Modelagem Conjunta de Dados Longitudinais e de Sobrevivência via Polinômios de Bernstein

Juliana Freitas de Mello e Silva

Orientador: Vinícius Diniz Mayrink
Co-orientador: Fábio Nogueira Demarqui

Departamento de Estatística
Universidade Federal de Minas Gerais

03 de agosto de 2018

Conteúdo

1. Introdução
   - Dados Longitudinais
   - Dados de Sobrevivência
2. Objetivos
3. Modelagem Conjunta de Dados Longitudinais e de Sobrevivência
4. Polinômios de Bernstein
   - PB para modelar a componente longitudinal
   - PB para modelar a componente de sobrevivência
5. Aplicações
   - Estudo de Crescimento
   - Tratamento Alternativo para Pacientes HIV Positivos
6. Próximos Passos
7. Referências

Outline

1 Introdução

Conceitos básicos - dados longitudinais

De acordo com Fitzmaurice et al (2012):

- dados longitudinais são caracterizados por apresentarem medidas repetidas de uma mesma variável ao longo do tempo;
- mais especificamente, uma *variável longitudinal* é aquela acompanhada ao longo do estudo com repetidas medições; sendo assim, cada elemento da amostra apresenta uma ou mais medidas dessa variável;
- dessa forma, torna-se evidente que os valores dessa variável variam com o tempo e, portanto, ela é tempo-dependente.

Conceitos básicos - dados longitudinais

Além disso,

- há *independência entre* elementos e *dependência intra*-elementos;
- os dados podem ser balanceados ou desbalanceados.

Em relação à correlação intra-indivíduos, Fitzmaurice et al. (2004) afirmam que:

- geralmente as correlações são positivas;
- a correlação diminui com o espaçamento do tempo;
- mesmo com um grande espaçamento, dificilmente ficam próximas de zero;
- de maneira análoga, raramente ficam próximas de 1.

Conceitos básicos - dados longitudinais

Acompanhar indivíduos ao longo do tempo tem particularidades como:

- informação sobre a evolução de cada indivíduo;
- permite comparações de mesma natureza, eliminando possíveis efeitos de confusão;
- dados faltantes (em muitos casos);
- erro de medição.

Modelo Linear Misto

É uma forma de modelar dados longitudinais, essa especificação contempla efeitos fixos e aleatórios:

$$\begin{aligned} W_i(t_{ij}) &= X_i(t_{ij}) + \epsilon_i(t_{ij}),\ i = 1, 2, \ldots, n \text{ e } j = 1, 2, \ldots, n_i \\ &= \boldsymbol{\alpha}_i^\top \mathbf{V}_i + \boldsymbol{\beta}^\top \mathbf{V}_i^* + \epsilon_i(t_{ij}). \end{aligned} \tag{1}$$

Em que,

- $W_i(t_{ij})$ é o valor observado da variável longitudinal no tempo t_{ij};
- assumindo erro de medição, $X_i(t_{ij})$ representa o valor real (não observado) dessa variável;
- $\epsilon_i(t_{ij}) \sim N(0, \sigma_W^2)$ é erro de medição;
- $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$ é vetor de coeficientes (efeitos fixos);
- $\mathbf{V}_i^* = (1, V_{i1}^*, \ldots, V_{ip}^*)^\top$ é vetor de covariáveis;
- $\boldsymbol{\alpha}_i$ é vetor de efeitos aleatórios a nível de indivíduo (indep. de ϵ_i);
- $\mathbf{V}_i = (1, V_{i1}, \ldots, V_{iq})^\top$ é vetor de covariáveis associado ao efeito aleatório.

Modelo Linear Misto

Sobre esse modelo, destaca-se que:

- os vetores $\mathbf{V}_i^*$ e $\mathbf{V}_i$ não necessariamente são os mesmos;
- geralmente $\boldsymbol{\alpha}_i \sim N_q(\mathbf{a}_i, \Sigma_{\alpha_i})$ para $i = 1, 2, \ldots, n$;
- pode-se introduzir uma estrutura de correlação a partir de Σ_{α_i};
- $\boldsymbol{\beta}$ fornece interpretações a nível de população enquanto que $\boldsymbol{\alpha}_i$ indica mudanças mais específicas, a nível de indivíduo.

Além disso, assumindo uma distribuição Normal para o erro de medição tem-se, dado os efeitos aleatórios, que $W_i(t_{ij})$ também é Normal, com

$$\mathbb{E}\left[W_i(t_{ij})|\boldsymbol{\alpha}_i\right] = X_i(t_{ij}) \text{ e } \mathbb{V}ar\left[W_i(t_{ij})|\boldsymbol{\alpha}_i\right] = \sigma_W^2.$$

Conceitos básicos - análise de sobrevivência

Alguns conceitos fundamentais são:

- é utilizada quando se deseja estimar o tempo até a ocorrência de certo evento de interesse (T);
- contém características intrínsecas como falhas/censuras e assimetria;
  - há três tipos de censura: **à direita**, à esquerda e intervalar;
  - mecanismo causador da censura: informativo ou **não informativo**;

$$\delta_i = \begin{cases} 1, \text{ se indivíduo } i \text{ é uma falha} \\ 0, \text{ se indivíduo } i \text{ é uma censura (à direita)} \end{cases}$$

- para cada indivíduo se observa (u_i, δ_i), para $i = 1, 2, \ldots, n$.

Funções básicas

Em análise de sobrevivência, considera-se as seguintes funções que são relacionadas entre si:

- função de sobrevivência

$$S(u) = \mathbb{P}(T > u) = 1 - \mathbb{P}(T \leq u) = 1 - F(u);$$

- função risco

$$h(u) = \lim_{du \to 0} \frac{\mathbb{P}(u \leq T < u + du | T \geq u)}{du};$$

- função de risco acumulado

$$H(u) = \int_0^u h(t)dt.$$

Função de verossimilhança

No caso de modelagem paramétrica, censura à direita e não informativa, a função de verossimilhança é dada por:

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{\Phi}; D) &\propto \prod_{i=1}^{n} f(u_i|\boldsymbol{\Phi})^{\delta_i} (S(u_i|\boldsymbol{\Phi}))^{1-\delta_i} \\
&= \prod_{i=1}^{n} h(u_i|\boldsymbol{\Phi})^{\delta_i} \exp\left\{-H(u_i|\boldsymbol{\Phi})\right\}. \qquad (2)
\end{aligned}
$$

Em que:

- $\boldsymbol{\Phi}$ é o vetor de parâmetros a ser estimado;
- D representa os dados disponíveis;
- n é o número de indivíduos;
- u_i é o i-ésimo tempo de sobrevivência observado;
- δ_i é a indicadora de falha/censura.

Modelo de riscos proporcionais

Na presença de covariáveis, pode-se utilizar o modelo proposto por Cox (1972):

$$h(u|\mathbf{Z}) = h_0(u)\exp\left\{\boldsymbol{\eta}^\top \mathbf{Z}\right\}, \text{ para } u > 0, \tag{3}$$

no qual $h_0(u)$ é a função risco de base, $\mathbf{Z}$ é o vetor de covariáveis e $\boldsymbol{\eta}$ é o vetor de coeficientes.

Uma propriedade importante é que a razão dos riscos de dois elementos da amostra são proporcionais no tempo. Isto é,

$$\frac{h(u|\mathbf{Z}_1)}{h(u|\mathbf{Z}_2)} = \frac{h_0(u)\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_1\}}{h_0(u)\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_2\}} = \frac{\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_1\}}{\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_2\}} = \exp\left\{\boldsymbol{\eta}^\top (\mathbf{Z}_1 - \mathbf{Z}_2)\right\},$$

em que $\mathbf{Z}_1$ e $\mathbf{Z}_2$ são vetores de covariáveis para dois elementos da amostra.

