Introdução aos Modelos de Fragilidade

I Workshop de Verão em Matemática da UFCG

Prof. Dr. Tiago Almeida de Oliveira

Departamento de Estatística
Universidade Estadual da Paraíba

Sumário

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;
- Observa-se o tempo até a ocorrência de um evento de interesse ;

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;
- Observa-se o tempo até a ocorrência de um evento de interesse ;
- Útil para análise de dados com observações incompletas ou perdidas ao longo do tempo.

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;
- Observa-se o tempo até a ocorrência de um evento de interesse ;
- Útil para análise de dados com observações incompletas ou perdidas ao longo do tempo.
- Cálculo de probabilidades pertinentes à variável resposta ;

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;
- Observa-se o tempo até a ocorrência de um evento de interesse ;
- Útil para análise de dados com observações incompletas ou perdidas ao longo do tempo.
- Cálculo de probabilidades pertinentes à variável resposta ;
- Uso de técnicas não paramétricas, semiparamétricas e paramétricas para obtenção de estimativas ;

Introdução

Análise de Sobrevivência

- Área da estatística com alta aplicabilidade nas ciências da saúde, além das mais diversas áreas ;
- Observa-se o tempo até a ocorrência de um evento de interesse ;
- Útil para análise de dados com observações incompletas ou perdidas ao longo do tempo.
- Cálculo de probabilidades pertinentes à variável resposta ;
- Uso de técnicas não paramétricas, semiparamétricas e paramétricas para obtenção de estimativas ;
- Ajuste de modelos de regressão.

Introdução

Objetivo do estudo

- Introduzir os modelos de fragilidade e aplicá-los afim de realizar comparação com a modelagem convencional.

Sobrevivência

Em que tipo de desenho de estudo se aplica a Análise de Sobrevivência ?

- Coorte - observacional ou de intervenção (ensaio clínico) - pressupõe o acompanhamento dos indivíduos ao longo do tempo.

Softwares : R, python, SAS, STATA ...

Sobrevivência

As perguntas passíveis de resposta neste tipo de abordagem são :

- Qual o efeito de um determinado anticancerígeno sobre o tempo de sobrevivência ?
- Quais os fatores associados ao tempo de duração da amamentação ?
- Quais os fatores preditivos para reinternação hospitalar, considerando o tempo entre internações ?
- Qual o efeito da unidade assistencial na sobrevivência após um infarto agudo do miocárdio ?

Sobrevivência

exemplos considerando a possível perda de seguimento (censura) :

- Tempo até a menopausa ;
- Tempo até remissão de câncer ;
- Tempo até surgimento de cárie ;
- Tempo de falha de equipamentos industriais (engenharia) ;
- Tempo de duração do período de desemprego ou greve (economia).

Revisão de Literatura

Conceitos introdutórios

Tempo

T uma variável aleatória, não-negativa, absolutamente contínua, cujo tempo de sobrevida $T \geq 0$ é expresso por meio de várias funções matematicamente equivalentes. A saber : a função densidade de probabilidade, $f(t)$, a função de sobrevivência, $S(t)$, função de risco, $\lambda(t)$ e a função de risco acumulado, $\Lambda(t)$.

Revisão de Literatura

Conceitos introdutórios

Estas funções são usadas para descrever diferentes aspectos apresentado pelo conjunto de dados e utilizadas para caracterizar o comportamento de dados de tempo de sobrevivência, em que t representa o tempo de falha especificada em análise de sobrevivência, cuja distribuição pode ser caracterizada por qualquer umas das seguintes funções :

Revisão de Literatura

Conceitos introdutórios

Estas funções são usadas para descrever diferentes aspectos apresentado pelo conjunto de dados e utilizadas para caracterizar o comportamento de dados de tempo de sobrevivência, em que t representa o tempo de falha especificada em análise de sobrevivência, cuja distribuição pode ser caracterizada por qualquer umas das seguintes funções :

- A função densidade de probabilidade, $f(t)$

Revisão de Literatura

Conceitos introdutórios

Estas funções são usadas para descrever diferentes aspectos apresentado pelo conjunto de dados e utilizadas para caracterizar o comportamento de dados de tempo de sobrevivência, em que t representa o tempo de falha especificada em análise de sobrevivência, cuja distribuição pode ser caracterizada por qualquer umas das seguintes funções :

- A função densidade de probabilidade, $f(t)$
- A função de sobrevivência, $S(t)$

Revisão de Literatura

Conceitos introdutórios

Estas funções são usadas para descrever diferentes aspectos apresentado pelo conjunto de dados e utilizadas para caracterizar o comportamento de dados de tempo de sobrevivência, em que t representa o tempo de falha especificada em análise de sobrevivência, cuja distribuição pode ser caracterizada por qualquer umas das seguintes funções :

- A função densidade de probabilidade, $f(t)$
- A função de sobrevivência, $S(t)$
- A função de risco, $\lambda(t)$

Revisão de Literatura

Conceitos introdutórios

Estas funções são usadas para descrever diferentes aspectos apresentado pelo conjunto de dados e utilizadas para caracterizar o comportamento de dados de tempo de sobrevivência, em que t representa o tempo de falha especificada em análise de sobrevivência, cuja distribuição pode ser caracterizada por qualquer umas das seguintes funções :

- A função densidade de probabilidade, $f(t)$
- A função de sobrevivência, $S(t)$
- A função de risco, $\lambda(t)$
- A função de risco acumulado, $\Lambda(t)$.

Revisão de Literatura

Funções importantes

A função densidade de probabilidade

É caracterizada pelo evento de interesse ao observar um indivíduo no intervalo de tempo $[t, t + \Delta t]$ por unidade de tempo. E é definida como limite da probabilidade, (DINIZ ; LOUZADA, 2012). Expressa por,

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \tag{1}$$

em que, $f(t) \geq 0$ para todo t, e tem a área abaixo da curva igual a 1.

Revisão de Literatura

Funções importantes

A função de sobrevivência

Pode ser definida como a probabilidade de sobreviver acima de um tempo t. Formalmente ou matematicamente é escrita como

$$S(t) = P(T > t) = 1 - P(T \leq t) = 1 - F(t) \tag{2}$$

Esta função recebe valor 1 no tempo 0, e decresce ou permanece constante ao longo do tempo.

Revisão de Literatura

Funções importantes

A função de riso ou taxa de falha

Segundo Colosimo e Giolo (2006), a função de risco $\lambda(t)$ é a taxa de falha instantânea no tempo t condicionada à sobrevivência até o tempo t. Deste modo, fixando um intervalo de tempo t e $t + \Delta t$. A função taxa de falha ou função de risco é então definida da seguinte maneira

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t \mid T \geq t)}{\Delta t} \tag{3}$$

Revisão de Literatura

Funções importantes

A função de risco acumulado

Assim como o próprio nome sugere, esta função nos fornece o risco acumulado ou taxa de falha acumulada dos indivíduos, e é definida como

$$\Lambda(t) = \int_0^t \lambda(u)du \tag{4}$$

Revisão de Literatura

Conceitos introdutórios

Censura

Censura pode ser vista como uma observação incompleta, ou perdida ao longo de um tempo de estudo. A ocorrência de dados censurados se deve ao fato que em um estudo nem todos os indivíduos, por diversos motivos, o evento de interesse ocorre.

Características do Dados de sobrevivência

1. Distribuição Assimétrica Positiva.
2. Presença de Censura : Observação parcial da resposta.
   - Esquerda
   - Direita (tipo I, Tipo II, Aleatória)
   - Intervalar
3. Presença de Covariáveis

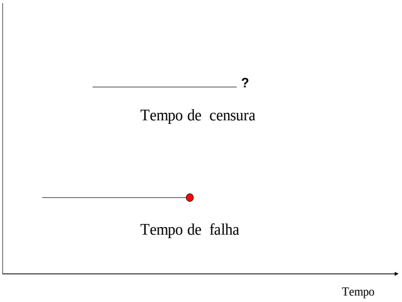


FIGURE – Exemplo

Causas de Informação Incompleta (censura) :

- óbito por outras causas - morte do paciente por causas externas ;
- término do estudo ;
- perda de contato - mudança de residência ;
- recusa em continuar participando do estudo ;
- mudança de procedimento - esquema de tratamento ;
- abandono devido a efeitos adversos de tratamento ;
- uso de dados prevalentes - óbitos antes do início do estudo.

Revisão de Literatura

Conceitos introdutórios

Segundo Strapasson (2007), para análise de sobrevivência é necessário que as observações sejam representadas por um vetor $(t_i,\ \delta_i,\ x_i)$ em que, t_i é o tempo observado de falha ou censura e δ_i uma variável indicadora de falha, em que $\delta_i = 1$, o tempo observado corresponde a uma falha ou $\delta_i = 0$, corresponde a uma censura. Para cada indivíduo observado tem-se uma covariável x_i, em que $i, i = 1, \ldots, n$ são observações representadas pelo um par $(t_i,\ \delta_i)$.

Revisão de Literatura

Conceitos introdutórios

Segundo Strapasson (2007), para análise de sobrevivência é necessário que as observações sejam representadas por um vetor $(t_i, \, \delta_i, \, x_i)$ em que, t_i é o tempo observado de falha ou censura e δ_i uma variável indicadora de falha, em que $\delta_i = 1$, o tempo observado corresponde a uma falha ou $\delta_i = 0$, corresponde a uma censura. Para cada indivíduo observado tem-se uma covariável x_i, em que $i, i = 1, \ldots, n$ são observações representadas pelo um par $(t_i, \, \delta_i)$.

