arquivos de rotinas em R
- ▶ **Exercícios**
  - ▶ Listas de exercícios
  - ▶ Para casa
  - ▶ Questionários do Moodle
- ▶ **Canais de comunicação:**
  - ▶ Durante as aulas
  - ▶ Moodle: aulas, materiais, listas de exercícios
  - ▶ Sala de aula virtual: notas das avaliações
  - ▶ e-mail do professor

# Aulas e material didático



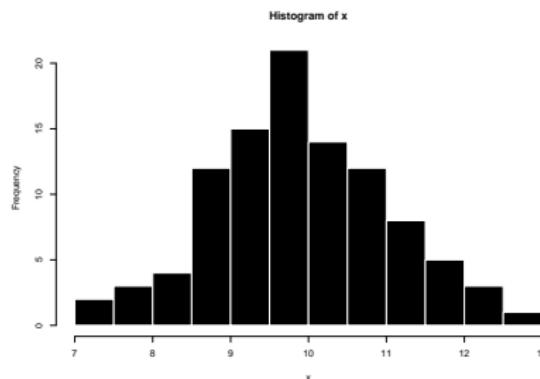
- ▶ **Aulas:** Quintas-feiras, das 15hs às 18hs, no MConf do Moodle da disciplina.
  - ▶ As aulas serão realizadas de maneira **síncrona** com **gravação** e disponibilizadas para posterior consulta.

# Aulas e material didático



- ▶ Exemplos e exercícios com o apoio do computador:
  - ▶ R e RStudio

```
hist(x, col = 'black', border = 'white')
```



# Conteúdo programático

- ▶ **Área 1**
  - ▶ Introdução e conceitos básicos
  - ▶ Especificando o tempo de sobrevida
  - ▶ Estimação e comparação das curvas de sobrevida
- ▶ **Área 2**
  - ▶ Análise de regressão utilizando o modelo de riscos proporcionais
  - ▶ Diagnóstico do modelo
  - ▶ Extensões do modelo de Cox
- ▶ **Área 3**
  - ▶ Tópicos especiais: eventos múltiplos, riscos competitivos, modelos de fragilidade

# Avaliação

- ▶ Será realizado uma avaliação de cada área por meio de questionários ou tarefas do Moodle
- ▶ Cada atividade de avaliação vale 10 pontos
- ▶ Será realizado um teste no Moodle (individual) como atividade de recuperação (*TR*)
  - ▶ Para os alunos que não atingirem o conceito mínimo para aprovação
  - ▶ **Este teste abrange todo o conteúdo da disciplina**

# Avaliação

$$MF = \frac{A_1 + A_2 + A_3}{3}$$

- ▶ **A:**  $9 \leq MF \leq 10$
- ▶ **B:**  $7,5 \leq MF < 9$
- ▶ **C:**  $6 \leq MF < 7,5$
- ▶ Se  $MF < 6$  o aluno poderá realizar o teste de recuperação e neste caso

$$MF' = MF \times 0,4 + TR \times 0,6$$

- ▶ **C:**  $MF' \geq 6$
- ▶ **D:**  $MF' < 6$

# Referências bibliográficas



## Principais

- Colosimo, E. A. e Giolo, S. R. **Análise de sobrevida aplicada**, Blucher, 2006.
- Carvalho, M. S., Andreozzi, V. L., Codeço, C. T., Campos, D. P., Barbosa, M. T. S. e Shimakura, E. S. **Análise de sobrevida: teoria e aplicações em saúde**, 2<sup>a</sup> ed. Editora Fiocruz, 2011.

## Complementares

- Kleinbaum, D. G. e Klein, M. **Survival Analysis: A Self-Learning Text**, 3<sup>a</sup> ed. Springer, 2012.

- └ O que é a análise de sobrevida e como ela é aplicada?

