

## Análise de Sobrevivência Aplicada à Saúde

Lupércio França Bessegato  
Dep. Estatística/UFJF

- Prof. Lupércio França Bessegato  
Departamento de Estatística – UFJF  
E-mail: [lupercio.bessegato@ufjf.edu.br](mailto:lupercio.bessegato@ufjf.edu.br)  
Site: [www.ufjf.br/lupercio\\_bessegato](http://www.ufjf.br/lupercio_bessegato)

Análise de Sobrevivência Aplicada à Saúde -- 2014

## Material de Apoio

- R:  
√ [www.r-project.org](http://www.r-project.org)
- Tutorial online do R:  
√ <http://www.leg.ufpr.br/Rtutorial>  
√ <http://www.leg.ufpr.br/~paulojus/embrapa/Rembrapa>
- Conjuntos de dados e material Análise de Sobrevivência – Carvalho et al. (Fiocruz)  
√ <http://sobrevida.fiocruz.br>

Análise de Sobrevivência Aplicada à Saúde -- 2014

## Roteiro

1. Conceitos Básicos
2. Técnicas Não Paramétricas
3. Modelos Probabilísticos e Inferência
4. Modelos de Regressão Paramétricos
5. Modelos de Regressão de Cox
6. Extensões do Modelo de Cox
7. Tópicos Adicionais
8. Referências

Análise de Sobrevivência Aplicada à Saúde -- 2014

## Introdução

## Modelo de Sobrevida

- Classe de modelos estocásticos usados para analisar características e fatores associados ao tempo até ocorrência de evento de interesse
- Evento de interesse:
  - √ Falha ou desfecho
    - Óbito
    - Nascimento
    - Cura
    - Diagnóstico de uma doença
    - Aparecimento de tumor
    - Aparecimento de efeito colateral de medicamento

Análise de Sobrevida Aplicada à Saúde -- 2014

## Modelos Estatísticos

- Estruturas de causalidade fundamentais em raciocínio epidemiológico e evidências empíricas
- Associação de efeitos principais e de interação a cada variável
- Possibilidade de afirmações probabilísticas para descrição e representação do conhecimento
- Variabilidade associada a predições produzidas pelo modelo (IC's)

Análise de Sobrevida Aplicada à Saúde -- 2014

## Mecanismos de Incerteza

- Erro amostral
  - √ Discrepâncias entre o valor estimado e o valor dos parâmetros são devidos à variação amostral
- Erro de medição:
  - √ Valores de uma mesma característica medida repetidas vezes exibem uma distribuição

Análise de Sobrevida Aplicada à Saúde -- 2014

### Análise de Sobrevivência

- Tempo é o objeto de interesse:
  - √ Tempo até a ocorrência de evento de interesse
  - √ Risco de ocorrência por unidade de tempo
- Exemplos;
  - √ Qual o efeito de anticancerígeno sobre o tempo de sobrevivência?
  - √ Quais os fatores associados ao tempo de duração da amamentação?
  - √ Quais os fatores preditivos para reinternação hospitalar, considerando o tempo entre internações?

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Dimensões de Tempo

- Idade do indivíduo:
  - √ Mortalidade é claramente dependente da idade
- Duração em uma determinada condição
  - √ Dimensão natural em ensaios clínicos
  - √ Ex.:
    - Duração de tempo desde o diagnóstico de Aids até o óbito em pacientes tratados com terapia antirretroviral
    - Tempo desde um evento cirúrgico até recidiva
- Tempo calendário
  - √ Associado ao efeito de período

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Modelo Estatístico de Sobrevivência

- Descrição da força de incidência como função de variáveis explicativas
  - √ Procura-se conhecer o efeito de cada uma delas e suas interações, livres da presença de confusão
  - √ É um modelo de regressão, com os seguintes elementos:
    - Variável resposta
    - Covariáveis explicativas
    - Função de ligação
    - Estrutura de erro

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Variável Resposta

- Probabilidade de sobrevivência:
  - √ Probabilidade de não ocorrer o evento de interesse no período  $[t_0, t]$
  - √  $P\{\text{sobreviver por mais que } t \text{ unidades de tempo}\}$
- Taxa de incidência (ou de *hazard*)
  - √ Taxa instantânea de ocorrência do evento de interesse no tempo  $t$ , condicionada à sua não ocorrência anterior à  $t$ .
- Taxa de incidência acumulada

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Planejamento de Estudo

- Paradigma de planejamento que gera informação de Análise de Sobrevida:
  - √ Delineamento pressupõe o acompanhamento de indivíduos de uma coorte ao longo do tempo
  - √ Variável resposta de excelência:
    - Tempo até o desfecho (categórico)
- Estudos clínicos que envolvem uma resposta temporal são frequentemente prospectivos e de longa duração

Análise de Sobrevida Aplicada à Saúde -- 2014

### Caracterização de Dados de Sobrevida

### Exemplo

- Hemodiálise em pacientes com insuficiência renal
  - √ Registro mensal de sessões de hemodiálise
  - √ Período: Janeiro/1988 a Outubro/2001
  - √ Evento de interesse:
    - Óbito
  - √ Variável tempo:
    - Intervalo entre o início tratamento e o óbito
  - √ Dados: *dialise.csv*

Análise de Sobrevida Aplicada à Saúde -- 2014

- Tempo de sobrevivência (meses) de 10 pacientes

Paciente (i)	Tempo (T <sub>i</sub> )
1	22
2	6
3	12
4	43
5	23
6	10
7	35
8	18
9	36
10	29

√ Tempos conhecidos para todos os indivíduos

Análise de Sobrevida Aplicada à Saúde -- 2014

- Uma característica desses estudos é a presença de informações incompletas ou parciais.

