

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

Estimação dos Parâmetros da Mistura de
Densidades Estáveis Simétricas

por

Raucélio Coelho Cardoch Valdes

Orientador: Profa. Cira Etheowalda Guevara Otiniano

Junho de 2011

Agradecimentos

- A Deus, em tudo e por tudo, sou grato.
- A minha amada esposa, Flávia, por sonhar comigo e por seu apoio incondicional.
- A minha filha Eliza Maria por ser o meu mais belo sonho.
- Aos meus pais Raul e Cecília por torcerem pela minha felicidade.
- Ao meus irmãos por , cada um a seu modo, serem exemplos para mim.
- A Professora Cira Etheowalda Guevara Otiniano por me aceitar como orientando e compartilhar comigo a sua competência acadêmica.
- A Professora Cibele e Professor Geraldo por acreditarem na minha capacidade para enfrentar o desafio do mestrado ao me concederem as cartas de recomendação.
- Aos Professores Rathie e Peter pelos conhecimentos transmitidos.

Sumário

Agradecimentos

Lista de Figuras 4

Lista de Tabelas 5

Resumo 6

Abstract 7

Introdução 8

1 Conceitos e Notações Preliminares 12

1.1 Introdução 12

1.2 Variáveis Aleatórias Estáveis 12

1.2.1 Propriedades 17

1.3 Modelos de Misturas Finitas 18

1.3.1 Definição 18

1.4 Função Característica 20

2 Estimação via Função Característica 22

2.1 Introdução 22

2.2 Função Distância Objetiva 23

2.3 Propriedades Assintóticas 24

2.4 Estimador EFC Para Mistura de Normais 26

2.5 Estimador EFC Para Mistura de Cauchy 28

2.6 Estimador EFC Para Mistura de Estáveis Simétricas 29

2.7	Aplicação do Estimador EFC em dados reais	31
2.7.1	Aplicação do Estimador EFC em misturas de Cauchy	32
3	Estimação pelo Algoritmo EM	36
3.1	Introdução	36
3.2	EM Para Mistura Finita de Variáveis Aleatórias	36
3.3	EM Para Mistura Finita de Densidades Estáveis	39
3.4	Aplicações do Algoritmo EM	40
3.4.1	Aplicação do EM a Dados Reais	40
3.4.2	Aplicação do EM a Dados Simulados	42
4	Considerações Numéricas	44
4.1	Introdução	44
4.2	Sobre os Programas Utilizados	44
4.3	Densidade de Uma Variável Aleatória Estável	45
4.3.1	Algoritmo I - Fórmula de Inversão	45
4.3.2	Algoritmo II - Representação pela Função H	46
4.3.3	Comentários	48
4.3.4	Simulação de uma Mistura Finita	49
4.4	Influência de b na Matriz $\Lambda^{-1}\Sigma\Lambda^{-1}$ na Estimação EFC	51
5	Conclusões	54
	Referências Bibliográficas	56
A	Tabelas	59
A.1	Tabelas citadas no Capítulo 2	60
A.2	Tabelas citadas no Capítulo 3	62
B	Cálculo de D_n para componentes $N(\mu, \sigma^2)$	63
C	Cálculo de D_n para componentes $Cauchy(\mu, \sigma)$	69
D	Cálculo de D_n para componentes Estáveis Simétricas	74

E	Programas para Cálculo da Densidade de uma Distribuição Estável	80
E.1	Programa R - Algoritmo I	80
E.2	Programa R - Algoritmo II	81
F	Programa R para Gerar Valores de uma Distribuição Estável	88

Lista de Figuras

1.1	Efeito dos parâmetros de uma v.a. estável sobre a sua função de densidade.	14
1.2	Efeito da parametrização da função característica sobre a função de densidade de uma v.a. estável.	16
2.1	Densidades propostas pelos estimadores EFC para o consumo de petróleo com componentes Normais.	32
2.2	Densidades propostas pelos estimadores EFC para o consumo de petróleo com componentes Cauchy.	35
3.1	Densidades propostas pelo algoritmo EM para o consumo de petróleo com componentes normais e componentes estáveis.	43
4.1	Log de densidades de v.a's estáveis obtidas pela fórmula de inversão e expansão por série da função H.	49
4.2	Efeito de b sobre variância dos estimadores EFC.	53
A.1	Densidades das misturas de Cauchy para os caso C1 e C2	61

Lista de Tabelas

2.1	Experimento de simulação para os estimadores EFC para componentes Cauchy.	33
2.2	Média das estimativas EFC para componentes Cauchy.	34
2.3	Estimativas EFC para o consumo de petróleo com componentes Cauchy.	34
3.1	Estimativas EM para o consumo de petróleo com componentes normais e componentes estáveis.	42
3.2	Experimento de estimação EM para misturas finitas de variáveis aleatórias estáveis.	42
3.3	Estimativas EM para misturas finitas de variáveis aleatórias estáveis.	43
A.1	EQM das estimativas EFC para componentes Cauchy.	60
A.2	Viés das estimativas EFC para componentes Cauchy.	60
A.3	Valor do parâmetro, média , EQM e viés das estimativas EFC para os casos C1 e C2.	61
A.4	EQM das estimativas EM para misturas finitas de variáveis aleatórias estáveis.	62
A.5	Viés das estimativas EM para misturas finitas de variáveis aleatórias estáveis.	62

Resumo

Misturas de finitas de densidades de variáveis aleatórias α -estáveis é uma alternativa de modelagem para dados multimodais com origem em subpopulações que apresentam assimetria e caudas pesadas. Na família das variáveis aleatórias α -estáveis os únicos membros com densidade em forma fechada simples são as Normais, Cauchy e Lévy $_{\frac{1}{2}}$. Entretanto, sua função característica é bem conhecida e envolve quatro parâmetros.

Neste trabalho, foram aplicados dois métodos de estimação dos parâmetros de uma mistura finita de densidades α -estáveis. O primeiro método é o Estimador via Função Característica (EFC) que minimiza a distância entre a função característica teórica e a função característica empírica, quando as componentes são Normais, Cauchy e α -estáveis simétricas. O segundo método é uma adaptação do algoritmo EM (Expectation-Maximization), devido ao cálculo numérico das densidades α -estáveis simétricas. Ilustrações numéricas dos resultados com dados reais e simulados também são apresentados.

Palavras Chave: *Misturas finitas; Densidades α -estáveis; EFC; Algoritmo EM.*

Abstract

Finite mixtures of α -stable random variables' densities is an alternative model for multimodal data stemming from subpopulations with skewed and heavy-tailed distributions. Regarding the family of α -stable random variables, the only members with simple closed-form density are the Normal, Cauchy and Lévy $_{\frac{1}{2}}$ distributions. However, its characteristic function is well-known, and encompasses four parameters.

In this study, two methods of parameter estimation of a finite mixture of α -stable densities have been applied. The first method is the Characteristic Function Estimator (CFE), which minimizes the distance between the theoretical characteristic function and the empirical characteristic function whenever there are Normal, Cauchy and symmetric α -stable components. The second method consists of an adaptation of the Expectation-Maximization (EM) algorithm, due to numerical calculations of symmetric α -stable densities. Numerical illustrations of results, consisting of real and simulated data, are presented in this study as well.

Keywords: Finite mixtures; α -stable densities; CFE; EM Algorithm.

Introdução

O objetivo deste trabalho é estimar os parâmetros de uma mistura finita de densidades, cujas componentes são variáveis aleatórias estáveis. Essa mistura é uma alternativa para modelar dados multimodais com origem em subpopulações distintas, nas quais a presença de valores extremos não é rara, assim, podem apresentar assimetria e caudas pesadas.

Uma variável aleatória X tem distribuição estável, se a soma de cópias independentes dessa variável ainda mantém a distribuição de X , exceto por locação e escala, isto é,

$$X_1 + \dots + X_n = n^{\frac{1}{\alpha}}X + D_n,$$

onde $X_1, \dots, X_n$ são cópias independentes de X . Essas distribuições foram apresentadas no trabalho *Calcul des Probabilités*, escrito por Paul Lévy em 1925, o qual pesquisou sobre a soma de variáveis aleatórias independentes.

Até o momento, não há uma fórmula fechada simples para a densidade de uma variável aleatória estável, exceto para os casos particulares de Normal, Cauchy e Lévy_{1/2}. Entretanto, há meios numéricos para obter valores da função de densidade e outras quantidades de interesse relacionadas a essas variáveis aleatórias.

A função característica de uma variável aleatória estável é bem conhecida e envolve o vetor de quatro parâmetros $\theta = (\alpha, \sigma, \mu, \beta)$ na relação

$$\varphi_X(t; \theta) = \begin{cases} \exp \{ i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \operatorname{sign}(t) \tan(\alpha \frac{\pi}{2})] \} & , \text{ se } \alpha \neq 1 \\ \exp \{ i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \frac{\pi}{2} \operatorname{sign}(t) \log(t)] \} & , \text{ se } \alpha = 1 \end{cases},$$

com $\operatorname{sign}(t) = \frac{t}{|t|}$, $\forall t \in \mathbb{R}$, $0 < \alpha \leq 2$ é o índice de estabilidade; $-1 \leq \beta \leq 1$ é o parâmetro de simetria; $\sigma > 0$ é o parâmetro de escala e $\mu \in \mathbb{R}$ é o parâmetro de locação. Por ser bem definida, a função característica de uma variável aleatória estável

é usada em métodos de estimação dos seus parâmetros desde, aproximadamente, 1965. Dentre os métodos propostos, os mais conhecidos são apresentados em Famma e Roll (1968), DuMouchel (1971), Press (1972), Paulson *et al.* (1975), Hill (1975), Koutrouvelis (1980), Dorea *et al.* (2007).

Ainda sobre distribuições estáveis, sempre que as quantidades observadas puderem ser aproximadas por somas normalizadas de pequenos termos e as densidades empíricas dos dados exibirem caudas pesadas e assimetria, o Teorema do Limite Central Generalizado junto com a evidência empírica forte indicam que um modelo mais realístico pode ser obtido com essa família de distribuições.

Os modelos de mistura são utilizados para modelar populações, cujos dados têm origem em K subpopulações distintas. Cada subpopulação, ou componente da mistura, é bem descrita por uma família paramétrica. As densidades associadas a cada subpopulação são ponderadas por suas frequências relativas, p_k , na população. Assim, a densidade da população, ou da mistura, é a combinação convexa

$$h(x; \Theta) = \sum_{j=1}^K p_j f_j(x; \theta_j),$$

onde, $p_1, \dots, p_K$ são os pesos da mistura, $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)'$, θ_j são os parâmetros da j -ésima componente e $f_1, \dots, f_K$, são as componentes da mistura.

Misturas finitas de densidades têm aplicações práticas em diversas áreas. Os livros Titterington *et al.* (1985), MacLachlan e Peel (2000) descrevem modelos, propriedades e aplicações em áreas como: pesca, economia, medicina, psicologia, paleontologia, botânica, agricultura, zoologia e confiabilidade, entre outras. Como aplicações indiretas incluem: outliers, mistura de normais para teste de robustez, análise de clusters, modelos de estrutura latentes, inferência bayesiana, estimação não-paramétrica de densidades e aproximação de modelos de não mistura com modelos da mistura.

A mistura de distribuições normais é utilizada em aplicações, nas quais os dados são multimodais e provenientes de subpopulações normais, por exemplo, processamento de imagens e de sinais para radares. Entretanto, a mistura de normais não leva a uma modelagem satisfatória quando as densidades das componentes apresentam assimetria ou caudas pesadas. Nessa situação, a utilização de componentes

estáveis é uma boa alternativa para a modelagem de dados não normais. Um modelo multivariado de sobrevivência com mistura de distribuições estáveis positivas foi apresentado por Ravishanker e Dey (2000). Salas-Gonzalez *et al.* (2009), Xu e Knight (2011) aplicaram misturas de distribuições estáveis para modelar dados de finanças.

Para estimar os parâmetros de misturas finitas os métodos de estimação mais utilizados são: Máxima Verossimilhança, Algoritmo EM e métodos Bayesianos. O algoritmo EM se destaca de outros métodos de estimação pela sua eficiência em modelos de mistura, porém, o algoritmo pode convergir lentamente devido à falta de uma forma fechada dos estimadores ou pelo do valor inicial para o procedimento. Swami e How (2000) e Hatzinakos (1988) mostram que o EM é eficiente e converge rapidamente para mistura de componentes Normais, Cauchy e Lévy $\frac{1}{2}$. Um outro método numérico alternativo de estimação dos parâmetros de uma mistura de componentes normais é minimizar a distância entre a função característica teórica e a função característica empírica.

Paulson *et al.* (1975) propuseram esse método, que consiste em determinar o estimador $\tilde{\Theta}$ que minimiza a função distância objetiva

$$D_n(\Theta, x) = \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \theta)|^2 dW(t),$$

onde $\varphi_X(t; \theta) = E(e^{itX})$ é a função característica teórica, $\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j}$ é a função característica empírica e $W(t) = \exp(-bt^2)$. Heathcote (1977) obteve propriedades de consistência e normalidade assintóticas para $\tilde{\Theta}$ nos casos univariado e multivariado e, também, sugeriu que a escolha de $W(t)$ pode influenciar na eficiência do estimador e, como exemplo, mostrou que os estimadores têm baixa eficiência quando $W(t) = \exp(-t^2)$. Com base nesse trabalho, Xu e Knight (2011) propõem um método iterativo para aumentar a eficiência dos estimadores de uma mistura de componentes normais. Esse método tem como base o cálculo da matriz de variância-covariância dos estimadores obtidos pela minimização da distância objetiva da mistura.

Neste trabalho, são tratados dois métodos de estimação dos parâmetros da mistura $h(x, \Theta)$, cujas componentes são estáveis simétricas. O primeiro método foi baseado no trabalho de Xu e Knight (2011) e o segundo método é o EM.

No Capítulo 1, são apresentadas as notações e os conceitos preliminares utilizados neste trabalho, como: variáveis aleatórias estáveis, modelos de misturas fini-

tas e de função característica. No Capítulo 2, defini-se o estimador EFC, $\tilde{\Theta}$, como aquele que minimiza a distância entre a função característica e, a sua contrapartida amostral, a função característica empírica; citam-se as propriedades assintóticas de $\tilde{\Theta}$, demonstradas por Heathcote (1977); revisa-se $\tilde{\Theta}$ para misturas de componentes normais proposto por Xu e Knight (2011); é deduzida a forma fechada de $\tilde{\Theta}$ para misturas de componentes Cauchy; apresentam-se os resultados alcançados a fim de se determinar as formas fechadas de $\tilde{\Theta}$ para misturas de variáveis estáveis simétricas ; e, por fim, os estimadores $\tilde{\Theta}$, para misturas de Normais e para misturas de Cauchy, são aplicados a dados reais e simulados. No Capítulo 3, o algoritmo EM para misturas finitas é descrito e adaptado para estimar os parâmetros de uma mistura de variáveis aleatórias estáveis, assim como, são apresentadas aplicações em dados reais e simulados. No Capítulo 4, são feitas considerações sobre os programas utilizados, cálculos de densidades de variáveis aleatórias estáveis, geração de valores aleatórios de uma distribuição estável e sobre a matriz de covariância do estimador $\tilde{\Theta}$. No Capítulo 5, são apresentas as conclusões deste trabalho e propostas para trabalhos futuros.

Capítulo 1

Conceitos e Notações Preliminares

1.1 Introdução

Neste capítulo são apresentados os conceitos e notações utilizadas nesta dissertação. Na seção 1.2, apresentam-se definições e propriedades de variáveis aleatórias estáveis. Na seção 1.3, descreve-se os modelos de mistura finita. Na seção 1.4, defini-se a função característica de uma mistura finita, em particular, aquelas com componentes estáveis.

1.2 Variáveis Aleatórias Estáveis

As variáveis aleatórias α -estáveis, também conhecidas por estáveis, foram introduzidas pelo matemático Paul Lévy por meio do seu trabalho *Calcul des Probabilités* no ano de 1925, no qual estudou o comportamento da soma de variáveis aleatórias independentes. Em palavras, pode-se dizer que uma variável aleatória é estável se for invariante por somas exceto por locação e escala. Uma variável aleatória normal é um caso particular dessa classe de variáveis aleatórias, pois a soma de normais resulta em uma normal com mudanças nos seus parâmetros de locação e escala.

Definição 1.2.1. *Uma variável aleatória Z é estável, se existem as constantes $A > 0$, $B > 0$, $C > 0$ e $D \in \mathbb{R}$, tais que*

$$AZ_1 + BZ_2 = CZ + D,$$

onde Z_1 e Z_2 são cópias independentes de Z .

Definição 1.2.2. *Sejam $Z_1, \dots, Z_n$ cópias independentes de Z . Então Z é chamada α -estável se existirem constantes $D_n \in \mathbb{R}$ tais que*

$$S_n = Z_1 + \dots + Z_n = n^{\frac{1}{\alpha}} Z + D_n,$$

onde $\alpha \in (0, 2]$ é chamado de índice de estabilidade ou expoente característico.

Não existe forma fechada simples da distribuição de uma variável aleatória estável X qualquer, porém, sua função característica é conhecida e bem definida por Lévy (1925).

A definição seguinte traz a representação da função característica de uma variável aleatória estável, conforme Zolotarev (1986) e Samorodnitsky e Taqqu (1994), chamada de forma canônica da função característica.

Definição 1.2.3. *Uma variável aleatória Z com função de distribuição de probabilidade*

$$F(x; \alpha, \sigma, \beta, \mu) = S_\alpha(\sigma, \beta, \mu)(x)$$

e função de densidade

$$f(x; \theta) = F'(x; \theta),$$

com $\theta = (\alpha, \sigma, \beta, \mu)$, é considerada estável, quando sua função característica for

$$\varphi_Z(t; \theta) = \begin{cases} \exp \{ i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \text{sign}(t) \tan(\alpha \frac{\pi}{2})] \} & , \text{ se } \alpha \neq 1, \\ \exp \{ i\mu t - \sigma^\alpha |t|^\alpha [1 - i\beta \frac{\pi}{2} \text{sign}(t) \log(t)] \} & , \text{ se } \alpha = 1, \end{cases} \quad (1.1)$$

onde $\text{sign}(t) = \frac{t}{|t|}$, $\forall t \in \mathbb{R}$; $0 < \alpha \leq 2$ é chamado de índice de estabilidade; $-1 \leq \beta \leq 1$ é o parâmetro de simetria; $\sigma > 0$ é o parâmetro de escala e $\mu \in \mathbb{R}$ é o parâmetro de localização.

Uma variável aleatória Z é α -estável simétrica se

$$Z \sim S_\alpha(\sigma, 0, \mu), \quad (1.2)$$

cuja função característica é

$$\varphi_Z(t; \theta) = \exp \{ i\mu t - \sigma^\alpha |t|^\alpha \}. \quad (1.3)$$

A Figura 1.1 apresenta os gráficos das funções de densidade da variável aleatória $X \sim S_\alpha^0(\sigma, \beta, \mu_0)$, definida em (1.1), onde apenas um dos seus parâmetros varia

e os demais permanecem fixos. O gráfico (a) mostra o efeito da variação μ_0 ; o (b), o efeito de σ na escala; o (c), o efeito de β na simetria, e o (d) mostra o efeito de α nas caudas. Tais gráficos foram obtidos pelo Algoritmo I, descrito no Capítulo 4.

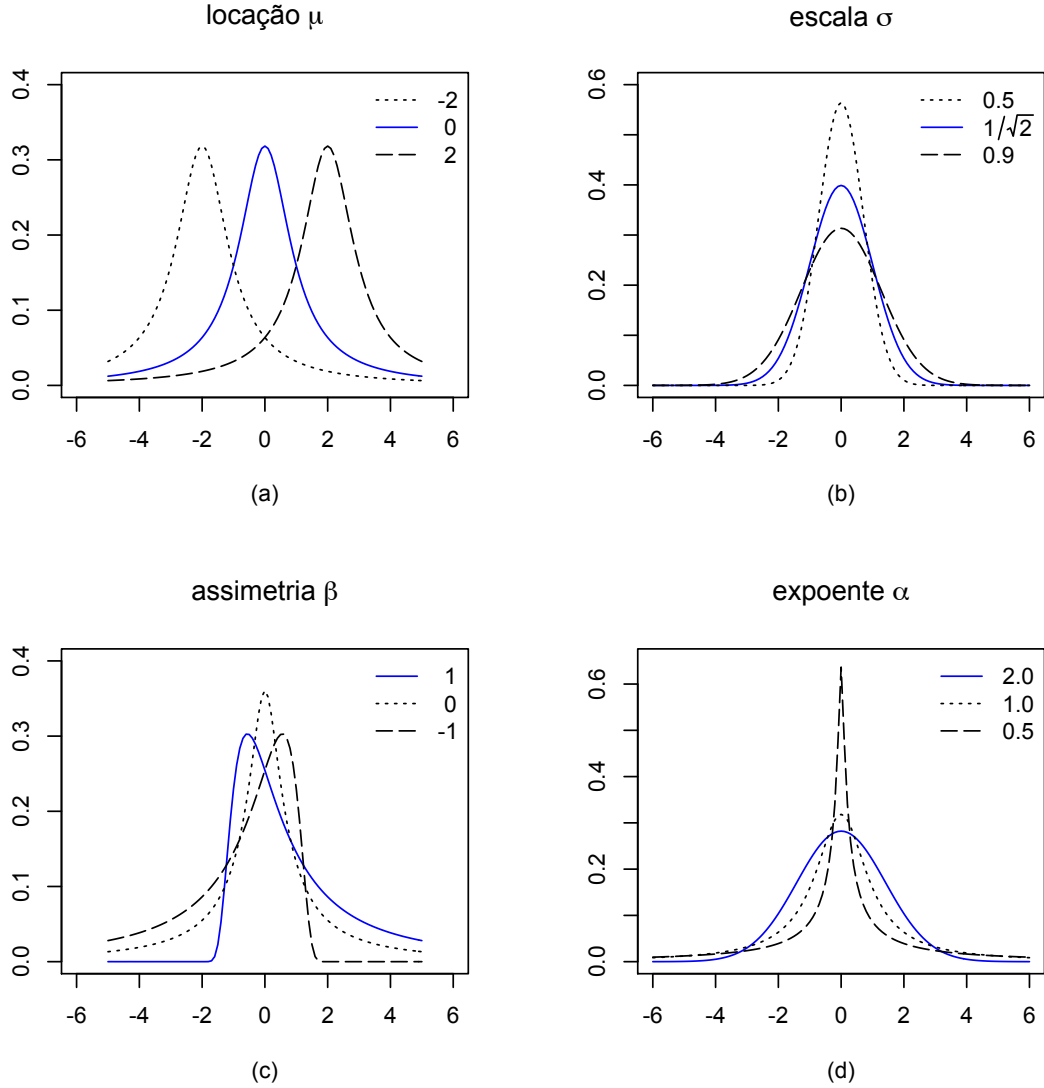


Figura 1.1: Efeito dos parâmetros μ_0 , σ , β e α no comportamento da função de densidade de $X \sim S_\alpha^0(\sigma, \beta, \mu_0)$. Em (a), $\mu_0 \in \{-2, 0, 2\}$, $\alpha = 1, \beta = 0, \sigma = 1$; em (b), $\sigma \in \{0.5, \frac{1}{\sqrt{2}}, 0.9\}$, $\alpha = 2, \beta = 0$ e $\mu_0 = 0$; em (c), $\beta \in \{-1, 0, 1\}$, $\alpha = 0.8, \sigma = 1$ e $\mu_0 = 0$; em (d), $\alpha \in \{0.5, 1.5, 2\}$, $\beta = 0, \sigma = 1$ e $\mu_0 = 0$.

