

Milton Pifano Soares Ferreira

Inferência em Grafos Aleatórios  
Exponenciais através do ABC

**Belo Horizonte**  
**2016**



Milton Pifano Soares Ferreira

# Inferência em Grafos Aleatórios Exponenciais através do ABC

Dissertação apresentada ao Departamento de Estatística, do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, para a obtenção de Título de Mestre em Estatística.

Orientador: Marcos Oliveira Prates  
Coorientador: Adrian Pablo H. Luna

**Belo Horizonte**  
**2016**

Soares Ferreira, Milton P.

Inferência em Grafos Aleatórios Exponenciais através do ABC

90 páginas

Dissertação (Mestrado) - Instituto de Ciências Exatas da Universidade de Federal de Minas Gerais. Departamento de Estatística.

1. Ergm, Bergm
2. Erdos-Renyi
3. Grafos Aleatórios Exponenciais
4. ABC

I. Universidade de Federal de Minas Gerais. Instituto de Ciências Exatas. Departamento de Estatística.

## Comissão Julgadora:

---

Prof. Dr.  
Remy de Paiva Sanchis

---

Profa. Dra.  
Magda Carvalho Pires

---

Prof. Dr.  
Marcos Oliveira Prates

---

Prof. Dr.  
Adrian Pablo Hinojosa Luna

*Dedico esse trabalho à minha esposa, Aglaia, e ao meu filho, Rodrigo, pelo apoio e paciência que tanto demonstraram por todo o período de estudo.*

## Agradecimentos

A realização deste trabalho em muito se deve à colaboração e apoio de diversas pessoas. Cronologicamente agradeço, primeiramente, ao professor do Departamento de Estatística da UFMG, Gregório Saravia Atuncar, que muito me apoiou, incentivou e ajudou para que eu tivesse condições de cursar o Mestrado em Estatística. Sem esse apoio provavelmente eu não teria chegado até aqui.

Em segundo lugar, agradeço ao professor Osvaldo Farhat, do Departamento de Ciência da Computação da UFMG, pela oportunidade de participar do projeto InfoSAS. Sem a participação nesse projeto não teria sido possível completar o Mestrado em Estatística, pois esse projeto viabilizou trabalho e estudo no mesmo local físico além da flexibilidade de dedicação de acordo com as necessidades de cada atividade.

Em terceiro lugar, e, claro, não menos importante, agradeço ao meu orientador e ao meu co-orientador, que realmente me deram a orientação e os ensinamentos necessários para que eu pudesse desenvolver a pesquisa.

Finalmente, agradeço ao Departamento de Estatística da UFMG por trabalhar para oferecer uma infra-estrutura de ensino de alta qualidade para que possamos desenvolver pesquisas e aprimorar nosso conhecimento, contribuindo para o crescimento individual e do Brasil de modo geral.

## *Resumo*

Modelos de Grafos Aleatórios Exponenciais (ERGMs) são modelos estatísticos paramétricos para a distribuição de probabilidade de estruturas de rede a partir da análise de configurações baseadas na presença ou ausência de arestas, tais como k-estrelas e triângulos. Essas configurações são ponderadas por parâmetros do modelo.

ERGMs são principalmente utilizados para modelar redes sociais, mas também são utilizados na física e na biologia. Eles são fortemente embasados pelos fenômenos que os descrevem, pois a sua utilização requer que o pesquisador leve em consideração as razões teóricas (sociais, biológicas, etc) que levam à formação das arestas.

O foco deste trabalho é apresentar o estudo comparativo de dois métodos de estimação de parâmetros para os modelos ERGMs quando o número de nós cresce assintoticamente.

Os métodos considerados foram o Bayesiano, proposto por [Caimo and Friel 2011](#), baseado no algoritmo Markov Chain Monte Carlo (MCMC), e o método *Approximate Bayesian Computation* (ABC) proposto por [Del Moral et al. 2012](#), [Beaumont et al. 2009](#), [Drovandi e Pettitt 2011](#) e [Lenormand et al 2012](#). Os resultados mostram que o método ABC, especialmente o proposto por Lenormand, foi bem superior tanto no quesito **qualidade do ajuste** quanto no de **performance**.

Apesar do estudo em questão estar restrito a um número limitado de parâmetros, a partir dos experimentos efetuados temos fortes convicções de que o método ABC proposto por [Lenormand et al 2012](#) é consistentemente melhor que o método Bayesiano proposto por [Caimo e Friel 2011](#).

**Palavras-chave:** ABC, MCMC, Grafos Aleatórios Exponenciais, ERGMs, Bergm

## *Abstract*

Exponential random graph models are parametric statistical methods for probability distributions of network structures through the analyses of configurations based on the presence (or absence) of edges, such as k-stars and triangles. These configurations are weighted by model parameters.

These models are principled statistical approaches to model social networks, but they are also applied in physics and biology. They are theory driven in such a way that their use requires the researcher to consider the complexity of why edges (social, biological or whatever) are formed.

The focus of this work is to present a comparative study between two parameter estimation methods for the ERGMs models not taking into account the theories behind the edges formation.

The methods considered were the one proposed by [Caimo and Friel 2011](#), a Bayesian method based on Markov Chain Monte Carlo (MCMC), and the Approximate Bayesian Computation (ABC) method proposed by [Del Moral et al. 2012](#), [Beaumont et al. 2009](#), [Drovandi and Pettitt 2011](#) e

[Lenormand et al 2012](#). The results show that the ABC method, specially the one proposed by Lenormand, surpassed the Bayesian one considering both the goodness-of-fit and performance.

Despite the fact that this study was restricted to a limited number of parameters, based on the experiments done so far, we are strongly convinced that the ABC approach, proposed by [Lenormand et al 2012](#), is consistently better than the Bayesian method proposed by [Caimo and Friel 2011](#).

**Keywords:** ABC, MCMC, Exponential Random Graph Models, ERGM, Bergm



# Lista de Figuras

1.1	Grafo de Markov . . . . .	5
1.2	Grafo de Dependência . . . . .	6
1.3	Exemplo de triângulo e 6-estrelas . . . . .	7
1.4	Transição de fase (Diaconis e Chatterjee 2013). . . . .	13
1.5	Linhas de <i>contour</i> para a figura 1.4 (Diaconis e Chatterjee 2013). . . . .	14
2.1	Gráfico ilustrativo da movimentação referente ao algoritmo <i>exchange</i> adaptativo. . . . .	19
4.1	Grafos aleatórios Bernoulli <b>50 nós</b> : Beaumont, Bergm e Delmoral . . . . .	28
4.2	Grafos aleatórios Bernoulli <b>50 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	29
4.3	Grafos aleatórios Bernoulli <b>50 nós</b> : Rejeição . . . . .	30
4.4	Grafos aleatórios Bernoulli <b>2.250 nós</b> : Beaumont, Bergm e Delmoral . . . . .	31
4.5	Grafos aleatórios Bernoulli <b>2.250 nós</b> : Drovandi e Lenorman . . . . .	32
4.6	Tempo de processamento com todos os métodos. . . . .	33
4.7	Tempo de processamento sem o método ABC-Rejeição, pois ele apresentou tempos de processamento muito elevados, próximos de 24h, com 1.250 vértices, alterando a escala do gráfico e dificultando a visualização comparativa entre os métodos Bayesiano e ABC. . . . .	33
A.1	Grafos aleatórios Bernoulli <b>50 nós</b> : Beaumont, Bergm e Delmoral . . . . .	48
A.2	Grafos aleatórios Bernoulli <b>50 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	49
A.3	Grafos aleatórios Bernoulli <b>50 nós</b> : Rejeição . . . . .	50
A.4	Grafos aleatórios Bernoulli <b>100 nós</b> : Beaumont, Bergm e Delmoral . . . . .	51
A.5	Grafos aleatórios Bernoulli <b>100 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	52
A.6	Grafos aleatórios Bernoulli <b>100 nós</b> : Rejeição . . . . .	53
A.7	Grafos aleatórios Bernoulli <b>200 nós</b> : Beaumont, Bergm e Delmoral . . . . .	54

A.8	Grafos aleatórios Bernoulli <b>200 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	55
A.9	Grafos aleatórios Bernoulli <b>200 nós</b> : Rejeição . . . . .	56
A.10	Grafos aleatórios Bernoulli <b>500 nós</b> : Beaumont, Bergm e Delmoral . . . . .	57
A.11	Grafos aleatórios Bernoulli <b>500 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	58
A.12	Grafos aleatórios Bernoulli <b>500 nós</b> : Rejeição . . . . .	59
A.13	Grafos aleatórios Bernoulli <b>750 nós</b> : Beaumont, Bergm e Delmoral . . . . .	60
A.14	Grafos aleatórios Bernoulli <b>750 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	61
A.15	Grafos aleatórios Bernoulli <b>750 nós</b> : Rejeição . . . . .	62
A.16	Grafos aleatórios Bernoulli <b>1.000 nós</b> : Beaumont, Bergm e Delmoral . . . . .	63
A.17	Grafos aleatórios Bernoulli <b>1.000 nós</b> : Drovandi, Lenorman e Marjoram . . . . .	64
A.18	Grafos aleatórios Bernoulli <b>1.000 nós</b> : Rejeição . . . . .	65
A.19	Grafos aleatórios Bernoulli <b>1.250 nós</b> : Beaumont, Bergm e Delmoral . . . . .	66
A.20	Grafos aleatórios Bernoulli <b>1.250 nós</b> : Drovandi, Lenorman e Rejeição . . . . .	67
A.21	Grafos aleatórios Bernoulli <b>1.500 nós</b> : Beaumont, Bergm e Delmoral . . . . .	68
A.22	Grafos aleatórios Bernoulli <b>1.500 nós</b> : Drovandi e Lenorman . . . . .	69
A.23	Grafos aleatórios Bernoulli <b>1.750 nós</b> : Beaumont, Bergm e Delmoral . . . . .	70
A.24	Grafos aleatórios Bernoulli <b>1.750 nós</b> : Drovandi e Lenorman . . . . .	71
A.25	Grafos aleatórios Bernoulli <b>2.000 nós</b> : Beaumont, Bergm e Delmoral . . . . .	72
A.26	Grafos aleatórios Bernoulli <b>2.000 nós</b> : Drovandi e Lenorman . . . . .	73
A.27	Grafos aleatórios Bernoulli <b>2.250 nós</b> : Beaumont, Bergm e Delmoral . . . . .	74
A.28	Grafos aleatórios Bernoulli <b>2.250 nós</b> : Drovandi e Lenorman . . . . .	75
B.1	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : <u><math>\theta_1 = 0,8</math> e <math>\theta_2 = 0,02</math></u> , <b>50</b> nós . . . . .	77
B.2	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : <u><math>\theta_1 = 0,8</math> e <math>\theta_2 = 0,02</math></u> , <b>50</b> nós . . . . .	78
B.3	Grafos aleatórios exponenciais método Bergm : <u><math>\theta_1 = 0,8</math> e <math>\theta_2 = 0,02</math></u> , <b>50</b> nós . . . . .	79
B.4	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : <u><math>\theta_1 = 0,2</math> e <math>\theta_2 = 0,2</math></u> , <b>50</b> nós . . . . .	80
B.5	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : <u><math>\theta_1 = 0,2</math> e <math>\theta_2 = 0,2</math></u> , <b>50</b> nós . . . . .	81

B.6	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : $\theta_1 = 0,5$ e $\theta_2 = 0,35$ , <u>50</u> nós . . . . .	82
B.7	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : $\theta_1 = 0,5$ e $\theta_2 = 0,35$ , <u>50</u> nós . . . . .	83
B.8	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : $\theta_1 = 0,9$ e $\theta_2 = 0,02$ , <u>50</u> nós . . . . .	84
B.9	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : $\theta_1 = 0,9$ e $\theta_2 = 0,02$ , <u>50</u> nós . . . . .	85
B.10	Grafos aleatórios exponenciais método Bergm : $\theta_1 = 0,9$ e $\theta_2 = 0,02$ , <u>50</u> nós . . . . .	86
B.11	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : $\theta_1 = 0,6$ e $\theta_2 = 0,1$ , <u>50</u> nós . . . . .	87
B.12	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : $\theta_1 = 0,6$ e $\theta_2 = 0,1$ , <u>50</u> nós . . . . .	88
B.13	Grafos aleatórios exponenciais método ABC Lernormand, versão assintótica : $\theta_1 = 0,7$ e $\theta_2 = 0,3$ , <u>50</u> nós . . . . .	89
B.14	Grafos aleatórios exponenciais método ABC Lernormand, versão Bayesiana : $\theta_1 = 0,7$ e $\theta_2 = 0,3$ , <u>50</u> nós . . . . .	90

# Lista de Tabelas

4.1	Resultado grafo aleatório Bernoulli para <b>50</b> nós . . . . .	26
4.2	Resultado grafo aleatório Bernoulli para <b>2.250</b> nós . . . . .	26
4.3	Resultado grafo exponencial $\theta_1 = 0,8$ e $\theta_2 = 0,02$ : <b>50</b> nós . . . . .	36
4.4	Resultado grafo exponencial $\theta_1 = 0,2$ e $\theta_2 = 0,2$ : <b>50</b> nós . . . . .	36
4.5	Resultado grafo exponencial $\theta_1 = 0,5$ e $\theta_2 = 0,35$ : <b>50</b> nós . . . . .	36
4.6	Resultado grafo exponencial $\theta_1 = 0,9$ e $\theta_2 = 0,02$ : <b>50</b> nós . . . . .	37
4.7	Resultado grafo exponencial $\theta_1 = 0,6$ e $\theta_2 = 0,1$ : <b>50</b> nós . . . . .	37
4.8	Resultado grafo exponencial $\theta_1 = 0,7$ e $\theta_2 = 0,3$ : <b>50</b> nós . . . . .	37
A.1	Resultado grafo aleatório Bernoulli para <b>50</b> nós . . . . .	45
A.2	Resultado grafo aleatório Bernoulli para <b>100</b> nós . . . . .	45
A.3	Resultado grafo aleatório Bernoulli para <b>200</b> nós . . . . .	46
A.4	Resultado grafo aleatório Bernoulli para <b>500</b> nós . . . . .	46
A.5	Resultado grafo aleatório Bernoulli para <b>750</b> nós . . . . .	46
A.6	Resultado grafo aleatório Bernoulli para <b>1.000</b> nós . . . . .	46
A.7	Resultado grafo aleatório Bernoulli para <b>1.250</b> nós . . . . .	46
A.8	Resultado grafo aleatório Bernoulli para <b>1.500</b> nós . . . . .	47
A.9	Resultado grafo aleatório Bernoulli para <b>1.750</b> nós . . . . .	47
A.10	Resultado grafo aleatório Bernoulli para <b>2.000</b> nós . . . . .	47
A.11	Resultado grafo aleatório Bernoulli para <b>2.250</b> nós . . . . .	47

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Modelos de Grafos Aleatórios Exponenciais (ERGM)	4
1.2	Método de estimação ABC - Approximate Bayesian Computation	8
1.3	Teoria assintótica para Grafos Aleatórios Exponenciais	13
<b>2</b>	<b>Inferência para Modelos ERGM</b>	<b>15</b>
2.1	Inferência Clássica	15
2.2	Inferência Bayesiana	16
<b>3</b>	<b>Metodologia ABC para ERGM</b>	<b>21</b>
<b>4</b>	<b>Estudo de Simulação</b>	<b>23</b>
4.1	Grafos Bernoulli : Caso Erdős - Renyi	23
4.2	Grafos Exponenciais : Arestas e Triângulos	34
<b>5</b>	<b>Conclusões e Trabalhos Futuros</b>	<b>39</b>
	<b>Referências Bibliográficas</b>	<b>41</b>
<b>A</b>	<b>Resultados do Estudo de Simulação para Grafos Bernoulli : Caso Erdős-Renyi</b>	<b>45</b>
<b>B</b>	<b>Gráficos do Estudo de Simulação para Grafos Exponenciais : Arestas e Triângulos</b>	<b>76</b>



# Capítulo 1

## Introdução

O objetivo desse trabalho é comparar dois métodos de inferência para modelos de grafos aleatórios exponenciais para grafos com grande número de nós, a saber : o método Bayesiano proposto por [Caimo e Friel 2011](#) e a nossa proposta de utilização do método *Approximate Bayesian Computation*(ABC) em conjunto com a teoria assintótica de [Diaconis e Chatterjee 2013](#). Apresentamos a seguir os conceitos que são os pilares dessa dissertação : "Grafos", "Modelos de Grafos Aleatórios Exponenciais (ERGM)", "Método de estimação ABC" e "Teoria assintótica para Grafos Aleatórios Exponenciais".

Modelos de grafos aleatórios exponenciais são os modelos estatísticos mais utilizados no contexto de redes sociais, além disso também são utilizados na física, biologia, dentre outros. Para apresentação das idéias, introduziremos o modelo no contexto de redes sociais, pois a grande maioria dos artigos descreve esse modelo nesse contexto, bem como o livro de [Koskinen et al 2012](#), que é uma das melhores fontes para o entendimento da análise estatística de redes.

Uma rede pode ser representada de diversas maneiras. Uma das representações mais úteis é como um *grafo*, que consiste de *nós* conectados por *arestas* (dependendo do contexto, usaremos o termo *rede* como sinônimo de *grafo*; por exemplo, *redes sociais*, ao invés de, *grafos sociais*).

A teoria de grafos tem sido útil na análise de redes sociais por diversos motivos. Primeiramente, a teoria de grafos provê um vocabulário que pode ser utilizado para rotular e representar muitas propriedades de estruturas sociais. Este vocabulário também disponibiliza um conjunto de conceitos básicos que permite nos referir a essas propriedades com certa precisão. Em segundo lugar, a teoria de grafos disponibiliza operações matemáticas e idéias, com as quais muitas das propriedades podem ser quantificadas e medidas. Finalmente, dado o vocabulário e o arcabouço matemático, a teoria de grafos possibilita provar teoremas sobre grafos, e conseqüentemente, sobre representações da estrutura social ([Wasserman e Faust 1994](#)).

Além de sua utilidade como um sistema matemático, a teoria de grafos disponibiliza uma representação de uma rede social como um modelo de um sistema social composto de um conjunto de atores e ligações entre eles. Por *modelo* entende-se como uma representação simplificada de uma situação que contém alguns, mas não todos, os elementos de uma situação que ele representa. Quando um grafo é utilizado como um modelo de uma rede social, *nós/vértices* representam os atores e as *ligações/arestas* conectando os nós representam os relacionamentos entre eles ([Wasserman e Faust 1994](#)).

Sendo assim, temos que um *grafo* é uma representação matemática de uma rede binária. O número de nós de um grafo é denotado por  $n$ . Uma variável aleatória  $Y_{ij}$  indica se há uma ligação entre os

nós  $i$  e  $j$  ( $Y_{ij} = 1$ ) ou não ( $Y_{ij} = 0$ ). Auto-ligação de nós não é permitida nesta dissertação, ou seja, ( $Y_{ii} = 0$ ) para todo  $i$ .

Um grafo aleatório é representado por sua matriz de adjacência  $Y$  com elementos  $Y_{ij}$ . Grafos são, por definição, não-direcionados (ou seja,  $Y_{ij} = Y_{ji}$  para todo  $i, j$ ).

Continuando a notação representativa de *grafos*, a substituição de um índice  $i$  ou  $j$  por um sinal de  $+$  denotará uma soma em relação ao índice em questão, ou seja,  $Y_{i+} = \sum_j Y_{ij}$  e  $Y_{+i} = \sum_j Y_{ji}$ ; variáveis aleatórias são denotadas por letras maiúsculas ( $Y$ ); e uma observação da variável aleatória por letras minúsculas ( $y$ ).

Redes são estruturas relacionais, e redes sociais representam estruturas de ligações diádicas entre atores (nós/vértices) sociais, tais como : amizade entre indivíduos, parceria entre empresas, comércio entre países, etc. A natureza das redes conduz a uma dependência entre os atores (nós/vértices), como também uma dependência entre as ligações (arestas). Mas modelos estatísticos são baseados em suposições de independência, e a natureza complexa de dependência nas redes tem retardado o desenvolvimento desses modelos para estruturas de redes (Snijders 2011).

Redes sociais são caracterizadas por inúmeros tipos de dependência que tem sido encontradas tanto empírica, quanto teoricamente. O entendimento dessas dependências é importante, pois os diferentes modelos estatísticos existentes na literatura divergem justamente na forma de representá-las. A seguir descrevemos apenas algumas dessas dependências, pois existem dezenas delas (Snijders 2011) :

- **Reciprocidade** de ligações direcionadas ocorre quando há um tratamento recíproco entre dois atores (nós) , ou seja, ocorre uma dependência entre  $Y_{ij}$  e  $Y_{ji}$ .
- **Homofilia**, tendência de atores (nós) de se relacionarem com atores com característica similares (cor, raça, sexo, curso, etc).
- **Transitividade ou clusterização** de ligações (arestas) são caracterizadas quando ocorre situações do tipo "amigos de meus amigos, são meus amigos". Se houver uma tendência de ocorrer transitividade, a existência de duas ligações  $Y_{ij} = Y_{jh} = 1$  irá conduzir a um aumento da probabilidade da ligação  $Y_{ih}$  ser igual a 1 ( $Y_{ih} = 1$ ), ou seja, a ocorrência de um triângulo. A concatenação de triângulos pode criar grupos conectados maiores.
- **Grau**, informação que caracteriza atores altamente conectados daqueles pouco conectados, gerando modelos que são centrados em nós (Freeman 1979).

Conforme Snijders 2011, os três modelos estatísticos mais abrangentes que contemplam dependências existentes em grafos são :

- **Incorporação de estruturas de rede por meio de covariáveis**. Essa abordagem utiliza um modelo com resíduos independentes e tenta representar as dependências de rede por meio de variáveis explicativas. Para que isso seja plausível, isso pode ser considerado principalmente em configurações longitudinais onde observações anteriores da rede podem ser usadas para produzir covariáveis (Gulati and Gargiulo 1999).
- **Controle de estruturas da rede, ou seja, análise de uma estrutura da rede enquanto mantém as demais estruturas constantes**. Essa abordagem estabelece o controle de certas dependências existentes na rede sem explicitamente modelá-las. O exemplo mais conhecido dessa



abordagem é um procedimento de permutação, onde nós da rede são permutados (pode-se dizer, também, que linhas e colunas da matriz de adjacência são permutadas, simultaneamente, de forma que a estrutura da rede é mantida intacta). Esta abordagem é chamada de *Quadratic Assignment Procedure (QAP)* proposta por [Krackhardt 1987](#). Nessa abordagem ainda temos os métodos : *Multiple Regression QAP* ou *MRQAP* proposto por [Krackhardt 1988](#) e [Dekker et al 2007](#); estimador sanduíche de variância desenvolvido por [Lindgren 2010](#), que usa a idéia de estimador consistente heteroscedasticamente para uma matriz de covariância, elaborado por [White 1980](#); e a abordagem de condicionar a estatísticas que expressam dependências da rede, onde se assume que a distribuição é uniforme, daí o nome **modelos condicionamente uniformes**.

- **Modelagem de estruturas de rede estocasticamente.** Esse é o modelo em que se enquadra os modelos de grafos aleatórios exponenciais (ERGMs). Nesta abordagem, explicitamente se modela as dependências estruturais entre as ligações/arestas, que são consideradas variáveis aleatórias, onde hipóteses são formuladas em termos de parâmetros que representam dependências mais complexas, tal como a **transitividade**, que é uma dependência que envolve três variáveis ao mesmo tempo.

Conforme descrito acima, a terceira abordagem é onde se enquadra os modelos ERGMs, pois nela as dependências da rede são representadas, explicitamente, como um modelo estocástico.

Antes de apresentarmos os modelos ERGMs é importante destacarmos o uso da probabilidade e da estatística para modelar redes. Muitos estudos sobre redes consideram apenas uma única rede e, nesse sentido, temos  $N = 1$ , o que leva à discussão de porque métodos estatísticos são aplicáveis a esse tipo de dado. Importante destacar que fazer inferência estatística para  $N = 1$  não é incomum. Situações similares ocorrem, por exemplo, em análises econômicas para longas séries temporais, análises espaciais e espaço-temporais. Como pode um pesquisador, de posse de apenas uma rede observada, inferir princípios organizacionais para ela? Por causa da relativa constância dos princípios de organização de uma rede, uma única rede observada captura o acúmulo de processos sociais, como um traço arqueológico ([Scott et al](#)). Princípios organizacionais estáveis resultarão em padrões de ligações de rede que podem ser observadas nos dados, mesmo quando os dados são de uma única instância no tempo. Esses padrões das ligações são, de fato, uma assinatura estrutural da rede e proveem evidências a partir das quais pode-se fazer inferência a respeito do processo social que gerou a rede ([Scott et al](#)).

Dessa forma, procura-se por padrões na rede. Não é suficiente verificar apenas um padrão de cada vez, pois as estruturas são construídas umas a partir de outras. Por exemplo, um triângulo contém 3 ligações de rede, então, segue-se que, se uma rede tem muitas ligações (arestas), mais triângulos deverão ser formados. A questão então é a seguinte : dada a densidade de uma rede, ocorrerá a existência de mais triângulos do que o esperado? Se sim, talvez deva-se considerar um processo social específico que conduza à geração de triângulos (por exemplo : transitividade). Se não, não há a necessidade de se definir um mecanismo de triangulação, pois apenas a densidade (ou seja, indivíduos em uma rede estão propensos a criar relações/ligações com qualquer outro) é uma característica suficiente e parcimoniosa.

Um importante aspecto relativo à inferência em redes é que, embora se tenha apenas um *sistema* no sentido de que, potencialmente, há uma interdependência entre todos os atores e ligações, existe um grande número de variáveis (atores e ligações) e, por isso, pode-se desenvolver inferência estatística porque suposições razoáveis de independência condicional podem ser consideradas ([Snijders 2011](#)).

Não se pode negar que tal conhecimento só pode ser sólido se houver um acúmulo de resultados por meio de replicações, isto é, estudos amplos do mesmo fenômeno ou processo com diferentes grupos ou contextos. Muitos progressos científicos podem ser alcançados quando dados de várias redes, que podem ser consideradas como replicações umas das outras, ou seja, várias escolas, empresas, municípios, etc; mas essa ainda é uma área da modelagem de rede muito pouco desenvolvida (Snijders 2011).

Grafos aleatórios exponenciais (ERGMs) - algumas vezes chamados de modelos  $p^*$  - são uma família de modelos estatísticos para redes que permite inferir sobre padrões proeminentes nos dados, dado a presença de outras estruturas de rede.

## 1.1 Modelos de Grafos Aleatórios Exponenciais (ERGM)

Modelos aleatórios para grafos foram inicialmente estudados por Erdos-Renyi 1959 que propuseram o modelo Bernoulli onde as arestas são consideradas independentes e são criadas aleatoriamente de acordo com uma propabilidade  $p$  fixa.

Holland and Leinhardt 1981 introduziram o modelo  $p_1$  onde os pares de nós, e não as arestas, são considerados independentes. Para grafos não direcionados, esse modelo é idêntico ao modelo Bernoulli.

Frank e Strauss 1986 propuseram os modelos de grafos aleatórios Markovianos onde cada par de aresta é, dado o restante do grafo, condicionalmente dependente quando eles tem um nó em comum.

Wasserman e Pattison 1996 propuseram uma generalização do modelo anterior, inicialmente batizado como o modelo do tipo  $p^*$ . A ênfase colocada na forma log-linear do modelo  $p^*$  facilitou criar extensões ao modelo básico de grafos de Markov. Essa proposta popularizou o modelo ERGM, uma vez que ele é comumente atribuído como sendo da classe de modelos  $p^*$ .

Mas é o trabalho de Frank e Strauss 1986 que é considerado a base do modelo ERGM atual.

De acordo com Snijders 2011, Frank e Strauss 1986 deram início a um tipo de modelagem onde a dependência entre as ligações eram representadas diretamente. Eles definiram dependências de Markov para distribuições em rede de forma análoga a distribuições de processos estocásticos, ou seja, condicionado a outras variáveis aleatórias, duas variáveis aleatórias são independentes desde que elas não estejam ligadas (onde uma ligação, no caso de um processo estocástico, seria definido como uma sequência no tempo). No caso de redes, a definição é mais precisamente definida como se segue. A matriz  $Y = (Y_{ij})$  de variáveis aleatórias, que podem ser consideradas como um grafo estocástico, é um *grafo de Markov* se para cada conjunto de 4 nós distintos  $i, j, h, k$ , as variáveis aleatórias  $Y_{ij}$  e  $Y_{hk}$  são condicionalmente independentes, dado todas as demais variáveis aleatórias em  $Y$ . Essa parece ser um tipo plausível de independência condicional, apropriada para redes sociais.

É importante, nesse momento, descrevermos as propriedades de independência condicional em grafos de Markov (ou modelos de grafos não-direcionados, ou campos aleatórios de Markov).

Grafos de Markov definem relacionamentos condicionalmente independentes por meio de separação de grafos simples como se segue : dado conjuntos de nós  $A, B$ , e  $C$ , dizemos que  $Y_A \perp_G Y_B | Y_C$  ( $A$  é condicionalmente independente de  $B$  no grafo  $G$ , dado  $C$ ), se e somente se,  $C$  separa  $A$  de  $B$  no grafo  $G$ . Isso significa que, ao se remover todos os nós de  $C$ , se não restar nenhum caminho conectando qualquer nó em  $A$  a qualquer nó em  $B$ , então a propriedade de independência condicional é

válida. Isso é chamado de **propriedade global de Markov** para modelos de grafos não-direcionados. Para ilustrar essa propriedade, na Figura 1.1 temos que  $\{1,2\} \perp \{6,7\} \mid \{3,4,5\}$ .

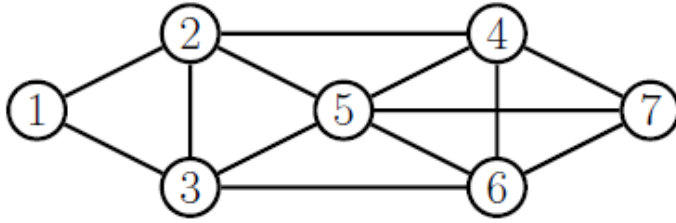


Figura 1.1: Grafo de Markov

Um nó é condicionalmente independente de todos os demais, dado seus vizinhos. Essa propriedade é chamada de **propriedade local de Markov**.

Finalmente, a partir da propriedade local de Markov, pode-se verificar que 2 nós são condicionalmente independentes, dado os demais, se não houver uma aresta entre eles. Essa propriedade é chamada de **propriedade de emparelhamento de Markov (pairwise Markov property)**.

Baseado nessas propriedades, podemos derivar as seguintes propriedades de independência condicional (dentre outras) a partir do grafo da Figura 1.1 :

- **Emparelhamento (Pairwise)** :  $1 \perp 7 \mid \text{demais}$
- **Local** :  $1 \perp \text{demais} \mid 2,3$
- **Global** :  $1,2 \perp 6,7 \mid 3,4,5$

Para provar o teorema que é a base dos modelos de grafos aleatórios exponenciais, [Frank e Strauss 1986](#) utilizaram grafos de dependência para representar estruturas de dependência em grafos aleatórios. Nesse contexto, as variáveis aleatórias são variáveis indicadoras para as arestas de um grafo aleatório  $G$ . Conforme já apresentamos, as variáveis indicadoras de arestas são elementos da matriz de adjacência  $Y = (Y_{ij})$ , de  $G$ .

O grafo de dependência  $D$ , de  $G$ , é um grafo não-aleatório que especifica a estrutura de dependência entre as  $M = \binom{n}{2}$  variáveis aleatórias  $Y_{ij} (1 \leq i \leq j \leq n)$ , onde  $n$  é o número de nós em  $G$ . Os nós de  $D$  são as possíveis arestas de  $G$ , e as arestas de  $D$  ligam os pares de arestas de  $G$  que são condicionalmente dependentes. Isso significa que,  $D$  terá uma aresta entre  $\{i,j\}$  e  $\{k,h\}$ , em  $M$ , se e somente se,  $Y_{ij}$  e  $Y_{kh}$  são condicionalmente dependentes das demais arestas em  $Y$ .

Um grafo  $G$  é um **grafo de Markov**, ou tem **dependência de Markov**, se  $D$  não possui arestas entre conjuntos disjuntos  $\{i,j\}$  e  $\{k,h\}$  em  $M$ . Isso significa que, arestas não-incidentes em  $G$  são condicionalmente independentes. Essa definição de **dependência de Markov** é consistente com a definição de vizinhos mais próximos de [Besag 1974](#) se os vizinhos mais próximos de  $\{i,j\} \in M$  são  $\{i,v\}$  e  $\{j,v\}$ , para todo  $v \in N = \{1, \dots, n\}$  diferente de  $i$  e  $j$  ([Frank e Strauss 1986](#)). A figura 1.2 ilustra o grafo de dependência  $D$  para um **grafo de Markov** de ordem  $n = 4$ .

De acordo com o teorema de Hammersley-Clifford ([Frank e Strauss 1986](#)), obtém-se a seguinte caracterização de uma função de distribuição de probabilidade para um grafo aleatório  $G$  :

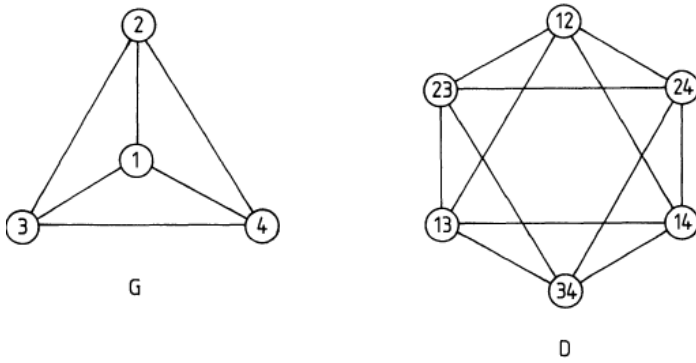


Figura 1.2: Grafo de Dependência

**Teorema HC:** Qualquer grafo não-direcionado  $G$  em  $N = \{1, \dots, n\}$  com estrutura de dependência  $D$  tem probabilidade

$$P(G) = \frac{\exp \sum_{A \subseteq G} \alpha_A}{c},$$

onde  $c$  é uma constante de normalização dada por

$$c = \sum_{G, P(G) > 0} \exp \sum_{A \subseteq G} \alpha_A$$

e  $\alpha_A$  é uma constante arbitrária se  $A$  é um clique de  $D$  e  $\alpha_A = 0$ , caso contrário.

Frank e Strauss 1986 provaram que a dependência de Markov para grafos não-direcionados e o requisito adicional de que a distribuição de probabilidade é invariante sob a permutação dos nós, é equivalente à expressar a distribuição de probabilidade de um grafo por

$$P(Y = y) = \frac{\exp(\theta L(y) + \sum_{k=2}^{n-1} \sigma_k S_k(y) + \tau T(y))}{k(\theta, \sigma, \tau)} \tag{1.1}$$

onde  $L(y) = \sum_{i < j} y_{ij}$  é o número de arestas,  $T = \sum_{i < j < h} y_{ij} y_{jh} y_{ih}$  é o número de triângulos, e  $S_k = \sum_i \sum_{j_1 < j_2 < \dots < j_k} y_{ij_1} y_{ij_2} \dots y_{ij_k}$  é o contador  $k$ -estrela (com  $S_1(y) = L(y)$ ) (Figura 1.3).

Os parâmetros estatísticos nesse modelo são  $\theta, \sigma_2, \dots, \sigma_{n-1}, \tau$ ; e  $k(\theta, \sigma, \tau)$  é a constante de normalização. O fato do logaritmo da probabilidade ser uma combinação linear de parâmetros, essa distribuição faz parte da família de distribuições exponenciais (Lehmann e Romano 2005), onde as estatísticas são estatísticas suficientes que são contagens de sub-grafos, ou seja, a frequência no grafo de pequenas configurações (nesse caso : arestas,  $k$ -estrelas, e triângulos).

Frank 1991 e Wasserman e Pattison 1996 generalizaram o grafo de Markov propondo que o expoente em (1.1) pode ser, a princípio, qualquer estatística, dessa forma permitindo qualquer tipo de dependência entre as variáveis aleatórias compostas pelas arestas, ou seja :

$$P(Y = y) = \frac{\exp(\sum_k \theta_k S_k(y))}{k(\theta)}, \tag{1.2}$$

onde  $S_k(y)$  pode ser qualquer estatística dependente da rede e de covariáveis observadas. Elas podem ser especificadas para refletir os questionamentos da pesquisa e para se obter um bom ajuste do modelo e dos dados. A princípio esse modelo pode representar qualquer distribuição no espaço de grafos que atribua uma probabilidade positiva para cada possível grafo, embora tal representa-

ção possa não ser necessariamente parcimoniosa ou tratável. Esse modelo foi chamado de  $p^*$  por Wasserman e Pattison 1996; sendo mais recentemente chamado de *Grafos Aleatórios Exponenciais (ERGM)*.



Figura 1.3: Exemplo de triângulo e 6-estrelas

Uma característica particularmente importante dos ERGMs é que arestas dependem umas das outras; assim, a existência de uma aresta pode afetar a existência de outras.

O modelo ERGM é relativamente aberto, uma vez que ele possibilita avaliar uma gama de teorias de acordo com o interesse do pesquisador. Uma vez que o pesquisador consiga traduzir a teoria em termos quantitativos e relacionados, o modelo ERGM provê uma estrutura com a qual hipóteses podem ser estatisticamente examinadas.

A base do modelo está na hipótese de que as conexões entre os atores, que são consideradas como variáveis aleatórias, se auto-organizam em estruturas pequenas chamadas de configurações, tais como : grau,  $k$ -estrela, triângulo, etc.

Estatísticas de rede são incorporadas aos modelos ERGMs como estatísticas suficientes associadas a parâmetros, que medem a importância de cada uma delas. Essas estatísticas são somatórios do número de configurações no grafo observado ou alguma função desses somatórios.

Conforme descrito anteriormente, a estrutura de um grafo aleatório não-direcionado pode ser representada por uma matriz de adjacência (aleatória)  $\mathbf{Y}$  da seguinte forma (computacionalmente é representada por outras estruturas de dados mais eficientes) :  $n$  nós e um conjunto de arestas  $\{Y_{ij} : i = 1, 2, \dots, n; j = 1, 2, \dots, n\}$ , onde  $Y_{ij} = 1$  se o par de nós  $(i, j)$  está conectado e  $Y_{ij} = 0$ , caso contrário. Auto-relacionamentos não são permitidos ( $Y_{ii} = 0$ ). Seja  $\Upsilon$  o conjunto de todas as possíveis matrizes de adjacência (todos os possíveis grafos não-direcionados) de  $n$  nós e  $\mathbf{y}$  uma realização de  $\mathbf{Y}$ . O modelo ERGM representa a distribuição conjunta de todas as arestas do grafo e, conseqüentemente, a distribuição de probabilidade de  $\mathbf{Y}$  com a seguinte função de verossimilhança :

$$\pi(\mathbf{y}|\theta) = \frac{q_{\theta}(\mathbf{y})}{z(\theta)} = \frac{\exp\{\theta^t s(\mathbf{y})\}}{z(\theta)} \quad (1.3)$$

Onde :

- $s(\mathbf{y})$  é um vetor conhecido de estatísticas suficientes (ex.: número de arestas, número de triângulos, etc).
- $\theta \in \Theta$  : é um vetor de parâmetros do modelo. Um número positivo para para  $\theta_i \in \Theta$  resulta em uma maior probabilidade de se observar a estatística  $s(\mathbf{y})$  nos grafos.
- $z(\theta) = \sum_{\mathbf{y} \in \Upsilon} \exp\{\theta^t s(\mathbf{y})\}$  a constante de normalização.

O modelo ERGM apresenta algumas dificuldades que requerem estratégias específicas para se permitir a inferência dos parâmetros. As principais dificuldades são :

- Cálculo da constante de normalização ( $z(\theta)$ ):

A constante de normalização é função dos parâmetros e consiste na soma de  $2^{\binom{n}{2}}$  possíveis grafos não-direcionados. Ex.:  $n=34$  nós  $\Rightarrow 2^{\binom{34}{2}}=2^{561}=7.5479 \times 10^{168}$ .

Pode-se ver que para um grafo pequeno (poucos nós) seu cálculo é computacionalmente intenso. Se considerarmos redes reais de milhares e milhões de nós esse cálculo torna-se impraticável (veja o *web site* <http://snap.stanford.edu/snap/index.html> da plataforma SNAP - *Stanford Network Analysis Platform*).

- Degeneração:

A degeneração é um dos importantes aspectos em modelos de grafos aleatórios

(Caimo e Friel 2011). Refere-se a um modelo probabilístico que coloca muita massa de probabilidade em um número pequeno de grafos, como por exemplo, grafos vazios ou grafos completos. Quando o modelo é quase degenerado, métodos baseados na estimação por máxima verossimilhança (MC-MLE: *Monte Carlo Maximum Likelihood Estimation*) podem falhar e retornar uma estimativa para  $\theta$  que tem pouca probabilidade de se gerar grafos semelhantes ao grafo observado.

Isso pode ocorrer porque a convergência do algoritmo pode ser afetada por valores degenerados de parâmetros que durante o processo de simulação da rede podem gerar grafos completos ou vazios (Handcock 2003).

Novas especificações propostas por Snijders et al. 2006 podem mitigar a degeneração e prover um razoável ajuste dos dados.

## 1.2 Método de estimação ABC - Approximate Bayesian Computation

Conforme apresentado, o modelo ERGM pertence à classe de modelos com constante de normalização intratável, o que requer a utilização de métodos específicos para a inferência de seus parâmetros. O método *Approximate Bayesian Computation* (ABC) é um dos métodos muito utilizado nessas situações.

