



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Estimação da Função de Distribuição Conjunta
para Dados com Censura Intervalar Bivariados

por

Marcos Paulo da Rocha Eirado

Orientador: Prof. Dr. Antônio Eduardo Gomes

Julho de 2010

Marcos Paulo da Rocha Eirado

Estimação da Função de Distribuição Conjunta 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 2010

Termo de Aprovação

Marcos Paulo da Rocha Eirado

Estimação da Função de Distribuição Conjunta 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 de defesa: 28 de junho de 2010

Orientador:

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

Comissão Examinadora:

Prof. Dr. George Freitas von Borries
Departamento de Estatística, UnB

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

Brasília, Julho de 2010

Ficha Catalográfica

EIRADO, MARCOS PAULO DA ROCHA

Estimação da Função de distribuição conjunta para dados com censura intervalar bivariados, (UnB - IE, Mestre em Estatística, 2010).

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

1. Censura intervalar 2. Distribuição bivariada 3. Análise de sobrevivência 4. Estimador não paramétrico de máxima verossimilhança

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 desta monografia de Projeto Final pode ser reproduzida sem a autorização por escrito do autor.

Marcos Paulo da Rocha Eirado

*À minha esposa, Fernanda,
e aos meus pais, Fátima
e Dato Eirado*

Agradecimentos

- Primeiramente a DEUS, pois tudo que tenho e sou foi dado por Ele. Todo louvor seja dado a Ele!
- À minha esposa, pelo apoio em todas as vezes que precisei de ajuda. Por me compreender em todas as noites que passei em claro e por todas as vezes que abdiquei de nosso lazer para concluir essa tarefa.
- A meus pais, pelo carinho com que me criaram e por me ensinarem a importância que a educação tem na vida de um cidadão.
- A meus irmãos, colegas de profissão, por todos os bons momentos que vivemos.
- Ao professor Antônio Eduardo Gomes, por estar sempre acessível e disposto a me ajudar, em todas as vezes que seu auxílio foi necessário.
- À família de minha esposa, por ter me acolhido e ter me ajudado.
- À minha cunhada Fabrícia, por ter me emprestado o seu notebook por um mês para que as simulações fossem concluídas a tempo.
- A todos os professores e servidores do programa de pós-graduação do Departamento de Estatística da UnB, por todo o suporte e estrutura necessária.
- Aos meus colegas e todas as outras pessoas que contribuíram para que esse objetivo fosse alcançado.

Sumário

Lista de Figuras	4
Lista de Tabelas	7
Resumo	8
Abstract	9
Introdução	10
1 Conceitos em Análise de Sobrevida	12
1.1 Censura à direita	13
1.2 Censura Intervalar	15
1.3 Caso Bivariado	18
2 Estimação da função de distribuição univariada	26
2.1 Método de Imputação Múltipla	27
2.2 Estimador não paramétrico de máxima verossimilhança	29
2.2.1 Estimação de F no Caso 1 de censura intervalar	31
2.2.2 Algoritmo ICM	33
3 Estimação da função de distribuição bivariada	41
4 Cópulas	46
4.1 Simulação com cópulas	49
4.2 Medidas de dependência	51
4.2.1 τ de Kendall	51

4.2.2	ρ de Spearman	52
4.3	Famílias de Cópulas	53
5	Suavização	57
5.1	Caso Univariado	58
5.1.1	Bootstrap	61
5.1.2	Validação Cruzada de Verossimilhança	62
5.2	Caso Bivariado	63
6	Simulações e aplicação	66
6.1	Simulações	66
6.2	Aplicação	102
7	Conclusões e Trabalhos Futuros	107
7.1	Conclusões	107
7.2	Trabalhos Futuros	108

Lista de Figuras

1.1	Dados com censura intervalar bivariada	18
2.1	Estimativas com algoritmo ICM	40
3.1	Estimador não paramétrico de máxima verossimilhança de H	45
4.1	Modelo de cópula Clayton	55
4.2	Modelo de cópula Frank	55
6.1	Densidade estimada para $\hat{\alpha}$ - Caso 1	69
6.2	Densidade estimada para $\hat{\alpha}$ - Caso 2	69
6.3	Densidade estimada para $\hat{\alpha}$ - Caso 3	70
6.4	Densidade conjunta estimada para janelas Caso 1, $n = 50$	70
6.5	Densidade conjunta estimada para janelas Caso 1, $n = 100$	71
6.6	Densidade conjunta estimada para janelas Caso 1, $n = 200$	71
6.7	Comparativo - Densidade das janelas - Estimador Cópula - Caso 1	72
6.8	Comparativo - Densidade das janelas - Estimador ENPMV - Caso 1	72
6.9	Densidade conjunta estimada para janelas Caso 2, $n = 50$	73
6.10	Densidade conjunta estimada para janelas Caso 2, $n = 100$	73
6.11	Densidade conjunta estimada para janelas Caso 2, $n = 200$	74
6.12	Comparativo - Densidade das janelas - Estimador Cópula - Caso 2	74
6.13	Comparativo - Densidade das janelas - Estimador ENPMV - Caso 2	75
6.14	Densidade conjunta estimada para janelas Caso 3, $n = 50$	75
6.15	Densidade conjunta estimada para janelas Caso 3, $n = 100$	76
6.16	Densidade conjunta estimada para janelas Caso 3, $n = 200$	76
6.17	Comparativo - Densidade das janelas - Estimador Cópula - Caso 3	77

6.18	Comparativo - Densidade das janelas - Estimador ENPMV - Caso 3 .	77
6.19	Valores esperados para o Caso 1	96
6.20	Valores esperados para o Caso 2	96
6.21	Valores esperados para o Caso 3	97
6.22	Valores esperados - Caso 1, $n = 50$	97
6.23	Valores esperados - Caso 1, $n = 100$	98
6.24	Valores esperados - Caso 1, $n = 200$	98
6.25	Valores esperados - Caso 2, $n = 50$	99
6.26	Valores esperados - Caso 2, $n = 100$	99
6.27	Valores esperados - Caso 2, $n = 200$	100
6.28	Valores esperados - Caso 3, $n = 50$	100
6.29	Valores esperados - Caso 3, $n = 100$	101
6.30	Valores esperados - Caso 3, $n = 200$	101
6.31	Estimativas das funções de distribuição marginais	102
6.32	Estimativa com cópula Clayton	103
6.33	Estimativa com cópula Frank	104
6.34	Estimativa com cópula Gumbel-Hougaard	104
6.35	Estimativa - estimador ENPMV	105

Lista de Tabelas

4.1	Algumas famílias clássicas de cópulas Arquimedianas	54
6.1	Parâmetros utilizados nas simulações ($i = 1, 2$)	66
6.2	Número de casos em que $\hat{\alpha}$ não foi obtido	67
6.3	Vício - estimador Cópula - caso 1, $n = 50$	78
6.4	Variância $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 50$	78
6.5	EQM $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 50$	78
6.6	Vício - estimador ENPMV - caso 1, $n = 50$	79
6.7	Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 50$	79
6.8	EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 50$	79
6.9	Vício - estimador Cópula - caso 1, $n = 100$	80
6.10	Variância $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 100$	80
6.11	EQM $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 100$	80
6.12	Vício - estimador ENPMV - caso 1, $n = 100$	81
6.13	Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 100$	81
6.14	EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 100$	81
6.15	Vício - estimador Cópula - caso 1, $n = 200$	82
6.16	Variância $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 200$	82
6.17	EQM $\times 10^{-3}$ - estimador Cópula - caso 1, $n = 200$	82
6.18	Vício - estimador ENPMV - caso 1, $n = 200$	83
6.19	Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 200$	83
6.20	EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, $n = 200$	83
6.21	Vício - estimador Cópula - caso 2, $n = 50$	84
6.22	Variância $\times 10^{-3}$ - estimador Cópula - caso 2, $n = 50$	84
6.23	EQM $\times 10^{-3}$ - estimador Cópula - caso 2, $n = 50$	84

6.24	Vício - estimador ENPMV - caso 2, n = 50	85
6.25	Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 50	85
6.26	EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 50	85
6.27	Vício - estimador Cópula - caso 2, n = 100	86
6.28	Variância $\times 10^{-3}$ - estimador Cópula - caso 2, n = 100	86
6.29	EQM $\times 10^{-3}$ - estimador Cópula - caso 2, n = 100	86
6.30	Vício - estimador ENPMV - caso 2, n = 100	87
6.31	Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 100	87
6.32	EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 100	87
6.33	Vício - estimador Cópula - caso 2, n = 200	88
6.34	Variância $\times 10^{-3}$ - estimador Cópula - caso 2, n = 200	88
6.35	EQM $\times 10^{-3}$ - estimador Cópula - caso 2, n = 200	88
6.36	Vício - estimador ENPMV - caso 2, n = 200	89
6.37	Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 200	89
6.38	EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 200	89
6.39	Vício - estimador Cópula - caso 3, n = 50	90
6.40	Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 50	90
6.41	EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 50	90
6.42	Vício - estimador ENPMV - caso 3, n = 50	91
6.43	Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 50	91
6.44	EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 50	91
6.45	Vício - estimador Cópula - caso 3, n = 100	92
6.46	Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 100	92
6.47	EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 100	92
6.48	Vício - estimador ENPMV - caso 3, n = 100	93
6.49	Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 100	93
6.50	EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 100	93
6.51	Vício - estimador Cópula - caso 3, n = 200	94
6.52	Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 200	94
6.53	EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 200	94
6.54	Vício - estimador ENPMV - caso 3, n = 200	95
6.55	Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 200	95

6.56	EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, $n = 200$	95
6.57	Estimativas para α	103
6.58	Estimativas para o estimador ENPMV	105
6.59	Dados do ACTG 181	106

Resumo

Estudamos dois métodos diferentes para estimar a função de distribuição H de um vetor aleatório (X, Y) com censura intervalar em ambas as variáveis. No primeiro, usamos o algoritmo proposto por Maathuis (2005) para obter o estimador não paramétrico de máxima verossimilhança (ENPMV) bivariado de H e, depois, suavizamos-no com o núcleo estimador bivariado. No segundo método, combinamos os ENPMV's das marginais, suavizados pelo núcleo estimador, com um modelo de cópula para obtermos uma estimativa suave de H . Por fim, fazemos a comparação desses métodos em relação a vício, variância e EQM.

Palavras Chave: *Censura intervalar, distribuição bivariada, análise de sobrevivência, estimador não paramétrico de máxima verossimilhança.*

Abstract

We study two different methods to estimate the joint distribution function H of a bivariate interval censored random vector (X, Y) . In the first one, we use the algorithm proposed by Maathuis (2005) to get the bivariate nonparametric maximum likelihood estimator (NPMLE) of H and, then, smooth it with a bivariate kernel estimator. In the second approach, we combine the kernel smoothed NPMLE estimator of the marginal distributions with copulas to get a kernel smoothed estimate of H . After that, we compare these estimates using bias, variance and MSE as criteria.

key words: *Interval censoring, bivariate distribution, survival analysis, nonparametric maximum likelihood estimator.*

Introdução

O homem moderno está sempre tentando aperfeiçoar seus processos. Por isso, busca conhecer o mundo que o cerca, estudando fenômenos biológicos, físicos, sociais, entre outros. Em qualquer um dos casos, a análise de dados experimentais é um etapa fundamental para que se chegue a conclusões.

Todavia, nem sempre é possível observar esses dados da melhor forma possível. Às vezes, temos restrições financeiras, barreiras tecnológicas ou algum outro fato que impede que mensuremos a gradeza estudada com perfeição. Isso acontece, por exemplo, quando observamos o tempo de desenvolvimento de algumas doenças letais, em que só é possível saber se houve o desenvolvimento citado se o animal usado no estudo for sacrificado.

Nesse contexto, surge a Análise de Sobrevivência como ferramenta estatística capaz de aproveitar, de acordo com algumas premissas, dados que são observados parcialmente, chamados de dados censurados, na análise estatística. Desenvolver metodologias de análise fica mais difícil dependendo de como acontece essa “censura” na observação do fenômeno.

Na investigação dos fenômenos de que falamos, muitas vezes é importante encontrar relações entre dois fatores. Se esses fatores podem ser mensurados, o pesquisador pode querer saber como eles interagem, isto é, atuam conjuntamente. Em termos matemáticos, o pesquisador tem interesse em estimar a distribuição conjunta desses dois fatores. No entanto, se há censura na observação dos dois fatores, cria-se um problema maior.

Seja H a função de distribuição conjunta de um vetor aleatório (X, Y) , em que X e Y são variáveis aleatórias positivas, ambas com censura intervalar. No presente esforço, estudamos dois métodos diferentes para estimar H , a fim de compará-los em

relação ao vício e varibilidade, por meio de simulações.

No primeiro método, usamos o algoritmo proposto por Maathuis (2005) para estimar não-parametricamente H de forma direta e depois suavizamos essa estimativa com o núcleo estimador bivariado. O outro método estima H fazendo o uso de cópulas para unir as estimativas não paramétricas das funções de distribuição marginais, suavizadas pelo núcleo estimador, obtidas com o algoritmo iterativo da minorante convexa máxima apresentado por Groeneboom e Wellner (1992).

O trabalho está organizado como segue. No capítulo 1, apresentamos alguns conceitos básicos de Análise de Sobrevivência. No capítulo 2, tratamos de dois estimadores não paramétricos da função de distribuição univariada. No capítulo 3, tratamos do estimador não paramétrico de máxima verossimilhança da função de distribuição bivariada. No capítulo 4, discutimos Cópulas. No capítulo 5, apresentamos o método de suavização para funções de distribuição por meio do núcleo estimador. No capítulo 6, são exibidos os resultados encontrados nas simulações realizadas. Finalmente, o capítulo 7 traz as conclusões de nossos estudos.

A inspiração para essa dissertação veio do trabalho desenvolvido por Romeo, Tanaka e Lima (2006), que usam uma abordagem bayesiana e modelos de cópulas para estimar a função de sobrevivência em dados com censura à direita.

Capítulo 1

Conceitos em Análise de Sobrevivência

Começamos esse texto definindo alguns conceitos importantes e fixando a notação que será utilizada ao longo da leitura. Supomos que o leitor tem conhecimentos básicos em Teoria de Probabilidades e está familiarizado com alguns métodos e procedimentos estatísticos, como, por exemplo, a estimação pelo método de máxima verossimilhança.

Em Análise de Sobrevivência, geralmente estudamos o tempo T *transcorrido até a ocorrência de um evento de interesse*, que chamamos genericamente de “falha”. Por isso, chamamos a variável aleatória T de *tempo de falha*. Queremos descrever a distribuição de probabilidades da variável aleatória T , que aqui vamos supor contínua. Isso pode ser feito se conhecermos a sua função de distribuição, que é definida por

$$F(t) = P(T \leq t), \quad t \in \mathbb{R}.$$

A função de Sobrevivência de T , S , é dada por

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

A função taxa de falha é

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} = \frac{f(t)}{S(t)},$$

em que f é a função densidade de probabilidades de T , ou seja, $f(t) = \frac{dF}{dt}(t)$.

Mesmo com um bom delineamento amostral, às vezes é difícil ou até mesmo impossível ter certeza de quando ocorre a falha em uma unidade amostral. Assim sendo, quando acontece a observação parcial de T , dizemos que houve *censura* em T . Se não há censura, dizemos que a observação é *exata*. Portanto, a característica principal dos dados em Análise de Sobrevivência é a presença de censura. De fato, as observações censuradas de T são importantes e não podem ser descartadas. Queremos incorporá-las em nossas análises, visto que há informação nelas e omiti-las pode levar a estimativas viciadas e conclusões equivocadas.

Algumas hipóteses são assumidas para prosseguirmos com nossas conclusões. Vamos supor que o mecanismo que produz a censura é aleatório e independente de T . Dessa maneira, podemos obter a função de verossimilhança para um dado tipo de censura. Para tornar o conceito de censura mais claro e delimitar o campo de estudo em que se pode atuar, é essencial caracterizar como se deu a ocorrência de censura. A literatura classifica a censura em três tipos: censura à esquerda, censura à direita e censura intervalar. Na censura à esquerda, T assume um valor inferior ao tempo registrado (tempo de censura). Na censura à direita, T assume um valor superior ao tempo de censura. Na censura intervalar, sabemos que o tempo de falha assume um valor entre dois instantes de observação.

A estrutura mais comum em dados de sobrevivência apresenta valores com censura à direita e valores exatos para o tempo de falha.

1.1 Censura à direita

Na censura à direita, sabemos que o tempo de falha T ocorre após o tempo efetivamente observado, que denotamos por X . Seja C a variável aleatória que representa o tempo transcorrido até que ocorra a censura na observação de T . Conforme já foi dito, supomos que T e C são independentes. Nesse caso de censura, temos que

$$X = \min\{C, T\}.$$

Esse tipo de censura é comum em estudos clínicos, em que o paciente abandona a

pesquisa antes de desenvolver um determinada doença.

Com esta configuração, obtemos, portanto, uma amostra aleatória X_1, X_2, \dots, X_n de X , onde X_i pode ser uma observação exata ou censurada. Precisamos encontrar a função de verossimilhança para estes dados coletados. Antes, definimos a variável

$$\delta = \begin{cases} 1, & \text{se } T \leq C; \\ 0, & \text{se } T > C. \end{cases} \quad (1.1)$$

A variável δ indica se houve censura à direita ($\delta = 0$) ou observação exata ($\delta = 1$) na unidade amostral. Dada uma amostra $(X_1, \delta_1), (X_2, \delta_2), \dots, (X_n, \delta_n)$ de dados com censura à direita, pode-se mostrar (Colosimo e Giolo (2006)) que a verossimilhança para esses dados é

$$L(F) = \prod_{i=1}^n [f(X_i)]^{\delta_i} [S(X_i)]^{1-\delta_i}, \quad (1.2)$$

onde F é a função de distribuição de T . Podemos ver na expressão acima que a verossimilhança conta com a contribuição dos dados exatos, com os termos $f(X_i)$, e com a contribuição dos dados censurados, por meio dos termos $S(X_i)$. O estimador não paramétrico de máxima verossimilhança (Colosimo e Giolo (2006)) de $S(t)$ é o estimador de Kaplan-Meier (ou estimador limite-produto), definido por

$$\hat{S}(t) = \prod_{i: y_i \leq t} \left(\frac{n_i - d_i}{n_i} \right),$$

onde y_1, y_2, \dots, y_n são os valores ordenados de X_1, X_2, \dots, X_n , d_i é o número de falhas em y_i e n_i é o número de observações sob risco até y_i , ou seja, que não falharam nem foram censuradas até y_i .

O estimador de Kaplan-Meier é muito conhecido e está implementado em muitos pacotes estatísticos, o que o torna fácil de usar.

1.2 Censura Intervalar

Em dados com censura intervalar, só temos conhecimento de que T pertence a um intervalo $[L, U]$. Essa situação é típica em estudos em que ocorrem inspeções periódicas nas unidades amostrais. Nesse caso, não podemos observar exatamente quando ocorre a falha, mas sabemos que ela ocorreu entre as duas últimas inspeções. Supomos que as variáveis aleatórias L e U que são efetivamente observadas são independentes de T . Também ocorre que $P(L \leq U) = 1$.

Podemos generalizar as outras formas de censura usando a estrutura de dados da censura intervalar: na censura à direita, tomamos $L = C$ e $U = \infty$ para obtermos $T \in [C, \infty)$; na censura à esquerda, colocamos $L = 0$ e $U = C$ para termos $T \in (0, C]$; se observamos T de fato, escrevemos $L = T$ e $U = T$, para termos $T \in [T, T]$. Dessa maneira, é possível colocar todas as observações com a estrutura $[L_i, U_i]$.

Suponhamos, então, que não foi possível obter uma amostra de observações exatas T_1, T_2, \dots, T_n de T e, por assim ser, possuímos apenas uma amostra de pares $[L_1, U_1], [L_2, U_2], \dots, [L_n, U_n]$, sendo que pode ocorrer $L_i = 0$ ou $U_i = \infty$, $i = 1, 2, \dots, n$, de acordo com a observação que fizemos no parágrafo anterior. Definimos as variáveis

$$\delta = \begin{cases} 1, & \text{se } T \leq L; \\ 0, & \text{caso contrário.} \end{cases}$$

e

$$\gamma = \begin{cases} 1, & \text{se } 0 \neq L < T \leq U < \infty \\ 0, & \text{caso contrário} \end{cases},$$

Como podemos ver, a variável δ é indicadora da existência de censura à esquerda, isto é, $\delta = 1_{\{T \leq L\}}$. Analogamente, γ indica a presença de censura intervalar na unidade amostral, o que significa que $\gamma = 1_{\{L < T \leq U\}}$. A verossimilhança para dados com censura intervalar (ver Groeneboom e Wellner (1992)) é dada por

$$L(F) = \prod_{i=1}^n [F(L_i)]^{\delta_i} [F(U_i) - F(L_i)]^{\gamma_i} [1 - F(U_i)]^{1 - \delta_i - \gamma_i}. \quad (1.3)$$

Como $f_{L,U}(x, y) = 0$ para $x > y$, pois $P(L \leq U) = 1$, a demonstração desse

resultado é a que segue:

$$\begin{aligned}
P(L \leq l, U \leq u, \delta = 1, \gamma = 0) &= P(L \leq l, U \leq u, T \leq L) \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_{-\infty}^x f_{L,U,T}(x, y, t) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_{-\infty}^x f_T(t) f_{L,U}(x, y) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l F_T(x) f_{L,U}(x, y) dx dy,
\end{aligned}$$

pela independência entre T e (L, U) , segue que

$$\begin{aligned}
f(l, u, 1, 0) &= \frac{\partial^2}{\partial l \partial u} P(L \leq l, U \leq u, \delta = 1, \gamma = 0) \\
&= F_T(l) f_{L,U}(l, u), \quad \text{para } l < u.
\end{aligned}$$

Também,

$$\begin{aligned}
P(L \leq l, U \leq u, \delta = 0, \gamma = 1) &= P(L \leq l, U \leq u, L \leq T \leq U) \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_x^y f_{L,U,T}(x, y, t) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_x^y f_T(t) f_{L,U}(x, y) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l [F_T(y) - F_T(x)] f_{L,U}(x, y) dx dy,
\end{aligned}$$

o que nos dá

$$\begin{aligned}
f(l, u, 0, 1) &= \frac{\partial^2}{\partial l \partial u} P(L \leq l, U \leq u, \delta = 0, \gamma = 1) \\
&= [F_T(u) - F_T(l)] f_{L,U}(l, v), \quad \text{para } l < u.
\end{aligned}$$

Para o último caso, com $\delta = \gamma = 0$, tem-se

$$\begin{aligned}
P(L \leq l, U \leq u, \delta = 0, \gamma = 0) &= P(L \leq l, U \leq u, T > U) \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_u^{\infty} f_{L,U,T}(x, y, t) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l \int_u^{\infty} f_T(t) f_{L,U}(x, y) dt dx dy \\
&= \int_{-\infty}^u \int_{-\infty}^l [1 - F_T(x)] f_{L,U}(x, y) dx dy,
\end{aligned}$$

de onde segue que

$$\begin{aligned}
f(l, u, 0, 0) &= \frac{\partial^2}{\partial l \partial u} P(L \leq l, U \leq u, \delta = 0, \gamma = 0) \\
&= [1 - F_T(u)] f_{L,U}(l, u), \quad \text{para } l < u.
\end{aligned}$$

□

Convém observar que a expressão da verossimilhança em (1.3) tem essa forma se estamos considerando a notação com a quádrupla $\{(L_i, U_i, \delta_i, \gamma_i)\}$. Na prática, usando essa notação, quando $\delta_i = 1$, não observamos de fato U_i . Da mesma forma, quando $\delta_i + \gamma_i = 0$, não registramos L_i .

Por outro lado, se usarmos a notação (L_i, U_i) com L_i e U_i podendo assumir $L_i = 0$ ($\delta_i = 1$) e $U_i = \infty$ ($\delta_i + \gamma_i = 0$), a verossimilhança (1.3) pode ser reescrita como

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

Existe ainda um caso particular de censura intervalar, chamado de **caso 1** de censura intervalar ou dados de estado corrente, em que só ocorrem censura à direita ou censura à esquerda. Neste caso, há apenas um instante de observação (tempo de censura). Se assim for, usando a notação da censura intervalar, a verossimilhança é dada por

$$L(F) = \prod_{i=1}^n [F(L_i)]^{\delta_i} [1 - F(U_i)]^{1-\delta_i}. \quad (1.4)$$

Diferentemente do caso em que só há censura à direita, quando temos dados com censura intervalar, o estimador não-paramétrico de máxima verossimilhança de S não

tem forma fechada. O capítulo 2 trata esse problema com detalhe.

Uma vez definidas quantidades importantes e apresentadas as funções de verossimilhança de que precisaremos em nossos estudos para o caso univariado, podemos agora nos ater ao caso bivariado, tema de estudo nesse trabalho.

1.3 Caso Bivariado

Sejam T_1 e T_2 duas variáveis aleatórias contínuas e positivas. A função de distribuição conjunta do vetor aleatório (T_1, T_2) é definida por

$$H(t_1, t_2) = P(T_1 \leq t_1, T_2 \leq t_2), \quad t_1, t_2 \in \mathbb{R}.$$

Se temos censura intervalar nos valores de ambas as variáveis T_1 e T_2 , teremos uma amostra $[L_{11}, U_{11}] \times [L_{21}, U_{21}], [L_{12}, U_{12}] \times [L_{22}, U_{22}], \dots, [L_{1n}, U_{1n}] \times [L_{2n}, U_{2n}]$ de retângulos em \mathbb{R}^2 . Apresentamos um exemplo para esse tipo de dados na figura 1.1.

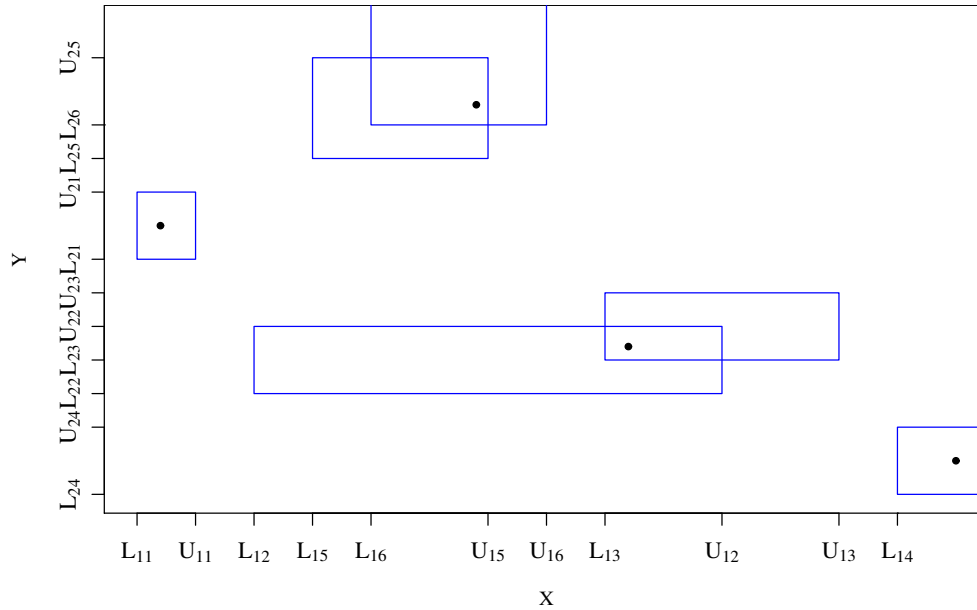


Figura 1.1: Dados com censura intervalar bivariada

Assim, a função de verossimilhança para esses dados, que depende de H , é (ver Gentleman e Vandal (2002))

$$L(H) = \prod_{i=1}^n [H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i})], \quad (1.5)$$

em que empregamos a notação compacta (L_i, U_i) , isto é, pode ocorrer $L_i = 0$ (quando $\delta_i = 1$) e $U_i = \infty$ (sempre que $\delta_i + \gamma_i = 0$).

Demonstraremos como obter a expressão (1.5). Temos nove possibilidades diferentes, de acordo com os valores de $\delta_1, \gamma_2, \delta_1$ e γ_2 . Mostraremos como se obtém a verossimilhança caso a caso:

1º caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 < L_1, L_2 \leq l_2, U_2 \leq u_2, T_2 < L_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{x_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{-\infty}^{x_1} f_{L_1, U_1, T_1, L_2, U_2, T_2}(x_1, y_1, t_1, x_2, y_2, t_2) \\
&\hspace{25em} \times dt_1 dx_1 dy_1 dt_2 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f_{T_1, T_2}(t_1, t_2) \\
&\hspace{25em} \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} H(x_1, x_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2,
\end{aligned}$$

pela independência dos vetores aleatórios (T_1, T_2) e (L_1, U_1, L_2, U_2) . Diferenciando, segue que

$$\begin{aligned}
& f(l_1, u_1, 1, 0, l_2, u_2, 1, 0) = \\
& \frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= H(l_1, l_2) f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2), \quad \text{para } l_1 < u_1, l_2 < u_2.
\end{aligned}$$

2º caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 < L_1, L_2 \leq l_2, U_2 \leq u_2, L_2 \leq T_2 \leq U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{x_2}^{y_2} \int_{-\infty}^{x_1} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\quad \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [H(x_1, y_2) - H(x_1, x_2)] f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\quad \times dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Daí,

$$\begin{aligned}
& f(l_1, u_1, 1, 0, l_2, u_2, 0, 1) = \\
& \frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= [H(l_1, u_2) - H(l_1, l_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2), \quad \text{para } l_1 < u_1, l_2 < u_2.
\end{aligned}$$

3º caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 < L_1, L_2 \leq l_2, U_2 \leq u_2, T_2 > U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{y_2}^{\infty} \int_{-\infty}^{x_1} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\quad \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [F_{T_1}(x_1) - H(x_1, y_2)] \\
&\quad \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Portanto,

$$\begin{aligned}
f(l_1, u_1, 1, 0, l_2, u_2, 0, 0) &= \\
\frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 1, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= [F_{T_1}(l_1) - H(l_1, u_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2), \quad \text{para } l_1 < u_1, l_2 < u_2.
\end{aligned}$$

4º caso

$$\begin{aligned}
P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, L_1 \leq T_1 \leq U_1, L_2 \leq l_2, U_2 \leq u_2, T_2 < L_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{-\infty}^{x_2} \int_{x_1}^{y_1} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\hspace{20em} \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [H(y_1, x_2) - H(x_1, x_2)] \\
&\hspace{20em} \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Logo,

$$\begin{aligned}
f(l_1, u_1, 0, 1, l_2, u_2, 1, 0) &= \\
\frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= [H(u_1, l_2) - H(l_1, l_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2), \quad \text{para } l_1 < u_1, l_2 < u_2.
\end{aligned}$$

5° caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, L_1 \leq T_1 \leq U_1, L_2 \leq l_2, U_2 \leq u_2, L_2 \leq T_2 \leq U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{x_2}^{y_2} \int_{x_1}^{y_1} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\quad \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [H(y_1, y_2) - H(y_1, x_2) - H(x_1, y_2) + H(x_1, x_2)] \\
&\quad \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Assim,

$$\begin{aligned}
& f(l_1, u_1, 0, 1, l_2, u_2, 0, 1) = \\
& \frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= [H(u_1, u_2) - H(u_1, l_2) - H(l_1, u_2) + H(l_1, l_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2),
\end{aligned}$$

para $l_1 < u_1, l_2 < u_2$.

6° caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, L_1 \leq T_1 \leq U_1, L_2 \leq l_2, U_2 \leq u_2, T_2 > U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{y_2}^{\infty} \int_{x_1}^{y_1} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\quad \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [F_{T_1}(y_1) - F_{T_1}(x_1) - H(y_1, y_2) + H(x_1, y_2)] \\
&\quad \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Então,

$$\begin{aligned}
f(l_1, u_1, 0, 1, l_2, u_2, 0, 0) &= \\
&\frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 1, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= [F_{T_1}(u_1) - F_{T_1}(l_1) - H(u_1, u_2) + H(l_1, u_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2),
\end{aligned}$$

para $l_1 < u_1, l_2 < u_2$.

7º caso

$$\begin{aligned}
&P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 > U_1, L_2 \leq l_2, U_2 \leq u_2, T_2 < L_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{-\infty}^{x_2} \int_{y_1}^{\infty} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\hspace{25em} \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [F_{T_2}(x_2) - H(y_1, x_2)] \\
&\hspace{15em} \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Com isso,

$$\begin{aligned}
f(l_1, u_1, 0, 0, l_2, u_2, 1, 0) &= \\
&\frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 1, \gamma_2 = 0) \\
&= [F_{T_2}(l_2) - H(u_1, l_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2), \quad \text{para } l_1 < u_1, l_2 < u_2.
\end{aligned}$$

8° caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 > U_1, L_2 \leq l_2, U_2 \leq u_2, L_2 \leq T_2 \leq U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{x_2}^{y_2} \int_{y_1}^{\infty} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\hspace{20em} \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [F_{T_2}(y_2) - F_{T_2}(x_2) - H(y_1, y_2) + H(y_1, x_2)] \\
&\hspace{20em} \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Logo,

$$\begin{aligned}
& f(l_1, u_1, 0, 0, l_2, u_2, 0, 1) = \\
& \frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 1) \\
&= [F_{T_2}(u_2) - F_{T_2}(l_2) - H(u_1, u_2) + H(u_1, l_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2),
\end{aligned}$$

para $l_1 < u_1, l_2 < u_2$.

