Universidade de Brasília
Departamento de Estatística

Inferência Bayesiana do modelo Pareto tipo IV em Análise de Sobrevivência

Guilherme Moura Foly de Freitas
10/0012019

Monografia apresentada para obtenção do título de Bacharel em Estatística.

Brasília
2015

Guilherme Moura Foly de Freitas
10/0012019

Inferência Bayesiana do modelo Pareto tipo IV em Análise de Sobrevivência

Orientador:
Prof. **Dr. Eduardo Yoshio Nakano**

Monografia apresentada para obtenção do título de Bacharel em Estatística.

Brasília
2015

AGRADECIMENTOS

Agradeço primeiro à Deus, por me guiar pelo caminho que eu trilhei. Sou grato aos professores que me ofereceram à formação necessária para chegar até aqui. Sou grato também à família, aos amigos e aos colegas, por me acompanharem por todo o caminho. E, em particular, sou grato ao professor Eduardo Nakano pela sua atenção e dedicação na realização deste trabalho.

RESUMO

Inferência Bayesiana do modelo Pareto tipo IV em Análise de Sobrevivência

Este trabalho pretende apresentar uma inferência bayesiana para os parâmetros da distribuição Pareto do tipo IV em dados de sobrevivência. Para isso, será utilizado um teste de significância genuinamente bayesiano (FBST - *Full Bayesian Significance test*), para testar os parâmetros dessa distribuição. Será proposto o uso de simulações via *Markov Chain Monte Carlo* (MCMC) para a obtenção da posteriori. Essa metodologia será aplicada, posteriormente, em um conjunto de dados reais. Todas as simulações e estimativas foram produzidas e geradas pelo software free, R.

Palavras-chave: Inferência Bayesiana, MCMC, Pareto IV, Análise de Sobrevivência

SUMÁRIO

LISTA DE TABELAS **6**

1 INTRODUÇÃO **7**

1.1 Objetivos 8

2 METODOLOGIA **9**

2.1 A Análise de Sobrevivência 9

2.1.1 A Censura 9

2.1.2 Os tipos de Censura 10

2.1.3 Representação dos Dados Sobrevivência 11

2.1.4 Especificando o Tempo de Sobrevivência 12

2.1.5 O método de Máxima Verossimilhança 13

2.2 A Inferência Bayesiana 15

2.2.1 O Teorema de Bayes 15

2.2.2 O FBST 17

2.2.3 Regra de Decisão e Validação no FBST 20

2.2.4 A Implementação do $Ev(\mathbf{\Theta_0}; \mathbf{x})$ 21

2.3 Simulação Monte Carlo 22

2.3.1 O MCMC 22

3 O MODELO PARETO TIPO IV **24**

3.1 A verossimilhança da Pareto Tipo IV 25

3.2 Obtenção da Posteriori 25

4 RESULTADOS **29**

4.1 Simulações 29

4.2 Influência na escolha da priori 33

5 APLICAÇÃO EM DADOS REAIS **37**

6 CONCLUSÃO **41**

REFERÊNCIAS **43**

A APÊNDICE **45**

A.1 Conjunto de dados 45

A.2 Programação no R 47

LISTA DE TABELAS

Tabela 1a -Inferência Bayesiana para os parâmetros da distribuição Pareto IV com dados simulados utilizando diferentes tamanhos de amostra e percentuais de censura 30

Tabela 1b -Inferência Bayesiana para os parâmetros da distribuição Pareto IV com dados simulados utilizando diferentes tamanhos de amostra e percentuais de censura (Continuação) 31

Tabela 2a -Influência na escolha da priori de α e θ na inferência dos parâmetros do modelo Pareto IV 34

Tabela 2b -Influência na escolha da priori de α e θ na inferência dos parâmetros do modelo Pareto IV(Continuação) 35

Tabela 3 - Inferência Bayesiana para os parâmetros da distribuição Pareto IV aplicada aos dados referentes ao tempo de sobrevivência de pacientes com câncer de pulmão 38

Tabela 4a -Conjunto de dados referentes ao trabalho de Lagakos (1978), que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão. 45

Tabela 4b -Conjunto de dados referentes ao trabalho de Lagakos (1978), que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão.(Continuação) 46

1 INTRODUÇÃO

Em diversas áreas do conhecimento, é comum ter o tempo como um objeto de interesse do seu estudo. Em muitos casos deseja-se, por exemplo, observar o tempo a que um paciente sobrevive a um determinado tratamento. Por esse motivo, podemos dizer que estamos interessados no tempo até o óbito desse paciente, ou ainda no risco que esse paciente têm de vir a óbito.

Na Estatística, utilizamos a Análise de Sobrevivência para estudar os tempos até a ocorrência de um evento de interesse. Dentro dessa área, existem diversos tipos de modelos que podem ser utilizados para representar esses tempos de morte (falhas). Podemos citar, por exemplo, a Distribuição de Pareto tipo II (Barry, 1983) que assume a existência de um risco constante ao longo do tempo (Exponencial) para todos os indivíduos da população, sendo que esse risco varia de indivíduo para indivíduo segundo uma distribuição Gama. Desta forma, tem-se que a distribuição de Pareto tipo II acomoda apenas riscos decrescentes. Existe uma série de distribuições de Pareto, denominadas Pareto do tipo I, II, III e IV (Barry, 1983 e Johnson et al., 1994) e a distribuição generalizada de Feller-Pareto (Feller, 1971).

A proposta deste trabalho será trabalhar com o modelo Pareto tipo IV de forma obter uma maior flexibilidade no ajustamento da função de risco. A escolha desse modelo se deve ao fato do mesmo apresentar, no contexto de análise de sobrevivência, uma grande flexibilidade em sua função de risco. A Pareto tipo IV apresenta uma expressão matemática fechada, algo que não acontece com o modelo generalizado de Feller-Pareto. O modelo Pareto tipo IV acomoda, como casos especiais, os modelos Pareto tipo I, II e III.

Sendo assim, o objetivo desse trabalho é realizar a inferência bayesiana do modelo Pareto tipo IV dentro do contexto de análise de sobrevivência. Serão apresentadas estimativas pontuais e intervalares dos parâmetros e também um teste de significância genuinamente bayesiano (FBST, *Full Bayesian Significance Test*) para testar seus parâmetros. Com isso, teremos uma medida de evidência que mantém as mesmas propriedades desejáveis dos p-valores da abordagem clássica, facilitando a nossa abordagem em um contexto bayesiano.

Essa metodologia será aplicada tanto em um conjunto de dados simulado quanto em um conjunto de dados reais, referentes a um estudo sobre pacientes com câncer de pulmão.

O FBST é um teste genuinamente bayesiano, pois depende exclusivamente da distribuição à posteriori dos parâmetros (Pereira e Stern, 1999). O interesse deste trabalho era testar três hipóteses: $H_0 : \alpha = 1$, $H_0 : \theta = 1$ e $H_0 : \theta$=1 e α=1. A hipótese $H_0 : \alpha = 1$ equivale a admitir que o modelo Pareto tipo III se ajusta aos dados. A hipótese $H_0 : \theta = 1$ equivale a admitir que o modelo Pareto tipo II se ajusta. Por fim, a hipótese $H_0 : \theta$=1 e α=1 é um caso particular da Pareto tipo II. Se essas hipóteses forem verificadas, a opção se dará pelo modelo mais simples (com menos parâmetros).

As distribuições à posteriori dos parâmetros foram obtidas por meio de simulações via MCMC - *Markov Chain Monte Carlo* (Gamerman, 1997), implementado através do pacote *MCMCPack* do software free, R (R Core Team, 2014), de onde todas as simulações e estimativas apresentadas nesse trabalho foram obtidas.

1.1 Objetivos

Objetivo Geral:

O objetivo deste trabalho é apresentar uma inferência bayesiana dos parâmetros do modelo Pareto tipo IV no contexto de análise de sobrevivência.

Objetivos Específicos:

1. Formular o modelo Pareto tipo IV dentro de um contexto de análise de sobrevivência;
2. Estimação pontual e intervalar dos parâmetros do modelo.
3. Realizar um teste de significância genuinamente bayesiano (FBST) para testar os parâmetros do modelo
4. Ilustração do modelo em dados reais e simulados.

2 METODOLOGIA

Temos como interesse do nosso estudo o desenvolvimento de inferências para os parâmetros da distribuição Pareto tipo IV em um contexto de Análise de Sobrevivência. Para isso, uma revisão de literatura se faz necessária.

2.1 A Análise de Sobrevivência

Em muitos casos, na Estatística, temos o tempo como uma variável de interesse do nosso estudo. Na análise de Sobrevivência, a variável resposta é, em geral, o tempo até a ocorrência de um evento de interesse. Esse tempo é denominado **tempo de falha**, podendo ser o tempo até a morte de um paciente, bem como o tempo até a cura de uma doença específica *(Colosimo, 2006)*.

2.1.1 A Censura

Dentro da análise de sobrevivência, encontramos dados com uma característica especial. Entre os dados observados existem observações parciais da resposta, chamadas de **censuras**. É comum que, em estudos de sobrevivência, alguns dos pacientes selecionados para o estudo venham a se mudar de cidade. Pode ser que o estudo tenha terminado e o paciente ainda não tenha sofrido o evento de interesse ou o paciente também pode ter morrido por causas diferentes das de interesse.

Toda a informação que foi coletada desse paciente não pode, simplesmente, ser desconsiderada do modelo. Essa informação deve receber um tratamento diferenciado, visto que as técnicas clássicas de Análise de Regressão ou de Planejamento de Experimentos não são aplicáveis nesse contexto.

2.1.2 Os tipos de Censura

Um conjunto de dados em Análise de Sobrevivência pode apresentar diversos tipos de Censura. Afinal, como estamos interessados no tempo em que um evento de interesse ocorra, há diversos motivos que podem levar o evento de interesse a não ocorrer naquele indivíduo.

Há situações em que, depois de passado todo o tempo estabelecido no estudo, alguns indivíduos ainda não experimentaram o evento de interesse. Também podem ocorrer situações em que, após o inicio do estudo, o pesquisador é obrigado a parar de coletar informações de um determinado indivíduo por algum motivo (mudança ou causa da morte diferente da de interesse, por exemplo). Nesses casos, dizemos que o tempo em que o evento de interesse ocorre de estar à direita do tempo observado, ou seja, essa observação será considerada uma **censura à direita**.

Em outras situações, o evento de interesse já ocorreu antes mesmo do estudo ter começado. Por esse motivo, quando o pesquisador coletou as informações sobre aquele indivíduo, só o que ele sabe é que o evento de interesse ocorreu aintes do inicio da pesquisa. Nesse tipo de situação, dizemos que o tempo em que o evento de interesse ocorre está à esquerda do tempo observado. Por esse motivo, esse tipo de observação será considerada uma **censura à esquerda**.

Em muitos casos, também não é possivel obter-se o tempo exato da ocorrencia o evento de interesse. Em uma situação de visitas periódicas à algum profissional de saúde, por exemplo, só o que se sabe é que o evento de interesse ocorreu dentro de um intervalo de tempo. Isso caracteriza uma **censura intervalar**.

Dentre esses três tipos de censura, a censura à direita é o tipo mais comum em Análise de Sobrevivência. Esse tipo de censura se subdivide em outras três categorias, sendo elas apresentadas à seguir:

- **Tipo I**: Após o fim do estudo, o paciente não experimenta o evento de interesse.

- **Tipo II**: Determina-se previamente um número de falhas desejado. Após observar exatamente essa quantidade de falhas, o estudo acaba.

- **Aleatória**: O evento de interesse não ocorre até o final do estudo por algum motivo. O pesquisador é obrigado, portanto, a interromper a observação daquele paciente em qualquer que seja o momento do estudo. Engloba também como caso particular as censuras do tipo I.

Como é esperado, o mecanismo de censura mais comum utilizado nos estudos de Análise de Sobrevivência envolvem as **Censuras à direita aleatórias**. Esse será o tipo de censura que será considerada neste trabalho.

Obs: Dentro de um contexto bayesiano, o tipo de censura à direita (tipo I, II ou aleatória) será irrelevante para a inferência. Isso ocorre porque todos os tipos de censura à direita irão gerar funções de verossimilhança proporcionais que, consequentemente, irão gerar a mesma distribuição a posteriori dos parâmetros, respeitando assim o princípio da verossimilhança

2.1.3 Representação dos Dados Sobrevivência

Os dados de sobrevivência do indivíduo i $(i = 1, ..., n)$ sob estudo são representados, em geral, pelo par (t_i, δ_i), sendo t_i o tempo observado (de falha ou de censura) e δ_i será uma variável indicadora de falha ou censura, isto é:

$$\delta_i = \begin{cases} 1 & \text{se } t_i \text{ é um tempo de falha;} \\ 0 & \text{se } t_i \text{ é um tempo de censura.} \end{cases}$$

2.1.4 Especificando o Tempo de Sobrevivência

Na teoria de Análise de Regressão e Planejamento de Experimentos, uma variável aleatória pode ser especificada de acordo com a sua função densidade de probabilidades, caso a distribuição seja contínua ou de acordo com a funçao de probabilidade caso a distribuição seja discreta. Em Analise de Sobrevivência, outras duas funções complementam esse tipo de especificação.

Assumimos que uma variável aleatória contínua T, não negativa ,representa o tempo de falha. Essa variável T é representada por sua **função densidade**, a sua **função de sobrevivência** e sua **função taxa de falha (ou função risco**). Estudos que consideram tempos de sobrevivência discretos podem ser vistos em Nakano e Carrasco(2006), Carrasco, et al.(2015) e Brunello e Nakano(2015).