A Definição 1.2.3 para variável aleatória estável permite estabelecer as relações abaixo com outras variáveis aleatórias:

- $S_{\frac{1}{2}}(\sigma, \beta, \mu) \sim Levy_{\frac{1}{2}}(\mu, \sigma)$,

- $S_1(\sigma, 0, \mu) \sim \text{Cauchy}(\mu, \sigma)$ e
- $S_2\left(\frac{\sigma}{\sqrt{2}}, 0, \mu\right) \sim \text{Normal}(\mu, \sigma)$.

Da definição 1.2.3 tem-se que a distribuição de Z é a distribuição de somas normalizadas de cópias independentes de Z ,

$$\frac{S_n - b_n}{n^{\frac{1}{\alpha}}} \stackrel{d}{=} Z.$$

E sugere que uma distribuição estável pode ser obtida como a distribuição limite de somas normalizadas de variáveis aleatórias i.i.d.. Formalmente, isto corresponde ao Teorema Limite Central Generalizado. Sejam $X_1, X_2, \dots$ variáveis aleatórias i.i.d. com variância não necessariamente finita. Então,

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - b_n}{n^{\frac{1}{\alpha}} \gamma(n)} \leq Z\right) = S_\alpha(\sigma, \beta, \mu)(z), \quad (1.4)$$

onde, $S_n = X_1 + \dots + X_n$, γ é uma função tal que

$$\lim_{t \rightarrow \infty} \frac{\gamma(tx)}{\gamma(t)} = 1 \quad (1.5)$$

e

$$b_n = \begin{cases} EX & , \text{ se } 1 < \alpha \leq 2 \\ 0 & , \text{ se } 0 < \alpha < 1 \\ 0 & , \text{ se } \alpha = 1 \text{ e } X \text{ simétrica.} \end{cases} \quad (1.6)$$

Nolan (2011) propôs uma parametrização de (1.1), com o objetivo de contornar as dificuldades numéricas encontradas no cálculo do termo $\beta\sigma \tan(\alpha\frac{\pi}{2})$ quando $\alpha \rightarrow 1$ e $\beta \neq 0$. Essa parametrização facilita a obtenção dos valores de interesse relacionados a uma variável aleatória estável como função de densidade ou função distribuição.

Definição 1.2.4. *Uma variável aleatória Z com função de distribuição de probabilidade*

$$F(x; \alpha, \sigma, \beta, \mu) = S_\alpha^0(\sigma, \beta, \mu)(x)$$

e função de densidade

$$f(x; \theta) = F'(x; \theta),$$

com $\theta = (\alpha, \sigma, \beta, \mu)$, é considerada estável, quando sua função característica for

$$\varphi_Z(t; \theta) = \begin{cases} \exp\{i\mu_0 t - \sigma^\alpha |t|^\alpha [1 - i\beta \text{sign}(t) \tan(\alpha\frac{\pi}{2}) (|\sigma t|^{1-\alpha} - 1)]\}, & \text{se } \alpha \neq 1, \\ \exp\{i\mu_0 t - \sigma^\alpha |t|^\alpha [1 - i\beta \frac{\pi}{2} \text{sign}(t) \log(t)]\} & , \text{ se } \alpha = 1, \end{cases} \quad (1.7)$$

onde $\text{sign}(t) = \frac{t}{|t|}$, $\forall t \in \mathbb{R}$; $0 < \alpha \leq 2$ é chamado de índice de estabilidade; $-1 \leq \beta \leq 1$ é o parâmetro de simetria; $\sigma > 0$ é o parâmetro de escala e $\mu \in \mathbb{R}$ é o parâmetro de locação.

Na Definição 1.2.3 e na Definição 1.2.4 os parâmetros $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\sigma > 0$ têm significados iguais; mas, os parâmetros $\mu \in \mathbb{R}$ e $\mu_0 \in \mathbb{R}$ se relacionam por meio de

$$\mu = \begin{cases} \mu_0 - \beta\sigma \tan(\alpha\frac{\pi}{2}) & , \text{ se } \alpha \neq 1, \\ \mu_0 - \beta\frac{2}{\pi}\sigma \ln(\sigma) & , \text{ se } \alpha = 1. \end{cases} \quad (1.8)$$

Na figura 1.2, O gráfico (a) corresponde a função de densidade correspondente a Definição 1.2.3 e o (b) corresponde à Definição 1.2.2.

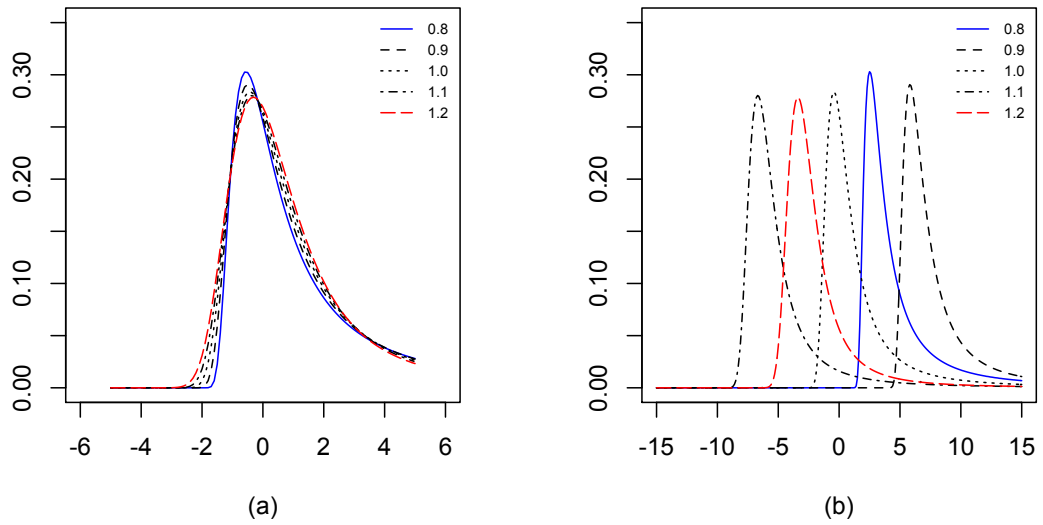


Figura 1.2: A função de densidade de uma v.a. estável X com $\alpha \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$, $\sigma = 1$, $\beta = 1$ e $\mu = 0$. Em (a), segundo a Definição 1.2.3, $X \sim S_{\alpha}^0(1, 1, 0)$, e, em (b), segundo a Definição 1.2.2, $X \sim S_{\alpha}(1, 1, 0)$.

Os trabalhos Zolotarev (1986) e Samorodnitsky e Taqqu (1994) apresentam, em detalhes, as seguintes propriedades.

1.2.1 Propriedades

Propriedade 1.2.1. *Seja $X \sim S_\alpha(\sigma, \beta, \mu)$ ou $X \sim S_\alpha^0(\sigma, \beta, \mu)$, em (1.1) ou em (1.7), então,*

$$S_\alpha(1, \beta, 0) = -S_\alpha(1, -\beta, 0) \quad (1.9)$$

Desse modo, a densidade e a função de distribuição satisfazem, respectivamente, as relações

$$\begin{aligned} f(x|\alpha, \beta, \sigma = 1, \mu = 0) &= f(-x|\alpha, -\beta, \sigma = 1, \mu = 0), \\ F(x|\alpha, \beta, \sigma = 1, \mu = 0) &= 1 - F(-x|\alpha, -\beta, \sigma = 1, \mu = 0). \end{aligned}$$

O lado esquerdo de (1.9) representa uma distribuição estável padrão. Desta forma, a utilidade dessa propriedade é a possibilidade de implementar algoritmos que calculam valores de densidades ou de funções de distribuição apenas no intervalo $[0, \infty)$ e usá-los para determinar essas quantidades quando $x \in (-\infty; 0)$.

Propriedade 1.2.2. *Sejam $\mu \in \mathbb{R}$ e $0 < \sigma < \infty$ a locação e a escala, respectivamente. A relação entre uma densidade de uma variável aleatória estável qualquer e a estável padrão, $\mu = 0$ e $\sigma = 1$, é dada por*

$$f(x; \alpha, \beta, \sigma, \mu) = \frac{1}{\sigma} f\left(\frac{x - \mu}{\sigma}; \alpha, \beta, 1, 0\right).$$

Nolan (1999), apresenta um algoritmo para obter a densidade de uma variável estável padrão e usa a Propriedade 1.2.2 para obter a densidade de uma variável aleatória estável mais geral.

Propriedade 1.2.3. *Seja $X \sim S_\alpha^0(\sigma, \beta, \mu)$, com $0 < \alpha < 2$ e $-1 < \beta \leq 1$, então, para $x \rightarrow \infty$,*

$$\lim_{x \rightarrow \infty} x^\alpha P(X > x) = \sigma^\alpha c_\alpha (1 + \beta)x, \quad (1.10)$$

$$\lim_{x \rightarrow \infty} x^\alpha P(X < -x) = \sigma^\alpha c_\alpha (1 - \beta)x. \quad (1.11)$$

$$\text{com } c_\alpha = \frac{\sin\left(\frac{\pi\alpha}{2}\right)\Gamma(\alpha)}{2}.$$

As expressões (1.10) e (1.11) podem ser escritas em termos de densidades e são:

$$\lim_{x \rightarrow \infty} x^{\alpha+1} f(x|\alpha, \beta, \sigma, \mu) = \alpha \sigma^\alpha c_\alpha (1 + \beta), \quad (1.12)$$

$$\lim_{x \rightarrow \infty} x^{\alpha+1} f(-x|\alpha, \beta, \sigma, \mu) = \alpha \sigma^\alpha c_\alpha (1 - \beta). \quad (1.13)$$

As expressões (1.10) - (1.13) permitem aproximar as caudas da densidade de uma variável aleatória estável para valores suficientemente grandes, já que nessa situação podem ocorrer dificuldades numéricas. As expressões (1.12) - (1.13) indicam que o decaimento caudal da densidade de uma variável aleatória estável é uma função potência de ordem α , e quanto menor o valor de α mais pesada é a cauda.

Propriedade 1.2.4. *Seja X uma variável aleatória estável, então para $0 < p < \alpha$ o momento absoluto de ordem p existe, isto é, $E|X|^p < \infty$. Assim, para $1 < \alpha \leq 2$, $E|X| < \infty$ e para $0 < \alpha \leq 2$, $E|X|^2 = EX^2 = \infty$.*

Essa propriedade indica que $Var(X) = \infty$ quando $\alpha < 2$ e $EX = \infty$ quando $\alpha < 1$. Assim, esses momentos estão relacionados como os pesos das caudas. Cabe notar que $E|X|^2 = EX^2 < \infty$ para $\alpha = 2$.

O objetivo desta dissertação é estimar os parâmetros de misturas finitas de densidades estáveis, em particular as simétricas como em (1.2). Para tal, na próxima seção é definida uma mistura finita de densidades e sua função característica.

1.3 Modelos de Misturas Finitas

1.3.1 Definição

Suponha que $F(x|\theta)$ representa a função de distribuição de uma variável aleatória X dado θ e $G(\theta)$ representa a função distribuição da variável aleatória θ . A distribuição marginal de (X, θ) , denotada por

$$H(x) = \int_{-\infty}^{+\infty} F(x|\theta) dG(\theta), \quad (1.14)$$

é a mistura de F e G , onde $F(x|\theta)$ é conhecida como núcleo, e $G(\theta)$ é a distribuição de mistura.

Se toda a massa correspondente a $G(\Theta)$ está concentrada em um conjunto finito de pontos $\theta_1, \theta_2, \dots, \theta_K$, então, (1.14) se torna uma mistura finita de K componentes, cuja distribuição é dada por

$$H(x) = \sum_{j=1}^K F_j(x|\theta_j) G(\theta_j). \quad (1.15)$$

Para simplificar a notação, $p_j \equiv G(\theta_j)$ e $F_j(x) = F(x|\theta_j)$, então, (1.15) pode ser escrito como

$$H(x) = \sum_{j=1}^K p_j F_j(x) \quad (1.16)$$

onde $p_j \geq 0$, para $j = 1, 2, \dots, K$ e

$$\sum_{j=1}^K p_j = 1.$$

Em (1.16), p_j é conhecido com a j -ésima componente de proporção, $F_j(x)$ a j -ésima componente de mistura e K o número de componentes.

Definição 1.3.1. *Uma variável aleatória X é dita ser uma mistura finita de K componentes, caso sua função de distribuição for expressa por*

$$H(x; \Theta) = \sum_{j=1}^K p_j F_j(x; \theta_j), \quad (1.17)$$

onde $j = 1, \dots, K$, $F_j(x)$ é uma função de distribuição qualquer e $p_j \geq 0$, com

$$\sum_{j=1}^K p_j = 1,$$

é a probabilidade de ocorrência dos pontos de massa θ_j de uma variável aleatória $G(\Theta)$, onde $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e θ_j são os parâmetros da j -ésima componente F_j .

Se, em (1.14), $G(\theta)$ é absolutamente contínua, a densidade $g(\theta)$ existe de tal forma que $g(\theta) \equiv G'(\theta)$ e se $h(x)$ e $f(x|\theta)$ são as densidades correspondentes às distribuições $H(x)$ e $F(x|\theta)$, a partir de (1.14), tem-se

$$h(x) = \int_{-\infty}^{+\infty} f(x|\theta) g(\theta) d\theta. \quad (1.18)$$

Analogamente, a expressão (1.18) pode ser escrita como

$$h(x) = \sum_{j=1}^K p_j f_j(x) \quad (1.19)$$

onde a densidade $f_j(x)$ é a j -ésima componente de mistura correspondente à distribuição $F_j(x)$. Assim, a função de densidade de uma mistura $H(x)$, em (1.17), pode ser escrita como

$$h(x; \Theta) = \sum_{j=1}^K p_j f_j(x; \theta_j), \quad (1.20)$$

onde $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e θ_j são os parâmetros da j -ésima componente f_j , para $j = 1, \dots, K$.

1.4 Função Característica

A função característica é uma ferramenta estatística na estimação de parâmetros quando a função de densidade não tem forma fechada ou é de difícil manipulação.

Definição 1.4.1. *Seja X uma variável aleatória. A função característica de X é a função $\varphi_X : \mathbb{R} \rightarrow \mathbb{C}$ definida por*

$$\varphi_X(t; \theta) = E(e^{itX}), \quad (1.21)$$

com $t \in \mathbb{R}$, $i = \sqrt{-1}$ e θ sendo o vetor dos parâmetros de X .

Em termos da integral de Reimann-Stieltjes (1.21) é

$$\varphi_X(t; \theta) = \int_{-\infty}^{+\infty} e^{itX} dF_X(x; \theta), \quad (1.22)$$

onde $F_X(x)$ é uma função de distribuição de X . Quando X é absolutamente contínua com densidade $f(x; \theta)$, a função característica dessa variável é dada por

$$\varphi_X(t; \theta) = \int_{-\infty}^{+\infty} e^{itX} f(x; \theta) dx$$

e por

$$\varphi_X(t; \theta) = \sum_{j=1}^n e^{itX} p(x_j; \theta)$$

quando X for discreta.

A contrapartida amostral da função característica é a função característica empírica.

Definição 1.4.2. *Seja $X_1, \dots, X_n$ uma amostra aleatória independente e identicamente distribuída (i.i.d.) de uma variável aleatória X . A função característica empírica de X é definida por*

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j}. \quad (1.23)$$

As funções características (1.22) e (1.23) são as bases do método de estimação estudado no Capítulo 2, que em síntese, minimiza a distância entre essas funções.

No contexto de mistura de densidades, a função característica da mistura finita (1.20) é

$$\begin{aligned} \varphi_H(t; \theta) &= E(e^{itX}) \\ &= \int_{-\infty}^{+\infty} e^{itX} h(x; \theta) dx \\ &= \int_{-\infty}^{+\infty} e^{itX} \sum_{j=1}^K p_j f_j(x) dx \\ &= \sum_{j=1}^K p_j \int_{-\infty}^{+\infty} e^{itX} f_j(x) dx \\ &= \sum_{j=1}^K p_j \varphi_{X_j}(t; \theta). \end{aligned} \quad (1.24)$$

A seguinte proposição apresenta a função característica (1.24), quando $f_j(x)$ são densidades de variáveis aleatórias estáveis.

Proposição 1.4.1. *Dada uma mistura finita H de $S_{\alpha_i}(\mu_i, \beta_i, \sigma_i)$ com $i = 1, \dots, K$. A função característica de H é*

$$\begin{aligned} \varphi_H(t; \theta) &= \sum_{\substack{i=1 \\ \alpha_i \neq 1}}^K p_i \exp \left\{ i\mu_i t - \sigma_i^{\alpha_i} |t|^{\alpha_i} [1 - i\beta_i \operatorname{sign}(t) \tan(\alpha_i \frac{\pi}{2})] \right\} \\ &+ \sum_{\substack{i=1 \\ \alpha_i = 1}}^K p_i \exp \left\{ i\mu_i t - \sigma_i^{\alpha_i} |t|^{\alpha_i} [1 - i\beta_i \frac{\pi}{2} \operatorname{sign}(t) \log(t)] \right\}. \end{aligned} \quad (1.25)$$

Capítulo 2

Estimação via Função Característica

2.1 Introdução

Na estimação dos parâmetros de densidades estáveis a falta de uma forma fechada simples para essas densidades levaram autores como Paulson *et al.* (1975), Heathcote (1977) e Xu e Knight (2011) a contornarem esse problema mediante a minimização da distância objetiva da função característica $\varphi_X(t; \theta)$ com a função característica empírica $\tilde{\varphi}_n(t)$.

Neste capítulo, apresenta-se esse método de estimação. Na seção 2.2, é apresentado o Estimador via Função Característica (EFC) de Θ , pela minimização da função distância objetiva.

A consistência e normalidade assintótica do estimador EFC para Θ são apresentadas, para os casos univariado e multivariado, na seção 2.3.

O estimador EFC não faz restrição ao tipo de variável aleatória para a qual é válido. Assim, na seção 2.4, esse estimador é aplicado às misturas finitas de normais. Revisam-se os procedimentos adotados em Xu e Knight (2011) para encontrar a forma fechada de $D_n(\Theta; x)$ e o seu processo iterativo para melhorar a eficiência dos estimadores de Θ .

Com o mesmo procedimento de Xu e Knight (2011), na seção 2.5, chega-se a uma forma fechada para a função distância objetiva correspondente à mistura finita de densidades Cauchy, que permite obter estimadores EFC da mistura.

Na seção 2.6, apresenta-se os resultados alcançados na tentativa de obter esti-

madores EFC para misturas de densidades de variáveis aleatórias estáveis simétricas.

Os estimadores descritos nas seções 2.4 e 2.5 são aplicados a dados reais e simulados nas seção 2.7.

2.2 Função Distância Objetiva

Definição 2.2.1. *Seja $X_1, \dots, X_n$ uma amostra aleatória de uma variável aleatória X . A função distância objetiva é definida por*

$$D_n(\Theta, x) = \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \theta)|^2 dW(t), \quad (2.1)$$

onde $\tilde{\varphi}_n(t)$ é a função característica empírica definida em (1.23), $\varphi_X(t; \theta)$ é a função característica definida em (1.22) e $W(t)$ é uma função integrável.

Aplicando propriedades dos números complexos e das integrais de Reimann-Stieltjes em (2.1) obtém-se

$$\tilde{\varphi}_n(t) = U_n(t) + iV_n(t)$$

e

$$\varphi_X(t) = u(t; \theta) + iv(t; \theta).$$

Assim, (2.1) torna-se

$$\begin{aligned} D_n(\Theta, x) &= \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \theta)|^2 dW(t) \\ &= \int_{-\infty}^{+\infty} |U_n(t) + iV_n(t) - u(t; \theta) - iv(t; \theta)|^2 dW(t) \\ &= \int_{-\infty}^{+\infty} |(U_n(t) - u(t; \theta)) + i(V_n(t) - v(t; \theta))|^2 dW(t) \\ &= \int_{-\infty}^{+\infty} ((U_n(t) - u(t; \theta))^2 + (V_n(t) - v(t; \theta))^2) dW(t) \\ &= \int_{-\infty}^{+\infty} ((U_n(t) - u(t; \theta))^2 + (V_n(t) - v(t; \theta))^2) w(t) dt. \end{aligned} \quad (2.2)$$

A função $w(t) = W'(t)$, em (2.2), designada neste trabalho por função peso, tem como objetivo garantir a convergência de (2.2). Paulson *et al.* (1975) usa $w(t) = \exp(-t^2)$ por conveniência computacional e sugere, em casos específicos, $w(t) = \exp(-|t|)$. Heathcote (1977) descreve de forma genérica $W(t)$ como uma

função não decrescente independente dos parâmetros e dos dados amostrais. Contudo, considera a conveniência computacional como um fator relevante nessa escolha. Xu e Knight (2011) escolheu $w(t) = \exp(-bt^2)$, com $b > 0$, a fim de encontrar estimadores mais eficientes. Nesses trabalhos, o estimador de Θ a ser alcançado é $\tilde{\Theta}$, que minimiza (2.2), e é tal que

$$\tilde{\Theta} = \underset{\Theta \in \Theta}{\operatorname{argmim}} \left[\int_{-\infty}^{+\infty} ((U_n(t) - u(t; \theta))^2 + (V_n(t) - v(t; \theta))^2) w(t) dt. \right], \quad (2.3)$$

que é designado por estimador via função característica ou EFC.

