

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Estimação bayesiana via cópulas para
dados com censura intervalar bivariados

por

Fábio De Araújo Jesus Paixão

Orientador: Prof. Dr. Antônio Eduardo Gomes

Julho de 2015

Fábio De Araújo Jesus Paixão

Estimação bayesiana via cópulas para dados com censura intervalar bivariados

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Universidade de Brasília
Brasília, Julho de 2015

TERMO DE APROVAÇÃO

Fábio de Araújo Jesus Paixão

ESTIMAÇÃO BAYESIANA VIA CÓPULAS PARA
DADOS COM CENSURA INTERVALAR BIVARIADOS

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito parcial à obtenção do título de Mestre em Estatística.

Data da defesa: 30 de 06 de 2015

Orientador:

Prof. Dr. Antônio Eduardo Gomes - Orientador
Departamento de Estatística, UnB

Comissão Examinadora:

Prof. Dra. Cira Etheowalda Guevara Otiniano - Membro da Banca
Departamento de Estatística, UnB

Prof. Dr. Frederico Rodrigues Borges da Cruz - Membro da Banca
Departamento de Estatística, UFMG

Brasília, Julho de 2015

Ficha Catalográfica

PAIXÃO, FÁBIO DE ARAÚJO JESUS

Estimação bayesiana via cópulas para dados com censura intervalar bivariados,
(UnB - IE, Mestre em Estatística, 2015).

Dissertação de Mestrado - Universidade de Brasília. Departamento de Estatística
- Instituto de Ciências Exatas.

1. Análise de Sobrevivência
2. Análise Bayesiana
3. Censura intervalar bivariada
4. Distribuição Weibull
5. Função cópulas
6. Métodos MCMC

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho pode ser reproduzida sem a autorização por escrito do autor.

Fábio de Araújo Jesus Paixão

Sumário

Lista de Figuras	3
Lista de Tabelas	4
Resumo	5
Abstract	6
1 Introdução	7
2 Conceitos Básicos em Análise de Sobrevida	9
2.1 Introdução	9
2.2 Censura	12
2.2.1 Presença de Variáveis Explicativas	15
2.3 Modelos Paramétricos	16
2.3.1 Distribuição Weibull	16
2.3.2 Distribuição Gama	18
2.3.3 Distribuição Log-Normal	18
2.4 Tempos de Falhas Multivariados	19
2.4.1 Caso Bivariado	19
3 Cópulas	22
3.1 Função Cópula	22
3.1.1 Cópulas Arquimedianas	24
3.1.2 Associação	27
3.1.3 Distribuições Marginais Weibull	28

4	Métodos Bayesianos	30
4.1	Paradigma Bayesiano	30
4.1.1	Amostrador de Gibbs	32
4.1.2	Metropolis-Hastings	33
4.2	Metodologia Bayesiana versus Frequentista	34
5	Simulação	36
5.1	Simulação dos tempos de falha dependentes	36
5.1.1	Estimação MCMC em duas etapas	39
6	Resultados e aplicação à dados reais	41
6.1	Simulações	41
6.2	Aplicação	47
7	Conclusões e trabalhos futuros	50
7.1	Conclusões	50
7.2	Trabalhos Futuros	51
	Referências Bibliográficas	52

Lista de Figuras

2.1	Representação gráfica Censura Intervalar Bivariada	20
3.1	Densidade das cópulas de Independência, Gumbel, Clayton e Frank (Lee et al., 2011)	26
6.1	Dependência Caudal das cópulas para amostras $n = 1000$	46
6.2	Ajuste não paramétrico e curva da Weibull ajustada - CMV Sangue .	48
6.3	Ajuste não paramétrico e curva da Weibull ajustada - CMV Urina . .	48

Lista de Tabelas

6.1	Média e [desvio padrão] dos parâmetros das marginais - Clayton . . .	43
6.2	Média e [desvio padrão] dos parâmetros das marginais - Gumbel . . .	43
6.3	Média e [desvio padrão] dos parâmetros das marginais - Frank	44
6.4	Média do parâmetro de associação, segundo o modelo Cópula	45
6.5	Desvio Padrão do parâmetro de associação, segundo o modelo Cópula	45
6.6	Estimativas dos parâmetros das marginais	49
6.7	Estimativas de α	49

Resumo

Neste trabalho apresentamos uma metodologia bayesiana em dois estágios para estimar a função de distribuição conjunta entre tempos de sobrevivência bivariados, com censura intervalar. A estrutura de dependência das variáveis foi representada por um modelo cópula, em particular usando uma função da família das cópulas arquimedianas. As distribuições marginais foram modeladas assumindo tempos de falha simulados com distribuição Weibull e ajustados seguindo o contexto bayesiano, utilizando o método de simulação MCMC (Monte Carlo via Cadeias de Markov) Metropolis-Hastings. Para a estimação da função de distribuição conjunta foram simuladas amostras da distribuição a posteriori conjunta, obtida via metodologia bayesiana com uso de função cópula. Os resultados experimentais demonstram que o método proposto fornece estimativas satisfatórias quando os modelos marginais e de cópulas são supostos corretamente. Todo o código foi implementado utilizando o software estatístico livre R.

Palavras Chave: *Análise de Sobrevivência, análise Bayesiana, censura intervalar bivariada, distribuição Weibull, função cópulas, métodos MCMC.*

Abstract

We present a Bayesian methodology in two stages to estimate the joint distribution function of bivariate interval censored survival data. The dependence structure of the variables was represented by a copula model, in particular by using a Archimedean copula. The marginal distributions were modeled assuming simulated failure times with Weibull distribution and adjusted using the Bayesian framework, using the Metropolis-Hastings MCMC simulation method (Markov chain Monte Carlo). To get the bivariate distribution function we need sample from the posterior distribution, gotten by bayesian methodology using copulas. The experimental results demonstrate that the proposed method provides good estimates when the marginal and copulas models are supposed correctly. All the code was implemented using the free statistical software R.

key words: *Survival analysis, Bayesian analysis , bivariate interval censored, Weibull distribution, Copulas function, MCMC methods.*

1. Introdução

A análise do tempo até a ocorrência de determinado evento, comumente chamada de Análise de Sobrevivência está presente em várias áreas do conhecimento, como na medicina, biologia, engenharia e economia. Esse tipo de estudo caracteriza-se principalmente por apresentar dados censurados ou informação parcial da variável resposta.

O expressivo avanço computacional nas últimas décadas permitiu o surgimento na literatura de diversos métodos para lidar com dados de sobrevivência, entre eles podemos citar modelos paramétricos, não-paramétricos, modelos bayesianos e até mesmo modelos mistos que envolvem mais de uma técnica.

Quando lidamos com dados de sobrevivência bivariados, a suposição de independência entre os tempos de falha pode não ser adequada, e portanto abordagens clássicas usuais podem levar a conclusões equivocadas. Diversos métodos tem sido considerados para lidar com dados de sobrevivência multivariados, com destaque para o modelo de fragilidade. Nesse modelo, um efeito aleatório (fragilidade) é incluído na função de risco para tentar descrever a possível interação entre os tempos de falha.

Como um meio alternativo, podemos modelar marginalmente cada tempo de falha, em um primeiro estágio, e depois considerar o uso de funções cópulas para obter a distribuição conjunta H . Dado que em um modelo cópula as distribuições marginais não dependem da escolha da estrutura de dependência, e portanto podemos proceder a estimação em duas etapas. Na análise de sobrevivência diversos autores estudaram modelos baseados em funções cópulas, por exemplo, Hougaard (1989), Oakes (1989), Shih & Louis (1995) e Gustafson et. al. (2003).

Nesta dissertação, iremos modelar dados de sobrevivência bivariados, com censura intervalar em ambos os tempos de falha, utilizando funções cópulas em um contexto

bayesiano, em que os resultados a posteriori serão obtidos via simulação MCMC (Monte Carlo via Cadeias de Markov). Para captar a estrutura de correlação entre as variáveis será considerada uma função pertencente à classe das cópulas arquimedianas. A estimação será realizada em duas etapas, na primeira será realizado o ajuste bayesiano das distribuições marginais, e a segunda etapa consiste na estimação bayesiana conjunta, via cópulas.

O trabalho está estruturado da seguinte forma: No capítulo 2, são apresentados os conceitos básicos em análise de sobrevivência, no capítulo 3, introduzimos a parte conceitual da metodologia bayesiana e descrevemos os dois métodos de simulação MCMC mais conhecidos, Amostrador de Gibbs e Metropolis-Hastings. No capítulo 4, são apresentados os modelos de cópulas, em especial a família de cópulas arquimedianas, e as medidas de associação. No capítulo 5, são descritos os procedimentos de simulação dos tempos de falha dependentes e estimação da distribuição conjunta H . No capítulo 6, são apresentados os resultados obtidos através dos procedimentos de simulação realizados. Por fim, o capítulo 7 traz as conclusões do nosso estudo. Os programas desenvolvidos estão apresentados no apêndice.

2. Conceitos Básicos em Análise de Sobrevivência

2.1 Introdução

A análise de sobrevivência pode ser definida como um conjunto de técnicas e procedimentos para análise de dados relacionados ao tempo. O seu objeto de estudo ou variável resposta é o tempo T transcorrido até a ocorrência de um evento de interesse, comumente denominado de "falha", a partir de um tempo inicial pré-estabelecido. A principal característica de dados de sobrevivência é a presença de censura, que é a observação parcial da resposta.

O termo análise de sobrevivência refere-se basicamente a situações da área médica que envolvem dados censurados. Entretanto, encontramos em outras áreas, dados que necessitam do mesmo tipo de tratamento. Em engenharia, por exemplo, podemos estar interessados em estimar o tempo de vida de determinado componente, ou mesmo, o tempo até que este componente apresente alguma falha. Também é comum encontrar situações semelhantes nas ciências sociais.

Na área médica, esse tempo, denominado de tempo de falha, pode ser o tempo até o óbito de um paciente, bem como até a cura ou recidiva de uma doença. Na área industrial podemos estar interessados no tempo decorrido até a falha de um equipamento, nesse caso denominamos este estudo como análise de confiabilidade.

Em dados biomédicos, normalmente, o início da medição do tempo coincide com o final do tempo de recrutamento dos indivíduos, enquanto que o final do tempo de estudo depende dos interesses e recursos do pesquisador. Porém, em muitas situações o tempo de início coincide com o instante de recrutamento do indivíduo.

Para pesquisas industriais, o início da medição do tempo coincide, na maioria das vezes, com o momento em que o produto em estudo é colocado em funcionamento para o teste, enquanto que o tempo final coincide com a falha do produto ou com o final do estudo.

A variável de interesse, o tempo transcorrido até o momento de falha, é estritamente positiva e apresenta-se, muitas vezes, em escala contínua. Algumas particularidades da análise de sobrevivência e confiabilidade devem-se a características dos tipos de dados que normalmente estão disponíveis. As duas principais características dos dados de sobrevivência são: a presença de censura e a presença de covariáveis.

O tempo de sobrevivência ou de falha T , pode ser descrito pela sua distribuição de probabilidades, que é expressa através das seguintes funções matematicamente equivalentes, ou seja, a especificação de uma delas é suficiente para se obter as outras. As seguintes funções serão descritas a seguir: função densidade de probabilidade $f(t)$, função de sobrevivência $S(t)$ e função taxa de falha instantânea ou de risco $h(t)$.

Seja T um variável aleatória contínua e não negativa que representa o tempo de sobrevivência de indivíduos de uma população. Todas as funções, exceto quando indicado de outra forma, são definidas no intervalo $[0, \infty)$.

A função densidade de probabilidade (fdp) é expressa como o limite da probabilidade de um indivíduo vir a experimentar o evento de interesse no intervalo de tempo $[t, t + dt)$ por unidade de tempo e é expressa como,

$$f(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt}, \quad (2.1)$$

em que $f(t) \geq 0$ para todo t , e portanto a área abaixo da curva é igual a 1. Seja $F(t)$ a função de distribuição, ou a probabilidade de um indivíduo experimentar o evento de interesse até o tempo t , temos

$$F(t) = P(T \leq t) = \int_0^t f(u) du \quad (2.2)$$

A probabilidade de sobrevivência de um indivíduo pelo menos até o instante de tempo t é dada pela função de sobrevivência:

$$S(t) = 1 - F(t) = P(T > t)$$

Notam-se que $S(t)$ é uma função monótona decrescente com $S(0) = 1$ e $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$. A função taxa de falha ou de risco, $h(t)$, é a taxa instantânea de falha no tempo t , e é definida por

$$\begin{aligned} h(t) &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} \\ &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{P(T \geq t)dt} \\ &= \lim_{dt \rightarrow 0} \left[\frac{F(t + dt) - F(t)}{dt} \right] \frac{1}{S(t)} \\ &= \left(\frac{d}{dt} F(t) \right) \frac{1}{S(t)} \\ &= \frac{f(t)}{S(t)} \end{aligned} \tag{2.3}$$

Em particular, $h(t)dt$ é a probabilidade do indivíduo falhar no intervalo de tempo $[t, t + dt)$, dado que ele tenha sobrevivido até o instante t . As funções $f(t)$, $F(t)$, $S(t)$ e $h(t)$ fornecem especificações matematicamente equivalentes sobre T . É fácil derivar expressões para $S(t)$ e $f(t)$ em termos de $h(t)$. Temos que $f(t) = -\frac{d}{dt}S(t)$, portanto implica que

$$h(t) = -\frac{d}{dt} \log(S(t)) \tag{2.4}$$

Agora integrando os dois lados de (2.3), e exponencializando, chegamos a

$$S(t) = \exp\left(-\int_0^t h(u)du\right) \tag{2.5}$$

A função taxa de falha acumulada $H(t)$, é definida como $H(t) = \int_0^t h(u)du$, a qual é relacionada com a função de sobrevivência por $S(t) = \exp(-H(t))$. Já que $S(\infty) = 0$, segue que $H(\infty) = \lim_{t \rightarrow \infty} H(t) = \infty$. Portanto, a função de risco, $h(t)$, tem as seguintes propriedades

$$h(t) \geq 0 \quad \text{e} \quad \int_0^{\infty} h(t)d(t) = \infty$$

E finalmente a partir de (2.4) e (2.5), temos

$$f(t) = h(t) \exp\left(-\int_0^t h(u)du\right) \quad \text{ou} \quad f(t) = h(t)S(t) \quad (2.6)$$

2.2 Censura

Estudos clínicos que envolvem uma resposta temporal são normalmente prospectivos e de longa duração. Apesar desses estudos serem longos, muitas vezes há perda de informação porque a pesquisa é encerrada antes de se observar falha em todos os pacientes, ou porque alguns pacientes abandonam a pesquisa antes de ocorrer o evento de interesse. Note que a única informação obtida para estes indivíduos é que o tempo de ocorrência do evento de interesse é maior que o tempo registrado.

