

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Métodos Bayesianos para seleção de modelos de mistura de distribuições normais e t de Student assimétricas

Walkiria Maria de Oliveira Macerau

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Walkiria Maria de Oliveira Macerau

Métodos Bayesianos para seleção de modelos de mistura de distribuições normais e t de Student assimétricas

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-USFCar, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Luis Aparecido Milan

**USP - USFCar – São Carlos
Julho de 2023**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

M141m Macerau, Walkiria Maria de Oliveira
Métodos Bayesianos para seleção de modelos de
mistura de distribuições normais e t de Student
assimétricas / Walkiria Maria de Oliveira Macerau;
orientador Luis Aparecido Milan. -- São Carlos,
2023.
140 p.

Tese (Doutorado - Programa Interinstitucional de
Pós-graduação em Estatística) -- Instituto de Ciências
Matemáticas e de Computação, Universidade de São
Paulo, 2023.

1. Modelos de mistura. 2. Distribuições
assimétricas. 3. Inferência Bayesiana. 4. Métodos
MCMC. 5. Data driven Reversible jump. I. Milan,
Luis Aparecido, orient. II. Título.

Walkiria Maria de Oliveira Macerau

Bayesian methods for selection of the skew normal and
skew t mixture models

Thesis submitted to the Institute of Mathematics
and Computer Science – ICMC-USP and to the
Department of Statistics – DEs-UFSCar – in
accordance with the requirements of the Statistics
Interagency Graduate Program, for the degree of
Doctor in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Luis Aparecido Milan

USP - UFSCar – São Carlos
July 2023



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Walkiria Maria de Oliveira Macerau, realizada em 28/06/2023.

Comissão Julgadora:

Prof. Dr. Luis Aparecido Milan (UFSCar)

Profa. Dra. Vera Lucia Damasceno Tomazella (UFSCar)

Prof. Dr. Erlandson Ferreira Saraiva (UFMS)

Profa. Dra. Miriam Harumi Tsunemi (UNESP)

Profa. Dra. Roseli Aparecida Leandro (ESALQ/USP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

Para Rodolfo, Pedro e Lavínia, com todo meu amor.

AGRADECIMENTOS

A Deus por infinitas graças recebidas.

Ao meu esposo Rodolfo, amor da minha vida, pelo apoio de todos os dias.

Aos meus filhos Pedro e Lavínia, meus presentes de Deus, pela compreensão da minha ausência.

À minha mãe Otília (*in memoriam*), meu eterno amor e minha eterna gratidão.

À minha família pelo apoio e compreensão principalmente nos dois primeiros anos do doutorado.

Aos colegas do curso de Mestrado e Doutorado pela ajuda de qualquer forma.

Aos colegas do Departamento de Estatística da Universidade Estadual de Maringá pelo apoio na conclusão do curso de Doutorado em Estatística.

Ao Instituto de Ciências Matemáticas e de Computação (ICMC-USP) e ao Departamento de Estatística da Universidade Federal de São Carlos (DEs-USFCar), por me oferecerem a oportunidade de poder concluir o curso de Doutorado em Estatística.

Aos professores do Programa Interinstitucional de Pós-Graduação em Estatística USP-UFSCar, que ministraram aulas em nosso curso, por transmitirem os seus conhecimentos com empenho e dedicação.

Aos professores Vera Tomazella e Erlandson F. Saraiva, membros da banca do exame de qualificação e defesa, e as professoras Miriam Harumi Tsunemi e Roseli Aparecida Leandro, membros da banca do exame de defesa, pelas sugestões feitas para melhoria deste trabalho.

Em especial ao professor Luis Aparecido Milan pelas ideias, sugestões, apoio e amizade.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

MACERAU, W. M. O. **Métodos Bayesianos para seleção de modelos de mistura de distribuições normais e t de Student assimétricas**. 2023. 140 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo e Departamento de Estatística, Universidade Federal de São Carlos, São Carlos – SP, 2023.

Neste trabalho, consideramos modelos de mistura cujas componentes da mistura são modelados pelas distribuições normal assimétrica e t de Student assimétrica. Para a estimação desses modelos de mistura de distribuições assimétricas, consideramos a abordagem Bayesiana, via métodos de Monte Carlo em Cadeias de Markov (MCMC), uma vez que essa abordagem permite o desenvolvimento de um procedimento de estimação conjunta do número de componentes da mistura e dos parâmetros associados as componentes da mistura. Também empregamos o método *Reversible jump* e propomos o método *Data-driven Reversible jump*, para modelar a mistura das distribuições normais assimétricas e t de Student assimétricas, tanto para ajustar quanto para selecionar o número de componentes da mistura. Comparamos o desempenho desses dois métodos (*Reversible jump* e *Data-driven Reversible jump*) para selecionar o melhor modelo por meio de simulações. O método *Data-driven Reversible jump* foi mais preciso em apontar o melhor modelo nos estudos de simulação realizados.

Palavras-chave: Modelos de mistura, distribuições assimétricas, Inferência Bayesiana, MCMC, *Data-driven Reversible jump*.

ABSTRACT

MACERAU, W. M. O. **Bayesian methods for selection of the skew normal and skew t mixture models**. 2023. 140 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo e Departamento de Estatística, Universidade Federal de São Carlos, São Carlos – SP, 2023.

In this work, we consider mixture models whose components of the mixture are modeled by the skew normal and skew t distributions. For the estimation of these skew mixtures models, we used a Bayesian approach, via Markov Chain Monte Carlo methods (MCMC), since this approach allows the development of a joint estimation procedure of the mixture components number and the associated parameters of the mixture components. We also use the *Reversible jump* method and propose the use of the *Data-driven Reversible jump* method for modelling the mixture of skew normal and skew t distributions, both to adjust and select the number of components of the mixture. We compare the performances of these two methods (*Reversible jump* and *Data-driven Reversible jump*) for selecting the best model through simulations. The *Data-driven Reversible jump* method was more accurate in pointing out the best model in the simulation studies carried out.

Keywords: Mixture models, asymmetric distributions, Bayesian Inference, MCMC, *Data-driven Reversible jump*.

LISTA DE ILUSTRAÇÕES

Figura 1	– Exemplo de uma distribuição de probabilidade simétrica.	8
Figura 2	– Formas apresentadas por uma distribuição de probabilidade: (a) Distribuição assimétrica à esquerda. (b) Distribuição simétrica. (c) Distribuição assimétrica à direita. Média: linha vermelha, mediana: linha verde e moda: linha azul.	8
Figura 3	– Densidades da distribuição normal assimétrica com parâmetros $\xi = 0$, e (a) $\alpha = 1$; $\omega = 0.5$ (linha em preto), $\omega = 1$ (linha em vermelho), $\omega = 1.5$ (linha em verde), $\omega = 2$ (linha em azul), e (b) $\omega = 1$; $\alpha = -1$ (linha em preto), $\alpha = 0$ (linha em vermelho - distribuição normal padronizada), $\alpha = 0.5$ (linha em verde), $\alpha = 1$ (linha em azul).	11
Figura 4	– Densidades da distribuição t de Student assimétrica com parâmetros $\xi = 0$, $\omega = 1$, e (a) $\alpha = 2$; $\nu = 1$ (linha em preto), $\nu = 2$ (linha em vermelho), $\nu = 3$ (linha em verde), $\nu = 10$ (linha em azul), e (b) $\nu = 3$; $\alpha = 0$ (linha em preto - distribuição t de Student com 3 graus de liberdade), $\alpha = 0.5$ (linha em vermelho), $\alpha = 1$ (linha em verde), $\alpha = 2$ (linha em azul).	15
Figura 5	– Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	67
Figura 6	– Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	67
Figura 7	– Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	68
Figura 8	– Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	68
Figura 9	– Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	70
Figura 10	– Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	71

Figura 11 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	71
Figura 12 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	72
Figura 13 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	74
Figura 14 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	75
Figura 15 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	75
Figura 16 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.	76
Figura 17 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados de escores globais.	80
Figura 18 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do IMC.	83
Figura 19 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do tempo de duração das erupções do vulcão.	86
Figura 20 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do tempo de espera entre as erupções do vulcão.	88
Figura 21 – Dados da criptomoeda <i>Bitcoin</i> variável último preço: (a) Gráfico da série original, (b) Gráfico da função de autocorrelação da série original, (c) Gráfico do retorno da série. (d) Gráfico da função de autocorrelação do retorno da série, (e) Gráfico do logaritmo do retorno da série. (f) Gráfico da função de autocorrelação do logaritmo do retorno da série.	89

Figura 22 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 1$ componente e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).	127
Figura 23 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 2$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).	128
Figura 24 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 3$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).	128
Figura 25 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 4$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).	129
Figura 26 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 1$ componente e $N = 1000$	129
Figura 27 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 2$ componentes e $N = 1000$	130
Figura 28 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 3$ componentes e $N = 1000$	130
Figura 29 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 4$ componentes e $N = 1000$	131
Figura 30 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 2$ componentes, dados dos escores globais. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).	131
Figura 31 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 2$ componentes, dados dos escores globais.	132

LISTA DE QUADROS

Quadro 1 – Casos particulares das distribuições em estudo	16
---	----

LISTA DE TABELAS

Tabela 1	– Valores dos parâmetros das componentes dos modelos de mistura, empregados nos estudos de simulação.	62
Tabela 2	– Estimativas das probabilidades <i>a posteriori</i> de K obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições normais, dados gerados com $K = 1, 2, 3$ e 4 componentes, e tamanhos de amostras $N = 100, 200, 500$ e 1000.	63
Tabela 3	– Estimativas das probabilidades <i>a posteriori</i> de K obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições normais assimétricas, dados gerados com $K = 1, 2, 3$ e 4 componentes e tamanhos de amostras $N = 100, 200, 500$ e 1000.	64
Tabela 4	– Estimativas das probabilidades <i>a posteriori</i> de K obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições t de Student assimétricas, dados gerados com $K = 1, 2, 3$ e 4 componentes e tamanhos de amostras $N = 100, 200, 500$ e 1000.	65
Tabela 5	– Médias <i>a posteriori</i> e Intervalos de 95% de Credibilidade dos parâmetros dos modelos de mistura de distribuições normais, obtidas pelo método <i>Data-driven Reversible jump</i> , com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$	66
Tabela 6	– Médias <i>a posteriori</i> e Intervalos de 95% de Credibilidade dos parâmetros dos modelos de mistura de distribuições normais assimétricas, obtidas pelo método <i>Data-driven Reversible jump</i> , com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$	69
Tabela 7	– Médias <i>a posteriori</i> e Intervalo de 95% de Credibilidade para parâmetros dos modelos de mistura de distribuições t de Student assimétricas, obtidas pelo método <i>Data-driven Reversible jump</i> , com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$	73
Tabela 8	– Estimativas das probabilidades <i>a posteriori</i> do número de componentes K , obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados dos escores globais.	78

Tabela 9 – Médias <i>a posteriori</i> , Medianas <i>a posteriori</i> e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas com $K = 2$ componentes, obtidas pelo método <i>Data-driven Reversible jump</i> , para os dados de escores globais.	79
Tabela 10 – Estimativas das probabilidades <i>a posteriori</i> do número de componentes K , obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas para os dados do IMC.	81
Tabela 11 – Médias <i>a posteriori</i> , Medianas <i>a posteriori</i> e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas com $K = 2$ componentes, obtidas pelo método <i>Data-driven Reversible jump</i> , para os dados do IMC.	82
Tabela 12 – Estimativas das probabilidades <i>a posteriori</i> do número de componentes K , obtidas pelos métodos <i>Data-driven Reversible jump</i> (DDRJ) e <i>Reversible jump</i> (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas para os dados de erupção do vulcão.	84
Tabela 13 – Médias <i>a posteriori</i> , Medianas <i>a posteriori</i> e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas com $K = 2$ componentes, obtidas pelo método <i>Data-driven Reversible jump</i> , para os dados do tempo de duração das erupções do vulcão.	85
Tabela 14 – Médias <i>a posteriori</i> , Medianas <i>a posteriori</i> e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e <i>t</i> de Student assimétricas com $K = 2$ componentes, para os dados do tempo de espera entre as erupções do vulcão.	87

LISTA DE ABREVIATURAS E SIGLAS

AIC	Critério de informação de Akaike
BIC	Critério de informação Bayesiano
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDC	Center for Disease Control
DDRJ	<i>Data-driven Reversible jump</i>
DEs	Departamento de Estatística
DIC	Critério de desvio de informação
EM	<i>Expectation-Maximization</i> (Esperança maximizada)
EMV	Estimativa de máxima verossimilhança
EUA	Estados Unidos da América
GS	<i>Gibbs sampling</i>
IC	Intervalos de credibilidade
ICMC	Instituto de Ciências Matemáticas e de Computação
IMC	Índice de massa corporal
MCMC	Monte Carlos via cadeias de Markov
MH	Metropolis-Hastings
NCHS	National Center Health Statistics
PIB	Produto interno bruto
R	<i>Software R</i>
RJ	<i>Reversible jump</i>
TRV	Teste da razão de verossimilhança
UFSCar	Universidade Federal de São Carlos
USP	Universidade de São Paulo

LISTA DE SÍMBOLOS

X — Variável aleatória X

γ — Medida de assimetria

μ_X — Média ou valor esperado da variável aleatória X

$med(X)$ — Mediana da variável aleatória X

σ_X — Desvio padrão da variável aleatória X

γ_P — Coeficiente de Assimetria de Pearson

$mo(X)$ — Moda da variável aleatória X

x — Qualquer valor da reta real ou um estado qualquer

$f_X(x)$ — Função densidade de probabilidade da variável aleatória X

\in — Pertence

\Re — Conjunto dos reais com dimensão 1

$\tilde{\mu}_3$ — Medida de assimetria de Edgeworth (1904) e Charlier (1905)

μ_3 — Terceiro momento em torno da média ou valor esperado da variável aleatória X

$f_Z(z|\alpha)$ — Função densidade de probabilidade de uma distribuição normal assimétrica padrão

$\phi(\cdot)$ — Densidade de uma distribuição normal padrão

$\Phi(\cdot)$ — Função de distribuição acumulada

α — Parâmetro de assimetria da distribuição normal assimétrica e da t de Student assimétrica

Z — Variável aleatória com distribuição normal assimétrica padrão

\sim — Segue a distribuição

$NA(\cdot)$ — Distribuição normal assimétrica

$N(\cdot)$ — Distribuição normal

\rightarrow — Tende para

$+\infty$ — Mais infinito

\xrightarrow{d} — Convergência em distribuição

$-\infty$ — Menos infinito

Z^2 — Distribuição normal ao quadrado

χ_r^2 — Distribuição Qui-quadrado com r graus de liberdade, com $r > 1$

ξ — Parâmetro de locação da distribuição normal assimétrica e da t de Student assimétrica

ω^2 — Parâmetro de escala da distribuição normal assimétrica e da t de Student assimétrica

Y — Variável aleatória Y

$f_Z(z|\alpha, \nu)$ — Função densidade de probabilidade de uma distribuição t de Student assimétrica com ν graus de liberdade

$t_\nu(\cdot)$ — Densidade de uma distribuição t de Student com ν graus de liberdade

$T_\nu(\cdot)$ — Função de distribuição acumulada de uma distribuição t de Student com ν graus de liberdade

ν — Parâmetro dos graus de liberdade da distribuição t de Student simétrica e assimétrica

G — Variável aleatória G com distribuição Gama(ψ, λ)

V — Variável aleatória V com distribuição Qui-quadrado com ν graus de liberdade

\perp — Independência

$T_\nu(\cdot)$ — Distribuição t de Student com ν graus de liberdade

$TA_\nu(\cdot)$ — Distribuição t de Student assimétrica com ν graus de liberdade

$\pm\infty$ — Mais ou menos infinito

K — Número de componentes de mistura

$\boldsymbol{\theta}$ — Vetor de parâmetros

\boldsymbol{p} — Vetor de parâmetros do pesos ou probabilidade de mistura

k — k -ésima componente de mistura

p_k — k -ésimo peso ou proporção da k -ésima componente de mistura

$f_Y(y_i|\boldsymbol{\theta}_k)$ — densidades de mistura pertencentes à uma mesma família de distribuições

$\boldsymbol{\theta}_k$ — k -ésimo parâmetro (vetor ou escalar) associado a k -ésima componente de mistura

N — Total de observações

\boldsymbol{Y} — Vetor aleatório ou variável aleatória contínua independente

\boldsymbol{S} — Vetor de variáveis aleatórias discreta não observáveis

\boldsymbol{S}_i — Variável aleatória com distribuição Multinomial com K resultados e uma observação

$Y_i|\boldsymbol{S}_i = k, \boldsymbol{\theta}$ — Variável aleatória que segue a distribuição condicional $f_{Y_i}(y|\boldsymbol{S}_i = s_i, \boldsymbol{\theta})$

\Re^K — Conjunto dos reais com dimensão K

Pr — Probabilidade

iid — Independente e identicamente distribuído

$I_{\{s_i\}}(k)$ — Função indicadora que associa Y_i a k -ésima componente de mistura

\mathbf{y} — Uma realização de \mathbf{Y}

\mathbf{s} — Uma realização de \mathbf{S}

n_k — Número de observações da k -ésima componente de mistura

\mathcal{Y} — Espaço amostral de observações

Θ — Espaço paramétrico

φ — Vetor de parâmetros

\mathfrak{R}^+ — Conjunto dos reais positivo com dimensão 1

$\pi(\cdot)$ — Distribuição *a priori*

$\pi(\cdot|\cdot)$ — Distribuição *a posteriori*

$\mathcal{L}(\cdot|\cdot)$ — Função de verossimilhança

\propto — Proporcional

$q(\cdot, \cdot)$ — Distribuição proposta

$\alpha(\cdot, \cdot)$ — Probabilidade de aceitação do algoritmo *Metropolis-Hastings*

u — Valor gerado de uma distribuição Uniforme no intervalo 0 a 1

$U(\cdot, \cdot)$ — Distribuição Uniforme

d — Dimensão d

\mathcal{M}_k — k -ésimo modelo de mistura

\mathfrak{R}^{m_k} — Conjunto dos reais com dimensão m_k

m_k — Dimensão m_k

$p(\cdot)$ — Distribuição de probabilidade

Θ — Espaço de estados

$q_{d_{k \rightarrow k'}}(\cdot)$ — Densidade conhecida

$g_{k \rightarrow k'}(\cdot)$ — Função de mapeamento

$Beta(\cdot, \cdot)$ — Distribuição Beta

τ — Parâmetro da distribuição de Dirichlet

k_0 — Número de componentes vazias antes do movimento nascimento (*birth*)

b_K — Probabilidade da escolha de um movimento *split*

d_K — Probabilidade da escolha de um movimento *merge*
 K_{max} — Valor máximo especificado para K
 μ — Parâmetro de locação da distribuição normal (média da distribuição)
 σ — Parâmetro do desvio padrão da distribuição normal
 σ^2 — Parâmetro de escala da distribuição normal (variância da distribuição)
 μ_k — Parâmetro ou hiperparâmetro de locação da distribuição normal da k -ésima componente de mistura (média da k -ésima componente de mistura)
 σ_k — Parâmetro ou hiperparâmetro do desvio padrão da distribuição normal da k -ésima componente de mistura
 σ_k^2 — Parâmetro ou hiperparâmetro de escala da distribuição normal da k -ésima componente de mistura (variância da k -ésima componente de mistura)
 P_{aloc} — Probabilidade de alocação das componentes
 $g(\cdot)$ — Densidade da distribuição *Beta*
 T — Número de iterações
 B — *burn in*
 N_K — Número de vezes que o número de componentes é igual a k
 P_K — probabilidade *a posteriori* de K
 \tilde{K} — Estimativa da probabilidade *a posteriori* de K
 δ — Vetor de probabilidades da distribuição Multinomial com K resultados e uma observação
 ξ_k — Parâmetro de locação da distribuição normal assimétrica e da t de Student assimétrica da k -ésima componente de mistura
 ω_k^2 — Parâmetro de escala da distribuição normal assimétrica e t de Student assimétrica da k -ésima componente de mistura
 α_k — Parâmetro de assimetria da distribuição normal e da t de Student assimétrica da k -ésima componente de mistura
 \bar{y}_k — Média amostral da k -ésima componente
 S_k^2 — Variância amostral da k -ésima componente
 a_k — Hiperparâmetro da distribuição Gama inversa da k -ésima componente de mistura
 b_k — Hiperparâmetro da distribuição Gama inversa da k -ésima componente de mistura
 τ_k — Hiperparâmetro da distribuição de Dirichlet da k -ésima componente de mistura
 $U_d(\cdot)$ — Distribuição Uniforme discreta

v_k — Parâmetro dos graus de liberdade da distribuição t de Student assimétrica da k -ésima componente de mistura

τ_k — Hiperparâmetro da distribuição de Dirichlet da k -ésima componente de mistura

c_k — Hiperparâmetro da distribuição Gama da k -ésima componente de mistura

d_k — Hiperparâmetro da distribuição Gama da k -ésima componente de mistura

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Introdução e Revisão Bibliográfica	1
1.2	Objetivo e organização da tese	6
2	ASSIMETRIA E DISTRIBUIÇÕES ASSIMÉTRICAS	7
2.1	Assimetria de uma distribuição	7
2.2	Distribuições assimétricas	9
2.3	Distribuição normal assimétrica	10
2.3.1	<i>Medidas características da distribuição normal assimétrica</i>	12
2.4	Distribuição <i>t</i> de Student assimétrica	12
2.4.1	<i>Medidas características da distribuição t de Student assimétrica</i>	15
2.4.2	<i>Casos particulares das distribuições normais e t de Student assimétricas</i>	16
3	MODELOS DE MISTURA	17
3.1	Modelos de mistura de distribuições	17
3.1.1	<i>FUNÇÃO DE VEROSSIMILHANÇA</i>	20
3.2	Número de componentes	21
3.3	Identificabilidade de um modelo de mistura de distribuições	23
3.3.1	<i>NÃO IDENTIFICABILIDADE DEVIDO À INVARIÂNCIA DE RE-CLASSIFICAÇÃO DAS COMPONENTES DE MISTURA</i>	23
3.3.2	<i>NÃO IDENTIFICABILIDADE DEVIDO AO POTENCIAL DE SOBRE AJUSTE</i>	24
4	ABORDAGEM BAYESIANA	25
4.1	Distribuições <i>a priori</i> e <i>a posteriori</i>	25
4.2	Algoritmo <i>Metropolis-Hastings</i>	26
4.3	Algoritmo <i>Gibbs sampling</i>	27
4.4	Método <i>Reversible jump</i>	28
4.4.1	<i>MÉTODO REVERSIBLE JUMP PARA MODELOS DE MISTURA</i>	30
4.4.2	<i>ATUALIZAÇÃO DE s</i>	33
4.5	Método <i>Data-driven Reversible jump</i>	34
5	MODELO DE MISTURA DE DISTRIBUIÇÕES NORMAIS ASSIMÉTRICAS	39

5.1	O modelo	39
5.2	Função de verossimilhança	40
5.3	Abordagem Bayesiana	40
5.4	Estimação	41
6	MODELO DE MISTURA DE DISTRIBUIÇÕES T DE STUDENT ASSIMÉTRICAS	49
6.1	O modelo	49
6.2	Função de verossimilhança	50
6.3	Abordagem Bayesiana	50
6.4	Estimação	52
7	RESULTADOS	61
7.1	Simulação	61
7.2	Aplicação em dados reais	76
7.2.1	<i>DADOS DE ESCORES DE COGNIÇÃO</i>	77
7.2.2	<i>DADOS DE ÍNDICE DE MASSA CORPORAL E DADOS DE ERUPÇÃO DE VULCÃO</i>	80
7.2.3	<i>DADOS DE CRIPTOMOEDAS</i>	88
8	CONCLUSÃO	91
	REFERÊNCIAS	93
	APÊNDICE A JACOBIANO DA TRANSFORMAÇÃO	97
A.1	Jacobiano da proposta <i>split</i> para o modelo mistura de distribuições normais	97
A.2	Jacobiano da proposta <i>split</i> para o modelo mistura de distribuições normais assimétricas	98
A.3	Jacobiano da proposta <i>split</i> para o modelo de mistura de distribuições t de Student assimétricas	100
	APÊNDICE B CÁLCULO DAS FUNÇÕES DE VEROSSIMILHANÇA 103	
B.1	Cálculo da função de verossimilhança do modelo de mistura de distribuições normais assimétricas	103
B.2	Cálculo da função de verossimilhança do modelo de mistura de distribuições t de Student assimétricas	105
	APÊNDICE C CÁLCULO DAS DISTRIBUIÇÕES A POSTERIORI . 107	
C.1	Cálculo da distribuição <i>a posteriori</i> do modelo de mistura de distribuições normais assimétricas	107

C.2	Cálculo da distribuição <i>a posteriori</i> do modelo de mistura de distribuições <i>t</i> de Student assimétricas	109
APÊNDICE D CÁLCULO DAS PROBABILIDADES DE ACEITAÇÃO DO MÉTODO RJ 111		
D.1	Cálculo da probabilidade de aceitação do movimento “nascimento” para o modelo de mistura de distribuições normais assimétricas	111
D.2	Cálculo da probabilidade de aceitação do movimento “nascimento” para o modelo de mistura de distribuições <i>t</i> de Student assimétricas	112
D.3	Cálculo da probabilidade de aceitação do movimento <i>split</i> para o modelo de mistura de distribuições normais assimétricas	113
D.4	Cálculo da probabilidade de aceitação do movimento <i>split</i> para o modelo de mistura de distribuições <i>t</i> de Student assimétricas	115
APÊNDICE E CÁLCULO DAS PROBABILIDADES DE ACEITAÇÃO DO MÉTODO DDRJ 117		
E.1	Cálculo da probabilidade de aceitação do movimento <i>split</i> para o modelo de mistura de distribuições normais assimétricas	117
E.2	Cálculo da probabilidade de aceitação do movimento <i>split</i> para o modelo de mistura de distribuições <i>t</i> de Student assimétricas	122
APÊNDICE F VERIFICAÇÃO DAS CADEIAS 127		
F.1	Estudo de simulação	127
F.2	Análise de dados reais	131
APÊNDICE G PROGRAMA 133		

INTRODUÇÃO

Neste capítulo, apresentamos uma introdução em conjunto com uma revisão bibliográfica sobre modelos de mistura, com ênfase nos modelos em que as distribuições das componentes de mistura são as distribuições normais assimétricas e t de Student assimétricas. Também apresentamos a nossa proposta de trabalho, os objetivos e a organização dessa tese.

1.1 Introdução e Revisão Bibliográfica

Modelos de mistura são utilizados quando assumimos que as observações são provenientes de uma população composta por subpopulações, denominadas de componentes de mistura. Essas componentes de mistura são modeladas por uma distribuição de probabilidade pertencente a uma mesma família de distribuições, mas com valores dos parâmetros diferentes.

Os modelos de mistura são aplicáveis nas mais diversas áreas, tais como, astronomia, biologia, genética, medicina, psiquiatria, economia, engenharia, marketing e ciências sociais. Nessas aplicações, são abrangidas uma variedade de técnicas nas principais áreas da estatística incluindo, por exemplo, as análises de clusters, de discriminante e de sobrevivência (MCLACHLAN; PEEL, 2000).

Na análise dos modelos de mistura buscamos distinguir qual distribuição é mais indicada para ajustar as componentes e também procuramos identificar qual o número de componentes é mais apropriado para modelar os dados em questão. Sendo assim, a busca por distribuições que consigam um melhor ajuste para as componentes do modelo mistura e a procura por metodologias que consigam estimar o número de componentes recebem atenção, despertam o interesse dos estatísticos e de toda a comunidade científica.

A distribuição normal é a distribuição pioneira utilizada em modelos de mistura, como descrito em Titterington *et al.* (1985), McLachlan e Basford (1988) e McLachlan e Peel (2000). Entretanto, na presença de caudas pesadas dentro das componentes de mistura Peel e McLachlan

(2000), Shoham (2002), Shoham *et al.* (2003), Lin e Lee (2004) e Wang *et al.* (2004) aplicam a distribuição t de Student. Os modelos de mistura de distribuições normais e t de Student resolvem muitos problemas, porém se as componentes de mistura contêm caudas pesadas e assimetria, nestes casos é mais indicado utilizar uma distribuição assimétrica para modelar as componentes de mistura com esses comportamentos.

Lin *et al.* (2007b) introduz um modelo de mistura baseado na distribuição normal assimétrica, proposta por Azzalini (1985), com o objetivo de contornar a assimetria dentro das componentes de mistura. Segundo os autores, o modelo de mistura de normais assimétricas oferece grande flexibilidade na modelagem de componentes com comportamento assimétrico, porém na presença de caudas pesadas ele não é tão eficiente. Em outro trabalho, Lin *et al.* (2007a) propõe um modelo de mistura baseado na distribuição t de Student assimétrica, introduzida por Azzalini e Capitanio (2003), esse modelo consegue ajustar as caudas pesadas e a assimetria. Além disso, o modelo de mistura de distribuições t de Student assimétricas possui a vantagem de conter os modelos de mistura de distribuições normais, t de Student e normais assimétricas como casos particulares.

Encontramos aplicações de modelos de mistura de distribuições assimétricas nas mais diversas áreas de pesquisa, tais como, economia, climatologia, engenharia e ciências biomédicas, sendo que nessas aplicações o número de componentes dos modelos de mistura é um número conhecido. Por exemplo, mistura de normais assimétricas são aplicadas em dados de erupção de vulcão e bancários, respectivamente, em Lin *et al.* (2007b) e Lin (2009), mistura de t de Student assimétricas são empregadas em dados geocronológicos, índice de massa corporal (IMC) e citometria de fluxo, respectivamente, em Jasra *et al.* (2006), Lin *et al.* (2007a) e Lee e McLachlan (2016), e mistura de normais assimétricas e t de Student assimétricas são adotadas para modelar dados de IMC e produto interno bruto (PIB) em Cabral *et al.* (2008), e em dados de citometria de fluxo em Lee e McLachlan (2013). No caso multivariado, Lin (2009) faz uso de misturas de normais assimétricas em dados bancários e Wang *et al.* (2004) aplica misturas de t de Student assimétricas em dados botânicos.

Na literatura são apresentadas algumas abordagens para obter as estimativas dos parâmetros das componentes de mistura e para selecionar o número de componentes dos modelos, que é outro papel importante da análise de modelos de mistura. Sendo que esses procedimentos de estimação tem o objetivo de estimar um modelo de mistura que apresenta o melhor ajuste e a maioria considera o número de componentes como sendo um valor conhecido.

Variações do algoritmo EM (*Expectation-Maximization*) (DEMPSTER *et al.*, 1977) para obter as estimativas de máxima verossimilhança (EMV's) dos parâmetros do modelo de mistura são apresentadas em Lin *et al.* (2007b). Os autores também empregam uma abordagem Bayesiana baseada na simulação de Monte Carlo via cadeia de Markov (MCMC) por meio do algoritmo *Metropolis-Hasting* (MH) (METROPOLIS *et al.*, 1953), para obter as estimativas dos parâmetros do modelo, e para comparar o ajuste dos modelos aplicam a log-verossimilhança

e os critérios de informação de Akaike (AIC) (AKAIKE, 1974) e de informação bayesiano (BIC) (SCHWARZ, 1978). O algoritmo MCMC *Gibbs sampling* (GS) é proposto em Frühwirth-Schnatter e Pyne (2010), para obter as estimativas do modelo de mistura de distribuições normais assimétricas e t de Student assimétricas univariada e multivariada. Nesta abordagem, os autores verificam o ajuste dos modelos por meio da log-verossimilhança *a posteriori*, variações do BIC (SCHWARZ, 1978) e do desvio de informação (DIC) (SPIEGELHALTER *et al.*, 2002). O algoritmo EM (DEMPSTER *et al.*, 1977) é empregado, para obter as EMV's do modelo de mistura de distribuições normais assimétricas multivariada em Lin (2009) e o ajuste dos modelos é verificado por meio da log-verossimilhança, do teste da razão de verossimilhança (TRV), do AIC (AKAIKE, 1974) e do BIC (SCHWARZ, 1978). Nas abordagens citadas todos os métodos fixam o número de componentes de mistura, em seguida as estimativas dos parâmetros associados as componentes de mistura são obtidas e finalmente os modelos são comparados por meio de algum critério de seleção baseado na estimação frequentista.

Na abordagem de estimação frequentista, como o algoritmo EM (DEMPSTER *et al.*, 1977), o método apresenta, em geral, problemas de convergência a um máximo local. Segundo Frühwirth-Schnatter (2006, p. 52) as dificuldades práticas com a estimação do máximo da função de verossimilhança em modelos de mistura são para encontrar o máximo global da função de verossimilhança numericamente, além dos problemas de convergência quando o tamanho da amostra é pequeno e quando as componentes de mistura não estão bem separadas. Um outro problema do algoritmo EM ocorre quando uma das componentes fica vazia, pois neste caso as estimativas de máxima verossimilhança não existem. Com relação aos métodos de comparação de ajustes frequentistas, como por exemplo, os critérios de AIC (AKAIKE, 1974) e BIC (SCHWARZ, 1978), apesar de serem conceitualmente diferente, eles também utilizam o máximo da função de verossimilhança como medida de ajustamento.

Outro fato nos procedimentos de estimação frequentistas é que as estimativas dos parâmetros do modelo são obtidas para um número de componentes conhecido. Já os bayesianos são baseados em simulação MCMC, que permitem propor alterações no valor do número de componentes e estimar os parâmetros do modelo, durante as iterações. Desta forma, a abordagem Bayesiana é mais atrativa, pois permite escolher o modelo com melhor ajuste e estimar os parâmetros do modelo ao mesmo tempo.

Em nosso trabalho, propomos uma abordagem Bayesiana tanto para obter as estimativas dos parâmetros, quanto para estimar o número de componentes dos modelos de mistura. Mais precisamente, neste trabalho empregamos os seguintes métodos de estimação MCMC: o algoritmo *Metropolis-Hastings* ((METROPOLIS *et al.*, 1953) e (HASTINGS, 1970)) aplicado para atualizar as estimativas dos parâmetros associados as componentes de mistura, o método *Reversible jump* (RJ) (GREEN, 1995) utilizado para estimar o número de componentes de mistura e os parâmetros associados as componentes, e propomos o método *Data-driven Reversible jump* (DDRJ) também para estimar o número de componentes de mistura e os parâmetros associados

as componentes. Nestas duas últimas abordagens, a estimativa da probabilidade *a posteriori* do número de componentes é a medida usada para verificar qual o modelo de mistura apresenta melhor ajuste.

O algoritmo *Metropolis-Hastings* ((METROPOLIS *et al.*, 1953) e (HASTINGS, 1970)) é um método MCMC para obter uma sequência de amostras aleatórias a partir de uma distribuição de probabilidade da qual a amostragem direta é difícil. Desta forma, essa sequência pode ser usada para uma aproximação desta distribuição, que em nosso caso é a distribuição *a posteriori*.

O método *Reversible jump* (GREEN, 1995) é uma generalização do algoritmo *Metropolis-Hastings* ((METROPOLIS *et al.*, 1953) e (HASTINGS, 1970)), porém permite transições do modelo atual para modelos de dimensões superiores ou inferiores. Richardson e Green (1997) desenvolvem o método *Reversible jump* (GREEN, 1995) para modelos de mistura de distribuições normais. Segundo os autores, a cada iteração é proposta uma alteração do número de componentes de mistura por meio de modificações nos parâmetros de uma componente do modelo, e condicional aos “novos” parâmetros é feita uma realocação das observações, sendo que esta proposta é aceita de acordo com uma certa probabilidade. Desta forma, condicional no valor atualizado do número de componentes, os parâmetros dos modelos são atualizados de acordo com suas distribuições *a posteriori*. Ao final das iterações, selecionamos o número de componentes de acordo a sua probabilidade *a posteriori*. O método *Reversible jump* (GREEN, 1995) pode apresentar problemas em transitar entre os diferentes modelos (ZUANETTI; MILAN, 2017).

No método *Data-driven Reversible jump*, a atualização do número de componentes é feita por movimentos baseados nos dados, ou seja, os dados de uma componente são particionados e então as estimativas dos valores dos parâmetros são obtidos condicional nos dados e, portanto, evita transformações aleatórias dos parâmetros. Assim, o método torna-se mais simples e acelera o procedimento de busca do modelo que apresente melhor ajuste, uma vez que, candidatos mais adequados são gerados usando os dados. O *Data-driven Reversible jump* apresenta um desempenho melhor comparado a outras abordagens, pois para estimar o número de componentes de mistura o método simula um número de componentes utilizando os dados, escolhe uma componente a ser excluída de acordo com sua importância no modelo atual, ou mescla os efeitos de duas componentes. Desta forma, o modelo candidato é mais fácil de ser aceito e o espaço de modelos possíveis são mais facilmente explorados.

Jain e Neal (2004, 2007) apresentam um método orientado aos dados para analisar modelos de mistura de distribuições normais, e empregam a metodologia proposta em dados biológicos e simulados. Os autores consideram que as mudanças de agrupamento não incrementais do método evitam o problema de ficar preso em máximos locais, permitindo que a distribuição *a posteriori* seja totalmente explorada. Além disso, em dimensões mais altas, os procedimentos continuam funcionando bem e a convergência para a distribuição *a posteriori* do modelo de mistura é relativamente rápida.

Saraiva (2009) propõe os algoritmos *split-merge* MCMC e *birth-split-merge* MCMC para a estimação conjunta dos parâmetros e do número de componentes de um modelo de mistura de distribuições normais, e aplica a metodologia em dados de galáxia, de expressão gênica e simulados. O autor conclui que os algoritmos exploram melhor os grupos de observações, uma vez que as observações são alocadas de acordo com os parâmetros gerados das distribuições *a posteriori* do número de componentes dado as observações. Desta forma, os algoritmos tornam-se mais eficientes computacionalmente, em relação aos métodos alternativos, pois cada movimento pode ser rapidamente proposto e testado.

Saraiva e Milan (2012) aplicam o método *birth-split-merge* MCMC (SARAIVA, 2009) para estimar o modelo de mistura de distribuições normais, e empregam a metodologia em dados de expressão gênica e simulados. Os autores concluem que métodos orientados aos dados apresentam melhor desempenho do que os métodos que não utilizam essa abordagem, pois evitam considerar vários modelos que têm pouca ou nenhuma chance de serem selecionados como modelos apropriados, além de ser mais eficiente do ponto de vista computacional.

Zuanetti (2016) desenvolve o método *Data-driven Reversible jump*, baseado no algoritmo *birth-split-merge* MCMC proposto por Saraiva (2009), para modelos de mistura de distribuições binomiais, e aplica a metodologia em dados simulados e de genética. A autora compara o desempenho do *Data-driven Reversible jump* com o tradicional *Reversible jump* (GREEN, 1995) e conclui que este último se move com maior dificuldade entre os possíveis modelos e pode manter-se em um modelo específico por mais tempo, mesmo que este não seja um modelo adequado. Já o *Data-driven Reversible jump* mostra melhor desempenho, além de fornecer informações de incerteza para todos os parâmetros por meio dos intervalos de credibilidade.

Por sua vez, Zuanetti e Milan (2017) também empregam o método *Data-driven Reversible jump*, em modelos de mistura de distribuições binomiais dependentes, e utilizam a metodologia em dados de incidência de diabetes e simulados. Os autores concluem que o método *Data-driven Reversible jump* mostra melhor desempenho para identificar o modelo mais adequado e para estimar seus parâmetros.

Em resumo, o método *Data-driven Reversible jump* apresenta as seguintes vantagens:

1. Não há necessidade de transformações arbitrárias nos parâmetros do modelo;
2. O procedimento de busca por um modelo com melhor ajuste é considerado rápido;
3. Flutua entre os modelos mais próximo do verdadeiro modelo;

Desta forma, propomos o método *Data-driven Reversible jump* para analisar modelos de mistura de distribuições assimétricas, mais precisamente analisar modelos em que as componentes de mistura seguem as distribuições normal assimétrica e a *t* de Student assimétrica, caso univariado. Uma vez o método *Data-driven Reversible jump* possui as vantagens já mencionadas acima e ainda não foi proposto para modelos de mistura cujas distribuições se afasta da condição de simetria.

1.2 Objetivo e organização da tese

Os objetivos de pesquisa relatados nessa tese são:

1. Propor a utilização da abordagem Bayesiana com a estimação e a seleção de modelos por meio do método *Data-driven Reversible jump*, para os modelos de mistura com distribuições assimétricas, mais especificamente as distribuições normal assimétrica e t de Student assimétrica;
2. Descrever e aplicar um método de simulação para verificar o desempenho do método *Data-driven Reversible jump*, tanto na seleção de modelos de mistura quanto na estimação dos parâmetros dos modelos de mistura;
3. Aplicar a metodologia proposta em conjuntos de dados simulados e reais;

A contribuição teórica presente neste trabalho é o desenvolvimento do método *Data-driven Reversible jump*, que é baseado nos dados, para os modelos de mistura de distribuições normais assimétricas e t de Student assimétricas. O método *Data-driven Reversible jump* proposto permite buscar informações sobre a estimativa no número de componentes do modelo de mistura de distribuições normais assimétricas e t de Student assimétricas, quando esse valor é desconhecido, a partir da estimativa da distribuição *a posteriori* para o número de componentes. Além disso, o método fornece a estimação conjunta dos parâmetros associados às componentes de mistura.

Todos os procedimentos de estimação e as simulações computacionais foram implementadas no *Software R* (R Core Team, 2022).

Esta tese está organizada em 8 capítulos. No Capítulo 2, apresentamos a definição de assimetria e fazemos uma breve descrição das distribuições normal assimétrica e t de Student assimétrica. No Capítulo 3, definimos mistura de distribuições, descrevemos a estrutura de um modelo de mistura de distribuições, suas respectivas funções de verossimilhança e também evidenciamos a questão da estimação do número de componentes e da identificabilidade de modelos de mistura. No Capítulo 4, apresentamos a abordagem Bayesiana empregada neste estudo, por meio do esquema MCMC por meio dos algoritmos *Metropolis-Hastings*, *Gibbs sampling*, e os métodos *Reversible jump* e *Data-driven Reversible jump*. Nos Capítulos 5 e 6, descrevemos, respectivamente, os modelos de misturas de distribuições normais assimétricas e t de Student assimétricas, suas funções de verossimilhança, e também a abordagem Bayesiana empregada. No Capítulo 7, apresentamos, os resultados encontrados no estudo de simulação e nas análises dos dados reais. Finalmente, no Capítulo 8, apresentamos as considerações finais deste trabalho.

ASSIMETRIA E DISTRIBUIÇÕES ASSIMÉTRICAS

Neste capítulo, definimos assimetria de uma distribuição e fazemos uma breve descrição das distribuições normal assimétrica e t de Student assimétrica.

2.1 Assimetria de uma distribuição

A assimetria ou a simetria é uma característica da distribuição dos dados, que pode ser verificada por meio das medidas de centralidade, isto é, a média, a moda e a mediana dos dados. Algumas medidas para verificar a assimetria de observações podem ser encontradas em [Macerau \(2012\)](#).

Considerando uma variável aleatória X , [Mood et al. \(1974, p. 76\)](#) utilizam a seguinte medida para verificar assimetria

$$\gamma = \frac{\mu_X - med(X)}{\sigma_X},$$

em que $\mu_X = \mathbb{E}[X]$, $med(X)$ e σ_X são, respectivamente, a média ou o valor esperado, a mediana e o desvio padrão da variável aleatória X .

Temos que $-1 \leq \gamma \leq 1$, além disso, se $\gamma < 0$ a distribuição é assimétrica à esquerda, se $\gamma > 0$ a distribuição é assimétrica à direita e se $\gamma = 0$ a distribuição é simétrica. [Pearson \(1895\)](#) introduziu uma medida de assimetria denominada Coeficiente de Assimetria de Pearson, dada por

$$\gamma_P \cong \frac{\mu_X - mo(X)}{\sigma_X}, \quad (2.1)$$

em que μ_X , $mo(X)$ e σ_X são, respectivamente, a média ou o valor esperado, a moda e o desvio padrão da variável aleatória X .

Na Expressão (2.1) dividimos pelo desvio padrão para obter uma medida adimensional. Sob assimetria à direita, $\mu_X > mo(X)$ e conseqüentemente $\gamma_P > 0$, e sob assimetria à esquerda o oposto ocorre.

Identificar se a variável aleatória X apresenta ou não simetria é um fator importante, pois, uma vez que apontamos a assimetria, podemos procurar por distribuições de probabilidade que consigam modelar o comportamento assimétrico da variável aleatória.

Apresentamos a Figura 1 com um exemplo de uma distribuição de probabilidade simétrica.

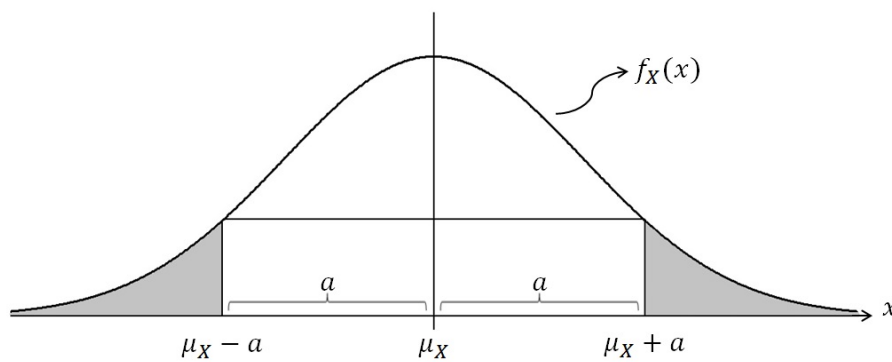


Figura 1 – Exemplo de uma distribuição de probabilidade simétrica.

Na Figura 1, temos que μ_X representa a média ou o valor esperado da variável aleatória X , isto é, a média da distribuição de probabilidade, a é uma distância qualquer, $\mu_X - a$ e $\mu_X + a$ são pontos no eixo dos valores x , que representam os valores que a variável aleatória contínua X pode assumir, e $f_X(x)$ representa a função densidade de probabilidade da variável aleatória X , ou seja, a probabilidade que a variável aleatória X pode assumir para cada intervalo da reta x , com $x \in \mathfrak{R}$.

A Figura 2 apresenta três ilustrações de tipos de assimetria apresentados por uma distribuição de probabilidade.

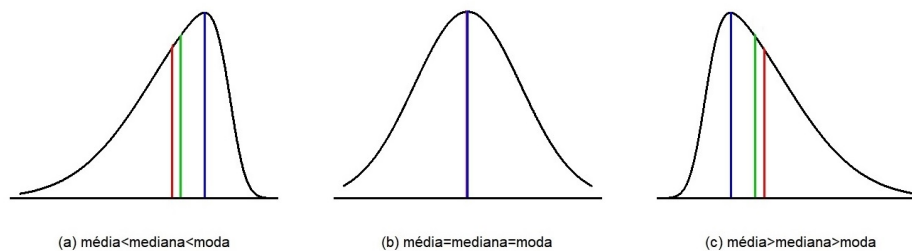


Figura 2 – Formas apresentadas por uma distribuição de probabilidade: (a) Distribuição assimétrica à esquerda. (b) Distribuição simétrica. (c) Distribuição assimétrica à direita. Média: linha vermelha, mediana: linha verde e moda: linha azul.

Podemos observar que a distribuição de probabilidade apresentada na [Figura 2\(a\)](#) possui uma assimetria à esquerda, ou seja, o gráfico da função densidade tem cauda mais longa à esquerda do que à direita, com relação à reta passando pela moda. Neste caso, temos que a média (linha vermelha) é menor que a mediana (linha verde) que por sua vez é menor que a moda (linha azul), também observamos que a distribuição de probabilidade apresentada na [Figura 2\(c\)](#) possui uma assimetria à direita, isto é, o gráfico da função densidade tem cauda mais longa à direita do que à esquerda, com relação à reta passando moda. Desta forma, temos que a média (linha vermelha) é maior que a mediana (linha verde) que por sua vez é maior que a moda (linha azul), e a [Figura 2\(b\)](#) apresenta uma distribuição de probabilidade simétrica, em que a média (linha vermelha) é igual a mediana (linha verde) que é igual a moda (linha azul).

Outra medida de assimetria proposta por [Edgeworth \(1904\)](#) e [Charlier \(1905\)](#), é dada por

$$\tilde{\mu}_3 = E \left[\left(\frac{X - \mu_X}{\sigma_X} \right)^3 \right] = \frac{\mu_3}{\sigma_X^3}, \quad (2.2)$$

em que μ_X e σ_X são, respectivamente, a média ou valor esperado e o desvio-padrão da variável aleatória X , e μ_3 é o terceiro momento em torno da média ou valor esperado, μ_X .

Na Expressão (2.2), se $\tilde{\mu}_3 < 0$, temos que a distribuição é assimétrica à esquerda, se $\tilde{\mu}_3 > 0$, temos que a distribuição é assimétrica à direita, e se $\tilde{\mu}_3 = 0$ a distribuição é simétrica. Esta medida só é válida quando o terceiro momento da variável aleatória X existe.

[Mood et al. \(1974, p. 75\)](#) também apresenta o terceiro momento em torno da média, μ_3 , como uma medida de assimetria e a Expressão (2.2) como um coeficiente de assimetria. Outros resultados e medidas de assimetria podem ser encontrados em [Arnold e Groeneveld \(1995\)](#).

2.2 Distribuições assimétricas

Na literatura existem várias distribuições assimétricas, tais como, as distribuições normal generalizada, exponencial negativa, Gama, log-normal, Qui-quadrado e Weibull, e o interesse da comunidade científica em estudar distribuições assimétricas vem aumentando a cada dia. Por exemplo, [Azzalini \(1985\)](#) propõe uma classe de distribuições assimétricas que dependem da introdução de um parâmetro com capacidade de modelar a assimetria, a distribuição normal assimétrica pertence à essa classe de distribuições. [Arnold e Groeneveld \(1995\)](#) estudam famílias de distribuições ordenadas por assimetria com respeito à moda, as distribuições Gama, log-normal e Weibull estão compreendidas nesse estudo. [Fernández et al. \(1995\)](#) propõe um método para transformar uma distribuição simétrica em assimétrica, e escrevem a distribuição t de Student assimétrica conforme esse método. [Azzalini e Capitanio \(2003\)](#) apresentam dentro da estrutura de distribuições assimétricas construídas com a introdução do parâmetro de assimetria, a distribuição t de Student assimétrica.

Neste trabalho, empregamos as distribuições normal assimétrica introduzida por [Azzalini \(1985\)](#) e t de Student assimétrica proposta por [Azzalini e Capitanio \(2003\)](#). A escolha da distribuição normal assimétrica é devida pelo fato dessa distribuição pertencer à uma classe de distribuições assimétricas analiticamente tratáveis e que possui como caso particular a distribuição normal. Já a escolha da distribuição t de Student assimétrica é devido a ela ser uma extensão da distribuição normal assimétrica e conter as distribuições t de Student, normal e normal assimétrica como casos particulares.

2.3 Distribuição normal assimétrica

A distribuição normal assimétrica, introduzida por [Azzalini \(1985\)](#), é uma distribuição de probabilidade contínua que estende a distribuição normal permitindo a presença de assimetria.

A função densidade de probabilidade da distribuição normal assimétrica padrão é dada por

$$f_Z(z|\alpha) = 2\phi(z)\Phi(\alpha z), \quad (2.3)$$

em que $\phi(z)$ é a densidade de uma distribuição normal padrão e $\Phi(\alpha z)$ é a função de distribuição acumulada de uma distribuição normal padrão em αz , e $\alpha \in \mathfrak{R}$ é o parâmetro de assimetria.

A distribuição normal padronizada pode ser obtida na Equação (2.3) quando $\alpha = 0$. Quando $\alpha > 0$ a distribuição é assimétrica à direita e quando $\alpha < 0$ é assimétrica à esquerda.

Suponha uma variável aleatória Z com distribuição normal assimétrica padrão com parâmetro α , denotado por $Z \sim NA(\alpha)$. Algumas propriedades básicas da distribuição normal assimétrica padrão, dadas por [Azzalini \(1985\)](#) são:

1. $NA(0) = N(0, 1)$;
2. Se $Z \sim NA(\alpha)$, então $-Z \sim NA(-\alpha)$;
3. Quando $\alpha \rightarrow +\infty$, $Z \xrightarrow{d} |X|$, e quando $\alpha \rightarrow -\infty$, $Z \xrightarrow{d} -|X|$, em que $X \sim N(0, 1)$;
4. Se $Z \sim NA(\alpha)$, então $Z^2 \sim \chi_1^2$.

Para adicionar o parâmetro de locação ξ e o parâmetro de escala ω , utilizamos a transformação linear $Y = \xi + \omega Z$, em que Y é uma variável aleatória com distribuição normal assimétrica com parâmetros (ξ, ω^2, α) , ou seja, $Y \sim NA(\xi, \omega^2, \alpha)$, e Z é uma variável aleatória com função densidade dada por (2.3).

A função densidade de probabilidade de Y é dada por

$$f_Y(y|\xi, \omega^2, \alpha) = \frac{2}{\omega} \phi\left(\frac{y-\xi}{\omega}\right) \Phi\left(\alpha \left(\frac{y-\xi}{\omega}\right)\right), \quad (2.4)$$

em que ξ é o parâmetro de locação, $\xi \in (-\infty, +\infty)$; ω^2 é o parâmetro de escala, $\omega^2 > 0$; α é o parâmetro de forma, $\alpha \in (-\infty, +\infty)$, também denominado parâmetro de assimetria.

A Expressão (2.4) pode ser reescrita como

$$f_Y(y|\xi, \omega^2, \alpha) = \frac{2}{\sqrt{2\pi}\omega^2} \exp\left\{-\frac{1}{2}\left(\frac{y-\xi}{\omega}\right)^2\right\} \int_{-\infty}^{\alpha\left(\frac{y-\xi}{\omega}\right)} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du. \quad (2.5)$$

Usando o *software* R (R Core Team, 2022), construímos a Figura 3, que mostra o comportamento da função densidade de probabilidade de uma distribuição normal assimétrica com parâmetros $\xi = 0$, e com (a) $\alpha = 1$ e ω assumindo os valores 0.5; 1.0; 1.5 e 2.0, e (b) $\omega = 1$ e α assumindo os valores -1.0; 0.0; 0.5 e 1.0.

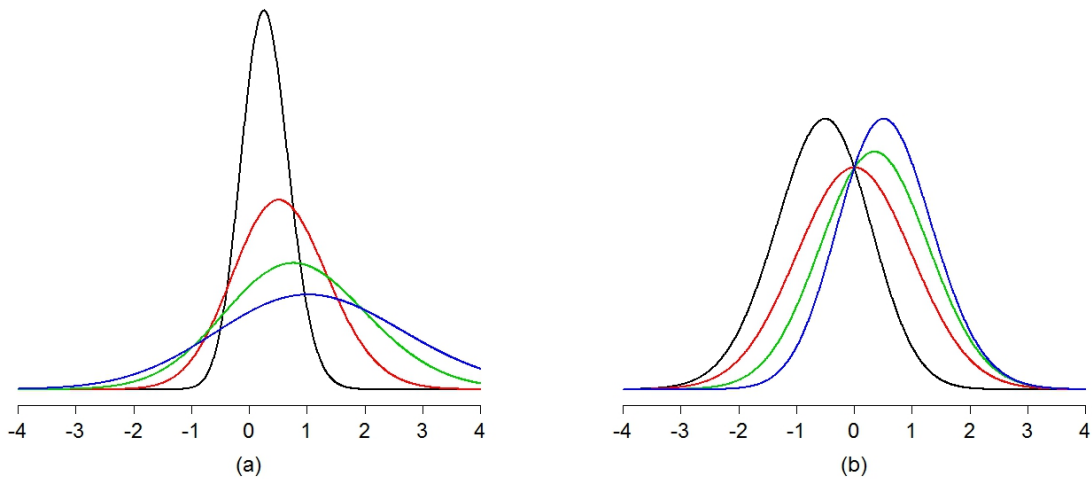


Figura 3 – Densidades da distribuição normal assimétrica com parâmetros $\xi = 0$, e (a) $\alpha = 1$; $\omega = 0.5$ (linha em preto), $\omega = 1$ (linha em vermelho), $\omega = 1.5$ (linha em verde), $\omega = 2$ (linha em azul), e (b) $\omega = 1$; $\alpha = -1$ (linha em preto), $\alpha = 0$ (linha em vermelho - distribuição normal padronizada), $\alpha = 0.5$ (linha em verde), $\alpha = 1$ (linha em azul).

Observamos, por meio da Figura 3(a), que o parâmetro ω controla a dispersão, e, por meio da Figura 3(b), que o parâmetro α controla a assimetria da distribuição. A Figura 3(b), também apresenta a densidade de uma distribuição normal simétrica (linha em vermelho).