Modelo para variáveis tempo-dependentes

No caso de variáveis dependentes do tempo, pode-se usar uma extensão do modelo anterior:

$$h(u|\mathbf{Z}(u)) = h_0(u)\exp\left\{\boldsymbol{\eta}^\top \mathbf{Z}(u)\right\}, u > 0, \tag{4}$$

em que $\mathbf{Z}(u)$ é tal que apresenta mudanças ao longo do tempo.

Sobre esse modelo, tem-se que:

- é necessário o conhecimento dos valores de $\mathbf{Z}(u)$ para todo o tempo de seguimento (Klein & Moeschberger, 2003) (variáveis externas x internas);
- a propriedade de riscos proporcionais não é mais atendida:

$$\frac{h(u|\mathbf{Z}_1(u))}{h(u|\mathbf{Z}_2(u))} = \frac{h_0(u)\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_1(u)\}}{h_0(u)\exp\{\boldsymbol{\eta}^\top \mathbf{Z}_2(u)\}} = \exp\left\{\boldsymbol{\eta}^\top(\mathbf{Z}_1(u) - \mathbf{Z}_2(u))\right\},$$

em que $\mathbf{Z}_1$ e $\mathbf{Z}_2$ são vetores de covariáveis de dois elementos.

Outline

2 Objetivos

Objetivos

De forma geral, o objetivo deste trabalho é modelar conjuntamente dados longitudinais e de sobrevivência fazendo uso do Polinômio de Bernstein, por meio de inferência Bayesiana.

Com essa proposta, pretende-se encontrar estimativas mais precisas e robustas; obtendo também informações ricas e relevantes, ao aplicar essas metodologias à dados reais.

Outline

3 Modelagem Conjunta de Dados Longitudinais e de Sobrevivência

Exemplos

Como exemplos de dados longitudinais e de sobrevivência, tem-se:

- estudos com pacientes soropositivos
  - variável longitudinal: contagem de células CD4;
  - variável de sobrevivência: tempo até a progressão para AIDS ou óbito.

- estudos envolvendo pacientes com câncer
  - variável longitudinal: qualidade de vida desses pacientes (índice), tamanho do tumor;
  - variável de sobrevivência: tempo até o óbito ou recidiva.

Inicialmente, ao tentar analisar dados longitudinais e de sobrevivência se fez uso, principalmente, de duas abordagens.

- Modelagem separada - priorizava-se uma das duas variáveis:
  - no caso da variável longitudinal, pode-se usar um Modelo Linear Misto;
  - para a sobrevivência, modelo de Cox para variáveis tempo-dependentes.

- Modelo de Dois-Estágios (proposto por Tsiatis et al. (1995));
  - primeiro estágio: modela-se a variável longitudinal através de uma função trajetória;
  - segundo estágio: essa estimativa é imputada no modelo de Cox para variável tempo-dependente.

Considerações sobre as abordagens

Sobre a **modelagem separada**:

- variável longitudinal
  - nos Modelos Lineares Mistos, a informação de sobrevivência não é considerada;
- variável de sobrevivência
  - como já mencionado, é requerido o conhecimento dos valores da variável longitudinal para todo o tempo de seguimento;
  - dados longitudinais geralmente apresentam erro, assim a utilização dessa informação diretamente no modelo de Cox pode levar à estimativas viciadas (Ibrahim et al., 2010).

Considerações sobre as abordagens

Em relação ao **modelo de dois estágios**:

- ao atribuir um modelo para a variável longitudinal, pode-se contornar a questão de erro de medição e dados faltantes.

Além disso, de acordo com Wu et al. (2012):

- a estimativa da componente longitudinal não diferencia ocorrência de censura ou evento de interesse;
- a incerteza associada à estimativa da componente longitudinal não é levada em consideração no segundo estágio;
- essa modelagem subestima os parâmetros que ligam os dois sub-modelos;
- o vício relacionado à componente longitudinal depende da força da associação entre as duas variáveis, enquanto que para a sobrevivência, esse vício depende do erro de medição.

Modelagem conjunta

Considerando o que já foi exposto, tem-se que a modelagem conjunta:

- foi motivada por estudos com pacientes HIV positivos (Ibrahim et al., 2010);
- é utilizada em casos nos quais há dados do tipo "tempo até evento", além de uma ou mais variáveis longitudinais;
- é indicada principalmente para o caso de variáveis longitudinais do tipo interna (ou endógena);
- fornece estimativas mais precisas e robustas ao utilizar toda informação disponível simultaneamente.

Modelagem conjunta

Além do mais, ao utilizar essa modelagem é possível obter informações do tipo:

- trajetória da variável longitudinal para cada indivíduo;
- sobrevivência de cada paciente, e;
- a relação entre essas duas variáveis.

Assim, a modelagem conjunta é dada em três passos:

1. especificação da chamada função trajetória;
2. sub-modelo para a componente longitudinal;
3. sub-modelo para a componente de sobrevivência.

Modelagem conjunta - componente longitudinal

Atribui-se um modelo para o processo longitudinal com o intuito de representar a trajetória "real" (não necessariamente observada) dessa componente:

$$X_i(t) = \mathbf{f(t)}^\top \boldsymbol{\alpha_i} \;, \; i = 1, 2, \ldots, n. \tag{5}$$

Nesse caso, $\mathbf{f(t)}^\top$ é um vetor de funções do tempo t e $\boldsymbol{\alpha}_i$ é vetor de efeitos aleatórios. Esta forma inclui:

- especificação polinomial;
- *splines*;
- outras funções não-lineares (em t).

A função trajetória deve ser especificada levando em consideração, tanto a fidelidade ao processo biológico quanto o conceito de parcimônia.

Modelagem conjunta - componente longitudinal

Considerando a especificação da função trajetória $X_i(u)$, os dados longitudinais observados são conectados com a função da trajetória da forma:

$$
\begin{aligned}
W_i(t_{ij}) &= X_i(t_{ij}) + \epsilon_i(t_{ij}),\ i = 1, 2, \ldots, n \text{ e } j = 1, 2, \ldots, n_i \\
&= \mathbf{f(t)}^\top \boldsymbol{\alpha_i} + \epsilon_i(t_{ij}). \qquad (6)
\end{aligned}
$$

Em que:

- $W_i(t_{ij})$ é o valor observado da variável longitudinal do item i no tempo t_{ij};
- $\epsilon_i \sim N(0, \sigma_W^2)$ é o erro de medida;
- ϵ_i é independente de $\boldsymbol{\alpha_i}$;
- pode-se assumir estrutura de correlação em ϵ_i.

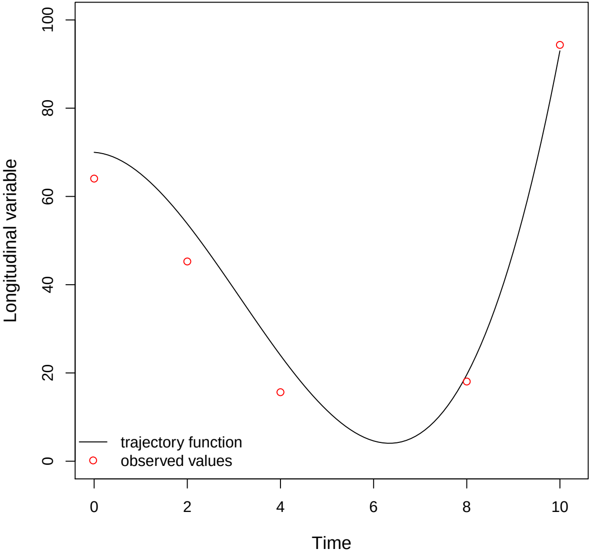


Figura: Representação da função trajetória e dos valores observados.

