

UNIVERSIDADE FEDERAL DE ALFENAS

**PABLO CESCÓN PORTES**

**MODELAGEM BAYESIANA DOS NÍVEIS MÁXIMOS DO ÍNDICE DE PREÇOS AO  
CONSUMIDOR**

Alfenas/MG

2017

**PABLO CESCÓN PORTES**

**MODELAGEM BAYESIANA DOS NÍVEIS MÁXIMOS DO ÍNDICE DE PREÇOS AO  
CONSUMIDOR**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, área de concentração em Estatística Aplicada e Biometria da Universidade Federal de Alfenas, MG, como parte dos requisitos para a obtenção do título de Mestre. Linha de Pesquisa: Modelagem Estatística e Estatística Computacional.  
Orientador: Prof. Dr. Luiz Alberto Beijo.  
Coorientador: Prof. Dr. Fabricio Goecking Avelar.

Alfenas/MG

2017

Dados Internacionais de Catalogação-na-Publicação (CIP)  
Biblioteca Central da Universidade Federal de Alfenas

Portes, Pablo Cescon.

Modelagem Bayesiana dos níveis máximos do Índice de Preços ao Consumidor. / Pablo Cescon Portes. -- Alfenas/MG, 2017.

73 f.

Orientador: Luiz Alberto Beijo.

Dissertação (Mestrado em Estatística Aplicada e Biometria) -  
Universidade Federal de Alfenas, 2017.

Bibliografia.

1. Inflação. 2. Teoria bayesiana de decisão estatística. 3. Previsão estatística. 4. Países do BRICS. 5. Índices de preços ao consumidor.  
I. Beijo, Luiz Alberto. II. Título.

CDD-519.54



MINISTÉRIO DA EDUCAÇÃO  
Universidade Federal de Alfenas / UNIFAL-MG  
Programa de Pós-graduação em Estatística Aplicada e Biometria

Rua Gabriel Monteiro da Silva, 700. Alfenas - MG CEP 37130-000  
Fone: (35) 3299-1392 (Secretaria) (35) 3299-1121(Coordenação)  
<https://www.unifal-mg.edu.br/ppgeab/>



**PABLO CESCO PORTES**

**“MODELAGEM BAYESIANA DOS NÍVEIS MÁXIMOS DO ÍNDICE DE PREÇOS AO CONSUMIDOR”**


A Banca Examinadora, abaixo assinada, aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Área de Concentração: Estatística Aplicada e Biometria

Aprovado em: 10 de fevereiro de 2017.

Prof. Dr. Luiz Alberto Beijo  
Instituição: UNIFAL-MG

Assinatura: 

Profa. Dra. Cássia Helena Marchon  
Instituição: UNIFAL-MG

Assinatura: 

Prof. Dr. Manoel Vitor de Souza Veloso  
Instituição: UNIFAL-MG

Assinatura: 

## AGRADECIMENTOS

Agradeço, primeiramente, a Deus por ter me proporcionado capacidade para concluir os meus objetivos, entre eles a conclusão da Pós-Graduação em Estatística Aplicada e Biometria.

Também agradeço ao meu orientador Prof. Luiz Beijo, por todo ensinamento, pela paciência e por acreditar no trabalho, desde o princípio. E ao meu coorientador Prof. Fabrício Avelar, por todo conhecimento a mim transmitido. Ao colaborador Prof. Reinaldo Marques, por todo apoio e auxílio em todo o trabalho, além de toda disponibilidade em me ajudar.

Agradeço ao meu pai, Paulo, e minha mãe, Eliane, por todo apoio emocional, por sempre estarem do meu lado até mesmo nas situações ruins, e por acreditarem em mim. A minha irmã, Paula, com a qual sempre pude contar para compartilhar todas as emoções, boas e ruins, além de ser uma grande companheira.

A minha avó, Edith, e a todos os meus tios e tias, por me ajudarem quando precisei em diversas situações.

A todos os meus amigos, os que deixei de ver com frequência em Jundiaí, os que fiz em Varginha durante a graduação, e os mais recentes que pude conhecer em Alfenas. Também agradeço por todos os amigos e colegas do Mestrado, por todo conhecimento trocado e pelo espírito de ajuda mútua durante esses dois anos.

E por último, a minha namorada Laís, a quem, embora tenha conhecido no fim dessa caminhada, foi importante me apoiando e me ajudando.

*"It is clear that we are just an advanced breed of primates, on a minor planet orbiting around a very average star, in the outer suburb of one among a hundred billion galaxies... But, ever since the dawn of civilization, people have craved for an understanding of the underlying order of the world. There ought to be something very special about the boundary conditions of the universe. And what can be more special than that there is no boundary? And there should be no boundary to human endeavor. We are all different. However bad life may seem, there is always something you can do, and succeed at. While there is life, there is hope."*

(THE THEORY OF EVERYTHING, 2014)

## RESUMO

O Índice de Preços ao Consumidor Amplo (IPCA) é o índice utilizado pelo Banco Central do Brasil ao estabelecer suas metas inflacionárias. Por servir como uma referência de inflação, o IPCA é atentamente monitorado, tanto por investidores estrangeiros e brasileiros, quanto por gestores públicos. Sabe-se que uma inflação alta e descontrolada causa distorções e perdas econômicas no país, assim há um interesse por parte de administradores e gestores financeiros em prever a inflação máxima para um determinado período de tempo. Dessa forma, o objetivo do trabalho foi modelar os níveis máximos de IPCA, que podem ocorrer em um quadrimestre. A escolha de quadrimestres visa equiparar a análise com os intervalos entre as apresentações dos demonstrativos de cumprimento das metas fiscais por parte do Poder Executivo. Foi utilizada a distribuição Generalizada de Valores Extremos (do inglês *Generalized Extreme Values* - GEV) para modelagem. Para a estimação dos parâmetros da distribuição GEV utilizou-se o método da Máxima Verossimilhança e a Inferência Bayesiana. Na elicitação de informação para construção das distribuições *a priori*, foram utilizados dados de países economicamente semelhantes ao Brasil, a Rússia, China e Índia, os quais pertencem ao BRICS. Além disso, foram criadas diferentes combinações de distribuição *a priori*, usando informações desses países com diferentes estruturas de variância. Para avaliar qual melhor metodologia de estimação foram analisadas a acurácia e precisão das estimativas dos níveis máximos de inflação para determinados tempos de retorno. Os resultados permitiram observar que a abordagem Bayesiana, que utilizou como informação a média de dados dos países do BRICS para construção da distribuição *a priori* Normal Trivariada, levou a predições mais precisas e acuradas.

Palavras-chave: Distribuição Generalizada de Valores Extremos. Inflação. Inferência Bayesiana. Níveis de Retorno. BRICS.

## ABSTRACT

The Brazilian Consumer Price Index (CPI) is the index used by the Central Bank of Brazil in establishing its inflation targets. By serving as an inflation reference, the Brazilian CPI is closely monitored by foreign and Brazilian investors as well as by public managers. It is known that high and uncontrolled inflation causes distortions and economic losses in the country, so there is an interest on the part of managers and financial managers to predict maximum inflation for a certain period of time. Thus, the objective of the work was to model the maximum Brazilian CPI levels, which can occur in a four-month period. The choice of four-month periods aims to equate the analysis with the intervals between the presentations of the statements of compliance with the fiscal targets by the government. The Generalized Extreme Values (GEV) distribution was used for modeling. For the estimation of the parameters of the GEV distribution the maximum likelihood method and the Bayesian Inference were used. In the elicitation of information for the construction of the prior distributions, we used data from countries economically similar to Brazil: Russia, China and India, which belong to BRICS. In addition, different combinations of prior distribution were created, using information from these countries with different variance structures. In order to evaluate the best estimation methodology, the accuracy and precision of the estimates of the maximum inflation levels for certain return times were analyzed. The results showed that the Bayesian approach, which used as information the mean data of the BRICS countries for construction of the Normal Trivariate prior distribution, led to more accurate and accurate predictions.

Keywords: Generalized Extreme Values Distribution. Inflation. Bayesian Inference. Return Levels. BRICS.



## LISTA DE FIGURAS

Figura 1 - Evolução do IPCA entre os anos de 1980 e 2015.....	15
Figura 2 - Construção de uma série de máximos por POT.....	22
Figura 3 - Construção de uma série de máximos por Blocos de Máximos.....	23
Figura 4 - Série de níveis máximos quadrimestrais de IPCA.....	49

## LISTA DE TABELAS

Tabela 1 - Hiperinflação nos países da Europa entre as décadas de 1920 e 1940.....	16
Tabela 2 - Estimativas dos parâmetros da distribuição GEV calculadas via método de Máxima Verossimilhança, para os dados de níveis máximos de IPC da Rússia, China, Índia e a média dos três países (MRCI).....	36
Tabela 3 - Variâncias e Covariâncias das estimativas dos parâmetros da GEV calculadas via Método de Máxima Verossimilhança, para os dados de níveis máximos de IPC da Rússia, China, Índia e a média dos três países (MRCI).....	37
Tabela 4 - Estimativas, calculadas via método da Máxima Verossimilhança, e Desvios Padrão dos parâmetros da distribuição GEV e seus respectivos intervalos (I.C.) com 95% confiança.....	40
Tabela 5 - Intervalos (I.C.) com 95% confiança e amplitudes para os níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, a partir das distribuições Gumbel e GEV e seus respectivos EMAP e EQMP.....	41
Tabela 6 - Estimativas, desvios padrão (D.P.) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos, utilizando a distribuição a <i>priori</i> Normal I com diferentes matrizes de variâncias e covariâncias (Var), além dos critérios de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).....	42
Tabela 7 - Estimativas, desvios padrão (D.P.) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos, utilizando a distribuição a <i>priori</i> Normal II com diferentes matrizes de variâncias e covariâncias (Var), além dos critérios de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).....	43
Tabela 8 - Critérios de convergência Raftery-Lewis (R-L) e Geweke ( $Z_g$ ) das cadeias dos níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres dos modelos com distribuições a <i>priori</i> Normal I (N I) e Normal II (N II) e diferentes matrizes de variâncias e covariâncias (Var).....	44
Tabela 9 - Intervalos HPD com 95% de credibilidade dos níveis de retorno para o tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, calculados dos modelos	

	utilizando a distribuição a <i>priori</i> Normal I com diferentes matrizes de variâncias e covariâncias (Var).....	45
Tabela 10-	Intervalos HPD com 95% de credibilidade dos níveis de retorno para o tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, calculados dos modelos utilizando a distribuição a <i>priori</i> Normal II com diferentes matrizes de variâncias e covariâncias (Var).....	46
Tabela 11-	Erro Médio Absoluto de Predição (EMAP) e Erro Quadrático Médio de Predição (EQMP) dos modelos Bayesianos com distribuição a <i>priori</i> Normal I e Normal II e diferentes matrizes de variâncias e covariâncias (Var).....	47
Tabela 12-	Estimativas, desvios padrão (D.P.) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos selecionados, utilizando a distribuição a <i>priori</i> Normal Trivariada com diferentes matrizes de variâncias e covariâncias (Var) e respectivos resultados dos critérios de convergência de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).....	49
Tabela 13-	Critérios de convergência Raftery-Lewis (R-L) e Geweke ( $Z_g$ ) das cadeias dos níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres dos modelos Bayesianos selecionados.....	50
Tabela 14-	Níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, seus respectivos intervalos HPD com 95% de credibilidade, calculados a partir dos modelos Bayesianos selecionados.....	50

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	11
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> .....	13
2.1	ÍNDICE DE PREÇOS AO CONSUMIDOR.....	13
2.2	ANÁLISE DE EVENTOS EXTREMOS.....	21
2.3	INFERÊNCIA BAYESIANA.....	26
2.3.1	Aplicações na Análise de Eventos Extremos.....	28
2.4	TESTES AUXILIARES NA ANÁLISE DE EVENTOS EXTREMOS.....	29
2.4.1	Teste de Ljung-Box.....	30
2.4.2	<i>Runs-Test</i> .....	31
2.4.3	Teste de Kolmogorov-Smirnov.....	31
2.4.4	Teste de Razão de Verossimilhança.....	32
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	33
3.1	DETERMINAÇÃO DO TAMANHO DOS BLOCOS.....	33
3.2	DADOS.....	33
3.2	MODELAGEM.....	34
3.2.1	Análises Preliminares.....	34
3.2.2	Estimação via Máxima Verossimilhança.....	35
3.2.3	Estimação via Inferência Bayesiana.....	35
3.2.4	Avaliação dos Modelos.....	38
3.3	RECURSOS COMPUTACIONAIS.....	39
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	40
4.1	SELEÇÃO DOS MODELOS.....	40
4.2	AJUSTES DOS MODELOS.....	48
<b>5</b>	<b>CONCLUSÕES</b> .....	53
	<b>REFERÊNCIAS</b> .....	54
	<b>APÊNDICES</b> .....	63
	<b>ANEXOS</b> .....	71

## 1 INTRODUÇÃO

Índices macroeconômicos são importantes para se analisar a situação econômica, política e social em geral de um país. No Brasil, o Índice de Preços ao Consumidor Amplo (IPCA) é utilizado pelo Banco Central do Brasil (BACEN) para estipular as metas inflacionárias. É imprescindível para investidores e administradores trabalhar com boas previsões de inflação. Porém, muitas vezes o interesse está atrelado à previsão do índice máximo de inflação, para um determinado período de tempo.

A análise da inflação máxima quadrimestral é importante pois concide com os intervalos entre as divulgações dos demonstrativos de cumprimento das metas fiscais por parte do Poder Executivo. O cumprimento das metas fiscais, assim como das metas inflacionárias, podem indicar uma instabilidade econômica ou desequilíbrio na condição das políticas macroeconômicas, daí sua relevância para os agentes econômicos.

Uma das formas de se modelar máximos é utilizando distribuição generalizada de valores extremos (GEV). Há diversos métodos de estimação dos parâmetros da GEV, um deles é o método da Máxima Verossimilhança, que possui boas propriedades assintóticas. Uma metodologia que pode ser usada é a Inferência Bayesiana que consegue incorporar informações *a priori* do conjunto de dados, o que permite diminuir as incertezas acerca dos parâmetros a serem estudados, fazendo com que essa abordagem para estimação seja eficaz também para pequenos tamanhos amostrais.

No Brasil o IPCA, assim como outros índices econômicos, não dispõe de uma série extensa de dados, além disso, quando se extrai os máximos do conjunto por blocos acaba-se reduzindo mais ainda o tamanho da série, o que pode tornar o método da Máxima Verossimilhança menos preciso para estimação dos parâmetros. Nesse caso, a Inferência Bayesiana pode apresentar maior eficiência para a estimação dos parâmetros e predição dos níveis de inflação máxima.

Na análise de níveis máximos de IPCA uma possibilidade seria usar informações *a priori* a partir de dados de países economicamente similares ao Brasil, especificamente Rússia, Índia e China, os quais pertencem ao BRICS. Utilizando a abordagem Bayesiana e considerando diferentes combinações de distribuição *a priori*, foi possível avaliar se houve melhora na eficiência e precisão das previsões de máximos quadrimestrais de inflação.

Com isso, o trabalho objetivou modelar os níveis máximos quadrimestrais de IPCA

utilizando a distribuição Generalizada de Valores Extremos, aplicando, para a estimação dos parâmetros, o método da Máxima Verossimilhança e a Inferência Bayesiana. Pretende-se, também, baseando-se em dados dos países do BRICS, avaliar diferentes formas de elicitar informações considerando uma *priori* Normal Trivariada e comparar os resultados obtidos pelo método da Máxima Verossimilhança e pela Inferência Bayesiana, avaliando a predição da inflação máxima para os tempos de 2, 3, 4, 5 e 10 quadrimestres.

A utilização da Inferência Bayesiana se mostrou melhor quando a distribuição *a priori* foi construída com base na média dos dados dos países dos BRICS. Essa análise pode ser feita através da verificação dos erros de predição e estimativas intervalares das inflações máximas quadrimestrais.

## 2 REFERENCIAL TEÓRICO

Nessa seção serão apresentados os conceitos sobre Índice de Preços ao Consumidor, análise de eventos extremos, Inferência Bayesiana e dos testes auxiliares utilizados durante o trabalho.

### 2.1 ÍNDICE DE PREÇOS AO CONSUMIDOR

Os índices macroeconômicos são indicadores utilizadas pelo governo e agentes econômicos para análise da situação econômica do país. Muitos são esses indicadores, como o Produto Interno Bruto (PIB), que é a soma do valor de tudo aquilo que é produzido no país em um determinado período de tempo, a Taxa de Câmbio, que mede o preço da moeda nacional em relação à outra, a Taxa do Sistema Especial de Liquidação e de Custódia (Selic), a taxa média dos financiamentos diários apurados para títulos federais, entre outros. Alguns índices buscam acompanhar os preços gerais do país como o Índice Geral de Preços do Mercado (IGP-M) e o Índice de Preços ao Consumidor Amplo (IPCA), sendo este último utilizado para fixar as metas inflacionárias do governo.

O IPCA mede a variação de preços dos bens e serviços consumidos por famílias com renda entre 1 e 40 salários mínimos e que habitam as principais capitais do Brasil: São Paulo, Rio de Janeiro, Belo Horizonte, Recife, Salvador, Porto Alegre, Belém, Curitiba, Fortaleza, Brasília, Goiânia, Vitória e Campo Grande (IPEA, 2016). A cesta de bens e serviços que compõe o cálculo do IPCA é definida a partir das Pesquisas de Orçamentos Familiares (POFs) do Instituto Brasileiro de Geografia e Estatística (IBGE). Geralmente essas pesquisas são realizadas de seis em seis anos e o IPCA é medido mensalmente pelo IBGE.

Sabe-se que o IPCA é um índice de Laspeyres, ou seja, mede a variação do custo de compra da cesta de bens e serviços entre o período atual e o período base (HILL, 2006). Dessa forma o IPCA pode ser escrito conforme:

$$L_{0,t} = \frac{\sum_{i=1}^n p_i^{(t)} q_i^{(0)}}{\sum_{i=1}^n p_i^{(0)} q_i^{(0)}} = \sum_{i=1}^n \left( \frac{p_i^{(0)} q_i^{(0)}}{\sum_{j=1}^n p_j^{(0)} q_j^{(0)}} \right) \left( \frac{p_i^{(t)}}{p_i^{(0)}} \right) = \sum_{i=1}^n w_i \left( \frac{p_i^{(t)}}{p_i^{(0)}} \right) \quad (2.1)$$

em que  $L_{0,t}$  é o índice de Laspeyres para o tempo  $t$  tendo como base o período 0,  $p_i$  e  $q_i$  são, respectivamente, preço e quantidade comprada do subitem  $i$  e  $w_i$  é o peso do subitem  $i$  na cesta total de bens da população-alvo no período de referência.

De acordo com IPEA (2016) a POF de 2008/2009 teve uma diminuição no número de itens em relação às duas últimas (1995-1996 e 2002-2003). Em que as duas primeiras tinham 93 e 94 itens cada, a última contempla 84 itens, isso após a atualização de janeiro de 2014. Nos Anexos podem ser vistas as três POFs e seus respectivos itens.

De acordo com Bryan e Cecchetti (1993) é cada vez mais importante analisar em quais índices de preços a política monetária de um país se baseia. O IPC ainda é o índice mais comum e popularmente usado para basear a inflação. Quanto ao IPCA, segundo Assaf Neto (2008) é a medida mais relevante para a inflação, do ponto de vista político e econômico, pois, é esse índice que é selecionado pelo Conselho Monetário Nacional como referência para o sistema de metas de inflação, em vigor desde 1999 no Brasil.

Assaf Neto (2008) explica que uma das consequências da inflação é a queda do poder de compra da moeda, dessa forma a inflação pode ser um grande problema para um país, conforme desestimula investimentos e diminui o poder de compra da população. No Brasil muitas foram as medidas tomadas por governantes para tentar frear a inflação, como: congelamento de preços e salários, abertura econômica, ajustes fiscais, congelamento de depósitos bancários, e os planos econômicos implementados: Plano Cruzado (1986), Plano Bresser (1987), Plano Verão (1989), Plano Collor (1990), plano Collor 2 (1991) e Plano Real (1994).

O último plano citado merece uma maior atenção, pois, de acordo com Filgueiras (2006), o Plano Real foi diferente, visto que os demais planos nunca completaram um ano de existência. Implementado durante o governo Itamar Franco, no final do ano de 1993, inicialmente se chamaria "Plano FHC", porém foi batizado como "Plano Real" em 1994. Teve como intuito ser um plano de estabilização econômica no Brasil, na tentativa de conter a hiperinflação dos preços de bens e serviços vividos nessa época.

Desde o anúncio de um novo plano (em dezembro de 1993) até a implementação da moeda brasileira atual, o Real, em julho de 1994, decorreu-se três etapas, para que não acontecesse como no Plano Cruzado (implementação abrupta, resultando no fracasso): o ajuste fiscal, a criação da Unidade de Referência do Valor (URV) e a instituição da nova moeda (FILGUEIRAS, 2006). É possível averiguar na Figura 1 que realmente o Plano Real estabilizou a hiperinflação dos preços de bens e serviços:



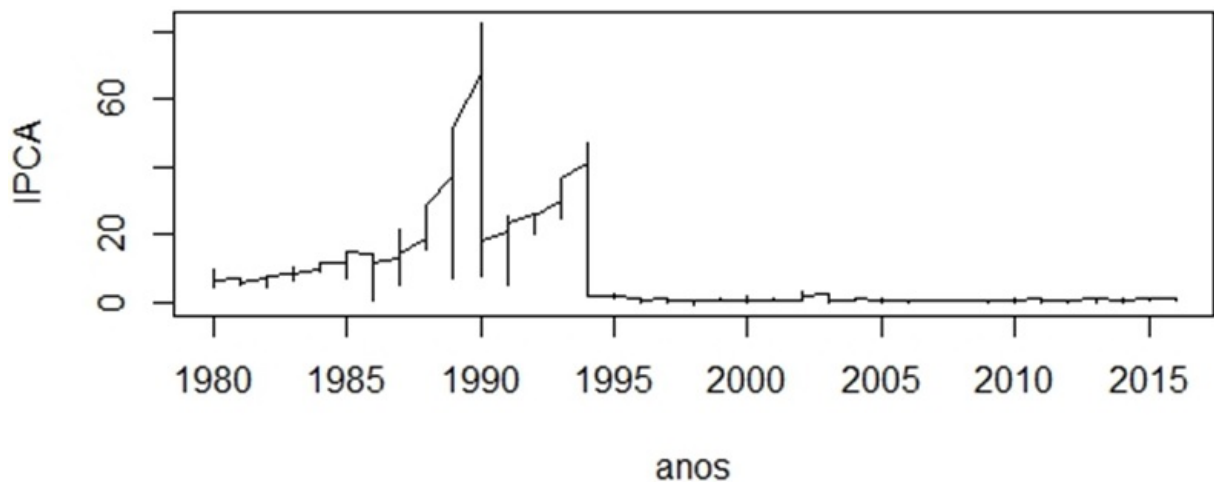


Figura 1 – Evolução do IPCA entre os anos de 1980 e 2015.  
 Fonte: Construída pelo autor, utilizando dados do IPEA-DATA (2016).