Sem a presença de censura, as técnicas estatísticas clássicas usuais, como análise de regressão e planejamento de experimentos, poderiam ser utilizadas na análise desse tipo de dados. Podemos citar o exemplo em que o interesse seja comparar o tempo médio de vida de três grupos de pacientes. Se não houver censuras, pode-se usar a análise de variância para se fazer tal comparação. Se houver censuras, não poderíamos usar as técnicas usuais como a análise de variância, pois estas técnicas necessitam de todos os tempos de falha. Portanto, é necessário o uso dos métodos de análise de sobrevivência, pois estes possibilitam incorporar na análise estatística a informação contida nos dados censurados.

É importante que, mesmo censurados, todos os resultados provenientes de um estudo de sobrevivência devem ser usados na análise estatística. Duas razões justificam esse procedimento: (i) mesmo sendo incompletas, as observações censuradas fornecem informações sobre o tempo de vida de pacientes; (ii) a omissão das censuras no cálculo das estatísticas de interesse pode acarretar estimativas viciadas e conclusões equivocadas.

Alguns tipos de censuras são diferenciados na análise de sobrevivência. A censura do tipo I ocorre quando a pesquisa tem uma duração pré-estabelecida independentemente do número de falhas observadas. A censura do tipo II é aquela em que define-se no início do estudo quantas falhas devem ser observadas e, atingido este número, termina-se o estudo. Um terceiro tipo de censura, o do tipo aleatório, ocorre quando um paciente é retirado do estudo antes de ter ocorrido falha, ou também, por exemplo se o paciente falecer por outra causa.

As censuras podem aparecer de três formas, nos diferentes tipos de censura:

Censura à Direita: Caracteriza-se pelo fato do tempo de ocorrência da falha ser maior que o tempo registrado.

Censura à Esquerda: Ocorre quando o tempo de ocorrência do evento de interesse é menor que o tempo registrado, ou seja, a falha ocorreu antes do indivíduo ser observado.

Censura Intervalar: Este é o tipo mais geral de censura que acontece, por exemplo, quando são feitos exames periódicos nos pacientes, é conhecido somente que o evento de interesse ocorreu num determinado intervalo de tempo. Pelo fato do tempo T não ser conhecido exatamente, mas sim pertencer a um intervalo, isto é, $T \in (U, V]$, estes dados são denominados por *sobrevivência intervalar* ou, usualmente, por *dados de censura intervalar*.

Quando somente uma observação do tempo é registrada, tem-se uma estrutura de dados conhecida como *caso 1 de censura intervalar*. Tal caso é comum quando a realização de dois exames é impossível ou inviável, por exemplo, quando deseja-se estimar o prazo de validade de um determinado produto embalado. Não se sabe o instante de início da deterioração, mas somente depois que o produto é aberto, é possível verificar se já houve deterioração.

Quando são feitos exames periódicos, ou seja, pode-se ter duas informações de tempo para um mesmo caso, tem-se uma estrutura de dados conhecida como *caso geral de censura intervalar*. Este caso é observado, por exemplo, quando deseja-

se identificar o tempo que determinado equipamento leva para apresentar problemas. Como muitos problemas não causam falhas fatais ao equipamento, não é possível identificá-los imediatamente quando ocorrem. Mas é possível fazer vistorias periódicas e identificar se o equipamento apresenta algum problema. Estes dados são representados com as funções indicadoras:

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

e

$$\gamma_i = \begin{cases} 1, & \text{se } L_i < T_i \leq U_i \\ 0, & \text{se } T_i > U_i \text{ ou } T_i \leq L_i. \end{cases}$$

em que L_i e U_i são, respectivamente, os limites inferior e superior do intervalo observado para o i -ésimo indivíduo.

Vale ressaltar que a censura intervalar generaliza as outras formas de censura (esquerda ou direita). No caso da censura a esquerda temos $T \in (0, L]$, ou seja, $L = 0$. Para a censura a direita temos que $T \in [U, \infty)$, isto é, $U = \infty$.

Censura Intervalar

São estudados dois casos de censura intervalar:

Caso 1. Seja $(T_1, L_1), \dots, (T_n, L_n)$ uma amostra de variáveis aleatórias no espaço \mathbb{R}_+^2 em que os T_i e L_i são variáveis independentes (não negativas) com função de distribuição F e G , respectivamente. Neste caso, a única variável que é observada é L_i (“observações de tempo”) e seja $\delta_i = I_{\{T_i \leq L_i\}}$ em que I_A é a variável indicadora de ocorrência do evento A . Tem-se a função de log-verossimilhança de F igual a:

$$\mathcal{L}(F) = \sum_{i=1}^n \{\delta_i \log F(L_i) + (1 - \delta_i) \log(1 - F(L_i))\}, \quad (2.7)$$

em que $F(t) = P(T_i \leq t)$ é estritamente contínua.

Caso 2. Seja $(T_1, L_1, U_1), \dots, (T_n, L_n, U_n)$ uma amostra aleatória no espaço \mathbb{R}_+^3 , em que T_i é uma variável aleatória (não negativa) com função de distribuição F , e L_i

e U_i são variáveis aleatórias (não negativas), independentes de T_i , com distribuição conjunta H e $P(L_i \leq U_i) = 1$. Neste caso, as variáveis que podem ser observadas são (L_i, U_i) (as “observações do tempo”) e $\delta = I_{\{T_i \leq L_i\}}$, $\gamma_i = I_{\{T_i \in (L_i, U_i]\}}$. Nesse caso, as funções verossimilhança e log-verossimilhança de F são dadas pelas funções:

$$L(F) = \prod_{i=1}^n \{[F(L_i)]^{\delta_i} + [F(U_i) - F(L_i)]^{\gamma_i} + [1 - F(U_i)]^{1-\delta_i-\gamma_i}\}. \quad (2.8)$$

e

$$\mathcal{L}(F) = \sum_{i=1}^n \{\delta_i \log F(L_i) + \gamma_i \log(F(U_i) - F(L_i)) + (1 - \delta_i - \gamma_i) \log(1 - F(U_i))\}. \quad (2.9)$$

2.2.1 Presença de Variáveis Explicativas

Além do tempo de sobrevivência e da variável indicadora de censura, também podemos observar nos dados algumas variáveis que explicam parte da variação no tempo de sobrevivência, como: sexo, idade, tratamento, dentre outras. Estas variáveis são chamadas de “variáveis explicativas” ou “covariáveis”. Quando os tempos de sobrevivência estão relacionados com essas variáveis, dizemos que a população em estudo é heterogênea, caso contrário ela é homogênea.

Muitas vezes, o objetivo da pesquisa não é apenas estudar o tempo que certo evento leva para ocorrer, mas também possíveis tratamentos ou características que podem alterar esse tempo. Assim, temos na análise de sobrevivência a variável tempo de sobrevivência, a variável indicadora de censura e um vetor de variáveis explicativas disponíveis para análise. Um complicador para essa análise ocorre quando as variáveis explicativas dependem do tempo, ou seja, no fim do experimento a variável explicativa pode ter valores diferentes dos valores do início do estudo. Isso acontece quando, por exemplo, um tratamento aumenta a dose de certo remédio com o decorrer do tempo.

2.3 Modelos Paramétricos

2.3.1 Distribuição Weibull

A família de distribuições Weibull foi proposta originalmente em 1939 e sua ampla aplicabilidade foi também discutida em 1951 e 1954, pelo físico sueco Waloddi Weibull (Weibull, 1951). Devido à sua grande flexibilidade e simplicidade em se adaptar a diversos fenômenos, essa distribuição tem sido muito utilizada para explicar a variabilidade de fenômenos temporais e não-temporais, tendo grande aplicabilidade na modelagem de dados biomédicos e industriais. Alguns exemplos de ocorrência de variáveis aleatórias não-negativas e temporais podem ser observados em tempos de sobrevivência, tempos de espera, duração de epidemias *etc.* Exemplos não-temporais de variáveis não-negativas incluem resistência de materiais, dimensões de partículas, níveis pluviométricos *etc.* Embora as distribuições Exponencial e Gama forneçam ajustes razoáveis para a distribuição de frequência de algumas dessas variáveis aleatórias, em alguns casos esse ajuste pode ser insatisfatório. A distribuição de Weibull se mostra mais eficiente para o estudo desses fenômenos. Sua função de densidade de probabilidade tem a seguinte forma:

$$f(t) = \frac{\beta}{\lambda} \left(\frac{t}{\lambda}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\lambda}\right)^\beta\right], t \geq 0, \quad (2.10)$$

em que $\beta > 0$ e $\lambda > 0$ são os parâmetros. O parâmetro β determina a forma da curva da função densidade. Alguns valores de β podem resultar em outras distribuições. Como exemplo temos, quando $\beta = 1$, a distribuição exponencial com parâmetro $\theta = \frac{1}{\lambda}$ e, sendo assim, a distribuição exponencial é um caso particular da distribuição de Weibull. O parâmetro β é chamado de parâmetro de forma. Temos que λ é o parâmetro de escala. Ao avaliarmos graficamente a função densidade de probabilidade, alterar o valor de λ equivale a mudar a escala do eixo das abscissas.

A função de distribuição é dada por:

$$F(t) = 1 - \exp\left(-\frac{t}{\lambda}\right)^\beta. \quad (2.11)$$

As funções de risco e de sobrevivência são dadas, respectivamente, por:

$$h(t) = \frac{\beta}{\lambda} \left(\frac{t}{\lambda}\right)^{\beta-1} \quad (2.12)$$

e

$$S(t) = \exp\left[-\left(\frac{t}{\lambda}\right)^\beta\right] \quad (2.13)$$

Uma característica relevante para o uso da distribuição Weibull na modelagem de tempos de sobrevivência é sua abrangência com relação às diferentes formas de funções de risco. Para $\beta < 1$, têm-se função de risco monótona decrescente (ocorre em processos onde o número de falhas é grande no início e decai com o tempo, como por exemplo, mortalidade infantil). Para $\beta > 1$, a função será monótona crescente (geralmente em processos em que há materiais submetidos a desgaste). E para $\beta = 1$, têm-se a distribuição exponencial com função de risco constante.

As expressões para a média e a variância da distribuição Weibull incluem o uso da função gama, isto é,

$$E(T) = \lambda \Gamma\left(1 + \frac{1}{\beta}\right), \quad (2.14)$$

$$Var(T) = \lambda^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2 \right], \quad (2.15)$$

sendo a função gamma, $\Gamma(k)$, definida por $\Gamma(k) = \int_0^\infty x^{k-1} \exp(-x) dx$.

2.3.2 Distribuição Gama

A distribuição gama é caracterizada por dois parâmetros λ e β , e pertence a família das distribuições de probabilidade contínuas. As distribuições exponencial e Qui-quadrado são casos especiais da distribuição gama.

Se $X \sim Gama(\lambda, \beta)$ então:

$$f(x) = \frac{\beta^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\beta x}; \quad E(X) = \frac{\lambda}{\beta}; \quad Var(X) = \frac{\lambda}{\beta^2}$$

em que x, λ e $\beta > 0$.

2.3.3 Distribuição Log-Normal

Assim como a distribuição Weibull, a distribuição log-normal é muito utilizada na caracterização de tempos de sobrevivência.

Uma variável aleatória $Y \sim LN(\mu, \sigma^2)$, se seu logaritmo $X = \log(Y)$ segue uma distribuição normal. Logo sua função de densidade é

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{\log(x) - \mu}{2\sigma^2}\right]; \quad x > 0$$

Sua função de sobrevivência é denotada por

$$S(x) = 1 - \Phi\left[\frac{\log(x) - \mu}{\sigma}\right]$$

em que $\Phi(\cdot)$ é a função de distribuição acumulada de uma variável normal padrão.