√ Situações frequentes:

- Indivíduos de uma coorte que não são acompanhados até a ocorrência do evento de interesse
- Não se conhece a data de início do acompanhamento do coorte

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Informações parciais

√ Causas de perda da informação temporal completa

- Óbitos por causas não relacionadas ao estudo
- Término do estudo
- Perda de contato
- Recusa em continuar participando
- Mudança de procedimento
- Abandono devido a efeitos adversos de tratamento
- Desconhecimento do início do tempo da modelagem

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Perda da informação temporal completa – Causas

√ Óbitos por causas não relacionadas ao estudo

- Ex.: causas externas

√ Término do estudo

√ Perda de contato

- Ex.: mudança de residência

√ Recusa em continuar participando

(relativamente comum em estudos de coorte)

√ Mudança de procedimento

- Ex.: paciente recebe transplante e sai da coorte de acompanhamento de hemodiálise

√ Abandono devido a efeitos adversos de tratamento

Análise de Sobrevivência Aplicada à Saúde -- 2014

√ Desconhecimento do início do tempo da modelagem

- Ex.: data de infecção em estudo de tempo entre contaminação pelo HIV e o diagnóstico de Aids

√ Utilização de dados prevalentes:

- Estudo de sobrevivência em diálise, pacientes que morreram antes da implementação do sistema de acompanhamento estão excluídos

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Comentários

- As técnicas de Análise de Sobrevivência são frequentemente utilizadas nos estudos de coorte e ensaios clínicos
  - √ É possível seu uso nos demais estudos,
    - os tempos até a ocorrência do desfecho devem estar claramente definidos e obtidos.

Análise de Sobrevivência Aplicada à Saúde --  
2014

### Tipos de Estudos mais Frequentes

- Estudos envolvendo somente uma amostra:
  - √ Descritivo, pois não existe grupo de comparação
  - √ Em geral, seu objetivo é a identificação de fatores de prognóstico para a doença em estudo
- Estudos envolvendo dois ou mais grupos
  - √ Estudos comparativos

Análise de Sobrevivência Aplicada à Saúde --  
2014

### Elementos da Resposta

- Tempo inicial
  - √ Indivíduos devem ser comparáveis na origem do estudo  
(exceto por diferenças medidas por covariáveis)
- Escala de medida
  - √ Em geral, tempo real ou “de relógio”

Análise de Sobrevivência Aplicada à Saúde --  
2014

- Evento de interesse (desfecho)
  - √ Devem ser definidos de forma clara e precisa
  - √ Pode ocorrer devido
    - a uma única causa
    - duas ou mais causas
  - √ Riscos competitivos:
    - Situação em que causas de desfecho competem entre si

Análise de Sobrevivência Aplicada à Saúde --  
2014

## Censura

- Perda de informação decorrente de não se ter observado a data de ocorrência do desfecho
  - √ Mesmos censurados, todos os resultados de um estudo de sobrevivência devem ser usados na análise
    - Mesmo incompletas, as observações censuradas fornecem informações sobre o tempo de vida de pacientes
    - Omissão das censuras pode acarretar conclusões viciadas
- Observação parcial da resposta
  - √ Tempo até desfecho é superior ao observado

Análise de Sobrevivência Aplicada à Saúde -- 2014

## Mecanismos de Censura

- Censura tipo I
  - √ Estudo será concluído após um período de tempo pré-determinado (tempo do estudos não é aleatório)
- Censura tipo II
  - √ Estudo será concluído após a ocorrência do desfecho em um número pré-estabelecido de indivíduos (tempo do estudo é aleatório)

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Censura aleatória:
  - √ É o que mais ocorre na prática:
  - √ Ex.:
    - Paciente é retirado do estudo sem ter ocorrido o desfecho
    - Paciente morre por razão diferente da estudada
  - √ Representação matemática:
    - $T_i$ : tempo até o desfecho do paciente  $i$
    - $C_i$ : tempo de censura desse paciente
    - $T_i$  e  $C_i$  são variáveis aleatórias independentes
    - Observa-se  $t_i = \min\{T_i, C_i\}$
    - Indicadora de desfecho  $\delta = \begin{cases} 1 & , \text{ se } T \leq C \\ 0 & , \text{ se } T > C \end{cases}$

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Censura tipo I é caso particular de censura aleatória
- Em uma amostra  $(T_i, C_i)$ ,  $i = 1, 2, \dots, n$ 
  - √ Se todo  $C_i = C$ , constante sob controle do pesquisador
  - √  $t_i = \min\{T_i, C_i\}$  é uma variável aleatória mista
    - Um componente contínuo e outro discreto (tem probabilidade  $> 0$  em  $t = C$ )

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Tipos de Censura

- Censura à direita
- Censura à esquerda
- Censura intervalar

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Censura à Direita

- Tempo de ocorrência do desfecho está à direita do tempo registrado
  - √ Aproveita-se a informação do tempo durante o qual a pessoa esteve sob observação sem que ocorresse o evento

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Hemodiálise em pacientes com insuficiência renal
  - √ Registro mensal de sessões de hemodiálise
  - √ Período: Janeiro/1988 a Outubro/2001
  - √ Evento de interesse:
    - Óbito
  - √ Variável tempo:
    - Intervalo entre o início tratamento e o óbito
  - √ Dados: *dialise.csv*

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Tempos de observação (em meses) de 10 pacientes em diálise

Paciente (i)	Tempo ( $T_i$ )	Status ( $\delta_i$ )
1	22	1
2	6	0
3	12	1
4	43	0
5	23	1
6	10	1
7	35	1
8	18	0
9	36	1
10	29	1

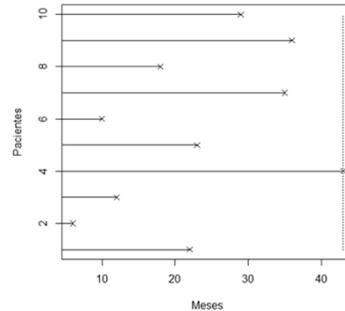
$T_i$ : tempo de observação do indivíduo i  
 $\delta_i$ : variável indicadora do status  
 1, desfecho antes do final do estudo  
 0, indivíduo deixou de ser observado

$(T_3, \delta_3)$ : tempo de sobrevivência é o tempo de acompanhamento  
 $(T_8, \delta_8)$ : tempo de sobrevivência é maior do que tempo observado (informação incompleta)