É possível observar na Figura 1 a estabilização considerável da inflação na economia brasileira nos períodos que sucedem a implementação do Plano Real. Porém, essa estabilização da inflação imposta pelo governo acarretou em algumas consequências para a economia. Carneiro (2002) apresenta uma delas como sendo o aumento da sensibilidade da economia brasileira à economia internacional, além do grande aumento do déficit da balança comercial, ou seja, as importações começaram a crescer em relação às exportações. Para Gremaud et al. (2009), a diminuição do IPCA resultou no aumento da demanda agregada e da atividade econômica, pelo fato de ter ocorrido elevação do poder aquisitivo das classes de baixa e média rendas. Além disso, o mesmo autor explicita que houve também um crescimento das operações de crédito, pois o consumidor se sentira atraído a tomar empréstimos, o que aconteceu por causa da “ilusão monetária”, ou seja, uma diminuição da taxa nominal de juros, acompanhada de um aumento excessivo da taxa real de juros. Garcia e Perron (1996) explicam como isso ocorre. A taxa real de juros normalmente é calculada conforme a seguinte equação:

$$i_r = i_n - \pi \quad (2.2)$$

em que  $i_r$  é a taxa real de juros,  $i_n$  é a taxa nominal de juros e  $\pi$  a inflação.

Mundell (1963) afirma que a inflação está negativamente relacionada à taxa real de juros. Tanto Garcia e Perron (1996) como Fama e Gibbons (1982) se dedicam a comprovar essa relação. No primeiro trabalho, os autores modelaram tanto a taxa de IPC quanto a taxa real de juros, utilizando séries temporais, dessa maneira conseguiram averiguar que as duas

variáveis estão correlacionadas. No segundo artigo, os autores criaram modelos de regressão linear e, sabendo que o modelo foi bem ajustado, concluíram a mesma relação encontrada no primeiro.

Nesses contextos, é compreendida a importância de se estudar o IPC e IPCA para análises da conjuntura econômica no Brasil. Porém esse incidente de hiperinflação e alta dos preços não foi um processo exclusivo do Brasil da década de 80 e 90. Um dos casos mais famosos é o da Alemanha entre as duas Guerras Mundiais, conforme Webb (1989), em 1913 toda a moeda em circulação na Alemanha totalizava 6 bilhões de marcos, após 10 anos os mesmos 6 bilhões de marcos não conseguiam comprar um pão de centeio. Nesse mesmo período, ocorreram outras situações de hiperinflações na Europa, como se pode perceber pela Tabela 1.

Tabela 1 – Hiperinflação nos países da Europa entre as décadas de 1920 e 1940.

País	Início	Fim	Taxa mensal de inflação (%)
Áustria	Out. 1921	Ago. 1922	47
Alemanha	Ago. 1922	Nov. 1923	322
Grécia	Nov. 1943	Nov. 1944	365
Hungria I	Mar. 1923	Fev. 1924	46
Hungria II	Ago. 1945	Jul. 1946	19800
Polônia	Jan. 1923	Jan. 1924	82
Rússia	Dez. 1921	Jan. 1924	57

Fonte: Cagan ed. Por Friedman (1956).

A Europa viveu grandes hiperinflações na primeira metade do século XX: um produto que custava 1 unidade monetária na Hungria em 1945, por exemplo, iria custar  $3,8 \times 10^{27}$  unidades monetárias um ano depois (CAGAN, 1956). Hiperinflações também aconteceram em países da América Latina nas décadas de 1980 e 1990. Países como a Bolívia, Argentina, Nicarágua e Peru sofreram períodos de hiperinflação que variaram de 15% a 45% por mês (BLANCHARD, 1999).

O Brasil não foi derrotado em uma guerra como os países europeus que viveram hiperinflação no passado, apesar das condições gerais serem as mesmas, como grande dívida externa em grande parte dívida pública, o período inflacionário brasileiro teve um forte componente inercial, assim a inflação futura estava fortemente vinculada à inflação passada (PEREIRA e NAKANO, 1991). Nos países como a Alemanha, a dívida externa cresceu muito rápido devido a Primeira Guerra Mundial, dessa forma houve um aumento na oferta de moeda para

pagamento da dívida o que, de acordo com a equação quantitativa da moeda, acarreta em um grande aumento dos preços, como pode ser visto a partir da identidade:

$$MV = PY \quad (2.3)$$

em que  $M$  é a quantidade de moeda,  $V$  a velocidade de circulação da moeda,  $P$  é o preço geral dos produtos e  $Y$  o PIB do país. Assumindo-se que  $V$  é relativamente estável em um período curto de tempo e que a economia encontra-se próxima do pleno emprego de seus recursos produtivos, um aumento de  $M$  acarretará em um aumento de  $P$ .

O IPC também tem sua importância no âmbito de análise de investimentos, Glebova et al. (2013) encontraram grande relação entre o IPC da Rússia com a atratividade de investimentos nas indústrias químicas da região do Tartaristão (parte da macrozona econômica da Rússia Europeia e dos Urais), Fischer (2013) utiliza-se da técnica de dados em painel para comprovar que altas inflações inibem os investimentos de pequenas empresa. Além disso, o mesmo expõe que a incerteza sobre a inflação futura tem consequências negativas sobre o crescimento econômico.

Tang e Xiong (2012) afirmam que o IPC, bem como a inflação, tem relação com o mercado de commodities dos Estados Unidos, assim como Diewert (2001) que também apresenta o IPC como medida para estabelecer preços de um grupo de *commodities*. Por outro lado Blomberg e Harris (1995) admitem que esta relação não é causal. IPC também é usado para tentar prever preços de bens de consumo, como fizeram Jo e Lusk (2016), ao utilizar o IPC para prever preços de comidas e bebidas, e Sharma (2016), que tendo dados de IPC de 54 países conseguiu prever o preço do ouro.

A importância da análise de IPCA e Inflação no Brasil também está atrelada ao cálculo de taxa de juros de ativos de renda fixa. Quanto a isso, Berger (2015) comenta que alguns brasileiros tinham a falsa impressão de estar lucrando com uma poupança que pagava 30% de juros ao mês na época de hiperinflação, mas na verdade essas mesmas pessoas estariam perdendo dinheiro, visto que os preços geralmente subiam numa proporção maior do que a poupança.

Há Títulos Públicos Federais, ativos de renda fixa, que levam em consideração o valor da inflação no cálculo de seus juros. A administração desses títulos está nas mãos da Secretaria do Tesouro Nacional (STN), criada em 1986 dentro da estrutura do Ministério da Fazenda. Os mesmos são emitidos em mercado e formam a dívida pública federal interna (BERGER, 2015).

Existem muitos tipos de títulos públicos federais como: Letras do Tesouro Nacional (LTN), Letras Financeiras do Tesouro (LTF), Nota de Tesouro Nacional, Série “B”(NTN-B), Nota do Tesouro Nacional, Série “C”(NTN-C), entre muitos outros. O decreto nº 3859, de 4 de julho de 2001 apresenta algumas características para a NTN-B:

- a) prazo: definido pelo ministro da Fazenda, quando da emissão do título;
- b) taxa de Juros: definida pelo ministro de Estado da Fazenda, quando da emissão em porcentagem ao ano, sobre o valor nominal;
- c) valor Nominal: múltiplos de R\$ 1000 (nominativa e negociável);
- d) atualização do valor nominal: pela variação do IPCA do mês anterior ao resgate e do mês anterior à data-base do título;
- e) pagamento de juros: semestralmente;
- f) resgate da parcela principal: em parcela única na data do vencimento.

Existem NTN em que a atualização do valor nominal é feita, por exemplo, pela variação do IGP-M, que é o caso da NTN-C, ou ainda pela variação da cotação do dólar americano (US\$), como a NTN-D. Existem, além dessas NTN citadas, mais sete: NTN-A, NTN-F, NTN-H, NTN-I, NTN-M, NTN-P, NTN-R e uma sub-série NTN-R2 (DECRETO nº 3859).

Conforme o Ministério da Fazenda (2016), o cálculo de preço do título público NTN-B é feito seguindo a equação:

$$\text{Preço} = \text{Cotação}(\%) * \text{VNA} \quad (2.4)$$

em que VNA é Valor Nominal Atualizado (no caso da NTN-B pelo IPCA) e a Cotação (%) é encontrada de acordo com a equação:

$$\text{Cotação}(\%) = \sum_{i=1}^n \frac{100((1+j)^{0,5} - 1)}{(1 + TIR)^{\frac{d_{ui}}{252}}}, \quad (2.5)$$

em que  $j$  é a taxa de juros,  $TIR$  é a taxa interna de retorno (padrão  $du/252$  % ao ano (a.a.)) e  $d_{ui}$  é o número de dias úteis entre a data de liquidação  $i$  (inclusive) e a data de vencimento (exclusive).

A Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (AN-BIMA) mostra que há três técnicas para se calcular o VNA. Sabendo que o IPCA é divulgado pelo IBGE na primeira quinzena do mês, a data do cálculo muda a técnica de cálculo do VNA. Se a data do cálculo coincide com décimo quinto dia do mês (mês fechado), a fórmula para o

cálculo pode ser dada conforme a equação:

$$\text{VNA} = \frac{\text{IPCA}_t}{\text{IPCA}_0} * \text{VN}_{db}, \quad (2.6)$$

em que o  $\text{VN}_{db}$  é o valor nominal na data-base (*db*),  $\text{IPCA}_0$  é o IPCA divulgado pelo IBGE no mês anterior à data-base e  $\text{IPCA}_t$  é o último IPCA divulgado pelo IBGE. Caso a data do cálculo seja anterior ao décimo quinto dia do mês a equação seguinte mostra o cálculo do VNA:

$$\text{VNA} = \frac{\text{IPCA}_t}{\text{IPCA}_0} * \text{VN}_{db} * (1 + \text{Variação do mês de referência})^{\frac{du1}{du2}} \quad (2.7)$$

em que  $du1$  é o número de dias úteis entre a data de divulgação e o dia 15 do mês de referência e  $du2$  é o número de dias úteis entre o dia 15 do mês seguinte e o dia 15 do mês de referência. Caso a data do cálculo seja posterior ao décimo quinto dia do mês, o cálculo será conforme a seguinte equação:

$$\text{VNA} = \frac{\text{IPCA}_t}{\text{IPCA}_0} * \text{VN}_{db} * (1 + \text{IPCA}_{proj})^{\frac{du1}{du2}} \quad (2.8)$$

em que  $\text{IPCA}_{proj}$  é a projeção do IPCA apurada junto ao comitê de acompanhamento macroeconômico da ANDIMA (Associação Nacional das Instituições do Mercado Financeiro). Percebe-se pelas fórmulas acima que o IPCA tem grande impacto para os cálculos do preço da NTN-B.

Muitos artigos abordam o tema de ativos de renda fixa e mais afundo de títulos públicos federais, pois pesquisas nessa área são muito importantes para gestores e administradores de entidades financeiras e investidores nacionais e internacionais. Caldeira et al. (2012) utilizou a técnica introduzida por Markowitz (1952), inicialmente utilizada para otimização de carteiras compostas por ativos de renda variável, para maximizar ganhos em carteiras de ativos de renda fixa, para isso os autores utilizaram os títulos públicos NTN-F, LTN, NTN-B, NTN-C e LFT, os mesmos obtiveram resultados satisfatórios.

Graça et al. (2015) também fazem uma análise do contexto de títulos públicos, poupança, certificados de depósito bancário, entre outros para averiguar quais desses fundos são mais rentáveis para os pequenos investidores. Os autores concluíram que o fundo com maior retorno foi o NTN-B, quando comparado aos outros fundos de renda fixa.

Franklin Junior et al. (2012) desenvolvem outro ponto relevante no estudo de títulos públicos e outros ativos de renda fixa. O mercado segurador brasileiro, principalmente entidades

aberta e fechada de previdência complementar (EAPC e EFPC) que têm que garantir seus compromissos futuros com investimentos a longo prazo trazidos a valor presente (MARQUES et al., 2012). Sabendo que um dos títulos públicos mais importantes é o NTN-B, o estudo do IPCA se torna imprescindível por ser o índice utilizado como indexador desse tipo de ativo.

O IPC de alguns países podem ser parecidos por vários fatores econômicos e financeiros, como é o caso do BRICS. Ciani (2015) afirma que o termo BRIC foi criado pelo economista O'Neil, chefe de pesquisa em economia global do grupo financeiro Goldman Sachs, em 2001 para se referir aos quatro países emergentes: Brasil, Rússia, Índia e China. Em 2011 a África do Sul passou a fazer parte do grupo e a sigla teve o acréscimo da letra "S" (*South Africa*) resultando em BRICS. Esses países têm o maior PIB entre os países emergentes, conforme Ciani (2015), e, apesar de grandes diferenças culturais e políticas, são países com grande dimensão territorial e com grande população, além disso, têm o Índice de Desenvolvimento Humano (IDH) similares e sofrem de alguns problemas sociais semelhantes, como desigualdade social e mortalidade infantil (IBGE, 2014).

De acordo com o Ministério das Relações Exteriores (2016), em 2007, em reunião informal, verificou-se interesse dos países do BRIC (ainda sem a África do Sul) em aprofundar a construção de ideias à margem da Assembleia Geral da ONU. Em 2008, foi feita a primeira reunião formal em Ecatimburgo, na Rússia, onde foi criada uma entidade político-diplomática. Desde então os Chefes de Estado e de Governo do BRICS se encontram anualmente em reuniões de Cúpulas. As atividades intra-BRICS já abrangem cerca de 30 áreas, como: agricultura, ciência e tecnologia, cultura, espaço exterior, governança e segurança da internet, previdência social, propriedade intelectual, saúde, turismo, entre outras. Porém a área econômico-financeira do grupo continua sendo o tema mais discutido, na VI Cúpula do BRICS foram assinados dois acordos: o do Arranjo Contingente de Reservas (ACR), voltado para ajuda mútua dos países do BRICS quando houver flutuação no balanço de pagamentos, e o do Novo Banco de Desenvolvimento (NBD), voltado para o financiamento de projetos e estruturas sustentáveis em países emergentes, os dois totalizaram um investimento de 200 bilhões de dólares.

Esse segundo acordo seria de grande importância, tanto para o BRICS como para outros países emergentes, pois Pao e Tse (2010) mostraram que juntos os países do BRICS são responsáveis por quase 30% da emissão de  $CO_2$  do planeta. Os mesmos autores tentam encontrar determinantes para essa emissão, outros trabalhos também analisam a sustentabilidade desses países como o de Gomes et al (2015) que tenta encontrar relações entre as características

econômico-financeiras das empresas e seu enfoque na sustentabilidade nos países do BRICS.

É vista uma conjuntura econômica mundial favorável aos países do BRICS, O'Neil (2001) argumenta que o BRICS será mais forte economicamente que o grupo G6 (Estados Unidos, Japão, Alemanha, França, Itália e Reino Unido) nos próximos 30 a 40 anos. Outros trabalhos contextualizaram conclusões parecidas, como: Cooper e Antkeiwicz (2008) e Mahbubani (2008).

Como foi vista nessa seção, a importância de criar modelos com intuito de prever a inflação interessa, não interessa só administradores e gestores públicos, como também investidores, de todas as escalas, e gestores financeiros de fundos de pensão e seguradoras. Porém, em alguns casos, esses agentes econômicos estão de fato interessados na previsão da inflação máxima dentro de um período de tempo. Uma das maneiras de se modelar o IPCA máximo é utilizando os métodos encontrados na Teoria de Valores Extremos.

## 2.2 ANÁLISE DE EVENTOS EXTREMOS

A Teoria dos Valores Extremos (TVE) é um dos importantes ramos da Estatística, pois possibilita a análise de eventos extremos que apresentam baixa frequência relativa, ou seja, são eventos que raramente acontecem, mas que podem resultar em catástrofes ambientais, sociais e econômicas. O estudo desses eventos é de grande importância em áreas como engenharia, economia, climatologia, atuária e administração. Existem muitos trabalhos publicados que utilizam modelagem de valores extremos, como modelagem das ondas do oceano (DAWSON, 2000); falha na memória celular (MCNULTY et al., 2000); engenharia eólica (HARRIS, 2001); estratégias administrativas (DAHAN; MENDELSON, 2001); processamentos de dados biológicos (ROBERTS, 2000); precipitações máximas (BEIJO et al., 2003). Há também várias aplicações em Ciências Atuariais para cálculo de riscos que resultam em grandes perdas econômicas para as seguradoras, resseguradoras, instituições financeiras ou entidades de previdências entre elas: risco financeiro nas bolsas de valores (ARRAES; ROCHA, 2006) e perda máxima operacional (ERGASHEV et al. 2013).

De acordo com Embrechts et al. (1997), em 1992 um furacão nos Estados Unidos chamado "Andrew" causou uma perda de 16 bilhões de dólares às seguradoras. Dois anos depois, em 1994, um terremoto no nordeste da Califórnia resultou em um prejuízo de mais de

11 bilhões de dólares. Esses dois eventos foram, economicamente, os mais prejudiciais, entre 1970 e 1995, no mundo. Em um âmbito social, no mesmo período, dois eventos marcaram o mundo: um furacão em Bangladesh em 1970 onde aproximadamente 300 mil pessoas foram a óbito; e um terremoto na cidade de Tangshan na China que resultou em 250 mil vítimas fatais em 1976.

Entende-se que os eventos extremos estão associados aos valores extremos (máximos e mínimos) das distribuições, ou seja, são valores que se encontram longe do aglomerado de dados. De acordo com Coles (2001) quando tem-se uma amostra aleatória  $X_1, X_2, \dots, X_n$  independente e identicamente distribuída (iid) de uma distribuição  $F$  qualquer, existem duas formas distintas de se encontrar uma série de máximos (uma distribuição extrema máxima), extraindo os valores das variáveis aleatórias que ultrapassem um limiar  $u$  suficientemente alto. Essa técnica é comumente chamada de POT (*Picks Over Threshold*), como mostra a Figura 2.

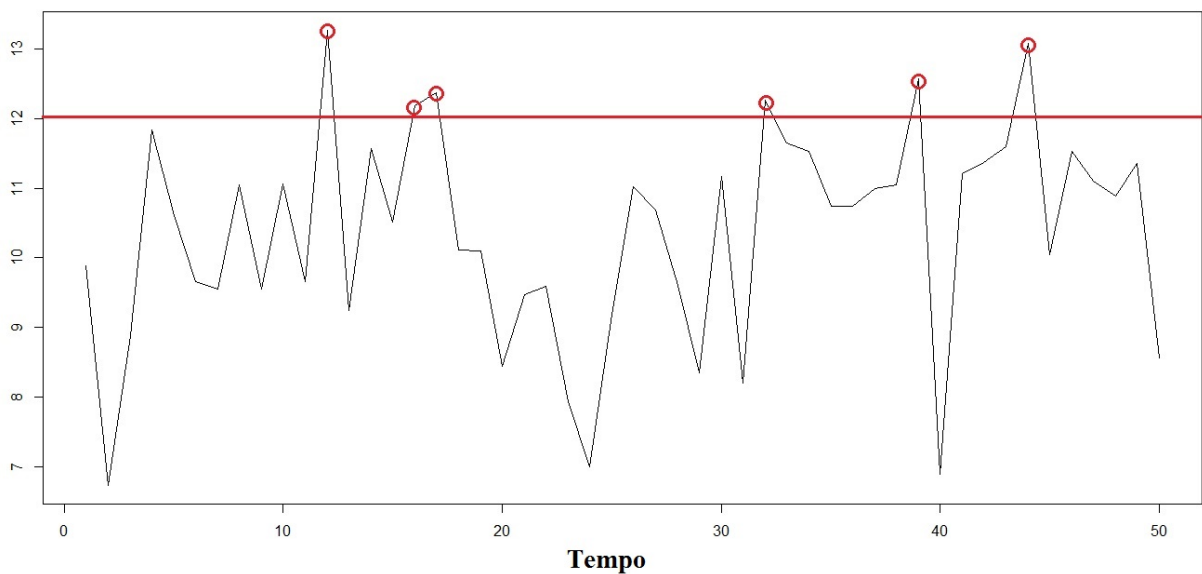


Figura 2 – Construção de uma série de máximos por POT.

Fonte: Elaborado pelo autor.

Essa série de máximo se aproximará da distribuição generalizada de Pareto (GPD - *Generalized Pareto Distribution*). Porém o trabalho focará em uma outra técnica de construção de série de máximos, o método de blocos de máximos, que consiste em extrair o valor máximo de cada uma das variáveis aleatórias. Um exemplo pode ser visto na Figura 3.