2.4 Tempos de Falhas Multivariados

No caso univariado de dados de sobrevivência, é considerado que os tempos de sobrevivência de indivíduos distintos são independentes. Embora essa suposição seja adequada para muitos estudos, ela pode não ser válida para outros. Algumas vezes, os tempos de sobrevivência são observados em grupos homogêneos de indivíduos, e portanto os tempos, dentro de cada grupo, podem não ser mutuamente independentes. Por exemplo quando são observados tempos de sobrevivência em indivíduos de uma mesma família, espera-se que esses tempos apresentem certa semelhança que não seriam observadas entre indivíduos sem laços familiares. Portanto é razoável supor que existe associação entre os tempos de um mesmo grupo.

Outra situação em que podemos observar associação entre os tempos é quando, por exemplo, cada indivíduo de um estudo está sujeito a múltiplos eventos do mesmo tipo, conhecidos como eventos recorrentes, tais como ataques epiléticos ou ataques cardíacos. Podem-se observar também eventos de tipos diferentes no mesmo indivíduo, tais como múltiplas sequelas em pacientes com doenças crônicas.

Situações como as citadas, em que é razoável supor a existência de associação entre os tempos de sobrevivência, caracterizam dados de sobrevivência multivariados. Neste trabalho serão considerados apenas os casos de tempos de sobrevivência bivariados.

Os tempos de sobrevivência multivariados são obtidos quando:

- Dois ou mais eventos, do mesmo tipo ou de tipos diferentes, ocorrem no mesmo indivíduo;
- Eventos ocorrem em indivíduos que estão agrupados pelo planejamento usado no estudo, e existe razão para assumir uma possível estrutura de correlação entre os tempos de falha de indivíduos do mesmo grupo, sendo que existe independência entre os grupos.

2.4.1 Caso Bivariado

Sejam T_1 e T_2 variáveis aleatórias contínuas e positivas, e H sua função de distribuição conjunta. Suponha que existe um mecanismo de censura, independente de (T_1, T_2) ,

portanto o par (T_1, T_2) não é observado diretamente. Então, em vez de uma observação (t_1, t_2) , observaremos uma região retangular $R \subset \mathbb{R}^2$. Portanto os dados consistem de N observações de retângulos R_1, \dots, R_N , e o objetivo é calcular o estimador de máxima verossimilhança (EMV) \hat{H}_n de H .

Nossa amostra será composta pelos tempos de observação ou vértices dos retângulos R_1, \dots, R_N representados por $[L_{11}, U_{11}] \times [L_{21}, U_{21}], \dots, [L_{1N}, U_{1N}] \times [L_{2N}, U_{2N}]$ em que $T_{1i} \in [L_{1i}, U_{1i}]$ e $T_{2i} \in [L_{2i}, U_{2i}]$, ressaltando que podemos ter $L_{.i} = 0$ e $U_{.i} = \infty$. Na figura (2.1) apresentamos um exemplo gráfico para esse tipo de dados.

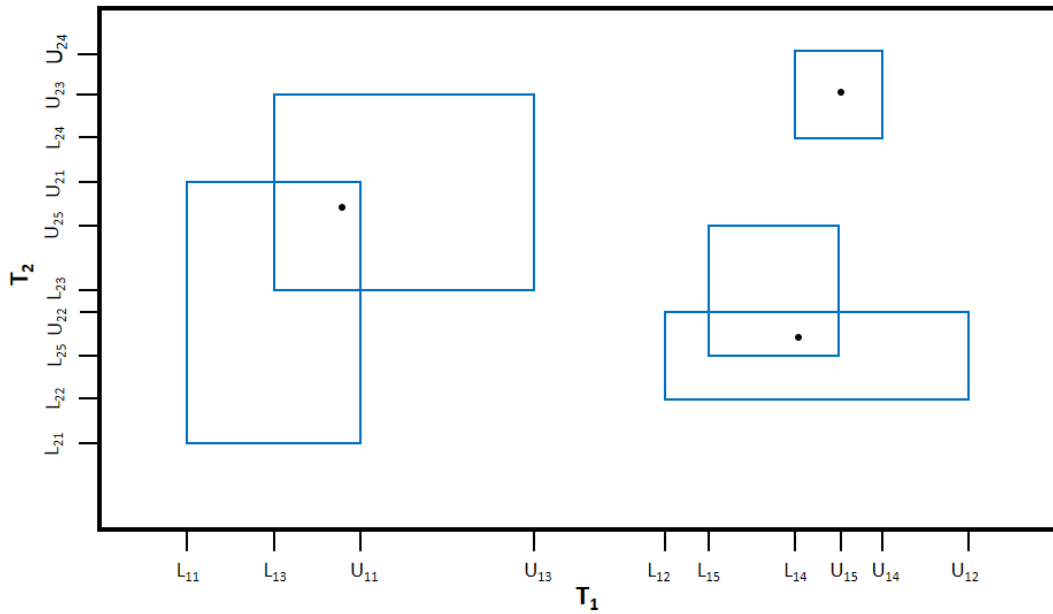


Figura 2.1: Representação gráfica Censura Intervalar Bivariada

Então a função de verossimilhança para esses dados, que depende de H é definida por:

$$L(H) = \prod_{i=1}^N [H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i})] \quad (2.16)$$

Calcular o estimador de máxima verossimilhança paramétrico \hat{H}_n diretamente a partir de (2.16) pode não ser uma tarefa trivial, pois não sabemos se estamos lidando

com tempos de falhas independentes, e portanto obter a estrutura de dependência entre T_1 e T_2 exige o conhecimento do tipo e grau de dependência entre as variáveis. Para tentar lidar com esse problema, utilizaremos os modelos de cópulas. Assim como no caso univariado o estimador de máxima verossimilhança não-paramétrico de H não possui forma fechada.

3. Cópulas

3.1 Função Cópula

O uso de cópulas se mostrou uma ferramenta muito útil para modelar uma função de distribuição conjunta de variáveis de interesse. Em particular, ganhou importância por ser uma função relativamente simples para descrever estrutura de dependência entre variáveis aleatórias com distribuição conjunta. Como um modelo de estrutura de dependência, as cópulas possuem diversas vantagens sobre outras medidas de dependência, como por exemplo, o coeficiente de correlação linear. O uso de cópulas permite modelar dependências lineares e não lineares, e também mensurar o grau de dependência na cauda de distribuições.

Cópula é uma função de distribuição multivariada com funções de distribuição marginais uniformes, denotada por \mathbf{C} . Seja T_1 e T_2 os tempos de falha, e seja (S_1, S_2) e (F_1, F_2) as respectivas funções de sobrevivência e distribuição, então as funções de sobrevivência e distribuição conjuntas de T_1 e T_2 (Nelsen 2006), são dadas por

$$S(t_1, t_2) = C(S_1(t_1), S_2(t_2))$$

$$F(t_1, t_2) = C(F_1(t_1), F_2(t_2))$$

Para o caso bivariado, a forma da cópula é a maneira mais fácil de expressar e gerar distribuições conjuntas. No caso bivariado, a cópula é uma função $C : [0, 1]^2 \rightarrow [0, 1]$ tal que:

Para todo u e v em $[0, 1]$,

$$\lim_{u_i \rightarrow 0} C(u_1, u_2) = 0 \quad \text{para algum } i,$$

$$\lim_{u_i \rightarrow 1} C(u_1, u_2) = 1 \quad \text{para } i = 1, 2,$$

e

$$C(u, 1) = u \quad \text{e} \quad C(1, v) = v$$

Para todo u_1, u_2, v_1, v_2 em $[0, 1]$ tal que $0 \leq u_1 \leq u_2 \leq 1$ e $0 \leq v_1 \leq v_2 \leq 1$,

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$$

Quando $F_1(t_1) = u$ e $F_2(t_2) = v$, a função cópula $C(F_1(t_1), F_2(t_2))$ é a função de distribuição bivariada. Inversamente qualquer função de distribuição bivariada $F(t_1, t_2)$ com funções marginais contínuas F_1 e F_2 podem ser unicamente expressas por uma função cópula.

$$C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v))$$

Teorema de Sklar: Seja F uma função de distribuição conjunta bivariada das variáveis contínuas T_1 e T_2 com marginais F_1 e F_2 respectivamente. Existe uma cópula C (isto é, uma função de distribuição bivariada em $[0, 1]^2$ com funções de distribuição marginais uniformes) de tal modo que, para $-\infty < t_1 < \infty, -\infty < t_2 < \infty$,

$$F(t_1, t_2) = P(T_1 < t_1, T_2 < t_2) = C(F_1(t_1), F_2(t_2)) \quad (3.1)$$

Note que o teorema de Sklar implica simplesmente que $C(u, v) = P(U < u, V < v)$ para variáveis aleatórias uniformes U e V em $[0, 1]$.

Teorema (Limites de Fréchet-Hoeffding) Para toda cópula C e todo (u, v) em $[0, 1]^2$,

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v)$$

Note que pelo teorema de Sklar's $F(t_1, t_2) = C(F_1(t_1), F_2(t_2))$, em que $C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v))$ e $u = F_1(t_1)$, $v = F_2(t_2)$. Então

$$\max(F_1(t_1) + F_2(t_2) - 1, 0) \leq F(t_1, t_2) \leq \min(F_1(t_1), F_2(t_2))$$

Para se obter a função de distribuição conjunta multivariada a partir de uma cópula, é necessária apenas uma simples transformação em cada variável marginal, em que cada marginal transformada tem distribuição uniforme. Portanto a estrutura de dependência pode ser expressa como uma distribuição multivariada sobre as uniformes obtidas, e uma cópula é precisamente uma distribuição multivariada a partir de variáveis aleatórias marginalmente uniformes.

Especificar a dependência entre os tempos de falha T_1 e T_2 , é o mesmo que especificar a dependência entre U e V , o que pode ser uma tarefa infinitamente mais simples.

Neste caso, há muitas famílias de cópulas que diferem no detalhe da dependência que elas representam.

3.1.1 Cópulas Arquimedianas

A Cópula Arquimediana é um método conveniente para modelar um distribuição bivariada, devido à sua forma simples e uma variedade de estruturas de dependência. O uso de transformadas de Laplace leva a construção de cópulas Arquimedianas. Especificamente, seja φ uma função monotonicamente decrescente de $[0, 1] \rightarrow [0, \infty)$ tal que $\varphi(1) = 0$ e $\varphi''(x) > 0$. Defina a pseudo inversa como: $\varphi^{[-1]}(x) = \varphi^{-1}(x)$ para $0 \leq x \leq \varphi(0)$ e zero para $\varphi(0) \leq x \leq \infty$. Note que se $\varphi(0) = \infty$, então

$\varphi^{[-1]} = \varphi^{-1}$. Para números reais u e v , uma cópula Arquimediana C , de variáveis aleatórias bivariadas U e V é definida:

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (3.2)$$

em que $\varphi : [0, 1] \rightarrow [0, \infty)$, e φ é chamado de gerador da cópula.

Cópula Arquimedianas envolvem um parâmetro de dependência caudal, também conhecido com parâmetro de associação. Este parâmetro descreve a magnitude da dependência na cauda inferior ou superior de uma distribuição multivariada e pode ser utilizado para analisar a dependência entre valores extremos. O gerador φ contém todas a informação sobre a estrutura de dependência de uma distribuição multivariada de variáveis aleatórias em termos do parâmetro de associação θ .

Baseado no nível da estrutura de dependência caudal, serão consideradas quatro famílias de cópulas Arquimedianas: Gumbel, Clayton, Frank e cópula da independência. Como pode-se observar na figura 3.1:

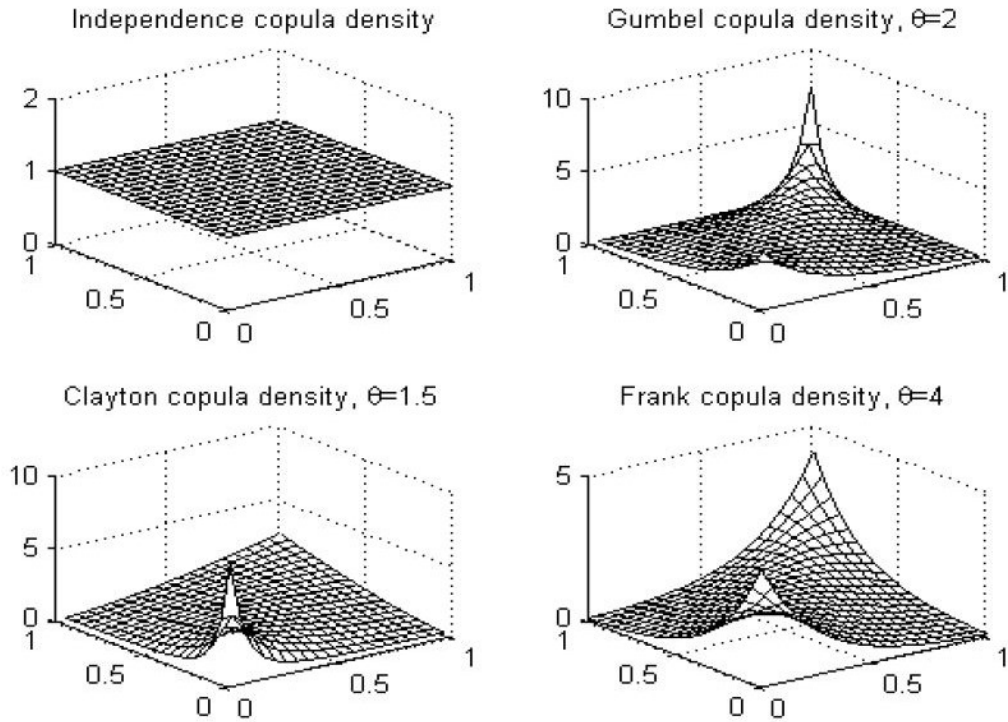


Figura 3.1: Densidade das cópulas de Independência, Gumbel, Clayton e Frank (Lee et al., 2011)

Gumbel:

$$C(u, v; \theta) = \exp \left\{ - \left[(-\log u)^\theta + (-\log v)^\theta \right]^{\frac{1}{\theta}} \right\}$$

e $\varphi(t) = (-\log t)^\theta$ para $\theta \geq 1$. $\theta = 1$ implica independência entre as distribuições. Quando $\theta \rightarrow \infty$, a cópula de Gumbel atinge o limite inferior de Fréchet, portanto a distribuição é caracterizada por valores extremos. Isso implica alta dependência na cauda superior.