Censura

$$\delta_i = \begin{cases} 1, & \text{quando } T \leq C \\ 0, & \text{quando } T > C \end{cases}$$

Exemplo : Sobrevida em Aids

T = tempo entre diagnóstico e o óbito.

Paciente (i)	Tempo (T_i)	δ_i	Terapia	Doença Apresentada
1	25	0	sem	PPC
2	2	1	mono	Kaposi
3	34	0	potente	Toxoplasmose
4	12	1	comb	Duas ou mais
5	11	1	sem	PPC
6	10	1	mono	Toxoplasmose
7	5	1	mono	Duas ou mais
8	13	1	potente	Kaposi
9	9	1	comb	PPC
10	19	1	potente	Tuberculose

Revisão de Literatura

O estimador de Kaplan-Meier

Como as funções citadas anteriormente não permitem a presença de observações censuradas. Kaplan e Meier (1958) propuseram um estimador em que as estimativas podem ser obtidas a partir de métodos não-paramétricos. Sendo assim, o estimador de Kaplan-Meier é definido por

$$\widehat{S(t)} = \prod_{j, t_j < t} \left(1 - \frac{d_j}{n_j}\right) \tag{5}$$

Revisão de Literatura

Comparação entre as curvas de sobrevivência

O teste Log-rank

Hipóteses de interesse :

$$\begin{cases} H_0 : & S_1(t) = S_2(t) \\ H_1 : & S_1(t) \neq S_2(t) \end{cases}$$

Estatística do teste :

$$T = \frac{\left[\sum_{j=1}^{k} (d_{2j} - w_{2j})\right]^2}{\sum_{j=1}^{k} \left(V_j\right)^2} \sim \chi^2_{(1)} \tag{6}$$

em que d_{2j} caracteriza a falha dos indivíduos do grupo 2 no tempo j. w_{2j} e V_j é a média e a variância de d_{2j} respectivamente.

Revisão de Literatura

Comparação entre as curvas de sobrevivência

O teste Peto

Hipóteses de interesse :

$$\begin{cases} H_0: & S_1(t) = S_2(t) \\ H_1: & S_1(t) \neq S_2(t) \end{cases}$$

Estatística do teste :

$$S = \frac{\left[\sum_{j=1}^{k} u_j (d_{2j} - w_{2j})\right]^2}{\sum_{j=1}^{k} u_j \left(V_j\right)^2} \sim \chi^2_{(1)}, \qquad (7)$$

em que,

$$u_j = \widetilde{S}(t_j - 1)\frac{n_j}{n_j + 1}$$

onde, $\widetilde{S}(.)$ é um estimador da função de sobrevivência conhecido como Peto-Prentice.

Revisão de Literatura

Distribuições de Probabilidade

Algumas distribuições de probabilidade para o tempo de falha.

- Exponencial ;
- Weibull ;
- Log-normal ;
- Log-logística ;
- Normal inversa ;
- Gompertz ;
- Gama ;
- Log-gama ;
- Hayleigh ;
- Gama Generalizada.

Revisão de Literatura

Modelos probabilísticos

O modelo Exponencial

De acordo com Wienke (2010), a variável aleatória tempo de falha T seguem uma distribuição exponencial ($T \sim Exp(\alpha)$). Então segue que :

Função densidade de probabilidade :

$$f(t) = \alpha \exp^{-\alpha t} \tag{8}$$

Função de sobrevivência :

$$S(t) = \exp^{-\alpha t} \tag{9}$$

Função de risco :

$$\lambda(t) = \alpha \tag{10}$$

Função de risco acumulado :

$$\Lambda(t) = \alpha t \tag{11}$$

Revisão de Literatura

O modelo Log-normal

No modelo log-normal ($T \sim logN(\mu, \sigma^2)$), o logarítimo natural $ln(T)$ do tempo de vida T é assumido ser normalmente distribuído ($ln(T) \sim N(\mu, \sigma^2)$). A função densidade de probabilidade de uma variável T que segue uma distribuição log-normal é dada por

$$f(t) = \frac{1}{\sqrt{2\pi}t\sigma} exp\left\{-\frac{1}{2}\left(\frac{log(t)-\mu}{\sigma}\right)^2\right\} \tag{12}$$

Revisão de Literatura

Modelos probabilísticos

O modelo Log-normal

Segundo Wienke (2010) no modelo log-normal as funções de sobrevivência e de risco não segue um forma fechada e inclui um integral normal incompleta dada da seguinte maneira

$$\Phi(t) = \int_{-\infty}^{t} \phi(\sigma) d\sigma, \tag{13}$$

em que $\phi(t)$ denota a função densidade de probabilidade de uma distribuição normal padrão.

Distribuição Log-normal

- Densidade $\rightarrow f(t) = \frac{1}{\sqrt{2\pi}t\sigma} \exp\left\{-\frac{1}{2}\left[\frac{\log(t)-\mu}{\sigma}\right]^2\right\}$

- Sobrevivência $\rightarrow S(t) = 1 - \Phi\left[\frac{\log(t)-\mu}{\sigma}\right] = \Phi\left[\frac{-\log(t)+\mu}{\sigma}\right]$

- Função risco $\rightarrow h(t) = \frac{\frac{1}{\sqrt{2\pi}t\sigma} \exp\left\{-\frac{1}{2}\left[\frac{\log(t)-\mu}{\sigma}\right]^2\right\}}{\Phi\left[\frac{-\log(t)+\mu}{\sigma}\right]}$

- Em que $\mu \in \mathbb{R}$ é a média do logaritmo do tempo de falha, $\sigma > 0$ é o desvio padrão e $\Phi(.)$ é a função de distribuição acumulada de uma normal padrão.

Revisão de Literatura

Modelos probabilísticos

O modelo Log-logístico

Função densidade de probabilidade :

$$f(t) = \frac{\alpha\gamma t^{\gamma-1}}{(1+\alpha t^{\gamma})^2} \tag{14}$$

Função de sobrevivência :

$$S(t) = \frac{1}{1+\alpha t^{\gamma}} \tag{15}$$

Função de risco :

$$\lambda(t) = \frac{\alpha\gamma t^{\gamma-1}}{1+\alpha t^{\gamma}} \tag{16}$$

Função de risco acumulado :

$$\Lambda(t) = ln(1+\alpha t^{\gamma}) \tag{17}$$

Distribuição Gama

- Densidade $g(t) = \frac{1}{\alpha\Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} \exp\left[-\left(\frac{t}{\alpha}\right)\right]$

- Sobrevivência $\rightarrow S(t) = 1 - G(t) = 1 - \gamma_1\left[k, \frac{t}{\alpha}\right]$

- Função risco $\rightarrow h(t) = \frac{\frac{1}{\alpha\Gamma(k)}\left(\frac{t}{\alpha}\right)^{k-1}\exp\left[-\left(\frac{t}{\alpha}\right)\right]}{1-\gamma_1\left[k, \frac{t}{\alpha}\right]}$

- Em que $k > 0$ é o parâmetro de forma, $\alpha > 0$ o parâmetro de escala e $\Gamma(k)$ é a função gama, definida por :

$$\Gamma(k) = \int_0^\infty t^{k-1}\exp(-t)dt.$$

- $\gamma(k, x) = \int_0^x w^{k-1}e^{-w}dw$ é a função gama incompleta e $\gamma_1(k, x)$ é a razão da função gama incompleta, definida por $\gamma_1(k, x) = \gamma(k, x)/\Gamma(k)$

Estimação por Máxima Verossimilhança

Função de verossimilhança na sobrevivência

- Sem censura :

$$L \propto \prod_{i \epsilon F} f(t_i)$$

- Com censura à direita :

$$L \propto \prod_{i \epsilon F} f(t_i) \prod_{i \epsilon D} S(t_i)$$

- Com censura à esquerda :

$$L \propto \prod_{i \epsilon F} f(t_i) \prod_{i \epsilon D} S(t_i) \prod_{i \epsilon E} [1 - S(t_i)]$$

- Com censura intervalar :

$$L \propto \prod_{i \epsilon F} f(t_i) \prod_{i \epsilon D} S(t_i) \prod_{i \epsilon E} [1 - S(t_i)] \prod_{i \epsilon I} [S(t_i^-) - S(t_i^+)]$$

Estimação por Máxima Verossimilhança

- Função de verossimilhança na sobrevivência forma geral para todos os mecanismos de censura :

$$L(\theta) = \prod_{i=1}^{r} f(t_i;\theta) \prod_{i=r+1}^{n} S(t_i;\theta)$$

ou equivalentemente por :

$$L(\theta) = \prod_{i=1}^{n} \left[f(t_i;\theta)\right]^{\delta_i} \left[S(t_i;\theta)\right]^{1-\delta_i}$$

Estimação por Máxima Verossimilhança

- A função de verossimilhança avalia o quanto os dados apoiam, concordam ou suportam cada valor possível do parâmetro a ser estimado.