ABC é uma família de técnicas computacionais em estatística Bayesiana. Essas técnicas permitem ajustar um modelo aos dados sem se basear no cálculo da verossimilhança do modelo (Beaumont et al. 2002) dependendo, ao contrário, de simulações dos dados a partir do modelo. Se o modelo produz a observação  $x \sim f(x|\theta)$  e  $\pi(\theta)$  é a distribuição a priori do parâmetro  $\theta$ , o algoritmo ABC consiste em simular conjuntamente  $\theta_i \sim \pi(\theta_i)$  e  $x \sim f(x|\theta_i)$  e em aceitar o  $\theta_i$  simulado, se e somente se,  $\rho(x, y) < \epsilon$ , onde  $\rho$  é uma medida de distância entre o dado observado  $y$  e o simulado  $x$ ;  $\epsilon > 0$  é chamado de nível de tolerância. O nível de tolerância representa o nível de acurácia na aproximação da verossimilhança. De fato, os  $\{\theta_i\}_{1 \leq i \leq N}$  gerados tem distribuição de probabilidade proporcional a  $\pi(\theta)P_\theta(\rho(x, y) < \epsilon)$ , onde  $P_\theta(z)$  representa a distribuição de probabilidade de  $z$  dado o parâmetro  $\theta$ . Essa densidade é uma aproximação da distribuição a posteriori do modelo  $\pi(\theta)P_\theta(f(x|\theta) = y)$ . Técnicas ABC requerem que o modelo seja simulado um grande número de vezes o que o torna um esquema computacionalmente intenso.

Algumas melhorias ao esquema ABC original tem sido propostas para agilizar seu processamento. Elas incluem o uso de regressões locais para melhorar a inferência do parâmetro (Beaumont et al. 2002; Blum e Francois 2010); a incorporação de passo MCMC para melhor explorar o espaço paramétrico

(Marjoram et al. 2003); e o sequenciamento de distribuições a posteriori, inspirado pelos métodos sequenciais de Monte Carlo (Sisson et al. 2007; Toni et al. 2009; Beaumont et al. 2009). Essa última classe de métodos consiste em se processar as simulações do ABC em vários passos usando as simulações do passo anterior para construir uma proposta melhor do parâmetro para o próximo passo. Segundo Lenormand et al 2012 essa estratégia evita as áreas de baixa probabilidade no espaço paramétrico, focando o esforço computacional nas áreas de grande massa de probabilidade. Beaumont et al. 2009 propuseram um algoritmo chamado de **Population Monte Carlo ABC (PMC-ABC)** que melhora o *mixing* dos parâmetros em relação à estratégia original **Sequencial Monte Carlo ABC (SMC-ABC)**.

De acordo com Lenormand et al 2012, o algoritmo **PMC-ABC** tem duas deficiências básicas. A primeira diz respeito à sequência dos níveis de tolerância  $\{\epsilon_1, \dots, \epsilon_T\}$  que tem que ser gerada pelo algoritmo ABC, o que, na prática, implica em simulações preliminares do modelo, que é um passo computacionalmente caro para modelos complexos. E a segunda deficiência está no fato de que o algoritmo não possui um critério específico para definir se ocorreu convergência.

A proposta de Lenormand et al 2012 consiste em uma modificação do algoritmo **PMC-ABC** que eles chamaram de **Adaptative Population Monte Carlo ABC (APMC-ABC)**. Nesse algoritmo a sequência de níveis de tolerância é determinada pelo próprio algoritmo conforme proposto por Drovandi e Pettitt 2011 e Del Moral et al. 2012, onde se constroi os níveis de tolerância *on-line* avaliando a distância  $\rho(x, y)$  das simulações geradas anteriormente; o problema de duplicação das partículas existente no algoritmo de Drovandi e Pettitt 2011 e Del Moral et al. 2012 é evitado; e o algoritmo estabelece um critério de convergência.

Resumindo, métodos ABC se baseiam na simulação de observações da distribuição conjunta dos dados e da distribuição a priori dos parâmetros, comparando os dados simulados com os dados observados. A probabilidade de se simular dados  $x$  com uma pequena distância de  $y$  diminui a medida que se aumenta a dimensionalidade dos dados. Como consequência ocorre uma diminuição substancial na eficiência do algoritmo ABC. Uma estratégia comum para minimizar esse problema é substituir  $y$  e  $x$  por estatísticas sumárias  $S(y)$  e  $S(x)$ , que são selecionadas para capturar informação relevante de  $y$  e  $x$ . Se essas estatísticas sumárias forem **estatísticas suficientes** com relação aos parâmetros do modelo ( $\theta$ ), o aumento da eficiência obtida não ocorre em erro, pois, por definição, suficiência implica que toda a informação de  $y$  a respeito de  $\theta$  é capturada por  $S(y)$  (Wilkinson 2008).

Para mostrar a intuição por trás dos métodos ABC, a seguir apresentamos o algoritmo *naive* de 3 técnicas, a saber : ABC Rejeição (Lenormand et al 2012), ABC MCMC (Marjoram et al. 2003) e Population Monte Carlo ABC (PMC-ABC) (Lenormand et al 2012).

**ABC Rejeição**

- Dado  $N$  números de partículas

- Para  $i = 1$  até  $N$  faça

Repita

Simule  $\theta^*$  a partir de  $\pi(\theta)$

Simule  $x^*$  a partir de  $\pi(x|\theta^*)$

Até  $\rho(S(y), S(x^*)) < \epsilon$

Faça  $\theta_i = \theta^*$

- Fim para

onde :

- $\rho$  é uma função de distância.
- $\pi(\theta)$  é a distribuição a priori de  $\theta$ .
- $\pi(\cdot|\theta)$  é a distribuição conjunta dos dados.
- $y$  são os dados observados.



**ABC MCMC - Metropolis-Hasting**

- 1) Inicialize  $\theta_1, i = 1$ .
- 2) Gere um candidato  $\theta^*$  a partir de uma distribuição proposta  $q(\theta, \theta_i)$
- 3) Gere um conjunto de dados  $x^*$  a partir de  $\pi(x|\theta^*)$
- 4) Simule  $u \sim U(0,1)$
- 5) Calcule

$$\alpha = \min \left( 1, \frac{\pi(\theta^*)q(\theta_i, \theta^*)}{\pi(\theta_i)q(\theta^*, \theta_i)} 1(\rho(S(x^*), S(y)) < \epsilon) \right)$$

- 6) Se  $u < \alpha$ , faça  $\theta_{i+1} = \theta^*$ , caso contrário,  $\theta_{i+1} = \theta_i$
- Se  $i < N$ , faça  $i = i + 1$  e vá para o passo 2.

onde :

- $\rho$  é uma função de distância.
- $q(\theta_i|\theta_{i-1})$  é uma distribuição proposta para varrer o espaço paramétrico de  $\theta$ .
- $\pi(.|\theta)$  é a distribuição conjunta dos dados.
- $y$  são os dados observados.

**Population Monte Carlo ABC (PMC-ABC)**

Baseado em [Lenormand et al 2012](#) apresentamos as ideias e o algoritmo PMC-ABC, lembrando que o algoritmo proposto por [Lenormand et al 2012](#) é uma modificação deste algoritmo como base.

Este método consiste em se gerar  $N$  amostras dos parâmetros a cada iteração do algoritmo, ou seja,  $S^{(t)} = (\theta_i^{(t)})_{i=1,\dots,N}$ . Cada partícula  $\theta_i$  da amostra  $S^{(t)}$  deve satisfazer o nível de tolerância predefinido  $\epsilon_t$ , onde  $\epsilon_1 \geq \epsilon_t \geq \epsilon_T$ . Satisfazer o nível de tolerância significa  $\rho_i^t = \rho(x,y) \leq \epsilon_t$ , onde  $x$  foi gerado a partir da distribuição  $f(x|\theta_i^t)$ . No passo  $t$  a amostra  $S^{(t)}$  é gerada a partir da amostra anterior  $S^{(t-1)}$  usando a metodologia de filtro de partícula, ou seja : a primeira amostra  $S^{(1)}$  é gerada usando o passo ABC normal; no passo  $t$ , as partículas são geradas a partir de uma distribuição proposta Gaussiana  $K_t$ , com variância igual ao dobro da variância empírica ponderada da amostra anterior  $S^{(t-1)}$  ( $\theta_i^{(t)} \sim K_t(\theta|\theta^*)$ ), até que  $\theta_i^{(t)}$  satisfaça o nível de tolerância  $\epsilon_t$ , onde  $\theta^*$  é amostrado aleatoriamente da amostra anterior  $S^{(t-1)}$  com probabilidade  $(w_i^{(t-1)})_{i=1,\dots,N}$ , sendo  $w_i^{(t-1)}$  proporcional ao inverso de sua importância na amostra anterior  $S^{(t-1)}$ ; o algoritmo termina quando a amostra  $S^{(T)}$  é gerada, isto é, quando o nível de tolerância  $\epsilon_T$  é alcançado. A seguir, os detalhes do algoritmo :

- Atribua valores iniciais para  $N$  partículas e defina uma sequência decrescente de níveis de tolerância  $\epsilon_1 \geq \epsilon_t \geq \epsilon_T$ . Faça  $t = 1$ .

- Para  $i = 1$  até  $N$  faça

Repita

Simule  $\theta_i^{(1)} \sim \pi(\theta)$  e  $x \sim f(x|\theta_i^{(1)})$

Até  $\rho(S(x),S(y)) \leq \epsilon_1$

$$w_i^{(1)} = \frac{1}{N}$$

- Fim para.
- $\sigma_2^2 = 2$  vezes a variância empírica ponderada de  $(\theta_i^{(1)})_{1 \leq i \leq N}$
- Para  $t = 2$  até  $T$  faça

Para  $i = 1$  até  $N$  faça

Repita

.. Amostre  $\theta_i^*$  de  $\theta_j^{(t-1)}$  com probabilidades  $w_j^{(t-1)}$

.. Gere  $\theta_i^{(t)}|\theta_i^* \sim N(\theta_i^*, \sigma_t^2)$  e  $x \sim f(x|\theta_i^{(t)})$

Até  $\rho(S(x),S(y)) \leq \epsilon_t$

$$w_i^{(t)} \propto \frac{\pi(\theta_i^{(t)})}{\sum_{j=1}^N w_j^{(t-1)} \sigma_t^{-1} \phi(\sigma_t^{-1}(\theta_i^{(t)} - \theta_j^{(t-1)}))}, \text{ onde } \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Fim para

$$\sigma_{t+1}^2 = 2 \text{ vezes a variância empírica ponderada de } (\theta_i^{(1)})_{1 \leq i \leq N}$$

- Fim para

### 1.3 Teoria assintótica para Grafos Aleatórios Exponenciais

Dado o modelo da família exponencial para grafos

$$P_\beta(G) = \exp\left(\sum_{i=1}^k \beta_i T_i(G) + \phi(\beta)\right),$$

onde  $\beta = (\beta_1, \dots, \beta_k)$  é um vetor de parâmetros reais,  $T_1, T_2, \dots, T_k$  são funções no espaço de grafos (exemplo : número de arestas, número de triângulos,  $k$ -estrelas, etc ...), e  $\phi$  é uma constante de normalização, [Diaconis e Chatterjee 2013](#) provam que para  $\beta_1 \in \mathbb{R}$  e  $\beta_2 > 0$ , onde  $T_1$  é o número de arestas e  $T_2$  é o número de triângulos, no limite, ou seja, quando o número de nós é suficientemente grande ( $> 30$ ), o modelo converge para o modelo de Erdos-Rényi, isto é, o grafo converge para um grafo aleatório de Bernoulli onde as arestas são formadas com probabilidade  $u^*(\beta_1, \beta_2)$ , onde

$$u^*(\beta_1, \beta_2) = \max\left(\beta_1 u + \beta_2 u^3 - \frac{1}{2} u \log u - \frac{1}{2} (1-u) \log(1-u)\right).$$

Na verdade, [Diaconis e Chatterjee 2013](#) provam que tal comportamento ocorre de forma geral, onde essa convergência vale para quase todos os grafos.

Os gráficos a seguir mostram a relação entre  $u^*$  e  $(\beta_1, \beta_2)$ .

A Figura 1.4 destaca que há uma descontinuidade à esquerda onde  $u^*$  salta de próximo de 0 para próximo de 1. Isso corresponde à transição de fase na região  $\beta_1 < 0$  e  $\beta_2 > 0$ .

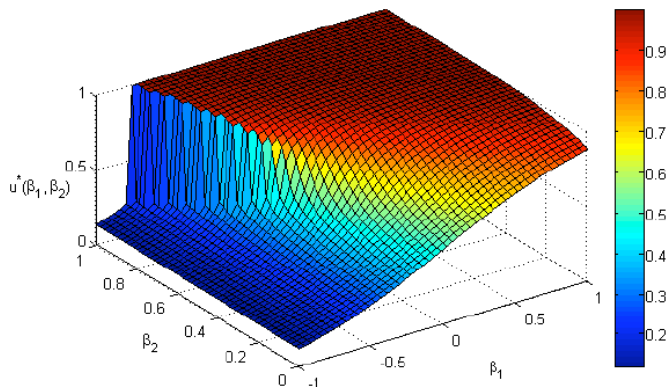


Figura 1.4: Transição de fase ([Diaconis e Chatterjee 2013](#)).

O gráfico da Figura 1.5 destaca as linhas de *contour* para a Figura 1.4. Todos os pares  $(\beta_1, \beta_2)$  na mesma linha correspondem ao mesmo valor de  $u^*$  e, conseqüentemente, no limite, esses modelos corresponderão ao mesmo modelo Erdos-Rényi. A região de transição de fase é vista na parte superior esquerda onde todas as linhas convergem.

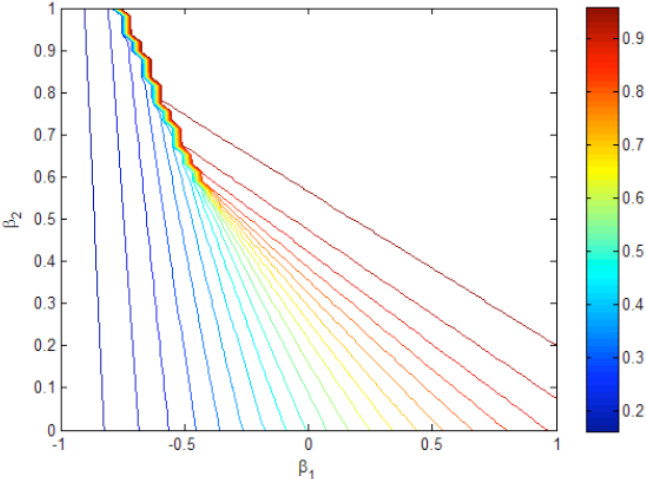


Figura 1.5: Linhas de *contour* para a figura 1.4 (Diaconis e Chatterjee 2013).

## Capítulo 2

# Inferência para Modelos ERGM

### 2.1 Inferência Clássica

#### Estimativa por Pseudo-Máxima Verossimilhança (MPLE)

Inicialmente proposto por [Besag 1974](#) uma estratégia geral para aproximar distribuições de um *campo aleatório de Markov* e adaptado para modelos de rede por [Strauss e Ikeda 1990](#), a estratégia de estimação chamada de **Máxima pseudo-verossimilhança** consiste de um produto de distribuições condicionais completas, a saber :

$$\begin{aligned}\pi(y|\theta) &\approx \pi_{pseudo}(y|\theta) = \prod_{i \neq j} \pi(y_{ij}|y_{-ij}, \theta) \\ &= \prod_{i \neq j} \frac{\pi(y_{ij} = 1|y_{-ij}, \theta)^{y_{ij}}}{[1 - \pi(y_{ij} = 1|y_{-ij}, \theta)]^{1-y_{ij}}}\end{aligned}$$

onde  $y_{-ij}$  representa todos os pares de nós do grafo, excluindo o par  $y_{ij}$ .

A idéia básica por trás desse modelo é a suposição de baixa dependência entre as variáveis no grafo de forma que a verossimilhança possa ser aproximada pela pseudo-verossimilhança, o que torna a estimação rápida e viável, pois não envolve a constante de normalização que é intratável para grafos com mais de 6 vértices ([Strauss e Ikeda 1990](#)). Mesmo assim essa estratégia mostrou-se de forma geral inadequada uma vez que ela utiliza apenas informações locais enquanto que a estrutura de um grafo é afetada por interações globais ([Caimo e Friel 2011](#)). Além disso, [Corander et al. 1998](#) apresentaram estudos de simulação que sugerem a inconsistência do método.

#### Estimativa de Máxima Verossimilhança por Monte Carlo (MC-MLE)

[Geyer e Thompson 1992](#) propuseram uma estratégia de aproximar a verossimilhança de modelos da família exponencial utilizando  $m$  amostras  $\{y_i\} \sim \pi(y|\theta_0)$  para  $\theta_0$  conhecido, por meio de algoritmos MCMC. A verossimilhança aproximada é, então, maximizada para aproximar o estimador de máxima verossimilhança (EMV).

Para se obter uma aproximação da verossimilhança, escreve-se o log da verossimilhança de  $\theta$  e  $\theta_0$  como

$$w_{\theta_0}(\theta) = l(\theta) - l(\theta_0) \approx (\theta - \theta_0)^t s(y) - \log \left[ \frac{1}{m} \sum_{i=1}^m \exp\{(\theta - \theta_0)^t s(y_i)\} \right],$$

onde

$l(\cdot)$  é a log-verossimilhança e

$$\log \left[ \frac{1}{m} \sum_{i=1}^m \exp\{(\theta - \theta_0)^t s(y_i)\} \right] = \frac{1}{m} \frac{\hat{z}(\theta)}{\hat{z}(\theta_0)} = \frac{1}{m} \sum_{i=1}^m \frac{q_{\theta}(y_i)}{q_{\theta_0}(y_i)}$$

Veja que  $(\theta - \theta_0)^t s(y)$  depende dos parâmetros  $\theta$  e  $\theta_0$  e dos dados  $y$ ;

$\log \left[ \frac{1}{m} \sum_{i=1}^m \exp\{(\theta - \theta_0)^t s(y_i)\} \right]$  usa os parâmetros  $\theta$  e  $\theta_0$  e a amostra  $\{y_i\}$ , mas não usa os dados  $y$ ; e  $w_{\theta_0}(\theta)$  é uma função de  $\theta$ , sendo seu máximo um estimador de Monte Carlo para o EMV.

A escolha de  $\theta_0$  é crucial sendo que o ideal é que ele seja próximo do EMV de  $\theta$ .  $w_{\theta_0}(\theta)$  é muito sensível ao valor de  $\theta_0$  que, se não for bem escolhido, pode-se não obter uma função que tenha máximo (Caimo e Friel 2011).

## 2.2 Inferência Bayesiana

A inferência Bayesiana para modelos ERGM está disponível no pacote do R **Bergm**. O tratamento Bayesiano para este problema (Caimo e Friel 2011), onde uma distribuição  $\pi(\theta)$  é atribuída para  $\theta$ , o interesse é a distribuição *a posteriori*

$$\pi(\theta|y) \propto \pi(y|\theta)\pi(\theta)$$

Uma implementação simples (*naive*) do algoritmo Metropolis-Hastings que propõe mover de  $\theta$  para  $\theta^*$  segue abaixo :

### Algoritmo Metropolis-Hastings *naive* para o modelo ERGM

Inicializa  $\theta$  e o número de iterações  $T$ ;

```

while  $t \leq T$  do
  Amostre  $\theta^* \sim h(\cdot|\theta)$ ;
  Compute  $\alpha = \min \left( 1, \frac{q_{\theta^*}(y)\pi(\theta^*)}{q_{\theta}(y)\pi(\theta)} \times \frac{z_{\theta}}{z_{\theta^*}} \right)$ ;
  Amostre  $r \sim U[0,1]$ ;
  if ( $r < \alpha$ ) then
    |  $\theta = \theta^*$ ;
  end
   $t = t + 1$ 
end

```

onde :

- $q_{\theta}(\mathbf{y}) = \exp\{\theta^t s(\mathbf{y})\}$  é o núcleo da função de verossimilhança.
- $h(\cdot|\theta)$  é a distribuição proposta para movimentação de  $\theta$ .

A razão  $\frac{z_\theta}{z_{\theta^*}}$  é impraticável devido à presença de 2 constantes de normalização :  $z_\theta$  e  $z_{\theta^*}$ . O que torna o processo duplamente não tratável.

O algoritmo implementado no pacote **Bergm** do **R** e apresentado no artigo de [Caimo e Friel 2011](#) contorna esse problema a partir da utilização de uma versão adaptada, para o contexto ERGM, do algoritmo *exchange* descrito por [Murray et al. 2006](#).

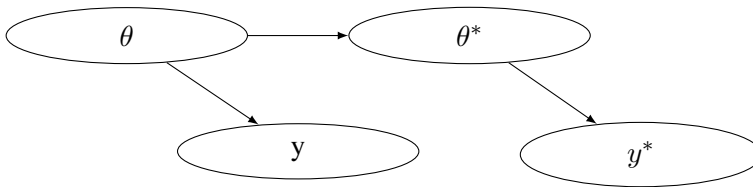
O algoritmo *exchange* amostra de uma distribuição estendida, a saber :

$$\pi(\theta^*, y^*, \theta | y) \propto \pi(y | \theta) \pi(\theta) h(\theta^* | \theta) \pi(y^* | \theta^*)$$

onde :

- $\pi(y^* | \theta^*)$  é a mesma distribuição na qual  $y$  está definida.
- $h(\theta^* | \theta)$  é uma distribuição arbitrária para os parâmetros estendidos que podem depender dos parâmetros  $\theta$  (ex.: um passeio aleatório centrado em  $\theta$ ),

cuja representação gráfica é a seguinte :



A seguir apresentamos os passos do algoritmo *exchange*:

Inicializa  $\theta$  e o número de iterações  $T$ ;

**while**  $t \leq T$  **do**

    Amostre  $\theta^* \sim h(\cdot | \theta)$ ;

    Gere o grafo auxiliar  $y^* \sim \pi(\cdot | \theta^*)$ ;

    Compute  $\alpha = \min \left( 1, \frac{q_\theta(y^*) \pi(\theta^*) h(\theta | \theta^*) q_{\theta^*}(y)}{q_\theta(y) \pi(\theta) h(\theta^* | \theta) q_{\theta^*}(y^*)} \times \frac{z(\theta) z(\theta^*)}{z(\theta) z(\theta^*)} \right)$ ;

    Amostre  $r \sim U[0,1]$ ;

**if**  $(r < \alpha)$  **then**

        |  $\theta = \theta^*$ ;

**end**

$t = t + 1$

**end**

Pode-se verificar que todas as constantes não calculáveis se cancelam. A movimentação de troca (*exchange*) ocorre quando é oferecido ao grafo observado  $y$  o parâmetro auxiliar  $\theta^*$  e, de forma análoga, quando é oferecido ao grafo auxiliar  $y^*$  o parâmetro  $\theta$ .

A equação para cálculo de  $\alpha$  ainda pode ser simplificada se optarmos por uma distribuição simétrica para  $h(\cdot)$  :

$$\alpha = \min \left( 1, \frac{q_\theta(y^*) \pi(\theta^*) q_{\theta^*}(y)}{q_\theta(y) \pi(\theta) q_{\theta^*}(y^*)} \right)$$

A razão  $\frac{q_\theta(y^*)}{q_{\theta^*}(y^*)}$  pode ser vista como uma estimativa de amostragem por importância da razão  $\frac{z(\theta)}{z(\theta^*)}$  uma vez que :

$$E_{y^* | \theta^*} \left[ \frac{q_\theta(y^*)}{q_{\theta^*}(y^*)} \right] = \sum_{y^* \in \Upsilon} \frac{q_\theta(y^*)}{q_{\theta^*}(y^*)} \frac{q_{\theta^*}(y^*)}{z(\theta^*)} = \frac{z(\theta)}{z(\theta^*)}$$

É importante destacar que o passo que gera o grafo auxiliar é um passo difícil, uma vez que requer que se simule da distribuição  $\pi(y^*|\theta^*)$ . Essa simulação é efetuada a partir da utilização de procedimentos MCMC (Metropolis-Hastings) onde a cada iteração é efetuada a comparação da probabilidade do grafo proposto com a probabilidade do grafo observado, e se decide se o grafo proposto será aceito ou não. O grafo proposto é selecionado a cada passo a partir da alteração do estado corrente de um par de nós, isto é, criando uma nova aresta ou excluindo uma já existente. Esse processo é computacionalmente caro. Com o objetivo de se melhorar o *mixing* da cadeia de Markov, utiliza-se a estratégia *tie no tie (TNT)*.

A estratégia TNT funciona da seguinte forma : a cada iteração, ao invés de se selecionar um par de nós aleatoriamente, primeiro se seleciona com igual probabilidade o conjunto de pares de nós com arestas ou o conjunto sem arestas; uma vez selecionado o conjunto, propõe-se, aleatoriamente, criar ou excluir uma aresta, dependendo de qual conjunto foi selecionado. Como na prática os grafos são normalmente esparsos, a probabilidade de se selecionar um par de nós sem arestas é menor e, com isso, o amostrador não perde tanto tempo propondo a criação de arestas que tem grande chance de não serem aceitas (Morris et al. 2008)

Caimo e Friel 2011 propuseram o uso de um método adaptativo (*adaptive direction sampling* - ADS - (Roberts e Gilks 1994) através da estratégia *population* MCMC que consiste de  $H$  cadeias que interagem entre si. O espaço amostral é  $\theta_1, \dots, \theta_H$  (cada  $\theta_i$  é um vetor) e a distribuição a posteriori é  $\pi(\theta_1|y) \otimes \dots \otimes \pi(\theta_H|y)$ . O algoritmo *exchange* adaptativo, descrito a seguir, detalhe o uso dessas cadeias :

### Algoritmo *exchange* adaptativo

Para cada cadeia  $h = 1, \dots, H$ , faça :

1. Selecione aleatoriamente de  $\{1, \dots, H\}$ , e sem reposição,  $h_1$  e  $h_2$ .
2. Amostre  $\epsilon$  de uma distribuição simétrica em  $\mathbb{R}$ .
3. Proponha  $\theta_{h_1}^* = \theta_{h_1}^i + \gamma(\theta_{h_2}^i - \theta_{h_1}^i) + \epsilon$ , onde  $\gamma$  é um fator de movimentação.
4. Amostre  $y^*$  de  $\pi(\cdot|\theta_{h_1}^*)$  pelos métodos MCMC (TNT, por exemplo), retirando uma realização de muitas execuções da cadeia como uma amostra aproximada dessa distribuição.
5. Aceite a movimentação de  $\theta_{h_1}^i$  para  $\theta_{h_1}^{i+1} = \theta_{h_1}^*$  com probabilidade

$$\alpha = \min \left( 1, \frac{q_{\theta_{h_1}^*}(y^*)\pi(\theta_{h_1}^*)q_{\theta_{h_1}^i}(y)}{q_{\theta_{h_1}^i}(y)\pi(\theta_{h_1}^i)q_{\theta_{h_1}^*}(y^*)} \right)$$

A Figura 2.1 mostra, esquematicamente, como ocorre a movimentação das partículas na implementação do algoritmo *exchange* adaptativo.



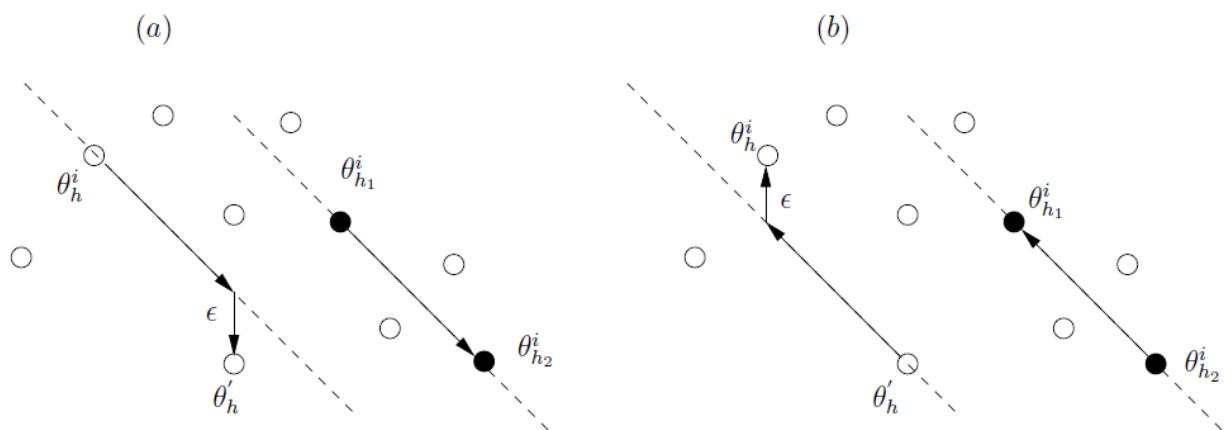


Figura 2.1: Gráfico ilustrativo da movimentação referente ao algoritmo *exchange* adaptativo.



## Capítulo 3

# Metodologia ABC para ERGM

Conforme apresentado, o modelo ERGM possui a seguinte função de verossimilhança

$$\pi(\mathbf{y}|\theta) = \frac{q_{\theta}(\mathbf{y})}{z(\theta)} = \frac{\exp\{\theta^t s(\mathbf{y})\}}{z(\theta)}$$

Onde :

- $s(\mathbf{y})$  é um vetor conhecido de estatísticas suficientes.
- $\theta \in \Theta$  : é um vetor de parâmetros do modelo. Um número positivo para  $\theta_i \in \Theta$  resulta em uma maior probabilidade de se observar a estatística  $s(\mathbf{y})$  nos grafos.
- $z(\theta) = \sum_{y \in Y} \exp\{\theta^t s(\mathbf{y})\}$  a constante de normalização.

Os modelos ERGM pertencem à classe de modelos com constante de normalização intratável, ou seja, a função de verossimilhança não é conhecida, requerendo, com isso, a utilização de métodos específicos para a inferência de seus parâmetros. Os métodos ABC permitem ajustar um modelo aos dados sem se basear no cálculo da verossimilhança (Beaumont et al. 2002) dependendo, ao contrário, de simulações dos dados a partir do modelo.

Podemos ver, a seguir, que as principais características dos métodos ABC o tornam um bom método a ser aplicado no contexto dos modelos ERGM :

- o método consiste em gerar  $\theta^*$  da distribuição a priori do parâmetro  $\pi(\theta)$  e gerar grafos  $y^* \sim \pi(y|\theta^*)$ , aceitando o  $\theta^*$  simulado, se e somente se,  $\rho(y, y^*) < \epsilon$ , onde  $\rho$  é uma medida de distância entre o dado observado  $y$  e o simulado  $y^*$ ;
- os  $\{\theta_i^*\}_{1 \leq i \leq N}$  gerados tem distribuição de probabilidade proporcional a  $\pi(\theta)\pi(\rho(y, y^*) < \epsilon|\theta)$ , onde  $\pi(x|\theta)$  representa a distribuição de probabilidade de  $x$  dado o parâmetro  $\theta$ . Essa densidade é uma aproximação da distribuição a posteriori do modelo ( $\pi(\theta)\pi(y|\theta)$ ).
- a comparação  $\rho(y, y^*) < \epsilon$ , onde  $\epsilon$  é o nível de tolerância, ocorre a partir de estatísticas suficientes com relação aos parâmetros do modelo ( $\theta$ ) e o aumento da eficiência obtida não ocorre em erro, pois, por definição, suficiência implica que toda a informação de  $y$  a respeito de  $\theta$  é capturada por  $s(y)$  (Wilkinson 2008), ou seja,  $\rho(s(y), s(y^*)) < \epsilon$ .
- no modelo ERGM, estatísticas suficientes são contagens de estruturas do grafo, tais como : número de arestas, número de triângulos, número de  $k$ -estrelas, etc.

- a utilização de níveis de tolerância é devido ao fato da comparação exata requerer um grande número de simulações com custo computacional elevado, e também porque para dados contínuos a comparação exata não ser possível.

A seguir apresentamos um algoritmo para geração de distribuições *a posteriori* dos parâmetros do modelo ERGM com utilização da técnica ABC.

#### **Estratégia ABC para modelos ERGM**

- 1) Dado um grafo observado  $y$ .
- 2) Calcule as estatísticas suficientes para  $y$ .
- 3) Configure a distribuição *a priori*.
- 4) Configure o modelo a ser utilizado pelo método ABC.
- 5) Configure demais argumentos do método ABC, tais como : tamanho da amostra a ser gerada, critério de parada, proporção de partículas a ser repassada de uma iteração para outra, etc.
- 6) Execute método ABC informando : a distribuição *a priori*, as estatísticas suficientes do grafo observado, o modelo, os demais argumentos.

O resultado do Item 6 é a amostra *a posteriori* dos parâmetros que é o passo no qual se implementa o algoritmo ABC. Listamos, a seguir, alguns dos principais métodos existentes:

- ABC Rejeição ([Pritchard et al. 1999](#));
- ABC MCMC ([Marjoram et al. 2003](#));
- ABC Monte Carlo Sequencial (SMC-ABC) ([Del Moral et al 2006](#));
- ABC Monte Carlo Sequencial Adaptativo (ASMC-ABC) ([Del Moral et al. 2012](#));
- ABC Population Monte Carlo (PMC-ABC) ([Beaumont et al. 2009](#));
- ABC Monte Carlo Sequencial Renovado (*replenishment*) (RSMC-ABC); ([Drovandi e Pettitt 2011](#));
- ABC Adaptive Population Monte Carlo (APMC-ABC) ([Lenormand et al 2012](#)).

## Capítulo 4

# Estudo de Simulação

De acordo com o trabalho de [Diaconis e Chatterjee 2013](#), grafos aleatórios exponenciais com dois parâmetros são assintoticamente (número de nós maior que 30) equivalentes a grafos aleatórios Bernoulli. A aplicação dos resultados desse trabalho e a utilização de métodos ABC na inferência de parâmetros desses modelos são as grandes contribuições do nosso estudo.

Sendo assim, nosso estudo de simulação foi composto dos seguintes modelos :

- Modelo para grafos Bernoulli : Caso Erdős - Renyi
- Modelo para grafos Exponenciais : Arestas e Triângulos

Conforme apresentado no Item 1.1, os grafos aleatórios Bernoulli foram os primeiros grafos aleatórios a serem propostos na literatura por [Erdos-Renyi 1959](#). Esses grafos são um caso especial de grafos aleatórios exponenciais, onde considera-se apenas um parâmetro  $\theta$ , associado ao número de arestas, sendo que as arestas ocorrem independentemente com a mesma probabilidade

$$P(Y_{ij} = 1) = \frac{e^\theta}{1 + e^\theta}$$

para  $i \neq j$ .

Para os estudos de simulação para grafos aleatórios exponenciais optamos pelo modelo que inclui parâmetros para arestas e triângulos por ser o modelo detalhado por [Diaconis e Chatterjee 2013](#).

Todos os resultados aqui apresentados foram executados em uma máquina configurada com Linux Ubuntu 14.04.02, Processador Intel Core i5-3330, CPU 3GHz e 15G de RAM.

### 4.1 Grafos Bernoulli : Caso Erdős - Renyi

O modelo estatístico para grafos aleatórios exponenciais com apenas um parâmetro associado ao número de arestas é o seguinte :

$$\pi(\mathbf{y}|\theta) = \frac{q_\theta(\mathbf{y})}{z(\theta)} = \frac{\exp\{\theta_1 s_1(\mathbf{y})\}}{z(\theta_1)}$$

Onde :

- $s_1(\mathbf{y}) = \sum_{i < j} y_{ij}$ , é a estatística suficiente igual ao número de arestas.

- $\theta_1 \in \Theta$  : é o parâmetro do modelo.
- $z(\theta_1) = \sum_{y \in Y} \exp\{\theta_1 s_1(y)\}$  é a constante de normalização.

Neste modelo a probabilidade  $p$  de criação das arestas é constante e igual a  $p = \frac{e^{\theta_1}}{1+e^{\theta_1}}$  e resolvendo a equação de  $p$  para  $\theta_1$  obtemos  $\theta_1 = -\log(\frac{1}{p} - 1)$ . Esse resultado é obtido considerando um grafo com apenas 2 nós conectados e, como a suposição é de que as variáveis aleatórias são independentes, a verossimilhança é o produto das probabilidades.

Utilizamos distribuições *a priori* não informativas conforme descrito a seguir :

- Método Bayesiano : Normal (0,100)
- Métodos ABC : Unif (0,1)

Por restrições do pacote **R Bergm**, a distribuição *a priori* é sempre uma distribuição Normal. Para utilizarmos outra distribuição é necessário alterar o código fonte, o que foi feito no teste de simulação para grafos aleatórios exponenciais com dois parâmetros (item 4.2, a seguir), mas que não consideramos necessário para este caso.

Para análise da qualidade do ajuste, consideramos amostras geradas das probabilidades  $p$  e não amostras dos parâmetros. O motivo para tal escolha é devido ao fato da distribuição *a posteriori* de  $p$  ter forma fechada com rotinas para geração de amostra teórica implementadas na linguagem **R**. Ou seja, a distribuição *a posteriori* para um modelo Bernoulli com distribuição *a priori* Beta (1, 1) = Unif (0, 1), é uma distribuição Beta ( $\alpha$  = número de pares de nós conectados + 1,  $\beta$  = número de pares de nós não conectados + 1). A qualidade do ajuste foi analisada a partir dos testes de Kullback–Leibler e Kolmogorov–Smirnov, também implementados pela linguagem **R**.

**Kullback–Leibler** é uma medida de diferença entre duas distribuições de probabilidade  $P$  e  $Q$  que equivale ao valor esperado da diferença logarítmica entre essas distribuições, onde a esperança é calculada usando a probabilidade de referência  $P$  ou  $Q$ . Ou seja, considerando  $P(x)$  a densidade de referência, matematicamente temos :

- Distribuições discretas :  $\sum_x P(x) \log \frac{P(x)}{Q(x)}$ .
- Distribuições contínuas :  $\int p(x) \log \frac{p(x)}{q(x)} dx$ , onde  $p(x)$  e  $q(x)$  são as densidades de  $P(x)$  e  $Q(x)$ , respectivamente.

Na prática, informamos na função do **R** a amostra teórica e a amostra gerada pelos métodos de estimação e, internamente, a função calcula a probabilidade de cada elemento da amostra a partir da proporção (item da amostra) / (soma dos itens da amostra), e retorna o resultado da esperança descrita acima.

O teste **Kolmogorov–Smirnov** tenta determinar se dois conjuntos de dados diferem significativamente. Ele tem a vantagem de não fazer distinção a respeito da distribuição dos dados. Tecnicamente falando, ele é não-paramétrico e independente de distribuição. Ele pode ser utilizado para comparar uma amostra com uma distribuição de probabilidade de referência (teste KS para uma amostra), ou para comparar duas amostras (teste KS para duas amostras). A estatística Kolmogorov–Smirnov quantifica a distância entre a função de distribuição empírica da amostra e a função de distribuição cumulativa da distribuição de referência (teste KS para uma amostra), ou entre as funções de distribuição

empíricas das duas amostras (teste KS para duas amostras). A hipótese nula estabelece que a amostra é gerada a partir da distribuição de referência (teste KS para uma amostra), ou as duas amostras são extraídas da mesma distribuição (teste KS para duas amostras). No nosso caso, utilizamos o teste Kolmogorov-Smirnov para duas amostras, uma vez que obtemos uma amostra da distribuição teórica e geramos outra amostra a partir dos métodos de estimação.

Importante destacar que para esses testes (Kullback–Leibler e Kolmogorov-Smirnov) quanto **menor** a estatística de teste, melhor; mas para o *P-value* do teste Kolmogorov-Smirnov, uma vez que a hipótese nula estabelece que as duas amostras são extraídas da mesma distribuição, quanto **maior**, melhor.

Como já descrito, grafos aleatórios exponenciais são muito utilizados no contexto de redes sociais. O tamanho dessas redes pode variar de poucos nós a milhões de nós (<http://snap.stanford.edu/data/index.html>). Para uma análise mais realística, consideramos redes com os seguintes números de nós : 50, 100, 200, 500, 750, 1.000, 1.250, 1.500, 1.750, 2.000 e 2.250. O número limite de 2.250 foi obtido definindo-se o tempo de processamento de 24 horas para estimação de parâmetros considerando o método Bayesiano implementado no pacote **R Bergm**.

A análise comparativa realizada ocorreu entre o método Bayesiano implementado pelo pacote **R Bergm**, descrito no Item 2.2, e os seguintes métodos ABC, descrito no Item 3, implementados pelo pacote **R EasyABC**:

- ABC Rejeição ([Pritchard et al. 1999](#))
- ABC MCMC ([Marjoram et al. 2003](#))
- ABC Monte Carlo Sequencial Adaptativo (ASMC-ABC) ([Del Moral et al. 2012](#))
- ABC Population Monte Carlo (PMC-ABC) ([Beaumont et al. 2009](#))
- ABC Monte Carlo Sequencial Renovado (*replenishment*) (RSMC-ABC) ([Drovandi e Pettitt 2011](#))
- ABC Adaptive Population Monte Carlo (APMC-ABC) ([Lenormand et al 2012](#))

A seguir apresentamos os resultados para 50 e 2.250 nós, e  $p = 0.731$ . Os demais resultados podem ser verificados no **Apêndice A**.

Os resultados estão ordenados por método, o tempo de processamento está em segundos, KL = Kullback–Leibler e KS = Kolmogorov-Smirnov.