2.3 Propriedades Assintóticas

Observada a condição de regularidade exigida para a prova da consistência de $\tilde{\Theta}$ que é

$$\frac{\partial}{\partial \Theta} D_n(\Theta, x) = \int_{-\infty}^{+\infty} \frac{\partial}{\partial \Theta} |\tilde{\varphi}_n(t) - \varphi_X(t; \theta)|^2 dW(t), \quad (2.4)$$

Heathcote (1977) prova a consistência e a normalidade assintótica do estimador $\tilde{\Theta}$. O resultado alcançado é o seguinte.

Proposição 2.3.1 (Heathcote(1977) - caso univariado). *Mantida a condição de regularidade (2.4): (i) há um estimador $\tilde{\Theta}$ consistente, para Θ , que minimiza (2.2); (ii) se $u''(t; \theta)$ e $v''(t; \theta)$ são uniformemente dominados por uma função integrável então*

$$\sqrt{n} (\tilde{\Theta}_n - \Theta) \rightarrow N \left[0; \frac{\operatorname{Var}(K(\Theta_0))}{\lambda^2(\Theta_0)} \right]$$

com

$$\lambda(\Theta_0) = \int_{-\infty}^{+\infty} |\varphi'(t, \Theta_0)|^2 dW(t)$$

e

$$\begin{aligned} \operatorname{Var}(K(\Theta)) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [\operatorname{cov}\{\cos(tX), \cos(sX)\} u'(t, \Theta) u'(s, \Theta) \\ &+ 2\operatorname{cov}\{\cos(tX), \sin(sX)\} u'(t, \Theta) u'(s, \Theta) \\ &+ \operatorname{cov}\{\sin(tX), \sin(sX)\} s'(t, \Theta) s'(s, \Theta)] dW(t) dW(s), \end{aligned}$$

onde

$$K(\Theta) = \int_{-\infty}^{+\infty} [\cos(tX) - u(t; \Theta)] u'(t; \Theta) + [\sin(tX) - v(t; \Theta)] v'(t; \Theta) dW(t).$$

A Proposição 2.3.1 foi diretamente estendida para o caso multivariado no mesmo trabalho, na seguinte forma

Proposição 2.3.2 (Heathcote (1977) - caso multivariado). *Mantida a condição de regularidade (2.4):(i) há um estimador $\tilde{\Theta}$ consistente, para $\Theta = (\theta_1, \dots, \theta_i)$, que minimiza (2.2); (ii) se $u''(t; \Theta)$ e $v''(t; \Theta)$ são uniformemente dominados por uma função integrável então*

$$\sqrt{n} \left(\tilde{\Theta}_n - \Theta \right) \rightarrow N \left[0; \Lambda^{-1} \Sigma \Lambda^{-1} \right]$$

onde, $\Lambda^{-1} \Sigma \Lambda^{-1}$ é a matriz de covariância das variáveis aleatórias.

$$\begin{aligned} K^{(i)}(\Theta) &= \int_{-\infty}^{+\infty} [\cos(tX) - u(t; \Theta)] \frac{\partial u(t; \Theta)}{\partial \theta_i} dW(t) \\ &+ \int_{-\infty}^{+\infty} [\sin(tX) - v(t; \Theta)] \frac{\partial v(t; \Theta)}{\partial \theta_i} dW(t), \end{aligned}$$

para $i = 1, \dots, k$ e Λ é uma matriz simétrica $k \times k$ na qual o elemento (i, j) é

$$\lambda_{(i,j)} = \int_{-\infty}^{+\infty} \left\{ \frac{\partial u(t; \Theta)}{\partial \theta_i} \frac{\partial u(t; \Theta)}{\partial \theta_j} + \frac{\partial v(t; \Theta)}{\partial \theta_i} \frac{\partial v(t; \Theta)}{\partial \theta_j} \right\} dW(t). \quad (2.5)$$

A dificuldade na aplicação da Proposição 2.3.1 e da Proposição 2.3.2 consiste nos cálculos envolvidos para determinar, analiticamente, $Var(K(\Theta_0))$, no caso univariado, e a matriz $\Lambda^{-1} \Sigma \Lambda^{-1}$, no caso multivariado. A proposição a seguir, expressa os elementos da matriz de covariância em termos de esperança matemática.

Proposição 2.3.3 (Xu e Knight (2011)). *Mantida a condição de regularidade (2.4):(i) há um estimador $\tilde{\Theta}$ consistente, para $\Theta = (\theta_1, \dots, \theta_i)$, que minimiza (2.2); (ii) se adicionalmente, $u''(t; \Theta)$ e $v''(t; \Theta)$ são uniformemente dominados por uma função integrável, então*

$$\sqrt{n} \left(\tilde{\Theta}_n - \Theta \right) \rightarrow N \left[0; \Lambda^{-1} \Sigma \Lambda^{-1} \right],$$

cujos elementos da matriz Λ e Σ são

$$\Lambda_{ij} = E \left(\frac{\partial D_n^2(\Theta; x)}{\partial \Theta_i \partial \Theta_j} \right) \quad (2.6)$$

e

$$\Sigma_{ij} = E \left(\frac{\partial D_n(\Theta; x)}{\partial \Theta_i} \frac{\partial D_n(\Theta; x)}{\partial \Theta_j} \right). \quad (2.7)$$

onde Θ_i e Θ_j são os elementos i e j do vetor de parâmetros Θ .

A Proposição 2.3.3 permite a estimação numérica da matriz de covariância $\Lambda^{-1}\Sigma\Lambda^{-1}$ quando há uma forma fechada para $D_n(\Theta; x)$.

Heathcote (1977) alerta para a baixa eficiência dos estimadores EFC gerados ao usar a função peso $w(t) = e^{-t^2}$ em (2.2). Xu e Knight (2011) usa o valor do traço do inverso da matriz de informação sobre o traço da matriz $\Lambda^{-1}\Sigma\Lambda^{-1}$ como a medida de eficiência para os estimadores EFC para uma mistura de normais. A essa medida chamaram de Eficiência Relativa Assintótica. Com a escolha por $w(t) = e^{-bt^2}$, $b > 0$, todas as entradas da matriz $\Lambda^{-1}\Sigma\Lambda^{-1}$ dependem de b e, por isso, propôs um processo iterativo para escolher o valor ótimo para b , com o objetivo de melhorar a eficiência relativa assintótica dos estimadores obtidos ao minimizar (2.11). O processo iterativo compõe-se das seguintes fases:

1. Determine um valor inicial para b , digamos b_0 ;
2. Faça $b = b_0$, execute o procedimento que minimize $D_n(\Theta, x)$ para estimar Θ_0 , ou seja, $\hat{\Theta}_0 = \underset{\hat{\Theta} \in \Theta}{\operatorname{argmim}} \left[D_n(\Theta; x) \right]$;
3. Use $\hat{\Theta}_0$ para calcular a matriz de covariância $\Lambda^{-1}\Sigma\Lambda^{-1}$ da Proposição 2.3.3, digamos S_0 , na qual cada entrada está em função de b ;
4. Atualize b através da minimização de uma certa medida (MS), tal como o traço ou o determinante da matriz de covariância estimada no passo 2, ou seja, $b_1 = \underset{b_1 \in \mathbb{R}^+}{\operatorname{argmim}} \left[MS(S_0) \right]$;
5. Usando o valor atualizado de b no passo 4, repita os passos de 2 a 4 até que $|b_j - b_{j-1}|$ seja suficientemente pequeno.

2.4 Estimador EFC Para Mistura de Normais

Quando as K componentes de uma mistura finita, em (1.20), são

$$S_2 \left(\frac{\sigma_j}{\sqrt{(2)}}, 0, \mu_j \right) \sim N(\mu_j, \sigma_j^2), \quad (2.8)$$

ou

$$f_j(x; \theta_j) = \frac{1}{\sqrt{2}\sigma_j} e^{-\frac{1}{2} \left(\frac{x - \mu_j}{\sigma_j} \right)^2},$$

então,

$$h(x; \Theta) = \sum_{j=1}^K p_j \frac{1}{\sqrt{2\sigma_j}} e^{-\frac{1}{2} \left(\frac{x - \mu_j}{\sigma_j} \right)^2}, \quad (2.9)$$

onde, $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e $\theta_j = (\mu_j, \sigma_j^2)$ são os parâmetros da j -ésima componente f_j , é uma mistura finita de K componentes normais.

Para estimar $\tilde{\Theta}$, de (2.3), com a função peso $w(t) = e^{-bt^2}$, $b > 0$, minimiza-se

$$D_n(\Theta, x) = \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 \exp(-bt^2) dt, \quad (2.10)$$

onde

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j}$$

é a função característica empírica, e

$$\varphi(t, \Theta) = \sum_{i=1}^k \exp(i\mu_k t - \frac{\sigma_k^2 t^2}{2})$$

é a função característica da mistura. A solução é apresentada na proposição a seguir.

Proposição 2.4.1 (Xu e Knight (2011)). *Se uma amostra aleatória é gerada a partir de uma mistura finita de K variáveis aleatórias com distribuição Normal($\mu_j; \sigma_j^2$), diferentes, a função distância objetiva $D_n(\Theta, x)$ é*

$$\begin{aligned} D_n(\Theta, x) &= \frac{1}{n^2} \sqrt{\frac{\pi}{b}} \sum_{i=1}^n \sum_{j=1}^n \exp\left(-\frac{1}{4b}(x_i - x_j)^2\right) \\ &+ \sum_{k=1}^K p_k^2 \sqrt{\frac{\pi}{b + \sigma_k^2}} \\ &+ 2 \sum_{k=1}^K \sum_{\substack{h=1 \\ h \neq k}}^K p_k p_h \sqrt{\frac{\pi}{b + 0.5(\sigma_k^2 + \sigma_h^2)}} \exp\left(-\frac{(\mu_k - \mu_h)^2}{4b + 2(\sigma_k^2 + \sigma_h^2)}\right) \\ &- \frac{2}{n} \sum_{k=1}^K p_k \sqrt{\frac{\pi}{0.5\sigma_k^2 + b}} \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu_k)^2}{4b + 2\sigma_k^2}\right). \end{aligned} \quad (2.11)$$

A prova da Proposição 2.4.1 está no Apêndice B. O resultado (2.11) é de simples implementação computacional, característica que se mantém independente do número $K \geq 2$ de componentes da mistura estudada. Além disso, depende do valor escolhido para a variável $b > 0$ da função peso $w(t) = e^{-bt^2}$.

No capítulo 4 é apresentado um experimento que mostra como o valor de b influencia na variância dos estimadores EFC gerados para estimar os parâmetros $\theta = (\mu, \sigma^2)$ de uma variável aleatória normal. O resultado de uma aplicação com dados reais de consumo de petróleo, do estimador de EFC, para uma mistura de duas componentes normais é apresentado na seção 2.7.

O mesmo procedimento utilizado para obter (2.2) é aplicado para a mistura de componentes Cauchy.

2.5 Estimador EFC Para Mistura de Cauchy

Quando as K componentes de uma mistura finita, em (1.20), são

$$S_1(\sigma_j, 0, \mu_j) \sim \text{Cauchy}(\mu_j, \sigma_j), \quad (2.12)$$

ou

$$f_j(x; \theta_j) = \frac{1}{\sigma_j \pi \left[1 + \left(\frac{x - \mu_j}{\sigma_j} \right)^2 \right]},$$

então

$$h(x; \Theta) = \sum_{j=1}^K p_j \frac{1}{\sigma_j \pi \left[1 + \left(\frac{x - \mu_j}{\sigma_j} \right)^2 \right]}, \quad (2.13)$$

onde, $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e $\theta_j = (\mu_j, \sigma_j)$ são os parâmetros da j -ésima componente f_j , tem-se uma mistura finita de K componentes Cauchy.

Para estimar $\tilde{\Theta}$, em (2.2), com a função peso $w(t) = e^{-b|t|}$, $b > 0$, minimiza-se

$$D_n(\Theta, x) = \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 \exp(-b|t|) dt, \quad (2.14)$$

onde

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j} \quad (2.15)$$

é a função característica empírica, e

$$\varphi_H(t; \Theta) = \sum_{k=1}^2 p_k \exp(-\sigma_k |t| + i\mu_k t) \quad (2.16)$$

é a função característica da mistura. A solução apresentada está na proposição a seguir.

Proposição 2.5.1. *Se uma amostra aleatória é gerada a partir de uma mistura finita de K distribuições $\text{Cauchy}(\mu_j; \sigma_j)$, diferentes, então a função distância objetiva $D_n(\Theta, x)$ é*

$$\begin{aligned}
D(\Theta, x) &= \frac{2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{b}{b^2 + (x_i - x_j)^2} \\
&+ 2 \sum_{k=1}^K p_k^2 \frac{1}{2\sigma_k + b} \\
&+ 4 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \frac{\sigma_k + \sigma_j + b}{(\sigma_k + \sigma_j + b)^2 + (\mu_k - \mu_j)^2} \\
&- \frac{4}{n} \sum_{k=1}^K \sum_{j=1}^n p_k \frac{\sigma_k + b}{(\sigma_k + b)^2 + (\mu_k - x_j)^2}.
\end{aligned} \tag{2.17}$$

A prova da Proposição 2.5.1 está no Apêndice C. Estimativas para dados simulados, ao minimizar (2.17), são mostradas na seção 2.7.

O principal resultado deste capítulo é apresentado na próxima seção, quanto à tentativa de obter o estimador EFC para uma mistura de densidades estáveis simétricas.

2.6 Estimador EFC Para Mistura de Estáveis Simétricas

Quando as K componentes de uma mistura finita, em (1.20), são

$$F_j(x; \alpha_j, \sigma_j, \beta_j, \mu_j) = S_{\alpha_j}(\sigma_j, \beta_j, \mu_j)(x), j = 1, \dots, K$$

com

$$f_j(x_j; \theta_j) = F'_j(x; \theta_j), j = 1, \dots, K$$

então

$$h(x; \Theta) = \sum_{j=1}^K p_j f_j(x; \theta_j), \tag{2.18}$$

onde $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e $\theta_j = (\alpha_j, \sigma_j, \beta_j = 0, \mu_j)$, é uma mistura finita de K componentes estáveis simétricas.

Para estimar $\tilde{\Theta}$ de (2.2), com a função peso $w(t)$, minimiza-se

$$D_n(\Theta, x) = \int_{-\infty}^{+\infty} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 w(t) dt, \quad (2.19)$$

onde

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itX_j} \quad (2.20)$$

é a função característica empírica, e

$$\varphi_X(t) = \sum_{i=1}^K p_i \exp \{i\mu_i t - \sigma_i^\alpha |t|^\alpha\}$$

é a função característica da mistura. Procedemos a obter uma forma fechada de (2.19), como no caso Normal e Cauchy.

Nesta situação, não foi possível chegar a uma forma simples para (2.19), mas, chegou-se a uma representação por série para a mistura com $K = 2$ densidades estáveis simétricas com o mesmo índice de estabilidade α . Para chegar nos resultados (2.2) e (2.19) utilizam-se de cálculos de integrais que envolvam a função Gama. A expressão encontrada é

$$\begin{aligned} D_n(\Theta, x) &= \frac{2p^2 \Gamma(\frac{1-b}{\alpha})}{\alpha (2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2 \Gamma(\frac{1-b}{\alpha})}{\alpha (2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\ &+ \frac{4p(1-p)}{\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k (\mu_1 - \mu_2)^{2k} \Gamma(\frac{2k+1-b}{\alpha})}{(\sigma_1^\alpha + \sigma_2^\alpha)^{\frac{2k+1-b}{\alpha}}} \\ &- \frac{4p}{\alpha n} \sum_{j=1}^n \left[\sum_{k=0}^{\infty} \frac{(-1)^k (\mu_1 - x_j)^{2k} \Gamma(\frac{2k+1-b}{\alpha})}{\sigma_1^{2k+1-b}} \right] \\ &- \frac{4(1-p)}{\alpha n} \sum_{j=1}^n \left[\sum_{k=0}^{\infty} \frac{(-1)^k (\mu_2 - x_j)^{2k} \Gamma(\frac{2k+1-b}{\alpha})}{\sigma_2^{2k+1-b}} \right], \quad (2.21) \end{aligned}$$

quando, em (2.19), $w(t) = |t|^{-b}$, $0 < b < 1$. Outra representação para (2.19) é

$$\begin{aligned} D_n(\Theta, x) &= 2p^2 \sum_{k=0}^{\infty} \frac{(-1)^k (2\sigma_1^{\alpha k}) \Gamma(\alpha k + 1)}{k!} + 2(1-p)^2 \sum_{k=0}^{\infty} \frac{(-1)^k (2\sigma_2^{\alpha k}) \Gamma(\alpha k + 1)}{k!} \\ &+ 8p(1-p) \sum_{k=0}^{\infty} \frac{(-1)^k (\sigma_1^\alpha + \sigma_2^\alpha)^k \Gamma(\alpha k + 1) \cos[(\alpha k + 1) \arctan(\mu_1 - \mu_2)]}{k! (\alpha k + 1) [(\mu_1 - \mu_2)^2 + 1(\alpha k + 1)]} \\ &- \frac{8p}{n} \sum_{j=1}^n \sum_{k=0}^{\infty} \frac{(-1)^k \sigma_1^{\alpha k} \Gamma(\alpha k + 1) \cos[(\alpha k + 1) \arctan(\mu_1 - x_j)]}{k! (\alpha k + 1) [(\mu_1 - x_j)^2 + 1(\alpha k + 1)]} \\ &- \frac{8(1-p)}{n} \sum_{j=1}^n \sum_{k=0}^{\infty} \frac{(-1)^k \sigma_2^{\alpha k} \Gamma(\alpha k + 1) \cos[(\alpha k + 1) \arctan(\mu_2 - x_j)]}{k! (\alpha k + 1) [(\mu_2 - x_j)^2 + 1(\alpha k + 1)]} \quad (2.22) \end{aligned}$$

quando, em (2.19), $w(t) = e^{-|t|}$.

Tanto em (2.21) como em (2.22) não se obteve sucesso nos procedimentos de minimização, entretanto não se deve descartar a possibilidade de minimizá-las.

2.7 Aplicação do Estimador EFC em dados reais

A fim de ilustrar o uso de (2.11), o log do consumo *per capita* de petróleo, em 135 países no ano de 2001, disponíveis no site do Banco Mundial no endereço <http://data.worldbank.org/indicator/UE.USE.PCAP.KG.OE>, foi modelado como uma mistura finita $H_n(X)$ de duas normais.

Para isso, considere $K = 2$, em (2.11), para obter

$$\begin{aligned}
D_n(\Theta, x) &= \frac{1}{n^2} \sqrt{\frac{\pi}{b}} \sum_{i=1}^n \sum_{j=1}^n \exp\left(-\frac{1}{4b}(x_i - x_j)^2\right) \\
&+ p_1^2 \sqrt{\frac{\pi}{b + \sigma_1^2}} \\
&+ (1 - p_1)^2 \sqrt{\frac{\pi}{b + \sigma_2^2}} \\
&+ 2p_1(1 - p_1) \sqrt{\frac{\pi}{b + 0.5(\sigma_1^2 + \sigma_2^2)}} \exp\left(-\frac{(\mu_1 - \mu_2)^2}{4b + 2(\sigma_1^2 + \sigma_2^2)}\right) \\
&- \frac{2}{n} p_1 \sqrt{\frac{\pi}{0.5\sigma_1^2 + b}} \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu_1)^2}{4b + 2\sigma_1^2}\right) \\
&- \frac{2}{n} (1 - p_1) \sqrt{\frac{\pi}{0.5\sigma_2^2 + b}} \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu_2)^2}{4b + 2\sigma_2^2}\right). \tag{2.23}
\end{aligned}$$

A estimativa gerada pela minimização de (2.23), com $b = 1$, valor sugerido por Paulson *et al.* (1975), resulta na mistura

$$H(x) = 0.447 \times N(6.278, 0.255) + 0.553 \times N(8.020, 0.500).$$

A Figura 2.1 mostra o histograma dos dados e a linha contínua é a densidade de $H(x)$.