Clayton:

$$C(u, v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$$

e $\varphi(t) = \frac{t^{-\theta} - 1}{\theta}$ para $\theta > 0$. Quando $\theta \rightarrow \infty$, a cópula de Clayton atinge o limite superior de Fréchet, portanto implica alta dependência na cauda inferior. E $\theta \rightarrow 0$ implica independência entre as distribuições.

Frank:

$$C(u, v; \theta) = -\frac{1}{\theta} \cdot \log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right]$$

e $\varphi(t) = -\log \left[\frac{e^{-\theta t} - 1}{e^{-\theta} - 1} \right]$ para $\theta \neq 0$. Quando $\theta \rightarrow \infty$, a cópula de Frank atinge o limite superior de Fréchet, e quando $\theta \rightarrow -\infty$ atinge o limite inferior de Fréchet. Quando $\theta \rightarrow 0$ implica independência entre as distribuições.

Independência:

$$C(u, v) = u.v$$

e $\varphi(t) = -\log t$

A dependência caudal de cópulas pode ser ilustrada pela sua função densidade.

$$f(x_1, x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2)$$

em que c é a densidade de C e f_1 e f_2 são as marginais.

3.1.2 Associação

Uma medida de associação muito utilizada é o Kendall's τ :

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 \quad (3.3)$$

Pela expressão (4.3) Kendall's τ é calculado pela cópula que contém o parâmetro de associação θ . Por outro lado, o parâmetro de associação θ pode ser obtido pelo Kendall's τ calculado a partir dos dados. Para cópulas Arquimedianas bivariadas, em que as duas variáveis são absolutamente contínuas, o coeficiente τ pode ser calculado pela seguinte expressão:

$$\tau = 1 + \left(\int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt \right) \quad (3.4)$$

O gerador da cópula possui o parâmetro de associação, e a partir de (4) o gerador pode ser expresso através de τ . Portanto θ pode ser obtido resolvendo a expressão (4).

Gumbel:

$$\tau = \frac{\theta - 1}{\theta}$$

Clayton:

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

Frank:

$$\tau = 1 - \frac{4}{\theta} \left(1 - \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt \right)$$

3.1.3 Distribuições Marginais Weibull

Dadas duas distribuições marginais Weibull

$$F_i(t_i) = 1 - e^{-\left(\frac{t_i}{\beta_i}\right)^{\lambda_i}}, i = 1, 2$$

É possível construir uma distribuição bivariada $F(t_1, t_2)$ tal que $F(t_1, t_2) = C(F_1(t_1), F_2(t_2))$. Utilizando as cópulas Arquimedianas a distribuição conjunta é dada

por,

Gumbel:

$$F(t_1, t_2) = \exp \left\{ - \left[(-\log F_1(t_1))^\theta + (-\log F_2(t_2))^\theta \right]^{\frac{1}{\theta}} \right\}$$

Clayton:

$$F(t_1, t_2) = \left(F_1(t_1)^{-\theta} + F_2(t_2)^{-\theta} - 1 \right)^{-\frac{1}{\theta}}$$

Frank:

$$F(t_1, t_2) = -\frac{1}{\theta} \cdot \log \left[1 + \frac{(e^{-\theta F_1(t_1)} - 1)(e^{-\theta F_2(t_2)} - 1)}{(e^{-\theta} - 1)} \right]$$

Independência:

$$F(t_1, t_2) = F_1(t_1)F_2(t_2)$$

4. Métodos Bayesianos

4.1 Paradigma Bayesiano

O paradigma Bayesiano consiste em especificar um modelo de probabilidade para os dados observados D , dado um vetor de parâmetros desconhecidos $\boldsymbol{\theta}$, levando em conta a função de verossimilhança $L(\boldsymbol{\theta}|D)$. Então assume-se que o vetor $\boldsymbol{\theta}$ é aleatório e tem distribuição a *priori* denotada por $\pi(\boldsymbol{\theta})$. A inferência quanto a $\boldsymbol{\theta}$ é em seguida baseada na distribuição a *posteriori*, obtida pelo teorema de Bayes. A distribuição a *posteriori* de $\boldsymbol{\theta}$ é dada por

$$\pi(\boldsymbol{\theta}|D) = \frac{L(\boldsymbol{\theta}|D)\pi(\boldsymbol{\theta})}{\int_{\Theta} L(\boldsymbol{\theta}|D)\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (4.1)$$

em que Θ denota o espaço paramétrico de $\boldsymbol{\theta}$. A partir de (3.1) nota-se que $\pi(\boldsymbol{\theta}|D)$ é proporcional a verossimilhança multiplicada pela *priori*,

$$\pi(\boldsymbol{\theta}|D) \propto L(\boldsymbol{\theta}|D)\pi(\boldsymbol{\theta})$$

e portanto envolve a contribuição dos dados observados através de $L(\boldsymbol{\theta}|D)$, e uma contribuição da informação a *priori* quantificada em $\pi(\boldsymbol{\theta})$. A quantidade $m(D) = \int_{\Theta} L(\boldsymbol{\theta}|D)\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$ é a constante normalizadora de $\pi(\boldsymbol{\theta}|D)$, e é comumente chamada de distribuição marginal dos dados ou distribuição a *priori* preditiva.

Na maioria dos modelos e aplicações, $m(D)$ não possui uma forma analítica fechada, e portanto $\pi(\boldsymbol{\theta}|D)$ não possui forma fechada. No entanto, existem diversos métodos computacionais para simulação de amostras da distribuição a *posteriori*

$\pi(\boldsymbol{\theta}|D)$ assim como métodos para estimar $m(D)$. Um dos mais conhecidos métodos computacionais para simulação de amostras de $\pi(\boldsymbol{\theta}|D)$ é chamado de Amostrador de Gibbs. O Amostrador de Gibbs é um poderoso algoritmo de simulação que permite amostrar a partir de $\pi(\boldsymbol{\theta}|D)$ sem precisar do conhecimento da constante normalizadora $m(D)$. Outro importante método computacional é o algoritmo de Metropolis-Hastings. Existem também diversos outros algoritmos híbridos.

Entre os métodos de Monte Carlo via Cadeias de Markov (MCMC), os mais utilizados são o Amostrador de Gibbs e Metropolis-Hastings. Quando a geração não-iterativa da distribuição da qual se deseja obter uma amostra for muito complicada ou custosa, recomenda-se em geral o uso de algoritmos MCMC.

Os métodos MCMC consistem em simular observações diretas de alguma distribuição complexa de interesse, no caso bayesiano a distribuição a *posteriori*, que são ligeiramente dependentes. A idéia desse método é transformar um problema estático, em um problema de natureza dinâmica, construindo um processo estocástico temporal artificial, simples de simular e que convirja na distribuição de interesse. Este processo temporal artificial simulado é em geral uma Cadeia de Markov homogênea.

Para melhor entendimento desse método, precisamos de algumas definições.

Processo Estocástico: É uma família de variáveis aleatórias $\{X_t\}_{t \geq 0}$ definidas sobre o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, no qual t é um parâmetro linear ao longo de um conjunto de índices T .

Cadeia de Markov: É um caso particular de um processo estocástico, de tempo discreto e variáveis discretas, no qual os estados futuros são independentes dos estados passados, dado o estado presente.

$$p(\theta^{(t+1)} | \theta^{(1)}, \theta^{(2)}, \dots, \theta^{(t)}) = p(\theta^{(t+1)} | \theta^{(t)})$$

A propriedade que a distribuição do estado seguinte é conhecida, sabendo-se apenas o estado atual, é conhecida como "perda de memória", ou propriedade de Markov.

Integração de Monte Carlo: Simulações que aproximam numericamente integrais.

Se queremos resolver uma integral qualquer I , podemos aproximar I a partir do estimador \hat{I}_n e $\hat{I}_n \rightarrow I$ é um estimador simulado consistente se $n \rightarrow \infty$. Isso deve ser verdade a partir da Lei Forte dos Grandes Números para Cadeias de Markov, porém é necessário que a sequência de observações seja independente. Não é possível gerar uma sequência de amostras independentes, mas é possível simular uma sequência ligeiramente dependente utilizando uma Cadeia de Markov, e então obter quantidades de interesse a partir das amostras simuladas (*posteriori*).

É importante salientar que os valores gerados a partir de uma simulação iterativa não são considerados independentes. Desse modo, é usual obter muitas iterações do método, e descartar uma quantidade suficiente das primeiras observações simuladas, para obter amostras independentes da distribuição, processo conhecido como *Burn-in*.

4.1.1 Amostrador de Gibbs

O Amostrador de Gibbs é um dos mais conhecidos algoritmos MCMC na literatura computacional Bayesiana. Como discutido em Besag & Green (1993), o Amostrador de Gibbs foi baseado nas idéias de Grenander (1983), e formalmente desenvolvido por Geman & Geman (1984). Foi primeiramente utilizado em problemas de inferência Bayesiana por Gelfand & Smith (1990).

O método consiste em obter amostras de uma distribuição conjunta $p(\theta_1, \dots, \theta_k)$, no nosso caso a distribuição a posteriori, a partir de um processo iterativo de amostragem de uma cadeia de Markov. Para utilizar o método, precisamos conhecer as distribuições condicionais completas dos k parâmetros da distribuição conjunta, representadas por

$$\begin{aligned} \pi(\theta_1 \mid \theta_2, \theta_3, \dots, \theta_{k-1}, \theta_k), \\ \pi(\theta_2 \mid \theta_1, \theta_3, \dots, \theta_{k-1}, \theta_k), \\ \vdots \\ \pi(\theta_k \mid \theta_1, \theta_2, \theta_3, \dots, \theta_{k-1}) \end{aligned}$$

As etapas do algoritmo são:

1. Escolha arbitrariamente um vetor de valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$, e defina $i = 0$
2. Gere $\theta^{(i+1)} = (\theta_1^{(i+1)}, \dots, \theta_k^{(i+1)})$, como a seguir:
 - Gere $\theta_1^{(i+1)} \sim \pi(\theta_1 | \theta_2^{(i)}, \theta_3^{(i)}, \dots, \theta_k^{(i)}, D)$;
 - Gere $\theta_2^{(i+1)} \sim \pi(\theta_2 | \theta_1^{(i+1)}, \theta_3^{(i)}, \dots, \theta_k^{(i)}, D)$;
 - ⋮
 - Gere $\theta_k^{(i+1)} \sim \pi(\theta_k | \theta_1^{(i+1)}, \theta_2^{(i+1)}, \dots, \theta_{k-1}^{(i+1)}, D)$.
3. Defina $i=i+1$ e volte ao passo 2.

Assim cada componente do vetor $\boldsymbol{\theta}$ é visitado em sua ordem natural e um ciclo nesse esquema requer a geração de k amostras aleatórias. Gelfand & Smith (1990) mostrou que, sob certas condições de regularidade, a sequência $\{\theta^{(i)}, i = 1, 2, \dots\}$ tem distribuição estacionária $\pi(\boldsymbol{\theta} | D)$. Schervish & Carlin (1992) forneceram condição suficiente que garantiu a convergência geométrica da sequência.

4.1.2 Metropolis-Hastings

O algoritmo Metropolis-Hastings foi desenvolvido por Metropolis et al. (1953) e generalizado por Hastings (1970). Assim como o Amostrador de Gibbs, esse é um dos métodos mais importantes e difundidos dentre os algoritmos MCMC, e particular é utilizado quando as distribuições condicionais a posteriori são complexas, não possuem forma fechada, impossibilitando a geração direta a partir dessas distribuições.

Para aplicação do algoritmo é necessária a escolha de uma densidade de transição $q(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ que representará a regra de passagem que define a cadeia. Também define-se $U \sim \text{Uniforme}(0, 1)$.

Então, uma versão geral do algoritmo Metropolis-Hastings para amostragem da distribuição a posteriori $\pi(\boldsymbol{\theta} | D)$ pode ser descrita como a seguir:

1. Escolha arbitrariamente um ponto de partida $\boldsymbol{\theta}^{(0)}$ e defina $i = 0$.
2. Gere um valor candidato $\boldsymbol{\theta}^*$ a partir de $q(\boldsymbol{\theta}^{(i)}, \cdot)$ e $u \sim U(0, 1)$.