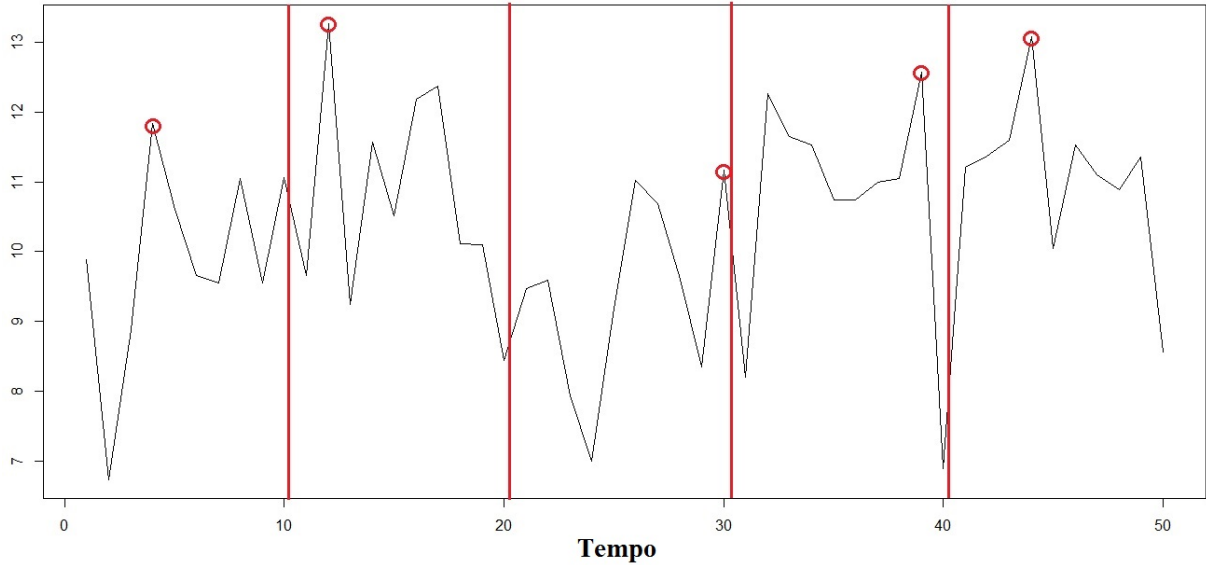


Figura 3 – Construção de uma série de máximos por Blocos de Máximos.  
Fonte: Elaborado pelo autor.

Dessa forma, ao extrair o máximo de cada uma das variáveis aleatórias, cria-se uma nova série de dados  $Z$  dessa maneira:

$$Z = \max\{X_1, X_2, \dots, X_n\} \quad (2.9)$$

Como a amostra é independente e indenticamente distribuída (iid), é possível encontrar a função distribuição dos máximos da seguinte forma:

$$\begin{aligned} Pr\{Z \leq z\} &= Pr\{X_1 \leq z, X_2 \leq z, \dots, X_n \leq z\} \\ &= Pr\{X_1 \leq z\}Pr\{X_2 \leq z\} \dots Pr\{X_n \leq z\} \\ &= \{F(z)\}^n \end{aligned} \quad (2.10)$$

A princípio, o resultado obtido na equação (2.10) não é imediatamente útil, pois a distribuição  $F(\cdot)$  continua desconhecida. Além disso trata-se de uma função distribuição degenerada, ou seja, assumindo apenas os valores 0 ou 1. Essa dificuldade é superada quando normalizamos a variável  $Z$ :

$$\frac{Z - b}{a}. \quad (2.11)$$

de forma que  $a > 0$ . Com isso  $b$  e  $a$  são os parâmetros de posição e escala de  $Z$ . Dessa forma, Fisher e Tippett (1928) e posteriormente Gnedenko (1943), mostraram que:

$$Pr\{(M_n - b)/a \leq z\} \xrightarrow{d} G(z), \quad \text{para } n \rightarrow \infty. \quad (2.12)$$

Se  $G$  não convergir para uma função distribuição degenerada,  $G$  convergerà para um dos três tipos de função distribuição:

$$I : G(z) = \exp \left\{ -\exp \left[ - \left( \frac{z - b}{a} \right) \right] \right\}, \quad -\infty < z < \infty, \quad (2.13)$$

$$II : G(z) = \exp \left\{ - \left( \frac{z - b}{a} \right)^{-\alpha} \right\}, \quad z > b, \quad (2.14)$$

$$III : G(z) = \exp \left\{ - \left[ - \left( \frac{z - b}{a} \right)^\alpha \right] \right\}, \quad z < b. \quad (2.15)$$

em que  $a > 0$ ,  $-\infty < b < \infty$  e  $\alpha > 0$ , sendo os parâmetros de escala, posição e forma. Essas funções distribuições correspondem às distribuições de valores extremos Gumbel (2.13), Weibull (2.14) e Fréchet (2.15).

Diferentes aplicações podem ser feitas com essas distribuições na análise de valores extremos, definidas pelos seus comportamentos diferenciados na direção dos extremos. Entretanto, uma das dificuldades dessa diferenciação é escolher qual das três distribuições é mais adequada para determinado conjunto de dados (COLES,2001). Para isso, Jenkinson (1955) estruturou uma composição das três distribuições de valores extremos, a distribuição generalizada de valores extremos, ou GEV (*Generalized Extreme Values*). A distribuição GEV tem a função distribuição descrita da seguinte forma:

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2.16)$$

em que  $-\infty < z < \mu - \sigma/\xi$  para  $\xi < 0$ ,  $-\infty < z < \infty$  para  $\xi \rightarrow 0$ ,  $\mu - \sigma/\xi < z < \infty$  para  $\xi > 0$  e  $\sigma > 0$ , sendo  $\mu$ ,  $\sigma$  e  $\xi$  os parâmetros de posição, escala e forma, respectivamente. Derivando-se a equação (2.16) em relação a  $z$ , obtêm-se a função densidade de probabilidade (fdp) da distribuição GEV, dada pela equação:

$$g(z) = \frac{1}{\sigma} \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-\left(\frac{1 + \xi}{\xi}\right)} \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right] \right\}. \quad (2.17)$$

As funções distribuições de valores extremos de Fréchet e Weibull correspondem aos casos particulares que  $\xi > 0$  e  $\xi < 0$ , respectivamente. Quando é feito o limite de (2.16) com  $\xi$  tendendo a zero, a mesma tende para a função distribuição da Gumbel. Dessa forma, é visto que  $\xi$  indica para qual das três formas das distribuições a GEV tende.

Sabe-se que a variável extrema máxima é construída utilizando bloco de máximos, ou seja, a mesma não leva consigo uma possível dependência que a variável aleatória, da qual foi originada, poderia ter. Porém, Finkenstädt e Rootzén (2001) mostram que pode haver tendência na série de máximo mesmo sendo independente, dessa forma utiliza-se o modelo GEV com tendência, como pode ser visto:

$$G(z) = \exp \left\{ - \left[ 1 + \xi \left( \frac{z - \mu(t)}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad (2.18)$$

em que  $\mu(t)$  é o parâmetro de posição que, ao contrário da distribuição GEV convencional, é uma função indexada no tempo  $t$ . Essa função no tempo pode assumir funções lineares de primeiro grau, por exemplo  $\mu_0 + \mu_1 t$  ou até mesmos funções quadráticas:  $\mu_0 + \mu_1 t + \mu_2 t^2$ . Trabalhos como os de Coles e Tawn (1990) e Dixon e Tawn (1992) utilizaram a distribuição GEV com tendência, ambos aplicaram em dados de níveis máximos do mar.

Nadarajah e Choi (2007) sugerem sempre a utilização da GEV, pois ela possui toda a flexibilidade contida em seus casos particulares, entretanto, algumas vezes na modelagem de valores extremos é visto que a utilização da distribuição Gumbel é preferível à distribuição GEV, pelo fato do parâmetro de forma não ser significativo para explicar o conjunto de dados extremos. Bozdogan (1987) explica que, em geral, busca-se um modelo mais parcimonioso, ou seja, um modelo menos complexo, com menos parâmetros e que explique bem o comportamento da variável, no âmbito da análise de valores extremos. A Gumbel é uma distribuição mais parcimoniosa que a GEV. Esse tipo de preferência também se refere aos modelos com e sem tendência. Para analisar se o modelo com menos parâmetros não perde na explicação da variável estudada existem muitos métodos, como: critério de Akaike (AIC), critério bayesiano de Schwarz (BIC), análise de *deviance*, entre outros.

Existem muitos métodos de estimação dos parâmetros da GEV, como: método da máxima verossimilhança, método dos momentos, método de estimação de Hill, método de Kernel, método PMW, entre outros (GOMES; GUILLOU, 2015). Para Smith (1985), o método da máxima verossimilhança é um método particularmente eficiente pela sua capacidade de adaptação à distribuições com funções complexas. A função de verossimilhança da distribuição GEV pode

ser vista a seguir:

$$l(\mu, \sigma, \xi | z_1, \dots, z_n) = -n \log \sigma - \left( \frac{1 + \xi}{\xi} \right) \sum_{i=1}^n \log \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] \quad (2.19)$$

Entretanto, uma das dificuldades, no caso da GEV, desse método está nas condições de regularidade, quando  $-1 \leq \xi \leq -0,5$  os estimadores de Máxima Verossimilhança não satisfazem as propriedades assintóticas e quando  $\xi < -1$  os estimadores não existem (SMITH, 1985). Outra dificuldade está no âmbito matemático, isto é, a sensibilidade numérica da função verossimilhança: ou seja, pequenas mudanças nos valores dos dados algumas vezes afetam muito as estimativas, além disso, para encontrar os estimadores por Máxima Verossimilhança é preciso utilizar métodos numéricos, pois os mesmos não são encontrados analiticamente. Coles e Dixon (1999), argumentam contrariamente ao método da Máxima Verossimilhança pelo fato de suas propriedades não serem satisfeitas para pequenas amostras. Uma alternativa ao método da Máxima Verossimilhança seria a Inferência Bayesiana (BEIJO et al., 2009).

### 2.3 INFERÊNCIA BAYESIANA

Para Coles (2001), alguns motivos levaram pesquisadores a utilizar Inferência Bayesiana na análise de valores extremos, talvez o principal deles seja a escassez de dados extremos, gerando a necessidade de incluir novas fontes de informação para o modelo. Paulino et al. (2003) apresenta outro motivo pelo qual o uso de técnicas Bayesianas é importante, a distribuição *a posteriori*, que muitas vezes fornece estimativas mais completas, por justamente conseguir incorporar informação *a priori*, ou seja, informações passadas, juntamente com a informação do conjunto de dados.

Para Box e Tiao (1973) a distribuição *a priori* ajuda a diminuir as incertezas dos parâmetros estudados. Quando a informação *a priori* não é palpável ou não é significativa à amostra, existem alternativas que visam encontrar técnicas para construção das distribuições *a priori* (KASS; WASSERMAN, 1996). Esses métodos conduzem a distribuições *a priori* não informativas.

Porém, a elicitación de informação *a priori* pode ser feita a partir do conhecimento de

especialistas da área, sobre o fenômeno estudado. Para West e Crosse (1992), qualquer forma de informação fornecida pelo especialista deve ser tratada como um conjunto de dados. Assim, pode-se utilizar como informação a *priori* estudos anteriores e conjuntos de dados existentes.

Choy et al. (2009) apresentam duas formas de elicitaco de informao para construo de distribuo a *priori*, com base em estudos anteriores e de conhecimentos de especialistas na rea estudada. Os autores mostram que as duas formas so eficientes. Alguns trabalhos como os de Anholt et al. (2000), Fleischman e Burwen (2003) e Wintle et al. (2003), utilizaram informao a *priori*, baseados em estudos anteriores e no em pareceres de especialistas.

Aps a construo da distribuo a *priori*  calculada a distribuo a *posteriori*. Sabe-se que a *posteriori* dos parmetros  proporcional  multiplicaco da distribuo a *priori* dos parmetros e a funo de verossimilhana do modelo, como  visto na equao a seguir:

$$p(\theta|z_1, \dots, z_n) \propto p(\theta)L(\theta|z_1, \dots, z_n), \quad (2.20)$$

em que  $\theta$  so os parmetros da distribuo ou modelo,  $z_1, \dots, z_n$  os valores observados da amostra,  $p(\theta|z_1, \dots, z_n)$  a distribuo a *posteriori*,  $p(\theta)$  a distribuo a *priori* de  $\theta$  e  $L(\theta|z_1, \dots, z_n)$  a funo de verossimilhana.

A distribuo a *posteriori* de um parmetro contm toda informao probabilística do mesmo, assim se integra a *posteriori* conjunta em relao aos outros parmetros para obteno da distribuo a *posteriori* marginal para cada parmetro. Dessa forma,  possvel calcular as estatísticas para cada parmetro, como mdia, mediana e varincia.

Alm das estimativas pontuais, podem ser calculadas as estimativas intervalares, obtidas de uma regio do espao paramtrico que contenha maior parte da massa probabilística. Essas estimativas intervalares so chamadas de intervalos de credibilidade. De acordo com Paulino et al. (2003), dentre os infinitos intervalos de credibilidade possveis, aquele com densidade a *posteriori* mxima  chamado de intervalo HPD (*Highest Posterior Density*).

Em situaes em que no  possvel se obter a distribuo marginal do parmetro de forma analítica, como  o caso das distribues marginais dos parmetros da GEV, utilizam-se tcnicas de simulaes probabilísticas, ou seja, simulaes baseadas em distribues de probabilidade. Uma dessas tcnicas  a de simulao Monte Carlo via Cadeias de Markov (MCMC), que so possvel com a evoluo computacional. Esse tipo de simulao gera variveis aleatrias cujas densidades so desconhecidas. Dessa maneira, h um passeio aleatrio no espao paramtrico estudado. Assim, converge-se para uma distribuo estacionria do

parâmetro, que nesse caso é a distribuição a *posteriori* (NOGUEIRA et al., 2004). Gamerman (1997) explica que a metodologia MCMC visa construir uma densidade para cada parâmetro.

Como as cadeias de Markov geradas devem convergir para uma distribuição estacionária, existem alguns critérios para verificação de não convergência delas. Um desses critérios é o de Raftery-Lewis, proposto por Raftery e Lewis (1992), que estima o número de iterações necessárias para a convergência à distribuição estacionária, dessa forma, quando o cálculo do critério resultar no valor 1, significa que a cadeia convergiu. Geweke (1992) também estruturou um critério para analisar independência da cadeia, ou seja, quando o valor  $p$  desse critério estiver dentro da área de aceitação, a cadeia é independente e, conseqüentemente, não há evidências para afirmar que houve falta de convergência. Com esses testes analisa-se se há ou não ausência de convergência.

### 2.3.1 Aplicações na Análise de Eventos Extremos

No âmbito da análise valores extremos, Parent e Bernier (2003) utilizam informação de especialistas da área de hidrologia para construção de uma distribuição a *priori* informativa, porém, os mesmos também fizeram uso de uma distribuição a *priori* não informativa, com intuito de modelar o risco de inundações extremas no rio Garonne (abrange Espanha e França). Os intervalos de credibilidade foram mais precisos com a distribuição a *priori* informativa.

Behrens et al. (2004) analisaram o uso da Inferência Bayesiana para a GPD, por meio de simulações e fazendo uma aplicação utilizando conjunto de dados da Nasdaq 100, os autores obtiveram informação a *priori* perguntando a peritos qual o valor da mediana e do quantil de 90% para as estimativas dos parâmetros. Os autores analisaram que as predições dos níveis máximos da Nasdaq foram melhores utilizando a metodologia Bayesiana, em comparação à metodologia frequentista.

Coles e Powell (1996) utilizaram como informação a *priori* os níveis máximos da velocidade dos ventos dos Estados Unidos, para modelar os níveis máximos da velocidade dos ventos de San Diego, Portland e Minnesota. Os autores também se propuseram a formular uma distribuição a *priori*, a fim de melhorar as estimativas dos parâmetros e suas predições, assim, estipularam uma distribuição a *priori* para os três parâmetros da GEV, em que os mesmos seguem uma distribuição Normal Trivariada, como pode-se ver a seguir:

$$\theta_1 \sim \text{MVN}(\Phi_0, \Sigma_0), \quad (2.21)$$

em que MVN é a sigla para Normal Multivariada (*Multivariate Normal*),  $\theta_1 = (\mu, \log(\sigma), \xi)$  é vetor dos parâmetros,  $\Phi_0$  é o vetor de médias dos hiperparâmetros e  $\Sigma_0$  a matriz de variâncias e covariâncias dos hiperparâmetros. Utiliza-se o logaritmo natural ( $\log$ ) do parâmetro de escala pelo fato do mesmo seguir uma distribuição assimétrica. Dessa forma, assume-se que o parâmetro de escala segue uma distribuição Lognormal, ou seja, o logaritmo do mesmo segue uma distribuição Normal (BEAULIEU et al., 1995).

Stephenson e Ribatet (2006) mostram que é possível utilizar a distribuição Normal Trivariada em que, além de calcular o logaritmo natural do parâmetro de escala, utiliza-se o logaritmo natural do parâmetro de posição, pois assume-se que  $\mu$  também segue uma distribuição assimétrica. A seguir é vista a distribuição *a priori* dessa Normal Trivariada.

$$\theta_2 \sim \text{MVN}(\Phi_0, \Sigma_0), \quad (2.22)$$

em que  $\theta_2 = (\log(\mu), \log(\sigma), \xi)$  é o vetor dos parâmetros,  $\Phi_0$  é o vetor de médias e  $\Sigma_0$  a matriz de variâncias e covariâncias.

## 2.4 TESTES AUXILIARES NA ANÁLISE DE EVENTOS EXTREMOS

Para análise de valores extremos alguns pressupostos devem ser verificados. A série de máximos deve ser independente, para que o arcabouço teórico, que envolve a modelagem utilizando a distribuição GEV, seja respeitado. Além disso, deve-se verificar se a série é estacionária ou não, se não, o modelo GEV com tendência deverá ser utilizado. Para isso, foi utilizado o teste de Ljung-Box e *Runs-Test*, respectivamente.

Após a modelagem, o teste de Kolmogorov-Smirnov é utilizado para analisar a aderência das distribuições à séries de máximos. Além disso, também é usado o teste de Razão de Verossimilhança para investigar se existem diferenças estatísticas entre duas distribuições ajustadas, como a GEV e Gumbel.

Os conceitos e a metodologia de todos esses testes podem ser vistos nesta seção.

### 2.4.1 Teste de Ljung-Box

A princípio Box e Pierce (1970) propuseram o teste Box-Pierce com intuito de analisar os resíduos de modelos de séries temporais, pois os mesmos deviam ser independentes, dessa forma o teste tinha como hipótese nula que os resíduos eram independentes. Porém para esse teste, os resíduos  $(a_i)$  devem seguir uma distribuição Normal com média 0 e variância constante  $(a_i \sim N(0, \sigma^2))$ .

Para generalizar o teste quando não se observava normalidade nesses resíduos, Ljung e Box (1978) sugeriram um novo teste, em que a hipótese nula também é a independência dos resíduos. Para calcular a estatística do teste Ljung-Box primeiramente são calculadas as autocorrelações dos resíduos, conforme a equação:

$$r_k = \sum_{i=k+1}^n \frac{a_i a_{i-k}}{a_i^2}, \quad (2.23)$$

em que  $a_i$  é o  $i$ -ésimo resíduo do modelo e  $r_k$  a  $k$ -ésima autocorrelação.

Anderson (1942) mostra que:  $E(r_k) = 0$ ,  $Var(r_k) = (n-k)/[n(n+2)]$  e  $Cov(r_k, r_l) = 0$  (para  $k \neq l$ ). Além disso  $r_k$  segue uma distribuição Normal, com isso calcula-se a estatística do teste de acordo com a equação:

$$Q(r) = n(n-2) \sum_{i=k+1}^n \frac{r_k^2}{(n-k)} \quad (2.24)$$

em que  $m$  é o número de iterações tomada na função de autocorrelação. Dessa forma os autores concluem que  $Q(r)$  segue uma distribuição Qui-Quadrado com  $(m-p-q)$  graus de liberdade onde  $p$  e  $q$  são as ordens do modelo de séries temporais ajustado. Caso  $p$  e  $q$  sejam nulos, a distribuição será uma Qui-Quadrado com  $m$  graus de liberdade. O teste rejeita a hipótese nula quando  $Q(r) > \chi_{1-\alpha, m-p-q}^2$  com um nível  $\alpha$  de significância.

Pesquisas feitas na área de Teoria dos Valores Extremos normalmente usam séries de dados temporais, dessa forma muitos trabalhos utilizam o teste Ljung-Box para averiguar se a série é independente. Entre alguns desses trabalhos está Shahzadi (2013), que utiliza série de chuvas máximas anuais no Paquistão, Kliester et al. (2006) em seu trabalho sobre previsão de VaR, Dahlen et al. (2015) em que tenta estimar o valor máximo de risco de gestão, Barbosa et al. (2014) trabalha com dados extremos de temperatura e Reis et al. (2012) também utiliza o teste de Ljung-Box em sua pesquisa sobre umidade relativa do ar mínima.



### 2.4.2 *Runs-Test*

Um teste não-paramétrico para averiguar tendência na série de dados é o *Runs-Test* ou teste de Wald-Wolfowitz (WALD; WOLFOWITZ, 1948). A hipótese nula consiste na ideia de não existir tendência na série. Sua estatística é calculada considerando  $M$  a mediana de uma série  $X_1, X_2, \dots, X_n$  no tempo, para cada  $X_i$  se atribui o símbolo “a” se ele for maior ou igual a  $M$  e “b” se ele for menor que  $M$ , então tem-se  $n = n_a + n_b$ , sendo assim  $T$  é a estatística do teste, o número total de grupos com símbolos iguais. Rejeita-se a hipótese nula caso  $T$  seja relativamente pequeno a um nível de significância  $\alpha$ .