Lee (1992), definiu a **função densidade de probabilidade**, $f(t)$, como o limite da probabilidade de um indivíduo experimentar o evento de interesse em um intervalo de tempo $(t; t + \Delta t)$ por unidade de Δt (comprimento do intervalo), ou simplesmente por unidade de tempo. É expressa da forma:

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

Onde $f(t) \geq 0$ para todo t e a área abaixo da curva de $f(t)$ é igual a 1.

Colosimo (2006), define a **função de sobrevivência**,$S(t)$, como sendo a probabilidade de um indivíduo não falhar até um certo tempo, ou seja, é a probabilidade do indivíduo sobreviver até o tempo t. A função de sobrevivência, $S(t)$, é uma função monótona, decrescente e pode ser descrita da forma:

$$S(t) = P(T > t) = \int_t^{\infty} f(x)dx;$$

Uma consequência dessa definição é que a função de sobrevivência pode ser escrita em função da função densidade acumulada, através da fórmula: $F(t) = 1 - S(t)$.

De posse das definições de probabilidade de falha e probabilidade de sobrevivência, existe uma função derivada desses mesmos conceitos. A **função taxa de falha**(ou **função risco**) é definida como o limite da probabilidade de um indivíduo falhar no intervalo $(t; t + \Delta t)$ condicionada ao evento de que esse mesmo indivíduo sobreviveu até o tempo t, dividida pelo comprimento do intervalo. Pode ser expressa da seguinte forma:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t};$$

A função $h(t)$ representa o risco instantâneo do indivíduo experimentar o evento de interesse. O esperado é que, à medida que o tempo passe, o risco do indivíduo sofrer o evento de interesse possa ser diferente. Funções de risco **crescentes** indicam que conforme o tempo transcorre, os indivíduos possuem um risco maior de experimentar o evento de interesse. Um pensamento análogo pode ser utilizado para interpretar as Funções de risco **decrescentes**. Uma função de risco **constante** indica que a chance do indivíduo sofrer o evento de interesse seja a mesma, independente do tempo.

Uma propriedade interessante dessa função, se a distribuição for contínua, é que ela pode ser relacionada com as funções de densidade e sobrevivência por meio da equação: $h(t) = \frac{f(t)}{S(t)}$.

2.1.5 O método de Máxima Verossimilhança

Uma vez que se tenha a informação das funções que definem o nosso conjunto de dados, é de nosso interesse encontrar uma forma de estimar os parâmetros dos modelos que estamos estabelecendo.

Seguimos a idéia do método de Máxima Verossilimilhança usual, calculando o produtório da função densidade. Ou seja:

$$L(\theta) = \prod_{i=1}^{n} [f(t_i; \theta)]$$

sendo θ um parâmetro genérico do modelo probabilístico de interesse.

Esse método, porém, não é válido para os dados de sobrevivência, uma vez que a informação das censuras precisam ser consideradas no modelo. O cálculo da função de verossimilhança levará em consideração não só a função densidade como também a função de sobrevivência e/ou a função de risco.

Teremos agora dois tipos de tempos que devem ser considerados. Os tempos de falha, em que o evento de interesse ocorreu, que irão contribuir na função de verossimilhança com com a sua respectiva função densidade de probabilidades $f(t)$. Já os tempos de censura, em que os indivíduos não vieram a falhar, contribuirão na verossimilhança com a sua função de sobrevivência $S(t)$, visto que a única informação que temos de um indivíduo censurado é que o seu tempo de falha é sabidamente maior que o tempo censurado.

A partir disso, a função de verossimilhança ficará da forma:

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

Onde δ_i é o indicador de censura, sendo que $\delta_i = 1$ indica que o tempo t_i é um tempo de falha e $\delta_i = 0$ indica que o tempo t_i é um tempo de censura.

2.2 A Inferência Bayesiana

Na Estatística, frequentemente utilizamos probabilidades para expressar a informação que temos sobre quantidades desconhecidas, assim como nossas crenças sobre as mesmas. Mesmo que esse tipo de tratamento seja, na maioria das vezes, informal, pode-se provar que existe uma relação entre probabilidade e informação. A regra de Bayes utiliza-se da probabilidade para representar um conjunto de crenças, de forma a desenvolver um método onde à medida que obtemos mais informação sobre aquela quantidade desconhecida, as crenças que temos sobre aquela quantidade também se modificam.

O processo de aprendizado indutivo via a regra de Bayes é conhecido como **Inferência Bayesiana**.

2.2.1 O Teorema de Bayes

Pode-se afirmar que ferramenta mais básica para realizar inferências do ponto de vista da Estatística Bayesiana é o Teorema de Bayes. Consideramos dois eventos, A e B. Assumimos o axioma da teoria da probabilidade que diz:

$$P(A \cap B) = P(B|A)P(A)$$

E assumindo também que:

$$P(A \cap B) = P(B \cap A)$$

A consequência imediata que se têm é que:

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

Se rearranjarmos a equação acima, irêmos obter:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

Tal situação pode ser interpretada da seguinte forma:

Suponha que estamos interessados no evento A. Por meio de algum estudo anterior ou até mesmo do conhecimento geral sobre o assunto, podemos estabelecer $P(A)$ como sendo a **probabilidade à priori** para o nosso modelo. No nosso contexto, essa probabilidade representa todas as crenças que temos sobre A antes de observarmos qualquer evidência relevante.

Suponha que, então, nós passamos a observar B. Essa nova informação pode fazer com que as crenças que tinhamos sobre A mudem. Dizemos que estamos interessados em obter a **probabilidade à posteriori** de A. A equação (1), portanto, nos diz que queremos revisar as crenças sobre A a partir da informação que obtivemos de B. Essa probabilidade à posteriori, $P(A|B)$, pode ser escrita como o produto da probabilidade à priori,$P(A)$, pela informação amostral que obtivemos após o início do estudo, dada por $\frac{P(B|A)}{P(B)}$.

Estendendo esse conceito para além de simples eventos, a mesma ideia pode ser aplicada para o conceito de **variáveis aleatórias**. Matematicamente, a variável aleatória é uma função que associa elementos do espaço amostral Ω a valores numéricos. Em termos formais, uma variável aleatória é uma função que associa a todo evento do espaço amostral Ω um único valor real.

Essa variável, que chamaremos sem perda de generalidade de X, assume valores x dentro de um espaço qualquer R, ou seja: $x \in R$. X terá uma distribuição, também chamada de **distribuição à priori**, $P(X)$, onde especificamos que: $P(x) = P(X = x)$, $x \in R$. Essa distribuição, da mesma forma que a probabilidade correspondente, reflete todo o conhecimento e as crenças que temos a respeito de X.

Assumindo outra variável aleatória Y, dizemos que para qualquer valor de y a expressão $P(y|X)$, com valores $P(y|x) = P(Y = y|X = x)$, é considerada função de x e pode ser chamada de **função de verossimilhança** de X para os dados y. Essa função, $P(y|X)$, têm o papel de representar a nossa informação amostral, ou tudo o que é possivel retirar de informação à partir dos dados observados.

De maneira análoga, podemos inferir que $P(X|y)$ será a nossa **distribuição à posteriori** conforme os dados, representando as novas crenças que temos a respeito de X à partir dos valores observados em Y.

Sendo assim, é possivel expressar a equação (1) pela relação:

$$\underbrace{P(X|y)}_{\text{Posteriori}} \propto \underbrace{P(y|X)}_{\text{Verossimilhança}} \times \underbrace{P(X)}_{\text{Priori}} \tag{2}$$

Onde diremos que a distribuição à posteriori será proporcional ao produto da sua função de verossimilhança pela distribuição à priori, à menos de uma constante de proporcionalidade.

Vale lembrar que as definições acima foram dadas assumindo que X e Y são variáveis aleatórias discretas. Um pensamento análogo também pode ser aplicado para variáveis aleatórias contínuas, desde que consideremos P(X) como sendo a **densidade de probabilidades** de X. Todos os outros cálculos seguem de maneira análoga para o caso contínuo.

2.2.2 O FBST

Na inferência clássica, a medida de evidência denominada p-valor é amplamente utilizadas em diversas áreas do conhecimento. Porém discute-se muito à respeito do uso dessa medida, uma vez que ela acaba entrando em conflito com as medidas de evidência Bayesianas. Isso nos mostra que, em algumas situações, o p-valor pode não ser uma boa medida de evidência para alguma hipótese estatística. (Cella, 2013)

Diversos pesquisadores da área, a começar com Jeffreys em 1961, alertaram para o conflito entre as duas medidas de evidência. Seus estudos mostravam que a evidência contra uma hipótese nula (medida através da probabilidade posterior, fator de Bayes ou da verossimilhança comparativa) pode diferir drasticamente do p-valor, sendo que, para grandes amostras, esse conflito se mostra cada vez mais evidente. De forma geral, os estudos

Berger and Selke (1987) e Berger and Delampady(1987) sugerem que os p-valores podem ser medidas altamente enganosas acerca da evidência trazida pelos dados contra a hipótese nula precisa, em determinadas situações.

O desentendimento entre as duas abordagens se dá por diversos motivos. Por um lado, o p-valor não considera em seu cálculo a hipótese alternativa, enquanto a abordagem Bayesiana considera em seus cálculos as duas hipóteses envolvidas. O argumento frequentista é que, ao se atribuir uma probabilidade à priori para a hipótese nula, podemos distribuir a probabilidade remanescente da hipótese alternativa de maneira difusa, o que resultaria em uma probabilidade posterior pequena para essa hipótese, favorecendo assim a hipótese nula.

Devido a esse tipo de conflito é que viu-se a necessidade de se desenvolver uma nova medida de evidência. Proposta por Pereira e Stern (1999) a medida de evidência genuinamente Bayesiana é obtida por meio de um procedimento denominado *Full Bayesian Signicance Test* **(FBST)**. O termo "genuinamente bayesiano" decorre do fato da medida de evidência proposta ser baseada somente na distribuição à posteriori e poder ser caracterizada dentro de um contexto de teoria de decisão. (Pereira e Stern, 1999)

O FBST testa as hipóteses baseando-se no cálculo da probabilidade a posteriori da região HPD (Highest Posterior Density) que é "tangente"ao conjunto que define a hipótese nula.

- **Definição**: Seja $\mathbf{y}$ uma variável aleatória e a quíntupla (Ω,A,F,Θ,P) o modelo estatístico paramétrico associado a essa variável, sendo Ω é o espaço amostral (conjunto dos possíveis valores de $\boldsymbol{y}$), A é uma sigma-álgebra conveniente de subconjuntos de Ω, F é uma classe de distribuições de probabilidade em A, indexadas no espaço paramétrico Θ e P é uma densidade a priori em Θ. Admita também que, após observar $\mathbf{y}$, obtêm-se a densidade à posteriori de θ, P($\theta|\boldsymbol{y}$), restringindo essa a uma função densidade de probabilidade.

Seja T_ϕ o subconjunto do espaço paramétrico onde a densidade posterior é maior do que ϕ:

$$T_\phi = \{\theta \in \Theta | P(\theta|\boldsymbol{y})) > \phi\}$$

A probabilidade de T_ϕ é a sua probabilidade à posteriori, $\kappa = \int_{T_\phi} P(\theta|x)\, d\theta$. Agora, seja P* o máximo da densidade superior sob a hipótese nula (H_0), obtida através de θ^*:

$$\theta^* \in sup_{\theta \in \Theta_0} P(\theta), \quad P^* = P(\theta^*)$$

Assim, $T^* = T_{P^*}$ é o conjunto tangente à hipótese nula com credibilidade κ^*.

Temos, portanto, que a medida de evidência proposta por Pereira-Stern em favor de H_0, sendo H_0 um subconjunto Θ_0 do espaço paramétrico Θ (com $dim(\Theta_0) < dim(\Theta)$, será igual à probabilidade complementar do conjunto T^*:

$$ev(\Theta_0; \boldsymbol{y}) = 1 - \kappa^* = 1 - P(\theta \in T^*(\boldsymbol{y})|\boldsymbol{y})$$

e o procedimento (ou teste) FBST consiste em aceitar H_0 sempre que $ev(\Theta_0; \boldsymbol{y})$ for grande.

Em poucas palavras, essa medida considera todos os pontos do espaço paramétrico cujos valores da densidade à posteriori são menores ou iguais ao seu supremo na região Θ_0. Se $ev(\Theta_0; \boldsymbol{y})$ for grande, então o subconjunto Θ_0 cai em uma região do espaço paramétrico de alta probabilidade à posteriori. Isso significa que T^* terá uma probabilidade à posteriori "pequena", o que nos leva a crer que os dados suportam a hipótese nula,H_0. Por outro lado, se um valor pequeno da evidência levaria à rejeição da hipótese nula (Pereira and Stern,1999).

2.2.3 Regra de Decisão e Validação no FBST

De acordo com Cella (2013), devemos encontrar uma maneira de determinar um valor de K (ponto crítico cujo valor depende da função perda escolhida) de modo que possamos determinar à partir de que ponto devemos rejeitar ou aceitar H_0, ou seja:

- Rejeitar H_0 se $ev(\Theta_0; \boldsymbol{y}) \leq K$
- Aceitar H_0 se $ev(\Theta_0; \boldsymbol{y}) > K$

De acordo com Madruga et al. (2001), podemos considerar $\mathbf{D}$ o espaço de decisões tal que $\mathbf{D}$=$(AceitarH_0(d_0), RejeitarH_0(d_1))$ e sendo L a função de $D \times \Theta \rightarrow \Re^+$ definida por:

$$L(RejeitarH_0, \theta) = a[1 - \mathbf{1}(\theta \in T^*(x))] \qquad \text{e}$$
$$L(AceitarH_0, \theta) = b + c\mathbf{1}(\theta \in T^*(x)) \qquad \text{, com} \qquad a, b, c > 0$$

Dados esses pressupostos, podemos assumir que o ponto de corte será $K = \frac{b+c}{a+c}$. (Madruga et al.,2001).