9° caso

$$\begin{aligned}
& P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= P(L_1 \leq l_1, U_1 \leq u_1, T_1 > U_1, L_2 \leq l_2, U_2 \leq u_2, T_2 > U_2) \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} \int_{y_2}^{\infty} \int_{y_1}^{\infty} f_{T_1, T_2}(t_1, t_2) f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) \\
&\hspace{20em} \times dt_1 dt_2 dx_1 dy_1 dx_2 dy_2 \\
&= \int_{-\infty}^{u_2} \int_{-\infty}^{l_2} \int_{-\infty}^{u_1} \int_{-\infty}^{l_1} [1 - F_{T_1}(y_1) - F_{T_2}(x_2) + H(y_1, y_2)] \\
&\hspace{20em} \times f_{L_1, U_1, L_2, U_2}(x_1, y_1, x_2, y_2) dx_1 dy_1 dx_2 dy_2.
\end{aligned}$$

Portanto,

$$\begin{aligned}
f(l_1, u_1, 0, 0, l_2, u_2, 0, 0) &= \\
\frac{\partial^4}{\partial l_1 \partial u_1 \partial l_2 \partial u_2} P(L_1 \leq l_1, U_1 \leq u_1, \delta_1 = 0, \gamma_1 = 0, L_2 \leq l_2, U_2 \leq u_2, \delta_2 = 0, \gamma_2 = 0) \\
&= [1 - F_{T_1}(u_1) - F_{T_2}(u_2) + H(u_1, u_2)] f_{L_1, U_1, L_2, U_2}(l_1, u_1, l_2, u_2),
\end{aligned}$$

para $l_1 < u_1, l_2 < u_2$. Como teríamos uma expressão muito extensa para a função de verossimilhança no caso bivariado, preferimos adotar a notação com as duplas $(L_{.i}, U_{.i})$.

Se admitirmos alguma família paramétrica para H ou F , no caso univariado, podemos estimar H (ou F) maximizando a verossimilhança em (1.5) (ou a verossimilhança em (1.3)) em relação aos parâmetros de H (ou F). Mesmo assim, assumir alguma estrutura de dependência entre T_1 e T_2 pode ser uma tarefa difícil. Por isso, faremos uso de modelos de cópulas para tentar lidar com esse desafio.

Assim como no caso univariado, o estimador não-paramétrico de máxima verossimilhança de H não possui forma fechada. Na verdade, muitos avanços feitos nesse sentido são recentes e serão considerados aqui.

Conforme já mencionamos, nosso interesse, nesse trabalho, está em estimar H . Faremos isso usando técnicas expostas nos capítulos que seguirão. Uma dessas técnicas consiste em estimar as funções de distribuição marginais de T_1 e T_2 para estimar H pelo uso de modelos de cópulas, que são modelos paramétricos de distribuição conjunta com distribuições marginais $U(0, 1)$. Assim sendo, apresentamos agora métodos para estimar a função de distribuição de uma variável aleatória cujas observações possuem censura intervalar.

Capítulo 2

Estimação da função de distribuição univariada

A distribuição de uma variável aleatória T é totalmente caracterizada por sua função de distribuição F . Por isso, a sua estimação é de suma importância. Com dados de censura intervalar, a tarefa de estimar F não parametricamente é mais complexa que para dados completos.

Vários estudos foram desenvolvidos ao longo do tempo para estimar F não parametricamente, sendo notáveis os trabalhos de Peto (1973), Turnbull (1976), entre outros, que procuraram formas de computar o estimador de máxima verossimilhança não paramétrico \hat{F} de F . Peto (1973) caracteriza as regiões onde \hat{F} irá ter saltos, enquanto Turnbull (1976) obtém equações de auto-consistência que \hat{F} deve satisfazer, além de desenvolver o algoritmo de auto-consistência para o cálculo de \hat{F} . Uma contribuição muito importante foi dada por Groeneboom e Wellner (1992), com o algoritmo *Iterative Convex Minorant* (ICM), que calcula \hat{F} consideravelmente mais rápido que o algoritmo de auto-consistência dado por Turnbull (1976).

Além do estimador de máxima de verossimilhança não paramétrico, existem outros métodos não paramétricos para estimar F . Entre eles, está o Método de Imputação Múltipla proposto por Pan (2000), que faz a imputação de dados e uso do estimador de Kaplan-Meier para alcançar seus fins. Apresentamos esses dois métodos aqui, começando com a metodologia de Pan (2000).

2.1 Método de Imputação Múltipla

Nesta seção, apresentamos o método de imputação múltipla para dados com censura intervalar proposto por Pan (2000). Na verdade, fazemos uma adaptação da idéia original, pois Pan (2000) estuda o efeito de covariáveis na estimação da função de sobrevivência S . Em nosso trabalho, contudo, não abrangemos o estudo de covariáveis.

Suponhamos, portanto, que não observamos diretamente os valores de uma variável aleatória T que tem função de sobrevivência S e, por isso, obtemos apenas uma amostra de intervalos $[L_i, U_i]$ que contém T_i , para $i = 1, 2, \dots, n$. Lembremos que o caso $U_i = \infty$ é o caso de censura à direita e $L_i = 0$ corresponde a censura à esquerda. A idéia desse método é a seguinte: gerar m conjuntos de dados com censura à direita, com os vetores $\mathbf{T}_j = (T_{j1}, T_{j2}, \dots, T_{jn})$ e $\boldsymbol{\delta}_j = (\delta_{j1}, \delta_{j2}, \dots, \delta_{jn})$, $j = 1, 2, \dots, m$, onde $T_{ji} \in [L_i, U_i]$, e, a partir desses dados imputados, obter uma estimativa \bar{S} de S usando métodos para dados com censura à direita, que são amplamente difundidos e facilmente computados.

Para usar os estimadores para dados censurados à direita, é necessário termos a estrutura para dados com esse tipo de censura, que é aquela composta pelos pares $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$, onde a variável δ , conforme definimos na expressão (1.1), indica a presença de censura à direita ($\delta = 0$) ou falha ($\delta = 1$). Faremos essa conversão da estrutura dos dados por meio do algoritmo Poor Man's Data Augmentation (PDMA), apresentado por Wei e Tanner (1991), descrito a seguir:

1. Suponha que a atual estimativa de S é $\bar{S}_{(k)}$.
2. Gere m conjuntos de dados possivelmente censurados à direita $\{(\mathbf{T}_1, \boldsymbol{\delta}_1), (\mathbf{T}_2, \boldsymbol{\delta}_2), \dots, (\mathbf{T}_m, \boldsymbol{\delta}_m)\}$ da seguinte maneira: Fixe $j \in \{1, 2, \dots, m\}$; Se $U_i < \infty$, gere T_{ji} da distribuição dada por $\bar{S}_{(k)}$ dado que $T_{ji} \in [L_i, U_i]$ e faça $\delta_{ji} = 1$; Se $U_i = \infty$, faça $T_{ji} = L_i$ e $\delta_{ji} = 0$.
3. Para cada $j = 1, 2, \dots, m$, obtenha uma estimativa $\bar{S}_j^{(k)}$ de S usando o conjunto de dados $(\mathbf{T}_j, \boldsymbol{\delta}_j)$.
4. Faça $\bar{S}_{(k+1)} = \frac{1}{m} \sum_{j=1}^m \bar{S}_j^{(k)}$. Se $|\bar{S}_{(k+1)} - \bar{S}_{(k)}| < \varepsilon$, pare. Caso contrário, faça $k = k + 1$ e volte ao passo 1.

Analisando o algoritmo, vemos que a mudança da estrutura de dados se dá da seguinte maneira: nas observações em que não há censura à direita, é gerada uma observação de uma variável aleatória $T_{ji} \sim \bar{S}_{(k)}(\cdot | L_i < T_{ji} \leq U_i)$ e essa observação simulada é tratada como se fosse exata; por outro lado, o algoritmo mantém as observações censuradas à direita como elas são originalmente. Com esses novos dados, podemos estimar S facilmente, uma vez que agora temos uma amostra de tempos de falha censurados à direita ou exatos.

Ainda pela descrição do algoritmo, é necessário que tenhamos uma estimativa inicial de S , $\bar{S}_{(0)}$. Para obter isso, PPan (2000) recomenda que geremos T_{ij} como no passo 2, mas teremos $T_{ij} \sim U(L_i, U_i)$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$, quando não há censura à direita, ao invés de fazermos $T_{ji} \sim \bar{S}_{(k)}(\cdot | L_i < T_{ji} \leq U_i)$. Com essa amostra, calculamos $\bar{S}_{(0)}$ como nos passos 3 e 4. O cálculo de $S_j^{(k)}$, no passo 3, pode ser feito com o estimador de Kaplan-Meier.

É importante notarmos que a estimativa $\bar{S}_{(k)}$ obtida da forma estipulada pelo algoritmo é uma função escada, ou seja, é discreta. Assim sendo, existe um número finito de saltos estimados $\{s_1, s_2, \dots, s_{k_i}\}$ em pontos $\{t_1, t_2, \dots, t_{k_i}\}$ contidos em cada intervalo $[L_i, U_i]$ em que não há censura à direita. Isso faz com que o valor de T_{ji} gerado seja igual a um dos valores em $\{t_1, t_2, \dots, t_{k_i}\}$. Com isso, a cada iteração, os elementos de $\{t_1, t_2, \dots, t_{k_i}\}$ que não forem amostrados no processo de imputação não serão pontos de salto de $\bar{S}_{(k+1)}$ e, portanto, não serão considerados nas imputações das iterações seguintes. Isto pode fazer com que \bar{S} evolua para uma estimativa com cada vez menos saltos. Para contornarmos esse problema, pode-se suavizar $\bar{S}_{(k)}$ com o uso de núcleo estimadores, a fim de que tenhamos uma amostragem contínua no processo de imputação.

Pan (2000) afirma que o valor de m não precisa ser grande para que bons resultados sejam obtidos. Em suas simulações, foi usado $m = 10$. É essencial registrarmos que apesar de estudos de simulação apresentarem bons resultados, não há provas ou critérios que determinam a distribuição de \bar{S} ou sua convergência para S . Na verdade, Paz (2005) observou indícios de não convergência do algoritmo quando aplicado ao ajuste do modelo de Cox para dados simulados com a presença de covariável.

2.2 Estimador não paramétrico de máxima verossimilhança

Uma vez apresentada a estimação pela imputação múltipla, atemo-nos agora ao estimador não paramétrico de máxima verossimilhança de F , que denotaremos por \hat{F} . Queremos encontrar, na classe \mathcal{C} de todas as funções de distribuição de v.a.'s positivas, uma função de distribuição \hat{F} que maximiza a função de verossimilhança para dados com censura intervalar, que é dada, conforme a expressão (1.3), por

$$L(F) = \prod_{i=1}^n [F(L_i)]^{\delta_i} [F(U_i) - F(L_i)]^{\gamma_i} [1 - F(U_i)]^{1-\delta_i-\gamma_i}, \quad F \in \mathcal{C}. \quad (1.3)$$

Conforme já observamos, $L(F)$ é o produto das diferenças $F(U_i) - F(L_i)$, em que pode ocorrer $L_i = 0$ ou $U_i = \infty$. Isto significa que o valor da função de verossimilhança é indiferente à forma de uma função de distribuição qualquer no interior dos intervalos $[L_i, U_i]$. Mais ainda, essa observação nos leva a concluir que todas as funções de distribuição cujos valores $F(U_i) - F(L_i)$ são iguais terão o mesmo valor para L . Assim sendo, duas funções de distribuição F_1 e F_2 podem ser diferentes, mas é possível ocorrer $L(F_1) = L(F_2)$.

Os argumentos do parágrafo anterior nos mostram que \hat{F} pode não ser única no interior dos intervalos $[L_i, U_i]$, em que pode ocorrer $L_i = 0$ ou $U_i = \infty$. Isso acontece porque o valor da verossimilhança avaliada em \hat{F} é indiferente à forma de \hat{F} em (L_i, U_i) .

De qualquer forma, encontrar \hat{F} da maneira como temos definido até agora é bastante difícil, uma vez que a classe \mathcal{C} definida como acima tem dimensão infinita. Além disso, há infinitas formas de redefinirmos uma função que maximiza (1.3) para encontrarmos outro estimador de máxima verossimilhança para F , bastando modificar sua forma nos intervalos (L_i, U_i) . Portanto, pelo que acabamos de explicar, é necessário caracterizarmos melhor \hat{F} , de forma que consigamos calcular \hat{F} de forma única, em um espaço “menor” de funções de distribuição. Nesse sentido, o trabalho de Peto (1973) é fundamental e, por esse motivo, resumiremos sua contribuição para este assunto aqui.

Peto (1973) mostra que o cálculo de \hat{F} consiste, na verdade, na estimação de um

número finito de parâmetros. Para tanto, define um conjunto de intervalos disjuntos fechados $\{[a_j, b_j] : j = 1, \dots, J\}$, que são os intervalos cujos extremos inferiores a_j pertencem ao conjunto $\{L_i, i = 1, 2, \dots, n\}$ e os extremos superiores b_j pertencem ao conjunto $\{U_i, i = 1, 2, \dots, n\}$ e que não contêm nenhum ponto de $\{L_i\}$ ou de $\{U_i\}$, a não ser em seus pontos extremos inferiores ou superiores, respectivamente.

Com essa definição dos intervalos $[a_j, b_j]$, que chamamos de *interseções máximas*, Peto (1973) prova a seguinte

Proposição 2.1. *Na procura por \hat{F} , é suficiente considerar apenas as funções de distribuição que são horizontais em todos os pontos exceto nos intervalos $[a_j, b_j]$ e que crescem em alguns ou em todos os intervalos $[a_j, b_j]$.*

Turnbull (1976) mostrou também que qualquer função de sobrevivência que cresce fora de $\bigcup_{j=1}^J [a_j, b_j]$ não pode ser uma estimativa de máxima verossimilhança de F .

Pela Proposição 2.1, vemos que \hat{F} terá crescimento apenas nos intervalos $[a_j, b_j]$, ou seja, haverá um “salto” em \hat{F} nessas regiões. Por isso, para unificarmos o cálculo de \hat{F} , vamos convencionar que esses saltos em \hat{F} , isto é, os incrementos de \hat{F} em $[a_j, b_j]$, se darão totalmente nos pontos b_j e, assim, \hat{F} será constante no interior dos intervalos $[a_j, b_j]$. Note-se que, pela proposição anterior e com essa convenção, saímos de uma procura de uma função em um espaço infinito de funções para a estimação de um número finito de parâmetros. Por isso, o cálculo dos intervalos $[a_j, b_j]$ é chamado de passo da redução de dimensão.

Postas as considerações acima sobre \hat{F} , falta-nos saber em quais dos intervalos $[a_j, b_j]$ \hat{F} terá saltos e quanto valerá cada salto desse. Em termos matemáticos, o salto p_j no intervalo $[a_j, b_j]$ é dado por $p_j = F(b_j) - F(a_j)$, $j = 1, 2, \dots, J$.

Podemos reescrever a verossimilhança (1.3) em função do vetor $\mathbf{p} = (p_1, p_2, \dots, p_J)$. Antes, definimos, para $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, J$, as variáveis

$$\alpha_{ij} = \begin{cases} 1, & \text{se } [a_j, b_j] \subseteq [L_i, U_i]; \\ 0, & \text{caso contrário.} \end{cases}$$

Mostraremos que

$$L(F) = \mathcal{L}(\mathbf{p}) = \prod_{i=1}^n \left(\sum_{j=1}^J \alpha_{ij} p_j \right).$$

Fixe i . Suponha que temos $[a_{k_1}, b_{k_1}], [a_{k_2}, b_{k_2}], \dots, [a_{k_i}, b_{k_i}] \subseteq [L_i, U_i]$, o que significa que $\alpha_{k_1} = \alpha_{k_2} = \dots = \alpha_{k_i} = 1$ e os outros a_{ij} são iguais a zero. Suponha, sem perda de generalidade, que esses intervalos estão ordenados. Assim, tem-se

$$\begin{aligned} F(U_i) - F(L_i) &= F(U_i) - F(b_{k_i}) + F(b_{k_i}) - F(a_{k_i}) + F(a_{k_i}) - F(b_{k_{i-1}}) \\ &\quad + \dots + F(a_{k_1}) - F(L_i) \\ &= F(U_i) - F(b_{k_i}) + p_{k_i} + F(a_{k_i}) - F(b_{k_{i-1}}) + p_{k_{i-1}} \\ &\quad + \dots + p_{k_1} + F(a_{k_1}) - F(L_i) \end{aligned}$$

Como \hat{F} é horizontal (constante) fora dos intervalos $[a_j, b_j]$, pela proposição 2.1, teremos que $\hat{F}(U_i) - \hat{F}(L_i) = p_{k_i} + p_{k_{i-1}} + \dots + p_{k_1} = \sum_{j=1}^J \alpha_{ij} p_j$. Por isso, reescrevemos a função de verossimilhança dependendo apenas dos saltos p_j da seguinte maneira:

$$\mathcal{L}(\mathbf{p}) = \prod_{i=1}^n \left(\sum_{j=1}^J \alpha_{ij} p_j \right). \quad (2.1)$$

Logo, estimaremos F maximizando (2.1) em relação aos parâmetros p_1, p_2, \dots, p_J , sujeitos à restrição $\sum_{j=1}^J p_j = 1$. Qualquer algoritmo de otimização, como o de Newton-Raphson, pode ser usado para resolver este problema. Entre alguns algoritmos se destaca, por sua importância histórica, o algoritmo de auto-consistência proposto por Turnbull (1976), que é, na verdade, uma aplicação do algoritmo EM. Atualmente, o algoritmo mais eficiente para encontrar \hat{F} é o algoritmo iterativo da minorante convexa máxima, que abreviaremos aqui por algoritmo ICM (*iterative convex minorant*). Esse algoritmo usa a teoria de regressões isotônicas na estimação de F . Expomos brevemente aqui a aplicação de regressões isotônicas no Caso 1 de censura intervalar para nos familiarizarmos com o algoritmo ICM posteriormente.

2.2.1 Estimação de F no Caso 1 de censura intervalar

Conforme já colocamos, ocorre o **caso 1** de censura intervalar (ou dados de estado corrente) quando, ao colhermos nossos dados, só sabemos se o real tempo de falha ocorreu antes ou depois do tempo registrado. Isto significa que há presença apenas de censura à direita ou à esquerda em nossos dados.

No caso 1, nossos dados são os pares $(Y_1, \delta_1), (Y_2, \delta_2), \dots, (Y_n, \delta_n)$, em que δ indica

a presença de censura à esquerda ($\delta = 1$) ou de censura à direita ($\delta = 0$). Usando a notação da censura intervalar, temos $Y_i = U_i$, se $\delta = 1$ e $Y_i = L_i$, se $\delta = 0$. Reescrevendo a verossimilhança (1.4), temos

$$\begin{aligned} L(F) &= \prod_{i=1}^n [F(L_i)]^{\delta_i} [1 - F(U_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n [F(Y_i)]^{\delta_i} [1 - F(Y_i)]^{1-\delta_i}. \end{aligned}$$

Para nossa exposição ficar clara, vamos definir alguns conceitos preliminares necessários para a apresentação da regressão isotônica. Dizemos que uma função $f : X \rightarrow \mathbb{R}$, $X \subseteq \mathbb{R}$, é isotônica se $f(x) \leq f(y)$ sempre que $x < y$. Uma função de distribuição F é isotônica, por exemplo.

Sejam $X \subseteq \mathbb{R}$ um conjunto finito e $g, w : X \rightarrow \mathbb{R}$ funções, sendo que $w(x) > 0$ para todo $x \in X$. Então, dizemos que g^* é uma regressão isotônica de g com pesos w se g^* minimiza, na classe de funções isotônicas $f : X \rightarrow \mathbb{R}$, a soma

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x),$$

isto é, $g^* = \arg \min \{ \sum_{x \in X} [g(x) - f(x)]^2 w(x) : f \text{ é isotônica em } X \}$.

No contexto do Caso 1 de censura intervalar, um resultado importante é o

Teorema 2.2.1. *Se $f : X \rightarrow I$ é isotônica, $\Phi : I \rightarrow \mathbb{R}$ é estritamente convexa e g, w e g^* são como na definição acima de regressão isotônica, então g^* é a única função que maximiza*

$$Q(f) = \sum_{x \in X} \{ \Phi(f(x)) + [g(x) - f(x)] \phi(f(x)) \} w(x),$$

em que $\phi(y) = \Phi'(y)$.

Portanto, pelo Teorema acima, a regressão isotônica de g com pesos w maximiza o logaritmo de (2.2.1), se tomarmos $g(Y_i) = \delta_i$, $w(Y_i) = 1$ e $\Phi(y) = y \log(y) + (1 - y) \log(1 - y)$. Assim, $\hat{F} = g^*$. Logo, pode-se calcular \hat{F} usando a função minorante

convexa máxima, apresentada em Groeneboom e Wellner (1992) com

$$\hat{F}(Y_{(m)}) = \max_{i \leq m} \min_{k \geq m} \frac{\sum_{i \leq j \leq k} \delta_{(j)}}{k - i + 1},$$

onde $\delta_{(j)}$ é valor de δ associado a $Y_{(j)}$, a j -ésima estatística de ordem de (Y_1, Y_2, \dots, Y_n) .

Uma interpretação geométrica é que $\hat{F}(Y_{(m)})$ também é dada pela derivada à esquerda no m -ésimo ponto do diagrama de somas acumuladas formado pelos pontos $(0, 0)$ e $P_i = \left(\sum_{j=1}^i w(Y_{(j)}), \sum_{j=1}^i w(Y_{(j)})g(Y_{(j)}) \right) = (i, \sum_{j=1}^i \delta_{(j)})$, $i = 1, 2, \dots, n$.

2.2.2 Algoritmo ICM

Uma vez salientada a importância da teoria de regressões isotônicas para a estimação de F no caso 1, voltamos nossa atenção agora para o caso geral de censura intervalar. Como já vimos, o estimador não paramétrico de máxima verossimilhança é encontrado se calcularmos o vetor \mathbf{p} que maximiza a verossimilhança em (2.1) ou então se encontrarmos diretamente a função \hat{F} que maximiza (1.3). Fazer isso diretamente não é fácil e, por isso, buscamos resultados que possam auxiliar nossa busca.

O estimador não paramétrico de máxima verossimilhança de F é calculado pelo algoritmo ICM. Para apresentá-lo aqui, usaremos a linha de raciocínio exibida em Jongbloed (1998).

De fato, existe um resultado importante para formas quadráticas muito usado na teoria de regressões isotônicas que é relacionado aos nossos interesses. Para colocá-lo aqui, primeiro definamos o conjunto $\mathcal{K} = \{(x_1, x_2, \dots, x_N) \in \mathbb{R}^N : x_1 \leq x_2 \leq \dots \leq x_N\}$, o conjunto $\mathcal{K}_1 = \{\mathbf{x} \in \mathcal{K} : x_1 \geq 0\}$ e a forma quadrática $\psi : \mathbb{R}^N \rightarrow (-\infty, \infty]$, isto é, um função com a forma

$$\psi(x) = \frac{1}{2}(x - y)^T M(x - y), \quad (2.2)$$

em que $y = (y_1, y_2, \dots, y_N) \in \mathbb{R}^N$ e a matriz diagonal positiva definida $M = \text{diag}(m_i)$ são ambos fixos. Seja também o vetor $\hat{x} = \arg \min_{x \in \mathcal{K}} \psi(x)$.

Defina ainda o Diagrama de Somas Acumuladas (DSA) como o conjunto de pontos

P_i em \mathbb{R}^2 , ou seja, $DSA = \{P_i : 0 \leq i \leq N\}$, dados por $P_0 = (0, 0)$ e

$$P_j = \left(\sum_{l=1}^j m_l, \sum_{l=1}^j m_l y_l \right), \quad \text{para } 1 < j < N.$$

Um resultado usado no desenvolvimento do algoritmo ICM (ver Jongbloed (1998)) é o seguinte:

Proposição 2.2. *Sejam \mathcal{K} , \mathcal{K}_1 , ψ e o DSA como definimos acima. Então, a componente \hat{x}_i do vetor \hat{x} é calculada como a derivada à esquerda, avaliada em P_i , da função minorante convexa máxima (isto é, a maior função convexa situada abaixo) do DSA.*

Se o espaço em que minimizamos ϕ é \mathcal{K}_1 , então, as componentes negativas de \hat{x} têm seu valor alterado para zero.

Podemos reescrever a função ψ em (2.2) como $\psi(x) = \frac{1}{2} \sum_{i=1}^N (x_i - y_i)^2 m_i$, para $x \in \mathcal{K}$. Com essa forma, vemos que ψ é, na realidade uma função que atribui a cada elemento $x \in \mathcal{K}$ uma soma quadrada de erros, com pesos m_i , em relação ao vetor $y \in \mathbb{R}^N$, que é fixo. Como os elementos de $\mathcal{K} \subset \mathbb{R}^N$ têm suas suas componentes ordenadas, temos que ψ é a soma quadrática dos erros que minimizamos no problema da regressão isotônica.

A função de verossimilhança dos dados com censura intervalar em (1.3), contudo, não é uma forma quadrática e, portanto, não podemos aplicar diretamente a Proposição 2.2. Por isso, precisamos usar o resultado da Proposição 2.2 de forma mais geral. Assim sendo, seja ϕ uma função definida em \mathcal{K} satisfazendo o seguinte:

- (i) ϕ é convexa, contínua, e atinge seu mínimo em \mathcal{K} em um único ponto \hat{x} ;
- (ii) ϕ é continuamente diferenciável no conjunto $\{x \in \mathbb{R}^N : \phi(x) < \infty\}$.

Uma aproximação de Taylor de segunda ordem, em torno de $x^{(k)}$ para $\phi(x)$, $x \in \mathcal{K}$,

é dada por

$$\begin{aligned}\phi(x) &\approx \phi(x^{(k)}) + (x - x^{(k)})^T \nabla \phi(x^{(k)}) + \frac{1}{2} (x - x^{(k)})^T H_\phi(x^{(k)}) (x - x^{(k)}) \\ &\approx c + \frac{1}{2} \left(x - x^{(k)} + H_\phi^{-1}(x^{(k)}) \nabla \phi(x^{(k)}) \right)^T H_\phi(x^{(k)}) \\ &\quad \times \left(x - x^{(k)} + H_\phi^{-1}(x^{(k)}) \nabla \phi(x^{(k)}) \right),\end{aligned}$$

em que c não depende de x e H_ϕ é a matriz hessiana de ϕ .

Como calcular H_ϕ^{-1} pode tomar muito tempo, é comum substituir a matriz H_ϕ por outro tipo de matriz semidefinida positiva M mais fácil de ser invertida. Por isso, usualmente, toma-se M diagonal com entradas positivas. Isso posto, podemos fazer aproximações locais para ϕ em torno de um ponto $x^{(k)}$ com uma forma quadrática do tipo

$$\begin{aligned}q_\phi(x, x^{(k)}) &= \frac{1}{2} \left(x - x^{(k)} + M(x^{(k)})^{-1} \nabla \phi(x^{(k)}) \right)^T \\ &\quad \times M(x^{(k)}) \left(x - x^{(k)} + M(x^{(k)})^{-1} \nabla \phi(x^{(k)}) \right),\end{aligned}$$

onde $M(x^{(k)})$ é uma matriz diagonal positiva definida que depende de $x^{(k)}$. Toma-se usualmente $M(x^{(k)})$ como a diagonal da matriz hessiana de ϕ . De acordo com a Proposição 2.2, a partir dessa aproximação, podemos obter o ponto $x^{(k+1)}$ que maximiza $q_\phi(\cdot, x^{(k)})$ como a derivada à esquerda do Diagrama de Somas Acumuladas dado pelos pontos $P_0 = (0, 0)$ e

$$P_j = \left(\sum_{l=1}^j m_l^{(k)}, \sum_{l=1}^j \left(m_l^{(k)} x_l^{(k)} - \frac{\partial \phi}{\partial x_l}(x^{(k)}) \right) \right), \quad \text{para } 1 < j < N,$$

sendo $m_l^{(k)}$ o l -ésimo elemento da diagonal de $M(x^{(k)})$.

A idéia do algoritmo ICM é, portanto, tomar $\phi = -\log L(F)$ e, com o procedimento que acabamos de exibir, encontrar o estimador não paramétrico de máxima verossimilhança de F iterativamente, dado um valor inicial $x^{(0)}$. Tornemos essa idéia mais clara. Tomando o logaritmo da expressão (1.3), obtemos a função de

log-verossimilhança para dados com censura intervalar como sendo

$$-\log L(F) = - \sum_{i=1}^n \left(\delta_i \log F(L_i) + \gamma_i \log [F(U_i) - F(L_i)] + (1 - \delta_i - \gamma_i) \log [1 - F(U_i)] \right). \quad (2.3)$$

Olhando essa expressão, vemos que, em princípio, temos $2n$ pontos em que \hat{F} pode ter saltos e, portanto, há $2n$ parâmetros a estimar, que são os valores $F(L_1), F(U_1), \dots, F(L_n), F(U_n)$. Entretanto, ao fazermos uma análise mais detalhada, vemos que alguns desses parâmetros na verdade não influem na log-verossimilhança ou não precisam ser calculados. É o que ocorre quando o menor dos $L_i, L_{(1)}$, corresponde a um $\delta = 0$ ou o maior dos $U_i, U_{(n)}$, corresponde a $\delta + \gamma = 1$. Nesses casos, podemos colocar diretamente $\hat{F}(L_{(1)}) = 0$ e $\hat{F}(U_{(n)}) = 1$. Quando não temos censura intervalar na unidade i , isto é, quando $\gamma_i = 0$, um dos tempos L_i ou U_i não tem influência em (2.3) e, portanto, não é necessário calcular $\hat{F}(L_i)$ ou $\hat{F}(U_i)$. Na prática, isto significa que não é necessário calcular \hat{F} nos pontos L_i , se $\delta_i = \gamma_i = 0$ ou U_i , se $\delta_i = 1$.

Depois de averiguarmos quais dos valores L_i e U_i influenciam de fato na log-verossimilhança (2.3), teremos os tempos $t_1 \leq t_2 \leq \dots \leq t_p, p = n + \sum_1^n \gamma_i$, em que teremos que calcular $\hat{F}(t_j)$. Além disso, podemos particionar o conjunto de índices $\{1, 2, \dots, p\}$ da seguinte maneira:

- I_1 é o conjunto de índices em que há censura à esquerda, isto é, $\delta = 1$;
- I_{2a} é o conjunto dos índices em que há censura intervalar ($\gamma = 1$) e t_j é um tempo do tipo L_i ;
- I_{2b} é o conjunto dos índices em que há censura intervalar ($\gamma = 1$) e t_j é um tempo do tipo U_i ;
- I_3 é o conjunto de índices em que há censura à direita, isto é, $\delta + \gamma = 0$;

Um primeiro ponto a observar é que quando ocorre $\gamma_i = 1$, os pontos L_i e U_i são potenciais pontos de salto de \hat{F} , por termos $L_i \neq 0$ e $U_i < \infty$. Portanto, existem dois índices $k \in I_{2a}$ e $l \in I_{2b}$ tais que $t_k = L_i$ e $t_l = U_i$. Isso implica que $k < l$, pois $L_i < U_i$ e o conjunto $\{t_1, t_2, \dots, t_p\}$ é ordenado.

Mas os conjuntos I_{2a} e I_{2b} são formados justamente pelos pontos L_i e U_i , respectivamente, com $\gamma_i = 1$. Por isso, para todo elemento $k \in I_{2a}$, existe um elemento associado $l \in I_{2b}$, determinado pela relação única entre L_i e U_i . Além disso, sempre ocorre $k < l$.

Portanto, podemos definir uma bijeção $r : I_{2a} \rightarrow I_{2b}$ entre os conjuntos I_{2a} e I_{2b} e, portanto, podemos determinar os elementos de I_{2b} por intermédio dos elementos de I_{2a} . Mais ainda, temos $r(i) > i$ para todo $i \in I_{2a}$. Daí, todo termo $F(U_i) - F(L_i)$ da função de verossimilhança em (1.3) pode ser substituído por $F(t_j) - F(t_{r(j)})$, para algum $j \in I_{2a}$.