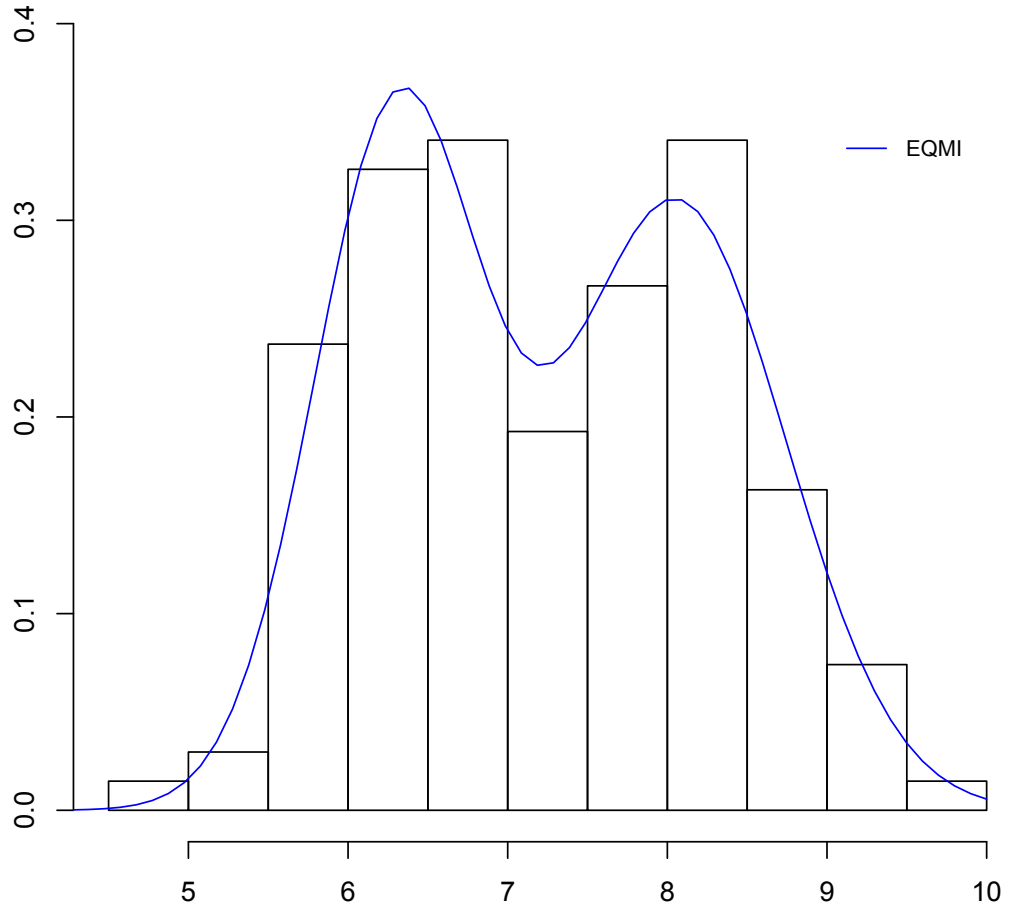


Figura 2.1: Histograma do consumo de petróleo em 135 países no ano de 2001 e a densidade $H(x)$ estimada por (2.23).

2.7.1 Aplicação do Estimador EFC em misturas de Cauchy

A fim de obter estimativas EFC para misturas com duas componentes Cauchy pela uso de (2.17), foram replicadas 1000 (mil) vezes as misturas da Tabela 2.1. Essas misturas foram escolhidas para verificar o desempenho do EFC quando há: diferenças nos tamanhos das amostras, $A1$ e $A2$; nos pesos das componentes, $B1$ e $B2$; nos parâmetros de locação, $C1$ e $C2$; e nos parâmetros de escala, $D1$ e $D2$.

Para executar esses experimentos é necessário exprimir $D(\Theta, x)$, em (2.17),

Tabela 2.1: Pesos e parâmetros, das componentes Cauchy, e tamanho das amostras utilizadas nos experimentos.

Experimento	Componente 1			Componente 2			n
	p_1	μ_1	σ_1	p_2	μ_2	σ_2	
A1	0.50	-3	1	0.50	3	1	100
A2	0.50	-3	1	0.50	3	1	200
B1	0.50	0	1	0.50	4	1	100
B2	0.75	0	1	0.25	4	1	100
C1	0.50	0	1	0.50	6	1	100
C2	0.50	0	1	0.50	10	1	100
D1	0.50	0	1	0.50	0	5	100
D2	0.50	0	1	0.50	0	10	100

para $K = 2$, que resulta em

$$\begin{aligned}
D(\Theta, x) &= \frac{2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{b}{b^2 + (x_i - x_j)^2} \\
&+ 2p_1^2 \frac{1}{2\sigma_1 + b} \\
&+ 2(1 - p_1)^2 \frac{1}{2\sigma_2 + b} \\
&+ 4p_1(1 - p_1) \frac{\sigma_1 + \sigma_2 + b}{(\sigma_1 + \sigma_2 + b)^2 + (\mu_1 - \mu_2)^2} \\
&- \frac{4}{n} \sum_{j=1}^n p_1 \frac{\sigma_1 + b}{(\sigma_1 + b)^2 + (\mu_1 - x_j)^2} \\
&- \frac{4}{n} \sum_{j=1}^n (1 - p_1) \frac{\sigma_2 + b}{(\sigma_2 + b)^2 + (\mu_2 - x_j)^2}.
\end{aligned} \tag{2.24}$$

A Tabela 2.2 apresenta as médias das estimativas EFC. Nessa tabela, chama a atenção o mau desempenho do método nos casos D1 e D2, onde os parâmetros de locação são iguais. Também, foram calculados o erro quadrático médio (EQM) e o viés médio das estimativas obtidas, esses resultados estão nas Tabelas A.1 e A.2 do Apêndice A.

Também foi aplicado (2.24) ao log do consumo *per capita* de petróleo em 135 países no ano de 2001. As estimativas obtidas estão na Tabela 2.3. A Figura 2.2 contém o histograma desses dados e a densidade estimada da mistura, representada

Tabela 2.2: Média das estimativas EFC geradas nos experimentos descritos da tabela (2.1).

Experimento	p_1	μ_1	σ_1	μ_2	σ_2
A1	0.50053	-2.99598	0.99678	3.00464	1.00435
A2	0.49976	-3.00125	1.00765	3.00210	0.99352
B1	0.50690	0.00744	1.00115	3.98398	0.97763
B2	0.72005	-0.02182	0.95188	3.87715	1.05853
C1	0.50256	-0.00014	1.00789	6.00745	0.98205
C2	0.50052	-0.00260	1.01660	9.99094	0.99291
D1	0.51771	-0.87086	3.26692	0.95838	3.15718
D2	0.50287	-1.14071	5.66872	1.08797	5.65397

pela linha contínua.

Tabela 2.3: Estimativas EFC do log do consumo de petróleo com componentes Cauchy.

p_1	μ_1	σ_1	p_2	μ_2	σ_2
0.5041649	6.4152190	0.2816420	0.4958310	8.0439988	0.3317306

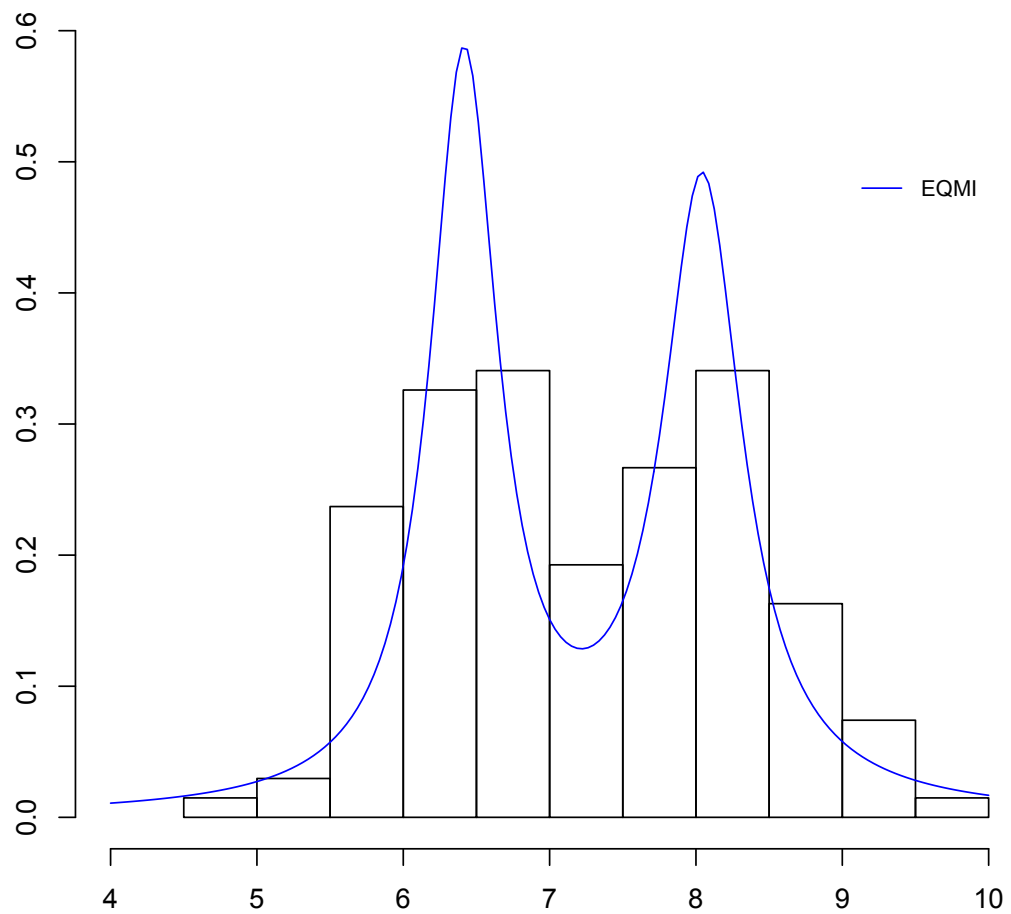


Figura 2.2: Histograma do consumo de petróleo em 135 países no ano de 2001 e a densidade de $H(x)$ estimada por (2.24)

Capítulo 3

Estimação pelo Algoritmo EM

3.1 Introdução

O algoritmo EM é um dos métodos mais usados para estimar os parâmetros de modelos de misturas finitas, sempre que as componentes satisfaçam certas condições de regularidade. O EM foi utilizado neste trabalho por apresentar um algoritmo que, em tese, pode estimar qualquer mistura finita de variáveis aleatórias, em particular, de estáveis.

Na seção 3.2 é apresentado o algoritmo EM para mistura finita de variáveis aleatórias. Na seção 3.3, o EM é aplicado a misturas finitas de densidades estáveis, enfatizando-se o cálculo numérico dessas densidades. Na seção 3.4 esse método de estimação é aplicado a dados reais e a dados simulados.

3.2 EM Para Mistura Finita de Variáveis Aleatórias

A estimação dos parâmetros da mistura de $K \geq 2$ densidades é uma das aplicações mais utilizadas do algoritmo EM. Nesse caso, considere a seguinte densidade

$$h(x; \Theta) = \sum_{j=1}^K p_j f_j(x; \theta_j), \quad (3.1)$$

onde $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ é o vetor de parâmetros, $\sum_{j=1}^K p_j = 1$, e cada componente da mistura f_j é uma densidade com parâmetro θ_j . A expressão da log-

verossimilhança incompleta de Θ dado $\mathcal{X} = \{x_1, \dots, x_N\}$ é

$$\log(L(\Theta|\mathcal{X})) = \log \prod_{i=1}^N h(x_i; \Theta) = \sum_{i=1}^N \log \left(\sum_{j=1}^K p_j f_j(x_i; \theta_j) \right). \quad (3.2)$$

Considere $\mathcal{X}$ como incompleto e $\mathcal{Y} = \{y_i\}_{i=1}^N$ como uma variável aleatória desconhecida, cujos valores informam qual componente gerou cada dado, isto é, $y_i \in \{1, \dots, K\}$ para valor em $\mathcal{X}$, e $y_i = j$ se o i -ésimo dado foi gerado pela j -ésima componente da mistura. Nessa situação, a expressão da verossimilhança, em (3.2), é simplificada. Assim,

$$\begin{aligned} \log(L(\Theta|x, y)) &= \log(f(\mathcal{X}, \mathcal{Y}|\Theta)) \\ &= \log \left(\prod_{i=1}^N f(x_i, y_i|\Theta) \right) \\ &= \sum_{i=1}^N \log(f(x_i|y_i, \Theta)f(y_i|\Theta)) \\ &= \sum_{i=1}^N \log(p_i f(y_i|\theta_{y_i})) \end{aligned} \quad (3.3)$$

com

$$f(y_i|\theta) = P(Y = y_i) = p_{y_i}, \quad (3.4)$$

e

$$f(x_i|y_i, \Theta) = f_{y_i}(y_i|\theta_{y_i}) \quad (3.5)$$

que pode ser otimizada utilizando várias técnicas.

Para contornar o fato da não existência de $\mathcal{Y}$, esses valores devem ser estimados. Primeiro, deve-se encontrar uma expressão para a distribuição dos dados faltantes. Assuma-se que os parâmetros da mistura $\Theta^{(q)} = (p_1^{(q)}, \dots, p_k^{(q)}, \theta_1^{(q)}, \dots, \theta_k^{(q)})$ são apropriados para a verossimilhança $L(\Theta^{(q)}|\mathcal{X}, \mathcal{Y})$. Dado $\Theta^{(q)}$, calcule $f_j(x_i|\theta_j^{(q)})$ para cada i e j . As proporções da mistura, p_j , podem ser consideradas como sendo probabilidades *a priori* para cada componente da mistura. Então, pela regra de Bayes,

$$f(y_i|x_i, \Theta^{(q)}) = \frac{p_{y_i}^{(q)} f_{y_i}(x_i|\theta_{y_i}^{(q)}, y_i)}{h(x_i|\Theta^{(q)})} = \frac{p_{y_i}^{(q)} f_{y_i}(x_i|\theta_{y_i}^{(q)}, y_i)}{\sum_{j=1}^{(q)} p_j^{(q)} f_j(x_i|\theta_j^{(q)})} \quad (3.6)$$

e

$$f(\dagger|\mathcal{X}, \Theta^{(q)}) = \prod_{i=1}^N f(y_i|x_i, \Theta^{(q)}) \quad (3.7)$$

onde $\dagger = (y_1, \dots, y_N)$ é um exemplo de dado faltante com espaço paramétrico Υ . Com isso, foi possível encontrar a densidade marginal assumindo a existência de variáveis e escolhendo valores iniciais para os parâmetros da distribuição.

Neste caso, define a equação

$$\begin{aligned} Q(\Theta, \Theta^{(q)}) &= E [\log f((\mathcal{X}, \mathcal{Y}|\Theta)|\mathcal{X}, \Theta^{(q)})] \\ &= \sum_{\ell=1}^K \sum_{i=1}^N \log(p_\ell f_\ell(x_i|\theta_\ell)) f(\ell|x_i, \Theta^{(q)}) \\ &= \sum_{\ell=1}^K \sum_{i=1}^N \log(p_\ell) f(\ell|x_i, \Theta^{(q)}) \end{aligned} \quad (3.8)$$

$$+ \sum_{\ell=1}^K \sum_{i=1}^N \log(f_\ell(x_i|\theta_\ell)) f(\ell|x_i, \Theta^{(q)}) \quad (3.9)$$

onde $\ell \in \{1, \dots, K\}$.

Essa expressão é maximizada em termos de p_ℓ e de θ_ℓ separadamente. Atento à restrição $\sum_{\ell=1}^K p_\ell = 1$, com a finalidade de encontrar a expressão de p_ℓ pelo método do multiplicador de Lagrange λ_p , chega-se a

$$p_\ell = \frac{1}{N} \sum_{i=1}^N f(\ell|x_i, \Theta^{(q)}).$$

Para encontrar as atualizações dos parâmetros θ_ℓ basta maximizar a segunda parcela de (3.9) em relação à θ_ℓ , assim

$$\theta_\ell = \underset{\theta_\ell \in \Theta}{\operatorname{argmax}} \left[\sum_{\ell=1}^K \sum_{i=1}^N \log(f_\ell(x_i|\theta_\ell)) f(\ell|x_i, \Theta^{(q)}) \right].$$

Em síntese, o algoritmo EM para estimar os parâmetros de uma mistura de distribuições é a repetição dos seguintes passos até que algum critério de convergência seja alcançado.

- Passo E

Calcular

$$Q(\Theta, \Theta^{(q)}) = \sum_{\ell=1}^K \sum_{i=1}^N \log(p_\ell f_\ell(x_i|\theta_\ell^{(q)})) f(\ell|x_i, \Theta^{(q)}) \quad (3.10)$$

onde

$$f(\ell|x_i, \Theta^{(q)}) = \frac{p_\ell^{(q)} f_\ell(x_i|\theta_\ell^{(q)})}{\sum_{j=1}^K p_j^{(q)} f_j(x_i|\theta_j^{(q)})} \quad (3.11)$$

- Passo M

Determinar os estimadores de p_ℓ e θ_ℓ , usando

$$p_\ell = \frac{1}{N} \sum_{i=1}^N f(\ell|x_i, \Theta^{(q)}) \quad (3.12)$$

e

$$\theta_\ell = \underset{\theta_\ell \in \Theta}{\operatorname{argmax}} \left[\sum_{\ell=1}^K \sum_{i=1}^N \log(f_\ell(x_i|\theta_\ell^{(q)})) f(\ell|x_i, \Theta^{(q)}) \right]. \quad (3.13)$$

É frequente encerrar o algoritmo quando um número de iterações $q = 1, \dots, L$ alcança o valor L , ou quando

$$|Q(\Theta, \Theta^{(q+1)}) - Q(\Theta, \Theta^{(q)})| < \epsilon,$$

para um valor de $\epsilon > 0$ pequeno.

3.3 EM Para Mistura Finita de Densidades Estáveis

Quando as K componentes, em (1.20), de uma mistura finita de distribuições são

$$F_j(x; \alpha_j, \sigma_j, \beta_j, \mu_j) = S_{\alpha_j}(\sigma_j, \beta_j, \mu_j)(x), \quad j = 1, \dots, K$$

com

$$f_j(x_j; \theta_j) = F'_j(x; \theta_j), \quad j = 1, \dots, K$$

então,

$$h(x; \Theta) = \sum_{j=1}^K p_j f_j(x; \theta_j), \quad (3.14)$$

onde $\sum_{j=1}^K p_j = 1$, $\Theta = (p_1, \dots, p_{K-1}, \theta_1, \dots, \theta_K)$ e $\theta_j = (\alpha_j, \sigma_j, \beta_j = 0, \mu_j)$ é uma mistura finita de K componentes estáveis.

As seguintes etapas descrevem o algoritmo EM para estimar (3.14).

1. Faça $q = 1, \dots, L$ um contador de iterações.
2. Faça $\ell = 1, \dots, K$ um identificador de componentes.
3. Determine $\Theta^{(0)}$ como valor inicial para $\Theta^{(q)}$.

4. Passo E

Calcular

$$Q(\Theta, \Theta^{(q)}) = \sum_{\ell=1}^K \sum_{i=1}^N \log(p_{\ell} f_{\ell}(x_i | \theta_{\ell}^{(q)})) f(\ell | x_i, \Theta^{(q)}), \quad (3.15)$$

onde

$$f(\ell | x_i, \Theta^{(q)}) = \frac{p_{\ell}^{(q)} f_{\ell}(x_i | \theta_{\ell}^{(q)})}{\sum_{j=1}^M p_{\ell}^{(q)} f_{\ell}(x_i | \theta_{\ell}^{(q)})}, \quad (3.16)$$

com $f_{\ell}(x | \theta_{\ell}^{(q)})$ gerados pelos algoritmos I ou II na seção 4.3.

5. Passo M

Determinar os estimadores de p_{ℓ} e θ_{ℓ} , com

$$p_{\ell} = \frac{1}{N} \sum_{i=1}^N f(\ell | x_i, \Theta^{(q)}) \quad (3.17)$$

e

$$\theta_{\ell} = \underset{\theta_{\ell} \in \Theta}{\operatorname{argmax}} \left[\sum_{\ell=1}^K \sum_{i=1}^N \log(f_{\ell}(x | \theta_{\ell}^{(q)})) f(\ell | x_i, \Theta^{(q)}) \right]. \quad (3.18)$$

6. Atualize $\Theta^{(q+1)}$ com as estimativas obtidas no passo M em 5.

7. Se o critério de convergência não foi alcançado retorne à etapa 4.

Devido ao fato de não haver uma forma simples para a densidade da maioria dos membros da família das distribuições estáveis. Neste trabalho, os valores da densidade $f_{\ell}(x | \theta)$, em (3.15) e (3.16), são obtidos, por meio numérico, pelos algoritmos apresentados na seção 4.3. O primeiro, proposto em Nolan (1999), utiliza a fórmula de inversão, o segundo foi proposto por Otiniano *et al.* (2011) e utiliza a função especial H. O segundo algoritmo possibilita tratar melhor o cálculo de densidade para valores extremos, uma dificuldade apresentada nos dois procedimentos.

Em casos onde todos os componentes da mistura são uma distribuição *Normal*, ou uma *Cauchy* ou uma *Levy*_{1/2} podem ser encontrados estimadores exatos para θ_{ℓ} .

3.4 Aplicações do Algoritmo EM

3.4.1 Aplicação do EM a Dados Reais

Para ilustrar o uso do Método EM em dados reais, analisou-se o log do consumo *per capita* de petróleo em 135 países no ano de 2001. Esses dados foram ajustados

por uma mistura

$$H_n(x) = pN(\mu_1, \sigma_1^2) + (1 - p)N(\mu_2, \sigma_2^2), \quad (3.19)$$

cujas componentes são duas distribuições normais e pela mistura

$$H_s(x) = pS_2(\sigma_1, 0, \mu_1) + (1 - p)S_2(\sigma_2, 0, \mu_2), \quad (3.20)$$

cujas componentes são duas distribuições estáveis simétricas. As estimativas encontradas estão na Tabela 3.1, pág. 42.

A equivalência entre $H_n(x)$ e $H_s(x)$ pode ser melhor observada, lembrando que

$$S_2(\sigma, 0, \mu) \approx N(\mu, \frac{\sigma}{\sqrt{2}}),$$

assim, o modelo $H_s(x)$, em (3.20), sendo uma variável aleatória simétrica com valor de $\alpha = 2$, pode ser expresso por meio de componentes com distribuição normal mantendo-se os valores de locação e multiplicando os parâmetros de escala por $\sqrt{2}$. Logo

$$\begin{aligned} H_s(x) &= 0.4823223 \times S_2(0.3753994, 0, 6.3301055) \\ &+ 0.5176777 \times S_2(0.4713789, 0, 8.0929957) \\ &= 0.4823223 \times N(6.3301055, 0.5308949) \\ &+ 0.5176777 \times N(8.0929957, 0.6666317). \end{aligned}$$

A Figura 3.1, pág. 43, mostra as densidades de $H_n(x)$ e $H_s(x)$ junto com o histograma dos dados analisados.

Tabela 3.1: Estimativas EM para o consumo *per capita* de petróleo em 135 países, no ano de 2001, considerando componentes normais e componentes estáveis.

