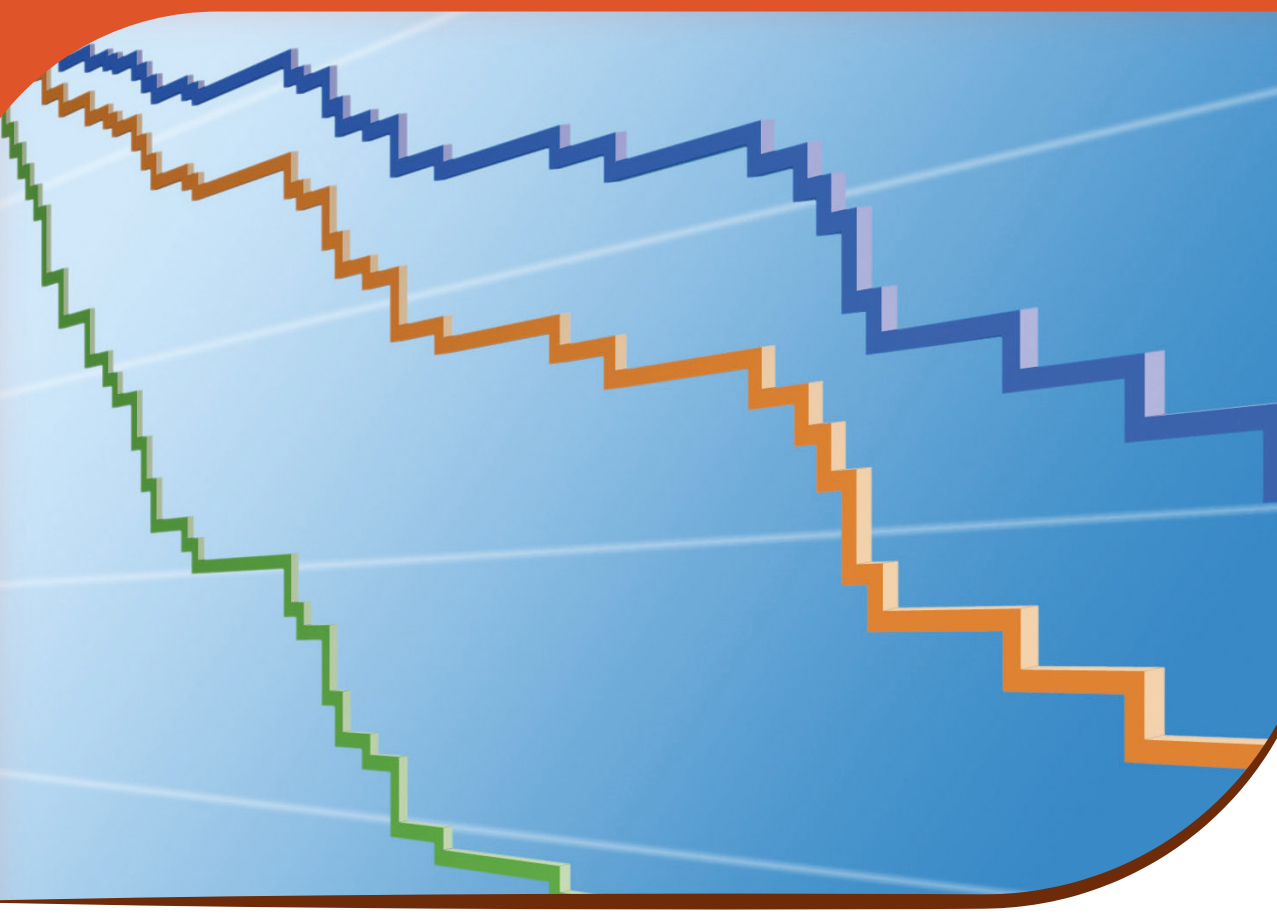
ENRICO ANTÔNIO COLOSIMO
SUELY RUIZ GIOLO

ANÁLISE DE SOBREVIVÊNCIA APLICADA



Blucher

ABE - PROJETO FISHER

2ª edição

Análise de sobrevivência aplicada

2ª edição

Enrico Antônio Colosimo

Departamento de Estatística
Universidade Federal de Minas Gerais

Suely Ruiz Giolo

Departamento de Estatística
Universidade Federal do Paraná

Análise de sobrevivência aplicada

2ª edição – 2024
Editora Edgard Blücher Ltda.

Publisher Edgard Blücher
Editores Eduardo Blücher e Jonatas Eliakim
Coordenação editorial Andressa Lira
Produção editorial Thaís Pereira
Revisão de texto Renata Truyts
Adaptação de capa Laércio Flenic
Imagem da capa Suely Ruiz Giolo

Blucher

Rua Pedroso Alvarenga, 1245, 4º andar
04531-934 - São Paulo - SP - Brasil
Tel.: 55 11 3078-5366
contato@blucher.com.br
www.blucher.com.br

Segundo o Novo Acordo Ortográfico, conforme 6. ed. do *Vocabulário Ortográfico da Língua Portuguesa*, Academia Brasileira de Letras, julho de 2021.

Dados Internacionais de Catalogação na Publicação (CIP)
Angélica Ilacqua CRB–8/7057

Colosimo, Enrico Antônio
Análise de sobrevivência aplicada / Enrico Antônio Colosimo, Suely Ruiz Giolo. – 2. ed. – São Paulo : Blucher, 2024.
380 p.

Bibliografia
ISBN 978-85-212-2199-9

1. Análise de sobrevivência (Biometria) 2. Pesquisa médica – Métodos estatísticos I. Título II. Giolo, Suely Ruiz

23-4859 CDD 610.72

Índices para catálogo sistemático:
1. Análise de sobrevivência (Biometria)

Conteúdo

Capítulo 1

Conceitos básicos e exemplos

1.1 Introdução

A análise de sobrevivência é uma das áreas da estatística que mais cresceu nos últimos tempos, o que se deve ao desenvolvimento e aprimoramento de técnicas estatísticas combinados com computadores cada vez mais velozes. Uma evidência quantitativa deste crescimento é o número de aplicações de métodos de análise de sobrevivência em medicina. Segundo Bailar III e Mosteller (1992), o uso desses métodos cresceu de 11% em 1979 para 32% em 1989 nos artigos do conceituado periódico *The New England Journal of Medicine*. De acordo com os autores, esta foi a área da estatística que mais se destacou no período avaliado. Stigler (1994) também constatou que o artigo do estimador de Kaplan-Meier (KAPLAN; MEIER, 1958) e o do modelo de Cox (COX, 1972) foram os dois artigos mais citados na literatura estatística no período de 1987 a 1989. Em 2014, estes artigos foram classificados nas posições 11ª e 24ª entre os 100 mais citados de todos os tempos, com 38.600 e 28.439 citações, respectivamente (VAN NOORDEN et al., 2014). Em 2017, com 44.319 citações na Web of Science®, o artigo de Kaplan e Meier tornou-se a publicação de estatística mais citada na literatura científica. Um grande impulso para a utilização crescente do método de Kaplan-Meier foi, sem dúvida, o próprio artigo de Cox (1972) em que ele deixou clara a importância deste método (STALPERS; KAPLAN, 2018).

Em análise de sobrevivência, a variável resposta é usualmente o tempo medido desde uma data base até a ocorrência de um evento de interesse, denominado *tempo até o evento* ou *tempo de sobrevida*. Em estudos de câncer, pode ser o tempo desde o diagnóstico da doença até a morte do paciente ou até a remissão (quando não há sinais de câncer nos exames clínicos, laboratoriais e de imagem), ou, então, o tempo desde a remissão do câncer até a recidiva, conhecida como recaída, recorrência ou retorno da doença.

A principal característica dos dados de sobrevivência é a presença de *censuras*, definidas como observações parciais da resposta. Isso se refere às situações em que o seguimento do paciente foi interrompido porque, dentre outros motivos, ele mudou de cidade, ou foi a óbito por alguma outra razão não associada àquela em análise, ou o estudo terminou para a análise dos dados. Portanto, a informação referente à resposta se resume ao conhecimento de que o tempo até o evento é superior ao observado.