Temos que estimar, portanto, o vetor $\boldsymbol{\beta} \in \mathbb{R}^p$, cujas componentes são os valores $\beta_i = F(t_i)$, sendo que ocorre $0 \leq \beta_1 \leq \beta_2 \leq \dots \leq \beta_p$. Feitas todas as considerações acima, podemos reescrever a log-verossimilhança (2.3) como

$$\phi(\boldsymbol{\beta}) = - \sum_{i \in I_1} \log \beta_i - \sum_{i \in I_{2a}} \log [\beta_{r(i)} - \beta_i] - \sum_{i \in I_3} \log [1 - \beta_i]. \quad (2.4)$$

Agora, podemos ver claramente que o problema de encontrar o estimador não paramétrico de máxima verossimilhança de F é, na verdade, uma aplicação do procedimento que descrevemos anteriormente: a maximização de uma função côncava no conjunto \mathcal{K}_1 . Temos, então, que calcular as derivadas parciais de primeira e segunda ordens de $\phi(\boldsymbol{\beta})$ para obtermos os pontos P_j do Diagrama de Somas Acumuladas e, com o resultado da 2.2, achamos o vetor $\boldsymbol{\beta}^{(k+1)}$ que maximiza a aproximação quadrática $q_\phi(\cdot, x^{(k)})$. Assim, temos:

$$\frac{\partial \phi}{\partial \beta_i}(\boldsymbol{\beta}) = \begin{cases} -\beta_i^{-1}, & \text{se } i \in I_1; \\ [\beta_{r(i)} - \beta_i]^{-1}, & \text{se } i \in I_{2a}; \\ -[\beta_i - \beta_{r^{-1}(i)}]^{-1}, & \text{se } i \in I_{2b}; \\ [1 - \beta_i]^{-1}, & \text{se } i \in I_3. \end{cases}$$

e também

$$\frac{\partial^2 \phi}{\partial \beta_i^2}(\boldsymbol{\beta}) = m_i = \begin{cases} \beta_i^{-2}, & \text{se } i \in I_1; \\ [\beta_{r(i)} - \beta_i]^{-2}, & \text{se } i \in I_{2a}; \\ [\beta_i - \beta_{r^{-1}(i)}]^{-2}, & \text{se } i \in I_{2b}; \\ [1 - \beta_i]^{-2}, & \text{se } i \in I_3. \end{cases}$$

Com as derivadas parciais calculadas e lembrando que $\beta_i = F(t_i)$, onde t_i pode ser igual a L_j ou U_j , podemos escrever as componentes dos pontos P_j em termos dos valores $\{L_i, U_i, \delta_i, \gamma_i\}$ da amostra. Mas antes, vamos definir algumas funções que facilitarão a notação aqui empregada. São as seguintes:

$$G_F(t) = \sum_{L_i \leq t} \frac{\delta_i}{[F(L_i)]^2} + \sum_{L_i \leq t} \frac{\gamma_i}{[F(U_i) - F(L_i)]^2} + \sum_{U_i \leq t} \frac{\gamma_i}{[F(U_i) - F(L_i)]^2} + \sum_{U_i \leq t} \frac{1 - \delta_i - \gamma_i}{[1 - F(U_i)]^2},$$

$$W_F(t) = \sum_{L_i \leq t} \frac{\delta_i}{F(L_i)} - \sum_{L_i \leq t} \frac{\gamma_i}{F(U_i) - F(L_i)} + \sum_{U_i \leq t} \frac{\gamma_i}{F(U_i) - F(L_i)} - \sum_{U_i \leq t} \frac{1 - \delta_i - \gamma_i}{1 - F(U_i)},$$

e

$$V_F(t) = W_F(t) + \sum_{t_i \leq t} F(t_i) [G_F(t_i) - G_F(t_{i-1})].$$

Calculamos a primeira componente de P_j :

$$\begin{aligned}
\sum_{l=1}^j m_l &= \sum_{l=1}^j \frac{\partial^2 \phi}{\partial \beta_l^2}(\boldsymbol{\beta}) = \sum_{i \in I_1}^{\beta_i \leq F(t_j)} \beta_i^{-2} + \sum_{i \in I_{2a}}^{\beta_i \leq F(t_j)} [\beta_{r(i)} - \beta_i]^{-2} \\
&\quad + \sum_{i \in I_{2b}}^{\beta_i \leq F(t_j)} [\beta_i - \beta_{r^{-1}(i)}]^{-2} + \sum_{i \in I_3}^{\beta_i \leq F(t_j)} [1 - \beta_i]^{-2} \\
&= \sum_{L_i \leq t_j} \frac{\delta_i}{[F(L_i)]^2} + \sum_{L_i \leq t_j} \frac{\gamma_i}{[F(U_i) - F(L_i)]^2} \\
&\quad + \sum_{U_i \leq t_j} \frac{\gamma_i}{[F(U_i) - F(L_i)]^2} + \sum_{U_i \leq t_j} \frac{1 - \delta_i - \gamma_i}{[1 - F(U_i)]^2} \\
&= G_F(t_j).
\end{aligned}$$

É interessante observarmos, pelo que acabamos de mostrar, que

$$m_j = \sum_{l=1}^j m_l - \sum_{l=1}^{j-1} m_l = G_F(t_j) - G_F(t_{j-1}).$$

Analogamente, pode-se obter a segunda componente de P_j , $P_{j,2}$, como

$$\begin{aligned}
P_{j,2} &= \sum_{l=1}^j \left(m_l \beta_l - \frac{\partial \phi}{\partial \beta_l}(\boldsymbol{\beta}) \right) = - \sum_{l=1}^j \frac{\partial \phi}{\partial \beta_l}(\boldsymbol{\beta}) + \sum_{l=1}^j m_l \beta_l \\
&= \sum_{i \in I_1}^{\beta_i \leq F(t_j)} \beta_i^{-1} - \sum_{i \in I_{2a}}^{\beta_i \leq F(t_j)} [\beta_{r(i)} - \beta_i]^{-1} + \sum_{i \in I_{2b}}^{\beta_i \leq F(t_j)} [\beta_i - \beta_{r^{-1}(i)}]^{-1} \\
&\quad - \sum_{i \in I_3}^{\beta_i \leq F(t_j)} [1 - \beta_i]^{-1} + \sum_{l=1}^j \beta_l m_l \\
&= \sum_{L_i \leq t_j} \frac{\delta_i}{F(L_i)} - \sum_{L_i \leq t_j} \frac{\gamma_i}{F(U_i) - F(L_i)} + \sum_{U_i \leq t_j} \frac{\gamma_i}{F(U_i) - F(L_i)} \\
&\quad - \sum_{U_i \leq t_j} \frac{1 - \delta_i - \gamma_i}{1 - F(U_i)} + \sum_{l=1}^j F(t_l) [G_F(t_l) - G_F(t_{l-1})] \\
&= W_F(t_j) + \sum_{t_l \leq t_j} F(t_l) [G_F(t_l) - G_F(t_{l-1})] \\
&= V_F(t_j)
\end{aligned}$$

Feitos todos esses cálculos, temos todos os elementos necessários para proceder à estimação de F pelo algoritmo ICM. Defina o vetor

$$\mathbf{F}^{(k)} = (F^{(k)}(t_1), F^{(k)}(t_2), \dots, F^{(k)}(t_p)).$$

O algoritmo ICM é o seguinte:

Passo 1: $F^{(0)}(t_j) = \frac{1}{p}$ para $1 \leq j \leq p$;

Passo 2: Construir o Diagrama de Somas Acumuladas com $P_0 = (0, 0)$ e, para $j = 1, 2, \dots, p$, $P_j = (G_{F^{(k)}}(t_j), V_{F^{(k)}}(t_j))$;

Passo 3: Obtenha $F^{(k+1)}(t_j)$ como sendo a derivada à esquerda em $G_{F^{(k)}}(t_j)$ da função minorante convexa máxima do diagrama de somas do passo anterior;

Passo 4: Se $\|F^{(k+1)} - F^{(k)}\| < \varepsilon$, pare. Caso contrário, faça $k = k + 1$ e vá para o passo 2.

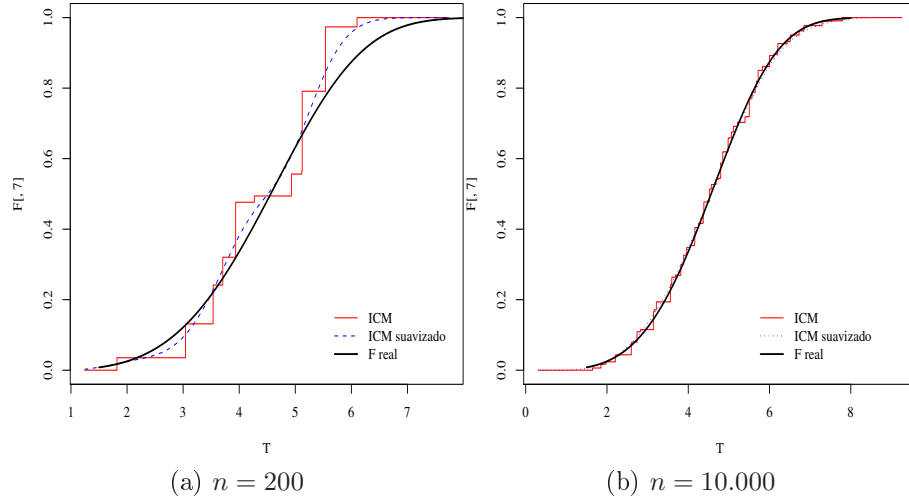


Figura 2.1: Estimativas com algoritmo ICM

O algoritmo ICM possui convergência mais rápida que o algoritmo EM de auto consistência, proposto por Turnbull (1976). Encontrar \hat{F} para o caso univariado é um problema com uma solução bem adequada, visto que o algoritmo ICM é eficiente. Já para o caso da estimação não paramétrica da função de distribuição bivariada, não há ainda um procedimento tão consagrado com o algoritmo ICM. De qualquer forma, muitos avanços têm sido conquistados, sendo que alguns deles são bastante recentes. Por isso, apresentamos no próximo capítulo a estimação não paramétrica da função de distribuição bivariada pelo método de máxima verossimilhança.

Capítulo 3

Estimação da função de distribuição bivariada

A estimação da distribuição conjunta de duas variáveis aleatórias T_1 e T_2 , ambas com censura intervalar, é o nosso objetivo principal neste trabalho. Neste capítulo, estamos interessados mais precisamente em um estimador não paramétrico de máxima verossimilhança de H , função de distribuição de T_1 e T_2 .

Algumas propostas foram apresentadas, entre as quais destacamos os trabalhos de Betensky e Finkelstein (1999) e Gentleman e Vandal (2002). Para encontrar o ENPMV \hat{H} de H , Betensky e Finkelstein (1999) aplicam as idéias de Peto (1973) e Turnbull (1976), que servem para o caso univariado e que foram apresentados no capítulo 2. Dessa forma, seguimos um raciocínio análogo ao caso univariado para obtermos os resultados do caso bivariado, que apresentamos como segue.

Suponha que temos uma amostra de retângulos $R_1 = [L_{11}, U_{11}] \times [L_{21}, U_{21}]$, $R_2 = [L_{12}, U_{12}] \times [L_{22}, U_{22}]$, ..., $R_n = [L_{1n}, U_{1n}] \times [L_{2n}, U_{2n}]$, em que $T_{1i} \in [L_{1i}, U_{1i}]$ e $T_{2i} \in [L_{2i}, U_{2i}]$. A parte da verossimilhança que depende de H para esses dados foi exibida em (1.5) e é dada por

$$L(H) = \prod_{i=1}^n [H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i})], \quad (1.5)$$

lembrando que podemos ter $L_{.i} = 0$ ou $U_{.i} = \infty$.

Ao observarmos a expressão (1.5) acima, vemos que o valor da verossimilhança

$L(H)$ não depende do comportamento de H no interior de cada retângulo R_i . Por isso, o estimador não paramétrico de máxima verossimilhança \hat{H} pode não ser único. Além disso, procuramos \hat{H} no espaço de todas as funções de distribuição bivariadas, ou seja, a dimensão de nosso problema é infinito. Precisamos, então, de resultados que possam simplificar nossa busca.

Como no caso univariado, vamos definir um conjunto importante de retângulos $\xi_1, \xi_2, \dots, \xi_J \subset \mathbb{R}^2$, que chamaremos de *interseções máximas*. Usaremos a definição de Wong e Yu (1999) para as interseções máximas: $\xi_j \neq \emptyset$ é uma interseção máxima se é uma interseção finita dos R_i tal que, para todo i tem-se $\xi_j \cap R_i = \xi_j$ ou $\xi_j \cap R_i = \emptyset$. Na verdade, explicando em palavras, ξ_j é a menor interseção de retângulos da amostra $R_{k_1}, R_{k_2}, \dots, R_{k_l}$, $k_l \leq n$, cuja interseção é não vazia, isto é, $R_{k_s} \cap R_{k_i} \neq \emptyset$.

Calcular as interseções máximas $\xi_1, \xi_2, \dots, \xi_J$ é computacionalmente dispendioso, pois, por definição, temos que verificar quais dos retângulos R_i se interseccionam. A proposição de algoritmos para esta tarefa tem sido tema de alguns trabalhos na literatura, dentre os quais destacamos os algoritmos de Betensky e Finkelstein (1999), Gentleman e Vandal (2002), Bogaerts e Lesaffre (2004) e Maathuis (2005). O mais recente e rápido é o algoritmo *HeightMap* de Maathuis (2005), que usa a idéia de “mapas de altura” para encontrar as interseções máximas. O algoritmo HeightMap calcula o número de retângulos R_i da amostra que passam por um ponto (x, y) do plano. De fato, as interseções máximas ξ_j são as regiões de máximo local desse mapa de alturas.

Há propriedades importantes das interseções máximas. Primeiramente, temos que ξ_j é também um retângulo, pois é uma interseção de retângulos e, portanto, podemos escrever $\xi_j = [a_j, b_j] \times [c_j, d_j]$. Cabe observarmos que ξ_j pode ser um ponto ou um intervalo, o que quer dizer que ξ_j pode ter dimensão menor que dois.

Um segundo fato é fundamental: Gentleman e Vandal (2002) mostram que o estimador não paramétrico de máxima verossimilhança de H só pode ter crescimento nos retângulos $\xi_1, \xi_2, \dots, \xi_J$. Isto significa que se um estimador H_1 de H atribui salto positivo a regiões disjuntas de $\xi_1, \xi_2, \dots, \xi_J$, então H_1 não pode ser um estimador de máxima verossimilhança de H . Por isso, o estimador não paramétrico de máxima verossimilhança \hat{H} é constante fora de $\bigcup_{j=1}^J \xi_j$. Com isso em mente, estimar os valores $p_j = H(b_j, d_j) - H(b_j, c_j) - H(a_j, d_j) + H(a_j, c_j)$ é fundamental para que nossos

propósitos sejam alcançados. É importante enfatizarmos ainda que nada impede que alguns dos p_j possam ser iguais a zero.

Ainda em analogia com o caso univariado, se H é uma função de distribuição que é constante fora dos ξ_j , pode-se mostrar que

$$H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i}) = \sum_{j=1}^J \alpha_{ij} p_j,$$

onde

$$\alpha_{ij} = \begin{cases} 1, & \text{se } \xi_j \subseteq R_i; \\ 0, & \text{caso contrário.} \end{cases}$$

Feitas todas essas considerações, podemos restringir nossa busca de \hat{H} à classe de funções de distribuição que são constantes fora dos ξ_j . Definindo $\mathbf{p} = (p_1, p_2, \dots, p_J)$, obtemos a seguinte verossimilhança:

$$\begin{aligned} L(H) &= \prod_{i=1}^n [H(U_{1i}, U_{2i}) - H(U_{1i}, L_{2i}) - H(L_{1i}, U_{2i}) + H(L_{1i}, L_{2i})] \\ &= \prod_{i=1}^n \left(\sum_{j=1}^J \alpha_{ij} p_j \right) \\ &= L_1(\mathbf{p}). \end{aligned} \tag{3.1}$$

Portanto, ao calcularmos os ξ_j , temos os p_j bem definidos e, portanto, temos J parâmetros a estimar. Logo, reduzimos a dimensão infinita de nosso problema para uma dimensão finita J . Por isso, é comum dizer que esse passo da estimação de H é o passo da redução de dimensão.

Com os valores de p_j estimados, a estimativa de H no ponto (x, y) é calculada da seguinte maneira (ver Wong e Yu (1999)):

$$\hat{H}(x, y) = \sum_{\xi_j \subseteq [0, x] \times [0, y]} \hat{p}_j. \tag{3.2}$$

Para estimarmos o vetor \mathbf{p} , temos que encontrar o argumento máximo de $L_1(\mathbf{p})$, como em (3.1). Como \hat{H} não terá saltos fora de $\bigcup_{j=1}^J \xi_j$, segue que $\sum_{j=1}^J p_j = 1$. Assim sendo, nosso problema é o seguinte:

$$\max \mathcal{L}(\mathbf{p}) = \max \log L_1(\mathbf{p})$$

sujeito a:

$$p_1 + p_2 + \dots + p_J = 1,$$

$$p_j \geq 0, \quad j = 1, 2, \dots, J.$$

Temos, portanto, um problema não linear de otimização côncava com restrições lineares. Gentleman e Geyer (1994) discutem esse problema, analisando as condições de Karush - Kuhn - Tucker (KKT), afirmando que essas condições são necessárias e suficientes para a otimalidade de $\mathcal{L}(\mathbf{p})$. De fato, os multiplicadores de Lagrange envolvidos nas condições de KKT são importantes para determinar a unicidade de \hat{p} , como veremos adiante.

A unicidade de \hat{p} depende da concavidade de \mathcal{L} , que é determinada por sua matriz Hessiana no ponto \hat{p} , $\mathcal{H}(\hat{p})$. Se $\mathcal{H}(\hat{p})$ é negativa definida, então \mathcal{L} é estritamente côncava e, portanto, \hat{p} é único. Gentleman e Geyer (1994) mostram que $\mathcal{H}(\hat{p}) = A'DA$, em que $A = (\alpha_{ij})$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, J$ e D é matriz diagonal com elementos $-1/\sum_{j=1}^J \alpha_{ij}p_j$, $i = 1, 2, \dots, n$. Assim, se $\text{posto}(A) = J$, então $\mathcal{H}(\hat{p})$ é de posto completo e segue que \hat{p} é único.

Essa condição não é necessária e, dessa forma, pode ocorrer de \mathcal{L} não ser estritamente côncava, mas \hat{p} ser único. Com efeito, Gentleman e Geyer (1994) dão uma condição para que \hat{p} seja único, mesmo que \mathcal{L} não seja estritamente côncava. Seja A_2 a matriz formada pelas colunas da matriz A correspondentes aos multiplicadores de Lagrange nulos. Gentleman e Geyer (1994) mostram que se o posto de A_2 é igual ao seu número de colunas, então \hat{p} é único.

O passo do cálculo de \hat{p} na estimação de H é chamado de passo da otimização. É importante lembrarmos também que estimamos a matriz de variâncias e covariâncias de \hat{p} por

$$\widehat{Var}(\hat{p}) = -[\mathcal{H}(\hat{p})]^{-1},$$

o que nos permite calcular a variância de \hat{H} , que é uma combinação linear de p_1, p_2, \dots, p_J , conforme (3.2).

O cálculo de \hat{H} está implementado no pacote MLEcens do software R, que usa o

algoritmo HeightMap para calcular as interseções máximas $\xi_1, \xi_2, \dots, \xi_J$ e, em seguida, estima p maximizando $\mathcal{L}(p)$. Segundo a documentação desse pacote, o passo de otimização é feito usando uma combinação de programação quadrática sequencial e o algoritmo de redução do suporte de Groeneboom, Jongbloed e Wellner (2008).

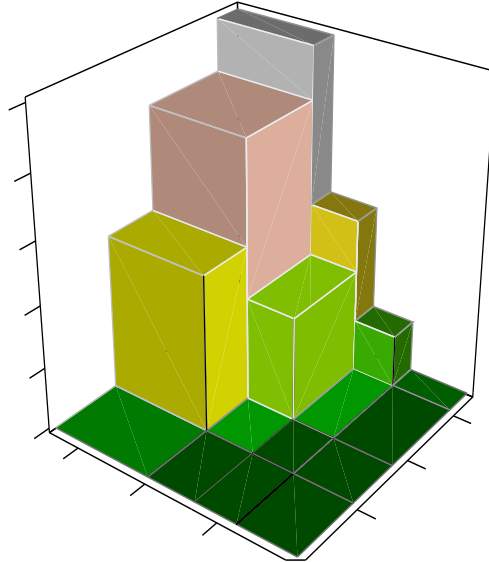


Figura 3.1: Estimador não paramétrico de máxima verossimilhança de H

Neste capítulo, vimos como estimar H diretamente a partir dos dados observados. Entretanto, um dos objetos de nosso estudo é a estimação de H usando estimativas das funções de distribuição marginais de T_1 e T_2 . Tal abordagem é possível graças aos uso de modelos de cópulas, que é o assunto do próximo capítulo.

Capítulo 4

Cópulas

O capítulo 2 foi destinado à exposição do estimador não paramétrico de máxima verossimilhança da função de distribuição univariada da variável aleatória T_1 (ou T_2). Contudo, nosso objetivo central é estimar a função de distribuição conjunta do vetor aleatório (T_1, T_2) , que definimos anteriormente como H . De fato, é possível estimar H usando a informação contida nas funções de distribuição marginais por meio de modelos de cópulas, que apresentamos neste capítulo.

A idéia dos modelos de cópulas é simples: obter a função de distribuição conjunta H de T_1 e T_2 a partir de suas funções de distribuição marginais F e G , respectivamente, por meio de uma função que “ligue” F e G . Essa função que liga F e G é que denominamos cópula. Podemos fazer isso por causa de um resultado muito importante, chamado Teorema de Sklar, que garante que essa cópula sempre existe e, se T_1 e T_2 são contínuas, a cópula que liga F e G é única.

O interesse por modelos de cópulas vem crescendo bastante nos últimos anos. Isso se deve muito à facilidade para implementação computacional e praticidade em aplicações. Bem sabemos que analisar variáveis conjuntamente pode nos levar a conclusões muito mais ricas que aquelas a que chegamos em análises em que consideramos as variáveis em estudo isoladamente. A execução destas análises conjuntas pode ser facilitada pela aplicação de modelos de cópulas.

Apresentaremos aqui nesse capítulo a teoria de que precisamos para fazermos a aplicação de modelos de cópulas, tendo como referência básica Nelsen (2006). Para tanto, vamos estabelecer alguns conceitos necessários e expor alguns resultados im-

portantes. Começamos com a definição formal de cópula.

Definição 4.0.1. *Uma cópula bidimensional é uma função $C : I^2 \rightarrow I$, $I = [0, 1]$, com as seguintes propriedades:*

1. $C(u, 0) = C(0, v) = 0$, $C(u, 1) = u$ e $C(1, v) = v$ para todos $u, v \in I$.
2. Para todo retângulo $[u_1, v_1] \times [u_2, v_2] \subseteq I^2$, temos

$$C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0.$$

□

Analisando a definição de cópula, podemos verificar que uma cópula bidimensional é, na verdade, uma função de distribuição conjunta bivariada cujas marginais são uniformes no intervalo $[0, 1]$. Usaremos nesse texto o termo cópula para fazermos referência a cópulas bidimensionais, ressaltando que uma cópula pode ser definida em qualquer dimensão finita.

Posta essa definição, pode-se mostrar alguns fatos a respeito de uma cópula C qualquer. O primeiro fato é que uma cópula é sempre não decrescente em cada um de seus argumentos. Isto quer dizer que, se $u_1 \leq u_2$, $v_1 \leq v_2$ e t pertencem a I , então

$$C(u_1, t) \leq C(u_2, t) \text{ e } C(t, v_1) \leq C(t, v_2).$$

Daí segue ainda que $C(u_1, v_1) \leq C(u_2, v_2)$.

O segundo fato concerne a continuidade de C : vale, para quaisquer $(u_1, v_1), (u_2, v_2) \subseteq I^2$, que:

$$|C(u_1, v_1) - C(u_2, v_2)| \leq |u_1 - u_2| + |v_1 - v_2|,$$

isto é, C é uniformemente contínua.

Em terceiro lugar, para qualquer cópula C e para todo vetor $(u, v) \in I^2$, tem-se

$$W(u, v) \stackrel{\text{def}}{=} \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) \stackrel{\text{def}}{=} M(u, v).$$

As funções W e M acima também são cópulas e são chamadas de limite inferior

e limite superior de Fréchet-Hoeffding, respectivamente. Outra cópula muito importante é a cópula produto, $\Pi(u, v) = uv$.

Os três fatos acima são propriedades básicas de cópulas que decorrem de sua definição. Definido o que é uma cópula, podemos passar para suas aplicações em variáveis aleatórias. Conforme já colocamos aqui, queremos obter uma forma de ligar as funções de distribuição marginais à função de distribuição bivariada. Nesse contexto, o principal resultado é o

Teorema 4.0.2 (Teorema de Sklar). *Sejam X e Y variáveis aleatórias com funções de distribuição F e G , respectivamente, e função de distribuição conjunta H . Então, existe uma cópula C tal que*

$$H(x, y) = C(F(x), G(y)), \quad \forall (x, y) \in \mathbb{R}^2.$$

Se X e Y são contínuas, então C é única. Caso contrário, C é unicamente determinada em $Im(F) \times Im(G)$. \square

Um corolário imediato do Teorema de Sklar é que podemos encontrar C diretamente pela relação

$$C(u, v) = H(F^{-1}(u), G^{-1}(v))$$

para todo $(u, v) \in I^2$. Chamamos a cópula associada à distribuição conjunta de duas variáveis aleatórias X e Y de “cópula de X e Y ”.

Com o Teorema de Sklar, podemos reescrever a verossimilhança bivariada (1.5) em termos de uma cópula. Suponha que a cópula de T_1 e T_2 é C_α , que depende do vetor de parâmetros α . Então, podemos reescrever (1.5) como

$$L(H) = \prod_{i=1}^n \left[C_\alpha(F(U_{1i}), G(U_{2i})) - C_\alpha(F(U_{1i}), G(L_{2i})) \right. \\ \left. - C_\alpha(F(L_{1i}), G(U_{2i})) + C_\alpha(F(L_{1i}), G(L_{2i})) \right]. \quad (4.1)$$

Pela expressão da função de verossimilhança (4.1) acima, vemos que podemos estimar H a partir de estimativas das marginais de T_1 e T_2 , F e G , respectivamente,

por meio da maximização de (4.1) em relação ao parâmetro α . É isso que iremos fazer, uma vez que podemos obter estimativas de F e G pelo método de imputação múltipla ou pelo estimador não paramétrico de máxima verossimilhança.

É preciso notar que as estimativas de F e G devem ser avaliadas nos pontos L_i e U_i . Contudo, conforme já colocamos anteriormente, tanto o método de imputação múltipla como o estimador não paramétrico de máxima verossimilhança fornecem como estimativa funções escada para F (e G). Também por esse motivo, usaremos núcleo estimadores para obtermos estimativas suavizadas de F e G . Apresentamos os núcleo estimadores no próximo capítulo.

Como temos mais de um método para estimar a função de distribuição H , surge o interesse em fazer a comparação para saber qual método é mais eficaz. Para isso, estudos de simulação foram realizados. Por isso, apresentamos na próxima seção um método de simulação de vetores aleatórios bivariados usando modelos de cópulas.

4.1 Simulação com cópulas

Sabemos que a simulação de dados tem importância primordial no estudo de novas técnicas e métodos estatísticos. Com dados simulados, podemos saber se estimadores estudados possuem algumas propriedades desejadas, visto que conhecemos completamente a distribuição dos dados que simulamos.

A geração univariada de uma amostra de uma variável aleatória é facilmente executada pelo método da transformação integral de probabilidade. Dado $Z \sim U(0, 1)$ e uma função de distribuição F , fazendo $T_1 = F^{-1}(Z)$, teremos que T_1 será uma variável aleatória com função de distribuição F .

Podemos usar também a transformada integral de probabilidade para gerarmos uma amostra (x, y) de um vetor aleatório (T_1, T_2) cuja função de distribuição conjunta é H e que tem marginais F e G . Para tanto, precisamos de um par (Z, V) de variáveis aleatórias $Z \sim U(0, 1)$ e $V \sim U(0, 1)$ para que façamos $T_1 = F^{-1}(Z)$ e $T_2 = G^{-1}(V)$. Entretanto, esse par (Z, V) precisa manter a estrutura de dependência, se ela existe, do vetor aleatório (T_1, T_2) .

Logo, a questão a ser resolvida é saber como obter um par (Z, V) de variáveis aleatórias $Z \sim U(0, 1)$ e $V \sim U(0, 1)$ que preserva a dependência de (T_1, T_2) . Isso pode

ser feito com o uso de cópulas e o método da distribuição condicional. Suponhamos que (T_1, T_2) tem cópula C associada a H . Se obtivermos um par (Z, V) com função de distribuição conjunta C , podemos obter a amostra (x, y) com o uso da transformada integral de probabilidade. Assim,

$$\begin{aligned} P(T_1 \leq x, T_2 \leq y) &= P(F^{-1}(Z) \leq x, G^{-1}(V) \leq y) = P(Z \leq F(x), V \leq G(y)) \\ &= C(F(x), G(y)) = H(x, y). \end{aligned}$$

Para gerarmos a amostra (x, y) de (T_1, T_2) pelo método da distribuição condicional, necessitamos da distribuição de V dado $Z = z$. Com isso, dado $Z = z$, obtemos a amostra v de V . Suponhamos que a densidade de Z é f_Z e a densidade conjunta de (Z, V) é f . A função de distribuição de $(V|Z = z)$, $F_{V|Z}(v|z)$, é

$$\begin{aligned} F_{V|Z}(v|z) &= \int_0^v f(t|z) dt = \int_0^v \frac{f(z, t)}{f_Z(z)} dt = \int_0^v \frac{f(z, t)}{1} dt \\ &= \int_0^v \frac{\partial^2 C}{\partial t \partial z}(z, t) dt = \frac{\partial C}{\partial z}(z, v). \end{aligned}$$

Assim, para gerarmos o par (x, y) do vetor aleatório (T_1, T_2) , usamos o método da distribuição condicional com o seguinte algoritmo:

1. Gere $z \sim U(0, 1)$ e $t \sim U(0, 1)$ independentes.
2. Faça $v = F_V^{-1}(t|z) = \frac{\partial C}{\partial z}^{-1}(z, t)$.
3. Calcule $x = F^{-1}(z)$ e $y = G^{-1}(v)$.

O par desejado é (x, y) , que tem distribuição conjunta H , unicamente determinada pela cópula C .

Esse procedimento apresentado gera observações exatas para o vetor (T_1, T_2) . Todavia, estamos trabalhando com dados censurados. Assim sendo, geraremos também amostras para as variáveis aleatórias L_1, U_1, L_2 e U_2 , que são independentes de T_1 e T_2 . Em posse dessa amostra de L_1, U_1, L_2, U_2, T_1 e T_2 , podemos definir as variáveis $\delta_1, \gamma_1, \delta_2$ e γ_2 . Com isso, obtemos nossa amostra simulada de dados com censura intervalar, composta pelas quádruplas $(L_{1i}, U_{1i}, \delta_{1i}, \gamma_{1i})$ e $(L_{2i}, U_{2i}, \delta_{2i}, \gamma_{2i})$.

4.2 Medidas de dependência

A relação entre duas variáveis aleatórias é tema de estudo de muitas pesquisas. Geralmente, quer-se determinar como uma variável X influencia outra variável Y , ou seja, como é a dependência entre elas. Caracterizar o conceito de dependência é difícil, a não ser no caso da dependência linear. Na verdade, há dependência sempre que a função de distribuição conjunta de (X, Y) não é o produto das funções de distribuição marginais. Por isso, a única forma de caracterizar completamente a dependência é descrevendo a distribuição conjunta de (X, Y) .

Da mesma maneira, é difícil mensurar o grau de dependência entre (X, Y) , quando esta não é linear. Contudo, existem algumas medidas que buscam resumir, de alguma forma, a dependência entre duas variáveis. Essas medidas geralmente estão ligadas aos parâmetros da cópula, pois a cópula descreve a distribuição conjunta de (X, Y) .

As medidas de dependências mais famosas são o τ de Kendall e ρ de Spearman. Existem outras medidas, que não serão abordadas aqui, mas podem ser encontradas em Nelsen (2006). Exibiremos aqui essas duas medidas e algumas relações delas com a cópula de (X, Y) .

Ressaltamos que não há estatísticas que estimem τ ou ρ para dados com censura intervalar e, portanto, apresentamos a definição populacional dessas quantidades. Para serem construídas, ambas as medidas usam a idéia de concordância, que é, em poucas palavras, a tendência de que os valores de X cresçam quando os valores de Y crescem (ou X decresce quando Y decresce). Há discordância quando X tende a crescer quando Y decresce (ou X decresce quando Y cresce).

4.2.1 τ de Kendall

Sejam (X_1, Y_1) e (X_2, Y_2) dois vetores aleatórios iid com função de distribuição H conjunta e cópula C . Definimos o τ de Kendall por

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

Na verdade, τ é igual à probabilidade de concordância menos a probabilidade de discordância entre as v.a.'s X e Y . Por isso, se $\tau = -1$, há discordância total entre

X e Y e, assim, uma dependência negativa. Por outro lado, se $\tau = 1$, então ocorre o contrário: concordância total e, portanto, dependência positiva.

