

Análise de Sobrevivência: Uma abordagem unificada via Processo de Contagem

Josemar Rodrigues
Universidade Federal de São Carlos

Escola de Verão 2012,
16/01/2012-10/02/2012-Natal-RGN

Tópicos

- **Porquê uma Análise de Sobrevida Unificada?**
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevivência Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevivência integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevivência
- Conclusões: Pesquisas em andamento
- Referências

Tópicos

- Porquê uma Análise de Sobrevida Unificada?
- Processo de contagem no intervalo $[0, t]$
- Análise de recorrência
- Análise de Sobrevida integrada (Clássico e Bayesiano)
- Teoria Unificada de Longa Duração
- Modelo destrutivo e destrutivo correlacionado
- Modelo Híbrido: Flexibilização do modelo em dois estágios baseado no processo da carcinogênese.
- Um mecanismo unificado e composto para formulação de novas distribuições de probabilidades em Análise de Sobrevida
- Conclusões: Pesquisas em andamento
- Referências

Porquê uma Análise de Sobrevivência Unificada?

- Análise de Sobrevivência é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevivência: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevivência Unificada?

- Análise de Sobrevivência é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevivência: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevivência Unificada?

- Análise de Sobrevivência é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevivência: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevivência Unificada?

- Análise de Sobrevivência é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevivência: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevida Unificada?

- Análise de Sobrevida é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevida: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevivência Unificada?

- Análise de Sobrevivência é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevivência: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Porquê uma Análise de Sobrevida Unificada?

- Análise de Sobrevida é uma das mais antigas áreas da Estatística (Século 17, Ciências Atuariais e Demografia). Sugestão: Ler Andersen (1985).
- Antes da 2a. guerra mundial: AS preocupava-se com a taxa de morte.
- Análise de Sobrevida: Análise do tempo de ocorrência de um evento de interesse sob censura.
- Artigo de impacto: Artigo de Kaplan-Meier (1958) com mais de 33.000 citações até 04/2009.
- Artigo de impacto (Introdução de covariadas): Artigo de Sir David Cox (1972) com 24.000 citações até 04/2009.
- Pacientes imunes ao evento de interesse: Boag, J.W. (1949) e Berkson/Gage (1952). Maior impacto a partir de 1996 com o texto de Maller/Zhou.
- Pergunta: Porquê somente agora a AS está preocupada com os imunes e com modelos probabilísticos alternativos?

Modelo de Cox

- Porquê Sir David Cox não introduziu a taxa de cura em seu famoso modelo de riscos proporcionais?.
- É possível justificar o interesse atual pelos modelos de longa duração e a generalização dos clássicos modelos de AS através de uma abordagem unificada com interpretação biológica?



Figura: David Cox

Notações

- **Fatores de risco:** $N =$ no. de fatores de risco que estão competindo para a ocorrência de um evento durante o intervalo $[0, \tau_0]$, onde τ_0 é o tempo limite (conhecido) da unidade em estudo.
- **Distribuição de probabilidade de N :**

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots$$

- **Tempo de vida da unidade:** T
- **Tempo de promoção:** $Y_i(t)$ = tempo de promoção para o i -ésimo fator de risco causar o evento de interesse no intervalo $[0, t]$, para $i = 1, \dots, N$.

Notações

- **Fatores de risco:** $N =$ no. de fatores de risco que estão competindo para a ocorrência de um evento durante o intervalo $[0, \tau_0]$, onde τ_0 é o tempo limite (conhecido) da unidade em estudo.
- **Distribuição de probabilidade de N :**

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots$$

- **Tempo de vida da unidade:** T
- **Tempo de promoção:** $Y_i(t)$ = tempo de promoção para o i -ésimo fator de risco causar o evento de interesse no intervalo $[0, t]$, para $i = 1, \dots, N$.

Notações

- **Fatores de risco:** N = no. de fatores de risco que estão competindo para a ocorrência de um evento durante o intervalo $[0, \tau_0]$, onde τ_0 é o tempo limite (conhecido) da unidade em estudo.
- **Distribuição de probabilidade de N :**

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots$$

- **Tempo de vida da unidade:** T
- **Tempo de promoção:** $Y_i(t)$ = tempo de promoção para o i -ésimo fator de risco causar o evento de interesse no intervalo $[0, t]$, para $i = 1, \dots, N$.

Notações

- **Fatores de risco:** $N =$ no. de fatores de risco que estão competindo para a ocorrência de um evento durante o intervalo $[0, \tau_0]$, onde τ_0 é o tempo limite (conhecido) da unidade em estudo.
- **Distribuição de probabilidade de N :**

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots$$

- **Tempo de vida da unidade:** T
- **Tempo de promoção:** $Y_i(t) =$ tempo de promoção para o i -ésimo fator de risco causar o evento de interesse no intervalo $[0, t]$, para $i = 1, \dots, N$.

Suposições

- **Tempo de promoção :**
Dado $N(t) = n$ (“shared frailty”),

$$Y_i \sim F(t), i = 1, 2, \dots, n. \text{ (v.a.i).}$$

- Tempo de promoção ordenado gerado por $F(t)$ (latente ou observado):
Dado $N(t) = n$, vamos supor que os n fatores de riscos ocorrem nos tempos ordenados

$$Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)}.$$

Suposições

- **Tempo de promoção :**
Dado $N(t) = n$ (“shared frailty”),

$$Y_i \sim F(t), i = 1, 2, \dots, n. \text{ (v.a.i).}$$

- **Tempo de promoção ordenado gerado por $F(t)$ (latente ou observado):**
Dado $N(t) = n$, vamos supor que os n fatores de riscos ocorrem nos tempos ordenados

$$Y_{(1)} \leq Y_{(2)} \leq \dots, Y_{(n)}.$$

Processo de Contagem

- **Processo de contagem:** Consideremos a seguinte v.a. composta

$$N(t) = \begin{cases} Z_1(t) + Z_2(t) + \cdots + Z_N & \text{se } N > 0, \\ 0, & \text{se } N = 0, \end{cases} \quad (1)$$

onde $Z_i(t)$ indica a presença do i -ésimo fator de risco durante o intervalo de tempo $[0, t]$.

Suposição (mecanismo de ativação):

$N(t) \mid N = n \sim \text{Binomial}[n, p(t)]$, onde $p(t)$ é a probabilidade da presença do i -ésimo fator de risco durante um intervalo de tempo $[0, t]$.

- **Modelos de longa duração (taxa de cura):** $p_0 = P[N = 0]$.
- $\{N(t), p(t), 0 \leq t \leq \tau_0\}$: Expressa o comportamento competitivo entre os fatores de risco para causar o evento de interesse durante o intervalo $[0, t]$.

Processo de Contagem

- **Processo de contagem:** Consideremos a seguinte v.a. composta

$$N(t) = \begin{cases} Z_1(t) + Z_2(t) + \cdots + Z_N & \text{se } N > 0, \\ 0, & \text{se } N = 0, \end{cases} \quad (1)$$

onde $Z_i(t)$ indica a presença do i -ésimo fator de risco durante o intervalo de tempo $[0, t]$.

Suposição (mecanismo de ativação):

$N(t) \mid N = n \sim \text{Binomial}[n, p(t)]$, onde $p(t)$ é a probabilidade da presença do i -ésimo fator de risco durante um intervalo de tempo $[0, t]$.

- **Modelos de longa duração (taxa de cura):** $p_0 = P[N = 0]$.
- $\{N(t), p(t), 0 \leq t \leq \tau_0\}$: Expressa o comportamento competitivo entre os fatores de risco para causar o evento de interesse durante o intervalo $[0, t]$.

Processo de Contagem

- **Processo de contagem:** Consideremos a seguinte v.a. composta

$$N(t) = \begin{cases} Z_1(t) + Z_2(t) + \cdots + Z_N & \text{se } N > 0, \\ 0, & \text{se } N = 0, \end{cases} \quad (1)$$

onde $Z_i(t)$ indica a presença do i -ésimo fator de risco durante o intervalo de tempo $[0, t]$.

Suposição (mecanismo de ativação):

$N(t) \mid N = n \sim \text{Binomial}[n, p(t)]$, onde $p(t)$ é a probabilidade da presença do i -ésimo fator de risco durante um intervalo de tempo $[0, t]$.

- **Modelos de longa duração (taxa de cura):** $p_0 = P[N = 0]$.
- $\{N(t), p(t), 0 \leq t \leq \tau_0\}$: Expressa o comportamento competitivo entre os fatores de risco para causar o evento de interesse durante o intervalo $[0, t]$.

Mecanismo de ativação

- Composição do mecanismo de ativação:

$$P[N(t) = j; p(t)] = \sum_{n=j}^{\infty} P[N = n] \overbrace{P[N(t) = j | N = n; p(t)]}^{\text{mecanismo de ativação}}. \quad (2)$$

- A probabilidade de ativação $p(t)$ no instante t unifica diferentes mecanismo de ativação existentes na literatura.
- Primeira ativação no instante t (Rodrigues *et al.*, 2011b): $[N(t) = 1] \Rightarrow P[N(t) = 1; p(t)]$ (" incidence function").
- Última ativação no instante t : $[N(t) = N] \Rightarrow P[N(t) = N; p(t)]$ ("net survival").

Mecanismo de ativação

- Composição do mecanismo de ativação:

$$P[N(t) = j; p(t)] = \sum_{n=j}^{\infty} P[N = n] \overbrace{P[N(t) = j | N = n; p(t)]}^{\text{mecanismo de ativação}}. \quad (2)$$

- A probabilidade de ativação $p(t)$ no instante t unifica diferentes mecanismo de ativação existentes na literatura.
- Primeira ativação no instante t (Rodrigues *et al.*, 2011b): $[N(t) = 1] \Rightarrow P[N(t) = 1; p(t)]$ (“incidence function”).
- Última ativação no instante t : $[N(t) = N] \Rightarrow P[N(t) = N; p(t)]$ (“net survival”).

Mecanismo de ativação

- Composição do mecanismo de ativação:

$$P[N(t) = j; p(t)] = \sum_{n=j}^{\infty} P[N = n] \overbrace{P[N(t) = j | N = n; p(t)]}^{\text{mecanismo de ativação}}. \quad (2)$$

- A probabilidade de ativação $p(t)$ no instante t unifica diferentes mecanismo de ativação existentes na literatura.
- Primeira ativação no instante t (Rodrigues *et al.*, 2011b): $[N(t) = 1] \Rightarrow P[N(t) = 1; p(t)]$ (" incidence function").
- Última ativação no instante t : $[N(t) = N] \Rightarrow P[N(t) = N; p(t)]$ ("net survival").

Mecanismo de ativação

- Composição do mecanismo de ativação:

$$P[N(t) = j; p(t)] = \sum_{n=j}^{\infty} P[N = n] \overbrace{P[N(t) = j | N = n; p(t)]}^{\text{mecanismo de ativação}}. \quad (2)$$

- A probabilidade de ativação $p(t)$ no instante t unifica diferentes mecanismo de ativação existentes na literatura.
- Primeira ativação no instante t (Rodrigues *et al.*, 2011b): $[N(t) = 1] \Rightarrow P[N(t) = 1; p(t)]$ (" incidence function").
- Última ativação no instante t : $[N(t) = N] \Rightarrow P[N(t) = N; p(t)]$ ("net survival").

Casos específicos do mecanismo de ativação

- Processo de promoção: $p(t) = F(t)$, onde $Y_i(t) \sim F(t)$.
- Processo de reparação ou destrutivo: $p(t) = p \Rightarrow N(t) = D \leq N$.
- TULD: $p(t) = 1 \Rightarrow t = \tau_0 \Rightarrow N(\tau_0) = N$. Não temos reparação ou eliminação de fatores de risco (células iniciadas)
- Processo de Poisson Ponderado Unitário (Balakrishnan and Kozubowski, 2008):
 $X_i(t) \sim \text{Uniforme}[0, t], \tau_0 = 1 \Rightarrow p(t) = t, 0 \leq t \leq 1$.
- Processo binário (AS): $N(t) \sim \text{Bernoulli}(F(t))$.

Casos específicos do mecanismo de ativação

- Processo de promoção: $p(t) = F(t)$, onde $Y_i(t) \sim F(t)$.
- Processo de reparação ou destrutivo: $p(t) = p \Rightarrow N(t) = D \leq N$.
- TULD: $p(t) = 1 \Rightarrow t = \tau_0 \Rightarrow N(\tau_0) = N$. Não temos reparação ou eliminação de fatores de risco (células iniciadas)
- Processo de Poisson Ponderado Unitário (Balakrishnan and Kozubowski, 2008):
 $X_i(t) \sim \text{Uniforme}[0, t], \tau_0 = 1 \Rightarrow p(t) = t, 0 \leq t \leq 1$.
- Processo binário (AS): $N(t) \sim \text{Bernoulli}(F(t))$.

Casos específicos do mecanismo de ativação

- Processo de promoção: $p(t) = F(t)$, onde $Y_i(t) \sim F(t)$.
- Processo de reparação ou destrutivo: $p(t) = p \Rightarrow N(t) = D \leq N$.
- TULD: $p(t) = 1 \Rightarrow t = \tau_0 \Rightarrow N(\tau_0) = N$. Não temos reparação ou eliminação de fatores de risco (células iniciadas)
- Processo de Poisson Ponderado Unitário (Balakrishnan and Kozubowski, 2008):
 $X_i(t) \sim \text{Uniforme}[0, t], \tau_0 = 1 \Rightarrow p(t) = t, 0 \leq t \leq 1$.
- Processo binário (AS): $N(t) \sim \text{Bernoulli}(F(t))$.

Casos específicos do mecanismo de ativação

- Processo de promoção: $p(t) = F(t)$, onde $Y_i(t) \sim F(t)$.
- Processo de reparação ou destrutivo: $p(t) = p \Rightarrow N(t) = D \leq N$.
- TULD: $p(t) = 1 \Rightarrow t = \tau_0 \Rightarrow N(\tau_0) = N$. Não temos reparação ou eliminação de fatores de risco (células iniciadas)
- Processo de Poisson Ponderado Unitário (Balakrishnan and Kozubowski, 2008):
 $X_i(t) \sim \text{Uniforme}[0, t], \tau_0 = 1 \Rightarrow p(t) = t, 0 \leq t \leq 1$.
- Processo binário (AS): $N(t) \sim \text{Bernoulli}(F(t))$.

Casos específicos do mecanismo de ativação

- Processo de promoção: $p(t) = F(t)$, onde $Y_i(t) \sim F(t)$.
- Processo de reparação ou destrutivo: $p(t) = p \Rightarrow N(t) = D \leq N$.
- TULD: $p(t) = 1 \Rightarrow t = \tau_0 \Rightarrow N(\tau_0) = N$. Não temos reparação ou eliminação de fatores de risco (células iniciadas)
- Processo de Poisson Ponderado Unitário (Balakrishnan and Kozubowski, 2008):
 $X_i(t) \sim \text{Uniforme}[0, t], \tau_0 = 1 \Rightarrow p(t) = t, 0 \leq t \leq 1$.
- Processo binário (AS): $N(t) \sim \text{Bernoulli}(F(t))$.

Continuação....

- Teoria de filas (área carcinogênica, Klebanov *et al.* (1993)): Seja m o número de servidores ou unidades de reparação com uma taxa de serviço μ_0 e $h_0 = E[N]/\tau_0$:

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!}.$$

- Teoria de Filas ($m = \infty$)(Yakovlev and Polig, 1996):

$$U | T = t \sim G(u)I_{[0 \leq u \leq t]}(u) \Rightarrow p(t) = \int_0^t F(t-u)dG(u).$$

- Teoria de filas (Yakovlev and Polig, 1996) :

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!},$$

onde $h_0 = \theta h$, h : quantidade de droga e θ : sensibilidade a droga.

Continuação....

- Teoria de filas (área carcinogênica, Klebanov *et al.* (1993)): Seja m o número de servidores ou unidades de reparação com uma taxa de serviço μ_0 e $h_0 = E[N]/\tau_0$:

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!}.$$

- Teoria de Filas ($m = \infty$)(Yakovlev and Polig, 1996):

$$U | T = t \sim G(u)I_{[0 \leq u \leq t]}(u) \Rightarrow p(t) = \int_0^t F(t-u)dG(u).$$

- Teoria de filas (Yakovlev and Polig, 1996) :

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!},$$

onde $h_0 = \theta h$, h : quantidade de droga e θ : sensibilidade a droga.

Continuação....

- Teoria de filas (área carcinogênica, Klebanov *et al.* (1993)): Seja m o número de servidores ou unidades de reparação com uma taxa de serviço μ_0 e $h_0 = E[N]/\tau_0$:

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!}.$$

- Teoria de Filas ($m = \infty$)(Yakovlev and Polig, 1996):

$$U | T = t \sim G(u)I_{[0 \leq u \leq t]}(u) \Rightarrow p(t) = \int_0^t F(t-u)dG(u).$$

- Teoria de filas (Yakovlev and Polig, 1996) :

$$p(t) = p = \frac{(h_0/\mu_0)^m/m!}{\sum_{k=0}^m (h_0/\mu_0)^k/k!},$$

onde $h_0 = \theta h$, h : quantidade de droga e θ : sensibilidade a droga.

Mecanismo de seleção (Rodrigues *et al.*, 2011b)

- Variável seletiva: Divide as unidades em dois grupos. Os imunes que não estão sujeitos ao evento de interesse para qualquer instante t e os não imunes.

$$U(t) = \begin{cases} 1 & \text{se } N(t) > 1, \\ 0, & \text{se } N(t) = 0. \end{cases} \quad (3)$$

- Distribuição principal: $Y \sim G(t)$
Problema: $T \sim F(t) = ?$ (versão ponderada da G).

$$\begin{aligned} f_T(t) &= \frac{g(t)P[U(t) = 1 \mid Y \leq t]}{P[U(t) = 1]} = \frac{g(t)P[U(t) = 1 \mid Y \leq t]}{1 - p_0} \\ &= \frac{w(t)g(t)}{E[w(Y)]} \text{ (Distribuição ponderada),} \end{aligned}$$

onde

$$1 - w(t) = P[U(t) = 0 \mid Y \leq t] \text{ "Conditional personal probability"}$$

Mecanismo de seleção (Rodrigues *et al.*, 2011b)

- Variável seletiva: Divide as unidades em dois grupos. Os imunes que não estão sujeitos ao evento de interesse para qualquer instante t e os não imunes.

$$U(t) = \begin{cases} 1 & \text{se } N(t) > 1, \\ 0, & \text{se } N(t) = 0. \end{cases} \quad (3)$$

- Distribuição principal: $Y \sim G(t)$

Problema: $T \sim F(t) = ?$ (versão ponderada da G).

$$\begin{aligned} f_T(t) &= \frac{g(t)P[U(t) = 1 \mid Y \leq t]}{P[U(t) = 1]} = \frac{g(t)P[U(t) = 1 \mid Y \leq t]}{1 - p_0} \\ &= \frac{w(t)g(t)}{E[w(Y)]} \text{ (Distribuição ponderada),} \end{aligned}$$

onde

$1 - w(t) = P[U(t) = 0 \mid Y \leq t]$ "Conditional personal probability"

Continuação...

- 1a. ativação:

$$N(t) = 1 \Rightarrow T = \min\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\min}(t)g(t)}{E[w_{\min}(Y)]}.$$

- Última ativação:

$$N(t) = N \Rightarrow T = \max\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\max}(t)g(t)}{E[w_{\max}(Y)]}.$$

- Composto ou híbrido: $N_j(t), j = 1, \dots, N_t \Rightarrow T =$

$$\min_{1 \leq j \leq N_t} \max\{Y_{j1}, \dots, Y_{j, N_j(t)}\} \Rightarrow f_T(t) = \frac{w_{\min\max}(t)g(t)}{E[w_{\min\max}(Y)]}.$$

Continuação...

- 1a. ativação:

$$N(t) = 1 \Rightarrow T = \min\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\min}(t)g(t)}{E[w_{\min}(Y)]}.$$

- Última ativação:

$$N(t) = N \Rightarrow T = \max\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\max}(t)g(t)}{E[w_{\max}(Y)]}.$$

- Composto ou híbrido: $N_j(t), j = 1, \dots, N_t \Rightarrow T =$

$$\min_{1 \leq j \leq N_t} \max\{Y_{j1}, \dots, Y_{j, N_j(t)}\} \Rightarrow f_T(t) = \frac{w_{\min\max}(t)g(t)}{E[w_{\min\max}(Y)]}.$$

Continuação...