Modelagem conjunta - componente de sobrevivência

O seguinte modelo é atribuído à função risco:

$$\begin{aligned} h_i(u_i) &= \lim_{du \to 0} \frac{\mathbb{P}(u_i \leq T_i < u_i + du | T_i \geq u_i, X_i^H(u_i), \mathbf{Z}_i)}{du} \\ &= h_0(u_i) \exp\left\{\gamma X_i(u_i) + \boldsymbol{\eta}^\top \mathbf{Z}_i\right\}, \; i = 1, 2, \ldots, n, \end{aligned} \tag{7}$$

Em que:

- $X_i^H(u_i) = \{X_i(t),\ 0 \leq t < u_i\}$ é o histórico do processo longitudinal até o tempo u_i;
- $\mathbf{Z}_i$ vetor de covariáveis medidas no tempo inicial;
- $h_0(u_i)$ é a função risco de base;
- γ mede a relação entre a componente longitudinal $(X_i(u_i))$ e o tempo até o evento;
- $\boldsymbol{\eta}^\top$ é vetor de coeficientes.

Modelagem conjunta - função de verossimilhança

Sejam X e Y duas variáveis aleatórias contínuas. Sabe-se que a função densidade conjunta pode ser escrita da forma: $f_{(X,Y)}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)$.

A partir disso, a função de verossimilhança no caso de censura à direita e não informativa é dada por:

$$
\begin{aligned}
L(\boldsymbol{\Phi}; D) &= \prod_{i=1}^{n} \int \left\{ h(u_i)^{\delta_i} \exp\left\{-H(u_i)\right\} \left(\prod_{j=1}^{n_i} p(W_i(t_{ij})|(\boldsymbol{\alpha}_i, \sigma^2)) \right) p(\boldsymbol{\alpha}_i|.) \right\} d\boldsymbol{\alpha}_i \\
&= \prod_{i=1}^{n} \int \left[h_0(u_i) \exp\{\gamma X_i(u_i) + \boldsymbol{\eta}^\top \mathbf{Z}_i\} \right]^{\delta_i} \exp\left\{ - \int_0^{u_i} h_0(w) \exp\left\{ \gamma X_i(w) + \boldsymbol{\eta}^\top \mathbf{Z}_i \right\} dw \right\} \\
&\quad \frac{1}{(2\pi\sigma^2)^{n_i/2}} \exp\left\{ - \sum_{j=1}^{n_i} \frac{\{W_i(t_{ij}) - X_i(t_{ij})\}^2}{2\sigma^2} \right\} p(\boldsymbol{\alpha}_i|.) d\alpha_i,
\end{aligned}
$$

A seguir, será apresentada forma de aproximar as quantidades de interesse.

Outline

4 Polinômios de Bernstein

Aproximações por polinômios

De forma geral, algumas características atrativas de polinômios são:

- podem ser representados de maneira simples, uma vez que é possível escrevê-los em forma de somatório - isso facilita cálculos como derivadas, vetor gradientes, entre outros (Osman & Ghosh, 2012);

- além disso, polinômios são funções bem comportadas e infinitamente deriváveis (de Figueiredo, 1996).

Polinômios de Bernstein

Sobre os Polinômios de Bernstein, especificamente, tem-se que:

- foram propostos por Sergei Natanovich Bernstein, em 1913;
- surgiram a partir de uma demonstração simples para um caso especial do Teorema de Weiestrass baseado na teoria de probabilidades (Lorentz, 1986; Bernstein, 1913).

O Teorema de Weierstrass é formalmente definido como (Bartle & Sherbert, 2011):

Theorem (Weierstrass Approximation Theorem)

Let $I = [a, b]$ *and let* $f : I \to \mathbb{R}$ *be a continuous function. If* $\varepsilon > 0$ *is given, then there exists a polynomial function* p_ε *such that* $|f(x) - p_\varepsilon(x)| < \varepsilon$ *for all* $x \in I$*.*

A ideia da demostração tida por Bernstein foi a seguinte:

- considere um evento A tal que a probabilidade de sua ocorrência seja x (isto é, $\mathbb{P}(A) = x$);

- em seguida, assuma que serão realizados m ensaios desse experimento de forma que a quantidade $f(k/m)$ será paga a um jogador hipotético se o evento A ocorrer k vezes;

- definindo uma variável aleatória como o número de sucessos (ocorrer evento A) em m tentativas, fica claro que se tem uma distribuição $Binomial(m, x)$.

Portanto, a probabilidade de que o evento A ocorra k vezes será

$$\binom{m}{k} x^k (1-x)^{m-k},$$

enquanto que o valor esperado da quantidade a ser paga é

$$E_m(x) = \sum_{k=0}^{m} f\left(\frac{k}{m}\right) \binom{m}{k} x^k (1-x)^{m-k}. \tag{8}$$

A partir disso, Bernstein demonstrou que $|f(x) - E_m(x)| < \varepsilon$, para um $\varepsilon > 0$. Ou seja,

$$f(x) = \lim_{m\to\infty} \sum_{k=0}^{m} f\left(\frac{k}{m}\right) \left(\begin{array}{c} m \\ k \end{array}\right) x^k (1-x)^{m-k}.$$

Assim, o Polinômio de Bersntein de grau m para aproximar a função f é dado por $E_m(x)$ em (8) e a quantidade

$$B_{(k,m)}(x) = \left(\begin{array}{c} m \\ k \end{array}\right) x^k (1-x)^{m-k} \tag{9}$$

é chamada de base de Bernstein.

Considerações sobre o PB

Algumas considerações importantes:

- não é necessário saber muito sobre a função f para aproximá-la via PB;

- matematicamente, a aproximação pelo PB requer o conhecimento (exato ou aproximado) da função f em $m+1$ pontos de seu domínio (sendo m o grau do PB);

- as bases de Bernstein podem ser vistas como pesos, uma vez que $0 < B_{(k,m)}(x) < 1$ para todo $k = 0, 1, \ldots, m$ e

$$\sum_{k=0}^{m} B_{(k,m)}(x) = \sum_{k=0}^{m} \binom{m}{k} x^k (1-x)^{m-k} = 1.$$

Considerações sobre o PB

Além disso,

- como visto em Lorentz (1986) e Osman & Ghosh (2012), os PB se destacam por sua capacidade de apresentarem a melhor aproximação, no sentido de preservar a forma da função que está sendo aproximada;
- são bastante utilizados para aproximar funções densidade (Osman & Ghosh, 2012);
- ao impor restrições em sua formulação, é possível obter propriedades de funções monótonas, côncava/convexa, dentre outras.

PB para modelar a componente longitudinal

Para modelar a variável longitudinal, seguiu-se o que foi proposto em Wang & Ghosh (2013).

Tal proposta consistiu em utilizar os PB para modelar parte da *função trajetória* - que representa o valor real da variável longitudinal - da seguinte forma:

$$X_i\left(t_{ij}\right) = \sum_{k=1}^{m} b_{k,m}\left(t_{ij}\right)\xi_{i,k},$$

em que

$$b_{k,m}\left(t_{ij}\right) = \left(\begin{array}{c} m-1 \\ k-1 \end{array}\right)\left(\frac{t_{ij}}{\tau}\right)^{k-1}\left(1-\frac{t_{ij}}{\tau}\right)^{m-k}$$

é a base de Bernstein e $\boldsymbol{\xi}_i = (\xi_{i,1}, \xi_{i,2}, \ldots, \xi_{i,m})$ é um vetor de coeficientes (a nível de indivíduo) associado à base $\mathbf{b}_m(.) = (b_{1,m}(.), b_{2,m}(.), \ldots, b_{m,m}(.))^\top$ podendo assumir qualquer valor real.