De uma forma geral, nos casos em que a evidência obtida é muito próxima de zero (ou de 1), a decisão natural é rejeitar (ou aceitar) H_0. Nas demais situações, pode-se estabelecer o nível de signicância do teste, como é feito nos testes clássicos, e buscar a validação do resultado obtido.

A validação dos resultados do FBST em estudos de simulação pode ser feita através de medidas empíricas, tais como o nível de significância e o poder empírico do teste, sendo possível a comparação entre resultados clássicos e o FBST. Madruga et al.(2001) apresentam as seguintes medidas:

- O nível de significância empírico de um teste de hipótese, com base em um grande número de repetições, é dado pela proporção de vezes em que a hipótese nula é rejeitada quando ela é verdadeira;

- O poder empírico de um teste de hipótese, com base em um grande número de repetições, é dado pela proporção de vezes em que a hipótese nula é rejeitada quando ela é falsa;

2.2.4 A Implementação do $Ev(\mathbf{\Theta_0}; \mathbf{x})$

Como visto em Cella (2013), o cálculo do valor de $Ev(\mathbf{\Theta_0}; \mathbf{x})$ envolve as duas etapas que são descritas à seguir:

- **1ª Etapa** - *Etapa de Otimização*: Consiste em maximizar a densidade à posteriori $\pi(\theta|\mathbf{x})$ sob H_0. Em outras palavras, consiste em obter o valor de θ^* que maximiza a densidade posterior, ou seja:

$$\pi(\theta^*|x) = sup_{\Theta_0}\pi(\theta|x)$$

- **2ª Etapa** - *Etapa de Integração*: Consiste em integrar a densidade à posteriori $\pi(\theta|\mathbf{x})$ sob o conjunto complementar a $T^*(x)$, ou seja:

$$I = \int_{T^{*C}(x)} \pi(\theta|\mathbf{x})\mathbf{d}\theta$$

com $T^{*C}(x) = \{\theta \in \Theta : \pi(\theta|\mathbf{x}) \leq \pi(\theta^*|\mathbf{x})\}$

Acontece que, na maioria dos casos, não é possivel obter essa integral da posteriori de maneira analítica. Tornam-se necessários, portanto, métodos numéricos para a obtenção dessa integral, como o *Método de MonteCarlo* (nos casos de baixa dimensão do espaço paramétrico) e os *Métodos MCMC* (nos casos de alta dimensão do espaço paramétrico).

2.3 Simulação Monte Carlo

Normalmente,queremos tirar conclusões sobre uma população de interesse com base na informação que obtemos de uma amostra. Para a maioria dos casos, admite-se que a distribuição estatística é conhecida, o que nos permite encontrar estimativas pontuais e intervalares para os parâmetros, testar hipóteses e estabelecer um modelo.

Há casos, porém, em que a distribuição dos parâmetros de interesse não é conhecida. Em outros casos, algumas suposições do modelo que queremos utilizar são violadas. Para esses casos, desenvolveu-se uma técnica de simulação conhecida como Monte Carlo. (Gamerman, 1997)

O processo consiste em obter diversas amostras sucessivas da mesma população e calcular as estimativas para os parâmetros de interesse em cada uma dessas amostras. Em alguns casos, isso significa simular repetidas vezes observações de uma mesma distribuição. Na prática, uma vez que tenhamos um número muito grande de amostras, a média de todas as estimativas obtidas serve como uma boa aproximação o parâmetro real de nosso interesse.

A técnica nos permite realizar inferências quando a distribuição da estatística do teste não é conhecida. Também se pode utilizar esse método para estimar o desempenho de métodos de inferência quando alguma suposição for violada, além de servir como uma forma de avaliar o poder do teste.

2.3.1 O MCMC

Para a maioria dos casos, onde as distribuições são conhecidas à priori, os métodos de Monte Carlo convencionais geram uma boa estimativa para os parâmetros de interesse. Porém, à medida que aumentamos a dimensão do problema e incluimos uma quantidade maior de parâmetros a serem estudados, os algoritimos convencionais passam a gerar estimativas cada vez piores, se afastando muito do parâmetro real.

Além disso, os métodos de Monte Carlo convencionais não são aplicaveis em distribuições que não tinham uma fórmula fechada ou que não obtinham uma solução analítica. Por esses motivos, faz-se necessária uma técnica mais geral, que atenda a maioria dos casos e não tenha muitas restrições quanto ao seu uso. Criou-se então, a técnica conhecida como MCMC (*Markov Chain Monte Carlo*). (Gamerman, 1997)

Assim como nos métodos convencionais, a idéia é obter um número grande de amostras e, a partir delas, encontrar estimativas para os parâmetros de interesse. Esse método, porém, se utiliza de um algorítimo de simulação mais complexo, onde cada iteração do algorítmo depende exclusivamente da iteração anterior, seguindo a mesma ideia das cadeias de Markov.

Isso permite que, dadas algumas condições bastante gerais, as cadeias geradas tenham um comportamento limite bem definido. Em outras palavras, se o número de iterações do algorítmo for suficientemente grande, a cadeia se aproxima muito da distribuição de interesse.

Para o nosso caso, o MCMC gera uma boa aproximação para a distribuição à posteriori dos parâmetros. Por esse motivo, esse será o algorítmo utilizado para as simulações que serão realizadas nesse trabalho.

3 O MODELO PARETO TIPO IV

Assuma que o tempo T até a ocorrência de um evento de interesse possa ser mensurado de maneira contínua. Assuma também que cada indivíduo da população estudada possui um risco λ de experimentar o evento de interesse ao longo do tempo. A variável aleatória T será modelada por uma distribuição Weibull (θ, λ). Assim, dado a chance λ, T segue uma distribuição Weibull com densidade: (Paulo et al., 2013)

$$f(t|\lambda) = \lambda\theta t^{\theta-1}e^{-\lambda t^{\theta}}; \tag{3}$$

e função de risco:

$$h(t|\lambda) = \lambda\theta t^{\theta-1} = \lambda\varphi(t;\theta);$$

Desse modo, cada indivíduo possui o seu prórpio risco de experimentar o evento de interesse. Essa risco, aqui denotado por λ, varia com o tempo de acordo com a função $\varphi(t;\theta)$. Considere também que esse risco individual, λ, é distribuído de acordo com a distribuição Gama(α, β), ou seja:

$$g(\lambda) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}; \tag{4}$$

De (3) e (4), temos que a função de densidade de Probabilidades da distribuição de Pareto do tipo IV é dada por:

$$f(t) = \int_0^{\infty} f(t,\lambda)d\lambda = \int_0^{\infty} g(\lambda)f(t \mid \lambda)d\lambda = \frac{\theta\alpha\beta^{\alpha}t^{\theta-1}}{(t^{\theta}+\beta)^{\alpha+1}}; \tag{5}$$

Para $t \geq 0$, α, β e θ podendo assumir qualquer valor positivo. Uma informação relevante é que (5) se reduz para a Pareto tipo II (ou pareto Lomax) quando θ=1 e para Pareto tipo III quando α=1. (Arnold, 1983)

A função de sobrevivência da distribuição Pareto tipo IV é dada por:

$$S(t) = \left(\frac{\beta}{t^{\theta}+\beta}\right)^{\alpha}, t \geq 0; \tag{6}$$

E a função de risco:

$$h(t) = \frac{\theta\alpha t^{\theta-1}}{t^{\theta}+\beta}, t \geq 0; \tag{7}$$

3.1 A verossimilhança da Pareto Tipo IV

Assumindo que a contribuição para a verossimilhança dos tempos de falha sejam dados pela função de densidade (5) e a contribuição dos tempos de censura sejam dados pela função de sobrevivência (6) temos que a função de verossimilhança é dada da seguinte forma: (Lawless,1982).

$$L\left(t,\delta \mid \theta,\alpha,\beta\right) \propto \prod_{i=1}^{n}[f(t)]^{\delta i}[S(t)]^{1-\delta i} \propto \prod_{i=1}^{n}[h(t)]^{\delta i}S(t) \propto \prod_{i=1}^{n}\left[\frac{(\theta\alpha t_i^{\theta-1})^{\delta i}\beta^{\alpha}}{(t_i^{\theta}+\beta)^{\alpha+\delta i}}\right]$$

Onde α, β e θ são os parâmetros a serem estimados com α>0, β>0, e θ>0; $t = (t_1, t_2, ..., t_n)$ é o vetor de valores observados, com seus respectivos indicadores de censuras dados por $\delta = (\delta_1, \delta_2, ..., \delta_n)$. δ_i=1 indica que o tempo t_i foi tempo de falha e δ_i=0 indica que o tempo t_i foi tempo de censura.

3.2 Obtenção da Posteriori

Assumindo que a população de interesse possa ser modelada por uma distribuição Pareto tipo IV e as condições impostas no capítulo 4 deste trabalho sejam satistfeitas, a função de verossimilhança da Pareto IV é dada por:

$$L\left(t,\delta \mid \alpha,\beta,\theta\right) = \prod_{i=1}^{n}\left[\frac{(\theta\alpha t_i^{\theta-1})^{\delta i}\beta^{\alpha}}{(t_i^{\theta}+\beta)^{\alpha+\delta i}}\right] = \frac{\theta^{\sum\delta_i}\alpha^{\sum\delta_i}\beta^{n\alpha}[\prod_{i=1}^{n}t_i^{\delta_i(\theta-1)}]}{\prod_{i=1}^{n}(t_i^{\theta}+\beta)^{\alpha+\delta_i}}$$

Considerando também que cada parâmetro pode ser modelado por uma distribuição Gama, cada uma com seus respectivos hiperparâmetros, ou seja:

Suas respectivas distribuições à priori seriam proporcionais à:

$$\pi(\alpha) \propto \alpha^{a_1-1}e^{-b_1\alpha} \quad ; \quad \pi(\beta) \propto \beta^{a_2-1}e^{-b_2\beta} \quad ; \quad \pi(\theta) \propto \theta^{a_3-1}e^{-b_3\theta}$$

Supondo independência a priori de α, β e θ, tem-se que :

$$\pi(\alpha,\beta,\theta) \propto \alpha^{a_1-1}e^{-b_1\alpha}\beta^{a_2-1}e^{-b_2\beta}\theta^{a_3-1}e^{-b_3\theta}$$

Assim,a distribuição a posteriori conjunta é proporcional à:

$$\pi(\alpha,\beta,\theta \mid t,\delta) \propto L\left(t,\delta \mid \alpha,\beta,\theta\right)\pi(\alpha,\beta,\theta)$$

$$= \frac{[\prod_{i=1}^{n} t_i^{\delta_i(\theta-1)}]}{\prod_{i=1}^{n}(t_i^{\theta}+\beta)^{\alpha+\delta_i}}\ \alpha^{a_1+\sum\delta_i-1}\ \beta^{a_2+n\alpha-1}\ \theta^{a_3+\sum\delta_i-1}\ e^{-b_1\alpha}\ e^{-b_2\beta}\ e^{-b_3\theta}$$

É possível manipular o termo $\prod_{i=1}^{n} t_i^{\delta_i(\theta-1)}$, de forma que simplifique a forma final da posteriori. Pela propriedade do logarítmo e da função exponencial, podemos assumir que:

$$\prod_{i=1}^{n} t_i^{\delta_i(\theta-1)} = e^{log(\prod_{i=1}^{n} t_i^{\delta_i(\theta-1)})} = e^{\sum log(t_i^{\delta_i(\theta-1)})} = e^{\sum\delta_i(\theta-1)log(t_i)}$$

$$= e^{\sum\delta_i\theta log(t_i)}e^{-\sum\delta_i log(t_i)} \tag{8}$$

Como o segundo termo de (8) não depende de nenhum parâmetro, pode entrar na constante de proporcionalidade. Assim, o numerador daquela fração será proporcional à:

$$e^{\theta\sum\delta_i log(t_i)}$$

Em outras palavras, a densidade à posteriori conjunta, $\pi\left(\alpha,\beta,\theta \mid t,\delta\right)$, será proporcional à:

$$\pi(\alpha,\beta,\theta \mid t,\delta) \propto \frac{1}{\prod_{i=1}^{n}(t_i^{\theta}+\beta)^{\alpha+\delta_i}}\alpha^{a_1+\sum\delta_i-1}\beta^{a_2+n\alpha-1}\theta^{a_3+\sum\delta_i-1}e^{-b_1\alpha}e^{-b_2\beta}e^{-\theta(b_3-\sum\delta_i log(t_i)} \tag{9}$$

Note que a espressão fechada da posteriori (9) não pode ser calculada analiticamente. No entanto, ela pode ser estimada empiricamente por meio dos métodos MCMC(*Markov Chain Monte Carlo*). Neste trabalho, os valores da posteriori (9) foram gerados por meio do comando "MCMCmetrop1R" da biblioteca MCMCPack do software R (R Core Team, 2014)

4 RESULTADOS

Inicialmente, a ideia é obter a distribuição à posteriori da Pareto tipo IV. De posse dessa informação, é possível obter as estimativas pontuais e intervalares para os parâmetros. Uma primeira impressão é que o tamanho n da amostra inicial gerada e o percentual de censura da amostra interfiram diretamente na precisão das estimativas encontradas.