- 1a. ativação:

$$N(t) = 1 \Rightarrow T = \min\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\min}(t)g(t)}{E[w_{\min}(Y)]}.$$

- Última ativação:

$$N(t) = N \Rightarrow T = \max\{Y_1, \dots, Y_{N(t)}\} \Rightarrow f_T(t) = \frac{w_{\max}(t)g(t)}{E[w_{\max}(Y)]}.$$

- Composto ou híbrido: $N_j(t), j = 1, \dots, N_t \Rightarrow T =$

$$\min_{1 \leq j \leq N_t} \max\{Y_{j1}, \dots, Y_{j, N_j(t)}\} \Rightarrow f_T(t) = \frac{w_{\min\max}(t)g(t)}{E[w_{\min\max}(Y)]}.$$

Conclusão: Análise de Sobrevivência Unificada (ASU)

$$w(t) = \overbrace{P[U(t) = 1 \mid Y \leq t]}^{\text{Mecanismo prob. de seleção}}$$

$$ASU = \begin{cases} U \not\perp X(\text{ASFC}) \Rightarrow \begin{cases} \text{ASI} \Rightarrow f_T(t) = g(t)w(t), \\ \text{ASP} \Rightarrow f_T(t) = \frac{w(t)g(t)}{E[w(Y)]}, \end{cases} \\ U \perp Y(\text{ASC}) \Rightarrow f_T(t) = g(t). \end{cases} \quad (4)$$

- ASFC: Análise de Sobrevivência com Fração de Cura;
- ASI: Análise de Sobrevivência Imprópria;
- ASP: Análise de Sobrevivência Ponderada.
- ASC: Análise de Sobrevivência Clássica (dois estados: vivo ou morto e $w(t) = 1 - p_0 \Rightarrow f_T(t) = g(t)$).

Conclusão: Análise de Sobrevivência Unificada (ASU)

$$w(t) = \overbrace{P[U(t) = 1 \mid Y \leq t]}^{\text{Mecanismo prob. de seleção}}$$

$$ASU = \begin{cases} U \not\perp X(\text{ASFC}) \Rightarrow & \begin{cases} \text{ASI} \Rightarrow f_T(t) = g(t)w(t), \\ \text{ASP} \Rightarrow f_T(t) = \frac{w(t)g(t)}{E[w(Y)]}, \end{cases} \\ U \perp Y(\text{ASC}) \Rightarrow & f_T(t) = g(t). \end{cases} \quad (4)$$

- ASFC: Análise de Sobrevivência com Fração de Cura;
- ASI: Análise de Sobrevivência Imprópria;
- ASP: Análise de Sobrevivência Ponderada.
- ASC: Análise de Sobrevivência Clássica (dois estados: vivo ou morto e $w(t) = 1 - p_0 \Rightarrow f_T(t) = g(t)$).

Conclusão: Análise de Sobrevivência Unificada (ASU)

$$w(t) = \overbrace{P[U(t) = 1 \mid Y \leq t]}^{\text{Mecanismo prob. de seleção}}$$

$$ASU = \begin{cases} U \not\perp X(\text{ASFC}) \Rightarrow & \begin{cases} \text{ASI} \Rightarrow f_T(t) = g(t)w(t), \\ \text{ASP} \Rightarrow f_T(t) = \frac{w(t)g(t)}{E[w(Y)]}, \end{cases} \\ U \perp Y(\text{ASC}) \Rightarrow & f_T(t) = g(t). \end{cases} \quad (4)$$

- ASFC: Análise de Sobrevivência com Fração de Cura;
- ASI: Análise de Sobrevivência Imprópria;
- ASP: Análise de Sobrevivência Ponderada.
- ASC: Análise de Sobrevivência Clássica (dois estados: vivo ou morto e $w(t) = 1 - p_0 \Rightarrow f_T(t) = g(t)$).

Conclusão: Análise de Sobrevivência Unificada (ASU)

$$w(t) = \overbrace{P[U(t) = 1 \mid Y \leq t]}^{\text{Mecanismo prob. de seleção}}$$

$$ASU = \begin{cases} U \not\perp X(\text{ASFC}) \Rightarrow & \begin{cases} \text{ASI} \Rightarrow f_T(t) = g(t)w(t), \\ \text{ASP} \Rightarrow f_T(t) = \frac{w(t)g(t)}{E[w(Y)]}, \end{cases} \\ U \perp Y(\text{ASC}) \Rightarrow & f_T(t) = g(t). \end{cases} \quad (4)$$

- ASFC: Análise de Sobrevivência com Fração de Cura;
- ASI: Análise de Sobrevivência Imprópria;
- ASP: Análise de Sobrevivência Ponderada.
- ASC: Análise de Sobrevivência Clássica (dois estados: vivo ou morto e $w(t) = 1 - p_0 \Rightarrow f_T(t) = g(t)$).

Problemas..

- Modelo latente (ML): $\{\overbrace{Y, N(t), t > 0}^{\text{latente.}}\}$ Como formular $w(t)$ e qual esquema de ativação deveremos utilizar para associar T com Y e $N(t)$.?
- Modelo recorrente (MR): $\{\overbrace{T, Y, N(t), t > 0}^{\text{conhecido.}}\}$ Como introduzir o mecanismo de seleção quando evento é recorrente e censurado por T ?
- Para unificar os ML e MR do ponto de vista do processo pontual $\{N(t), t > 0\}$ é fundamental introduzir a idéia de dados históricos ou condições de monitoramento (Engenharia), que chamaremos de **evento histórico** (Singpurwalla., 2007).

Problemas..

- Modelo latente (ML): $\{\overbrace{Y, N(t), t > 0}^{\text{latente.}}\}$ Como formular $w(t)$ e qual esquema de ativação deveremos utilizar para associar T com Y e $N(t)$.?
- Modelo recorrente (MR): $\{\overbrace{T, Y, N(t), t > 0}^{\text{conhecido.}}\}$ Como introduzir o mecanismo de seleção quando evento é recorrente e censurado por T ?
- Para unificar os ML e MR do ponto de vista do processo pontual $\{N(t), t > 0\}$ é fundamental introduzir a idéia de dados históricos ou condições de monitoramento (Engenharia), que chamaremos de **evento histórico** (Singpurwalla., 2007).

Problemas..

- Modelo latente (ML): $\overbrace{\{Y, N(t), t > 0\}}^{\text{latente.}}$ Como formular $w(t)$ e qual esquema de ativação deveremos utilizar para associar T com Y e $N(t)$.?
- Modelo recorrente (MR): $\overbrace{\{T, Y, N(t), t > 0\}}^{\text{conhecido.}}$ Como introduzir o mecanismo de seleção quando evento é recorrente e censurado por T ?
- Para unificar os ML e MR do ponto de vista do processo pontual $\{N(t), t > 0\}$ é fundamental introduzir a idéia de dados históricos ou condições de monitoramento (Engenharia), que chamaremos de **evento histórico** (Singpurwalla., 2007).

Evento histórico (Singpurwalla., 2007)

- **História do processo pontual $N(t)$:**
 $H(t) \Rightarrow$ informação acumulada até o instante t .
- Eventos históricos usuais:

$$H(t) = \{N(s), U(s), s \leq t\}, H(t) = \{U(t) = 1\}, U(t) = I_{\{T > t\}}$$

$$H(t) = \{N(s), U(s) = 1, X(s)\}, H(t) = \{N(s), s \leq t\},$$

$$H(t) = \{N(s), (Y_{(1)}, \dots, Y_{N(s)}), s \leq t\}$$

Evento histórico (Singpurwalla., 2007)

- **História do processo pontual $N(t)$:**
 $H(t) \Rightarrow$ informação acumulada até o instante t .
- **Eventos históricos usuais:**

$$H(t) = \{N(s), U(s), s \leq t\}, H(t) = \{U(t) = 1\}, U(t) = I_{[T > t]}$$

$$H(t) = \{N(s), U(s) = 1, X(s)\}, H(t) = \{N(s), s \leq t\},$$

$$H(t) = \{N(s), (Y_{(1)}, \dots, Y_{N(s)}), s \leq t\}$$

Função de intensidade do processo de contagem no instante t dado $H(t)$

- Função de intensidade do processo de contagem no instante t dado $H(t)$.

$$\begin{aligned}\lambda(t | H(t)) &= \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) = 1 | H(t)]}{\Delta t}, \\ &= \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t) | H(t)]}{\Delta t}.\end{aligned}$$

- Se $\lambda(t | H_t) = \lambda(t)$, temos que

$$\lambda(t) = \frac{dE[N(t)]}{dt} \rightarrow E[N(t)] = \int_0^t \lambda(u) du.$$

- Processo de renovação: $\lambda(t | H_t) = \lambda(t - T_{N(t)})$
- Modelo de Cox: $\lambda(t) = \lambda_0(t) \exp(\beta^t X(t)) \Rightarrow$
 “Martingale Theory of point processes” (Andersen, 1985).

Função de intensidade do processo de contagem no instante t dado $H(t)$

- Função de intensidade do processo de contagem no instante t dado $H(t)$.

$$\begin{aligned}\lambda(t | H(t)) &= \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) = 1 | H(t)]}{\Delta t}, \\ &= \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t) | H(t)]}{\Delta t}.\end{aligned}$$

- Se $\lambda(t | H_t) = \lambda(t)$, temos que

$$\lambda(t) = \frac{dE[N(t)]}{dt} \rightarrow E[N(t)] = \int_0^t \lambda(u) du.$$

- Processo de renovação: $\lambda(t | H_t) = \lambda(t - T_{N(t)})$
- Modelo de Cox: $\lambda(t) = \lambda_0(t) \exp(\beta^t X(t)) \Rightarrow$
 “Martingale Theory of point processes” (Andersen, 1985).

Função de intensidade do processo de contagem no instante t dado $H(t)$

- Função de intensidade do processo de contagem no instante t dado $H(t)$.

$$\begin{aligned}\lambda(t | H(t)) &= \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) = 1 | H(t)]}{\Delta t}, \\ &= \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t) | H(t)]}{\Delta t}.\end{aligned}$$

- Se $\lambda(t | H_t) = \lambda(t)$, temos que

$$\lambda(t) = \frac{dE[N(t)]}{dt} \rightarrow E[N(t)] = \int_0^t \lambda(u) du.$$

- Processo de renovação: $\lambda(t | H_t) = \lambda(t - T_{N(t)})$
- Modelo de Cox: $\lambda(t) = \lambda_0(t) \exp(\beta^t X(t)) \Rightarrow$
 “Martingale Theory of point processes” (Andersen, 1985).

Função de intensidade do processo de contagem no instante t dado $H(t)$

- Função de intensidade do processo de contagem no instante t dado $H(t)$.

$$\begin{aligned}\lambda(t | H(t)) &= \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) = 1 | H(t)]}{\Delta t}, \\ &= \lim_{\Delta t \rightarrow 0} \frac{E[N(t + \Delta t) - N(t) | H(t)]}{\Delta t}.\end{aligned}$$

- Se $\lambda(t | H_t) = \lambda(t)$, temos que

$$\lambda(t) = \frac{dE[N(t)]}{dt} \rightarrow E[N(t)] = \int_0^t \lambda(u) du.$$

- Processo de renovação: $\lambda(t | H_t) = \lambda(t - T_{N(t)})$
- Modelo de Cox: $\lambda(t) = \lambda_0(t) \exp(\beta^t X(t)) \Rightarrow$
 “Martingale Theory of point processes” (Andersen, 1985).

Função Geradora de probabilidades (FGP)(Feller, 1968)

- FGP de N :

$$A_N(s) = E[s^N], \quad 0 \leq s \leq 1.$$

- FGP de $N(t)$:

$$A_{N(t)}(s) = A_N(1 - p(t) + p(t)s), \quad 0 \leq s \leq 1.$$

- Probabilidade de não ocorrência do evento de interesse durante o intervalo $[0, t]$:

$$P[N(t) = 0] = A_N(1 - p(t))$$

- Se $p(t) = F(t) \Rightarrow \lim_{t \rightarrow \infty} P[N(t) = 0] = A_N(0) = p_0$.

Função Geradora de probabilidades (FGP)(Feller, 1968)

- FGP de N :

$$A_N(s) = E[s^N], \quad 0 \leq s \leq 1.$$

- FGP de $N(t)$:

$$A_{N(t)}(s) = A_N(1 - p(t) + p(t)s), \quad 0 \leq s \leq 1.$$

- Probabilidade de não ocorrência do evento de interesse durante o intervalo $[0, t]$:

$$P[N(t) = 0] = A_N(1 - p(t))$$

- Se $p(t) = F(t) \Rightarrow \lim_{t \rightarrow \infty} P[N(t) = 0] = A_N(0) = p_0$.

Função Geradora de probabilidades (FGP)(Feller, 1968)

- FGP de N :

$$A_N(s) = E[s^N], \quad 0 \leq s \leq 1.$$

- FGP de $N(t)$:

$$A_{N(t)}(s) = A_N(1 - p(t) + p(t)s), \quad 0 \leq s \leq 1.$$

- Probabilidade de não ocorrência do evento de interesse durante o intervalo $[0, t]$:

$$P[N(t) = 0] = A_N(1 - p(t))$$

- Se $p(t) = F(t) \Rightarrow \lim_{t \rightarrow \infty} P[N(t) = 0] = A_N(0) = p_0$.

Função Geradora de probabilidades (FGP)(Feller, 1968)

- FGP de N :

$$A_N(s) = E[s^N], \quad 0 \leq s \leq 1.$$

- FGP de $N(t)$:

$$A_{N(t)}(s) = A_N(1 - p(t) + p(t)s), \quad 0 \leq s \leq 1.$$

- Probabilidade de não ocorrência do evento de interesse durante o intervalo $[0, t]$:

$$P[N(t) = 0] = A_N(1 - p(t))$$

- Se $p(t) = F(t) \Rightarrow \lim_{t \rightarrow \infty} P[N(t) = 0] = A_N(0) = p_0$.

Outras questões importantes

- $H_t = \{T > t\}$ (TULD) : Como associar o tempo de vida T com as quantidades latentes $(Y_{(1)}, \dots, Y_{(N(t))})$? (Rodrigues *et al.*, 2008; Cooner *et al.*, 2007)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))}), T\}$ (Análise de Recorrência): Como modelar a dependência entre $N(t)$ e $(Y_{(1)}, \dots, Y_{(N(t))})$ com o tempo terminal T (censura informativa)? (Wang *et al.*, 2001)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))})\}$ (Análise de Recorrência) : Como modelar a dependência entre os eventos recorrentes? (“Shared Frailty Models” Wienke (2010))

Outras questões importantes

- $H_t = \{T > t\}$ (TULD) : Como associar o tempo de vida T com as quantidades latentes $(Y_{(1)}, \dots, Y_{(N(t))})$? (Rodrigues *et al.*, 2008; Cooner *et al.*, 2007)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))}), T\}$ (Análise de Recorrência): Como modelar a dependência entre $N(t)$ e $(Y_{(1)}, \dots, Y_{(N(t))})$ com o tempo terminal T (censura informativa)? (Wang *et al.*, 2001)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))})\}$ (Análise de Recorrência) : Como modelar a dependência entre os eventos recorrentes? (“Shared Frailty Models” Wienke (2010))

Outras questões importantes

- $H_t = \{T > t\}$ (TULD) : Como associar o tempo de vida T com as quantidades latentes $(Y_{(1)}, \dots, Y_{(N(t))})$? (Rodrigues *et al.*, 2008; Cooner *et al.*, 2007)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))}), T\}$ (Análise de Recorrência): Como modelar a dependência entre $N(t)$ e $(Y_{(1)}, \dots, Y_{(N(t))})$ com o tempo terminal T (censura informativa)? (Wang *et al.*, 2001)
- $H_t = \{N(t), (Y_{(1)}, \dots, Y_{(N(t))})\}$ (Análise de Recorrência) : Como modelar a dependência entre os eventos recorrentes? (“Shared Frailty Models” Wienke (2010))

Processo de Poisson não homogêneo (Feller, 1968)

- Processo de Poisson não homogêneo:

$$N \sim \text{Poisson}(\lambda) \Rightarrow N(t) \sim \text{Poisson}(M(t)),$$

$$M(t) = E[N(t)] = \lambda p(t) \Rightarrow \lambda(t) = \frac{dM(t)}{dt} = \lambda \frac{dp(t)}{dt},$$

ou seja $M(t) = \int_0^t \lambda(u) du$ (função de intensidade acumulada).

- Se $p(t) = tI_{[0,1]} \Rightarrow M(t) = \lambda t$: processo de Poisson homogêneo unitário com $\lambda(t) = \lambda I_{[0,1]}$.

Processo de Poisson não homogêneo (Feller, 1968)

- Processo de Poisson não homogêneo:

$$N \sim \text{Poisson}(\lambda) \Rightarrow N(t) \sim \text{Poisson}(M(t)),$$

$$M(t) = E[N(t)] = \lambda p(t) \Rightarrow \lambda(t) = \frac{dM(t)}{dt} = \lambda \frac{dp(t)}{dt},$$

ou seja $M(t) = \int_0^t \lambda(u) du$ (função de intensidade acumulada).

- Se $p(t) = tI_{[0,1]} \Rightarrow M(t) = \lambda t$: processo de Poisson homogêneo unitário com $\lambda(t) = \lambda I_{[0,1]}$.

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo (“Multivariate failure time data”(Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo (“Multivariate failure time data”)(Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo ("Multivariate failure time data" (Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo (“Multivariate failure time data”(Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo ("Multivariate failure time data" (Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo ("Multivariate failure time data" (Lim *et al.*, 2007).

Análise de recorrência, (Cook and Lawless, 2002).

- Processo de Poisson: $N \sim \text{Poisson}(g(X, \beta))$: número de tumores no intervalo $[0, \tau_0]$, $X^t = (X_1, \dots, X_p)$, $\beta^t = (\beta_1, \dots, \beta_p)$.
- Evento histórico: $H(t) = \{N(t) = n, (Y_{(1)}, \dots, Y_{(n)}), X\}$.
- $p(t) = F(t)$
- Modelo de Cox: $\lambda(t) = f(t)g(X, \beta)$
- Objetivo: Estimar (procedimento semiparamétrico)

$$M(t) = E[N(t)] = F(t)g(X, \beta).$$

- Pesquisa: Introduzir probabilidade de cura e modelar a dependência entre as ocorrências de tumores e o tempo terminal via fragilidade (censura informativa) (Rondeau *et al.*, 2011).
- AR: Os eventos são naturalmente ordenados e ocorrem sequencialmente no tempo (“Multivariate failure time data”)(Lim *et al.*, 2007).

Análise de Sobrevida

$$(\tau_0 = \infty, P[N = 1] = 1 - p_0, U \perp Y \Rightarrow T = Y)$$

- Processo com dois estados (vivo ou morto):

$$N(t) = \begin{cases} 1 & T \leq t \text{ (não está vivo até o instante } t\text{),} \\ 0 & T \geq t \text{ (está vivo até o instante } t\text{),} \end{cases} \quad (5)$$

absorvente

- Tempo de vida : $T =$ Como é desconhecida para $t = 0 \Rightarrow T$ é uma v. aleatória. Como $T = Y \Rightarrow T \sim F(t)$.
- $p_n(t) = F(t), n = 0, 1$
- Evento Histórico:
 $H(t) = [T \geq t] \Rightarrow S(t) = P[T \geq t]$ (Função de Sobrevida).

Análise de Sobrevida

$$(\tau_0 = \infty, P[N = 1] = 1 - p_0, U \perp Y \Rightarrow T = Y)$$

- Processo com dois estados (vivo ou morto):

$$N(t) = \begin{cases} 1 & T \leq t \text{ (não está vivo até o instante } t\text{),} \\ 0 & T \geq t \text{ (está vivo até o instante } t\text{),} \end{cases} \quad (5)$$

absorvente

- Tempo de vida : $T =$ Como é desconhecida para $t = 0 \Rightarrow T$ é uma v. aleatória. Como $T = Y \Rightarrow T \sim F(t)$.
- $p_n(t) = F(t), n = 0, 1$
- Evento Histórico:
 $H(t) = [T \geq t] \Rightarrow S(t) = P[T \geq t]$ (Função de Sobrevida).

Análise de Sobrevida

$$(\tau_0 = \infty, P[N = 1] = 1 - p_0, U \perp Y \Rightarrow T = Y)$$

- Processo com dois estados (vivo ou morto):

$$N(t) = \begin{cases} 1 & T \leq t \text{ (não está vivo até o instante } t\text{),} \\ 0 & T \geq t \text{ (está vivo até o instante } t\text{),} \end{cases} \quad (5)$$

absorvente

- Tempo de vida : $T =$ Como é desconhecida para $t = 0 \Rightarrow T$ é uma v. aleatória. Como $T = Y \Rightarrow T \sim F(t)$.
- $p_n(t) = F(t), n = 0, 1$
- Evento Histórico:
 $H(t) = [T \geq t] \Rightarrow S(t) = P[T \geq t]$ (Função de Sobrevida).

Análise de Sobrevida

$$(\tau_0 = \infty, P[N = 1] = 1 - p_0, U \perp Y \Rightarrow T = Y)$$

- Processo com dois estados (vivo ou morto):

$$N(t) = \begin{cases} 1 & T \leq t \text{ (não está vivo até o instante } t\text{),} \\ 0 & T \geq t \text{ (está vivo até o instante } t\text{),} \end{cases} \quad (5)$$

absorvente

- Tempo de vida : $T =$ Como é desconhecida para $t = 0 \Rightarrow T$ é uma v. aleatória. Como $T = Y \Rightarrow T \sim F(t)$.
- $p_n(t) = F(t), n = 0, 1$
- Evento Histórico:
 $H(t) = [T \geq t] \Rightarrow S(t) = P[T \geq t]$ (Função de Sobrevida).