PB para modelar a componente longitudinal

Mais explicitamente, o modelo proposto por Wang & Ghosh (2013) é dado por:

$$
\begin{aligned}
W_i\left(t_{ij}\right) &= X_i\left(t_{ij}\right) + \epsilon_i\left(t_{ij}\right), i = 1, 2, \ldots, n \text{ and } j = 1, 2, \ldots, n_i \qquad (10)\\
&= \sum_{k=1}^{m} b_{k,m}\left(t_{ij}\right)\xi_{i,k} + \epsilon_i\left(t_{ij}\right)\\
&= \sum_{k=1}^{m} \binom{m-1}{k-1}\left(\frac{t_{ij}}{\tau}\right)^{k-1}\left(1-\frac{t_{ij}}{\tau}\right)^{m-k}\xi_{i,k} + \epsilon_i\left(t_{ij}\right),
\end{aligned}
$$

em que

- $W_i\left(t_{ij}\right)$ é o valor *observado* da variável longitudinal no tempo t_{ij};
- τ é o tempo (máximo) de seguimento, de forma que $t_{ij}/\tau \in (0, 1)$;
- $X_i(t_{ij})$ é o valor "real" da variável longitudinal;
- $\epsilon_i\left(t_{ij}\right) \sim N(0, \sigma_W^2)$ é o erro de medida.

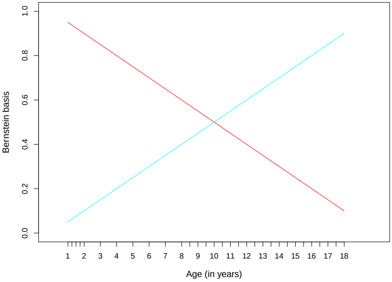
PB para modelar a componente longitudinal

Vale ressaltar que é possível introduzir informação proveniente de covariáveis de uma forma simples.

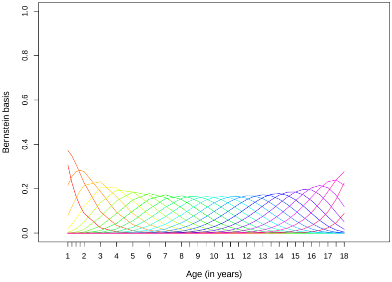
$$\begin{aligned} W_i\left(t_{ij}\right) &= X_i\left(t_{ij}\right) + \epsilon_i\left(t_{ij}\right),\, i = 1, 2, \ldots, n \text{ e } j = 1, 2, \ldots, n_i \\ &= \sum_{k=1}^{m} b_{k,m}\left(t_{ij}\right) \xi_{i,k} + \boldsymbol{\beta}^\top \mathbf{V}_i^* + \epsilon_i\left(t_{ij}\right). \end{aligned}$$

Aqui, $\mathbf{V}_i^* = (V_{i1}^*, V_{i2}^*, \ldots, V_{iq}^*)^\top$ é o vetor de covariáveis medidas na linha de base para a pessoa i e $\boldsymbol{\beta}$ é o vetor de coeficientes associado à $\mathbf{V}_i^*$.

Tanto no trabalho original quanto no atual, assumiu-se que $\boldsymbol{\xi}_i \sim N_m(\boldsymbol{\mu}_{\boldsymbol{\xi}_i}, \Sigma_{\boldsymbol{\xi}_i})$. Consequentemente, tem-se que a função $X_i(.)$ é aproximada por um Processo Gaussiano.



(a) $m = 2.$



(b) $m = 24.$

Figura: Base de Bernstein.

Considerações

Na especificação proposta por Wang & Ghosh (2013), assume-se que

- o valor do grau do PB (m) deve ser maior do que 1 a fim se esquivar de um Processo Gaussiano degenerado;
- além disso, m deve ser menor que o número máximo de medidas para evitar problemas de multicolinearidade;
- logo $m \in [2, \max_i n_i)$, em que n_i é o número de medidas de cada item da amostra.

Assim, as quantidades desconhecidas a serem estimadas são $\boldsymbol{\mu}_{\boldsymbol{\xi}_i}$, $\Sigma_{\boldsymbol{\xi}_i}$ $(i = 1, 2, \ldots, n)$ e σ_W^2. Caso seja necessário, pode-se impor restrições no vetor de médias $\boldsymbol{\mu}_{\boldsymbol{\xi}_i}$.

PB para modelar a componente de sobrevivência

No contexto de sobrevivência,

- os PB serão utilizados para modelar a função risco de base, $h_0(.)$, do modelo de riscos proporcionais - com ou sem variáveis dependentes do tempo;
- o trabalho base a ser seguido é o de Osman & Ghosh (2012), que consideraram, sobretudo, o caso de curvas de sobrevivência que se cruzam ao longo do tempo.

PB para modelar a componente de sobrevivência

Primeiro, Osman & Ghosh (2012) propuseram uma estrutura para modelar a função risco, como segue:

$$h_m(u, \boldsymbol{\gamma}) = \sum_{k=1}^{m} \gamma_k \mathbf{g}_{m,k}(u) = \boldsymbol{\gamma}^\top \mathbf{g}_m(u),\ 0 \le u < \infty, \quad (11)$$

em que

- m é um inteiro positivo;
- $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_m)^\top$ é um vetor de coeficientes tal que $\gamma_k \ge 0$ $(k = 0, 1, \ldots, m)$;
- $\mathbf{g}_m(u) = (g_{m,0}(u), g_{m,1}(u), \ldots, g_{m,m}(u))^\top$ é um vetor de funções de base satisfazendo $g_{m,k}(.) \ge 0$ e $\int_0^\infty g_{m,k}(u)du < \infty$, para todo $k \in \{0, 1, \ldots, m\}$.

PB para modelar a componente de sobrevivência

Consequentemente, o modelo para a função de risco acumulado será

$$\begin{aligned} H_m(u,\boldsymbol{\gamma}) &= \int_0^u h_m(v,\boldsymbol{\gamma})dv = \int_0^u \sum_{k=1}^{m} \gamma_k \mathsf{g}_{m,k}(v)dv \\ &= \sum_{k=1}^{m} \gamma_k \mathsf{G}_{m,k}(u) = \boldsymbol{\gamma}^\top \mathbf{G}_m(u),\ 0 \leq u < \infty, \end{aligned} \tag{12}$$

em que $\mathbf{G}_m(u) = (G_{m,0}(u), G_{m,1}(u), \ldots, G_{m,m}(u))$ é um vetor de funções de base no qual cada componente é igual a $G_{m,k}(u) = \int_0^u g_{m,k}(v)dv$.

PB para modelar a componente de sobrevivência

Em seguida, considere a aproximação pelo PB para a função de risco acumulado $H(.)$:

$$\begin{aligned}\tilde{H}(u;m) &= \sum_{k=0}^{m} H\left(\frac{k}{m}\tau\right)\binom{m}{k}\left(\frac{u}{\tau}\right)^{k}\left(1-\frac{u}{\tau}\right)^{m-k} \\ &= \sum_{k=1}^{m} H\left(\frac{k}{m}\tau\right)\binom{m}{k}\left(\frac{u}{\tau}\right)^{k}\left(1-\frac{u}{\tau}\right)^{m-k}, \qquad (13)\end{aligned}$$

em que

- m é o grau do PB;
- $\tau = \{u : S(u) = 0\}$;
- $S(.)$ é a função de sobrevivência.