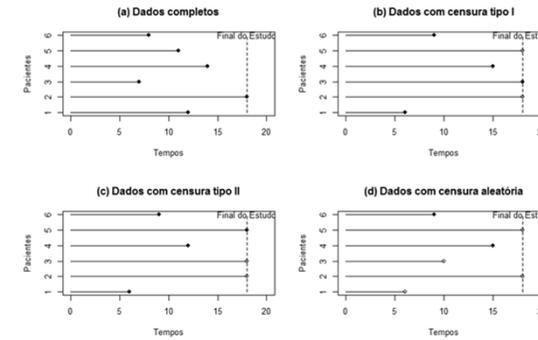
Análise de Sobrevivência Aplicada à Saúde -- 2014

- Tempos de observação (em meses) de 10 pacientes em diálise

Paciente (i)	Tempo ( $T_i$ )	Status ( $\delta_i$ )
1	22	1
2	6	0
3	12	1
4	43	0
5	23	1
6	10	1
7	35	1
8	18	0
9	36	1
10	29	1



Análise de Sobrevivência Aplicada à Saúde -- 2014



- Desfecho ocorrido para todos antes do final do estudo
- Desfecho não ocorreu para alguns até o final do estudo
- Término de estudo após número determinado de falhas
- Interrupções e alguns pacientes sem desfecho

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Censura à Esquerda

- Não se conhece o momento da ocorrência do desfecho, mas sabemos que ele ocorreu antes do tempo observado
  - √ Tempo observado é maior que o tempo de falha

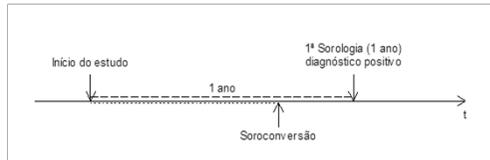
Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estudo para investigar os fatores associados à infecção por leptospirose em comunidade de baixa renda, recém-criada.
  - √ Inclusão de participantes quando de mudança para a área
  - √ Anualmente coleta-se sangue dos participantes para verificar se houve soroconversão

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Censura à Esquerda



- √ Alguns participantes já estavam soropositivos logo na 1ª coleta
  - Tempo para soroconversão (tempo de sobrevivência) é menor que o tempo para o primeiro exame

Análise de Sobrevivência Aplicada à Saúde -- 2014

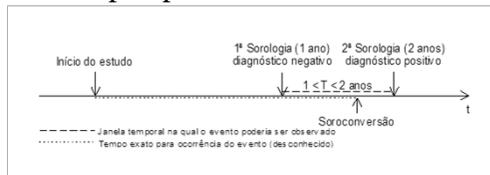
### Censura Intervalar

- Tipo mais geral de censura
  - √ Ex.: estudo em que os pacientes são acompanhados em visitas periódicas e conhece-se apenas que o evento de interesse ocorreu em um certo intervalo de tempo
- Representação:  $(t^- < T < t^+)$ 
  - √  $t^-$  e  $t^+$ : tempos observados

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estudo de leptospirose



- √ Participante em que a soroconversão foi observada no segundo exame
  - Tempo exato de sobrevivência é maior que o tempo até o primeiro exame e menor do que o tempo até o segundo exame
  - $(t^- < T < t^+)$

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estudo de leptospirose
  - √ Tempo exato de sobrevivência é maior que o tempo até o primeiro exame e menor que o tempo até o segundo exame
  - √  $(t^- < T < t^+)$ 
    - $t^-$ : tempo observado entre o início do estudo e o 1º exame
    - $t^+$ : tempo observado do 2º exame

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estudo do tempo entre diagnóstico de Aids e óbito
  - √ 193 pacientes acompanhados em ambulatório
  - √ Período: 1986 a 2000
  - √ Observados 92 óbitos com 101 pacientes vivos
    - Eventos: 92
    - Censuras: 101
  - √ Banco de dados: *ipecc.csv*

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Censura à Direita



- √ Paciente estava vivo pelo menos até dezembro/2000, data do último contato registrado

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Censura Informativa e Não Informativa

- Censura não informativa:
  - √ Motivo da perda de informação não está relacionado com o desfecho
  - √ Ex.:
    - Data final do estudo não foi definida em função do desfecho
- Censura informativa:
  - √ Motivo da perda de informação está relacionado com desfecho
  - √ Ex.:
    - Abandono do tratamento devido à piora do paciente
    - Óbito não conhecido por falha no acompanhamento

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Censuras informativas devem ser evitadas:
  - √ Introduzem um viés de seleção
- Métodos para evitar viés:
  - √ Predefinir todos os passos na busca de pacientes
  - √ Analisar as causas de censura
  - √ Comparar perfis dos pacientes com tempos censurados e não censurados
  - √ Etc.

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Truncamento

- É caracterizado por uma condição que exclui certos indivíduos do estudo
  - √ Indivíduos são excluídos por motivo relacionado à ocorrência do estudo
- Os pacientes não são acompanhados a partir do tempo inicial, mas somente após experimentarem um certo evento
  - √ Estudo inclui somente indivíduos em que o evento ocorreu dentro de uma janela temporal pré-estabelecida ( $T_E$ ,  $T_D$ )

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Truncamento à Esquerda

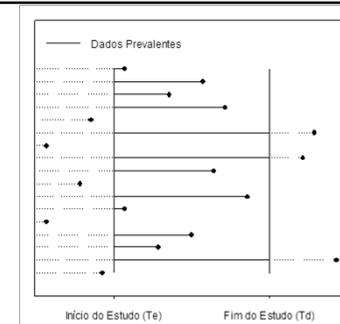
- Inclui somente observações em que o desfecho ocorreu após o limite inferior da janela temporal de observação
  - √  $T \geq T_E$
  - √ Só ocorre quando a perda de informação está relacionada a indivíduos que foram excluídos porque já tinham experimentado o evento antes do início do estudo e não podiam ser observados