## O que é a análise de sobrevida e como ela é aplicada?

# O que é a análise de sobrevida?

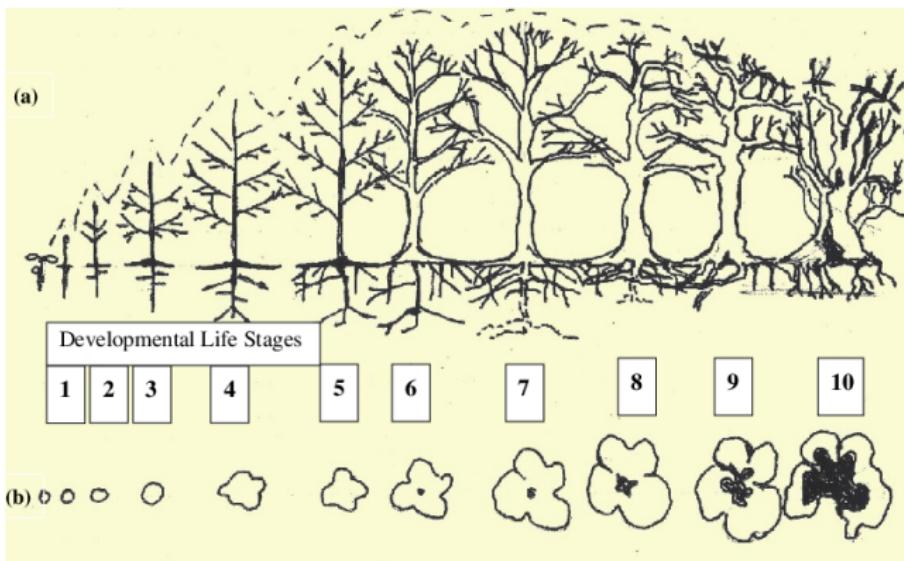
- ▶ Antes de abordarmos esta questão, vamos pensar sobre o que “sobrevive”.
- ▶ Nos casos considerados aqui, estamos falando sobre coisas que têm um **tempo de vida**.
  - ▶ Aquelas coisas que “nascem”, **vivem, mudam de status** enquanto vivem e depois **morrem**.
- ▶ Portanto, “**sobrevida**” é a **descrição de um período de vida** ou um processo de vida antes da ocorrência de uma mudança de status<sup>1</sup>.

---

<sup>1</sup>Em análise de sobrevida, antes da ocorrência de um **evento**

# O que é a análise de sobrevida?

- ▶ Em termos de “sobrevida”, o que pensamos primeiro são organismos como várias espécies animais e outras formas de vida.
- ▶ Após o nascimento, uma entidade viva cresce, passa por um processo de envelhecimento e então se decompõe gradualmente.



# O que é a análise de sobrevida?

- ▶ As mudanças graduais e desenvolvimentos ao longo de um curso de vida refletem o **processo de sobrevida**.
- ▶ Para os seres humanos em particular, **sobrevivemos da morte, doença e deficiência funcional**.

# O que é a análise de sobrevida?

- ▶ Embora a biologia seja sua base primária, o significado da sobrevida é amplamente social.
- ▶ Em diferentes fases da vida, frequentamos a escola, casamos, desenvolvemos uma carreira profissional e nos aposentamos ao envelhecer.



# O que é a análise de sobrevida?

- ▶ Nesse ínterim, muitos de nós vivenciamos desestruturação familiar, envolvemo-nos em atividades sociais, cultivamos hábitos e passatempos pessoais e fazemos ajustes em nossa vida diária de acordo com as condições físicas e mentais.
- ▶ Essas facetas sociais são coisas que não são organismos, mas seu tempo de vida é como o de um ser vivo: coisas que vivem, coisas que têm começos, transformações e depois mortes.

# O que é a análise de sobrevida?

- ▶ Em um contexto mais amplo, a sobrevida também pode incluir eventos como quebra de um automóvel, o colapso de um sistema político em um país ou a realocação de uma unidade de trabalho.
- ▶ Em casos como esse e em outros, a existência dita processos de sobrevida e sua mudança de status, indicada pela ocorrência de eventos.