Na ausência de censuras, as técnicas estatísticas clássicas, como análise de regressão e análise de variância, podem eventualmente ser consideradas para a análise de dados de sobrevivência, provavelmente utilizando uma transformação para a resposta. No entanto, na presença de censuras, tais técnicas tornam-se inviáveis. Assim, são necessários métodos estatísticos para a análise de dados de sobrevivência que possibilitem incorporar na análise a informação contida tanto nas observações completas quanto nas parciais (censuras).

O termo análise de sobrevivência refere-se basicamente às situações na área da saúde que envolvem dados censurados. Todavia, condições similares ocorrem em outras áreas que usam as mesmas técnicas de análise de dados. Em engenharia, por exemplo, são comuns os estudos em que componentes são colocados sob teste com a finalidade de se estimar características relacionadas aos seus tempos de vida, tais como o tempo médio ou a probabilidade de certo componente durar mais de cinco anos. Exemplos podem ser encontrados em Nelson (1990), Freitas e Colosimo (1997) e Meeker e Escobar (1998). Os engenheiros denominam esta área confiabilidade. O mesmo ocorre em Ciências Sociais, em que, para várias situações, a resposta de

interesse é o tempo entre eventos (ELANDT-JONHSON; JONHSON, 1980, ALLISON, 1984). Os cientistas sociais denominam esta área análise de história de eventos (BOX-STEFFENSMEIER; JONES, 2004). Ainda, criminalistas estudam o tempo entre a liberação de presos e a ocorrência de crimes; estudiosos do trabalho se concentram nos tempos relativos às mudanças de empregos, desempregos, promoções e aposentadorias; e os demógrafos nos tempos associados a nascimentos, mortes, casamentos, divórcios e migrações. O crescimento observado no número de aplicações na área da saúde também pode ser observado nestas outras áreas.

Este texto foi motivado por ilustrações essencialmente da área da saúde, de modo que os exemplos e colocações são, em geral, conduzidos para esta área. Contudo, as técnicas estatísticas apresentadas são de ampla utilização em várias outras áreas do conhecimento, como enfatizado anteriormente.

Este capítulo é dedicado à apresentação de conceitos básicos e funções importantes para a análise de dados de sobrevivência. Na Seção 1.2, são apresentados os objetivos e o planejamento de alguns estudos clínicos e industriais. A Seção 1.3 trata a caracterização dos dados de sobrevivência. Exemplos que envolvem dados de sobrevivência são descritos na Seção 1.4. A Seção 1.5 finaliza o capítulo mostrando as principais funções utilizadas para a análise de dados de sobrevivência, bem como algumas relações probabilísticas importantes entre elas.

1.2 Objetivo e planejamento dos estudos

Os estudos clínicos e epidemiológicos são investigações científicas conduzidas com o objetivo de estimar quantidades desconhecidas ou realizar predições. Tais investigações são realizadas coletando dados e analisando-os por meio de métodos estatísticos. Em geral, esses estudos apresentam três etapas, que são comuns a qualquer situação envolvendo a análise estatística de dados. São elas:

1. Formulação da pergunta de interesse (*research question*).
2. Planejamento e coleta dos dados.
3. Análise estatística dos dados para responder à pergunta de interesse.

A primeira etapa de um estudo é gerada pela curiosidade científica do pesquisador. Identificar fatores de risco para uma doença é o objetivo que aparece com mais frequência em estudos observacionais. Por outro lado, a comparação de drogas ou diferentes opções terapêuticas caracteriza-se como o principal objetivo dos ensaios clínicos controlados.

Os textos estatísticos concentram todo o esforço na terceira etapa, ou seja, na análise estatística dos dados, mesmo admitindo a importância de um adequado planejamento do estudo. Este texto não é diferente dos demais, mas traz a seguir uma breve descrição da segunda etapa.

Em análise de sobrevivência, a resposta é por natureza longitudinal. O delineamento de estudos com respostas dessa natureza pode ser observacional ou experimental, assim como retrospectivo ou prospectivo. As quatro formas básicas de estudos epidemiológicos são: descritivo, caso-controle, coorte e ensaio clínico aleatorizado. Os três primeiros são observacionais, e o último é experimental devido à intervenção do pesquisador em alocar tratamento ao paciente de forma aleatória. O uso das técnicas de análise de sobrevivência é mais frequente nos estudos de coorte e ensaios clínicos, mas o seu uso também é possível nos demais estudos, desde que os tempos até o evento de interesse possam ser claramente definidos e mensurados.

Os estudos envolvendo somente uma amostra, usualmente de doentes, são descritivos, pois não há grupos de comparação. Nestes estudos, o objetivo é usualmente a identificação de fatores prognósticos para a doença em estudo. Os outros tipos de estudo são comparativos, o que significa que o objetivo deles é a comparação entre dois ou mais grupos.

Os estudos caso-controle são retrospectivos. Neles, dois grupos de indivíduos, um de doentes (casos) e outro de não doentes (controles), são comparados em relação à exposição a um ou mais fatores de interesse. Estes estudos são simples, de baixo custo e rápidos, pois a informação já se encontra disponível. Contudo, eles têm algumas limitações por estarem sujeitos a alguns tipos de vieses relacionados à informação disponível sobre a história da exposição, bem como à incerteza sobre a escolha do grupo controle. Uma discussão mais ampla sobre esses estudos pode ser encontrada,

dentre outros, em Breslow e Day (1980) e Rothman et al. (2012).

As limitações dos estudos caso-controle podem ser superadas pelos estudos de coorte. Neles, dois grupos de indivíduos, um exposto e outro não exposto ao fator de interesse, são acompanhados por um período de tempo registrando-se a ocorrência da doença ou do evento de interesse. As vantagens dos estudos de coorte são: poder avaliar a comparabilidade dos grupos no início do estudo e identificar as variáveis de interesse a serem medidas. Por outro lado, é um estudo longo que, em geral, envolve custos elevados, pois os indivíduos são acompanhados por um período de tempo muitas vezes superior a alguns anos. Além disso, eles não são indicados para doenças raras. Uma importante referência é Breslow e Day (1987).

A forma mais consagrada de pesquisa clínica é o ensaio clínico aleatorizado, dito experimental devido à intervenção direta do pesquisador ao alocar tratamento ao paciente de forma aleatória. O objetivo da alocação aleatória de tratamentos é garantir a comparabilidade dos grupos. Ensaios clínicos duplo-cegos são realizados sempre que possível a fim de evitar vieses por parte dos envolvidos no estudo. Para mais detalhes podem ser consultados Pocock (1983), Friedman et al. (1998) e Giolo (2017).

Quanto aos estudos industriais, eles são usualmente de campo ou realizados na própria empresa simulando situações de campo. Entretanto, há também os estudos denominados testes de vida acelerados. Neles, os itens amostrais são estressados para falhar mais rápido com o objetivo de reduzir o tempo de coleta dos dados. As estimativas das quantidades de interesse nas condições de uso são obtidas, nesses estudos, por extrapolações. Mais detalhes são encontrados em Nelson (1990) e Freitas e Colosimo (1997).

Os testes de degradação, que podem ou não ser acelerados, são extensões dos testes de vida acelerados que vêm ganhando espaço na literatura de engenharia. Neles, uma variável numérica associada ao tempo até a falha é registrada ao longo do período de acompanhamento. Com base nos valores desta variável, podem-se obter as estimativas de interesse mesmo quando nenhuma falha tenha sido registrada. Detalhes são encontrados em Meeker e Escobar (1998), Oliveira e Colosimo (2004) e Freitas et al. (2009).

1.3 Caracterização dos dados de sobrevivência

Os dados de sobrevivência são caracterizados pelos tempos até o evento ou tempos de sobrevida e, frequentemente, pelas censuras. Em geral, covariáveis também são registradas para os indivíduos. Os três elementos que caracterizam o tempo de sobrevida são: o tempo inicial, a escala de medida e o evento de interesse. Esses elementos e a censura são tratados a seguir.