A relação importante entre τ e as cópulas é a seguinte:

Proposição 4.1. *Sejam X e Y duas variáveis contínuas cuja cópula é C . Então,*

$$\tau = \tau_{X,Y} = 4 \int_0^1 \int_0^1 C(u,v) dC(u,v) - 1.$$

Por essa relação, podemos ver que o valor de τ provavelmente vai depender dos parâmetros da cópula C . Pode acontecer também que a integral dupla não possa ser avaliada analiticamente. Esse é o caso do modelo de cópula Frank com parâmetro α (que apresentaremos adiante), em que temos

$$\tau = 1 - \frac{4}{\alpha} \left[1 - \frac{1}{\alpha} \int_0^\alpha \frac{t}{e^t - t} dt \right].$$

Ocorre sempre $-1 \leq \tau \leq 1$, sendo que valores positivos indicam dependência positiva e valores negativos indicam dependência negativa, no sentido que aqui expusimos.

4.2.2 ρ de Spearman

Suponha agora que (X_1, Y_1) , (X_2, Y_2) e (X_3, Y_3) são três vetores aleatórios iid cuja função de distribuição é H e cópula C . Define-se o ρ de Spearman por

$$\rho = \rho_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]).$$

A construção de ρ é parecida com a do τ de Kendall. Nelsen (2006) explica que a diferença é que ρ é a diferença das probabilidades de concordância e discordância de (X_1, Y_1) e (X_2, Y_3) , respectivamente.

O resultado que relaciona ρ e a cópula C é a

Proposição 4.2. *Sejam X e Y duas variáveis contínuas cuja cópula é C . Então,*

$$\rho = \rho_{X,Y} = 12 \int_0^1 \int_0^1 C(u,v) dudv - 3.$$

Assim como acontece com τ , temos que $-1 \leq \rho \leq 1$, sendo que se $\rho > 0$, então há tendência de crescimento de X e Y na mesma direção. Se $\rho < 0$ a tendência de crescimento de X e Y é oposta.

4.3 Famílias de Cópulas

Já definimos formalmente o que é uma cópula e mostramos quais serão as suas aplicações concretas no problema que estamos estudando. Contudo, uma definição geral pode não ser diretamente aplicável e, assim sendo, precisamos de algumas funções que satisfazem a definição de cópula.

Muitas funções já foram definidas de forma a satisfazer os requisitos de uma cópula, sendo que uma importante classe de funções desse tipo é a classe de cópulas Arquimedianas. Esta classe de cópulas se destaca por sua simplicidade e por sua grande capacidade de representar várias estruturas de dependência.

Definição 4.3.1. *Uma cópula Arquimediana $C : I^2 \rightarrow I$, $I = [0, 1]$ tem a seguinte forma:*

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v)),$$

onde $\phi : I \rightarrow [0, \infty]$ é uma função contínua, convexa e estritamente decrescente tal que $\phi(1) = 0$. □

Portanto, se encontrarmos uma função ϕ com as propriedades acima, podemos gerar um cópula Arquimediana. Quando isso ocorre com uma cópula C , dizemos que C é uma cópula Arquimediana gerada por ϕ . Podemos também tomar ϕ dependendo de algum vetor de parâmetros α , se for necessário. A cópula produto Π , que é gerada se tomarmos $\phi(t) = -\ln t$, é um exemplo simples de cópula Arquimediana.

Três propriedades algébricas, que seguem da definição de uma cópula Arquimediana C gerada por ϕ , são as seguintes:

1. C é simétrica, isto é, $C(u, v) = C(v, u)$ para todos $u, v \in I$;
2. C é associativa, isto é, $C(C(u, v), w) = C(u, C(v, w))$ para todos $u, v, w \in I$;
3. Se $c > 0$ é uma constante, então $c\phi$ é também um gerador de C .

As cópulas Arquimedianas possuem outras muitas propriedades e resultados interessantes, que não abrangem o escopo de nossas aplicações. O leitor interessado em tais resultados pode consultar Nelsen (2006).

Por exemplo, um resultado interessante que relaciona cópulas Arquimedianas e τ de Kendall é a

Proposição 4.3. *Sejam X e Y duas variáveis contínuas cuja cópula C é Arquimediana, gerada por ϕ . Então, o valor de τ de Kendall é dado por*

$$\tau = 1 + 4 \int_0^1 \frac{\phi'(t)}{\phi(t)} dt.$$

Para nossas aplicações, faremos uso de famílias de cópulas com apenas um parâmetro α . Apresentamos agora algumas famílias mais famosas de cópulas Arquimedianas na tabela 4.1, com seu respectivo espaço paramétrico Ω :

Tabela 4.1: Algumas famílias clássicas de cópulas Arquimedianas

Nome da família	$\phi(t)$	C	Ω
Clayton	$\frac{1}{\alpha}(t^{-\alpha} - 1)$	$[\max(u^{-\alpha} + v^{-\alpha} - 1, 0)]^{-\frac{1}{\alpha}}$	$[-1, \infty) \setminus \{0\}$
Frank	$-\ln \left[\frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1} \right]$	$-\frac{1}{\alpha} \ln \left[1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1} \right]$	$(-\infty, \infty) \setminus \{0\}$
Gumbel-Hougaard	$(-\ln t)^\alpha$	$\exp \left\{ - [(-\ln u)^\alpha + (-\ln v)^\alpha]^{\frac{1}{\alpha}} \right\}$	$[1, \infty)$
Gumbel-Barnett	$\ln(1 - \alpha \ln t)$	$uv \exp(-\alpha \ln u \ln v)$	$(0, 1]$
Ali-Mikhail-Haq	$-\ln \left[\frac{1 - \alpha(1-t)}{t} \right]$	$\frac{uv}{1 - \alpha(1-u)(1-v)}$	$[-1, 1)$

Nem todos os modelos de cópulas pertencem à classe Arquimediana. Por exemplo, a família de cópulas de Farlie-Gumbel-Morgenstern, dada por

$$C_\alpha(u, v) = uv + \alpha uv(1-u)(1-v),$$

não pertence à classe Arquimediana, pois não é associativa.

Naturalmente, existem outras muitas famílias de cópulas Arquimedianas que não estão relacionadas na tabela 4.1. Além disso, existem muitos outros modelos de cópulas que também não são Arquimedianos.

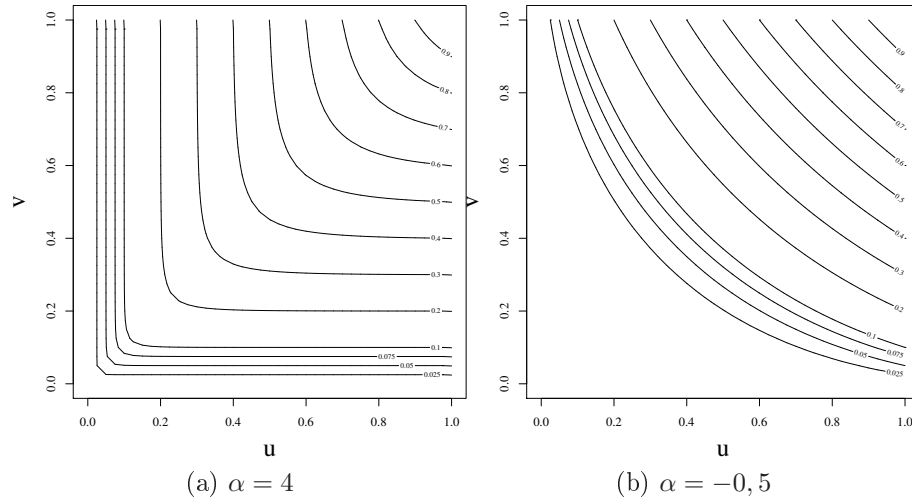


Figura 4.1: Modelo de cópula Clayton

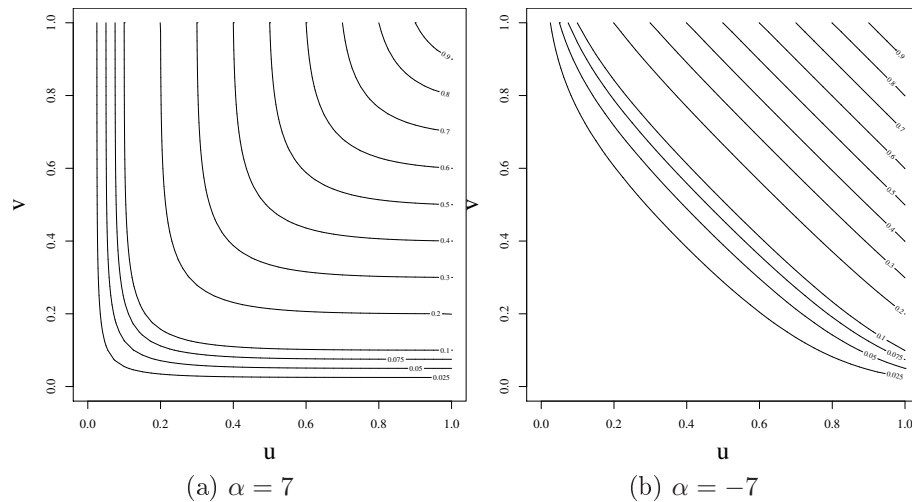


Figura 4.2: Modelo de cópula Frank

Em nossas simulações, usamos as famílias de cópulas de Clayton e Frank, combinadas com a escolha das marginais para T_1 e para T_2 . A geração de pares advindos do modelo de Clayton, entre outros modelos, está implementada no pacote QRMLib do software R. Para a geração de pares da cópula Frank, usamos o procedimento descrito na seção 4.1. Nesse último caso, dados dois números aleatório u e t , encontramos o

par v de u por

$$v = -\frac{1}{\alpha} \ln \left(1 + \frac{te^{\alpha u}(e^{-\alpha} - 1)}{1 + t(e^{\alpha u} - 1)} \right).$$

Apresentadas as aplicações dos modelos de cópulas em nossos problemas, faremos agora a exposição do método de suavização dos estimadores da função de distribuição univariada e bivariada por meio de núcleo estimadores, que serão necessários para o cálculo da verossimilhança (4.1).

Capítulo 5

Suavização

Conforme vimos nos capítulos 2 e 3, os estimadores não paramétricos de máxima verossimilhança das funções de distribuição univariada F e bivariada H são indefinidos no interior das interseções máximas, que definem o suporte desses respectivos estimadores. Isso ocorre porque apenas podemos estimar qual o valor do salto p_j , mas não temos como saber como essa massa p_j se distribui no interior da interseção máxima correspondente.

Algumas convenções podem ser feitas para definirmos onde se dará o incremento p_j em \hat{F} ou \hat{H} . No caso univariado, podemos convencionar, por exemplo, que o salto p_j se dará totalmente no ponto $x = b_j$. Isso nos dará uma estimativa “otimista” da curva de sobrevivência, se T é o tempo de vida de pessoas com uma doença. Por esse motivo, pode ser definido que o salto se dá no ponto $x = a_j$. De qualquer forma, vemos que existem infinitas maneiras de distribuímos p_j no interior da interseção máxima e essa distribuição é arbitrária, visto que a verossimilhança não depende dessa distribuição escolhida.

Tendo em vista os argumentos apresentados acima e considerando também que \hat{F} e \hat{H} são constantes nas regiões disjuntas das interseções máximas, estamos interessados em obter uma estimativa suave de F e de H . Algumas propostas para dados com censura intervalar foram apresentadas na literatura, sendo a maioria para o caso univariado. No caso univariado, lembramos, por exemplo, os trabalhos de Kooperberg e Stone (1992), Pan (2000b) e Braun, Duchesne e Stafford (2005). Já no caso bivariado, encontramos apenas contribuições para a suavização na presença de censura bivariada

à direita, como no trabalho de Bae et al. (2005).

Para o caso univariado, usaremos a suavização dada por núcleo estimadores como em Pan (2000b), uma vez que este estimador usa a informação do estimador de F para que a suavização seja realizada. Usar a informação do estimador de F na suavização é algo desejável, visto que estamos comparando justamente estes estimadores. Para o caso bivariado, iremos usar o estimador análogo àquele do caso de censura univariada. Apresentamos agora esses estimadores, começando com o caso univariado.

5.1 Caso Univariado

Podemos descrever completamente o comportamento de uma variável aleatória X se conhecemos sua função densidade f , sua função de distribuição F ou a função característica φ . Por isso, é tão importante estimar alguma dessas três funções. No entanto, nenhuma delas é diretamente observável e, por isso, não conhecemos sua forma ou propriedades.

Mesmo que f ou F não sejam diretamente observáveis, temos informações sobre estas funções quando obtemos amostras de X e, assim sendo, podemos estimá-las. Um método bastante difundido e eficiente é a estimação de f ou F com o uso do núcleo estimador. Primeiramente, vamos definir esse estimador para dados exatos e, em seguida, apresentaremos o núcleo estimador para dados com censura intervalar.

Suponha que temos observações exatas X_1, X_2, \dots, X_n de uma variável aleatória X . O núcleo estimador de f no ponto $x \in \mathbb{R}$ é definido por

$$\tilde{f}_k(x; h) = \int_{-\infty}^{\infty} k_h(x - t) d\hat{F}(t) = \frac{1}{n} \sum_{i=1}^n k_h(x - X_i),$$

em que k_h é uma função densidade que depende do parâmetro de escala h , que chamamos de janela. Integrando em relação a x , obtemos o núcleo estimador de F , dado por

$$\tilde{F}_K(x; h) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i).$$

onde K_h é a função de distribuição associada a k_h . As funções k_h e K_h são denominadas funções núcleo.

No caso de dados censurados em intervalos, temos uma amostra $[L_1, U_1], [L_2, U_2], \dots, [L_n, U_n]$ e precisamos supor que já temos uma estimativa discreta de F , cujos incrementos se dão nos pontos $0 = X_0 < X_1 < X_2 < \dots < X_r$. Com isso, podemos definir o núcleo estimador de f como (ver Pan (2000b)):

$$\tilde{f}_k(x; h) = \int_{-\infty}^{\infty} k_h(x-t) d\hat{F}(t) = \sum_{i=1}^r k_h(x - X_i) [\hat{F}(X_i) - \hat{F}(X_{i-1})]$$

e, conseqüentemente, obtemos o núcleo estimador de F por

$$\tilde{F}_K(x; h) = \sum_{i=1}^r K_h(x - X_i) [\hat{F}(X_i) - \hat{F}(X_{i-1})],$$

onde $\hat{F}(X_i)$ é uma estimativa de F no ponto X_i , sendo que $\hat{F}(X_0) = 0$ sempre. As estimativas $\hat{F}(X_i)$ podem ser obtidas, por exemplo, pelo estimador não paramétrico de máxima verossimilhança ou pelo método de imputação múltipla de Pan (2000), que foram apresentados no capítulo 2.

É interessante notarmos que o núcleo estimador na censura intervalar é, na verdade, uma generalização do núcleo estimador para dados não censurados. Isso ocorre porque no caso de dados não censurados, o estimador de máxima verossimilhança de $F(X_k)$ é a função de distribuição empírica e, portanto, temos que $\hat{F}(X_k) = k/n$. Dessa forma, temos sempre que $\hat{F}(X_i) - \hat{F}(X_{i-1}) = 1/n$.

Podemos ver também que o estimador acima depende da escolha de K_h e h , além do fato de necessitarmos de estimativas de $F(X_i)$. Wand e Jones (1995) afirmam que, geralmente, a escolha de K_h não afeta muito a estimativa e é, de certa forma, arbitrária. É usual escolher k_h simétrica em torno de zero e satisfazendo as seguintes propriedades:

1. $\int k_h(t) dt = 1$;
2. $\int t k_h(t) dt = 0$;
3. $\int t^2 k_h(t) dt \neq 0$.

Essas propriedades são necessárias para obtermos uma aproximação de um valor ótimo para h (ver Silverman (1986)) para dados não censurados.

Wand e Jones (1995) alertam ainda que uma boa escolha de h , entretanto, é fundamental para a bondade do núcleo estimador. Um valor de h muito grande acarreta uma suavização excessiva e h muito pequeno faz com que \tilde{F}_k apresente grande variabilidade, no sentido de que não haverá muita suavidade.

Outra complicação que podemos notar é que, para variáveis aleatórias positivas, existe a possibilidade de atribuir, por esse estimador, probabilidade positiva para pontos menores que zero. Esse tipo de problema é comumente chamado de “*problema de fronteira*”. Com isso, ocorre que $\tilde{F}_K(\infty; h) - \tilde{F}_K(0; h) < 1$. Para contornar esse problema, podemos modificar nosso estimador tomando (ver Silverman (1986))

$$\tilde{F}_K(x; h) = \sum_{i=1}^r [K_h(x - X_i) - K_h(-x - X_i)] [\hat{F}(X_i) - \hat{F}(X_{i-1})],$$

em que K_h é simétrico em torno de zero.

Como todo estimador, $\tilde{F}_K(x; h)$ está sujeito a variabilidade aleatória. Mais do que isso, $\tilde{F}_K(x; h)$ possui mais um fator que pode afetar sua performance: a janela h . Há, portanto, alguns critérios para avaliar o desempenho de $\tilde{F}_K(x; h)$. Para um ponto fixado $x \in \mathbb{R}$, verificamos a qualidade de $\tilde{F}_K(x; h)$ pelo Erro Quadrático Médio (EQM):

$$EQM[\tilde{F}_K(x; h)] = E \left\{ \tilde{F}_K(x; h) - F(x) \right\}^2.$$

Para uma avaliação global do desempenho de $\tilde{F}_K(x; h)$, define-se o Erro Quadrático Médio Integrado (EQMI) por

$$EQMI[\tilde{F}_K(\cdot; h)] = \int E \left\{ \tilde{F}_K(x; h) - F(x) \right\}^2 dx.$$

Conforme já dissemos, o desempenho de $\tilde{F}_K(x; h)$ está intimamente ligado ao valor h . A escolha não subjetiva, automática, de um valor ótimo para h , segundo algum critério, é tema de vasta literatura. A maioria das propostas surge da otimização de alguma medida de desempenho de $\tilde{F}_K(x; h)$, como o *EQMI*. Quando temos dados não censurados, por exemplo, um estimador não viesado para $EQMI[\tilde{f}_k(x; h)]$ (ver

Wand e Jones (1995)) é:

$$VCMQ(h) = \int \tilde{f}_k(x; h)^2 dx - \frac{2}{n} \sum_{i=1}^n \tilde{f}_{-i}(X_i; h),$$

onde \tilde{f}_{-i} é \tilde{f}_k calculada sem a observação X_i . A idéia do método de Validação Cruzada de Mínimos Quadrados (para dados não censurados), por exemplo, é tomar o valor ótimo de h como sendo aquele que minimiza $VCMQ$.

Existem ainda métodos que propõem a escolha de uma janela $h(x)$ para cada $x \in \mathbb{R}$. Há também técnicas em que é escolhido um valor h_i para cada X_i . Esses métodos são bem estudados quando se tem uma amostra de dados não censurados.

Por outro lado, para dados com censura intervalar, não há um estimador fechado para o EQMI de $\tilde{F}_K(\cdot; h)$. Por isso, usa-se o método de bootstrap para estimar $EQMI[\tilde{F}_K(\cdot; h)]$ e, assim, estimamos a janela ótima h minimizando essa estimativa bootstrap do $EQMI$. Além desse método, há uma técnica apresentada por Pan (2000b) que faz uma adaptação do método de validação cruzada de verossimilhança de Silverman (1986). Exibiremos agora esses dois métodos de obter h automaticamente, começando pelo método de bootstrap. Para facilitarmos nossa linguagem, denotaremos o valor ótimo de h que buscamos por h_{opt} .

5.1.1 Bootstrap

Como já expusemos, queremos estimar um valor ótimo, segundo algum critério, para h . O critério mais comum para essa otimalidade é, na maioria das formulações, a minimização do $EQMI[\tilde{F}_K(\cdot; h)]$ ou de uma estimativa sua. Contudo, não há uma expressão analítica ou estimador fechado para essa quantidade. Por isso, vamos fazer essa estimação usando o bootstrap para obter o valor de h_{opt} .

Para tanto, usaremos como referência básica o trabalho de Hall (1990), que apresenta várias aplicações da técnica de bootstrap para estimar parâmetros de suavidade em alguns problemas não paramétricos. Hall (1990) faz seus estudos para dados não censurados, mas suas idéias são facilmente estendidas para o estimador com o qual estamos trabalhando para dados com censura intervalar. Faraway e Jhun (1990) também fazem contribuições para esse problema com dados não censurados, mas se

concentram apenas nos núcleo estimadores.

Conforme já colocamos, O EQMI de $\tilde{F}_K(\cdot; h)$ é

$$EQMI[\tilde{F}_K(\cdot; h)] = \int E \left\{ \tilde{F}_K(x; h) - F(x) \right\}^2 dx.$$

Uma estimativa bootstrap é dada por

$$BEQMI(h, h_0) = \widehat{EQMI}[\tilde{F}_K(\cdot; h)] = \frac{1}{B} \sum_{j=1}^B \int \left\{ \tilde{F}_K^{(j)}(x; h) - \tilde{F}_K(x; h_0) \right\}^2 dx,$$

onde:

- h_0 é um valor inicial para a janela h ;
- $\tilde{F}_K(x; h_0)$ é o núcleo estimador de F usual;
- $\tilde{F}_K^{(j)}(x; h)$, $j = 1, 2, \dots, B$, é o núcleo estimador de F calculado a partir da j -ésima reamostra obtida com reposição, de tamanho n_1 , dos dados originais $\{(L_i, U_i, \delta_i, \gamma_i), i = 1, 2, \dots, n\}$.

A reamostragem acima é feita B vezes, tomando-se $n_1 \leq n$. Tomamos h_{opt} como o valor de h que minimiza $BEQMI$. Para diminuir o efeito da escolha de h_0 , o procedimento pode ser repetido iterativamente com o valor ótimo de h anterior.

O outro método que estudamos de escolha de h_{opt} é o de Validação Cruzada de Verossimilhança, que segue na próxima subseção.

5.1.2 Validação Cruzada de Verossimilhança

A maioria dos métodos de escolha de h se baseia na otimização de um critério. Boa parte desses métodos considera como critério a minimização do EQM ou EQMI. Contudo, podemos utilizar como critério a própria verossimilhança. A ideia central do método de Validação Cruzada de Verossimilhança é obter h_{opt} como o valor de h que maximiza a verossimilhança (1.3), tendo como argumento $\tilde{F}_K(x; h)$.

Esse critério é bem conveniente na suavização de F em censura intervalar, dado que nesta situação a verossimilhança para dados coletados depende diretamente da função de distribuição F . Some-se a essa conveniência o fato de não termos estimadores ou

aproximações assintóticas para o EQM ou EQMI. Por isso, esse foi o método que usamos para a escolha de h_{opt} .

Conforme (1.3), a verossimilhança para os dados com censura intervalar univariada é

$$L(F) = \prod_{i=1}^n [F(L_i)]^{\delta_i} [F(U_i) - F(L_i)]^{\gamma_i} [1 - F(U_i)]^{1-\delta_i-\gamma_i}, \quad (1.3)$$

que depende de F e pode ser avaliada em $\tilde{F}_K(\cdot; h)$.

Pan (2000b) propõe uma modificação no método usual de Validação Cruzada de Verossimilhança de Silverman (1986). Primeiramente, dividimos aleatoriamente nosso conjunto de dados \mathcal{D} em V conjuntos de tamanhos iguais $\mathcal{D}^{(v)}$, $j = 1, 2, \dots, V$. Suponha que $\tilde{F}_K^{(-v)}(\mathcal{D}^{(v)}; h)$ é o núcleo estimador de F obtido com o conjunto de dados $\mathcal{D} \setminus \mathcal{D}^{(v)}$ e avaliado nos pontos pertencentes ao conjunto $\mathcal{D}^{(v)}$. Obtemos h_{opt} como

$$h_{opt} = \arg \max_h \sum_{v=1}^V \log L \left(\tilde{F}_K^{(-v)}(\mathcal{D}^{(v)}; h) \right).$$

Em nossas simulações, usamos $V = 10$. Essa metodologia nos parece simples e elimina a escolha de muitos parâmetros, como no método de bootstrap. Além disso, não é necessário estimar nenhuma quantidade ou fazer reamostragens, que muitas vezes podem ter um peso computacional grande em estudos de simulação como o nosso.

Já exibimos o núcleo estimador para o caso univariado. Passaremos agora para o caso bivariado.

5.2 Caso Bivariado

Vimos no capítulo 3 que os incrementos do estimador não paramétrico de máxima verossimilhança da função de distribuição bivariada H ocorrem em retângulos ξ_j , que chamamos de *interseções máximas*. Assim sendo, naquela oportunidade escrevemos $\xi_j = [a_j, b_j] \times [c_j, d_j]$ e associamos a cada ξ_j o valor do incremento p_j de \hat{H} , que é estimado por máxima verossimilhança.

Essa construção acima é importante para que definamos o núcleo estimador de H . É necessário que já tenhamos a estimativa não paramétrica de máxima verossi-

lança \hat{H} , cujo suporte se dá nas interseções máximas $\xi_1, \xi_2, \dots, \xi_J$, com incrementos são $\hat{p}_1, \hat{p}_2, \dots, \hat{p}_J$, respectivamente. O núcleo estimador de H avaliado em $(x, y) \in \mathbb{R}^2$ é semelhante ao caso univariado e é definido analogamente àquele do caso de censura à direita bivariada (ver Wells e Yeo (1996)) como

$$\begin{aligned}\tilde{H}_K(x, y; R) &= \int \int_{\mathbb{R}^2} K_R(x - t, y - u) d\hat{H}(t, u) \\ &= \sum_{j=1}^J K_R(x - b_j, y - d_j) \hat{p}_j,\end{aligned}$$

onde K_R é uma função de distribuição bivariada elíptica que depende da matriz simétrica definida positiva 2x2 R , chamada de matriz janela.

A escolha de R é um problema ainda mais complexo que no caso univariado. Aqui, temos primeiramente que escolher alguma parametrização de R para, depois, procurarmos valores “ótimos”, em algum sentido, para suas entradas. É comum tomar R diagonal para simplificar o problema, como fazem Bae et al. (2005) e Gürler (2000). Wells e Yeo (1996), por exemplo, definem o núcleo estimador bivariado com R diagonal e com entradas iguais.

Com dados não censurados, extensões do caso univariado foram feitas para a escolha de R . Duong e Hazelton (2005), por exemplo, apresentam um método de Validação Cruzada para a escolha de R .

Ainda não há contribuições para a escolha de R no caso de censura intervalar bivariada. De qualquer forma, o método de Validação Cruzada de Verossimilhança apresentado na subseção 5.1.2 deste capítulo pode ser facilmente estendido para o caso bivariado, uma vez que nesse método só precisamos avaliar a verossimilhança com a estimativa $\tilde{H}_K(\cdot; R)$ para proceder à escolha de R .

Com esse método, procedemos da seguinte maneira: dividimos aleatoriamente nosso conjunto de dados \mathcal{D} em V conjuntos de tamanhos iguais $\mathcal{D}^{(v)}$, $j = 1, 2, \dots, V$, e obtemos $\tilde{H}_K^{(-v)}(\mathcal{D}^{(v)}; R)$, que é o núcleo estimador de H obtido com o conjunto de dados $\mathcal{D} \setminus \mathcal{D}^{(v)}$ e avaliado nos pontos pertencentes ao conjunto $\mathcal{D}^{(v)}$. Assim, o valor ótimo de R , R_{opt} , é dado por

$$R_{opt} = \arg \max_R \sum_{v=1}^V \log L \left(\tilde{H}_K^{(-v)}(\mathcal{D}^{(v)}; R) \right), \quad (5.1)$$

sendo que a função L na expressão (5.1) acima é a função de verossimilhança de dados com censura intervalar para o caso bivariado, descrita em (1.5).

Em nossas simulações, tomamos R diagonal e com entradas h_1 e h_2 . Além disso, consideramos $V = 10$, como no caso univariado.

Capítulo 6

Simulações e aplicação

6.1 Simulações

Fizemos a exposição dos métodos que usaremos em nossos estudos de simulação e, agora, podemos expor os resultados encontrados em nossas simulações. Anotamos na tabela 6.1 os parâmetros utilizados.

Tabela 6.1: Parâmetros utilizados nas simulações ($i = 1, 2$)

Caso	T_i	L_i	$U_i - L_i$	Cópula
1	Weibull(4, 0,2)	Weibull(4, 1/3,8)	Weibull(4, 1/2,5)	Clayton(4)
2	Weibull(4, 0,2)	Weibull(4, 1/3,8)	Weibull(4, 1/2,5)	Frank(7)
3	Weibull(4, 0,2)	Weibull(4, 1/3,8)	Weibull(4, 1/2,5)	Frank(-7)

Em cada caso de simulação, foram adotados três tamanhos de amostra diferentes $n = 50, 100, 200$ e foram feitas 10.000 repetições em cada uma dessas configurações. Por isso, tínhamos 10.000 amostras de tamanho $n = 50$ para o caso 1, 10.000 amostras de tamanho $n = 50$ para o caso 2 e assim por diante, ocorrendo o mesmo com $n = 100$ e $n = 200$.

Embora tenhamos tomado o mesmo modelo de cópula para os casos 2 e 3, o sinal do parâmetro escolhido em cada caso indica o sentido da dependência entre T_1 e T_2 , usando τ de Kendall como medida de dependência. No caso 2, temos $\tau = 0,562256$, enquanto $\tau = -0,562256$ no caso 3. No caso 1, temos $\tau = 2/3$.

Usamos o núcleo gaussiano, univariado e bivariado, sempre que núcleo estimadores foram utilizados. Conforme já expomos, a escolha da janela foi feita segundo o método

de validação cruzada de verossimilhança proposto por Pan (2000b). Todos os cálculos foram feitos no software livre R, sendo que algumas funções foram implementadas na linguagem C e invocadas dentro do R por meio da função “.C()”.

Daqui para frente, para evitar confusão e simplificar a redação, chamaremos o estimador que utiliza o modelo de cópulas com marginais suavizadas de “Cópula”, enquanto denominamos o outro método de “ENPMV”.

Para avaliar os dois estimadores, calculamos a estimativa nos pontos do produto cartesiano entre os decís de T_1 e T_2 . Como tomamos T_1 e T_2 com a mesma distribuição, os decís de T_1 e T_2 são iguais. Feitos esses cálculos, estimamos o vício, variância e EQM em cada um desses pontos, usando as repetições de cada caso.

Antes de partirmos para as estimativas de vício, variância e EQM, mostraremos alguns resultados referentes às estimativas do parâmetro da cópula, α , e às janelas do núcleo estimador, para os casos univariados e bivariados.

É fundamental dizermos que nem sempre foi possível estimar o parâmetro α da cópula, visto que nem sempre a função de verossimilhança (4.1), que é avaliada na cópula, tinha um ponto de máximo. Por isso, a tabela 6.2 lista o número de vezes em que isso ocorreu. Nesses casos, para estimar o vício, variância e EQM do estimador que utiliza o modelo de cópulas, utilizamos as amostras em que foi possível estimar α .

Tabela 6.2: Número de casos em que $\hat{\alpha}$ não foi obtido

Caso	n	Casos	%
1	50	1005	10,05
1	100	127	1,27
1	200	4	0,04
2	50	405	4,05
2	100	25	0,25
2	200	1	0,01
3	50	31	0,31
3	100	2	0,02

É nítido que o número de casos em que $\hat{\alpha}$ não existe decresce à medida que aumenta o tamanho de amostra. Além disso, podemos ver que a cópula Clayton foi a que apresentou mais vezes esse tipo de problema. Nas figuras de 6.1 a 6.3, que seguem mais abaixo, temos uma estimativa não paramétrica, dada pelo núcleo estimador, da

densidade de $\hat{\alpha}$ para cada caso e tamanho de amostra, da qual retiramos os *outliers* para que o gráfico fosse melhor visualizado. Pode-se ver que a distribuição de $\hat{\alpha}$ vai se tornando mais simétrica e concentrada em torno do verdadeiro valor de α à medida que o tamanho da amostra aumenta. Essa é uma característica dos estimadores de máxima verossimilhança.

Expomos também os gráficos que ilustram a distribuição das janelas (parâmetros de suavização) estimadas (\hat{h}_1, \hat{h}_2) do núcleo estimador para ambos os estimadores. Analisando-os, verifica-se uma propriedade desejável: os valores da janela estimada tendem a diminuir à medida que o tamanho de amostra aumenta. Além disso, a distribuição conjunta de (\hat{h}_1, \hat{h}_2) parece ser elíptica e simétrica.

Finalmente, apresentamos nas tabelas 6.3 a 6.56 abaixo as estimativas de vício, variância e EQM para os dois métodos que estudamos, calculadas nos decis de T_1 e T_2 (que tomamos com mesma distribuição). Seguem também os gráficos das curvas de nível (figuras 6.19 a 6.30) dos valores esperados para ambos os estimadores, em cada caso e tamanho de amostra. Consideramos os níveis iguais a 0,025, 0,050, 0,075, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7 e 0,8.

A análise dessas tabelas nos mostra o seguinte: Nos casos 1 e 2, para todos os tamanhos de amostra, o vício (em valor absoluto) do estimador ENPMV só não é maior que o vício do estimador Cópula quando os valores de T_1 ou T_2 são pequenos. Os valores do vício decrescem, principalmente ao longo da bissetriz $x = y$, e ficam cada vez mais negativos à medida que T_1 ou T_2 crescem. No caso 3, para todos os tamanhos de amostra, a diferença do vício dos dois estimadores é menor que nos dois casos anteriores. Além disso, no caso 2, o vício do estimador ENPMV é maior nos pares dos decis cujas ordens estão no conjunto $\{(k, i) : 1 \leq k < i \leq 9\}$.