Se substituirmos o parâmetro $\alpha = 0$ na Equação (2.5), temos que a função densidade será

$$f_Y(y|\xi, \omega^2, 0) = \frac{2}{\sqrt{2\pi}\omega^2} \exp\left\{-\frac{1}{2}\left(\frac{y-\xi}{\omega}\right)^2\right\} \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du. \quad (2.6)$$

Das propriedades da distribuição normal padrão, temos que o valor da integral da Equação (2.6) é igual a $1/2$. Portanto, obtemos a distribuição normal com média ξ e variância ω^2 .

2.3.1 Medidas características da distribuição normal assimétrica

Suponha Y uma variável aleatória com distribuição $NA(\xi, \omega^2, \alpha)$, a esperança de Y é dada por

$$E[Y] = \xi + \omega\sqrt{2/\pi}\delta(\alpha),$$

a variância de Y é

$$\text{Var}[Y] = \omega^2 (1 - (2/\pi)\delta^2(\alpha)),$$

os coeficientes de assimetria e curtose são, respectivamente, dados por

$$\gamma_1 = \frac{4 - \pi}{2} \frac{(\sqrt{2/\pi}\delta(\alpha))^3}{(1 - (2/\pi)\delta^2(\alpha))^{3/2}},$$

e

$$\gamma_2 = 2(\pi - 3) \frac{(\sqrt{2/\pi}\delta(\alpha))^4}{(1 - (2/\pi)\delta^2(\alpha))^4},$$

em que $\delta(\alpha) = \frac{\alpha}{\sqrt{1+\alpha^2}}$.

Maiores detalhes sobre os cálculos das medidas características da distribuição normal assimétrica ver [Azzalini \(2013, p. 30-31\)](#).

2.4 Distribuição t de Student assimétrica

A distribuição t de Student assimétrica é uma extensão da distribuição normal assimétrica, isto é, uma distribuição de probabilidade contínua assimétrica, e sua função densidade de probabilidade é dada por

$$f_Z(z|\alpha, \nu) = 2t_\nu(z)T_\nu(\alpha z), \quad (2.7)$$

em que $t_\nu(z)$ é a densidade de uma distribuição t de Student com ν graus de liberdade e $T_\nu(\alpha z)$ é a função de distribuição acumulada de uma distribuição t de Student em αz com ν graus de liberdade, e $\alpha \in \Re$ é o parâmetro de assimetria, ([AZZALINI; CAPITANIO, 2003](#)).

Podemos recuperar a distribuição t de Student em (2.7) quando $\alpha = 0$.

Sabemos que a distribuição t de Student é a razão de uma variável aleatória que segue a distribuição normal padrão, pela raiz quadrada de uma variável aleatória Qui-quadrado dividida pelos seus graus de liberdade ([MOOD et al., 1974, p. 249](#)). Segundo [Azzalini e Capitanio \(2003\)](#) se quisermos introduzir a assimetria é só substituir a variável aleatória normal padrão por uma normal assimétrica padrão. Para fazer essa transformação um resultado preliminar sobre as variáveis aleatórias que seguem a distribuição Gama é necessário.

Lema 1. (AZZALINI; CAPITANIO, 2003) Se uma variável aleatória $G \sim \text{Gama}(\psi, \lambda)$, com média ψ/λ e variância ψ/λ^2 , então para qualquer $a, b \in \mathfrak{R}$

$$E \left[\Phi \left(a\sqrt{G} + b \right) \right] = \text{Pr} \left(T \leq a\sqrt{\psi/\lambda} \right),$$

em que T é uma variável aleatória que possui distribuição t de Student não central com 2ψ graus de liberdade e parâmetro de não centralidade $-b$.

Prova do Lema 1: Se $U \sim N(0, 1)$, então

$$\begin{aligned} E \left[\Phi \left(a\sqrt{G} + b \right) \right] &= E \left[\text{Pr} \left(U \leq a\sqrt{g} + b \mid G = g \right) \right] \\ &= E \left[\text{Pr} \left(U - b \leq a(g)^{1/2} \mid G = g \right) \right] \\ &= E \left[\text{Pr} \left(\frac{U - b}{(g\lambda/\psi)^{1/2}} \leq a(\psi/\lambda)^{1/2} \mid G = g \right) \right] \\ &= \text{Pr} \left(T' \leq a\sqrt{\psi/\lambda} \right). \end{aligned}$$

em que $T = \frac{U-b}{(G\lambda/\psi)^{1/2}}$, é uma variável aleatória que possui distribuição t de Student não central com 2ψ graus de liberdade e parâmetro de não centralidade $-b$. \square

Azzalini e Capitanio (2003) definem a distribuição t de Student assimétrica com a correspondente transformação

$$Y = \frac{Z}{\sqrt{V/v}},$$

em que Z possui densidade dada por (2.3) e V é uma variável aleatória que segue a distribuição Qui-quadrado com v graus de liberdade, com função densidade dada por

$$f_V(v|v) = \frac{1}{\Gamma(v/2)} \left(\frac{1}{2} \right)^{v/2} v^{(v/2)-1} e^{-\frac{1}{2}v}, \quad v > 0, \quad e \quad v > 0, \quad (2.8)$$

e $V \perp Z$.

Aplicando o Lema 1 à uma variável que segue a distribuição $\text{Gama}(v/2, v/2)$, Azzalini e Capitanio (2003), definem a função densidade

$$f_X(x|\alpha, v) = \frac{2}{\omega} t_v(x) T_{v+1} \left(\alpha x \sqrt{\frac{v+1}{v+x^2}} \right), \quad (2.9)$$

em que X é uma variável aleatória que possui distribuição t de Student assimétrica, com parâmetro de assimetria $\alpha \in (-\infty, +\infty)$, $t_v(\cdot)$ e $T_{v+1}(\cdot)$ denotam, respectivamente, a função densidade de probabilidade e a função de distribuição acumulada de uma distribuição t de Student com, respectivamente, v e $v+1$ graus de liberdade, isto é, $X \sim TA_v(\alpha)$.

Considerando a transformação linear $Y = \xi + \omega X$, e a função densidade de X dada em (2.9), temos que Y segue a distribuição t de Student assimétrica com função densidade dada por

$$f_Y(y|\xi, \omega^2, \alpha, v) = \frac{2}{\omega} t_v \left(\frac{y-\xi}{\omega} \right) T_{v+1} \left(\alpha \left(\frac{y-\xi}{\omega} \right) \sqrt{\frac{v+1}{v + \left(\frac{y-\xi}{\omega} \right)^2}} \right), \quad \text{com } y \in \mathfrak{R}, \quad (2.10)$$

em que ξ é o parâmetro de locação, $\xi \in (-\infty, +\infty)$; ω^2 é o parâmetro de escala, $\omega^2 > 0$; α é o parâmetro de assimetria, $\alpha \in (-\infty, +\infty)$; e ν são os graus de liberdade, com $\nu \geq 1$; desta forma, podemos denotar que $Y \sim TA_\nu(\xi, \omega^2, \alpha)$.

Podemos reescrever a Expressão (2.10) como

$$f_Y(y|\xi, \omega^2, \alpha, \nu) = \frac{2}{\omega} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} \left(1 + \frac{\left(\frac{y-\xi}{\omega}\right)^2}{\nu} \right)^{-\left(\frac{\nu+1}{2}\right)} \\ \times \int_{-\infty}^{\alpha\left(\frac{y-\xi}{\omega}\right)} \frac{\sqrt{\frac{\nu+2}{\nu+1+\left(\frac{y-\xi}{\omega}\right)^2}}}{\Gamma(\frac{\nu+1}{2})\sqrt{(\nu+1)\pi}} \frac{\Gamma(\frac{\nu+2}{2})}{\Gamma(\frac{\nu+1}{2})\sqrt{(\nu+1)\pi}} \left(1 + \frac{u^2}{\nu+1} \right)^{-\left(\frac{\nu+2}{2}\right)} du. \quad (2.11)$$

Algumas propriedades da distribuição t de Student assimétrica, dadas por [Azzalini e Capitanio \(2003\)](#) são:

1. $TA_\nu(0) = t_\nu(0, 1)$.
2. Quando $\alpha \rightarrow \pm\infty$, $TA_\nu(\alpha)$ tende para uma densidade t de Student truncada em valores positivos, ou seja, se $\alpha \rightarrow \pm\infty$, então $X \sim |t_\nu|$.
3. Se $X \sim TA_\nu(\alpha)$, então $\xi + \omega X \sim TA_\nu(\xi, \omega^2, \alpha)$.

Usando o *software* R ([R Core Team, 2022](#)), construímos a [Figura 4](#), que mostra o comportamento da função densidade de probabilidade de uma distribuição t de Student assimétrica com parâmetros $\xi = 0$, $\omega = 1$, e com (a) $\alpha = 2$ e ν assumindo os valores 1; 2; 3 e 10, e (b) $\nu = 3$ e α assumindo os valores 0; 0.5; 1.0 e 2.0.

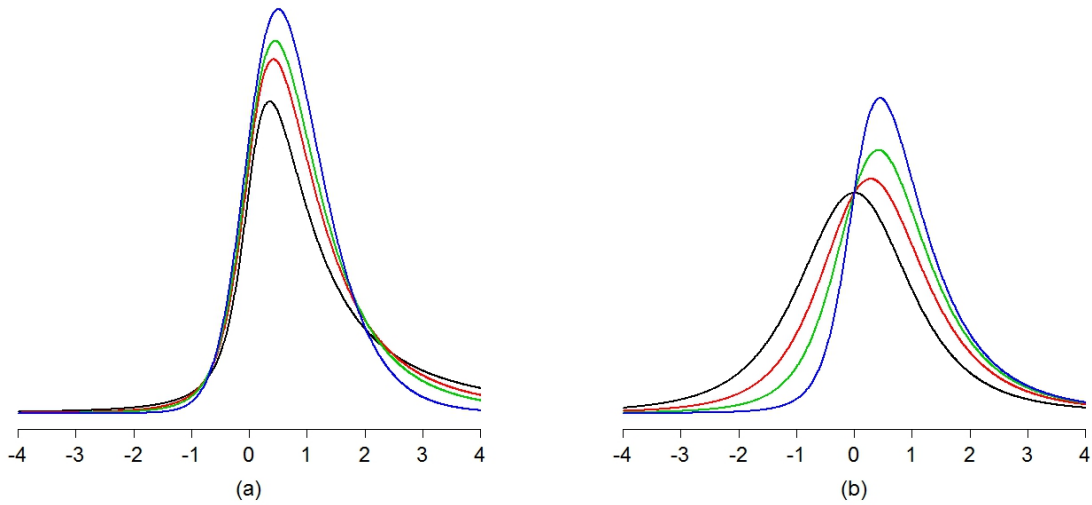


Figura 4 – Densidades da distribuição t de Student assimétrica com parâmetros $\xi = 0$, $\omega = 1$, e (a) $\alpha = 2$; $\nu = 1$ (linha em preto), $\nu = 2$ (linha em vermelho), $\nu = 3$ (linha em verde), $\nu = 10$ (linha em azul), e (b) $\nu = 3$; $\alpha = 0$ (linha em preto - distribuição t de Student com 3 graus de liberdade), $\alpha = 0.5$ (linha em vermelho), $\alpha = 1$ (linha em verde), $\alpha = 2$ (linha em azul).

Observamos, por meio da Figura 4(a), que o parâmetro graus de liberdade, ν , controla a dispersão, e por meio da Figura 4(b), que o parâmetro α controla a assimetria da distribuição. A Figura 4(b), também apresenta a densidade de uma distribuição t de Student com três graus de liberdade (linha em vermelho).

As distribuições normal assimétrica, normal e Cauchy também são casos particulares da t de Student assimétrica, ou seja:

- (i) Se os graus de liberdade $\nu \rightarrow \infty$ em (2.9), então $X \sim NA(\alpha)$, como dada em (2.3);
- (ii) Se os graus de liberdade $\nu \rightarrow \infty$ em (2.9), então $X \sim NA(\alpha)$, como dada em (2.3), e se $\alpha = 0$ em (2.3), então $X \sim N(0, 1)$, como descrito na Propriedade 1 da Seção 2.3.
- (iii) Se $\alpha = 0$ em (2.9), temos que $X \sim t_\nu(0, 1)$, como descrito na Propriedade 1 da Seção 2.4, e se $\nu = 1$ então $X \sim t_1(0, 1)$, logo $X \sim Cauchy(0, 1)$ de acordo com Mood *et al.* (1974, p. 251).

2.4.1 Medidas características da distribuição t de Student assimétrica

Suponha Y uma variável aleatória com distribuição $TA_\nu(\xi, \omega^2, \alpha)$, a esperança de Y é dada por

$$E[Y] = \xi + \omega b_\nu \delta(\alpha), \text{ com } \nu > 1,$$

a variância de Y é

$$Var[Y] = \omega^2 \left[\frac{\nu}{\nu - 2} - b_\nu^2 \delta^2(\alpha) \right], \text{ com } \nu > 2,$$

os coeficientes de assimetria e curtose são, respectivamente, dados por

$$\gamma_1 = \frac{b_v \delta(\alpha)}{(1 - (2/\pi)\delta^2(\alpha))^{3/4}} \left[\frac{v(3 - \delta^2(\alpha))}{v-3} - \frac{3v}{v-2} + 2b_v^2 \delta^2(\alpha) \right], \text{ com } v > 3,$$

e

$$\gamma_2 = \frac{1}{(1 - (2/\pi)\delta^2(\alpha))^2} \times \left[\frac{3v^2}{(v-2)(v-4)} - \frac{4b_v^2 \delta^2(\alpha)v(3 - \delta^2(\alpha))}{v-3} + \frac{6b_v^2 \delta^2(\alpha)v}{v-2} - 3b_v^4 \delta^4(\alpha) \right] - 3, \text{ com } v > 4,$$

em que $b_v = \frac{\sqrt{v} \Gamma(\frac{v-1}{2})}{\sqrt{\pi} \Gamma(\frac{v}{2})}$ e $\delta(\alpha) = \frac{\alpha}{\sqrt{1+\alpha^2}}$.

Maiores detalhes sobre os cálculos das medidas características da distribuição t de Student assimétrica ver [Azzalini \(2013, p. 103-104\)](#).

2.4.2 Casos particulares das distribuições normais e t de Student assimétricas

Sabemos que as distribuições normal e t de Student assimétricas compreendem algumas distribuições como casos particulares, o [Quadro 1](#) apresenta esses casos.

Quadro 1 – Casos particulares das distribuições em estudo

Distribuições	Normal assimétrica $NA(\xi, \omega^2, \alpha)$	t de Student assimétrica $TA_v(\xi, \omega^2, \alpha)$
Cauchy	-	$TA_1(\xi, \omega^2, 0)$
Normal	$NA(\xi, \omega^2, 0)$	$TA_\infty(\xi, \omega^2, 0)$
Normal assimétrica	-	$TA_\infty(\xi, \omega^2, \alpha)$
t de Student	-	$TA_v(\xi, \omega^2, 0)$

Observamos por meio do [Quadro 1](#) que a distribuição de Cauchy é um caso particular da distribuição t de Student assimétrica quando $\alpha = 0$ e $v = 1$, a distribuição normal é um caso particular da distribuição normal assimétrica quando $\alpha = 0$ e também da distribuição t de Student assimétrica quando $\alpha = 0$ e $v \rightarrow \infty$, a distribuição normal assimétrica é um caso particular da distribuição t de Student assimétrica $v \rightarrow \infty$ e a distribuição t de Student é um caso particular da distribuição t de Student assimétrica quando $\alpha = 0$.

MODELOS DE MISTURA

Neste capítulo, definimos mistura de distribuições, apresentamos a estrutura de um modelo de mistura de distribuições e também a sua função de verossimilhança. Além disso, evidenciamos a questão da estimação do número de componentes e da identificabilidade de modelos de mistura.

3.1 Modelos de mistura de distribuições

Um modelo de mistura é um modelo probabilístico utilizado para representar a existência de subgrupos ou subpopulações dentro de uma população, sem exigir que o conjunto de dados observados identifique o grupo à qual pertence cada observação individual. Mais precisamente um modelo de mistura corresponde à distribuição de mistura que representa a distribuição de probabilidade das observações da população.

A distribuição associada à cada subgrupo da população é denominada de componente de mistura, e é ponderada pela frequência relativa desse subgrupo. Essa frequência relativa é denominada de proporções ou pesos da componente de mistura.

A seguir, apresentamos uma definição formal de mistura de distribuições.

Definição 1. (SARAIVA, 2009) Qualquer combinação linear convexa

$$\sum_{k=1}^K p_k f_Y(y|\theta_k), \text{ com } p_k > 0 \text{ e } \sum_{k=1}^K p_k = 1. \quad (3.1)$$

em que p_k são as proporções ou pesos da componentes de mistura, $f_Y(y|\theta_k)$ são as densidades de mistura, pertencentes à uma família de distribuições indexadas pelo parâmetro θ_k , que pode ser um vetor ou um escalar, é denominada de uma mistura de distribuições.

Segundo Frühwirth-Schnatter (2006)[p. 1] para descrever um modelo de mistura, consideramos uma população composta por K subgrupos, também denominados de componentes

de mistura, com proporções ou pesos p_1, p_2, \dots, p_K relativos ao tamanho das componentes, e supomos que o interesse está em uma variável aleatória \mathbf{Y} que é heterogênea dentro da população e homogênea dentro das componentes. Devido à essa heterogeneidade, \mathbf{Y} tem uma distribuição de probabilidade diferente em cada componente, mas pertencente a uma mesma família paramétrica de distribuições, $f_Y(\mathbf{y}|\boldsymbol{\theta})$, porém com valores do parâmetro $\boldsymbol{\theta}$ (vetor ou escalar) diferente para cada componente.

As componentes podem ser caracterizadas por uma variável indicadora discreta \mathbf{S} , com $\mathbf{s} \in \{1, \dots, K\}$, assim ao fazermos uma amostragem aleatória de tal população, podemos registrar não apenas \mathbf{Y} , mas também a variável indicadora \mathbf{S} . A probabilidade de amostragem da componente s é denotada por p_s , ao passo que condicional ao conhecimento de s , a variável aleatória \mathbf{Y} segue a distribuição $f_Y(\mathbf{y}|\boldsymbol{\theta}_s)$, com $\boldsymbol{\theta}_s$ sendo o parâmetro da componente s . Então, a densidade conjunta de $f_Y(\mathbf{y}, \mathbf{s})$ é dada por

$$f_Y(\mathbf{y}, \mathbf{s}) = f_S(\mathbf{s})f_Y(\mathbf{y}|\mathbf{s}) = p_s f_Y(\mathbf{y}|\boldsymbol{\theta}_s).$$

Logo, temos uma mistura finita de distribuição, pois o que observamos é apenas a variável aleatória \mathbf{Y} , e a densidade marginal $f_Y(\mathbf{y})$ é dada pela seguinte densidade de mistura

$$f_Y(\mathbf{y}) = \sum_{s=1}^K f_Y(\mathbf{y}, \mathbf{s}) = \sum_{s=1}^K p_s f_Y(\mathbf{y}|\boldsymbol{\theta}_s) = p_1 f_Y(\mathbf{y}|\boldsymbol{\theta}_1) + \dots + p_K f_Y(\mathbf{y}|\boldsymbol{\theta}_K),$$

em que $f_Y(\mathbf{y}|\boldsymbol{\theta}_s)$ é a distribuição da s -ésima componente de mistura com parâmetro $\boldsymbol{\theta}_s$ e p_s é o s -ésimo peso ou probabilidade de mistura, com $0 \leq p_s \leq 1$, para $s \in \{1, \dots, K\}$.

Para construir um modelo de mistura de distribuições [Saraiva \(2009\)](#) considera $\mathbf{y} = (y_1, \dots, y_N)'$ uma amostra aleatória de uma variável aleatória \mathbf{Y} , com $\mathbf{Y} = (Y_1, \dots, Y_N)'$, proveniente de uma densidade com mistura de distribuições com K componentes, ou seja,

$$f_Y(y_i|\boldsymbol{\theta}, \mathbf{p}) = \sum_{k=1}^K p_k f_Y(y_i|\boldsymbol{\theta}_k), \quad (3.2)$$

em que $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K)'$ é o vetor de parâmetros e $\mathbf{p} = (p_1, \dots, p_K)'$ é o vetor de pesos ou probabilidade de mistura, p_k e $f_Y(y_i|\boldsymbol{\theta}_k)$ são, respectivamente, o peso e a distribuição de probabilidade indexada pelo parâmetro $\boldsymbol{\theta}_k$ (vetor ou escalar), ambos associados a k -ésima componente de mistura, com $p_k > 0$ e $\sum_{k=1}^K p_k = 1$, para $i = 1, \dots, N$, em que N é o total de observações, com $k = 1, \dots, K$ e $K > 1$.

Para explorar a estrutura “oculta” do modelo de mistura, [Saraiva \(2009\)](#) considera \mathbf{S} como uma variável aleatória latente ou não observável $\mathbf{S} = (\mathbf{S}_1, \dots, \mathbf{S}_N)'$, com $i = 1, \dots, N$ e cada $\mathbf{S}_i = (S_{i1}, \dots, S_{iK})'$, com dimensão K , indica qual componente a observação y_i é proveniente, ou seja,

$$S_{ik} = \begin{cases} 1, & \text{se } y_i \text{ é proveniente da componente } k \\ 0, & \text{caso contrário,} \end{cases}$$

com $\sum_{k=1}^K S_{ik} = 1$, para $i = 1, \dots, N$ e $k = 1, \dots, K$.

Assim,

$$y_i | S_{ik} = 1, \boldsymbol{\theta}_k \sim f_Y(y_i | \boldsymbol{\theta}_k) \text{ e } \mathbf{S}_i \sim \text{Multinomial}_K(1, \mathbf{p} = (p_1, \dots, p_K)), \quad (3.3)$$

ou seja, $y_i | S_{ik} = 1, \boldsymbol{\theta}_k$ segue uma distribuição de probabilidade $f_Y(y_i | \boldsymbol{\theta}_k)$, e \mathbf{S}_i segue uma distribuição multinomial com K modalidades e uma única observação, em que $\mathbf{p} = (p_1, \dots, p_K) \in \mathfrak{R}_+^K$ é o vetor de pesos.

Os valores S'_i s devem ou não ser parte das quantidades de interesse, e condicional em S_i , as observações independentes $\mathbf{y} = (y_1, \dots, y_N)'$ são provenientes das densidades

$$P(y_i | S_{ik} = 1, \boldsymbol{\theta}_k) = f_Y(y_i | \boldsymbol{\theta}_k), \quad (3.4)$$

para $i = 1, \dots, N$ e $k = 1, \dots, K$.

Integrando sobre as variáveis não observáveis S_1, \dots, S_N Saraiva (2009) obtém o modelo dado em (3.2)

$$P(y_i | \boldsymbol{\theta}, \mathbf{p}) = \sum_{k=1}^K P(S_{ik} = 1) P(y_i | S_{ik} = 1, \boldsymbol{\theta}_k) = \sum_{k=1}^K p_k f_Y(y_i | \boldsymbol{\theta}_k). \quad (3.5)$$

Neste trabalho, para descrever o modelo de mistura de distribuições com um número de componentes K conhecido, consideramos \mathbf{Y} como um vetor aleatório ou uma variável aleatória contínua independente univariada, com $\mathbf{Y} = (Y_1, \dots, Y_N)'$, e \mathbf{S} uma variável indicadora ou um vetor de variáveis aleatórias discreta não observável com $\mathbf{S} = (S_1, \dots, S_N)'$, em que cada S_i possui distribuição multinomial, isto é, $S_i \sim \text{Multinomial}_K(1, \mathbf{p} = (p_1, \dots, p_K))$, com K resultados possíveis e um observação, com $i = 1, \dots, N$ e $\sum_{k=1}^K p_k = 1$.

Desta forma, temos que a distribuição de cada variável aleatória independente Y_i é controlada pelo valor assumido da variável aleatória não observável S_i correspondente, isto é, $Y_i | S_i = k, \boldsymbol{\theta}$ tem distribuição condicional $f_{Y_i}(y | S_i = s_i, \boldsymbol{\theta})$, em que $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_K)'$ é o vetor de parâmetros com $\boldsymbol{\theta} \in \mathfrak{R}^K$, e $\boldsymbol{\theta}_k$ é o parâmetro (vetor ou escalar) associado a k -ésima componente de mistura, para $k = 1, \dots, K$.

Logo, a distribuição marginal de Y_i pode ser escrita como

$$f_{Y_i}(\mathbf{y} | \boldsymbol{\theta}, \mathbf{p}) = \sum_{k=1}^K f_{Y_i}(\mathbf{y} | S_i = k, \boldsymbol{\theta}_k) Pr(S_i = k | \mathbf{p}), \quad (3.6)$$

em que $f_{Y_i}(\mathbf{y} | S_i = k, \boldsymbol{\theta}_k)$ é a função densidade de probabilidade de Y_i condicional a $S_i = k$, que em geral, pertencem à mesma família paramétrica de distribuições, mas com valores de parâmetros diferentes, também denominada de distribuição de probabilidade da k -ésima componente do modelo de mistura, e $Pr(S_i = k | \mathbf{p}) = p_k$ é o peso ou a probabilidade de mistura associado a k -ésima componente mistura para $k = 1, \dots, K$, com $\sum_{k=1}^K p_k = 1$ e $\mathbf{p} = (p_1, \dots, p_K)'$ é o vetor de pesos.

Note que a distribuição marginal de Y_i dada em (3.6) é uma mistura de distribuições com K componentes, em que $f_{Y_i}(\mathbf{y}|S_i = k, \boldsymbol{\theta}_k)$ corresponde as densidades de mistura e $Pr(S_i = k|\mathbf{p})$ corresponde ao p_k , ambos dados na Expressão (3.1), do tradicional modelo de mistura de distribuições.

3.1.1 FUNÇÃO DE VEROSSIMILHANÇA

Para construir a função de verossimilhança, consideramos a sequência de variáveis aleatórias não observáveis S_i , como descrito na Seção 3.1, ou seja, como uma sequência de variáveis aleatórias idênticas independentes, com mesma distribuição (*iid*) multinomial com K resultados possíveis, uma observação e $\mathbf{p} = (p_1, \dots, p_k)'$, tal que $S_i = s_i$, com $s_i \in \{1, \dots, K\}$ e $i = 1, \dots, N$.

Logo, a distribuição multinomial de S_i pode ser expressa por

$$Pr(S_i = s_i|\mathbf{p}) = \prod_{k=1}^K p_k^{I_{\{s_i\}}(k)}, \quad (3.7)$$

em que $I_{\{s_i\}}(k)$ é a função indicadora que associa Y_i a k -ésima componente de mistura, isto é, assume *um* se $(s_i = k)$ for verdadeira e *zero* caso contrário, e p_k é o peso associado a k -ésima componente mistura, com $\sum_{k=1}^K p_k = 1$ para $k = 1, \dots, K$.

Além disso, temos também que

$$f_{Y_i}(\mathbf{y}|S_i = s_i, \boldsymbol{\theta}) = \prod_{k=1}^K f_{Y_i}(\mathbf{y}|S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)}. \quad (3.8)$$

A distribuição conjunta de Y_i e S_i é dada pelo produto das expressões (3.7) e (3.8), isto é,

$$\begin{aligned} f_{Y_i}(\mathbf{y}, S_i = s_i|\boldsymbol{\theta}, \mathbf{p}) &= f_{Y_i}(\mathbf{y}|S_i = s_i, \boldsymbol{\theta})Pr(S_i = s_i|\mathbf{p}) \\ &= \prod_{k=1}^K p_k^{I_{\{s_i\}}(k)} f_{Y_i}(\mathbf{y}|S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)}. \end{aligned}$$

Logo, a distribuição conjunta de $\mathbf{Y} = (Y_1, \dots, Y_N)'$ e $\mathbf{S} = (S_1, \dots, S_N)'$ é dada por

$$f_{\mathbf{Y}}(\mathbf{y}, \mathbf{s}|\boldsymbol{\theta}, \mathbf{p}) = \prod_{i=1}^N \left\{ \prod_{k=1}^K p_k^{I_{\{s_i\}}(k)} f_{Y_i}(\mathbf{y}|S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)} \right\}.$$

Considerando os dados observados $\mathbf{y} = (y_1, \dots, y_N)'$ uma realização de $\mathbf{Y} = (Y_1, \dots, Y_N)'$, e $\mathbf{s} = (s_1, \dots, s_N)'$ uma realização de $\mathbf{S} = (S_1, \dots, S_N)'$, temos que a função de verossimilhança

da mistura de distribuições de \mathbf{y} e \mathbf{s} é dada por

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) &= \prod_{i=1}^N \left\{ \prod_{k=1}^K p_k^{I_{\{s_i\}}(k)} f_{Y_i}(y | S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)} \right\} \\
&= \prod_{k=1}^K \left\{ \prod_{i=1}^N p_k^{I_{\{s_i\}}(k)} f_{Y_i}(y | S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)} \right\} \\
&= \prod_{k=1}^K p_k^{\sum_{i=1}^N I_{\{s_i\}}(k)} \left\{ \prod_{i=1}^N f_{Y_i}(y | S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)} \right\} \\
&= \prod_{k=1}^K p_k^{n_k} \left\{ \prod_{i=1}^N f_{Y_i}(y | S_i = k, \boldsymbol{\theta}_k)^{I_{\{s_i\}}(k)} \right\}, \tag{3.9}
\end{aligned}$$

em que $n_k = \sum_{i=1}^N I_{\{s_i\}}(k)$ é o número de observações da k -ésima componente de mistura.

A obtenção dos EMV's utiliza a função log-verossimilhança e o multiplicador de Lagrange, devido a restrição $\sum_{k=1}^K p_k = 1$. Logo a função de verossimilhança é dada por

$$\begin{aligned}
l(\boldsymbol{\theta}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) &= \log(\mathcal{L}(\boldsymbol{\theta}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K)) \\
&= \sum_{k=1}^K n_k \log(p_k) + \sum_{i=1}^N \sum_{k=1}^K I_{\{s_i\}}(k) \log(f_{Y_i}(y | S_i = k, \boldsymbol{\theta}_k)) + \lambda \left(1 - \sum_{k=1}^K p_k \right).
\end{aligned}$$

Assim, para um número de componentes K conhecido é usual utilizar o algoritmo EM (DEMPSTER *et al.*, 1977) para obter os EMV's para os parâmetros do modelo. No entanto, o algoritmo EM apresenta, em geral, problemas de convergência a um máximo local, uma vez que a estimação do máximo da função de verossimilhança ou da função log verossimilhança em modelos de mistura são para encontrar o máximo global da função numericamente (FRÜHWIRTH-SCHNATTER, 2006).

Já para o número de componentes K desconhecido é usual fixar diferentes valores para K , e comparar os modelos segundo algum critério, como AIC e BIC. Apesar desses critérios frequentistas serem conceitualmente diferentes, eles também utilizam o máximo da função de verossimilhança ou da função log verossimilhança como medida de ajustamento.

Desta forma, nos procedimentos de estimação frequentistas as estimativas dos parâmetros dos modelos são obtidas para K conhecido, já os métodos bayesianos são baseados em simulação MCMC, sendo possível propor alterações no valor de K e estimar os parâmetros do modelo, durante as iterações, o que torna a abordagem Bayesiana mais atrativa.

3.2 Número de componentes

O número de componentes K de um modelo de mistura pode ser conhecido ou não, e metodologias para estimar o valor de K quando ele é desconhecido, são de interesse da comunidade científica.

Gelman *et al.* (2013)[p. 536] sugere, no contexto clássico, aplicar o algoritmo EM para obter uma estimativa de máxima verossimilhança para os parâmetros segundo cada escolha de K , e relatar o resultado de K como aquele com a melhor probabilidade marginal estimada. Na abordagem Bayesiana, os autores indicam a distribuição *a priori* para K como uma Poisson truncada e o algoritmo MCMC *Reversible jump* (GREEN, 1995) para calcular a distribuição *a posteriori* para K . Os autores afirmam que tal abordagem é computacionalmente intensiva, e que na prática, é mais comum ajustar o modelo para várias opções de K , e em seguida usar um ajuste adequado.

Sob o enfoque bayesiano, encontramos na literatura algumas metodologias para estimar K . Diebolt e Robert (1994) apresentam métodos de aproximação que avaliam a distribuição *a posteriori* de K pelo algoritmo MCMC *Gibbs sampling* (GELMAN *et al.*, 2013), os autores consideram que K é um valor fixo. Richardson e Green (1997) propõem o algoritmo MCMC *Reversible jump* (GREEN, 1995), segundo os autores esse algoritmo é capaz de saltar entre os subespaços de parâmetros correspondentes a diferentes números de componentes na mistura, e nesta metodologia o valor de K é desconhecido. Stephens (2000) descreve um algoritmo MCMC alternativo ao *Reversible jump* (GREEN, 1995), que vê os parâmetros do modelo como um processo de ponto (marcado), criando um processo de *nascimento-morte* de uma cadeia de Markov com uma distribuição estacionária apropriada. Saraiva (2009) propõe o algoritmo MCMC *split-merge*, nesse algoritmo o movimento *split* faz uma partição nos dados, e condicional a esta partição novos parâmetros são propostos para as componentes, já o movimento *merge* junta duas componentes e também novos parâmetros são propostos. Neste dois últimos trabalhos o valor de K também é considerado desconhecido.

Em resumo, para estimar K temos duas alternativas a estimação frequentista e a abordagem Bayesiana. A vantagem da abordagem Bayesiana é que esta é mais flexível e podemos desenvolver um procedimento de estimação conjunta de K e dos parâmetros θ_K , o que pelos procedimentos frequentistas não é possível. Isto é, na abordagem Bayesiana, os métodos MCMC permitem propor alterações no valor de K durante as iterações, e então os parâmetros das componentes e as variáveis não observáveis são atualizados condicional em um valor de K . Dessa forma, é possível fazer inferências sobre K , quando este valor é desconhecido, e ao mesmo tempo verificar qual é o melhor ajuste tomando como base a distribuição *a posteriori* para K .

Neste trabalho, utilizamos a abordagem Bayesiana para fazer inferências sobre K , pois a estimativa da probabilidade *a posteriori* para K proporciona uma forma direta de selecionar um modelo de um conjunto de modelos competidores, de diferentes dimensões. Mais precisamente, aplicamos o algoritmo MCMC *Reversible jump* (GREEN, 1995) e propomos o algoritmo MCMC *Data-driven Reversible jump*. Descrevemos estas metodologias, respectivamente, na Seção 4.4 e na Seção 4.5 do Capítulo 4.

3.3 Identificabilidade de um modelo de mistura de distribuições

Uma questão importante no processo de estimação dos modelos de mistura é a identificabilidade do modelo.

Segundo Frühwirth-Schnatter (2006)[p. 15] uma família paramétrica de distribuições, indexada por um parâmetro $\boldsymbol{\theta} \in \Theta$ definido sobre um espaço amostral de observações \mathcal{Y} , é dito, identificável se para quaisquer dois parâmetros $\boldsymbol{\theta} \in \Theta$ e $\boldsymbol{\theta}^* \in \Theta$ definem a mesma lei de probabilidade em \mathcal{Y} , então $\boldsymbol{\theta}$ e $\boldsymbol{\theta}^*$ são idênticos.

Em outras palavras, se a função densidade $f(\mathbf{y}|\boldsymbol{\theta})$ e a função densidade $f(\mathbf{y}|\boldsymbol{\theta}^*)$ são idênticas para quase todo $\mathbf{y} \in \mathcal{Y}$, então os parâmetros $\boldsymbol{\theta}$ e $\boldsymbol{\theta}^*$ precisam ser idênticos, isto é,

$$f(\mathbf{y}|\boldsymbol{\theta}) = f(\mathbf{y}|\boldsymbol{\theta}^*), \text{ para quase todo } \mathbf{y} \in \mathcal{Y} \rightarrow \boldsymbol{\theta} = \boldsymbol{\theta}^*.$$

Assim, se para quaisquer dois parâmetros $\boldsymbol{\theta}$ e $\boldsymbol{\theta}^*$ distintos, e as distribuições $f(\mathbf{y}|\boldsymbol{\theta})$ e $f(\mathbf{y}|\boldsymbol{\theta}^*)$ são idênticas para quase todo $\mathbf{y} \in \mathcal{Y}$, então esta família de distribuições não é identificável.

Em geral, para um modelo de mistura de distribuição finita, existem dois tipos de não identificabilidade.

1. Devido à invariância de reclassificação das componentes de mistura;
2. Devido ao potencial sobre ajuste de componentes;

3.3.1 NÃO IDENTIFICABILIDADE DEVIDO À INVARIÂNCIA DE RECLASSIFICAÇÃO DAS COMPONENTES DE MISTURA

A não identificabilidade devido à invariância de reclassificação é causada pela invariância de uma distribuição de mistura para reclassificar as componentes, em outras palavras, é a invariância com relação à mudança na etiqueta das componentes.

Para elucidar, vamos considerar um modelo de mistura de distribuições normais com duas componentes, isto é,

$$f_Y(y|\boldsymbol{\varphi}) = p_1 f_Y(y|\mu_1, \sigma_1^2) + p_2 f_Y(y|\mu_2, \sigma_2^2),$$

em que $\boldsymbol{\theta}_k = (\mu_k, \sigma_k^2)' \in \Theta = \mathfrak{R} \times \mathfrak{R}^+$, $p_k \in (0, 1)$, com $k = 1, 2$ e $\boldsymbol{\varphi} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, p_1, p_2)' \in \Theta_2$.

Tomemos um parâmetro $\boldsymbol{\varphi}^* \in \Theta_2$ arbitrário, com $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$, e vamos definir o parâmetro $\boldsymbol{\varphi}^* = (\boldsymbol{\theta}_2, \boldsymbol{\theta}_1, p_2, p_1)'$, ou seja, trocando a ordem dos componentes.

Temos que a distribuição de $\boldsymbol{\varphi}$ e $\boldsymbol{\varphi}^*$ é a mesma, embora evidentemente os dois parâmetros sejam distintos, isto é,

$$\begin{aligned} f_Y(y|\boldsymbol{\varphi}^*) &= p_2 f_Y(y|\mu_2, \sigma_2^2) + p_1 f_Y(y|\mu_1, \sigma_1^2) \\ &= p_1 f_Y(y|\mu_1, \sigma_1^2) + p_2 f_Y(y|\mu_2, \sigma_2^2) = f_Y(y|\boldsymbol{\varphi}). \end{aligned}$$

Devido à essa invariância, uma mistura de duas distribuições normais não é identificável (FRÜHWIRTH-SCHNATTER, 2006)[p. 16].

3.3.2 NÃO IDENTIFICABILIDADE DEVIDO AO POTENCIAL DE SOBRE AJUSTE

Suponha uma mistura finita de distribuições com K componentes definida num espaço paramétrico Θ_K . Qualquer mistura finita com $K - 1$ componentes definida no mesmo espaço paramétrico, Θ_K , ou uma das componentes está vazia ou duas componentes são iguais.

Vamos exemplificar novamente, considerando um modelo de mistura de distribuições normais com duas componentes. Tomemos um parâmetro $\boldsymbol{\varphi}_2 = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, p_1, p_2)' \in \Theta_2$ arbitrário. Sabemos que qualquer mistura de distribuições com duas componentes pode ser escrita como uma mistura de distribuições com 3 componentes, em que a terceira componente possui peso ou proporção nula, $p_3 = 0$, ou seja,

$$\begin{aligned} f_Y(y|\boldsymbol{\varphi}) &= p_1 f_Y(y|\mu_1, \sigma_1^2) + p_2 f_Y(y|\mu_2, \sigma_2^2) \\ &= p_1 f_Y(y|\mu_1, \sigma_1^2) + p_2 f_Y(y|\mu_2, \sigma_2^2) + 0 \times f_Y(y|\mu_3, \sigma_3^2). \end{aligned}$$

Temos que o parâmetro $\boldsymbol{\varphi} = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \mu_3, \sigma_3^2, p_1, p_2, 0)' \in \Theta_3$ correspondente a mistura dada acima, encontra-se em um conjunto de não identificabilidade, pois a densidade $f_Y(y|\boldsymbol{\varphi})$ é a mesma para valores arbitrários de μ_3 e σ_3^2 .

O mesmo conjunto de não identificabilidade pode ser encontrado em uma mistura de distribuições normais com duas componentes, sendo que a terceira é gerada da divisão de uma componente em duas, isto é

$$\begin{aligned} f_Y(y|\boldsymbol{\varphi}) &= p_1 f_Y(y|\mu_1, \sigma_1^2) + p_2 f_Y(y|\mu_2, \sigma_2^2) \\ &= p_1 f_Y(y|\mu_1, \sigma_1^2) + (p_2 - p_3) f_Y(y|\mu_2, \sigma_2^2) + p_3 f_Y(y|\mu_3, \sigma_3^2). \end{aligned}$$

Novamente, o parâmetro $\boldsymbol{\varphi} = (\mu_1, \sigma_1^2, \mu_2, \sigma_2^2, \mu_3, \sigma_3^2, p_1, p_2 - p_3, p_3)' \in \Theta_3$ encontra-se em um conjunto de não identificabilidade, pois a densidade $f_Y(y|\boldsymbol{\varphi})$ é a mesma para valores arbitrários de p_3 , com $0 \leq p_3 \leq p_2$ (FRÜHWIRTH-SCHNATTER, 2006)[p. 17].

Maiores detalhes sobre problemas de identificabilidade de modelos de mistura de distribuições podem ser encontrados em Teicher (1963), Yakowitz e Spragins (1968), Chandra (1977), Redner e Walker (1984) e Crawford (1994).

ABORDAGEM BAYESIANA

Neste capítulo, descrevemos a abordagem Bayesiana por meio dos métodos de Monte Carlo via cadeias de Markov (MCMC) empregados para fazer inferências, para os modelos de mistura de distribuições assimétricas em estudo, compreendendo o processo de estimação conjunta do número de componentes dos modelos de mistura e dos parâmetros associados as componentes de mistura.

4.1 Distribuições *a priori* e *a posteriori*

A distribuição *a priori* conjunta para os parâmetros de um modelo de mistura, com um número de componentes K desconhecido e com independência na variável não observável é estabelecida considerando independência entre os parâmetros $\boldsymbol{\theta}$ e \mathbf{p} , isto é,

$$\begin{aligned}\pi(\boldsymbol{\theta}, \mathbf{p}, K) &= \pi(\boldsymbol{\theta}|\mathbf{p}, K)\pi(\mathbf{p}|K)\pi(K) \\ &= \pi(\boldsymbol{\theta}|K)\pi(\mathbf{p}|K)\pi(K).\end{aligned}\tag{4.1}$$

Considerando a k -ésima componente do modelo de mistura, a distribuição *a priori* pode ser escrita como

$$\pi(\boldsymbol{\theta}_k, p_k, k) = \pi(\boldsymbol{\theta}_k)\pi(p_k)\pi(k)$$

A distribuição *a posteriori* a menos de valores constantes é construída da seguinte forma

$$\begin{aligned}\pi(\boldsymbol{\theta}, \mathbf{p}, K|\mathbf{y}, \mathbf{s}) &\propto P(\mathbf{y}, \mathbf{s}; \boldsymbol{\theta}, \mathbf{p}, K) \\ &\propto P(\mathbf{y}, \mathbf{s}|\boldsymbol{\theta}, \mathbf{p}, K)\pi(\boldsymbol{\theta}, \mathbf{p}, K) \\ &\propto P(\mathbf{y}, \mathbf{s}|\boldsymbol{\theta}, \mathbf{p}, K)\pi(\boldsymbol{\theta}|K)\pi(\mathbf{p}|K)\pi(K) \\ &\propto P(\mathbf{y}|\boldsymbol{\theta}, \mathbf{p}, K)P(\mathbf{S} = \mathbf{s}|\mathbf{p}, K)\pi(\boldsymbol{\theta}|K)\pi(\mathbf{p}|K)\pi(K)\end{aligned}$$

em que $P(\mathbf{y}|\boldsymbol{\theta}, \mathbf{p}, K)P(\mathbf{S} = \mathbf{s}|\mathbf{p}, K) = \mathcal{L}(\boldsymbol{\theta}, \mathbf{p}|\mathbf{y}, \mathbf{s}, K)$ é a função de verossimilhança para os parâmetros $\boldsymbol{\theta}$ e \mathbf{p} , condicional nos dados completos.

No caso dos modelos de mistura de distribuições, a função de verossimilhança é igual à distribuição conjunta como dada na Expressão (3.9) do Capítulo 3, porém considerada como uma função dos parâmetros desconhecidos.

Desta forma, a distribuição *a posteriori* pode ser escrita como

$$\pi(\boldsymbol{\theta}, \mathbf{p}, K | \mathbf{y}, \mathbf{s}) \propto \mathcal{L}(\boldsymbol{\theta}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) \pi(\boldsymbol{\theta} | K) \pi(\mathbf{p} | K) \pi(K). \quad (4.2)$$

A distribuição *a posteriori* da k -ésima componente do modelo de mistura é dada por

$$\pi(\theta_k, p_k | \mathbf{y}, \mathbf{s}) \propto \mathcal{L}(\theta_k, p_k | \mathbf{y}, \mathbf{s}) \pi(\theta_k) \pi(p_k) \pi(k).$$

A distribuição *a posteriori*, em geral, é muito complexa analiticamente, mas pode ser aproximada utilizando técnicas de simulação como o método de Monte Carlo via cadeias de Markov (MCMC).

Quando trabalhamos com cadeias de Markov construímos uma cadeia que tem a distribuição desejada como distribuição de equilíbrio, e então, podemos obter uma amostra dessa distribuição registrando os estados da cadeia.

Existem vários processos para a construção de uma cadeia de Markov, sendo que neste trabalho utilizamos os seguintes métodos MCMC: o algoritmo *Metropolis-Hastings* para atualizar as estimativas dos parâmetros do modelo de mistura de distribuições normais, normais assimétricas e t de Student assimétricas; o algoritmo *Gibbs sampling* para atualizar as variáveis não observáveis; e os métodos *Reversible jump* e *Data-driven Reversible jump*, ambos para selecionar o número de componentes e conjuntamente estimar os parâmetros associados as componentes dos modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas. Sendo o método *Data-driven Reversible jump* a nossa proposta alternativa de aperfeiçoamento do método *Reversible jump*.

4.2 Algoritmo *Metropolis-Hastings*

O algoritmo *Metropolis-Hastings*, proposto por [Metropolis et al. \(1953\)](#) e [Hastings \(1970\)](#), é uma adaptação de um passeio aleatório ([CHIB; GRRENBURG, 1995](#)) com uma regra de aceitação ou rejeição de convergência para a distribuição de equilíbrio, que em nosso caso é a distribuição *a posteriori*.

No algoritmo *Metropolis-Hastings*, um valor é gerado de uma distribuição auxiliar, também denominada de proposta, e aceito com uma dada probabilidade, garantindo a convergência da cadeia para a distribuição de equilíbrio.

Suponha que a cadeia está no estado $\boldsymbol{\theta}$ e um valor $\boldsymbol{\theta}'$ é gerado de uma *distribuição proposta*, $q(\boldsymbol{\theta}' | \boldsymbol{\theta})$, em que essa distribuição pode depender ou não do estado atual da cadeia. O

novo valor θ' gerado de $q(\theta'|\theta)$ é aceito com probabilidade $\alpha(\theta, \theta') = \min\{1, A\}$, em que A é dado pela expressão

$$A = \frac{\pi(\theta')q(\theta|\theta')}{\pi(\theta)q(\theta'|\theta)}, \quad (4.3)$$

e $\pi(\theta)$ dado em (4.3) é a distribuição de interesse.

O algoritmo *Metropolis-Hastings* pode ser descrito nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$, e especifique um valor inicial para $\theta^{(0)}$;
2. Gere um valor θ' da distribuição $q(\theta'|\theta)$;
3. Calcule a probabilidade de aceitação $\alpha(\theta, \theta')$, e gere um valor $u \sim U(0, 1)$;
4. Se $u \leq \alpha(\theta, \theta')$ aceite o novo valor, e faça $\theta^{(t+1)} = \theta'$, caso contrário rejeite e faça $\theta^{(t+1)} = \theta^{(t)}$;
5. Mude o contador de t para $t + 1$, e volte ao passo 2 para $t = 1, \dots, T$.

Como, em geral, a distribuição de interesse é a distribuição *a posteriori*, temos que a expressão de A na probabilidade de aceitação, $\alpha(\theta, \theta') = \min\{1, A\}$, pode ser reescrita como

$$A = \frac{\pi(\theta')\mathcal{L}(\theta'|y)q(\theta|\theta')}{\pi(\theta)\mathcal{L}(\theta|y)q(\theta'|\theta)}, \quad (4.4)$$

em que $\pi(\theta)$ é a distribuição *a priori* e $\mathcal{L}(\theta|y)$ é a função de verossimilhança.

Neste trabalho, consideramos o caso particular do algoritmo *Metropolis-Hastings* em que as distribuições propostas são simétricas, isto é quando $q(\theta'|\theta) = q(\theta|\theta')$ para todos os valores de θ e θ' , e neste caso a expressão de A na probabilidade de aceitação, $\alpha(\theta, \theta') = \min\{1, A\}$, se reduz a

$$A = \frac{\pi(\theta')\mathcal{L}(\theta'|y)}{\pi(\theta)\mathcal{L}(\theta|y)}, \quad (4.5)$$

ou seja, o valor de A na Expressão (4.5) equivale ao quociente da distribuição *a posteriori* calculada sob o valor gerado θ' pelo valor da distribuição *a posteriori* calculado na t -ésima iteração.

4.3 Algoritmo Gibbs sampling

O algoritmo *Gibbs sampling* proposto por [Gelman et al. \(2013\)](#), é um método MCMC utilizado para obter amostras de uma distribuição de probabilidade.

A ideia do *Gibbs sampling* é gerar amostras da distribuição *a posteriori* para cada parâmetro ou em blocos de parâmetros com os demais parâmetros fixados em seus valores atuais. Neste algoritmo, a cadeia de Markov irá se mover para um novo valor, isto é, não existe mecanismo de aceitação ou rejeição ([GELMAN et al., 2013](#))[p. 276].

Suponha que o vetor de parâmetros $\boldsymbol{\theta}$ possui d componentes, ou seja, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)'$. A atualização de $\boldsymbol{\theta}$ é realizada para cada componente de $\boldsymbol{\theta}$ condicionado ao valor de todos os outros, logo existem d etapas na iteração t , e em cada iteração uma ordem dos d componentes de $\boldsymbol{\theta}$ é escolhido, e por sua vez, cada θ_j , ou seja, o j -ésimo componente de $\boldsymbol{\theta}$ é amostrado a partir da distribuição condicional completa, dado todos os outros componentes.

Em geral, cada um dos componentes θ_j pode ser uni ou multidimensional. Portanto, a distribuição condicional completa de θ_j , dada por $\pi(\theta_j | \boldsymbol{\theta}_{-j})$, em que $\boldsymbol{\theta}_{-j} = (\theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_d)'$, é obtida a partir da distribuição conjunta.

O algoritmo *Gibbs sampling* pode ser descrito nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$;
2. Especifique um valor inicial para $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})'$;
3. Obtenha um novo valor para $\boldsymbol{\theta}^{(t)}$ a partir de $\boldsymbol{\theta}^{(t-1)}$ por meio de gerações sucessivas dos valores

$$\begin{aligned} \theta_1^{(t)} &\sim \pi(\theta_1 | \theta_2^{(t-1)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}) \\ \theta_2^{(t)} &\sim \pi(\theta_2 | \theta_1^{(t)}, \theta_3^{(t-1)}, \dots, \theta_d^{(t-1)}) \\ &\vdots \\ \theta_d^{(t)} &\sim \pi(\theta_d | \theta_1^{(t)}, \theta_2^{(t)}, \dots, \theta_{d-1}^{(t)}) \end{aligned}$$

4. Mude o contador de t para $t + 1$, e volte ao passo 2 para $t = 1, \dots, T$.

O algoritmo *Gibbs sampling* é um caso particular do algoritmo *Metropolis-Hastings* em que os componentes de $\boldsymbol{\theta}$ são atualizados um de cada vez ou em blocos, tomando a distribuição condicional completa como proposta e a probabilidade de aceitação igual a 1 (GELMAN *et al.*, 2013)[p. 281].

4.4 Método *Reversible jump*

O método *Reversible jump*, introduzido por Green (1995), é um algoritmo utilizado para selecionar modelos e também para estimar o número de componentes dos modelos de mistura de distribuições. A vantagem deste algoritmo consiste na capacidade de saltar entre espaços paramétricos de dimensionalidade diferente, baseado na criação de uma cadeia de Markov que se move entre modelos com diferentes números de parâmetro.

Suponha que temos uma coleção contável de modelos candidatos $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_K$, e que cada modelo \mathcal{M}_k , com $k \in \{1, \dots, K\}$, possui vetor de parâmetros $\boldsymbol{\theta}_k \in \mathfrak{R}^{m_k}$, de dimensão m_k , em que m_k toma diferentes valores para diferentes modelos. Considerando os dados observados

\mathbf{y} , a distribuição conjunta de $(k, \boldsymbol{\theta}_k, \mathbf{y})$ é dada por

$$p(k, \boldsymbol{\theta}_k, \mathbf{y}) = p(\mathbf{y}|k, \boldsymbol{\theta}_k)p(\boldsymbol{\theta}_k|k)p(k).$$

Observamos que a distribuição conjunta de $(k, \boldsymbol{\theta}_k, \mathbf{y})$ é o produto da distribuição *a priori* pela função de verossimilhança, ou seja, a menos de valores constantes a distribuição conjunta de $(k, \boldsymbol{\theta}_k, \mathbf{y})$ pode ser vista como uma distribuição *a posteriori* de $(k, \boldsymbol{\theta}_k|\mathbf{y})$, isto é,

$$\pi(k, \boldsymbol{\theta}_k|\mathbf{y}) \propto \mathcal{L}(\mathbf{y}|k, \boldsymbol{\theta}_k)\pi(\boldsymbol{\theta}_k|k)\pi(k) \quad (4.6)$$

em que $\mathcal{L}(\mathbf{y}|k, \boldsymbol{\theta}_k)$ é a função de verossimilhança e $\pi(\boldsymbol{\theta}_k|k)\pi(k) = \pi(k, \boldsymbol{\theta}_k)$ é a distribuição *a priori* de $\boldsymbol{\theta}_k$ sobre o modelo \mathcal{M}_k .

O método *Reversible jump* usa a distribuição *a posteriori* como alvo de uma simulação MCMC sobre o espaço de estado $\Theta = \cup_{k \in K} (\{k\} \times \mathfrak{R}^{m_k})$, em que os estados da cadeia são da forma $(k, \boldsymbol{\theta}_k)$, cuja dimensão pode variar sobre o espaço de estados, isto é, o método *Reversible jump* pode ser visto como uma extensão do algoritmo *Metropolis-Hastings*, porém para espaços de estado mais gerais.

Para descrever o método *Reversible jump*, vamos supor que estamos atualmente no estado $(k, \boldsymbol{\theta}_k)$ no modelo \mathcal{M}_k e desejamos propor um movimento para um estado $(k', \boldsymbol{\theta}'_{k'})$ no modelo $\mathcal{M}_{k'}$, que possui dimensão maior, isto é, $m_{k'} > m_k$. Para fazer uma correspondência entre os dois estados dos modelos, geramos um vetor aleatório \mathbf{u} de dimensão $d_{k \rightarrow k'} = m_{k'} - m_k$ a partir de uma densidade conhecida $q_{d_{k \rightarrow k'}}(\mathbf{u})$. O estado atual $\boldsymbol{\theta}_k$ e o vetor aleatório \mathbf{u} são então mapeados para o novo estado $\boldsymbol{\theta}'_{k'} = g_{k \rightarrow k'}(\boldsymbol{\theta}_k, \mathbf{u})$ por meio de uma função de mapeamento $g_{k \rightarrow k'} : \mathfrak{R}^{m_k} \rightarrow \mathfrak{R}^{m_{k'}}$.

A probabilidade de aceitação desta proposta de mudança de estado é dado por $\min\{1, A\}$, em que A é dado por

$$A = \frac{\pi(k', \boldsymbol{\theta}'_{k'}|\mathbf{y}) q(k|k')}{\pi(k, \boldsymbol{\theta}_k|\mathbf{y}) q(k'|k)} \frac{1}{q_{d_{k \rightarrow k'}}(\mathbf{u})} \left| \frac{\partial g_{k \rightarrow k'}(\boldsymbol{\theta}_k, \mathbf{u})}{\partial(\boldsymbol{\theta}_k, \mathbf{u})} \right|. \quad (4.7)$$

Na Expressão (4.7) temos que $\pi(k', \boldsymbol{\theta}'_{k'}|\mathbf{y})$ e $\pi(k, \boldsymbol{\theta}_k|\mathbf{y})$ denotam, respectivamente, a distribuição *a posteriori*, dada em (4.6), do modelo $\mathcal{M}_{k'}$ (modelo proposto) e do modelo \mathcal{M}_k (modelo atual); $q(k|k')$ e $q(k'|k)$ denotam, respectivamente, a probabilidade de propor um movimento do modelo \mathcal{M}_k para o modelo $\mathcal{M}_{k'}$, e do modelo $\mathcal{M}_{k'}$ para o modelo \mathcal{M}_k ; $q_{d_{k \rightarrow k'}}(\mathbf{u})$ é a função densidade do vetor aleatório \mathbf{u} ; e o último termo é o determinante da matriz Jacobiana decorrente da mudança de variável $(\boldsymbol{\theta}_k, \mathbf{u})$ para $(\boldsymbol{\theta}'_{k'})$.