1.3.1 Tempo até o evento ou tempo de sobrevida

O tempo de início do estudo deve ser sempre precisamente estabelecido. Nele, os indivíduos devem ser comparáveis, com exceção de diferenças medidas pelas covariáveis. Em um ensaio clínico aleatorizado, a data da aleatorização é a escolha natural para o tempo inicial do estudo. Contudo, outras escolhas são possíveis, tais como a data de início do tratamento da doença ou a data do diagnóstico da doença.

A escala de medida é quase sempre o tempo de relógio (horas, dias, semanas, meses etc.), apesar de existirem alternativas. Em testes de engenharia, por exemplo, escalas usuais de medida são: o número de ciclos, a quilometragem de veículos, ou qualquer outra medida de carga.

O terceiro elemento é o evento de interesse que, na maioria das vezes, consiste em eventos indesejáveis e, usualmente, chamados de falha. Nos estudos de sobrevivência, é importante definir qual é o evento de forma clara e precisa. Para algumas situações, a definição do evento é clara (são exemplos, morte e recidiva), mas para outras pode assumir termos ambíguos. Por exemplo, fabricantes de produtos alimentícios têm interesse no tempo desde a chegada do produto ao supermercado até ele ficar inapropriado ao consumo. Neste caso, o evento *inapropriado ao consumo* precisa ser claramente definido antes do início do estudo. Pode ser, por exemplo, quando a área do produto atingir certa concentração de microorganismos por mm^2.

Quanto ao tempo registrado em análise de sobrevivência, ele pode ser de natureza cronológica, idade ou de duração. O tempo cronológico é usual em estudos de coorte em que todos os indivíduos iniciam o estudo na mesma

data. O objetivo desses estudos consiste, em geral, em investigar o desenvolvimento e a evolução de patologias devido ao envelhecimento de populações. A coorte do estudo de Framingham, iniciado em 1948, se destaca como a mais importante envolvendo doenças cardiovasculares (MAHMOOD et al., 2014). No Brasil, a coorte do Estudo Longitudinal de Saúde de Adultos (ELSA) vem acompanhando cerca de 15 mil funcionários de seis instituições públicas desde 2008 com o objetivo de investigar a incidência de diabetes e doenças cardiovasculares (AQUINO et al., 2012).

Já o tempo cuja natureza é a idade surge para caracterizar padrões particulares ao longo da vida de seres humanos, cobaias ou plantas. A idade de crianças ao desmame e o tempo de vida de mangueiras, mencionados em dois dos exemplos da Seção 1.4, ilustram esta situação.

Por fim, o tempo cuja natureza é a duração caracteriza-se como o exemplo mais usual em análise de sobrevivência. Neste caso, a data de entrada no estudo não é a mesma para todos os indivíduos. Na área da saúde, por exemplo, é usual que os pacientes não sejam recrutados na mesma data, mas sim que a amostra seja formada ao longo de um período de tempo, que pode durar meses ou até mesmo anos. Isso significa que o tempo inicial de estudo dos pacientes ocorre em datas cronológicas diferentes, como ilustra a Figura 1.1. Após a inclusão no estudo, cada paciente é acompanhado até a ocorrência do evento de interesse ou então até a perda de seguimento ou final do estudo (censura), quando os dados são analisados.

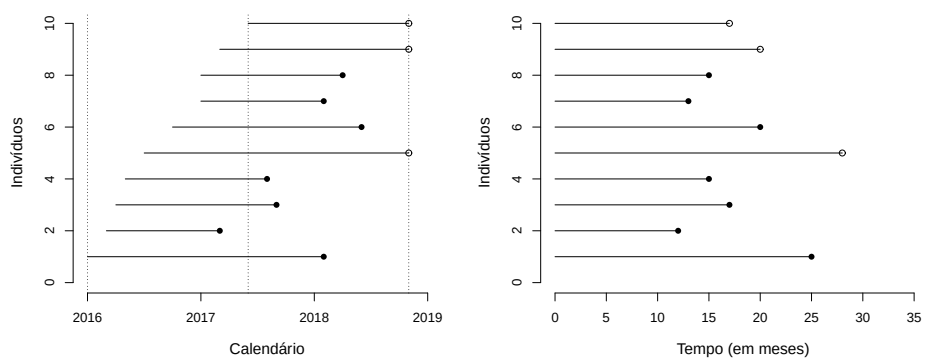


Figura 1.1 – Data cronológica de entrada no estudo (à esquerda) e tempo de seguimento (à direita) medido a partir da entrada até o evento • ou até a censura ∘.

A representação à esquerda na Figura 1.1 ilustra as datas cronológicas de recrutamento de dez indivíduos. Nela, é possível observar que o primeiro recrutamento ocorreu em 01/01/2016 e o último em 31/05/2017, bem como que a finalização do estudo se deu em 30/10/2019. A representação à direita mostra os tempos de seguimento dos indivíduos medidos a partir do tempo inicial, que corresponde à data cronológica de entrada no estudo (referencial $t = 0$ na representação à direita), até a ocorrência do evento ou da censura.

1.3.2 Censura e dados truncados

Os estudos que envolvem uma resposta temporal são frequentemente prospectivos e de longa duração. Mesmo sendo longos, os estudos de sobrevivência geralmente terminam antes de todos os indivíduos terem apresentado o evento de interesse. Assim, uma característica desses estudos é a presença de observações incompletas ou parciais, denominadas censuras, que podem ocorrer por várias razões, dentre elas, a necessidade de finalizar o estudo, a perda de seguimento do paciente e a ocorrência de outro evento que impede observar aquele de interesse. Para os indivíduos cujo evento não aconteceu, tudo o que se sabe é que o tempo até a ocorrência do evento de interesse é superior ao tempo até a última data de contato.

Ressalta-se o fato de que todas as observações provenientes de um estudo de sobrevivência (completas e parciais) devem ser utilizadas na análise estatística. Duas razões justificam tal procedimento: (i) mesmo sendo incompletas, as censuras fornecem informações relevantes sobre o tempo de sobrevida dos pacientes; e (ii) a omissão das censuras no cálculo das estatísticas de interesse pode acarretar conclusões viesadas.

Os estudos de sobrevivência podem apresentar diferentes mecanismos de censura. Censuras tipo I ocorrem nos estudos que, ao serem finalizados após um período preestabelecido, registram em seu término alguns indivíduos que ainda não apresentaram o evento de interesse. Por outro lado, censuras tipo II resultam de estudos que são finalizados após a ocorrência do evento de interesse para um número preestabelecido de indivíduos. Um terceiro mecanismo, o de censuras aleatórias, é o mais comum na área da saúde.

Dentre outros motivos, ele ocorre quando há a necessidade de remover um paciente do estudo sem ter ocorrido o evento, ou quando o paciente morre por algum motivo não associado àquele em análise. A Figura 1.2 ilustra os três mecanismos de censura mencionados.