Quanto à variância, temos que nos casos 1 e 2, o estimador Cópula tem variância menor em quase todos os pontos, exceto na região dos maiores decis de T_1 e T_2 . Isso ocorre porque a variância do estimador ENPMV sofre um abrupto decréscimo na região dos maiores decis de T_1 e T_2 . O estimador Cópula, ao contrário, tem um crescimento mais “uniforme” da variância à medida que T_1 e T_2 crescem. Já no caso 3, o comportamento da variância é similar ao que ocorre com o vício.

O comportamento do EQM segue basicamente o que ocorre com o vício, o que nos leva a concluir que o estimador Cópula é, segundo o EQM, um estimador melhor que

o estimador ENPMV. É importante enfatizarmos que o vício, variância e EQM, em ambos os estimadores, diminuem quando aumentamos o tamanho da amostra, como esperado.

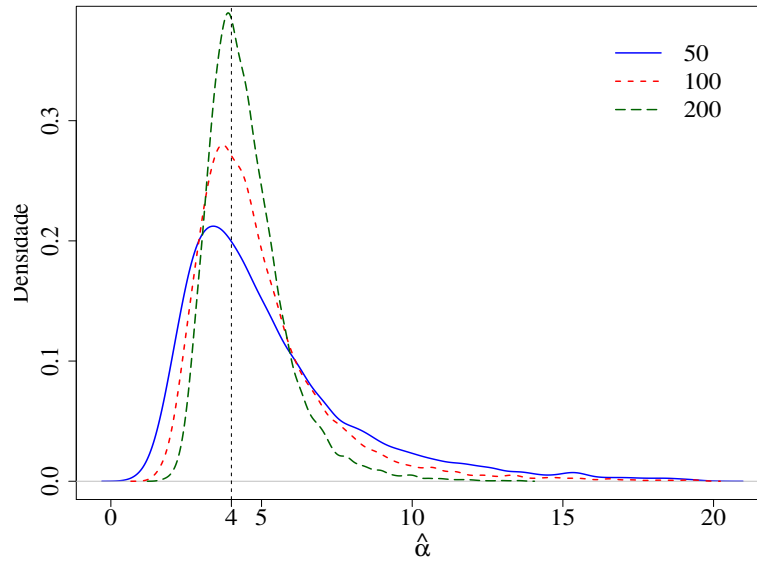


Figura 6.1: Densidade estimada para $\hat{\alpha}$ - Caso 1

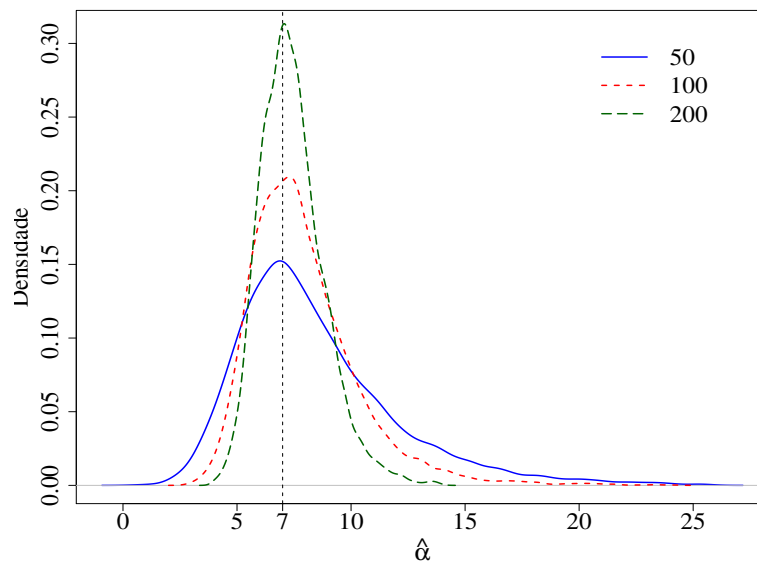


Figura 6.2: Densidade estimada para $\hat{\alpha}$ - Caso 2

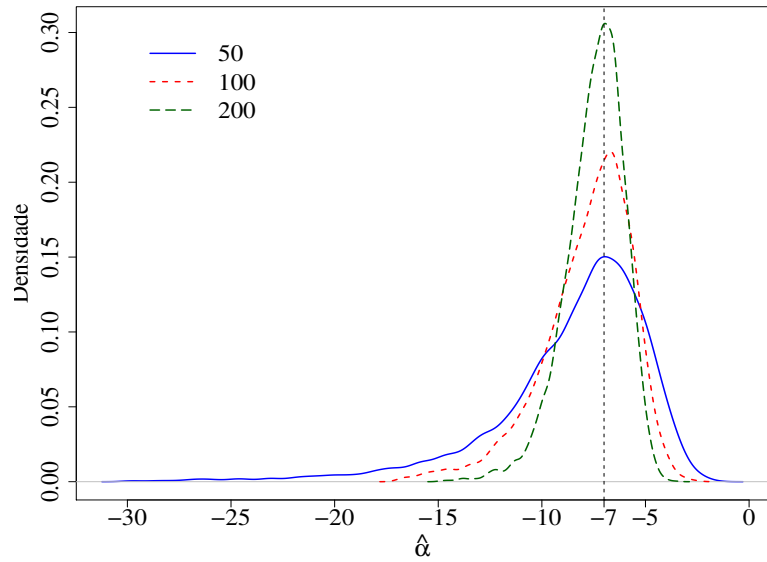


Figura 6.3: Densidade estimada para $\hat{\alpha}$ - Caso 3

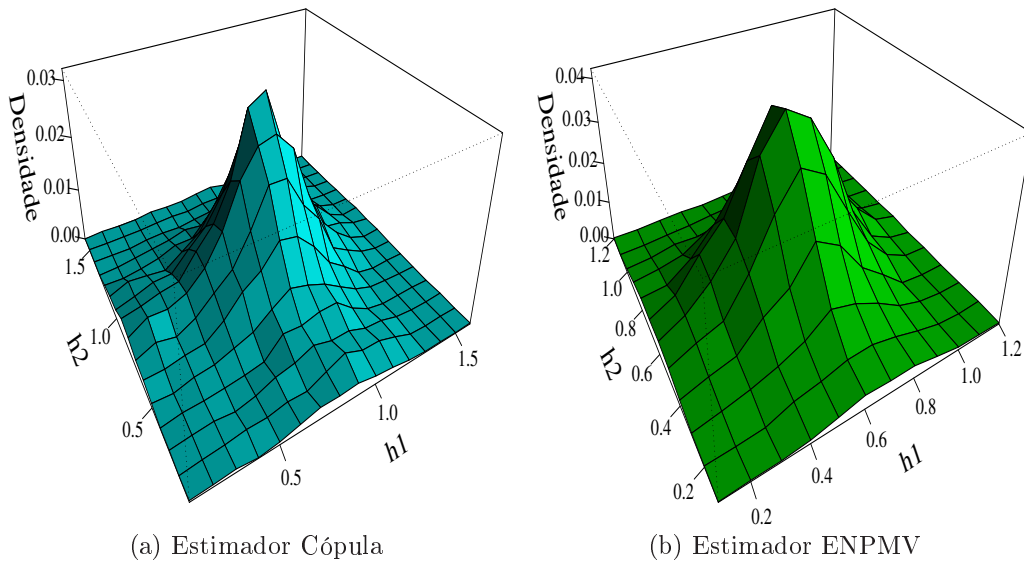


Figura 6.4: Densidade conjunta estimada para janelas Caso 1, $n = 50$

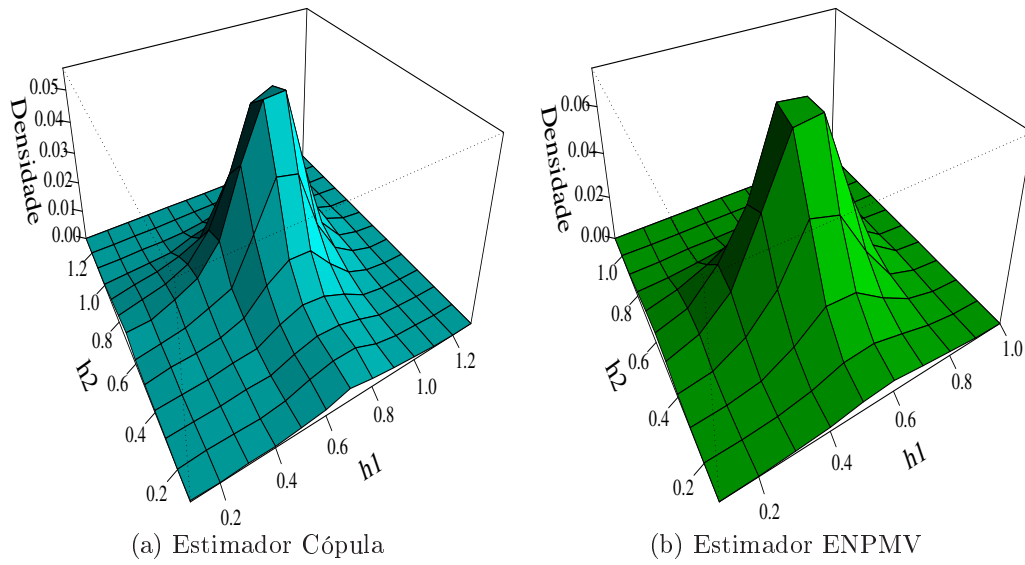


Figura 6.5: Densidade conjunta estimada para janelas Caso 1, $n = 100$

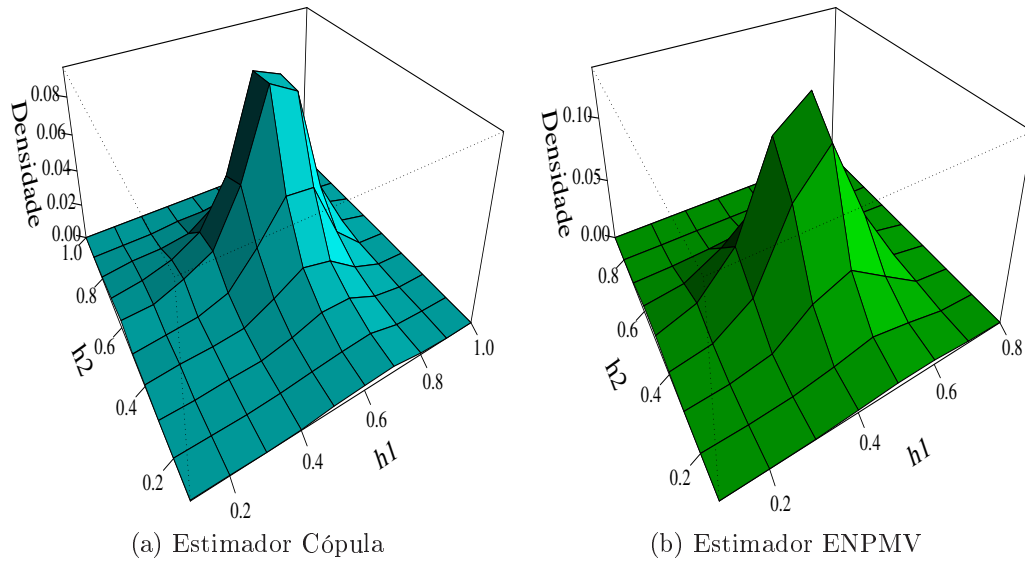


Figura 6.6: Densidade conjunta estimada para janelas Caso 1, $n = 200$

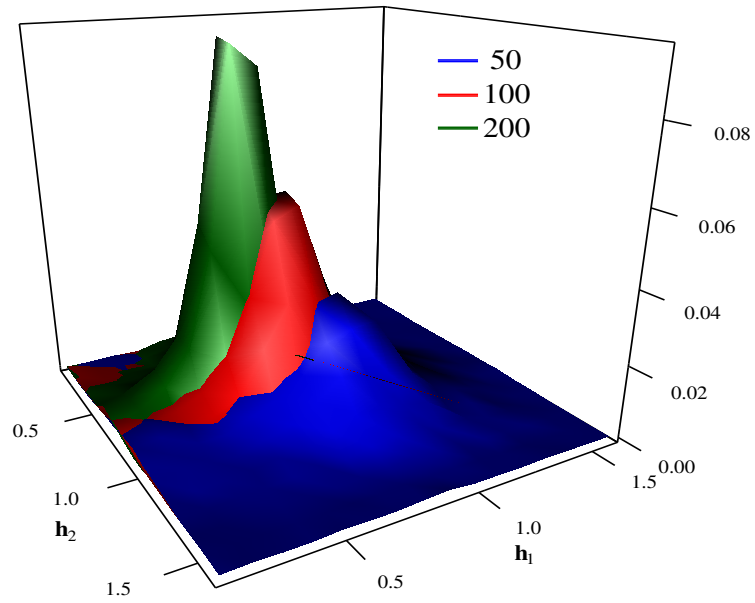


Figura 6.7: Comparativo - Densidade das janelas - Estimador Cópula - Caso 1

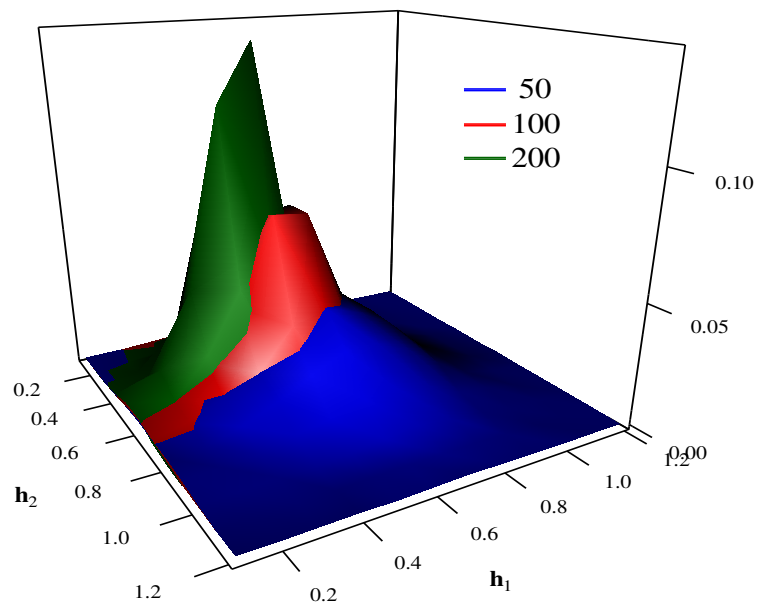
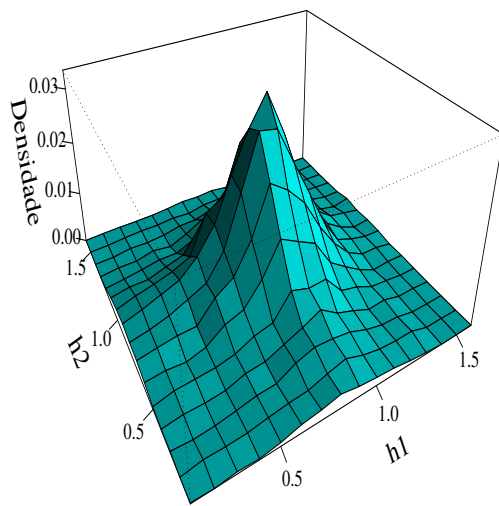
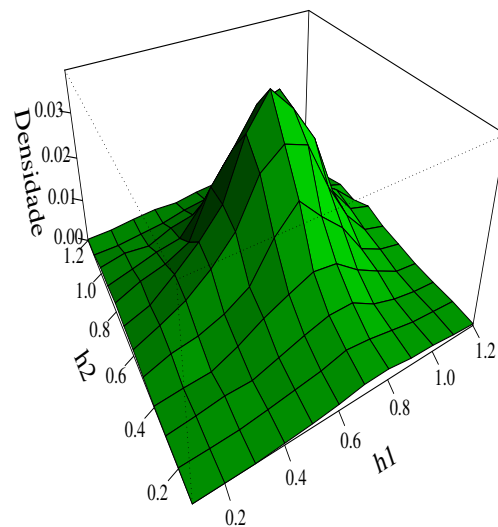


Figura 6.8: Comparativo - Densidade das janelas - Estimador ENPMV - Caso 1

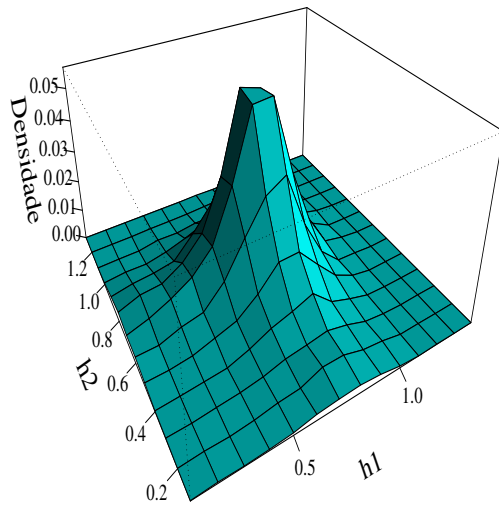


(a) Estimador Cópula

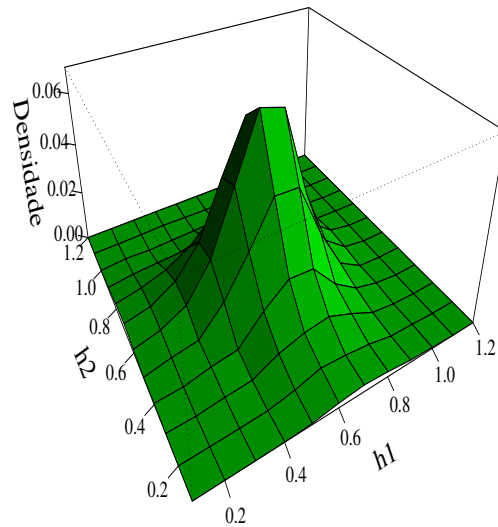


(b) Estimador ENPMV

Figura 6.9: Densidade conjunta estimada para janelas Caso 2, $n = 50$



(a) Estimador Cópula



(b) Estimador ENPMV

Figura 6.10: Densidade conjunta estimada para janelas Caso 2, $n = 100$

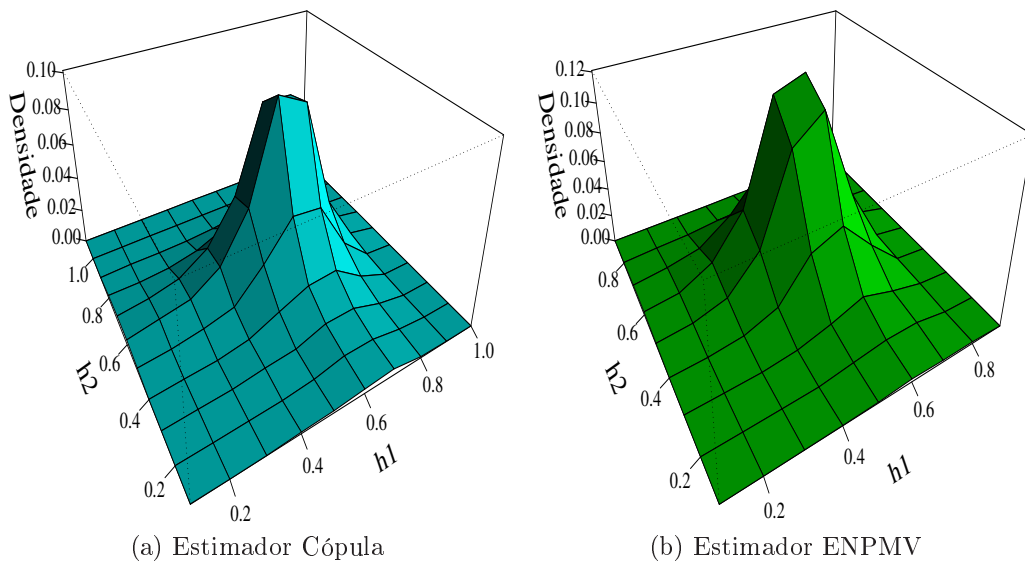


Figura 6.11: Densidade conjunta estimada para janelas Caso 2, $n = 200$

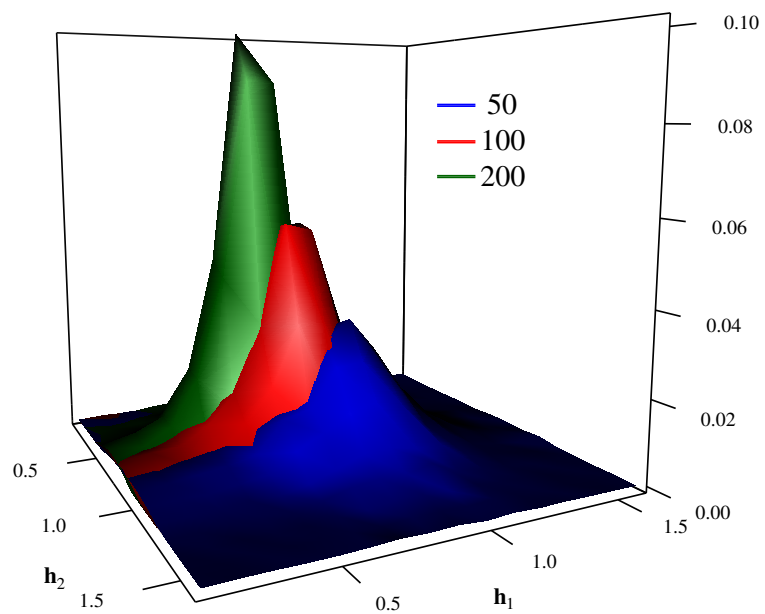


Figura 6.12: Comparativo - Densidade das janelas - Estimador Cópula - Caso 2

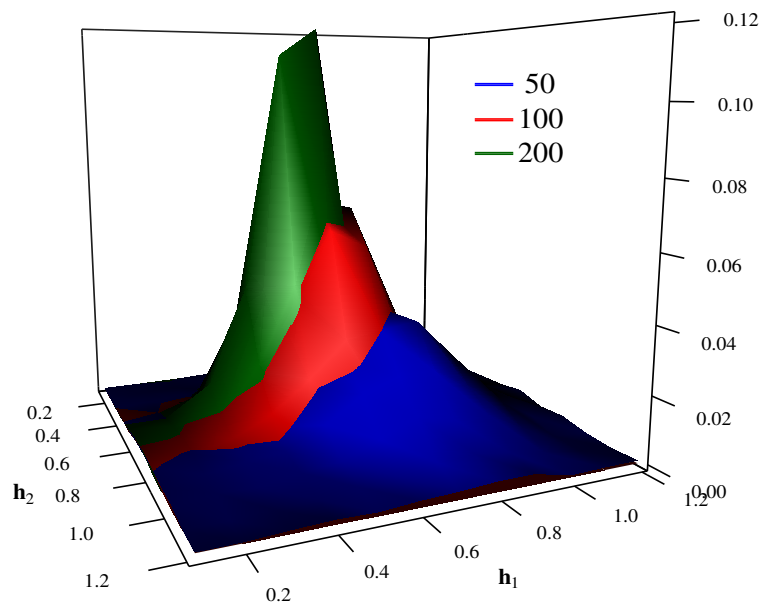


Figura 6.13: Comparativo - Densidade das janelas - Estimador ENPMV - Caso 2

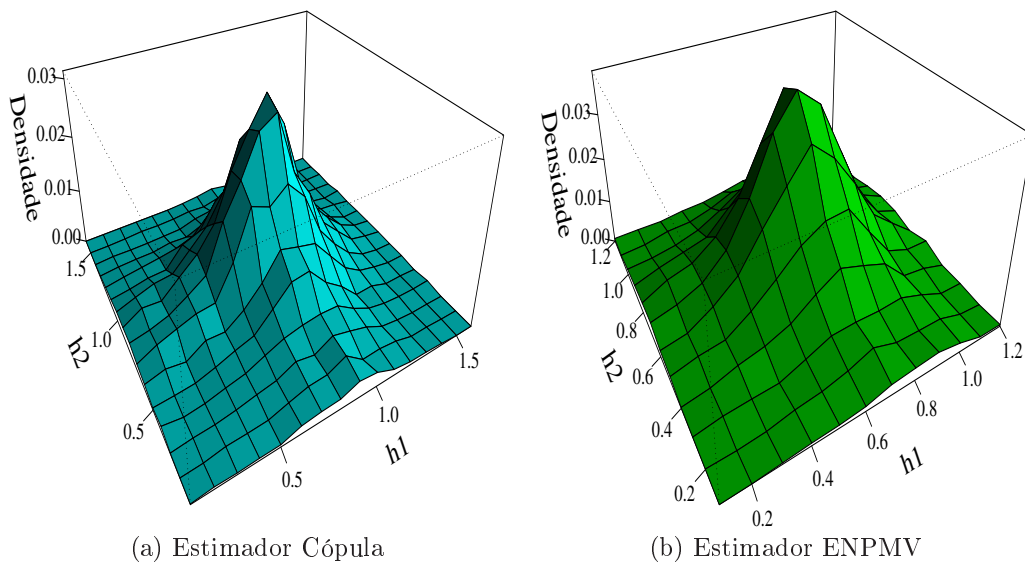


Figura 6.14: Densidade conjunta estimada para janelas Caso 3, $n = 50$

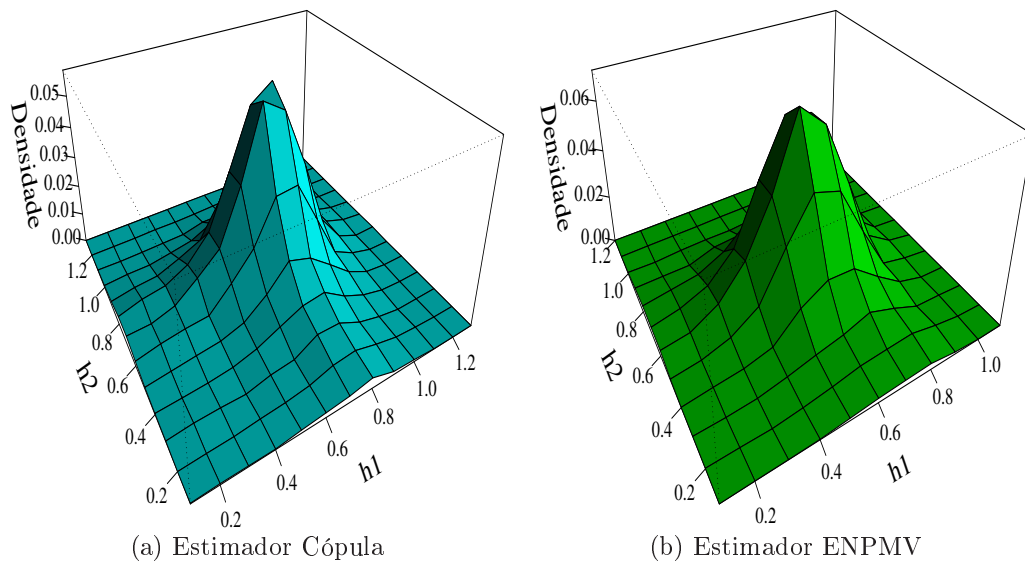


Figura 6.15: Densidade conjunta estimada para janelas Caso 3, $n = 100$

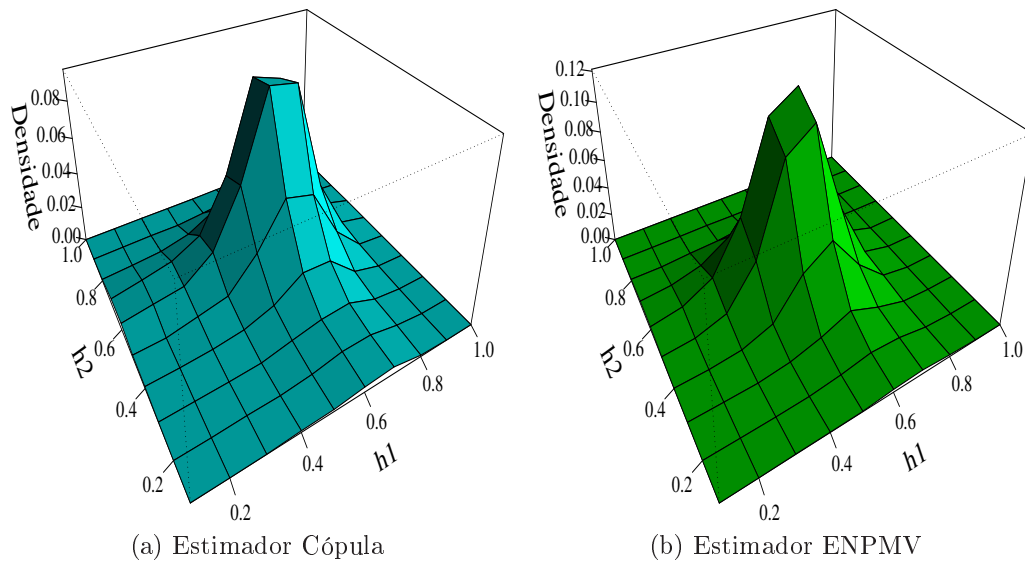


Figura 6.16: Densidade conjunta estimada para janelas Caso 3, $n = 200$

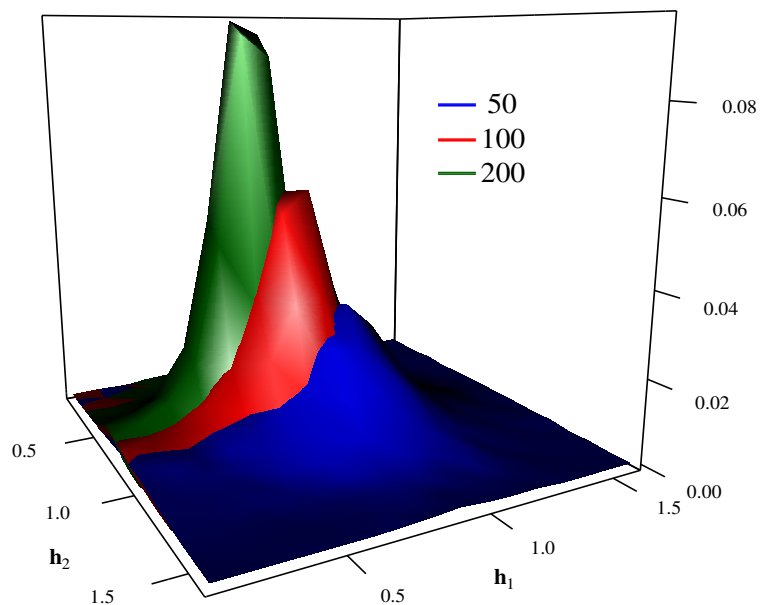


Figura 6.17: Comparativo - Densidade das janelas - Estimador Cópula - Caso 3

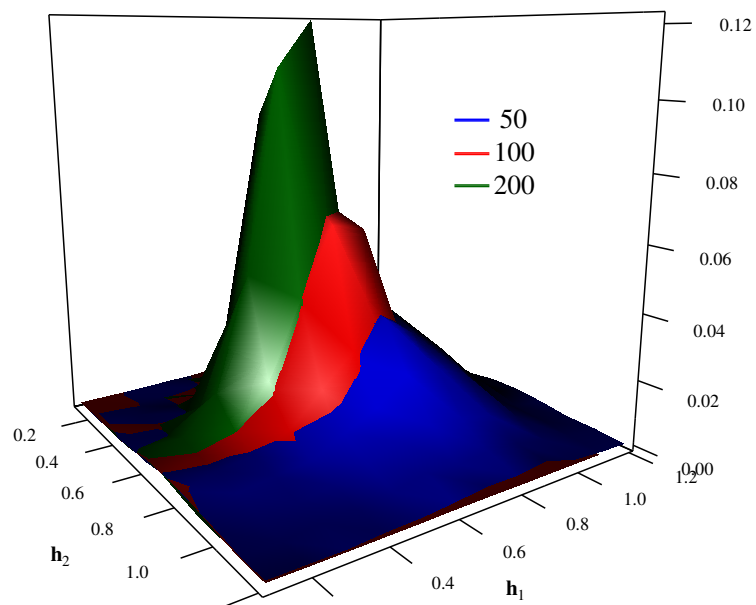


Figura 6.18: Comparativo - Densidade das janelas - Estimador ENPMV - Caso 3

Tabela 6.3: Vício - estimador Cópula - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	-0.002	0.008	0.011	0.012	0.012	0.013	0.013	0.013	0.013
2	0.008	0.008	0.013	0.015	0.017	0.018	0.018	0.019	0.019
3	0.012	0.012	0.011	0.010	0.012	0.013	0.014	0.016	0.016
4	0.012	0.015	0.010	0.005	0.002	0.002	0.004	0.005	0.006
5	0.013	0.016	0.011	0.002	-0.005	-0.008	-0.009	-0.008	-0.007
6	0.013	0.017	0.013	0.002	-0.008	-0.015	-0.020	-0.021	-0.021
7	0.013	0.018	0.014	0.003	-0.009	-0.020	-0.027	-0.031	-0.032
8	0.013	0.018	0.015	0.004	-0.009	-0.021	-0.031	-0.037	-0.040
9	0.013	0.018	0.016	0.006	-0.007	-0.021	-0.032	-0.040	-0.043