- Pressupostos do método de Máxima Verossimilhança :

  $\rightarrow$ Observações independentes
  $\rightarrow$ Tempos de sobrevivência independentes
  $\rightarrow$ Censuras independentes

- Os estimadores de máxima verossimilhança são os valores de θ que maximizam $L(\theta)$

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta} = 0$$

Estatísticas AIC, BIC e CAIC

Considerando uma amostra aleatória $X_1, \ldots, X_n$ de tamanho n e o vetor de parâmetros $\boldsymbol{\theta}$, as estatísticas AIC (critério de informação de Akaike), BIC (critério de informação bayesiano) e CAIC (AIC corrigido) podem ser calculadas por :

$$
\begin{aligned}
AIC &= -2l(\boldsymbol{\theta}) + 2d, \\
BIC &= -2l(\boldsymbol{\theta}) + d\log(n), \\
CAIC &= \text{AIC} + \frac{2(d+2)(d+3)}{n-d-3},
\end{aligned}
$$

em que $l(\boldsymbol{\theta})$ é o logaritmo da função de verossimilhança e d representa o número de parâmetros estimados do modelo.

Modelo de Regressão Paramétrica

- Nos modelos paramétricos, a **inclusão de covariáveis** segue a forma utilizada em modelos lineares generalizados, podendo ser tanto contínuas - pressão sanguínea, idade, dosagens bioquímicas - como categóricas - gênero, tratamento, comportamentos.
- O objetivo de um modelo de regressão é o de **estimar o efeito** de covariáveis (ou variáveis independentes ou preditores), $\boldsymbol{X_1, X_2, \ldots, X_p}$, sobre uma variável resposta (ou variável dependente), $\boldsymbol{Y}$.
- Supondo uma distribuição da família exponencial para a variável resposta teremos um modelo linear generalizado.
- Ainda que a distribuição exponencial e a Weibull sejam parte desta família, os modelos de regressão paramétricos para tempo de sobrevivência não são parte dos GLM por causa de **dados censurados**.

Modelo de Regressão Paramétrico

- T –> tempo até o evento ou censura, variável resposta
- $\boldsymbol{x}$ –> vetor de covariáveis
- Função de risco : $\lambda(t|\boldsymbol{x}) = \lambda_0(t)g(\boldsymbol{x}^{\boldsymbol{t}}\boldsymbol{\beta})$:

  $\boldsymbol{\beta}$ –> coeficientes estimados
  $g(.)$ –> função de ligação, positiva e contínua (exponencial, Weibull)
- Razão de riscos λ/λ_0 é função das covariáveis e não depende do tempo –> riscos **proporcionais**

Modelo de Regressão Paramétrica

- Assumimos que o parâmetro da distribuição depende de covariáveis segundo uma função
- Exemplo : $\alpha(\boldsymbol{x}) = \exp(\boldsymbol{x}\boldsymbol{\beta})$
- Modelo Exponencial :

$$S(t|\boldsymbol{x}) = \exp(-\alpha(\boldsymbol{x})t) = \exp(-\exp(\boldsymbol{x}\boldsymbol{\beta})t)$$

$$\lambda(t|\boldsymbol{x}) = \alpha(\boldsymbol{x}) = \exp(\boldsymbol{x}\boldsymbol{\beta})$$

- Modelo Weibull :

$$S(t) = \exp(-(\alpha(\boldsymbol{x})t)^{\eta}) = \exp(-(\exp(\boldsymbol{x}\boldsymbol{\beta})t)^{\eta})$$

$$\lambda(t) = \eta\alpha(\boldsymbol{x})^{\eta}t^{\eta-1} = \eta(\exp(\boldsymbol{x}\boldsymbol{\beta}))^{\eta}t^{\eta-1}$$

O modelo de riscos proporcionais de Cox

Colosimo e Giolo (2006) definem o modelo de Cox da seguinte maneira. Considere p covariáveis, de modo que $\boldsymbol{x}$ seja um vetor com os componentes $\boldsymbol{x} = (x_1, \ldots, x_p)^t$. A expressão geral do modelo de regressão de Cox considera :

$$\lambda(t) = \lambda_0(t) g(\boldsymbol{x}^t \boldsymbol{\beta}), \tag{18}$$

em que g é uma função não-negativa que deve ser especificada, tal que $g(0) = 1$.

Material e Métodos

O modelo de riscos proporcionais de Cox

O componente não-paramétrico, $\lambda_0(t)$, não é especificado e é uma função não-negativa do tempo. Este componente é usualmente chamado de função de base, pois $\lambda(t) = \lambda_0(t)$ quando $\boldsymbol{x} = \mathbf{0}$. O componente paramétrico é frequentemente usado na seguinte forma multiplicativa :

$$g(\boldsymbol{x}^t\boldsymbol{\beta}) = \exp(\boldsymbol{x}^t\boldsymbol{\beta}) = \exp(x_1\beta_1 + \ldots + x_p\beta_p), \tag{19}$$

Método da Máxima Verossimilhança Parcial

Considere que, em uma amostra de n indivíduos, existam $k \leq n$ falhas distintas nos tempos $t_1 < t_2 < ... < t_k$. A função de verossimilhança parcia $L(\beta)$:

$$L(\beta) = \prod_{i=1}^{k} \frac{\exp\left(X_i'\beta\right)}{\sum_{j \in R(t_i)} \exp\left(X_j'\beta\right)} = \prod_{i=1}^{k} \left(\frac{\exp\left(X_i'\beta\right)}{\sum_{j \in R(t_i)} \exp\left(X_j'\beta\right)} \right)^{\delta_i},$$

em que $\delta_i = 0$ ou 1.

Estimação de Parâmetros

Estimadores para o vetor de parâmetros β podem ser obtidos maximizando-se o logaritmo da função de verossimilhança parcial. Isto é,

$$U(\beta) = \sum_{i=1}^{n} \delta_i \left[x_i - \frac{\sum_{j \in R(t_i)} x_j \exp(X_j' \widehat{\beta})}{\sum_{j \in R(t_i)} \exp(X_j' \widehat{\beta})} \right]$$

Razão de Riscos

- $\lambda(t) = \lambda_0(t)\exp(\beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p)$
- Considere a covariável X_1 :

$$\text{RR} = \frac{\text{Risco} \quad (\text{grupo} \quad 1)}{\text{Risco} \quad (\text{grupo} \quad 0)}.$$

$$\text{RR} = \exp(\beta)$$

Revisão de Literatura

Modelos de fragilidade

Fragilidade

O que é o modelo de fragilidade ? ?

Para que serve ? ?

Quais as vantagens desses modelos em relação aos modelos convencionais de sobrevivência ? ?

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.

Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

- Em essência o conceito de fragilidade vem do trabalho de Greenwood e Yule (1920) sobre "propensão para acidentes";

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

- Em essência o conceito de fragilidade vem do trabalho de Greenwood e Yule (1920) sobre "propensão para acidentes" ;
- Clayton (1978) desenvolveu um modelo bivariado para doenças cronicas em família ;

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

- Em essência o conceito de fragilidade vem do trabalho de Greenwood e Yule (1920) sobre "propensão para acidentes" ;
- Clayton (1978) desenvolveu um modelo bivariado para doenças cronicas em família ;
- Vaupel et al. (1979) usou o termo fragilidade (Fragilidade Gama) para a heterogeneidade não observada em um contexto univariado ;

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

- Em essência o conceito de fragilidade vem do trabalho de Greenwood e Yule (1920) sobre "propensão para acidentes" ;
- Clayton (1978) desenvolveu um modelo bivariado para doenças cronicas em família ;
- Vaupel et al. (1979) usou o termo fragilidade (Fragilidade Gama) para a heterogeneidade não observada em um contexto univariado ;
- Hougaard (1986) fragilidade estável positiva ;

Revisão de Literatura

O modelo de fragilidade

Suposição : A função de risco para cada indivíduo é proporcional para a função de risco base, $\lambda(t)$. Essa suposição implica que a função de risco é completamente determinada por um vetor de covariáveis.
Problema : Pode haver covariáveis não observadas que faça com que essa suposição seja violada.

Histórico dos Modelos de Fragilidade

- Em essência o conceito de fragilidade vem do trabalho de Greenwood e Yule (1920) sobre “propensão para acidentes" ;
- Clayton (1978) desenvolveu um modelo bivariado para doenças cronicas em família ;
- Vaupel et al. (1979) usou o termo fragilidade (Fragilidade Gama) para a heterogeneidade não observada em um contexto univariado ;
- Hougaard (1986) fragilidade estável positiva ;
- Oakes (1989) Expandiu para outras famílias e distribuições.

Fragilidade Univariada

- Modelos de fragilidade podem ajudar a explicar heterogeneidade não contabilizada.

$$\lambda_j(t) = \lambda_0(t) \exp(\boldsymbol{x}_j^t \boldsymbol{\beta} + W_j \psi), j = 1, ..., n \tag{20}$$

em que W_j é o termo de fragilidade para a distribuição de probabilidade com média 0 e variância 1.
Se W_j pode ser mensurado e incluso no modelo, então ψ poderá ser 0 e pode-se obter o modelo de riscos proporcionais padrão.
A função de risco condicional sob ambas as covariáveis e a fragilidade pode ser reescrita como,

$$\lambda_j(t) = \lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}), j = 1, ..., n \tag{21}$$

em que $z_j = \exp(W_j \psi)$. Desta forma o risco de um indivíduo também depende de uma variável aleatória não observada, Z_j, a qual age multiplicativamente sobre o risco. Se a fragilidade não é levada em conta então $z_j = 1$.