Continuação...

- Função de risco (prob. de transição instantânea no instante t): A função de risco é uma ferramenta essencial à análise de dados em AS.

$$\lambda(t) = h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t \mid H(t)]}{\Delta t}. \quad (6)$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (7)$$

- Processo simples de sobrevivência (dois estados sendo um absorvente)



Continuação...

- Função de risco (prob. de transição instantânea no instante t): A função de risco é uma ferramenta essencial à análise de dados em AS.

$$\lambda(t) = h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t \mid H(t)]}{\Delta t}. \quad (6)$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (7)$$

- Processo simples de sobrevivência (dois estados sendo um absorvente)



Continuação...

- Função de risco (prob. de transição instantânea no instante t): A função de risco é uma ferramenta essencial à análise de dados em AS.

$$\lambda(t) = h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t \mid H(t)]}{\Delta t}. \quad (6)$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (7)$$

- Processo simples de sobrevivência (dois estados sendo um absorvente)

$h(t)$: prob. de transição

O: estado vivo \rightleftharpoons 1: estado morto

Função verossimilhança com censura aleatória

Censura aleatória. A i -ésima unidade pode sofrer censura em um instante aleatório L_i com função densidade $f_R(\cdot)$ e FS denotada por $S_R(\cdot)$, $i = 1, \dots, n$. Neste cenário os dados são representados por (t_i, δ_i) , em que

$$\delta_i = \begin{cases} 1, & \text{se } t_i \leq L_i, \\ 0, & \text{se } t_i > L_i, \end{cases}$$

$i = 1, \dots, n$.

Continuação....

Teorema (1.4)

A função densidade conjunta de (t_i, δ_i) , $i = 1, \dots, n$, em AS com censura aleatória tem expressão

$$\prod_{i=1}^n \{f(t_i)S_R(t_i)\}^{\delta_i} \{f_R(t_i)S(t_i)\}^{1-\delta_i}.$$

Função verossimilhança com censura aleatória não-informativa

Qualquer que seja o tipo de censura, supondo que $f_R(\cdot)$ não depende de ϑ (censura não-informativa), a função densidade conjunta do par (t_i, δ_i) assume expressão proporcional a

$$\{f(t_i)\}^{\delta_i} \{S(t_i)\}^{1-\delta_i},$$

$i = 1, \dots, n$, sendo que para as censuras tipo I, II e aleatória, se $\delta_i = 0$, tomamos $S(t_i) = S(L)$, $S(t_i) = S(t_{(d)})$ e $S(t_i) = S(L_i)$, respectivamente.

Continuação.....

- Função de verossimilhança com censura não-informativa:

$$\mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}}) = \prod_{i=1}^n \{f(t_i; \vartheta)\}^{\delta_i} \{S(t_i; \vartheta)\}^{1-\delta_i}.$$

- Estimação pontual (EMV): $\hat{\vartheta}$
- Intervalos de confiança: Censura II (exato) e Censura I (assintótico).
- Testes de hipóteses: Razão de verossimilhança (assintótico).

Continuação.....

- Função de verossimilhança com censura não-informativa:

$$\mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}}) = \prod_{i=1}^n \{f(t_i; \vartheta)\}^{\delta_i} \{S(t_i; \vartheta)\}^{1-\delta_i}.$$

- Estimação pontual (EMV): $\hat{\vartheta}$
- Intervalos de confiança: Censura II (exato) e Censura I (assintótico).
- Testes de hipóteses: Razão de verossimilhança (assintótico).

Continuação.....

- Função de verossimilhança com censura não-informativa:

$$\mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}}) = \prod_{i=1}^n \{f(t_i; \vartheta)\}^{\delta_i} \{S(t_i; \vartheta)\}^{1-\delta_i}.$$

- Estimaco pontual (EMV): $\hat{\vartheta}$
- Intervalos de confiana: Censura II (exato) e Censura I (assinttico).
- Testes de hipteses: Razo de verossimilhana (assinttico).

Continuação.....

- Função de verossimilhança com censura não-informativa:

$$\mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}}) = \prod_{i=1}^n \{f(t_i; \vartheta)\}^{\delta_i} \{S(t_i; \vartheta)\}^{1-\delta_i}.$$

- Estimação pontual (EMV): $\hat{\vartheta}$
- Intervalos de confiança: Censura II (exato) e Censura I (assintótico).
- Testes de hipóteses: Razão de verossimilhança (assintótico).

Abordagem Bayesiana: Distribuição *a posteriori*

Sob o princípio bayesiano, ao vetor ϑ de dimensão q^* atribuímos uma distribuição de probabilidade *a priori* com função densidade $\pi(\vartheta)$. As inferências sobre ϑ estão baseadas na distribuição *a posteriori* de ϑ , cuja função densidade é obtida pelo teorema de Bayes, conduzindo a

$$\pi(\vartheta | \mathbf{D}_{\text{obs}}) = \frac{\mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}})\pi(\vartheta)}{\int_{\Theta} \mathcal{L}(\vartheta; \mathbf{D}_{\text{obs}})\pi(\vartheta)d\vartheta}.$$

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - ...
 - Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - ...
 - Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - 3 ...
 - 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - 3 ...
 - 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - 3 ...
 - 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\vartheta^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\vartheta^{(j)}$ a partir de $\vartheta^{(j-1)}$ percorrendo os seguintes passos:
 - 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \vartheta_{-1}^{(j-1)})$;
 - 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \vartheta_{-2}^{(j-1)})$;
 - 3 ...
 - 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \vartheta_{-q^*}^{(j-1)})$,

em que $\vartheta_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \vartheta_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\vartheta_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\boldsymbol{\vartheta}^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\boldsymbol{\vartheta}^{(j)}$ a partir de $\boldsymbol{\vartheta}^{(j-1)}$ percorrendo os seguintes passos:
 - 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \boldsymbol{\vartheta}_{-1}^{(j-1)})$;
 - 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \boldsymbol{\vartheta}_{-2}^{(j-1)})$;
 - 3 ...
 - 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \boldsymbol{\vartheta}_{-q^*}^{(j-1)})$,

em que $\boldsymbol{\vartheta}_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \boldsymbol{\vartheta}_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\boldsymbol{\vartheta}_{-i}^{(j-1)}$.

Amostrador de Gibbs

- Atribuir valores iniciais arbitrários ao vetor paramétrico de dimensão q^* , ou seja, $\boldsymbol{\vartheta}^{(0)} = (\vartheta_1^{(0)}, \dots, \vartheta_{q^*}^{(0)})^\top$;
- Na j -ésima iteração do algoritmo, gerar $\boldsymbol{\vartheta}^{(j)}$ a partir de $\boldsymbol{\vartheta}^{(j-1)}$ percorrendo os seguintes passos:

- 1 Gerar $\vartheta_1^{(j)}$ de $p(\vartheta_1 | \boldsymbol{\vartheta}_{-1}^{(j-1)})$;
- 2 Gerar $\vartheta_2^{(j)}$ de $p(\vartheta_2 | \boldsymbol{\vartheta}_{-2}^{(j-1)})$;
- 3 ...
- 4 Gerar $\vartheta_{q^*}^{(j)}$ de $p(\vartheta_{q^*} | \boldsymbol{\vartheta}_{-q^*}^{(j-1)})$,

em que $\boldsymbol{\vartheta}_{-i}^{(j-1)} = (\vartheta_1^{(j-1)}, \dots, \vartheta_{i-1}^{(j-1)}, \vartheta_{i+1}^{(j-1)}, \dots, \vartheta_{q^*}^{(j-1)})^\top$ e $p(\vartheta_i^{(j)} | \boldsymbol{\vartheta}_{-i}^{(j-1)})$ é a densidade condicional (conhecida) de $\vartheta_i^{(j)}$ dado $\boldsymbol{\vartheta}_{-i}^{(j-1)}$.

Regressão exponencial

- Devido a sua importância histórica, simplicidade e aplicabilidade, daremos ênfase ao modelo exponencial.
- Supondo $p_0 = 0$ temos que $S_{\text{pop}}(t) = S(t)$, $f_{\text{pop}}(t) = f(t)$,
 $h_{\text{pop}}(t) = h(t) = \frac{f(t)}{S(t)}$.
- Nesta seção desenvolveremos um procedimento estatístico integrador dos métodos clássicos e bayesianos. O procedimento de MV do ponto de vista clássico será baseado nos textos de Lawless (2002) e Colosimo and Giolo (2006), em que o parâmetro de interesse é $\alpha^* = E[T]$.
- O procedimento bayesiano será apoiado em Ibrahim *et al.* (2001), em que o parâmetro de interesse é $\lambda^* = \frac{1}{\alpha^*}$ (função de risco constante).

Regressão exponencial

- Devido a sua importância histórica, simplicidade e aplicabilidade, daremos ênfase ao modelo exponencial.
- Supondo $p_0 = 0$ temos que $S_{\text{pop}}(t) = S(t)$, $f_{\text{pop}}(t) = f(t)$,
 $h_{\text{pop}}(t) = h(t) = \frac{f(t)}{S(t)}$.
- Nesta seção desenvolveremos um procedimento estatístico integrador dos métodos clássicos e bayesianos. O procedimento de MV do ponto de vista clássico será baseado nos textos de Lawless (2002) e Colosimo and Giolo (2006), em que o parâmetro de interesse é $\alpha^* = E[T]$.
- O procedimento bayesiano será apoiado em Ibrahim *et al.* (2001), em que o parâmetro de interesse é $\lambda^* = \frac{1}{\alpha^*}$ (função de risco constante).

Regressão exponencial

- Devido a sua importância histórica, simplicidade e aplicabilidade, daremos ênfase ao modelo exponencial.
- Supondo $p_0 = 0$ temos que $S_{\text{pop}}(t) = S(t)$, $f_{\text{pop}}(t) = f(t)$,
 $h_{\text{pop}}(t) = h(t) = \frac{f(t)}{S(t)}$.
- Nesta seção desenvolveremos um procedimento estatístico integrador dos métodos clássicos e bayesianos. O procedimento de MV do ponto de vista clássico será baseado nos textos de Lawless (2002) e Colosimo and Giolo (2006), em que o parâmetro de interesse é $\alpha^* = E[T]$.
- O procedimento bayesiano será apoiado em Ibrahim *et al.* (2001), em que o parâmetro de interesse é $\lambda^* = \frac{1}{\alpha^*}$ (função de risco constante).

Regressão exponencial

- Devido a sua importância histórica, simplicidade e aplicabilidade, daremos ênfase ao modelo exponencial.
- Supondo $p_0 = 0$ temos que $S_{\text{pop}}(t) = S(t)$, $f_{\text{pop}}(t) = f(t)$,
 $h_{\text{pop}}(t) = h(t) = \frac{f(t)}{S(t)}$.
- Nesta seção desenvolveremos um procedimento estatístico integrador dos métodos clássicos e bayesianos. O procedimento de MV do ponto de vista clássico será baseado nos textos de Lawless (2002) e Colosimo and Giolo (2006), em que o parâmetro de interesse é $\alpha^* = E[T]$.
- O procedimento bayesiano será apoiado em Ibrahim *et al.* (2001), em que o parâmetro de interesse é $\lambda^* = \frac{1}{\alpha^*}$ (função de risco constante).

Função verossimilhança no caso clássico

$$\begin{aligned}\mathcal{L}(\alpha^*; \mathbf{D}_{\text{obs}}) &= \prod_{i=1}^n \left\{ \frac{1}{\alpha^*} \exp\left(-\frac{t_i}{\alpha^*}\right) \right\}^{\delta_i} \left\{ \exp\left(-\frac{1}{\alpha^*} t_i\right) \right\}^{1-\delta_i} \\ &= \prod_{i=1}^n \left(\frac{1}{\alpha^*}\right)^{\delta_i} \exp\left(-\frac{t_i}{\alpha^*}\right) = \left(\frac{1}{\alpha^*}\right)^d \exp\left(-\frac{1}{\alpha^*} \sum_{i=1}^n t_i\right),\end{aligned}$$

em que $d = \sum_{i=1}^n \delta_i$.

EMV para diferentes mecanismos de censura

$$\widehat{\alpha}^* = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n \delta_i} = \frac{\sum_{i=1}^n \delta_i t_i + \sum_{i=1}^n (1 - \delta_i) t_i}{\sum_{i=1}^n \delta_i} = \frac{V}{d}. \quad (8)$$

- Censura tipo II (d fixo):

$$\widehat{\alpha}^* = \frac{\sum_{i=1}^d t_{(i)} + (n - d)t_{(d)}}{d}.$$

EMV para diferentes mecanismos de censura

$$\widehat{\alpha}^* = \frac{\sum_{i=1}^n t_i}{\sum_{i=1}^n \delta_i} = \frac{\sum_{i=1}^n \delta_i t_i + \sum_{i=1}^n (1 - \delta_i) t_i}{\sum_{i=1}^n \delta_i} = \frac{V}{d}. \quad (8)$$

- Censura tipo II (d fixo):

$$\widehat{\alpha}^* = \frac{\sum_{i=1}^d t_{(i)} + (n - d)t_{(d)}}{d}.$$

EMV para diferentes mecanismos de censura

- Censura tipo I (d aleatório):

$$\widehat{\alpha}^* = \frac{\sum_{i \in \bar{C}} t_i + (n - d)L}{d}.$$

- Censura aleatória (d aleatório):

$$\widehat{\alpha}^* = \frac{\sum_{i \in \bar{C}} t_i + \sum_{i \in C} L_i}{d}.$$

EMV para diferentes mecanismos de censura

- Censura tipo I (d aleatório):

$$\widehat{\alpha}^* = \frac{\sum_{i \in \bar{C}} t_i + (n - d)L}{d}.$$

- Censura aleatória (d aleatório):

$$\widehat{\alpha}^* = \frac{\sum_{i \in \bar{C}} t_i + \sum_{i \in C} L_i}{d}.$$

Função verossimilhança

A função verossimilhança para o caso bayesiano tem expressão

$$\mathcal{L}(\lambda^*; \mathbf{D}_{\text{obs}}) = \lambda^{*d} \exp \left(-\lambda^* \sum_{i=1}^n t_i \right), \quad (9)$$

com

$$\lambda^* = \frac{1}{\alpha^*}.$$

Distribuições *a priori* e *a posteriori* para λ^*

- *A priori* conjugada:

$$\pi(\lambda^*) \propto \lambda^{*\alpha_0^*-1} \exp\{-\lambda_0^*\lambda^*\}.$$

- *A posteriori* de λ^* :

$$\pi(\lambda^* | D_{\text{obs}}) \propto \lambda^{*d+\alpha_0^*-1} \exp\left\{-\lambda^*\left(\lambda_0^* + \sum_{i=1}^n t_i\right)\right\},$$

que identificamos como sendo uma distribuição gama com parâmetros $d + \alpha_0^*$ e $\lambda_0^* + \sum_{i=1}^n t_i$.

Distribuições *a priori* e *a posteriori* para λ^*

- *A priori* conjugada:

$$\pi(\lambda^*) \propto \lambda^{*\alpha_0^* - 1} \exp\{-\lambda_0^* \lambda^*\}.$$

- *A posteriori* de λ^* :

$$\pi(\lambda^* | D_{\text{obs}}) \propto \lambda^{*d + \alpha_0^* - 1} \exp\left\{-\lambda^* \left(\lambda_0^* + \sum_{i=1}^n t_i\right)\right\},$$

que identificamos como sendo uma distribuição gama com parâmetros $d + \alpha_0^*$ e $\lambda_0^* + \sum_{i=1}^n t_i$.

Esperança *a posteriori* de λ^*

$$E[\lambda^* | \mathbf{D}_{\text{obs}}] = \frac{d + \alpha_0^*}{\lambda_0^* + \sum_{i=1}^n t_i}, \quad \text{Var}[\lambda^* | \mathbf{D}_{\text{obs}}] = \frac{d + \alpha_0^*}{(\lambda_0^* + \sum_{i=1}^n t_i)^2}.$$

$$\pi(\lambda^*) \propto \frac{1}{\lambda^*} \implies E[\lambda^* | \mathbf{D}_{\text{obs}}] = \frac{d}{\underbrace{\sum_{i=1}^n t_i}_{\text{MV}}}.$$

Distribuições *a priori* e *a posteriori* para $\alpha^* = \frac{1}{\lambda^*}$

- *A priori* conjugada:

$$\pi(\alpha^*) \propto \alpha^{*-(\alpha_0^*+1)} \exp\{-\lambda_0^*/\alpha^*\}.$$

- Distribuição *a posteriori* de α^* :

$$\pi(\alpha^* | D_{\text{obs}}) \propto \alpha^{*-(d+\alpha_0^*+1)} \exp\left\{-\left(\lambda_0^* + \sum_{i=1}^n t_i\right)/\alpha^*\right\},$$

que identificamos como sendo uma distribuição gama inversa com parâmetros $d + \alpha_0^*$ e $\lambda_0^* + \sum_{i=1}^n t_i$.

Distribuições *a priori* e *a posteriori* para $\alpha^* = \frac{1}{\lambda^*}$

- *A priori* conjugada:

$$\pi(\alpha^*) \propto \alpha^{*-(\alpha_0^*+1)} \exp\{-\lambda_0^*/\alpha^*\}.$$

- Distribuição *a posteriori* de α^* :

$$\pi(\alpha^* | D_{\text{obs}}) \propto \alpha^{*-(d+\alpha_0^*+1)} \exp\left\{-\left(\lambda_0^* + \sum_{i=1}^n t_i\right)/\alpha^*\right\},$$

que identificamos como sendo uma distribuição gama inversa com parâmetros $d + \alpha_0^*$ e $\lambda_0^* + \sum_{i=1}^n t_i$.

Esperança *a posteriori* de α_*

$$\begin{aligned} E[\alpha^* | \mathbf{D}_{\text{obs}}] &= \frac{\lambda_0^* + \sum_{i=1}^n t_i}{d + \alpha_0^* - 1} \\ &= \left(\frac{\alpha_0^* - 1}{d + \alpha_0^* - 1} \right) \frac{\lambda_0^*}{\alpha_0^* - 1} + \left(\frac{d}{d + \alpha_0^* - 1} \right) \frac{\sum_{i=1}^n t_i}{d}, \end{aligned}$$

$$d + \alpha_0^* > 1.$$

Densidade preditiva *a posteriori* de t_f (tempo de falha futuro)

$$\begin{aligned} \pi(t_f | \mathbf{D}_{\text{obs}}) &= \int_0^{\infty} \lambda^* \exp(-\lambda^* t_f) \lambda^{*d+\alpha_0^*-1} \exp\left\{-\lambda^*(\lambda_0^* + \sum_{i=1}^n t_i)\right\} d\lambda^* \\ &= \frac{(d + \alpha_0^*)(\lambda_0^* + \sum_{i=1}^n t_i)^{d+\alpha_0^*}}{(\lambda_0^* + \sum_{i=1}^n t_i + t_f)^{d+\alpha_0^*+1}}, \end{aligned}$$

se $t_f > 0$ e igual a 0, caso contrário. Esta função densidade é conhecida como beta inversa (Aitchison and Dunsmore, 1975).

Também pode ser mostrado que

$$E[t_f | \mathbf{D}_{\text{obs}}] = \frac{\lambda_0^* + \sum_{i=1}^n t_i}{d + \alpha_0^* - 1}$$

e

$$\text{Var}[t_f | \mathbf{D}_{\text{obs}}] = \frac{(\lambda_0^* + \sum_{i=1}^n t_i)^2 (d + \alpha_0^*)}{(d + \alpha_0^* - 2)^2 (d + \alpha_0^* - 1)},$$

$$d + \alpha_0^* - 2 > 0.$$

Um intervalo de credibilidade de $100\gamma\%$ para t_f

$$LI = \left(\lambda_0^* + \sum_{i=1}^n t_i \right) \left\{ \left(\frac{2}{1+\gamma} \right)^{1/(d+\alpha_0^*)} - 1 \right\} \quad e$$

$$LS = \left(\lambda_0^* + \sum_{i=1}^n t_i \right) \left\{ \left(\frac{2}{1-\gamma} \right)^{1/(d+\alpha_0^*)} - 1 \right\}.$$

A priori de Jeffreys para λ^*

$$\pi(\lambda^*) \propto \frac{1}{\lambda^*} \sqrt{\frac{\sum_{i=1}^n \{1 - \exp(-\lambda^* L_i)\}}{n}} \leq \frac{1}{\lambda^*} .$$

Com censura

Sem censura

$$\alpha_0^* = \lambda_0^* = 0 :$$

$$\pi(\lambda^*) \propto \frac{1}{\lambda^*} \implies E[\lambda^* | \mathbf{D}_{\text{obs}}] = \frac{d}{\underbrace{\sum_{i=1}^n t_i}_{\text{MV}}} .$$

Exemplo (2.1)

Dados referentes aos tempos de reincidência T (em meses) de 20 pacientes com câncer de bexiga que passaram por um procedimento cirúrgico, havendo três observações censuradas (Colosimo and Giolo, 2006, p. 101). Adotamos uma distribuição a priori de Jeffreys para λ^ :*

$$\pi(\lambda^*) \propto 1/\lambda^*.$$

Medidas de posição e dispersão da distribuição a posteriori de $\alpha^ = 1/\lambda^*$ são apresentadas na próxima tabela.*

Tabela: Medidas de posição e dispersão das distribuições *a posteriori* da função de risco e de um tempo de falha futuro no modelo exponencial – Exemplo 2.1.