Tabela 6.4: Variância $\times 10^{-3}$ - estimador Cópula - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	2.11	2.66	2.88	2.95	2.99	3.00	3.01	3.01	3.02
2	2.70	3.12	3.45	3.73	3.88	3.96	4.00	4.03	4.05
3	2.91	3.42	3.40	3.67	3.87	3.98	4.07	4.13	4.17
4	2.98	3.67	3.64	3.84	4.02	4.13	4.23	4.33	4.42
5	3.01	3.80	3.81	3.99	4.14	4.19	4.27	4.39	4.54
6	3.03	3.87	3.93	4.12	4.24	4.22	4.22	4.31	4.49
7	3.04	3.91	4.01	4.23	4.35	4.24	4.15	4.13	4.24
8	3.04	3.94	4.08	4.34	4.49	4.34	4.15	4.00	3.96
9	3.05	3.96	4.12	4.43	4.66	4.54	4.34	4.04	3.63

Tabela 6.5: EQM $\times 10^{-3}$ - estimador Cópula - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	2.11	2.73	3.00	3.10	3.14	3.16	3.17	3.17	3.18
2	2.78	3.19	3.61	3.96	4.16	4.27	4.33	4.38	4.41
3	3.04	3.57	3.51	3.78	4.00	4.16	4.28	4.37	4.44
4	3.14	3.89	3.74	3.86	4.03	4.13	4.24	4.35	4.46
5	3.18	4.06	3.94	4.00	4.16	4.26	4.35	4.46	4.58
6	3.20	4.17	4.09	4.12	4.31	4.45	4.61	4.76	4.91
7	3.21	4.23	4.21	4.24	4.44	4.63	4.87	5.09	5.28
8	3.22	4.27	4.30	4.36	4.57	4.79	5.11	5.36	5.54
9	3.22	4.30	4.37	4.46	4.71	4.97	5.39	5.64	5.45

Tabela 6.6: Vício - estimador ENPMV - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	-0.047	-0.037	-0.022	-0.012	-0.004	0.001	0.004	0.007	0.009
2	-0.037	-0.062	-0.053	-0.036	-0.022	-0.011	-0.002	0.004	0.009
3	-0.022	-0.053	-0.068	-0.062	-0.048	-0.032	-0.019	-0.008	0.001
4	-0.011	-0.037	-0.063	-0.075	-0.070	-0.057	-0.042	-0.027	-0.014
5	-0.004	-0.022	-0.048	-0.070	-0.080	-0.076	-0.064	-0.048	-0.031
6	0.001	-0.011	-0.033	-0.058	-0.077	-0.084	-0.080	-0.067	-0.049
7	0.005	-0.003	-0.019	-0.042	-0.065	-0.080	-0.085	-0.079	-0.064
8	0.007	0.004	-0.008	-0.027	-0.049	-0.068	-0.079	-0.081	-0.071
9	0.009	0.009	0.001	-0.014	-0.032	-0.050	-0.064	-0.071	-0.068

Tabela 6.7: Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.74	1.28	1.69	2.01	2.28	2.48	2.64	2.78	2.88
2	1.32	1.96	2.36	2.71	3.00	3.25	3.46	3.64	3.81
3	1.74	2.35	2.66	2.97	3.25	3.46	3.63	3.79	3.94
4	2.07	2.68	2.94	3.28	3.61	3.84	3.99	4.10	4.21
5	2.32	2.96	3.18	3.55	3.93	4.20	4.33	4.41	4.45
6	2.53	3.20	3.38	3.75	4.17	4.48	4.65	4.72	4.70
7	2.69	3.41	3.56	3.91	4.33	4.66	4.86	4.94	4.86
8	2.82	3.59	3.72	4.03	4.43	4.75	4.98	5.12	5.01
9	2.92	3.75	3.87	4.14	4.47	4.73	4.95	5.09	4.92

Tabela 6.8: EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	2.97	2.67	2.18	2.15	2.29	2.48	2.66	2.82	2.96
2	2.68	5.82	5.18	4.03	3.49	3.37	3.46	3.66	3.89
3	2.20	5.17	7.35	6.87	5.52	4.51	3.99	3.85	3.95
4	2.19	4.01	6.85	8.86	8.56	7.14	5.76	4.84	4.39
5	2.34	3.45	5.46	8.51	10.36	10.04	8.47	6.75	5.43
6	2.53	3.32	4.44	7.07	10.05	11.56	11.05	9.25	7.13
7	2.71	3.41	3.93	5.70	8.51	11.09	12.14	11.25	8.94
8	2.87	3.61	3.79	4.78	6.81	9.31	11.29	11.71	10.08
9	3.00	3.83	3.87	4.34	5.50	7.20	9.05	10.17	9.50

Tabela 6.9: Vício - estimador Cópula - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.002	0.009	0.011	0.011	0.011	0.011	0.011	0.011	0.011
2	0.009	0.013	0.015	0.016	0.017	0.017	0.018	0.018	0.018
3	0.011	0.014	0.013	0.012	0.013	0.013	0.014	0.015	0.015
4	0.011	0.016	0.012	0.008	0.006	0.005	0.006	0.007	0.007
5	0.011	0.017	0.012	0.006	0.001	-0.002	-0.003	-0.003	-0.002
6	0.011	0.017	0.013	0.006	-0.002	-0.007	-0.011	-0.013	-0.013
7	0.011	0.018	0.014	0.006	-0.003	-0.011	-0.017	-0.021	-0.022
8	0.011	0.018	0.014	0.007	-0.003	-0.013	-0.021	-0.026	-0.029
9	0.011	0.018	0.015	0.007	-0.002	-0.013	-0.022	-0.029	-0.030

Tabela 6.10: Variância $\times 10^{-3}$ - estimador Cópula - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	1.17	1.50	1.62	1.65	1.66	1.67	1.67	1.67	1.67
2	1.50	1.51	1.72	1.87	1.94	1.97	1.99	2.00	2.01
3	1.62	1.77	1.84	1.98	2.10	2.17	2.22	2.25	2.27
4	1.66	1.93	2.01	2.12	2.20	2.28	2.35	2.40	2.45
5	1.67	2.01	2.12	2.23	2.29	2.32	2.37	2.44	2.53
6	1.68	2.04	2.20	2.31	2.35	2.33	2.32	2.36	2.47
7	1.68	2.07	2.25	2.40	2.42	2.36	2.28	2.24	2.33
8	1.68	2.08	2.28	2.47	2.52	2.43	2.28	2.15	2.16
9	1.69	2.09	2.31	2.53	2.62	2.56	2.37	2.14	2.01

Tabela 6.11: EQM $\times 10^{-3}$ - estimador Cópula - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	1.18	1.59	1.73	1.77	1.79	1.80	1.80	1.80	1.81
2	1.59	1.67	1.94	2.13	2.22	2.28	2.31	2.33	2.34
3	1.73	1.97	2.00	2.13	2.25	2.35	2.42	2.46	2.50
4	1.78	2.18	2.15	2.18	2.24	2.31	2.38	2.45	2.51
5	1.80	2.28	2.27	2.26	2.29	2.33	2.38	2.45	2.54
6	1.80	2.34	2.37	2.34	2.35	2.38	2.44	2.52	2.64
7	1.81	2.37	2.44	2.43	2.43	2.48	2.56	2.67	2.82
8	1.81	2.39	2.49	2.51	2.52	2.59	2.70	2.82	2.97
9	1.81	2.41	2.53	2.59	2.62	2.72	2.86	2.96	2.94

Tabela 6.12: Vício - estimador ENPMV - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	-0.039	-0.026	-0.012	-0.003	0.002	0.006	0.008	0.009	0.010
2	-0.026	-0.045	-0.036	-0.021	-0.009	-0.000	0.006	0.010	0.013
3	-0.012	-0.036	-0.050	-0.045	-0.031	-0.018	-0.008	0.000	0.006
4	-0.003	-0.021	-0.045	-0.056	-0.052	-0.040	-0.027	-0.015	-0.005
5	0.002	-0.009	-0.032	-0.052	-0.060	-0.056	-0.045	-0.032	-0.019
6	0.005	-0.000	-0.019	-0.040	-0.056	-0.061	-0.057	-0.047	-0.032
7	0.007	0.006	-0.008	-0.027	-0.045	-0.057	-0.061	-0.056	-0.043
8	0.009	0.010	-0.000	-0.015	-0.032	-0.046	-0.056	-0.058	-0.049
9	0.009	0.012	0.006	-0.006	-0.019	-0.032	-0.043	-0.049	-0.046

Tabela 6.13: Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.58	0.90	1.12	1.28	1.39	1.47	1.53	1.57	1.60
2	0.90	1.16	1.31	1.45	1.56	1.66	1.74	1.81	1.87
3	1.12	1.35	1.48	1.63	1.75	1.85	1.93	2.00	2.07
4	1.28	1.50	1.64	1.83	1.97	2.06	2.13	2.18	2.24
5	1.40	1.63	1.76	1.98	2.15	2.26	2.32	2.36	2.38
6	1.48	1.73	1.86	2.09	2.29	2.42	2.48	2.50	2.48
7	1.54	1.82	1.94	2.17	2.37	2.52	2.59	2.61	2.57
8	1.58	1.88	2.02	2.24	2.42	2.57	2.64	2.66	2.60
9	1.61	1.94	2.09	2.30	2.45	2.55	2.58	2.58	2.51

Tabela 6.14: EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	2.10	1.58	1.26	1.29	1.39	1.50	1.59	1.65	1.69
2	1.58	3.22	2.62	1.89	1.64	1.66	1.78	1.91	2.03
3	1.27	2.67	4.03	3.64	2.74	2.18	1.98	2.00	2.11
4	1.29	1.96	3.66	4.94	4.62	3.64	2.84	2.41	2.27
5	1.40	1.71	2.76	4.63	5.70	5.35	4.33	3.36	2.73
6	1.51	1.73	2.20	3.67	5.39	6.19	5.77	4.67	3.53
7	1.59	1.85	2.01	2.89	4.38	5.79	6.33	5.76	4.46
8	1.65	1.98	2.02	2.46	3.42	4.71	5.77	5.97	5.04
9	1.69	2.09	2.12	2.33	2.80	3.58	4.47	5.02	4.63

Tabela 6.15: Vício - estimador Cópula - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.005	0.009	0.010	0.010	0.011	0.011	0.011	0.011	0.011
2	0.010	0.012	0.013	0.014	0.014	0.015	0.015	0.015	0.015
3	0.010	0.013	0.012	0.011	0.011	0.011	0.012	0.012	0.013
4	0.011	0.014	0.011	0.008	0.006	0.006	0.006	0.006	0.007
5	0.011	0.015	0.011	0.006	0.002	-0.000	-0.001	-0.002	-0.001
6	0.011	0.015	0.012	0.006	-0.000	-0.005	-0.008	-0.009	-0.010
7	0.011	0.015	0.012	0.006	-0.001	-0.008	-0.013	-0.016	-0.017
8	0.011	0.015	0.013	0.007	-0.001	-0.009	-0.016	-0.020	-0.023
9	0.011	0.016	0.013	0.007	-0.001	-0.009	-0.017	-0.023	-0.025

Tabela 6.16: Variância $\times 10^{-3}$ - estimador Cópula - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.56	0.78	0.85	0.86	0.87	0.87	0.87	0.87	0.87
2	0.76	0.78	0.94	1.03	1.07	1.09	1.10	1.10	1.11
3	0.83	0.91	1.01	1.10	1.18	1.22	1.25	1.27	1.28
4	0.84	1.00	1.11	1.18	1.25	1.30	1.34	1.38	1.41
5	0.85	1.03	1.18	1.23	1.29	1.33	1.36	1.41	1.47
6	0.85	1.05	1.22	1.27	1.31	1.33	1.34	1.38	1.46
7	0.85	1.06	1.25	1.31	1.32	1.30	1.27	1.25	1.32
8	0.85	1.06	1.26	1.35	1.37	1.33	1.24	1.16	1.15
9	0.85	1.07	1.28	1.38	1.42	1.40	1.30	1.15	1.02

Tabela 6.17: EQM $\times 10^{-3}$ - estimador Cópula - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.59	0.87	0.95	0.97	0.98	0.98	0.99	0.99	0.99
2	0.86	0.93	1.11	1.23	1.28	1.31	1.32	1.33	1.34
3	0.93	1.09	1.14	1.22	1.30	1.35	1.39	1.42	1.44
4	0.95	1.20	1.23	1.24	1.28	1.33	1.37	1.41	1.45
5	0.96	1.25	1.31	1.27	1.29	1.33	1.36	1.41	1.47
6	0.96	1.28	1.36	1.31	1.31	1.35	1.40	1.47	1.55
7	0.97	1.29	1.40	1.35	1.33	1.36	1.43	1.51	1.61
8	0.97	1.30	1.43	1.39	1.37	1.41	1.49	1.58	1.66
9	0.97	1.31	1.45	1.43	1.42	1.48	1.58	1.66	1.63

Tabela 6.18: Vício - estimador ENPMV - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	-0.030	-0.016	-0.004	0.003	0.006	0.008	0.009	0.010	0.010
2	-0.016	-0.034	-0.025	-0.012	-0.002	0.004	0.008	0.011	0.012
3	-0.004	-0.025	-0.038	-0.033	-0.021	-0.010	-0.002	0.003	0.007
4	0.002	-0.012	-0.033	-0.042	-0.038	-0.028	-0.017	-0.008	-0.001
5	0.006	-0.002	-0.021	-0.038	-0.044	-0.041	-0.031	-0.021	-0.011
6	0.008	0.005	-0.010	-0.027	-0.040	-0.045	-0.041	-0.032	-0.022
7	0.009	0.009	-0.002	-0.016	-0.031	-0.041	-0.044	-0.040	-0.030
8	0.009	0.011	0.004	-0.007	-0.020	-0.032	-0.039	-0.041	-0.035
9	0.010	0.012	0.007	-0.001	-0.011	-0.021	-0.030	-0.035	-0.033

Tabela 6.19: Variância $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.39	0.55	0.65	0.71	0.76	0.78	0.80	0.81	0.82
2	0.54	0.64	0.73	0.81	0.86	0.91	0.95	0.98	0.99
3	0.64	0.72	0.83	0.92	0.98	1.03	1.07	1.11	1.13
4	0.70	0.78	0.91	1.02	1.11	1.16	1.20	1.22	1.24
5	0.74	0.84	0.97	1.11	1.22	1.29	1.33	1.34	1.34
6	0.77	0.88	1.02	1.16	1.29	1.39	1.43	1.44	1.42
7	0.78	0.92	1.06	1.18	1.31	1.41	1.47	1.47	1.42
8	0.79	0.95	1.10	1.20	1.30	1.39	1.45	1.45	1.37
9	0.80	0.96	1.13	1.22	1.29	1.35	1.39	1.35	1.24

Tabela 6.20: EQM $\times 10^{-3}$ - estimador ENPMV - caso 1, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	1.29	0.81	0.66	0.72	0.79	0.85	0.88	0.90	0.91
2	0.81	1.76	1.37	0.95	0.87	0.93	1.02	1.09	1.14
3	0.65	1.36	2.29	2.00	1.42	1.13	1.08	1.12	1.18
4	0.70	0.93	1.98	2.79	2.54	1.92	1.48	1.28	1.25
5	0.78	0.84	1.41	2.54	3.19	2.94	2.31	1.77	1.47
6	0.83	0.90	1.12	1.91	2.93	3.42	3.14	2.49	1.89
7	0.86	0.99	1.07	1.46	2.27	3.09	3.40	3.04	2.33
8	0.88	1.07	1.11	1.26	1.71	2.41	3.01	3.10	2.58
9	0.89	1.12	1.18	1.22	1.41	1.81	2.28	2.56	2.30

Tabela 6.21: Vício - estimador Cópula - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.009	0.012	0.012	0.012	0.012	0.012	0.012	0.013	0.013
2	0.012	0.017	0.017	0.016	0.015	0.016	0.016	0.017	0.018
3	0.012	0.017	0.017	0.014	0.012	0.012	0.013	0.014	0.015
4	0.011	0.016	0.014	0.009	0.004	0.002	0.002	0.004	0.006
5	0.011	0.015	0.012	0.004	-0.004	-0.009	-0.011	-0.010	-0.008
6	0.011	0.015	0.012	0.002	-0.009	-0.017	-0.022	-0.023	-0.021
7	0.012	0.016	0.013	0.002	-0.011	-0.022	-0.030	-0.034	-0.033
8	0.012	0.017	0.014	0.004	-0.010	-0.023	-0.034	-0.040	-0.041
9	0.012	0.018	0.015	0.005	-0.008	-0.021	-0.033	-0.041	-0.043

Tabela 6.22: Variância $\times 10^{-3}$ - estimador Cópula - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	1.17	1.89	2.34	2.61	2.78	2.88	2.94	2.99	3.01
2	1.87	2.67	3.12	3.45	3.65	3.79	3.88	3.94	3.99
3	2.31	3.11	3.41	3.66	3.82	3.93	4.02	4.10	4.16
4	2.58	3.44	3.68	3.90	4.03	4.12	4.21	4.31	4.42
5	2.74	3.65	3.85	4.03	4.14	4.19	4.27	4.39	4.54
6	2.84	3.78	3.95	4.07	4.12	4.10	4.12	4.24	4.45
7	2.90	3.88	4.04	4.14	4.15	4.05	3.94	3.96	4.21
8	2.94	3.94	4.13	4.25	4.28	4.17	3.93	3.75	3.81
9	2.97	3.99	4.20	4.36	4.46	4.41	4.19	3.85	3.37

Tabela 6.23: EQM $\times 10^{-3}$ - estimador Cópula - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	1.25	2.04	2.49	2.75	2.91	3.02	3.09	3.14	3.18
2	2.02	2.96	3.42	3.70	3.89	4.03	4.15	4.24	4.31
3	2.45	3.41	3.71	3.86	3.97	4.07	4.19	4.30	4.40
4	2.71	3.69	3.88	3.98	4.05	4.12	4.22	4.33	4.45
5	2.86	3.87	3.99	4.05	4.15	4.27	4.38	4.48	4.60
6	2.96	4.01	4.08	4.07	4.20	4.41	4.61	4.76	4.91
7	3.04	4.14	4.20	4.15	4.26	4.55	4.85	5.09	5.32
8	3.09	4.23	4.32	4.26	4.37	4.70	5.07	5.32	5.49
9	3.12	4.31	4.43	4.39	4.52	4.87	5.30	5.52	5.21

Tabela 6.24: Vício - estimador ENPMV - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	-0.013	-0.017	-0.015	-0.011	-0.006	-0.002	0.002	0.005	0.007
2	-0.017	-0.028	-0.030	-0.027	-0.020	-0.012	-0.005	0.002	0.007
3	-0.016	-0.030	-0.040	-0.043	-0.038	-0.029	-0.018	-0.008	0.001
4	-0.011	-0.027	-0.043	-0.054	-0.056	-0.050	-0.039	-0.025	-0.012
5	-0.007	-0.021	-0.038	-0.056	-0.068	-0.069	-0.061	-0.047	-0.030
6	-0.002	-0.013	-0.029	-0.050	-0.069	-0.080	-0.079	-0.068	-0.049
7	0.001	-0.005	-0.019	-0.039	-0.061	-0.079	-0.087	-0.083	-0.066
8	0.004	0.002	-0.008	-0.026	-0.047	-0.067	-0.083	-0.087	-0.076
9	0.007	0.007	0.001	-0.012	-0.030	-0.048	-0.066	-0.076	-0.073

Tabela 6.25: Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.49	0.97	1.40	1.77	2.08	2.33	2.54	2.70	2.83
2	0.95	1.59	2.07	2.49	2.84	3.14	3.40	3.62	3.81
3	1.38	2.07	2.47	2.84	3.15	3.40	3.62	3.81	3.97
4	1.74	2.50	2.86	3.22	3.54	3.78	3.98	4.11	4.22
5	2.05	2.86	3.17	3.53	3.86	4.12	4.30	4.40	4.44
6	2.29	3.15	3.42	3.76	4.08	4.35	4.54	4.63	4.62
7	2.49	3.40	3.62	3.92	4.22	4.50	4.71	4.82	4.80
8	2.66	3.62	3.80	4.05	4.32	4.58	4.79	4.96	4.94
9	2.79	3.80	3.97	4.17	4.38	4.60	4.77	4.95	4.81

Tabela 6.26: EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.65	1.26	1.63	1.89	2.12	2.33	2.54	2.72	2.88
2	1.25	2.37	2.99	3.22	3.26	3.30	3.42	3.63	3.87
3	1.62	2.99	4.07	4.65	4.60	4.25	3.96	3.87	3.97
4	1.88	3.24	4.67	6.10	6.67	6.29	5.48	4.76	4.37
5	2.09	3.28	4.64	6.70	8.44	8.89	8.03	6.60	5.34
6	2.30	3.31	4.28	6.30	8.85	10.70	10.78	9.19	7.00
7	2.50	3.42	3.97	5.45	7.94	10.73	12.34	11.68	9.12
8	2.68	3.62	3.87	4.71	6.50	9.13	11.66	12.55	10.73
9	2.83	3.86	3.97	4.32	5.26	6.94	9.06	10.67	10.12

Tabela 6.27: Vício - estimador Cópula - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.009	0.012	0.012	0.011	0.011	0.011	0.011	0.012	0.012
2	0.012	0.018	0.018	0.017	0.017	0.017	0.017	0.018	0.018
3	0.011	0.018	0.017	0.015	0.014	0.013	0.014	0.014	0.015
4	0.011	0.017	0.015	0.011	0.007	0.006	0.006	0.007	0.008
5	0.011	0.016	0.013	0.007	0.001	-0.002	-0.004	-0.004	-0.002
6	0.011	0.016	0.013	0.005	-0.003	-0.009	-0.013	-0.015	-0.014
7	0.011	0.017	0.013	0.005	-0.004	-0.013	-0.020	-0.023	-0.023
8	0.011	0.017	0.014	0.006	-0.004	-0.014	-0.023	-0.029	-0.030
9	0.011	0.018	0.015	0.007	-0.003	-0.013	-0.023	-0.030	-0.031

Tabela 6.28: Variância $\times 10^{-3}$ - estimador Cópula - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.63	1.03	1.30	1.46	1.55	1.60	1.64	1.66	1.67
2	1.01	1.36	1.62	1.79	1.89	1.95	2.00	2.03	2.06
3	1.26	1.59	1.86	2.03	2.12	2.18	2.24	2.29	2.33
4	1.42	1.74	1.99	2.17	2.25	2.29	2.34	2.42	2.49
5	1.51	1.84	2.07	2.24	2.32	2.35	2.40	2.51	2.64
6	1.57	1.91	2.12	2.28	2.34	2.33	2.34	2.43	2.61
7	1.60	1.96	2.17	2.32	2.34	2.27	2.17	2.17	2.33
8	1.62	1.99	2.22	2.39	2.42	2.32	2.14	1.98	2.02
9	1.64	2.02	2.26	2.46	2.53	2.47	2.29	1.99	1.77

Tabela 6.29: EQM $\times 10^{-3}$ - estimador Cópula - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.71	1.17	1.43	1.58	1.67	1.73	1.77	1.79	1.81
2	1.15	1.69	1.95	2.08	2.17	2.24	2.30	2.35	2.39
3	1.39	1.91	2.16	2.26	2.30	2.35	2.42	2.50	2.56
4	1.54	2.02	2.21	2.29	2.30	2.32	2.38	2.46	2.55
5	1.63	2.11	2.24	2.29	2.33	2.36	2.42	2.52	2.64
6	1.69	2.18	2.28	2.31	2.34	2.42	2.52	2.65	2.80
7	1.72	2.24	2.35	2.35	2.36	2.44	2.57	2.72	2.88
8	1.75	2.30	2.42	2.42	2.43	2.52	2.67	2.81	2.92
9	1.76	2.33	2.48	2.51	2.54	2.64	2.81	2.87	2.73

Tabela 6.30: Vício - estimador ENPMV - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	-0.009	-0.011	-0.009	-0.005	-0.001	0.002	0.004	0.006	0.007
2	-0.010	-0.016	-0.017	-0.014	-0.008	-0.002	0.004	0.009	0.012
3	-0.009	-0.017	-0.025	-0.027	-0.022	-0.015	-0.006	0.002	0.008
4	-0.005	-0.014	-0.027	-0.035	-0.037	-0.031	-0.022	-0.011	-0.002
5	-0.002	-0.008	-0.022	-0.037	-0.046	-0.047	-0.039	-0.028	-0.015
6	0.001	-0.002	-0.015	-0.031	-0.047	-0.055	-0.054	-0.044	-0.029
7	0.004	0.004	-0.006	-0.022	-0.039	-0.054	-0.061	-0.057	-0.043
8	0.006	0.008	0.002	-0.011	-0.027	-0.044	-0.057	-0.061	-0.052
9	0.007	0.011	0.008	-0.002	-0.015	-0.029	-0.043	-0.052	-0.050

Tabela 6.31: Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.37	0.67	0.92	1.12	1.28	1.40	1.48	1.55	1.59
2	0.67	0.99	1.23	1.42	1.58	1.71	1.82	1.90	1.97
3	0.91	1.21	1.43	1.63	1.80	1.93	2.04	2.13	2.21
4	1.10	1.40	1.62	1.83	2.01	2.15	2.24	2.31	2.36
5	1.25	1.55	1.77	2.00	2.21	2.36	2.46	2.50	2.52
6	1.37	1.68	1.90	2.13	2.35	2.52	2.62	2.63	2.61
7	1.45	1.79	2.00	2.21	2.42	2.59	2.68	2.67	2.60
8	1.51	1.87	2.08	2.26	2.44	2.58	2.66	2.63	2.53
9	1.55	1.94	2.15	2.31	2.45	2.54	2.58	2.50	2.37

Tabela 6.32: EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.45	0.78	1.00	1.15	1.28	1.40	1.50	1.58	1.64
2	0.78	1.24	1.52	1.61	1.65	1.71	1.83	1.98	2.11
3	0.99	1.51	2.06	2.34	2.30	2.15	2.08	2.13	2.26
4	1.13	1.60	2.33	3.09	3.38	3.13	2.71	2.43	2.37
5	1.26	1.62	2.27	3.36	4.33	4.53	4.00	3.26	2.73
6	1.37	1.69	2.11	3.11	4.52	5.55	5.53	4.60	3.48
7	1.47	1.80	2.03	2.68	3.96	5.49	6.38	5.94	4.48
8	1.54	1.94	2.08	2.39	3.20	4.53	5.92	6.36	5.26
9	1.59	2.07	2.21	2.31	2.66	3.40	4.43	5.21	4.85

Tabela 6.33: Vício - estimador Cópula - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.008	0.010	0.010	0.010	0.010	0.010	0.010	0.010	0.010
2	0.010	0.015	0.015	0.014	0.014	0.014	0.014	0.015	0.015
3	0.010	0.015	0.014	0.012	0.011	0.011	0.011	0.012	0.013
4	0.010	0.014	0.012	0.009	0.006	0.005	0.005	0.006	0.007
5	0.010	0.014	0.011	0.006	0.001	-0.002	-0.003	-0.002	-0.001
6	0.011	0.014	0.011	0.005	-0.002	-0.007	-0.010	-0.011	-0.010
7	0.011	0.015	0.011	0.005	-0.003	-0.010	-0.016	-0.018	-0.018
8	0.011	0.015	0.012	0.006	-0.002	-0.011	-0.018	-0.023	-0.024
9	0.011	0.015	0.013	0.007	-0.001	-0.010	-0.017	-0.023	-0.025

Tabela 6.34: Variância $\times 10^{-3}$ - estimador Cópula - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.30	0.51	0.65	0.74	0.79	0.82	0.84	0.85	0.86
2	0.50	0.69	0.83	0.91	0.97	1.00	1.03	1.04	1.06
3	0.65	0.84	0.99	1.07	1.12	1.16	1.19	1.22	1.24
4	0.74	0.93	1.08	1.17	1.22	1.25	1.28	1.33	1.37
5	0.79	0.98	1.13	1.22	1.26	1.26	1.28	1.34	1.41
6	0.82	1.02	1.17	1.25	1.28	1.25	1.23	1.28	1.37
7	0.83	1.05	1.20	1.29	1.31	1.26	1.18	1.17	1.26
8	0.84	1.07	1.23	1.33	1.37	1.30	1.17	1.04	1.04
9	0.85	1.08	1.26	1.37	1.44	1.41	1.27	1.06	0.90

Tabela 6.35: EQM $\times 10^{-3}$ - estimador Cópula - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.36	0.61	0.75	0.84	0.89	0.92	0.94	0.96	0.97
2	0.61	0.90	1.04	1.11	1.16	1.20	1.24	1.27	1.29
3	0.76	1.05	1.17	1.22	1.24	1.28	1.32	1.36	1.40
4	0.84	1.13	1.22	1.25	1.26	1.27	1.31	1.36	1.42
5	0.89	1.18	1.25	1.25	1.26	1.27	1.29	1.35	1.41
6	0.93	1.23	1.29	1.27	1.28	1.30	1.34	1.39	1.47
7	0.95	1.27	1.33	1.31	1.32	1.36	1.42	1.50	1.59
8	0.96	1.30	1.38	1.37	1.37	1.42	1.49	1.55	1.61
9	0.97	1.32	1.42	1.42	1.44	1.50	1.57	1.60	1.51

Tabela 6.36: Vício - estimador ENPMV - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	-0.006	-0.006	-0.004	-0.002	0.001	0.003	0.005	0.006	0.006
2	-0.006	-0.010	-0.011	-0.008	-0.003	0.002	0.006	0.009	0.011
3	-0.004	-0.011	-0.017	-0.018	-0.014	-0.008	-0.001	0.005	0.009
4	-0.002	-0.008	-0.018	-0.025	-0.026	-0.021	-0.013	-0.004	0.002
5	0.001	-0.003	-0.014	-0.026	-0.033	-0.033	-0.027	-0.017	-0.007
6	0.003	0.002	-0.007	-0.021	-0.033	-0.040	-0.039	-0.030	-0.019
7	0.005	0.006	-0.001	-0.013	-0.027	-0.039	-0.044	-0.041	-0.030
8	0.006	0.009	0.005	-0.004	-0.017	-0.030	-0.041	-0.044	-0.037
9	0.006	0.011	0.009	0.003	-0.007	-0.019	-0.029	-0.037	-0.035

Tabela 6.37: Variância $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.27	0.43	0.55	0.64	0.70	0.75	0.78	0.80	0.81
2	0.43	0.58	0.69	0.78	0.85	0.90	0.95	0.98	1.01
3	0.55	0.70	0.81	0.91	0.99	1.04	1.09	1.13	1.16
4	0.64	0.80	0.92	1.04	1.13	1.19	1.22	1.24	1.26
5	0.70	0.87	1.00	1.13	1.24	1.31	1.33	1.33	1.33
6	0.75	0.93	1.06	1.19	1.31	1.39	1.43	1.41	1.37
7	0.78	0.97	1.10	1.22	1.34	1.44	1.48	1.45	1.38
8	0.80	1.01	1.14	1.25	1.36	1.43	1.46	1.41	1.30
9	0.81	1.04	1.17	1.28	1.37	1.40	1.40	1.32	1.17

Tabela 6.38: EQM $\times 10^{-3}$ - estimador ENPMV - caso 2, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.30	0.47	0.57	0.64	0.70	0.76	0.80	0.83	0.85
2	0.46	0.67	0.80	0.84	0.86	0.91	0.99	1.07	1.13
3	0.56	0.81	1.09	1.23	1.19	1.10	1.09	1.15	1.23
4	0.64	0.85	1.24	1.67	1.81	1.62	1.38	1.26	1.27
5	0.70	0.88	1.20	1.80	2.35	2.41	2.04	1.62	1.38
6	0.76	0.93	1.11	1.62	2.42	3.00	2.92	2.33	1.72
7	0.80	1.01	1.10	1.38	2.06	2.94	3.44	3.12	2.28
8	0.83	1.09	1.16	1.26	1.64	2.35	3.12	3.34	2.67
9	0.85	1.16	1.25	1.28	1.42	1.75	2.27	2.66	2.42

Tabela 6.39: Vício - estimador Cópula - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.000	0.000	0.001	0.001	0.001	0.000	-0.001	-0.003	-0.004
2	0.000	0.001	0.002	0.002	0.002	0.000	-0.003	-0.007	-0.007
3	0.001	0.002	0.003	0.003	0.002	-0.002	-0.008	-0.013	-0.012
4	0.001	0.002	0.003	0.002	-0.001	-0.008	-0.016	-0.022	-0.021
5	0.001	0.002	0.002	-0.001	-0.007	-0.017	-0.027	-0.035	-0.034
6	0.001	0.001	-0.002	-0.007	-0.016	-0.028	-0.041	-0.049	-0.048
7	-0.001	-0.003	-0.007	-0.015	-0.026	-0.040	-0.053	-0.061	-0.060
8	-0.003	-0.006	-0.012	-0.021	-0.033	-0.048	-0.061	-0.068	-0.066
9	-0.004	-0.007	-0.011	-0.021	-0.034	-0.048	-0.061	-0.067	-0.065