Pelo Teorema de Weierstrass é garantido que há convergência uniforme em $[0, \tau]$. Ou seja, $H(.) = \lim_{m\to\infty} \tilde{H}(.; m)$.

PB para modelar a componente de sobrevivência

Derivando a expressão (13), obtém-se uma aproximação para a função risco $h(.)$:

$$\begin{aligned} \tilde{h}(u;m) &= \frac{\partial}{\partial u}\tilde{H}(u;m) \qquad (14) \\ &= \sum_{k=1}^{m}\left\{H\left(\frac{k}{m}\tau\right) - H\left(\frac{k-1}{m}\tau\right)\right\}\frac{f_\beta\left(u/\tau;k,m-k+1\right)}{\tau}, \end{aligned}$$

em que $f_\beta\left(u/\tau;k,m-k+1\right)$ é a função de densidade de uma $Beta(k,m-k+1)$ no ponto u/τ.

Analogamente, vale pelo Teorema de Weierstrass que $\tilde{h}(.;m) \longrightarrow h(.)$ uniformemente no intervalo $(0,\tau]$ à medida que $m \longrightarrow \infty$.

PB para modelar a componente de sobrevivência

Por fim, Osman & Ghosh (2012) conectaram os modelos propostos (em (11) e (12)) com as aproximações pelo PB (em (14) e (13)). Isto é:

$$\begin{aligned} h_m(u, \boldsymbol{\gamma}) &= \sum_{k=1}^{m} \gamma_k \mathbf{g}_{m,k}(u) \\ &\approx \sum_{k=1}^{m} \left\{ H\left(\frac{k}{m}\tau\right) - H\left(\frac{k-1}{m}\tau\right) \right\} \frac{f_\beta\left(u/\tau; k, m-k+1\right)}{\tau}. \end{aligned}$$

PB para modelar a componente de sobrevivência

Portanto,

$$
\begin{aligned}
G_{m,k}(u) &= \int_0^u g_{m,k}(v)dv \\
&= \int_0^u \frac{f_\beta\left(v/\tau; k, m-k+1\right)}{\tau} dv \\
&= \int_0^u f_\beta\left(v/\tau; k, m-k+1\right) d(v/\tau).
\end{aligned}
$$

Ou seja, $G_{m,k}(.)$ é a função de distribuição acumulada de uma distribuição $Beta(k, m-k+1)$.

A quantidade desconhecida a ser estimada é o vetor de coeficientes γ.

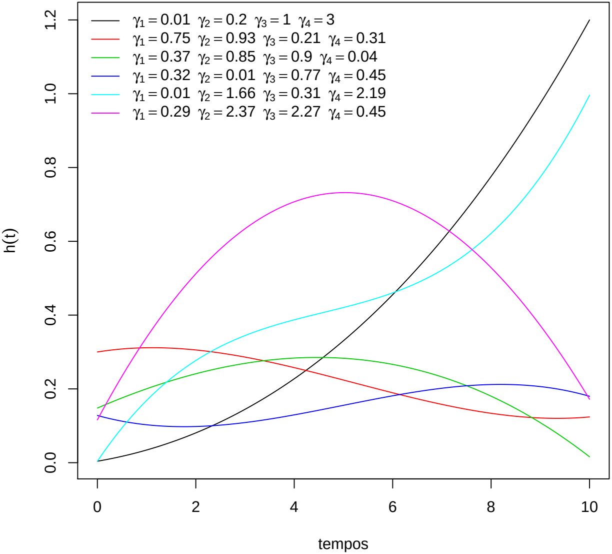


Figura: Ilustração da flexibilidade PB para modelar a função risco, $m = 4$.

Considerações

É importante ressaltar que:

- ao impor as restrições $\gamma_k \geq 0$ e $g_{m,k}(.) \geq 0$, para $k = 0, 1, \ldots, m$, assegura-se que o modelo para a função de risco acumulado provê uma forma não-decrescente;

- matematicamente, $h_m(.)$ converge para a função $h(.)$ quando $m \to \infty$; contudo, na prática, considera-se um valor finito para m;

- verificou-se através de estudos de simulação que, em alguns casos, um valor relativamente baixo de m fornece uma boa aproximação.

Considerações

Ainda,

- Osman & Ghosh (2012) sugerem utilizar $m = \lceil \sqrt{n} \rceil$ no contexto de sobrevivência;
- Wang & Ghosh (2013) propuseram um critério para escolher esse valor, no modelo para a variável longitudinal.

Outline

5 Aplicações

Aplicação

Serão apresentadas duas aplicações:

1. dados de crescimento, baseado no trabalho de Wang & Ghosh (2013);
2. dados de pacientes soropositivos, seguindo Guo & Carlin (2004).

Para ambas aplicações utilizou-se os *softwares* `JAGS`, `R` e o pacote `R2jags`. As especificações do MCMC foram:

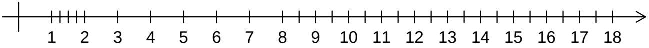
- *burn-in*: 50000;
- *lag*: 10;
- amostra *a posteriori*: 5000.

Para comparação de modelos, foram usados os critérios DIC, -2 LPML e -2 WAIC. Assim, quanto menor esses valores, melhor é o modelo.

Estudo de Crescimento - descrição

- consiste em alturas de 93 pessoas (39 meninos e 54 meninas) medidas 31 vezes ao longo de 18 anos;
- não há valores faltantes;
- todas as pessoas foram medidas 31 vezes, nos mesmos pontos no tempo.

Além disso, a escala do tempo era a idade de cada pessoa, e as alturas foram medidas nos seguintes pontos:



Estudo de Crescimento - descrição do modelo

Em Wang & Ghosh (2013):

- um modelo para cada sexo;

$$W_i\left(\frac{t_{ij}}{\tau}\right) = \sum_{k=1}^{m} b_{k,m}\left(\frac{t_{ij}}{\tau}\right)\xi_{i,k} + \epsilon_i\left(\frac{t_{ij}}{\tau}\right).$$

Trabalho atual:

- uma única estrutura englobou ambos os sexos;
- foi incluído um termo de intercepto;
- efeito o sexo variou ao longo do tempo;

$$W_i\left(\frac{t_{ij}}{\tau}\right) = \beta_0 + \sum_{k=1}^{m} b_{k,m}\left(\frac{t_{ij}}{\tau}\right)\xi_{i,k} + \beta_j\ \text{Sexo}_i + \epsilon_i\left(\frac{t_{ij}}{\tau}\right).$$

Estudo de Crescimento

As distribuições *a priori* para os parâmetros desconhecidos foram:

- $\beta_0 \sim N(60, 10^2)$;
- $\boldsymbol{\beta} \sim N_{31}(\mathbf{0}, 20^2\mathbb{I}_{31})$;
- $\boldsymbol{\xi}_i \sim N_m(\boldsymbol{\mu}_{\boldsymbol{\xi}_i}, \sigma^2_{\xi}\mathbb{I}_m)$;
- $\mu_{\boldsymbol{\xi}_{i,1}} \sim Ga(1, 1/10)$, $(\mu_{\boldsymbol{\xi}_{i,k}} - \mu_{\boldsymbol{\xi}_{i,k-1}}) \sim Ga(1, 1/10)$, $k = 2, 3, \ldots, m$;
- $(1/\sigma^2_{\xi}) \sim Ga(0.1, 0.1)$ e $\sigma^2_i \sim Ga(0.1, 0.1)$.