## Aplicação da análise de sobrevida

- ▶ A prática da análise de sobrevida é o uso da razão para **descrever, medir e analisar características de eventos** para fazer previsões sobre não apenas a sobrevida, mas também sobre **processos de tempo até o evento**<sup>2</sup> - o período de tempo até a mudança de status ou a ocorrência de um evento - como:
  - ▶ de vivos a mortos;
  - ▶ de solteiros a casados;
  - ▶ de saudáveis doentes.

---

<sup>2</sup> Time-to- event processes.

## Aplicação da análise de sobrevida

- ▶ Como a “passagem da vida”, genética, biológica ou mecânica, pode ser interrompida por doença, violência, fatores ambientais ou outros, muitas pesquisas em análise de sobrevida envolvem a realização de **comparações entre grupos** ou **categorias de uma população**;
  - ▶ ou ainda, examinar as **variáveis que influenciam** seus processos de **sobrevida**.
- ▶ No campo acadêmico, a análise de sobrevida é amplamente utilizada em uma longa lista de ciências aplicadas, devido à considerável disponibilidade de **dados longitudinais** que registram históricos de diversos processos de sobrevida e as ocorrências de diferentes eventos.

## Exemplos

- ▶ Na **pesquisa médica**, os **ensaios clínicos** são usados regularmente para avaliar a eficácia de novos medicamentos ou tratamentos de doenças.
  - ▶ Os pesquisadores utilizam a análise de sobrevida para comparar o **risco de óbito** ou **de recuperação de doença** entre grupos populacionais que recebem diferentes medicamentos ou tratamentos.
  - ▶ Os resultados dessa análise, por sua vez, podem fornecer informações importantes com implicações para a definição de políticas.

O que é a análise de sobrevida e como ela é aplicada?

# Exemplos

THE NEW ENGLAND JOURNAL OF MEDICINE

ORIGINAL ARTICLE

## Tofacitinib in Patients Hospitalized with Covid-19 Pneumonia

Pereira O, Gonçalves M.D., Ph.D., Daniel Quirk, M.D., M.P.H., Renzo H. Forteles, M.D., Ph.D., Luis N. Mata, M.D., Ph.D., José F. Soriano, M.D., Ph.D., Marilice O. Antunes, M.D., Ph.D., Roberto Kátil Filho, M.D., Ph.D., Wagner M. Lacerda, M.D., Alexandre P. Tognoni, M.D., Ph.D., Vilmar G. Cunha, M.D., Ph.D., Priscila M. Soares, M.D., Ph.D., Ana Paula G. Góes, M.D., Ph.D., Bruno S. Sampaio, B.Sc., Silvia R.L. Azeiz, M.S., Ronald V.P. Soares, Ph.D., Luciana P.A. Pinto, Ph.D., Kleber Corrêa, M.B.A., Roberto G.R.A.P. Moreira, Ph.D., Frederico Moreira, M.Sc., Helio P. Gonçalves, M.D., Ph.D., Darciel G. da Cunha, M.D., Ph.D., Mariana P. Freitas, M.D., Leover M. Gouveia, M.D., Júper Dourado, Ph.D., Luiz V. Rezende, M.D., Ph.D., Tatiana Koncz, M.D., Ph.D., and Cláudio Bewerunge, M.D., Ph.D.,  
for the STOP-COVID Trial Investigators\*

oxygen devices), 7 (status of being hospitalized while receiving invasive mechanical ventilation or ECMO), or 8 (death) on the eight-level National Institute of Allergy and Infectious Diseases (NIADDK) ordinal scale of disease severity (on a scale from 1 to 8, with higher scores indicating a worse condition) (Table 1; see the Supplementary Appendix). Patients who were enrolled in the trial while they were receiving oxygen through high-flow devices (category 6) were considered to have met the criteria for the primary outcome if they presented with clinical worsening to category 7 or 8. The occurrence of the primary out-

come results for the primary efficacy outcome were analyzed by means of binary regression with Firth correction, with trial group and anti-viral therapy for Covid-19 as covariates, and are expressed as a risk ratio. The antiviral treatments on day 1 were used in the statistical model. Death or respiratory failure through day 28 was analyzed in a manner similar to that used for the primary outcome. The effect of the intervention on death through day 28 is expressed as a hazard ratio derived from Cox regression. For ordinal data, a proportional-odds model with adjustment for baseline antiviral therapy was used. An odds ratio

of less than 1.0 represents a clinical improvement as assessed on the ordinal scale. Odds proportionality was assessed with the use of the method of Pullensteiner-Robinson.<sup>17</sup> We created Kaplan-Meier survival curves to express the time until the occurrence of the primary outcome, both overall and stratified according to the use of supplemental oxygen at baseline, and the occurrence of death through 28 days.