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Hemodiálise em pacientes com insuficiência renal
  - √ Registro mensal de sessões de hemodiálise
  - √ Período: Janeiro/1988 a Outubro/2001
  - √ Ingresso no estudo:
    - Pacientes em tratamento em janeiro de 1988 ou que iniciaram o tratamento numa data posterior
  - √ Dados truncados:
    - Pacientes diagnosticados e mortos antes de Janeiro/1988
  - √ Truncamento à esquerda devido a uso de dados prevalentes

Análise de Sobrevivência Aplicada à Saúde -- 2014



- Parte do tempo é observado prospectivamente e parte retrospectivamente
  - √ Informações dos indivíduos que já sofreram o desfecho antes o início do estudo não podem ser incluídos no estudo, embora se esteja considerando os sobreviventes

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Viés de seleção:
  - √ Indivíduos com sobrevivência mais longa têm maior probabilidade de entrar no estudo (subestima risco de ocorrência do evento)
- Correção do viés:
  - √ Pacientes que iniciaram a diálise antes do início do estudo só terão considerados os tempos após a data de início do estudo

Análise de Sobrevida Aplicada à Saúde -- 2014

- Interpretação adequada da modelagem com dados prevalentes:
  - √ Resultados encontrados estão condicionados ao fato de o paciente estar vivo em Janeiro/1988

Análise de Sobrevida Aplicada à Saúde -- 2014

### Truncamento à Direita

- Ocorre quando o critério de seleção inclui somente os indivíduos que sofreram o evento
  - √ Data de ocorrência do evento é sempre menor que o limite superior da janela temporal ( $T \leq T_D$ )
  - √ Comum em estudos de sobrevivência que partem do dado do óbito para selecionar as observações

Análise de Sobrevida Aplicada à Saúde -- 2014

### Exemplo

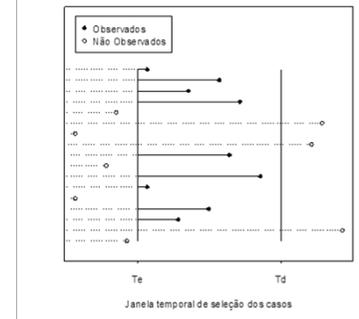
- Análise para estimar fatores associados à sobrevivência de pacientes com Aids que partem do registro de óbito
  - √ Indivíduos que não sofrem o evento não são incluídos no estudo, mesmo tendo todos os fatores de risco presentes.

Análise de Sobrevida Aplicada à Saúde -- 2014

### Exemplo

- Análise para estimar fatores associados à sobrevivência de pacientes com Aids que partem do registro de óbito
  - √ Não haverá censura à direita
  - √ Indivíduos que não sofrem o evento não são incluídos no estudo, mesmo tendo todos os fatores de risco presentes.

Análise de Sobrevivência Aplicada à Saúde -- 2014



- √ Amostra definida pelos eventos na janela temporal
- √ Trajetórias em pontilhado não são observadas (mesmo que cruzem a janela – evento não ocorreu)
- √ Nunca há censura à direita

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Viés de seleção:
  - √ Risco de ocorrência do evento é superestimado
  - √ Em doenças de curta duração e muito bem caracterizadas este viés de seleção será pouco importante

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estimação de tempo de vida de habitantes de localidade:
  - √ Amostra retirada do banco de dados da previdência local
  - √ Fazem parte da amostra apenas moradores que atingiram a aposentadoria
  - √ Observações truncadas à esquerda

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo

- Estudos de tempo entre infecção e desenvolvimento de AIDS
  - √ Fazem parte da amostra somente pacientes que têm comprovada a soropositividade
  - √ Número de pacientes infectados é desconhecido
    - Indivíduos já infectados e que ainda não desenvolveram a doença são desconhecidos para o pesquisador e não são incluídos na amostra
  - √ Informações truncadas à direita

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Coortes Abertas

- Momento de entrada dos pacientes na coorte varia
  - √ Exemplo: Pacientes em hemodiálise
- Representação:
  - √  $(I_i, F_i, \delta_i)$ : informação sobrevivência indivíduo  $i$ 
    - $I_i$ : tempo calendário de início de observação
    - $F_i$ : tempo calendário de término de observação
  - √  $(I_5, F_5, \delta_5) = (10, 33, 1) \rightarrow (T_i, \delta_i) = (23, 1)$

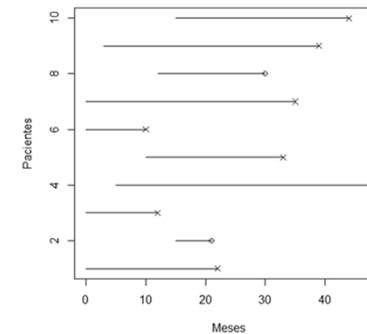
Análise de Sobrevivência Aplicada à Saúde -- 2014

- Exemplo: Pacientes em hemodiálise

Paciente (i)	Tempo inicial	Tempo final	Tempo ( $T_i$ )	Status ( $\delta_i$ )
1	0	22	22	1
2	15	21	6	0
3	0	12	12	1
4	5	48	43	0
5	10	33	23	1
6	0	10	10	1
7	0	35	35	1
8	12	30	18	0
9	3	39	36	1
10	15	44	29	1

Dados em formato de processo de contagem

Análise de Sobrevivência Aplicada à Saúde -- 2014



- Eixo das abcissas:
  - √ Tempo calendário
- Ponto inicial:
  - √ Data do 1º registro ou data de início de estudo prospectivo

Análise de Sobrevivência Aplicada à Saúde -- 2014

## Tempo de Sobrevida no R

```
library(survival)
Surv(tempo, status)
Surv(inicio, fim, status)
```

### • Comando Surv( )

√ Combina, em uma única variável as informações sobre o tempo de sobrevivência e o status de cada indivíduo

- Status = 1, se ocorreu o evento de interesse
- Status = 0, se o tempo foi censurado