### 2.4.3 Teste de Kolmogorov-Smirnov

O teste não-paramétrico de Kolmogorov-Smirnov, proposto por Kolmogorov (1933) e Smirnov (1939), tem como intuito analisar se, para um nível de confiança, uma função distribuição acumulada assumida se adere a função de distribuição acumulada empírica de um conjunto de dados. Dessa maneira a hipótese nula do teste é a de que os dados seguem a distribuição estipulada na pesquisa. A estatística do teste é calculada conforme a equação:

$$D_n = \sup_x |F(x) - F_n(x)|, \quad (2.25)$$

em que  $D_n$  é a distância absoluta máxima vertical entre os gráficos de  $F(x)$ , que corresponde a função distribuição acumulada assumida, e  $F_n(x)$ , a função distribuição acumulada empírica para os dados. Se essa  $D_n$  exceder o quantil para  $1 - \alpha$  dado através da tabela de quantis para o teste de Kolmogorov-Smirnov, então rejeita-se a hipótese nula com  $\alpha$  nível de significância.

Existem alguns trabalhos como de Razali e Wah (2011), Saculinggan e Balase (2013) e Torman et al. (2012) afirmam que existem testes mais poderosos e que cometem menos erros que o Kolmogorov-Smirnov para testar normalidade. Porém, outros trabalhos: Chicheportiche e Bouchaud (2012) e Crutcher (1975) mostram que o teste é eficaz para valores extremos. Além disso, vários trabalhos na área de análise de valores extremos utiliza o teste para averiguar aderência, como Grimm et al. (2016) e Coumou et al. (2014) que trabalham com dados de precipitação e vento respectivamente. Também existe todo um arcabouço teórico já apresentado

no trabalho que afirma que ao extrair-se os máximos das variáveis aleatórias essa nova variável extrema máxima seguirá uma das três distribuições de valores extremos.

#### 2.4.4 Teste de Razão de Verossimilhança

Para McCullagh e Nelder (1989), quando se ajusta um modelo, a ideia principal é substituir um conjunto de dados observados por um conjunto de valores estimados utilizando para isso o menor número de parâmetros possíveis. Pois quanto menos parâmetros, menos complexo é o modelo, e de interpretação fácil. O que ocorre é que muitas vezes um modelo mais simples acaba por não explicar de forma adequada os dados, necessitando, dessa maneira, da inclusão de novos parâmetros.

A fim de determinar se a adição de um parâmetro é significativa ou não, para um modelo, Nelder e Wedderburn (1972) propuseram um cálculo para determinar qual modelo tem menor discrepância na sua função de verossimilhança. Casella e Berger (2010) mostram que a diferença entre as discrepâncias de dois modelos é igual à razão de verossimilhança, e essa diferença segue uma distribuição Qui-Quadrado com um grau de liberdade, como é visto em:

$$-2\log\left(\frac{L_s(\Theta|z_i)}{L_c(\Theta|z_i)}\right) = -2\log(L_s(\Theta|z_i)) + 2\log(L_c(\Theta|z_i)) \quad (2.26)$$

em que  $L_s$  é a função de verossimilhança do modelo sem o parâmetro estudado e  $L_c$  é a função de verossimilhança do modelo com o parâmetro.

Sabendo que essa diferença segue assintoticamente uma distribuição Qui-Quadrado com um grau de liberdade, é possível averiguar se a diferença entre as funções de verossimilhança dos modelos é significativa, ou seja, se o parâmetro incluído em um dos modelos é significativo. Esse teste é chamado de Teste de Razão de Verossimilhança (TRV), e tem como hipótese nula que o parâmetro é igual a zero e hipótese alternativa que o parâmetro é diferente de zero. Dessa maneira, o valor  $p$  do teste é calculado como a probabilidade da distribuição Qui-Quadrado, com um grau de liberdade, assumir o valor encontrado na diferença das funções de verossimilhança dos modelos estudados. Se essa probabilidade for menor que um  $\alpha$  de significância estipulado, rejeita-se a hipótese nula e conclui-se que o parâmetro é significativo.

### 3 MATERIAL E MÉTODOS

Nesta seção serão apresentados os dados utilizados no trabalho, bem como suas coletas e organizações. Também são vistas as metodologias para modelagem: como foram encontrados os estimadores de Máxima Verossimilhança; a forma de elicitación das informações e construção das distribuições *a priori*; a forma de avaliação dos modelos.

#### 3.1 DETERMINAÇÃO DO TAMANHO DOS BLOCOS

Seria interessante para gestores, investidores, administradores públicos e privados a análise e previsão da inflação máxima para períodos de tempo como um ano ou semestre. Porém um intervalo de tempo interessante a ser analisado seria o máximo quadrimestral, pois de acordo com o 4º parágrafo do 9º artigo da Lei complementar nº 101 (Lei de Responsabilidade Fiscal), de 4 de maio de 2000, o Poder Executivo deverá demonstrar e avaliar o cumprimento das metas fiscais quadrimestrais, em audiência fiscal, até o final dos meses de maio, setembro e fevereiro, na Comissão Mista de Planos, Orçamentos Públicos e Fiscalização do Congresso Nacional, ou equivalente nas Casas Legislativas estaduais e municipais. Um dos parâmetros macroeconômicos analisados nessas metas é o IPCA. Outros tempos de retorno como semestral e anual também poderiam ser analisados, porém deixaria-se de utilizar mais dados, pois o bloco de máximos seria maior, resultando em um tamanho de série temporal de máximos menor.

#### 3.2 DADOS

Os dados do IPCA foram coletados no site do IPEA-DATA (2016) e fazem referência ao período de Janeiro de 2003 a Abril de 2016. As séries de IPC mensal da Rússia, China e Índia, utilizadas como informação *a priori* na Inferência Bayesiana, foram extraídas do site da OECD (2016) (*Organisation for Economic Co-operation and Development*) e são referentes ao mesmo período utilizado para o IPCA (Janeiro de 2003 a Abril de 2016).

Após extrair os máximos de inflação por quadrimestre, formaram-se as séries de máximos, referentes ao primeiro quadrimestre de 2003 ao primeiro quadrimestre de 2016. As séries

completas de máximos foram divididas em duas:

- a) séries para ajuste: primeiro quadrimestre de 2003 ao último quadrimestre de 2012;
- b) séries para verificação: primeiro quadrimestre de 2013 ao primeiro quadrimestre de 2016.

As séries para ajuste foram utilizadas para calcular as estimativas de níveis máximos de IPCA, para os tempos de retorno de 2, 3, 4, 5 e 10 quadrimestres. Baseando-se nas séries para verificação foi possível calcular a acurácia e precisão das estimativas de inflação máxima.

### 3.3 MODELAGEM

Aqui são expostas as metodologias utilizados para modelagem dos níveis máximos quadrimestrais de IPCA e a forma de avaliação dos modelos.

#### 3.3.1 Análises Preliminares

Foi utilizado o teste de Ljung-Box (seção 2.4.1) para averiguar se há dependência na série de níveis máximos de IPCA. O mesmo também foi aplicado nas séries usadas como informação *a priori* na Inferência Bayesiana.

As distribuições GEV e Gumbel foram usadas para modelagem dos níveis máximos quadrimestrais de IPCA. Caso houvesse ausência de aleatoriedade na série, seriam utilizados os modelos GEV com tendência e Gumbel com tendência. A análise de tendência foi verificada por meio do *Runs-Test* (seção 2.4.2).

O teste de Komolgorov-Smirnov (seção 2.4.3) verificou a aderência das distribuições GEV e Gumbel, com os parâmetros estimados via Máxima Verossimilhança, à série de níveis máximos quadrimestrais de IPCA.

O Teste de Razão de Verossimilhança (seção 2.4.4) calculou a diferença entre as variabilidades das distribuições ajustadas e verificou se essa diferença era significativa. Dessa forma, avaliou a significância do parâmetro  $\xi$ , ou seja, se deveria ser usada a distribuição GEV (com  $\xi$ ) ou a distribuição Gumbel (sem  $\xi$ ).

### 3.3.2 Estimação via Máxima Verossimilhança

Os parâmetros das distribuições GEV e Gumbel foram estimados utilizando o método da Máxima Verossimilhança. Com base na função de Verossimilhança da GEV:

$$l(\mu, \sigma, \xi | z_1, \dots, z_n) = -n \log \sigma - \left( \frac{1 + \xi}{\xi} \right) \sum_{i=1}^n \log \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right], \quad (3.1)$$

derivou-se a mesma em relação a  $\mu$ ,  $\sigma$  e  $\xi$ . Igualando a zero, encontra-se o seguinte sistema:

$$\begin{cases} \frac{1}{\hat{\sigma}} \sum_{i=1}^n \left( \frac{1 + \hat{\xi} - w_i^{-\frac{1}{\hat{\xi}}}}{\hat{\sigma}} \right) = 0 \\ -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n \left\{ \frac{(z_i - \hat{\mu}) \left[ (1 + \hat{\xi}) - w_i^{-\frac{1}{\hat{\xi}}} \right]}{w_i} \right\} = 0 \\ \sum_{i=1}^n \left\{ \left( 1 - w_i^{-\frac{1}{\hat{\xi}}} \right) \left[ \frac{1}{\hat{\xi}^2} \ln(w_i) - \frac{(z_i - \hat{\mu})}{\hat{\xi} \hat{\sigma} w_i} \right] - \frac{(z_i - \hat{\mu})}{\hat{\sigma} w_i} \right\} = 0, \end{cases} \quad (3.2)$$

em que  $w_i = 1 + \hat{\xi} \left( \frac{z_i - \hat{\mu}}{\hat{\sigma}} \right)$  e  $\hat{\mu}$ ,  $\hat{\sigma}$  e  $\hat{\xi}$  são os estimadores de Máxima Verossimilhança dos parâmetros da GEV.

Visto que o sistema de equações (3.2) não tem solução analítica, utilizou-se o método quasi-Newton para obtenção da solução numérica do sistema.

### 3.3.3 Estimação via Inferência Bayesiana

Para estimação dos parâmetros via Inferência Bayesiana, utilizou-se, fundamentando-se em Coles e Powell (1996) e Stephenson e Ribatet (2006), a distribuição a priori Normal Trivariada:

$$p(\mu, \sigma, \xi) \propto |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\theta - \Phi_0)^T \Sigma_0^{-1} (\theta - \Phi_0) \right\}, \quad (3.3)$$

em que  $\theta = (\mu, \sigma, \xi)$  é o vetor de parâmetros da GEV,  $\Phi_0$  é o vetor de médias dos hiperparâmetros, em que  $\Phi_0 = (\mu_0, \sigma_0, \xi_0)$ ,  $\mu$ ,  $\sigma$  e  $\xi$  são os parâmetros da distribuição GEV e  $\Sigma_0$  a matriz de

variâncias e covariâncias, explicitada da seguinte forma:

$$\Sigma_0 = \begin{bmatrix} Var(\mu_0) & Cov(\mu_0, \sigma_0) & Cov(\mu_0, \xi_0) \\ Cov(\mu_0, \sigma_0) & Var(\sigma_0) & Cov(\sigma_0, \xi_0) \\ Cov(\mu_0, \xi_0) & Cov(\sigma_0, \xi_0) & Var(\xi_0) \end{bmatrix}. \quad (3.4)$$

Para elicitación da informação para construção das distribuições *a priori*, foram feitas pesquisas com intuito de encontrar economias equivalentes à do Brasil. Com base nessas pesquisas (seção 2.1), concluiu-se que os demais países pertencentes ao BRICS eram os que tinham tal equivalência, com exceção da África do Sul, que é o país que mais destoa desse grupo. Com isso fez-se uso do IPC máximo quadrimestral da Rússia, China e Índia como informação utilizada nas distribuições *a priori*.

A elicitación das distribuições *a priori* foi feita da seguinte forma:

- a) foi ajustada a distribuição GEV, pelo método da Máxima Verossimilhança, para os níveis máximos quadrimestrais de IPC dos três países;
- b) foi calculada a média do IPC máximo quadrimestral dos três países (MRCI);
- c) foi ajustada a distribuição GEV para os níveis máximos quadrimestrais de IPC da média (MRCI);
- d) dessa forma foram obtidas as estimativas, variâncias e covariâncias dos parâmetros da GEV, conforme a Tabela 2 e a Tabela 3, para cada conjunto de dados.

Tabela 2 – Estimativas dos parâmetros da distribuição GEV calculadas via método de Máxima Verossimilhança, para os dados de níveis máximos de IPC da Rússia, China, Índia e a média dos três países (MRCI).

Parâmetro Estimado	Rússia	China	Índia	MRCI
$\hat{\mu}_{EM}$	0,897	0,698	1,040	1,032
$\hat{\sigma}_{EM}$	0,461	0,427	0,461	0,208
$\hat{\xi}_{EM}$	0,198	-0,075	0,118	0,138

Fonte: Elaborado pelo autor.

As estimativas apresentadas na Tabela 2 foram utilizadas no vetor de médias ( $\Phi_0$ ), que é o hiperparâmetro, tanto da distribuição *a priori* Normal de Coles e Powell (1996), chamada no trabalho de Normal I, como para a distribuição *a priori* Normal de Stephenson e Ribatet (2006), nomeada de Normal II. Para a Normal I  $\mu_0 = \hat{\mu}_{EM}$ ,  $\sigma_0 = \log(\hat{\sigma}_{EM})$  e  $\xi_0 = \hat{\xi}_{EM}$ , já para Normal II  $\mu_0 = \log(\hat{\mu}_{EM})$ ,  $\sigma_0 = \log(\hat{\sigma}_{EM})$  e  $\xi_0 = \hat{\xi}_{EM}$ , em que  $\hat{\mu}_{EM}$  é a estimativa de máxima verossimilhança (*EM*) do parâmetro  $\mu$ ,  $\hat{\sigma}_{EM}$  é a estimativa do parâmetro  $\sigma$  e  $\hat{\xi}_{EM}$  é a estimativa do parâmetro  $\xi$ . Ou seja, em ambas as distribuições utiliza-se o logaritmo natural

da estimativa de  $\sigma$ , e na distribuição Normal II utiliza-se o logaritmo natural da estimativa de  $\mu$  também. Após isso foram calculadas as variâncias e covariâncias das estimativas, como podem ser vistos:

Tabela 3 – Variâncias e Covariâncias das estimativas dos parâmetros da GEV calculadas via Método de Máxima Verossimilhança, para os dados de níveis máximos de IPC da Rússia, China, Índia e a média dos três países (MRCI).

	Rússia	China	Índia	MRCI
$Var(\mu_0)$	0,006812229	0,0170332759	0,006791435	0,0016217703
$Var(\sigma_0)$	0,0043364572	0,002744743	0,003926540	0,0010337162
$Var(\xi_0)$	0,0170332759	0,009196113	0,014226016	0,033642872
$Cov(\mu_0, \sigma_0)$	0,003050444	0,001067108	0,002571373	0,0007691314
$Cov(\mu_0, \xi_0)$	-0,003268730	-0,002376313	-0,003146090	-0,0035083671
$Cov(\sigma_0, \xi_0)$	-0,0005958221	-0,001654709	-0,001122958	-0,0023156966

Fonte: Elaborado pelo autor.

As variâncias e covariâncias encontradas na Tabela 3 foram utilizadas para construção do hiperparâmetro  $\Sigma_0$ . Para averiguar diferentes estruturas de variabilidade, e conseqüentemente colocar mais ou menos peso na distribuição *a priori*, essas matrizes foram multiplicadas por 1, 2 e 4 ( $1\Sigma_0$ ,  $2\Sigma_0$ ,  $4\Sigma_0$ ), sendo assim resultou-se em vinte e quatro modelos Bayesianos analisados (dois tipos de distribuição *a priori*, quatro conjuntos de dados diferentes e três diferentes pesos no hiperparâmetro  $\Sigma_0$ ).

Com a construção das distribuições *a priori*, pode-se encontrar a distribuição *a posteriori* dos parâmetros da GEV, sendo a proporção da multiplicação das equações (3.1) e (3.3):

$$p(\mu, \sigma, \xi | z_1, \dots, z_n) \propto |\Sigma_0|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\theta - \Phi_0)^T \Sigma_0^{-1} (\theta - \Phi_0) \right\} \times \\ \times \left\{ -n \log \sigma - \left( \frac{1 + \xi}{\xi} \right) \sum_{i=1}^n \log \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left[ 1 + \xi \left( \frac{z_i - \mu}{\sigma} \right) \right] \right\}, \quad (3.5)$$

Como as distribuições marginais, para cada parâmetro, da distribuição *a posteriori* não podem ser encontradas analiticamente, foi utilizada simulação MCMC. Para que as simulações fossem eficientes foram criadas cadeias com 10 milhões de valores (para todos os modelos), excluindo os 100 mil primeiros valores calculados (*burn-in* de 100 mil) e, além disso, foram dados saltos a cada 100 valores (*thin* de 100), isso resultava em cadeias finais com 99 mil valores.

Os critérios para analisar se houve desvio de convergência dessas cadeias foram o de Raftery-Lewis (1992) e Geweke (1992). Dessa forma, sempre que o critério de Raftery-Lewis

calculado para as cadeias fosse próxima a 1, seria um indício de convergência das mesmas. Já para analisar o critério de Geweke foi avaliado o valor normalizado, ou seja, como se utilizou 95% de confiança para o critério, o valor normalizado não poderia ser maior que 1,96 ou menor que -1,96.

Para calcular as estimativas intervalares utilizou-se os intervalos HPD com 95% de credibilidade.

### 3.3.4 Avaliação dos Modelos

Para avaliar os modelos foram analisados os níveis de retorno preditos, os quais foram obtidos conforme Naghettini e Pinto (2007):

$$z(T) = \mu + \frac{\sigma}{\xi} \left[ (z_p)^{-\xi} - 1 \right], \quad (3.6)$$

em que  $z(T)$  é o nível de retorno,  $z_p = -\log(1 - p_t)$ ,  $p_t = 1/T$  e  $T$  é o tempo de retorno desejado. A interpretação prática do nível de retorno é que, espera-se que em um tempo médio de retorno de  $T$  quadrimestres, o nível de inflação máximo seja maior ou igual a  $z(T)$  em pelo menos um mês. Foram utilizados cinco tempos de retorno, para 2, 3, 4, 5 e 10 quadrimestres.

A fim de estudar a acurácia e a precisão dos valores preditos, foram utilizadas as séries para ajuste e para verificação (seção 4.1). Logo, pôde-se calcular os intervalos de confiança, para as distribuições com parâmetros estimados pelo método da Máxima Verossimilhança, e os intervalos HPD, para os modelos Bayesianos. Através das estimativas intervalares, foi possível investigar suas amplitudes (precisão) e a cobertura dos valores reais observados (acurácia), ou seja, se o intervalo englobava o valor observado dentro de seus limites.

Além disso foram calculados os Erros Médios Absolutos de Predição (EMAP), para analisar a acurácia, e o Erro Quadrático Médio de Predição, para examinar a variabilidade das predições. O cálculo do Erro Médio Absoluto de Predição (EMAP) foi feito conforme a equação:

$$\text{EMAP} = \sum_{i=1}^N \frac{|VO_i - VP_i|}{N} \quad (3.7)$$

em que  $N$  é o número de predições feitas, sendo o  $i$ -ésimo Valor Observado ( $VO$ ) e o  $i$ -ésimo



Valor Predito ( $VP$ ) referente ao tempo de retorno  $T$ . O cálculo do Erro Quadrático Médio de Predição (EQMP) foi feito conforme a equação:

$$\text{EQMP} = \sum_{i=1}^N \frac{(VO_i - VP_i)^2}{N} \quad (3.8)$$

em que  $N$  é o número de predições, o  $i$ -ésimo Valor Observado ( $VO$ ) e o  $i$ -ésimo Valor Predito ( $VP$ ) referente ao tempo de retorno  $T$ .

### 3.4 RECURSOS COMPUTACIONAIS

Todas as análises estatísticas foram realizadas utilizando o *software* R Core Team (2016) e seus respectivos pacotes: *coda* (PLUMMER et al., 2015), *evd* (STEPHENSON, 2012), *evd-bayes* (STEPHENSON; RIBATET, 2006) e *extRemes* (GILLELAND; KATZ, 2015).

## 4 RESULTADOS E DISCUSSÃO

Primeiro foi feita a modelagem dos níveis máximos quadrimestrais de IPCA utilizando a série para ajuste. Dessa forma, foi possível calcular a acurácia e precisão das predições dos modelos ajustados, baseando-se na série para verificação.

Com base na acurácia e precisão da predição dos níveis de retorno, foi possível selecionar os modelos Bayesianos que obtiveram melhores resultados. Assim, foi feita a modelagem da série completa de níveis máximos quadrimestrais de IPCA, usando os modelos selecionados.

### 4.1 SELEÇÃO DOS MODELOS

Inicialmente foi feito o teste de Ljung-Box na série para ajuste dos níveis máximos de IPCA, o teste demonstrou que a série é independente ( $p=0,8069$ ). Além disso foi feito o *Runs-test* que indicou que a série é aleatória ( $p=0,5745$ ). Com base nos resultados desses testes foi possível ajustar as distribuições GEV e Gumbel, utilizando o método da Máxima Verossimilhança para estimação dos parâmetros. Podemos ver os ajustes a seguir:

Tabela 4 – Estimativas, calculadas via método da Máxima Verossimilhança, e Desvios Padrão dos parâmetros da distribuição GEV e seus respectivos intervalos (I.C.) com 95% confiança.

Distribuição	Parâmetro	Estimativa	Desvio Padrão	I.C. 95%
GEV	$\mu$	0,5303	0,0405	[ 0,4508 ; 0,6098]
	$\sigma$	0,2013	0,0291	[0,1442 ; 0,2583]
	$\xi$	0,0902	0,0926	[ -0,0912 ; 0,2717]
Gumbel	$\mu$	0,5411	0,0402	[0,4624 ; 0,6198]
	$\sigma$	0,2070	0,0297	[0,1488 ; 0,2652]

Fonte: Elaborado pelo autor.

O teste de Kolmogorov-Smirnov indicou aderência das distribuições GEV ( $p=0,6240$ ) e Gumbel ( $p=0,6728$ ) à série de níveis máximos quadrimestrais de IPCA.