3. Defina $\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^*$ se $u \leq \alpha(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^*)$ ou $\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)}$ caso contrário, em que a probabilidade de aceitação é definida por:

$$\alpha(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^*) = \min \left\{ \frac{\pi(\boldsymbol{\theta}^*|D) q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(i)})}{\pi(\boldsymbol{\theta}^{(i)}|D) q(\boldsymbol{\theta}^{(i)}, \boldsymbol{\theta}^*)}, 1 \right\}$$

4. Defina $i = i + 1$, e volte para o passo 2.

Repita o procedimento acima até atingir a convergência da cadeia. Geralmente é escolhido um número N de iterações e depois verificada se a cadeia (amostra de $\pi(\boldsymbol{\theta}|D)$) gerada é estacionária.

4.2 Metodologia Bayesiana versus Frequentista

Uma pergunta natural sobre a metodologia bayesiana é o que ela oferece de vantagem em relação aos métodos frequentistas na análise de sobrevivência. Existem diversas vantagens como:

- Modelos de sobrevivência são em geral difíceis de se ajustar, especialmente na presença de esquemas complexos de censura. Utilizando o Amostrador de Gibbs ou qualquer outra técnica MCMC, ajustar modelos de sobrevivência complexos se torna uma tarefa bem mais simples. Além disso, simulações MCMC nos permitem realizar inferências para qualquer tamanho de amostra sem precisar recorrer à cálculos assintóticos. No paradigma frequentista, estimativas de variância, por exemplo, usualmente requerem argumentos assintóticos o que pode ser bastante complicado ou mesmo não possível, ou seja, necessita-se que o tamanho da amostra seja suficiente para que a aproximação assintótica seja válida. Enquanto isso, na metodologia bayesiana basta obter a distribuição a posteriori seja analiticamente ou por simulação.
- O método bayesiano permite incorporar informação a priori ao modelo, por exemplo, no caso de dados históricos. Para muitos modelos a inferência clássica pode ser considerada como um caso especial da inferência bayesiana com prioris não informativas, como uma priori uniforme. Nesse caso a posteriori será a estimativa de máxima verossimilhança.

- A análise bayesiana fornece ferramentas simples para comparação de ajuste entre modelos via BIC (*bayesian information criterion*) ou critério de seleção do modelo obtido pelo amostrador de Gibbs. Na estatística frequentista não existe um método unificado para comparação entre modelos e muitas vezes são requeridos argumentos assintóticos.
- Variáveis com não resposta ou valores *missings* são tratados como parâmetros na inferência bayesiana e resultam apenas em pequenas mudanças nos algoritmos computacionais, enquanto na metodologia clássica exigem métodos computacionais mais sofisticados, o que pode se tornar algo complicado.

5. Simulação

5.1 Simulação dos tempos de falha dependentes

Trabalhar com dados simulados tem grande importância no cenário científico atual, em que dispomos de equipamentos e softwares robustos para realizar diversos tipos de análises, nos permitindo estudar novos modelos matemáticos ou técnicas estatísticas, antes de aplicá-los a dados reais. Também nos permite comparar novos métodos com os já existentes, quanto às suas propriedades assintóticas ou qualidade do ajuste.

A partir de amostras univariadas uniformes independentes pode-se obter amostras de uma distribuição univariada qualquer, através de vários procedimentos. Um desses métodos é conhecido como método da função de distribuição inversa ou método da transformação integral de probabilidade. Para se obter uma observação t de uma variável aleatória T com função de distribuição F , gere uma variável $U \sim U(0, 1)$ e faça $t = F^{(-1)}(u)$, em que $F^{(-1)}$ é a função quasi-inversa de F .

Nosso objetivo é gerar uma amostra (t_1, t_2) de um vetor aleatório dos tempos de falha (T_1, T_2) com distribuição conjunta H , em que T_1 e T_2 possuem distribuições marginais F e G respectivamente. Poderíamos simular um par (t_1, t_2) a partir das variáveis aleatórias $U \sim U(0, 1)$ e $V \sim U(0, 1)$ e realizar as seguintes transformações $t_1 = F^{(-1)}(u)$ e $t_2 = F^{(-1)}(v)$, porém não teríamos, caso exista, a representação da estrutura de dependência entre os tempos (t_1, t_2) .

Então para simular amostras do vetor (T_1, T_2) com dependência entre os tempos de falha, utilizaremos o método da distribuição condicional por cópulas. Pelo teorema de Sklar, temos que o par (T_1, T_2) de variáveis aleatórias com função de distribuição conjunta $H(t_1, t_2)$ possui uma cópula associada $C(u, v)$, em que $u = F(t_1)$ e $v = G(t_2)$.

Para simular a amostra (t_1, t_2) pelo método da distribuição condicional, pre-

precisamos gerar primeiramente o par (u, v) de observações uniformes $(0, 1)$ das variáveis aleatórias (U, V) com distribuição conjunta C . Para tal necessitamos da distribuição condicional de V dado $U = u$, denotada por $F_{V|U}(v|u)$:

$$\begin{aligned}
F_{V|U}(v|u) &= P(V \leq v | U = u) = \int_0^v f(x|u)dx \\
&= \int_0^v \frac{f(u, x)}{f(u)}dx = \int_0^v f(u, x)dx \\
&= \int_0^v \frac{\partial^2 C(u, x)}{\partial x \partial u} dx = \frac{\partial C(u, v)}{\partial u}
\end{aligned} \tag{5.1}$$

Então o seguinte algoritmo é utilizado para gerar os tempos de falha (t_1, t_2) :

1. Gere $u \sim U(0, 1)$ e $x \sim U(0, 1)$ independentes;
2. Faça $v = F_{X|U}^{(-1)}(x|u) = \frac{\partial C^{(-1)}(u, x)}{\partial u}$;
3. Faça $t_1 = F_U^{(-1)}(u)$ e $t_2 = G_V^{(-1)}(v)$.

O par (t_1, t_2) possui função de distribuição conjunta H unicamente determinada pela cópula C_α , com parâmetro de dependência α .

O procedimento descrito gera tempos exatos (T_1, T_2) com distribuições marginais F e G respectivamente, e distribuição conjunta H . Porém o interesse está em como trabalhar com dados censurados, então também iremos gerar observações das variáveis aleatórias dos tempos de observação L_1, U_1, L_2 e U_2 , que são independentes de T_1 e T_2 . A partir das variáveis T_1, L_1, U_1, T_2, L_2 e U_2 obtemos as variáveis indicadoras de censura $\delta_1, \gamma_1, \delta_2$ e γ_2 . Para o desenvolvimento deste trabalho serão utilizadas as amostras simuladas compostas pelas variáveis $(L_{1i}, U_{1i}, \delta_{1i}, \gamma_{1i})$ e $(L_{2i}, U_{2i}, \delta_{2i}, \gamma_{2i})$.

As cópulas C_α utilizadas para geração das amostras do estudo são as cópula de Clayton, Frank e Gumbel.

$$C(u, v; \theta) = -\frac{1}{\theta} \cdot \log \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} \right]$$

Derivando $C(u, v; \theta)$ em relação a variável u , obtemos:

$$\frac{\partial C(u, v)}{\partial u} = \frac{(e^{-\theta v} - 1)e^{-\theta u}}{(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)}$$

Finalmente calculamos $F_{X|U}^{(-1)}(x|u)$:

$$\begin{aligned} \frac{\partial C^{(-1)}(u, x)}{\partial u} &= \frac{(e^{-\theta v} - 1)e^{-\theta u}}{(e^{-\theta} - 1) + (e^{-\theta u} - 1)(e^{-\theta v} - 1)} = x \\ x(e^{-\theta} - 1) + x(e^{-\theta u} - 1)(e^{-\theta v} - 1) &= (e^{-\theta v} - 1)e^{-\theta u} \\ xe^{\theta u}(e^{-\theta} - 1) + xe^{\theta u}(e^{-\theta u} - 1)(e^{-\theta v} - 1) &= (e^{-\theta v} - 1) \\ xe^{\theta u}(e^{-\theta u} - 1)(e^{-\theta v} - 1) - (e^{-\theta v} - 1) &= -xe^{\theta u}(e^{-\theta} - 1) \\ (e^{-\theta v} - 1) [xe^{\theta u}(e^{-\theta u} - 1) - 1] &= -xe^{\theta u}(e^{-\theta} - 1) \\ (e^{-\theta v} - 1) &= \frac{-xe^{\theta u}(e^{-\theta} - 1)}{-(x(e^{\theta u} - 1) + 1)} \\ v &= -\frac{1}{\theta} \log \left[1 + \frac{xe^{\theta u}(e^{-\theta} - 1)}{1 + x(e^{\theta u} - 1)} \right] \end{aligned}$$

Utilizaremos a distribuição Weibull F e G , e com isso temos:

$$F_U^{(-1)}(u) = \lambda[-\log(1 - u)]^{\frac{1}{\beta}} \quad \text{e} \quad G_V^{(-1)}(v) = \lambda[-\log(1 - v)]^{\frac{1}{\beta}}$$

5.1.1 Estimação MCMC em duas etapas

Após a geração dos tempos de falha dependentes utilizando o procedimento descrito acima, foram simulados de forma independente, observações das variáveis aleatórias dos tempos de observação L_1 , U_1 , L_2 e U_2 . Primeiro foi realizada a estimação Bayesiana dos parâmetros das distribuições marginais, utilizando o algoritmo Metropolis-Hastings.

Como não existe priori conjunta conjugada ao assumirmos os parâmetros β e λ desconhecidos. Consideraremos β e λ independentes, em que β tem distribuição gama e λ tem distribuição log-normal.

No primeiro estágio, temos as prioris para os vetores de parâmetros θ_1 e θ_2 dadas por $\pi_j^\theta(\cdot)$, $j = 1, 2$, e a função de máxima verossimilhança

$$L(\mathbf{L}_j, \mathbf{U}_j, \boldsymbol{\delta}_j, \boldsymbol{\gamma}_j | \boldsymbol{\theta}_j) = \prod_{i=1}^n \{ [F(L_{ij})]^{\delta_{ij}} [F(U_{ij}) - F(L_{ij})]^{\gamma_{ij}} \times [1 - F(U_{ij})]^{1 - \delta_{ij} - \gamma_{ij}} \}, \quad j = 1, 2 \quad (5.2)$$

e a distribuição a posteriori fica,

$$\pi(\boldsymbol{\theta}_j | \mathbf{L}_j, \mathbf{U}_j, \boldsymbol{\delta}_j, \boldsymbol{\gamma}_j) \propto L(\mathbf{L}_j, \mathbf{U}_j, \boldsymbol{\delta}_j, \boldsymbol{\gamma}_j | \boldsymbol{\theta}_j) \pi_j^\theta(\boldsymbol{\theta}_j) \quad (5.3)$$

Se $\tilde{\boldsymbol{\theta}}_j$ é o estimador de $\boldsymbol{\theta}_j$ então na segunda etapa, o estimador do parâmetro de associação α é obtido pela posteriori

$$\pi(\alpha | \mathbf{L}_1, \mathbf{U}_1, \mathbf{L}_2, \mathbf{U}_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \tilde{\boldsymbol{\theta}}_1, \tilde{\boldsymbol{\theta}}_2) \propto L(\mathbf{L}_1, \mathbf{U}_1, \mathbf{L}_2, \mathbf{U}_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \tilde{\boldsymbol{\theta}}_1, \tilde{\boldsymbol{\theta}}_2 | \alpha) \pi^\alpha(\alpha) \quad (5.4)$$

em que $\pi^\alpha(\alpha)$ é a priori do parâmetro de associação.

No caso bivariado a verossimilhança será

$$L(H) = \prod_{i=1}^N [H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i})] \quad (5.5)$$

com a função de distribuição conjunta $H(t_1, t_2)$ para o caso da cópula de Frank

$$H(t_1, t_2) = -\frac{1}{\alpha} \log \left[1 + \frac{\left(e^{-\alpha \left[1 - e^{-\left(\frac{t_1}{\lambda_1} \right)^{\beta_1}} \right]} - 1 \right) \left(e^{-\alpha \left[1 - e^{-\left(\frac{t_2}{\lambda_2} \right)^{\beta_2}} \right]} - 1 \right)}{(e^{-\alpha} - 1)} \right] \quad (5.6)$$

6. Resultados e aplicação à dados reais

6.1 Simulações

Nos capítulos anteriores foram expostos os métodos utilizados para estimação dos parâmetros das distribuições marginais com censura intervalar e para estimação do parâmetro de associação entre essas distribuições, a partir de um modelo de cópula. Nesse capítulo serão descritos os parâmetros utilizados nas simulações e os seus resultados.

Um dos objetivos desse estudo de simulação é verificar se o algoritmo consegue recuperar o valor real dos parâmetro. Ao lidar com dados simulados, quando sabemos o valor real de cada parâmetro, podemos realizar tal verificação. Além disso também iremos verificar como se comportam esses estimadores com relação a sua variabilidade e consistência.

Foram geradas amostras com distribuição Weibull e distribuição conjunta determinada por três modelos de cópulas: Clayton, Gumbel e Frank. Para cada cópula foram considerados três valores para o parâmetro de associação $\alpha = 1, 2 e 5$ e três tamanhos de amostras diferentes $n = 50, 100 e 200$. Para os parâmetros de ambas marginais foram adotados os valores $\lambda = 5$ (parâmetro de escala) e $beta = 2, 5$ (parâmetro de forma). Esse esquema resulta em 9 casos de simulação para cada modelo de cópula utilizado. Além de trabalhar com o modelo de cópula certo para ajustar os valores dos parâmetros em cada caso, também ajustaremos os dados simulados utilizando os outros dois modelos de cópula, com isso podemos ter uma ideia dos possíveis problemas que a escolha de um modelo de cópula errado pode acarretar nas estimativas.