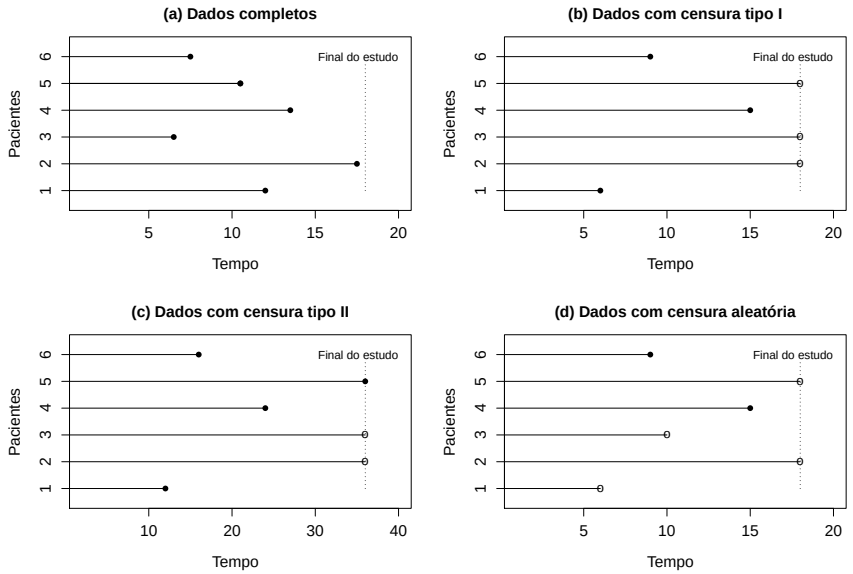


Figura 1.2 – Mecanismos de censura com • falha e ∘ censura. (a) todos experimentam o evento antes do final do estudo, (b) alguns não experimentam o evento até o término do estudo, (c) o estudo é finalizado após um número prefixado de falhas, e (d) alguns não experimentam o evento por razões diversas.

A representação probabilística do mecanismo de censuras aleatória pode ser feita considerando duas variáveis aleatórias independentes T e C, em que T é o tempo até o evento e C é o tempo até a censura. Para cada indivíduo, observa-se $t = \min(T, C)$ e δ tal que

$$\delta = \begin{cases} 1 & \text{se } T \leq C \\ 0 & \text{se } T > C. \end{cases}$$

Quando C é uma constante fixa sob o controle do pesquisador, tem-se a censura tipo I. Ou seja, a censura tipo I é um caso particular da aleatória. Neste caso, a variável aleatória t tem probabilidade maior que zero em $t = C$, o que significa, no caso de censura tipo I, que t é uma variável aleatória mista com um componente contínuo e outro discreto.

Os mecanismos de censura descritos e representados na Figura 1.2 são conhecidos como censura à direita devido ao tempo até a ocorrência do evento de interesse estar à direita do tempo registrado. Esta situação é a mais frequente em estudos envolvendo dados de sobrevivência. Contudo, outras duas formas de censura podem ocorrer: censura à esquerda e intervalar.

A censura à esquerda ocorre quando o tempo registrado é maior do que o tempo até o evento, o que indica que o evento de interesse já havia ocorrido quando o indivíduo foi observado. Para ilustrar esta situação, pode ser citado um estudo realizado em uma comunidade para determinar a idade em que as crianças aprendem a ler. Quando o estudo foi iniciado, algumas crianças já sabiam ler, mas não se lembravam a idade exata em que isto havia ocorrido, caracterizando a presença de censuras à esquerda.

Para esse mesmo estudo, observa-se a possibilidade de se ter simultaneamente a presença de censuras à direita, visto que algumas crianças podem não aprender a ler até o término do estudo. Na presença simultânea de dois tipos de censura em um mesmo estudo, diz-se que os dados de sobrevivência são duplamente censurados (TURNBULL, 1974).

A censura intervalar é um tipo mais geral de censura que ocorre, por exemplo, em estudos nos quais os pacientes são monitorados em visitas periódicas, sendo conhecido somente que o evento de interesse ocorreu em um intervalo de tempo. Devido ao tempo exato até o evento T não ser conhecido, mas somente o intervalo a que ele pertence, isto é, $T \in (L, U]$, tais dados são denominados *dados de sobrevivência intervalar* ou, ainda, *dados de censura intervalar*. Lindsey e Ryan (1998) observam que tempos exatos de sobrevida, bem como censuras à direita e à esquerda, são casos especiais de dados de sobrevivência intervalar, com $L = U$ para os tempos exatos, $U = \infty$ para censuras à direita e $L = 0$ para censuras à esquerda.

Outra característica presente em alguns estudos de sobrevivência é o truncamento, muitas vezes confundido com censura. De modo geral, o truncamento é caracterizado por uma condição que exclui certos indivíduos do estudo. Por exemplo, se para estimar a distribuição do tempo de vida dos moradores de uma certa localidade for extraída uma amostra do banco

de dados da previdência local, somente farão parte da amostra os moradores aposentados vivos. Ou seja, os que faleceram antes da amostra ser extraída são excluídos do estudo, e a amostra é dita truncada à esquerda. Por outro lado, em estudos sobre a Aids, a data da infecção é bastante utilizada como tempo inicial, com o desenvolvimento da Aids sendo usualmente o evento de interesse. Neste caso, o número de pacientes infectados é desconhecido. Então, indivíduos já infectados e que ainda não desenvolveram a doença são desconhecidos para o pesquisador, não sendo incluídos na amostra. Somente fazem parte da amostra os pacientes com comprovação da doença, e a amostra é dita truncada à direita. Outros exemplos podem ser encontrados em Nelson (1990a), Kalbfleisch e Lawless (1992) e Klein e Moeschberger (2003).

A presença de censuras traz dificuldades para a análise estatística. Para a censura tipo II, há métodos exatos de inferência estatística para situações simples, que ocorrem raramente em estudos de sobrevivência (LAWLESS, 2003). Na prática, resultados assintóticos são utilizados para a análise de dados de sobrevivência. Como esses resultados não exigem o reconhecimento do mecanismo de censura, as mesmas técnicas estatísticas são utilizadas na análise de dados oriundos dos três mecanismos de censura à direita.

Neste texto, o foco se concentra nos dados de sobrevivência com censuras à direita, muito frequentes em medicina, engenharia e ciências sociais. Assim, quando for simplesmente mencionado *censura*, entenda-se censura à direita. Para o tratamento de dados censurados e truncados, recomenda-se Turnbull (1976) e Klein e Moeschberger (2003). Quanto aos dados de sobrevivência intervalar, algumas técnicas são abordadas no Capítulo 8.

1.4 Exemplos de dados de sobrevivência

As técnicas de análise de sobrevivência são úteis em várias situações. Na área da saúde, por exemplo, elas são utilizadas para a identificação de fatores prognósticos de uma doença, bem como para a comparação de tratamentos. Em oncologia, qualquer nova terapêutica ou droga para o combate ao câncer requer a realização de estudos em que a resposta é, em geral, o

tempo de sobrevida dos pacientes, denominada pelos oncologistas de sobrevida global. Estudos epidemiológicos sobre a Aids também têm utilizado essas técnicas, como o publicado por Jacobson et al. (1993).

A seguir, são apresentados alguns exemplos utilizados no decorrer do texto para ilustrar as técnicas estatísticas descritas. Vários deles retratam situações provenientes de assessorias estatísticas realizadas pelos autores, enquanto outros foram extraídos da literatura.

1.4.1 Dados de hepatite viral aguda

Um ensaio clínico foi realizado para investigar o efeito de uma terapia com esteroide no tratamento de hepatite viral aguda. Os 29 pacientes com a doença foram aleatorizados para receber um placebo ou o tratamento com esteroide. O acompanhamento de cada paciente foi de no máximo 16 semanas. Os tempos até a morte ou perda de acompanhamento ou, ainda, até o final do estudo estão na Tabela 1.1, com + indicando censura.

Tabela 1.1 – Dados registrados em um ensaio clínico sobre hepatite viral

Grupo	Tempo de sobrevida em semanas
Controle	1^+ 2^+ 3 3 3^+ 5^+ 5^+ 16^+ 16^+ 16^+ 16^+ 16^+ 16^+ 16^+ 16^+
Esteroide	1 1 1 1^+ 4^+ 5 7 8 10 10^+ 12^+ 16^+ 16^+ 16^+

Fonte: Gregory et al. (1976).

1.4.2 Dados de camundongos infectados pela malária

Um estudo experimental foi realizado com 44 camundongos no Centro de Pesquisas René Rachou, Fiocruz-MG, com o objetivo de investigar a eficácia da imunização pela malária. Os camundongos foram aleatorizados em três grupos e infectados pela malária (*Plasmodium berghei*). No entanto, os camundongos do grupo 1 foram imunizados 30 dias antes da infecção. Além da infecção pela malária, os camundongos dos grupos 1 e 3 também foram infectados pela esquistossomose (*Schistosoma mansoni*). A duração do estudo após a infecção foi de 30 dias. Os tempos medidos desde a infecção até a morte estão na Tabela 1.2, com + indicando censura.