Abordagem de verossimilhança para Modelos de Fragilidade

Dados de sobrevivência consistem de uma combinação de observações de tempos de falha e censura. A função de verossimilhança para dados de sobrevivência é dada por,

$$L = \prod_{j=1}^{n} [(1 - G_j(t))f_j(t)]^{\delta_j} [(1 - F_j(t))g_j(t)]^{1-\delta_j} \tag{22}$$

em que, δ_j é o indicador de ceunsura, $g(.)$ e $G(.)$ são funções de densidade de probabilidade e função acumulada para o tempo de censura, $f(.)$ e $F(.)$ são funções de densidade e acumulada para o tempo de falha.

Revisão de Literatura

Assumindo Censura a direita,

$$L = \prod_{j=1}^{n} [f_j(t)]^{\delta_j} [S_j(t)]^{1-\delta_j}$$

Desde que $\lambda_j(t) = f_j(t)/S_j(t)$, pode-se reescrever $f_j(t)$ na função de verossimilhança,

$$L = \prod_{j=1}^{n} [\lambda_j(t)]^{\delta_j} [S_j(t)]$$

Pode-se derivar as formas da função de verossimilhança condicional e marginal do modelo de fragilidade,

$$\begin{aligned} \lambda_j(t) &= \lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) \\ \frac{f_j(t)}{S_j(t)} &= \lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) \end{aligned}$$

Integrando ambos os lados,

$$\begin{aligned} \int_0^\infty \frac{f_j(t)}{S_j(t)}dt &= \int_0^\infty \lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) dt \\ -ln(S(t)) &= \Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) \end{aligned}$$

Ademais,

$$S_j(t) = \exp[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_i^t \boldsymbol{\beta})]$$

A verossimilhança condicional é dada por,

$$L(\psi, \beta | z_j) = \prod_{j=1}^{n} [\lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})]^{\delta_j} \exp\left[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})\right]$$

em que ψ é um vetor de parâmetros da função de risco.

Integrando ambos os lados,

$$\begin{aligned} \int_0^\infty \frac{f_j(t)}{S_j(t)}dt &= \int_0^\infty \lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) dt \\ -ln(S(t)) &= \Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}) \end{aligned}$$

Ademais,

$$S_j(t) = \exp[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_i^t \boldsymbol{\beta})]$$

A verossimilhança condicional é dada por,

$$L(\psi, \beta | z_j) = \prod_{j=1}^{n} [\lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})]^{\delta_j} \exp\left[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})\right]$$

em que ψ é um vetor de parâmetros da função de risco.
A função Marginal de versossimilhança é,

$$L(\psi, \theta, \beta) = \prod_{j=1}^{n} \left[\int_0^\infty \lambda_0(t) z_j \exp(\boldsymbol{x}_i^t \boldsymbol{\beta})\right]^{\delta_j} \exp\left[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})\right] g(z) dz$$

em que, g(z) é a função de distibuição de probabilidade das fragilidades $z_1, ..., z_n$

Modelo de Fragilidade Gama

Para obter a verossimilhança marginal para o modelo de fragilidade Gama, deve-se proceder da seguinte forma :

Modelo de Fragilidade Gama

Para obter a verossimilhança marginal para o modelo de fragilidade Gama, deve-se proceder da seguinte forma :

Dado z_i ser uma amostra independente e identicamente distribuída (i.i.d) de uma variável aleatória gama com função de densidade.

$$g(z) = \frac{z^{1/\theta-1}e^{-z/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}}, z > 0, \theta > 0,$$

com $E(Z) = 1$ e $Var(Z) = \theta$. Maiores valores de θ implicam que há um maior grau de heterogeneidade entre indivíduos.

Primeiro, a fragilidade gama pode ser integrada na verossimilhança condicional de sobrevivência. Isto deixa explicito a função marginal de verossimilhança,

$$L(\psi, \theta, \boldsymbol{\beta}) = \prod_{j=1}^{n} \int_0^{\infty} \left[\lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})\right]^{\delta_j} \exp\left[-\Lambda_0(t) z_j \exp(\boldsymbol{x}_j^t \boldsymbol{\beta})\right] \frac{z^{1/\theta - 1} e^{-z/\theta}}{\Gamma(1/\theta)\theta^{1/\theta}} dz \quad (23)$$

em que ψ contém os parâmetros da função de risco de base. $\psi = (\alpha)$ para risco base exponencial e $\psi = (\eta, \alpha)$ para risco base Weibull.

Revisão de Literatura

Rearranjando os termos na equação 23, nós obtemos a seguinte expressão,

$$L(\psi,\theta,\boldsymbol{\beta}) = \prod_{j=1}^{n} \lambda_0(t)^{\delta_j} \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}\delta_j) \int_0^\infty \frac{z^{1/\theta+d-1} e^{-z/\theta} exp(-\sum_{j=1}^{n} \Lambda_0(t) z e^{\boldsymbol{x}^t \boldsymbol{\beta}})}{\Gamma(1/\theta)\theta^{1/\theta}} dz \tag{24}$$

$$= \prod_{j=1}^{n} \lambda_0(t)^{\delta_j} \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}\delta_j) \frac{\Gamma(1/\theta+d)\theta^{(1/\theta+d)}}{\Gamma(1/\theta)\theta^{1/\theta}} \int_0^\infty \frac{z^{1/\theta+d-1} exp(-z(1/\theta + \sum_{j=1}^{n} \Lambda_0(t) e^{\boldsymbol{x}^t \boldsymbol{\beta}})}{\Gamma(1/\theta+d)\theta^{(1/\theta+d)}} dz \tag{25}$$

em que $d = \sum_{j=1}^{n} \delta_j$.
Para fazer o problema tratável, integra-se no termo de fragilidade z. O termo sob a integral é a função geradora de momentos (f.g.m) da distribuição gama com função de densidade de probabilidade $\Gamma(1/\theta + d, 1/\theta)$. Utilizando este fato, nós podemos derivar a expressão para a função de verossimilhança marginal como,

Revisão de Literatura

$$L(\psi,\theta,\boldsymbol{\beta}) = \prod_{j=1}^{n} \lambda_0(t)^{\delta_j} \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}\delta_j) \frac{\Gamma(1/\theta + d)\theta^{(1/\theta+d)}}{\Gamma(1/\theta)\theta^{1/\theta}\theta^{(1/\theta+d)}(1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}})^{(1/\theta+d)}} \times$$

$$\times \int_0^\infty \frac{z^{1/\theta+d-1} exp(-z(1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}}) \left[1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}}\right]^{(1/\theta+d)}}{\Gamma(1/\theta + d)} dz$$

$$L(\psi,\theta,\boldsymbol{\beta}) = \prod_{j=1}^{n} \lambda_0(t)^{\delta_j} \exp(\boldsymbol{x}_j^t \boldsymbol{\beta}\delta_j) \frac{\Gamma(1/\theta + d)}{\Gamma(1/\theta)\theta^{1/\theta}(1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}})^{(1/\theta+d)}} \times$$

$$\times \int_0^\infty \frac{z^{1/\theta+d-1} exp(-z(1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}}) \left[1/\theta + \sum_{j=1}^{n} \Lambda_0(t)e^{\boldsymbol{x}^t\boldsymbol{\beta}}\right]^{(1/\theta+d)}}{\Gamma(1/\theta + d)} dz$$

O termo sob a integral é a função de densidade e probabilidade de $\Gamma(1/\theta + d, 1/\theta + \sum_{j=1}^{n} \Lambda_0(t) e^{\boldsymbol{x}_j^t \boldsymbol{\beta}})$, a qual integra para 1. Ademais, a função de verossimilhança marginal é,

$$L(\boldsymbol{\psi}, \theta, \boldsymbol{\beta}) = \frac{\Gamma(1/\theta + d) \prod_{j=1}^{n} \lambda_0(t)^{\delta_j} \exp(\boldsymbol{x}_j^t \boldsymbol{\beta} \delta_j)}{(1/\theta + \sum_{j=1}^{n} \Lambda_0(t) e^{\boldsymbol{x}^t \boldsymbol{\beta}})^{(1/\theta + d)} \Gamma(1/\theta) \theta^{1/\theta}}.$$

Aplicando o logaritmo na expressão, obtemos a função de log-verossimilhança marginal, $l(\boldsymbol{\psi}, \theta, \boldsymbol{\beta})$.

$$l(\boldsymbol{\psi}, \theta, \boldsymbol{\beta}) = \sum_{j=1}^{n} \Bigg[d \log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d)) - (1/\theta + d) \log(1 + \theta \sum_{j=1}^{n} \Lambda_0(t) \exp(\boldsymbol{x}^t \boldsymbol{\beta})) + \\ + \sum_{j=1}^{n} \delta_j (\boldsymbol{x}^t \boldsymbol{\beta} + \log(\lambda_0(t))) \Bigg]$$

Maximizando a função de log-verossimilhança, pode-se obter o estimador de máxima verossimilhança para $\boldsymbol{\psi}, \theta$ e $\boldsymbol{\beta}$.