Parâmetros	Misturas	
	$H_n(x)$	$H_s(x)$
p_1	0.4769274	0.4823223
p_2	0.5230726	0.5176777
α	-	2.00
μ_1	6.3210750	6.3301055
μ_2	8.0809060	8.0929957
σ_1	0.5259982	0.3753994
σ_2	0.6735426	0.4713789

3.4.2 Aplicação do EM a Dados Simulados

Para fazer uma aplicação do Método EM com dados simulados, foram geradas 100 amostras, com 200 observações, de misturas com duas componentes estáveis simétricas, e com o mesmo valor para α . Nas sete situações analisadas, o valor de α é diferente, mas, os demais parâmetros permanecem constantes. A Tabela 3.3 apresenta os valores médios das estimativas alcançadas em cada experimento. As Tabelas A.5 e A.4 apresentam, respectivamente, o viés e o EQM dos parâmetros estimados.

Tabela 3.2: Experimento de estimação EM para misturas finitas de variáveis aleatórias estáveis.

Experimento	p	α	μ_1	μ_2	σ_1	σ_2
A	0.50	2.0	-3	3	1	1
B	0.50	1.8	-3	3	1	1
C	0.50	1.6	-3	3	1	1
D	0.50	1.4	-3	3	1	1
E	0.50	1.2	-3	3	1	1
F	0.50	1.0	-3	3	1	1
G	0.50	0.8	-3	3	1	1

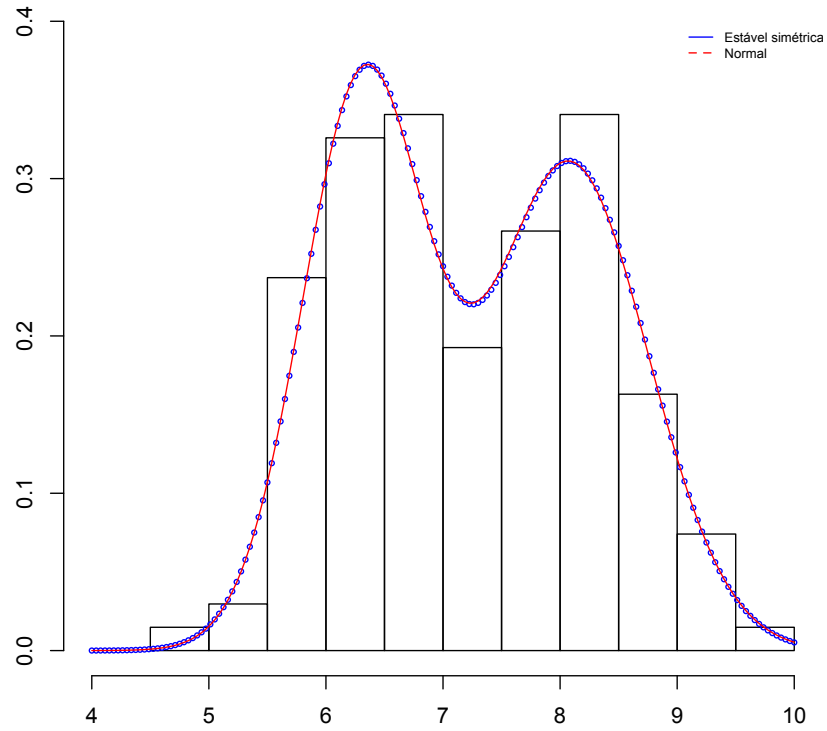


Figura 3.1: Histograma do consumo *per capita* de petróleo em 135 países no ano de 2001 e as densidades propostas pelo Método EM, considerando as componentes - distribuições estáveis e distribuições normais.

Tabela 3.3: Estimativas EM para misturas finitas de variáveis aleatórias estáveis.

Experimento	p	α	μ_1	μ_2	σ_1	σ_2
A	0.49860	1.98986	-3.00380	3.01856	0.99465	0.98742
B	0.49913	1.82035	-2.99848	3.00830	0.99067	0.97044
C	0.50332	1.60558	-3.02056	2.99551	0.97801	0.99503
D	0.50283	1.41429	-2.99396	3.00051	0.99767	0.98565
E	0.49539	1.20096	-3.01940	2.98301	0.98991	1.00026
F	0.50275	1.00280	-2.99881	3.01236	0.99090	1.00507
G	0.50314	0.79295	-3.00071	3.00181	0.99423	0.93766

Capítulo 4

Considerações Numéricas

4.1 Introdução

Neste capítulo são apresentados detalhes computacionais importantes que deixaram de ser explicitados nos capítulos anteriores.

4.2 Sobre os Programas Utilizados

Nesta dissertação todas as operações numéricas de integração, otimização e produção das figuras foram executadas no programa R. A parte algébrica do trabalho que envolve as integrais em (2.2), no caso de misturas de Normais e Cauchy, e o cálculo das derivadas parciais da Proposição 2.3.3 foram conferidas no programa Maple 14.

As otimizações foram realizadas pela função OPTIM. Essa escolha, deve-se ao fato dessa função disponibilizar, dentre outros métodos de otimização, uma adaptação do método Quasi-Newton, na qual é possível determinar intervalos de valores possíveis para as variáveis das funções otimizadas. Assim, pode-se impor as restrições $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$ e $\sigma > 0$ aos métodos de estimação EM e EFC.

As integrações numéricas realizadas utilizaram a função INTEGRATE. Durante o cálculo das densidades utilizadas pelo método EM, observou-se que essa função de integração alcançava com maior frequência o número máximo de subdivisões permitidas para o procedimento, à medida que o valor de α diminuía. Esse fato, gera um alerta que interrompia a execução do programa R. Para contornar esse problema,

usa-se a opção STOP.ON.ERRO=FALSE como uma opção ao aumento do número de subdivisões permitidas na opção SUBDIVISION, pois a situação acima não indica erro de estimativa.

Outro fato observado é que, quanto menor for α , maior é o tempo computacional para o algoritmo EM atingir o critério de convergência

$$|\Theta^{q+1} - \Theta^q| \leq 0.001,$$

e maior o número de interações realizadas no procedimento.

4.3 Densidade de Uma Variável Aleatória Estável

4.3.1 Algoritmo I - Fórmula de Inversão

Nolan (1999) apresenta como a função de densidade de uma variável aleatória $X \sim S_\alpha^0(1, \beta, 0)$ padrão ($\sigma = 1$ e $\mu = 0$), obtida por meio da fórmula de inversão, a função

$$f^0(x; \alpha, \beta) = \frac{1}{\pi} \int_0^\infty \cos(h(x, t; \alpha, \beta)) e^{-t^\alpha} dt, \quad (4.1)$$

onde,

$$h(x, t; \alpha, \beta) = \begin{cases} xt + \beta \left(\tan \frac{\pi\alpha}{2} \right) (t - t^\alpha), & \text{se } \alpha \neq 1 \\ xt + \beta \frac{2}{\pi} t \ln t, & \text{se } \alpha = 1 \end{cases},$$

cujo domínio é

$$x \in \begin{cases} [\mu - \sigma \left(\tan \frac{\pi\alpha}{2} \right), +\infty), & \text{se } \alpha \leq 1, \beta = 1 \\ [-\infty, \mu + \sigma \left(\tan \frac{\pi\alpha}{2} \right)], & \text{se } \alpha \leq 1, \beta = -1 \\ (-\infty, +\infty), & \text{caso contrário.} \end{cases}$$

Para calcular a densidade de uma v.a. estável $X \sim S_\alpha^0(\sigma, \beta, \mu)$ com $\sigma > 0$ e $-\infty < \mu < +\infty$ usa-se a relação

$$f^0(x; \alpha, \beta, \sigma, \mu) = \frac{1}{\sigma} f^0\left(\frac{x - \mu}{\sigma}; \alpha, \beta\right),$$

onde σ e μ são parâmetros de escala e localização, respectivamente.

Em situações onde a parametrização utilizada é a de Samorodnitsky e Taqqu (1994), isto é, $X \sim S_\alpha(\sigma, \beta, \mu)$ com $\sigma > 0$ e $-\infty < \mu < +\infty$, pode-se utilizar (4.1)

assim:

$$f(x; \alpha, \beta, \sigma, \mu) = \begin{cases} \frac{1}{\sigma} f^0 \left(\frac{x - (\sigma \beta \tan(\frac{\pi \alpha}{2}) + \mu)}{\sigma}; \alpha, \beta \right), & \text{se } \alpha \neq 1 \\ \frac{1}{\sigma} f^0 \left(\frac{x - ((\frac{2}{\pi}) \sigma \beta \ln(\sigma) + \mu)}{\sigma}; \alpha, \beta \right), & \text{se } \alpha = 1, \end{cases}$$

com domínio

$$x \in \begin{cases} [\mu, +\infty), & \text{se } \alpha \leq 1, \beta = 1 \\ [-\infty, \mu), & \text{se } \alpha \leq 1, \beta = -1 \\ (-\infty, +\infty), & \text{caso contrário.} \end{cases}$$

Apesar de fácil implementação, o uso de (4.1) apresenta algumas dificuldades numéricas:

1. Para $\alpha < 0.75$, a região de integração torna-se maior e é difícil alcançar resultados precisos da integração numérica.
2. Para $\beta \neq 0$ e $0 < |\alpha - 1| < 0.001$, há problemas numéricos com o termo $(\tan \frac{\pi \alpha}{2}) (t - t^\alpha)$.
3. Para x muito grande, o integrando oscila muito rápido.

4.3.2 Algoritmo II - Representação pela Função H

Usando a função especial H , Otiniano *et al.* (2011) propõem uma f.d.p. para $X \sim S_\alpha(\sigma, \beta, \mu)$, com $\mu = 0$, $0 < \alpha < 2$, mas $\alpha \neq 1$, e para $x > 0$. Assim, definindo

$$\tilde{\sigma} = \sigma \left(1 - \beta^2 \tan^2 \left(\alpha \frac{\pi}{2} \right) \right)^{-\frac{1}{2\alpha}} \quad (4.2)$$

e

$$\tilde{\beta} = \begin{cases} \frac{2}{\pi \alpha} \arctan \left(\beta \tan \frac{\alpha \pi}{2} \right), & \text{se } 0 < \alpha < 1 \\ \frac{2}{\pi(\alpha-2)} \arctan \left(\beta \tan \frac{(\alpha-2)\pi}{2} \right), & \text{se } 1 < \alpha < 2 \end{cases} \quad (4.3)$$

a f.d.p. proposta, para $1 < \alpha < 2$, é

$$f \left(x; \alpha, \tilde{\sigma}, \tilde{\beta} \right) = \frac{1}{\alpha \tilde{\sigma}} \sum_{\nu=0}^{\infty} \frac{\Gamma \left(\frac{1}{\alpha} + \frac{1}{\alpha} \nu \right) (-1)^\nu \left(\frac{x}{\tilde{\sigma}} \right)^\nu}{\nu! \Gamma \left[\left(\frac{1+\tilde{\beta}}{2} - \frac{\tilde{\beta}}{\alpha} \right) \left(1 + \nu \right) \right] \Gamma \left[\frac{1-\tilde{\beta}}{2} + \frac{\tilde{\beta}}{\alpha} - \left(\frac{1+\tilde{\beta}}{2} - \frac{\tilde{\beta}}{\alpha} \right) \nu \right]}, \quad (4.4)$$

e, no caso de $1 < \alpha < 2$, a f.d.p. é

$$f \left(x; \alpha, \tilde{\sigma}, \tilde{\beta} \right) = \frac{1}{\alpha \tilde{\sigma}} \sum_{\nu=0}^{\infty} \frac{\Gamma \left(1 + \frac{1}{\alpha} \nu \right) (-1)^\nu \left(\frac{x}{\tilde{\sigma}} \right)^{-1-\frac{\nu}{\alpha}}}{\nu! \Gamma \left[\frac{1-\tilde{\beta}}{2} - \frac{1+\tilde{\beta}}{2} \left(1 + \frac{\nu}{\alpha} \right) \right] \Gamma \left[\frac{1+\tilde{\beta}}{2} - \left(1 + \frac{\nu}{\alpha} \right) \right]}. \quad (4.5)$$

As densidades (4.4) e (4.5) podem ser escritas como duas séries específicas. A série $h(x)$ para gerar os valores centrais da densidade, e a série $c(x)$ para gerar, as de caudas. Essas séries são definidas em quatro situações, que dependem se $0 < \alpha < 1$ ou $1 < \alpha < 2$ e se $x < \mu$ ou $x \geq \mu$, como a seguir.

1. Se $\alpha > 1$ e $x < \mu$

• centro

$$h(x) = \sum_{\nu=0}^{v_1} \frac{\sin\left(\frac{1}{2} \frac{\pi(\nu+1)(\alpha-\tilde{\beta}\alpha+2\tilde{\beta})}{\alpha}\right) \Gamma\left(\frac{\nu+1}{\alpha}\right) \left(\frac{x-\nu}{\tilde{\sigma}}\right)^{\nu}}{\alpha\tilde{\beta}\pi\Gamma(\nu+1)} \quad (4.6)$$

• cauda

$$c(x) = \sum_{\nu=0}^{v_2} \frac{\sin\left(\frac{1}{2}\pi(2+\nu\alpha+\tilde{\beta}\nu\alpha)\right) (-1)^{\nu} \left(\frac{\tilde{\sigma}}{x-\mu}\right)^{\nu\alpha} \Gamma(1+\nu\alpha)}{-(x-\mu)\pi\Gamma(\nu+1)} \quad (4.7)$$

2. Se $\alpha > 1$ e $x \geq \mu$

• centro

$$h(x) = \sum_{\nu=0}^{v_1} \frac{\sin\left(\frac{1}{2} \frac{\pi(\nu+1)(\alpha+\tilde{\beta}\alpha-2\tilde{\beta})}{\alpha}\right) \Gamma\left(\frac{\nu+1}{\alpha}\right) \left(-\frac{x-\nu}{\tilde{\sigma}}\right)^{\nu}}{\alpha\tilde{\beta}\pi\Gamma(\nu+1)} \quad (4.8)$$

• cauda

$$c(x) = \sum_{\nu=0}^{v_2} \frac{\sin\left(\frac{1}{2}\pi(2+\nu\alpha+\tilde{\beta}\nu\alpha-2\nu\tilde{\beta})\right) (-1)^{\nu} \left(\frac{\tilde{\sigma}}{x-\mu}\right)^{\nu\alpha} \Gamma(1+\nu\alpha)}{(x-\mu)\pi\Gamma(\nu+1)} \quad (4.9)$$

3. Se $\alpha < 1$ e $x < \mu$

• centro

$$h(x) = \sum_{\nu=0}^{v_2} \frac{\sin\left(\frac{1}{2}\pi(1-\tilde{\beta}+\nu-\nu\tilde{\beta})\right) \Gamma\left(\frac{\nu+1}{\alpha}\right) \left(\frac{x-\mu}{\tilde{\sigma}}\right)^{\nu}}{\alpha\tilde{\beta}\pi\Gamma(\nu+1)} \quad (4.10)$$

• cauda

$$c(x) = \sum_{\nu=0}^{v_1} \frac{\sin\left(\frac{1}{2}\pi(2+\nu\alpha-\tilde{\beta}\nu\alpha)\right) (-1)^{\nu} \left(\frac{\tilde{\sigma}}{x-\mu}\right)^{\nu\alpha} \Gamma(1+\nu\alpha)}{-(x-\mu)\pi\Gamma(\nu+1)} \quad (4.11)$$

4. Se $\alpha < 1$ e $x \geq \mu$

- centro

$$h(x) = \sum_{\nu=0}^{v_2} \frac{\sin\left(\frac{1}{2}\pi(1 + \tilde{\beta} + \nu + \nu\tilde{\beta})\right) \Gamma\left(\frac{\nu+1}{\alpha}\right) \left(-\frac{x-\mu}{\tilde{\sigma}}\right)^{\nu}}{\alpha\tilde{\beta}\pi\Gamma(\nu+1)} \quad (4.12)$$

- cauda

$$c(x) = \sum_{\nu=0}^{v_1} \frac{\sin\left(\frac{1}{2}\pi(2 + \nu\alpha + \tilde{\beta}\nu\alpha)\right) (-1)^{\nu} \left(\frac{\tilde{\sigma}}{x-\mu}\right)^{\nu\alpha} \Gamma(1 + \nu\alpha)}{(x - \mu)\pi\Gamma(\nu+1)} \quad (4.13)$$

Ao utilizar esse procedimento de cálculo no algoritmo EM, a densidade de x é o menor resultado gerado pelas séries $h(x)$ e $c(x)$. Neste trabalho, os argumentos ν_1 e ν_2 , definidos de modo empírico, assumem os valores 100 e 10, respectivamente.

4.3.3 Comentários

O Algoritmo I é de simples implementação computacional. Entretanto, ao ser aplicado em valores extremos, gera valores que assumem um comportamento oscilatório. Esse comportamento pode resultar em valores negativos, que no caso do método EM impossibilita o uso desses valores. Um modo de contornar esse problema é aplicar a Propriedade 1.2.3, a dificuldade encontrada nessa solução é determinar a partir de quando a aproximação deve ser utilizada.

O Algoritmo II é de simples implementação computacional. Aplicado em valores extremos, os valores gerados por $h(x)$ não oscilam, crescem. Esse comportamento facilita determinar o momento que deve ser usado os valores de $c(x)$, para aproximar as caudas.

A Figura 4.1 mostra o log da densidade gerada pelo Algoritmo I, linha pontilhada, e, pelo Algoritmo II, linha contínua, para $\alpha \in \{0.5, 1.1, 1.5, 1.9\}$. Nota-se que, no centro, as alternativas são equivalentes, mas nas caudas, a segunda não oscila.

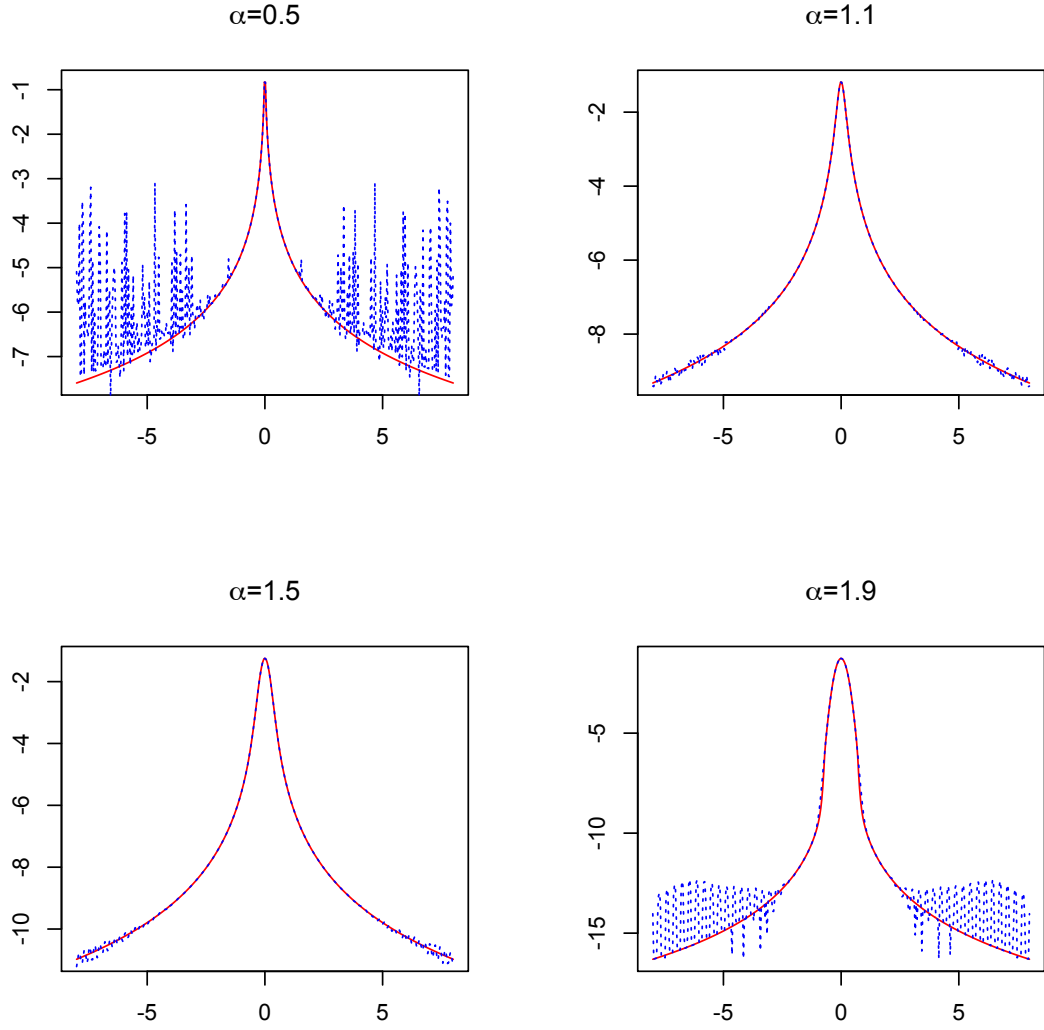


Figura 4.1: Log das densidades geradas pelo Algoritmo I, linha pontilhada, e, pelo Algoritmo II, linha contínua, para $\alpha \in \{0.5, 1.1, 1.5, 1.9\}$.

4.3.4 Simulação de uma Mistura Finita

A fim de gerar amostras de valores aleatórios de uma mistura finita, utilizou-se o algoritmo a seguir, simples de ser implementado, mas que exige ser possível gerar valores aleatórios para as componentes da mistura.

O procedimento para uma mistura $H(x)$, em (1.17), de K componentes F_j , $j = 1, \dots, K$, e pesos p_j , tais que $\sum_{j=1}^K p_j = 1$, é o seguinte:

1. para $i = 1, \dots, N$, gere $u_i \sim U(0, 1)$;

2. se $u_i \leq p_1$, gere $H_i \sim F_1$;
3. para $q = 2, \dots, K$, se $\sum_{j=1}^{q-1} p_j < u_i \leq \sum_{j=1}^q p_j$, gere $H_i \sim F_q$;
4. se $u_i \geq \sum_{j=1}^K p_j$, gere $H_i \sim F_K$;

Uma alternativa ao algoritmo acima, caso seja fácil simular valores inteiros $i = 1, \dots, K$, com probabilidade de escolha p_i , é o seguinte:

1. gere uma amostra de tamanho N dos valores de $i = 1, \dots, K$, com reposição e probabilidade de escolha p_i ;
2. faça m_i a frequência de cada valor de i na amostra;
3. simule de $X \sim F_i$ uma amostra de tamanho m_i .