Análise de Sobrevida Aplicada à Saúde -- 2014

### • Formato clássico

```
> library(survival)
> ipec = read.csv2(file = "dados/ipec.csv",header = T)
> ipec[1:8, c("id", "tempo", "status")]
  id tempo status
1  1   852     1
2  2   123     1
3  3  1145     1
4  4  2755     0
5  5  2117     0
6  6   329     0
7  7    60     1
8  8   151     1

> Surv(ipec$tempo, ipec$status)
[1] 852 123 1145 2755+ 2117+ 329+ 60 151 1563 1247
```

√ + ao lado do tempo de observação indica censura

Análise de Sobrevida Aplicada à Saúde -- 2014

### • Formato de processo de contagem

```
> ipec[1:8, c("id", "ini", "fim", "tempo", "status")]
  id ini  fim tempo status
1  1 1243 2095   852     1
2  2 2800 2923   123     1
3  3 1250 2395  1145     1
4  4 1915 4670  2755     0
5  5 2653 4770  2117     0
6  6    3   332   329     0
7  7   36    96    60     1
8  8    1   152   151     1

> Surv(ipec$ini, ipec$fim, ipec$status)
[1] (1243,2095] (2800,2923] (1250,2395] (1915,4670+]
[5] (2653,4770+] ( 3, 332+] ( 36, 96] ( 1, 152]
```

- √ + ao lado da data de saída da coorte indica censura
- √ Tempo início da observação não faz parte do intervalo
- √ Último tempo é parte do intervalo

Análise de Sobrevida Aplicada à Saúde -- 2014

## Funções Básicas de Sobrevida

### Introdução

- Exemplo: Sobrevivência ao diagnóstico de Aids
  - √ Coorte de 50 pacientes acompanhados por 4 anos (1460 dias)
  - √ Observados 32 óbitos
  - √ Objetivo: descrição do processo
    - Taxa de mortalidade média:  $32/50 = 64\%$
    - Taxa de mortalidade anual:  $(0,64)(100)/4 = 16$  óbitos por 100 pessoas-ano
  - √ Taxa não é homogênea no tempo
    - A probabilidade de um paciente vir a morrer de Aids é função que depende do tempo

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Questões em Análise de Sobrevivência

- √ Qual o **risco** de um paciente com Aids vir a falecer em até **três anos** após o diagnóstico?
- √ Qual a **probabilidade** de um paciente sobreviver por mais de 2 anos após o diagnóstico?
- √ Qual o **número esperado** de óbitos em uma coorte de pacientes acompanhados por 5 anos?
- √ Qual o **tempo mediano** de sobrevivência?

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Função de Densidade de Probabilidade

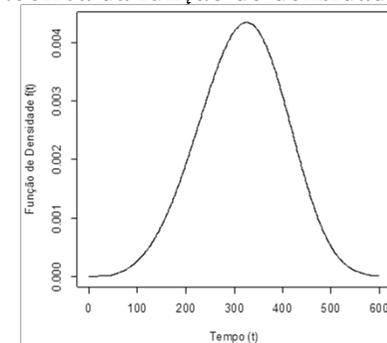
- $T$ : tempo de sobrevivência (tempo até desfecho)
  - √ Variável aleatória contínua e positiva
- $f$ : função de densidade de probabilidade de  $T$

$$f(t) = \lim_{\epsilon \rightarrow 0^+} \frac{P\{t \leq T \leq t + \epsilon\}}{\epsilon}$$

- √ Pode ser interpretada como a probabilidade de um indivíduo sofrer um evento em um intervalo instantâneo de tempo

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Curva teórica da função de densidade:



- √ Weibull(4, 350) → (forma, escala)
  - Esperança: 317,2
  - Mediana: 319,4

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo – Sem Censura

- Estudo do tempo entre diagnóstico de Aids e óbito

√ 193 pacientes acompanhados em ambulatório

√ Tempos de sobrevivência de 32 pacientes com Aids que morreram antes do fim do estudo (sem censura)

```
> aids.sem
[1] 3 18 29 54 60 84 110 112 116 123 134 145 151 151 158
[16] 173 194 214 329 331 371 408 490 514 541 555 688 780 801 858
[31] 887 998
```

$$\hat{f}_x(t) = \frac{N_x(t)}{\text{total ocorrências} \times \Delta_x}$$

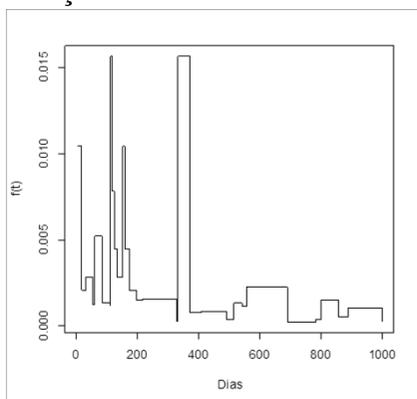
Análise de Sobrevivência Aplicada à Saúde -- 2014

Tempo (dias)	Intervalo	R <sub>x</sub> (t) (risco)	N <sub>x</sub> (t) (eventos)	D <sub>x</sub> (amplitude)	f <sub>x</sub> (t) (densidade)
3	(0, 3]	32	1	3	0,01042
18	(3, 18]	31	1	15	0,00208
29	(18, 29]	30	1	11	0,00284
54	(29, 54]	29	1	25	0,00125
60	(54, 60]	28	1	6	0,00521
84	(60, 84]	27	1	24	0,00130
110	(84, 110]	26	1	26	0,00120
112	(110, 112]	25	1	2	0,01563
116	(112, 116]	24	1	4	0,00781
123	(116, 123]	23	1	7	0,00446
134	(123, 134]	22	1	11	0,00284
145	(134, 145]	21	1	11	0,00284

$$\hat{f}_{(18,29]}(t) = \frac{N_{(18,29]}(t)}{(32)\Delta_x} = \frac{1}{(32)(11)} = 0,00284$$

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Função de densidade estimada



√ Suavização pode melhorar estimação

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Função de Sobrevivência

- Qual a probabilidade de um paciente sobreviver por mais de 365 dias?

√ Qual é o valor de P{T > 365}?