sis of Covid-19 to randomization was 5 days. The median body-mass index (the weight in kilograms divided by the square of the height in meters) was 29.7. A total of 50.2% of the patients had hypertension, and 23.5% had diabetes mellitus. At baseline, 75.4% of the patients were receiving supplemental oxygen, 78.5% were being treated with glucocorticoids, 77.9% were receiving prophylactic anticoagulation, and 20.8% were re-

### TOFACITINIB IN COVID-19 PNEUMONIA

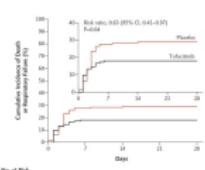


Figure 2. Cumulative incidence of the primary outcome.

The primary outcome was death or respiratory failure through day 28. The hazard ratio for the value of the primary outcome was calculated by means of binary regression with Firth correction, with trial group and ordinal day of antiviral therapy for Covid-19 as covariates. The inset shows the same data on an expanded scale.

N ENGL J MED 2021; 385: 21-30

DOI: 10.1056/NEJMoa2034222

The New England Journal of Medicine  
Downloaded from nejm.org by Rodriguez Rios on August 25, 2021. For personal use only. No other uses without permission.  
Copyright © 2021 Massachusetts Medical Society. All rights reserved.

# Exemplos

## Articles ■

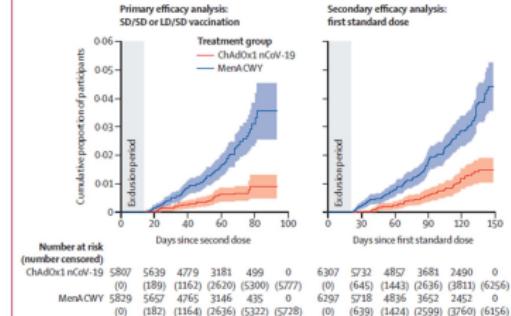
### Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK



63

Mariah Ringer\*, Susan Costa-Chavez\*, Shabir Madhi\*, Li Y-Medley\*, Pedro M Almagro†, Penelope Edes, Bruno Argote-Vieira‡, Sofia Pacheco, Pedro M Almagro, Rosalind Cole-Jones, Oliver Corah, Thomas J Dorton, Karsten Driessé, Christopher J Duncan, Katherine B Ewer, Farzana Iqbal, Jason Lee Failey, Saad Laftiti, Shashank Mehta, Michael J McElroy, Peter D Maitland, Goodwin Catherine M Green, Christopher A Greaves, Paul F Heath, Catherine Helliwell, Helen Hill, Ian Hirsh, Steven Hollingshead, Paul Hunter, Helen Ibbotson, Heather Jackson, Karen Janssen, Daniel Kelleher, James Kirkham, Jennifer Larmore, Alison M Lees, Alice Lubell, Vanessa Lutje, Parikh Lilja, Julian Mallory, Ana V A Mendes, Eudice P Miller, Angela M Mousavifar, Alastair McGregor, David Morrison, Yemisi Mugidde, Anushka Nana, Peter O'Reilly, Sharmin Palupolu, Ana Pinto, Emma Plowd, Francesco Pizzetti, Lorraine Quinn, Shafiq Rizvi, Samia Sabir, Daniel Salazar, Farzana Suleman, Sophie Tsiatis, Zainab Umar, Elizabeth Vlajicic, Edoardo Spizzo, Rebekah Sutherland, Richard Turner, Jennifer Tuite, Michael Turchar, David F Turner, John Vilemavsky, Target Villanueva, Marion E Weston, Christopher J Williams, Alexander D Ziegler\*, Adrián Yáñez\*, Teresa Zamora\*, Sarah C Gilbert\*, Andrew J Fisher\* on behalf of the Oxford COVID Vaccine Trial Group