Tabela 4.1: Resultado grafo aleatório Bernoulli para **50** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Beaumont	0,725973	63	0,000502	0,134	0
Bergm	0,726381	40	0,000297	0,013640553	0,265007
Delmoral	0,7286	5	0,000283	0,082	<u>0,002403</u>
Drovandi	0,726654	83	0,000330	0,025	0,913476
Lenormand	0,727054	199	0,000278	0,062	<u>0,042815</u>
Marjoram	0,726422	207	0,000319	0,0201875	0,002946
Rejeicao	0,726294	171	0,000295	0,026	0,887939

Tabela 4.2: Resultado grafo aleatório Bernoulli para **2.250** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Beaumont	0,092073	3.344	0,222879	1	0
Bergm	0,731024	65.547	0,000035	0,458986	0
Delmoral	0,730905	1.765	0,000009	0,430693	0
Drovandi	<b>0,730872</b>	<b>13.736</b>	<b>0</b>	<b>0,103</b>	<b>0,000049</b>
Lenormand	<b>0,731</b>	<b>11.260</b>	<b>0</b>	<b>0,058</b>	<b>0,069190</b>

#### Análise dos Resultados :

Podemos verificar pelas Tabelas (4.1, 4.2 , A.1 - A.11) e Figuras (4.1, 4.2, 4.6, 4.7, A.1-A28) que :

- Os métodos ABC Sequencial Lenormand e Drovandi se destacam como os que melhor ajustam o modelo.
- O ajuste pelo método Bayesiano (**Bergm**) tem uma piora significativa a medida que se aumenta o número de nós.
- Com o aumento do número de nós, o tempo de processamento do método Bayesiano (**Bergm**) aumenta exponencialmente .
- Os métodos ABC Lenormand e Drovandi são os mais consistentes.
- O método ABC Lenormand apresenta resultados melhores, “perdendo” para o método ABC Drovandi apenas no cenário com 1.750 nós.

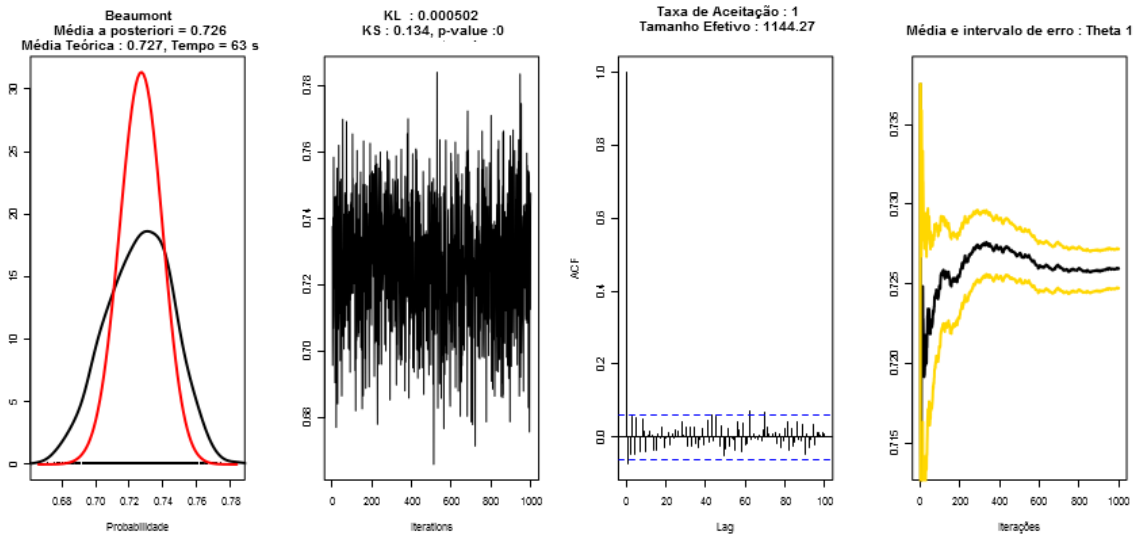
Com relação à configuração dos métodos para execução é importante destacar :

- A calibragem dos parâmetros foi direcionada tanto pela qualidade do ajuste, quanto pelo tamanho efetivo da amostra.
- O método Bayesiano implementado no pacote **Bergm**, baseado no algoritmo *exchange*, mostrou-se de difícil calibragem dos parâmetros de variância/covariância e do parâmetro “gama”, ambos associados à movimentação das partículas nas *H* cadeias (“Population” MCMC).

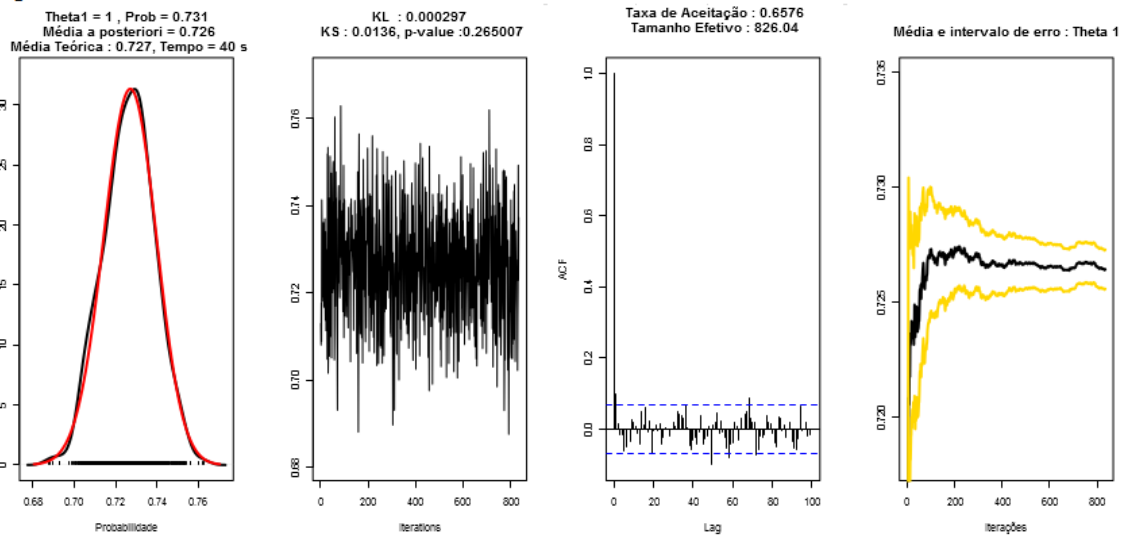


- Os métodos ABC necessitam da especificação de uma função, denominada de função do modelo, a partir do qual os dados são simulados, dados os valores dos parâmetros (gerados pela distribuição *a priori*). A função do modelo deve retornar o cálculo das estatísticas suficientes dos dados simulados. No nosso caso, a função simula um grafo e calcula o número de arestas, no caso do modelo Erdős-Renyi; e o número de arestas e o número de triângulos, no caso do modelo de grafos exponenciais.
- O método ABC de Lenormand requer uma configuração mínima, e sobressai em relação aos demais métodos, especialmente em relação ao método Bayesiano pela facilidade de calibragem de seus parâmetros, praticamente não necessitando alteração dos mesmos quando do aumento do número de nós do grafo.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

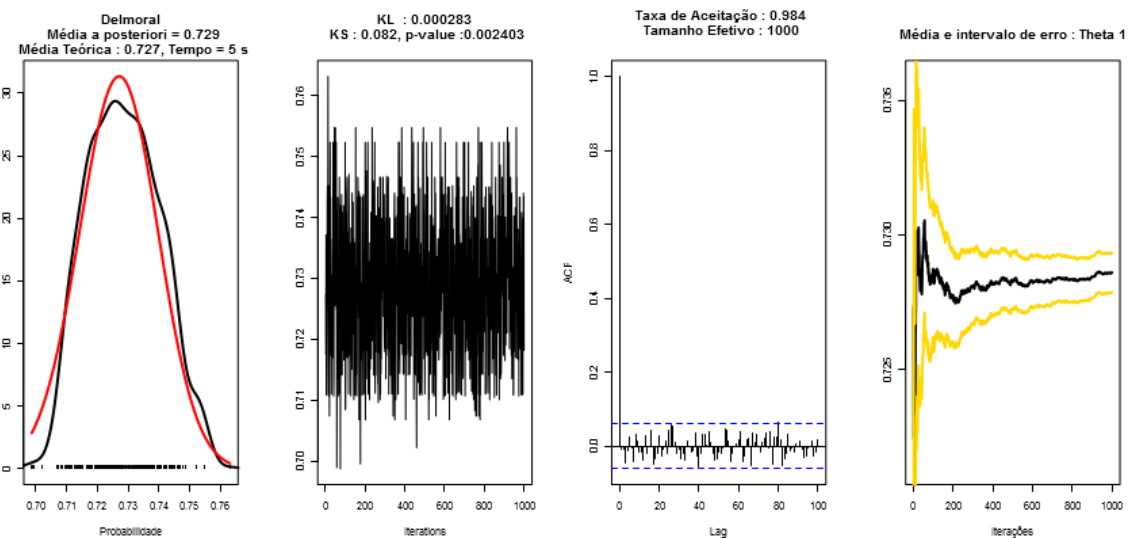
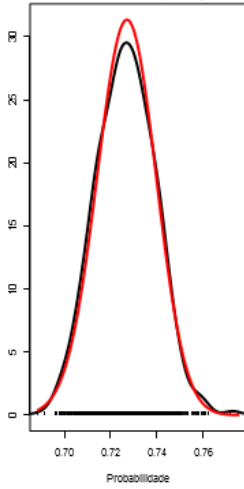


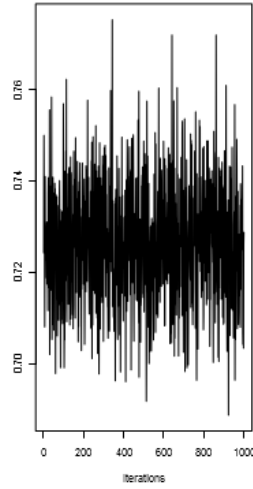
Figura 4.1: Gráficos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi

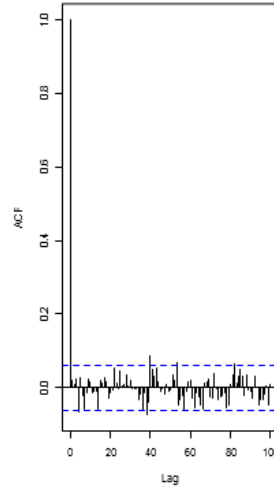
Drovandi  
Média a posteriori = 0.727  
Média Teórica : 0.727, Tempo = 83 s



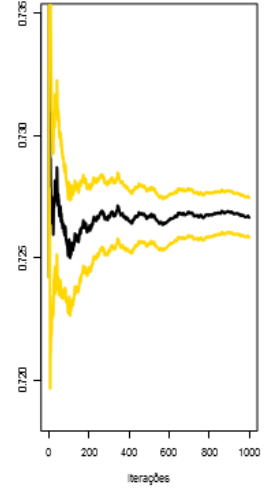
KL : 0.00033  
KS : 0.025, p-value : 0.913476



Taxa de Aceitação : 1  
Tamanho Efetivo : 1000

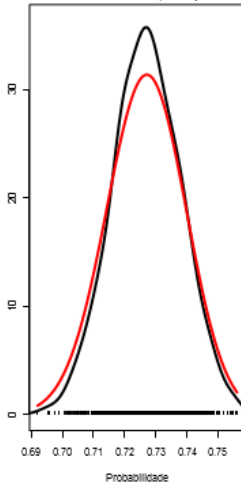


Média e intervalo de erro : Theta 1

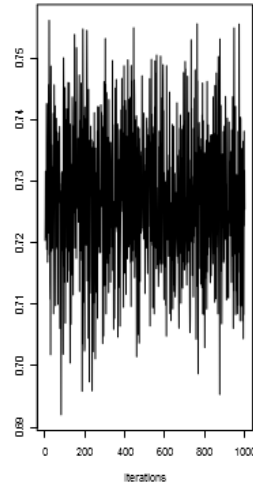


## ABC Sequencial-Lenormand

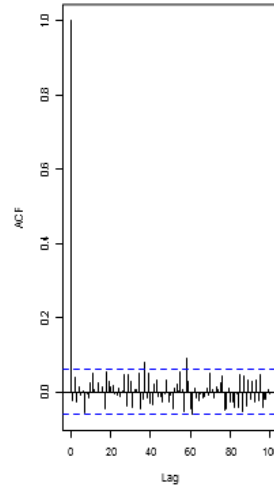
Lenormand  
Média a posteriori = 0.727  
Média Teórica : 0.727, Tempo = 199 s



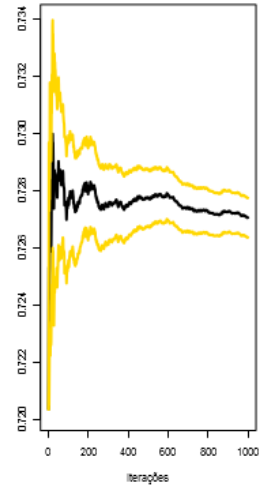
KL : 0.000278  
KS : 0.062, p-value : 0.042815



Taxa de Aceitação : 1  
Tamanho Efetivo : 1000

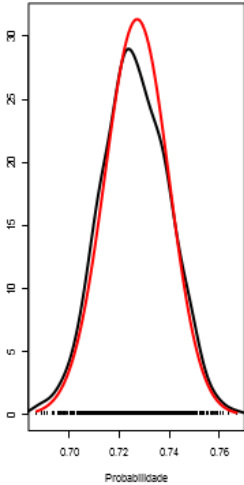


Média e intervalo de erro : Theta 1

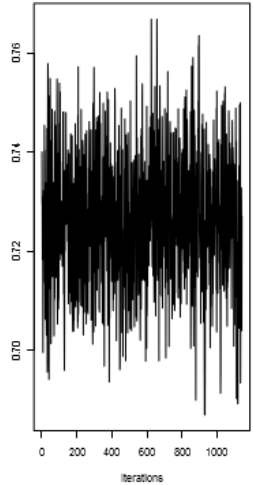


## ABC MCMC-Marjoram

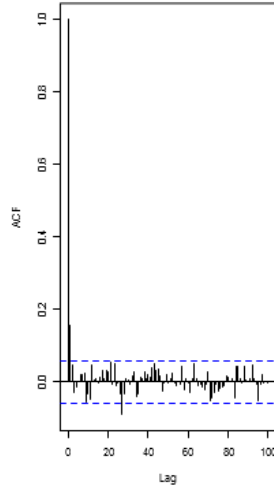
Marjoram  
Média a posteriori = 0.726  
Média Teórica : 0.727, Tempo = 207 s



KL : 0.000319  
KS : 0.0202, p-value : 0.002946



Taxa de Aceitação : 0.9256  
Tamanho Efetivo : 1109.96



Média e intervalo de erro : Theta 1

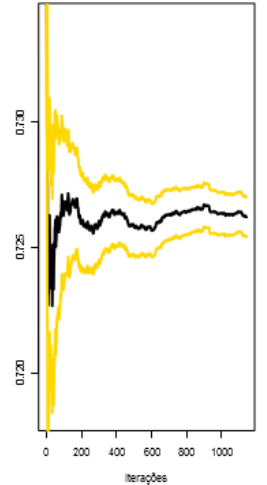


Figura 4.2: Grafos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenormand e Marjoram.

ABC Rejeicao

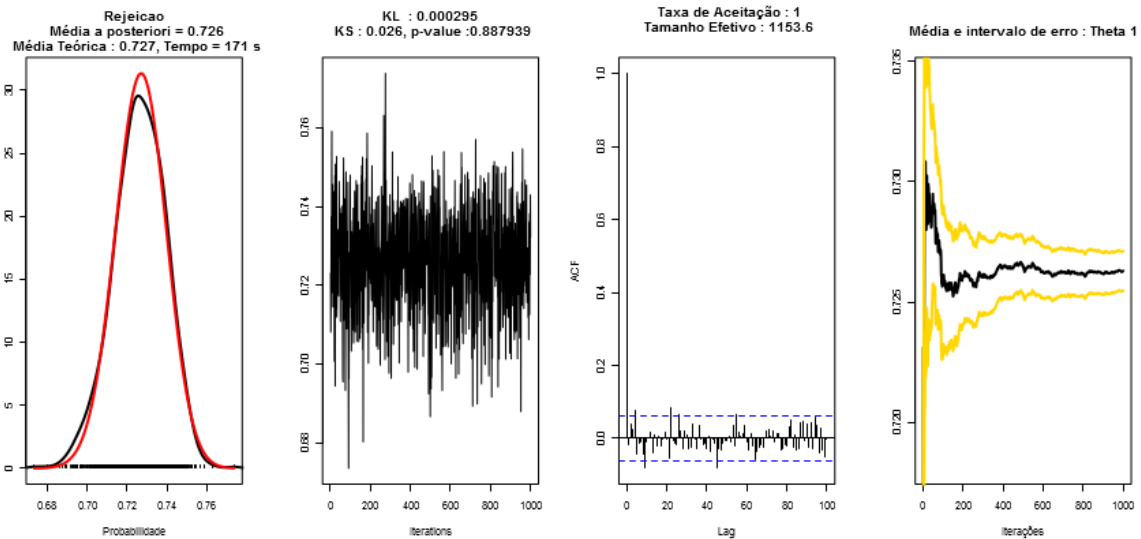
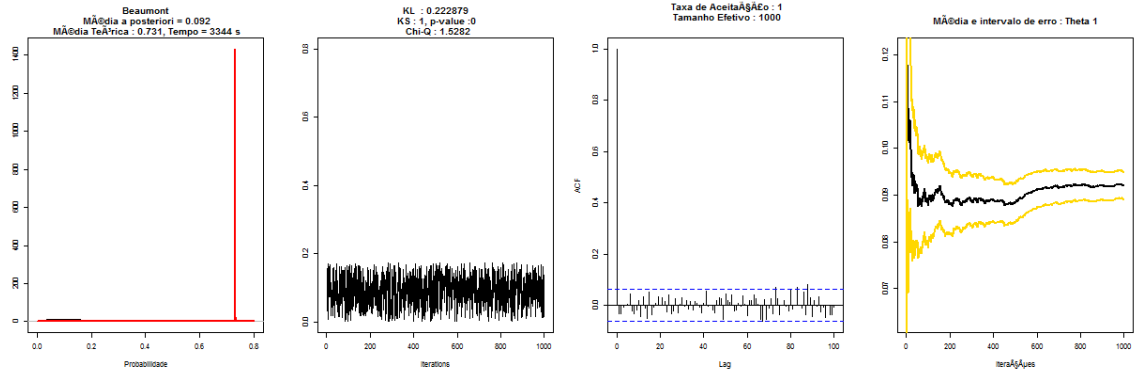
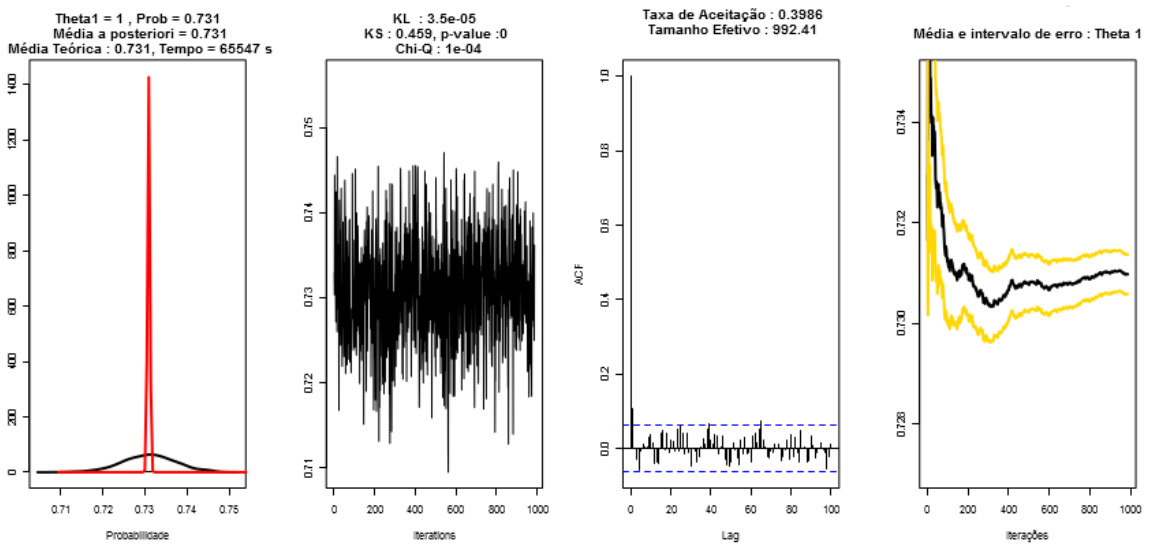


Figura 4.3: Grafos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

## ABC Sequential-Beaumont



## Bergm



## ABC Sequential-Delmoral

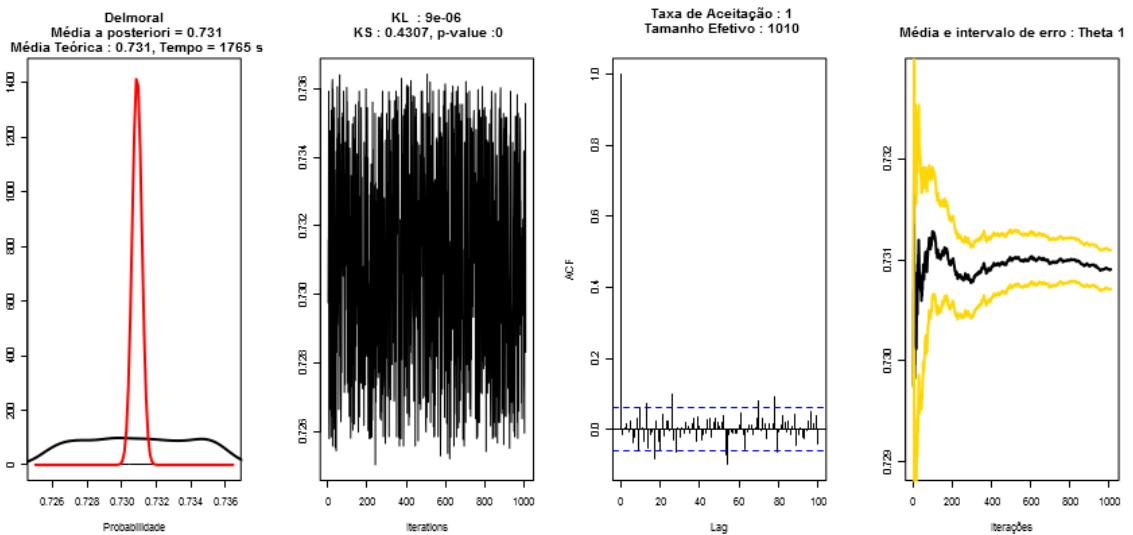
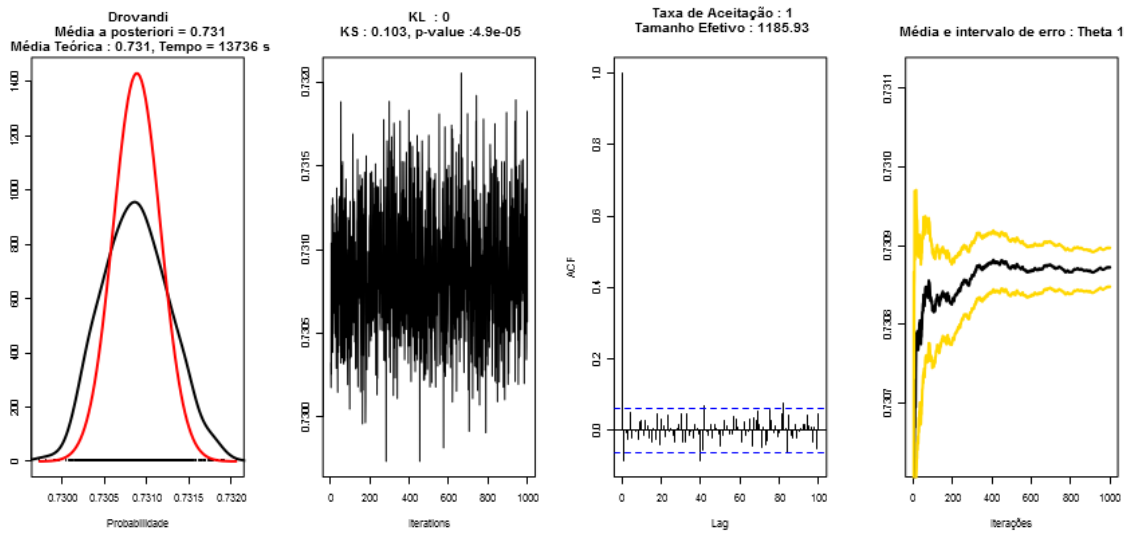


Figura 4.4: Grafos aleatórios Bernoulli **2.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

ABC Sequencial-Drovandi



ABC Sequencial-Lenormand

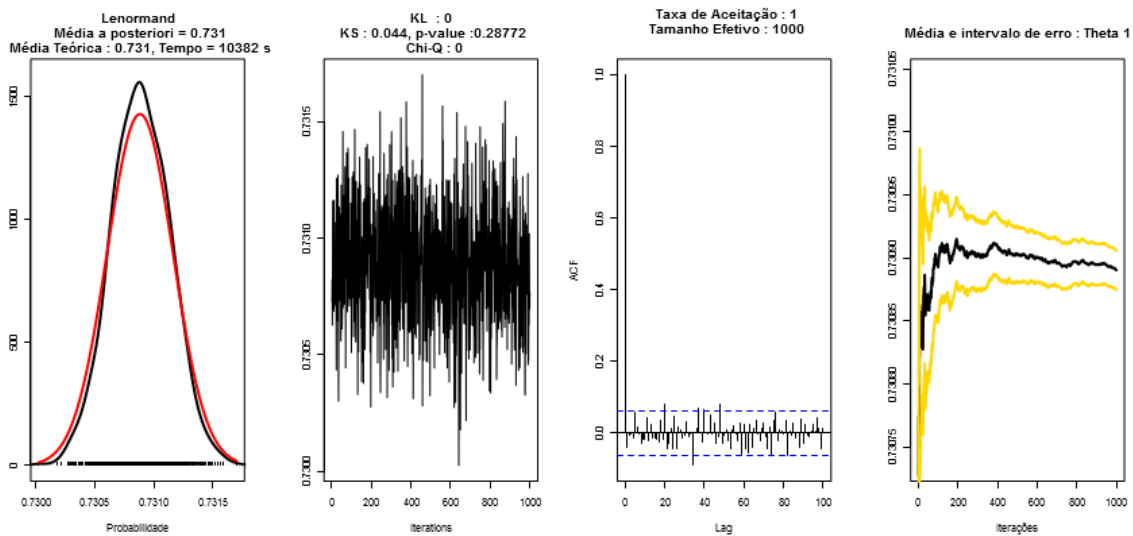


Figura 4.5: Grafos aleatórios Bernoulli **2.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com lag de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi e Lenorman. Não é apresentado o resultado para os métodos Marjoram e Rejeição, pois o tempo de processamento ultrapassou 24 horas.

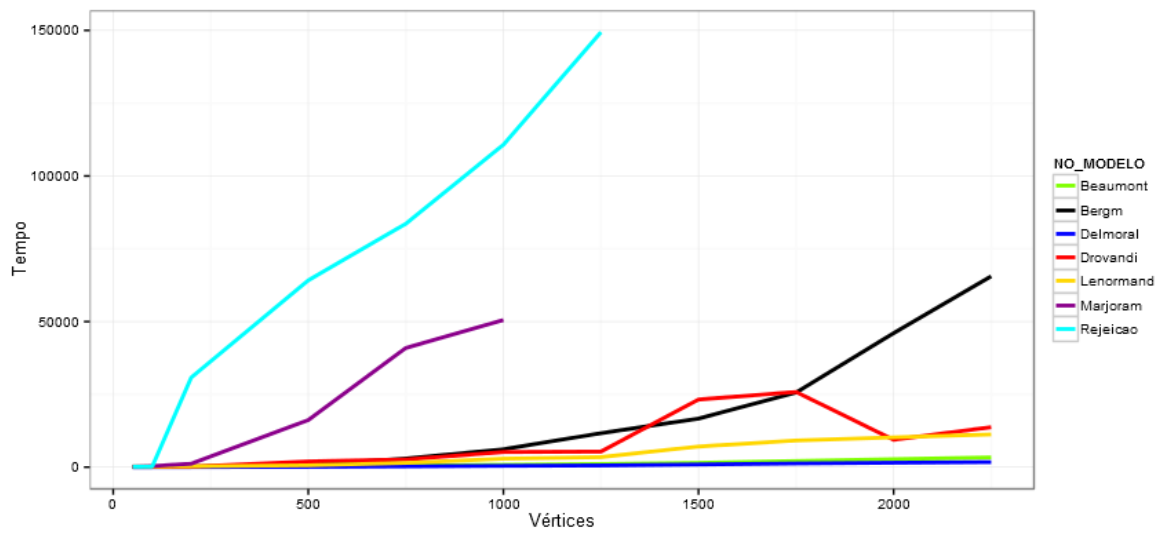


Figura 4.6: Tempo de processamento com todos os métodos.

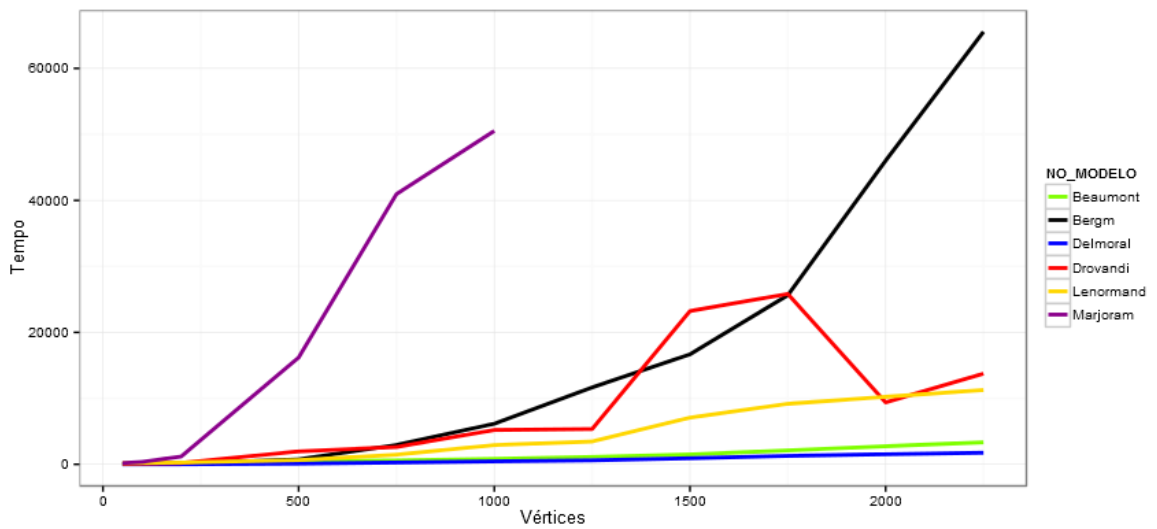


Figura 4.7: Tempo de processamento sem o método ABC-Rejeição, pois ele apresentou tempos de processamento muito elevados, próximos de 24h, com 1.250 vértices, alterando a escala do gráfico e dificultando a visualização comparativa entre os métodos Bayesiano e ABC.

## 4.2 Grafos Exponenciais : Arestas e Triângulos

O modelo estatístico para grafos aleatórios exponenciais com parâmetros associados ao número de arestas e ao número de triângulos é o seguinte :

$$\pi(\mathbf{y}|\theta) = \frac{q_{\theta}(\mathbf{y})}{z(\theta)} = \frac{\exp\{\theta_1 s_1(\mathbf{y}) + \theta_2 s_2(\mathbf{y})\}}{z(\theta_1, \theta_2)}$$

Onde :

- $s_1(\mathbf{y}) = \sum_{i<j} y_{ij}$ , é a estatística suficiente igual ao número de arestas.
- $s_2(\mathbf{y}) = \sum_{i<j<h} y_{ij}y_{jh}y_{ih}$ , é a estatística suficiente igual ao número de triângulos.
- $\theta_1, \theta_2 \in \Theta$  : são os parâmetros do modelo.
- $z(\theta_1, \theta_2) = \sum_{y \in Y} \exp\{\theta_1 s_1(y) + \theta_2 s_2(y)\}$  é a constante de normalização.

Conforme descrito no item anterior, para o método ABC é necessário informar a função do modelo a partir do qual os dados são simulados, dado os valores dos parâmetros. Para este estudo de simulação implementamos duas versões para essa função. Uma versão utilizando os resultados assintóticos de [Diaconis e Chatterjee 2013](#) e outra versão utilizando a rotina de simulação de grafos disponibilizada pelo pacote **R Bergm**. O objetivo de implementarmos essas duas versões foi podermos comparar as estimativas do método Bayesiano com a nossa proposta do método ABC tanto pela geração de amostras de grafos Bernoulli ([Diaconis e Chatterjee 2013](#)), como baseado no método de simulação Bayesiano. A versão assintótica, conforme descrito no Item 1.3, se baseia na relação

$$p^* = u^*(\theta_1, \theta_2) = \max(\theta_1 u + \theta_2 u^3 - \frac{1}{2}u \log u - \frac{1}{2}(1-u) \log(1-u)).$$

onde, para  $\theta_1 \in \mathbb{R}$  e  $\theta_2 > 0$ ,  $T_1$  sendo o número de arestas e  $T_2$  o número de triângulos, o grafo converge para um grafo aleatório Bernoulli onde as arestas são formadas com probabilidade  $p^*$ .

Com isso, a especificação das duas versões da função do modelo para nossa proposta do método ABC podem ser assim resumidas :

- Versão ABC assintótica : a função recebe os valores de  $\{\theta_1, \theta_2\}$  gerados pela distribuição *a priori*; calcula o valor de  $p^*$ ; gera um grafo aleatório Bernoulli com probabilidade  $p^*$ ; e retorna o número de arestas e o número de triângulos do grafo gerado.
- Versão ABC Bayesiana : a função recebe os valores de  $\{\theta_1, \theta_2\}$  gerados pela distribuição *a priori*; simula um grafo aleatório exponencial via método Bayesiano a partir dos valores de  $\{\theta_1, \theta_2\}$ ; e retorna o número de arestas e o número de triângulos do grafo gerado.

[Diaconis e Chatterjee 2013](#) destaca a existência de transição de fase na região do espaço paramétrico onde  $\theta_1 < 0$ , ou seja, região na qual ocorre uma descontinuidade nos valores de  $p^*$  ocorrendo um salto de 0 para 1. Devido a isso, restringimos a região paramétrica em  $\theta = \{\theta_1, \theta_2 : \theta_1 > 0, \theta_2 > 0\}$  utilizando a distribuição *a priori* não informativa LogNormal com  $\mu = -\log(4)/2$  e  $\sigma = \log(4)$ , de tal forma que o parâmetro *a priori* tem média igual a 1 e variância igual a 3. O aprofundamento de estudos dessa situação consta da lista do Capítulo 5 .



Conforme já citado no item anterior, por restrições do pacote **R Bergm** tivemos que alterar o código fonte do mesmo para poder utilizar uma função de distribuição *a priori* diferente da distribuição Normal.

Para obtermos um grafo “observado” base para a inferência, ao invés de gerarmos um grafo aleatório Bernoulli com probabilidade  $u^*(\theta_1, \theta_2)$  e com a possibilidade desse grafo pertencer a uma região de baixa massa probabilística dificultando a convergência da cadeia, utilizamos uma estratégia para evitar tal situação conforme descrito pelo algoritmo a seguir :

```
. Defina os valores para  $\theta_1$  e  $\theta_2$ ;
. Calcule o valor de  $u(\theta_1, \theta_2)$ ;
. Defina um intervalo I para  $u(\theta_1, \theta_2)$  de 10% para mais e para menos;
while o valor de  $p^*$  não estiver no intervalo I do
  . Gere um grafo Bernoulli  $G^*$ , com probabilidade  $u(\theta_1, \theta_2)$ ;
  . Estime os parâmetros  $\theta_1^*, \theta_2^*$  do grafo gerado utilizando método de pseudo- máxima
  verossimilhança (pacote R ergm);
  . Calcule o valor de  $u^*(\theta_1^*, \theta_2^*)$ ;
end
. A partir do grafo  $G^*$ , simule 1.000 grafos;
. Calcule as estatísticas diâmetro, número de arestas, número de triângulos, transitividade e
grau de cada um dos 1.000 grafos;
. Calcule a média e a variância para cada uma dessas estatísticas;
. Calcule intervalos  $IC^*$  para cada uma das estatísticas considerando 2 vezes o desvio padrão
para mais e para menos;
while as estatísticas do grafo  $G^{**}$  não estiverem dentro dos intervalos  $IC^*$  do
  . Gere outro grafo  $G^{**}$  a partir do grafo  $G^*$ ;
  . Calcule as estatísticas diâmetro, número de arestas, número de triângulos, transitividade e
grau para o grafo  $G^{**}$ ;
end
. Salve o grafo  $G^{**}$  para uso na inferência;
```

Baseado nos resultados de simulação dos grafos Bernoulli, dentre os métodos ABC, optamos por efetuar os estudos de simulação para os grafos exponenciais com arestas e triângulos apenas para o método ABC Lenormand, pois foi o que apresentou melhores resultados, conforme apresentado no Item 4.1.

A seguir apresentamos os resultados para 50 nós com tamanho da amostra gerada de aproximadamente 1.000. Fato relevante para se destacar é não termos conseguido gerar amostras para o método Bayesiano em 75% dos cenários simulados, daí o fato de não apresentarmos dados nas tabelas abaixo para esse método e, também, não termos estudos de simulação para um número maior de nós. Na tentativa de gerar as amostras efetuamos, em média, para cada um dos cenários, 6 configurações diferentes para os parâmetros de calibragem do método, sem obter êxito. As amostras geradas foram amostras degeneradas, o que impede de se efetuar as análises. Os gráficos estão disponibilizados no **Apêndice B**.

Para analisarmos os resultados apresentados a seguir é relevante reportarmos ao trabalho de [Diaconis e Chatterjee 2013](#) descrito no Item 1.3, especificamente na Figura 1.5, onde temos que o modelo é não identificável uma vez que pode-se obter o mesmo valor de  $p^*$  para diferentes valores de  $\theta_1$  e  $\theta_2$ , o que é demonstrado pelas curvas de nível da Figura 1.5. Pelos dados das tabelas vemos que

o método **ABC Lenormand, versão assintótica**, mostra exatamente esse fato.

Constatamos que, também para grafos aleatórios exponenciais, nossa proposta de utilização do método ABC aplicando a teoria assintótica de [Diaconis e Chatterjee 2013](#) supera o método Bayesiano tanto na estimativa dos parâmetros, quanto no tempo de processamento.

- Nos cenários  $\{\theta_1 = 0,2, \theta_2 = 0,2\}$  e  $\{\theta_1 = 0,5, \theta_2 = 0,35\}$  as estimativas da versão ABC assintótica estão bem próximas do real valor tanto para  $p^*$ , quanto para os parâmetros. Já as estimativas de  $p^*$  e dos parâmetros para a versão ABC Bayesiana variam de distante para muito distante dos valores reais com tempo de processamento 6 a 8 vezes maior. Para o método Bayesiano do pacote **R Bergm** não foi nem possível gerar amostras.
- Nos cenários  $\{\theta_1 = 0,8, \theta_2 = 0,02\}$ ,  $\{\theta_1 = 0,9, \theta_2 = 0,02\}$ ,  $\{\theta_1 = 0,6, \theta_2 = 0,1\}$  e  $\{\theta_1 = 0,7, \theta_2 = 0,3\}$  apesar da versão ABC assintótica apresentar as estimativas para os parâmetros distantes do valor real, a estimativa para  $p^*$  ficou bem próximo do real valor de  $p^*$ , fato, esse, explicitado pela Figura 1.5. Ou seja, o método ABC assintótico foi capaz de recuperar valores para  $p^*$  próximos do real valor de  $p^*$  para diferentes valores dos parâmetros. Mais uma vez os outros métodos, ou não geraram amostras, ou as estimativas dos parâmetros ficaram bem distantes, e as estimativas de  $p^*$  apresentaram divergências no resultado de 16% a 20% do real valor de  $p^*$ .
- O tempo de processamento para o método ABC Lenormand, versão assintótica, é no mínimo 4 vezes menor que o tempo dos demais métodos considerando todos os cenários.