- S(t): função de sobrevivência

$$S(t) = P\{T > t\} = 1 - P\{T \leq t\} = 1 - F(t)$$

√ F(t): função de distribuição acumulado de T

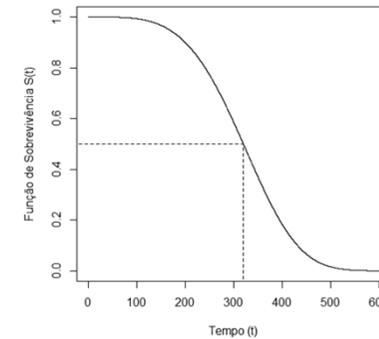
Análise de Sobrevivência Aplicada à Saúde -- 2014

### Propriedades

- $S(0) = 1$
- $S(+\infty) = 0$
- $S(t)$  é não crescente
- A partir da função de sobrevivência  $S(t)$  é possível calcular os percentis de  $T$ .

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Curva teórica da função de sobrevivência:



- √ Tempo de sobrevivência mediano: 319,4
- √  $S(319,4) = 0,5$

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Estimação Função Sobrevivência – Sem Censura

- Estimação de  $S(t)$  na ausência de censura:

$$\hat{S}_x(t) = \frac{R_x(t)}{n^{\circ} \text{ total de pacientes}}$$

√  $R_x(t)$ : número de pessoas sob risco no início do intervalo de tempo  $x$ .

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Exemplo – Sem Censura

- Estudo do tempo entre diagnóstico de Aids e óbito
- √ 193 pacientes acompanhados em ambulatório
- √ Tempos de sobrevivência de 32 pacientes com Aids que morreram antes do fim do estudo (sem censura)

```
> aids.sem
[1] 3 18 29 54 60 84 110 112 116 123 134 145 151 151 158
[16] 173 194 214 329 331 371 408 490 514 541 555 688 780 801 858
[31] 887 998
```

$$\hat{F}(18) = P\{T \leq 18\} = \frac{2}{32} = 0,06250$$

$$\hat{S}(18) = 1 - \hat{F}(18) = 0,93750$$

Análise de Sobrevivência Aplicada à Saúde -- 2014

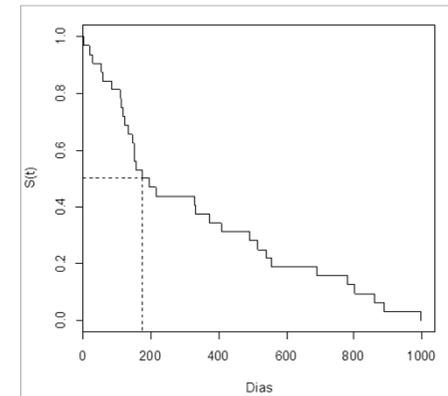
Tempo (dias)	Intervalo	$R_x(t)$ (risco)	$N_x(t)$ (eventos)	$S_x(t)$ (sobrevivência)
3	(0, 3]	32	1	1,00000
18	(3, 18]	31	1	0,96875
29	(18, 29]	30	1	0,93750
54	(29, 54]	29	1	0,90625
60	(54, 60]	28	1	0,87500
84	(60, 84]	27	1	0,84375
110	(84, 110]	26	1	0,81250
112	(110, 112]	25	1	0,78125
116	(112, 116]	24	1	0,75000
123	(116, 123]	23	1	0,71875
134	(123, 134]	22	1	0,68750
145	(134, 145]	21	1	0,65625

$\hat{S}_{(18,29]}(t) = \frac{R_{(18,29]}(t)}{32} = 0,93750$

P{sobreviver mais de 116 dias} = 0,71875  
 P{sobreviver mais de 145 dias} = 0,62500

Análise de Sobrevivência Aplicada à Saúde -- 2014

• Função de sobrevivência estimada



√ Tempo de sobrevivência mediano: 173 dias

Análise de Sobrevivência Aplicada à Saúde -- 2014

**Função de Risco**

- Qual o risco de um paciente com Aids vir a óbito após sobreviver 365 dias?  
 √ Esse risco de morrer aumenta ou diminui com o tempo?
- $\lambda(t)$ : risco instantâneo de um indivíduo sofrer o evento entre o tempo  $t$  e  $t + \epsilon$ , dado que ele sobreviveu até o tempo  $t$   
 √  $\epsilon$  é infinitamente pequeno  $\rightarrow \lambda(t)$  expressa risco instantâneo de ocorrência de uma evento, dado que até então o evento não tenha ocorrido

Análise de Sobrevivência Aplicada à Saúde -- 2014

$$\lambda(t) = \lim_{\epsilon \rightarrow 0^+} \frac{P\{t \leq T \leq t + \epsilon | T \geq t\}}{\epsilon}$$

√  $\lambda(t)$  também é denominada:

- função ou taxa de incidência,
- força de infecção,
- taxa de falha,
- força de mortalidade,
- força de mortalidade condicional.

√  $\lambda(t)$  é uma taxa (tempo<sup>-1</sup>).

- √ Pode assumir qualquer valor positivo
- (não é probabilidade)

Análise de Sobrevivência Aplicada à Saúde -- 2014

$$\lambda(t) = \frac{f(t)}{S(t)} \quad \lambda(t) = -\frac{d \ln(S(t))}{dt}$$

- $\lambda(t)$  e  $S(t)$  são inversamente proporcionais
  - √ Quando o risco aumenta, a probabilidade de sobrevivência diminui e vice-versa

Análise de Sobrevida Aplicada à Saúde -- 2014

### Comentário

- A função de risco (taxa de falha) é mais informativa que a função de sobrevivência
  - √ Diferentes funções de sobrevivência podem ter formas diferentes enquanto suas funções de risco podem diferir drasticamente
  - √ A modelagem da função de risco (taxa de falha) é um método importante para os dados de sobrevivência.