Revisão de Literatura

As formas paramétricas da função de risco então que a verossimilhança marginal é também paramétrica e nós podemos usar as técnicas clássicas de máximo verossimilhança para estimar os parâmetros de interesse. O risco e a função de risco acumulada para a distribuição exponencial são dadas por,

$$\lambda(t) = \frac{f(t)}{S(t)} = \frac{\alpha e^{-\alpha t}}{e^{-\alpha t}} = \alpha \tag{26}$$

e

$$\Lambda(t) = \alpha t, \tag{27}$$

Revisão de Literatura

A função de log verossimilhança marginal para a fragilidade gama com função de risco base exponencial é dada por,

$$l(\psi, \theta, \beta) = \sum_{j=1}^{n} \left[d\log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d)) - (1/\theta + d)\log(1 + \theta\sum_{j=1}^{n} \alpha t e^{\boldsymbol{x}^t\boldsymbol{\beta}}) + \right.$$
$$\left. + \sum_{j=1}^{n} \delta_j(\boldsymbol{x}^t\boldsymbol{\beta}) + \log(\alpha) \right]$$

A função de risco e a função acumulada de risco para a distribuição Weibull é,

$$\lambda(t) = \frac{\eta t^{\eta-1}}{\alpha^{\eta}} \quad e \quad \Lambda(t) = \left(\frac{t}{\alpha}\right)^{\eta}$$

A função de log verossimilhança marginal para a fragilidade gama com função de risco base Weibull,

$$l(\psi, \theta, \beta) = \sum_{j=1}^{n} \left[d\log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d)) - (1/\theta + d)\log(1 + \theta\sum_{j=1}^{n} \left(\frac{t}{\alpha}\right)^{\eta} e^{\boldsymbol{x}^t\boldsymbol{\beta}}) + \right.$$
$$\left. + \sum_{j=1}^{n} \delta_j(\boldsymbol{x}^t\boldsymbol{\beta}) + \log\left(\frac{\eta t^{\eta-1}}{\alpha^{\eta}}\right) \right].$$

Revisão de Literatura

Como um exemplo, a primeira derivada para o modelo de fragilidade gama com função de risco base weibull e uma covariável é,

$$l(\psi, \theta, \beta) = \sum_{j=1}^{n} \left[d \log(\theta) - \log(\Gamma(1/\theta)) + \log(\Gamma(1/\theta + d)) - (1/\theta + d) \log(1 + \theta \sum_{j=1}^{n} \left(\frac{t}{\alpha}\right)^{\eta} e^{\beta^t x_j}) + \sum_{j=1}^{n} \delta_j \left(\beta^t X_j + \log\left(\frac{\eta t^{\eta-1}}{\alpha^{\eta}}\right)\right) \right]$$

$$\frac{\partial l(\eta, \alpha, \theta, \beta)}{\partial \alpha} = \sum_{j=1}^{n} \left[\frac{(1 + \theta d)\, \frac{\eta}{\alpha} \sum_{j=1}^{n} \left(\frac{t}{\alpha}\right)^{\eta} e^{\beta^t x_j} x_j}{\left(1 + \sum_{j=1}^{n} \left(\frac{t}{\alpha}\right)^{\eta} e^{\beta^t x_j}\right)} - d\,\frac{\eta}{\alpha} \right]$$

Revisão de Literatura

$$\frac{\partial l\left(\eta,\alpha,\theta,\beta\right)}{\partial\eta}=\sum_{j=1}^{n}\left[\frac{-\left(1+\theta d\right)\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}x_{j}}\log\left(t\right)}{\left(1+\theta\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}x_{j}}\right)}+\sum_{j=1}^{n}\delta_{j}\left(\frac{1}{\eta}+\log(t)\right)\right]$$

$$\frac{\partial l\left(\eta,\alpha,\theta,\beta\right)}{\partial\theta}=\sum_{j=1}^{n}\left[\frac{-\left(\frac{1}{\theta}+d\right)\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}x_{j}}}{\left(1+\theta\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}x_{j}}\right)}+\frac{1}{\theta^{2}}\log\left(1+\theta\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}x_{j}}\right)+\frac{d}{\theta}+\frac{\Gamma'\left(\frac{1}{\theta}\right)}{\theta^{2}\Gamma\left(\frac{1}{\theta}\right)}\right]$$

Revisão de Literatura

$$\frac{\partial l\left(\eta,\alpha,\theta,\beta\right)}{\partial\beta}=\sum_{j=1}^{n}\left[\frac{-(1+\theta d)\displaystyle\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}\boldsymbol{x}_{j}}\boldsymbol{X_{j}}}{\left(1+\theta\displaystyle\sum_{j=1}^{n}\left(\frac{t}{\alpha}\right)^{\eta}e^{\beta^{t}}\boldsymbol{X_{j}}\right)}+\sum_{j=1}^{n}\delta_{j}\boldsymbol{X_{j}}\right]$$

As estimativas de Máxima Verossimilhança podem ser obtidas pelo conjunto de cada uma das primeiras derivadas igualadas a zero e resolvidas nos parâmetros de interesse.

Matriz de Variância e Covariância Assintótica

Matriz de Variância e Covariância Assintótica dos modelos pode ser derivada da expressão da log verossimilhança. Dado $\boldsymbol{H}$ ser a Matriz Hessiana da segunda derivadas da função de logverossimilhança marginal $l(\psi, \theta, \beta)$. O negativo no valor esperado da matriz Hessiana é conhecido como Informação de Fisher. A letra $\boldsymbol{I}$ denota a Matriz de Informação de Fisher.

$$\boldsymbol{I}(\psi, \theta, \beta) = -E(H(\psi, \theta, \beta)) \tag{28}$$

A matriz de informação observada I, é então o negativo da matriz Hessiana.

$$I(\psi, \theta, \beta) = -H(\psi, \theta, \beta) \tag{29}$$

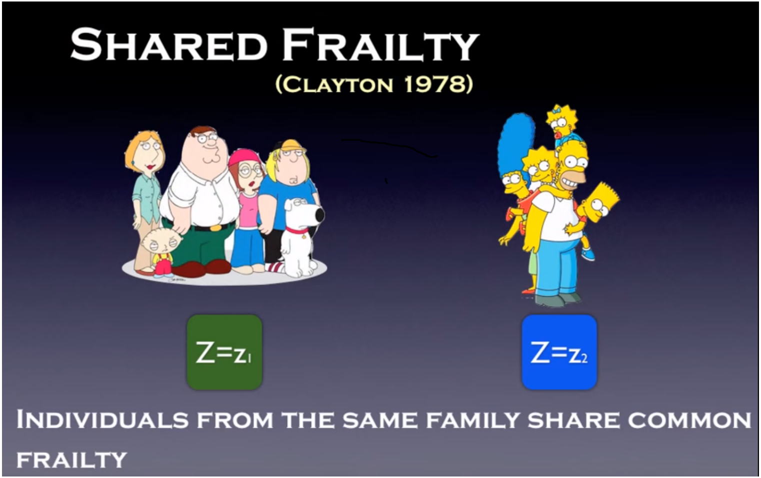
Portanto, pode-se obter a matriz de variância-covariância assintótica das estimativas tomando o inverso da matriz de informação de Fisher, e a matriz de variância-covariância estimada pode ser obtida do inverso da matriz de informação observada, avaliando-a nos valores reais dos estimadores de máxima verossimilhança.

Revisão de Literatura

O modelo de fragilidade compartilhada

Revisão de Literatura

O modelo de fragilidade compartilhada



Revisão de Literatura

O modelo de fragilidade compartilhada

Wienke (2010) define o modelo de fragilidade compartilhada da seguinte maneira, suponha que exitem n grupos e que o i-ésimo grupo possui n_i indivíduos, os quais estão associados com uma fragilidade não observada Z_i, $(1 \leq i \leq n)$. Um vetor $\boldsymbol{X}_{ij}$, $(1 \leq i \leq n, 1 \leq j \leq n_i)$ está associado com o j-ésimo indivíduo do i-ésimo grupo.
O modelo de fragilidade compartilhada é dado por

$$\lambda_{ij}(t) = z_i \lambda_0(t) \exp(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}) = z_i \lambda_{ij,c}(t) \tag{30}$$

onde, $\lambda_{ij,c}(t)$ é o risco do indivíduo j do grupo i depois do efeito da fragilidade do grupo ter sido avaliado (DUCHATEAU ; JANSSEN, 2008).

Revisão de Literatura

O modelo de fragilidade compartilhada

Quando a informação da covariável do indivíduo j do grupo i, $\boldsymbol{x}_{ij}$ é igual a $\boldsymbol{x}$, podemos reescrever a equação (30) da seguinte forma

$$\lambda_{ij}(t) = z_i \lambda_{x,c}(t) \tag{31}$$

Revisão de Literatura

Função de sobrevivência conjunta

De acordo com Duchateau e Janssen (2008), este modelo induz correlação entre os tempos de eventos de indivíduos de um mesmo grupo. Neste caso, é investigada a função de sobrevivência conjunta para um grupo.
Assumindo $j = 1, \ldots, n_i$, então a função de sobrevivência condicional conjunta para o grupo i é dada por

$$S_i(\boldsymbol{t}_{ni}) = \exp\left[-z_i(\Lambda_0(t_1)\exp(\boldsymbol{x}_{i1}^t\boldsymbol{\beta}) + \ldots + \Lambda_0(t_{ni})\exp(\boldsymbol{x}_{in_i}^t\boldsymbol{\beta}))\right] \quad (32)$$

com $\boldsymbol{t}_{ni} = (t1, \ldots, t_{ni})$.