Tabela 4.3: Resultado grafo exponencial  $\theta_1 = 0,8$  e  $\theta_2 = 0,02$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,361</b>	<b>0,223</b>	<b>0,844</b>	<b>0,841</b>	0,999	<b>28 min</b>
Lenormand, versão Bayesiana	0,344	0,031	0,844	0,685	1	2h 48min
Bergm	0,374	0,037	0,844	0,702	0,117	2h 18min

Tabela 4.4: Resultado grafo exponencial  $\theta_1 = 0,2$  e  $\theta_2 = 0,2$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,268</b>	<b>0,166</b>	<b>0,743</b>	<b>0,749</b>	1	<b>32 min</b>
Lenormand, versão Bayesiana	0,365	0,032	0,743	0,695	1	2h 50min
Bergm	-	-	0,743	-	-	+2h

Tabela 4.5: Resultado grafo exponencial  $\theta_1 = 0,5$  e  $\theta_2 = 0,35$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,495</b>	<b>0,403</b>	<b>0,947</b>	<b>0,962</b>	1	<b>34 min</b>
Lenormand, versão Bayesiana	0,43	0,072	0,947	0,751	0,999	3h 51min
Bergm	-	-	0,947	-	-	+2h

Tabela 4.6: Resultado grafo exponencial  $\theta_1 = 0,9$  e  $\theta_2 = 0,02$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,376</b>	<b>0,225</b>	<b>0,869</b>	<b>0,849</b>	0,91	<b>29 min</b>
Lenormand, versão Bayesiana	0,369	0,032	0,869	0,696	1	3h 17min
Bergm	0,419	0,037	0,869	0,722	0,481	2h 13min

Tabela 4.7: Resultado grafo exponencial  $\theta_1 = 0,6$  e  $\theta_2 = 0,1$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,42</b>	<b>0,268</b>	<b>0,835</b>	<b>0,889</b>	0,977	<b>27 min</b>
Lenormand, versão Bayesiana	0,377	0,04	0,835	0,706	1	2h 53min
Bergm	-	-	0,835	-	-	+2h

Tabela 4.8: Resultado grafo exponencial  $\theta_1 = 0,7$  e  $\theta_2 = 0,3$  : **50** nós

Método	$\theta_1$	$\theta_2$	$p^*$ definido	$p^*$ estimado	Taxa	Tempo
Lenormand, versão assintótica	<b>0,518</b>	<b>0,412</b>	<b>0,954</b>	<b>0,966</b>	0,91	<b>30 min</b>
Lenormand, versão Bayesiana	0,444	0,075	0,954	0,759	1	3h 39min
Bergm	-	-	0,954	-	-	+2h



## Capítulo 5

# Conclusões e Trabalhos Futuros

Até o momento final de desenvolvimento desta dissertação não identificamos nenhum artigo que apresente um estudo sobre o uso do método ABC para estimação de modelos ERGM. Nosso trabalho explorou o uso dos métodos ABC mais recentes da literatura e registrou a sua eficácia nesse contexto. Essa é, sem dúvida, uma das nossas grandes contribuições.

Outra grande contribuição foi a identificação do método ABC como um método mais eficiente que o método Bayesiano tanto no que diz respeito à qualidade do ajuste, quanto ao tempo de processamento. Além disso, o método ABC se mostrou eficaz em situações onde o método Bayesiano não foi capaz de apresentar resultados.

Os resultados demonstram que a nossa proposta de utilização do método ABC aplicando a a teoria assintótica de [Diaconis e Chatterjee 2013](#) se mostrou um grande diferencial em relação aos métodos Bayesiano (Bergm) e ao método ABC que utiliza a simulação de dados Bayesiana, pois mesmo quando os parâmetros estimados estão distantes dos parâmetros reais, o valor de  $p^*$  estimado pelo nosso algoritmo está bem próximo do valor real de  $p^*$ , que é um fato conhecido ([Diaconis e Chatterjee 2013](#)) e explicitado pelas curvas de nível da Figura 1.5. Com esse ganho obtido por nossa proposta torna-se viável a aplicação do método em redes com grande número de arestas, qualificando-o como um método para ser aplicado em redes reais, pois as redes reais podem ser compostas de milhares e, até, milhões de nós.

Considerando os resultados do nosso estudo, identificamos as seguintes atividades para análise futura :

- Estudar a inferência do modelo para os casos onde  $(\theta_1 < 0 \text{ e } \theta_2 > 0)$ , pois [Diaconis e Chatterjee 2013](#) destacam a existência de transição de fase na região do espaço paramétrico onde  $\theta_1 < 0$ , ou seja, região na qual ocorre uma descontinuidade nos valores de  $u^*$  ocorrendo um salto de 0 para 1.
- Identificar métricas para qualidade do ajuste para grafos aleatórios exponenciais com mais de um parâmetro (BIC - *Bayesian Information Criterion*, AIC - *Akaike Information Criterion*, DIC - *Deviance Information Criterion*, KL - *kullback Leibler Divergence*, etc);
- Efetuar estudos das técnicas utilizadas nesse trabalho em redes **reais**.
- Aplicar o método ABC para estimativa de parâmetros de outros modelos que também tenham a constante de normalização intratável, uma vez que ele se mostrou eficiente para modelos ERGM.

- Aprofundar o estudo do método Bayesiano implementado pelo pacote **R Bergm** para identificar os motivos pelos quais não conseguimos gerar amostras para a maioria dos cenários.
- Aplicar o método ABC considerando outras configurações do modelo ERGM, ou seja, grau, transitividade,  $k$ -estrelas, etc.

# Referências Bibliográficas

- [Beaumont et al. 2002] Beaumont, M. A., Zhang, W., Balding, D. J.: Approximate Bayesian computation in population genetics. *Genetics*. 162(4),2025-2035 (2002)
- [Beaumont et al. 2009] Beaumont, M. A., Cornuet, J., Marin, J., and Robert, C P (2009) "Adaptive approximate Bayesian computation." *Biometrika*, **96**, 983-990.
- [Begum-Munni et al 2014] Begum, Munni, et al. "Network motif identification and structure detection with exponential random graph models." *Network Biology* 4.4 (2014): 155.
- [Besag 1974] Besag, J. E. (1974), "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *Journal of the Royal Statistical Society, Series B*, 36, 192-236.
- [Blum e Francois 2010] Blum, M. G. B. François, O.: Non-linear regression models for approximate Bayesian computation. *Stat. Comput.* 20(1),63-73 (2010)
- [Caimo e Friel 2011] Caimo, A. and Friel, N. (2011). "Bayesian inference for exponential random graph models." *Social Networks* **33** 41–55.
- [Caimo and Friel 2011] Caimo, A. and Friel, N. (2011). "Bayesian inference for exponential random graph models." *Social Networks* **33** 41–55.
- [Cranmer-Skyler and Bruce 2011] Cranmer, Skyler J., and Bruce A. Desmarais. "Inferential network analysis with exponential random graph models." *Political Analysis* 19.1 (2011): 66-86.
- [Del Moral et al 2006] Del Moral, Pierre, Arnaud Doucet, and Ajay Jasra. "Sequential monte carlo samplers." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68.3 (2006): 411-436.
- [Del Moral et al. 2012] Del Moral, P., Doucet, A., and Jasra, A. (2012) "An adaptive sequential Monte Carlo method for approximate Bayesian computation". *Statistics and Computing*, 22, 1009-1020.
- [Diaconis e Chatterjee 2013] P. Diaconis and S. Chatterjee (2013), Estimating and Understanding Exponential Random Graph Models. *Ann. Statist.*, 41(5):2428-2461
- [Drovandi e Pettitt 2011] Drovandi, C. C., Pettitt, A. N.: Estimation of parameters for macroparasite population evolution using approximate Bayesian computation. *Biometrics*. 67(1),225-233 (2011)
- [Drovandi and Pettitt 2011] Drovandi, C. C., Pettitt, A. N.: Estimation of parameters for macroparasite population evolution using approximate Bayesian computation. *Biometrics*. 67(1),225-233 (2011)

- [Erdos-Renyi 1959] Erdos, P., Renyi, A. (1959). On random graphs. *Publicationes Mathematicae (Debrecen)*, 6, 290– 297.
- [Fienberg 2012] Fienberg, S.E. (2012). A Brief History of Statistical Models for Network Analysis and Open Challenges. *Journal of Computational and Graphical Statistics*, 21: 825- 839
- [Frank e Strauss 1986] Frank, O., Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81( 395), 832– 842.
- [Frank 1991] Frank, O. (1991). Statistical analysis of change in networks. *Statistica Neerlandica* 45, 283-293.
- [Freeman 1979] Freeman, L. C. (1979). Centrality in networks: I. conceptual clarification. *Social Networks* 1, 215-239.
- [Geyer e Thompson 1992] Geyer, C. J. and Thompson, E. A. (1992), "Constrained Monte Carlo maximum likelihood for dependent data (with discussion)," *Journal of the Royal Statistical Society, Series B*, 54, 657-699.
- [Gulati and Gargiulo 1999] Gulati, R. and M. Gargiulo (1999). "Where do interorganizational networks come from?" *American Journal of Sociology* 104, 1439-1493.
- [Handcock 2003] Handcock, M. S. (2003), "Assessing Degeneracy in Statistical Models of Social Networks," Working Paper no.39, Center for Statistics and the Social Sciences, University of Washington.
- [Holland and Leinhardt 1981] Holland PW, Leinhardt S (1981). "An Exponential Family of Probability Distributions for Directed Graphs." *Journal of the American Statistical Association*, **76**(373), 33-65.
- [Hunter et al. 2008] Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M (2008b). "ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks." *Journal of Statistical Software*, 24(3), 1-29. URL <http://www.jstatsoft.org/v24/i03>.
- [Corander et al. 1998] J. Corander, K. Dahmstrom, and P. Dahmstrom. Maximum likelihood estimation for markov graphs. *Research Report, Department of Statistics, University of Stockholm*, 8, 1998.
- [Scott et al] John Scott, Peter J. Carrington. "The SAGE Handbook of Social Network Analysis". Capítulo 32, por Garry Robins, 484-486.
- [Lenormand et al 2012] Lenormand, M., Jabot, F., Deffuant G. (2012) "Adaptive approximate Bayesian computation for complex models". <http://arxiv.org/pdf/1111.1308.pdf>
- [Kolaczyk and Csardi 2009] Erik D. Kolaczyk, Gabor Csardi, "Statistical Analysis of Network Data with R "(Springer, London, 2009)
- [Koskinen 2004] Koskinen, J. H. (2004), "Bayesian Analysis of Exponential Random Graphs - Estimation of Parameters and Model Selection," *Research Report 2004:2, Department of Statistics, Stockholm University*.



- [Koskinen et al 2012] Lusher, Dean; Koskinen, Johan; Robins, Garry (2012-11-30). Exponential Random Graph Models for Social Networks (Structural Analysis in the Social Sciences). Cambridge University Press.
- [Krackhardt 1987] Krackhardt, D. (1987). "Qap partialling as a test of spuriousness". *Social Networks* 9, 171-186.
- [Krackhardt 1988] Krackhardt, D. (1988). "Predicting with networks - nonparametric multiple regression analysis of dyadic data". *Social Networks* 10 (4), 359-381.
- [Dekker et al 2007] Dekker, D., D. Krackhardt, and T. A. B. Snijders (2007). "Sensitivity of mrqap tests to collinearity and autocorrelation conditions". *Psychometrika* 72, 563-581.
- [Lazega et al. 2004] Lazega, E., Mounier, L., Stofer, R., Tripiet, A., 2004. Discipline scientifique et discipline sociale: Réseaux de conseil, apprentissage collectif et innovation dans la recherche française sur le cancer (1997–1999). *Recherches Sociologiques* 35, 3–27
- [Lehmann e Romano 2005] Lehmann, E. and J. Romano (2005). *Testing Statistical Hypotheses* (3rd ed.). New York: Springer.
- [Lindgren 2010] Lindgren, K.-O. (2010). "Dyadic regression in the presence of heteroscedasticity - an assessment of alternative approaches". *Social Networks* 32, 279-289.
- [Lusher-Koskinen and Robins 2012] D. Lusher, J. Koskinen, G. Robins, "Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications "(Cambridge University Press, Cambridge, 2012)
- [Marjoram et al. 2003] Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S.: Markov chain Monte Carlo without likelihoods. *P. Natl. Acad. Sci. USA*. 100(26),15324-15328 (2003)
- [Møller et al. 2006] Møller, Jesper, et al. "An efficient Markov chain Monte Carlo method for distributions with intractable normalising constants." *Biometrika* 93.2 (2006): 451-458.
- [Morris et al. 2008] Morris, M., Handcock, M. S., and Hunter, D. R. (2008), "Specification of Exponential- Family Random Graph Models: Terms and Computational Aspects," *Journal of Statistical Software*, 24.
- [Murray et al. 2006] Murray, I., Ghahramani, Z., and MacKay, D. (2006), "MCMC for doubly-intractable distributions," in *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence (UAI-06)*, Arlington, Virginia: AUAI Press.
- [Pritchard et al. 1999] Pritchard, J.K., and M.T. Seielstad and A. Perez-Lezaun and M.W. Feldman (1999) Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, 16, 1791-1798.
- [Roberts e Gilks 1994] Roberts, G. O. and Gilks, W. R. (1994), "Convergence of adaptive direction sampling," *Journal of Multivariate Analysis*, 49, 287-298.
- [Robins et al 2007] Robins G, Pattison P, Kalish Y, Lusher D (2007a). "An Introduction to Exponential Random Graph (p\*) Models for Social Networks." *Social Networks*, 29(2), 173-191.

- [Sisson et al. 2007] Sisson S.A, Fan Y, Tanaka M.M 2007. "Sequential Monte Carlo without likelihoods." *Proc. Natl Acad. Sci. USA*. 104, 1760–1765
- [Snijders 2011] Snijders TAB. 2011. Statistical models for social networks. *Annu. Rev. Sociol.* 37:131–53
- [Snijders et al. 2006] Snijders, T. A. B., Pattison, P. E., Robins, G. L. and Handcock, M. S. (2006). "New specification for exponential random graph models." *Sociological Methodology* **36** 99–153.
- [Strauss e Ikeda 1990] Strauss, D. and Ikeda, M. (1990), "Pseudolikelihood estimation for social networks," *Journal of the American Statistical Association*, 5, 204-212.
- [Toivonen-Riitta et al 2009] Toivonen, Riitta, et al. "A comparative study of social network models: Network evolution models and nodal attribute models." *Social Networks* 31.4 (2009): 240-254.
- [Toni et al. 2009] Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M. P. H.: Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. Roy. Soc. Interface*. 6,187 (2009)
- [van Duijn et al. 2004] van Duijn, M. A. J., Snijders, T. A. B., Zijlstra, B. (2004). A random effects model with covariates for directed graphs. *Statistica Neerlandica*, 58( 2), 234– 254.
- [Wasserman e Pattison 1996] Wasserman, S. and P. Pattison (1996). Logit models and logistic regression for social networks: I. An introduction to Markov graphs and  $p^*$ . *Psychometrika* 61, 401-425.
- [Wang e Yves 2014] Wang, Jing, and Yves F. Atchadé. "Approximate Bayesian computation for exponential random graph models for large social networks." *Communications in Statistics-Simulation and Computation* 43.2 (2014): 359-377.
- [Wasserman e Faust 1994] Wasserman, Stanley, and Katherine Faust. *Social network analysis: Methods and applications*. Vol. 8. Cambridge university press, 1994.
- [White 1980] White, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity". *Econometrica* 48, 817-830.
- [Wikipedia KL] Wikipedia, Kullback–Leibler divergence, "[https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler\\_divergence](https://en.wikipedia.org/wiki/Kullback%E2%80%93Leibler_divergence)"
- [Wikipedia KS] Wikipedia, Kolmogorov–Smirnov test, "[https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov\\_test](https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test)"
- [Wilkinson 2008] Richard D. Wilkinson (2008) "Approximate Bayesian computation (ABC) gives exact results under the assumption of model error". <http://arxiv.org/abs/0811.3355>

## Apêndice A

# Resultados do Estudo de Simulação para Grafos Bernoulli : Caso Erdős-Renyi

Apresentamos, a seguir, todos os resultados para o estudo de simulação dos "Grafos Bernoulli".

Os resultados estão ordenados por método.

Tabela A.1: Resultado grafo aleatório Bernoulli para **50** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Beaumont	0,725973	63	0,000502396	0,134	0,000000032
Bergm	0,726381	40	0,000297269	0,013640553	0,265006713
Delmoral	0,7286	5	0,000282759	0,082	0,002403443
Drovandi	0,726654	83	0,000330253	0,025	0,913475513
Lenormand	0,727054	199	0,000277959	0,062	0,042815177
Marjoram	0,726422	207	0,000319168	0,0201875	0,002945681
Rejeicao	0,726294	171	0,000294566	0,026	0,887938607

Tabela A.2: Resultado grafo aleatório Bernoulli para **100** nós

Modelo	Média a post.	Tempo(s)	KL	KS	KS-pvalue	ChiQ
Beaumont	0,730211	104	0,000337	0,262	0	0,000675
Bergm	0,730186	49	0,000092	0,054378	0	0,000184
Delmoral	0,730199	8	0,000076	0,069307	0,015634	0,000153
Drovandi	<b>0,730046</b>	114	<b>0,000074</b>	0,039	<b>0,432432</b>	<b>0,000149</b>
Lenormand	0,729674	210	0,000067	0,048	0,199518	0,000133
Marjoram	0,729727	373	0,000079	0,043938	0	0,000158
Rejeicao	<b>0,730077</b>	212	<b>0,000076</b>	0,029	<b>0,794394</b>	<b>0,000153</b>

Para completar a análise é necessário apresentar os .

Tabela A.3: Resultado grafo aleatório Bernoulli para **200** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,732618	17	0,00002337	0,085	0,001456305
Bergm	0,732017	80	0,000045945	0,17235023	0
Drovandi	0,73211	204	0,000018313	0,026	0,887938607
Beaumont	0,731653	264	0,000258131	0,374	0
Lenormand	0,732148	277	0,000016888	0,051	0,148338062
Marjoram	0,732015	1191	0,000026063	0,0839375	0
Rejeicao	0,732175	30792	0,000017663	0,028	0,827956861

Tabela A.4: Resultado grafo aleatório Bernoulli para **500** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,730809	117	0,000012227	0,337	0
Beaumont	0,730252	503	0,002432379	0,482	0
Lenormand	0,730445	650	0,000002789	0,072	0,011211077
Bergm	0,730649	758	0,00003501	0,348294931	0
Drovandi	0,730468	1966	0,000003291	0,024	0,935580992
Marjoram	0,73052	16185	0,000010176	0,2421875	0
Rejeicao	0,730466	64155	0,00000329	0,047	0,219330006

Tabela A.5: Resultado grafo aleatório Bernoulli para **750** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,730474	282	0,000015223	0,369	0
Beaumont	0,733812	604	0,002526835	0,515	0
Lenormand	0,730605	1470	0,000001071	0,048	0,199518349
Drovandi	0,730624	2632	0,000001371	0,035	0,57265485
Bergm	0,730785	2940	0,000033913	0,381013825	0
Marjoram	0,730687	40949	0,000008921	0,3183125	0
Rejeicao	0,730637	83599	0,000001419	0,043	0,313577535

Tabela A.6: Resultado grafo aleatório Bernoulli para **1.000** nós

Modelo	Média a post.	Tempo(s)	KL	KS	KS-pvalue
Beaumont	0,411085	838	0,008056	1	0
Bergm	0,730602	6.154	0,000036	0,4	0
Delmoral	<b>0,730156</b>	466	<b>0,000007</b>	0,433	0
Drovandi	0,730683	5.208	0,000001	0,043	0,313578
Lenormand	0,730698	2.921	0,000001	0,04	0,400474
Marjoram	0,73064	50.525	0,000009	0,363062	0
Rejeicao	<b>0,730681</b>	<b>110.745</b>	<b>0,000001</b>	0,047	<b>0,219330</b>

Tabela A.7: Resultado grafo aleatório Bernoulli para **1.250** nós

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,730922	614	0,000005968	0,38	0
Beaumont	0,263931	1131	0,019603734	1	0
Lenormand	0,730871	3468	0,000000449	0,052	0,133834304
Drovandi	0,730869	5355	0,000000561	0,067	0,022463708
Bergm	0,730698	11654	0,000034972	0,439539171	0
Rejeicao	0,730895	149385	0,00000056	0,04	0,400473666

Tabela A.8: Resultado grafo aleatório Bernoulli para **1.500 nós**

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,730239	936	0,000010815	0,439	0
Beaumont	0,183776	1500	0,040520403	1	0
Lenormand	0,730485	7080	0,000000301	0,024	0,935580992
Bergm	0,730664	16670	0,000034119	0,440184332	0
Drovandi	0,730484	23231	0,000000317	0,029	0,79439442

Tabela A.9: Resultado grafo aleatório Bernoulli para **1.750 nós**

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,730584	1289	0,000008302	0,423762376	0
Beaumont	0,133143	2119	0,087848962	1	0
Lenormand	0,730664	9174	0,000000222	0,052	0,133834304
Bergm	0,730827	25622	0,000033075	0,455668203	0
Drovandi	0,730668	25825	0,000000232	0,042	0,341001034

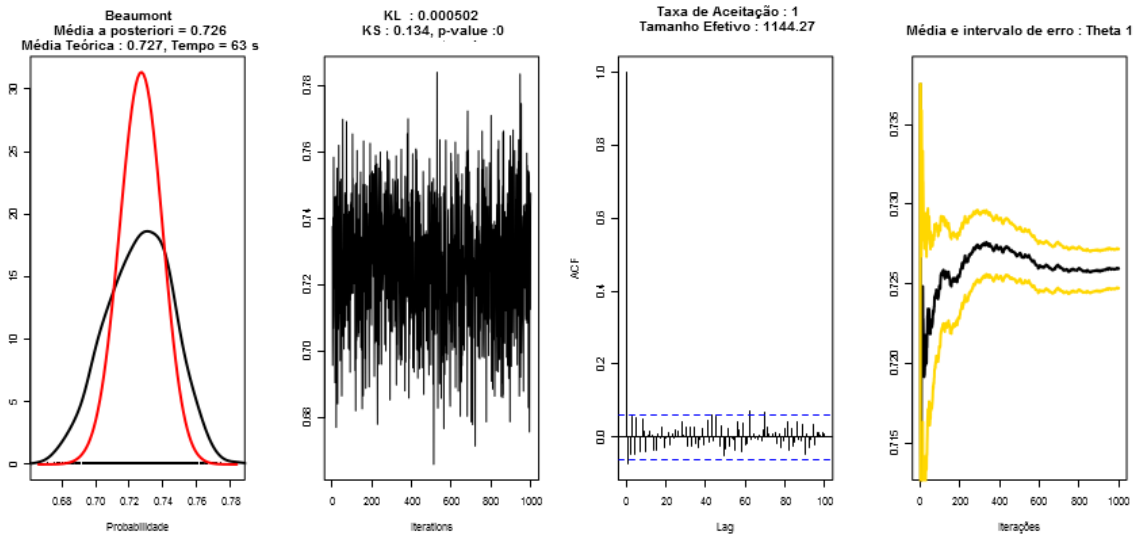
Tabela A.10: Resultado grafo aleatório Bernoulli para **2.000 nós**

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Delmoral	0,731434	1535	0,000010001	0,491089109	0
Beaumont	0,103624	2751	0,183073073	1	0
Drovandi	0,730972	9388	0,000000236	0,085	0,001456305
Lenormand	0,731	10241	0	0,058	0,069190326
Bergm	0,731076	46022	0,000032958	0,454654378	0

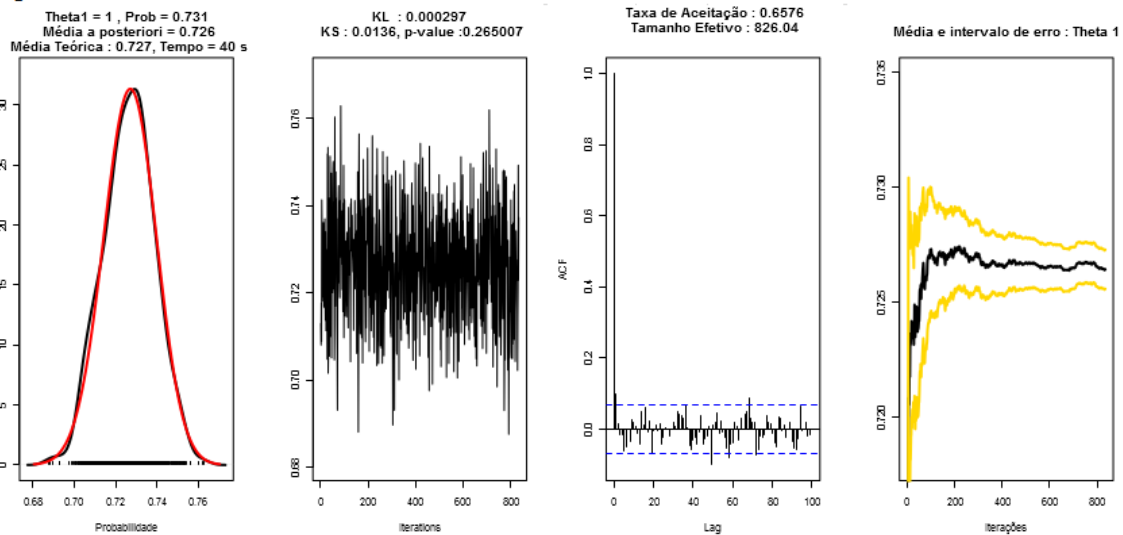
Tabela A.11: Resultado grafo aleatório Bernoulli para **2.250 nós**

Modelo	Média a post.	Tempo	KL	KS	KS-pvalue
Beaumont	0,092073	3344	0,222878529	1	0
Bergm	0,731024	65547	0,000035151	0,458986175	0
Delmoral	0,730905	1765	0,000008956	0,430693069	0
Drovandi	0,730872	13736	0,000000208	0,103	0,000049386
Lenormand	0,731	11260	0	0,058	0,069190326

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

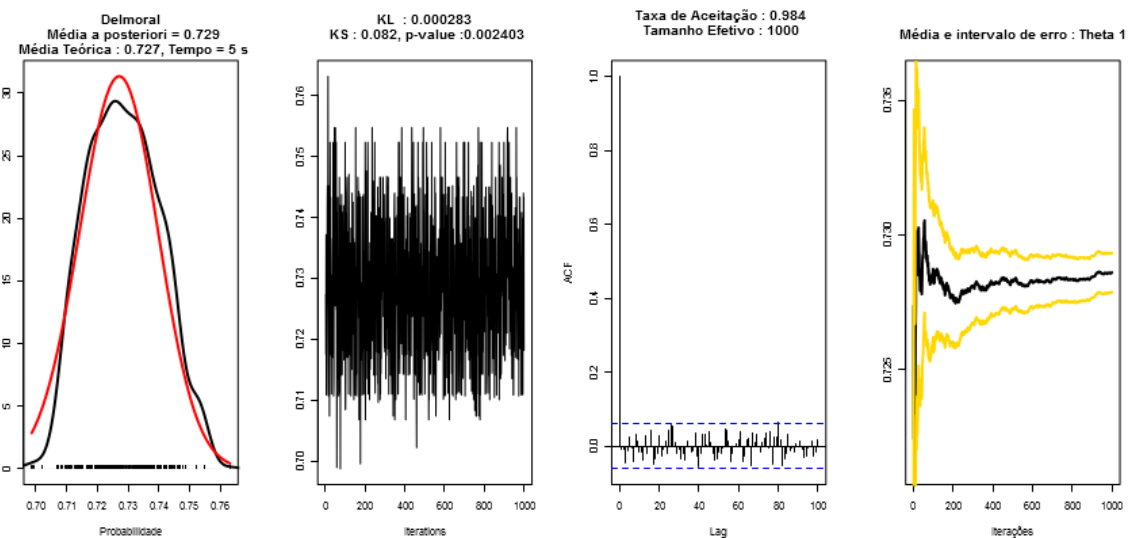
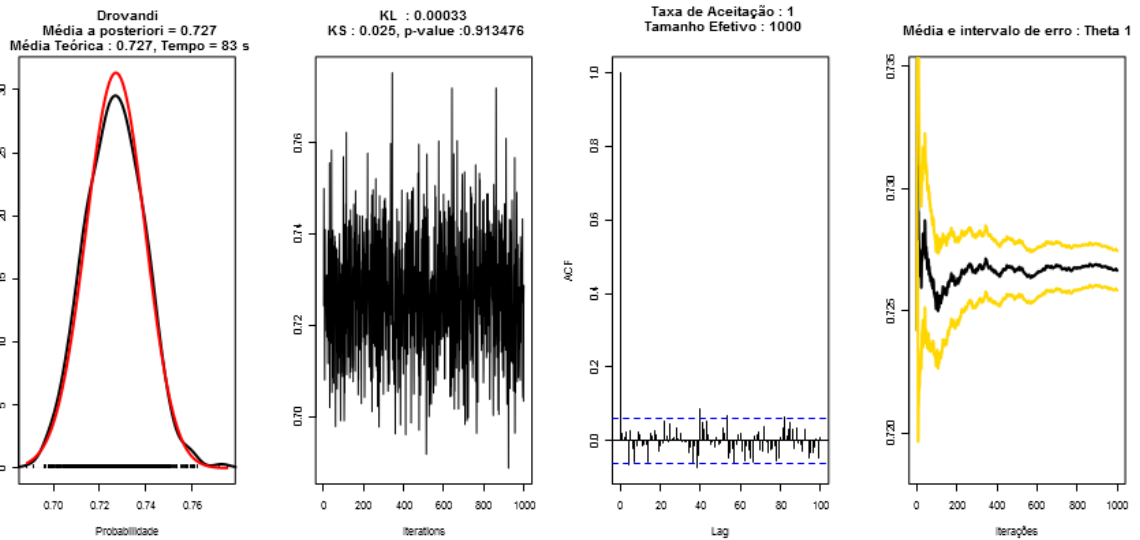
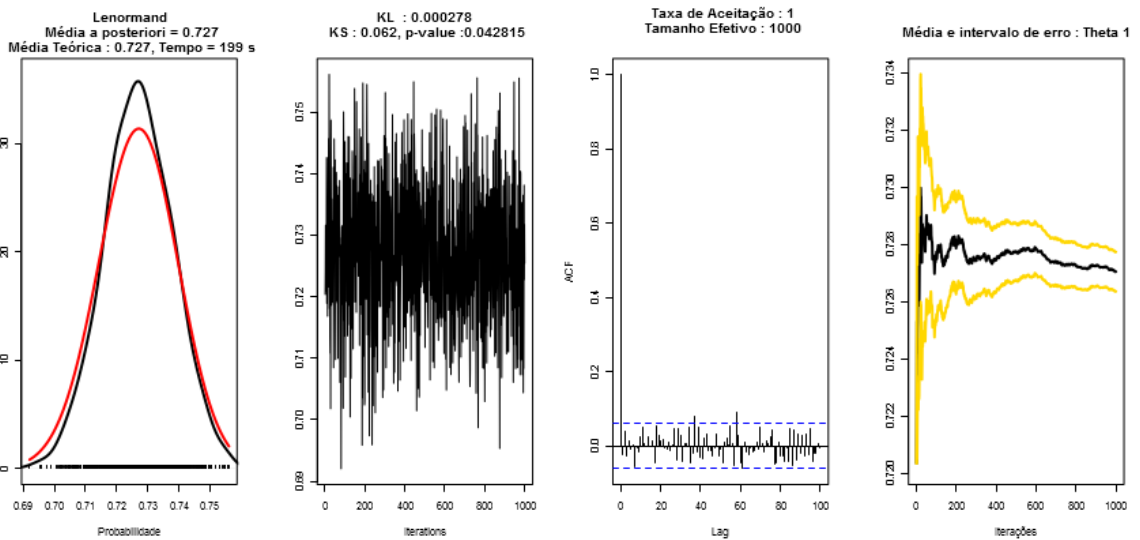


Figura A.1: Grafos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand



## ABC MCMC-Marjoram

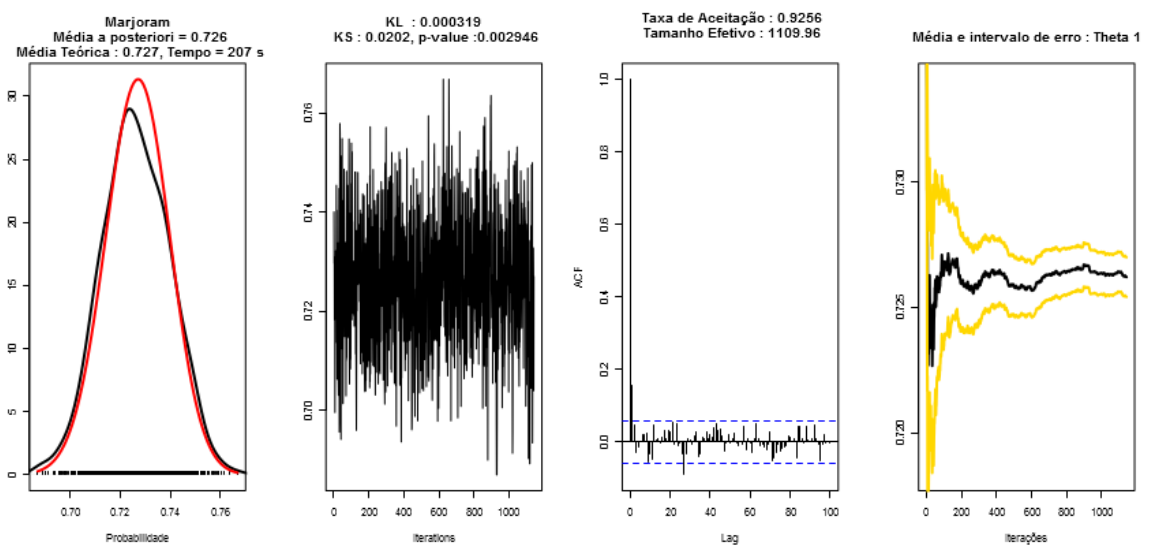


Figura A.2: Grafos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenorman e Marjoram.

ABC Rejeicao

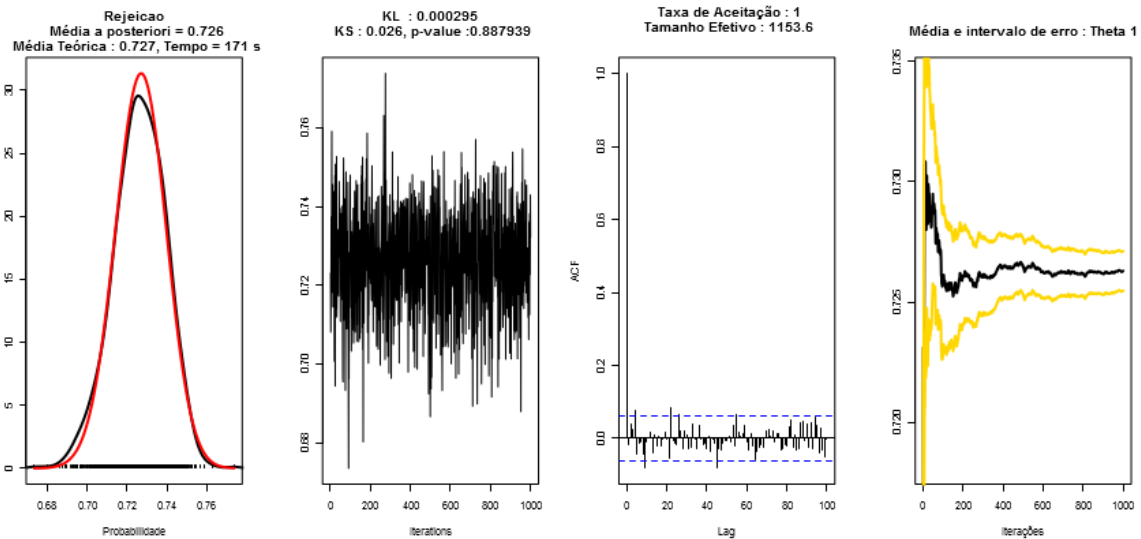
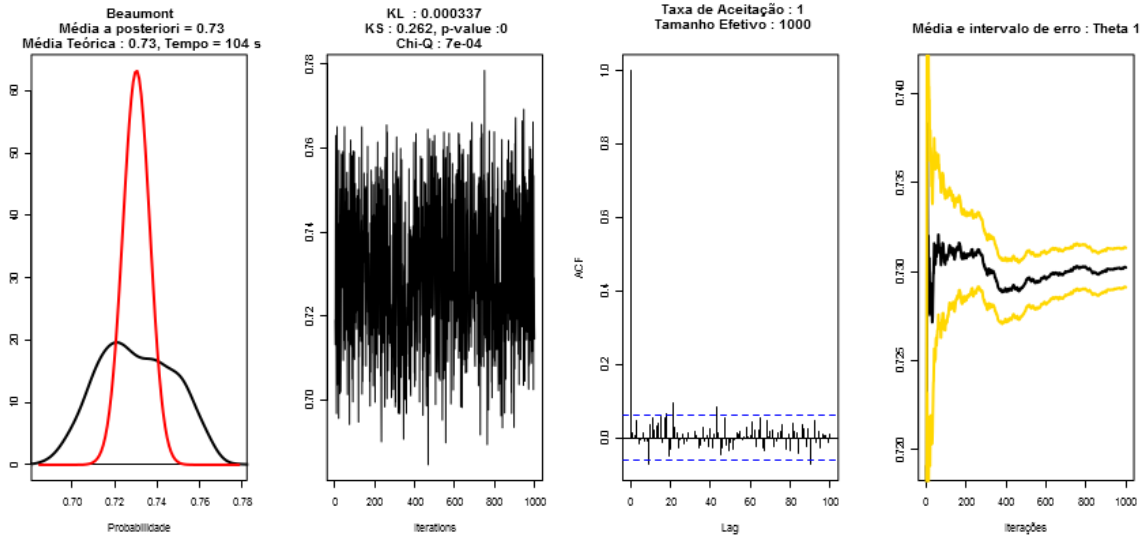


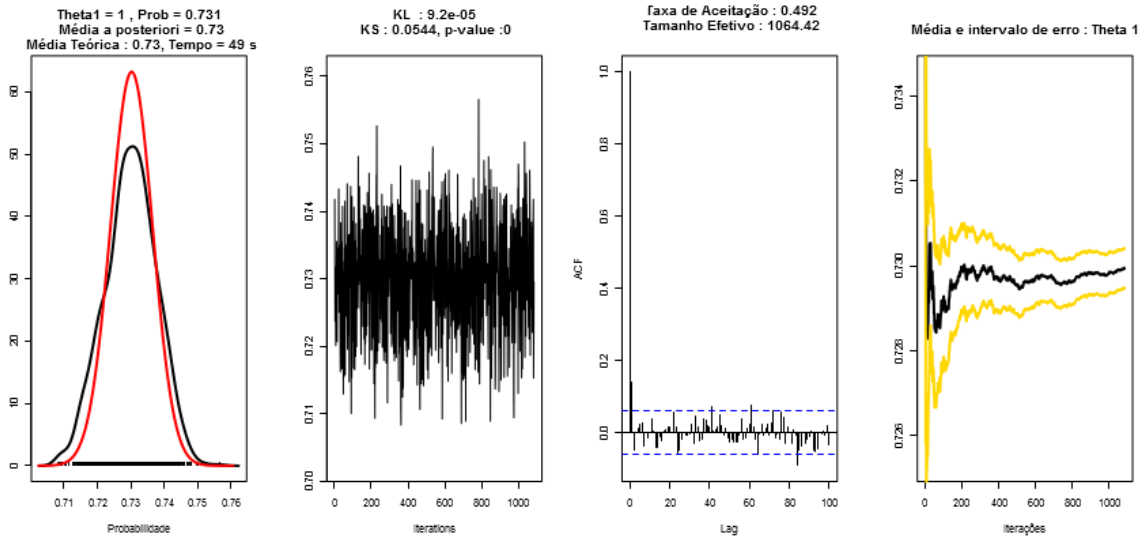
Figura A.3: Grafos aleatórios Bernoulli **50 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.