Análise de Sobrevida Aplicada à Saúde -- 2014

### Estimação Função de Risco – Sem Censura

- Estimação de  $\lambda(t)$  na ausência de censura:

$$\hat{\lambda}_x(t) = \frac{N_x(t)}{R_x(t)\Delta_x}$$

- √  $N_x(t)$ : quantidade de ocorrências no intervalo de tempo  $x$ .
- √  $R_x(t)$ : número de pessoas sob risco no início do intervalo de tempo  $x$ .
- √  $\Delta_x$ : amplitude do intervalo  $x$ .

Análise de Sobrevida Aplicada à Saúde -- 2014

### Observação

- Uma alternativa é utilizar as relações entre as funções  $S(t)$ ,  $f(t)$  e  $\lambda(t)$ .

Análise de Sobrevida Aplicada à Saúde -- 2014

• Estimativa da função de risco

Tempo (dias)	Intervalo	$R_x(t)$ (risco)	$N_x(t)$ (eventos)	$\Delta_x$ (amplitude)	$S_x(t)$ (sobrevivência)	$\lambda_x(t)$ (função risco)
3	(0, 3]	32	1	3	1,00000	0,01042
18	(3, 18]	31	1	15	0,96875	0,00215
29	(18, 29]	30	1	11	0,93750	0,00303
54	(29, 54]	29	1	25	0,90625	0,00138
60	(54, 60]	28	1	6	0,87500	0,00595
84	(60, 84]	27	1	24	0,84375	0,00154
110	(84, 110]	26	1	26	0,81250	0,00148
112	(110, 112]	25	1	2	0,78125	0,02000
116	(112, 116]	24	1	4	0,75000	0,01042
123	(116, 123]	23	1	7	0,71875	0,00621
134	(123, 134]	22	1	11	0,68750	0,00413
145	(134, 145]	21	1	11	0,65625	0,00433
151	(145, 151]	20	2	6	0,62500	0,01667

$$\hat{\lambda}_{(18,29]}(t) = \frac{N_{(18,29]}(t)}{R_{(18,29]}(t) \times \Delta_{(18,29]}} = \frac{1}{(30)(11)} = 0,00303$$

Análise de Sobrevida Aplicada à Saúde -- 2014

• Uma alternativa é utilizar as relações entre as funções  $S(t)$ ,  $f(t)$  e  $\lambda(t)$ .

• Exemplo:

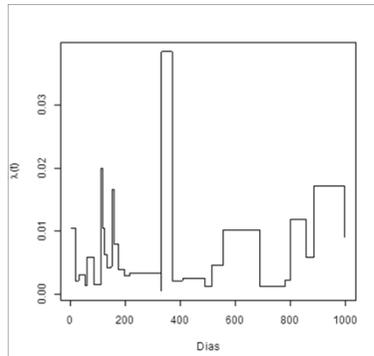
√ Risco no intervalo (18, 29]

– Estimativa por meio da relação  $f(t)/S(t)$ :

$$\hat{\lambda}_{(18,29]}(t) = \frac{f(18)}{S(18)}(t) = \frac{0,00284}{0,93750} = 0,00303$$

Análise de Sobrevida Aplicada à Saúde -- 2014

• Função de risco estimada



√ Função de risco estimada com dados esparsos flutua muito

– Não permite identificar padrões

Análise de Sobrevida Aplicada à Saúde -- 2014

**Comportamento da Função de Risco**

• Permite avaliar o risco de um indivíduo sofrer um evento em um determinado tempo  $t$ , dado que ele já sobreviveu até aquele momento

• Qual é o risco de um paciente falecer após uma cirurgia?

√ O risco de vir a óbito no 1º dia pós-cirúrgico é o mesmo de vir a óbito no 30º dia?

Análise de Sobrevida Aplicada à Saúde -- 2014

- Pode-se supor o risco constante no caso da Aids?

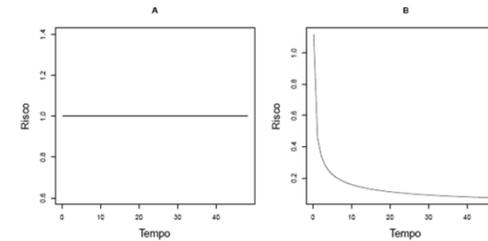
√ A suposição não é razoável pois ela é uma doença que vai comprometendo o sistema imunológico

- Antes da existência do retroviral, a hipótese de risco de óbito crescente a partir do diagnóstico parecia razoável. E agora?

√ Alternativa mais complexa:

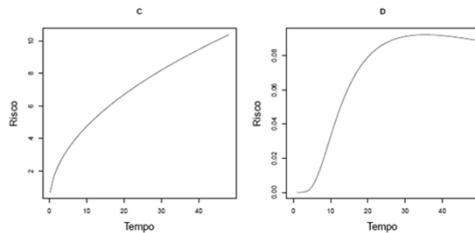
- Risco seria alto logo após diagnóstico (presença de doenças oportunistas). Nos 1<sup>os</sup> anos o risco decresceria (tratamento), voltando a aumentar devido à diminuição do efeito do tratamento

Análise de Sobrevivência Aplicada à Saúde -- 2014



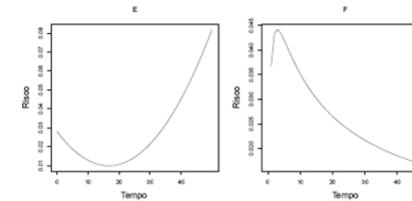
- Curva de risco constante
  - √ Ex: Ocorrência de fratura em população escolar em 12 meses  
(caso fatores de exposição a risco forem os mesmos)
- Curva decrescente de risco
  - √ Ex.: Óbito pós-cirúrgico  
- (risco é maior no primeiros dias, caindo rapidamente a zero com o passar do tempo)
- A ocorrência do evento é **sem memória**

Análise de Sobrevivência Aplicada à Saúde -- 2014



- Curva de risco crescente
  - √ Ex: Estudo sobre o tempo até solidificação de fratura óssea  
- Desfecho: solidificação de fratura  
- Risco nos primeiros dias é nulo até que praticamente todas as fraturas estejam solidificadas
- Exemplo de possível descrição de metástase
  - √ Período inicial: risco é baixo e não varia (quimioterapia)
  - √ Aumenta em seguida
  - √ Após um tempo fica aproximadamente constante até decair novamente