Tabela 1.2 – Dados registrados em camundongos infectados pela malária

Grupo	n	Tempo de sobrevida em dias
1	16	7 8 8 8 8 12 12 17 18 22 30^+ 30^+ 30^+ 30^+ 30^+ 30^+
2	15	8 8 9 10 10 14 15 15 18 19 21 22 22 23 25
3	13	8 8 8 8 8 8 9 10 10 10 11 17 19

1.4.3 Dados de leucemia pediátrica

A leucemia aguda é a neoplasia de maior incidência na população com idade inferior a 15 anos, com a maioria dos casos sendo de Leucemia Linfoblástica Aguda (LLA). O objetivo do tratamento de crianças com LLA é a obtenção de períodos longos de sobrevida livre da doença, o que, eventualmente, significa a cura. Assim, para identificar os fatores que afetam o tempo de sobrevida de crianças brasileiras com LLA, um estudo foi realizado pelo Grupo Cooperativo Mineiro para Tratamento de Leucemias Agudas. Neste estudo, 128 crianças com idade inferior a 15 anos, todas com LLA, foram acompanhadas de 1988 a 1992. A variável resposta foi o tempo, em anos, desde a remissão (ausência de sinais da doença) até a recidiva ou a morte, a que ocorreu primeiro. Das 128 crianças, 120 entraram em remissão e são elas que compõem o conjunto de dados. Os fatores registrados no estudo foram: idade, peso, estatura, contagem de leucócitos, porcentagem de linfoblastos, porcentagem de vacúolos, fator de risco e indicador de sucesso da remissão. Detalhes adicionais sobre o estudo podem ser encontrados em Colosimo et al. (1992) e Viana et al. (1994).

1.4.4 Dados de sinusite em pacientes infectados pelo HIV

Há vários trabalhos na literatura sobre a Aids, a maioria com foco na sobrevida de pacientes infectados pelo HIV. Outras investigações, contudo, são de interesse, tal como o estudo conduzido pela Profa. Denise Gonçalves, da UFMG, cujo objetivo consistia em investigar se a infecção pelo HIV aumenta o risco de ocorrência de sinusite. Nesse estudo, 112 pacientes foram acompanhados no período de março de 1993 a fevereiro de 1995, sendo 91 HIV positivo e 21 HIV negativo. A classificação quanto à infecção pelo HIV se-

guiu os critérios do *Center for Disease Control* (CDC, 1987), sendo ela: HIV soronegativo (não possui o HIV), HIV soropositivo assintomático (possui o vírus, mas não desenvolveu o quadro clínico de Aids), com ARC (*Aids Related Complex*: apresenta baixa imunidade e outros indicadores clínicos que antecedem o quadro clínico de Aids), ou com Aids (apresenta infecções oportunistas que definem Aids). Esta classificação foi reavaliada a cada consulta, sendo elas trimestrais e cuja frequência mediana ao longo do estudo foi 4. Esta é a principal variável a ser considerada no estudo, sendo caracterizada como dependente do tempo, pois os pacientes mudam de classificação ao longo do estudo. Esta característica requer o uso das técnicas tratadas no Capítulo 6. Variáveis como contagem de células CD4 e CD8 também são dependentes do tempo, mas elas foram registradas somente no início do estudo e ocorreu a falta de registro de ambas para 37% dos pacientes, o que inviabilizou seu uso nas análises. A variável resposta foi o tempo, medido em dias, desde a 1ª consulta até a ocorrência de sinusite, e o objetivo foi identificar fatores de risco para a sinusite. Mais detalhes podem ser encontrados em Gonçalves (1995) e Colosimo e Vieira (1996).

1.4.5 Dados de aleitamento materno

A Organização Mundial da Saúde (OMS) recomenda o leite materno como a única fonte de alimentação para crianças entre 4 e 6 meses de vida. Desse modo, identificar fatores associados ao aleitamento materno em diferentes populações é fundamental para alcançar esta recomendação.

Neste contexto, os Profs. Eugênio Goulart e Cláudia Lindgren, ambos da UFMG, realizaram um estudo no Centro de Saúde São Marcos de Belo Horizonte, MG, para conhecer a prática de aleitamento materno das mães que utilizam este centro, bem como os possíveis fatores de risco ou de proteção para o desmame precoce. Para tanto, aplicaram um inquérito epidemiológico com questões demográficas e comportamentais a 150 mães de crianças com idade inferior a 2 anos. A variável resposta foi o tempo, em meses, desde o nascimento até o desmame completo da criança, definido como sendo o primeiro dia em que a criança não mais se alimenta de leite materno.

1.4.6 Dados de câncer de mama

Um estudo foi realizado com 94 mulheres com diagnóstico precoce de câncer de mama com o objetivo de pesquisar duas terapias: (*i*) somente radioterapia e (*ii*) radioterapia em conjunto com quimioterapia. Um total de 46 delas recebeu a primeira terapia, e as demais a segunda. As pacientes foram monitoradas a cada 4 a 6 meses, registrando-se, em cada visita, a ocorrência de retração da mama (nenhuma, moderada ou severa) e o tempo (em meses) até o aparecimento de retração moderada ou severa da mama.

Como as visitas foram realizadas com espaçamentos de 4 a 6 meses, não se sabe com exatidão quando a primeira retração da mama ocorreu; sabe-se somente que ela se deu entre duas das visitas realizadas. Por outro lado, o que se sabe a respeito das pacientes com ausência de retração da mama até a última visita é que o evento não aconteceu até aquele momento e que, caso venha a ocorrer, será daquele momento em diante. Mais detalhes podem ser encontrados em Klein e Moeschberger (2003). Para este estudo, nota-se que os dados são de sobrevivência intervalar, tratados no Capítulo 8.

1.4.7 Dados de tempo de vida de mangueiras

Um estudo foi conduzido na Esalq-USP com o objetivo de verificar a resistência de mangueiras a uma praga denominada seca da mangueira, que mata a planta. O objetivo foi identificar novas mangueiras, obtidas a partir de enxertos, que fossem resistentes à praga citada. Para tanto, um experimento fatorial completamente aleatorizado foi realizado com 6 copas enxertadas sobre 7 porta-enxertos (fatorial 6×7). As 42 combinações foram replicadas em 5 blocos diferentes, totalizando 210 unidades experimentais.

O experimento foi instalado em 1971 e visitado 12 vezes entre 1972 e 1992. Em cada visita, registrou-se a condição (viva ou morta) de todas as mangueiras. A resposta de interesse foi o tempo (em anos) até a morte das mangueiras, que é de natureza intervalar, visto que a morte da mangueira ocorre entre duas visitas consecutivas, sendo o tempo exato desconhecido. Detalhes adicionais sobre este estudo podem ser encontrados em Chalita et al. (1999) e Giolo et al. (2009).

1.4.8 Dados de Covid-19 em crianças e adolescentes

O surgimento da Covid-19 foi reconhecido em janeiro de 2020 e, dada à sua propagação pelo mundo, foi declarada pandemia pela OMS em março de 2020. O Brasil, um dos países mais afetados, registrou mais de 16 milhões de casos e 450 mil mortes por Covid-19 até maio de 2021. Embora a Covid-19 pode ser observada em todas as faixas etárias, as crianças apresentam, tipicamente, um quadro menos severo da doença do que os adultos. No entanto, com a progressão da pandemia, manifestações mais severas da doença surgiram em pacientes pediátricos. Neste contexto, torna-se importante avaliar os fatores associados à morte por Covid-19 em crianças e adolescentes hospitalizados com diagnóstico confirmado de SARS-CoV-2.