## ABC Sequencial-Beaumont



## Bergm



## ABC Sequencial-Delmoral

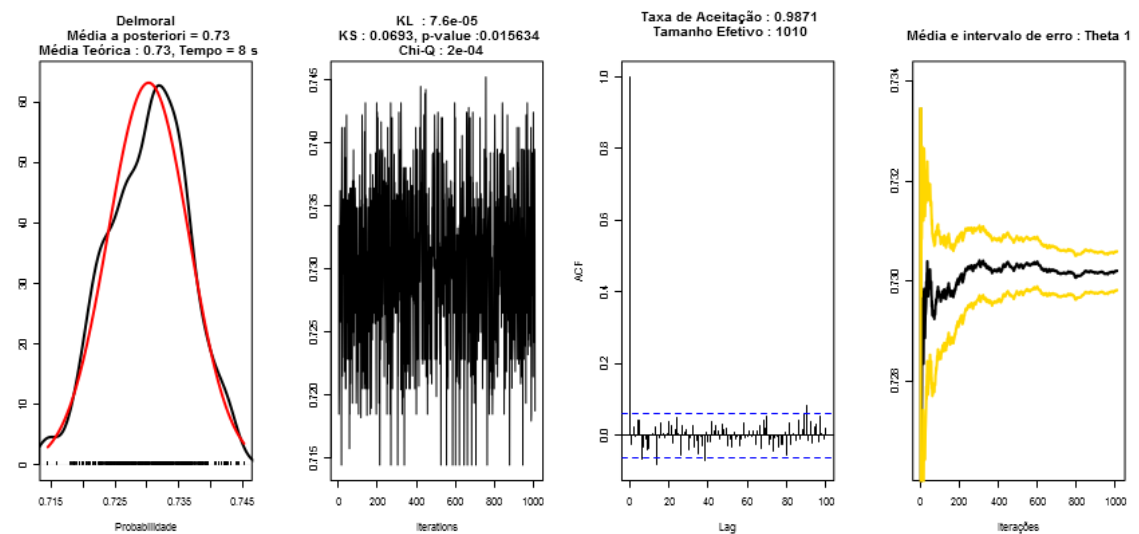
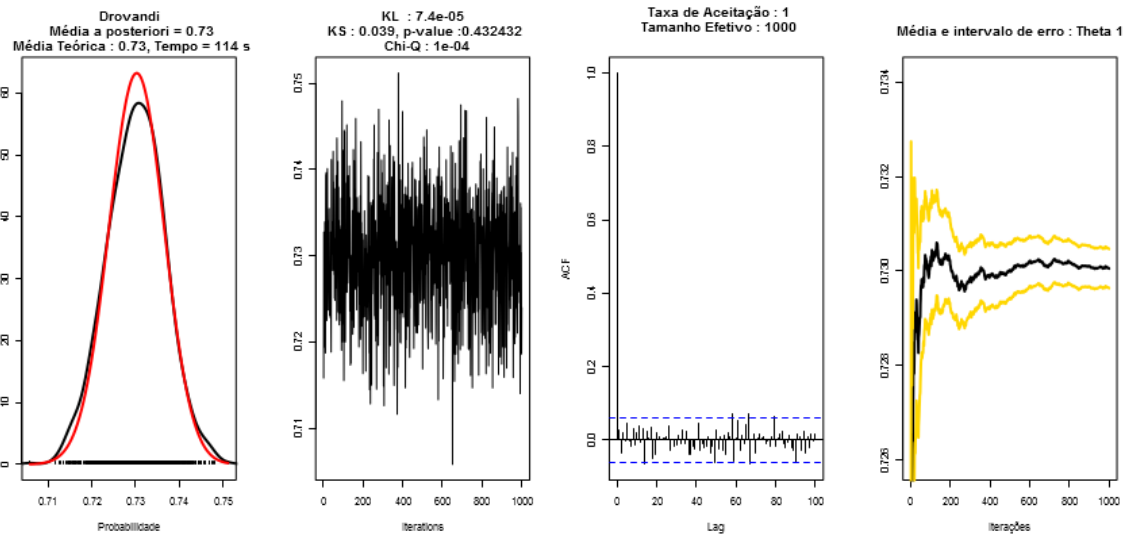
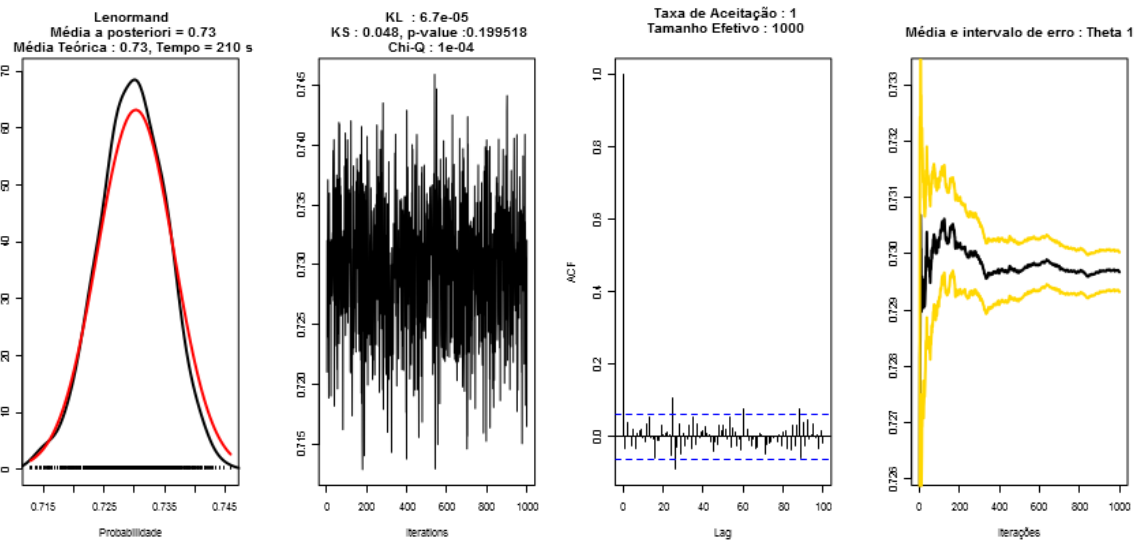


Figura A.4: Gráficos aleatórios Bernoulli **100 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

ABC Sequencial-Drovandi



ABC Sequencial-Lenormand



ABC MCMC-Marjoram

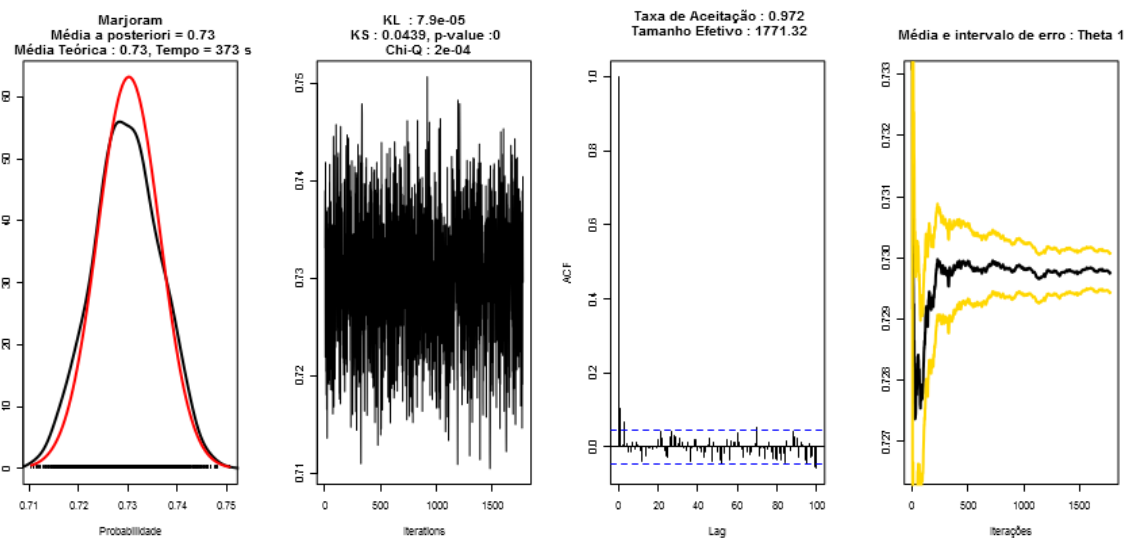


Figura A.5: Grafos aleatórios Bernoulli **100 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenormand e Marjoram.

## ABC Rejeicao

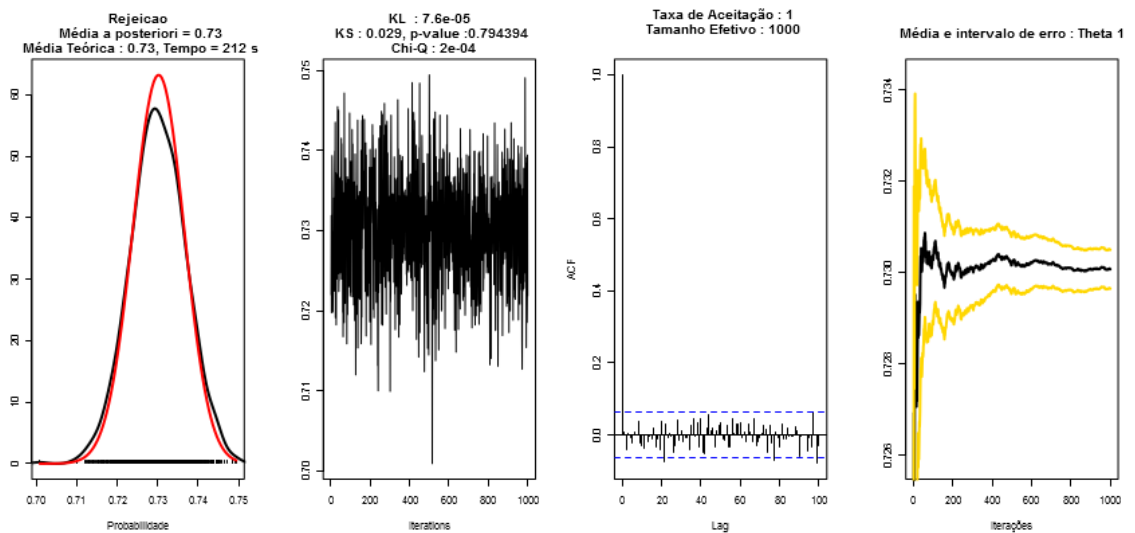
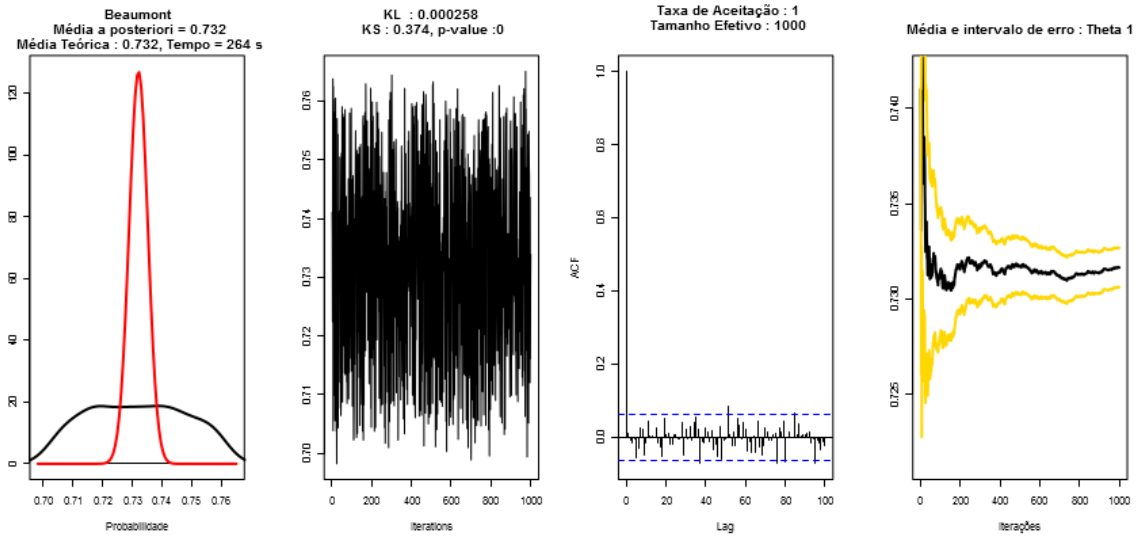
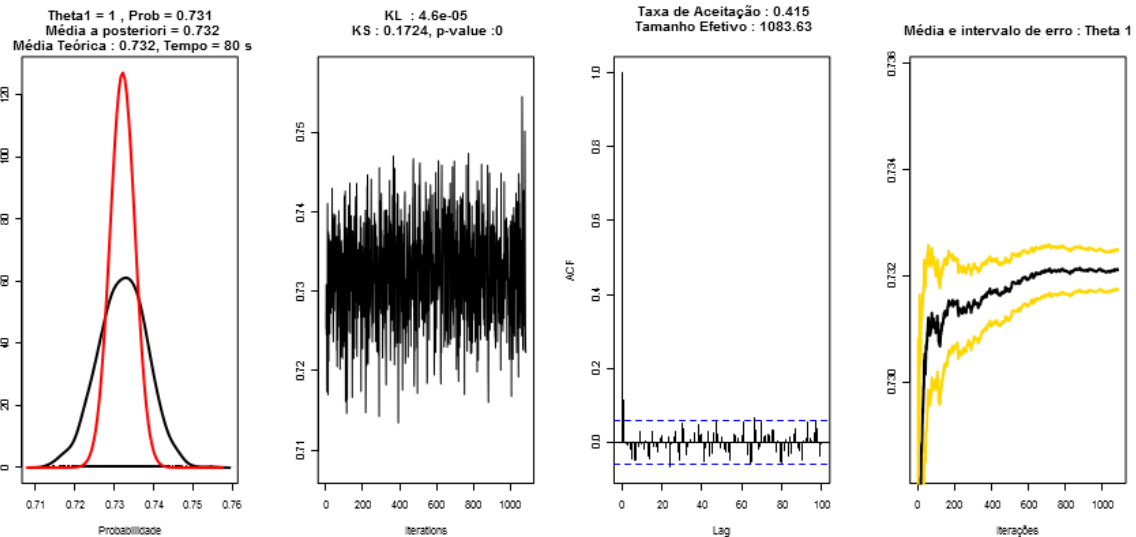


Figura A.6: Grafos aleatórios Bernoulli **100 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

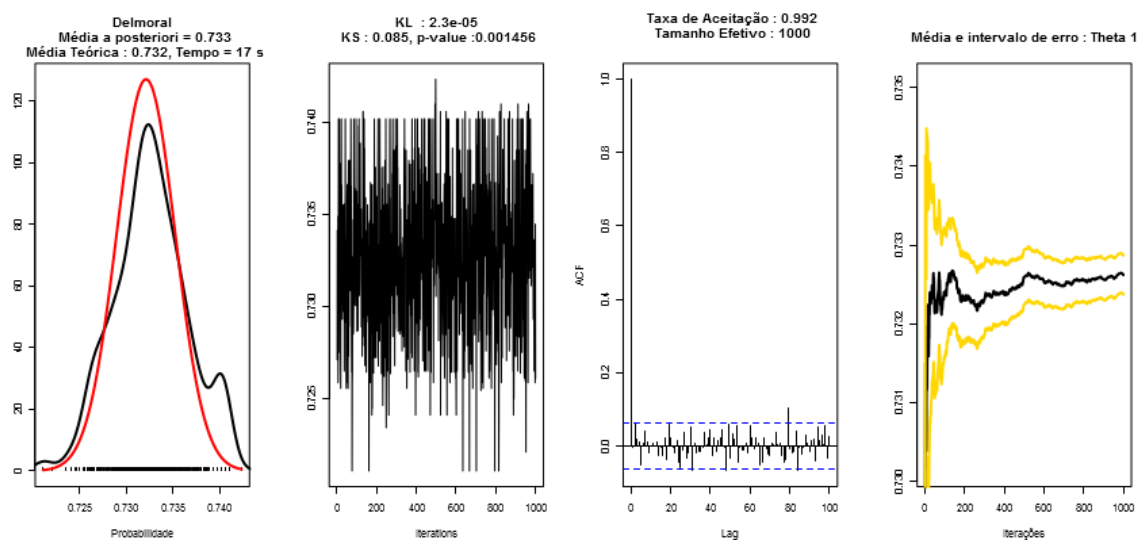
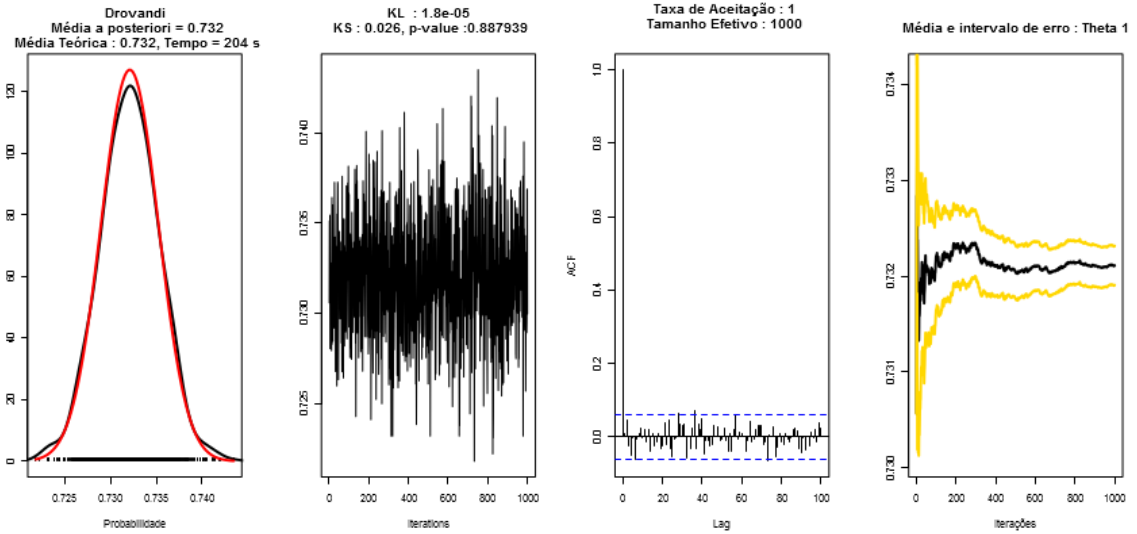
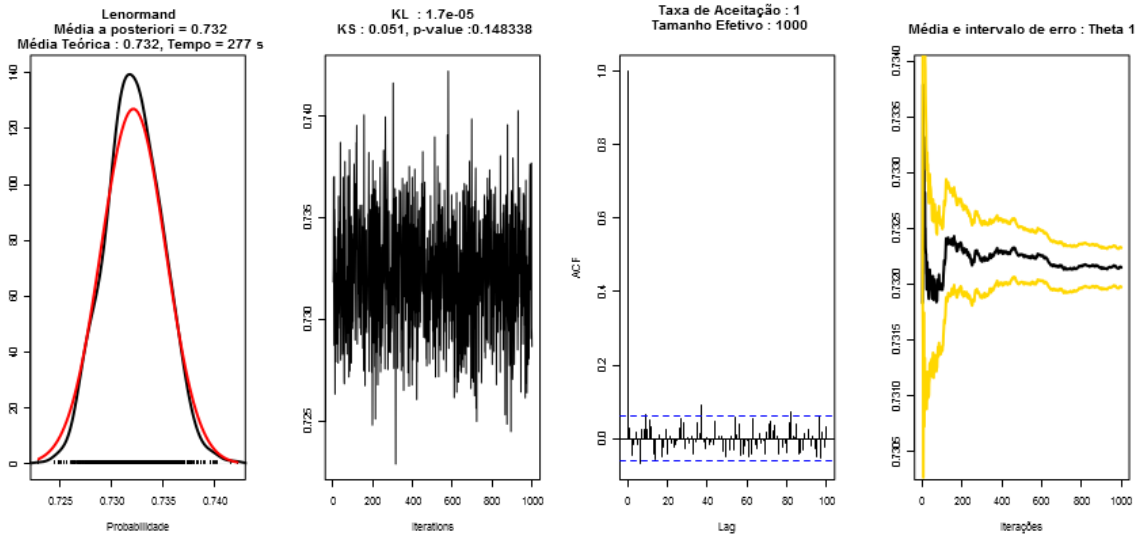


Figura A.7: Grafos aleatórios Bernoulli **200 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand



## ABC MCMC-Marjoram

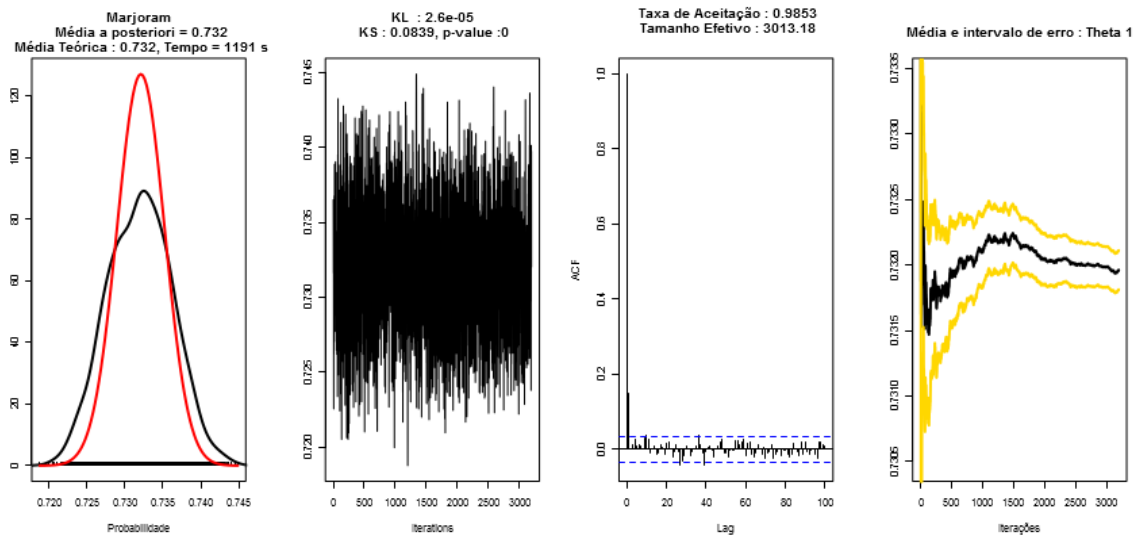


Figura A.8: Grafos aleatórios Bernoulli **200 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenorman e Marjoram.

ABC Rejeicao

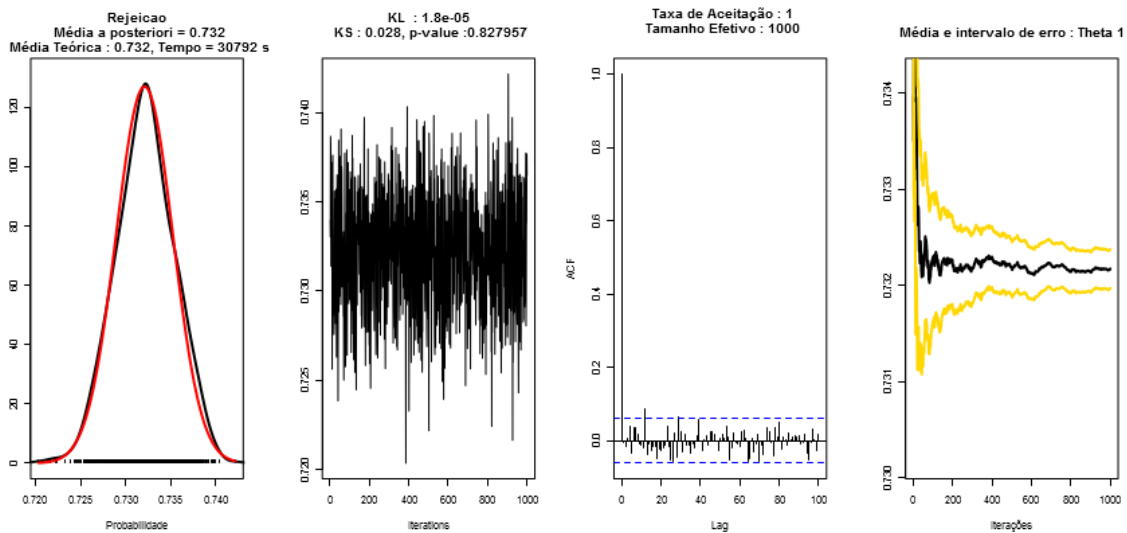
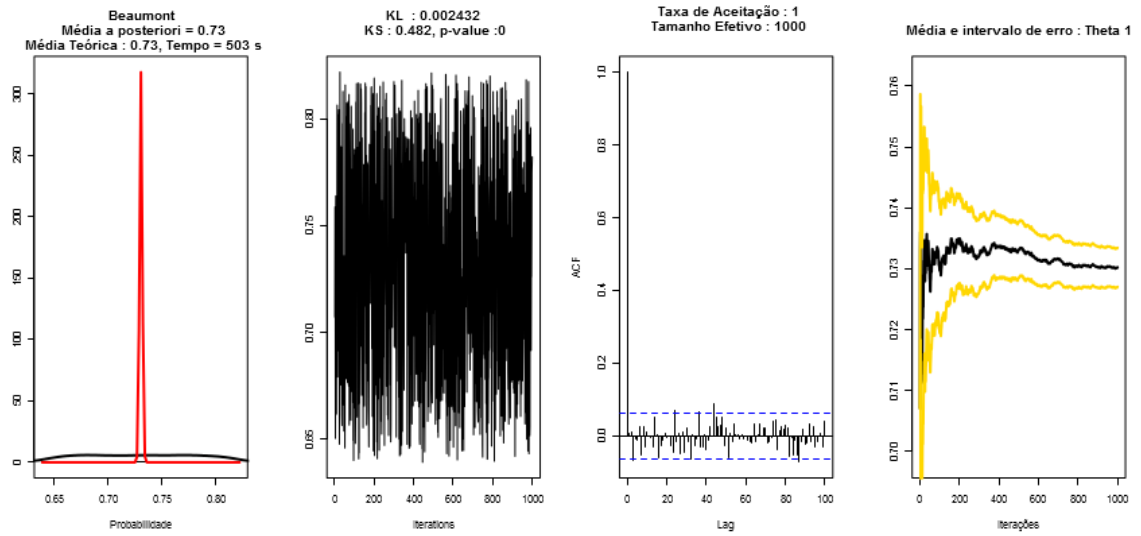
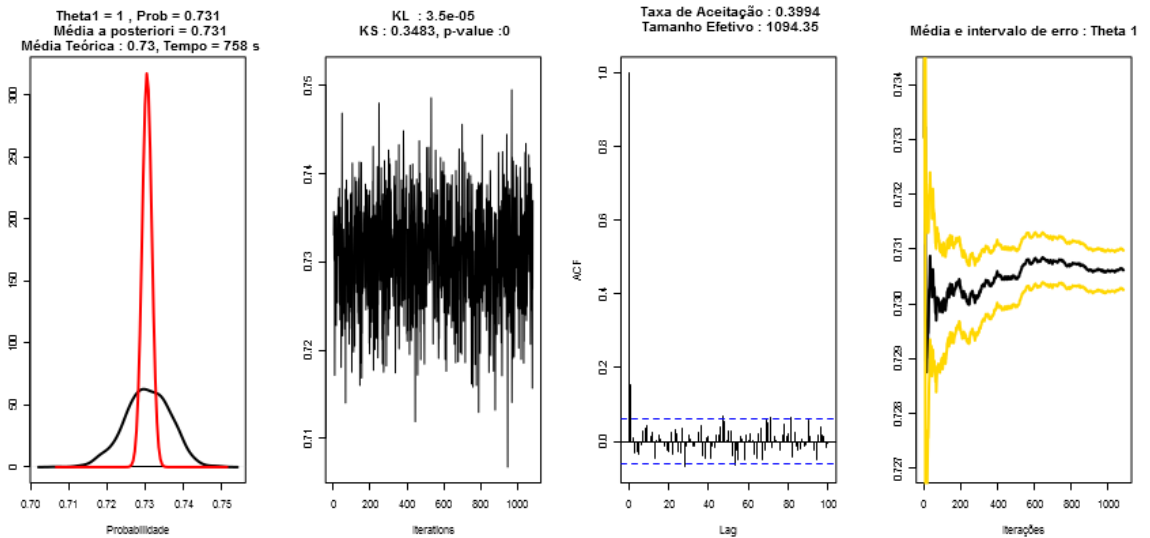


Figura A.9: Grafos aleatórios Bernoulli **200 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

## ABC Sequencial-Beaumont



## Bergm



## ABC Sequencial-Delmoral

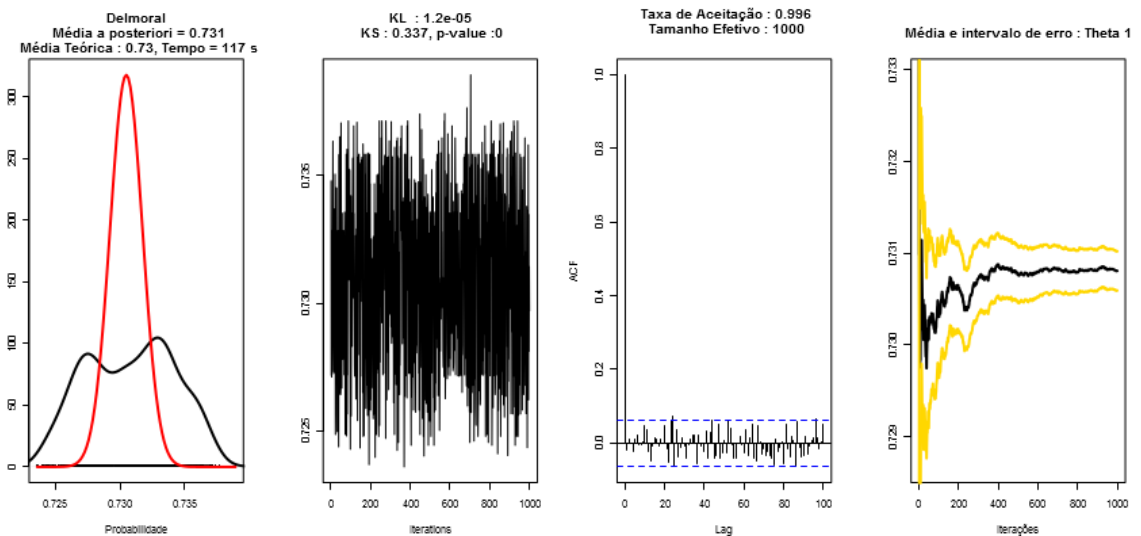
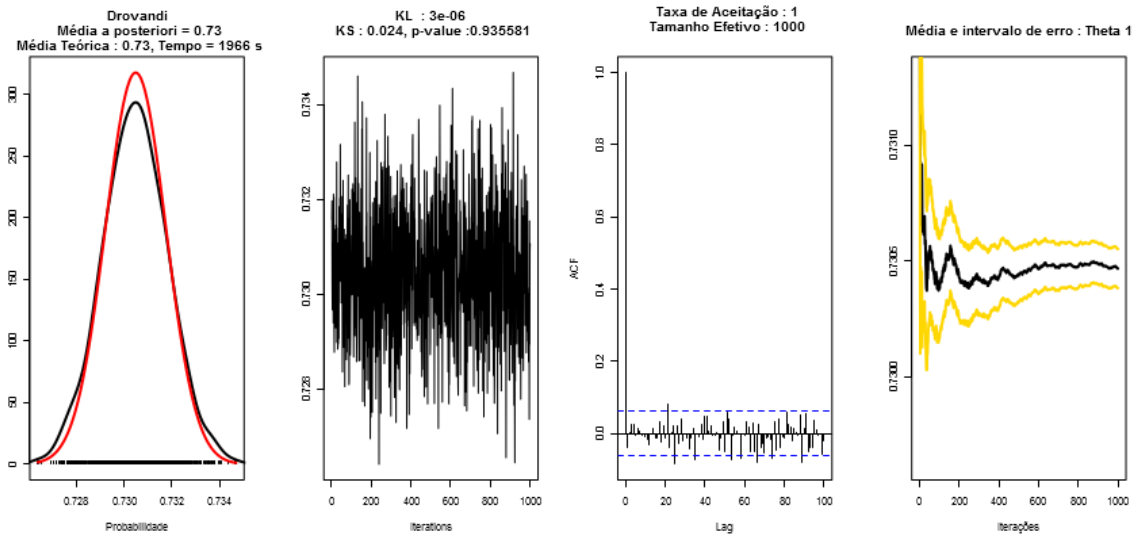
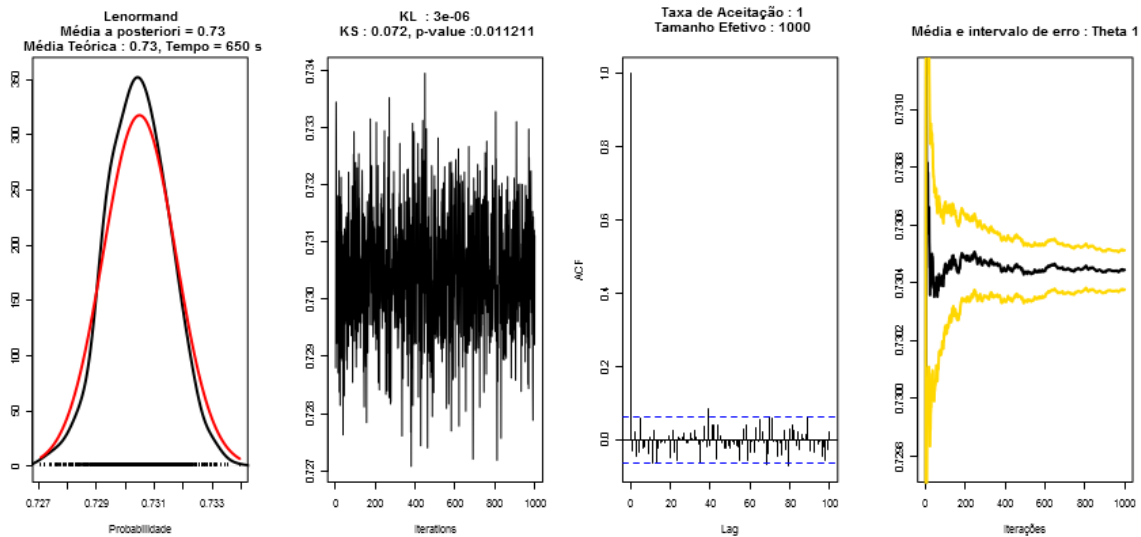


Figura A.10: Grafos aleatórios Bernoulli **500 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

ABC Sequencial-Drovandi



ABC Sequencial-Lenormand



ABC MCMC-Marjoram

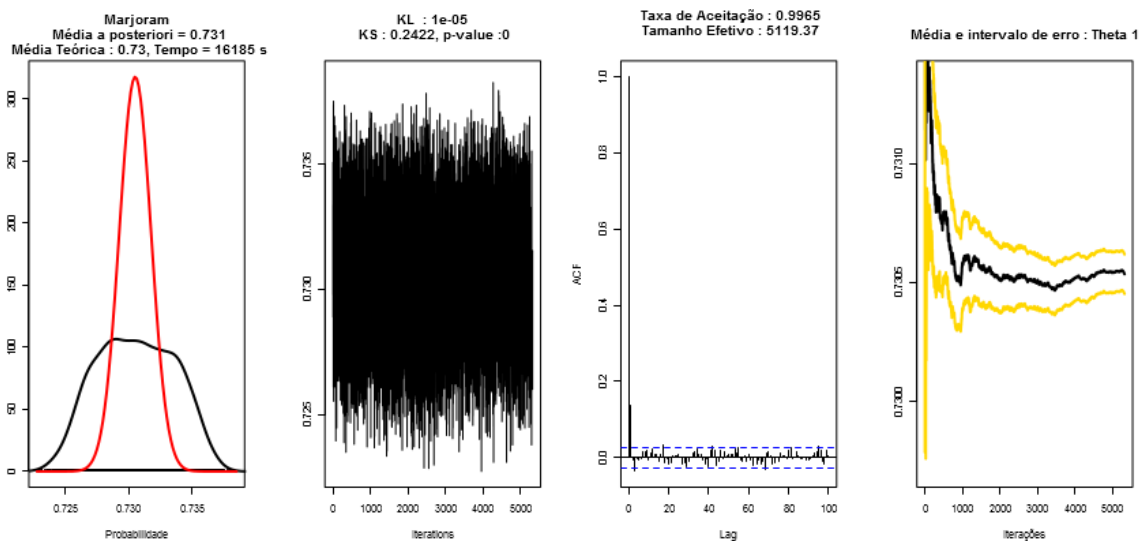


Figura A.11: Gráficos aleatórios Bernoulli **500 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenorman e Marjoram.



## ABC Rejeicao

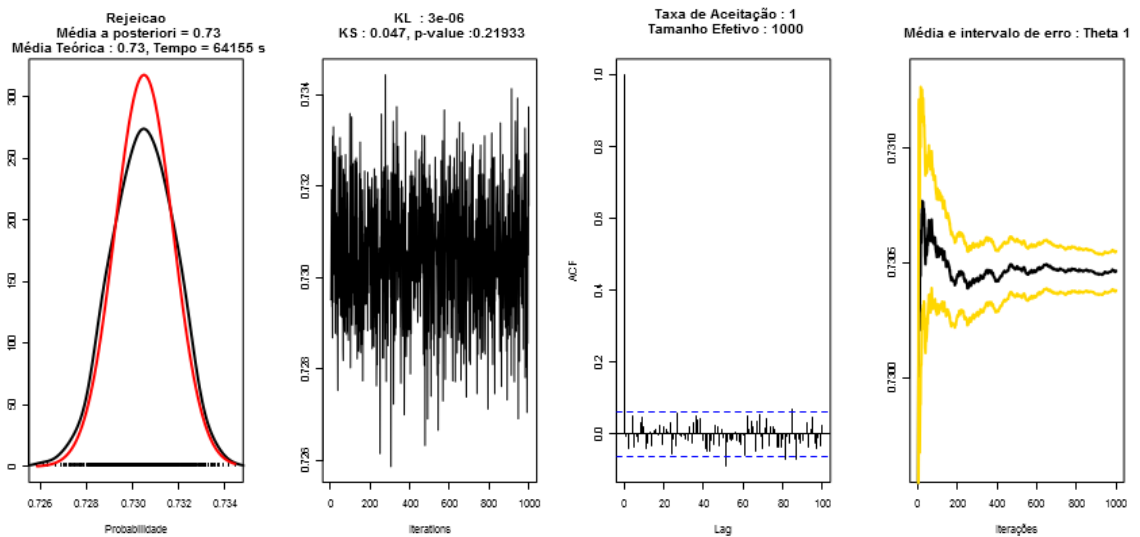
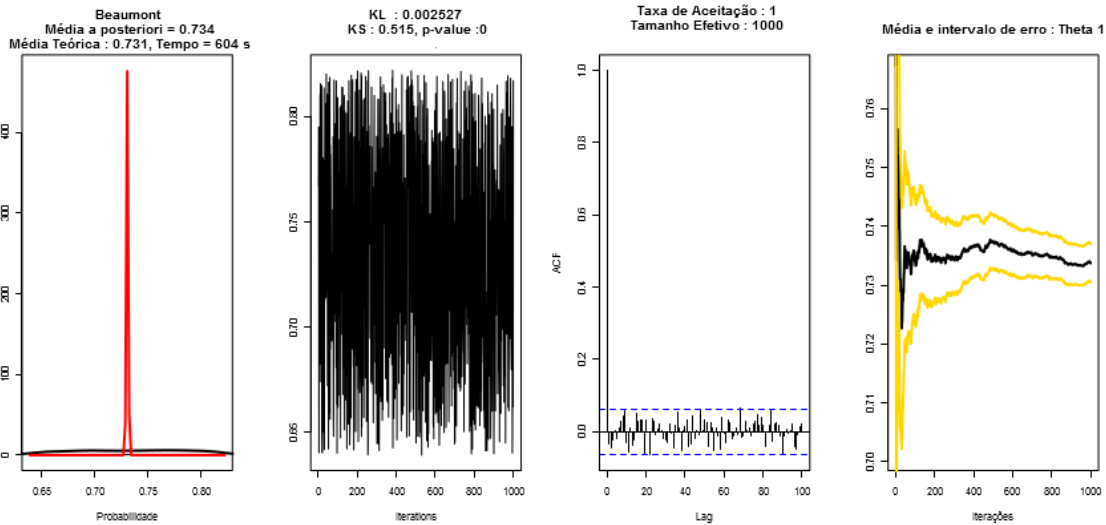
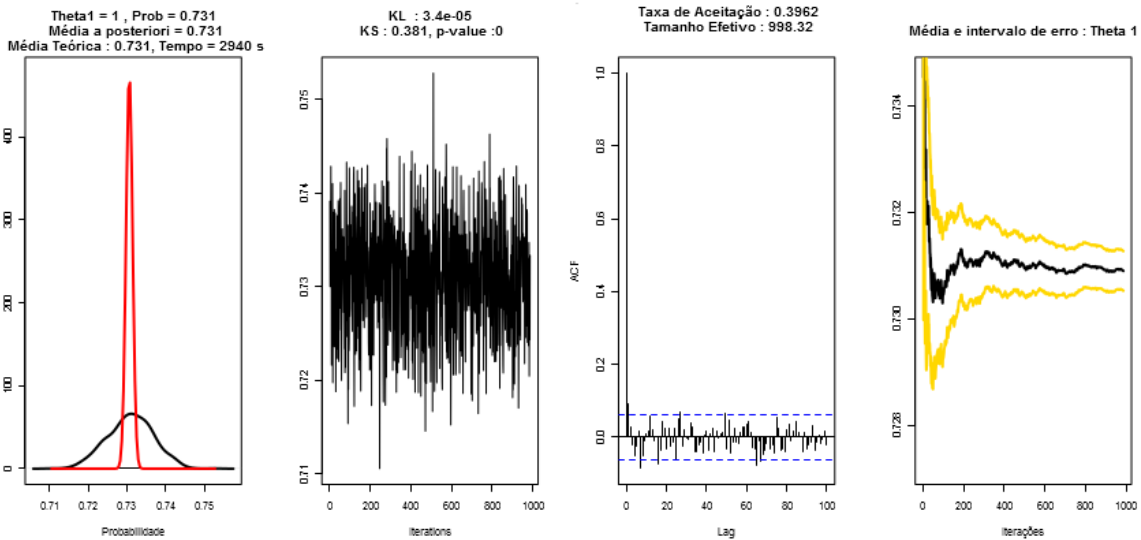


Figura A.12: Grafos aleatórios Bernoulli **500 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