A fim de gerar valores aleatórios de $X \sim S_\alpha(\sigma, \beta, \mu)$, em (1.1), usa-se a proposição seguinte para gerar $Z \sim S_\alpha(1, \beta, 0)$, padrão ($\mu = 0, \sigma = 1$).

Proposição 4.3.1 (Chambers *et al.* (1976)). *Seja Y uniformemente distribuída no intervalo $(-\frac{\pi}{2}, \frac{\pi}{2})$ e W é exponencialmente distribuída com média 1 (um) e $0 < \alpha < 1$. A variável aleatória simétrica*

$$Z = \begin{cases} \frac{\sin(\alpha Y)}{\cos(Y)^{\frac{1}{\alpha}}} \left[\frac{\cos((\alpha-1)Y)}{W} \right]^{\frac{1-\alpha}{\alpha}}, & \text{se } \alpha = 1 \\ \tan(Y), & \text{se } \alpha = 1. \end{cases},$$

tem distribuição $S_\alpha(1, 0, 0)$.

No caso de assimetria, para $-1 \leq \beta \leq 1$, defina

$$\theta_0 = \frac{\arctan\left(\beta \tan\left(\frac{\pi\alpha}{2}\right)\right)}{\alpha}.$$

Então,

$$Z = \begin{cases} \frac{\sin(\alpha(\theta_0+Y))}{(\cos(\alpha\theta_0)\cos(Y))^{\frac{1}{\alpha}}} \left[\frac{\cos(\alpha\theta_0+(\alpha-1)Y)}{W} \right]^{\frac{1-\alpha}{\alpha}}, & \text{se } \alpha = 1 \\ \frac{2}{\pi} \left[\left(\frac{\pi}{2} + \beta Y \right) \tan(Y) - \beta \log\left(\frac{\frac{\pi}{2} W \cos(Y)}{\frac{\pi}{2} + \beta Y} \right) \right], & \text{se } \alpha = 1. \end{cases}$$

tem distribuição $S_\alpha(1, \beta, 0)$.

Uma vez gerados os valores para $Z \sim S_\alpha(1, \beta, 0)$, pode-se simular $X \sim S_\alpha(\sigma, \beta, \mu)$ por meio das relações

$$X = \begin{cases} \sigma Z + \mu & , \text{se } \alpha \neq 1 \\ \sigma Z + \mu + \beta \frac{2}{\pi} \sigma \ln(\sigma) & , \text{se } \alpha = 1. \end{cases}$$

Para gerar valores aleatórios de $X^0 \sim S_\alpha^0(\mu, \beta, \sigma)$, em (1.7), calcule $Z \sim S_\alpha(1, \beta, 0)$, na Proposição 4.3.1, e use as relações

$$X^0 = \begin{cases} \sigma (Z - \beta \frac{2}{\pi} \sigma \ln(\sigma)) + \mu & , \text{se } \alpha \neq 1 \\ \sigma Z + \mu & , \text{se } \alpha = 1. \end{cases}$$

4.4 Influência de b na Matriz $\Lambda^{-1}\Sigma\Lambda^{-1}$ na Estimação EFC

Xu e Knigh (2011) apresenta que as variâncias assintóticas dos estimadores EFC do parâmetro $\Theta = (\mu, \sigma^2)$ para uma única distribuição Normal podem ser escritas como

$$var(\hat{\mu}) = \frac{\sigma^2}{n} \left(\frac{b^2 + 2b\sigma^2 + \sigma^4}{b^2 + 2b\sigma^2 + 0.75\sigma^4} \right)^{\frac{3}{2}} \quad (4.14)$$

$$var(\hat{\sigma}^2) = \frac{16(b + \sigma^2)^2}{9n} \left(\frac{16(b + \sigma^2)^2(b^2 + 2b\sigma^2 + \sigma^4)}{(4b^2 + 8b\sigma^2 + 3\sigma^4)^{\frac{5}{2}}} - 1 \right) \quad (4.15)$$

e $cov(\hat{\mu}, \hat{\sigma}) = 0$, pela Proposição 2.3.1.

A Proposição 2.3.3 expressa os elementos da matriz $\Lambda^{-1}\Sigma\Lambda^{-1}$ de

$$\sqrt{n} (\tilde{\Theta}_n - \Theta) \rightarrow N [0; \Lambda^{-1}\Sigma\Lambda^{-1}]$$

em forma de esperança matemática

$$\Lambda_{ij} = E \left(\frac{\partial D_n^2(\Theta; x)}{\partial \Theta_i \partial \Theta_j} \right) \quad (4.16)$$

e

$$\Sigma_{ij} = E \left(\frac{\partial D_n(\Theta; x)}{\partial \Theta_i} \frac{\partial D_n(\Theta; x)}{\partial \Theta_j} \right). \quad (4.17)$$

onde Θ_i e Θ_j são os elementos i, j do vetor de parâmetros Θ .

Desse modo, é possível estimar numericamente matriz de covariância assintótica $\Lambda^{-1}\Sigma\Lambda^{-1}$ quando há uma forma fechada para $D_n(\Theta; x)$. Essa estimativa pode ser calculada da seguinte maneira:

1. Determine os elementos da matriz do produto das primeiras derivadas.
2. Determine os elementos da matriz das segundas derivadas.
3. Gere um número, J, grande de amostras.
4. Para cada amostra determine os valores em 1 e 2.
5. A média desses valores determina as estimativas das matrizes $\hat{\Sigma}$ e $\hat{\Lambda}$.

A fim de experimentar o algoritmo acima, para os estimadores EFC de uma Normal $(1, 2^2)$, para $b = 1$, obteve-se a função distância objetiva (2.11) para $K = 1$, assim

$$\begin{aligned}
D_n(\Theta, x) &= \frac{1}{100^2} \sqrt{\frac{\pi}{b}} \sum_{i=1}^{100} \sum_{j=1}^{100} \exp\left(-\frac{1}{4b}(x_i - x_j)^2\right) \\
&+ \sqrt{\frac{\pi}{b + \sigma^2}} \\
&- \frac{2}{100} \sqrt{\frac{\pi}{0.5\sigma^2 + b}} \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu)^2}{4b + 2\sigma^2}\right), \tag{4.18}
\end{aligned}$$

e calculou-se por (4.14) e (4.15) a variância assintótica $var(\hat{\mu}) = 0.05195664$ e $var(\hat{\sigma}) = 0.4627235$. Depois, com 10.000 amostras de tamanho 100, estimou-se $var(\hat{\mu}) = 0.05162112$, $var(\hat{\sigma}^2) = 0.4629334$ e $cov(\hat{\mu}, \hat{\sigma}) = 0.003286041$ pelo algoritmo acima. As estimativas das $var(\hat{\mu})$ e $var(\hat{\sigma})$ foram bem próximas, com as diferenças surgindo a partir da terceira casa decimal. A $cov(\hat{\mu}, \hat{\sigma})$ obtida pelo algoritmo diminui com o aumento das amostras usadas no algoritmo.

A Figura 4.2 mostra a influência que o valor de b tem na variação conjunta, $var(\hat{\mu}) + var(\hat{\sigma})$, dos estimadores EFC. Percebe-se que a variância conjunta tende a 0.36, à medida que b cresce, e esse valor é menor que 0.52 quando $b = 1$.

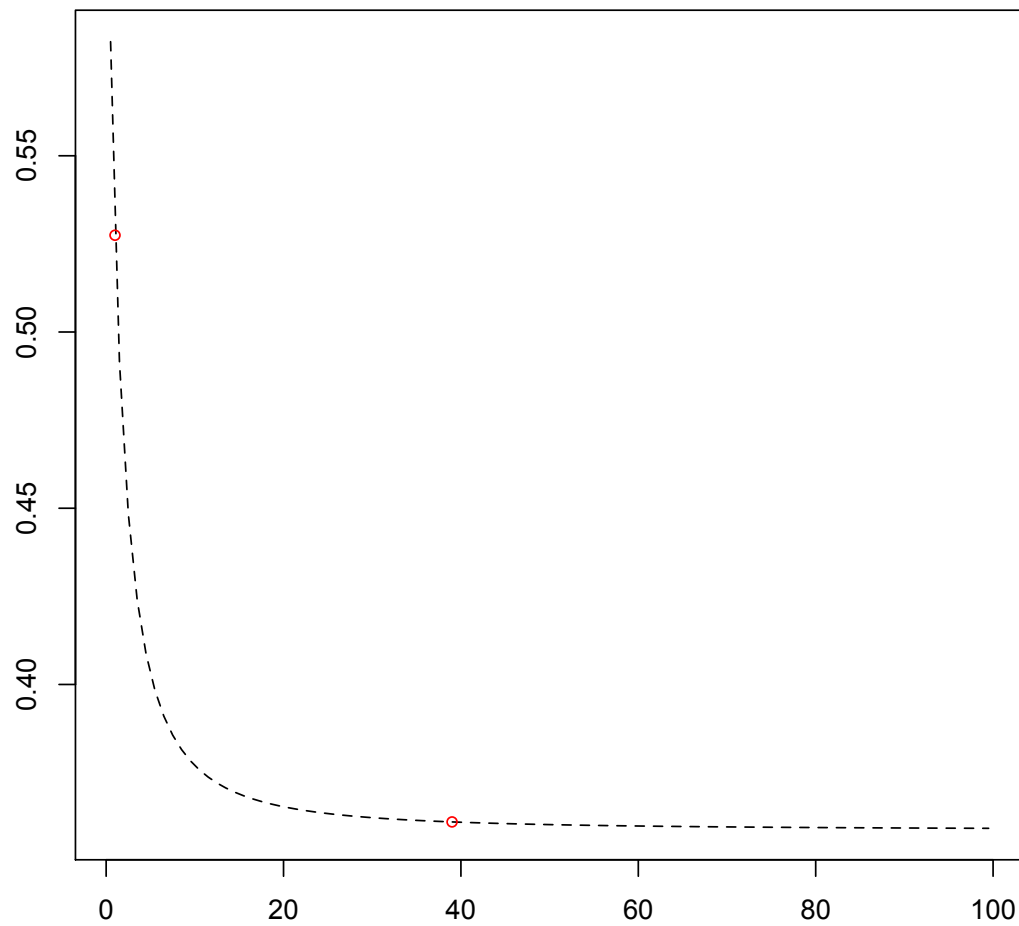


Figura 4.2: Variação conjunta dos estimadores EFC para uma $N(\mu, \sigma^2)$ em função do valor de b .

Capítulo 5

Conclusões

O algoritmo EM para misturas de densidades estáveis é mais lento conforme o valor de α diminui e o tamanho da amostra aumenta, precisando de mais iterações para alcançar os critérios de convergência. Isso pode ocorrer devido a maior presença de valores extremos.

Outra dificuldade encontrada ao aplicar o algoritmo EM é o cálculo das densidades de variáveis aleatórias estáveis pelo Algoritmo I. Em testes, esse método de cálculo, além de gerar densidades que oscilam nas caudas, em alguns casos apresentou valores de densidades negativos para valores extremos. Além disso, a aproximação para valores extremos na Propriedade 1.2.3 não é fácil de ser aplicada. Uma solução é utilizar o menor valor gerado pelas duas séries proposta em Otiniano *et al.* (2011), como um método de cálculo para densidades. Essa solução também tem a vantagem de não envolver a operação de integração, apesar de usar duas somas.

A virtude do algoritmo EM para estimar os parâmetros de uma mistura finita de distribuições estáveis é a sua generalidade, isto é, pode ser aplicado a qualquer mistura finita dessas distribuições. As dificuldades a serem enfrentadas são as otimizações e a escolha dos valores iniciais, principalmente para α e β .

Para trabalhos futuros, em relação ao algoritmo EM, propõe-se estabelecer procedimentos para determinar valores iniciais para os parâmetros, com foco em α e β . Verificar o desempenho desse método de estimação para misturas estáveis quaisquer. Também, é importante testar, de modo mais exaustivo, os procedimentos de cálculo de densidades em Otiniano *et al.* (2011), visto o bom desempenho que apresentou no

caso de distribuições estáveis simétricas.

O método de estimação EFC, para modelar misturas de densidades simétricas, envolve séries e mostrou-se de difícil otimização.

Outra dificuldade é determinar de forma analítica a matriz de covariância $\Lambda^{-1}\Sigma\Lambda^{-1}$ da Proposição 2.3.2, pois as operações de derivação e integração tornam-se cada vez mais trabalhosas à medida que o número K de componentes da mistura cresce.

Como virtudes dos estimadores EFC citam-se: menor custo computacional comparado com o EM; a consistência e normalidade assintótica; possibilidade de melhorar a eficiência dos estimadores por meio de uma escolha ótima para $b > 0$ para $w(t) = e^{-bt^2}$, para a mistura de normais, e para $w(t) = e^{-b|t|}$ no caso de misturas de Cauchy; simples implementação computacional; e a existência de procedimento numérico de estimação para matriz de covariância $\Lambda^{-1}\Sigma\Lambda^{-1}$ da Proposição 2.3.2.

Como continuação deste trabalho, com foco no processo de estimação EFC, sugere-se otimizar as formas de $D_n(\Theta; x)$, em (2.21) e (2.22), que foram definidas para estimar os parâmetros de misturas finitas de duas distribuições simétricas, cujos valores para α são constantes e, generalizá-las.

Referências Bibliográficas

- [1] Casarin, R.. Bayesian Inference for Mixture of Stable Distributions. Working paper n. 0428 CEREMADE. University Paris IX, (2004).
- [2] Chambers, J.M., Mallows, C.L., Stuck, B.W.. A method for simulating stable random variables. Journal of the American Statistical Association. **71** , 340-344, (1976).
- [3] Dempster, A. P., Laird, N. M., Rubin, D. B.. Maximum likelihood estimation from incomplete data via EM algorithm (with discussion). Journal of the Royal Statistical Society. Série **B**, 39, 1-38, (1977).
- [4] Dorea, C. C. Y., Otiniano, C. E. G., Matsushita, R., Rathie, P. N.. Lévy flight approximations for scaled transformations of random walks. Computational Statistics Data Analysis. Volume 51, Issue **12**, 6343-6354, (2007).
- [5] DuMouchel, W. H.. Stable Distributions in Statistical Inference. Ph. D. dissertation, Department of Statistics, Yale University, (1971).
- [6] Fama, E. F., Roll R.. Some Properties of Symmetric Stable Distributions. Journal of the American Statistical Association. **63**, (323). 817-836, (1968).
- [7] Hatzinakos D., How J.. Applications of Empirical Characteristic Functions to Estimation and Detection Problems. Signal Process. **65** (**2**), 199-219, (1988).
- [8] Heathcote, C. R.. Integrated Mean Square Error Estimation of Parameters. Biometrika. **64**, 255-264, (1977).
- [9] Hill, B. M.. A simple general approach to inference about the tail of distribution. Annals of Statistics. **3**, 1163-1174, (1975).

- [10] Koutrouvelis I. A.. Regression-Type Estimation of the Parameters of Stable Laws
Journal of the American Statistical Association. **75**, n. 372, (1980).
- [11] McLachlan, G. e Peel D..Finite Mixture Models. New York. Wiley, (2000),
- [12] Nolan, J. P.. Stable Distributions: Models for Heavy Data. Boston. Birkhauser (2011). Em edição, Capítulo 1 disponível em <http://academic2.american.edu/~jpnolan/stable/stable.html>
- [13] Nolan, J. P.. An algorithm for evaluating stable in Zolotarev's (M) parametrization. Mathematical and Computer Modelling. **29**, 229-233, (1999).
- [14] Nolan, J. P.. Numerical calculation of stable densities and distribution functions. Communications in Statistics - Stochastic Models. **13**, 759-774, (1977).
- [15] Otiniano, C.E. G., Rathie, P. N., Sousa, T. R., Dorea, C.C.Y.. Stable Lévy density: convolution and numerical illustrations. Mathematical and Computer Modelling. Preprint, janeiro (2011).
- [16] Paulson, A. S., Holcomb E. W., The Estimation of the Parameters of the Stable Laws. Biometrika **62** (1), 163-170, (1975).
- [17] Press, S. J.. Estimation in Univariate and Multivariate Stable Distributions. Journal of the American Statistical Association. **67**, 842-846, (1972).
- [18] Ravishanker, N. e Dey, D. K.. Multivariate Survival Models with a Mixture of Positive Stable Frailties. Methodology And Computing In Applied Probability, **2:3**, 293-308. (2000).
- [19] Salas-Gonzalez, D., Kuruoglu, E. E. e Ruiz, D. P.. Finite mixture of α -stable distributions. Digital Signal Processing **19**, 250-264, (2009).
- [20] Samorodnitsky, G. e ,Taqqu, M. S.. Stable Non-Gaussian Random Processes. Chapman & Hall (1994).
- [21] Saraiva, E. F. Modelo de Mistura com Número de Componentes Desconhecido: Estimação via Método Split-Merge. Tese de Doutorado, Departamento de Estatística, Universidade Federal de São Carlos. (2009)

- [22] Swami A.. Non-Gaussian mixture models for detection and estimation in heavy-tailed noise. IEEE International Conference on Acoustics, Speech, and Signal Processing. Istambul, Turkey, June 2000.
- [23] Titterington D. M., Smith A. F. M., e Makov U. E.. Statistical Analysis of Finite Mixture Models. Wiley, (1985).
- [24] Xu, D. e Knight, J. L.. Continuous Empirical Characteristic Function Estimation of Mixtures of Normal Parameters. Econometric Reviews. **30 (1)**, 25-50, (2011).
- [25] Yu, J.. Empirical Characteristic Function Estimation and Its Applications. Econometric Reviews. **23**, 93-123, (2004).
- [26] Zolotarev, V. M.. One-dimensional stable distributions. Amer. Math. Soc. Transl. of Math. Monographs, Vol. **65**, (Translation of the original 1983 Russian), Amer. Math. Soc., Providence, RI, (1986).

Apêndice A

Tabelas

Os estimadores EFC e EM foram aplicados a dados simulados como forma de se observar os seus desempenhos, mas apenas as médias das 1.000 (mil) estimativas dos parâmetro obtidas em cada uma das simulações realizadas foram apresentadas. Neste apêndice são apresentados o erro quadrático médio (EQM) e o viés associados as estimativas obtidas.

Para cada modelo simulados foram geradas 1.000 (mil) amostras de determinado tamanho, e para cada amostra foram obtidas as estimativas e interesse. No contexto de simulação os valores procurados são conhecidos, assim toma-se

$$EQM = \sum_{i=1}^{\nu} \left(\theta - \hat{\theta}_i \right)^2 \quad (A.1)$$

e

$$viés = \sum_{i=1}^{\nu} \left(\theta - \hat{\theta}_i \right), \quad (A.2)$$

onde ν é o número de de simulações realizadas, θ é o valor conhecido do parâmetro e $\hat{\theta}$ é o valor estimado do parâmetro.

A.1 Tabelas citadas no Capítulo 2

As Tabelas A.1 e A.2 apresentam o EQM e o viés associados as estimativas EFC do experimento descrito na Tabela 2.1, pag. 33, no Capítulo 2.

Tabela A.1: EQM das estimativas EFC para componentes Cauchy.

Experimento	p_1	μ_1	σ_1	μ_2	σ_2
A1	0.00518	0.04444	0.07164	0.05468	0.06986
A2	0.00240	0.02377	0.03488	0.02545	0.03275
B1	0.00878	0.06021	0.08312	0.06543	0.08899
B2	0.01520	0.04080	0.05692	0.44283	0.30917
C1	0.00486	0.05525	0.06991	0.04826	0.07464
C2	0.00343	0.04733	0.06281	0.04388	0.05699
D1	0.05037	4.92619	15.95894	12.42316	17.79740
D2	0.02070	9.98003	55.83803	9.33180	52.36811

Tabela A.2: Viés das estimativas EFC para componentes Cauchy.

Experimento	p_1	μ_1	σ_1	μ_2	σ_2
A1	0.00053	0.00402	-0.00322	0.00464	0.00435
A2	-0.00024	-0.00125	0.00765	0.00210	-0.00648
B1	0.00690	0.00744	0.00115	-0.01602	-0.02237
B2	-0.02995	-0.02182	-0.04812	-0.12285	0.05853
C1	0.00256	-0.00014	0.00789	0.00745	-0.01795
C2	0.00052	-0.00260	0.01660	-0.00906	-0.00709
D1	0.01771	-0.87086	2.26692	0.95838	-1.84282
D2	0.00287	-1.14071	4.66872	1.08797	-4.34603

A Tabela A.3 apresenta em forma compacta o valor real dos parâmetros, a média das estimativas, o EQM e o viés para os casos C1 e C2 descritos na Tabela 2.1, pag. 33.

Tabela A.3: Valor do parâmetro, média , EQM e viés das estimativas EFC para os casos C1 e C2.

Estatística	Caso	p_1	μ_1	σ_1	μ_2	σ_2
Valores	C1	0.5	0.0	1.0	6.0	1.0
	C2	0.5	0.0	1.0	10.0	1.0
Média	C1	0.5027	-0.0001	1.0079	6.0075	0.9821
	C2	0.5005	-0.0026	1.0166	9.9909	0.9929
EQM	C1	0.0049	0.0553	0.0699	0.0483	0.0746
	C2	0.00343	0.0473	0.0628	0.0439	0.0570
Viés	C1	0.0026	-0.0001	0.0079	0.0075	-0.0180
	C2	0.0005	-0.0026	0.0166	-0.0091	-0.0071

A Figura A.1 mostra as densidades, considerando as estimativas médias, e os histogramas os valores das 1.000 (mil) amostras dos casos C1 e C2.