Análise de Sobrevivência Aplicada à Saúde -- 2014



- Comportamento do risco:
  - √ Risco começa alto
  - √ Decai durante certo tempo
  - √ Volta a subir
- Comportamento do risco?
  - √ Sobee rapidamente no início
  - √ Cai rapidamente depois
  - √ Exemplo: comportamento possível da tuberculose  
(após introdução dos tuberculostáticos, risco de óbito decai rapidamente)

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Função de Risco Acumulado

- Qual é o risco de um paciente com Aids vir a óbito no 1º ano após diagnóstico? E nos primeiros dois anos?

$$\Lambda(t) = \int_0^t \lambda(u) du$$

$\sqrt{\Lambda(t)}$ : função de risco acumulado

- mede o risco de ocorrência do evento até o tempo t
- é uma taxa e não está restrito ao intervalo [0,1]

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Estimação Função de Risco Acumulado – Sem Censura

- Estimação de  $\Lambda(t)$  na ausência de censura:

$$\hat{\Lambda}_x(t) = \sum_{k=1}^{x-1} \hat{\lambda}_k(t) \Delta_k, \text{ com } \hat{\Lambda}_1 = 0$$

$\sqrt{\lambda_x(t)}$ : função de risco no intervalo de tempo x.

$\sqrt{\Delta_x}$ : amplitude do intervalo x.

Análise de Sobrevivência Aplicada à Saúde -- 2014

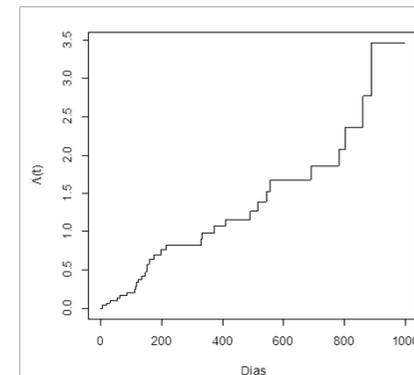
- Estimativa da função de risco acumulado

Tempo (dias)	Intervalo	$R_x(t)$ (risco)	$N_x(t)$ (eventos)	$\Delta_x$ (amplitude)	$S_x(t)$ (sobrevivência)	$\lambda_x(t)$ (função risco)	$\Lambda_x(t)$ (risco acumulado)
3	(0, 3]	32	1	3	1,00000	0,01042	0
18	(3, 18]	31	1	15	0,96875	0,00215	0,03125
29	(18, 29]	30	1	11	0,93750	0,00303	0,06351
54	(29, 54]	29	1	25	0,90625	0,00138	0,09684
60	(54, 60]	28	1	6	0,87500	0,00595	0,13132
84	(60, 84]	27	1	24	0,84375	0,00154	0,16704
110	(84, 110]	26	1	26	0,81250	0,00148	0,20408
112	(110, 112]	25	1	2	0,78125	0,02000	0,24254
116	(112, 116]	24	1	4	0,75000	0,01042	0,28254
123	(116, 123]	23	1	7	0,71875	0,00621	0,32420
134	(123, 134]	22	1	11	0,68750	0,00413	0,36768
145	(134, 145]	21	1	11	0,65625	0,00433	0,41314
151	(145, 151]	20	2	6	0,62500	0,01667	0,46076
158	(151, 158]	18	1	7	0,56250	0,00794	0,56076
173	(158, 173]	17	1	15	0,53125	0,00392	0,61631

$$\hat{\Lambda}_{(29)}(t) = 0,00215 \times 15 + 0,03125 = 0,063508$$

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Função de risco acumulado estimada:



Análise de Sobrevivência Aplicada à Saúde -- 2014

### Relação entre Funções de Sobrevivência

- As funções  $f(t)$ ,  $S(t)$ ,  $\lambda(t)$ , e  $\Lambda(t)$  são equivalentes
  - √ Diferentes formas de representar o mesmo fenômeno
- Respondem a três perguntas básicas:
  - √ Qual a probabilidade de sobreviver por mais de  $t$  unidades de tempo?
  - √ Qual o risco de sofrer o evento no tempo  $t$  se sabemos que o paciente sobreviveu até aquele momento?
  - √ Qual o risco de sofrer o evento até um determinado tempo  $t$ ?

Análise de Sobrevivência Aplicada à Saúde -- 2014

- Relações entre as funções básicas: de sobrevivência:

$$\begin{aligned}S(t) &= 1 - F(t) \\S(t) &= \exp(-\Lambda(t)) \\ \lambda(t) &= -\frac{d \ln(S(t))}{dt} \\ \lambda(t) &= \frac{f(t)}{S(t)} \\ \lambda(t) &= \frac{f(t)}{1 - F(t)} \\ \Lambda(t) &= -\ln(S(t))\end{aligned}$$

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Referências

### Bibliografia

- Carvalho, M. S. et al. *Análise de Sobrevivência: Teoria e Aplicações em Saúde*. (Fiocruz)
- Colosimo, E. A. e Giolo, S. R. *Análise de Sobrevivência Aplicada*. (Edgard Blucher)
- Klein, J. P. e Moeschberger, M. L. *Survival Analysis: Techniques for Censored and Truncated Data*. (Springer)
- Kleinbaum, D. G. *Survival Analysis: a Self-Learning Text*

Análise de Sobrevivência Aplicada à Saúde -- 2014

### Material de Apoio

- R:  
√ [www.r-project.org](http://www.r-project.org)
- Tutorial online do R:  
√ <http://www.leg.ufpr.br/Rtutorial>  
√ <http://www.leg.ufpr.br/~paulojus/embrapa/Rembrapa>
- Conjuntos de dados e material Análise de Sobrevivência – Carvalho et al.  
√ <http://sobrevida.fiocruz.br>

Análise de Sobrevivência Aplicada à Saúde -- 2014