Partindo dessa ideia, deseja-se testar se existe ou não alguma distribuição mais simples do que a Pareto IV capaz de representar os dados de maneira fidedigna, sem perda de informação.

Os resultados obtidos por este trabalho serão apresentados à seguir.

4.1 Simulações

Para fins ilustrativos, obteve-se uma amostra de dados de sobrevivência a partir de simulações, utilizando-se o software R.

As amostras da distribuição Pareto tipo IV foram geradas a partir dos parâmetros iniciais α_0 =0.5, β_0=2 e θ_0 =3 e por meio do método de transformação inversa da função de distribuição acumulada. O mecanismo de censura utilizado foi aleatório, isto é, independente do tempo de falha, de acordo com os respectivos tamanhos de amostra e dos níveis de censura selecionados no estudo.

A partir dos tamanhos de amostra n=50, 100, 200 e 500 com níveis de censura de 0%, 10%, 20% e 30%, obteve-se os vetores de tempos de sobrevivência, com seus respectivos indicadores de censura.

Assumindo que a função densidade e a função de sobrevivência dos valores gerados respeitem as condições impostas pelo capítulo 4 deste trabalho, as estimativas pontuais de máxima verossimilhança, as estimativas Bayesianas (média a posteriori), com seus respectivos intervalos HPD e os e-valores para os três testes de interesse nesse trabalho serão apresentados na tabela (1). Todas as estimativas foram realizadas adotando-se prioris não informativas (difusas) para os parâmetros α, β e θ ($a_1 = b_1 = a_2 = b_2 = a_3 = b_3$=0.001).

Tabela 1a – Inferência Bayesiana para os parâmetros da distribuição Pareto IV com dados simulados utilizando diferentes tamanhos de amostra e percentuais de censura

n	Censura	Parâmetro	EMV	Média à posteriori	HPD (95%)		*e*-valor		
							H_0: $\theta = 1$	H_0: $\alpha = 1$	H_0: θ=1 e α=1
50	0%	α	0.3389	0.3497	0.0855	0.6967	-	0.0563	-
		β	1.4981	1.5952	0.0008	3.7103	-	-	<0.001
		θ	3.7444	4.1535	1.9408	6.8313	<0.001	-	-
	10%	α	0.3188	0.3303	0.089	0.6593	-	0.0462	-
		β	1.3739	1.4785	0.0042	3.5347	-	-	<0.001
		θ	3.8185	4.2168	1.8691	6.8728	<0.001	-	-
	20%	α	0.3342	0.3541	0.085	0.7386	-	0.0883	-
		β	2.3236	2.5268	0.0332	5.7694	-	-	<0.001
		θ	3.6652	4.0653	1.795	6.7952	<0.001	-	-
	30%	α	0.2942	0.3064	0.0424	0.6597	-	0.0741	-
		β	2.1818	2.3654	0.0013	5.8243	-	-	<0.001
		θ	3.4428	3.9913	1.5499	6.9672	<0.001	-	-
100	0%	α	0.5102	0.5302	0.2325	0.8939	-	0.1287	-
		β	2.1586	2.2803	0.6179	4.2899	-	-	<0.001
		θ	3.091	3.1933	2.0709	4.4178	<0.001	-	-
	10%	α	0.4527	0.465	0.2014	0.7936	-	0.0728	-
		β	2.1204	2.218	0.5542	4.239	-	-	<0.001
		θ	3.0821	3.202	2.0798	4.4839	<0.001	-	-
	20%	α	0.4036	0.4176	0.1728	0.7251	-	0.045	-
		β	2.2982	2.4271	0.5977	4.6664	-	-	<0.001
		θ	3.2235	3.3553	2.0929	4.738	<0.001	-	-
	30%	α	0.2381	0.2297	0.0669	0.4319	-	0.0012	-
		β	1.6812	1.6323	0.0115	3.5333	-	-	<0.001
		θ	4.5008	5.1634	2.5627	8.4814	<0.001	-	-

Dados simulados de uma Pareto IV com parâmetros α = 0,5; β = 2 e θ= 3;

Hiperparâmetros da priori: $a_1 = b_1 = a_2 = b_2 = a_3 = b_3$=0.001.

Estimativas obtidas por MCMC de tamanho 500.000,considerando um burn-in de 100.000.

Tabela 1b – Inferência Bayesiana para os parâmetros da distribuição Pareto IV com dados simulados utilizando diferentes tamanhos de amostra e percentuais de censura (Continuação)

n	Censura	Parâmetro	EMV	Média à posteriori	HPD (95%)		*e*-valor		
							H_0: $\theta = 1$	H_0: $\alpha = 1$	H_0: θ=1 e α=1
200	0%	α	0.4138	0.4172	0.2544	0.5959	-	0.001	-
		β	2.1071	2.1406	1.0258	3.3767	-	-	<0.001
		θ	3.7061	3.7755	2.8313	4.7978	<0.001	-	-
	10%	α	0.4043	0.4094	0.2397	0.6054	-	0.0016	-
		β	2.2186	2.2652	1.0068	3.6478	-	-	<0.001
		θ	3.4633	3.5317	2.5807	4.5293	<0.001	-	-
	20%	α	0.3948	0.4014	0.225	0.6072	-	0.0029	-
		β	2.5424	2.6074	1.1786	4.2739	-	-	<0.001
		θ	3.3611	3.4298	2.4763	4.4355	<0.001	-	-
	30%	α	0.3873	0.3977	0.209	0.6194	-	0.0065	-
		β	2.9176	3.0243	1.2733	5.0641	-	-	<0.001
		θ	3.1281	3.19	2.2732	4.1797	<0.001	-	-
500	0%	α	0.3965	0.3974	0.2992	0.5035	-	<0.001	-
		β	1.5883	1.5975	1.0184	2.2088	-	-	<0.001
		θ	3.5427	3.5702	2.9987	4.1764	<0.001	-	-
	10%	α	0.3532	0.354	0.262	0.4504	-	<0.001	-
		β	1.458	1.4666	0.9051	2.0616	-	-	<0.001
		θ	3.5406	3.5703	2.9679	4.1993	<0.001	-	-
	20%	α	0.3138	0.315	0.2273	0.4107	-	<0.001	-
		β	1.5578	1.5718	0.9289	2.2605	-	-	<0.001
		θ	3.5249	3.5588	2.9074	4.2406	<0.001	-	-
	30%	α	0.289	0.2902	0.2064	0.3835	-	<0.001	-
		β	1.6495	1.6642	0.9596	2.4128	-	-	<0.001
		θ	3.4183	3.4531	2.8135	4.1491	<0.001	-	-

Dados simulados de uma Pareto IV com parâmetros α = 0,5; β = 2 e θ= 3;
Hiperparâmetros da priori: $a_1 = b_1 = a_2 = b_2 = a_3 = b_3$=0.001. Estimativas obtidas por MCMC de tamanho 500.000,considerando um burn-in de 100.000.

As tabelas (1a) e (1b) nos mostram alguns resultados interessantes. O primeiro, e mais simples de observar, é que se o tamanho de amostra for fixado e variarmos o percentual de censura, nota-se que as melhores estimativas são aquelas onde o percentual de censura é menor. Isso acontece porque à medida que aumentamos o percentual de censura, estamos perdendo parte da informação. Isso se reflete numa piora na estimativa dos nossos parâmetros, tanto nas estimativas pontuais, quanto no aumento da amplitude do intervalo HPD.

De forma similar, se fixarmos um percentual de censura qualquer e observarmos as estimativas pontuais para os parâmetros nos diferentes tamanhos de amostra observa-se que as estimativas não estão próximas do valor teórico fornecido inicialmente. Esse fenômeno não é causado pela imprecisão das estimativas e sim pela variabilidade ao gerar os dados aleatórios. Note que as estimativas bayesianas são próximas dos EMV's, dado evidências que o desvio foi devido aos dados simulados.

Por esse motivo, devemos observar o intervalo HPD. Por mais que as estimativas pontuais não sigam o comportamento esperado (de que se aumentarmos o tamanho de amostra com o mesmo percentual de censura, a estimativa deveria melhorar), o intervalo HPD diminui de amplitude conforme aumentamos o tamanho da amostra. Em outras palavras, esse intervalo HPD indica que a estimativa encontrada estará mais próxima do valor real sempre que o tamanho da amostra for maior. Isso é um indicativo que o tamanho da amostra melhora a precisão das nossas estimativas, independente do percentual de censura escolhido.

Devemos, porém, avaliar os *e*-valores encontrados para os testes. A ideia é verificar se modelos mais simples do que o modelo Pareto IV se ajustam bem ao mesmo conjunto de dados (nesse caso, dados simulados). Se for esse o caso, então os modelos com menos parâmetros serão utilizados.

O que a tabela mostra nesse sentido é que rejeitamos as hipóteses $H_0 : \theta = 1$ e $H_0 : \theta$=1 e α=1, uma vez que todos os *e*-valores foram pequenos . Isso significa que, independente do tamanho de amostra utilizado, o modelo Pareto do tipo II parece não se ajustar para esse conjunto de dados.

Para a hipótese $H_0 : \alpha = 1$, encontramos resultados diferentes. Com o tamanho de amostra n=50, não há evidência para a rejeição da hipótese nula a um nível de significancia de 5%, considerando os niveis de censura iguais a 0%, 20% e 30%.

Mesmo que para o nível de censura igual a 10% a hipótese nula tenha sido rejeitada, o que se observa é que o modelo com menos parâmetros parece ser mais adequado do que o modelo Pareto IV. A diferença entre os dois não é tão significativa, o que nos permite trabalhar com menos parâmetros e ter uma boa representatividade dos dados.

Para o tamanho de amostra n=100, continuamos não rejeitando ($H_0 : \alpha = 1$) nos dois primeiros niveis de censura (0% e 10%). Já para os outros dois níveis (20% e 30%), o modelo mais simples parece não ser o suficiente para representar o nosso conjunto de dados, uma vez que a hipótese nula é rejeitada.

Daí em diante, para os tamanhos de amostra n=200 e n=500, todos os e-valores nos levam à rejeição a hipótese nula. Em outras palavras, para qualquer nível de censura, o modelo Pareto IV representa melhor os dados e deve, portanto, ser utilizado.

O que obtemos então é uma indicação de que, para tamanhos de amostra pequenos, o teste não tem poder suficiente de indicar uma diferença significativa entre o modelo Pareto IV e o modelo com menos parâmetros (α=1). Nesses casos, podemos optar por modelos mais simples. Porém, quando o tamanho de amostra é muito grande, os testes ganham poder e podem comprovar que os modelos mais simples não representam de maneira fidedigna a complexidade do comportamento do nosso banco de dados, o que nos leva a optar pelo modelo mais complexo.

4.2 Influência na escolha da priori

O objetivo desta seção é estudar o quanto a escolha dos hiperparâmetros das prioris de α e θ influenciam na inferência dos parâmetros do modelo Pareto IV. Para isso, o valor dos hiperparâmetros da priori de β (a_2 e b_2) foram mantidos constantes, iguais a 0.001. Seria equivalente dizer que a priori atribuida à Beta foi não-informativa (difusa).

Os valores iniciais para a simulação foram mantidos (α_0 =0.5, β_0=2 e θ_0 =3). Foi gerada uma amostra de tamanho n=100 e percentual de censura igual a 10%. A partir dela, variou-se os valores de a_1 e b_1 , sendo esses os valores dos hiperparâmetros para a priori de α e variou-se também os valores de a_3 e b_3, sendo esses os hiperparâmetros para a priori de θ. Os resultados dessas simulações encontram-se na tabela (2):

Tabela 2a – Influência na escolha da priori de α e θ na inferência dos parâmetros do modelo Pareto IV

a_1	b_1	a_3	b_3	Parâmetro	Média à posteriori	HPD (95%)		e-valor H_0: $\theta = 1$	e-valor H_0: $\alpha = 1$	e-valor H_0: θ=1 e α=1
0.001	0.001	0.001	0.001	α	0.465	0.2014	0.7936	-	0.0728	-
				β	2.218	0.5542	4.239	-	-	<0.001
				θ	3.202	2.0798	4.4839	<0.001	-	-
0.01	0.01	0.01	0.01	α	0.4706	0.1938	0.804	-	0.0783	-
				β	2.2431	0.5503	4.3163	-	-	<0.001
				θ	3.1967	2.0192	4.4823	<0.001	-	-
0.1	0.1	0.1	0.1	α	0.4739	0.1981	0.7957	-	0.0799	-
				β	2.2675	0.5979	4.3662	-	-	<0.001
				θ	3.1665	2.0652	4.4209	<0.001	-	-
1	1	1	1	α	0.5447	0.2505	0.9205	-	0.1511	-
				β	2.5718	0.7117	4.7251	-	-	<0.001
				θ	2.8974	1.9477	3.9583	<0.001	-	-
1	1	0.001	0.001	α	0.493	0.2116	0.8346	-	0.0953	-
				β	2.3584	0.6068	4.4362	-	-	<0.001
				θ	3.1071	2.0365	4.3333	<0.001	-	-
0.001	0.001	1	1	α	0.5173	0.2348	0.8696	-	0.1214	-
				β	2.4459	0.7033	4.5813	-	-	<0.001
				θ	2.9706	1.988	4.0568	<0.001	-	-

Dados simulados de uma Pareto IV com parâmetros α = 0,5; β = 2; θ= 3; n=100 e Percentual de censura=10%. Estimativas obtidas por MCMC de tamanho 500.000,considerando um burn-in de 100.000.