Além disso, o Teste de Razão de Verossimilhança mostrou que a diferença entre as duas distribuições não foi significativa ( $p=0,2583$ ), ou seja, o parâmetro  $\xi$  não é significativo para a distribuição GEV. Dessa forma, a distribuição Gumbel, por ser menos complexa e explicar o conjunto de dados da mesma forma que a distribuição GEV, seria mais adequada. Entretanto, analisou-se também a acurácia e precisão, das duas distribuições na predição dos níveis de

retorno, para avaliar qual das duas é mais apropriada para o trabalho.

Foram calculados os intervalos de confiança para os níveis de retorno de acordo com tempos estipulados, as amplitudes dos intervalos, o Erro Médio Absoluto de Predição (EMAP) e o Erro Quadrático Médio de Predição (EQMP), como pode ser visto a seguir:

Tabela 5 – Intervalos (I.C.) com 95% confiança e amplitudes para os níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, a partir das distribuições Gumbel e GEV e seus respectivos EMAP e EQMP.

Distribuição	T.R.	Valor Observado	I.C.	Amplitude	EMAP	EQMP
Gumbel	2	0,86	[0,529 ; 0,704]	0,175	0,1754	0,0387
	3	0,86	[0,621 ; 0,834]	0,213		
	4	0,92	<b>[0,677 ; 0,920]</b>	0,243		
	5	0,92	<b>[0,718 ; 0,985]</b>	0,267		
	10	1,32	[0,835 ; 1,178]	0,342		
GEV	2	0,86	[0,515 ; 0,695]	0,180	0,1561	0,0311
	3	0,86	[0,607 ; 0,832]	0,225		
	4	0,92	<b>[0,664 ; 0,927]</b>	0,263		
	5	0,92	<b>[0,705 ; 1,001]</b>	0,296		
	10	1,32	[0,818 ; 1,247]	0,429		

Fonte: Elaborado pelo autor.

Nota: em negrito os intervalos que englobaram os valores reais.

Analisando os resultados apresentados na Tabela 5, percebe-se que a acurácia e precisão dos níveis de retorno calculados a partir das duas distribuições, foram parecidas. Exceto o nível de retorno para 10 quadrimestres, nesse caso a distribuição Gumbel apresentou menor amplitude, sendo mais precisa. Ambas englobaram em seus intervalos de confiança apenas dois dos cinco valores observados, para 4 e 5 quadrimestres. Entretanto, o EMAP e EQMP foram menores para a distribuição GEV. Dessa forma é visto que os níveis de retorno calculados a partir da distribuição GEV, como um todo, tiveram maior acurácia e precisão que os calculados a partir da distribuição Gumbel.

Além disso, observando os resultados apresentados na Tabela 5, é possível averiguar que os intervalos de confiança dos níveis de retorno calculados a partir das duas distribuições estão subestimando os valores reais, ou seja, em todos os casos em que os valores observados não foram englobados pelos intervalos, os mesmos foram maiores que o limite superior dos intervalos. Isso acarreta um prejuízo para os investidores, por exemplo. Pois seriam calculados máximos de ganhos menores que os reais, fazendo-os investir em outros títulos de renda fixa, que poderiam vir a ser menos rentáveis que o NTN-B (indexado no IPCA).

Também foi ajustada a distribuição GEV utilizando a Inferência Bayesiana para a estimação dos parâmetros, com propósito de verificar qual das abordagens proporciona maior

precisão e acurácia das estimativas dos níveis de retorno. Para isso, primeiramente foi analisado a independência das séries para ajuste de níveis máximos quadrimestrais de IPC da Rússia, China, Índia e da média (MRCI). Assim como na série para ajuste de níveis máximos de IPCA, foi utilizado o teste de Ljung-Box. O mesmo indicou que todas as séries de máximos eram independentes: Rússia ( $p=0,5245$ ), China ( $p=0,1621$ ), Índia ( $p=0,8894$ ) e MRCI ( $p=0,7010$ ). Sabendo disso foram ajustados os modelos. Na Tabela 6 podem ser vistas as estimativas dos parâmetros dos modelos que utilizaram a distribuição a *priori* Normal I, além dos critérios de convergência e intervalos HPD com 95% de credibilidade:

Tabela 6 – Estimativas, desvios padrão (D.P) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos, utilizando a distribuição a *priori* Normal I com diferentes matrizes de variâncias e covariâncias (Var) e respectivos resultados dos critérios de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).

Modelo (Var)	Parâmetro	Estimativa	D.P.	HPD	R-L	$Z_g$
Rússia (1Var)	$\mu$	0,7030	0,0501	[0,6040 ; 0,8005]	1,00	-0,05
	$\sigma$	0,4099	0,0254	[0,3614 ; 0,4611]	1,00	-0,41
	$\xi$	0,2890	0,1094	[0,0762 ; 0,5033]	0,99	0,89
Rússia (2Var)	$\mu$	0,6510	0,0544	[0,5451 ; 0,7587]	0,99	-1,30
	$\sigma$	0,3829	0,0320	[0,3227 ; 0,4471]	1,01	-1,14
	$\xi$	0,2802	0,1386	[0,0076 ; 0,5462]	1,00	-0,07
Rússia (4Var)	$\mu$	0,6074	0,0565	[0,4962 ; 0,7190]	1,01	1,60
	$\sigma$	0,3481	0,0378	[0,2772 ; 0,4236]	0,99	0,47
	$\xi$	0,2369	0,1577	[-0,0607 ; 0,5395]	1,00	0,60
China (1Var)	$\mu$	0,6191	0,0581	[0,5053 ; 0,7330]	1,01	-1,15
	$\sigma$	0,4206	0,0264	[0,3695 ; 0,4726]	0,99	1,87
	$\xi$	0,0157	0,0912	[-0,1626 ; 0,1922]	1,01	0,06
China (2Var)	$\mu$	0,6008	0,0614	[0,4786 ; 0,7200]	0,99	0,49
	$\sigma$	0,3918	0,0333	[0,3287 ; 0,4581]	0,99	-0,77
	$\xi$	0,0708	0,1169	[-0,1518 ; 0,2957]	1,00	0,68
China (4Var)	$\mu$	0,5860	0,0597	[0,4678 ; 0,7025]	1,00	0,59
	$\sigma$	0,3523	0,0387	[0,2793 ; 0,4290]	1,01	1,12
	$\xi$	0,1268	0,1363	[-0,1263 ; 0,3933]	1,00	-0,86
Índia (1Var)	$\mu$	0,7019	0,0519	[0,5987 ; 0,8014]	1,00	1,22
	$\sigma$	0,3851	0,0283	[0,3316 ; 0,4415]	0,99	1,11
	$\xi$	0,3234	0,1177	[0,0916 ; 0,5520]	1,00	-0,03
Índia (2Var)	$\mu$	0,6416	0,0538	[0,5365 ; 0,7474]	1,00	-0,33
	$\sigma$	0,3540	0,0341	[0,2898 ; 0,4223]	1,01	-1,85
	$\xi$	0,2983	0,1432	[0,0159 ; 0,5729]	1,00	-1,91
Índia (4Var)	$\mu$	0,5963	0,0536	[0,4907 ; 0,7018]	1,01	-0,97
	$\sigma$	0,3172	0,0385	[0,2447 ; 0,3934]	1,00	0,30
	$\xi$	0,2337	0,1531	[-0,0549 ; 0,5277]	1,00	1,08
MRCI (1Var)	$\mu$	0,7353	0,0336	[0,6703 ; 0,8022]	1,00	-1,20
	$\sigma$	0,2224	0,0083	[0,2064 ; 0,2389]	1,00	-0,70
	$\xi$	-0,0044	0,0529	[-0,1029 ; 0,0996]	1,00	-0,04
MRCI (2Var)	$\mu$	0,6518	0,0352	[0,5829 ; 0,7203]	1,00	0,65
	$\sigma$	0,2155	0,0107	[0,1953 ; 0,2369]	1,01	1,17
	$\xi$	0,0548	0,0658	[-0,0675 ; 0,1843]	1,01	-1,49
MRCI (4Var)	$\mu$	0,5978	0,0369	[0,5260 ; 0,6709]	1,00	-0,05
	$\sigma$	0,2114	0,0137	[0,1850 ; 0,2389]	1,01	0,25
	$\xi$	0,0883	0,0757	[-0,0535 ; 0,2344]	0,99	-0,23

Fonte: Elaborado pelo autor.

Observa-se na Tabela 6, que os critérios de convergência indicaram que não houve indícios de que as cadeias não convergiram, pois o critério de Raftery-Lewis se aproximou de 1

em todos os casos e os resultados do critério de Geweke não foram significativos para nenhuma cadeia ( $|Z_g| < 1,96$ ). Os resultados do ajuste dos modelos que utilizaram a distribuição *a priori* Normal II estão apresentados na Tabela 7.

Tabela 7 – Estimativas, desvios padrão (D.P.) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos, utilizando a distribuição *a priori* Normal II com diferentes matrizes de variâncias e covariâncias (Var) e respectivos resultados dos critérios de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).

Modelo (Var)	Parâmetro	Estimativa	D.P.	HPD	R-L	$Z_g$
Rússia (1Var)	$\mu$	0,7352	0,0430	[0,6521 ; 0,8207]	1,00	-1,58
	$\sigma$	0,4131	0,0260	[0,3627 ; 0,4643]	1,00	0,51
	$\xi$	0,2844	0,1059	[0,0764 ; 0,4895]	1,00	0,64
Rússia (2Var)	$\mu$	0,6802	0,0481	[0,5879 ; 0,7762]	1,01	-1,35
	$\sigma$	0,3838	0,0332	[0,3199 ; 0,4491]	1,00	-1,03
	$\xi$	0,2846	0,1325	[0,0266 ; 0,5427]	1,00	0,84
Rússia (4Var)	$\mu$	0,6267	0,0507	[0,5284 ; 0,7267]	1,00	0,62
	$\sigma$	0,3442	0,0389	[0,2718 ; 0,4224]	1,01	-1,22
	$\xi$	0,2460	0,1507	[-0,0482 ; 0,5298]	1,01	1,85
China (1Var)	$\mu$	0,6429	0,0451	[0,5555 ; 0,7319]	1,00	0,39
	$\sigma$	0,4200	0,0267	[0,3684 ; 0,4730]	1,01	0,46
	$\xi$	0,0253	0,0928	[-0,1551 ; 0,2059]	1,00	-1,48
China (2Var)	$\mu$	0,6180	0,0509	[0,5190 ; 0,7180]	1,00	-0,06
	$\sigma$	0,3899	0,0334	[0,3279 ; 0,4581]	1,00	-1,58
	$\xi$	0,0861	0,1181	[-0,1385 ; 0,3161]	0,99	1,11
China (4Var)	$\mu$	0,5944	0,0530	[0,4908 ; 0,6985]	1,00	1,18
	$\sigma$	0,3507	0,0384	[0,2792 ; 0,4282]	1,01	0,71
	$\xi$	0,1403	0,1364	[-0,1161 ; 0,4050]	1,00	-0,27
Índia (1Var)	$\mu$	0,7418	0,0467	[0,6505 ; 0,8330]	1,01	1,18
	$\sigma$	0,3926	0,0291	[0,3361 ; 0,4495]	1,00	0,62
	$\xi$	0,2971	0,1113	[0,0717 ; 0,5058]	0,99	-0,04
Índia (2Var)	$\mu$	0,6736	0,0492	[0,5794 ; 0,7724]	1,00	0,40
	$\sigma$	0,3575	0,0356	[0,2909 ; 0,4298]	1,01	-0,18
	$\xi$	0,2860	0,1338	[0,0223 ; 0,5446]	1,00	0,20
Índia (4Var)	$\mu$	0,6156	0,0501	[0,5166 ; 0,7133]	1,00	0,49
	$\sigma$	0,3154	0,0399	[0,2412 ; 0,3948]	1,01	1,16
	$\xi$	0,2323	0,1456	[-0,0439 ; 0,5155]	1,01	0,89
MRCI (1Var)	$\mu$	0,7951	0,0286	[0,7404 ; 0,8522]	1,00	0,75
	$\sigma$	0,2300	0,0084	[0,2135 ; 0,2466]	1,01	1,10
	$\xi$	-0,0531	0,0417	[-0,1289 ; 0,0293]	1,00	1,24
MRCI (2Var)	$\mu$	0,7162	0,0310	[0,6551 ; 0,7764]	0,99	1,30
	$\sigma$	0,2239	0,0112	[0,2025 ; 0,2462]	1,00	1,37
	$\xi$	0,0005	0,0536	[-0,1003 ; 0,1041]	1,01	0,65
MRCI (4Var)	$\mu$	0,6526	0,0333	[0,5878 ; 0,7182]	1,00	0,31
	$\sigma$	0,2184	0,0145	[0,1911 ; 0,2477]	1,00	1,02
	$\xi$	0,0436	0,0642	[-0,0742 ; 0,1706]	0,99	1,12

Fonte: Elaborado pelo autor.

De acordo com os resultados apresentados na Tabela 7, também não houve evidências para concluir falta de convergência para essas cadeias, pois para todas elas o critério de Raftery-Lewis se aproximou de 1 e os valores normalizados do critério de Geweke não foram significativos para nenhum caso ( $|Z_g| < 1,96$ ). É possível observar que os modelos que utilizaram como informação *a priori* a média dos três países (MRCI) apresentaram os menores desvios padrão, consequentemente intervalos HPD mais precisos. Isso acontece tanto para os modelos com distribuição *a priori* Normal I, como para Normal II.

Na sequência, foram calculados os níveis de retorno para os tempos de 2, 3, 4, 5 e 10 quadrimestres. No entanto, para avaliar os níveis de retorno e seus intervalos, primeiramente foram investigados os critérios de convergência das cadeias dos níveis de retorno.

Tabela 8 – Critérios de convergência Raftery-Lewis (R-L) e Geweke ( $Z_g$ ) das cadeias dos níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres dos modelos com distribuições *a priori* Normal I (N I) e Normal II (N II) e diferentes matrizes de variâncias e covariâncias (Var).

T.R.	2		3		4		5		10	
Modelo (Var)	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$
Rússia N I (1Var)	0,99	-0,07	1,00	0,11	1,01	0,30	1,00	0,43	0,98	0,65
Rússia N I (2Var)	0,99	-1,43	0,99	-1,47	1,00	-1,39	0,99	-1,29	1,00	-1,03
Rússia N I (4Var)	1,01	1,54	1,01	1,41	1,02	1,29	1,020	1,21	1,01	1,04
China N I (1Var)	0,99	-0,81	0,99	-0,33	0,99	-0,04	1,00	0,13	1,01	0,47
China N I (2Var)	0,99	0,35	0,99	0,25	0,99	0,24	0,99	0,25	0,99	0,33
China N I (4Var)	0,99	0,95	1,00	1,17	1,00	1,08	1,00	0,96	1,01	0,49
Índia N I (1Var)	1,00	1,29	0,99	1,16	0,99	1,06	0,99	0,90	1,00	0,51
Índia N I (2Var)	0,99	-0,79	1,01	-1,59	1,00	-1,96	0,99	-2,14	0,99	-2,30
Índia N I (4Var)	1,00	-0,72	1,00	-0,21	1,01	0,10	1,00	0,29	0,99	0,66
MRCI N I (1Var)	0,99	-1,53	0,99	-1,38	1,01	-1,21	1,00	-1,04	0,99	-0,32
MRCI N I (2Var)	0,99	0,71	0,99	0,66	0,99	0,54	0,99	0,42	1,00	-0,05
MRCI N I (4Var)	0,99	-0,02	0,99	-0,01	0,98	-0,03	0,99	-0,06	0,99	-0,15
Rússia N II (1Var)	0,99	-1,23	1,01	-0,58	1,01	-0,13	0,99	0,14	0,99	0,57
Rússia N II (2Var)	1,00	-1,36	1,01	-1,07	1,01	-0,76	1,01	-0,54	1,00	-0,08
Rússia N II (4Var)	0,99	0,32	0,99	0,32	1,00	0,48	1,00	0,62	1,00	0,98
China N II (1Var)	0,99	0,38	0,99	0,09	0,99	-0,17	0,99	-0,38	0,99	-0,88
China N II (2Var)	0,99	-0,34	0,99	-0,44	0,99	-0,34	0,99	-0,22	0,99	0,14
China N II (4Var)	0,99	1,23	0,99	1,11	1,00	0,95	1,00	0,81	0,99	0,43
Índia N II (1Var)	1,00	1,15	1,01	1,00	1,01	0,83	1,00	0,70	1,00	0,40
Índia N II (2Var)	0,98	0,30	0,98	0,17	0,99	0,08	0,99	0,02	0,99	-0,15
Índia N II (4xVar)	0,99	0,75	1,01	1,10	1,01	1,24	1,01	1,30	1,01	1,37
MRCI N II (1Var)	0,99	0,67	0,99	1,21	0,99	1,08	1,01	1,54	1,01	1,47
MRCI N II (2Var)	1,01	1,03	0,99	0,07	0,99	1,32	0,99	0,32	1,00	1,30
MRCI N II (4Var)	0,99	0,97	1,01	0,87	0,99	0,87	0,99	0,78	0,99	1,62
	Observado	0,86	0,86	0,92	0,92	1,32				

Fonte: Elaborado pelo autor.

Observando os resultados apresentados na Tabela 8, é visto que também não houve evidências que mostrassem falta de convergência das cadeias dos níveis de retorno, pois o critério de Raftery-Lewis se aproximou de 1 para todas elas e o critério de Geweke não foi significativo para nenhum caso ( $|Z_g| < 1,96$ ). Dessa forma, puderam ser calculados os intervalos HPD com 95% de credibilidade dos níveis de retorno.

Tabela 9 – Intervalos HPD com 95% de credibilidade dos níveis de retorno para o tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, calculados dos modelos utilizando a distribuição a *priori* Normal I com diferentes matrizes de variâncias e covariâncias (Var).

Modelo (Var)		T.R.					Acertos
		2	3	4	5	10	
Rússia (1Var)	HPD	<b>[0,746 ; 0,991]</b>	[1,008 ; 1,308]	[1,179 ; 1,556]	[1,303 ; 1,773]	[1,669 ; 2,654 ]	1
	Amp	0,245	0,300	0,377	0,470	0,985	
Rússia (2Var)	HPD	<b>[0,673 ; 0,933]</b>	[0,898 ; 1,248]	[1,039 ; 1,500]	[1,144 ; 1,728]	[1,427 ; 2,643]	1
	Amp	0,260	0,349	0,460	0,585	1,216	
Rússia (4Var)	HPD	<b>[0,612 ; 0,877]</b>	<b>[0,788 ; 1,162]</b>	<b>[0,900 ; 1,397]</b>	[0,981 ; 1,606]	<b>[1,191 ; 2,418]</b>	4
	Amp	0,265	0,374	0,497	0,625	1,227	
China (1Var)	HPD	<b>[0,655 ; 0,891]</b>	[0,874 ; 1,131]	[1,007 ; 1,292]	[1,103 ; 1,418]	[1,359 ; 1,834]	1
	Amp	0,235	0,257	0,285	0,315	0,474	
China (2Var)	HPD.	<b>[0,619 ; 0,874]</b>	<b>[0,820 ; 1,114]</b>	[0,946 ; 1,285]	[1,035 ; 1,423]	<b>[1,267 ; 1,889]</b>	3
	Amp	0,255	0,294	0,339	0,388	0,622	
China (4Var)	HPD.	[0,593 ; 0,848]	<b>[0,771 ; 1,084]</b>	<b>[0,877 ; 1,255]</b>	[0,957 ; 1,402]	<b>[1,172 ; 1,922]</b>	3
	Amp	0,255	0,313	0,378	0,445	0,750	
Índia (1Var)	HPD	<b>[0,728 ; 0,963]</b>	[0,932 ; 1,235]	[1,062 ; 1,456]	[1,148 ; 1,644]	[1,383 ; 2,385]	1
	Amp	0,235	0,303	0,394	0,495	1,002	
Índia (2Var)	HPD	<b>[0,646 ; 0,894]</b>	<b>[0,817 ; 1,153]</b>	[0,929 ; 1,367]	[1,005 ; 1,551]	<b>[1,200 ; 2,248]</b>	3
	Amp	0,248	0,336	0,438	0,546	0,1,048	
Índia (4Var)	HPD	[0,590 ; 0,837]	<b>[0,738 ; 1,072]</b>	<b>[0,836 ; 1,265]</b>	<b>[0,898 ; 1,421]</b>	<b>[1,082 ; 2,032]</b>	4
	Amp	0,247	0,334	0,429	0,523	0,950	
MRCI (1Var)	HPD	<b>[0,747 ; 0,884]</b>	[0,865 ; 1,005]	[0,941 ; 1,083]	[0,997 ; 1,140]	[1,154 ; 1,317]	1
	Amp	0,137	0,140	0,142	0,143	0,163	
MRCI (2Var)	HPD	[0,660 ; 0,805]	<b>[0,777 ; 0,929]</b>	<b>[0,849 ; 1,008]</b>	<b>[0,901 ; 1,072]</b>	[1,067 ; 1,277]	3
	Amp	0,145	0,153	0,159	0,166	0,210	
MRCI (4Var)	HPD	[0,601 ; 0,754]	<b>[0,714 ; 0,881]</b>	<b>[0,791 ; 0,970]</b>	<b>[0,845 ; 1,037]</b>	[0,997 ; 1,263]	3
	Amp	0,153	0,166	0,179	0,192	0,266	
	Obs	0,86	0,86	0,92	0,92	1,32	

Fonte: Elaborado pelo autor.

Nota: em negrito os intervalos que englobaram os valores reais.

Com base nos resultados apresentados na Tabela 9, podem ser vistos os intervalos HPD com 95% de credibilidade dos níveis de retorno calculados a partir dos modelos que utilizaram a distribuição a *priori* Normal I. Analisando tal tabela, é possível ver que os intervalos calculados a partir dos modelos, que utilizaram como informação a *priori* apenas um país, superestimaram o máximo de IPCA. Porém, quando a variabilidade da distribuição a *priori* foi aumentada, ou seja, o peso da distribuição a *priori* foi menor, os intervalos passaram a englobar uma maior parte dos valores reais de IPCA máximo. Todavia, é possível perceber também que a amplitude dos intervalos foi grande, ou seja, as estimativas intervalares tiveram baixa precisão.