Oliveira et al. (2021) analisaram os dados de todos os pacientes com idade inferior a vinte anos registrados entre 16/02/2020 e 09/01/2021 no Sistema de Informação de Vigilância Epidemiológica da Gripe (Sivep-Gripe). Este sistema monitora os dados de pacientes admitidos em hospitais brasileiros com quadro respiratório grave. No período mencionado, 11.613 pacientes com idade inferior a vinte anos foram registrados com diagnóstico de Covid-19. O principal objetivo do estudo foi identificar e quantificar os fatores associados ao tempo até a morte do paciente, medido em dias a partir da data de internação. Nesse estudo, o tempo até a alta hospitalar é um evento competitivo, visto que a ocorrência da alta impede a ocorrência da morte. A amostra de pacientes que recebem alta não representa àqueles que estão sob risco de morte; esta é outra forma de caracterizar um evento competitivo. Estes conceitos são explorados no Capítulo 9, assim como a análise de um recorte de 600 pacientes desta base de dados envolvendo três covariáveis: idade, sexo e presença de comorbidades.

1.4.9 Dados de câncer de pele

O câncer de pele melanoma pode aparecer em qualquer parte do corpo na forma de manchas, pintas ou sinais. Embora a cirurgia seja o tratamento mais indicado, radioterapia e quimioterapia também podem ser utilizadas,

dependendo do estadiamento do câncer. Neste cenário, um estudo foi realizado entre 2005 e 2015 em um Centro de Diagnóstico e Tratamento de Câncer. O objetivo foi analisar a sobrevida de 1.117 pacientes diagnosticados com melanoma levando-se em conta o sexo, a idade e o estadiamento da doença. A variável resposta foi o tempo, em meses, desde o diagnóstico do câncer até o óbito. Em razão da melhora expressiva na sobrevivência de pacientes com melanoma, em parte devido ao diagnóstico precoce, há, neste estudo, um percentual elevado de pacientes com períodos longos de sobrevida livre do câncer, o que geralmente indica a cura da doença. Assim, modelos propostos para dados de sobrevivência com essa característica, denominados modelos com fração de cura, ou com fração de imunes, ou, ainda, com sobreviventes de longa duração, foram utilizados para a análise dos dados desse estudo. Esses modelos são tratados no Capítulo 10.

1.5 Especificando o tempo de sobrevivência

Em análise de sobrevivência, a variável aleatória T, não negativa e contínua, é geralmente especificada pela função de sobrevivência ou pela função taxa de falha apresentadas a seguir.

1.5.1 Função de sobrevivência

A função de sobrevivência para T contínua é definida como a probabilidade de um indivíduo sobreviver ao tempo t, isto é, além do tempo t. Em termos probabilísticos, expressa-se por

$$S(t) = P(T > t), \quad t \geq 0,$$

tal que $S(0) = 1$ e $\lim_{t\to\infty} S(t) = 0$. Em consequência, a função de distribuição acumulada, expressa por $F(t) = P(T \leq t) = 1-S(t)$, define a probabilidade de o indivíduo falhar até o tempo t, ou seja, não sobreviver ao tempo t.

A Figura 1.3 apresenta as curvas de sobrevivência associadas a dois grupos de pacientes. A partir dessa figura, é possível notar que o tempo de sobrevida dos pacientes do grupo 1 é superior ao dos pacientes do grupo 2 ao longo da maior parte do tempo de acompanhamento. Para os pacientes

do grupo 1, o tempo em que cerca de 50% deles morrem (tempo mediano) é de 20 anos, enquanto para os pacientes do grupo 2 é de 10 anos. Outra informação que pode ser extraída é o percentual de pacientes vivos até um determinado tempo de interesse. Por exemplo, para os pacientes do grupo 1, é possível observar que cerca de 90% deles estão vivos após 10 anos do início do estudo, enquanto para os do grupo 2 tem-se apenas 50%.

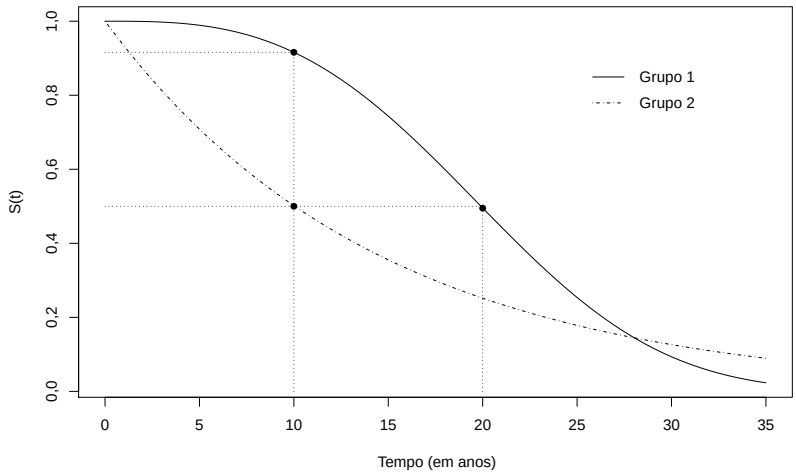


Figura 1.3 – Curvas de sobrevivência para dois grupos de pacientes.

1.5.2 Função taxa de falha

A probabilidade de falha durante o intervalo de tempo $(t_1, t_2]$ pode ser expressa por $P(t_1 < T \leq t_2) = F(t_2) - F(t_1) = S(t_1) - S(t_2)$. Já a taxa de falha no intervalo $(t_1, t_2]$ é definida como a probabilidade de falha neste intervalo, condicional à sobrevivência do indivíduo ao tempo t_1, dividida pelo comprimento do intervalo. Ou seja,

$$\lambda\big((t_1, t_2]\big) = \frac{P(t_1 < T \leq t_2 \,|\, T > t_1)}{(t_2 - t_1)} = \frac{S(t_1) - S(t_2)}{(t_2 - t_1)\, S(t_1)}. \tag{1.1}$$

As taxas de falha são números positivos, mas sem limite superior, isto é, $\lambda(\cdot) \geq 0$. Assim, se $(t_1, t_2] = (10, 11]$ horas com $\lambda\big((10, 11]\big) = 0{,}10$/hora, conclui-se que se há 100 indivíduos sob risco no tempo $t = 10$ horas, então é esperado que 10% deles falhem até $t = 11$ horas.

Redefinindo o intervalo como $(t, t+\Delta t]$, a expressão (1.1) fica dada por

$$\lambda((t, t+\Delta t]) = \frac{F(t+\Delta t) - F(t)}{\Delta t\, S(t)} = \frac{S(t) - S(t+\Delta t)}{\Delta t\, S(t)},$$

de modo que assumindo Δt bem pequeno, o comprimento do intervalo fica bem estreito, resultando na, assim denominada, taxa *instantânea* de falha.

A função taxa de falha é bastante útil para descrever a distribuição do tempo de sobrevida de pacientes, já que ela descreve a forma em que a taxa instantânea de falha muda com o tempo. Formalmente, é definida por

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t \,|\, T > t)}{\Delta t}, \tag{1.2}$$

podendo apresentar várias formas, dentre elas as exibidas na Figura 1.4.

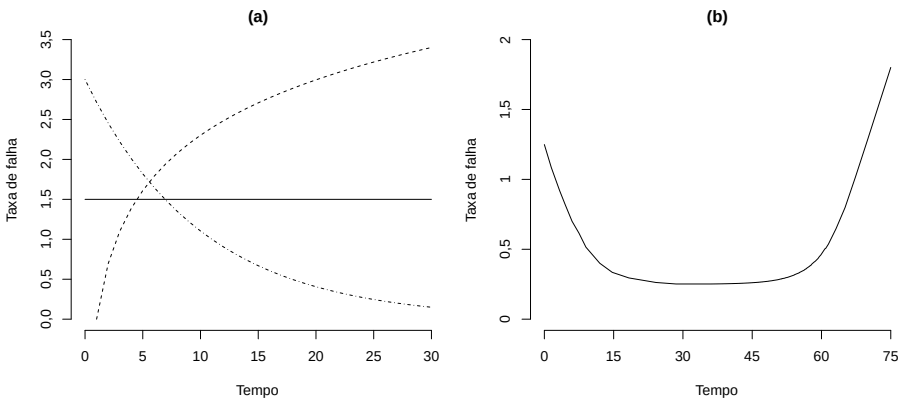


Figura 1.4 – (*a*) Função taxa de falha - - crescente – constante -·- decrescente, e (*b*) função taxa de falha conhecida como curva da banheira.