Estudo de Crescimento

Tabela: Medidas de comparação variando o grau do PB (m).

m	DIC	-2 LPML	-2 WAIC
2	13383.5917	13460.7712	13440.7828
7	9294.6963	8483.4376	8397.8259
10	8852.1138	7420.0102	7258.9831
14	7818.8343	6289.0428	6075.1684
17	7495.2776	5627.2639	5386.2563
21	6508.6014	4998.2957	4664.6427
24	**6117.4511**	**4577.4924**	**4195.4561**

Estudo de Crescimento

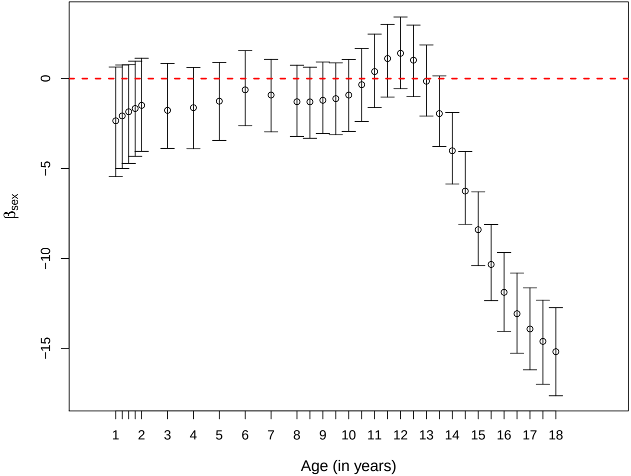
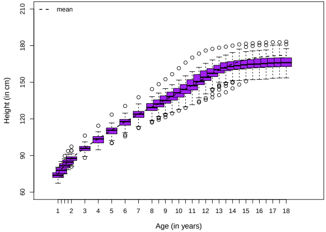
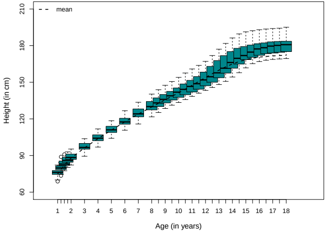


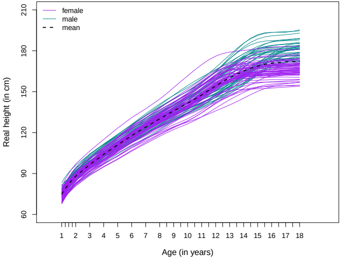
Figura: Mediana *a posteriori* e intervalo HPD para β_j, $j = 1, 2, \ldots, 31$, $m = 24$.



(a) Meninas.



(b) Meninos.



(c) Gráfico de perfis - dados “reais”.

Tratamento Alternativo para Pacientes HIV_+ - descrição

Objetivo principal: comparar dois tratamentos alternativos para pacientes HIV positivos que falharam ou eram intolerantes à zidovudine (AZT).

Para isso, um ensaio clínico com $n = 467$ pacientes foi realizado. As variáveis foram:

- tempo até o óbito (com 40,26% de falha);
- contagem de célula CD4;

- droga usada no tratamento (zalcitabine - ddC, didanosine - ddI);
- gênero (feminino, masculino);
- doenças oportunistas no início do estudo (sim, não);
- AZT (falhou, intolerante).

Os tempos de medição foram pré-especificados em: linha de base, 2º mês, 6º mês, 12º mês e 18º mês. Aqui trabalharemos na escala de raiz quadrada.

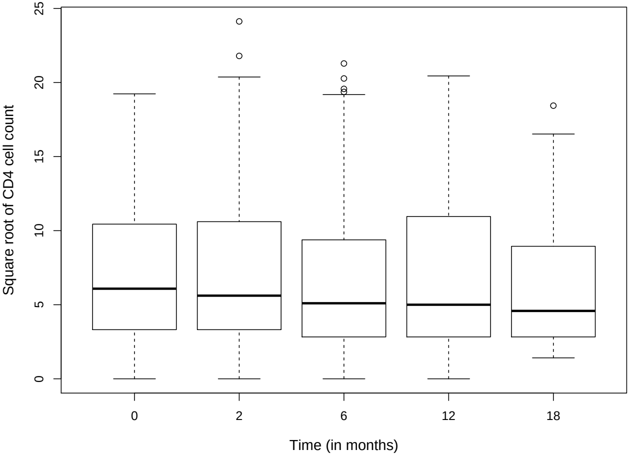


Figura: Boxplot da raiz quadrada da contagem de células CD4 para cada tempo de medição.

Tratamento Alternativo para Pacientes HIV_+ - valores faltantes

Tabela: Número de medidas observadas e possíveis, para cada tempo de medição.

Tempo	Num. de med. observadas (%)	Num. de med. possíveis (%)
Linha de base	467 (100%)	467 (100%)
2º mês	368 (78.80%)	453 (97.00%)
6º mês	310 (66.38%)	404 (86.51%)
12º mês	226 (48.39%)	318 (68.09%)
18º mês	37 (7.92%)	58 (12.42%)

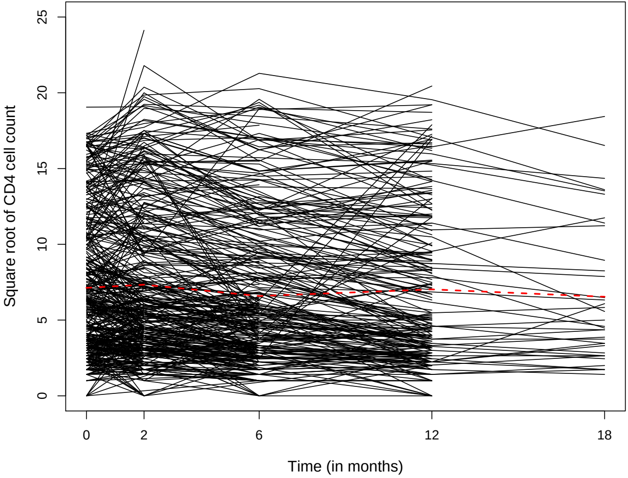


Figura: Gráfico de perfis mostrando o comportamento da contagem observada de células CD4 para cada paciente.

Tratamento Alternativo para Pacientes HIV$_+$

Para esse banco de dados, foram feitas duas comparações.

- formas de lidar com a imputação automática do `JAGS`: nenhuma, **parcial** e total;
- diferentes especificações de modelos (foram retiradas 14 observações)
  - componente longitudinal: Normal, PB (com graus $m = 2, 4$);
  - componente de sobrevivência: Weibull, MEP (número de intervalos foram $m = 4, 10, 16, 22$), PB (os graus foram $m = 4, 10, 16, 22$).

Ressalta-se que todos os modelos foram ajustados de forma **separada**. As especificações são detalhadas a seguir.