A proposta de movimento reverso, ou seja, do modelo $\mathcal{M}_{k'}$ para \mathcal{M}_k é aceita com probabilidade de $\min\{1, A^{-1}\}$ em que A é dado na Expressão (4.7).

Logo, podemos descrever o método *Reversible jump* nos seguintes passos:

1. Proponha um movimento do estado $(k, \boldsymbol{\theta}_k)$ do modelo \mathcal{M}_k para o estado $(k', \boldsymbol{\theta}'_{k'})$ do modelo $\mathcal{M}_{k'}$, com uma probabilidade $q(k|k')$;

2. Gere $\mathbf{u} \sim q_{d_{k \rightarrow k'}}(\mathbf{u})$ com dimensão $d_{k \rightarrow k'} = m_{k'} - m_k$;
3. Faça $\boldsymbol{\theta}'_{k'} = g_{k \rightarrow k'}(\boldsymbol{\theta}_k, \mathbf{u})$;
4. Aceite a mudança para o estado $(k', \boldsymbol{\theta}'_{k'})$ com probabilidade $\min\{1, A\}$, em que A é dado pela Expressão (4.7).

4.4.1 MÉTODO REVERSIBLE JUMP PARA MODELOS DE MISTURA

Richardson e Green (1997) utilizam o método *Reversible jump* para estimar o número de componentes K do modelo de mistura de distribuições normais univariadas. Neste algoritmo, os autores propõem os movimentos “nascimento” (*birth*), “morte” (*death*), *split* e *merge* de componentes de mistura.

No movimento “nascimento” (*birth*) uma componente vazia, k^* , é adicionada na mistura, com valores do vetor de parâmetros, $\boldsymbol{\theta}_{k^*}$, gerados da distribuição *a priori*, o peso, p_{k^*} , gerado de uma distribuição $Beta(1, K)$, e então os pesos são reescalados de tal forma que a soma de 1, isto é, $p'_k = p_k(1 - p_{k^*})$. Já no movimento “morte” (*death*), uma componente “vazia” da mistura é removida e os pesos também são reescalados para somar 1, ou seja, $p'_k = p_k/(1 - p_{k^*})$.

A probabilidade de aceitação da proposta do movimento “nascimento” (*birth*) é dada por $\min\{1, A\}$ e do reverso “morte” (*death*) é aceita com probabilidade $\min\{1, A^{-1}\}$, em que A é dado por

$$\begin{aligned}
 A &= \frac{\pi(\mathcal{M}_{K+1}) p_{k^*}^{\tau-1} (1 - p_{k^*})^{N+K\tau-K} (K+1)}{\pi(\mathcal{M}_K) Beta(K\tau, \tau)} \\
 &\times \frac{q(K|K+1)}{q(K+1|K)} \frac{1}{(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} (1 - p_{k^*})^K.
 \end{aligned} \tag{4.8}$$

Na Expressão (4.8) temos na primeira linha, o primeiro quociente que é a razão das probabilidades *a priori* do modelo com o movimento “nascimento” (*birth*) de uma componente, \mathcal{M}_{K+1} , pelo atual, \mathcal{M}_K , e o segundo quociente é a razão das distribuições *a priori* de *Dirichlet* com parâmetro τ , em que N é o número total de observações, $Beta(\cdot, \cdot)$ é a função beta e o termo $(K+1)$ é devido a razão $\frac{(K+1)!}{K!}$, proveniente da suposição de permutabilidade para as componentes de mistura, pois existem, respectivamente, $K!$ e $(K+1)!$ formas de escrever $\boldsymbol{\theta}_K$ e $\boldsymbol{\theta}_{K+1}$. Na segunda linha, a primeira divisão é a razão das probabilidades dos movimentos “nascimento” (*birth*), pelo “morte” (*death*); no segundo quociente o termo (k_0+1) surge da probabilidade de se escolher uma componente para o movimento “morte” (*death*), em que k_0 é o número de componentes vazias antes do movimento “nascimento” (*birth*); o terceiro quociente é o inverso da função densidade de probabilidade da distribuição $Beta(1, K)$ do qual p_{k^*} foi gerado; e o último termo é o jacobiano da transformação. A razão de verossimilhança é 1 devido as propostas dos movimentos envolverem componentes vazias.

Para os movimentos de *split* e *merge* é necessário o mecanismo de salto reversível, pois precisamos projetar esses movimentos em conjunto. A estratégia é escolher distribuições propostas que sugerem uma probabilidade razoável de aceitação, mas estritamente sujeita à exigência de correspondência de dimensão, e a condição de equilíbrio que é determinada pela probabilidade de aceitação dada em (4.7) (RICHARDSON; GREEN, 1997).

A escolha entre *split* (*sp*) ou *merge* (*me*) é aleatória e são calculadas com probabilidades dadas por $Pr(sp) = b_K$ e $Pr(me) = d_K = 1 - b_K$, respectivamente, sendo que $d_1 = 0$ e $b_{K_{max}} = 0$, em que K_{max} é o valor máximo que K pode assumir, e para $K = 2, 3, \dots, K_{max} - 1$ temos que $b_K = d_K = \frac{1}{2}$.

O movimento *merge* consiste em escolher um par de componentes (k_1, k_2) , aleatoriamente e adjacentes, e essas duas componentes são mescladas em uma componente, denotada por k^* , e então as observações y_i com s_{i1} e s_{i2} são realocadas.

A realocação é feita definindo um $s_{i^*} = k^*$ e os valores dos parâmetros da componente k^* são atribuídos por uma regra de identificabilidade, como por exemplo, a ordenação de médias das componentes, já o peso da nova componente, k^* , é a soma dos pesos das componentes combinadas k_1 e k_2 .

Considerando um modelo de mistura de distribuições com dois parâmetros, (μ, σ^2) , por exemplo, o modelo de mistura de distribuições normais, Richardson e Green (1997) propõe o movimento *merge* escolhendo o par de componentes aleatoriamente e adjacentes em termos do valor de suas médias, isto é, $\mu_1 < \mu_2$. A nova componente k^* possui vetor de parâmetros $(p_{k^*}, \mu_{k^*}, \sigma_{k^*}^2)'$, e os parâmetros dessa nova componente são dados por

$$\left. \begin{aligned} p_{k^*} &= p_{k_1} + p_{k_2}, \\ p_{k^*} \mu_{k^*} &= p_{k_1} \mu_{k_1} + p_{k_2} \mu_{k_2}, \\ p_{k^*} (\mu_{k^*}^2 + \sigma_{k^*}^2) &= p_{k_1} (\mu_{k_1}^2 + \sigma_{k_1}^2) + p_{k_2} (\mu_{k_2}^2 + \sigma_{k_2}^2). \end{aligned} \right\} \quad (4.9)$$

A proposta *merge*, neste caso, é determinística, pois as escolhas de k_1 e k_2 já foram feitas, e portanto, a probabilidade de aceitação é dada pela Expressão (4.7).

Já o movimento *split* consiste em dividir uma componente da mistura em duas, isto é, uma componente k^* é escolhida aleatoriamente e dividida em duas, denominadas k_1 e k_2 , com pesos e parâmetros de acordo com as equações dadas em (4.9), mas a dimensão do espaço paramétrico atual para o proposto não é a mesma, sendo necessário gerar um vetor \mathbf{u} para propor valores para os novos parâmetros.

Richardson e Green (1997) consideram o vetor $\mathbf{u} = (u_1, u_2, u_3)'$, as seguintes distribuições: $u_1 \sim Beta(2, 2)$, $u_2 \sim Beta(2, 2)$ e $u_3 \sim Beta(1, 1)$, e as equações

$$p_{k_1} = p_{k^*} u_1 \quad , \quad p_{k_2} = p_{k^*} (1 - u_1), \quad (4.10)$$

$$\mu_{k_1} = \mu_{k^*} - u_2 \sigma_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}} \quad , \quad \mu_{k_2} = \mu_{k^*} + u_2 \sigma_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}}, \quad (4.11)$$

$$\sigma_{k_1}^2 = u_3(1 - u_2^2)\sigma_{k^*}^2 \frac{p_{k^*}}{p_{k_1}}, \quad \sigma_{k_2}^2 = (1 - u_3)(1 - u_2^2)\sigma_{k^*}^2 \frac{p_{k^*}}{p_{k_2}}, \quad (4.12)$$

que definem os 6 novos parâmetros.

A probabilidade de aceitação da proposta do movimento *split* é dada por $\min\{1, A^{SP}\}$, e do reverso *merge* é aceita com probabilidade $\min\{1, A^{SP^{-1}}\}$, em que A^{SP} é dado por

$$\begin{aligned} A^{SP} &= \frac{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_1, \boldsymbol{\theta}_{k_1})]^{I_{\{s_i\}}(k_1)} \prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_2, \boldsymbol{\theta}_{k_2})]^{I_{\{s_i\}}(k_2)}}{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k^*, \boldsymbol{\theta}_{k^*})]^{I_{\{s_i\}}(k^*)}} \\ &\times (K+1) \frac{\pi(\mathcal{M}_{K+1})}{\pi(\mathcal{M}_K)} \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_1+n_2} \text{Beta}(\boldsymbol{\tau}, K\boldsymbol{\tau})} \\ &\times \frac{\pi(\boldsymbol{\mu}_{k_1})\pi(\boldsymbol{\mu}_{k_2})}{\pi(\boldsymbol{\mu}_{k^*})} \times \frac{\pi(\boldsymbol{\sigma}_{k_1}^{-1})\pi(\boldsymbol{\sigma}_{k_2}^{-1})}{\pi(\boldsymbol{\sigma}_{k^*}^{-1})} \\ &\times \frac{d_{K+1}}{b_K P_{aloc}} \frac{1}{g(u_1)g(u_2)g(u_3)} \\ &\times \frac{p_{k^*} |\boldsymbol{\mu}_{k_1} - \boldsymbol{\mu}_{k_2}| \sigma_{k_1}^2 \sigma_{k_2}^2}{u_2(1 - u_2^2)u_3(1 - u_3)\sigma_{k^*}^2} \end{aligned} \quad (4.13)$$

Na Expressão (4.13) temos na primeira linha a razão do produto das funções de verossimilhança das componentes k_1 e k_2 pela função de verossimilhança da componente k^* . Na segunda linha temos o termo $(K+1)$ proveniente das permutabilidades, o primeiro quociente é a razão das probabilidades *a priori* do modelo *split* pelo *merge*, e o segundo quociente é a razão das distribuições *a priori* de *Dirichlet*, em que n_j é o número de observações alocadas na componente j , com $j = 1, 2$. Os termos $\pi(\boldsymbol{\mu}_{k_j})$ e $\pi(\boldsymbol{\sigma}_{k_j}^{-1})$ são, respectivamente, as distribuições *a priori* de $\boldsymbol{\mu}_{k_j}$ e $\boldsymbol{\sigma}_{k_j}^{-1}$, para $j = 1, 2$ e $*$; d_{K+1} e b_K são, respectivamente, as probabilidades da escolha de um *split* ($K+1$) e um *merge* (K); P_{aloc} é a probabilidade de alocação das componentes k_1 e k_2 em k^* ; $g(\cdot)$ é a densidade da distribuição *Beta* das componentes do vetor \boldsymbol{u} ; e a última linha é o Jacobiano da transformação.

Maiores detalhes sobre o cálculo do Jacobiano ver [Seção A.1 do Apêndice A](#).

Desta forma, temos que o método *Reversible jump* para modelos de misturas de distribuições normais pode ser descrito nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$, e especifique um valor inicial para $\boldsymbol{\theta}^{(0)}$ e $\boldsymbol{p}^{(0)}$;
2. Para $t = 1, 2, \dots, T$:
 - (i) atualize $\boldsymbol{\theta}$ da distribuição *a posteriori*;
 - (ii) atualize \boldsymbol{p} da distribuição *a posteriori*;
 - (iii) atualize \boldsymbol{s} (ver [Subseção 4.4.2](#));

- (iv) escolha entre os movimentos *split* ou *merge* com probabilidade, respectivamente, dadas por $Pr(sp)$ e $Pr(me)$. Aceite o movimento *split* com a probabilidade de aceitação $\min\{1, A^{sp}\}$ e o *merge* com a probabilidade $\min\{1, A^{sp-1}\}$ em que A^{sp} é dado em (4.13).
- (v) escolha entre os movimentos “nascimento” (*birth*) ou “morte” (*death*) com probabilidade, respectivamente, dadas por $Pr(n)$ e $Pr(m)$. Aceite o movimento “nascimento” (*birth*) com probabilidade $\min\{1, A\}$ e o “morte” (*death*) com a probabilidade $\min\{1, A^{-1}\}$ em que A é dado em (4.8).
3. Considere um *burn in* de B iterações, e calcule o número de vezes que o número de componentes é igual a K , denotado por N_K para $K = 1, \dots, K_{max}$ nas $T - B$ iterações;
4. Calcule a probabilidade *a posteriori* de K que é dada por

$$P_K = \frac{N_K}{T - B}, \quad (4.14)$$

para $K = 1, \dots, K_{max}$.

Assumimos como estimativa da probabilidade *a posteriori* para o número de componentes K o valor $\tilde{K} = \frac{P_K}{\sum_{1 \leq K \leq K_{max}} P_K}$.

4.4.2 ATUALIZAÇÃO DE \mathbf{s}

Seja a variável aleatória \mathbf{S} como descrito na Seção 3.1 do Capítulo 3 e \mathbf{s} uma realização de $\mathbf{S} = (S_1, \dots, S_N)$, temos que os valores de S_i são simulados de sua distribuição *a posteriori* condicional dada por

$$Pr(S_i = k | \mathbf{y}, \boldsymbol{\theta}, \mathbf{p}) = \frac{p_k f_{Y_i}(y_i | S_i = k, \boldsymbol{\theta}_k)}{\sum_{k=1}^K p_k f_{Y_i}(y_i | S_i = k, \boldsymbol{\theta}_k)}, \quad (4.15)$$

em que p_k é o peso e $f_{Y_i}(y_i | S_i = k, \boldsymbol{\theta}_k)$ é a densidade de $Y_i | S_i = k, \boldsymbol{\theta}_k$, ambos associados a k -ésima componente de S_i .

Logo $S_i \sim \text{Multinomial}_K(1, \boldsymbol{\delta}_i = (\delta_1, \dots, \delta_K)')$, com K resultados e uma observação, tal que $\delta_k = P(S_i = k | \mathbf{y}, \boldsymbol{\theta}, \mathbf{p})$, que é dada em (4.15)

Como a distribuição condicional de \mathbf{s} é completa sua atualização é por meio do algoritmo *Gibbs sampling* da seguinte forma:

1. Inicialize o contador de iterações $t = 0$, e especifique um valor inicial para $\mathbf{s}^{(0)}$;
2. Obtenha um novo valor de $\mathbf{s}^{(t)}$ a partir de $\mathbf{s}^{(t-1)}$, por meio da geração sucessiva do valor $S_i \sim \text{Multinomial}_K(1, \boldsymbol{\delta}_i = (\delta_1, \dots, \delta_K)')$, para $i = 1, \dots, N$;
3. Mude o contador de t para $t + 1$, e retorne ao Passo 2 para $t = 1, 2, \dots, T$.

Observe no Passo 2, que a cada iteração t , o valor de S_i , com $i = 1, \dots, N$ é atualizado.

4.5 Método *Data-driven Reversible jump*

O método *Data-driven Reversible jump* é um procedimento MCMC para selecionar modelos de mistura por meio da probabilidade *a posteriori* para o número de componentes K de um modelo de mistura.

O algoritmo *Data-driven Reversible jump* é direcionado pelos dados, e a atualização do número de componentes de mistura é feita por meio dos movimentos *birth*, *split* e *merge*, ou seja, as mesmas propostas *split* e *merge* do método *Reversible jump*. Porém, no método *Data-driven Reversible jump* ao atualizar uma variável não observável esta é capaz de determinar o “nascimento” (*birth*) de uma nova componente, e o passo “morte” (*death*) ocorre sempre que o número de observações pertencentes a uma componente “cai” para zero. A distribuição *a priori* para os pesos das componentes é diferente da assumida por Richardson e Green (1997). Nesta proposta, a distribuição de Dirichlet tem parâmetros $\frac{\delta}{K}$ e em seguida é assumido $K \rightarrow \infty$, e este procedimento determina uma abordagem semelhante ao processo de Dirichlet (SARAIVA, 2009, p. 4).

A implementação do algoritmo *Data-driven Reversible jump* é feita em 2 passos. No primeiro passo os valores atuais dos parâmetros são atualizados usando algum algoritmo MCMC, como o *Metropolis-Hastings* ou *Gibbs sampling*, e no segundo passo o número de componentes do modelo de mistura é atualizado usando o movimento *split* ou *merge*.

Considere o estado atual $x = (\boldsymbol{\theta}, \mathbf{p}, K)$ do modelo \mathcal{M}_K , com K componentes e o movimento proposto para o estado $x^* = (\boldsymbol{\theta}^*, \mathbf{p}^*, K^*)$ do modelo \mathcal{M}_{K^*} , em que “*” significa um *split* ou um *merge*. Este movimento é aceito segundo a probabilidade de aceitação do algoritmo *Metropolis-Hastings*, que denotaremos por $\Psi(x^*|x) = \min\{1, A^*\}$, em que A^* é dado por

$$A^* = \frac{\mathcal{L}(\boldsymbol{\theta}^*, \mathbf{p}^* | \mathbf{y}, \mathbf{s}^*, K^*) \pi(\boldsymbol{\theta}^*, \mathbf{p}^*, K^*) q(x|x^*)}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) \pi(\boldsymbol{\theta}, \mathbf{p}, K) q(x^*|x)}. \quad (4.16)$$

Na Expressão (4.16) temos que o primeiro quociente é a razão de verossimilhanças do modelo \mathcal{M}_{K^*} pelo modelo \mathcal{M}_K , o segundo quociente é a razão de prioris do modelo \mathcal{M}_{K^*} pelo modelo \mathcal{M}_K e $q(\cdot|\cdot)$ é a distribuição da proposta de transição, a partir da qual deduzimos o movimento *split* ou *merge*. Aqui, mesmo estando propondo um movimento entre modelos com diferentes dimensões, usamos uma probabilidade de aceitação do *Metropolis-Hastings* para avaliar a aceitação da transição, por que os parâmetros do modelo proposto são gerados a partir de uma distribuição proposta que é independente dos parâmetros do modelo atual e o Jacobiano é igual a 1 (ZUANETTI, 2016).

Desta forma, a escolha entre os movimentos *split* ou *merge* é realizada em duas etapas:

1. Escolhemos duas observações distintas, i_1 e i_2 , por exemplo, utilizando uma distribuição uniforme discreta $(1, \dots, N)$, sem reposição.

2. Se $s_{i_1} = s_{i_2}$, temos que ambas as observações pertencem à mesma componente, então o movimento *split* é proposto, caso contrário, se $s_{i_1} \neq s_{i_2}$, ambas as observações pertencem a componentes diferentes e o movimento *merge* é proposto.

A probabilidade da escolha de um movimento *split*, $s_{i_*} = s_{i_1} = s_{i_2}$, é dada por

$$b_{s_{i_*}} = \frac{n_{s_{i_*}}(n_{s_{i_*}} - 1)}{N(N - 1)}, \quad (4.17)$$

em que $n_{s_{i_*}} = \sum_{i=1}^N I_{\{s_i\}}(s_{i_*})$ é o número de observações alocadas na componente s_{i_*} , e a probabilidade da escolha de um *merge* de componentes s_{i_1} e s_{i_2} é dada por

$$d_{s_{i_1}s_{i_2}} = \frac{2n_{s_{i_1}}n_{s_{i_2}}}{N(N - 1)}, \quad (4.18)$$

em que $n_{s_{i_j}} = \sum_{i=1}^N I_{\{s_i\}}(s_{i_j})$ é o número de observações alocadas na componente s_{i_j} , com $j = 1, 2$.

Observe que $d_{s_{i_1}s_{i_2}} + b_{s_{i_*}}$ não é obrigatoriamente igual a 1, se $K > 1$ temos $K!(K - 1)!$ pares de componentes para o *merge* e K componentes para o *split* e um movimento entre essas possibilidades é escolhido (ZUANETTI, 2016, p. 12).

No movimento *split* realocamos as observações em duas componentes e simulamos parâmetros para elas. Neste caso, a realocação das observações é feita, utilizando a notação de \mathcal{S} como $\mathbf{s}^{sp} = (s_1^{sp}, \dots, s_N^{sp})'$. Assim, para $i = 1, \dots, N$ temos que:

- Se $i = i_1$, distribuimos y_{i_1} para a $(K + 1)$ -ésima componente, $s_{i_1}^{sp} = K + 1$.
- Se $i = i_2$, mantemos y_{i_2} na mesma componente, $s_{i_2}^{sp} = s_{i_2}$.
- Para as demais observações, $s_i = s_{i_*}$, com $s_{i_*} \neq s_{i_1}$ e $s_{i_*} \neq s_{i_2}$, distribuimos y_i para a componente $K + 1$ com probabilidade $Pr_{s_{i_1}^{sp}}(i)$ ou mantemos y_i na componente s_{i_*} com probabilidade $1 - Pr_{s_{i_1}^{sp}}(i)$, em que a $Pr_{s_{i_1}^{sp}}(i)$ é escolhida arbitrariamente.
- Se $s_i \neq s_{i_*}$, mantemos y_i na mesma componente, $s_i^{sp} = s_i$.

O procedimento *split* pode ser descrito da seguinte forma: inicialmente é proposto um movimento ou uma mudança no estado do modelo. Se esse movimento é aceito, então duas observações são sorteadas e se as variáveis não observáveis correspondentes à essas duas observações tem o mesmo valor, então o movimento será um *split*. A variável aleatória não observável será verificada para $i = 1, \dots, N$. Para o primeiro índice da variável não observável sorteado, o valor observável correspondente será associado a uma nova componente $K + 1$. Para o segundo índice da variável não observável sorteado, o valor observável correspondente a esse índice continuará associado a sua componente atual. Depois, para todos os outros índices onde o estado do índice é igual a s_{i_*} , temos que o valor observável correspondente a esse índice será associado a nova componente $K + 1$, com uma probabilidade arbitrária, ou seja, continuará sendo associado ao estado s_{i_*} com o complemento dessa probabilidade arbitrária.

A probabilidade da nova configuração \mathbf{s}^{SP} de \mathbf{S} é dada por

$$Pr(\mathbf{s}^{SP}|\mathbf{s}) = \prod_{i:s_i^{SP}=s_{i_1}^{SP}} Pr_{s_{i_1}^{SP}}(i) \prod_{i:s_i^{SP}=s_{i_2}^{SP}} \left(1 - Pr_{s_{i_1}^{SP}}(i)\right), \quad (4.19)$$

em que $Pr_{s_{i_1}^{SP}}(i_1) = 1$, $Pr_{s_{i_1}^{SP}}(i_2) = 0$ e $Pr_{s_{i_1}^{SP}}(i) = 1/2$ para $i \notin \{i_1, i_2\}$ o que resulta em

$$Pr(\mathbf{s}^{SP}|\mathbf{s}) = \left(\frac{1}{2}\right)^{n_{s_{i_1}^{SP}} + n_{s_{i_2}^{SP}} - 2}. \quad (4.20)$$

Para definir parâmetros para as novas componentes simulamos valores candidatos $x^{SP} = (\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1)$ da distribuição condicional *a posteriori*, e atualizamos somente $\boldsymbol{\theta}^{SP}$ para $\boldsymbol{\theta}_{s_{i_1}^{SP}}^{SP}$ e $\boldsymbol{\theta}_{s_{i_2}^{SP}}^{SP}$, condicional sobre $\mathbf{S} = \mathbf{s}^{SP}$.

A distribuição da proposta de um movimento *split* é dada por

$$\begin{aligned} q(x^{SP}|x) &= b_{s_{i^*}} Pr(\mathbf{s}^{SP}|\mathbf{s}) \\ &\times \pi(\boldsymbol{\theta}_{s_{i_1}^{SP}}^{SP}, \mathbf{p}^{SP}, K+1|\mathbf{y}, \mathbf{s}^{SP}) \\ &\times \pi(\boldsymbol{\theta}_{s_{i_2}^{SP}}^{SP}, \mathbf{p}^{SP}, K+1|\mathbf{y}, \mathbf{s}^{SP}), \end{aligned} \quad (4.21)$$

em que $\pi(\cdot|\cdot)$ é a distribuição *a posteriori*.

Desta forma, a probabilidade de aceitação do movimento *split* pode ser escrita como $\Psi(x^{SP}|x) = \min\{1, A^{SP}\}$, em que A^{SP} é dado por

$$A^{SP} = \frac{\mathcal{L}(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}|\mathbf{y}, \mathbf{s}^{SP}, K+1) \pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1) q(x|x^{SP})}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p}|\mathbf{y}, \mathbf{s}, K) \pi(\boldsymbol{\theta}, \mathbf{p}, K) q(x^{SP}|x)}. \quad (4.22)$$

Na Expressão (4.22) temos que a razão de verossimilhança do movimento *split* é dada por

$$\frac{\mathcal{L}(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1|\mathbf{y}, \mathbf{s}^{SP})}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p}, K|\mathbf{y}, \mathbf{s})} = \frac{\prod_{t:s_t^{SP}=s_{i_1}^{SP}} [f_{Y_i}(y_i|S_i = s_{i_1}^{SP}, \boldsymbol{\theta}_{s_{i_1}^{SP}}^{SP})]^{I_{s_i}(s_{i_1}^{SP})} \prod_{t:s_t^{SP}=s_{i_2}^{SP}} [f_{Y_i}(y_i|S_i = s_{i_2}^{SP}, \boldsymbol{\theta}_{s_{i_2}^{SP}}^{SP})]^{I_{s_i}(s_{i_2}^{SP})} \frac{\prod_{k=1}^{K+1} p_k^{s_k^{SP}}}{\prod_{k=1}^K p_k^{n_k}}}{\prod_{t:s_t=s_{i^*}} [f_{Y_i}(y_i|S_i = s_{i^*}, \boldsymbol{\theta}_{s_{i^*}})]^{I_{s_i}(s_{i^*})} \frac{\prod_{k=1}^K p_k^{n_k}}{\prod_{k=1}^K p_k^{n_k}}}, \quad (4.23)$$

a razão de prioris é dada por

$$\frac{\pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1)}{\pi(\boldsymbol{\theta}, \mathbf{p}, K)} = \frac{\left(\frac{\Gamma\left(\sum_{k=1}^{K+1} \tau_k\right)}{\prod_{k=1}^{K+1} \Gamma(\tau_k)} \prod_{k=1}^{K+1} p_k^{\tau_k - 1} \right) \pi(\boldsymbol{\theta}_{s_{i_1}^{SP}}) \pi(\boldsymbol{\theta}_{s_{i_2}^{SP}})}{\left(\frac{\Gamma\left(\sum_{k=1}^K \tau_k\right)}{\prod_{k=1}^K \Gamma(\tau_k)} \prod_{k=1}^K p_k^{\tau_k - 1} \right) \pi(\boldsymbol{\theta}_{s_{i^*}})}, \quad (4.24)$$

e a razão de distribuições propostas é dada por

$$\begin{aligned} \frac{q(x|x^{SP})}{q(x^{SP}|x)} &= \frac{2n_{s_{i_1}^{SP}}n_{s_{i_2}^{SP}}}{n_{s_{i_*}}(n_{s_{i_*}} - 1) \left(\frac{1}{2}\right)^{n_{s_{i_1}^{SP}} + n_{s_{i_2}^{SP}} - 2}} \\ &\times \frac{\pi(\boldsymbol{\theta}_{s_{i_*}}, \mathbf{p}, K | \mathbf{y}, \mathbf{s})}{\pi(\boldsymbol{\theta}_{s_{i_1}^{SP}}, \mathbf{p}^{SP}, K + 1 | \mathbf{y}, \mathbf{s}^{SP}) \pi(\boldsymbol{\theta}_{s_{i_2}^{SP}}, \mathbf{p}^{SP}, K + 1 | \mathbf{y}, \mathbf{s}^{SP})}. \end{aligned} \quad (4.25)$$

Se $\tau_k = 1$, para $k = 1, \dots, K$ a razão de prioris pode ser simplificada por

$$\frac{\pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K + 1)}{\pi(\boldsymbol{\theta}, \mathbf{p}, K)} = K \times \frac{\pi(\boldsymbol{\theta}_{s_{i_1}^{SP}}) \pi(\boldsymbol{\theta}_{s_{i_2}^{SP}})}{\pi(\boldsymbol{\theta}_{s_{i_*}})}. \quad (4.26)$$

No movimento *merge* mesclamos as observações de duas componentes em uma, e também simulamos parâmetros para a nova componente. A mesclagem das observações é feita, utilizando a notação de \mathcal{S} como $\mathbf{s}^{mg} = (s_1^{mg}, \dots, s_N^{mg})'$. Deste modo, para $i = 1, \dots, N$, temos que:

- Se $s_i < \max(s_{i_1}, s_{i_2})$, fazemos $s_i^{mg} = s_i$.
- Se $s_i = \max(s_{i_1}, s_{i_2})$, fazemos $s_i^{mg} = \min(s_{i_1}, s_{i_2})$.
- Se $s_i > \max(s_{i_1}, s_{i_2})$, fazemos $s_i^{mg} = s_i - 1$.

Desta forma, o procedimento irá eliminar o último componente do vetor de componentes, e com isso o valor de K será atualizado para $K - 1$.

A probabilidade da nova configuração \mathbf{s}^{mg} é dada por

$$Pr(\mathbf{s}^{mg} | \mathbf{s}) = 1. \quad (4.27)$$

Definimos os parâmetros da nova componente simulando valores candidatos $x^{mg} = (\boldsymbol{\theta}^{mg}, \mathbf{p}^{mg}, K - 1)$ da distribuição condicional *a posteriori*, atualizando somente $\boldsymbol{\theta}^{mg}$ para $\boldsymbol{\theta}_{s_{i_*}}^{mg}$, condicional sobre $\mathcal{S} = \mathbf{s}^{mg}$.

A distribuição da proposta de um movimento *merge* é dada por

$$q(x^{mg}|x) = d_{s_{i_1} s_{i_2}} \times \pi(\boldsymbol{\theta}_{s_{i_*}}^{mg}, \mathbf{p}^{mg} | \mathbf{y}, \mathbf{s}^{mg}, K - 1), \quad (4.28)$$

em que $\pi(\cdot | \cdot)$ é a distribuição *a posteriori* condicional utilizada para atualizar os parâmetros.

A probabilidade de aceitação do movimento *merge* é dada por $\Psi(x^{mg}|x) = \min\{1, A^{mg}\}$, em que $A^{mg} = 1/A^{SP}$.

Portanto, a probabilidade de propor um movimento *split*, ou seja, um movimento do estado x para o estado x^{SP} é a mesma que estar no estado x^{mg} e propor um movimento *split*, isto é, $q(x^{SP}|x) = q(x|x^{mg})$ e analogamente $q(x^{mg}|x) = q(x|x^{SP})$ (ZUANETTI, 2016, p. 14)

Assim, o algoritmo *Data-driven Reversible jump* para modelos de misturas pode ser descrito nos seguintes passos:

1. Especifique um valor inicial para K e $\mathbf{s} = \{s_1, \dots, s_N\}$ tal que $s_i \in \{1, \dots, K\}$, para $i = 1, \dots, N$;
2. Atualize $(\boldsymbol{\theta}_k, p_k)'$ para $k = 1, \dots, K$ da distribuição *a posteriori*;
3. *Data driven Reversible jump*: Atualize K .
Para $t = 1, \dots, T$:
 - (a) Amostre duas observações distintas, i_1 e i_2 , e decida entre os movimentos *split* ou *merge*;
 - (b) Construa o s^* candidato e os valores candidatos x^* , em que $*$ é um movimento *split* (*sp*) ou um movimento *merge* (*mg*);
 - (c) Aceite a proposta com probabilidade $\Psi(x^*|x) = \min\{1, A^*\}$, em que A^* é dado em (4.16);
 - i. Se o movimento *split* é aceito, faça $K^{(t)} = K^{(t-1)} + 1$, e considere x^{sp} ;
 - ii. Se o movimento *merge* é aceito, faça $K^{(t)} = K^{(t-1)} - 1$, e considere x^{mg} ;
 - iii. Se nenhum dos movimentos são aceitos, faça $K^{(t)} = K^{(t-1)}$, e considere x ;
4. Atualize s_i , para $i = 1, \dots, N$ de uma distribuição $Multinomial_{K^{(t)}}(1, \boldsymbol{\delta}_i = (\delta_1, \dots, \delta_{K^{(t)}})')$, em que $\delta_k = Pr(S_i = k | \mathbf{y}, \boldsymbol{\theta}, \mathbf{p}, K^{(t)})$ (Ver Subseção 4.4.2).
5. Atualize $(\boldsymbol{\theta}_k, p_k)'$, para $k = 1, \dots, K^{(t)}$ da distribuição *a posteriori*.

No método *Data driven Reversible jump*, a estimativa da probabilidade *a posteriori* para o número de componentes K é obtida da mesma forma que no método *Reversible jump*.

A distribuição *a priori* e a função de verossimilhança são invariantes sob permutações de rótulos das componentes (ZUANETTI, 2016, p. 16).

Maiores detalhes sobre não identificabilidade dos parâmetros de modelos de mistura de distribuições, também chamados de *label switching* podem ser encontrados em Dahl (2006), Jasra *et al.* (2005) e Stephens (2000).

MODELO DE MISTURA DE DISTRIBUIÇÕES NORMAIS ASSIMÉTRICAS

Neste capítulo, descrevemos o modelo de mistura de distribuições normais assimétricas com um número de componentes conhecido e a sua função de verossimilhança. Além disso, também apresentamos os procedimentos utilizados para estimar o número de componentes juntamente com os parâmetros associados as componentes de mistura do modelo de mistura de distribuições normais assimétricas.

5.1 O modelo

Suponha que as densidades das componentes de mistura seguem a distribuição normal assimétrica como definida em (2.5), com vetor de parâmetros $\boldsymbol{\theta}_k = (\xi_k, \omega_k^2, \alpha_k)'$. Temos que a função densidade de probabilidade condicional, $f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k)$, é dada por

$$f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k) = \frac{2}{\sqrt{2\pi}\omega_k^2} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_k}{\omega_k}\right)^2\right\} \int_{-\infty}^{\alpha_k\left(\frac{y_i - \xi_k}{\omega_k}\right)} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du, \quad (5.1)$$

em que ξ_k é o parâmetro de locação, ω_k^2 é o parâmetro de escala, α_k é o parâmetro de assimetria, da k -ésima componente de mistura, com $k = 1, \dots, K$.

Podemos escrever o modelo de mistura de distribuições normais assimétricas com um número de componentes K conhecido, correspondente a Expressão (3.6) do Capítulo 3, como

$$\begin{aligned} f_{Y_i}(y|\boldsymbol{\theta}, \mathbf{p}) &= \sum_{k=1}^K f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k) Pr(S_i = k|\mathbf{p}) \\ &= \sum_{k=1}^K p_k \frac{2}{\sqrt{2\pi}\omega_k^2} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_k}{\omega_k}\right)^2\right\} \\ &\quad \times \int_{-\infty}^{\alpha_k\left(\frac{y_i - \xi_k}{\omega_k}\right)} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du, \end{aligned} \quad (5.2)$$

em que $(\xi_k, \omega_k^2, \alpha_k, p_k)'$ são os parâmetros da k -ésima componente do modelo a serem estimados, para $k = 1, \dots, K$.

Denotamos o modelo de mistura de distribuições normais assimétricas com o número de componentes K conhecido por $Y_i|S_i = k, \boldsymbol{\theta}, \mathbf{p} \sim NA(\xi_k, \omega_k^2, \alpha_k)$, com $k = 1, \dots, K$.

5.2 Função de verossimilhança

A função de verossimilhança conjunta de \mathbf{y} e \mathbf{s} , para mistura de distribuições normais assimétricas, de acordo com a Expressão (3.9) do Capítulo 3, é dada por

$$\begin{aligned} \mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{p}|\mathbf{y}, \mathbf{s}, K) &= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\sqrt{2\pi\omega_k^2}} \right)^{n_k} \exp \left\{ - \left[\frac{(S_k^2)(n_k - 1) + n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} \right] \right\} \right. \\ &\quad \left. \times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ -\frac{u^2}{2} \right\} du \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \right\}, \end{aligned} \quad (5.3)$$

em que $n_k = \sum_{i=1}^N I_{\{s_i\}}(k)$ é o número de observações da k -ésima componente de mistura, $\bar{y}_k = \frac{1}{n_k} \sum_{i=1}^N y_i I_{\{s_i\}}(k)$ é a média amostral da componente k , e $S_k^2 = \frac{1}{n_k - 1} \sum_{i=1}^N (y_i - \bar{y}_k)^2 I_{\{s_i\}}(k)$ é a variância amostral da componente k .

5.3 Abordagem Bayesiana

A distribuição *a priori* conjunta para os parâmetros do modelo de mistura de distribuições normais assimétricas, $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha})'$, de acordo com a Expressão (4.1) do Capítulo 4 é dada por

$$\begin{aligned} \pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{p}, K) &= \pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}|K) \pi(\mathbf{p}|K) \pi(K) \\ &= \pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K) \pi(\boldsymbol{\omega}^2|K) \pi(\boldsymbol{\alpha}|K) \pi(\mathbf{p}|K) \pi(K), \end{aligned} \quad (5.4)$$

pois $\boldsymbol{\xi} \perp \boldsymbol{\alpha}$ e $\boldsymbol{\omega}^2 \perp \boldsymbol{\alpha}$, em que o símbolo \perp indica independência entre os parâmetros.

Considerando a k -ésima componente do modelo de mistura de distribuições normais assimétricas, a distribuição *a priori* pode ser escrita por

$$\pi(\boldsymbol{\theta}_k, p_k, k) = \pi(\xi_k|\omega_k^2) \pi(\omega_k^2) \pi(\alpha_k) \pi(p_k) \pi(k). \quad (5.5)$$

Neste trabalho, utilizamos distribuições *a priori* vagas para todos os parâmetros dos modelos, ou seja, distribuições *a priori* com pouca informação. Assim, tomamos as distribuições *a priori* descritas abaixo para os seguintes parâmetros:

1. $\xi_k|\omega_k^2 \sim Normal(\mu_k, \omega_k^2)$, em que μ_k é conhecida, com $\mu_k \in (-\infty, +\infty)$, e ω_k^2 é condicionada, com $\omega_k^2 > 0$, para $k = 1, \dots, K$;

2. $\omega_k^2 \sim \text{Gama inversa}(a_k, b_k)$, em que a_k e b_k são conhecidos, com $a_k > 0$ e $b_k > 0$, para $k = 1, \dots, K$;
3. $\alpha_k \sim \text{Normal}(\mu_k, \sigma_k^2)$ em que μ_k e σ_k^2 são conhecidas, com $\mu_k \in (-\infty, +\infty)$, e $\sigma_k^2 > 0$, para $k = 1, \dots, K$;
4. $p_k \sim \text{Dirichlet}(\tau_k)$, em que τ_k é conhecido, com $\tau_k > 0$, para $k = 1, \dots, K$;
5. $K \sim U_d(K_{max})$, em que U_d representa a distribuição Uniforme discreta no conjunto $\{1, \dots, K_{max}\}$ e K_{max} é o valor máximo especificado para K .

Sabemos de acordo com a [Seção 4.1](#) do [Capítulo 4](#) que a distribuição *a posteriori*, a menos de valores constantes, é produto da função de verossimilhança pela distribuição *a priori*, então para mistura de distribuições normais assimétricas em que $\theta = (\xi, \omega^2, \alpha)'$, a Expressão (4.2) pode ser escrita como

$$\begin{aligned} \pi(\xi, \omega^2, \alpha, p, K | \mathbf{y}, \mathbf{s}) &\propto \mathcal{L}(\xi, \omega^2, \alpha, p | \mathbf{y}, \mathbf{s}, K) \\ &\times \pi(\xi | \omega^2, K) \pi(\omega^2 | K) \pi(\alpha | K) \pi(p | K) \pi(K). \end{aligned} \quad (5.6)$$

Logo, a distribuição *a posteriori* da k -ésima componente do modelo de mistura de distribuições normais assimétricas é dada por

$$\begin{aligned} \pi(\xi_k, \omega_k^2, \alpha_k, p_k | \mathbf{y}, \mathbf{s}) &\propto \mathcal{L}(\xi_k, \omega_k^2, \alpha_k, p_k | \mathbf{y}, \mathbf{s}) \\ &\times \pi(\xi_k | \omega_k^2) \pi(\omega_k^2) \pi(\alpha_k) \pi(p_k) \pi(k). \end{aligned} \quad (5.7)$$

As distribuições condicionais para os parâmetros do modelo de mistura de distribuições normais assimétricas não são conjugadas, com exceção do parâmetro peso da componente (ver [Seção C.1](#) do [Apêndice C](#)).

Maiores detalhes sobre os cálculos da função de verossimilhança e da distribuição *a posteriori* do modelo de mistura de distribuições normais assimétricas ver, respectivamente, a [Seção B.1](#) do [Apêndice B](#) e a [Seção C.1](#) do [Apêndice C](#).

5.4 Estimação

As probabilidades *a priori* assumidas para os parâmetros dos modelos de mistura de distribuições normais assimétricas estão descritas na [Seção 5.3](#) e escolhemos os seguintes valores para os hiperparâmetros das distribuições *a priori*: $K \sim U_d(10)$; $p_k \sim \text{Dirichlet}(1)$; $\xi_k \sim \text{Normal}(0, \omega_k^2)$; $\omega_k^2 \sim \text{Gama inversa}(0.01, 0.01)$; e $\alpha_k \sim \text{Normal}(0, 100)$, para $k = 1, \dots, 10$.

Para estimar o número de componentes K e os parâmetros associados as componentes do modelo de mistura de distribuições utilizamos os métodos *Reversible jump* e o *Data-driven Reversible jump*, ou seja, aplicamos os movimentos “nascimento” (*birth*), “morte” (*death*), *split*

e *merge*, como descrito, respectivamente, na [Subseção 4.4.1](#) e na [Seção 4.5](#) do [Capítulo 4](#), considerando que as componentes de mistura seguem a distribuição normal assimétrica.

Assim, no movimento “nascimento” (*birth*) do método *Reversible jump*, nós adicionamos uma componente “vazia” com vetor de parâmetros $(\xi_{k^*}, \omega_{k^*}^2, \alpha_{k^*}, p_{k^*})'$ gerados a partir das distribuições *a priori*, descritas no início dessa seção, exceto para o parâmetro p_{k^*} que foi gerado de uma distribuição Beta $(1, K)$, de acordo com [Richardson e Green \(1997\)](#). No movimento de “morte” (*death*) removemos uma componente “vazia” na mistura, e em ambos os movimentos reescalamos os pesos como descrito na [Subseção 4.4.1](#) do [Capítulo 4](#).

Atribuímos as seguintes probabilidades para os movimentos “nascimento” (*birth*), $Pr(n)$, e “morte” (*death*), $Pr(m)$:

$$\begin{cases} Pr(n) = 1, & \text{se } K = 1; \\ Pr(n) = Pr(m) = \frac{1}{2}, & \text{se } 1 < K < K_{max} - 1; \\ Pr(m) = 1, & \text{se } K = K_{max}. \end{cases} \quad (5.8)$$

Desta forma, temos que se $K = 1$, um movimento de “nascimento” é escolhido com probabilidade $Pr(n) = 1$, se $1 < K \leq K_{max} - 1$ um “nascimento” (*birth*) ou uma “morte” (*death*) é escolhido aleatoriamente com probabilidade $Pr(n) = Pr(m) = \frac{1}{2}$, e se $K = K_{max}$ um movimento de “morte” (*death*) é escolhido com probabilidade $Pr(m) = 1$.

A probabilidade de aceitação da proposta do movimento “nascimento” (*birth*) é dada por $\min\{1, A\}$ e do reverso “morte” (*death*) é aceita com probabilidade $\min\{1, A^{-1}\}$, em que A é dado por

$$\begin{aligned} A &= \frac{\pi(\xi|\omega^2, K+1)\pi(\omega^2|K+1)\pi(\alpha|K+1)}{\pi(\xi|\omega^2, K)\pi(\omega^2|K)\pi(\alpha|K)} \\ &\times \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{Beta(K\tau, \tau)} \\ &\times \frac{q(K|K+1)}{q(K+1|K)} \frac{1}{(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} \frac{1}{(1-p_{k^*})^K}. \end{aligned} \quad (5.9)$$

Na Expressão (5.9) temos na primeira linha, a razão das probabilidades *a priori* do modelo com o movimento “nascimento” (*birth*) de uma componente pelo atual. Na segunda linha, a razão das distribuições *a priori* de *Dirichlet* com parâmetro τ , em que N é o número total de observações, $Beta(\cdot, \cdot)$ é a função beta e o termo $(K+1)$ é devido a razão $\frac{(K+1)!}{K!}$, proveniente da suposição de permutabilidade para as componentes de mistura, pois existem, respectivamente, $K!$ e $(K+1)!$ formas de escrever θ_K e θ_{K+1} . Na terceira linha, a primeira divisão é a razão das probabilidades dos movimentos “nascimento” (*birth*), pelo “morte” (*death*), o termo (k_0+1) surge da probabilidade de se escolher uma componente para o movimento “morte” (*death*), em que k_0 é o número de componentes “vazias” antes do movimento “nascimento” (*birth*); o terceiro quociente é o inverso da função densidade de probabilidade da distribuição $Beta(1, K)$ do qual p_{k^*} foi gerado; e o último termo é o jacobiano da transformação. A razão de verossimilhança é 1 devido as propostas dos movimentos envolverem componentes “vazias”.

No movimento *merge* escolhemos as componentes baseado na [Subseção 4.4.1](#) do [Capítulo 4](#), isto é, propomos o movimento *merge* escolhendo um par de componentes aleatoriamente e adjacentes em termos do valor do parâmetro de locação, ou seja, $\xi_1 < \xi_2$, e as probabilidades dos movimentos *split* e *merge* tomamos como iguais, $Pr(sp) = Pr(me) = \frac{1}{2}$, para $K = 2, 3, \dots, K_{max} - 1$.

Portanto, a nova componente k^* possui vetor de parâmetros dado por $(\xi_{k^*}, \omega_{k^*}^2, \alpha_{k^*}, p_{k^*})'$ e os parâmetros da nova componente são calculados por

$$\left. \begin{aligned} p_{k^*} &= p_{k_1} + p_{k_2}, \\ p_{k^*} \xi_{k^*} &= p_{k_1} \xi_{k_1} + p_{k_2} \xi_{k_2}, \\ p_{k^*} (\xi_{k^*}^2 + \omega_{k^*}^2) &= p_{k_1} (\xi_{k_1}^2 + \omega_{k_1}^2) + p_{k_2} (\xi_{k_2}^2 + \omega_{k_2}^2), \\ p_{k^*} \alpha_{k^*} &= p_{k_1} \alpha_{k_1} + p_{k_2} \alpha_{k_2}. \end{aligned} \right\} \quad (5.10)$$

Tomamos o vetor $\mathbf{u} = (u_1, u_2, u_3, u_4)'$ e consideramos as seguintes distribuições para as componentes do vetor \mathbf{u} , $u_1 \sim Beta(2, 2)$, $u_2 \sim Beta(2, 2)$, $u_3 \sim Beta(1, 1)$ e $u_4 \sim Beta(2, 2)$, e as seguintes equações

$$p_{k_1} = p_{k^*} u_1, \quad p_{k_2} = p_{k^*} (1 - u_1), \quad (5.11)$$

$$\xi_{k_1} = \xi_{k^*} - u_2 \omega_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}}, \quad \xi_{k_2} = \xi_{k^*} + u_2 \omega_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}}, \quad (5.12)$$

$$\omega_{k_1}^2 = u_3 (1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{p_{k_1}}, \quad \omega_{k_2}^2 = (1 - u_3) (1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{p_{k_2}}, \quad (5.13)$$

$$\alpha_{k_1} = \alpha_{k^*} u_4, \quad \alpha_{k_2} = \alpha_{k^*} (1 - u_4), \quad (5.14)$$

que fornecem os oito novos parâmetros necessários para a transformação, satisfazendo (5.10).

As transformações necessárias para a obtenção dos parâmetros da componente k^* (5.10), as distribuições das componentes do vetor \mathbf{u} e a Equações de (5.11) a (5.14) são baseadas em [Richardson e Green \(1997\)](#).

A probabilidade de aceitação da proposta do movimento *split* é dada por $\min\{1, A^{sp}\}$, e do reverso *merge* é aceita com probabilidade $\min\{1, A^{sp^{-1}}\}$, em que A para o modelo de mistura

de distribuições normais assimétricas é dado por

$$\begin{aligned}
A^{sp} &= \left[\left(\frac{2}{\sqrt{2\pi}\omega_{k_1}^2} \right)^{n_{k_1}} \exp \left\{ - \left[\frac{(S_{k_1}^2)(n_{k_1} - 1) + n_{k_1}(\bar{y}_{k_1} - \xi_{k_1})^2}{2\omega_{k_1}^2} \right] \right\} \right] \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_1}} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k_1)} \\
&\times \left[\left(\frac{2}{\sqrt{2\pi}\omega_{k_2}^2} \right)^{n_{k_2}} \exp \left\{ - \left[\frac{(S_{k_2}^2)(n_{k_2} - 1) + n_{k_2}(\bar{y}_{k_2} - \xi_{k_2})^2}{2\omega_{k_2}^2} \right] \right\} \right] \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2}} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k_2)} \\
&\times \frac{1}{\left[\left(\frac{2}{\sqrt{2\pi}\omega_{k^*}^2} \right)^{n_{k^*}} \exp \left\{ - \left[\frac{(S_{k^*}^2)(n_{k^*} - 1) + n_{k^*}(\bar{y}_{k^*} - \xi_{k^*})^2}{2\omega_{k^*}^2} \right] \right\} \right]} \\
&\times \frac{1}{\prod_{i=1}^N \int_{-\infty}^{\alpha_{k^*}} \left(\frac{y_i - \xi_{k^*}}{\omega_{k^*}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k^*)} \\
&\times \frac{(K+1) \pi(\xi | \omega^2, K+1) \pi(\omega^2 | K+1) \pi(\alpha | K+1)}{\pi(\xi | \omega^2, K) \pi(\omega^2 | K) \pi(\alpha | K)} \\
&\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{Beta(\tau, K\tau)} \\
&\times \frac{\pi(\xi_{k_1}) \pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2) \pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1}) \pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \\
&\times \frac{1}{P_{aloc} g(u_1) g(u_2) g(u_3) g(u_4)} \times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1}}{u_2(1-u_2) u_3(1-u_3) u_4 \omega_{k^*}^2}. \tag{5.15}
\end{aligned}$$

Na Expressão (5.15) temos a razão do produto das funções de verossimilhança das componentes k_1 e k_2 , pela função de verossimilhança da componente k^* ; o termo $(K+1)$ proveniente das permutabilidades; o quociente da distribuição *a priori* do modelos com o movimento *split* pelo *merge*; p_{k_j} é o peso da componente k_j , com $j = 1, 2$; τ é o hiperparâmetro da distribuição *a priori* de p_{k_j} ; n_j é o número de observações alocadas na componente j , com $j = 1, 2$ e $*$; $\pi(\xi_{k_j})$, $\pi(\omega_{k_j}^2)$ e $\pi(\alpha_{k_j})$ são, respectivamente, as distribuições *a priori* para ξ_{k_j} , $\omega_{k_j}^2$ e α_{k_j} , para $j = 1, 2$ e $*$; P_{aloc} é a probabilidade de alocação das componentes k_1 e k_2 em k^* ; $g(\cdot)$ é a densidade da distribuição *Beta*; e o último termo é o Jacobiano da transformação.

Maiores detalhes sobre o cálculo da probabilidade de aceitação dos movimentos “nascimento” (*birth*) e *split* ver, respectivamente, a Seção D.1 e a Seção D.3 do Apêndice D, e sobre o Jacobiano da transformação ver a Seção A.2 do Apêndice A.

Assim, o método *Reversible jump* para o modelo de misturas de distribuições normais assimétricas pode ser resumido nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$, e especifique um valor inicial para $\xi^{(0)}$, $\omega^{2(0)}$, $\alpha^{(0)}$ e $p^{(0)}$;
2. Para $t = 1, 2, \dots, T$:
 - (i) atualize ξ , ω^2 e α da distribuição *a posteriori*;
 - (ii) atualize p da distribuição *a posteriori*;
 - (iii) atualize s conforme descrito na [Subseção 4.4.2 do Capítulo 4](#);
 - (iv) escolha entre os movimentos *split* ou *merge* com probabilidade $Pr(sp) = Pr(me) = \frac{1}{2}$. Aceite o movimento *split*, de acordo com a probabilidade de aceitação dada por $\min\{1, A^{sp}\}$, ou aceite o movimento *merge* com a probabilidade de aceitação dada por $\min\{1, A^{sp-1}\}$ em que A^{sp} é dado em (5.15).
 - (v) escolha entre os movimentos “nascimento” (*birth*) ou “morte” (*death*) com probabilidade $Pr(n) = Pr(m) = \frac{1}{2}$. Aceite o movimento “nascimento” (*birth*), de acordo com probabilidade de aceitação dada por $\min\{1, A\}$, ou aceite o movimento “morte” (*death*), com a probabilidade de aceitação dada por $\min\{1, A^{-1}\}$ em que A é dado em (5.9).
3. Considere um *burn in* de B iterações, e calcule o número de vezes que o número de componentes é igual a K , denotado por N_K para $K = 1, \dots, K_{max}$ nas $T - B$ iterações;
4. Calcule a probabilidade *a posteriori* de K que é dada pela Expressão (4.14).