Parâmetro	Média	Mediana	Desvio padrão	Intervalo de credibilidade de 95%
α^*	21,688	20,818	1,446	(13,355; 35,040)
t_f	21,688	14,441	23,845	(0,517; 84,090)

Teoria Unificada de Longa Duração

- Condições iniciais: $N(\tau_0) = N \rightarrow M(\tau_0) = E[N]$
- Distribuição de probabilidades: $p_n = P[N = n], n = 0, 1, \dots, \dots,$
- Taxa de cura: $p_0 = P[N = 0]$
- $p(t) = 1, \forall t > 0 \Rightarrow N(t) = N(\tau_0)$
- $\tau_0 = \infty$.

Teoria Unificada de Longa Duração

- Condições iniciais: $N(\tau_0) = N \rightarrow M(\tau_0) = E[N]$
- Distribuição de probabilidades: $p_n = P[N = n], n = 0, 1, \dots, \dots,$
- Taxa de cura: $p_0 = P[N = 0]$
- $p(t) = 1, \forall t > 0 \Rightarrow N(t) = N(\tau_0)$
- $\tau_0 = \infty$.

Teoria Unificada de Longa Duração

- Condições iniciais: $N(\tau_0) = N \rightarrow M(\tau_0) = E[N]$
- Distribuição de probabilidades: $p_n = P[N = n], n = 0, 1, \dots, \dots,$
- Taxa de cura: $p_0 = P[N = 0]$
- $p(t) = 1, \forall t > 0 \Rightarrow N(t) = N(\tau_0)$
- $\tau_0 = \infty$.

Teoria Unificada de Longa Duração

- Condições iniciais: $N(\tau_0) = N \rightarrow M(\tau_0) = E[N]$
- Distribuição de probabilidades: $p_n = P[N = n], n = 0, 1, \dots, \dots,$
- Taxa de cura: $p_0 = P[N = 0]$
- $p(t) = 1, \forall t > 0 \Rightarrow N(t) = N(\tau_0)$
- $\tau_0 = \infty$.

Teoria Unificada de Longa Duração

- Condições iniciais: $N(\tau_0) = N \rightarrow M(\tau_0) = E[N]$
- Distribuição de probabilidades: $p_n = P[N = n], n = 0, 1, \dots, \dots,$
- Taxa de cura: $p_0 = P[N = 0]$
- $p(t) = 1, \forall t > 0 \Rightarrow N(t) = N(\tau_0)$
- $\tau_0 = \infty$.

O que é um modelo de longa duração?

- Com o rápido desenvolvimento dos tratamentos médicos, em geral os dados revelam que uma parte substancial dos pacientes podem ser de curados.
- Os modelos tradicionais em Análise de Sobrevida não incluem esta possibilidade. Os pacientes estão em risco durante o tratamento com probabilidade um.
- Os modelos que incluem esta possibilidade são conhecidos na literatura como modelos de longa duração (MLD).
- Áreas de aplicação dos MLD (Ibrahim *et al.*, 2001; Hoggart and Griffin, 2001; Rodrigues *et al.*, 2008): médica, financeira (Análise de Risco, "Customer Attrition"), testes acelerados, prevenção de câncer, etc.

O que é um modelo de longa duração?

- Com o rápido desenvolvimento dos tratamentos médicos, em geral os dados revelam que uma parte substancial dos pacientes podem ser de curados.
- Os modelos tradicionais em Análise de Sobrevida não incluem esta possibilidade. Os pacientes estão em risco durante o tratamento com probabilidade um.
- Os modelos que incluem esta possibilidade são conhecidos na literatura como modelos de longa duração (MLD).
- Áreas de aplicação dos MLD (Ibrahim *et al.*, 2001; Hoggart and Griffin, 2001; Rodrigues *et al.*, 2008): médica, financeira (Análise de Risco, "Customer Attrition"), testes acelerados, prevenção de câncer, etc.

O que é um modelo de longa duração?

- Com o rápido desenvolvimento dos tratamentos médicos, em geral os dados revelam que uma parte substancial dos pacientes podem ser de curados.
- Os modelos tradicionais em Análise de Sobrevida não incluem esta possibilidade. Os pacientes estão em risco durante o tratamento com probabilidade um.
- Os modelos que incluem esta possibilidade são conhecidos na literatura como modelos de longa duração (MLD).
- Áreas de aplicação dos MLD (Ibrahim *et al.*, 2001; Hoggart and Griffin, 2001; Rodrigues *et al.*, 2008): médica, financeira (Análise de Risco, "Customer Attrition"), testes acelerados, prevenção de câncer, etc.

O que é um modelo de longa duração?

- Com o rápido desenvolvimento dos tratamentos médicos, em geral os dados revelam que uma parte substancial dos pacientes podem ser de curados.
- Os modelos tradicionais em Análise de Sobrevida não incluem esta possibilidade. Os pacientes estão em risco durante o tratamento com probabilidade um.
- Os modelos que incluem esta possibilidade são conhecidos na literatura como modelos de longa duração (MLD).
- Áreas de aplicação dos MLD (Ibrahim *et al.*, 2001; Hoggart and Griffin, 2001; Rodrigues *et al.*, 2008): médica, financeira (Análise de Risco, "Customer Attrition"), testes acelerados, prevenção de câncer, etc.

O que é um modelo de longa duração?

- Com o rápido desenvolvimento dos tratamentos médicos, em geral os dados revelam que uma parte substancial dos pacientes podem ser de curados.
- Os modelos tradicionais em Análise de Sobrevida não incluem esta possibilidade. Os pacientes estão em risco durante o tratamento com probabilidade um.
- Os modelos que incluem esta possibilidade são conhecidos na literatura como modelos de longa duração (MLD).
- Áreas de aplicação dos MLD (Ibrahim *et al.*, 2001; Hoggart and Griffin, 2001; Rodrigues *et al.*, 2008): médica, financeira (Análise de Risco, "Customer Attrition"), testes acelerados, prevenção de câncer, etc.

Continuação.....

- **AS:** Assume-se que todas as unidades experimentais ("indivíduos") apresentam o evento de interesse.
- **ASLD:** Considera a possibilidade de indivíduos sobreviventes por longo período, proporção de curados (Boag(1949), Berkson e Gage(1952) - dados de câncer de mama).

Algumas referências importantes

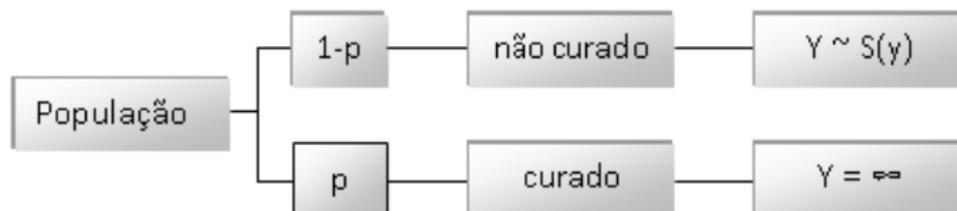
- Boag (1949),
- Berkson and Gage (1952),
- Maller and Zhou (1996),
- Yakovlev and Tsodikov (1996),
- Chen *et al.* (1999),
- Hoggart and Griffin (2001),
- Ibrahim *et al.* (2001),
- Tsodikov *et al.* (2003),
- Rodrigues *et al.* (2011a),
- Cooner *et al.* (2007) e
- Rodrigues *et al.* (2008).
- de Castro *et al.* (2009); Cancho *et al.* (2012)

Modelo de mistura

- Berkson & Gage (1952) introduziu o modelo de mistura padrão. Neste modelo é assumido que uma certa fração p da população é "curado", e o restante $1 - p$ são não curados. A função de sobrevivência para a população ($S_p(y)$) é dada por

$$S_p(y) = p + (1 - p)S(y)$$

$S(y)$: Função de sobrevivência própria (grupo dos não curados)



Especificando Tempo de Sobrevivência

Seja $T \in \mathfrak{R}^+$ uma variável aleatória que denota o tempo de sobrevivência. A Função de distribuição de T pode ser escrita como:

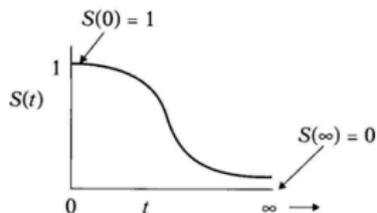
$$F(t) = P(T \leq t) = \int_0^t f(u) du$$

em que f é a *f.d.p* de T .

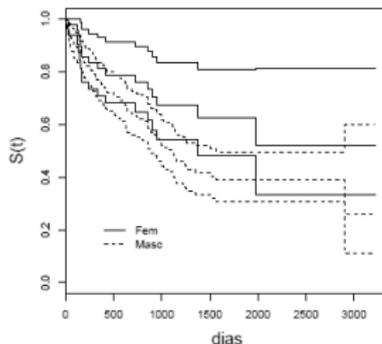
Define-se a Função de Sobrevivência, $S(t)$, como a probabilidade de um individuo sobreviver a um tempo maior do que t , ou seja,

$$S(t) = 1 - F(t)$$

Forma da Função de Sobrevivência



Quando $\lim_{t \rightarrow \infty} S(t) \neq 0$



Desvantagens do modelo de mistura padrão

- **Freqüentista.** Na presença de covariáveis não satisfaz à propriedade de riscos proporcionais.
- **Bayesiano.** Na presença de covariáveis as distribuições *a priori* não-informativas impróprias em geral implicam em distribuições *a posteriori* impróprias.
- **Interpretação biológica.** Não descreve o mecanismo biológico que gera o tempo de ocorrência do evento de interesse.

Modelo alternativo: modelo de tempo de promoção (ou ativação) (Yakovlev and Tsodikov, 1996; Chen *et al.*, 1999).

Desvantagens do modelo de mistura padrão

- **Freqüentista.** Na presença de covariáveis não satisfaz à propriedade de riscos proporcionais.
- **Bayesiano.** Na presença de covariáveis as distribuições *a priori* não-informativas impróprias em geral implicam em distribuições *a posteriori* impróprias.
- **Interpretação biológica.** Não descreve o mecanismo biológico que gera o tempo de ocorrência do evento de interesse.

Modelo alternativo: modelo de tempo de promoção (ou ativação) (Yakovlev and Tsodikov, 1996; Chen *et al.*, 1999).

Desvantagens do modelo de mistura padrão

- **Freqüentista.** Na presença de covariáveis não satisfaz à propriedade de riscos proporcionais.
- **Bayesiano.** Na presença de covariáveis as distribuições *a priori* não-informativas impróprias em geral implicam em distribuições *a posteriori* impróprias.
- **Interpretação biológica.** Não descreve o mecanismo biológico que gera o tempo de ocorrência do evento de interesse.

Modelo alternativo: modelo de tempo de promoção (ou ativação) (Yakovlev and Tsodikov, 1996; Chen *et al.*, 1999).

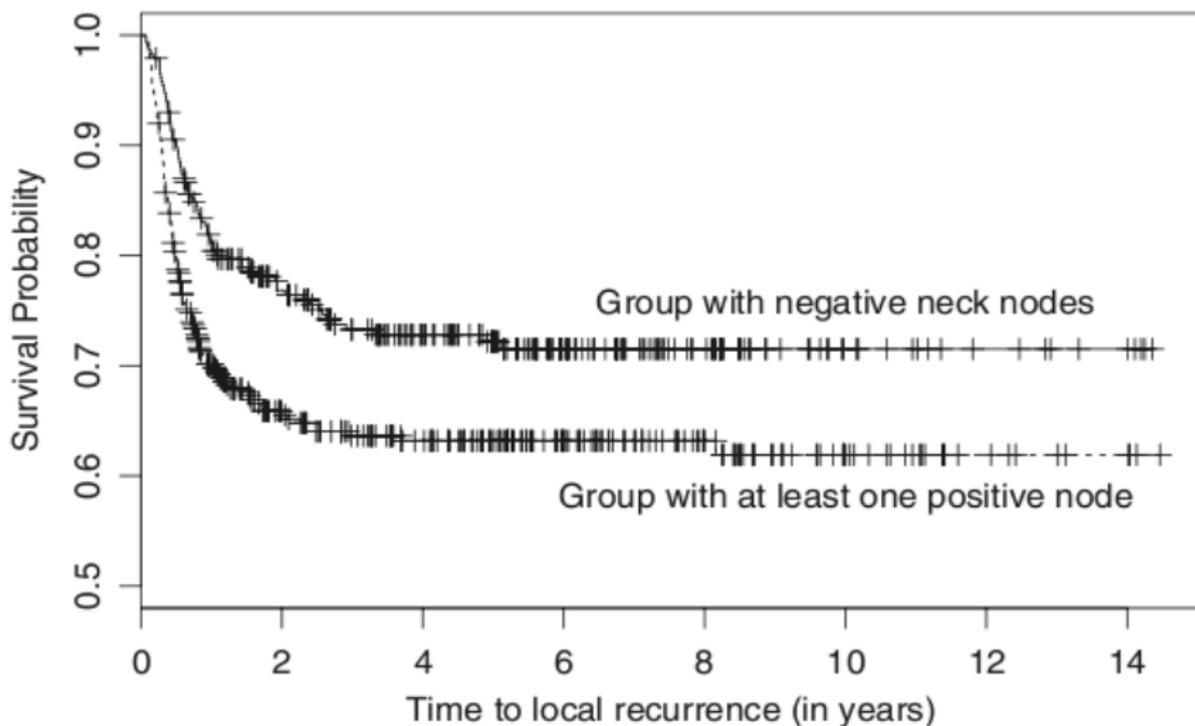


Figura: Estimativas de Kaplan-Meier da função de sobrevivência, dados de Peng *et al.* (2007) ($n = 672$).

Duas questões importantes

Duas perguntas importantes:

- Como estender a FS tal que seja possível modelar o mecanismo que conduz a possibilidade de cura ou a morte do paciente durante o tratamento?
- Como desenvolver procedimentos clássicos e bayesianos integrados para estimar a fração de cura?

Duas questões importantes

Duas perguntas importantes:

- Como estender a FS tal que seja possível modelar o mecanismo que conduz a possibilidade de cura ou a morte do paciente durante o tratamento?
- Como desenvolver procedimentos clássicos e bayesianos integrados para estimar a fração de cura?

Duas questões importantes

Duas perguntas importantes:

- Como estender a FS tal que seja possível modelar o mecanismo que conduz a possibilidade de cura ou a morte do paciente durante o tratamento?
- Como desenvolver procedimentos clássicos e bayesianos integrados para estimar a fração de cura?

Propriedades da FSI

- 1 Se $\Upsilon = 1$, então $S_{\text{pop}}(t) = S(t)$, isto é, esta classe contém a FS usual da AS,
- 2 $S_{\text{pop}}(0) = 1$,
- 3 $S_{\text{pop}}(t)$ é decrescente em t e
- 4 $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 - \Upsilon$.

Propriedades da FSI

- 1 Se $\Upsilon = 1$, então $S_{\text{pop}}(t) = S(t)$, isto é, esta classe contém a FS usual da AS,
- 2 $S_{\text{pop}}(0) = 1$,
- 3 $S_{\text{pop}}(t)$ é decrescente em t e
- 4 $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 - \Upsilon$.

Propriedades da FSI

- 1 Se $\Upsilon = 1$, então $S_{\text{pop}}(t) = S(t)$, isto é, esta classe contém a FS usual da AS,
- 2 $S_{\text{pop}}(0) = 1$,
- 3 $S_{\text{pop}}(t)$ é decrescente em t e
- 4 $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 - \Upsilon$.

Propriedades da FSI

- 1 Se $\Upsilon = 1$, então $S_{\text{pop}}(t) = S(t)$, isto é, esta classe contém a FS usual da AS,
- 2 $S_{\text{pop}}(0) = 1$,
- 3 $S_{\text{pop}}(t)$ é decrescente em t e
- 4 $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 - \Upsilon$.

Propriedades da FSI

- 1 Se $\Upsilon = 1$, então $S_{\text{pop}}(t) = S(t)$, isto é, esta classe contém a FS usual da AS,
- 2 $S_{\text{pop}}(0) = 1$,
- 3 $S_{\text{pop}}(t)$ é decrescente em t e
- 4 $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 - \Upsilon$.

Flexibilização da FS

- Levando em conta a propriedade 4,

$$p_0 = 1 - \Upsilon$$

é a fração de cura.

- A FSI $S_{\text{pop}}(t)$ também pode ser vista como um modelo de mistura da forma

$$S_{\text{pop}}(t) = p_0 + (1 - p_0)S(t), \quad (10)$$

$$S(t) = \int_t^{\infty} \frac{f(u)}{\Upsilon} du.$$

- Como formular um procedimento objetivo para obter $S_{\text{pop}}(t)$ que unifique a AS com os MLD? (Cap. 3)

Flexibilização da FS

- Levando em conta a propriedade 4,

$$p_0 = 1 - \Upsilon$$

é a fração de cura.

- A FSI $S_{\text{pop}}(t)$ também pode ser vista como um modelo de mistura da forma

$$S_{\text{pop}}(t) = p_0 + (1 - p_0)S(t), \quad (10)$$

$$S(t) = \int_t^{\infty} \frac{f(u)}{\Upsilon} du.$$

- Como formular um procedimento objetivo para obter $S_{\text{pop}}(t)$ que unifique a AS com os MLD? (Cap. 3)

Flexibilização da FS

- Levando em conta a propriedade 4,

$$p_0 = 1 - \Upsilon$$

é a fração de cura.

- A FSI $S_{\text{pop}}(t)$ também pode ser vista como um modelo de mistura da forma

$$S_{\text{pop}}(t) = p_0 + (1 - p_0)S(t), \quad (10)$$

$$S(t) = \int_t^{\infty} \frac{f(u)}{\Upsilon} du.$$

- Como formular um procedimento objetivo para obter $S_{\text{pop}}(t)$ que unifique a AS com os MLD? (Cap. 3)

Função de Sobrevivência Imprópria - FSI

$S_{pop}(t) \equiv$ função de sobrevivência da população (FSI)

$$S_{pop}(t) = 1 - \gamma + \int_t^\infty f(u) du \quad , \quad \gamma \leq 1$$

Propriedades:

- 1 Se $\gamma = 1 \Rightarrow S_{pop}(t) = S(t)$, isto é, esta classe contém a FS usual da AS;
- 2 $S_{pop}(0) = 1$;
- 3 $S_{pop}(t) \downarrow t$;
- 4 $\lim_{t \rightarrow \infty} S_p(t) = 1 - \gamma = p_0 \equiv$ fração de cura. .

Função de risco

- A função de risco é uma ferramenta essencial à análise de dados em AS. É definida como

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t | T \geq t]}{\Delta t}.$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (11)$$

- Motivados por (11), definimos a função de risco imprópria correspondente à FSI como

$$h_{\text{pop}}(t) = \frac{f_{\text{pop}}(t)}{S_{\text{pop}}(t)}, \quad \text{onde} \quad f_{\text{pop}}(t) = -\frac{d}{dt} S_{\text{pop}}(t).$$

Função de risco

- A função de risco é uma ferramenta essencial à análise de dados em AS. É definida como

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t | T \geq t]}{\Delta t}.$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (11)$$

- Motivados por (11), definimos a função de risco imprópria correspondente à FSI como

$$h_{\text{pop}}(t) = \frac{f_{\text{pop}}(t)}{S_{\text{pop}}(t)}, \quad \text{onde} \quad f_{\text{pop}}(t) = -\frac{d}{dt} S_{\text{pop}}(t).$$

Função de risco

- A função de risco é uma ferramenta essencial à análise de dados em AS. É definida como

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T \leq t + \Delta t | T \geq t]}{\Delta t}.$$

- Pode ser provado (Colosimo and Giolo, 2006) que

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log(S(t)). \quad (11)$$

- Motivados por (11), definimos a função de risco imprópria correspondente à FSI como

$$h_{\text{pop}}(t) = \frac{f_{\text{pop}}(t)}{S_{\text{pop}}(t)}, \quad \text{onde} \quad f_{\text{pop}}(t) = -\frac{d}{dt} S_{\text{pop}}(t).$$

Modelo Unificado de Longa Duração

- **Estágio inicial de ativação.** Seja N o número de fatores de risco de ocorrência de um evento de interesse com a seguinte distribuição de probabilidade e f.g.p.:

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots, \quad A(s) = \sum_{n=0}^{\infty} p_n s^n.$$

- **Estágio final de ativação (maturação).** Seja Y_i o tempo para a ativação do i -ésimo fator de risco, $i = 1, \dots, n$, com $F(y) = 1 - S(y)$ e independentes de N .

Problema crucial: como relacionar a variável aleatória observável T (tempo de ocorrência do evento de interesse) com as variáveis latentes do estágio final de ativação ou maturação?