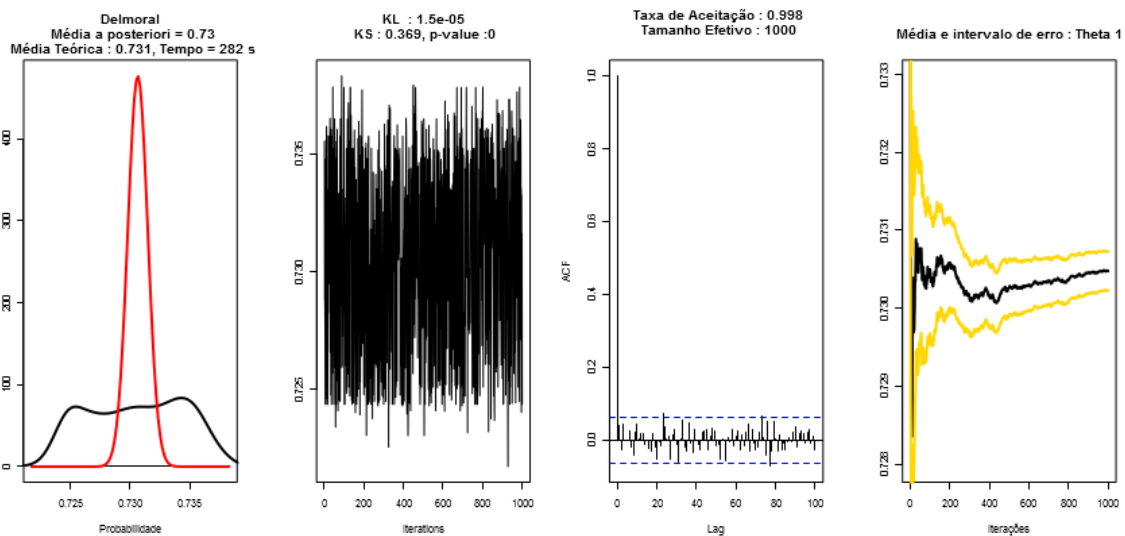
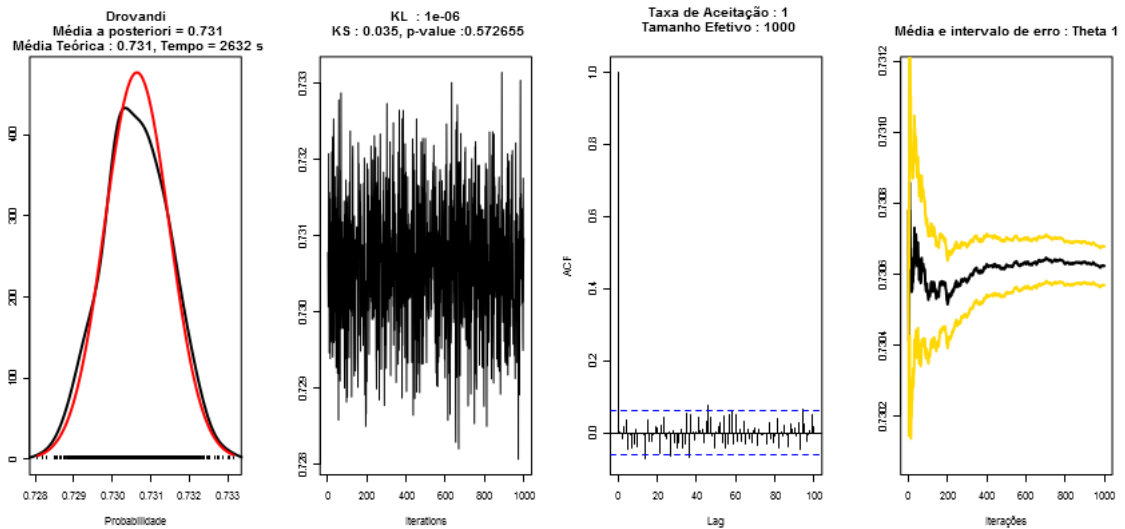
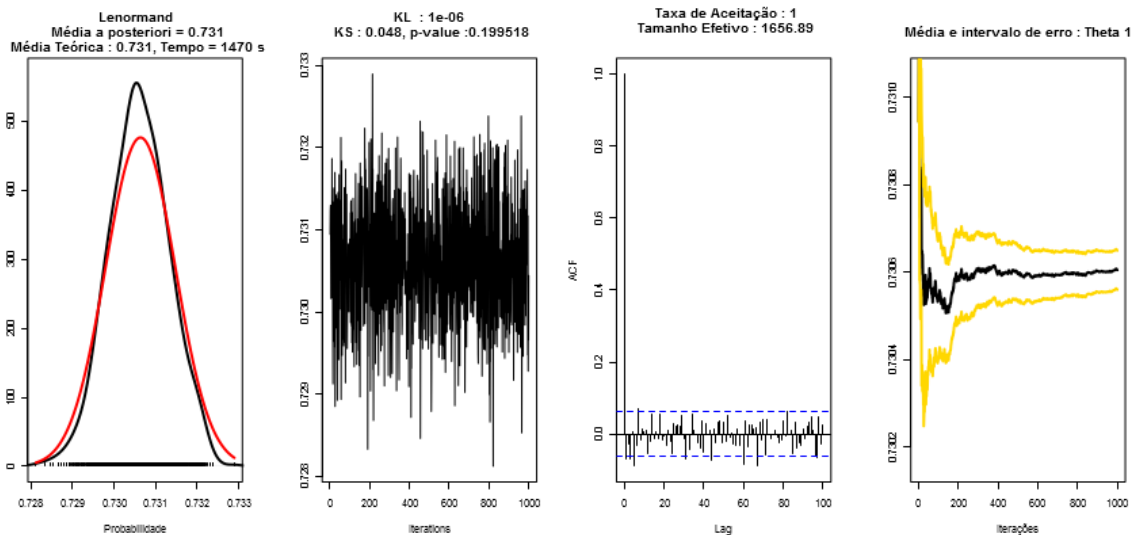


Figura A.13: Grafos aleatórios Bernoulli **750 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand



## ABC MCMC-Marjoram

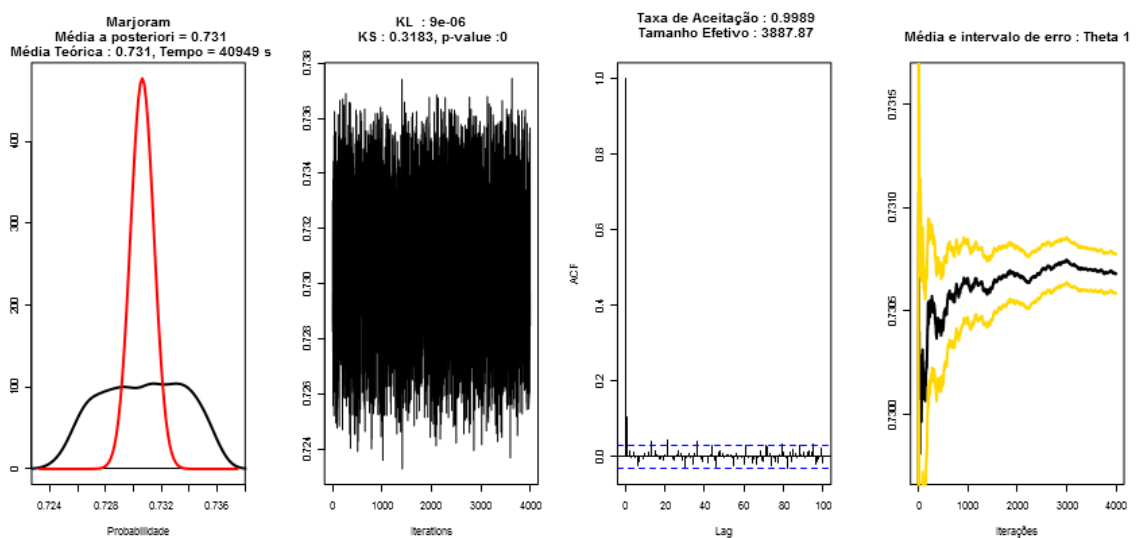


Figura A.14: Grafos aleatórios Bernoulli **750 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenormand e Marjoram.

ABC Rejeicao

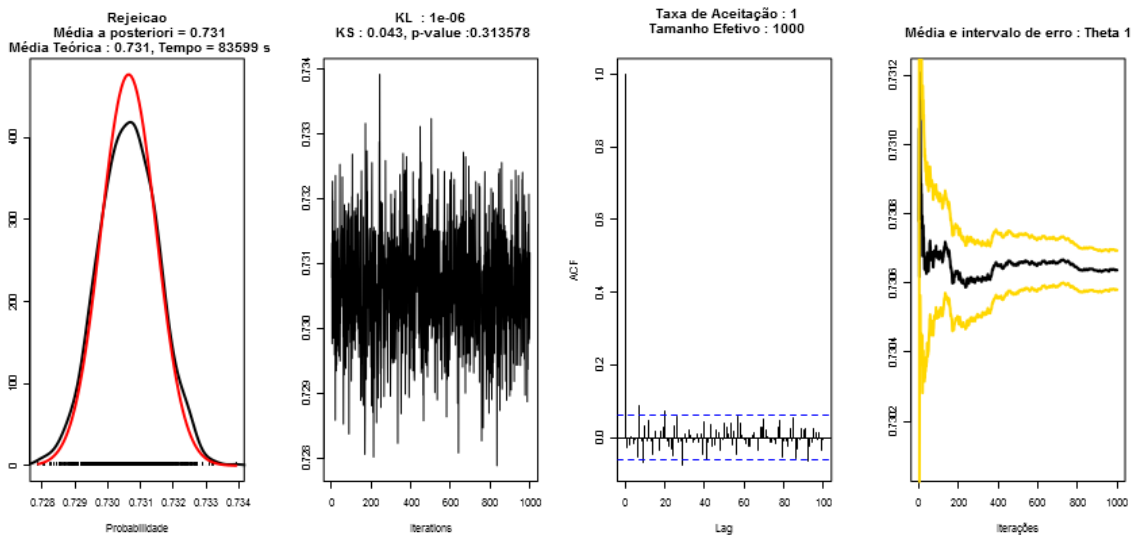
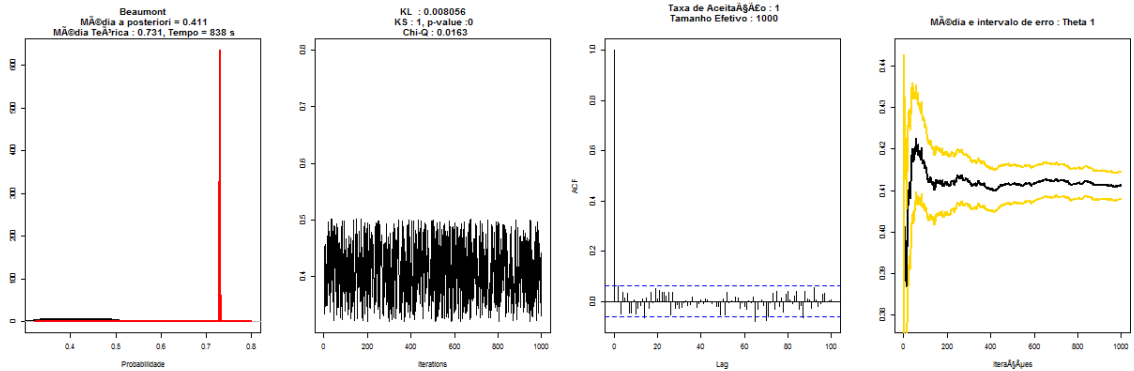
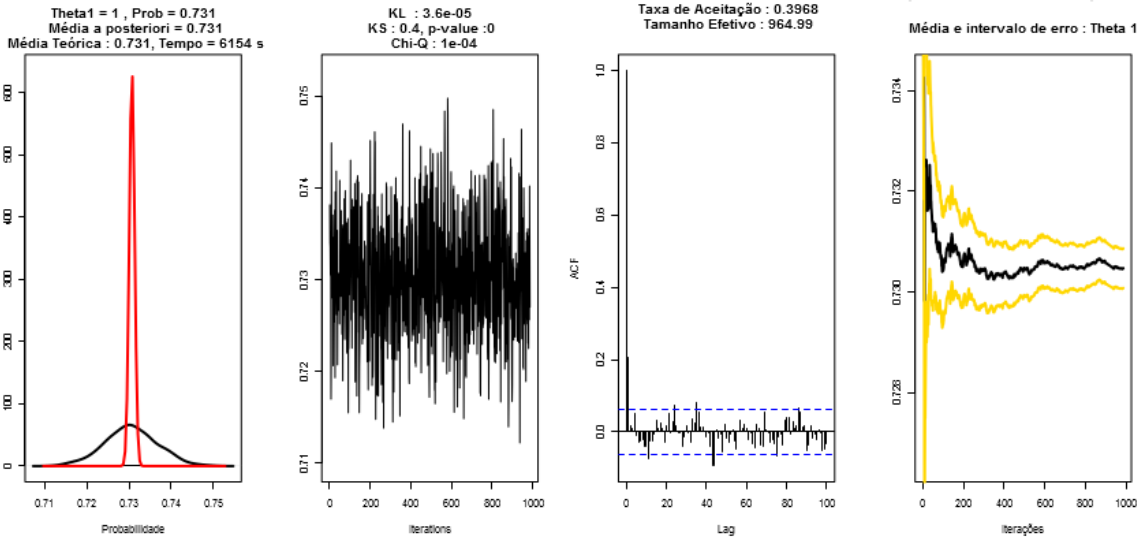


Figura A.15: Grafos aleatórios Bernoulli **750 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

## ABC Sequencial-Beaumont



## Bergm



## ABC Sequencial-Delmoral

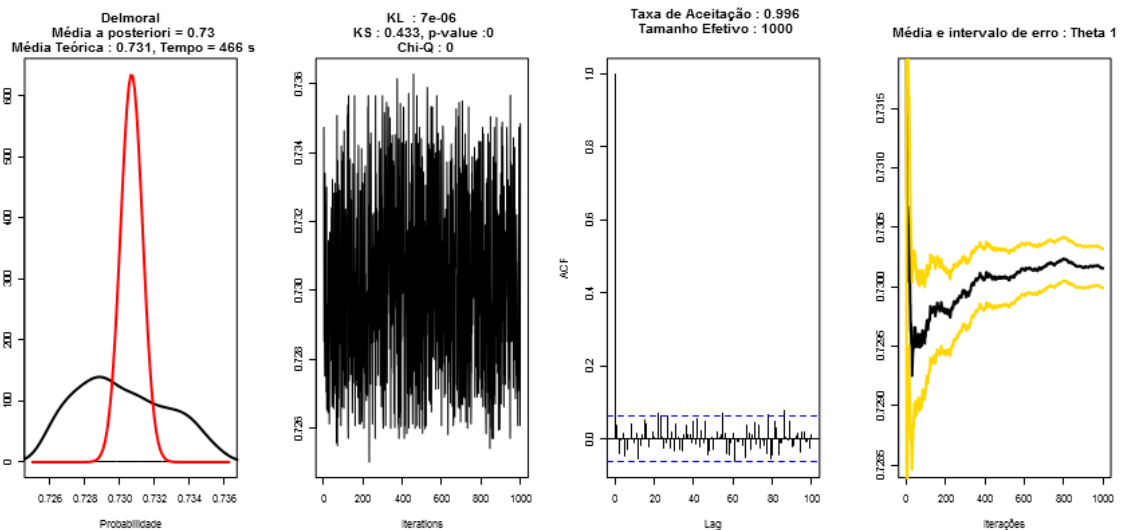
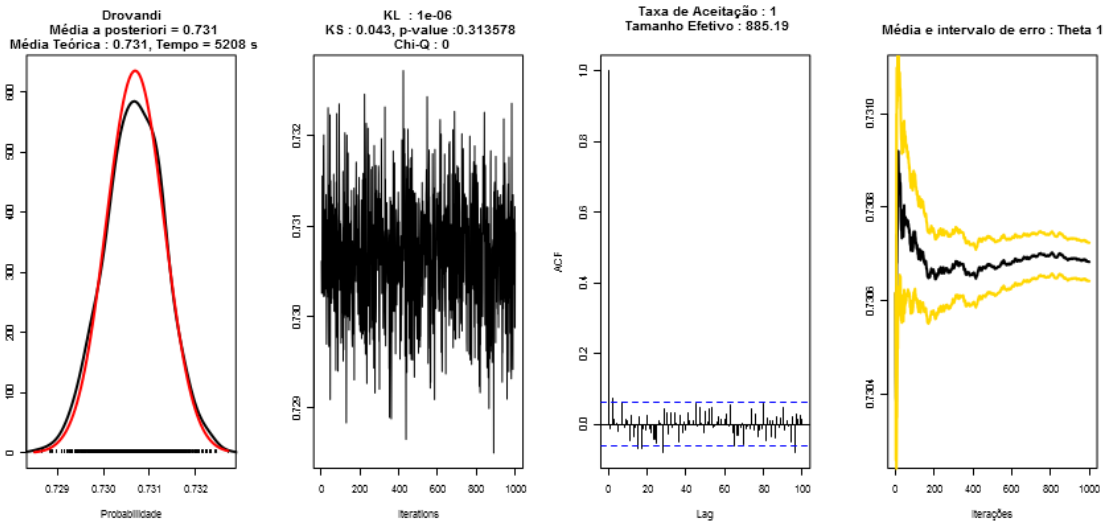
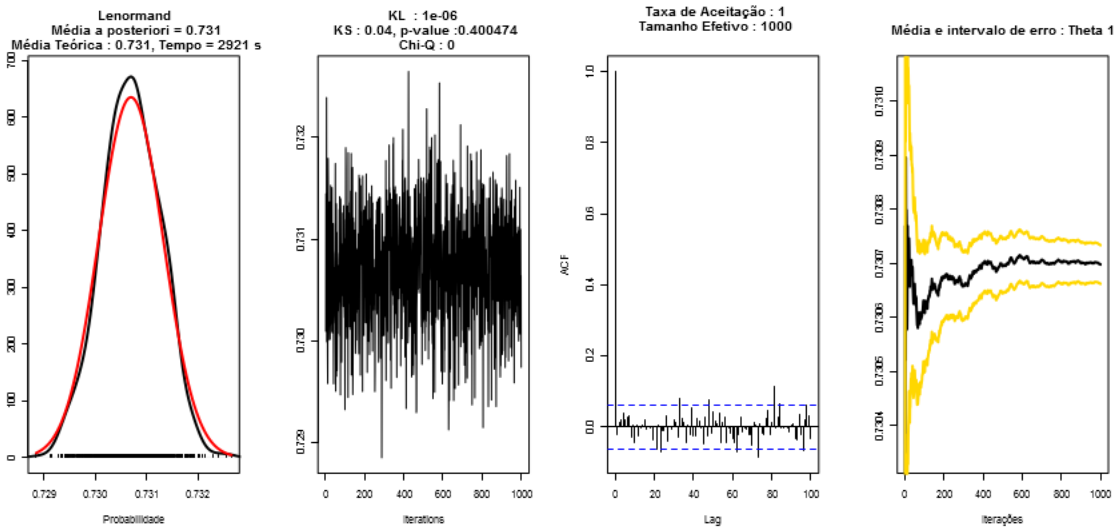


Figura A.16: Gráficos aleatórios Bernoulli **1.000 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

ABC Sequencial-Drovandi



ABC Sequencial-Lenormand



ABC MCMC-Marjoram

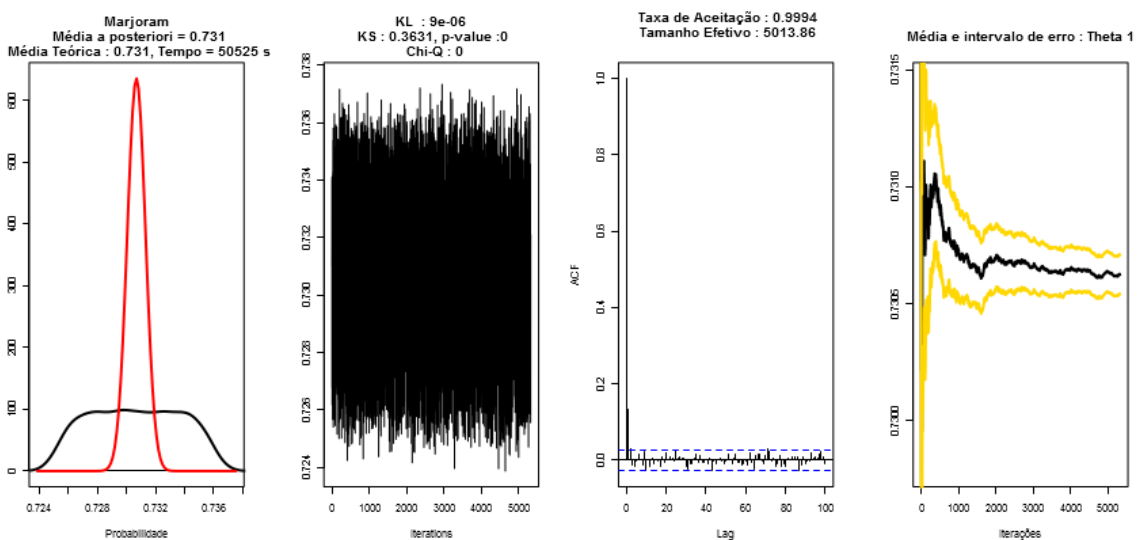


Figura A.17: Grafos aleatórios Bernoulli **1.000 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenorman e Marjoram.

## ABC Rejeicao

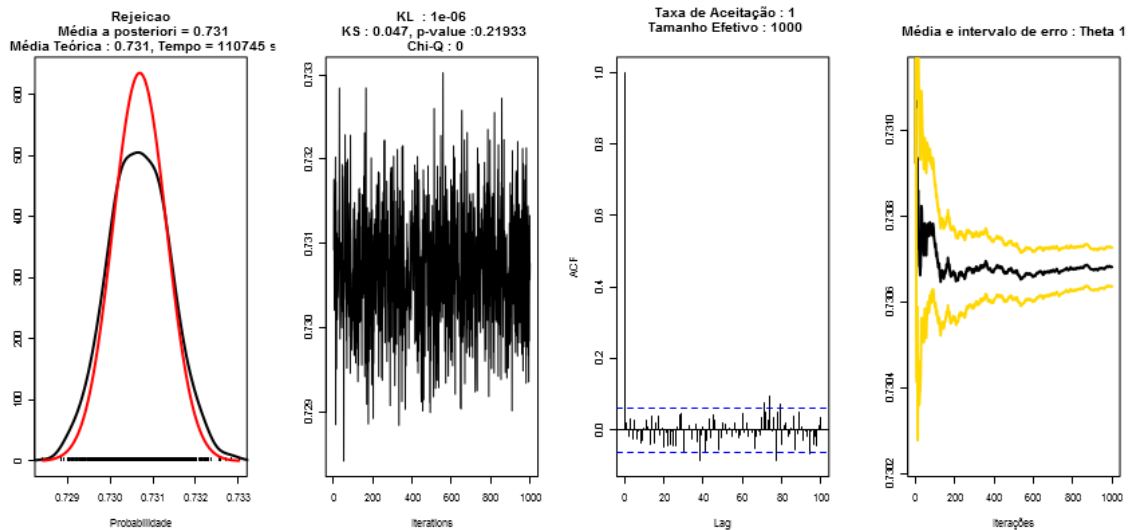


Figura A.18: Gráficos aleatórios Bernoulli **1.000 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para o método Rejeição.

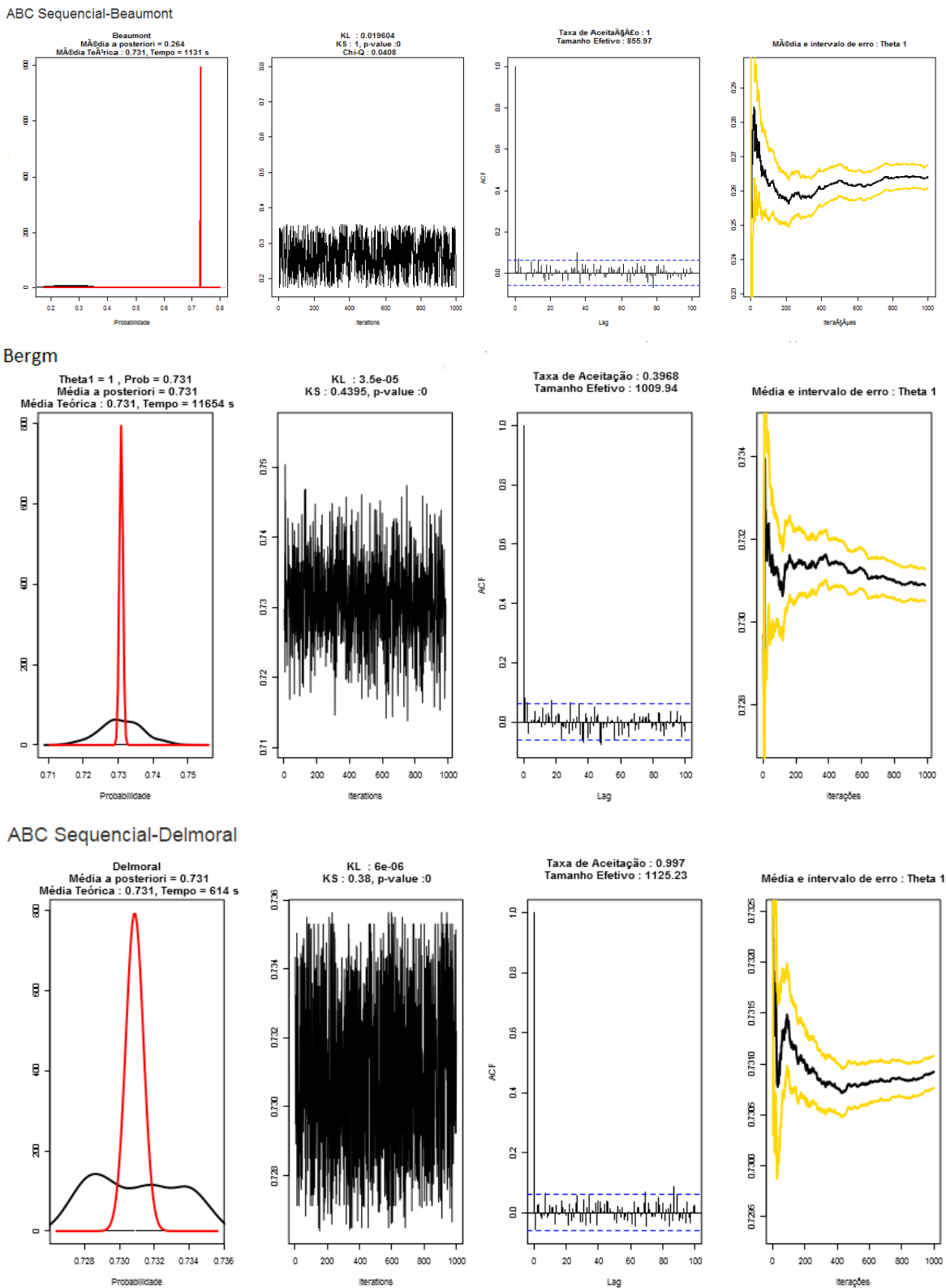
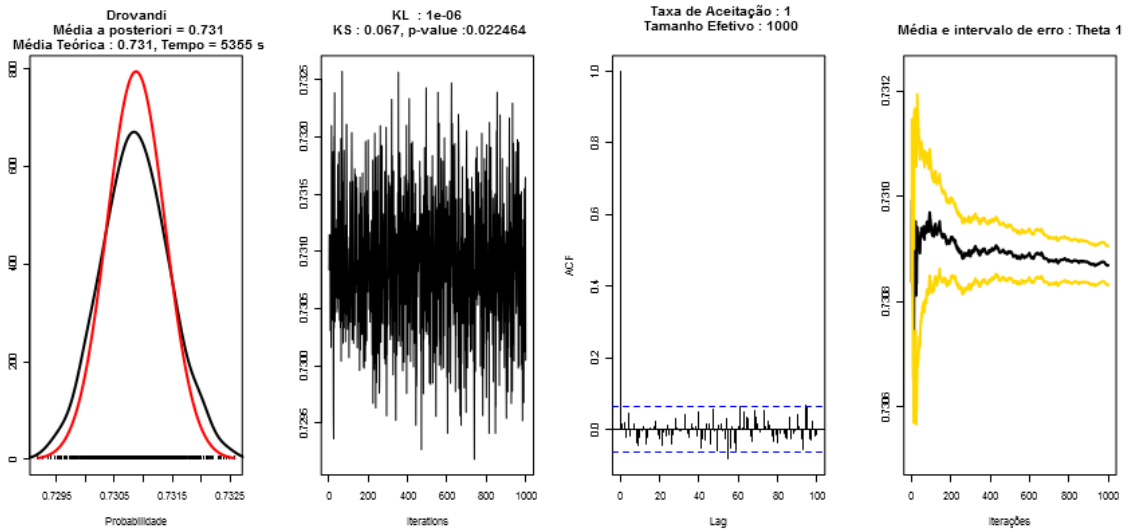


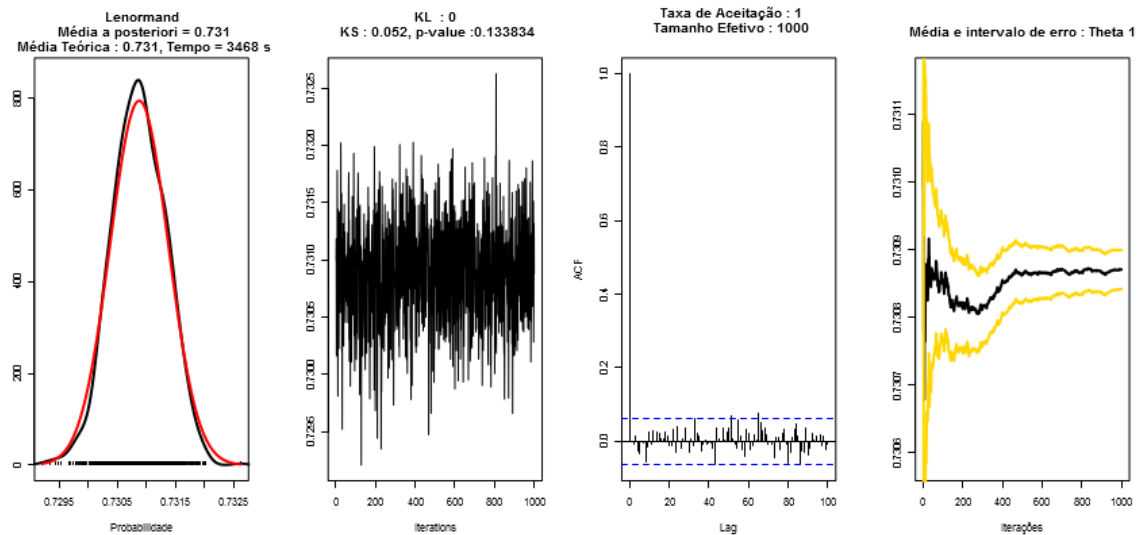
Figura A.19: Grafos aleatórios Bernoulli **1.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com lag de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.



## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand



## ABC Rejeicao

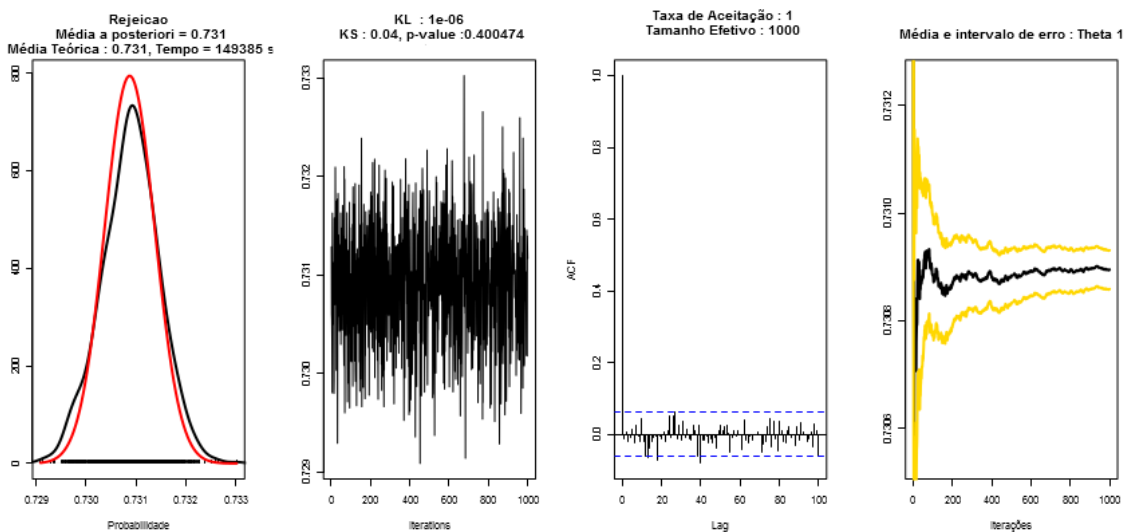
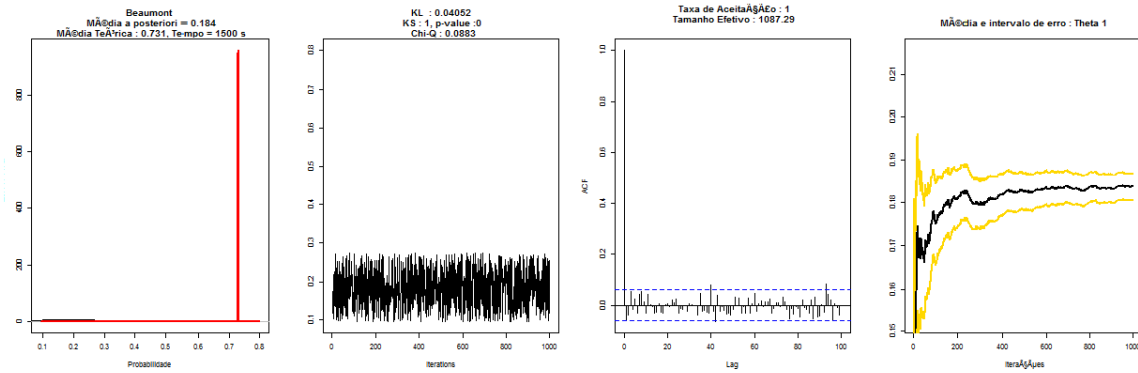
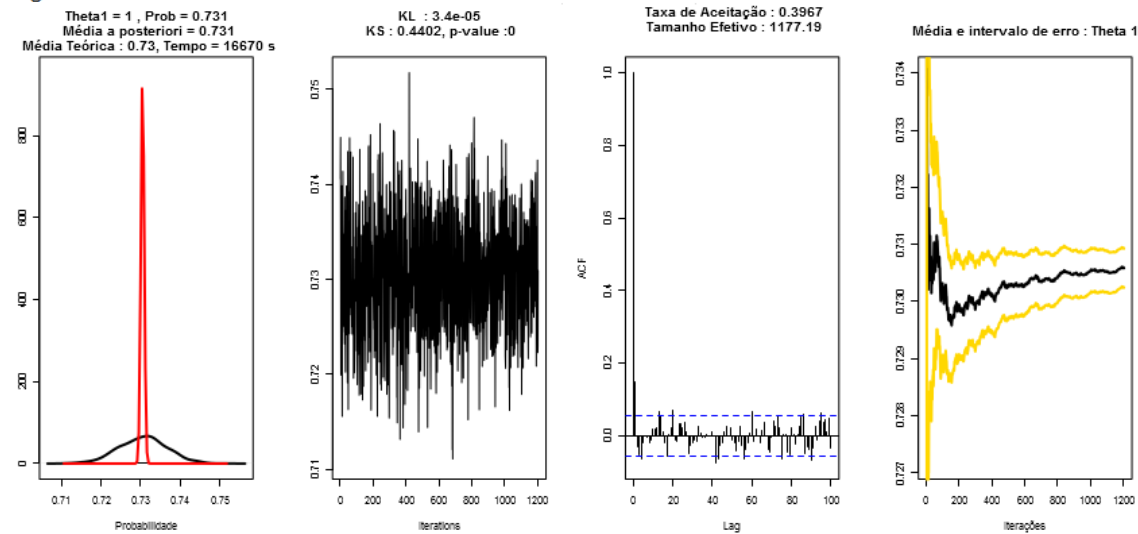


Figura A.20: Grafos aleatórios Bernoulli **1.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi, Lenorman e Rejeição. Não é apresentado o resultado para o método Marjoram, pois o tempo de processamento ultrapassou 24 horas.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

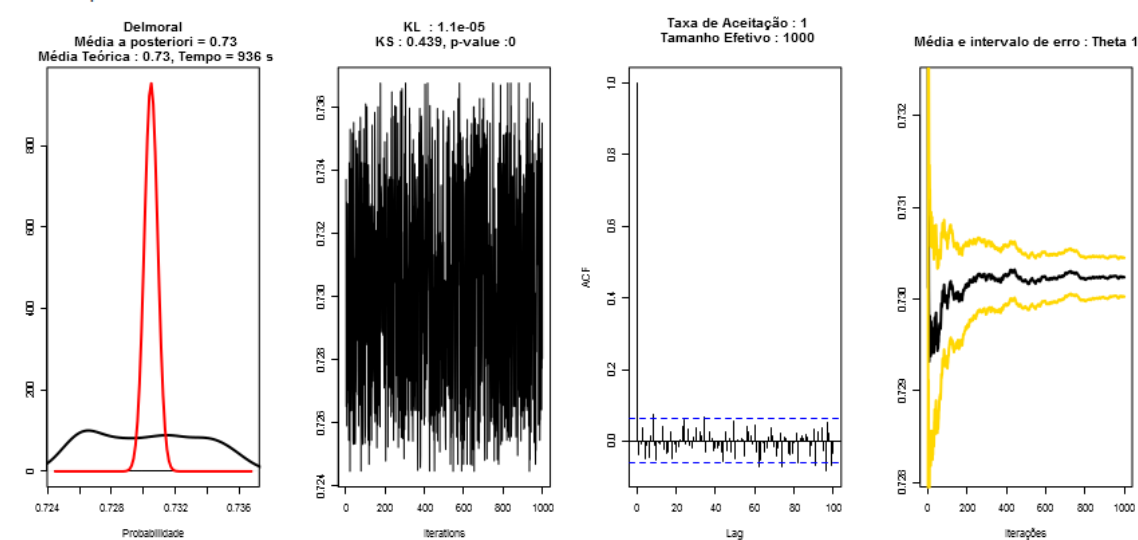
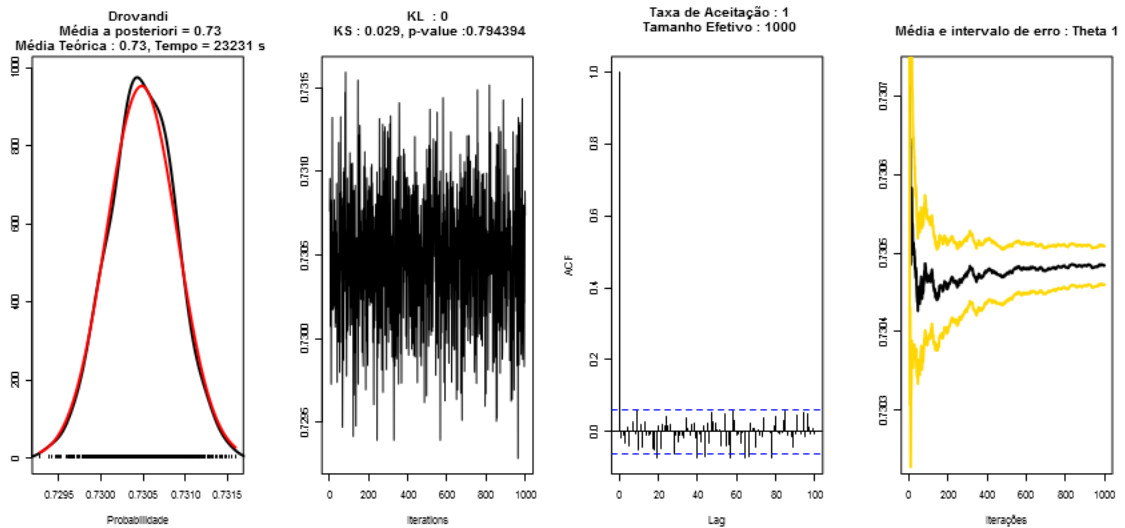


Figura A.21: Grafos aleatórios Bernoulli **1.500 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand

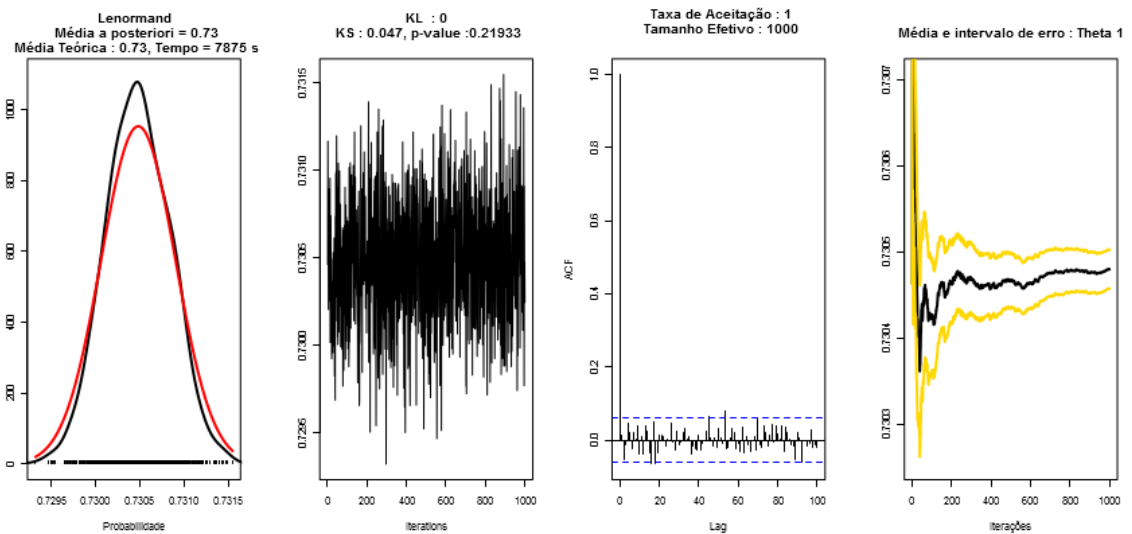
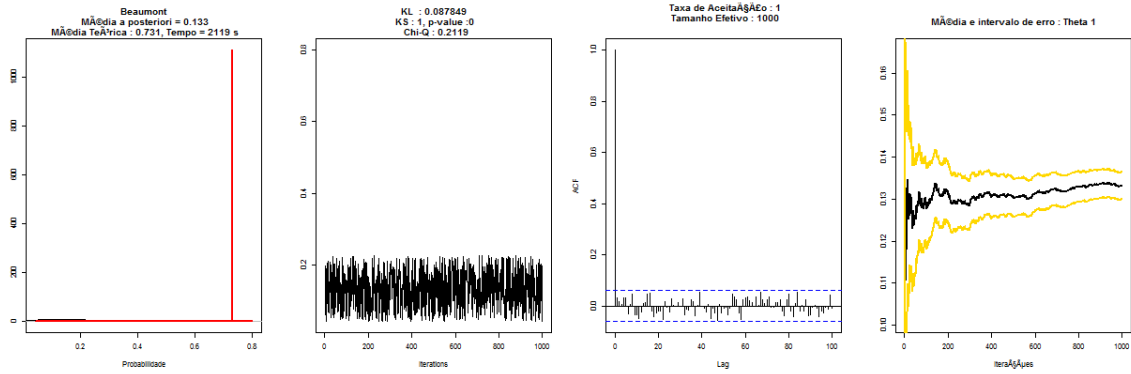
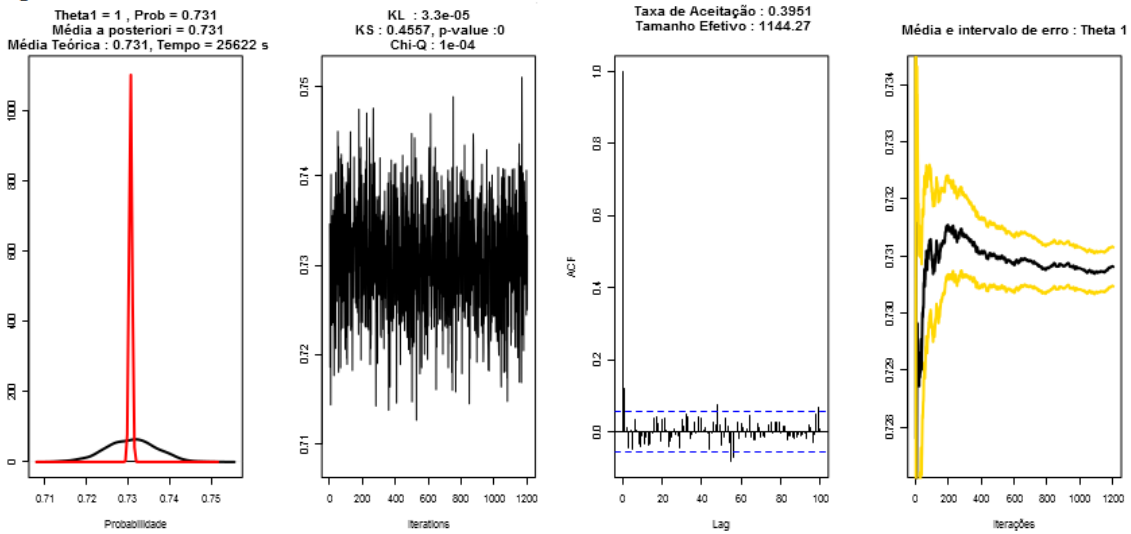


Figura A.22: Grafos aleatórios Bernoulli **1.500 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi e Lenormand. Não é apresentado o resultado para os métodos Marjoram e Rejeição, pois o tempo de processamento ultrapassou 24 horas.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

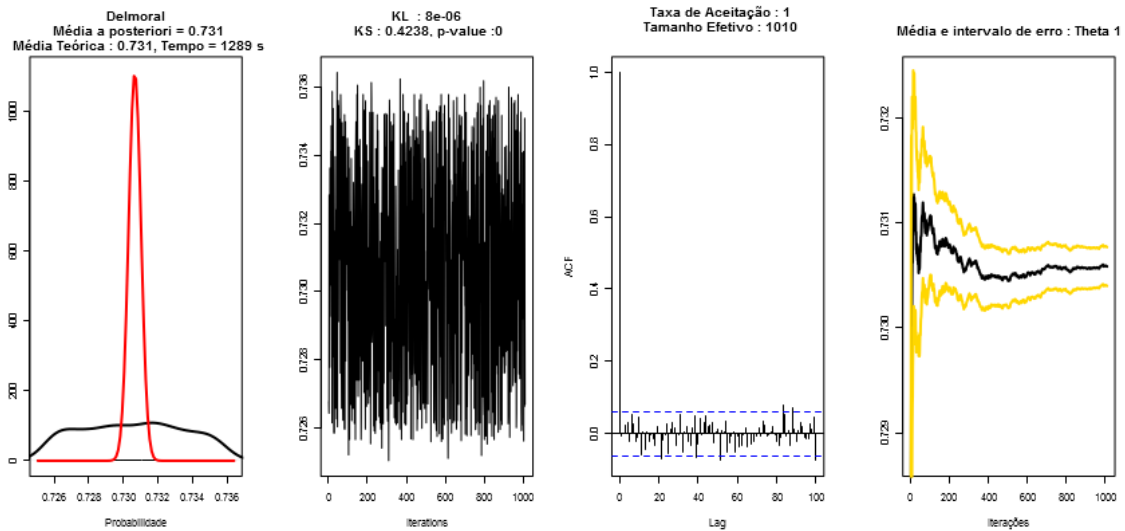
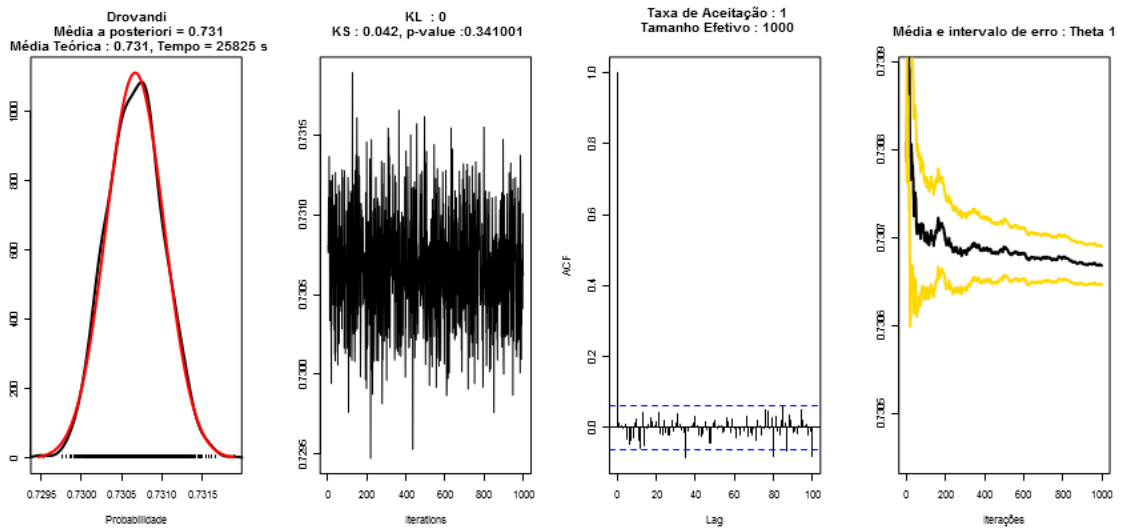


Figura A.23: Grafos aleatórios Bernoulli **1.750 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com lag de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand

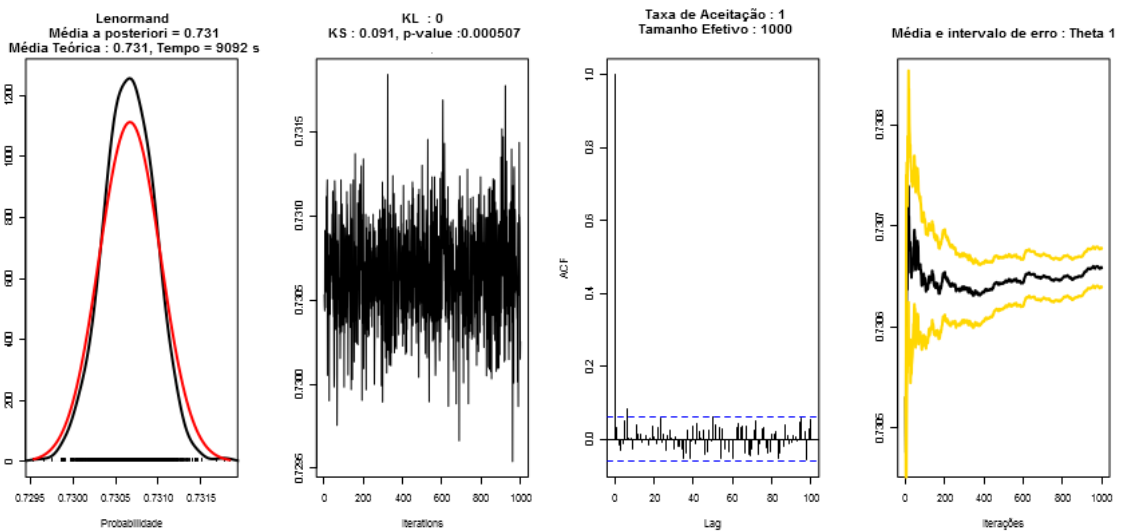
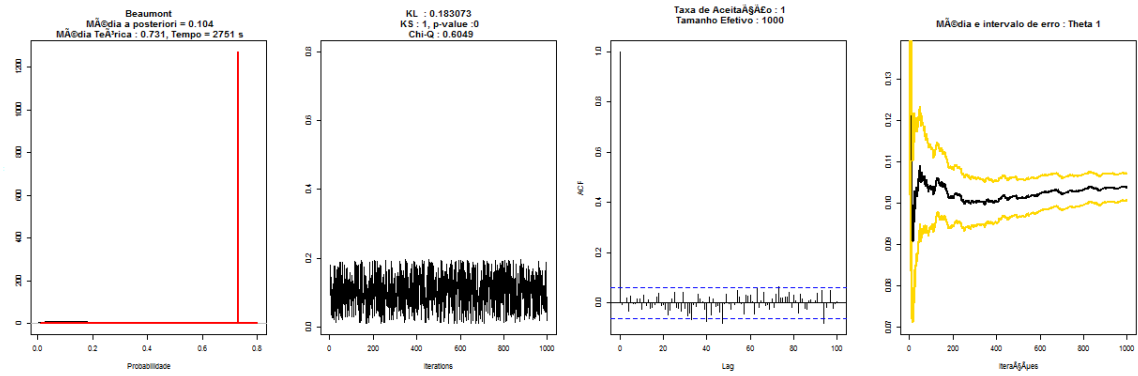
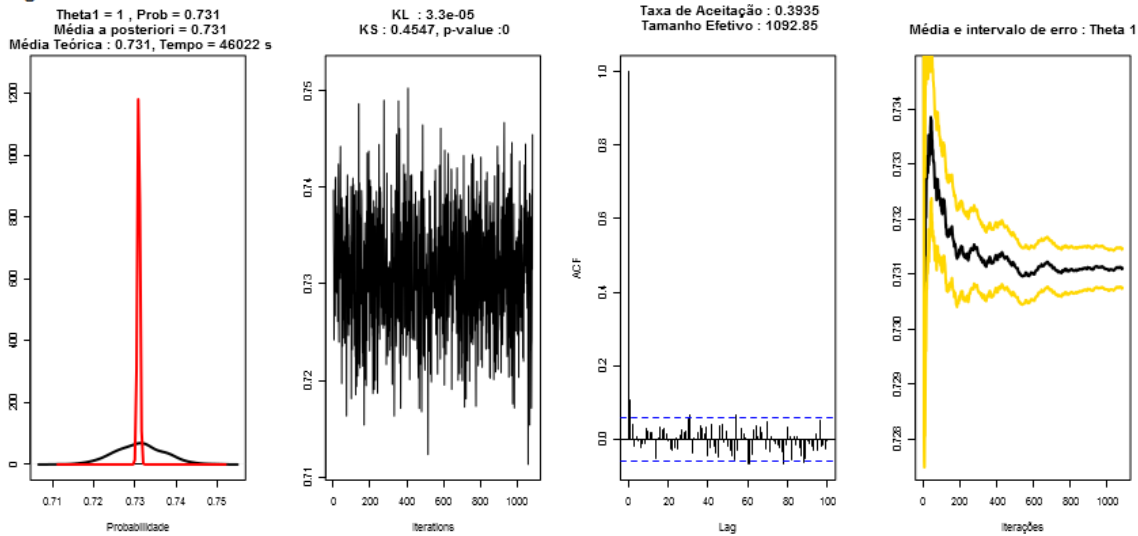


Figura A.24: Gráficos aleatórios Bernoulli **1.750 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi e Lenormand. Não é apresentado o resultado para os métodos Marjoram e Rejeição, pois o tempo de processamento ultrapassou 24 horas.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

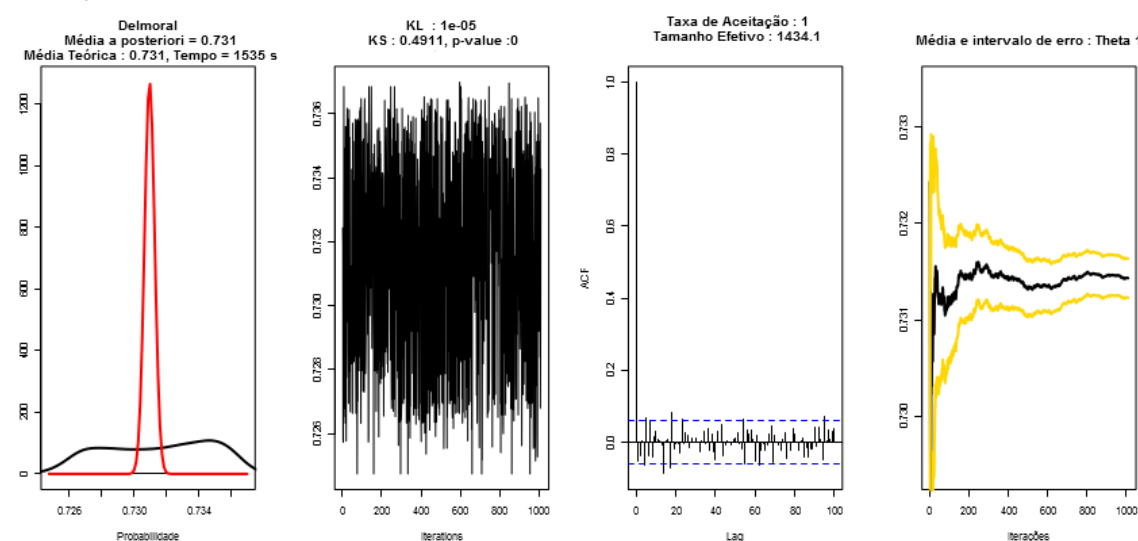
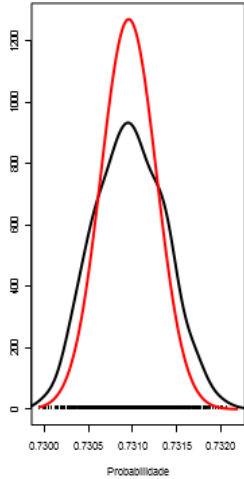


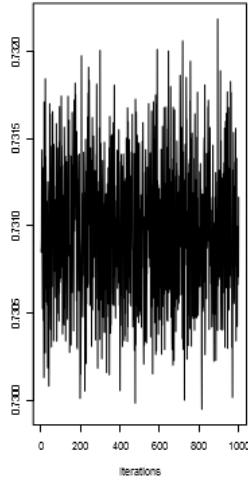
Figura A.25: Grafos aleatórios Bernoulli **2.000 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.

## ABC Sequencial-Drovandi

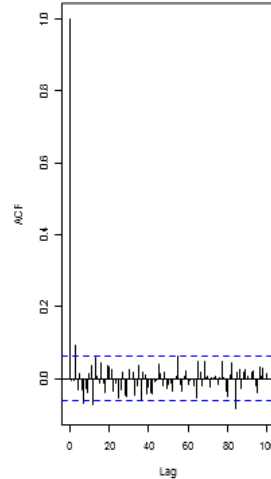
Drovandi  
 Média a posteriori = 0.731  
 Média Teórica : 0.731, Tempo = 9388 s



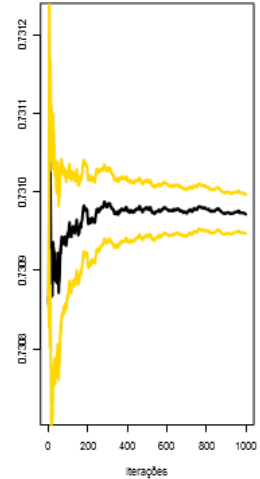
KL : 0  
 KS : 0.085, p-value : 0.001456



Taxa de Aceitação : 1  
 Tamanho Efetivo : 853.01

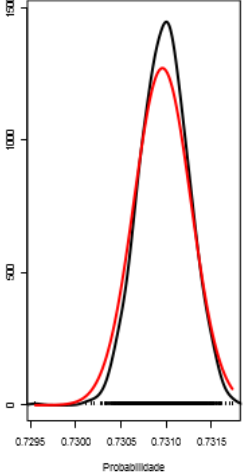


Média e intervalo de erro : Theta 1

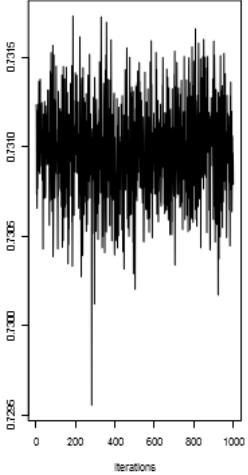


## ABC Sequencial-Lenormand

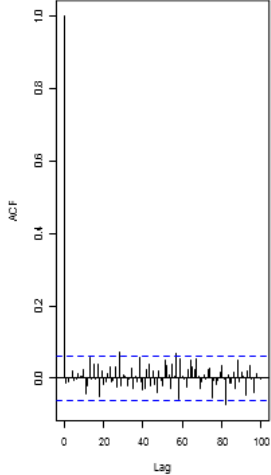
Lenormand  
 Média a posteriori = 0.731  
 Média Teórica : 0.731, Tempo = 10605 s



KL : 0  
 KS : 0.054, p-value : 0.108283



Taxa de Aceitação : 1  
 Tamanho Efetivo : 1000



Média e intervalo de erro : Theta 1

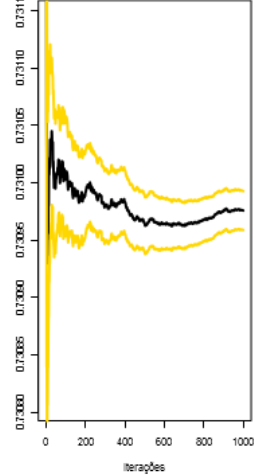
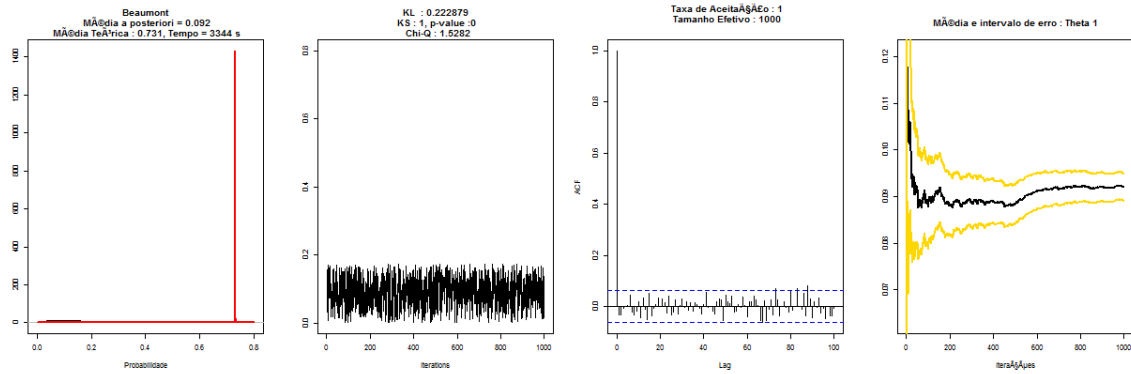
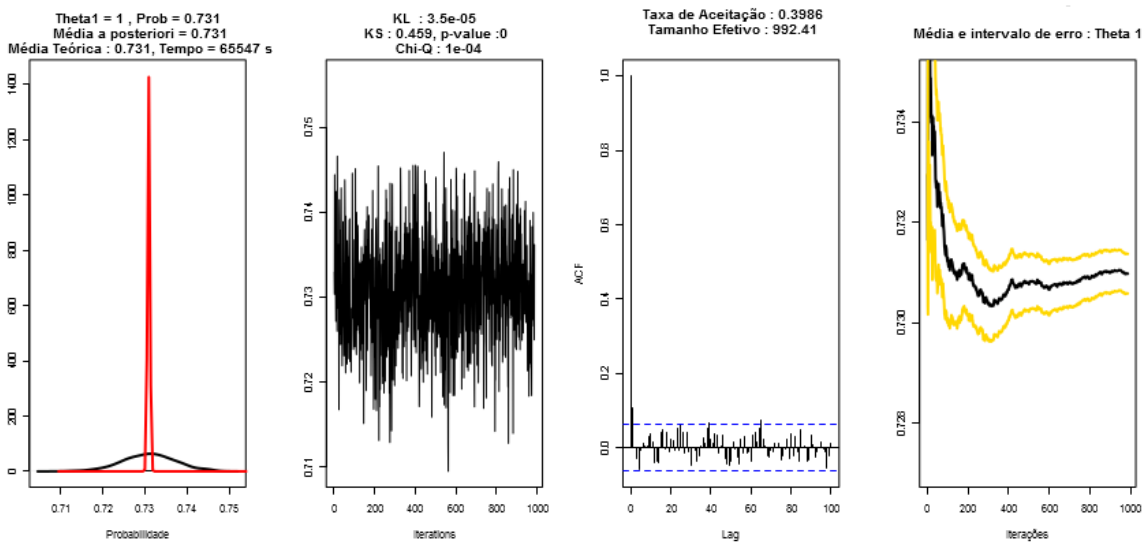


Figura A.26: Gráficos aleatórios Bernoulli **2.000 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi e Lenormand. Não é apresentado o resultado para os métodos Marjoram e Rejeição, pois o tempo de processamento ultrapassou 24 horas.

ABC Sequencial-Beaumont



Bergm



ABC Sequencial-Delmoral

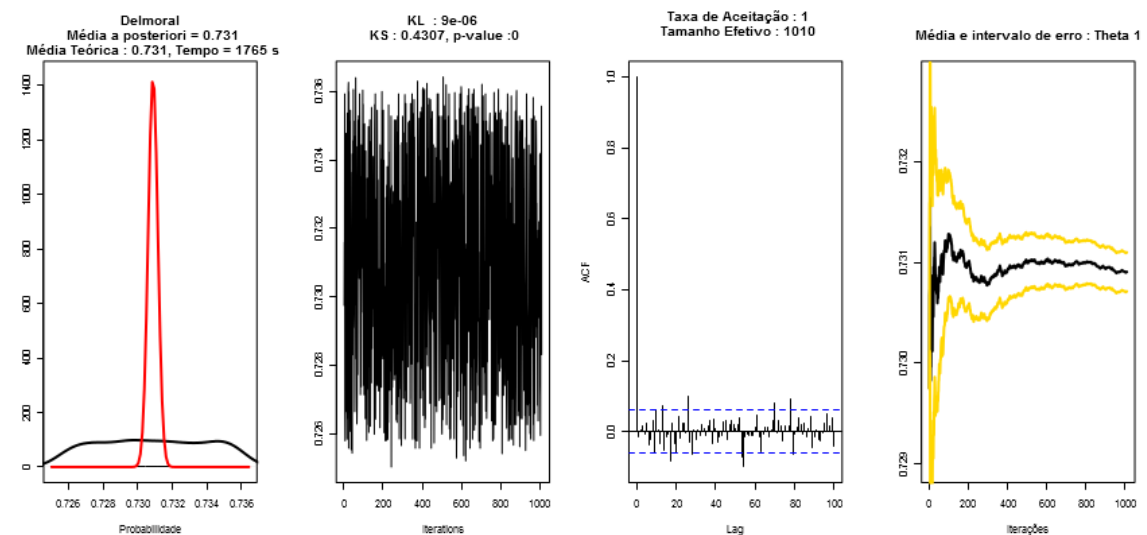
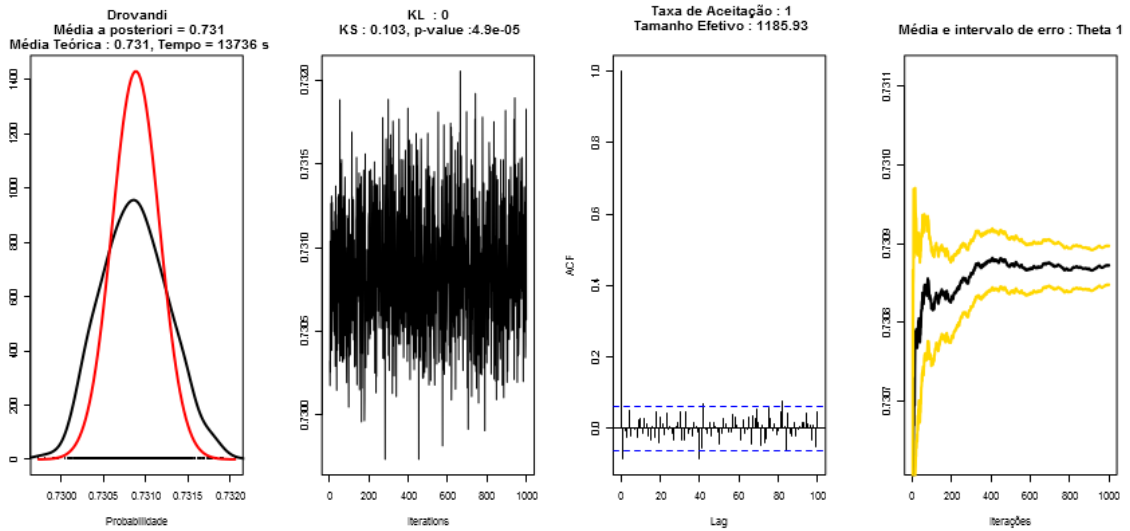


Figura A.27: Grafos aleatórios Bernoulli **2.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Beaumont, Bergm e Delmoral.



## ABC Sequencial-Drovandi



## ABC Sequencial-Lenormand

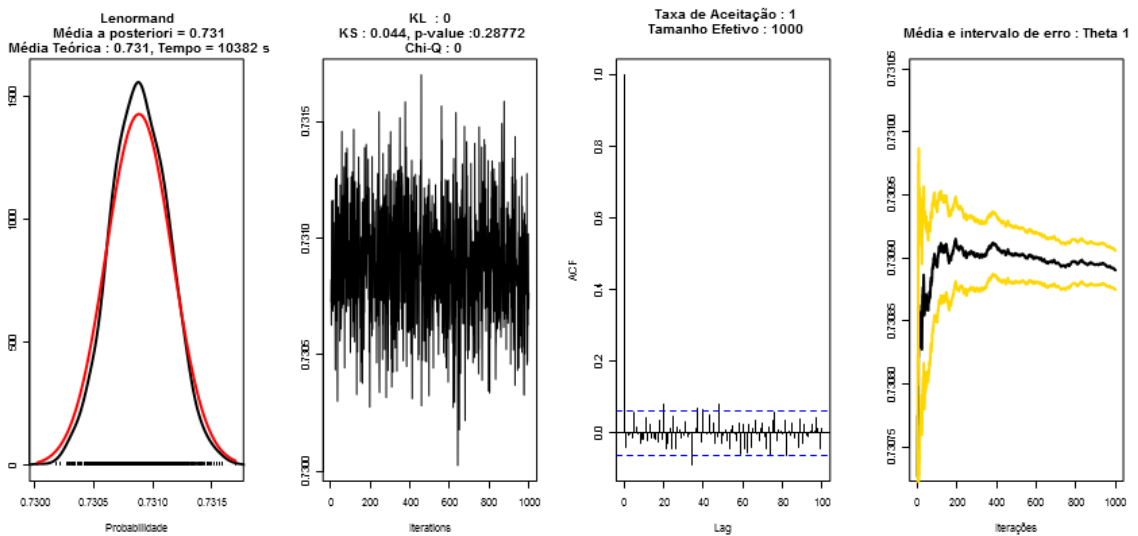


Figura A.28: Gráficos aleatórios Bernoulli **2.250 nós** : gráfico da distribuição *a posteriori* (linha preta) versus gráfico da distribuição teórica (linha vermelha), gráfico da convergência da cadeia, correlograma das amostras com *lag* de tamanho 100 e gráfico da média da amostra com intervalo de erro quadrático médio (linha amarela) para os métodos Drovandi e Lenormand. Não é apresentado o resultado para os métodos Marjoram e Rejeição, pois o tempo de processamento ultrapassou 24 horas.

## **Apêndice B**

# **Gráficos do Estudo de Simulação para Grafos Exponenciais : Arestas e Triângulos**

Apresentamos, a seguir, todos os gráficos para o estudo de simulação dos "Grafos Exponenciais : Arestas e Triângulos".

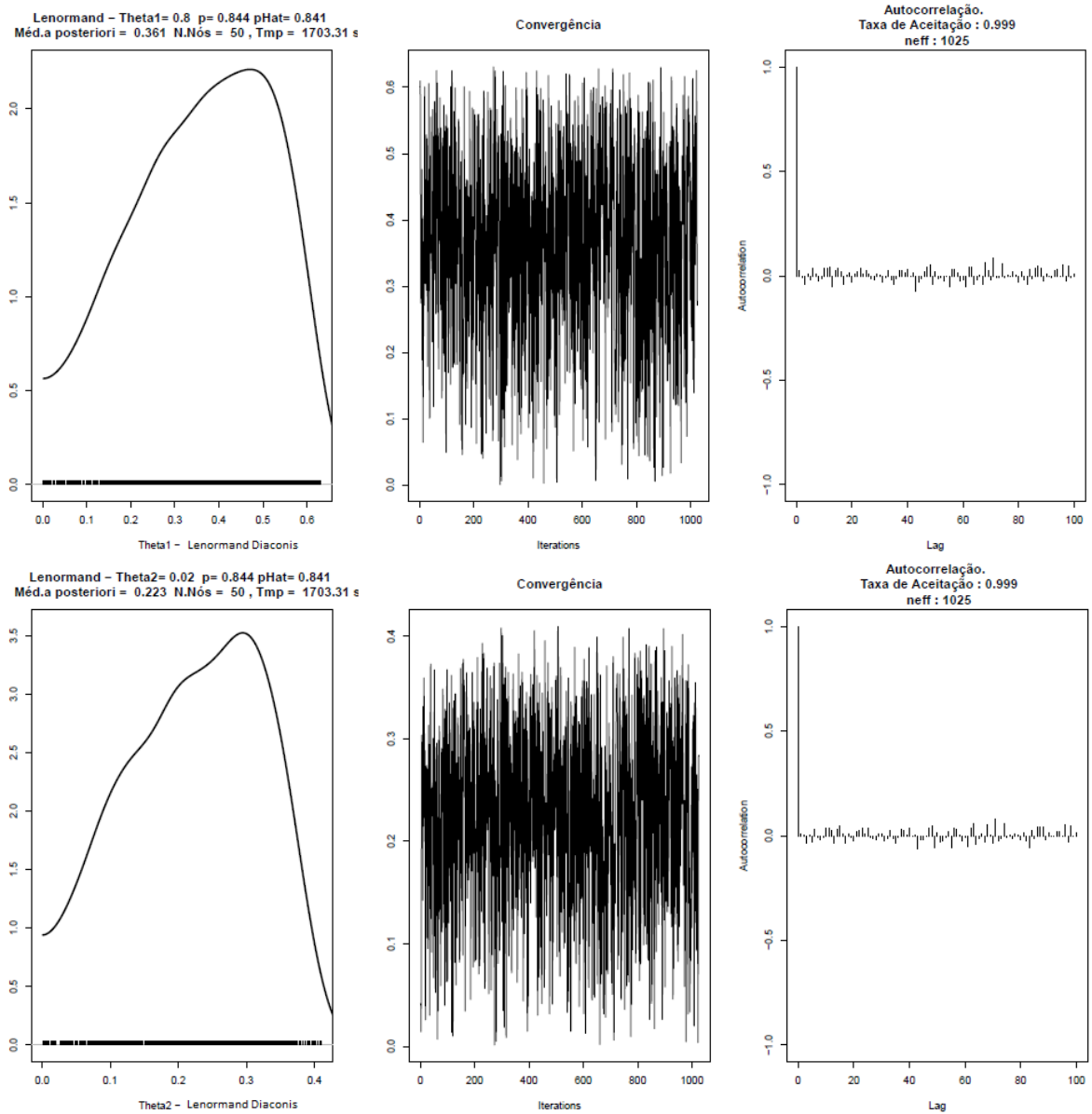


Figura B.1: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,8$  e  $\theta_2 = 0,02$ , **50** nós.

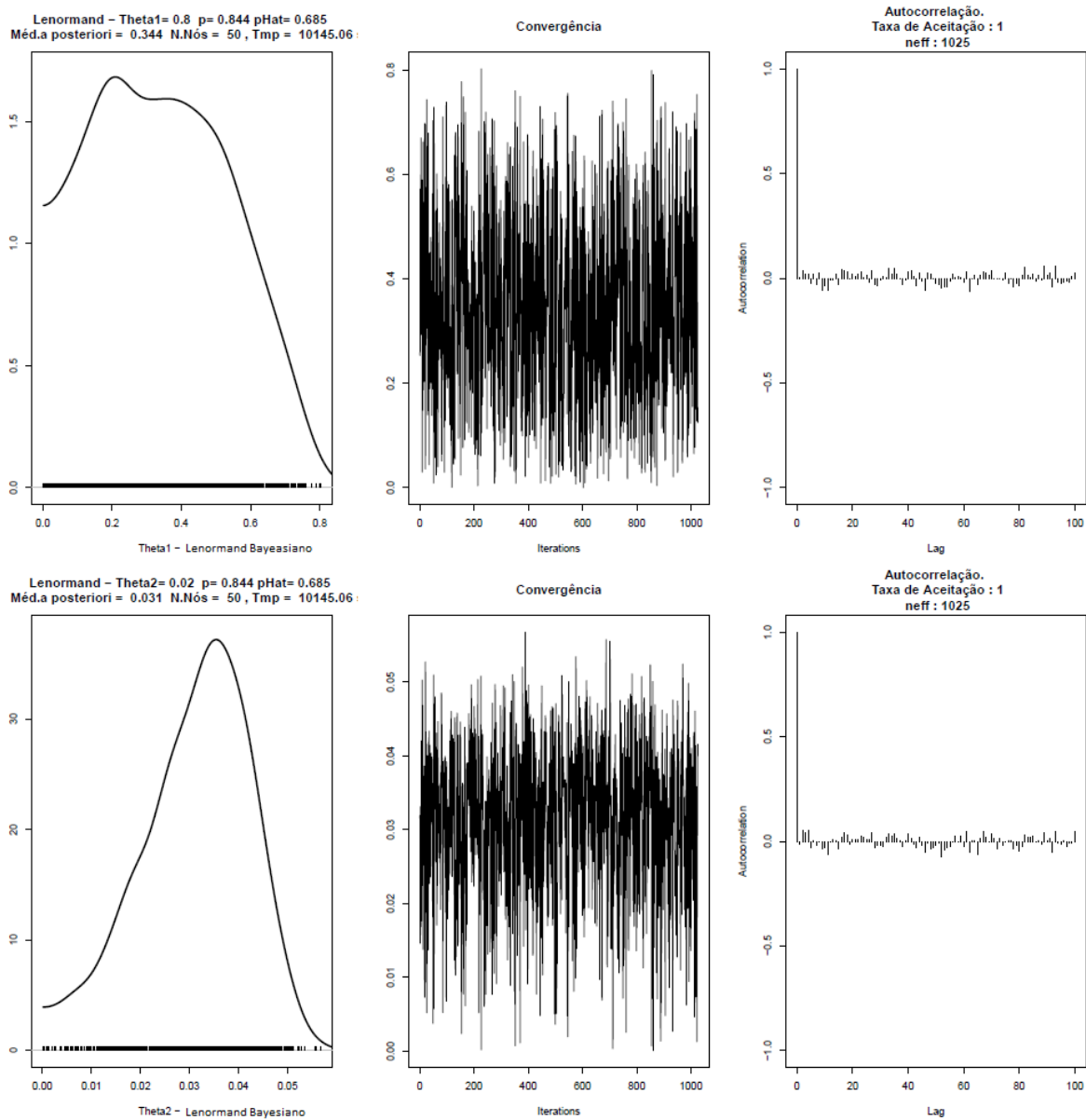


Figura B.2: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,8$  e  $\theta_2 = 0,02$ , 50 nós.

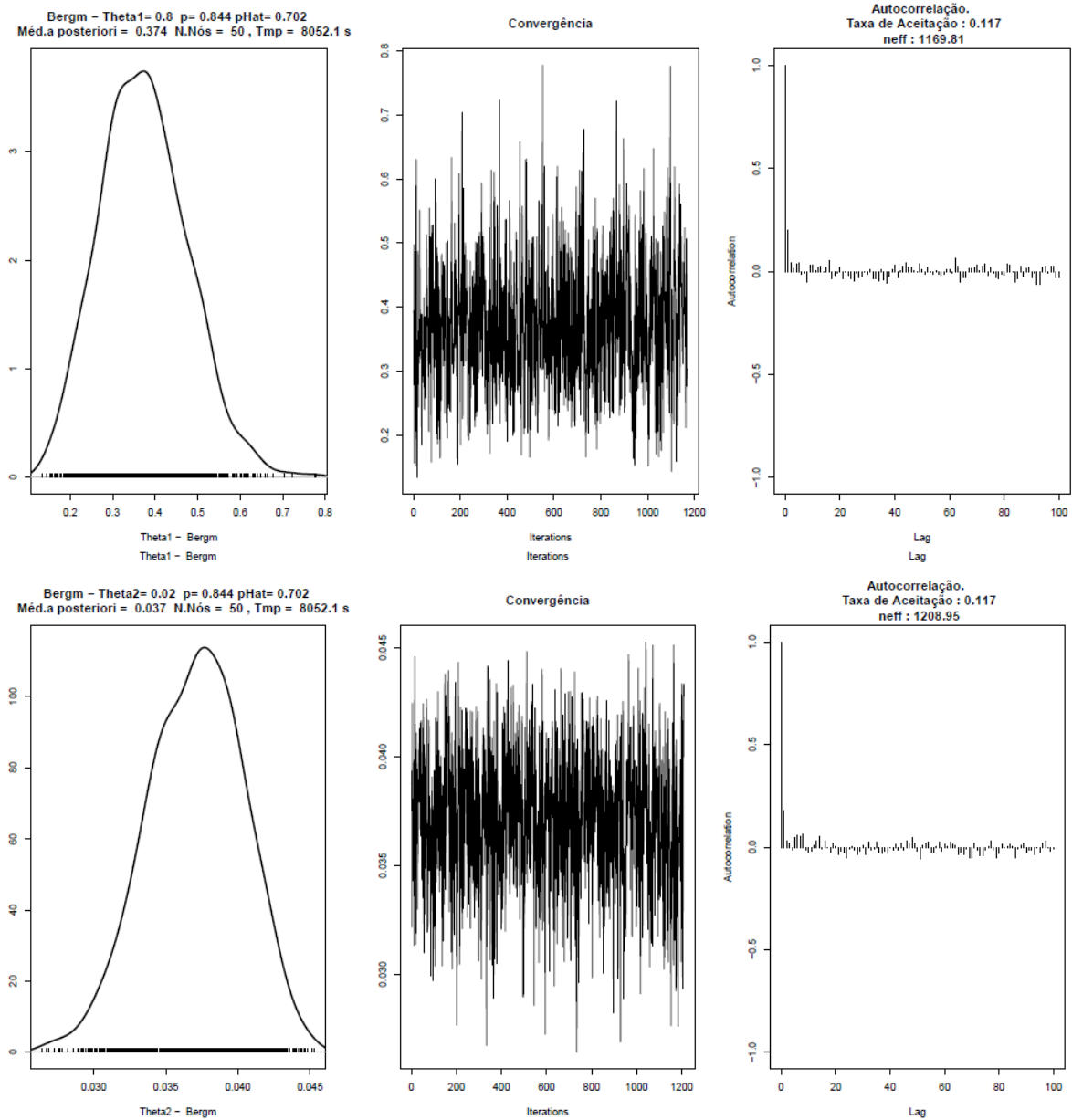


Figura B.3: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método Bergm :  $\theta_1 = 0,8$  e  $\theta_2 = 0,02$ , **50** nós.