No método *Data-driven Reversible jump* a probabilidade de aceitação da proposta do movimento *split* é dada por $\Psi(x^{sp}|x) = \min\{1, A^{sp}\}$, em que A^{sp} para o modelo de mistura de distribuições normais assimétricas é dada por

$$A^{sp} = \frac{\mathcal{L}(\xi^{sp}, \omega^{2sp}, \alpha^{sp}, p^{sp} | y, s^{sp}, K+1) \pi(\xi^{sp}, \omega^{2sp}, \alpha^{sp}, p^{sp}, K+1) q(x|x^{sp})}{\mathcal{L}(\xi, \omega^2, \alpha, p | y, s, K) \pi(\xi, \omega^2, \alpha, p, K) q(x^{sp}|x)}. \quad (5.16)$$

Na Expressão (5.16) o primeiro quociente é a razão de verossimilhança do movimento *split*, dada por

$$\frac{\prod_{j=1}^2 \left[\frac{2}{\sqrt{2\pi}\omega_{s_{ij}}^{2sp}} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{sp}}\right)^2\right\} \int_{-\infty}^{\alpha_{s_{ij}}^{sp}} \left(\frac{y_i - \xi_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{sp}}\right) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du \right]}{\frac{2}{\sqrt{2\pi}\omega_{s_{i*}}^2} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_{s_{i*}}}{\omega_{s_{i*}}}\right)^2\right\} \int_{-\infty}^{\alpha_{s_{i*}}} \left(\frac{y_i - \xi_{s_{i*}}}{\omega_{s_{i*}}}\right) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du} \frac{\prod_{k=1}^{K+1} p_k^{n_k^{sp}}}{\prod_{k=1}^K p_k^{n_k}}, \quad (5.17)$$

o segundo quociente é a razão de distribuições *a priori*, dada por

$$K \times \frac{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi\omega_{s_{i_j}}^{2sp}}} \exp \left\{ -\frac{(\xi_{s_{i_j}}^{sp} - \mu_{s_{i_j}}^{sp})^2}{2\omega_{s_{i_j}}^{2sp}} \right\} \frac{b_{s_{i_j}}^{sp}}{\Gamma(a_{s_{i_j}}^{sp})} \left(\frac{1}{\omega_{s_{i_j}}^{2sp}} \right)^{-(a_{s_{i_j}}^{sp}+1)} \right]}{\frac{1}{\sqrt{2\pi\omega_{s_{i_*}}^2}} \exp \left\{ -\frac{(\xi_{s_{i_*}} - \mu_{s_{i_*}})^2}{2\omega_{s_{i_*}}^2} \right\} \frac{b_{s_{i_*}}}{\Gamma(a_{s_{i_*}})} \left(\frac{1}{\omega_{s_{i_*}}^2} \right)^{-(a_{s_{i_*}}+1)}} \exp \left\{ -\frac{b_{s_{i_j}}^{sp}}{\omega_{s_{i_j}}^{2sp}} \right\} \frac{1}{\sqrt{2\pi\sigma_{s_{i_j}}^{2sp}}} \exp \left\{ -\frac{(\alpha_{s_{i_j}}^{sp} - \mu_{s_{i_j}}^{sp})^2}{2\sigma_{s_{i_j}}^{2sp}} \right\}}{\exp \left\{ -\frac{b_{s_{i_*}}}{\omega_{s_{i_*}}^2} \right\} \frac{1}{\sqrt{2\pi\sigma_{s_{i_*}}^2}} \exp \left\{ -\frac{(\alpha_{s_{i_*}} - \mu_{s_{i_*}})^2}{2\sigma_{s_{i_*}}^2} \right\}}, \quad (5.18)$$

e finalmente, o terceiro quociente é a razão de distribuições propostas dada por

$$\frac{2n_{s_{i_1}^{sp}}n_{s_{i_2}^{sp}}}{n_{s_{i_*}}(n_{s_{i_*}}-1)} \left(\frac{1}{2} \right)^{\frac{n_{s_{i_1}^{sp}}+n_{s_{i_2}^{sp}}-2}{s_{i_1}^{sp}}}$$

$$\times \frac{2}{\sqrt{2\pi\omega_{s_{i_*}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_*}}}{\omega_{s_{i_*}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_*}}} \left(\frac{y_i - \xi_{s_{i_*}}}{\omega_{s_{i_*}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \prod_{k=1}^K p_k^{n_k}$$

$$\times \frac{1}{\sqrt{2\pi\omega_{s_{i_*}}^2}} \exp \left\{ -\frac{(\xi_{s_{i_*}} - \mu_{s_{i_*}})^2}{2\omega_{s_{i_*}}^2} \right\} \frac{b_{s_{i_*}}}{\Gamma(a_{s_{i_*}})} \left(\frac{1}{\omega_{s_{i_*}}^2} \right)^{-(a_{s_{i_*}}+1)} \exp \left\{ -\frac{b_{s_{i_*}}}{\omega_{s_{i_*}}^2} \right\} \frac{1}{\sqrt{2\pi\sigma_{s_{i_*}}^2}} \exp \left\{ -\frac{(\alpha_{s_{i_*}} - \mu_{s_{i_*}})^2}{2\sigma_{s_{i_*}}^2} \right\}$$

$$\times \frac{1}{\prod_{j=1}^2 \left[\frac{2}{\sqrt{2\pi\omega_{s_{i_j}}^{2sp}}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_j}}^{sp}}{\omega_{s_{i_j}}^{sp}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_j}}^{sp}} \left(\frac{y_i - \xi_{s_{i_j}}^{sp}}{\omega_{s_{i_j}}^{sp}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \right] \prod_{k=1}^{K+1} p_k^{n_k^{sp}}}$$

$$\times \frac{1}{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi\omega_{s_{i_j}}^{2sp}}} \exp \left\{ -\frac{(\xi_{s_{i_j}}^{sp} - \mu_{s_{i_j}}^{sp})^2}{2\omega_{s_{i_j}}^{2sp}} \right\} \frac{b_{s_{i_j}}^{sp}}{\Gamma(a_{s_{i_j}}^{sp})} \left(\frac{1}{\omega_{s_{i_j}}^{2sp}} \right)^{-(a_{s_{i_j}}^{sp}+1)} \exp \left\{ -\frac{b_{s_{i_j}}^{sp}}{\omega_{s_{i_j}}^{2sp}} \right\} \frac{1}{\sqrt{2\pi\sigma_{s_{i_j}}^{2sp}}} \exp \left\{ -\frac{(\alpha_{s_{i_j}}^{sp} - \mu_{s_{i_j}}^{sp})^2}{2\sigma_{s_{i_j}}^{2sp}} \right\} \right]}.$$
(5.19)

Na probabilidade de aceitação do movimento *split* as constantes normalizadoras não são conhecidas, pois as condicionais obtidas na [Seção 5.3](#) não são conjugadas para os parâmetros de locação, escala e assimetria (ver [Seção C.1 do Apêndice C](#)).

Maiores detalhes sobre as Expressões (5.17), (5.18) e (5.19) ver [Seção E.1 do Apêndice E](#).

Desta forma, descrevemos o método *Data-driven Reversible jump* para modelos de misturas de distribuições normais assimétricas nos seguintes passos:

1. Especifique um valor inicial para K e $\mathbf{s}^{(0)} = (s_1^{(0)}, \dots, s_N^{(0)})$ tal que $s_i^{(0)} \in \{1, \dots, K\}$, para $i = 1, \dots, N$;
2. Atualize $(\xi_k, \omega_k^2, \alpha_k, p_k)^t$ para $k = 1, \dots, K$ da distribuição *a posteriori* dada em (5.7);
3. *Data driven Reversible jump*: Atualize K .

Para $t = 1, \dots, T$:

- (a) Amostre duas observações, i_1 e i_2 , e decida entre os movimentos *split* ou *merge*;
- (b) Construa o s^* candidato e os valores candidatos x^* ;

- (c) Aceite a proposta com probabilidade $\Psi(x^*|x) = \min\{1, A^*\}$, em que $*$ é um movimento *split* (*sp*) ou um movimento *merge* (*mg*). Temos que A^{sp} para o modelo de mistura de normais assimétrica é dado em (5.16);
- i. Se o movimento *split* é aceito, faa $K^{(t)} = K^{(t-1)} + 1$, e considere x^{sp} ;
 - ii. Se o movimento *merge* é aceito, faa $K^{(t)} = K^{(t-1)} - 1$, e considere x^{mg} ;
 - iii. Se nenhum dos movimentos so aceitos, faa $K^{(t)} = K^{(t-1)}$, e considere x ;
4. Atualize s_i , para $i = 1, \dots, N$ de uma distribuio $Multinomial_{K^{(t)}}(1, \boldsymbol{\delta}_i = (\delta_1, \dots, \delta_{K^{(t)}})')$, em que $\delta_k = Pr(S_i = k | \mathbf{y}, \boldsymbol{\theta}, \mathbf{p}, K^{(t)})$ (Ver Subseo 4.4.2 do Capítulo 4).
5. Atualize $(\xi_k, \omega_k^2, \alpha_k, p_k)'$, para $k = 1, \dots, K^{(t)}$ da distribuio *a posteriori* dada em (5.7).

As estimativas dos parmetros do modelo de mistura de distribuies normais assimétricas foram atualizadas por meio do algoritmo *Metropolis-Hastings*, considerando o caso particular em que as distribuies propostas so simétricas, como descrito na Seo 4.2 do Capítulo 4.

A justificativa para a escolha desse algoritmo é que as distribuies condicionais obtidas na Seo 5.3 no so conjugadas para os parmetros de locao, escala e assimetria.

Desta forma, a partir do valor atualizado para K , o algoritmo *Metropolis-Hastings* para o modelo de mistura de distribuies normais assimétricas consiste nos seguintes passos:

1. Inicialize o contador de iteraes $t = 0$ e especifique um valor inicial para $\xi_k^{(0)}$, $\omega_k^{2(0)}$, $\alpha_k^{(0)}$ e $p_k^{(0)}$;
2. Gere um novo valor para:
 - (i) $\xi_k'^{(t)}$ da distribuio $q(\xi_k'^{(t)} | \xi_k)$, isto é, $\xi_k'^{(t)} \sim Normal(0, \omega_k^{2'(t-1)})$;
 - (ii) $\omega_k^{2'(t)}$ da distribuio $q(\omega_k^{2'(t)} | \omega_k^2)$, isto é, $\omega_k^{2'(t)} \sim Gama\ inversa(0.01, 0.01)$;
 - (iii) $\alpha_k'^{(t)}$ da distribuio $q(\alpha_k'^{(t)} | \alpha_k)$, isto é, $\alpha_k'^{(t)} \sim Normal(0, 100)$;
 - (iv) $p_k'^{(t)}$ da distribuio $q(p_k'^{(t)} | p_k)$, isto é, $p_k'^{(t)} \sim Dirichlet(1)$;
3. Calcule a probabilidade de aceitao $\alpha(\boldsymbol{\theta}_k, \boldsymbol{\theta}_k') = \min\{1, A\}$ em que A é dado por

$$A = \frac{\pi(\xi_k'^{(t)}, \omega_k^{2'(t)}, \alpha_k'^{(t)}, p_k'^{(t)} | \mathbf{y}, \mathbf{s})}{\pi(\xi_k^{(t-1)}, \omega_k^{2(t-1)}, \alpha_k^{(t-1)}, p_k^{(t-1)} | \mathbf{y}, \mathbf{s})}, \quad (5.20)$$

e gere um valor $u \sim U(0, 1)$;

4. Se $u \leq \alpha(\boldsymbol{\theta}_k, \boldsymbol{\theta}_k')$ aceite o novo valor, e faa $\xi_k^{(t+1)} = \xi_k'$, $\omega_k^{2(t+1)} = \omega_k^{2'}$, $\alpha_k^{(t+1)} = \alpha_k'$ e $p_k^{(t+1)} = p_k'$, caso contrrio, rejeite e faa $\xi_k^{(t+1)} = \xi_k^{(t)}$, $\omega_k^{2(t+1)} = \omega_k^{2(t)}$, $\alpha_k^{(t+1)} = \alpha_k^{(t)}$ e $p_k^{(t+1)} = p_k^{(t)}$.
5. Mude o contador de t para $t + 1$, e volte ao Passo 2, para $t = 1, 2, \dots, T$.

MODELO DE MISTURA DE DISTRIBUIÇÕES T DE STUDENT ASSIMÉTRICAS

Neste capítulo, apresentamos o modelo de mistura de distribuições t de Student assimétricas com um número de componentes conhecido e a sua função de verossimilhança. Também descrevemos os procedimentos utilizados para estimar o número de componentes juntamente com os parâmetros associados as componentes de mistura do modelo de mistura de distribuições t de Student assimétricas.

6.1 O modelo

Considerando que as densidades das componentes de mistura seguem a distribuição t de Student assimétrica como definida em (2.11), com vetor de parâmetros $\boldsymbol{\theta}_k = (\xi_k, \omega_k^2, \alpha_k, \nu_k)'$. Temos que a função densidade de probabilidade condicional, $f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k)$, pode ser escrita como

$$\begin{aligned}
 f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k) &= \frac{2}{\omega_k} \frac{\Gamma\left(\frac{\nu_k+1}{2}\right)}{\Gamma\left(\frac{\nu_k}{2}\right) \sqrt{\nu_k} \pi} \left(1 + \frac{\left(\frac{y_i - \xi_k}{\omega_k}\right)^2}{\nu_k}\right)^{-\left(\frac{\nu_k+1}{2}\right)} \\
 &\times \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \frac{\sqrt{\frac{\nu_k+2}{\nu_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}}}{\Gamma\left(\frac{\nu_k+1}{2}\right) \sqrt{(\nu_k+1)} \pi} \Gamma\left(\frac{\nu_k+2}{2}\right) \left(1 + \frac{u^2}{\nu_k+1}\right)^{-\left(\frac{\nu_k+2}{2}\right)} du.
 \end{aligned} \tag{6.1}$$

em que ξ_k é o parâmetro de locação, ω_k^2 é o parâmetro de escala, α_k é o parâmetro de assimetria e ν_k são os graus de liberdade, da k -ésima componente de mistura, com $k = 1, \dots, K$.

Logo, podemos escrever o modelo de mistura de distribuições t de Student assimétricas com um número de componente K conhecido, correspondente a Expressão (3.6) do [Capítulo 3](#),

por

$$\begin{aligned}
f_{Y_i}(y|\boldsymbol{\theta}, \mathbf{p}) &= \sum_{k=1}^K f_{Y_i}(y|S_i = k, \boldsymbol{\theta}_k) Pr(S_i = k|\mathbf{p}) \\
&= \sum_{k=1}^K p_k \frac{2}{\omega_k} \frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2})\sqrt{v_k\pi}} \left(1 + \frac{(y_i - \xi_k)^2}{v_k}\right)^{-\left(\frac{v_k+1}{2}\right)} \\
&\quad \times \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2})\sqrt{(v_k+1)\pi}} \left(1 + \frac{u^2}{v_k+1}\right)^{-\left(\frac{v_k+2}{2}\right)} du,
\end{aligned} \tag{6.2}$$

em que $(\xi_k, \omega_k^2, \alpha_k, v_k, p_k)'$ são os parâmetros da k -ésima componente do modelo a serem estimados, para $k = 1, \dots, K$.

Assim, denotamos o modelo de mistura de distribuições t de Student assimétricas com o número de componentes K conhecido por $Y_i|S_i = k, \boldsymbol{\theta}, \mathbf{p} \sim TA_{v_k}(\xi_k, \omega_k^2, \alpha_k)$, com $k = 1, \dots, K$.

6.2 Função de verossimilhança

Para misturas de distribuições t de Student assimétricas, temos que a função de verossimilhança conjunta de \mathbf{y} e \mathbf{s} , conforme a Expressão (3.9) do Capítulo 3, é dada por

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p}|\mathbf{y}, \mathbf{s}, K) &= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\omega_k}\right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2})\sqrt{v_k\pi}}\right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k}\right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
&\quad \times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \left\{ \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2})} \right. \\
&\quad \left. \left. \times \sqrt{(v_k+1)\pi} \left(1 + \frac{u^2}{v_k+1}\right)^{-\left(\frac{v_k+2}{2}\right)} du \right\} \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \right\}
\end{aligned} \tag{6.3}$$

em que $n_k = \sum_{i=1}^N I_{\{s_i\}}(k)$ é o número de observações da k -ésima componente de mistura.

6.3 Abordagem Bayesiana

Para misturas de t de Student assimétricas temos que $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v})'$. Como $\boldsymbol{\xi} \perp \boldsymbol{\alpha}$, $\boldsymbol{\xi} \perp \mathbf{v}$, $\boldsymbol{\omega}^2 \perp \boldsymbol{\alpha}$, $\boldsymbol{\omega}^2 \perp \mathbf{v}$ e $\boldsymbol{\alpha} \perp \mathbf{v}$, logo a distribuição *a priori* conjunta se resume em

$$\begin{aligned}
\pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p}, K) &= \pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}|K)\pi(\mathbf{p}|K)\pi(K) \\
&= \pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K)\pi(\boldsymbol{\omega}^2|K)\pi(\boldsymbol{\alpha}|K)\pi(\mathbf{v}|K)\pi(\mathbf{p}|K)\pi(K).
\end{aligned} \tag{6.4}$$

Considerando a k -ésima componente do modelo de mistura de distribuições t de Student assimétricas, a distribuição *a priori* pode ser escrita como

$$\pi(\boldsymbol{\theta}_k, p_k, k) = \pi(\xi_k | \omega_k^2) \pi(\omega_k^2) \pi(\alpha_k) \pi(\mathbf{v}_k) \pi(p_k) \pi(k). \quad (6.5)$$

Neste trabalho, utilizamos distribuições *a priori* vagas para todos os parâmetros dos modelos, ou seja, distribuições *a priori* com pouca informação. Desta forma, tomamos as distribuições *a priori* descritas abaixo para os seguintes parâmetros:

1. $\xi_k | \omega_k^2 \sim Normal(\mu_k, \omega_k^2)$, em que μ_k é conhecida, com $\mu_k \in (-\infty, +\infty)$, e ω_k^2 é condicionada, com $\omega_k^2 > 0$, para $k = 1, \dots, K$;
2. $\omega_k^2 \sim Gama\ inversa(a_k, b_k)$, em que a_k e b_k são conhecidos, com $a_k > 0$ e $b_k > 0$, para $k = 1, \dots, K$;
3. $\alpha_k \sim Normal(\mu_k, \sigma_k^2)$ em que μ_k e σ_k^2 são conhecidas, com $\mu_k \in (-\infty, +\infty)$, e $\sigma_k^2 > 0$, para $k = 1, \dots, K$;
4. $\mathbf{v}_k \sim Gama(c_k, d_k)$, em que c_k e d_k são conhecidos, com $c_k > 0$ e $d_k > 0$, para $k = 1, \dots, K$.
5. $p_k \sim Dirichlet(\boldsymbol{\tau}_k)$, em que $\boldsymbol{\tau}_k$ é conhecido, com $\tau_k > 0$, para $k = 1, \dots, K$;
6. $k \sim U_d(K_{max})$, em que U_d representa uma distribuição Uniforme discreta no conjunto $\{1, \dots, K_{max}\}$ e K_{max} é o valor máximo especificado para K .

A distribuição *a posteriori* para o modelo de misturas de distribuições t de Student assimétricas, com vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v})'$, correspondente a Expressão (4.2) do Capítulo 4, pode ser escrita como

$$\begin{aligned} \pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p}, K | \mathbf{y}, \mathbf{s}) &\propto \mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) \\ &\times \pi(\boldsymbol{\xi} | \boldsymbol{\omega}^2 | K) \pi(\boldsymbol{\omega}^2 | K) \pi(\boldsymbol{\alpha} | K) \pi(\mathbf{v} | K) \pi(\mathbf{p} | K) \pi(K). \end{aligned} \quad (6.6)$$

Desta forma, a distribuição *a posteriori* da k -ésima componente do modelo de mistura de distribuições t de Student assimétricas é dada por

$$\begin{aligned} \pi(\xi_k, \omega_k^2, \alpha_k, \mathbf{v}_k, p_k | \mathbf{y}, \mathbf{s}) &\propto \mathcal{L}(\xi_k, \omega_k^2, \alpha_k, \mathbf{v}_k, p_k | \mathbf{y}, \mathbf{s}) \\ &\times \pi(\xi_k | \omega_k^2) \pi(\omega_k^2) \pi(\alpha_k) \pi(\mathbf{v}_k) \pi(p_k) \pi(k). \end{aligned} \quad (6.7)$$

As distribuições condicionais para os parâmetros do modelo de mistura de distribuições t de Student assimétricas também não são conjugadas, exceto para o parâmetro peso da componente (ver Seção C.2 do Apêndice C).

Maiores detalhes sobre os cálculos da função de verossimilhança e da distribuição *a posteriori* do modelo de mistura de distribuições t de Student assimétricas ver, respectivamente, a Seção B.2 do Apêndice B e a Seção C.2 do Apêndice C.

6.4 Estimação

De acordo com as distribuições *a priori* descritas na Seção 6.3, escolhemos os seguintes valores para os hiperparâmetros: $K \sim U_d(10)$; $p_k \sim Dirichlet(1)$; $\xi_k \sim Normal(0, \omega_k^2)$; $\omega_k^2 \sim Gama\ inversa(0.01, 0.01)$; $\alpha_k \sim Normal(0, 100)$; e $v_k \sim Gama(100, 100)$, para $k = 1, \dots, 10$.

Utilizamos os métodos *Reversible jump* e o *Data-driven Reversible jump* para estimar o número de componentes K do modelo de mistura de distribuições, ou seja, aplicamos os movimentos “nascimento” (*birth*), “morte” (*death*), *split* e *merge*, como descrito na Subseção 4.4.1 e na Seção 4.5 do Capítulo 4, considerando que as componentes de mistura seguem a distribuição t de Student assimétrica.

Desta forma, no movimento “nascimento” (*birth*) do método *Reversible jump*, nós adicionamos uma componente “vazia” com vetor de parâmetros $(\xi_{k^*}, \omega_{k^*}^2, \alpha_{k^*}, v_{k^*}, p_{k^*})'$ gerados a partir das distribuições *a priori*, descritas no início dessa seção, exceto para o parâmetro p_{k^*} que foi gerado de uma distribuição Beta $(1, K)$, conforme Richardson e Green (1997). No movimento de “morte” (*death*) removemos uma componente “vazia” na mistura, e em ambos os movimentos reescalamos os pesos como descrito na Subseção 4.4.1 do Capítulo 4.

Atribuímos as probabilidades para os movimentos “nascimento” (*birth*), $Pr(n)$, e “morte” (*death*), $Pr(m)$, de acordo com a Expressão (5.8) do Capítulo 5, e a probabilidade de aceitação da proposta do movimento “nascimento” (*birth*) é dada por $\min\{1, A\}$ e do reverso “morte” (*death*) é aceita com probabilidade $\min\{1, A^{-1}\}$, em que A é dado por

$$\begin{aligned}
 A &= \frac{\pi(\xi|\omega^2, K+1)\pi(\omega^2|K+1)\pi(\alpha|K+1)\pi(\mathbf{v}|K+1)}{\pi(\xi|\omega^2, K)\pi(\omega^2|K)\pi(\alpha|K)\pi(\mathbf{v}|K)} \\
 &\times \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{Beta(K\tau, \tau)} \\
 &\times \frac{q(K|K+1)}{q(K+1|K)} \frac{1}{(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} \frac{1}{(1-p_{k^*})^K}.
 \end{aligned} \tag{6.8}$$

Na Expressão (6.8) temos na primeira linha, a razão das probabilidades *a priori* do modelo com o movimento “nascimento” (*birth*) de uma componente pelo atual. Na segunda linha, a razão das distribuições *a priori* de *Dirichlet* com parâmetro τ , em que N é o número total de observações, $Beta(\cdot, \cdot)$ é a função beta e o termo $(K+1)$ é devido a razão $\frac{(K+1)!}{K!}$, proveniente da suposição de permutabilidade para as componentes de mistura, pois existem, respectivamente, $K!$ e $(K+1)!$ formas de escrever θ_K e θ_{K+1} . Na terceira linha, a primeira divisão é a razão das probabilidades dos movimentos “nascimento” (*birth*), pelo “morte” (*death*), o termo (k_0+1) surge da probabilidade de se escolher uma componente para o movimento “morte” (*death*), em que k_0 é o número de componentes “vazias” antes do movimento “nascimento” (*birth*); o terceiro quociente é o inverso da função densidade de probabilidade da distribuição $Beta(1, K)$ do qual p_{k^*} foi gerado; e o último termo é o jacobiano da transformação. A razão de verossimilhança é 1

devido as propostas dos movimentos envolverem componentes “vazias”.

Para o movimento *merge* escolhemos as componentes de mistura de distribuições t de Student assimétricas adotando a mesma metodologia de escolha de componentes, do modelo de mistura de normais assimétricas, como descrito na [Seção 5.4](#) do [Capítulo 5](#).

Assim, a nova componente k^* com vetor de parâmetros dado por $(\xi_{k^*}, \omega_{k^*}^2, \alpha_{k^*}, \nu_{k^*}, p_{k^*})'$, possui parâmetros calculados de acordo com [Richardson e Green \(1997\)](#), isto é,

$$\left. \begin{aligned} p_{k^*} &= p_{k_1} + p_{k_2}, \\ p_{k^*} \xi_{k^*} &= p_{k_1} \xi_{k_1} + p_{k_2} \xi_{k_2}, \\ p_{k^*} (\xi_{k^*}^2 + \omega_{k^*}^2) &= p_{k_1} (\xi_{k_1}^2 + \omega_{k_1}^2) + p_{k_2} (\xi_{k_2}^2 + \omega_{k_2}^2), \\ p_{k^*} \alpha_{k^*} &= p_{k_1} \alpha_{k_1} + p_{k_2} \alpha_{k_2}, \\ p_{k^*} \nu_{k^*} &= p_{k_1} \nu_{k_1} + p_{k_2} \nu_{k_2}. \end{aligned} \right\} \quad (6.9)$$

O vetor $\mathbf{u} = (u_1, u_2, u_3, u_4, u_5)'$ possui as seguintes distribuições para suas componentes $u_1 \sim \text{Beta}(2, 2)$, $u_2 \sim \text{Beta}(2, 2)$, $u_3 \sim \text{Beta}(1, 1)$, $u_4 \sim \text{Beta}(2, 2)$ e $u_5 \sim \text{Beta}(1, 1)$ e as equações abaixo fornecem os dez novos parâmetros necessários para a transformação, satisfazendo (6.9), isto é,

$$p_{k_1} = p_{k^*} u_1 \quad , \quad p_{k_2} = p_{k^*} (1 - u_1), \quad (6.10)$$

$$\xi_{k_1} = \xi_{k^*} - u_2 \omega_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}}, \quad \xi_{k_2} = \xi_{k^*} + u_2 \omega_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}}, \quad (6.11)$$

$$\omega_{k_1}^2 = u_3 (1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{p_{k_1}}, \quad \omega_{k_2}^2 = (1 - u_3) (1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{p_{k_2}}, \quad (6.12)$$

$$\alpha_{k_1} = \alpha_{k^*} u_4 \quad , \quad \alpha_{k_2} = \alpha_{k^*} (1 - u_4), \quad (6.13)$$

$$\nu_{k_1} = \nu_{k^*} u_5 \quad , \quad \nu_{k_2} = \nu_{k^*} (1 - u_5). \quad (6.14)$$

As transformações necessárias para a obtenção dos parâmetros da componente k^* (6.9), as distribuições das componentes do vetor \mathbf{u} e a Equações de (6.10) a (6.14), também são baseadas em [Richardson e Green \(1997\)](#).

A probabilidade de aceitação da proposta do movimento *split* é dada por $\min\{1, A^{sp}\}$, e do reverso *merge* é aceita com probabilidade $\min\{1, A^{sp^{-1}}\}$, em que A para o modelo de mistura

de distribuições t de Student assimétricas é dado por

$$\begin{aligned}
A^{SP} &= \left[\left(\frac{2}{\omega_{k_1}} \right)^{n_{k_1}} \left(\frac{\Gamma(\frac{v_{k_1}+1}{2})}{\Gamma(\frac{v_{k_1}}{2})\sqrt{v_{k_1}\pi}} \right)^{n_{k_1}} \prod_{i=1}^{N} \left(1 + \frac{(y_i - \xi_{k_1})^2}{\omega_{k_1}^2 v_{k_1}} \right)^{-\left(\frac{v_{k_1}+1}{2}\right)} \right. \\
&\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_{k_1} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right) \sqrt{\frac{v_{k_1}+2}{v_{k_1}+1 + \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_1}+2}{2})}{\Gamma(\frac{v_{k_1}+1}{2})\sqrt{(v_{k_1}+1)\pi}} \left(1 + \frac{u^2}{v_{k_1}+1} \right)^{-\left(\frac{v_{k_1}+2}{2}\right)} du \right\} \Big]^{I_{\{s_i\}}(k_1)} \\
&\times \left[\left(\frac{2}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}\pi}} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right. \\
&\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right) \sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \Big]^{I_{\{s_i\}}(k_2)} \\
&\times \frac{1}{\left[\left(\frac{1}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}\pi}} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right]} \\
&\frac{\prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right) \sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \Big]^{I_{\{s_i\}}(k_2)}} \\
&\times (K+1) \frac{\pi(\xi|\omega^2, K+1)\pi(\omega^2|K+1)\pi(\alpha|K+1)\pi(\mathbf{v}|K+1)}{\pi(\xi|\omega^2, K)\pi(\omega^2|K)\pi(\alpha|K)\pi(\mathbf{v}|K)} \\
&\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{\text{Beta}(\tau, k\tau)} \\
&\times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \times \frac{\pi(v_{k_1})\pi(v_{k_2})}{\pi(v_{k^*})} \\
&\times \frac{1}{P_{alloc}g(u_1)g(u_2)g(u_3)g(u_4)g(u_5)} \times \frac{p_{k^*}|\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1} v_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4u_5\omega_{k^*}^2}. \tag{6.15}
\end{aligned}$$

Na Expressão (6.15) temos a razão do produto das funções de verossimilhança das componentes k_1 e k_2 pela função de verossimilhança da componente k^* ; o termo $(K+1)$ proveniente das permutabilidades, o quociente das distribuições *a priori* do modelo com o movimento *split* pelo *merge*; p_{k_j} é o peso da componente k_j , com $j = 1, 2$; τ é o hiperparâmetro da distribuição *a priori* de p_{k_j} ; n_j é o número de observações alocadas na componente j , com $j = 1, 2$ e $*$; $\pi(\xi_{k_j})$, $\pi(\omega_{k_j}^2)$, $\pi(\alpha_{k_j})$ e $\pi(v_{k_j})$ são, respectivamente, as distribuições *a priori* para ξ_{k_j} , $\omega_{k_j}^2$, α_{k_j} e v_{k_j} , para $j = 1, 2$ e $*$; P_{alloc} é a probabilidade de alocação das componentes k_1 e k_2 em k^* ; $g(\cdot)$ é a densidade da distribuição *Beta*; e o último termo é o Jacobiano da transformação.

Maiores detalhes sobre o cálculo da probabilidade de aceitação dos movimentos “nascimento” (*birth*) e *split* ver, respectivamente, a Seção D.2 e a Seção D.4 do Apêndice D, e sobre o Jacobiano da transformação ver a Seção A.3 do Apêndice A.

Desta forma, o método *Reversible jump* para o modelo de misturas de distribuições t de Student assimétricas pode ser descrito nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$, e especifique um valor inicial para $\xi^{(0)}$, $\omega^{2(0)}$, $\alpha^{(0)}$, $\mathbf{v}^{(0)}$ e $\mathbf{p}^{(0)}$;
2. Para $t = 1, 2, \dots, T$:
 - (i) atualize ξ , ω^2 , \mathbf{v} e α da distribuição *a posteriori*;
 - (ii) atualize \mathbf{p} da distribuição *a posteriori*;
 - (iii) atualize \mathbf{s} conforme descrito na [Subseção 4.4.2](#) do [Capítulo 4](#);
 - (iv) escolha entre os movimentos *split* ou *merge* com probabilidade $Pr(sp) = Pr(me) = \frac{1}{2}$. Aceite o movimento *split*, de acordo com a probabilidade de aceitação dada por $\min\{1, A^{sp}\}$, ou aceite o movimento *merge* com a probabilidade de aceitação dada por $\min\{1, A^{sp-1}\}$ em que A^{sp} é dado em (6.15).
 - (v) escolha entre os movimentos “nascimento” (*birth*) ou “morte” (*death*) com probabilidade $Pr(n) = Pr(m) = \frac{1}{2}$. Aceite o movimento “nascimento” (*birth*), de acordo com probabilidade de aceitação dada por $\min\{1, A\}$, ou aceite o movimento “morte” (*death*), com a probabilidade de aceitação dada por $\min\{1, A^{-1}\}$ em que A é dado em A é dado em (6.8).
3. Considere um *burn in* de B iterações, e calcule o número de vezes que o número de componentes é igual a K , denotado por N_K para $K = 1, \dots, K_{max}$ nas $T - B$ iterações;
4. Calcule a probabilidade *a posteriori* de K que é dada pela Expressão (4.14).

No método *Data-driven Reversible jump* a probabilidade de aceitação da proposta do movimento *split* é dada por $\Psi(x^{sp}|x) = \min\{1, A^{sp}\}$, em que A^{sp} para o modelo de mistura t de Student assimétricas é dada por

$$A^{sp} = \frac{\mathcal{L}(\xi^{sp}, \omega^{2sp}, \alpha^{sp}, \mathbf{v}^{sp}, \mathbf{p}^{sp} | \mathbf{y}, \mathbf{s}^{sp}, K+1) \pi(\xi^{sp}, \omega^{2sp}, \alpha^{sp}, \mathbf{p}^{sp}, K+1) q(x|x^{sp})}{\mathcal{L}(\xi, \omega^2, \alpha, \mathbf{v}, \mathbf{p} | \mathbf{y}, \mathbf{s}, K) \pi(\xi, \omega^2, \alpha, \mathbf{p}, \mathbf{v}, K) q(x^{sp}|x)}. \quad (6.16)$$

Na Expressão (6.16) o primeiro quociente é a razão de verossimilhança do movimento

split, dada por

$$\begin{aligned}
& \prod_{j=1}^2 \left[\frac{2}{\omega_{s_{ij}}^{sp}} \frac{\Gamma\left(\frac{v_{s_{ij}}^{sp}+1}{2}\right)}{\Gamma\left(\frac{v_{s_{ij}}^{sp}}{2}\right)} \frac{1}{\sqrt{v_{s_{ij}}^{sp}} \pi} \left(1 + \frac{\left(\frac{y_i - \xi_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{sp}}\right)^2}{v_{s_{ij}}^{sp}}\right)^{-\left(\frac{v_{s_{ij}}^{sp}+1}{2}\right)} \right. \\
& \times \left. \int_{-\infty}^{\infty} \alpha_{s_{ij}}^{sp} \left(\frac{y_i - \xi_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{sp}}\right) \sqrt{\frac{v_{s_{ij}}^{sp}+2}{v_{s_{ij}}^{sp}+1 + \left(\frac{y_i - \xi_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{sp}}\right)^2}} \frac{\Gamma\left(\frac{v_{s_{ij}}^{sp}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{ij}}^{sp}+1}{2}\right)} \frac{1}{\sqrt{(v_{s_{ij}}^{sp}+1)\pi}} \left(1 + \frac{u^2}{v_{s_{ij}}^{sp}+1}\right)^{-\left(\frac{v_{s_{ij}}^{sp}+2}{2}\right)} du \right] \\
& \times \frac{1}{\frac{2}{\omega_{s_{i*}}^{sp}} \frac{\Gamma\left(\frac{v_{s_{i*}}^{sp}+1}{2}\right)}{\Gamma\left(\frac{v_{s_{i*}}^{sp}}{2}\right)} \frac{1}{\sqrt{v_{s_{i*}}^{sp}} \pi} \left(1 + \frac{\left(\frac{y_i - \xi_{s_{i*}}^{sp}}{\omega_{s_{i*}}^{sp}}\right)^2}{v_{s_{i*}}^{sp}}\right)^{-\left(\frac{v_{s_{i*}}^{sp}+1}{2}\right)}} \\
& \times \frac{1}{\int_{-\infty}^{\infty} \alpha_{s_{i*}}^{sp} \left(\frac{y_i - \xi_{s_{i*}}^{sp}}{\omega_{s_{i*}}^{sp}}\right) \sqrt{\frac{v_{s_{i*}}^{sp}+2}{v_{s_{i*}}^{sp}+1 + \left(\frac{y_i - \xi_{s_{i*}}^{sp}}{\omega_{s_{i*}}^{sp}}\right)^2}} \frac{\Gamma\left(\frac{v_{s_{i*}}^{sp}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{i*}}^{sp}+1}{2}\right)} \frac{1}{\sqrt{(v_{s_{i*}}^{sp}+1)\pi}} \left(1 + \frac{u^2}{v_{s_{i*}}^{sp}+1}\right)^{-\left(\frac{v_{s_{i*}}^{sp}+2}{2}\right)} du} \frac{\prod_{k=1}^{K+1} p_k^{sp n_k^{sp}}}{\prod_{k=1}^K p_k^{n_k}}, \tag{6.17}
\end{aligned}$$

o segundo quociente é a razão de distribuições *a priori*, dada por

$$\begin{aligned}
& K \times \frac{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi\omega_{s_{ij}}^{2sp}}} \exp\left\{-\frac{(\xi_{s_{ij}}^{sp} - \mu_{s_{ij}}^{sp})^2}{2\omega_{s_{ij}}^{2sp}}\right\} \frac{b_{s_{ij}}^{sp}}{\Gamma(a_{s_{ij}}^{sp})} \left(\frac{1}{\omega_{s_{ij}}^{2sp}}\right)^{-(a_{s_{ij}}^{sp}+1)} \right]}{\frac{1}{\sqrt{2\pi\omega_{s_{i*}}^2}} \exp\left\{-\frac{(\xi_{s_{i*}} - \mu_{s_{i*}})^2}{2\omega_{s_{i*}}^2}\right\} \frac{b_{s_{i*}}^{a_{s_{i*}}}}{\Gamma(a_{s_{i*}})} \left(\frac{1}{\omega_{s_{i*}}^2}\right)^{-(a_{s_{i*}}+1)}} \\
& \frac{\exp\left\{-\frac{b_{s_{ij}}^{sp}}{\omega_{s_{ij}}^{2sp}}\right\} \frac{1}{\sqrt{2\pi\sigma_{s_{ij}}^{2sp}}} \exp\left\{-\frac{(\alpha_{s_{ij}}^{sp} - \mu_{s_{ij}}^{sp})^2}{2\sigma_{s_{ij}}^{2sp}}\right\}}{\exp\left\{-\frac{b_{s_{i*}}}{\omega_{s_{i*}}^2}\right\} \frac{1}{\sqrt{2\pi\sigma_{s_{i*}}^2}} \exp\left\{-\frac{(\alpha_{s_{i*}} - \mu_{s_{i*}})^2}{2\sigma_{s_{i*}}^2}\right\}}. \tag{6.18}
\end{aligned}$$

e finalmente, o terceiro quociente é a razão de distribuições propostas dada por

$$\begin{aligned}
& \frac{2n_{s_1^{sp}}n_{s_2^{sp}}}{n_{s_{i^*}}(n_{s_{i^*}}-1)} \left(\frac{1}{2}\right)^{n_{s_1^{sp}}+n_{s_2^{sp}}-2} \omega_{s_{i^*}} \frac{2}{\Gamma\left(\frac{v_{s_{i^*}}}{2}\right)} \frac{\Gamma\left(\frac{v_{s_{i^*}}+1}{2}\right)}{\sqrt{v_{s_{i^*}}}\pi} \left(1 + \frac{\left(\frac{y_i - \xi_{s_{i^*}}}{\omega_{s_{i^*}}}\right)^2}{v_{s_{i^*}}}\right)^{-\left(\frac{v_{s_{i^*}}+1}{2}\right)} \\
& \times \int_{-\infty}^{\infty} \alpha_{s_{i^*}} \left(\frac{y_i - \xi_{s_{i^*}}}{\omega_{s_{i^*}}}\right) \sqrt{\frac{v_{s_{i^*}}+2}{v_{s_{i^*}}+1 + \left(\frac{y_i - \xi_{s_{i^*}}}{\omega_{s_{i^*}}}\right)^2}} \frac{\Gamma\left(\frac{v_{s_{i^*}}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{i^*}}+1}{2}\right)} \frac{1}{\sqrt{(v_{s_{i^*}}+1)\pi}} \left(1 + \frac{u^2}{v_{s_{i^*}}+1}\right)^{-\left(\frac{v_{s_{i^*}}+2}{2}\right)} du \prod_{k=1}^{K+1} p_k^{n_k} \\
& \times \frac{1}{\prod_{j=1}^2 \left[\frac{2}{\omega_{s_{i_j}^{sp}}} \frac{\Gamma\left(\frac{v_{s_{i_j}^{sp}}+1}{2}\right)}{\Gamma\left(\frac{v_{s_{i_j}^{sp}}}{2}\right)} \frac{1}{\sqrt{v_{s_{i_j}^{sp}}}\pi} \left(1 + \frac{\left(\frac{y_i - \xi_{s_{i_j}^{sp}}}{\omega_{s_{i_j}^{sp}}}\right)^2}{v_{s_{i_j}^{sp}}}\right)^{-\left(\frac{v_{s_{i_j}^{sp}}+1}{2}\right)} \right]} \\
& \times \frac{\int_{-\infty}^{\infty} \alpha_{s_{i_j}^{sp}} \left(\frac{y_i - \xi_{s_{i_j}^{sp}}}{\omega_{s_{i_j}^{sp}}}\right) \sqrt{\frac{v_{s_{i_j}^{sp}}+2}{v_{s_{i_j}^{sp}}+1 + \left(\frac{y_i - \xi_{s_{i_j}^{sp}}}{\omega_{s_{i_j}^{sp}}}\right)^2}} \frac{\Gamma\left(\frac{v_{s_{i_j}^{sp}}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{i_j}^{sp}}+1}{2}\right)} \frac{1}{\sqrt{(v_{s_{i_j}^{sp}}+1)\pi}} \left(1 + \frac{u^2}{v_{s_{i_j}^{sp}}+1}\right)^{-\left(\frac{v_{s_{i_j}^{sp}}+2}{2}\right)} du \prod_{k=1}^K p_k^{s_p^{n_k}}}{1} \\
& \times \frac{\frac{1}{\sqrt{2\pi}\omega_{s_{i^*}}^2} \exp\left\{-\frac{(\xi_{s_{i^*}} - \mu_{s_{i^*}})^2}{2\omega_{s_{i^*}}^2}\right\} \frac{b_{s_{i^*}}^{a_{s_{i^*}}}}{\Gamma(a_{s_{i^*}})} \left(\frac{1}{\omega_{s_{i^*}}^2}\right)^{-(a_{s_{i^*}}+1)}}{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi}\omega_{s_{i_j}^{sp}}^2} \exp\left\{-\frac{(\xi_{s_{i_j}^{sp}} - \mu_{s_{i_j}^{sp}})^2}{2\omega_{s_{i_j}^{sp}}^2}\right\} \frac{b_{s_{i_j}^{sp}}^{a_{s_{i_j}^{sp}}}}{\Gamma(a_{s_{i_j}^{sp}})} \left(\frac{1}{\omega_{s_{i_j}^{sp}}^2}\right)^{-(a_{s_{i_j}^{sp}}+1)} \right]} \\
& \times \frac{\exp\left\{-\frac{b_{s_{i^*}}}{\omega_{s_{i^*}}^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_{s_{i^*}}^2} \exp\left\{-\frac{(\alpha_{s_{i^*}} - \mu_{s_{i^*}})^2}{2\sigma_{s_{i^*}}^2}\right\}}{\exp\left\{-\frac{b_{s_{i_j}^{sp}}}{\omega_{s_{i_j}^{sp}}^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_{s_{i_j}^{sp}}^2} \exp\left\{-\frac{(\alpha_{s_{i_j}^{sp}} - \mu_{s_{i_j}^{sp}})^2}{2\sigma_{s_{i_j}^{sp}}^2}\right\}}}. \tag{6.19}
\end{aligned}$$

Na probabilidade de aceitação do movimento *split* as constantes normalizadoras não são conhecidas, pois as condicionais obtidas na Seção 6.3 não são conjugadas para os parâmetros de localização, escala e assimetria (ver Seção C.2 do Apêndice C).

Maiores detalhes sobre as Expressões (6.17), (6.18) e (6.19) ver Seção E.2 do Apêndice E.

Desta forma, descrevemos o método *Data-driven Reversible jump* para modelos de misturas de distribuições t de Student assimétricas nos seguintes passos:

1. Especifique um valor inicial para K e $\mathbf{s}^{(0)} = (s_1^{(0)}, \dots, s_N^{(0)})$ tal que $s_i^{(0)} \in \{1, \dots, K\}$, para $i = 1, \dots, N$;
2. Atualize $(\xi_k, \omega_k^2, \alpha_k, v_k, p_k)'$ para $k = 1, \dots, K$ da distribuição *a posteriori* dada em (6.7);
3. *Data driven Reversible jump*: Atualize K .
Para $t = 1, \dots, T$:

- (a) Amostre duas observações, i_1 e i_2 , e decida entre os movimentos *split* ou *merge*;
- (b) Construa o s^* candidato e os valores candidatos x^* ;
- (c) Aceite a proposta com probabilidade $\Psi(x^*|x) = \min\{1, A^*\}$, em que $*$ é um movimento *split* (*sp*) ou um movimento *merge* (*mg*). Temos que A^{sp} para o modelo de mistura de t de Student assimétrica é dado em (6.16);
 - i. Se o movimento *split* é aceito, faça $K^{(t)} = K^{(t-1)} + 1$, e considere x^{sp} ;
 - ii. Se o movimento *merge* é aceito, faça $K^{(t)} = K^{(t-1)} - 1$, e considere x^{mg} ;
 - iii. Se nenhum dos movimentos são aceitos, faça $K^{(t)} = K^{(t-1)}$, e considere x ;
4. Atualize s_i , para $i = 1, \dots, N$ de uma distribuição $Multinomial_{K^{(t)}}(1, \boldsymbol{\delta}_i = (\delta_1, \dots, \delta_{K^{(t)}})')$, em que $\delta_k = Pr(S_i = k | \mathbf{y}, \boldsymbol{\theta}, \mathbf{p}, K^{(t)})$ (Ver Subseção 4.4.2 do Capítulo 4).
5. Atualize $(\xi_k, \omega_k^2, \alpha_k, \nu_k, p_k)'$, para $k = 1, \dots, K^{(t)}$ da distribuição *a posteriori* dada em (6.7).

As estimativas dos parâmetros do modelo de mistura de distribuições t de Student assimétricas também foram atualizadas por meio do algoritmo *Metropolis-Hastings*, considerando o caso particular em que as distribuições propostas são simétricas, como descrito na Seção 4.2 do Capítulo 4, pois também as distribuições *a posteriori* obtidas na Seção 6.3 não são conjugadas para os parâmetros de locação, escala, assimetria e graus de liberdade, sendo essa a justificativa para a escolha desse algoritmo.

Desta forma, a partir do valor atualizado para K , o algoritmo *Metropolis-Hastings* para o modelo de mistura de distribuições t de Student assimétricas pode ser descrito nos seguintes passos:

1. Inicialize o contador de iterações $t = 0$ e especifique um valor inicial para $\xi_k^{(0)}$, $\omega_k^{2(0)}$, $\alpha_k^{(0)}$, $\nu_k^{(0)}$ e $p_k^{(0)}$;
2. Gere um novo valor para:
 - (i) $\xi_k^{(t)}$ da distribuição $q(\xi_k^{(t)} | \xi_k)$, isto é, $\xi_k^{(t)} \sim Normal(0, \omega_k^{2(t-1)})$;
 - (ii) $\omega_k^{2(t)}$ da distribuição $q(\omega_k^{2(t)} | \omega_k^2)$, isto é, $\omega_k^{2(t)} \sim Gama\ inversa(0.01, 0.01)$;
 - (iii) $\alpha_k^{(t)}$ da distribuição $q(\alpha_k^{(t)} | \alpha_k)$, isto é, $\alpha_k^{(t)} \sim Normal(0, 100)$;
 - (iv) $\nu_k^{(t)}$ da distribuição $q(\nu_k^{(t)} | \nu_k)$, isto é, $\nu_k^{(t)} \sim Gama(100, 100)$;
 - (v) $p_k^{(t)}$ da distribuição $q(p_k^{(t)} | p_k)$, isto é, $p_k^{(t)} \sim Dirichlet(1)$;
3. Calcule a probabilidade de aceitação $\alpha(\boldsymbol{\theta}_k, \boldsymbol{\theta}_k') = \min\{1, A\}$ em que A é dado por

$$A = \frac{\pi(\xi_k^{(t)}, \omega_k^{2(t)}, \alpha_k^{(t)}, \nu_k^{(t)}, p_k^{(t)} | \mathbf{y}, \mathbf{s})}{\pi(\xi_k^{(t-1)}, \omega_k^{2(t-1)}, \alpha_k^{(t-1)}, \nu_k^{(t-1)}, p_k^{(t-1)} | \mathbf{y}, \mathbf{s})}, \quad (6.20)$$

e gere um valor $u \sim U(0, 1)$;

4. Se $u \leq \alpha(\boldsymbol{\theta}_k, \boldsymbol{\theta}'_k)$ aceite o novo valor, e faça $\xi_k^{(t+1)} = \xi'_k$, $\omega_k^{2(t+1)} = \omega_k^{2'}$, $\alpha_k^{(t+1)} = \alpha'_k$, $\mathbf{v}_k^{(t+1)} = \mathbf{v}'_k$ e $p_k^{(t+1)} = p'_k$, caso contrário, rejeite e faça $\xi_k^{(t+1)} = \xi_k^{(t)}$, $\omega_k^{2(t+1)} = \omega_k^{2(t)}$, $\alpha_k^{(t+1)} = \alpha_k^{(t)}$, $\mathbf{v}_k^{(t+1)} = \mathbf{v}_k^{(t)}$ e $p_k^{(t+1)} = p_k^{(t)}$.

5. Mude o contador de t para $t + 1$, e volte ao Passo 2 para $t = 1, 2, \dots, T$.

RESULTADOS

Neste capítulo, descrevemos os resultados encontrados no estudo de simulação e na análises de dados reais.

7.1 Simulação

Neste estudo de simulação, geramos observações da variável aleatória \mathbf{Y} considerando os modelos de mistura de distribuições normais assimétricas e t de Student assimétricas utilizando, respectivamente, as funções *rsn* e *rst* da *library sn*, as quais estão disponíveis no *Software R* (R Core Team, 2022). Também geramos observações da variável aleatória \mathbf{Y} considerando o modelo de mistura de distribuições normais usando a função *rnorm* do pacote básico do *Software R* (R Core Team, 2022).

Estimamos o número de componentes K dos modelos de mistura de distribuições em estudo, por meio da estimativa da probabilidade *a posteriori* de K , obtida pelos métodos *Data-driven Reversible jump* e *Reversible jump*. Atualizamos as estimativas dos parâmetros associados as componentes dos modelos de misturas empregando o algoritmo *Metropolis-Hastings*, pelo motivo já descrito na [Seção 5.4](#) do [Capítulo 5](#), e também atualizamos os valores da variável não observável \mathbf{S} , aplicando o algoritmo *Gibbs sampling*. Nas simulações, consideramos 55.000 iterações, um *burn-in* de 5.000 iterações e saltos de 100 iterações.

Assim, simulamos dados de modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas com $K = 1, 2, 3$ e 4 componentes, com os seguintes vetores de proporções ou pesos $\mathbf{p} = (1)'$, $\mathbf{p} = (0.3, 0.7)'$, $\mathbf{p} = (0.5, 0.2, 0.3)'$, $\mathbf{p} = (0.3, 0.1, 0.4, 0.2)'$, respectivamente, e tamanhos de amostras $N = 100, 200, 500$ e 1000.

Descrevemos na Tabela 1, os valores tomados para os parâmetros, correspondente a cada componente, dos modelos de mistura em estudo.

Tabela 1 – Valores dos parâmetros das componentes dos modelos de mistura, empregados nos estudos de simulação.

Modelos de mistura	Componente 1	Componente 2	Componente 3	Componente 4
Normal	$\mu_1 = -1.0$	$\mu_2 = 3.0$	$\mu_3 = 7.0$	$\mu_4 = 12.0$
	$\sigma_1 = 0.7$	$\sigma_2 = 1.2$	$\sigma_3 = 1.0$	$\sigma_4 = 1.5$
Normal assimétrica	$\xi_1 = -1.0$	$\xi_2 = 5.0$	$\xi_3 = 10.0$	$\xi_4 = 17.0$
	$\omega_1 = 1.7$	$\omega_2 = 1.0$	$\omega_3 = 1.5$	$\omega_4 = 2.0$
	$\alpha_1 = -1.0$	$\alpha_2 = -0.5$	$\alpha_3 = 1.0$	$\alpha_4 = 1.5$
<i>t</i> de Student assimétrica	$\xi_1 = -1.0$	$\xi_2 = 5.0$	$\xi_3 = 10.0$	$\xi_4 = 19.0$
	$\omega_1 = 1.5$	$\omega_2 = 0.5$	$\omega_3 = 2.0$	$\omega_4 = 1.7$
	$\alpha_1 = 1.0$	$\alpha_2 = -0.5$	$\alpha_3 = 1.0$	$\alpha_4 = -1.5$
	$\nu_1 = 4.0$	$\nu_2 = 4.0$	$\nu_3 = 4.0$	$\nu_4 = 4.0$

Por exemplo, simulamos dados de um modelo de mistura de distribuições normais com $K = 2$ componentes, com vetor de proporções $\mathbf{p} = (p_1, p_2)' = (0.3, 0.7)'$; com os seguintes vetores de parâmetros $\boldsymbol{\mu} = (\mu_1, \mu_2)' = (-1.0, 3.0)'$ e $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)' = (0.7, 1.2)'$, com tamanho de amostras $N = 100$, em seguida, calculamos a estimativa da probabilidade *a posteriori* de K pelos métodos *Data-driven Reversible jump* e *Reversible jump*. Repetimos o mesmo procedimento para os tamanhos de amostras $N = 200, 500$ e 1000 .

Apresentamos as Tabelas 2, 3 e 4 com as estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* e *Reversible jump* para os respectivos modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas.

Tabela 2 – Estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições normais, dados gerados com $K = 1, 2, 3$ e 4 componentes, e tamanhos de amostras $N = 100, 200, 500$ e 1000.

K simulado	K	N=100		N=200		N=500		N=1000	
		DDRJ	RJ	DDRJ	RJ	DDRJ	RJ	DDRJ	RJ
1	1	0.999	0.982	0.999	0.965	0.999	0.975	0.999	0.988
	2	0.001	0.018	0.001	0.033	0.001	0.024	0.001	0.012
	≥ 3	-	-	-	0.002	-	0.001	-	-
2	1	0.013	0.001	0.001	-	-	-	-	-
	2	0.987	0.786	0.999	0.543	0.999	0.854	0.965	0.554
	3	-	0.171	-	0.336	0.001	0.106	0.035	0.288
	≥ 4	-	0.042	-	0.121	-	0.040	-	0.158
3	1	-	-	-	-	-	-	-	-
	2	0.053	0.020	0.005	0.001	0.001	0.003	0.002	0.238
	3	0.947	0.571	0.995	0.476	0.999	0.422	0.998	0.411
	4	-	0.255	-	0.320	-	0.320	-	0.248
	≥ 5	-	0.154	-	0.203	-	0.255	-	0.103
4	1	-	-	-	-	-	-	-	-
	2	0.015	0.014	0.002	0.003	0.002	0.002	0.003	0.001
	3	0.022	0.026	0.004	0.029	0.101	0.012	0.002	0.009
	4	0.963	0.488	0.994	0.466	0.897	0.406	0.995	0.624
	5	-	0.348	-	0.335	-	0.362	-	0.310
	≥ 6	-	0.124	-	0.167	-	0.218	-	0.056

Observamos por meio da [Tabela 2](#) que o método *Data-driven Reversible jump* apresenta maior valor para as estimativas da probabilidade *a posteriori* de K , para todos os valores reais simulados de K e tamanhos de amostras dos modelos de mistura de distribuições normais. Observamos também que o método *Data-driven Reversible jump* se move entre os modelos mais próximos do verdadeiro modelo simulado, sendo essa uma das vantagens da aplicação desse método.

Tabela 3 – Estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições normais assimétricas, dados gerados com $K = 1, 2, 3$ e 4 componentes e tamanhos de amostras $N = 100, 200, 500$ e 1000.

K simulado	K	N=100		N=200		N=500		N=1000	
		DDRJ	RJ	DDRJ	RJ	DDRJ	RJ	DDRJ	RJ
1	1	0.999	0.606	0.999	0.949	0.996	0.961	0.996	0.749
	2	0.001	0.249	0.001	0.047	0.004	0.036	0.004	0.237
	≥ 3	-	0.145	-	0.004	-	0.003	-	0.014
2	1	0.046	0.001	0.001	-	-	-	-	-
	2	0.954	0.726	0.999	0.681	0.999	0.412	0.999	0.391
	3	-	0.198	-	0.211	0.001	0.291	0.001	0.383
	4	-	0.053	-	0.067	-	0.203	-	0.163
	≥ 5	-	0.022	-	0.041	-	0.094	-	0.063
3	1	-	-	-	-	-	-	-	-
	2	0.044	0.010	0.007	0.006	0.005	0.005	0.001	-
	3	0.956	0.439	0.993	0.587	0.995	0.431	0.999	0.340
	4	-	0.314	-	0.267	-	0.304	-	0.310
	5	-	0.160	-	0.106	-	0.169	-	0.217
	≥ 6	-	0.077	-	0.034	-	0.091	-	0.133
4	1	-	-	-	-	-	-	-	-
	2	0.014	0.007	0.004	0.005	0.001	0.004	0.002	0.003
	3	0.032	0.033	0.009	0.090	0.004	0.327	0.013	0.008
	4	0.954	0.365	0.987	0.367	0.995	0.368	0.985	0.507
	5	-	0.336	-	0.326	-	0.291	-	0.372
	≥ 6	-	0.259	-	0.212	-	0.010	-	0.110

Observamos por meio da [Tabela 3](#) que o método *Data-driven Reversible jump* apresenta maior valor para as estimativas da probabilidade *a posteriori* de K , para todos os valores reais simulados de K e tamanhos de amostras dos modelos de mistura de distribuições normais assimétricas. Novamente observamos que o método *Data-driven Reversible jump* se move entre os modelos mais próximos do verdadeiro modelo simulado.

Tabela 4 – Estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições t de Student assimétricas, dados gerados com $K = 1, 2, 3$ e 4 componentes e tamanhos de amostras $N = 100, 200, 500$ e 1000 .