Modelo Unificado de Longa Duração

- **Estágio inicial de ativação.** Seja N o número de fatores de risco de ocorrência de um evento de interesse com a seguinte distribuição de probabilidade e f.g.p.:

$$p_n = P[N = n], \quad n = 0, 1, 2, \dots, \quad A(s) = \sum_{n=0}^{\infty} p_n s^n.$$

- **Estágio final de ativação (maturação).** Seja Y_i o tempo para a ativação do i -ésimo fator de risco, $i = 1, \dots, n$, com $F(y) = 1 - S(y)$ e independentes de N .

Problema crucial: como relacionar a variável aleatória observável T (tempo de ocorrência do evento de interesse) com as variáveis latentes do estágio final de ativação ou maturação?

Flexibilização do mecanismo de ativação em dois estágios

- **Transição inicial (determinística).** O primeiro fator de risco ativado é responsável pela ocorrência do evento de interesse, ou seja

$$T = \min \{ Y_1, Y_2, \dots, Y_N \} = Y_{(1)}, \quad \text{com} \quad P[Y_0 = \infty] = 1.$$

- **Transição aleatória (Cooner *et al.*, 2007; Kim *et al.*, 2011).** A ordem de ativação dos fatores de risco é aleatorizada, isto é, a r -ésima ativação do fator de risco é uma v.a.d. tal que $N \geq r \geq 1$ e

$$T = Y_{(r)}.$$

Flexibilização do mecanismo de ativação em dois estágios

- **Transição inicial (determinística).** O primeiro fator de risco ativado é responsável pela ocorrência do evento de interesse, ou seja

$$T = \min \{ Y_1, Y_2, \dots, Y_N \} = Y_{(1)}, \quad \text{com} \quad P[Y_0 = \infty] = 1.$$

- **Transição aleatória (Cooner *et al.*, 2007; Kim *et al.*, 2011).** A ordem de ativação dos fatores de risco é aleatorizada, isto é, a r -ésima ativação do fator de risco é uma v.a.d. tal que $N \geq r \geq 1$ e

$$T = Y_{(r)}.$$

Flexibilização do mecanismo de ativação em dois estágios

- **Transição inicial (determinística).** O primeiro fator de risco ativado é responsável pela ocorrência do evento de interesse, ou seja

$$T = \min \{ Y_1, Y_2, \dots, Y_N \} = Y_{(1)}, \quad \text{com} \quad P[Y_0 = \infty] = 1.$$

- **Transição aleatória (Cooner *et al.*, 2007; Kim *et al.*, 2011).** A ordem de ativação dos fatores de risco é aleatorizada, isto é, a r -ésima ativação do fator de risco é uma v.a.d. tal que $N \geq r \geq 1$ e

$$T = Y_{(r)}.$$

Resultados sobre o MLD2E (p. 34)

Teorema (3.1)

A função de sobrevivência da v.a. T correspondente a um MLD2E é dada por

$$S_{pop}(t) = A_p(S(t)) = \sum_{n=0}^{\infty} p_n \{S(t)\}^n, \quad (12)$$

sendo que $A_p(\cdot)$ é a função geradora da seqüência $p = \{p_n\}$.

Demonstração.

Temos que

$$\begin{aligned} S_{\text{pop}}(t) &= P[N = 0] + P[Y_1 > t, Y_2 > t, \dots, Y_N > t, N \geq 1] \\ &= P[N = 0] + \sum_{n=1}^{\infty} P[N = n] P[Y_1 > t, Y_2 > t, \dots, Y_n > t] \\ &= p_0 + \sum_{n=1}^{\infty} p_n \{S(t)\}^n = A_p(S(t)), \end{aligned}$$

encerrando a prova. □

Obs. Este resultado responde à pergunta inicial relacionada com a modelagem do mecanismo de ocorrência do evento de interesse.

Resultados sobre o MLD2E (p. 34)

Teorema (3.2)

A função de sobrevivência de longa duração $S_{pop}(t)$ é tal que

$$\lim_{t \rightarrow \infty} S_{pop}(t) = P[N = 0] = p_0, \quad (13)$$

em que p_0 denota a proporção de não ocorrências do evento de interesse na população.

Resultados sobre o MLD2E (p. 36)

Teorema (3.3)

A função de sobrevivência de longa duração em dois estágios é dada por

$$S_{pop}(t) = p_0 + (1 - p_0)S^*(t), \quad (14)$$

em que

$$S^*(t) = \sum_{n=1}^{\infty} p_n^* \{S(t)\}^n \quad e \quad p_n^* = \frac{p_n}{1 - p_0}.$$

Resultados sobre o MLD2E (p. 37)

- Função densidade imprópria associada à função de sobrevivência de longa duração em dois estágios é dada por

$$f_{\text{pop}}(t) = f(t) \frac{dA_p(s)}{ds} \Big|_{s=S(t)}.$$

- A função de risco imprópria de longa duração em dois estágios é dada por

$$h_{\text{pop}}(t) = \frac{f_{\text{pop}}(t)}{S_{\text{pop}}(t)} = \frac{f(t)}{S_{\text{pop}}(t)} \frac{dA_p(s)}{ds} \Big|_{s=S(t)}. \quad (15)$$

Resultados sobre o MLD2E (p. 37)

- Função densidade imprópria associada à função de sobrevivência de longa duração em dois estágios é dada por

$$f_{\text{pop}}(t) = f(t) \frac{dA_p(s)}{ds} \Big|_{s=S(t)}.$$

- A função de risco imprópria de longa duração em dois estágios é dada por

$$h_{\text{pop}}(t) = \frac{f_{\text{pop}}(t)}{S_{\text{pop}}(t)} = \frac{f(t)}{S_{\text{pop}}(t)} \frac{dA_p(s)}{ds} \Big|_{s=S(t)}. \quad (15)$$

Resultados sobre o MLD2E (p. 38)

Teorema (3.4)

A função de risco imprópria de longa duração, $h_{pop}(t)$, tem a propriedade de riscos proporcionais se, e somente se, $p_n = \frac{\theta^n e^{-\theta}}{n!}$, para $n = 0, 1, \dots$, isto é, $N \sim \text{Poisson}(\theta)$.

Modelos de Longa Duração Existentes

- Uma ampla discussão de métodos de inferência freqüentista para o modelo de mistura padrão pode ser encontrado em Maller & Zhou (1996).
- Yakovlev & Tsodikov (1996) propõem um modelo de longa duração paramétrico baseado em características do crescimento do tumor.
- Chen, Ibrahim & Sinhá (1999) desenvolvem uma formulação bayesiana do modelo de Yakovlev & Tsodikov (1996).
- Desde então, diversas variações e extensões dos modelos de longa duração foram desenvolvidos, incluindo Chen, Ibrahim, & Sinha (2004), Kim, Chen, Dey, & Gamerman (2007), Cooner, Banerjee, Carlin & Sinha (2007), Rodrigues, de Castro, Cancho & Louzada-Neto (2009), e muitos outros.
- Modelo destrutivo: ver Rodrigues *et al.* (2011a).

Tabela: Alguns modelos de longa duração

	Distribuição de N		
	Binomial negativa (η, θ)	Poisson (θ)	Bernoulli (θ)
$S_{\text{pop}}(t)$	$\{1 + \eta\theta F(t)\}^{-1/\eta}$	$\exp\{-\theta F(t)\}$	$1 - \theta F(t)$
$f_{\text{pop}}(t)$	$\{1 + \eta\theta F(t)\}^{-1/\eta-1} \theta f(t)$	$\theta f(t) \exp\{-\theta F(t)\}$	$\theta f(t)$
$h_{\text{pop}}(t)$	$\{1 + \eta\theta F(t)\}^{-1} \theta f(t)$	$\theta f(t)$	$\frac{\theta}{1 - \theta F(t)} f(t)$
ρ_0	$(1 + \eta\theta)^{-1/\eta}$	$\exp(-\theta)$	$1 - \theta$

Modelo Destrutivo

- **Estágio de Promoção**

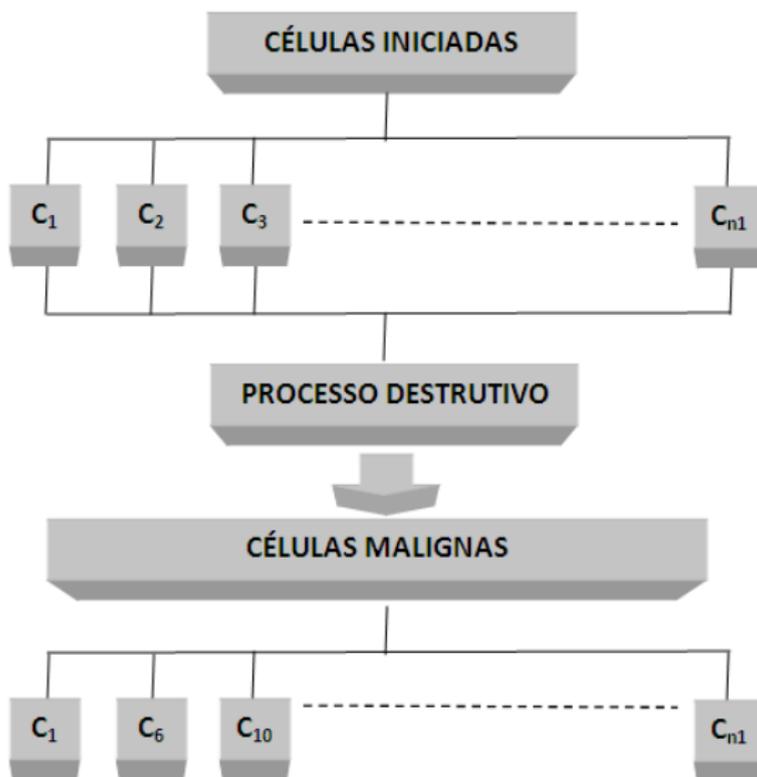
Rodrigues et al.(2011) - Destructive weighted Poisson cure rate models

PROCESSO DESTRUTIVO ($p(t) = p$)

D : o número total de células malignas que surgem como resultado da promoção entre as $N = n$ células iniciadas

$$D = \begin{cases} Z_1 + Z_2 + \cdots + Z_N & , N > 0 \\ 0 & , N = 0 \end{cases}$$

- Dado $N = n$, $Z_j, j = 1, \dots, N$, v.a.i e, independentes de N , seguindo uma Bernoulli(p);
- $D \leq N$.



Continuação.....

- **Estágio de Promoção**

Cenário de Risco Competitivos - Cox e Oakes (1984)

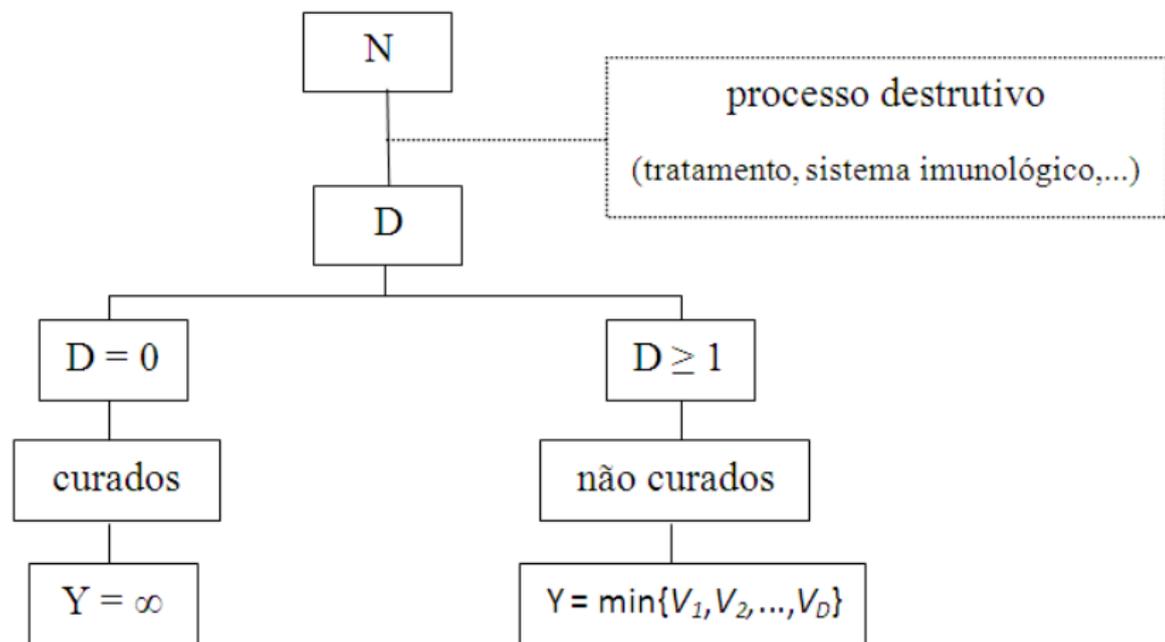
Dado $D=d$. Sejam $Y_j, j = 1, 2, \dots, d$, v.a.i.c com função de distribuição $F(y) = 1 - S(y)$ e independentes de D .

$Y_j \equiv$ Representa o tempo de promoção devido à j -ésima célula maligna.

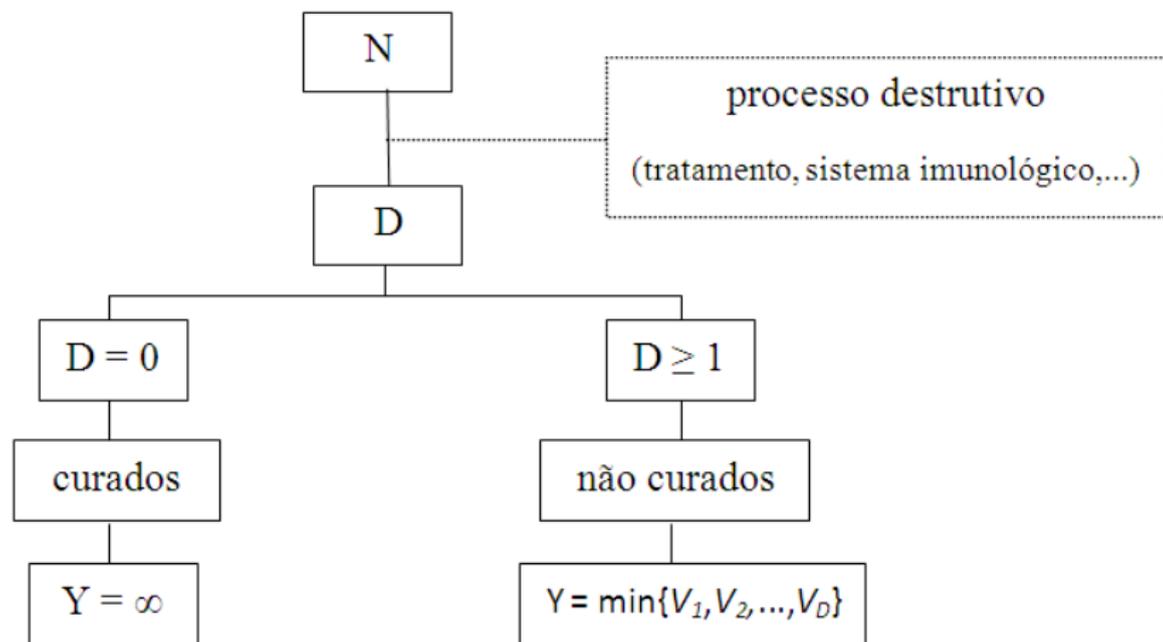
$Y \equiv$ Tempo de ocorrência do evento de interesse.

$$T = \min\{Y_i\}_{i=0}^D, \quad \mathbb{P}[Y_0 = \infty] = 1$$

Representação do MDC em termos de um diagrama



Representação do MDC em termos de um diagrama



Continuação....

Função de sobrevivência de longa duração Destrutivo (Rodrigues et. al. (2008,2011))

$$S_{pop}(t) = P[T \geq t] = \mathbb{A}_D(S(t)) = \sum_{d=0}^{\infty} P[D = d] \{S(t)\}^d = \mathbb{A}_N\left(\mathbb{A}_{Z_j}(S(t))\right)$$

$\mathbb{A}(\cdot) \equiv$ função geradora de probabilidade (f.g.p.) (Feller, 1968) .

Surge a Pergunta: Como introduzir a dependência entre as células?

- Distribuição Série de Potência Generalizada Inflada (SPGI) - Kolev et al. (2000)
- Haynatzk et.al (2000) - Importância da dependência biológica entre as células.

Continuação....

Função de sobrevivência de longa duração Destrutivo (Rodrigues et. al. (2008,2011))

$$S_{pop}(t) = P[T \geq t] = \mathbb{A}_D(S(t)) = \sum_{d=0}^{\infty} P[D = d] \{S(t)\}^d = \mathbb{A}_N\left(\mathbb{A}_{Z_j}(S(t))\right)$$

$\mathbb{A}(\cdot) \equiv$ função geradora de probabilidade (f.g.p.) (Feller, 1968) .

Surge a Pergunta: Como introduzir a dependência entre as células?

- Distribuição Série de Potência Generalizada Inflada (SPGI) - Kolev et al. (2000)
- Haynatzk et.al (2000) - Importância da dependência biológica entre as células.

Continuação....

Função de sobrevivência de longa duração Destrutivo (Rodrigues et. al. (2008,2011))

$$S_{pop}(t) = P[T \geq t] = \mathbb{A}_D(S(t)) = \sum_{d=0}^{\infty} P[D = d] \{S(t)\}^d = \mathbb{A}_N\left(\mathbb{A}_{Z_j}(S(t))\right)$$

$\mathbb{A}(\cdot) \equiv$ função geradora de probabilidade (f.g.p.) (Feller, 1968) .

Surge a Pergunta: Como introduzir a dependência entre as células?

- Distribuição Série de Potência Generalizada Inflada (SPGI) - Kolev et al. (2000)
- Haynatzk et.al (2000) - Importância da dependência biológica entre as células.

Continuação....

Função de sobrevivência de longa duração Destrutivo (Rodrigues et. al. (2008,2011))

$$S_{pop}(t) = P[T \geq t] = \mathbb{A}_D(S(t)) = \sum_{d=0}^{\infty} P[D = d] \{S(t)\}^d = \mathbb{A}_N\left(\mathbb{A}_{Z_j}(S(t))\right)$$

$\mathbb{A}(\cdot) \equiv$ função geradora de probabilidade (f.g.p.) (Feller, 1968) .

Surge a Pergunta: Como introduzir a dependência entre as células?

- Distribuição Série de Potência Generalizada Inflada (SPGI) - Kolev et al. (2000)
- Haynatzk et.al (2000) - Importância da dependência biológica entre as células.

Distribuição Série de Potência Generalizada Inflada (SPGI) - Kolev et al. (2000)

$$p_n = \mathbb{P}[N = n; \theta, \rho] = \frac{1}{g(\theta)} \sum_{n_1, n_2, \dots} a_n [\theta(1 - \rho)]^{\sum_{i=1}^{\infty} n_i} \rho^{\sum_{i=2}^{\infty} (i-1)n_i}, \quad n = 0, 1, \dots$$

- $\rho \in [0, 1]$ e a_n depende somente de n ;
- $g(\theta) = \sum_{n=0}^{\infty} a_n \theta^n$ é uma função diferenciável, finita e positiva;
- $\theta \in (0, s)$ (s pode ser ∞) é tal que $g(\theta)$ é finita;
- O somatório é sobre o conjunto de todos os inteiros não negativos n_1, n_2, \dots , tal que $\sum_{i=1}^{\infty} i n_i = n$.

Tabela: As escolhas de a_n , $g(\theta)$ e o parâmetro θ para alguns casos especiais da distribuição SPGI.