Tabela 6.40: Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.01	0.02	0.04	0.10	0.23	0.59	1.39
2	0.00	0.01	0.03	0.08	0.19	0.42	0.88	1.79	3.02
3	0.01	0.03	0.10	0.24	0.54	1.07	1.92	3.13	4.21
4	0.02	0.08	0.24	0.57	1.17	2.02	3.08	4.32	5.20
5	0.04	0.20	0.54	1.17	2.12	3.16	4.25	5.43	6.10
6	0.10	0.43	1.07	2.03	3.18	4.24	5.26	6.33	6.77
7	0.24	0.91	1.95	3.14	4.31	5.30	6.21	7.15	7.33
8	0.60	1.84	3.18	4.42	5.53	6.35	7.10	7.85	7.75
9	1.40	3.07	4.24	5.27	6.12	6.68	7.18	7.65	7.22

Tabela 6.41: EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.01	0.02	0.04	0.10	0.23	0.60	1.41
2	0.00	0.01	0.03	0.09	0.19	0.42	0.89	1.84	3.07
3	0.01	0.04	0.11	0.25	0.55	1.08	1.99	3.30	4.34
4	0.02	0.09	0.25	0.57	1.17	2.08	3.34	4.81	5.66
5	0.04	0.20	0.54	1.17	2.17	3.44	4.98	6.62	7.26
6	0.10	0.43	1.07	2.09	3.45	5.05	6.91	8.70	9.06
7	0.25	0.92	2.00	3.36	5.00	6.92	9.03	10.86	10.88
8	0.61	1.88	3.32	4.85	6.65	8.66	10.79	12.48	12.11
9	1.42	3.12	4.36	5.70	7.28	9.03	10.85	12.17	11.41

Tabela 6.42: Vício - estimador ENPMV - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.001	0.003	0.006	0.009	0.013	0.017	0.020	0.020	0.014
2	0.003	0.008	0.014	0.021	0.028	0.034	0.035	0.029	0.016
3	0.006	0.014	0.024	0.034	0.042	0.045	0.040	0.026	0.008
4	0.009	0.021	0.034	0.045	0.050	0.045	0.031	0.011	-0.008
5	0.013	0.029	0.042	0.050	0.048	0.034	0.013	-0.010	-0.026
6	0.017	0.034	0.045	0.046	0.034	0.014	-0.011	-0.033	-0.046
7	0.020	0.036	0.040	0.032	0.013	-0.011	-0.035	-0.054	-0.062
8	0.020	0.030	0.026	0.012	-0.009	-0.033	-0.054	-0.069	-0.073
9	0.014	0.016	0.008	-0.008	-0.027	-0.046	-0.063	-0.074	-0.076

Tabela 6.43: Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.02	0.04	0.09	0.18	0.32	0.55	0.94	1.59
2	0.02	0.06	0.14	0.28	0.48	0.77	1.19	1.82	2.70
3	0.04	0.14	0.31	0.54	0.87	1.29	1.84	2.60	3.49
4	0.09	0.28	0.55	0.90	1.35	1.89	2.56	3.43	4.34
5	0.18	0.49	0.88	1.36	1.93	2.58	3.35	4.29	5.17
6	0.33	0.80	1.33	1.93	2.61	3.34	4.17	5.16	5.93
7	0.57	1.24	1.91	2.63	3.42	4.22	5.07	6.03	6.63
8	0.96	1.89	2.70	3.55	4.44	5.25	6.05	6.90	7.24
9	1.60	2.77	3.61	4.47	5.33	6.02	6.62	7.22	7.15

Tabela 6.44: EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 50

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.02	0.08	0.18	0.35	0.62	0.95	1.33	1.78
2	0.02	0.12	0.34	0.72	1.29	1.92	2.42	2.66	2.95
3	0.08	0.34	0.88	1.70	2.64	3.31	3.41	3.26	3.55
4	0.18	0.74	1.71	2.91	3.82	3.96	3.54	3.56	4.40
5	0.36	1.32	2.69	3.86	4.21	3.75	3.52	4.39	5.86
6	0.64	1.98	3.39	4.03	3.80	3.53	4.29	6.24	8.02
7	0.98	2.51	3.53	3.65	3.60	4.33	6.27	8.91	10.51
8	1.36	2.76	3.40	3.69	4.53	6.31	8.93	11.63	12.62
9	1.80	3.02	3.67	4.53	6.06	8.17	10.62	12.75	12.86

Tabela 6.45: Vício - estimador Cópula - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.000	0.000	0.000	0.001	0.001	0.000	-0.000	-0.002	-0.002
2	0.000	0.001	0.001	0.002	0.002	0.001	-0.001	-0.004	-0.002
3	0.000	0.001	0.002	0.002	0.002	-0.000	-0.004	-0.008	-0.006
4	0.001	0.002	0.002	0.002	0.000	-0.004	-0.010	-0.015	-0.013
5	0.001	0.002	0.002	0.000	-0.004	-0.010	-0.019	-0.025	-0.023
6	0.001	0.001	-0.000	-0.004	-0.010	-0.020	-0.030	-0.036	-0.035
7	-0.000	-0.001	-0.005	-0.011	-0.020	-0.030	-0.041	-0.048	-0.045
8	-0.002	-0.004	-0.009	-0.017	-0.027	-0.038	-0.049	-0.055	-0.053
9	-0.002	-0.002	-0.007	-0.016	-0.026	-0.037	-0.047	-0.053	-0.050

Tabela 6.46: Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.00	0.01	0.02	0.05	0.12	0.30	0.76
2	0.00	0.00	0.01	0.04	0.10	0.22	0.46	0.89	1.57
3	0.00	0.01	0.05	0.12	0.29	0.58	1.02	1.60	2.28
4	0.01	0.04	0.13	0.31	0.64	1.13	1.69	2.31	2.91
5	0.02	0.10	0.29	0.64	1.15	1.75	2.31	2.86	3.35
6	0.05	0.23	0.58	1.11	1.73	2.32	2.81	3.26	3.63
7	0.12	0.46	1.01	1.66	2.28	2.81	3.18	3.52	3.75
8	0.30	0.87	1.56	2.23	2.80	3.22	3.47	3.67	3.76
9	0.76	1.55	2.26	2.84	3.31	3.58	3.68	3.69	3.58

Tabela 6.47: EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.00	0.01	0.02	0.05	0.12	0.31	0.76
2	0.00	0.00	0.02	0.04	0.10	0.22	0.46	0.90	1.57
3	0.00	0.02	0.05	0.13	0.29	0.58	1.04	1.66	2.32
4	0.01	0.04	0.13	0.31	0.64	1.14	1.79	2.54	3.09
5	0.02	0.10	0.29	0.64	1.16	1.86	2.66	3.49	3.90
6	0.05	0.23	0.58	1.13	1.84	2.71	3.68	4.58	4.82
7	0.12	0.46	1.03	1.77	2.66	3.74	4.85	5.79	5.82
8	0.30	0.88	1.64	2.51	3.54	4.70	5.84	6.71	6.52
9	0.76	1.56	2.31	3.09	3.98	4.94	5.86	6.48	6.07

Tabela 6.48: Vício - estimador ENPMV - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.000	0.001	0.003	0.005	0.008	0.012	0.015	0.015	0.011
2	0.001	0.004	0.009	0.014	0.021	0.026	0.028	0.024	0.015
3	0.003	0.009	0.016	0.025	0.032	0.036	0.033	0.022	0.009
4	0.005	0.014	0.025	0.035	0.040	0.038	0.027	0.011	-0.003
5	0.009	0.021	0.032	0.040	0.039	0.029	0.012	-0.006	-0.017
6	0.012	0.026	0.036	0.037	0.029	0.012	-0.007	-0.024	-0.031
7	0.015	0.028	0.032	0.025	0.011	-0.008	-0.026	-0.040	-0.045
8	0.015	0.023	0.020	0.008	-0.008	-0.026	-0.041	-0.052	-0.054
9	0.011	0.014	0.007	-0.005	-0.019	-0.034	-0.046	-0.054	-0.053

Tabela 6.49: Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.01	0.04	0.08	0.16	0.29	0.52	0.90
2	0.00	0.02	0.06	0.13	0.25	0.43	0.68	1.04	1.54
3	0.01	0.06	0.15	0.29	0.50	0.76	1.10	1.54	2.09
4	0.04	0.13	0.29	0.51	0.79	1.11	1.50	1.99	2.57
5	0.08	0.25	0.49	0.78	1.11	1.47	1.88	2.40	2.97
6	0.16	0.43	0.76	1.10	1.47	1.84	2.26	2.76	3.27
7	0.29	0.69	1.09	1.48	1.86	2.24	2.63	3.06	3.46
8	0.52	1.05	1.54	1.96	2.35	2.71	3.03	3.34	3.57
9	0.91	1.55	2.09	2.51	2.89	3.18	3.37	3.50	3.49

Tabela 6.50: EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 100

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.01	0.02	0.06	0.15	0.30	0.51	0.74	1.03
2	0.01	0.04	0.13	0.33	0.67	1.10	1.46	1.61	1.76
3	0.02	0.13	0.41	0.91	1.55	2.06	2.16	2.00	2.16
4	0.07	0.34	0.90	1.71	2.39	2.54	2.22	2.10	2.58
5	0.15	0.68	1.53	2.37	2.65	2.32	2.03	2.43	3.25
6	0.30	1.10	2.02	2.48	2.29	2.00	2.31	3.33	4.25
7	0.51	1.44	2.09	2.12	1.98	2.30	3.33	4.70	5.45
8	0.74	1.59	1.95	2.03	2.42	3.37	4.75	6.08	6.44
9	1.04	1.75	2.14	2.54	3.27	4.32	5.48	6.40	6.28

Tabela 6.51: Vício - estimador Cópula - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.000	0.000	0.000	0.000	0.001	0.001	0.000	-0.001	-0.000
2	0.000	0.000	0.001	0.001	0.002	0.001	-0.000	-0.002	-0.001
3	0.000	0.001	0.001	0.002	0.002	0.001	-0.002	-0.005	-0.004
4	0.000	0.001	0.002	0.002	0.002	-0.001	-0.006	-0.011	-0.010
5	0.001	0.001	0.002	0.001	-0.001	-0.007	-0.013	-0.019	-0.018
6	0.001	0.001	0.001	-0.001	-0.006	-0.014	-0.022	-0.028	-0.027
7	0.000	0.000	-0.002	-0.006	-0.013	-0.021	-0.030	-0.036	-0.035
8	-0.001	-0.001	-0.005	-0.010	-0.018	-0.027	-0.036	-0.042	-0.040
9	-0.000	-0.000	-0.004	-0.009	-0.017	-0.026	-0.035	-0.040	-0.039

Tabela 6.52: Variância $\times 10^{-3}$ - estimador Cópula - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.00	0.00	0.01	0.03	0.07	0.17	0.39
2	0.00	0.00	0.01	0.02	0.05	0.12	0.26	0.49	0.82
3	0.00	0.01	0.02	0.07	0.16	0.32	0.57	0.89	1.22
4	0.00	0.02	0.06	0.17	0.36	0.63	0.96	1.29	1.56
5	0.01	0.05	0.15	0.35	0.64	0.98	1.32	1.61	1.79
6	0.03	0.12	0.32	0.62	0.99	1.35	1.65	1.86	1.95
7	0.07	0.26	0.56	0.94	1.33	1.66	1.91	2.02	2.02
8	0.17	0.48	0.86	1.26	1.60	1.86	2.03	2.06	1.96
9	0.39	0.81	1.19	1.51	1.76	1.92	2.01	1.97	1.79

Tabela 6.53: EQM $\times 10^{-3}$ - estimador Cópula - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.00	0.00	0.01	0.03	0.07	0.17	0.39
2	0.00	0.00	0.01	0.02	0.05	0.12	0.26	0.49	0.82
3	0.00	0.01	0.02	0.07	0.16	0.33	0.58	0.91	1.24
4	0.00	0.02	0.07	0.17	0.36	0.63	0.99	1.41	1.67
5	0.01	0.05	0.16	0.35	0.64	1.02	1.49	1.97	2.14
6	0.03	0.12	0.32	0.62	1.03	1.53	2.12	2.64	2.69
7	0.07	0.26	0.57	0.98	1.50	2.12	2.80	3.32	3.24
8	0.17	0.49	0.89	1.36	1.93	2.60	3.31	3.79	3.59
9	0.39	0.81	1.20	1.60	2.07	2.62	3.21	3.59	3.30

Tabela 6.54: Vício - estimador ENPMV - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.000	0.000	0.001	0.003	0.005	0.008	0.010	0.011	0.008
2	0.000	0.002	0.005	0.009	0.014	0.019	0.021	0.018	0.011
3	0.001	0.005	0.010	0.017	0.024	0.028	0.026	0.017	0.007
4	0.003	0.009	0.017	0.026	0.032	0.030	0.022	0.008	-0.002
5	0.005	0.014	0.024	0.032	0.032	0.024	0.010	-0.004	-0.012
6	0.008	0.019	0.028	0.030	0.024	0.011	-0.004	-0.017	-0.023
7	0.010	0.021	0.026	0.022	0.011	-0.004	-0.018	-0.029	-0.032
8	0.011	0.018	0.017	0.009	-0.004	-0.017	-0.029	-0.037	-0.037
9	0.009	0.012	0.007	-0.001	-0.012	-0.023	-0.032	-0.038	-0.036

Tabela 6.55: Variância $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.00	0.01	0.04	0.08	0.16	0.30	0.51
2	0.00	0.01	0.03	0.06	0.13	0.24	0.40	0.61	0.88
3	0.00	0.03	0.07	0.16	0.28	0.45	0.65	0.89	1.17
4	0.02	0.06	0.16	0.30	0.47	0.67	0.89	1.14	1.41
5	0.04	0.13	0.28	0.47	0.68	0.89	1.11	1.35	1.57
6	0.08	0.24	0.45	0.67	0.89	1.11	1.33	1.55	1.71
7	0.16	0.40	0.65	0.89	1.12	1.34	1.57	1.75	1.83
8	0.30	0.60	0.88	1.12	1.35	1.55	1.74	1.85	1.84
9	0.50	0.86	1.15	1.38	1.57	1.72	1.84	1.86	1.74

Tabela 6.56: EQM $\times 10^{-3}$ - estimador ENPMV - caso 3, n = 200

Decil	1	2	3	4	5	6	7	8	9
1	0.00	0.00	0.01	0.02	0.06	0.14	0.26	0.41	0.58
2	0.00	0.01	0.05	0.14	0.33	0.60	0.84	0.94	1.02
3	0.01	0.05	0.18	0.46	0.87	1.22	1.30	1.18	1.22
4	0.02	0.15	0.46	0.98	1.47	1.59	1.36	1.21	1.41
5	0.06	0.33	0.87	1.47	1.71	1.47	1.22	1.37	1.72
6	0.14	0.60	1.23	1.60	1.49	1.23	1.35	1.86	2.23
7	0.27	0.84	1.31	1.36	1.24	1.36	1.91	2.60	2.83
8	0.42	0.94	1.17	1.20	1.37	1.85	2.59	3.23	3.24
9	0.57	1.00	1.20	1.38	1.72	2.23	2.84	3.27	3.07