K simulado	K	N=100		N=200		N=500		N=1000	
		DDRJ	RJ	DDRJ	RJ	DDRJ	RJ	DDRJ	RJ
1	1	0.999	0.357	0.998	0.332	0.998	0.335	0.999	0.327
	2	0.001	0.257	0.002	0.281	0.002	0.259	0.001	0.299
	3	-	0.194	-	0.230	-	0.231	-	0.214
	≥ 4	-	0.192	-	0.157	-	0.175	-	0.160
2	1	0.024	0.001	0.001	-	-	-	-	-
	2	0.976	0.361	0.999	0.437	0.999	0.515	0.999	0.360
	3	-	0.258	-	0.245	0.001	0.334	0.001	0.276
	4	-	0.236	-	0.191	-	0.112	-	0.246
	≥ 5	-	0.144	-	0.127	-	0.039	-	0.118
3	1	-	-	-	-	-	-	-	-
	2	0.037	0.008	0.004	0.004	0.180	0.007	0.001	-
	3	0.963	0.352	0.996	0.490	0.820	0.644	0.999	0.891
	4	-	0.299	-	0.316	-	0.274	-	0.096
	5	-	0.260	-	0.155	-	0.062	-	0.010
	≥ 6	-	0.081	-	0.035	-	0.013	-	0.003
4	1	-	-	-	-	-	-	-	-
	2	0.017	0.008	0.001	0.004	0.002	0.003	0.001	0.005
	3	0.021	0.023	0.003	0.007	0.003	0.006	0.038	0.006
	4	0.962	0.415	0.996	0.808	0.995	0.513	0.961	0.570
	5	-	0.362	-	0.166	-	0.299	-	0.260
	≥ 6	-	0.192	-	0.015	-	0.179	-	0.159

Observamos por meio da [Tabela 4](#) que o método *Data-driven Reversible jump* continua apresentando maior valor para as estimativas da probabilidade *a posteriori* de K , para todos os valores reais simulados de K e tamanhos de amostras dos modelos de mistura de distribuições t de Student assimétricas. Vemos também por meio da [Tabela 4](#) que o método *Data-driven Reversible jump* permanece se movendo entre os modelos mais próximos do verdadeiro modelo simulado.

Apresentamos [Tabela 5](#) com as médias *a posteriori* e os intervalos de 95% de credibilidade (IC) para os parâmetros, obtidos pelo método *Data-driven Reversible jump* para os modelos de mistura de distribuições normais, para os tamanhos de amostras $N = 500$ e $N = 1000$.

Tabela 5 – Médias *a posteriori* e Intervalos de 95% de Credibilidade dos parâmetros dos modelos de mistura de distribuições normais, obtidas pelo método *Data-driven Reversible jump*, com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$.

	Parâmetros	Médias (Intervalos de Credibilidade)			
		$N = 500$		$N = 1000$	
$K = 1$	$p_1 = 1.0$	1.00	(1.00; 1.00)	1.00	(1.00; 1.00)
	$\mu_1 = -1.0$	-1.06	(-1.12; -1.00)	-1.02	(-1.07; -0.98)
	$\sigma_1 = 0.7$	0.70	(0.66; 0.75)	0.69	(0.66; 0.72)
$K = 2$	$p_1 = 0.3$	0.31	(0.26; 0.35)	0.31	(0.28; 0.34)
	$\mu_1 = -1.0$	-0.94	(-1.08; -0.79)	-1.00	(-1.09; -0.88)
	$\sigma_1 = 0.7$	0.75	(0.68; 0.88)	0.75	(0.69; 0.90)
	$p_2 = 0.7$	0.69	(0.65; 0.74)	0.69	(0.66; 0.72)
	$\mu_2 = 3.0$	2.84	(2.71; 3.00)	3.01	(2.90; 3.10)
	$\sigma_2 = 1.2$	1.15	(1.02; 1.25)	1.18	(1.09; 1.26)
$K = 3$	$p_1 = 0.5$	0.50	(0.46; 0.55)	0.50	(0.47; 0.54)
	$\mu_1 = -1.0$	-0.99	(-1.09; -0.90)	-1.04	(-1.11; -0.97)
	$\sigma_1 = 0.7$	0.75	(0.68; 0.86)	0.72	(0.66; 0.83)
	$p_2 = 0.2$	0.19	(0.15; 0.23)	0.19	(0.16; 0.22)
	$\mu_2 = 3.0$	2.67	(2.40; 3.00)	3.03	(2.81; 3.25)
	$\sigma_2 = 1.2$	1.02	(0.88; 1.21)	1.10	(0.96; 1.25)
	$p_3 = 0.3$	0.31	(0.27; 0.36)	0.31	(0.27; 0.34)
	$\mu_3 = 7.0$	6.74	(6.54; 7.01)	6.96	(6.80; 7.11)
$\sigma_3 = 1.0$	1.02	(0.91; 1.14)	1.01	(0.92; 1.12)	
$K = 4$	$p_1 = 0.3$	0.31	(0.27; 0.35)	0.31	(0.27; 0.33)
	$\mu_1 = -1.0$	-0.96	(-1.09; -0.83)	-1.03	(-1.11; -0.94)
	$\sigma_1 = 0.7$	0.77	(0.68; 0.91)	0.74	(0.67; 0.90)
	$p_2 = 0.1$	0.08	(0.06; 0.12)	0.09	(0.07; 0.12)
	$\mu_2 = 3.0$	2.66	(2.16; 3.18)	2.81	(2.50; 3.10)
	$\sigma_2 = 1.2$	1.10	(0.91; 1.26)	1.04	(0.88; 1.21)
	$p_3 = 0.4$	0.41	(0.37; 0.46)	0.41	(0.38; 0.45)
	$\mu_3 = 7.0$	6.72	(6.55; 7.02)	6.99	(6.87; 7.10)
	$\sigma_3 = 1.0$	1.03	(0.94; 1.13)	1.05	(0.99; 1.13)
	$p_4 = 0.2$	0.20	(0.17; 0.24)	0.19	(0.17; 0.22)
	$\mu_4 = 12.0$	11.60	(11.20; 11.90)	12.10	(11.90; 12.30)
	$\sigma_4 = 1.5$	1.32	(1.13; 1.55)	1.31	(1.08; 1.52)

Verificamos por meio da [Tabela 5](#) que as médias *a posteriori* dos parâmetros dos modelos de mistura de distribuições normais com $K = 1, 2, 3, 4$ componentes, estão bem próximas do verdadeiro valor dos parâmetros, e que os intervalos de 95% credibilidade contém os verdadeiros valores dos parâmetros, indicando que o método *Data-driven Reversible jump* consegue fornecer estimativas da probabilidade *a posteriori* do número de componentes dos modelos de mistura de distribuições normais, e também de estimar os parâmetros do modelo conjuntamente.

Esboçamos as Figuras 5, 6, 7 e 8 com os histogramas e as distribuições empírica dos dados, ambas com as densidades reais (linhas pretas) e estimadas (linha azul) dos modelos de mistura de distribuições normais, para os dados gerados, respectivamente, com $K = 1, 2, 3, 4$ componentes, $N = 1000$, e com o número de componentes estimados pelas probabilidades *a posteriori* de K , obtidas pelos métodos *Data-driven Reversible jump*, de acordo com a Tabela 2.

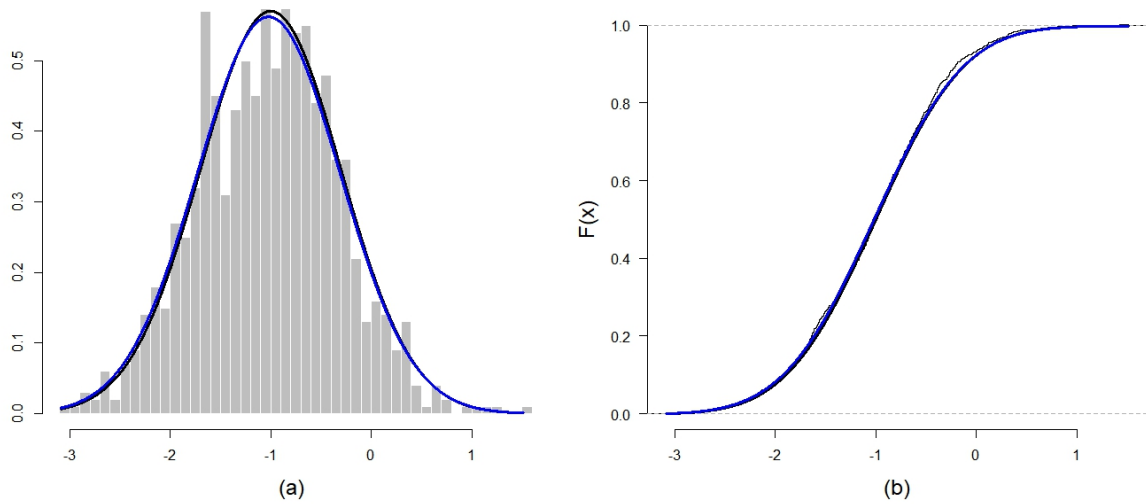


Figura 5 – Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

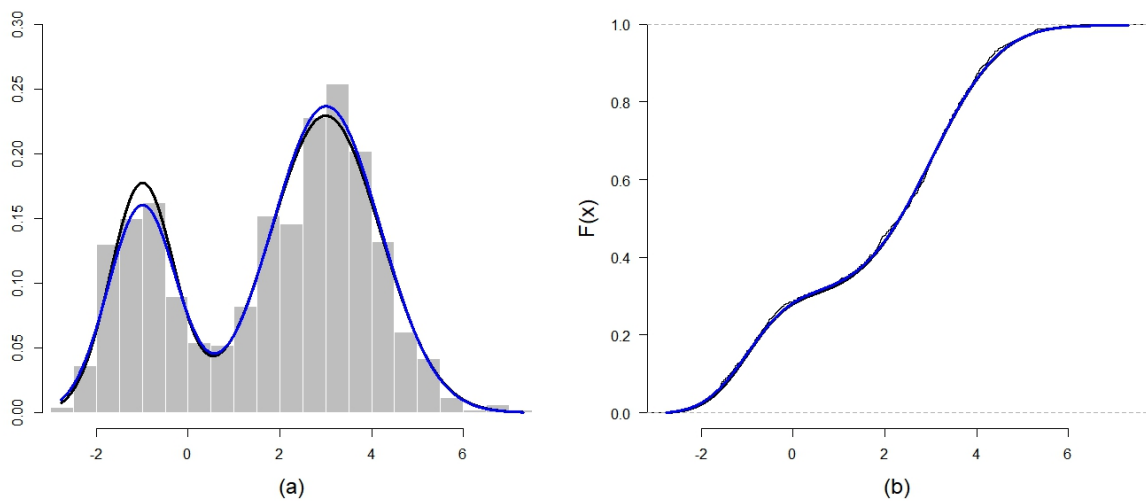


Figura 6 – Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

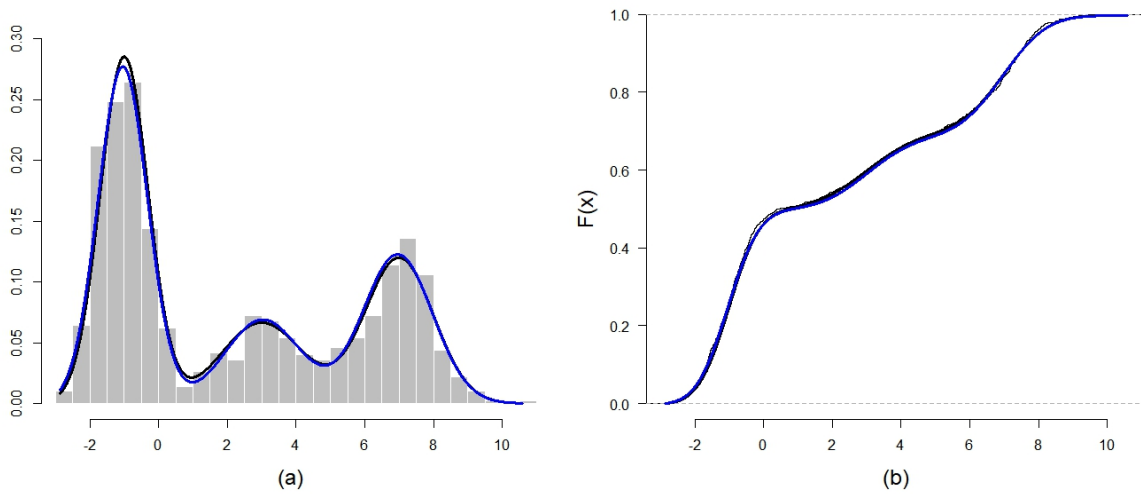


Figura 7 – Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

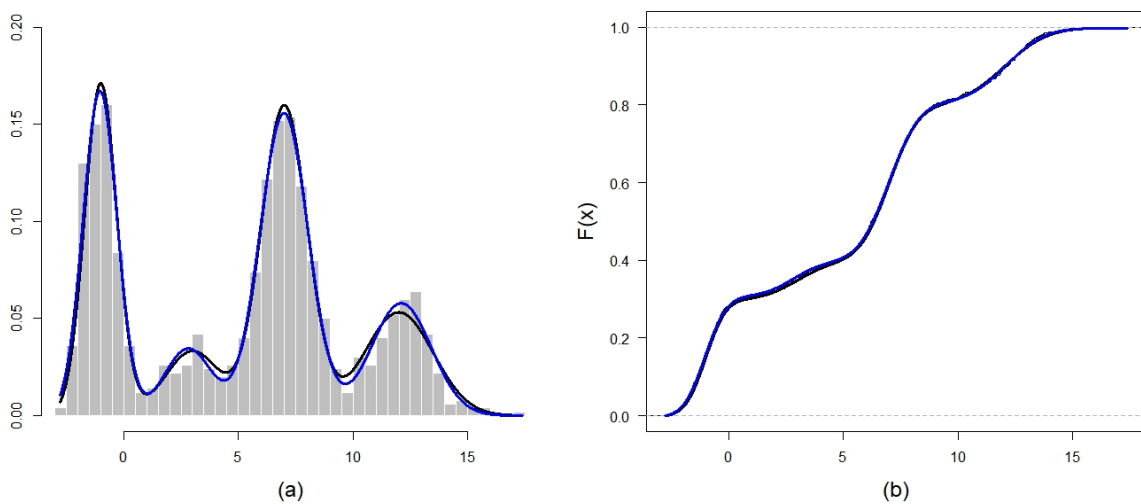


Figura 8 – Densidade real (linha preta) e estimada (linha azul) dos modelos de mistura de distribuições normais, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

Podemos observar por meio das Figuras 5, 6, 7 e 8 que tanto nos histogramas quanto nas distribuições empíricas, as densidades dos modelos estimados estão bem próximas dos modelos reais, indicando novamente que o método *Data-driven Reversible jump* fornece a estimativa da probabilidade *a posteriori* para o número de componentes do modelos de mistura de distribuições normais, e conjuntamente estimativas para os parâmetros dos modelos que sugerem um bom ajuste para os dados.

Apresentamos Tabela 6 com as médias *a posteriori* e os intervalos de 95% de credibilidade (IC) para os parâmetros, obtidos pelo método *Data-driven Reversible jump* para os modelos de mistura de distribuições normais assimétricas, para os tamanhos de amostras $N = 500$ e $N = 1000$.

Tabela 6 – Médias *a posteriori* e Intervalos de 95% de Credibilidade dos parâmetros dos modelos de mistura de distribuições normais assimétricas, obtidas pelo método *Data-driven Reversible jump*, com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$.

Parâmetros		Médias (Intervalos de Credibilidade)			
		$N = 500$		$N = 1000$	
$K = 1$	$p_1 = 1.0$	1.00	(1.00; 1.00)	1.00	(1.00, 1.00)
	$\xi_1 = -1.0$	-2.26	(-3.04; -1.16)	-2.39	(-3.06; -1.31)
	$\omega_1 = 1.7$	1.53	(1.36; 1.79)	1.54	(1.36; 1.80)
	$\alpha_1 = -1.0$	0.33	(-0.72; 1.29)	0.46	(-0.50; 1.25)
$K = 2$	$p_1 = 0.3$	0.30	(0.25; 0.34)	0.30	(0.27; 0.33)
	$\xi_1 = -1.0$	-1.88	(-2.49; -1.35)	-1.82	(-2.43; -1.25)
	$\omega_1 = 1.7$	1.25	(1.12; 1.41)	1.35	(1.22; 1.53)
	$\alpha_1 = -1.0$	-0.14	(-0.71; 0.48)	-0.20	(-0.88; 0.39)
	$p_2 = 0.7$	0.70	(0.66; 0.75)	0.70	(0.67; 0.73)
	$\xi_2 = 2.0^*$	2.25	(1.83; 2.90)	2.22	(1.91; 2.82)
	$\omega_2 = 1.0$	1.08	(0.94; 1.25)	1.05	(0.92; 1.20)
	$\alpha_2 = -0.5$	-0.58	(-0.34; 1.40)	0.72	(-0.21; 1.35)
$K = 3$	$p_1 = 0.5$	0.50	(0.45; 0.54)	0.50	(0.47; 0.53)
	$\xi_1 = -1.0$	-1.88	(-2.34; -1.41)	-1.87	(-2.36; -1.37)
	$\omega_1 = 1.7$	1.32	(1.20; 1.45)	1.37	(1.27; 1.48)
	$\alpha_1 = -1.0$	-0.11	(-0.58; 0.33)	-0.10	(-0.59; 0.36)
	$p_2 = 0.2$	0.20	(0.17; 0.24)	0.20	(0.18; 0.23)
	$\xi_2 = 5.0$	4.06	(3.61; 4.69)	4.26	(3.89; 4.91)
	$\omega_2 = 1.0$	1.04	(0.87; 1.33)	1.07	(0.93; 1.24)
	$\alpha_2 = -0.5$	0.78	(-0.23; 1.79)	0.62	(-0.30; 1.29)
	$p_3 = 0.3$	0.30	(0.26; 0.34)	0.30	(0.27; 0.33)
	$\xi_3 = 10.0$	10.00	(9.58; 10.60)	10.00	(9.69; 10.60)
	$\omega_3 = 1.5$	1.39	(1.17; 1.65)	1.48	(1.22; 1.71)
	$\alpha_3 = 1.0$	0.92	(0.22; 1.76)	1.03	(0.26; 1.67)
$K = 4$	$p_1 = 0.3$	0.30	(0.26; 0.34)	0.30	(0.27; 0.33)
	$\xi_1 = -1.0$	-1.87	(-2.47; -1.37)	-1.86	(-2.32; -1.45)
	$\omega_1 = 1.7$	1.28	(1.13; 1.42)	1.33	(1.12; 1.42)
	$\alpha_1 = -1.0$	-0.05	(-0.53; 0.52)	-0.16	(-0.57; 0.26)
	$p_2 = 0.1$	0.10	(0.07; 0.13)	0.10	(0.08; 0.12)
	$\xi_2 = 5.0$	4.05	(3.71; 4.74)	4.16	(3.79; 4.74)
	$\omega_2 = 1.0$	1.00	(0.87; 1.12)	0.98	(0.89; 1.14)
	$\alpha_2 = -0.5$	0.98	(-0.50; 1.61)	0.48	(-0.35; 1.12)
	$p_3 = 0.4$	0.40	(0.36; 0.45)	0.40	(0.37; 0.43)
	$\xi_3 = 10.0$	10.10	(9.67; 10.60)	10.30	(9.85; 10.80)
	$\omega_3 = 1.5$	1.34	(1.13; 1.55)	1.39	(1.09; 1.64)
	$\alpha_3 = 1.0$	0.77	(0.14; 1.46)	0.72	(0.14; 1.33)
	$p_4 = 0.2$	0.20	(0.16; 0.24)	0.20	(0.18; 0.23)
	$\xi_4 = 17.0$	17.20	(16.80; 17.80)	17.20	(16.70; 18.40)
	$\omega_4 = 2.0$	1.60	(1.18; 1.85)	1.76	(1.12; 2.13)
	$\alpha_4 = 1.5$	1.19	(0.39; 1.89)	1.28	(-0.08; 2.21)

(*) Alteramos o valor desse parâmetro para uma melhor visualização gráfica.

Verificamos por meio da [Tabela 6](#) que a maioria das médias *a posteriori* dos parâmetros dos modelos de mistura de distribuições normais assimétricas com $K = 1, 2, 3, 4$ componentes, estão bem próximas do verdadeiro valor dos parâmetros, e que a maioria dos intervalos de 95% credibilidade contém os verdadeiros valores dos parâmetros. Esse resultados indicam mais uma vez que o método *Data-driven Reversible jump* consegue fornecer estimativas da probabilidade *a posteriori* do número de componentes dos modelos de mistura de distribuições normais assimétricas, e também de estimar os parâmetros dos modelos conjuntamente.

Esboçamos as [Figuras 9, 10, 11 e 12](#) com os histogramas e as distribuições empírica dos dados, ambas com as densidades reais (linhas pretas) e estimadas (linhas vermelhas) dos modelos de mistura de distribuições normais assimétricas, para os dados gerados, respectivamente, com $K = 1, 2, 3, 4$ componentes, $N = 1000$, e com o número de componentes estimados pelas probabilidades *a posteriori* de K , obtidas pelos métodos *Data-driven Reversible jump*, de acordo com a [Tabela 3](#).

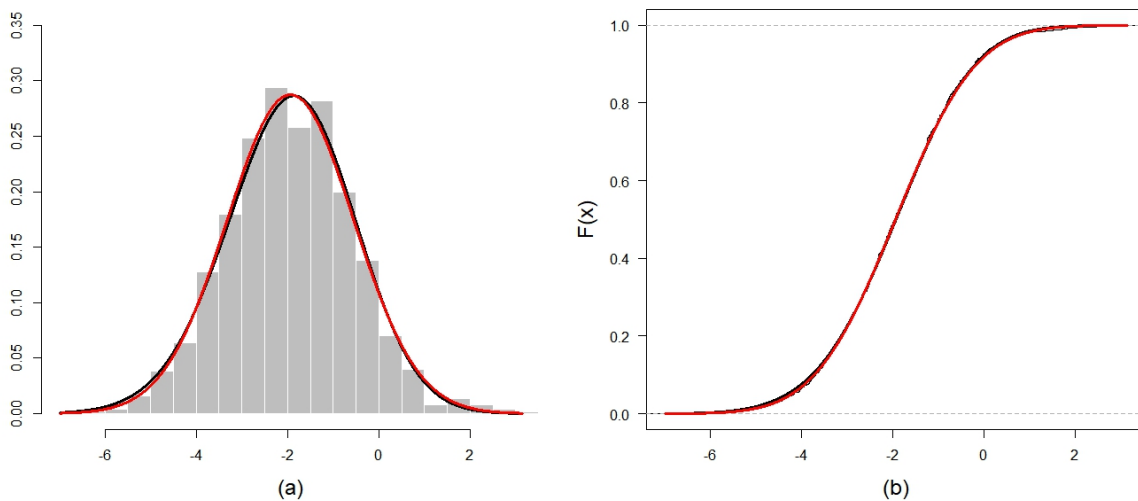


Figura 9 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

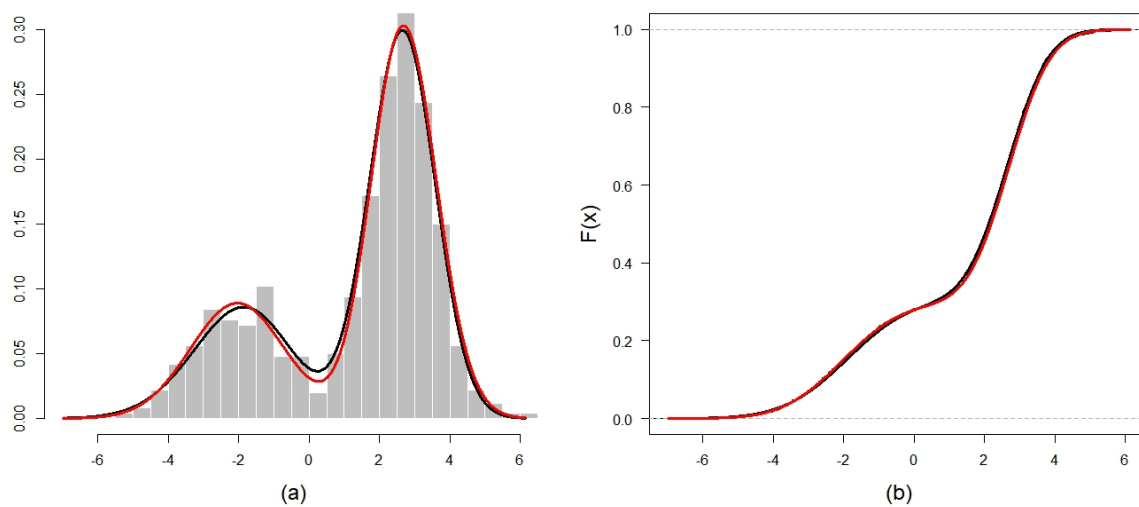


Figura 10 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

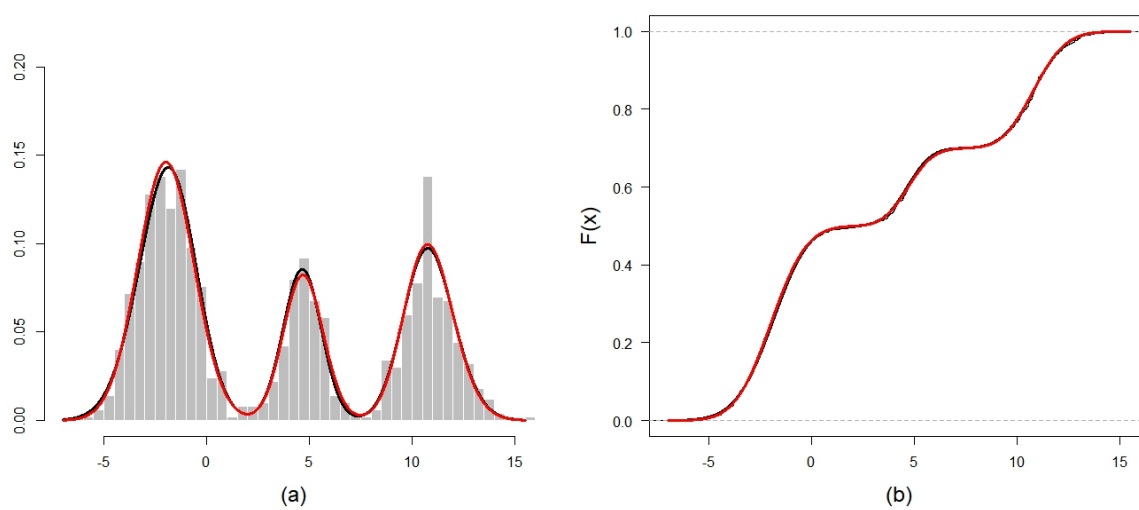


Figura 11 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

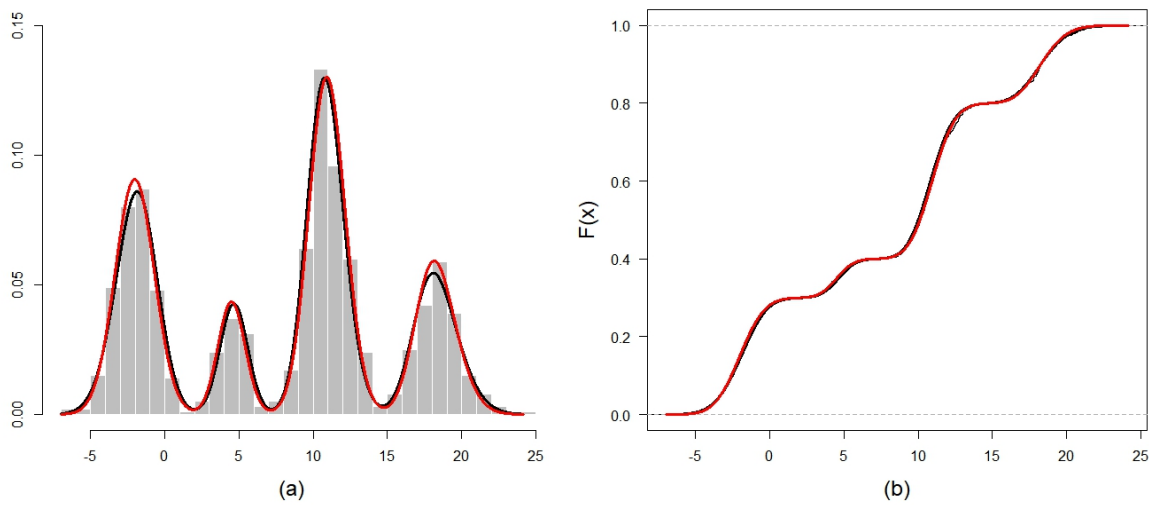


Figura 12 – Densidade real (linha preta) e estimada (linha vermelha) dos modelos de mistura de distribuições normais assimétricas, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

Podemos observar por meio das Figuras 9, 10, 11 e 12 que tanto nos histogramas quanto nas distribuições empíricas, as densidades dos modelos estimados estão bem próximas dos modelos reais. Esses resultados sugerem mais uma vez que o método *Data-driven Reversible jump* além de fornecer a estimativa da probabilidade *a posteriori* para o número de componentes do modelos de mistura de distribuições normais assimétrica, também fornece conjuntamente estimativas para os parâmetros dos modelos que sugerem um bom ajuste para os dados.

Apresentamos Tabela 7 com as médias *a posteriori* e os intervalos de 95% de credibilidade (IC) para os parâmetros, obtidos pelo método *Data-driven Reversible jump* para os modelos de mistura de distribuições *t* de Student assimétricas, para os tamanhos de amostras $N = 500$ e $N = 1000$.

Tabela 7 – Médias *a posteriori* e Intervalo de 95% de Credibilidade para parâmetros dos modelos de mistura de distribuições *t* de Student assimétricas, obtidas pelo método *Data-driven Reversible jump*, com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 500$ e $N = 1000$.

	Parâmetros	Médias (Intervalos de Credibilidade)			
		$N = 500$		$N = 1000$	
$K = 1$	$p_1 = 1.0$	1.00	(1.00; 1.00)	1.00	(1.00; 1.00)
	$\xi_1 = -1.0$	-0.97	(-1.27; -0.58)	-0.74	(-0.99; -0.44)
	$\omega_1 = 1.5$	1.36	(1.14; 1.59)	1.32	(1.19; 1.49)
	$\alpha_1 = 1.0$	0.89	(0.39; 1.36)	0.68	(0.34; 1.02)
	$v_1 = 4.0$	3.55	(2.45; 4.75)	3.26	(2.73; 4.13)
$K = 2$	$p_1 = 0.3$	0.29	(0.24; 0.33)	0.28	(0.26; 0.31)
	$\xi_1 = -1.0$	-0.79	(-1.53; 0.59)	-0.42	(-1.01; 0.35)
	$\omega_1 = 1.5$	1.35	(1.00; 1.72)	1.15	(0.98; 1.38)
	$\alpha_1 = 1.0$	0.60	(-1.30; 1.78)	0.20	(-0.84; 0.99)
	$v_1 = 4.0$	8.48	(4.90; 15.79)	3.95	(3.16; 4.95)
	$p_2 = 0.7$	0.71	(0.67; 0.76)	0.72	(0.69; 0.75)
	$\xi_2 = 5.0$	4.87	(4.63; 5.08)	5.03	(4.86; 5.20)
	$\omega_2 = 0.5$	0.46	(0.40; 0.55)	0.50	(0.44; 0.60)
	$\alpha_2 = -0.5$	-0.15	(-0.72; 0.47)	-0.59	(-1.10; -0.07)
	$v_2 = 4.0$	2.84	(2.21; 3.80)	3.78	(2.55; 5.99)
$K = 3$	$p_1 = 0.5$	0.51	(0.47; 0.56)	0.50	(0.46; 0.53)
	$\xi_1 = -1.0$	-1.16	(-1.50; -0.75)	-0.96	(-1.35; -0.42)
	$\omega_1 = 1.5$	1.46	(1.09; 1.91)	1.33	(1.09; 1.67)
	$\alpha_1 = 1.0$	1.39	(0.63; 2.20)	0.88	(0.15; 1.52)
	$v_1 = 4.0$	2.95	(1.72; 5.61)	3.24	(2.26; 4.33)
	$p_2 = 0.2$	0.20	(0.16; 0.24)	0.20	(0.17; 0.23)
	$\xi_2 = 5.0$	4.56	(4.30; 5.15)	5.03	(4.54; 5.40)
	$\omega_2 = 0.5$	0.60	(0.42; 0.84)	0.64	(0.44; 0.91)
	$\alpha_2 = -0.5$	1.84	(-0.86; 4.61)	-0.65	(-2.13; 0.84)
	$v_2 = 4.0$	4.30	(2.50; 7.12)	7.14	(5.01; 10.10)
	$p_3 = 0.3$	0.29	(0.25; 0.33)	0.30	(0.27; 0.36)
	$\xi_3 = 10.0$	9.67	(9.18; 10.21)	10.78	(9.88; 11.88)
	$\omega_3 = 2.0$	2.40	(1.85; 3.09)	1.74	(1.39; 2.07)
$\alpha_3 = 1.0$	2.49	(1.04; 4.82)	0.37	(-0.54; 1.29)	
$v_3 = 4.0$	4.83	(2.61; 8.38)	3.12	(2.27; 4.44)	
$K = 4$	$p_1 = 0.3$	0.30	(0.26; 0.33)	0.31	(0.27; 0.34)
	$\xi_1 = -1.0$	-1.05	(-1.39; -1.09)	-1.18	(-1.47; -0.83)
	$\omega_1 = 1.5$	1.30	(1.04; 1.63)	1.47	(1.17; 1.77)
	$\alpha_1 = 1.0$	1.14	(0.33; 1.98)	1.37	(0.69; 2.08)
	$v_1 = 4.0$	3.12	(1.99; 4.22)	3.63	(2.81; 4.69)
	$p_2 = 0.1$	0.10	(0.07; 0.12)	0.09	(0.07; 0.11)
	$\xi_2 = 5.0$	5.04	(4.64; 5.33)	5.00	(4.64; 5.33)
	$\omega_2 = 0.5$	0.48	(0.32; 0.74)	0.49	(0.34; 0.74)
	$\alpha_2 = -0.5$	-0.90	(-2.86; 0.54)	-0.63	(-3.05; 0.53)
	$v_2 = 4.0$	1.99	(1.16; 3.05)	6.75	(4.17; 8.75)
	$p_3 = 0.4$	0.39	(0.035; 0.43)	0.41	(0.38; 0.45)
	$\xi_3 = 10.0$	10.42	(9.58; 11.33)	10.55	(9.80; 11.36)
	$\omega_3 = 2.0$	1.68	(1.31; 2.26)	1.68	(1.40; 2.17)
	$\alpha_3 = 1.0$	0.50	(-0.28; 1.43)	0.40	(-0.30; 1.23)
	$v_3 = 4.0$	2.56	(1.44; 3.80)	2.24	(1.72; 2.89)
	$p_4 = 0.2$	0.21	(0.18; 0.25)	0.20	(0.17; 0.23)
	$\xi_4 = 19.0$	19.20	(18.67; 19.61)	19.13	(18.35; 19.67)
$\omega_4 = 1.7$	1.73	(1.23; 2.23)	1.66	(1.09; 2.43)	
$\alpha_4 = -1.5$	-2.16	(-3.97; -0.84)	-1.93	(-4.26; -0.31)	
$v_4 = 4.0$	2.54	(1.59; 3.48)	5.32	(3.77; 6.80)	

Verificamos por meio da [Tabela 7](#) que a maioria das médias *a posteriori* dos parâmetros dos modelos de mistura de distribuições *t* de Student assimétricas com $K = 1, 2, 3, 4$ componentes, estão bem próximas do verdadeiro valor dos parâmetros, e que a maioria dos intervalos de 95% credibilidade contém os verdadeiros valores dos parâmetros, indicando novamente que o método *Data-driven Reversible jump* consegue fornecer estimativas da probabilidade *a posteriori* do número de componentes dos modelos de mistura de distribuições *t* de Student assimétricas e também de estimar os parâmetros dos modelos conjuntamente.

Esboçamos as Figuras [13](#), [14](#), [15](#) e [16](#) com os histogramas e as distribuições empírica dos dados, ambas com as densidades reais (linhas preta) e estimadas (linha verde) dos modelos de mistura de distribuições *t* de Student assimétricas, para os dados gerados, respectivamente, com $K = 1, 2, 3, 4$ componentes, $N = 1000$, e com o número de componentes estimados pelas probabilidades *a posteriori* de K , obtidas pelos métodos *Data-driven Reversible jump*, de acordo com a [Tabela 4](#).

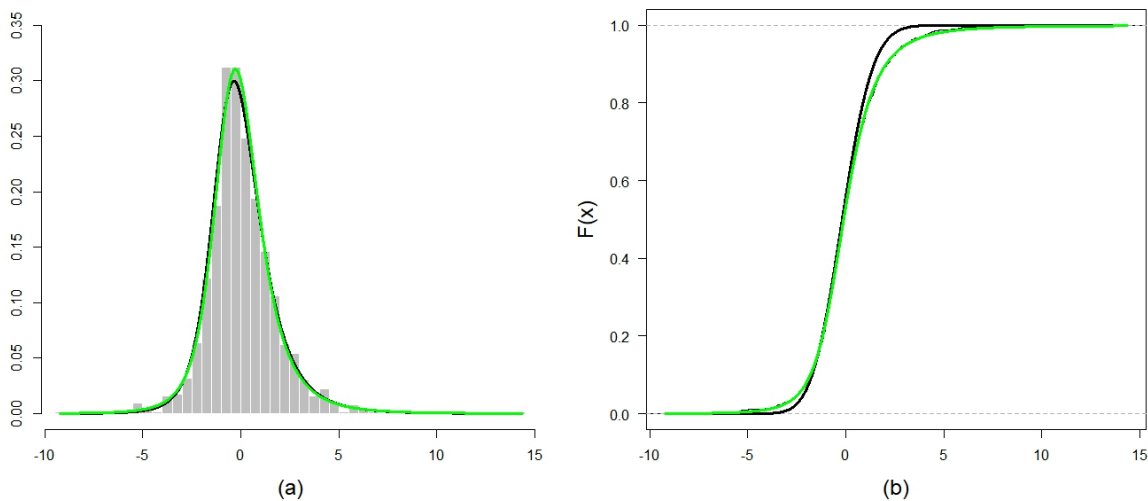


Figura 13 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições *t* de Student assimétricas, com $K = 1$ componente e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

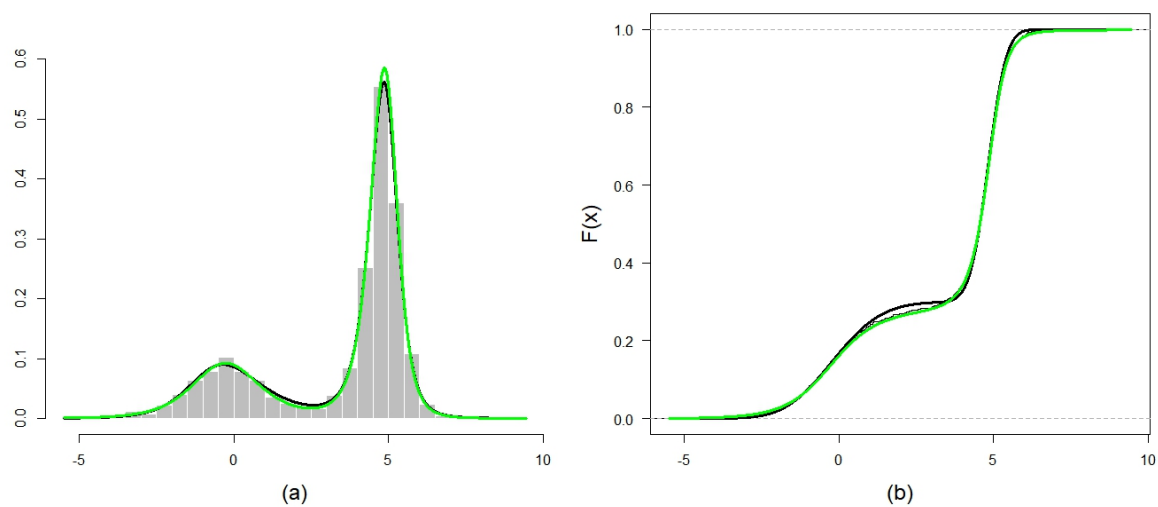


Figura 14 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 2$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

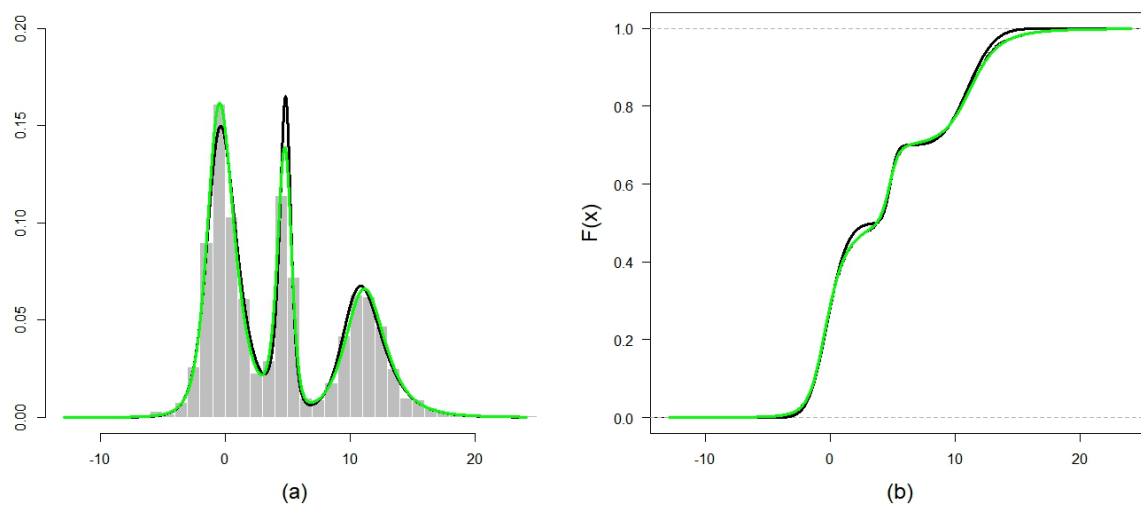


Figura 15 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 3$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

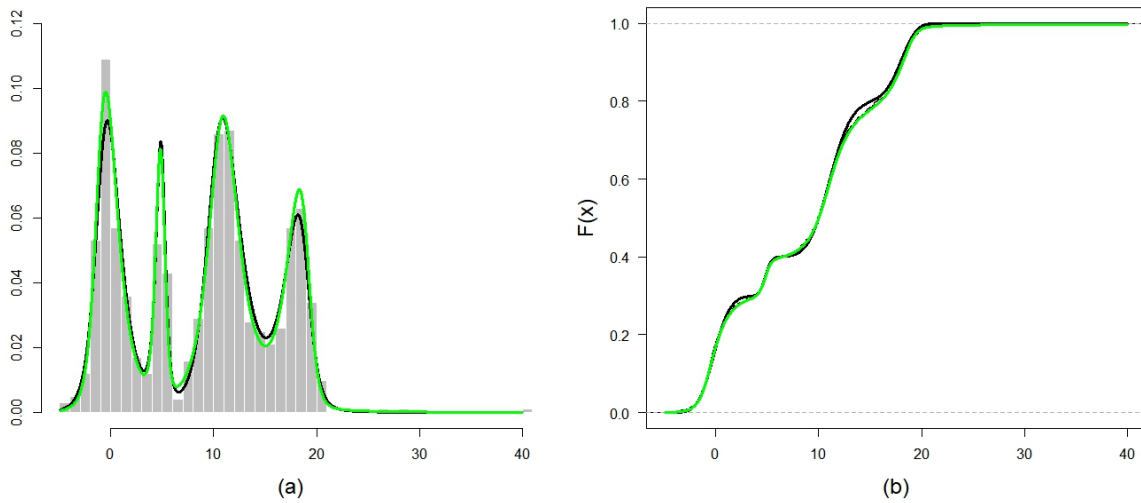


Figura 16 – Densidade real (linha preta) e estimada (linha verde) dos modelos de mistura de distribuições t de Student assimétricas, com $K = 4$ componentes e $N = 1000$. (a) Histograma e (b) Distribuição empírica.

Podemos observar por meio das Figuras 13, 14, 15 e 16 que tanto nos histogramas quanto nas distribuições empíricas, as densidades dos modelos estimados estão bem próximas dos modelos reais, recomendando novamente que o método *Data-driven Reversible jump* proporciona a estimativa da probabilidade *a posteriori* para o número de componentes do modelos de mistura de distribuições t de Student, e conjuntamente estimativas para os parâmetros dos modelos que sugerem um bom ajuste para os dados.

Inicializamos todas as cadeias com $K = 1$ e verificamos resultados semelhantes de convergência das cadeias para todos os cenários do estudo de simulação. Apresentamos na seção F.1 do Apêndice F os gráficos da verificação da convergência das cadeias para os modelos de mistura de distribuições normais para $K = 1, 2, 3, 4$ e tamanho de amostra $N = 1000$, os demais como já dito são similares.

7.2 Aplicação em dados reais

Na aplicação da metodologia proposta neste trabalho, utilizamos diferentes conjuntos de dados reais. O primeiro conjunto de observações refere-se a dados de escores de cognição, ou seja, dados de um experimento com pacientes que possuem a doença de Alzheimer. O segundo composto por dados de índice de massa corporal e o terceiro por dados da gêiser de vulcão, sendo que ambos estão disponíveis no *Software R* (R Core Team, 2022). O quarto e último conjunto de observações são dados financeiros da criptomoeda *Bitcoin*.

7.2.1 DADOS DE ESCORES DE COGNIÇÃO

A palavra cognição significa processo ou faculdade de adquirir um conhecimento. A cognição envolve diversos fatores como o pensamento, a linguagem, a percepção, a memória e o raciocínio, e faz parte do desenvolvimento intelectual. Em geral, a atividade intelectual está ligada ao funcionamento do próprio organismo, ao desenvolvimento biológico e ao envelhecimento de cada pessoa.

Com o envelhecimento, vem a perda das habilidades cognitivas que é uma característica da doença de Alzheimer. A doença de Alzheimer também pode estar relacionada a outros fatores de risco, como os genéticos e ambientais (FRÜHWIRTH-SCHNATTER; PYNE, 2010).

Para este trabalho, utilizamos dados da doença Alzheimer provenientes de um experimento com 451 indivíduos de um estudo de coorte das Ordens Religiosas (WILSON *et al.*, 2004) e do Projeto Memória e Envelhecimento (BENNETT *et al.*, 2005).

Nesse experimento, o nível de cognição dos indivíduos foi clinicamente calculado próximo à sua morte, com base em testes de funções cognitivas e resumidos por um escore médio global de cognição, em que escores mais altos sugerem melhores capacidades de cognição.

Esse conjunto de dados também foi analisado em Frühwirth-Schnatter e Pyne (2010), sendo que foi a própria Senhora Frühwirth-Schnatter que forneceu esse conjunto de observações para nossa pesquisa.

Como a distribuição dos dados de escores globais de cognição sugere assimetria, aplicamos modelos de mistura com distribuições normais assimétricas e t de Student assimétricas, e também o modelo de mistura de distribuições normais, com $K = 1, \dots, 10$ componentes, por meio da abordagem bayesiana apresentada no Capítulo 4. Seleccionamos o número de componentes de mistura e estimamos os parâmetros dos modelos, conforme descrito na Seção 7.1 desse mesmo capítulo.

Apresentamos a Tabela 8 com as estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* e *Reversible jump* para os, respectivos, modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados dos escore globais.

Tabela 8 – Estimativas das probabilidades *a posteriori* do número de componentes K , obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados dos escores globais.

Modelos de mistura	K	DDRJ	RJ
Normal	1	0.001	-
	2	0.999	0.563
	3	-	0.368
	≥ 4	-	0.069
Normal assimétrica	1	0.001	-
	2	0.999	0.639
	3	-	0.313
	≥ 4	-	0.048
t de Student assimétrica	1	0.001	-
	2	0.999	0.433
	3	-	0.375
	≥ 4	-	0.192

Observamos por meio da [Tabela 8](#) que o método *Data-driven Reversible jump* indica o modelo de mistura com $K = 2$ componentes, com estimativas das probabilidades *a posteriori* de 0.999, para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, e o método *Reversible jump* também indica o modelo de mistura com $K = 2$ componentes, com estimativas das probabilidades *a posteriori* de 0.563, 0.639 e 0.433, respectivamente, para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas. O método *Data-driven Reversible jump* fornece estimativas da probabilidades *a posteriori* maiores e também se move num número menor de modelos, quando comparado com o método *Reversible jump*.

Descrevemos na [Tabela 9](#) as médias *a posteriori*, as medianas *a posteriori* e os intervalos de 95% de credibilidade para os parâmetros dos modelos de mistura com distribuições normais, normais assimétricas e t de Student assimétricas, com o número de componentes $K = 2$, obtidas pelo método *Data-driven Reversible jump*, conforme descrito na [Tabela 8](#), para os dados de escores globais.

Tabela 9 – Médias *a posteriori*, Medianas *a posteriori* e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas com $K = 2$ componentes, obtidas pelo método *Data-driven Reversible jump*, para os dados de escores globais.

Modelos de mistura	Parâmetros	Médias	Medianas	I.C.
Normal	p_1	0.32	0.32	(0.20; 0.48)
	μ_1	-2.35	-2.36	(-2.89; -1.71)
	σ_1	1.04	1.04	(0.70; 1.34)
	p_2	0.68	0.68	(0.52; 0.80)
	μ_2	-0.30	-0.30	(-0.44; -0.17)
	σ_2	0.61	0.61	(0.51; 0.71)
Normal assimétrica	p_1	0.24	0.24	(0.18; 0.34)
	ξ_1	-2.41	-2.34	(-3.34; -1,83)
	ω_1	1.00	1.00	(0.82; 1.18)
	α_1	-0.36	-0.41	(-1.34; 0.81)
	p_2	0.76	0.76	(0.66; 0.83)
	ξ_2	0.26	0.28	(0.04; 0.41)
	ω_2	0.96	0.97	(0.74; 1.12)
	α_2	-1.82	-1.8	(-2.64; -0.88)
<i>t</i> de Student assimétrica	p_1	0.10	0.09	(0.05; 0.18)
	ξ_1	-4.10	-4.09	(-4.27; -3.98)
	ω_1	0.82	0.74	(0.36; 1.52)
	α_1	10.18	10.27	(3.43; 17.27)
	ν_1	3.86	3.97	(2.02; 6.29)
	p_2	0.90	0.91	(0.82; 0.95)
	ξ_2	0.44	0.45	(0.32; 0.55)
	ω_2	1.41	1.40	(1.21; 1.59)
	α_2	-3.98	-3.87	(-5.55; -2.76)
	ν_2	16.40	16.47	(7.11; 28.28)

Observamos por meio dos intervalos de 95% de credibilidade apresentados na [Tabela 9](#), que o parâmetro de assimetria α_1 do modelo de mistura de distribuições normais assimétricas não é significativo, pois o valor “0” (zero) pertence ao intervalo de credibilidade.

Exibimos a [Figura 17](#) com o histograma (a) e a distribuição empírica (b) dos dados, ambos com as densidades estimadas dos modelos de mistura de distribuições normais (linha azul), normais assimétricas (linha vermelha) e *t* de Student assimétricas (linha verde), para os dados dos escores globais.

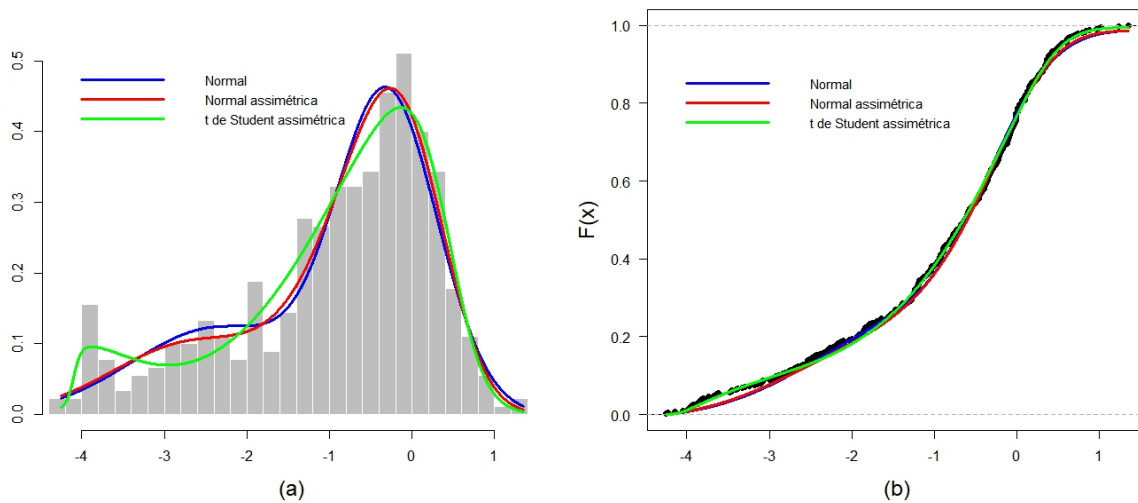


Figura 17 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados de escores globais.

Podemos observar por meio da [Figura 17](#) que os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, com $K = 2$ componentes, sugerem um bom ajuste para os dados dos escores globais. Podemos observar que o método *Data-driven Reversible jump* consegue fornecer além da estimativa da probabilidade *a posteriori* para o número de componentes do modelo, também as estimativas dos parâmetros que apresentam um bom ajuste para os dados.

7.2.2 DADOS DE ÍNDICE DE MASSA CORPORAL E DADOS DE ERUPÇÃO DE VULCÃO

Aplicamos também a metodologia desenvolvida neste trabalho em dois conjuntos de dados disponibilizado na *library mixsmsn* do *Software R* ([R Core Team, 2022](#)), sendo o primeiro, um conjunto dados de índice de massa corporal e o segundo, um conjunto de dados com duas variáveis sobre erupção de vulcão.

Para ambos os conjuntos de dados aplicamos os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, com $K = 1, \dots, 10$ componentes, por meio da abordagem bayesiana apresentada no [Capítulo 4](#). Também selecionamos o número de componentes de mistura e estimamos os parâmetros dos modelos, conforme descrito na [Seção 7.1](#) desse mesmo capítulo.

Devido a preocupação com os problemas de saúde e a forte relação da obesidade com muitas doenças crônicas o National Center Health Statistics (NCHS) pertencente ao Center for Disease Control (CDC) dos Estados Unidos da América (EUA) realizou uma pesquisa nacional sobre saúde e nutrição. Nessa pesquisa, dados do Índice de Massa Corporal (IMC) que

é calculado pela divisão do peso (em kg) do indivíduo dividido pela sua altura (em metros) ao quadrado, de um total de 4.579 participantes foram obtidos, uma vez que o IMC é uma medida padrão para sobrepeso e a obesidade (BASSO *et al.*, 2010).

Uma amostra desse conjunto de dados de IMC também foi analisado por Lin *et al.* (2007a). Porém, para explorar o padrão de mistura, os autores consideraram apenas os participantes com peso entre (39,50 kg a 70,00 kg) e (95,01 a 196,80kg) como dois subgrupos, sendo 1.069 participantes no primeiro subgrupo e 1.054 no segundo subgrupo.

Já a amostra dessas observações de medida do IMC disponibilizadas no *Software R* (R Core Team, 2022), a mesma analisada nesse trabalho, consiste em um total de 2.107 indivíduos do sexo masculino com idade entre 18 e 80 anos. Esse conjunto de dados também foi analisado por Basso *et al.* (2010).

Apresentamos a Tabela 10 com estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* e *Reversible jump* para os, respectivos, modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados do IMC.

Tabela 10 – Estimativas das probabilidades *a posteriori* do número de componentes K , obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados do IMC.

Modelos de mistura	K	DDRJ	RJ
Normal	1	0.001	-
	2	0.999	0.950
	3	-	0.035
	≥ 4	-	0.015
Normal assimétrica	1	0.001	-
	2	0.999	0.862
	3	-	0.094
	≥ 4	-	0.044
t de Student assimétrica	1	0.001	-
	2	0.999	0.974
	3	-	0.018
	≥ 4	-	0.008

Observamos por meio da Tabela 10 que o método *Data-driven Reversible jump* indica o modelo de mistura com $K = 2$ componentes, com estimativas das probabilidades *a posteriori* de 0.999 para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, e o método *Reversible jump* indica também o modelo de mistura com $K = 2$ componentes, com estimativa das probabilidades *a posteriori* de 0.950, 0.862 e 0.974, respectivamente, para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas. O método *Data-driven Reversible jump* fornece estimativas das probabilidades *a posteriori* maiores e também se move num número menor de modelos.

Descrevemos na [Tabela 11](#) as médias *a posteriori*, as medianas *a posteriori* e os intervalos de 95% credibilidade para os parâmetros dos modelos de mistura com distribuições normais, normais assimétricas e *t* de Student assimétricas, com o número de componentes $K = 2$, indicadas pelos métodos *Data-driven Reversible jump* e *Reversible jump*, conforme descrito na [Tabela 10](#), para os dados do IMC.

Tabela 11 – Médias *a posteriori*, Medianas *a posteriori* e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas com $K = 2$ componentes, obtidas pelo método *Data-driven Reversible jump*, para os dados do IMC.