Tabela 2b – Influência na escolha da priori de α e θ na inferência dos parâmetros do modelo Pareto IV(Continuação)

a_1	b_1	a_3	b_3	Parâmetro	Média à posteriori	HPD (95%)		e-valor H_0: $\theta = 1$	e-valor H_0: $\alpha = 1$	e-valor H_0: θ=1 e α=1
				α	0.6456	0.3461	0.9994	-	0.2428	-
1	5	1	5	β	2.7895	1.0983	4.8198	-	-	<0.001
				θ	2.423	1.7798	3.1149	<0.001	-	-
				α	0.4075	0.1969	0.6588	-	0.0122	-
1	5	0.001	0.001	β	1.9292	0.5038	3.6277	-	-	<0.001
				θ	3.3981	2.2624	4.6676	<0.001	-	-
				α	0.8347	0.3756	1.4349	-	0.7098	-
0.001	0.001	1	5	β	3.5785	1.1727	6.4793	-	-	<0.001
				θ	2.2157	1.5605	2.911	<0.001	-	-
				α	0.7161	0.2824	1.2863	-	0.4935	-
5	1	5	1	β	3.5122	0.9914	6.7162	-	-	<0.001
				θ	2.6681	1.7386	3.6692	<0.001	-	-
				α	0.8376	0.3028	1.5622	-	0.7254	-
5	1	0.001	0.001	β	3.962	1.097	7.5858	-	-	<0.001
				θ	2.4591	1.5909	3.4272	<0.001	-	-
				α	0.4218	0.1899	0.7015	-	0.0301	-
0.001	0.001	5	1	β	2.0313	0.4663	3.8804	-	-	<0.001
				θ	3.3983	2.2185	4.6668	<0.001	-	-

Dados simulados de uma Pareto IV com parâmetros α = 0,5; β = 2; θ= 3; n=100 e Percentual de censura=10%. Estimativas obtidas por MCMC de tamanho 500.000,considerando um burn-in de 100.000.

Se observarmos esses resultados, a primeira seleção de hiperparâmetros gera os mesmos resultados encontrados na tabela (1a), onde todos os hiperparâmetros foram iguais a 0.001, com n=100 e percentual de censura igual a 10%. Nesse caso, chegamos à conclusão de que havia evidência para a rejeição da primeira hipótese ($H_0 : \theta = 1$) e da terceira ($H_0 : \theta$=1 e α=1). Porém, não encontramos evidência suficiente para a rejeição de $H_0 : \alpha = 1$.

A escolha das prioris de α e θ parecem estar interferindo nas estimativas pontuais bayesianas para os parâmetros, na amplitude dos seus respectivos intervalos HPD e nos e-valores de somente em uma das hipóteses testadas ($H_0 : \alpha = 1$). Os estimadores de máxima verossimilhança não dependem da priori selecionada, o que gera a estimativa de α=0.4527, β=2.1204 e θ=3.0821 em todas as simulações.

Em contrapartida, os e-valores para $H_0 : \theta = 1$ e $H_0 : \theta$=1 e α=1 parecem não se alterar com a escolha da priori, uma vez que todas as combinações de hiperparâmetros geraram e-valores muito proximos de 0 para ambos os testes, o que é uma evidência forte de que essas hipóteses devam ser rejeitadas.

Ao observarmos os e-valores da hipótese $H_0 : \alpha = 1$ (tabela 2a), nota-se que à medida que se fornece mais informação à priori parecemos ter uma evidência maior de que essa hipótese não seja rejeitada.

Um outro exemplo da tabela (2b) nos mostra que, para a seleção a_1=1, b_1=5, a_3= 0.001 e b_3=0.001 a hipótese $H_0 : \alpha = 1$ é rejeitada, considerando um ponto de corte de 5%. Significa que, para essa seleção de hiperparâmetros, o modelo Pareto IV parece se ajustar melhor aos dados do que o modelo Pareto III ($\alpha = 1$). Esse resultado é coerente pois essa combinação de hiperparâmetros supõe, a priori, que a média de α é 0.2 , resultando em um menor valor da estimativa (mais distante de 1).

Isso é um reflexo da importância da informação à priori para as inferências realizadas nesse trabalho. Como o exemplo anterior mostra, dependendo da priori escolhida e de seus hiperparâmetros, pode-se encontrar resultados diferentes.

5 APLICAÇÃO EM DADOS REAIS

Este capítulo apresenta uma aplicação da metodologia inferencial do modelo Pareto tipo IV em um conjunto de dados reais. Os dados fazem parte do trabalho de Lagakos (1978), que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão. Os dados estão anexados no apêndice, apresentado ao final do trabalho.

Propôs-se ajustar um modelo Pareto tipo IV para esse conjunto de dados, de modo que seja possível encontrar as estimativas pontuais e intervalares para os parâmetros. Em seguida, seriam apresentados os e-valores para os três testes de hipóteses de interesse, para que pudessemos analisar se algum modelo pareto mais simples do que o modelo Pareto IV poderia ser utilizado. Também foi proposto um gráfico com as funções de sobrevivência estimadas, a fim de se apresentar uma comparação visual entre os possíveis modelos estabelecidos para esse conjunto de dados.

Os resultados são apresentados na figura (1) e na tabela (3), apresentadas à seguir.

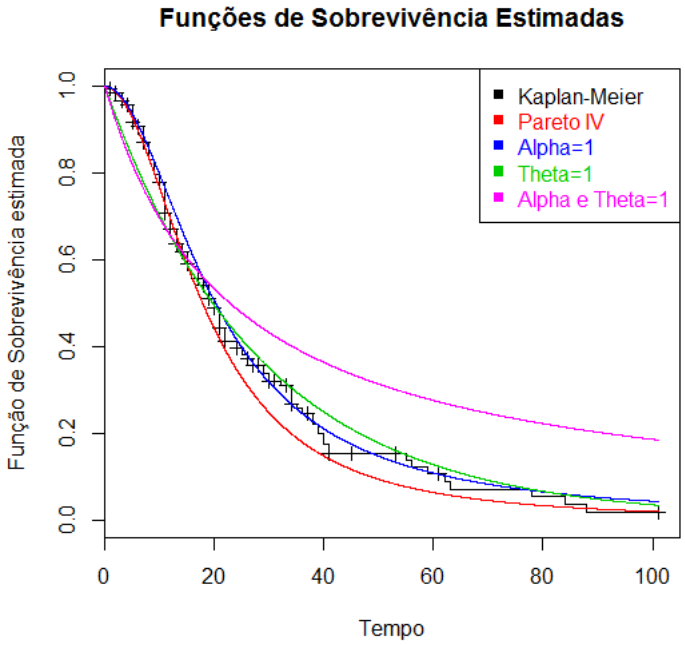

Figura 1 – Funções de Sobrevivência Estimadas para os dados referentes ao tempo de sobrevivência de pacientes de com câncer de pulmão

Tabela 3 – Inferência Bayesiana para os parâmetros da distribuição Pareto IV aplicada aos dados referentes ao tempo de sobrevivência de pacientes com câncer de pulmão

Parâmetro	Média à posteriori	HPD (95%)		e-valor		
				H_0: $\theta = 1$ (Pareto II)	H_0: $\alpha = 1$ (Pareto III)	H_0: θ=1 e α=1
α	1.2909	0.3247	2.9767	-	0.9962	-
β	377.8626	116.7	727.1538	-	-	<0.001
θ	1.9419	1.4011	2.515	<0.001	-	-

Dados referentes ao tempo de sobrevivência de n=194 pacientes de com câncer de pulmão; Estimativas obtidas por MCMC de tamanho 500.000,considerando um burn-in de 100.000.

A tabela (3) apresenta as estimativas pontuais e intervalares para os parâmetros, utilizando as técnicas propostas por esse trabalho. Analisando as informações que ela apresenta, verificamos que existe uma forte evidência de que o modelo Pareto tipo III não deve ser rejeitado. Em outras palavras, assumir que o parâmetro α é igual a 1 parece razoavel nesse conjunto de dados, o que nos permitiria trabalhar com um parâmetro à menos na nossa modelagem.

Essa informação é coerente com o que observamos na figura (1). Ela mostra algumas funções de sobrevivência estimadas para o conjunto de pacientes com cancer de pulmão. A primeira delas, em preto, seria a função de sobrevivência estimada pelo método de Kaplan-Meier. Para as outras funções estimadas, admitiu-se que todas as condições impostas no capítulo 3 deste trabalho foram satisfeitas e que esse conjunto de dados poderia ser modelado de acordo com uma distribuição Pareto IV, com os respectivos limitadores em seus parâmetros.

A função em vermelho, por exemplo, não possui restrição nos parâmetros. Seria então a função de sobrevivência estimada a partir do modelo mais geral, onde α, β e θ estão variando. A função estimada em azul admite que o parâmetro α é igual a 1, e varia os outros

dois parâmetros. A função em verde faz o mesmo, mas fixando o parâmetro θ ao invés de α. A função em roxo admite que tanto α como θ sejam iguais a 1, e varia somente o parâmetro β.

Podemos usar como referência a função estimada por Kaplan Meier, uma vez que o algoritimo utilizado para obtê-la é não paramétrico e leva em consideração a informação sobre as censuras no nosso banco de dados. Por ser baseada em um estimador de máxima verossimilhança, essa função é uma boa aproximação do que esperamos que seja a função de sobrevivência para o nosso conjunto de dados.

Se compararmos as funções, veremos que a distribuição Pareto III (que admite α=1) também gera estimativas mais próxima da função estimada por Kaplan-Meier. O modelo Pareto que gera a melhor estimativa para a função de sobrevivência nesse conjunto de dados é o modelo Pareto IV. Porém, como o ajuste do modelo Pareto III não pode ser rejeitado, então pelo princípio da parcimônia, o modelo Pareto III, que é mais simples que o modelo Pareto IV (um parâmetro a menos) pode ser considerado.

6 CONCLUSÃO

O modelo Pareto IV se mostrou acessível, podendo ser utilizado de forma prática para a modelagem de conjuntos de dados presentes na literatura. Mesmo que a posteriori encontrada não se assemelhe nenhuma distribuição conhecida, a mesma pôde ser facilmente obtida utilizando-se dos algoritimos de MCMC implantados pelo software free, R. O FBST provou-se uma boa ferramenta para se testar os parâmetros do modelo Pareto IV, fornecendo uma forma eficiente de se comparar modelos mais simples, nos permitindo optar pelo modelo que melhor represente o nosso conjunto de dados.

A metodologia apresentada neste trabalho se mostrou eficaz no ajuste e seleção de modelo para representar os dados sobre o tempo até a morte de pacientes com câncer de pulmão.

REFERÊNCIAS

BARRY C.A.(1983). **Pareto Distributions**. International Co-operative Publishing House.

BRUNELLO, G. H. V; NAKANO, E. Y. (2015) **Inferência bayesiana no modelo weibull discreto em dado com presença de censura.** TEMA: Tendências em Matemática Aplicada e Computacional, V.16, (no prelo).

CARRASCO, C. G; TUTIA, M. H.; NAKANO, E.Y. (2012) **Intervalos de confiança para os parâmetros do modelo geométrico com inflação de zeros.** TEMA: Tendências em Matemática Aplicada e Computacional, v.7, n.3, p. 247-255.

CELLA,L.O.G. (2013). **Regressão Ordinal Bayesiana**. Dissertação de Mestrado. - Universidade de Brasília - UnB

COLOSIMO,E.A. (2006). **Analise de Sobrevivência Aplicada**. Enrico Antonio Colosimo, Suely Ruiz Giolo. - São Paulo:Edgard Blücher.

COWELL, R.G.; DAWID, A.P. ; LAURITZEN, S.L.; SPIEGELHALTER, D.J. (1999). **Probabilistic Networks and Expert Systems**.Springer-Verlag New York, Inc.

DOURADO, P.H. ; CARVALHO, T.M. ; OTINIANO, C.E.G. ; NAKANO,E.Y. (2013) **Propriedades estatísticas e matemáticas da distribuição Pareto tipo IV e sua aplicação em dados censurados**.Departamento de Estatística, Universidade de Brasília - UnB.

FELLER, W. (1971).**An Introduction to Probability Theory and its Applications II** (2nd ed.). New York: Wiley. p. 50.

GAMERMAN, D. (1997).**Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.** Chapman & Hall . New York , ISBN: 0412818205.

HAMMERSLEY, J.M.; HANDSCOMB, D.C. (1964).**Monte Carlo methods**. London: Methuen.

HOFF, P.D.(2009). **A First Course in Bayesian Statistical Methods**. Springer Science+Business Media, LLC.

JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. (1994) **Continuous univariate distributions** Vol 1. Wiley Series in Probability and Statistics

LAGAKOS, S.W; (1978). **A covariate model for partially censored data subject to competing causes of failure**. Appl.Statist, 27, n.3 , p.235-241.

LAWLESS, J.F. (1982). **Statistical Models and Methods for Lifetime Data**, John Wiley and Sons, New York.

LEE, E.T. **Statistical Methods for Survival Data Analysis**. Lifetime Learning Publications, New York, 1992

NAKANO, E. Y.; CARRASCO, C. G. (2006) **Uma avaliação do uso de um modelo contínuo na análise de dados discretos de sobrevivência.** TEMA: Tendências em Matemática Aplicada e Computacional, v.7, n.1, p.91-100.