Distribuições	a_n	$g(\theta)$	θ	s
PI	$\frac{1}{n_1!n_2!\dots}$	e^θ	η	∞
BI	$\binom{m}{m-n_1-n_2-\dots, n_1, n_2, \dots}$	$(1 + \theta)^m$	$\frac{\pi}{1-\pi}$	1
BNI	$\frac{\Gamma(\phi^{-1} + \sum_{i=1}^{\infty} n_i)}{\Gamma(\phi^{-1}) [\sum_{i=1}^{\infty} n_i]!}$	$(1 - \theta)^{-\phi^{-1}}$	$\frac{\phi\eta}{1+\phi\eta}$	∞
SLI	$\frac{(-1+n_1+n_2+\dots)!}{n_1!n_2!\dots}$	$-\log(1 - \theta)$	$1 - \pi$	1

- Função geradora de probabilidade da v.a. SPGI

$$\mathbb{A}_{\rho_n}(z) = \frac{g(\theta z(1-\rho)(1-z\rho)^{-1})}{g(\theta)} \quad \text{para } 0 \leq z \leq 1$$

- Função de sobrevivência e fração de cura para v.a.

$$Y = \min\{V_i\}_{i=0}^D$$

$$S_{pop}(y) = \frac{g\left(\theta(1-\rho)(1-\rho F(y)) [1-(1-\rho F(y))\rho]^{-1}\right)}{g(\theta)}$$

$$\rho_0 = \lim_{y \rightarrow \infty} S_{pop}(y) = \frac{g(\theta(1-\rho)(1-\rho) [1-(1-\rho)\rho]^{-1})}{g(\theta)}$$

- Função geradora de probabilidade da v.a. SPGI

$$\mathbb{A}_{\rho_n}(z) = \frac{g(\theta z(1-\rho)(1-z\rho)^{-1})}{g(\theta)} \quad \text{para } 0 \leq z \leq 1$$

- Função de sobrevivência e fração de cura para v.a.

$$Y = \min\{V_i\}_{i=0}^D$$

$$S_{pop}(y) = \frac{g\left(\theta(1-\rho)(1-\rho F(y)) [1-(1-\rho F(y))\rho]^{-1}\right)}{g(\theta)}$$

$$\rho_0 = \lim_{y \rightarrow \infty} S_{pop}(y) = \frac{g(\theta(1-\rho)(1-\rho) [1-(1-\rho)\rho]^{-1})}{g(\theta)}$$

Casos especiais do modelo

- Modelo destrutivo correlacionado Poisson (DCP)

$$S_{pop}(y) = \exp \left\{ -\frac{\eta p F(y)}{1 - \rho(1 - p F(y))} \right\}$$

- Modelo destrutivo correlacionado binomial (DCB)

$$S_{pop}(y) = \left[1 - \frac{\pi p F(y)}{1 - \rho(1 - p F(y))} \right]^m$$

- Modelo destrutivo correlacionado binomial negativa (DCBN)

$$S_{pop}(y) = \left[\frac{1 - \rho(1 - p F(y))}{1 + \phi \eta p F(y) - \rho(1 - p F(y))} \right]^{\phi^{-1}}$$

- Modelo destrutivo correlacionado série logarítmica (DCSL)

$$\frac{(-\log(\pi))^{-1}}{(1 - p F(y))} \log \left\{ \frac{1 - \rho(1 - p F(y))}{1 - (1 - p F(y)) \rho} \right\}$$

Casos especiais do modelo

- Modelo destrutivo correlacionado Poisson (DCP)

$$S_{pop}(y) = \exp \left\{ -\frac{\eta p F(y)}{1 - \rho(1 - p F(y))} \right\}$$

- Modelo destrutivo correlacionado binomial (DCB)

$$S_{pop}(y) = \left[1 - \frac{\pi p F(y)}{1 - \rho(1 - p F(y))} \right]^m$$

- Modelo destrutivo correlacionado binomial negativa (DCBN)

$$S_{pop}(y) = \left[\frac{1 - \rho(1 - p F(y))}{1 + \phi \eta p F(y) - \rho(1 - p F(y))} \right]^{\phi^{-1}}$$

- Modelo destrutivo correlacionado série logarítmica (DCSL)

$$\frac{(-\log(\pi))^{-1}}{(1 - p F(y))} \log \left\{ \frac{1 - \rho(1 - p F(y))}{1 - (1 - p F(y))\rho} \right\}$$

Casos especiais do modelo

- Modelo destrutivo correlacionado Poisson (DCP)

$$S_{pop}(y) = \exp \left\{ -\frac{\eta p F(y)}{1 - \rho(1 - p F(y))} \right\}$$

- Modelo destrutivo correlacionado binomial (DCB)

$$S_{pop}(y) = \left[1 - \frac{\pi p F(y)}{1 - \rho(1 - p F(y))} \right]^m$$

- Modelo destrutivo correlacionado binomial negativa (DCBN)

$$S_{pop}(y) = \left[\frac{1 - \rho(1 - p F(y))}{1 + \phi \eta p F(y) - \rho(1 - p F(y))} \right]^{\phi^{-1}}$$

- Modelo destrutivo correlacionado série logarítmica (DCSL)

$$\frac{(-\log(\pi))^{-1}}{(1 - p F(y))} \log \left\{ \frac{1 - \rho(1 - p F(y))}{1 - (1 - p F(y))\rho} \right\}$$

Casos especiais do modelo

- Modelo destrutivo correlacionado Poisson (DCP)

$$S_{pop}(y) = \exp \left\{ -\frac{\eta p F(y)}{1 - \rho(1 - p F(y))} \right\}$$

- Modelo destrutivo correlacionado binomial (DCB)

$$S_{pop}(y) = \left[1 - \frac{\pi p F(y)}{1 - \rho(1 - p F(y))} \right]^m$$

- Modelo destrutivo correlacionado binomial negativa (DCBN)

$$S_{pop}(y) = \left[\frac{1 - \rho(1 - p F(y))}{1 + \phi \eta p F(y) - \rho(1 - p F(y))} \right]^{\phi^{-1}}$$

- Modelo destrutivo correlacionado série logarítmica (DCSL)

$$\frac{(-\log(\pi))^{-1}}{(1 - p F(y))} \log \left\{ \frac{1 - \rho(1 - p F(y))}{1 - (1 - p F(y))(1 - \pi(1 - \rho))} \right\}$$

Inferência

- A função de verossimilhança sob censura não-informativa é dada por

$$L(\vartheta; \mathbf{D}_{obs}) = \prod_{j=1}^m \{f_{pop}(t_j; \gamma)\}^{\delta_j} \{S_{pop}(t_j; \gamma)\}^{1-\delta_j}.$$

- Assumimos distribuição Weibull para o tempo de progressão (V)

$$F(v; \gamma) = 1 - \exp(-v^{\gamma_1} e^{\gamma_2}) \quad \text{e} \quad f(v; \gamma) = \gamma_1 v^{\gamma_1 - 1} \exp(\gamma_2 - v^{\gamma_1} e^{\gamma_2}),$$

para $v > 0$, $\gamma' = (\gamma_1, \gamma_2)$, com $\gamma_1 > 0$, e $\gamma_2 \in \mathfrak{R}$.

- Os modelos DCP, DCB e DCBN não são identificáveis no sentido de Li et al. (2001)., isto é, existem ϑ e ϑ^* , $\vartheta \neq \vartheta^*$, tais que $S_{pop}(y; \vartheta) = S_{pop}(y; \vartheta^*)$.

Inferência

- A função de verossimilhança sob censura não-informativa é dada por

$$L(\vartheta; \mathbf{D}_{obs}) = \prod_{j=1}^m \{f_{pop}(t_j; \gamma)\}^{\delta_j} \{S_{pop}(t_j; \gamma)\}^{1-\delta_j}.$$

- Assumimos distribuição Weibull para o tempo de progressão (V)

$$F(v; \gamma) = 1 - \exp(-v^{\gamma_1} e^{\gamma_2}) \quad \text{e} \quad f(v; \gamma) = \gamma_1 v^{\gamma_1 - 1} \exp(\gamma_2 - v^{\gamma_1} e^{\gamma_2}),$$

para $v > 0$, $\gamma' = (\gamma_1, \gamma_2)$, com $\gamma_1 > 0$, e $\gamma_2 \in \mathfrak{R}$.

- Os modelos DCP, DCB e DCBN não são identificáveis no sentido de Li et al. (2001)., isto é, existem ϑ e ϑ^* , $\vartheta \neq \vartheta^*$, tais que $S_{pop}(y; \vartheta) = S_{pop}(y; \vartheta^*)$.

Inferência

- A função de verossimilhança sob censura não-informativa é dada por

$$L(\boldsymbol{\vartheta}; \mathbf{D}_{obs}) = \prod_{j=1}^m \{f_{pop}(t_j; \boldsymbol{\gamma})\}^{\delta_j} \{S_{pop}(t_j; \boldsymbol{\gamma})\}^{1-\delta_j}.$$

- Assumimos distribuição Weibull para o tempo de progressão (V)

$$F(v; \boldsymbol{\gamma}) = 1 - \exp(-v^{\gamma_1} e^{\gamma_2}) \quad \text{e} \quad f(v; \boldsymbol{\gamma}) = \gamma_1 v^{\gamma_1 - 1} \exp(\gamma_2 - v^{\gamma_1} e^{\gamma_2}),$$

para $v > 0$, $\boldsymbol{\gamma}' = (\gamma_1, \gamma_2)$, com $\gamma_1 > 0$, e $\gamma_2 \in \Re$.

- Os modelos DCP, DCB e DCBN não são identificáveis no sentido de Li et al. (2001)., isto é, existem $\boldsymbol{\vartheta}$ e $\boldsymbol{\vartheta}^*$, $\boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}^*$, tais que $S_{pop}(y; \boldsymbol{\vartheta}) = S_{pop}(y; \boldsymbol{\vartheta}^*)$.

setbeamercoveredtransparent

- Para evitar este problema, propomos relacionar os parâmetros ρ e η (ou π) dos modelos DCP, DCB e DCBN com os vetores de covariáveis $\mathbf{x}'_j = (x_{j1}, \dots, x_{jk_1})$ e $\mathbf{w}'_j = (w_{j1}, \dots, w_{jk_1})$, respectivamente, sem elementos comuns.
- As funções de ligação

$$\log\left(\frac{\rho_j}{1 - \rho_j}\right) = \mathbf{x}'_j \beta_1, \quad \text{e} \quad \log(\eta_j) = \mathbf{w}'_j \beta_2 \quad \text{ou} \quad \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \mathbf{w}'_j \beta_2$$

$j = 1, \dots, n$, sendo $\beta'_1 = (\beta_{11}, \dots, \beta_{1k_1})$ e $\beta'_2 = (\beta_{21}, \dots, \beta_{2k_2})$ vetores com k_1 e k_2 coeficientes de regressão.

- $\vartheta = (\beta'_1, \beta'_2, \gamma', \rho, \phi, m)$
- As estimativas de máxima verossimilhança de $\hat{\vartheta}$ são obtidas maximizando o logaritmo da função de verossimilhança $\ell(\vartheta; \mathbf{D}_{obs}) = \log(L(\vartheta; \mathbf{D}_{obs}))$.
- A maximização foi efetuada numericamente aplicando o método L-BFGS-B, implementado na função *optim* do sistema R (R Development Core Team).

- Para evitar este problema, propomos relacionar os parâmetros ρ e η (ou π) dos modelos DCP, DCB e DCBN com os vetores de covariáveis $\mathbf{x}'_j = (x_{j1}, \dots, x_{jk_1})$ e $\mathbf{w}'_j = (w_{j1}, \dots, w_{jk_1})$, respectivamente, sem elementos comuns.
- As funções de ligação

$$\log\left(\frac{\rho_j}{1 - \rho_j}\right) = \mathbf{x}'_j \beta_1, \quad \text{e} \quad \log(\eta_j) = \mathbf{w}'_j \beta_2 \quad \text{ou} \quad \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \mathbf{w}'_j \beta_2$$

$j = 1, \dots, n$, sendo $\beta'_1 = (\beta_{11}, \dots, \beta_{1k_1})$ e $\beta'_2 = (\beta_{21}, \dots, \beta_{2k_2})$ vetores com k_1 e k_2 coeficientes de regressão.

- $\vartheta = (\beta'_1, \beta'_2, \gamma', \rho, \phi, m)$
- As estimativas de máxima verossimilhança de $\hat{\vartheta}$ são obtidas maximizando o logaritmo da função de verossimilhança $\ell(\vartheta; \mathbf{D}_{obs}) = \log(L(\vartheta; \mathbf{D}_{obs}))$.
- A maximização foi efetuada numericamente aplicando o método L-BFGS-B, implementado na função *optim* do sistema R (R Development Core Team).

- Para evitar este problema, propomos relacionar os parâmetros ρ e η (ou π) dos modelos DCP, DCB e DCBN com os vetores de covariáveis $\mathbf{x}'_j = (x_{j1}, \dots, x_{jk_1})$ e $\mathbf{w}'_j = (w_{j1}, \dots, w_{jk_1})$, respectivamente, sem elementos comuns.
- As funções de ligação

$$\log\left(\frac{\rho_j}{1 - \rho_j}\right) = \mathbf{x}'_j \beta_1, \quad \text{e} \quad \log(\eta_j) = \mathbf{w}'_j \beta_2 \quad \text{ou} \quad \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \mathbf{w}'_j \beta_2$$

$j = 1, \dots, n$, sendo $\beta'_1 = (\beta_{11}, \dots, \beta_{1k_1})$ e $\beta'_2 = (\beta_{21}, \dots, \beta_{2k_2})$ vetores com k_1 e k_2 coeficientes de regressão.

- $\vartheta = (\beta'_1, \beta'_2, \gamma', \rho, \phi, m)$
- As estimativas de máxima verossimilhança de $\hat{\vartheta}$ são obtidas maximizando o logaritmo da função de verossimilhança $\ell(\vartheta; \mathbf{D}_{obs}) = \log(L(\vartheta; \mathbf{D}_{obs}))$.
- A maximização foi efetuada numericamente aplicando o método L-BFGS-B, implementado na função *optim* do sistema R (R Development Core Team).

- Para evitar este problema, propomos relacionar os parâmetros ρ e η (ou π) dos modelos DCP, DCB e DCBN com os vetores de covariáveis $\mathbf{x}'_j = (x_{j1}, \dots, x_{jk_1})$ e $\mathbf{w}'_j = (w_{j1}, \dots, w_{jk_1})$, respectivamente, sem elementos comuns.
- As funções de ligação

$$\log\left(\frac{\rho_j}{1 - \rho_j}\right) = \mathbf{x}'_j \beta_1, \quad \text{e} \quad \log(\eta_j) = \mathbf{w}'_j \beta_2 \quad \text{ou} \quad \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \mathbf{w}'_j \beta_2$$

$j = 1, \dots, n$, sendo $\beta'_1 = (\beta_{11}, \dots, \beta_{1k_1})$ e $\beta'_2 = (\beta_{21}, \dots, \beta_{2k_2})$ vetores com k_1 e k_2 coeficientes de regressão.

- $\vartheta = (\beta'_1, \beta'_2, \gamma', \rho, \phi, m)$
- As estimativas de máxima verossimilhança de $\hat{\vartheta}$ são obtidas maximizando o logaritmo da função de verossimilhança $\ell(\vartheta; \mathbf{D}_{obs}) = \log(L(\vartheta; \mathbf{D}_{obs}))$.
- A maximização foi efetuada numericamente aplicando o método L-BFGS-B, implementado na função *optim* do sistema R (R Development Core Team).

- Para evitar este problema, propomos relacionar os parâmetros ρ e η (ou π) dos modelos DCP, DCB e DCBN com os vetores de covariáveis $\mathbf{x}'_j = (x_{j1}, \dots, x_{jk_1})$ e $\mathbf{w}'_j = (w_{j1}, \dots, w_{jk_1})$, respectivamente, sem elementos comuns.
- As funções de ligação

$$\log\left(\frac{\rho_j}{1 - \rho_j}\right) = \mathbf{x}'_j \beta_1, \quad \text{e} \quad \log(\eta_j) = \mathbf{w}'_j \beta_2 \quad \text{ou} \quad \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \mathbf{w}'_j \beta_2$$

$j = 1, \dots, n$, sendo $\beta'_1 = (\beta_{11}, \dots, \beta_{1k_1})$ e $\beta'_2 = (\beta_{21}, \dots, \beta_{2k_2})$ vetores com k_1 e k_2 coeficientes de regressão.

- $\vartheta = (\beta'_1, \beta'_2, \gamma', \rho, \phi, m)$
- As estimativas de máxima verossimilhança de $\hat{\vartheta}$ são obtidas maximizando o logaritmo da função de verossimilhança $\ell(\vartheta; \mathbf{D}_{obs}) = \log(L(\vartheta; \mathbf{D}_{obs}))$.
- A maximização foi efetuada numericamente aplicando o método L-BFGS-B, implementado na função *optim* do sistema R (R Development Core Team).

Aplicação: Dados de Melanoma Maligno

Os dados fazem parte de um estudo de melanoma maligno do hospital universitário de Odense (dinamarca) por K. T. Drzewiecki. O conjunto de dados inclui 205 pacientes observados após operação para a remoção de melanoma maligno no período de 16 anos. Estes dados estão disponíveis no pacote *timereg* no R. O tempo observado (T) varia de 10 a 5565 dias (de 0,0274 a 15,25 anos, com média = 5,9 e desvio-padrão = 3,1 anos) e se refere ao tempo até a morte do paciente ou o tempo de censura. Pacientes que morreram de outras causas, bem como pacientes que ainda estavam vivos ao final do estudo são observações censuradas (72%).

Tomamos o estado de úlcera (x) (ausente, $m = 115$; presente, $m = 90$) e espessura do tumor (w) (em mm, média = 2,92 e desvio padrão = 2,96), como covariáveis.

$$\log\left(\frac{p_j}{1-p_j}\right) = \beta_{1,intecp} + \beta_{1,esp}w, \quad e \quad \log(\eta_j) = \beta_{2,ulc:pres}x + \beta_{2,ulc:aus}(1-x)$$

Aplicação (Continuação)

Tabela : Os valores do $\max \log L(\cdot)$ e as estatísticas AIC e BIC para os sete modelos ajustados: DCP, DCB, DCBN, DCG, DCSL, binomial negativa e geométrico.

Critério	DCP	DCB	DCBN	DCG	DCSL	Binomial negativa	Geométrico
$\max \log L(\cdot)$	-198,60	-198,61	-198,12	-198,52	-197,96	-201,52	-205,42
AIC	411,21	413,21	412,24	411,06	413,92	415,04	420,83
BIC	434,47	439,80	438,82	434,32	443,83	435,00	437,45

Tabela: As estatísticas DIC, EAIC, EBIC e B para os sete modelos ajustados: DCP, DCB, DCBN, DCG, DCSL, binomial negativa e geométrico.

Critério	DCP	DCB	DCBN	DCG	DCSL	Binomial negativa	Geométrico
DIC	406,21	407,73	407,01	406,56	415,52	413,63	416,31
EAIC	419,60	421,11	421,40	417,90	425,54	420,51	427,10
EBIC	442,86	447,68	447,98	441,16	448,76	440,44	443,72
B	-206,49	-205,92	-205,84	-206,33	-208,76	-206,97	-212,54

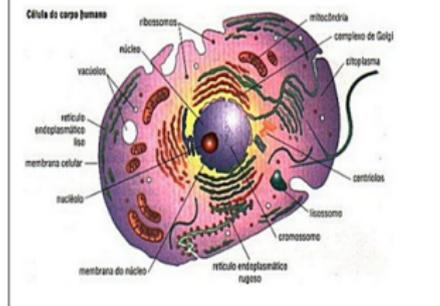
Aplicação (Continuação)

Tabela: Estimativas de máxima verossimilhança dos parâmetros do modelo DCG, seus desvios padrão e seus intervalos de confiança assintóticos de 95% (IC 95%).

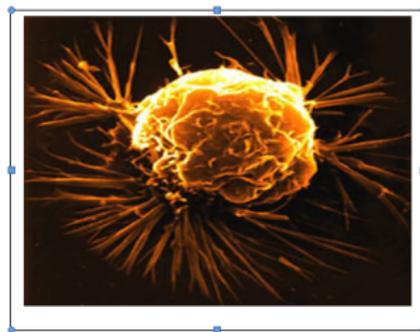
Parâmetro	Estimativa	desvio padrão	IC 95%
γ_1	2,46	0,34	(1,79 ; 3,12)
γ_2	-5,54	1,16	(-7,81 ; -3,26)
ρ	0,95	0,06	(0,83 ; 1,00)
$\beta_{1,intercepto}$	-4,84	0,95	(-6,70 ; -2,98)
$\beta_{1,espessura}$	0,95	0,27	(0,42 ; 1,48)
$\beta_{2,ulc:presente}$	0,53	0,30	(-0,06 ; 1,12)
$\beta_{2,ulc:ausente}$	-0,48	0,41	(-1,28 ; 0,32)

Modelo Híbrido (Patrick, 2011)

Célula: É a menor porção do nosso organismo capaz de criar sua própria energia, crescer e se multiplicar.



Câncer: O conjunto de mais de 100 doenças que têm em comum o crescimento desordenado de células que invadem



Objetivo: Propor modelos de longa duração para acomodarem características dos estágios (iniciação, promoção e **progressão**) da carcinogênese.