Transformação de Laplace da densidade de fragilidade

Função Característica

$\psi(s) = E(\exp(isZ))$

Função Geradora de Momentos

$E(\exp(sZ))$

Transformação de Laplace para variável aleatória positiva

$\mathcal{L}(s) = E(\exp(-sZ)) = \int_0^\infty \exp(-sZ) f(z) dz$

Geradora do n-ésimo momento

Transformação de Laplace da n-ésima derivada

$$\mathcal{L}^{(n)}(s) = (-1)^n \int_0^\infty Z^n \exp(-sZ) f(z) dz$$

Avaliada em $s = 0$

$$\mathcal{L}^{(n)}(s = 0) = (-1)^n \int_0^\infty Z^n f(z) dz = (-1)^n E(Z^n)$$

Transformação de Laplace Gama

$$\mathcal{L}(s) = (1 + \theta s)^{-1/\theta}$$

Revisão de Literatura

Função de sobrevivência conjunta

Usando a notação da equação (31) a função de sobrevivência conjunta para um grupo de tamanho n com informação da covariável $\mathbb{X} = (\boldsymbol{x}_1^t, \ldots, \boldsymbol{x}_n^t)^t$ é obtida da função de sobrevivência condicional conjunta integrando sob a fragilidade com sua respectiva distribuição, da seguinte maneira

$$\begin{aligned} S_{\mathbb{X},f}(\boldsymbol{t}_n) &= \int_{-\infty}^{+\infty} \exp(-z\Lambda_{\mathbb{X},c}(\boldsymbol{t}_n)) f_Z(z) dz \\ &= E[\exp(-Z\Lambda_{\mathbb{X},c}(\boldsymbol{t}_n))] \end{aligned} \tag{33}$$

com $Z \sim f_Z$.

Revisão de Literatura

Função de sobrevivência conjunta

É possível reescrevermos a função de sobrevivência conjunta em função da transformação de Laplace. A última linha da equação (33) é a transformação de Laplace de Z, $\mathcal{L}(s) = E[\exp(-Zs)]$, em que $s = \Lambda_{\mathbb{X},c}(\boldsymbol{t}_n)$, assim podemos escrever

$$S_{\mathbb{X},f}(\boldsymbol{t}_n) = \mathcal{L}(\Lambda_{\mathbb{X},c}(\boldsymbol{t}_n)) \tag{34}$$

Revisão de Literatura

Função de sobrevivência populacional

A função de sobrevivência populacional para os indivíduos tendo informação da covariável $\mathbb{X}$ é obtida a partir da função de sobrevivência conjunta $S_{ij}(t) = \exp(-z_i \Lambda_{\boldsymbol{x},c}(t))$ integrando fora da fragilidade com a respectiva função densidade da fragilidade

$$S_{\boldsymbol{x},f}(t) = \int_0^{\infty} \exp(-z_i \Lambda_{\boldsymbol{x},c}(t)) f_Z(z) dz \tag{35}$$

Revisão de Literatura

Função de sobrevivência populacional

A função de sobrevivência populacional pode também ser obtida por meio da transformação de Laplace da seguinte forma

$$S_{\boldsymbol{x},f}(t) = \mathcal{L}(\Lambda_{\boldsymbol{x},c}(t)) \tag{36}$$

Revisão de Literatura

A estatística τ de Kendall's

Para dois grupos i e k de tamanho dois aleatoriamente escolhidos, os tempos de eventos são (T_{i1} e T_{i2}) e (T_{k1} e T_{k2}). Sendo assim, a estatística τ de Kendall's (KENDALL, 1938), é definida da seguinte maneira

$$\tau = E[sign((T_{i1} - T_{k1})(T_{i2} - T_{k2}))] \quad (37)$$

em que $sign(x) = -1, 0, 1$ para $x < 0, x = 0, x > 0$.

Revisão de Literatura

Modelos de fragilidade paramétricos

O modelo de fragilidade Gama

De acordo com Duchateau e Janssen (2008) é usual que a escolha dos parâmetros da distribuição para modelagem da fragilidade é feita de maneira tal que $E(Z) = 1$. Assim sendo, o modelo gama apresenta a seguinte parametrização $(1/\theta, 1/\theta)$, resultando na seguinte função densidade de probabilidade :

$$f_Z(z) = \frac{z^{1/\theta - 1}\exp(-z/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \tag{38}$$

com a seguinte transformação de Laplace

$$\mathcal{L}(s) = (1 + \theta s)^{-1/\theta} \tag{39}$$

Revisão de Literatura

Modelos de fragilidade paramétricos

O modelo de fragilidade Gama

As funções de sobrevivência conjunta e populacional são dadas respectivamente por

$$S_{\mathbb{X},f}(t_1,\ldots,t_n) = (1+\theta\Lambda_{\mathbb{X},c}(\boldsymbol{t}_n))^{-1/\theta} \tag{40}$$

$$S_{\boldsymbol{x},f}(t) = \mathcal{L}(\Lambda_{\boldsymbol{x},c}(t)) = (1+\theta\Lambda_{\boldsymbol{x},c}(t))^{-1/\theta} \tag{41}$$

Para o modelo Gama o τ de Kendall's que é uma medida de dependência global é calculado utilizando a primeira e a segunda derivada da transformação de Laplace. Neste caso o τ de Kendall's é dado por

$$\tau = \frac{\theta}{\theta+2} \tag{42}$$

Revisão de Literatura

Modelos de fragilidade paramétricos

O modelo de fragilidade Lognormal

O modelo de fragilidade lognormal é definido da seguinte forma

$$f_Z(z) = \frac{1}{z\sqrt{2\pi\gamma}} \exp\left(-\frac{(\log z)^2}{2\gamma}\right) \tag{43}$$

com $\gamma > 0$. A média e a variância da fragilidade são dadas respectivamente por

$$\begin{aligned} E(Z) &= \exp(\gamma/2) \\ Var(Z) &= \exp(2\gamma) - \exp(\gamma) \end{aligned}$$

Revisão de Literatura

Modelos de fragilidade semi-paramétricos

McGilchrist e Aisbett (1991) usaram um expressão de representação alternativa para o modelo de fragilidade semi-paramétrico, o qual é representado da seguinte forma

$$\lambda_{ij}(t) = \lambda_0(t) \exp\left(\boldsymbol{x}_{ij}^t \boldsymbol{\beta} + w_i\right) \tag{44}$$

em que $w_i = \log z_i$. Onde w_i denota o efeito aleatório presente no modelo com variância γ e z_i denota a fragilidade com variância θ.

Revisão de Literatura

Modelos de fragilidade semi-paramétricos

Fragilidades Gama

No modelo de fragilidade Gama as fragilidades z_i seguem a seguinte função densidade de probabilidade, com parâmetros $\nu > 0$ e $\eta > 0$.

$$f_Z(z) = \frac{\eta^\nu z^{\nu-1} \exp(-\eta z)}{\Gamma(\nu)} \qquad 0 \leq z < \infty \tag{45}$$

No entanto, é comum adotarmos uma parametrização em que consideramos uma distribuição Gama uni-paramétrica com média 1 e variância θ. E assim, a densidade passa a ser como em (38).

Revisão de Literatura

Modelos de fragilidade semi-paramétricos

Fragilidades Gama

Além disso, a modelagem é baseada no efeito aleatório $W = \log(Z)$ ao invés das fragilidades. Dada a função densidade para os $Z_i's$ em (38), então a função densidade para os $W_i's$ corresponde a

$$f_W(w) = \frac{(\exp(w))^{1/\theta}\exp(-\exp(w)/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \tag{46}$$

Revisão de Literatura

Modelos de fragilidade semi-paramétricos

Fragilidades Lognormal

Considerando a situação em que as fragilidades $z_i s$ seguem uma distribuição lognormal com parâmetros μ e γ representada pela função densidade de probabilidade dada pela equação (47) (DUCHATEAU ; JANSSEN, 2008).

$$f_Z(z) = \frac{1}{z\sqrt{2\pi\gamma}} \exp\left(-\frac{1}{2\gamma}(\log z - \mu)^2 \right) \tag{47}$$

Então os efeitos aleatórios $w_i' s$ seguem uma distribuição normal com média μ e variância γ.

Revisão de Literatura

Estimação dos parâmetros no modelo paramétrico

No ajuste paramétrico, a estimação é basada na verossimilhança marginal em que as fragilidades tem sido integradas fora pela média da verossimilhança condicional com a respectiva distribuição da fragilidade. Sob algumas suposições, a log-verossimilhança marginal dos dados observados $\boldsymbol{u} = \{\boldsymbol{u}_{ij}; i \in I, j \in J_i\}$ pode ser escrita como

$$
\begin{aligned}
\ell_{marg}(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\gamma}; \boldsymbol{u}|\boldsymbol{\tau}) &= \sum_{i=1}^{n} \Bigg\{ \left[\sum_{j=1}^{n_i} \delta_{ij} \left(\log(\lambda_0(y_{ij})) + \boldsymbol{x}_{ij}^t \boldsymbol{\beta} \right) \right] \\
&+ \log \left[(-1)^{d_i} \mathcal{L}^{(d_i)} \left(\sum_{j=1}^{n_i} \Lambda_0(y_{ij}) \exp(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}) \right) \right] \\
&- \log \left[\mathcal{L} \left(\sum_{j=1}^{n_i} \Lambda_0(\tau_{ij}) \exp(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}) \right) \right] \Bigg\}
\end{aligned}
\tag{48}
$$

Revisão de Literatura

Estimação dos parâmetros no modelo paramétrico

com $d_i = \sum_{j=1}^{n_i} \delta_{ij}$ o número de eventos no i-ésimo grupo, e $\mathcal{L}^{(q)}(.)$ a q-ésima derivada da transformação de Laplace da distribuição da fragilidade.

As estimativas de $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são obtidas maximizando a log-verossimilhança, isto pode ser facilmente feito se for possível calcular as derivadas de ordem superior da transformação de Laplace $\mathcal{L}^{(q)}(.)$ até $q = \max\{d_1, \ldots, d_n\}$ (MUNDA et al., 2012).