Modelos de mistura	Parâmetros	Médias	Medianas	I.C.
Normal	p_1	0.29	0.29	(0.26; 0.33)
	μ_1	21.10	21.10	(20.90; 21.30)
	σ_1	1.63	1.63	(1.48; 1.78)
	p_2	0.71	0.71	(0.67; 0.74)
	μ_2	30.20	30.20	(29.6; 30.70)
	σ_2	6.85	6.85	(65.80; 7.19)
Normal assimétrica	p_1	0.53	0.53	(0.52; 0.54)
	ξ_1	19.80	19.62	(19.45; 19.90)
	ω_1	3.60	3.66	(3.37; 3.84)
	α_1	1.65	1.77	(1.42; 2.00)
	p_2	0.47	0.48	(0.47; 0.48)
	ξ_2	28.70	28.72	(28.61; 28.82)
	ω_2	7.98	7.99	(7.84; 8.10)
	α_2	10.3	9.23	(8.05; 11.40)
<i>t</i> de Student assimétrica	p_1	0.50	0.49	(0.47; 0.54)
	ξ_1	19.80	19.80	(19.10; 20.30)
	ω_1	2.99	2.94	(2.58; 3.95)
	α_1	1.40	1.36	(0.93; 2.44)
	ν_1	20.10	19.40	(14.20; 32.70)
	p_2	0.50	0.51	(0.46; 0.53)
	ξ_2	29.50	29.50	(28.70; 29.90)
	ω_2	4.34	4.22	(3.83; 6.21)
	α_2	2.58	1.90	(1.42; 10.80)
	ν_2	1.93	1.84	(1.63; 3.63)

Observamos por meio dos intervalos de 95% de credibilidade apresentados na [Tabela 11](#), que os parâmetros de assimetria α_1 e α_2 dos modelos de mistura de distribuições normais assimétricas e *t* de Student assimétricas são significativos, pois o valor “0” (zero) não pertence aos intervalos de credibilidade.

Exibimos a [Figura 18](#) com o histograma (a) e a distribuição empírica (b) dos dados, ambos com as densidades estimadas dos modelos de mistura de distribuições normais (linha azul), normais assimétricas (linha vermelha) e *t* de Student assimétricas (linha verde) para os dados do IMC.

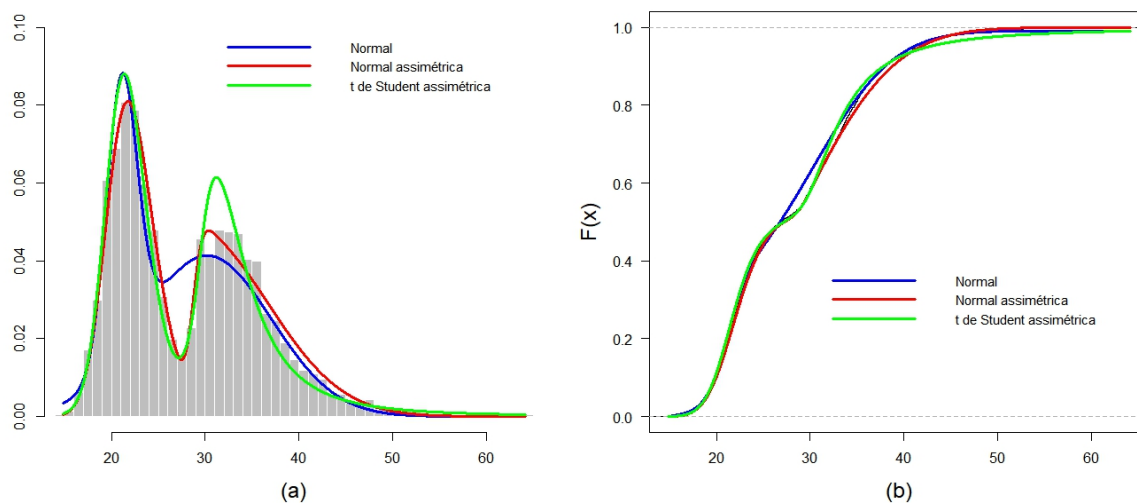


Figura 18 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do IMC.

Podemos observar por meio da [Figura 18](#) que os modelos de mistura de distribuições normais assimétricas e t de Student assimétricas, com $K = 2$ componentes, sugerem um bom ajuste para os dados dos IMC. Mais uma vez observamos que o método *Data-driven Reversible jump* além de fornecer as estimativas da probabilidade *a posteriori* para o número de componentes do modelo, também fornece conjuntamente as estimativas dos parâmetros que indicam um bom ajuste para os dados.

O segundo conjunto de dados disponibilizado no *Software R* ([R Core Team, 2022](#)) consiste em 272 observações das variáveis tempo de espera entre as erupções e a duração da erupção do gêiser *Old Faithful*, que fica no Parque Nacional de Yellowstone, Wyoming, EUA. Esse conjunto de dados também foi analisado por [Lin et al. \(2007b\)](#).

Uma gêiser é uma nascente termal que entra em erupção periodicamente, lançando uma coluna de água à vapor. A formação de gêiseres requer uma hidrogeologia favorável pois são fenômenos, em geral, raros. Existem cerca de mil em todo o mundo, e a metade destes ficam no Parque Nacional de Yellowstone, entre os quais um dos mais conhecidos é o *Old Faithful* (ou “Velho Fiel”) ([AZZALINI; BOWMAN, 1990](#)).

Apresentamos a [Tabela 12](#) com as estimativas das probabilidades *a posteriori* de K obtidas pelos métodos *Data-driven Reversible jump* e *Reversible jump* para os, respectivos, modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados de erupção do vulcão.

Tabela 12 – Estimativas das probabilidades *a posteriori* do número de componentes K , obtidas pelos métodos *Data-driven Reversible jump* (DDRJ) e *Reversible jump* (RJ) para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados de erupção do vulcão.

Modelos de mistura	K	Duração das erupções		Espera entre as erupções	
		DDRJ	RJ	DDRJ	RJ
Normal	1	-	-	-	0.003
	2	0.975	0.538	0.999	0.605
	3	0.025	0.431	0.001	0.279
	≥ 4	-	0.031	-	0.113
Normal assimétrica	1	-	-	-	0.002
	2	0.999	0.520	0.999	0.581
	3	0.001	0.450	0.001	0.320
	≥ 4	-	0.030	-	0.097
t de Student assimétrica	1	-	-	-	0.002
	2	0.999	0.504	0.999	0.581
	3	0.001	0.460	0.001	0.283
	≥ 4	-	0.036	-	0.134

Observamos por meio da [Tabela 12](#) que o método *Data-driven Reversible jump* indica o modelo de mistura com $K = 2$ componentes, com estimativa das probabilidades *a posteriori* de 0.975 para o modelo de mistura de distribuições normais, e de 0.999 para os modelos de mistura de distribuições normais assimétricas e t de Student assimétricas, e o método *Reversible jump* indica também o modelo de mistura com $K = 2$ componentes, com estimativas das probabilidades *a posteriori* de 0.538, 0.520 e 0.504, respectivamente, para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados do tempo de duração das erupções.

Por meio da [Tabela 12](#) também observamos que o método *Data-driven Reversible jump* indica o modelo de mistura com $K = 2$ componentes, com estimativa das probabilidades *a posteriori* de 0.999 para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, e o método *Reversible jump* indica o modelo de mistura com $K = 2$ componentes, com estimativas das probabilidades *a posteriori* de 0.605, 0.581 e 0.581, respectivamente, para os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas para os dados do tempo de espera entre as erupções

Descrevemos nas [Tabela 13](#) as médias *a posteriori*, as medianas *a posteriori* e os intervalos de 95% de credibilidade para os parâmetros dos modelos de mistura com distribuições normais, normais assimétricas e t de Student assimétricas, com o número de componentes $K = 2$, para os dados do tempo de duração das erupções do vulcão, indicadas pelos métodos *Data-driven Reversible jump* e *Reversible jump*, conforme descrito na [Tabela 12](#).

Tabela 13 – Médias *a posteriori*, Medianas *a posteriori* e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas com $K = 2$ componentes, obtidas pelo método *Data-driven Reversible jump*, para os dados do tempo de duração das erupções do vulcão.

Modelos de mistura	Parâmetros	Médias	Medianas	I.C.
Normal	p_1	0.35	0.35	(0.30; 0.41)
	μ_1	2.02	2.02	(1.97; 2.08)
	σ_1	0.24	0.24	(0.20; 0.29)
	p_2	0.65	0.65	(0.59; 0.70)
	μ_2	4.28	4.28	(4.21; 4.34)
	σ_2	0.44	0.44	(0.38; 0.49)
Normal assimétrica	p_1	0.35	0.36	(0.30; 0.41)
	ξ_1	1.71	1.71	(1.67; 1.75)
	ω_1	0.43	0.44	(0.33; 0.51)
	α_1	7.50	7.74	(4.68; 10.00)
	p_2	0.65	0.64	(0.59; 0.70)
	ξ_2	4.73	4.75	(4.21; 4.85)
	ω_2	0.62	0.62	(0.49; 0.73)
	α_2	-2.63	-2.67	(-4.65; 0.33)
<i>t</i> de Student assimétrica	p_1	0.35	0.36	(0.28; 0.41)
	ξ_1	1.74	1.74	(1.71; 1.78)
	ω_1	0.29	0.29	(0.23; 0.36)
	α_1	6.28	5.62	(3.74; 10.46)
	ν_1	2.28	2.27	(1.84; 2.76)
	p_2	0.65	0.64	(0.60; 0.72)
	ξ_2	4.71	4.71	(4.60; 4.81)
	ω_2	0.51	0.51	(0.43; 0.63)
	α_2	-2.23	-2.12	(-3.94; -1.21)
	ν_2	5.01	4.79	(3.07; 7.85)

Observamos por meio dos intervalos de 95% de credibilidade apresentados na [Tabela 13](#), que o parâmetro de assimetria α_2 do modelo de mistura de distribuições normais assimétricas não é significativo, pois o valor “0” (zero) pertence ao intervalo de credibilidade.

Exibimos a [Figura 19](#) com o histograma (a) e a distribuição empírica (b) dos dados, ambos com as densidades estimadas dos modelos de mistura de distribuições normais (linha azul), normais assimétricas (linha vermelha) e *t* de Student assimétricas (linha verde) para os dados do tempo de duração das erupções do vulcão.

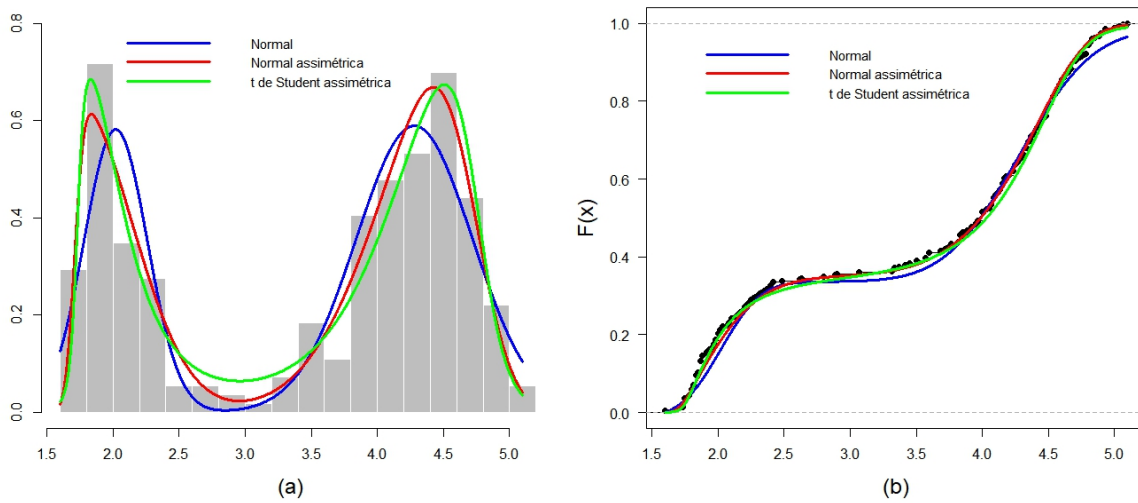


Figura 19 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do tempo de duração das erupções do vulcão.

Podemos observar por meio da [Figura 19](#) que o modelo de mistura de distribuições normais assimétricas e t de Student assimétricas, com $K = 2$ componentes, sugerem um bom ajuste para os dados do tempo de duração das erupções do vulcão. Novamente a estimativa da probabilidade *a posteriori* para o número de componentes do modelo, juntamente com as estimativas dos parâmetros obtidas pelo método *Data-driven Reversible jump* forneceram um bom ajuste para os dados.

Descrevemos nas [Tabela 14](#) as médias *a posteriori*, as medianas *a posteriori* e os intervalos de 95% de credibilidade para os parâmetros dos modelos de mistura com distribuições normais, normais assimétricas e t de Student assimétricas, com o número de componentes $K = 2$, para os dados do tempo de espera entre as erupções do vulcão, indicadas pelos métodos *Data-driven Reversible jump* e *Reversible jump*, conforme descrito na [Tabela 12](#).

Tabela 14 – Médias *a posteriori*, Medianas *a posteriori* e Intervalos de 95% de Credibilidade (I.C.) para os parâmetros dos modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas com $K = 2$ componentes, para os dados do tempo de espera entre as erupções do vulcão.

Modelos de mistura	Parâmetros	Médias	Mediana	I.C.
Normal	p_1	0.36	0.36	(0.30; 0.43)
	μ_1	54.69	54.66	(53.44; 56.11)
	σ_1	6.01	5.94	(5.05; 7.24)
	p_2	0.64	0.64	(0.58; 0.70)
	μ_2	80.13	80.14	(79.08; 81.14)
	σ_2	5.89	5.84	(5.20; 6.73)
Normal assimétrica	p_1	0.40	0.40	(0.31; 0.49)
	ξ_1	46.39	46.07	(44.78; 50.02)
	ω_1	12.91	13.14	(7.32; 16.53)
	α_1	6.00	6.29	(1.28; 10.92)
	p_2	0.60	0.60	(0.52; 0.69)
	ξ_2	80.94	81.39	(74.91; 85.79)
	ω_2	6.41	6.17	(4.98; 8.54)
	α_2	-0.01	-0.20	(-1.54; 2.39)
<i>t</i> de Student assimétrica	p_1	0.38	0.38	(0.32; 0.44)
	ξ_1	53.10	53.30	(51.40; 54.30)
	ω_1	1.77	1.68	(1.30; 2.47)
	α_1	0.28	0.23	(-0.10; 0.78)
	ν_1	0.93	0.91	(0.74; 1.19)
	p_2	0.62	0.62	(0.56; 0.69)
	ξ_2	80.00	80.10	(78.40; 81.50)
	ω_2	4.23	4.21	(3.70; 4.85)
	α_2	0.01	1.09	(-0.31; 0.36)
	ν_2	12.20	11.70	(10.80; 18.60)

Observamos por meio dos intervalos de 95% de credibilidade apresentados na [Tabela 14](#), que o parâmetro de assimetria α_2 do modelo de mistura de distribuições normais assimétricas não é significativo, pois o valor “0” (zero) pertence ao intervalo de credibilidade.

Exibimos a [Figura 20](#) com o histograma (a) e a distribuição empírica (b) dos dados, ambos com as densidades estimadas dos modelos de mistura de distribuições normais (linha azul), normais assimétricas (linha vermelha) e *t* de Student assimétricas (linha verde) para os dados do tempo de espera entre as erupções do vulcão.

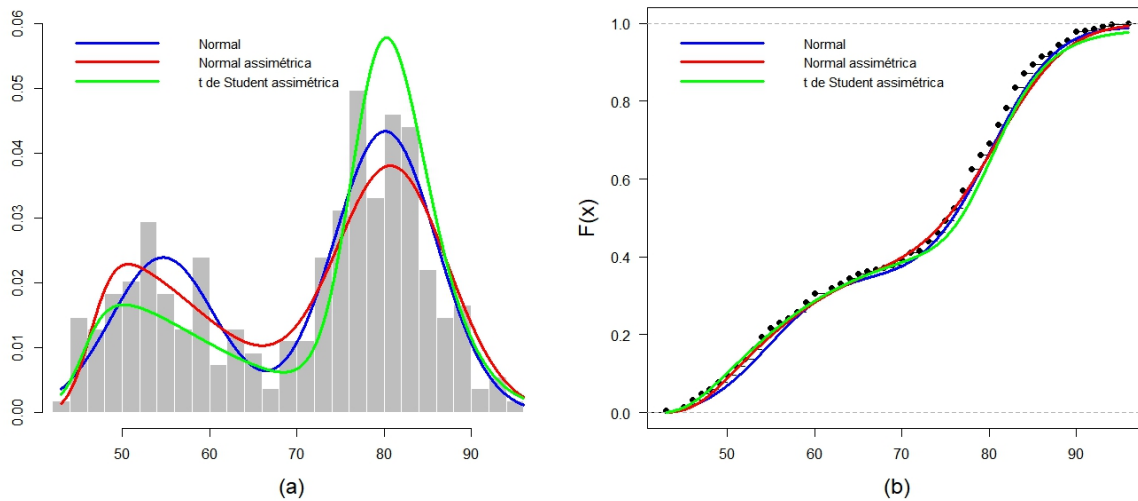


Figura 20 – Densidades estimadas dos modelos de mistura distribuições normais (linha azul), normais assimétricas (linha vermelha) e t de Student assimétricas (linha verde), com $K = 2$ componentes: (a) Histograma e (b) Distribuição empírica, para os dados do tempo de espera entre as erupções do vulcão.

Podemos observar por meio da [Figura 20](#) que os modelos de mistura de distribuições normais, normais assimétricas e t de Student assimétricas, com $K = 2$ componentes, sugerem um bom ajuste para os dados do tempo de espera entre as erupções do vulcão. Esse resultado também indica que o método *Data-driven Reversible jump* consegue fornecer a estimativa da probabilidade *a posteriori* para o número de componentes do modelo, conjuntamente com as estimativas dos parâmetros que apresentam um bom ajuste para os dados.

7.2.3 DADOS DE CRIPTOMOEDAS

As criptomoedas são moedas digitais ou virtuais descentralizadas. Uma moeda descentralizada é uma moeda em que o governo não controla sua oferta, isto é, não pode produzir unidades e nem dar suporte para entidades que possuem ativos medidos por ela.

Sendo assim, o controle descentralizado de cada criptomoeda funciona por meio de tecnologia contábil, geralmente um *blockchain* e uma criptografia de prova. O *blockchain* e a prova criptográfica funcionam como um banco de dados de transações financeiras para validar as transações, criar novas unidades monetárias e garantir a segurança, integridade e equilíbrio dos registros das criptomoedas.

As principais vantagens de usar criptomoedas em relação aos métodos de pagamento tradicionais, como Visa e Mastercard, são a privacidade e os baixos custos de transação ([CHAN et al., 2017](#)). De acordo com [site INVESTING.COM](#), o número atual de criptomoedas no mercado é de 8.713, sendo que a primeira criptomoeda criada, o Bitcoin, é a mais popular, e seu valor, atualmente, está na casa dos 22 mil dólares.

Neste trabalho, utilizamos dados da criptomoeda *Bitcoin*, que foram retirados do *site INVESTING.COM*, <<https://www.br.investing.com/crypto/bitcoin/historical-data>>. Esses dados são referentes a uma cotação histórica diária, no período de 01 de janeiro de 2013 a 31 de janeiro de 2023.

O *site INVESTING.COM* possui informações do preço, em dólares, de abertura, preço máximo e mínimo, e também do último preço das criptomoedas, além disso, oferece a possibilidade de *download* de dados de várias criptomoedas em diversos períodos de tempo. Em nosso trabalho, utilizamos nas análises a variável último preço, da criptomoeda *Bitcoin*, que se refere ao preço da criptomoeda no final do dia.

Sabemos que dados financeiros são correlacionados, desta forma inicialmente calculamos o retorno da série dos dados, ou seja, a razão entre o dia atual e o dia anterior, conforme descrito em [Morettin \(2008\)](#) e [Taylor \(1986\)](#), e depois calculamos o logaritmo do retorno. Após essas transformações obtivemos a estacionariedade das séries na média, [Figura 21\(c\)](#) dados do retorno e [Figura 21\(e\)](#) dados do logaritmo retorno, segundo o teste de Dickey-Fuller (valor $p < 0,001$). O comportamento da função de autocorrelação dos dados de retorno da série, [Figura 21\(d\)](#), e dos dados do logaritmo do retorno da série, [Figura 21\(f\)](#), também indicam a ausência de autocorrelação.

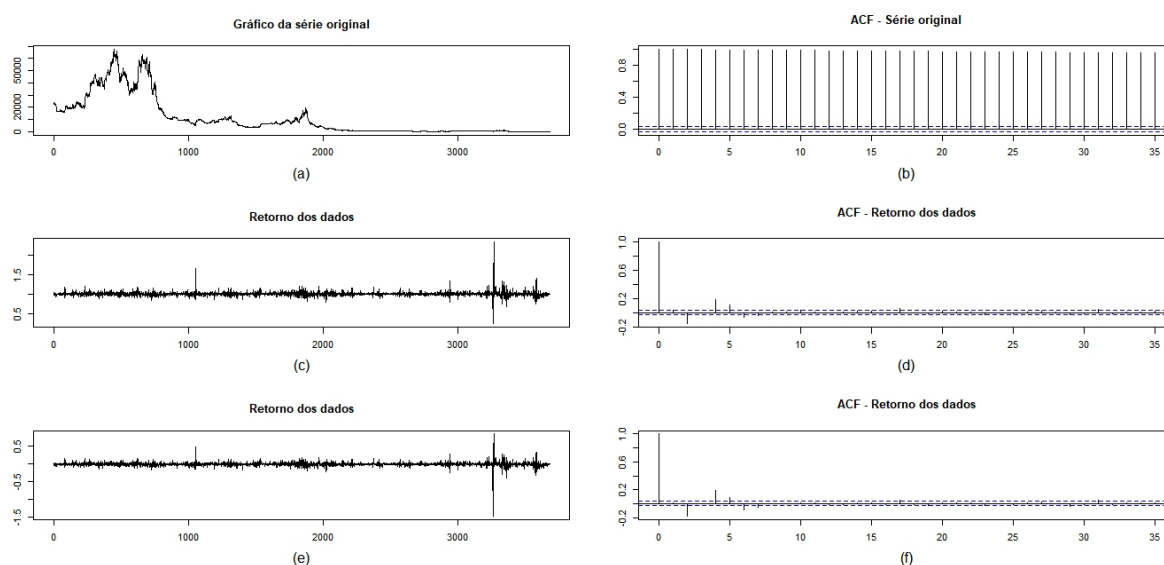


Figura 21 – Dados da criptomoeda *Bitcoin* variável último preço: (a) Gráfico da série original, (b) Gráfico da função de autocorrelação da série original, (c) Gráfico do retorno da série. (d) Gráfico da função de autocorrelação do retorno da série, (e) Gráfico do logaritmo do retorno da série. (f) Gráfico da função de autocorrelação do logaritmo do retorno da série.

Aplicamos os modelos de mistura de distribuições normais, normais assimétricas e *t* de Student assimétricas, com $K = 1, \dots, 10$ componentes, por meio da abordagem bayesiana apresentada no [Capítulo 4](#), e selecionamos o número de componentes de mistura utilizando e estimamos os parâmetros dos modelos de mistura, conforme descrito na [Seção 7.1](#) no início desse capítulo. Porém, a modelagem por meio dos modelos tratados neste estudo não foi adequada,

por isso decidimos retirar esse conjunto de dados deste trabalho.

As cadeias também foram verificadas para todos os conjunto de dados reais analisados, e todas apresentaram resultados similares de convergência. Apresentamos na [Seção F.2 do Apêndice F](#) os gráficos da verificação de convergência das cadeias para os modelos de mistura de distribuições normais para os dados de escores de cognição, os demais como já dito são similares.

CONCLUSÃO

Neste trabalho, propomos um método de estimação bayesiano para modelos de mistura de distribuições assimétricas, mais precisamente propomos o método *Data-driven Reversible jump* para estimar o número de componentes e os parâmetros associados as componentes dos modelos de mistura de distribuições normais assimétricas e *t* de Student assimétricas.

Também verificamos a performance do método *Data-driven Reversible jump* e comparamos o seu desempenho com o método tradicional *Reversible jump*, por meio de um estudo de simulação e análise de dados reais.

No estudo de simulação, identificamos que o método *Data-driven Reversible jump* fornece a estimativa do número de componentes de mistura por meio de altas taxas da estimativa da probabilidade *a posteriori* para o número de componentes, apontando assim o modelo de mistura que melhor se ajusta aos dados. Observamos também que o método *Data-driven Reversible jump* se move entre os modelos mais próximos do verdadeiro modelo, evitando encaixar os modelos que tem pouca ou nenhuma chance de serem selecionados e utilizando pouco esforço computacional.

Observamos também no estudo de simulação que ambos os métodos *Data-driven Reversible jump* e *Reversible jump* apresentaram um bom funcionamento quando o tamanho das amostras eram pequenos, apesar de [McLachlan e Peel \(2000, p.68\)](#) alertar que na análise de modelos de mistura o tamanho da amostra tem que ser muito grande, pois conjuntos de dados pequenos podem apresentar misturas com pesos de componentes pequenos e sobre ajuste de várias componentes.

Na análise de dados reais o método *Data-driven Reversible jump* foi eficaz para estimar o número de componentes, apresentando também maiores taxas da estimativa da probabilidade *a posteriori* para o número de componentes e se movendo entre um número menor de modelos quando comparado com o método tradicional *Reversible jump*. O método *Data-driven Reversible jump* também foi eficaz no processo de estimação dos parâmetro associados as componentes

de mistura dos modelos em estudo, uma vez que os modelos estimados apresentaram um bom ajuste aos dados.

Verificamos também que o algoritmo MCMC *Metropolis-Hastings*, utilizado na atualização das estimativas dos parâmetros do modelo foi eficiente e com pouco esforço computacional, contribuindo assim para que o método *Data-driven Reversible jump* fornecesse as estimativas da probabilidade *a posteriori* para o número de componentes dos modelos de mistura, e também as estimativas pontuais e intervalares para todos os parâmetros associados as componentes de mistura dos modelos em estudo, tanto para os dados simulados quanto para os dados reais.

Com os resultados obtidos nesta pesquisa evidenciamos que o método *Data-driven Reversible jump* quando comparado ao tradicional *Reversible jump* é uma alternativa eficiente para:

1. Identificar o número de componentes de um modelo de mistura de distribuições normais assimétricas e *t* de Student assimétricas;
2. Escolher um melhor ajuste para os modelos de mistura de distribuições normais assimétricas e *t* de Student assimétricas;
3. Fornecer estimativas para os parâmetros associados as componentes de mistura, que apresentem um bom ajuste para os dados, baseado na estimativa da probabilidade *a posteriori* para o número de componentes dos modelos de mistura de distribuições normais assimétricas e *t* de Student assimétricas.

Apresentamos como proposta de trabalhos futuros a aplicação do método *Data-driven Reversible jump* em modelos de regressão com mistura de distribuições normais assimétricas e *t* de Student assimétricas.

REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, p. 716–723, 1974. Citado na página 3.
- ARNOLD, B. C.; GROENEVELD, R. A. Measuring skewness with respect to the mode. **The American Statistician**, v. 49, p. 34–38, 1995. Citado na página 9.
- AZZALINI, A. A class of distributions which includes the normal ones. **Scandinavian Journal of Statistica**, v. 12, p. 171–178, 1985. Citado nas páginas 2, 9 e 10.
- AZZALINI, A. **The Skew-Normal and Related Families**. Institute of Mathematical Statistics Monographs: Cambridge University Press, 2013. Citado nas páginas 12 e 16.
- AZZALINI, A.; BOWMAN, A. W. A look at some data on the old faithful geyser. **Applied Statistics**, v. 39, p. 357–365, 1990. Citado na página 83.
- AZZALINI, A.; CAPITANIO, A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution. **Journal of the Royal Statistical Society**, v. 65, n. B, p. 367–389, 2003. Citado nas páginas 2, 9, 10, 12, 13 e 14.
- BASSO, R. M.; LACHOS, V. H.; CABRAL, C. R. B.; P., G. Robust mixture modeling based on scale mixtures of skew-normal distributions. **Computational Statistics & Data Analysis**, v. 54, n. 12, p. 2926–2941, 2010. Citado na página 81.
- BENNETT, D. A.; SCHNEIDER, J. A.; BUCHMAN A. S. AND DE LEON, C. M.; BIENIAS, J. L.; WILSON, R. S. The rush memory and aging project: study design and baseline characteristics of the study cohort. **Neuroepidemiology**, v. 25, p. 163–175, 2005. Citado na página 77.
- CABRAL, C. R. B.; BOLFARINE, H.; PEREIRA, J. R. G. Bayesian density estimation using skew student-t-normal mixtures. **Computational Statistics & Data Analysis**, v. 52, n. 12, p. 5075–5090, 2008. Citado na página 2.
- CHAN, S.; CHU, J.; NADARAJAH, S.; OSTERRIEDER, J. A statistical analysis of cryptocurrencies. **Journal of Risk and Financial Managemen**, v. 10, n. 2, p. 12, 2017. Citado na página 88.
- CHANDRA, S. On the mixtures of probability distributions. **Scandinavian Journal of Statistics**, v. 4, n. 3, p. 105–112, 1977. Citado na página 24.
- CHARLIER, C. V. L. Uber das fehlergesetz. **Arkiv för Matematik, Astronomi och Fysik**, v. 2, n. 8, p. 9, 1905. Citado nas páginas 25 e 9.
- CHIB, S.; GRRENBERG, E. Understanding the metropolis-hastings algorithm. **The American Statistician**, v. 49, n. 4, p. 327–335, 1995. Citado na página 26.
- CRAWFORD, S. L. An application of the laplace method to finite mixture distributions. **Journal of the American Statistical Association**, v. 89, n. 425, p. 259–267, 1994. Citado na página 24.

- DAHL, D. B. Model-based clustering for expression data via a dirichlet process mixture model. **In M. Vannucci, K.A. Do, & P. Müller, editors, Bayesian Inference for Gene Expression and Proteomics**, Cambridge University Press, 2006. Citado na página 38.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm (with discussion). **Journal of the Royal Statistical Society**, v. 39, n. B, p. 1–38, 1977. Citado nas páginas 2, 3 e 21.
- DIEBOLT, J.; ROBERT, C. P. Estimation of finite mixture distributions through bayesian sampling. **Journal of the Royal Statistical Society**, v. 56, n. B, p. 363–375, 1994. Citado na página 22.
- EDGEWORTH, F. Y. The law of error. **Transactions of the Cambridge Philosophical Society**, v. 20, p. 36–65, 113–113, 1904. Citado nas páginas 25 e 9.
- FERNÁNDEZ, C.; OSIEWALSKI, J.; STEEL, M. F. J. Modeling and inference with ν -spherical distribution. **Journal of the American Statistical Association**, v. 90, p. 1331–1340, 1995. Citado na página 9.
- FRÜHWIRTH-SCHNATTER, S. **Finite Mixture and Markov Switching Models**. New York: Springer Science, 2006. Citado nas páginas 3, 17, 21, 23, 24 e 77.
- FRÜHWIRTH-SCHNATTER, S.; PYNE, S. Bayesian inference for finite mixtures of univariate and multivariate skew-normal and skew- t distributions. **Bioestatiscs**, v. 11, n. 2, p. 317–336, 2010. Citado nas páginas 3 e 77.
- GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. **Bayesian Data Analysis**. A Chapman & Hall Book: CRC Press Taylor & Francis Group, 2013. Citado nas páginas 22, 27 e 28.
- GREEN, P. J. Reversible jump markov chain monte carlo computation and bayesian model determination. **Biometrika**, v. 82, n. 4, p. 711–732, 1995. Citado nas páginas 3, 4, 5, 22 e 28.
- HASTINGS, W. K. Monte carlo sampling methods usind markov chains and their applications. **Biometrika**, v. 57, n. 1, p. 97–109, 1970. Citado nas páginas 3, 4 e 26.
- INVESTING.COM. **Dados de criptomoedas Bitcoin**: Acesso em 06 de fevereiro. [S.l.], 2023. Disponível em: <<https://www.br.investing.com/crypto/bitcoin/historical-data>>. Citado nas páginas 88 e 89.
- JAIN, S.; NEAL, R. M. A split-merge markov chain monte carlo procedure for the dirichlet process mixture model. **Journal of Computational and Graphical Statistics**, v. 13, n. 1, p. 158–182, 2004. Citado na página 4.
- JAIN, S.; NEAL, R. M. Splitting and merging components of a nonconjugate dirichlet process mixture model. **Bayesian Analysis**, v. 2, n. 3, p. 445–472, 2007. Citado na página 4.
- JASRA, A.; HOLMES, C.; STEPHENS, D. A. Markov chain monte carlo methods and the label switching problem in bayesian mixture modeling. **Statistical Science**, v. 20, n. 1, p. 50–67, 2005. Citado na página 38.
- JASRA, A.; STEPHENS, D. A.; GALLAGHER, K.; HOLMES, C. C. Bayesian mixture modelling in geochronology via markov chain monte carlo. **Mathematical Geology**, v. 38, n. 3, p. 269–300, 2006. Citado na página 2.

- LEE, S. X.; MCLACHLAN, G. J. On mixtures of skew normal and skew t -distributions. **Advances in Data Analysis and Classification**, v. 7, n. 3, p. 241–266, 2013. Citado na página 2.
- LEE, S. X.; MCLACHLAN, G. J. Finite mixture of canonical fundamental skew t -distributions. **Statistics and Computing**, v. 26, n. 3, p. 573–589, 2016. Citado na página 2.
- LIN, T. I. Maximum likelihood estimation for multivariate skew normal mixture models. **Journal of Multivariate Analysis**, v. 100, p. 257–265, 2009. Citado nas páginas 2 e 3.
- LIN, T. I.; LEE, J. C. Bayesian analysis of mixture modelling using the multivariate t distribution. **Statistics and Computing**, v. 14, p. 119–130, 2004. Citado na página 2.
- LIN, T. I.; LEE, J. C.; HSIEH, W. J. Robust mixture modeling using the skew t distribution. **Statistics and Computing**, v. 17, p. 81–92, 2007. Citado nas páginas 2 e 81.
- LIN, T. I.; LEE, J. C.; YEN, S. Y. Finite mixture modelling using the skew normal distribution. **Statistica Sinica**, v. 17, p. 909–927, 2007. Citado nas páginas 2 e 83.
- MACERAU, W. M. O. **Comparação das Distribuições α -estável, Normal, t de Student e Laplace Assimétricas**. [S.l.]: Universidade Federal de São Carlos, 2012. Citado na página 7.
- MCLACHLAN, G. J.; BASFORD, K. E. **Mixture Models: Inference and Application to Clustering**. New York: Marcel Dekker, 1988. Citado na página 1.
- MCLACHLAN, G. J.; PEEL, D. **Finite Mixture Models**. New York: Wiley, 2000. Citado nas páginas 1 e 91.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLEB, E. Equations of state calculations by fast computing machines. **J. Chem. Phys.**, v. 21, p. 1087–1092, 1953. Citado nas páginas 2, 3, 4 e 26.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. International Editions: McGraw-Hill, 1974. Citado nas páginas 7, 9, 12 e 15.
- MORETTIN, P. A. **Econometria Financeira**. [S.l.]: Editora Blucher, 2008. Citado na página 89.
- PEARSON, K. Contributions to the mathematical theory of evolution, ii: Skew variation in homogeneous material. **Transactions of the Royal Philosophical Society, A**, p. 257–343, 1895. Citado na página 7.
- PEEL, D.; MCLACHLAN, G. J. Robust mixture modeling using the t -distribution. **Statistics and Computing**, v. 10, p. 339–348, 2000. Citado na página 2.
- R Core Team. **R: A Language and Environment for Statistical Computing**. São Paulo, Brasil, 2022. Disponível em: <<http://www.R-project.org/>>. Citado nas páginas 6, 11, 14, 61, 76, 80, 81, 83 e 133.
- REDNER, R. A.; WALKER, H. F. Mixture densities, maximum likelihood and the em algorithm. **SIAM Review**, v. 26, n. 2, p. 195–239, 1984. Citado na página 24.
- RICHARDSON, S.; GREEN, P. J. On bayesian analysis of mixtures with an unknown number of components. **Journal of the Royal Statistical Society**, v. 59, n. B, p. 731–792, 1997. Citado nas páginas 4, 22, 30, 31, 34, 42, 43, 52 e 53.

SARAIVA, E. F. **Modelo de mistura com número de componentes desconhecidos: Estimação via Método Split-Merge**. Tese (Doutorado) — Universidade Federal de São Carlos, UFSCar, 2009. Citado nas páginas 5, 17, 18, 19, 22 e 34.

SARAIVA, E. F.; MILAN, L. A. Clustering gene expression data using a posterior split-merge-birth procedure. **Scandinavia Journal of Statistics**, v. 39, n. 3, p. 399–415, 2012. Citado na página 5.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v. 6, n. 2, p. 461–464, 1978. Citado na página 3.

SHOHAM, S. Robust clustering by deterministic agglomeration em of mixtures of multivariate t -distributions. **Pattern Recognition**, v. 35, n. 5, p. 1127–1142, 2002. Citado na página 2.

SHOHAM, S.; FELLOWS, M. R.; NORMANN, R. A. Robust, automatic spike sorting using mixtures of multivariate t -distributions. **Journal of Neuroscience Methods**, v. 127, n. 2, p. 111–122, 2003. Citado na página 2.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society**, v. 64, p. 583–639, 2002. Citado na página 3.

STEPHENS, M. Dealing with label switching in mixture models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, v. 62, n. 4, p. 795–809, 2000. Citado nas páginas 22 e 38.

TAYLOR, S. **Modelling Financial Time Series**. New York: John Wiley & Sons, 1986. Citado na página 89.

TEICHER, H. Identifiability of finite mixtures. **The Annals of Mathematical Statistics**, v. 34, n. 4, p. 1265–1269, 1963. Citado na página 24.

TITTERINGTON, D. M.; SMITH, A. F. M.; MARKOV, U. E. **Statistical Analysis of Finite Mixture Distributions**. New York: Wiley, 1985. Citado na página 1.

WANG, H. X.; ZHANG, Q. B.; LUO, B.; WEI, S. Robust mixture modelling using multivariate t -distribution with missing information. **Pattern Recognition Letters**, v. 25, n. 6, p. 701–710, 2004. Citado na página 2.

WILSON, R.; BIENIAIS, J.; EVANS, D.; BENNETT, D. The religious orders study: overview and change in cognitive and motor speed. **Aging, Neuropsychology and Cognition**, v. 11, p. 280–303, 2004. Citado na página 77.

YAKOWITZ, S. J.; SPRAGINS, J. D. On the identifiability of finite mixture. **The Annals of Mathematical Statistics**, v. 39, n. 1, p. 209–214, 1968. Citado na página 24.

ZUANETTI, D. A. **Efficient Bayesian methods for mixture models with genetic applications**. Tese (Doutorado) — Universidade Federal de São Carlos, UFSCar, 2016. Citado nas páginas 5, 34, 35, 37 e 38.

ZUANETTI, D. A.; MILAN, L. A. A generalized mixture model applied to diabetes incidence data. **Biometrical Journal**, v. 59, n. 4, p. 826–842, 2017. Citado nas páginas 4 e 5.

JACOBIANO DA TRANSFORMAÇÃO

A.1 Jacobiano da proposta *split* para o modelo mistura de distribuições normais

$$|Jac^{SP}| = \det \begin{vmatrix} u_1 & p_{k^*} & 0 & 0 & 0 & 0 \\ 1-u_1 & -p_{k^*} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\sigma_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}} & \frac{-u_2 \sqrt{\frac{p_{k_2}}{p_{k_1}}}}{2\sigma_{k^*}} & 0 \\ 0 & 0 & 1 & \sigma_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}} & \frac{u_2 \sqrt{\frac{p_{k_1}}{p_{k_2}}}}{2\sigma_{k^*}} & 0 \\ 0 & \frac{u_3(-1+u_2^2)\sigma_{k^*}^2}{u_1^2} & 0 & \frac{-2u_3u_2\sigma_{k^*}^2}{u_1} & \frac{-u_3(-1+u_2^2)}{u_1} & -\frac{(-1+u_2^2)\sigma_{k^*}^2}{u_1} \\ 0 & \frac{(-1+u_3)(-1+u_2^2)\sigma_{k^*}^2}{(-1+u_1^2)} & 0 & \frac{-2(-1+u_3)u_2\sigma_{k^*}^2}{(-1+u_1)} & \frac{-(-1+u_3)(-1+u_2^2)}{(-1+u_1)} & -\frac{(-1+u_2^2)\sigma_{k^*}^2}{(-1+u_1)} \end{vmatrix}$$

$$|Jac^{SP}| = \frac{p_{k^*}(\sigma_{k^*}^2)^2(1-u_2^2) \left[\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} \right]}{u_1(1-u_1)} \quad (\text{A.1})$$

Da Equação (4.11), temos que

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} = \frac{\mu_{k_2} - \mu_{k^*}}{u_2 \sigma_{k^*}} \quad \text{e} \quad \sqrt{\frac{p_{k_2}}{p_{k_1}}} = -\frac{\mu_{k_1} - \mu_{k^*}}{u_2 \sigma_{k^*}},$$

então

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} = \frac{|\mu_{k_1} - \mu_{k_2}|}{u_2 \sigma_{k^*}}. \quad (\text{A.2})$$

Substituindo (A.2) em (A.1), temos

$$|Jac^{SP}| = \frac{p_{k^*} \sigma_{k^*}^2 (1-u_2^2) |\mu_{k_1} - \mu_{k_2}|}{u_1(1-u_1)u_2}. \quad (\text{A.3})$$

De (4.12), temos

$$(1 - u_2^2) = \frac{\sigma_{k_1}^2 p_{k_1}}{u_3 \sigma_{k^*}^2 p_{k^*}}. \quad (\text{A.4})$$

Então,

$$|Jac^{SP}| = \frac{p_{k_1} \sigma_{k_1}^2 |\mu_{k_1} - \mu_{k_2}|}{u_1 (1 - u_1) u_2 u_3}. \quad (\text{A.5})$$

De (4.10), temos

$$u_1 = \frac{p_{k_1}}{p_{k^*}} \text{ e } 1 - u_1 = \frac{p_{k_2}}{p_{k^*}}. \quad (\text{A.6})$$

Logo,

$$|Jac^{SP}| = \frac{p_{k^*}^2 \sigma_{k_1}^2 |\mu_{k_1} - \mu_{k_2}|}{p_{k_2} u_2 u_3}. \quad (\text{A.7})$$

De (4.12), também temos que

$$p_{k_2} = (1 - u_3)(1 - u_2^2) \sigma_{k^*}^2 \frac{p_{k^*}}{\sigma_{k_2}^2}. \quad (\text{A.8})$$

Portanto,

$$|Jac^{SP}| = \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \sigma_{k_1}^2 \sigma_{k_2}^2}{u_2 (1 - u_2^2) u_3 (1 - u_3) \sigma_{k^*}^2}. \quad (\text{A.9})$$

A.2 Jacobiano da proposta *split* para o modelo mistura de distribuições normais assimétricas

$$|Jac^{SP}| = \det \begin{vmatrix} u_1 & p_{k^*} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 - u_1 & -p_{k^*} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\omega_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}} & \frac{-u_2 \sqrt{\frac{p_{k_2}}{p_{k_1}}}}{2\omega_{k^*}} & 0 & 0 & 0 \\ 0 & 0 & 1 & \omega_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}} & \frac{u_2 \sqrt{\frac{p_{k_1}}{p_{k_2}}}}{2\omega_{k^*}} & 0 & 0 & 0 \\ 0 & \frac{u_3(-1+u_2^2)\omega_{k^*}^2}{u_1^2} & 0 & \frac{-2u_3 u_2 \omega_{k^*}^2}{u_1} & \frac{-u_3(-1+u_2^2)}{u_1} & \frac{-(-1+u_2^2)\omega_{k^*}^2}{u_1} & 0 & 0 \\ 0 & \frac{(-1+u_3)(-1+u_2^2)\omega_{k^*}^2}{(-1+u_1^2)} & 0 & \frac{-2(-1+u_3)u_2 \omega_{k^*}^2}{(-1+u_1)} & \frac{-(-1+u_3)(-1+u_2^2)}{(-1+u_1)} & \frac{-(-1+u_2^2)\omega_{k^*}^2}{(-1+u_1)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & u_4 & \alpha_{k^*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 - u_4 & -\alpha_{k^*} \end{vmatrix}$$

$$|Jac^{SP}| = \frac{p_{k^*} \omega_{k^*}^3 (1 - u_2^2) \alpha_{k^*} \left[\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} \right]}{u_1 (1 - u_1)} \quad (\text{A.10})$$

Da Equação (5.12), temos que

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} = \frac{\xi_{k_2} - \xi_{k^*}}{u_2 \omega_{k^*}} \text{ e } \sqrt{\frac{p_{k_2}}{p_{k_1}}} = -\frac{\xi_{k_1} - \xi_{k^*}}{u_2 \omega_{k^*}},$$

então

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} = \frac{|\xi_{k_1} - \xi_{k_2}|}{u_2 \omega_{k^*}}. \quad (\text{A.11})$$

Substituindo (A.11) em (A.10), temos

$$|Jac^{sp}| = \frac{p_{k^*} \omega_{k^*}^2 (1 - u_2^2) \alpha_{k^*} |\mu_{k_1} - \mu_{k_2}|}{u_1 (1 - u_1) u_2}. \quad (\text{A.12})$$

De (5.13), temos

$$(1 - u_2^2) = \frac{\omega_{k_1}^2 p_{k_1}}{u_3 \omega_{k^*}^2 p_{k^*}}. \quad (\text{A.13})$$

Então,

$$|Jac^{sp}| = \frac{p_{k_1} \omega_{k_1}^2 \alpha_{k^*} |\xi_{k_1} - \xi_{k_2}|}{u_1 (1 - u_1) u_2 u_3}. \quad (\text{A.14})$$

De (5.11), temos

$$u_1 = \frac{p_{k_1}}{p_{k^*}} \text{ e } 1 - u_1 = \frac{p_{k_2}}{p_{k^*}}. \quad (\text{A.15})$$

Logo,

$$|Jac^{sp}| = \frac{p_{k^*}^2 \omega_{k_1}^2 \alpha_{k^*} |\mu_{k_1} - \mu_{k_2}|}{p_{k_2} u_2 u_3}. \quad (\text{A.16})$$

De (5.13), também temos que

$$p_{k_2} = (1 - u_3)(1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{\omega_{k_2}^2}. \quad (\text{A.17})$$

Portanto,

$$|Jac^{sp}| = \frac{p_{k^*} \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k^*} |\mu_{k_1} - \mu_{k_2}|}{u_2 (1 - u_2^2) u_3 (1 - u_3) \sigma_{k^*}^2}. \quad (\text{A.18})$$

De (5.14), temos que

$$\alpha_{k^*} = \frac{\alpha_{k_1}}{u_4}. \quad (\text{A.19})$$

Finalmente,

$$|Jac^{sp}| = \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1}}{u_2 (1 - u_2^2) u_3 (1 - u_3) u_4 \omega_{k^*}^2}. \quad (\text{A.20})$$

A.3 Jacobiano da proposta *split* para o modelo de mistura de distribuições *t* de Student assimétricas

$$|Jac^{SP}| = \det \begin{vmatrix} u_1 & p_{k^*} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-u_1 & -p_{k^*} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\omega_{k^*} \sqrt{\frac{p_{k_2}}{p_{k_1}}} & -\frac{u_2 \sqrt{\frac{p_{k_2}}{p_{k_1}}}}{2\omega_{k^*}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_{k^*} \sqrt{\frac{p_{k_1}}{p_{k_2}}} & \frac{u_2 \sqrt{\frac{p_{k_1}}{p_{k_2}}}}{2\omega_{k^*}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{u_3(-1+u_2^2)\omega_{k^*}^2}{(-1+u_3)(-1+u_2^2)\omega_{k^*}^2} & 0 & -\frac{2u_3u_2\omega_{k^*}^2}{u_1} & -\frac{u_3(-1+u_2^2)}{2\omega_{k^*}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{u_3^2}{(-1+u_3)(-1+u_2^2)\omega_{k^*}^2} & 0 & -\frac{2(-1+u_3)u_2\omega_{k^*}^2}{(-1+u_3)} & -\frac{u_1}{(-1+u_3)(-1+u_2^2)} & -\frac{(-1+u_2^2)\omega_{k^*}^2}{(-1+u_3)} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & u_4 & \alpha_{k^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1-u_4 & -\alpha_{k^*} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & u_5 & V_{k^*} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-u_5 & -V_{k^*} \end{vmatrix}$$

$$|Jac^{SP}| = \frac{p_{k^*}^3 \omega_{k^*}^3 (1-u_2^2) \alpha_{k^*} V_{k^*} \left[\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} \right]}{u_1(1-u_1)} \quad (\text{A.21})$$

Da Equação (6.11), temos que

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} = \frac{\xi_{k_2} - \xi_{k^*}}{u_2 \omega_{k^*}} e \sqrt{\frac{p_{k_2}}{p_{k_1}}} = -\frac{\xi_{k_1} - \xi_{k^*}}{u_2 \omega_{k^*}},$$

$$\sqrt{\frac{p_{k_1}}{p_{k_2}}} + \sqrt{\frac{p_{k_2}}{p_{k_1}}} = \frac{|\xi_{k_1} - \xi_{k^*}|}{u_2 \omega_{k^*}}.$$

então

Substituindo (A.22) em (A.21), temos

$$|Jac^{SP}| = \frac{p_{k^*}^2 \omega_{k^*}^2 (1-u_2^2) \alpha_{k^*} V_{k^*} |\mu_{k_1} - \mu_{k_2}|}{u_1(1-u_1)u_2}. \quad (\text{A.23})$$

De (6.12), temos

$$(1 - u_2^2) = \frac{\omega_{k_1}^2 p_{k_1}}{u_3 \omega_{k^*}^2 p_{k^*}}. \quad (\text{A.24})$$

Então,

$$|Jac^{SP}| = \frac{p_{k_1} \omega_{k_1}^2 \alpha_{k^*} v_{k^*} |\xi_{k_1} - \xi_{k_2}|}{u_1 (1 - u_1) u_2 u_3}. \quad (\text{A.25})$$

De (6.10), temos

$$u_1 = \frac{p_{k_1}}{p_{k^*}} \text{ e } 1 - u_1 = \frac{p_{k_2}}{p_{k^*}}. \quad (\text{A.26})$$

Logo,

$$|Jac^{SP}| = \frac{p_{k^*}^2 \omega_{k_1}^2 \alpha_{k^*} v_{k^*} |\mu_{k_1} - \mu_{k_2}|}{p_{k_2} u_2 u_3}. \quad (\text{A.27})$$

De (6.12), também temos que

$$p_{k_2} = (1 - u_3)(1 - u_2^2) \omega_{k^*}^2 \frac{p_{k^*}}{\omega_{k_2}^2}. \quad (\text{A.28})$$

Portanto,

$$|Jac^{SP}| = \frac{p_{k^*} \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k^*} v_{k^*} |\mu_{k_1} - \mu_{k_2}|}{u_2 (1 - u_2^2) u_3 (1 - u_3) \sigma_{k^*}^2}. \quad (\text{A.29})$$

De (6.13) e (6.14), respectivamente, temos que

$$\alpha_{k^*} = \frac{\alpha_{k_1}}{u_4} \text{ e } v_{k^*} = \frac{v_{k_1}}{u_5}, \quad (\text{A.30})$$

Finalmente,

$$|Jac^{SP}| = \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1} v_{k_1}}{u_2 (1 - u_2^2) u_3 (1 - u_3) u_4 u_5 \omega_{k^*}^2}. \quad (\text{A.31})$$

CÁLCULO DAS FUNÇÕES DE VEROSSIMILHANÇA

B.1 Cálculo da função de verossimilhança do modelo de mistura de distribuições normais assimétricas

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \boldsymbol{p}, K | \boldsymbol{y}, \boldsymbol{s}) &= \prod_{k=1}^K \left\{ p_k^{n_k} \left[\prod_{i=1}^N \left(\frac{2}{\sqrt{2\pi}\omega_k^2} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_k}{\omega_k} \right)^2 \right\} \right. \right. \right. \\
&\quad \times \left. \left. \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \right]^{I_{\{s_i\}}(k)} \right\} \\
&= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\sqrt{2\pi}\omega_k^2} \right)^{n_k} \prod_{i=1}^N \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_k}{\omega_k} \right)^2 \right\} \right. \\
&\quad \times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}}(k) \right\} \\
&= \prod_{k=1}^K \left[p_k^{n_k} \left(\frac{2}{\sqrt{2\pi}\omega_k^2} \right)^{n_k} \exp \left\{ - \left[\frac{(S_k^2)(n_k - 1) + n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} \right] \right\} \right. \\
&\quad \times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ -\frac{u^2}{2} \right\} du \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}}(k) \right], \tag{B.1}
\end{aligned}$$

em que

$$\begin{aligned}
\prod_{i=1}^N \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_k}{\omega_k} \right)^2 \right\} &= \exp \left\{ \sum_{i=1}^N \left[-\frac{1}{2} \left(\frac{y_i - \xi_k}{\omega_k} \right)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N (y_i - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N (y_i^2 - 2\xi_k y_i + \xi_k^2) \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N y_i^2 - 2\xi_k \sum_{i=1}^N y_i + \sum_{i=1}^N \xi_k^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N y_i^2 - 2n_k \xi_k \bar{y}_k + n_k \xi_k^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N y_i^2 + n_k (-2\xi_k \bar{y}_k + \xi_k^2) \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N y_i^2 + n_k (-2\xi_k \bar{y}_k + \xi_k^2 + \bar{y}_k^2 - \bar{y}_k^2) \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\sum_{i=1}^N y_i^2 - n_k \bar{y}_k^2 + n_k (\bar{y}_k^2 - 2\xi_k \bar{y}_k + \xi_k^2) \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N y_i^2 - n_k \bar{y}_k^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N y_i^2 - 2n_k \bar{y}_k^2 + n_k \bar{y}_k^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N y_i^2 - 2\bar{y}_k n_k \bar{y}_k + n_k \bar{y}_k^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N y_i^2 - 2\bar{y}_k \sum_{i=1}^N y_i + n_k \bar{y}_k^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N y_i^2 - \sum_{i=1}^N 2\bar{y}_k y_i + \sum_{i=1}^N \bar{y}_k^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N (y_i^2 - 2\bar{y}_k y_i + \bar{y}_k^2) \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[\left(\sum_{i=1}^N (y_i - \bar{y}_k)^2 \right) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[S_k^2 (n_k - 1) + n_k (\bar{y}_k - \xi_k)^2 \right] \right\}. \tag{B.2}
\end{aligned}$$

B.2 Cálculo da função de verossimilhança do modelo de mistura de distribuições t de Student assimétricas

$$\begin{aligned}
 \mathcal{L}(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p}, K | \mathbf{y}, \mathbf{s}) &= \prod_{k=1}^K \left\{ p_k^{n_k} \left[\prod_{i=1}^N \left(\frac{2}{\omega_k} \frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2}) \sqrt{v_k \pi}} \left(1 + \frac{(y_i - \xi_k)^2}{v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \right. \right. \\
 &\times \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2}) \sqrt{(v_k+1)\pi}} \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
 &\times \left. \left. \left. I_{(-\infty, +\infty)}(y_i) \right)^{I_{\{s_i\}}(k)} \right] \right\} \\
 &= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\omega_k} \right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2}) \sqrt{v_k \pi}} \right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
 &\times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2}) \sqrt{(v_k+1)\pi}} \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
 &\times \left. \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}}(k) \right\} \\
 &= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\omega_k} \right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2}) \sqrt{v_k \pi}} \right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
 &\times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k}\right)} \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2}) \sqrt{(v_k+1)\pi}} \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
 &\times \left. \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}}(k) \right\}. \tag{B.3}
 \end{aligned}$$

CÁLCULO DAS DISTRIBUIÇÕES A POSTERIORI

C.1 Cálculo da distribuição *a posteriori* do modelo de mistura de distribuições normais assimétricas

$$\begin{aligned}
\pi(\xi, \omega^2, \alpha, p, K | y, s, \tau, a, b, \mu, \sigma^2) &= \prod_{k=1}^K \left[p_k^{n_k} \left(\frac{2}{\sqrt{2\pi\omega_k^2}} \right)^{n_k} \exp \left\{ - \left[\frac{(S_k^2)(n_k - 1) + n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} \right] \right\} \right. \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ - \frac{t^2}{2} \right\} dt \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \times \frac{1}{K_{max}} I_{\{1, 2, \dots, K_{max}\}}(K) \\
&\times \frac{\Gamma \left(\sum_{k=1}^K \tau_k \right)}{\prod_{k=1}^K \Gamma(\tau_k)} \prod_{k=1}^K p_k^{\tau_k - 1} I_{\{\tau_k\}}(p) \times \frac{1}{\sqrt{2\pi\omega_k^2}} \exp \left\{ - \frac{(\xi_k - \mu_k)^2}{2\omega_k^2} \right\} I_{(-\infty, +\infty)}(\xi_k) \\
&\times \frac{b_k^{a_k}}{\Gamma(a_k)} \left(\frac{1}{\omega_k^2} \right)^{-(a_k+1)} \exp \left\{ - \frac{b_k}{\omega_k^2} \right\} I_{(0, +\infty)}(\omega_k^2) \times \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp \left\{ - \frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} I_{(-\infty, +\infty)}(\alpha_k) \\
&\propto \prod_{k=1}^K \left[p_k^{n_k} \left(\frac{1}{\omega_k^2} \right)^{n_k} \exp \left\{ - \left[\frac{(S_k^2)(n_k - 1) + n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} \right] \right\} \right. \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ - \frac{t^2}{2} \right\} dt \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \\
&\times p_k^{\tau_k - 1} I_{\{\tau_k\}}(p) \left(\frac{1}{\omega_k^2} \right)^{1/2} \exp \left\{ - \frac{(\xi_k - \mu_k)^2}{2\omega_k^2} \right\} I_{(-\infty, +\infty)}(\xi_k) \\
&\times \left(\frac{1}{\omega_k^2} \right)^{-(a_k+1)} \exp \left\{ - \frac{b_k}{\omega_k^2} \right\} I_{(0, +\infty)}(\omega_k^2) \exp \left\{ - \frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} I_{(-\infty, +\infty)}(\alpha_k) \\
&\propto \prod_{k=1}^K \left[p_k^{n_k + \tau_k - 1} \left(\frac{1}{\omega_k^2} \right)^{n_k + 1/2 - (a_k + 1)} \right. \\
&\times \exp \left\{ - \left[\frac{(S_k^2)(n_k - 1) + n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} \right] - \frac{(\xi_k - \mu_k)^2}{2\omega_k^2} - \frac{b_k}{\omega_k^2} - \frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} \\
&\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ - \frac{t^2}{2} \right\} dt \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \right].
\end{aligned}$$

Isto é

$$\begin{aligned}
\pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \boldsymbol{p}, K | \boldsymbol{y}, \boldsymbol{s}, \boldsymbol{\tau}, \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2) &\propto \prod_{k=1}^K \left[p_k^{n_k + \tau_k - 1} \left(\frac{1}{\omega_k^2} \right)^{n_k + 1/2 - (a_k + 1)} \exp \left\{ -\frac{S_k^2(n_k - 1)}{2\omega_k^2} - \frac{b_k}{\omega_k^2} \right\} \right. \\
&\times \exp \left\{ -\frac{n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} - \frac{(\xi_k - \mu_k)^2}{2\omega_k^2} \right\} \exp \left\{ -\frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ -\frac{t^2}{2} \right\} dt \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \left. \right] \\
&\propto \prod_{k=1}^K \left[p_k^{n_k + \tau_k - 1} \left(\frac{1}{\omega_k^2} \right)^{-(a_k - (n_k + 1/2) + 1)} \exp \left\{ -\frac{b_k + \frac{S_k^2(n_k - 1)}{2}}{\omega_k^2} \right\} \right. \\
&\times \exp \left\{ -\frac{\left(\xi_k - \frac{n_k \bar{y}_k + \mu_k}{n_k + 1} \right)^2}{\frac{2\omega_k^2}{n_k + 1}} \right\} \exp \left\{ -\frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} \\
&\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right)} \exp \left\{ -\frac{t^2}{2} \right\} dt \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \right], \tag{C.1}
\end{aligned}$$

em que

$$\begin{aligned}
\exp \left\{ -\frac{n_k(\bar{y}_k - \xi_k)^2}{2\omega_k^2} - \frac{(\xi_k - \mu_k)^2}{2\omega_k^2} \right\} &= \exp \left\{ -\frac{n_k(\bar{y}_k - \xi_k)^2 - (\xi_k - \mu_k)^2}{2\omega_k^2} \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} [n_k(\bar{y}_k - \xi_k)^2 + (\xi_k - \mu_k)^2] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} [n_k(\bar{y}_k^2 - 2\xi_k \bar{y}_k + \xi_k^2) + (\xi_k^2 - 2\xi_k \mu_k + \mu_k^2)] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} [n_k \bar{y}_k^2 - 2n_k \xi_k \bar{y}_k + n_k \xi_k^2 + \xi_k^2 - 2\xi_k \mu_k + \mu_k^2] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} [\xi_k^2(n_k + 1) - 2\xi_k(n_k \bar{y}_k + \mu_k) + \mu_k \bar{y}_k^2 + \mu_k^2] \right\} \\
&= \exp \left\{ -\frac{1}{2\omega_k^2} \left[(n_k + 1) \left(\xi_k^2 - \frac{2\xi_k(n_k \bar{y}_k + \mu_k)}{(n_k + 1)} + \frac{\mu_k \bar{y}_k^2 + \mu_k^2}{(n_k + 1)} \right) \right] \right\} \\
&= \exp \left\{ -\frac{\xi_k^2 - \frac{2\xi_k(n_k \bar{y}_k + \mu_k)}{(n_k + 1)} + \frac{\mu_k \bar{y}_k^2 + \mu_k^2}{(n_k + 1)}}{\frac{2\omega_k^2}{(n_k + 1)}} \right\} \\
&= \exp \left\{ -\frac{\left(\xi_k - \frac{n_k \bar{y}_k + \mu_k}{n_k + 1} \right)^2}{\frac{2\omega_k^2}{(n_k + 1)}} \right\}. \tag{C.2}
\end{aligned}$$

Desta forma, as distribuições condicionais para os parâmetros da k -ésima componente do modelo de mistura de distribuições normais assimétricas são dadas por

$$\begin{aligned}
\xi_k | \omega_k^2, \mu_k, \boldsymbol{y}, \boldsymbol{s} &\sim \phi \left(\frac{n_k \bar{y}_k + \mu_k}{n_k + 1}, \frac{\omega_k^2}{n_k + 1} \right) \Phi_{n_k} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \right), \\
\omega_k^2 | a_k, b_k, \boldsymbol{y}, \boldsymbol{s} &\sim \text{Gamma Inversa} \left(a_k + n_k + \frac{1}{2}, b_k + \frac{S_k^2(n_k - 1)}{2} \right) \Phi_{n_k} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \right), \\
\alpha_k | \mu_k, \sigma_k^2, \boldsymbol{y}, \boldsymbol{s} &\sim \phi \left(\mu_k, \sigma_k^2 \right) \Phi_{n_k} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \right), \\
p_k | \boldsymbol{\tau}_k, \boldsymbol{y}, \boldsymbol{s} &\sim \text{Dirichlet}(n_k + \boldsymbol{\tau}_k),
\end{aligned}$$

em que n_k é o número de observações da componente k , \bar{y}_k é a média amostral da componente k , S_k^2 é a variância amostral da componente k , $\boldsymbol{\tau}_k$ é o hiperparâmetro da distribuição Dirichlet, μ_k e σ_k^2 são os hiperparâmetros da distribuição Normal, a_k e b_k são os hiperparâmetros da distribuição Gama inversa.