Então no total desse estudo, teremos 81 casos de simulação. Para cada caso foram geradas 200 amostras, totalizando 16.200 simulações.

Como estamos trabalhando com tempos de sobrevivência na presença de censuras, precisamos também simular os instantes de observação do fenômeno de interesse. E como pressuposto, esses tempos de observação devem ser independentes dos tempos de falha. No algoritmo esses instantes de observação são gerados a partir de uma distribuição uniforme. Para o primeiro instante são considerados como limites da distribuição uniforme os quantis 0,2 e 0,4 da distribuição Weibull utilizada para simular a amostra, já o segundo instante utiliza como limites os quantis 0,6 e 0,8. Ao realizar esse procedimento buscou-se manter um equilíbrio na quantidade de tipos de censura (esquerda, intervalar e direita) para os dados simulados. Todo o restante da metodologia utiliza essas amostras de instantes de observação para a estimação dos parâmetros.

Para a estimação dos parâmetros em cada amostra gerada, foram simuladas a partir da distribuição a posteriori, utilizando o algoritmo Metropolis-Hastings, uma cadeia com 22 mil elementos. Então, são descartadas os primeiros 2 mil elementos, procedimento conhecido como *burn-in*, e são selecionados 1 a cada 20 dentre os 20 mil elementos restantes, esse método visa minimizar a correlação entre os elementos da cadeia (tornar a série estacionária), resultando em uma amostra pseudo aleatória de tamanho $n = 1000$. A média dessa amostra é a estimativa do parâmetro de interesse. Foram consideradas as seguintes *prioris* vagas para os parâmetros das marginais: $Gama(1, 0.001)$ para o parâmetro de forma β e $Lognormal(0, 100)$ para o parâmetro de escala λ . Para os parâmetros de associação do modelo de cópulas foram utilizadas as *prioris* $Gama(1, 0.001)$ para as cópulas de Clayton e Frank e $Normal(0, 100)$ para a cópula de Frank.

As tabelas a 6.1, 6.2 e 6.3 contém o valor esperado e o desvio padrão das estimativas dos parâmetros das distribuições marginais. O modelo exposto nesse trabalho possui duas etapas independentes entre si, ou seja, não há associação entre as marginais e o modelo cópula adotado. Sendo assim, independente do modelo de cópula utilizado, não se espera diferenças nas estimativas dos parâmetros das marginais, fato que pode ser comprovado observando essas tabelas.

Tabela 6.1: Média e [desvio padrão] dos parâmetros das marginais - Clayton

n	$\theta_j, j = 1, 2$	$\alpha = 1$		$\alpha = 2$		$\alpha = 5$	
50	$\lambda_1 = 5$	5,15	[0,17]	5,08	[0,15]	5,10	[0,13]
	$\beta_1 = 2, 5$	2,57	[0,25]	2,38	[0,19]	2,50	[0,21]
	$\lambda_2 = 5$	5,12	[0,15]	5,15	[0,14]	5,05	[0,13]
	$\beta_2 = 2, 5$	2,63	[0,31]	2,47	[0,24]	2,50	[0,29]
100	$\lambda_1 = 5$	5,05	[0,07]	5,03	[0,07]	5,08	[0,07]
	$\beta_1 = 2, 5$	2,63	[0,15]	2,56	[0,11]	2,59	[0,11]
	$\lambda_2 = 5$	5,11	[0,07]	5,06	[0,06]	5,04	[0,06]
	$\beta_2 = 2, 5$	2,48	[0,14]	2,55	[0,10]	2,63	[0,12]
200	$\lambda_1 = 5$	5,01	[0,03]	5,02	[0,05]	5,02	[0,04]
	$\beta_1 = 2, 5$	2,49	[0,04]	2,55	[0,07]	2,50	[0,05]
	$\lambda_2 = 5$	5,07	[0,02]	4,99	[0,04]	5,03	[0,03]
	$\beta_2 = 2, 5$	2,57	[0,06]	2,56	[0,09]	2,51	[0,06]

Tabela 6.2: Média e [desvio padrão] dos parâmetros das marginais - Gumbel

n	$\theta_j, j = 1, 2$	$\alpha = 1$		$\alpha = 2$		$\alpha = 5$	
50	$\lambda_1 = 5$	5,10	[0,11]	5,04	[0,18]	5,07	[0,19]
	$\beta_1 = 2, 5$	2,73	[0,30]	2,60	[0,20]	2,67	[0,47]
	$\lambda_2 = 5$	5,08	[0,15]	5,20	[0,26]	5,16	[0,23]
	$\beta_2 = 2, 5$	2,59	[0,19]	2,47	[0,32]	2,62	[0,22]
100	$\lambda_1 = 5$	5,09	[0,10]	5,05	[0,07]	5,02	[0,08]
	$\beta_1 = 2, 5$	2,49	[0,12]	2,50	[0,14]	2,59	[0,09]
	$\lambda_2 = 5$	5,05	[0,06]	5,06	[0,06]	5,05	[0,09]
	$\beta_2 = 2, 5$	2,58	[0,10]	2,44	[0,13]	2,56	[0,10]
200	$\lambda_1 = 5$	4,99	[0,03]	5,07	[0,04]	5,02	[0,04]
	$\beta_1 = 2, 5$	2,55	[0,05]	2,49	[0,06]	2,53	[0,08]
	$\lambda_2 = 5$	5,02	[0,02]	5,11	[0,04]	5,01	[0,04]
	$\beta_2 = 2, 5$	2,53	[0,05]	2,47	[0,06]	2,56	[0,09]

Tabela 6.3: Média e [desvio padrão] dos parâmetros das marginais - Frank

n	$\theta_j, j = 1, 2$	$\alpha = 1$		$\alpha = 2$		$\alpha = 5$	
50	$\lambda_1 = 5$	5,01	[0,16]	5,16	[0,11]	5,13	[0,21]
	$\beta_1 = 2, 5$	2,56	[0,30]	2,53	[0,26]	2,54	[0,28]
	$\lambda_2 = 5$	5,00	[0,13]	5,22	[0,18]	5,12	[0,11]
	$\beta_2 = 2, 5$	2,58	[0,26]	2,52	[0,29]	2,57	[0,20]
100	$\lambda_1 = 5$	5,15	[0,07]	5,07	[0,07]	5,07	[0,07]
	$\beta_1 = 2, 5$	2,55	[0,17]	2,56	[0,10]	2,49	[0,11]
	$\lambda_2 = 5$	5,00	[0,06]	5,05	[0,08]	5,04	[0,06]
	$\beta_2 = 2, 5$	2,52	[0,12]	2,53	[0,13]	2,61	[0,11]
200	$\lambda_1 = 5$	5,03	[0,03]	5,05	[0,05]	5,03	[0,03]
	$\beta_1 = 2, 5$	2,47	[0,04]	2,53	[0,07]	2,46	[0,10]
	$\lambda_2 = 5$	5,00	[0,03]	5,00	[0,03]	4,97	[0,04]
	$\beta_2 = 2, 5$	2,44	[0,05]	2,50	[0,06]	2,47	[0,07]

Como pôde ser observado, em todos os casos o método apresentou bons resultados quanto ao valor das estimativas, além de ter seu vício e variância reduzidos para amostras maiores.

Os resultados das simulações para a estimação do parâmetro de associação α estão disponíveis nas tabelas 6.4 e 6.5. Em algumas situações o método apresentou estimativas significativamente distantes do valor real do parâmetro, esse problema ocorreu quando os dados foram gerados a partir da Cópula de Gumbel com $\alpha = 5$, para esse valor o coeficiente de dependência τ de *Kendall* é $\tau = 0,8$, e representa forte associação entre as variáveis. Esse problema ocorreu em cerca de 18% das estimativas, e estas foram descartadas.

Tabela 6.4: Média do parâmetro de associação, segundo o modelo Cópula

Cópula geradora	Cópula Ajustada									
	Clayton			Gumbel			Frank			
	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	
Clayton	$n = 50$	1,17	2,72	7,16	1,35	2,31	8,21	3,67	6,35	26,33
	$n = 100$	1,22	2,08	6,05	1,09	2,27	6,12	3,31	5,96	15,29
	$n = 200$	1,03	1,94	5,20	1,09	2,04	5,42	3,14	5,78	12,95
Gumbel	$n = 50$	1,13	2,14	12,65	1,13	2,51	9,01	0,92	6,00	36,26
	$n = 100$	1,06	2,05	9,88	1,08	2,17	6,82	0,64	6,03	31,61
	$n = 200$	1,03	1,89	8,73	1,06	2,13	5,62	0,43	5,52	24,35
Frank	$n = 50$	0,42	0,69	1,98	1,19	1,36	2,13	1,44	2,12	5,70
	$n = 100$	0,27	0,60	1,86	1,17	1,29	2,01	1,33	2,26	5,33
	$n = 200$	0,26	0,56	1,71	1,15	1,28	1,97	1,22	2,07	5,26

Tabela 6.5: Desvio Padrão do parâmetro de associação, segundo o modelo Cópula

Cópula geradora	Cópula Ajustada									
	Clayton			Gumbel			Frank			
	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 5$	
Clayton	$n = 50$	0,71	1,16	3,33	0,68	1,10	5,19	1,50	2,13	12,86
	$n = 100$	0,39	0,48	2,23	0,40	0,63	1,82	0,94	1,20	8,74
	$n = 200$	0,27	0,37	0,90	0,23	0,40	1,30	0,62	0,98	4,37
Gumbel	$n = 50$	0,06	0,89	7,38	0,07	0,64	5,34	0,49	1,53	6,36
	$n = 100$	0,01	0,75	5,30	0,04	0,29	2,75	0,30	1,53	10,45
	$n = 200$	0,00	0,48	2,56	0,02	0,19	1,17	0,22	0,71	10,55
Frank	$n = 50$	0,22	0,40	0,78	0,19	0,19	0,48	0,88	0,98	1,82
	$n = 100$	0,13	0,21	0,59	0,12	0,11	0,26	0,65	0,75	0,86
	$n = 200$	0,11	0,16	0,34	0,06	0,09	0,20	0,44	0,52	0,79

Algumas considerações podem ser feitas a partir dos resultados apresentados nas tabelas acima. Quando o modelo de cópula escolhido para ajustar os dados é o correto (o mesmo da cópula que gerou os dados), o método se mostra eficiente em retornar valores próximos ao do parâmetro real. Porém para valores de α elevados, o método apresenta alguns resultados imprecisos, como evidenciado para a cópula de Gumbel com $\alpha = 5$. Na prática valores elevados do parâmetro de associação não são tão comuns e portanto o método desenvolvido nesse trabalho pode ser uma boa ferramenta para o estudo de dados dessa natureza.

Todas as estimativas se mostram consistentes, a variação é reduzida com o aumento do tamanho da amostra, com exceção para os dados gerados pela cópula de Gumbel e ajustados com a cópula de Frank.

A escolha errada do modelo de cópula pode resultar em estimativas imprecisas, fato evidenciado pelo estudo de simulação realizado. Dentre os 3 modelos de cópulas considerados, apenas as cópulas de Clayton e Gumbel, quando utilizadas para ajustar os dados gerados pela outra cópula, apresentaram resultados próximos aos valores reais dos parâmetros, com ressalva para os valores altos de α . Apesar do bom ajuste no valor do parâmetro, vale ressaltar que a cópula de Gumbel apresenta dados com dependência na cauda superior, enquanto que na cópula de Clayton a dependência é na cauda inferior, abaixo segue a representação gráfica dessas dependências:

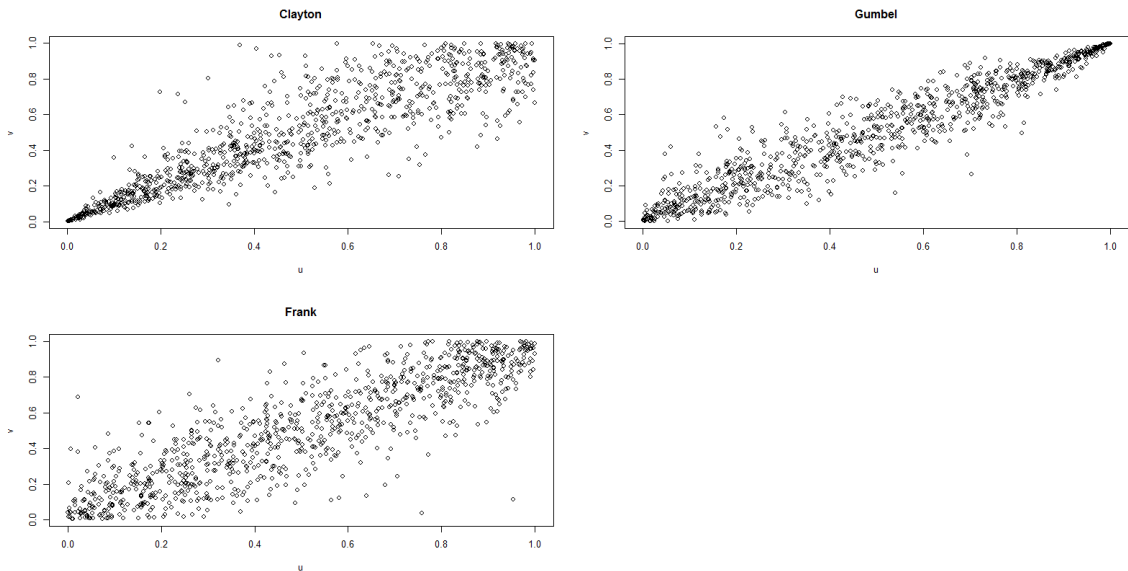


Figura 6.1: Dependência Caudal das cópulas para amostras $n = 1000$

6.2 Aplicação

Nesta seção iremos utilizar o algoritmo Bayesiano via cópulas, considerado neste trabalho, para estimar os parâmetros de interesse do conjunto de dados (Betensky, 1999) de um estudo observacional sobre AIDS, que foi conduzido pelo AIDS Clinical Trials Group (ACTG). A proposta do estudo ACTG 181 é determinar o comportamento natural de infecções causadas pelo cytomegalovirus (CMV) em pacientes portadores de HIV. Como as infecções causadas pelo CMV não estão associadas à sintomas nos pacientes, foram realizados exames clínicos de sangue e urina para detectar a presença desta infecção. Uma importante questão a ser respondida pelo estudo é acerca da relação entre o tempo de infecção no sangue e na urina. O conhecimento da distribuição conjunta dos tempos desses eventos é necessária para determinar a associação entre eles.