Este fato pode ser prejudicial para qualquer agente econômico que esteja interessado nos valores máximos de IPCA. Conforme IPEA-DATA (2016) a média de inflação máxima mensal para os três quadrimestres de 2015 foi de 0,75%, dessa forma, uma amplitude de 1%, por exemplo, pode ser considerada alta para utilizar esses modelos na predição da inflação máxima.

Os intervalos HPD dos níveis de retorno, calculados a partir dos modelos que utilizaram como informação a *priori* a média do IPC máximo (MRCI) dos três países, tiveram maior precisão. Além disso, os intervalos também conseguiram englobar maior quantidade dos valores

reais de IPCA máximo quando a variabilidade da distribuição *a priori* era aumentada.

Devido ao fato da média dos três países amenizar quedas ou altas do IPC que podem ter sido acarretadas por acontecimentos políticos, econômicos ou sociais próprios de cada país e cada época. Carstensen et al. (2009) criaram modelos com objetivo de prever a demanda da moeda e inflação da Alemanha, Itália, França e Espanha. Quando estimaram a previsão para o agregado dos países, os resultados foram mais satisfatórios do que para cada um deles em particular.

Tabela 10 – Intervalos HPD com 95% de credibilidade dos níveis de retorno para o tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, calculados dos modelos utilizando a distribuição *a priori* Normal II com diferentes matrizes de variâncias e covariâncias (Var).

Modelo (Var)		T.R.					Acertos
		2	3	4	5	10	
Rússia (1Var)	HPD	<b>[0,791 ; 1,012]</b>	[1,053 ; 1,328]	[1,224 ; 1,575]	[1,352 ; 1,796]	[1,726 ; 2,67]	1
	Amp	0,221	0,276	0,351	0,444	0,947	
Rússia (2Var)	HPD	<b>[0,710 ; 0,951]</b>	[0,937 ; 1,268]	[1,082 ; 1,525]	[1,185 ; 1,752]	[1,493 ; 2,691]	1
	Amp	0,241	0,331	0,443	0,567	1,198	
Rússia (4Var)	HPD	<b>[0,634 ; 0,886]</b>	<b>[0,814 ; 1,178]</b>	[0,926 ; 1,418]	[1,001 ; 1,627]	<b>[1,220 ; 2,450]</b>	3
	Amp	0,252	0,365	0,492	0,622	1,230	
China (1Var)	HPD	<b>[0,705 ; 0,892]</b>	[0,920 ; 1,133]	[1,058 ; 1,300]	[1,152 ; 1,428]	[1,408 ; 1,854]	1
	Amp	0,187	0,212	0,242	0,276	0,446	
China (2Var)	HPD	<b>[0,659 ; 0,873]</b>	<b>[0,858 ; 1,115]</b>	[0,983 ; 1,288]	[1,071 ; 1,428]	<b>[1,308 ; 1,914]</b>	3
	Amp	0,215	0,256	0,305	0,356	0,607	
China (4Var)	HPD	[0,613 ; 0,842]	<b>[0,792 ; 1,821]</b>	<b>[0,900 ; 1,257]</b>	[0,976 ; 1,403]	<b>[1,189 ; 1,934]</b>	3
	Amp	0,230	1,029	0,358	0,426	0,746	
Índia (1Var)	HPD	<b>[0,780 ; 0,994]</b>	[0,981 ; 1,266]	[1,109 ; 1,482]	[1,198 ; 1,666]	[1,436 ; 2,354]	1
	Amp	0,214	0,285	0,373	0,468	0,919	
Índia (2Var)	HPD	<b>[0,685 ; 0,916]</b>	<b>[0,859 ; 1,181]</b>	[0,966 ; 1,387]	[1,040 ; 1,561]	<b>[1,242 ; 2,227]</b>	3
	Amp	0,231	0,322	0,421	0,521	0,985	
Índia (4Var)	HPD	[0,612 ; 0,846]	<b>[0,758 ; 1,083]</b>	<b>[0,843 ; 1,262]</b>	<b>[0,906 ; 1,414]</b>	<b>[1,084 ; 1,986]</b>	4
	Amp	0,234	0,325	0,419	0,508	0,902	
MRCI (1Var)	HPD	<b>[0,819 ; 0,938]</b>	[0,935 ; 1,059]	[1,007 ; 1,135]	[1,060 ; 1,192]	<b>[1,210 ; 1,350]</b>	2
	Amp	0,118	0,125	0,129	0,132	0,149	
MRCI (2Var)	HPD	<b>[0,734 ; 0,861]</b>	<b>[0,848 ; 0,988]</b>	<b>[0,921 ; 1,070]</b>	[0,974 ; 1,130]	<b>[1,125 ; 1,318]</b>	4
	Amp	0,129	0,140	0,148	0,156	0,193	
MRCI (4Var)	HPD	[0,664 ; 0,805]	<b>[0,773 ; 0,932]</b>	<b>[0,847 ; 1,021]</b>	<b>[0,900 ; 1,087]</b>	[1,048 ; 1,297]	3
	Amp	0,141	0,159	0,173	0,187	0,250	
	Obs	0,86	0,86	0,92	0,92	1,32	

Fonte: Elaborado pelo autor.

Nota: em negrito os intervalos que englobaram os valores reais.

Na Tabela 10 estão apresentados os intervalos HPD com 95% de credibilidade dos níveis de retorno calculados com base nos modelos com distribuição *a priori* Normal II. O mesmo padrão encontrado nos intervalos HPD calculados para os modelos com distribuição *a priori* Normal I (Tabela 9) pode ser verificado nos intervalos dos níveis de retorno a partir dos modelos com distribuição *a priori* Normal II.

Porém foi observado, conforme a Tabela 10, que os modelos que utilizaram como informação *a priori* a média dos três países (MRCI) tiveram resultados mais precisos nos intervalos HPD dos níveis de retorno calculados a partir dos modelos com distribuição *a priori*



Normal II. A precisão dos intervalos foi maior, ou seja, a amplitude foi menor. Além de que os intervalos englobaram maior quantidade dos valores observados de IPCA máximo para os tempos de retorno estipulados. O hiperparâmetro  $\mu_0$  seguir uma distribuição assimétrica (como a Lognormal) pode ser uma das razões para que os modelos que usaram a distribuição *a priori* Normal II apresentem intervalos para as predições dos níveis de retorno com maior acurácia.

Além dos intervalos HPD, também foram analisados os Erros Médios Absolutos de Predição (EMAP) e os Erros Quadráticos Médios de Predição (EQMP).

Tabela 11 – Erro Médio Absoluto de Predição (EMAP) e Erro Quadrático Médio de Predição (EQMP) dos modelos Bayesianos com distribuição *a priori* Normal I e Normal II e diferentes matrizes de variâncias e covariâncias (Var).

Modelo (Var)	EMAP		EQMP	
	Normal I	Normal II	Normal I	Normal II
Rússia (1Var)	0,4258	0,2171	0,2477	0,0573
Rússia (2Var)	0,3334	0,3609	0,1491	0,1789
Rússia (4Var)	0,1951	0,2167	0,0452	0,0570
China (1Var)	0,2136	0,2309	0,0538	0,0647
China (2Var)	0,1927	0,2078	0,0429	0,0508
China (4Var)	0,1599	0,1681	0,0295	0,0327
Índia (1Var)	0,3160	0,3628	0,1342	0,1651
Índia (2Var)	0,2279	0,2402	0,0640	0,0725
Índia (4Var)	0,1391	0,1398	0,0229	0,0227
MRCI (1Var)	0,0889	0,1104	0,0091	0,0173
MRCI (2Var)	<b>0,0733</b>	<b>0,0853</b>	<b>0,0088</b>	<b>0,0080</b>
MRCI (4Var)	0,1002	<b>0,0734</b>	0,0154	<b>0,0088</b>

Fonte: Elaborado pelo autor.

Nota: em negrito os três menores valores para cada Erro.

De acordo com os resultados apresentados na Tabela 11, pode-se concluir que os modelos que utilizaram como informação *a priori* a média (MRCI) do IPC máximo quadrimestral dos três países obtiveram maior precisão e acurácia no cálculo dos níveis de retorno preditos, ou seja, menor EMAP (acurácia) e menor EQMP (precisão).

As predições para os níveis de retorno calculados a partir dos modelos que levavam como informação *a priori* o IPC máximo de apenas um dos três países tiveram precisão e acurácia semelhantes às predições feitas pelas distribuições GEV e Gumbel, com os parâmetros estimados via Método da Máxima Verossimilhança (Tabela 5).

Analisando os resultados apresentados nas Tabelas 9, 10 e 11, pode-se dizer que o modelo que reuniu maior acurácia e precisão dos níveis de retorno foi o modelo Bayesiano que utilizou como distribuição *a priori* a Normal II, como elicitação a média dos três países

(MRCI) e matriz de variâncias e covariâncias multiplicada por dois (2Var). Outros modelos que utilizaram MRCI como informação *a priori* também obtiveram precisão e acurácia satisfatórias.

## 4.2 AJUSTES DOS MODELOS

Considerando-se os EMAP e EQMP, a amplitude dos intervalos e a quantidade de valores reais englobados pelos intervalos, selecionou-se os modelos Bayesianos que usaram como informação *a priori* a MRCI, sendo com distribuição *a priori* Normal I (2Var), Normal II (2Var) e Normal II (4Var). Estes modelos foram utilizados para ajustar as séries de níveis máximos quadrimestrais de IPCA, referentes a Janeiro de 2003 a Março de 2016.

Primeiramente foi aplicado o teste de Ljung-Box na série completa dos níveis máximos de IPCA. O teste demonstrou que a série é independente ( $p=0,7520$ ). Sabendo disso também foi feito o teste *Runs-test*, o mesmo indicou que a série é aleatória ( $p=0,6291$ ). Com base nesse teste e analisando a Figura 4, pode-se concluir que não houve tendência na série, dessa forma, foi desnecessário ajustar os modelos com tendência. Além disso, pode-se afirmar que a série de IPCA máximo quadrimestral é estacionária.

A estacionaridade da inflação é tratada por autores e pesquisadores de formas diferentes. Ang et al. (2007) e Lee e Wu (2001) comentam que não existe consenso na literatura sobre a inflação ser estacionária ou não-estacionária. Os primeiros autores afirmam que é difícil a inflação ser não-estacionária em modelos econômicos de natureza monetária e natureza neokeynesiana. Já Lee e Wu (2001) explicam que o fato da oferta de moeda ter crescimento estável faz uma economia ter a inflação estacionária.

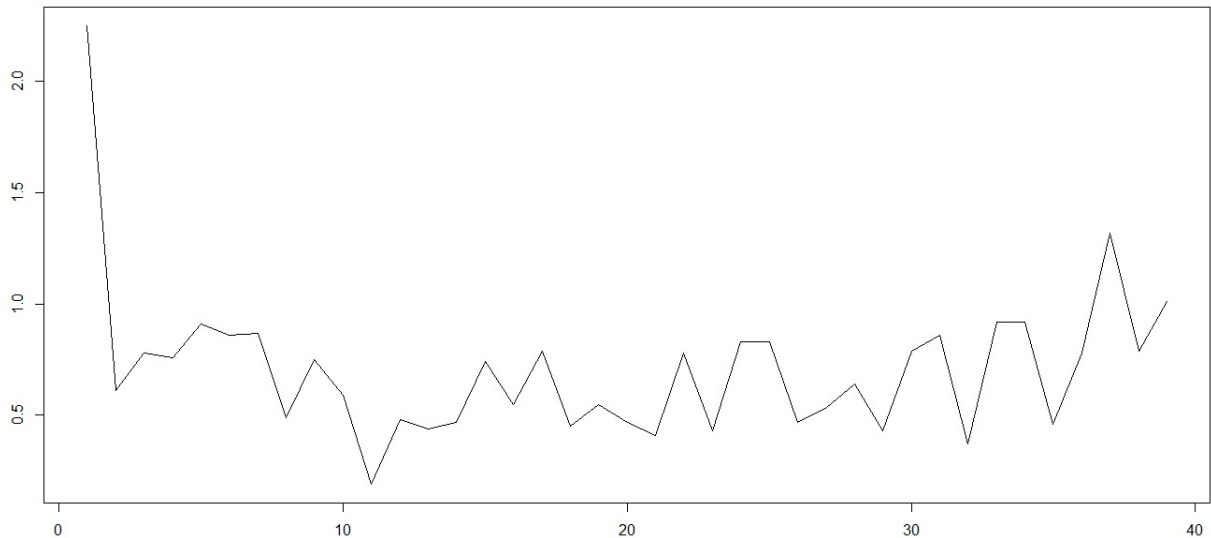


Figura 4 – Série de níveis máximos quadrimestrais de IPCA.  
Fonte: Elaborado pelo autor.

Foram ajustados os modelos Bayesianos selecionados conforme a seção 4.1. Assim, foi analisado se as séries completas de níveis máximos quadrimestrais de IPC da MRCI era independente. O teste de Ljung-Box indicou que a série de máximos quadrimestrais da MRCI é independente ( $p=0,9132$ ). Na Tabela 12 são apresentadas as estimativas dos modelos Bayesianos.

Tabela 12 – Estimativas, desvios padrão (D.P.) e os Intervalos HPD com 95% de credibilidade dos parâmetros dos modelos Bayesianos selecionados, utilizando a distribuição *a priori* Normal Trivariada com diferentes matrizes de variâncias e covariâncias (Var) e respectivos resultados dos critérios de convergência de Raftery-Lewis (R-L) e Geweke ( $Z_g$ ).

Modelo (Var)	Parâmetro	Estimativa	D.P.	HPD	R-L	$Z_g$
MRCI N I (2Var)	$\mu$	0,676	0,028	[0,620 ; 0,733]	1,01	1,077
	$\sigma$	0,192	0,007	[0,178 ; 0,207]	1,00	-0,486
	$\xi$	-0,002	0,043	[-0,083 ; 0,082]	1,02	-1,606
MRCI N II (2Var)	$\mu$	0,739	0,025	[0,690 ; 0,788]	0,99	-0,852
	$\sigma$	0,200	0,008	[0,185 ; 0,216]	1,01	0,654
	$\xi$	-0,045	0,327	[-0,107 ; 0,018]	0,99	0,227
MRCI N II (4Var)	$\mu$	0,679	0,028	[0,624 ; 0,732]	0,99	1,115
	$\sigma$	0,197	0,010	[0,177 ; 0,218]	0,99	0,301
	$\xi$	-0,009	0,042	[-0,089 ; 0,073]	1,01	0,458

Fonte: Elaborado pelo autor.

Os critérios indicaram que não houve indícios para suspeitar de falta de convergência das cadeias, pois o critério de Raftery-Lewis se aproximou de 1 em todos os casos e os resultados do critério de Geweke não foram significativos pra nenhum caso ( $|Z_g| < 1,96$ ). Foram calculados os níveis de retorno para os tempos de 2, 3, 4, 5 e 10 quadrimestre, mas antes os critérios de convergência das cadeias dos níveis de retorno foram analisados.

Tabela 13 – Critérios de convergência Raftery-Lewis (R-L) e Geweke ( $Z_g$ ) das cadeias dos níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres dos modelos Bayesianos selecionados.

T.R.	2		3		4		5		10	
Modelo (Var)	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$	R-L	$Z_g$
MRCI N I (2Var)	1,00	0,56	1,00	0,21	0,99	-0,25	1,02	-0,35	0,99	0,95
MRCI N II (2Var)	1,01	1,02	1,00	-0,35	1,00	0,12	1,00	-1,00	1,00	0,56
MRCI N II (4Var)	1,02	0,14	1,01	0,87	1,01	0,74	1,01	0,28	1,00	0,77

Fonte: Elaborado pelo autor.

Conforme os resultados apresentados na Tabela 13, é visto que não há evidência que indicam a falta de convergência das cadeias, ou seja, o critério de Raftery-Lewis se aproximou de 1 para todas as cadeias calculadas e os valores normalizados, obtidos pelo critério de Geweke, não foram significativos pra nenhum caso ( $|Z_g| < 1,96$ ). Dessa forma, seguiu-se a análise dos níveis de retorno.

Tabela 14 – Níveis de retorno para os tempos (T.R.) de 2, 3, 4, 5 e 10 quadrimestres, seus respectivos intervalos HPD com 95% de credibilidade, calculados a partir dos modelos Bayesianos selecionados.

Modelos	T.R.	2	3	4	5	10
MRCI N I (2x Var)	Estimativas	0,7475	0,8506	0,9165	0,9654	1,1095
	HPD	[0,6885 ; 0,8076]	[0,7886 ; 0,9133]	[0,8530 ; 0,9809]	[0,8988 ; 1,0295]	[1,0364 ; 1,1825]
MRCI N II (2x Var)	Estimativas	0,8117	0,9158	0,9812	1,0289	1,1667
	HPD	[0,7601 ; 0,8653]	[0,8594 ; 0,9730]	[0,9218 ; 1,0406]	[0,9695 ; 1,0924]	[1,0984 ; 1,2360]
MRCI N II (4x Var)	Estimativas	0,7509	0,8561	0,9232	0,9727	1,1185
	HPD	[0,6924 ; 0,8096]	[0,7926 ; 0,9214]	[0,8549 ; 0,9916]	[0,9034 ; 1,0466]	[1,0361 ; 1,2055]

Fonte: Elaborado pelo autor.

Sabendo que a série completa de níveis máximos quadrimestrais de IPCA abrange do primeiro quadrimestre de 2003 ao primeiro quadrimestre de 2016, até a data da publicação deste trabalho passaram-se dois quadrimestre. De acordo com IPEA-DATA (2016), a maior inflação mensal dentro desses dois quadrimestre foi 0,78. Todos os modelos Bayesianos selecionados englobaram esse valor em seus intervalos HPD para o tempo de retorno de 2 quadrimestres (Tabela 14). Esse resultado reforça os bons resultados obtidos com a Inferência Bayesiana na análise da predição dos níveis de retorno para inflação máxima quadrimestral no Brasil.

Analisando tanto os modelos que utilizaram o método da Máxima Verossimilhança, quanto os modelos que fizeram uso da Inferência Bayesiana, é possível apontar os modelos Bayesianos, que usaram como informação *a priori* a média do IPC máximo quadrimestral dos três países (MRCI), como os modelos de maior acurácia e maior precisão. Estes resultados corroboram com os obtidos por Beijo e Avelar (2010), que aplicam a TVE em dados climáticos, e Rios-Alejandro (2011), que aplicou a distribuição GEV na modelagem das temperaturas máximas da cidade de Monterrey no México. Usando Inferência Bayesiana, ambos os trabalhos,

apontam essa metodologia como mais adequada que o Método da Máxima Verossimilhança quando se trata de distribuições de valores extremos.

Outros autores apresentam a Inferência Bayesiana como importante método para modelagem de índices econômicos, como Otrok e Whiteman (1998), que utilizam a modelagem Bayesiana para a predição das condições econômicas do estado de Iowa nos Estados Unidos. Mais recentemente, Giannone et al. (2014) mostrou que a Inferência Bayesiana pode melhorar os resultados na projeção da inflação. Wright (2009), também avaliou alguns modelos para previsão da inflação dos Estados Unidos, o autor comparou o EQMP de vários modelos. Entre eles, Autoregressivos (AR), Autoregressivos com médias móveis (ARMA) e Bayesianos. Concluindo que os modelos Bayesianos foram mais precisos em suas estimativas.

Para um investidor ou gestor financeiro, o uso da metodologia correta é importante, pois a utilização de um modelo que forneça estimativas com baixa acurácia e/ou baixa precisão pode acarretar dois problemas: superestimação ou subestimação na predição da inflação máxima. O primeiro acarreta em uma perda no valor esperado caso os gestores tenham investido em títulos atualizados pela inflação, como NTN-B; o segundo leva a investimentos em outros títulos que venham a render menos que NTN-B.

Quando o investidor, no caso, é uma Entidade de Previdência Complementar, a escolha do modelo torna-se fundamental, pois no Brasil o segmento de renda fixa é o único que a legislação permite que os fundos invistam 100% de suas aplicações. Dessa forma, administradores desses fundos previdenciários buscam sempre otimizar a rentabilidade da entidade, como fizeram Boulier et al. (2001), os quais utilizaram métodos estocásticos para otimização de investimentos de fundos que oferecem previdências com contribuição definida. Além disso, Souza et al. (2015) afirmam que a renda fixa representa o maior percentual de aplicações dentro da política de investimento destes fundos devido ao baixo risco, principalmente dos títulos públicos.

Para as seguradoras é dada a mesma importância, no que tange seguros de vida, pois também são seguros com vigência de muitos anos e precisam de investimentos rentáveis a longo prazo. Grosen e Jorgensen (2000) afirmam que quando a rentabilidade cai muito, as seguradoras do ramo vida acabam tendo um estreitamento dramático da margem de segurança financeira. Assim, a gestão financeira das seguradoras deve contar com modelos bem ajustados para tranquilizar os administradores, ajudando a calcular taxas de juros futuras com maior precisão.

Considerando esses fatos, para a predição do IPCA máximo quadrimestral, os modelos Bayesianos, que utilizaram como informação *a priori* a média do IPC máxima da Rússia, China e Índia, são os mais adequados para descrever essa variável, de forma a dar subsídio para o planejamento e elaboração das políticas de investimento dos fundos atuariais.

## 5 CONCLUSÕES

Foi constatado que a série temporal de níveis máximos quadrimestrais de IPCA, de Janeiro de 2003 a Abril de 2016, não apresentou tendência.

Os parâmetros da distribuição GEV estimados via Inferência Bayesiana apresentaram maior acurácia e precisão na predição dos níveis máximos quadrimestrais de IPCA. Dentre os modelos Bayesianos, o que obteve melhor precisão e acurácia foi o que utilizou como informação *a priori* a média do IPC máximo quadrimestral da Rússia, China e Índia e o dobro da matriz de variâncias e covariâncias, para construção da distribuição *a priori* Normal Trivariada proposta por Stephenson e Ribatet (2006).