Sub-modelo longitudinal

Para $i = 1, 2, \ldots, n$ e $j = 1, 2, \ldots, n_i$, as especificações de modelos foram:

- modelo Normal

$$\begin{aligned} W_i(t_{ij}) &= \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij} \times \text{ Droga}_i + \beta_3 \text{ Gênero}_i + \beta_4 \text{ PrevIO}_i \\ &+ \beta_5 \text{ AZT}_i + \alpha_{0i} + \alpha_{1i} t_{ij} + \epsilon_i(t_{ij}) \end{aligned}$$

- aproximação via PB ($m = 2, 4$)

$$\begin{aligned} W_i(t_{ij}) &= \beta_0 + \beta_1 t_{ij} + \beta_2 t_{ij} \times \text{ Droga}_i + \beta_3 \text{ Gênero}_i + \beta_4 \text{ PrevIO}_i \\ &+ \beta_5 \text{ AZT}_i + \sum_{k=1}^{m} b_{k,m}(t_{ij}) \xi_{i,k} + \epsilon_i \left(\frac{t_{ij}}{\tau} \right) \end{aligned}$$

Sub-modelo de sobrevivência

Para $i = 1, 2, \ldots, n$:

- Weibull para os tempo de falha

$$\begin{aligned} T_i &\sim Weib(1, \zeta_i) \\ \zeta_i &= \psi_0 + \psi_1 \text{ Droga}_i + \psi_2 \text{ Gênero}_i + \psi_3 \text{ PrevIO}_i + \psi_4 \text{AZT}_i, \end{aligned}$$

- MEP para os tempos de falha ($m = 4, 10, 16, 22$)

$$\begin{aligned} T_i &\sim PE(\boldsymbol{\lambda}_i, \boldsymbol{\rho}), \\ \boldsymbol{\lambda}_i &= \lambda_{0k} \exp\{\psi_1 \text{ Droga}_i + \psi_2 \text{ Gênero}_i + \psi_3 \text{ PrevIO}_i + \psi_4 \text{ AZT}_i\}, \\ & \quad k = 1, 2, \ldots, m; \end{aligned}$$

- aproximando pelo PB ($m = 4, 10, 16, 22$)

$$\begin{aligned} h(u) &\approx \sum_{k=1}^{m} \gamma_k \mathsf{g}_{m,k}(u) \\ & \quad \exp\{\psi_1 \text{ Droga}_i + \psi_2 \text{ Gênero}_i + \psi_3 \text{ PrevIO}_i + \psi_4 \text{ AZT}_i\} \end{aligned}$$

Model	DIC	-2 LPML	-2 WAIC
$\mathcal{M}_{W}^{N}$	9353.6214	8011.1156	7417.3933
$\mathcal{M}_{PE_4}^{N}$	9442.5262	8047.5062	7419.0756
$\mathcal{M}_{PE_{10}}^{N}$	9260.8474	8018.6984	7419.1594
$\mathcal{M}_{PE_{16}}^{N}$	9281.9028	8034.1428	7419.3481
$\mathcal{M}_{PE_{22}}^{N}$	9344.0505	8036.5340	7420.7621
$\mathcal{M}_{BP_4}^{N}$	9442.5262	8047.5062	7419.0756
$\mathcal{M}_{BP_{10}}^{N}$	9260.8474	8018.6984	7419.1594
$\mathcal{M}_{BP_{16}}^{N}$	9281.9028	8034.1428	7419.3481
$\mathcal{M}_{BP_{22}}^{N}$	9344.0505	8036.5340	7420.7621
$\mathcal{M}_{W}^{BP_2}$	9249.8027	7944.4631	7461.2203
$\mathcal{M}_{W}^{BP_4}$	11581.6323	8031.8499	7368.8280
$\mathcal{M}_{PE_4}^{BP_2}$	**9078.8104**	7944.9195	7459.1482
$\mathcal{M}_{PE_{10}}^{BP_2}$	9209.1102	7936.7814	7462.4873
$\mathcal{M}_{PE_{16}}^{BP_2}$	9149.1778	7936.0419	7463.4823
$\mathcal{M}_{PE_{22}}^{BP_2}$	9199.8430	7928.5876	7461.5026

Model	DIC	-2 LPML	-2 WAIC
$\mathcal{M}_{PE_4}^{BP_4}$	11427.3900	8023.2179	**7363.3080**
$\mathcal{M}_{PE_{10}}^{BP_4}$	11541.4714	8036.4840	7370.7238
$\mathcal{M}_{PE_{16}}^{BP_4}$	11568.8448	8023.8398	7370.5032
$\mathcal{M}_{PE_{22}}^{BP_4}$	11543.0889	8048.5471	7367.4144
$\mathcal{M}_{BP_4}^{BP_2}$	9216.1899	7959.6367	7459.5862
$\mathcal{M}_{BP_{10}}^{BP_2}$	9088.2340	**7909.5476**	7459.0944
$\mathcal{M}_{BP_{16}}^{BP_2}$	9239.3201	7967.8001	7462.3679
$\mathcal{M}_{BP_{22}}^{BP_2}$	9212.8176	7923.3480	7461.6959
$\mathcal{M}_{BP_4}^{BP_4}$	11516.5412	7997.5721	7364.3042
$\mathcal{M}_{BP_{10}}^{BP_4}$	11452.2513	8005.4197	7367.1217
$\mathcal{M}_{BP_{16}}^{BP_4}$	11784.9150	7979.9541	7365.9464
$\mathcal{M}_{BP_{22}}^{BP_4}$	11454.3071	8037.7372	7367.8663

Figura: Comparação dos sub-modelos para a variável *longitudinal*.

Model	DIC	-2 LPML	-2 WAIC
$\mathcal{M}_{W}^{N}$	1538.1301	1537.7357	1537.7338
$\mathcal{M}_{PE_4}^{N}$	1538.9728	1538.0534	1538.0490
$\mathcal{M}_{PE_{10}}^{N}$	1539.0259	1538.0694	1538.0656
$\mathcal{M}_{PE_{16}}^{N}$	1538.6578	1538.0770	1538.0751
$\mathcal{M}_{PE_{22}}^{N}$	1538.7057	1538.0015	1537.9991
$\mathcal{M}_{BP_4}^{N}$	**1504.5523**	1498.5435	1498.5419
$\mathcal{M}_{BP_{10}}^{N}$	1531.9652	1511.5175	1511.5101
$\mathcal{M}_{BP_{16}}^{N}$	1582.6753	1531.9095	1531.8908
$\mathcal{M}_{BP_{22}}^{N}$	1630.2557	1556.3009	1556.2806
$\mathcal{M}_{W}^{BP_2}$	1537.9516	1537.7066	1537.7062
$\mathcal{M}_{W}^{BP_4}$	1538.3951	1537.8691	1537.8653
$\mathcal{M}_{PE_4}^{BP_2}$	1538.9781	1538.0420	1538.0385
$\mathcal{M}_{PE_{10}}^{BP_2}$	1538.9078	1538.0403	1538.0374
$\mathcal{M}_{PE_{16}}^{BP_2}$	1538.5358	1537.8367	1537.8370
$\mathcal{M}_{PE_{22}}^{BP_2}$	1539.0230	1538.2468	1538.2453

Model	DIC	-2 LPML	-2 WAIC
$\mathcal{M}_{PE_4}^{BP_4}$	1539.1776	1538.1922	1538.1911
$\mathcal{M}_{PE_{10}}^{BP_4}$	1539.4302	1538.2609	1538.2562
$\mathcal{M}_{PE_{16}}^{BP_4}$	1538.6243	1537.8401	1537.8360
$\mathcal{M}_{PE_{22}}^{BP_4}$	1539.7737	1538.4683	1538.4626
$\mathcal{M}_{BP_4}^{BP_2}$	1504.6387	**1498.4037**	**1498.4014**
$\mathcal{M}_{BP_{10}}^{BP_2}$	1533.4782	1511.4879	1511.4836
$\mathcal{M}_{BP_{16}}^{BP_2}$	1580.5349	1531.2795	1531.2667
$\mathcal{M}_{BP_{22}}^{BP_2}$	1625.8323	1555.4329	1555.4143
$\mathcal{M}_{BP_4}^{BP_4}$	1505.0525	1498.5384	1498.5340
$\mathcal{M}_{BP_{10}}^{BP_4}$	1534.4938	1511.2390	1511.2277
$\mathcal{M}_{BP_{16}}^{BP_4}$	1576.4428	1531.2379	1531.2168
$\mathcal{M}_{BP_{22}}^{BP_4}$	1617.8107	1554.2922	1554.2445

Figura: Comparação dos sub-modelos *de sobrevivência*.