Os dados contém o tempo em meses de 204 pacientes até que a presença do CMV no sangue ou na urina fosse detectada em algum exame clínico, portanto temos as informações censuradas em ambos eventos de interesse do estudo.

O ajuste das marginais foi realizado utilizando um método não paramétrico e o método paramétrico Bayesiano, em que foram ajustados dois modelos Weibull para as marginais.

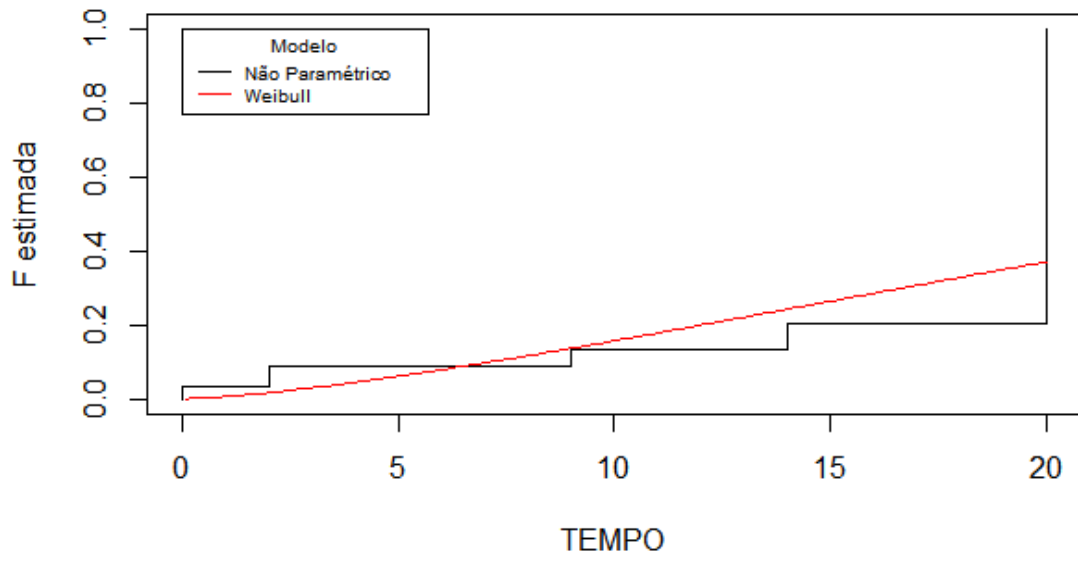


Figura 6.2: Ajuste não paramétrico e curva da Weibull ajustada - CMV Sangue

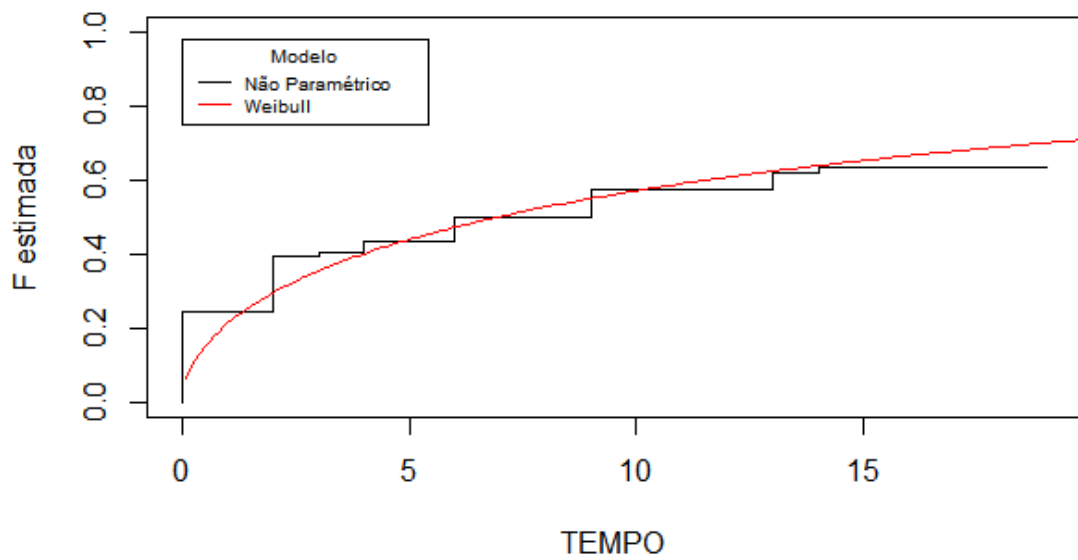


Figura 6.3: Ajuste não paramétrico e curva da Weibull ajustada - CMV Urina

O ajuste da Weibull nas marginais são bem próximas às curvas obtidas pelo método não paramétrico, de maneira geral o ajuste paramétrico é satisfatório para os dados do estudo. Abaixo seguem as estimativas dos parâmetros das marginais e das cópulas.

Tabela 6.6: Estimativas dos parâmetros das marginais

Parâmetro	CMV Urina	CMV Sangue
$\hat{\lambda}$	11,307	32,649
$\hat{\beta}$	0,625	1,465

Tabela 6.7: Estimativas de α

Cópula	$\hat{\alpha}$
Clayton	0,465
Gumbel	1,627
Frank	3,141

7. Conclusões e trabalhos futuros

7.1 Conclusões

Dados de sobrevivência bivariados com presença de censuras são comuns principalmente na área médica, em estudos longitudinais em que o evento de interesse pode ser observado em mais de um instante de tempo. Portanto a variável resposta poderá ter censura à esquerda, direita ou intervalar. O uso do modelo Bayesiano via cópulas se mostra uma ferramenta poderosa para se obter estimativas de interesse.

Para o estudo de simulação foi implementado o modelo bayesiano em duas etapas para dados com censura intervalar bivariado. Foram simuladas duas variáveis aleatórias com distribuição Weibull e seus respectivos instantes de observação, a dependência entre os tempos foi simulada via cópulas (Clayton, Gumbel e Frank).

A partir dos resultados do estudo de simulação, observou-se que o ajuste feito considerando a mesma cópula utilizada para simular os dados, apresentou resultados satisfatórios quanto ao vício e variância, em especial quando a dependência entre os tempos de falha não é muito alta. Porém quando trabalhamos com dados reais, a qualidade do ajuste ficará limitado a escolha adequada da cópula, podendo então levar a resultados equivocados.

Para os dados reais analisados neste trabalho, obtivemos bons resultados, em comparação com o método não paramétrico, para o ajuste das distribuições marginais modeladas segundo o modelo paramétrico Weibull. Ficamos limitados ainda a um método que permita a escolha correta do modelo de cópula para dados em que não conhecemos os parâmetros ou modelos reais. Essa limitação será trabalhada em estudos futuros. O algoritmo utilizado foi todo desenvolvido no software R e encontra-se no apêndice deste trabalho.

7.2 Trabalhos Futuros

Trabalhos futuros incluem estender o algoritmo a outros métodos de estimação das marginais, como por exemplo, um modelo não paramétrico suavizado, para dar maior liberdade quanto adequamento do ajuste dos dados. Além disso, incorporar outros modelos de cópula ao código, pois os tipos de dependência não se limitam apenas aos modelos de cópulas apresentados.

É essencial para uma escolha de um modelo cópula candidato ao ajuste dos dados, alguma solução gráfica para dados censurados que permita a visualização do tipo de dependência. Ressaltando que para esse tipo de dados não observamos pontos, e sim retângulos que contém o evento de interesse.

Estabelecer alguma medida que comprove a qualidade do ajuste do modelo cópula, podendo ser trabalhada em comparação com algum ajuste não paramétrico da distribuição conjunta dos dados.

Bibliografia

- [1] Betensky, R. A. & Finkelstein, D. M., 1999. “A Non-parametric maximum likelihood estimator for bivariate interval censored data”. Department of Biostatistics, Harvard School of Public Health, Boston.
- [2] Besag, J. & Green, P. J., 1993. “Spatial Statistics and Bayesian computation”. Journal of the Royal Statistical Society, Series B 55, 25-37.
- [3] Calle, M. L. & Rosingana, 1996. “The Analysis of Interval-Censored Survival Data. From a Nonparametric Perspective to a Nonparametric Bayesian Approach”. Dept. d’Estadística i Investigació Operativa Universitat Politècnica de Catalunya.
- [4] Colosimo, E. A. & Giolo, S. R., 2006. “Análise de Sobrevivência Aplicada.” Editora Blucher, São Paulo.
- [5] Eirado, M. P. R., 2010. “Estimação da função de distribuição conjunta para dados com censura intervalar bivariados”. Dissertação de mestrado, Departamento de Estatística UNB.
- [6] Flores, A. Q., 2008. “Copula functions and bivariate distributions for survival analysis: An application to political survival”. Wilf Department of Politics, New York University.
- [7] Gelfand, A. E. & Smith, A. F. M., 1990. “Sampling-based approaches to calculating marginal densities”. Journal of the American Statistical Association, 85, 398-409.

- [8] Geman, S. & Geman, D., 1984. "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.
- [9] Goggins, W. B. & Filkelstein, D. M., 2000. "A Proportional Hazards Model for Multivariate Interval-Censored Failure Time Data". *Biometrics* 56, 940-943.
- [10] Grenander, U., 1983. "Tutorial in pattern theory. Technical Report.". Providence, R.I.: Division of Applied Mathematics, Brown University.
- [11] Gustafson, P.; Aeschliman, D. & Levy, A. R., 2003. "A simple approach to fitting Bayesian survival models". *Lifetime Data Analysis*, 9, 5-19.
- [12] Guure, C. B.; Ibrahim, N. A. & Adam, M. B., 2013. "Bayesian Inference of the Weibull Model Based on Interval-Censored Survival Data". *Computational and Mathematical Methods in Medicine Volume 2013*.
- [13] Hougaard, P., 1989. "Fitting a multivariate failure time distribution". *IEEE Transactions on Reliability*, 38, 444-448.
- [14] Huard, D.; Évin, G. & Favre, A-C., 2006. "Bayesian copula selection". Institut National de la Recherche Scientifique, Centre Eau, Terre & Environnement, Qué., Canada.
- [15] Ibrahim, J. G.; Chen, M. & Sinha, D., 2001. "Bayesian Survival Analysis". Springer-Verlag, New York.
- [16] Kelly, D. L., 2007. "Using Copulas to Model Dependence in Simulation Risk Assessment". ASME International Mechanical Engineering Congress and Exposition.
- [17] Lambert, P., 2005. "Archimedean copula estimation using Bayesian splines smoothing techniques". Institut de Statistique, Université catholique de Louvain, Louvain-la-Neuve, Belgium.
- [18] Lee, E-J.; Kim, C-H. & Lee, S-H., 2011. "Life Expectancy Estimate With Bivariate Weibull Distribution Using Archimedean Copula". *International Journal of Biometrics and Bioinformatics (IJBB)*, Volume (5) : Issue (3).

- [19] Maathuis, M. H., 2003. “Nonparametric maximum likelihood estimation for bivariate censored data”. Master’s thesis, Delft University of Technology.
- [20] Maathuis, M. H., 2005. “Reduction algorithm for the NPMLE for the distribution function of bivariate interval censored data”. *Journal of Computational and Graphical Statistics*.
- [21] McNeil, A. J.; Frey, R. & Embrechts, P., 2005. “Quantitative Risk Management - Concepts, Techniques and Tools”. Princeton University Press.
- [22] Metropolis, N.; Rosenbluth, A. H.; Rosenbluth, M. N.; Teller, A. H. & Teller, E., 1953. “Equations of state calculations by fast computing machine”. *Journal of Chemical Physics*, 21, 1087-1091.
- [23] Nelsen, R. B., 2006. “An Introduction to Copulas”. Springer-Verlag, New York.
- [24] Nicoloutsopoulos, D., 2005. “Parametric and Bayesian non-parametric estimation of copulas”. University College London.
- [25] Oakes, D., 1989. “Bivariate survival models induced by frailties”. *Journal of the American Statistical Society Series B*, 44, 414-422.
- [26] Romeo, J. S.; Tanaka, N. I. & Pedroso-de-Lima, A. C., 2006. “Bivariate survival modeling: a Bayesian approach based on Copulas”. *Lifetime Data Analysis*.
- [27] Santos, C. A., 2010. “Dados de sobrevivência multivariados na presença de covariáveis e observações censuradas: uma abordagem bayesiana”. Tese de doutorado, Universidade Federal de São Carlos - DEs/UFSCar.
- [28] Schervish, M. J. & Carlin, B. P., 1992. “On the convergence of successive substitution sampling”. *Journal of Computational and Graphical Statistics* 1, 111-127.
- [29] Shih, J. H. & Louis, T. A., 1995. “Inferences on the association parameter in copula models for bivariate survival data”. *Biometrics*, 51, 1384-1399.
- [30] Sparling, Y. H.; Younes, N. & Lachin, J. M., 2006. “Parametric survival models for interval-censored data with time-dependent covariates”. *Biostatistics*, 7, 4, pp. 599-614.