C.2 Cálculo da distribuição a posteriori do modelo de mistura de distribuições t de Student assimétricas

$$\begin{aligned}
 \pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, p, K | \mathbf{y}, \mathbf{s}, \boldsymbol{\tau}, \mathbf{a}, \mathbf{b}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{c}, \mathbf{d}) &= \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{2}{\omega_k} \right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2})\sqrt{v_k\pi}} \right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
 &\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma\left(\frac{v_k+2}{2}\right)}{\Gamma\left(\frac{v_k+1}{2}\right)\sqrt{(v_k+1)\pi}} \\
 &\times \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
 &\times \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}^{(k)}} \left. \right\} \\
 &\times \frac{1}{K_{max}} I_{\{1, 2, \dots, K_{max}\}}(K) \\
 &\times \frac{\Gamma\left(\sum_{k=1}^K \tau_k\right)}{\prod_{k=1}^K \Gamma(\tau_k)} \prod_{k=1}^K p_k^{\tau_k-1} I_{\{\tau_k\}}(p) \\
 &\times \frac{1}{\sqrt{2\pi\omega_k^2}} \exp\left\{-\frac{(\xi_k - \mu_k)^2}{2\omega_k^2}\right\} I_{(-\infty, +\infty)}(\xi_k) \\
 &\times \frac{b_k^{a_k}}{\Gamma(a_k)} \left(\frac{1}{\omega_k^2}\right)^{-(a_k+1)} \exp\left\{-\frac{b_k}{\omega_k^2}\right\} I_{(0, +\infty)}(\omega_k^2) \\
 &\times \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2}\right\} I_{(-\infty, +\infty)}(\alpha_k) \\
 &\times \frac{1}{\Gamma(c_k)} \frac{1}{d_k^{c_k}} v_k^{c_k-1} \exp\left\{-\frac{v_k}{d_k}\right\} I_{(0, +\infty)}(v_k) \\
 &\propto \prod_{k=1}^K \left\{ p_k^{n_k} \left(\frac{1}{\omega_k} \right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2})\sqrt{v_k\pi}} \right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
 &\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma\left(\frac{v_k+2}{2}\right)}{\Gamma\left(\frac{v_k+1}{2}\right)\sqrt{(v_k+1)\pi}} \\
 &\times \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
 &\times \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i) I_{\{s_i\}^{(k)}} \left. \right\} \\
 &\times p_k^{\tau_k-1} I_{\{\tau_k\}}(p) \\
 &\times \left(\frac{1}{\omega_k^2}\right)^{1/2} \exp\left\{-\frac{(\xi_k - \mu_k)^2}{2\omega_k^2}\right\} I_{(-\infty, +\infty)}(\xi_k) \\
 &\times \left(\frac{1}{\omega_k^2}\right)^{-(a_k+1)} \exp\left\{-\frac{b_k}{\omega_k^2}\right\} I_{(0, +\infty)}(\omega_k^2) \\
 &\times \exp\left\{-\frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2}\right\} I_{(-\infty, +\infty)}(\alpha_k) \\
 &\times v_k^{(c_k-1)} \exp\left\{-\frac{v_k}{d_k}\right\} I_{(0, \infty)}(v_k).
 \end{aligned}$$

Isto é

$$\begin{aligned}
\pi(\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v}, \mathbf{p}, K | \mathbf{y}, \mathbf{s}, \boldsymbol{\tau}, \mathbf{a}, \mathbf{b}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{c}, \mathbf{d}) &\propto \prod_{k=1}^K \left\{ p_k^{n_k + \tau_k - 1} \left(\frac{1}{\omega_k} \right)^{n_k} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2}) \sqrt{v_k \pi}} \right)^{n_k} \right. \\
&\times \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \\
&\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2}) \sqrt{(v_k+1)\pi}} \\
&\times \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \left. \right\} \\
&\times \left(\frac{1}{\omega_k^2} \right)^{1/2} \exp \left\{ -\frac{(\xi_k - \mu_k)^2}{2\omega_k^2} \right\} I_{(-\infty, +\infty)}(\xi_k) \\
&\times \left(\frac{1}{\omega_k^2} \right)^{-(a_k+1)} \exp \left\{ -\frac{b_k}{\omega_k^2} \right\} I_{(0, +\infty)}(\omega_k^2) \\
&\times \exp \left\{ -\frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} \right\} I_{(-\infty, +\infty)}(\alpha_k) \\
&\times v_k^{(c_k-1)} \exp \left\{ -\frac{v_k}{d_k} \right\} I_{(0, \infty)}(v_k) \\
&\propto \prod_{k=1}^K \left\{ p_k^{n_k + \tau_k - 1} \left(\frac{\Gamma(\frac{v_k+1}{2})}{\Gamma(\frac{v_k}{2}) \sqrt{v_k \pi}} \right)^{n_k} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_k)^2}{\omega_k^2 v_k} \right)^{-\left(\frac{v_k+1}{2}\right)} \right. \\
&\times \prod_{i=1}^N \int_{-\infty}^{\infty} \alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+2}{v_k+1 + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \frac{\Gamma(\frac{v_k+2}{2})}{\Gamma(\frac{v_k+1}{2}) \sqrt{(v_k+1)\pi}} \times \left(1 + \frac{u^2}{v_k+1} \right)^{-\left(\frac{v_k+2}{2}\right)} du \\
&\times \left. \prod_{i=1}^N I_{(-\infty, +\infty)}(y_i)^{I_{\{s_i\}}(k)} \right\} \\
&\times \left(\frac{1}{\omega_k^2} \right)^{(a_k+2n_k+1/2+1)} v_k^{(c_k-1)} \exp \left\{ -\frac{(\xi_k - \mu_k)^2}{2\omega_k^2} - \frac{b_k}{\omega_k^2} - \frac{(\alpha_k - \mu_k)^2}{2\sigma_k^2} - \frac{v_k}{d_k} \right\}.
\end{aligned} \tag{C.3}$$

Portanto, as condicionais para os parâmetros do modelo de mistura de distribuições t de Student assimétricas podem resumidas em

$$\begin{aligned}
\xi_k | \omega_k^2, \mu_k, \mathbf{y}, \mathbf{s} &\sim \phi \left(\mu_k, \omega_k^2 \right) t_{v_k} \left(\frac{y_i - \xi_k}{\omega_k} \right) T_{v_k+1} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+1}{v_k + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \right), \\
\omega_k^2 | a_k, b_k, \mathbf{y}, \mathbf{s} &\sim \text{Gamma Inversa} \left(a_k + 2n_k + \frac{1}{2}, b_k \right) t_{v_k} \left(\frac{y_i - \xi_k}{\omega_k} \right) T_{v_k+1} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+1}{v_k + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \right), \\
\alpha_k | \mu_k, \sigma_k^2, \mathbf{y}, \mathbf{s} &\sim \phi \left(\mu_k, \sigma_k^2 \right) t_{v_k} \left(\frac{y_i - \xi_k}{\omega_k} \right) T_{v_k+1} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+1}{v_k + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \right), \\
v_k | c_k, d_k, \mathbf{y}, \mathbf{s} &\sim \text{Gamma} (c_k, d_k) t_{v_k} \left(\frac{y_i - \xi_k}{\omega_k} \right) T_{v_k+1} \left(\alpha_k \left(\frac{y_i - \xi_k}{\omega_k} \right) \sqrt{\frac{v_k+1}{v_k + \left(\frac{y_i - \xi_k}{\omega_k}\right)^2}} \right), \\
p_k | \boldsymbol{\tau}_k, \mathbf{y}, \mathbf{s} &\sim \text{Dirichlet} (n_k + \tau_k),
\end{aligned}$$

em que n_k é o número de observações da componente k .

CÁLCULO DAS PROBABILIDADES DE ACEITAÇÃO DO MÉTODO RJ

D.1 Cálculo da probabilidade de aceitação do movimento “nascimento” para o modelo de mistura de distribuições normais assimétricas

Temos que

$$\pi(\mathcal{M}_{K+1}) = \pi(\boldsymbol{\theta}, K+1) = \pi(\boldsymbol{\theta}|K+1). \quad (\text{D.1})$$

Como $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha})'$ temos que

$$\begin{aligned} \pi(\mathcal{M}_{K+1}) &= \pi(\boldsymbol{\theta}, K+1) = \pi(\boldsymbol{\theta}|K+1) \\ &= \pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K+1)\pi(\boldsymbol{\omega}^2|K+1)\pi(\boldsymbol{\alpha}|K+1). \end{aligned} \quad (\text{D.2})$$

Logo

$$\begin{aligned} A &= \frac{\pi(\mathcal{M}_{K+1})}{\pi(\mathcal{M}_K)} \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{\text{Beta}(K\tau, \tau)} \frac{q(K|K+1)}{q(K+1|K)(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} (1-p_{k^*})^K \\ &= \frac{\pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K+1)\pi(\boldsymbol{\omega}^2|K+1)\pi(\boldsymbol{\alpha}|K+1)}{\pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K)\pi(\boldsymbol{\omega}^2|K)\pi(\boldsymbol{\alpha}|K)} \\ &\times \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{\text{Beta}(K\tau, \tau)} \\ &\times \frac{q(K|K+1)}{q(K+1|K)} \frac{1}{(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} (1-p_{k^*})^K \end{aligned} \quad (\text{D.3})$$

D.2 Cálculo da probabilidade de aceitação do movimento “nascimento” para o modelo de mistura de distribuições t de Student assimétricas

Neste caso, temos que $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\omega}^2, \boldsymbol{\alpha}, \mathbf{v})'$ temos que

$$\begin{aligned}\pi(\mathcal{M}_{K+1}) &= \pi(\boldsymbol{\theta}, K+1) = \pi(\boldsymbol{\theta}|K+1) \\ &= \pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K+1)\pi(\boldsymbol{\omega}^2|K+1)\pi(\boldsymbol{\alpha}|K+1)\pi(\mathbf{v}|K+1).\end{aligned}\quad (\text{D.4})$$

Logo

$$\begin{aligned}A &= \frac{\pi(\mathcal{M}_{K+1})}{\pi(\mathcal{M}_K)} \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{\text{Beta}(K\tau, \tau)} \frac{q(K|K+1)}{q(K+1|K)(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} (1-p_{k^*})^K \\ &= \frac{\pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K+1)\pi(\boldsymbol{\omega}^2|K+1)\pi(\boldsymbol{\alpha}|K+1)\pi(\mathbf{v}|K+1)}{\pi(\boldsymbol{\xi}|\boldsymbol{\omega}^2, K)\pi(\boldsymbol{\omega}^2|K)\pi(\boldsymbol{\alpha}|K)\pi(\mathbf{v}|K)} \\ &\times \frac{p_{k^*}^{\tau-1}(1-p_{k^*})^{N+K\tau-K}(K+1)}{\text{Beta}(K\tau, \tau)} \\ &\times \frac{q(K|K+1)}{q(K+1|K)} \frac{1}{(k_0+1)} \frac{1}{g_{1,K}(p_{k^*})} (1-p_{k^*})^K\end{aligned}\quad (\text{D.5})$$

D.3 Cálculo da probabilidade de aceitação do movimento split para o modelo de mistura de distribuições normais assimétricas

$$\begin{aligned}
 A^{SP} &= \frac{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_1, \boldsymbol{\theta}_{k_1})]^{I_{\{s_i\}}(k_1)} \prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_2, \boldsymbol{\theta}_{k_2})]^{I_{\{s_i\}}(k_2)}}{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k^*, \boldsymbol{\theta}_{k^*})]^{I_{\{s_i\}}(k^*)}} \\
 &\times (K+1) \frac{\pi(\mathcal{M}_{K+1}) p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{\pi(\mathcal{M}_K) p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{\text{Beta}(\tau, K\tau)} \\
 &\times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \\
 &\times \frac{d_{K+1}}{b_K P_{\text{alloc}}(u_1)g(u_2)g(u_3)g(u_4)} \\
 &\times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4 \omega_{k^*}^2} \\
 &= \prod_{i=1}^N \left[\frac{2}{\sqrt{2\pi}\omega_{k_1}^2} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)^2 \right\} \int_{-\infty}^{\alpha_{k_1} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \right]^{I_{\{s_i\}}(k_1)} \\
 &\times \prod_{i=1}^N \left[\frac{2}{\sqrt{2\pi}\omega_{k_2}^2} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)^2 \right\} \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \right]^{I_{\{s_i\}}(k_2)} \\
 &\times \frac{1}{\prod_{i=1}^N \left[\frac{2}{\sqrt{2\pi}\omega_{k^*}^2} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{k^*}}{\omega_{k^*}} \right)^2 \right\} \int_{-\infty}^{\alpha_{k^*} \left(\frac{y_i - \xi_{k^*}}{\omega_{k^*}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \right]^{I_{\{s_i\}}(k^*)}} \\
 &\times (K+1) \frac{\pi(\boldsymbol{\xi} | \boldsymbol{\omega}^2, K+1) \pi(\boldsymbol{\omega}^2 | K+1) \pi(\boldsymbol{\alpha} | K+1)}{\pi(\boldsymbol{\xi} | \boldsymbol{\omega}^2, K) \pi(\boldsymbol{\omega}^2 | K) \pi(\boldsymbol{\alpha} | K)} \\
 &\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{\text{Beta}(\tau, K\tau)} \\
 &\times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \\
 &\times \frac{(1/2)}{(1/2) P_{\text{alloc}}(u_1)g(u_2)g(u_3)g(u_4)} \\
 &\times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4 \omega_{k^*}^2},
 \end{aligned}$$

utilizando alguns resultados da Expressão (B.1) do Apêndice B, temos

$$\begin{aligned}
A^{sp} &= \left[\left(\frac{2}{\sqrt{2\pi}\omega_{k_1}^2} \right)^{n_{k_1}} \exp \left\{ - \left[\frac{(S_{k_1}^2)(n_{k_1} - 1) + n_{k_1}(\bar{y}_{k_1} - \xi_{k_1})^2}{2\omega_{k_1}^2} \right] \right\} \right] \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_1} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k_1)} \\
&\times \left[\left(\frac{2}{\sqrt{2\pi}\omega_{k_2}^2} \right)^{n_{k_2}} \exp \left\{ - \left[\frac{(S_{k_2}^2)(n_{k_2} - 1) + n_{k_2}(\bar{y}_{k_2} - \xi_{k_2})^2}{2\omega_{k_2}^2} \right] \right\} \right] \\
&\times \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k_2)} \\
&\times \frac{1}{\left[\left(\frac{2}{\sqrt{2\pi}\omega_{k^*}^2} \right)^{n_{k^*}} \exp \left\{ - \left[\frac{(S_{k^*}^2)(n_{k^*} - 1) + n_{k^*}(\bar{y}_{k^*} - \xi_{k^*})^2}{2\omega_{k^*}^2} \right] \right\} \right]} \\
&\times \frac{1}{\prod_{i=1}^N \int_{-\infty}^{\alpha_{k^*} \left(\frac{y_i - \xi_{k^*}}{\omega_{k^*}} \right)} \frac{1}{\sqrt{2\pi}} \exp \left\{ - \frac{u^2}{2} \right\} du I_{(-\infty, +\infty)}(y_i) \Big]^{I_{\{s_i\}}(k^*)} \\
&\times (K+1) \frac{\pi(\xi | \omega^2, K+1) \pi(\omega^2 | K+1) \pi(\alpha | K+1)}{\pi(\xi | \omega^2, K) \pi(\omega^2 | K) \pi(\alpha | K)} \\
&\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{Beta(\tau, K\tau)} \\
&\times \frac{\pi(\xi_{k_1}) \pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2) \pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1}) \pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \\
&\times \frac{1}{P_{aloc} g(u_1) g(u_2) g(u_3) g(u_4)} \times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1}}{u_2(1-u_2^2) u_3(1-u_3) u_4 \omega_{k^*}^2}. \tag{D.6}
\end{aligned}$$

D.4 Cálculo da probabilidade de aceitação do movimento split para o modelo de mistura de distribuições t de Student assimétricas

$$\begin{aligned}
 A^{SP} &= \frac{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_1, \boldsymbol{\theta}_{k_1})]^{I_{\{s_i\}}(k_1)} \prod_{i=1}^N [f_{Y_i}(y_i | S_i = k_2, \boldsymbol{\theta}_{k_2})]^{I_{\{s_i\}}(k_2)}}{\prod_{i=1}^N [f_{Y_i}(y_i | S_i = k^*, \boldsymbol{\theta}_{k^*})]^{I_{\{s_i\}}(k^*)}} \\
 &\times (K+1) \frac{\pi(\mathcal{M}_{K+1})}{\pi(\mathcal{M}_K)} \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{\text{Beta}(\tau, K\tau)} \\
 &\times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \times \frac{\pi(v_{k_1})\pi(v_{k_2})}{\pi(v_{k^*})} \\
 &\times \frac{d_{K+1}}{b_K P_{\text{alloc}} g(u_1)g(u_2)g(u_3)g(u_4)g(u_5)} \times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1} v_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4 u_5 \omega_{k^*}^2},
 \end{aligned}$$

então

$$\begin{aligned}
 A^{SP} &= \left[\left(\frac{2}{\omega_{k_1}} \right)^{n_{k_1}} \left(\frac{\Gamma(\frac{v_{k_1}+1}{2})}{\Gamma(\frac{v_{k_1}}{2})\sqrt{v_{k_1}}\pi} \right)^{n_{k_1}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_1})^2}{\omega_{k_1}^2 v_{k_1}} \right)^{-\left(\frac{v_{k_1}+1}{2}\right)} \right. \\
 &\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_1} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)} \frac{\sqrt{\frac{v_{k_1}+2}{v_{k_1}+1 + \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}}\right)^2}}}{\sqrt{v_{k_1}+1 + \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_1}+2}{2})}{\Gamma(\frac{v_{k_1}+1}{2})\sqrt{(v_{k_1}+1)\pi}} \left(1 + \frac{u^2}{v_{k_1}+1} \right)^{-\left(\frac{v_{k_1}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_1)} \\
 &\times \left[\left(\frac{2}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}}\pi} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right. \\
 &\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \frac{\sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}}}{\sqrt{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_2)} \\
 &\times \frac{1}{\left[\left(\frac{2}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}}\pi} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right.} \\
 &\left. \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \frac{\sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}}}{\sqrt{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}}\right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_2)} \\
 &\times (K+1) \frac{\pi(\boldsymbol{\xi} | \boldsymbol{\omega}^2, K+1) \pi(\boldsymbol{\omega}^2 | K+1) \pi(\boldsymbol{\alpha} | K+1) \pi(\mathbf{v} | K+1)}{\pi(\boldsymbol{\xi} | \boldsymbol{\omega}^2, K) \pi(\boldsymbol{\omega}^2 | K) \pi(\boldsymbol{\alpha} | K) \pi(\mathbf{v} | K)} \\
 &\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{\text{Beta}(\tau, K\tau)} \times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \times \frac{\pi(v_{k_1})\pi(v_{k_2})}{\pi(v_{k^*})} \\
 &\times \frac{(1/2)}{(1/2) P_{\text{alloc}} g(u_1)g(u_2)g(u_3)g(u_4)g(u_5)} \times \frac{p_{k^*} |\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1} v_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4 u_5 \omega_{k^*}^2},
 \end{aligned}$$

portanto,

$$\begin{aligned}
A^{SP} &= \left[\left(\frac{2}{\omega_{k_1}} \right)^{n_{k_1}} \left(\frac{\Gamma(\frac{v_{k_1}+1}{2})}{\Gamma(\frac{v_{k_1}}{2})\sqrt{v_{k_1}\pi}} \right)^{n_{k_1}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_1})^2}{\omega_{k_1}^2 v_{k_1}} \right)^{-\left(\frac{v_{k_1}+1}{2}\right)} \right. \\
&\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_1} \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)} \sqrt{\frac{v_{k_1}+2}{v_{k_1}+1 + \left(\frac{y_i - \xi_{k_1}}{\omega_{k_1}} \right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_1}+2}{2})}{\Gamma(\frac{v_{k_1}+1}{2})\sqrt{(v_{k_1}+1)\pi}} \left(1 + \frac{u^2}{v_{k_1}+1} \right)^{-\left(\frac{v_{k_1}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_1)} \\
&\times \left[\left(\frac{2}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}\pi}} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right. \\
&\times \left. \prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_2)} \\
&\times \frac{1}{\left[\left(\frac{1}{\omega_{k_2}} \right)^{n_{k_2}} \left(\frac{\Gamma(\frac{v_{k_2}+1}{2})}{\Gamma(\frac{v_{k_2}}{2})\sqrt{v_{k_2}\pi}} \right)^{n_{k_2}} \prod_{i=1}^N \left(1 + \frac{(y_i - \xi_{k_2})^2}{\omega_{k_2}^2 v_{k_2}} \right)^{-\left(\frac{v_{k_2}+1}{2}\right)} \right]} \\
&\times \frac{\prod_{i=1}^N \int_{-\infty}^{\alpha_{k_2} \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)} \sqrt{\frac{v_{k_2}+2}{v_{k_2}+1 + \left(\frac{y_i - \xi_{k_2}}{\omega_{k_2}} \right)^2}} \left\{ \frac{\Gamma(\frac{v_{k_2}+2}{2})}{\Gamma(\frac{v_{k_2}+1}{2})\sqrt{(v_{k_2}+1)\pi}} \left(1 + \frac{u^2}{v_{k_2}+1} \right)^{-\left(\frac{v_{k_2}+2}{2}\right)} du \right\} \right]^{I_{\{s_i\}}(k_2)}} \\
&\times (K+1) \frac{\pi(\xi|\omega^2, K+1)\pi(\omega^2|K+1)\pi(\alpha|K+1)\pi(v|K+1)}{\pi(\xi|\omega^2, K)\pi(\omega^2|K)\pi(\alpha|K)\pi(v|K)} \\
&\times \frac{p_{k_1}^{\tau-1+n_1} p_{k_2}^{\tau-1+n_2}}{p_{k^*}^{\tau-1+n_{k^*}}} \frac{1}{Beta(\tau, K\tau)} \\
&\times \frac{\pi(\xi_{k_1})\pi(\xi_{k_2})}{\pi(\xi_{k^*})} \times \frac{\pi(\omega_{k_1}^2)\pi(\omega_{k_2}^2)}{\pi(\omega_{k^*}^2)} \times \frac{\pi(\alpha_{k_1})\pi(\alpha_{k_2})}{\pi(\alpha_{k^*})} \times \frac{\pi(v_{k_1})\pi(v_{k_2})}{\pi(v_{k^*})} \\
&\times \frac{1}{P_{alloc}g(u_1)g(u_2)g(u_3)g(u_4)g(u_5)} \times \frac{p_{k^*}|\mu_{k_1} - \mu_{k_2}| \omega_{k_1}^2 \omega_{k_2}^2 \alpha_{k_1} v_{k_1}}{u_2(1-u_2^2)u_3(1-u_3)u_4 u_5 \omega_{k^*}^2}. \tag{D.7}
\end{aligned}$$

CÁLCULO DAS PROBABILIDADES DE ACEITAÇÃO DO MÉTODO DDRJ

- E.1** Cálculo da probabilidade de aceitação do movimento *split* para o modelo de mistura de distribuições normais assimétricas

A razão de verossimilhanças para a proposta *split* é dada por

$$\begin{aligned}
& \frac{\mathcal{L}(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP})}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p}, K | \mathbf{y}, \mathbf{s})} \\
&= \frac{\prod_{i: s_i^{SP} = s_{i_1}^{SP}} [f_{Y_i}(y_i | S_i = s_{i_1}^{SP}, \boldsymbol{\theta}^{SP})]^{I_{\{s_i\}}(s_{i_1}^{SP})} \prod_{i: s_i^{SP} = s_{i_2}^{SP}} [f_{Y_i}(y_i | S_i = s_{i_2}^{SP}, \boldsymbol{\theta}^{SP})]^{I_{\{s_i\}}(s_{i_2}^{SP})} \prod_{k=1}^{K-1} p_k^{SP}}{\prod_{i: s_i = s_{i_k}} [f_{Y_i}(y_i | S_i = s_{i_k}, \boldsymbol{\theta}_{s_{i_k}})]^{I_{\{s_i\}}(s_{i_k})} \prod_{k=1}^K p_k^{Rk}} \\
&= \prod_{i: s_i^{SP} = s_{i_1}^{SP}} \left[\frac{2}{\sqrt{2\pi\omega_{s_{i_1}^{SP}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_1}^{SP}}}{\omega_{s_{i_1}^{SP}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_1}^{SP}}} \left(\frac{y_i - \xi_{s_{i_1}^{SP}}}{\omega_{s_{i_1}^{SP}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \right] \\
&\times \prod_{i: s_i^{SP} = s_{i_2}^{SP}} \left[\frac{2}{\sqrt{2\pi\omega_{s_{i_2}^{SP}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_2}^{SP}}}{\omega_{s_{i_2}^{SP}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_2}^{SP}}} \left(\frac{y_i - \xi_{s_{i_2}^{SP}}}{\omega_{s_{i_2}^{SP}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \right] \\
&\times \prod_{i: s_i = s_{i_k}} \frac{1}{\prod_{k=1}^{K+1} p_k^{SP}} \left[\frac{2}{\sqrt{2\pi\omega_{s_{i_k}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_k}}}{\omega_{s_{i_k}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_k}}} \left(\frac{y_i - \xi_{s_{i_k}}}{\omega_{s_{i_k}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \right] \\
&\times \prod_{j=1}^2 \left[\frac{2}{\sqrt{2\pi\omega_{s_{i_j}^{SP}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_j}^{SP}}} \left(\frac{y_i - \xi_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du \right] \\
&= \frac{\prod_{i: s_i = s_{i_k}} \frac{2}{\sqrt{2\pi\omega_{s_{i_k}}^2}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i - \xi_{s_{i_k}}}{\omega_{s_{i_k}}} \right)^2 \right\} \int_{-\infty}^{\alpha_{s_{i_k}}} \left(\frac{y_i - \xi_{s_{i_k}}}{\omega_{s_{i_k}}} \right) \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{u^2}{2} \right\} du}{\prod_{k=1}^{K+1} p_k^{SP} \prod_{k=1}^K p_k^{Rk}}, \tag{E.1}
\end{aligned}$$

e a razão de distribuições *a priori* é dada por

$$\begin{aligned}
 & \frac{\pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1)}{\pi(\boldsymbol{\theta}, \mathbf{p}, K)} \\
 &= K \times \frac{\pi(\boldsymbol{\theta}_{S_1}^{SP})\pi(\boldsymbol{\theta}_{S_2}^{SP})}{\pi(\boldsymbol{\theta}_{S_{1k}})} \\
 &= K \times \frac{\pi(\xi_{S_1}^{SP}, \omega_{S_1}^{2SP}, \alpha_{S_1}^{SP})\pi(\xi_{S_2}^{SP}, \omega_{S_2}^{2SP}, \alpha_{S_2}^{SP})}{\pi(\xi_{S_{1k}}, \omega_{S_{1k}}^2, \alpha_{S_{1k}})}, \\
 &= K \times \frac{\pi(\xi_{S_1}^{SP} | \omega_{S_1}^{2SP})\pi(\omega_{S_1}^{2SP})\pi(\alpha_{S_1}^{SP})\pi(\omega_{S_2}^{2SP})\pi(\xi_{S_2}^{SP} | \omega_{S_2}^{2SP})\pi(\omega_{S_2}^{2SP})\pi(\alpha_{S_2}^{SP})}{\pi(\xi_{S_{1k}} | \omega_{S_{1k}}^2)\pi(\omega_{S_{1k}}^2)\pi(\alpha_{S_{1k}})}, \\
 &= K \times \\
 &\times \frac{1}{\sqrt{2\pi}\omega_{S_1}^{2SP}} \exp\left\{-\frac{(\xi_{S_1}^{SP} - \mu_{S_1}^{SP})^2}{2\omega_{S_1}^{2SP}}\right\} \frac{b_{S_1}^{SP}}{\Gamma(a_{S_1}^{SP})} \left(\frac{1}{\omega_{S_1}^{2SP}}\right)^{-(a_{S_1}^{SP}+1)} \exp\left\{-\frac{b_{S_1}^{SP}}{\omega_{S_1}^{2SP}}\right\} \frac{1}{\sqrt{2\pi}\sigma_{S_1}^{2SP}} \exp\left\{-\frac{(\alpha_{S_1}^{SP} - \mu_{S_1}^{SP})^2}{2\sigma_{S_1}^{2SP}}\right\} \\
 &\times \frac{1}{\sqrt{2\pi}\omega_{S_2}^{2SP}} \exp\left\{-\frac{(\xi_{S_2}^{SP} - \mu_{S_2}^{SP})^2}{2\omega_{S_2}^{2SP}}\right\} \frac{b_{S_2}^{SP}}{\Gamma(a_{S_2}^{SP})} \left(\frac{1}{\omega_{S_2}^{2SP}}\right)^{-(a_{S_2}^{SP}+1)} \exp\left\{-\frac{b_{S_2}^{SP}}{\omega_{S_2}^{2SP}}\right\} \frac{1}{\sqrt{2\pi}\sigma_{S_2}^{2SP}} \exp\left\{-\frac{(\alpha_{S_2}^{SP} - \mu_{S_2}^{SP})^2}{2\sigma_{S_2}^{2SP}}\right\} \\
 &\times \frac{1}{\sqrt{2\pi}\omega_{S_{1k}}^2} \exp\left\{-\frac{(\xi_{S_{1k}} - \mu_{S_{1k}})^2}{2\omega_{S_{1k}}^2}\right\} \frac{b_{S_{1k}}^{SP}}{\Gamma(a_{S_{1k}}^{SP})} \left(\frac{1}{\omega_{S_{1k}}^2}\right)^{-(a_{S_{1k}}^{SP}+1)} \exp\left\{-\frac{b_{S_{1k}}^{SP}}{\omega_{S_{1k}}^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_{S_{1k}}^2} \exp\left\{-\frac{(\alpha_{S_{1k}} - \mu_{S_{1k}})^2}{2\sigma_{S_{1k}}^2}\right\} \\
 &= K \times \frac{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi}\omega_{S_j}^{2SP}} \exp\left\{-\frac{(\xi_{S_j}^{SP} - \mu_{S_j}^{SP})^2}{2\omega_{S_j}^{2SP}}\right\} \frac{b_{S_j}^{SP}}{\Gamma(a_{S_j}^{SP})} \left(\frac{1}{\omega_{S_j}^{2SP}}\right)^{-(a_{S_j}^{SP}+1)} \exp\left\{-\frac{b_{S_j}^{SP}}{\omega_{S_j}^{2SP}}\right\} \frac{1}{\sqrt{2\pi}\sigma_{S_j}^{2SP}} \exp\left\{-\frac{(\alpha_{S_j}^{SP} - \mu_{S_j}^{SP})^2}{2\sigma_{S_j}^{2SP}}\right\} \right]}{\frac{1}{\sqrt{2\pi}\omega_{S_{1k}}^2} \exp\left\{-\frac{(\xi_{S_{1k}} - \mu_{S_{1k}})^2}{2\omega_{S_{1k}}^2}\right\} \frac{b_{S_{1k}}^{SP}}{\Gamma(a_{S_{1k}}^{SP})} \left(\frac{1}{\omega_{S_{1k}}^2}\right)^{-(a_{S_{1k}}^{SP}+1)} \exp\left\{-\frac{b_{S_{1k}}^{SP}}{\omega_{S_{1k}}^2}\right\} \frac{1}{\sqrt{2\pi}\sigma_{S_{1k}}^2} \exp\left\{-\frac{(\alpha_{S_{1k}} - \mu_{S_{1k}})^2}{2\sigma_{S_{1k}}^2}\right\}}}. \tag{E.2}
 \end{aligned}$$

Utilizando os resultados das Expressões (E.1) e (E.2) temos que a razão da distribuição proposta é dada por

$$\begin{aligned}
& \frac{q(x|x^{SP})}{q(x^{SP}|x)} \\
&= \frac{2n_{s_1}^{SP}n_{s_2}^{SP}}{n_{s_{1k}}(n_{s_{1k}}-1)\left(\frac{1}{2}\right)^{n_{s_1}^{SP}+n_{s_2}^{SP}-2}} \frac{\pi(\boldsymbol{\theta}_{s_1}, \mathbf{p}, K|y, \mathbf{s})}{\pi(\boldsymbol{\theta}_{s_1}^{SP}, \mathbf{p}^{SP}, K+1|y, \mathbf{s}^{SP})\pi(\boldsymbol{\theta}_{s_2}^{SP}, \mathbf{p}^{SP}, K+1|y, \mathbf{s}^{SP})} \\
&= \frac{2n_{s_1}^{SP}n_{s_2}^{SP}}{n_{s_{1k}}(n_{s_{1k}}-1)\left(\frac{1}{2}\right)^{n_{s_1}^{SP}+n_{s_2}^{SP}-2}} \frac{\mathcal{L}(\boldsymbol{\theta}_{s_1}, \mathbf{p}, K|y, \mathbf{s})\pi(\boldsymbol{\theta}_{s_1}, \mathbf{p}, K)}{\mathcal{L}(\boldsymbol{\theta}_{s_1}^{SP}, \mathbf{p}^{SP}, K+1|y, \mathbf{s}^{SP})\pi(\boldsymbol{\theta}_{s_1}, \mathbf{p}^{SP}, K+1)\mathcal{L}(\boldsymbol{\theta}_{s_2}^{SP}, \mathbf{p}^{SP}, K+1|y, \mathbf{s}^{SP})\pi(\boldsymbol{\theta}_{s_2}, \mathbf{p}^{SP}, K+1)} \\
&= \frac{2n_{s_1}^{SP}n_{s_2}^{SP}}{n_{s_{1k}}(n_{s_{1k}}-1)\left(\frac{1}{2}\right)^{n_{s_1}^{SP}+n_{s_2}^{SP}-2}} \\
&\times \frac{2}{\sqrt{2\pi}\omega_{s_{1k}}^2} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_{s_{1k}}}{\omega_{s_{1k}}}\right)^2\right\} \int_{-\infty}^{\alpha_{s_{1k}}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du \prod_{k=1}^{n_k} p_k^{n_k} \\
&\times \frac{1}{\sqrt{2\pi}\omega_{s_{1k}}^2} \exp\left\{-\frac{(\xi_{s_{1k}} - \mu_{s_{1k}})^2}{2\omega_{s_{1k}}^2}\right\} \frac{b_{s_{1k}}^{\alpha_{s_{1k}}}}{\Gamma(d_{s_{1k}})} \left(\frac{1}{\omega_{s_{1k}}}\right)^{-(d_{s_{1k}}+1)} \exp\left\{-\frac{(\alpha_{s_{1k}} - \mu_{s_{1k}})^2}{2\sigma_{s_{1k}}^2}\right\} \\
&\times \prod_{j=1}^2 \left[\frac{2}{\sqrt{2\pi}\omega_{s_{1j}}^{2SP}} \exp\left\{-\frac{1}{2}\left(\frac{y_i - \xi_{s_{1j}}^{SP}}{\omega_{s_{1j}}^{SP}}\right)^2\right\} \int_{-\infty}^{\alpha_{s_{1j}}^{SP}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du \prod_{k=1}^{n_k} p_k^{n_k} \right] \\
&\times \prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi}\omega_{s_{1j}}^{2SP}} \exp\left\{-\frac{(\xi_{s_{1j}}^{SP} - \mu_{s_{1j}}^{SP})^2}{2\omega_{s_{1j}}^{2SP}}\right\} \frac{b_{s_{1j}}^{d_{s_{1j}}^{SP}}}{\Gamma(d_{s_{1j}}^{SP})} \left(\frac{1}{\omega_{s_{1j}}^{SP}}\right)^{-(d_{s_{1j}}^{SP}+1)} \exp\left\{-\frac{(\alpha_{s_{1j}}^{SP} - \mu_{s_{1j}}^{SP})^2}{2\sigma_{s_{1j}}^{2SP}}\right\} \right]
\end{aligned} \tag{E.3}$$

E.2 Cálculo da probabilidade de aceitação do movimento *split* para o modelo de mistura de distribuições *t* de Student assimétricas

A razão de verossimilhanças para a proposta *split* é dada por

$$\begin{aligned}
 & \frac{\mathcal{L}(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP})}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p}, K | \mathbf{y}, \mathbf{s})} \\
 &= \frac{\prod_{i: s_i^{SP} = s_i^{SP}} [f_{Y_i}(y_i | S_i = s_i^{SP}, \boldsymbol{\theta}^{SP})]^{I_{\{s_i\}}(s_i^{SP})} \prod_{i: s_i^{SP} = s_i^{SP}} [f_{Y_i}(y_i | S_i = s_i^{SP}, \boldsymbol{\theta}^{SP})]^{I_{\{s_i\}}(s_i^{SP})} \prod_{k=1}^{K+1} p_k^{SP}}{\prod_{i: S_i = s_{i_k}} [f_{Y_i}(y_i | S_i = s_{i_k}, \boldsymbol{\theta}_{s_{i_k}})]^{I_{\{s_i\}}(s_{i_k})} \prod_{k=1}^K p_k^{nk}} \\
 &= \prod_{j=1}^2 \left[\frac{2}{\omega_{S_{ij}}^{SP}} \frac{\Gamma\left(\frac{v_{S_{ij}}^{SP}+1}{2}\right)}{\Gamma\left(\frac{v_{S_{ij}}^{SP}}{2}\right)} \sqrt{v_{S_{ij}}^{SP} \pi} \left(1 + \frac{\left(\frac{y_j - s_{S_{ij}}^{SP}}{\omega_{S_{ij}}^{SP}}\right)^2}{v_{S_{ij}}^{SP}}\right)^{-\left(\frac{v_{S_{ij}}^{SP}+1}{2}\right)} \right. \\
 &\quad \times \left. \alpha_{S_{ij}}^{SP} \left(\frac{y_j - s_{S_{ij}}^{SP}}{\omega_{S_{ij}}^{SP}}\right) \sqrt{\frac{v_{S_{ij}}^{SP}+2}{v_{S_{ij}}^{SP}+1 + \left(\frac{y_j - s_{S_{ij}}^{SP}}{\omega_{S_{ij}}^{SP}}\right)^2}} \frac{\Gamma\left(\frac{v_{S_{ij}}^{SP}+2}{2}\right)}{\Gamma\left(\frac{v_{S_{ij}}^{SP}+1}{2}\right)} \sqrt{\frac{v_{S_{ij}}^{SP}+1}{v_{S_{ij}}^{SP}+1}} \pi \left(1 + \frac{u^2}{v_{S_{ij}}^{SP}+1}\right)^{-\left(\frac{v_{S_{ij}}^{SP}+2}{2}\right)} du \right] \\
 &\quad \times \frac{1}{\frac{2}{\omega_{S_{ik}}}} \frac{\Gamma\left(\frac{v_{S_{ik}}+1}{2}\right)}{\Gamma\left(\frac{v_{S_{ik}}}{2}\right)} \sqrt{v_{S_{ik}} \pi} \left(1 + \frac{\left(\frac{y_j - s_{S_{ik}}}{\omega_{S_{ik}}}\right)^2}{v_{S_{ik}}}\right)^{-\left(\frac{v_{S_{ik}}+1}{2}\right)} \\
 &\quad \times \frac{1}{\int_{-\infty}^{\infty} \alpha_{S_{ik}} \left(\frac{y_j - s_{S_{ik}}}{\omega_{S_{ik}}}\right) \sqrt{\frac{v_{S_{ik}}+2}{v_{S_{ik}}+1 + \left(\frac{y_j - s_{S_{ik}}}{\omega_{S_{ik}}}\right)^2}} \frac{\Gamma\left(\frac{v_{S_{ik}}+2}{2}\right)}{\Gamma\left(\frac{v_{S_{ik}}+1}{2}\right)} \sqrt{\frac{v_{S_{ik}}+1}{v_{S_{ik}}+1}} \pi \left(1 + \frac{u^2}{v_{S_{ik}}+1}\right)^{-\left(\frac{v_{S_{ik}}+2}{2}\right)} du} \prod_{k=1}^{K+1} p_k^{SP} \prod_{k=1}^K p_k^{nk} \tag{E.5}
 \end{aligned}$$

e a razão de distribuições *a priori* é dada por

$$\begin{aligned}
 & \frac{\pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1)}{\pi(\boldsymbol{\theta}, \mathbf{p}, K)} \\
 &= K \times \frac{\pi(\boldsymbol{\theta}_{S_1}^{SP})\pi(\boldsymbol{\theta}_{S_2}^{SP})}{\pi(\boldsymbol{\theta}_{S_{1k}})} \\
 &= K \times \frac{\pi(\xi_{S_1}^{SP}, \omega_{S_1}^{2SP}, \alpha_{S_1}^{SP}, \nu_{S_1}^{SP})\pi(\xi_{S_2}^{SP}, \omega_{S_2}^{2SP}, \alpha_{S_2}^{SP}, \nu_{S_2}^{SP})}{\pi(\xi_{S_{1k}}^2, \omega_{S_{1k}}^2, \alpha_{S_{1k}}, \nu_{S_{1k}})}, \\
 &= K \times \frac{\pi(\xi_{S_1}^{SP} | \omega_{S_1}^{2SP})\pi(\alpha_{S_1}^{SP})\pi(\nu_{S_1}^{SP})\pi(\omega_{S_2}^{2SP})\pi(\xi_{S_2}^{SP})\pi(\alpha_{S_2}^{SP})\pi(\nu_{S_2}^{SP})}{\pi(\xi_{S_{1k}} | \omega_{S_{1k}}^2)\pi(\omega_{S_{1k}}^2)\pi(\alpha_{S_{1k}})\pi(\nu_{S_{1k}})} \\
 &= K \times \frac{1}{\sqrt{2\pi}\omega_{S_1}^{2SP}} \exp\left\{-\frac{(\xi_{S_1}^{SP} - \mu_{S_1}^{SP})^2}{2\omega_{S_1}^{2SP}}\right\} \frac{b_{S_1}^{SP}}{\Gamma(\alpha_{S_1}^{SP})} \left(\frac{1}{\omega_{S_1}^{2SP}}\right)^{-(\alpha_{S_1}^{SP}+1)} \exp\left\{-\frac{b_{S_1}^{SP}}{\omega_{S_1}^{2SP}}\right\} \sqrt{2\pi}\sigma_{S_1}^{2SP} \frac{1}{\Gamma(c_{S_1}^{SP})} \frac{1}{d_{S_1}^{SP}} \exp\left\{-\frac{(\alpha_{S_1}^{SP} - \mu_{S_1}^{SP})^2}{2\sigma_{S_1}^{2SP}}\right\} \frac{1}{\Gamma(c_{S_1}^{SP})} \frac{1}{d_{S_1}^{SP}} \exp\left\{-\frac{\nu_{S_1}^{SP}}{d_{S_1}^{SP}}\right\} \\
 &\times \frac{1}{\sqrt{2\pi}\omega_{S_2}^{2SP}} \exp\left\{-\frac{(\xi_{S_2}^{SP} - \mu_{S_2}^{SP})^2}{2\omega_{S_2}^{2SP}}\right\} \frac{b_{S_2}^{SP}}{\Gamma(\alpha_{S_2}^{SP})} \left(\frac{1}{\omega_{S_2}^{2SP}}\right)^{-(\alpha_{S_2}^{SP}+1)} \exp\left\{-\frac{b_{S_2}^{SP}}{\omega_{S_2}^{2SP}}\right\} \sqrt{2\pi}\sigma_{S_2}^{2SP} \frac{1}{\Gamma(c_{S_2}^{SP})} \frac{1}{d_{S_2}^{SP}} \exp\left\{-\frac{(\alpha_{S_2}^{SP} - \mu_{S_2}^{SP})^2}{2\sigma_{S_2}^{2SP}}\right\} \frac{1}{\Gamma(c_{S_2}^{SP})} \frac{1}{d_{S_2}^{SP}} \exp\left\{-\frac{\nu_{S_2}^{SP}}{d_{S_2}^{SP}}\right\} \\
 &\times \frac{1}{\sqrt{2\pi}\omega_{S_{1k}}^2} \exp\left\{-\frac{(\xi_{S_{1k}} - \mu_{S_{1k}})^2}{2\omega_{S_{1k}}^2}\right\} \frac{b_{S_{1k}}^{SP}}{\Gamma(\alpha_{S_{1k}}^{SP})} \left(\frac{1}{\omega_{S_{1k}}^2}\right)^{-(\alpha_{S_{1k}}^{SP}+1)} \exp\left\{-\frac{b_{S_{1k}}^{SP}}{\omega_{S_{1k}}^2}\right\} \sqrt{2\pi}\sigma_{S_{1k}}^2 \frac{1}{\Gamma(c_{S_{1k}}^{SP})} \frac{1}{d_{S_{1k}}^{SP}} \exp\left\{-\frac{(\alpha_{S_{1k}}^{SP} - \mu_{S_{1k}})^2}{2\sigma_{S_{1k}}^2}\right\} \frac{1}{\Gamma(c_{S_{1k}}^{SP})} \frac{1}{d_{S_{1k}}^{SP}} \exp\left\{-\frac{\nu_{S_{1k}}^{SP}}{d_{S_{1k}}^{SP}}\right\} \\
 &= K \times \frac{\prod_{j=1}^2 \left[\frac{1}{\sqrt{2\pi}\omega_{S_{1j}}^{2SP}} \exp\left\{-\frac{(\xi_{S_{1j}}^{SP} - \mu_{S_{1j}}^{SP})^2}{2\omega_{S_{1j}}^{2SP}}\right\} \frac{b_{S_{1j}}^{SP}}{\Gamma(\alpha_{S_{1j}}^{SP})} \left(\frac{1}{\omega_{S_{1j}}^{2SP}}\right)^{-(\alpha_{S_{1j}}^{SP}+1)} \exp\left\{-\frac{b_{S_{1j}}^{SP}}{\omega_{S_{1j}}^{2SP}}\right\} \sqrt{2\pi}\sigma_{S_{1j}}^{2SP} \frac{1}{\Gamma(c_{S_{1j}}^{SP})} \frac{1}{d_{S_{1j}}^{SP}} \exp\left\{-\frac{(\alpha_{S_{1j}}^{SP} - \mu_{S_{1j}}^{SP})^2}{2\sigma_{S_{1j}}^2}\right\} \frac{1}{\Gamma(c_{S_{1j}}^{SP})} \frac{1}{d_{S_{1j}}^{SP}} \exp\left\{-\frac{\nu_{S_{1j}}^{SP}}{d_{S_{1j}}^{SP}}\right\} \right]}{\frac{1}{\sqrt{2\pi}\omega_{S_{1k}}^2} \exp\left\{-\frac{(\xi_{S_{1k}} - \mu_{S_{1k}})^2}{2\omega_{S_{1k}}^2}\right\} \frac{b_{S_{1k}}^{SP}}{\Gamma(\alpha_{S_{1k}}^{SP})} \left(\frac{1}{\omega_{S_{1k}}^2}\right)^{-(\alpha_{S_{1k}}^{SP}+1)} \exp\left\{-\frac{b_{S_{1k}}^{SP}}{\omega_{S_{1k}}^2}\right\} \sqrt{2\pi}\sigma_{S_{1k}}^2 \frac{1}{\Gamma(c_{S_{1k}}^{SP})} \frac{1}{d_{S_{1k}}^{SP}} \exp\left\{-\frac{(\alpha_{S_{1k}}^{SP} - \mu_{S_{1k}})^2}{2\sigma_{S_{1k}}^2}\right\} \frac{1}{\Gamma(c_{S_{1k}}^{SP})} \frac{1}{d_{S_{1k}}^{SP}} \exp\left\{-\frac{\nu_{S_{1k}}^{SP}}{d_{S_{1k}}^{SP}}\right\}}}. \tag{E.6}
 \end{aligned}$$

A razão da distribuição proposta é dada por

$$\begin{aligned}
 q(x|x^{SP}) &= \frac{2n_{S_{1j}}^{SP} n_{S_{1j}}^{SP}}{n_{S_{1k}}(n_{S_{1k}} - 1) \left(\frac{1}{2}\right)_{-1}^{n_{S_{1j}}^{SP} + n_{S_{1j}}^{SP} - 2}} \pi(\boldsymbol{\theta}_{S_1}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP}) \pi(\boldsymbol{\theta}_{S_2}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP}) \\
 q(x^SP|x) &= \frac{2n_{S_{1j}}^{SP} n_{S_{1j}}^{SP}}{n_{S_{1k}}(n_{S_{1k}} - 1) \left(\frac{1}{2}\right)_{-1}^{n_{S_{1j}}^{SP} + n_{S_{1j}}^{SP} - 2}} \mathcal{L}(\boldsymbol{\theta}_{S_{1k}} \cdot \mathbf{p}, K | \mathbf{y}, \mathbf{s}) \pi(\boldsymbol{\theta}_{S_{1k}} \cdot \mathbf{p}, K) \\
 &= \frac{2n_{S_{1j}}^{SP} n_{S_{1j}}^{SP}}{n_{S_{1k}}(n_{S_{1k}} - 1) \left(\frac{1}{2}\right)_{-1}^{n_{S_{1j}}^{SP} + n_{S_{1j}}^{SP} - 2}} \mathcal{L}(\boldsymbol{\theta}_{S_1}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP}) \pi(\boldsymbol{\theta}_{S_1}^{SP}, \mathbf{p}^{SP}, K+1) \mathcal{L}(\boldsymbol{\theta}_{S_2}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP}) \pi(\boldsymbol{\theta}_{S_2}^{SP}, \mathbf{p}^{SP}, K+1).
 \end{aligned}$$

Finalmente a probabilidade de aceitação do movimento *split* é dada por

$$\begin{aligned}
 A^{SP} &= \frac{\mathcal{L}(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1 | \mathbf{y}, \mathbf{s}^{SP}) \pi(\boldsymbol{\theta}^{SP}, \mathbf{p}^{SP}, K+1) q(\mathbf{x} | \mathbf{x}^{SP})}{\mathcal{L}(\boldsymbol{\theta}, \mathbf{p}, K | \mathbf{y}, \mathbf{s}) \pi(\boldsymbol{\theta}, \mathbf{p}, K) q(\mathbf{x}^{SP} | \mathbf{x})} \\
 &= \prod_{j=1}^2 \left[\frac{\Gamma\left(\frac{v_{S_{1j}}^{SP}+1}{2}\right)}{\omega_{S_{1j}}^{SP} \Gamma\left(\frac{v_{S_{1j}}^{SP}}{2}\right) \sqrt{v_{S_{1j}}^{SP}} \pi} \left(1 + \frac{\left(\frac{y_j - \xi_{S_{1j}}^{SP}}{\omega_{S_{1j}}^{SP}}\right)^2}{v_{S_{1j}}^{SP}} \right)^{-\left(\frac{v_{S_{1j}}^{SP}+1}{2}\right)} \right. \\
 &\quad \times \left. \int_{-\infty}^{\infty} \frac{\alpha_{S_{1j}}^{SP} \left(\frac{y_j - \xi_{S_{1j}}^{SP}}{\omega_{S_{1j}}^{SP}}\right) \sqrt{\frac{v_{S_{1j}}^{SP}+2}{v_{S_{1j}}^{SP}+1} \left(\frac{y_j - \xi_{S_{1j}}^{SP}}{\omega_{S_{1j}}^{SP}}\right)^2}}{\Gamma\left(\frac{v_{S_{1j}}^{SP}+1}{2}\right) \sqrt{(v_{S_{1j}}^{SP}+1)\pi}} \left(1 + \frac{u^2}{v_{S_{1j}}^{SP}+1} \right)^{-\left(\frac{v_{S_{1j}}^{SP}+2}{2}\right)} du \right] \\
 &\quad \times \frac{1}{\frac{2}{\omega_{S_{1k}}^{SP}} \Gamma\left(\frac{v_{S_{1k}}^{SP}+1}{2}\right) \sqrt{v_{S_{1k}}^{SP}} \pi \left(1 + \frac{\left(\frac{y_j - \xi_{S_{1k}}^{SP}}{\omega_{S_{1k}}^{SP}}\right)^2}{v_{S_{1k}}^{SP}} \right)^{-\left(\frac{v_{S_{1k}}^{SP}+1}{2}\right)}} \\
 &\quad \times \frac{1}{\int_{-\infty}^{\infty} \frac{\alpha_{S_{1k}}^{SP} \left(\frac{y_j - \xi_{S_{1k}}^{SP}}{\omega_{S_{1k}}^{SP}}\right) \sqrt{\frac{v_{S_{1k}}^{SP}+2}{v_{S_{1k}}^{SP}+1} \left(\frac{y_j - \xi_{S_{1k}}^{SP}}{\omega_{S_{1k}}^{SP}}\right)^2}}{\Gamma\left(\frac{v_{S_{1k}}^{SP}+1}{2}\right) \sqrt{(v_{S_{1k}}^{SP}+1)\pi}} \left(1 + \frac{u^2}{v_{S_{1k}}^{SP}+1} \right)^{-\left(\frac{v_{S_{1k}}^{SP}+2}{2}\right)} du} \\
 &\quad \times \prod_{j=1}^K \left[\frac{1}{\sqrt{2\pi} \omega_{S_{1j}}^{SP}} \exp\left\{-\frac{(\xi_{S_{1j}}^{SP} - \mu_{S_{1j}}^{SP})^2}{2\omega_{S_{1j}}^{SP}}\right\} \Gamma\left(\frac{a_{S_{1j}}^{SP}}{2}\right) \left(\frac{1}{\omega_{S_{1j}}^{SP}}\right)^{-\left(a_{S_{1j}}^{SP}+1\right)} \right. \\
 &\quad \times \left. \frac{b_{S_{1j}}^{SP}}{\sqrt{2\pi} \sigma_{S_{1j}}^{SP}} \exp\left\{-\frac{(\alpha_{S_{1j}}^{SP} - \mu_{S_{1j}}^{SP})^2}{2\sigma_{S_{1j}}^{SP}}\right\} \right] \\
 &\quad \times \frac{1}{\sqrt{2\pi} \alpha_{S_{1k}}^{SP}} \exp\left\{-\frac{(\xi_{S_{1k}} - \mu_{S_{1k}})^2}{2\alpha_{S_{1k}}^{SP}}\right\} \Gamma\left(\frac{a_{S_{1k}}}{2}\right) \left(\frac{1}{\alpha_{S_{1k}}}\right)^{-\left(a_{S_{1k}}+1\right)} \exp\left\{-\frac{(\alpha_{S_{1k}} - \mu_{S_{1k}})^2}{2\sigma_{S_{1k}}^2}\right\}
 \end{aligned}$$

$$\begin{aligned}
& \times \frac{2n_{s_{i_1}^{SP}}n_{s_{i_2}^{SP}}}{n_{s_{i_k}^{SP}}(n_{s_{i_k}^{SP}}-1)\left(\frac{1}{2}\right)^{n_{s_{i_1}^{SP}}+n_{s_{i_2}^{SP}}-2}} \\
& \times \frac{2}{\omega_{s_{i_k}^{SP}}}\frac{\Gamma\left(\frac{v_{s_{i_k}^{SP}}+1}{2}\right)}{\Gamma\left(\frac{v_{s_{i_k}^{SP}}}{2}\right)}\sqrt{v_{s_{i_k}^{SP}}\pi}\left(1+\frac{\left(\frac{y_i-\xi_{s_{i_k}^{SP}}}{\omega_{s_{i_k}^{SP}}}\right)^2}{v_{s_{i_k}^{SP}}}\right)^{-\left(\frac{v_{s_{i_k}^{SP}}+1}{2}\right)} \\
& \times \frac{\alpha_{s_{i_k}^{SP}}\left(\frac{y_i-\xi_{s_{i_k}^{SP}}}{\omega_{s_{i_k}^{SP}}}\right)\sqrt{\frac{v_{s_{i_k}^{SP}}-2}{v_{s_{i_k}^{SP}}+1+\left(\frac{y_i-\xi_{s_{i_k}^{SP}}}{\omega_{s_{i_k}^{SP}}}\right)^2}}}{\int_{-\infty}^{\infty}\frac{\Gamma\left(\frac{v_{s_{i_k}^{SP}}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{i_k}^{SP}}+1}{2}\right)}\sqrt{(v_{s_{i_k}^{SP}}+1)\pi}\left(1+\frac{u^2}{v_{s_{i_k}^{SP}}+1}\right)^{-\left(\frac{v_{s_{i_k}^{SP}}+2}{2}\right)}du\prod_{k=1}^{K+1}p_k^{n_k}} \\
& \times \frac{1}{\prod_{j=1}^2\left[\frac{2}{\omega_{s_{i_j}^{SP}}}\frac{\Gamma\left(\frac{v_{s_{i_j}^{SP}}+1}{2}\right)}{\Gamma\left(\frac{v_{s_{i_j}^{SP}}}{2}\right)}\sqrt{v_{s_{i_j}^{SP}}\pi}\left(1+\frac{\left(\frac{y_j-\xi_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}}\right)^2}{v_{s_{i_j}^{SP}}}\right)^{-\left(\frac{v_{s_{i_j}^{SP}}+1}{2}\right)}\right]} \\
& \times \frac{\alpha_{s_{i_j}^{SP}}\left(\frac{y_j-\xi_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}}\right)\sqrt{\frac{v_{s_{i_j}^{SP}}+2}{v_{s_{i_j}^{SP}}+1+\left(\frac{y_j-\xi_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}}\right)^2}}}{\int_{-\infty}^{\infty}\frac{\Gamma\left(\frac{v_{s_{i_j}^{SP}}+2}{2}\right)}{\Gamma\left(\frac{v_{s_{i_j}^{SP}}+1}{2}\right)}\sqrt{(v_{s_{i_j}^{SP}}+1)\pi}\left(1+\frac{u^2}{v_{s_{i_j}^{SP}}+1}\right)^{-\left(\frac{v_{s_{i_j}^{SP}}+2}{2}\right)}du\prod_{k=1}^{n_k}p_k^{n_k}} \\
& \times \frac{1}{\prod_{j=1}^2\left[\frac{1}{\sqrt{2\pi}\omega_{s_{i_j}^{SP}}^2}\exp\left\{-\frac{(\xi_{s_{i_j}^{SP}}-\mu_{s_{i_j}^{SP}})^2}{2\omega_{s_{i_j}^{SP}}^2}\right\}\frac{b_{s_{i_j}^{SP}}}{\Gamma(a_{s_{i_j}^{SP}})}\left(\frac{1}{\omega_{s_{i_j}^{SP}}}\right)^{-a_{s_{i_j}^{SP}}}\exp\left\{-\frac{b_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}^2}u\right\}\frac{1}{\sqrt{2\pi}\sigma_{s_{i_j}^{SP}}^2}\exp\left\{-\frac{(\alpha_{s_{i_j}^{SP}}-\mu_{s_{i_j}^{SP}})^2}{2\sigma_{s_{i_j}^{SP}}^2}\right\}}\right]} \\
& \times \frac{1}{\prod_{j=1}^2\left[\frac{1}{\sqrt{2\pi}\omega_{s_{i_j}^{SP}}^2}\exp\left\{-\frac{(\xi_{s_{i_j}^{SP}}-\mu_{s_{i_j}^{SP}})^2}{2\omega_{s_{i_j}^{SP}}^2}\right\}\frac{b_{s_{i_j}^{SP}}}{\Gamma(a_{s_{i_j}^{SP}})}\left(\frac{1}{\omega_{s_{i_j}^{SP}}}\right)^{-a_{s_{i_j}^{SP}}}\exp\left\{-\frac{b_{s_{i_j}^{SP}}}{\omega_{s_{i_j}^{SP}}^2}u\right\}\frac{1}{\sqrt{2\pi}\sigma_{s_{i_j}^{SP}}^2}\exp\left\{-\frac{(\alpha_{s_{i_j}^{SP}}-\mu_{s_{i_j}^{SP}})^2}{2\sigma_{s_{i_j}^{SP}}^2}\right\}}\right]}
\end{aligned}
\tag{E.8}$$

VERIFICAÇÃO DAS CADEIAS

F.1 Estudo de simulação

As Figuras 22, 23, 24 e 25 mostram, respectivamente, o desenvolvimento das cadeias ao longo da geração das observações dos modelos de mistura de distribuições normais com $K = 1, 2, 3, 4$ componentes e tamanho de amostra $N = 1000$. Podemos observar por meio dessas figuras que as representações gráficas indicam convergência das cadeias.

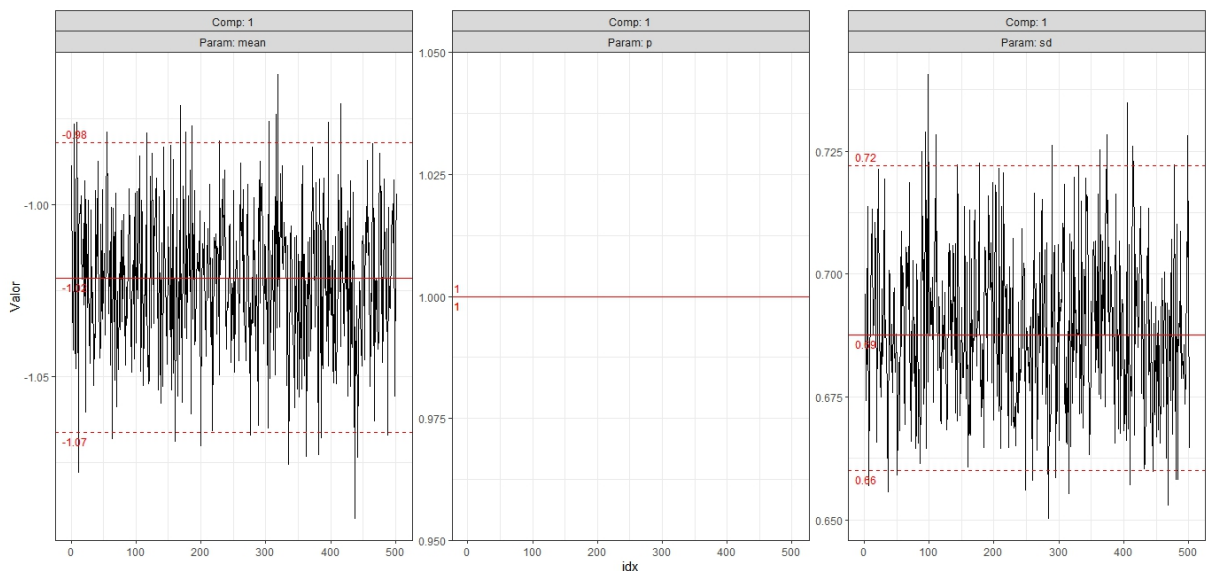


Figura 22 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 1$ componente e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).

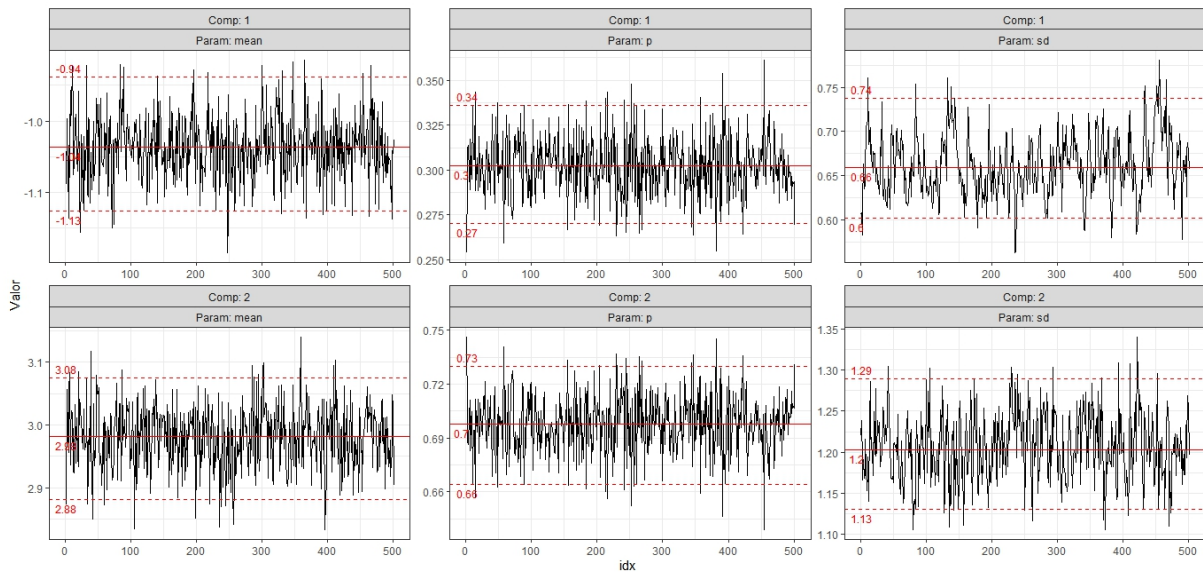


Figura 23 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 2$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).

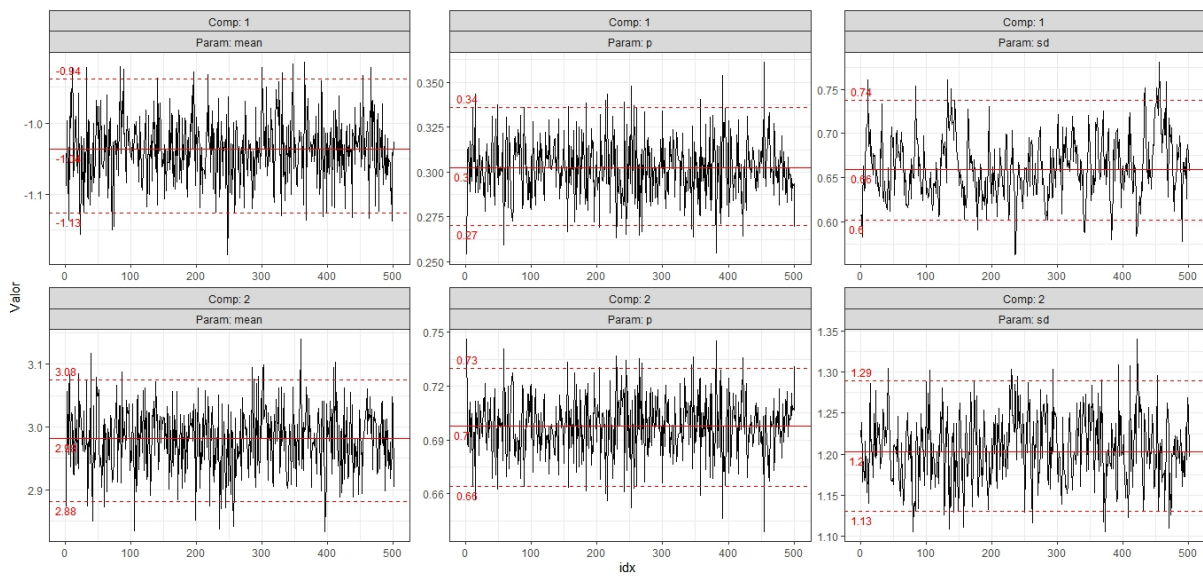


Figura 24 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 3$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).

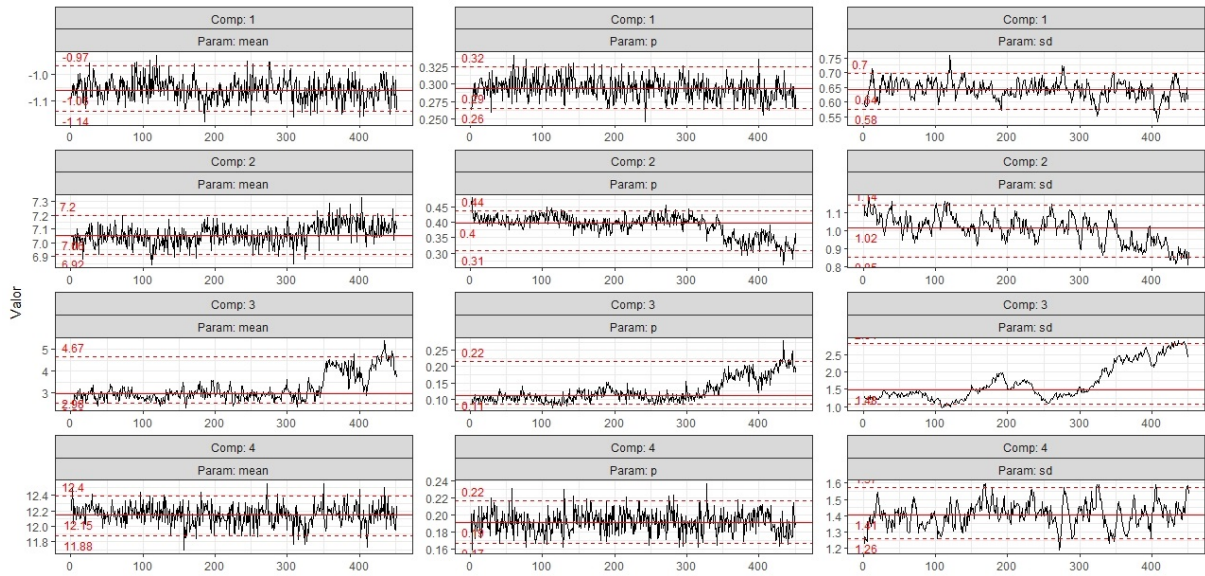


Figura 25 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 4$ componentes e $N = 1000$. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).

As Figuras 26, 27, 28 e 29 apresentam, respectivamente, os gráficos da autocorrelação parcial das observações dos modelos de mistura de distribuições normais gerados com $K = 1, 2, 3, 4$ componentes e $N = 1000$. Podemos observar por meio dessas figuras que as representações gráficas indicam ausência de autocorrelação.

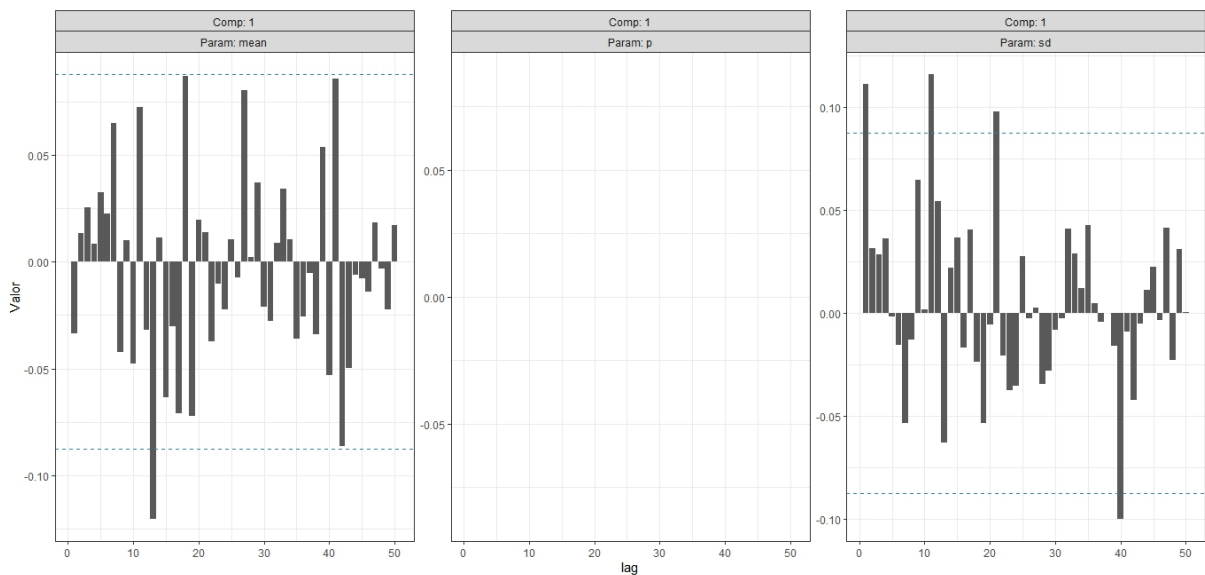


Figura 26 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 1$ componente e $N = 1000$.

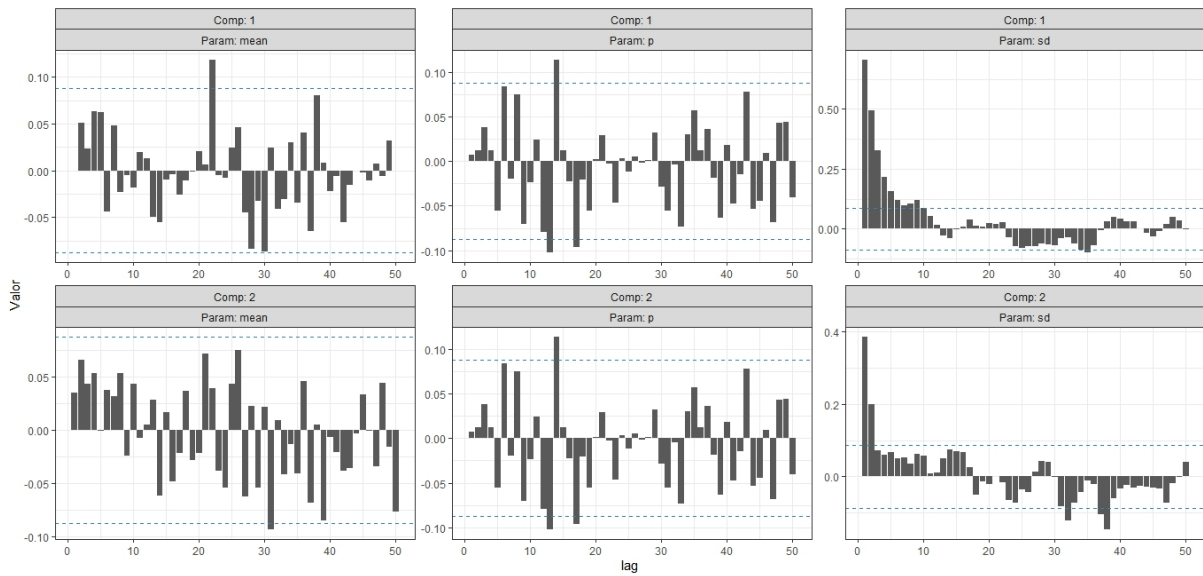


Figura 27 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 2$ componentes e $N = 1000$.

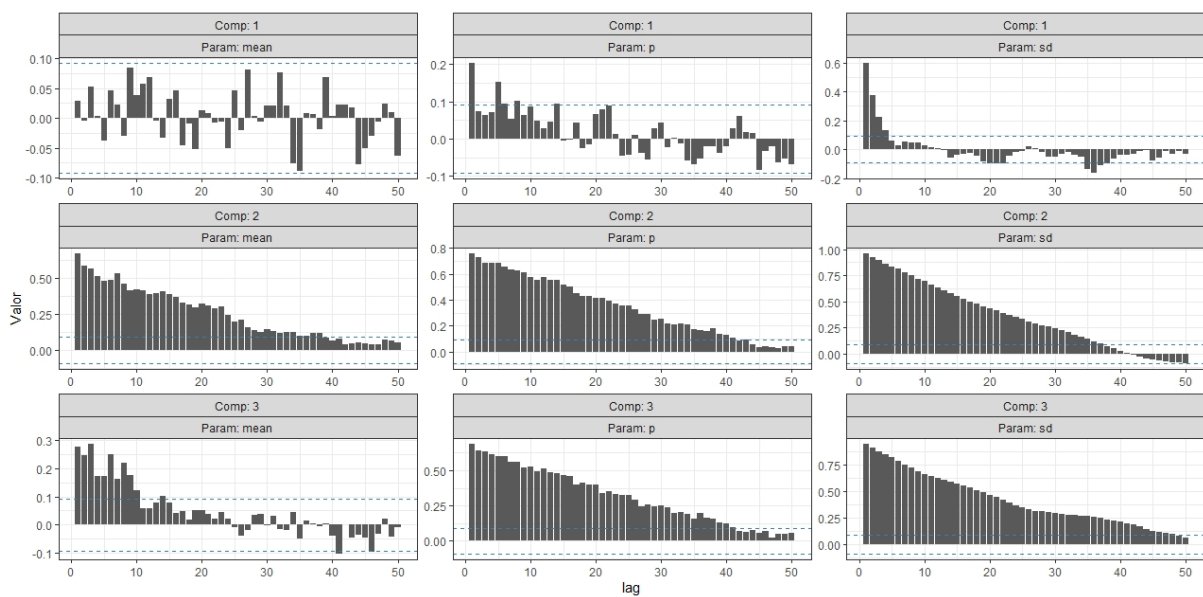


Figura 28 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 3$ componentes e $N = 1000$.

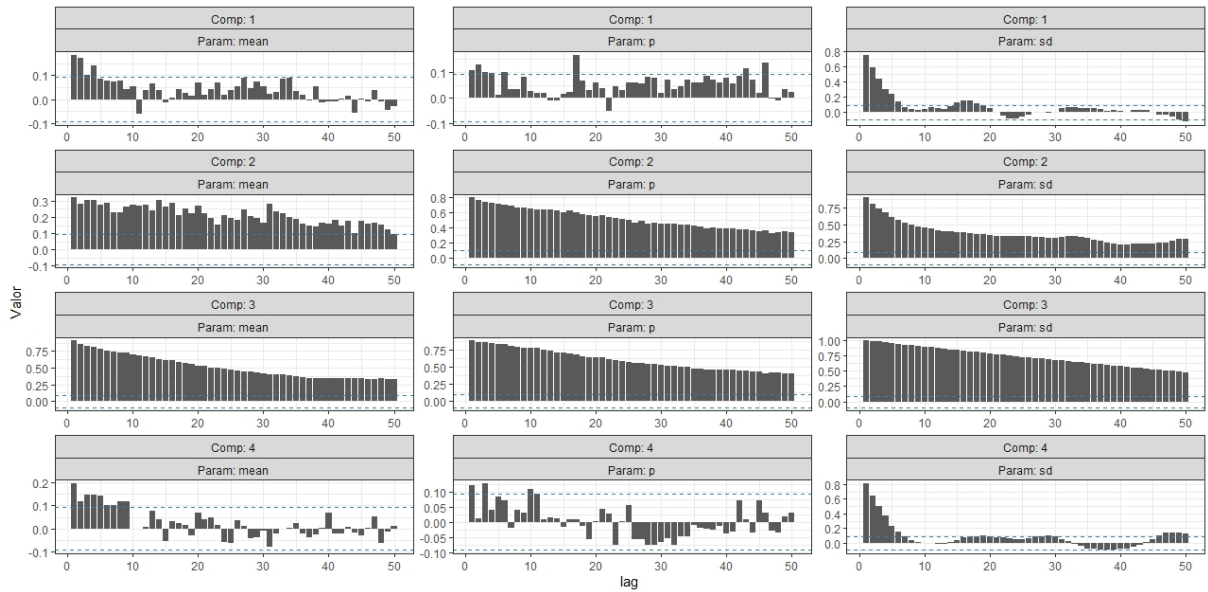


Figura 29 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 4$ componentes e $N = 1000$.

F.2 Análise de dados reais

A Figura 30 mostra o desenvolvimento das cadeias ao longo da geração das observações dos modelos de mistura de distribuições normais com $K = 2$ componentes para os dados de escores globais. Podemos observar por meio dessa figura que as representações gráficas indicam convergência das cadeias.

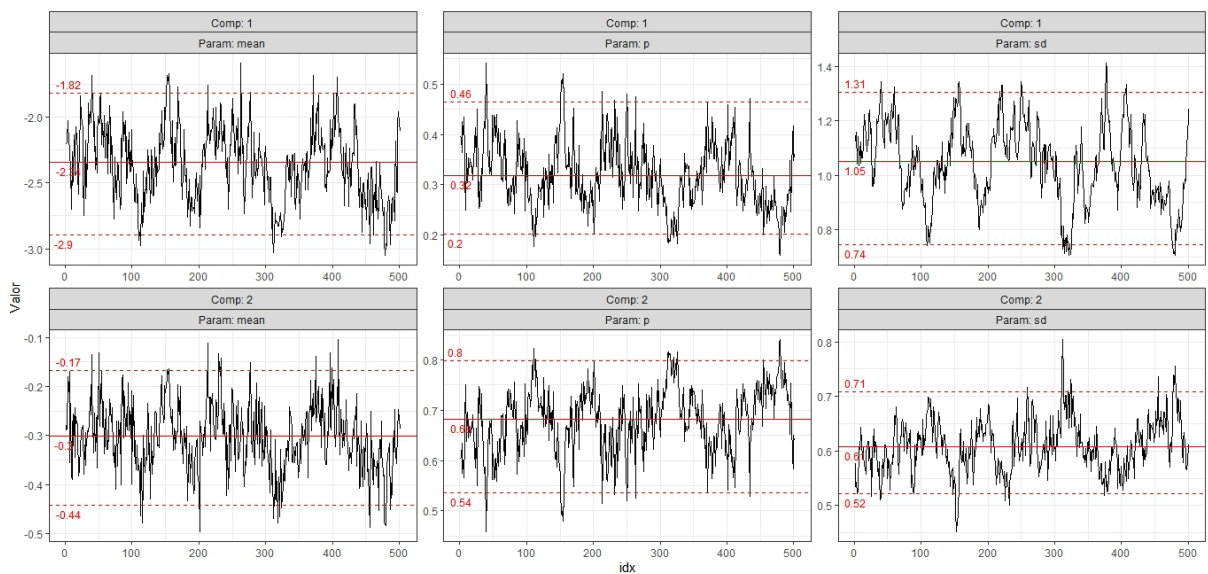


Figura 30 – Gráficos das cadeias - Modelo de mistura de distribuições normais com $K = 2$ componentes, dados dos escores globais. Valor mediano (linha sólida) e intervalo de credibilidade (linha tracejada).

A Figura 31 apresenta os gráficos da autocorrelação parcial das observações dos modelos de mistura de distribuições normais com $K = 2$ componentes para os dados dos escores globais. Podemos observar por meio dessa figura que as representações gráficas indicam ausência de autocorrelação para $lag=25$.

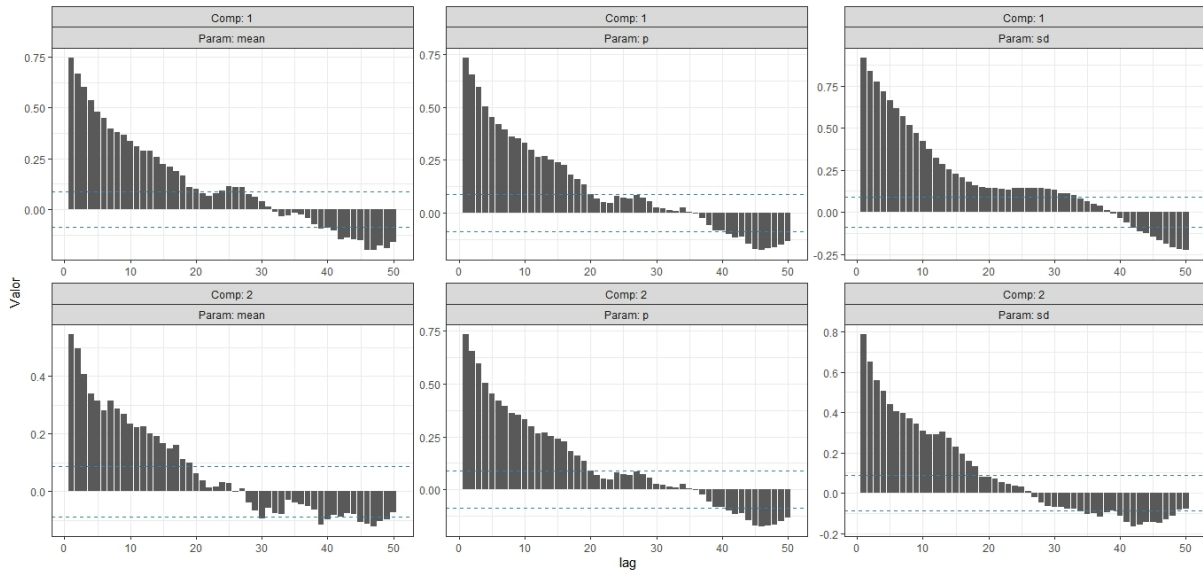


Figura 31 – Gráficos das funções de autocorrelações - Modelo de mistura de distribuições normais com $K = 2$ componentes, dados dos escores globais.

PROGRAMA

Códigos em R Core Team (2022)

```

densidade <- function(x, theta, distr, log=FALSE){
  f <- get(distr)
  d = do.call(f, c(list(x),log=TRUE,as.list(theta)))
  if ( log==TRUE ) return(d) else return(exp(d))}

densidade_k <- function(x, theta, distr, log=FALSE){
  K = length(theta)
  d = purrr::map(.x=1:K, function(k){
  f <- get(distr[k])
  theta[[k]][,1,drop=TRUE] * do.call(f, c(list(x),as.list(theta[[k]])[-1]))}) %>%
  set_names(paste0('C',1:K)) %>%
  bind_cols() %>%
  apply(MARGIN=1, FUN=sum)
  return(d)}

geradora_k <- function(n, theta, distr){
  K = length(theta)
  d = purrr::map(.x=1:K, function(k){
  f <- get(distr[k])
  do.call(f, c(list(theta[[k]][,1,drop=TRUE]*n),as.list(theta[[k]])[-1]))}) %>%
  unlist()
  return(sample(d))}

logVero_k <- function(y, theta, distr){
  ll = densidade_k(y, theta, distr) %>%
  log() %>%
  sum()
  if ( !is.finite(ll) ) return(-1e5) else return(ll)}

logPrio_k <- function(theta, par, distr){
  k1 = length(theta)
  k2 = length(par)
  k3 = length(distr)
  KJ = length(theta %>% unlist())

```

```

if ( k1 != k2 ) {
stop('"x" e "par" devem ter o mesmo comprimento')
} else if ( k1 != k3 ) {
stop('"x" e "dist" devem ter o mesmo comprimento')
} else if ( k2 != k3 ) {
stop('"par" e "dist" devem ter o mesmo comprimento')}
lp = purrr::map(.x=1:k1, .f=function(k)
J = length(distr[[k]])
purrr::map(.x=1:J, .f=function(j){
f <- get(distr[[k]][j])
logPrio = log(
do.call( f, c( list(x=theta[[k]][,j,drop=TRUE]), par[[k]][[j]] ) ) )
if ( !is.finite(logPrio) ) return( log(1e-5) ) else return( logPrio )}) %>%
set_names(paste0('P',1:J)) %>%
bind_cols()) %>%
set_names(paste0('K',1:k1)) %>%
unlist() %>%
sum(); return(lp)}

logPost_k <- function(y, param.priori.theta0, distr.priori.theta0, theta, distr.y){
logPost = logVero_k(y=y, theta=theta, distr=distr.y) +
logPrio_k(theta=theta, par=param.priori.theta0, distr=distr.priori.theta0)
return(logPost)}

geracaoCandidatos <- function(theta, espacoParametrico, sdPasso=NULL){
K = length(theta)
purrr::map(.x=1:ncol(theta), .f=function(i){
if ( any(sdPasso<0) ) {
stop('"sdPasso" deve ser um valor ou vetor de valores positivos.')}
if ( is.null(sdPasso) ) { sd = 0.2 } else {
sd = sdPasso[i]}
if ( length(espacoParametrico)!=length(sdPasso) ) {
stop('"sdPasso" e "espacoParametrico" devem ter o mesmo numero de elementos.')}
if ( length(theta)!=length(sdPasso) ) {
stop('"sdPasso" e "theta00" devem ter o mesmo numero de elementos.')}
if ( length(theta)!=length(espacoParametrico) ) {
stop('"espacoParametrico" e "theta00" devem ter o mesmo numero de elementos.')}
if ( espacoParametrico[i]=='positivo' ) {
if ( theta[,i,drop=TRUE]<0 ) {
stop('"0 valor fornecido deve ser estritamente maior do que zero.')}
m = theta[,i,drop=TRUE]
v = sdPasso**2
return( exp( log(m)+rnorm(n=1, mean=0, sd=v) ) )
} else if ( espacoParametrico[i]=='real' ) {
m = theta[,i,drop=TRUE]
return( rnorm(n=1, mean=m, sd=sdPasso) )
} else if ( espacoParametrico[i]=='0-1' ) {
if ( sd>0.25 ) {
stop('"Por favor, para o parametro limitado no intervalo (0,1) selecione um passo menor
do que 0,25.')}
if ( theta[,i,drop=TRUE]>=1 | theta[,i,drop=TRUE]<=0 ) {
stop('"0 valor fornecido deve pertencer ao intervalo (0,1).')}
m = theta[,i,drop=TRUE]
v = sdPasso**2
z = log(m/(1-m)) + rnorm(n=1, mean=0, sd=v)
return( 1/( 1+exp(-z) ) )}) %>%
set_names(names(theta)) %>%

```

```

bind_rows()

geracaoProporcoesCandidatas <- function(p, sdPasso=NULL){
  if ( any(sdPasso<0) ) {
    stop('"sdPasso" deve ser um valor ou vetor de valores positivos.')}
  if ( is.null(sdPasso) ) { sd = 0.2 } else {
    sd = sdPasso}
  m = p
  z = log(m/(1-m)) + rnorm(n=length(p), mean=0, sd=sd)
  z = 1/(1+exp(-z))
  return( z/sum(z))}

metropolis_MCMC <- function(y, theta0, distr.y, distr.priori.theta0, param.priori.theta0,
  niter, sdPassoProp, sdPasso, espacoParametrico){
  K = length(theta0)
  theta = theta0
  nRow = nrow(theta0[[1]])
  SEQ = tryCatch( (nRow+1):(nRow+niter), error=function(x)NULL )
  if ( is.null(SEQ) ) browser()
  if ( length(SEQ)==0 ) browser()
  for ( t in SEQ ) {
    p = geracaoProporcoesCandidatas(
    p=purrr::map(.x=theta, .f=function(x){ x[t-1,1] }) %>% unlist(),
    sdPasso=sdPassoProp)
    for ( k in 1:K ) {
      theta[[k]][t,1] = p[k]
      theta[[k]][t,-1] = geracaoCandidatos(
      theta=theta[[k]][t-1,-1,drop=FALSE],
      espacoParametrico=espacoParametrico,
      sdPasso=c(sdPasso))}
    A0 = purrr::map(.x=theta, .f=function(x){ head(tail(x,2),1) })
    A1 = purrr::map(.x=theta, .f=function(x){ tail(x,1) })
    logPost_T1 = logPost_k(y=y, param.priori.theta0=param.priori.theta0,
    distr.priori.theta0=distr.priori.theta0, theta=A1, distr.y=distr.y)
    logPost_T0 = logPost_k(y=y, param.priori.theta0=param.priori.theta0,
    distr.priori.theta0=distr.priori.theta0, theta=A0, distr.y=distr.y)
    probAceitacao = logPost_T1 - logPost_T0
    if ( log(runif(1)) >= probAceitacao ){
      for (k in 1:K) {
        theta[[k]][t,] = theta[[k]][t-1,]
        theta[[k]]}}}}
    return( theta )}

mode <- function(x){
  ux <- unique(x)
  ux[which.max(tabulate(match(x, ux)))]}

amostra_post_S <- function(y,theta,distr){
  K = length(theta)
  theta = purrr::map(.x=theta, .f=function(x){ x$p = ifelse(is.na(x$p),0,x$p); x })
  S = purrr::map_dbl(.x=y, .f=function(Y){
  d = purrr::map(.x=1:K, function(k){
  f <- get(distr[k])
  theta[[k]][,1,drop=TRUE] * do.call(f, c(list(Y),as.list(theta[[k]])[-1]))}) %>% unlist()
  d = ifelse(d==0,1e-10,d)
  p = d/sum(d)
  s = rmultinom(n=1, size=1, prob=p)

```

```

s = which (s==1); s}); return(S)}

amostra_post_P <- function(S,gamma=1){
K = tryCatch( max(1:max(S)), error=function(x)NULL )
if ( is.null(K) ) browser()
N = purrr::map_dbl(.x=1:K, .f=function(c){ length(S[S==c]) })
P = c(bmixture::rdirichlet(n=1, alpha=1+N))
if ( any(is.na(P)) ) browser()
return(P)}

N = 1000; K = 4
distr.y = c('dsn','dsn','dsn','dsn')
theta0 = list(data.frame(p=0.3, xi=-1.0, omega=1.7, alpha=-1.0)
, data.frame(p=0.1, xi=5.0, omega=1.0, alpha=-0.5)
, data.frame(p=0.4, xi=10.0, omega=1.5, alpha=1.0)
, data.frame(p=0.2, xi=17.0, omega=2.0, alpha=1.5))
y = geradora_k(n=N, theta=theta0, distr=rep('rsn',K))

probabilidadeSplit <- function(y, S, theta0,pos1, pos2){
K0 = length(theta0)
S0 = S
C0 = S0[pos1]
M0 = purrr::map_dbl(.x=1:K0, .f=function(k){ length(S0[S0==k]) })
P0 = M0/sum(M0)
K1 = K0+1
S1 = S0
aux0 = median(y[S0==C0])
if ( y[C0]<aux0 ) { aux00 = c(C0,K1) } else { aux00 = c(K1,C0) }
aux10 = median(y[S0==C0 & y<aux0])
aux11 = sd(y[S0==C0 & y<aux0])
aux12 = dnorm(x=y[S0==C0], mean=aux10, sd=aux11)
aux20 = median(y[S0==C0 & y>=aux0])
aux21 = sd(y[S0==C0 & y>=aux0])
aux22 = dnorm(x=y[S0==C0], mean=aux20, sd=aux21)
aux3 = aux12/(aux12+aux22)
S1[S0==C0] = purrr::map_dbl(.x=aux3, .f=function(p) sample(x=aux00, size=1, prob=c(p,1-p)))
S1[pos1] = C0
S1[pos2] = K1
M1 = purrr::map_dbl(.x=1:K1, .f=function(k){ length(S1[S1==k]) })
C1 = c(C0,K1)
P1 = M1/sum(M1)
theta1 = theta0
theta1[[K1]] = theta1[[C0]]
theta1[[C0]]$xi = mean(y[S1==C0])
theta1[[C0]]$omega = sd(y[S1==C0])
theta1[[C0]]$alpha = (mean(y[S1==C0])-mode(y[S1==C0]))/sd(y[S1==C0])
theta1[[C0]] = metropolis_MCMC(y=y[S1==C0], theta0=theta1[C0],
distr.y=distr.y, distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=c(0.1,0.05,0.1),
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)
theta1[[C0]]$p = P1[C0]
theta1[[K1]]$xi = mean(y[S1==K1])
theta1[[K1]]$omega = sd(y[S1==K1])
theta1[[K1]]$alpha = (mean(y[S1==K1])-mode(y[S1==K1]))/sd(y[S1==K1])
theta1[[K1]] = metropolis_MCMC(
y=y[S1==K1], theta0=theta1[K1],

```

```

distr.y=distr.y, distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=c(0.1,0.05,0.1),
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)
theta1[[K1]]$p = P1[K1]
P1 = amostra_post_P(S=S1)
theta1 = purrr::map(.x=1:K1, .f=function(k)
{ theta1[[k]]$p = ifelse(is.na(P1[k]),0,P1[k]); theta1[[k]] })
logPost1 = logPost_k(y=y, param.priori.theta0=rep(param.priori.theta0,K1),

distr.priori.theta0=rep(distr.priori.theta0,K1),
theta=theta1, distr.y=rep(distr.y,K1))
logPost0 = logPost_k(y=y, param.priori.theta0=rep(param.priori.theta0,K0),

distr.priori.theta0=rep(distr.priori.theta0,K0),
theta=theta0, distr.y=rep(distr.y,K0))
PA = exp(logPost1-logPost0)
PA = ifelse(is.na(PA),1e5,PA); print(PA)
list(K0=K0, S0=S0, C0=C0, theta0=theta0,
K1=K1, S1=S1, C1=C1, theta1=theta1,PA = PA)}

probabilidadeMerge <- function(y, S, theta0, pos1, pos2){
K0 = length(theta0)
S0 = S
C0 = sort(S0[c(pos1,pos2)])
M0 = purrr::map_dbl(.x=1:K0, .f=function(k){ length(S0[S0==k]) })
P0 = M0/sum(M0)
K1 = K0-1
S1 = S0
S1[S1==C0[2]] = C0[1]
S1[S1>C0[2]] = S1[S1>C0[2]]-1
C1 = C0[1]
M1 = purrr::map_dbl(.x=1:K1, .f=function(k){ length(S1[S1==k]) })
P1 = M1/sum(M1)
theta1 = theta0
theta1[[C1]]$p = theta0[[C0[1]]]$p + theta0[[C0[2]]]$p
theta1[[C1]]$xi = (theta0[[C0[1]]]$p*theta0[[C0[1]]]$xi+
  theta0[[C0[2]]]$p*theta0[[C0[2]]]$xi)/theta1[[C1]]$p
theta1[[C1]]$omega = sqrt(abs((theta0[[C0[1]]]$p*(theta0[[C0[1]]]$xi**2+t
  heta0[[C0[1]]]$omega**2)+theta0[[C0[2]]]$p*(theta0[[C0[2]]]$xi**2+
  theta0[[C0[2]]]$omega**2))/theta1[[C1]]$p-theta1[[C1]]$xi**2))
theta1[[C1]]$alpha = (theta0[[C0[1]]]$p*theta0[[C0[1]]]$alpha+
  theta0[[C0[2]]]$p*theta0[[C0[2]]]$alpha)/theta1[[C1]]$p
for ( j in C0[2]:K0 ) theta1[j] = theta0[j+1]
theta1 = theta1[1:K1]
theta1[[C1]] = metropolis_MCMC(
y=y[S1%in%C1], theta0=theta1[C1],
distr.y=distr.y, distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=c(0.1,0.05,0.1),
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)
theta1[[C1]]$p = theta0[[C0[1]]]$p + theta0[[C0[2]]]$p
P1 = amostra_post_P(S=S1)
theta1 = purrr::map(.x=1:K1, .f=function(k)
{ theta1[[k]]$p = ifelse(is.na(P1[k]),0,P1[k]); theta1[[k]] })
logPost1 = logPost_k(y=y, param.priori.theta0=rep(param.priori.theta0,K1),

```

```

distr.priori.theta0=rep(distr.priori.theta0,K1),
theta=theta1, distr.y=rep(distr.y,K1))
logPost0 = logPost_k(y=y, param.priori.theta0=rep(param.priori.theta0,K0),

distr.priori.theta0=rep(distr.priori.theta0,K0),
theta=theta0, distr.y=rep(distr.y,K0))
PA = exp(logPost1-logPost0) # exp(logVero1-logVero0) # exp(RV) #
PA = ifelse(is.na(PA),1e5,PA); print(PA)
list(K0=K0, S0=S0, C0=C0, theta0=theta0,
K1=K1, S1=S1, C1=C1, theta1=theta1, PA=PA)}

distr.y = c('dsn')
distr.priori.theta0 = list(c('ddirichlet','dnorm','dinvgamma','dnorm'))
param.priori.theta0 = list(list(c('alpha'=1), c('mean'=0,'sd'=1e2),c('shape'=1e-2,'rate'=1e-2),
c('mean'=0,'sd'=1e2)))
theta0 = list(data.frame(p=1, xi=mean(y), omega=1, alpha=0.5))

reversible_MCMC <- function(kmax, y, theta0, distr.y, distr.priori.theta0, param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=c(0.1,0.05,0.1), espacoParametrico){
B = niter
MOV = c()
auxSM = NA
thetaGeral = NULL
G.Hist = list()
u = NULL
N = length(y)
K0 = length(theta0)
S0 = tryCatch( amostra_post_S(y=y, theta=theta0, distr=rep(distr.y,K0)), error=function(x)NULL )
P0 = amostra_post_P(S=S0)
theta0 = purrr::map(.x=1:K0, .f=function(k){
aux = metropolis_MCMC(
y=y[S0==k], theta0=theta0[k],
distr.y=rep(distr.y,1), distr.priori.theta0=rep(distr.priori.theta0,1),

param.priori.theta0=rep(param.priori.theta0,1),
niter=1, sdPassoProp=0.1, sdPasso=c(0.1,0.05,0.1),
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)
aux$p = ifelse(is.na(P0[k]),0,P0[k])
aux})
theta_hat = theta0
nK = length(theta_hat)
dados.aux =
data.frame(x=seq(min(y),max(y),l=250)) %>%
mutate(
y=densidade_k(x=x, theta=theta_hat, distr=rep('dsn',nK)),
yc=densidade_k(x=x, theta=theta_hat, distr=rep('psn',nK)))
dados.aux.2 = tibble(y=y, S=factor(S0))
G.Hist[[1]] = tibble(y=y, S=factor(S0)) %>%
ggplot() +
geom_histogram(aes(x=y, fill=S), binwidth=0.5, alpha=0.5) +
geom_rug(aes(x=y, y=0, colour=S)) +
geom_line(data=dados.aux, aes(x=x, y=0.5*N*y), size=0.75, colour='black') +
scale_colour_manual(values = RColorBrewer::brewer.pal(9,'Set1')) +
scale_fill_manual(values = RColorBrewer::brewer.pal(9,'Set1')) +
guides(fill=guide_legend(nrow=1), colour=guide_legend(nrow=1)) +
labs(y='Densidade', x=NULL, colour=NULL, fill=NULL) +

```

```

coord_cartesian(ylim=c(0,85*N/1000)) +
theme_minimal() +
theme(legend.position='bottom'); print(G.Hist[[1]])

thetaGeral = tibble(
b=1, KO=NA, K1=NA, CO=NA, C1=NA,
S0=NA, S1=NA, theta0=list(theta0), theta1=NA)
for ( b in 2:B ) {
pos = sort(sample(seq(1,N), 2, replace=FALSE))
AAA = tryCatch( S0[pos[1]]==S0[pos[2]], error=function(x)NULL )
if ( is.null(AAA) | is.na(AAA) ) browser()
if ( KO==kmax ) { auxSM[b]=1; while ( S0[pos[1]]==S0[pos[2]] ) { pos=sort(sample(1:N,2)) } } else
if ( KO==1 ) { auxSM[b]=0; while ( S0[pos[1]]!=S0[pos[2]] ) { pos=sort(sample(1:N,2)) } } else
if ( AAA==TRUE ) { auxSM[b]=0 } else { auxSM[b]=1 }
if ( auxSM[b]==0 ) {
MOV[b] = 'split'
candidato = tryCatch(probabilidadeSplit(
y=y, S=S0, theta0=theta0,
pos1=pos[1], pos2=pos[2]
), error=function(x)NULL)
A1 = candidato$theta1
CO = candidato$CO
KO = candidato$KO
probAceitacao = candidato$PA
AAA = tryCatch( runif(1) < probAceitacao, error=function(x)NULL )
if ( length(AAA)==0 ) browser() else
if ( is.null(AAA) ) browser() else
if ( is.na(AAA) ) browser()
} else
if ( auxSM[b]==1 ) {
MOV[b] = 'merge'
candidato = tryCatch(probabilidadeMerge(
y=y, S=S0, theta0=theta0,
pos1=pos[1], pos2=pos[2]
), error=function(x)NULL)
A1 = candidato$theta1
CO = candidato$CO
KO = candidato$KO
probAceitacao = candidato$PA
AAA = tryCatch( runif(1) < probAceitacao, error=function(x)NULL )
if ( length(AAA)==0 ) browser() else
if ( is.null(AAA) ) browser() else
if ( is.na(AAA) ) browser()}
if ( AAA==TRUE ) {
theta1 = A1
K1 = length(theta1)
C1 = candidato$C1
S1 = tryCatch( amostra_post_S(y=y, theta=theta1, distr=rep(distr.y,K1)), error=function(x)NULL )
theta1 = purrr::map(.x=1:K1, .f=function(k){
metropolis_MCMC(
y=y[S1==k], theta0=theta1[k],
distr.y=distr.y, distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=sdPasso,
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)})
P1 = amostra_post_P(S=S1)
theta1 = purrr::map(.x=1:K1, .f=function(k){ theta1[[k]]$p = P1[k]; theta1[[k]] })

```

```

thetaGeral = bind_rows(
thetaGeral, tibble(b=b, K0=K0, K1=K1, C0=list(C0), C1=list(C1), S0=list(S0), S1=list(S1),
  theta0=list(theta0), theta1=list(theta1)))
theta0 = theta1
K0 = K1; S0 = S1
} else {
MOV[b] = '-'
theta1 = theta0
C1 = C0; K1 = K0
S1 = tryCatch( amostra_post_S(y=y, theta=theta1, distr=rep(distr.y,K1)), error=function(x)NULL )
theta1 = purrr::map(.x=1:K1, .f=function(k){
metropolis_MCMC(
y=y[S1==k], theta0=theta1[k],
distr.y=distr.y, distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
niter=1000, sdPassoProp=0.1, sdPasso=sdPasso,
espacoParametrico=c('real','positivo','real')
)[[1]] %>% tail(1)})
P1 = amostra_post_P(S=S1)
theta1 = purrr::map(.x=1:K1, .f=function(k){ theta1[[k]]$p = ifelse(is.na(P1[k]),0,P1[k]);

theta1[[k]] })
thetaGeral = bind_rows(
thetaGeral,
tibble(b=b, K0=K0, K1=K1, C0=list(C0), C1=list(C1),
S0=list(S0), S1=list(S1), theta0=list(theta0), theta1=list(theta1)))
theta0 = theta1
K0 = K1; S0 = S1}
cat(paste0('# --- Passo: ',b,' --- #\n'))
print(theta1)
theta0 = theta1
theta_hat = theta0
nK = length(theta_hat)
dados.aux =
data.frame(x=seq(min(y),max(y),l=250)) %>%
mutate(
y=densidade_k(x=x, theta=theta_hat, distr=rep('dsn',nK)),
yc=densidade_k(x=x, theta=theta_hat, distr=rep('psn',nK)))
dados.aux.2 = tibble(y=y, S=factor(S1))
G.Hist[[b+1]] = tibble(y=y, S=factor(S1)) %>%
ggplot() +
geom_histogram(aes(x=y, fill=S), binwidth=0.5, alpha=0.5) +
geom_rug(aes(x=y, y=0, colour=S)) +
geom_line(data=dados.aux, aes(x=x, y=0.5*N*y), size=0.75, colour='black') +
geom_point(data=dados.aux.2 %>% slice(pos), aes(x=y, y=75, colour=S), size=3) +
  scale_colour_manual(values = RColorBrewer::brewer.pal(9,'Set1')) +
  scale_fill_manual(values = RColorBrewer::brewer.pal(9,'Set1')) +
  guides(fill=guide_legend(nrow=1), colour=guide_legend(nrow=1)) +
  labs(y='Densidade', x=NULL, colour=NULL, fill=NULL, title=MOV[b]) +
  coord_cartesian(ylim=c(0,90*N/1000)) + theme_minimal() +
  theme(legend.position='bottom'); print(G.Hist[[b+1]])}
return( list(thetaGeral,G.Hist) )}

niter = 50000
DDRJ = reversible_MCMC(kmax=10, y=y, theta0=theta0, distr.y=distr.y,
  distr.priori.theta0=distr.priori.theta0, param.priori.theta0=param.priori.theta0,
  niter=niter, sdPassoProp=0.1, sdPasso=c(0.05,0.01,0.05), espacoParametrico=c('real','positivo','real'))

```