Vaccine efficacy was calculated as 1– adjusted relative risk (ChAdOx1 nCoV-19 vs control groups) computed using a Poisson regression model with robust variance.<sup>9</sup> The model contained terms for study, treatment group, and age group (18–55, 56–69, and  $\geq 70$  years) at randomisation. A reduced model that did not contain a term for age was used for models affected by convergence issues due to having few cases in the older age groups. The logarithm of the period at risk for the primary endpoint for pooled analysis was used as an offset variable in the model to adjust for volunteers having different follow-up times during which the events occurred. Cumulative incidence is presented using the Kaplan-Meier method.



**Figure:** Kaplan-Meier cumulative incidence of primary symptomatic, NAAT-positive COVID-19. Cumulative incidence of symptomatic COVID-19 after two doses (left) or after first standard dose in participants receiving only standard-dose vaccines (right). Grey shaded areas show the exclusion period after each dose in which cases were excluded from the analysis. Blue and red shaded areas show 95% CIs. LD/LD=low-dose prime plus standard-dose boost. MenA/CWY=meningococcal group A, C, W, and Y conjugate vaccine. NAAT=nucleic acid amplification test. SD/SD=two standard-dose vaccines given.

## Exemplos

- ▶ Os dados de sobrevida são comumente coletados e analisados nas **ciências sociais**, com tópicos que variam amplamente, desde desemprego até reincidência no uso de drogas, ruptura conjugal, carreiras ocupacionais e outros processos sociais.
- ▶ Na **demografia**, além da análise da mortalidade, os pesquisadores se preocupam com processos de sobrevida como o início do uso de anticoncepcionais, a migração interna e internacional e intervalos entre os primeiros nascidos vivos.

## Exemplos

- ▶ No campo da **saúde pública**, a análise de sobrevida pode ser aplicada à análise da utilização de assistência à saúde.
  - ▶ Essa avaliação é de especial importância para planejadores e acadêmicos, pois os sistemas de serviços de saúde refletem a organização política e econômica de uma sociedade e se preocupa com questões filosóficas fundamentais que envolvem vida, morte e qualidade de vida.

## Exemplos

- ▶ A análise de sobrevida também teve amplas aplicações em algumas outras disciplinas, como **engenharia, ciência política, gestão de negócios e economia**.
- ▶ Por exemplo, na engenharia, os cientistas aplicam a análise de sobrevida para realizar testes de vida na durabilidade de produtos/componentes mecânicos ou elétricos<sup>3</sup>.
  - ▶ Os resultados de tais estudos podem ser usados para a melhoria da qualidade dos produtos.

---

<sup>3</sup>Geralmente, os métodos de análise de sobrevida, quando aplicados em problemas de engenharia, são incorporados aos **métodos de confiabilidade**.

# A história da análise de sobrevida e seu progresso

## Um breve histórico

- ▶ Originalmente, a análise de sobrevida era usada exclusivamente para investigações de mortalidade e morbidade nos registros de estatísticas vitais.
- ▶ A primeira análise aritmética dos processos de sobrevida humana remonta ao **século XVII**, quando o estatístico inglês **John Graunt** publicou a primeira **tábuas de vida** em **1662**.

# Um breve histórico



*Natural and Political*  
**OBSERVATIONS**  
 Mentioned in a following INDEX,  
 and made upon the  
**Bills of Mortality.**  
 By JOHN GRAUNT,  
 Citizen of  
**LONDON.**

With reference to the Government, Religion, Trade,  
 Growth, Ayre, Diseases, and the several Charge of the  
 said CITY.

— Non, me amittere Turba, latere.  
*Centaurus parvus LeBaronis* —

LONDON,  
 Printed by Thos: Roper, for John Martin, James Allardyce  
 and Tis: Dicau, at the Sign of the Bell in St. Paul's  
 Church-yard, MDCLXII.