Como é o Processo da Carcinogênese

O processo de carcinogênese, ou seja, de formação de câncer, em geral se dá lentamente, podendo levar vários anos para que uma célula cancerosa prolifere e dê origem a um tumor visível. Esse processo passa por vários estágios antes de chegar ao tumor. São eles:



Pressupostos Básicos

- (i) O evento de iniciação no processo da carcinogênese é a formação de uma lesão pré-cancerosa que, no longo prazo, é capaz de produzir um tumor evidente. Pensamos nessas lesões pré-cancerosas como as células iniciadas. Tratamos o número de células iniciadas como uma variável aleatória N_1 ;
- (ii) Todas as lesões primárias podem ser consideradas como estando sujeitas a processos de reparo ou eliminadas depois de algum tratamento prolongado, mas algumas delas ainda não serão reconhecidas pelo sistema de reparo ou tratamento não funcionou e, conseqüentemente, não reparadas;
- (iii) Uma lesão pré-cancerosa não reparada permanece dormente enquanto ela prossegue com o estágio de promoção de desenvolvimento do tumor. Todas as lesões estão sujeitas a promoção independentemente (ou dependentemente) umas das outras;

Pressupostos Básicos

- (i) O evento de iniciação no processo da carcinogênese é a formação de uma lesão pré-cancerosa que, no longo prazo, é capaz de produzir um tumor evidente. Pensamos nessas lesões pré-cancerosas como as células iniciadas. Tratamos o número de células iniciadas como uma variável aleatória N_1 ;
- (ii) Todas as lesões primárias podem ser consideradas como estando sujeitas a processos de reparo ou eliminadas depois de algum tratamento prolongado, mas algumas delas ainda não serão reconhecidas pelo sistema de reparo ou tratamento não funcionou e, conseqüentemente, não reparadas;
- (iii) Uma lesão pré-cancerosa não reparada permanece dormente enquanto ela prossegue com o estágio de promoção de desenvolvimento do tumor. Todas as lesões estão sujeitas a promoção independentemente (ou dependentemente) umas das outras;

Pressupostos Básicos

- (i) O evento de iniciação no processo da carcinogênese é a formação de uma lesão pré-cancerosa que, no longo prazo, é capaz de produzir um tumor evidente. Pensamos nessas lesões pré-cancerosas como as células iniciadas. Tratamos o número de células iniciadas como uma variável aleatória N_1 ;
- (ii) Todas as lesões primárias podem ser consideradas como estando sujeitas a processos de reparo ou eliminadas depois de algum tratamento prolongado, mas algumas delas ainda não serão reconhecidas pelo sistema de reparo ou tratamento não funcionou e, conseqüentemente, não reparadas;
- (iii) Uma lesão pré-cancerosa não reparada permanece dormente enquanto ela prossegue com o estágio de promoção de desenvolvimento do tumor. Todas as lesões estão sujeitas a promoção independentemente (ou dependentemente) umas das outras;

Pressupostos Básicos

- (i) O evento de iniciação no processo da carcinogênese é a formação de uma lesão pré-cancerosa que, no longo prazo, é capaz de produzir um tumor evidente. Pensamos nessas lesões pré-cancerosas como as células iniciadas. Tratamos o número de células iniciadas como uma variável aleatória N_1 ;
- (ii) Todas as lesões primárias podem ser consideradas como estando sujeitas a processos de reparo ou eliminadas depois de algum tratamento prolongado, mas algumas delas ainda não serão reconhecidas pelo sistema de reparo ou tratamento não funcionou e, conseqüentemente, não reparadas;
- (iii) Uma lesão pré-cancerosa não reparada permanece dormente enquanto ela prossegue com o estágio de promoção de desenvolvimento do tumor. Todas as lesões estão sujeitas a promoção independentemente (ou dependentemente) umas das outras;

Continuação...

- (iv) Uma vez que a célula maligna ou clonogênica surge como resultado da promoção da célula iniciada, começa o estágio de progressão produzindo uma colônia de descendentes (células tumorais), chamada de clone ou tumor. Tratamos o número de células malignas resultantes do estágio de promoção como uma variável aleatória N_2 . O tempo que uma célula maligna leva para se transformar em um tumor detectável é considerado como uma variável aleatória com função de distribuição $F(\cdot) = 1 - S(\cdot)$, sendo $S(\cdot)$ função de sobrevivência. Todas células malignas estão sujeitas a progressão independentemente (ou dependentemente) umas das outras.
- (v) Um tumor torna-se detectável quando o seu tamanho atinge um valor limite (proliferações de células tumorais). Tratamos o número de células tumorais como uma variável aleatória N_3 .

Continuação...

- (iv) Uma vez que a célula maligna ou clonogênica surge como resultado da promoção da célula iniciada, começa o estágio de progressão produzindo uma colônia de descendentes (células tumorais), chamada de clone ou tumor. Tratamos o número de células malignas resultantes do estágio de promoção como uma variável aleatória N_2 . O tempo que uma célula maligna leva para se transformar em um tumor detectável é considerado como uma variável aleatória com função de distribuição $F(\cdot) = 1 - S(\cdot)$, sendo $S(\cdot)$ função de sobrevivência. Todas células malignas estão sujeitas a progressão independentemente (ou dependentemente) umas das outras.
- (v) Um tumor torna-se detectável quando o seu tamanho atinge um valor limite (proliferações de células tumorais). Tratamos o número de células tumorais como uma variável aleatória N_3 .

Continuação...

- (iv) Uma vez que a célula maligna ou clonogênica surge como resultado da promoção da célula iniciada, começa o estágio de progressão produzindo uma colônia de descendentes (células tumorais), chamada de clone ou tumor. Tratamos o número de células malignas resultantes do estágio de promoção como uma variável aleatória N_2 . O tempo que uma célula maligna leva para se transformar em um tumor detectável é considerado como uma variável aleatória com função de distribuição $F(\cdot) = 1 - S(\cdot)$, sendo $S(\cdot)$ função de sobrevivência. Todas células malignas estão sujeitas a progressão independentemente (ou dependentemente) umas das outras.
- (v) Um tumor torna-se detectável quando o seu tamanho atinge um valor limite (proliferações de células tumorais). Tratamos o número de células tumorais como uma variável aleatória N_3 .

Modelo Híbrido

- **Estágio de Progressão**

$N_3 \equiv$ Número de células tumorais originadas de cada célula maligna (nível "threshold").

$$p_{n_3} = P[N_3 = n_3] \quad , \quad n_3 = 0, 1, \dots$$

Dado $N_k = n_k$, $k = 1, 2, 3$. Sejam Z_{ij} , $i = 1, \dots, N_1$ e $j = 1, \dots, N_3$, v.a.i.c com função de distribuição $F(z) = 1 - S(z)$ e independentes de N_k .

$Z_{ij} \equiv$ O tempo que a (i, j) -ésima célula maligna transformar-se em um tumor detectável, denominado tempo de progressão.

$Y \equiv$ Tempo de ocorrência do evento de interesse.

Modelo Híbrido....

Como incorporar o nível threshold no cenário de riscos competitivos?

ESQUEMA DE ATIVAÇÃO HÍBRIDO

$$T = \min \left\{ \max \left\{ Y_{ij} \right\}_{j=0}^{N_3} \right\}_{i=0}^{N_2}$$

$\mathbb{P}[Y_{0j} = \infty] = 1 \equiv$ Fração de cura

$\mathbb{P}[Y_{i0} = \infty] = 1 \equiv$ Proporção de células malignas que "morrem" antes da indução do tumor.

Modelo Híbrido....

Como incorporar o nível threshold no cenário de riscos competitivos?

ESQUEMA DE ATIVAÇÃO HÍBRIDO

$$T = \min \left\{ \max \left\{ Y_{ij} \right\}_{j=0}^{N_3} \right\}_{i=0}^{N_2}$$

$\mathbb{P}[Y_{0j} = \infty] = 1 \equiv$ Fração de cura

$\mathbb{P}[Y_{i0} = \infty] = 1 \equiv$ Proporção de células malignas que "morrem" antes da indução do tumor.

Modelo Híbrido....

Como incorporar o nível threshold no cenário de riscos competitivos?

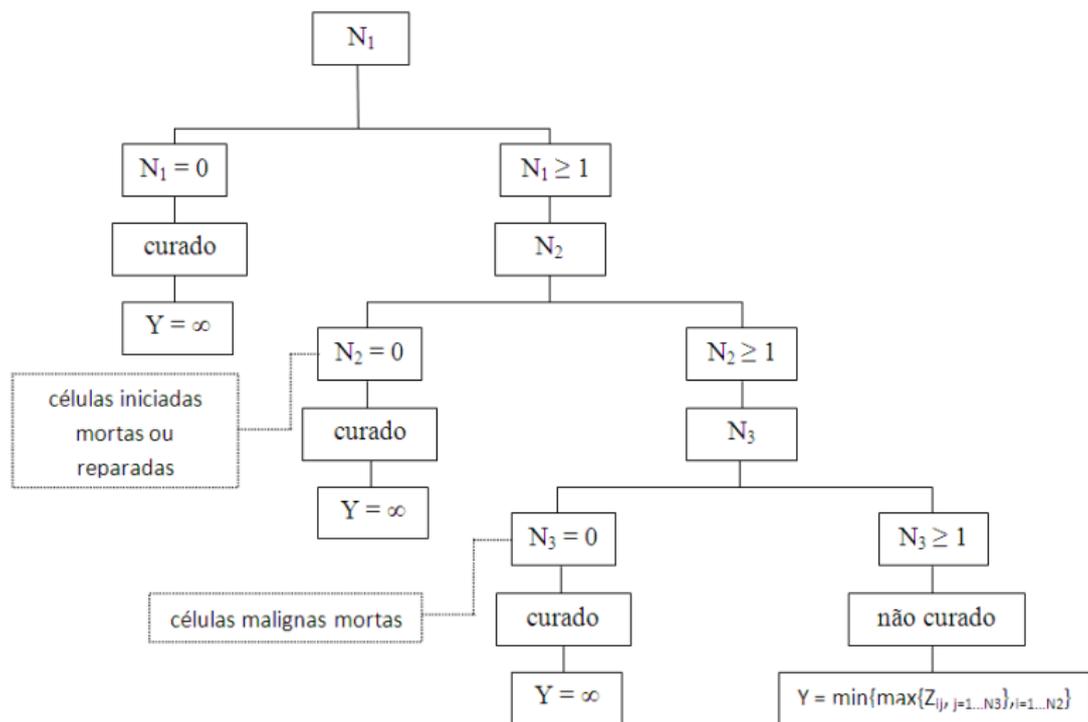
ESQUEMA DE ATIVAÇÃO HÍBRIDO

$$T = \min \left\{ \max \left\{ Y_{ij} \right\}_{j=0}^{N_3} \right\}_{i=0}^{N_2}$$

$\mathbb{P}[Y_{0j} = \infty] = 1 \equiv$ Fração de cura

$\mathbb{P}[Y_{i0} = \infty] = 1 \equiv$ Proporção de células malignas que "morrem" antes da indução do tumor.

Representação do modelo híbrido em termos de um diagrama



Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
 - Alguns modelos especiais.
 - Um exemplo ilustrativo.
 - Conclusões
 - Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Tópicos: Um mecanismo unificado para formulação de novas distribuições de probabilidades em Análise de Sobrevivência

- Mecanismo de seleção: processo carcinogênico.
- Probabilidade individual.
- Mecanismo composto pela primeira e última ativação de células iniciadas.
- Alguns resultados importantes.
- Alguns modelos especiais.
- Um exemplo ilustrativo.
- Conclusões
- Referências

Mecanismo de seleção: processo carcinogênico.

- Utilizar o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível em Análise de Sobrevivência.
- Este modelo contém vários modelos para tempos de vida recentemente introduzidos (interpretação biológica).

Mecanismo de seleção: processo carcinogênico.

- Utilizar o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível em Análise de Sobrevivência.
- Este modelo contém vários modelos para tempos de vida recentemente introduzidos (interpretação biológica).

Processo carcinogênico

- Mecanismo de seleção: Processo carcinogênico

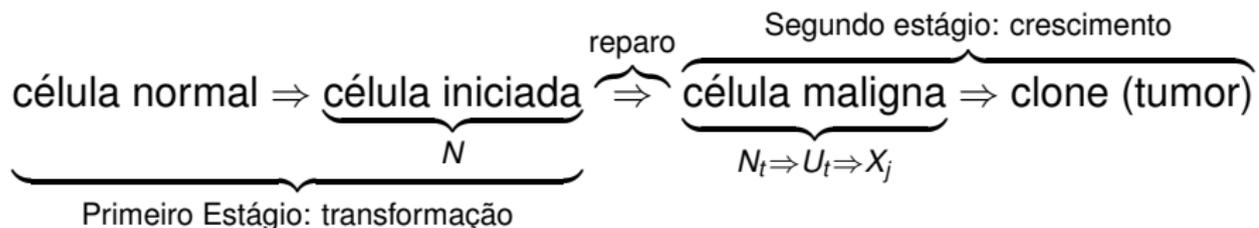


Figura: Processo carcinogênico em dois estágios

U_t : variável de seleção

X_j : tempo de promoção da j -ésima célula iniciada (clonogens).

Modelo latente com taxa de cura

- **Primeiro estágio (transformação) $\Rightarrow N$ (células danificadas)**

$$p_n = P(N = n), \quad n = 0, 1, \dots \quad (16)$$

$$A_N(s) = \sum_{n=0}^{\infty} p_n s^n \Rightarrow (\text{pgf}) \Rightarrow p_0 = P[N = 0]. \quad (17)$$

Feller (1967): "The power and the possibilities of the pgf are rarely fully utilized."

p_0 : taxa de cura

Modelo latente com taxa de cura

- **Segundo estágio (crescimento)** Dado

$N = n \Rightarrow X_j \perp X_k \mid N = n \quad (j \neq k)$ com f.d.p $g(x)$ e $S(x) = 1 - G(x)$.

$$N_t = \begin{cases} Z_1 + Z_2 + \dots + Z_N, & \text{se } N > 0, \\ 0, & \text{if } N = 0, \end{cases} \quad (18)$$

$Z_j = \begin{cases} 1, & \text{se } X_j \leq t \Leftrightarrow j\text{-ésima célula é ativada durante o tempo } t, \\ 0, & \text{se } X_j > t \Leftrightarrow j\text{-ésima célula não é ativada durante o tempo } t, \end{cases}$

$Z_j \approx \text{Bern}[G(t)], j = 1, \dots, n$: presença do j -ésimo clone durante o tempo t

N_t : variável latente (destrutiva)

$$(Z_j \perp N)$$

Modelo flexível para o tempo de vida T .

- Esquema de ativação no instante t : R

$$N_t = R \quad (R \geq 1) \Rightarrow T = X_{(R)}$$

- Primeira ativação durante o tempo t : $R = 1 \Rightarrow T = X_{(1)}$.
- Última ativação no instante t : $R = N \Rightarrow T = X_{(N)}$.
- **Problema:** Flexibilizar a fd $g(t)$ de $T(R = 1$ ou $R = N)$ para pacientes expostos ao agente cancerígeno durante o tempo t .

Modelo flexível para o tempo de vida T .

- Esquema de ativação no instante t : R

$$N_t = R \quad (R \geq 1) \Rightarrow T = X_{(R)}$$

- Primeira ativação durante o tempo t : $R = 1 \Rightarrow T = X_{(1)}$.
- Última ativação no instante t : $R = N \Rightarrow T = X_{(N)}$.
- **Problema:** Flexibilizar a fd $g(t)$ de $T(R = 1$ ou $R = N)$ para pacientes expostos ao agente cancerígeno durante o tempo t .

Modelo flexível para o tempo de vida T .

- Esquema de ativação no instante t : R

$$N_t = R \quad (R \geq 1) \Rightarrow T = X_{(R)}$$

- Primeira ativação durante o tempo t : $R = 1 \Rightarrow T = X_{(1)}$.
- Última ativação no instante t : $R = N \Rightarrow T = X_{(N)}$.
- Problema:** Flexibilizar a fd $g(t)$ de $T(R = 1$ ou $R = N)$ para pacientes expostos ao agente cancerígeno durante o tempo t .

Modelo flexível para o tempo de vida T .

- Esquema de ativação no instante t : R

$$N_t = R \quad (R \geq 1) \Rightarrow T = X_{(R)}$$

- Primeira ativação durante o tempo t : $R = 1 \Rightarrow T = X_{(1)}$.
- Última ativação no instante t : $R = N \Rightarrow T = X_{(N)}$.
- Problema:** Flexibilizar a fd $g(t)$ de $T(R = 1$ ou $R = N)$ para pacientes expostos ao agente cancerígeno durante o tempo t .

Continuação...

- Segue da fórmula fundamental de prob. condicional que

$$P(N_t = j) = \sum_{n=j}^{\infty} p_n \overbrace{P(N_t = j | N = n)}^{\text{Binomial}(n, G(t)) : \text{mecanismo de destruição}},$$

correspondente a fgp (Feller, 1968) dada por

$$A_{N_t}(s) = A_N[1 - (1 - s)G(t)]. \quad (19)$$

A função de sobrevivência de longa duração (Rodrigues *et al.*, 2008) é obtida de (19) como

$$S_{\text{Pop}}(t) = P(T \geq t) = P[N_t = 0] = A_{N_t}(0) = A_N[S(t)]. \quad (20)$$

Continuação...

- Motivado pelo trabalho de Arellano-Valle *et al.* (2006) vamos definir a variável de seleção e conectar com a fgp $A_{N_t}(s)$ e a fd $g(x)$ do tempo de promoção X . Inicialmente, vamos supor que a população de pacientes está dividida em duas sub-populações (curados e não curados) através da seguinte v. de Bernoulli (**mecanismo de seleção**) durante o tempo t :

$$U_t = \begin{cases} 1, & \text{se } N_t \geq 1, \\ 0, & \text{se } N_t = 0, \end{cases} \quad (21)$$

onde $P(U_t = 1) = 1 - P(N_t = 0) = 1 - p_0$.

Definições

Definição

Definimos **como mecanismo probabilístico de seleção de T** a distribuição condicional de Y dado $U_t = 1$.

- Esta definição simplesmente afirma que o mecanismo de seleção da va T é a distribuição de probabilidades X , restrita aos pacientes não curados.

$$f_T(t) = \frac{g(t) P(U_t = 1 \mid Y \leq t)}{P(U_t = 1)} = \frac{g(t) P(U_t = 1 \mid Y \leq t)}{1 - p_0}. \quad (22)$$

- De fato, (22) pode ser expressa como **”weighted distribution”** (Bayarri and DeGroot, 1992)

$$f_T(t) = \frac{w(t) g(t)}{E[w(Y)]}, \quad w(t) = P(U_t = 1 \mid Y \leq t). \quad (23)$$

Definições

Definição

Definimos **como mecanismo probabilístico de seleção de T** a distribuição condicional de Y dado $U_t = 1$.

- Esta definição simplesmente afirma que o mecanismo de seleção da va T é a distribuição de probabilidades X , restrita aos pacientes não curados.



$$f_T(t) = \frac{g(t) P(U_t = 1 | Y \leq t)}{P(U_t = 1)} = \frac{g(t) P(U_t = 1 | Y \leq t)}{1 - p_0}. \quad (22)$$

- De fato, (22) pode ser expressa como **”weighted distribution”** (Bayarri and DeGroot, 1992)

$$f_T(t) = \frac{w(t) g(t)}{E[w(Y)]}, \quad w(t) = P(U_t = 1 | Y \leq t). \quad (23)$$

Definições

Definição

Definimos **como mecanismo probabilístico de seleção de T** a distribuição condicional de Y dado $U_t = 1$.

- Esta definição simplesmente afirma que o mecanismo de seleção da va T é a distribuição de probabilidades X , restrita aos pacientes não curados.

-

$$f_T(t) = \frac{g(t) P(U_t = 1 | Y \leq t)}{P(U_t = 1)} = \frac{g(t) P(U_t = 1 | Y \leq t)}{1 - p_0}. \quad (22)$$

- De fato, (22) pode ser expressa como **”weighted distribution”** (Bayarri and DeGroot, 1992)

$$f_T(t) = \frac{w(t) g(t)}{E[w(Y)]}, \quad w(t) = P(U_t = 1 | Y \leq t). \quad (23)$$

Probabilidade individual

Definição

(Probabilidade individual de não cura sob o esquema de 1a. ativação)

Probabilidade do paciente morrer de uma célula danificada (clonogens), dado que $Y \leq t$ (“conditional personal non-cure rate”), é definida como

$$\gamma_{np}(t) = w(t) = P(U_t = 1 \mid Y \leq t). \quad (24)$$

$$\gamma_p = 1 - \gamma_{np} \Rightarrow \text{“personal probability”}$$

Teorema

“The crude cumulative distribution” e “the net survival at time t ” são dadas por

$$P(N_t = 1) = \frac{G(t) dA_N(s)}{ds} \Big|_{s=S(t)}, \quad (25)$$

$$P(N_t = N) = A_N[G(t)], \text{ respectivamente.}$$

Continuação: Rodrigues *et al.* (2011b)

Teorema

Sob a 1a. ativação e última ativação temos que

$$\gamma_{np} = \left. \frac{dA_N(s)}{ds} \right|_{s=S(t)}, \quad (26)$$

$$\gamma_{np} = \left. \frac{dA_N(s)}{ds} \right|_{s=G(t)}, \text{ respectivamente.}$$

Continuação...

Tabela: Mecanismo de seleção e probabilidade pessoal.