A função crescente na Figura 1.4 indica que a taxa de falha aumenta com o tempo, o que mostra um efeito gradual de envelhecimento. A constante indica que a taxa de falha não se altera com o tempo, enquanto a decrescente mostra que a taxa de falha diminui à medida que o tempo passa. Por sua vez, a taxa de falha associada ao tempo de vida dos seres humanos corresponde a uma combinação das três curvas mencionadas, sendo conhecida como curva da banheira por ter um comportamento decrescente no período inicial, representando a mortalidade infantil, constante na faixa intermediária e crescente no período final (efeito do envelhecimento).

A função taxa de falha é, em geral, mais informativa do que a função de sobrevivência. Isso porque diferentes funções de sobrevivência podem apresentar formas semelhantes, enquanto suas correspondentes funções taxa de falha podem diferir drasticamente. Dessa forma, a modelagem da função taxa de falha é um importante método para dados de sobrevivência.

1.5.3 Função taxa de falha acumulada

A função taxa de falha acumulada fornece, como o próprio nome sugere, a taxa de falha acumulada do indivíduo, sendo definida por

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

A inclinação da função $\Lambda(t)$ corresponde à taxa de falha $\lambda(t)$. Portanto, a interpretação de $\Lambda(t)$ tem como foco avaliar a inclinação desta função em cada tempo t ao longo de todo o eixo do tempo.

A partir da Figura 1.5, que mostra três formas para $\Lambda(t)$ e suas respectivas funções $\lambda(t)$, nota-se que a função taxa de falha acumulada $\Lambda_1(t)$ apresenta inclinação igual em todo o eixo do tempo, o que implica $\lambda_1(t)$ constante. Por sua vez, é possível observar que a inclinação da função $\Lambda_2(t)$ aumenta e a de $\Lambda_3(t)$ diminui ao longo do eixo do tempo, o que implica $\lambda_2(t)$ crescente e $\lambda_3(t)$ decrescente.

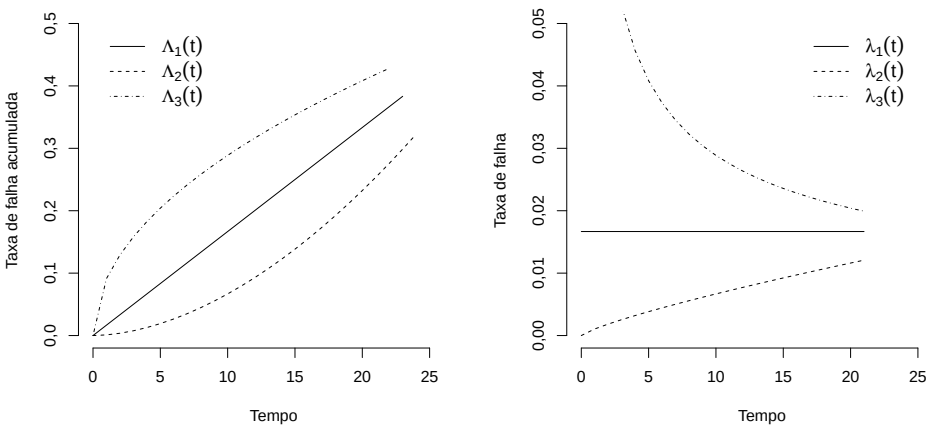


Figura 1.5 – Função taxa de falha acumulada e respectiva função taxa de falha.

1.5.4 Tempo médio e vida média residual

Outras duas quantidades de interesse em análise de sobrevivência são o tempo médio de vida e a vida média residual. Para T contínua, a primeira é obtida pela área sob a função de sobrevivência. Isto é,

$$t_m = E(T) = \int_0^\infty t f(t)\, dt = \int_0^\infty S(t)\, dt.$$

Já a vida média residual (ou tempo médio restante de vida) é definida condicional a um certo tempo t. Assim, para os indivíduos que sobreviverem ao tempo t, essa quantidade fornece o tempo que, em média, eles ainda têm de vida, sendo obtida pela área sob a curva de sobrevivência à direita do tempo t dividida por $S(t)$. Isto é,

$$\text{vmr}(t) = E(T - t \mid T > t) = \frac{\int_t^\infty (u-t) f(u) du}{S(t)} = \frac{\int_t^\infty S(u) du}{S(t)},$$

com $f(\cdot)$ a função densidade de probabilidade de T dada por

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t)}{\Delta t}.$$

A partir da expressão $\text{vmr}(t)$, observa-se que $\text{vmr}(0) = t_m$.

1.5.5 Relações entre as funções

Algumas relações probabilísticas importantes entre as funções definidas anteriormente são apresentadas a seguir.

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}\Big(\log[S(t)]\Big),$$

$$\Lambda(t) = \int_0^t \lambda(u)\, du = -\log\big[S(t)\big] \tag{1.3}$$

e

$$S(t) = \exp(-\Lambda(t)) = \exp\Big(-\int_0^t \lambda(u)\, du\Big). \tag{1.4}$$

Tais relações mostram que o conhecimento de uma das funções, por exemplo $S(t)$, implica no conhecimento de $F(t)$, $f(t)$, $\lambda(t)$ e $\Lambda(t)$.

Outras relações envolvendo estas funções são as seguintes:

$$S(t) = \frac{\text{vmr}(0)}{\text{vmr}(t)} \exp\left(-\int_0^t \frac{du}{\text{vmr}(u)}\right)$$

e

$$\lambda(t) = \left(\frac{d\,\text{vmr}(t)}{dt} + 1\right) / \text{vmr}(t).$$

1.5.6 Funções no caso discreto

As funções $S(t)$, $\lambda(t)$ e $\Lambda(t)$ foram definidas para uma variável aleatória T não negativa e contínua. Entretanto, há situações em que é razoável assumir que T seja uma variável aleatória discreta. Por exemplo, quando os indivíduos são avaliados nas mesmas datas ao longo de um período de seguimento e nelas ocorre o registro do evento para diversos deles, resultando em vários tempos até o evento iguais a t_1, t_2, ..., t_k. Ou seja, há a presença de empates em t_j, $j = 1, \ldots, k$. Essa situação é tratada no Capítulo 8.

Para as situações em que é assumido que T é variável aleatória discreta, são apresentadas, a seguir, as expressões das funções $S(t)$, $\lambda(t)$ e $\Lambda(t)$. Considere que T assume os valores t_1, t_2, ..., t_k, com $0 \leq t_1 < \ldots < t_k$, e que sua função de probabilidade seja $p(t_j) = P(T = t_j)$. Neste caso,

$$S(t) = P(T > t) = \sum_{t_j > t} p(t_j).$$

Ainda, a função taxa de falha fica definida, para $j = 1, 2, \ldots, k$, por

$$\lambda(t_j) = P(T = t_j \,|\, T > t_{j-1}) = \frac{p(t_j)}{S(t_{j-1})}.$$

Como $P(T \geq t_j) = P(T > t_{j-1}) = S(t_{j-1})$ e $p(t_j) = S(t_{j-1}) - S(t_j)$, segue que a função taxa de falha pode ser escrita por

$$\lambda(t_j) = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} = 1 - \frac{S(t_j)}{S(t_{j-1})}.$$

Utilizando esta última relação e indução matemática, a função de sobrevivência pode ser escrita em termos da função taxa de falha como

$$S(t) = P(T > t) = \prod_{t_j \leq t} [1 - \lambda(t_j)].$$

Quanto à função taxa de falha acumulada, uma definição intuitiva é

$$\Lambda(t) = \sum_{t_j \leq t} \lambda(t_j).$$

Contudo, ela não preserva a relação (1.4) vista para o caso contínuo. Uma alternativa que preserva esta relação é dada, para o caso discreto, por

$$\Lambda(t) = -\sum_{t_j \leq t} \log[1 - \lambda(t_j)] = -\log[S(t)].$$

Quando a taxa de falha em cada tempo for pequena, essas duas definições produzem resultados próximos.