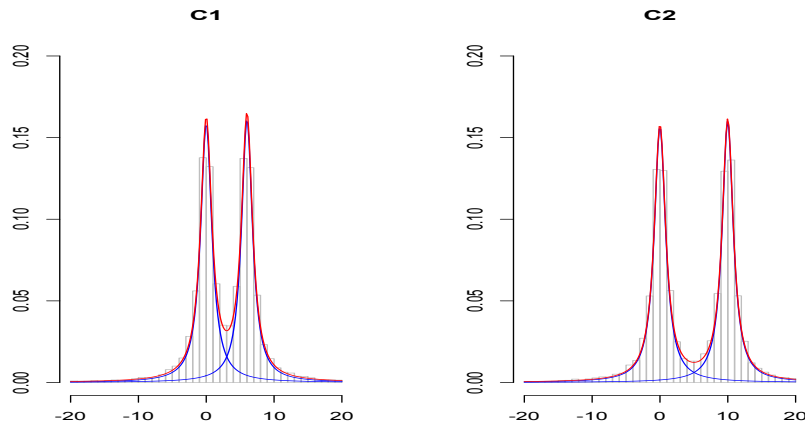


Figura A.1: Densidades das misturas de Cauchy para os caso C1 e C2

A.2 Tabelas citadas no Capítulo 3

As Tabelas A.4 e A.5 apresentam o EQM e o viés das estimativas do experimento de simulação descrito na Tabela 3.2, pag. 42, no Capítulo 3.

Tabela A.4: EQM das estimativas EM para misturas finitas de variáveis aleatórias estáveis.

Experimento	p	α	μ_1	μ_2	σ_1	σ_2
A	0.00144	0.00078	0.02521	0.02845	0.00785	0.00630
B	0.00137	0.00611	0.02331	0.02842	0.00896	0.00933
C	0.00165	0.00532	0.02284	0.02821	0.00878	0.00820
D	0.00140	0.00456	0.02071	0.01798	0.01341	0.01178
E	0.00152	0.00386	0.01019	0.01536	0.01394	0.01606
F	0.00163	0.00223	0.00443	0.00825	0.02013	0.02409
G	0.00234	0.00096	0.00857	0.00257	0.02903	0.02490

Tabela A.5: Viés das estimativas EM para misturas finitas de variáveis aleatórias estáveis.

Experimento	p	α	μ_1	μ_2	σ_1	σ_2
A	-0.00140	-0.01014	-0.00380	0.01856	-0.00535	-0.01258
B	-0.00087	0.02035	0.00152	0.00830	-0.00933	-0.02956
C	0.00332	0.00558	-0.02056	-0.00449	-0.02199	-0.00497
D	0.00283	0.01429	0.00604	0.00051	-0.00233	-0.01435
E	-0.00461	0.00096	-0.01940	-0.01699	-0.01009	0.00026
F	0.00275	0.00280	0.00119	0.01236	-0.00910	0.00507
G	0.00314	-0.00705	-0.00071	0.00181	-0.00577	-0.06234

Apêndice B

Cálculo de D_n para componentes

$$N(\mu, \sigma^2)$$

Demonstração da Proposição 2.4.1

parte I

A primeira parte da demonstração consiste em transformar

$$|\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2$$

em uma soma de 4 parcelas. Assim, com

$$|\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 = [U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2$$

onde

$$U_n - u(t; \theta) = \frac{1}{n} \sum_{j=1}^n \cos(tx_j) - \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2}$$

e

$$V_n - v(t; \theta) = \frac{1}{n} \sum_{j=1}^n \sin(tx_j) - \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2}.$$

Então,

$$\begin{aligned}
[U_n - u(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \cos(tx_j) \right]^2 \\
&- \frac{2}{n} \left[\sum_{j=1}^n \cos(tx_j) \times \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right] \\
&+ \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right]^2
\end{aligned}$$

e

$$\begin{aligned}
[V_n - v(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \sin(tx_j) \right]^2 \\
&- \frac{2}{n} \left[\sum_{j=1}^n \sin(tx_j) \times \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right] \\
&+ \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right]^2.
\end{aligned}$$

Daí, obtemos que $[U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2$ é igual a

$$\begin{aligned}
&\frac{1}{n^2} \left[\left(\sum_{j=1}^n \cos(tx_j) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] \\
&- \frac{2}{n} \sum_{k=1}^K p_k e^{-\frac{1}{2}\sigma_k^2 t^2} \left[\cos(\mu_k t) \sum_{j=1}^n \cos(tx_j) - \sin(\mu_k t) \sum_{j=1}^n \sin(tx_j) \right] \\
&+ \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right]^2 \tag{B.1}
\end{aligned}$$

$$+ \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\frac{1}{2}\sigma_k^2 t^2} \right]^2. \tag{B.2}$$

Expandindo (B.1) e (B.2), temos que

$$\begin{aligned}
(B.1) &= p_1^2 \cos^2(\mu_1) e^{-\sigma_1^2 t^2} + \dots + p_K^2 \cos^2(\mu_K) e^{-\sigma_K^2 t^2} \\
&+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \cos(\mu_k t) \cos(\mu_j t) e^{-\frac{1}{2}(\sigma_k^2 + \sigma_j^2) t^2}
\end{aligned}$$

e

$$\begin{aligned}
(B.2) &= p_1^2 \sin^2(\mu_1) e^{-\sigma_1^2 t^2} + \dots + p_K^2 \sin^2(\mu_K) e^{-\sigma_K^2 t^2} \\
&+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \sin(\mu_k t) \sin(\mu_j t) e^{-\frac{1}{2}(\sigma_k^2 + \sigma_j^2) t^2}.
\end{aligned}$$

Assim, adicionando (B.1) a (B.2) temos

$$\begin{aligned}
& p_1^2 e^{-\sigma_1^2 t^2} \left[\cos^2(\mu_1 t) + \sin^2(\mu_1 t) \right] + \dots + p_K^2 e^{-\sigma_K^2 t^2} \left[\cos^2(\mu_K t) + \sin^2(\mu_K t) \right] \\
& + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-\frac{1}{2}(\sigma_k^2 + \sigma_j^2)t^2} \left[\cos(\mu_k t) \cos(\mu_j t) + \sin(\mu_k t) \sin(\mu_j t) \right] \\
& = \sum_{k=1}^K p_k^2 e^{-\sigma_k^2 t^2} + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-\frac{1}{2}(\sigma_k^2 + \sigma_j^2)t^2} \left[\cos(\mu_k t - \mu_j t) \right].
\end{aligned}$$

Finalmente, temos que $|\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2$ é

$$\begin{aligned}
& \frac{1}{n^2} \left[\left(\sum_{j=1}^n \cos(tx_j) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] \\
& + \sum_{k=1}^K p_k^2 e^{-\sigma_k^2 t^2} \\
& + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-\frac{1}{2}(\sigma_k^2 + \sigma_j^2)t^2} \left[\cos(\mu_k t - \mu_j t) \right] \\
& - \frac{2}{n} \sum_{k=1}^K p_k e^{-\frac{1}{2}\sigma_k^2 t^2} \left[\cos(\mu_k t) \sum_{j=1}^n \cos(tx_j) - \sin(\mu_k t) \sum_{j=1}^n \sin(tx_j) \right]. \quad (\text{B.3})
\end{aligned}$$

Demonstração Parte II

A segunda parte da demonstração consiste em integrar (de $-\infty$ a ∞) cada parcela do resultado final da parte I, multiplicada pela função $w(t) = \exp(-bt^2)$.

Cálculo Auxiliar

A integral

$$\int_{-\infty}^{\infty} \exp(-c_1 t^2) \cos(c_2 t) \exp(-c_3 t^2) dt = \sqrt{\frac{\pi}{c_1 + c_3}} \times \exp\left(-\frac{c_2^2}{4(c_1 + c_3)}\right) \quad (\text{B.4})$$

é a forma geral das quatro que deverão ser calculadas, assim esse resultado pode ser usado nas demais. Para o cálculo da integral acima, observe que o lado esquerdo de

(B.4) pode ser expresso como

$$\underbrace{\int_{-\infty}^{\infty} \exp(-(c_1 + c_3)t^2) \left[\frac{\exp(ic_2) + \exp(-ic_2)}{2} \right] dt}_{I} + \underbrace{\frac{1}{2} \int_{-\infty}^{\infty} \exp(-(c_1 + c_3)t^2 + ic_2t) dt + \frac{1}{2} \int_{-\infty}^{\infty} \exp(-(c_1 + c_3)t^2 - ic_2t) dt}_{II}$$

Logo, com

$$\begin{aligned} I &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left(t^2 - \frac{ic_2t}{(c_1 + c_3)} \right) \right\} dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[\left(t - \frac{ic_2}{2(c_1 + c_3)} \right)^2 + \left(\frac{ic_2}{2(c_1 + c_3)} \right)^2 \right] \right\} dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[t - \frac{ic_2}{2(c_1 + c_3)} \right]^2 \right\} dt \times \exp \left[\left(\frac{ic_2}{2(c_1 + c_3)} \right)^2 \right] \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[t - \frac{ic_2}{2(c_1 + c_3)} \right]^2 \right\} dt \times \exp \left(-\frac{c_2^2}{4(c_1 + c_3)} \right) \\ &= \frac{1}{2} \sqrt{\frac{\pi}{c_1 + c_3}} \times \exp \left(-\frac{c_2^2}{4(c_1 + c_3)} \right) \end{aligned}$$

e

$$\begin{aligned} II &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left(t^2 + \frac{ic_2t}{(c_1 + c_3)} \right) \right\} dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[\left(t + \frac{ic_2}{2(c_1 + c_3)} \right)^2 + \left(\frac{ic_2}{2(c_1 + c_3)} \right)^2 \right] \right\} dt \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[t + \frac{ic_2}{2(c_1 + c_3)} \right]^2 \right\} dt \times \exp \left[\left(\frac{ic_2}{2(c_1 + c_3)} \right)^2 \right] \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp \left\{ -(c_1 + c_3) \left[t + \frac{ic_2}{2(c_1 + c_3)} \right]^2 \right\} dt \times \exp \left(-\frac{c_2^2}{4(c_1 + c_3)} \right) \\ &= \frac{1}{2} \sqrt{\frac{\pi}{c_1 + c_3}} \times \exp \left(-\frac{c_2^2}{4(c_1 + c_3)} \right) \end{aligned}$$

Assim, finalmente, chegamos ao resultado

$$\sqrt{\frac{\pi}{c_1 + c_3}} \times \exp \left(-\frac{c_2^2}{4(c_1 + c_3)} \right).$$

Integrando B.3

A integral do primeiro termo de (B.3) é

$$\begin{aligned} & \frac{1}{n^2} \int_{-\infty}^{\infty} \left[\left(\sum_{i=1}^n \cos(tx_i) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] \times \exp(-bt^2) dt \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \int_{-\infty}^{\infty} \cos[(x_i - x_j)t] \times \exp(-bt^2) dt. \end{aligned}$$

Usando (B.4) com $c_1 = 0$, $c_2 = x_i - x_j$ e $c_3 = b$, temos o resultado

$$= \sqrt{\frac{\pi}{b}} \times \exp\left(-\frac{(x_i - x_j)^2}{4b}\right) \quad (\text{B.5})$$

A integral do segundo termo de (B.3) é

$$\int_{-\infty}^{\infty} \sum_{k=1}^K p_k^2 \exp(-\sigma_k^2 t^2) \times \exp(-bt^2) dt.$$

Usando (B.4) com $c_1 = \sigma_k^2$, $c_2 = 0$ e $c_3 = b$, temos o resultado

$$\begin{aligned} & \sqrt{\frac{\pi}{\sigma_k^2 + b}} \times \exp\left(-\frac{0}{4(\sigma_k^2 + b)}\right) \\ &= \sqrt{\frac{\pi}{\sigma_k^2 + b}}. \end{aligned} \quad (\text{B.6})$$

A integral do terceiro termo de (B.3) é

$$\begin{aligned} & \int_{-\infty}^{\infty} 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \exp\left(-\frac{1}{2}(\sigma_k^2 + \sigma_j^2)t^2\right) \left[\cos(\mu_k t - \mu_j t) \right] \times \exp(-bt^2) dt \\ &= 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}(\sigma_k^2 + \sigma_j^2)t^2\right) \left[\cos(\mu_k t - \mu_j t) \right] \times \exp(-bt^2) dt. \end{aligned}$$

Usando (B.4) com $c_1 = 0.5(\sigma_k^2 + \sigma_j^2)$, $c_2 = (\mu_k - \mu_j)$ e $c_3 = b$, temos o resultado

$$2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \sqrt{\frac{\pi}{0.5(\sigma_k^2 + \sigma_j^2) + b}} \times \exp\left(-\frac{(\mu_k - \mu_j)^2}{2(\sigma_k^2 + \sigma_j^2) + 4b}\right) \quad (\text{B.7})$$

A integral do quarto termo de (B.3) é

$$\begin{aligned} & - \int_{-\infty}^{\infty} \frac{2}{n} \sum_{k=1}^K p_k \exp\left(-\frac{1}{2}\sigma_k^2 t^2\right) \left[\cos(\mu_k t) \sum_{j=1}^n \cos(tx_j) - \sin(\mu_k t) \sum_{j=1}^n \sin(tx_j) \right] \times \exp(-bt^2) dt \\ &= -\frac{2}{n} \sum_{k=1}^K p_k \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sigma_k^2 t^2\right) \sum_{j=1}^n \cos((x_j - \mu_k)t) \times \exp(-bt^2) dt. \end{aligned}$$

Usando (B.4) com $c_1 = 0.5\sigma_k^2$, $c_2 = (x_j - \mu_k)$ e $c_3 = b$, temos o resultado

$$-\frac{2}{n} \sum_{k=1}^K p_k \sqrt{\frac{\pi}{0.5\sigma_k^2 + b}} \times \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu_k)^2}{2\sigma_k^2 + 4b}\right). \quad (\text{B.8})$$

Somando as parcelas de (B.5) a (B.8) chegamos ao resultado esperado

$$\begin{aligned} D(\theta, r) &= \int_{-\infty}^{+\infty} |\widetilde{\varphi}_n(t) - \varphi_X(t; \theta)|^2 dG(t) \\ &= \frac{1}{n^2} \sqrt{\frac{\pi}{b}} \sum_{i=1}^n \sum_{j=1}^n \exp\left(-\frac{1}{4b}(x_i - x_j)^2\right) \\ &\quad + \sum_{k=1}^K p_k^2 \sqrt{\frac{\pi}{b + \sigma_k^2}} \\ &\quad + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_h \sqrt{\frac{\pi}{b + 0.5(\sigma_k^2 + \sigma_h^2)}} \exp\left(-\frac{(\mu_k - \mu_h)^2}{4b + 2(\sigma_k^2 + \sigma_h^2)}\right) \\ &\quad - \frac{2}{n} \sum_k^K \left[p_k \sqrt{\frac{\pi}{0.5\sigma_k^2 + b}} \sum_{j=1}^n \exp\left(-\frac{(x_j - \mu_k)^2}{4b + 2\sigma_k^2}\right) \right] \end{aligned} \quad (\text{B.9})$$

Apêndice C

Cálculo de D_n para componentes

$Cauchy(\mu, \sigma)$

Demonstração da Proposição 2.5.1

Parte I

Seja X uma mistura discreta finita, cujas K componentes são distribuições $Cauchy(\mu_k, \sigma_k)$ com $k = 1, \dots, K$ e $x_1, \dots, x_n$ é uma amostra aleatória de X . Neste contexto, a função característica de X é

$$\varphi_X(t; \Theta) = \sum_{k=1}^K p_k e^{i\mu_k t - \sigma_k |t|},$$

e, a sua contra-partida amostral, a função característica empírica, é

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itx_j}.$$

Com o uso da igualdade $e^{iy} = \cos(y) + i \sin(y)$, chega-se a

$$\varphi_X(t; \Theta) = \sum_{k=1}^K \cos(\mu_k t) e^{-\sigma_k |t|} + i \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} = u(t; \theta) + iv(t; \theta)$$

e a

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n \cos(tx_j) + i \frac{1}{n} \sum_{j=1}^n \sin(tx_j) = U_n + iV_n$$

Afim de determinar D_n , a primeira ação é expressar

$$\begin{aligned} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 &= |(U_n + iV_n) - (u(t; \theta) + iv(t; \theta))|^2 \\ &= |(U_n - u(t; \theta)) + i(V_n - v(t; \theta))|^2 \\ &= [U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2. \end{aligned}$$

como parcelas de uma soma. Para isso, com

$$U_n - u(t; \theta) = \frac{1}{n} \sum_{j=1}^n \cos(tx_j) - \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|}$$

e

$$V_n - v(t; \theta) = \frac{1}{n} \sum_{j=1}^n \sin(tx_j) - \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|},$$

tem-se

$$\begin{aligned} [U_n - u(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \cos(tx_j) \right]^2 \\ &\quad - \frac{2}{n} \left[\sum_{j=1}^n \cos(tx_j) \times \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|} \right] \\ &\quad + \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|} \right]^2 \end{aligned}$$

e

$$\begin{aligned} [V_n - v(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \sin(tx_j) \right]^2 \\ &\quad - \frac{2}{n} \left[\sum_{j=1}^n \sin(tx_j) \times \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} \right] \\ &\quad + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} \right]^2 \end{aligned}$$

Disso, tem-se que

$$\begin{aligned} &[U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2 \\ &= \frac{1}{n^2} \left[\left(\sum_{j=1}^n \cos(tx_j) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] \\ &\quad - \frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k |t|} \left[\cos(\mu_k t) \cos(tx_j) + \sin(\mu_k t) \sin(tx_j) \right] \end{aligned} \quad (\text{C.1})$$

$$+ \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|} \right]^2 + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} \right]^2. \quad (\text{C.2})$$

Escrevendo os somatórios em (C.1) e (C.2) como

$$\begin{aligned} \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|} \right]^2 &= p_1^2 \cos^2(\mu_1) e^{-2\sigma_1 |t|} + \dots + p_K^2 \cos^2(\mu_K) e^{-2\sigma_K |t|} \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \cos(\mu_k t) \cos(\mu_j t) e^{-(\sigma_k + \sigma_j) |t|} \end{aligned}$$

e

$$\begin{aligned} \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} \right]^2 &= p_1^2 \sin^2(\mu_1) e^{-2\sigma_1 |t|} + \dots + p_K^2 \sin^2(\mu_K) e^{-2\sigma_K |t|} \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \sin(\mu_k t) \sin(\mu_j t) e^{-(\sigma_k + \sigma_j) |t|}, \end{aligned}$$

além disso aplicando as propriedades trigonométricas $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ e $\cos^2(x) + \sin^2(x) = 1$ obtém-se

$$\begin{aligned} &\left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k |t|} \right]^2 + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k |t|} \right]^2 \\ &= p_1^2 e^{-2\sigma_1 |t|} \left[\cos^2(\mu_1 t) + \sin^2(\mu_1 t) \right] + \dots + p_K^2 e^{-2\sigma_K |t|} \left[\cos^2(\mu_K t) + \sin^2(\mu_K t) \right] \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k + \sigma_j) |t|} \left[\cos(\mu_k t) \cos(\mu_j t) + \sin(\mu_k t) \sin(\mu_j t) \right] \\ &= \sum_{k=1}^K p_k^2 e^{-2\sigma_k |t|} + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k + \sigma_j) |t|} \left[\cos(\mu_k t - \mu_j t) \right]. \end{aligned}$$

Também escrevendo

$$\begin{aligned} &\frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k |t|} \left[\cos(\mu_k t) \cos(tx_j) + \sin(\mu_k t) \sin(tx_j) \right] \\ &= -\frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k |t|} \cos[(\mu_k - x_j)t] \end{aligned} \quad (\text{C.3})$$

e

$$\frac{1}{n^2} \left[\left(\sum_{i=1}^n \cos(tx_i) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \cos[(x_i - x_j)t]$$

Finalmente, chega-se a

$$\begin{aligned}
|\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \cos[(x_i - x_j)t] \\
&+ \sum_{k=1}^K p_k^2 e^{-2\sigma_k|t|} \\
&+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k + \sigma_j)|t|} \cos[(\mu_k - \mu_j)t] \\
&- \frac{2}{n} \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^n p_k e^{-\sigma_k|t|} \cos[(\mu_k - x_j)t]. \tag{C.4}
\end{aligned}$$

Demonstração Parte II

A segunda parte da demonstração consiste em integrar, em relação a t , de $(-\infty$ a $\infty)$ cada parcela, desenvolvida na parte I, multiplicada pela função $w(t) = \exp(-b|t|)$. Com a mudança de variável $s = |t|$ e organizando os expoentes tem-se

$$\begin{aligned}
D(\Theta, x) &= \frac{2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \int_0^\infty \cos[(x_i - x_j)t] e^{-bt} dt \\
&+ 2 \sum_{k=1}^K p_k^2 \int_0^\infty e^{-(2\sigma_k + b)t} dt \\
&+ 4 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \int_0^\infty e^{-(\sigma_k + \sigma_j + b)t} \cos[(\mu_k - \mu_j)t] dt \\
&- \frac{4}{n} \sum_{k=1}^K \sum_{j=1}^n p_k \int_0^\infty e^{-(\sigma_k + b)t} \cos[(\mu_k - x_j)t] dt \tag{C.5}
\end{aligned}$$

Usando a igualdade

$$\int_0^\infty e^{-at} \cos(bt) dt = \frac{a}{a^2 + b^2}, \tag{C.6}$$

onde $a > 0$ e $b > 0$, nas parcelas de (C.5), obtém-se o seguinte resultado

$$\begin{aligned}
D(\Theta, x) &= \frac{2}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{b}{b^2 + (x_i - x_j)^2} \\
&+ 2 \sum_{k=1}^K p_k^2 \frac{1}{2\sigma_k + b} \\
&+ 4 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \frac{\sigma_k + \sigma_j + b}{(\sigma_k + \sigma_j + b)^2 + (\mu_k - \mu_j)^2} \\
&- \frac{4}{n} \sum_{k=1}^K \sum_{j=1}^n p_k \frac{\sigma_k + b}{(\sigma_k + b)^2 + (\mu_k - x_j)^2}.
\end{aligned} \tag{C.7}$$

Apêndice D

Cálculo de D_n para componentes Estáveis Simétricas

Demonstração do Resultado (2.21)