PEREIRA, C.A.B.; STERN, JM. (1999). Evidence and credibility: full bayesian significance test of precise hypothesis. **Entropy**, 1:99-110.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org/.

A APÊNDICE

A.1 Conjunto de dados

Tabela 4a – Conjunto de dados referentes ao trabalho de Lagakos (1978), que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão.

Paciente	Tempo	Censura	Paciente	Tempo	Censura	Paciente	Tempo	Censura
1	11	1	36	19	1	71	13	0
2	11	1	37	12	0	72	11	1
3	5	1	38	4	1	73	1	0
4	22	0	39	17	1	74	8	1
5	11	1	40	21	1	75	19	1
6	8	1	41	5	1	76	2	0
7	1	0	42	9	1	77	3	0
8	21	1	43	1	1	78	26	1
9	13	0	44	2	0	79	11	1
10	4	0	45	7	1	80	4	0
11	4	0	46	21	1	81	5	1
12	5	0	47	28	0	82	18	0
13	17	0	48	14	0	83	22	1
14	30	1	49	34	1	84	6	0
15	1	0	50	13	1	85	5	1
16	21	1	51	1	0	86	24	1
17	8	1	52	12	0	87	7	1
18	14	0	53	88	1	88	17	1
19	8	1	54	4	0	89	37	0
20	1	0	55	2	0	90	59	1
21	84	1	56	21	0	91	13	1
22	16	1	57	10	1	92	41	1
23	29	1	58	11	1	93	17	1
24	22	1	59	8	1	94	40	1
25	15	0	60	8	1	95	56	1
26	9	1	61	12	1	96	27	0
27	12	1	62	2	1	97	19	0
28	10	0	63	3	1	98	20	1
29	11	1	64	4	0	99	14	0
30	30	0	65	7	0	100	5	1
31	18	0	66	2	1	101	11	1
32	13	1	67	12	1	102	11	1
33	5	1	68	3	1	103	9	1
34	6	1	69	11	0	104	14	0
35	20	1	70	55	1	105	24	1

Tabela 4b – Conjunto de dados referentes ao trabalho de Lagakos (1978), que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão.(Continuação)

Paciente	Tempo	Censura	Paciente	Tempo	Censura	Paciente	Tempo	Censura
106	33	0	141	14	1	176	17	1
107	28	0	142	11	1	177	63	1
108	22	1	143	29	1	178	6	0
109	34	1	144	8	1	179	13	1
110	22	1	145	7	1	180	7	1
111	39	1	146	15	0	181	53	0
112	3	1	147	27	1	182	15	1
113	61	0	148	7	1	183	38	1
114	10	1	149	19	1	184	15	1
115	24	0	150	27	1	185	10	1
116	10	0	151	7	1	186	11	0
117	12	1	152	4	1	187	20	0
118	26	0	153	21	1	188	10	0
119	12	1	154	9	1	189	11	0
120	31	0	155	40	1	190	7	0
121	4	0	156	6	0	191	15	1
122	15	1	157	27	0	192	32	1
123	34	1	158	34	1	193	34	0
124	39	1	159	78	1	194	18	1
125	101	0	160	11	1			
126	11	1	161	41	0			
127	36	1	162	6	0			
128	25	1	163	14	1			
129	34	0	164	21	1			
130	12	0	165	45	0			
131	62	1	166	9	1			
132	9	1	167	5	1			
133	41	1	168	20	1			
134	38	1	169	25	1			
135	18	0	170	12	0			
136	30	1	171	18	1			
137	6	1	172	35	1			
138	14	0	173	31	0			
139	19	1	174	14	1			
140	12	1	175	13	1			