1.6 Exercícios

1. Seis ratos foram expostos a um material cancerígeno com o objetivo de se observar o tempo até o desenvolvimento do tumor. Ao final de 30 semanas de estudo, o seguinte cenário foi registrado: os ratos A, B e C desenvolveram o tumor em 10, 15 e 25 semanas, respectivamente, o rato D morreu acidentalmente sem tumor na 20ª semana e os ratos E e F permaneceram livres do tumor até o término do estudo.

(a) Defina a resposta para este estudo.

(b) Identifique o tipo de resposta (falha ou censura) observada para cada um dos seis ratos no estudo.

2. Indivíduos que ingressaram em um estudo ao longo de sua realização foram acompanhados com o intuito de se observar o aparecimento de um certo sintoma. A resposta considerada foi a idade em que o sintoma apareceu pela primeira vez. Para os seis indivíduos descritos a seguir, identifique o tipo de censura registrado para cada um deles.

(a) O primeiro indivíduo ingressou no estudo aos 25 anos de idade já apresentando o sintoma.

(b) Outros dois indivíduos ingressaram no estudo aos 20 e 28 anos de idade e não apresentaram o sintoma até o término do estudo.

(c) Outros dois indivíduos ingressaram aos 35 e 40 anos de idade e apresentaram o sintoma respectivamente no 2º e 6º exames após terem entrado no estudo. Os exames foram realizados a cada dois anos.

(d) O sexto indivíduo, que ingressou aos 36 anos de idade, mudou de cidade após 4 anos no estudo sem apresentar o sintoma.

3. Mostre que $\lambda(t) = \dfrac{f(t)}{S(t)} = -\dfrac{d}{dt}\Big(\log[S(t)]\Big)$.

4. Mostre que $\Lambda(t) = \int_0^t \lambda(u)du = -\log[S(t)]$.

5. Mostre que $\text{vmr}(t) = \dfrac{\int_t^\infty (u-t)f(u)du}{S(t)} = \dfrac{\int_t^\infty S(u)du}{S(t)}$.

6. Suponha que a taxa de falha associada à variável T seja expressa pela função linear $\lambda(t) = \beta_0 + \beta_1 t$, com β_0 e $\beta_1 > 0$. Obtenha $S(t)$ e $f(t)$.

7. Suponha que a vida média residual associada à variável T seja dada por $\text{vmr}(t) = t + 10$. Obtenha $E(T)$, $\lambda(t)$ e $S(t)$.

8. Para cada um dos exemplos descritos na Seção 1.4, identifique o tempo inicial, a escala de medida e o evento de interesse.

9. Para uma variável aleatória T discreta, mostre que:

(a) $S(t) = \prod\limits_{t_j \leq t} S(t_j)/S(t_{j-1})$. Use indução matemática.

(b) $S(t) = \prod\limits_{t_j \leq t} [1 - \lambda(t_j)]$. Use o resultado anterior.

Enrico Antônio Colosimo

Professor do Departamento de Estatística da UFMG. Ph.D. em Estatística pela University of Wisconsin-Madison, Estados Unidos. Foi editor-chefe do *Brazilian Journal of Probability and Statistics* no período de 2019 a 2020. É coautor do livro *Confiabilidade: análise de tempo de falha e testes de vida acelerados*. Suas áreas de interesse são: métodos estatísticos em análise de sobrevivência, confiabilidade/sistemas reparáveis e dados longitudinais.

Suely Ruiz Giolo

Professora do Departamento de Estatística da UFPR. Licenciada em Matemática e Bacharel em Estatística pela Unesp, Mestre em Estatística pela Unicamp e Doutora em Estatística e Experimentação Agronômica pela Esalq-USP e pela Lancaster University, Inglaterra. Realizou pós-doutorado em Ciências Biológicas no InCor-FMUSP e no IME-USP. É autora do livro *Introdução à análise de dados categóricos com aplicações*.

Sobre o livro

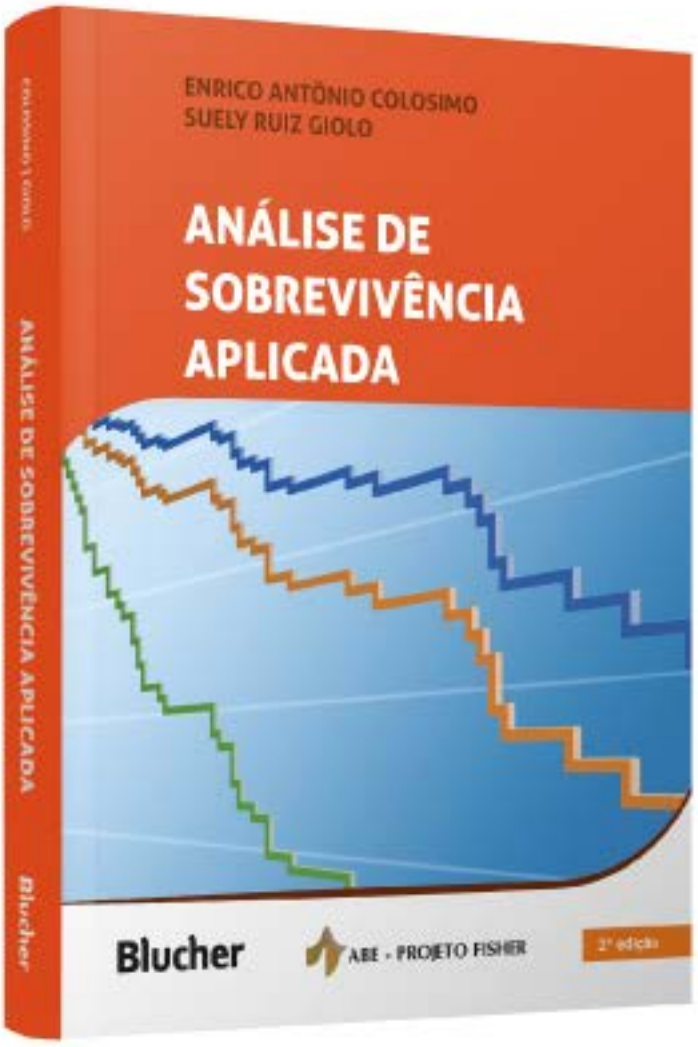
Este livro aborda conceitos básicos, técnicas não paramétricas e modelos de regressão para a análise de dados de sobrevivência. São apresentados os modelos de Cox, de Aalen e o de fragilidade, bem como modelos para a análise de dados de sobrevivência intervalar e grupados. Os tópicos adicionais contemplados nesta segunda edição compreendem métodos e modelos para a análise de dados nos contextos de riscos competitivos e de fração de imunes. Com o propósito de ilustrar as metodologias expostas, vários exemplos reais são analisados no texto. Para a execução das análises foi adotado o *software* estatístico R, de domínio público. Os comandos utilizados estão no endereço eletrônico https://docs.ufpr.br/~giolo/asa. O livro é voltado para alunos de cursos de Estatística, bem como para alunos, profissionais e pesquisadores de outras áreas (Saúde, Biológicas, Ciências Humanas, Engenharias etc.), que tenham interesse em metodologias para a análise de dados de sobrevivência.

ISBN 978-85-212-2199-9

ABE - PROJETO FISHER

www.blucher.com.br

Blucher



Clique aqui e:

VEJA NA LOJA

Análise de sobrevivência aplicada

Enrico Antônio Colosimo, Suely Ruiz Giolo

ISBN: 9788521221999

Páginas: 362

Formato: 17 x 24 cm

Ano de Publicação: 2024