Outline

6 Próximos Passos

Próximos passos

Os próximos passos a serem abordados são:

- implementação do modelo de dois estágios;
- implementação da modelagem conjunta para dados longitudinais e de sobrevivência;
- estudo de simulação.

Além disso, há outros pontos interessantes que podem ser explorados como, por exemplo:

- utilizar toda a informação do histórico da variável longitudinal ($X_i^H(.)$);
- tratar os dados faltantes.

Outline

7 Referências

Referências I

- Bernstein, S. N. (1913) Démonstration du théorème de Weiertrass fondée sur le calcul des probabilités. *Kharkov Mathematical Society*, **13**.
- Fitzmaurice, G. M., Laird, N. M. e Ware, J. H. (2012) *Applied Longitudinal Analysis*. Wiley Series in Probability and Statistics - Applied Probability and Statistics Section Series. John Wiley & Sons, 2nd edn.
- Guo, X. & Carlin, B. P. (2004) Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician*, **58**, 16–24.
- Ibrahim, J. G., Chu, H. e Chen, L. M. (2010) Basic concepts and methods for joint models of longitudinal and survival data. *Journal of Clinical Oncology*, **28**, 2796–2801.
- Lorentz, G. G. (1986) Bernstein Polynomials, vol. 323 of AMS Chelsea Publishing. American Mathematical Society.

Referências I I

Osman, M. & Ghosh, S. K. (2012) Nonparametric regression models for right-censored data using Bernstein polynomials. *Computational Statistics & Data Analysis*, 56, 559– 573.

Tsiatis, A. A. & Davidian, M. (2004) Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica*, **14**, 809–834.

Tsiatis, A. A., Degruttola, V. e Wulfsohn, M. S. (1995) Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association*, **90**, 27–37.

Wang, L. & Ghosh, S. K. (2013) Nonparametric models for longitudinal data using Bernstein polynomial sieve. *Relatório técnico*, Departmento de Estatística, North Carolina State University.

Wu, L., Liu, W., Yi, G. Y. e Huang, Y. (2012) Analysis of longitudinal and survival data: Joint modeling, inference methods, and issues. *Journal of Probability and Statistics*, **2012**, 1–17.

Obrigada! :)

Relações das funções de sobrevivência

$$S(u) = \exp(-H(u)) \qquad \text{e} \qquad h(u) = \frac{f(u)}{S(u)}.$$

Funções de sobrevivência - MEP

Primeiramente, considere a quantidade t_j, $j = 1, 2, \ldots, b$:

$$t_j = \begin{cases} s_{j-1}, \text{ se } t < s_{j-1} \\ t, \text{ se } t \in (s_{j-1}, s_j] \\ s_j, \text{ se } t > s_j \end{cases}$$

A partir dela, define-se a função de risco acumulado:

$$H(t|\boldsymbol{\lambda}) = \sum_{j=1}^{b} \lambda_j (t_j - s_{j-1})$$

A função densidade de probabilidade:

$$f(t|\boldsymbol{\lambda}) = \lambda_j \exp\left\{ -\sum_{j=1}^{b} \lambda_j (t_j - s_{j-1}) \right\}, \, t \in I_j \ , \, \lambda_j > 0, \, j = 1, 2, \ldots, b$$

A função de sobrevivência:

$$S(t|\boldsymbol{\lambda}) = \exp\left\{ -\sum_{j=1}^{b} \lambda_j (t_j - s_{j-1}) \right\}$$

Função risco - MEP

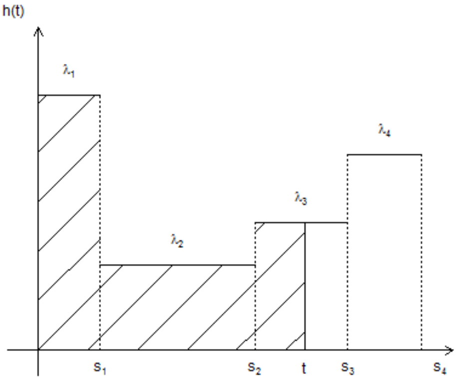


Figura: Modelo Exponencial por Partes.

Tabela: Comparison measures varying the form of imputation - complete, partial and no-imputation.

Complete inputation

	DIC	-2 LPML	-2 WAIC
Longitudinal	15343.7365	11605.1941	10408.7312
Survival	1622.7828	1622.4815	1622.4781

Partial imputation

	DIC	-2 LPML	-2 WAIC
Longitudinal	9516.3405	8112.5174	7485.8851
Survival	1622.5017	1622.3043	1622.3003

No imputation

	DIC	-2 LPML	-2 WAIC
Longitudinal	7116.9886	6503.7523	6168.7855
Survival	1622.5585	1622.3967	1622.3934

Tabela: Posterior estimates for the coefficients from both longitudinal and survival sub-models.

Longitudinal sub-model

	Mean	Median	Mode	Std. Dev.	HPD 95%
Intercept	8.0420	8.0420	8.0500	0.3686	[7.2914 , 8.7550]
Time	-0.1524	-0.1524	-0.1548	0.0238	[-0.2013 , -0.1059]
Time $\times$ Drug (ddI)	0.0455	0.0459	0.0488	0.0304	[-0.0161 , 0.1051]
Gender (male)	-0.2291	-0.2334	-0.2704	0.3212	[-0.8462 , 0.4102]
PrevOI (yes)	-2.2530	-2.2530	-2.2560	0.2211	[-2.7038 , -1.8479]
AZT (failure)	-0.1304	-0.1323	-0.1374	0.2216	[-0.5494 , 0.3214]

Survival sub-model

	Mean	Median	Mode	Std. Dev.	HPD 95%
Drug (ddI)	0.1367	0.1369	0.1091	0.1484	[-0.1293 , 0.4544]
Gender (male)	-0.1569	-0.1597	-0.1549	0.1268	[-0.4047 , 0.0919]
PrevOI (yes)	0.5692	0.5666	0.5653	0.1088	[0.3755 , 0.7962]
AZT (failure)	0.1028	0.1019	0.1095	0.0846	[-0.0699 , 0.2656]

Cálculo WAIC

$$\widehat{elpd}_{waic} = \widehat{lpd} - \widehat{p}_{waic}, \tag{15}$$

onde $\widehat{p}_{waic}$ é o número efetivo de parâmetros.

$$\widehat{lpd} = \sum_{i=1}^{n} \log\left(\frac{1}{M}\sum_{l=1}^{M} p(t_i|\boldsymbol{\beta}^l, \boldsymbol{\lambda}^l_{(\rho)}, \boldsymbol{\psi}^l)\right) \tag{16}$$

$$\widehat{p}_{waic} = \sum_{i=1}^{n} V_{l=1}^{M}(\log(p(t_i|\boldsymbol{\beta}^l, \boldsymbol{\lambda}^l_{(\rho)}, \boldsymbol{\psi}^l))) \tag{17}$$

Cálculo LPML

A quantidade CPO associada à i-ésima observação é definida como a preditiva de t_i condicional em $D_{obs}^{(-i)}$:

$$CPO_i = f(t_i|D_{obs}^{(-i)}). \tag{18}$$

É possível estimar a equação (18) através da amostra *a posteriori*:

$$\widehat{CPO}_i = L\left(\sum_{l=1}^{M}\left[f(t_i|\boldsymbol{\beta}^l, \boldsymbol{\lambda}_{(\rho)}^l, \boldsymbol{\psi}^l, D_{obs})\right]\right)^{-1}, \tag{19}$$

onde M é o tamanho da amostra *a posteriori*.

$$LPML = \sum_{i=1}^{n}\log\left(CPO_i\right) \tag{20}$$

De acordo com essa medida, quanto maior melhor.