A.2 Programação no R

```
set.seed(123)

ParetoIV<-function(alpha0,beta0,theta0,n,p.cens,a1,b1,a2,b2,a3,b3){
  estimativas<-c(1:15)



############################
##### GERANDO OS DADOS #####
############################
#alpha0<-alpha0
#beta0<-beta0
#theta0<-theta0

#n<-n  ### tamanho da amostra

set.seed(123)
u<-runif(n,0,1)
tempo<-(beta0/((1-u)^(1/alpha0)) - beta0)^(1/theta0)

p.cens<-p.cens   ### percentual de censura
delta<-rbinom(n,1,1-p.cens)



##############################
### log DENSIDADE E SOBREV ###
##############################
log.f<-function(alpha,beta,theta,tempo){
  ( log(theta) + log(alpha) + alpha*log(beta) + (theta-1)*log(tempo) -
      (alpha+1)*log(tempo^theta + beta) )
}

log.s<-function(alpha,beta,theta,tempo){
  ( alpha*log(beta) - alpha*log(tempo^theta + beta) )
}
```

```
#######################################
#### FUNÇÃO DE VEROSSIMILHANÇA (-) ###
#######################################

log.vero<-function(parametros,tempo,delta){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-parametros[3]

  if ( (alpha>0) && (beta>0) && (theta>0) )
    return (
      -1*sum( delta*log.f(alpha,beta,theta,tempo) +
                (1-delta)*log.s(alpha,beta,theta,tempo) )
    )
  else return (-Inf)
}

############################
### OBTENÇÃO DOS EMVs  ###
############################
emv<-nlm(log.vero,c(1,1,1),tempo=tempo,delta=delta)
est.EMV<-emv$estimate



########################
#### log-POSTERIORI ####
########################
#a1<-a2<-a3<-b1<-b2<-b3<-0.001   ## hiperparâmetros da priori

log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-parametros[3]

  if ( (alpha>0) && (beta>0) && (theta>0) )
```

```
    return (
      sum( delta*log.f(alpha,beta,theta,tempo) +
             (1-delta)*log.s(alpha,beta,theta,tempo) )
      + (a1-1)*log(alpha) - b1*alpha
      + (a2-1)*log(beta)  - b2*beta
      + (a3-1)*log(theta) - b3*theta
    )
  else return (-Inf)
}


#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=c(1,1,1),mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
                      delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

##################################
### obtenção das estimativas ###
##################################
mean(amost[,1])
mean(amost[,2])
mean(amost[,3])
est.bayes<-c(mean(amost[,1]),mean(amost[,2]),mean(amost[,3]))









################################
#### VERIFICAÇÃO DO AJUSTE ###
################################
```

```
require(survival)
km<-survfit(Surv(tempo,delta)~1)
plot(km,conf.int=F,xlab="Tempo",ylab="Função de Sobrevivência estimada")

tt<-seq(0,max(tempo),0.01)

sobrev.EMV<-exp(log.s(est.EMV[1],est.EMV[2],est.EMV[3],tt))
points(tt,sobrev.EMV,type="l",col=2)

sobrev.bayes<-exp(log.s(est.bayes[1],est.bayes[2],est.bayes[3],tt))
points(tt,sobrev.bayes,type="l",col=4)

legend(30,1,c("Kaplan-Meier","EMV","Bayes"),text.col=c(1,2,4),bty="n")



##############################################################################
#### INTERVALO HPD PARA ALPHA, BETA E THETA - (APROXIMADO, CONFIANÇA=95%)###
##############################################################################
require(TeachingDemos)
hpdalpha<-emp.hpd(amost[,1], conf=0.95)

hpdbeta<-emp.hpd(amost[,2], conf=0.95)

hpdtheta<-emp.hpd(amost[,3], conf=0.95)



##########################################################
#### CALCULANDO A LOG POSTERIORI DOS VALORES SIMULADOS##
##########################################################

m<-length(amost[,1])
valores.log.post<-numeric(m)
for (i in 1:m){
  valores.log.post[i]<-log.post(amost[i,],tempo,delta,a1,a2,a3,b1,b2,b3)
}
```

```
##########################################################
#### CALCULANDO  E VALOR PARA H1:THETA=1##
##########################################################

log.post.H1<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-1

  if ( (alpha>0) && (beta>0))
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                 (1-delta)*log.s(alpha,beta,theta,tempo) )
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}

H1<-nlm(log.post.H1,c(1,1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H1<- -1*H1$min


#e-valor H1:THETA=1
e1<-sum(valores.log.post<max.post.H1)/m


##########################################################
#### CALCULANDO  E VALOR PARA H2:ALPHA=1##
##########################################################

log.post.H2<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
```

```
  alpha<-1
  beta<-parametros[1]
  theta<-parametros[2]

  if ( (beta>0) && (theta>0))
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                  (1-delta)*log.s(alpha,beta,theta,tempo) )
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}

H2<-nlm(log.post.H2,c(1,1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H2<- -1*H2$min



#e-valor H2:ALPHA=1
e2<-sum(valores.log.post<max.post.H2)/m

#########################################################
#### CALCULANDO  E VALOR PARA H2:ALPHA=1 E THETA=1 ##
#########################################################

log.post.H3<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-1
  beta<-parametros[1]
  theta<-1

  if ( (beta>0) )
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                  (1-delta)*log.s(alpha,beta,theta,tempo) )
```

```
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}

H3<-nlm(log.post.H3,c(1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H3<- -1*H3$min



#e-valor H3:ALPHA=1 E THETA=1
e3<-sum(valores.log.post<max.post.H3)/m


estimativas[1:3]<-(est.EMV)
estimativas[4:6]<-(est.bayes)
estimativas[7:8]<-(hpdalpha)
estimativas[9:10]<-(hpdbeta)
estimativas[11:12]<-(hpdtheta)
estimativas[13]<-(e1)
estimativas[14]<-(e2)
estimativas[15]<-(e3)

return(estimativas)
estimativas

}


##########################################################
#### MONTANDO AS TABELAS ##
##########################################################

aa<-ParetoIV(0.5,2,3,50,0,0.001,0.001,0.001,0.001,0.001,0.001)
```

```
ab<-ParetoIV(0.5,2,3,50,0.1,0.001,0.001,0.001,0.001,0.001,0.001)
ac<-ParetoIV(0.5,2,3,50,0.2,0.001,0.001,0.001,0.001,0.001,0.001)
ad<-ParetoIV(0.5,2,3,50,0.3,0.001,0.001,0.001,0.001,0.001,0.001)
ba<-ParetoIV(0.5,2,3,100,0,0.001,0.001,0.001,0.001,0.001,0.001)
bb<-ParetoIV(0.5,2,3,100,0.1,0.001,0.001,0.001,0.001,0.001,0.001)
bc<-ParetoIV(0.5,2,3,100,0.2,0.001,0.001,0.001,0.001,0.001,0.001)
bd<-ParetoIV(0.5,2,3,100,0.3,0.001,0.001,0.001,0.001,0.001,0.001)
ca<-ParetoIV(0.5,2,3,200,0,0.001,0.001,0.001,0.001,0.001,0.001)
cb<-ParetoIV(0.5,2,3,200,0.1,0.001,0.001,0.001,0.001,0.001,0.001)
cc<-ParetoIV(0.5,2,3,200,0.2,0.001,0.001,0.001,0.001,0.001,0.001)
cd<-ParetoIV(0.5,2,3,200,0.3,0.001,0.001,0.001,0.001,0.001,0.001)
da<-ParetoIV(0.5,2,3,500,0,0.001,0.001,0.001,0.001,0.001,0.001)
db<-ParetoIV(0.5,2,3,500,0.1,0.001,0.001,0.001,0.001,0.001,0.001)
dc<-ParetoIV(0.5,2,3,500,0.2,0.001,0.001,0.001,0.001,0.001,0.001)
dd<-ParetoIV(0.5,2,3,500,0.3,0.001,0.001,0.001,0.001,0.001,0.001)

require(xtable)

#########################################################
####  TABELA PARA N=50  ####
#########################################################
cens<-c("0%","0%","0%","10%","10%","10%","20%","20%","20%","30%","30%","30%")
par<-c("alpha","Beta","theta")
EMV<-round(as.numeric(c(aa[1],aa[2],aa[3],
                        ab[1],ab[2],ab[3],
                        ac[1],ac[2],ac[3],
                        ad[1],ad[2],ad[3])),digits = 4)
Mpost<-round(as.numeric(c(aa[4],aa[5],aa[6],
                          ab[4],ab[5],ab[6],
                          ac[4],ac[5],ac[6],
                          ad[4],ad[5],ad[6])),digits = 4)
LIHPD<-round(as.numeric(c(aa[7],aa[9],aa[11],
                          ab[7],ab[9],ab[11],
                          ac[7],ac[9],ac[11],
                          ad[7],ad[9],ad[11])),digits = 4)
LSHPD<-round(as.numeric(c(aa[8],aa[10],aa[12],
                          ab[8],ab[10],ab[12],
```

```
                        ac[8],ac[10],ac[12],
                        ad[8],ad[10],ad[12])),digits = 4)
evalor1<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor1[3]<-round(as.numeric(aa[13]),digits = 5)
evalor1[6]<-round(as.numeric(ab[13]),digits = 5)
evalor1[9]<-round(as.numeric(ac[13]),digits = 5)
evalor1[12]<-round(as.numeric(ad[13]),digits = 5)

evalor2<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor2[1]<-round(as.numeric(aa[14]),digits = 4)
evalor2[4]<-round(as.numeric(ab[14]),digits = 4)
evalor2[7]<-round(as.numeric(ac[14]),digits = 4)
evalor2[10]<-round(as.numeric(ad[14]),digits = 4)

evalor3<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor3[2]<-round(as.numeric(aa[15]),digits = 4)
evalor3[5]<-round(as.numeric(ab[15]),digits = 4)
evalor3[8]<-round(as.numeric(ac[15]),digits = 4)
evalor3[11]<-round(as.numeric(ad[15]),digits = 4)

x<-as.matrix(cbind(cens,par,EMV,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))

x<-data.frame(x)

x<-xtable(x)

print(x)

#########################################################
####  TABELA PARA N=100  ####
#########################################################

cens<-c("0%","0%","0%","10%","10%","10%","20%","20%","20%","30%","30%","30%")
par<-c("alpha","Beta","theta")
EMV<-round(as.numeric(c(ba[1],ba[2],ba[3],
                        bb[1],bb[2],bb[3],
```

```
                        bc[1],bc[2],bc[3],
                        bd[1],bd[2],bd[3])),digits = 4)
Mpost<-round(as.numeric(c(ba[4],ba[5],ba[6],
                          bb[4],bb[5],bb[6],
                          bc[4],bc[5],bc[6],
                          bd[4],bd[5],bd[6])),digits = 4)
LIHPD<-round(as.numeric(c(ba[7],ba[9],ba[11],
                          bb[7],bb[9],bb[11],
                          bc[7],bc[9],bc[11],
                          bd[7],bd[9],bd[11])),digits = 4)
LSHPD<-round(as.numeric(c(ba[8],ba[10],ba[12],
                          bb[8],bb[10],bb[12],
                          bc[8],bc[10],bc[12],
                          bd[8],bd[10],bd[12])),digits = 4)
evalor1<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor1[3]<-round(as.numeric(ba[13]),digits = 5)
evalor1[6]<-round(as.numeric(bb[13]),digits = 5)
evalor1[9]<-round(as.numeric(bc[13]),digits = 5)
evalor1[12]<-round(as.numeric(bd[13]),digits = 5)

evalor2<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor2[1]<-round(as.numeric(ba[14]),digits = 4)
evalor2[4]<-round(as.numeric(bb[14]),digits = 4)
evalor2[7]<-round(as.numeric(bc[14]),digits = 4)
evalor2[10]<-round(as.numeric(bd[14]),digits = 4)

evalor3<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor3[2]<-round(as.numeric(ba[15]),digits = 4)
evalor3[5]<-round(as.numeric(bb[15]),digits = 4)
evalor3[8]<-round(as.numeric(bc[15]),digits = 4)
evalor3[11]<-round(as.numeric(bd[15]),digits = 4)

y<-as.matrix(cbind(cens,par,EMV,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))

y<-data.frame(y)
```

```
y<-xtable(y)

print(y)

##########################################################
####  TABELA PARA N=200  ####
##########################################################

cens<-c("0%","0%","0%","10%","10%","10%","20%","20%","20%","30%","30%","30%")
par<-c("alpha","Beta","theta")
EMV<-round(as.numeric(c(ca[1],ca[2],ca[3],
                        cb[1],cb[2],cb[3],
                        cc[1],cc[2],cc[3],
                        cd[1],cd[2],cd[3])),digits = 4)
Mpost<-round(as.numeric(c(ca[4],ca[5],ca[6],
                          cb[4],cb[5],cb[6],
                          cc[4],cc[5],cc[6],
                          cd[4],cd[5],cd[6])),digits = 4)
LIHPD<-round(as.numeric(c(ca[7],ca[9],ca[11],
                          cb[7],cb[9],cb[11],
                          cc[7],cc[9],cc[11],
                          cd[7],cd[9],cd[11])),digits = 4)
LSHPD<-round(as.numeric(c(ca[8],ca[10],ca[12],
                          cb[8],cb[10],cb[12],
                          cc[8],cc[10],cc[12],
                          cd[8],cd[10],cd[12])),digits = 4)
evalor1<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor1[3]<-round(as.numeric(ca[13]),digits = 5)
evalor1[6]<-round(as.numeric(cb[13]),digits = 5)
evalor1[9]<-round(as.numeric(cc[13]),digits = 5)
evalor1[12]<-round(as.numeric(cd[13]),digits = 5)

evalor2<-c("-","-","-","-","-","-","-","-","-","-","-","-")
evalor2[1]<-round(as.numeric(ca[14]),digits = 4)
evalor2[4]<-round(as.numeric(cb[14]),digits = 4)
evalor2[7]<-round(as.numeric(cc[14]),digits = 4)
evalor2[10]<-round(as.numeric(cd[14]),digits = 4)
```

```
evalor3<-c("-","-","-","-","-","-","-","-","-","-","-","-","-")
evalor3[2]<-round(as.numeric(ca[15]),digits = 4)
evalor3[5]<-round(as.numeric(cb[15]),digits = 4)
evalor3[8]<-round(as.numeric(cc[15]),digits = 4)
evalor3[11]<-round(as.numeric(cd[15]),digits = 4)

z<-as.matrix(cbind(cens,par,EMV,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))

z<-data.frame(z)

z<-xtable(z)

print(z)

#########################################################
####  TABELA PARA N=500  ####
#########################################################

cens<-c("0%","0%","0%","10%","10%","10%","20%","20%","20%","30%","30%","30%")
par<-c("alpha","Beta","theta")
EMV<-round(as.numeric(c(da[1],da[2],da[3],
                        db[1],db[2],db[3],
                        dc[1],dc[2],dc[3],
                        dd[1],dd[2],dd[3])),digits = 4)
Mpost<-round(as.numeric(c(da[4],da[5],da[6],
                          db[4],db[5],db[6],
                          dc[4],dc[5],dc[6],
                          dd[4],dd[5],dd[6])),digits = 4)
LIHPD<-round(as.numeric(c(da[7],da[9],da[11],
                          db[7],db[9],db[11],
                          dc[7],dc[9],dc[11],
                          dd[7],dd[9],dd[11])),digits = 4)
LSHPD<-round(as.numeric(c(da[8],da[10],da[12],
                          db[8],db[10],db[12],
                          dc[8],dc[10],dc[12],
```

```
                        dd[8],dd[10],dd[12])),digits = 4)
evalor1<-c("-","-","-","-","-","-","-","-","-","-","-","-","-")
evalor1[3]<-round(as.numeric(da[13]),digits = 5)
evalor1[6]<-round(as.numeric(db[13]),digits = 5)
evalor1[9]<-round(as.numeric(dc[13]),digits = 5)
evalor1[12]<-round(as.numeric(dd[13]),digits = 5)

evalor2<-c("-","-","-","-","-","-","-","-","-","-","-","-","-")
evalor2[1]<-round(as.numeric(da[14]),digits = 4)
evalor2[4]<-round(as.numeric(db[14]),digits = 4)
evalor2[7]<-round(as.numeric(dc[14]),digits = 4)
evalor2[10]<-round(as.numeric(dd[14]),digits = 4)

evalor3<-c("-","-","-","-","-","-","-","-","-","-","-","-","-")
evalor3[2]<-round(as.numeric(da[15]),digits = 4)
evalor3[5]<-round(as.numeric(db[15]),digits = 4)
evalor3[8]<-round(as.numeric(dc[15]),digits = 4)
evalor3[11]<-round(as.numeric(dd[15]),digits = 4)

w<-as.matrix(cbind(cens,par,EMV,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))

w<-data.frame(w)

w<-xtable(w)

print(w)


##########################################################
#### PARTE 2 - INFLUENCIA DA ESCOLHA DA PRIORI  ##
##########################################################


x1<-ParetoIV(0.5,2,3,100,0.1,0.001,0.001,0.001,0.001,0.001,0.001)
x2<-ParetoIV(0.5,2,3,100,0.1,0.01,0.01,0.001,0.001,0.01,0.01)
x3<-ParetoIV(0.5,2,3,100,0.1,0.1,0.1,0.001,0.001,0.1,0.1)
```

```
x4<-ParetoIV(0.5,2,3,100,0.1,1,1,0.001,0.001,1,1)
x5<-ParetoIV(0.5,2,3,100,0.1,1,1,0.001,0.001,0.001,0.001)
x6<-ParetoIV(0.5,2,3,100,0.1,0.001,0.001,0.001,0.001,1,1)
x7<-ParetoIV(0.5,2,3,100,0.1,1,5,0.001,0.001,1,5)
x8<-ParetoIV(0.5,2,3,100,0.1,1,5,0.001,0.001,0.001,0.001)
x9<-ParetoIV(0.5,2,3,100,0.1,0.001,0.001,0.001,0.001,1,5)