- [31] Sun, L.; Wang, L. & Sun, J., 2006. “Estimation of the Association for Bivariate Interval-censored Failure Time Data”. Board of the Foundation of the Scandinavian Journal of Statistics.
- [32] Suzuki, A. K., 2012. “Modelos de sobrevivência bivariados baseados na cópula FGM: uma abordagem bayesiana”. Tese de doutorado, Universidade Federal de São Carlos - DEs/UFSCar.
- [33] Weibull, W., 1951. “A statistical distribution function of wide applicability”. Journal of Applied Mechanics, 292-297.

Apêndice 1

```
#####
## ESTIMAÇÃO BAYESIANA PARA DADOS COM CENSURA INTERVALAR BIVARIADOS ##
#####

require(QRM) #pacote Cópulas

#set.seed(2345)

### n = tamanho da amostra ###
n = 1000

### Parâmetro de associação alfa (cópula) ###
alfa = 4

### beta = valor real do parametro beta (forma) ###
beta1 = 2.5
beta2 = 2.5

### lambda = valor real do parametro lambda (escala) ###
lambda1 = 5
lambda2 = 5

#####
# Geração de distribuições marginais de t1 e t2 Weibull com dependência (Cópula de Frank) #
#####

# Formulação matematica para obtenção de u e v
#u = runif(n,0,1)
#x = runif(n,0,1)
#v = -(1/alfa)*log(1+((x*exp(alfa*u))*(exp(-alfa)-1))/(1+x*(exp(alfa*u)-1)))

# Gerando u e v diretamente pelo pacote QRM
r1 = rAC("frank", n = n, d = 2, theta = alfa)

#r1 = rAC("clayton", n = n, d = 2, theta = alfa)

#r1 = rAC("gumbel", n = n, d = 2, theta = alfa)

u = r1[,1]
v = r1[,2]

t1 = qweibull(u,beta1,lambda1)
t2 = qweibull(v,beta2,lambda2)

## Geração dos instantes de observação da variável t1 ##
inf1 <- runif(n,qweibull(0.2,beta1,lambda1),qweibull(0.4,beta1,lambda1))
sup1 <- runif(n,qweibull(0.6,beta1,lambda1),qweibull(0.8,beta1,lambda1))
```

```

## Geração dos instantes de observação da variável t2 ##
inf2 <- runif(n,qweibull(0.2,beta2,lambda2),qweibull(0.4,beta2,lambda2))
sup2 <- runif(n,qweibull(0.6,beta2,lambda2),qweibull(0.8,beta2,lambda2))

### delta = indicadora de t < l ###
delta1 <- 1*(t1<=inf1)
delta2 <- 1*(t2<=inf2)

## gama = indicadora de l <= t < u ###
gama1 <- (t1>inf1)*(t1<=sup1)
gama2 <- (t2>inf2)*(t2<=sup2)

s1 <- sup1
i1 <- inf1
s1 <- inf1*(delta1==1)*(gama1==0)+999*(delta1==0)*(gama1==0)+sup1*(gama1==1)*(delta1==0)
i1 <- 0*(delta1==1)*(gama1==0)+sup1*(delta1==0)*(gama1==0)+inf1*(gama1==1)*(delta1==0)

inf1 <- i1
sup1 <- s1

s2 <- sup2
i2 <- inf2
s2 <- inf2*(delta2==1)*(gama2==0)+999*(delta2==0)*(gama2==0)+sup2*(gama2==1)*(delta2==0)
i2 <- 0*(delta2==1)*(gama2==0)+sup2*(delta2==0)*(gama2==0)+inf2*(gama2==1)*(delta2==0)

inf2 <- i2
sup2 <- s2

#####
#cbind(t1,inf1,sup1,delta1,gama1)
#cbind(t2,inf2,sup2,delta2,gama2)

#table(delta1,gama1)
#table(delta2,gama2)

#####
## Cálculo da função de distribuição F no ponto t ##
#####

Fd = function(t,forma,escala){pweibull(t,forma,escala)}

#####
## Funcao "lik" calcula a verossimilhanca ##
#####

lik = function(l,u,delta,gama,forma,escala){

  ep = 1e-6

  lik = prod(((Fd(u,forma,escala)-Fd(l,forma,escala)+ep)^delta)*
    ((Fd(u,forma,escala)-Fd(l,forma,escala)+ep)^gama)*
    ((Fd(u,forma,escala)-Fd(l,forma,escala)+ep)^(1-delta-gama)))

  return(lik)

}

#####
## MCMC ##
#####

#####
## Função qtr calcula o valor do kernel de transição (comum para lambda e beta) ##
#####

```

```

# phi = valor atual do parâmetro
# phi.line = novo valor do parâmetro

qtr <- function(phi,phi.line,arg) {
  dgamma(phi,arg*phi.line,arg)
}

#####
### Função pri.lambda calcula a densidade de lambda ###
#####
pri.lambda = function(lambda,mu,sigma) {
  dlnorm(lambda,mu,sigma)
}

#####
### Função "pri.beta" calcula a densidade de beta ###
#####
pri.beta = function(beta,forma,escala) {
  dgamma(beta,forma,escala)
}

#####
### Função "post" calcula a densidade a posteriori ###
#####
#pb.a & pb.b: parâmetro de forma e escala respectivamente da priori de beta
#pl.a & pl.b: parâmetro média e desvio respectivamente da priori de lambda

post = function(l,u,delta,gama,forma,escala,pb.a,pb.b,pl.a,pl.b){
  post = lik(l,u,delta,gama,forma,escala)*pri.beta(forma,pb.a,pb.b)
  pri.lambda(escala,pl.a,pl.b)
  return(post)
}

#####
### Algoritmo Metropolis-Hastings MH ###
#####

lambda.MH = beta.MH = numeric(0)
mu = 0 #média da priori de lambda
sd = 100 #desvio da priori de lambda
bf = 1 #parâmetro de forma da priori de beta
be = 0.001 #parâmetro de escala da priori de beta

b = 10
j = 0
lambdaj = 2
betaj = 1

MH1 = function(l,u,delta,gama){
while (j < 22000){
  j = j+1
  #print(j)

  lambda.star = rgamma(1,b*lambdaj,b)

  aux.l = (post(l,u,delta,gama,betaj,lambda.star,bf,be,mu,sd)*qtr(lambdaj,lambda.star,b))/
    (post(l,u,delta,gama,betaj,lambdaj,bf,be,mu,sd)*qtr(lambda.star,lambdaj,b))

  p.aceit = min(c(1,aux.l))
  z = runif(1,0,1)
  if(z<=p.aceit) lambdaj = lambda.star else {lambdaj = lambdaj}
  lambda.MH = c(lambda.MH,lambdaj)
}
}

```

```

beta.star = rgamma(1,b*betaj,b)

aux.b = (post(l,u,delta,gama,beta.star,lambdaj,bf,be,mu,sd)*qtr(betaj,beta.star,b))/
        (post(l,u,delta,gama,betaj,lambdaj,bf,be,mu,sd)*qtr(beta.star,betaj,b))

p.aceit = min(c(1,aux.b))
z = runif(1,0,1)
if(z<=p.aceit) betaj = beta.star else {betaj = betaj}
beta.MH = c(beta.MH,betaj)

#print(c(lambdaj,betaj))

}
return(cbind(lambda.MH,beta.MH))
}

```

```

ind = seq(1,20000,by=20) #índices que farão parte da amostra
MH1.t1 = MH1(inf1,sup1,delta1,gama1)
MH1.t1 = MH1.t1[2001:22000,1:2] #Burn-in 2000
MH1.t1 = MH1.t1[ind,1:2] #Seleção de 1 a cada 20

```

```

MH1.t2 = MH1(inf2,sup2,delta2,gama2)
MH1.t2 = MH1.t2[2001:22000,1:2] #Burn-in 2000
MH1.t2 = MH1.t2[ind,1:2] #Seleção de 1 a cada 20

```

```

lambda1.MH = mean(MH1.t1[,1])
beta1.MH = mean(MH1.t1[,2])
lambda2.MH = mean(MH1.t2[,1])
beta2.MH = mean(MH1.t2[,2])

```

```

c(lambda1.MH,beta1.MH)
c(lambda2.MH,beta2.MH)

```

```

#####
### Estimação Bayesiana Conjunta de H(t1,t2) ###
#####

#####
### Funcao "H" calcula a densidade conjunta de t1 e t2 ###
#####

```

```

### transformação de volta para marginais U(0,1)

```

```

lower1 <- pweibull(inf1,beta1.MH,lambda1.MH)
upper1 <- pweibull(sup1,beta1.MH,lambda1.MH)

```

```

lower2 <- pweibull(inf2,beta2.MH,lambda2.MH)
upper2 <- pweibull(sup2,beta2.MH,lambda2.MH)

```

```

##### Cópula de Clayton #####

```

```

#H <- function(alfa,tempo1,tempo2) {
  #a = tempo1^(-alfa)
  #b = tempo2^(-alfa)

  #cop = (a+b-1)^(-1/alfa)

```

```

  #return(cop)
#}

```

```

#####

```

```

##### Cópula de Frank #####

```

```

#H <- function(alfa,tempo1,tempo2) {
  #a = (exp(-alfa*tempo1)-1)
  #b = (exp(-alfa*tempo2)-1)
  #c = (exp(-alfa)-1)

  #cop = -(1/alfa)*log(1+(a*b/c))

  #return(cop)
#}
#####

##### Cópula de Gumbel #####
H <- function(alfa,tempo1,tempo2) {
a = (-log(tempo1))^alfa
b = (-log(tempo2))^alfa

cop = exp(-(a+b)^(1/alfa))

return(cop)
}
#####

##### Funcao "lik.biv" calcula a verossimilhanca para o caso Bivariado ###
#####

lik.biv = function(alfa){

  lik.biv = prod(H(alfa,upper1,upper2)-H(alfa,lower1,upper2)-H(alfa,upper1,lower2)+H(alfa,lower1,lower2))

  return(lik.biv)

}

#####

##### Funcao "pri.alfa" calcula a densidade do parâmetro de associação alfa ###
#####

##### cópula de Clayton #####
pri.alfa = function(alfa,forma,escala) {
dgamma(alfa,forma,escala)
}
#####

##### cópula de Frank #####
#pri.alfa = function(alfa,mu,sd) {
#dnorm(alfa,mu,sd)
#}
#####

##### cópula de Gumbel #####
#pri.alfa = function(alfa,li,ls) {
# dunif(alfa,li,ls)
#}
#####

##### Função "post.biv" calcula a densidade a posteriori ###
#####
#pa.a & pa.b: parâmetros de forma e escala da priori gamma de alfa

post.biv = function(alfa,pa.a,pa.b){

```



```

post.biv = lik.biv(alfa)*pri.alfa(alfa,pa.a,pa.b)

return(post.biv)
}

#####

#####
### Algoritmo Metropolis-Hastings MH ###
#####

alfa.MH = numeric(0)
##### Cópula de Clayton #####
af1 = 1 #parâmetro de escala da priori de alfa
af2 = 0.001 #parâmetro de forma da priori de alfa

##### Cópula de Frank #####
#af1 = 0 #média da priori de alfa
#af2 = 100 #desvio da priori de alfa

##### Cópula de Gumbel #####
#af1 = 0 #média da priori de alfa
#af2 = 100 #desvio da priori de alfa

b = 10
j = 0
alfaj = 1

MH2 = function(){
while (j < 22000){
j = j+1
#print(j)

alfa.star = rgamma(1,b*alfaj,b)

aux.a = (post.biv(alfa.star,af1,af2)*qtr(alfaj,alfa.star,b))/
(post.biv(alfaj,af1,af2)*qtr(alfa.star,alfaj,b))

p.aceit = min(c(1,aux.a))
if(is.na(p.aceit)==TRUE) p.aceit=0 else {p.aceit = p.aceit}
z = runif(1,0,1)
if(z<=p.aceit) alfaj = alfa.star else {alfaj = alfaj}
alfa.MH = c(alfa.MH,alfaj)

#print(alfaj)
}
return(alfa.MH)
}

#####

ind = seq(1,20000,by=20) #índices que farão parte da amostra
MH2 = MH2()
MH2 = MH2[2001:22000] #Burn-in 2000
MH2 = MH2[ind] #Seleção de 1 a cada 20

alfa.MH = mean(MH2)
#alfa.MH #valor final estimado do parâmetro de associação da cópula

```