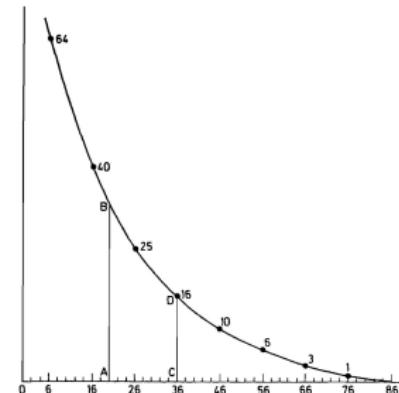
A TABLE of the Number of CHRISTENINGS and MORTALITY						
For the Year 1605 and 1606.*						
Weeks.	Days of the Month.	Christ. Bur.	Bur.	Parish.	Weeks.	Days of the Month.
1 Dec.	26	102	116	5	5	28 July
2 January	1	117	151	6	5	29
3	2	120	154	4	4	30
4	3	121	157	4	4	31
5	4	143	154	6	4	1
6	5	143	161	3	2	2
7	6	121	154	5	5	3
8	7	121	151	5	5	4
9	8	136	169	12	6	5
10	9	103	117	9	3	6
11	10	110	95	7	4	7
12	11	106	97	7	4	8
13	12	120	131	14	11	9
14	13	124	133	17	8	10
15	14	127	134	13	9	11
16	15	121	131	13	9	12
17	16	139	139	14	8	13
18	17	115	110	11	7	14
19	18	1	92	136	17	10
20	19	115	120	11	7	15
21	20	115	94	13	8	16
22	21	113	131	14	9	17
23	22	89	98	9	2	18
24	23	117	121	19	14	19
25	24	127	112	19	14	20
26	25	131	119	15	10	21
27	26	132	126	34	16	22
Christened — — — — —						
The Totals — — — — —						
Buried — — — — —						
Whereof of the Plague — — — — —						
Clarifred — — — — —						
6614						
7920						
1124						
* BELL London's Remembrance.						
A TABLE						
Original from THE OHIO STATE UNIVERSITY						

Digitized by Google

Original from  
THE OHIO STATE UNIVERSITY

# Um breve histórico

Christian Huygens (1629-1695)



## Um breve histórico

- ▶ Por um longo período de tempo, a análise de sobrevida foi considerada um instrumento analítico, principalmente nos estudos biomédicos e demográficos.
- ▶ Numa fase posterior, expandiu-se gradualmente para o domínio da engenharia para descrever/avaliar o curso dos produtos industriais.

## Um breve histórico

- ▶ Nos últimos cinquenta anos, o escopo da análise de sobrevida cresceu tremendamente como consequência do rápido **desenvolvimento da ciência da computação**, particularmente o avanço de poderosos pacotes de software estatístico.
- ▶ A conveniência de usar software de computador para criar e utilizar modelos estatísticos complexos levou cientistas de muitas disciplinas a começar a usar modelos de sobrevida.

## Um breve histórico

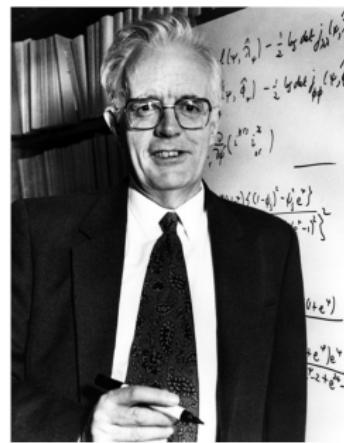
- ▶ Como as aplicações da análise de sobrevida cresceram rapidamente, a inovação metodológica se acelerou em um ritmo sem precedentes nas últimas décadas.
- ▶ O advento do **modelo de Cox** e a perspectiva da **verossimilhança parcial** em **1972** desencadeou o avanço de um grande número de métodos e técnicas estatísticas caracterizadas por modelagem de regressão na análise de dados de sobrevida.