A elicitação da informação *a priori* com base em dados de países economicamente semelhantes, juntamente a Inferência Bayesiana, se mostrou eficaz para predição do IPCA máximo. Essa forma de elicitação pode ser utilizada por pesquisadores com aplicações em múltiplas áreas.

O resultados apresentados no trabalho podem ser usados por gestores financeiros, investidores e administradores, que tenham interesse em analisar a inflação máxima quadrimestral, tanto para análise de investimento, quanto para análise econômica.

Pesquisas e estudos futuros podem ser realizados como aprofundamento do tema o presente trabalho: considerar blocos de tempo diferentes (trimestral, semestral, anual); modelagem de outros índices macroeconômicos com a mesma forma de elicitação de informação; e utilização de modelos de valores extremos bivariados, com a incorporação dos máximos de outras variáveis econômicas.

## REFERÊNCIAS

- ANG, A.; BEKAERT, G.; WEI, M. Do macro variables, asset markets, our surveys forecast inflation better? **Journal of Monetary Economics**, v. 54, n. 4, p. 1163-1212, 2007.
- ANHOLT, B. R.; WERNER, E.; SKELLY, D. K. Effect of food and predators on the activity of four larval ranid frogs. **Ecology**, v. 81, n. 12, p. 3509-3521, 2000.
- ASSOCIAÇÃO BRASILEIRA DAS ENTIDADES DOS MERCADOS FINANCEIRO E DE CAPITAIS. **Metodologia de Cálculo dos Valores Nominais Atualizados**. 2008. Disponível em: < [http://www.anbima.com.br/vna/arqs/vna\\_metodologia.pdf](http://www.anbima.com.br/vna/arqs/vna_metodologia.pdf) >. Acesso em: 25 maio 2016.
- ANDERSON, R. L. Distribution of the Serial Correlation Coefficient. **The Annals of Mathematical Statistics**, v. 13, n. 1, p. 1-13, 1942.
- ARRAES, R. A.; ROCHA, A. S. Perdas extremas em mercados de risco. **Revista Contabilidade e Finanças**, São Paulo, v. 17, n. 42, dez. 2006.
- ASSAF NETO, A. **Mercado financeiro**. São Paulo: Editora Atlas, 2008.
- BARBOSA, E. C. et al. Distribuição Generalizada de Valores Extremos (GVE): um estudo aplicado a valores de temperatura mínima da cidade de Viçosa-MG. **Revista da Estatística da UFOP**, Ouro Preto, v. 3, n. 3, p. 387-391, 2014.
- BEHRENS, C. N.; LOPES, H. F.; GAMERMAN, D. Bayesian analysis of extreme events with threshold estimation. **Statistical Modelling**, v. 4, n. 3, p. 227-244, 2004.
- BEIJO, L. A.; AVELAR, F. G. Distribuição Generalizada de Valores Extremos no estudo de dados climáticos: uma breve revisão e aplicação. **Revista da Estatística UFOP**, Ouro Preto, v. 1, 2010.
- BEIJO, L. A. et al. Estudo da precipitação máxima em Jaboticabal (SP) pela distribuição de Gumbel utilizando dois métodos de estimação dos parâmetros. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 11, n. 1, p. 141-147, 2003.
- BEIJO, L. A.; VIVANCO, M. J. F.; MUNIZ, J. A. Análise Bayesiana no estudo do tempo de retorno das precipitações pluviométricas máximas em Jaboticabal (SP). **Ciência e Agrotecnologia**, Lavras, v. 33, n. 1, p. 261-270, 2009.
- BERGER, P. L. **Mercado de Renda Fixa no Brasil: ênfase em títulos públicos**. Rio de Janeiro: Editora Interciência, 2015.
- BLANCHARD, O. **Macroeconomia: teoria e política econômica**. Rio de Janeiro: Editora Campus, 1999.



BLOMBERG, S.; HARRIS, B. Ethan, S. The Commodity-Consumer Price Connection: Fact or Fable? **Economic Policy Review**, v. 1, n. 3, Oct. 1995.

BOULIER, J. F.; HUANG, S. J.; TAILLARD, G. Optimal management under stochastic interest rates: the case of a protected defined contribution pension fund. **Insurance: Mathematics and Economics**, v. 28, n. 2, p. 173-189, 2001.

BOX, G. E. P.; PIERCE, D. A. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Model. **Journal of the American Statistical Association**, v. 65, n. 332, p. 1509-1526, 1970.

BOX, G. E. P.; TIAO, G. **Bayesian Inference in Statistical Analysis**. [S.l]: Addison-Wesley, 1973.

BOZDOGAN, H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, v. 52, n. 3, p. 345-370, 1987.

BRASIL. **Decreto nº 3859**, de 4 de julho de 2001. Estabelece as características dos Títulos da Dívida Pública Mobiliária Federal interna e dá outras providências. Brasília, 2001. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/decreto/2001/D3859.htm](http://www.planalto.gov.br/ccivil_03/decreto/2001/D3859.htm)>. Acessado em: 2 abr. 2016.

BRASIL. **Lei Complementar nº 101**, de 4 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. Brasília, 2000. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/leis/LCP/Lcp101.htm](http://www.planalto.gov.br/ccivil_03/leis/LCP/Lcp101.htm)>. Acessado em: 8 jun. 2016.

BRASIL. Ministério da Fazenda. **Títulos Públicos Federais: metodologia de cálculo dos títulos públicos federais ofertados nos leilões primários**. 2016.

BRASIL. Ministério das Relações Exteriores **BRICS: Brasil, Rússia, Índia, China e África do Sul**. 2016. Disponível em: <<http://www.itamaraty.gov.br/pt-BR/politica-externa/mecanismos-inter-regionais/3672-brics>>. Acessado em: 23 maio 2016.

BRYAN, M. F.; CECCHETTI, S. G. The consumer price index as a measure of inflation. **National Bureau of Economic Research Working Paper Series**, n. 4505, Oct. 1993.

CAGAN, P. "The Monetary Dynamics of Hyperinflation". In Friedman, Milton (ed.). **Studies in the Quantity Theory of Money**. Chicago: University of Chicago Press, 1956.

CALDEIRA, J. F.; MOURA, G. V. S. Otimização de carteiras de títulos públicos. **Advances in Scientific and Applied Accounting**, São Paulo, v. 5, n. 3, p. 349-376, 2012.

CARNEIRO, R. **Desenvolvimento em crise: a economia brasileira no último quarto de século XX**. São Paulo: Editora UNESP, IE-Unicamp, 2002.

CARSTENSEN, K.; HAGEN, J.; HOSSFELD, O.; NEAVES, A. S. Money Demand Stability

and Inflation Prediction in the Four Largest EMU Countries. **Scottish Journal of Political Economy**, v. 56, n. 1, p. 73-93, 2009.

CASELLA, G.; BERGER, R. L. **Inferência estatística**. São Paulo: Cengage Learning, 2010.

CHICHEPORTICHE, R. BOUCHAD, J. Weighted Kolmogorov-Smirnov test: Accounting for the tails. **Physical Review**, v. 86, n. 4, 2012.

CHOY, S. L.; O'LEARY, R.; MENGERSEN, K. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. **Ecological Society of America**, v. 90, n. 1, p. 265-277, 2009.

CIANI, T. A.; PIMENTA JUNIOR, T.; OLIVEIRA, R. L. O Desempenho de grandes empresas do BRIC, EUA, Japão e Alemanha: uma comparação com base na geração de valor. **Gestão e Produção**, São Carlos, v. 22, n. 4, out./dez. 2015.

COLES, S. G. **An Introduction to Statistical Modeling of Extreme Values**. Bristol: Springer, 2001.

COLES, S. G.; DIXON, M. J. Likelihood-based inference for extreme value models. **Extremes**, v. 2, n. 1, p. 5-23, 1999.

COLES, S. G.; POWELL, E. A. Bayesian methods in extreme value modelling: a review and new developments. **International Statistics Review**, Voorburg, v. 64, n. 1, p. 119-136, 1996.

COLES, S. G.; TAWN, J. A. Statistics of coastal flood prevention. **Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**, v. 332, n. 1627, p. 457-476, 1990.

COLES, S. G.; TAWN, J. A. A Bayesian analysis of extreme rainfall data. **Applied Statistics**, Oxford, v. 45, n. 4, p. 463-478, 1996.

COOPER, A. F.; ANTKIEWICZ, A. **Emerging powers in global governance: lessons from the Heiligendamm process**. Waterloo: Wilfrid Laurier Univ. Press, 2008.

COUMOU, D. et al. Quase-resonant circulation regimes and hemispheric synchronization of extreme values weather in boreal summer. **Potsdam Institute for Climate Impact Research**. v. 111, n. 34, p. 12331-12336, 2014.

CRUTCHER, H. L. A note on the Possible Missue of the Kolmogorov-Smirnov Test. **Journal of Applied Metereology**, v. 14, p. 1600-1603, 1975.

DAHAN, E.; MENDELSON, H. An extreme value model of concept testing. **Management Science**, v. 47, p. 102-116, 2001.

DAHLEN, K. E. et al. On the estimation of extreme values for risk assessment and Management: The ACER Method. **International Journal of Business**. v. 20, n. 1, p. 33-51, 2015.

- DAWSON, T. H. Maximum wave crests in heavy seas. **Journal of Offshore Mechanics and Arctic Engineering - Transactions of the AMSE**, v. 122, n. 3, p. 222-224, 2000.
- DIWERT, W. E. The Consumer Price Index and index number purpose. **Journal of Economic and Social Measurement**. v. 27 n. 3-4, p. 167-248, 2001.
- DIXON, M. J.; TAWN, J. A. Trends in UK extreme sea-levels: a spatial approach. **Geophysical Journal International**, v. 111, n. 3, p. 607-616, 1992.
- EMBRECHTS, P.; KLUPPELBERG, C.; MIKOSCH, T. **Modelling Extremal Events: for Insurance and Finance**. New York: Springer, 1997.
- ERGASHEV, B.; MITTNIK, S.; SEKERIS, E. A Bayesian Approach to Extreme Value Estimation in Operational Risk Modeling. **Center For Quantitative Risk Analysis**, Munique, v. 10, p. 1-27, Oct. 2013.
- FAMA, E. F.; GIBBONS, M. R. Inflation, Real returns and capital investment. **Journal of Monetary Economics**, v. 9, n. 3, p. 297-323, 1982.
- FILGUEIRAS, L. **História do Plano Real: fundamentos, impactos e contradições**. 3. ed. São Paulo: Boitempo, 2006.
- FINKENSTADT, B.; ROOTZÉN, H. **Extreme Values in Finance, Telecommunications, and the Environment**. [S.l]:Chapman & Hall/CRC, 2001.
- FISCHER, G. **Investment choice and inflation uncertainty**. London: The London School of economics and political Science, 2013.
- FISHER, R. A.; TIPPETT, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. **Proceedings of the Cambridge Philosophical Society**, v. 24, p. 180-190, 1928.
- FLEISCHMAN, S. J.; BURWEN, D. L. Mixture models for the species apportionment of hydroacoustic data, with echo-envelope length as the discriminatory variable. **ICES Journal of Marine Science: Journal du Conseil**, v. 60, n. 3, p. 592-598, 2003.
- FRANKLIN Jr, S. L.; DUARTE, T. B.; NEVES, C. R.; MELO, E. F. L. A estrutura a termo de taxas de juros no Brasil: modelos, estimação e testes. **Economia Aplicada**, v. 16, n.2, p. 255-290, 2012.
- GAMERMAN, D. Sampling from the posterior distribution in generalized linear mixed models. **Statistics and Computing**, v. 7, n. 1, p. 57-68, 1997.
- GARCIA, R.; PERRON, P. An Analysis of the Real Interest Rate Under Regime Shifts. **The Review of Economics and Statistics**, v. 78, n. 1, p. 111-125, Feb. 1996.
- GEWEKE, J. **Evaluating the accuracy of sampling-based approaches to the calculation of**

**posterior moments.** Minneapolis: Federal Reserve Bank of Minneapolis, Research Department, 1992.

GIANNONE, D. et al. Short-term inflation projections: A Bayesian vector autoregressive approach. **International Journal of Forecasting**, v.30, n.3, p. 635-644, 2014.

GILLELAND, E.; KATZ, R. W. **extRemes:** extreme value analysis. R package version 2.0-7. 2015.

GLEBOVA, I. S.; SADYRTDINOV, R.; RODNYANSKY, D. Impact analysis of investment attractiveness of the Republic of Tatarstan on fixed investments of its leading companies. **World applied sciences Journal**, v. 26, n. 7, p. 911-916, 2013.

GNEDENKO, B. V. Sur la distribution limite du terme maximum d'une serie aleatoire. **Annals of Mathematics**, v. 44, p. 423-453, 1943.

GOMES, I.; GUILLOU, A. Extreme Value Theory and Statistics of Univariate Extremes: A Review. **International Statistical Review**, Garsington, v. 83, n. 2, p. 263-292, Jan 2015.

GOMES, P. H. V. et al. Fatores determinantes do disclosure voluntário sob o enfoque da sustentabilidade: uma análise das empresas dos países do BRIC. **Revista de Gestão Social e Ambiental**, v. 9, n. 2, p. 70-87, 2015.

GRAÇA, W. G.; JANUZZI, F. V.; SOARES, F. P. O dilema do pequeno investidor perante as alternativas conservadoras de mercado: títulos públicos versus fundos referenciados DI. **Revista de Administração de Roraima-UFRR**, Boa Vista, v. 5, n. 2, p.189-214, jul./dez. 2015.

GREMAUD, A. P.; VASCONCELLOS Jr, M. A. S.; TONETO, R. **Economia brasileira: contemporânea.** 7. ed. São Paulo: Editora Atlas, 2009.

GRIMM, A. M. et al. Interdecadal variability and extreme precipitation events in South America during the monsoon season. **Climate Research**, v. 68, p. 277-294, 2016.

GROSEN, A.; JORGENSEN, P. L. Fair valuation of life insurance liabilities: The impact of interest rate guarantees, surrender options, and bonus policies. **Insurance: Mathematics and Economics**, v. 26, n. 1, p. 37-57, 2000.

HARRIS, R. I. The accuracy of design values predicted from extreme value analysis. **Journal of Wind Engineering and Industrial Aerodynamics**, v. 89, n. 2, p. 153-164, 2001.

HILL, R. J. When does chaining reduce the Paasche-Laspeyres spread? An application to scanner data. **Review of Income and Wealth**, v. 52, n. 2, p. 309-325, 2006.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **BRICS Joint Statistical Publication.** Rio de Janeiro, 2014. Disponível em:<[http://www.brics.ibge.gov.br/downloads/BRICS\\_Join t\\_ Statistcal\\_ Publication\\_ 2014.pdf](http://www.brics.ibge.gov.br/downloads/BRICS_Join_t_Statistical_Publication_2014.pdf)>. Acessado em: 21 mar. 2016.

IPEA-DATA. **Inflação**: IPCA. 2016. Disponível em:

<<http://www.ipeadata.gov.br/Default.aspx>>. Acessado em: 30 maio 2016.

JENKINSON, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. **Quarterly Journal of the Royal Meteorological Society**, v. 81, n. 348, p. 158-171, Apr. 1955.

JO, J.; LUSK, J. L. Predicting food prices using Data from consumer surveys and search. **Southern Agricultural Economics Association Annual Meeting**, Santo Antonio, v. 1, n. 1, 2016.

KASS, R. E.; WASSERMAN, L. The Selection of Prior Distributions by Formal Rules. **Journal of the American Statistical Association**, v. 91, n. 435, p. 1343-1370, 1996.

KLIESTER, K.; MITTNIK, S.; PAOLELLA, M. S. Value-at-Risk Prediction: A Comparison of Alternativa Strategies. **Journal of Financial Econometrics**, v. 4, n. 1, p. 53-89, 2006.

KOLMOGOROV, A. Sulla Determinazione Empirica di una Legge di Distribuzione. **Giornale dell'Istituto Italiano degli Attuari**, v. 4, p. 1-11, 1933.

LEE, H.; WU, J. Mean Reversion of Inflation Rates: Evidence from 13 OECD Countries. **Journal of Macroeconomics**, v. 23, n. 3, p. 477-487, 2001.

LJUNG, G. M.; BOX, G. E. P. On a Measure of Lack Fit in Time Series Models. **Biometrika**, v. 65, n. 2, p. 297-303, 1978.

MAHBUBANI, K. **The New Asian Hemisphere**: The Irresistible Shift of Global Power to the East. New York: Public Affairs, 2008.

MARKOWITZ, H. Portfolio selection. **Journal of Finance**, v. 7, n. 1, p. 77-91, 1952.

MARQUES, R.; PIZZINGA, A.; VEREDA, L. Restricted Kalman filter applied to dynamic style analysis of actuarial funds. **Applied Stochastic Models in Business and Industry**, v. 28, n. 6, p. 558-570, 2012.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**. [S.l]: CRC press, 1989.

MCNULTY, P. J., SCHEICK, L. Z., ROTH, D. R., DAVIS, M. G., TORTORA, M. R. S. First failure predictions for EPROMs of the type flown on the MPTB satellite. **IEEE Transactions on Nuclear Science**, v. 47, p. 2237-2243, 2000.

MUNDELL, R. Inflation and Real Interest. **Journal of Political Economy**, v. 71, n. 3, p. 280-283. Jun. 1963.

NADARAJAH, S.; CHOI, D., Maximum daily rainfall in South Korea. **Journal of Earth System Science**, v. 116, n. 4, p. 311-320, 2007.

NAGHETTINI, M.; PINTO, E. J. A. **Correlação e regressão. Hidrologia e Estatística.** Belo Horizonte: Serviço Geológico do Brasil-CPRM, 2007.

NELDER J. A.; WEDDERBURN, R. W. Generalized Linear Models. **Journal of the Royal Statistical Society. Series A (General)**, v. 135, n. 3, p.370-384, 1972.

NOGUEIRA, D. A.; SÁFADI, T.; FERREIRA, D. F. Avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov. **Revista brasileira de Estatística**, v. 65, n. 224, p. 59-88, 2004.

O'NEIL, J. Building better global economic BRICs. **Global Economics Paper**, v. 66, 2001.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Inflation (CPI)**. Disponível em: < <https://data.oecd.org/price/inflation-cpi.htm>>. Acessado em: 15 jun. 2016.

OTROK, C.; WHITEMAN, C. H. Bayesian Leading Indicators: Measuring and Predicting Economic Conditions in Iowa. **International Economic Review**, v. 39, n. 4, p. 997-1014, 1998.

PAO, H.; TSAI, C. CO2 emissions, energy consumption and economic growth in BRIC countries. **Energy Policy**, v.38, n. 12, p. 7850-7860, 2010.

PARENT, E.; BERNIER, J. Encoding prior experts judgments to improve risk analysis of extreme hydrological events via POT modeling. **Journal of Hydrology**, v. 283, n. 1, p. 1-18, 2003.

PAULINO, C. D.; TURKMAN, M. A. A.; BENTO, M. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003.

PEREIRA, L. C. B.; NAKANO, Y. Hiperinflação e estabilização no Brasil: o primeiro plano Collor. **Revista de Economia Política**, v. 11, n.4, out./dez. 1991.

PLUMMER, M. et al. **CODA: output Analysis and Diagnostics for MCMC**. R package version 0.18-1. 2015.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016.

RAFTERY, A. E.; LEWIS, S. Comment: One long run with diagnostics: implementation strategies for markov chain monte carlo. **Statistical Science**, Hayward, v. 7, n. 4, p. 493-497, 1992.

RAZALI, N. M.; WAH, Y. B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. **Journal of Statistical Modeling and Analytics**, v. 2, n. 1, p. 21-33, 2011.

REIS, C. J.; BEIJO, L. A.; AVELAR, F. G. Análise dos níveis mínimos de umidade relativa do

ar via distribuição Generalizada de Valores Extremos. **Sigmae**, Alfenas, v. 1, n. 1, p. 57-64, 2012.

RÍOS-ALEJANDRO, J. G. Temperaturas extremas en la ciudad de Monterrey NL México. **Revista Chapingo. Serie ciencias forestales y del ambiente**, v. 17, n. 2, p. 225-230, 2011.

ROBERTS, S. J. Extreme value statistics for novelty detection in biomedical data processing. **IEE Proceedings - Science Measurement and Technology**, v. 147, p. 363-367, 2000.

SACULLINGGAN, M.; BALASE, E. A. Empirical Power Comparison Of Goodness of Fit Tests for Normality In The Presence of Outliers. **Journal of Physics: Conference Series**, v. 435, 2013.

SHAHZADI, A. A review: regional frequency analysis of anual maximum rainfall in monsoon region of Pakistan using L-moments. **International Journal of Advanced Statistics and Probability**, v. 1, n. 3, p. 97-101, 2013.

SHARMA, S. S. Can consumer price index predict gold price returns? **Economic Modelling**, v. 55, p. 269-278, Jun. 2016.

SMIRNOV, H., Sur les fucarts de la Courbe de Distribution Empirique. **In Recueil Mathematique: Matemnticeskii Sbornik**, v. 48, n. 1, p. 3-26, 1939.

SMITH, R. L. Maximum likelihood estimation in a class of nonregular cases. **Biometrika**, London, v. 72, n. 1, p. 67-90, 1985.

SOUZA, H. C. et al. Avaliação do impacto de alterações na marcação de papéis de renda fixa para um fundo de pensão mineiro. **Revista da Universidade Vale do Rio Verde**, v. 14, n. 1, p. 484-504, 2015.

STEPHENSON, A. **evd**: functions for extreme value distributions. R package version 2.3-0. 2012.

STEPHENSON, A.; RIBATET, M. **evdbayes**: Bayesian Analysis in Extreme Value Theory. R package version 1.0-6. 2006.