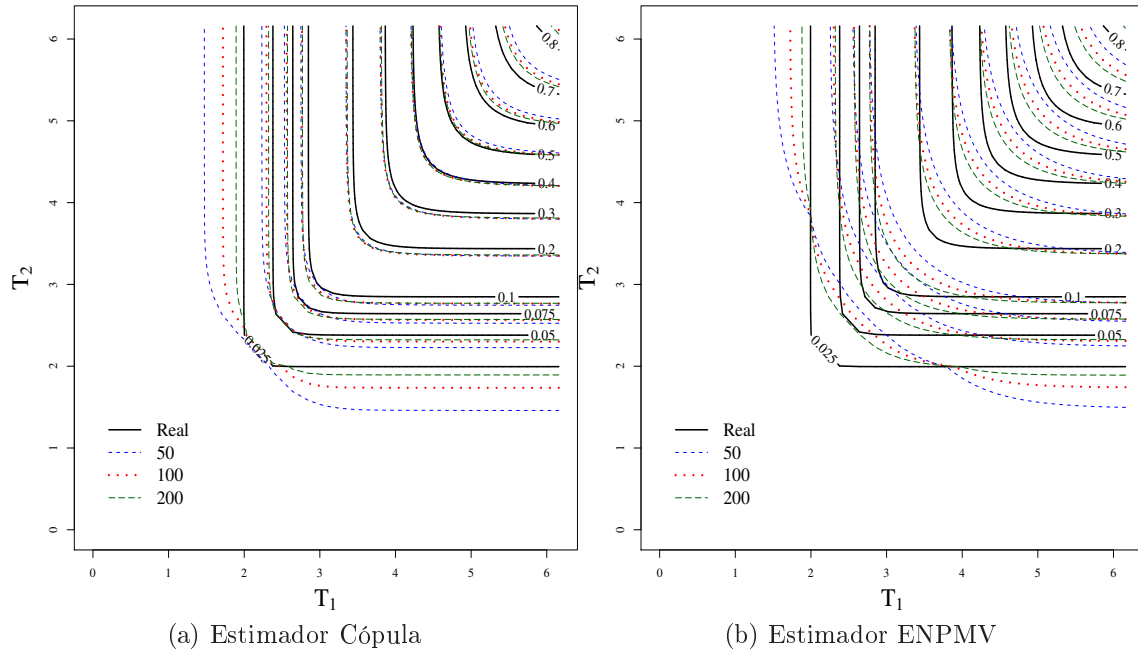


Figura 6.19: Valores esperados para o Caso 1

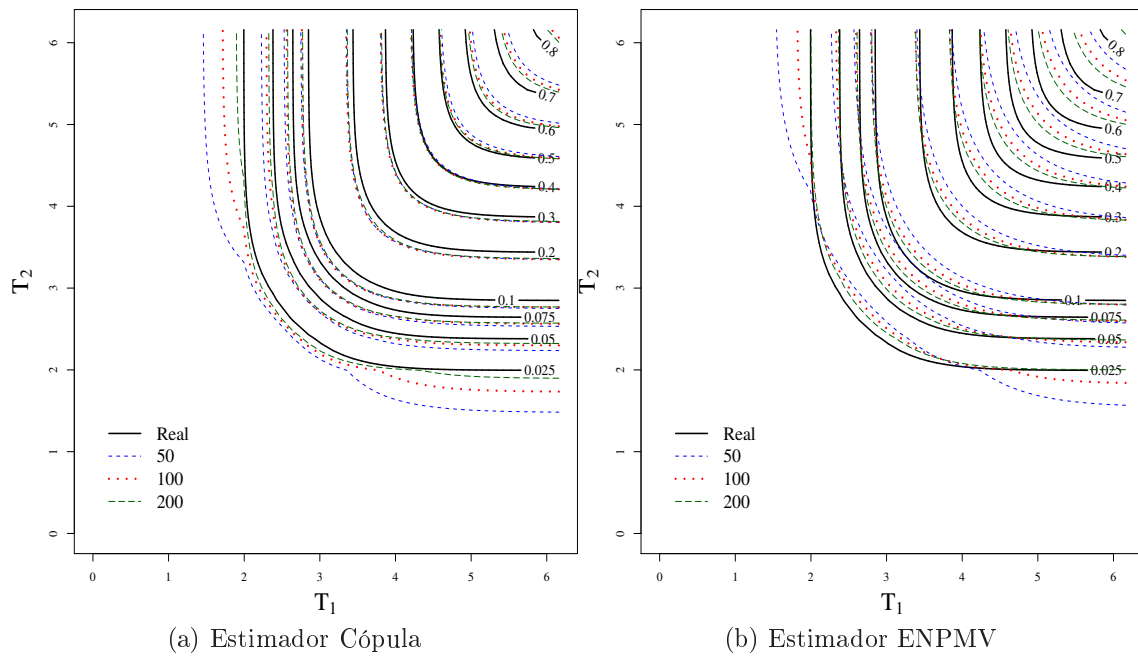


Figura 6.20: Valores esperados para o Caso 2

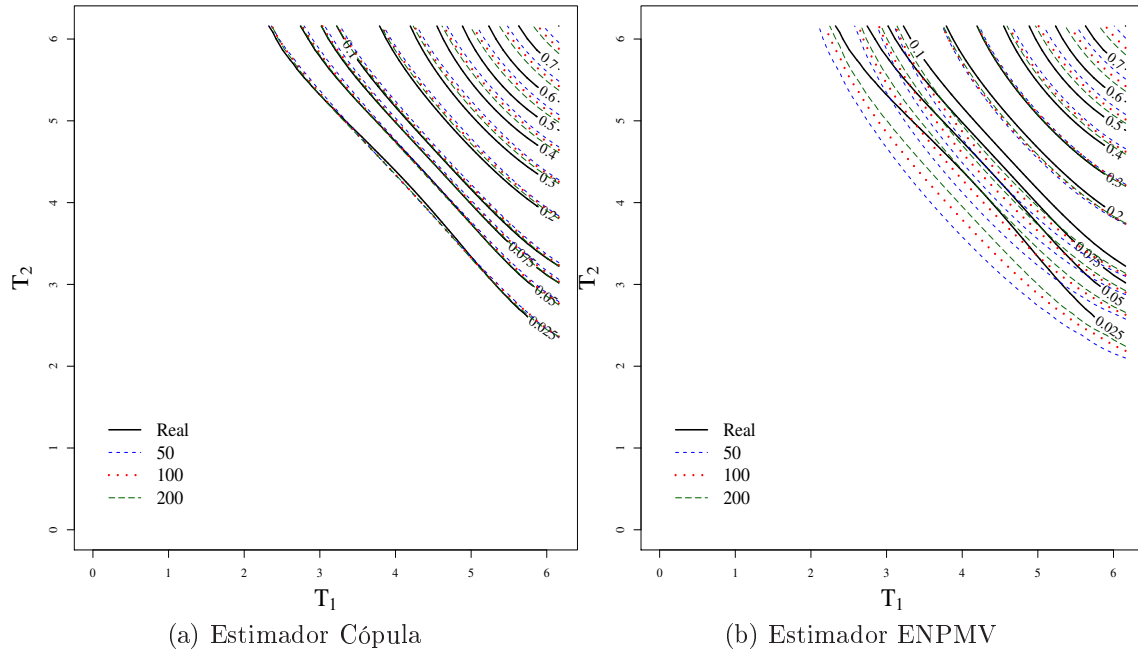


Figura 6.21: Valores esperados para o Caso 3

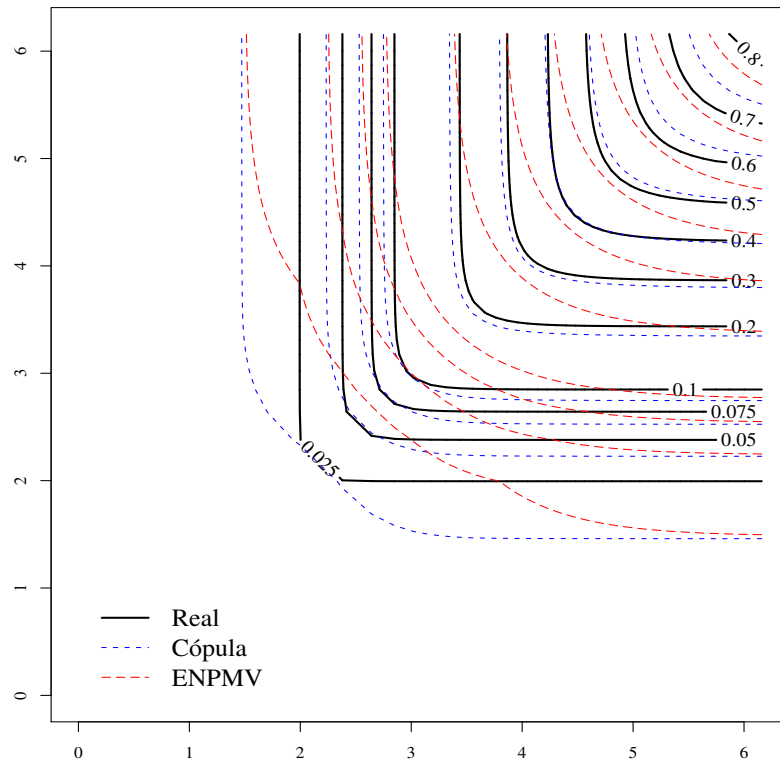


Figura 6.22: Valores esperados - Caso 1, $n = 50$

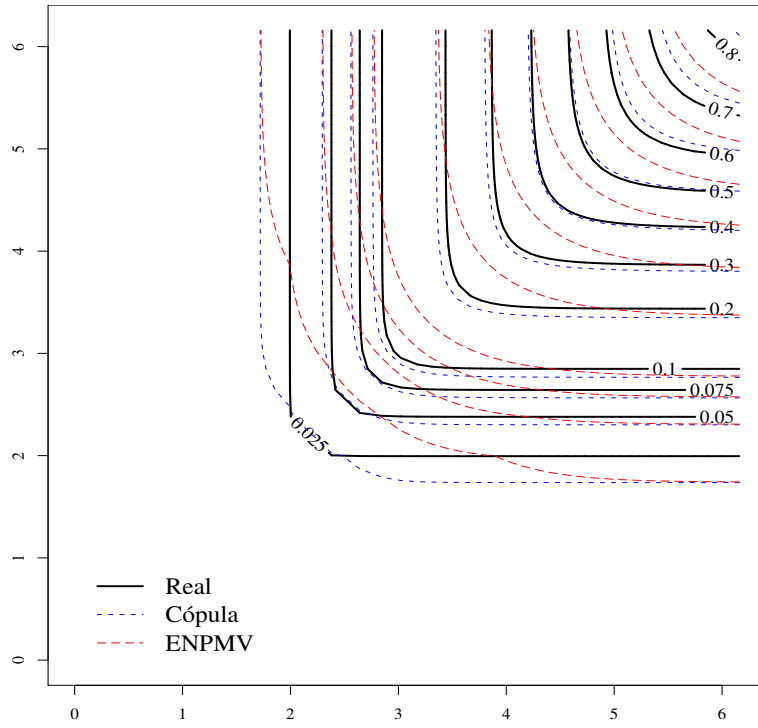


Figura 6.23: Valores esperados - Caso 1, $n = 100$

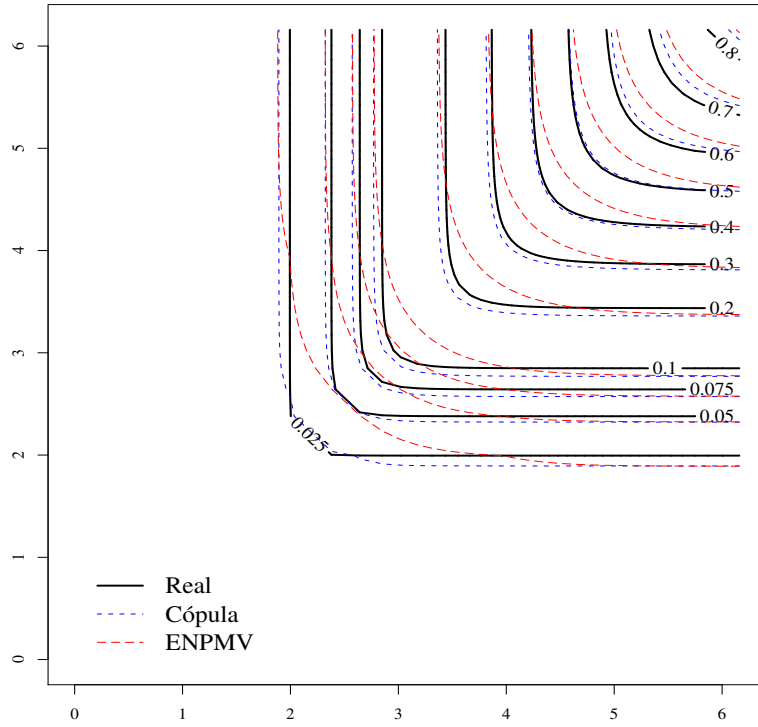


Figura 6.24: Valores esperados - Caso 1, $n = 200$

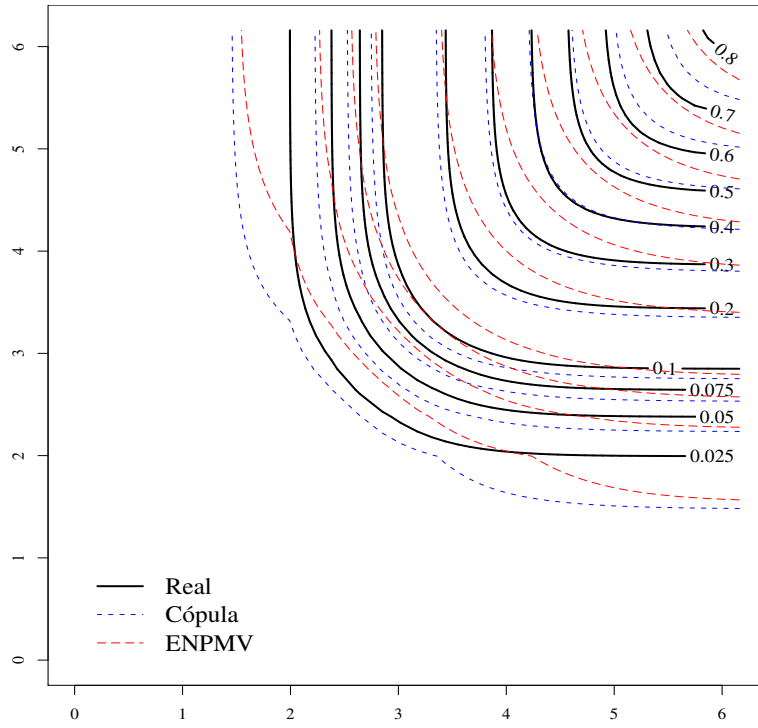


Figura 6.25: Valores esperados - Caso 2, $n = 50$

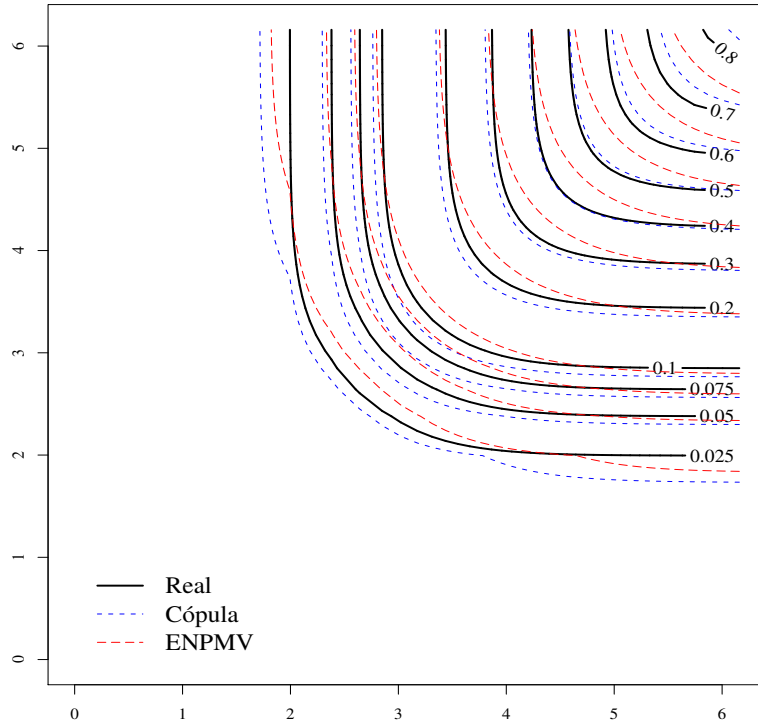


Figura 6.26: Valores esperados - Caso 2, $n = 100$

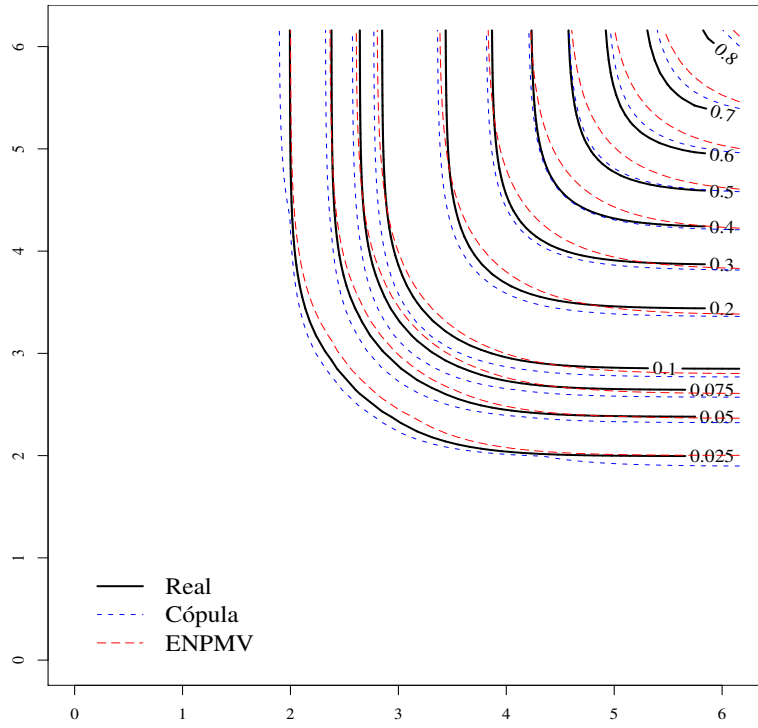


Figura 6.27: Valores esperados - Caso 2, $n = 200$

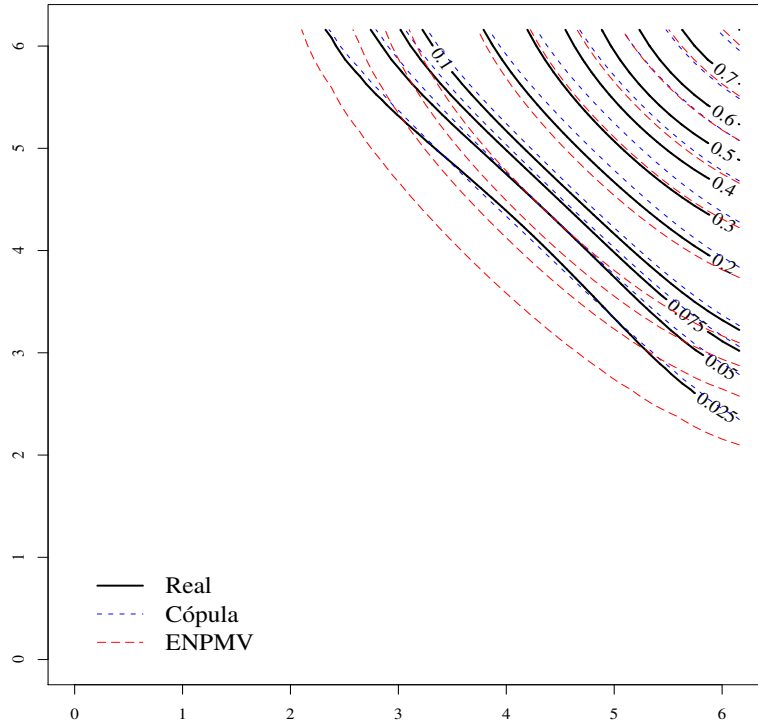


Figura 6.28: Valores esperados - Caso 3, $n = 50$

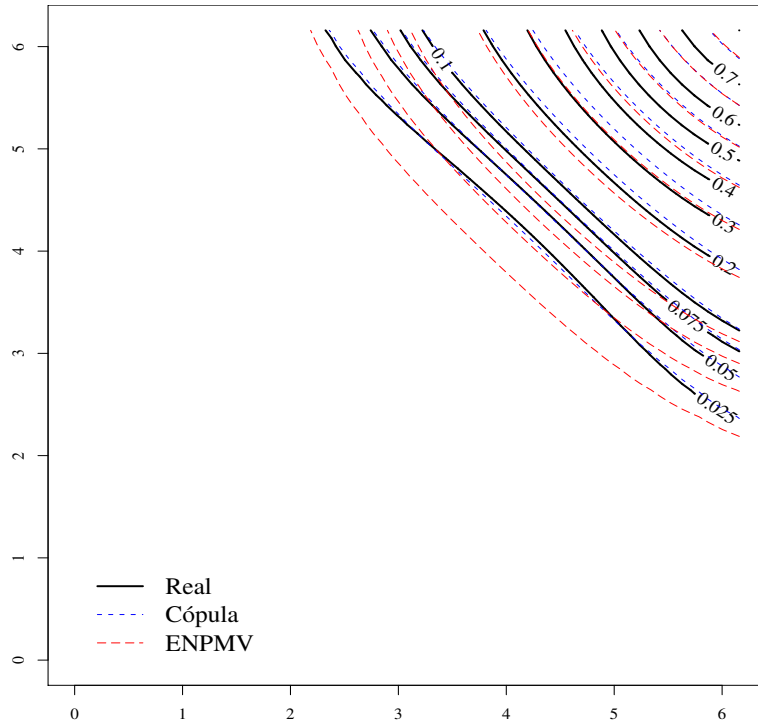


Figura 6.29: Valores esperados - Caso 3, $n = 100$

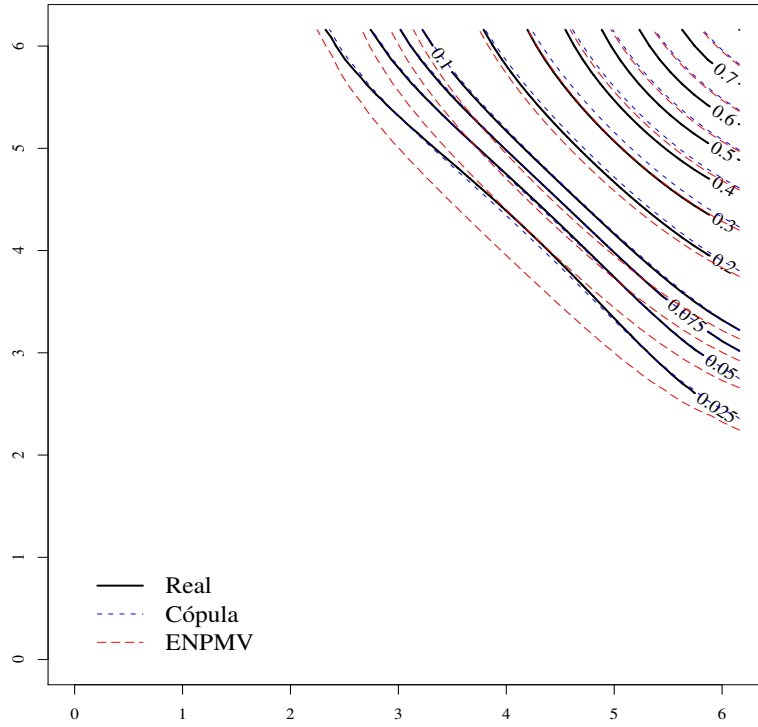


Figura 6.30: Valores esperados - Caso 3, $n = 200$

6.2 Aplicação

Goggins e Finkelstein (2000) discutem um conjunto de dados originados de um experimento clínico sobre AIDS, o AIDS Clinical Trial Group (ACTG) 181, com indivíduos infectados com HIV. Os dados contêm o tempo (em semanas) até a ocorrência de infecções oportunistas por citomegalovirus (CMV). Durante o estudo, entre outras atividades, foram colhidas amostras de sangue e de urina dos pacientes em suas visitas clínicas, sendo as amostras testadas para detecção da presença de CMV. Os dados, relacionados na tabela 6.59 no fim deste capítulo, contêm o intervalo entre visitas sucessivas de cada paciente durante o qual ocorreu a detecção de CMV no sangue e na urina. Denotaremos o tempo até a detecção de CMV no sangue por CMV_s , enquanto CMV_u designa o tempo até a ocorrência da infecção pelo CMV na urina.

As estimativas (escada e suavizada) para as funções de distribuição marginais estão na figura 6.31. As janelas estimadas foram $h_s = 1,4$ e $h_u = 1$. O método univariado de escolha da janela não funcionou bem para os dados de CMV_u .

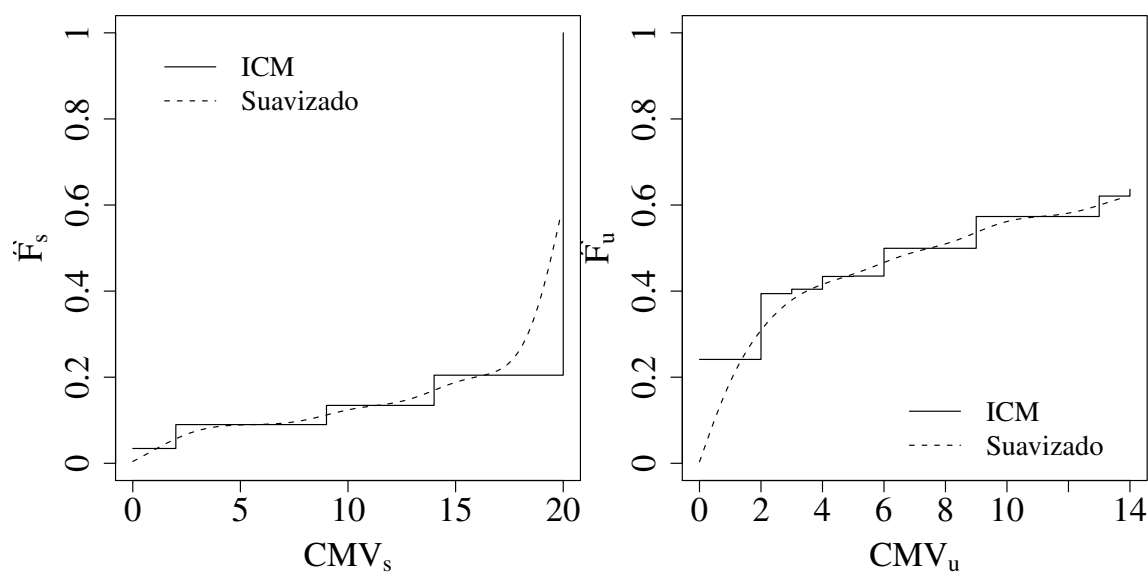


Figura 6.31: Estimativas das funções de distribuição marginais

Para os intervalos de censura de CMV_u , não houve um ponto de máximo para que h_{opt} fosse escolhido. Isso pode ter ocorrido porque a estimativa de F_u , que está na figura 6.31, atinge o valor máximo 0,636, por existirem muitas amostras com censura à direita depois do ponto $CMV_u = 20$. Dessa forma, a função que determina o valor

de h_{opt} não tinha um ponto de máximo.

Foram feitos os ajustes com as cópulas de Clayton, Frank e Gumbel-Hougaard. A tabela 6.57 relaciona os valores estimados para o parâmetro α da cópula em cada caso.

Tabela 6.57: Estimativas para α

Cópula	$\hat{\alpha}$
Clayton	0,1513928
Frank	2,687302
Gumbel-Hougaard	1,367143

As figuras 6.32 a 6.34 exibem as curvas de nível para os ajustes em cada modelo de cópula, depois que as funções de distribuição de marginais foram suavizadas e avaliadas nos pontos da amostra.

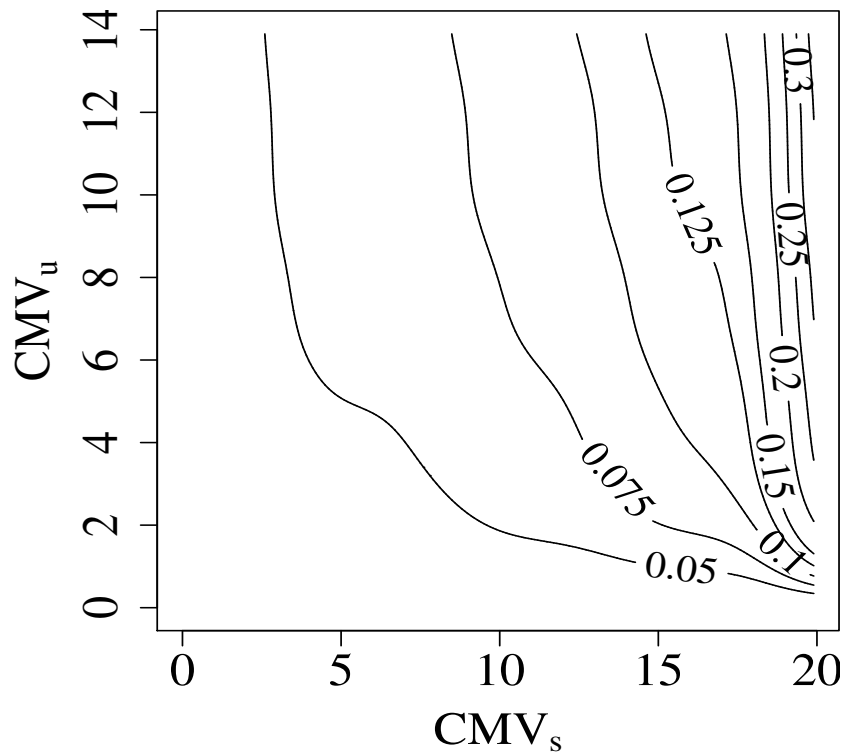


Figura 6.32: Estimativa com cópula Clayton

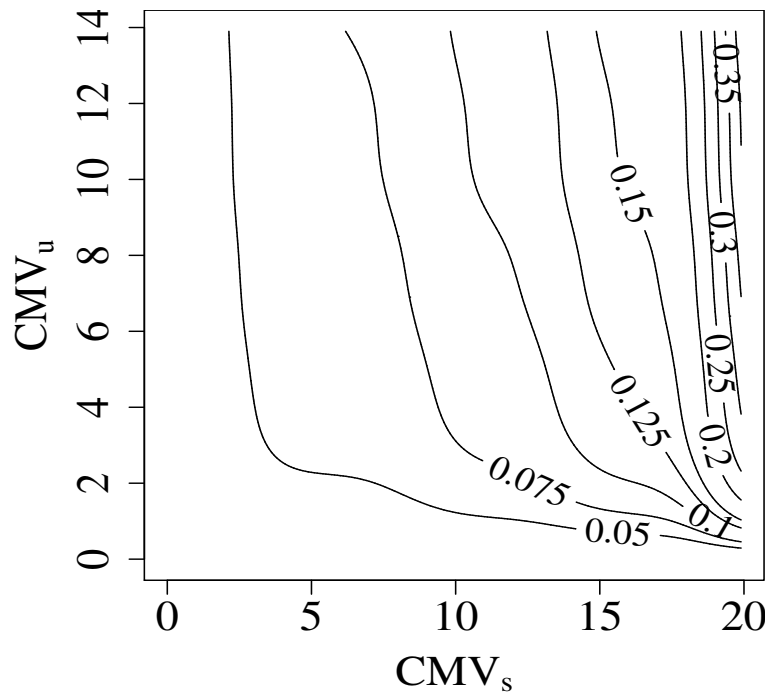


Figura 6.33: Estimativa com cópula Frank

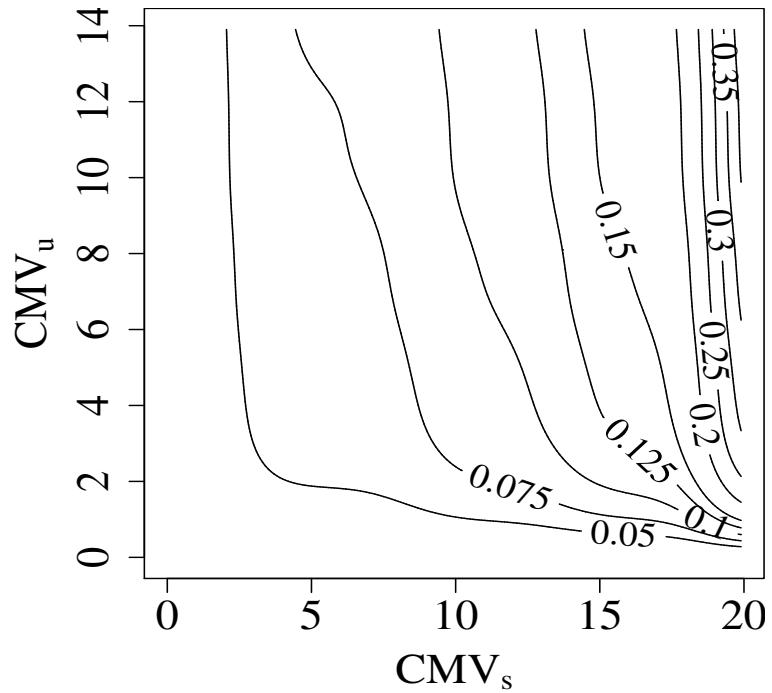


Figura 6.34: Estimativa com cópula Gumbel-Hougaard

Por outro lado, para o estimador ENPMV, a tabela 6.58 mostra as intersecções máximas $\xi_j = [a_j, b_j] \times [c_j, d_j]$, com seu respectivo salto estimado \hat{p}_j .

O método bivariado para a escolha das janelas não teve boa performance, pois

Tabela 6.58: Estimativas para o estimador ENPMV

a_j	b_j	c_j	d_j	\hat{p}_j
$-\infty$	0	$-\infty$	0	0,0245098
$-\infty$	0	1	2	0,0098039
2	2	2	2	0,0080436
2	2	$-\infty$	0	0,0308123
3	3	11	11	0,0173165
6	6	$-\infty$	0	0,0231092
6	6	6	6	0,0090596
12	12	9	9	0,0357672
14	14	9	9	0,0178366
14	14	13	13	0,0451156
14	14	14	14	0,0354233
13	14	4	4	0,0253821
19	20	$-\infty$	0	0,1617647
19	∞	19	∞	0,3422901
19	∞	6	6	0,0649892
20	∞	2	2	0,1181730
20	∞	3	3	0,0303448
16	∞	4	4	0,0002583

não havia um ponto de máximo estrito. Utilizou-se $h_s = 3$ e $h_u = 0,8$ para que fosse obtida a estimativa de H cujas curvas de nível estão na figura 6.35.

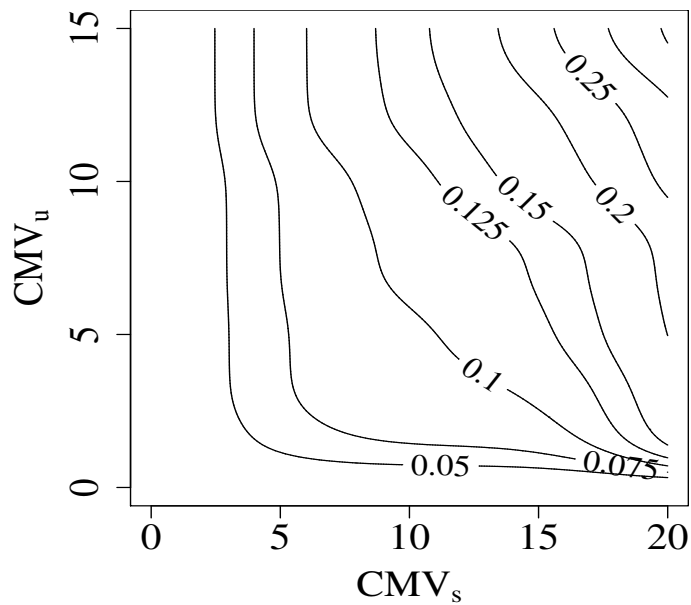


Figura 6.35: Estimativa - estimador ENPMV

Tabela 6.59: Dados do ACTG 181

$CMV_s \times CMV_u$			
(11,∞) × (11,∞)	(1,∞) × (1,∞)	(13,∞) × (−∞,0)	(2,∞) × (1,3)
(11,∞) × (11,∞)	(9,∞) × (4,∞)	(16,∞) × (3,4)	(8,∞) × (−∞,0)
(11,∞) × (11,∞)	(12,∞) × (11,∞)	(15,∞) × (−∞,0)	(13,∞) × (13,∞)
(11,∞) × (8,10)	(2,∞) × (1,4)	(16,∞) × (5,7)	(11,∞) × (−∞,0)
(7,∞) × (6,8)	(10,∞) × (11,∞)	(8,∞) × (−∞,0)	(13,∞) × (13,∞)
(11,∞) × (12,∞)	(8,∞) × (11,∞)	(1,2) × (1,2)	(16,∞) × (−∞,0)
(8,12) × (8,10)	(12,∞) × (12,∞)	(1,4) × (−∞,0)	(11,∞) × (4,6)
(10,∞) × (10,∞)	(12,∞) × (12,∞)	(15,∞) × (1,2)	(1,∞) × (−∞,0)
(6,∞) × (6,∞)	(5,∞) × (11,∞)	(5,9) × (−∞,0)	(8,∞) × (15,∞)
(2,9) × (9,11)	(2,∞) × (3,∞)	(16,∞) × (16,∞)	(8,14) × (9,10)
(11,∞) × (11,∞)	(6,10) × (−∞,0)	(1,2) × (−∞,0)	(13,∞) × (12,15)
(5,∞) × (−∞,0)	(2,∞) × (2,∞)	(15,∞) × (1,3)	(8,∞) × (16,∞)
(16,∞) × (16,∞)	(13,∞) × (13,∞)	(14,∞) × (−∞,0)	(8,∞) × (7,10)
(−∞,0) × (−∞,0)	(15,∞) × (−∞,0)	(8,∞) × (1,4)	(2,11) × (11,13)
(18,∞) × (2,2)	(8,∞) × (−∞,0)	(15,∞) × (2,4)	(11,∞) × (11,∞)
(16,∞) × (−∞,0)	(16,∞) × (6,9)	(13,∞) × (13,∞)	(5,∞) × (10,∞)
(16,∞) × (−∞,0)	(5,∞) × (−∞,0)	(4,∞) × (−∞,0)	(6,∞) × (6,∞)
(10,∞) × (10,∞)	(2,∞) × (−∞,0)	(13,∞) × (9,13)	(14,∞) × (14,∞)
(10,∞) × (−∞,0)	(13,∞) × (−∞,0)	(4,∞) × (−∞,0)	(−∞,0) × (−∞,0)
(19,∞) × (1,3)	(13,∞) × (13,∞)	(5,∞) × (4,6)	(13,∞) × (−∞,0)
(2,14) × (9,11)	(5,∞) × (−∞,0)	(12,∞) × (12,∞)	(14,∞) × (14,∞)
(−∞,0) × (1,3)	(5,∞) × (9,∞)	(13,∞) × (13,∞)	(13,∞) × (13,∞)
(18,∞) × (18,∞)	(17,∞) × (−∞,0)	(1,∞) × (13,∞)	(13,∞) × (14,∞)
(11,∞) × (9,∞)	(8,∞) × (1,3)	(8,∞) × (9,∞)	(13,∞) × (13,∞)
(15,∞) × (1,5)	(16,∞) × (3,5)	(9,∞) × (11,∞)	(14,∞) × (6,7)
(5,∞) × (1,2)	(−∞,0) × (−∞,0)	(1,∞) × (1,∞)	(14,∞) × (14,∞)
(11,14) × (12,14)	(2,6) × (−∞,0)	(13,∞) × (13,∞)	(12,∞) × (9,11)
(13,∞) × (−∞,0)	(14,∞) × (13,∞)	(2,∞) × (4,∞)	(12,∞) × (13,∞)
(14,∞) × (16,∞)	(16,∞) × (17,∞)	(13,∞) × (1,3)	(7,∞) × (9,∞)
(−∞,0) × (1,2)	(13,∞) × (13,∞)	(3,∞) × (5,∞)	(14,∞) × (1,2)
(16,∞) × (17,∞)	(1,∞) × (1,∞)	(10,∞) × (10,∞)	(14,∞) × (14,∞)
(16,∞) × (1,2)	(14,∞) × (−∞,0)	(9,∞) × (9,∞)	(13,∞) × (3,5)
(16,∞) × (17,∞)	(12,∞) × (6,9)	(6,∞) × (6,∞)	(13,∞) × (13,∞)
(5,9) × (4,6)	(−∞,0) × (−∞,0)	(1,∞) × (10,∞)	(14,∞) × (12,∞)
(11,∞) × (13,∞)	(14,∞) × (12,14)	(13,∞) × (1,2)	(13,∞) × (13,∞)
(12,∞) × (13,∞)	(19,∞) × (19,∞)	(11,∞) × (11,∞)	(11,∞) × (11,∞)
(8,∞) × (8,∞)	(19,∞) × (1,3)	(11,∞) × (6,8)	(7,∞) × (7,∞)
(12,∞) × (8,9)	(16,20) × (−∞,0)	(2,∞) × (−∞,0)	(14,∞) × (−∞,0)
(4,∞) × (4,∞)	(1,∞) × (1,∞)	(10,∞) × (1,3)	(14,∞) × (14,∞)
(11,∞) × (2,4)	(10,∞) × (8,10)	(11,∞) × (−∞,0)	(4,∞) × (4,∞)
(11,∞) × (1,3)	(6,∞) × (−∞,0)	(1,∞) × (1,∞)	(13,∞) × (13,∞)
(1,∞) × (1,∞)	(19,∞) × (6,8)	(19,∞) × (19,∞)	(1,3) × (−∞,0)
(11,15) × (10,14)	(14,∞) × (6,9)	(1,3) × (8,11)	(11,∞) × (5,7)
(5,9) × (−∞,0)	(19,∞) × (5,7)	(11,15) × (1,4)	(11,∞) × (4,5)
(19,∞) × (−∞,0)	(19,∞) × (1,2)	(−∞,0) × (−∞,0)	(6,∞) × (6,∞)
(16,∞) × (1,2)	(17,∞) × (17,∞)	(17,∞) × (17,∞)	(10,∞) × (−∞,0)
(20,∞) × (1,3)	(8,∞) × (3,6)	(11,∞) × (2,6)	(1,∞) × (−∞,0)
(17,∞) × (17,∞)	(9,∞) × (15,∞)	(11,14) × (2,6)	(8,∞) × (1,2)
(1,4) × (−∞,0)	(19,∞) × (1,3)	(19,∞) × (−∞,0)	(3,∞) × (1,2)
(15,∞) × (−∞,0)	(18,∞) × (18,∞)	(1,∞) × (−∞,0)	(13,∞) × (13,∞)
(19,∞) × (19,∞)	(1,6) × (1,6)	(1,4) × (−∞,0)	(14,∞) × (1,2)

Capítulo 7

Conclusões e Trabalhos Futuros

7.1 Conclusões

A partir dos cálculos e gráficos obtidos no capítulo anterior, podemos afirmar que a estimação da função de distribuição bivariada é melhor feita pelo estimador Cópula, o qual primeiramente estima as funções de distribuição marginais que, após suavizadas, são unidas por um modelo de cópulas correto, cujo parâmetro é estimado por máxima verossimilhança.

Fazemos essa afirmação com base nas estimativas de vício, variância e EQM que obtivemos para os casos que simulamos. De fato, em todos os dois casos que simulamos, o referido estimador, que chamamos de “Cópula”, tem desempenho muito melhor. No caso 3 de simulação, em que há dependência negativa entre T_1 e T_2 , os estimadores quase se equivalem, mas há vantagem em favor do estimador Cópula na comparação das estimativas de vício.

O estimador cópula mostrou-se melhor nas simulações. Contudo, fizemos o ajuste com o modelo de cópulas que era conhecido. Essa pode ser uma limitação em aplicações com dados reais, como foi na aplicação que aqui fizemos. Realmente, a escolha de um modelo de cópula ruim pode levar a resultados de desempenho ruim. Além disso, nossas simulações indicaram que, mesmo usando o modelo verdadeiro de cópulas, o estimador de máxima verossimilhança do parâmetro da cópula pode não existir, por não haver pontos de máximo na função de verossimilhança conjunta com marginais suavizadas.

Por outro lado, o estimador ENPMV, que suaviza, por meio do núcleo estimador bivariado, o estimador não paramétrico de máxima verossimilhança da função de distribuição conjunta de T_1 e T_2 , H , pode ser também uma boa opção para estimarmos H . Afirmamos isso porque nesse estimador temos uma abordagem totalmente não paramétrica, diferentemente do estimador Cópula, em que temos que lançar mão de um modelo paramétrico.

A seleção das janelas (parâmetros de suavização) do núcleo estimador pelo método de Pan (2000b) surgiu como alternativa viável no casos univariado e bivariado. A viabilidade vem do fato de que não há como estimar critérios de desempenho do núcleo estimador em dados com censura intervalar (univariados ou bivariados).

As curvas de nível do valor esperado de ambos os estimadores sempre apontam a superioridade do estimador Cópula. De qualquer forma, ambos os estimadores melhoram em relação a vício, variância e EQM quando aumentamos o tamanho da amostra. Isso pode indicar consistência local dos estimadores.

7.2 Trabalhos Futuros

As simulações feitas apontaram a superioridade do estimador Cópula na estimação de H . Um dos passos para o cálculo do estimador cópula é a estimação das funções de distribuição marginais. Por isso, seria interessante estudar a conjunção de outros métodos de estimação das marginais, como o método de imputação que aqui expusemos, com o ajuste de modelos de cópulas.

É necessário também realizar simulações exaustivas com vários modelos de cópulas. Verificar o efeito da escolha de um modelo de cópula inadequado sobre a estimação de H parece ser essencial. Além disso, é importante estudar também a escolha de várias distribuições marginais.

Quanto ao estimador ENPMV, é prudente averiguar se uma boa escolha da função K_R do núcleo estimador influencia as estimativas. O modelo de Clayton, por exemplo, tem curvas de nível que formam linhas perpendiculares. Uma função núcleo elíptica pode não captar bem esse tipo de distribuição bivariada. Nossas simulações apontaram também que o estimador ENPMV teve desempenho melhor no caso em que T_1 e T_2 tinham dependência negativa. Pode-se verificar a veracidade disso.

Referências Bibliográficas

BAE, W. et al. Smoothing techniques for the bivariate kaplan-meier estimator. *Communications in Statistics - Theory and Methods*, 2005.

BARLOW, R. et al. *Statistical Inference Under Order Restrictions*. New York: Wiley, 1972.

BETENSKY, R.; FINKELSTEIN, D. A non-parametric maximum likelihood estimator for bivariate interval censored data. *Statistics in Medicine*, 1999.

BOGAERTS, K.; LESAFFRE, E. A new fast algorithm to find the regions of possible support for bivariate interval censored data. *Journal of Computational and Graphical Statistics*, 2004.

BRAUN, J.; DUCHESNE, T.; STAFFORD, J. E. Likelihood density estimation for interval censored data. *The Canadian Journal of Statistics*, 2005.

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

DUONG, T.; HAZELTON, M. Cross-validation bandwidth matrices for multivariate kernel density estimation. *Scandinavian Journal of Statistics*, 2005.

FARAWAY, J.; JHUN, M. Bootstrap choice of bandwidth for density estimation. *Journal of the American Statistical Association*, 1990.

GENTLEMAN, R.; GEYER, C. J. Maximum likelihood for interval censored data: Consistency and computation. *Biometrika*, 1994.

GENTLEMAN, R.; VANDAL, A. C. Nonparametric estimation of the bivariate cdf for arbitrarily censored data. *The Canadian Journal of Statistics*, 2002.

- GOGGINS, W.; FINKELSTEIN, D. A proportional hazards model for multivariate interval-censored failure time data. *Biometrics*, 2000.
- GROENEBOOM, P.; JONGBLOED, G.; WELLNER, J. The support reduction algorithm for computing nonparametric function estimates in mixture models. *Scandinavian Journal of Statistics*, 2008.
- GROENEBOOM, P.; WELLNER, J. *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Boston: Birkhäuser Verlag, 1992.
- GUO, J. *The Statistical Analysis of Interval-censored Failure Time Data*. New York: Springer, 2006.
- GÜRLER, U. Bivariate density estimation with randomly truncated data. *Journal of Multivariate Analysis*, 2000.
- HALL, P. Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, 1990.
- JONGBLOED, G. The iterative convex minorant algorithm for nonparametric estimation. *Journal of Computational and Graphical Statistics*, 1998.
- KOOPERBERG, C.; STONE, C. J. Logsplines density estimation for censored data. *Journal of Computational and Graphical Statistics*, 1992.
- KULASEKERA, K.; PADGETT, W. Bayes bandwidth selection in kernel density estimation with censored data. *Journal of Nonparametric Statistics*, 2006.
- MAATHUIS, M. Reduction algorithm for the npml for the distribution function of bivariate interval censored data. *Journal of Computational and Graphical Statistics*, 2005.
- MARRON, J. S.; PADGETT, W. J. Asymptotically optimal bandwidth selection for kernel density estimators from randomly right-censored samples. *The Annals of Statistics*, 1987.
- NELSEN, R. *An Introduction to Copulas*. 2. ed. New York: Springer, 2006.

- PAN, W. A multiple imputation approach to cox regression with interval censored data. *Biometrics*, 2000.
- PAN, W. Smooth estimation of the survival function for interval censored data. *Statistics in Medicine*, 2000b.
- PAZ, L. *Modelo de Cox para Dados com Censura Intervalar*. Dissertação (Mestrado) — Universidade Federal de Minas Gerais, Brasil, 2005.
- PETO, R. Experimental survival curves for interval-censored data. *Applied Statistics*, 1973.
- ROMEO, J.; TANAKA, N.; LIMA, A. Pedroso-de. Bivariate survival modeling: a bayesian approach based on copulas. *Lifetime Data Analysis*, 2006.
- SILVERMAN, B. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall, 1986.
- TURNBULL, B. W. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B, Methodological*, 1976.
- WAND, M. P.; JONES, M. C. *Kernel Smoothing*. Londres: Chapman and Hall, 1995.
- WEI, G. C. G.; TANNER, M. A. Applications of multiple imputation to the analysis of censored regression data. *Biometrics*, 1991.
- WELLS, M.; YEO, K. Density estimation with bivariate censored data. *Journal of the American Statistical Association*, 1996.
- WONG, G. Y.; YU, Q. Generalized mle of a joint distribution function with multivariate interval-censored data. *Journal of Multivariate Analysis*, 1999.