# Um breve histórico

$$\lambda(t) = \lambda_0(t)e^{(x^T \beta)}$$

---

$$\begin{aligned}\ell(\beta|X) &= \sum_j \ln \frac{\exp x_j^T \beta}{\sum_{i:t_i \geq t_j} \exp x_i^T \beta} \\ &= \sum_j \left[ x_j^T \beta - \ln \sum_{i:t_i \geq t_j} \exp x_i^T \beta \right]\end{aligned}$$



## Um breve histórico

- ▶ A principal contribuição do modelo de Cox, dada sua capacidade de gerar procedimentos de estimativa simplificados na análise de dados de sobrevida, é o fornecimento de uma **abordagem estatística flexível** para modelar os processos de sobrevida complexos associados a covariáveis mensuráveis.

## Um breve histórico

- ▶ Mais recentemente, o surgimento da **teoria dos processos de contagem** destaca o início de uma nova era na análise de sobrevida devido ao seu tremendo poder inferencial e alta flexibilidade para modelar **eventos repetidos** para a mesma observação e alguns outros processos de sobrevida complexos.
  - ▶ Em particular, esta perspectiva moderna combina elementos da **teoria de grandes amostras**, a **teoria de martingales** e a **teoria de integração estocástica**, fornecendo um novo conjunto de procedimentos estatísticos e regras na modelagem de dados de sobrevida.

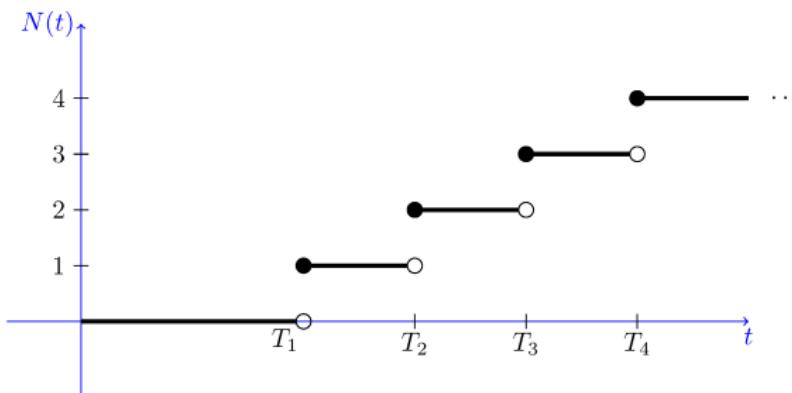
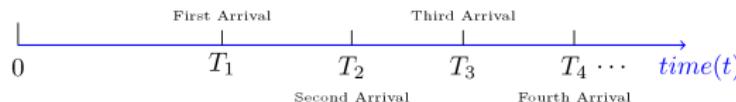
## Características gerais da estrutura de dados de sobrevida

## Estrutura dos dados

- ▶ Em essência, um processo de sobrevida descreve um período de vida de um tempo de início especificado até a ocorrência de um evento particular.
- ▶ Portanto, a **principal característica** dos dados de sobrevida é a descrição de uma **mudança no status** como medida de resultado subjacente.
- ▶ Mais formalmente, uma mudança de status é a ocorrência de um **evento** que designa o fim de um período de vida ou o término de um processo de sobrevida.

# Estrutura dos dados

- ▶ Por exemplo, uma mudança de status ocorre quando uma pessoa morre, se casa ou quando um automóvel quebra.



## Estrutura dos dados

- ▶ Esta característica de um “salto” de status torna a análise de sobrevida um tanto semelhante a algumas perspectivas estatísticas mais convencionais sobre dados de desfechos qualitativos, como o **modelo logístico**.
- ▶ Esses modelos tradicionais também podem ser usados para examinar uma mudança de status ou a ocorrência de um determinado evento, comparando o status no início e o status no final de um intervalo de observação.
- ▶ Entretanto, essas abordagens **ignoram o tempo de ocorrência** desse evento e, portanto, não possuem a capacidade de descrever um processo de tempo até o evento.
  - ▶ A falta dessa capacidade pode ser prejudicial à qualidade dos resultados analíticos, gerando conclusões enganosas.