Parte I

Seja X uma mistura discreta finita, cujas K componentes são distribuições $S_\alpha(\sigma, 0, \mu)$ com $k = 1, \dots, K$ e $x_1, \dots, x_n$. Neste contexto, a função característica de X é

$$\varphi_X(t; \Theta) = \sum_{k=1}^K p_k e^{i\mu_k t - \sigma_k^\alpha |t|^\alpha},$$

e, a sua contra-partida amostral, a função característica empírica, é

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itx_j}.$$

Com o uso da igualdade $e^{iy} = \cos(y) + i \sin(y)$, chega-se a

$$\varphi_X(t; \Theta) = \sum_{k=1}^K \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} + i \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} = u(t; \theta) + iv(t; \theta)$$

e a

$$\tilde{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n \cos(tx_j) + i \frac{1}{n} \sum_{j=1}^n \sin(tx_j) = U_n + iV_n \quad (\text{D.1})$$

A fim de determinar D_n , a primeira ação é expressar

$$\begin{aligned} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 &= |(U_n + iV_n) - (u(t; \theta) + iv(t; \theta))|^2 \\ &= |(U_n - u(t; \theta)) + i(V_n - v(t; \theta))|^2 \\ &= [U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2. \end{aligned}$$

como parcelas de uma soma. Para isso, com

$$U_n - u(t; \theta) = \frac{1}{n} \sum_{j=1}^n \cos(tx_j) - \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha}$$

e

$$V_n - v(t; \theta) = \frac{1}{n} \sum_{j=1}^n \sin(tx_j) - \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha},$$

tem-se

$$\begin{aligned} [U_n - u(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \cos(tx_j) \right]^2 \\ &\quad - \frac{2}{n} \left[\sum_{j=1}^n \cos(tx_j) \times \sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right] \\ &\quad + \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 \end{aligned}$$

e

$$\begin{aligned} [V_n - v(t; \theta)]^2 &= \frac{1}{n^2} \left[\sum_{j=1}^n \sin(tx_j) \right]^2 \\ &\quad - \frac{2}{n} \left[\sum_{j=1}^n \sin(tx_j) \times \sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right] \\ &\quad + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 \end{aligned}$$

Disso, tem-se que

$$\begin{aligned} &[U_n - u(t; \theta)]^2 + [V_n - v(t; \theta)]^2 \\ &= \frac{1}{n^2} \left[\left(\sum_{j=1}^n \cos(tx_j) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] \\ &\quad - \frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k^\alpha |t|^\alpha} \left[\cos(\mu_k t) \cos(tx_j) + \sin(\mu_k t) \sin(tx_j) \right] \end{aligned} \quad (D.2)$$

$$+ \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2. \quad (D.3)$$

Escrevendo os somatórios em (D.2) e (D.3) como

$$\begin{aligned} \left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 &= p_1^2 \cos^2(\mu_1) e^{-2\sigma_1^\alpha |t|^\alpha} + \dots + p_K^2 \cos^2(\mu_K) e^{-2\sigma_K^\alpha |t|^\alpha} \\ &\quad + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \cos(\mu_k t) \cos(\mu_j t) e^{-(\sigma_k^\alpha + \sigma_j^\alpha) |t|^\alpha} \end{aligned}$$

e

$$\begin{aligned} \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 &= p_1^2 \sin^2(\mu_1) e^{-\sigma_1^\alpha |t|^\alpha} + \dots + p_K^2 \sin^2(\mu_K) e^{-\sigma_K^\alpha |t|^\alpha} \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j \sin(\mu_k t) \sin(\mu_j t) e^{-(\sigma_k^\alpha + \sigma_j^\alpha) |t|^\alpha} \end{aligned}$$

,além disso aplicando as propriedade trigonométricas $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ e $\cos^2(x) + \sin^2(y) = 1$ obtém-se

$$\begin{aligned} &\left[\sum_{k=1}^K p_k \cos(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 + \left[\sum_{k=1}^K p_k \sin(\mu_k t) e^{-\sigma_k^\alpha |t|^\alpha} \right]^2 \\ &= p_1^2 e^{-2\sigma_1^\alpha |t|^\alpha} \left[\cos^2(\mu_1 t) + \sin^2(\mu_1 t) \right] + \dots + p_K^2 e^{-2\sigma_K^\alpha |t|^\alpha} \left[\cos^2(\mu_K t) + \sin^2(\mu_K t) \right] \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k^\alpha + \sigma_j^\alpha) |t|^\alpha} \left[\cos(\mu_k t) \cos(\mu_j t) + \sin(\mu_k t) \sin(\mu_j t) \right] \\ &= \sum_{k=1}^K p_k^2 e^{-\sigma_k^\alpha |t|^\alpha} + 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k^\alpha + \sigma_j^\alpha) |t|^\alpha} \left[\cos(\mu_k t - \mu_j t) \right]. \end{aligned}$$

Também escrevendo

$$\begin{aligned} &\frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k^\alpha |t|^\alpha} \left[\cos(\mu_k t) \cos(tx_j) + \sin(\mu_k t) \sin(tx_j) \right] \\ &= -\frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k^\alpha |t|^\alpha} \cos[(\mu_k - x_j)t] \end{aligned} \quad (D.4)$$

e

$$\frac{1}{n^2} \left[\left(\sum_{i=1}^n \cos(tx_i) \right)^2 + \left(\sum_{j=1}^n \sin(tx_j) \right)^2 \right] = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \cos[(x_i - x_j)t]$$

Finalmente, chega-se a

$$\begin{aligned} |\tilde{\varphi}_n(t) - \varphi_X(t; \Theta)|^2 &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \cos[(x_i - x_j)t] \\ &+ \sum_{k=1}^K p_k^2 e^{-2\sigma_k^\alpha |t|^\alpha} \\ &+ 2 \sum_{k=1}^K \sum_{\substack{j=1 \\ j \neq k}}^K p_k p_j e^{-(\sigma_k^\alpha + \sigma_j^\alpha) |t|^\alpha} \cos[(\mu_k - \mu_j)t] \\ &- \frac{2}{n} \sum_{k=1}^K \sum_{j=1}^n p_k e^{-\sigma_k^\alpha |t|^\alpha} \cos[(\mu_k - x_j)t] \end{aligned} \quad (D.5)$$

Demonstração Parte II

A segunda parte da demonstração consiste em integrar, em relação a t , de $(-\infty$ a $\infty)$ cada parcela, desenvolvida na parte I, multiplicada pela função $w(t) = |t|^\alpha$. Para $K = 2$, com a mudança de variável $s = |t|$, ignorando a primeira parcela em (D.5) a função a ser minimizada é

$$\begin{aligned}
 E_n &= 2p^2 \int_0^\infty e^{-2\sigma_1^\alpha t^\alpha} t^{-b} dt + 2(1-p)^2 \int_0^\infty e^{-2\sigma_2^\alpha t^\alpha} t^{-b} dt \\
 &+ 4p(1-p) \int_0^\infty e^{-(\sigma_1^\alpha + \sigma_2^\alpha)t^\alpha} \cos[(\mu_1 - \mu_2)t] t^{-b} dt \\
 &- \frac{4p}{n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_1^\alpha t^\alpha} \cos[(\mu_1 - x_j)t] t^{-b} dt \\
 &- \frac{4(1-p)}{n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_2^\alpha t^\alpha} \cos[(\mu_2 - x_j)t] t^{-b} dt
 \end{aligned} \tag{D.6}$$

Com $t^\alpha = y$, tem-se que $t = y^{\frac{1}{\alpha}}$ e $dt = \frac{1}{\alpha} y^{\frac{1}{\alpha}-1} dy$, declara-se

$$\begin{aligned}
 \int_0^\infty e^{-2\sigma^\alpha t^\alpha} &= \int_0^\infty e^{-2\sigma^\alpha y} y^{\frac{-b}{\alpha}} \frac{1}{\alpha} y^{\frac{1}{\alpha}-1} dy \\
 &= \frac{1}{\alpha} \int_0^\infty e^{-2\sigma^\alpha y} y^{\frac{1-b}{\alpha}-1} dy \\
 &= \frac{1}{\alpha} \frac{\Gamma\left(\frac{1-b}{\alpha}\right)}{(2\sigma^\alpha)^{\frac{1-b}{\alpha}}},
 \end{aligned} \tag{D.7}$$

$1 - b > 0$, isto é $b < 1$. Com (D.7) aplicada em (D.6), resulta que

$$\begin{aligned}
 E_n &= \frac{2p^2 \Gamma\left(\frac{1-b}{\alpha}\right)}{\alpha (2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2 \Gamma\left(\frac{1-b}{\alpha}\right)}{\alpha (2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\
 &+ 4p(1-p) \int_0^\infty e^{-(\sigma_1^\alpha + \sigma_2^\alpha)t^\alpha} \cos[(\mu_1 - \mu_2)t] t^{-b} dt \\
 &- \frac{4p}{n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_1^\alpha t^\alpha} \cos[(\mu_1 - x_j)t] t^{-b} dt \\
 &- \frac{4(1-p)}{n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_2^\alpha t^\alpha} \cos[(\mu_2 - x_j)t] t^{-b} dt.
 \end{aligned} \tag{D.8}$$

A mesma substituição é aplicada nas três ultimas integrais. Assim,

$$\begin{aligned}
E_n &= \frac{2p^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\
&+ \frac{4p(1-p)}{\alpha} \int_0^\infty e^{-(\sigma_1^\alpha + \sigma_2^\alpha)y} \cos\left[(\mu_1 - \mu_2)y^{\frac{1}{\alpha}}\right] y^{\frac{1-b}{\alpha}-1} dy \\
&- \frac{4p}{\alpha n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_1^\alpha y} \cos\left[(\mu_1 - x_j)y^{\frac{1}{\alpha}}\right] y^{\frac{1-b}{\alpha}-1} dy \\
&- \frac{4(1-p)}{\alpha n} \sum_{j=1}^n \int_0^\infty e^{-\sigma_2^\alpha y} \cos\left[(\mu_2 - x_j)y^{\frac{1}{\alpha}}\right] y^{\frac{1-b}{\alpha}-1} dy. \tag{D.9}
\end{aligned}$$

Em (D.9), substitua

$$\cos(ax) = \sum_{k=0}^{\infty} (-1)^k \frac{(ax)^{2k}}{2k!}, \forall x \tag{D.10}$$

O resultado é

$$\begin{aligned}
E_n &= \frac{2p^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\
&+ \frac{4p(1-p)}{\alpha} \int_0^\infty \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_1 - \mu_2)^{2k}}{2k!} e^{-(\sigma_1^\alpha + \sigma_2^\alpha)y} y^{\frac{2k+1-b}{\alpha}-1} dy \\
&- \frac{4p}{\alpha n} \sum_{j=1}^n \int_0^\infty \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_1 - x_j)^{2k}}{2k!} e^{-\sigma_1^\alpha y} y^{\frac{2k+1-b}{\alpha}-1} dy \\
&- \frac{4(1-p)}{\alpha n} \sum_{j=1}^n \int_0^\infty \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_2 - x_j)^{2k}}{2k!} e^{-\sigma_2^\alpha y} y^{\frac{2k+1-b}{\alpha}-1} dy. \tag{D.11}
\end{aligned}$$

Organizando os somatórios e integrais, chega-se

$$\begin{aligned}
E_n &= \frac{2p^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\
&+ \frac{4p(1-p)}{\alpha} \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_1 - \mu_2)^{2k}}{2k!} \int_0^\infty e^{-(\sigma_1^\alpha + \sigma_2^\alpha)t^\alpha} y^{\frac{2k+1-b}{\alpha}-1} dy \\
&- \frac{4p}{\alpha n} \sum_{j=1}^n \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_1 - x_j)^{2k}}{2k!} \int_0^\infty e^{-\sigma_1^\alpha t^\alpha} y^{\frac{2k+1-b}{\alpha}-1} dy \\
&- \frac{4(1-p)}{\alpha n} \sum_{j=1}^n \sum_{k=0}^{\infty} (-1)^k \frac{(\mu_2 - x_j)^{2k}}{2k!} \int_0^\infty e^{-\sigma_2^\alpha t^\alpha} y^{\frac{2k+1-b}{\alpha}-1} dy. \tag{D.12}
\end{aligned}$$

cs Comparando a integral em (D.12) com (D.7), obtém-se o resultado desejado

$$\begin{aligned}
D_n(\Theta, x) &= \frac{2p^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_1^\alpha)^{\frac{1-b}{\alpha}}} + \frac{2(1-p)^2\Gamma(\frac{1-b}{\alpha})}{\alpha(2\sigma_2^\alpha)^{\frac{1-b}{\alpha}}} \\
&+ \frac{4p(1-p)}{\alpha} \sum_{k=0}^{\infty} \frac{(-1)^k(\mu_1 - \mu_2)^{2k}\Gamma(\frac{2k+1-b}{\alpha})}{(\sigma_1^\alpha + \sigma_2^\alpha)^{\frac{2k+1-b}{\alpha}}} \\
&- \frac{4p}{\alpha n} \sum_{j=1}^n \left[\sum_{k=0}^{\infty} \frac{(-1)^k(\mu_1 - x_j)^{2k}\Gamma(\frac{2k+1-b}{\alpha})}{\sigma_1^{2k+1-b}} \right] \\
&- \frac{4(1-p)}{\alpha n} \sum_{j=1}^n \left[\sum_{k=0}^{\infty} \frac{(-1)^k(\mu_2 - x_j)^{2k}\Gamma(\frac{2k+1-b}{\alpha})}{\sigma_2^{2k+1-b}} \right].
\end{aligned}$$

Apêndice E

Programas para Cálculo da Densidade de uma Distribuição Estável

E.1 Programa R - Algoritmo I

```
### Objetivo: calcular a f.d.p. de uma v.a. estável X,  
### com parâmetros a, b, u , s.  
### O algoritmo utilizado é proposto por Nolan (2011) e  
### utiliza a parametrização 0 padronizada (u=0,b=1).  
### Parâmetros: alpha = a , beta = b, mu = u, sigma = s e  
### X o vetor de valores a ter a f.d.p. calculada.  
  
destavel<-function (x,a,s=1,b,u=0){  
  ### para a != 1  
  fdp1 <- function (t)  
  {  
    res <- (1/(pi*s))*cos(((x[i]-(u))/s)* t + b*tan(pi*a/2)*(t-t^a))*exp(-t^a)  
  }  
  ### para a = 1  
  fdp2 <- function (t)
```

```

{
res <- (1/pi*s)*cos (((x[i]-(u))/s) * t+b*(2/pi)*t*log(t))*exp(-t)
}

### calculo
aux2 <- rep(0,length(x))
for (i in 1 : length(x))
{
if (a!=1) aux1<-integrate (fdp1,0,Inf, subdivisions = 1000)
else aux1 <- integrate (fdp2,0,Inf, subdivisions = 1000)
  aux2[i]<-aux1$value
}
return (aux2)
}

```

E.2 Programa R - Algoritmo II

```

### Objetivo: calcular a f.d.p. de uma v.a. estavel Y,
### com parâmetros a, b, u , s.
### O algoritmo é proposto por Otiniano et al. (2010)
### e utiliza parametrização S0 padronizada (u=0,b=1).
### Parâmetros:
### caracteriza a f.d.p: alpha=a,beta =b,mu=u,sigma=s
### vetor de valores para ser calculada a f.d.p.: y
### Retorna:
### uma matriz 4 * n, n é numero de elementos
### do vetor y. Nessa matriz as coluna armazenam
### col. 1 - valor de y
### col. 2 - valor do centro da f.d.p
### col. 3 - valor da cauda da f.d.p
### col. 4 - valor da f.d.p

stable2 <- function (y, a, m, s)

```

```

{
## variaveis auxiliares
x <- y
# o valor de b = 0 para calculo de f.d.p estáveis
# simétricas, mas pode assumir valores no inter-
# valo [-1,1]
b <- 0

# Números de termos dos somatórios
v1<- 100
v2<- 10

# tamanho da amostra
n <- length (x)

## variaveis locais
# variavel s2
s2 <- s * ( 1 + (b ^ 2) * ( ( tan (a * pi / 2) ) ^2 ) ) ^ ( 1 / ( 2 * a ) )

# variavel b2
if ( a > 0 & a < 1 )
{
b2 <- ( 2 / ( pi * a ) ) * atan ( b * tan ( ( pi * a ) / 2 ) )
} else if ( a > 1 & a < 2 )
{
b2 <- ( 2 / ( pi * ( a - 2 ) ) ) * atan ( b * tan ( ( pi * a ) / 2 ) )
}

### definição das séries para o centro e caudas
### da f.d.p. que depende do valore de 'a' se
### y >= m ou y < m

```

```

## Centro da densidade: a > 1
# f1 quando x >= m
f1 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * ( v + 1 ) * (a + b2 * a - 2 * b2) / a )
                  * gamma ( (v + 1) / a ) * ( - ( x - m ) / s2) ^ v )
                  / ( a * s2 * pi * gamma (v + 1))
  }
  return ( aux )
}

# f2 quando x < m
f2 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * ( v + 1 ) * (a - b2 * a + 2 * b2) / a )
                  * gamma ( (v + 1) / a ) * ( ( x - m ) / s2) ^ v )
                  / ( a * s2 * pi * gamma (v + 1))
  }
  return ( aux )
}

f <- function (x)
{
  if ( x < m ) aux <- f2 (x,v1)
  else if ( x >= m ) aux <- f1 (x,v1)
  return (aux)
}

```

```

}

## Caudas da densidade: a > 1
# r1 quando x >= 0
r1 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * ( 2 + v * a + b2 * v * a - 2 * v * b2 ) )
                  * ( ( -1 ) ^ v )
                  * (s2/( x - m)) ^ (v * a ) * gamma (1 + v*a) )
                  / ( (x - m) * pi * gamma (v + 1) )
  }
  return ( aux )
}

# r2 quando x < 0.
r2 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * ( 2 + v * a - b2 * v * a + 2 * v * b2 ) )
                  * ( ( -1 ) ^ v )
                  * (-s2/( x - m)) ^ (v * a ) * gamma (1 + v*a) )
                  / ( -(x - m) * pi * gamma (v + 1) )
  }
  return ( aux )
}

r <- function (x)
{
  if ( x < m ) aux <- r2 (x,v2)

```

```

else if ( x >= m ) aux <- r1 (x,v2)
return (aux)
}

## centro da densidade: a < 1
# a1 quando x >= m
a1 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * (1 + b2 + v + v * b2) )
                  * gamma ( (v + 1) / a ) * ( - ( x - m ) / s2) ^ v )
                  / ( a * s2 * pi * gamma (v + 1))
  }
  return ( aux )
}
# a2 quando x < m
a2 <- function (x , lim)
{
  aux <- 0
  for (v in 0 : lim )
  {
    aux <- aux +( sin ( 0.5 * pi * (1 - b2 + v - v * b2) )
                  * gamma ( (v + 1) / a ) * ( ( x - m ) / s2) ^ v )
                  / ( a * s2 * pi * gamma (v + 1))
  }
  return ( aux )
}

a3 <- function (x)
{

```

```

if ( x < m ) aux <- a2 (x,v2)
else if ( x >= m ) aux <- a1 (x,v2)
return (aux)
}

##caudas da densidade: a < 1
# h1 quando x >= mu 0 .
h1 <- function (x , lim)
{
aux <- 0
for (v in 0 : lim )
{
aux <- aux +( sin ( 0.5 * pi * ( 2 + v * a + b2 * v * a ) ) * ( ( -1 ) ^ v )
               * (s2/( x - m)) ^ ( v * a ) * gamma (1 + v*a) )
               / ( (x - m) * pi * gamma (v + 1) )
}
return ( aux )
}

# h2 quando x < 0
h2 <- function (x , lim)
{
aux <- 0
for (v in 0 : lim )
{
aux <- aux +( sin ( 0.5 * pi * ( 2 + v * a - b2 * v * a ) ) * ( ( -1 ) ^ v )
               * (-s2/( x - m)) ^ (v * a ) * gamma (1 + v*a) )
               / ( -(x - m) * pi * gamma (v + 1) )
}
return ( aux )
}

h <- function (x)
{

```



```

if ( x < m ) aux <- h2 (x,v1)
else if ( x >= m ) aux <- h1 (x,v1)
return (aux)
}

## montando a desnsidade
g <- matrix (0 , n, 4, byrow = TRUE )
if (a > 1 )
{
  for (i in 1 : n)
  {
    g [i,1] <- y[i]
    g [i,2] <- f(y[i])
    g [i,3] <- r(y[i])
    g [i,4] <- min(abs(g [i,2]),abs(g[i,3]))
  }
}
else if ( a < 1 )
{
  for (i in 1 : n)
  {
    g [i,1] <- y[i]
    g [i,2] <- a3(y[i])
    g [i,3] <- h(y[i])
    g [i,4] <- min(abs(g [i,2]),abs(g[i,3]))
  }
}
return (g)
}#fim da função

```

Apêndice F

Programa R para Gerar Valores de uma Distribuição Estável

```
### Simula uma v.a. estavel X com parametros a, b, u , s.  
### O algoritmo utilizado é proposto por Nolan.  
### Se p == 0, Y tem a parametrização S0 (Nolan);  
### se p != 0, Y tem a parametrização S (canonica Taqqu).  
### Parâmetros: alpha = a ,beta = b, mu = u, sigma = s, p=0.
```

```
restavel <- function(n=1,a,s=1,b,u=0,p=0){  
  # Distribuições auxiliares  
  V = runif (n,-pi/2,pi/2)  
  W = rexp (n,1)  
  # Função auxiliar  
  theta <- function (a,b)  
  {  
    atan(b*tan(pi*a/2))/a  
  }  
  if (a != 1) {  
    B<- theta (a,b)  
    X1<- sin(a*(V+B)) /(cos(a*B)*cos(V))^(1/a)  
    X2<- (cos(a*B + (a-1)*V)/W)^((1-a)/a)
```

```

X <- X1 * X2
Y0<- s*(X-b*tan(pi*a/2))+u
Y1<- s*X + u
      } else {
X<- (2/pi)*((pi/2+b*V)*tan(V)-b*log(pi/2*W*cos(V)/(pi/2+b*V)))
Y0<- s*X+u
Y1<- s*X + (u + b*tan(pi*a/2))
      }
if (p==0) return (Y0) else return (Y1)
}

```