Revisão de Literatura

Estimação dos parâmetros no modelo semi-paramétrico

Verossimilhança parcial penalizada

De cordo com Duchateau e Janssen (2008) na abordagem da verossimilhança parcial penalizada a verossimilhança completa dos dados consiste de duas partes. A primeira parte é a verossimilhança condicional dos dados dado as fragilidades, enquanto que a segunda parte corresponde a distribuição das fragilidades.

$$l_{ppl}(\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{w}) = l_{part}(\boldsymbol{\beta}, \boldsymbol{w}) - l_{pen}(\gamma, \boldsymbol{w}) \tag{49}$$

em que, com $\eta_{ij} = \boldsymbol{x}_{ij}^t \boldsymbol{\beta} + w_i$ e $\boldsymbol{\eta} = (\eta_{11}, \ldots, \eta_{sn_s})$,

Revisão de Literatura

Estimação dos parâmetros no modelo semi-paramétrico

Verossimilhança parcial penalizada

$$l_{part}(\boldsymbol{\beta}, \boldsymbol{w}) = \sum_{i=1}^{s} \sum_{j=1}^{n_i} \delta_{ij} \left[\eta_{ij} - \log \left(\sum_{qw \in R(y_{ij})} \exp(\eta_{qw}) \right) \right] \tag{50}$$

e

$$l_{pen}(\gamma, \boldsymbol{w}) = - \sum_{i=1}^{s} \log f_W(w_i) \tag{51}$$

Revisão de Literatura

Estimação dos parâmetros no modelo semi-paramétrico

Função de penalidade no modelo com efeito aleatório normal

Para os efeitos aleatórios $w_i, i = 1, \ldots, s$, tendo uma densidade normal com média zero e variância γ, temos que

$$l_{pen}(\gamma, \boldsymbol{w}) = \frac{1}{2} \sum_{i=1}^{s} \left(\frac{w_i^2}{\gamma} + \log(2\pi\gamma) \right) \tag{52}$$

Segundo Duchateau e Janssen (2008) a maximização da log-verossimilhança parcial penalizada é feita por meio de uma algoritmo iterativo em que é avaliado em dois passos : *"inner loop"* e *"outer loop"*.

Revisão de Literatura

Estimação dos parâmetros no modelo semi-paramétrico

Função de penalidade no modelo com fragilidade gama

Considerando que os efeitos aleatórios são representados pela função densidade $f_W(w)$ dada em (46). A função de penalidade no modelo tem a seguinte representação

$$l_{pen}(\theta, \boldsymbol{w}) = -\frac{1}{\theta}\sum_{i=1}^{s}\Big(w_i - \exp(w_i)\Big) \tag{53}$$

A maximização da log-verossimilhança no modelo com fragilidade gama é similar à maximização vista para o modelo com efeitos aleatórios seguindo uma distribuição normal.

Introdução

Diabetes Mellitus

Introdução

Diabetes Mellitus

- Doença caracterizada pela elevação da glicose no sangue (hiperglicemia).

Introdução

Diabetes Mellitus

- Doença caracterizada pela elevação da glicose no sangue (hiperglicemia).
- Causada devido à defeitos na secreção ou na ação do hormônio insulina.

Introdução

Diabetes Mellitus

- Doença caracterizada pela elevação da glicose no sangue (hiperglicemia).
- Causada devido à defeitos na secreção ou na ação do hormônio insulina.
- Dois tipos de manifestação : Diabetes tipo 1 e Diabetes tipo 2.

Introdução

Diabetes Mellitus

- Doença caracterizada pela elevação da glicose no sangue (hiperglicemia).
- Causada devido à defeitos na secreção ou na ação do hormônio insulina.
- Dois tipos de manifestação : Diabetes tipo 1 e Diabetes tipo 2.

Diabetes tipo 1

- Resultado da formação de anticorpos pelo próprio organismo contra as células beta pancreáticas, levando a deficiência de insulina.

Introdução

Diabetes Mellitus

- Doença caracterizada pela elevação da glicose no sangue (hiperglicemia).
- Causada devido à defeitos na secreção ou na ação do hormônio insulina.
- Dois tipos de manifestação : Diabetes tipo 1 e Diabetes tipo 2.

Diabetes tipo 1

- Resultado da formação de anticorpos pelo próprio organismo contra as células beta pancreáticas, levando a deficiência de insulina.

Diabetes tipo 2

- É observado um quadro de resistência insulínica. Ocorre em cerca de 90% dos pacientes notificados com o Diabetes.

Introdução

Retinopatia Diabética

Introdução

Retinopatia Diabética

- Complicação que ocorre quando o excesso de glicose no sangue danifica os vasos sanguíneos dentro da retina, geralmente está ligada à maneira inadequada de se tratar a diabetes.

Introdução

Retinopatia Diabética

- Complicação que ocorre quando o excesso de glicose no sangue danifica os vasos sanguíneos dentro da retina, geralmente está ligada à maneira inadequada de se tratar a diabetes.
- Compromete a visão do portador, podendo chegar a cegueira caso não seja tratada.

Introdução

Retinopatia Diabética

- Complicação que ocorre quando o excesso de glicose no sangue danifica os vasos sanguíneos dentro da retina, geralmente está ligada à maneira inadequada de se tratar a diabetes.
- Compromete a visão do portador, podendo chegar a cegueira caso não seja tratada.
- Um tratamento à *laser* é necessário na fase mais agressiva da doença.

Material e Métodos

Material

O conjunto de dados utilizado neste trabalho é derivado do trabalho feito por Blair et al. em 1976, na Irlanda do Norte. A base de dados contém 394 observações de 197 pacientes com retinopatia diabética que faziam um tratamento de fotocoagulação à *laser*. As variáveis presentes no banco de dados são : indicador do indivíduo, olho, status, tratamento, idade, tipo de *laser*, tipo de diabetes e escore de risco de cegueira (valores maiores que 6 indicam alto risco). Esse conjunto de dados é apenas uma amostra aleatória do conjunto de dados original. É possível ter acesso à esses dados através do comando *data(rms)* no software R.

Resultados e Discussão

Análise descritiva

Table – Resumo das estimativas de Kaplan-Meier para os grupos tratado e controle.

Grupo	**N**	**Eventos**	**Mediana**	$\boldsymbol{IC_{95\%}}$
Controle	197	101	43,7	$[31,6; 59,8]$
Tratado	197	54	NA	NA

Table – Tempo de acompanhamento mediano.

Grupo	**N**	**Eventos**	**Mediana**	$\boldsymbol{IC_{95\%}}$
Controle	197	96	51,1	[46,3 ; 55,3]
Tratado	197	143	50,6	[46,2 ; 54,6]

Resultados e Discussão

Análise descritiva

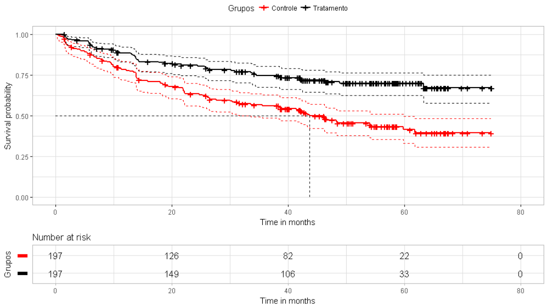


FIGURE – Curvas de sobrevivência de Kaplan-Meier para os grupos tratado e controle.

Resultados e Discussão

Análise descritiva

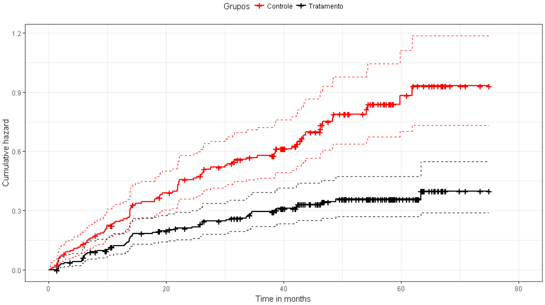


FIGURE – Curvas de risco para os grupos tratado e controle.

Resultados e Discussão

Análise descritiva

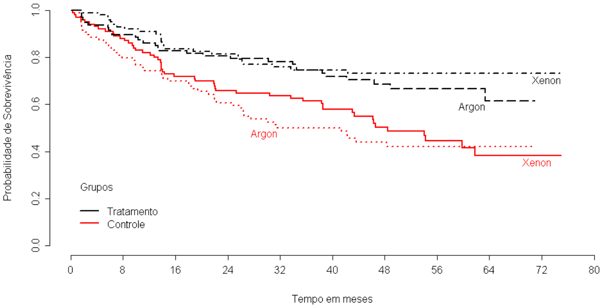


FIGURE – Curvas de sobrevivência para os tipos de *lasers*.

Resultados e Discussão

Análise descritiva

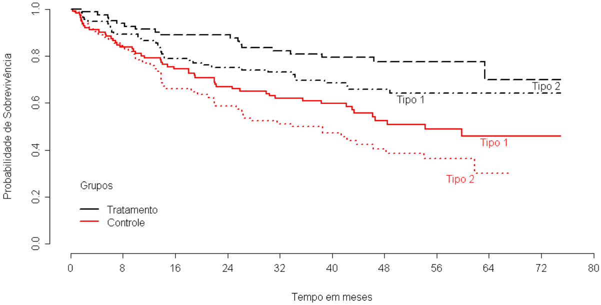


FIGURE – Curvas de sobrevivência para os pacientes com diabetes do tipo 1 e do tipo 2.