# Estrutura dos dados

## Exemplo

- ▶ Suponha que dois grupos da população têm a mesma taxa de ocorrência de um determinado evento até o final de um período de observação, mas os membros de um grupo costumam experimentar o evento significativamente mais tarde do que aqueles no outro.
- ▶ O primeiro grupo populacional tem um padrão de sobrevida vantajoso porque sua vida média é estendida.
- ▶ Obviamente, a regressão logística ignora esse fator de tempo, não fornecendo informações precisas.

## Estrutura dos dados

- ▶ A maioria dos modelos de sobrevida leva em consideração o **fator tempo** no salto de status.
- ▶ Dada essa capacidade, a **segunda característica** dos dados de sobrevida é a **Descrição de um processo de tempo até o evento**.
- ▶ Na literatura de análise de sobrevida, **o tempo de ocorrência de um determinado evento** é considerado uma **variável aleatória**, referida como **tempo do evento**, **tempo de falha** ou **tempo de sobrevida**.

## Estrutura dos dados

- ▶ A **terceira característica** principal da estrutura de dados de sobrevida é a **censura**.
- ▶ Os dados de sobrevida são geralmente coletados por um intervalo de tempo no qual as ocorrências de um determinado evento são observadas.
- ▶ Como resultado, os pesquisadores só podem observar os eventos que ocorrem dentro de uma janela de monitoramento entre dois limites de tempo<sup>4</sup>.

---

<sup>4</sup>Tempo de seguimento (*follow up time*).

## Estrutura dos dados

- ▶ Consequentemente, os tempos de sobrevida completos para muitas unidades sob exame não são observados, com perda de informações ocorrendo antes do início ou após o final do intervalo do estudo.
- ▶ Algumas unidades podem ser perdidas para observação no meio de uma investigação devido a vários motivos.
- ▶ Na análise de sobrevida, esse status ausente em tempos de evento é chamado de censura, que pode ser dividido em uma variedade de tipos.

## Estrutura dos dados

- ▶ Para a maioria dos tipos de censura, uma porção do tempo de sobrevida para observações censuradas é observável e pode ser utilizada no cálculo do risco de experimentar um determinado evento.
  - ▶ Na análise de sobrevida, esta porção dos tempos observados é referida como **tempos de sobrevida censurados**.
- ▶ Como a censura ocorre com frequência, a maioria das análises de sobrevida lida literalmente com **dados de sobrevida incompletos** e, consequentemente, os cientistas encontraram maneiras de usar essas informações limitadas para analisar corretamente os dados de sobrevida incompletos com base em algumas suposições restritivas sobre a distribuição dos tempos de sobrevida censurados.

## Estrutura dos dados

- ▶ Como os processos de sobrevida variam essencialmente com base nas características básicas das observações e condições ambientais, um campo considerável de análise de sobrevida é conduzido por meio de **modelagem de regressão de dados censurados** envolvendo uma ou mais variáveis preditoras.

## Estrutura dos dados

- ▶ Dada a adição de **covariáveis**, a **estrutura de dados de sobrevida** pode ser vista como consistindo em informações sobre três fatores principais:
  - ▶ tempos de sobrevida;
  - ▶ status de censura;
  - ▶ e covariáveis.
- ▶ Dada uma **amostra aleatória** de  $n$  unidades, a estrutura de dados para análise de sobrevida na verdade contém  $n$  destas triplas.
- ▶ A maioria dos modelos de sobrevida é construída sobre essa estrutura de dados.

## Estrutura dos dados

- Dadas as diferentes ênfases na variedade de recursos, a análise de sobrevida também é conhecida como **análise de duração**, **análise de tempo até o evento**, **análise de histórico de eventos** ou **análise de dados de confiabilidade**.

## Próxima aula

- ▶ Caracterização dos dados de sobrevida
- ▶ Exemplos
- ▶ Especificação dos tempos de sobrevida

# Por hoje é só!

Sejam tod@s bem-vind@s!