x10<-ParetoIV(0.5,2,3,100,0.1,5,1,0.001,0.001,5,1)
x11<-ParetoIV(0.5,2,3,100,0.1,5,1,0.001,0.001,0.001,0.001)
x12<-ParetoIV(0.5,2,3,100,0.1,0.001,0.001,0.001,0.001,5,1)



num<-c(1:36)
par<-c("alpha","Beta","theta")
Mpost<-round(as.numeric(
  c(x1[4],x1[5],x1[6],x2[4],x2[5],x2[6],
    x3[4],x3[5],x3[6],x4[4],x4[5],x4[6],
    x5[4],x5[5],x5[6],x6[4],x6[5],x6[6],
    x7[4],x7[5],x7[6],x8[4],x8[5],x8[6],
    x9[4],x9[5],x9[6],x10[4],x10[5],x10[6],
    x11[4],x11[5],x11[6],x12[4],x12[5],x12[6])
  ),digits = 4)
LIHPD<-round(as.numeric(
  c(x1[7],x1[9],x1[11],x2[7],x2[9],x2[11],
    x3[7],x3[9],x3[11],x4[7],x4[9],x4[11],
    x5[7],x5[9],x5[11],x6[7],x6[9],x6[11],
    x7[7],x7[9],x7[11],x8[7],x8[9],x8[11],
    x9[7],x9[9],x9[11],x10[7],x10[9],x10[11],
    x11[7],x11[9],x11[11],x12[7],x12[9],x12[11])
),digits = 4)
LSHPD<-round(as.numeric(
  c(x1[8],x1[10],x1[12],x2[8],x2[10],x2[12],
    x3[8],x3[10],x3[12],x4[8],x4[10],x4[12],
    x5[8],x5[10],x5[12],x6[8],x6[10],x6[12],
    x7[8],x7[10],x7[12],x8[8],x8[10],x8[12],
    x9[8],x9[10],x9[12],x10[8],x10[10],x10[12],
    x11[8],x11[10],x11[12],x12[8],x12[10],x12[12])
),digits = 4)
```

```
evalor1<-c(rep("-",36))
evalor1[3]<-round(as.numeric(x1[13]),digits = 4)
evalor1[6]<-round(as.numeric(x2[13]),digits = 4)
evalor1[9]<-round(as.numeric(x3[13]),digits = 4)
evalor1[12]<-round(as.numeric(x4[13]),digits = 4)
evalor1[15]<-round(as.numeric(x5[13]),digits = 4)
evalor1[18]<-round(as.numeric(x6[13]),digits = 4)
evalor1[21]<-round(as.numeric(x7[13]),digits = 4)
evalor1[24]<-round(as.numeric(x8[13]),digits = 4)
evalor1[27]<-round(as.numeric(x9[13]),digits = 4)
evalor1[30]<-round(as.numeric(x10[13]),digits = 4)
evalor1[33]<-round(as.numeric(x11[13]),digits = 4)
evalor1[36]<-round(as.numeric(x12[13]),digits = 4)
evalor1

for (i in 1:36){

  if(evalor1[i]>0.9){
    evalor1[i]<-"-"}

  else
    if(evalor1[i]<0.001){
      evalor1[i]="<0.001"}
}



evalor2<-c(rep(1,36))
evalor2[1]<-round(as.numeric(x1[14]),digits = 4)
evalor2[4]<-round(as.numeric(x2[14]),digits = 4)
evalor2[7]<-round(as.numeric(x3[14]),digits = 4)
evalor2[10]<-round(as.numeric(x4[14]),digits = 4)
evalor2[13]<-round(as.numeric(x5[14]),digits = 4)
evalor2[16]<-round(as.numeric(x6[14]),digits = 4)
evalor2[19]<-round(as.numeric(x7[14]),digits = 4)
evalor2[22]<-round(as.numeric(x8[14]),digits = 4)
```

```
evalor2[25]<-round(as.numeric(x9[14]),digits = 4)
evalor2[28]<-round(as.numeric(x10[14]),digits = 4)
evalor2[31]<-round(as.numeric(x11[14]),digits = 4)
evalor2[34]<-round(as.numeric(x12[14]),digits = 4)




evalor3<-c(rep(1,36))
evalor3[2]<-round(as.numeric(x1[15]),digits = 4)
evalor3[5]<-round(as.numeric(x2[15]),digits = 4)
evalor3[8]<-round(as.numeric(x3[15]),digits = 4)
evalor3[11]<-round(as.numeric(x4[15]),digits = 4)
evalor3[14]<-round(as.numeric(x5[15]),digits = 4)
evalor3[17]<-round(as.numeric(x6[15]),digits = 4)
evalor3[20]<-round(as.numeric(x7[15]),digits = 4)
evalor3[23]<-round(as.numeric(x8[15]),digits = 4)
evalor3[26]<-round(as.numeric(x9[15]),digits = 4)
evalor3[29]<-round(as.numeric(x10[15]),digits = 4)
evalor3[32]<-round(as.numeric(x11[15]),digits = 4)
evalor3[35]<-round(as.numeric(x12[15]),digits = 4)


for (i in 1:36){

  if(evalor3[i]>0.9){
    evalor3[i]<-"-"}

  else
    if(evalor3[i]<0.001){
      evalor3[i]="<0.001"}
}


z<-as.matrix(cbind(num,par,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))
z
```

```
z<-data.frame(z)

z<-xtable(z)

print(z)

#################################
##### EXEMPLO DA APLICAÇÃO #####
#################################

##Este capítulo apresenta uma aplicação da metodologia inferencial do modelo Pareto tipo IV
##em um conjunto de dados reais. Os dados fazem parte do trabalho de Lagakos (1978),
##que analisou o tempo de sobrevivência de n=194 pacientes com câncer de pulmão.



tempo1<-c(11,11,5,22,11,8,1,21,13,4,4,5,17,30,1,21,8,14,8,1,84,16,29,22,15,9,
          12,10,11,30,18,13,5,6,20,19,12,4,17,21,5,9,1,2,7,21,28,14,34,13,1,12,
          88,4,2,21,10,11,8,8,12,2,3,4,7,2,12,3,11,55,13,11,1,8,19,2,3,26,11,4,5,18,22,6,5)
delta1<-c(1,2,2,0,1,1,0,1,0,0,0,0,0,1,0,2,1,0,2,0,2,2,2,1,0,1,2,0,1,0,0,1,1,1,
          2,1,0,1,2,1,1,1,1,0,1,1,0,0,1,2,0,0,1,0,0,0,1,1,1,1,1,2,1,0,0,1,1,2,
          0,2,0,2,0,1,2,0,0,1,2,0,1,0,1,0,2)
tempo2<-c(24,7,17,37,59,13,41,17,40,56,27,19,20,14,5,11,11,9,14,24,33,28,22,34,
          22,39,3,61,10,24,10,12,26,12,31,4,15,34,39,101,11,36,25,34,12,62,9,41,
          38,18,30,6,14,19,12,14,11,29,8,7,15,27,7,19,27,7,4,21,9,40,6,27,34,78,11,
          41,6,14,21,45,9,5,20,25,12,18,35,31,14,13,17,63,6,13,7,53,15,38,15,10,11,
          20,10,11,7,15,32,34,18)
delta2<-c(2,1,2,0,1,1,2,2,1,1,0,0,1,0,1,2,1,2,0,2,0,0,2,2,1,1,2,0,2,0,0,1,0,2,0,0,2,2,
          1,0,2,1,2,0,0,1,1,2,1,0,1,2,0,1,2,2,2,1,1,1,0,1,1,1,1,1,1,1,2,1,0,0,1,1,1,0,0,
          1,1,0,1,1,1,1,0,1,1,0,1,2,1,2,0,1,1,0,1,2,1,1,0,0,0,0,0,1,2,0,1)
tempo<-c(tempo1,tempo2)
delta<-c(delta1,delta2)
delta[which(delta==2)]<-1

require(xtable)
```

```
#Gerar tabela dos dados#
x<-cbind(tempo[1:35],delta[1:35],36:70,
         tempo[36:70],delta[36:70],71:105,
         tempo[71:105],delta[71:105])
x<-as.data.frame(x)
x<-xtable(x,digits=0)
x

y<-cbind(106:140,tempo[106:140],delta[106:140],
         141:175,tempo[141:175],delta[141:175],
         176:194,tempo[176:194],delta[176:194])
y<-as.data.frame(y)
y<-xtable(y,digits=0)
y

estimativas<-c(1:15)

##############################
### log DENSIDADE E SOBREV ###
##############################
log.f<-function(alpha,beta,theta,tempo){
  ( log(theta) + log(alpha) + alpha*log(beta) + (theta-1)*log(tempo) -
      (alpha+1)*log(tempo^theta + beta) )
}

log.s<-function(alpha,beta,theta,tempo){
  ( alpha*log(beta) - alpha*log(tempo^theta + beta) )
}


#####################################
#### FUNÇÃO DE VEROSSIMILHANÇA (-) ###
#####################################

log.vero<-function(parametros,tempo,delta){
  alpha<-parametros[1]
```

```
  beta<-parametros[2]
  theta<-parametros[3]

  if ( (alpha>0) && (beta>0) && (theta>0) )
    return (
      -1*sum( delta*log.f(alpha,beta,theta,tempo) +
                (1-delta)*log.s(alpha,beta,theta,tempo) )
    )
  else return (-Inf)
}

############################
### OBTENÇÃO DOS EMVs  ###
############################
emv<-nlm(log.vero,c(1,1,1),tempo=tempo,delta=delta)
est.EMV<-emv$estimate



#########################
#### log-POSTERIORI ####
#########################
a1<-a2<-a3<-b1<-b2<-b3<-0.001   ## hiperparâmetros da priori

log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-parametros[3]

  if ( (alpha>0) && (beta>0) && (theta>0) )
    return (
      sum( delta*log.f(alpha,beta,theta,tempo) +
             (1-delta)*log.s(alpha,beta,theta,tempo) )
      + (a1-1)*log(alpha) - b1*alpha
      + (a2-1)*log(beta)  - b2*beta
      + (a3-1)*log(theta) - b3*theta
    )
```

```
  else return (-Inf)
}


#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=c(1,1,1),mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
                      delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

################################
### obtenção das estimativas ###
################################
mean(amost[,1])
mean(amost[,2])
mean(amost[,3])
est.bayes<-c(mean(amost[,1]),mean(amost[,2]),mean(amost[,3]))









#############################################################################
#### INTERVALO HPD PARA ALPHA, BETA E THETA - (APROXIMADO, CONFIANÇA=95%)###
#############################################################################
require(TeachingDemos)
hpdalpha<-emp.hpd(amost[,1], conf=0.95)
```

```
hpdbeta<-emp.hpd(amost[,2], conf=0.95)

hpdtheta<-emp.hpd(amost[,3], conf=0.95)



############################################################
#### CALCULANDO A LOG POSTERIORI DOS VALORES SIMULADOS##
############################################################

m<-length(amost[,1])
valores.log.post<-numeric(m)
for (i in 1:m){
  valores.log.post[i]<-log.post(amost[i,],tempo,delta,a1,a2,a3,b1,b2,b3)
}



############################################################
#### CALCULANDO  E VALOR PARA H1:THETA=1##
############################################################

log.post.H1<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-1

  if ( (alpha>0) && (beta>0))
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                  (1-delta)*log.s(alpha,beta,theta,tempo) )
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}
```

```
H1<-nlm(log.post.H1,c(1,1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H1<- -1*H1$min


#e-valor H1:THETA=1
e1<-sum(valores.log.post<max.post.H1)/m


##########################################################
#### CALCULANDO  E VALOR PARA H2:ALPHA=1##
##########################################################

log.post.H2<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-1
  beta<-parametros[1]
  theta<-parametros[2]

  if ( (beta>0) && (theta>0))
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                  (1-delta)*log.s(alpha,beta,theta,tempo) )
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}

H2<-nlm(log.post.H2,c(1,1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H2<- -1*H2$min



#e-valor H2:ALPHA=1
```

```
e2<-sum(valores.log.post<max.post.H2)/m

############################################################
#### CALCULANDO  E VALOR PARA H2:ALPHA=1 E THETA=1 ##
############################################################

log.post.H3<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-1
  beta<-parametros[1]
  theta<-1

  if ( (beta>0) )
    return (
      -1*(sum( delta*log.f(alpha,beta,theta,tempo) +
                 (1-delta)*log.s(alpha,beta,theta,tempo) )
          + (a1-1)*log(alpha) - b1*alpha
          + (a2-1)*log(beta)  - b2*beta
          + (a3-1)*log(theta) - b3*theta)
    )
  else return (-Inf)
}

H3<-nlm(log.post.H3,c(1),tempo=tempo,delta=delta,
        a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)
max.post.H3<- -1*H3$min



#e-valor H3:ALPHA=1 E THETA=1
e3<-sum(valores.log.post<max.post.H3)/m


estimativas[1:3]<-(est.EMV)
estimativas[4:6]<-(est.bayes)
estimativas[7:8]<-(hpdalpha)
estimativas[9:10]<-(hpdbeta)
estimativas[11:12]<-(hpdtheta)
```

```
estimativas[13]<-(e1)
estimativas[14]<-(e2)
estimativas[15]<-(e3)




cens<-c("0%","0%","0%")
par<-c("alpha","Beta","theta")
EMV<-round(as.numeric(c(estimativas[1],estimativas[2],estimativas[3])),digits = 4)

Mpost<-round(as.numeric(c(estimativas[4],estimativas[5],estimativas[6])),digits = 4)

LIHPD<-round(as.numeric(c(estimativas[7],estimativas[9],estimativas[11])),digits = 4)

LSHPD<-round(as.numeric(c(estimativas[8],estimativas[10],estimativas[12])),digits = 4)

evalor1<-c("-","-","-")
evalor1[3]<-round(as.numeric(estimativas[13]),digits = 5)


evalor2<-c("-","-","-")
evalor2[1]<-round(as.numeric(estimativas[14]),digits = 4)


evalor3<-c("-","-","-")
evalor3[2]<-round(as.numeric(estimativas[15]),digits = 4)

x<-as.matrix(cbind(cens,par,EMV,Mpost,LIHPD,LSHPD,
                   evalor1,evalor2,evalor3))

x<-data.frame(x)

x<-xtable(x)

print(x)
```

```
################################
#### VERIFICAÇÃO DO AJUSTE ###
################################
require(survival)
km<-survfit(Surv(tempo,delta)~1)
plot(km,conf.int=F,xlab="Tempo",ylab="Função de Sobrevivência estimada")

tt<-seq(0,max(tempo),0.01)

#TEste

a1<-a2<-a3<-b1<-b2<-b3<-0.001   ## hiperparâmetros da priori

log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-parametros[3]

  if ( (alpha>0) && (beta>0) && (theta>0) )
    return (
      sum( delta*log.f(alpha,beta,theta,tempo) +
```

```
             (1-delta)*log.s(alpha,beta,theta,tempo) )
      + (a1-1)*log(alpha) - b1*alpha
      + (a2-1)*log(beta)  - b2*beta
      + (a3-1)*log(theta) - b3*theta
    )
  else return (-Inf)
}

#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=c(1,1,1),mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
                      delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

################################
### obtenção das estimativas ###
################################

est.bayes<-c(mean(amost[,1]),mean(amost[,2]),mean(amost[,3]))

sobrev.bayes<-exp(log.s(est.bayes[1],est.bayes[2],est.bayes[3],tt))
points(tt,sobrev.bayes,type="l",col=2)









####   monte carlo alpha =1
########################
#### log-POSTERIORI ####
########################
a1<-a2<-a3<-b1<-b2<-b3<-0.001    ## hiperparâmetros da priori
```

```
log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-1
  beta<-parametros[1]
  theta<-parametros[2]

  if ((beta>0) && (theta>0) )
    return (
      sum( delta*log.f(1,beta,theta,tempo) + (1-delta)*log.s(1,beta,theta,tempo) )
      + (a2-1)*log(beta)  - b2*beta
      + (a3-1)*log(theta) - b3*theta
    )
  else return (-Inf)
}


#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=c(1,1),mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
                      delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

#################################
### obtenção das estimativas ###
#################################
mean(amost[,1])
mean(amost[,2])
est.bayes<-c(1,mean(amost[,1]),mean(amost[,2]))


sobrev.bayes<-exp(log.s(est.bayes[1],est.bayes[2],est.bayes[3],tt))
points(tt,sobrev.bayes,type="l",col=4)
```

```
####  monte carlo theta =1

########################
#### log-POSTERIORI ####
########################
a1<-a2<-a3<-b1<-b2<-b3<-0.001    ## hiperparâmetros da priori

log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-parametros[1]
  beta<-parametros[2]
  theta<-1

  if ( (alpha>0) && (beta>0) )
    return (
      sum( delta*log.f(alpha,beta,theta,tempo) + (1-delta)*log.s(alpha,beta,theta,tempo) )
      + (a1-1)*log(alpha) - b1*alpha
      + (a2-1)*log(beta)  - b2*beta
    )
  else return (-Inf)
}


#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=c(1,1),mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
```

```
                    delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

#################################
### obtenção das estimativas ###
#################################
est.bayes<-c(mean(amost[,1]),mean(amost[,2]),1)




sobrev.bayes<-exp(log.s(est.bayes[1],est.bayes[2],est.bayes[3],tt))
points(tt,sobrev.bayes,type="l",col=3)









####  monte carlo alpha=1 e theta=1
########################
#### log-POSTERIORI ####
########################
a1<-a2<-a3<-b1<-b2<-b3<-0.001    ## hiperparâmetros da priori

log.post<-function(parametros,tempo,delta,a1,a2,a3,b1,b2,b3){
  alpha<-1
  beta<-parametros
  theta<-1
```

```
  if (beta>0)
    return (
      sum( delta*log.f(1,beta,1,tempo) + (1-delta)*log.s(1,beta,1,tempo) )
      + (a2-1)*log(beta)  - b2*beta
    )
  else return (-Inf)
}



#############
### MCMC ###
#############
require(MCMCpack)
amost <- MCMCmetrop1R(log.post, theta.init=1,mcmc=500000,
                      burnin=100000,logfun = TRUE,tempo=tempo,
                      delta=delta,a1=a1,a2=a2,a3=a3,b1=b1,b2=b2,b3=b3)

################################
### obtenção das estimativas ###
################################
mean(amost[,1])
mean(amost[,2])
est.bayes<-c(1,mean(amost),1)


sobrev.bayes<-exp(log.s(est.bayes[1],est.bayes[2],est.bayes[3],tt))
points(tt,sobrev.bayes,type="l",col=6)









legend('topright',
       c("Kaplan-Meier","Pareto IV","Alpha=1","Theta=1","Alpha e Theta=1"),
       col=c(1,2,4,3,6),
       text.col=c(1,2,4,3,6),pch=15)
```

```
title(main="Funções de Sobrevivência Estimadas")
```