Resultados e Discussão

Análise descritiva

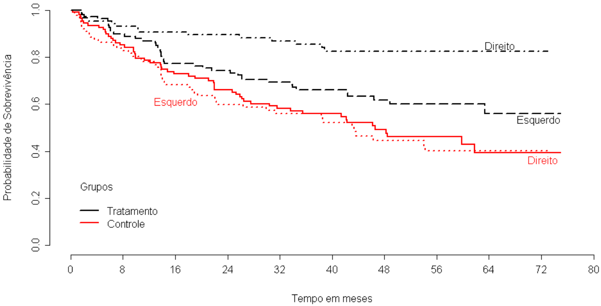


FIGURE – Curvas de sobrevivência para os pacientes que faziam tratamento no olho direito e no olho esquerdo.

Resultados e Discussão

Análise descritiva

TABLE – Testes Log-rank e Peto para diferença entre as curvas de sobrevivência dos grupos estudados.

	Log-rank		**Peto**	
Grupo	χ^2	**P-valor**	χ^2	**P-valor**
Tratado vs *Controle*	22,2	$< 0,0001$	20,7	$< 0,0001$
Laser xenon vs *Laser argon*	22,4	$< 0,0001$	20,9	$< 0,0001$
DM tipo 1 vs *DM tipo 2*	22,5	$< 0,0001$	20,8	$< 0,0001$
Olho direito vs *Olho esquerdo*	24,0	$< 0,0001$	23,1	$< 0,0001$

Resultados e Discussão

Ajuste dos modelos Univariados Semiparamétricos

TABLE – Valores de AIC para os modelos ajustados para o conjunto de dados de Retinopatia Diabética.

Modelos	**AIC**
Cox clássico	1707,931
Fragilidade gama	1436,263
Fragilidade inv. gaussiana	1541,557
Fragilidade t	1549,709

Resultados e Discussão

Modelo selecionado

TABLE – Estimativas dos parâmetros do modelo de fragilidade gama.

Covariáveis	**Coef**	**exp(Coef)**	**SE**	χ^2	**DF**	**P-valor**
Tratamento	$-0,5714$	0,5647	0,2365	5,8400	1,0	0,0160
Risco	0,1932	1,2131	0,0827	5,4500	1,0	0,0200
Tratamento*Tipo2	$-1,1427$	0,3190	0,3679	9,1500	1,0	0,0025
Fragilidade				188,64	128,3	$< 0,001$

Avaliação da qualidade do ajuste

Concordância do modelo	Variância do efeito aleatório (θ)
0,881	2,000

Conclusão Parcial do Estudo - Ajuste Semiparamétrico Univariado

- O tratamento realmente surtiu efeito sob a cegueira dos pacientes ;
- Os indivíduos que tinham diabetes do tipo 2 eram mais suscetíveis a obter a cegueira, quando esses não faziam o tratamento.
- Os pacientes que faziam tratamento com o *laser xenon* tinham maior probabilidade de sobrevivência, ou seja, de não cegar.
- Os indivíduos que faziam tratamento no olho esquerdo tinham maiores chances de cegar.

- Mediante os resultados obtidos, é possível afirmarmos que os modelos de fragilidade univariados semiparamétricos compõem um ferramenta estatística poderosa para tratamento de dados de sobrevivência na presença de heterogeneidade entre os indivíduos, quando comparados à modelagem convencional de Cox.

Resultados e Discussão

Ajuste paramétrico - Fragilidade Compartilhada

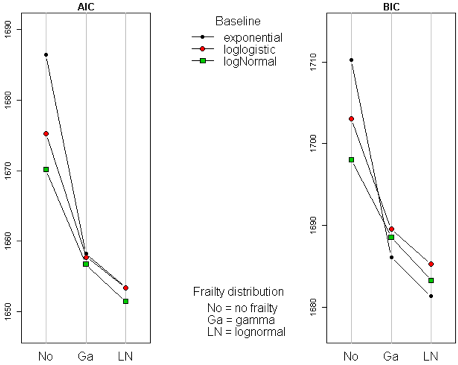


FIGURE – Valores de AIC para as distribuições do risco e da fragilidade.

Resultados e Discussão

Ajuste paramétrico

Table – Estimativas dos parâmetros para o modelo paramétrico sem a fragilidade e o modelo de fragilidade compartilhada.

MPSF (AIC = 1664,885)				
Covariável	**Coef**	**exp(Coef)**	**SE**	**P-valor**
Tratamento_tratado	-0,491	0,612	0,219	0,025 (<0,05)
Olho_esquerdo	0,319	1,376	0,163	0,050
Diabetes_tipo 2	0,310	1,363	0,199	0,119
Tratado_diabetes tipo 2	-0,782	0,457	0,352	0,026 (<0,05)
MPFCL (AIC = 1645,542), $\tau = 0,316$				
Covariável	**Coef**	**exp(Coef)**	**SE**	**P-valor**
Tratamento_tratado	-0,656	0,519	0,240	0,006 (<0,01)
Olho_esquerdo	0,490	1,632	0,185	0,008 (<0,01)
Diabetes_tipo 2	0,389	1,475	0,268	0,147
Tratado_diabetes tipo 2	-0,908	0,403	0,371	0,015 (<0,05)

Resultados e Discussão

Ajuste semi-paramétrico

TABLE – Estimativas dos parâmetros para o modelo semi-paramétrico sem a fragilidade e o modelo de fragilidade compartilhada semi-paramétrico gama.

MSSF (AIC = 1711,551)				
Covariáveis	**Coef**	**exp(Coef)**	**SE**	**P-valor**
Tratamento_tratado	-0,482	0,617	0,220	0,011 (<0,05)
Olho_esquerdo	0,322	1,380	0,163	0,022 (<0,05)
Diabetes_tipo 2	0,311	1,365	0,200	0,116
Tratado_diabetes tipo 2	-0,785	0,456	0,352	0,0097 (<0,01)
MFCSG (AIC = 1697,694), $\boldsymbol{\tau = 0,348}$				
Covariáveis	**Coef**	**exp(Coef)**	**SE**	**P-valor**
Tratamento_tratado	-0,635	0,530	0,233	0,008 (<0,01)
Olho_esquerdo	0,467	1,595	0,179	0,011 (<0,05)
Diabetes_tipo 2	0,366	1,442	0,270	0,174
Tratado_diabetes tipo 2	-0,916	0,400	0,366	0,013 (<0,05)

Resultados e Discussão

Adequação do modelo paramétrico

Não foi possível avaliar o ajuste do modelo paramétrico. Sendo esta, uma perspectiva futura do estudo.

Resultados e Discussão

Adequação do modelo semi-paramétrico

Foi avaliado os resíduos do modelo ajustado a fim de verificar se os valores preditos pelo modelo se aproximaram bem dos valores observados.

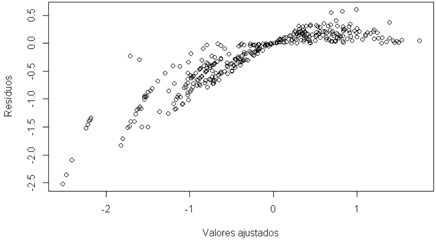


FIGURE – Resíduos *vs* valores ajustados para o modelo de fragilidade semi-paramétrico gama.

Conclusão

Diante do exposto pode-se concluir que os modelos de fragilidade compõe uma ferramenta estatística poderosa do que os modelos convencionais de sobrevivência para poder-se avaliar a influência de covariáveis na variável resposta.

Por meio dos testes Log-rank e Peto foi possível comprovar que o *laser xenon* teve maior eficácia para diminuir o risco da cegueira nos pacientes e que a diabetes do tipo 2 quando tratada teve melhores chances de não obter a cegueira.

Apesar de não ter sido avaliado os resíduos do modelo paramétrico de fragilidade compartilhada, este se mostrou mais competitivo, gerando o menor valor de AIC. Portanto, esse modelo tem grandes chances de ter sido o melhor ajuste obtido para essa base de dados.

Referências I

COLOSIMO, E. A. ; GIOLO, S. R. *Análise de Sobrevivência Aplicada*. 1. ed. São Paulo : Edgard Blucher, 2006.

DINIZ, C. ; LOUZADA, F. Modelagem estatística para risco de crédito. *ABE, São Paulo-SP*, 2012.

DUCHATEAU, L. ; JANSSEN, P. *The Frailty Models*. 1. ed. New York : Springer Science & Business Media, 2008.

KAPLAN, E. L. ; MEIER, P. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958.

KENDALL, M. G. A new measure of rank correlation. *Biometrika*, JSTOR, v. 30, n. 1/2, p. 81–93, 1938.

MCGILCHRIST, C. ; AISBETT, C. Regression with frailty in survival analysis. *Biometrics*, JSTOR, v. 47, p. 461–466, 1991.

Referências II

MUNDA, M. et al. Parfm : parametric frailty models in r. *Journal of Statistical Software*, Foundation for Open Access Statistics, v. 51, n. 11, p. 1–20, 2012.

PAULA, G. A. *Modelos de regressão : com apoio computacional*. São Paulo : IME-USP, 2004.

STRAPASSON, E. *Comparaçao de modelos com censura intervalar em análise de sobrevivência*. Tese (Doutorado) — Universidade de São Paulo, 2007.

WIENKE, A. *Frailty Models in Survival Analysis*. Boca Raton : CRC Press, 2010.