TANG, K.; XIONG, W. Index investment and the financialization of commodities. **Financial Analysts Journal**, v. 68, n. 5, p. 54-74, 2012.

THE THEORY OF EVERYTHING. Director: James Marsh. Production: Tim Bevan. United Kingdom, 123 min. 2014.

TORMAN, V. B. L.; COSTER, R.; RIBOLDI, J. Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação. **Clinical & Biomedical Research**, v. 32, n. 2, p. 227-234, 2012.

UNIVERSIDADE FEDERAL DE ALFENAS. **Manual de normalização para elaboração de trabalhos acadêmicos, dissertações e teses da Unifal-MG**. 2016.

WALD, A.; WOLFOWITZ, J. Optimum Character of the Sequential Probability Ratio Test. **The Annals of Mathematical Statistics**, v. 19, n. 3, p.326-399, 1948.

WEBB, S. B. **Hyperinflation and stabilization in Weimar Germany**. Oxford: Oxford university press, 1989.

WEST, M.; CROSSE, J. Modelling Probabilistic Agent Opinion. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 54, n. 1, p. 285-299, 1992.

WINTLE, B. A. et al. The Use of Bayesian Model Averaging to Better Represent Uncertainty in Ecological Models. **Conservation Biology**, v. 17, n. 6, p. 1579-1590, 2003.

WRIGHT, J. H. Forecasting US Inflation by Bayesian Model Averaging. **Journal of Forecasting**, v. 28, n. 2, p. 131-144, 2009.



## APÊNDICES

### APÊNDICE A - Rotina no R utilizada para calcular as estimativas dos parâmetros e níveis de retorno via Máxima Verossimilhança

```
##### Cria o vetor dos níveis máximos quadrimestrais de IPCA.
```

```
dados_ipca_max<-c(2.25, 0.61, 0.78, 0.76, 0.91, 0.86, 0.87,
0.49, 0.75, 0.59, 0.19, 0.48, 0.44, 0.47, 0.74, 0.55, 0.79,
0.45, 0.55, 0.47, 0.41, 0.78, 0.43, 0.83, 0.83, 0.47, 0.53,
0.64, 0.43, 0.79, 0.86, 0.37, 0.92, 0.92, 0.46, 0.78, 1.32,
0.79, 1.01)
```

```
##### Gera gráfico dos níveis máximos quadrimestrais de IPCA.
```

```
plot(dados_ipca_max, type = "l")
```

```
##### Calcula o valor p do teste de Ljung-Box para o vetor
##### dos máximos de IPCA.
```

```
Box.test(dados_ipca_max, type = "Ljung-Box")
```

```
##### Calcula o valor p do Runs-Test para o vetor dos máximos
##### de IPCA.
```

```
md<- median(dados_ipca_max) # Mediana
z_ab<- as.factor(ifelse(dados_ipca_max<md,"A","B"))
runs.test(z_ab)
```

```
##### Encontra as estimativas e intervalos de confiança dos
##### parâmetros da distribuição GEV e Gumbel (Shape = 0) via
##### Método da Máxima Verossimilhança.
```

```
modelo1<-fgev(dados_max, corr=T) # Dist. GEV
modelo1
confint(modelo1) # Intervalo de Confiança GEV
modelo2<-fgev(dados_max, shape=0, corr = T) # Dist Gumbel
modelo2
```

```

confint(modelo2) # Intervalo de Confiança Gumbel
}

##### Teste de Razão de Verossimilhança.

anova(modelo1,modelo2)

##### Encontra os valores p dos testes de
##### Kolmogorov-Smirnov, para as distribuições
##### GEV e Gumbel.

mu<-modelo2$estimate[1] # Param Posição Gumbel
sig<-modelo2$estimate[2] # Param Escala Gumbel
mu2<-modelo1$estimate[1] # Param Posição GEV
sig2<-modelo1$estimate[2] # Param Escala GEV
xi<-modelo1$estimate[3] # Param Forma GEV
ks.test(dados_max,"pgumbel",mu,sig,alternative="greater")
ks.test(dados_max,"pgev",mu2,sig2,xi,alternative="greater")

##### Retira os 10 últimos valores do vetor de máximos
##### de IPCA para construção da série para ajuste

dados_ipca_max2<-dados_ ipca_max[1:29]
modelo21<-fgev(dados_ ipca_max2, shape=0) # Dist Gumbel
modelo21
modelo22<-fgev(dados_ ipca_max2) # Dist GEV
modelo22

##### Estima os níveis de retornos e construindo
##### intervalos de confiança para os mesmos, para os
##### tempos de retorno de 2, 3, 4, 5 e 10 quadrimestres.

mu3<-modelo21$estimate[1] # Param Posição Gumbel
sig3<-modelo21$estimate[2] # Param Escala Gumbel
mu4<-modelo31$estimate[1] # Param Posição GEV
sig4<-modelo31$estimate[2] # Param Escala GEV
xi3<-modelo31$estimate[3] # Param Forma GEV
px<-c(2,3,4,5,10) # Tempos de Retorno

```

```
prob<-(1/px)
ret_gev<-qgev(prob,loc=mu4,scale=sig4,shape=xi3,lower.tail = F)
ret_gev # Niveis de Retorno GEV
ret_gumbel<-qgev(prob,loc=mu3,scale=sig3,lower.tail = F)
ret_gumbel # Niveis de Retorno Gumbel
q_fit <- fgev(dados_max2, prob = 1/2)
confint(q_fit) # Intervalo de Confiança para tempo 2 GEV
q_fit <- fgev(dados_max2, prob = 1/3)
confint(q_fit) # Intervalo de Confiança para tempo 3 GEV
q_fit <- fgev(dados_max2, prob = 1/4)
confint(q_fit) # Intervalo de Confiança para tempo 4 GEV
q_fit <- fgev(dados_max2, prob = 1/5)
confint(q_fit) # Intervalo de Confiança para tempo 5 GEV
q_fit <- fgev(dados_max2, prob = 1/10)
confint(q_fit) # Intervalo de Confiança para tempo 10 GEV
q_fit <- fgev(dados_max2, , shape=0, prob = 1/2)
confint(q_fit) # Intervalo de Confiança para tempo 2 Gumbel
q_fit <- fgev(dados_max2, , shape=0 ,prob = 1/3)
confint(q_fit)# Intervalo de Confiança para tempo 3 Gumbel
q_fit <- fgev(dados_max2, , shape=0 ,prob = 1/4)
confint(q_fit)# Intervalo de Confiança para tempo 4 Gumbel
q_fit <- fgev(dados_max2, , shape=0, prob = 1/5)
confint(q_fit)# Intervalo de Confiança para tempo 5 Gumbel
q_fit <- fgev(dados_max2, , shape=0,prob = 1/10)
confint(q_fit)# Intervalo de Confiança para tempo 10 Gumbel
```

## APÊNDICE B - Rotina no R utilizada para calcular as estimativas dos parâmetros e níveis de retorno via Inferência Bayesiana

```
##### Cria os vetores dos níveis máximos quadrimestrais do
##### IPC da Rússia, China, Índia e a média dos três
##### países (MRCI). Além disso encontra os valores p
##### dos testes de Ljung-Box para cada um dos vetores.

dados_max_chi<-c(0.89999906, 1.19999479, 1.10000023, 0.50000478,
1.09999506, 0.60000416, 1.79999480, 0.70000608, 1.29999990,
0.50000358, 0.49999577, 1.39999973, 0.99999656, 1.20000039,
1.20001507, 2.59999390, 0.09998985, 0.89998743, 0.00000000,
0.49999860, 1.00008306, 1.20579804, 0.60107158, 1.09838317,
1.27844418, 0.47215849, 1.50094758, 0.18503823, 0.65236594,
0.99956235, 1.09967352, 0.79705221, 0.99998841, 0.56701249,
0.53095499, 0.35309153, 1.22805634, 0.52317275, 0.52172071,
1.80621814) # Série de máximos da China
dados_max_rus<-c(1.6299909, 0.7999974, 1.7499931, 0.9900076,
0.9200025, 2.6199885, 1.3400060, 0.6400053, 2.4299969, 1.6599964,
0.6700070, 1.6799946, 1.1100055, 0.9500014, 2.3100090, 1.4199997,
0.9699946, 2.3699958, 1.6499952, 0.6300035, 1.6400064, 0.8599921,
0.8400801, 2.3699904, 0.7799645, 0.2300152, 0.4999968, 0.5799995,
1.2300329, 0.9700145, 0.6599968, 0.8199895, 0.5899622, 1.0199748,
0.6499930, 3.8500077, 2.2199750, 0.7999883, 1.0000102, 0.5999820)
# Série de máximos da Rússia
dados_max_ind<-c(1.2320381, 0.8048214, 0.8015957, 0.7936434,
0.9765613, 0.9596918, 0.7618976, 1.7013305, 1.1070158, 0.8403346,
1.6528897, 1.6000114, 0.7874002, 1.5384588, 0.7518784, 1.4814790,
2.1428534, 1.3698607, 1.3513610, 4.5751670, 1.8181787, 1.1764685,
2.2988770, 1.6483001, 0.5405873, 2.1163992, 0.5076582, 1.9900471,
1.9230737, 0.9302307, 0.9049756, 1.7315984, 1.2605393, 1.2551910,
2.4390209, 0.3952910, 0.7873999, 1.1627542, 1.1277842, 1.4760124)
# Série de máximos da Índia

Box.test(dados_max_chi, type = "Ljung-Box")
#Teste Ljung-Box para série da China
Box.test(dados_max_rus, type = "Ljung-Box")
#Teste Ljung-Box para série da Rússia
Box.test(dados_max_ind, type = "Ljung-Box")
```

```

#Teste Ljung-Box para série da Índia
dados_max_bric<-dados_max_chi+dados_max_ind+dados_max_rus
dados_max_bric<-dados_max_bric/3
dados_max_bric # Série de máximos da MRCI
Box.test(dados_max_bric, type = "Ljung-Box")
#Teste Ljung-Box para série da MRCI

##### Modela a distribuição GEV via Método da Máxima
##### Verossimilhança para cada vetor e mostra
##### as estimativas e as matrizes de variâncias
##### e covariâncias de cada modelo.

modelo_rus<-fevd(dados_max_rus)
modelo_rus #Dist GEV para Rússia
modelo_chi<-fevd(dados_max_chi)
modelo_chi #Dist GEV para China
modelo_ind<-fevd(dados_max_ind)
modelo_ind #Dist GEV para Índia
modelo_bric<-fevd(dados_max_bric)
modelo_bric #Dist GEV para MRCI

##### Constrói as distribuições a priori e posteriori,
##### nesse caso muda-se as matrizes de variâncias e
##### covariâncias e os vetores de acordo com os
##### observados nos modelos ajustados de cada país
##### pela Máxima Verossimilhança. As matrizes são
##### multiplicadas por 1, 2 ou 4 (i). Na construção da
##### distribuição a posteriori chama-se um dos dois
##### tipos de distribuição a priori, Normal I ou Normal
##### II. Depois disso os valores encontrados são
##### armazenados em um objeto MCMC.

mat_cov_priori<-matrix(c(11,12,13,21,22,23,31,32,33,ncol=3))
# Matriz de Variâncias e Covariâncias
vet_med_priori<-c(Posição,log(Escala),Forma)
# Vetor de médias para Normal I
vet_med_priori2<-c(log(Posição),log(Escala),Forma)
# Vetor de médias para Normal II
pn1<-prior.norm(mean=vet_med_priori, cov = (i*mat_cov_priori))

```

```

# Construção da distribuição a priori Normal I
pn2<-prior.loglognorm(mean=vet_med_priori2, cov = (i*mat_cov_priori))
# Construção da distribuição a priori Normal II
modelo3<-posterior(10000000, init=c(0,1,0), prior = pnj, lh="gev",
burn = 100000,thin = 100, data = dados_max, psd = c(.1,.2,.2))
# Construção da distribuição a posteriori
modelo3<-as.mcmc(modelo3)
# Armazena os valores em um objeto MCMC

##### Encontra as estimativas e desvios padrão dos
##### parâmetros, constrói os intervalos de
##### credibilidade e averigua os critérios de
##### convergência das cadeias.

summary(modelo3)
# Estimativas e desvios padrão
HPDinterval(modelo3)
# Intervalos HPD
raftery.diag(modelo3)
# Critério de Raftery-Lewis
geweke.diag(modelo3)
# Critério de Geweke

##### Retira os 10 últimos valores de cada um
##### dos vetores de níveis máximos quadrimestrais
##### de IPC, para compor as séries para ajuste

modelo_rus2<-fevd(dados_max_rus[1:29])
modelo_rus2 # Dist GEV para Rússia
modelo_chi2<-fevd(dados_max_chi[1:29])
modelo_chi2 # Dist GEV para China
modelo_ind2<-fevd(dados_max_ind[1:29])
modelo_ind2 # Dist GEV para Índia
modelo_bric2<-fevd(dados_max_bric[1:29])
modelo_bric2 # Dist GEV para MRCI

##### Igualmente aos modelos completos, contrói
##### todas as distribuições a priori e posteriori,
##### seus intervalos de credibilidade e os

```

```

##### critérios de convergência.

mat_cov_priori1<-matrix(c(11,12,13,21,22,23,31,32,33,ncol=3))
# Matriz de Variâncias e Covariâncias
vet_med_priori1<-c(Posição,log(Escala),Forma)
# Vetor de médias para Normal I
vet_med_priori12<-c(log(Posição),log(Escala),Forma)
# Vetor de médias para Normal II
pn11<-prior.norm(mean=vet_med_priori1, cov = (i*mat_cov_priori1))
# Construção da distribuição a priori Normal I
pn12<-prior.loglognorm(mean=vet_med_priori12, cov = (i*mat_cov_priori1))
# Construção da distribuição a priori Normal II
modelo4<-posterior(10000000, init=c(0,1,0), prior = pn1j, lh="gev",
burn = 100000,thin = 100, data = dados_max2, psd = c(.1,.2,.2))
# Construção da distribuição a posteriori
modelo4<-as.mcmc(modelo4)
# Armazena os valores em um objeto MCMC
summary(modelo4)
# Estimativas e desvios padrão
HPDinterval(modelo4)
# Intervalos HPD
raftery.diag(modelo4)
# Critério de Raftery-Lewis
geweke.diag(modelo4)
# Critério de Geweke

##### Calcula as estimativas dos níveis de retorno
##### para os tempos de 2, 3, 4, 5 e 10 quadrimestres,
##### seus respectivos intervalos de credibilidade e
##### os critérios de convergência para as cadeias dos
##### níveis de retorno.

k<-length(modelo4[,1])
loc <- numeric(k)
escala <- numeric(k)
forma <- numeric(k)
prob=c(1/2,1/3,1/4,1/5,1/10) # Tempos de retorno
np<-length(prob)

```

```

q1 <- numeric(k)
q2 <- numeric(k)
q3 <- numeric(k)
q4 <- numeric(k)
q5 <- numeric(k)
for (i in 1:k)
{
  loc[i] <- modelo5[i,1]
  escala[i] <- modelo5[i,2]
  forma[i] <- modelo5[i,3]
  q1[i]<-(loc[i]+((escala[i]/forma[i])* (((-log(1-prob[1])) ^
                                          (-forma[i]))-1)))
  q2[i]<-(loc[i]+((escala[i]/forma[i])* (((-log(1-prob[2])) ^
                                          (-forma[i]))-1)))
  q3[i]<-(loc[i]+((escala[i]/forma[i])* (((-log(1-prob[3])) ^
                                          (-forma[i]))-1)))
  q4[i]<-(loc[i]+((escala[i]/forma[i])* (((-log(1-prob[4])) ^
                                          (-forma[i]))-1)))
  q5[i]<-(loc[i]+((escala[i]/forma[i])* (((-log(1-prob[5])) ^
                                          (-forma[i]))-1)))
}
nit<-length(q1)
qo<-matrix(0,nit,5)
qo[,1]<-q1
qo[,2]<-q2
qo[,3]<-q3
qo[,4]<-q4
qo[,5]<-q5
nivel_ret_mcmc <- as.mcmc(qo)
summary(nivel_ret_mcmc)
# Estimativas e Desvios Padrão do Níveis de Retorno
HPDinterval(nivel_ret_mcmc)
# Intervalos HPD para os Níveis de Retorno
raftery.diag(nivel_ret_mcmc)
# Critério de Raftery-Lewis para as cadeias
geweke.diag(nivel_ret_mcmc)
# Critério de Geweke para as cadeias

```



## ANEXOS

## ANEXO A - Pesquisas de Orçamento Familiar (POF) de 95-96

Alimentação e bebidas	refeição pronta; lanche para viagem; refeição fora de casa (fc); lanche (fc); café da manhã (fc); doces (fc); refrigerante (fc); cafezinho (fc); caldos (fc); cerveja (fc); chopp (fc); aguardente (fc); e outras bebidas alcoólicas (fc)
Habitação	aluguel residencial; condomínio; imposto predial; taxa de água e esgoto; mudança;. gás encanado; e energia elétrica residencial.
Artigos de residência	conserto de geladeira, freezer, aparelho de som, videocassete, máquina de lavar/secar roupa, condicionador de ar, bomba d'água; reforma de estofado; e manutenção de caixa d'água, fossa etc
Transportes (públicos)	ônibus urbano, intermunicipal e interestadual; táxi; trem; ferry-boat; avião; metrô; navio; barco; e transporte escolar
Transportes (particulares)	emplacamento e licença; seguro voluntário de veículo; conserto de automóvel; estacionamento; pedágio; lubrificação e lavagem; e pintura de veículo
Saúde e cuidados pessoais	médico; dentista; tratamento psicológico e fisioterápico; exame de laboratório; hospitalização e cirurgia; eletrodiagnósticos; radiografia; e plano de saúde
Serviços pessoais	costureira; tinturaria e lavanderia; manicure e pedicure; barbeiro; cabeleireiro; empregado doméstico; depilação; massagem e sauna; cartório; despachante; serviço funerário; alfaiate; serviço bancário; e conselho de classe
Recreação	cinema; ingresso para jogo; clube; teatro; aluguel de fita de videocassete; boate, danceteria e discoteca; jogos lotéricos; aluguel de fita de videogame; motel; tele sena; bingo; hotel; revelação e cópia; e excursões
Educação	creche; pré-escola; primeiro grau; segundo grau; terceiro grau; e cursos diversos
Comunicação	correio, telefone fixo; telefone público; telefone celular; e TV a cabo

**ANEXO B - Pesquisas de Orçamento Familiar (POF) de 02-03**

Alimentação e bebidas	refeição (fc); lanche (fc); café da manhã (fc); doces (fc); refrigerante e água mineral (fc); cafezinho (fc); cerveja (fc); chopp (fc); e outras bebidas alcoólicas (fc)
Habitação	aluguel residencial; condomínio; taxa de água e esgoto; mudança; mão de obra para reparos; gás encanado; e energia elétrica residencial
Artigos de residência	conserto de refrigerador, freezer, aparelho de som, televisor, máquina de lavar/secar roupa, bomba d'água; reforma de estofado; e conserto de bomba d'água
Transportes (públicos)	ônibus urbano, intermunicipal e interestadual; táxi; trem; ferry-boat; avião; metrô; barco; e transporte escolar
Transportes (particulares)	emplacamento e licença; seguro voluntário de veículo; conserto de automóvel; estacionamento; multa; pedágio; lubrificação e lavagem; pintura de veículo; reboque; e aluguel de veículo
Saúde e cuidados pessoais	médico; dentista; tratamento psicológico e fisioterápico; exame de laboratório; hospitalização e cirurgia; eletrodiagnósticos; radiografia; e plano de saúde
Serviços pessoais	costureira; manicure e pedicure; barbeiro; cabeleireiro; empregado doméstico; depilação; cartório; despachante; serviço bancário; e conselho de classe
Recreação	cinema; ingresso para jogo; clube; compra e tratamento de animais; aluguel de DVD e fita de videocassete; boate, danceteria e discoteca; jogos de azar; revelação e cópia; motel; hotel; e excursão
Educação	creche; educação infantil; ensino fundamental; ensino médio; ensino superior; pós-graduação; supletivo; curso preparatório; curso técnico; curso de idioma; curso de informática; autoescola; ginástica; natação, balé; e escolinha de esporte
Comunicação	correio, telefone fixo; telefone público; celular; TV a cabo; e acesso à internet

**ANEXO C - Pesquisas de Orçamento Familiar (POF) de 08-09**

Alimentação e bebidas	refeição (fc); lanche (fc); café da manhã (fc); doces (fc); refrigerante e água mineral (fc); cafezinho (fc); cerveja (fc); e outras bebidas alcoólicas (fc)
Habitação	aluguel residencial; condomínio; taxa de água e esgoto; mudança; mão de obra para reparos; e energia elétrica residencia
Artigos de residência	conserto de refrigerador, aparelho de som, televisor, máquina de lavar roupa, reforma de estofado; e manutenção de microcomputador
Transportes (públicos)	ônibus urbano, intermunicipal e interestadual; táxi; trem; passagem aérea; metrô; transporte hidroviário; e transporte escolar
Transportes (particulares)	emplacamento e licença; seguro voluntário de veículo; conserto de automóvel; estacionamento; pedágio; multa; lubrificação e lavagem; pintura de veículo; e aluguel de veículo
Saúde e cuidados pessoais	médico; dentista; fisioterapeuta; psicólogo; exame de laboratório; hospitalização e cirurgia; exame de imagem; e plano de saúde
Serviços pessoais	costureira; manicure; cabeleireiro; empregado doméstico; depilação; despachante; serviço bancário; e conselho de classe
Recreação	cinema; ingresso para jogo; clube; tratamento de animais; locação de DVD; boate e danceteria; jogos de azar; motel; hotel; excursão; e revelação e cópia
Educação	creche; educação infantil; ensino fundamental; ensino médio; ensino superior; pós-graduação; curso preparatório; curso técnico; curso de idioma; curso de informática; e atividades físicas
Comunicação	Correio; telefone fixo; telefone público; celular; acesso à internet; telefone com internet; e TV por assinatura com internet