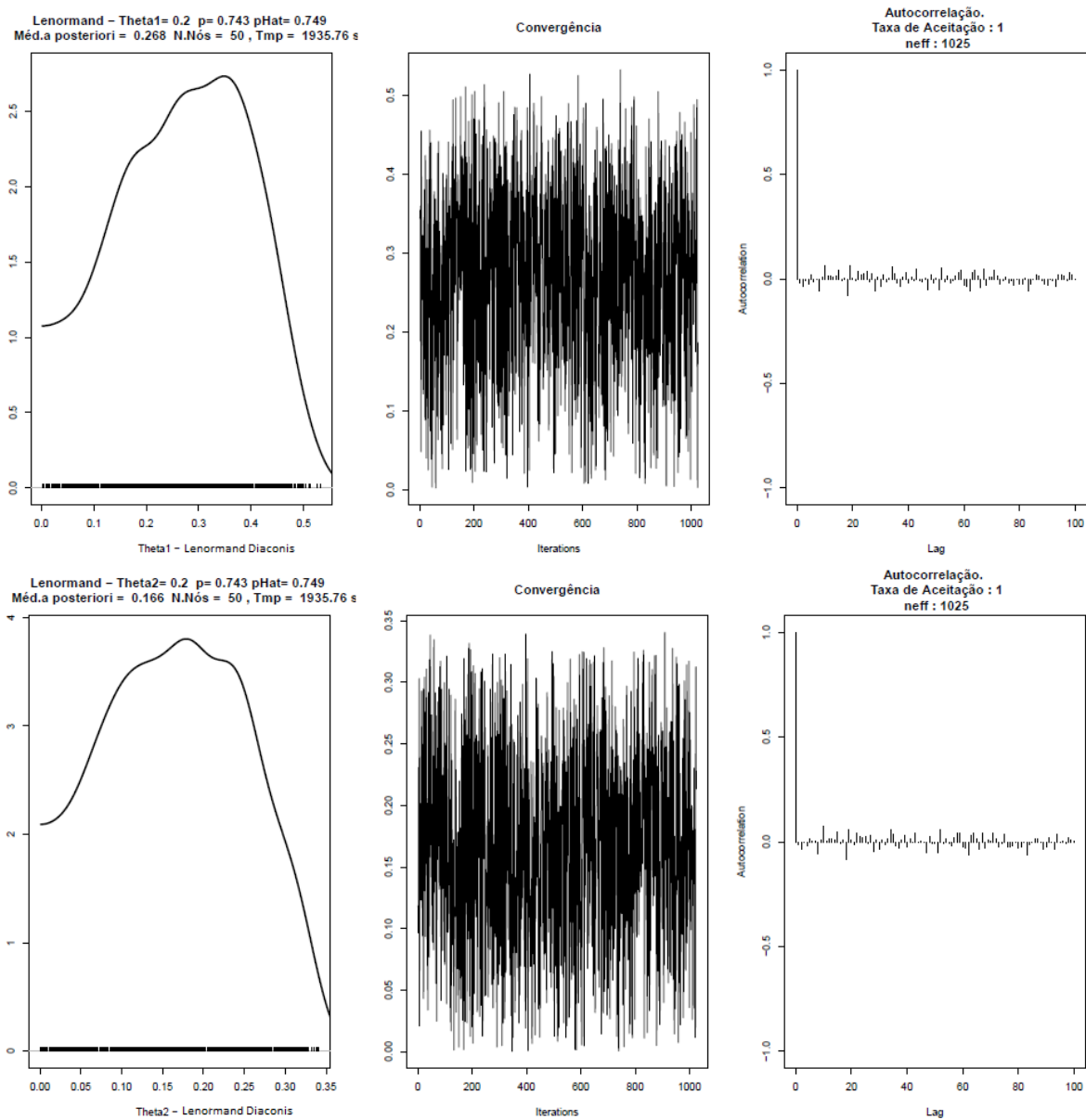


Figura B.4: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,2$  e  $\theta_2 = 0,2$ , **50** nós.

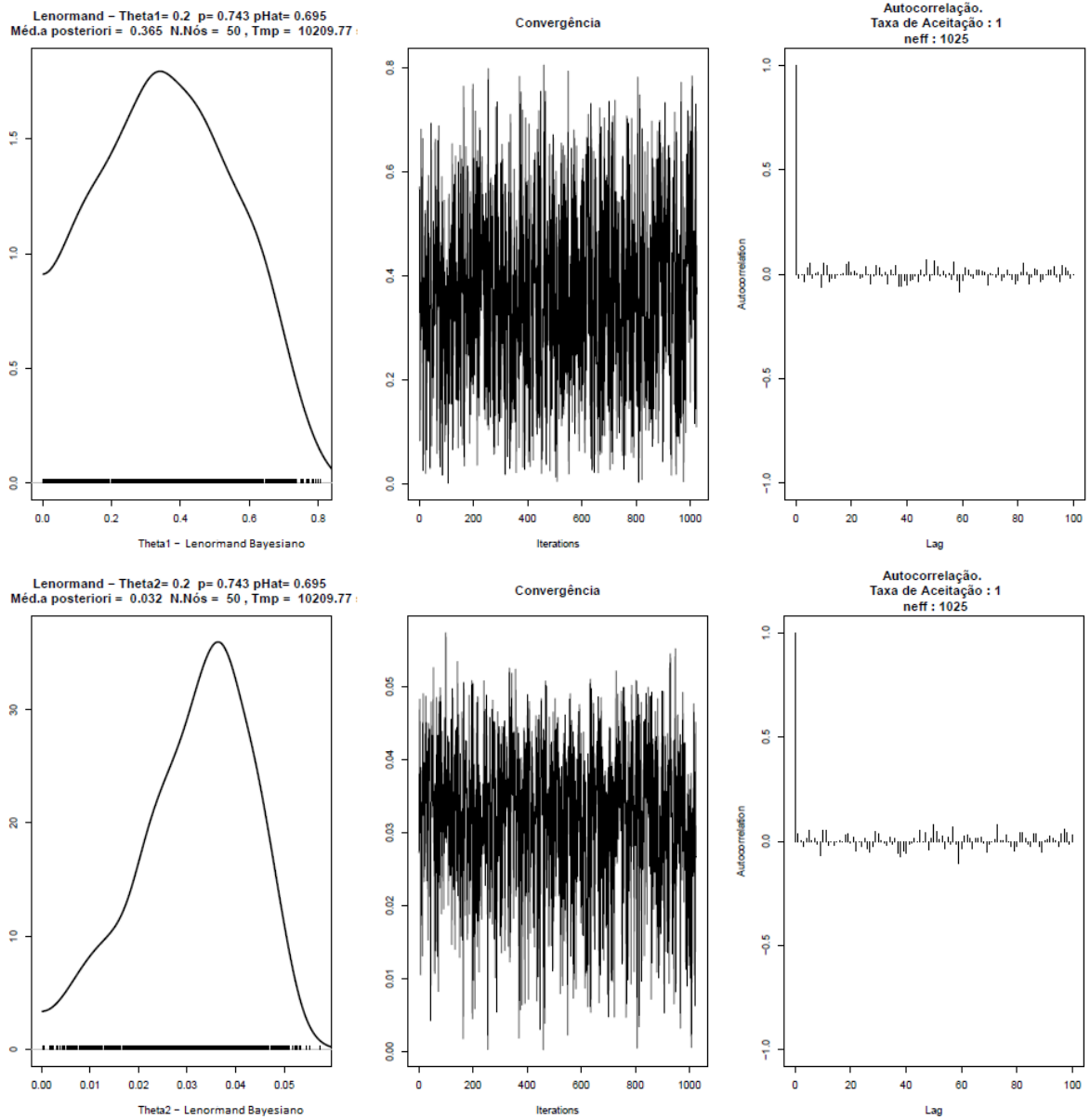


Figura B.5: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,2$  e  $\theta_2 = 0,2$ , **50** nós.

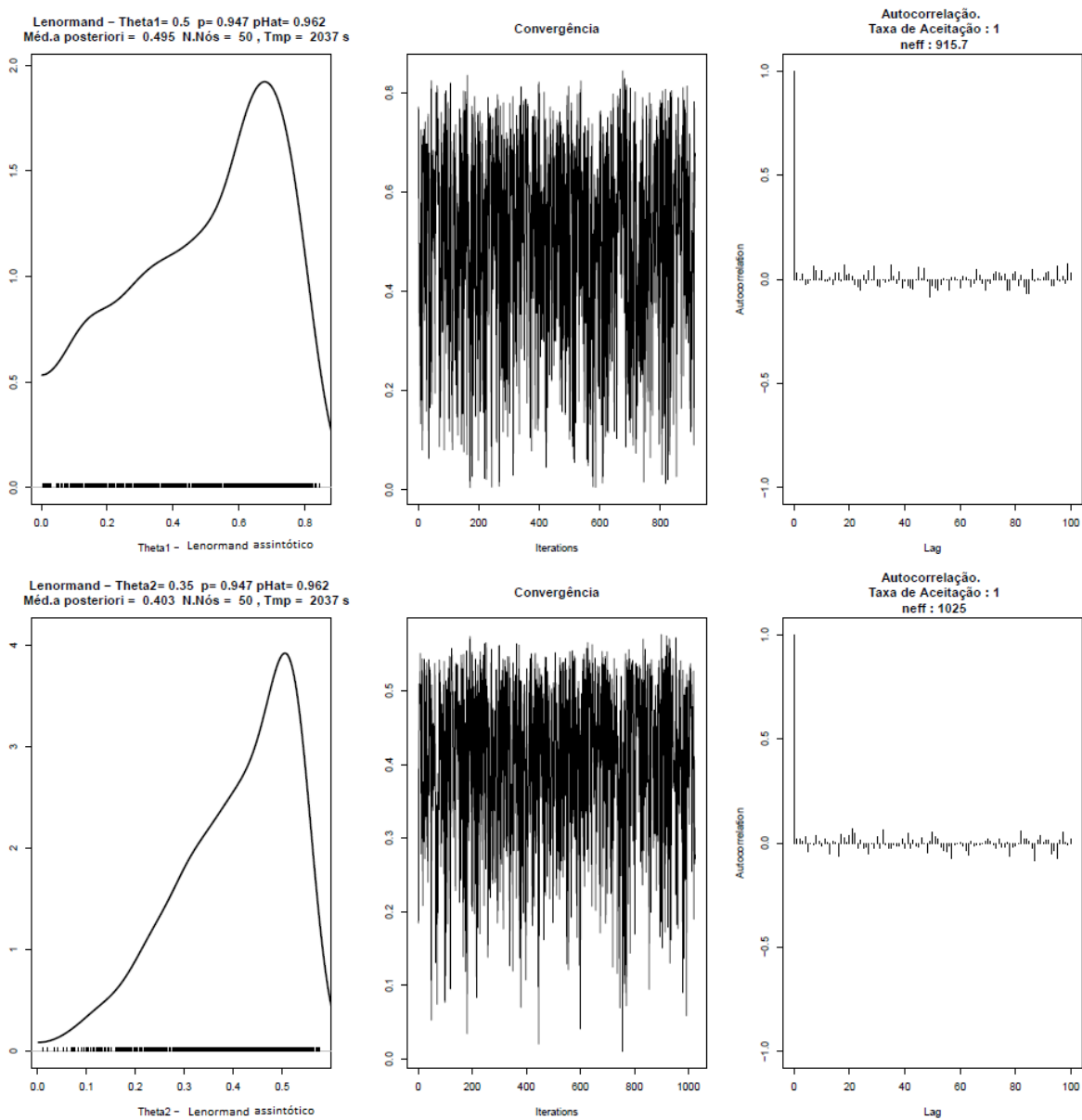


Figura B.6: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,5$  e  $\theta_2 = 0,35$ , **50** nós.



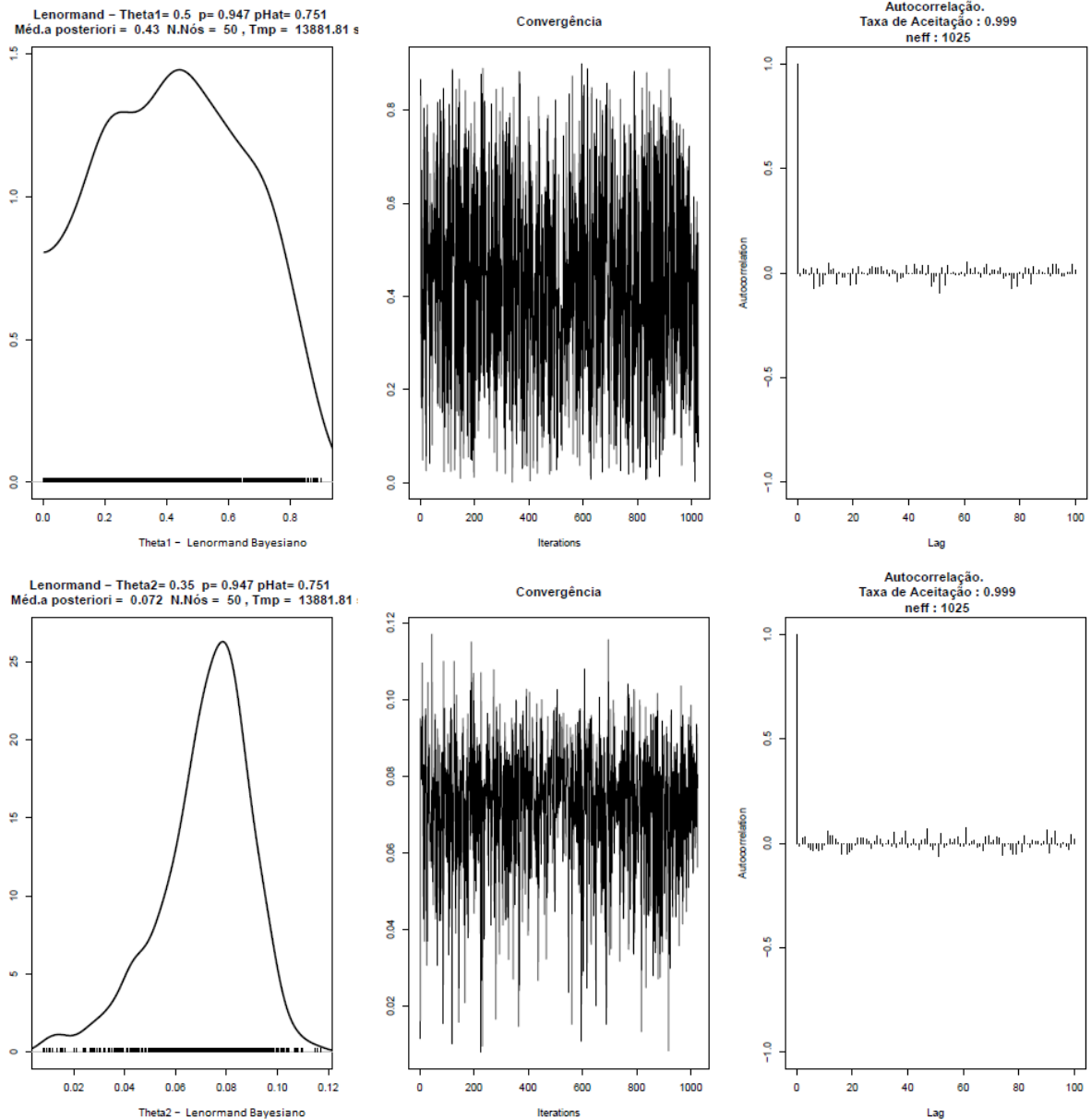


Figura B.7: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,5$  e  $\theta_2 = 0,35$ , **50** nós.

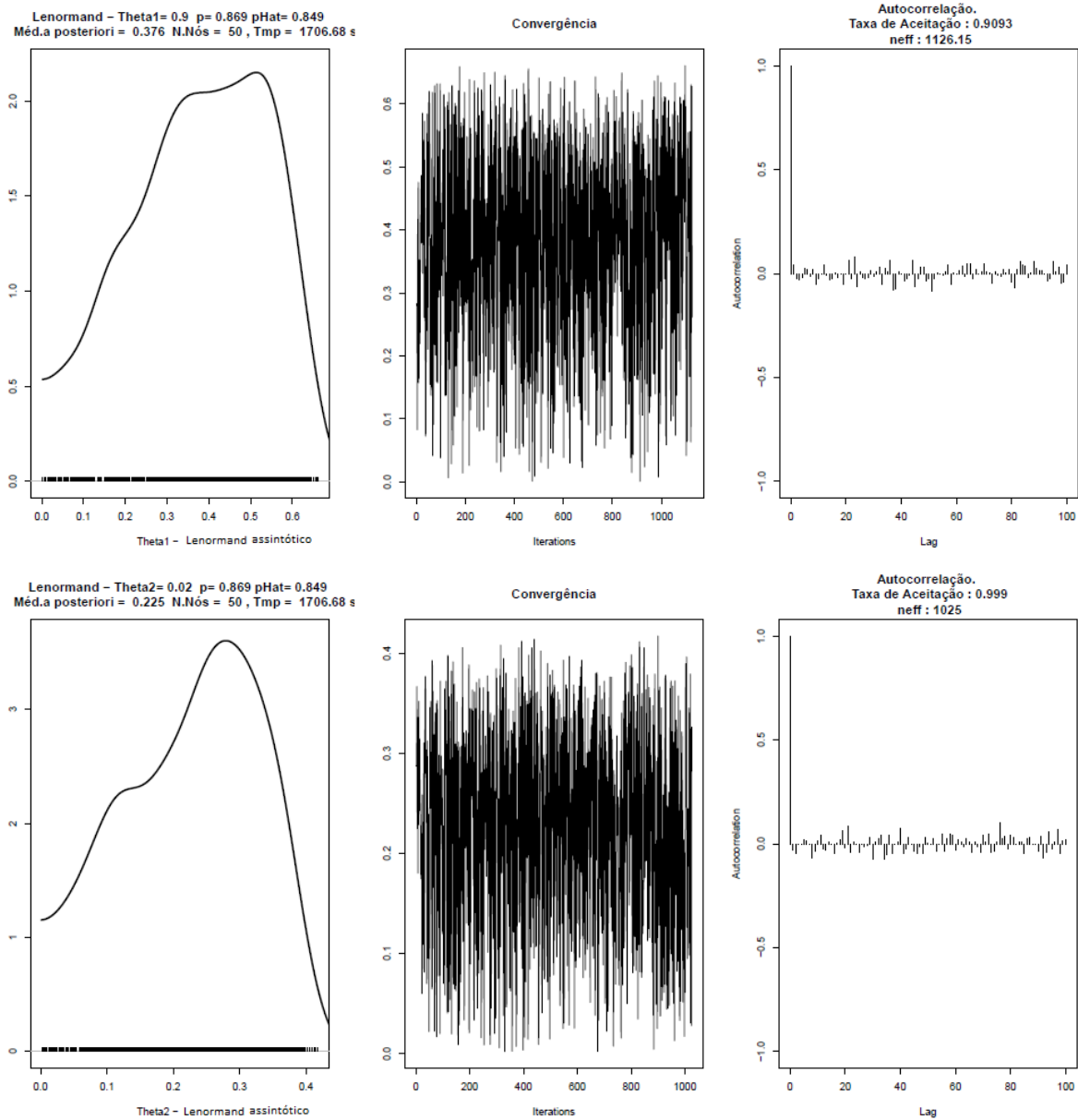


Figura B.8: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,9$  e  $\theta_2 = 0,02$ , **50** nós.

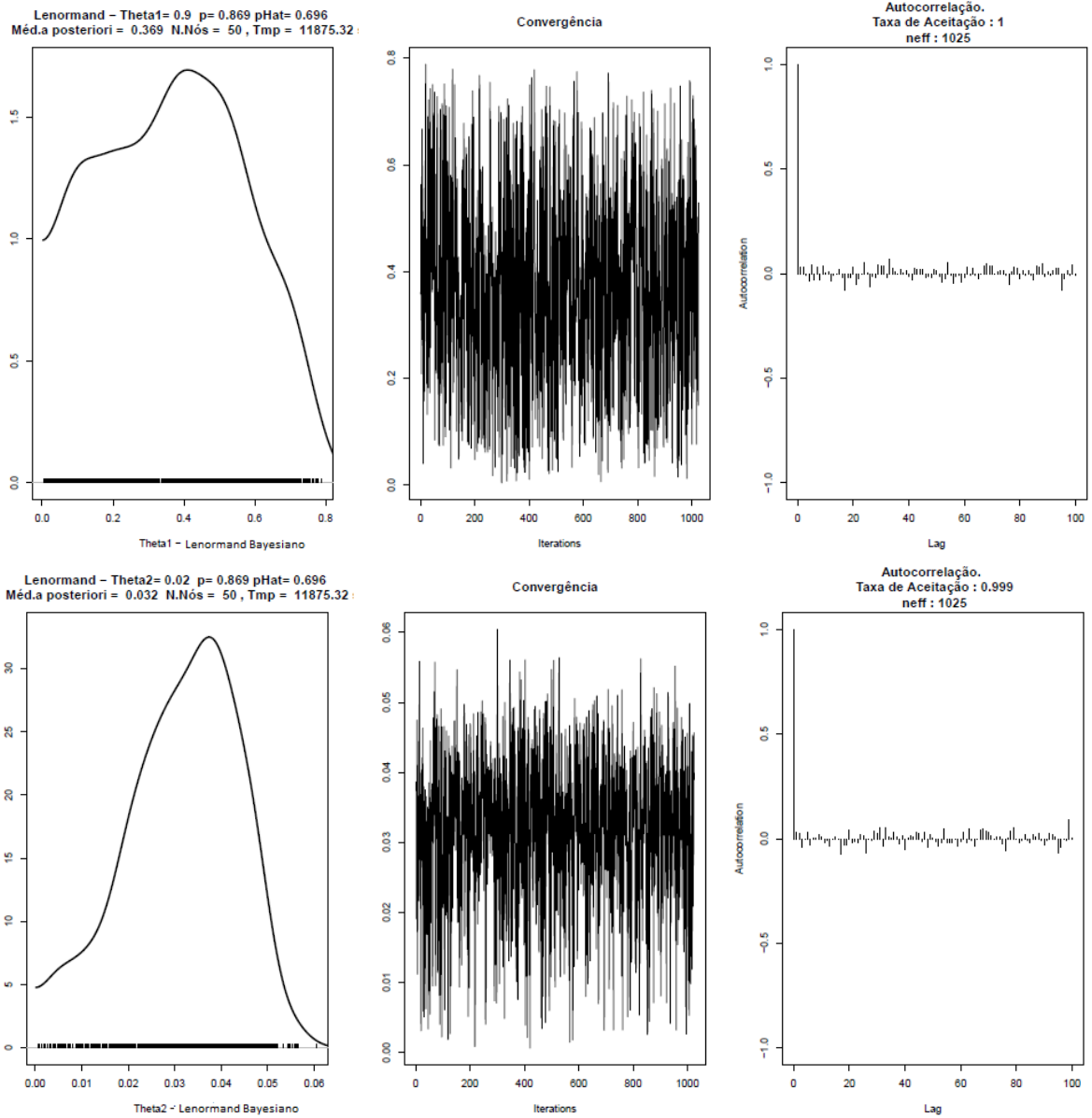


Figura B.9: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com lag de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,9$  e  $\theta_2 = 0,02$ , **50** nós.

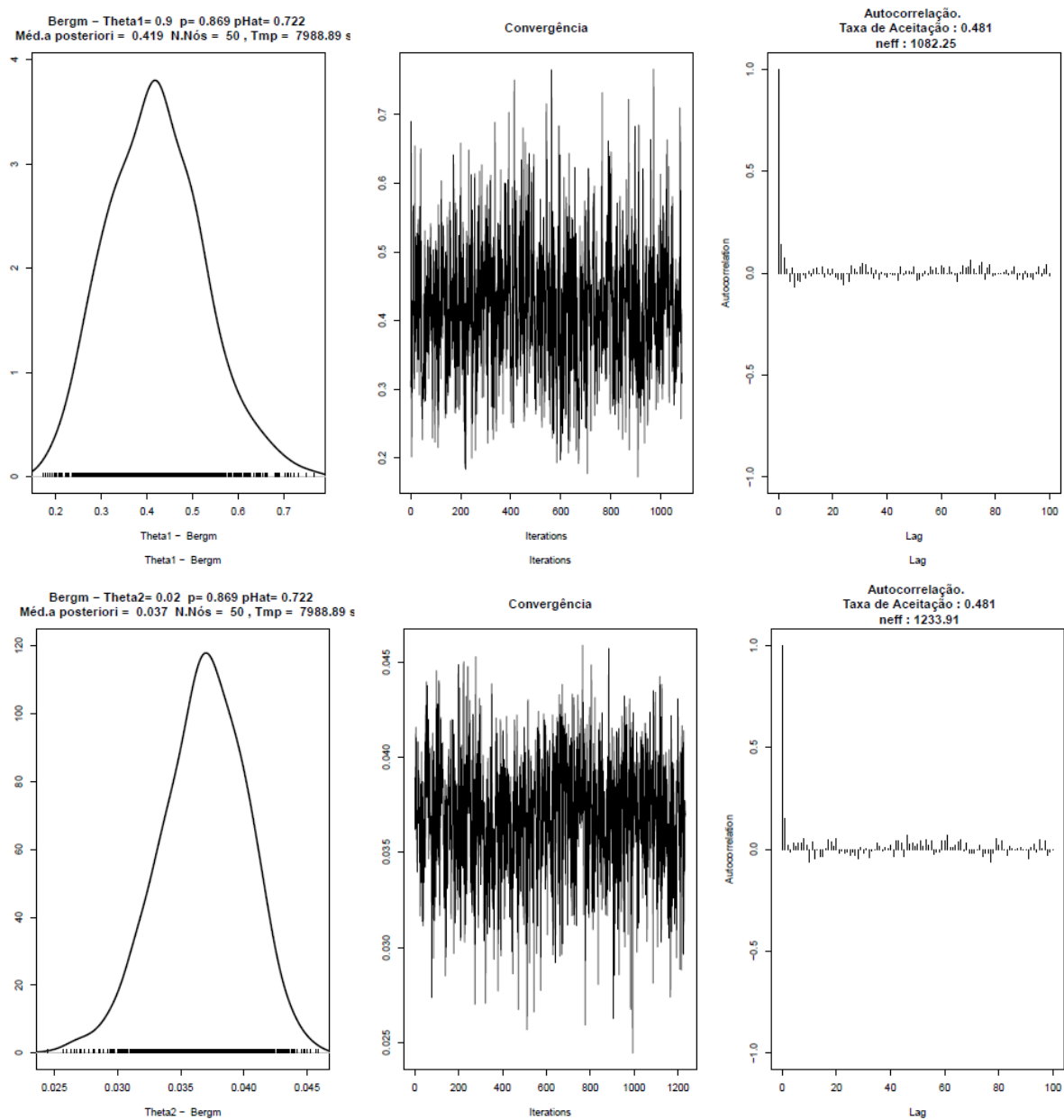


Figura B.10: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método Bergm :  $\theta_1 = 0,9$  e  $\theta_2 = 0,02$ , 50 nós.

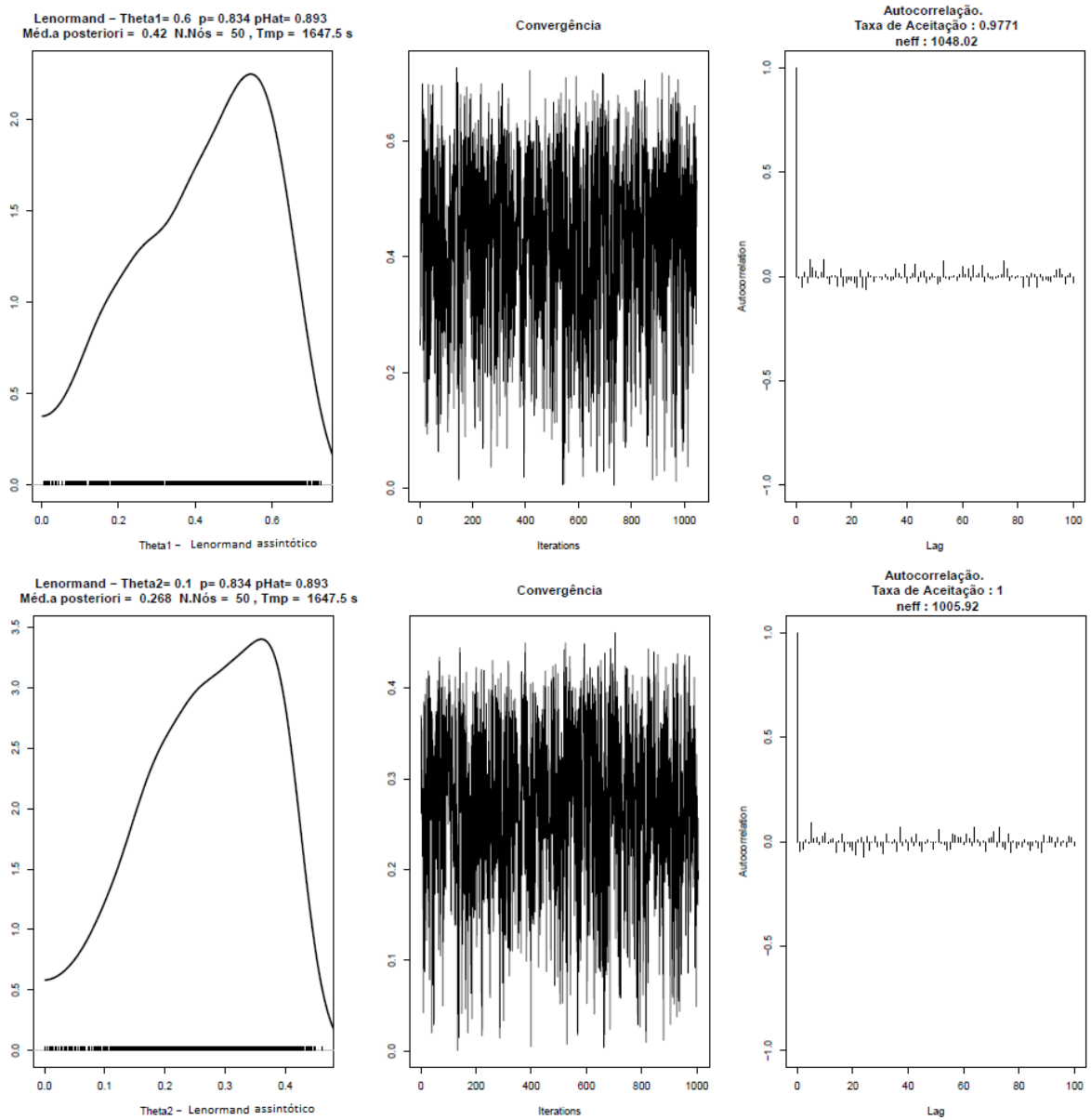


Figura B.11: Gráficos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,6$  e  $\theta_2 = 0,1$ , **50** nós.

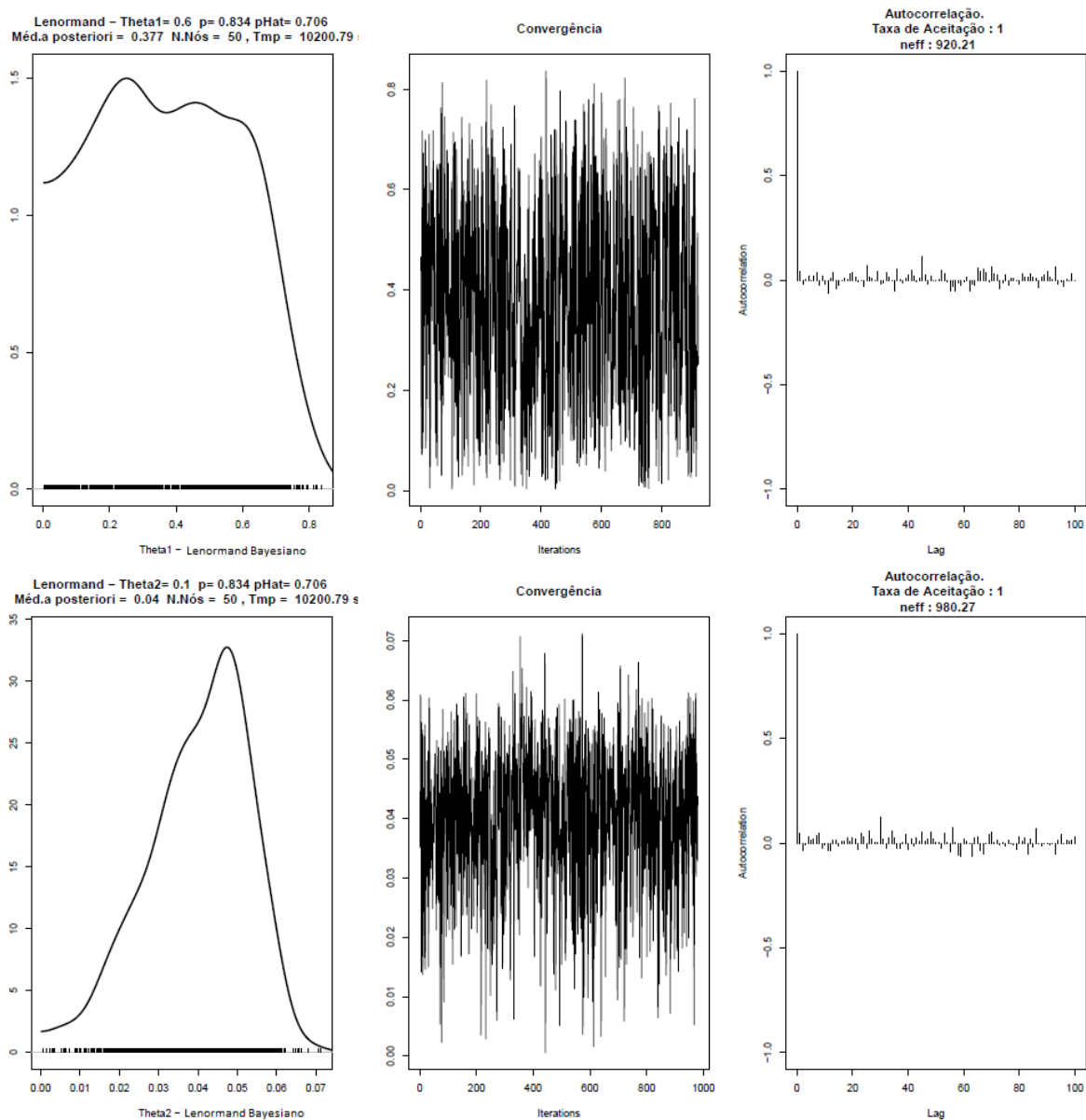


Figura B.12: Grafos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,6$  e  $\theta_2 = 0,1$ , 50 nós.

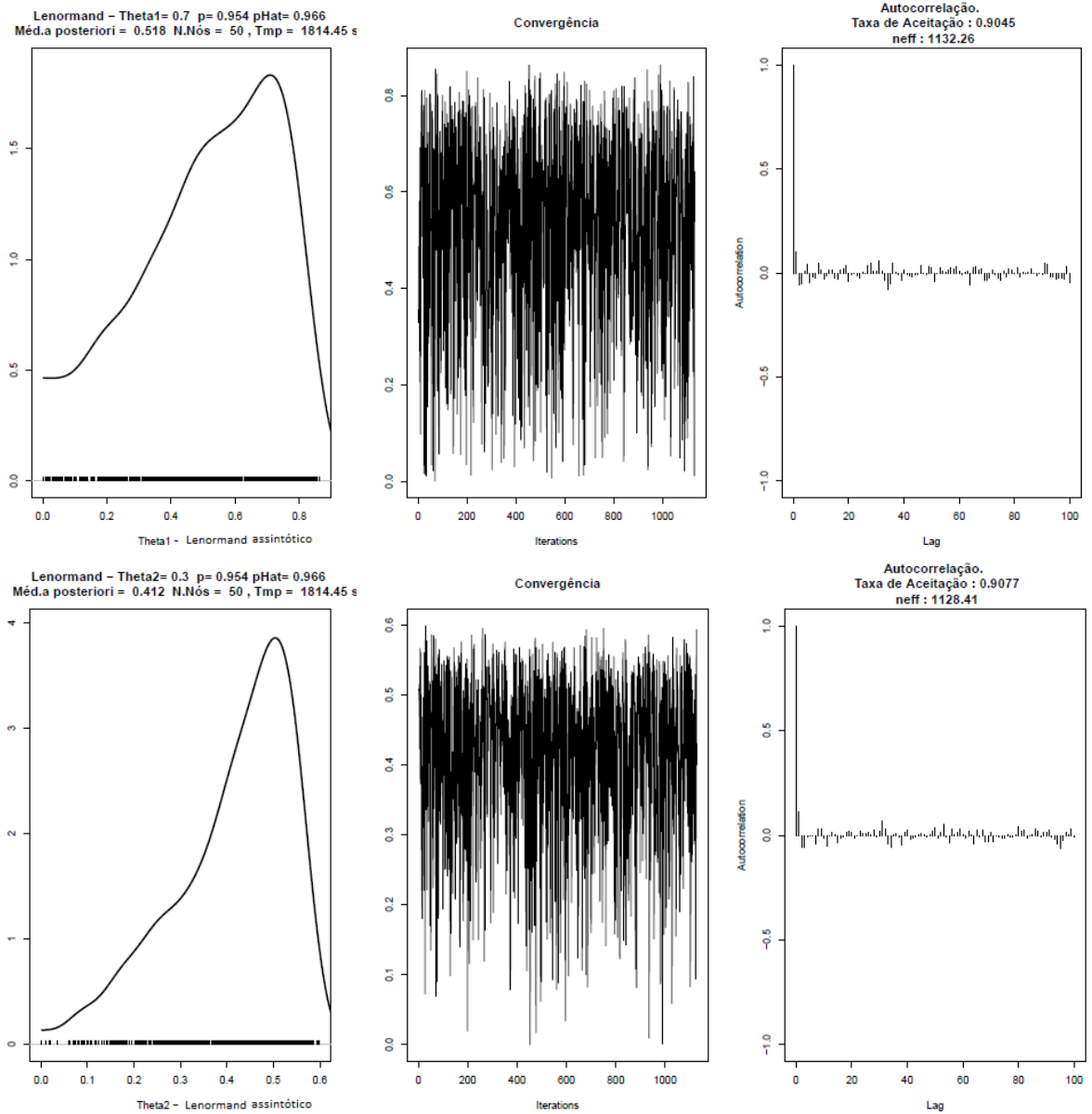


Figura B.13: Gráficos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com lag de tamanho 100 para o método ABC Lenormand, versão assintótica :  $\theta_1 = 0,7$  e  $\theta_2 = 0,3$ , **50** nós.

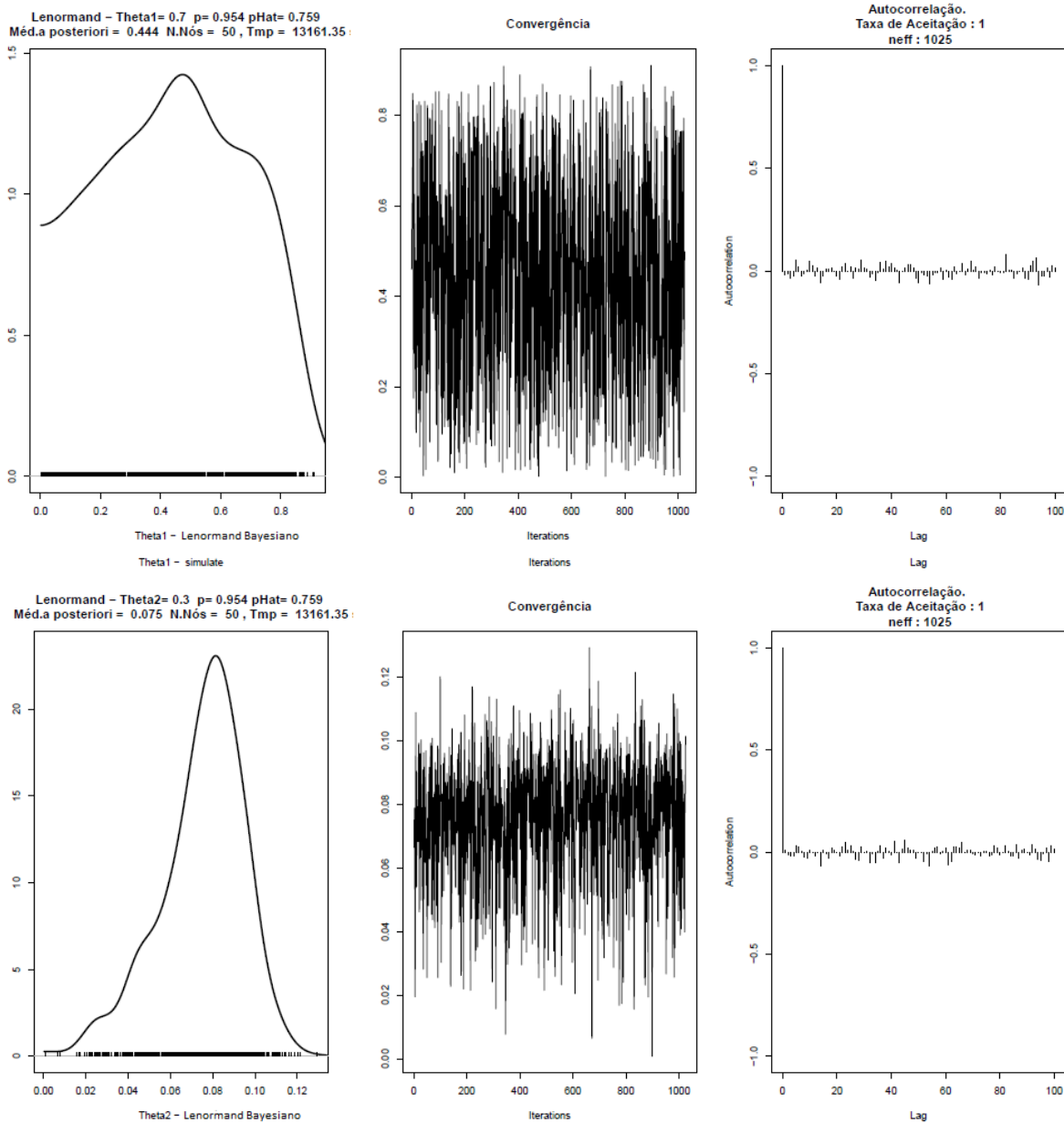


Figura B.14: Gráficos aleatórios exponenciais para arestas e triângulos : gráfico da distribuição *a posteriori*, gráfico da convergência da cadeia e correlograma das amostras com *lag* de tamanho 100 para o método ABC Lenormand, versão Bayesiana :  $\theta_1 = 0,7$  e  $\theta_2 = 0,3$ , **50** nós.