Mecanismo de seleção	1a. ativação	última ativação
$f_T(t)$	$\frac{g(t)}{1-p_0} \left\{ \frac{dA_N(s)}{ds} \Big _{s=S(t)} \right\}$	$\frac{g(t)}{1-p_0} \left\{ \frac{dA_N(s)}{ds} \Big _{s=G(t)} \right\}$
$S_T(t)$	$\frac{A_N[S(t)]-p_0}{1-p_0}$	$\frac{1-A_N[G(t)]}{1-p_0}$
$h_T(t)$	$\frac{g(t) \left\{ \frac{dA_N(s)}{ds} \Big _{s=S(t)} \right\}}{A_N[S(t)]-p_0}$	$\frac{g(t) \left\{ \frac{dA_N(s)}{ds} \Big _{s=G(t)} \right\}}{1-A_N[G(t)]}$
$\gamma_p(t)$	$1 - \frac{dA_N(s)}{ds} \Big _{s=S(t)}$	$1 - \frac{dA_N(s)}{ds} \Big _{s=G(t)}$

Modelo composto ou híbrido

A função de sobrevivência composta de Y no instante t é obtida através da composição do modelo unificado sob 1a. ativação com a última ativação (ver detalhes em Rodrigues *et al.*, 2011b) como segue:

$$S_T(t) = \frac{A_{N_1} \left\{ \frac{1 - A_{N_2}[G(t)]}{1 - p_0} \right\} - p_0}{1 - p_0}. \quad (27)$$

A função de densidade correspondente em (27) é dada por

$$f_T(t) = \frac{g(t)}{(1 - p_0)^2} \left\{ \frac{dA_{N_1}(s)}{ds} \Big|_{s = \frac{1 - A_{N_2}[G(t)]}{1 - p_0}} \right\} \left\{ \frac{dA_{N_2}(s)}{ds} \Big|_{s = G(t)} \right\}, \quad (28)$$

onde $p_0 = P[N_1 = 0] = P[N_2 = 0]$ e λ é a taxa de cura.

Modelo composto: “*Exponentiated exponential Poisson (EEP) distribution*” (Ristić and Nadarajah, 2010)

Seja N_1 o número de células iniciadas durante o período t , onde N_1 é um distribuição de Poisson truncada em zero com fdp

$$P[N_1 = n_1] = (e^\lambda - 1)^{-1} \frac{\lambda^{n_1}}{n_1!}. \quad (29)$$

Além de (29) suponhamos que $P[N_2 = \alpha]$ e o tempo de promoção $X_{i,j}, i = 1, \dots, N_1, j = 1, \dots, \alpha$, variáveis exponenciais com parâmetros de escala β . Logo, de (27), temos uma interpretação biológica a função de sobrevivência de EEP dada por

$$S_T(t) = (e^\lambda - 1)^{-1} \{ \exp[\lambda - \lambda(1 - e^{-\beta t})^\alpha] - 1 \}. \quad (30)$$

Modelo unificado: "The exponential-power series (EPS) distribution", Chahkandi and Ganjali (2009)

- Distribuição série de potências:

$$p_n(\alpha) = P(N = n; \alpha) = \frac{a_n \alpha^n}{A(\alpha)}, n = 0, 1, \dots, \quad (31)$$

onde $a_n > 0$, $A(\alpha) = \sum_n a_n \alpha^n$ and $\alpha > 0$. Esta família de distribuições inclui a binomial, Poisson, negative binomial, entre outras.

- A fgp é dada por $A_N(s; \alpha) = \frac{A(\alpha s)}{A(\alpha)}$ and $p_0 = \frac{a_0}{A(\alpha)}$.
- Supondo que o tempo de promoção é uma exponencial com parâmetro β , sob o esquema de 1a. ativação temos da Tabela 5 que

$$f_T(t; \theta) = \frac{\alpha \beta \exp(-\beta t) \frac{dA_N(s; \alpha)}{ds} |_{s=\exp(-\beta t)}}{A(\alpha) - a_0}, \quad (32)$$

onde $\theta = (\alpha, \beta)$.

- Estimaco de estes parmetros via mv e o algoritmo EM

Modelo unificado: "The exponential-power series (EPS) distribution", Chahkandi and Ganjali (2009)

- Distribuição série de potências:

$$p_n(\alpha) = P(N = n; \alpha) = \frac{a_n \alpha^n}{A(\alpha)}, n = 0, 1, \dots, \quad (31)$$

onde $a_n > 0$, $A(\alpha) = \sum_n a_n \alpha^n$ and $\alpha > 0$. Esta família de distribuições inclui a binomial, Poisson, negative binomial, entre outras.

- A fgp é dada por $A_N(s; \alpha) = \frac{A(\alpha s)}{A(\alpha)}$ and $p_0 = \frac{a_0}{A(\alpha)}$.
- Supondo que o tempo de promoção é uma exponencial com parâmetro β , sob o esquema de 1a. ativação temos da Tabela 5 que

$$f_T(t; \theta) = \frac{\alpha \beta \exp(-\beta t) \frac{dA_N(s; \alpha)}{ds} |_{s=\exp(-\beta t)}}{A(\alpha) - a_0}, \quad (32)$$

onde $\theta = (\alpha, \beta)$.

- Estimacoo destes parmetros via mv e o algoritmo EM

Modelo unificado: "The exponential-power series (EPS) distribution", Chahkandi and Ganjali (2009)

- Distribuição série de potências:

$$p_n(\alpha) = P(N = n; \alpha) = \frac{a_n \alpha^n}{A(\alpha)}, n = 0, 1, \dots, \quad (31)$$

onde $a_n > 0$, $A(\alpha) = \sum_n a_n \alpha^n$ and $\alpha > 0$. Esta família de distribuições inclui a binomial, Poisson, negative binomial, entre outras.

- A fgp é dada por $A_N(s; \alpha) = \frac{A(\alpha s)}{A(\alpha)}$ and $p_0 = \frac{a_0}{A(\alpha)}$.
- Supondo que o tempo de promoção é uma exponencial com parâmetro β , sob o esquema de 1a. ativação temos da Tabela 5 que

$$f_T(t; \theta) = \frac{\alpha \beta \exp(-\beta t) \frac{dA_N(s; \alpha)}{ds} |_{s=\exp(-\beta t)}}{A(\alpha) - a_0}, \quad (32)$$

onde $\theta = (\alpha, \beta)$.

Modelo unificado: "The exponential-power series (EPS) distribution", Chahkandi and Ganjali (2009)

- Distribuição série de potências:

$$p_n(\alpha) = P(N = n; \alpha) = \frac{a_n \alpha^n}{A(\alpha)}, n = 0, 1, \dots, \quad (31)$$

onde $a_n > 0$, $A(\alpha) = \sum_n a_n \alpha^n$ and $\alpha > 0$. Esta família de distribuições inclui a binomial, Poisson, negative binomial, entre outras.

- A fgp é dada por $A_N(s; \alpha) = \frac{A(\alpha s)}{A(\alpha)}$ and $p_0 = \frac{a_0}{A(\alpha)}$.
- Supondo que o tempo de promoção é uma exponencial com parâmetro β , sob o esquema de 1a. ativação temos da Tabela 5 que

$$f_T(t; \theta) = \frac{\alpha \beta \exp(-\beta t) \frac{dA_N(s; \alpha)}{ds} |_{s=\exp(-\beta t)}}{A(\alpha) - a_0}, \quad (32)$$

onde $\theta = (\alpha, \beta)$.

- Estimação destes parâmetros via mv e o algoritmo FM algorithm

Exemplo ilustrativo

- Primeiro conjunto de dados contém 213 observações e foi analisado por Adamidis and Loukas (1998), Kuş (2007) e Chahkandi and Ganjali (2009).
- Intervalos em horas entre falhas consecutivas de um sistema de ar condicionado em vôos do Boeing 720 (Proschan, 1963).
- Segundo conjunto de dados, apresentado por Cox and Lewis (1966), contém 109 observações sobre o número de dias entre "coal-mining" desastres.
- Os cálculos foram obtidos via R.

Exemplo ilustrativo

- Primeiro conjunto de dados contém 213 observações e foi analisado por Adamidis and Loukas (1998), Kuş (2007) e Chahkandi and Ganjali (2009).
- Intervalos em horas entre falhas consecutivas de um sistema de ar condicionado em vôos do Boeing 720 (Proschan, 1963).
- Segundo conjunto de dados, apresentado por Cox and Lewis (1966), contém 109 observações sobre o número de dias entre "coal-mining" desastres.
- Os cálculos foram obtidos via R.

Exemplo ilustrativo

- Primeiro conjunto de dados contém 213 observações e foi analisado por Adamidis and Loukas (1998), Kuş (2007) e Chahkandi and Ganjali (2009).
- Intervalos em horas entre falhas consecutivas de um sistema de ar condicionado em vôos do Boeing 720 (Proschan, 1963).
- Segundo conjunto de dados, apresentado por Cox and Lewis (1966), contém 109 observações sobre o número de dias entre “coal-mining” desastres.
- Os cálculos foram obtidos via R.

Exemplo ilustrativo

- Primeiro conjunto de dados contém 213 observações e foi analisado por Adamidis and Loukas (1998), Kuş (2007) e Chahkandi and Ganjali (2009).
- Intervalos em horas entre falhas consecutivas de um sistema de ar condicionado em vôos do Boeing 720 (Proschan, 1963).
- Segundo conjunto de dados, apresentado por Cox and Lewis (1966), contém 109 observações sobre o número de dias entre “coal-mining” desastres.
- Os cálculos foram obtidos via R.

Tabela: Estimativas dos parâmetros e ajuste das distribuições (ordenado segundo A^*).

Dados	Distribuição	Estimativas	W^*	A^*
Boeing ($n = 213$)	ECOMP	$(7.37 \times 10^{-3}, 1.37, 0.983)$	0.068	0.449
	EG	$(7.99 \times 10^{-3}, 0.429)$	0.074	0.484
	EP	$(7.49 \times 10^{-3}, 1.34)$	0.070	0.461
	Exponencial	1.07×10^{-2}	0.165	1.018
Disastres ($n = 109$)	ECOMP	$(1.55 \times 10^{-3}, 3.90, 1.08)$	0.067	0.432
	EP	$(1.65 \times 10^{-3}, 3.26)$	0.068	0.439
	EG	$(2.36 \times 10^{-3}, 0.619)$	0.068	0.480
	Exponencial	4.15×10^{-3}	0.070	0.658

Outras recentes distribuições de tempo de vida

- Distribuição "Classical Lehmann alternative" .
- Distribuição Exponencial Conway–Maxwell Poisson (ECOMP) .
- Distribuição "Exponentiated Weibull" (EW) .
- Família de distribuições de Kumaraswamy .
- Distribuição de "Kumaraswamy Weibull (KwW)".
- Distribuição "exponential-power series (EPS)".
- Família "Beta generalized (BG)".

Conclusões

- Usamos o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível para AS.
- O modelo unificado contém a maioria dos modelos recentemente introduzidos com interpretação biológica.
- Mais recentemente generalizamos o modelo unificado introduzindo o modelo composto unificado.
- Esperamos ter motivado uma linha de pesquisa que vem sendo desenvolvida com sucesso no Brasil e novas idéias são bem-vindas.

Conclusões

- Usamos o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível para AS.
- O modelo unificado contém a maioria dos modelos recentemente introduzidos com interpretação biológica.
- Mais recentemente generalizamos o modelo unificado introduzindo o modelo composto unificado.
- Esperamos ter motivado uma linha de pesquisa que vem sendo desenvolvida com sucesso no Brasil e novas idéias são bem-vindas.

Conclusões

- Usamos o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível para AS.
- O modelo unificado contém a maioria dos modelos recentemente introduzidos com interpretação biológica.
- Mais recentemente generalizamos o modelo unificado introduzindo o modelo composto unificado.
- Esperamos ter motivado uma linha de pesquisa que vem sendo desenvolvida com sucesso no Brasil e novas idéias são bem-vindas.

Conclusões

- Usamos o mecanismo de seleção proposto por Arellano-Valle *et al.* (2006) para formular um modelo flexível para AS.
- O modelo unificado contém a maioria dos modelos recentemente introduzidos com interpretação biológica.
- Mais recentemente generalizamos o modelo unificado introduzindo o modelo composto unificado.
- Esperamos ter motivado uma linha de pesquisa que vem sendo desenvolvida com sucesso no Brasil e novas idéias são bem-vindas.

Pesquisas em andamento

- **Análise de recorrência com variável terminal (censura informativa)**
- Análise de recorrência baseado no processo Poisson-Ponderado (Rodrigues *et al.*, 2011a).
- Latent cure rate modeling under repair system and threshold effect.

Pesquisas em andamento

- Análise de recorrência com variável terminal (censura informativa)
- Análise de recorrência baseado no processo Poisson-Ponderado (Rodrigues *et al.*, 2011a).
- Latent cure rate modeling under repair system and threshold effect.

Pesquisas em andamento

- Análise de recorrência com variável terminal (censura informativa)
- Análise de recorrência baseado no processo Poisson-Ponderado (Rodrigues *et al.*, 2011a).
- Latent cure rate modeling under repair system and threshold effect.

Continuação....

- Processo carcinogênico:

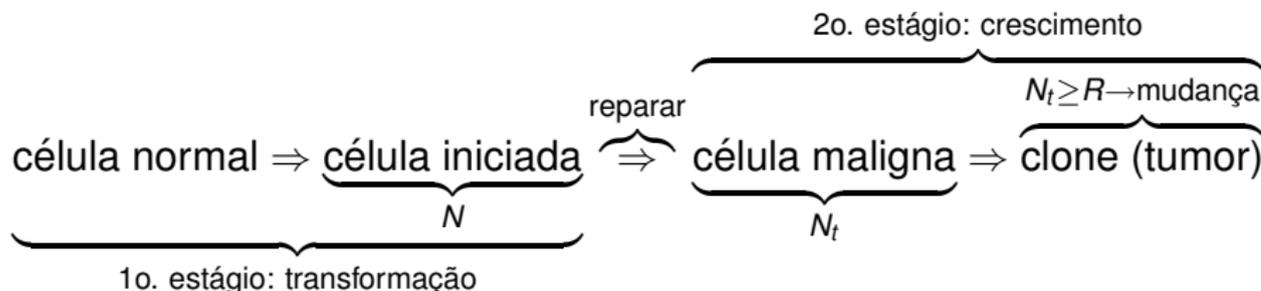


Figura: Processo carcinogênico em dois estágios

- **Problema:** Formular um modelo latente com taxa de cura e mecanismo de reparo de células expostas a irradiação para estimar a taxa de cura, volume do tumor e a primeira passagem do ponto crítico.

Continuação....

- Processo carcinogênico:

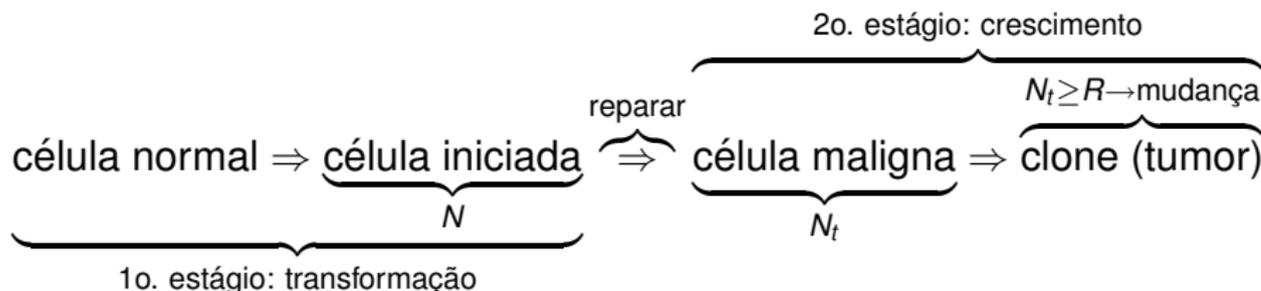


Figura: Processo carcinogênico em dois estágios

- **Problema:** Formular um modelo latente com taxa de cura e mecanismo de reparo de células expostas a irradiação para estimar a taxa de cura, volume do tumor e a primeira passagem do ponto crítico.

Mensagem final e agradecimento

A principal mensagem deste minicurso é como a teoria de processos de contagem é extremamente útil para unificar modelos estatísticos baseado em eventos históricos.

MUITO OBRIGADO

- Adamidis, K. and Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics and Probability Letters*, **39**, 35–42.
- Aitchison, J. and Dunsmore, I. R. (1975). *Statistical Prediction Analysis*. Cambridge University Press, Cambridge, UK.
- Andersen, P. (1985). Counting process models for life history data: A review. *Scandinavian Journal of Statistics*, **2**, 97–158.
- Arellano-Valle, R. B., Branco, M. D., and Genton, M. G. (2006). A unified view on skewed distributions arising from selections. *The Canadian Journal of Statistics*, **34**, 581–601.
- Balakrishnan, N. and Kozubowski, T. (2008). A class of weighted Poisson processes. *Statistics & Probability Letters*, **78**, 2346–2352.
- Bayarri, M. J. and DeGroot, M. (1992). A bad view of weighted distributions and selection models. In: *Bayesian Statistics 4*, Eds: J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, London: Oxford University Press, pages 17–29.
- Berkson, J. and Gage, R. P. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **47**(259), 501–515.

- Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society B*, **11**(1), 15–53.
- Cancho, V. G., Castro, M., and Rodrigues, J. (2012). A Bayesian analysis of the Conway-Maxwell-Poisson cure rate model. *Statistical Papers*, **53**, 165–176.
- Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, **53**, 4433–4440.
- Chen, M.-H., Ibrahim, J. G., and Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**(447), 909–919.
- Colosimo, E. A. and Giolo, S. R. (2006). *Análise de Sobrevida Aplicada*. Edgard Blücher, São Paulo–SP.
- Cook, R. and Lawless, J. (2002). Analysis of repeated events. *Statistical Methods in Medical Research*, **11**, 141–166.
- Cooner, F., Banerjee, S., Carlin, B. P., and Sinha, D. (2007). Flexible

- cure rate modeling under latent activation schemes. *Journal of the American Statistical Association*, **102**, 560–572.
- Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*. Methuen & Co. Ltd., London.
- de Castro, M., Cancho, V.-G., and Rodrigues, J. (2009). A Bayesian long-term survival model parametrized in the cure fraction. *Biometrical Journal*, **51**, 443–455.
- Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, volume I. Wiley, John Wiley & Sons, New York, 3rd edition.
- Hoggart, C. J. and Griffin, J. E. (2001). A Bayesian partition model for customer attrition. In E. I. George, editor, *Bayesian Methods with Applications to Science, Policy, and Official Statistics (Selected Papers from ISBA 2000, Crete, Greece)*, pages 61–70, Luxembourg. International Society for Bayesian Analysis, Eurostat. Proceedings of the Sixth World Meeting of the International Society for Bayesian Analysis.

- Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York, NY.
- Kim, S., Chen, M., and Dey, D. (2011). A new threshold regression model for survival data with a cure fraction. *Lifetime Data Analysis*, (17), 101–122.
- Klebanov, L. B., Rachev, S. T., and Yakovlev, A. (1993). A stochastic-model of radiation carcinogenesis: latent time distributions and their properties. *Mathematical Biosciences*, **18**(113), 51–75.
- Kuş, C. (2007). A new lifetime distribution. *Computational Statistics and Data Analysis*, **51**, 4497–4509.
- Lawless, J. F. (2002). *Statistical Models and Methods for Lifetime Data*. Wiley, New York, NY, 2nd edition.
- Lim, H., Liu, J., and Melzer-Lange, M. (2007). Comparison of methods for analysing recurrent event data: application to emergency department visits of firearms victims. *Accident Analysis and Prevention*, **39**(5), 290–299.

- Maller, R. A. and Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York, NY.
- Peng, Y. W., Taylor, J. M. G., and Yu, B. B. (2007). A marginal regression model for multivariate failure time data with a surviving fraction. *Lifetime Data Analysis*, **13**(3), 351–369.
- Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, **5**, 375–383.
- Ristić, M. M. and Nadarajah, S. (2010). A new lifetime distribution. Research Report 21, Probability and Statistics Group, School of Mathematics, The University of Manchester, Manchester, UK.
- Rodrigues, J., Cancho, V. G., de Castro, M., and Louzada-Neto, F. (2008). On the unification of the long-term survival models. *Statistics and Probability Letters*, **39**, 753–759.
- Rodrigues, J., Castro, M., Balakrishnan, N., and Cancho, V. (2011a). Destructive weighted Poisson cure rate model. *Lifetime Data Analysis*, **17**, 333–346.
- Rodrigues, J., Balakrishnan, N., Cordeiro, G. M., and de Castro, M. (2011b). Unified view on lifetime distributions arising from selection

- mechanisms. *Computation Statistics and Data Analysis*, doi: 10.1016/j.csda2011.06018.
- Rondeau, V., Schaffner, E., Corbiere, F., Gonzalez, J., and Mathoulin-Pelissier, S. (2011). Cure frailty models for survival data: Application to recurrences for breast cancer and to hospital readmi. *Statistical Methods in Medical Research : DOI: 10.1177/0962280210395521*, pages 1–18.
- Singpurwalla., N. D. (2007). *Reliability and Risk: A Bayesian Perspective*. Wiley, England.
- Tsodikov, A. D., Ibrahim, J. G., and Yakovlev, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, **98**(464), 1063–1078.
- Wang, M., Quin, J., and Chiang, C. (2001). Analyzing recurrent event data with informative censoring. *American Statistical Association Journal of the American Statistical Association*, **96**(455), –185.
- Wienke, A. (2010). *Frailty models in survival analysis*. Chapman & Hall, US.

Yakovlev, A. and Polig, E. (1996). A diversity of response displayed by stochastic model of radiation carcinogenesis allowing for cell death. *Mathematical Biosciences*, (132), 1–33.

Yakovlev, A. Y. and Tsodikov, A. D. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore.