

**UNIVERSIDADE DE SÃO PAULO**

Instituto de Ciências Matemáticas e de Computação

**Seleção de modelos de associação RC utilizando reversible  
jump MCMC**

**Flávio Fagundes Ferreira**

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



SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

**Flávio Fagundes Ferreira**

## Seleção de modelos de associação RC utilizando reversible jump MCMC

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, 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 – São Carlos**  
**Setembro de 2025**

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

Ferreira, Flávio Fagundes  
FF383s Seleção de modelos de associação RC utilizando  
s reversible jump MCMC / Flávio Fagundes Ferreira;  
orientador Luis Aparecido Milan. -- São Carlos,  
2025.  
228 p.

Tese (Doutorado - Programa Interinstitucional de  
Pós-graduação em Estatística) -- Instituto de  
Ciências Matemáticas e de Computação, Universidade  
de São Paulo, 2025.

1. Modelo de associação RC. 2. Tabela de  
contingência. 3. Metropolis-Hastings. 4. Reversible  
jump MCMC. 5. Análise de dados educacionais. I.  
Milan, Luis Aparecido, orient. II. Título.

**Flávio Fagundes Ferreira**

Model selection for the RC association model using  
reversible jump MCMC

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar, in partial fulfillment of the requirements for the degree of the Doctorate Interagency Program Graduate in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Luis Aparecido Milan

**USP – São Carlos**  
**September 2025**





# 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 do candidato Flavio Fagundes Ferreira, realizada em 09/06/2025.

### Comissão Julgadora:

Prof. Dr. Luis Aparecido Milan (UFSCar)

Prof. Dr. Erlandson Ferreira Saraiva (UFMS)

Prof. Dr. Anderson Luiz Ara Souza (UFPR)

Profa. Dra. Lia Hanna Martins Morita (UFMT)

Profa. Dra. Clarice Garcia Borges Demétrio (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 Maria Ester (in memoriam) e Elpídio (in memoriam), Flávia e Mateus.*



# AGRADECIMENTOS

---

---

Em primeiro lugar agradeço a Deus pela vida e pelas oportunidades que me concedeu e que permitiu que eu chegasse até aqui com saúde, muita luta e força de vontade.

À minha mãe, Maria Ester (*in memorian*), pelo amor incondicional e que sempre me apoiou e acreditou na educação antes de tudo, apesar de pouco estudo, mas de uma imensa sabedoria. Ao meu pai, Elpídio (*in memorian*) que desde cedo me colocou no caminho da escola e nunca me deixou desistir, meu amor e minha gratidão eterna.

Ao meu irmão Flaviano (*in memorian*) que não pode estar comigo nesta caminhada, mas certamente estaria torcendo por mim a cada dia. A meu irmão Fernando, pelo apoio nas horas difíceis, sempre um entusiasta da vida acadêmica e sempre incentivou meu retorno aos estudos.

À minha esposa, Flávia e ao meu filho, Mateus, com apenas 9 anos de idade, pelo carinho e compreensão incondicional, pelos inúmeros momentos que precisei estar ausente nesses longos anos do doutorado.

Ao meu orientador Prof. Dr. Luis A. Milan, que aceitou o desafio de orientação, pela paciência, compreensão, incentivo, pelas ideias, sugestões, pela amizade e compreensão, principalmente na reta final.

Aos meus amigos Júlio César, Victor Hugo e João Batista, que puderam me ouvir, incentivar e me direcionar para o caminho certo nos momentos de aperto e apreensão.

Aos professores do Programa Interinstitucional de Pós-Graduação em Estatística USP-UFSCar, que ministraram aulas à distância, on-line, em uma situação inédita, nunca vivenciada, e mesmo com todas as dificuldades puderam transmitir seus conhecimentos com empenho, dedicação e maestria. E aos colegas do curso de Doutorado pela ajuda ao longo do curso.

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

Ao INEP, por conceder o afastamento durante o curso de Doutorado e investido no meu crescimento profissional e acreditado no meu potencial, meu muito obrigado. A todos(as) os(as) colegas da Coordenação Geral de Monitoramento e Integração Logística (CGMIL) do INEP, em especial da Coordenação de Monitoramento e Gestão de Riscos (CMGR), por acreditarem que eu conseguiria chegar até o fim de pé, meus calorosos agradecimentos.



*“All our dreams can come true, if we have the courage to pursue them.”*  
*(Walt Disney)*



# RESUMO

FERREIRA, F. F. **Seleção de modelos de associação RC utilizando reversible jump MCMC**. 2025. 228 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2025.

O modelo de associação RC (Rows and Columns) de ordem  $K$  aplicado em análises de tabela de contingência fornece os valores das estimativas dos parâmetros que avaliam o grau de associação entre as categorias das variáveis dispostas em linhas e colunas da tabela. Propomos uma nova metodologia para a estimação da ordem  $K$  do modelo de associação utilizando inferência Bayesiana e o reversible jump via MCMC (RJMCMC). Nos gráficos ilustramos os agrupamentos através de intervalos de credibilidade para confirmar os resultados obtidos via RJMCMC. Aplicamos a metodologia proposta em dados simulados para validação do método e também em dados de estudos anteriores para fins de comparação e os resultados foram convergentes. Posteriormente aplicamos a metodologia proposta em três bancos de dados reais na avaliação do desempenho de participantes do Enem 2023, nas provas de linguagens, matemática e redação segundo a escolaridade do pai e da mãe. Concluímos que existe associação entre algumas categorias do desempenho dos estudantes e a escolaridade. Houve indicação gráfica de agrupamento de linhas e colunas tanto nos resultados com dados simulados quanto nos resultados do estudo comparativo.

**Palavras-chave:** modelo de associação RC, tabela de contingência, inferência Bayesiana, procrustes, biplot, Metroplis-Hastings, Reversible Jump MCMC.



# ABSTRACT

FERREIRA, F. F. **Model selection for the RC association model using reversible jump MCMC**. 2025. 228 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2025.

The RC (Rows and Columns) association model of order  $K$  applied in contingency table analysis provides the values of the parameter estimates that evaluate the degree of association between the categories of the variables arranged in rows and columns of the table. We propose a new methodology for estimating the order  $K$  of the association model using Bayesian inference and reversible jump by MCMC (RJMCMC). In the graphics, we illustrated the groupings through credible intervals to confirm the results obtained by RJMCMC. We applied the proposed methodology to simulated data to validate the method and also to data from previous studies for comparison purposes, and the results were convergent. We subsequently we applied the proposed methodology to three real databases to assess the performance of Enem 2023 participants in the language, mathematics and writing tests according to their father's and mother's education. We conclude that there is an association between some categories performance and education. There is a graphical indication of grouping of rows and columns both in the results with simulated data and in the results of the comparative study.

**Keywords:** RC association models, contingency table, Bayesian inference, procrustes, biplot, Metroplis-Hastings, Reversible Jump MCMC.



# LISTA DE ILUSTRAÇÕES

---

---

Figura 1 – Exemplo gráfico de pontos antes da aplicação do procrustes e depois da aplicação do procrustes . . . . .	67
Figura 2 – Exemplo de um gráfico biplot das projeções da matriz U na matriz V . . . . .	70
Figura 3 – Gráfico exemplificativo de regiões de credibilidade construídos após a aplicação do procrustes . . . . .	72
Figura 4 – Representação gráfica das regiões de credibilidade de $\mu_{ij}$ e $v_{ij}$ . . . . .	77
Figura 5 – Representação gráfica das regiões de credibilidade de $\mu_{ij}$ e $v_{ij}$ com segmentos de retas . . . . .	77
Figura 6 – Biplot para a amostra . . . . .	78
Figura 7 – Biplot para a amostra . . . . .	78
Figura 8 – Representação gráfica das regiões de credibilidades de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado . . . . .	80
Figura 9 – Representação gráfica do biplot para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado comparativo . . . . .	81
Figura 10 – Representação gráfica do biplot para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado comparativo . . . . .	82
Figura 11 – Proporções simuladas para o valor de K - Prova de linguagens . . . . .	88
Figura 12 – Representação gráfica das estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado do desempenho na prova de linguagens . . . . .	89
Figura 13 – Representação do biplot para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado do desempenho na prova de linguagens. . . . .	90
Figura 14 – Representação do biplot na forma de vetores para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado sobre os dados do desempenho na prova de linguagens. . . . .	91
Figura 15 – Proporções simuladas para o valor de K - Prova de matemática . . . . .	93
Figura 16 – Representação gráfica das estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado do desempenho na prova de matemática. . . . .	94
Figura 17 – Representação do biplot para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado no desempenho na prova matemática. . . . .	95
Figura 18 – Representação do biplot na forma de vetores para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado sobre os dados do desempenho na prova de matemática. . . . .	96
Figura 19 – Proporções simuladas para o valor de K - Prova de redação . . . . .	98

Figura 20 – Representação gráfica das estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado do desempenho na prova de Redação . . . . .	99
Figura 21 – Representação do biplot para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado no desempenho na prova de Redação. . . . .	100
Figura 22 – Representação do biplot na forma de vetores para as estimativas de $\mu_{ij}$ e $v_{ij}$ para o modelo estimado sobre os dados do desempenho na prova de redação. . . . .	101
Figura 23 – Trace plot dos resultados do MH aplicados na amostra simulada para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ e $\lambda_3^R$ . . . . .	123
Figura 24 – Trace plot dos resultados do MH aplicados na amostra simulada para $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ e $\mu_{21}$ . . . . .	124
Figura 25 – Trace plot dos resultados do MH aplicados na amostra simulada para $\mu_{32}$ , $\mu_{42}$ , $v_{11}$ e $v_{21}$ . . . . .	125
Figura 26 – Trace plot dos resultados do MH aplicados na amostra simulada para $v_{22}$ , $v_{32}$ , $v_{42}$ e $v_{52}$ . . . . .	126
Figura 27 – Trace plot dos resultados do MH aplicados na amostra simulada para o log-verossimilhança . . . . .	127
Figura 28 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ e $\lambda_3^R$ . . . . .	129
Figura 29 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ e $\mu_{21}$ . . . . .	130
Figura 30 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para $\mu_{32}$ , $\mu_{42}$ , $v_{11}$ e $v_{12}$ . . . . .	131
Figura 31 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para $v_{22}$ , $v_{23}$ , $v_{24}$ e $v_{25}$ . . . . .	132
Figura 32 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para log-verossimilhança . . . . .	133
Figura 33 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_1^C$ , $\lambda_2^C$ e $\lambda_3^C$ . . . . .	135
Figura 34 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ , $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ , $\mu_{12}$ e $\mu_{22}$ . . . . .	136
Figura 35 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para $\mu_{32}$ , $\mu_{42}$ , $v_{11}$ , $v_{21}$ , $v_{31}$ , $v_{41}$ , $v_{51}$ e $v_{12}$ . . . . .	137
Figura 36 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para $v_{22}$ , $v_{32}$ , $v_{42}$ , $v_{52}$ e log-verossimilhança . . . . .	138
Figura 37 – Distribuições <i>a posteriori</i> do MH aplicados aos dados da amostra simulada para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_1^C$ , $\lambda_2^C$ e $\lambda_3^C$ . . . . .	140
Figura 38 – Distribuições <i>a posteriori</i> do MH aplicados aos dados da amostra simulada para $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ , $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ , $\mu_{12}$ e $\mu_{22}$ . . . . .	141

Figura 39 – Distribuições <i>a posteriori</i> do MH aplicados aos dados da amostra simulada para $\mu_{32}, \mu_{42}, v_{11}, v_{21}, v_{31}, v_{41}, v_{51}$ e $v_{12}$ . . . . .	142
Figura 40 – Distribuições <i>a posteriori</i> do MH aplicados aos dados da amostra simulada para log-verossimilhança . . . . .	143
Figura 41 – Trace plot dos resultados do MH aplicados no estudo comparativo para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	145
Figura 42 – Trace plot dos resultados do MH aplicados no estudo comparativo para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	146
Figura 43 – Trace plot dos resultados do MH aplicados no estudo comparativo para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	147
Figura 44 – Trace plot dos resultados do MH aplicados no estudo comparativo para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	148
Figura 45 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para $\lambda, \lambda_1^R, \lambda_2^R$ e $\lambda_3^R$ . . . . .	150
Figura 46 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para $\lambda_4^R, \lambda_5^R, \lambda_1^C, \lambda_2^C, \lambda_3^C, \lambda_4^C, \lambda_5^C$ e $\mu_{11}$ . . . . .	151
Figura 47 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$ e $\mu_{42}$ . . . . .	152
Figura 48 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para $\mu_{52}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{21}$ e $v_{22}$ . . . . .	153
Figura 49 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para $v_{23}, v_{24}, v_{25}$ e log-verossimilhança . . . . .	154
Figura 50 – Distribuições <i>a posteriori</i> do MH aplicados aos dados do estudo comparativo para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	156
Figura 51 – Distribuições <i>a posteriori</i> do MH aplicados aos dados do estudo comparativo para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	157
Figura 52 – Distribuições <i>a posteriori</i> do MH aplicados aos dados do estudo comparativo para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	158
Figura 53 – Distribuições <i>a posteriori</i> do MH aplicados aos dados do estudo comparativo para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	159
Figura 54 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	161
Figura 55 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	162
Figura 56 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	163
Figura 57 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	164

Figura 58 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_5^R$ , $\lambda_1^C$ e $\lambda_2^C$ . . . . .	166
Figura 59 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para $\lambda_3^C$ , $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ , $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ e $\mu_{51}$ . . . . .	167
Figura 60 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para $\mu_{12}$ , $\mu_{22}$ , $\mu_{32}$ , $\mu_{42}$ , $\mu_{52}$ , $v_{11}$ , $v_{21}$ e $v_{31}$ . . . . .	168
Figura 61 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para $v_{41}$ , $v_{51}$ , $v_{12}$ , $v_{22}$ , $v_{32}$ , $v_{42}$ , $v_{52}$ e log-verossimilhança . . . . .	169
Figura 62 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ e $\lambda_3^R$ . . . . .	171
Figura 63 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para $\lambda_4^R$ , $\lambda_5^R$ , $\lambda_1^C$ , $\lambda_2^C$ , $\lambda_3^C$ , $\lambda_4^C$ , $\lambda_5^C$ e $\mu_{11}$ . . . . .	172
Figura 64 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ , $\mu_{51}$ , $\mu_{12}$ , $\mu_{22}$ , $\mu_{32}$ e $\mu_{42}$ . . . . .	173
Figura 65 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para $\mu_{52}$ , $v_{11}$ , $v_{12}$ , $v_{13}$ , $v_{14}$ , $v_{15}$ , $v_{21}$ e $v_{22}$ . . . . .	174
Figura 66 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para $v_{23}$ , $v_{24}$ , $v_{25}$ e log-verossimilhança . . . . .	175
Figura 67 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de linguagens para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_5^R$ , $\lambda_1^C$ e $\lambda_2^C$ . . . . .	176
Figura 68 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de linguagens para $\lambda_3^C$ , $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ , $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ e $\mu_{51}$ . . . . .	177
Figura 69 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de linguagens para $\mu_{12}$ , $\mu_{22}$ , $\mu_{32}$ , $\mu_{42}$ , $\mu_{52}$ , $v_{11}$ , $v_{21}$ e $v_{31}$ . . . . .	178
Figura 70 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de linguagens para $v_{41}$ , $v_{51}$ , $v_{12}$ , $v_{22}$ , $v_{32}$ , $v_{42}$ , $v_{52}$ e log-verossimilhança . . . . .	179
Figura 71 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_5^R$ , $\lambda_1^C$ e $\lambda_2^C$ . . . . .	180
Figura 72 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para $\lambda_3^C$ , $\lambda_4^C$ , $\lambda_5^C$ , $\mu_{11}$ , $\mu_{21}$ , $\mu_{31}$ , $\mu_{41}$ e $\mu_{51}$ . . . . .	181
Figura 73 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para $\mu_{12}$ , $\mu_{22}$ , $\mu_{32}$ , $\mu_{42}$ , $\mu_{52}$ , $v_{11}$ , $v_{21}$ e $v_{31}$ . . . . .	182
Figura 74 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para $v_{41}$ , $v_{51}$ , $v_{12}$ , $v_{22}$ , $v_{32}$ , $v_{42}$ , $v_{52}$ e log-verossimilhança . . . . .	183
Figura 75 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para $\lambda$ , $\lambda_1^R$ , $\lambda_2^R$ , $\lambda_3^R$ , $\lambda_4^R$ , $\lambda_5^R$ , $\lambda_1^C$ e $\lambda_2^C$ . . . . .	185

Figura 76 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	186
Figura 77 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	187
Figura 78 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	188
Figura 79 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para $\lambda, \lambda_1^R, \lambda_2^R$ e $\lambda_3^R$ . . . . .	190
Figura 80 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para $\lambda_4^R, \lambda_5^R, \lambda_1^C, \lambda_2^C, \lambda_3^C, \lambda_4^C, \lambda_5^C$ e $\mu_{11}$	191
Figura 81 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$ e $\mu_{42}$ . . . . .	192
Figura 82 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para $\mu_{52}, v_{11}, v_{12}, v_{13}$ . . . . .	193
Figura 83 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para $v_{23}, v_{24}, v_{25}$ e log-verossimilhança	194
Figura 84 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de matemática para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	195
Figura 85 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de matemática para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	196
Figura 86 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de matemática para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	197
Figura 87 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de matemática para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	198
Figura 88 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	199
Figura 89 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	200
Figura 90 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	201
Figura 91 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	202
Figura 92 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	204
Figura 93 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}$ . . . . .	205
Figura 94 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	206

Figura 95 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	207
Figura 96 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para $\lambda, \lambda_1^R, \lambda_2^R$ e $\lambda_3^R$ . . . . .	209
Figura 97 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para $\lambda_4^R, \lambda_5^R, \lambda_1^C, \lambda_2^C, \lambda_3^C, \lambda_4^C, \lambda_5^C$ e $\mu_{11}$	210
Figura 98 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$ e $\mu_{42}$	211
Figura 99 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para $\mu_{52}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{21}$ e $v_{22}$	212
Figura 100 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para $v_{23}, v_{24}, v_{25}$ e log-verossimilhança	213
Figura 101 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de redação para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	214
Figura 102 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de redação para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	215
Figura 103 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de redação para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	216
Figura 104 – Distribuições <i>a posteriori</i> do MH aplicados no desempenho na prova de redação para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	217
Figura 105 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para $\lambda, \lambda_1^R, \lambda_2^R, \lambda_3^R, \lambda_4^R, \lambda_5^R, \lambda_1^C$ e $\lambda_2^C$ . . . . .	218
Figura 106 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$ e $\mu_{51}$ . . . . .	219
Figura 107 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para $\mu_{12}, \mu_{22}, \mu_{32}, \mu_{42}, \mu_{52}, v_{11}, v_{21}$ e $v_{31}$ . . . . .	220
Figura 108 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para $v_{41}, v_{51}, v_{12}, v_{22}, v_{32}, v_{42}, v_{52}$ e log-verossimilhança . . . . .	221

# LISTA DE QUADROS

---

---

Quadro 1 – Descrição das Áreas de Conhecimento e Componentes Curriculares do Enem	84
Quadro 2 – Critérios para avaliação do desempenho na prova de redação no Enem 2023.	85
Quadro 3 – Quantidade de candidatos selecionados para aplicação da metodologia com dados reais por área de conhecimento. . . . .	87
Quadro 4 – Competências da área 1 - Linguagens, Códigos e suas Tecnologias. . . . .	223
Quadro 5 – Níveis de desempenho que são utilizados para avaliar a competência II nas redações do Enem 2023. . . . .	224
Quadro 6 – Níveis de desempenho que serão utilizados para avaliar a Competência III nas redações do Enem 2023. . . . .	224
Quadro 7 – Níveis de desempenho que serão utilizados para avaliar a Competência IV nas redações do Enem 2023. . . . .	225
Quadro 8 – Níveis de desempenho que serão utilizados para avaliar a Competência V nas redações do Enem 2023. . . . .	225
Quadro 9 – Quantidade de municípios selecionados por Undiade da Federação. . . . .	227
Quadro 10 – Municípios selecionados por Unidade da Federação. . . . .	228



# LISTA DE TABELAS

---

---

Tabela 1 – Exemplo de uma tabela de contingência 3x4 . . . . .	32
Tabela 2 – Tipos de modelos de associação . . . . .	52
Tabela 3 – Graus de liberdade por tipo de modelo de associação RC(K) . . . . .	53
Tabela 4 – Valores dos parâmetros utilizados para geração da tabela de contingência . . . . .	74
Tabela 5 – Tabela de contingência simulada para validar o código computacional . . . . .	74
Tabela 6 – Resultado da comparação entre as estimativas obtidas por metodologias distintas . . . . .	75
Tabela 7 – Visitas dos modelo para estimar K . . . . .	76
Tabela 8 – Classificação de adultos de acordo com a faixa etária e escolaridade . . . . .	79
Tabela 9 – Resultado da comparação entre as estimativas obtidas por metodologias distintas . . . . .	80
Tabela 10 – Desempenho dos candidatos do Enem 2023 na prova de linguagens segundo a escolaridade da mãe. . . . .	87
Tabela 11 – Parâmetros do modelo de associação referentes às notas da prova de linguagens estimados via MH. . . . .	88
Tabela 12 – Desempenho dos candidatos do Enem 2023 na prova de matemática segundo a escolaridade do pai. . . . .	92
Tabela 13 – Parâmetros do modelo de associação referentes às notas da prova de matemática estimados via MH. . . . .	92
Tabela 14 – Desempenho dos candidatos do Enem 2023 na prova de redação segundo a escolaridade da mãe. . . . .	97
Tabela 15 – Parâmetros do modelos de associação para Redação estimadas via MH. . . . .	97
Tabela 16 – Exemplo de transformação de tabela de contingência de tripla entrada para dupla entrada. . . . .	119
Tabela 17 – Comparativo dos valores dos parâmetros reais com os intervalos de credibilidade dos valores estimados utilizando dados da tabela de continguência simulada utilizando o MH . . . . .	121
Tabela 18 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados aos dados da amostra simulada . . . . .	128
Tabela 19 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados aos dados do artigo comparativo . . . . .	149
Tabela 20 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens . . . . .	170
Tabela 21 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática . . . . .	189

Tabela 22 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação . . . . .	208
---	-----

# SUMÁRIO

---

---

1	INTRODUÇÃO . . . . .	31
1.1	Tabela de contingência . . . . .	31
1.2	Estudo da associação em tabelas de contingência . . . . .	32
1.2.1	<i>Associação entre variáveis categóricas</i> . . . . .	32
1.2.2	<i>Medidas de associação</i> . . . . .	33
1.3	Aplicações . . . . .	36
1.4	Objetivos . . . . .	36
2	REVISÃO . . . . .	39
2.1	<i>Odds ratio</i> . . . . .	39
2.1.1	<i>Odds ratio em tabelas de contingência <math>2 \times 2</math></i> . . . . .	40
2.1.2	<i>Odds ratios em tabelas (<math>I \times J</math>)</i> . . . . .	41
2.2	Decomposição em valores singulares . . . . .	41
2.3	Biplot . . . . .	42
2.4	Procrustes . . . . .	43
2.4.1	<i>Descrição da metodologia do procrustes</i> . . . . .	44
2.5	Identificabilidade de parâmetros . . . . .	45
2.6	Simulação . . . . .	46
2.6.1	<i>Algoritmo Metropolis-Hastings</i> . . . . .	47
2.6.2	<i>Método reversible jump MCMC</i> . . . . .	47
2.6.3	<i>Avaliação da convergência no método MCMC</i> . . . . .	49
3	MODELO DE ASSOCIAÇÃO RC . . . . .	51
3.1	Definições . . . . .	51
3.1.1	<i>Definição de tabela de contingência</i> . . . . .	51
3.2	Primeiros modelos de associação . . . . .	52
3.3	Modelo de associação RC . . . . .	54
3.3.1	<i>Representação matricial do modelo de associação RC</i> . . . . .	55
3.3.2	<i>SVD e o modelo de associação</i> . . . . .	56
3.4	Identificabilidade no modelo de associação RC . . . . .	56
3.5	Outras interpretações do modelo de associação . . . . .	57
3.6	Concordâncias entre o modelo de associação e o <i>odds ratio</i> com biplot . . . . .	57

<b>3.7</b>	<b>Estimação</b> . . . . .	<b>59</b>
<b>3.7.1</b>	<b><i>Função de verossimilhança</i></b> . . . . .	<b>59</b>
<b>3.7.2</b>	<b><i>Abordagem Bayesiana</i></b> . . . . .	<b>60</b>
3.7.2.1	<i>Distribuições a priori do modelo de associação RC</i> . . . . .	60
3.7.2.2	<i>A distribuição a posteriori no modelo de associação RC</i> . . . . .	61
<b>3.8</b>	<b>Simulação</b> . . . . .	<b>61</b>
<b>3.8.1</b>	<b><i>Metropolis-Hastings no modelo de associação RC</i></b> . . . . .	<b>62</b>
<b>3.8.2</b>	<b><i>Reversible Jump MCMC no modelo de associação</i></b> . . . . .	<b>63</b>
<b>3.8.3</b>	<b><i>Reversible Jump MCMC na seleção do melhor valor de K</i></b> . . . . .	<b>64</b>
<b>3.8.4</b>	<b><i>Aplicação do procrustes no modelo de associação</i></b> . . . . .	<b>65</b>
3.8.4.1	<i>Análise do procrustes no modelo de associação</i> . . . . .	66
3.8.4.2	<i>Movimento de rotação no procrustes</i> . . . . .	67
3.8.4.3	<i>Exemplo gráfico da aplicação do procrustes</i> . . . . .	67
<b>3.8.5</b>	<b><i>Aplicação do biplot no modelo de associação</i></b> . . . . .	<b>68</b>
<b>3.8.6</b>	<b><i>Exemplo de um biplot</i></b> . . . . .	<b>68</b>
<b>3.8.7</b>	<b><i>Regiões de credibilidade no modelo de associação</i></b> . . . . .	<b>71</b>
<b>4</b>	<b>SELEÇÃO DE MODELOS</b> . . . . .	<b>73</b>
<b>4.1</b>	<b>Aplicação</b> . . . . .	<b>73</b>
<b>4.1.1</b>	<b><i>Comparação do desempenho do método proposto com uma metodologia conhecida (um estudo de caso)</i></b> . . . . .	<b>79</b>
<b>4.2</b>	<b>Aplicação em dados reais</b> . . . . .	<b>82</b>
<b>4.2.1</b>	<b><i>Exame Nacional do Ensino Médio - Enem - contextualização</i></b> . . . . .	<b>83</b>
<b>4.2.2</b>	<b><i>Microdados do Enem 2023</i></b> . . . . .	<b>84</b>
4.2.2.1	<i>Uso da TRI para calcular as notas do Enem</i> . . . . .	84
4.2.2.2	<i>Nota da prova discursiva (redação) no Enem</i> . . . . .	85
4.2.2.3	<i>Preparação da base de dados e das tabelas de contingências</i> . . . . .	86
<b>4.2.3</b>	<b><i>Aplicação 1: Desempenho na prova de linguagens.</i></b> . . . . .	<b>87</b>
<b>4.2.4</b>	<b><i>Procrustes das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de linguagens.</i></b> . . . . .	<b>89</b>
<b>4.2.5</b>	<b><i>Biplot das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de linguagens.</i></b> . . . . .	<b>89</b>
<b>4.2.6</b>	<b><i>Aplicação 2: Desempenho na prova de matemática.</i></b> . . . . .	<b>91</b>
<b>4.2.7</b>	<b><i>Procrustes das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de matemática.</i></b> . . . . .	<b>93</b>
<b>4.2.8</b>	<b><i>Biplot das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de matemática.</i></b> . . . . .	<b>94</b>
<b>4.2.9</b>	<b><i>Aplicação 3: Desempenho na prova de redação.</i></b> . . . . .	<b>96</b>
<b>4.2.10</b>	<b><i>Procrustes das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de redação.</i></b> . . . . .	<b>98</b>

4.2.11	<i>Biplot das estimativas de <math>\mu_{ij}</math> e <math>v_{ij}</math> para as notas da prova de redação.</i>	99
5	CONCLUSÃO	103
5.1	Conclusões do trabalho	103
	REFERÊNCIAS	105
APÊNDICE A	CÁLCULO DA MATRIZ JACOBIANO PARA O REVERSIBLE JUMP	111
A.1	Um exemplo - Movimento de $M_1$ para $M_2$ - Matriz $2 \times 2$ para $3 \times 2$ .	112
A.2	Outro exemplo - Movimento de $M_1$ para $M_2$ - Matriz $2 \times 2$ para $2 \times 3$ .	115
APÊNDICE B	COMBINAÇÃO DE DOIS NÍVEIS DE UMA TABELA DE TRIPLA ENTRADA	119
APÊNDICE C	TABELAS COMPARATIVAS	121
APÊNDICE D	VERIFICAÇÃO DAS CADEIAS	123
D.1	Gráficos trace plot dos resultados do MH aplicados aos dados da amostra simulada	123
D.2	Verificação da convergência através do método de Gelman-Rubin para os dados da amostra simulada	128
D.3	Gráficos de auto correlação dos resultados do MH aplicados aos dados da amostra simulada	134
D.4	Gráficos das distribuições <i>a posteriori</i> dos parâmetros do MH aplicados aos dados da amostra simulada	139
D.5	Gráficos trace plot dos resultados da cadeia aplicados aos dados do artigo comparativo	144
D.6	Verificação da convergência através do método de Gelman-Rubin para os dados do artigo comparativo	149
D.7	Gráficos das distribuições <i>a posteriori</i> dos resultados do MH aplicado aos dados do artigo comparativo	155
D.8	Gráficos de auto correlação dos resultados do MH aplicados nos dados do artigo comparativo	160
D.9	Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de linguagens	165
D.10	Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de linguagens	170
D.11	Gráficos das distribuições <i>a posteriori</i> dos resultados do MH aplicado no desempenho na prova de linguagens	176

D.12	Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de linguagens . . . . .	180
D.13	Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de matemática . . . . .	184
D.14	Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de matemática . . . . .	189
D.15	Gráficos das distribuições <i>a posteriori</i> dos resultados do MH aplicado no desempenho na prova de matemática . . . . .	195
D.16	Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de matemática . . . . .	199
D.17	Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de redação. . . . .	203
D.18	Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de redação. . . . .	208
D.19	Gráficos das distribuições <i>a posteriori</i> dos resultados do MH aplicado no desempenho na prova de redação . . . . .	214
D.20	Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de redação . . . . .	218
APÊNDICE E	COMPETÊNCIAS AVALIADAS NAS PROVAS DO ENEM . . . . .	223
E.1	Competências avaliadas no Enem para prova de linguagens e suas tecnologias. . . . .	223
E.2	Competências avaliadas no Enem para prova de Redação. . . . .	223
APÊNDICE F	DESCRIÇÃO E QUANTIDADE DE MUNICÍPIOS SELECIONADOS PARA APLICAÇÃO EM DADOS REAIS	227

---

# INTRODUÇÃO

---

O uso do termo medida de associação ou não associação como uma medida da relação entre dois atributos pode ser encontrado no trabalho de [Yule \(1900\)](#) que propõe um método para quantificar a associação de variáveis categóricas. O uso de tabelas de contingência para analisar variáveis categóricas ou quantitativas discretizadas ou agrupadas em intervalos, é um procedimento comum em diversas áreas do conhecimento tais como, ciências sociais, biológicas e da saúde, na área da educação, psicologia entre outras e o modelo de associação RC pode ser utilizado em todos esses casos.

## 1.1 Tabela de contingência

Uma tabela de contingência é caracterizada pela disposição de variáveis categóricas ou quantitativas categorizadas (com ou sem ordenamento natural) em linhas e colunas. As variáveis são representadas por categorias em uma tabela de contingência. As tabelas podem ser de dupla entrada ou bidimensionais (two-way, na terminologia inglesa), tripla ou tridimensional (three-way, no idioma inglês) ou mais entradas. A dimensão de uma tabela de contingência é determinada pela quantidade de categorias de cada variável disposta em linhas e colunas, geralmente determinadas pelo pesquisador, de forma arbitrária e está diretamente ligada ao objeto de interesse do estudo. A característica mais usual de uma tabela de contingência é a exibição da síntese dos resultados observados de um experimento.

O termo tabela de contingência foi utilizado pela primeira vez por [Pearson \(1904\)](#) ao definir uma medida de contingência e propor a expressão do coeficiente de contingência  $\sqrt{\phi^2/(1+\phi^2)}$  como uma medida ou quantificação da relação entre duas variáveis qualitativas. A quantidade  $\phi^2$  se refere ao valor  $\chi^2/N$  em que,  $\chi^2 = \sum_n (O - E)^2/E$ , com  $n=1,2,\dots,N$ , onde  $N$  é o total de observações,  $O$  é a frequência observada e  $E$  é a frequência esperada de uma tabela de contingência.

Como exemplo de uma tabela de contingência de dupla entrada, considere A e B duas variáveis categóricas com I linhas e J colunas (dimensão  $I \times J$ ), respectivamente. Seja  $Y_{ij}$  uma variável aleatória de uma distribuição  $\mathcal{D}$  que representa a contagem das ocorrências em cada célula da tabela de contingência e seja  $y_{ij}$  a realização dessa variável aleatória. As somas de cada linha e cada coluna dessa tabela representam distribuições marginais e são representadas por  $Y_{i+} = \sum_{j=1}^J Y_{ij}$  e  $Y_{+j} = \sum_{i=1}^I Y_{ij}$  respectivamente, a quantidade ou número de ocorrências total da tabela é representada por  $N = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}$ .

A [Tabela 1](#) ilustra um exemplo de uma tabela de contingência de dupla entrada de dimensão  $3 \times 4$  (I=3 e J=4) onde cada contagem  $y_{ij}$  indica a quantidade observada resultante do cruzamento da linha  $i$  com a coluna  $j$  da tabela.

Tabela 1 – Exemplo de uma tabela de contingência 3x4

Variável B	Variável A				Total
	Categoria 1	Categoria 2	Categoria 3	Categoria 4	
Categoria 1	$y_{11}$	$y_{12}$	$y_{13}$	$y_{14}$	$y_{1+}$
Categoria 2	$y_{21}$	$y_{22}$	$y_{23}$	$y_{24}$	$y_{2+}$
Categoria 3	$y_{31}$	$y_{32}$	$y_{33}$	$y_{34}$	$y_{3+}$
Total	$y_{+1}$	$y_{+2}$	$y_{+3}$	$y_{+4}$	$n$

Fonte: Elaborada pelo autor.

Embora grande parte dos estudos que envolvam tabelas de contingência seja mais comumente em tabelas de dupla entrada, ainda assim é possível fazer a combinação de dois níveis de uma tabela de tripla entrada e torná-la bidimensional. Um exemplo dessa transformação pode ser encontrado no [Apêndice B](#).

## 1.2 Estudo da associação em tabelas de contingência

O uso da medida de associação como uma medida da relação entre categorias de variáveis dispostas numa tabela de contingência é uma prática desde os anos 1900 e a aplicação em estudos reais como aplicações em áreas como medicina, sociologia, engenharia, psicologia entre outros.

### 1.2.1 Associação entre variáveis categóricas

Medir a associação entre duas variáveis categóricas é o objeto de estudo de muitos pesquisadores que trabalham com experimentos ou utilizam tabelas de contingência para extrair informações de ensaios ou estudos laboratoriais.

Diversos autores tem publicados textos sobre métodos e metodologias para medir a associação de variáveis categóricas. Dentre os diversos encontrados na literatura, Agresti tem sido um dos autores com uma variedade de artigos sobre o tema. [Agresti \(1984\)](#) descreve

algumas medidas de associação para tabelas de contingências baseadas em exemplos. Algumas das metodologias tratadas pelo autor são a estatística qui-quadrado de Pearson, o teste da razão de verossimilhança (comumente denominado estatística G2) e o *odds ratio*.

### 1.2.2 Medidas de associação

A medida de associação entre linhas e colunas de uma tabela de contingência pode ser estimada através do modelo de associação RC, de ordem  $K$ , ou RC( $K$ ), introduzida por Goodman (1979a, 1981) utilizando procedimentos iterativos e métodos Bayesianos para a obtenção dos valores das estimativas dos parâmetros do modelo.

O valor de  $K$  representa os níveis dos efeitos das linhas e/ou colunas do modelo de associação. Na ótica matemática,  $K$  é o número de termos bilineares no modelo de associação.

Quando duas variáveis categóricas não são independentes, se torna interessante conhecer melhor a estrutura de associação entre essas variáveis, ou seja, conhecer como as categorias dessas variáveis se relacionam. Nesse caso, a medida de associação é a metodologia mais indicada para estudar a força de uma associação, quais categorias das variáveis estão envolvidas nessa associação e, no caso de variáveis ordinais, a direção da associação pode ser também uma característica de interesse nesse estudo.

Em tabelas de contingência de dimensão  $2 \times 2$  a associação é positiva ou negativa. À medida em que a dimensão da tabela vai aumentando para dimensões ( $I \times J$ ,  $I > 2$  e/ou  $J > 2$ ), a estrutura do modelo de associação vai se tornando mais complexa, porque o número de parâmetros do modelo aumenta. Por exemplo, na tabela de contingência de dimensão  $2 \times 2$ , a ocorrência do aumento de uma unidade em uma variável dessa tabela, implicará também um aumento de uma unidade na outra variável associada da tabela. Se a associação é negativa, a ocorrência do aumento de uma unidade em uma variável, implicará na redução de uma unidade na outra variável associada, ou vice-versa. Já para tabelas maiores a associação pode ser toda positiva (ou negativa) ou iniciar positiva e trocar de sinal para negativo (e vice versa), daí surge a necessidade de uma estrutura (conjunto de variáveis) de modelo mais complexa. A título de exemplo, podemos imaginar a aplicação de herbicidas na lavoura de laranja para combater pragas que trazem ferrugem para a fruta. Quanto maior a dose de herbicida mais eficiente será, mas pode atingir um ponto em que a fruta poderá ser afetada diretamente pelo composto aplicado e afetando a quantidade de frutos saudáveis para colheita. Nesse exemplo, podemos ter inicialmente uma associação positiva para o efeito do herbicida e terminar como negativa, já que afetou o fruto também. Essas hipóteses podem ser verificadas através do modelo de associação.

Um estudo com extensa discussão sobre algumas das principais formas de medir associação foi apresentado por Yule (1912) ao propor uma modificação de um coeficiente já descrito cerca de doze anos antes (coeficiente de contingência) para testar a presença ou ausência de associação entre dois caracteres ou atributos em tabelas de contingência. No artigo, além de

abordar o cálculo de um coeficiente de associação, o autor traz vários exemplos aplicados e comparações em termos de associação de diferentes tabelas utilizando diferentes meios de aferir proporções, argumentos que justificam o uso de uma ou outra equação para medir associação, além de um discurso sobre coeficiente de correlação em tabelas de contingência.

[Goodman e Kruskal \(1954\)](#) produziram os primeiros artigos na tentativa de medir associação em tabelas de contingência ao apresentarem uma série de medidas que podem ser consideradas úteis em determinadas situações de estudo. A medida de associação seria uma probabilidade ou uma simples função de probabilidade dentro de algum modelo específico. As medidas de associação apresentadas não devem ser consideradas como únicas, universais, pois outras medidas de associação podem ser implementadas (por exemplo, associação positiva ou negativa, associação linear e não linear, homogeneidade e heterogeneidade), mas podem ser consideradas como uma sugestão de uso em circunstâncias apropriadas. Dentre as medidas apresentadas, estão a medida baseada na predição ideal de ordem (nesse caso a medida não muda se as linhas ou colunas permutarem), medida baseada em modelo com função de perda e baseada em modelo de confiabilidade, medida baseada na predição proporcional, medida de associação parcial e medida de múltipla associação.

Cerca de meia década depois, [Goodman e Kruskal \(1959\)](#) rediscutiram o trabalho anterior com observações complementares, inclusão de uma variedade de novos registros bibliográficos e de novas medidas, como a associação baseada na estrutura latente e novas expressões para algumas das medidas de associação apresentadas anteriormente.

No estudo sobre medida de associação em tabela de dimensão  $2 \times 2$ , que tratou de um aspecto bem restrito do problema da medição de associação em uma tabela onde os atributos são emparelhados e os totais marginais não são fixos, foram propostas várias funções na tentativa de medir a associação ([EDWARDS, 1963](#)).

Dois métodos para análise de tabelas de contingência onde as categorias das respostas são ordenadas e permitem o uso de regressão ponderada foram apresentados por [Williams e Grizzle](#). Um dos métodos é similar ao uso de testes de classificação nos casos em que há muitos empates, com base em ranks obtidos das informações contidas nas respostas ordenadas, enquanto que o outro método é baseado em escala Thurstoniano (que mede a força das opiniões e atitudes). Nos dois métodos os resultados podem ser apresentados na forma de análise de variância ([WILLIAMS; GRIZZLE, 1972](#)).

Outro trabalho voltado para análise de associação em tabelas de contingência cujas colunas são representadas por categorias ordenadas foi publicado por [Simon \(1974\)](#). Para preservar a ordem das categorias, dois modelos foram elaborados fornecendo frequências estimadas das células e testes de hipóteses de igualdade dos parâmetros das linhas. O primeiro modelo foi do tipo log-linear e o segundo avalia o logit acumulado nas linhas. Os procedimentos propostos baseiam-se em reparametrizações da distribuição multinomial. Testes qui-quadrado para avaliar independência foram adotados.

Outros autores também trouxeram artigos relacionadas com a medida de associação em tabela de contingência, como [Mosteller \(1968\)](#). O texto chama a atenção para o fato de que uma primeira atitude da maioria dos cientistas sociais ao lidar com uma tabela de contingência é calcular a estatística qui-quadrado, às vezes é suficiente, outras vezes esclarecedor ou um desperdício, mas às vezes não avançam o suficiente.

Uma variedade de métodos para análise de tabelas de contingência, como um estudo com a utilização de modelos log-lineares onde as categorias (das linhas e colunas) são ordinais foi proposto por [Haberman \(1974\)](#). Foram apresentados métodos de seleção de modelos, cálculo de estimativas e testes da razão, ambos de máxima verossimilhança e cálculos de variâncias assintóticas. O trabalho ainda inclui restrições impostas na definição dos modelos log-lineares e uma classe de modelos que inclui modelos de independência, de equiprobabilidade e modelos logit, permitindo, assim, obter mais informações além de um simples teste qui-quadrado de independência.

Dentre outros trabalhos mais recentes que tratam de medida de associação, podemos citar [Forcina e Kateri \(2021\)](#) que propõe uma nova classe de modelos de associação RC inovando, principalmente, a permissão da escolha do tipo de logit (local, global, continuação) e da escala de interação (via medidas de divergência  $\phi$ ), combinando essas opções sob restrições de posto em matrizes de interações. [Espindiller \(2017\)](#) aborda em sua tese, a medição da associação em tabelas de contingência sob uma perspectiva da Teoria da Informação. Discute as limitações de medidas clássicas, tais como a razão de chances e seus problemas em amostras pequenas com zeros amostrais, propõe o uso das divergências  $\phi$  para generalizar estimadores, modelos e testes. Ainda, introduz medidas e testes para assimetria, como por exemplo, versões  $\phi$  – *generalizadas* dos testes de [McNemar \(1947\)](#) e [Bowker \(1948\)](#).

[Bartolucci e Forcina \(2020\)](#) implementaram um pacote no software R, Modelos RC Estendidos (ou Extended Row–Column Models, na terminologia inglesa). O objetivo do pacote é oferecer estimativas de máxima verossimilhança e ferramentas de diagnóstico com o intuito de avaliar a qualidade do ajuste. [Kateri \(2014b\)](#) trabalha a análise estatística de tabelas de contingência abordando os conceitos fundamentais, as medidas de associação, os testes de independência, modelos log-lineares e modelos de associação. Por fim, [Kateri \(2018\)](#) explora o uso das divergências- $\phi$  como base unificadora na análise de tabelas de contingência, mostrando que medidas tradicionais, como a divergência de Kullback–Leibler, são casos particulares dentro dessa família. A abordagem ainda permite definir estimadores, testes e medidas de associação mais gerais e robustos, além de discutir propriedades assintóticas e a interpretação estatística das novas medidas

## 1.3 Aplicações

Há várias possibilidades de aplicação da metodologia que estamos propondo nesse trabalho, como por exemplo avaliar a associação entre o tratamento de pacientes com alguma doença degenerativa e a evolução da doença, associação entre perfis sócio-demográficos e o desempenho de alunos em testes de avaliação educacional ou provas de acesso ao ensino superior, estudar a associação entre a aplicação de diferentes herbicidas e a evolução de plantas ou redução de pragas, estudar a relação entre nível sócio-cultural e participação em movimentos sociais.

Embora a aplicação do modelo de associação RC em tabelas de contingência seja mais comum em tabelas de dupla entrada (two-way) de dimensão igual ou superior dois, a metodologia proposta pode ser aplicada em tabelas de contingência de qualquer dimensão sem necessidade de adaptação, necessitando apenas combinar os níveis de entrada da linha ou da coluna da tabela de forma que fiquem representadas por duas entradas (vide exemplo apresentado na [Tabela 16](#) do [Apêndice B](#)).

São poucos os artigos encontrados na literatura que tratam da estimação dos parâmetros do modelo de associação RC utilizando a metodologia Bayesiana, principalmente com os métodos iterativos MCMC. Nesse sentido, identificamos uma oportunidade existente na metodologia na que se refere ao método de seleção do parâmetro de associação K utilizando o enfoque Bayesiano e o reversible jump via Monte Carlo Markov Chain - RJMCMC (na tradução livre para a língua portuguesa, saltos reversíveis via Monte Carlo e Cadeias de Markov). Essa oportunidade trouxe uma motivação e inspiração para o desenvolvimento do presente trabalho.

## 1.4 Objetivos

Os objetivos e principais contribuições dessa pesquisa abrangem: uma proposta de metodologia de seleção do valor de K (que será detalhado mais adiante) do modelo de associação RC(K) que possibilite mensurar ou quantificar a força da interação ou do cruzamento entre variáveis aleatórias dispostas em linhas e colunas de uma tabela de contingência, utilizando a inferência Bayesiana e o RJMCMC, a indicação da existência do agrupamento de linhas ou colunas da tabela de contingência utilizando intervalos de credibilidade que serão dispostos na forma de gráficos biplot. A proposta inclui também a utilização da ferramenta gráfica do procrustes, que conjuntamente com o biplot serão utilizados como instrumento auxiliares na interpretação dos resultados. Também iremos utilizar a decomposição em valores singulares (DVS) durante a aplicação do RJMCMC para construir o biplot.

A aplicação do método proposto se dará inicialmente em uma amostra simulada, com os valores dos parâmetros conhecidos para fins de calibração e avaliação da performance do método. Posteriormente, aplicaremos o método em uma tabela de contingência existente em artigos que aplicaram a técnica para fins comparativos de resultados. Por fim, aplicaremos a técnica proposta

em três bancos de dados reais relacionados à área educacional.

Este trabalho está dividido em 5 capítulos. No [Capítulo 2](#) faremos uma revisão dos métodos necessários para atingir os objetivos propostos, como revisão de *odds ratio*, decomposição em valores singulares, as ferramentas gráficas do procrustes e do biplot, incluindo exemplos gráficos. Também fazem parte dessa seção uma descrição sobre intervalos de credibilidade, os métodos MCMC, Metropolis-Hastings e reversible jump. No [Capítulo 3](#) definimos uma tabela de contingência associada a uma distribuição de probabilidade, um breve histórico sobre modelo de associação de Goodman, dissertamos sobre a questão da identificabilidade de parâmetros e a representação gráfica no modelo de associação. Na seção de estimação exploramos a abordagem Bayesiana e a construção das distribuições *a priori* e *a posteriori*. Descrevemos a aplicação dos métodos MCMC no modelo de associação bem como a aplicação das técnicas Metropolis-Hastings e reversible jump. Finalizamos comentando sobre as dificuldades existentes na elaboração do biplot devido à oscilação dos eixos, a interpretação das nuvens de pontos e indicação de agrupamentos. No [Capítulo 4](#) exploramos a seleção de modelos, agrupamento e desagrupamento de categorias e aplicação em dados reais. A conclusão será descrita no [Capítulo 5](#). O trabalho finaliza com as referências bibliográficas e os apêndices.



---

## REVISÃO

---

Nesse capítulo iremos fazer uma breve revisão de alguns métodos e recursos que servirão de suporte para a aplicação da metodologia que estamos propondo.

### 2.1 *Odds ratio*

Nessa seção iremos introduzir o conceito de *odds ratio* que será necessário em algum momento para compreensão do tema. Leitores familiarizados com a temática, podem seguir para as seções posteriores.

Quando estamos diante de uma variável binária do tipo sucesso e fracasso, é comum a interpretação direta do resultado (em percentuais ou não) ou em termos de probabilidade de sucesso e fracasso. O *odds ratio*, ou razão de chances (traduzido para o português) é uma medida que avalia a importância relativa do sucesso em relação ao fracasso.

Suponha uma variável binária qualquer e  $p$  e  $1-p$  as probabilidades de sucesso e fracasso, respectivamente, dessa variável. A chance (*odds*) de sucesso é

$$odds(p) = \frac{p}{1-p}, \quad (2.1)$$

em que  $p \in (0, 1)$ .

Como as probabilidades são positivas, no caso de  $odds(p) > 1$ , implica que a probabilidade de sucesso é maior que fracasso, por exemplo, se  $p=0.8$ , então  $odds(p) = 0.8/0.2=4$ , ou seja, a chance de sucesso é 4 vezes maior que de fracasso, por outro lado se  $odds(p)=0.2$ ,  $odds(p)=0.2/0.8=1/4$ , o fracasso é 4 vezes mais provável que um sucesso (AGRESTI, 2012).

Quando estamos de diante de duas populações independentes e o interesse é comparar a resposta, como em estudos que envolvem dois tratamentos, ou estudos de caso e controle, por exemplo, as suas probabilidades de sucesso podem ser comparadas através da razão de

probabilidades, ou *odds ratio* (também popularmente conhecida como *razão de chances* (KATERI, 2014b)). Seja  $\theta$  uma medida de *odds ratio* e  $p_1$  e  $p_2$  as probabilidades de sucesso de duas variáveis binárias e cada uma de populações independentes e suas respectivas probabilidades de fracasso  $(1-p_1)$  e  $(1-p_2)$ , então,

$$\text{odds ratio}(p_1, p_2) = \theta = \frac{\text{odds}(p_1)}{\text{odds}(p_2)} = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{p_1(1-p_2)}{p_2(1-p_1)}, \quad (2.2)$$

em que  $p_1$  e  $p_2 \in (0, 1)$ .

### 2.1.1 Odds ratio em tabelas de contingência $2 \times 2$

Em termos de distribuição conjunta de uma tabela de contingência  $2 \times 2$ , onde as probabilidades em cada uma das células é  $p_{ij}$ , uma definição equivalente para *odds* da linha  $i$  é dada por  $\text{odds}(p_i) = p_{i1}/p_{i2}$ , com  $i=j=1,2$  então  $\theta$  é (AGRESTI, 2013):

$$\theta = \frac{p_{11}/p_{12}}{p_{21}/p_{22}} = \frac{p_{11} p_{22}}{p_{12} p_{21}}. \quad (2.3)$$

Quando a medida de  $\theta=1$ , indica que as variáveis binárias da tabela são independentes, que é equivalente a dizer que  $p_1 = p_2$  na Equação 2.2, ou que as razões são iguais na Equação 2.3, para qualquer outro valor de  $\theta \neq 1$ , corresponde a uma associação positiva ou negativa.

Algumas propriedades sobre *odds ratio*:

- O valor de  $\theta$  é sempre positivo, sendo que para  $\theta = 1$  indica independência entre as variáveis de classificação binária;
- Quando  $1 < \theta < \infty$  significa que os indivíduos da linha 1 possuem maior probabilidade de sucesso que indivíduos da linha 2, isto implica em  $p_1 > p_2$ . Como exemplo, se  $\theta = 2$ , a chance de sucesso da linha 1 é 2 vezes que a chance de sucesso na linha 2, mas isso não significa que a probabilidade  $p_1 = 2 \times p_2$ , e sim um risco relativo igual a 2;
- Quanto mais distante  $\theta$  for de 1, em qualquer direção, mais forte será a associação entre as variáveis binárias;
- Ao inverter a ordem das linhas (ou colunas) da tabela de contingência, o valor de  $\theta$  será igual ao inverso do valor original;
- Às vezes é mais conveniente utilizar o  $\log(\theta)$  para inferência em tabelas de contingência. Nesse caso, independência corresponderá a  $\log(\theta) = 0$  e o  $\log(\theta)$  se mantém simétrico em relação à inversão do valor de linhas (ou colunas), ou seja, há uma inversão do sinal (de + para - ou de - para +), mas a força da associação permanece a mesma;

- Não há mudança de valor de  $\theta$  caso a orientação da tabela de contingência seja invertida (transposta);

O *odds ratio* é uma medida de associação para tabelas de contingência de dimensão  $2 \times 2$  e exerce um papel importante na elaboração e interpretação do modelo na análise em tabelas de contingência, tem ainda a característica de não depender das distribuições marginais das variáveis de classificação, sendo, portanto, uma boa medida da sua associação (KATERI, 2014b).

### 2.1.2 Odds ratios em tabelas ( $I \times J$ )

Diante de tabelas de contingência de dimensão ( $2 \times 2$ ) um único número representado por  $\theta$  tem condições de resumir a associação, mas para tabelas de dimensões maiores, ( $I, J > 2$ ), usualmente não é possível um único número  $\theta$  representar a associação sem que haja alguma perda de informação, nesse caso, seria necessário um conjunto de valores  $\theta$  para medir a associação entre categorias da tabela (AGRESTI, 2013).

Para tabelas de dimensão ( $I \times J$ ), é essencial a decomposição (que não é única) da tabela original em um conjunto de  $(I - 1)(J - 1)$  subtabelas de dimensão ( $2 \times 2$ ) e cada uma delas descreve as associações subjacentes, nesse caso, temos um subconjunto de *odds ratios* locais representado por:

$$\theta_{ij} = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i,j+1}\pi_{i+1,j}} \quad \text{com } i = 1, \dots, I-1 \quad j = 1, \dots, J-1 \quad (2.4)$$

## 2.2 Decomposição em valores singulares

A decomposição em valores singulares (em inglês, singular value decomposition, SVD) é uma ferramenta de análise numérica da álgebra linear com diversas aplicações que envolvem matrizes, sejam elas quadradas ou não, como por exemplo em soluções de sistemas de equações lineares, estudo de dependência entre linhas e colunas de uma matriz, avaliação do grau de singularidade de matrizes, determinação do posto de matriz entre outros. A SVD fatora ou decompõe uma matriz qualquer em três outras matrizes com propriedades importantes que torna sua exploração em diversas aplicações, principalmente em várias metodologias estatísticas.

Vamos adotar a seguinte definição: Seja  $\mathbf{Y}$  uma matriz real de dimensão  $I \times J$  e  $\mathbf{Y}^T$  sua transposta, então existem matrizes ortogonais  $\mathbf{U} = [u_1 | \dots | u_I] \in \mathbb{R}^{I \times I}$  e  $\mathbf{V} = [v_1 | \dots | v_J] \in \mathbb{R}^{J \times J}$  tais que  $\mathbf{U}^T \mathbf{Y} \mathbf{V} = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{I \times J}$ ,  $r = \min\{I, J\}$ , em que  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$ , portanto,  $\mathbf{Y} = \mathbf{U} \Sigma \mathbf{V}^T$ .

As matrizes  $\mathbf{U}^T$  e  $\mathbf{V}^T$  representam as matrizes transpostas  $\mathbf{U}$  e  $\mathbf{V}$ , respectivamente, além disso  $\mathbf{U}$  e  $\mathbf{V}$  são matrizes ortogonais. Os elementos  $\sigma_1, \sigma_2, \dots, \sigma_r$  da diagonal principal de  $\Sigma$  são chamados valores singulares de  $\mathbf{Y}$ , os elementos (colunas de  $\mathbf{U}$ )  $u_1, u_2, \dots, u_I$  são chamadas de

vetores singulares à esquerda de  $\mathbf{Y}$  e as linhas da matriz  $\mathbf{V}$ , isto é,  $v_1, v_2, \dots, v_J$  são chamados de vetores singulares à direita de  $\mathbf{Y}$ . A prova completa desse teorema pode ser encontrada em (GOLUB; VAN-LOAN, 2013).

A partir da definição acima, podemos descrever algumas das propriedades úteis da SVD,

- Se  $\mathbf{U}^T \mathbf{Y} \mathbf{V} = \Sigma$  é a SVD de  $\mathbf{Y} \in \mathbb{R}^{I \times J}$  e  $I \geq J$ , então para  $i=1, \dots, I$ ,  $\mathbf{Y} v_i = \sigma_i u_i$  e  $\mathbf{Y}^T u_i = \sigma_i v_i$ ;
- Se  $\mathbf{Y}$  tem  $r$  valores singulares positivos, então  $\text{rank}(\mathbf{Y}) = r$ .

A SVD é uma fatoração da matriz na forma do produto de duas matrizes intercaladas por um matriz retangular diagonal, ou seja, estamos nos referindo a um produto de três matrizes. Qualquer matriz pode ter uma SVD e uma das propriedades dessa decomposição é que esses valores singulares são únicos.

O fato das matrizes  $\mathbf{U}$  e  $\mathbf{V}$  serem ortogonais e a inversa de uma matriz ortogonal coincide com a sua transposta, isto é,  $\mathbf{U} \mathbf{U}^T = \mathbf{I}$  e  $\mathbf{V}^T \mathbf{V} = \mathbf{I}$  (com  $\mathbf{I}$  representando a matriz identidade) faz com que os vetores colunas formam um conjunto ortonormal em  $\mathbb{R}^n$ . Com esse resultado, o conjunto de vetores  $u_1, \dots, u_I$  e  $v_1, \dots, v_J$  são ortogonais ( $u_i^T u_j = 0$  para  $i \neq j$  e  $v_i^T v_j = 0$  para  $i \neq j$ , respectivamente).

## 2.3 Biplot

O biplot é uma técnica gráfica proposta por Gabriel (1971) que permite a representação gráfica de uma matriz de dados de dimensão  $I \times J$  através de pontos e vetores em um mesmo plano bidimensional. Entre as características do biplot destacamos a visualização gráfica da existência de possíveis agrupamentos de categorias dispostas em linhas ou colunas de uma tabela de contingência e também a indicação de possíveis interações entre essas categorias.

Qualquer matriz pode ser exibida como um biplot, desde que o posto dessa matriz seja igual ou superior a dois (ou seja, o número de linhas ou colunas linearmente independentes seja igual ou superior a dois), porém para matrizes de posto mais alto, a exibição não será exata, mas pode ser feita de forma aproximada utilizando uma matriz de posto dois que mais se aproxima da matriz original. O prefixo “bi” na terminologia biplot indica que é uma exibição conjunta de linhas e colunas de uma matriz (GABRIEL, 1980).

Um biplot exato de uma matriz só é possível se a matriz for de posto igual a dois, mas para uma matriz de posto superior a dois os seguintes passos podem ser aplicados para representar essa matriz em um biplot:

- O primeiro passo é aproximar a matriz  $\mathbf{Y}$  por uma matriz  $\mathbf{Y}'$  de posto igual a 2;
- O segundo passo é fatorar essa matriz  $\mathbf{Y}'$  como um produto  $\mathbf{U} \mathbf{V}^T$  de uma matriz  $\mathbf{U} (I \times 2)$  e uma matriz  $\mathbf{V}^T (2 \times J)$ ;

- O terceiro passo é rotular cada linha da matriz  $\mathbf{U}$  com um símbolo qualquer ( $u_i$ , por exemplo) e cada coluna da matriz  $\mathbf{V}^T$  como outro símbolo diferente ( $v_j$ , por exemplo);
- O quarto e último passo é fazer o gráfico desses pontos rotulados, resultando no biplot aproximado da matriz original.

Outra forma de representar uma matriz de rank igual a três, por exemplo, seria primeiro formar um biplot com a primeira e segunda dimensões da matriz, em seguida formar outro biplot utilizando a primeira e terceira dimensões e por fim, um terceiro biplot utilizando a segunda e a terceira dimensões da matriz. Mas em geral a primeira e segunda dimensões da matriz já explicam a maior parte da variabilidade dos dados. O leitor pode encontrar mais detalhes de cada um dos passos em (GABRIEL, 1980).

Por ser uma forma de exibição gráfica de dados multivariados, o gráfico biplot possui uma ampla aplicabilidade em diversos campos da pesquisa, como por exemplo na área biomédica e da saúde, como uma forma de inspeção preliminar à análise formal dos dados coletados. Alguns exemplos práticos da simplicidade do uso e métodos de construção podem ser encontrados em (GABRIEL; ODOROFF, 1990)

O biplot pode ser aplicado na análise de correlação canônica como um meio alternativo de reduzir a dificuldade de interpretação das variáveis canônicas, de forma que a exibição gráfica possa mostrar as relações entre os dois conjuntos de variáveis, utilizando a abordagem de aproximação matricial com a técnica biplot (BRAAK, 1990).

## 2.4 Procrustes

O procrustes é uma técnica gráfica multivariada que começou a ser desenvolvida por Gower (1971) mas foi no artigo de Gower (1975) que a terminologia procrustes foi adotada. Essa técnica aplica transformações de rotação, translação, reflexão e reescalonamento em dados na forma de matrizes de modo a tornar comparáveis as categorias dessas matrizes transformadas. O objetivo do método é tornar as medidas consensuais na tentativa de identificar padrões gráficos semelhantes e subsidiar a análise inferencial dos dados. A análise de procrustes possui variações e formas de aplicação, dentre elas a mais comum é a análise ortogonal generalizada de procrustes (sigla GPA em inglês) devido à natureza ortogonal da matriz de rotação.

A ideia da metodologia do procrustes é alinhar ao máximo as configurações multivariadas das figuras correspondentes em espaços bidimensionais dos dados e trata as dimensões incompatíveis como essencialmente arbitrárias e, em seguida, girar os espaços até obter a similaridade máxima usando a rotação de procrustes (GRICE; HARDY, 2009).

Estudar as relações entre vários conjuntos de dados utilizando algum tipo de análise combinada é um procedimento desejável e comum na pesquisa científica. O ajuste de duas

matrizes  $\mathbf{Y}_1$  e  $\mathbf{Y}_2$ , por exemplo, permite a aplicação de movimentos de rotação, translação, reflexão e estimativa de um fator de escala homogêneo de forma que o melhor ajuste seja aquele que atende aos critérios de mínimos quadrados da distância euclidiana entre dois pontos, nesse caso valores de duas matrizes (norma de Frobenius). Problemas assim têm solução analítica e o melhor ajuste ocorre quando os centróides de  $\mathbf{Y}_1$  e  $\mathbf{Y}_2$  estão sobrepostos. Uma outra forma de analisar um conjunto de matrizes ao invés de examinar todos os pares possíveis é utilização dessa metodologia, que utiliza dos movimentos de rotação, reflexão e redimensionamento de uma só vez, minimizando a soma de quadrados das distâncias entre os centróides de cada cluster (GOWER, 1975).

### 2.4.1 Descrição da metodologia do procrustes

Uma característica do procrustes é a possibilidade de estudar a relação ou proximidade entre dois subespaços simultâneos conhecidos. Esse estudo é baseado na rotação dos eixos dos subespaços e pode ser realizado através da SVD que foi apresentado na Seção 2.2. A obra de Golub e Van-Loan (2013) também fornece um material sobre essa metodologia e iremos utilizar como uma referência nesse trabalho.

Considere um conjunto de experimentos hipotéticos cujos resultados consistem em uma matriz  $\mathbf{A} \in \mathbb{R}^{I \times J}$  com  $I$  linhas e  $J$  colunas. Se esse mesmo conjunto de experimentos for novamente executado, resultará em uma matriz de dados diferente,  $\mathbf{B} \in \mathbb{R}^{I \times J}$ .

Utilizando a metodologia do procrustes, é possível rotacionar o subespaço formado pela matriz  $\mathbf{B}$  no subespaço  $\mathbf{A}$  resolvendo a Equação 2.5.

$$\text{minimizar } \|\mathbf{A} - \mathbf{BQ}\|_F, \quad \text{sujeito a } \mathbf{Q}^T \mathbf{Q} = \mathbf{I}_J. \quad (2.5)$$

O subscrito  $\|\cdot\|_F$  na Equação 2.5 representa a norma matricial Frobenius. Como exemplo, no caso da matriz  $\mathbf{A}$  é definida como,  $\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^I \sum_{j=1}^J |a_{ij}|^2}$ .

O fato de otimizar  $\mathbf{Q}$  pode ser especificado em termos da SVD do produto das matrizes  $\mathbf{B}^T \mathbf{A}$ . Um fato crítico nessa otimização é a derivação do traço da matriz, representada pela Equação 2.6.

$$\text{tr}(\mathbf{C}) = \sum_{s=1}^S c_{ss}, \quad \text{para } \mathbf{C} \in \mathbb{R}^{S \times S}. \quad (2.6)$$

Adicionalmente, se  $\mathbf{C}_1$  e  $\mathbf{C}_2$  têm a mesma dimensão de linhas e colunas, então,

$$\text{tr}(\mathbf{C}_1^T \mathbf{C}_2) = \mathbf{C}_2^T \mathbf{C}_1 \quad (2.7)$$

Ainda, se  $\mathbf{Q} \in \mathbb{R}^{J \times J}$ , for uma matriz ortogonal, então a otimização da Equação 2.5 tem como solução a Equação 2.8,

$$\begin{aligned}
\|\mathbf{A} - \mathbf{B} \mathbf{Q}\|_F^2 &= \sum_{l=1}^J \|\mathbf{A}(:,l) - \mathbf{B} \mathbf{Q}(:,l)\|_2^2 \\
&= \sum_{l=1}^J \{ \|\mathbf{A}(:,l)\|_2^2 + \|\mathbf{B} \mathbf{Q}(:,l)\|_2^2 - 2 \mathbf{Q}(:,l)^T \mathbf{B}^T \mathbf{A}(:,l) \} \\
&= \|\mathbf{A}\|_F^2 + \|\mathbf{B} \mathbf{Q}\|_F^2 - 2 \left[ \sum_{l=1}^J \mathbf{Q}^T (\mathbf{B}^T \mathbf{A}) \right]_{ll} \\
&= \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 - 2 \operatorname{tr}(\mathbf{Q}^T (\mathbf{B}^T \mathbf{A})).
\end{aligned} \tag{2.8}$$

Portanto, a Equação 2.5 é equivalente a Equação 2.9,

$$\max_{\mathbf{Q}^T \mathbf{Q} = \mathbf{I}_J} \operatorname{tr}(\mathbf{Q}^T \mathbf{B}^T \mathbf{A}) \tag{2.9}$$

Assim, se  $\mathbf{U}^T (\mathbf{B}^T \mathbf{A}) \mathbf{V} = \Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_J)$  for a decomposição em valores singulares de  $\mathbf{B}^T \mathbf{A}$ , definimos a matriz ortogonal  $\mathbf{Z}$  por  $\mathbf{Z} = \mathbf{V}^T \mathbf{Q}^T \mathbf{U}$ , e aplicando a Equação 2.7, resulta na Equação 2.10,

$$\operatorname{tr}(\mathbf{Q}^T \mathbf{B}^T \mathbf{A}) = \operatorname{tr}(\mathbf{Q}^T \mathbf{U} \Sigma \mathbf{V}^T) = \operatorname{tr}(\mathbf{Z} \Sigma) = \sum_{l=1}^J z_{ll} \sigma_l \leq \sum_{l=1}^J \sigma_l \tag{2.10}$$

O limite superior é alcançado definindo  $\mathbf{Z} = \mathbf{I}_J$ , ou seja, definindo  $\mathbf{Q} = \mathbf{U} \mathbf{V}^T$ .

Um algoritmo resumido do que abordamos até aqui pode ser escrito como: Considere  $\mathbf{A} \in \mathbb{R}^{I \times J}$  e  $\mathbf{B} \in \mathbb{R}^{I \times J}$ , os passos a seguir permitem encontrar uma matriz ortogonal  $\mathbf{Q} \in \mathbb{R}^{J \times J}$  tal que  $\|\mathbf{A} - \mathbf{B} \mathbf{Q}\|_F$  seja mínimo,

- Calcular  $\mathbf{C} = \mathbf{B}^T \mathbf{A}$ ;
- Encontrar a SVD de  $\mathbf{U}^T \mathbf{C} \mathbf{V} = \Sigma$  e guardar os valores de  $\mathbf{U}$  e  $\mathbf{V}$ ;
- Calcular  $\mathbf{Q} = \mathbf{U} \mathbf{V}^T$ .

Para maiores detalhes e provas da metodologia do procrustes, indicamos a leitura da seguinte referência: [Gower \(1971\)](#), [Gower e Dijkstra \(2004\)](#) e [Golub e Van-Loan \(2013\)](#).

## 2.5 Identificabilidade de parâmetros

Uma revisão sobre a abordagem da identificabilidade de parâmetros no contexto de modelagem aponta que esse tema não é consenso entre os pesquisadores. Modelos estatísticos complexos exigem esforços adicionais nas formulações de um modelo adequado. Mas podem

existir situações em que o modelo não é identificável, seja por motivo da ausência do estabelecimento de regras de restrição, pela incorreta especificação dos parâmetros do modelo ou mesmo estas estando alinhadas com as regras da literatura, permanece a falta de identificabilidade. As restrições de identificabilidade devem fazer parte da definição do modelo.

Diante da lacuna existente de um consenso acadêmico, iremos trabalhar com o conceito de identificabilidade de parâmetros apresentado por [Frühwirth-Schnatter \(2006\)](#).

Uma família paramétrica de distribuições é indexada por um parâmetro  $\psi \in \Psi$ , definido sobre todo o espaço amostral  $\Omega$ , então essa família é dita ser identificável se para quaisquer dois parâmetros,  $\psi$  e  $\psi^*$  (ambos pertencendo a  $\Psi$ ), definem a mesma lei de probabilidade em  $\Omega$ , portanto, ambos  $\psi$  e  $\psi^*$  são idênticos. Do ponto de vista de funções de densidades de probabilidade (fdp's), por exemplo,  $f(\omega|\psi)$  e  $f(\omega|\psi^*)$ , significa que se essas fdp's forem idênticas para quase todos  $\omega \in \Omega$ , então os parâmetros  $\psi$  e  $\psi^*$  precisam necessariamente serem iguais,

$$f(\omega|\psi) = f(\omega|\psi^*), \text{ para quase todo } \omega \in \Omega \rightarrow \psi = \psi^*. \quad (2.11)$$

Se para quaisquer dois parâmetros distintos  $\psi$  e  $\psi^* \in \Psi$ , têm as funções de densidade de probabilidade  $f(\omega|\psi)$  e  $f(\omega|\psi^*)$  idênticas para quase todo  $\omega \in \Omega$ , então essa família de distribuições não é identificável, dessa forma, qualquer subconjunto  $\mathcal{U}(\psi)$  de  $\Psi$ , definido como

$$\mathcal{U}(\psi) = \{ \psi^* \in \Psi : f(\omega|\psi^*) = f(\omega|\psi), \text{ para quase todo } \omega \in \Omega \}, \quad (2.12)$$

de forma que contenha dois ou mais pontos em  $\Omega$  é considerado um conjunto de não identificabilidade de parâmetro.

## 2.6 Simulação

Dentro do contexto da modelagem Bayesiana, a etapa da simulação estocástica é parte fundamental na estimação dos parâmetros do modelo. Os métodos de Monte Carlo via Cadeias de Markov (na literatura conhecidos como MCMC), são largamente empregados na inferência Bayesiana, principalmente pela dificuldade em obter soluções de forma analítica. O método teve início com o artigo [Metropolis \*et al.\* \(1953\)](#).

A ideia central do método MCMC é utilizar as propriedades da cadeia de Markov para simular amostras de uma distribuição de interesse. A metodologia MCMC é a principal ferramenta de estimação na inferência Bayesiana, tendo como sua fundamental característica a utilização das propriedades das cadeias de Markov para simular o comportamento de vetores aleatórios.

Dentre os métodos de MCMC existentes, os mais populares são o algoritmo de Metropolis-Hastings e o amostrador de Gibbs, que é um caso particular do algoritmo de Metropolis-Hastings,

apresentado por [Geman e Geman \(1984\)](#). Encontramos uma descrição introdutória mais acessível que tornou essa técnica mais difundida na comunidade científica em [Casella e George \(1992\)](#).

O foco do trabalho está concentrado no Metropolis-Hastings e no reversible jump MCMC (RJMCMC) que falaremos mais à frente.

### 2.6.1 Algoritmo Metropolis-Hastings

O algoritmo Metropolis Hastings (utilizaremos a escrita abreviada MH) surge da proposta generalizada por [Hastings \(1970\)](#) a partir do trabalho apresentado por Metropolis *et al* (1953). O artigo publicado por [Chib e Greenberg \(1995\)](#), forneceu um tutorial de introdução ao algoritmo, baseado na publicação do trabalho original de [Metropolis \*et al.\* \(1953\)](#) e [Hastings \(1970\)](#).

O trabalho é idealizado no contexto de simular uma função de densidade de probabilidade alvo contínua, embora a mesma ideia se aplica para distribuições discretas ou mistas.

A implementação do algoritmo de MH depende necessariamente da escolha adequada de uma função de densidade geradora candidata. O mais razoável é escolher uma função densidade que atue no mesmo espaço paramétrico da função de densidade alvo. Dentre as possibilidades de escolha de família de densidade candidata geradora da distribuição alvo apresentadas no trabalho de [Chib e Greenberg \(1995\)](#), em nosso trabalho optamos por utilizar o passeio aleatório, pela facilidade na implementação, interpretação e desempenho computacional. O passeio aleatório também foi utilizado no trabalho de [Metropolis \*et al.\* \(1953\)](#). Com relação a escolha da densidade geradora, a mais comum na literatura é a distribuição Normal.

### 2.6.2 Método reversible jump MCMC

O reversible jump MCMC (RJMCMC) é um método que propõe uma cadeia de Markov reversível onde é possível saltar em subespaços de dimensões diferentes e foi apresentado por [Green \(1995\)](#). Essa metodologia amplia a aplicabilidade do método MCMC incluindo a seleção de variáveis em um modelo e a seleção de modelos.

O objetivo do método de RJMCMC é construir uma cadeia de Markov que possa saltar entre os  $\mathcal{M}$  modelos e retornar valores simulados dos parâmetros estimados do(s) modelo(s) mais adequado(s) ao conjunto de dados  $y$ . Como resultado, ao final do processo, retorna a probabilidade *a posteriori* para cada um dos principais modelos. Com esses resultados, será possível estabelecer um julgamento do melhor, ou, na verdade, do modelo mais adequado dentre os  $K$  modelos possíveis.

Para entender o reversible jump no contexto Bayesiano, vamos supor que para um conjunto de valores observáveis  $y = (y_1, y_2, y_3, \dots)$  temos uma quantidade contável de modelos candidatos  $\mathcal{M} = \{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \dots, \mathcal{M}_{\mathcal{K}}\}$  todos indexados por um parâmetro que pode ser tratado como uma variável indicadora de um modelo auxiliar  $s \in \mathcal{K}$ . Dessa forma, um modelo  $\mathcal{M}_{s'}$ , por exemplo, passa a ser rotulado como um modelo ao qual  $s = s'$ . Para cada um dos

modelos  $\mathcal{M}_s$  há um vetor  $\zeta_s \in \mathbb{R}^{n_s}$  de parâmetros desconhecido, onde  $n_s$  representa a dimensão de cada vetor que pode apresentar diferentes valores para cada modelo  $s \in \mathcal{S}$ .

No processo de implementação do reversible jump, ocorrerão dois tipos de movimentos durante a simulação de uma cadeia Markov, o movimento entre os  $\mathcal{M}_s$  modelos, o que chamamos de movimento entre-modelos ( $Mov_{entre}$ ), é o momento em que um modelo  $\mathcal{M}_s$  é posto em teste contra outro modelo  $\mathcal{M}_{s'}$  e critérios estabelecidos previamente irão retornar o modelo selecionado. Nessa fase é preciso muita atenção pois há uma necessidade de se garantir a reversibilidade da cadeia durante a movimentação entre-modelos.

E também ocorrerá movimento intra-modelos ( $Mov_{intra}$ ), que nada mais é do que a aplicação do método de MH para simular uma cadeia de Markov do modelo  $\mathcal{M}_s$  (ou  $\mathcal{M}_{s'}$ ) selecionado no movimento entre-modelos. Nesse passo também que ocorre a atualização das variáveis do vetor aleatório  $\zeta_s$  em bloco único, sub-blocos através do Metropolis-Hastings, tendo como referência a distribuição *a posteriori* conjunta  $p(\zeta_s|y, \mathcal{M}_s)$ .

O método possibilita a alternância dos movimentos intra e inter-modelos simultaneamente. Esse processo de alternância termina quando um critério de convergência é atendido ou quando uma quantidade considerada suficientemente grande de amostra seja obtida.

O movimento entre diferentes modelos de diferentes dimensões é factível graças à possibilidade de equivalência entre essas dimensões, ou seja, à possibilidade de complementar dimensões de modelos menores quando, por exemplo,  $n_s \neq n_{s'}$ . Esse processo será apresentado mais adiante.

A movimentação entre modelos, sejam eles com mesma dimensão ou não, ocorre da seguinte forma:

Suponha que o modelo atual seja  $\mathcal{M}_s$  com espaço de estado  $(s, \zeta_s)$  e estamos interessados em testar um modelo proposto  $\mathcal{M}_{s'}$  pertencente ao espaço de estado  $(s', \zeta_{s'})$ , cujas dimensões sejam distintas,  $n_{s'} > n_s$ .

Em primeiro lugar, é preciso equalizar as dimensões dos dois modelos. A forma proposta por [Green \(1995\)](#) foi combinar um vetor  $\mathbf{u}$  de dimensão  $d_{s \rightarrow s'} = n_{s'} - n_s$ , gerada de uma densidade conhecida  $q_{d_{s \rightarrow s'}}(\mathbf{u})$ , passando então o vetor de parâmetros do estado atual  $\zeta_k$  para um novo estado comparável,  $\zeta_{s'} = g_{s \rightarrow s'}(\zeta_s, \mathbf{u})$  utilizando uma função de transição  $g_s \rightarrow g_{s'} : \mathbb{R}^{n_s} \times \mathbb{R}^{n_{s'}}$ .

Com isso, a probabilidade de aceitação do modelo proposto, para a distribuição conjunta *a posteriori* para o movimento  $\mathcal{M}_s \rightarrow \mathcal{M}_{s'}$  é dada pela Equação 2.13,

$$\alpha [(s, \theta_s), (s', \theta_{s'})] = \min \left\{ 1, \frac{\pi(s', \theta_{s'}|y)}{\pi(s, \theta_k|y)} \frac{q(s' \rightarrow s)}{q(s \rightarrow s')} \frac{1}{q_{d_{s \rightarrow s'}}(\mathbf{u})} \left| \frac{\partial g_{s \rightarrow s'}(\theta_k, \mathbf{u})}{\partial(\theta_k, \mathbf{u})} \right| \right\}, \quad (2.13)$$

onde  $\pi(s', \theta_{s'}|y)$  e  $\pi(s, \zeta_s|y)$  representam a distribuição *a posteriori* do modelo proposto  $\mathcal{M}_{s'}$  e do modelo atual  $\mathcal{M}_k$ , respectivamente, as funções  $q(s \rightarrow s')$  e  $q(s' \rightarrow s)$  representam a proba-

bilidade de propor uma mudança do modelo  $\mathcal{M}_s$  para o modelo  $\mathcal{M}_{s'}$  e  $\mathcal{M}_{s'}$  para o modelo  $\mathcal{M}_s$ , respectivamente, já o termo  $q_{d_{s \rightarrow s'}}(\mathbf{u})$  denota a função de densidade do vetor aleatório  $\mathbf{u}$ , e  $|\cdot|$  representa o determinante da matriz do Jacobiano da mudança de estado de  $(s, \theta_k)$  para o espaço de estado  $(s', \zeta_{s'})$ , ou Jacobiano, como é conhecido na literatura. O Jacobiano se faz necessário devido à mudança de variável quanto utilizamos a função de transição  $g_{s \rightarrow s'}$  para a mudança de variáveis de  $(\zeta_s, \mathbf{u}) \rightarrow \zeta_{s'}$ .

Já no caso de uma movimentação oposta, isto é, a probabilidade de aceitação do modelo  $\mathcal{M}_k$  ao invés do modelo  $\mathcal{M}_{s'}$ , (movimento inverso)  $\mathcal{M}_{s'} \rightarrow \mathcal{M}_k$  é dada por,

$$\alpha [(s', \zeta_{s'}), (s, \zeta_s)] = \alpha [(s, \zeta_k), (s', \zeta_{s'})]^{-1} \quad (2.14)$$

Como referências para métodos MCMC, MH e RJMCMC sugerimos a leitura das seguintes publicações: [Metropolis \*et al.\* \(1953\)](#), [Chib e Greenberg \(1995\)](#), [Green \(1995\)](#), [Fan e Sisson \(2010\)](#) e [Brooks \*et al.\* \(2011\)](#).

### 2.6.3 Avaliação da convergência no método MCMC

A avaliação da convergência da cadeia utilizando os métodos MCMC pode ser feita de diversas formas, utilizando as técnicas numéricas e as técnicas gráficas. Porém, as técnicas gráficas como gráfico de rastreamento (ou trace plot, na terminologia inglesa), por exemplo, se torna uma alternativa interessante na fase de desenvolvimento.

Um critério que consideramos interessante é o da estatística de [Gelman e Rubin \(1992\)](#) e trata-se de um método estatístico que compara a variabilidade inter e intracadeias, ou seja, o método consiste em uma análise de variância em várias subcadeias para cada parâmetro.



---

## MODELO DE ASSOCIAÇÃO RC

---

### 3.1 Definições

Nesse capítulo são exploradas as definições de tabela de contingência associada à distribuição de probabilidade e as definições do modelo de associação RC. Para estimação dos parâmetros do modelo, descrevemos a função de verossimilhança e o método Bayesiano, em que, apresentamos as distribuições *a priori* consideradas e a distribuição *a posteriori* conjunta para os parâmetros do modelo.

#### 3.1.1 Definição de tabela de contingência

Definimos uma tabela de contingência de dupla entrada  $\mathbf{Y}$  de dimensão  $I \times J$  como o resultado do cruzamento de duas variáveis aleatórias categóricas ou categorizadas, onde  $I$  é o número de linhas e  $J$  é o número de colunas dessa tabela. A variável aleatória  $Y_{ij}$  representa célula  $i,j$  da tabela de contingência, como sendo o cruzamento do elemento da linha  $i$  com o elemento da coluna  $j$ , para  $i=1,\dots,I$  e  $j=1,\dots,J$ . Seja  $\mathbf{Y}$  a notação na forma matricial dessa mesma tabela de contingência com elementos também representados pela variável aleatória  $Y_{ij}$ .

A contagem ou frequência das caselas resultantes do cruzamento entre as categorias dessas duas variáveis, pode ser representada por  $y_{ij}$  como uma realização da variável aleatória  $Y_{ij}$ , proveniente de uma distribuição de probabilidade representada por  $\mathcal{D}_{ij}$ , onde  $\mathcal{D}_{ij}$  pode ter distribuição Poisson, Multinomial (condicionada por linha ou coluna ou total da tabela inteira) ou Binomial Negativa. No modelo de associação que iremos trabalhar, optamos pela distribuição Poisson, ou seja,  $\mathcal{D}_{ij} \sim Poi(\xi_{ij})$ , onde  $\xi_{ij}$  representa o parâmetro da distribuição Poisson para cada célula da tabela de contingência, em que a média e a variância são iguais ao parâmetro desta distribuição. Os valores de cada célula  $(i,j)$ , tem distribuição Poisson tal que  $P(Y_{ij} = y_{ij} | \xi) = \frac{e^{-\xi_{ij}} \xi_{ij}^{y_{ij}}}{y_{ij}!}$ , em que,  $y_{ij} \in \Upsilon = \{0, 1, 2, \dots\}$ ,  $\xi_{ij} > 0$  e  $y_{ij}!$  representa o fatorial do valor  $y_{ij}$ , para  $i=1,\dots,I$ ,  $j=1,\dots,J$ .

## 3.2 Primeiros modelos de associação

A característica do modelo de associação é estimar as interações entre as variáveis aleatórias qualitativas representadas por categorias dispostas em linhas e colunas de uma tabela de contingência. Metodologias para medir a independência entre categorias dispostas em linhas e colunas em tabelas de contingência de dimensão 2x2 tem sido utilizada a vários anos, principalmente as métricas de *odds ratio* (já apresentado anteriormente).

Uma tabela de contingência  $Y$  com duas entradas onde  $y_{ij}$  representa a frequência observada para a célula  $i,j$  (com  $i=1,\dots,I$  e  $j=1,\dots,J$ ) e, seja  $F_{ij}$  a frequência esperada da variável aleatória  $Y_{ij}$  ou  $F_{ij} = E[Y_{ij}]$ , para essa mesma tabela de contingência. Para subtabelas de ordem 2x2 formadas pelas  $i$ -ésimas-1 linhas e  $j$ -ésimas-1 colunas adjacentes,  $(i, i+1)$  e  $(j, j+1)$ , respectivamente, seja  $\theta_{ij}$  o *odds ratio* correspondente para  $i=1,\dots,I-1$  e  $j=1,\dots,J-1$ . Baseada na frequência esperada, temos que  $\theta_{i,j}$  tem a seguinte forma (GOODMAN, 1979a):

$$\theta_{ij} = \frac{F_{i,j} F_{i+1,j+1}}{F_{i,j+1} F_{i+1,j}}. \quad (3.1)$$

Nesse caso temos  $(I - 1) \times (J - 1)$  *odds ratio*  $\theta_{ij}$  formado pelas linhas e colunas adjacentes da tabela de contingência. Testes usuais como a estatística qui-quadrado e a Análise de Variância, ANOVA, utilizados para avaliar a independência entre as categorias das linhas e colunas podem ser vistos como uma forma de medir a associação das  $(I-1) \times (J-1)$  subtabelas representadas pelos *odds ratio* ( $\theta_{ij}$ ). Tomando como referência a expressão (3.1), temos uma classe de modelos de associação. São eles: O modelo de independência estatística ou modelo de associação nula, o modelo de associação uniforme, o modelo de associação de efeito linha, o modelo de associação de efeito coluna, o modelo de efeito linhas e colunas (modelo que mede a associação, efetivamente, entre linhas e colunas) escrito de duas formas, na forma usual e na forma logarítmica (GOODMAN, 1979b).

O quadro 2 apresenta um resumo dessa classe de modelos de associação,

Tabela 2 – Tipos de modelos de associação

Id (item)	Odds-ratio	Condição de $\theta$	Modelo resultante
(1)	$\theta_{ij} = 1$	$\theta = 1$	Independência ou associação nula
(2)	$\theta_{ij} = \theta$	$\theta$ constante	Uniforme
(3)	$\theta_{ij} = \theta_i$	$\theta_i$ constante	Efeito linha
(4)	$\theta_{ij} = \theta_j$	$\theta_j$ constante	Efeito coluna
(5)	$\theta_{ij} = \theta_i \cdot \theta_j$	$\theta_i$ e $\theta_j$ desconhecidos	Efeito linha e coluna (associação)
(6)	$\log(\theta_{ij}^*) = \theta_i^* \cdot \theta_j^*$	$\theta_i^*$ e $\theta_j^*$ desconhecidos	Efeito linha e coluna (associação)

Fonte: Elaborada pelo autor.

Em termos de graus de liberdade a classe de modelo de associação se resume da seguinte forma,

Tabela 3 – Graus de liberdade por tipo de modelo de associação RC(K)

Tipo de modelo de associação	Graus de liberdade
(1) Modelo de associação nula	(I - 1) (J - 1)
(2) Modelo de associação uniforme	(IJ - I - J)
(3) Modelo de associação de efeito linha	(I - 1) (J - 2)
(4) Modelo de associação de efeito coluna	(I - 2) (J - 1)
(5) Modelo de associação de efeito linha e coluna	(I - 2) (J - 2)
(6) Modelo de associação de efeito linha e coluna	(I - 2) (J - 2)

Fonte: Adaptada de Goodman (1979b).

O Modelo (6) apresentado na Tabela 2 pode ser reescrito, em termos de frequência esperada,  $F_{ij}$  na forma de um modelo log-bilinear, com vistas à questão da identificabilidade de parâmetros (GOODMAN, 1979b),

$$E[Y_{ij}] = F_{ij} = \alpha_i \beta_j e^{\phi \mu_i \nu_j}, \quad \text{para todo } i, j \quad (3.2)$$

onde,  $\alpha$ ,  $\beta$  são parâmetros que medem os efeitos da linha e coluna, respectivamente,  $\phi$ ,  $\mu$  e  $\nu$  são os parâmetros das interações linhas e colunas. Cabe observar que esse modelo (3.2) considera  $K=1$ , uma vez que se o valor de  $k=0$ , temos um modelo de independência e  $E[Y_{ij}] = \alpha_i \beta_j$  (GOODMAN, 1979b).

A reescritura do modelo (3.2), considerando  $k=1, \dots, K$  é representada por,

$$E[Y_{ij}] = \alpha_i \beta_j \exp \left( \sum_{k=1}^K \phi_k \mu_{ik} \nu_{kj} \right), \quad (3.3)$$

sendo os parâmetros  $\alpha_i$  e  $\beta_j$  representam os efeitos das linhas e colunas, respectivamente,  $\phi_k$  representa o nível de cada interação  $k$ , e  $\mu_{ik}$  e  $\nu_{kj}$  medem os efeitos de interações das linhas e colunas, respectivamente,  $i=1, \dots, I$  e  $j=1, \dots, J$ .

As restrições (3.4) e (3.5) que fazem parte da especificação do modelo (3.3), são (GOODMAN, 1981):

$$\sum_{i=1}^I \alpha_i = \sum_{j=1}^J \beta_j = \sum_{i=1}^I \mu_{ik} = \sum_{j=1}^J \nu_{jk} = 0 \quad (3.4)$$

e

$$\begin{aligned} \sum_{i=1}^I \mu_{ik} &= \sum_{j=1}^J \nu_{jk} = \sum_{i=1}^I \mu_{ik} \mu_{ik'} = \sum_{j=1}^J \nu_{jk} \nu_{jk'} = 0 \\ \sum_{i=1}^I \mu_{ik}^2 &= \sum_{j=1}^J \nu_{jk}^2 = 1, \end{aligned} \quad (3.5)$$

em que o índice  $k, k' \in \{1, \dots, K\}$  e  $k \neq k'$ .

Considerando o modelo da Equação (3.2), o cálculo do *odds ratio* é dado pela Equação (3.6).

$$\log(\theta_{ij}) = \phi (\mu_{i+1} - \mu_i) (v_{j+1} - v_j). \quad (3.6)$$

As restrições impostas aos parâmetros da Equação (3.6) são (GOODMAN, 1979b),

$$\begin{aligned} \sum_{i=1}^I \mu_i &= \sum_{j=1}^J v_j = 0, \\ \sum_{i=1}^I \mu_i^2 &= \sum_{j=1}^J v_j^2 = 1 \end{aligned} \quad (3.7)$$

respectivamente.

No artigo posterior de Goodman (1981) em que comparou o modelo de associação e o modelo correlação canônica descrevendo as estimativas dos scores das linhas e colunas de uma tabela de contingência, onde hipóteses foram testadas utilizando a estatística qui-quadrado para verificar a hipótese nula de independência entre as categorias das variáveis em linha e colunas, apenas o modelo log-bilinear foi utilizado nos moldes do modelo de associação,

Para subtabelas 2x2 obtidas através das linhas e colunas adjacentes ( $i, i+1$  e  $j, j+1$ , respectivamente), a tabela de contingência ( $I \times J$ ) tem o seguinte logaritmo do *odds ratio*  $\theta_{ij}$  correspondente à Equação (3.6) (GOODMAN, 1981).

$$\log(\theta_{ij}) = \sum_{k=1}^K \phi_k (\mu_{ki} - \mu_{k,i+1}) (v_{kj} - v_{k,j+1}). \quad (3.8)$$

Prevalendo as mesmas restrições da Equação (3.7) para o modelo (3.8). O modelo (3.3) foi utilizado nas publicações posteriores de Goodman (1985, 1986, 1996).

### 3.3 Modelo de associação RC

O modelo de associação RC(K) de Goodman (1986), na sua forma saturada, para análises de tabelas de contingência que adotamos foi transformado em termos de logaritmo e é representado pela Equação (3.9),

$$\eta_{ij} = \text{Log}(E[Y_{ij}]) = \log(\alpha_i) + \log(\beta_j) + \sum_{k=1}^K \phi_k \mu_{ik} v_{kj} \quad (3.9)$$

para  $i=1, \dots, I, j=1, \dots, J$  e  $k=1, \dots, K$ , com restrições (3.5) e (3.11).

Tomando como referência a Equação (3.9) e rotulando os parâmetros, o modelo de associação RC(K), saturado, que iremos trabalhar será o modelo representado na Equação (3.10),

$$\eta_{ij} = \lambda + \lambda_i^R + \lambda_j^C + \sum_{k=1}^K \phi_k \mu_{ik} v_{jk} \quad (3.10)$$

em que  $\eta_{ij} = \log(E[Y_{ij}])$  é composto pelos componentes lineares e bi-linear do modelo,  $(\lambda + \lambda_i^R) = \log(\alpha_i)$ ,  $\lambda_j^C = \log(\beta_j)$ , sendo  $\lambda$  a média global do modelo. Os parâmetros  $\lambda_i^R$  e  $\lambda_j^C$  são parâmetros que representam os efeitos das linhas e das colunas, respectivamente, da tabela de contingência, cada  $\phi_k$  é uma medida de associação entre as categorias das linhas e colunas da tabela e,  $\mu_i = (\mu_{i1}, \dots, \mu_{iK})$  e  $v_j = (v_{j1}, \dots, v_{jK})$  são vetores de parâmetros que representam os efeitos das interações das linhas e das colunas, respectivamente, da tabela de contingência, para  $i=1, \dots, I$ ,  $j=1, \dots, J$ ,  $k=1, \dots, K$ , com  $K_{max} = \min(I, J) - 1$ .

Nesse trabalho, os valores do parâmetro  $\phi_k$  são positivos, ou seja,  $\phi_k \geq 0$ . A ordem do modelo de associação K tem um papel importante nas análises. Quando o valor de  $K = 0$  ou  $\phi_1 = 0$  para  $K = 1$  temos o modelo de associação nulo, ou modelo de independência. Já para  $K > 0$  (exceto o caso em que  $\phi_1 = 0$  quando  $K=1$ ), temos uma associação entre as categorias das variáveis em que a força dessa associação será medida através dos valores de  $\phi_k$ , para  $k=1, \dots, K$  (GOODMAN, 1981).

As restrições da Equação (3.10), descritas na Equação (3.11), que fazem parte da especificação do modelo, além das já especificadas em (3.5), são importantes para garantirem a identificabilidade dos parâmetros e proporcionam a interpretabilidade do modelo.

$$\sum_{i=1}^I \lambda_i^R = \sum_{j=1}^J \lambda_j^C = 0. \quad (3.11)$$

A restrições apresentadas em (3.5) e (3.11) foram propostas por Goodman (1981) e também foram utilizadas por Iliopoulos, Kateri e Ntzoufras (2007, 2009), Tsujitani (1987), Milan e Whittaker (1992), Kateri, Ahmad e Papaioannou (1998) e Becker e Clogg (1989).

### 3.3.1 Representação matricial do modelo de associação RC

A notação matricial da reparametrização do modelo de associação RC apresentado em (3.10) é,

$$\eta = \lambda + \lambda^R \mathbf{1}_J + \mathbf{1}_I^T \lambda^C + \mu_k \phi_k v_k^T \quad (3.12)$$

em que,

- $\eta = \log E[\mathbf{Y}]$ , é uma matriz de dimensão  $I \times J$  composta pelos elementos  $\eta_{ij}$  dado na Expressão (3.10), para  $i=1, \dots, I$  e  $j=1, \dots, J$ .

- $\lambda^{\mathbf{R}} = [\lambda_1^{\mathbf{R}}, \lambda_2^{\mathbf{R}}, \dots, \lambda_I^{\mathbf{R}}]^T$  é um vetor colunas de dimensão I,
- $\lambda^{\mathbf{C}} = [\lambda_1^{\mathbf{C}}, \lambda_2^{\mathbf{C}}, \dots, \lambda_J^{\mathbf{C}}]$  é vetor linhas de dimensão J,
- $\mathbb{1}_I = \mathbb{1}_J$  = são vetores linhas de dimensões I e J, respectivamente,
- $\mu_{\mathbf{K}}$  é uma matriz de dimensão  $I \times K$  composta pelos elementos  $\mu_{ik}$ , para  $i=1, \dots, I$ , e  $k=1, \dots, K$ ,
- $\nu_{\mathbf{K}}$  é uma matriz de dimensão  $J \times K$  composta pelos elementos  $\nu_{jk}$ , para  $j=1, \dots, J$ , e  $k=1, \dots, K$ ,
- $\phi_{\mathbf{K}} \text{diag} [\phi_1, \dots, \phi_K]$  representa uma matriz diagonal de dimensão  $K \times K$ ,

com,  $\phi_1 \geq \phi_2 \geq \dots \geq \phi_K$  e  $k=1, \dots, K$ .

A Equação (3.12) está sujeita às mesmas regras de indentificabilidade apresentadas nas Equações (3.5) e (3.11), com  $\mu\mu^T = I$ ,  $\nu^T\nu = I$  e  $\phi$ =diagonal.

Ao trabalharmos o modelo em sua forma matricial, vamos ter a vantagem de utilizarmos a SVD, em que as propriedades da decomposição, já apresentadas na Seção 2.2 atendem perfeitamente aos critérios de restrições que precisamos impor ao modelo para garantir a indentificabilidade.

### 3.3.2 SVD e o modelo de associação

No presente trabalho iremos utilizar a decomposição em valores singulares na estimação dos parâmetros do modelo de associação RC, primeiro porque as propriedades da SVD satisfazem os critérios das restrições (3.5) necessárias para garantir a indentificabilidade do modelo durante a fase de implementação na forma matricial, segundo pela possibilidade de construção da solução gráfica dos biplots.

A SVD neste trabalho é aplicada para atender as regras de indentificabilidade dos parâmetros do modelo de associação RC. Para cada simulação do modelo via métodos MCMC, aplicamos a SVD. Os resultados obtidos após aplicação da SVD serão utilizados na construção do biplot e aplicação do procrustes.

## 3.4 Identificabilidade no modelo de associação RC

O modelo de associação RC sem restrição ou com restrições insuficientes é um modelo não identificável. A indentificabilidade faz parte da definição do modelo, as propostas de restrições para tornar o modelo de associação RC identificável encontrados na literatura podem se diferenciar na sua estrutura mas atendem os critérios de indentificabilidade.

O modelo de associação RC foi utilizado no estudo das estimativas dos parâmetros de ordem restrita com imposição de restrições na ordem das estimativas dos parâmetros do modelo

de efeito linha (ou coluna). Essas estimativas dos parâmetros do modelo possibilitaram o agrupamento de categorias das variáveis dispostas em linhas (ou colunas) de uma tabela (AGRESTI; CHUANG; KEZOUH, 1987). Outra proposta de reparametrização alternativa sobre as restrições dos parâmetros do modelo de associação utilizou a abordagem Bayesiana e RJMCMC para estimar as probabilidades *a posteriori* do modelo de associação utilizando uma restrição de ordem das linhas e colunas da tabela de contingência (ILIOPOULOS; KATERI; NTZOUFRAS, 2009).

### 3.5 Outras interpretações do modelo de associação

Vários autores aplicaram e tiveram suas próprias interpretações sobre o modelo de associação de Goodman após 1981, inclusive sobre as semelhanças com outros modelos, como por exemplo, o modelo de correlação canônica em que Goodman já havia apresentado.

Becker (1989) apresenta uma família de modelos complementares aos log-lineares utilizado para analisar a associação em tabelas de contingência quando algumas ou todas as variáveis forem ordinais. A verificação do ajuste dos modelos foi feita a partir das estatísticas qui-quadrado de Pearson e de razão de verossimilhança. Foi utilizado um gráfico do tipo biplot para ilustrar a associação entre as variáveis. Aplicação em uma tabela de renda e educação, detectou a existência de uma associação positiva entre essas variáveis, correspondendo a maior renda ao ensino superior.

O artigo de Becker e Clogg (1989) abordam uma classe de modelos para a analisar diferenças na associação entre duas variáveis discretas. O estudo basicamente revisou o modelo de associação RC aplicado em tabelas bidirecionais e apresentou métodos de ponderação alternativos, como por exemplo,  $\sum_{i=1}^I g_i \mu_{ik} = \sum_{j=1}^J h_j \nu_{ik} = 0$  e  $\sum_{i=1}^I g_i \mu_{ik}^2 = \sum_{j=1}^J h_j \nu_{jk}^2 = 1$ , em que  $g_i$  e  $h_j$  são os pesos atribuídos à  $i$ -ésima linha e à  $j$ -ésima coluna, de uma tabela de contingência, respectivamente, e a escolha dos valores se torna importante na comparação entre duas tabelas de contingência quando a distribuição marginal de linhas e colunas são não homogêneas entre as tabelas. Inclusive, Goodman (1981, 1985, 1986) utilizou pesos marginais na demonstração da relação entre os modelos de associação e de correlação canônica convencional. Quando esses fatores,  $g_i = h_j = 1$ , para todo  $i$  e  $j$ , resulta na Equação (3.6).

### 3.6 Concordâncias entre o modelo de associação e o *odds ratio* com biplot

Uma interpretação da distância entre os pontos de linhas e colunas foi desenvolvido por Rooij e Heiser (2005). Esses modelos permitem expressões para o *odds ratio* em termos de distâncias. As representações gráficas podem ser de dois tipos, rotulados de gráfico do tipo I e tipo II (os gráficos estão disponíveis no artigo do autor). No primeiro, a relação entre um par

de conjuntos é representada por uma regra de distância entre pontos e no segundo, a relação é representada por uma regra de produto interno.

No gráfico do Tipo I, representado por pontos o espaço euclidiano, obtidos através da distância entre os pontos, que é a representação da relação entre as categorias de dois conjuntos (conjunto representado pelas linhas e outro pelas colunas). Já no caso do tipo II, é necessário que pelo menos um conjunto seja representado no gráfico como vetores e o outro conjunto, por pontos projetados sobre os vetores, representando o relacionamento entre os conjuntos, esse é o típico gráfico biplot. As representações gráficas do modelo de associação RC e do modelo de correlação canônica são baseadas em produtos internos, ou seja, biplot.

Os modelos baseados na distância foram rotulados pelos autores de modelos de associação de distância. Aplicando transformações utilizando a função de decaimento exponencial e função Gaussiana, as distâncias são então representadas pela Equação (3.13).

$$\theta_{ij} = \text{Exp}(-d_{ij}(X)) \quad \text{e} \quad \theta_{ij} = \text{Exp}(-d_{ij}^2(X)), \quad \text{respectivamente,} \quad (3.13)$$

em que  $\theta_{ij}$  representa o *odds-ratio* do modelo de associação apresentado na Equação (3.6).

No caso do decaimento exponencial a distância absoluta resulta na distância canônica, e no caso da transformação Gaussiana resulta na distância Euclidiana. Por ser a distância mais comum, a distância Euclidiana foi utilizada juntamente com a função de transformação Gaussiana.

No estudo, o modelo de associação de distância possui duas propostas: O modelo de associação de distância de modo único e o de duas vias. O modelo de associação de distância modo único, pode ser aplicado apenas a tabelas quadradas e tem forma,

$$F_{ij} = c \alpha_i \beta_j \text{Exp}(-d_{ij}^2(\mathbf{Y})), \quad (3.14)$$

em que  $c$  é uma constante, a média geral da tabela de contingência,  $\alpha_i$  e  $\beta_j$  são parâmetros marginais da mesma tabela de contingência, respectivamente.

Aplicando o logaritmo, a Equação (3.14) é reescrito como

$$\log(F_{ij}) = \lambda + \lambda_i^R + \lambda_j^C - d_{ij}^2(\mathbf{Y}), \quad (3.15)$$

em que  $\lambda = \text{Log}(c)$ ,  $\lambda_i^R = \text{Log}(\alpha_i)$ , e  $\lambda_j^C = \text{log}(\beta_j)$ .

A Equação (3.15) ainda pode expandida de forma que a função da distância seja

$$\text{Log}(F_{ij}) = \lambda + \lambda_i^R + \lambda_j^C - \sum_k (y_{ik} - y_{jk})^2 = \lambda + \lambda_i^R + \lambda_j^C - \text{tr}(\mathbf{Y}^T \mathbf{A}_{ij} \mathbf{Y}) \quad (3.16)$$

em que  $\text{tr}$  denota o traço de uma matriz. E  $A_{ij}$  é definido como  $A_{ij} = (e_i - e_j)(e_i - e_j)^T$ ,  $e_i$  constitui a  $i$ -ésima coluna da matriz identidade de ordem  $K$  (ROOIJ; HEISER, 2005).

A relação entre modelo de associação RC e modelos de análises de correspondência em análise de variáveis ordinais discretas em tabela de contingência já foi objeto de estudos de Goodman (1985,1986). Porém, Tsujitani (1988) explorou a análise gráfica para interpretar a abordagem do modelo de associação. As frequências observadas  $F_{ij}$  foram assumidas como uma distribuição Poisson, e o modelo de associação RC dado na Equação (3.1) e em termos de *odds ratio* local o modelo descrito foi dada na Equação (3.9).

O trabalho de Tsujitani (1988) propõe um método de estimação dos parâmetros para o modelo, baseado no método da máxima verossimilhança, quando as restrições de ordem dos parâmetros dos efeitos da associação são impostas no modelo de efeito coluna, e propôs também exibições gráficas dos efeitos da associação para explorar uma série de características dos dados em uma única imagem. Gráficos ilustrativos foram elaborados aplicando ponderações (fator  $\tau$ , com  $\tau \in [0, 1]$ ) nas estimativas dos efeitos da associação das linhas e colunas utilizando como fator de ponderação, parcela do valor dos parâmetros de associação  $\phi$ , ou seja,  $\mu_{ik}^* = \phi_k^\tau \mu_{ik}$  e  $v_{jk}^* = \phi_k^{1-\tau} v_{jk}$ , onde  $v_{jk}^*$  e  $v_{jk}^*$  são os efeitos da associação ponderados pelo fator  $\tau$ . As ilustrações gráficas apresentadas não se tratam de um é biplot puramente, foram plotados os scores das linhas e colunas em forma de pontos e concluiu-se que algumas eram distinguíveis, outras apresentavam grupos homogêneos e outros casos haviam grupos bem distintos de scores.

## 3.7 Estimação

Nessa seção vamos apresentar a aplicação da função de verossimilhança e o método de estimação do modelo de associação RC, utilizado na maioria dos artigos referentes ao tema e o método Bayesiano. Este último, daremos um enfoque maior por ser tema principal da proposta.

### 3.7.1 Função de verossimilhança

O modelo Poisson para contagens em tabela de contingência,  $Y_{ij} \sim \text{Poisson}(e^{\eta_{ij}})$ , onde  $i=1,\dots,I$ ,  $j=1,\dots,J$  e  $\eta_{ij} \geq 0$ . A média do modelo de associação apresentado em (3.3) e foi aplicada da função logarítmica resultando  $\eta_{ij} = \log(E[Y_{ij}])$ , (Equação 3.9), portanto,  $E[Y_{ij}] = e^{\eta_{ij}}$ , nesse caso, a distribuição de probabilidade da variável aleatória  $Y_{ij}$  é expressa pela Equação 3.17, com  $K$  conhecido,

$$Pr(Y_{ij} = y_{ij} | \eta, K) = \frac{e^{-e^{\eta_{ij}}} e^{\eta_{ij}^{y_{ij}}}}{y_{ij}!}, \quad (3.17)$$

em que  $y_{ij} \geq 0$ .

Para a distribuição de probabilidade Poisson a função de verossimilhança pode ser expressa da forma

$$L(\eta | y_{ij}) \equiv \prod_{i=1}^I \prod_{j=1}^J f(y_{ij}; \eta_{ij}) = \prod_{i=1}^I \prod_{j=1}^J \frac{e^{-\eta_{ij}} e^{y_{ij}\eta_{ij}}}{y_{ij}!}, \quad (3.18)$$

onde  $\eta_{ij}$  foi definido em (3.10).

A Expressão (3.18) da verossimilhança também pode ser reescrita como

$$L(\eta, y_{ij}) \propto e^{-\sum_{i=1}^I \sum_{j=1}^J e^{\eta_{ij}} + \sum_{i=1}^I \sum_{j=1}^J y_{ij} \eta_{ij}}, \quad (3.19)$$

### 3.7.2 Abordagem Bayesiana

Nessa seção vamos apresentar as distribuições *a priori* e *a posteriori* como parte integrante da modelagem Bayesiana do modelo de associação.

#### 3.7.2.1 Distribuições *a priori* do modelo de associação RC

Em inferência Bayesiana, quando não temos informações sobre as quantidades de interesse desconhecidas (no modelo de associação RC essas quantidades estão representadas pelos parâmetros  $\lambda, \lambda_i^R, \lambda_j^C, \phi_k, \mu_{ik}, v_{kj}$ ), nós atribuímos informações vagas para as distribuições *a priori*.

Nossa proposta nesse tema é definir que as distribuições *a priori* para os parâmetros do modelo de associação RC seguem uma distribuição Normal e são representadas pelas expressões

$$\lambda \sim Normal(0, \sigma_\lambda^2), \quad \lambda_i^R \sim Normal(0, \sigma_{\lambda_i^R}^2), \quad \lambda_j^C \sim Normal(0, \sigma_{\lambda_j^C}^2), \quad (3.20)$$

$$\phi_k \sim Lognormal(0, \sigma_{\phi_k}^2), \quad \mu_{ik} \sim Normal(0, \sigma_{\mu_{ik}}^2), \quad e \quad v_{kj} \sim Normal(0, \sigma_{v_{kj}}^2), \quad (3.21)$$

para  $i=1, \dots, I$ ,  $j=1, \dots, J$  e  $k=1, \dots, K$ .

Como não dispomos de informação alguma a respeito dos parâmetros, nossa intenção é trabalhar com *prioris* vagas, portanto vamos assumir variâncias grandes, por exemplo  $\sigma_\lambda^2 = \sigma_{\lambda_i^R}^2 = \sigma_{\lambda_j^C}^2 = \sigma_{\phi_k}^2 = \sigma_{\mu_{ik}}^2 = \sigma_{v_{kj}}^2 = d^2 = 10^3$ .

Adotamos a ideia de que as distribuições *a priori* do modelo de associação RC são representadas por

$$p(\lambda), \quad \prod_{i=1}^I p(\lambda_i^R), \quad \prod_{j=1}^J p(\lambda_j^C), \quad \prod_{k=1}^K p(\phi_k), \quad \prod_{i=1}^I p(\mu_{ik}) \quad e \quad \prod_{j=1}^J p(v_{kj}) \quad (3.22)$$

### 3.7.2.2 A distribuição a posteriori no modelo de associação RC

Pelo Teorema de Bayes a distribuição *a posteriori* é dada pelo produto da distribuição *a priori* e a função de verossimilhança.

Para  $K$  fixo e conhecido, o modelo de associação RC, tem a distribuição *a posteriori* expressa pela Equação (3.23),

$$p(\eta | y) \propto L(y|\eta) p(\lambda) \prod_{i=1}^I p(\lambda_i^R) \prod_{j=1}^J p(\lambda_j^C) \prod_{k=1}^K p(\phi_k) \prod_{k=1}^K \prod_{i=1}^I p(\mu_{ik}) \prod_{k=1}^K \prod_{j=1}^J p(\nu_{kj}), \quad (3.23)$$

onde  $L(y|\eta)$  é a função de verossimilhança definida em (3.18). Podemos reescrever a Equação (3.23), expandindo a função de verossimilhança e passando os termos constantes para o sinal de proporcionalidade

$$p(\eta | y) \propto \text{Exp} \left\{ \sum_{i=1}^I \sum_{j=1}^J y_{ij} \eta_{ij} - \sum_{i=1}^I \sum_{j=1}^J e^{\eta_{ij}} \right\} \\ \times p(\lambda) \prod_{i=1}^I p(\lambda_i^R) \prod_{j=1}^J p(\lambda_j^C) \prod_{k=1}^K p(\phi_k) \prod_{k=1}^K \prod_{i=1}^I p(\mu_{ik}) \prod_{k=1}^K \prod_{j=1}^J p(\nu_{kj}) \quad (3.24)$$

Na forma matricial a distribuição *a posteriori* da Expressão (3.24) pode ser representada por

$$p(\eta | \mathbf{y}) \propto \exp(\mathbb{1}_J^T \mathbf{y} \boldsymbol{\eta} \mathbb{1}_J - \mathbb{1}_J^T \exp(\boldsymbol{\eta}) \mathbb{1}_J) \\ \times \frac{1}{\sigma_\lambda^{n/2}} \exp(-\mathbf{y}' \sigma_\lambda^{-1} \mathbf{y}) \times \frac{1}{\Sigma_{\lambda_i^R}^{n/2}} \exp(-\mathbf{y}' [\Sigma_{\lambda_i^R}^{-1}] \mathbf{y}) \\ \times \frac{1}{\Sigma_{\lambda_j^C}^{n/2}} \exp(-\mathbf{y}' [\Sigma_{\lambda_j^C}^{-1}] \mathbf{y}) \times \frac{1}{\Sigma_{\phi_k}^{n/2}} \exp(-\mathbf{y}' [\Sigma_{\phi_k}^{-1}] \mathbf{y}) \\ \times \frac{1}{\Sigma_{\mu_{ik}}^{n/2}} \exp(-\mathbf{y}' [\Sigma_{\mu_{ik}}^{-1}] \mathbf{y}) \times \frac{1}{\Sigma_{\nu_{kj}}^{n/2}} \exp(-\mathbf{y}' [\Sigma_{\nu_{kj}}^{-1}] \mathbf{y}), \quad (3.25)$$

em que,  $\Sigma_{\lambda_i^R}$ ,  $\Sigma_{\lambda_j^C}$ ,  $\Sigma_{\phi_k}$ ,  $\Sigma_{\mu_{ik}}$ , e  $\Sigma_{\nu_{kj}}$  são matrizes de variâncias e covariâncias dos parâmetros  $\lambda_i^R$ ,  $\lambda_j^C$ ,  $\phi_k$ ,  $\mu_{ik}$  e  $\nu_{kj}$ , respectivamente.

## 3.8 Simulação

A etapa de simulação de valores aleatórios da distribuição *a posteriori* dos parâmetros do modelo de associação RC será descrita nessa seção. Iniciaremos uma breve descrição de como utilizaremos a metodologia do Metropolis-Hastings e posteriormente do RJMCMC.

### 3.8.1 Metropolis-Hastings no modelo de associação RC

A justificativa para utilizarmos o método de MH se deu pelo fato da necessidade de simularmos uma amostra da distribuição *a posteriori* e então obtermos as estimativas dos parâmetros do modelo de associação. Uma característica do método é que o algoritmo de MH simula valores de uma variável aleatória para uma distribuição de interesse utilizando uma distribuição auxiliar conhecida.

Os valores estimados dos parâmetros do modelo RC são obtidas durante a atualização dos parâmetros dentro do modelo, no momento em que reversible jump for utilizado para a escolha do melhor valor do parâmetro K. As estimativas finais serão obtidas depois de realizar a T iterações do algoritmo. Cabe ressaltar que nesse passo, o valor de K é fixo.

O algoritmo de Metropolis-Hastings.

Para um total de  $t=1, \dots, T$  iterações (T é o número total de iterações), o algoritmo de MH utilizado no modelo de associação será da seguinte forma:

Algoritmo de Metropolis-Hastings

1. Na primeira iteração ( $t=0$ ), atribuir um valor inicial para cada parâmetro do modelo de associação;
2. Propor novos valores para os parâmetros, utilizando a técnica do passeio aleatório, e aplicar a regra da identificabilidade de parâmetros;
3. Simular o valor da distribuição *a priori* do modelo utilizando valores do passo anterior;
4. Simular o valor da distribuição *a posteriori* do modelo do modelo utilizando valores propostos no passo 2;
5. Calcular a probabilidade de aceitação  $\alpha(y, y^*)$  para os valores propostos dos parâmetros do modelo de associação utilizando a expressão  $\alpha(y, y^*) = \min \left[ 1, \frac{\pi(y^*) q(y^*, y)}{\pi(y) q(y, y^*)} \right]$ , em que  $\pi(y)$  e  $\pi(y^*)$  representam as distribuições *a posteriori* de  $y$  e  $y^*$ , respectivamente e  $q(y, y^*)$  e  $q(y^*, y)$  representam a função de transição do estado  $y$  para  $y^*$  e de  $y^*$  para  $y$ , respectivamente;
6. Simular o valor  $u$  de uma variável aleatória de uma distribuição uniforme  $[0, 1]$ ;
7. Aceitar os valores propostos se  $u < \alpha(y, y^*)$ , que é a probabilidade de aceitação calculada no passo 5, caso contrário, manter os valores dos parâmetros;
8. Fazer  $t=t+1$  e repetir os passos de 2 a 7 até obter uma quantidade suficiente de valores necessários para obter a convergência ou estabilidade da cadeia.

### 3.8.2 Reversible Jump MCMC no modelo de associação

A utilização do RJMCMC possibilita trabalharmos com modelos que envolvem vetores com diferentes dimensões - mistura de diferentes dimensões de espaço amostral do modelo - tarefa que as abordagens usuais de MCMC (MH e Gibbs-Sampling) não atenderiam, uma vez que os mesmos precisam que o vetor de parâmetro mantenha sempre a mesma dimensão.

Dois destaques importantes no RJMCMC são: o primeiro diz respeito à equalizar a dimensão do espaço amostral dos modelos no momento em que formos testar um modelo proposto com um modelo atual e o outro destaque é a questão da reversibilidade da Cadeia de Markov, é necessária manter a cadeia reversível para a aplicação do método. Cada um desses pontos são descritos a seguir.

No que diz respeito à equalização das dimensões de dois modelos com espaços paramétricos de dimensões distintas, o modelo com espaço paramétrico menor precisa ser completado da seguinte forma: Suponha que o interesse seja testar a mudança do modelo  $\mathcal{M}_1$  para o modelo  $\mathcal{M}_2$ , cuja dimensão do modelo  $\mathcal{M}_1$  seja maior que a dimensão do modelo  $\mathcal{M}_2$ , ou seja,  $d_{\mathcal{M}_1} > d_{\mathcal{M}_2}$ , e o vetor de parâmetros do modelo  $\mathcal{M}_1$  seja representado por  $\delta_1$  e do modelo  $\mathcal{M}_2$  por  $\delta_2$ . Nesse caso, precisamos completar a dimensão do modelo  $\mathcal{M}_2$  com variáveis aleatórias de alguma distribuição de probabilidade, no caso desse trabalho, vamos utilizar variáveis aleatórias com distribuição Uniforme independente, e a partir daí, combinar um vetor  $\mathbf{u}$  de dimensão  $d_{\mathcal{M}_2} = d_{\mathcal{M}_1} - d_{\mathcal{M}_2}$ , gerada da função densidade conhecida  $q_{d_{\mathcal{M}_1} \rightarrow d_{\mathcal{M}_2}}(\mathbf{u})$ . Portanto, o vetor de parâmetros do estado atual da cadeia  $\delta_1$  será comparável com o novo estado da cadeia de mesma dimensão,  $\tilde{\delta}_2 = g_{\mathcal{M}_1 \rightarrow \mathcal{M}_2}(\delta_1, \mathbf{u})$  através da função de transição  $g_{\mathcal{M}_1 \rightarrow \mathcal{M}_2} : \mathbb{R}^{d_{\mathcal{M}_1}} \times \mathbb{R}^{d_{\mathcal{M}_2}}$ .

A probabilidade de aceitação do modelo proposto, para a distribuição conjunta *a posteriori* para o movimento  $\mathcal{M}_1 \rightarrow \mathcal{M}_2$  é dada por  $\alpha \left[ (\mathcal{M}_1, \delta_1), (\mathcal{M}_2, \tilde{\delta}_2) \right] = \min(1, \mathcal{A})$ , em que

$$\mathcal{A} = \left\{ \frac{\pi(\mathcal{M}_2 | \tilde{\delta}_2, y^*)}{\pi(\mathcal{M}_1 | \delta_1, y)} \frac{q_{(\mathcal{M}_2 \rightarrow \mathcal{M}_1)}}{q_{(\mathcal{M}_1 \rightarrow \mathcal{M}_2)}} \frac{1}{q_{d_{\mathcal{M}_1} \rightarrow d_{\mathcal{M}_2}}(\mathbf{u})} \left| \frac{\partial g_{\mathcal{M}_1 \rightarrow \mathcal{M}_2}(\delta_1, \mathbf{u})}{\partial(\tilde{\delta}_2, \mathbf{u})} \right| \right\}, \quad (3.26)$$

para  $\pi(\mathcal{M}_2, \tilde{\delta}_2 | y^*)$  e  $\pi(\mathcal{M}_1 | \delta_1, y)$  representam a distribuição *a posteriori* apresentada na expressão (3.26) do modelo proposto  $\mathcal{M}_2$  e do modelo atual  $\mathcal{M}_1$ , respectivamente, a função  $q_{(\mathcal{M}_2 \rightarrow \mathcal{M}_1)}$  representa a probabilidade de propor uma mudança do modelo  $\mathcal{M}_1$  para o modelo  $\mathcal{M}_2$  e o termo  $q_{d_{\mathcal{M}_1} \rightarrow d_{\mathcal{M}_2}}(\mathbf{u})$  denota a função de densidade do vetor aleatório  $\mathbf{u}$ , e  $|\cdot|$  representa o determinante da matriz do Jacobiano da mudança de estado de  $(\mathcal{M}_1, \delta_1)$  para o espaço de estado  $(\mathcal{M}_2, \tilde{\delta}_2)$ , ou Jacobiano, como é conhecido na literatura. O Jacobiano se faz necessário devido à mudança de variável quanto utilizamos a função de transição  $g_{\mathcal{M}_1 \rightarrow \mathcal{M}_2}(\delta_1, \mathbf{u})$  para a mudança de variáveis de  $(\delta_1, \mathbf{u}) \rightarrow \tilde{\delta}_2$ .

No caso de uma movimentação oposta, a probabilidade de aceitação do modelo  $\mathcal{M}_1$  frente ao modelo  $\mathcal{M}_2$ , (o movimento inverso)  $\mathcal{M}_2 \rightarrow \mathcal{M}_1$  é dada por,

$$\alpha \left[ (\mathcal{M}_2, \underline{\delta}_2), (\mathcal{M}_1, \underline{\delta}_1) \right] = \alpha \left[ (\mathcal{M}_1, \underline{\delta}_1), (\mathcal{M}_2, \underline{\delta}_2) \right]^{-1} \quad (3.27)$$

Outro aspecto a ser destacado é a questão da reversibilidade da cadeia de Markov nesse processo de transição de um modelo para outro.

Vale ressaltar que nesse trabalho iremos utilizar o reversible jump em dois momentos distintos, nessa ordem: em um primeiro momento para encontrar o melhor valor de K do modelo de associação, em um segundo momento, com K fixo e conhecido, para realizar o agrupamento e/ou desagrupamento de linhas (colunas) da tabela de contingência. Assunto a ser tratado na próxima seção.

### 3.8.3 Reversible Jump MCMC na seleção do melhor valor de K

No modelo de associação RC, é comum encontrarmos na literatura o uso dos métodos de MCMC para estimar os parâmetros desejados. Poucas publicações tratam de estimar o melhor valor do parâmetro K, e torna-se um desafio enorme encontrarmos qual seria esse valor.

No processo de estimação dos modelos RC(K), é inevitável a ocorrência de mudança para mais ou para menos, do número de parâmetros do modelo devido à mudança de K. Nesse sentido, o RJMCMC possibilita testarmos, de forma iterativa em cada rodada da simulação MCMC, qual o melhor valor do parâmetro K.

Ao lidarmos com o parâmetro K não fixo (em que,  $K=1, \dots, \min(I,J)-1$ ), nos deparamos com a necessidade de reavaliar os valores estimados de cada parâmetro do modelo de associação para diferentes valores de K. Nesse momento a utilização do método RJMCMC proposto por [Green \(1995\)](#) permite o salto de um modelo para outro e testar o melhor valor de K, obtido do modelo mais visitado.

Suponha que uma cadeia de Markov encontra-se no modelo  $\mathcal{M}_1$ , no espaço paramétrico de dimensão  $d_{\mathcal{M}_1}$ , no tempo t e nosso interesse é testar a transição para um modelo  $\mathcal{M}_2$ , no tempo t+1, no espaço paramétrico de dimensão  $d_{\mathcal{M}_2}$  e o subscripto represente a ordem K do modelo de associação. Uma sugestão do algoritmo do reversible jump com N iterações pode ser construído da seguinte forma:

1. No primeiro passo, defina  $k \in \{1, \dots, \mathcal{K}\}$  e  $\delta_k$  de forma que  $p(\delta_k|y, \mathcal{M}_k) > 0$ ;
2. Utilizar o movimento inter-modelo fixando  $\mathcal{M}_k$  e estimar os novos parâmetros  $\delta_k$  utilizando o algoritmo de MH descrito na [Subseção 3.8.1](#);

3. Propor um sorteio do valor de uma variável aleatória com distribuição Uniforme  $U[1,K]$  para definir qual o valor de  $k$  ( $k \in \{1, 2, \dots, \mathcal{K}\}$ , em que  $\mathcal{K} \leq \min((I-1), (J-1))$ );
4. Propor um movimento do estado  $(\mathcal{M}_k, \delta_k)$  do modelo  $\mathcal{M}_k$  para um novo estado  $(\mathcal{M}_{k'}, \delta_{k'})$  do modelo  $\mathcal{M}_{k'}$ , com uma probabilidade  $q(\mathcal{M}_k | \mathcal{M}_{k'})$ , em que  $k \neq k'$ ;
5. Equalizar os dimensões do espaço paramétrico, ou seja, gerar  $\mathbf{u} \sim q_{d_{\mathcal{M}_k} \rightarrow d_{\mathcal{M}_{k'}}}(\mathbf{u})$  com dimensão  $d_{\mathcal{M}_k \rightarrow \mathcal{M}_{k'}} = d_{\mathcal{M}_{k'}} - d_{\mathcal{M}_k}$ , para  $d_{\mathcal{M}_{k'}} \geq d_{\mathcal{M}_k}$ . Adotamos a distribuição Uniforme  $[0,1]$  para  $\mathbf{u}$ ;
6. Fazer  $\delta_{k'}' = g_{\mathcal{M}_k \rightarrow \mathcal{M}_{k'}}(\delta_k, \mathbf{u})$  (Calcular  $y = g_{kk'}(\delta_k, \mathbf{u}), y \in \delta_{k'}'$ );
7. Calcular a probabilidade de aceitar a proposta de mudança  $\alpha(\mathcal{M}_k, \mathcal{M}_{k'})$  dado pela Expressão (3.26);
8. Simular  $u \approx U(0, 1)$ ;
9. Aceite a mudança para o estado  $(\mathcal{M}_{k'}, \delta_{k'}')$  com probabilidade  $\alpha(\mathcal{M}_k, \mathcal{M}_{k'})$  calculada no passo (7), se  $u < \alpha(\mathcal{M}_k, \mathcal{M}_{k'})$  seguir com  $\mathcal{M}_{k'}$ , caso contrário, seguir com  $\mathcal{M}_k$ ;
10. Fazer  $t=t+1$ , enquanto  $t \leq N$ , voltar ao passo 2.

Após a simulação ser finalizada, a probabilidade *a posteriori* para cada modelo  $\mathcal{M}_k$  pode ser estimada pela proporção de aceitação de cada modelo,  $k \in \{1, \dots, \mathcal{K}\}$ ,

$$pr(\mathcal{M}_k | y) = \frac{\text{número de vezes que } \mathcal{M}_k \text{ foi aceito}}{\text{número de modelos visitados}}$$

O modelo  $\mathcal{M}_k$  que obtiver a maior proporção será o modelo final selecionado pelo RJMCMC. O valor de  $K$  é obtido nesse modelo final.

### 3.8.4 Aplicação do procrustes no modelo de associação

Condições de ortogonalidade para garantir a identificabilidade de modelos é uma das características da decomposição em valores singulares. Essas condições são dependentes das amostras utilizadas no cálculo da decomposição e isso acaba afetando os resultados e acarretar em distorções devido às restrições impostas na SVD. Outro problema que as distorções apresentam é a distorção da visualização gráfica correta de possíveis agrupamentos de valores resultantes das simulações. A proposta da utilização da análise de procrustes é necessária para minimizar os efeitos dessas distorções e fazer as correções das mudanças nas direções dos eixos dos valores no gráfico onde ocorrem essas distorções para cada resultado das simulações, pois mesmo que os valores estimados dos parâmetros permaneçam próximos do valor mais verdadeiro dos parâmetros, as estimativas de diferentes conjunto de dados acabam diferindo consideravelmente, isso é que chamamos de problemas de extraviabilidade nos diagramas reamostrados. No

momento em que são esboçados no gráfico em grande quantidade na mesma tela, a variabilidade extra das distorções se revela de três maneiras:

- Ocorre uma rotação dos eixos;
- Uma inversão da ordem (reordenação) dos valores singulares, quando dois ou mais valores singulares possuem grandezas similares;
- Uma reflexão nos eixos.

Uma consequência imediata da variabilidade extra é que a distribuição das replicações da simulação da variável de interesse observada está sujeita à distorção, isto é, muito tendenciosa ou muito sobredispersa. A única incerteza que temos interesse é a da variação amostral das estimativas dos parâmetros obtidas, estrutura com base em um modelo e erros aleatórios. Um segundo tipo de variação perturbadora que falamos acima é devido às restrições impostas no modelo para identificar as estimativas. Para minimizar os efeitos dessa variação é que a proposta de utilização da análise de procrustes servirá de filtro fazendo com que cada diagrama obtido em cada passo da simulação possa girar de forma a reduzir as variações impostas pelas restrições utilizadas (MILAN; WHITTAKER, 1995).

#### 3.8.4.1 Análise do procrustes no modelo de associação

A técnica do procrustes é aplicada a cada etapa da simulação no MCMC, da seguinte forma: dentro os primeiros resultados do processo de simulação, será escolhido um conjunto de valores dos parâmetros para servir como uma espécie de valores de referência (serão fixados), para os estimadores do parâmetro que compõem o modelo. A cada conjunto de novos valores das estimativas resultantes das simulações é feita uma comparação desses resultados com os valores de referências utilizando a norma de Frobenius para a diferença de duas matrizes:

$$\|U - V\|_F^2 = \sum_{i=1}^n \sum_{j=1}^p (u_{ij} - v_{ij})^2, \quad (3.28)$$

Vamos supor que  $\omega_0$  seja a matriz de coordenadas definida como de referência e  $\omega^{[s]}$  seja a s-ésima etapa de realização do processo de simulação. Para cada estimativa obtida,  $\omega^{[s]}$ , um conjunto  $W = \{\omega_z^{[s]}; z \in Z\}$  é comparado com a estimativa de referência. O conjunto  $Z$  pode ser finito ou infinito. A matriz de coordenadas,  $\omega_*^{[s]}$ , mais próxima da matriz de referência pela norma Frobenius, será considerado como conjunto identificado de estimativas. A ordem de  $\omega_0$  e  $\omega^{[s]}$  para o ajuste de um modelo bilinear é  $n \times M$  para exibir os pontos referentes à linhas da matriz, caso  $\omega$  corresponda a  $U\Sigma$ , ou  $p \times M$  para plotar os pontos da coluna, caso  $\omega$  corresponda a  $V\Sigma$  (lembrando que na DVS temos  $Y = U\Sigma V^T$ ). Vale ressaltar que os exemplos apresentados são para  $M = 2$ , mas as técnicas discutidas podem ser aplicadas para dimensões maiores que 2.

### 3.8.4.2 Movimento de rotação no procrustes

Essa última sequência de movimento da técnica de procrustes é obtida através de uma rotação das estimativas. Nessa etapa os pontos do gráfico identificados são  $\omega_*^{[s]} = \omega^{[s]}\tilde{\mathbf{Q}}$ , onde a matriz ortogonal representada por  $\tilde{\mathbf{Q}}$  gira os pontos (isto é, as linhas de  $\omega^{[s]}$ ) para aquela posição que está mais próxima do conjunto de pontos de referência. Essa técnica é conhecida como procrustes ortogonal e é utilizada para escolher a melhor rotação e pode ser descrita por,

$$\Delta = \min \|\omega_0 - \omega^{[s]}\mathbf{Q}\|_F^2 \quad \text{sujeito a } \mathbf{Q}^T\mathbf{Q} = \mathbf{I}, \quad (3.29)$$

onde  $\mathbf{Q}$  é uma matriz ortogonal de dimensão  $M \times M$  que gira os pontos para uma posição que mais se aproxima de  $\omega_0$ .

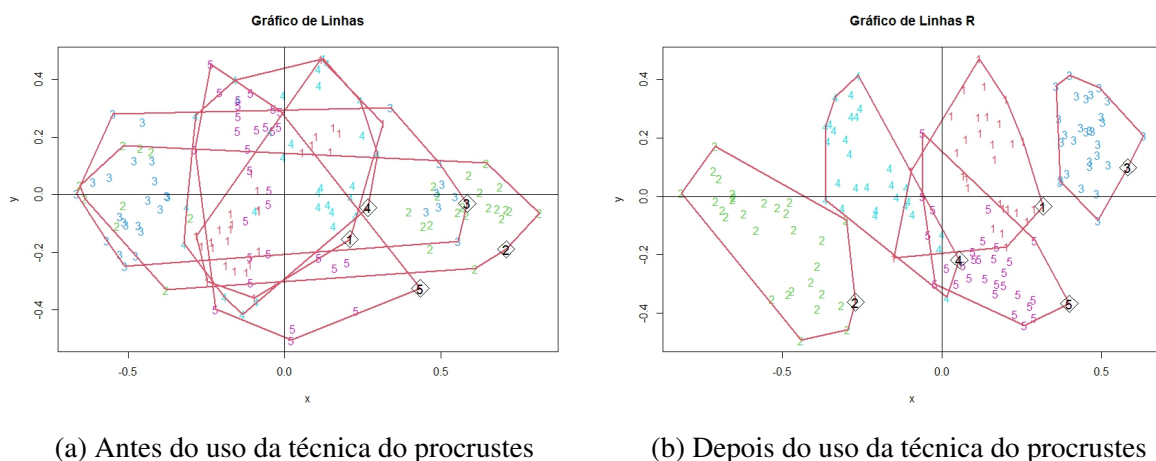
A solução para a Equação 3.29, a matriz  $\tilde{\mathbf{Q}}$ , é  $\tilde{\mathbf{Q}} = \mathbf{U}\mathbf{V}^T$  onde  $\mathbf{U}$  e  $\mathbf{V}$  são obtidos através do SVD e  $(\omega^{[s]})^T \omega_0 \mathbf{U} \Sigma_p \mathbf{V}^T$  (MILAN; WHITTAKER, 1995).

### 3.8.4.3 Exemplo gráfico da aplicação do procrustes

A aplicação da metodologia do procrustes no modelo de associação RC ocorrerá na etapa após o processo de simulação MCMC. Usaremos a primeira iteração da simulação pelo método MCMC (após o descarte ou *burn-in*) para selecionar um conjunto de estimativas para definirmos como amostra de referência. A partir dessa iteração, ao final de cada passo, o conjunto de valores estimados dos parâmetros dispostos na forma de matriz passará pela análise de procrustes e os resultados serão armazenados na forma de vetores para posteriormente representarmos na forma de biplot.

Para ilustrar a técnica do procrustes, vamos apresentar um exemplo com dados simulados antes de depois do movimento de rotação.

Figura 1 – Exemplo gráfico de pontos antes da aplicação do procrustes e depois da aplicação do procrustes



Fonte: Elaborada pelo autor.

Na [Figura 1\(a\)](#), notamos que os pontos estão todos espalhados pelo espaço bidimensional e nenhum padrão de agrupamento entre as variáveis é visível. Já na [Figura 1\(b\)](#), após a aplicação da técnica do procrustes, os pontos ficaram bem próximos, formando várias nuvens de pontos, sendo possível distinguir grupos formados nos quatro quadrantes do gráfico. No exemplo em questão podemos ver que o uso da técnica do procrustes foi essencial e preciso na detecção dos agrupamentos.

### 3.8.5 Aplicação do biplot no modelo de associação

A técnica para a construção do biplot não requer um algoritmo muito complexo, pode ser desenvolvido em qualquer plataforma de programação que gera resultados gráficos. Em [Udina \(2005\)](#), por exemplo, há uma descrição da construção de algoritmos para a construção de vários tipos de biplot, com formatos e qualidade de impressão ajustáveis. Utilizando a técnica da decomposição de valores singulares e o dimensionamento dos rótulos após a decomposição, em [Lipkovich e Smith \(2002\)](#) encontramos uma descrição de um conjunto de macros para o Microsoft *Excel*<sup>®</sup> que podem ser usadas para construir um gráfico biplot utilizando resultados de várias técnicas estatísticas multivariadas tais como a análise de componentes principais, análise de correlação canônica, análise discriminante, escalonamento multidimensional e análise de correspondência.

No estudo de [Salinas et al. \(2013\)](#) com a utilização do biplot em dados de indicadores de ciência e tecnologia na avaliação do desempenho das 25 universidades mais bem classificadas no ranking mundial baseada nas variáveis: Pesquisa (volume, receita e reputação), Citação (Influência em pesquisa), Perspectiva internacional (funcionários, alunos e pesquisadores) e Ensino (ambiente de aprendizagem), que são utilizadas pelo The World University Ranking, mostrou que citações e perspectiva internacional não se correlacionam com pesquisa e ensino; que pesquisa e ensino são altamente correlacionadas. E também encontrou quatro clusters distintos de universidades: um formado por universidades com valores mais elevados no ensino e pesquisa (por exemplo, Harvard, Yale, MIT), melhor desempenho em citações (Pennsylvania e California LA), as universidades que apresentam o desempenho mais baixo em todos os indicadores (Duke, Cornell e Michigan) e, no último cluster, formado por universidades caracterizadas pelo alto valor no cenário internacional (University College, Imperial College, Toronto e British Columbia).

### 3.8.6 Exemplo de um biplot

O primeiro passo para a construção de um biplot é decompor a matriz original como produto de duas outras matrizes que sejam ortogonais. Uma alternativa para obter essas matrizes é a decomposição em valores singulares (SVD, já apresentado anteriormente), que também foi utilizada por [Gabriel \(1971\)](#). Para ilustrar o método, vamos considerar como referência a matriz utilizada nessa pesquisa.

Suponha uma matriz  $\mathbf{A}$  de ordem  $4 \times 3$ , de posto dois, cuja forma fatorada seja

$$\mathbf{A} = \mathbf{U} \mathbf{V}^T$$

$$\mathbf{A} = \mathbf{U} \mathbf{V}^T$$

$$\begin{bmatrix} 2 & 2 & -4 \\ 2 & 1 & -3 \\ 0 & -3/2 & 3/2 \\ -1 & -1/2 & 3/2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 1 \\ 0 & -3/2 \\ -1 & -1/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

A matriz  $\mathbf{U}$  é composta por quatro vetores  $[u_1, \dots, u_4]$  e representa os casos da matriz  $\mathbf{A}$  e a matriz  $\mathbf{V}$  é composta de 3 vetores  $[v_1, \dots, v_3]$  e representa as variáveis ou colunas da matriz  $\mathbf{A}$ . Cada vetor representa um par ordenado que pode ser posicionado como um ponto no gráfico biplot.

Para quaisquer dois vetores,  $\vec{u}$  e  $\vec{v}$ , por exemplo, com  $n$  elementos cada um, o produto escalar entre  $u$  e  $v$  pode ser escrito como  $\vec{u}^T \vec{v} = u_1 v_1 + u_2 v_2, \dots, u_n v_n$ .

Portanto, cada elemento de  $\mathbf{A}$  é resultado do produto escalar entre os vetores  $u$  e  $v$ . Por exemplo,  $A_{11} = u_1 v_1 = [2 \ 2] [1 \ 0]^T = 2$ ,  $A_{23} = u_2 v_3 = [2 \ 1] [-1 \ -1]^T = -3$  e assim sucessivamente.

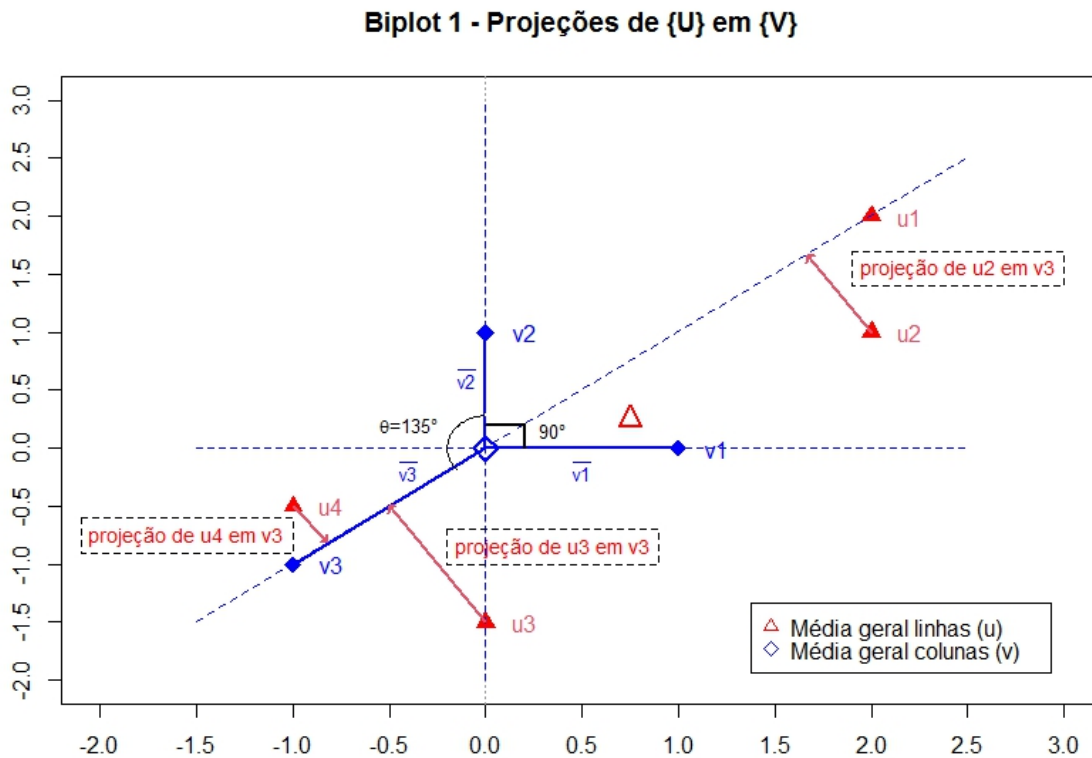
Na [Figura 2](#) temos uma representação da matriz  $\mathbf{A}$  no biplot. Os pontos na cor em vermelho em formato de triângulos são os vetores da matriz  $\mathbf{U} = [u_1, u_2, u_3, u_4]$  que representam as categorias da variável disposta nas linhas de  $\mathbf{A}$ , os pontos na cor azul (em formato do símbolo losango) são os vetores da matriz  $\mathbf{V}^T = [v_1, v_2, v_3]$ , que representam as categorias da variável disposta nas colunas da matriz  $\mathbf{A}$ .

A localização dos pontos no gráfico (em vermelho) indica que quanto mais próximos esses pontos estiverem um do outro, mais similares serão as categorias (da variável) representadas por esses pontos. Outro resultado do biplot a destacar é a indicação de possíveis interações entre linhas e colunas da matriz. Pontos concentrados na região na direção de um segmento de reta podem indicar a existência de associação entre as variáveis da matriz.

Os segmentos de reta (cor azul) na figura foram construídos a partir da origem, coordenadas  $(0,0)$ , e conectam a cada ponto do vetor  $\mathbf{V}$ . Por exemplo, o segmento de reta  $\vec{v}_1$  é formado pela linha que conecta a origem e o ponto da coordenada  $v_1=(1,0)$ ,  $\vec{v}_2$  é a linha entre a origem e o ponto da coordenada  $v_2=(1,0)$  e o segmento de reta  $\vec{v}_3$  é a linha entre a origem e o ponto da coordenada  $v_3=(-1,-1)$ .

Em um biplot, o comprimento de um segmento de reta indica uma medida de variabilidade da categoria da variável. O cosseno do ângulo formado entre dois segmentos de reta fornece

Figura 2 – Exemplo de um gráfico biplot das projeções da matriz U na matriz V



Fonte: Elaborada pelo autor.

uma medida da correlação entre essas categorias. Ângulos retos indicam ausência de correlação, enquanto obtusos e agudos indicam correlação negativa e positiva, respectivamente (SALINAS *et al.*, 2013).

No exemplo do biplot, note que o ângulo formado pelos segmentos de reta  $\bar{v}_1$  e  $\bar{v}_2$  forma  $90^\circ$ , portanto a correlação entre essas duas categorias é zero. No caso dos segmentos  $\bar{v}_2$  e  $\bar{v}_3$ , o ângulo formado é obtuso, mede  $130^\circ$ , portanto, a associação entre essas duas variáveis é negativa com valor  $-0.7071 = \cos(135^\circ)$ .

Os vetores (cor vermelha) perpendiculares ao segmento  $\bar{v}_3$  representam a projeção dos pontos (que são as linhas da matriz *A*) no segmento de reta  $\bar{v}_3$ . Essas projeções são ortogonais, ou seja, formam um ângulo de  $90^\circ$  com  $\bar{v}_3$ , e os valores da matriz original podem ser recuperados utilizando essa projeção.

A distância do segmento formado entre o ponto O e a projeção do ponto *u*<sub>2</sub> no segmento  $\bar{v}_3$  multiplicado pela distância do segmento  $\bar{v}_3$  é igual ao valor correspondente da matriz original *A*. Vamos a um exemplo, suponha que a distância entre o ponto O e o ponto *v*<sub>3</sub> (comprimento de  $\bar{v}_3$ ) é a norma de *v*<sub>3</sub>, ou seja,  $\|v_3\| = \sqrt{(-1)^2 + (-1)^2} = 1.4142$ . A distância entre a projeção do ponto *u*<sub>2</sub> no segmento  $\bar{v}_3$  e a origem é obtida através da Equação: projeção(*u*<sub>2</sub>) =  $\|u_2\| \cos(\theta)$ ,

onde  $\theta$  é o ângulo formado pelos segmentos  $u_2$  e  $v_3$ . Para encontrar o valor de  $\cos(\theta)$  usamos a expressão,  $\cos(\theta) = \frac{u_2 v_3^T}{\|u\| \|v\|}$ . Substituindo os valores, encontramos  $\cos(\theta) = -0.9487$ , então a projeção( $u_2$ ) =  $\sqrt{5} - 0.9487 = -2.1213$ . Portanto, o valor de  $A_{2,3} = 1.4142 * (-2.1213) = -3$ . Da mesma forma, todos os valores da terceira coluna de  $A_{ij}$  podem ser obtidos. Para obter os valores das outras colunas, é preciso fazer as projeções de todos os pontos de  $\{U\}$  nos segmentos  $\overline{v_1}$  e  $\overline{v_2}$ .

As combinações lineares entre linhas e colunas podem ser visualizadas no biplot utilizando os valores médios de cada variável calculados da seguinte forma (BRADU; GABRIEL, 1978):

$$u_{.1} = \frac{1}{I} \sum_{i=1}^I u_{i1} = \frac{1}{4}(2+2+0-1) = \frac{3}{4} \quad e \quad u_{.2} = \frac{1}{I} \sum_{i=1}^I u_{i2} = \frac{1}{4}(2+1-\frac{3}{2}-\frac{1}{2}) = \frac{1}{4}$$

e

$$v_{1.} = \frac{1}{J} \sum_{j=1}^J v_{1j} = \frac{1}{3}(1+0-1) = 0 \quad e \quad v_{2.} = \frac{1}{J} \sum_{j=1}^J v_{2j} = \frac{1}{3}(0+1-1) = 0$$

e através do produto escalar obter as médias das linhas, colunas e média geral, respectivamente,

$$(i) \quad a_{i.} = u_i' v. \implies \text{p.e. } a_{1.} = \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \quad \text{que corresponde à média da linha } A_{1.};$$

$$(ii) \quad a_{.j} = u' v_j \implies \text{p.e. } a_{.1} = \begin{bmatrix} \frac{3}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{3}{4} \quad \text{que corresponde à média da coluna } A_{.1};$$

$$(iii) \quad a_{..} = u' a. \implies \text{p.e. } a_{..} = \begin{bmatrix} \frac{3}{4} & \frac{3}{4} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0 \quad \text{que corresponde à média geral da matriz } A_{..}$$

A nossa proposta de trabalho nesse tema vai utilizar os resultados das estimativas dos parâmetros do modelo de associação RC(K) obtidos através das simulações MCMC e construir o gráfico biplot.

### 3.8.7 Regiões de credibilidade no modelo de associação

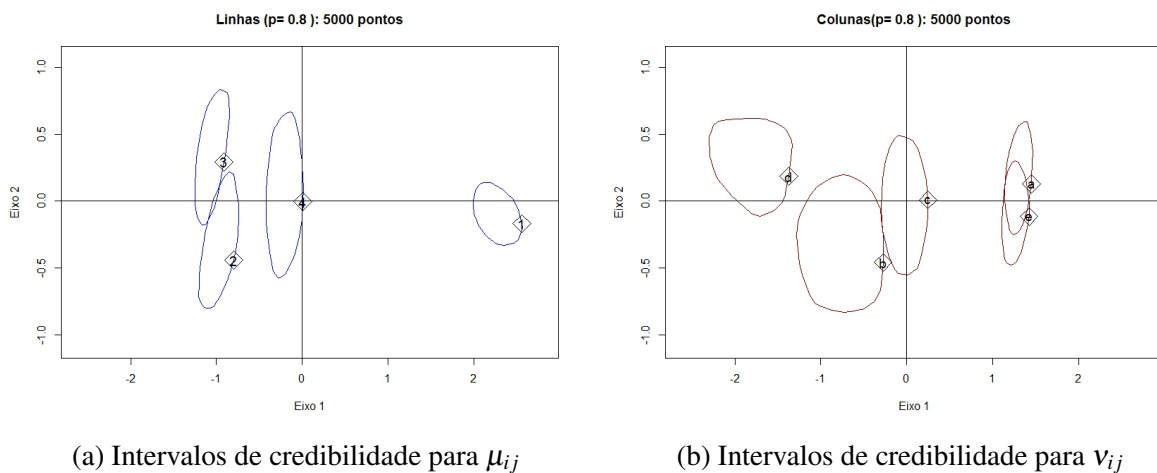
Problemas de incertezas nas exposições da análise de variáveis bidimensionais podem ocorrer em situações onde as estimativas são obtidas por meio de métodos de simulação. Em casos dessa natureza, é recomendável traçar círculos de tolerância em torno dos pontos, centrados na média desse conjunto de estimativas. Sob a suposição de normalidade das estimativas, esse procedimento faz com que a variação amostral das estimativas seja ignorada e os círculos de tolerância podem ser considerados com regiões de confiança ou de credibilidade (MILAN; WHITTAKER, 1995).

No modelo de associação RC vamos produzir regiões de credibilidade em torno das estimativas dos parâmetros ou nuvens de pontos, que serão originadas após a aplicação da

metodologia do procrustes. Essas regiões irão permitir, por exemplo, visualizar a intersecção das nuvens de pontos no gráfico bidimensional e conseqüentemente sugerir indícios variáveis semelhantes, agrupamentos ou diferenças entre variáveis. Em conjunto com o biplot, os resultados gráficos atuarão como um suporte nas análises das estimativas dos parâmetros do modelo de associação RC.

A Figura 3 ilustra dois exemplos gráficos de intervalos de credibilidade de 90% para uma tabela de contingência de ordem  $4 \times 5$ , obtida através de dados simulados, após a aplicação da metodologia de prócrustes.

Figura 3 – Gráfico exemplificativo de regiões de credibilidade construídos após a aplicação do procrustes



Fonte: Elaborada pelo autor.

---

## SELEÇÃO DE MODELOS

---

### 4.1 Aplicação

Nesse capítulo iremos tratar da aplicação da proposta de três formas distintas. Primeiro iremos aplicar a metodologia do trabalho em uma tabela de contingência simulada utilizando parâmetros conhecidos. Em um segundo momento, aplicaremos o modelo em uma tabela de contingência utilizada por outros autores, para fins de comparação dos resultados e, por fim, aplicaremos em três conjunto de dados reais na área de educação.

Na etapa de teste do algoritmo desenvolvido para atender as propostas do trabalho usaremos inicialmente uma tabela de contingência simulada com parâmetros da distribuição Poisson, previamente conhecidos, para a validação do código computacional desenvolvido e do modelo proposto. A justificativa é que diante do uso de parâmetros conhecidos, esperamos que o código desenvolvido possa espelhar a metodologia proposta do trabalho, e assim obtermos os resultados das estimativas simuladas próximas das estimativas reais utilizadas para gerar a tabela de contingência. Com o mesmo objetivo, iremos aplicar o método em um conjunto de dados já utilizado em trabalhos anteriores e comparar os resultados para sabermos se o código computacional está adequado para utilização em dados reais.

A [Tabela 4](#) apresenta os parâmetros utilizados para simular uma tabela de contingência de dimensão 4x5. Por simplificação, optamos por não colocar os valores de  $\phi$ 's na tabela, cujos valores foram fixados iguais a 1. Nesse exemplo, utilizamos um valor fixo para  $K=2$ .

Tabela 4 – Valores dos parâmetros utilizados para geração da tabela de contingência

$\lambda = 4,0$	$\lambda_1^C = -0,6$	$\mu_{11} = 0,86$	$\mu_{21} = 0,5$	$v_{11} = 0,55$	$v_{21} = -0,02$
$\lambda_1^R = -0,5$	$\lambda_2^C = -0,4$	$\mu_{21} = -0,34$	$\mu_{22} = 0,5$	$v_{12} = -0,29$	$v_{22} = 0,76$
$\lambda_2^R = -0,3$	$\lambda_3^C = 0,2$	$\mu_{31} = -0,37$	$\mu_{32} = 0,5$	$v_{13} = 0,07$	$v_{23} = 0,52$
$\lambda_3^R = 0,2$	$\lambda_4^C = 0,3$	$\mu_{41} = -0,06$	$\mu_{42} = 0,5$	$v_{14} = -0,55$	$v_{24} = 0,02$
$\lambda_4^R = 0,6$	$\lambda_5^C = 0,5$			$v_{15} = 0,53$	$v_{25} = 0,39$

Fonte: Elaborada pelo autor.

A tabela de contingência (Tabela 5) foi construída utilizando um algoritmo que desenvolvemos na linguagem R Core Team (2022), que simula uma distribuição Poisson para cada casela (i,j), utilizando os parâmetros descritos na Tabela 4. O resultado obtido é a tabela de contingência,

Tabela 5 – Tabela de contingência simulada para validar o código computacional

Categorias das linhas	Categorias das colunas					Total
	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5	
Linha 1	577	3	22	2	858	1.462
Linha 2	3	31	21	208	16	279
Linha 3	12	20	22	437	16	507
Linha 4	53	26	57	159	73	368
Total	645	80	122	806	963	2.616

Fonte: Elaborada pelo autor.

Os dados da tabela de contingência (Tabela 5) foram submetidos ao programa computacional desenvolvido nessa pesquisa. Para obtermos as estimativas dos parâmetros implementamos o algoritmo de Metropolis-Hastings. Utilizamos a distribuição Normal como densidade geradora de candidatas. Foram realizadas 5.000 simulações, com descarte (burn-in) dos 500 primeiros resultados logo após aplicar saltos de tamanho 50, para reduzir a autocorrelação. Para cada uma das 5.000 iterações, o algoritmo aplica o que chamamos de “delay”, ou retardo, isto é, a rotina calcula 50 estimativas mas retorna apenas o último valor calculado, totalizando 250.000 iterações.

A Tabela 6 apresenta comparação entre os valores dos parâmetros conhecidos e os estimados através da simulação MH, pela mediana (que representa o valor mais concentrado da nuvem de pontos) dos valores da distribuição *à posteriori*.

Tabela 6 – Resultado da comparação entre as estimativas obtidas por metodologias distintas

	Parâmetro utilizado	Método proposto			Parâmetro utilizado	Método proposto	
	Real	$IC_{0.025}$	$IC_{0.975}$		Real	$IC_{0.025}$	$IC_{0.975}$
$\lambda$	4,00	( 3.206	3.768 )	$\mu_{12}$	-0,437	(-0.313	0.201 )
$\lambda_1^R$	-0,50	(-0.918	0.392 )	$\mu_{22}$	-0,328	(-0.892	0.060 )
$\lambda_2^R$	-0,30	(-0.734	0.049 )	$\mu_{32}$	-0,613	(-0.253	0.858 )
$\lambda_3^R$	0,20	(-0.478	0.317 )	$\mu_{42}$	-0,569	(-0.553	0.718 )
$\lambda_4^R$	0,60	( 0.265	0.911 )	$v_{11}$	0,557	( 1.127	1.530 )
$\lambda_1^C$	-0,60	(-0.438	0.472 )	$v_{12}$	-0,296	(-1.402	0.319 )
$\lambda_2^C$	-0,40	(-1.736	-0.444 )	$v_{13}$	0,071	( 0.335	0.297 )
$\lambda_3^C$	0,20	(-0.610	0.168 )	$v_{14}$	-0,557	(-2.350	-1.307 )
$\lambda_4^C$	0,30	(-0.138	1.091 )	$v_{15}$	0,533	( 1.042	-1.446 )
$\lambda_5^C$	0,50	( 0.196	0.933 )	$v_{21}$	-0,019	(-0.329	0.768 )
$\mu_{11}$	0,861	( 1.957	2.646 )	$v_{22}$	0,757	(-0.839	0.271 )
$\mu_{21}$	-0,340	(-1.247	-0.730 )	$v_{23}$	0,520	(-0.579	0.570 )
$\mu_{31}$	-0,372	(-1.264	-0.845 )	$v_{24}$	0,019	(-0.088	0.628 )
$\mu_{41}$	-0,063	(-0.439	0.033 )	$v_{25}$	0,392	(-0.631	0.413 )

Fonte: Elaborada pelo autor.

Nota –  $IC_{0.025}$  e  $IC_{0.975}$  são os intervalos de credibilidade de 0.025 e 0.975, respectivamente.

Os resultados da [Tabela 6](#) mostram que as estimativas obtidas pelo método proposto estão próximas dos valores reais. Grande parte dos parâmetros estimados, estão contidos no intervalo de credibilidade (quantil 0.025 e 0.975). Embora a casos em que o valor real está fora do intervalo de credibilidade, uma possível explicação está no fato de que os valores reais estão submetidos a diferente regra de identificabilidade, a referência para o procrustes não são os valores reais dos parâmetros, que em geral não estão disponíveis, mas sim a primeira ocorrência dos parâmetros estimados na sequência de simulações. Feitas essas considerações, inferimos que o método está funcionando corretamente.

Os gráficos “trace plots” para todos os parâmetros poderão ser vistos na seção [Seção D.1](#) do [Apêndice D](#). Uma primeira inspeção visual desses gráficos “trace plots” sinalizam evidências da convergência do método, a média está estabilizada e poucos pontos estão fora dos limites de 5% e 95% dos dados.

A avaliação da convergência foi realizada através da estatística de [Gelman e Rubin \(1992\)](#). A tabela com os resultados, bem como os gráficos podem ser encontrados na sessão [Seção D.2](#) do [Apêndice D](#).

As distribuições *a posteriori* para os parâmetros estimados apresentaram um comportamento regular, com forma semelhante à de uma distribuição normal e sugerindo indícios de que as estimativas estão dentro dos padrões aceitáveis. Para fins de conferência, todos os gráficos das autocorrelações e das distribuições podem ser encontrados na [Seção D.3](#) e [Seção D.4](#) do

### Apêndice D.

Antes de entrarmos na etapa de visualização gráfica, precisamos checar se a ordem do modelo de associação é a mesma que utilizamos para gerar os dados. Nesse caso, aplicamos os algoritmos de RJMCMC que desenvolvemos, num total de 5.000 iterações com descarte dos 500 primeiros valores e salto de 20 unidades para reduzir a autocorrelação.

Ao término da simulação do RJMCMC, o modelo mais visitado foi que apresentou um valor de  $K = 2$ , com 69% do total de visitas, como mostra a [Tabela 7](#). Esse valor coincide com o mesmo valor de  $K$  que foi utilizado para gerar a amostra. Sendo assim, podemos concluir que para essa amostra, o método e o algoritmo para detecção do valor de  $K$  estão funcionando corretamente.

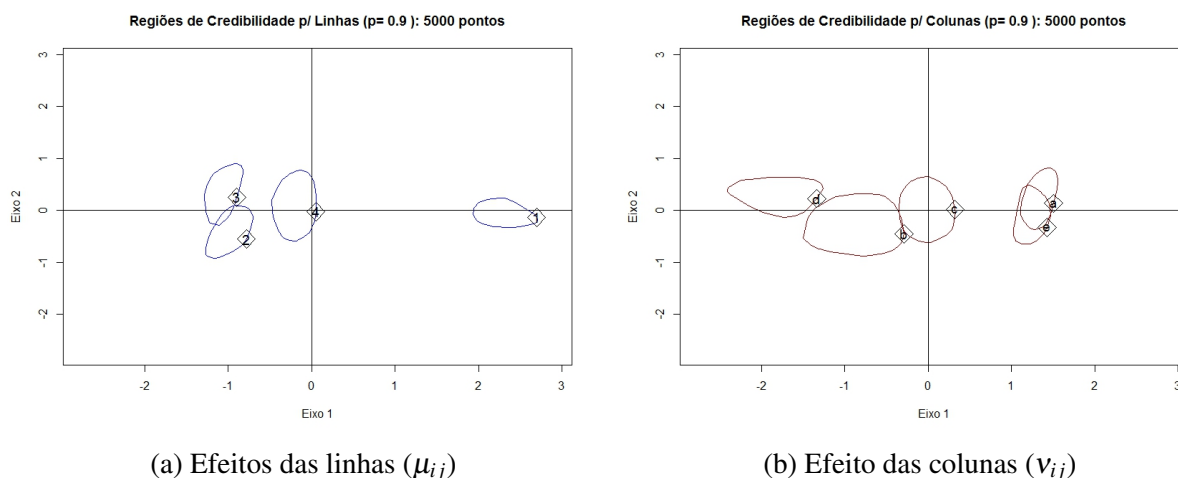
Tabela 7 – Visitas dos modelo para estimar  $K$

Valores de $K$	0	1	2	3
Quantidade de visitas	0	324	3.443	1.234
Proporção de visitas	0%	6%	69%	25%

Fonte: Elaborada pelo autor.

Constatado que o valor de  $K=2$ , temos as ferramentas para discutir a abordagem gráfica do modelo de associação. Com relação às regiões de credibilidade calculadas para a tabela de contingência simulada, as [Figura 4 \(a\)](#) e [Figura 4 \(b\)](#) mostram as regiões confiança tanto para os dados das linhas quanto para as colunas, respectivamente.

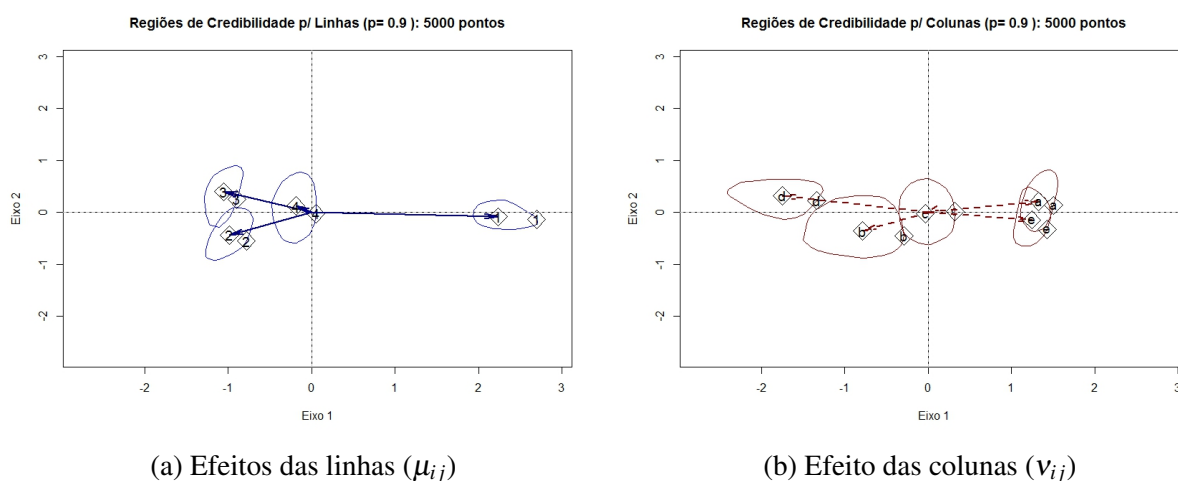
Podemos observar nessas imagens que além das regiões de confiança não estarem bem distribuídas no espaço bidimensional, elas estão concentradas ao logo do eixo  $x$  e parece não haver sinais de agrupamento a serem checados.

Figura 4 – Representação gráfica das regiões de credibilidade de  $\mu_{ij}$  e  $v_{ij}$ 

Fonte: Elaborada pelo autor.

Uma breve inspeção visual da Figura 4, pela nossa percepção é que com uma precisão de 80%, provavelmente nenhuma das linhas da Figura 4 (a) serão agrupadas, por outro lado, com as colunas da Figura 4 (b), temos um forte indício de que a primeira e a última coluna poderão ser agrupadas, pois as regiões estão interceptando-se quase que totalmente.

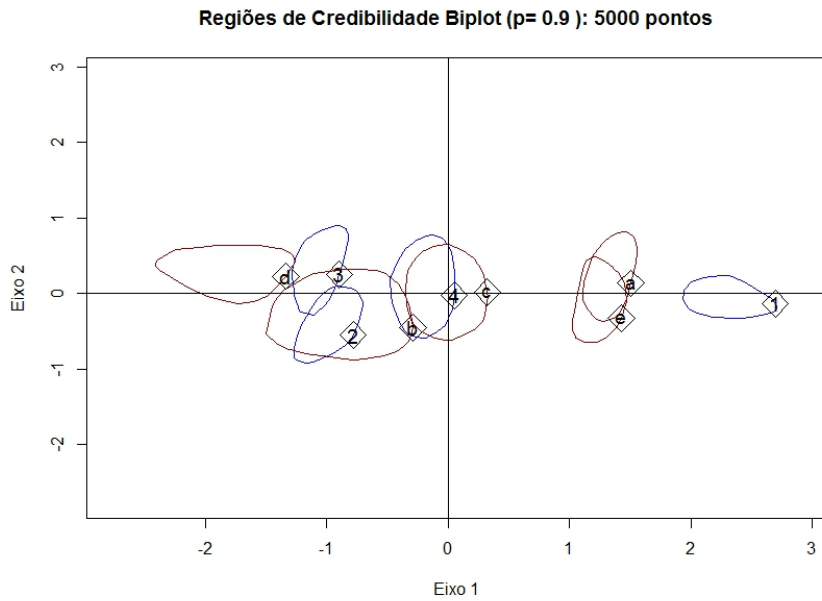
Ainda através dos mesmos gráficos, adicionamos segmentos de reta na forma de vetores partindo da origem até o centro de cada região e com a mesma escala da Figura 4.

Figura 5 – Representação gráfica das regiões de credibilidade de  $\mu_{ij}$  e  $v_{ij}$  com segmentos de retas

Fonte: Elaborada pelo autor.

Temos duas formas de representação do biplot, uma na forma de regiões de credibilidade e outra utilizando apenas os segmentos de retas (conforme já visto separadamente na figura anterior). No primeiro caso, temos a Figura 6 como resultado.

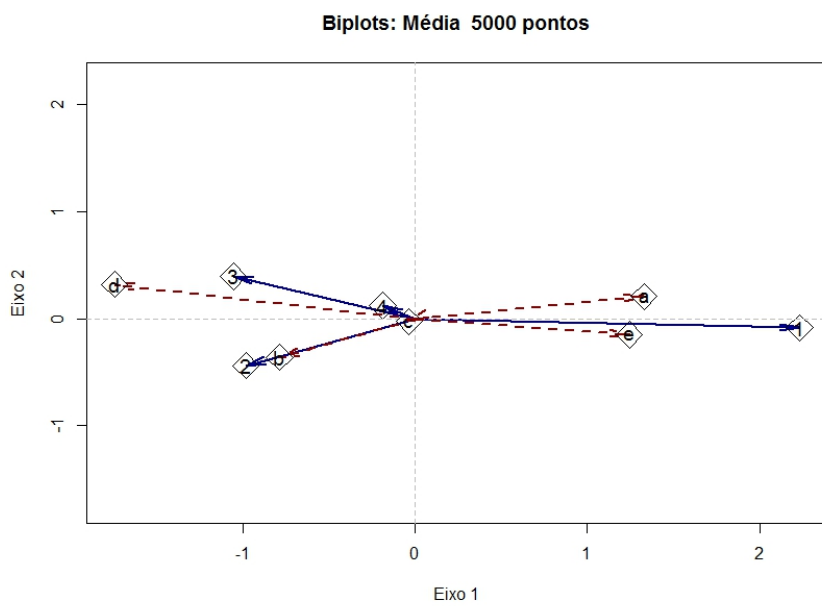
Figura 6 – Biplot para a amostra



Fonte: Elaborada pelo autor.

Para o segundo caso, a representação do biplot através de segmentos de reta está na [Figura 7](#). Essa forma de representação nos permite fazer inferências a respeito da associação existente entre linhas e colunas da tabela de contingência simulada.

Figura 7 – Biplot para a amostra



Fonte: Elaborada pelo autor.

Na [Figura 7](#) os segmentos na cor azul representam as estimativas da categoria das linhas da tabela de contingência e na cor vermelha representam as estimativas das categorias das colunas. Em termos de associação, por exemplo, temos uma forte associação entre as linhas 2 e coluna 2 da tabela de frequência (o ângulo formado entre esses dois segmentos é muito próximo de zero). As colunas 2 e 5 também possuem uma associação significativa com a linha 1. A linha 1 e a coluna 4 possuem associação negativa (ângulo de quase 180 graus, cujo cosseno está próximo de -1). A linha 1 possui a maior dispersão dos dados seguido da coluna 4.

#### 4.1.1 Comparação do desempenho do método proposto com uma metodologia conhecida (um estudo de caso)

Para compararmos o desempenho dos resultados obtidos através da metodologia proposta nesse trabalho, vamos aplicar o método em uma tabela de contingência de dimensão  $5 \times 5$  ( $I=5$ ,  $J=5$ ), originados do artigo de [Wermuth e Cox \(1998\)](#), que são resultados de pesquisas com adultos na Alemanha Ocidental, em que as contagens apresentadas foram obtidas pelo cruzamento do tipo de escolaridade forma e a faixa etária de todos os entrevistados durante os anos de 1991 e 1992. Esse mesmo exemplo também foi utilizado por [Kateri e Iliopoulos \(2003\)](#), [Jackson, Gray e Fienberg \(2008\)](#) e [Kateri \(2014a\)](#). As quantidades se encontram na [Tabela 8](#) a seguir.

Tabela 8 – Classificação de adultos de acordo com a faixa etária e escolaridade

Escolaridade	Faixa etária					Total
	18–29	30–44	45–59	60–74	> 74	
Básico incompleto	12	13	12	20	7	64
Básico completo	215	507	493	460	137	1.812
Ensino Médio	277	300	192	126	38	933
Médio superior	52	91	47	15	6	211
Intensivo	233	225	102	74	19	653
Total	789	1.136	846	695	207	3.673

Fonte: Adaptada de [Kateri e Iliopoulos \(2003\)](#).

Inicialmente os dados da tabela de contingência ([Tabela 8](#)) foram submetidos ao algoritmo de MH e os resultados foram comparados com os publicados no artigo de [Kateri e Iliopoulos \(2003\)](#). A [Tabela 9](#) apresenta a comparação dos valores estimados obtidos utilizando as duas metodologias. Por questão de visualização, iremos omitir as estimativas dos efeitos linhas e colunas, porque os resultados do estudo não contemplam essas estimativas, o que impede de realizar essas comparações.

Os resultados da [Tabela 9](#) indicam que os métodos são semelhantes, os valores das estimativas dos efeitos das linhas e das colunas então contidos, em sua maior parte, dentro do intervalo de credibilidade e não apresentam grandes discrepâncias significativamente grandes. O valor de K resultante da aplicação da metodologia proposta para os dados do artigo foi de  $K=2$ ,

Tabela 9 – Resultado da comparação entre as estimativas obtidas por metodologias distintas

	Método al- ternativo	Método proposto			Método al- ternativo	Método proposto	
	Mediana	$IC_{0.025}$	$IC_{0.975}$		Mediana	$IC_{0.025}$	$IC_{0.975}$
$\mu_{11}$	-0,575	(-1.186	0.140 )	$\nu_{11}$	0,529	( 0.552	1.103 )
$\mu_{21}$	-0,484	(-0.938	-0.381 )	$\nu_{12}$	0,428	( 0.083	0.710 )
$\mu_{31}$	0,168	(-0.065	0.528 )	$\nu_{13}$	0,073	(-0.445	0.264 )
$\mu_{41}$	0,517	( 0.009	0.998 )	$\nu_{14}$	-0,520	(-0.910	0.228 )
$\mu_{51}$	0,374	( 0.176	0.827 )	$\nu_{15}$	-0,511	(-1.111	0.068 )
$\mu_{12}$	-0,431	(-0.533	0.614 )	$\nu_{21}$	-0,684	( 0.072	0.694 )
$\mu_{22}$	0,605	(-0.642	0.155 )	$\nu_{22}$	0,400	(-0.507	0.327 )
$\mu_{32}$	-0,209	(-0.259	0.522 )	$\nu_{23}$	0,573	(-0.685	0.274 )
$\mu_{42}$	0,467	(-0.634	0.276 )	$\nu_{24}$	-0,181	(-0.441	0.459 )
$\mu_{52}$	-0,432	(-0.112	0.618 )	$\nu_{25}$	-0,107	(-0.606	0.490 )

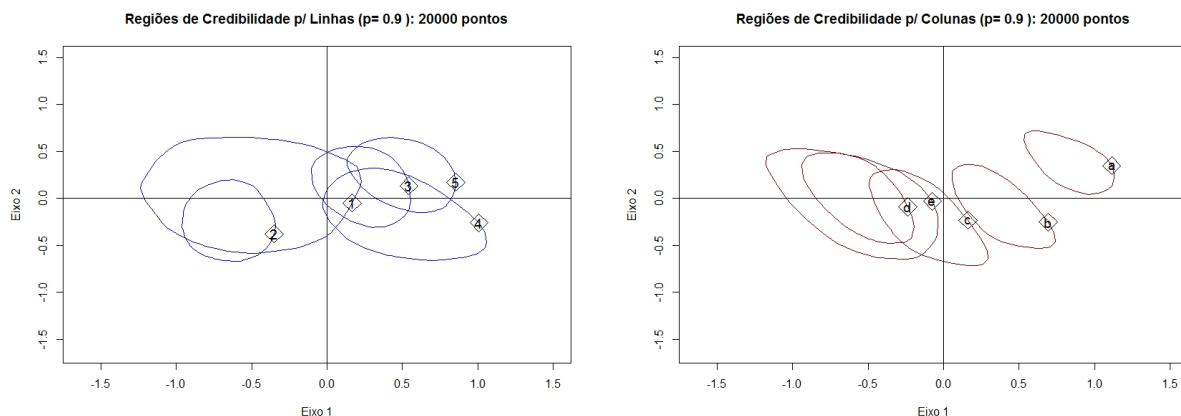
Fonte: Elaborada pelo autor.

Nota –  $IC_{0.025}$  e  $IC_{0.975}$  são os intervalos de credibilidade de 0.025 e 0.975, respectivamente.

o mesmo valor que foi considerado no artigo. Logo sugere-se que a metodologia proposta está consistente para esses dados comparativos.

Em termos de agrupamento de linhas e colunas, de acordo com o trabalho de [Kateri e Iliopoulos \(2003\)](#), o modelo alternativo sugeriu o agrupamento das linhas 1 e 2 e das colunas 4 e 5. Na metodologia proposta, o mesmo agrupamento também foi sugerido pelo gráfico do biplot. Isso indica que a metodologia proposta é válida. Graficamente, podemos visualizar esses resultados através da [Figura 8](#) a seguir,

Figura 8 – Representação gráfica das regiões de credibilidades de  $\mu_{ij}$  e  $\nu_{ij}$  para o modelo estimado

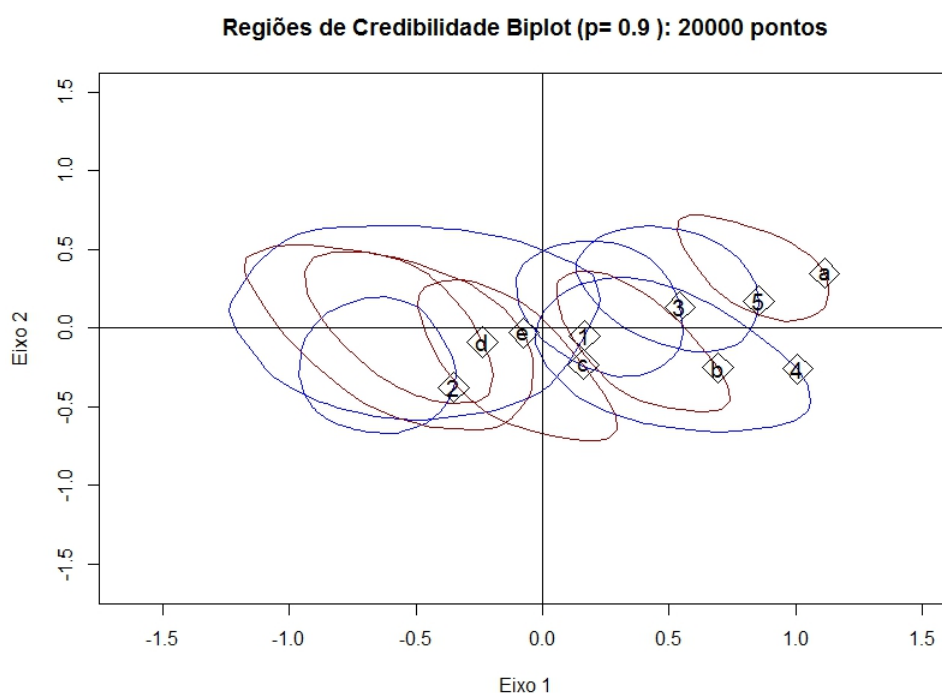


(a) Regiões de credibilidade para as linhas de  $\mu_{ij}$       (b) Regiões de credibilidade para as colunas de  $\nu_{ij}$

Fonte: Elaborada pelo autor.

A representação gráfica do biplot nesse caso é mostrada na [Figura 9](#). Como as regiões estão concentradas umas próximas das outras, a análise de associação pode ser mais complicada e como alternativa, vamos utilizar o biplot com os segmentos de retas, como o apresentado na [Figura 10](#). Embora os artigos não tragam uma análise mais pontual das associações, faremos aqui alguns comentários que achamos pertinentes.

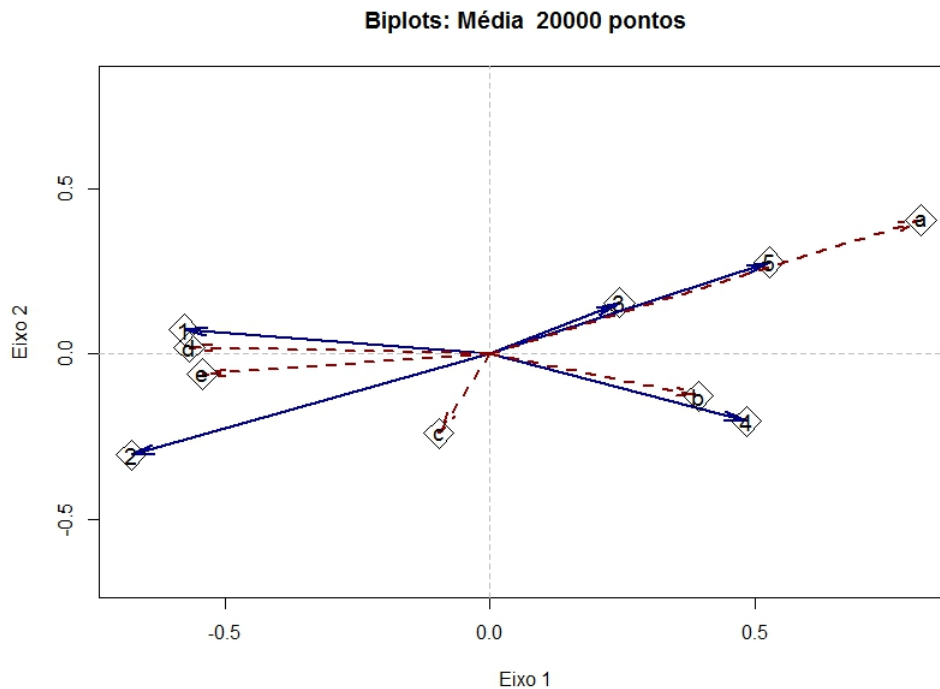
Figura 9 – Representação gráfica do biplot para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado comparativo



Fonte: Elaborada pelo autor.

A [Figura 9](#) indica que há uma associação forte entre as linhas 3 e 5 com a coluna 1 (escolaridades ensino médio e intensiva e a faixa etária de 18 a 29 anos), também entre o ensino médio superior com a faixa etária de 30 a 44 anos. O ensino básico completo está associado negativamente com a faixa etária de 18 a 29 anos. Há associação entre o ensino básico incompleto com adultos acima de 60 anos.

Figura 10 – Representação gráfica do biplot para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado comparativo



Fonte: Elaborada pelo autor.

O estudo comparativo serviu como balizador da metodologia proposta bem como para a validação dos códigos computacionais. Novamente, da mesma forma que a aplicação utilizando dados simulados, os resultados se mostraram satisfatórios, esse fato reitera a adequabilidade do modelo proposto e dos códigos desenvolvidos.

Os gráficos trace plots, de convergência da cadeia, bem como as distribuições *a posteriori* e de auto correlação, podem ser encontrados em [Seção D.5](#), [Seção D.6](#), [Seção D.7](#) e [Seção D.8](#) do [Apêndice D](#).

## 4.2 Aplicação em dados reais

Nesta seção, iremos aplicar a metodologia proposta em três diferentes tabelas de contingência, construídas com dados reais da área da educação. A ideia é avaliar a associação entre o desempenho no Enem 2023 (Exame Nacional do Ensino Médio) de estudantes da rede estadual de ensino e a escolaridade dos pais ou responsáveis. Elegemos os desempenhos nas provas de linguagens suas tecnologias (ou simplesmente linguagens), matemática e suas tecnologias (vamos utilizar apenas o termo matemática) e na prova de redação. Os detalhes da seleção dos dados e construção das tabelas serão apresentados a seguir.

### 4.2.1 Exame Nacional do Ensino Médio - Enem - contextualização

Criado em 1998, o Exame Nacional do Ensino Médio (Enem) tem como objetivo avaliar o desempenho dos concluintes do ensino médio no Brasil. A partir do ano de 2009, passou a ser um dos principais meios de ingresso nas Universidades Federais. O Exame mede as competências e habilidades dos estudantes por meio de uma redação e provas objetivas que abrangem quatro áreas do conhecimento: linguagens, ciências humanas, ciências da natureza e matemática.

As provas do Enem são aplicadas pelo [Inep \(2025\)](#) e os resultados do Enem são de grande importância tanto para os candidatos quanto para as universidades, para a sociedade, união estados e municípios. São várias as finalidades, dentre as principais podemos citar:

- Autoavaliação e inserção no mercado de trabalho;
- Referência nacional para aperfeiçoamento dos currículos do ensino médio;
- Acesso à educação superior e aos programas governamentais de financiamento;
- Instrumento de seleção para o mundo do trabalho desenvolvimento de estudos e indicadores sobre a educação brasileira.

Desde sua criação em 1998, o Enem passou por mudanças, mas manteve seu objetivo de avaliar os conhecimentos dos estudantes. Em 2009, o Enem passou por uma reformulação metodológica, alterando as suas Matrizes de Referência.

Essa reformulação visou melhorar a avaliação dos conhecimentos dos estudantes e sua preparação para a educação superior. A aplicação do Enem inclui a produção de diferentes cadernos para cada prova, com questões organizadas em ordem diferente, para prevenir fraudes. Os participantes respondem a um questionário socioeconômico, que em 2023 contava com 25 questões, dentre elas, o nível de escolaridade dos pais. Essas mudanças visam melhorar a segurança e a acessibilidade do Exame.

O [Quadro 1](#) apresenta a descrição dos componentes curriculares para cada área do conhecimento.

Quadro 1 – Descrição das Áreas de Conhecimento e Componentes Curriculares do Enem

Área do conhecimento	Componentes curriculares
Linguagens, códigos e suas tecnologias (linguagens)	Língua Portuguesa, Literatura, Língua Estrangeira (Inglês ou Espanhol), Artes, Educação Física e Tecnologias da Informação e Comunicação
Ciências humanas e suas tecnologias	História, Geografia, Filosofia e Sociologia
Ciências da natureza e suas tecnologias	Química, Física e Biologia
Matemática e suas tecnologias (matemática)	Matemática

Fonte: Adaptada de [Inep \(2023\)](#).

Para esse trabalho, escolhemos analisar o desempenho dos participantes do Enem 2023 das áreas de conhecimento de linguagens, matemática e redação, por serem as disciplinas mais cobradas em diversos exames nacionais e internacionais.

## 4.2.2 Microdados do Enem 2023

O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira [Inep \(2025\)](#) disponibiliza informações detalhadas sobre o Enem, incluindo as provas, os gabaritos, as informações sobre os itens e as notas, e o questionário socioeconômico respondido pelos inscritos no exame. Os dados são disponibilizados em formato “.csv”, os inputs para a leitura desses arquivos foram elaborados utilizando os softwares SAS, SPSS e R. Os inputs trazem a possibilidade de carregar os rótulos juntamente com os dados. Todas essas informações estão disponíveis no arquivo “Microdados-Enem-2023” e em uma planilha específica para os itens. Além disso, um dicionário fornece explicações sobre as variáveis contidas em cada base de dados.

### 4.2.2.1 Uso da TRI para calcular as notas do Enem

O Enem utiliza a Teoria de Resposta ao Item (TRI) para calcular as notas dos participantes e analisar os itens da prova. A TRI é um conjunto de modelos matemáticos que relaciona a probabilidade de um participante responder corretamente a uma questão com seu conhecimento na área avaliada e as características dos itens. Leitores familiarizados com o tema, podem se abster da leitura dessa subseção.

Um aspecto importante da TRI é que ela considera a particularidade de cada item, o que significa que as notas não dependem apenas do total de itens da prova, mas sim de cada item

individualmente. Isso permite que dois participantes com a mesma quantidade de acertos sejam avaliados de forma diferente, dependendo de quais itens estão certos e errados.

No Enem, a TRI é utilizada para calcular a proficiência dos participantes, considerando não apenas o número de acertos, mas também os parâmetros dos itens e o padrão de respostas. A TRI permite uma avaliação mais precisa e detalhada das habilidades dos participantes.

Em cada área do conhecimento são avaliadas 30 habilidades distribuídos em 9 competências em linguagens e 7 competências em matemática. A nota final é um valor entre 0 e 1.000 pontos. Para esse trabalho, foram criadas faixas de notas de forma a termos 5 categorias por área de conhecimento.

Maiores detalhes sobre TRI e suas aplicações podem ser consultados em [Lawley \(1943\)](#), [Lord \(1952\)](#), [Andrade, Tavares e Valle \(2000\)](#), [Klein \(2003\)](#), [Baker e Kim \(2004\)](#), [Jackson, Gray e Fienberg \(2008\)](#).

#### 4.2.2.2 Nota da prova discursiva (redação) no Enem

Para a nota da prova discursiva no Enem, é disponibilizado um Manual de Redação do Enem 2023 que está disponível para auxiliar os usuários, fornecendo informações sobre a Matriz de Referência e amostras comentadas de redações que receberam pontuação máxima no Enem no ano anterior (2022).

A prova discursiva de cada participante é avaliada por, pelo menos, dois professores graduados em letras ou linguística, de forma independente, sem que um conheça a nota atribuída pelo outro. O desempenho de cada participante será avaliado individualmente por cada um dos professores, de acordo com os critérios do [Quadro 2](#).

Quadro 2 – Critérios para avaliação do desempenho na prova de redação no Enem 2023.

Competência	Descrição
Competência I	Demonstrar domínio da modalidade escrita formal da língua portuguesa;
Competência II	Compreender a proposta de redação e aplicar conceitos das várias áreas de conhecimento para desenvolver o tema, dentro dos limites estruturais do texto dissertativo-argumentativo em prosa;
Competência III	Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista;
Competência IV	Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação;
Competência V	Elaborar proposta de intervenção para o problema abordado, respeitando os direitos humanos.

Fonte: Adaptada de [Inep \(2023\)](#).

Para cada uma das 5 competências, o avaliador irá atribuir uma nota entre 0 e 200 pontos e a soma desses pontos das 5 competências (de 0 e 1.000) formará a nota final de cada avaliador. A nota final da redação será a média aritmética das notas dos dois avaliadores.

#### 4.2.2.3 Preparação da base de dados e das tabelas de contingências

Os microdados do Enem 2023 possuem um total de 3.933.955 de registros com 77 variáveis e foi preciso aplicar alguns filtros para chegar a um conjunto de dados de nossa preferência.

Os primeiros filtros a serem aplicados foram: seleção apenas dos participantes que estão cursando o ensino médio e irão concluir até o final do ano corrente (2023), ser estudante de escola pública da rede estadual ou escola privada, ser estudante do ensino regular, não ser participante do Exame como treineiro, a escola atual estar localizado na zona urbana e o participante ser brasileiro. Foram mantidos no banco de dados apenas as variáveis: “Nota da prova de linguagens e Códigos”, “Nota da prova de matemática”, “Nota da prova de redação”, “Q001:Até que série seu pai, ou o homem responsável por você, estudou?”, “Q002:Até que série sua mãe, ou a mulher responsável por você, estudou?”, “Nome do município da escola”, “Código da Unidade da Federação da escola” e “Número de inscrição (trata-se de uma máscara e não o seu número de inscrição original no Enem)”. Com essas seleções a base de dados ficou reduzida a 654.402 participantes.

O passo seguinte foi realizar o cruzamento da base de dados reduzida com uma base de dados obtida no site do [IBGE \(2021\)](#) do Produto Interno Bruto per capita (PIB-pc) por municípios do Brasil. Nesse caso, a chave única de ligação entre as bases de dados foi o código do município. A necessidade da informação do PIB-pc é justificada pelo fato de precisarmos selecionar participantes das escolas de municípios semelhantes em alguma característica, por exemplo, de mesmo porte socioeconômico. Uma das formas de fazer isso é a seleção daqueles municípios com o valor PIB-pc próximos entre si.

Para a aplicação da metodologia, foram selecionados os municípios com número de habitantes entre 80.000 e 90.000 moradores e com PIB-pc entre R\$00,00 e R\$5.000,00 mensais. Dessa forma, teremos um universo de participantes diversificados. Após esses filtros, o conjunto de dados reduziu para 11.451 participantes.

Os dados foram separados em 3 conjuntos por área de conhecimento: linguagens, matemática e redação. Para evitar a ocorrência e valores ausentes, foram considerados apenas os participantes que estiveram presentes no dia da aplicação de cada prova que estamos considerando.

As tabelas de contingências finais para as três áreas de conhecimento foram construídas categorizando as variáveis em 5 grupos cada: “Nota da prova de linguagens”, “Nota da prova de matemática”, “Nota da prova de redação”, “Q001” e “Q002”. As tabelas de contingências resultantes serão apresentadas nas próximas seções.

O [Quadro 3](#) apresenta a quantidade de participantes que foram selecionados para aplicação com dados reais por área de conhecimentos.

Quadro 3 – Quantidade de candidatos selecionados para aplicação da metodologia com dados reais por área de conhecimento.

Área de conhecimento	Quantidade de candidatos
Linguagens	11.415
Matemática	10.897
Redação	10.859

Fonte: Elaborada pelo autor.

A descrição completa e quantitativo de municípios por unidade da federação selecionados para o estudo com dados reais podem ser consultados no [Quadro 9](#) e no [Quadro 10](#) do [Apêndice F](#).

### 4.2.3 Aplicação 1: Desempenho na prova de linguagens.

A primeira aplicação em dados reais será na avaliação do desempenho dos candidatos selecionados na prova de linguagens segundo a escolaridade da mãe. A escolha da figura feminina foi subjetiva. A [Tabela 10](#) detalha as categorias das linhas e das colunas da tabela de contingência utilizada.

Tabela 10 – Desempenho dos candidatos do Enem 2023 na prova de linguagens segundo a escolaridade da mãe.

Escolaridade da mãe	Desempenho da prova por faixa de valores					Total
	0 a 380	380 a 440	440 a 510	510 a 560	560 a 1.000	
Fundamental incompleto	162	307	752	448	272	1.941
Fundamental completo	89	191	522	405	271	1.478
Médio completo	189	417	1.199	1.320	1.189	4.314
Superior/pós-grad. completo	62	171	647	809	1.287	2.976
Não sabe/não estudou	83	144	251	148	80	706
Total	585	1.230	3.371	3.130	3.099	11.415

Fonte: Elaborada pelo autor.

Os dados da [Tabela 10](#) foram submetidos ao algoritmo do MH para estimar os parâmetros do modelo de associação proposto. No total foram realizadas 5.000 iterações, com saltos de tamanho 20 para redução da autocorrelação. Os descartes foram de (burn-in) das 500 observações iniciais. Os resultados obtidos estão dispostos na [Tabela 11](#).

Tabela 11 – Parâmetros do modelo de associação referentes às notas da prova de linguagens estimados via MH.

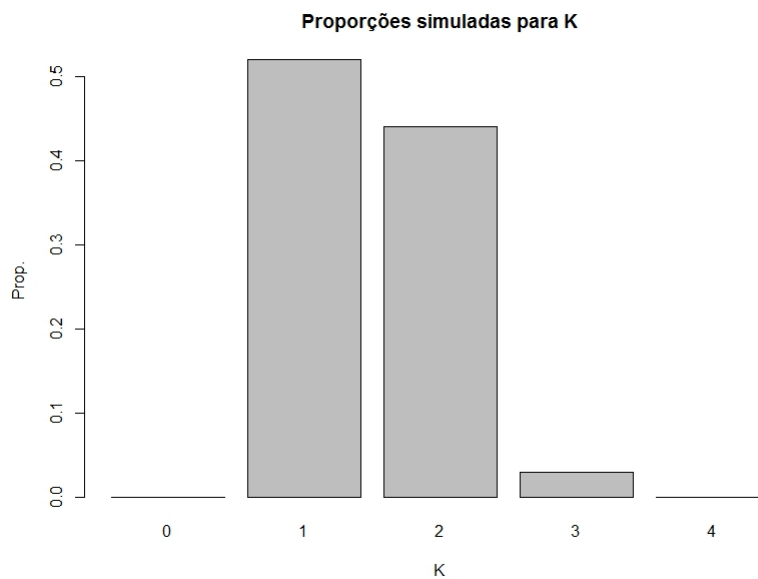
	Valor estimado		Valor estimado		Valor estimado		Valor estimado		Valor estimado
$\lambda$	5,725	$\lambda_1^R$	0,106	$\lambda_1^C$	-1,061	$\mu_{11}$	-0,456	$v_{11}$	-0,723
$\phi_{11}$	1,000	$\lambda_2^R$	-0,191	$\lambda_2^C$	-0,311	$\mu_{21}$	-0,117	$v_{12}$	-0,487
$\phi_{22}$	1,000	$\lambda_3^R$	0,796	$\lambda_3^C$	0,669	$\mu_{31}$	0,352	$v_{13}$	-0,142
$IL(\theta)^{(1)}$	-123,69	$\lambda_4^R$	0,194	$\lambda_4^C$	0,481	$\mu_{41}$	0,989	$v_{14}$	0,348
—	—	$\lambda_5^R$	-0,903	$\lambda_5^C$	0,225	$\mu_{51}$	-0,767	$v_{15}$	1,005
—	—	—	—	—	—	$\mu_{12}$	-0,0617	$v_{21}$	0,069
—	—	—	—	—	—	$\mu_{22}$	-0,113	$v_{22}$	0,009
—	—	—	—	—	—	$\mu_{32}$	0,0498	$v_{23}$	-0,121
—	—	—	—	—	—	$\mu_{42}$	0,038	$v_{24}$	-0,031
—	—	—	—	—	—	$\mu_{52}$	0,081	$v_{25}$	0,067

Fonte: Elaborada pelo autor.

Nota – (1) logaritmo da função de verossimilhança global

Após a obtenção das estimativas, os dados da tabela de contingência foram submetidos ao algoritmo de RJMCMC para obtenção do melhor valor de K. Pela [Figura 11](#), o modelo mais visitado foi o que apresentou valores empatados com  $K = 1$ , com 52% de visitas seguido de  $k = 2$  com 44% de visitas.

Figura 11 – Proporções simuladas para o valor de K - Prova de linguagens

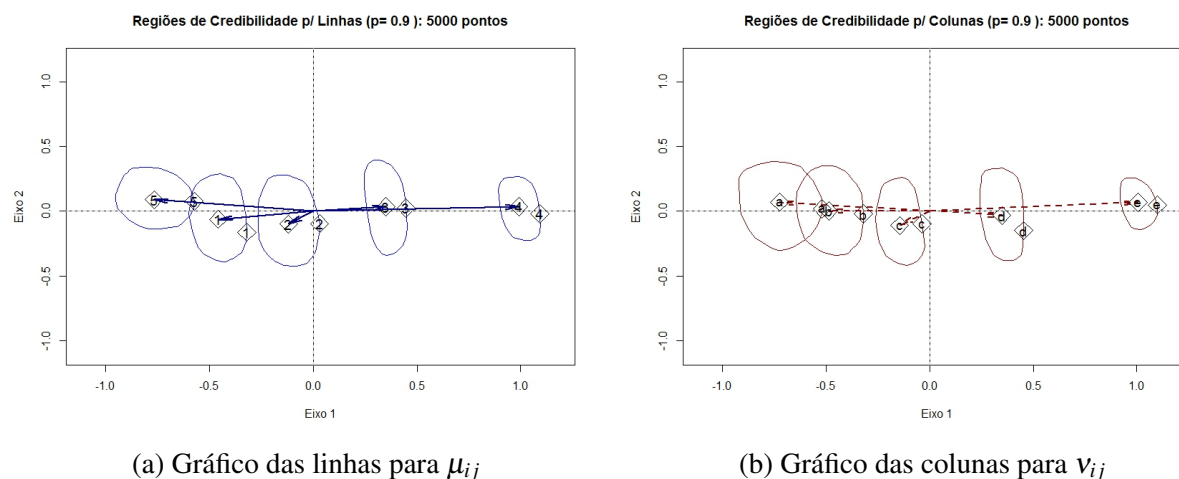


Fonte: Elaborada pelo autor.

#### 4.2.4 Procrustes das estimativas de $\mu_{ij}$ e $\nu_{ij}$ para as notas da prova de linguagens.

Aplicamos a técnica do procrustes para gerar as regiões de 90% de credibilidade tanto para linhas quanto para colunas. O resultado é apresentado na [Figura 12](#),

Figura 12 – Representação gráfica das estimativas de  $\mu_{ij}$  e  $\nu_{ij}$  para o modelo estimado do desempenho na prova de linguagens



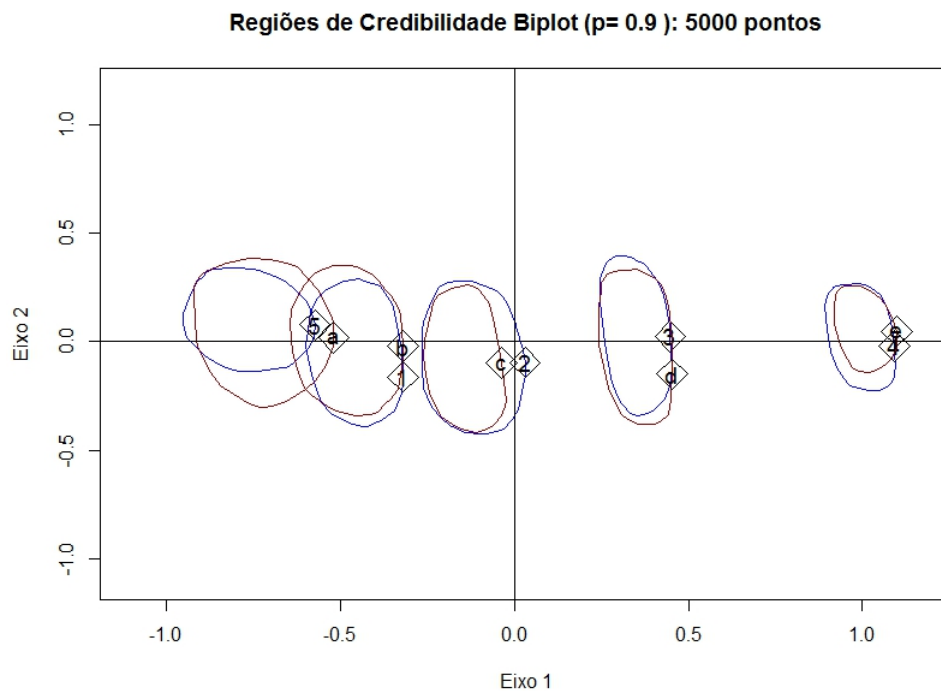
Fonte: Elaborada pelo autor.

Note que após a aplicação do procrustes, a [Figura 12](#) apresenta as regiões obtidas tanto para linhas como para colunas e seus respectivos segmentos de reta. De acordo com o gráfico, aparentemente não existe uma distinção das nuvens de dados do ponto de vista do eixo 2. De acordo com a [Figura 12](#), visualmente, não temos evidências de indicativos de agrupamento tanto nas linhas quanto nas colunas.

#### 4.2.5 Biplot das estimativas de $\mu_{ij}$ e $\nu_{ij}$ para as notas da prova de linguagens.

A representação do biplot para essa aplicação, com o valor de  $K$  igual a 1 não seria possível, porém, para fins didáticos, apresentamos o biplot, para o caso em que o modelo mais visitado fosse  $K=2$ . Na [Figura 13](#), as regiões de linhas e colunas se apresentam bem sobrepostas e com indicativo de associação, note que não há variabilidade com o eixo secundário, isso explica o modelo mais visitado ser com  $K=1$ . Portanto, existe associação entre as categorias das linhas e das colunas da tabela de contingência para o desempenho na prova de linguagens, mas a representação do biplot nesse caso é apenas para fins didáticos.

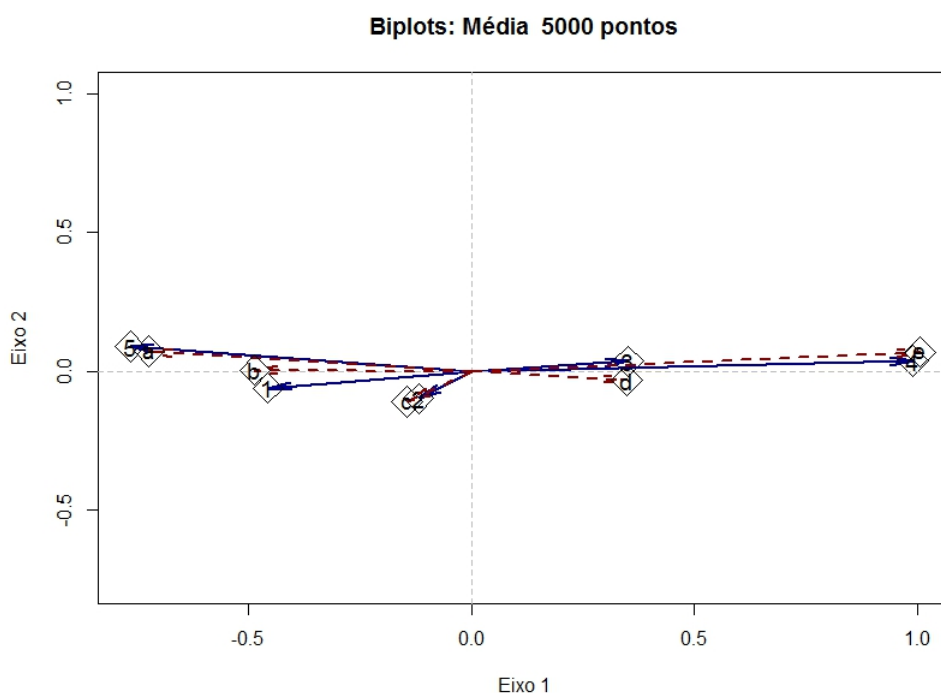
Figura 13 – Representação do biplot para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado do desempenho na prova de linguagens.



Fonte: Elaborada pelo autor.

Fazendo a representação do biplot da [Figura 13](#) no formato de segmentos de retas, temos os seguintes resultados: Existe uma associação significativa entre a linha 4 e a coluna 5, ou seja, maior desempenho com maior nível de escolaridade. Estão associadas a linha 5 com a coluna 1, isto é, escolaridade desconhecida sem escolaridade associada a um menor desempenho na prova de linguagens e suas tecnologias. A associação entre a linha 3 e a coluna 4 indica que o ensino médio completo está associado com o desempenho intermediário na prova e por fim, estão associadas negativamente as linhas 1 e 5, com as colunas 5 e 6, bem como as linhas 3 e 4 com as colunas 1 e 2, o que reforça que escolaridades mais baixas estão associadas negativamente (de forma oposta) com as notas mais altas. Se, por hipótese, considerando  $K = 2$ .

Figura 14 – Representação do biplot na forma de vetores para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado sobre os dados do desempenho na prova de linguagens.



Fonte: Elaborada pelo autor.

Os gráficos trace plots, de convergência da cadeia, bem como as distribuições *a posteriori* e de auto correlação, resultantes das simulações sobre os dados do desempenho na prova de linguagens, podem ser encontrados em [Seção D.9](#), [Seção D.10](#), [Seção D.11](#) e [Seção D.12](#) do [Apêndice D](#).

#### 4.2.6 Aplicação 2: Desempenho na prova de matemática.

A segunda aplicação em dados reais constituiu na avaliação do desempenho dos candidatos selecionados na prova de matemática segundo a escolaridade do pai. A escolha pela escolaridade do pai foi subjetiva. A tabela de contingência utilizada está representada pela [Tabela 12](#).

Tabela 12 – Desempenho dos candidatos do Enem 2023 na prova de matemática segundo a escolaridade do pai.

Escolaridade do pai	Desempenho da prova por faixa de valores					Total
	0 a 380	380 a 470	470 a 590	590 a 700	700 a 1.000	
Fundamental incompleto	354	874	877	447	120	2.672
Fundamental completo	123	354	448	318	89	1.332
Médio completo	272	752	1.197	998	384	3.603
Superior/pós-grad. completo	62	194	413	615	438	1.722
Não sabe/não estudou	1.063	544	477	230	65	1.568
Total	585	2.718	3412	2.608	1.096	10.897

Fonte: Elaborada pelo autor.

Os dados da Tabela 12 foram submetidos ao algoritmo do MH para estimar os parâmetros do modelo de associação proposto. No total foram realizadas 5.000 iterações, com saltos de tamanho 20, com descarte das primeiras 500 observações e um atraso de 100 rodadas para a guarda do valor, isso significa que para guardar um conjunto de resultados são executadas 100 iterações. Os resultados obtidos estão dispostos na Tabela 13.

Os resultados da Tabela 13 se referem à mediana da distribuição *a posteriori*. O maior efeito para a categoria linha foi o que representa a escolaridade superior/pós-graduação ( $\mu_{42} = 1.155$ ). Para as colunas, o maior efeito foi para categoria que representa o maior desempenho dos estudantes ( $\mu_{42} = 1.034$ ).

Tabela 13 – Parâmetros do modelo de associação referentes às notas da prova de matemática estimados via MH.

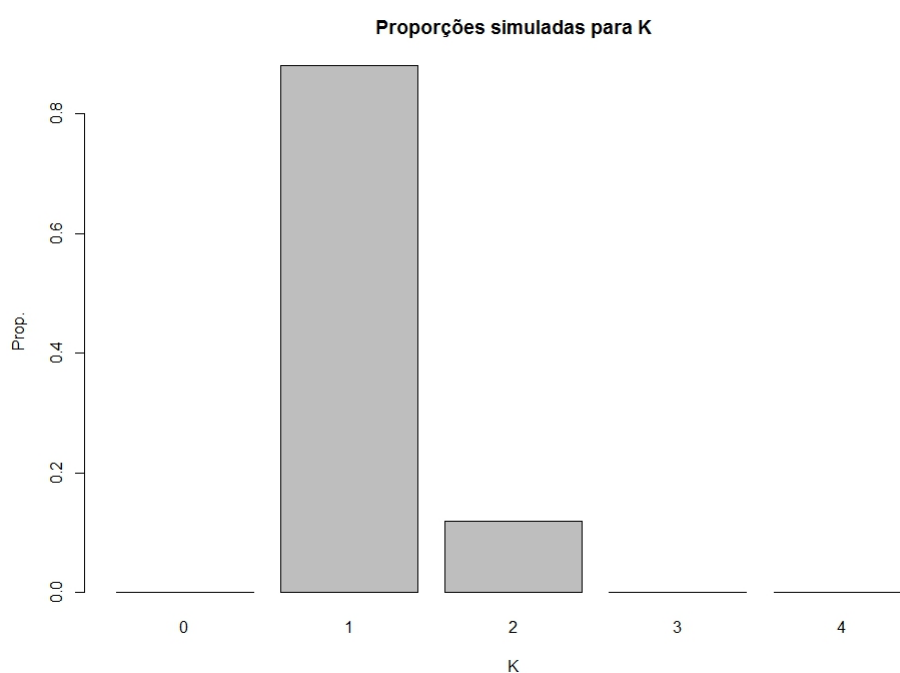
	Valor estimado		Valor estimado		Valor estimado		Valor estimado		Valor estimado
$\lambda$	5,796	$\lambda_1^R$	0,262	$\lambda_1^C$	-0,617	$\mu_{11}$	-0,565	$v_{11}$	-0,782
$\phi_{11}$	1,000	$\lambda_2^R$	-0,370	$\lambda_2^C$	0,364	$\mu_{21}$	-0,108	$v_{12}$	-0,599
$\phi_{22}$	1,000	$\lambda_3^R$	0,644	$\lambda_3^C$	0,641	$\mu_{31}$	0,255	$v_{13}$	-0,122
$lL(\theta)^{(1)}$	-122,860	$\lambda_4^R$	-0,229	$\lambda_4^C$	0,329	$\mu_{41}$	1,155	$v_{14}$	0,469
—	—	$\lambda_5^R$	-0,305	$\lambda_5^C$	-0,716	$\mu_{51}$	-0,738	$v_{15}$	1,034
—	—	—	—	—	—	$\mu_{12}$	0,019	$v_{21}$	0,099
—	—	—	—	—	—	$\mu_{22}$	-0,162	$v_{22}$	0,008
—	—	—	—	—	—	$\mu_{32}$	-0,093	$v_{23}$	-0,130
—	—	—	—	—	—	$\mu_{42}$	0,089	$v_{24}$	-0,110
—	—	—	—	—	—	$\mu_{52}$	0,113	$v_{25}$	0,107

Fonte: Elaborada pelo autor.

Nota – (1) logaritmo da função de verossimilhança global

Submetemos os dados da tabela de contingência [Tabela 12](#) no algoritmo do RJMCMC que desenvolvemos para estimar o valor de  $K$  e o resultado é apresentado na [Figura 15](#). Semelhante à primeira aplicação nos dados do desempenho na prova de linguagens, o modelo mais visitado também foi o que representa o valor de  $K=1$ , com 88% de visitas, seguido do modelo para  $K=2$  com apenas 12% de visitas. Para a construção do biplot, também nesse caso, para fins didáticos, aplicamos o procrustes.

Figura 15 – Proporções simuladas para o valor de  $K$  - Prova de matemática



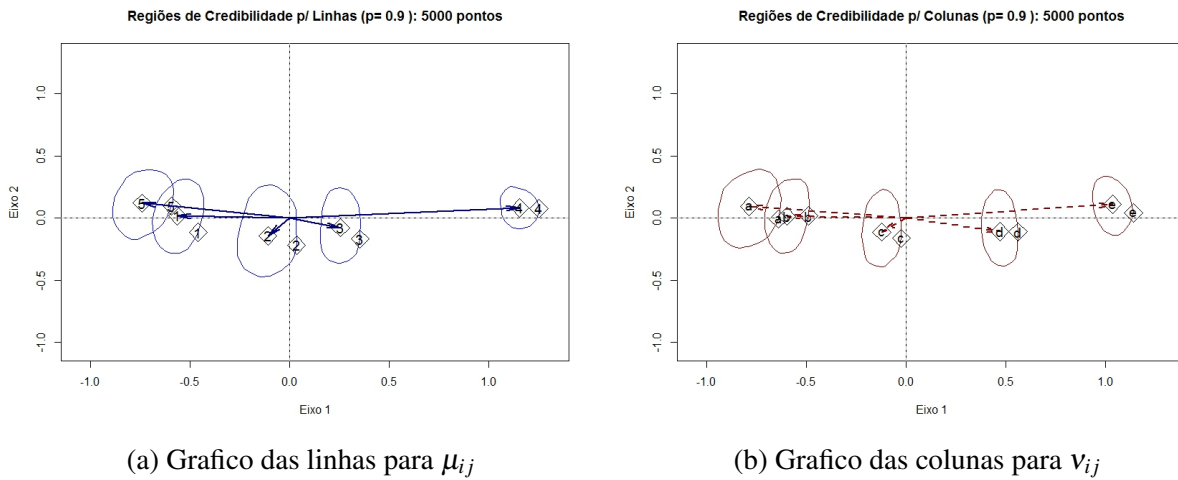
Fonte: Elaborada pelo autor.

#### 4.2.7 Procrustes das estimativas de $\mu_{ij}$ e $v_{ij}$ para as notas da prova de matemática.

Para a construção das regiões de credibilidade é fundamental a aplicação do procrustes. As regiões com 90% de credibilidade para 5.000 simulações para linhas estão na [Figura 16 \(a\)](#). Uma avaliação preliminar indica que pode haver um agrupamento entre as linhas 1 e 2 (ou seja, escolaridade até o ensino fundamental completo). Também notamos que há uma interseção entre as categorias 3, 4 e 5 (ensino médio completo ou mais, incluindo não sabe/não estudou). Um teste de agrupamento seria necessário para confirmar.

A [Figura 16 \(b\)](#) representa as regiões de credibilidade das categorias das colunas da tabela de frequência. Novamente, o gráfico indica um possível agrupamento das colunas “d” e “e” (desempenho acima de 510 pontos) e a intersecção dessas duas colunas com a coluna “c”(faixa de notas de 440 a 510).

Figura 16 – Representação gráfica das estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado do desempenho na prova de matemática.

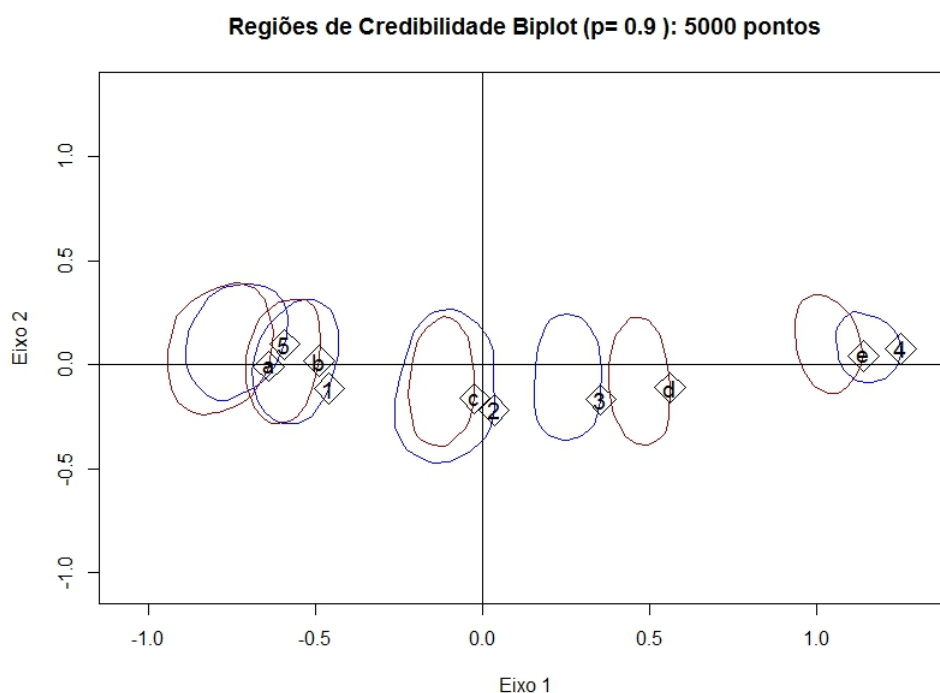


Fonte: Elaborada pelo autor.

#### 4.2.8 *Biplot das estimativas de $\mu_{ij}$ e $v_{ij}$ para as notas da prova de matemática.*

A representação do biplot em termos de regiões de 95% de credibilidade para esse caso é mostrada na [Figura 17](#). Para explorar as medidas de associação, vamos trabalhar com o biplot representado por segmentos de retas.

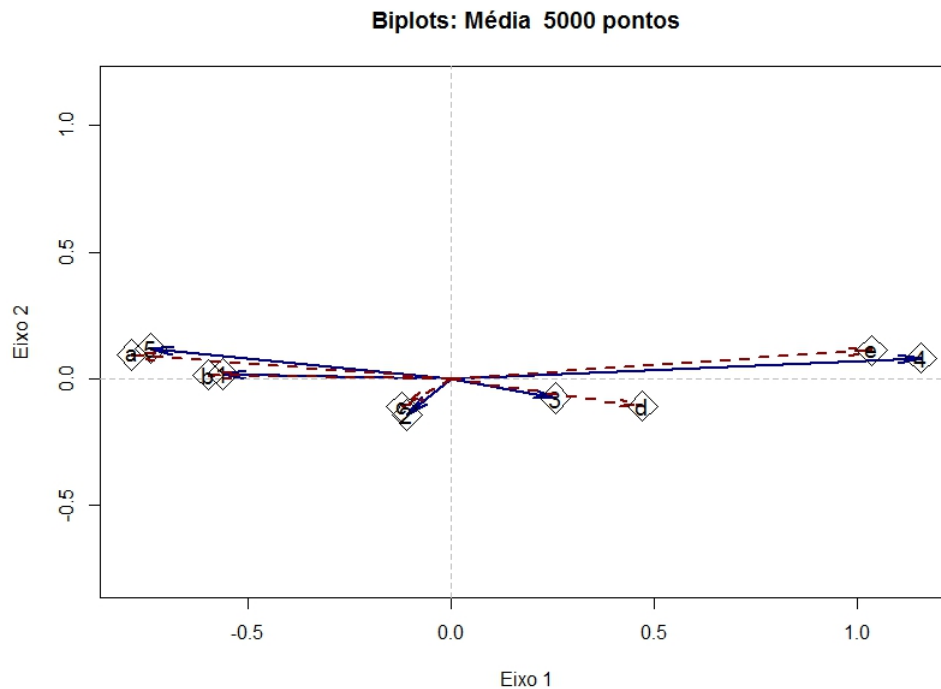
Figura 17 – Representação do biplot para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado no desempenho na prova matemática.



Fonte: Elaborada pelo autor.

A representação do biplot no formato de segmentos de retas, expandindo a escala é mostrada na [Figura 13](#). As maiores variabilidades dos dados estão no ensino fundamental completo (linha 2) e na menor classe de desempenho (coluna “a”). As maiores associações entre categorias dispostas em linhas e colunas estão entre ensino médio completo e não sabe/não estudou (linhas 3 e 5) com a menor classe de notas (coluna “a”), também entre as categorias superior completo (linha 4) com segunda menor faixa de desempenho.

Figura 18 – Representação do biplot na forma de vetores para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado sobre os dados do desempenho na prova de matemática.



Fonte: Elaborada pelo autor.

Os gráficos trace plots, de convergência da cadeia, bem como as distribuições *a posteriori* e de auto correlação, resultantes das simulações sobre os dados do desempenho na prova de língüegens, podem ser encontrados em [Seção D.13](#), [Seção D.14](#), [Seção D.15](#) e [Seção D.16](#) do [Apêndice D](#).

#### 4.2.9 Aplicação 3: Desempenho na prova de redação.

A terceira aplicação com dados reais será na avaliação do desempenho dos candidatos que foram selecionados e realizaram a prova de Redação, segundo a escolaridade da mãe. A escolha da figura materna também foi subjetiva. A [Tabela 14](#) apresenta as quantidades.

Tabela 14 – Desempenho dos candidatos do Enem 2023 na prova de redação segundo a escolaridade da mãe.

Escolaridade da mãe	Desempenho da prova por faixa de valores					Total
	0 até 430	430 até 570	570 até 750	750 até 900	900 até 1.000	
Fundamental incompleto	286	536	586	300	85	1.793
Fundamental completo	184	362	496	282	77	1.401
Médio completo	392	869	1.405	1.067	403	4.136
Superior/pós-grad. completo	131	405	877	1.029	463	2.905
Não sabe/não estudou	137	205	190	71	21	624
Total	1.130	2.377	3.554	2.749	1.049	10.859

Fonte: Elaborada pelo autor.

Os dados da Tabela 14 foram submetidos ao algoritmo do MH para estimar os parâmetros do modelo de associação proposto. No total foram realizadas 5.000 iterações, com saltos de tamanho 20 para redução das autocorrelações e com descarte das 500 primeiras simulações (burn-in). Os resultados obtidos estão dispostos na Tabela 15.

Tabela 15 – Parâmetros dos modelos de associação para Redação estimadas via MH.

	Valor estimado		Valor estimado		Valor estimado		Valor estimado		Valor estimado
$\lambda$	5,688	$\lambda_1^R$	0,003	$\lambda_1^C$	-0,361	$\mu_{11}$	-0,395	$v_{11}$	-0,856
$\phi_{11}$	1,000	$\lambda_2^R$	-0,211	$\lambda_2^C$	0,364	$\mu_{21}$	-0,137	$v_{12}$	-0,507
$\phi_{22}$	1,000	$\lambda_3^R$	0,898	$\lambda_3^C$	0,681	$\mu_{31}$	0,328	$v_{13}$	-0,048
$lL(\theta)^{(1)}$	-118,963	$\lambda_4^R$	0,461	$\lambda_4^C$	0,218	$\mu_{41}$	1,029	$v_{14}$	0,547
—	—	$\lambda_5^R$	-1,151	$\lambda_5^C$	-0,899	$\mu_{51}$	-0,828	$v_{15}$	0,862
—	—	—	—	—	—	$\mu_{12}$	0,004	$v_{21}$	0,046
—	—	—	—	—	—	$\mu_{22}$	-0,149	$v_{22}$	0,013
—	—	—	—	—	—	$\mu_{32}$	0,047	$v_{23}$	-0,077
—	—	—	—	—	—	$\mu_{42}$	0,023	$v_{24}$	-0,081
—	—	—	—	—	—	$\mu_{52}$	0,064	$v_{25}$	0,092

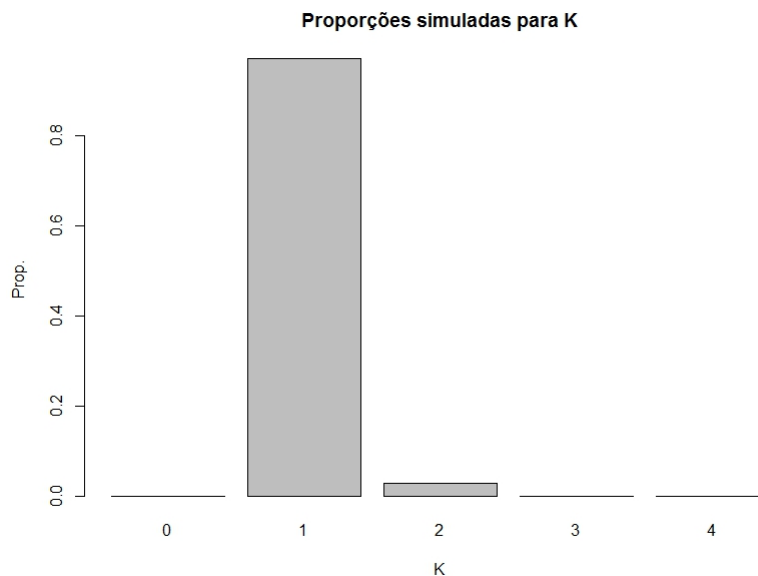
Fonte: Elaborada pelo autor.

Nota – (1) logaritmo da função de verossimilhança global

De acordo com a Tabela 15, o parâmetro que apresentou maior valor da mediana da distribuição *a posteriori* foi  $\mu_{41} = 1.029$ , que corresponde à categoria de escolaridade nível superior/pós-graduação. As medianas do segundo nível de K apresentaram os valores mais baixo para as duas categorias (linhas e colunas).

Para obtenção do melhor valor K, submetemos a tabela de contingência ao algoritmo do RJMCMC, num total de 5.000 simulações, com descarte dos 20 primeiros resultados, saltos de tamanho 20 e uma atraso de 500 rodadas na guarda do valor, ou seja, para cada simulação, executamos 500. O resultado para o melhor valor K está na [Figura 19](#) a seguir.

Figura 19 – Proporções simuladas para o valor de K - Prova de redação



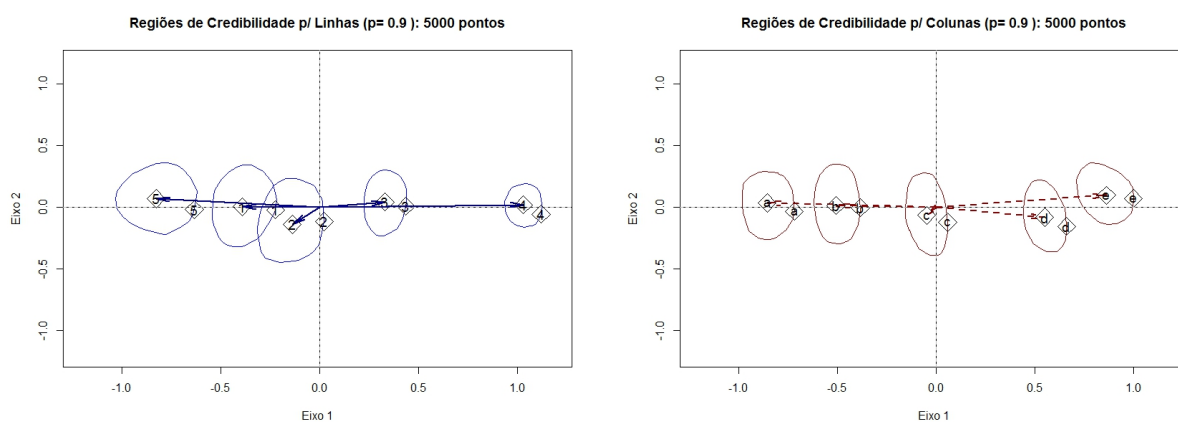
Fonte: Elaborada pelo autor.

Podemos notar que o modelo mais visitado foi o que representa o valor de K=1 com 97% de visitas, portanto, nesse caso, não temos como fazer a representação do biplot para dimensões inferiores a 1. Para fins de exemplificação, se assumíssemos o valor de K=2

#### **4.2.10 Procrustes das estimativas de $\mu_{ij}$ e $v_{ij}$ para as notas da prova de redação.**

Após a aplicação do procrustes, a [Figura 20](#) sugere que não há sinais visuais indicativos de agrupamento de linhas e/ou colunas. Apenas uma pequena interseção entre as linhas 1 e 2, mas não parece ser significativa (necessitaria de um teste específico para este caso para confirmar). Notamos também que o eixo 2 não discrimina as regiões, tanto para linhas quanto para as colunas, permanecendo o espalhamento dos pontos ao longo do eixo 1.

Figura 20 – Representação gráfica das estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado do desempenho na prova de Redação



(a) Grafico das linhas para  $\mu_{ij}$

(b) Grafico das colunas para  $v_{ij}$

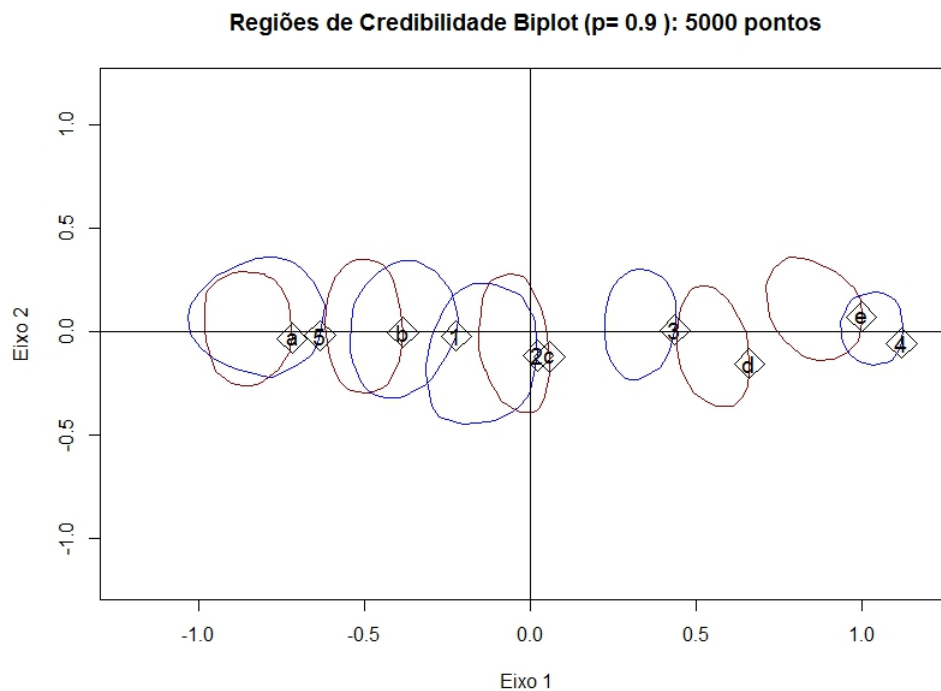
Fonte: Elaborada pelo autor.

#### 4.2.11 *Biplot das estimativas de $\mu_{ij}$ e $v_{ij}$ para as notas da prova de redação.*

A representação do biplot por regiões de credibilidade de 90% para as estimativas da prova de redação é mostrada na [Figura 21](#).

As regiões de credibilidade que representam as linhas 1, 2 e 5 da tabela de contingência aparecem quase que completamente sobrepostas às regiões que representam as colunas “a”, “b” e “c”. Visualmente, não identificamos a possibilidade de agrupamento de linhas e colunas.

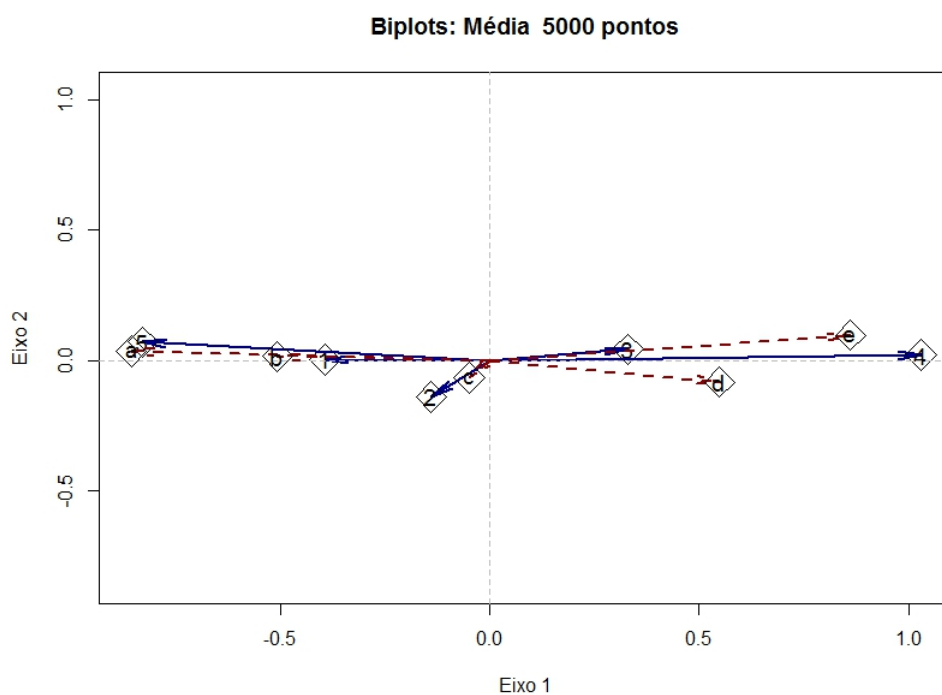
Figura 21 – Representação do biplot para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado no desempenho na prova de Redação.



Fonte: Elaborada pelo autor.

Fazendo a representação do biplot da [Figura 21](#) no formato de segmentos de retas, temos como resultado a [Figura 22](#). Podemos extrair as seguintes análises: Variáveis com maiores dispersões, escolaridade da mãe nível superior/Pós-graduação e não sabe/não respondeu (linhas 4 e 5 do gráfico) e a coluna com o valor do desempenho (coluna “a”. Em termos de associação, temos que as escolaridades da mãe médio completo e superior completo (linhas 3 e 4) estão fortemente associadas às maiores notas de desempenho (colunas “d” e “e”). As menores notas de desempenho (colunas “a” e “b”) estão associadas aos menores níveis de escolaridades (linhas 1 e 5). Por fim, as menores notas de desempenho possuem associação oposta aos maiores níveis de escolaridade, o que faz todo sentido.

Figura 22 – Representação do biplot na forma de vetores para as estimativas de  $\mu_{ij}$  e  $v_{ij}$  para o modelo estimado sobre os dados do desempenho na prova de redação.



Fonte: Elaborada pelo autor.

Os gráficos trace plots, de convergência da cadeia, bem como as distribuições *a posteriori* e de auto correlação, resultantes das simulações sobre os dados do desempenho na prova de linguagens, podem ser encontrados em [Seção D.17](#), [Seção D.18](#), [Seção D.19](#) e [Seção D.20](#) do [Apêndice D](#).



---

## CONCLUSÃO

---

### 5.1 Conclusões do trabalho

Propomos neste trabalho um método de estimação utilizando a abordagem Bayesiana da ordem para um modelo de associação RC(K) utilizando o reversible jump MCMC. Introduzimos a utilização do biplot para auxiliar na identificação de associação e agrupamento de linhas e/ou colunas, juntamente com o procrustes e os resultados foram satisfatórios. Utilizamos o procrustes para reduzir as oscilações devido à indenticabilidade de parâmetros e mostramos que o ganho é significativo quando aplicamos essa técnica. Não é possível eliminar toda a oscilação dos eixos, mas os resultados foram interessantes

A aplicação da metodologia proposta nos dados simulados se mostrou eficiente, os valores dos parâmetros foram próximos dos valores reais. Houve indicativos de agrupamento de colunas para o conjunto de dados simulados detectados através do biplot.

A aplicação da metodologia nos dados já utilizados em outros artigos que tratam de modelo de associação se mostrou consistente, inclusive no que diz respeito ao agrupamento de linhas e colunas, obtendo a mesma conclusão em consonância com outros autores, como [Wermuth e Cox \(1998\)](#).

Na aplicação em dados reais, não houve indicativos de agrupamento de linhas ou de colunas, mas o método se mostrou apropriado para medir a associação entre o desempenho dos participantes no Enem 2023 nas provas de linguagens, matemática e redação, segundo a escolaridade do pai ou da mãe. Os gráficos biplot revelaram algumas associações interessantes, enquanto que o procruste, nesse caso das tabelas de contingência teve uma atuação mais discreta, mas os gráficos de linhas e colunas revelaram a existência de um espelhamento em torno do eixo horizontal e a associação dos desempenhos melhores com a escolaridade dos pais maiores e desempenho piores com escolaridades baixa dos pais.

Podemos concluir que para os dados aqui testados, a metodologia se mostrou robusta e

adequada para estimar os parâmetros do modelo de associação RC(K), encontrar o melhor valor de K e fazer o agrupamento e desagrupamento de linhas e/ou colunas utilizando o RJMCMC.

Como proposta de trabalhos futuros, propomos uma metodologia para fazer o agrupamento e desagrupamento de linhas /ou colunas de uma tabela de contingência também utilizando o RJMCMC.

## REFERÊNCIAS

---

---

AGRESTI, A. Analysis of ordinal categorical data. Wiley, p. 5–23, 1984. Citado na página 32.

\_\_\_\_\_. Categorical data analysis. Wiley, 2012. Citado na página 39.

\_\_\_\_\_. Categorical data analysis. John Wiley & Sons, Inc., Hoboken, New Jersey, Gainesville, Florida, p. 37–39, 44–45, 54–55, 2013. ISSN 978-0-470-46363-5. Disponível em: <[https://mybiostats.files.wordpress.com/2015/03/3rd-ed-alan\\_agresti\\_categorical\\_data\\_analysis.pdf](https://mybiostats.files.wordpress.com/2015/03/3rd-ed-alan_agresti_categorical_data_analysis.pdf)>. Citado nas páginas 40 e 41.

AGRESTI, A.; CHUANG, C.; KEZOUH, A. Order-restricted score parameters in association models for contingency tables. **Journal of the American Statistical Association**, v. 82:398, p. 619–623, 1987. Citado na página 57.

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. **Teoria de resposta ao item: conceitos e aplicações**. 1st. ed. São Paulo-SP: Associação Brasileira de Estatística, 2000. 164 p. Citado na página 85.

BAKER, F. B.; KIM, S. H. **Item response theory: parameter estimation techniques**. 2nd. ed. New York: CRC Press, 2004. 528 p. Citado na página 85.

BARTOLUCCI, F.; FORCINA, A. **ExtRC: Extended RC Models for Contingency Tables**. [S.l.], 2020. R package version 1.0. Disponível em: <<https://CRAN.R-project.org/package=ExtRC>>. Citado na página 35.

BECKER, M. P. Models for the analysis of association in multivariate contingency tables. **Journal of the American Statistical Association**, v. 84, n. 408, p. 1014–1019, dec. 1989. Disponível em: <<http://www.jstor.org/stable/2290077>>. Citado na página 57.

BECKER, M. P.; CLOGG, C. C. Analysis of sets of two-way contingency tables using association models. **Journal of the American Statistical Association**, Taylor & Francis, v. 84, n. 405, p. 142–151, 1989. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01621459.1989.10478749>>. Citado nas páginas 55 e 57.

BOWKER, A. H. A test for symmetry in contingency tables. **Journal of the American Statistical Association**, v. 43, p. 572–574, 1948. Citado na página 35.

BRAAK, C. J. F. T. Interpreting canonical correlation analysis through biplots of structure correlations and weights. **Psychometrika**, v. 55, p. 519–531, 02 1990. Citado na página 43.

BRADU, D.; GABRIEL, K. R. The biplot as a diagnostic tool for models of two-way tables. **Technometrics**, Taylor & Francis, v. 20, n. 1, p. 47–68, 1978. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/00401706.1978.10489617>>. Citado na página 71.

BROOKS, S.; GELMAN, A.; JONES, G.; MENG, X.-L. **Handbook of Markov Chain Monte Carlo**. New York: Chapman & Hall/CRC, 2011. ISBN 9780429138508. Citado na página 49.

- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, Taylor & Francis, Ltd, v. 46, p. 167–174, 03 1992. Disponível em: <[https://biostat.jhsph.edu/~mmccall/articles/casella\\_1992.pdf](https://biostat.jhsph.edu/~mmccall/articles/casella_1992.pdf)>. Citado na página 47.
- CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. **The American Statistician**, JSTOR, v. 49, n. 4, p. 327–35, 1995. Disponível em: <<https://doi.org/10.2307/2684568>>. Acesso em: 01/11/2023. Citado nas páginas 47 e 49.
- EDWARDS, A. W. F. The measure of association in a 2 x 2 table. **Journal of the Royal Statistical Society. Series A (General)**, Oxford University Press, v. 126, n. 1, p. 109–114, 1963. ISSN 00359238. Disponível em: <<http://www.jstor.org/stable/2982448>>. Citado na página 34.
- ESPENDILLER, M. **Association in contingency tables : an informationtheoretic approach**. Tese (PhD thesis) — RWTH Aachen University, 2017. Disponível em: <<https://publications.rwth-aachen.de/record/688951>>. Citado na página 35.
- FAN, Y.; SISSON, S. A. Reversible jump markov chain monte carlo. **arXiv: Methodology**, 2010. Disponível em: <<https://api.semanticscholar.org/CorpusID:88512168>>. Citado na página 49.
- FORCINA, A.; KATERI, M. A new general class of rc association models: Estimation and main properties. **Journal of Multivariate Analysis**, v. 184, p. 104741, 2021. ISSN 0047-259X. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0047259X21000191>>. Citado na página 35.
- FRÜHWIRTH-SCHNATTER, S. Finite mixture and markov switching models. Springer Science, New York, 2006. Citado na página 46.
- GABRIEL, K. R. The biplot graphical display of matrices with application to principal component analysis. **Biometrika**, v. 58, n. 3, p. 453–467, 12 1971. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/58.3.453>>. Citado nas páginas 42 e 68.
- \_\_\_\_\_. Biplot display of multivariate matrices for inspection of data and diagnosis. at Sheffield, England, march 1980. Citado nas páginas 42 e 43.
- GABRIEL, K. R.; ODOROFF, C. L. Biplots in biomedical research. **Statistics in medicine**, v. 9 5, p. 469–85, 1990. Disponível em: <<https://doi.org/10.1002/sim.4780090502>>. Citado na página 43.
- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**, Institute of Mathematical Statistics, v. 7, n. 4, p. 457–472, 1992. ISSN 08834237. Disponível em: <<http://www.jstor.org/stable/2246093>>. Citado nas páginas 49 e 75.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. **Pattern Analysis and Machine Intelligence**, IEEE Transactions on, v. 6, p. 721–741, 1984. Citado na página 47.
- GOLUB, G. H.; VAN-LOAN, C. F. **Matrix Computations**. 4th ed. ed. Baltimore: The Johns Hopkins University Press, 2013. 76-80 e 327-328 p. Citado nas páginas 42, 44 e 45.
- GOODMAN, L. A. Multiplicative models for the analysis of occupational mobility tables and other kinds of cross-classification tables. **American Journal of Sociology**, v. 84, p. 804 – 819, 1979. Disponível em: <<https://www.jstor.org/stable/2778025>>. Citado nas páginas 33 e 52.

\_\_\_\_\_. Simple models for the analysis of association in cross-classifications having ordered categories. **Journal of the American Statistical Association**, American Statistical Association, v. 74, n. 367, p. 537–552, 1979. Disponível em: <<http://www.jstor.org/stable/2286971>>. Citado nas páginas 52, 53 e 54.

\_\_\_\_\_. Association models and canonical correlation in the analysis of cross-classifications having ordered categories. **Journal of the American Statistical Association**, American Statistical Association, Taylor & Francis, Ltd., v. 76, n. 374, p. 320–334, 1981. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2287833>>. Citado nas páginas 33, 53, 54, 55 e 57.

\_\_\_\_\_. The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries. **Ann. Statist.**, v. 13, p. 10–69, 1985. Citado nas páginas 54, 57 e 59.

\_\_\_\_\_. Some useful extensions of the usual correspondence analysis approach and the usual log-linear models approach in the analysis of contingency tables. **International Statistical Review / Revue Internationale de Statistique**, Wiley, International Statistical Institute (ISI), v. 54, n. 3, p. 243–270, 1986. ISSN 03067734, 17515823. Disponível em: <<http://www.jstor.org/stable/1403053>>. Citado nas páginas 54, 57 e 59.

\_\_\_\_\_. A single general method for the analysis of cross-classified data: Reconciliation and synthesis of some methods of pearson, yule, and fisher, and also some methods of correspondence analysis and association analysis. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 433, p. 408–428, 1996. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/01621459.1996.10476702>>. Citado na página 54.

GOODMAN, L. A.; KRUSKAL, W. H. Measures of association for cross classifications. **Journal of the American Statistical Association**, v. 49, n. 268, p. 732–764, 1954. Citado na página 34.

\_\_\_\_\_. Measures of association for cross classifications. ii: Further discussion and references. **Journal of The American Statistical Association - J AMER STATIST ASSN**, Springer, New York, NY, v. 54, p. 123–163, 03 1959. Citado na página 34.

GOWER, J. C. Statistical methods of comparing different multivariate analyses of the same data. **Journal Mathematics in the archaeological and historical science**, p. 138–149, 1971. Disponível em: <<https://repository.rothamsted.ac.uk/item/8wv79/statistical-methods-of-comparing-different-multivariate-analyses-of-the-same-data>>. Citado nas páginas 43 e 45.

\_\_\_\_\_. Generalized procrustes analysis. **Psychometrika**, v. 40, p. 33–51, 03 1975. Citado nas páginas 43 e 44.

GOWER, J. C.; DIJKSTERHUIS, G. B. **Procrustes Problems**. Oxford University Press, 2004. ISBN 9780198510581. Disponível em: <<https://doi.org/10.1093/acprof:oso/9780198510581.001.0001>>. Citado na página 45.

GRADSHTEYN, I. S.; RYZHIK, I. M. **Jacobian Determinant - in Tables of Integrals, Series, and Products**. 6th ed. ed. San Diego, CA: Academic Press, 2000. 1068-1069 p. Citado na página 111.

GREEN, P. J. Reversible jump markov chain monte carlo computation and bayesian model determination. v. 82, n. 4, p. 711–732, 1995. Citado nas páginas 47, 48, 49 e 64.

- GRICE, J.; HARDY, K. K. Generalized procrustes analysis: A tool for exploring aggregates and persons. **Applied Multivariate Research**, v. 13, p. 1–21, 08 2009. Citado na página 43.
- HABERMAN, S. J. Log-linear models for frequency tables with ordered classifications. **Biometrics**, v. 30, n. 4, p. 589–600, Dec. 1974. Disponível em: <<https://api.semanticscholar.org/CorpusID:124058692>>. Citado na página 35.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. **Biometrika**, v. 57, p. 97, 1970. Citado na página 47.
- IBGE. **Instituto Brasileiro de Geografia e Estatísticas**. 2021. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html?=&t=downloads>>. Acesso em: 27/02/2025. Citado na página 86.
- ILIOPOULOS, G.; KATERI, M.; NTZOUFRAS, I. Bayesian estimation of unrestricted and order-restricted association model for a two-way contingency table. **Computational Statistics and Data Analysis**, v. 51, p. 4643–4655, 2007. Citado na página 55.
- \_\_\_\_\_. Bayesian model comparison for the order restricted rc association model. **Psychometrika**, v. 74, p. 561–587, 2009. Citado nas páginas 55 e 57.
- INEP. **Microdados do Enem 2023**. Brasília, DF, Brasil: INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA, 2023. Disponível em: <<https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/enem>>. Acesso em: 25/02/2025. Citado nas páginas 84, 85, 224 e 225.
- \_\_\_\_\_. **INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA**. Brasília, DF, Brasil: [s.n.], 2025. Disponível em: <<https://www.gov.br/inep>>. Acesso em: 25/02/2025. Citado nas páginas 83 e 84.
- JACKSON, L. F.; GRAY, A. G.; FIENBERG, S. E. Sequential category aggregation and partitioning approaches for multi-way contingency tables based on survey and census data. **The Annals of Applied Statistics**, Institute of Mathematical Statistics, v. 2, n. 3, p. 955 – 981, 2008. Disponível em: <<https://doi.org/10.1214/08-AOAS175>>. Citado nas páginas 79 e 85.
- KATERI, M. Contingency table analysis. **Statistics for Industry and Technology**, Springer, v. 525, p. 169–170, 2014. Citado na página 79.
- \_\_\_\_\_. **Contingency Table Analysis: Methods and Implementation Using R**. [S.l.: s.n.], 2014. 25–26 p. ISBN 978-0-8176-4810-7. Citado nas páginas 35, 40 e 41.
- \_\_\_\_\_.  $\phi$ -divergence in contingency table analysis. **Entropy**, v. 20, n. 5, p. 324, 2018. Disponível em: <<https://doi.org/10.3390/e20050324>>. Citado na página 35.
- KATERI, M.; AHMAD, R.; PAPAIOANNOU, T. New features in the class of association models. **Applied Stochastic Models Data Analysis**, v. 14, n. 22, p. 125–136, 12 1998. Citado na página 55.
- KATERI, M.; ILIOPOULOS, G. On collapsing categories in two-way contingency tables. **Statistics**, Taylor & Francis, v. 37, n. 5, p. 443–455, 2003. Disponível em: <<https://doi.org/10.1080/0233188031000123780>>. Citado nas páginas 79 e 80.

KLEIN, R. Utilização da teoria da resposta ao item no Sistema Nacional de Avaliação da Educação Básica (SAEB). **Ensaio: Avaliação e Políticas Públicas em Educação**, scielo, v. 11, p. 283 – 296, 07 2003. ISSN 0104-4036. Disponível em: <[http://educa.fcc.org.br/scielo.php?script=sci\\_arttext&pid=S0104-40362003000300003&nrm=iso](http://educa.fcc.org.br/scielo.php?script=sci_arttext&pid=S0104-40362003000300003&nrm=iso)>. Citado na página 85.

LAWLEY, D. N. On problems connected with item selection and test construction. **Proceedings of the Royal Society of Edinburgh. Section A. Mathematical and Physical Sciences**, v. 61, p. 273 – 287, 1943. Disponível em: <<https://api.semanticscholar.org/CorpusID:122072511>>. Citado na página 85.

LIPKOVICH, I. A.; SMITH, E. P. Biplot and singular value decomposition macros for excel©. **Journal of Statistical Software**, v. 7, n. 5, p. 1–15, 2002. Disponível em: <<https://www.jstatsoft.org/index.php/jss/article/view/v007i05>>. Citado na página 68.

LORD, F. M. **A theory of test scores**. Psychometric Society, 1952. Disponível em: <<https://api.semanticscholar.org/CorpusID:118536487>>. Citado na página 85.

MCNEMAR, Q. Note on the sampling error of the difference between correlated proportions or percentages. **Psychometrika**, v. 20, n. 12, p. 153–157, 1947. Citado na página 35.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. *the journal of chemical physics*. v. 21(6), p. 1087–1092, 1953. Disponível em: <<https://doi.org/10.1063/1.1699114>>. Citado nas páginas 46, 47 e 49.

MILAN, L.; WHITTAKER, J. Estimation of the parameters of the bilinear association model using the gibbs sampler. In: FAHRMEIR, L.; FRANCIS, B.; GILCHRIST, R.; TUTZ, G. (Ed.). **Advances in GLIM and Statistical Modelling**. New York, NY: Springer New York, 1992. v. 78, p. 131–138. ISBN 978-1-4612-2952-0. Citado na página 55.

\_\_\_\_\_. Application of the parametric bootstrap to models that incorporate a singular value decomposition. **Journal of the Royal Statistical Society Series C: Applied Statistics**, [Wiley, Royal Statistical Society], v. 44, n. 1, p. 31–49, 12 1995. ISSN 00359254, 14679876. Disponível em: <<http://www.jstor.org/stable/2986193>>. Citado nas páginas 66, 67 e 71.

MOSTELLER, F. Association and estimation in contingency tables. **Journal of the American Statistical Association**, Taylor & Francis, Ltda, v. 63, n. 321, p. 1–28, Mar 1968. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2283825>>. Citado na página 35.

PEARSON, K. Mathematical contributions to the theory of evolution xiii: On the theory of contingency and its relation to association and normal correlation. Cambridge University Press, London, v. 1, p. 9–10, 1904. Disponível em: <<https://ia801300.us.archive.org/8/items/cu31924003064833/cu31924003064833.pdf>>. Acesso em: 12/03/2024. Citado na página 31.

R Core Team. **R: A Language and Environment for Statistical Computing**. São Paulo, Brasil: [s.n.], 2022. Disponível em: <<http://www.R-project.org/>>. Citado na página 74.

ROOIJ, M. d.; HEISER, W. Graphical representations and odds ratios in a distance model for the analysis of cross-classified data. **Psychometrika**, v. 70, n. 1, p. 99–122, 03 2005. Citado nas páginas 57 e 59.

SALINAS, D. T.; GARCÍA, N. R.; CONTRERAS, E. J.; HERRERA, F.; CÓZAR, E. D. L. On the use of biplot analysis for multivariate bibliometric and scientific indicators. **Journal of the American Society for Information Science and Technology**, v. 64, p. 1468–1479, 2013. Citado nas páginas 68 e 70.

SIMON, G. Alternative analyses for the singly-ordered contingency table. **Journal of the American Statistical Association**, Taylor & Francis, Ltda, v. 69, n. 348, p. 971–976, 1974. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2286174>>. Citado na página 34.

TSUJITANI, M. Maximum likelihood methods for association models in ordered categorical data. **Behaviormetrika**, Wiley, International Statistical Institute (ISI), v. 14, n. 22, p. 61–67, 07 1987. Disponível em: <[https://doi.org/10.2333/bhmk.14.22\\_61](https://doi.org/10.2333/bhmk.14.22_61)>. Citado na página 55.

\_\_\_\_\_. Graphical analysis of association in cross-classifications having ordered categories. **Behaviormetrika**, v. 15, p. 41–53, 1988. Citado na página 59.

UDINA, F. Interactive biplot construction. **Journal of Statistical Software**, v. 13, p. 519–531, 2005. Citado na página 68.

WERMUTH, N.; COX, D. R. On the application of conditional independence to ordinal data. **International Statistical Review / Revue Internationale de Statistique**, [Wiley, International Statistical Institute (ISI)], v. 66, n. 2, p. 181–199, 08 1998. ISSN 03067734, 17515823. Disponível em: <<http://www.jstor.org/stable/1403489>>. Citado nas páginas 79 e 103.

WILLIAMS, O. D.; GRIZZLE, J. E. Analysis of contingency tables having ordered response categories. **Journal of the American Statistical Association**, Taylor & Francis, v. 67, n. 337, p. 55–63, 1972. ISSN 01621459. Disponível em: <<https://doi.org/10.2307/2284701>>. Citado na página 34.

YULE, G. U. On the association of attributes in statistics: With illustrations from the material of the childhood society, &c. **Philosophical Transactions of the Royal Society of London. Containing Papers of a Mathematical or Physical Character**, The Royal Society, v. 194, p. 257–319, 1900. ISSN 02643952. Disponível em: <<http://www.jstor.org/stable/90759>>. Acesso em: 2024-03-12. Citado na página 31.

\_\_\_\_\_. On the methods of measuring association between two attributes. **Journal of the Royal Statistical Society**, v. 75, n. 6, p. 579–642, May 1912. ISSN 0952-8385. Citado na página 33.

# CÁLCULO DA MATRIZ JACOBIANO PARA O REVERSIBLE JUMP

---

O Jacobiano é uma matriz de derivadas parciais de primeira ordem de uma função vetorial. Ele é usada para descrever como uma função que mapeia um espaço em outro (geralmente de dimensão  $n$  para  $m$ ) se comporta localmente, ou seja, como pequenas variações nas variáveis de entrada afetam as variáveis de saída.

A matriz jacobiana generaliza o conceito de derivada para funções de várias variáveis, ela descreve a melhor aproximação linear da função  $\mathbf{F}$  em torno de um ponto. Se  $\mathbf{F}$  for uma função escalar ( $m = 1$ ), o jacobiano se reduz ao vetor gradiente.

Dentre algumas aplicações da matriz jacobiana, nesse trabalho estamos interessados no determinante dessa matriz, que é usado para mudanças de variáveis em integrais múltiplas.

No caso de saltos de um modelo  $M_{ij}^{(1)}$  para um modelo de dimensão maior,  $i'j'$ , com  $i < i'$  e/ou  $j < j'$ ,  $M_{ij}^{(2)}$ , temos o termo geral do vetor de parâmetros da forma:

$$\underline{\theta}^L = \{\lambda_R^k, \mu_{i+l,k}\} \quad e \quad \underline{\theta}^C = \{\lambda_C^k, \mu_{k,c+j}\} \quad (\text{A.1})$$

com  $l=1,2,\dots,L$ ,  $c=1,2,\dots,C$ ,  $k=1,2,\dots,K$ , sendo  $L$  e  $C$  o número de linhas ou colunas, respectivamente, adicionais da tabela de contingência.

Definição (GRADSHTEYN; RYZHIK, 2000): Se  $f_1, f_2, \dots, f_n$  são funções de  $n$  valores reais que são diferenciáveis, com relação a  $x_1, x_2, \dots, x_n$ , então o determinante do Jacobiano  $J_f(x)$

de  $f_i$  com relação à  $x_j$  é o determinante:

$$J_f(x) = \begin{vmatrix} \frac{\partial f_1}{\partial(x_1)} & \frac{\partial f_1}{\partial(x_2)} & \cdots & \frac{\partial f_1}{\partial(x_n)} \\ \frac{\partial f_2}{\partial(x_1)} & \frac{\partial f_2}{\partial(x_2)} & \cdots & \frac{\partial f_2}{\partial(x_n)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial(x_1)} & \frac{\partial f_n}{\partial(x_2)} & \cdots & \frac{\partial f_n}{\partial(x_n)} \end{vmatrix} \quad (\text{A.2})$$

A notação:  $\frac{\partial(f_1, f_2, \dots, f_n)}{\partial(x_1, x_2, \dots, x_n)}$ , é também utilizada para expressar o Jacobiano  $J_f(x)$   
 $\mu_{ij}$  e  $v_{ij}$

## A.1 Um exemplo - Movimento de $M_1$ para $M_2$ - Matriz $2 \times 2$ para $3 \times 2$ .

$$M_1 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix} \Rightarrow \text{completar com v.a. uniformes} \Rightarrow \begin{bmatrix} \eta_{11} & \eta_{12}, \\ \eta_{21} & \eta_{22} \\ u & u \end{bmatrix} \quad (\text{A.3})$$

com,

$$M_2 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ \eta_{31} & \eta_{32} \end{bmatrix} \quad (\text{A.4})$$

Na forma expansiva dos  $\eta$ 's, a matriz A.3 será:

$$M_1 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ u & u \end{bmatrix} = \begin{bmatrix} \lambda + \lambda_1^R + \lambda_1^C + \mu_{11}\phi_{11}v_{11} & \lambda + \lambda_1^R + \lambda_2^C + \mu_{11}\phi_{11}v_{12} \\ +\mu_{12}\phi_{22}v_{21} + \mu_{13}\phi_{33}v_{31} & +\mu_{12}\phi_{22}v_{22} + \mu_{13}\phi_{33}v_{32} \\ \lambda + \lambda_2^R + \lambda_1^C + \mu_{21}\phi_{11}v_{11} & \lambda + \lambda_2^R + \lambda_2^C + \mu_{21}\phi_{11}v_{12} \\ +\mu_{22}\phi_{22}v_{21} + \mu_{23}\phi_{33}v_{31} & +\mu_{22}\phi_{22}v_{22} + \mu_{23}\phi_{33}v_{32} \\ u & u \end{bmatrix} \quad (\text{A.5})$$

A matriz A.4 na forma expansiva será:

$$M_2 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \\ \eta_{31} & \eta_{32} \end{bmatrix} = \begin{bmatrix} \lambda + \lambda_1^R + \lambda_1^C + \mu_{11}\phi_{11}v_{11} & \lambda + \lambda_1^R + \lambda_2^C + \mu_{11}\phi_{11}v_{12} \\ +\mu_{12}\phi_{22}v_{21} + \mu_{13}\phi_{33}v_{31} & +\mu_{12}\phi_{22}v_{22} + \mu_{13}\phi_{33}v_{32} \\ \lambda + \lambda_2^R + \lambda_1^C + \mu_{21}\phi_{11}v_{11} & \lambda + \lambda_2^R + \lambda_2^C + \mu_{21}\phi_{11}v_{12} \\ +\mu_{22}\phi_{22}v_{21} + \mu_{23}\phi_{33}v_{31} & +\mu_{22}\phi_{22}v_{22} + \mu_{23}\phi_{33}v_{32} \\ \lambda + \lambda_3^R + \lambda_1^C + \mu_{31}\phi_{11}v_{11} & \lambda + \lambda_3^R + \lambda_2^C + \mu_{31}\phi_{11}v_{12} \\ +\mu_{32}\phi_{22}v_{21} + \mu_{33}\phi_{33}v_{31} & +\mu_{32}\phi_{22}v_{22} + \mu_{33}\phi_{33}v_{32} \end{bmatrix} \quad (\text{A.6})$$

Tomando a matriz A.6 como referência, o vetor de parâmetros do modelo  $M_1$  com  $K = 3$  para a matriz A.5 é dado pela equação A.7.

$$\theta_{\sim 1} = \{\lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, u_1, u_2, u_3, u_4\}, \quad (\text{A.7})$$

enquanto que o vetor de parâmetros do modelo  $M_2$  com  $K = 3$  é dado pela A.8,

$$\theta_{\sim 2} = \{\lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, \lambda_3^R, \mu_{31}, \mu_{32}, \mu_{33}\}. \quad (\text{A.8})$$

Isso implica que precisamos completar a matriz A.5 com realizações da distribuição uniforme que nos leve a uma relação bijetora com os parâmetros  $\lambda_3^R, \mu_{31}, \mu_{32}$  e  $\mu_{33}$ . Portanto a função de mudança é dada pela expressão A.9.

$$\theta_{\sim 2} = h_{12}(\theta_{\sim 1}, u) = (\theta_{\sim 1}, u_1, u_2, u_3, u_4). \quad (\text{A.9})$$

Portanto o Jacobiano da mudança de  $M_1 \implies M_2$  é dado pelo determinante da matriz

A.10,

$$J_{M_1} = \begin{vmatrix}
\frac{\partial \lambda}{\partial (\lambda)} & \frac{\partial \lambda}{\partial (\lambda_1^R)} & \frac{\partial \lambda}{\partial (\lambda_1^C)} & \frac{\partial \lambda}{\partial (\mu_{11})} & \cdots & \frac{\partial \lambda}{\partial (u_1)} & \frac{\partial \lambda}{\partial (u_2)} & \frac{\partial \lambda}{\partial (u_3)} & \frac{\partial \lambda}{\partial (u_4)} \\
\frac{\partial \lambda_1^R}{\partial (\lambda)} & \frac{\partial \lambda_1^R}{\partial (\lambda_1^R)} & \frac{\partial \lambda_1^R}{\partial (\lambda_1^C)} & \frac{\partial \lambda_1^R}{\partial (\mu_{11})} & \cdots & \frac{\partial \lambda_1^R}{\partial (u_1)} & \frac{\partial \lambda_1^R}{\partial (u_2)} & \frac{\partial \lambda_1^R}{\partial (u_3)} & \frac{\partial \lambda_1^R}{\partial (u_4)} \\
\frac{\partial \lambda_1^C}{\partial (\lambda)} & \frac{\partial \lambda_1^C}{\partial (\lambda_1^R)} & \frac{\partial \lambda_1^C}{\partial (\lambda_1^C)} & \frac{\partial \lambda_1^C}{\partial (\mu_{11})} & \cdots & \frac{\partial \lambda_1^C}{\partial (u_1)} & \frac{\partial \lambda_1^C}{\partial (u_2)} & \frac{\partial \lambda_1^C}{\partial (u_3)} & \frac{\partial \lambda_1^C}{\partial (u_4)} \\
\frac{\partial \mu_{11}}{\partial (\lambda)} & \frac{\partial \mu_{11}}{\partial (\lambda_1^R)} & \frac{\partial \mu_{11}}{\partial (\lambda_1^C)} & \frac{\partial \mu_{11}}{\partial (\mu_{33})} & \cdots & \frac{\partial \mu_{11}}{\partial (u_1)} & \frac{\partial \mu_{11}}{\partial (u_2)} & \frac{\partial \mu_{11}}{\partial (u_3)} & \frac{\partial \mu_{11}}{\partial (u_4)} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial \lambda_2^R}{\partial (\lambda)} & \frac{\partial \lambda_2^R}{\partial (\lambda_1^R)} & \frac{\partial \lambda_2^R}{\partial (\lambda_1^C)} & \frac{\partial \lambda_2^R}{\partial (\mu_{33})} & \cdots & \frac{\partial \lambda_2^R}{\partial (u_1)} & \frac{\partial \lambda_2^R}{\partial (u_2)} & \frac{\partial \lambda_2^R}{\partial (u_3)} & \frac{\partial \lambda_2^R}{\partial (u_4)} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\frac{\partial u_1}{\partial (\lambda)} & \frac{\partial u_1}{\partial (\lambda_1^R)} & \frac{\partial u_1}{\partial (\lambda_1^C)} & \frac{\partial u_1}{\partial (\mu_{11})} & \cdots & \frac{\partial u_1}{\partial (u_1)} & \frac{\partial u_1}{\partial (u_2)} & \frac{\partial u_1}{\partial (u_3)} & \frac{\partial u_1}{\partial (u_4)} \\
\frac{\partial u_2}{\partial (\lambda)} & \frac{\partial u_2}{\partial (\lambda_1^R)} & \frac{\partial u_2}{\partial (\lambda_1^C)} & \frac{\partial u_2}{\partial (\mu_{11})} & \cdots & \frac{\partial u_2}{\partial (u_1)} & \frac{\partial u_2}{\partial (u_2)} & \frac{\partial u_2}{\partial (u_3)} & \frac{\partial u_2}{\partial (u_4)} \\
\frac{\partial u_3}{\partial (\lambda)} & \frac{\partial u_3}{\partial (\lambda_1^R)} & \frac{\partial u_3}{\partial (\lambda_1^C)} & \frac{\partial u_3}{\partial (\mu_{11})} & \cdots & \frac{\partial u_3}{\partial (u_1)} & \frac{\partial u_3}{\partial (u_2)} & \frac{\partial u_3}{\partial (u_3)} & \frac{\partial u_3}{\partial (u_4)} \\
\frac{\partial u_4}{\partial (\lambda)} & \frac{\partial u_4}{\partial (\lambda_1^R)} & \frac{\partial u_4}{\partial (\lambda_1^C)} & \frac{\partial u_4}{\partial (\mu_{11})} & \cdots & \frac{\partial u_4}{\partial (u_1)} & \frac{\partial u_4}{\partial (u_2)} & \frac{\partial u_4}{\partial (u_3)} & \frac{\partial u_4}{\partial (u_4)}
\end{vmatrix} \quad (\text{A.10})$$

Para encontrar o valor do Jacobiano da matriz A.10, calculamos o valor das derivadas parciais e o resultado é dado pelo cálculo do determinante da matriz A.11, que será igual a 1.

$$\text{Det}(J_{M_1}) = \begin{vmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 \end{vmatrix} = 1 \quad (\text{A.11})$$

## A.2 Outro exemplo - Movimento de $M_1$ para $M_2$ - Matriz $2 \times 2$ para $2 \times 3$ .

$$M_1 = \begin{bmatrix} \eta_{11} & \eta_{12} \\ \eta_{21} & \eta_{22} \end{bmatrix} \Rightarrow \text{completar com v.a. uniformes} \Rightarrow \begin{bmatrix} \eta_{11} & \eta_{12} & u \\ \eta_{21} & \eta_{22} & u \end{bmatrix} \quad (\text{A.12})$$

Em termos de  $\eta$ 's, a matriz  $M_2$  é,

$$M_2 = \begin{bmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{21} & \eta_{22} & \eta_{23} \end{bmatrix} \quad (\text{A.13})$$

A matriz A.12 na forma expansiva dos  $\eta$ 's é dada pela matriz A.14.

$$M_1 = \begin{bmatrix} \eta_{11} & \eta_{12} & u \\ \eta_{21} & \eta_{22} & u \end{bmatrix} = \begin{bmatrix} \lambda + \lambda_1^R + \lambda_1^C + \mu_{11}\phi_{11}v_{11} & \lambda + \lambda_1^R + \lambda_2^C + \mu_{11}\phi_{11}v_{12} \\ +\mu_{12}\phi_{22}v_{21} + \mu_{13}\phi_{33}v_{31} & +\mu_{12}\phi_{22}v_{22} + \mu_{13}\phi_{33}v_{32} \\ \lambda + \lambda_2^R + \lambda_1^C + \mu_{21}\phi_{11}v_{11} & \lambda + \lambda_2^R + \lambda_2^C + \mu_{21}\phi_{11}v_{12} \\ +\mu_{22}\phi_{22}v_{21} + \mu_{23}\phi_{33}v_{31} & +\mu_{22}\phi_{22}v_{22} + \mu_{23}\phi_{33}v_{32} \end{bmatrix} \begin{matrix} u \\ u \end{matrix} \quad (\text{A.14})$$

A matriz A.13 na forma expansiva dos  $\eta$ 's é dada pela matriz A.15,

$$M_2 = \begin{bmatrix} \eta_{11} & \eta_{12} & \eta_{13} \\ \eta_{21} & \eta_{22} & \eta_{23} \end{bmatrix} = \begin{bmatrix} \lambda + \lambda_1^R + \lambda_1^C + \mu_{11}\phi_{11}v_{11} & \lambda + \lambda_1^R + \lambda_2^C + \mu_{11}\phi_{11}v_{12} & \lambda + \lambda_1^R + \lambda_3^C + \mu_{11}\phi_{11}v_{13} \\ +\mu_{12}\phi_{22}v_{21} + \mu_{13}\phi_{33}v_{31} & +\mu_{12}\phi_{22}v_{22} + \mu_{13}\phi_{33}v_{32} & +\mu_{12}\phi_{22}v_{23} + \mu_{13}\phi_{33}v_{33} \\ \lambda + \lambda_2^R + \lambda_1^C + \mu_{21}\phi_{11}v_{11} & \lambda + \lambda_2^R + \lambda_2^C + \mu_{21}\phi_{11}v_{12} & \lambda + \lambda_2^R + \lambda_3^C + \mu_{21}\phi_{11}v_{13} \\ +\mu_{22}\phi_{22}v_{21} + \mu_{23}\phi_{33}v_{31} & +\mu_{22}\phi_{22}v_{22} + \mu_{23}\phi_{33}v_{32} & +\mu_{22}\phi_{22}v_{23} + \mu_{23}\phi_{33}v_{33} \end{bmatrix} \quad (\text{A.15})$$

Os parâmetros do modelo  $M_1$  com  $K = 3$  são,

$$\underline{\theta}_1 = \{\lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, u_1, u_2, u_3, u_4\}. \quad (\text{A.16})$$

Analogamente, os parâmetros do modelo  $M_2$  com  $K = 3$ ,

$$\underline{\theta}_2 = \{\lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, \lambda_3^C, v_{13}, v_{23}, v_{33}\}, \quad (\text{A.17})$$

que leva à função de mudança da equação A.18.

$$\underline{\theta}_2 = h_{12}(\underline{\theta}_1, \underline{u}) = (\underline{\theta}_1, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9). \quad (\text{A.18})$$

Sendo assim, os parâmetros do modelo  $M_2$  com  $K = 3$ , matriz  $2 \times 2$  para  $4 \times 2$ ,

$$\begin{aligned} \underline{\theta}_2 = \{ & \lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, \\ & \lambda_3^R, \mu_{31}, \mu_{32}, \mu_{33}, \lambda_4^R, \mu_{41}, \mu_{42}, v_{43} \} \end{aligned} \quad (\text{A.19})$$

Já os parâmetros do modelo  $M_2$  com  $K = 3$  e matriz  $2 \times 2$  para  $2 \times 4$ .

$$\begin{aligned} \underline{\theta}_2 = \{ & \lambda, \lambda_1^R, \lambda_1^C, \mu_{11}, \phi_{11}, v_{11}, \mu_{12}, \phi_{22}, v_{21}, \mu_{13}, \phi_{33}, v_{31}, \lambda_2^R, \lambda_2^C, \mu_{21}, \mu_{22}, \mu_{23}, v_{12}, v_{22}, v_{32}, \\ & \lambda_3^C, v_{13}, v_{23}, v_{33}, \lambda_4^C, v_{14}, v_{24}, v_{34} \}, \end{aligned} \quad (\text{A.20})$$

que leva ao valor unitário do Jacobiano da transformação.

$$\text{Det}(J_{M_1})^{3 \times 2} = 1 \quad (\text{A.21})$$



## COMBINAÇÃO DE DOIS NÍVEIS DE UMA TABELA DE TRIPLA ENTRADA

Embora grande parte dos estudos que envolvem tabelas de contingência seja mais comumente em tabelas de dupla entrada, ainda assim é possível fazer a combinação de dois níveis de uma tabela de tripla entrada e torná-la bidimensional. Para fins didáticos, apresentamos um exemplo de uma possível combinação

Tabela 16 – Exemplo de transformação de tabela de contingência de tripla entrada para dupla entrada.

(a) Tabela de três entradas (3 níveis)					(b) Tabela reorganizada para dupla entrada.			
Gênero	Fumante	Consumo álcool		Total	Gênero	Consumo álcool		Total
		Sim	Não			Sim	Não	
Masc	Sim	$n_{111}$	$n_{112}$	$n_{11+}$	Masc. fuma	$n_{11}$	$n_{12}$	$n_{1+}$
	Não	$n_{121}$	$n_{122}$	$n_{12+}$	Masc. não fuma	$n_{21}$	$n_{22}$	$n_{2+}$
Fem	Sim	$n_{211}$	$n_{212}$	$n_{21+}$	Fem. fuma	$n_{31}$	$n_{32}$	$n_{3+}$
	Não	$n_{221}$	$n_{222}$	$n_{22+}$	Fem. não fuma	$n_{41}$	$n_{42}$	$n_{4+}$
Total		$n_{++1}$	$n_{++2}$	$n$	Total			

Fonte: Elaborada pelo autor.

Fonte: Elaborada pelo autor.



## TABELAS COMPARATIVAS

Exibimos na tabela a seguir um comparativo dos valores dos parâmetros de uma tabela de contingência simulada e os intervalos de credibilidade de 2.5% e 97.5% obtidos das distribuições *a posteriori* submetidas aos algoritmos de Metroplis-Hastings.

Tabela 17 – Comparativo dos valores dos parâmetros reais com os intervalos de credibilidade dos valores estimados utilizando dados da tabela de continguência simulada utilizando o MH

Parâmetro	Real	IC(2,5% - 97,5%)	Parâmetro	Real	IC(2,5% - 97,5%)	Parâmetro	Real	IC(2,5% - 97,5%)
$\lambda$	4,00	(3,20; 3,76)	$\mu_{11}$	0,861	(1,95; 2,64)	$v_{11}$	0,55	(1,12; 1,53)
$\lambda_1^R$	-0,50	(-0,92; 0,39)	$\mu_{21}$	-0,340	(-1,24; -0,73)	$v_{12}$	-0,29	(-1,40; -0,32)
$\lambda_2^R$	-0,30	(-0,73; 0,05)	$\mu_{31}$	-0,372	(-1,26; -0,84)	$v_{13}$	0,071	(-0,33; 0,29)
$\lambda_3^R$	0,20	(-0,48; 0,31)	$\mu_{41}$	-0,063	(-0,43; 0,03)	$v_{14}$	-0,56	(-2,35; -1,30)
$\lambda_4^R$	0,60	(0,26; 0,91)	$\mu_{12}$	-0,437	(-0,31; 0,20)	$v_{15}$	0,533	(1,04; 1,44)
$\lambda_1^C$	-0,60	(-0,43; 0,47)	$\mu_{22}$	-0,328	(-0,89; 0,06)	$v_{21}$	-0,02	(-0,32; 0,76)
$\lambda_2^C$	-0,40	(-1,73; -0,44)	$\mu_{32}$	-0,613	(-0,25; 0,85)	$v_{22}$	0,757	(-0,83; 0,27)
$\lambda_3^C$	0,20	(-0,61; 0,16)	$\mu_{42}$	-0,569	(-0,55; 0,71)	$v_{23}$	0,520	(-0,57; 0,57)
$\lambda_4^C$	0,30	(-0,13; 1,09)	—	—	—	$v_{24}$	0,019	(-0,08; 0,62)
$\lambda_5^C$	0,50	(0,19; 0,93)	—	—	—	$v_{25}$	0,392	(-0,63; 0,41)

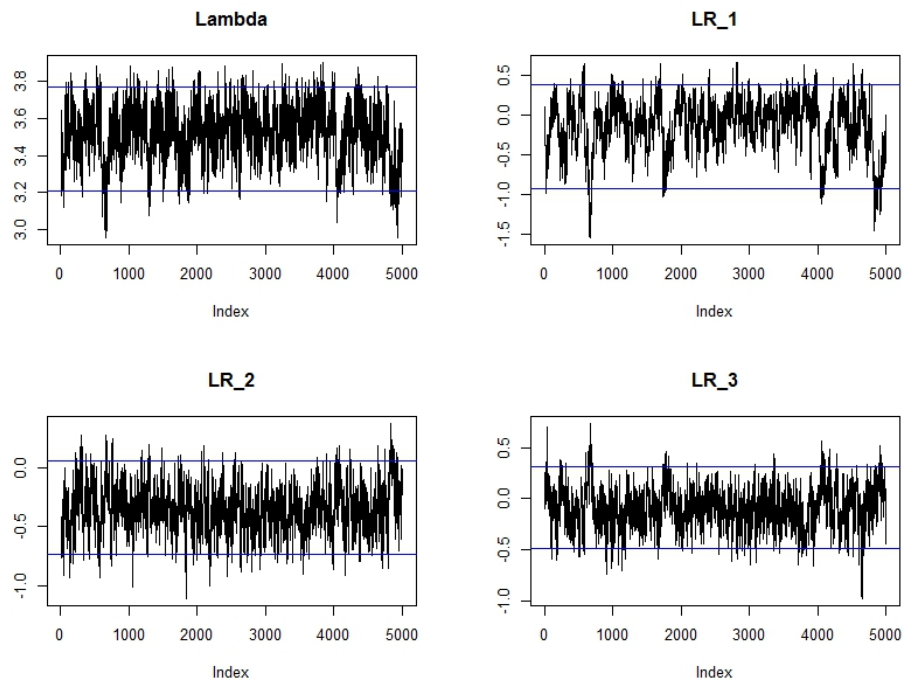
Fonte: Elaborada pelo autor.



## VERIFICAÇÃO DAS CADEIAS

### D.1 Gráficos trace plot dos resultados do MH aplicados aos dados da amostra simulada

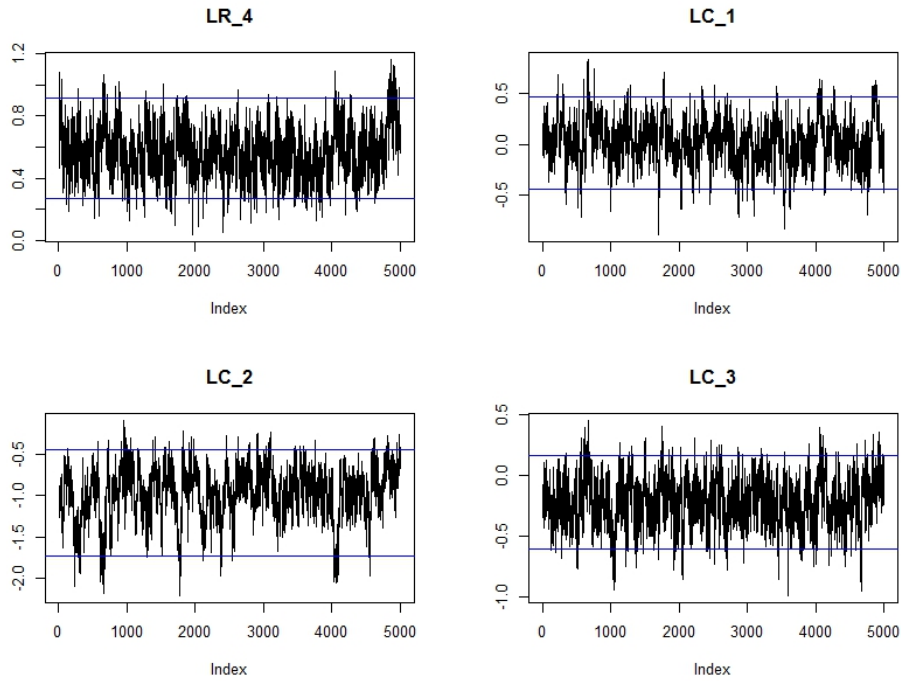
Figura 23 – Trace plot dos resultados do MH aplicados na amostra simulada para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



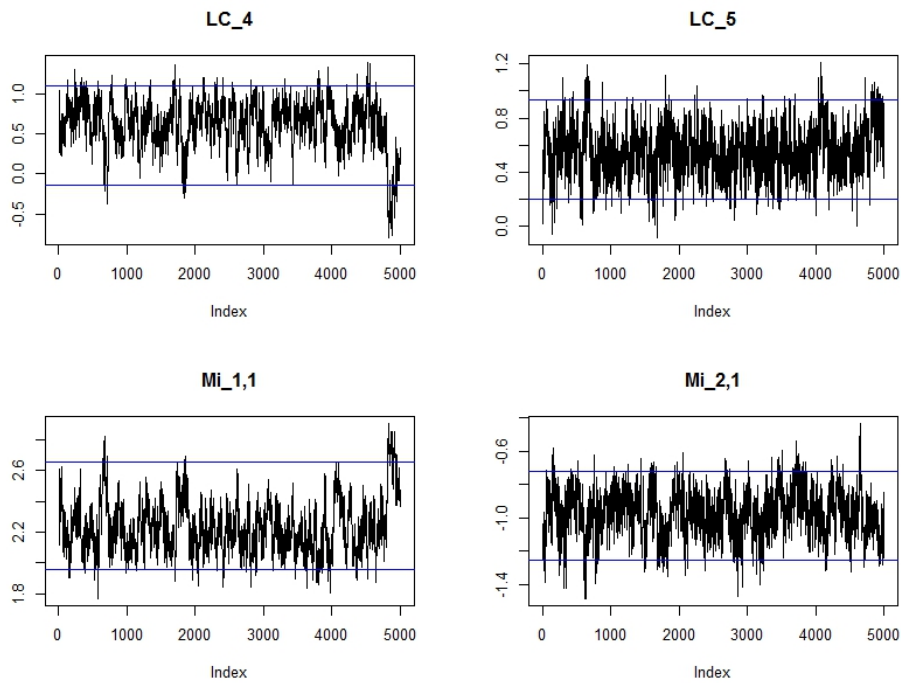
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 24 – Trace plot dos resultados do MH aplicados na amostra simulada para  $\lambda_4^C, \lambda_5^C, \mu_{11}$  e  $\mu_{21}$

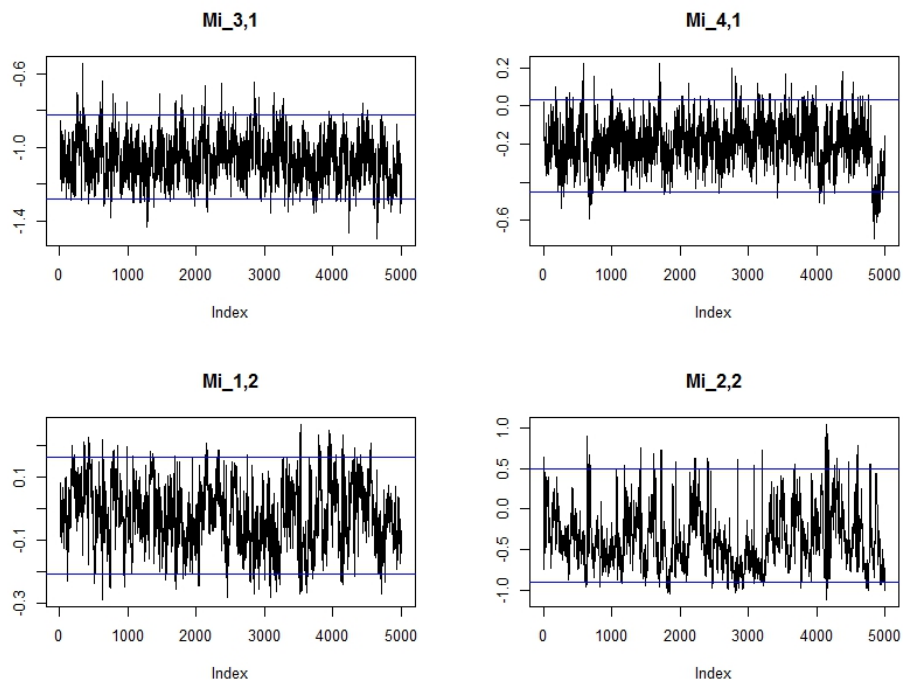
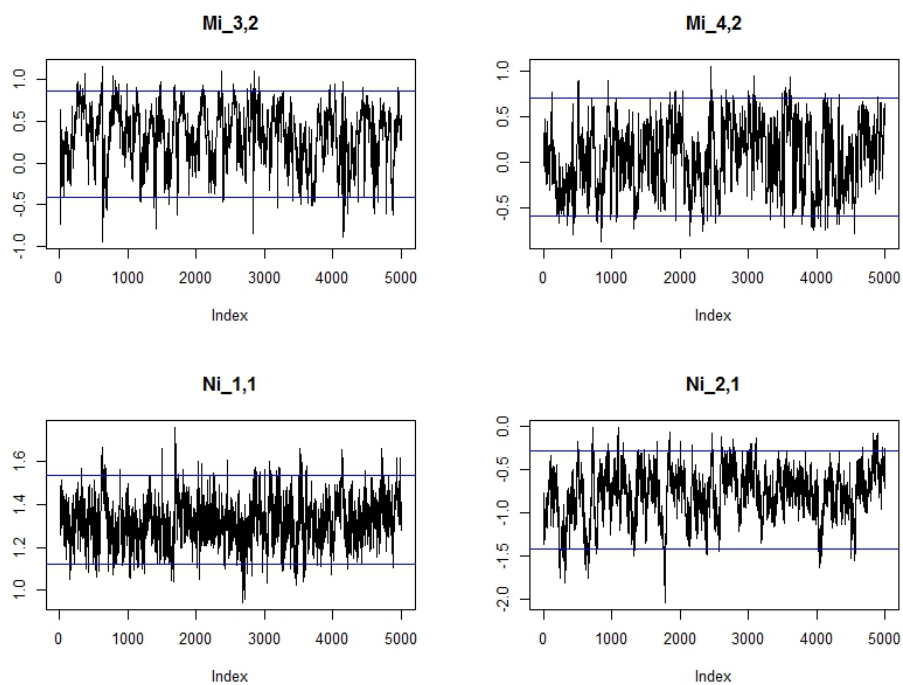


(a)  $\lambda_4^R, \lambda_1^C, \lambda_2^C$  e  $\lambda_3^C$

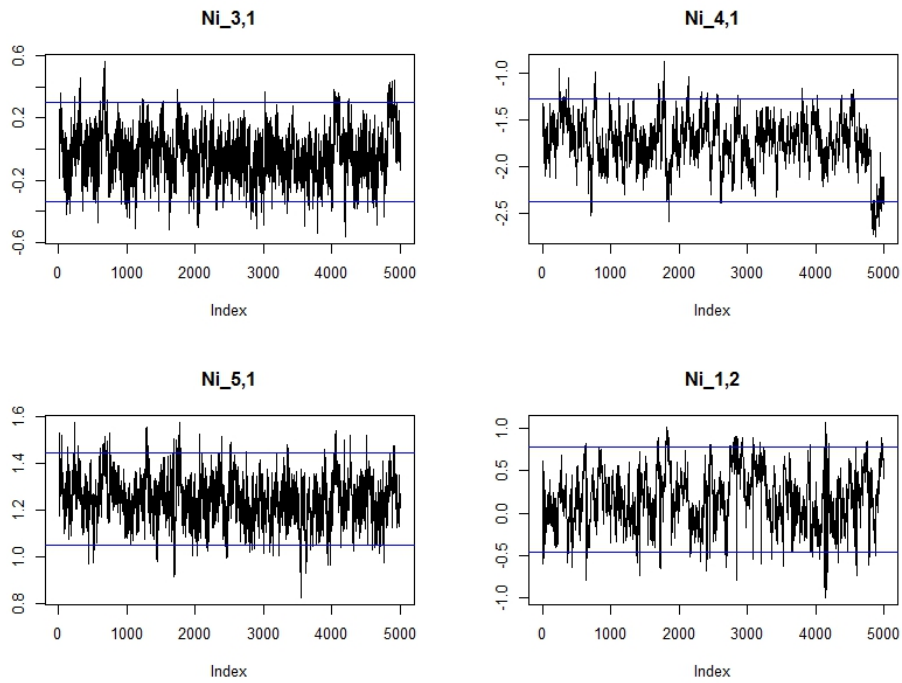
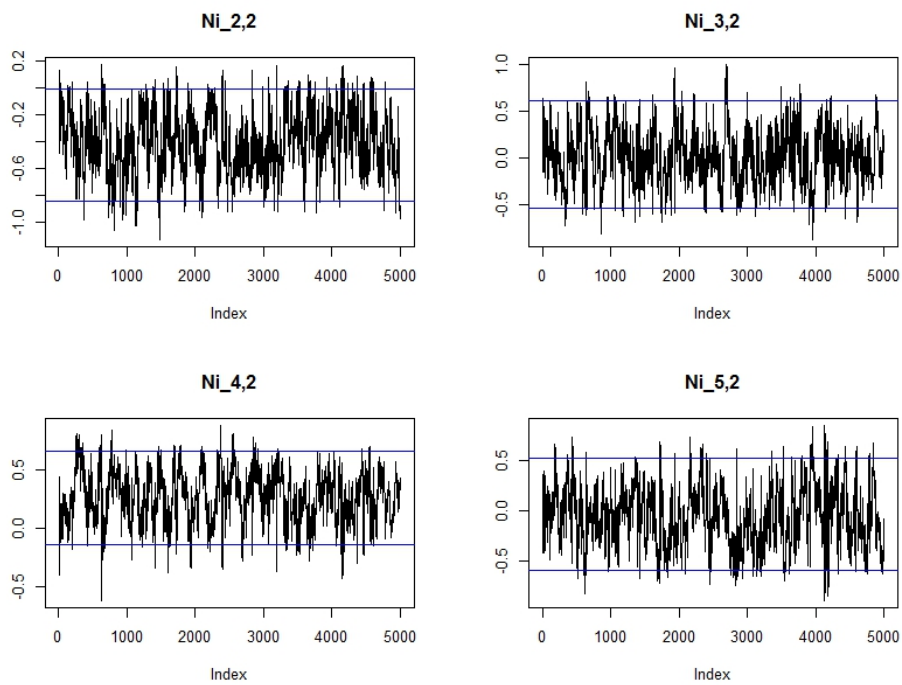


(b)  $\lambda_4^C, \lambda_5^C, \mu_{11}$  e  $\mu_{21}$

Fonte: Elaborada pelo autor.

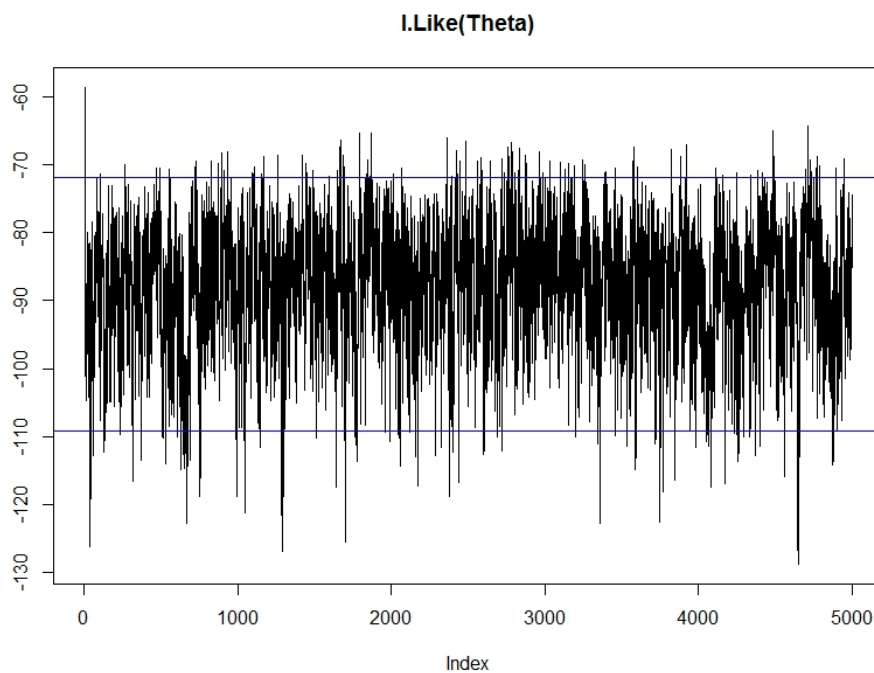
Figura 25 – Trace plot dos resultados do MH aplicados na amostra simulada para  $\mu_{32}, \mu_{42}, v_{11}$  e  $v_{21}$ (a)  $\mu_{31}, \mu_{41}, \mu_{12}$  e  $\mu_{22}$ (b)  $\mu_{32}, \mu_{42}, v_{11}$  e  $v_{21}$ 

Fonte: Elaborada pelo autor.

Figura 26 – Trace plot dos resultados do MH aplicados na amostra simulada para  $v_{22}, v_{32}, v_{42}$  e  $v_{52}$ (a)  $v_{31}, v_{41}, v_{51}$  e  $v_{12}$ (b)  $v_{22}, v_{32}, v_{42}$  e  $v_{52}$ 

Fonte: Elaborada pelo autor.

Figura 27 – Trace plot dos resultados do MH aplicados na amostra simulada para o log-verossimilhança



(a) log-verossimilhança

Fonte: Elaborada pelo autor.

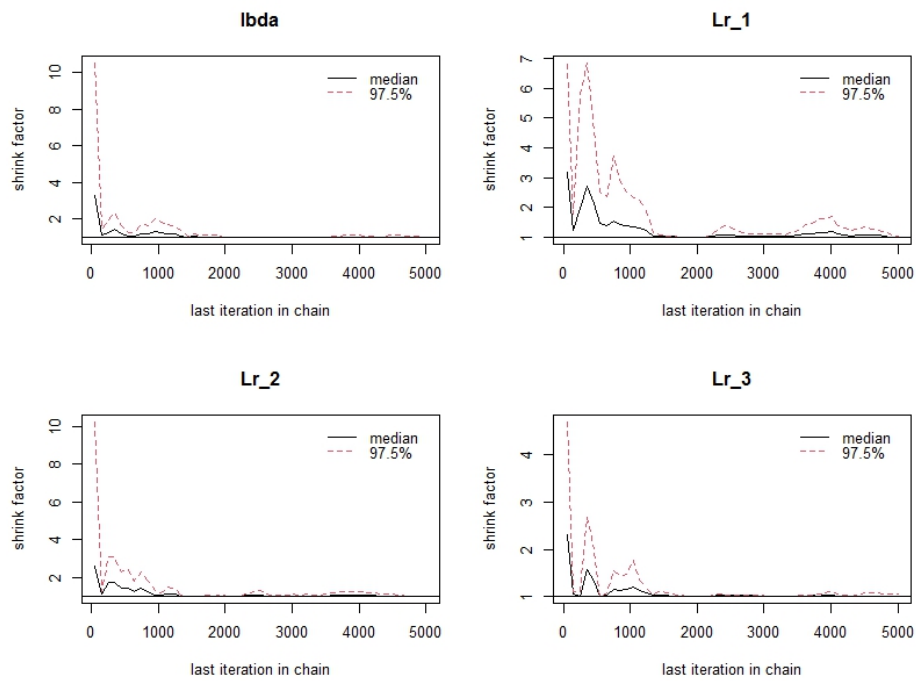
## D.2 Verificação da convergência através do método de Gelman-Rubin para os dados da amostra simulada

Tabela 18 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados aos dados da amostra simulada

	Point est.	Upper c.I.		Point est.	Upper c.I.		Point est.	Upper c.I.
$\lambda$	1.01	1.02	$\mu_{11}$	1.01	1.03	$v_{11}$	1.01	1.04
$\lambda_1^R$	1.01	1.05	$\mu_{21}$	1.01	1.02	$v_{12}$	1.00	1.01
$\lambda_2^R$	1.00	1.00	$\mu_{31}$	1.00	1.02	$v_{13}$	1.01	1.04
$\lambda_3^R$	1.01	1.05	$\mu_{41}$	1.02	1.08	$v_{14}$	1.01	1.01
$\lambda_4^R$	1.01	1.04	$\mu_{12}$	1.28	2.13	$v_{15}$	1.03	1.13
$\lambda_1^C$	1.03	1.13	$\mu_{22}$	1.27	1.90	$v_{21}$	1.08	1.32
$\lambda_2^C$	1.00	1.01	$\mu_{32}$	1.70	3.04	$v_{22}$	1.56	3.59
$\lambda_3^C$	1.01	1.04	$\mu_{42}$	1.25	1.92	$v_{23}$	1.02	1.10
$\lambda_4^C$	1.01	1.02	—	—	—	$v_{24}$	1.66	3.18
$\lambda_5^C$	1.01	1.01	—	—	—	$v_{25}$	1.02	1.02
—	—	—	—	—	—	$l.Log(\theta)$	1.02	1.08

Fonte: Elaborada pelo autor.

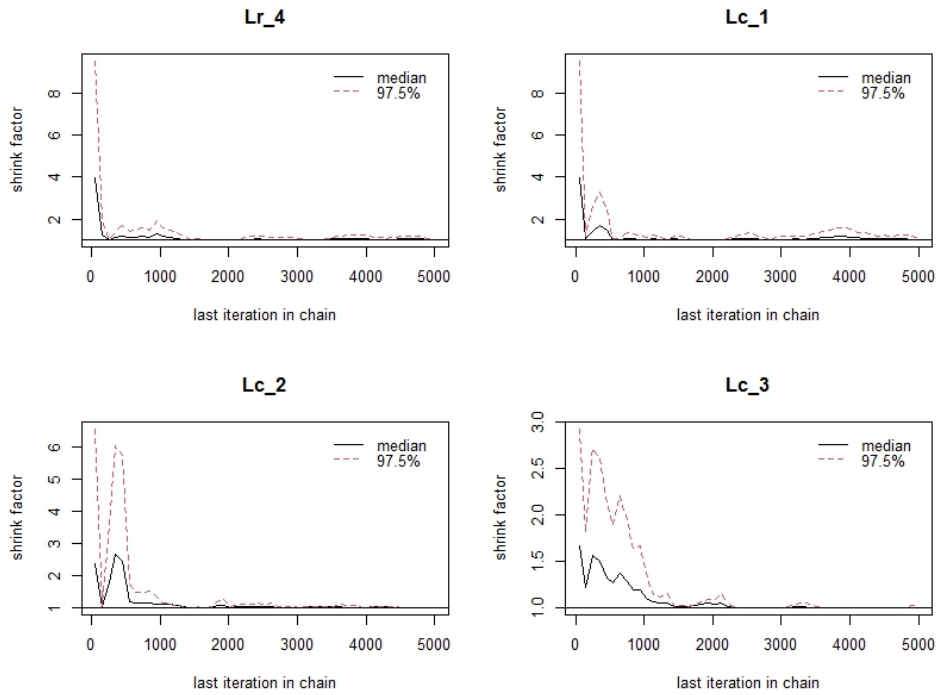
Figura 28 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



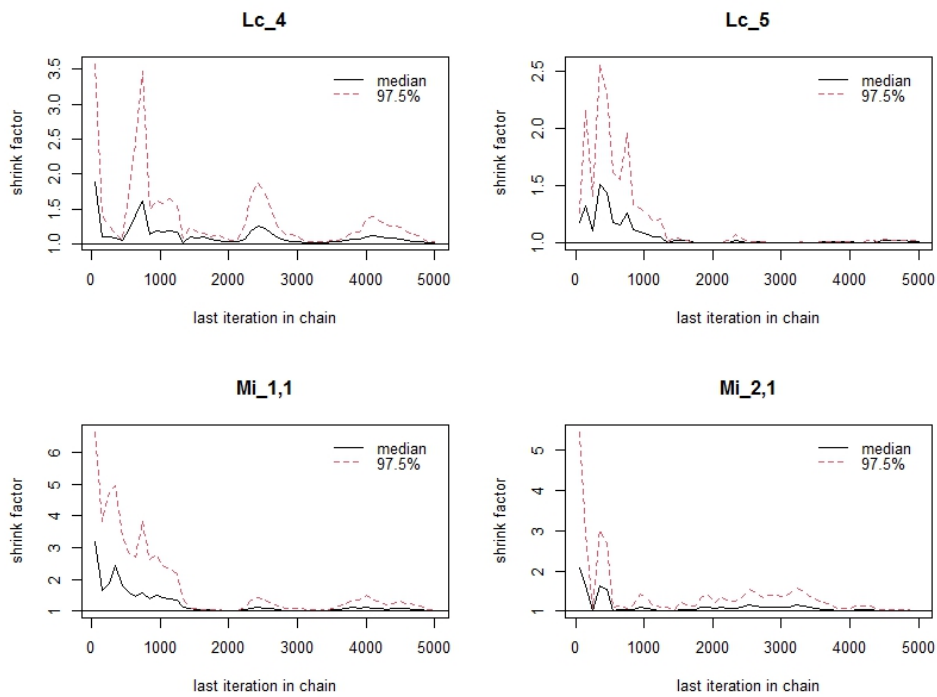
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 29 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$  e  $\mu_{21}$



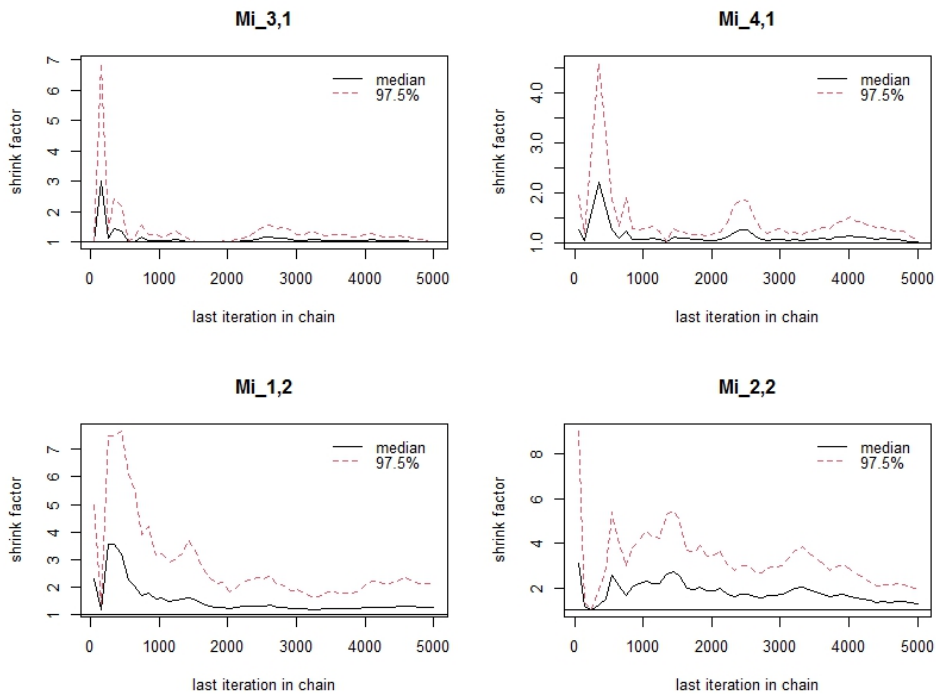
(a)  $\lambda_4^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$  e  $\lambda_3^C$



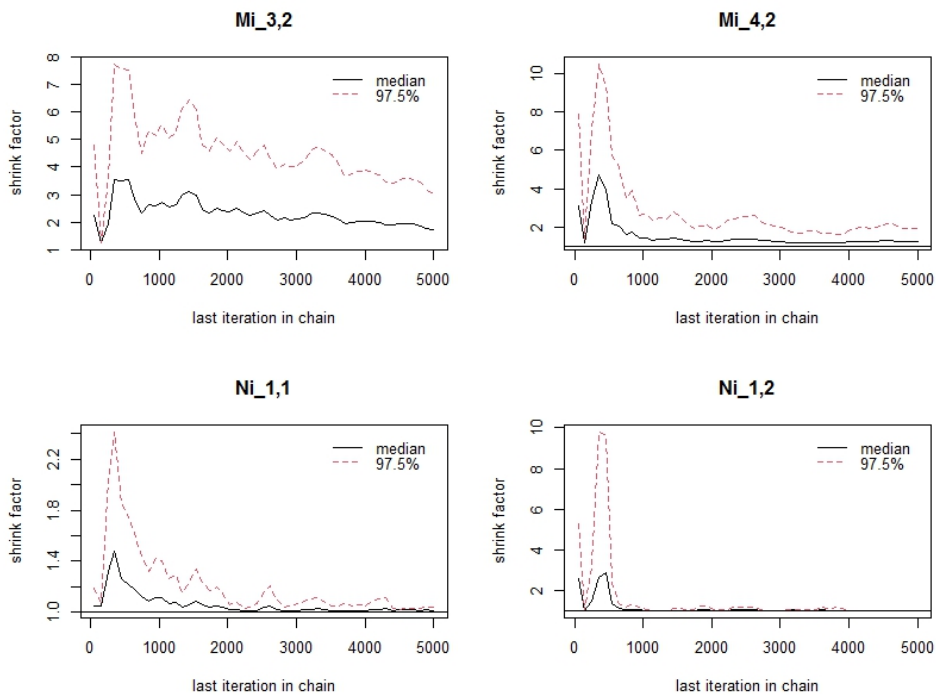
(b)  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$  e  $\mu_{21}$

Fonte: Elaborada pelo autor.

Figura 30 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para  $\mu_{32}, \mu_{42}, \nu_{11}$  e  $\nu_{12}$



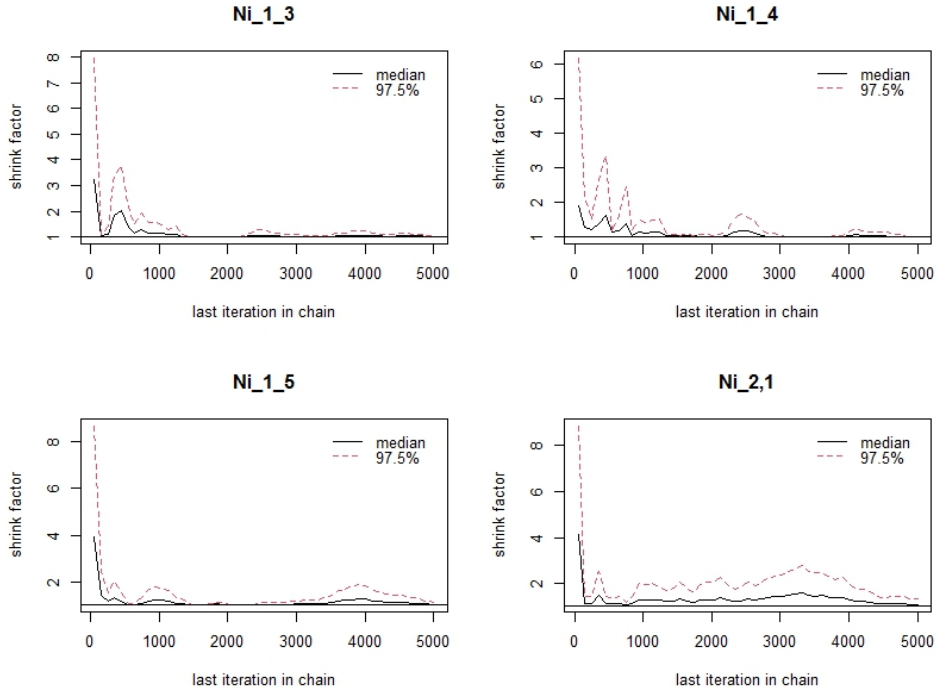
(a)  $\mu_{31}, \mu_{41}, \mu_{12}$  e  $\mu_{22}$



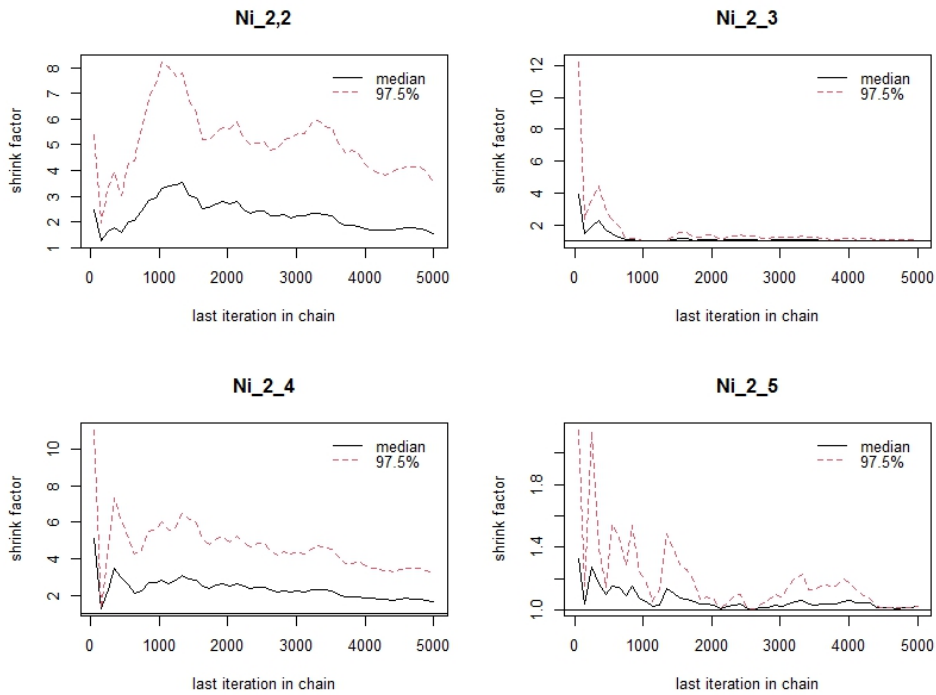
(b)  $\mu_{32}, \mu_{42}, \nu_{11}$  e  $\nu_{12}$

Fonte: Elaborada pelo autor.

Figura 31 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para  $v_{22}, v_{23}, v_{24}$  e  $v_{25}$



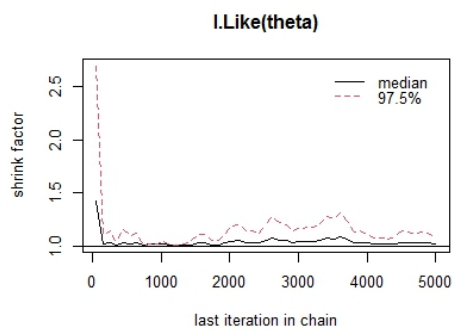
(a)  $v_{13}, v_{14}, v_{15}$  e  $v_{21}$



(b)  $v_{22}, v_{23}, v_{24}$  e  $v_{25}$

Fonte: Elaborada pelo autor.

Figura 32 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados da amostra simulada para log-verossimilhança

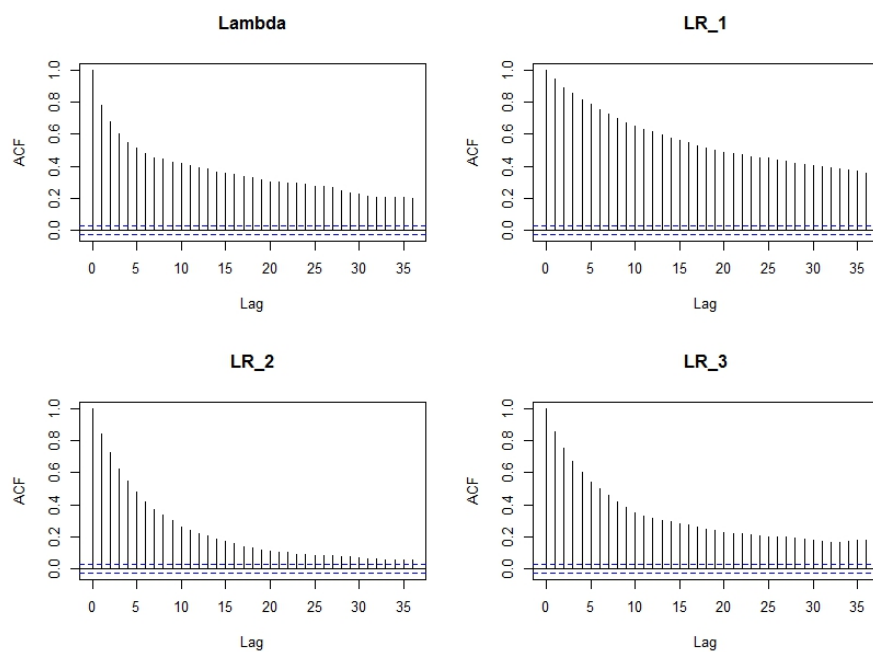


(a) log-verossimilhança

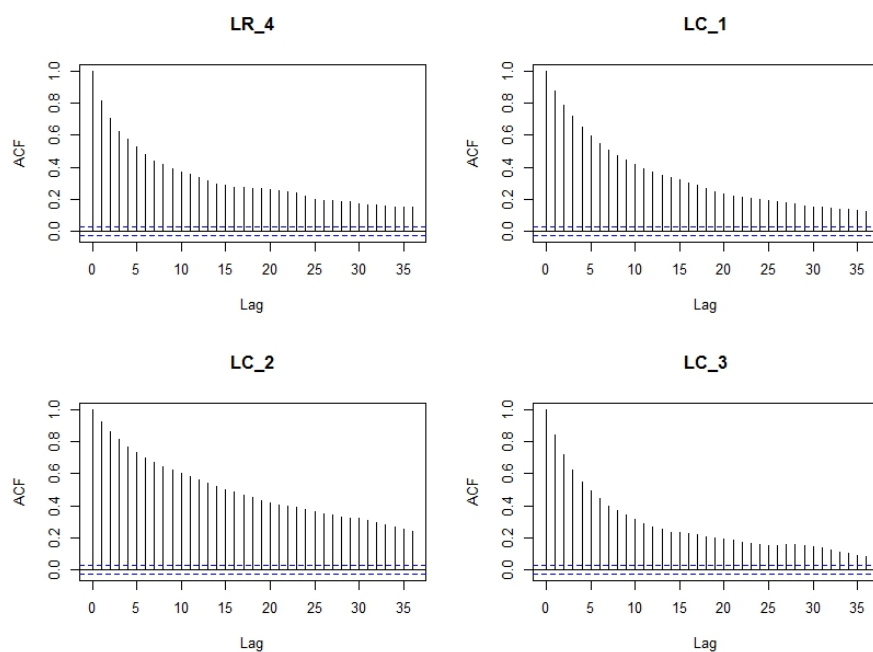
Fonte: Elaborada pelo autor.

### **D.3 Gráficos de auto correlação dos resultados do MH aplicados aos dados da amostra simulada**

Figura 33 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$  e  $\lambda_3^C$



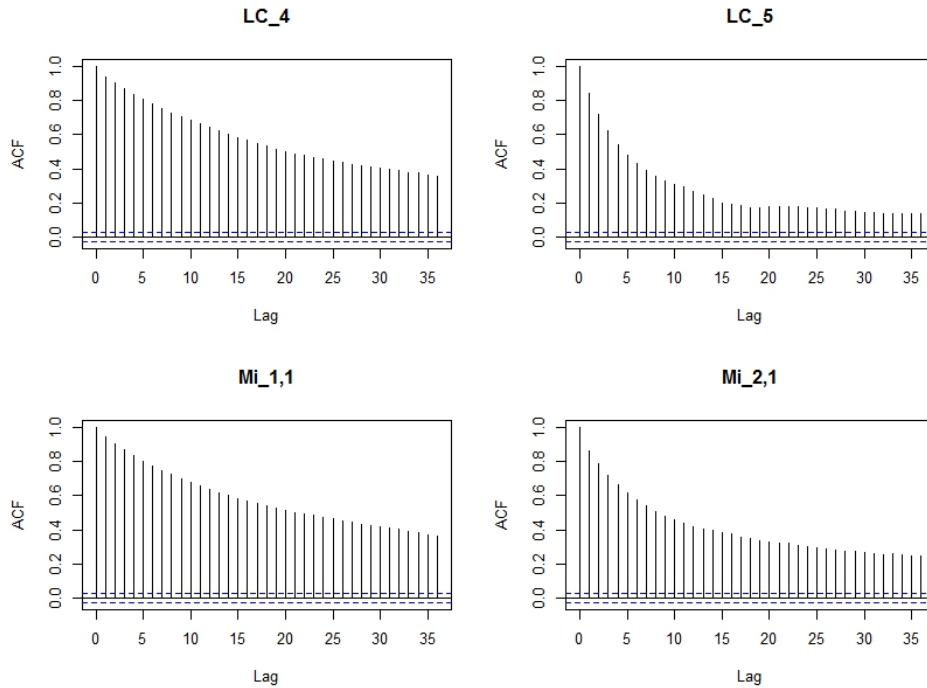
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



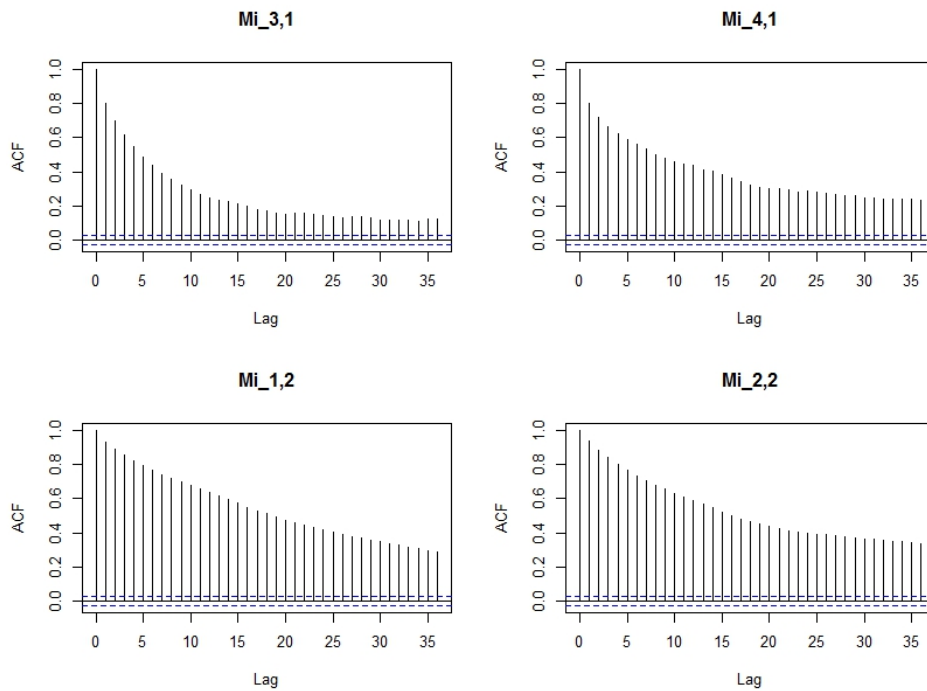
(b)  $\lambda_4^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$  e  $\lambda_3^C$

Fonte: Elaborada pelo autor.

Figura 34 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$ ,  $\mu_{12}$  e  $\mu_{22}$



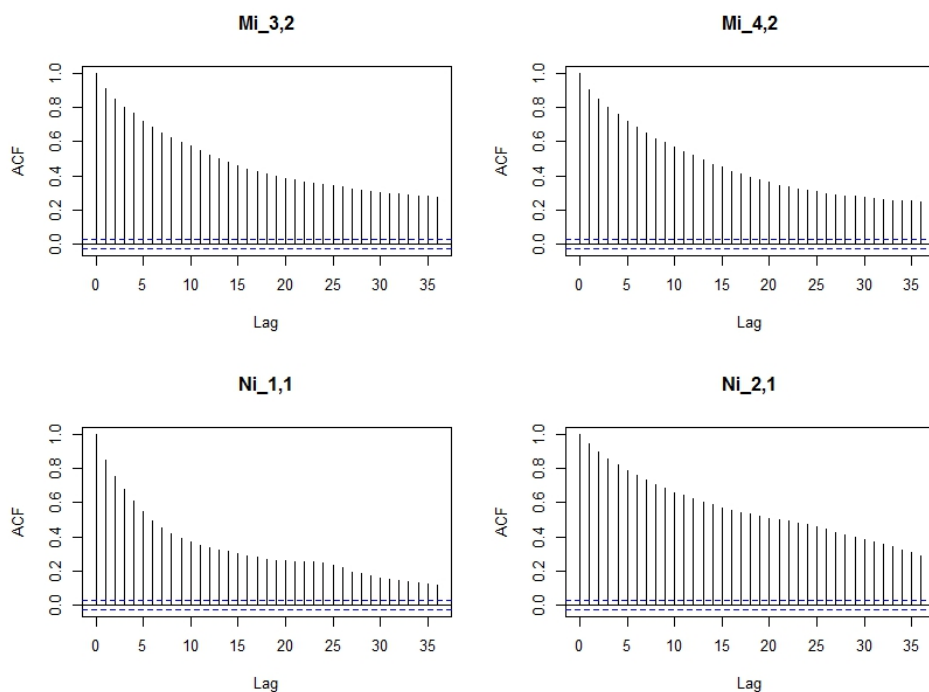
(a)  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$  e  $\mu_{21}$



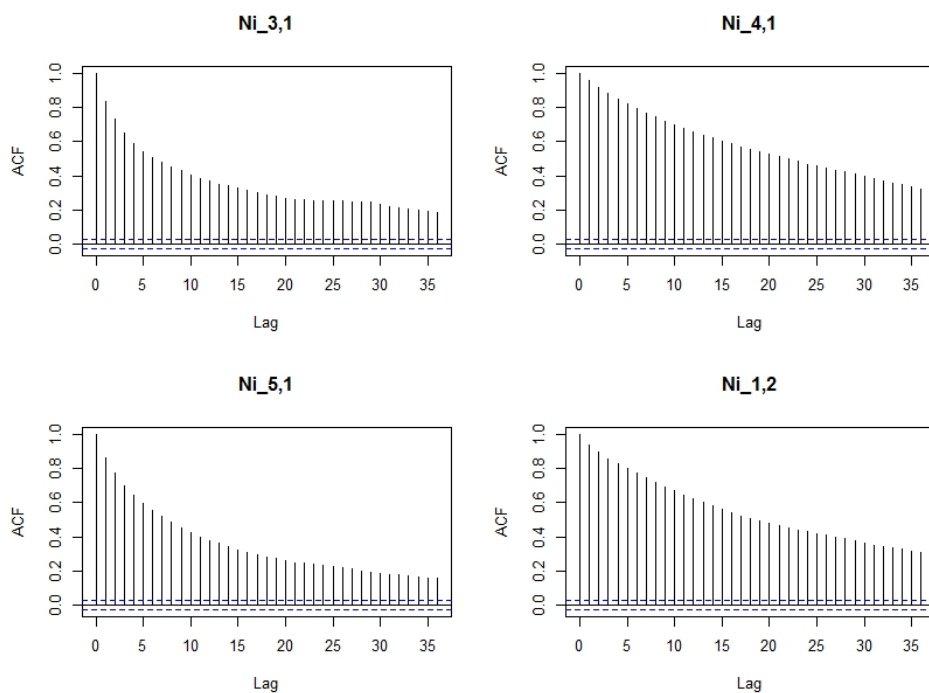
(b)  $\mu_{31}$ ,  $\mu_{41}$ ,  $\mu_{12}$  e  $\mu_{22}$

Fonte: Elaborada pelo autor.

Figura 35 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para  $\mu_{32}, \mu_{42}, v_{11}, v_{21}, v_{31}, v_{41}, v_{51}$  e  $v_{12}$



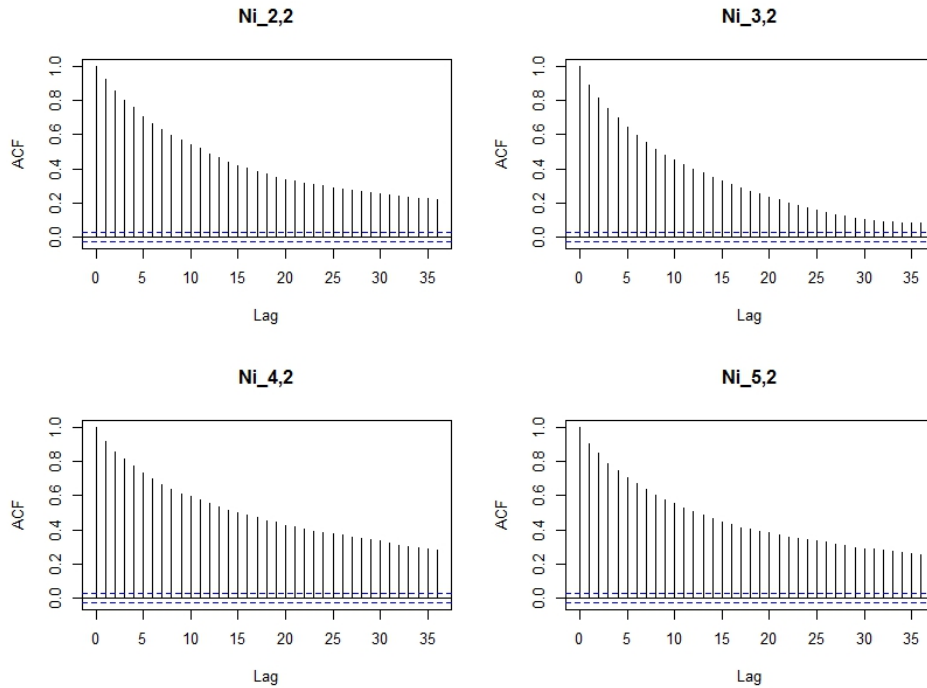
(a)  $\mu_{32}, \mu_{42}, v_{11}$  e  $v_{21}$



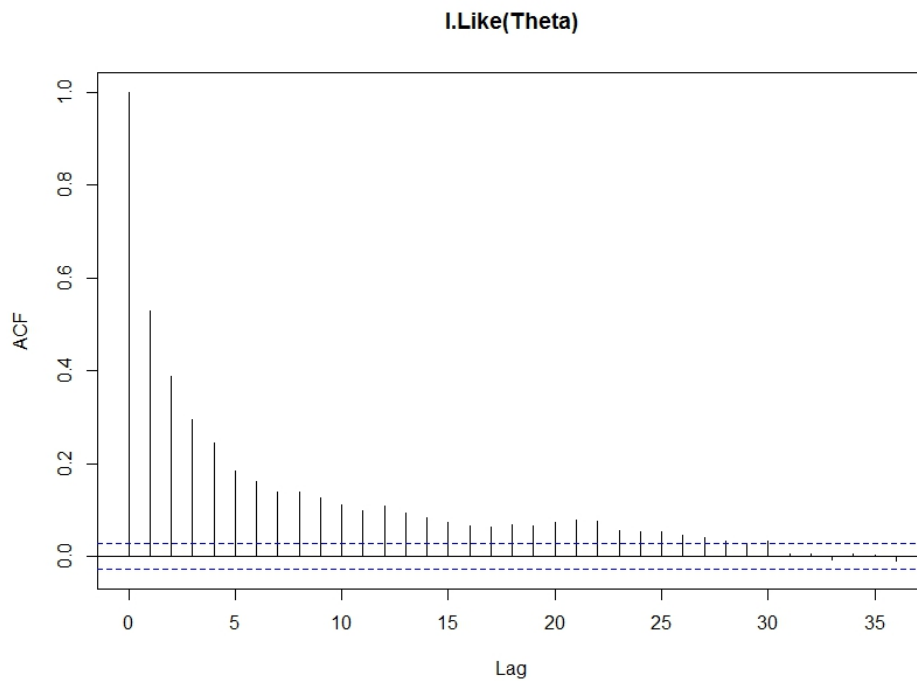
(b)  $v_{31}, v_{41}, v_{51}$  e  $v_{12}$

Fonte: Elaborada pelo autor.

Figura 36 – Auto-correlação dos resultados do MH aplicados aos dados da amostra simulada para  $v_{22}, v_{32}, v_{42}, v_{52}$  e log-verossimilhança



(a)  $v_{22}, v_{32}, v_{42}$  e  $v_{52}$

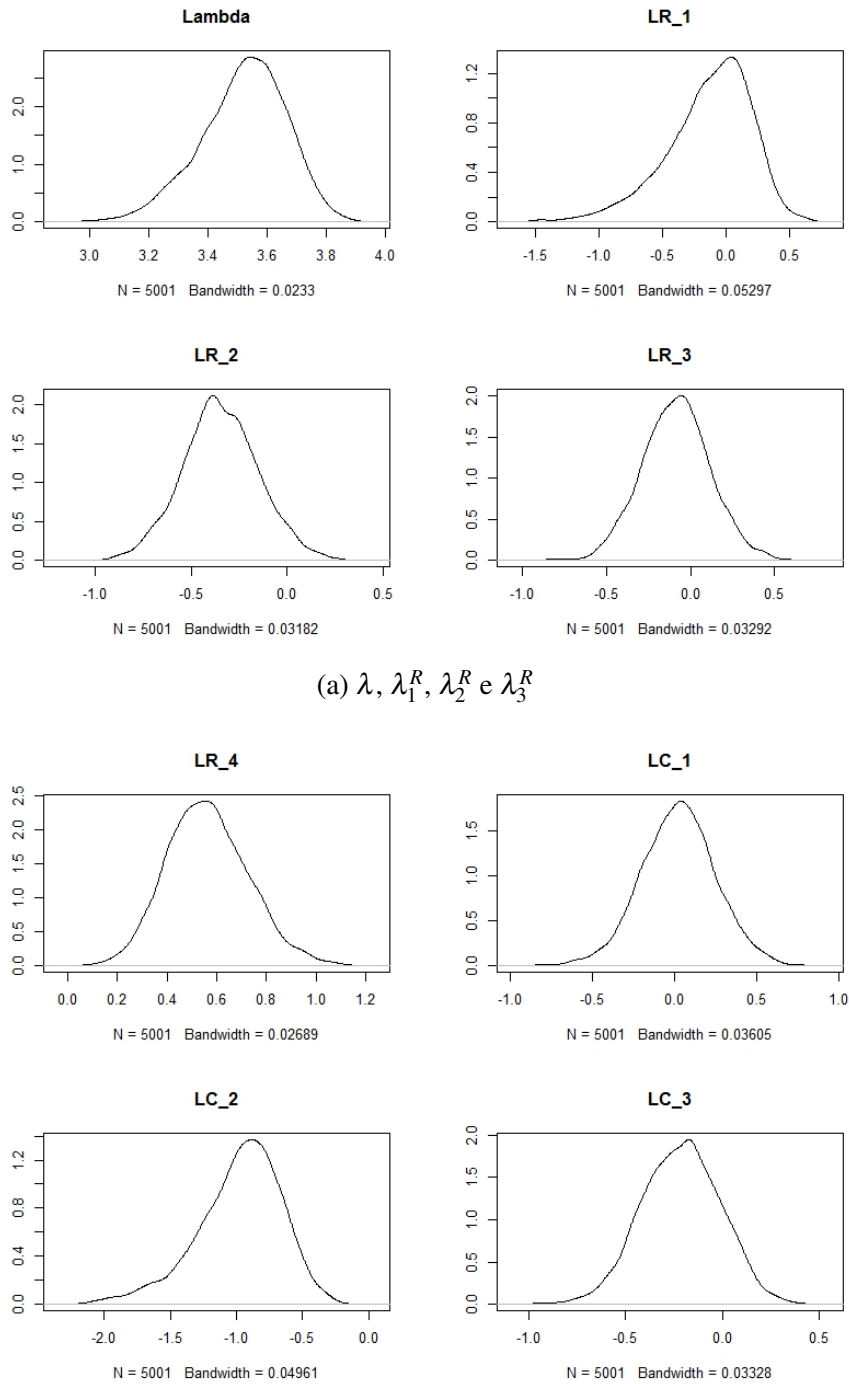


(b) log-verossimilhança

Fonte: Elaborada pelo autor.

## **D.4 Gráficos das distribuições *a posteriori* dos parâmetros do MH aplicados aos dados da amostra simulada**

Figura 37 – Distribuições *a posteriori* do MH aplicados aos dados da amostra simulada para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$  e  $\lambda_3^C$

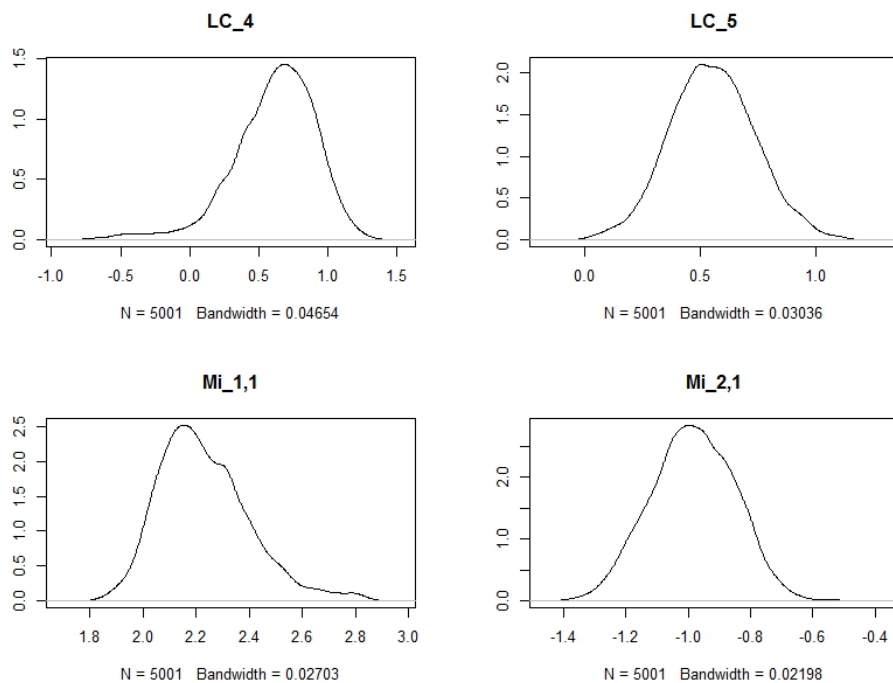


(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

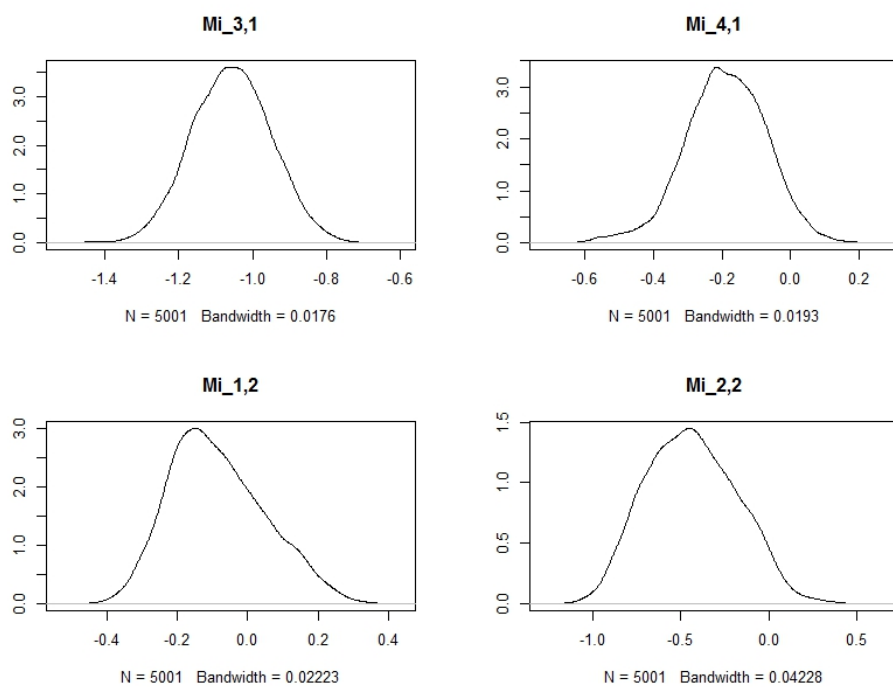
(b)  $\lambda_4^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$  e  $\lambda_3^C$

Fonte: Elaborada pelo autor.

Figura 38 – Distribuições *a posteriori* do MH aplicados aos dados da amostra simulada para  $\lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}, \mu_{12}$  e  $\mu_{22}$



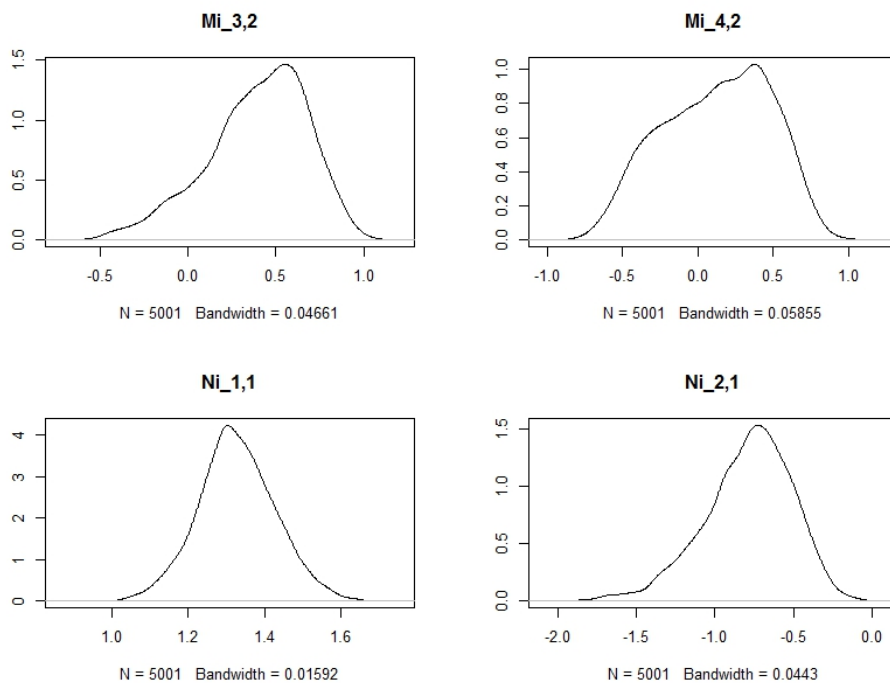
(a)  $\lambda_4^C, \lambda_5^C, \mu_{11}$  e  $\mu_{21}$



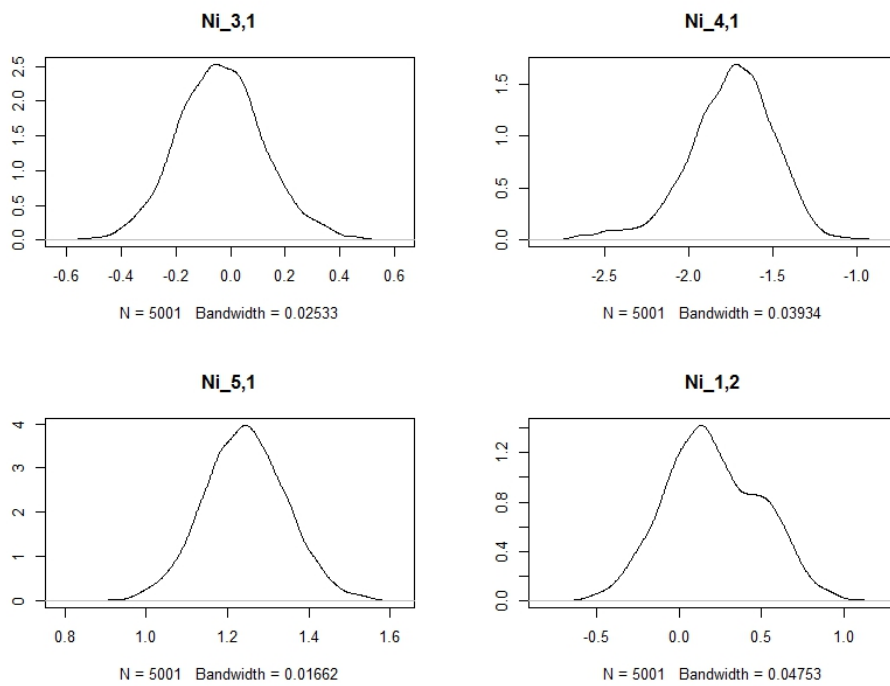
(b)  $\mu_{31}, \mu_{41}, \mu_{12}$  e  $\mu_{22}$

Fonte: Elaborada pelo autor.

Figura 39 – Distribuições *a posteriori* do MH aplicados aos dados da amostra simulada para  $\mu_{32}, \mu_{42}, \nu_{11}, \nu_{21}, \nu_{31}, \nu_{41}, \nu_{51}$  e  $\nu_{12}$



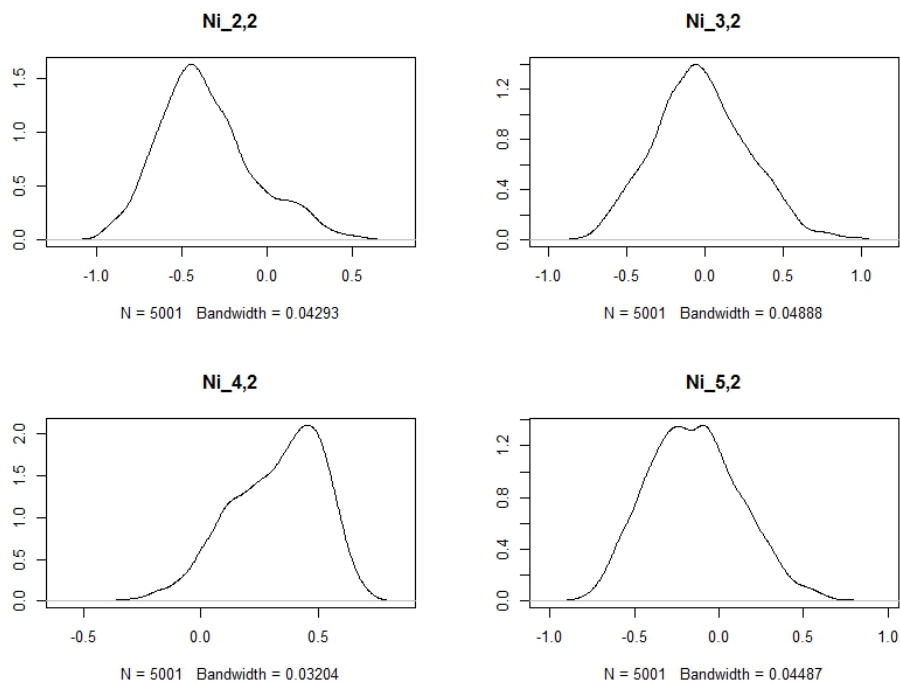
(a)  $\mu_{32}, \mu_{42}, \nu_{11}$  e  $\nu_{21}$



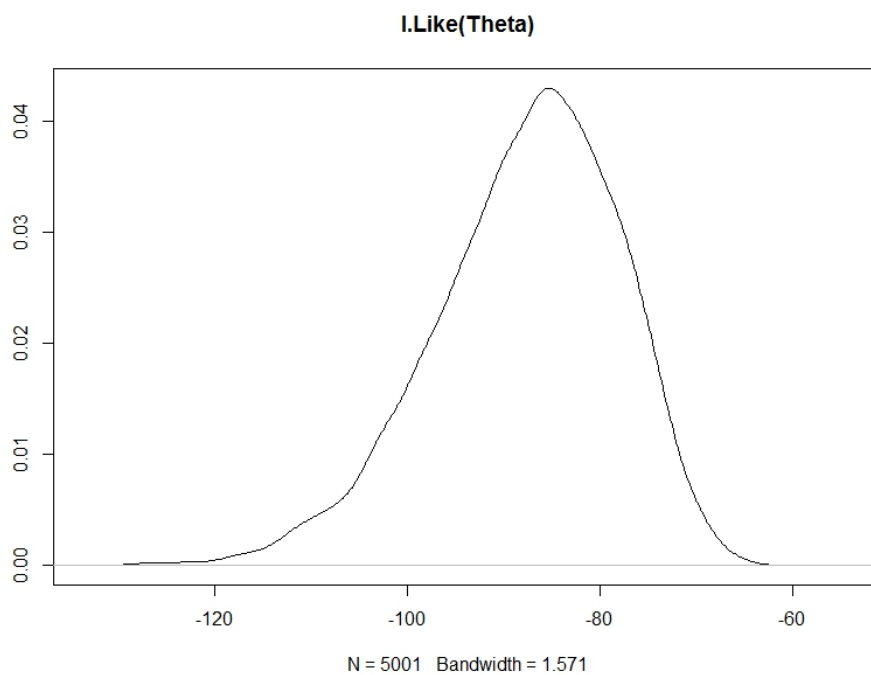
(b)  $\nu_{31}, \nu_{41}, \nu_{51}$  e  $\nu_{12}$

Fonte: Elaborada pelo autor.

Figura 40 – Distribuições *a posteriori* do MH aplicados aos dados da amostra simulada para log-verossimilhança



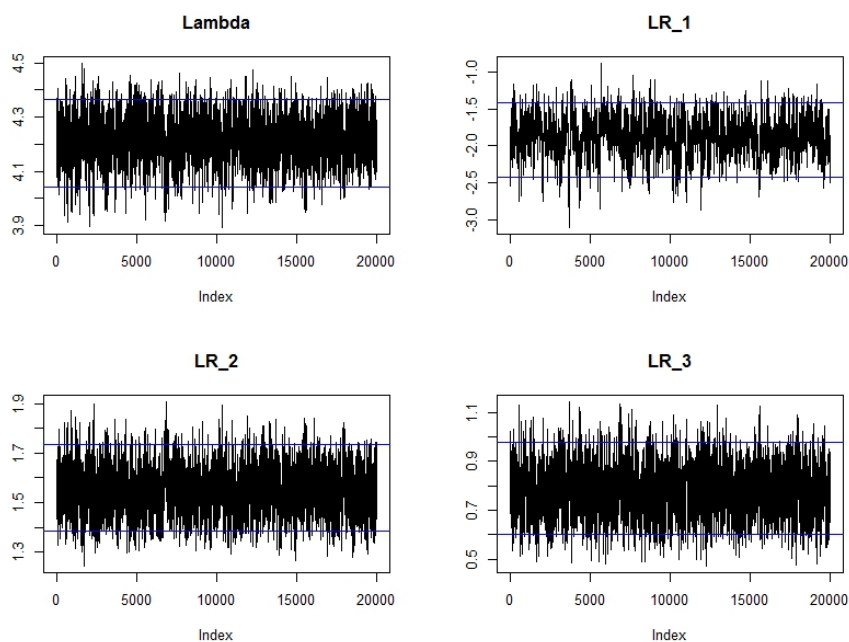
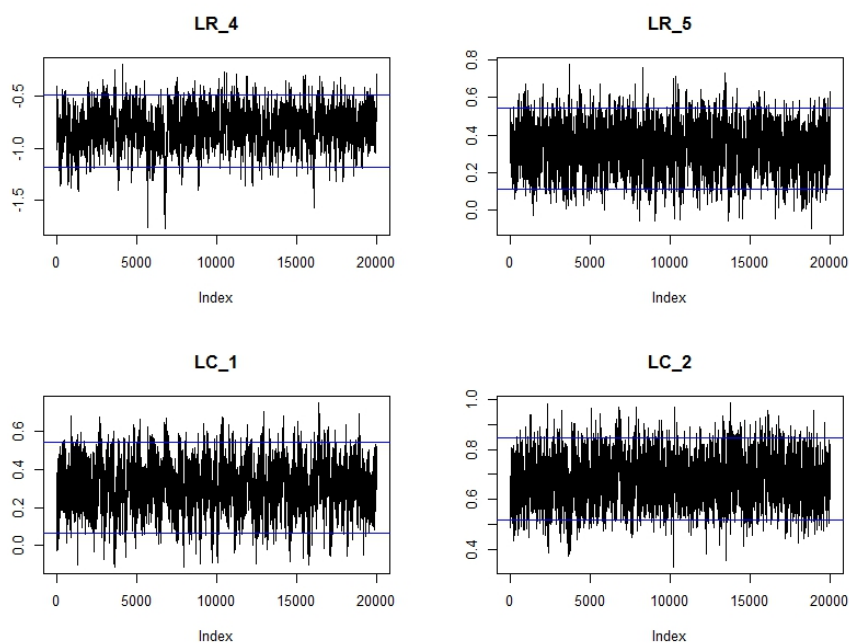
(a)  $v_{22}, v_{32}, v_{42}$  e  $v_{52}$



(b) log-verossimilhança

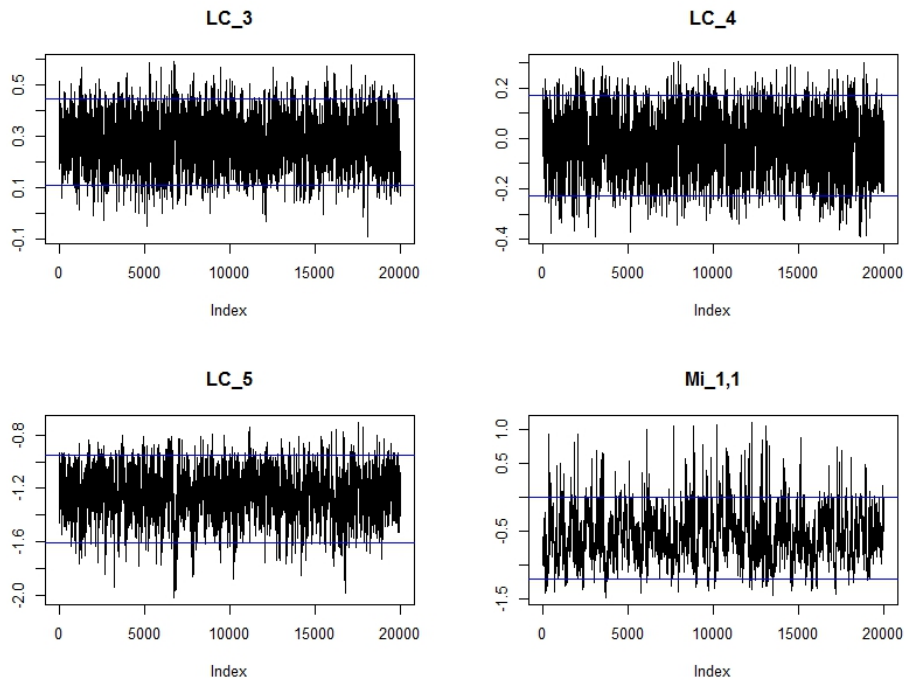
Fonte: Elaborada pelo autor.

## **D.5 Gráficos trace plot dos resultados da cadeia aplicados aos dados do artigo comparativo**

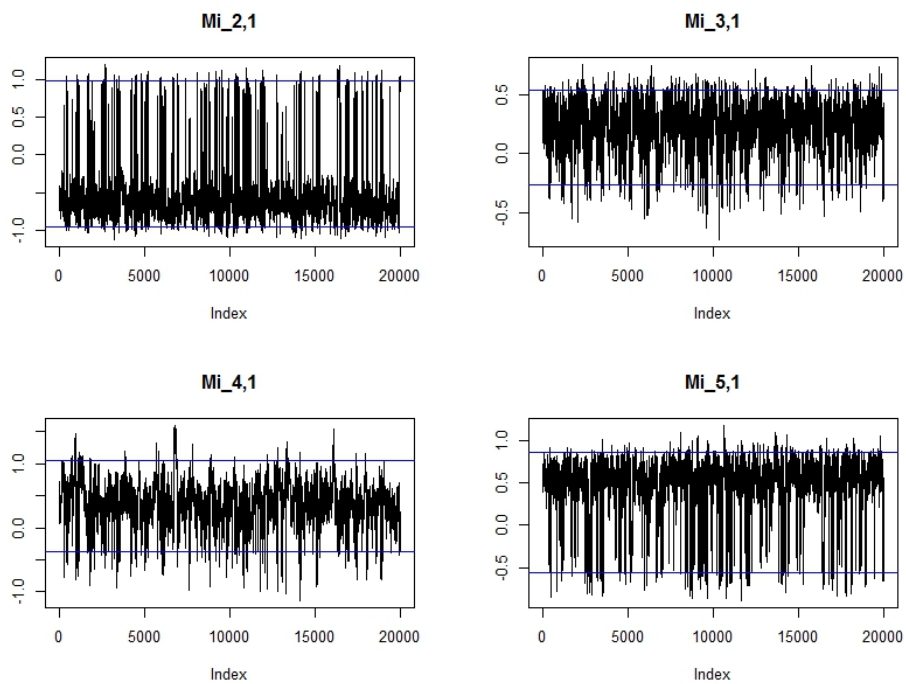
Figura 41 – Trace plot dos resultados do MH aplicados no estudo comparativo para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$ (a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$ (b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$ 

Fonte: Elaborada pelo autor.

Figura 42 – Trace plot dos resultados do MH aplicados no estudo comparativo para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

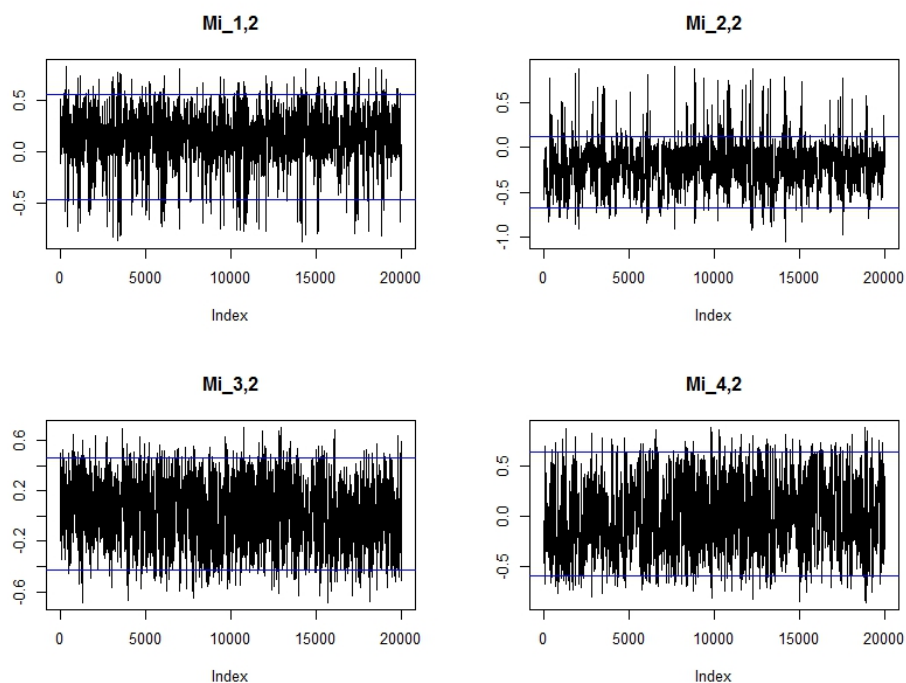
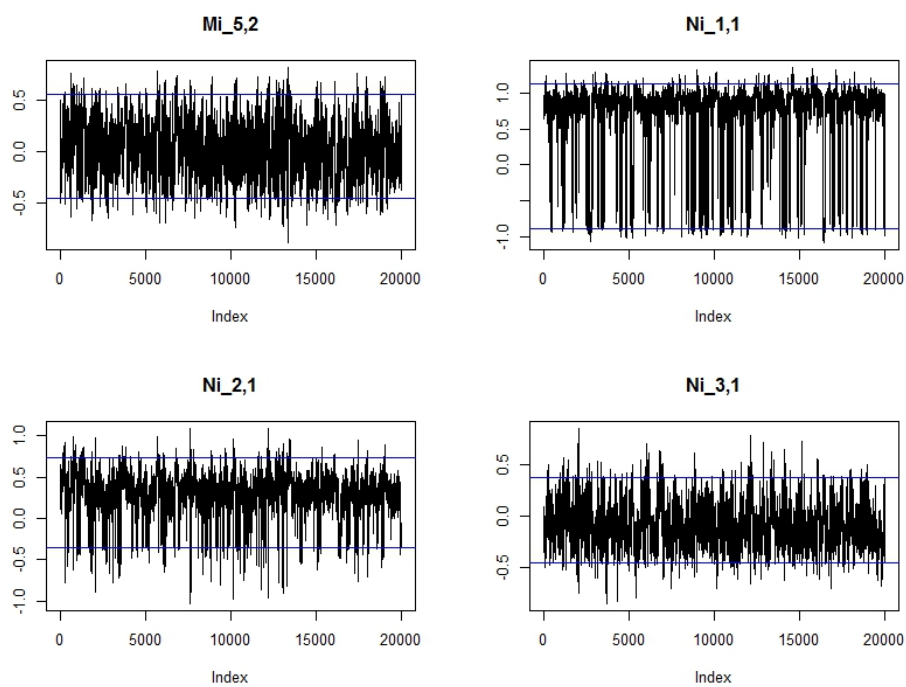


(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



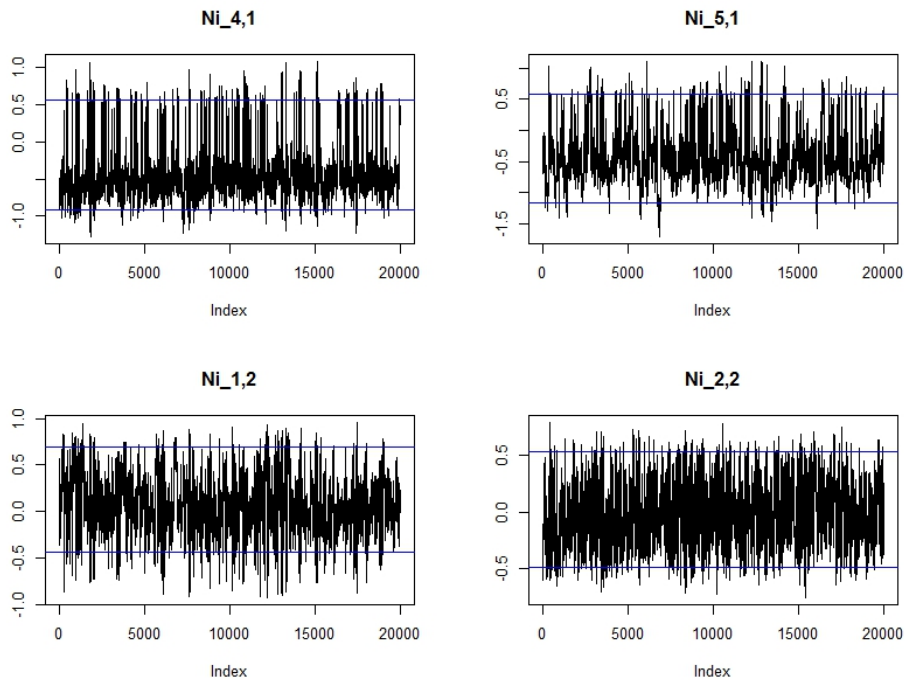
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

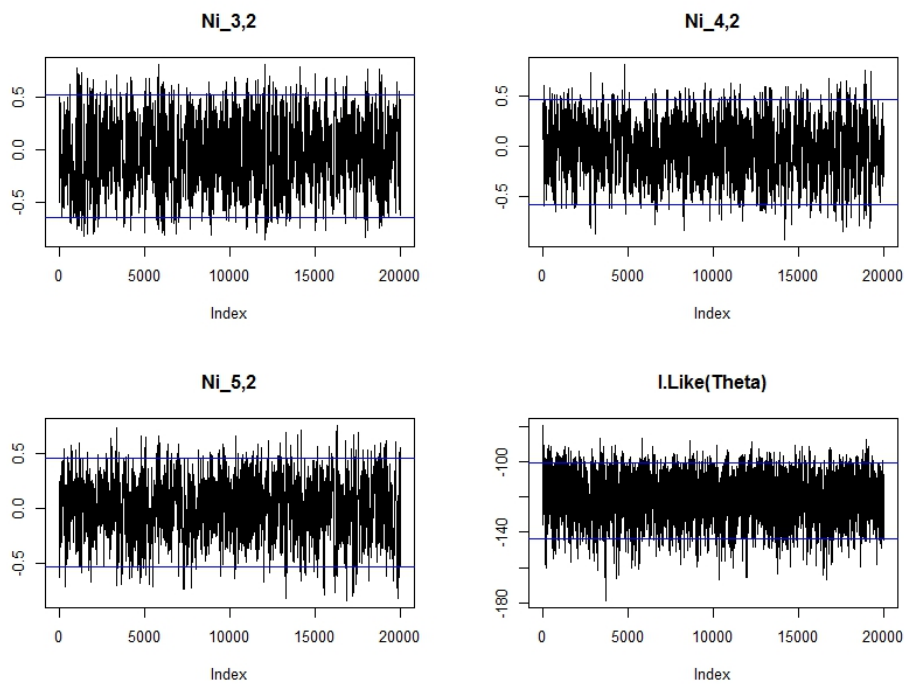
Figura 43 – Trace plot dos resultados do MH aplicados no estudo comparativo para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$ (a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$ (b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$ 

Fonte: Elaborada pelo autor.

Figura 44 – Trace plot dos resultados do MH aplicados no estudo comparativo para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

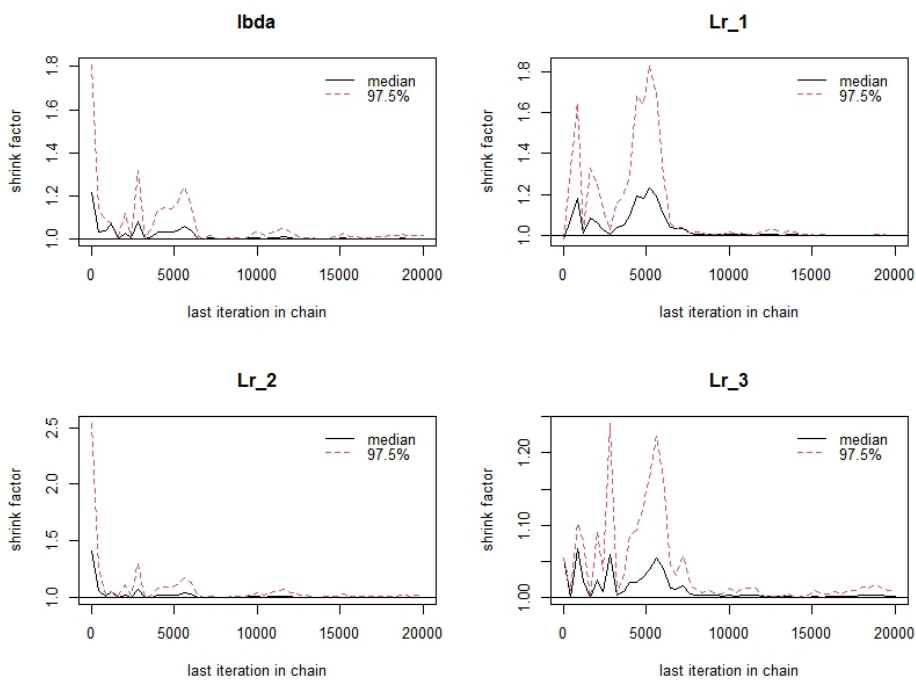
## D.6 Verificação da convergência através do método de Gelman-Rubin para os dados do artigo comparativo

Tabela 19 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados aos dados do artigo comparativo

	Point est.	Upper c.I.		Point est.	Upper c.I.		Point est.	Upper c.I.
$\lambda$	1.00	1.02	$\mu_{11}$	1.00	1.01	$v_{11}$	1.00	1.00
$\lambda_1^R$	1.00	1.00	$\mu_{21}$	1.00	1.00	$v_{12}$	1.00	1.00
$\lambda_2^R$	1.00	1.01	$\mu_{31}$	1.00	1.00	$v_{13}$	1.00	1.00
$\lambda_3^R$	1.00	1.01	$\mu_{41}$	1.01	1.03	$v_{14}$	1.00	1.01
$\lambda_4^R$	1.01	1.03	$\mu_{51}$	1.00	1.00	$v_{15}$	1.00	1.00
$\lambda_5^R$	1.00	1.01	$\mu_{12}$	1.00	1.00	$v_{21}$	1.00	1.00
$\lambda_1^C$	1.00	1.01	$\mu_{22}$	1.00	1.01	$v_{22}$	1.00	1.00
$\lambda_2^C$	1.00	1.01	$\mu_{32}$	1.00	1.00	$v_{23}$	1.00	1.00
$\lambda_3^C$	1.00	1.00	$\mu_{42}$	1.00	1.01	$v_{24}$	1.00	1.01
$\lambda_4^C$	1.00	1.01	$\mu_{52}$	1.00	1.02	$v_{25}$	1.00	1.01
$\lambda_5^C$	1.00	1.01	---	---	—	$l.Log(\theta)$	1.00	1.00

Fonte: Elaborada pelo autor.

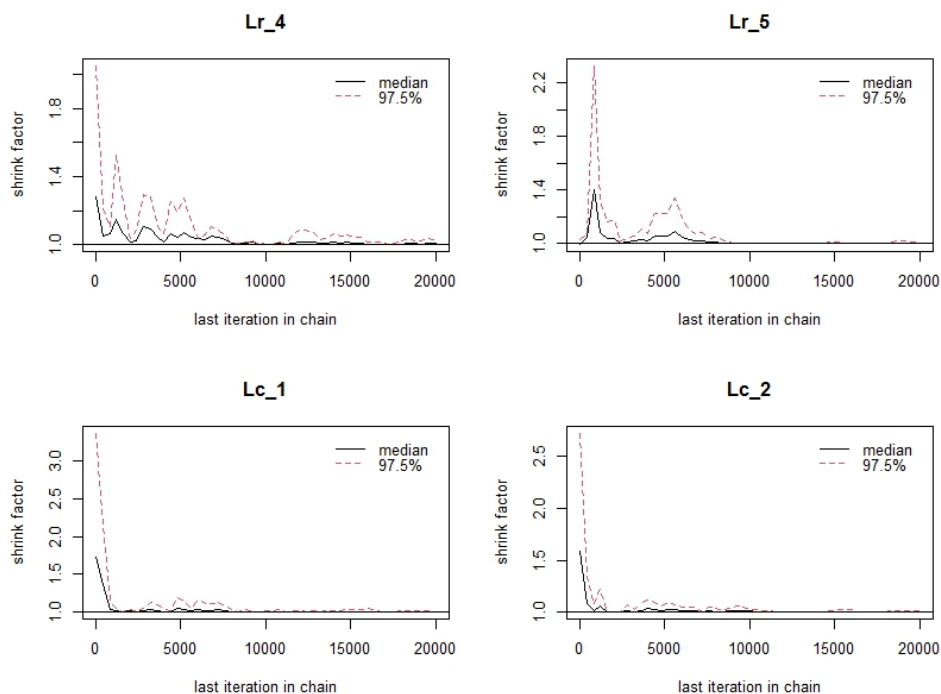
Figura 45 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



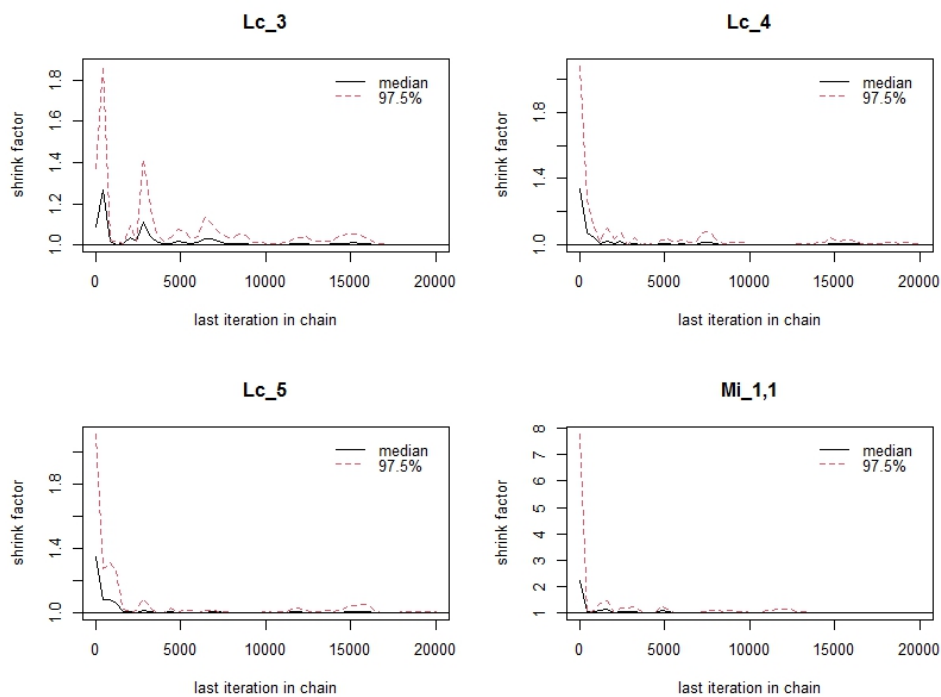
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 46 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$ ,  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



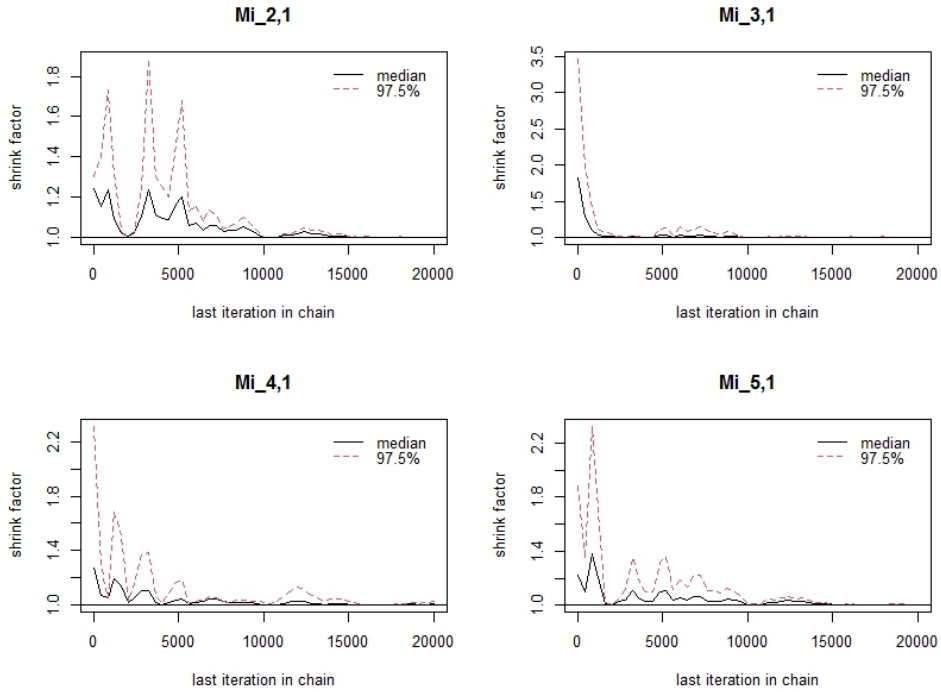
(a)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



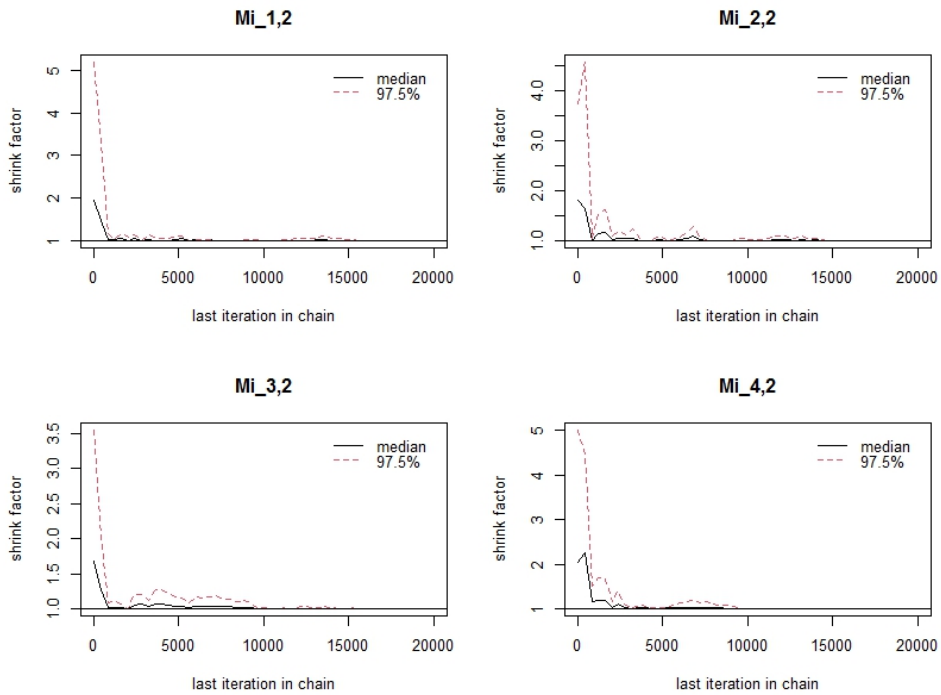
(b)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$

Fonte: Elaborada pelo autor.

Figura 47 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para  $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$



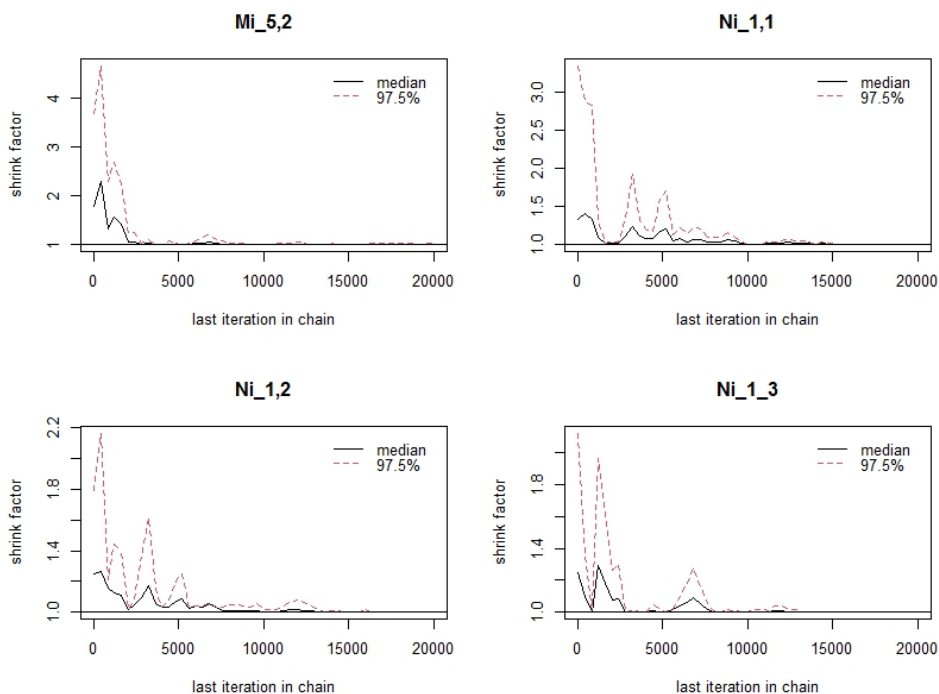
(a)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



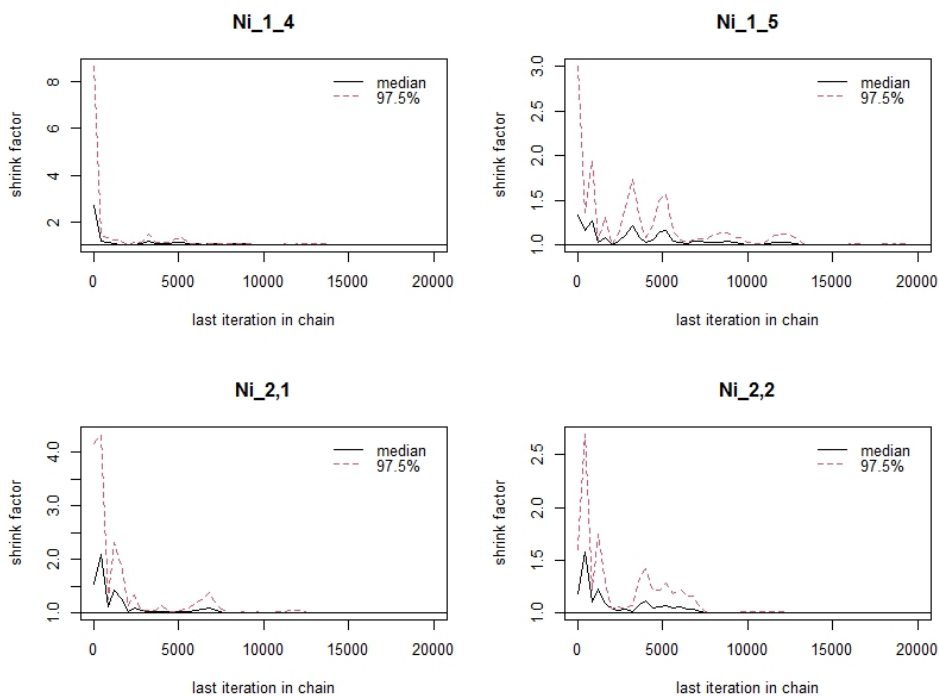
(b)  $\mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$

Fonte: Elaborada pelo autor.

Figura 48 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para  $\mu_{52}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{21}$  e  $v_{22}$



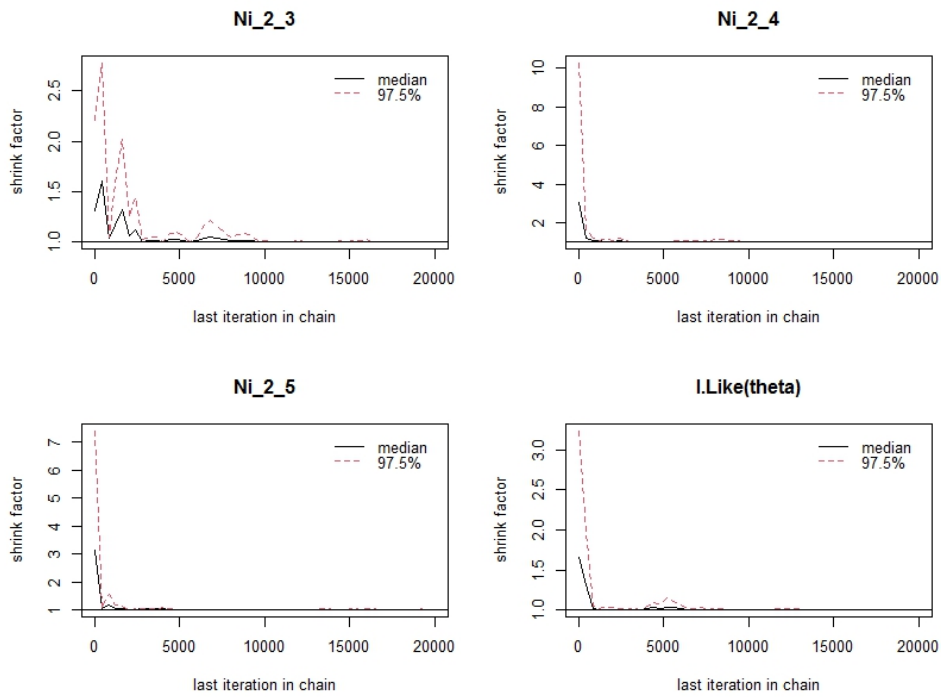
(a)  $\mu_{52}, v_{11}, v_{12}$  e  $v_{13}$



(b)  $v_{14}, v_{15}, v_{21}$  e  $v_{22}$

Fonte: Elaborada pelo autor.

Figura 49 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicado nos dados do estudo comparativo para  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

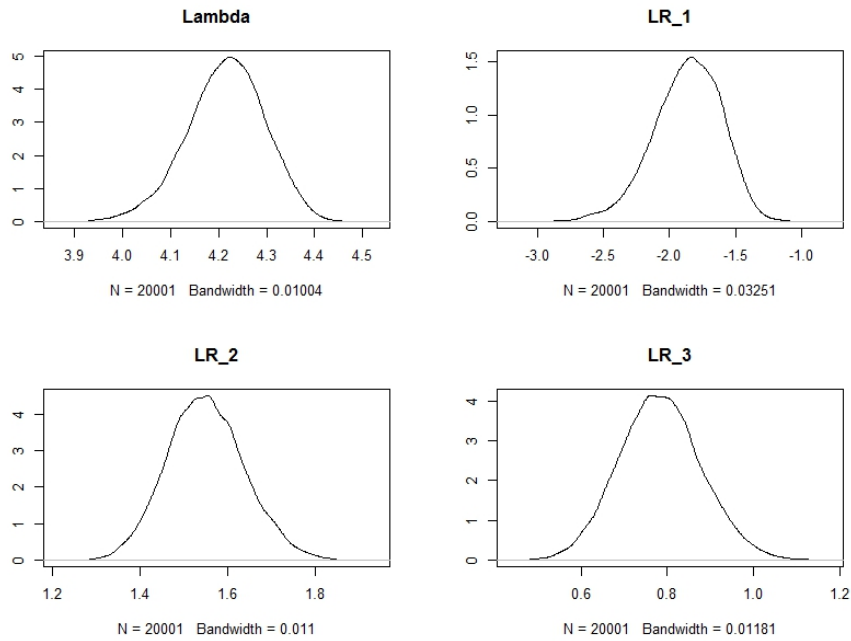


(a)  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e logLikelihood

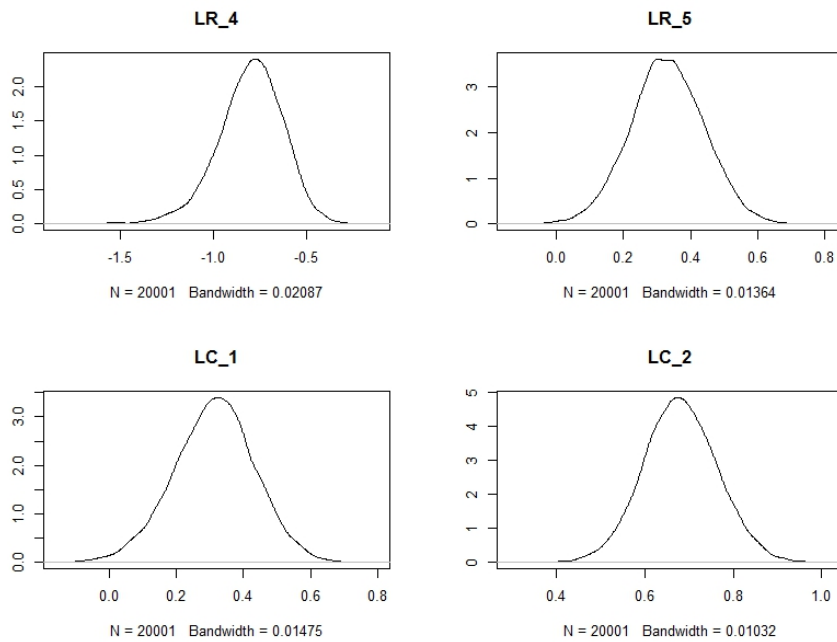
Fonte: Elaborada pelo autor.

## **D.7 Gráficos das distribuições *a posteriori* dos resultados do MH aplicado aos dados do artigo comparativo**

Figura 50 – Distribuições *a posteriori* do MH aplicados aos dados do estudo comparativo para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



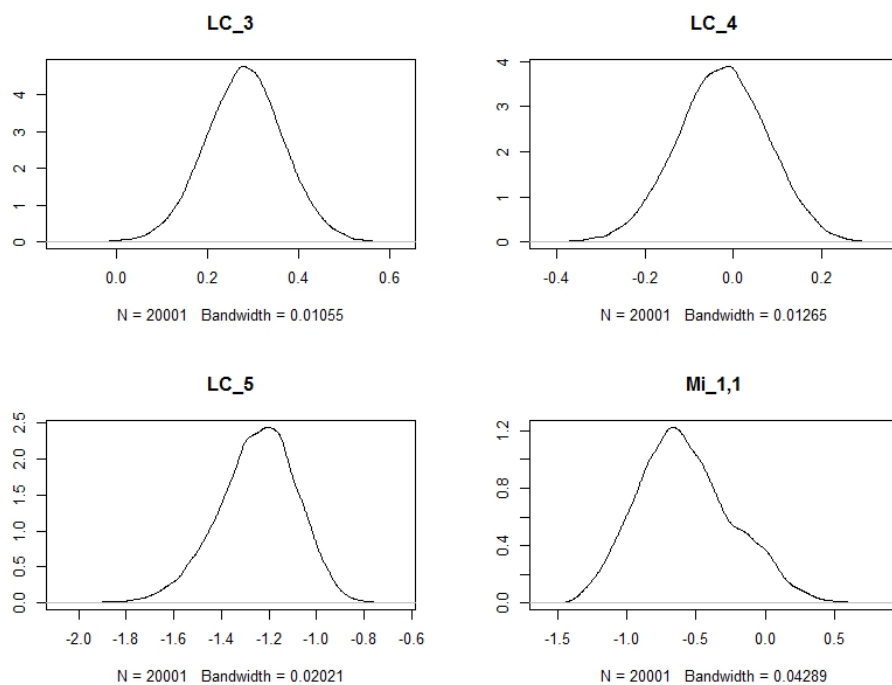
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



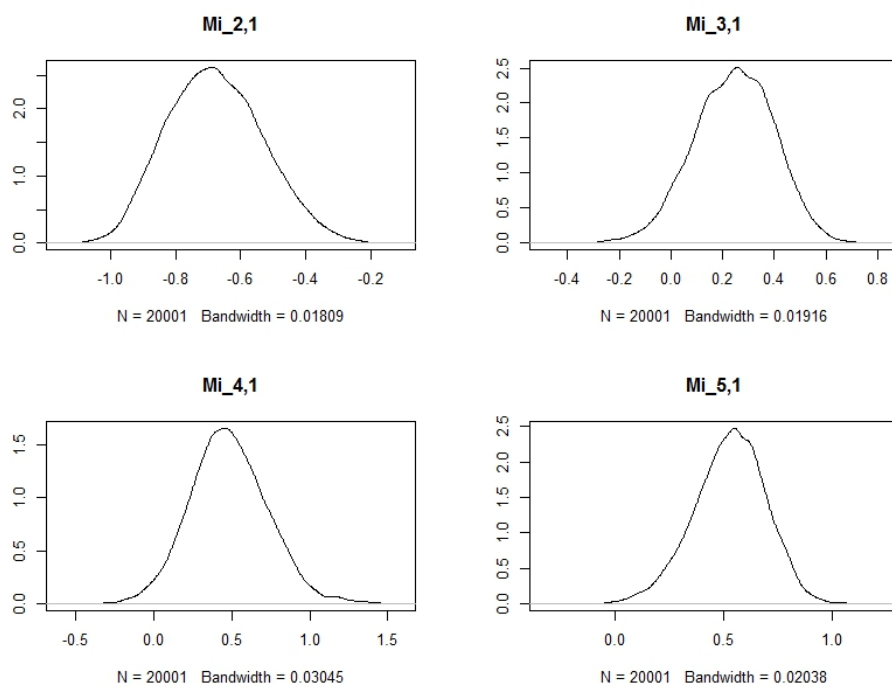
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 51 – Distribuições *a posteriori* do MH aplicados aos dados do estudo comparativo para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



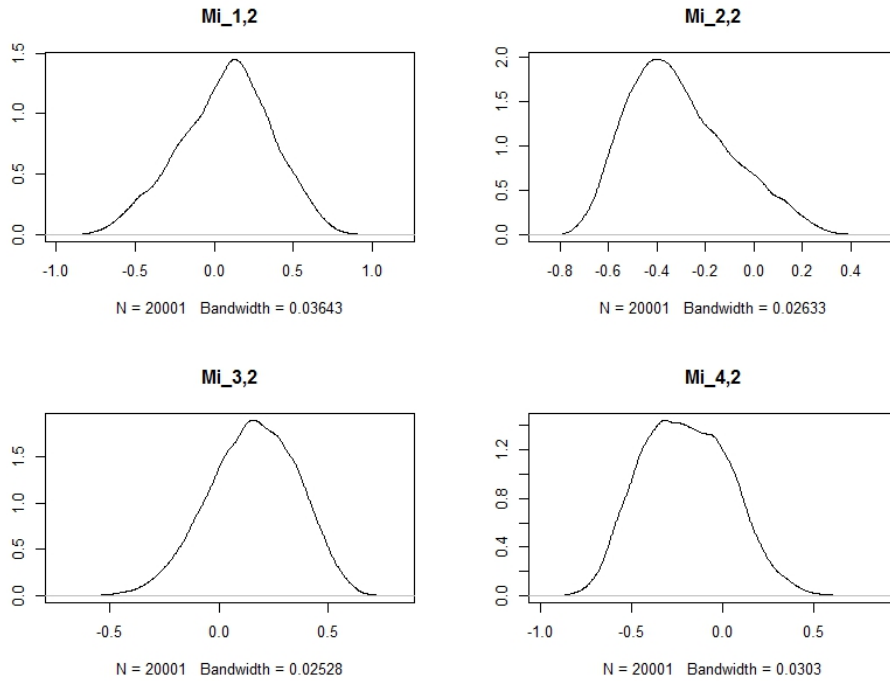
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



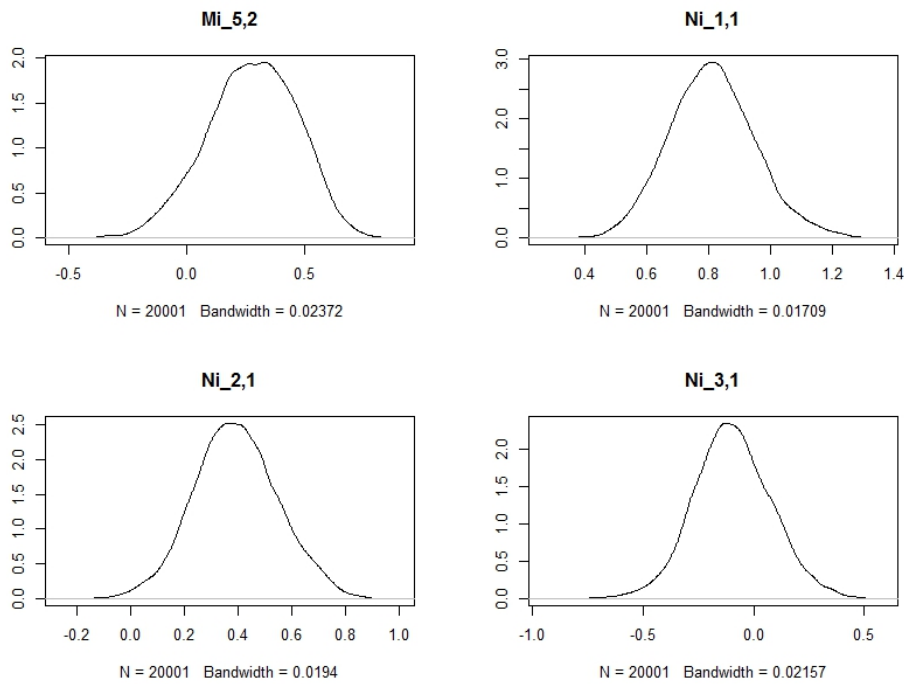
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 52 – Distribuições *a posteriori* do MH aplicados aos dados do estudo comparativo para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



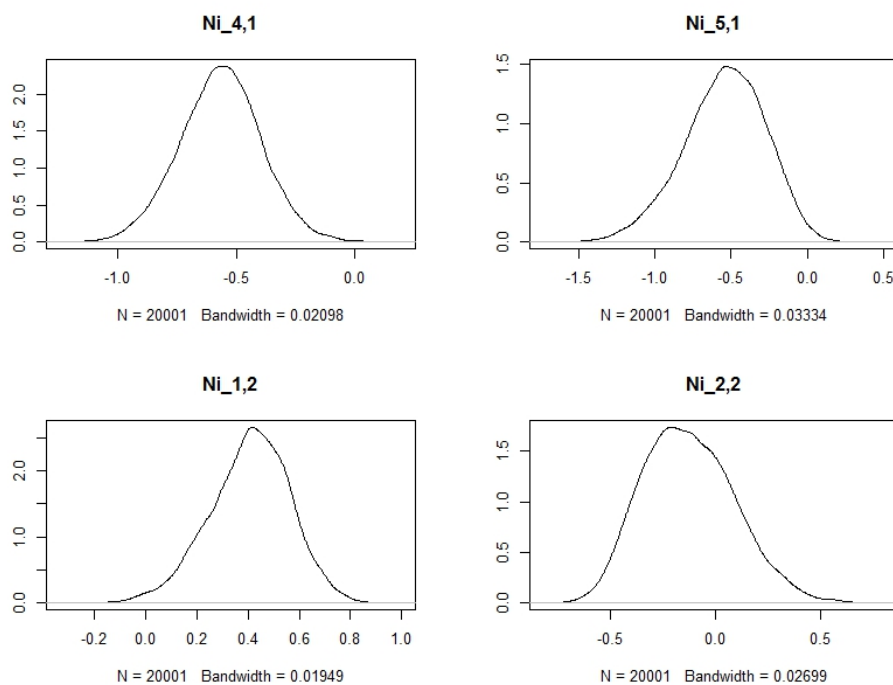
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



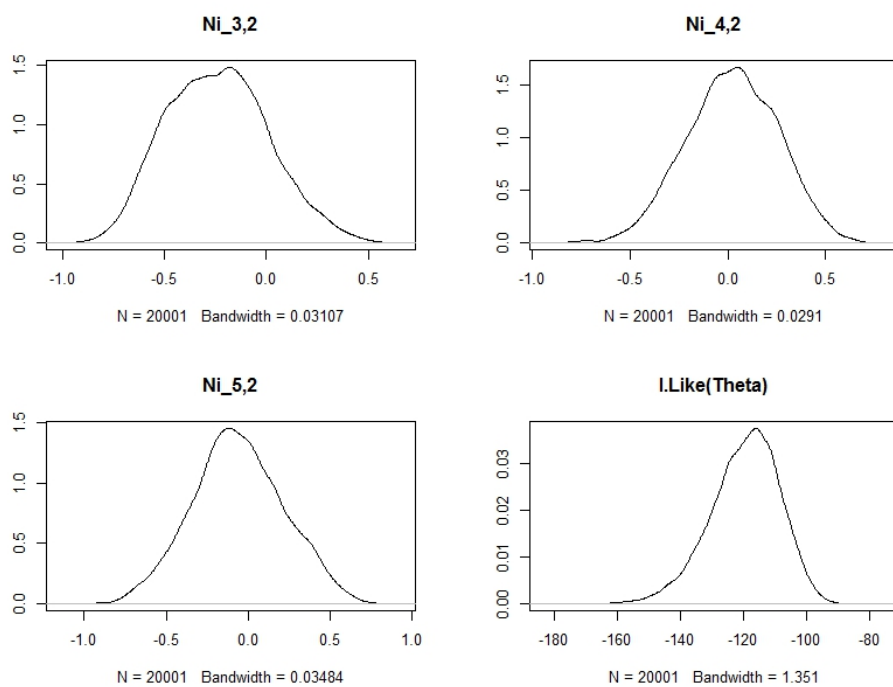
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 53 – Distribuições *a posteriori* do MH aplicados aos dados do estudo comparativo para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$

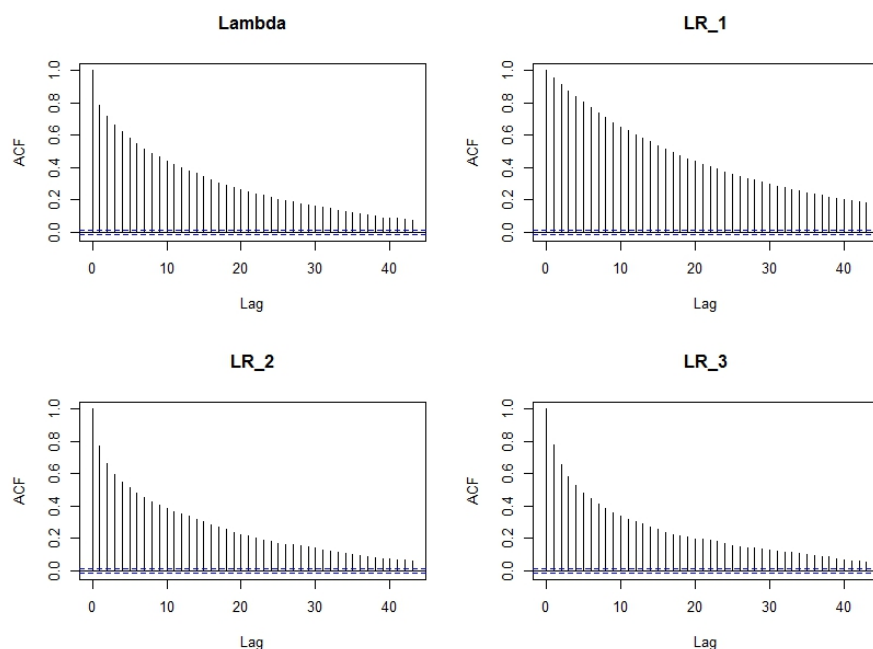


(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

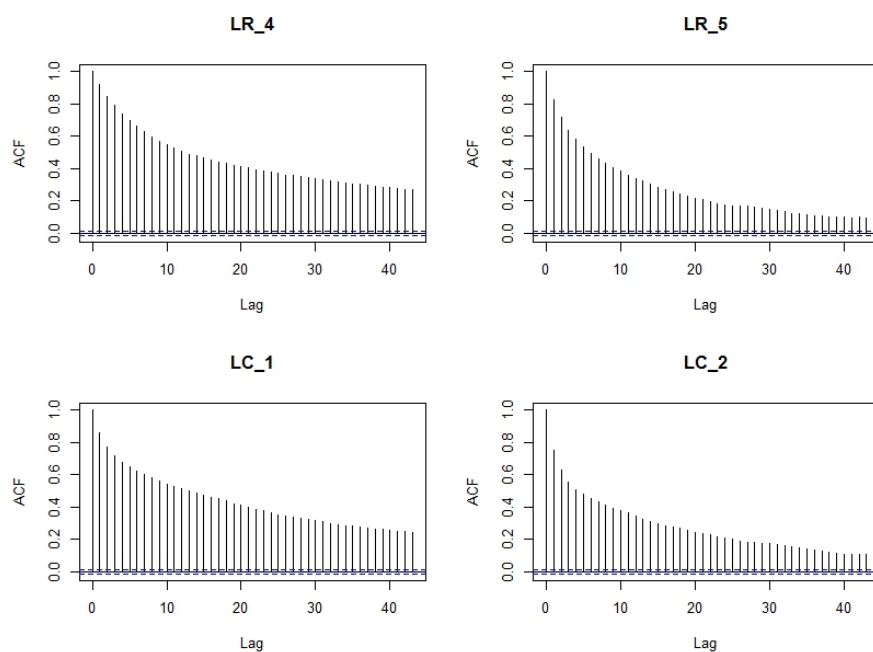
Fonte: Elaborada pelo autor.

## **D.8 Gráficos de auto correlação dos resultados do MH aplicados nos dados do artigo comparativo**

Figura 54 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



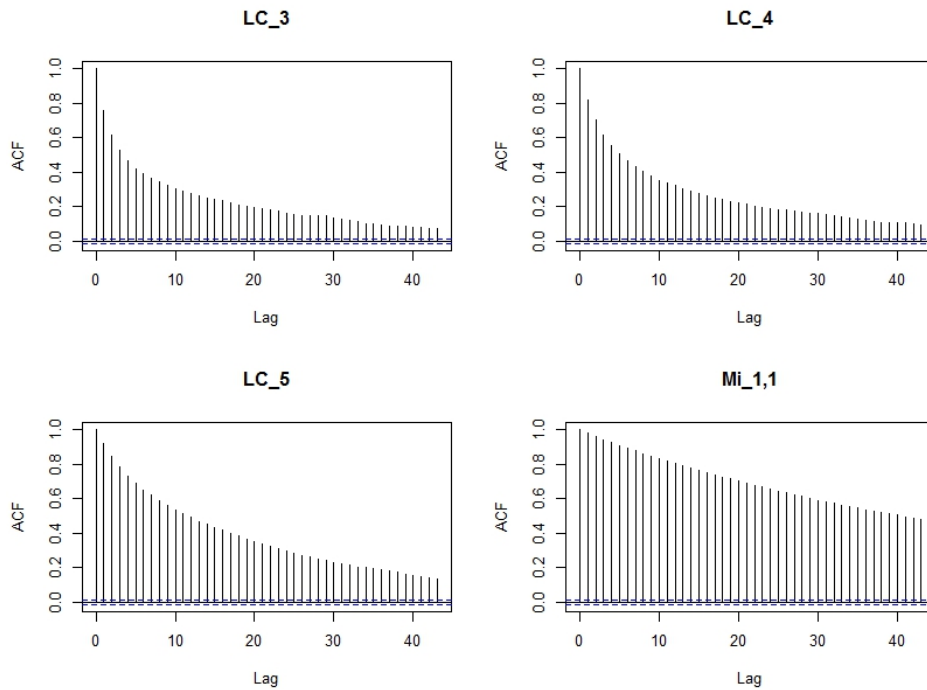
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



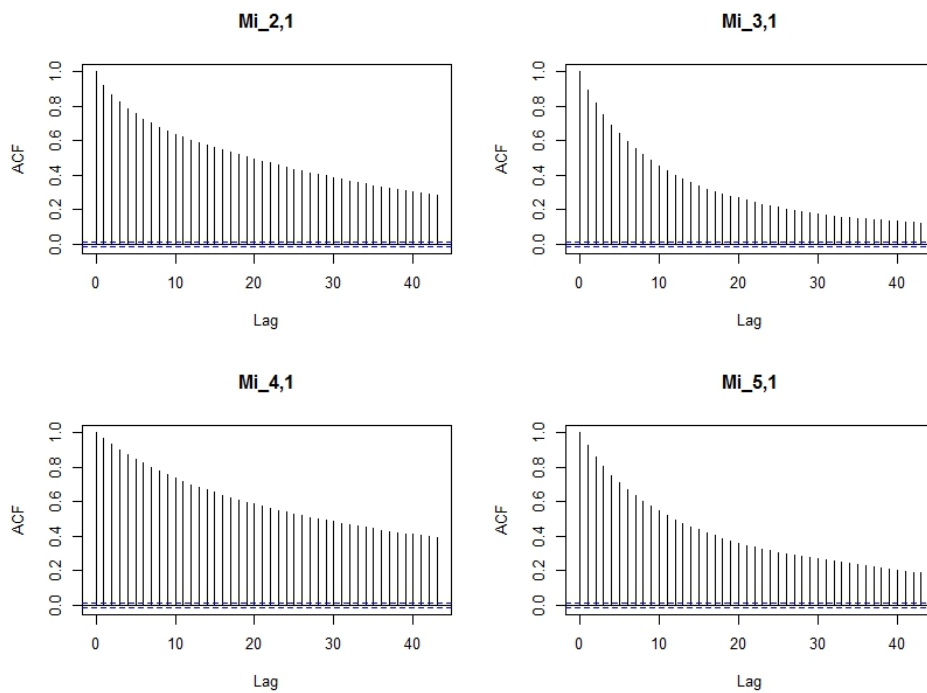
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 55 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



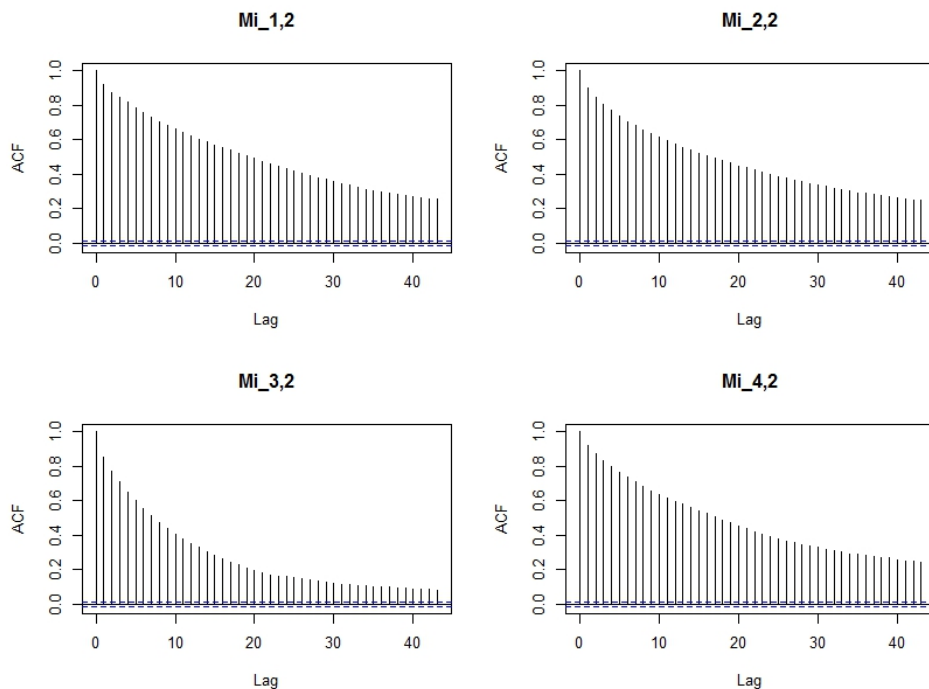
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



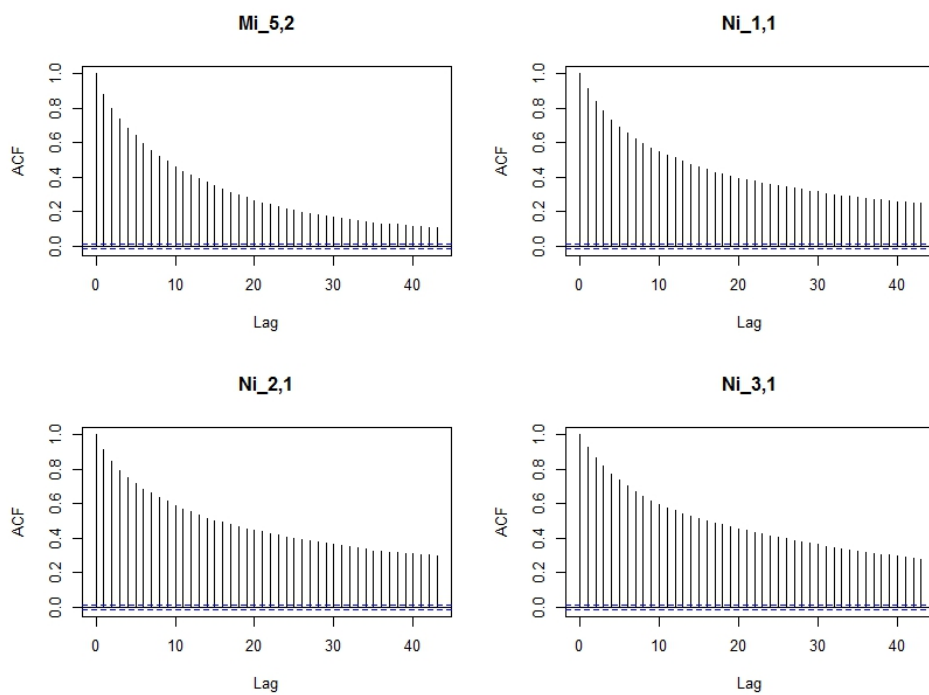
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 56 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



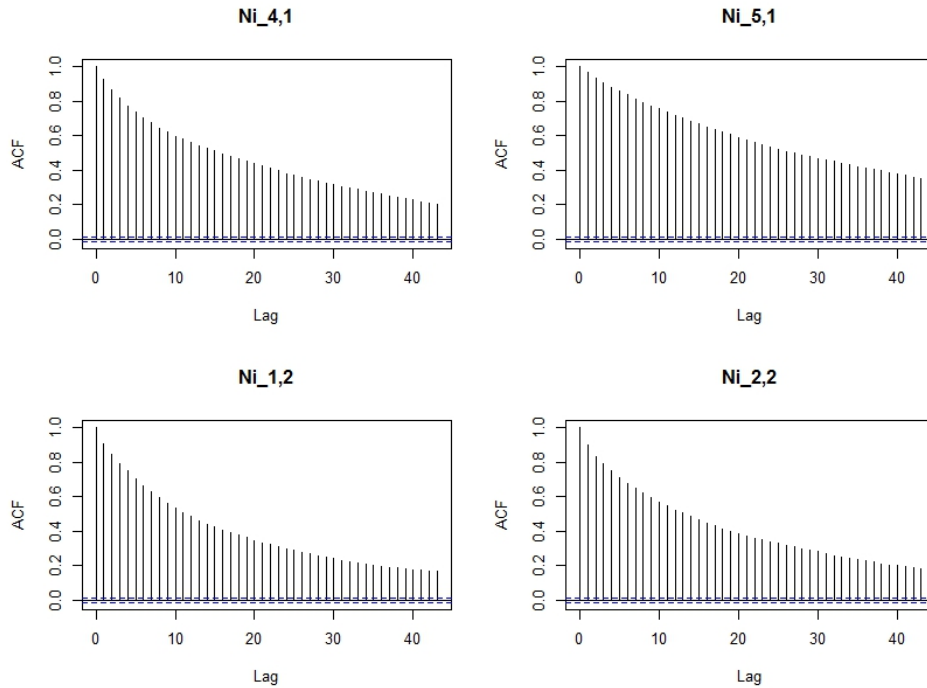
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



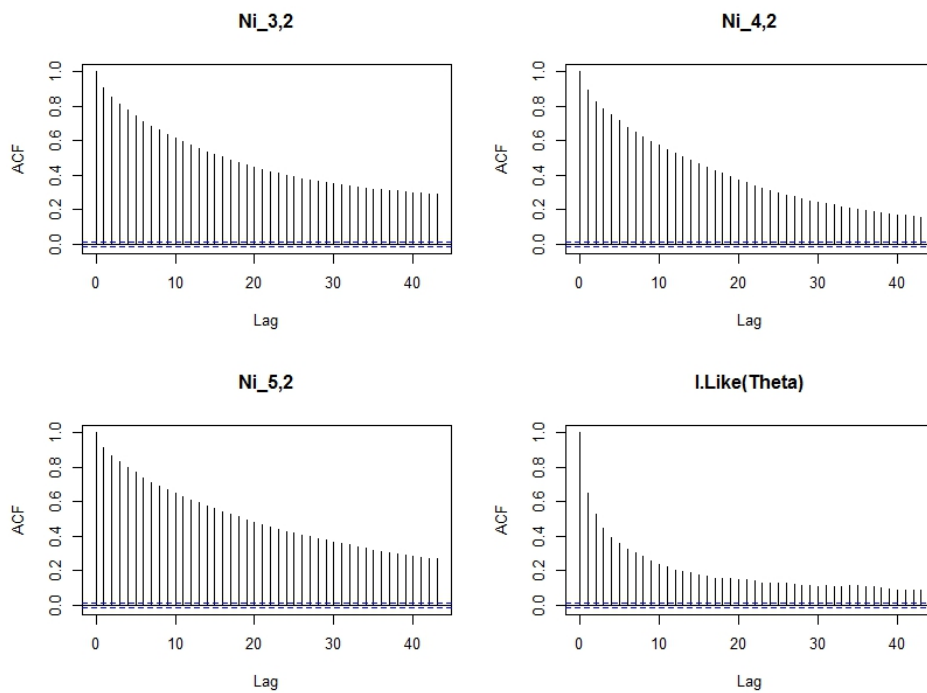
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 57 – Auto-correlação dos resultados do MH aplicados aos dados do artigo comparativo para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



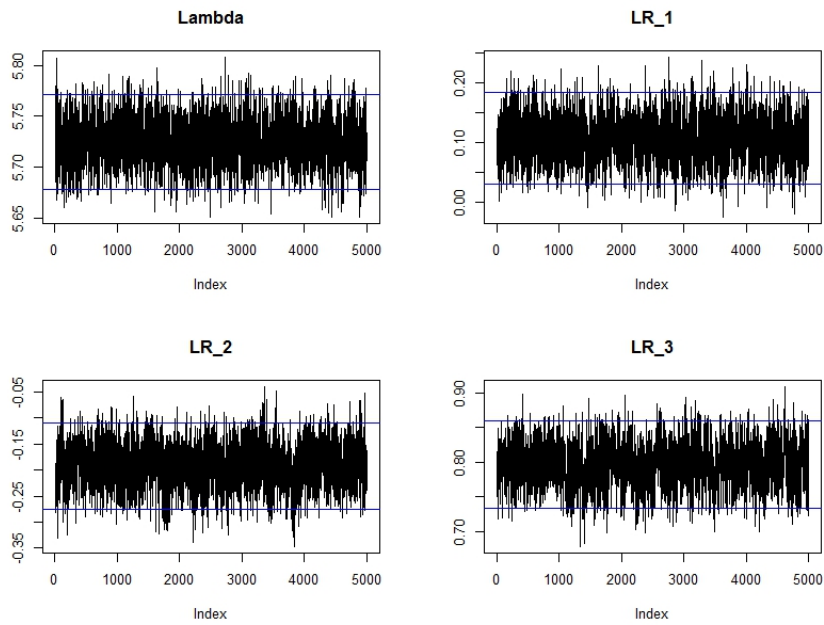
(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

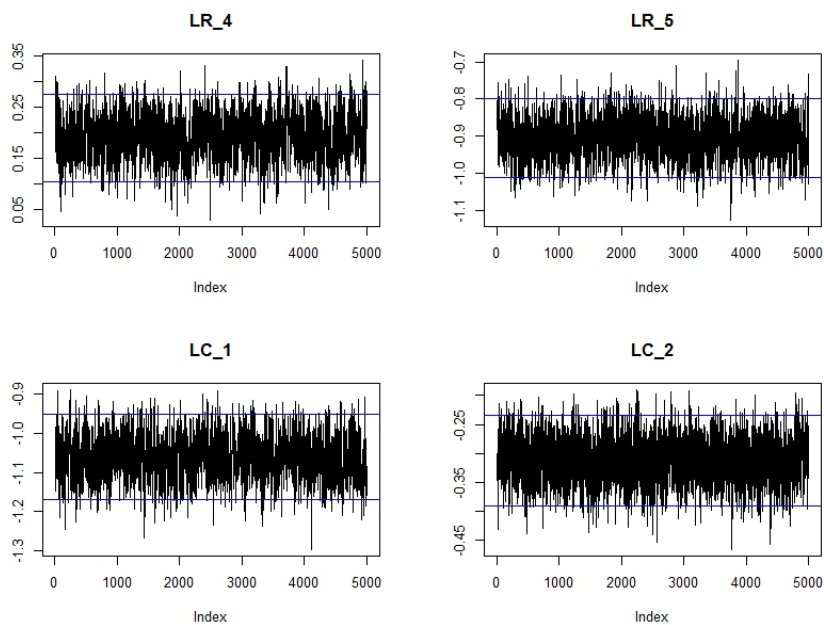
## **D.9 Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de linguagens**

Trace plots para a simulação MCMC na aplicação dos resultados da prova de linguagens e suas tecnologias.

Figura 58 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



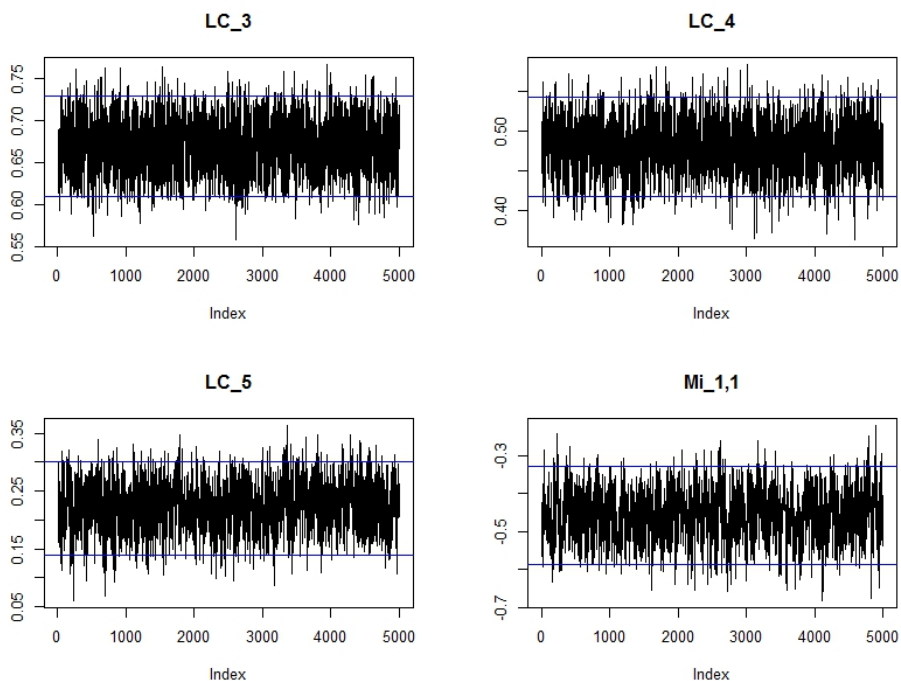
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



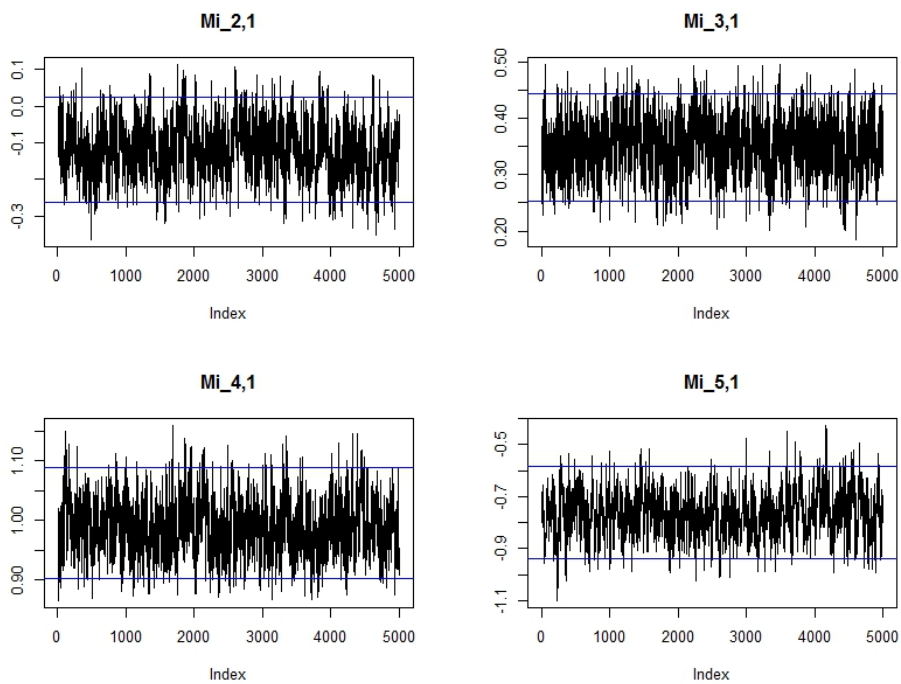
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 59 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



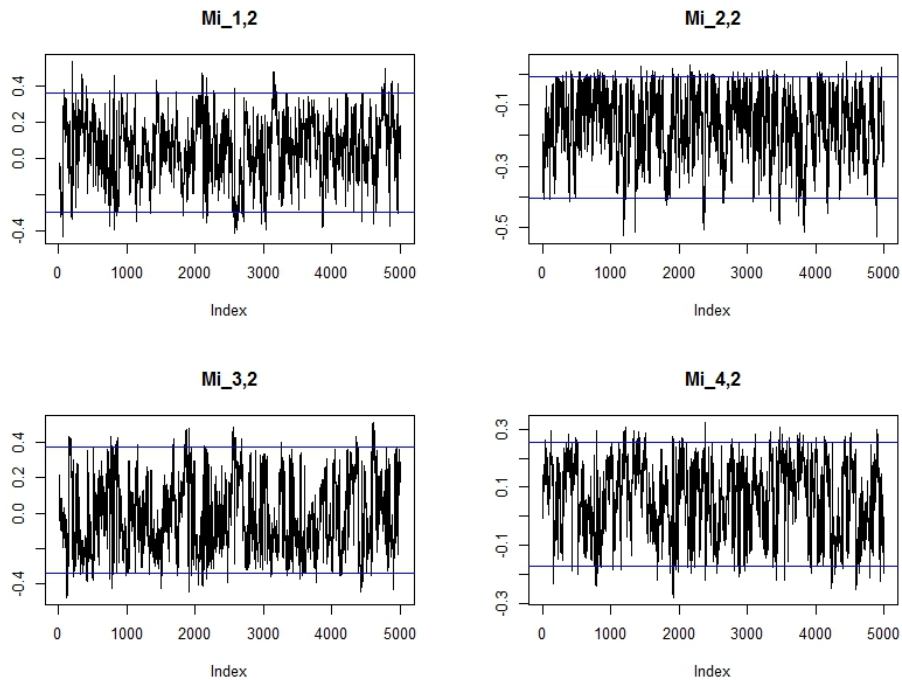
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



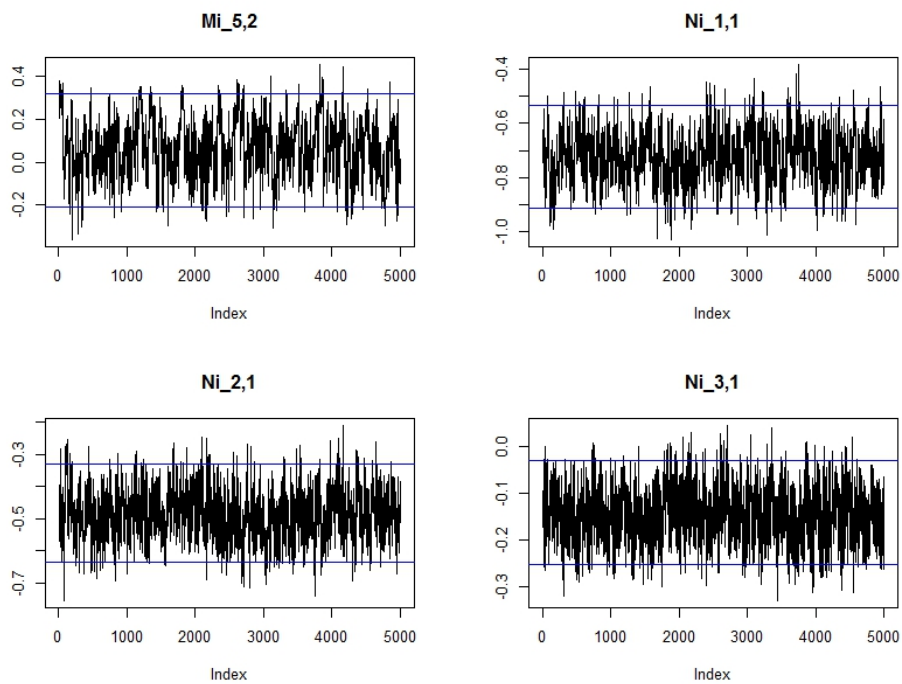
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 60 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



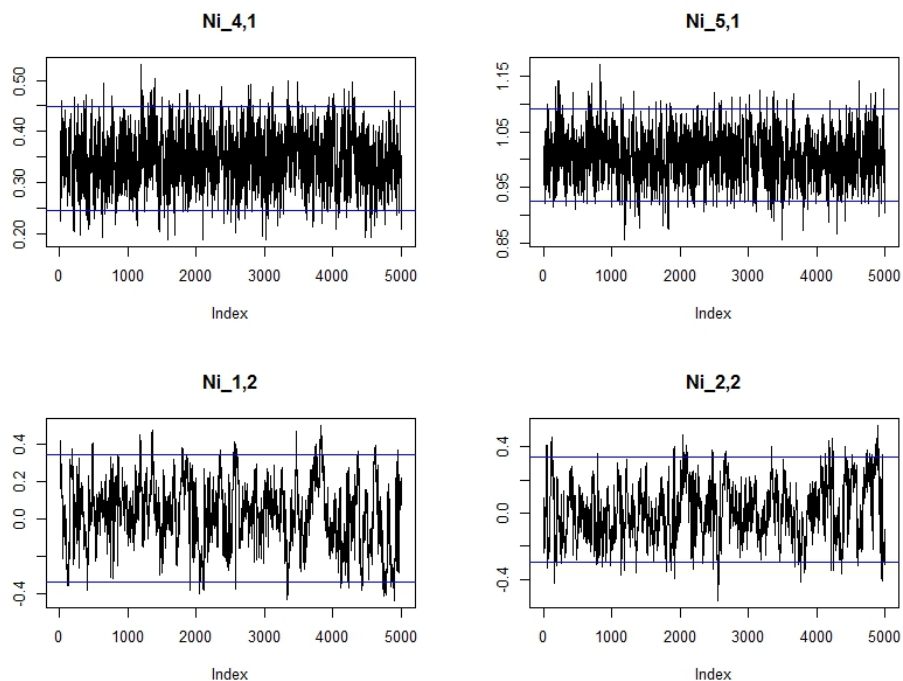
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



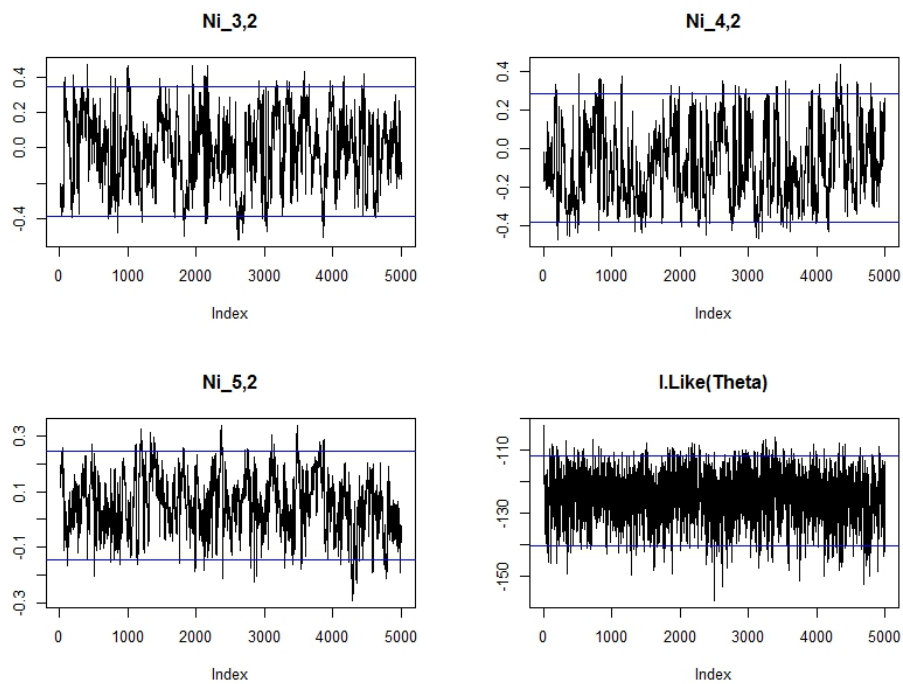
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 61 – Trace plot dos resultados do MH aplicados no desempenho na prova de linguagens para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

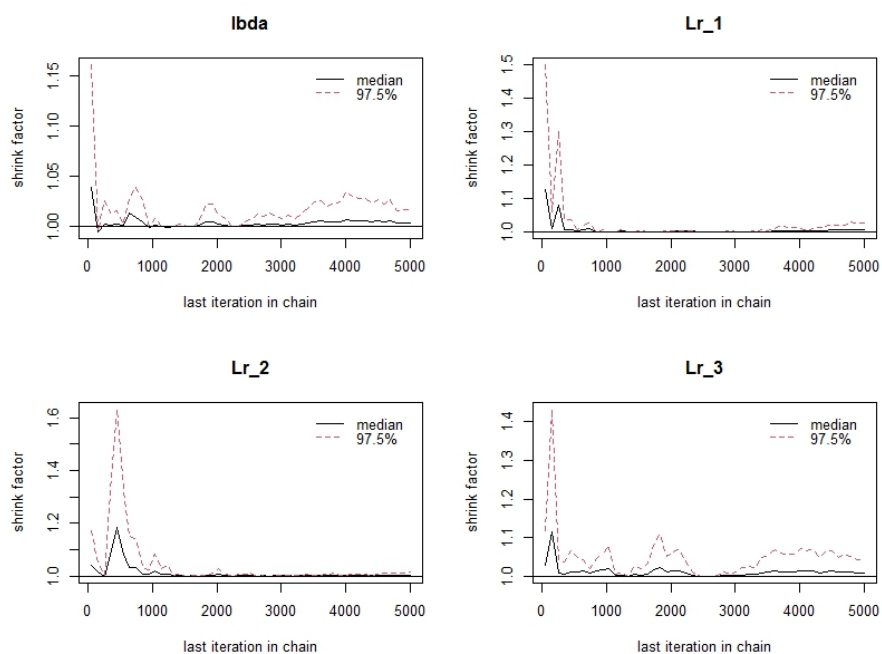
## D.10 Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de linguagens

Tabela 20 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens

	Point est.	Upper c.I.		Point est.	Upper c.I.		Point est.	Upper c.I.
$\lambda$	1.00	1.02	$\mu_{11}$	1.02	1.10	$v_{11}$	1.01	1.03
$\lambda_1^R$	1.01	1.03	$\mu_{21}$	1.00	1.00	$v_{12}$	1.01	1.03
$\lambda_2^R$	1.00	1.01	$\mu_{31}$	1.00	1.02	$v_{13}$	1.00	1.01
$\lambda_3^R$	1.01	1.05	$\mu_{41}$	1.00	1.00	$v_{14}$	1.00	1.02
$\lambda_4^R$	1.00	1.00	$\mu_{51}$	1.01	1.03	$v_{15}$	1.01	1.03
$\lambda_5^R$	1.00	1.02	$\mu_{12}$	1.02	1.10	$v_{21}$	1.01	1.03
$\lambda_1^C$	1.00	1.01	$\mu_{22}$	1.00	1.00	$v_{22}$	1.00	1.00
$\lambda_2^C$	1.00	1.00	$\mu_{32}$	1.00	1.00	$v_{23}$	1.02	1.07
$\lambda_3^C$	1.01	1.03	$\mu_{42}$	1.00	1.00	$v_{24}$	1.01	1.02
$\lambda_4^C$	1.00	1.01	$\mu_{52}$	1.03	1.11	$v_{25}$	1.01	1.01
$\lambda_5^C$	1.00	1.01	—	—	—	$l.Log(\theta)$	1.00	1.01

Fonte: Elaborada pelo autor.

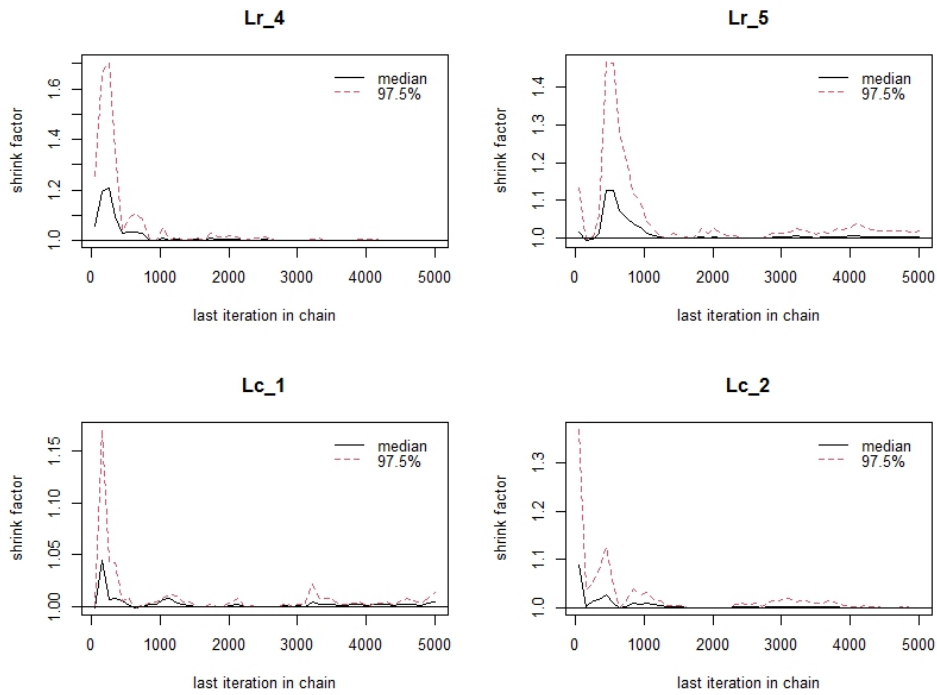
Figura 62 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



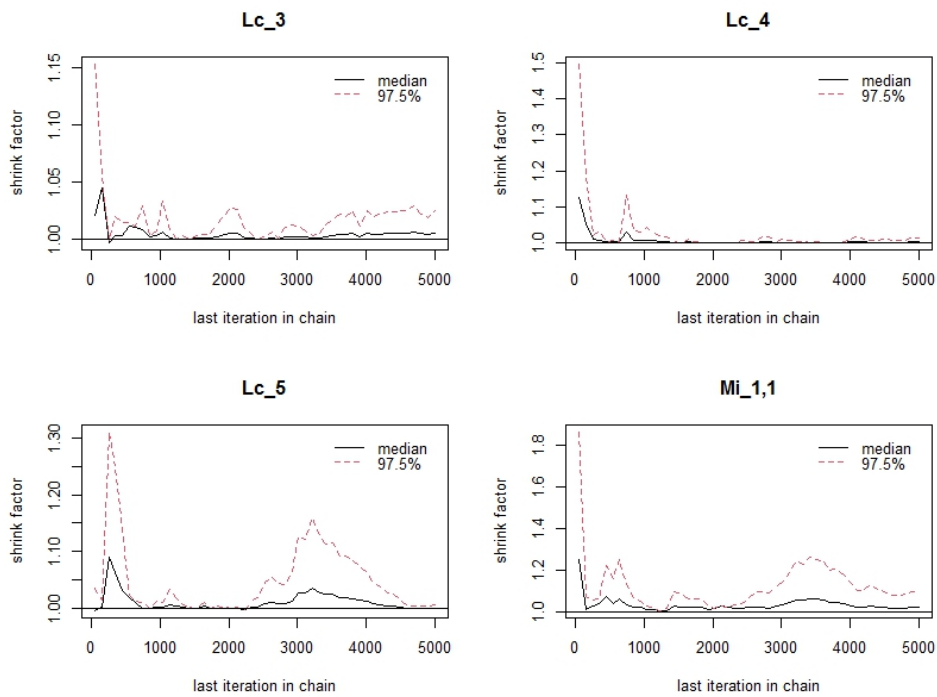
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 63 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda_4^R, \lambda_5^R, \lambda_1^C, \lambda_2^C, \lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$



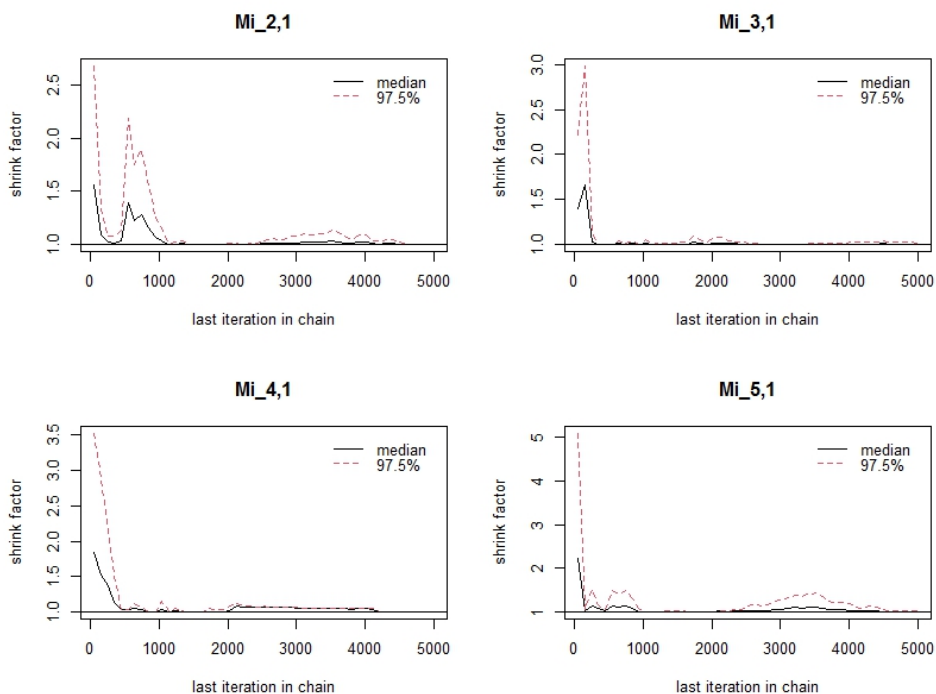
(a)  $\lambda_4^R, \lambda_5^R, \lambda_1^C$  e  $\lambda_2^C$



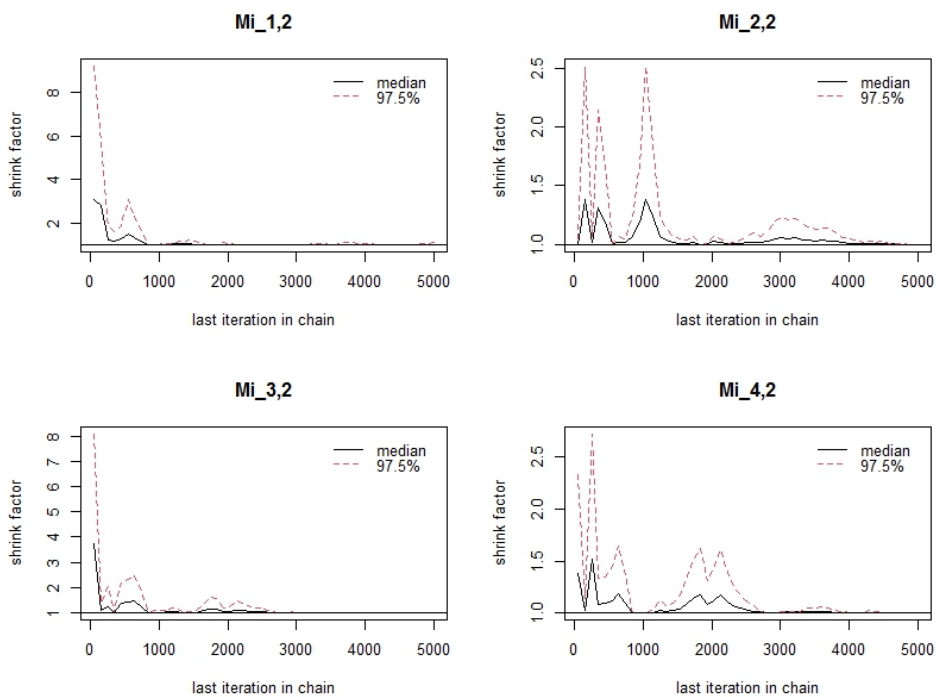
(b)  $\lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$

Fonte: Elaborada pelo autor.

Figura 64 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para  $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$



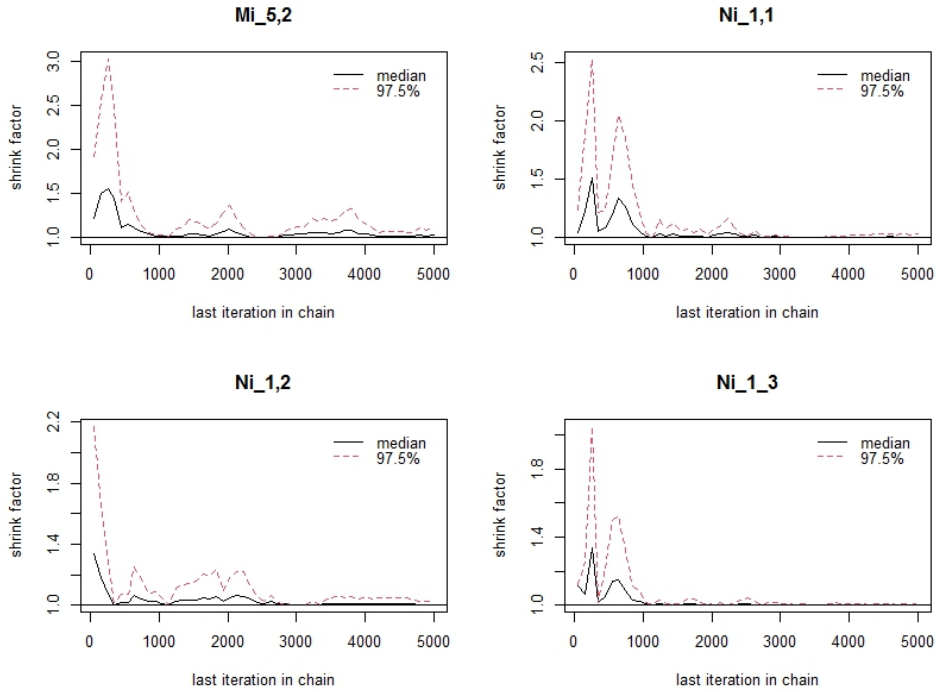
(a)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



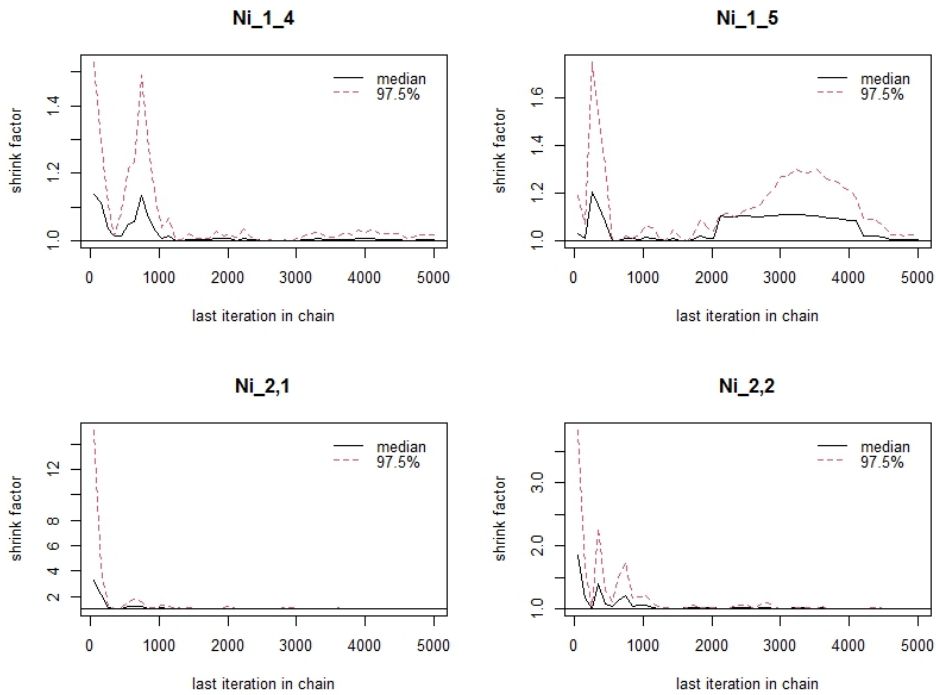
(b)  $\mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$

Fonte: Elaborada pelo autor.

Figura 65 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para  $\mu_{52}, v_{11}, v_{12}, v_{13}, v_{14}, v_{15}, v_{21}$  e  $v_{22}$



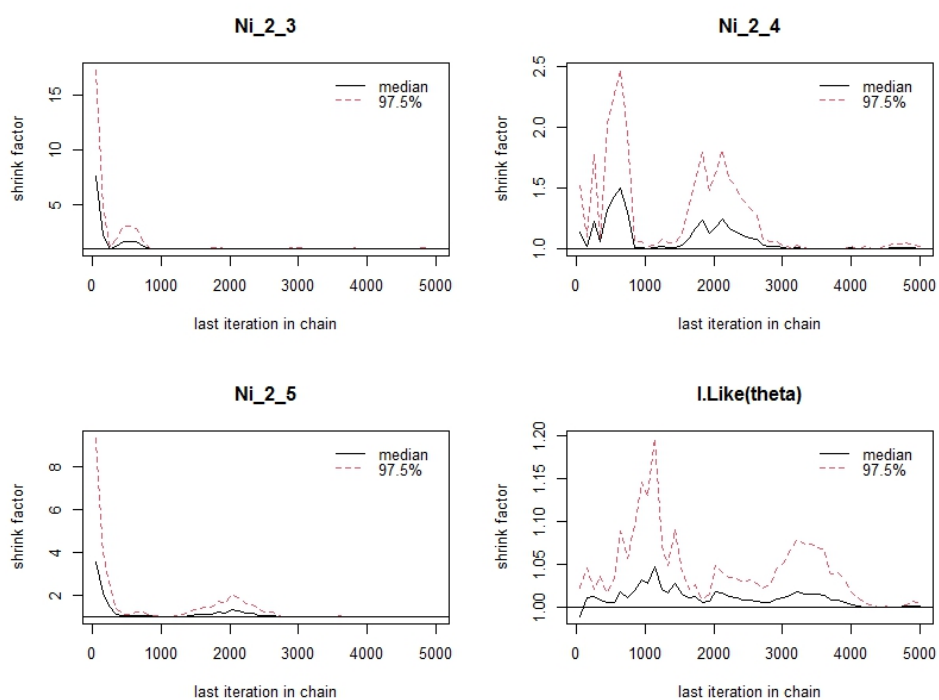
(a)  $\mu_{52}, v_{11}, v_{12}$  e  $v_{13}$



(b)  $v_{14}, v_{15}, v_{21}$  e  $v_{22}$

Fonte: Elaborada pelo autor.

Figura 66 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de linguagens para  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

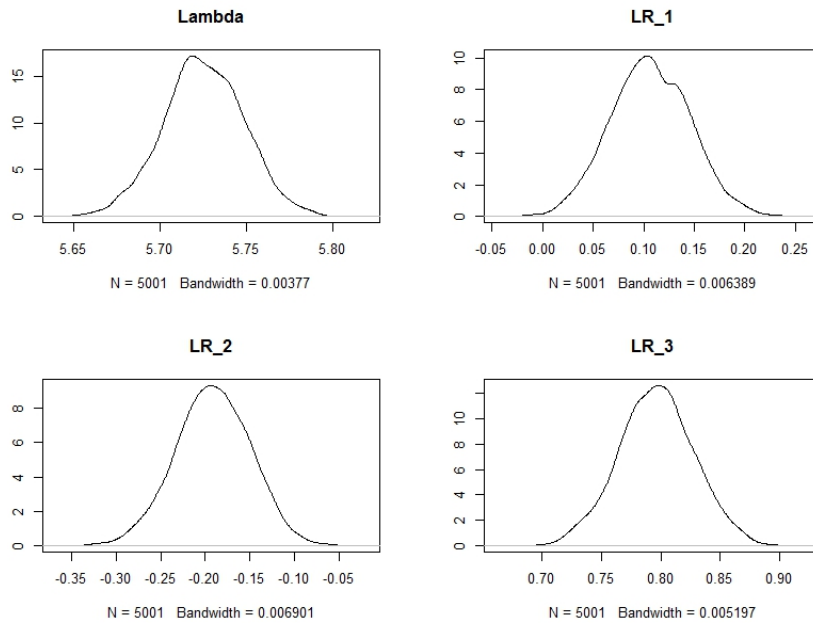


(a)  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

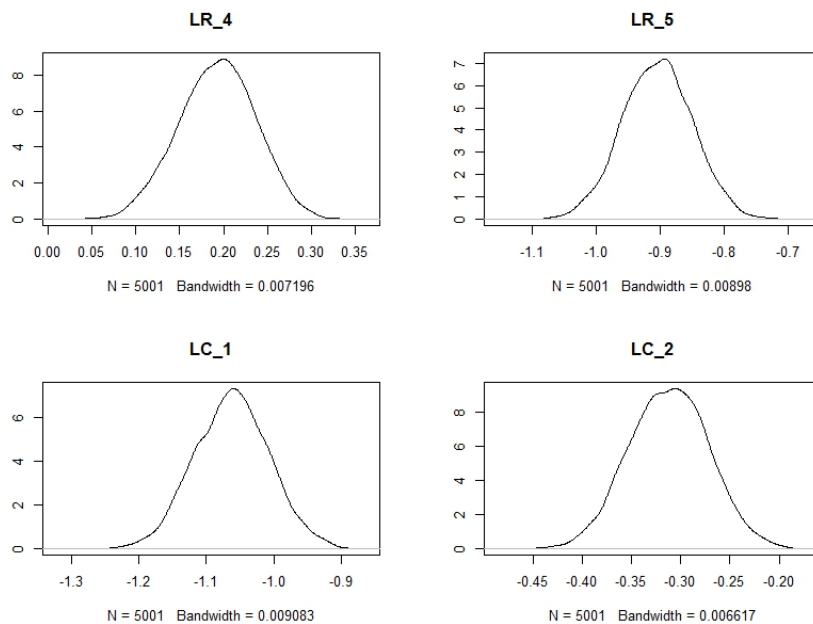
Fonte: Elaborada pelo autor.

## D.11 Gráficos das distribuições *a posteriori* dos resultados do MH aplicado no desempenho na prova de linguagens

Figura 67 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de linguagens para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



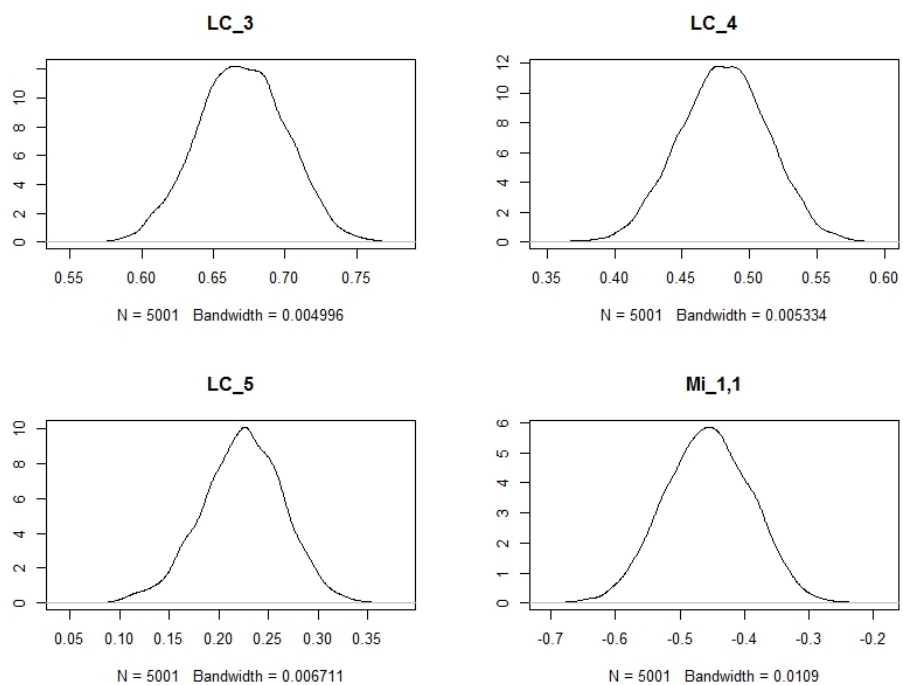
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



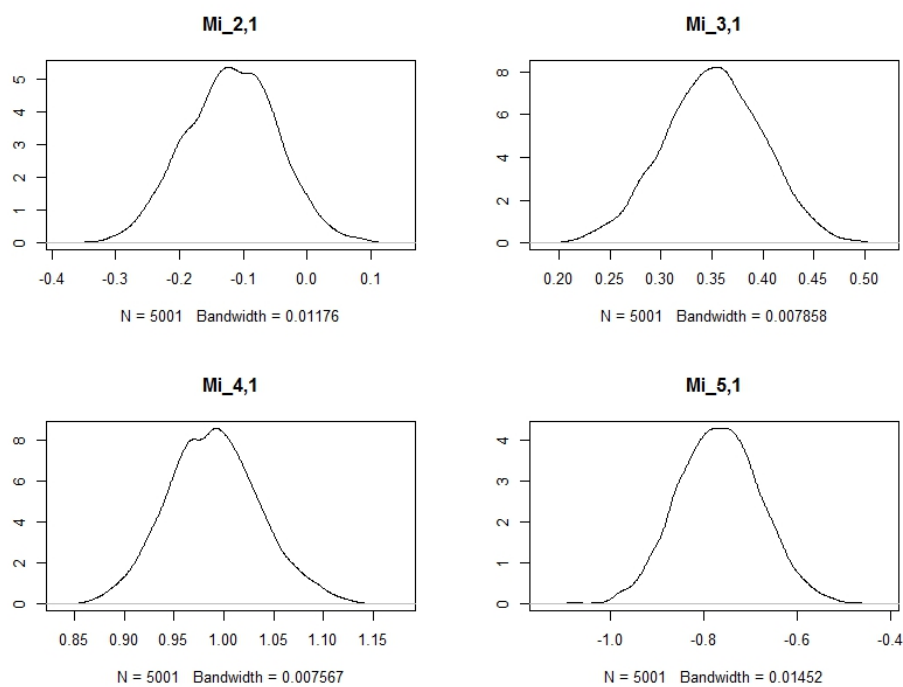
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 68 – Distribuições a posteriori do MH aplicados no desempenho na prova de linguagens para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



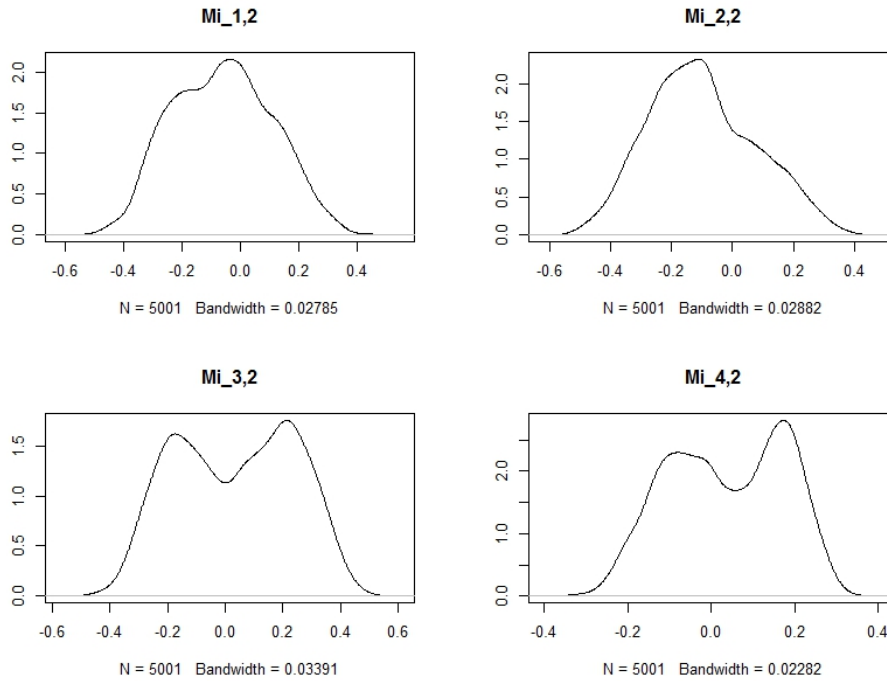
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



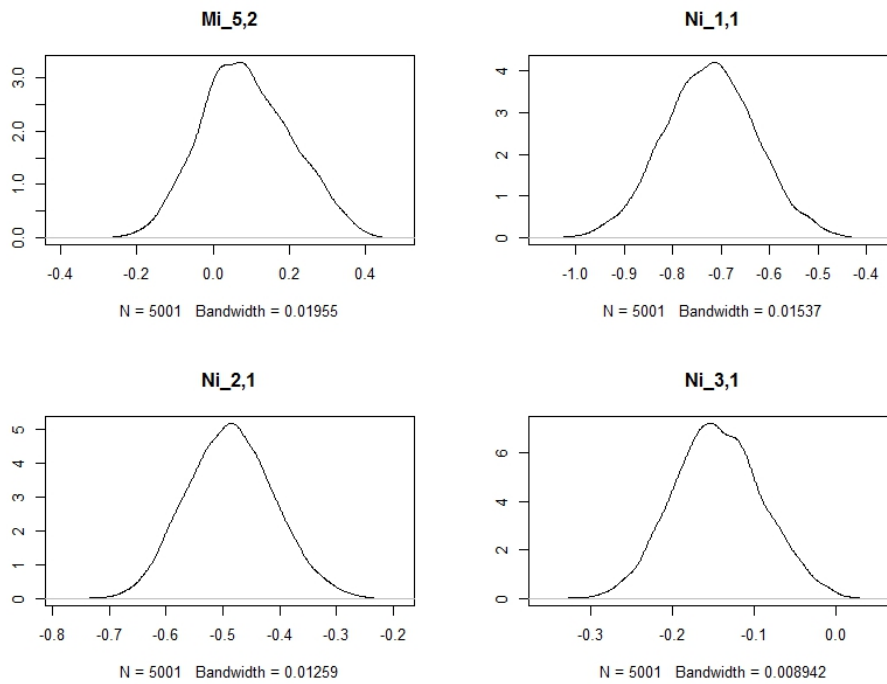
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 69 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de linguagens para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



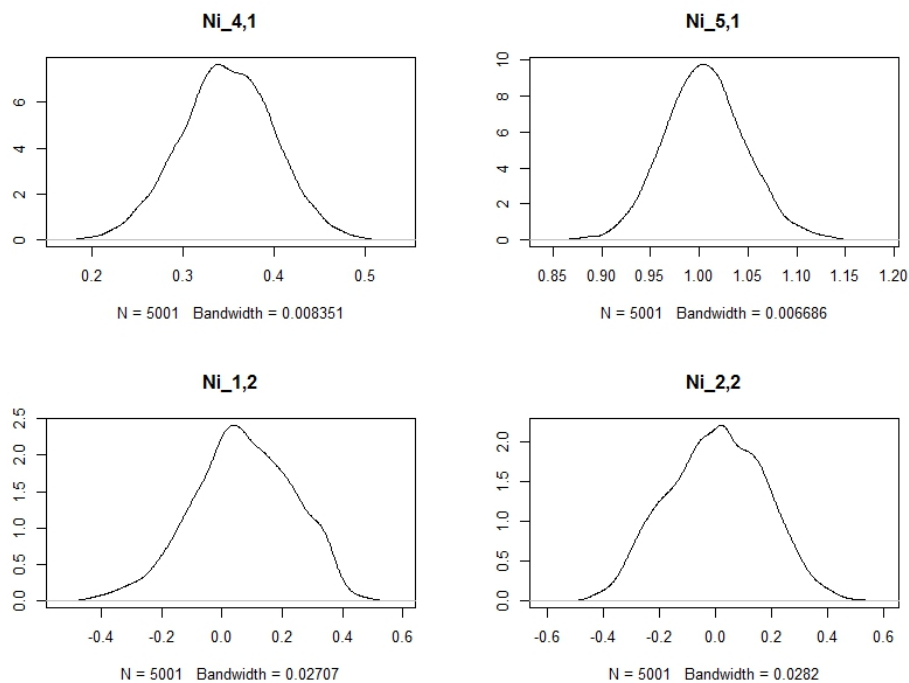
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



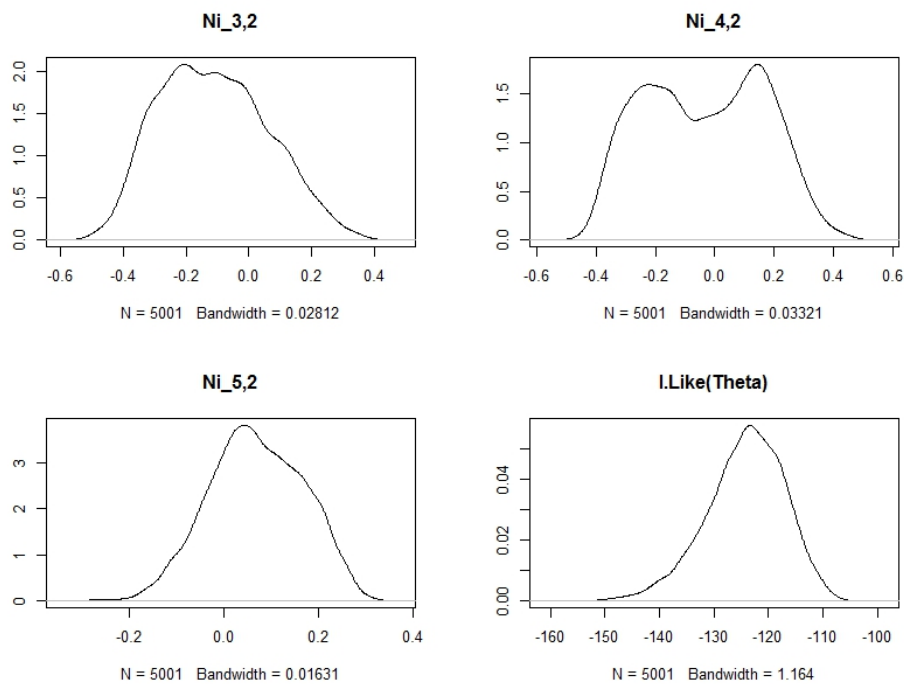
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 70 – Distribuições a posteriori do MH aplicados no desempenho na prova de linguagens para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$

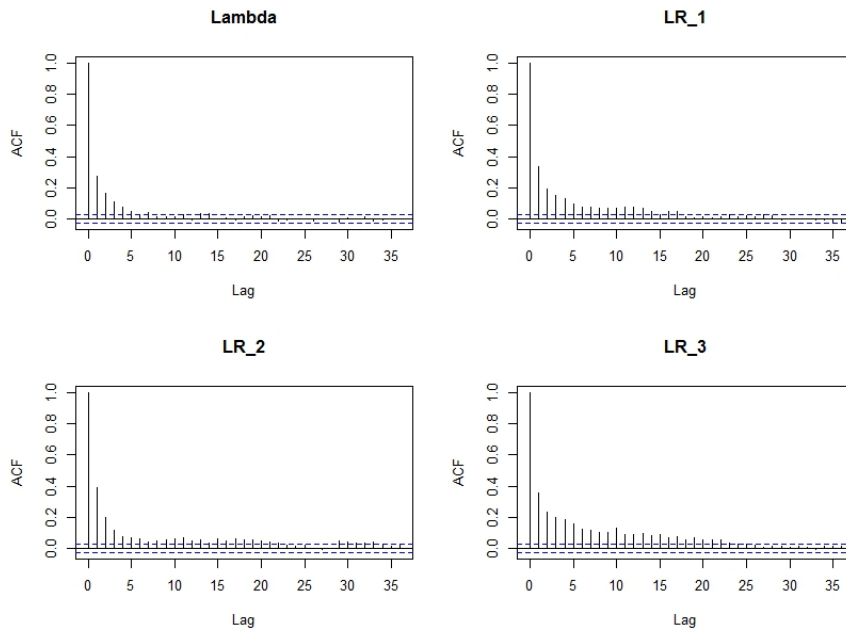


(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

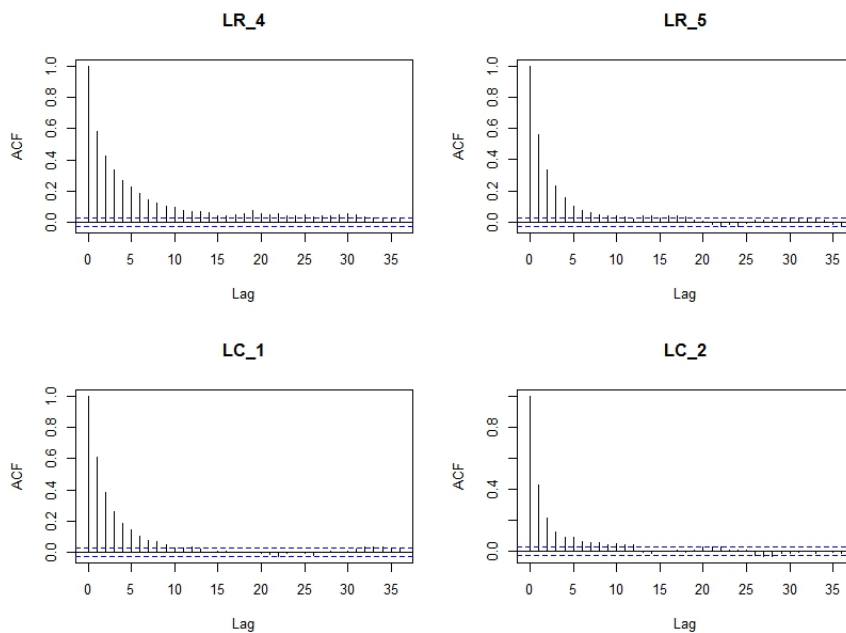
Fonte: Elaborada pelo autor.

## D.12 Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de linguagens

Figura 71 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



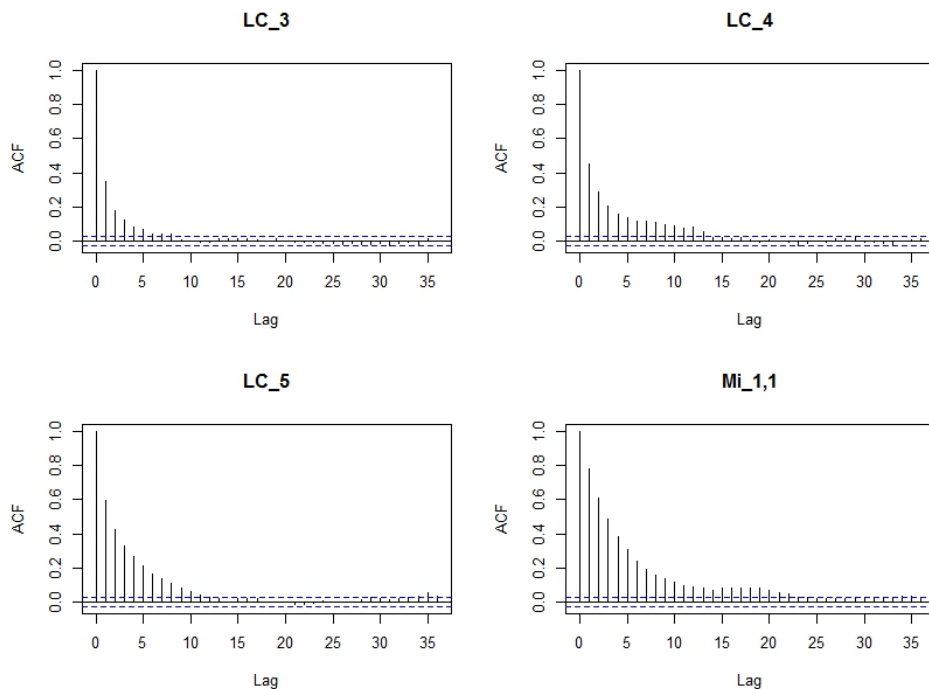
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



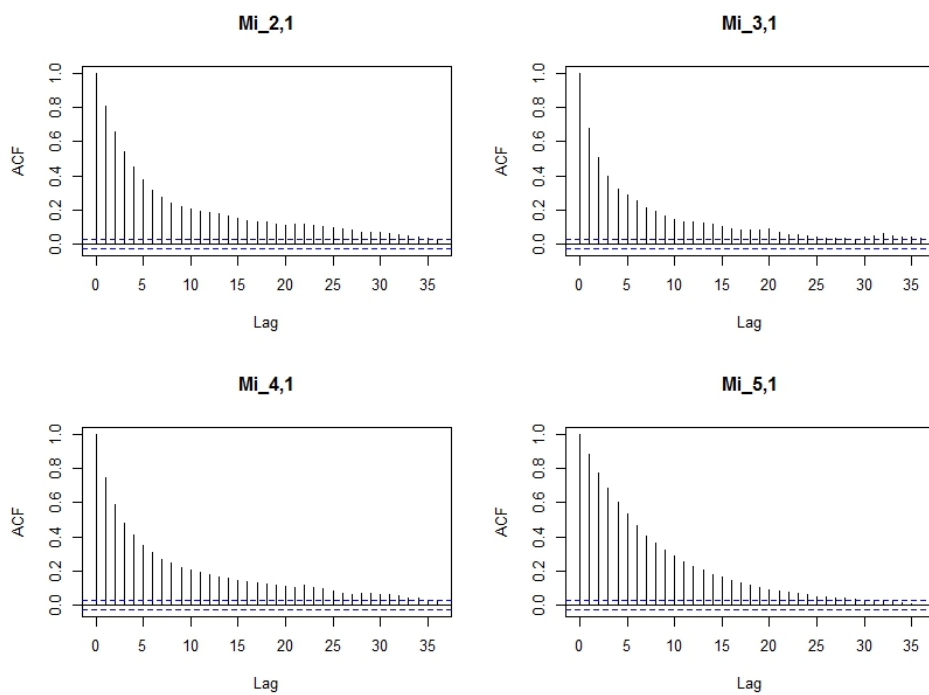
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 72 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para  $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



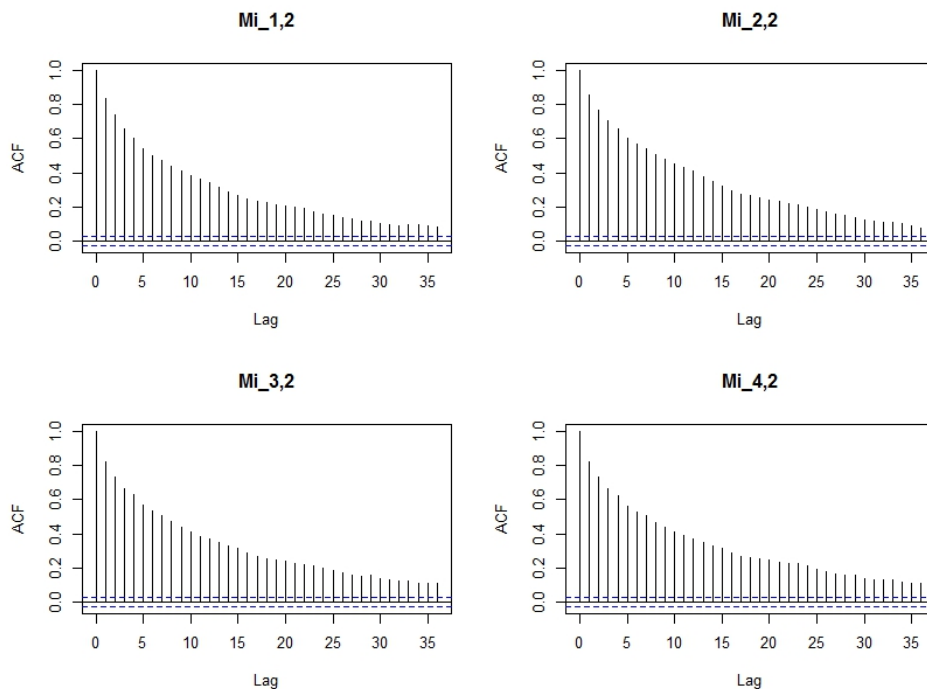
(a)  $\lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$



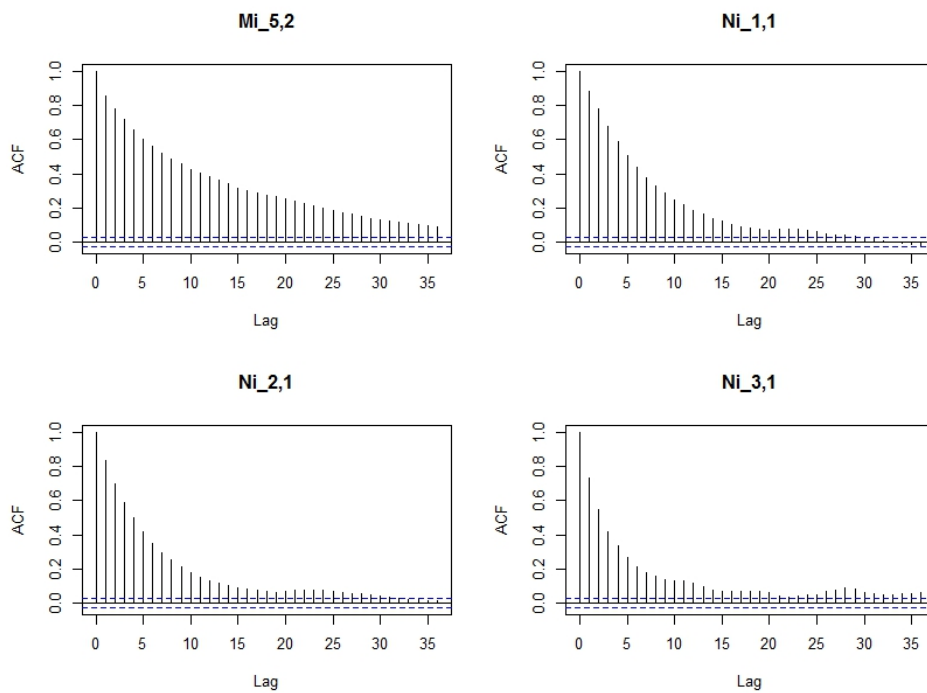
(b)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 73 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



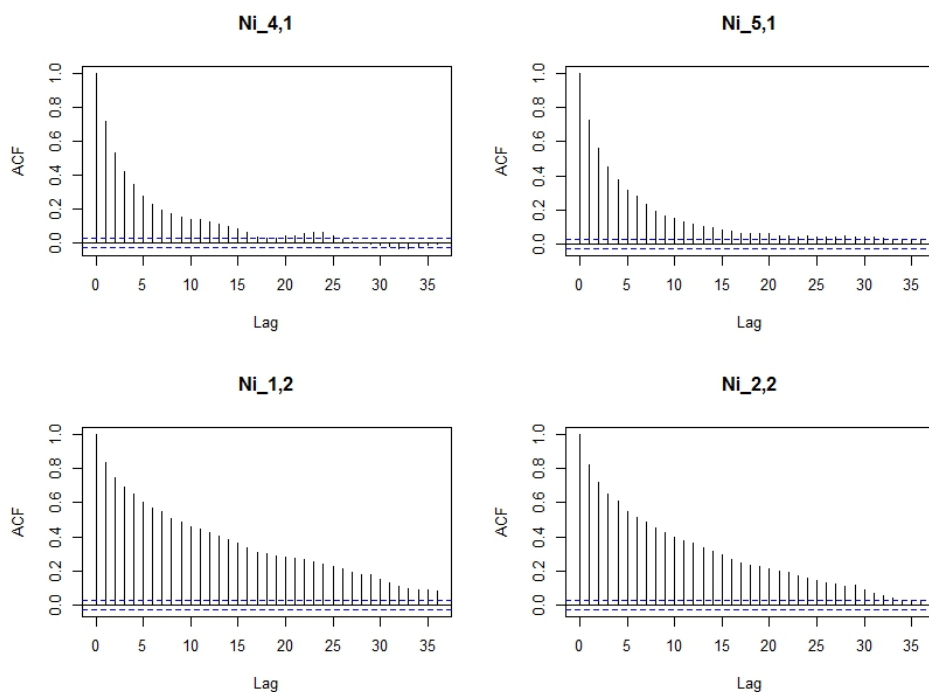
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



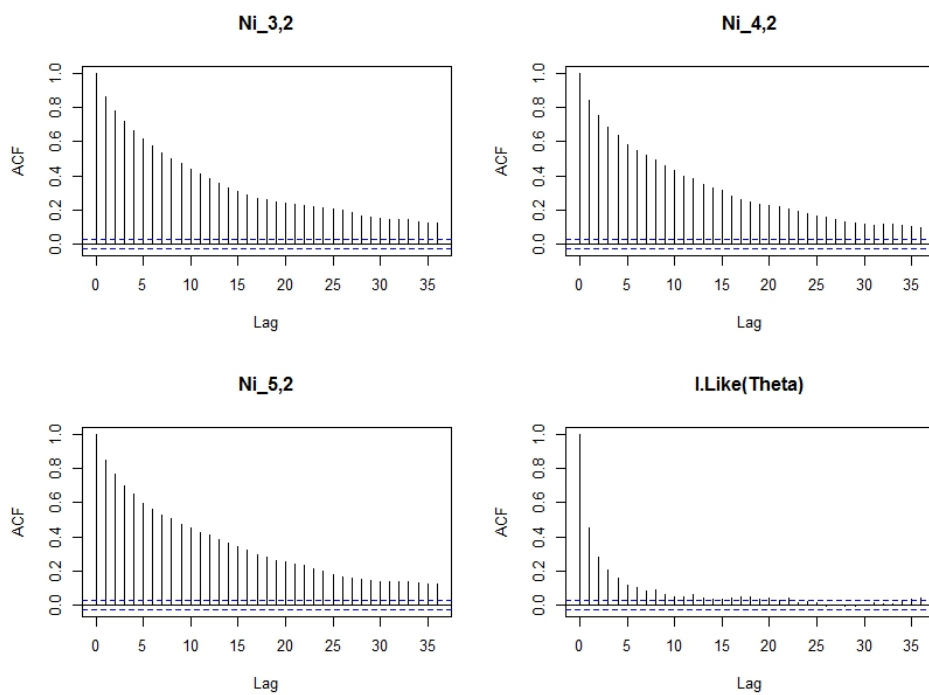
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 74 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de linguagens para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



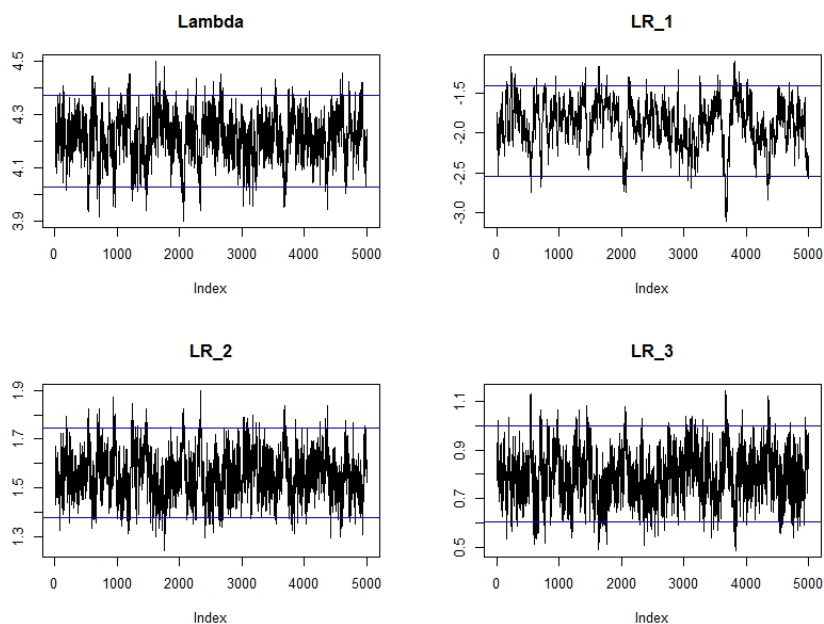
(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

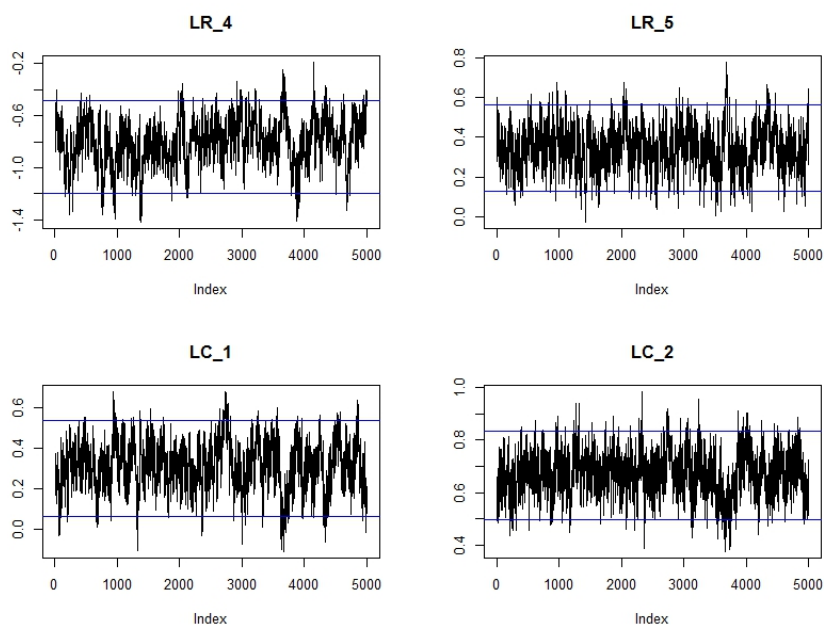
## **D.13 Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de matemática**

Trace plots para a simulação MCMC na aplicação dos dados da prova de matemática e suas tecnologias.

Figura 75 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



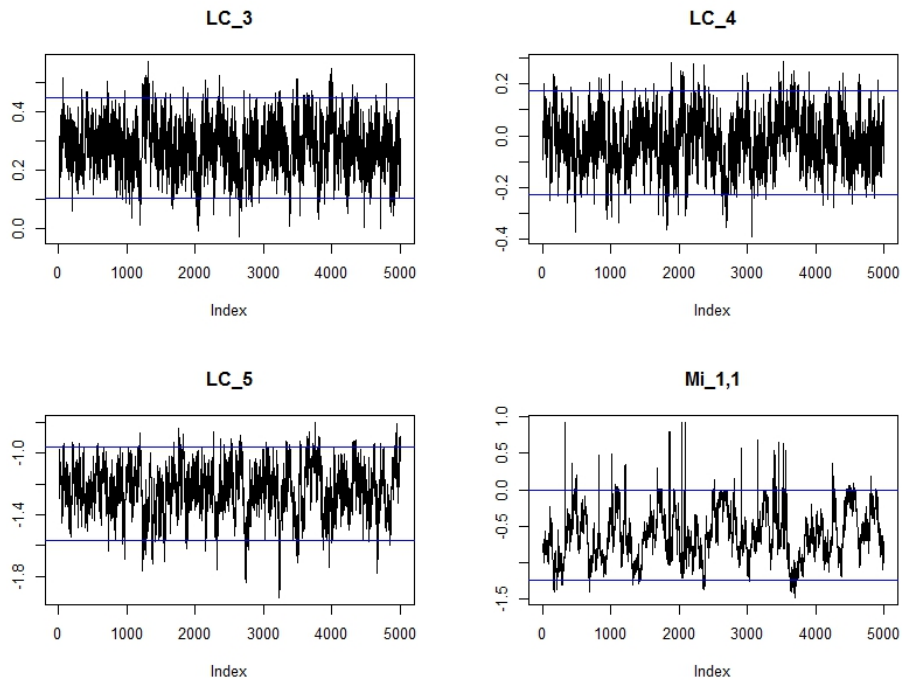
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



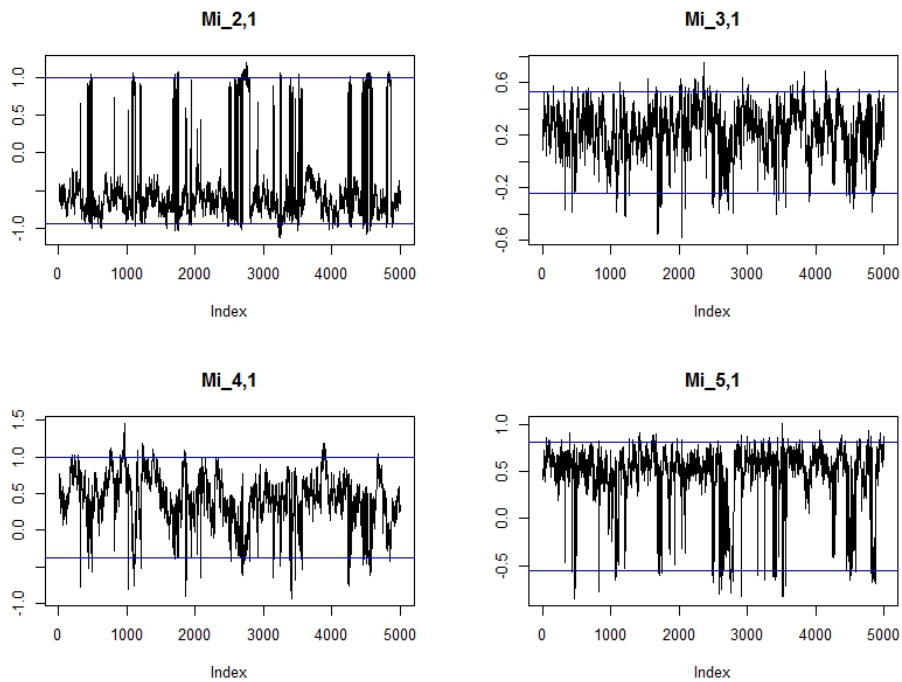
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 76 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



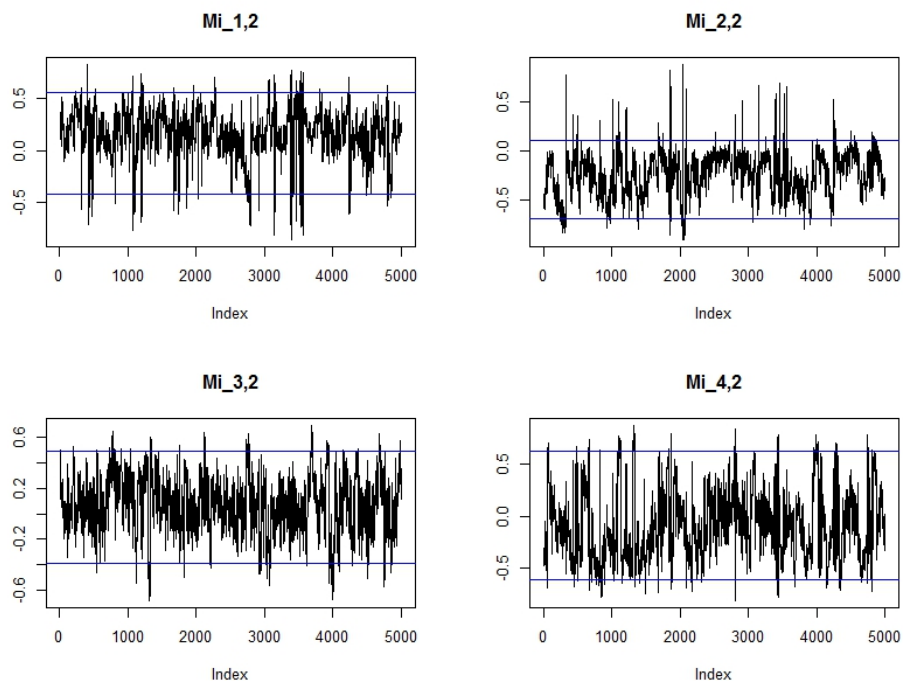
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



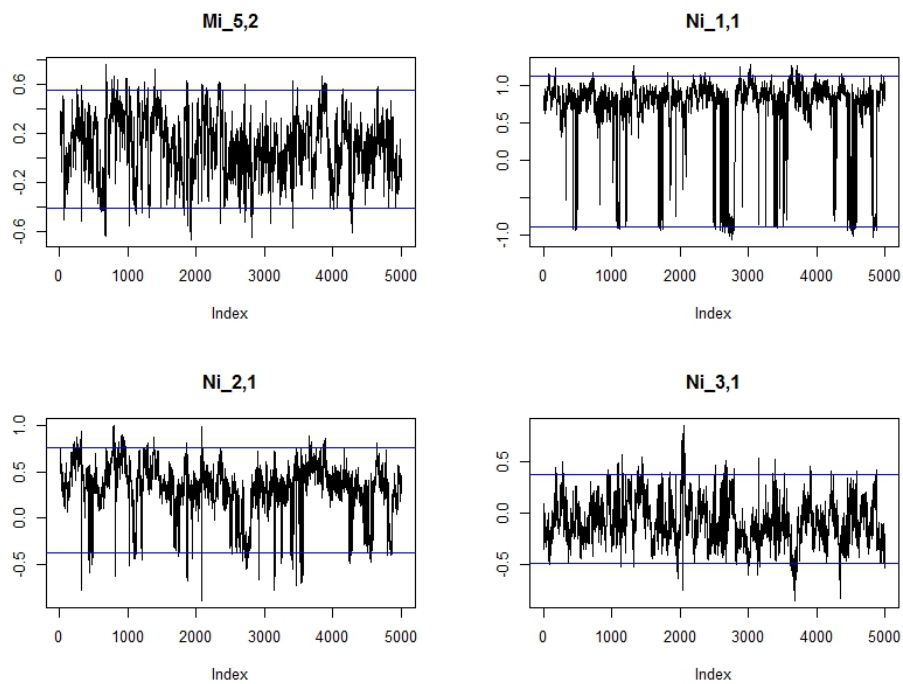
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 77 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



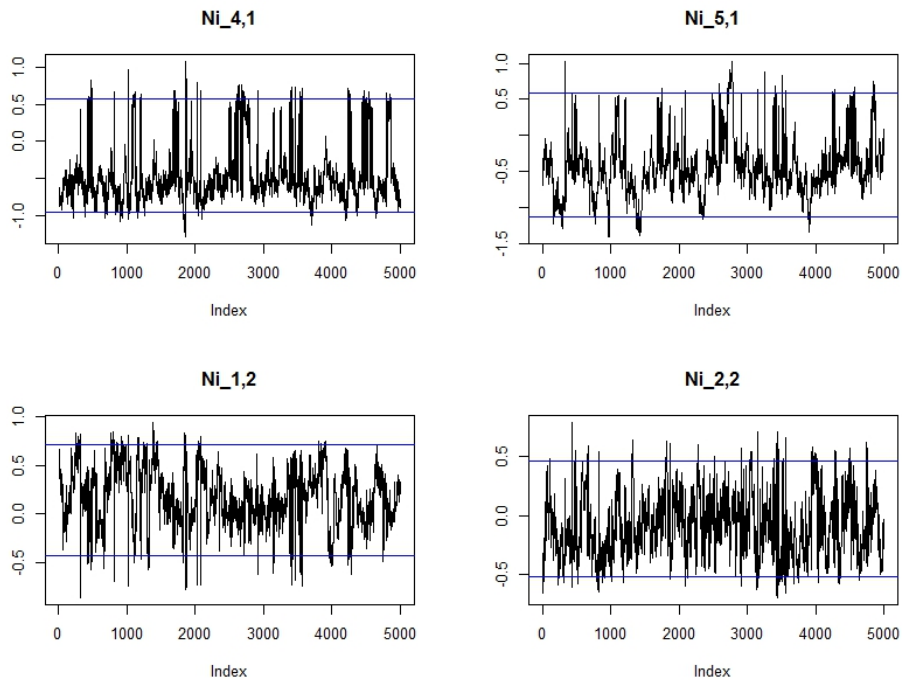
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



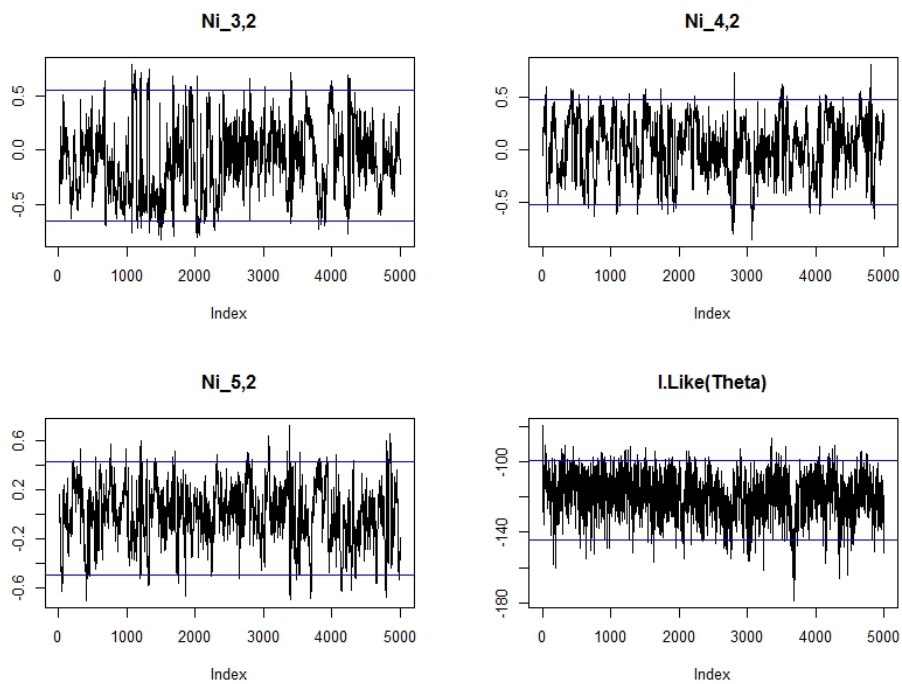
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 78 – Trace plot dos resultados do MH aplicados no desempenho na prova de matemática para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

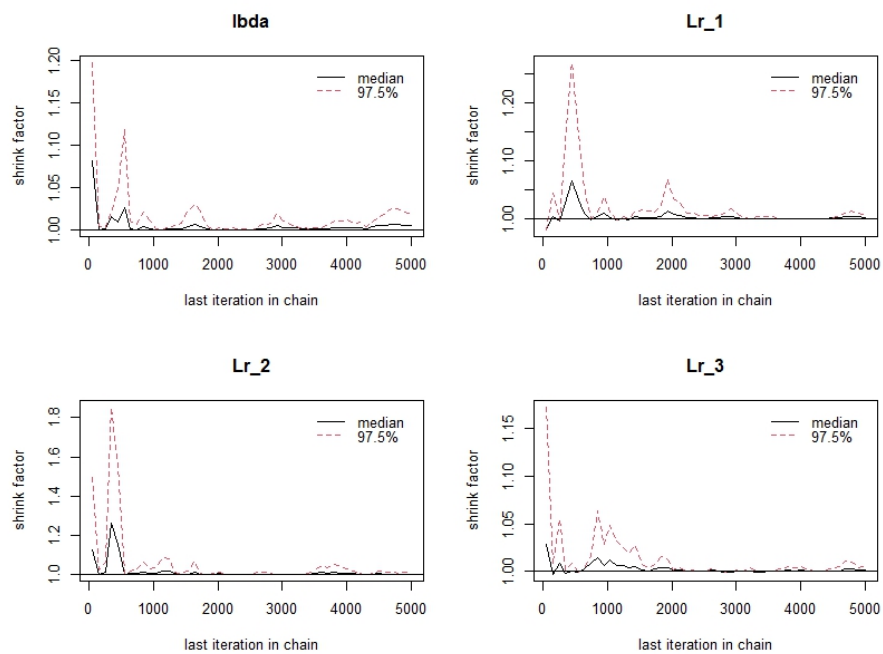
## D.14 Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de matemática

Tabela 21 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática

	Point est.	Upper c.I.		Point est.	Upper c.I.		Point est.	Upper c.I.
$\lambda$	1.00	1.02	$\mu_{11}$	1.01	1.03	$v_{11}$	1.01	1.04
$\lambda_1^R$	1.00	1.01	$\mu_{21}$	1.00	1.00	$v_{12}$	1.00	1.01
$\lambda_2^R$	1.00	1.01	$\mu_{31}$	1.00	1.02	$v_{13}$	1.00	1.01
$\lambda_3^R$	1.00	1.00	$\mu_{41}$	1.00	1.00	$v_{14}$	1.00	1.01
$\lambda_4^R$	1.00	1.01	$\mu_{51}$	1.02	1.10	$v_{15}$	1.01	1.04
$\lambda_5^R$	1.00	1.02	$\mu_{12}$	1.01	1.05	$v_{21}$	1.01	1.06
$\lambda_1^C$	1.00	1.01	$\mu_{22}$	1.00	1.01	$v_{22}$	1.00	1.00
$\lambda_2^C$	1.01	1.05	$\mu_{32}$	1.03	1.15	$v_{23}$	1.02	1.09
$\lambda_3^C$	1.00	1.01	$\mu_{42}$	1.04	1.17	$v_{24}$	1.02	1.10
$\lambda_4^C$	1.00	1.01	$\mu_{52}$	1.05	1.19	$v_{25}$	1.03	1.15
$\lambda_5^C$	1.00	1.02	---	---	—	$l.Log(\theta)$	1.00	1.02

Fonte: Elaborada pelo autor.

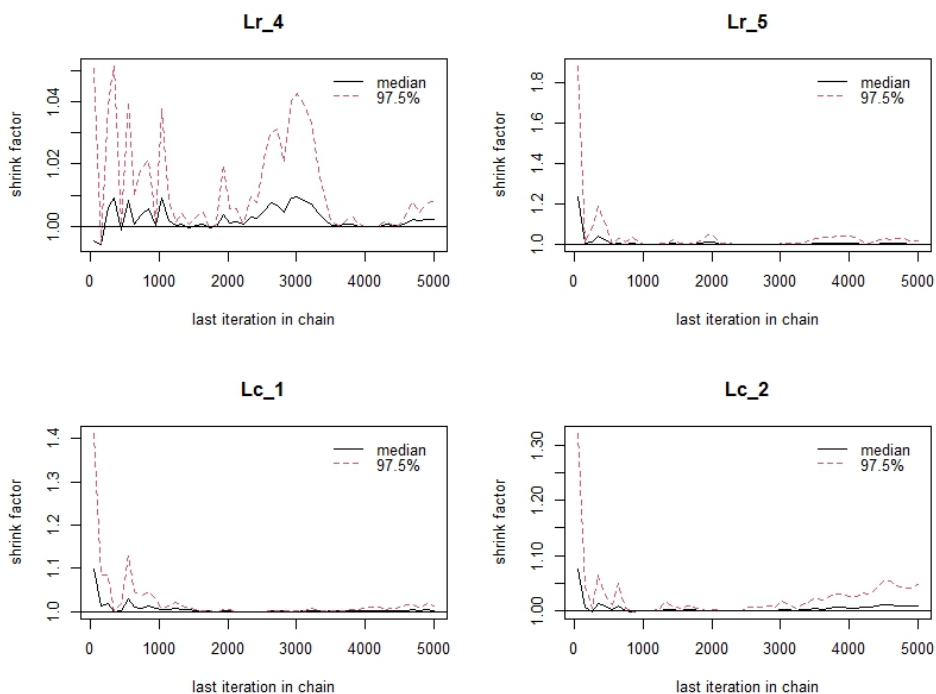
Figura 79 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



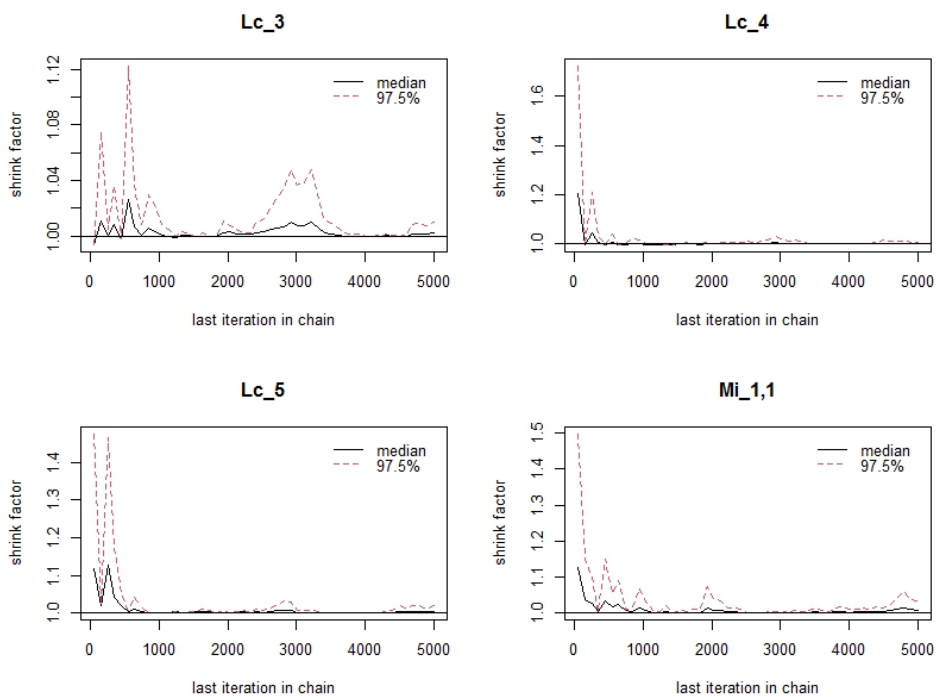
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 80 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$ ,  $\lambda_2^C$ ,  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



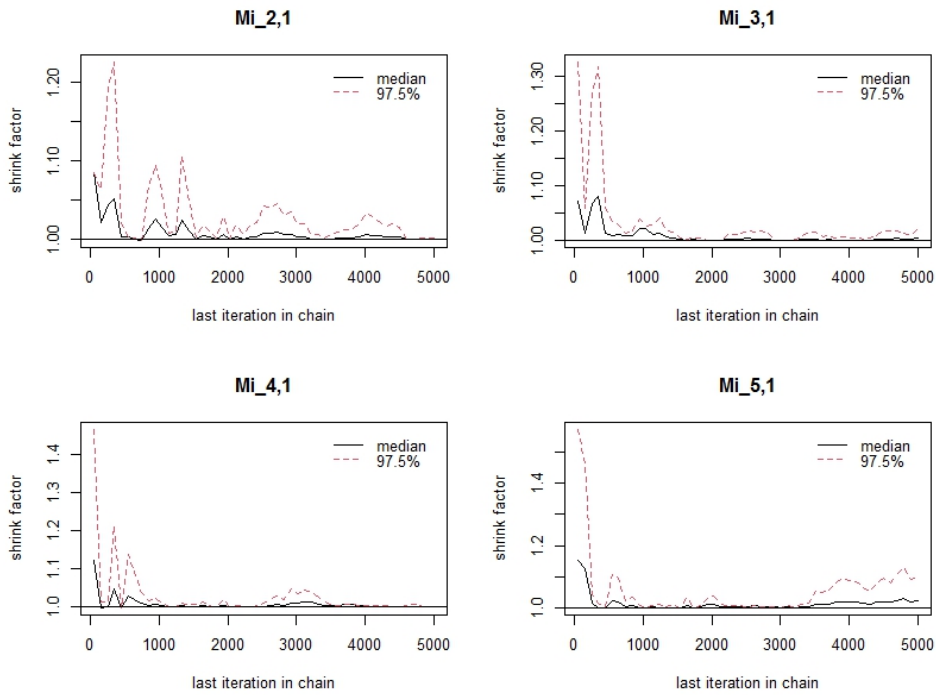
(a)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



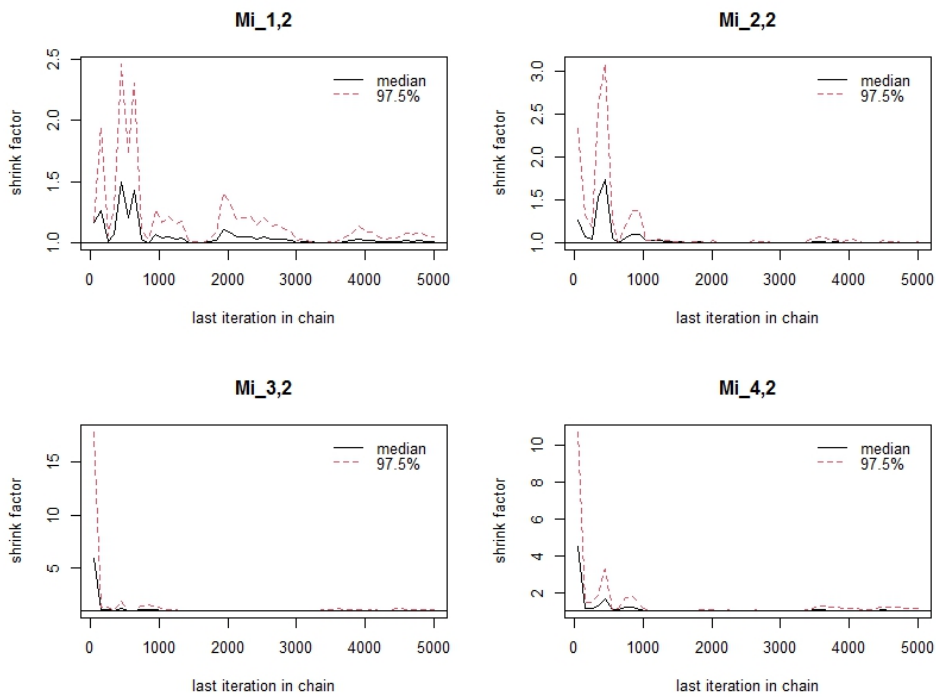
(b)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$

Fonte: Elaborada pelo autor.

Figura 81 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para  $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$



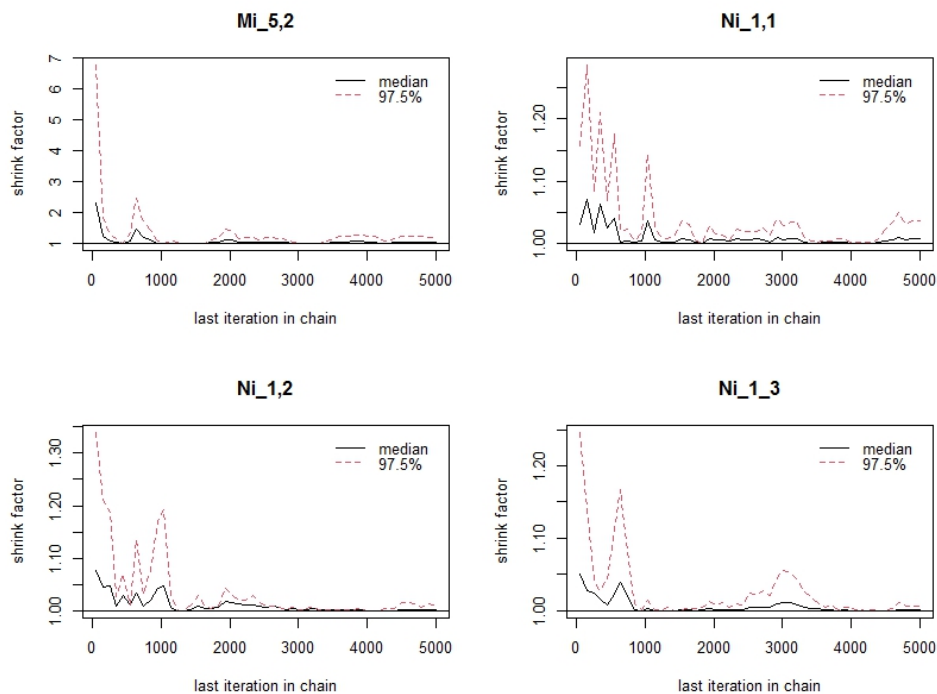
(a)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



(b)  $\mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$

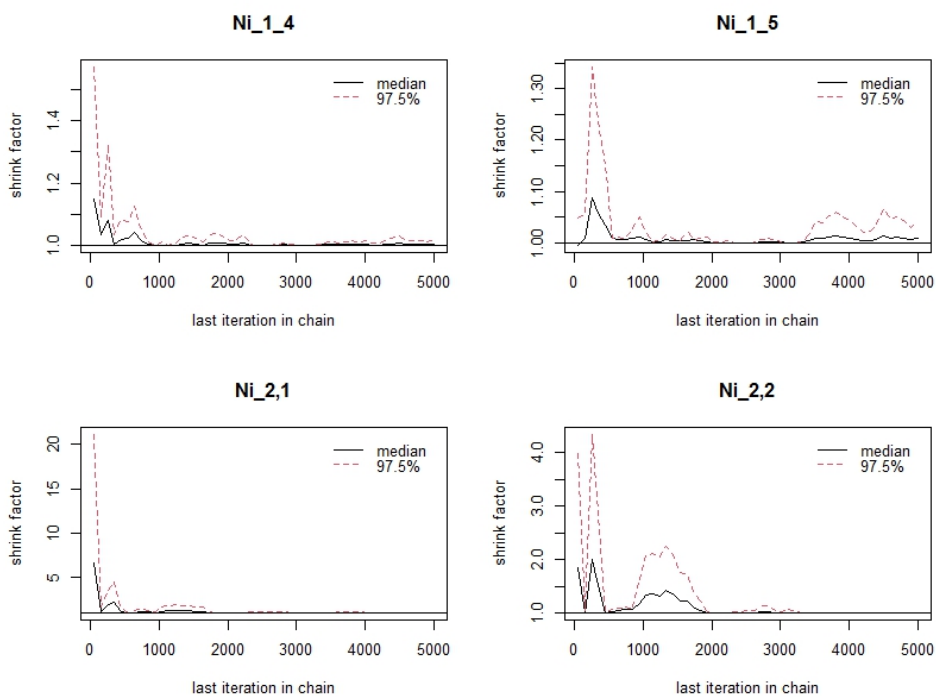
Fonte: Elaborada pelo autor.

Figura 82 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para  $\mu_{52}, v_{11}, v_{12}, v_{13}$



(a)  $\mu_{52}, v_{11}, v_{12}$  e  $v_{13}$

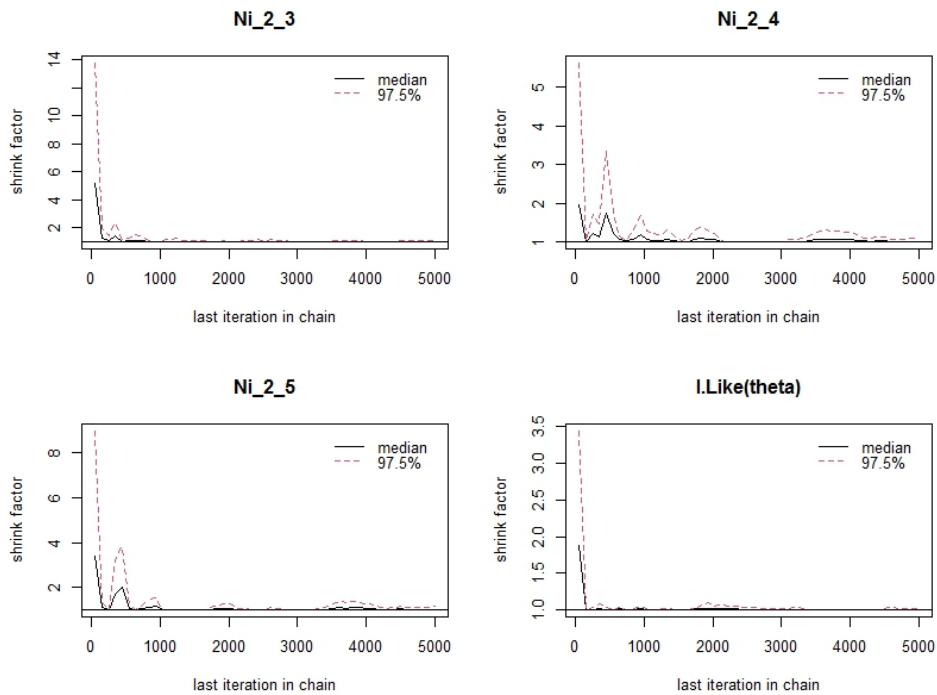
,  $v_{14}, v_{15}, v_{21}$  e  $v_{22}$



(b)  $v_{14}, v_{15}, v_{21}$  e  $v_{22}$

Fonte: Elaborada pelo autor.

Figura 83 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de matemática para  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

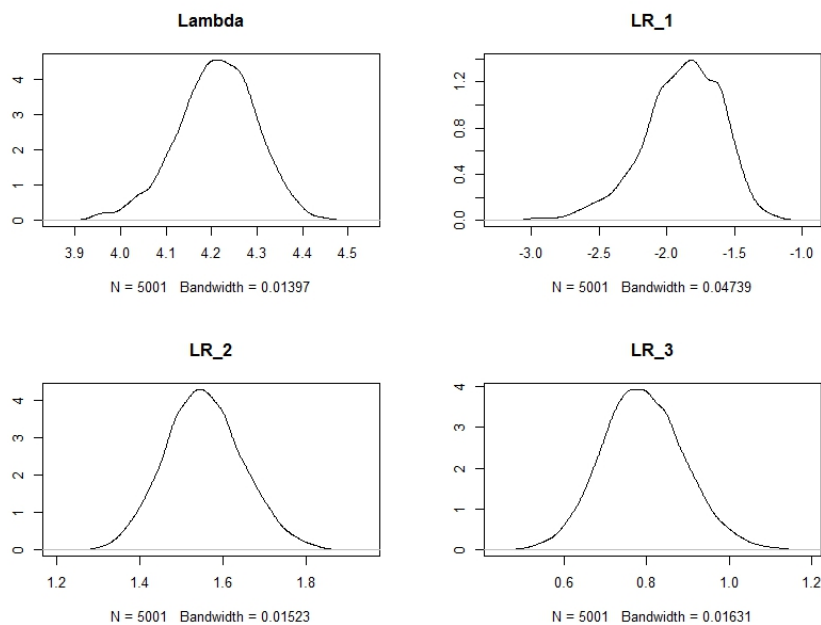


(a)  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

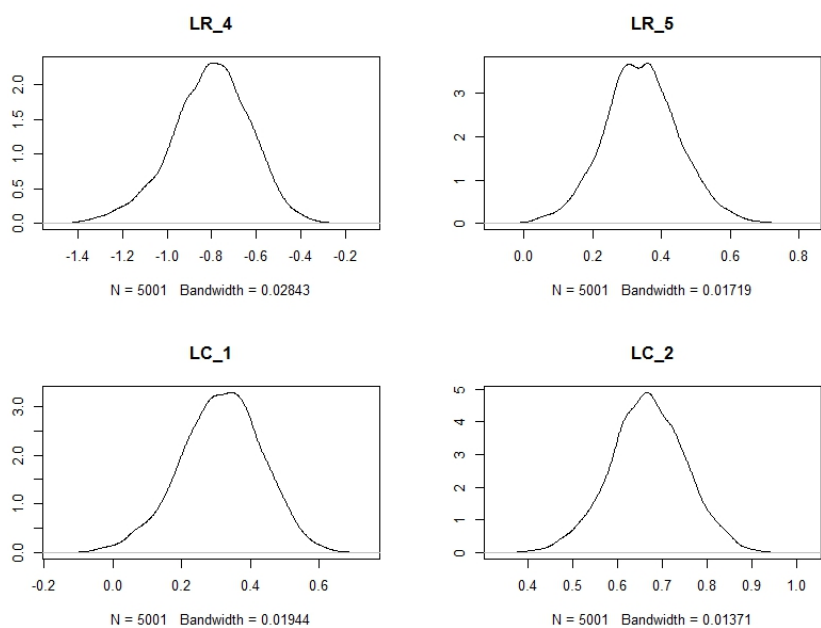
Fonte: Elaborada pelo autor.

## D.15 Gráficos das distribuições *a posteriori* dos resultados do MH aplicado no desempenho na prova de matemática

Figura 84 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de matemática para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



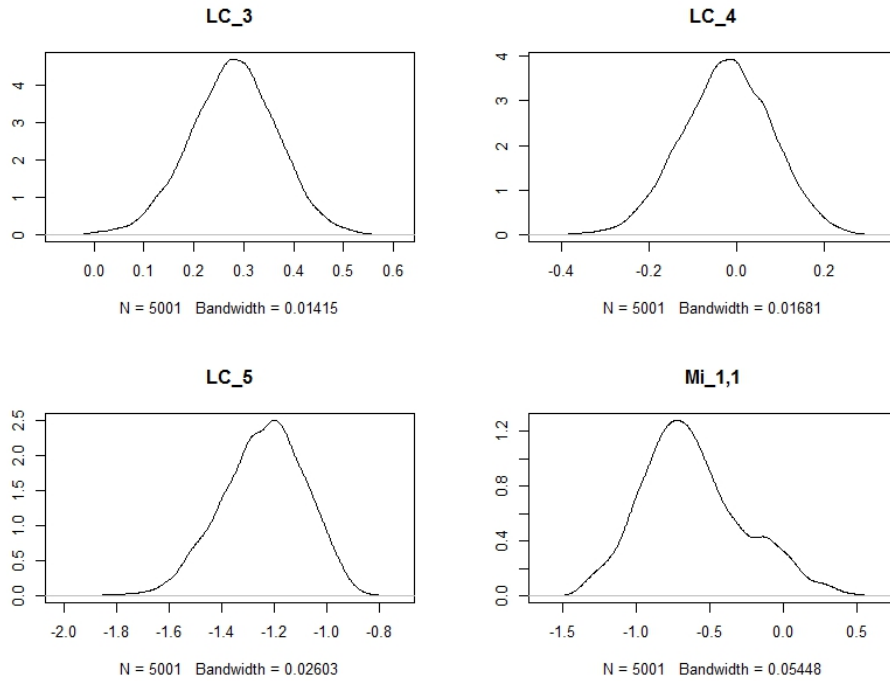
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



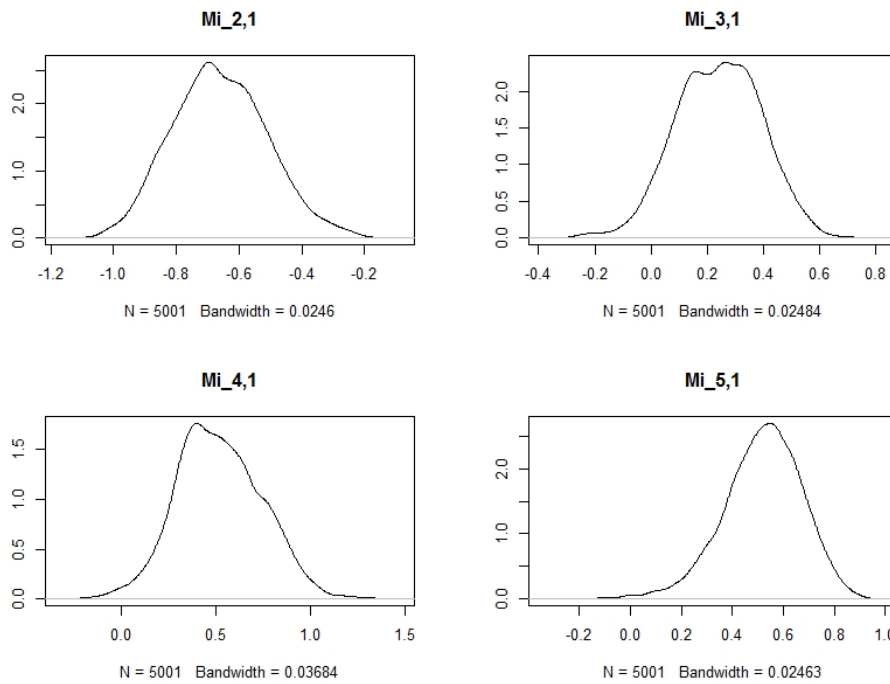
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 85 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de matemática para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



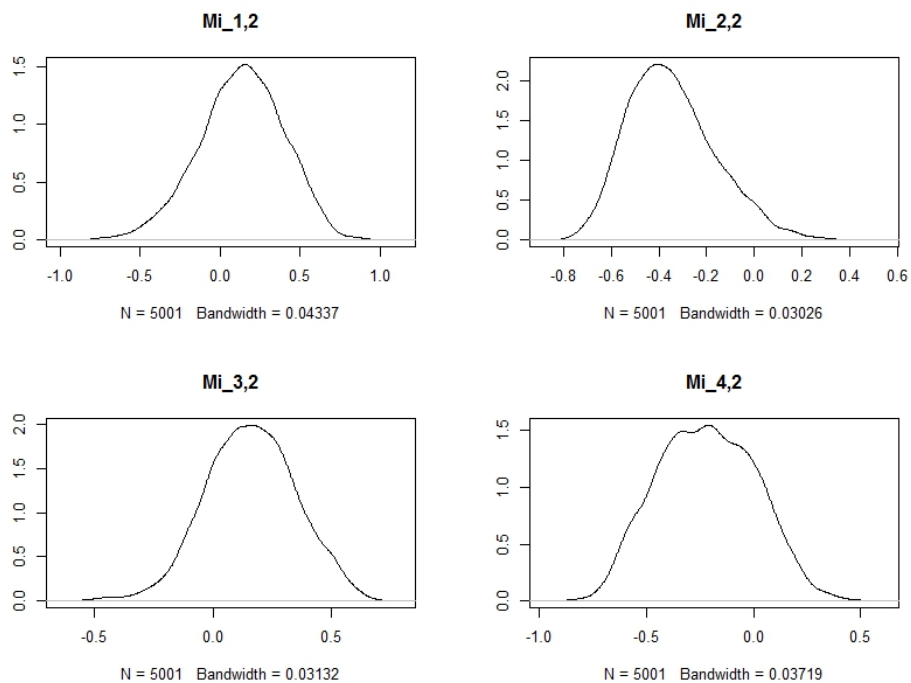
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



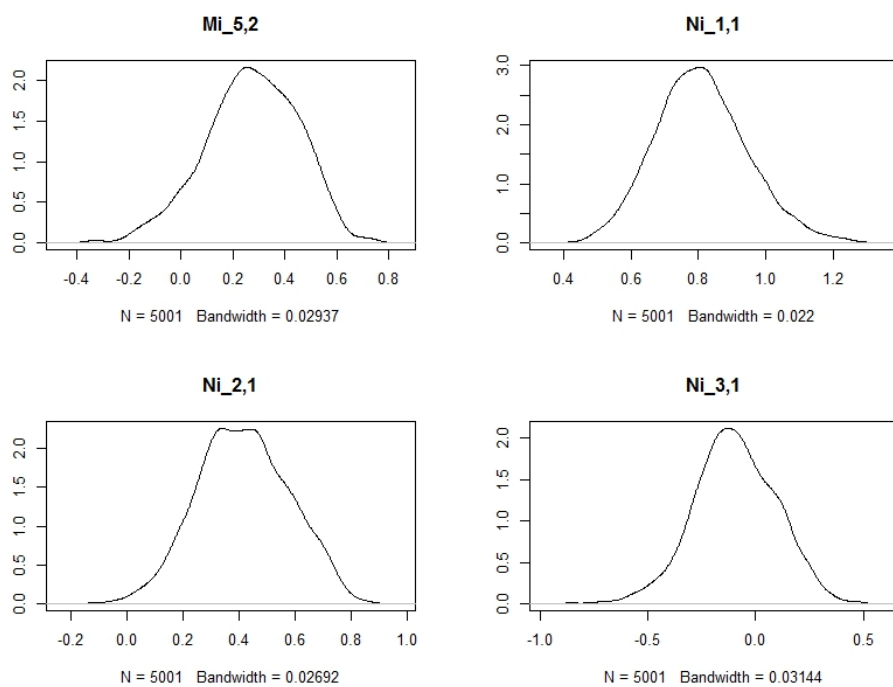
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 86 – Distribuições a posteriori do MH aplicados no desempenho na prova de matemática para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



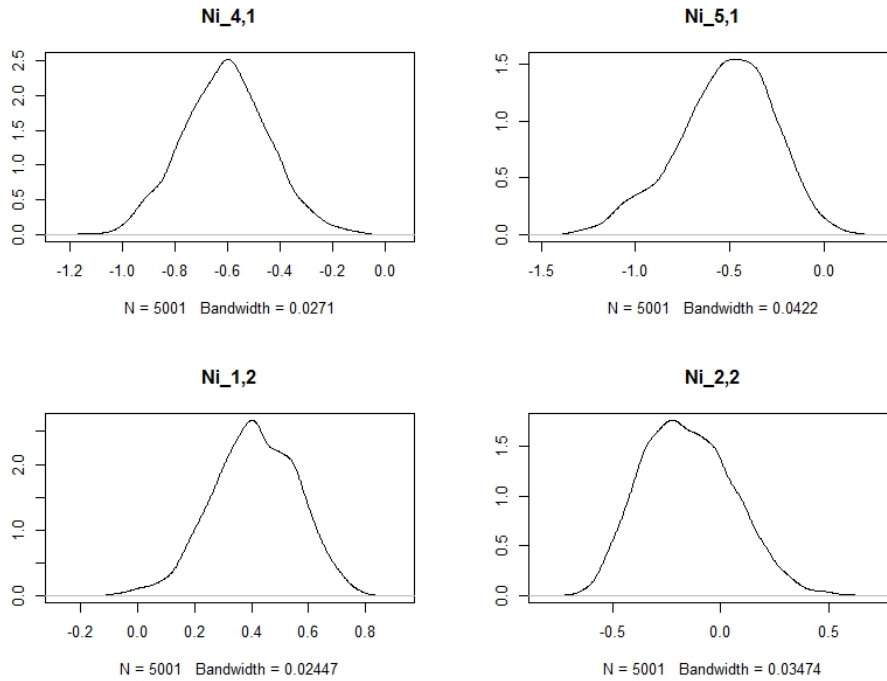
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



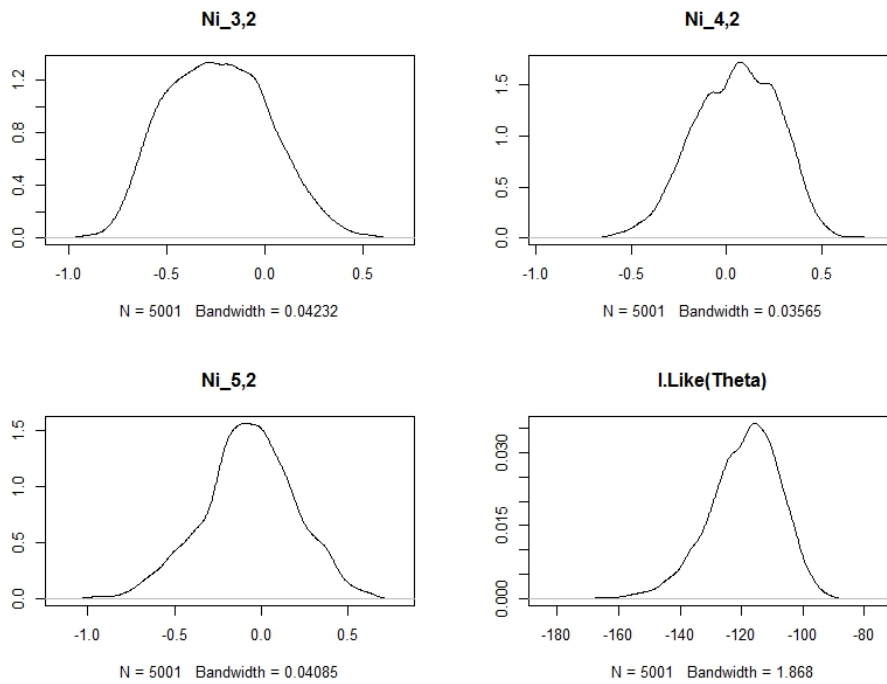
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 87 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de matemática para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$

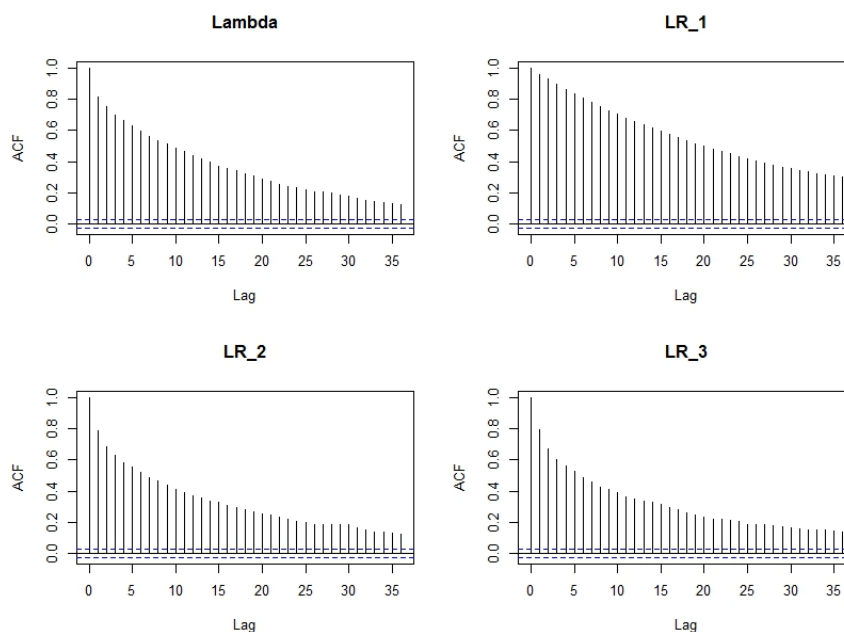


(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

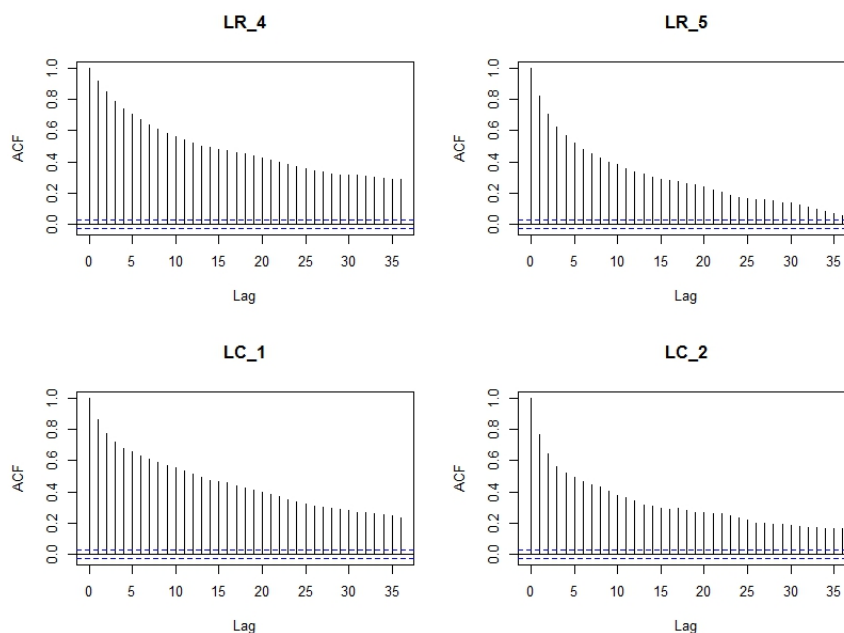
Fonte: Elaborada pelo autor.

## D.16 Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de matemática

Figura 88 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



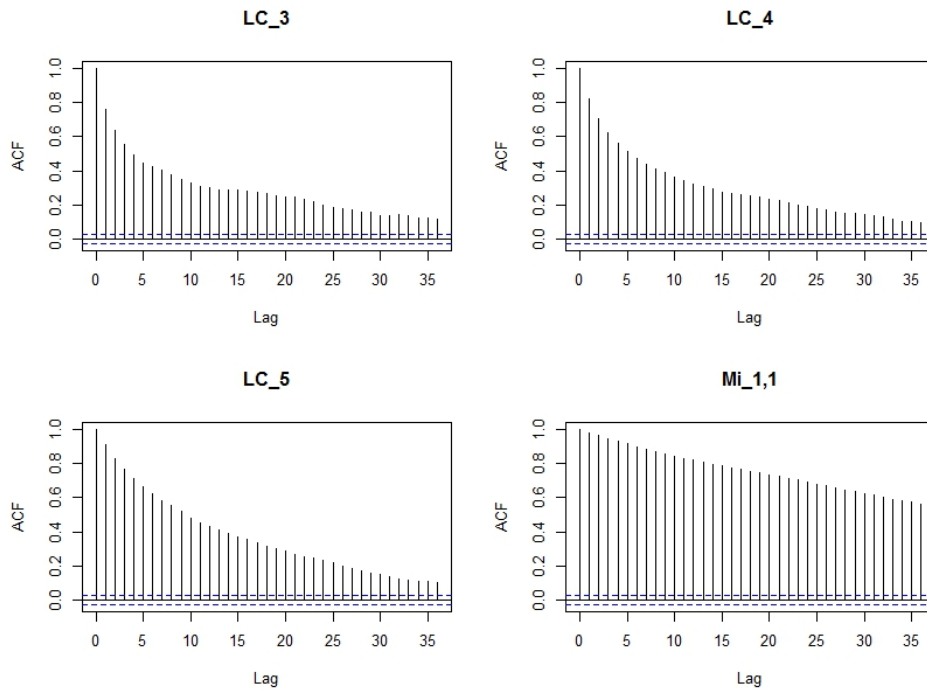
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



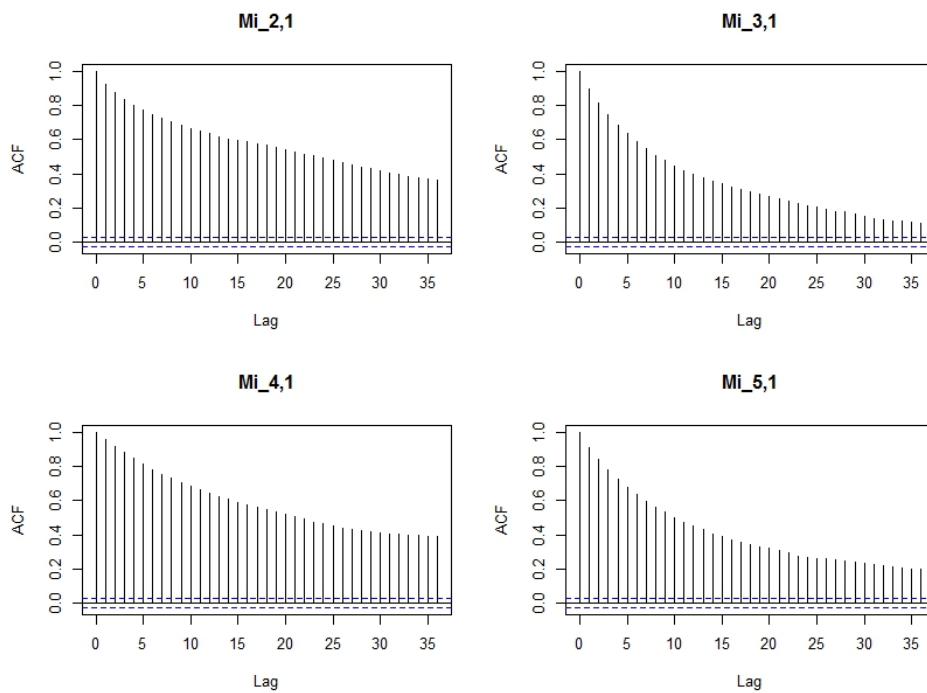
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 89 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



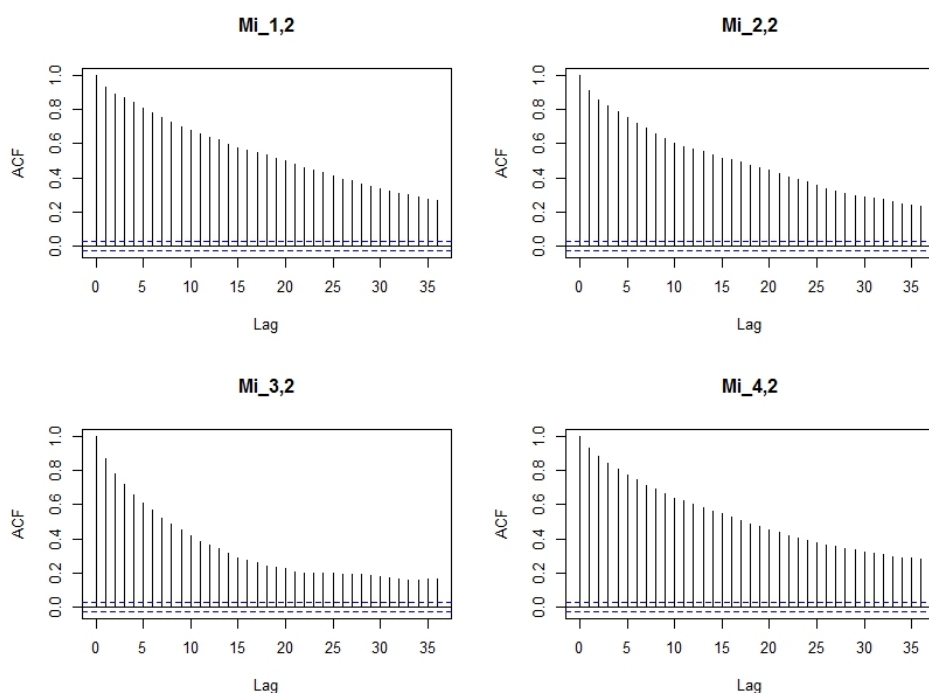
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



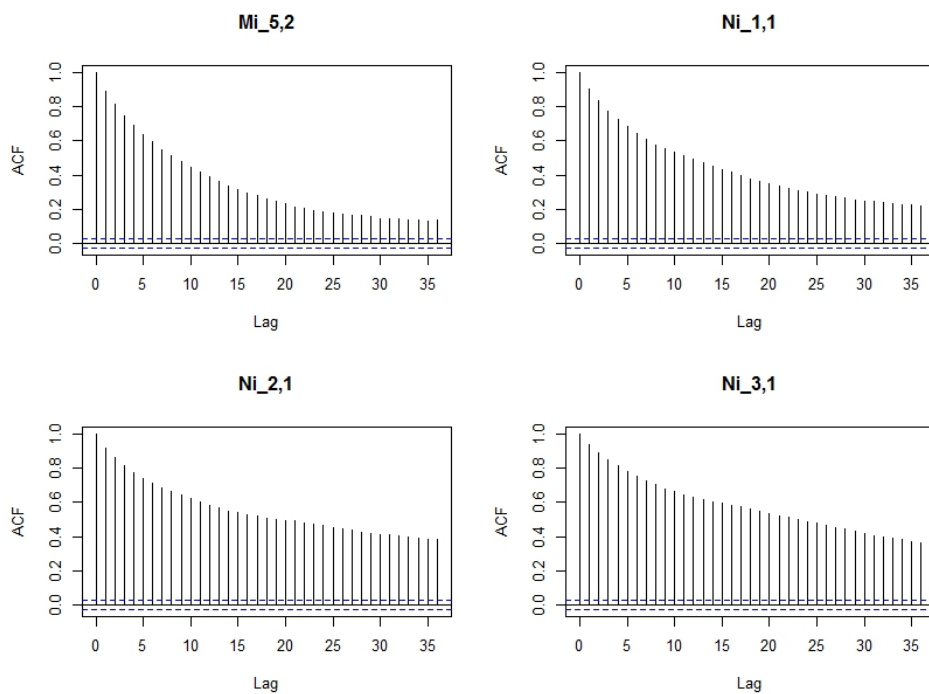
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 90 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $\nu_{11}$ ,  $\nu_{21}$  e  $\nu_{31}$



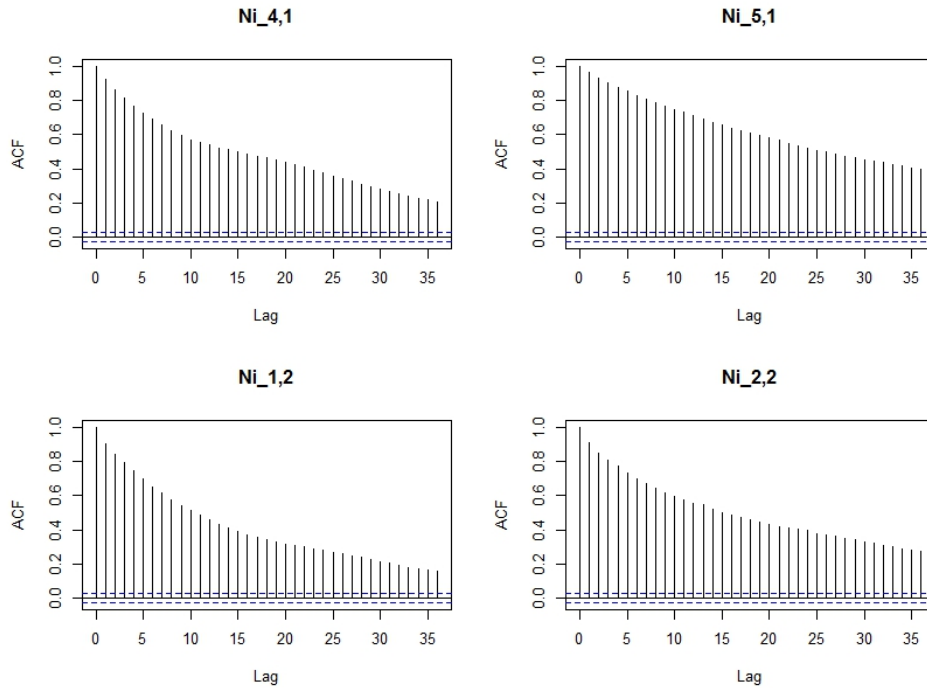
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



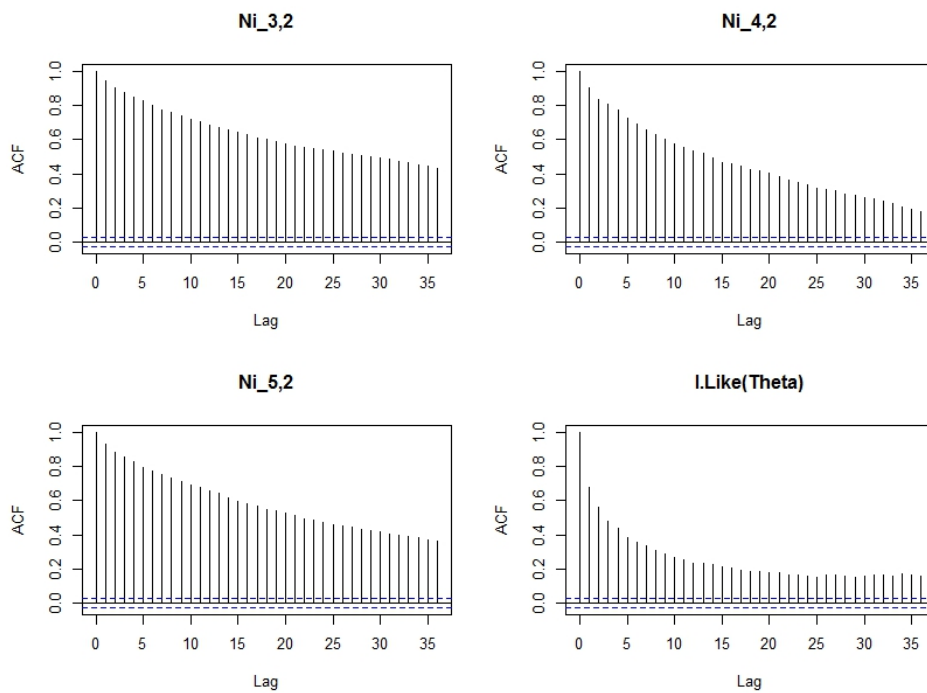
(b)  $\mu_{52}$ ,  $\nu_{11}$ ,  $\nu_{21}$  e  $\nu_{31}$

Fonte: Elaborada pelo autor.

Figura 91 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de matemática para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



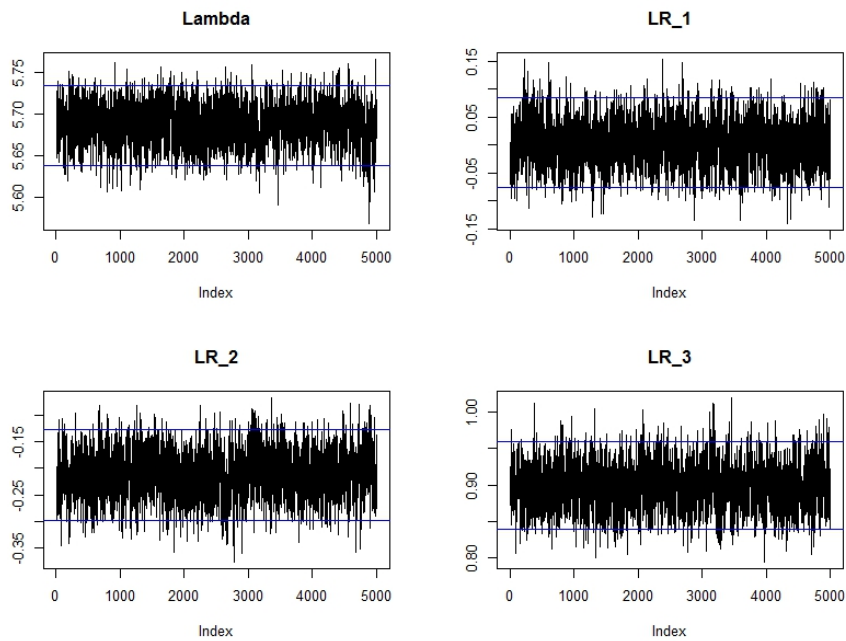
(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

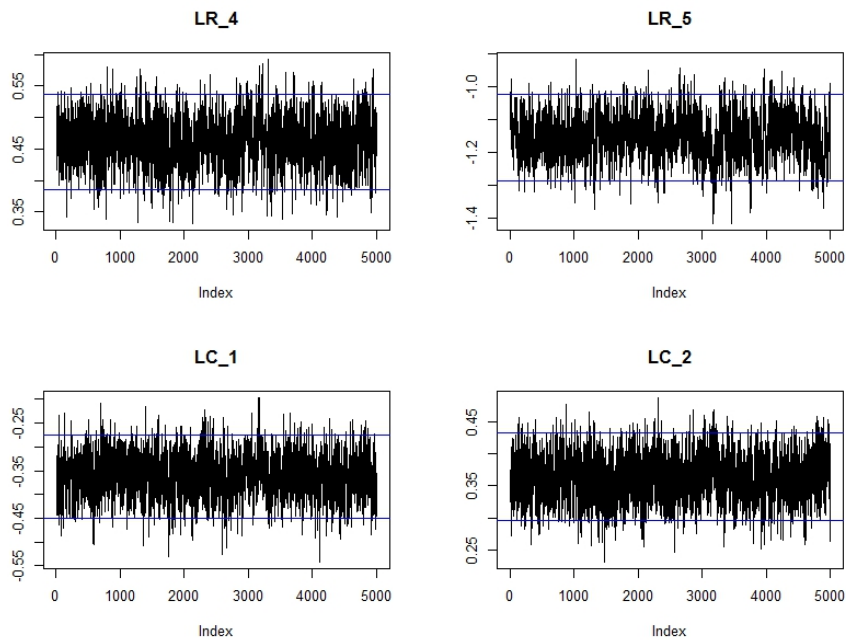
## **D.17 Gráficos trace plot dos resultados do MH aplicado no desempenho na prova de redação.**

Trace plots para a simulação MCMC na aplicação dos resultados da prova de redação.

Figura 92 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

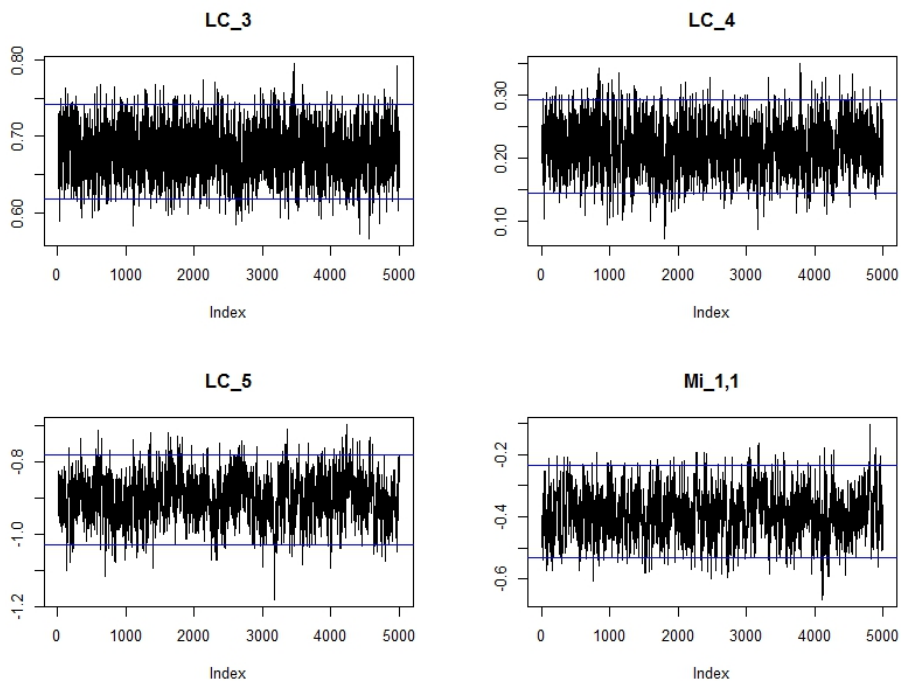


(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

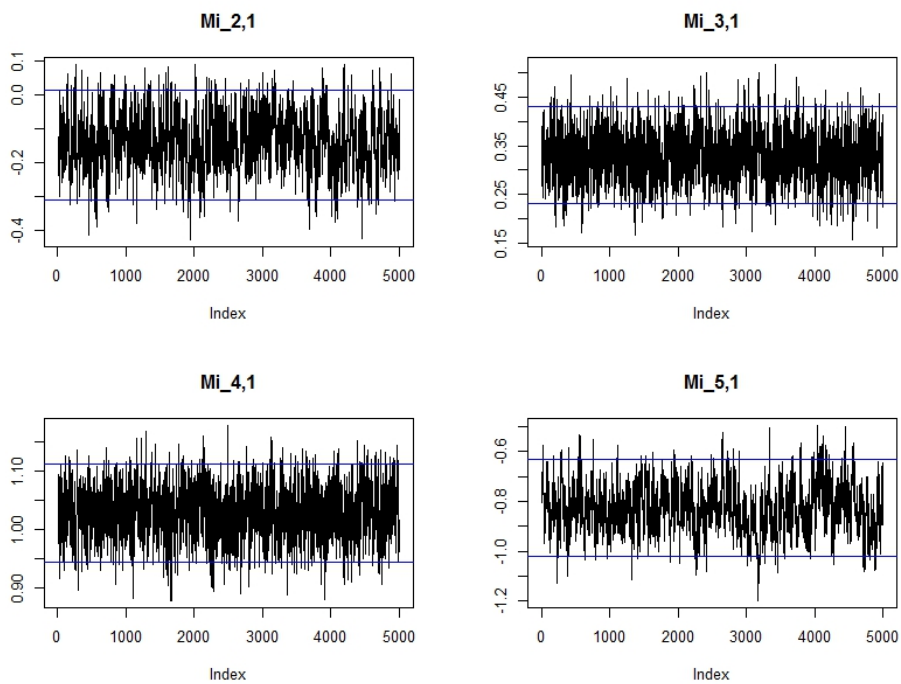
Fonte: Elaborada pelo autor.

Figura 93 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para  $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}$

,  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



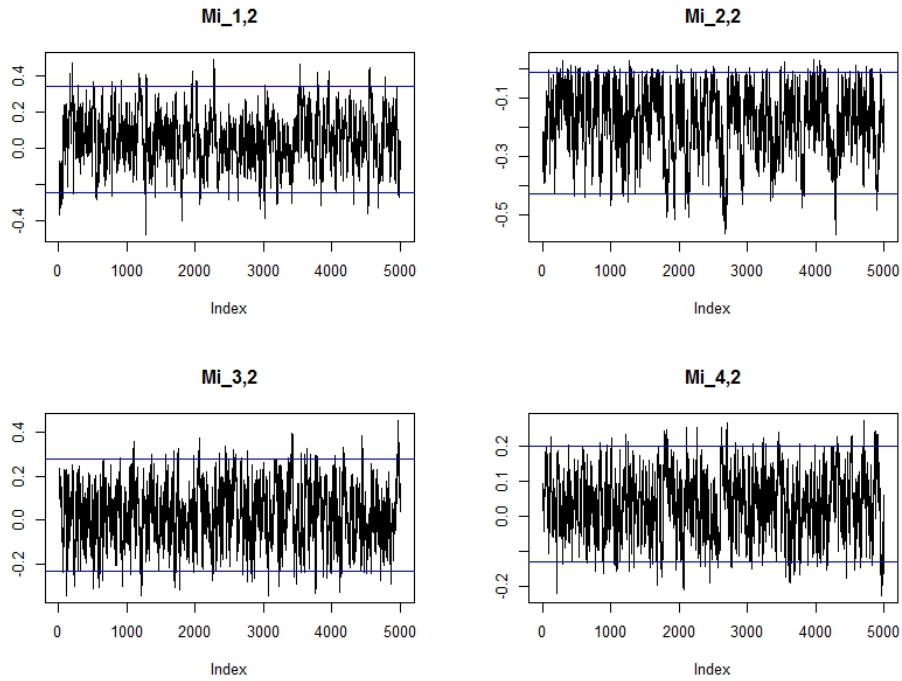
(a)  $\lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$



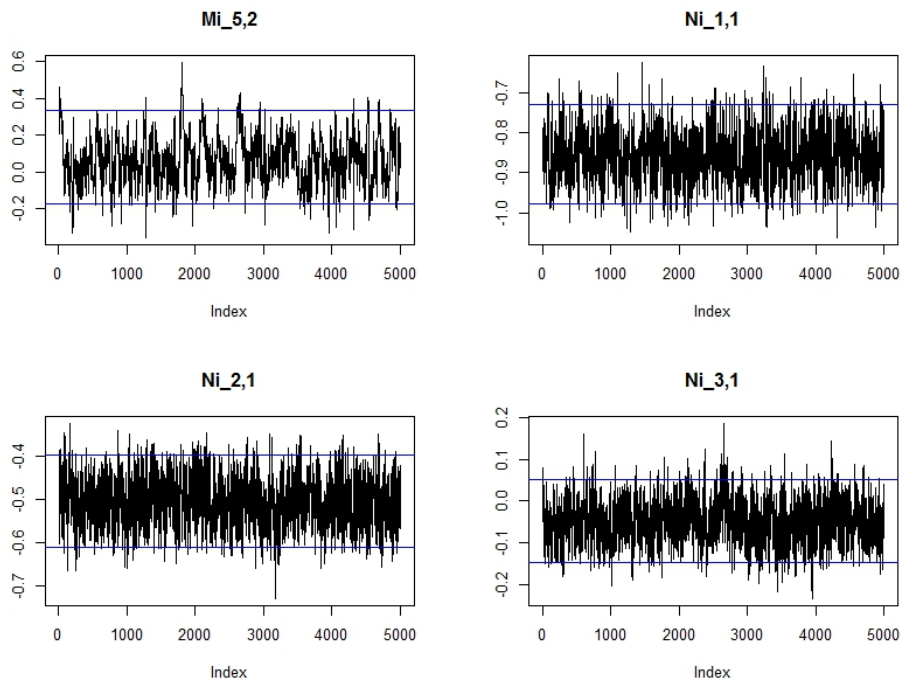
(b)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 94 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



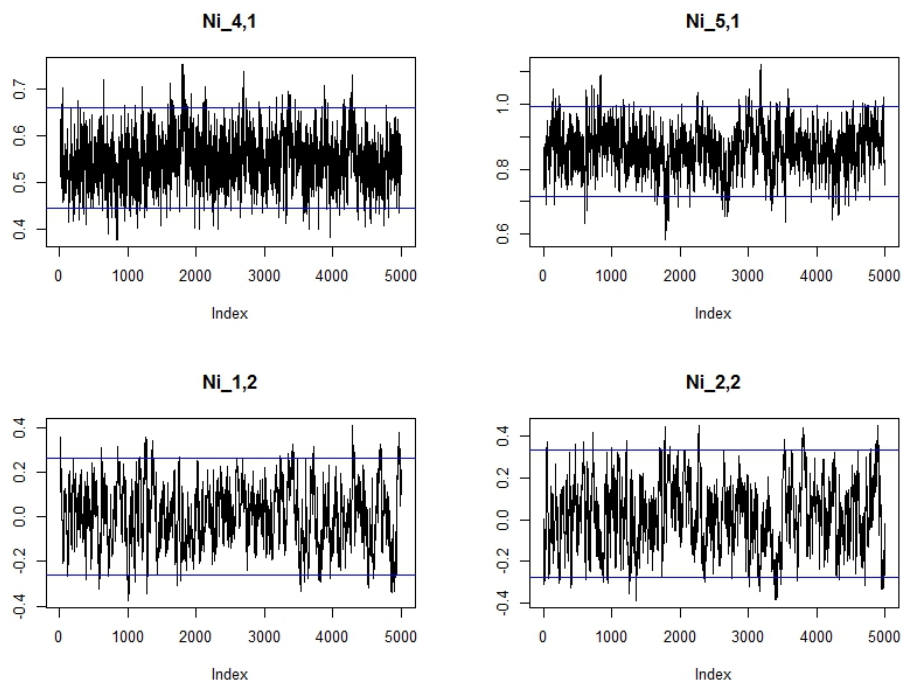
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



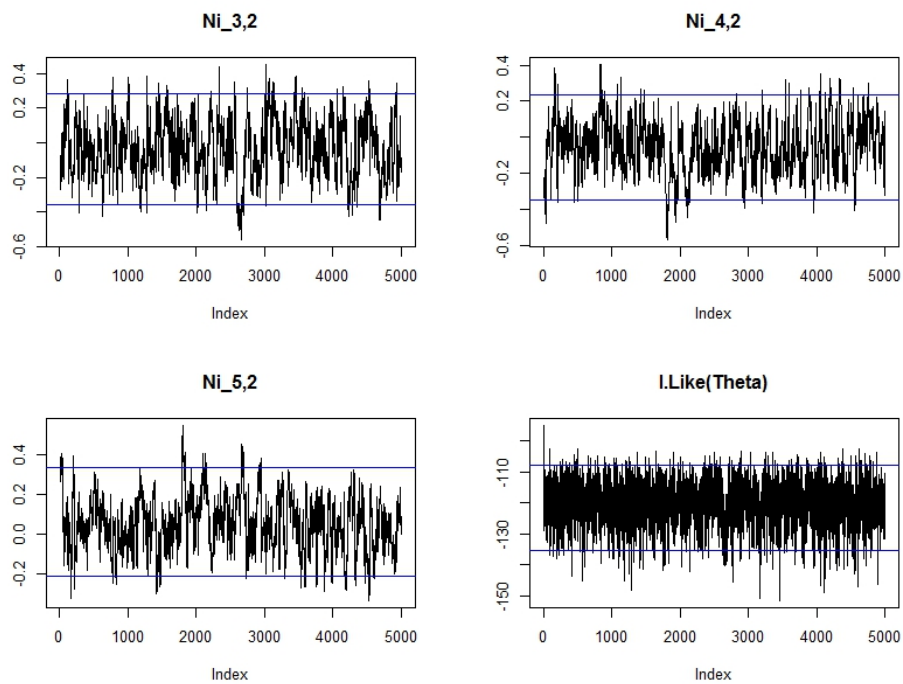
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 95 – Trace plot dos resultados do MH aplicados no desempenho na prova de redação para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

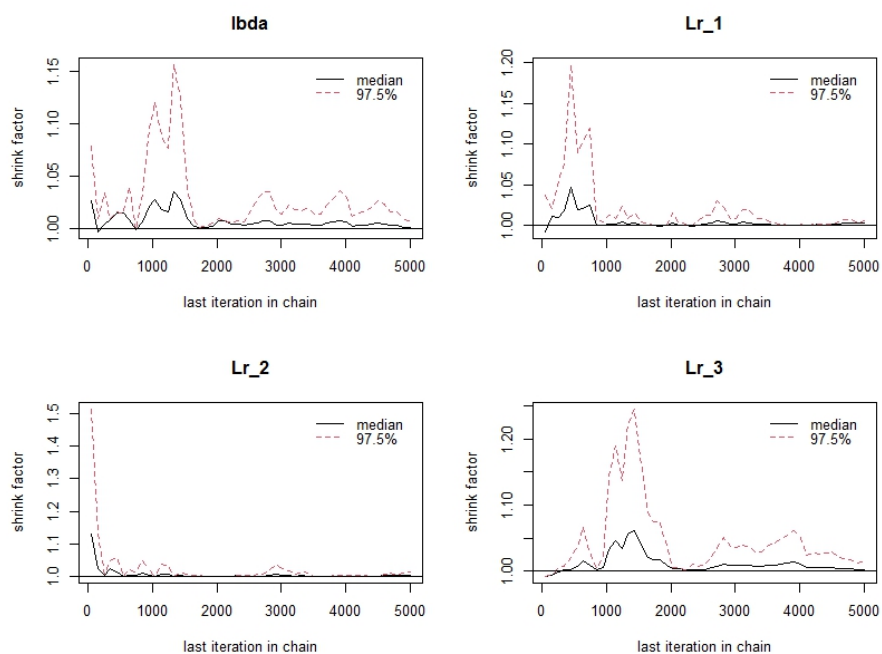
## D.18 Verificação da convergência através do método de Gelman-Rubin para os resultados da prova de redação.

Tabela 22 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação

	Point est.	Upper c.I.		Point est.	Upper c.I.		Point est.	Upper c.I.
$\lambda$	1.00	1.01	$\mu_{11}$	1.01	1.06	$v_{11}$	1.00	1.01
$\lambda_1^R$	1.00	1.01	$\mu_{21}$	1.00	1.00	$v_{12}$	1.00	1.00
$\lambda_2^R$	1.00	1.01	$\mu_{31}$	1.01	1.03	$v_{13}$	1.00	1.01
$\lambda_3^R$	1.00	1.01	$\mu_{41}$	1.00	1.00	$v_{14}$	1.00	1.01
$\lambda_4^R$	1.00	1.00	$\mu_{51}$	1.01	1.06	$v_{15}$	1.01	1.02
$\lambda_5^R$	1.00	1.00	$\mu_{12}$	1.02	1.07	$v_{21}$	1.02	1.06
$\lambda_1^C$	1.00	1.00	$\mu_{22}$	1.00	1.01	$v_{22}$	1.01	1.02
$\lambda_2^C$	1.00	1.01	$\mu_{32}$	1.05	1.20	$v_{23}$	1.00	1.00
$\lambda_3^C$	1.00	1.02	$\mu_{42}$	1.06	1.24	$v_{24}$	1.02	1.10
$\lambda_4^C$	1.00	1.00	$\mu_{52}$	1.02	1.09	$v_{25}$	1.03	1.09
$\lambda_5^C$	1.01	1.02	—	—	—	$l.Log(\theta)$	1.00	1.01

Fonte: Elaborada pelo autor.

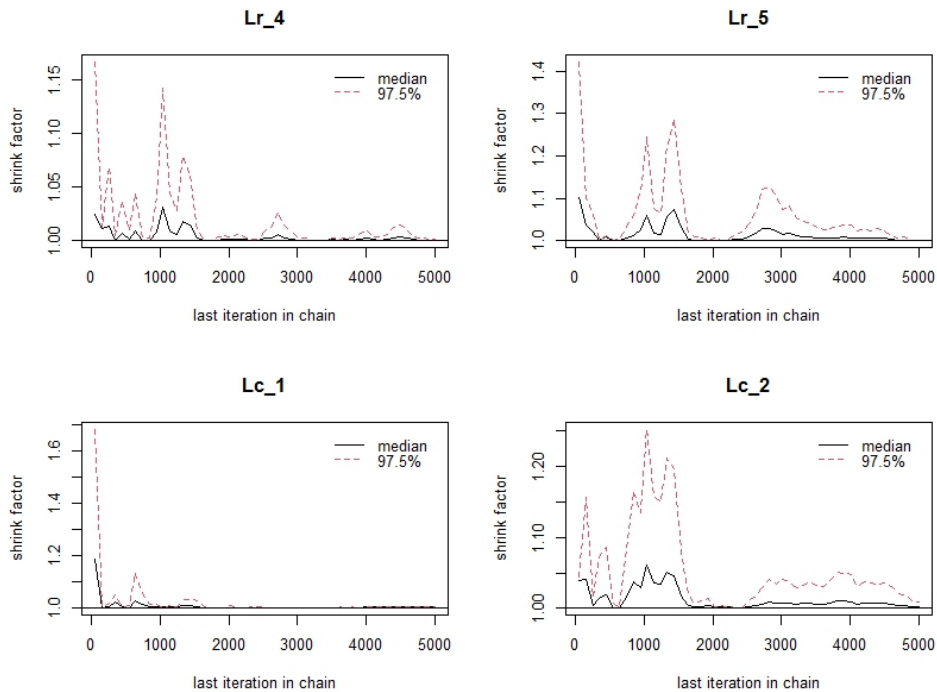
Figura 96 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



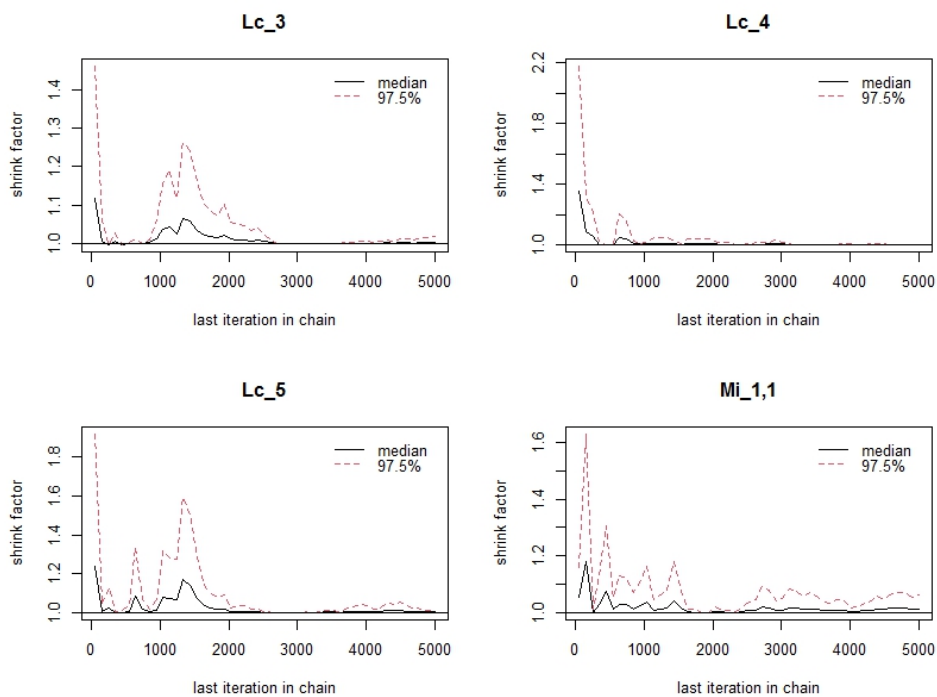
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

Fonte: Elaborada pelo autor.

Figura 97 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para  $\lambda_4^R, \lambda_5^R, \lambda_1^C, \lambda_2^C, \lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$



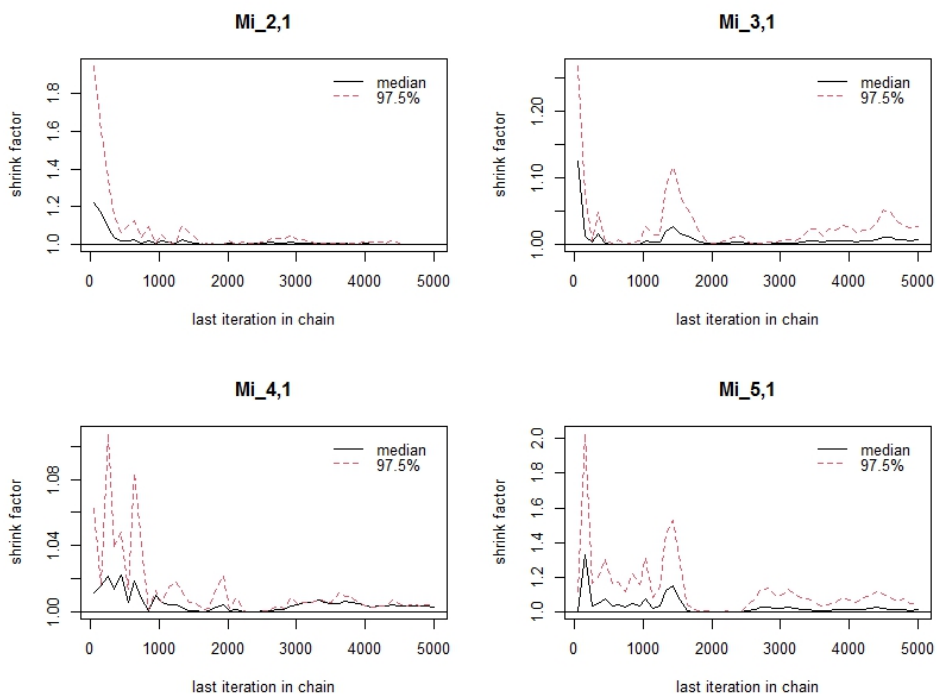
(a)  $\lambda_4^R, \lambda_5^R, \lambda_1^C$  e  $\lambda_2^C$



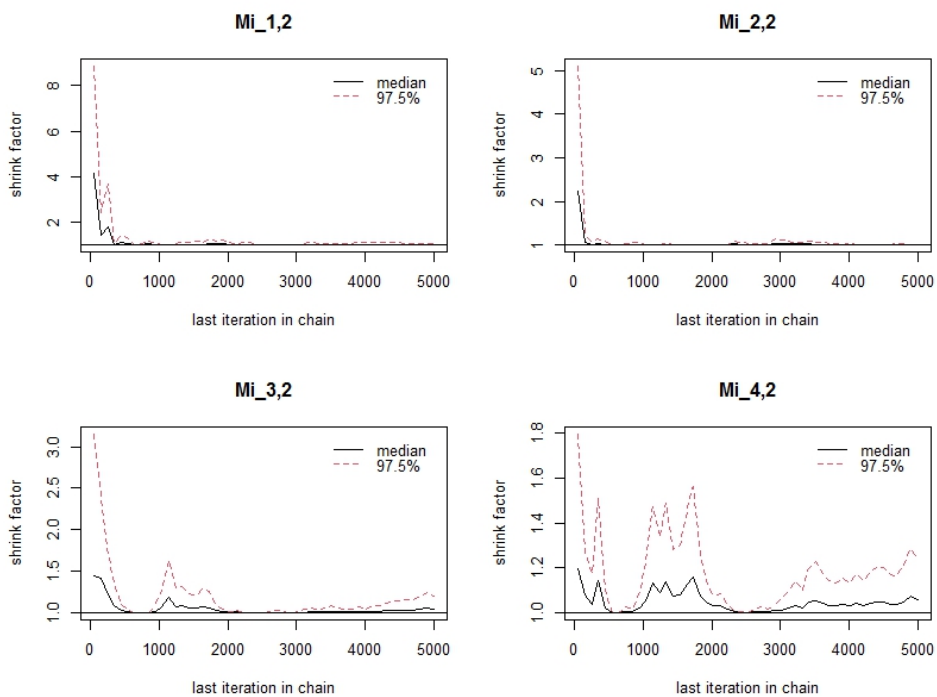
(b)  $\lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$

Fonte: Elaborada pelo autor.

Figura 98 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para  $\mu_{21}, \mu_{31}, \mu_{41}, \mu_{51}, \mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$



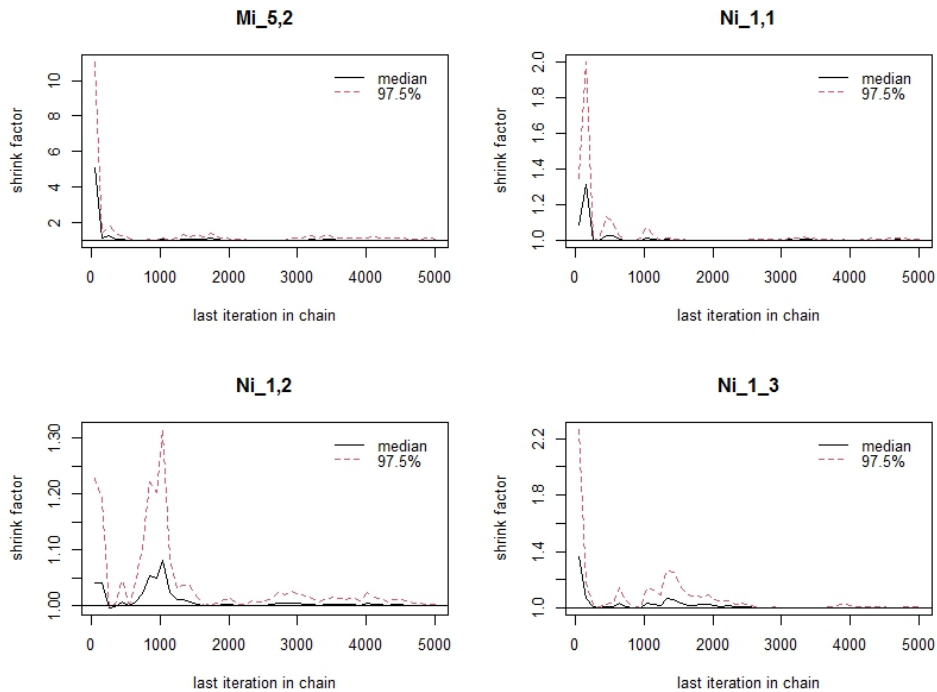
(a)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



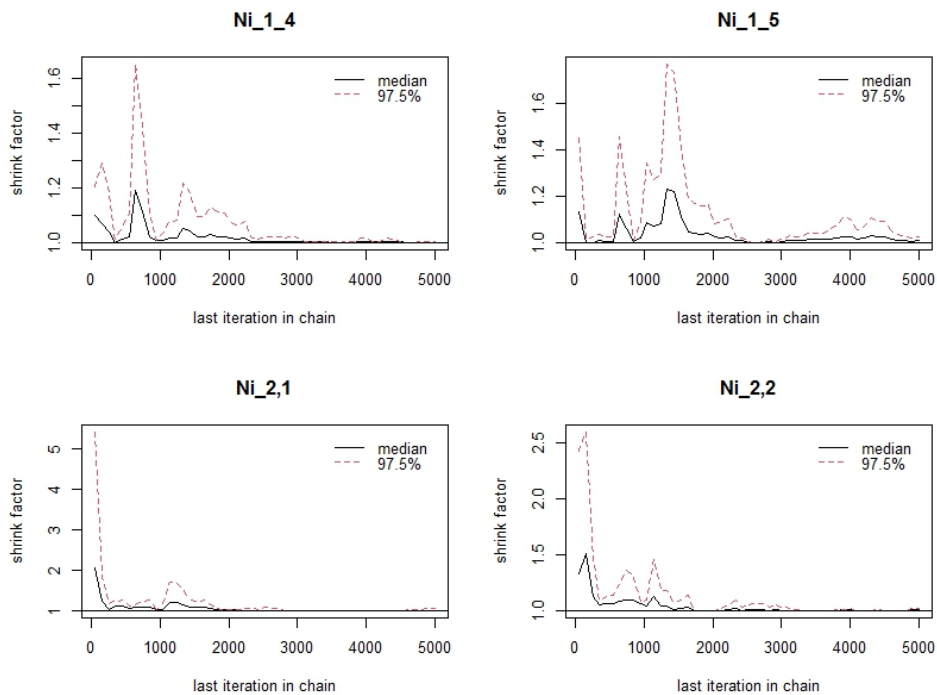
(b)  $\mu_{12}, \mu_{22}, \mu_{32}$  e  $\mu_{42}$

Fonte: Elaborada pelo autor.

Figura 99 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para  $\mu_{52}, \nu_{11}, \nu_{12}, \nu_{13}, \nu_{14}, \nu_{15}, \nu_{21}$  e  $\nu_{22}$



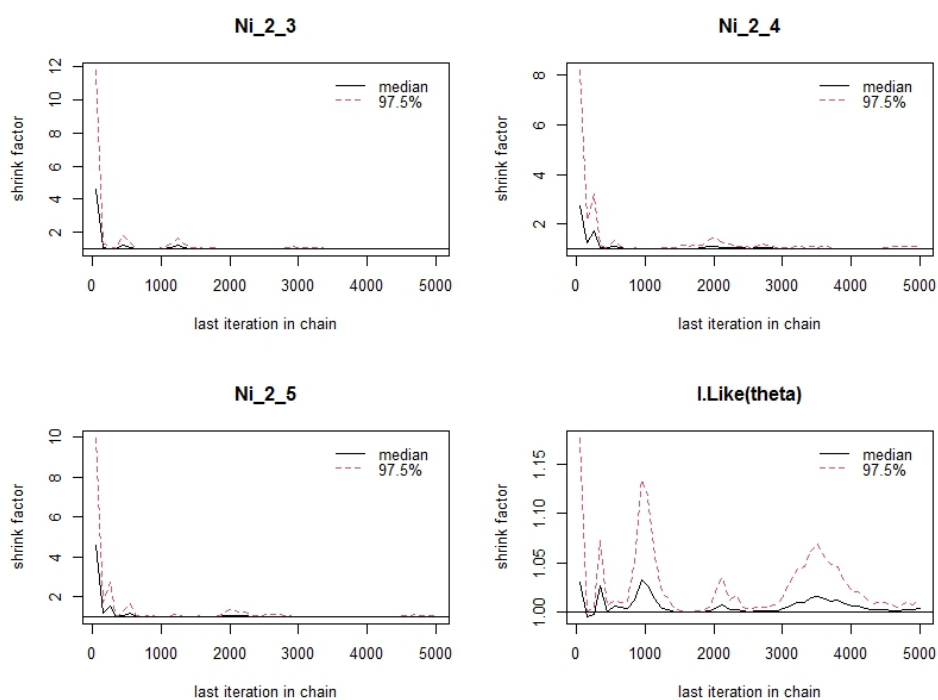
(a)  $\mu_{52}, \nu_{11}, \nu_{12}$  e  $\nu_{13}$



(b)  $\nu_{14}, \nu_{15}, \nu_{21}$  e  $\nu_{22}$

Fonte: Elaborada pelo autor.

Figura 100 – Teste de convergência de Gelman-Rubin para os resultados do MH aplicados no desempenho na prova de redação para  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

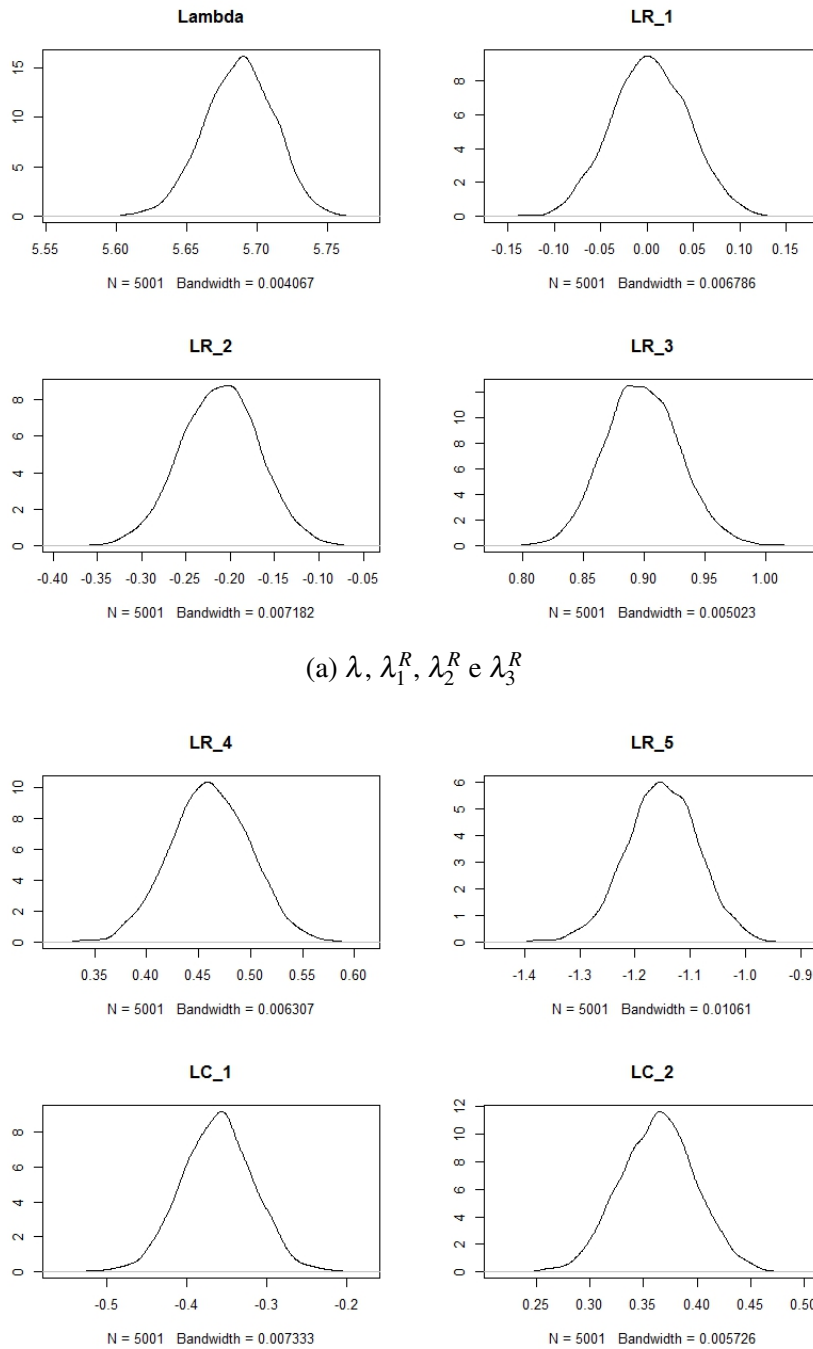


(a)  $v_{23}$ ,  $v_{24}$ ,  $v_{25}$  e log-verossimilhança

Fonte: Elaborada pelo autor.

## D.19 Gráficos das distribuições *a posteriori* dos resultados do MH aplicado no desempenho na prova de redação

Figura 101 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de redação para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

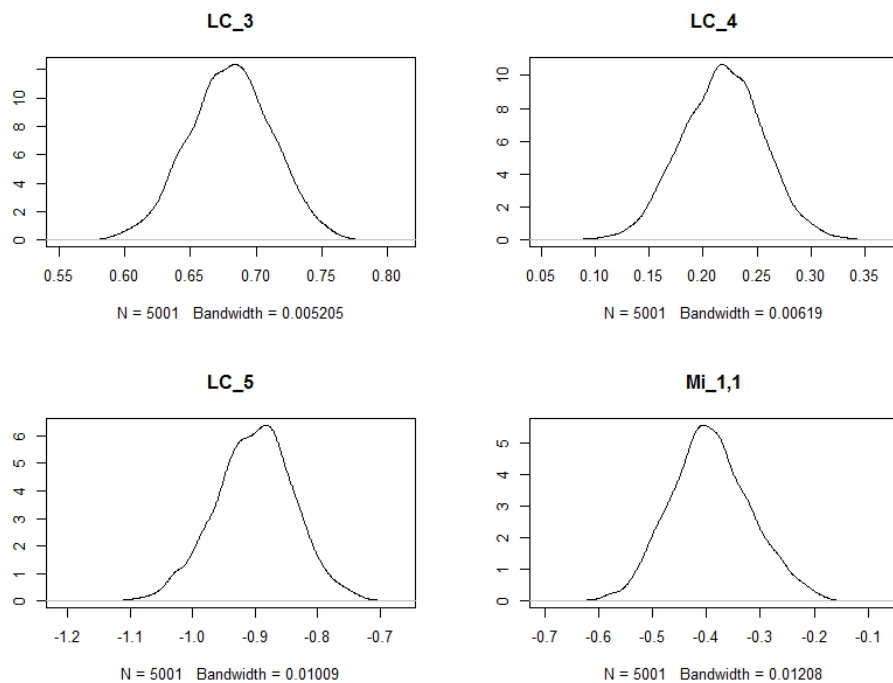


(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$

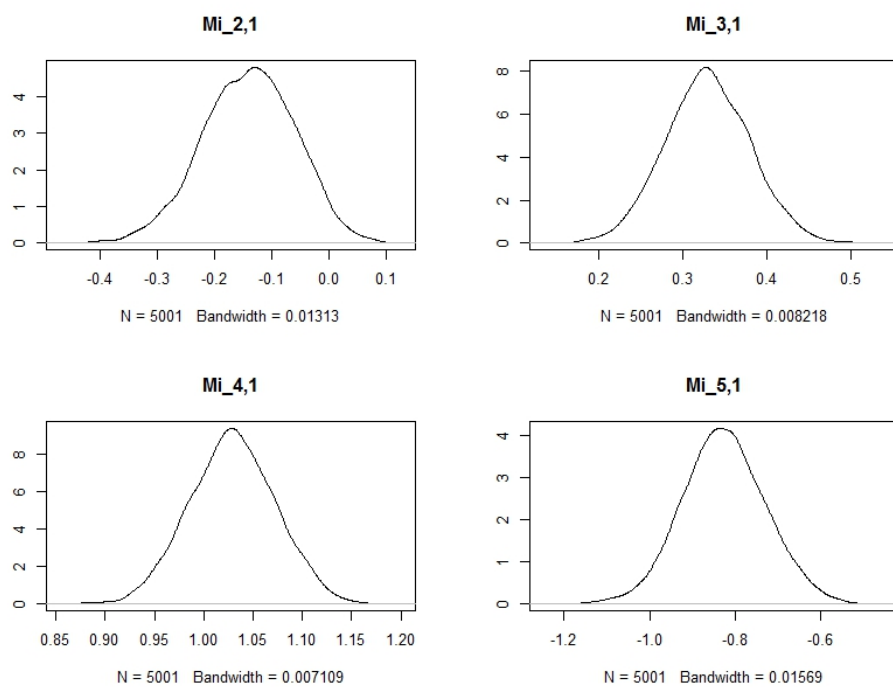
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 102 – Distribuições a posteriori do MH aplicados no desempenho na prova de redação para  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$ ,  $\mu_{11}$ ,  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$



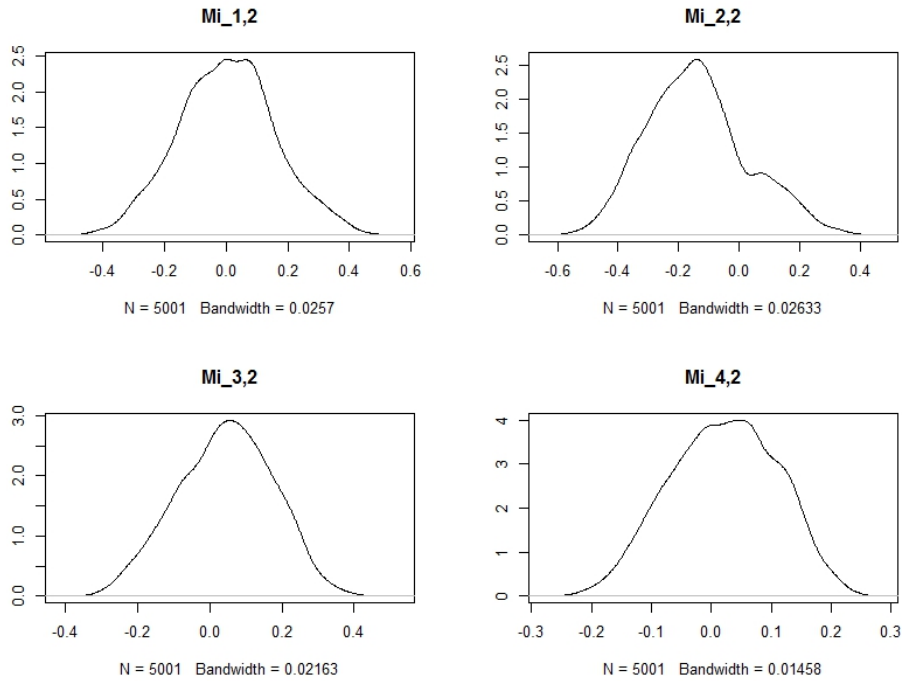
(a)  $\lambda_3^C$ ,  $\lambda_4^C$ ,  $\lambda_5^C$  e  $\mu_{11}$



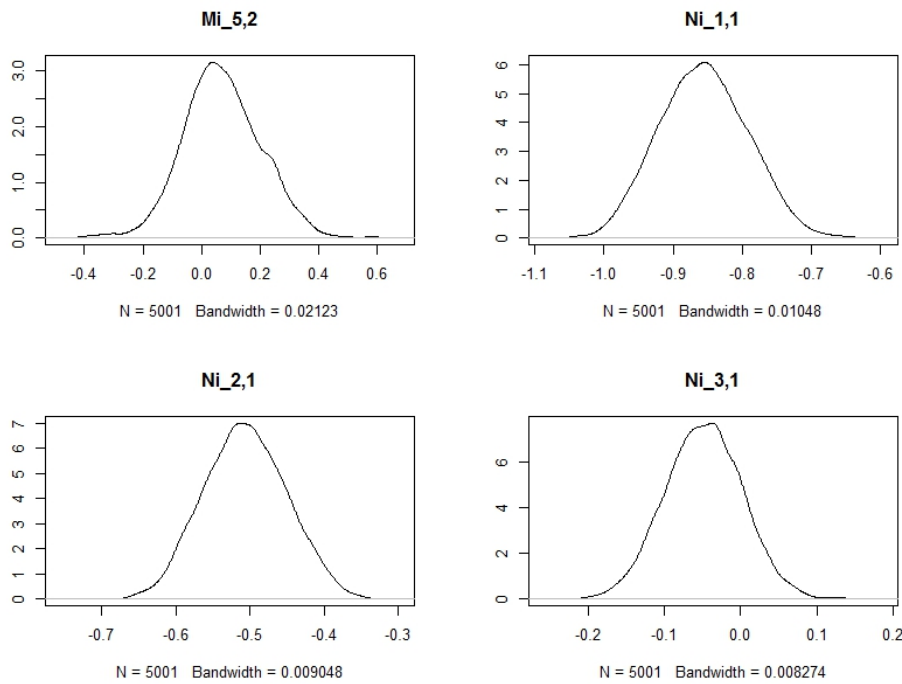
(b)  $\mu_{21}$ ,  $\mu_{31}$ ,  $\mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 103 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de redação para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



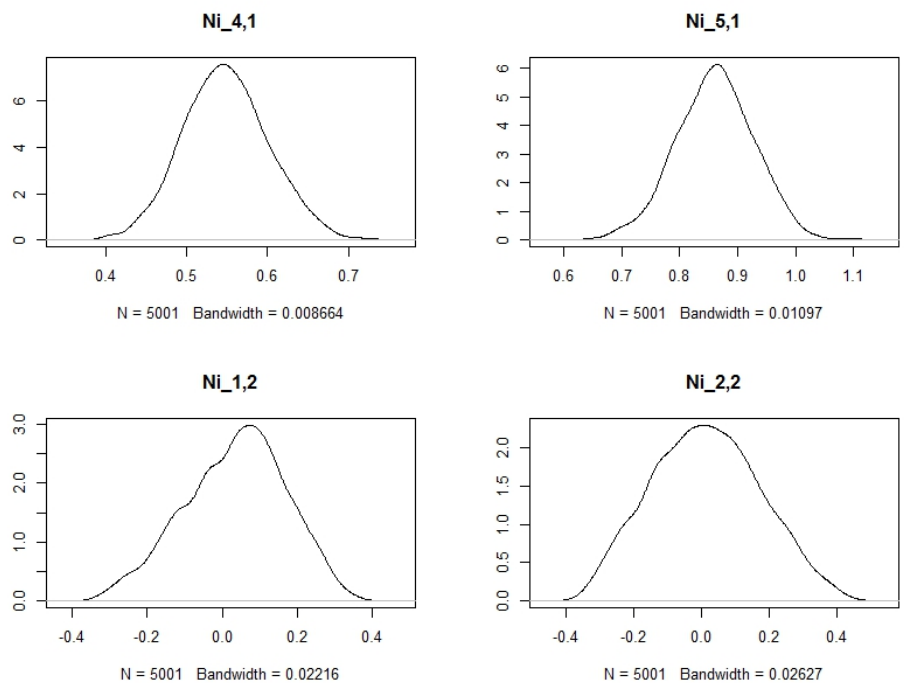
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



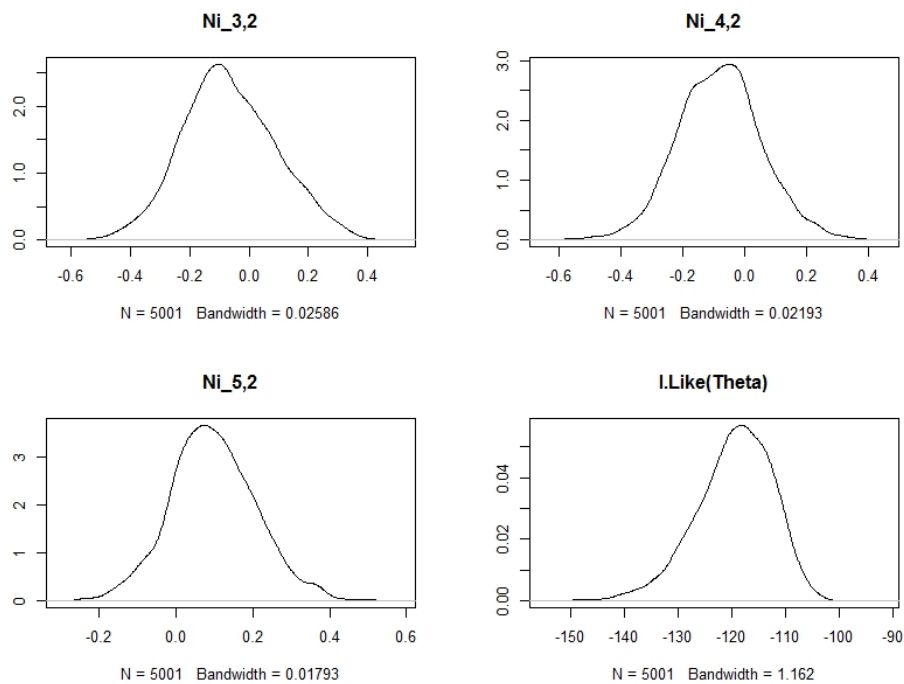
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 104 – Distribuições *a posteriori* do MH aplicados no desempenho na prova de redação para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$

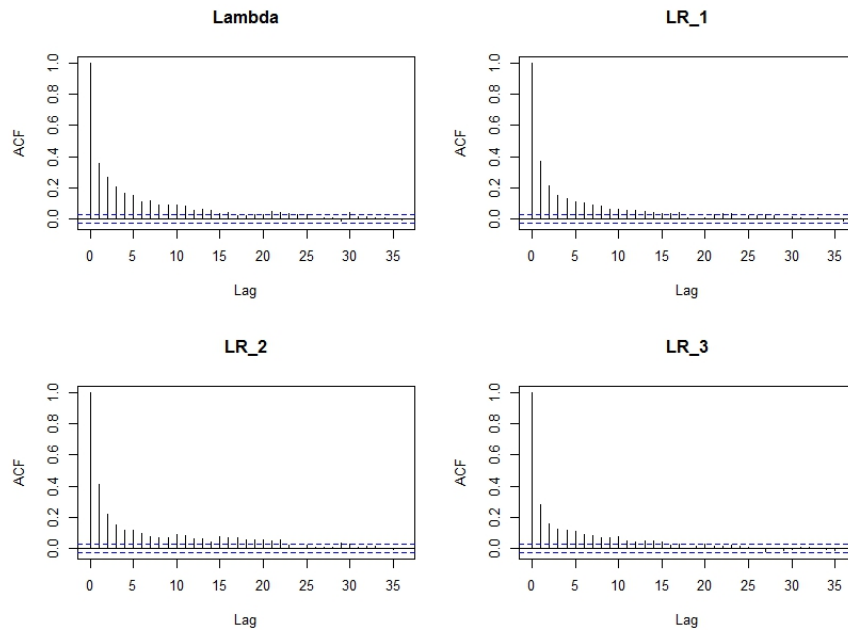


(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

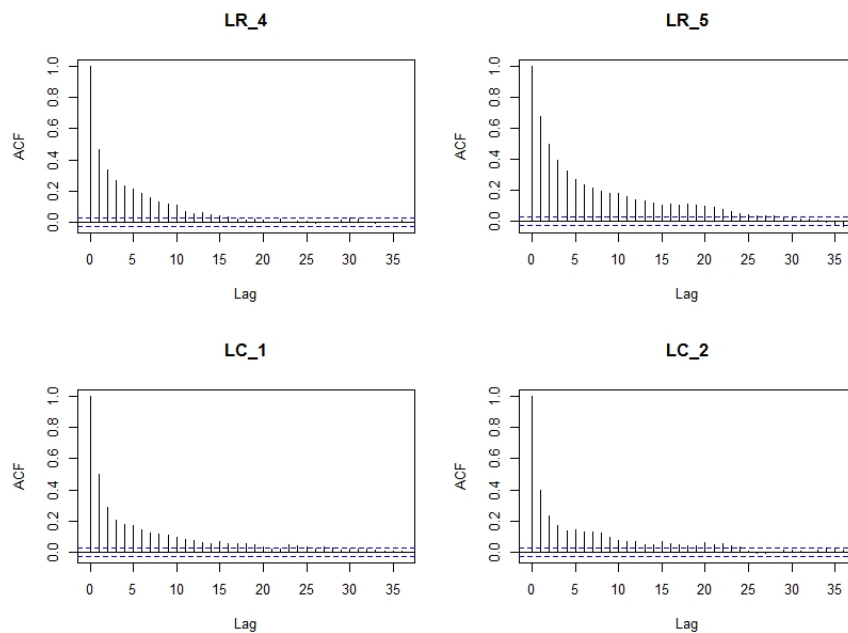
Fonte: Elaborada pelo autor.

## D.20 Gráficos de auto correlação dos resultados do MH aplicados no desempenho na prova de redação

Figura 105 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$ ,  $\lambda_3^R$ ,  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$



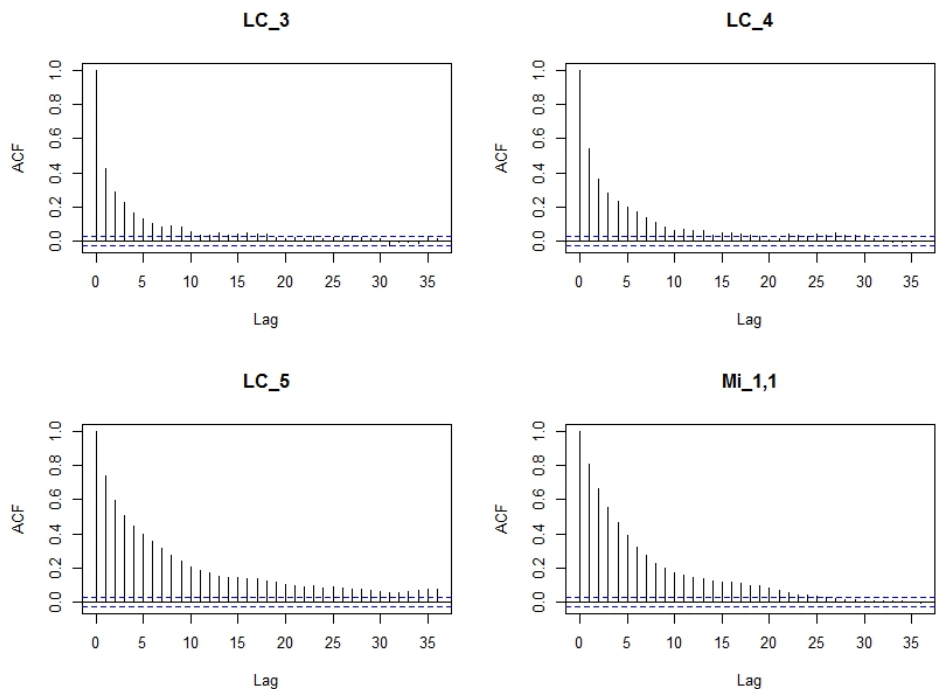
(a)  $\lambda$ ,  $\lambda_1^R$ ,  $\lambda_2^R$  e  $\lambda_3^R$



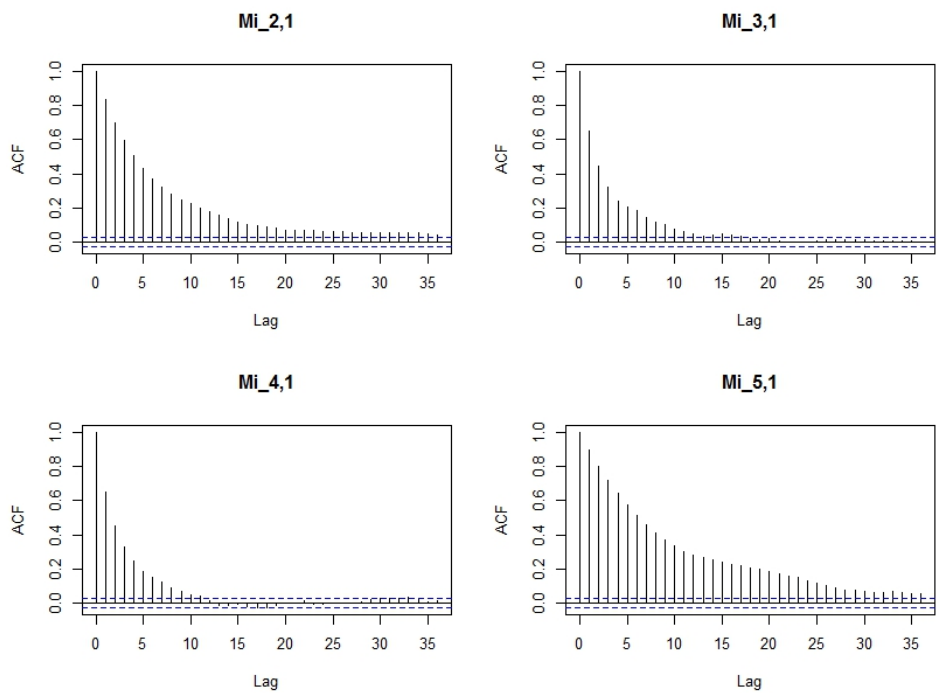
(b)  $\lambda_4^R$ ,  $\lambda_5^R$ ,  $\lambda_1^C$  e  $\lambda_2^C$

Fonte: Elaborada pelo autor.

Figura 106 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para  $\lambda_3^C, \lambda_4^C, \lambda_5^C, \mu_{11}, \mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$



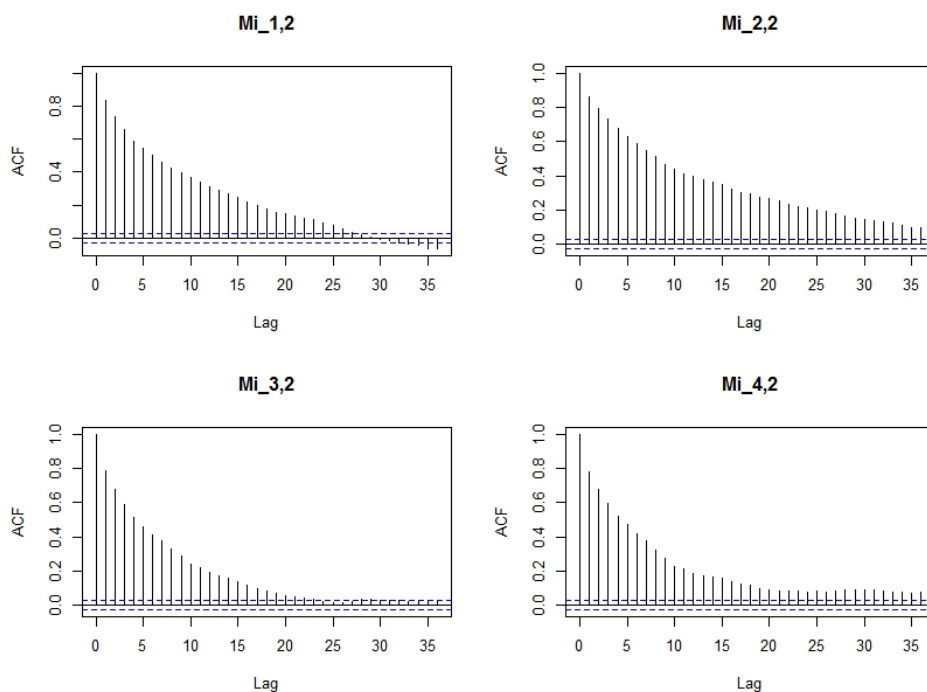
(a)  $\lambda_3^C, \lambda_4^C, \lambda_5^C$  e  $\mu_{11}$



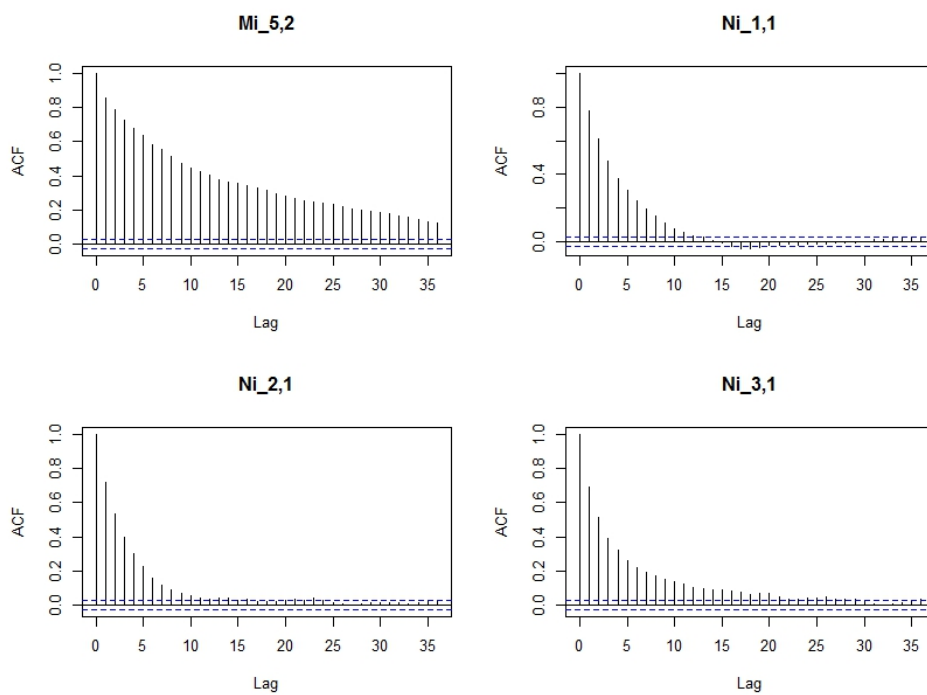
(b)  $\mu_{21}, \mu_{31}, \mu_{41}$  e  $\mu_{51}$

Fonte: Elaborada pelo autor.

Figura 107 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$ ,  $\mu_{42}$ ,  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$



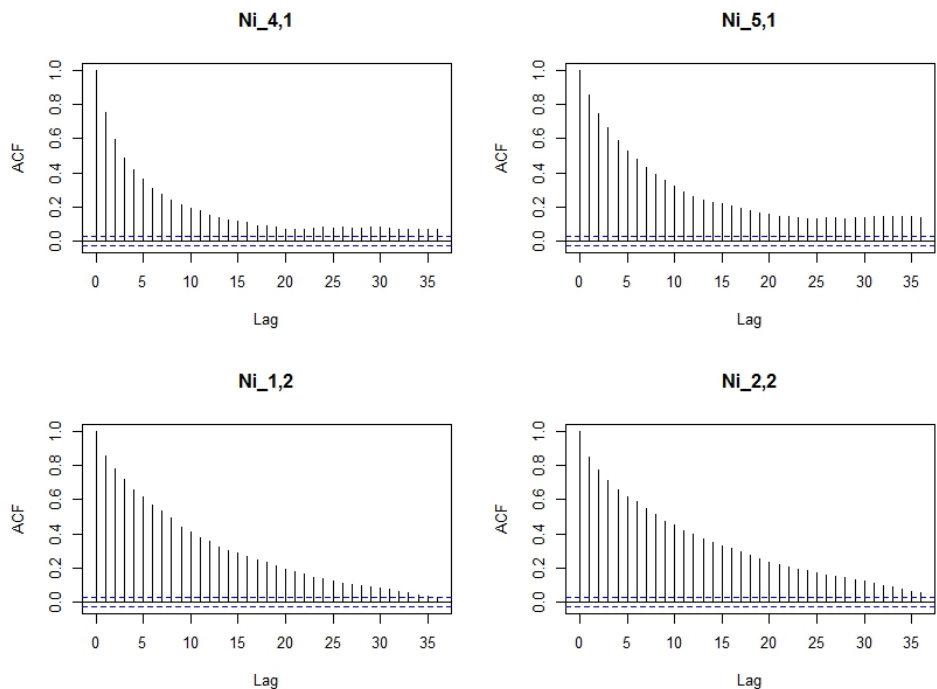
(a)  $\mu_{12}$ ,  $\mu_{22}$ ,  $\mu_{32}$  e  $\mu_{42}$



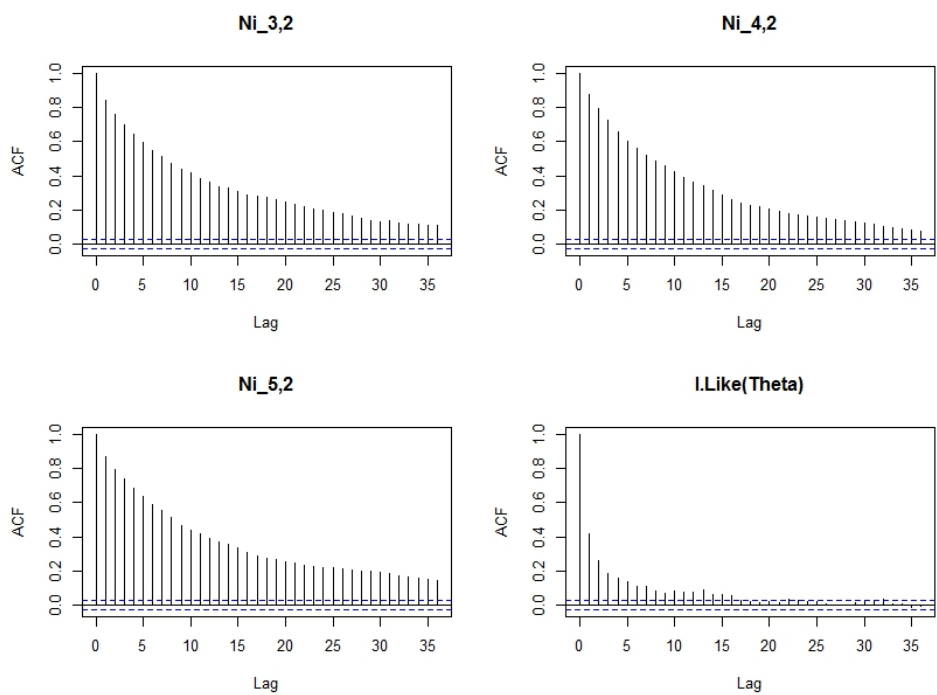
(b)  $\mu_{52}$ ,  $v_{11}$ ,  $v_{21}$  e  $v_{31}$

Fonte: Elaborada pelo autor.

Figura 108 – Auto-correlação dos resultados do MH aplicados no desempenho na prova de redação para  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$ ,  $v_{22}$ ,  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança



(a)  $v_{41}$ ,  $v_{51}$ ,  $v_{12}$  e  $v_{22}$



(b)  $v_{32}$ ,  $v_{42}$ ,  $v_{52}$  e log-verossimilhança

Fonte: Elaborada pelo autor.



## COMPETÊNCIAS AVALIADAS NAS PROVAS DO ENEM

### E.1 Competências avaliadas no Enem para prova de linguagens e suas tecnologias.

Quadro 4 – Competências da área 1 - Linguagens, Códigos e suas Tecnologias.

Item	Habilidade mensurada
H1	Identificar as diferentes linguagens e seus recursos expressivos como elementos de caracterização dos sistemas de comunicação;
H2	Recorrer aos conhecimentos sobre as linguagens dos sistemas de comunicação e informação para resolver problemas sociais;
H3	Relacionar informações geradas nos sistemas de comunicação e informação, considerando a função social desses sistemas;
H4	Reconhecer posições críticas aos usos sociais que são feitos das linguagens e dos sistemas de comunicação e informação.

Fonte: Elaborada pelo autor.

### E.2 Competências avaliadas no Enem para prova de Redação.

Quadro 5 – Níveis de desempenho que são utilizados para avaliar a competência II nas redações do Enem 2023.

Pontuação	Competência avaliada
200 pontos	Desenvolve o tema por meio de argumentação consistente, a partir de um repertório sociocultural produtivo, e apresenta excelente domínio do texto dissertativo-argumentativo;
160 pontos	Desenvolve o tema por meio de argumentação consistente e apresenta bom domínio do texto dissertativo-argumentativo, com proposição, argumentação e conclusão;
120 pontos	Desenvolve o tema por meio de argumentação previsível e apresenta domínio mediano do texto dissertativo-argumentativo, com proposição, argumentação e conclusão;
80 pontos	Desenvolve o tema recorrendo à cópia de trechos dos textos motivadores ou apresenta domínio insuficiente do texto dissertativo-argumentativo, não atendendo à estrutura com proposição, argumentação e conclusão;
40 pontos	Apresenta o assunto, tangenciando o tema, ou demonstra domínio precário do texto dissertativo-argumentativo, com traços constantes de outros tipos textuais;
0 ponto	Fuga ao tema/não atendimento à estrutura dissertativo-argumentativa. Nestes casos, a redação recebe nota zero e é anulada.

Fonte: Adaptada de [Inep \(2023\)](#).

Quadro 6 – Níveis de desempenho que serão utilizados para avaliar a Competência III nas redações do Enem 2023.

Pontuação	Competência avaliada
200 pontos	Apresenta informações, fatos e opiniões relacionados ao tema proposto, de forma consistente e organizada, configurando autoria, em defesa de um ponto de vista;
160 pontos	Apresenta informações, fatos e opiniões relacionados ao tema, de forma organizada, com indícios de autoria, em defesa de um ponto de vista;
120 pontos	Apresenta informações, fatos e opiniões relacionados ao tema, limitados aos argumentos dos textos motivadores e pouco organizados, em defesa de um ponto de vista;
80 pontos	Apresenta informações, fatos e opiniões relacionados ao tema, mas desorganizados ou contraditórios e limitados aos argumentos dos textos motivadores, em defesa de um ponto de vista;
40 pontos	Apresenta informações, fatos e opiniões pouco relacionados ao tema ou incoerentes e sem defesa de um ponto de vista;
0 ponto	Apresenta informações, fatos e opiniões não relacionados ao tema e sem defesa de um ponto de vista.

Fonte: Adaptada de [Inep \(2023\)](#).

Quadro 7 – Níveis de desempenho que serão utilizados para avaliar a Competência IV nas redações do Enem 2023.

Pontuação	Competência avaliada
200 pontos	Articula bem as partes do texto e apresenta repertório diversificado de recursos coesivos;
160 pontos	Articula as partes do texto, com poucas inadequações, e apresenta repertório diversificado de recursos coesivos.;
120 pontos	Articula as partes do texto, de forma mediana, com inadequações, e apresenta repertório pouco diversificado de recursos coesivos;
80 pontos	Articula as partes do texto, de forma insuficiente, com muitas inadequações, e apresenta repertório limitado de recursos coesivos;
40 pontos	Articula as partes do texto de forma precária;
0 ponto	Não articula as informações.

Fonte: Adaptada de [Inep \(2023\)](#).

Quadro 8 – Níveis de desempenho que serão utilizados para avaliar a Competência V nas redações do Enem 2023.

Pontuação	Competência avaliada
200 pontos	Elabora muito bem proposta de intervenção, detalhada, relacionada ao tema e articulada à discussão desenvolvida no texto;
160 pontos	Elabora bem proposta de intervenção relacionada ao tema e articulada à discussão desenvolvida no texto;
120 pontos	Elabora, de forma mediana, proposta de intervenção relacionada ao tema e articulada à discussão desenvolvida no texto;
80 pontos	Elabora, de forma insuficiente, proposta de intervenção relacionada ao tema, ou não articulada com a discussão desenvolvida no texto;
40 pontos	Apresenta proposta de intervenção vaga, precária ou relacionada apenas ao assunto;
0 ponto	Não apresenta proposta de intervenção ou apresenta proposta não relacionada ao tema ou ao assunto.

Fonte: Adaptada de [Inep \(2023\)](#).



## DESCRIÇÃO E QUANTIDADE DE MUNICÍPIOS SELECIONADOS PARA APLICAÇÃO EM DADOS REAIS

---



---

O [Quadro 9](#) apresenta a quantidade de municípios selecionados para a aplicação na metodologia segundo a Unidade da Federação,

Quadro 9 – Quantidade de municípios selecionados por Unidade da Federação.

UF	Qtde municípios	UF	Qtde municípios	UF	Qtde municípios
MG	5	RS	3	ES	1
MA	4	SC	3	RJ	1
CE	4	PR	3	AC	1
BA	4	PA	2	AM	1
PE	4	RO	1	TO	1
SP	4	RN	1	-	-

Fonte: Elaborada pelo autor.

O [Quadro 10](#) apresenta os nomes dos municípios selecionados para a aplicação com dados reais na metodologia proposta.

Quadro 10 – Municípios selecionados por Unidade da Federação.

UF	Nome do município	UF	Nome do município
AC	Cruzeiro do Sul	PE	Araripina
AM	Coari	PE	Carpina
BA	Dias d'Ávila	PE	Gravatá
BA	Guanambi	PE	Serra Talhada
BA	Jacobina	PR	Cianorte
BA	Serrinha	PR	Paranavaí
CE	Aquiraz	PR	Pato Branco
CE	Pacatuba	RJ	Seropédica
CE	Quixadá	RN	Macaíba
CE	Quixeramobim	RO	Cacoal
ES	Viana	RS	Cachoeira do Sul
MA	Barra do Corda	RS	Esteio
MA	Chapadinha	RS	Sapiranga
MA	Pinheiro	SC	Camboriú
MA	Santa Inês	SC	Caçador
MG	Alfenas	SC	São Bento do Sul
MG	Curvelo	SP	Campo Limpo Paulista
MG	João Monlevade	SP	Cruzeiro
MG	Três Corações	SP	Ibiúna
MG	Unai	SP	Lorena
PA	Moju	TO	Gurupi
PA	Redenção	—	—

Fonte: Elaborada pelo autor.

